



Tạp chí Khoa học Đại học Quốc gia Hà Nội:
Các Khoa học Trái đất và Môi trường

Website: <https://js.vnu.edu.vn/EES>



Tích hợp GIS và cây quyết định đánh giá thích nghi đất đai cây dứa trên địa bàn huyện Mỏ Cày Nam, tỉnh Bến Tre

Nguyễn Hữu Cường*

*Trường Đại học Tài nguyên và Môi trường Thành phố Hồ Chí Minh,
236B Lê Văn Sỹ, Tân Bình, Thành phố Hồ Chí Minh*

Nhận ngày 28 tháng 11 năm 2017

Chỉnh sửa ngày 25 tháng 01 năm 2018; Chấp nhận đăng ngày 26 tháng 3 năm 2018

Tóm tắt: Mục tiêu của nghiên cứu là tích hợp GIS và cây quyết định trong đánh giá đất đai nhằm cung cấp căn cứ cho quy hoạch sử dụng đất. GIS được sử dụng để thành lập các bản đồ đơn tính, đơn vị đất đai, thích nghi đất đai trong khi cây quyết định xác định mối quan hệ giữa các yếu tố đặc điểm đất đai với năng suất trung bình của cây trồng. Nghiên cứu áp dụng thử nghiệm trên địa bàn huyện Mỏ Cày Nam, tỉnh Bến Tre cho cây dứa. Mô hình hồi quy cây quyết định được thực hiện với: Biến mục tiêu là năng suất, các biến dự báo là: Loại đất, độ nhiễm mặn, độ nhiễm phèn, độ ngập nước và khả năng tưới. Kết quả mô hình cây quyết định cho thấy mức độ giải thích các biến dự báo là 99,09%. Diện tích cấp thích nghi cao chiếm 3.522,22 ha, cấp thích nghi trung bình chiếm 12.376,21 ha, cấp thích nghi kém chiếm 6.309,37 ha.

Từ khóa: Đánh giá đất đai, GIS, cây quyết định, cây dứa, huyện Mỏ Cày Nam.

1. Giới thiệu

Đánh giá đất đai là sự nhận định những tính năng của đất đai như một tài nguyên thiên nhiên, kinh tế và sản xuất nhằm mục đích xác định khả năng sản xuất của đất đai với chất lượng và giá trị khác nhau. Hiện nay có nhiều phương pháp được áp dụng thực hiện đánh giá thích nghi đất đai về mặt tự nhiên. FAO (1993) đã đưa ra khung đánh giá đất đai với phương pháp định tính kết hợp theo điều kiện hạn chế lớn nhất [1]. Các phương pháp bán định lượng

được đưa vào sử dụng đã đáp ứng được công tác quản lý tổng hợp nguồn tài nguyên đất đai bền vững [2]. Phương pháp phân tích thứ bậc Analytic Hierachy Process (AHP) được áp dụng để tính toán chỉ số thích nghi đất đai, trong đó có tính đến mối quan hệ giữa các yếu tố đặc điểm đơn tính của đất đai. Tuy nhiên phương pháp này lại có sự tham gia “nhận định” của các chuyên gia. Phương pháp Docuchaev đã lượng hóa, xây dựng mô hình hồi quy giữa lượng các vi chất trong đất với năng suất cây trồng nhưng chưa chú ý đến những yếu tố tự nhiên khác có ảnh hưởng đến tính thích nghi của cây trồng. Tương quan và phân tích hồi quy đã được sử dụng để xác định sự đóng góp tương đối của các đặc điểm đất được chọn [3].

* ĐT.: 84-987997088.

Email: nhcuong@hcmunre.edu.vn

<https://doi.org/10.25073/2588-1094/vnuees.4218>

Hiện nay, nhiều phương pháp khai phá dữ liệu (data mining) đã được áp dụng rộng rãi trong đánh giá đất đai [4] nhằm khắc phục những yếu tố mang tính chủ quan. Cây quyết định (decision tree) là một trong những thuật toán phân loại phổ biến nhất hiện nay trong khai phá dữ liệu [5]. Đã có nhiều nghiên cứu áp dụng phương pháp này trong đánh giá đất đai. van Lanen và cộng sự (1992) trong nghiên cứu đánh giá đất đai hỗn hợp định tính và định lượng đã tiến hành với cây khoai tây tại Hà Lan. Kết quả cho thấy khoảng 65% diện tích đất có khả năng phù hợp [6]. Bouma và cộng sự (1993) đã nghiên cứu đánh giá đất đai cho cây ngô ở cấp độ nông trại tại New York [7]. Jian Tian và cộng sự (2009) so sánh mức độ chính xác trong đánh giá đất đai ứng dụng khai phá dữ liệu với 3 kỹ thuật: cây quyết định, mạng nơ-ron và hồi quy. Kết quả cho thấy mô hình cây quyết định là mô hình tốt nhất [4]. Yang JingFeng và cộng sự (2010) trong nghiên cứu của mình về cây quyết định trong đánh giá đất đai tại tỉnh Quảng Đông, Trung Quốc đã chứng minh rằng phương pháp này thuận tiện để trích xuất các quy tắc phân loại với tỷ lệ chính xác 86,67% [8]. Cây quyết định cũng được sử dụng để thực hiện từ dữ liệu khảo sát đất đai vùng Maharashtra, Ấn Độ bởi Nirmal Kumar và cộng sự (2013). Việc kiểm tra chéo 10 lần cung cấp độ chính xác 100% [5]. Tại Việt Nam, Nguyễn Ánh Nga (2012) nghiên cứu khả năng ứng dụng khai phá dữ liệu trong đánh giá đất đai với kỹ thuật cây quyết định đối với cây điều và cây xoài trên địa bàn huyện Định Quán, Đồng Nai [9].

