

THÀNH LẬP BẢN ĐỒ THẨM THỰC VẬT TRÊN CƠ SỞ PHÂN TÍCH, XỬ LÝ ẢNH VIỄN THÁM

TS. Hoàng Xuân Thành

Đại học Thủy Lợi

Tóm tắt: *Thảm phủ thực vật là một trong nhiều yếu tố quan trọng được sử dụng để đánh giá các quá trình tự nhiên như xói lở, trượt lở, lũ lụt cũng như tốc độ phá hủy môi trường tự nhiên do các hoạt động nhân sinh. Đối với những khu vực miền núi hiểm trở, thành lập bản đồ thảm phủ gặp nhiều khó khăn do không thể tiến hành lấy mẫu phân tích đều khắp vùng. Thành lập sơ đồ thảm phủ thực vật từ phân tích ảnh viễn thám đã rút gọn thời gian và làm tăng độ chính xác của bản đồ.*

Kết quả phân tích và phân loại có kiểm định đối với ảnh Landsat vùng Tủa Chùa - Lai Châu có kết quả tốt với hệ số chính xác $K = 0.7$, nằm trong giới hạn độ chính xác cao theo phương pháp kiểm tra sau phân loại Kappa. Kết quả cho thấy thảm phủ thực vật vùng Tủa Chùa được phân ra bảy lớp khác nhau là rừng ổn định, rừng non, rừng hỗn hợp, lúa và hoa màu, cây bụi, đất trống đồi trọc và nước.

1. Tổng quan các phương pháp xử lý ảnh viễn thám nhằm phân tích, phân loại thảm phủ

Các phương pháp phân tích ảnh viễn thám rất đa dạng. Có thể liệt kê một số phương pháp phân tích ảnh như phương pháp phân ngưỡng (Manual thresholds), phương pháp phân loại không kiểm định (Unsupervised), phương pháp phân loại có kiểm định (Supervised), phương pháp Fuzzy (Fuzzy classification or Mixing models) nhưng hai phương pháp đang dùng phổ biến để phân loại thảm phủ hiện nay là phương pháp phân loại không kiểm định (Unsupervised) và phương pháp phân loại có kiểm định (Supervised).

Mỗi phương pháp phân loại đều sử dụng những thuật toán nhất định. Các thuật toán có giới hạn và khả năng ứng dụng trong các trường hợp khác nhau (Shrestha and Alfred, 2001). Những thuật toán thường được sử dụng phổ biến là khoảng cách nhỏ nhất (Minimum Distance), Parallelepiped và Maximum Likelihood (Richards, 1994). Trong số này, thuật toán Maximum Likelihood được các nhà phân loại sử dụng nhiều nhất trong các công trình nghiên cứu thảm phủ. (Keuchel et al., 2003; Shrestha and Alfred, 2001; Swain and Davis, 1978; Estes et al., 1983; Schowengerdt, 1983; Sabins, 1986; Lillesand and Kiefer, 2000; Jensen, 1996). Thuật toán Minimum Distance thường được áp dụng trong phương pháp phân loại không kiểm định,

còn hai thuật toán Maximum Likelihood và Parallelepiped thường được áp dụng trong phương pháp phân loại có kiểm định [6]. Ngoài ra, người ta còn sử dụng một số phương pháp làm nổi bật yếu tố thực vật như phương pháp phân tích chỉ số thực vật- NDVI và phép biến đổi Tasseled cap.

Trong phạm vi bài báo này, tác giả chỉ đề cập tới một số phương pháp thường được sử dụng cho phân tích và giải đoán thảm phủ thực vật mà được nhiều nhà nghiên cứu hay sử dụng hiện nay đó là phương pháp phân loại có kiểm định (Supervised).

Phân loại có kiểm định là một phương pháp xác suất có khả năng sắp xếp những pixel do người sử dụng định nghĩa thành những lớp khác nhau, trong đó tất cả các pixel trên một ảnh được nhận dạng thông qua ký hiệu phổ tương tự với mục đích nhận ra sự đồng nhất, những mẫu đại diện mang nét đặc trưng thể hiện khác nhau mà chúng ta muốn phân loại. Những mẫu này gọi là những *khu vực lấy mẫu* (training). Sự lựa chọn những khu vực lấy mẫu thích hợp dựa trên phạm vi quan sát và ý tưởng này sẽ được hỗ trợ bởi các nguồn tài liệu đáng tin cậy như ảnh hàng không, các bản đồ, hay những dữ liệu khảo sát thực địa.

