

Ứng dụng phương pháp khai phá dữ liệu đánh giá thích hợp đất trồng cam tại huyện Dầu Tiếng, tỉnh Bình Dương

Võ Quốc Khánh*, Nguyễn Văn Cường

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

Application of data mining techniques to evaluate the suitability of orange cultivation soil in Dau Tieng district, Binh Duong province

Vo Quoc Khanh*, Nguyen Van Cuong

Ho Chi Minh City University of Natural Resources and Environment

*Corresponding author: khanh692002@yahoo.com

<https://doi.org/10.55250/jo.vnuf.13.4.2024.099-107>

TÓM TẮT

Hiện nay, tại huyện Dầu Tiếng đang phát triển trồng cây cam nhưng chưa có những nghiên cứu đánh giá sự thích hợp đất đai một cách định lượng, nghiên cứu này được thực hiện với mục tiêu đánh giá từng mức thích hợp đất đai cho cây cam tại địa bàn. Phương pháp điều tra là khảo sát trực tiếp nông dân trồng cam theo bảng mẫu điều tra, tổng số điều tra 136 mẫu trên 11 đơn vị đất đai có trồng cây cam tại địa phương. Nghiên cứu sử dụng phương pháp hồi quy toán học bằng phần mềm DTREG thiết lập mô hình cây dựa trên năng suất cam được khảo sát từ phiếu điều tra và các biến dự báo là những thuộc tính nông học của đất đai như, độ dày đất, thành phần cơ giới, loại đất, mức gley và độ dốc để phân nút theo từng mức năng suất được đưa vào với những thuộc tính đất đai tương ứng. Sau đó, dựa theo phân cấp từng mức năng suất thích hợp của FAO để phân cấp sự thích hợp của cây cam với những thuộc tính đó. Kỹ thuật này thiết lập được 11 đơn vị đất đai, 6 tầng và 21 nút với 2 nhóm đất đai thích hợp S1 (cao), diện tích 68.961,7 ha chiếm 77,7%; 9 đơn vị đất đai có mức thích hợp S2 (vừa) với 16.945,8 ha, chiếm 19,1%; (còn 3,2 % diện tích không đánh giá). Nghiên cứu này bổ sung một kỹ thuật đánh giá đất đai cho cây cam theo định lượng, bổ sung vào đánh giá theo yếu tố hạn chế lớn nhất của FAO đề xuất, giúp cho công tác phát triển cây cam được tốt hơn, mang lại hiệu quả kinh tế cao.

Thông tin chung:

Ngày nhận bài: 05/04/2024

Ngày phản biện: 09/05/2024

Ngày quyết định đăng: 03/06/2024

Từ khóa:

Cây cam, cây quyết định, đánh giá đất đai, khai phá dữ liệu, thích hợp đất đai cho cây cam.

Keywords:

Data mining, decision tree, land evaluation, land suitability for orange tree, orange tree.

ABSTRACT

Currently, in Dau Tieng district, orange tree growing is being developed but there has been no research to assess land suitability quantitatively. This study was carried out with the goal of evaluating each level of land suitability for the orange tree in the area. The survey method was to directly interview farmers who are planting orange trees according to the survey form, with a total of 136 samples, distributed over 11 land units with grow orange trees locally. The study uses mathematical regression with DTREG software that establishes a tree model based on orange yield surveyed from the survey questionnaire and the predictive variables which are agronomic properties of the land such as soil thickness, mechanical composition, soil type, gley level and slope. The model is classified into nodes according to each productivity level. Then, based on FAO's hierarchy of appropriate productivity levels, we can classify the suitability of orange trees with those attributes. This technique establishes 11 land units, 6 floors and 21 nodes with 2 groups of suitable level S1 (high), an area of 68,961.7 hectares, accounting for 77.7%; 9 groups of suitable level S2 (moderate) with 16,945.8 hectares, accounting for 19.1%; (3.2% of the area not be evaluated). This study adds a quantitative land assessment technique for orange growing, supplemented with the assessment based on the biggest limiting factor proposed by FAO, helping to improve the development of orange trees, bringing more benefits, and high economic efficiency.

1. ĐẶT VẤN ĐỀ

Những năm qua tại địa phương, bên cạnh những loại cây chủ lực có truyền thống lâu nay như cao su, tiêu, điều đang ngày càng mất giá, nhiều nhà nông chuyển qua trồng cây cam [1]. Việc này đem lại một số kết quả khả quan trong trồng trọt nhưng bà con nông dân tại đây đang trồng cam theo phương pháp riêng của mỗi hộ gia đình bằng kinh nghiệm của họ chứ chưa có những đánh giá thích hợp của từng loại đất đai cho cây cam trên cơ sở định lượng một cách khoa học.

