ỨNG DỤNG PHÂN CỤM MỜ ĐÁNH GIÁ BIẾN ĐỘNG LỚP PHỦ KHU VỰC KHAI THÁC KHOÁNG SẢN HUYỆN QUỲ HỢP, TỈNH NGHỆ AN TỪ ẢNH LANDSAT

ĐẬU THANH BÌNH - *Sở Tài nguyên và Môi trường Nghệ An*
TRỊNH LÊ HÙNG - *Học viện Kỹ thuật Quân sự*

Do đặc điểm địa hình, khí hậu, Nghệ An là một tỉnh có nguồn tài nguyên khoáng sản dồi dào và phong phú so với các địa phương khác ở Việt Nam, trong đó có nhiều mỏ kim loại và đá quý, đá xây dựng. Tài nguyên khoáng sản là một trong những động lực chính giúp đẩy nhanh quá trình công nghiệp hóa-hiện đại hóa ở Nghệ An.

Trong nhiều năm qua, việc phát triển quá nóng ngành công nghiệp khai thác mỏ cũng như sử dụng công nghệ thủ công, lạc hậu đã ảnh hưởng nghiêm trọng tới môi trường sống, trong đó lớp phủ mặt đất là đối tượng dễ bị tổn thương nhất. Việc nghiên cứu, đánh giá biến động lớp phủ là một vấn đề có tính thực tiễn, giúp các nhà quản lý đưa ra các biện pháp nhằm giảm thiểu ảnh hưởng của quá trình khai khoáng đến cảnh quan và môi trường, cũng như phục vụ phát triển bền vững nền công nghiệp khai khoáng.

Phương pháp viễn thám với những ưu điểm nổi bật so với các phương pháp nghiên cứu truyền thống như diện tích phủ trùm rộng, thời gian cập nhật ngắn, dải phổ đa dạng,... đã được sử dụng hiệu quả trong giám sát biến động lớp phủ mặt đất [2]÷[6]. Ở Việt Nam cho đến nay cũng đã có một số nghiên cứu ứng dụng công nghệ viễn thám trong đánh giá biến động một số thành phần môi trường trong quá trình khai thác than ở Quảng Ninh [1] hay khai thác bauxit ở khu vực Tây Nguyên. Trong các nghiên cứu này chủ yếu sử dụng các phương pháp phân loại truyền thống như phân loại thống kê, phân loại hướng đối tượng để đánh giá hiện trạng và biến động lớp phủ mặt đất. Bài báo này trình bày một cách tiếp cận khác trong phân loại lớp phủ từ ảnh vệ tinh Landsat trên cơ sở phân cụm mờ (Fuzzy logic).

1. Phương pháp nghiên cứu

1.1. Cơ sở lý thuyết

Phân cụm là một công cụ toán học dùng để phát hiện cấu trúc hoặc các mẫu nào đó trong tập dữ liệu, theo đó các đối tượng bên trong cụm dữ liệu thể hiện bậc tương đồng nhất định [8], [9]. Thuật toán phân cụm có nhiều dạng khác nhau từ phân cụm rõ đơn thuần như k-Means và phát triển đến thuật toán phân cụm mờ loại một Fuzzy c-Means (FCM) bởi James C.Bezdek (1984).

Bài toán phân cụm FCM là bài toán tìm các độ thuộc u_{ij} nhằm tối thiểu hóa hàm mục tiêu:

$$J_m(U,v)=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m(d_{ik})^2, 1\le m\le\infty \quad (1)$$

Trong đó: dij= $||x_k-v_i||$ - Khoảng cách theo thước đo Euclide giữa mẫu dữ liệu x_k với trọng tâm cụm thứ i; $u_{ik}\in[0,1]$ - Bậc hay độ thuộc của dữ liệu mẫu x_k với cụm thứ i; $V=[v_{ij}]=[v_1,...,v_c]$ là ma trận biểu diễn các giá trị tâm của cụm.