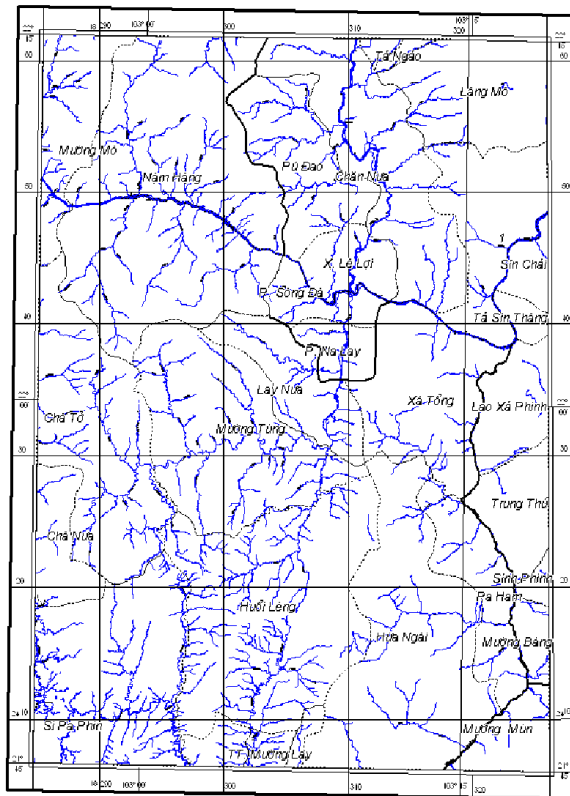
Những khu vực lấy mẫu trên ảnh số thường tham khảo dữ liệu khảo sát thực địa và ảnh hồng ngoại hàng không của khu vực đó. Những vùng

được mô tả trên ảnh số sẽ cũng là những nơi đại diện của khu vực lấy mẫu. Càng nhiều khu vực lấy mẫu thì độ chính xác toàn diện của ảnh phân loại sẽ càng cao. Những vị trí lấy mẫu sử dụng dấu hiệu phổ để đưa ra nét phác thảo của những khu vực đó. Hầu hết sự đồng nhất và những khu vực đại diện cho các vùng có đặc điểm thảm phủ khác nhau được so sánh với đặc tính phổ và sự khác nhau trong biểu đồ phân bố phổ. Sự phân loại ảnh số sử dụng thông tin phổ tiêu biểu bởi một hay nhiều kênh phổ của ảnh vệ tinh và cố gắng để phân loại mỗi pixel độc lập dựa trên thông tin phổ này. Trong trường hợp khác, các đối tượng được ấn định tất cả những pixel trong ảnh theo các lớp riêng biệt hay theo các chủ đề riêng (ví dụ như nước, rừng lá kim, rừng rụng lá...). Sự lựa chọn các loại thảm phủ phụ thuộc vào dữ liệu phổ, thuật toán phân loại và ý kiến của giới chuyên môn trong việc phân loại thảm phủ [3].

2. Khu vực nghiên cứu

Vùng nghiên cứu thuộc tỉnh Lai Châu có diện tích khoảng 2270km² với tọa độ địa lý từ 102°55'8" Đông – 21°45'06" Bắc tới 103°18'38" Đông – 22°15'01" Bắc. Đây là vùng núi cao hiểm trở, địa hình phân cắt mạnh, từ các thung lũng giữa núi có độ cao tuyệt đối chưa tới 200m (thung lũng Nậm Lay, Nậm Na) tới các dải núi cao hơn 1900m (dải núi Phu Dao, Nậm He). Nhiều dãy núi kéo dài theo phương Tây Bắc – Đông Nam hoặc gần trùng hướng Bắc Nam, tức là gần trùng với phương cấu trúc địa chất.