Kỹ thuật xây dựng mô hình bằng phương pháp hồi quy toán học chia tách dữ liệu được đề xuất là phương pháp chia tách những giá trị hữu ích từ nguồn dữ liệu đưa vào để phân tích dự báo về mối quan hệ của chúng. Bên cạnh một số phần mềm ứng dụng cho kỹ thuật này thì DTREG là phần mềm được thiết kế để đưa ra mô hình cây quyết định với các yếu tố nông học và năng suất cây trồng được đưa vào để phân tích hồi quy. Kỹ thuật này chia tách các dữ liệu đầu vào thành từng nhóm với các đặc điểm cho năng suất tương đồng [3]. Hiện nay, đã có nhiều nghiên cứu về vấn đề này. Trên thế giới, Andi Nurkholis đã thực hiện phương pháp này trên cây tỏi [2]. Tại Việt Nam, Nguyễn Hữu Cường đã thực hiện trên cây cao su tại huyện Phú Giáo [4] và cây dứa tại huyện Mỏ Cày Nam, Bến Tre [5].

Tại địa bàn nghiên cứu cho đến nay chưa có đánh giá đất đai nào sử dụng kỹ thuật này cho cây cam, chính vì thế việc sử dụng DTREG để đánh giá các mức thích hợp đất đai trong nghiên cứu này được tiến hành nhằm: 1) Trình bày các kết quả phân tích của mô hình cây quyết định; 2) Trình bày các kết quả đánh giá theo phương pháp của FAO; 3) So sánh kết quả của hai phương pháp trên, để từ đó đề xuất phương pháp này nên được ứng dụng. Những kết quả này sẽ là cơ sở khoa học quan trọng để xác định vị trí trồng cây cam thích hợp tại khu vực nghiên cứu, đồng thời quản lý và phát triển nguồn tài nguyên đất đai một cách hiệu quả trong tương lai.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Địa điểm nghiên cứu

Địa điểm nghiên cứu tại huyện Dầu Tiếng có

62.205,8 ha đất nông nghiệp với 56.903,42 ha là đất cây lâu năm [6]. Huyện Dầu Tiếng có địa hình, nơi cao nhất 25 m gần khu vực trung tâm huyện và nơi thấp nhất có độ cao 4,3 m. Dầu Tiếng có hai mùa mưa và nắng. Mùa tháng 5 qua tháng 12 và mùa nắng từ tháng 12 đến tháng 4. Nhiệt độ bình quân trong năm khoảng 27°C, nhiệt độ nóng nhất 32,7°C và mát nhất là 22,8°C. Trong năm có trung bình 2.888 giờ nắng, giữa các mùa tương đối ổn định về số giờ nắng, điều này giúp các thực vật phát triển thuận lợi. Lượng mưa bình quân 2.101,5 mm/năm, trung bình mỗi năm có 141 ngày mưa. Cây lâu năm trồng nhiều ở đây là cao su, điều, hồ tiêu và cây ăn quả. Các loại cây như điều, lúa, khoai mì chiếm diện tích lớn lại cho hiệu quả kinh tế thấp, trong khi các loại cây ăn quả, hồ tiêu, rau màu có tỷ trọng thấp lại cho hiệu quả kinh tế cao [7].

2.2. Phương pháp thu thập xử lý thông tin

2.2.1. Phương pháp thu thập số liệu

Số liệu được thu thập bằng phương pháp khảo sát người trồng cam tại địa bàn thực hiện nghiên cứu với 136 mẫu trên tất cả diện tích đất có trồng cam. Các số liệu được xử lý và đưa vào phần mềm DTREG.

2.2.2. Phương pháp xử lý số liệu

Mô hình cây quyết định được thiết lập sắp xếp các ô theo mức năng suất [7] (decision tree) được trình bày dưới dạng cây nhị phân (tách hai chiều) xác định giá trị của biến năng suất bằng cách dự đoán các yếu tố dự báo. Các biến dự báo được đưa vào nghiên cứu gồm có: độ dày tầng đất, độ dốc, loại đất, mức độ gley và thành phần cơ giới [3]. Các hộp hình chữ nhật được hiển thị trong cây được gọi là "nút". Mỗi nút đại diện cho một tập hợp từ tập dữ liệu gốc. Có nút gốc và các nút con "Interior". Các nút cuối không tiếp tục chia nút con gọi là nút lá (Leaf) [8]. Số liệu năng suất thu thập được tổng hợp trên file Excel, bảng tổng hợp này đưa vào phần mềm DTREG và cho chạy, phần mềm sẽ xuất ra file kết quả bao gồm các giá trị năng suất của từng nút lá (nút cuối) và mô hình hộp cho những giá trị biến mục tiêu theo những tổ hợp (nút) của các biến dự báo đưa vào. Dựa trên file kết quả, mô hình này và bảng phân cấp thích hợp theo năng suất do FAO đề xuất để

đánh giá khả năng thích hợp cho cây cam trên từng loại đất.

