

Article

Integrating GIS-MCDA and Machine Learning Approach to Identify Marginal and Underutilized Lands in Leon County, Florida

Tewodros A. Simret ¹, Sewunet A. Natae ², Gang Chen ³, Victor Ibeanusi ¹ and Hubert Hirwa ^{1,4,*}

¹ School of the Environment, Florida Agricultural and Mechanical University, Tallahassee, FL 32307, USA; tewodros1.simret@famou.edu (T.A.S.); victor.ibeanusi@famou.edu (V.I.)

² Department of Economics and Law, University of Cassino and Southern Lazio, 03043 Cassino, FR, Italy; sewunetabera.natae@studentmail.unicas.it

³ Department of Civil and Environmental Engineering, Florida A&M University–Florida State University College of Engineering, Tallahassee, FL 32307, USA; gchen@eng.famu.fsu.edu

⁴ Sustainability Institute, School of the Environment, Florida Agricultural and Mechanical University, Tallahassee, FL 32307, USA

* Correspondence: hubert.hirwa@famou.edu

Abstract

This study developed an integrated Geographical Information System (GIS) based Multi-Criteria Decision Analysis (GIS-MCDA) framework based on hydrological marginality, soil suitability, land cover marginality, slope, social marginality, and contaminated land indicators to identify and classify marginal and underutilized lands in Leon County, Florida, using the analytic hierarchy process (AHP) and weighted overlay analysis. The AHP results assigned the highest weights to hydrological conditions (36%) and soil suitability (24%), indicating their dominant influence on land marginality. The final marginality model classified the study area as low marginality/highly suitable (24.19%), slightly marginal (39.17%), moderately marginal (2.48%), high marginality (30.66%), and very high marginality and underutilization potential (3.5%). According to sensitivity analysis, hydrology, soil, and social emphasis models showed results that were relatively stable compared to the AHP-weighted baseline model. However, an equal-weight model yielded a lower estimate. An exploratory 2035 urban-expansion scenario was developed using a standardized 2-mile (3.2 km) buffer around existing developed areas. Under this scenario, the combined area classified as high and very high marginality increased from approximately 152,771 acres (34.14%) under baseline conditions to approximately 176,534 acres (39.45%), representing a net increase of about 23,763 acres. These findings show that combining environmental and socioeconomic indicators can improve the identification of marginal and underutilized land and support sustainable land-use planning, restoration targeting, and future agricultural or bioenergy suitability assessments.



Academic Editor: Yijun Yao

Received: 6 June 2026

Revised: 17 July 2026

Accepted: 20 July 2026

Published: 23 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: underutilized land; marginal land; land suitability; spatial analysis; analytical hierarchy process (AHP); socioeconomic vulnerability; weighted overlay analysis; random forest (RF)

1. Introduction

Marginal land is land with low productivity that provides little economic return due to various biophysical, social, and economic constraints [1]. Underutilized land refers to idle or abandoned land, which is currently not being used to its full productive potential [2].

Underutilized and marginal land is subject to economic constraints as its production is not cost-effective and profitable under current management techniques and policies. From biophysical perspective, underutilized and marginal land is also subject to natural constraints like steep slopes, weather conditions, and soil quality. In addition, social perspective including human factors such as workforce and distance from infrastructure also plays a crucial role.

The demand for food is increasing at unprecedented rates, while the agricultural land base is shrinking in many regions. This dynamics of population growth [3], land fragmentation [4], and urbanization and infrastructure have led to significant land-use change [5]. Global food production must increase by 50–60% by 2050 to meet the demand of the growing population [6]. Steering agricultural expansion toward marginal and underutilized lands is highly suggested to avoid the environmental consequences of clearing high-value ecosystems [1].

Globally, marginal land is estimated to cover approximately 2.7 billion hectares, of which 1.5 billion hectares is currently unused but suitable for agricultural cultivation [7]. On the global scale, marginal land is increasingly viewed as essential for a bio-based economy, with projections to supply 14–85% of the total global biomass demand by 2050 [8]. In the United States, estimates of marginal agricultural land are substantial but vary depending on the evaluation criteria, ranging from 10 million to 173 million hectares [2]. In a study focused on 10 Midwest U.S. states, approximately 11 million hectares of non-forested marginal land has been identified, which is suitable for bioenergy production [9]. Florida has 17.16 million acres of forested land, representing 50% of the state's total acreage [10]. In regions like the Florida Panhandle, the focus on underutilized or sensitive land often involves monitoring forest health to protect these resources from invasive species [10].

Marginal land plays a vital role in bioenergy production and is viewed as a win for growing dedicated energy crops like switchgrass, miscanthus, and short-rotation woody crops, as they avoid direct competition with food crops on prime land [2,9]. Marginal land also provides ecosystem services by repurposing and facilitating carbon sequestration, groundwater recharge, and wildlife habitat restoration [11].

Leon County, Florida, was selected as the study area due to its diverse range of landscapes, including urban, peri-urban, forested, agricultural, and wetland areas. In Leon County, Florida, the number of farms increased by 14%, from 325 to 375, between 2017 and 2022, respectively, but total farmland area decreased from 91,730 acres to 70,058 acres in the same period according to the USDA Census of Agriculture. Land fragmentation of Leon County positions it as a critical region for case study, as urban expansion exerts pressure on rural property values, declining the scale of local economies and accelerating farmland conversion [12]. In addition, the county faces several environmental constraints, including wetlands, hydrological limitations, fragmented land cover, and contaminated land. These factors limit conventional agricultural activities, creating opportunities for alternative land use, such as ecological restoration, carbon sequestration, and biomass production [13]. There is a need for localized research combining environmental and socioeconomic factors to identify marginalized and underutilized lands at the county level, which justifies the use of Leon County as the testbed for the case study. The result will help assess how GIS-based multiple-criteria methods facilitate sustainable land-use planning.

Geographic Information System (GIS) and Multi-Criteria Decision Analysis (MCDA) techniques are useful tools for integrating diverse datasets into an analytical framework. Among these methods, the Weighted Linear Combination (WLC) is commonly used to evaluate land suitability by combining multiple criteria into a single composite index. It allows decision makers to aggregate multiple, often conflicting criteria, such as physical and socioeconomic criteria, into a single suitability score for each location [14]. GIS is a tool for

data acquisition, storage, manipulation, and visualization in the form of suitability maps [5]. MCDA provides a structured set of techniques for weighting and ranking alternatives based on conflicting criteria, thereby effectively incorporating stakeholder values into the spatial model [5,14]. WLC is the most widely used MCDA technique for land suitability assessments [15,16]. This technique allows decision makers to evaluate land suitability by combining multiple spatial factors based on their relative importance [16]. The process typically follows a three-step workflow to integrate diverse datasets: standardization, weighting, and aggregation.

The goal of this study is to identify and classify marginal and underutilized land in Leon County, Florida, by integrating environmental and socioeconomic data within a GIS-based MCDA framework. The specific objectives are to apply AHP to (1) determine the relative importance of environmental and socioeconomic criteria influencing marginal and underutilized land distribution; (2) validate model performance using sensitivity analysis, correlation analysis, and random forest regression; and (3) simulate future changes in marginality under projected urban expansion conditions. The results aim to support policy decisions and land-use planning by highlighting areas where underutilized land overlaps with socially vulnerable communities.

2. Literature Review and Conceptual Contexts

2.1. The Concept of Marginal Underutilized Land

The concept of marginal and underutilized land, usually associated with economic theories, has evolved since the 19th century, which is defined through specific economic and social lenses and is often described as “idle,” “waste,” “abandoned,” or “degraded” land [3]. Specifically, marginal land refers to the dynamic land that depends on the specific purposes, time, and place of use. The term is defined based on economic and biophysical perspectives. Notably, David Ricardo’s Theory of Rent firstly introduced the definition of marginal [6]. According to this theory, marginal land is primarily defined as land where cost-effective production is not profitable under the current market conditions, cultivation techniques, and agricultural policies. This definition is not far from the modern one, in which, in modern modeling, it is often identified as land with potential profit equal to or less than zero [2].

From the biophysical perspective, land characterized by natural constraints such as high acidity, salinity, steep slopes, extreme aridity, or shallow soil depth is considered marginal land [1]. According to FAO and UNEP, a land can also be classified as marginal if it supports only 40% or less of its total productivity potential [6]. Marginality is not necessarily a permanent state as a land can be highly suitable for uses other than food crops, such as grazing, forestry, or dedicated bioenergy crops like *Arundo donax* [1,2].

Underutilized land refers to areas that are physically available but functionally neglected or sub-optimally employed [4]. The land has low human employment levels relative to its potential, often due to excessive exploitation followed by abandonment [14]. Underutilized land is often covered with groves, shrubs, weeds, grass, or open fields. For instance, in Indonesia, a large part of the land that remains unused for 3 to 5 years is turned over to growing grasses and shrubs, which is categorized as underutilized land [4]. Socioeconomic factors contribute to land underutilization, including land fragmentation, distance from modern infrastructure, rural depopulation, and aging workforces [1]. Although the terms “marginal” and “underutilized” are sometimes used interchangeably, there is a major difference between them. Natural factors often constrain marginal land, whereas underutilized land might be suitable for agriculture but remains idle due to external factors.

2.2. Land-Use Change and Agricultural Land Decline

The global agricultural industry is going through great transformations induced by a combination of demographic, economic, and environmental pressures, with a demand projected to double by 2050 [3]. Land-use change refers to the modification in utilization or management of a particular area of land, such as converting a forest into a farm or a rural field into an urban center [17]. Key concepts of land-use change include human-driven processes, particularly population growth, that drive demand of land for agriculture, housing, and energy [18]. Agricultural land decline occurs when the total land dedicated to farming decreases. While many areas of the world experience agricultural expansion at the expense of deforestation, other areas face decline due to urbanization, policy shifts, or environmental factors [17].

Primary drivers of land-use changes include the following: (1) Population growth and urbanization—in Indonesia, paddy fields are converted to non-agricultural uses at a rate of 90,000 to 110,000 hectares per year, with 75% of the fields shifted to housing [19]. (2) Economic marginality—land is usually abandoned when cost-effective production is no longer possible under the current market conditions. These conditions can be caused by fluctuating commodity prices, rising production costs, or inflation [2]. (3) Land fragmentation—agricultural productivity is threatened by fragmented ownership, often due to inheritance patterns, in countries such as China and Slovakia. Smallholder ownership makes efficient production difficult, which leads to an increase in abandoned lands [4]. (4) Environmental degradation—in the lower Mississippi Alluvial Valley, intensive row cropping and chemical fertilizer use have reduced the soil's ability to grow crops, leading to marginality [2].

Contaminated land, such as brownfields and former industrial sites, represents a specific category of constrained land, which is unsuitable for primary uses such as food production without significant remediation [11]. Identifying contaminated land requires an assessment of biophysical and anthropogenic factors that limit land productivity and usability [3]. Failure to account for contaminated land can lead to overestimating available land resources, resulting in inappropriate land-use recommendations.

