

Research Paper

Topic Model Analysis of Research Themes and Trends in the Journal of Economic and Environmental Geology

Taeyong Kim¹, Hyemin Park¹, Junyong Heo¹, Minjune Yang^{2*}

¹Division of Earth Environmental System Sciences, Pukyong National University, Busan 48514, South Korea

²Department of Earth and Environmental Sciences, Pukyong National University, Busan 48514, South Korea

*Corresponding author : minjune@pknu.ac.kr

ARTICLE INFORMATION

Manuscript received 17 April 2021

Received in revised form 14 June 2021

Manuscript accepted 16 June 2021

Available online 28 June 2021

DOI : <http://dx.doi.org/10.9719/EEG.2021.54.3.353>

Research Highlights

- Topic modeling based on Latent Dirichlet Allocation to identify research topics, historical trends, and network strengths in the Journal of Economic and Environmental Geology.
- 8 topics: 'Petrology and geochemistry', 'Hydrology and hydrogeology', 'Economic geology', 'Volcanology', 'Soil contaminant and remediation', 'General and structural geology', 'Geophysics and geophysical exploration', and 'Clay mineral'.
- Network strengths between each topics after 1994 has increased by approximately 3 times compared to those before 1994.

ABSTRACT

Since the mid-twentieth century, geology has gradually evolved as an interdisciplinary context in South Korea. The journal of Economic and Environmental Geology (EEG) has a long history of over 52 years and published interdisciplinary articles based on geology. In this study, we performed a literature review using topic modeling based on Latent Dirichlet Allocation (LDA), an unsupervised machine learning model, to identify geological topics, historical trends (classic topics and emerging topics), and association by analyzing titles, keywords, and abstracts of 2,571 publications in EEG during 1968-2020. The results showed that 8 topics ('petrology and geochemistry', 'hydrology and hydrogeology', 'economic geology', 'volcanology', 'soil contaminant and remediation', 'general and structural geology', 'geophysics and geophysical exploration', and 'clay mineral') were identified in the EEG. Before 1994, classic topics ('economic geology', 'volcanology', and 'general and structure geology') were dominant research trends. After 1994, emerging topics ('hydrology and hydrogeology', 'soil contaminant and remediation', 'clay mineral') have arisen, and its portion has gradually increased. The result of association analysis showed that EEG tends to be more comprehensive based on 'economic geology'. Our results provide understanding of how geological research topics branch out and merge with other fields using a useful literature review tool for geological research in South Korea.

Keywords: Latent Dirichlet Allocation, topic modeling, trend analysis, association analysis, economic and environmental geology

Citation: Kim, T., Park, H., Heo, J., Yang, M. (2021) Topic Model Analysis of Research Themes and Trends in the Journal of Economic and Environmental Geology. Korea Economic and Environmental Geology, v.54, p.353-364, doi:10.9719/EEG2021.54.3.353.

✉ Journal homepage: <http://www.kseeg.org/main.html>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided original work is properly cited. pISSN 1225-7281; eISSN 2288-7962/©2021 The KSEEG. Printed by Hanrimwon Publishing Company. All rights reserved.

연구논문

기계학습 기반 토픽모델링을 이용한 학술지 “자원환경지질”의 연구주제 분류 및 연구동향 분석

김태용¹ · 박혜민¹ · 허준용¹ · 양민준^{2*}

¹부경대학교 지구환경시스템과학부

²부경대학교 지구환경과학과

*책임저자 : minjune@pknu.ac.kr

요약

국내 지질학의 연구 분야는 20세기 중반 이후부터 꾸준히 발전되어왔다. 학술지 “자원환경지질”은 국내 지질학을 대표하는 역사가 긴 학술지로 지질학을 바탕으로 하는 융복합연구 논문이 게재되고 있다. 본 연구는 학술지 “자원환경지질”에 게재된 논문을 대상으로 문헌 고찰(literature review)을 수행하여 지질학의 역사와 발전에 대해 논의하고자 한다. 1968년부터 2020년까지 총 2,571편의 논문 제목, 주제어, 다국어 초록을 수집하였으며, Latent Dirichlet Allocation (LDA) 기반 토픽모델링을 실시하여 연구주제를 분류하고 연구 동향과 주제간 연관성을 확인하였다. 학술지 “자원환경지질”은 총 8개의 연구주제(암석학 및 지구화학, 수문학 및 수리지질학, 광상학, 화산학, 토양오염 및 복원학, 기초지질 및 구조지질학, 지구물리 및 물리탐사, 점토광물)로 분류할 수 있었다. 1994년 이전에는 ‘광상학’, ‘화산학’, ‘기초지질 및 구조지질학’의 연구주제들이 활발하게 연구되었으며, 이후 ‘수문학 및 수리지질학’, ‘토양오염 및 복원학’, ‘지구물리 및 물리탐사’, ‘점토광물’의 연구주제들이 성행하였다. 연관성분석(network analysis)결과, 학술지 “자원환경지질”은 ‘광상학’을 기반으로 융복합적 연구 논문들이 게재되었다는 것을 확인하였다. 본 연구의 결과는 지질학을 다루는 연구자들에게 문헌 고찰의 새로운 방법론을 제시하여 지질학의 역사에 대한 이해를 제공했음에 의의가 있다.

