

ÁP DỤNG PHƯƠNG PHÁP K - NEAREST NEIGHBORS ĐỂ ƯỚC LƯỢNG GIÁ TRỊ LÂM PHẦN LÁ RỘNG THƯỜNG XANH DỰA VÀO ẢNH VỆ TINH ĐA PHỔ SPOT 5

Nguyễn Thị Thanh Hương

Khoa Nông Lâm Nghiệp, Trường Đại học Tây Nguyên Buôn Ma Thuột

ABSTRACT

The continuous variables of forest such as biomass or stand volume are important in forest management planning. This study used the kNN method (k-Nearest Neighbors) to estimate stand volume of natural evergreen broadleaf forest which have been impacted in different degrees. Different approaches were tested on SPOT 5 images via the calculated NDVI (Normalized Difference Vegetation Index). The estimates were carried out on different stratum and the whole forest stand. The independent field data was then used to validate the results using Root Mean Squared Error (RMSEs) as a criterion of accuracy assessment. The result show the accuracy of estimation was significantly improved with separate stratum analysis in comparison to the whole stand volume. The better results were found in case of using SPOT 5 in comparison to those with the application of NDVI. The RMSE and RMSE% were of $\pm 27,7 \text{ m}^3\text{ha}^{-1}$ and 15,6% respectively while these were $37,7 \text{ m}^3\text{ha}^{-1}$ and 20,9% respectively when using the NDVI. However, in case the stand has been heavily disturbed as the Class 1, the NDVI turns to be the most suitable choice in comparison to others with RMSE and RMSE% of $25 \text{ m}^3\text{ha}^{-1}$ and 28%, respectively while these were $25 \text{ m}^3\text{ha}^{-1}$ and 28% for RMSE and RMSE%, respectively.

Thông tin chung:

Ngày nhận: 03/03/2016

Ngày chấp nhận: 24/05/2016

Title:

Using k - Nearest Neighbors (kNN) to estimate stand volume of tropical evergreen broadleaved forest using multi-spectral satellite imagery

Từ khóa:

Ảnh vệ tinh SPOT 5, NDVI, trữ lượng, K-nearest neighbour, sai số trung phương

Keywords:

SPOT 5, NDVI, stand volume, k-Nearest Neighbors, RMSE, RMSE%

TÓM TẮT

Các biến số liên tục như trữ lượng hay sinh khối là rất cần thiết trong quản lý rừng và kế hoạch sử dụng rừng. Nghiên cứu này đã sử dụng phương pháp kNN (k-Nearest Neighbors) để ước lượng trữ lượng lâm phần rừng tự nhiên lá rộng thường xanh đã bị tác động ở các mức độ khác nhau. Nghiên cứu đã thử nghiệm với các giải pháp khác nhau, bao gồm lâm phần được ước lượng chung và ước lượng theo các khối trạng thái rừng; dữ liệu ảnh được ước lượng là ảnh chỉ số khác biệt thực vật NDVI (Normalized difference vegetation index) được tính từ ảnh SPOT 5 và ảnh đa phổ SPOT 5. Kết quả việc ước lượng bằng phương pháp kNN với dữ liệu từng khối trạng thái và ước lượng chung cho toàn lâm phần cho thấy nếu lâm phần được chia thành các khối tương đối đồng nhất thì việc ước lượng sẽ đạt kết quả tốt hơn so với việc ước lượng chung cho toàn bộ lâm phần, đặc biệt là đối với các khối trạng thái ít bị tác động. Tính chung cho toàn lâm phần sai số ước lượng bằng dữ liệu SPOT 5 đạt kết quả tốt hơn so với sử dụng dữ liệu NDVI. Trong khi sai số trung phương RMSE = $27,7 \text{ m}^3\text{ha}^{-1}$ tương ứng với sai số trung phương tương đối RMSE% = 15,6% nếu sử dụng ảnh SPOT5 để ước lượng thì sai số này của ảnh NDVI là 37,7 và 20,9%. Tuy nhiên, với lâm phần đã bị tác động mạnh (lớp 1), kết quả ước lượng trữ lượng sử dụng dữ liệu NDVI cho thấy hiệu quả hơn so với ảnh đa phổ. Sai số ước lượng bằng ảnh đa phổ là $37 \text{ m}^3\text{ha}^{-1}$ tương ứng với RMSE% là 37,3%, trong khi sai số này khi ước lượng bằng ảnh NDVI là $25 \text{ m}^3\text{ha}^{-1}$ với RMSE% là 28%.

Trích dẫn: Nguyễn Thị Thanh Hương, 2016. Áp dụng phương pháp k - Nearest Neighbors để ước lượng giá trị lâm phần lá rộng thường xanh dựa vào ảnh vệ tinh đa phổ SPOT 5. Tạp chí Khoa học Trường Đại học Cần Thơ. 43a: 59-67.