2.3. Integration of Biophysical and Socioeconomic Factors

The combination of biophysical and socioeconomic factors as resource standards is recommended for robust land-use suitability analysis [20,21] which transitions from a purely technical, data-driven approach to a collaborative, multifaceted planning framework [14]. This approach combines objective biophysical factors with subjective human values and economic realities to yield accurate, acceptable results. GIS-MCDA and machine learning approaches are two ways of integration. GIS-MCDA provides tools for storing and manipulating spatial data. At the same time, MCDA, for example, the WLC method, multiplies normalized weights by standardized factor values to create a composite stability score [1].

Machine learning, such as with random forests or MaxEnt, is a modern approach that uses algorithms to integrate complex datasets. A typical study on land suitability modeling was conducted titled Land suitability evaluation for multiple crop agroforestry planning using GIS and multi-criteria decision analysis: A case study in Fiji [22], which integrated geospatial data and artificial intelligence by building a data cube of 100 geospatial layers, including climate, water, and accessibility, to predict land suitability for agro-silvopastoral systems with high accuracy. The Integration of biophysical and socioeconomic factors gives a clearer understanding of the land compared to the use of separate data alone. In addition, it helps to account for constraints like political, social, and environmental factors [3].

A GIS-based study of agricultural land suitability relies mainly on biophysical and accessibility criteria. For instance, Garbaba and Wolteji evaluated wheat and maize suitability using elevation, slope, soil properties, drainage, climate, land cover, and road proximity but did not incorporate community-level socioeconomic vulnerability or contaminated land indicators [23]. In another study, GIS-based agricultural suitability assessments in Sudan and India primarily considered climate, soil, topography, land cover, groundwater, and hydrological conditions [24,25]. This study demonstrates the effectiveness of GIS-MCDA for evaluating physical agricultural potential and assessing environmental and terrain-based criteria.

Recent studies on land-use applications have begun to include socioeconomic and infrastructural variables. However, they are commonly designed for specific facilities or urban development suitability rather than for identifying marginal and underutilized lands [23,26]. Therefore, the simultaneous treatment of biophysical limitations, socioeconomic vulnerability, land underutilization, and site-specific environmental contaminations remains insufficiently developed, particularly at the county scale.

2.4. Research Gap

GIS-based MCDA has been applied to agricultural land suitability assessment. Many of these studies define suitability in terms of biophysical conditions (soil properties, slope, climate, hydrology, land use, and access to infrastructure). As an example, recent assessments of wheat, maize, irrigated agriculture, and rain-fed farming have relied mainly on environmental and terrain-related variables and indicators, such as household income, unemployment rate, population vulnerability, or contaminated lands, which are not incorporated within the same analytical framework [24,25,27]. Resulting approaches for estimating physical agricultural potential may provide only a partial representation of land marginality, which may also result from urban pressure, social disadvantage, and anthropogenic environmental constraints.

A second limitation concerns the representation of contaminated and brownfields. Suitability models for agricultural use usually do not include developed land, water, or protected areas, and contaminated areas are rarely treated as a distinct constraint. If contaminated sites and brownfields are not identified independently, the amount of land available for immediate agricultural production may be overestimated because these areas typically require site investigation, remediation, or crop-use restrictions before they can be safely repurposed. A third limitation concerns scale and context. Many of the suitability studies under-taken are conducted at large scales: national, watershed, regional, or even crop-specific. While this provides a wealth of general knowledge, there is a risk of missing county-level interactions involving urban growth, loss of farmland and wetlands, soil constraints, social-economic vulnerability, and contamination. Such a problem would be especially relevant in Leon County, which has experienced an increase in the number of farms between 2017 and 2022, along with a decrease in farmland area from about 91,730 acres to 70,058 acres.

The current study attempts to fill these gaps through the development of a county-scale GIS-MCDA approach which considers six aspects of marginality, including hydrological limitations, soil quality, marginality of land cover, slope, socioeconomic vulnerability, and contaminated land. In contrast to traditional approaches that focus on defining land based on its agricultural potential, the current study assesses marginality as a multidimensional concept arising from interactions among various environmental, social, and anthropogenic constraints. Through the application of the AHP weights, sensitivity analysis, correlation analysis, and validation via the random forest technique, the selected criteria are evaluated in terms of their stability and impact on marginality. Leon County is thus not only used

as an illustrative example for the research question but as a site for the evaluation of a transferable framework for various land-use activities.

3. Materials and Methods

3.1. Study Area

This study covers Leon County, Florida, which is located in the United States' Florida Panhandle (latitude 30.44° N, longitude 84.28° W) (Figure 1). The county encompasses a mix of urban, peri-urban, and rural landscapes, with Tallahassee as the state's capital and primary urban center. According to the U.S. Census Bureau, the county covers 702 sq mi (1818.17 km²), of which 667 sq mi (1727.52 km²) is land, and the rest is water. As of the 2020 Census, the total population is 292,198. About 19% of the residents were under 18, and 14.7% were 65 or older. From an environmental and climatic standpoint, the county is shaped by extensive riparian wetlands, seasonal floodplains, karst topography, and a longleaf pine ecosystem, with mean annual precipitation exceeding 1400 mm.

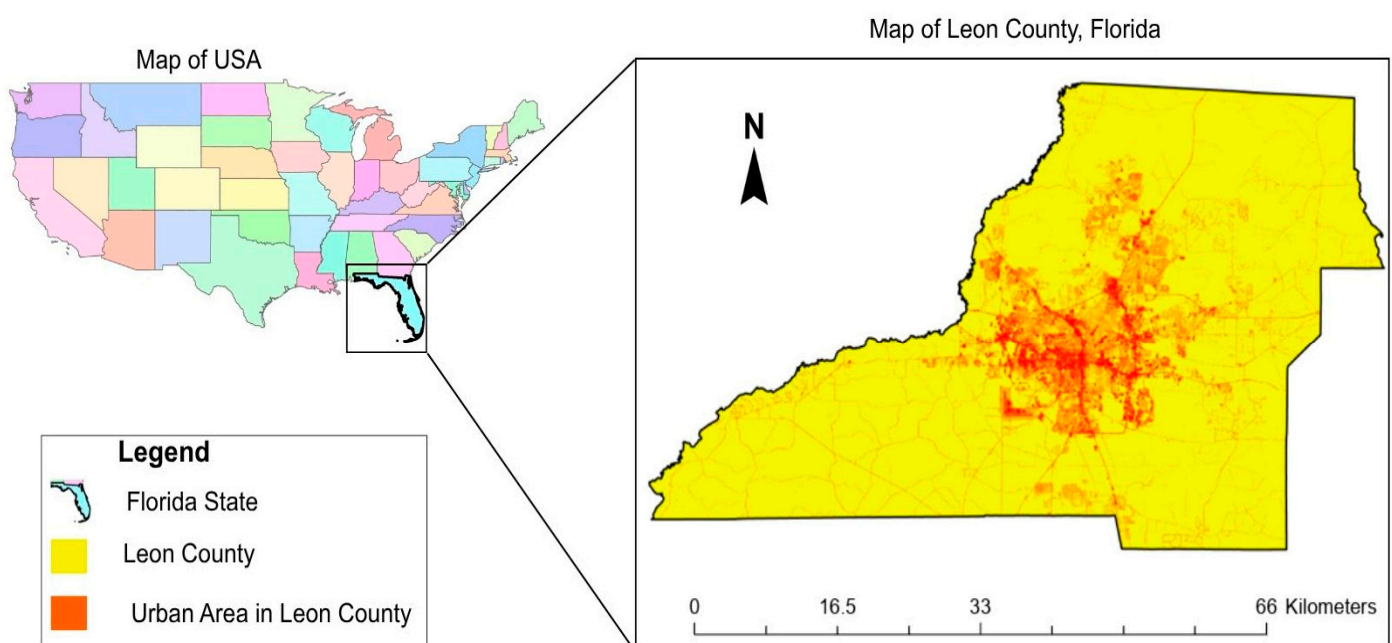


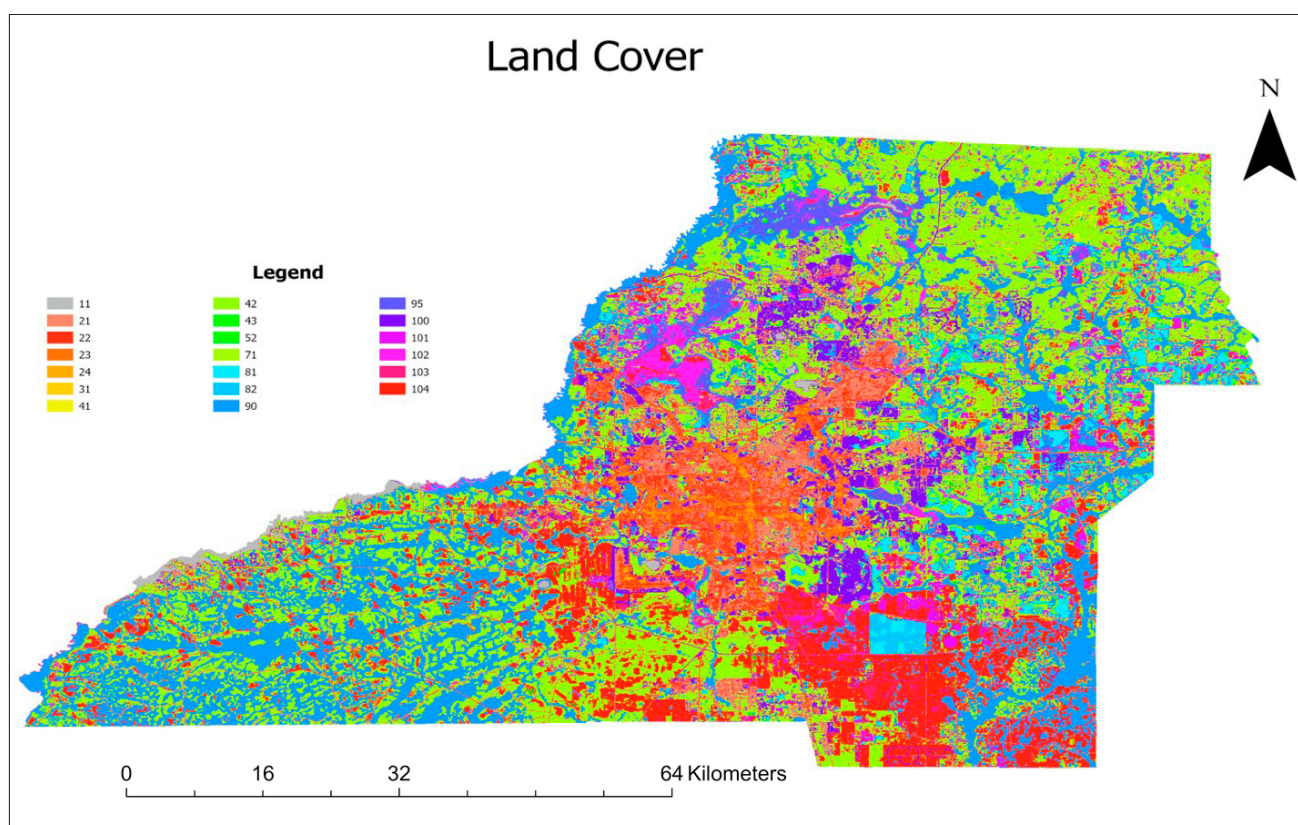
Figure 1. Study area: Leon County, Florida, USA.