Thuật toán FCM cung cấp một quá trình lặp qua lại giữa phương trình hàm thuộc:

$$u_{ik}=1/\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}, 1\le i\le c, 1\le k\le n \quad (2)$$

Và trọng tâm các cụm:

$$v_i=\sum_{k=1}^{n}(u_{ik})^m x_k / \sum_{k=1}^{n}(u_{ik})^m, 1\le i\le c \quad (3)$$

Một phân cụm, nghĩa là hàm mục tiêu (1) đạt giá trị tối thiểu, mà chủ yếu dựa trên đó độ tương tự giữa x_k và trọng tâm cụm v_i, điều này tương đương với hai điều kiện (2) và (3) phải thỏa mãn các ràng buộc. Sau mỗi vòng lặp, thuật toán tính toán và cập nhật các phần tử u trong ma trận phân hoạch U. Phép lặp sẽ dừng khi $\max\left\{\left\|u_{ij}^{(t+1)}-u_{ij}^{(t)}\right\|\right\}\le e$ trong đó ε là chuẩn kết thúc nằm trong khoảng [0,1] trong khi k là các bước lặp [8], [9].

1.2. Ứng dụng phân cụm mờ trong phân loại ảnh vệ tinh

Để áp dụng phương pháp phân cụm mờ FCM vào phân loại ảnh vệ tinh, cần phải chuyển các dữ liệu ảnh vệ tinh thành dữ liệu đầu vào theo cấu trúc

của thuật toán [7]. Giả sử ảnh vệ tinh Landsat có b kênh ảnh, trong nghiên cứu này nhóm tác giả bổ sung thêm dữ liệu ảnh chỉ số NDVI như là một kênh, như vậy dữ liệu dùng để phân loại gồm có b+1 kênh ảnh. Bài toán cần phân tách tập n vectơ đối tượng dữ liệu $X=\{x_1,x_2,...,x_n\}\in R^{b+1}$ vào C nhóm dựa trên tính toán tối thiểu hóa hàm mục tiêu, mỗi đối tượng x_k là một vectơ có b+1 thành phần. Trọng tâm ban đầu của các cụm được tính trung bình trên các đối tượng của mỗi lớp sau khi chúng được gán tạm thời về các lớp cần phân loại. Chi tiết các bước thực hiện như sau:

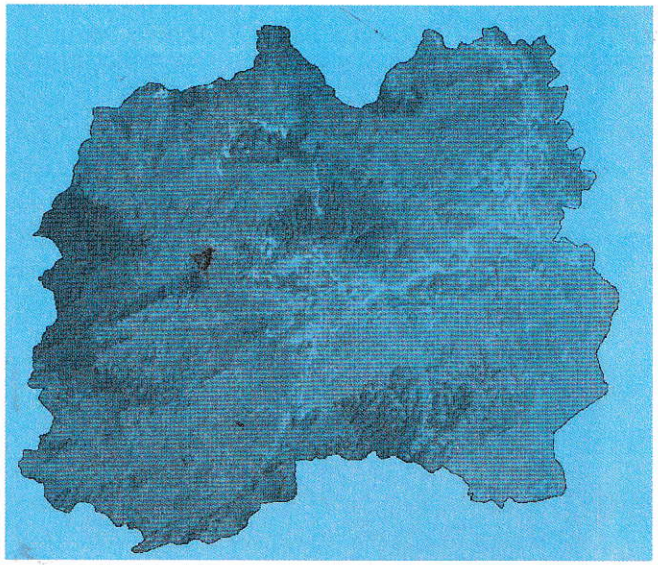
- Bước 1: Tiền xử lý ảnh vệ tinh;
- Bước 2: Lấy mẫu và xác định khoảng giá trị NDVI của các lớp đối tượng;
- Bước 3: Gán tạm thời các điểm ảnh về các lớp đối tượng;
- Bước 4: Tính trọng tâm cụm của các lớp $V=\{v_1,...,v_6\}$ từ các cụm ở giai đoạn 1, khởi tạo tham số mờ m và điều kiện dừng ε;
- Bước 5: Tính giá trị của hàm thuộc u_{ik} theo công thức (3);
- Bước 6: Cập nhật lại trọng tâm cụm $V=\{v_1,...,v_6\}$ theo công thức (4);
- Bước 7: Tính giá trị hàm mục tiêu, nếu điều kiện $\max\left\{\left\|u_{ij}^{(t+1)}-u_{ij}^{(t)}\right\|\right\}\le e$ thỏa mãn thì chuyển sang bước 8, ngược lại thì quay lại bước 5;
- Bước 8: Gán các điểm ảnh về các cụm, đánh giá độ chính xác, tô màu cho các lớp và đưa ra kết quả phân loại.