주요어 : 토픽 모델링, 잠재 디리클레 할당, 연구 동향 분석, 연관성 분석, 자원환경지질

1. 서론

지질학은 층서학, 퇴적학, 암석학, 광상학 등과 같이 고체 지구를 탐구하는 기초과학분야부터 지질공학, 수리지질학, 환경지질학 등 학문 간의 융합을 바탕으로 연구되는 응용과학분야까지 폭넓게 발전해왔다(Lee, 2020). 현대지질학의 범주는 서로 다른 학문 분야와 함께 연구하는 학제간 연구를 바탕으로 넓어지고 있으며, 학문간 경계는 흐려지고 있는 추세이다(Kim *et al.*, 2004). 국내 지질학을 대표하는 학술지는 암석학, 광물학, 지구물리학, 토양 및 수리지질학, 퇴적학, 지질공학 등 다양한 지질학적 연구 분야들을 기반으로 하고 있다. 특히, 학술지 “자원환경지질”은 1968년 “광산지질”로 처음 발행되었으며 발행 초기 자원 탐사 및 개발과 관련된 연구가 주를 이루었다. 최근 전 지구적 시스템의 관점에서 환경 문제가 대두됨에 따라 관련 연구들이 활발하게 진행되었으며, 1994년 학술지명을 “자원환경지질”로 변경하여 환경 분야 및 다양한 지질학 분야의 논문들이 게재되고 폭넓은 지질학적 연구를 다루고 있다.

문헌고찰(literature review)은 특정 연구 분야의 모든 연

구를 체계적(systematic)으로 분석하여 현재까지의 연구 동향과 향후 연구방향을 확인하는 중요한 단계이다(Kim *et al.*, 2020). 오랜 역사를 가진 학술지의 문헌고찰은 학술지가 추구하는 연구 분야 및 연구 동향을 확인할 수 있다는 점에서 연구의 목적과 의의를 찾을 수 있다. 기존의 문헌고찰 연구들은 연구자들이 직접 모든 문헌을 검토하면서 연구를 수행하였지만, 최근 연구들은 기계학습(machine learning)과 자연어처리(nature language processing) 분야에서 사용되고 있는 토픽모델링(topic modeling)을 이용하여 연구주제를 추론하고 그에 따른 동향 분석을 활발하게 진행되었다(Shin, 2020). 토픽모델링이란 문서 집합 내 주제(topic)를 발견하기 위한 통계 모델로, 그중 가장 널리 사용되고 있는 Latent Dirichlet allocation (LDA)는 주제별 단어들의 분포와 문헌별 주제의 분포를 추론할 수 있어 기존의 방법보다 빠르게 주제를 도출할 수 있기 때문에 문헌고찰 연구에 많이 사용되어왔다(Blei, 2012). 특히, LDA를 이용한 문헌고찰 연구는 기존에 사용하던 분석방법에 비해 빠르고, 간단하며, 연구자의 주관이 배제되기 때문에 현재까지 많은 분야에 사용되고 있다(Blei *et al.*, 2003).

문헌 고찰 연구 중 토픽모델링을 사용한 국외 연구로 Gatti *et al.* (2015)는 총 37개의 경영과학분야 학술지들을 대상으로 LDA를 사용하여 분야 내에 내재된 연구주제들을 추론하고 시간에 따른 변화를 파악하는 연구를 진행하였다. Sun and Yin (2017)은 토픽모델링을 교통분야에 접목해 총 22개의 학술지 내에 내재되어 있는 연구주제를 추론하고 그에 따른 연구 동향분석을 실시하였으며, 추가로 공동 출현 단어간 네트워크 분석(co-occurrence network analysis)을 통하여 단어 간 연관성(association)을 보여주었다. Xiong *et al.* (2019)는 총 43개의 제조분야의 학술지 내에 내재되어 있는 50개의 연구주제를 분류하였고, 연구 주제별 출현 확률을 바탕으로 연구 동향을 분석하였다. 또한, 인기도 점수(popularity score)와 동향 점수(trend score)를 사용하여 연구 주제별 인기도를 파악하고, 시간에 따른 연구 관심사의 변동을 확인하였다.

국내 또한 학술지 및 지질학의 특정 하위 분야와 관련된 연구가 활발히 진행되었다. Sang and Chung (1990)은 미국지질학협회(American Geological Institute; AGI)에서 제공하는 지질학의 분류체계를 이용하여 “대한지질학회”와 “대한광산지질학회” 그리고 연구소에서 발행한 논문을 대상으로 연구 동향을 분석한 결과, 1990년 이전에는 광물 자원조사 및 에너지 자원조사에 연구가 치중되어 있으며 기초분야 및 응용분야 연구의 저조함을 시사하였다. Heo and Kim (2018)은 학술지 “자원환경지질”에 게재된 논문을 대상으로 분석을 진행하였으며, 연구 분야의 논문 수를 정량적으로 제시하고 그에 따른 동향 및 해석을 하였다. 기존의 선행 연구들은 연구자가 대용량의 문헌을 직접 읽고 주제를 분류하여 해석하기 때문에 지질학의 전체적인 연구 동향을 확인하기에는 시간과 비용에 많은 제약이 있었다. 그러므로, 토픽모델링과 병행하여 연구를 진행한다면 분석에 소요되는 시간과 비용의 한계를 극복하는 연구 동향분석이 가능하다.

따라서, 본 연구는 학술지 “자원환경지질”을 대상으로 LDA 모델을 적용하여 학술지에 기고된 논문들의 연구분야를 규명하고 분류하여 학술지의 연구 범위를 확인하고, 연구 분야별 동향을 분석하여 추세를 파악하고, 연관성 분석(Network analysis)을 통해 연구 분야간 얼마나 관계가 있는지 정량적으로 평가하는 새로운 연구 동향 분석방법을 소개하고자 한다.