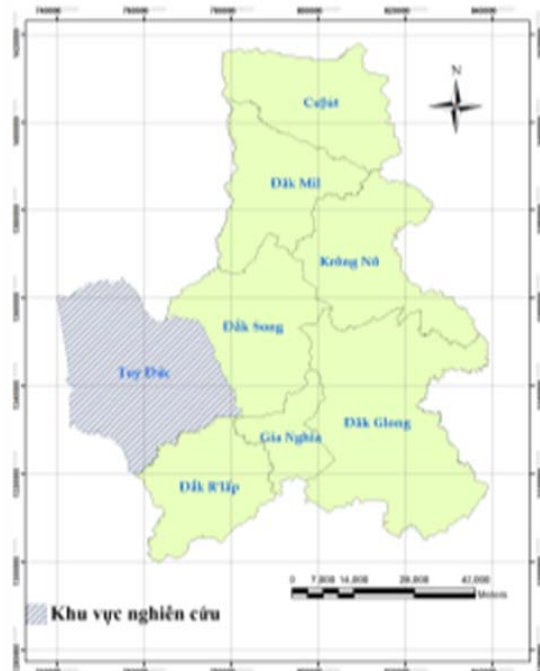
1 ĐẶT VẤN ĐỀ

Các biến số liên tục như trữ lượng hay sinh khối là rất cần thiết trong quản lý rừng và kế hoạch sử dụng rừng. Đây chính là cơ sở để định lượng được giá trị của tài nguyên rừng. Các hạn chế về thời gian cũng như tài chính không cho phép điều tra toàn bộ tổng thể lâm phần, chính vì vậy cần thiết kê lấy mẫu để ước lượng cho toàn bộ lâm phần. Trong những năm gần đây việc kết hợp điều tra thực địa với ảnh vệ tinh là mối quan tâm cao trong điều tra rừng vì đã tạo ra các khả năng sử dụng dữ liệu hiện trường chính xác cùng với ảnh vệ tinh phủ trong vùng khá rộng. Điều này cho phép phát triển các áp dụng về lâm nghiệp dựa trên phương pháp hiệu quả về chi phí và khách quan.

Nhiều nghiên cứu vẫn đang thực hiện để tìm các ứng dụng ảnh vệ tinh quang học trong ước lượng các nhân tố điều tra rừng ở cấp độ lâm phần với độ chính xác thích hợp cho xây dựng bản đồ rừng cũng như trong điều chế rừng với ảnh vệ tinh có độ phân giải cao hoặc ảnh hàng không (Holmgren and Thuresson, 1998). Tuy nhiên, cải thiện độ chính xác ước lượng có thể đạt được bằng việc phối hợp thông tin từ các nguồn khác nhau như Cohen and Spies (1992) hay bằng việc bổ sung thêm dữ liệu hỗ trợ như thông tin về chiều cao cây như nghiên cứu của Nilsson (1997) hoặc dữ liệu có sẵn của lâm phần như Tomppo *et al.* (1999), Holmgren *et al.* (2000). Việc phối hợp các nguồn dữ liệu không gian và phi không gian khác nhau trong điều tra và giám sát tài nguyên rừng như dữ liệu ảnh vệ tinh, dữ liệu điều tra mặt đất hay các nguồn dữ liệu số khác như các loại bản đồ được xem như là một nguồn đa dữ liệu. Việc phối hợp nguồn đa dữ liệu trong ước lượng tài nguyên đã chứng minh được tính hiệu quả của nó. Một số giải pháp toán học đã được sử dụng để ước lượng giá trị của lâm phần trong đó có phương pháp k-Nearest Neighbors (kNN). Phương pháp này đã được sử dụng lần đầu tiên trong điều tra rừng toàn quốc của Phần Lan (Tomppo, 1990). Phương pháp này đã tạo nên một công cụ ước lượng nhân tố lâm phần rừng một cách hiệu quả trong điều tra rừng bằng đa dữ liệu (Holmström &Fransson, 2003). Cơ sở của phương pháp này được mô tả bởi Tomppo (1990), McRoberts *et al.* (2007), Franco-Lopez *et al.* (2001). Việc ước lượng nhân tố điều tra rừng có thể thực hiện chung cho toàn bộ (Nguyen Thi Thanh Huong, 2011) hoặc có thể ước lượng trên các khối rừng tương đối đồng nhất. Trong tiếp cận này, ảnh vệ tinh được phân loại theo các lớp đồng nhất và các lớp này sau đó được sử dụng để ước

lượng các nhân tố thuộc tính. Các ô mẫu được bố trí trên các lớp đã được phân loại. Theo McRobert *et al.* (2002) việc bố trí ô mẫu trên từng khối trạng thái sẽ làm giảm đáng kể sai số ước lượng, thậm chí nếu ô mẫu độc lập với các lớp đã phân loại thì việc ước lượng theo từng lớp cũng làm tăng độ chính xác ước lượng.

Mục tiêu của nghiên cứu này là đánh giá độ chính xác ước lượng trữ lượng lâm phần sử dụng phương pháp phi tham số kNN trên các lâm phần được phân khối và không phân khối dựa vào ảnh vệ tinh SPOT 5.



Hình 1: Bản đồ khu vực nghiên cứu

Nghiên cứu được thực hiện tại các khu rừng lá rộng thường xanh ở huyện Tuy Đức, tỉnh Đắk Nông. Khu vực nghiên cứu nằm ở vĩ độ Bắc từ 11059' đến 12016' và kinh độ Đông từ 107013' đến 107028' với diện tích khoảng 500 km² (20x25km). Rừng lá rộng thường xanh nhiệt đới chiếm ưu thế ở khu vực này, tuy nhiên đã qua tác động của con người ở các cấp độ khác nhau, hình thành các trạng thái trữ lượng gỗ khác biệt.

