

Dempster-Shafer 이론을 이용한 목표 유도형 지질자료 통합

박노욱*¹, 지광훈¹, 권병두²

¹한국지질자원연구원 지질자원정보센터

²서울대학교 지구과학교육과

요약

이 논문에서는 기존 알려진 통합 목표가 존재하지 않거나 적은 경우에 적용 가능한 목표 유도형 공간통합 방법으로 Dempster-Shafer 이론을 적용하였다. 이 이론은 확률의 하한과 상한 개념을 도입하여 통합 결과의 불확실성을 표현할 수 있으며, 탐사자료의 범위가 일치하지 않은 경우에도 효율적으로 통합이 가능한 장점이 있다. 광물자원 부존가능성도 작성의 사례연구를 수행한 결과, 기존 광산 주변에서 높은 부존 가능성값을 얻을 수 있었으며, 추후 객관적인 자료 표현 및 목표 유도형 통합에의 적용이 필요하다.

주요어 : 공간통합, Dempster-Shafer 이론, 지질자료

1. 서론

지구과학의 모든 연구분야는 지구를 대상으로 하기 때문에, 대부분의 정보가 특정 위치와 연관된다. 즉 지구과학에서는 현장에서 얻어진 자료를 기반으로 하여 자료 획득 위치에 따른 다양한 공간 정보와 이를 설명하기 위한 데이터 베이스 속성정보와 시간에 따른 변화 양상을 수치 모델링, 시뮬레이션, 역산등과 같은 방법을 통하여 정량적으로 다루고 해석하고 있다. 이러한 지구과학의 각종 자료들은 위치정보와 함께 제시되고 분석되어야 한다는 특징을 갖기 때문에, 대용량 정보를 효과적으로 분석 및 관리할 수 있는 기능을 제공하는 GIS와 직접적인 연계가 가능하다. 또한 지구과학이 종합적 시각과 자료의 통합적 분석을 요구한다는 점에서 GIS의 적용범위는 현재 제한적으로 적용되고 있는 지구과학 활용분야를 확대할 수 있는 많은 가능성을 보이고 있다.

외국에서는 1990년대 이후 GIS를 기반으로 한 지구과학의 공간정보를 통합, 해석하려는 많은 시도가 있어 왔으며, 이러한 연구는 자연과학과 GIS를 연계시키려는 주요 동향 중 하나로 간주되고 있다. 실제적으로 이러한 연구는 통합방법론의 개발과 통합 정보 결과

의 분석 처리 기술의 연구로 다시 세분화 할 수 있다.

현재 다양한 공간적 정보를 효과적으로 다루기 위한 수학적 자료 통합에 대한 연구가 시도되고 있다. 현재까지 개발된 통합방법론으로는 Weight of evidence 방법(Bonham-Carter et al., 1989), Dempster-Shafer의 이론을 적용한 방법(Moon, 1990), 퍼지이론을 적용한 방법(Wright and Bonham-Carter, 1996) 등이 있다. 이러한 통합방법을 적용하여 통합 결과로 제시되는 결정정보자료는 지구과학의 연구목적에 따라 중요한 해석 가능 정보를 제공하게 된다.

이러한 자료 통합 방법론은 가용할 수 있는 기존 탐사대상의 수에 따라 목표유도형과 자료유도형으로 나눌 수 있다. 목표유도형 자료 통합은 기존 알려진 탐사 대상의 수가 적을 때 적용이 가능한 방법에 해당되며, 자료유도형 자료 통합은 탐사 대상의 수가 많아서 입력 자료와의 통계적 분석이 가능할 때 적용이 가능한 방법이다. 기존 알려진 탐사 대상의 수가 적거나 아직 개발이 이루어지지 않은 지역을 대상으로 탐사를 수행할 경우에 목표유도형 자료 통합은 적용성이 크다. 기존 개발된 자료 통합 방법론 중에서 퍼지 이론과 Dempster-Shafer 이론은 이러한 목적에 이용이 가능한 방법이다. 특히 대상 영역에 대해서 탐사가 이루어지는 영역의 차이가 있는 경우가 대부분이기 때문에 이러한 경우에 적용성이 크다고 할 수 있다.