2. 연구방법

2.1. 데이터 수집 및 전처리

본 연구에서 사용된 데이터는 국내 학술 DB 검색 사이트인 RISS (Research Information Sharing Service)에서

학술지 “자원환경지질”에 대하여 검색하였으며, 창간 년도인 1968년부터 2020년 53권 1호까지 총 52년간 출판한 논문 2,571편을 수집하여 사용하였다. “자원환경지질”의 전신인 “광산지질”은 1968년부터 1993년까지 연간 4회로 출간하였으며, 연간 평균 논문 수는 36.31편이었다. 1994년 이후로, 학술지 “광산지질”에서 “자원환경지질”로 이름을 바꾸면서 연간 6회로 출간하였으며, 연간 평균 논문 수는 60.25편으로 출간된 누적 논문의 수는 지속적으로 증가하는 추세(linear slope =53.6, $R^2 = 0.98$)를 보여주었다(Fig. 1).

데이터 전처리는 Fig. 2와 같이 수행하였다. 먼저, “자원환경지질”에서 발행된 논문들의 제목, 주제어, 초록에 대한 정보들을 수집하였으며, 발행 연도에 대한 메타데이터는 따로 처리하여 시간에 따른 주제의 분포를 확인하기 위해 사용하였다. 다음으로, 국문의 경우 영문의 표기법 및 연구자마다 사용하는 용어가 상이하기 때문에 모델의 성능이 떨어지는 경향을 보여 한계점을 보완하기 위해 영문 단어만 사용하였다. 준비된 데이터들을 띄어쓰기 및 구두점(punctuation mark)을 기준으로 문장을 단어로 변환해주는 토큰화(tokenization) 과정을 실시하였으며, 자주 등장하지만 분석에 필요하지 않는 단어들을 제거하기 위해 불용어처리를 실시하였다. 마지막으로, 단어의 시제 및 수 일치 등 같은 의미의 단어들을 원형으로 통일하여 모델의 성능을 높이기 위해 어간추출(stemming)을 사용하여 단어의 어미를 자르고 원형만 사용하는 방법을 적용하였다.

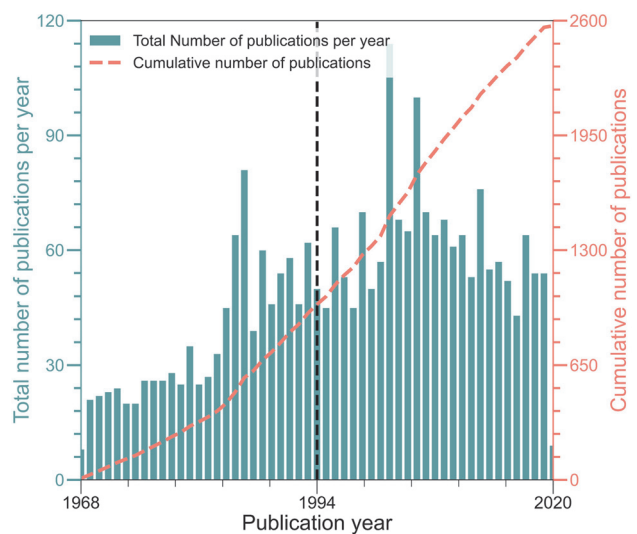


Fig. 1. Total number of publications per year and cumulative number of publications from 1968 to 2020 published in Economic and Environmental Geology (EEG). The turn point year (1994) when the journal was renamed is shown with a black dashed line.

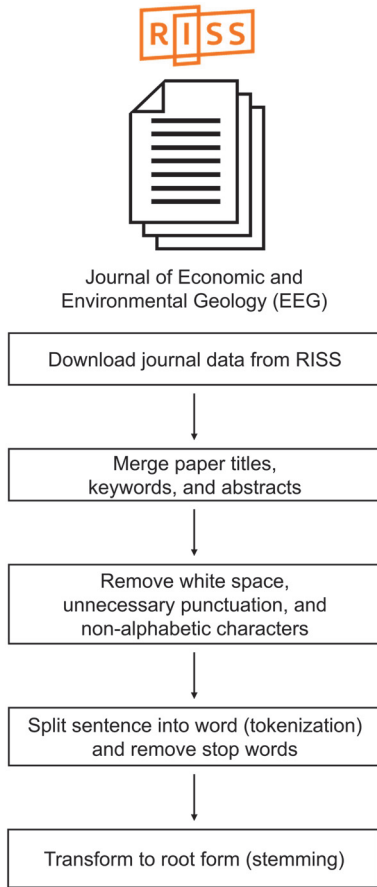


Fig. 2. Flowchart of data processing for topic modeling from Journal of Economic and Environmental Geology.

2.2. 분석방법

2.2.1. 잠재 디리클레 할당(Latent Dirichlet Allocation)

Latent Dirichlet Allocation (LDA)는 확률적 생성 모델로 관찰된 단어를 바탕으로 주제에 따른 단어의 분포를 추정하고, 문서별 주제의 분포를 추정한다(Blei *et al.*, 2003). LDA의 그래프 모델과 변수의 의미는 Fig. 3와 Table 1에서 확인할 수 있다. K 번째 주제에 대한 단어의 다항분포인 β_k 는 η 를 모수로 가지는 디리클레 분포 $Dirichlet(\eta)$ 에 의해 샘플링 된다. 문서 d 에 대한 주제의 다항분포인 θ_d 는 α 를 모수로 가지는 디리클레 분포 $Dirichlet(\alpha)$ 에 의해 샘플링 된다. 그런 다음, 문서에 대한 주제의 다항분포인 $Multi_K(\theta_d)$ 로부터 특정 문서에 대한 특정 단어는 특정 주제의 번호를 할당 받는 $z_{d,n}$ 로 표기된다. 마지막으로, 연구자가 관찰할 수 있는 단어 $w_{d,n}$ 는 구해진 단어에 대한 주제의 번호 $z_{d,n}$ 와 β_k 의 다항분포인 $Multi_V(\beta_{z_{d,n}})$ 따라 샘플링 된다. 이러한 과정은 식 (1)을 통해 표현 가능하다.