Vùng nằm trong lưu vực sông Đà và các sông nhánh của nó là sông Nậm Na, Nậm Lay, Nậm Múc. Các sông trên và hệ thống các suối nhánh của chúng thường có lòng hẹp, lởm thác ghềnh, lưu lượng nước biến đổi mạnh theo mùa. Trong khu vực này rừng nguyên thủy chiếm khoảng 20% diện tích phân bố ở vùng núi cao hoặc ven khe suối đi lại khó khăn, phần còn lại là nương rẫy và hầu như trọc hoá hoàn toàn do sự chặt phá, đốt cây cối làm nương rẫy. Nhiều loại gỗ quý như thông, pomu, lát, lim...ngày càng hiếm.



Hình 1: Diện tích khu vực nghiên cứu

3. Thành lập sơ đồ thảm phủ thực vật tỉ lệ 1:50.000 trên cơ sở phân tích, xử lý ảnh Landsat TM5

3.1. Dữ liệu sử dụng trong phân tích ảnh

Trong phạm vi nghiên cứu này, dữ liệu ảnh viễn thám Landsat TM5 chụp năm 2006 được sử dụng để phân tích và thành lập bản đồ thảm phủ. Ảnh có các đặc tính sau:

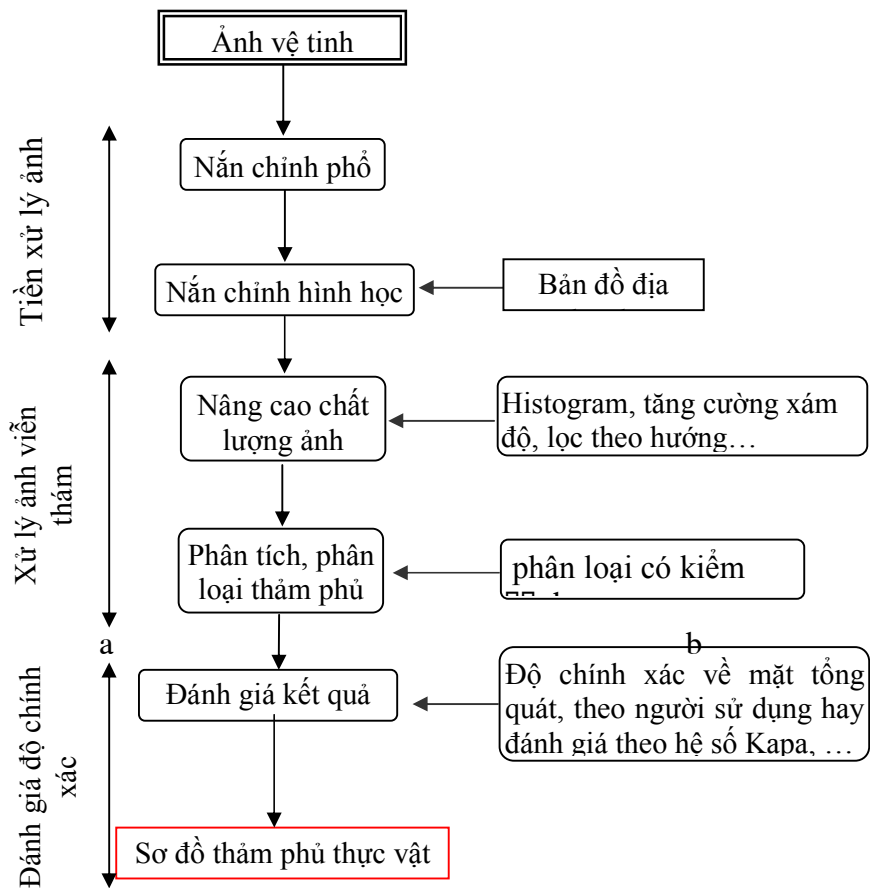
- Cảnh 1 có tuyến bay: 128, hàng: 45 (chụp năm 2006);
- Cảnh 2 có tuyến bay: 129, hàng: 45 (chụp năm 2006);

Các ảnh đều có độ phân giải 30m, độ phủ trùm 185×185km. Hệ tọa độ dùng để đăng ký cho ảnh là UTM, Elipxoid quy chiếu: WGS-84, đơn vị tính: mét, múi chiếu: 48.