Mục tiêu của nghiên cứu là tích hợp GIS và phương pháp cây quyết định trong khai phá dữ liệu đánh giá thích nghi đất đai tự nhiên đối với cây dừa trên địa bàn huyện Mỏ Cày Nam, tỉnh Bến Tre.

2. Phương pháp nghiên cứu và quy trình thực hiện

2.1. Phương pháp nghiên cứu

Mô hình cây quyết định

Cây quyết định được sử dụng để chia liên tiếp một tập dữ liệu lớn thành các tập con nhỏ bằng cách áp dụng một chuỗi các thuật toán [10]. Với mỗi phép chia liên tiếp, các tập con thu được trong tập kết quả sẽ ngày càng giống nhau.

Việc xây dựng cây quyết định được thực hiện dựa trên phép đo tính thuần nhất hay ngược lại là mức độ hỗn loạn (entropy) của một tập hợp. Có hai độ đo thường dùng để lựa chọn thuộc tính:

- *Độ lợi thông tin (Information Gain)*: Đo mức độ hỗn loạn của thông tin. Tại mỗi cấp, cây được phân nhánh theo thuộc tính có độ lợi thông tin lớn nhất (hay nói cách khác tối thiểu hóa mức độ hỗn loạn của thông tin).

Khi đó, với S là số lượng tập dữ liệu; s_i là số các mẫu của S nằm trong lớp C_i với $i = \{1, \dots, m\}$ thì thông tin cần biết để phân lớp một mẫu:

$$I(s_1, s_2, \dots, s_m) = - \sum_{i=1}^m \frac{s_i}{s} \log_2 \frac{s_i}{s}$$

Thuộc tính A có các giá trị $\{a_1, a_2, \dots, a_n\}$. Dùng thuộc tính A để phân chia tập dữ liệu thành n tập con $\{S_1, S_2, \dots, S_n\}$.

S_{ij} : số mẫu của lớp C_i thuộc tập con S_j ($A = a_j$).

Entropy của thuộc tính A:

$$E(A) = \sum_{j=1}^n \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{s} I(s_{1j} + s_{2j} + \dots + s_{mj})$$

Độ lợi thông tin dựa trên phân nhánh bằng thuộc tính A:

$$G(A) = I(s_1, s_2, \dots, s_m) - E(A)$$

- *Chỉ số Gini (Gini Index)*: Đo độ “không trong suốt” của thông tin, phép chia được chọn là phép chia có chỉ số Gini thấp nhất.

Chỉ số Gini của nút t:

$$GINI(t) = 1 - \sum_j p(j|t)^2$$

Với $p(j|t)$ là tần suất của lớp j trong nút t.

Khi phân chia nút p thành k nhánh, chất lượng của phép chia được tính bằng:

$$GINI_{chia} = \sum_{i=1}^k \frac{n_i}{n} GINI(i)$$

n_i : là số mẫu trong nút i

n : là số mẫu trong nút p

Ứng dụng cây quyết định trong đánh giá thích nghi đất đai bằng việc xác định những *tổ hợp các yếu tố đặc điểm đất đai* (Loại đất, độ nhiễm phèn, độ nhiễm mặn, độ ngập nước, khả năng tưới,...) và mức *sản lượng cây trồng tương ứng* [9]. Quy trình đánh giá đất đai theo các tiêu chí cây quyết định là dễ tiếp cận và minh bạch [7].

Nghiên cứu sử dụng phần mềm DTREG để xây dựng mô hình cây quyết định.

2.2. Phương pháp điều tra, khảo sát

Nghiên cứu thực hiện điều tra khảo sát nông hộ trên địa bàn huyện Mỏ Cày Nam, tỉnh Bến Tre. Thông tin điều tra gồm các đặc điểm tự nhiên đất đai (Loại đất, độ nhiễm mặn, độ nhiễm phèn, độ ngập nước, khả năng tưới) và năng suất (trái/ha/năm) cây dừa. Thông thường cây dừa được trồng với mật độ trung bình từ 250-300 cây/ha, tương ứng với khoảng cách (6-7)m x (6-7)m.

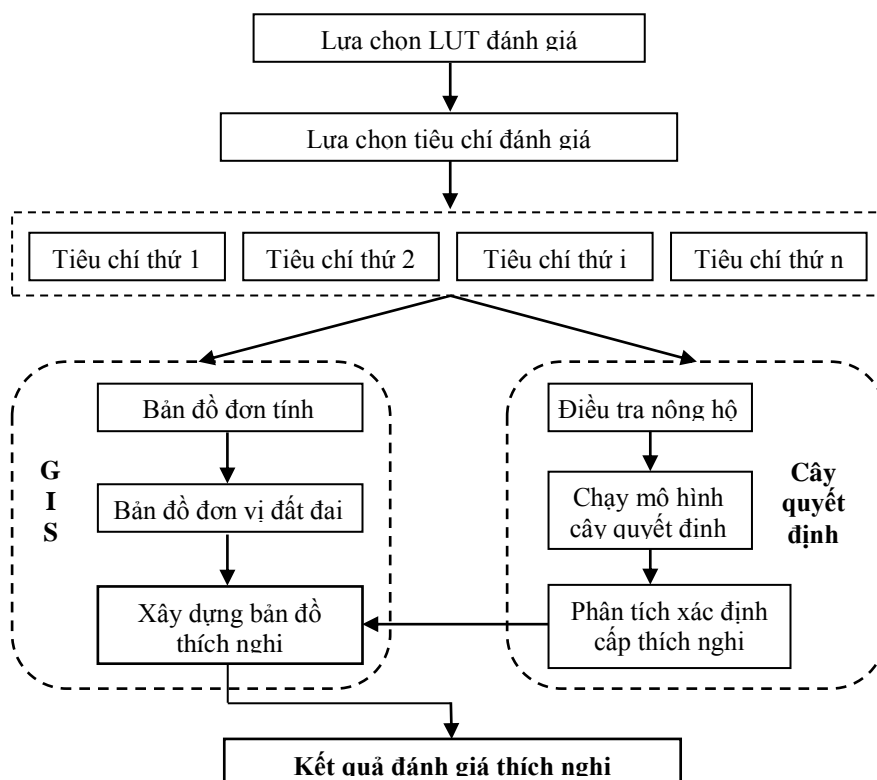
Kết quả điều tra được 86 phiếu. Tiến hành sàng lọc mẫu phiếu điều tra, số phiếu đạt yêu cầu là 80 phiếu, loại bỏ 6 phiếu không đạt yêu cầu.