2 DỮ LIỆU, PHẦN MỀM VÀ PHƯƠNG PHÁP

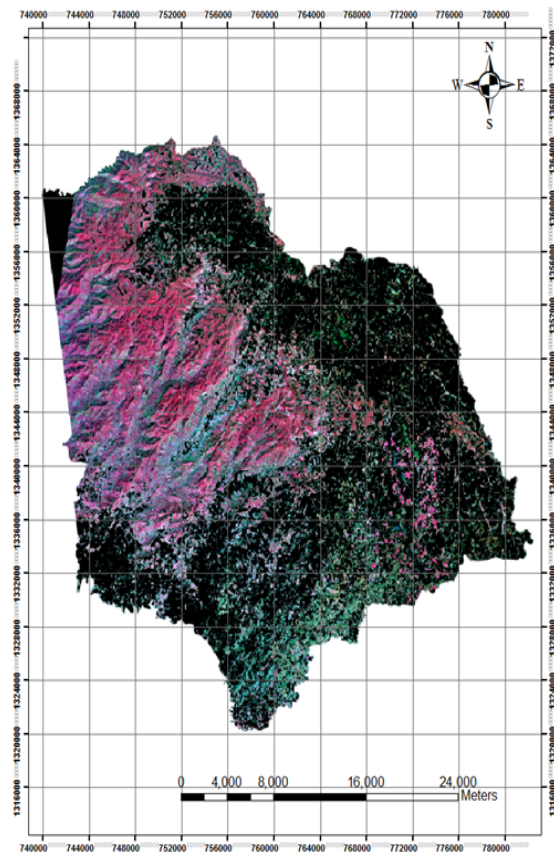
2.1 Dữ liệu và phần mềm

Các nguồn dữ liệu khác nhau bao gồm ảnh vệ tinh, dữ liệu số về rừng, bản đồ nền và các ô mẫu

được điều tra trên thực địa được sử dụng trong nghiên cứu này.

Ảnh vệ tinh quang học SPOT 5 được thu vào tháng 2 năm 2013. Ảnh vệ tinh đã được hiệu chỉnh dựa vào các điểm khống chế trên thực địa được thu thập bằng GPS và mô hình số độ cao (DEM). Mô hình này được tạo từ các đường đồng mức ở dạng GIS shapefile. Để phù hợp và liên kết được với các loại dữ liệu hiện có tại địa phương, ảnh SPOT 5 đã được tham chiếu hệ tọa độ UTM ở cột 48 (UTM 48N). Phương pháp tái chia mẫu “nearest neighbor” được áp dụng trong quá trình hiệu chỉnh hình học với độ phân giải không gian là 10m x 10m để bảo toàn giá trị của các pixel. Ngoài ra, ảnh vệ tinh sau hiệu chỉnh hình học cũng được hiệu chỉnh địa hình để giảm thiểu ảnh hưởng bề mặt địa hình đến giá trị độ sáng của ảnh.

Dữ liệu số liên quan được thu thập trong nghiên cứu này bao gồm bản đồ địa hình, sông suối, đường giao thông, bản đồ sử dụng đất, bản đồ hiện trạng rừng. Các dữ liệu này được sử dụng để tham khảo trong quá trình phân tích.



Tổng số có 111 ô mẫu được điều tra trên hiện trường, trong đó có 7 ô kế thừa từ nghiên cứu

trước. Các ô mẫu kế thừa đã được kiểm tra bằng chỉ số thực vật NDVI của năm 2009 và năm 2013 để đảm bảo tính ít thay đổi của hiện trạng rừng. Nghiên cứu đã sử dụng 84 ô mẫu để thực hiện các ước lượng và sử dụng 27 ô mẫu còn lại để đánh giá kết quả ước lượng. Các ô mẫu đã được điều tra trên thực địa với diện tích mỗi ô là 900 m² (30m x 30m) tương ứng với 1 cụm 3x3 pixel của ảnh vệ tinh SPOT 5. Các nhân tố lâm phần như đường kính ngang ngực (D_{1,3}), chiều cao cây (H), mật độ (N) được đo đếm toàn bộ trong ô mẫu. Vị trí trung tâm của ô được thu thập bằng GPS CSx 62. Mô hình tương quan thể tích với đường kính và chiều cao sử dụng từ kết quả nghiên cứu của Nguyen Thi Thanh Huong (2011). Mô hình có dạng:

$$\ln(V) = -10,0094 + 1,06589 \cdot \ln(D_{1,3}) + 1,93331 \cdot \ln(H) \quad (1)$$

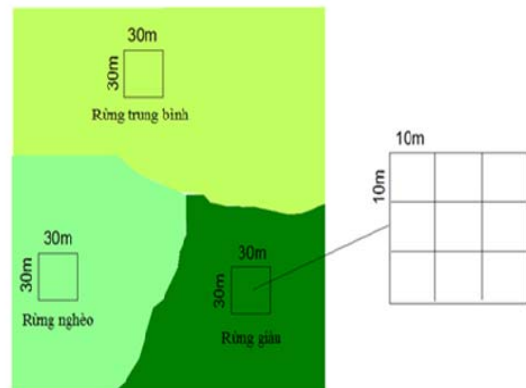
Với R² = 0,982; P < 0.

Tương ứng với mỗi ô mẫu với kích cỡ 30m x 30m, giá trị số trung bình (digital number-DN) của 9 pixel (3x3) của các band SPOT 5 cũng được tính toán.

2.2 Phương pháp

2.2.1 Phân chia rừng thành các khối đồng nhất

Nghiên cứu đã sử dụng kết hợp 2 phương pháp phân loại có kiểm định và không kiểm định để phân chia rừng thành các khối đồng nhất đồng thời loại bỏ những diện tích không phải rừng xen lẫn trong khu vực nghiên cứu. Ngoài ra, nghiên cứu còn sử dụng các dữ liệu có sẵn như bản đồ sử dụng đất, bản đồ hiện trạng rừng để phục vụ cho việc phân khối rừng (stratification).