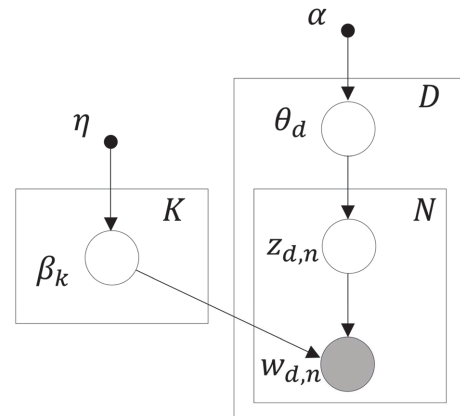


Fig. 3. Schematics of Latent Dirichlet Allocation (LDA) graphical model. Observed node (words of the documents) is shaded as grey. Rectangles are plate notation, which denotes replication.

Table 1. Notation of variables and parameters used in this study

Notation	Value	Description
<i>Indices</i>		
d		Index of documents
k		Index of topics
t		Index of years
<i>In LDA</i>		
α	0.01	Dirichlet prior on the per-document topic distribution (hyperparameter)
η	0.1	Dirichlet prior on the per-topic word distribution (hyperparameter)
θ_d		Topic distribution of document d
$\theta_{d,k}$		Distribution of topic k in document d
β_k		Word distribution of topic k
$z_{d,n}$		Topic assignment from θ_d in word n
$w_{d,n}$		Observed word n in document d
K	10	Number of topics
D	2,571	Number of documents
V	17,242	Number of unique words in the LDA vocabulary
N	314,248	Total number of words in all documents

$$\begin{aligned}
 & p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}) \\
 &= \prod_{i=1}^K p(\beta_i | \alpha) \sum_{d=1}^D p(\theta_d | \eta) \left(\sum_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta_{1:K}, z_{d,n}) \right)
 \end{aligned} \tag{1}$$

여기서, 매개변수 α 와 η 는 주제와 단어의 분포 모형에

관한 변수이고, 값이 작을수록 희소하게 나타난다. Griffiths *et al.* (2004)는 $\alpha = 50/K$, $\eta = 0.1$ 로 설정할 것을 제안하였으나, 한 문서 내에서 다양한 주제들이 균등하게 분포하여 주제의 대표성을 띠기 힘든 문제점이 존재한다. 따라서, 본 연구에서는 연구주제 간 차별성을 주기 위해 $\alpha=0.01$ 로 설정하여 희소성을 띠는 주제분포를 생성하였다. 또한, 모델의 반복 수행을 통해 변수들이 안정적인 값으로 수렴하는 깃스 샘플링(gibbs sampling) 방법을 이용하였으며(Porteous *et al.*, 2008), 100,000회 반복 샘플링 한 후 안정된 결과값을 사용하였다. 본 연구에서는 파이선 환경의 오픈소스 토픽모델링 라이브러리인 Tomotopy (<https://bab2min.github.io/tomotopy/v0.9.1/kr/>)의 LDA 모델을 이용하여 분석을 하였다.

LDA의 결과로 추론된 주제별 단어의 분포와 문서별 주제의 분포를 이용하여 연구 주제를 동정하고 시간에 따른 각 주제의 동향을 파악할 수 있다. 특정주제에 대한 연도별 분포는 식 (2)으로 계산하였다.

$$\theta_k^t = \frac{\sum_{d=1}^D \theta_{d,k}^t}{M^t} \quad (2)$$

여기서 θ_k^t 는 t 번째 연도에 해당하는 k 번째 주제에 대한 분포를 의미하고, $\theta_{d,k}^t$ 는 t 번째 연도의 d 번째 문서에서 k 번째 주제에 대한 분포를 의미한다. M^t 는 t 번째 연도에서의 총 문서의 수를 의미하고, 한 문서 내 모든 주제에 대한 확률 분포의 합은 1이다.

시간에 따른 주제별 동향을 확인하기 위해 단순회귀분석(simple linear regression analysis)을 실시하였다. 단순회귀분석의 독립변수를 학술지 “자원환경지질”의 출판 연도로 설정하였으며, 종속변수를 연구주제의 연도별 비중으로 설정하였다. 회귀계수의 유의성을 확인하기 위해 양측 t 검정(paired student t-test)을 실시하였다(Hsu and Lachenbruch, 2005). 유의확률(p-value)이 0.05 미만일 때, 회귀계수가 양수인 경우 최근에 연구가 활발하게 진행되는 것을 의미하는 Emerging topic으로 분류하였으며, 음수인 경우 학술지 발행 초창기에 연구가 활발하게 진행된 것을 의미하는 Classic topic으로 분류하였다(Griffiths *et al.*, 2004). 유의확률이 0.05 이상일 경우 회귀계수는 통계적으로 유의하지 않기 때문에 해당 연구주제는 동향을 파악할 때 제외하였다.

2.2.2. 동시 출현 확률에 따른 연관성 분석

LDA의 결과인 문서별 주제의 분포를 이용하여 주제별 연관성을 파악하기 위해 연관성 분석을 실시하였다. 동시 출현 확률에 따른 네트워크를 생성하기위해 식 (3)를

이용하여 문서에 따른 주제의 분포로 구성되어 있는 행렬 M 을 계산하였다.

$$M_{\text{period}} = \begin{bmatrix} \theta_{1,1} & \cdots & \theta_{1,n} \\ \vdots & \ddots & \vdots \\ \theta_{m,1} & \cdots & \theta_{m,n} \end{bmatrix} \text{ where period} = \begin{cases} \text{before 1994 (year} \leq 1994) \\ \text{after 1994 (year} > 1994) \end{cases} \quad (3)$$