2.3. Phương pháp ứng dụng GIS

Nghiên cứu sử dụng phần mềm GIS (MapInfo) xây dựng các bản đồ đơn tính, bản đồ đơn vị đất đai và bản đồ thích nghi đất đai.

Quy trình thực hiện đánh giá đất đai tích hợp GIS và mô hình cây quyết định

Quy trình nghiên cứu bắt đầu từ việc xác định loại hình sử dụng đất (LUT) để tiến hành đánh giá; xác định các tiêu chí đánh giá (cũng là các biến trong mô hình). Biến dự báo được đề xuất dựa trên đặc điểm tự nhiên đất đai của địa phương và yêu cầu sử dụng đất của cây trồng. Biến kết quả là năng suất thực tế cây trồng trên địa bàn nghiên cứu. Nhóm nghiên cứu đề xuất quy trình thực hiện theo hình 1.



Hình 1. Quy trình đánh giá đất đai tích hợp GIS và cây quyết định

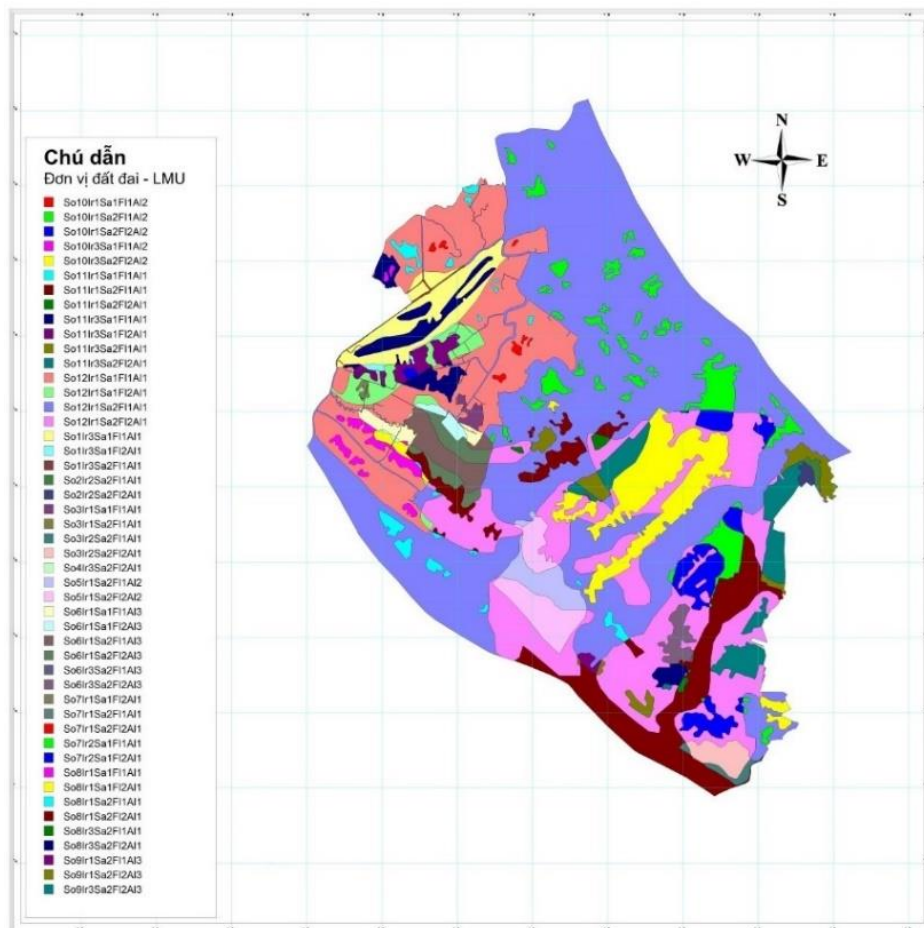
3. Kết quả và thảo luận

Xây dựng bản đồ đơn vị đất đai trên địa bàn huyện Mộ Cày Nam

Dựa vào dữ liệu thu thập được ở địa bàn nghiên cứu và đặc tính tự nhiên cây dừa, các đơn vị đất đai được xây dựng trên cơ sở mối quan hệ giữa tính chất loại hình thổ nhưỡng, độ nhiễm phèn, khả năng tưới, độ nhiễm mặn và độ ngập nước. Đây cũng là những yếu tố được lựa chọn là biến dự báo trong mô hình cây quyết định.