2. Thực nghiệm đánh giá biến động lớp phủ khu vực khai thác mỏ Quỳ Hợp, tỉnh Nghệ An

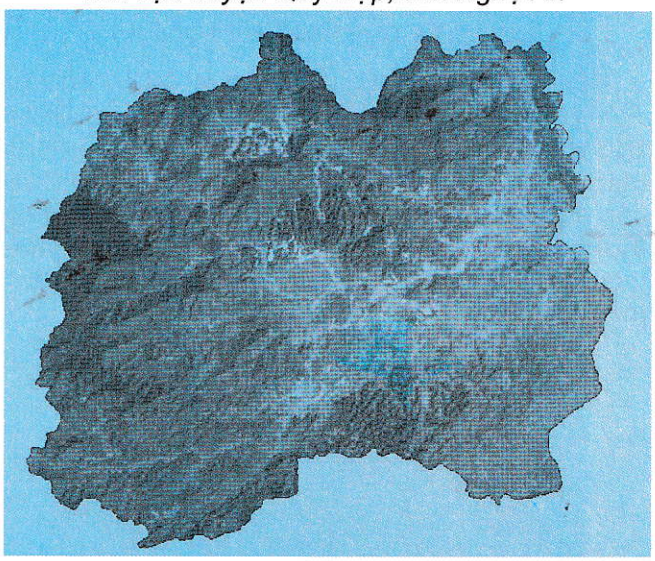
2.1. Khu vực nghiên cứu và tư liệu sử dụng

Khu vực thử nghiệm được lựa chọn là địa bàn huyện Quỳ Hợp, tỉnh Nghệ An. Đây là khu vực rất giàu tài nguyên khoáng sản với các mỏ thiếc, đá quý, đá xây dựng,... Hiện nay trên địa bàn huyện Quỳ Hợp có hàng chục điểm mỏ đang được khai thác. Quá trình khai thác mỏ bên cạnh những lợi ích to lớn về kinh tế-xã hội cũng gây ảnh hưởng nghiêm trọng tới cảnh quan và môi trường sống, đặc biệt là đối với lớp phủ thực vật. Tài nguyên rừng bị suy giảm một cách nhanh chóng, cả về diện tích lẫn trữ lượng.

Để đánh giá biến động lớp phủ khu vực huyện Quỳ Hợp, trong nghiên cứu sử dụng tư liệu ảnh vệ tinh Landsat giai đoạn 2000-2016. Tư liệu sử dụng bao gồm 3 cảnh ảnh: ảnh Landsat 7 ETM+ ngày 08-09-2000, ảnh Landsat 5 TM ngày 25-02-2010 và ảnh Landsat 8 ngày 10-02-2016. Các ảnh đều có chất lượng tốt, không bị ảnh hưởng bởi điều kiện thời tiết. Ảnh sau khi download từ website http://glovis.usgs.gov được hiệu chỉnh bức xạ, hiệu chỉnh hình học và cắt theo ranh giới hành chính huyện Quỳ Hợp (hình H.1÷H.3).