Leon County was selected as the study area due to the following reasons:

1. The county faces significant urban growth pressure—according to USDA data (2017–2021), the number of farms increased from 325 to 375 (approximately 14%). However, total farmland acreage declined from 91,730 acres to 70,058 acres (approximately 24%).
2. The county features multiple land cover features (Figure 2). Multiple environmental constraints reduce conventional agriculture use but create opportunities for sustainable land use, such as ecological restoration, carbon sequestration, and biomass production. Table 1 shows the land cover classification codes used in this study.
3. No prior study has combined environmental constraints with socioeconomic constraints to map marginal and underutilized land, making Leon County an ideal candidate for developing a transferable analytical framework.

Table 1. NLCD land cover classification codes used in this study.

Code	Land Cover Class	Code	Land Cover Class
11	Open Water	71	Grassland/Herbaceous
21	Developed, Open Space	81	Pasture/Hay
22	Developed, Low Intensity	82	Cultivated Crops
23	Developed, Medium Intensity	90	Woody Wetlands
24	Developed, High Intensity	95	Emergent Herbaceous Wetlands
31	Barren Land	100	Moss
41	Deciduous Forest	101	Lichens
42	Evergreen Forest	102	Palustrine Forested Wetland
43	Mixed Forest	103	Palustrine Scrub/Shrub Wetland
52	Shrub/Scrub	104	Estuarine Forested Wetland

**Figure 2.** Multiple land cover features' classification in Leon County.

3.2. Data Sources

The analysis integrated both biophysical and socioeconomic datasets to capture multiple dimensions of land suitability. The datasets used in this study are summarized in Table 2, as follows:

Table 2. Spatial datasets, data sources, spatial resolution, and analytical purpose used in the multi-criteria marginal land suitability analysis.

Datasets	Sources	Resolution	Uses
County and Census Boundaries	U.S. Census TIGER/Line database https://www.census.gov (accessed on 29 April 2026)	Vector	Define the study area

Table 2. Cont.

Datasets	Sources	Resolution	Uses
Land Cover Data	The National Land Cover Database (NLCD) 2021 is a 30 m resolution Landsat-based dataset. https://www.usgs.gov (accessed on 29 April 2026)	30 m	Identifying land-use classes, wetlands, water bodies, and developed areas.
Soil Data	Soil Survey Geographic Database (SSURGO) https://www.nrcs.usda.gov/ (accessed on 29 April 2026)	Vector/30 m	Providing detailed soil classification and farmland suitability information.
Elevation Data	USGS 3D Elevation Program (3DEP) at 1/3 arc-second resolution. https://www.usgs.gov/ (accessed on 29 April 2026)	1/3 arc-sec (~10 m)	Derive slope as a topographic constraint.
Socioeconomic Data	American Community Survey (ACS) https://www.census.gov/ (accessed on 29 April 2026)	Census tract	Provide information on population, income, and unemployment indicators.
Contaminated and brownfield site	Florida Department of Environmental Protection (FDEP) https://floridadep.gov/ (accessed on 29 April 2026)	Point/Vector	Identify environmental constraints.

All datasets were harmonized with respect to projection, spatial extent, and resolution prior to analysis. ESRI ArcGIS Pro Version 3.6.0 was used for data processing, raster standardization, terrain analysis, and multi-criteria modeling, including Project Raster, Extract by Mask, Reclassify, Raster Calculator, and Slope. These tools collectively allow for the development of a weighted marginality surface through spatial overlay analysis. A 30 m cell size, Common Snap raster for the NLCD dataset, and NAD 1983 State Plane Florida North FIPS 0930 feet were used to ensure projection consistency.

3.3. Marginality Criteria Development

Land marginality was developed and operationalized as a spatial layer based on six criteria. Higher values in each layer indicate stronger environmental and socioeconomic constraints, indicating higher marginality and reduced land productivity and usability.

3.4. Reclassification and Standardization

Variables, including soil suitability, slope, hydrological conditions, land cover, contamination, and socioeconomic factors, were standardized to a common marginality scale ranging from 1 (lowest marginality) to 10 (highest marginality) to enable the integration within a multi-criteria framework using ArcGIS Pro's reclassify tool. A cell with a score of 10 in any layer indicates the most constrained condition for that criterion, while a score of 1 indicates the least constrained. Table 3 describes the reclassification scheme for marginality standardization.

Table 3. Reclassification scheme for marginality standardization. All layers standardized to a 1–10 scale; higher values indicate greater marginality.

Variable	Original Condition	Marginality Score (1–10)	Interpretation
Slope	0–2%	2	Flat; minimal erosion risk
	2–5%	4	Gentle; low constraint
	5–10%	6	Moderate; farming efficiency reduced
	>10%	10	Steep; high erosion,

Table 3. *Cont.*

Variable	Original Condition	Marginality Score (1–10)	Interpretation
Soil	Prime farmland	2	Highest productivity
	Prime farmland if drained	5	Conditional productivity
	Not prime farmland	10	Severe soil constraints
Land Cover	Cultivated crops/Pasture	2	Actively productive land
	Shrub/Scrub, Grassland, Barren	5–7	Low-productivity cover
	Woody/Emergent Wetlands	10	Maximum hydrological constraint
	Developed medium/high intensity	Excluded	Not available for intervention
Contamination	Non-contaminated	1	No restriction
	Contaminated/Brownfield site	10	Remediation required
Social (Income, Unemployment Rate, and Population)	Highest median household income	1	Low vulnerability
	Lowest median household income	10	High vulnerability

Reclassification values were calculated through a combination of published suitability criteria for land, class-specific criteria, and subjective weight given to the degree to which each variable contributes to land marginality. Soil classes were ranked by agricultural productivity potential; land cover classes by agricultural suitability; hydrologic and slope classes by increasing physical constraints; contaminated land sites assigned by the maximum marginality values due to environmental constraints; and socioeconomic variables by the degree of socioeconomic disadvantage. The 1-to-10 normalized score measures relative, not absolute, marginality and allows integration of different datasets into the GIS-based MCDA model.

3.5. Analytical Hierarchy Process (AHP) Weighting

Analytic hierarchy process (AHP), a structured mathematical technique for multi-criteria decision-making (MCDM) to derive relative weights of importance for various criteria through human judgment, was used in this study [15,28]. AHP was developed by Thomas L. Saaty in the 1970s and is widely utilized in GIS to assess land suitability by integrating objective spatial data with subjective value judgments [29,30].

The pairwise comparison matrix was developed by the authors based on the published land suitability literature, established GIS-based Multi-Criteria Decision Analysis (MCDA) methodologies, and the objectives of the present study [15,28]. Reclassification values were calculated through a combination of published suitability criteria for land, class-specific criteria, and subjective weight given to the degree to which each variable contributes to land marginality. Soil classes were ranked by agricultural productivity potential; land cover classes by agricultural suitability; hydrologic and slope classes by increasing physical constraints; contaminated land sites assigned by the maximum marginality value due to environmental constraints; and socioeconomic variables by the degree of socioeconomic disadvantage. The 1-to-10 normalized score measures relative, not absolute, marginality and enables the integration of different datasets into the GIS-based MCDA model. Each cell in the matrix represents the relative importance of the row criterion compared to the column criterion. The matrix was then normalized by dividing each cell by its column sum, and criterion weights were derived as the row averages of the normalized matrix [31]. The detailed pairwise comparison is shown in Table 4, and Table 5 shows the AHP pairwise-derived criterion weights following matrix normalization.

Table 4. AHP pairwise comparison matrix showing relative criterion importance (Saaty scale).

Criteria	Hydrology	Soil	Land C	Slope	Social	Conta.
Hydrology	1.00	2.00	3.00	3.00	5.00	5.00
Soil	0.50	1.00	2.00	3.00	3.00	5.00
Land C.	0.33	0.50	1.00	3.00	3.00	5.00
Slope	0.33	0.33	0.33	1.00	1.00	3.00
Social	0.20	0.33	0.33	1.00	1.00	3.00
Conta.	0.20	0.20	0.20	0.33	0.33	1.00
SUM	2.57	4.37	6.87	11.33	13.33	22.00

Table 5. AHP pairwise-derived criterion weights following matrix normalization.

Criteria	Hydrology	Soil	Land C.	Slope	Social	Conta.	Average	Final Weight
Hydrology	0.39	0.46	0.44	0.26	0.38	0.23	0.36	36%
Soil	0.19	0.23	0.29	0.26	0.23	0.23	0.24	24%
Land C.	0.13	0.11	0.15	0.26	0.23	0.23	0.18	18%
Slope	0.13	0.08	0.05	0.09	0.08	0.14	0.09	9%
Social	0.08	0.08	0.05	0.09	0.08	0.14	0.08	8%
Conta.	0.08	0.05	0.03	0.03	0.03	0.05	0.04	4%

Consistency Analysis of the AHP Matrix

The responses were organized into a pairwise comparison matrix [32]. The weights were derived by calculating the principal eigenvector corresponding to the largest eigenvalue (λ_{max}) of the matrix using (Equation (1)) and then normalized so the weight summed to 1 [28]:

$$\lambda_{max} = \frac{\left(\sum \frac{Aw}{w} \right)}{n} \quad (1)$$

where Aw is the weighted sum vector, w is the weight vector, and n is the number of criteria. The consistency of the pairwise judgment was evaluated using the consistency ratio (CR) (Equation (2)), which provides a measure of the degree of logic and consistency in the expert judgment [33]. CR was calculated by comparing the inconsistency of the matrix to a random index. If CR is less than 0.1 (10%), the judgments are considered consistent, and the weights are reliable [34].

$$CR = \frac{CI}{RI} \quad (2)$$

where $CI = (\lambda_{max} - 1) / (n - 1)$ and RI is the Random Consistency Index for a matrix of size n (RI = 1.24 for $n = 6$). The consistency ratio (CR) was below the acceptable threshold of 0.10, $CR \approx 0.037$, indicating logical consistency in the weighting structure and that the pairwise judgments were not random.

3.6. Weighted Linear Combination (WLC) Model

In spatial analysis, derived weights were applied to map layers using the WLC method [1]. Each pixel's value in a criterion map was multiplied by its AHP weight, and the results were summed to produce a final suitability index map [32,35]. The Weighted Linear Combination algorithm (Equation (3)), within the class of local operators, was used to obtain the suitable marginal degree for each cell in the study area.

$$M = \sum_{i=0}^n W_i X_i \quad (3)$$

where M is the final weighted marginality score, W_i is the weight assigned to criterion I , X_i is the standardized marginality score for criterion I , and n is the number of criteria.