Loại hình thổ nhưỡng trên địa bàn huyện Mộ Cày Nam được chia thành 5 nhóm chính với tổng cộng 12 loại đất:

- Nhóm đất cát: Chủ yếu là đất cát giồng đã phân hóa phẫu diện (Cz2);
- Nhóm đất mặn: Gồm đất mặn trung bình và ít điển hình (M); Đất mặn trung bình và ít glây (Mg); Đất mặn trung bình và ít trên nền cát (M/c);
- Nhóm đất phèn: Gồm đất phèn hoạt động nông (Sj1); Đất phèn hoạt động sâu (Sj2);
- Nhóm đất phù sa: Gồm đất phù sa phân hóa yếu trung tính ít chua (P); Đất phù sa có tầng glây yếu hoặc trung tính, sâu (Pg); Đất có đóm loang lỗ chua, glây sâu (P(f)g); Đất có tầng loang lỗ chua, glây nông (Pfg); Đất có tầng loang lỗ trên nền cát (Pf/c);



Hình 2. Bản đồ đơn vị đất đai huyện Mộ Cày Nam.

- Nhóm đất nhân tác: Đất phù sa lên líp.

Độ nhiễm phèn trên địa bàn huyện được chia thành 3 cấp độ: Không phèn, độ nhiễm phèn nhỏ hơn 50 cm và độ nhiễm phèn lớn hơn 50 cm.

Khả năng tưới được chia thành 3 cấp độ: Khả năng tưới tốt, khả năng tưới trung bình và khả năng tưới kém.

Xâm nhập mặn được chia thành 2 cấp độ: Độ mặn từ 3-4‰, độ mặn lớn hơn 4‰.

Độ ngập nước được chia thành 2 cấp độ: Độ ngập nước nhỏ hơn 30 cm, độ ngập nước từ 30 – 60cm.

Ứng dụng GIS chồng xếp các chỉ tiêu đơn tính từ các bản đồ đơn tính cùng tỷ lệ: Bản đồ loại hình thổ nhưỡng, bản đồ nhiễm phèn, bản đồ khả năng tưới, bản đồ xâm nhập mặn và bản đồ ngập nước lập được bản đồ đơn vị đất đai

huyện Mỏ Cày Nam với 48 đơn vị đất đai (LMU). Xây dựng mô hình cây quyết định xác định cấp thích nghi cây dừa

Mô hình hồi quy cây quyết định đánh giá thích nghi đất đai cây dừa trên địa bàn huyện Mỏ Cày Nam được thực hiện với các biến:

- Biến mục tiêu (Target): Năng suất (nghìn trái/ha/năm).

- Các biến dự báo (Predictor): Loại đất (Loại đất), độ nhiễm mặn (Xâm nhập mặn), độ nhiễm phèn (Độ nhiễm phèn), độ ngập nước (Độ ngập nước), khả năng tưới (Khả năng tưới).

Mô hình cây quyết định sau khi chạy được xây dựng gồm 6 tầng với số nhóm phân chia là 12, cây đầy đủ có tất cả 13 nút lá (leaf nút), tổng số nút là 25. Kết quả “Phân tích phương sai” và “Tầm quan trọng các biến” được sử dụng để đánh giá mô hình.

```

===== Analysis of Variance =====
Variance in initial data sample = 33.656056
Residual (unexplained) variance after tree fitting = 0.3061744
Proportion of variance explained = 0.99090 (99.090%)
    
```

Hình 3. Kết quả phân tích phương sai của mô hình

Phương sai của tập dữ liệu trước khi xây dựng cây quyết định là: 33,66. Phương sai sau khi cây được ứng dụng vào tập dữ liệu để dự báo biến mục tiêu là: 0,30. Kết quả cho thấy một mức độ cải thiện phương sai đáng kể, cũng cho thấy tính thích hợp của mô hình cây quyết định được đưa ra. Khả năng được giải thích của biến mục tiêu bởi cây quyết định là 99,09%, còn lại 0,91% không thể giải thích được do chịu ảnh hưởng của các yếu tố khác. Cụ thể, các yếu tố loại đất, độ nhiễm mặn, độ nhiễm phèn, độ

ngập nước và khả năng tưới giải thích được 99,09% sự hình thành năng suất cây trồng. Như vậy có thể nói, mô hình cây quyết định được xây dựng có mức độ thích hợp và khả năng dự báo là khá cao.