H.1. Ảnh vệ tinh Landsat 7 ETM+ ngày 08-09-2000 khu vực huyện Quỳ Hợp, tỉnh Nghệ An



H.2. Ảnh vệ tinh Landsat 5 TM ngày 25-02-2010 khu vực huyện Quỳ Hợp, tỉnh Nghệ An

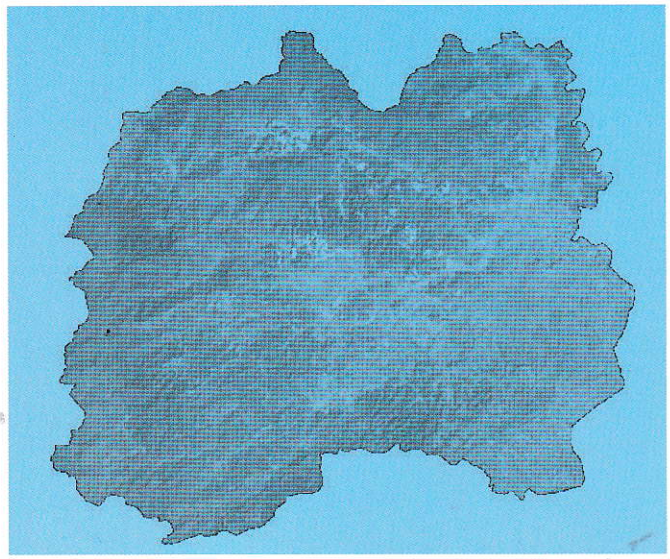
2.2. Kết quả phân loại và đánh giá biến động

Trong nghiên cứu này, chúng tôi tiến hành phân loại đối với 6 đối tượng lớp phủ, bao gồm: mặt nước (1); đất trống, đất ở (2); đồng cỏ, cây bụi (3); rừng nghèo (4); rừng trồng, rừng tái sinh (5); rừng tự nhiên (6). Ngôn ngữ lập trình C++ được sử dụng để cài đặt thuật toán và tiến hành phân loại.

Kênh đỏ và cận hồng ngoại (kênh 3, 4 đối với ảnh Landsat TM, ETM+; kênh 4, 5 đối với ảnh Landsat 8) được sử dụng để tính chỉ số thực vật NDVI theo công thức sau:

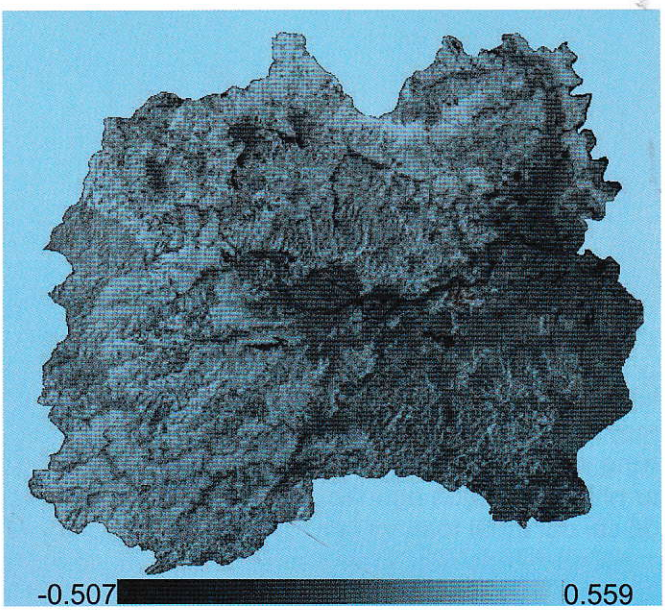
$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} - \rho_{RED}} \quad (4)$$

Trong đó: ρ_{NIR}, ρ_{RED} - Phản xạ phổ tại các kênh cận hồng ngoại và đỏ.