여기서, $\theta_{m,n}$ 은 m 번째 문서에서 n 번째 주제의 분포를 뜻한다. 또한, 학술지 이름의 변경 시점인 1994년을 기준으로 두 시대의 네트워크를 비교하기 위해 M_{period} 을 이용하여 1994년 이전의 문헌들로 구성된 행렬 $M_{\text{before 1994}}$ 과 이후의 문헌들로 구성된 행렬 $M_{\text{after 1994}}$ 을 이용하여 두 그룹 간 비교 분석하였다. 또한, 네트워크를 생성하기위해 식 (4)을 이용하여 주제별 동시 출현 확률을 곱한 연결 강도로 구성된 행렬 $\text{network strength}_{\text{period}}$ 를 정의하였다.

$$\text{network strength}_{\text{period}} = M_{\text{period}}^T \times M_{\text{period}} \quad (4)$$

여기서, $\text{network strength}_{\text{period}}$ 는 $K \times K$ 로 구성된 정방행렬로, 대각성분을 기준으로 대칭이 되는 구조이다. M_{period}^T 는 행렬 M_{period} 의 전치행렬을 의미하며 두 행렬 간 곱을 통해 주제별 동시 출현 확률을 바탕으로 계산된 연결 강도를 사용하였다. 또한, 두 시대별 비교 분석을 하기 위해 1994년 이전의 네트워크 행렬과 1994년 이후의 네트워크 행렬을 계산하였다.

식 (4)을 바탕으로 계산된 행렬 $\text{network strength}_{\text{period}}$ 을 네트워크 분석 툴인 Cytoscape를 사용하여 주제별 연관성 네트워크를 시각화하였다. 네트워크는 노드(node)와 엣지(edge)로 표현되며, 노드는 원(circle)으로 표현되는 특정 개체를 뜻하며, 엣지는 선(line)으로 표현되고 특정 개체 간 관계를 파악하는 데 사용된다. 본 연구는 1994년을 기준으로 연구주제 간 연관성을 파악하였으며, 노드는 각 연구주제를 의미하며, 엣지는 동시 출현 확률을 바탕으로 계산된 연결 강도를 사용하였다. 본 연구에서 각 노드 간 연결 강도(network strength)를 바탕으로 주제 간 연관성을 계산하였으며, 특정 노드에 대한 모든 노드의 평균 연결 강도(mean network strength)를 이용하여 특정 연구주제에 대한 모든 연구주제 간 연관성을 계산하였다. 또한, 노드의 사이즈와 방향성은 고려하지 않는 네트워크로 시각화하였다.

Table 2. Topics and top-10 related words per topics

Topic (ratio)	Topic label	Keyword (top-10)
Topic 1 (9.46%)	Petrology and geochemistry	granite element, foliated, rock, ppm, uranium, asbestos, content, geochemical, area
Topic 2 (12.08%)	Hydrology and hydrogeology	water, groundwater, sample, ph, reaction, metal, concentration, river, sample, stream
Topic 3 (25.08%)	Economic geology	stage, fluid, deposit, vein, ore, mineral, inclusion, skarn, temperature, quartz
Topic 4 (10.58%)	Volcanology	granite, rock, age, volcanic, tuff, formation, lava, island, alkali, basalt
Topic 5 (10.35%)	Soil contaminant and remediation	soil, contaminant, heavy, metal, groundwater, water, weather, removal, cd, pb
Topic 6 (11.72%)	General and structural geology	basin, formation, fault, okcheon, age, fold, structure, group, tectonic, direction
Topic 7 (13.02%)	Geophysics and geophysical exploration	fault, landslide, gravity, magnet, seismic, model, velocity, map, slope, anomaly
Topic 8 (7.72%)	Clay mineral	clay, layer, kaolin, kaolinite, medicine, mineral, deposit, rock, zone, heat

3. 결과 및 토의

3.1. 토픽모델링의 결과에 따른 연구주제 별 분류

토픽모델링은 LDA 모델이 주어진 자료에 대해 적절한 주제의 개수로 나누었는지 연구자가 판단하고, 결과를 효과적으로 해석하는 것이다. 사전 실험으로 주제의 개수를 5, 10, 15, 20, 30, 50으로 설정하여 모델링을 실시하였으며, 기존에 사용하는 지질학적 연구분류체계와 각 주제에 대한 단어의 연관성을 고려하였을 때 주제의 개수를 10으로 설정하는 것이 가장 적합한 것으로 판단되었다. 그중에서도 주제 내 단어 간 연관성을 파악하기 힘든 주제를 제외하고 유사한 주제를 같은 주제라 간주하여 총 8개의 연구 주제로 분류하였다. 주제별 단어들은 해당 주제에서 비중이 큰 단어 10개를 나열하였으며, 위 단어들을 바탕으로 연구 주제를 명명하였다(Table 2). 모델링 분석 결과를 바탕으로, ‘암석학 및 지구화학’, ‘수문학 및 수리지질학’, ‘광상학’, ‘화산학’, ‘토양오염 및 복원학’, ‘기초지질 및 구조지질학’, ‘지구물리 및 물리탐사학’, ‘점토광물’의 연구주제로 분류할 수 있었다. 전체 논문 중 ‘광상학’이 25.1%로 가장 높은 비중을 차지하였으며, ‘점토광물’이 8.1%로 가장 낮은 비중을 차지하였다. 이는 학술지 “자원환경지질”이 자원과 관련된 연구 분야들을 주로 다루고 있기 때문이라 해석된다. 또한, ‘점토광물’은 카올린, 카올리나이트(kaolin and kaolinite)와 같이 포괄적인 지질학적 주제가 아닌 점토광물을 대표하는 단어들을 포함하고 점토광물을 대상으로 하는 특정한 연구주제들로 구성되어 있기 때문에 학술지 내에서 국소적으로 분포하는 것으로 확인된다.