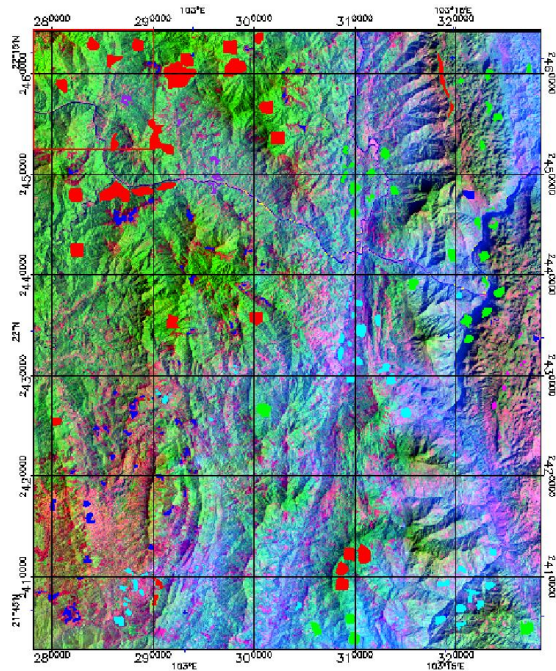
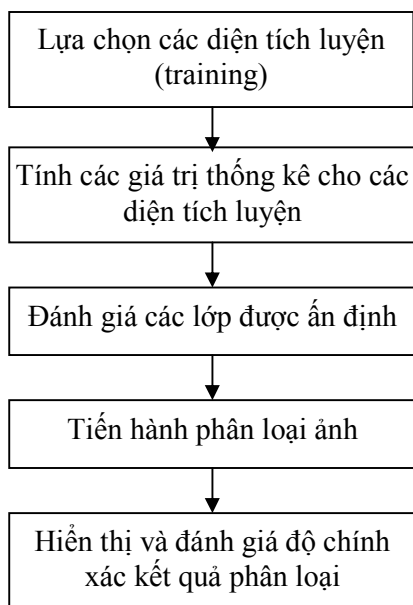
Ngoài ra, bản đồ địa hình tỷ lệ 1/50.000, hệ tọa độ VN-2000, Elipxoid quy chiếu: WGS-84, đơn vị: mét, múi chiếu: 48 cũng được sử dụng để phục vụ công tác nắn chỉnh hình học ảnh và loại bỏ các sai số về địa hình trên ảnh.

3.2. Phương pháp xử lý, phân loại ảnh viễn thám

Quy trình xử lý ảnh viễn thám để phân tích thảm phủ thực vật cho khu vực nghiên cứu nhìn chung tuân theo quy trình đã được mô tả trong hình 2.



Hình 2. Phương được sử dụng trong phân tích ảnh vệ tinh để thành lập sơ đồ thảm phủ thực vật



Hình 3: a- Quy trình phân loại có kiểm định của khu vực nghiên cứu; b- Các khu vực lấy mẫu khác nhau thể hiện trên trên ảnh Landsat TM5 bằng các polygon màu đỏ.