Hình 3: Minh họa thiết kế ô mẫu trên thực địa

Trong các chỉ tiêu đo đếm nhận thấy trữ lượng là nhân tố khá tổng hợp và thường có mối quan hệ với ảnh vệ tinh cao hơn các nhân tố khác như DBH

hay N. Do vậy, nghiên cứu đã sử dụng nhân tố này để làm các vùng mẫu (training sites) để phân loại. Kết quả phân loại cần được kiểm tra trước khi thành lập bản đồ thành quả cuối cùng. Sử dụng các mẫu độc lập (không trùng với mẫu dùng làm các vùng mẫu) để đánh giá kết quả phân loại. Kết quả đánh giá dựa trên các chỉ tiêu độ chính xác chung (Overall accuracy), độ chính xác người sản xuất (Producer accuracy) và độ chính xác người sử dụng (User accuracy).

Độ chính xác người sử dụng (User accuracy) là khả năng mà một loại thảm phủ nào đó trên bản đồ trùng khớp với loại thảm phủ tương ứng trên thực địa. Ngược lại, độ chính xác người sản xuất (Producer accuracy) là khả năng mà một loại thảm phủ nào trên thực địa trùng khớp với loại tương ứng trên bản đồ.

- Độ chính xác toàn bộ (Overall Accuracy):

$$T = \frac{\sum_{i=1}^k O_{ij}}{n} \times 100 \quad (2)$$

- Độ chính xác người sản xuất (Producer accuracy):

$$P_{i+} = \frac{S_{i+} - O_{ii}}{S_{i+}} \times 100 \quad (3)$$

- Độ chính xác người sử dụng (Use accuracy):

$$U_{ij} = \frac{S_{+j} - O_{ij}}{S_{+j}} \times 100 \quad (4)$$

Trong đó:

+ O_{ij} là giá trị thể hiện sự phù hợp ở hàng i và cột j ;

+ S_{i+} (loại thực) là tổng giá trị theo hàng ($i=1,2,\dots,K$): $S_{i+} = \sum_{j=1}^k O_{ij}$

+ S_{+j} (loại giải đoán) là tổng giá trị theo cột ($j=1,2,\dots,K$): $S_{+j} = \sum_{i=1}^k O_{ij}$

+ n là tổng số pixel trong bộ dữ liệu.

Ma trận sai số được tính toán trong chức năng thống kê của phần mềm ENVI.

Ngoài ra, việc đánh giá mức độ quan hệ giữa thực tế và kết quả phân loại còn dựa vào chỉ số Kappa với các mức độ phân biệt được căn cứ vào Bảng 1:

Bảng 1: Giá trị mức độ chặt chẽ theo chỉ số Kappa

Giá trị K	Mức độ phân biệt
< 0,00	Thấp
0,00 – 0,20	Nhẹ
0,21 – 0,40	Vừa
0,41 – 0,60	Tương đối chặt
0,61 – 0,80	Chặt
0,81 – 1,00	Rất chặt

(Nguồn: Navulur, 2007)

Phần mềm được sử dụng chủ yếu trong quá trình phân tích là phần mềm phi tham số kNN được phát triển bởi Stümer (2004). Ngoài ra, một số phần mềm cũng được sử dụng để hỗ trợ trong quá trình thành lập bản đồ và xử lý ảnh vệ tinh như ArcGIS và ENVI.

Dựa trên ảnh đã được phân khối rừng, tiến hành rút mẫu để thu thập số liệu trên thực địa. Số lượng ô mẫu tùy thuộc vào mỗi trạng thái rừng và đảm bảo tính đại diện cho từng trạng thái. Số ô mẫu này sau đó được tính dung lượng cần thiết trong từng khối trạng thái để đảm bảo tính khách quan của việc thu thập mẫu.

2.2.2 Phương pháp ước lượng trữ lượng bằng thuật toán kNN

Trong phương pháp này nhân tố điều tra rừng Y_i (trữ lượng) trong từng pixel cần được ước lượng được tính theo trọng số trung bình của ô mẫu k gần nhất. Khoảng cách không gian đối tượng (feature space) d được tính theo khoảng cách hình học Euclidean của không gian phổ (spectral space) của ảnh SPOT 5 từ ô mẫu đến các pixel cần ước lượng. Khoảng cách không gian phổ d_{p_i, p_j} được tính trong không gian đối tượng từ pixel p cần ước lượng đến pixel p_i được tham chiếu trên thực địa (dựa vào ô mẫu đã điều tra) được tính theo công thức sau:

$$d_{p(p_i)} = \left[\sum_{j=1}^n (x_{p,j} - x_{p_i,j})^2 \right]^{\frac{1}{2}} \quad (5)$$

Trong đó $x_{p,j}$ = giá trị độ sáng của một pixel trong một band nào đó của ảnh j , n = số pixel trong không gian phổ.