이 연구에서는 목표유도형 통합 방법으로 Dempster-Shafer 이론을 광물자원 탐사 가능성도 작성을 목적으로 적용하였다. 강원도 서백리 도폭을 대상으로 W/MO 광화대 탐지를 목적으로 사례연구를 수행하였다.

2. 적용 방법

Dempster(1967)에 의해 제창된 확률의 상한과 하한 이론을, Shafer(1976)가 확대 정리한 이 이론은 확률통계적 접근과는 달리 확신의 정도가 구간으로 표현되어 자료의 불확실성을 모사할 수 있다. 자료 융합 관점에서 이 이론은 자료 표현(data representation)을 위해 mass function 혹은 BPA(Basic Probability Assignment)를, 자료 결합(data combination)을 위해 Dempster's orthogonal sum을 적용한다.

이 이론에서는 우선 상호 배타적인 가설 n 개의 집합인 frame of discernment θ 를 설정하고, θ 의 모든 가능한 부분집합 즉 2^n 개의 가설에 대해 확신 정도를 할당하게 된다. θ 의 부분집합에 대한 개개 증거 혹은 자료들이 지지하는 확신 정도는 mass function 혹은 BPA로 표현되는데 식 (1)의 조건을 만족해야 한다.

$$\sum_{A \subset \theta} m(A) = 1, \quad m(\emptyset) = 0 \quad (1)$$

이렇게 표현되는 mass function에 기반하여 belief function(*Bel*)과 plausibility function (*Pls*)을 식 (2)로 정의한다.

$$\begin{aligned} Bel(A) &= \sum_{B \subset A} m(B) \\ Pls(A) &= 1 - Bel(\bar{A}) = \sum_{B \cap A = \emptyset} m(B) \end{aligned} \quad (2)$$

$$Bel(A) \leq P(A) \leq Pls(A)$$

여기서 A와 B는 가설을 각각 나타낸다.

식 (2)에서 belief function은 특정 가설에 부합되도록 할당되는 모든 mass function의 합으로 조건부 확률의 하한경계로, plausibility function은 상한경계로 해석할 수 있다. 일반적으로 $Bel(A) \neq Pls(A)$ 이기 때문에 실제 참값은 belief function과 plausibility function의 사이에 존재하게 된다. D-S 이론이 기존 확률 이론과 가장 큰 차이는 여기에 있으며, 신뢰구간으로 표현할 수 있는 $[Bel, Pls]$ 로 일종의 불확실성을 표현할 수 있다. 특히 탐사 자료에서 흔히 나타날 수 있는 탐사 영역의 불일치 구간에서는 belief function 값을 0으로, plausibility function 값을 1로 할당하여 불확실성이 1이 되도록 하여 모델링이 가능하다. 퍼지 이론에서도 탐사 영역의 불일치를 모델링 할 수 있으나 퍼지 소속 함수의 할당 과정에서 과추정 혹은 저추정 될 수 있다.

또 다른 확률 이론과의 차이점은 특정 가설에 대해 자료들이 지지하는 확신 정도가 표현되지 않는다면, 확신이나 불신을 할 수 없는 아무런 확신 정도가 형성되어 있지 않은 영역으로 인식한다. 예를 들어 $m(A)=0.6$ 일 경우 1에서 이 값을 빼준 0.4를 확률 이론에서는 A라는 가설에 반대되는 \bar{A} 에 대한 확신으로 표현하여 두 가지 상반된 가설에 대한 확신이 동시에 존재하는 것으로 간주한다. 그러나 D-S 이론에서는 만약 \bar{A} 에 대한 확신 정도를 자료로부터 알 수 없을 경우, 어떤 자료에 의해 확신되고 남은 값을 \bar{A} 에 할당하지 않는다. 대신 이를 아무런 확신이 형성되어 있지 않은 불확실성 영역 $A \cup \bar{A}$ 에 값을 할당하여 $m(A \cup \bar{A})=0.4$ 이 된다. 그러므로 모든 가설에서의 확률값의 합이 1이 되어야 하는 확률 이론과는 차이가 있다.