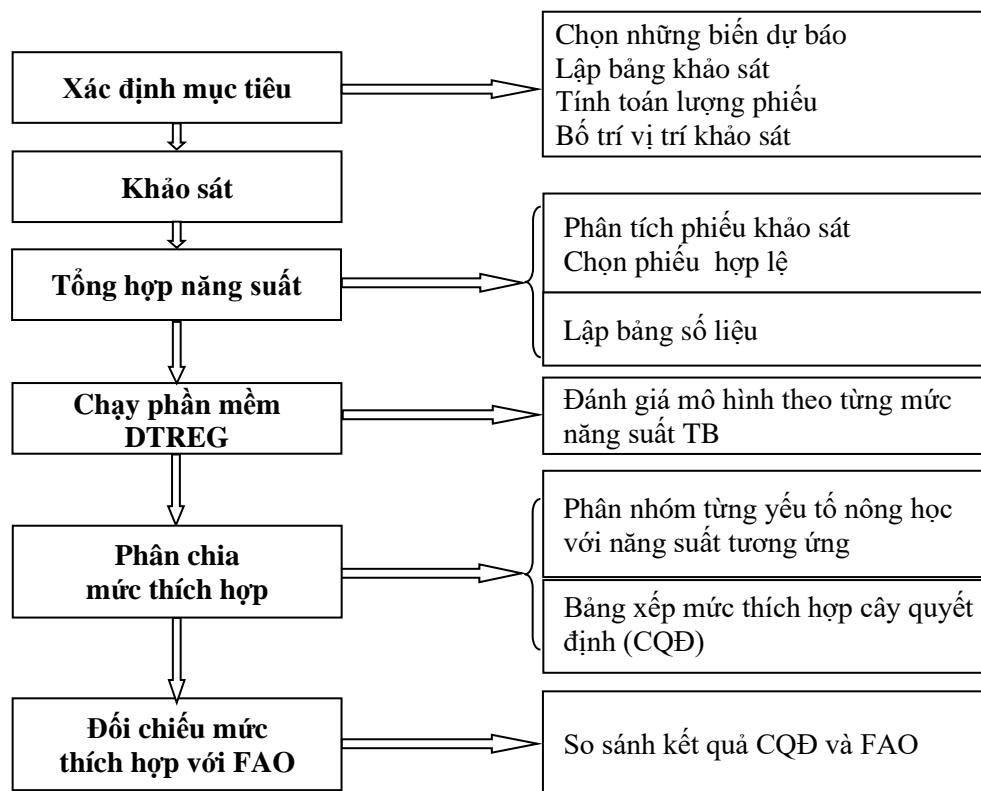
2.2.3. Phương pháp bản đồ

Sau khi đánh giá các mức thích hợp cho cây cam trên từng loại đất, đề tài chạy phần mềm ArcGis để thiết lập bản đồ trên từng diện tích đất thích hợp của cây cam tại địa điểm thực

hiện nghiên cứu.

2.3. Các bước thực hiện nghiên cứu

Xác định các yếu tố nông học đưa vào là độ dày tầng đất, độ dốc, nhóm đất, mức độ gley và thành phần cơ giới. Năng suất cam khảo sát tại địa phương được đưa vào làm biến độc lập. Các bước tiến hành theo trình tự ghi trong Hình 1.



Hình 1. Sơ đồ thực hiện đánh giá thích hợp cho cây cam theo cây quyết định

2.4. Số lượng mẫu điều tra

Các mẫu khảo sát được chia ra trên toàn diện tích đất trồng cam của huyện Dầu Tiếng,

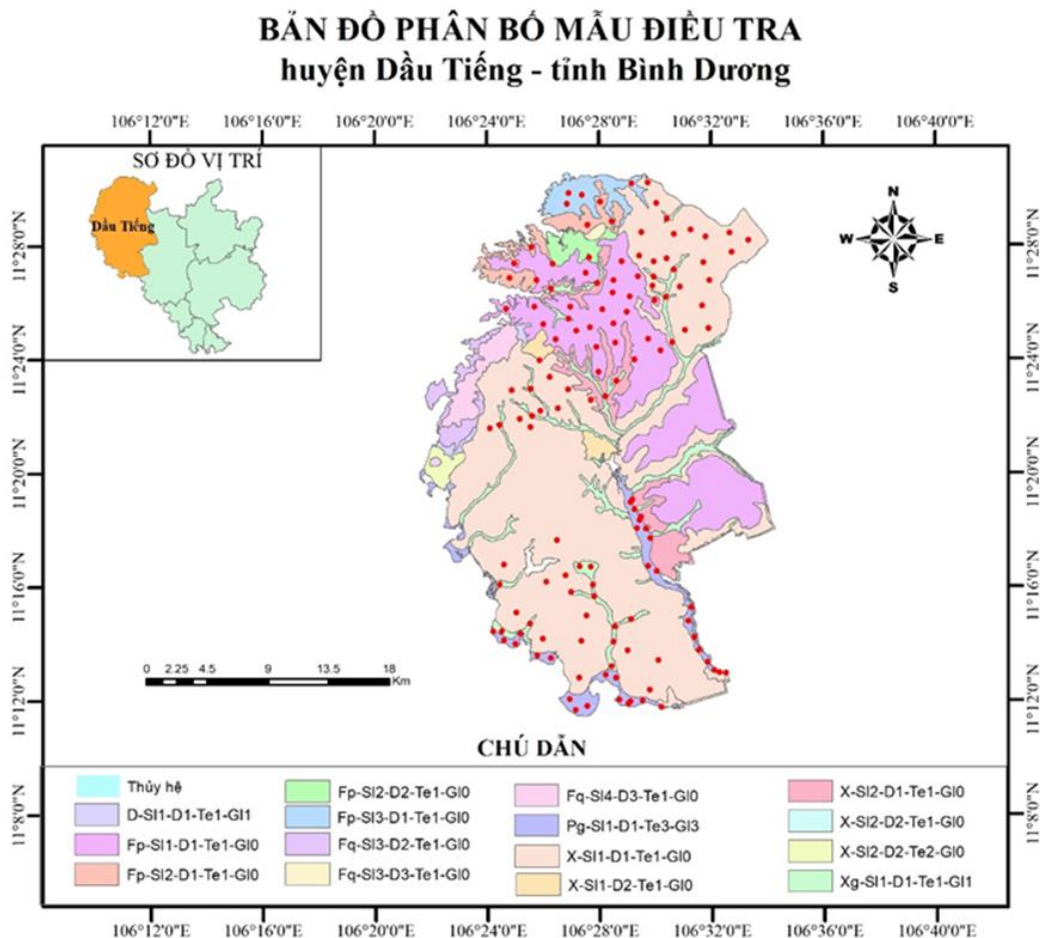
tổng cộng có 136 mẫu năng suất được khảo sát năng suất từ người trồng cam. Số mẫu thu thập từ các biến dự báo được trình bày dưới đây.