위 결과를 바탕으로, 각 연구 주제별 세부적인 내용을 확인하기 위해 주제별 논문의 출현 확률을 계산하였고 내림차순으로 논문을 정렬하여 각 주제를 대표하는 논문들을 선정하여 해석하였다. 첫째, ‘암석학 및 지구화학’

(Topic 1)은 특정 지역 내 지층에서 채취한 시료를 대상으로 암석학, 광물학 및 지구화학적으로 특성을 규명하는 논문들이 주를 이루고 있었다(Lee *et al.*, 1997; Lee and Lee, 1997; Na *et al.*, 1997; Yin *et al.*, 2013). 특히, 엽리상 화강암류를 대상으로 하는 많은 연구들이 존재하였으며(Kim and Kim, 1990; Kim *et al.*, 1994; Na *et al.*, 1997) 동위원소를 이용한 연구들도 활발히 진행되었다(Chon *et al.*, 1991; Kim, 1986; Park, 1997; Shin and Kagami, 1996). 둘째, ‘수문학 및 수리지질학’(Topic 2)은 지하수에 존재하는 오염물질의 제거기술과 광산수계의 오염물질 특성을 규명하는 논문들이 큰 비중을 차지하였다(Lee *et al.*, 2001; Yang *et al.*, 2008). 특히, 미생물의 활동에 의한 지하수 내 영향을 평가한 논문들이 존재하였으며(Kim *et al.*, 2014; Lee *et al.*, 2014), 지하수 내 수질특성 및 지하수 유동에 대해 다양한 방법으로 해석하였다(Hamm *et al.*, 2005; Ko *et al.*, 2005; Hyun *et al.*, 2006). 셋째, ‘광상학’(Topic 3)은 광상에서 산출되고 있는 광물 및 광상의 성인과 관련된 연구들이 주를 이루었다(Chi, 1975; Woo, 1988; Yang and Chi, 1986). 특히, 유체포유물과 금-은 광상을 대상으로 한 연구들이 진행되었으며(Choi *et al.*, 1988; Yoo *et al.*, 2002) 스카른 광상에 대한 논의가 이루어졌다(Han and Choi, 2012; Yun, 1983). 넷째, ‘화산학’(Topic 4)은 화산활동 및 화산암류와 관련하여 암석학, 지구화학, 지구연대학적 관점으로 해석한 논문들의 집합으로 나타났다(Hwang *et al.*, 2017; Lee, 2014; Yun and Hwang, 1988). 특히, ‘화산학’에서 ‘제주도’, ‘울릉도’, ‘우도’를 대상으로 하는 연구들이 활발히 진행된 것을 확인하였다(Hwang, 1993; Kim and Lee, 1983; Koh *et al.*, 2008; Koh and Park, 2010a, 2010b) 다섯째, ‘토양 오염 및 복원학’(Topic 5)은 토양 및 폐광산 내 존재하는 오염물질을 계면활성제 및 흡착제를 사용하여 정화하는 방법에 대한 논의가 이루어졌다. 위 주제에

서는 ‘독성원소’, ‘중금속’, ‘흡착’, ‘토양 오염’ 등의 주제를 대표하는 연구들이 진행되었다(Ahn and Lee, 2018; Ben and Yvette, 2000; Jeon *et al.*, 2006; Lee and Lee, 2012; Lee and Chon, 2004; Oh *et al.*, 2003). 또한, ‘수문학 및 수리지질학’과 같이 오염물질을 대상으로 복원기술을 이용한 정화 측면에서 유사한 경향의 논문들이 존재하였다. 여섯째, ‘기초지질 및 구조지질학’(Topic 6)은 야의 한반도를 대상으로 전반적인 기초지질조사 및 층서와 지질구조에 관한 논의가 진행되었다. 특히, 한반도 내 분지, 층군, 육괴를 대상으로 층서, 지질구조, 분지의 진화와 연관된 주제들이 논의되었다(Kim *et al.*, 1984; Kim, 1973, 1971; Kwon, 2012; Kwon *et al.*, 2019; Lee *et al.*, 1992; Na *et al.*, 1982; Son and Kim, 1994). 일곱째, ‘지구물리 및 물리탐사학’(Topic 7)은 특정 지역에 대한 중력탐사(Hwang *et al.*, 2003; Kim *et al.*, 2007; Min and Cho, 1995), 지진활동에 관한 연구(Kim and Bae, 2006; Kim *et al.*, 2006; Wee and Kim, 2008)들이 주를 이루었다. 또한, 탄성과 탐사를 이용한 연구들은 지질조사(Oh and Suh, 2005) 및 사면 안정성 평가(Um *et al.*, 2005) 등 다양한 분야에 적용되어 활발하게 연구되었다. 여덟째, ‘점토광물’(Topic 8)은 주로 점토광물의 특성규명과 그에 따른 다양한 연구들이 주를 이루었다(Kang, 2012; Kim, 1992; Lee *et al.*, 2011). 또한, 비금속광물을 약재로 활용하기 위한 연구들도 존재하였다(Bak and Kim, 2015; Kim and Park, 2015).

LDA로 분석한 학술지 “자원환경지질” 내 연구주제를 살펴본 결과 1) 지구화학, 암석학, 지구연대학적 관점으로 특정 지역 내 특성 규명 연구 (‘암석학 및 지구화학’, ‘화산학’), 2) 폐광산 및 부지 내 오염물질의 특성을 규명하고, 그에 따른 오염물질 정화 관련 연구 (‘수문학 및 수리지질학’, ‘토양오염 및 복원학’), 3) 한반도 및 해외 광상에 대한 성인 및 광화작용 연구 (‘광상학’), 4) 한반도를 대상으로 지질조사 및 구조적 해석 (‘기초지질 및 구조지질학’), 5) 중력탐사, 비저항탐사 등 지구물리학적 관점에서 지반조사, 해저탐사, 사면안정성 평가 등 다양한

분야에 대한 적용 (‘지구물리 및 물리탐사학’), 6) 점토광물의 특성을 규명하여 다양한 분야에 적용하려는 연구 (‘점토광물’) 로 요약이 가능하다. 특히, ‘암석학 및 지구화학’과 ‘지구물리 및 물리탐사’는 다양한 분야에서 적용되고 있어 다른 연구 주제들 속에서도 영향을 끼치고 있다는 것을 확인하였다. 이처럼 학술지 “자원환경지질”은 다양한 지질학적 요인의 집합체로 해석할 수 있다.