Nghiên cứu này đã sử dụng phần mềm kNN của Stümer (2004) (k Nearest Neighbors) để ước lượng trữ lượng. Ước lượng được thử nghiệm cho

toàn bộ lâm phần và theo từng khối trạng thái đồng nhất. Hai loại dữ liệu đầu vào là cần thiết trong quá trình này: i) file ảnh (image file) và ii) file dữ liệu thực địa (field sample file). Dữ liệu ảnh được sử dụng bao gồm ảnh SPOT 5 và ảnh NDVI. Các ảnh này được chuyển đổi thành dạng file ASCII. Quá trình ước lượng được thực hiện trên từng khối trạng thái riêng rẽ và ước lượng chung cho toàn bộ lâm phần. Để so sánh độ chính xác của các thử nghiệm, dữ liệu độc lập (dữ liệu không tham gia ước lượng) được dùng để đánh giá bằng tiêu chuẩn sai số trung phương (RMSE):

$$RMSE = \sqrt{\frac{\sum (X - X')^2}{n}} \tag{6}$$

Trong đó: RMSE là sai số trung phương; X là giá trị trữ lượng được nội suy; X' là giá trị trữ lượng được thu thập trên thực địa; n là ô mẫu dùng để đánh giá.

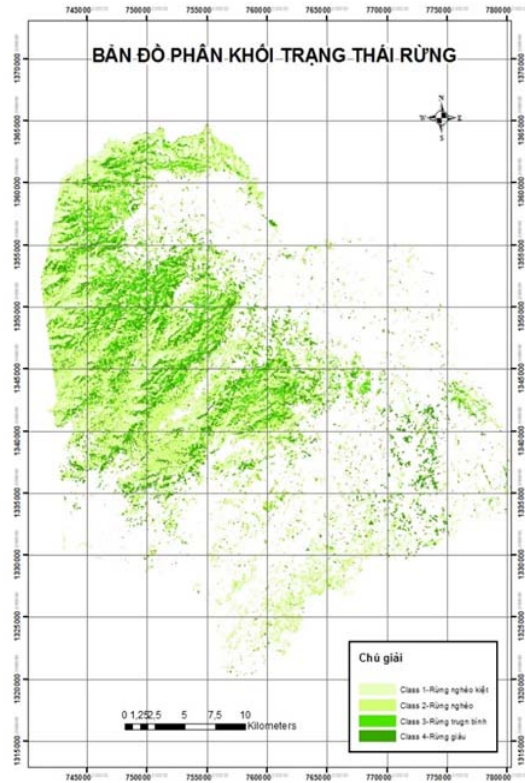
Dựa trên các giá trị Yi đã được ước lượng từ kNN, xây dựng bản đồ Yi trong phần mềm ArcGIS.

3 KẾT QUẢ VÀ THẢO LUẬN

3.1 Phân chia rừng thành các khối đồng nhất

Dựa trên thông tin từ ô mẫu điều tra thực địa, ô mẫu đã rơi vào các lớp phân loại phi giám định cùng với quan sát thực tế, các vùng mẫu được tạo để phân loại có giám sát và đánh giá kết quả phân loại. Trong khu vực nghiên cứu không chỉ có thảm phủ là rừng tự nhiên mà còn có các đối tượng khác như rừng trồng, đất nông nghiệp... Tuy nhiên, đối tượng quan tâm của nghiên cứu này là rừng tự nhiên, do vậy các đối tượng không phải là rừng

được loại ra trong quá trình phân loại bằng kỹ thuật tạo mặt nạ rừng. Ngoài các vùng mẫu được sử dụng để phân loại, các vùng mẫu độc lập tức là vùng mẫu không tham gia phân loại được sử dụng để đánh giá kết quả phân loại. Bản đồ phân loại được thể hiện ở Hình 3 và kết quả đánh giá được trình bày trong Bảng 2.



Hình 4: Bản đồ phân khối trạng thái rừng

Bảng 2: Bảng ma trận đánh giá kết quả phân loại rừng thành 4 lớp (lớp)

	Lớp 1	Lớp 2	Lớp 3	Lớp 4	Tổng hàng	Producer accuracy
Lớp 1	42	19	0	0	61	68,85
Lớp 2	12	138	23	2	175	78,86
Lớp 3	0	2	67	12	81	82,72
Lớp 4	0	0	0	172	172	100
Tổng cột	54	159	90	186	Độ chính xác toàn bộ = 85,69%	
Uer accuracy	77,78	86,79	74,44	92,47	Hệ số Kappa = ,0.7971	

Từ kết quả trên có thể dễ dàng nhận thấy đối với rừng bị tác động như khu vực nghiên cứu, trạng thái rừng được phân loại thành 4 lớp có độ chính xác toàn bộ khoảng 86% với hệ số Kappa 0,79 thể hiện mức độ chặt chẽ khá cao giữa kết quả phân loại và thực tế.

Từ dữ liệu thực địa và kết quả phân loại, sự khác biệt trữ lượng như giá trị trung bình, giá trị cực đại, cực tiểu cùng với đặc trưng và biến động của trữ lượng trong từng khối trạng thái được mô tả như trong Bảng 3.

Bảng 3: Đặc trưng và biến động trữ lượng của các lớp trạng thái

Mm^3ha^{-1}	Lớp 1	Lớp 2	Lớp 3	Lớp 4
Trung bình	92,22	152,66	197,85	286,80
Sai số số trung bình	4,51	5,24	5,37	13,07
Sai tiêu chuẩn - Standard Deviation	22,53	27,71	33,51	56,96
Phương sai - Sample Variance	507,73	767,57	1122,99	3244,92
Độ nhọn chuẩn - Kurtosis standard	0,517	-0,969	0,847	1,236
Độ lệch chuẩn - Skewness standard	-0,984	-0,087	-0,964	-0,398
Minimum – giá trị nhỏ nhất	45,76	110,63	142,77	212,97
Maximum – giá trị lớn nhất	135,34	204,76	279,02	412,70
Số ô mẫu	25,00	28,00	39,00	19,00
Biến động ở P = 95% - Confidence Level (95,0%)	9,30	10,74	10,86	27,46