Bảng 1. Số mẫu khảo sát năng suất cam

STT	Tổ hợp các yếu tố					Số mẫu điều tra
	Nhóm đất	Độ dày đất (cm)	TPCG	Độ dốc (°)	Mức độ gley	
1	Phù sa Gley	>100	Nặng	0-3	Mạnh	30
2	Xám	70-100	Cát pha	3-8	Không	5
3	Xám Gley	>100	Thịt nhẹ	0-3	Ít	9
4	Xám	70-100	Thịt nhẹ	3-8	Không	9
5	Đỏ vàng Fp	>100	Thịt nhẹ	8-15	Không	5
6	Xám	>100	Thịt Nhẹ	3-8	Không	8
7	Xám	70-100	Thịt nhẹ	0-3	Không	5
8	Đỏ vàng Fp	70-100	Thịt nhẹ	3-8	Không	7
9	Đỏ vàng Fp	>100	Thịt nhẹ	3-8	Không	4
10	Xám	>100	Thịt Nhẹ	0-3	Không	34
11	Đỏ vàng Fp	>100	Thịt nhẹ	0-3	Không	20
Tổng số						136

Sau khi tính toán số phiếu khảo sát trên từng loại đất theo Bảng 1, nhóm tác giả đã xây

dựng sơ đồ lấy mẫu điều tra cho địa bàn thực hiện nghiên cứu như Hình 2.



Hình 2. Sơ đồ các điểm khảo sát

2.5. Xác định các biến cho nghiên cứu

Số liệu thu thập cho chương trình thiết lập mô hình hồi quy cây quyết định theo các yếu tố sau:

Các biến dự báo (Predictor) là thành phần cơ giới, độ dày đất, loại đất, mức độ gley và độ dốc. Các biến trên được chia theo các thuộc tính:

Nhóm đất gồm có các loại đất phù sa, đỏ vàng trên phù sa cổ, đất xám và đất xám có Gley. Độ dày đất được tách ra 3 lớp: 50-70 cm; 70-100 cm, 100 cm. Thành phần cơ giới gồm có 3 cấp thịt nặng, thịt nhẹ và cát pha. Độ dốc có 3 mức: 0-3°, 3-8°, > 8° và cuối cùng là mức độ gley gồm có 3 mức: không có gley, gley nhẹ và gley mạnh.

Năng suất từ kết quả khảo sát được xác

định là biến mục tiêu.

Với các dữ liệu này, phần mềm DREG đã xây dựng cây quyết định theo từng nhóm năng suất cây cam gồm có 6 tầng, 21 nhóm trong đó có 11 nút cuối (nút lá).

3. KẾT QUẢ - THẢO LUẬN

3.1. Kết quả của mô hình

3.1.1. Sự phù hợp của mô hình

Từ năng suất thu thập được, phần mềm DTREG tính toán phương sai của dữ liệu đưa vào ban đầu là 85,5. Khi đã thiết lập mô hình thì phương sai là 7,54 (đây là phương sai phần mềm không giải thích được). Chương trình tính toán tỷ lệ phương sai giải thích được là 91,2%. Mức độ này thể hiện mô hình toán học phù hợp trong đánh giá định lượng sự thích hợp của cây cam trên từng loại đất.

Bảng 2. Tỷ lệ phương sai được giải thích

STT	Phương sai	Số liệu
1	Phương sai trước khi chạy mô hình	85,5
2	Phương sai khi đã thiết lập mô hình	7,54
3	Phần trăm phương sai giải thích được	91,2%

3.1.2. Độ quan trọng của biến dự báo

Trong quá trình chạy mô hình toán học, chương trình DTREG dựa vào ảnh hưởng của

các chỉ tiêu nông học đến năng suất cam đưa vào để tính toán tỷ lệ quan trọng của từng chỉ tiêu đến mô hình như Bảng 3.

