

NGHIÊN CỨU HOÀN THIỆN PHƯƠNG PHÁP TẦN SUẤT - NHẬN DẠNG TRONG XỬ LÝ VÀ PHÂN TÍCH SỐ LIỆU ĐỊA VẬT LÝ

VÕ THANH QUỲNH¹, NGUYỄN XUÂN BÌNH², NGUYỄN ĐỨC VINH¹

¹ Trường Đại học Khoa học Tự nhiên, ĐHQGHN

² Viện Vật lý địa cầu, Viện KH&CN Việt Nam

Tóm tắt: Các phương pháp phân tích nhận dạng đóng vai trò quan trọng trong xử lý và phân tích tài liệu địa vật lý, đặc biệt là đối với tài liệu địa vật lý hàng không trong tìm kiếm và dự báo triển vọng khoáng sản. Phương pháp tần suất - nhận dạng là một trong số các phương pháp nhận dạng mới, được xây dựng trên cơ sở ứng dụng phương pháp phân tích tần suất trong thuật toán đánh giá, lựa chọn thông tin và thuật toán phân tích đối sánh, xác định các đối tượng đồng dạng. Phương pháp này hiện đang được áp dụng bước đầu có hiệu quả trong xử lý và phân tích tài liệu phổ gamma hàng không phục vụ tìm kiếm và dự báo triển vọng khoáng sản. Tuy nhiên, phương pháp này vẫn còn một số hạn chế cần được tiếp tục nghiên cứu hoàn thiện và mở rộng hơn nữa phạm vi áp dụng của phương pháp. Bài báo này trình bày những kết quả nghiên cứu hoàn thiện nội dung phương pháp, cho phép giải quyết bài toán nhận dạng trong xử lý số liệu địa vật lý một cách triệt để hơn đối với cả hai trường hợp: biết trước và chưa biết trước các đối tượng đối sánh.

I. ĐẶT VẤN ĐỀ

Phương pháp tần suất - nhận dạng là một phương pháp phân tích nhận dạng mới, được xây dựng trên cơ sở ứng dụng phương pháp phân tích tần suất theo thuật toán Griffiths-Vinni trong lớp bài toán đánh giá và lựa chọn thông tin vào việc giải quyết bài toán nhận dạng trong địa vật lý. Phương pháp này đã được áp dụng thử nghiệm có hiệu quả trong xử lý và phân tích tài liệu phổ gamma hàng không, phục vụ tìm kiếm và dự báo triển vọng khoáng sản. Tuy nhiên, khi tiến hành áp dụng thực tế xử lý và phân tích các tài liệu địa vật lý hàng không, phương pháp tần suất - nhận dạng vẫn còn một số tồn tại cần được nghiên cứu, khắc phục để hoàn thiện và mở rộng hơn nữa phạm vi ứng dụng của phương pháp. Cụ thể là: phương pháp này mới đưa ra được thuật toán phân tích đối sánh, xác định mức độ đồng dạng của đối tượng đối sánh so với đối tượng mẫu, làm cơ sở cho việc đánh giá, dự báo mức độ triển vọng của chúng, khi đối tượng đối sánh đã biết mà chưa giải quyết được nhiệm vụ tìm kiếm, phát hiện, khoanh định ranh giới các đối tượng đồng dạng và mức độ đồng dạng của chúng khi chưa biết trước các đối tượng đối sánh. Phương pháp này cũng chỉ mới áp dụng có hiệu quả đối với tài liệu dị thường phổ gamma hàng không mà chưa được mở rộng cho các dạng tài liệu địa vật lý khác, kể cả tài liệu từ hàng không. Bài báo này trình bày kết quả nghiên cứu, hoàn thiện nội dung của phương pháp, bằng việc mở rộng thuật toán cho cả trường hợp biết trước và trường hợp chưa biết trước đối tượng đối sánh, cùng với việc xây dựng phần mềm xử lý trên máy tính tương ứng. Chương trình mới mang tên QTSM. Nó đã được áp dụng trong việc phân tích tài liệu địa vật lý hàng không (bao gồm cả tài liệu từ và tài liệu phổ gamma) trên một phần diện tích của vùng Tuy Hòa và thu được kết quả tốt. Hy vọng các kết quả nghiên cứu mới này có thể đưa vào áp dụng thực tế, góp phần đẩy nhanh và nâng cao chất lượng của công tác xử lý và

phân tích tài liệu địa vật lý hàng không, một nguồn tài liệu hết sức phong phú, nhưng chưa được khai thác triệt để ở nước ta hiện nay.