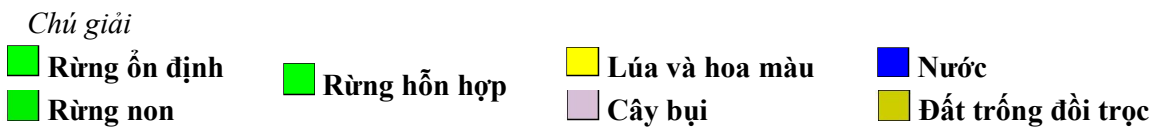
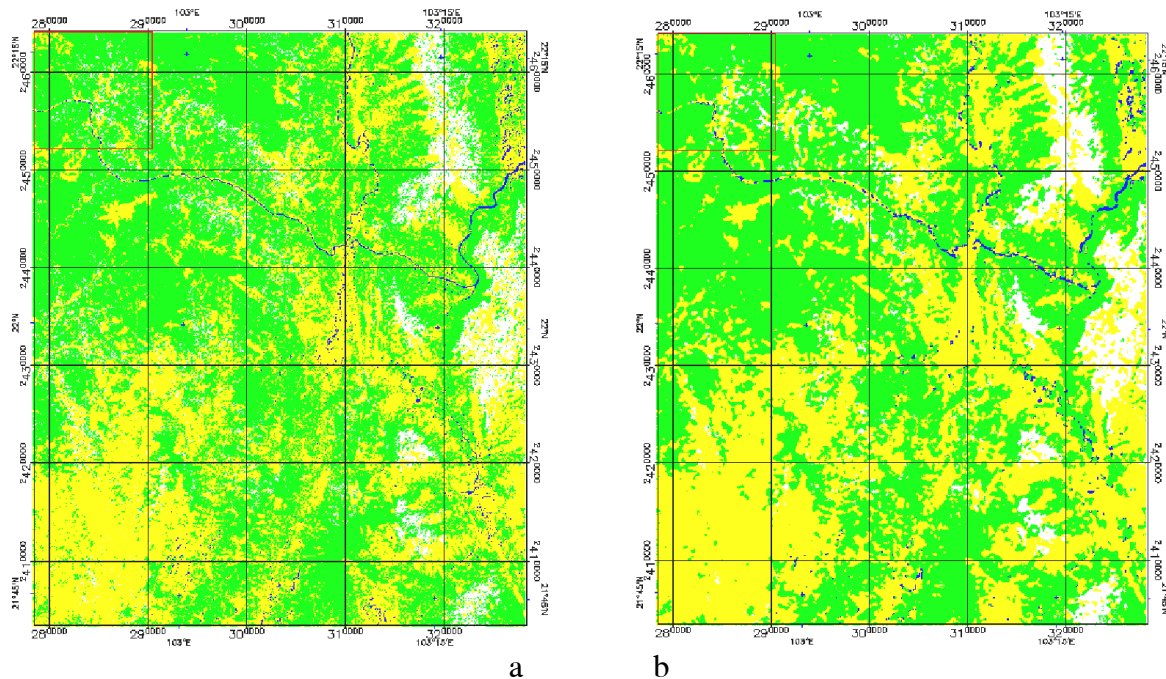
Tuy nhiên trong phần phân tích, phân loại thảm phủ thì chúng tôi đã lựa chọn phương pháp phân loại có kiểm định vì có nhiều ưu điểm [4] so với các phương pháp khác. Quy trình phân loại có kiểm định (chưa đánh giá độ chính xác) của khu vực nghiên cứu được mô tả ở hình 3a.

Trong quá trình lấy mẫu cho khu vực nghiên cứu đã xác định được 7 loại thảm phủ thực vật và loại khác (có kiểm tra ngoài thực địa) đó là rừng ổn định, rừng hỗn hợp, rừng non, lúa và màu, đất trồng đồi trọc, cây bụi và nước.

Trên hình 3b, các khu vực lấy mẫu khác nhau được thể hiện trên ảnh Landsat TM5 bằng các polygon với màu sắc khác nhau (thể hiện sự phân biệt giữa các loại đối tượng mà người sử dụng lựa chọn).

3.3. Kết quả phân loại

Sau khi việc lấy mẫu luyện và chỉ định tên cũng như màu cho cho các mẫu luyện thành một số lớp đối tượng riêng biệt thì quá trình tiến hành phân loại sẽ được thực hiện. Quá trình phân loại của khu vực nghiên cứu đã thu được sơ đồ ảnh với 7 lớp đối tượng khác nhau như trong hình 4a. Kết quả phân loại này mới ở dạng thô và trong đó có nhiều khu vực với diện tích quá nhỏ và không cần thiết cho nghiên cứu sau này. Do vậy để loại bỏ các đối tượng nằm đơn lẻ với diện tích nhỏ, chúng tôi đã sử dụng bộ lọc thành phần chủ yếu (Majority) để thực hiện quá trình lọc ảnh để thu được kết quả trơn hơn và đáp ứng với các yêu cầu đặt ra (hình 4b).



Hình 4. Quá trình phân loại của khu vực nghiên cứu đã thu được sơ đồ ảnh với: a- 7 lớp đối tượng khác nhau; b- sơ đồ thảm phủ gồm 7 lớp sau khi tiến hành lọc Majority

3.4. Đánh giá độ chính xác của phép phân loại có kiểm định

Đánh giá độ chính xác là thuật toán xác định độ tin cậy của sự phân loại ảnh. Độ chính xác của

ảnh được phân loại dựa vào khu vực mà nó đặt dữ liệu tham khảo (ground truth map). Hầu hết những phương pháp để đánh giá độ chính xác sự phân loại bao gồm một ma trận được xây dựng từ

2 loại dữ liệu (ví dụ: bản đồ sự phân loại viễn thám và dữ liệu tham khảo). Độ chính xác còn thể hiện mức độ phù hợp giữa những gì quan sát được và thực tế (thường là dưới dạng phần trăm).