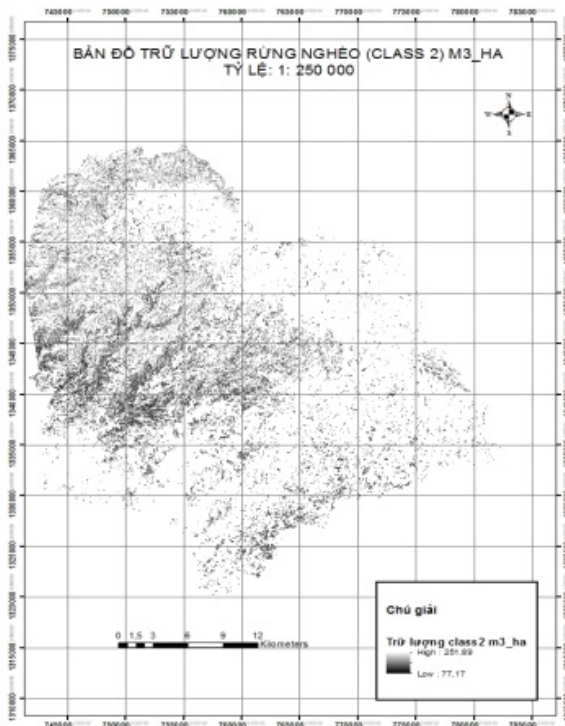
Kết quả phân tích đặc trưng trữ lượng 4 lớp ở bảng trên cho thấy:

Với độ lệch và độ nhọn chuẩn đều < 2 ở cả 4 lớp cho thấy số lượng ô mẫu thu được ở 4 lớp là đại diện và đạt phân bố chuẩn hay nói cách khác dữ liệu ô mẫu của 4 lớp đã đại diện cho các trạng thái rừng nghiên cứu.

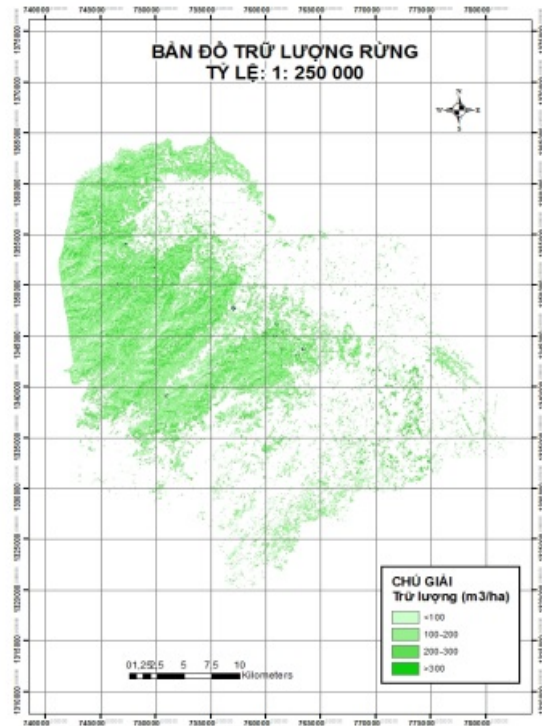
3.2 Kết quả ước lượng trữ lượng bằng phương pháp kNN

kNN là phương pháp phi tham số. Ước lượng

này chia làm 2 tiếp cận: i) ước lượng cho toàn lâm phần và ii) ước lượng theo từng khối trạng thái như đã phân tích. Dữ liệu ảnh dùng ước lượng là ảnh SPOT và ảnh NDVI. Dựa vào kết quả phân loại ở trên, ảnh SPOT 5 và NDVI được tách ra (subset) theo 4 trạng thái. Dựa trên các ảnh đã được subset tiến hành ước lượng trữ lượng cho từng khối và ước lượng chung cho toàn bộ khu vực. Các ảnh được ước lượng theo từng khối riêng rẽ sau đó được kết hợp lại để tạo bản đồ trữ lượng chung cho toàn bộ khu vực như minh họa trong Hình 5.



(a)



(b)

Hình 5: Bản đồ ước lượng trữ lượng rừng: a) rừng nghèo b) toàn bộ khu vực

3.3 Đánh giá độ chính xác của kết quả ước lượng

Kết quả ước lượng được so sánh với bộ dữ liệu độc lập bằng tiêu chuẩn sai số trung phương. Kết quả được tổng hợp trong Bảng 4.

Nhìn vào bảng kết quả cho thấy việc ước lượng rất khả quan khi sai số ước lượng thấp, đặc biệt khi lâm phần được phân chia thành các khối tương đối đồng nhất. Đối với ảnh SPOT, khi ước lượng chung cho toàn lâm phần, sai số trung phương là 46,86 m³ha⁻¹ tương ứng với sai số trung phương tương đối 25,21%. Tuy nhiên, kết quả được cải thiện rõ rệt khi ước lượng riêng cho từng khối trạng thái. Khi phân theo từng khối trạng thái kết quả ước lượng thấp nhất rơi vào lớp 1 tương ứng với rừng nghèo kiệt. Kết quả ước lượng với độ chính xác thấp nhất với RMSE gần 33 m³ha⁻¹ với sai số trung phương tương đối RMSE% đạt 37,27%.