II. NỘI DUNG PHƯƠNG PHÁP TẦN SUẤT - NHẬN DẠNG

Phương pháp tần suất - nhận dạng được xây dựng trên cơ sở ứng dụng phương pháp phân tích tần suất theo thuật toán Griffiths-Vinni. Nội dung của phương pháp này đã được trình bày trong bài báo “*Xây dựng một phương pháp nhận dạng mới trong xử lý tài liệu địa vật lý trên cơ sở vận dụng kết hợp phương pháp phân tích khoảng cách khái quát và phân tích tần suất* [8].

Sau đây là tóm tắt những nội dung chính của phương pháp.

1. Phương pháp xây dựng ma trận thông tin của đối tượng mẫu

Từ tập hợp số liệu của các chủng loại thông tin của đối tượng mẫu xây dựng các đường cong biến phân (đường cong mật độ phân bố). Từ các đường cong biến phân xác định khoảng giá trị đặc trưng cho từng tham số. Sau khi có được các khoảng giá trị đặc trưng, dùng nó làm “cửa sổ quét” để tạo ra các đơn vị thông tin cho từng chủng loại thông tin của từng phần tử. Đối với mỗi phần tử của mỗi chủng loại thông tin, nếu nó nằm trong khoảng giá trị đặc trưng sẽ nhận giá trị là 1, nằm ngoài sẽ nhận giá trị là 0. Bằng cách này sẽ chuyển được một ma trận thông tin với các số liệu địa chất, địa vật lý bất kỳ về ma trận thông tin chuẩn theo yêu cầu của thuật toán với các phần tử là các giá trị 1 hoặc 0.

2. Phương pháp đánh giá lựa chọn tổ hợp thông tin

Đánh giá lựa chọn tổ hợp thông tin được thực hiện bằng phương pháp phân tích tần suất theo thuật toán Griffiths-Vinni, thông qua tỷ trọng thông tin P_m .

P_m là cơ sở để lựa chọn tập hợp các tính chất đủ chứa tải những thông tin cần thiết theo yêu cầu nghiên cứu, nghĩa là khi cho P_m một giá trị tỷ lệ % nào đó, ta sẽ tìm được tập hợp m tính chất tương ứng. Số lượng chủng loại thông tin m được lựa chọn tùy thuộc vào giá trị ngưỡng của P_m cho trước.

3. Phương pháp phân tích đối sánh, xác định các đối tượng đồng dạng

a) Xây dựng ma trận thông tin cho đối tượng đối sánh tương tự như đối với đối tượng mẫu thông qua các khoảng giá trị đặc trưng của chính đối tượng mẫu.

b) Tiến hành đánh giá tỷ trọng thông tin cho tất cả các tính chất của đối tượng đối sánh bằng phương pháp phân tích tần suất theo thuật toán Griffiths-Vinni.

c) Tính tỷ trọng thông tin của tổ hợp thông tin đã được lựa chọn của đối tượng mẫu cho đối tượng đối sánh. Có thể xem giá trị này tương tự hệ số đồng dạng, được gọi là chỉ số đồng dạng, ký hiệu P_m^*

d) Đối tượng đối sánh được xem là đồng dạng với đối tượng mẫu khi P_m^* có giá trị đạt mức giá trị ngưỡng nào đó.

Phương pháp tần suất - nhận dạng đã được phân tích tự động trên máy tính bằng chương trình QTS. Đây là một phương pháp phân tích tổ hợp, về nguyên tắc có thể cho phép xử lý đối với các dạng số liệu địa chất - địa vật lý khác nhau. Tuy nhiên, trên thực tế phương pháp này mới chỉ bước đầu được áp dụng có hiệu quả đối với tài liệu dị thường phổ gamma hàng không trong tìm kiếm và dự báo triển vọng khoáng sản.

III. HOÀN THIỆN VÀ MỞ RỘNG PHẠM VI ÁP DỤNG PHƯƠNG PHÁP TẦN SUẤT - NHẬN DẠNG

1. Nghiên cứu hoàn thiện nội dung phương pháp

Như đã nói ở trên, phương pháp tần suất - nhận dạng đang dừng lại ở mức tiến hành phân tích đối sánh, xác định mức độ đồng dạng của đối tượng đối sánh so với đối tượng mẫu, khi các đối tượng đối sánh đã biết mà chưa giải quyết được nhiệm vụ tìm kiếm, phát hiện các đối tượng đồng dạng và đánh giá mức độ đồng dạng của chúng khi chưa biết trước các đối tượng đối sánh.