Weights were assigned based on the relative importance of each factor in determining land marginality. The weighting system was established based on the impact each criterion had on the degree of land marginality, as determined from past GIS-MCDA analyses of suitability studies. The factors related to hydrology and soils were assigned higher weights due to their significant impacts on limiting agriculture and protecting the environment. The applied model is shown in Equation (4):

$$MLI = (Hydro \times 0.36) + (Soil \times 0.24) + (LandCover \times 0.18) + (Slope \times 0.09) + (Social \times 0.08) + (Contam \times 0.04) \quad (4)$$

MLI was computed cell-by-cell across the full county extent using the ArcGIS Pro Raster Calculator. The resulting continuous surface ranges from approximately 2.4 to 8.9, representing a gradient from minimally constrained land (low MLI) to maximally constrained or underutilized land (high MLI). The composite raster was then classified into five marginality categories using equal-interval classification, ranging from Class 1 (very low marginality) to Class 5 (high marginality/high underutilization potential). Urban areas (medium- and high-intensity developed land) and open water bodies were masked from the final output using ArcGIS Pro's Extract by Mask tool, ensuring that the model focuses exclusively on realistically available land surfaces.

3.7. Sensitivity and Spatial Similarity Analysis

Sensitivity analysis (SA), a critical stage in evaluating MCDA models, was used to check the reliability and robustness of the results by assessing how uncertainties in input parameters affected the final output [36].

Four alternative weighting scenarios (Table 6) were developed and compared with the baseline AHP-derived scenario by evaluating (1) the total area classified as high marginality (Classes 4 and 5), (2) the spatial overlap of high-marginality areas between scenarios, and (3) changes in the spatial distribution of marginality hotspots across the county.

Table 6. Sensitivity analysis weighting scenarios. S1 represents the baseline AHP model; S2–S5 test alternative weighting assumptions.

Scenario	Hydro.	Soil	Land Cover	Slope	Social	Contam.	Planning Interpretation
S1: Baseline (AHP)	0.36	0.24	0.18	0.09	0.08	0.04	Expert-informed baseline
S2: Hydrology Emphasis	0.50	0.20	0.15	0.07	0.05	0.03	Watershed protection and flood risk planning
S3: Soil Emphasis	0.20	0.45	0.15	0.09	0.08	0.03	Agricultural capability and soil-productivity planning
S4: Social Emphasis	0.20	0.20	0.15	0.10	0.30	0.05	Social equity and vulnerability-oriented planning
S5: Equal Weights	0.166	0.166	0.166	0.166	0.166	0.166	Neutral benchmark without preferential weighting

Spatial similarity was evaluated using the Jaccard Similarity Index. For this analysis, each scenario was converted into a binary raster in which Classes 4 and 5 were assigned a value of 1, while Classes 1–3 were assigned a value of 0. The Jaccard Similarity Index was calculated using Equation (5):

$$J = \frac{|(A \cap B)|}{|(A \cup B)|} \quad (5)$$

where A represents the cells classified as Classes 4 and 5 in the baseline AHP scenario, B represents the cells classified as Classes 4 and 5 in an alternative scenario, $A \cap B$ represents

the cells classified as highly marginal in both scenarios, and $A \cup B$ represents the cells classified as highly marginal in either scenario. Jaccard values range from 0 to 1, where 0 indicates no spatial overlap and 1 indicates identical spatial distributions.

3.8. Correlation and Multicollinearity Analysis

Using ArcGIS Pro's Create Random Points tool, 1000 random points were distributed across the study area to assess the relationships among individual layers and the final composite MLI (Figure 3). The analysis is based on post-model diagnostics rather than as an independent validation of the AHP-derived marginality land index. Given that the final index was a result of a weighted combination of the same input criteria, a degree of relationship between these individual criteria and the final index was anticipated. Correlation analysis would help determine the nature and strength of these relationships and their consistency with the AHP weights, while also pointing out any redundancies among the input criteria.

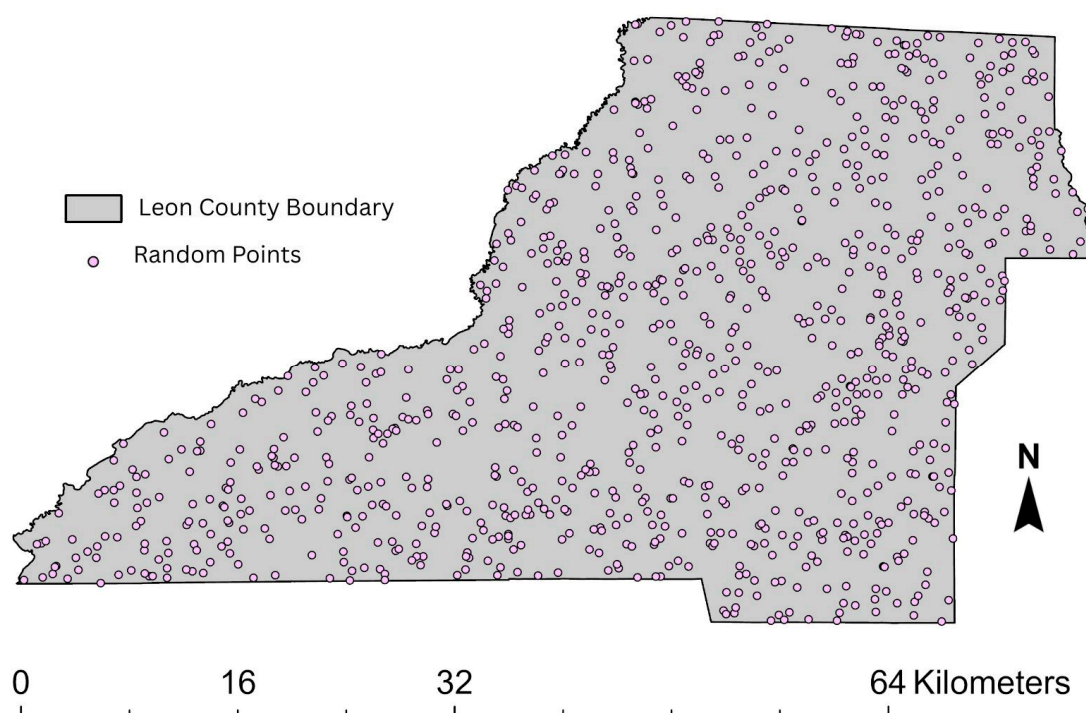


Figure 3. Distribution of 1000 random points across the study area.

The Pearson correlation coefficient was determined between hydrological marginality, soil marginality, land cover marginality, slope marginality, social marginality, contaminated land marginality, and the final AHP-based marginality index. Statistical significance of each correlation coefficient was determined at the $p < 0.05$ level using a two-tailed test.

Because pairwise correlation coefficients may not fully detect multicollinearity involving combinations of several predictors, the Variance Inflation Factor was also calculated for each of the six input criteria. VIF was calculated using Equation (6):

$$VIF_i = \frac{1}{1 - R_i^2} \quad (6)$$

where R_i^2 is the coefficient of determination obtained by regressing predictor i on all remaining predictor variables. A VIF below 5 was considered indicative of the absence of problematic multicollinearity.

3.9. Machine Learning Model Validation

GIS-based machine learning validation, a multi-step process that combines traditional statistical metrics with advanced spatial strategies, is essential to ensure model reliability [37,38]. Especially, random forest, a commonly used machine learning algorithm, was used in this study by constructing multiple decision trees from bootstrap samples of the training data and aggregating their predictions [37].

A Random Forest regression model was implemented in Google Colaboratory using Python version 3.12.13 and Scikit-learn version 1.5.2. The predictor variables included five standardized raster variables (Hydro, Soil, Slope, Social, and Contam) from 1000 random sample points, which served as the modeling dataset. The AHP-derived marginality index served as the response variable. Prior to model development, all predictor variables were standardized and checked for multicollinearity.

The dataset was randomly divided into 80% training and 20% testing subsets using a fixed random seed of 42 to ensure reproducibility. A random forest regressor consisting of 200 decision trees was trained on the training dataset. Model performance was evaluated using the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). In addition, feature importance scores were calculated to quantify the relative contribution of each predictor variable to the final marginality index.

3.10. Future Scenario Modeling

To evaluate how continued urban growth would unfold, a 2035-based scenario was developed to project future land expansion and land use. The year 2035 was selected as a medium-term planning horizon, providing sufficient time for measurable land-use changes while remaining relevant to current land management and policy decisions. The 2035 scenario was designed as a what-if planning exercise to assess the suitability model's sensitivity to plausible urban expansion, rather than as a prediction of future urban development.

Previously developed areas were extended by a 2-mile (3.2 km) buffer to account for potential future urban expansion. This buffer size was chosen based on assumptions about the radius of urban development around previously developed lands and did not aim to predict the actual boundaries of future urban expansion; rather, it is a standardized tool for assessing how further urban expansion might limit the availability of marginal and underutilized lands for agriculture.

These urban lands were excluded from the assessment of future suitability for agricultural development and were used in the GIS multi-criteria analysis to estimate the 2035 land suitability scenario. The comparison of these two scenarios helped evaluate the potential impact of urban expansion on land availability. The methodological workflow is shown in Figure 4.

3.11. Area Estimation

The area of each class was calculated using raster attribute tables. Cell counts were converted to area using the raster's spatial resolution, and results were expressed in acres using Equation (7) and as percentages of the total study area using Equation (8), which enabled quantitative comparison of land categories:

$$\text{Area (acres)} = \frac{\text{Count} \times \text{Area of one 30 m raster cell (900 m)}}{\text{Square meters in 1 acre (4046.85)}} \quad (7)$$

$$\text{Percent (\%) Area} = \frac{\text{Area Total}}{\text{Area Class}} \times 100 \quad (8)$$