Chênh lệch giữa thực tế và ước lượng ở các trạng thái khác đều rất khả quan với RMSE% <15%. Chênh lệch trữ lượng ở lớp 2 là 18,09 m³ha⁻¹, lớp 3 là 26,87 m³ha⁻¹ và lớp 4 gần 31 m³ha⁻¹ tương ứng với RMSE% lần lượt là 11,52%, 13,35% và 10,68%. Tính chung các lớp sai số ước lượng đạt gần 28 m³ha⁻¹ với RMSE% là 15,64%. Kết quả này cho thấy phương pháp phi tham số kNN khá hiệu quả để ứng dụng trong việc ước lượng trữ lượng lâm phần đặc biệt khi phân chia rừng thành các khối trạng thái tương đối đồng nhất. Ngoài ra, kết quả cũng chỉ ra với đối tượng rừng tương đối ổn định việc ước lượng trữ lượng mang lại kết quả chính xác hơn so với các đối tượng rừng đã bị tác động mạnh, thể hiện ở các lớp 2, 3, và 4 là các lớp rừng bị tác động ở mức vừa và mức thấp sai số ước lượng rất thấp đạt <15% trong khi đó rừng bị tác động mạnh như ở lớp 1 sai số ước lượng trên 25%.

Bảng 4: Đánh giá kết quả ước lượng bằng phương pháp kNN với dữ liệu SPOT và NDVI

	SPOT		NDVI	
	RMSE (m ³)	RMSE %	RMSE (m ³)	RMSE %
Ước lượng không phân lớp	46,86	25,21	53,30	28,57
Lớp 1	32,72	37,27	24,05	28,10
Lớp 2	18,09	11,52	21,05	14,06
Lớp 3	26,87	13,35	35,10	17,22
Lớp 4	30,87	10,68	64,24	20,99
Tính chung 4 lớp	27,70	15,64	37,72	20,95

Kết quả cũng tương tự khi ước lượng cho dữ liệu NDVI. Khi chia lâm phần thành các khối trạng thái, kết quả ước lượng đều được cải thiện đáng kể. RMSE% tính chung cho 4 lớp tương đương 21% trong khi sai số này ở kết quả không phân khối là gần 29%. Trong từng trạng thái độ chính xác của kết quả ước lượng thấp nhất vẫn là lớp 1, tuy vậy vẫn có sự sai khác nhỏ giữa các lớp còn lại so với ước lượng ảnh SPOT. Hầu hết sai số của ước lượng đối với ảnh SPOT đều cho sai số thấp hơn so với ảnh NDVI. Điều này cho thấy kết quả ước lượng khả quan hơn khi sử dụng nhiều band phổ để tính toán. Riêng lớp 1 cho kết quả sai số cao hơn khi ước lượng bằng ảnh SPOT 5. Điều này là do khi lâm phần bị tác động mạnh sự khác biệt thực vật trở nên quan trọng, vì vậy sử dụng ảnh NDVI để ước lượng đã cho kết quả tốt hơn so với sử dụng ảnh gốc SPOT.

Tóm lại, khi chia lâm phần thành các khối đồng nhất và các ước lượng được thực hiện riêng rẽ trên từng khối trạng thái cho kết quả chính xác cao hơn so với việc ước lượng chung cho toàn bộ lâm phần. Điều này có thể là do sự không đồng nhất về đặc điểm của lâm phần, do vậy khi áp dụng phương

pháp kNN ước lượng chung cho toàn bộ lâm phần đã cho kết quả thấp hơn so với việc ước lượng riêng cho từng khối trạng thái. Những lâm phần bị tác động ở mức vừa và thấp, việc ước lượng cũng thể hiện tốt hơn so với lâm phần bị tác động mạnh. Điều này là do lâm phần bị tác động mạnh dẫn đến cấu trúc rừng bị phá vỡ, rừng bị vỡ tán, do vậy việc ước lượng trở nên kém hiệu quả hơn so với các lâm phần ổn định.

4 KẾT LUẬN VÀ ĐỀ XUẤT

4.1 Kết luận

Mặc dù đã có một số nghiên cứu ứng dụng dữ liệu viễn thám vào việc quản lý nguồn tài nguyên rừng, tuy vậy hầu hết các nghiên cứu thường tập trung vào việc xác định số lượng như xây dựng các bản đồ hiện trạng, thăm phủ hơn là ước lượng chất lượng rừng như xây dựng bản đồ trữ lượng rừng. Vì vậy, tìm kiếm các giải pháp để lượng hóa được tài nguyên rừng là rất cần thiết đặc biệt trong bối cảnh hiện nay khi thế giới rất quan tâm đến vấn đề biến đổi khí hậu, và chương trình REDD (Reducing Emissions from Deforestation and Forest Degradation - Giảm phát thải từ việc suy thoái và mất rừng) hay chương trình PES (Payment

for Environmental Services - Chi trả dịch vụ môi trường) là một trong các giải pháp nhằm hướng đến việc thúc đẩy để quản lý rừng bền vững. Tài nguyên rừng giảm thiểu ảnh hưởng đến biến đổi khí hậu. Do vậy, ngoài việc xác định thâm phủ rừng hay trạng thái thì việc đề cập đến chất lượng rừng để có các giải pháp quản lý thích hợp để tác động hữu hiệu hơn trong quản lý cũng như lập kế hoạch sử dụng rừng.