Để mở rộng phạm vi áp dụng của phương pháp tần suất - nhận dạng cho trường hợp các đối tượng đối sánh chưa biết, bài toán được đặt ra là: trên một vùng nghiên cứu nào đó đã có các số liệu khảo sát địa vật lý cùng với các đối tượng mẫu đã biết (ví dụ như các điểm quặng, tụ khoáng hay các đối tượng địa chất cần nghiên cứu tìm kiếm nào đó v.v.), các đối tượng đối sánh chưa biết. Làm thế nào để phát hiện, khoanh định các diện tích đồng dạng với diện tích của đối tượng mẫu, cũng như đánh giá mức độ đồng dạng của chúng, làm cơ sở cho việc tìm kiếm, đánh giá triển vọng khoáng sản của các diện tích mới được phát hiện bằng phương pháp tần suất - nhận dạng. Để giải quyết nhiệm vụ này, chúng tôi đã đề xuất và xây dựng bổ sung một thuật toán phân tích đối sánh, xác định chỉ số đồng dạng như sau:

Bước 1. Nội suy các số liệu khảo sát địa vật lý thực tế lên mạng lưới đều (ô vuông hoặc chữ nhật) bằng các thuật toán nội suy hiện có.

Nội dung này được thực hiện dễ dàng bằng các thuật toán nội suy trong bộ chương trình phân tích phổ - thông kê COSCAD.

Bước 2. Thực hiện các nội dung như: a) Xây dựng ma trận thông tin của đối tượng mẫu; b) Đánh giá, lựa chọn tổ hợp thông tin trên đối tượng mẫu theo thuật toán Griffiths-Vinni.

Các nội dung này được thực hiện theo đúng phương pháp tần suất - nhận dạng đang làm.

Bước 3. Phân tích, tính toán chỉ số đồng dạng P_m^* cho tất cả các điểm trên các nút của của mạng lưới đều đã được nội suy, trên toàn bộ diện tích vùng nghiên cứu. Nội dung này được thực hiện như sau:

a) Dùng các “*cửa sổ quét*” để xác định ranh giới diện tích của các đối tượng đối sánh. Các “*cửa sổ quét*” có thể là các hình tròn, hình vuông, hình chữ nhật, hình ellipsoid với các kích thước và góc quay khác nhau. Các diện tích nằm trong cửa sổ quét được xem là các đối tượng đối sánh, cần tiến hành xử lý, phân tích đối sánh, xác định mức độ đồng dạng của chúng so với đối tượng mẫu thông qua chỉ số đồng dạng P_m^* . Nội dung này được thực hiện giống như trường hợp các đối tượng đối sánh đã biết của phương pháp tần suất - nhận dạng đã được trình bày. Giá trị P_m^* của đối tượng đối sánh vừa tính được sẽ được gán cho điểm trung tâm của cửa sổ quét. Điểm trung tâm này sẽ trùng với điểm nút của mạng lưới đều đã được nội suy.

b) Dịch chuyển cửa sổ quét khắp diện tích của vùng nghiên cứu, với bước dịch đều theo mạng lưới đã được nội suy. Kết quả ta được một file số liệu các chỉ số đồng dạng $P_m^*(x,y)$ theo tọa độ trùng với tọa độ của mạng lưới đã được nội suy trên khắp diện tích vùng nghiên cứu.