Bảng 3. Độ quan trọng của các yếu tố

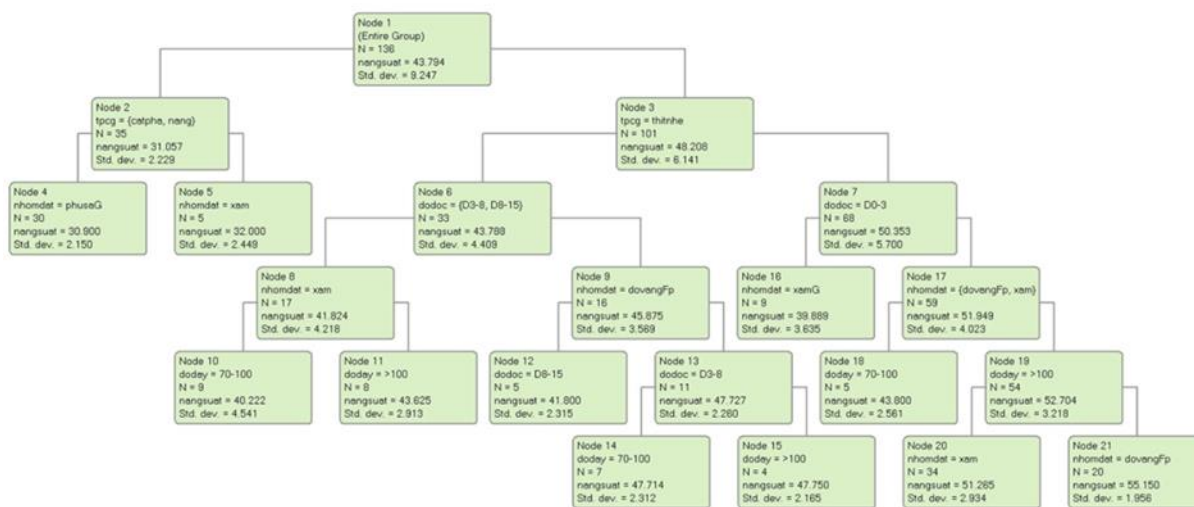
STT	Chỉ tiêu	Tỷ lệ quan trọng
1	Nhóm đất	100
2	TPCG	97,3
3	Mức độ Gley	95,9
4	Độ dốc	14,1
5	Độ dày đất	13,1

Như vậy, theo phân tích mối liên quan của năng suất cây cam và các yếu tố nông học thì yếu tố độ dày đất là ít quan trọng nhất với độ quan trọng là 13,1%, nhóm đất là chỉ tiêu quan trọng nhất với độ quan trọng là 100%. Các yếu tố còn lại như thành phần cơ giới, mức độ gley và độ dốc có các độ quan trọng lần lượt là 97,3; 95,9 và 14,1.

Phân tích năng suất cây cam theo CQĐ

Mô hình đã chia tách bộ dữ liệu năng suất cam thành 6 tầng và 11 nút cuối dựa trên sự tác động của các chỉ tiêu nông học đến năng suất.

Giá trị trung bình của tất cả năng suất đưa vào phần mềm là 43,79 tấn/ha được thể hiện ở nút gốc (nút đầu tiên). Sau đó, yếu tố thành phần cơ giới được chọn để chia tách thành 2 nút tiếp theo, nút 2 là cát pha và sét nặng, nút 3 là thịt nhẹ. Đất thịt nhẹ có năng suất cao hơn được chia theo bên phải của mô hình, có năng suất trung bình 48,2 tấn/ha trong 101 phiếu thu thập, còn node 2 gồm có đất cát pha và sét nặng, gồm 35 phiếu khảo sát cho năng suất 31,1 tấn/ha. Mô hình này phù hợp với yêu cầu sử dụng đất của cây cam [10].



Hình 3. mô hình cây quyết định
(Node: Nút; N: số mẫu; Std dev: Độ lệch chuẩn)

Theo FAO, năng suất tại vườn của cây trồng so với mức tối hảo đạt được của cây trồng đó quyết định mức thích hợp [11]. Từng mức thích

hợp cho cây cam theo năng suất tối hảo (60 tấn/ha/năm) được trình bày theo Bảng 4.

Bảng 4. Các mức thích hợp theo năng suất cam

STT	Mức thích hợp	Tỷ lệ (%)	Năng suất cam (tấn/ha/năm)
1	S1	> 80	> 48
2	S2	40 - 80	24 - 48
3	S3	20 - 40	12 - 24
4	N	< 20	< 12

Chú thích: S1: Thích hợp cao; S2: Thích hợp TB; S3: Thích hợp kém; N: Không thích hợp.

Theo sự phân cấp thích hợp theo hướng dẫn của FAO ở bảng trên, kết hợp với năng suất trung bình mà phần mềm DTREG đã phân

nhóm, đánh giá mức độ thích hợp cho cây cam như Bảng 5.