Phương pháp phi tham số kNN gần như chưa được ứng dụng để ước lượng các đặc điểm tài nguyên rừng ở Việt Nam, tuy nhiên kết quả nghiên cứu này cho thấy được tiềm năng của nó trong ước lượng trữ lượng rừng để giảm chi phí, thời gian cũng như nhân vật lực. Chất lượng ước lượng đạt khá cao với độ chính xác gần bằng hoặc >80%. Độ chính xác ước lượng khi dùng ảnh chỉ số thực vật NDVI gần 80% trong khi sử dụng ảnh SPOT độ chính xác đạt gần 85%. Khi làm phần được phân chia thành các khối tương đối đồng nhất, kết quả đã có sự cải thiện rõ rệt so với việc ước lượng chung cho lâm phần, cụ thể nếu ước lượng chung cho lâm phần sai số ước lượng >25% và 28% tương ứng với ảnh SPOT và ảnh NDVI, trong khi đó sai số tính chung cho 4 lớp đã được phân khối, sai số này đạt khoảng 15% đối với ảnh SPOT và 21% đối với ảnh NDVI. Nếu xét riêng cho từng khối trạng thái, ngoài trừ lâm phần đã bị tác động mạnh, sai số ước lượng khoảng 25% (đối với ảnh SPOT) và 28% (đối với ảnh NDVI), các lớp còn lại hầu hết đều đạt độ chính xác >80%, thậm chí có lớp đạt độ chính xác gần 90%. Kết quả này rất khả quan để quan tâm ứng dụng trong điều tra chất lượng rừng.

4.2 Đề xuất

Nghiên cứu này được thực hiện trong khu vực rừng đã bị tác động ở Tây Nguyên với kiểu rừng là rừng lá rộng thường xanh. Mặc dù độ chính xác của kết quả cho thấy đã cải thiện hơn khá nhiều so với các nghiên cứu trước tuy nhiên cũng cần có thêm dữ liệu ở các khu vực khác có cùng điều kiện để kiểm định tính khả thi của nó khi áp dụng, đồng thời cần có những thử nghiệm đối với các kiểu rừng khác như rừng khộp hoặc rừng hỗn giao là các kiểu rừng khá phổ biến ở Tây Nguyên để có thể phát triển và hệ thống hóa được phương pháp; ngoài ra, cũng cần thử nghiệm và so sánh các phương pháp khác nhau như phương pháp hồi quy, địa thống kê để tìm ra giải pháp tối ưu nhất có thể nâng cao tính hiệu quả khi áp dụng; đồng thời, cần thử nghiệm với các loại ảnh vệ tinh khác có độ

phân giải không gian cao hơn, hoặc có nhiều band phổ hơn cũng là hướng nghiên cứu cần quan tâm.

LỜI CẢM TẠ

Nghiên cứu này là một phần trong kết quả đề tài Nghiên cứu cấp Bộ - Bộ Giáo dục và Đào tạo “Xây dựng phương pháp sử dụng đa dữ liệu trong điều tra và giám sát rừng tự nhiên tại Tây Nguyên”, được thực hiện từ năm 2013-2015. Tác giả xin chân thành cảm ơn sự hỗ trợ về tài chính và sự quan tâm của Bộ Giáo dục và Đào tạo, sự quan tâm và tạo điều kiện của trường Đại học Tây Nguyên. Đây chính là yếu tố quan trọng góp phần tạo nên sự thành công của nghiên cứu.

TÀI LIỆU THAM KHẢO

- Cohen, W.B., Spies, T.A., 1992. Estimating structural attributes of Douglas-Fir/Western Hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment* 41: 1-17.
- Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*. 77: 251-274
- Holmgren, J., Joyce, S., Nilsson, M., and Olsson, H. 2000. Estimating stem volume and basal area in forest compartments by combining satellite image data with field data. *Scandinavian Journal of Forest Research*. 15: 103-111.
- Holmgren, P. and Thuresson, T. 1998. Satellite remote sensing for forestry planning – a review. *Scandinavian Journal of Forest Research*. 13: 90-110.
- Holmstrom, H., Fransson, J.E.S., 2003. Combining remotely sensed optical and radar data in kNN estimation of forest parameters. *Forest Science*. 49(3): 409 – 418.
- McRoberts, R. E., Nelson, M. D., & Wendt, D. G. (2002). Stratified estimation of forest area using satellite imagery, inventory data, and the k-nearest neighbours technique. *Remote Sensing of Environment*. 82: 457-468.
- McRoberts, R. E., Tomppo, E. O., Finley, A. O., & Heikkinen, J. (2007). Estimating areal means and variances of forest attributes using the k-nearest neighbours technique and satellite imagery. *Remote Sensing of Environment*. 111: 466-480.

- Navulur, K. 2007. Multispectral image analysis using object-oriented paradigm, CRC Press. 206p.
- Nguyen Thi Thanh Huong, 2011. Forestry Remote Sensing: Using multidata sources for unventory of Natural broad leaved ever-green forests in the Central Highlands of Vietnam. Lambert Academic Publishing. Germany. 143p.
- Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne Lidar. Acta Univ. Agricult. Suecia. Silvestria 17. Swedish University of Agricultural Sciences. Department of Forest Resource Management and Geomatics, Umeå, Sweden.
- Stümer, W., 2004. Kombination von terrestrischen Aufnahmen und Fernerkundungsdaten mit Hilfe der kNN methode zur Klassifizierung und Kartierung von Wäldern. Dissertation, Fakultät für Forst- Geo-und Hydrowissenschaften, Technischen Universität Dresden, 153p. (Combination of terrestrial inventory and remote sensing data using the kNN method for lópifying and mapping of forests).
- Tomppo, E., Goulding, C., and Katila, M. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scandinavian Journal of Forest Research. 14: 182–192.
- Tomppo E (1990) Satellite image-based National Forest Inventory of Finland. Photogramm J Finland. 12:115–120.