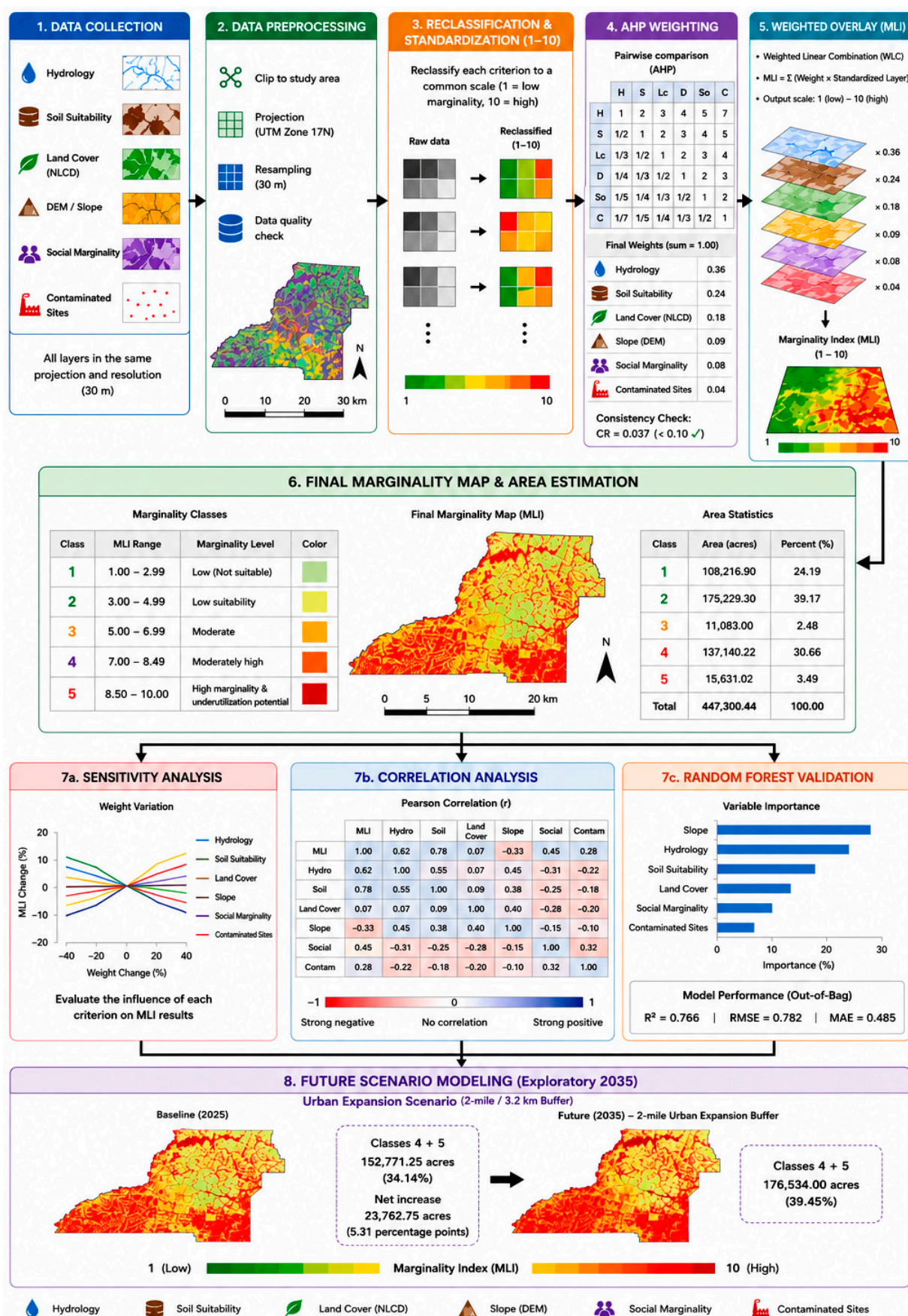


Figure 4. Overall workflow of the GIS-MCDA framework for identifying marginal and potentially underutilized lands in Leon County, Florida. Arrows indicate the sequence of processing steps, and the green-to-red color gradient represents increasing marginality from low to high.

4. Results

4.1. Individual Marginality Factors

Hydrological marginality is shown in Figure 5a, where darker areas indicate a marginality score of 10, and a score of 9 indicates higher hydrological constraints, including wetlands, saturated soils, flood-prone areas, and water-related limitations. A lighter color indicates drier, more stable conditions and is considered lower-marginality. The hydrological constraints are widespread across the county, especially along drainage networks and wetlands. Excessive moisture and flooding reduce land suitability and lead to land marginality.

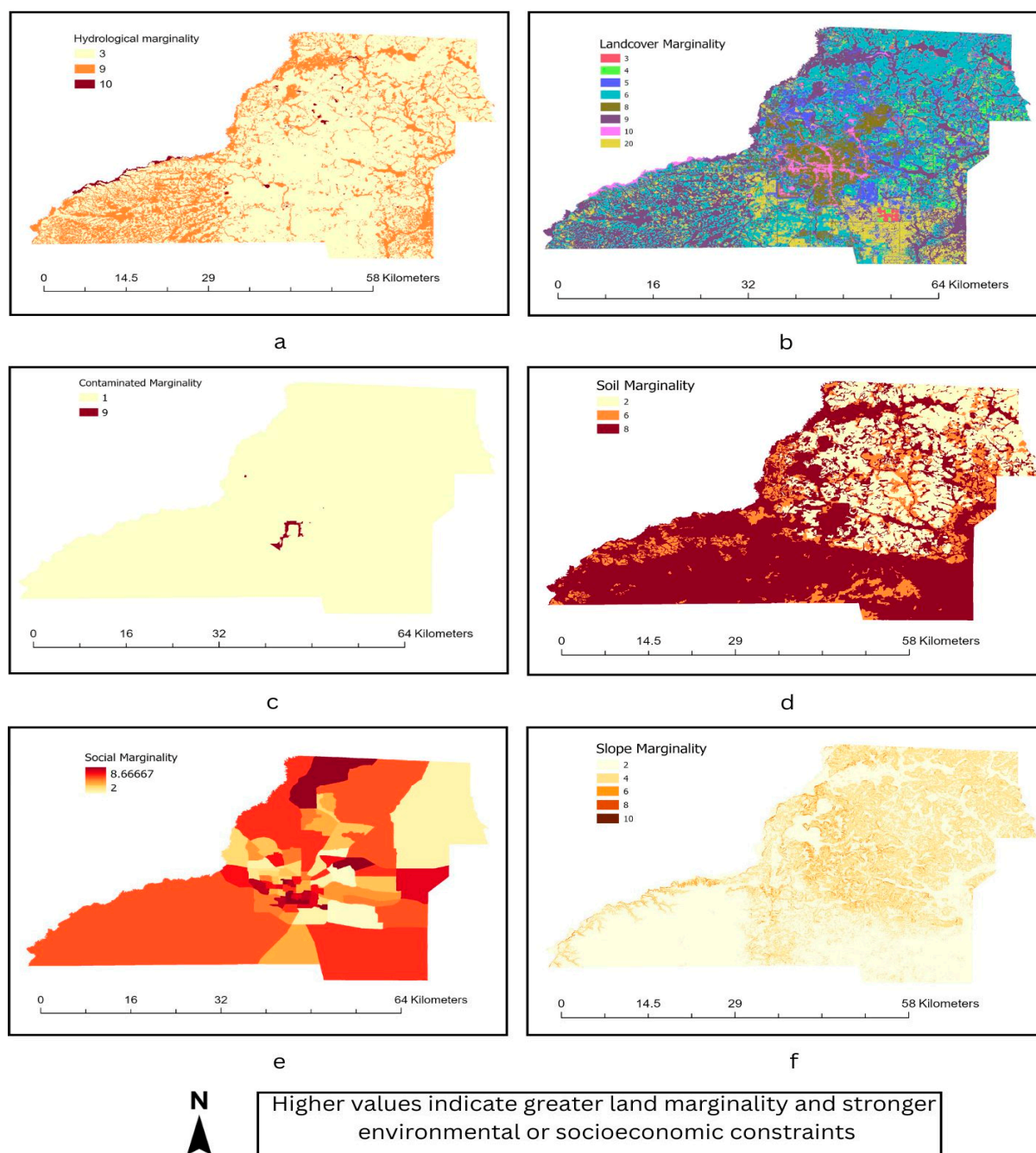


Figure 5. Input layers used in the GIS-based marginal land suitability assessment for Leon County, Florida, including (a) hydrological marginality, (b) land cover marginality, (c) contaminated land constraints, (d) soil marginality, (e) social marginality, and (f) slope marginality.

Figure 5b presents the land cover marginality. Higher marginality patterns are concentrated in urbanized and environmentally restricted landscapes, which are related to developed lands, dense forest cover, water bodies, and wetlands, while lower marginality levels are observed in agriculturally usable lands. Figure 5c shows the spatial distribution of contaminated and non-contaminated land within Leon County. The red shows the contaminated area, while the gray shows the non-contaminated area. The contaminated areas are concentrated primarily within the central portion of the county near the urbanized area. Figure 5d illustrates the spatial distribution of the soil. Areas with poor soil and drainage limitations are distributed in the southern part of the county and exhibit higher marginality. Lower marginality values are associated with more suitable agricultural soils.

Figure 5e shows the spatial distribution of social marginality across Leon County. Lighter yellow tones represent lower social marginality concentrated in eastern and central Leon County. These areas highly correspond to favorable socioeconomic conditions. Dark red and orange tones distribute across the central, northern, and southern parts of the study area, indicating greater social marginality and indicating multiple socioeconomic limitations. Figure 5f presents the spatial distribution of slope-derived marginality across the county based on the DEM-based terrain analysis. Orange and red tones in steeper terrain correspond to higher marginality, and yellow tones represent lower values in flatter areas. Steeper areas are considered more marginal because steeper slopes increase erosion risk, reduce mechanized farming efficiency, and limit land use.

4.2. Final Marginality Classification

The GIS-based weighted overlay analysis successfully identified the spatial pattern of marginal and underutilized land in Leon County, Florida. The outputs reveal spatial variation in marginality conditions across the county, influenced by both biophysical and socioeconomic constraints. The candidate land is classified into five distinct categories according to levels of marginality and potential underutilization. The categories are divided into Class 1, representing low marginality and high suitability, and Class 5, denoting very high marginality and high underutilization potential. Based on the results, high marginality is more commonly associated with hydrologic constraints, poor soil quality, fragmented land cover, and lower socioeconomic suitability.

The weighted marginality surface shown in Figure 6 represents the integrated effect of multiple biophysical and socioeconomic variables, including soil suitability, hydrological condition, slope, land cover, contamination areas, and socioeconomic indicators. The resulting raster shows a range of values from 2.4 to 8.9. Higher marginality values are observed in the south and southwest portions of the county, indicating an overlap of multiple constraints. On the other side, lower marginality values are shown in the north regions of the county, suggesting relatively favorable conditions for existing land use.

The map in Figure 6 shows discernible spatial distributions of the various land classifications in Leon County. Table 7 below shows the area and percentage distribution of the marginality derived for the AHP baseline-weighted overlay model in Leon County, Florida.

Table 7. Area and percentage distribution of marginality classes.

Class	Category	Acres	Percent
1	Low Marginality/Highly Suitable	108,216.88	24.19
2	Slightly Marginal	175,229.34	39.17
3	Moderately Marginal	11,083.04	2.48
4	High Marginality	137,140.22	30.66
5	Very High Marginality and Underutilization Potential	15,631.02	3.5

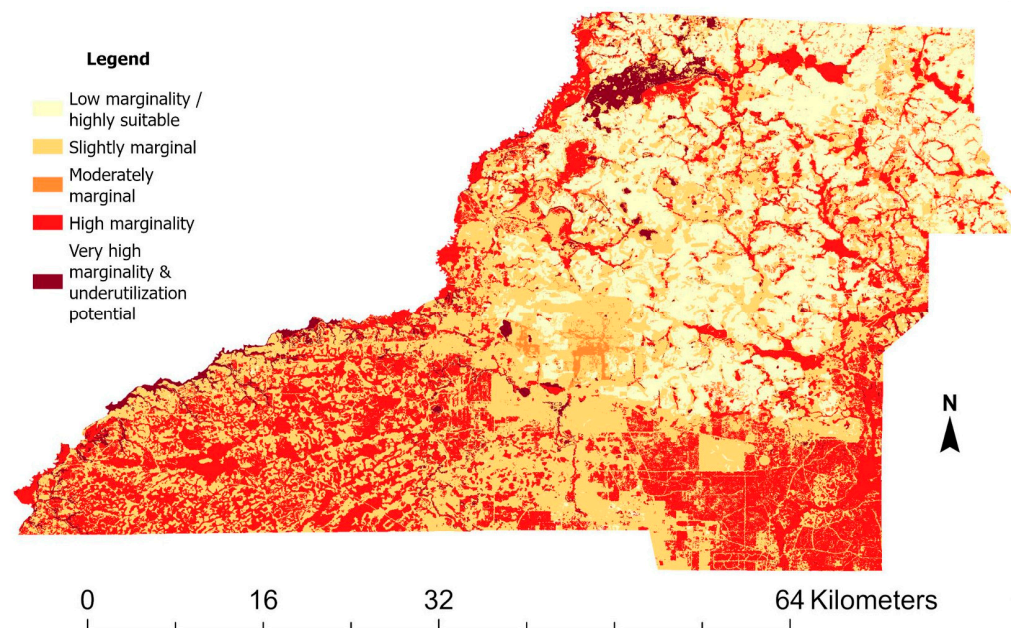


Figure 6. Spatial distribution of marginal and underutilized land suitability classes in Leon County, Florida.