Bảng 5. Các mức thích hợp đất đai của cây cam (theo mô hình)

STT	Điểm nút	Các mức chỉ tiêu					Năng suất cam	Mức thích hợp
		Nhóm đất	TPCG	Độ dày đất (cm)	Mức độ Gley	Độ dốc (°)		
1	4	Phù sa Gley	Nặng	>100	Mạnh	0-3	30,9	S2
2	5	Xám	Cát pha	70-100	Không	3-8	32,0	S2
3	16	Xám Gley	Thịt nhẹ	>100	Ít	0-3	39,9	S2
4	10	Xám	Thịt nhẹ	70-100	Không	3-8	40,2	S2
5	12	Đỏ vàng Fp	Thịt nhẹ	>100	Không	8-15	41,8	S2
6	11	Xám	Thịt Nhẹ	>100	Không	3-8	43,6	S2
7	18	Xám	Thịt nhẹ	70-100	Không	0-3	43,8	S2
8	14	Đỏ vàng Fp	Thịt nhẹ	70-100	Không	3-8	47,7	S2
9	15	Đỏ vàng Fp	Thịt nhẹ	>100	Không	3-8	47,8	S2
10	20	Xám	Thịt Nhẹ	>100	Không	0-3	51,3	S1
11	21	Đỏ vàng Fp	Thịt nhẹ	>100	Không	0-3	55,2	S1

Từ kết quả đánh giá trên, tác giả tính toán bảng tổng hợp diện tích cho từng mức thích

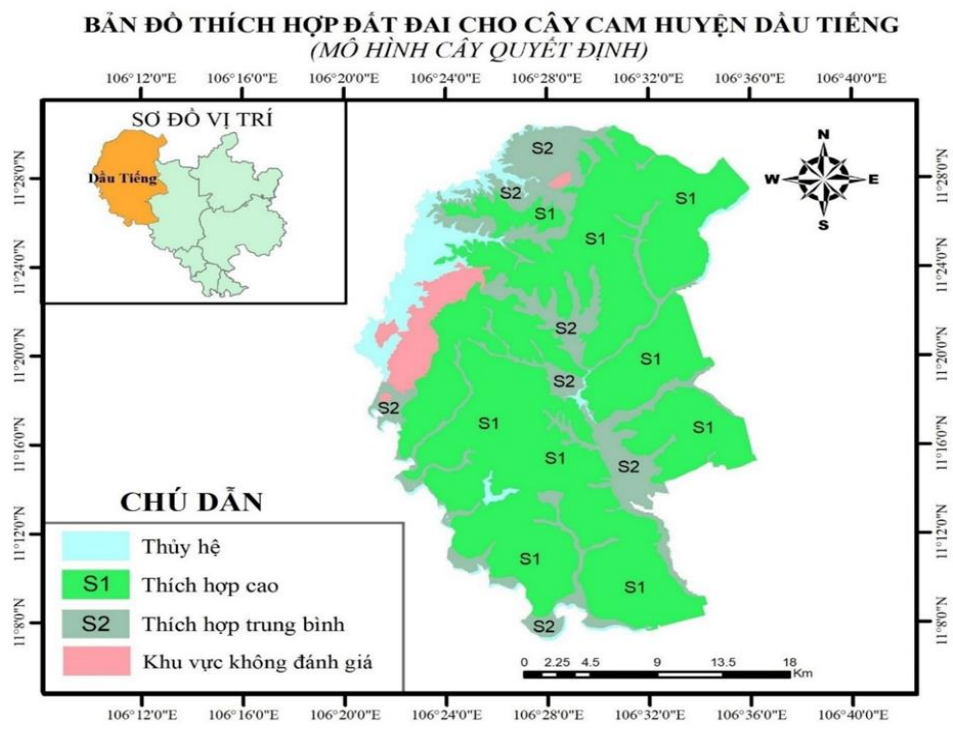
hợp (S1, S2) đất đai cho cây cam trên địa bàn huyện Dầu Tiếng, tỉnh Bình Dương như Bảng 6.

Bảng 6. Diện tích từng mức thích hợp đất đai với cây cam

Mức thích hợp	Diện tích (ha)	Tỷ lệ (%)
S1	68.961,7	77,7
S2	16.945,8	19,1
Không đánh giá	2.797,4	3,2
Tổng	88.705,0	100

Như vậy, với 136 phiếu điều tra của 11 đơn vị đất đai đưa vào đánh giá, kết quả có 2 đơn vị đất đai với diện tích 68.961,7 ha (chiếm 77,7%) thích hợp S1; 9 đơn vị đất đai được có diện tích 16.945,8 ha (chiếm 19,1%) thích hợp S2. Điều này được lý giải là người dân từ trước đã không trồng cam trên những diện tích đất có thích

hợp kém và không thích hợp. Kết quả này có tính tương đồng với những kết quả đã nghiên cứu tại huyện Bắc Tân Uyên [9]. Từ những mức thích hợp được đánh giá ở trên, bài báo này thiết lập bản đồ theo từng mức thích hợp cho cây cam tại huyện Dầu Tiếng như Hình 4.