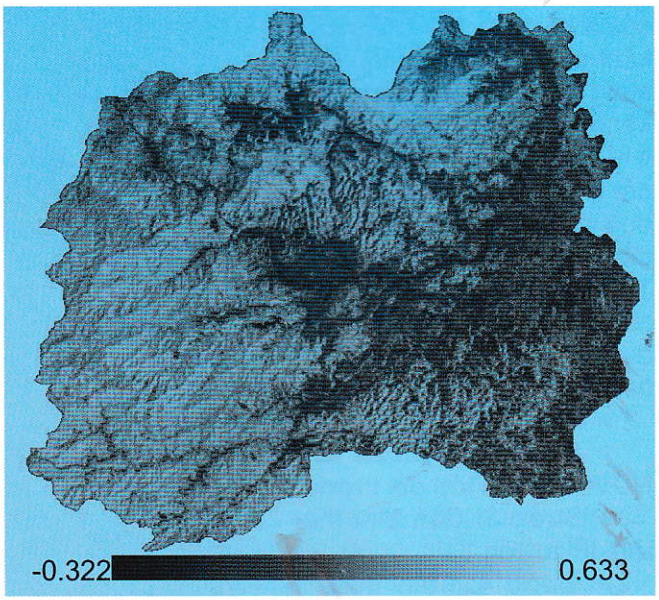


H.3. Ảnh vệ tinh Landsat 8 ngày 10-02-2016 khu vực huyện Quỳ Hợp, tỉnh Nghệ An

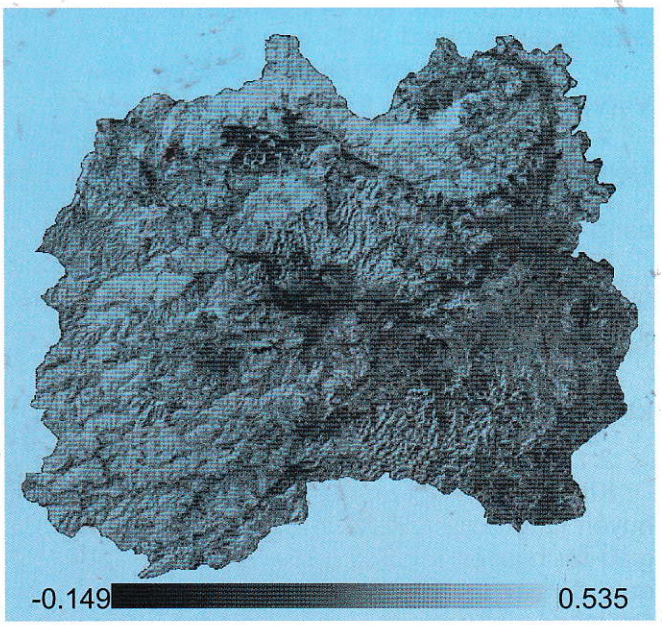
Kết quả xác định chỉ số thực vật NDVI đối với ảnh vệ tinh Landsat năm 2000, 2010 và 2016 khu vực huyện Quỳ Hợp được thể hiện trên các hình H.4÷H.6. Trên ảnh chỉ số thực vật NDVI, các pixel màu sáng đại diện cho các khu vực có thực vật phủ dày, trong khi các pixel màu tối thể hiện các khu vực không có thực vật hoặc thảm thực vật thưa.