The combined (high marginality) Class 4 and (very high marginality and underutilization potential) Class 5 categories represent approximately 152,771.24 acres, accounting for 34.16% of the total study area. In terms of their spatial location, they were scattered throughout the southern and eastern portions of the study area. Poor soil conditions, hydrological constraints, fragmented land cover, and low socioeconomic suitability often characterize these territories. The results of this final classification indicate that factors such as hydrologic limitations, land cover types, soil capabilities, and socioeconomics strongly influence land marginality in Leon County.

4.3. Sensitivity and Spatial Similarity Results

The sensitivity analysis showed that the estimated extent of high and very high marginality varied under the alternative weighting assumptions. The baseline AHP model classified approximately 152,771 acres as Classes 4 and 5. The soil-emphasis scenario identified approximately 151,887 acres, representing a difference of approximately 0.58% from the baseline. The social-emphasis scenario identified approximately 149,130 acres, representing a reduction of approximately 2.38%, while the hydrology-emphasis scenario identified approximately 142,258 acres, representing a reduction of approximately 6.88%. In contrast, the equal-weight scenario identified approximately 33,909 acres as Classes 4 and 5, representing a reduction of approximately 77.80% relative to the baseline.

Spatial agreement between the baseline AHP model and the alternative weighting scenarios was evaluated using the Jaccard Similarity Index (Table 8). The soil-emphasis scenario showed the highest spatial similarity with the baseline model ($J = 0.994$), followed by the social-emphasis scenario ($J = 0.976$) and the hydrology-emphasis scenario ($J = 0.913$). These high Jaccard values indicate that the locations classified as high and very high marginality remained highly consistent under plausible changes in soil, social, and hydrological weighting priorities.

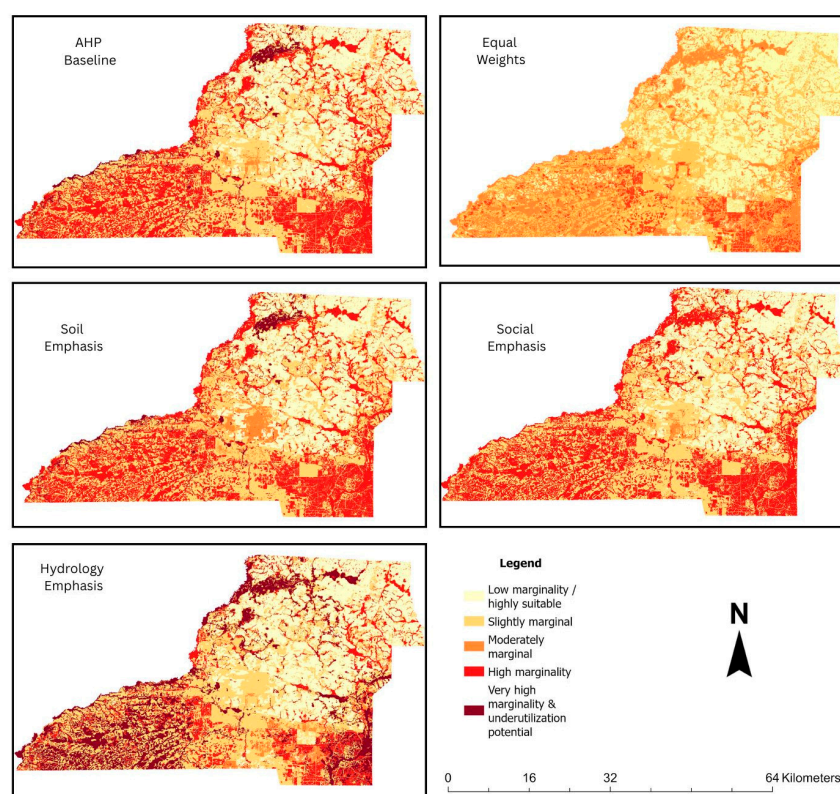
In contrast, the equal-weight scenario produced a substantially lower Jaccard value ($J = 0.219$), indicating limited spatial agreement with the baseline model. This result shows that assigning identical importance to all six criteria substantially altered both the estimated extent and spatial distribution of high-marginality land.

Table 8. Area variation and spatial similarity between the baseline AHP model and alternative weighting scenarios.

Scenario	Class 4 and 5 Area (Acres)	Difference from Baseline (%)	Intersection Cells	Jaccard Index
Baseline AHP	152,771.24	—	—	1.000
Hydrology emphasis	142,258	−6.88	633,088	0.913
Soil emphasis	151,887	−0.58	682,772	0.994
Social emphasis	149,130	−2.38	670,410	0.976
Equal weights	33,909	−77.80	150,745	0.219

Taken together, the acreage and Jaccard results indicate that the model was robust under the hydrology-, soil-, and social-emphasis scenarios but sensitive to the equal-weight configuration. The results also demonstrate that area-based comparisons alone are insufficient because scenarios with similar acreage may still differ spatially; the Jaccard analysis, therefore, provides an additional quantitative assessment of model stability. These results indicate that the estimated amount of high-marginality land remained relatively stable under the soil-, social-, and hydrology-emphasis scenarios but changed substantially under equal weighting. This suggests that the model is comparatively robust to plausible changes in the emphasis placed on major environmental and socioeconomic criteria, while the equal-weight configuration substantially altered the estimated extent of marginal land.

The scenario maps in Figure 7 also show changes in the geographic distribution of marginality classes. However, similar acreage values do not necessarily indicate that the same locations were classified in the same way. Therefore, the spatial comparison in this study is interpreted as descriptive rather than as a formal pixel-by-pixel agreement assessment. Quantitative spatial similarity measures, such as the Jaccard Index or Cohen's Kappa, would provide a stronger evaluation of map agreement and should be considered in future model development.

**Figure 7.** Sensitivity analysis results under different weighting scenarios, including the AHP baseline, equal weights, soil emphasis, social emphasis, and hydrology emphasis.

4.4. Correlation and Multicollinearity Results

Pearson correlation analysis reflects the relationship between each variable and the AHP-based final marginality model (Table 9). Hydrological marginality shows the strongest positive relation with the final model ($r = 0.743$, $p < 0.001$), followed by soil marginality ($r = 0.704$, $p < 0.001$) and land cover marginality ($r = 0.628$, $p < 0.001$). Slope marginality had a statistically significant negative association with the final index ($r = -0.332$, $p < 0.001$), reflecting the spatial distribution and standardized direction of slope constraints within the study area. In contrast, contaminated-land marginality showed a very weak and statistically nonsignificant association with the final index ($r = 0.019$, $p = 0.548$).

Table 9. Pearson correlation coefficients among the marginality criteria and the final AHP-derived marginality index.

Variable	Hydro.	Soil	Land Cover	Slope	Social	Contam.	Final AHP
Hydrology	1.00						
Soil	0.360 ***	1.00					
Land cover	0.070 *	0.260 ***	1.00				
Slope	−0.290 ***	−0.390 ***	−0.140 ***	1.00			
Social	0.120 ***	0.200 ***	0.100 **	−0.140 ***	1.00		
Contamination	−0.030	0.030	0.020	−0.030	0.140 ***	1.00	
Final AHP	0.743 ***	0.704 ***	0.628 ***	−0.332 ***	0.242 ***	0.019	1.00

Note: $n = 1000$. Two-tailed Pearson correlation tests were used. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Although the relationships among hydrology, soil, land cover, and the resulting index are fairly strong, they cannot be taken as independent evidence of the validity of the AHP model, since these parameters are included directly in the computation of the weighted index. The correlation analysis was carried out to determine whether the direction and intensity of the relationships corresponded to the weighting structure and to analyze the relationships among the criteria. It was revealed that the criteria with higher AHP weights had stronger associations with the resulting index, whereas the low-weight criterion of contaminated land showed practically no association.

According to the VIF analysis, there is no multicollinearity among the six input criteria (Table 10). The VIF values ranged from 1.02 to 1.37, well below the critical value of 5. The highest VIF, 1.37, was observed for soil marginality, followed by 1.22 and 1.19 for slope and hydrological marginality, respectively. The VIFs for land cover, social and contaminated land marginalities were 1.08, 1.07, and 1.02, respectively. These results indicate that none of the criteria can be explained well by a linear combination of the remaining predictors and support their simultaneous inclusion in the marginality assessment.

Table 10. Correlation and multicollinearity analysis of marginality criteria with the final AHP index.

Criterion	Pearson (r) with Final AHP Index	p-Value	VIF	Interpretation
Hydrological marginality	0.743	<0.001	1.19	Significant Positive Association; No Problematic Multicollinearity
Soil marginality	0.704	<0.001	1.37	Significant Positive Association; No Problematic Multicollinearity
Land-cover marginality	0.628	<0.001	1.08	Significant Positive Association; No Problematic Multicollinearity
Slope marginality	−0.332	<0.001	1.22	Significant Negative Association; No Problematic Multicollinearity

Table 10. Cont.

Criterion	Pearson (r) with Final AHP Index	p-Value	VIF	Interpretation
Social marginality	0.242	<0.001	1.07	Significant Positive Association; No Problematic Multicollinearity
Contaminated-land marginality	0.019	0.548	1.02	Nonsignificant Association; No Problematic Multicollinearity

4.5. Machine Learning Validation

The random forest regression model demonstrated high prediction accuracy when reproducing the AHP-based marginality index. Indeed, the coefficient of determination (R^2) was 0.766, indicating that approximately 76.6% of the variation in the marginality index was explained by the selected environmental and socioeconomic factors. The random forest regression model yielded an RMSE of 0.782 and an MAE of 0.485, indicating an acceptable prediction error for a spatial suitability model (Table 11).

Table 11. Random forest regression model configuration and performance.

Parameter	Model	Software	Training/Test Split	No. of Trees	Random Seed	R^2	RMSE	MAE
Value	Random Forest Regressor	Python (Scikit-learn)	80%/20%	200	42	0.766	0.782	0.485

Feature importance analysis showed that hydrological marginality was the most important variable among the predictors, while soil suitability and socioeconomic indicators came second, and contaminated sites were much less relevant in terms of their influence on the prediction performance of the model. Thus, the findings are consistent with the relative importance of the criteria identified using the AHP technique.

