

ỨNG DỤNG MÔ HÌNH CÂY QUYẾT ĐỊNH ĐÁNH GIÁ THÍCH HỢP ĐẤT TRỒNG CÂY CAM TRÊN ĐỊA BÀN HUYỆN BẮC TÂN UYÊN, TỈNH BÌNH DƯƠNG

Võ Quốc Khánh*, Nguyễn Văn Cường, Trương Thị Diệu Quân

Trường Đại học Tài nguyên và Môi trường TP. Hồ Chí Minh

<https://doi.org/10.55250/jo.vnuf.2022.4.088-095>

TÓM TẮT

Phương pháp khai phá dữ liệu với mô hình cây quyết định là phương pháp đánh giá thích hợp đất đai mang tính định lượng với biến mục tiêu là dữ liệu năng suất cây trồng được thu thập thực tế tại nông hộ trên địa bàn huyện Bắc Tân Uyên, tỉnh Bình Dương. Phương pháp này khắc phục được hạn chế của phương pháp đánh giá theo yếu tố hạn chế lớn nhất là có phân tích tính tương hỗ giữa các yếu tố thông qua chỉ tiêu năng suất thu được của cây trồng. Nghiên cứu đã thực hiện trên 127 phiếu điều tra có độ tin cậy cao được sử dụng để chạy phần mềm Dtree và phân tích dữ liệu xây dựng mô hình cây quyết định. Mô hình chia tách thành 18 đơn vị đất đai, trong đó có 6 đơn vị đất đai được đánh giá mức thích hợp S1, chiếm 33% và 12 đơn vị đất đai được đánh giá mức thích hợp S2, chiếm 67%. Theo diện tích, đánh giá thích hợp đất đai theo mô hình cây quyết định đã cho thích hợp cao (S1) có diện tích 25.796 ha, chiếm 66,2% và thích hợp trung bình (S2) có diện tích 9.652 ha, chiếm 24,8% so với diện tích đất tự nhiên. Kết quả này được dùng để so sánh với phương pháp đánh giá của FAO, nghiên cứu sử dụng mô hình cây quyết định có sự khác biệt với nhóm đất có mức thích nghi S1 và S2 lần lượt là 66,2% và 24,8%.

Từ khóa: Cây cam, cây quyết định, đánh giá đất đai, khai phá dữ liệu.

1. ĐẶT VẤN ĐỀ

Trong sản xuất nông nghiệp trên địa bàn huyện Bắc Tân Uyên, việc phát triển trồng cây cam tại địa phương còn mang tính tự phát, chưa có những nghiên cứu, đánh giá sự thích hợp của từng loại đất với những điều kiện tự nhiên khác nhau một cách khoa học, có hệ thống. Điều này rất dễ dẫn đến những hậu quả tiêu cực về sau nếu sự phát triển trồng cam ồ ạt ngay cả trên những vùng có điều kiện đất đai không phù hợp, đặc biệt là cây lâu năm như cây cam.

Phương pháp khai phá dữ liệu (data mining) đã được áp dụng rộng rãi trong đánh giá đất đai, nhưng mô hình cây quyết định là một trong những phương pháp sử dụng thuật toán để phân lớp phổ biến nhất hiện nay trong khai phá dữ liệu (Kumar et al., 2013). Đã có nhiều nghiên cứu áp dụng phương pháp này trong đánh giá đất đai. Trên thế giới, 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. Tại Việt Nam, Nguyễn Hữu Cường đã sử dụng phương pháp này để đánh giá thích hợp đất đai cho cây cao su tại huyện Phú Giáo tỉnh Bình Dương cho kết quả 99,12% diện tích đất có khả năng phù hợp (Nguyễn Hữu Cường, 2018).

Mục tiêu của nghiên cứu này là áp dụng phương pháp khai phá dữ liệu trong đánh giá thích hợp đất đai bằng mô hình cây quyết định và so sánh kết quả với phương pháp của FAO

*Corresponding author: khanh692002@yahoo.com

trên cây cam tại địa bàn huyện Bắc Tân Uyên, tỉnh Bình Dương.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Địa bàn nghiên cứu

Nghiên cứu được thực hiện trên địa bàn huyện Bắc Tân Uyên, là huyện mới tách ra từ 10 xã của huyện Tân Uyên (cũ) năm 2013. Bắc Tân Uyên có diện tích hơn 40 ngàn hecta đất tự nhiên, có vùng chuyên canh cây có múi dọc ven bờ hai con sông Đồng Nai và sông Bé. Phía Bắc của huyện có cao trình 40 – 50 m, một số đồi cao độc lập có cao trình 70 - 80 m, phía Nam của huyện có cao trình thấp khoảng 20 – 30 m, đất đai bằng phẳng, ít bị chia cắt, tạo thành vùng rộng lớn rất thuận lợi cho việc trồng cây công nghiệp lâu năm. Tổng dân số huyện năm 2021 là 87.532 người (UBND huyện BTU, 2021).