H.4. Ảnh chỉ số thực vật NDVI khu vực huyện Quỳ Hợp năm 2000

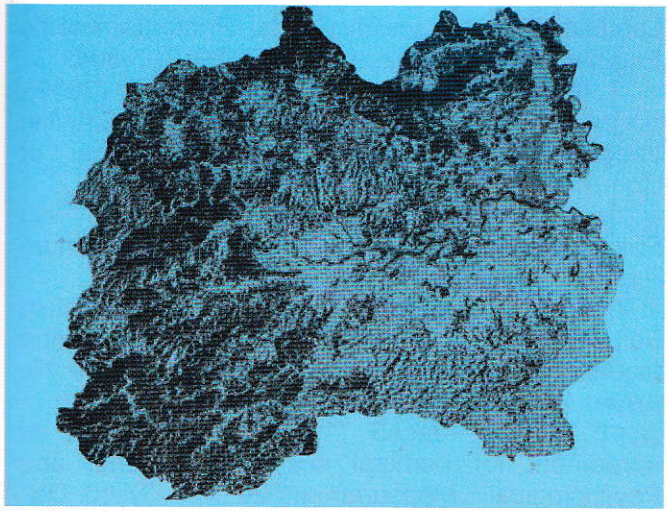


H.5. Ảnh chỉ số thực vật NDVI khu vực huyện Quỳ Hợp năm 2010

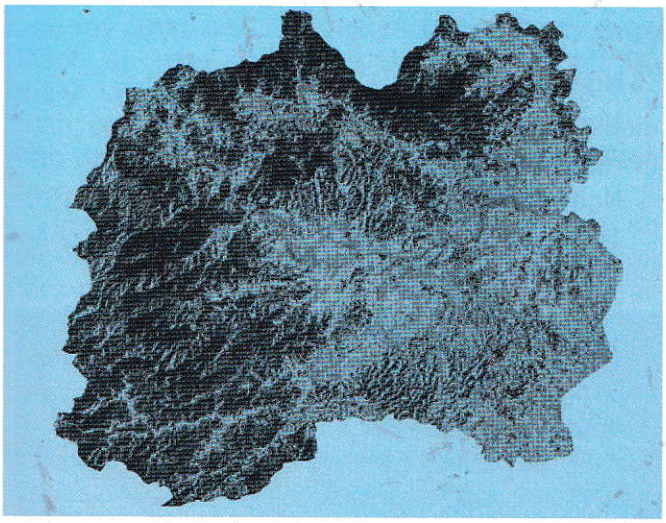


H.6. Ảnh chỉ số thực vật NDVI khu vực huyện Quỳ Hợp năm 2016

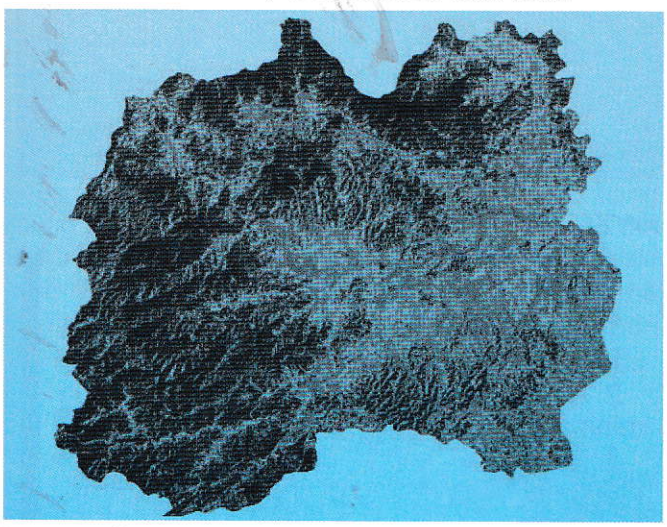
Ở giai đoạn đầu quá trình phân loại, chỉ số NDVI được sử dụng để lấy mẫu và xác định các khoảng giá trị cho các đối tượng, sau đó gán tạm thời các điểm ảnh về các đối tượng (bước 1÷3 trong phần1.2). Giai đoạn sau, phương pháp phân cụm mờ được áp dụng để thực hiện phân loại trên các kênh của ảnh vệ tinh đa phổ Landsat (bước 4-8 trong phần 1.2). Kết quả phân loại lớp phủ đối với tư liệu ảnh vệ tinh Landsat các năm 2000, 2010 và 2016 khu vực huyện Quỳ Hợp, tỉnh Nghệ An được thể hiện trên các hình H.7÷H.9 và Bảng 1.