개별 자료들에 대한 mass function이 정의되면 이들의 결합에는 Dempster's orthogonal sum을 이용하여 결합된 mass function을 얻게 되고, 이로부터 결합 belief function과 plausibility function을 구하게 된다. 일반적으로 p개의 자료 혹은 특징의 mass function m_i

($i = 1, 2, \dots, p$)의 가설 A에 대한 결합은 식 (3)을 이용하게 되며, Dempster's orthogonal sum은 교환법칙이 성립하기 때문에 결합의 순서에 상관없이 동일한 결과를 나타낸다.

$$m_1 \oplus m_2 \oplus \dots \oplus m_p(A) = \frac{\sum_{B_1 \cap \dots \cap B_p = A} \prod_{i=1, \dots, p} m_i(B_i)}{1 - K} \quad (3)$$

$$K = \sum_{B_1 \cap \dots \cap B_p = \emptyset} \prod_{i=1, \dots, p} m_i(B_i)$$

여기서 A와 B는 가설을 각각 나타내며, B_i 는 i 번째 자료 혹은 특징에 대한 가설을 나타낸다.

3. 사례 연구

이 연구에서는 Dempster-Shafer 이론의 적용가능성을 살펴보기 위해 광물자원 부존가능성도 작성을 위한 사례연구를 수행하였다. 연구지역은 텅스텐 광산인 상동광산이 있는 지질도폭영으로 서벽리이다. 이 지역에 대해서는 기존 Choi et al.(2000)이 퍼지 이론을 적용한 예가 있다.

연구지역에는 작은 광산들이 많이 존재하지만, 텅스텐 광산인 상동 광산이 존재하고 있다. 이 연구에서는 탐사 자료로 지질도, W, Sn, Mo, Cu, Pb, Zn 및 Ag 등의 지화학 자료, U/Th ratio의 항공지구물리탐사 자료를 이용하였다. 탐사 지역의 차이로 항공지구물리탐사자료는 연구지역의 하단부에 자료가 존재하지 않는다. 이 연구에서는 지질도를 직접 사용하는 대신에 Choi et al. (2000)의 선행 연구를 따라서 석회암, 화강암, 기타 지역, 화강암을 중심으로한 관입 접촉 버퍼 존을 새로이 구성하여 사용하였다.

이 연구의 자료 통합 목표는 "연구 지역에 W/Mo 광화대가 존재한다"로 설정하였으며, 선택된 광상 모델은 스카른 형태이다. Dempster-Shafer 이론을 적용할 때 mass function의 할당이 가장 중요한 부분인데, 이 연구지역에서는 알려진 광산이 상동 광산 1개뿐이어서 목표 유도형으로 mass function을 할당하였다. 통합 목표와 스카른 형태의 광상 모델을 고려하여 실험적으로 belief 및 plausibility function을 할당하였다. 특히 스카른 형태의 광상 모델을 고려하여 석회암과 화강암의 접촉 지역의 거리를 고려하였으며, U/Th ratio의 자료가 없는 지역은 belief function은 0으로, plausibility function은 1로 할당하였다.

각 자료별 belief 및 plausibility function을 Dempster's orthogonal sum을 이용하여 통합된 결과는 Fig. 1과 같다.