2.2. Phương pháp nghiên cứu

- Phương pháp khai phá dữ liệu và mô hình cây quyết định để xử lý và phân tích số liệu thu thập được từ việc điều tra khảo sát.

- Khai phá dữ liệu là quá trình trích xuất các thông tin có giá trị tiềm ẩn bên trong lượng lớn dữ liệu được lưu trữ trong các kho dữ liệu. Để đạt được những tri thức từ cơ sở dữ liệu hiện có, nhiều kỹ thuật khai phá dữ liệu khác nhau ra đời như: phân lớp dữ liệu, phân cụm dữ liệu, khai phá luật kết hợp, hồi quy, giải thuật di truyền, mạng nơ-ron, cây quyết định. Trong đó, kỹ thuật cây quyết định (decision tree) là một công cụ mạnh và hiệu quả trong việc phân lớp và dự báo.

- Cây quyết định là cấu trúc biểu diễn dưới

dạng cây. Trong đó, mỗi nút trong (internal node) biểu diễn một thuộc tính, nhánh (branch) biểu diễn giá trị có thể có của thuộc tính, mỗi lá (leaf node) biểu diễn các lớp quyết định và đỉnh trên cùng của cây gọi là gốc (root). Cây quyết định có thể được dùng để phân lớp bằng cách xuất phát từ gốc của cây và di chuyển theo các nhánh cho đến khi gặp nút lá (Nguyễn Ánh Nga, 2012).

- Ứng dụng cây quyết định trong đánh giá thích hợp đất đai bằng việc xác định những tổ hợp các yếu tố đặc điểm đất đai (độ dày tầng đất, độ dốc, thành phần cơ giới, khả năng tưới...) và mức sản lượng cây trồng tương ứng. Quy trình đánh giá đất đai theo các tiêu chí cây quyết định là dễ tiếp cận và minh bạch (Bouma et al., 1993).

- Sử dụng phần mềm DTREG để xây dựng mô hình cây quyết định.

- Biến dự báo là tập hợp dữ liệu đầu vào của các giá trị như loại đất, độ dày tầng đất, độ dốc địa hình, thành phần cơ giới, khả năng tưới.

- Biến mục tiêu là năng suất thu thập được từ thực tế nông hộ tại địa phương nghiên cứu.

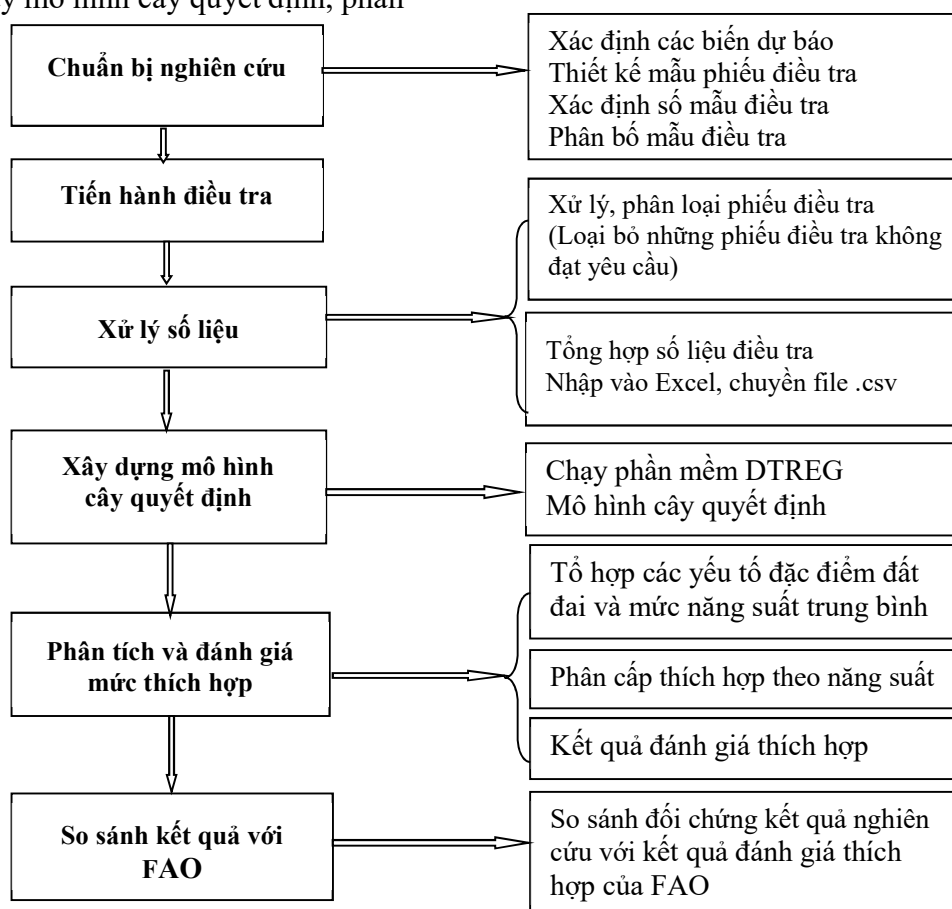
- Sau khi chạy mô hình cây quyết định, phần