3.2. 시간에 따른 연구주제의 확률 분포 변화

시간에 따른 학술지 “자원환경지질” 내 연구주제의 동향을 확인하기 위해 t검정을 실시하여 Emerging topic과 Classic topic으로 분류하였으며, 그에 따른 연구주제별 분포의 변화를 Fig. 4에 나타내었다. t검정 결과, Classic topic은 ‘광상학’(slope = -4.67×10^{-3}), ‘화산학’(slope = -2.97×10^{-3}), ‘기초지질 및 구조지질학’(slope = -2.43×10^{-3})의 연구주제들이 나타났다. Emerging topic은 ‘수문학 및 수리지질학’(slope = 4.91×10^{-3}), ‘토양오염 및 복원학’ (slope = 2.45×10^{-3}), ‘지구물리 및 물리탐사’(slope = 1.71×10^{-3}), ‘점토광물’(slope = 1.17×10^{-3})의 연구주제들로 분류되었다. 이러한 경향은 학술지 “광산지질”에서 주로 다루던 연구주제들이 학술지 “자원환경지질”으로 변경됨에 따라 다양한 지질학적 연구주제들로 연구범위를 넓혀 학술지가 추구하는 방향이 바뀌었다는 것으로 판단된다. 8개의 연구 주제 중 ‘암석학 및 지구화학’은 시간에 따른 확률 분포의 변화가 통계적으로 유의하지 않았다(slope = -1.59×10^{-4} , p-value > 0.05). 이는 학술지 “자원환경지질” 내에서 ‘암석학 및 지구화학’ 주제는 과거부터 현재까지 변화 없이 꾸준히 연구된 것으로 해석된다.

시간에 따른 연구주제의 확률 분포를 Fig. 5에 나타내었다. 학회지명의 변경 시점인 1994년을 기준으로 ‘광상학’, ‘화산학’, ‘기초지질 및 구조지질학’의 시대별 평균 확률 분포는 각각 32.3%, 15.9%, 14.6%에서 변경 시점 이후 18.2%, 5.5%, 8.9%로 감소하였다. 반면에, 1994년을 기준으로 이전의 ‘수문학 및 수리지질학’, ‘토양오염

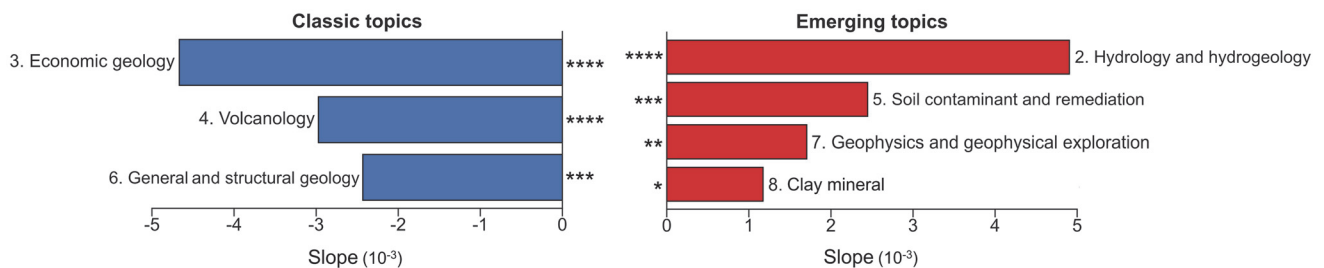


Fig. 4. Significant classic and emerging topics sorted by the slope of the simple linear regression model (* represents p = 0.05, ** represents p = 0.01, *** represents p = 0.001, and **** represents p = 0.0001).