Kết quả mô hình còn cho thấy tầm quan trọng (Mức độ ảnh hưởng) của mỗi biến dự báo (Loại đất, độ nhiễm mặn, độ nhiễm phèn, độ ngập nước và khả năng tưới) đến biến kết quả (Năng suất cây dừa) là khác nhau.

```

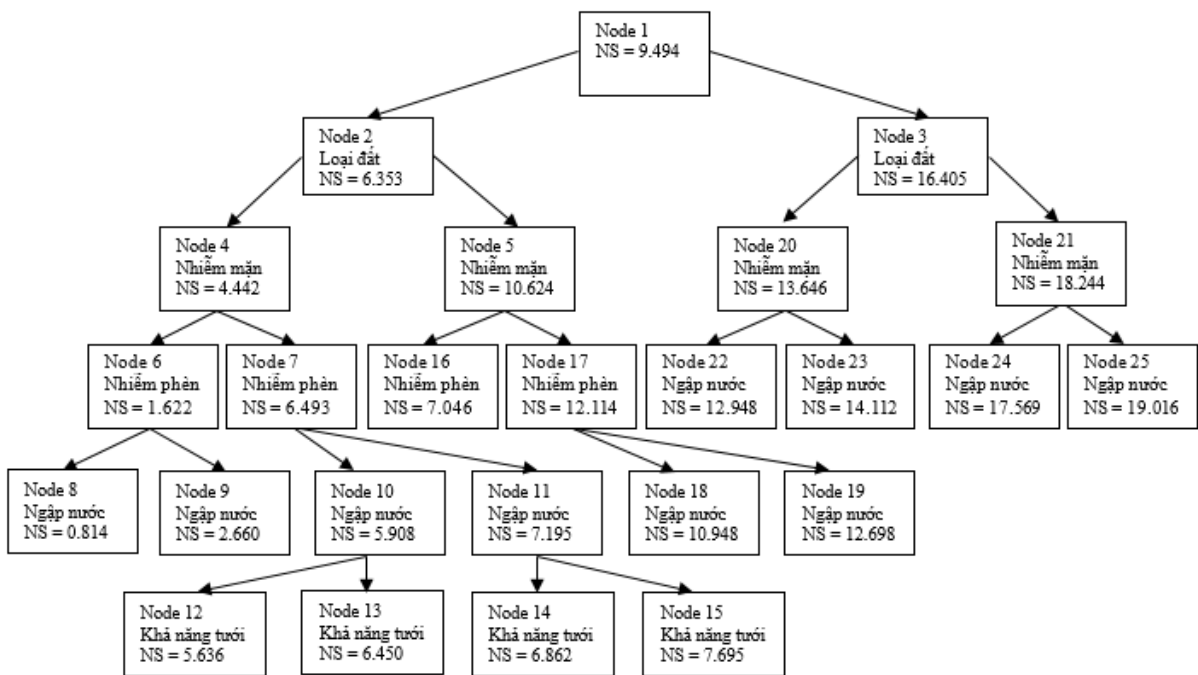
===== Overall Importance of Variables =====
Variable                                     Importance
-----
Loại đất                                     100.000
Xâm nhập mặn                                 33.147
Độ nhiễm phèn                               17.874
Độ ngập nước                                 2.400
Khả năng tưới                                0.198
    
```

Hình 4. Kết quả phân tích tầm quan trọng các biến theo mô hình

Năng suất cây dừa bị ảnh hưởng mạnh mẽ nhất bởi yếu tố loại đất. Tiếp theo, mức độ ảnh hưởng của yếu tố xâm nhập mặn đến năng suất bằng 33,15% so với loại đất ảnh hưởng đến năng suất. Tương tự, mức độ ảnh hưởng đến năng suất theo thứ tự tiếp theo là độ nhiễm phèn, độ ngập nước và khả năng tưới với các giá trị lần lượt là 17,87%, 2,40% và 0,20% so

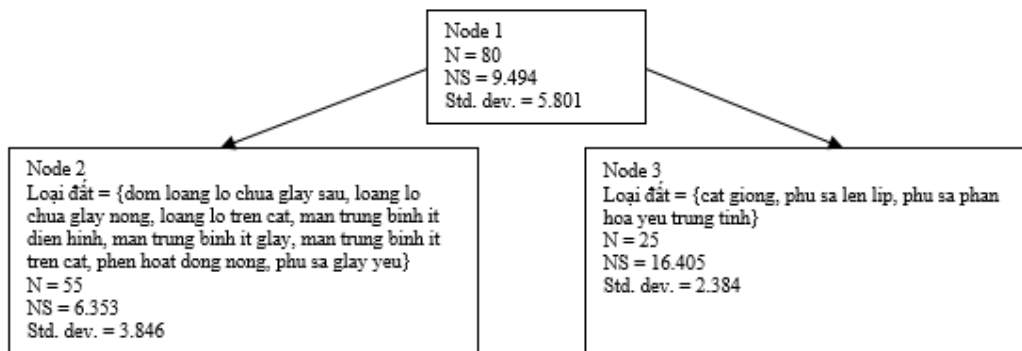
với loại đất ảnh hưởng đến năng suất. Mức độ ảnh hưởng của các yếu tố trên có thể được xem là phù hợp với thực tế địa bàn nghiên cứu cũng như đặc tính của cây dừa.

Mô hình cây quyết định đã được tạo ra với nút phân chia đầu tiên là biến Loại đất, sau đó đến các biến khác theo kết quả phân tích tầm quan trọng các biến.



Hình 5. Kết quả mô hình cây quyết định.

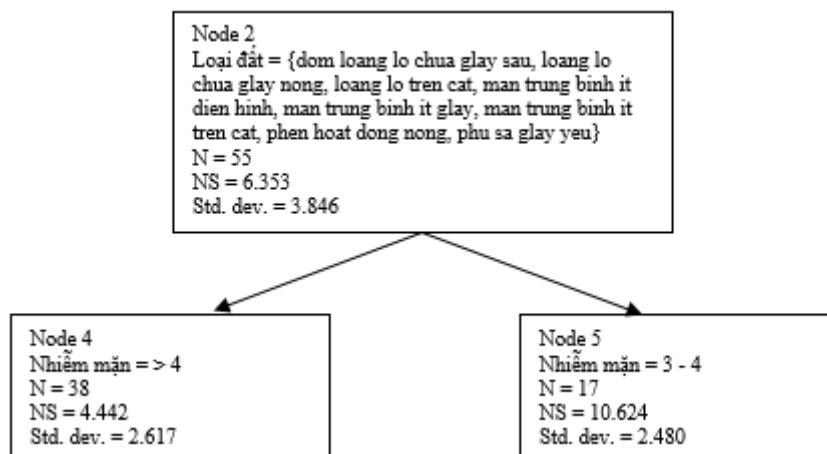
Ý nghĩa một “nhánh” của mô hình theo trình tự nút 1 – nút 2 – nút 4 – nút 7 – nút 10 – nút 12 được lý giải như sau:



Hình 6. Kết quả phân chia mô hình cây quyết định tại nút 1.