mềm sẽ phân theo các nhóm có tương đồng về năng suất ở các điểm node. Chúng tôi dựa vào đề xuất các mức thích hợp theo tỷ lệ năng suất tối hảo của FAO để đánh giá các mức thích hợp S1 (thích hợp cao), S2 (thích hợp trung bình), S3 (thích hợp kém) và N (không thích hợp) cho từng nhóm.

- Phương pháp ứng dụng GIS: Sử dụng phần mềm MapInfo phân tích dữ liệu thuộc tính và không gian để xây dựng bản đồ phân bố mẫu điều tra và bản đồ các mức thích hợp đất đai cho cây cam tại địa bàn nghiên cứu.

2.3. Quy trình thực hiện đánh giá đất đai theo mô hình cây quyết định

Bắt đầu bằng việc xác định các biến cho mô hình, gồm biến dự báo và biến mục tiêu. 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 mục tiêu là năng suất thực tế của cây trồng được trên địa bàn nghiên cứu. Mô hình nghiên cứu được xây dựng dựa trên các thông tin được thu thập thông qua điều tra nông hộ. Chúng tôi đề xuất quy trình thực hiện theo hình 1.



Hình 1. Khung quy trình thực hiện nghiên cứu

3. KẾT QUẢ VÀ THẢO LUẬN

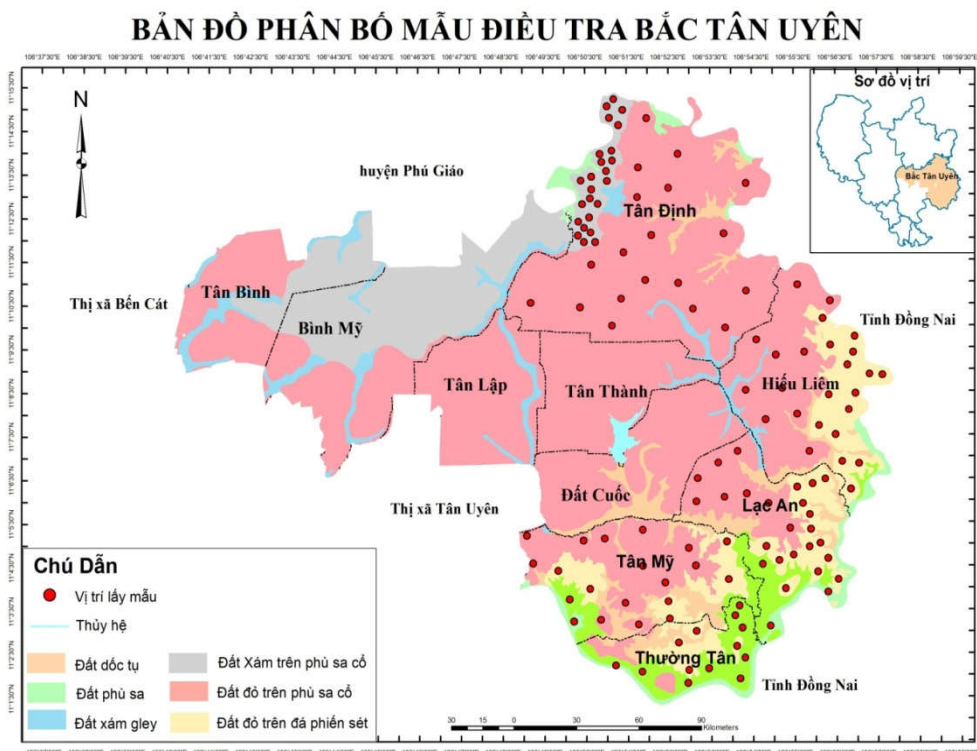
3.1. Số lượng mẫu điều tra

Số liệu thu thập được từ 140 điểm trong khu vực trồng cam tại huyện Tân Uyên, quá trình

sàng lọc dữ liệu đã có 13 phiếu điều tra không đạt yêu cầu khảo sát bị loại bỏ, còn lại là 127 số liệu đưa vào nghiên cứu.