H.7. Kết quả phân loại lớp phủ đối với ảnh vệ tinh Landsat năm 2000



H.8. Kết quả phân loại lớp phủ đối với ảnh vệ tinh Landsat năm 2010



H.9. Kết quả phân loại lớp phủ đối với ảnh vệ tinh Landsat năm 2016

Phân tích kết quả nhận được cho thấy, diện tích rừng (bao gồm cả rừng tự nhiên, rừng trồng và rừng nghèo) giảm một cách đáng kể trong giai đoạn từ 2000 đến 2016. Năm 2000, diện tích rừng tự nhiên chiếm 16,97 % tổng diện tích toàn huyện. Đến năm 2010, diện tích rừng tự nhiên giảm xuống còn 13,90 % và tiếp tục giảm xuống 11,99 % vào năm 2016. Diện tích rừng nghèo; rừng trồng, rừng tái sinh giảm từ 27,49 % và 25,33 % năm 2000 xuống còn 18,94 % và 15,68 % năm 2016. Như vậy, tính chung diện tích đất rừng huyện Quỳ Hợp đã giảm từ 69,79 % năm 2000 xuống 52,34 % năm 2010 và 46,61 % năm 2016.

Diện tích đất mặt nước có biến động không đáng kể, trong khi diện tích đất trống, đất ở tăng rất nhanh trong giai đoạn 2000-2016, từ 8,16 % lên 24,18 % (tương đương 3 lần).

Diện tích đất nông nghiệp, đồng cỏ cũng tăng một cách đáng kể với mức tăng khoảng 31,88 % trong giai đoạn 2000-2016 (Bảng 1).

Bảng 1. Kết quả phân loại lớp phủ trên ảnh vệ tinh Landsat năm 2000, 2010 và 2016

Tên đối tượng	Màu sắc	Diện tích (%)		
		2000	2010	2016
Mặt nước		2,29	4,10	3,15
Đất trống, đất ở		8,16	20,36	24,18
Đồng cỏ, đất nông nghiệp		19,77	23,20	26,07
Rừng nghèo		27,49	20,85	18,94
Rừng trồng, rừng tái sinh		25,33	17,59	15,68
Rừng tự nhiên		16,97	13,90	11,99

Bảng 2. Đánh giá độ chính xác kết quả phân loại

№	Chỉ số	Năm		
		2000	2010	2016
1	Kappa	0,8223	0,8031	0,8209
2	IQI	0,8892	0,8471	0,8752

Để đánh giá độ chính xác của kết quả phân loại, trong nghiên cứu sử dụng đồng thời hai chỉ số, bao gồm hệ số Kappa và chỉ số chất lượng hình ảnh IQI (Image Quality Index) [10], [11]. Giá trị các chỉ số này càng tiến gần đến 1 thể hiện độ chính xác của kết quả phân loại càng tăng. Đối với kết quả phân loại lớp phủ trên 03 cảnh ảnh vệ tinh Landsat khu vực huyện Quỳ Hợp, giá trị hệ số Kappa và IQI đều đạt trên 0,8 (Bảng 2). Như vậy, có thể khẳng định phương pháp phân cụm mờ có thể sử dụng hiệu quả trong phân loại lớp phủ với độ chính xác đảm bảo. Bên cạnh đó, việc áp dụng phân cụm mờ

trong phân loại ảnh vệ tinh cũng giúp nâng cao khả năng tự động hóa so với các phương pháp phân loại có giám sát thông dụng khác.

3. Kết luận

Quỳ Hợp là một địa phương có nguồn tài nguyên khoáng sản phong phú, đa dạng so với các khu vực khác ở Nghệ An. Ngành công nghiệp khai thác khoáng sản đã mang lại những lợi ích to lớn phục vụ phát triển kinh tế-xã hội ở địa phương. Tuy nhiên, quá trình khai khoáng cũng ảnh hưởng nghiêm trọng đến môi trường sống, làm biến đổi cảnh quan và lớp phủ mặt đất.