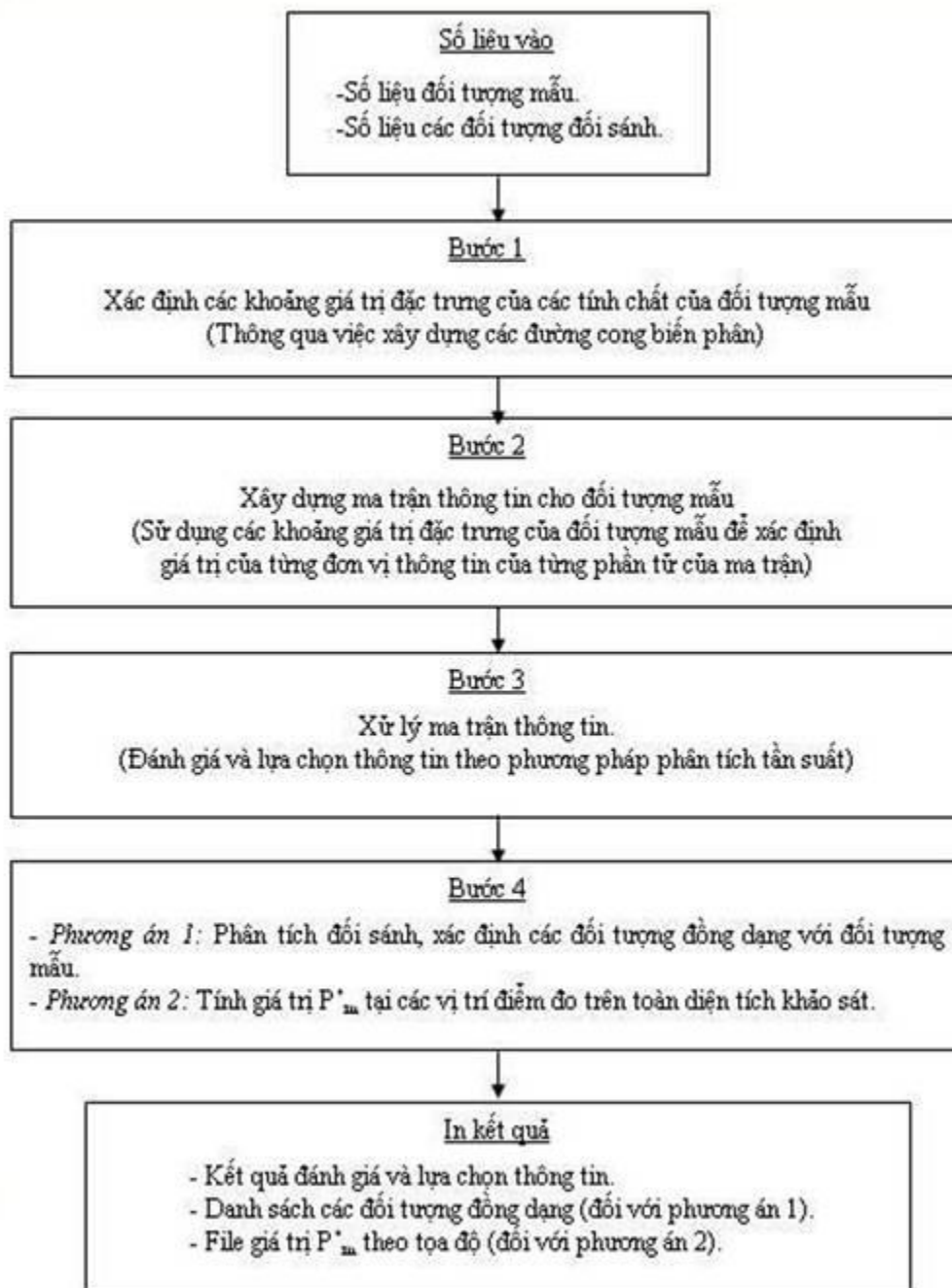
c) Khoanh định và đánh giá mức độ đồng dạng của các diện tích đồng dạng với đối tượng mẫu. Từ file số liệu này, với các mức giá trị ngưỡng cho trước ta sẽ có xác định được sự phân bố của các đối tượng đồng dạng cũng như mức độ đồng dạng của chúng so với đối tượng mẫu trên toàn diện tích nghiên cứu bằng việc xây dựng bản đồ đồng mức giá trị của $P_m^*(x,y)$.

Trên quan điểm của một phương pháp phân tích nhận dạng, đến đây bài toán nhận dạng cụ thể vừa đặt ra đã được giải quyết một cách triệt để như cách làm thông thường của các phương pháp phân tích nhận dạng trong địa vật lý.

2. Xây dựng chương trình và phân tích thử nghiệm

Theo nội dung phương pháp đã được bổ sung hoàn thiện như trình bày ở trên, chúng tôi đã xây dựng chương trình phân tích mới trên máy tính. Chương trình được viết bằng ngôn ngữ FORTRAN, trên cơ sở bổ sung, phát triển từ chương trình QTS hiện có, cho cả hai phương án: phương án biết trước và phương án chưa biết trước các đối tượng đối sánh. Chương trình có tên gọi mới: Chương trình QTSM. Các bước thực hiện của chương trình được thể hiện trên sơ đồ (Hình 1). Đây là một phương pháp phân tích tổ hợp, về nguyên tắc có thể cho phép xử lý đối với mọi dạng số liệu địa chất - địa vật lý bất kỳ với số lượng các tham số đầu vào đã được mở rộng đáng kể.