Bảng 1. Phân bố mẫu điều tra theo đơn vị đất đai

Đơn vị đất đai	Tổng hợp các yếu tố					Số mẫu điều tra
	Nhóm đất	Độ dày tầng đất (cm)	Thành phần cơ giới	Độ dốc (°)	Khả năng tưới	
1	Đỏ vàng Fs	50 - 70	Trung bình	D > 15	Có tưới	4
2	Đỏ vàng Fs	50 - 70	Trung bình	D8 - 15	Có tưới	5
3	Đỏ vàng Fp	50 - 70	Nhẹ	D8 - 15	Có tưới	7
4	Đỏ vàng Fs	70 - 100	Trung bình	D8 - 15	Có tưới	8
5	Đỏ vàng Fp	50 - 70	Nhẹ	D3 - 8	Có tưới	6
6	Đỏ vàng Fs	70 - 100	Trung bình	D3 - 8	Có tưới	8
7	Đỏ vàng Fp	70 - 100	Nhẹ	D8 - 15	Có tưới	7
8	Xám	> 100	Nhẹ	D3 - 8	Có tưới	10
9	Đỏ vàng Fs	70 - 100	Trung bình	D0 - 3	Có tưới	5
10	Đỏ vàng Fp	> 100	Nhẹ	D8 - 15	Có tưới	6
11	Xám	70 - 100	Nhẹ	D0 - 3	Có tưới	6
12	Đỏ vàng Fp	70 - 100	Nhẹ	D3 - 8	Có tưới	6
13	Xám	> 100	Nhẹ	D0 - 3	Có tưới	8
14	Phù sa	> 100	Nặng	D0 - 3	Có tưới	6
15	Đỏ vàng Fp	70 - 100	Nhẹ	D0 - 3	Có tưới	7
16	Đỏ vàng Fp	> 100	Nhẹ	D3 - 8	Có tưới	10
17	Đỏ vàng Fp	> 100	Nhẹ	D0 - 3	Có tưới	8
18	Phù sa	> 100	Trung bình	D0 - 3	Có tưới	10
Tổng số mẫu						127



Hình 2. Bản đồ phân bố mẫu điều tra

3.2. Xây dựng các biến cho mô hình cây quyết định

Mô hình hồi quy cây quyết định đánh giá tính thích hợp cho cây cam trên địa bàn huyện Bắc Tân Uyên được xây dựng bởi các biến:

Biến mục tiêu (Target): Năng suất (tấn/ha/năm) - được lấy từ kết quả điều tra nông hộ.

Các biến dự báo (Predictor): Nhóm đất, độ dày tầng đất, thành phần cơ giới, độ dốc, khả năng tưới với các mức như sau: nhóm đất: phù sa, đỏ vàng trên phù sa cổ, đỏ vàng trên đá phiến sét xám trên phù sa cổ. Độ dày tầng đất: 50 - 70; 70 - 100, 100 (cm). Thành phần cơ giới: nặng, trung bình, nhẹ. Độ dốc: 0 - 3°, 3 - 8°, 8 - 15°, > 15°. Khả năng tưới: có tưới.

Các dữ liệu thu thập của các biến này được

đưa vào để chạy phần mềm DREG và mô hình cây quyết định được xây dựng có 7 tầng, chia thành 18 nhóm và có tổng số node là 35.

3.3. Kết quả đánh giá thích hợp đất đai cho cây cam theo mô hình cây quyết định

Phân tích phương sai

Phương sai của dữ liệu đưa vào trước khi phân cảnh theo mô hình cây quyết định là 45,1 và sau khi xây dựng mô hình cây quyết định chỉ còn 1,89. Tỷ lệ phương sai được giải thích là 95,8% đã thể hiện mức độ tin cậy rất cao khi xây dựng mô hình cây quyết định. Điều này cũng cho thấy sự phù hợp khi sử dụng mô hình cây quyết định để đánh giá định lượng khả năng thích hợp đất đai cho cây cam dựa trên năng suất điều tra thực tế.

Bảng 2. Kết quả phân tích phương sai mô hình cây quyết định

STT	Thông số	Kết quả
1	Phương sai mẫu dữ liệu đầu vào	45,1
2	Phương sai sau khi tạo cây	1,89
3	Tỷ lệ phương sai được giải thích	95,8%

Mức độ quan trọng của các yếu tố tự nhiên đến năng suất cây cam

Khi khai phá chuỗi dữ liệu, phần mềm DTREG phân tích dữ liệu đầu vào đã đánh giá

mức độ ảnh hưởng của các biến dự báo đến biến mục tiêu (năng suất) để cho kết quả về mức độ quan trọng của từng yếu tố.