Hình 4. Bản đồ vị trí các mức thích hợp đất đai cho cây cam

3.2. Kết quả đánh giá theo phương pháp FAO

Dựa trên những yếu tố nông học tại địa bàn nghiên cứu, bài báo đã đánh giá thích hợp đất đai theo phương pháp FAO làm cơ sở so sánh

kết quả của hai phương pháp, từ đó có thể tìm ra sự khác biệt và tương đồng để có hướng bổ sung cho nhau trong đánh giá thích hợp đất đai.

Bảng 7. Yêu cầu sử dụng đất cho cây cam

Chỉ tiêu	Mức độ thích hợp			
	S1	S2	S3	N
Nhóm đất	Đất đỏ vàng trên phù sa cổ	Đất xám	Phù sa có Gley	Xám gley Đốc tụ Đất đỏ trên đá
TPCG	Thịt nhẹ	Cát pha	Nặng	Cát nhẹ
Mức độ Gley	Không	ít	Trung bình	Mạnh
Độ dày đất (cm)	> 100	70-100	50-70	<50
Độ dốc (°)	0-3	3-8	8 - 15	> 15

Nguồn: Vũ Cao Thái và cộng sự, 1997 [10].

Kết quả đánh giá thích hợp đất đai cho cây cam theo phương pháp hạn chế lớn nhất của FAO ở Bảng 7, cho diện tích từng mức thích hợp

(S1, S2, S3, N) đất đai cho cây cam tại địa bàn huyện Dầu Tiếng, tỉnh Bình Dương như Bảng 8.

Bảng 8. Diện tích các mức thích hợp theo FAO

Mức thích hợp	DT (ha)	(%)
S1	19.236,6	21,7
S2	57.432,5	64,7
S3	1.549,9	1,7
N	10.486,0	11,8
Tổng	88.705,0	100

Đánh giá theo phương pháp FAO hướng dẫn thì ở mức thích hợp S1 chỉ có một đơn vị đất (diện tích 19.236,6 ha) trong khi đó có 07 đơn vị đất thích hợp S2 (diện tích 57.432,5 ha); 01 đơn vị đất đai thích hợp S3 (diện tích 1.549,9 ha). Bên cạnh đó phương pháp này cũng cho

kết quả 06 đơn vị đất có mức N với diện tích 10.486,0 ha (chiếm 11,8%). Điều này cũng cho thấy phương pháp này có thể đánh giá được những diện tích đất không thích hợp chưa có trồng cây cam.

3.3. Kết quả giữa hai phương pháp đánh giá

Bảng 9. Diện tích các mức thích hợp đất đai của từng phương pháp

Mức thích hợp	Theo cây quyết định		Theo FAO	
	DT (ha)	(%)	DT (ha)	(%)
S1	68.961,7	77,7	19.236,6	21,7
S2	16.945,8	19,1	57.432,5	64,7
S3	-	-	1.549,9	1,7
N	-	-	7.688,6	8,7
Không so sánh	2.797,4	3,2	2.797,4	3,2
Tổng	88.705,0	100	88.705,0	100

Qua Bảng 9 trên cho thấy, cây cam có mức độ phù hợp trung bình và cao chiếm diện tích khá lớn, như vậy việc tập trung phát triển cây cam tại địa phương là phù hợp.

Đánh giá theo phương pháp cây quyết định thì có 02 đơn vị đất đai thích hợp S1 với 68.961,7 (chiếm đến 77,7%) và 9 đơn vị đất thích hợp S2 với 16.945,8 ha (chiếm 19,1%). Trong khi đó nếu đánh giá theo FAO chỉ có một ĐVĐĐ thích hợp S1 với diện tích 19.236,6 ha (tỷ lệ 21,7%). Điều này được lý giải là đánh giá theo phương pháp hạn chế lớn nhất của FAO chưa tính đến sự tương hỗ qua lại của các thuộc tính nông học đất đai mà chỉ dựa trên một yếu tố bất lợi nhất quyết định đến kết quả chung. Tuy nhiên trên thực tế nếu có một yếu tố hạn chế nhưng các yếu tố còn lại tốt hơn cũng có thể có quan hệ tương hỗ và vẫn cho năng suất cao, chính vì vậy nên ở hàng thứ 4 của Bảng 5, chỉ có loại đất xám là mức thích hợp S2 vẫn bị đánh giá là thích hợp S2 theo phương pháp FAO trong khi tất cả các yếu tố còn lại đều có mức thích hợp là S1. Nhưng khi đánh giá theo phương pháp định lượng dựa vào mức năng suất thu thập được, thì đơn vị đất này đạt kết quả S1 do cho năng suất cao (51,2 tấn/ha) bởi tác động qua lại của các yếu tố thuận lợi giúp

bổ khuyết cho yếu tố bất lợi. Vì đơn vị đất đai này chiếm diện tích rất lớn (49.725,1 ha) nên làm cho diện tích thích hợp S1 chiếm tỷ lệ rất cao (77,7%) khi đánh giá theo mô hình cây quyết định.