4.6. Future Scenario

The 2035 scenario should be interpreted as an exploratory planning scenario rather than a prediction of future land-use change. Its purpose is to evaluate the potential sensitivity of marginal land availability to continue urban expansion under a standardized set of assumptions.

An urban expansion scenario for the future was developed to assess potential shifts in marginal land distribution under anticipated development pressures in 2035. Table 9 illustrates scenarios that account for the impact of urban influence on existing developed lands and for a higher marginality rating in proposed urban expansion areas.

Based on the results of the future scenario, there was a marked increase in moderately high, highly marginal, and highly marginal lands in Leon County. The total of Class 4 and Class 5 land increased from about 152,771 acres (34.14%) to about 176,534 acres (36.59%) under the future urban expansion scenario in 2035, which indicates an estimated increase of about 23,762 acres of highly marginal lands (Table 12). Geographically, the increase in marginality occurred mostly in peri-urban areas near Tallahassee and along corridors of existing development (Figure 8). The urban expansion scenario shows that continued urban development could lead to increased land fragmentation and environmental/land-use restrictions in parts of the county.

Table 12. Comparison of current and projected 2035 marginality conditions showing changes in Class 4 and Class 5 land areas under the urban expansion scenario in Leon County, Florida.

Scenario	Class 4 Acres	Class 5 Acres	Class 4 + 5 Acres	Class 4 + 5%
Current AHP	137,140.22	15,631.02	152,771.25	34.14
Future 2035	1193.81	16,439.85	17,633.66	36.59

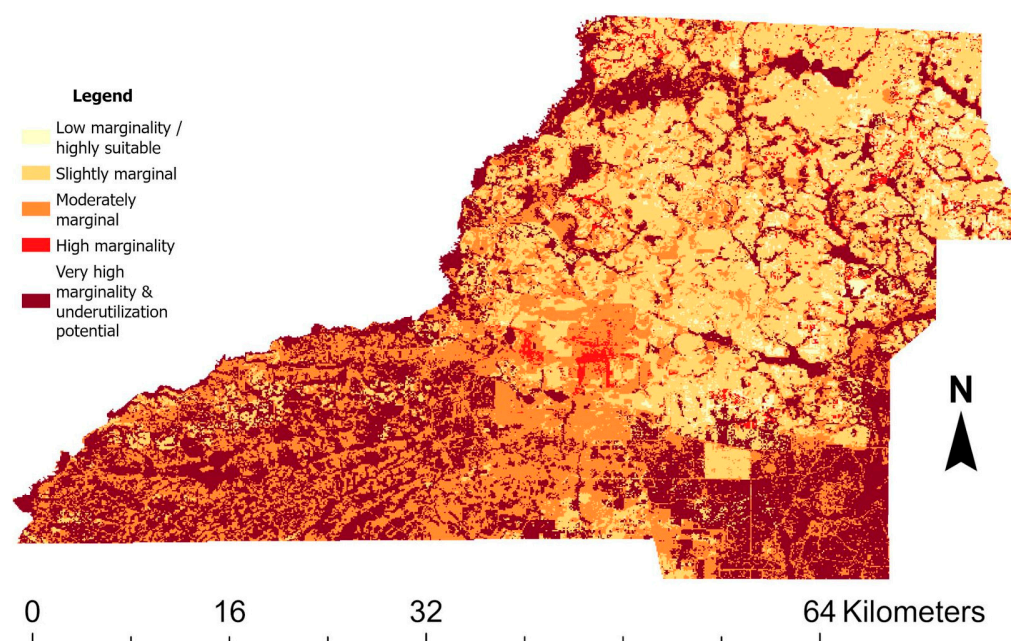


Figure 8. Future 2035 marginality scenario under projected urban expansion conditions in Leon County, Florida.

5. Discussion

The present study demonstrates the value of integrating environmental, socioeconomic, and contaminated-land information into a single GIS-based MCDA framework. The proposed framework incorporates socioeconomic vulnerability and environmental constraints, unlike many previous land suitability studies that emphasize biophysical variables. This integrated approach supports more balanced land-use planning by simultaneously considering agricultural suitability, environmental protection, and social equity.

This study uses environmental and socioeconomic indicators to identify marginal and underutilized lands in Leon County, Florida, by developing a GIS-MCDA framework. The results show that hydrological and social indicators are major drivers of marginality patterns across the study area. Areas characterized by wetlands, saturated soil, poor drainage, and flood-prone conditions show higher marginality, especially in the south and southwestern part of the study area. The study area was classified into five classes by the AHP weighted overlay baseline model: lower marginality/highly suitable (Class 1), slightly marginal (Class 2), moderately marginal (Class 3), high marginality and very high marginality (Class 4), and underutilization potential (Class 5). Approximately 34.14% of the study area falls under Classes 4 and 5. The most influential spatial constraint in the study area was hydrological marginality.

In the sensitivity analysis, four models were developed. Three models emphasized hydrology, soil, and social weights, resulting in relatively stable outcomes under alternative conditions to the AHP-weighted overlay baseline model. An equal-weight model for all indicators was also developed, resulting in a lower high marginality estimate, indicating that environmental and social scenarios cannot realistically be represented by equal weight.

A machine learning validation using random forest analysis further strengthens the dominance of hydrological and soil constraints in identifying hydrological marginality as a strong predictor, with approximately 92.2% of the combined total variable importance. The projection of marginal lands in 2035 was also developed and showed an increase in marginal land, suggesting that continued urban expansion may intensify fragmentation and environmental constraints in the study area.

5.1. Policy Implications and Sustainability

The GIS-based methodology framework has integrated multiple environmental and socioeconomic datasets to identify highly marginal and underutilized land. Higher marginality is associated with areas coinciding with wetlands, poorly drained soils, fragmented land cover, and lower socioeconomic activity. These areas are not suitable for conventional land use, but other forms of sustainable utilization may still be viable. This suggests opportunities for alternative and suitable land management strategies in Leon County, Florida. This study highlights that marginal land need not be interpreted as unusable.

This framework can support land-use planning, conservation prioritization, remediation efforts, and sustainable resource management at the local and regional levels. In relation to the future urban growth scenario, it is notable how much the peri-urban development pressures are beginning to affect land marginality. As seen in the predicted increase of around 23,762 acres of Class 4 and Class 5 lands, continued outward growth of urban areas near Tallahassee will lead to greater land fragmentation and decreased land-use flexibility over time. Growing development pressures can also lead to increased land sensitivity and underdevelopment. Based on sensitivity analysis, the developed GIS-MCDA model shows considerable resistance to realistic changes in the weighting scheme. Due to the stability within different weight variations, it is safe to conclude that the spatial distribution of marginal lands will not be affected in any noticeable way due to slight adjustments in the weighting scheme.

5.2. Strengths and Limitations of the Study

One of the strengths of this study was the combination of environmental and socioeconomic dimensions using the AHO weighted overlay analysis, sensitivity analysis, correlation analysis, and random forest validation analysis. A future scenario modeling approach further strengthened the final marginality assessment in the study area. In addition, this study examines both environmental and socioeconomic dimensions, unlike previous studies that primarily focus on biophysical constraints. There are several limitations to be acknowledged as well.

First, the AHP weighting approach depends partly on expert judgment and may therefore introduce subjectivity into the relative importance assigned to the criteria. Although the consistency ratio indicated that the pairwise comparisons were logically consistent, alternative judgments could produce different weighting structures. Sensitivity analysis and the Jaccard Similarity Index were used to evaluate the stability of the model under alternative weighting assumptions. The hydrology-, soil-, and social-emphasis scenarios showed high spatial agreement with the baseline model, whereas the equal-weight scenario produced substantially lower agreement. These findings indicate that the model is relatively robust under plausible changes in criterion priorities but remains sensitive to major changes in the overall weighting structure.

Second, the future urban-expansion scenario was developed using a simplified buffer-based approach. This method provides a transparent exploratory representation of potential development pressure but does not incorporate detailed drivers such as population growth, road accessibility, zoning regulations, land-market conditions, or historical land-use transi-

tion probabilities. Future studies could apply more advanced spatial simulation methods, such as cellular automata, CA–Markov modeling, or other land-use change models, to generate more realistic projections.

Third, the random forest analysis used five standardized environmental and socioeconomic predictors and excluded the land-cover marginality layer. As a result, the machine learning analysis did not directly estimate the predictive contribution of land cover, even though land cover remained an important component of the GIS-MCDA framework. This limits direct comparison between the random forest importance results and the complete six-criterion AHP model. Future analyses should evaluate alternative model specifications that include all six criteria and employ appropriate feature selection and spatial validation procedures.

6. Conclusions

This study shows the importance of integrating environmental and socioeconomic factors to identify marginal and underutilized land and inform future land-use planning and sustainable strategies. Areas that fall into the categories of low marginality/highly suitable, slightly marginal, and moderately marginal present opportunities for sustainable land use, such as biomass production, sustainable agriculture, or conservation. In contrast, areas under high marginality and very high marginality have underutilization potential that requires protection from intensified land-use pressure. Under the exploratory 2035 scenario, the combined extent of high and very high marginality increased from approximately 152,771 acres (34.14%) to 176,534 acres (39.45%), representing an estimated increase of about 23,763 acres. This result indicates the potential sensitivity of marginal land availability to continued urban expansion and supports proactive planning to reduce fragmentation and protect environmentally sensitive landscapes. Because the scenario is based on a standardized buffer, it should be interpreted as a planning exercise rather than a prediction of future development. Although the proposed GIS-MCDA framework provides a transparent and reproducible approach for identifying marginal and underutilized lands, this study is subject to certain limitations. The future urban expansion scenario represents an exploratory planning exercise rather than a prediction of future land-use change, and more advanced dynamic simulation models such as Cellular Automata (CA) or CA-Markov were not implemented because the primary objective of this study was to develop and validate a GIS-based land suitability framework rather than simulate future land-use dynamics. This study can be a valuable source of information for policymakers, land-use planners, and resource managers to balance economic and infrastructure development, agricultural preservation, future urban growth management, and environmental sustainability, thereby supporting informed decision-making. Future research should incorporate climate-related variables and dynamic land-use simulation models to further enhance regional sustainability planning and land management decision-making.

Author Contributions: Conceptualization, T.A.S. and H.H.; methodology, T.A.S., S.A.N. and H.H.; software, T.A.S. and H.H.; validation, T.A.S. and H.H.; formal analysis, T.A.S.; investigation, T.A.S. and H.H.; resources, V.I.; data curation, T.A.S. and H.H.; writing and editing T.A.S., H.H., S.A.N., G.C. and V.I.; supervision and resource, H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in this study are openly available. The county and census boundaries were obtained from the U.S. Census TIGER/Line database (<https://www.census.gov>, accessed on 29 April 2026), land cover data were obtained from the National Land Cover Database (NLCD) (<https://www.usgs.gov> accessed on 29 April 2026), soil data were obtained from the Soil Survey Geographic Database (SSURGO) (<https://www.nrcs.usda.gov/>, accessed on 29 April 2026), elevation data were obtained from the USGS 3D Elevation Program (3DEP) (<https://www.usgs.gov/>, accessed on 29 April 2026), socioeconomic data were obtained from the American Community Survey (ACS) (<https://www.census.gov/>, accessed on 29 April 2026), and contaminated and brownfield data were obtained from the Florida Department of Environmental Protection (FDEP), <https://floridadep.gov/>, accessed on 29 April 2026). The datasets used, processed, and analyzed and the supporting data for this study will be made available by the authors on request.