Bảng 3. Bảng mức độ quan trọng của các biến dự báo

STT	Biến mô hình	Mức độ quan trọng
1	Độ dốc	100,0
2	Nhóm đất	43,9
3	Độ dày	35,5
4	Thành phần cơ giới	1,2

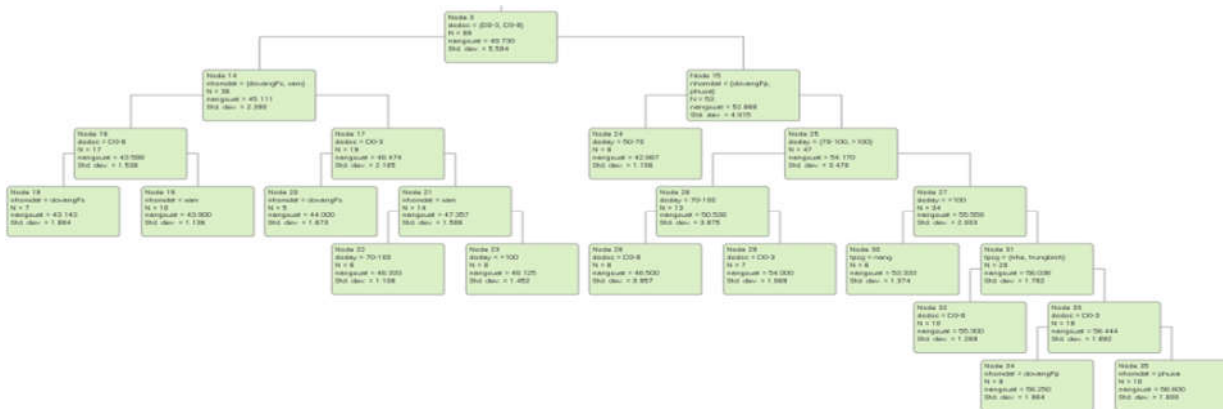
Theo đó, yếu tố độ dốc có ảnh hưởng mạnh nhất đến năng suất cây cam nên có mức độ quan trọng cao nhất, tiếp theo đó là yếu tố nhóm đất, độ dày và thành phần cơ giới.

Kết quả phân tích mô hình cây quyết định

Mô hình đã đưa ra 8 tầng và 35 nút (node) đánh giá thích hợp của cây cam với các yếu tố nghiên cứu. Trong 127 mẫu phiếu điều tra được

đưa vào phần mềm tính toán cho năng suất trung bình là 46,9 tấn/ha.

Nút đầu tiên phần mềm đã chọn biến độ dốc để chia tách dữ liệu. Như vậy, độ dốc là yếu tố quan trọng nhất ảnh hưởng đến sinh trưởng phát triển và năng suất cây cam theo các dữ liệu được đưa vào nghiên cứu.



Hình 3. Một nhánh của mô hình cây quyết định

Dựa trên kết quả mô hình cây quyết định, đánh giá trên 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 cam 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 so với năng suất tối hảo cây trồng (với năng suất tối hảo thu thập được trong nghiên cứu 60 tấn/ha/năm).

Bảng 4. Phân cấp thích nghi của FAO theo năng suất cây trồng

STT	Cấp thích hợp	Tỷ lệ phân cấp	Phân cấp trong nghiên cứu (tấn/ha/năm)
1	Thích hợp cao (S1)	> 80%	> 48
2	Thích hợp trung bình (S2)	40% - 80%	24 - 48
3	Thích hợp kém (S3)	20% - 40%	12 - 24
4	Không thích hợp	< 20%	< 12

Từ bảng trên, 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 5.

Bảng 5. Kết quả đánh giá thích hợp đất đai đối với cây cam theo mô hình cây quyết định