Đánh giá theo FAO có 07 ĐVĐĐ đạt S2 với diện tích 57.432,5 ha chiếm tỷ lệ 64,7%, nhưng do ĐVĐĐ nêu trên chiếm diện tích lớn khi đánh giá theo phương pháp đề xuất nó đã chuyển qua mức S1 vì vậy đánh giá theo phương pháp này diện tích thích hợp S2 chỉ còn diện tích 16.945,8 ha, chiếm tỷ lệ 19,1%.

Đánh giá theo phương pháp toán học, kỹ thuật phân tích hồi quy cây quyết định với phần mềm DTREG cho kết quả đánh giá định lượng. Tuy nhiên, phương pháp này vì cần phải thu thập năng suất cây trồng tại thực địa nên những diện tích đất chưa có trồng cây cam thì không thể đánh giá được. Vì vậy, cần sử dụng những phương pháp đánh giá khác để thực hiện cho diện tích đất này. Đây cũng là vấn đề hạn chế của phương pháp cây quyết định với mô hình hồi quy toán học.

4. KẾT LUẬN

Nghiên cứu đã thực hiện đánh giá đất đai bằng phương pháp hồi quy toán học, thiết lập mô hình cây quyết định và so sánh với phương

pháp của FAO trên cây cam tại huyện Dầu Tiếng, tỉnh Bình Dương. Kết quả nghiên cứu cho thấy việc đánh giá này mang lại nhiều ưu điểm hơn so với phương pháp truyền thống của FAO. Cụ thể, kết quả đánh giá mức độ thích hợp của cây cam trên địa bàn huyện Dầu Tiếng tập trung vào mức S1 và mức S2. Trong khi đó, phương pháp đánh giá truyền thống theo FAO thì phân bổ bao gồm 4 mức: S1, S2, S3 và không thích hợp (N). Trong đó, phương pháp đề xuất của nghiên cứu có đánh giá sự tương hỗ giữa các yếu tố một cách định lượng qua năng suất cây trồng, mức S1 chiếm tỷ lệ 77,7% và mức S2 chiếm 19,1% tổng diện tích cây cam trên địa bàn huyện. Kết quả này khác với phương pháp đánh giá của FAO mức S1 chiếm tỷ lệ 21,7% và mức S2 chiếm 64,7% còn lại là các mức S3 và N chiếm 10,4% về diện tích canh tác cam của địa phương.

Kết quả nghiên cứu đã đề xuất được phương pháp mới để đánh giá thích hợp cây cam trên địa bàn huyện Dầu Tiếng. Tuy nhiên để có kết quả đánh giá một cách chính xác, khách quan và triển khai vào thực tiễn thì nghiên cứu cần mở rộng khu vực cũng như cần được đánh giá, kiểm chứng nhiều hơn.

TÀI LIỆU THAM KHẢO

[1]. UBND huyện Dầu Tiếng (2019). Dầu Tiếng, nhiều Nông dân thành công với mô hình trồng cây có múi. UBND huyện Dầu Tiếng, Bình Dương, Việt Nam. Truy cập từ: <http://dautieng.binhduong.gov.vn/portal/Tin-tuc/Chi-tiet/Tin-kinh-te-196-3> ngày 15/3/2024.

[2]. Andi Nurkholis, Sukaesih Sitanggang Imas,

Annisa Annisa & Sobir Sobir (2021). Spatial decision tree model for garlic land suitability evaluation. *IAES International Journal of Artificial Intelligence*. 10(3): 666.

[3]. J Bouma, RJ Wagenet, MR Hoosbeek & JL Hutson (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.

[4]. 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): 84-93.

[5]. Nguyễn Hữu Cường (2018). Tích hợp GIS và cây quyết định đánh giá thích nghi đất đai cây dừa trên địa bàn huyện Mỏ Cày Nam, tỉnh Bến Tre. *Tạp chí Khoa học ĐHQGHN*. 34(1): 11.

[6]. UBND tỉnh Bình Dương (2022). Quyết định 381/QĐ-UBND ngày 05/4/2022 của UBND tỉnh Bình Dương về việc phê duyệt kế hoạch sử dụng đất năm 2022 huyện Dầu Tiếng.

[7]. UBND huyện Dầu Tiếng (2017). Quyết định về việc quy hoạch phát triển nông lâm ngư nghiệp huyện Dầu Tiếng đến năm 2025, tầm nhìn đến năm 2030.

[8]. Ashok N Srivastava & Mehran Sahami (2009). *Text mining: Classification, clustering, and applications*. Chapman and Hall/CRC.

[9]. Võ Quốc Khánh, Nguyễn Văn Cương & Trương Thị Diệu Quân (2022). 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. *Tạp chí Khoa học và Công nghệ Lâm nghiệp*. 4: 88-95.

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

[10]. 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. Nhà xuất bản Nông nghiệp.

[11]. TRE Chidley, J Elgy & J Antoine (1993). *Computerized systems of land resources appraisal for agricultural development*. Food & Agriculture Org. Rome, Italy.