Công nghệ viễn thám là một công cụ hiệu quả trong phân loại và giám sát biến động lớp phủ. Kết quả nhận được trong nghiên cứu cho thấy, phương pháp phân loại lớp phủ sử dụng phân cụm mờ có độ chính xác cao, giúp nâng cao khả năng tự động hóa trong phân loại. Kết quả nhận được cũng là một nguồn thông tin quan trọng giúp các nhà quản lý đưa ra các biện pháp bảo vệ cảnh quan, môi trường trong quá trình khai thác khoáng sản.❐

TÀI LIỆU THAM KHẢO

1. Võ Chí Mỹ, Trịnh Thị Hoài Thu, Phạm Thị Làn, Lê Thị Thu Hà (2010). Ứng dụng chỉ số thực vật giám sát biến động lớp phủ bề mặt do quá trình khai thác than, Hội nghị Khoa học và Kỹ thuật Mỏ quốc tế, trang 252-255.

2. Nguyen Dinh Duong (2006). Study land cover change in Vietnam in period 2001-2003 using MODIS 32days composite, truy cập ngày 4/11/2012, tại trang web www.geoinfo.com.vn.

3. Mai Đình Sinh, Ngô Thành Long, Fuzzy C-Means Clustering with Spatial Information for Satellite Image Classification, 2014 Workshop on Aerospace Technology and Applications, Vietnam, 12/2014, ISBN: 978-604-913-305-3, pp 365-378.

4. Mai Đình Sinh, Trịnh Lê Hùng, Ngô Thành Long (2016). Combining Fuzzy probability and Fuzzy clustering for multispectral satellite imagery classification, Tạp chí Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, số 54(3), trang 300 - 313.

5. Mai Đình Sinh, Trịnh Lê Hùng, Đào Khánh Hoài (2016). Phương pháp bán giám sát trong phân loại lớp phủ trên ảnh vệ tinh sử dụng thuật toán Mountain, Tạp chí Khoa học ĐHSP TPCHM, số 03(81), 132 - 143.

6. Rogan, Miller, J., Stow, D., Franklin, J., Levien, L., & Fischer, C. (2003). Land-cover change monitoring with classification trees using Landsat TM and ancillary data, Photogrammetric Engineering and Remote Sensing, 69: 793-804.

7. Rauf K. S., Valentin V. G.,Leonid P. (2009). Fuzzy clustering methods in Multispectral Satellite Image Segmentation, International Journal of Computing, Vol. 8, Issue 1, 87-94.

8. J. C. Bezdek (1981). Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York.

9. James C. Bezdek, Robert Ehrlich and William Full (1984). FCM: The fuzzy c-means clustering algorithm, Computers & Geosciences, Volume 10, Issues 2-3, Pages 191-203.

10. Cohen, Jacob (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement. **20** (1): 37-46. Doi: 10.1177/001316446002000104.

11. Wang Z. and Bovik A. C. (2002). A universal image quality index, IEEE signal processing letters, vol. 9, no. 3, March: 81-84.

Ngày nhận bài: 19-12-2016

Ngày gửi phản biện: 26-12-2016

Ngày nhận phản biện: 05-02-2017

Ngày chấp nhận đăng bài: 08-04-2017

Từ khóa: ảnh Landsat, phân cụm mờ, phân loại, lớp phủ, khoáng sản

SUMMARY

This article presents a result of land cover change assessment in mining area of Quỳ Hợp district, Nghệ An province from LANDSAT multispectral images period 2000-2016 based on Fuzzy logic. The results which are obtained in this study showed that the accuracy of fuzzy clustering methods in classification higher 80 %. This method can be used effectively for assessing and monitoring land cover change.

LỜI HAY Ý ĐẸP

1. Người hoàn thiện nhất là người có ích cho xã hội. *Kinh Coran.*

2. Người yêu nghề luôn thấy công việc của mình chưa hoàn hảo. *X. Pobli.*

3. Thời gian không đo lường bằng năm tháng mà bằng những gì chúng ta làm được. *H. Cason.*

VTH sưu tầm