Đơn vị đất đai	Điểm node	Tổ hợp các yếu tố					Khả năng tưới	Năng suất TB	Mức độ thích hợp
		Nhóm đất	Độ dày tầng đất (cm)	Thành phần cơ giới	Độ dốc (°)				
1	4	Đỏ vàng Fs	50 - 70	Trung bình	D > 15	Có tưới	29,8	S2	
2	8	Đỏ vàng Fs	50 - 70	Trung bình	D8 - 15	Có tưới	39,2	S2	
3	9	Đỏ vàng Fp	50 - 70	Nhẹ	D8 - 15	Có tưới	39,3	S2	
4	12	Đỏ vàng Fs	70 - 100	Trung bình	D8 - 15	Có tưới	42,5	S2	
5	24	Đỏ vàng Fp	50 - 70	Nhẹ	D3 - 8	Có tưới	42,7	S2	
6	18	Đỏ vàng Fs	70 - 100	Trung bình	D3 - 8	Có tưới	43,1	S2	
7	13	Đỏ vàng Fp	70 - 100	Nhẹ	D8 - 15	Có tưới	43,1	S2	
8	19	Xám	> 100	Nhẹ	D3 - 8	Có tưới	43,9	S2	
9	20	Đỏ vàng Fs	70 - 100	Trung bình	D0 - 3	Có tưới	44,0	S2	
10	11	Đỏ vàng Fp	> 100	Nhẹ	D8 - 15	Có tưới	44,2	S2	
11	22	Xám	70 - 100	Nhẹ	D0 - 3	Có tưới	46,3	S2	
12	28	Đỏ vàng Fp	70 - 100	Nhẹ	D3 - 8	Có tưới	46,5	S2	
13	23	Xám	> 100	Nhẹ	D0 - 3	Có tưới	48,1	S1	
14	30	Phù sa	> 100	Nặng	D0 - 3	Có tưới	53,3	S1	
15	29	Đỏ vàng Fp	70 - 100	Nhẹ	D0 - 3	Có tưới	54,0	S1	
16	32	Đỏ vàng Fp	> 100	Nhẹ	D3 - 8	Có tưới	55,3	S1	
17	34	Đỏ vàng Fp	> 100	Nhẹ	D0 - 3	Có tưới	56,3	S1	
18	35	Phù sa	> 100	Trung bình	D0 - 3	Có tưới	56,6	S1	

Kết quả bảng 5 cho thấy mức độ thích hợp cây cam theo các tổ hợp tính chất đất đai chịu sự ảnh hưởng lớn nhất của yếu tố độ dốc. Nhóm có độ dốc từ 8 - 15 và trên 15 độ, có năng suất trung bình là 40,5 tấn/ha, nhóm có độ dốc từ 0-3 và 3 - 8 độ, có năng suất trung bình là 49,7 tấn/ha. Như vậy, độ dốc thấp từ 0-3 độ và 3-8 độ phù hợp hơn cho sự phát triển của cây cam. Điều này có thể được lý giải do địa hình dốc cao thì khả năng giữ nước và dinh dưỡng bị hạn chế hơn do rửa trôi nên có ảnh hưởng không tốt đến năng suất của cây cam.

Tiếp theo là nhóm loại đất, ở nhóm này, đất đỏ vàng trên phù sa cổ và đất phù sa thích hợp nhất, cho năng suất đạt trung bình là 52,7 tấn/ha. Nhóm đất đỏ vàng trên đá phiến sét và đất xám

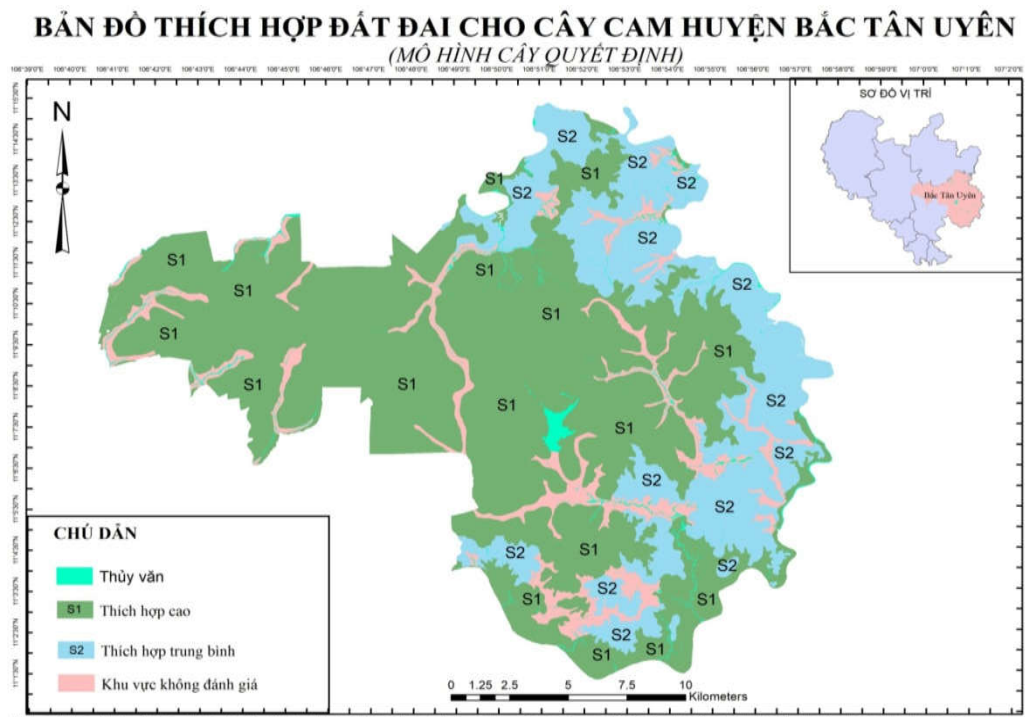
chỉ cho năng suất trung bình là 45,1 tấn/ha. Điều này cho chúng ta thấy đất phù sa và đỏ vàng trên phù sa cổ thích hợp với cây cam cao hơn đất đỏ vàng trên đá phiến sét và đất xám.