Một ma trận sai số là một ma trận vuông được sắp xếp theo hàng và cột chỉ rõ số lượng các mẫu pixel được gán cho một lớp riêng biệt liên quan tới các lớp hiện thời, được thực hiện bởi việc tham khảo dữ liệu. Độ chính xác toàn diện được tính bởi tổng pixel phân loại chính xác và tổng số pixel tách rời ra. Trong nghiên cứu này, phương pháp lấy mẫu ngẫu nhiên được dùng để đánh giá độ chính xác sự phân loại. Để đánh giá độ chính xác của sự phân loại thảm phủ, những mẫu ngẫu nhiên được mô tả cho mỗi lớp thực vật riêng biệt.

Độ chính xác rất cao của phép phân loại thường được chấp nhận phổ biến là trên 0.85 (85%), độ chính xác vừa phải thì nằm trong khoảng 0.4÷0.8. Các thông số này do Cục Địa chất Mỹ quy định [2]. Hệ số Kappa được sử dụng là thước đo đánh giá độ chính xác phân loại. Trái ngược hẳn với độ chính xác toàn diện được miêu tả ở trên, đây là hệ số tiện ích của tất cả các nguyên tố từ ma trận ở trên. Nó là sự khác nhau cơ bản giữa những gì có thực về sai số độ lệch của ma trận và tổng số thay đổi được chỉ ra bởi hàng và cột.

$$\hat{K} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} - x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} - x_{+i})} \quad (8)$$

r = số lượng cột trong ma trận ảnh

x_{ii} = số lượng pixel quan sát được tại hàng i và cột i (trên đường chéo chính)

x_{i+} = tổng pixel quan sát tại hàng i

x_{+i} = tổng pixel quan sát tại cột i

N = Tổng số pixel quan sát được trong ma trận ảnh

Hệ số Kappa thường nằm giữa 0 và 1, giá trị nằm trong khoảng này thì độ chính xác của sự phân loại được chấp nhận. Kappa có 3 nhóm giá trị:

- $K > 0.8$: độ chính xác cao
- $0.4 < K < 0.8$: độ chính xác vừa phải
- $K < 0.4$: độ chính xác thấp

Kết quả đánh giá mức độ chính xác sau khi phân loại ra 7 lớp thảm phủ bằng phương pháp phân loại có kiểm định chỉ ra rằng:

Độ chính xác toàn cục của phân loại (Overall Accuracy) = 77.3%

Hệ số thống kê Kappa = 0.7

Ngoài ra độ chính xác của phép phân loại còn được tính theo nhiều phương pháp khác nhau và kết quả tính toán được mô tả trong bảng 2. Nhìn chung kết quả phân loại thảm phủ thực vật là đạt mức chấp được.

Bảng 1. Ma trận đánh giá độ chính xác của phép phân loại theo %

Điểm quan sát (%)								
Lớp	Rừng ổn định	Rừng hỗn hợp	Nước	Lúa và màu	Cây bụi	Rừng non	Đất trống đồi trọc	Tổng số hàng
Rừng ổn định	83.88	0.66	0.00	0.58	2.36	3.19	7.67	38.46
Rừng hỗn hợp	3.46	78.78	27.71	19.27	5.59	2.57	0.32	20.52
Nước	0.00	0.81	33.42	2.52	0.00	0.83	0.00	1.01
Lúa và màu	0.77	8.19	36.83	64.47	14.66	9.13	15.85	16.19
Cây bụi	5.30	6.24	0.14	1.51	65.96	5.29	0.74	4.94
Rừng non	6.32	5.07	1.50	2.89	5.06	78.40	0.44	6.17
Đất trống đồi trọc	0.27	0.24	0.41	8.75	6.36	0.58	74.98	12.71
Tổng số cột	100	100	100	100	100	100	100	100

Bảng 2: Độ chính xác của phép phân loại được tính theo nhiều phương pháp khác nhau