Acknowledgments: The authors acknowledge the support of the School of the Environment at Florida Agricultural and Mechanical University for providing logistical support and software resources. We thank all institutions whose data were used in this study. The authors also thank all colleagues for their support throughout the research period.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cervelli, E.; Recchi, P.F.; Fagnano, M.; Scotto Di Pert, E.; Pindozi, S. Marginal lands between recovery and valorization. An inclusive definition to support bio-energy supply chains. The Southern Italy contexts case study. *Agric. Syst.* **2024**, *217*, 103951. [CrossRef]
2. Tiwari, P.; Poudel, K.P.; Yang, J.; Silva, B.; Yang, Y.; McConnell, M. Marginal agricultural land identification in the Lower Mississippi Alluvial Valley based on remote sensing and machine learning model. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *125*, 103568. [CrossRef]
3. Gibbs, H.K.; Salmon, J.M. Mapping the world's degraded lands. *Appl. Geogr.* **2015**, *57*, 12–21. [CrossRef]
4. Mulyani, A.; Mulyanto, B.; Barus, B.; Panuju, D.R.; Husnain. Geospatial Analysis of Abandoned Lands Based on Agroecosystems: The Distribution and Land Suitability for Agricultural Land Development in Indonesia. *Land* **2022**, *11*, 2071. [CrossRef]
5. Luan, C.; Liu, R.; Peng, S. Land-use suitability assessment for urban development using a GIS-based soft computing approach: A case study of Ili Valley, China. *Ecol. Indic.* **2021**, *123*, 107333. [CrossRef]
6. Jumaniyazov, I.; Juliev, M.; Reimov, M.; Orazbaev, A.; Reymov, T.; Bekanov, K. Marginal lands: A review of papers from the Scopus database published in English for the period of 1979–2022. *Soil Sci. Annu.* **2023**, *74*, 169657. [CrossRef]
7. Csikós, N.; Tóth, G. Concepts of agricultural marginal lands and their utilisation: A review. *Agric. Syst.* **2023**, *204*, 103560. [CrossRef]
8. Pancaldi, F.; Trindade, L.M. Marginal Lands to Grow Novel Bio-Based Crops: A Plant Breeding Perspective. *Front. Plant Sci.* **2020**, *11*, 227. [CrossRef] [PubMed]
9. Gelfand, I.; Sahajpal, R.; Zhang, X.; Izaurralde, R.C.; Gross, K.L.; Robertson, G.P. Sustainable bioenergy production from marginal lands in the US Midwest. *Nature* **2013**, *493*, 514–517. [CrossRef] [PubMed]
10. Robinson-Baker, A.M.S.; Haseeb, M.; Kanga, L.H.B. Diversity and Seasonal Abundance of the Pine Bark and Ambrosia Beetles in the Florida Panhandle. *Insects* **2025**, *16*, 1275. [CrossRef] [PubMed]
11. Mellor, P.; Lord, R.A.; João, E.; Thomas, R.; Hursthouse, A. Identifying non-agricultural marginal lands as a route to sustainable bioenergy provision—A review and holistic definition. *Renew. Sustain. Energy Rev.* **2021**, *135*, 110220. [CrossRef]
12. Solís, D.; Alvarez, S.; Nguyen, L. Agriculture Under Pressure: The Economic, Environmental, and Development Drivers Transforming Florida Agriculture. *Agriculture* **2026**, *16*, 661. [CrossRef]
13. FAO. *The Future of Food and Agriculture—Alternative Pathways to 2050*; FAO: Rome, Italy, 2018.
14. Malczewski, J. GIS-based land-use suitability analysis: A critical overview. *Prog. Plan.* **2004**, *62*, 3–65. [CrossRef]
15. Ahmad Zaki, Z.; Yusup, M.; Che Abdullah, I.; Abdullah, Y.A.; Abdul Halim Shah, A.S. Land Suitability Analysis for Urban Gardening Using Gis-Based Multi-Criteria Decision-Making Approach. *Plan. Malays.* **2023**, *21*. [CrossRef]
16. Karimi, H.; Amiri, S.; Huang, J.; Karimi, A. Integrating GIS and multi-criteria decision analysis for landfill site selection, case study: Javanrood County in Iran. *Int. J. Environ. Sci. Technol.* **2019**, *16*, 7305–7318. [CrossRef]
17. Abdeta, D.; Begna, R.; Zerfu, H.; Alemayehu, A.; Ayana, A.N.; Tsegaye, N.T. Analyzing the extent and drivers of land use/land cover change in Southwestern Ethiopia. *Discov. Sustain.* **2024**, *5*, 505. [CrossRef]
18. Msofe, N.; Sheng, L.; Lyimo, J. Land Use Change Trends and Their Driving Forces in the Kilombero Valley Floodplain, Southeastern Tanzania. *Sustainability* **2019**, *11*, 505. [CrossRef]

19. El Baroudy, A.A. Mapping and evaluating land suitability using a GIS-based model. *CATENA* **2016**, *140*, 96–104. [\[CrossRef\]](#)
20. Hirwa, H.; Ngwijabagabo, H.; Minani, M.; Tuyishime, S.P.; Habimana, I. Geospatial Assessment of Urban Flood Susceptibility using AHP-Based Multi-Criteria Technique: Case of Musanze, Rwanda. *Rwanda J. Eng. Sci. Technol. Environ.* **2023**, *5*. [\[CrossRef\]](#)
21. Ngwijabagabo, H.; Tsinda, A.; Dusengimana, A.; Minani, M.; Tuyishime, C.S.; Hirwa, H.; Habimana, I. Groundwater Potential Mapping using Geospatial and AHP Techniques in Eastern Province of Rwanda. *Rwanda J. Eng. Sci. Technol. Environ.* **2023**, *5*. [\[CrossRef\]](#)
22. Wotlolan, D.L.; Lowry, J.H.; Wales, N.A.; Glencross, K. Land suitability evaluation for multiple crop agro-forestry planning using GIS and multi-criteria decision analysis: A case study in Fiji. *Agrofor. Syst.* **2021**, *95*, 1519–1532. [\[CrossRef\]](#)
23. Suthar, N.; Das, D.; Mallik, J. Land-use suitability assessment for urban development using Multi-Criteria Decision-Making Analysis in the Himalayan districts of Shimla, Nainital, and Darjeeling, India. *Discov. Environ.* **2024**, *2*, 90. [\[CrossRef\]](#)
24. Kau, A.-S.; Gramlich, R.; Sewilam, H. Modeling land suitability to evaluate the potential for irrigated agriculture in the Nile region in Sudan. *Sustain. Water Resour. Manag.* **2023**, *9*, 10. [\[CrossRef\]](#)
25. Roy, S.; Hazra, S.; Chanda, A.; Das, S. Land suitability analysis using AHP-based multi-criteria decision model for sustainable agriculture in red and lateritic zones of West Bengal, India. *J. Earth Syst. Sci.* **2022**, *131*, 201. [\[CrossRef\]](#)
26. Zaniboni, A.; Tassinari, P.; Torreggiani, D. GIS-based land suitability analysis for the optimal location of integrated multi-trophic aquaponic systems. *Sci. Total Environ.* **2024**, *913*, 169790. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Gizawu Garbaba, F.; Negasa Wolteji, B. Application of geospatial technology on land suitability analysis for wheat and maize farming: In a case of Guder sub-watershed North West Oromia, Ethiopia. *Heliyon* **2024**, *10*, e33557. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Tavana, M.; Soltanifar, M.; Santos-Arteaga, F.J. Analytical hierarchy process: Revolution and evolution. *Ann. Oper. Res.* **2023**, *326*, 879–907. [\[CrossRef\]](#)
29. Desta, M.D.; Tesseme, T.; Yigezu, T.T.; Nigussie, A.B. Assessment of landfill site suitability using GIS, remote sensing, and the multi-criteria decision-making (AHP) approach, Ethiopia. *Geol. Ecol. Landsc.* **2025**, *9*, 662–675. [\[CrossRef\]](#)
30. Kayitete, L.; Bakolo, C.; Tomlinson, J.; Fawcett, J.; Tuyisenge, M.F.; De Dieu Tuyizere, J. Applying Multi-Criteria Analysis in GIS to predict suitability for recreational green space interventions in Kigali City, Rwanda. *Appl. Geomat.* **2025**, *17*, 163–175. [\[CrossRef\]](#)
31. Ligmann-Zielinska, A.; Jankowski, P. A Framework for Sensitivity Analysis in Spatial Multiple Criteria Evaluation. In *Geographic Information Science*; Cova, T.J., Miller, H.J., Beard, K., Frank, A.U., Goodchild, M.F., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5266, pp. 217–233. [\[CrossRef\]](#)
32. Hammami, S.; Zouhri, L.; Souissi, D.; Souei, A.; Zghibi, A.; Marzougui, A.; Dlala, M. Application of the GIS based multi-criteria decision analysis and analytical hierarchy process (AHP) in the flood susceptibility mapping (Tunisia). *Arab. J. Geosci.* **2019**, *12*, 653. [\[CrossRef\]](#)
33. Tuğaç, M.G.; Tercan, A.; Torunlar, H.; Karakurt, E.; Usul, M. Agricultural land suitability assessment with GIS-based multi-criteria decision analysis and geostatistical approach in semi-arid regions. *Soil Stud.* **2023**, *12*, 15–29. [\[CrossRef\]](#)
34. Chen, Y.; Yu, J.; Khan, S. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environ. Model. Softw.* **2010**, *25*, 1582–1591. [\[CrossRef\]](#)
35. AbdelRahman, M.A.E.; Saleh, A.M.; Arafat, S.M. Assessment of land suitability using a soil-indicator-based approach in a geomatics environment. *Sci. Rep.* **2022**, *12*, 18113. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Chen, Y.; Yu, J.; Xevi, E.; Shahbaz, K. A GIS-based sensitivity analysis of multi-criteria weights. In Proceedings of the 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences, Cairns, Australia, 13–17 July 2009.
37. Meyer, H.; Reudenbach, C.; Wöllauer, S.; Nauss, T. Importance of spatial predictor variable selection in machine learning applications—Moving from data reproduction to spatial prediction. *Ecol. Model.* **2019**, *411*, 108815. [\[CrossRef\]](#)
38. Ploton, P.; Mortier, F.; Réjou-Méchain, M.; Barbier, N.; Picard, N.; Rossi, V.; Dormann, C.; Cornu, G.; Viennois, G.; Bayol, N.; et al. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* **2020**, *11*, 4540. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.