Nhóm đất có độ dày tầng đất trên 100 cm cho năng suất đạt 55,6 tấn/ha, nhóm đất có tầng đất dày 70 – 100 cm cho năng suất 50,5 tấn/ha. Trong khi đó nhóm đất có độ dày 50 - 70 cm chỉ đạt năng suất 42,7 tấn/ha. Điều này cho ta thấy cây cam thích hợp hơn với những loại đất có độ dày tầng đất lớn hơn.

Mô hình tiếp tục phân nhánh theo thành phần cơ giới, thành phần cơ giới nhẹ và trung bình có năng suất trung bình đạt 56 tấn/ha, cao hơn nhánh có thành phần cơ giới nặng chỉ đạt 53,3 tấn/ha.

Bảng 6. Tổng hợp đánh giá mô hình cây quyết định theo diện tích

Mức thích hợp	Diện tích (ha)	Tỷ lệ (%)
S1	25.796,99	66,2
S2	9.652,03	24,8
Không đánh giá	3.495,41	9,0
Tổng	38.944,44	100



Hình 4. Bản đồ thích hợp đất đai cho cây cam theo mô hình cây quyết định

3.4. Đánh giá thích hợp đất đai theo yếu tố hạn chế lớn nhất của FAO

Để có cơ sở nghiên cứu, so sánh đồng thời thực hiện đánh giá thích hợp cây cam trên cùng

địa bàn, chúng tôi sử dụng phương pháp đánh giá truyền thống theo phương pháp yếu tố hạn chế lớn nhất của FAO.

Bảng 7. Mức thích hợp của các yếu tố tự nhiên với cây cam

Chỉ tiêu	Thích hợp			
	S1	S2	S3	N
Loại đất	Phù sa	Đất xám, Đất đỏ Fp(*)	Đất đỏ Fs(**)	Đất cát, nâu, xám gley
Độ dốc (°)	0 - 3	3 - 8	8 - 15	> 15
Độ dày tầng đất (cm)	> 100	70 - 100	50 - 70	< 50
Thành phần cơ giới	Trung bình	Nặng	Nhẹ	Cát nhẹ
Khả năng tưới		Có tưới		Không tưới

Nguồn: Vũ Cao Thái và cộng sự, 1997

(*): Đất đỏ Fp: Đất đỏ vàng trên phù sa cổ; (**): Đất đỏ Fs: Đất đỏ vàng trên đá phiến sét

Bảng 8. Tổng hợp đánh giá mô hình cây quyết định theo diện tích

Mức thích hợp	Diện tích (ha)	Tỷ lệ (%)
S1	968,04	2,5
S2	1.792,44	4,6
S3	29.921,11	76,8
N	2.767,43	7,1
Không đánh giá	3.495,41	9,0
TỔNG	38.944,44	100

3.5. So sánh kết quả đánh giá thích hợp theo phương pháp yếu tố hạn chế lớn nhất của

FAO và phương pháp khai phá dữ liệu theo mô hình cây quyết định

Bảng 9. So sánh các mức thích hợp về diện tích theo hai phương pháp

Mức thích hợp	FAO		CQĐ	
	Diện tích (ha)	Tỷ lệ (%)	Diện tích (ha)	Tỷ lệ (%)
S1	968,04	2,5	25.796,99	66,2
S2	1.792,44	4,6	9.652,03	24,8
S3	29.921,11	76,8	-	-
N	2.767,43	7,1	-	-
Không đánh giá	3.495,41	9,0	3.495,41	9,0
Tổng	38.944,44	100	38.944,44	100

Bảng 9 cho thấy đa số diện tích đất trên địa bàn huyện Bắc Tân Uyên thích hợp cho việc trồng cây cam theo cả hai phương pháp đánh giá. Phương pháp đánh giá định tính theo FAO có thích hợp cao (S1) chiếm tỷ lệ 2,5%, thích hợp trung bình (S2) chiếm tỷ lệ 4,6%, thích hợp kém (S3) chiếm tỷ lệ 76,8% và không thích hợp chỉ có 7,1%. Trong khi đánh giá định lượng bằng mô hình cây quyết định thì thích hợp cao (S1) chiếm tỷ lệ 66,2%, thích hợp trung bình (S2) chiếm tỷ lệ 24,8%, không có kết quả kém thích hợp và không thích hợp.