Tại nút 1 biến Loại đất phân chia thành nút 2 và nút 3. Ở nút 2 nếu dựa được trồng trên các loại đất: Đất mặn trung bình và ít điển hình; đất mặn trung bình và ít gầy; đất mặn trung bình và ít trên nền cát; đất phèn hoạt động nông; đất phèn hoạt động sâu; đất phù sa có tầng gầy yếu hoặc trung tính, sâu; đất có đóm loang lỗ chua,

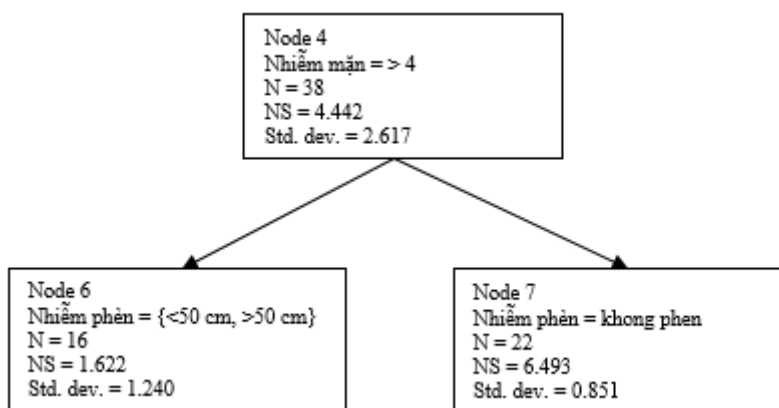
gây sâu; đất có tầng loang lỗ chua, gây nông; đất có tầng loang lỗ trên nền cát thì cho năng suất trung bình là 6.353 trái/ha. Ở nút 3 nếu cây dựa được trồng trên các loại đất: Đất cát giồng đã phân hóa phũ điện; đất phù sa phân hóa yếu trung tính ít chua và đất phù sa lên líp thì cho năng suất trung bình là 16.405 trái/ha.



Hình 7. Kết quả phân chia mô hình cây quyết định tại nút 2.

Tiếp theo tại nút 2 chia thành nút 4 và nút 5 theo yếu tố xâm nhập mặn. Với cùng các loại đất ở nút 2 nhưng nếu có độ xâm nhập mặn lớn hơn 4‰ thì cây dựa cho năng suất trung bình là

4.442 trái/ha (nút 4), nếu độ xâm nhập mặn từ 3-4‰ thì năng suất trung bình của cây dựa là 10.624 trái/ha (nút 5).



Hình 8. Kết quả phân chia mô hình cây quyết định tại nút 4.

Tương tự, nút 4 phân tách thành nút 6 và nút 7 dưới tác động của yếu tố nhiễm phèn. Với các loại đất như nút 2, độ nhiễm mặn như nút 4 và độ nhiễm phèn bé hơn hoặc lớn hơn 50 cm thì cây dứa cho năng suất trung bình 1.622 trái/ha (nút 6), nếu không phèn cây dứa cho năng suất trung bình là 6.493 trái/ha (nút 7).

Với cách lý giải tương tự, nút 7 phân tách thành nút 10 và 11 theo các giá trị khác nhau của Độ ngập nước. Nút 10 chia thành nút 12 và 13 theo các giá trị của Khả năng tưới. Như vậy, nếu trồng cây dứa trên vùng đất có các đặc điểm về loại đất, xâm nhập mặn, độ nhiễm phèn, độ ngập nước, khả năng tưới tương ứng như các nút 1 – nút 2 – nút 4 – nút 7 – nút 10 – nút 12 thì có năng suất trung bình là 5.636 trái/ha.

Dựa trên kết quả mô hình cây quyết định, tiến hành đi theo từng phân nhánh mô hình để xác định được *tổ hợp các yếu tố đặc điểm đất đai và mức năng suất trung bình của cây dứa tương ứng với tổ hợp đất*. Cấp thích nghi được phân chia theo gợi ý của FAO dựa trên tỷ lệ năng suất thực tế với năng suất tối hảo cây

trồng (với năng suất tối hảo thu thập được trong nghiên cứu 20.000 trái/ha/năm).

Bảng 1. Phân cấp thích nghi của FAO theo năng suất cây trồng

STT	Cấp thích nghi	Hướng dẫn phân cấp của FAO	Phân cấp trong nghiên cứu (nghìn trái/ha/năm)
1	S1 (Thích nghi cao)	> 80%	> 16,0
2	S2 (Thích nghi trung bình)	40% – 80%	8,0 – 16,0
3	S3 (Thích nghi kém)	20 – 40%	4,0 – 8,0
4	N (Không thích nghi)	< 20%	< 4,0

Từ đây xác định mức độ thích nghi cho từng tổ hợp dựa vào năng suất trung bình tương ứng của tổ hợp. Kết quả được thể hiện tại bảng 2.