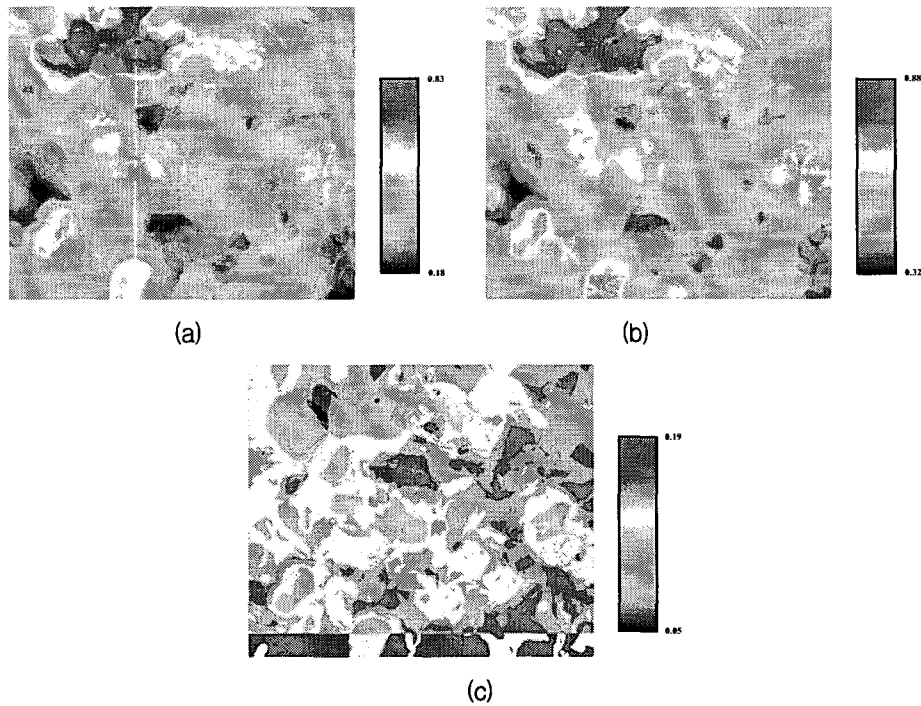


Fig. 1. 통합 결과: (a) belief function, (b) plausibility function, (c) ignorance function.

통합 목표에 대한 높은 지지 값은 연구지역의 북서쪽에 위치하고 있으며, 이 지역은 실제 알려진 상동 광산을 포함하고 있다. Fig. 1의 (b)에 제시된 plausibility function 값 역시 belief function과 유사한 양상을 나타내고 있다. 반면 ignorance function을 살펴보면, 불확실성은 탐사 목표에 대한 효율적인 정보를 제공한다고 볼 수 있는 W/Mo이 위치한 북서쪽에서 낮게 나타나고 있으며, 상대적으로 탐사 목표에 대한 높은 지지를 나타내지 않거나 탐사 자료가 존재하지 않는 연구지역 하단에서 높게 나타나고 있다.

3. 결론

이 연구에서는 Dempster-Shafer 이론의 적용가능성을 살펴보기 위해 광물자원 부존가능성도 작성을 위한 사례연구를 수행하였다. 적용 결과, 통합 목표인 상동 광산을 포함하는 주변에서 높은 부존 가능성 값을 얻을 수 있었다. 이 이론은 기본도내에서 탐사자료의 범위가 일치하지 않거나 누락된 경우에도 자료통합이 가능한 장점이 있다. 이는 전통적인 확률/통계적 접근방법에서는 이용되는 기본 자료의 범위가 일치해야 자료통합이 가

능하며 결과의 신뢰도가 증가된다. 반면 이 이론에서는 다소 주관적이기는 하나 자료내에서 정보가 없는 지역에서도 탐사목표에 대한 유의미한 mapping이 가능하다. 또한 이 이론에서는 확률론적 접근에서 가능한 결과의 불확실성 처리 및 해석보다는 구체적인 공간상의 불확실성을 처리할 수 있으므로 통합 결과의 해석에 중요한 근거를 제시할 수 있다. 그러나 이 이론의 경우 belief 및 plausibility function의 할당에 직접적으로 통합 결과가 영향을 받기 때문에 이에 대한 보다 많은 연구가 필요할 것으로 판단된다. 또한 목표 유도형 통합 이외에 자료 유도형 통합에의 적용 연구도 수행할 예정이다.

4. 참고문헌

- Bonham-Carter, G. F., Agterberg, F. P. and Wright, D. F., 1989, Weights of evidence modelling: a new approach to mapping mineral potential; in Agterberg, F. P. and Bonham-Carter, G. F.,(ed), Statistical applications in the earth sciences, Geological survey of canada Paper 89-9, p.171-183.
- Choi, S., Moon, W.M., and Choi, S.-G., 200, Fuzzy logic fusion of W-Mo exploration data from Seobyeong-ri, Korea, Geosciences Journal, 4(2), p.43-52.
- Dempster, A. P., 1967. Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics, 38(2): 325-339.
- Moon, W. M., 1990, Integration of geophysical and geological data using evidential belief function, IEEE Trans. Geosci. Remote Sensing, 28, p.711-720.
- Shafer, G., 1976. A mathematical theory of evidence, Princeton University Press.
- Rostirolla, S. P., Soares, P. C. and Chang, H. K., 1998, Bayesian and multivariate methods applied to favorability quantification in reconvavo basin and ribeira belt, brazil, Nonrenewable Resources, 7(1), p.7-23.