4. KẾT LUẬN

Tại khu vực nghiên cứu ở huyện Bắc Tân Uyên, tỉnh Bình Dương, mô hình cây quyết

định, mô hình chia tách thành 18 đơn vị đất đai, trong đó có 6 đơn vị đất đai được đánh giá mức thích hợp S1 chiếm 33% và 12 đơn vị đất đai được đánh giá mức thích hợp S2, chiếm 67%.

Theo diện tích, đánh giá thích hợp đất đai theo mô hình cây quyết định đã cho thích hợp cao (S1) có diện tích 25.796 ha chiếm 66,2% và thích hợp trung bình (S2) có diện tích 9.652 ha, chiếm 24,8% so với diện tích đất tự nhiên.

TÀI LIỆU THAM KHẢO

1. 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 and Management*. 9(4): 131-139.
2. Đào Châu Thu, Nguyễn Khang, 1998. *Đánh giá*

đất. NXB Nông nghiệp.

3. Hội Khoa học Đất Việt Nam, 2000. *Đất Việt Nam*. NXB Nông nghiệp.

4. 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.

5. 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.

6. 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 – Đại học Nông Lâm TP. Hồ Chí Minh.

7. Nguyễn Hữu Cường, 2018. Ứng dụng kỹ thuật khai phá dữ liệu đánh giá thích nghi đất đai cây cao su trên địa bàn huyện Phú Giáo, tỉnh bình dương. *Tạp chí Khoa học Trường Đại học Cần Thơ*, 54(3B), trang 84-93.

8. Nguyễn Xuân Nhiệm, Phạm Quang Khánh, 2014.

Nghiên cứu suy thoái và ô nhiễm môi trường đất nông nghiệp tỉnh Bình Dương. Báo cáo tổng kết đề tài khoa học, Sở Khoa học và Công nghệ tỉnh Bình Dương.

9. Phạm Quang Khánh, 1995. *Tài nguyên đất vùng Đông Nam Bộ, hiện trạng tiềm năng*. NXB Nông nghiệp.

10. Phạm Quang Khánh, 2011. Tài nguyên đất tỉnh Bình Dương. *Tạp chí Khoa học Đất*, số 36, trang 31-37.

11. Phan Liêu, 1992. *Đất Đông Nam Bộ*, NXB Nông nghiệp.

12. Tôn Thất Chiêu, Lê Thái Bạt, 1998. *Nghiên cứu ứng dụng phương pháp phân loại đất quốc tế FAO/UNESCO ở Việt Nam*, Hà Nội.

13. Trần An Phong, 1995. *Đánh giá hiện trạng sử dụng đất theo quan điểm sinh thái và phát triển lâu bền*. NXB Nông nghiệp, Hà Nội.

14. Trần Công Tấu, 2006. *Tài nguyên đất*. NXB Đại học Quốc gia, Hà Nội.

15. Vũ Cao Thái, Phạm Quang Khánh, Nguyễn Văn Khiêm, 1997. *Điều tra, đánh giá tài nguyên đất đai theo phương pháp FAO/UNESCO và qui hoạch sử dụng đất (trên địa bàn một tỉnh)*. NXB Nông nghiệp.

APPLICATION OF THE DECISION TREE MODEL TO EVALUATE THE SUITABILITY OF LAND FOR ORANGE TREES IN BAC TAN UYEN DISTRICT, BINH DUONG PROVINCE

Vo Quoc Khanh*, Nguyen Van Cuong, Truong Thi Dieu Quan
Ho Chi Minh City University of Natural Resources and Environment

SUMMARY

Currently, the area of orange cultivation has increased rapidly year by year in Bac Tan Uyen district, but there were not any systematic and scientific studies on suitable soil for orange trees there. Study: "Application of the decision tree method to evaluate the suitability of land for orange trees in Bac Tan Uyen district, Binh Duong province" aims to evaluate the suitability of land for orange trees. That supports the foundation of land use planning. The decision tree model is a quantitative land suitability evaluation method with two kinds of variables. The target variable is the crop yield data collected at the farms (t/ha) and the predictor variables consist of soil types, soil depth, slope and irrigation. This method determines the relationship between land characteristic factors and plant productivity in order to improve land evaluation methods. The study is performed on 127 highly reliable questionnaires used to run the software and analyze data to build a decision tree model, the model is divided into 18 land units, of which 6 land units evaluated as highly suitable (S1) accounted for 33% and 12 land units evaluated as suitable (S2), accounting for 67%. Counting on area, evaluating land by decision tree model gave high suitability (S1) with an area of 25,796 hectares, which accounted for 66.2% and suitability (S2) with an area of 9,652 hectares, which accounted for 24.8% of the natural land area. This result is different from the FAO's highest limitation factor method.

Keywords: Data mining, land evaluation, orange decision tree, orange tree.

Ngày nhận bài : 09/6/2022

Ngày phản biện : 10/7/2022

Ngày quyết định đăng : 29/7/2022