Bảng 2. Kết quả tổ hợp các yếu tố ảnh hưởng đến năng suất cây dứa theo mô hình cây quyết định

STT	Loại đất	Tổ hợp các yếu tố			Năng suất trung bình (nghìn trái/ha)	Cấp thích nghi	
		Xâm nhập mặn (%)	Độ nhiễm phèn (cm)	Độ ngập nước (cm)			Khả năng tưới
1	Đất phèn hoạt động sâu; Đất có đóm loang lỗ chua, glây sâu; Đất có tầng loang lỗ chua, glây nông	> 4	< 50; > 50	30 - 60	0,814 (±0,64)	N	
2	Đất phèn hoạt động nông; Đất phèn hoạt động sâu; Đất có đóm loang lỗ chua, glây sâu; Đất có tầng loang lỗ chua, glây nông	> 4	< 50; > 50	< 30	2,66 (±1,035)	N	
3	Đất mặn trung bình và ít điển hình; Đất mặn trung bình và ít trên nền cát; Đất phù sa có tầng glây yếu hoặc trung tính, sâu	> 4	Không phèn	30 - 60	Kém; Trung bình	5,636 (±0,35)	S3
4	Đất có tầng loang lỗ trên nền cát	> 4	Không phèn	30 - 60	Tốt	6,45 (±0,5)	S3
5	Đất mặn trung bình và ít điển hình; Đất mặn trung bình và ít glây; Đất phù sa có tầng glây yếu hoặc trung	> 4	Không phèn	< 30	Kém; Trung bình	6,862 (±0,345)	S3

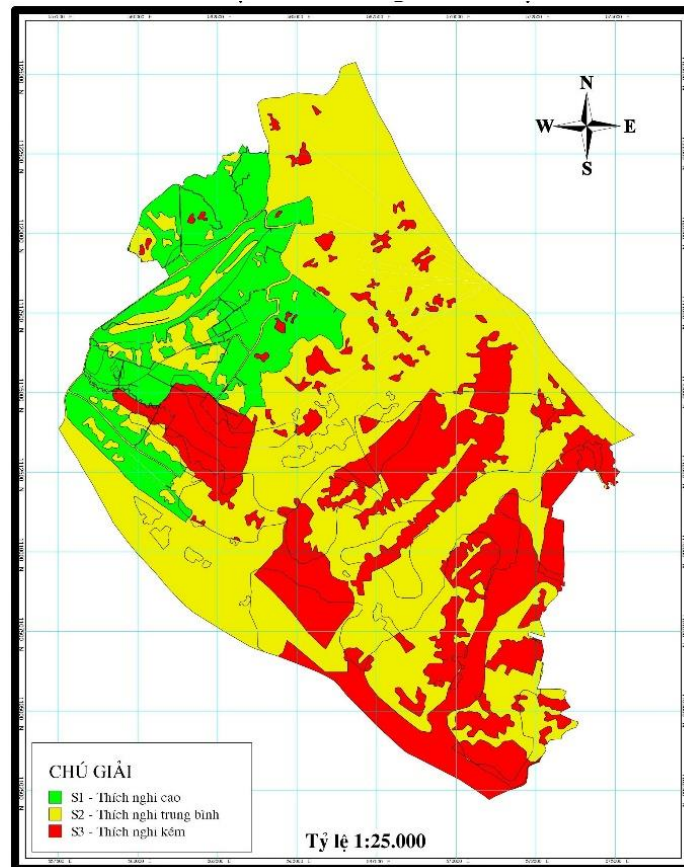
STT	Loại đất	Tổ hợp các yếu tố			Khả năng tưới	Năng suất trung bình (nghìn trái/ha)	Cấp thích nghi
		Xâm nhập mặn (‰)	Độ nhiễm phèn (cm)	Độ ngập nước (cm)			
	tính, sâu; Đất có tầng loang lỗ trên nền cát						
6	Đất mặn trung bình và ít glây; Đất phù sa có tầng glây yếu hoặc trung tính, sâu; Đất có tầng loang lỗ trên nền cát	> 4	Không phèn	< 30	Tốt	7,695 (±0,436)	S3
7	Đất phèn hoạt động sâu; Đất có tầng loang lỗ chua, glây nông; Đất phù sa có tầng glây yếu hoặc trung tính, sâu; Đất có tầng loang lỗ trên nền cát	3 - 4	< 50; > 50			7,046 (±0,917)	S3
8	Đất phù sa có tầng glây yếu hoặc trung tính, sâu; Đất có tầng loang lỗ trên nền cát	3 - 4	Không phèn	30 - 60		10,948 (±0,008)	S2
9	Đất mặn trung bình và ít glây; Đất phù sa có tầng glây yếu hoặc trung tính, sâu; Đất có tầng loang lỗ trên nền cát	3 - 4	Không phèn	< 30		12,698 (±0,434)	S2
10	Đất phù sa lên líp; Đất phù sa phân hóa yếu trung tính ít chua	> 4		30 - 60		12,948 (±0,018)	S2
11	Đất cát giồng đã phân hóa phần điện; Đất phù sa lên líp; Đất phù sa phân hóa yếu trung tính ít chua	> 4		< 30		14,112 (±0,482)	S2
12	Đất cát giồng đã phân hóa phần điện; Đất phù sa lên líp; Đất phù sa phân hóa yếu trung tính ít chua	3 - 4		30 - 60		17,569 (±0,435)	S1
13	Đất cát giồng đã phân hóa phần điện; Đất phù sa lên líp; Đất phù sa phân hóa yếu trung tính ít chua	3 - 4		< 30		19,016 (±0,421)	S1

Dựa vào kết quả tổ hợp các yếu tố ảnh hưởng đến năng suất cây dừa theo mô hình cây quyết định đối chiếu với các đơn vị đất đai

huyện Mô Cày Nam xác định được cấp thích nghi cho từng đơn vị đất đai (Bảng 3).