Lớp	Độ chính xác theo người số liệu (%)	Độ chính xác theo người sử dụng (%)
Rừng ôn định	83.88	96.23
Rừng hỗn hợp	78.78	73.88
Nước	33.42	39.47
Lúa và màu	64.47	68.68
Cây bụi	65.96	18.39
Rừng non	78.40	28.41
Đất trống đồi trọc	74.98	85.99

4. Kết luận

Vùng nghiên cứu nằm trong hai cảnh ảnh, vì vậy việc đầu tiên là phải tiến hành ghép ảnh. Công việc này đòi hỏi chất lượng của các ảnh phải tương đương nhau (khác nhau rất ít về mặt phổ

phản xạ) để khi ghép ảnh có sự khác biệt về phổ. Ảnh vùng nghiên cứu chụp hai mùa khác nhau năm 2006. Tuy nhiên về mặt phổ phản xạ thì hai ảnh có sự khác nhau không nhiều. Vùng nghiên cứu nằm gần như trọn vẹn trong một cảnh ảnh chỉ có một phần bên phải là nằm trong cảnh khác vì vậy việc tiến hành ghép ảnh và tiến hành cân bằng giá trị phổ cũng đảm bảo được chất lượng.

▪ Việc phân loại ảnh rất mất thời gian do phải tiến hành xử lý hiện tượng hơi nước che phủ trong ảnh. Sự bốc hơi của nước ở những khu vực thung lũng sông làm cho công tác phân loại khó khăn hơn.

▪ Sau khi phân loại có kiểm định tác giả đã tiến hành phân chia vùng nghiên cứu thành 7 loại thảm phủ khác nhau với chỉ số $K \sim 0.7$, giá trị này đã đáp ứng được mức độ nghiên cứu. Sở dĩ hệ số $K \sim 0.7$ là có thể chấp nhận được do điều kiện địa hình là vùng núi nên khả năng giải đoán cũng chỉ đạt được độ chính xác cao đến mức đó.

Tài liệu tham khảo

1. Daniel L. Civco, James D. Hurd, Emily H. Wilson, Mingjun Song, Zhenkui Zhang -Center for Land use Education And Research - Department of Natural Resources Management & Engineering - The University of Connecticut, 2002. A Comparison of Land Use and Land Cover Change Detection Methods. Presentation pages 2-74.
2. Freek D. van der Meer and Steven M. de Jong. Imaging Spectrometry, basic principles and prospective application. Kluwer Academic Publishers, 3-12, 115-130.
3. JAMES R. ANDERSON, ERNEST E. HARDY, JOHN T. ROACH, and RICHARD E. WITMER. A Land Use And Land Cover Classification System For Use With Remote Sensor Data, 7-37.
4. Ronald B. Smith, Laurent Bonneau - Yale Center for Earth Observation, February 2006. Landcover Classification using Satellite Images. Classification lecture 2-39.
5. Stanley Max. Using ERDAS Imagine 8.7 to conduct an Unsupervised Classification of US Army Corps of Engineers - Institute for Water Resources - Hydrologic Engineering Center (www.hec.usace.army.mil), June 1985. A tutorial on creating a grid cell land cover data file from remote sensing data, 23-38.
6. Shobha Sriharan, Virginia State University - Colleague: DeNeice Guest, Lockheed Martin-NASA Faculty Fellowship Program 2004 - Stennis Space Center, 2004. Analysis of Land Cover

Classes Using Unsupervised and Supervised Classification of Stennis Space Center (SSC) Image. Presentation pages 3-31.

Abstract

**VEGETATION COVER MAPPING BASED ON
REMOTELY SENSED IMAGE ANALYSIS**

Carpet vegetation cover is one of several important factors used to evaluate the natural processes like erosion, landslide, flood and speed to destroy the natural environment by human activities. For the dangerous mountain region, mapping the carpet difficult because can not take samples were analyzed throughout the region. Established carpet vegetation cover map from remote sensing image analysis has reduced time and increased accuracy of the map.

The results of analysis and testing classification for Landsat images Tua Chua region - Lai Chau have good results with exactly the coefficient $K = 0.7$, are in limited precision inspection methods after classifying Kappa . Results showed that vegetation cover area rugs Tua Chua assigned seven different layers of forest is stable, young forest, mixed forest, rice and vegetables, shrubs, bare soil and water.