Chúng tôi đã tiến hành phân tích thử nghiệm bằng chương trình phân tích mới trên tài liệu thực tế ở diện tích phần phía đông vùng bay Tuy Hòa cho cả tài liệu từ và phổ gamma hàng không và thu được kết quả tốt, phù hợp với các kết quả phân tích trước đó bằng tổ hợp của nhiều phương pháp nhận dạng khác nhau. Các đối tượng mẫu được lựa chọn ở đây là các cụm dị thường đã được kiểm tra đánh giá mặt đất và được xác định là rất có triển vọng khoáng sản. Ví dụ cụm 68 có triển vọng vàng, cụm 38 có triển vọng thiếc (Bảng 1).



Bảng 1. Kết quả phân tích theo phương pháp tần suất - nhận dạng (QTSM) và so sánh với kết quả kiểm tra mặt đất
(Đối tượng mẫu - cụm 68-K)

STT	Các đối tượng đồng dạng (cụm)	Chỉ số đồng dạng	Đã kiểm tra mặt đất	Kết quả đánh giá
1	10	82,55 %	*	T.V loại 1
2	20	80,68 %		

3	24	81,95 %		
4	39	84,24 %		
5	60	75,27 %	*	T.V loại 2
6	66	86,31 %		
7	79	79,96 %		
8	82	76,25 %		

IV. KẾT LUẬN

Bằng việc đề xuất và xây dựng bổ sung một thuật toán phân tích đối sánh, xác định chỉ số đồng dạng đối với trường hợp khi chưa biết trước các đối tượng đối sánh cho phương pháp tần suất - nhận dạng, chúng ta đã có được một phương pháp nhận dạng mới hoàn chỉnh hơn, có khả năng xử lý với các nguồn tài liệu đa dạng hơn, giải quyết một cách triệt để hơn yêu cầu của bài toán nhận dạng đặt ra.

Những kết quả nghiên cứu mới này hoàn toàn có thể đưa vào áp dụng thực tế để tham gia giải quyết nhiệm vụ tìm kiếm và dự báo triển vọng khoáng sản trong xử lý phân tích tài liệu địa vật lý máy bay, một nguồn tài liệu đồ sộ và hết sức phong phú nhưng chưa được khai thác triệt để ở nước ta hiện nay.

Đây là một phương pháp phân tích tổ hợp, về nguyên tắc này có thể mở rộng phân tích cho các tài liệu địa vật lý khác nhau. Tuy nhiên, hiện chưa có kết quả phân tích trên các dạng tài liệu địa vật lý khác. Trong thời gian tới, chúng tôi sẽ tiếp tục nghiên cứu để có thể hoàn thiện và mở rộng hơn nữa phạm vi áp dụng của phương pháp.

Bài báo này được hoàn thành với sự hỗ trợ kinh phí của Đề tài NCKH cấp ĐHQGHN mã số QG.10-14.

VĂN LIỆU

- 1. Davis J.C., 1995.** Statistics and data analysis in geology. *New York*.
- 2. Green A.A., 1987.** Leveling airborne gamma-radiation data using between-channel correlation information, *Geophysics*, 52/11.
- 3. Paratov G.C., 1987.** Các phương pháp toán trong tìm kiếm thăm dò khoáng sản. *Nxb Nedra, Moskva (tiếng Nga)*.
- 4. Võ Thanh Quỳnh (Chủ biên), 1996.** Báo cáo Kết quả bay đo từ phổ gamma tỷ lệ 1/25.000 vùng Tuy Hoà. *Lưu trữ Địa chất. Hà Nội*.
- 5. Võ Thanh Quỳnh, 1996.** Enhancement of effectiveness of extracting and using the information in analysing and processing the airborne gamma-spectrometric data by using new methods. *Intern. Worksh. & Exh. on Geophysics. Hà Nội*.
- 6. Võ Thanh Quỳnh, 2007.** Một cách tiếp cận mới giải quyết bài toán nhận dạng trong xử lý, phân tích tài liệu địa vật lý. *TC Địa chất, A/302 : 76-80. Hà Nội*.
- 7. Võ Thanh Quỳnh, 2008.** Phương pháp đánh giá và phân loại cụm dị thường trong xử lý - phân tích tài liệu phổ gamma hàng không. *TC Địa chất, A/304 : 70-75. Hà Nội*.

8. Võ Thanh Quỳnh, 2008. Xây dựng một phương pháp nhận dạng mới trong xử lý tài liệu địa vật lý trên cơ sở vận dụng kết hợp các phương pháp phân tích khoảng cách khái quát và phân tích tần suất. *TC Địa chất, A/305 : 61-66. Hà Nội.*