Bảng 3. Thống kê các đơn vị đất đai theo cấp thích nghi

STT	LMU	Cấp thích nghi	Diện tích (ha)	Tỷ lệ (%)
1	17, 18, 21, 40, 41, 44, 45	S1	3.522,22	15,86
2	3, 4, 6, 7, 10, 11, 12, , 19, 20, 24, 25, 39, 42, 43	S2	12.376,21	55,73
3	1, 2, 5, 8, 9, 13, 14, 15, 16, 22, 23, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 46, 47, 48	S3	6.309,37	28,41
Tổng			22.207,80	100,00



Hình 9. Bản đồ thích nghi cây dừa theo mô hình cây quyết định.

Dựa vào kết quả phân cấp thích nghi cây dừa trên địa bàn huyện Mỏ Cày Nam cho từng đơn vị đất đai cùng với bản đồ đơn vị đất đai ta xây dựng được bản đồ mức độ thích nghi đối với cây dừa.

Kết quả cho thấy diện tích đất ở huyện Mỏ Cày Nam phần lớn thích nghi trung bình đối với cây dừa chiếm 55,73% diện tích, thích nghi cao chiếm 15,86% diện tích, còn lại là thích nghi kém với 28,41% diện tích.

4. Kết luận

Nghiên cứu đã tích hợp GIS và cây quyết định để thực hiện đánh giá đất đai. GIS được sử dụng để thành lập các bản đồ đơn tính, đơn vị đất đai và thích nghi đất đai. Mô hình cây quyết định được xây dựng để xác định cấp thích nghi

cho từng đơn vị đất đai dựa vào việc phân chia từng tổ hợp tính chất và năng suất tương ứng. Phương pháp đánh giá theo cây quyết định bổ sung căn cứ định lượng cho phân cấp thích nghi, nhưng không hoàn toàn thay thế phương pháp luận đánh giá đất đai dựa trên phân cấp thứ bậc hay theo FAO. Kết quả nghiên cứu cho thấy diện tích đất ở huyện Mỏ Cày Nam phần lớn thích nghi trung bình đối với cây dừa và chiếm 55,73% diện tích.

Tài liệu tham khảo

- [1] FAO, 1993. FESLM: An international framework for evaluating sustainable land management. World soil resources report 73. FAO, Rome, Italy.
- [2] Lê Quang Trí và Phạm Thanh Vũ, 2010. Xác định một số tiêu chí cho đánh giá đất đai bản định

- lượng trên 02 vùng sinh thái khác nhau. Tạp chí Khoa học Trường Đại học Cần Thơ. 15b: 114-124.
- [3] de la Rosa, D., van Diepen, C.A., 2002. Qualitative and Quantitative Land Evaluation. In Willy H. Verheye. Land Use, Land Cover and Soil Sciences - Volume II: Land Evaluation. EOLSS, pp. 59-77.
- [4] Jian Tian, Yueming Hu, Jianmin Liu, Yanling Zhao, Changwei Wang, 2009. The comparative analysis of various classification models on land evaluation. Proc. SPIE 7492, International Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining, 74921A, 15 October 2009, Wuhan, China.
- [5] Kumar, N., Obi Reddy, G. P., Chatterji, S., 2013. Evaluation of Best First Decision Tree on Categorical Soil Survey Data for Land Capability Classification. International Journal of Computer Applications. 72(4): 5-8.
- [6] van Lanen, H.A.J., Hack-ten Broeke, M.J.D., Bouma, J., de Groot, W.J.M., 1992. A mixed qualitative/quantitative physical land evaluation methodology. Geoderma. 55(1-2): 37-54.
- [7] Bouma, J., Wagenet, R. J., Hoosbeek, M. R., Hutson, J. L., 1993. Using expert systems and simulation modelling for land evaluation at farm level: a case study from New York State. Soil Use Management. 9(4): 131-139.
- [8] Yang JingFeng, Li Ting, Chen ZhiMin, 2010. Land evaluation method based on decision tree produced by C4.5 and fuzzy decision. Agricultural Science & Technology – Hunan. 11(3): 1-3.
- [9] Nguyễn Ánh Nga, 2012. Ứng dụng kỹ thuật khai phá dữ liệu cho việc định lượng trong đánh giá đất đai trên địa bàn huyện Định Quán, tỉnh Đồng Nai. Luận văn thạc sĩ khoa học nông nghiệp. Trường Đại học Nông Lâm Thành phố Hồ Chí Minh.
- [10] Han, J., Kamper, M., 2006. Data Mining: Concepts and Techniques, Second Edition. Morgan Kaufmann Publishers, Elsevier Inc, 772 pages.

Integration of GIS and Decision Tree in Land Evaluation for Coconut Trees in Mo Cay Nam District, Ben Tre Province

Nguyen Huu Cuong

*Ho Chi Minh University of Resources and Environment,
236B Le Van Sy, Tan Binh, Ho Chi Minh City, Vietnam*

Abstract: In this research, the integrated model of GIS and decision tree (DT) was built for *land suitability analysis* that support the foundation of land use planning. GIS was used to create thematic maps and decision tree shows several factor combinations according to plant average productivity. This study is applied for coconut trees in Mo Cay Nam district, Ben Tre province. The target variable is the productivity and the predictor variables consist of soil types, salinity, acidity, flood and irrigation. The study shows that the interpretation level of the predictive variables is 99.09%. The area of highly suitable is 3,522.22 hectares, suitable is 12,376.21 hectares, moderately suitable is 6,309.37 hectares.

Keywords: Land evaluation, GIS, decision tree, coconut tree, Mo Cay Nam district.