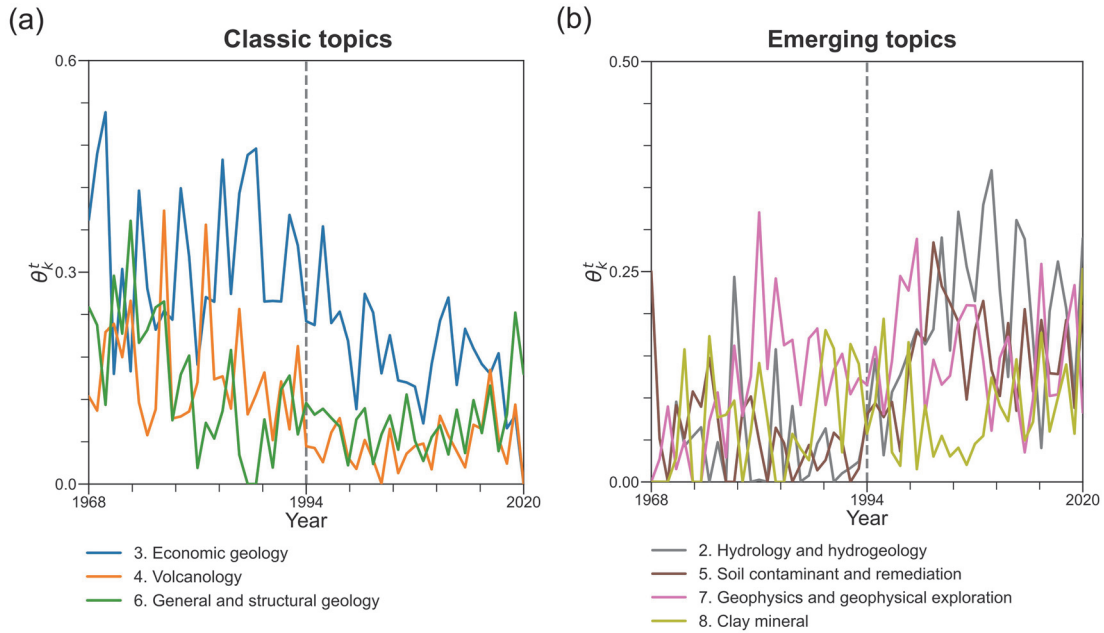


Fig. 5. Time series evolution of (a) classic topics and (b) emerging topics. Turn point, 1994 (dashed grey vertical line), serves as the dividing year between the two time periods.

및 복원학’, ‘지구물리 및 물리탐사’, ‘점토광물’의 시대별 평균 확률 분포는 각각 4.2%, 5.6%, 10.9%, 6.9%에서 이후 20.2%, 15.2%, 15.2%, 8.6%로 증가하였다. 특히, ‘수문학 및 수리지질학’과 ‘토양오염 및 복원학’의 연구 논문 수가 1994년 이후 크게 증가하였으며, 이는 경제개발을 위해 자원 개발이 필요하던 시기에 지질조사와 광상과 관련된 연구들이 활발하게 이루어졌으나, 1990년부터 경제 성장에 따른 환경 오염 문제 등이 야기되어 환경분야의 관심이 상대적으로 커졌기 때문에 이러한 경향을 띄는 것으로 확인된다(Heo and Kim, 2018).

총 52년을 대상으로 연구주제별 평균 확률 분포를 확

인한 결과, ‘광상학’(25.1%)과 ‘점토광물’(7.7%)을 제외한 나머지 연구주제(‘암석학 및 지구과학’, ‘수문학 및 수리지질학’, ‘화산학’, ‘토양오염 및 복원학’, ‘기초지질 및 구조지질학’, ‘지구물리 및 물리탐사학’)의 평균 확률 분포는 11.2%로 확인되었다. 이는 ‘광상학’을 기반으로 다양한 지질학적 분야들이 전반적으로 균형 있게 연구되었다는 것으로 해석된다.

3.3. 동시 출현 확률에 따른 연관성 분석 결과

동시 출현 확률에 따른 연관성 분석결과를 Fig. 6에 나타내었으며, 1994년을 기준으로 이전의 연구주제 간 연

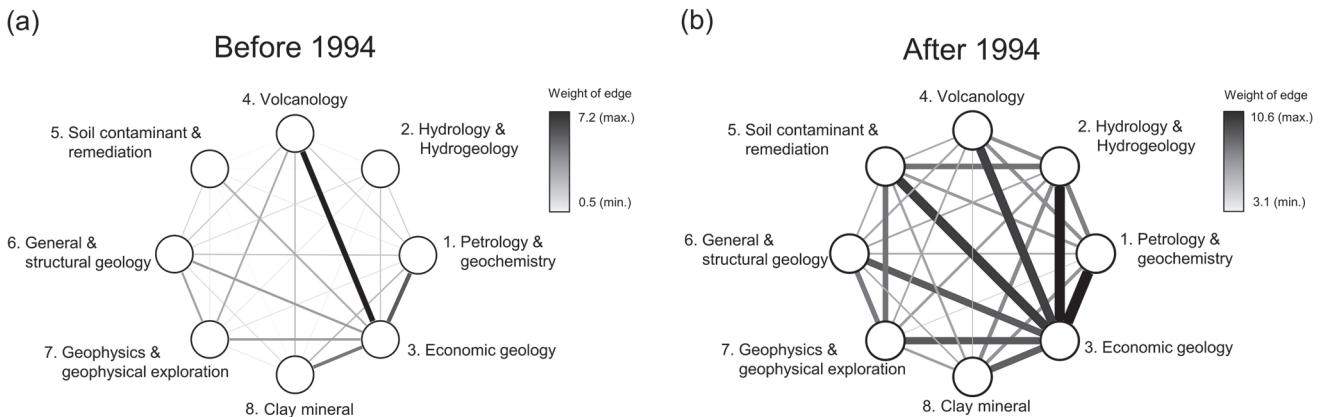


Fig. 6. Network structure of topics across two periods ((a) before 1994 and (b) after 1994). Nodes represent research topics and edges represent the degree of co-occurrence in abstracts and keyword sections.

관성(Fig. 6(a))과 이후의 연구 주제 간 연관성(Fig. 6(b))을 도식하였다. 또한, 연구주제 간 연관성을 시각화하기 위해 연결 강도(network strength)를 엣지의 굵기와 명도를 조절하여 표시하였다. 식 (3)을 이용한 1994년 이전의 네트워크 $M_{\text{before 1994}}$ 를 만드는 데 사용된 문헌은 총 944편이고, 이후의 네트워크 $M_{\text{after 1994}}$ 를 만드는 데 사용된 문헌은 총 1,627편이 사용되었다. 1994년 이전의 네트워크 평균 연결 강도는 1.81였으며(Fig. 6(a)), ‘화산학-광상학’(network strength = 5.42)였으며, ‘암석학 및 지구화학-광상학’(network strength = 1.79), ‘점토광물-광상학’(network strength = 1.52) 순으로 연구주제 간 연관성이 높았다. 이는 1994년 이전의 연구들이 ‘광상학’을 바탕으로 연구되었다고 해석된다. 또한, 특정 주제와 연결된 모든 주제에 대한 평균 연결강도는, ‘수문학 및 수리지질학’(mean network strength = 1.00), ‘토양오염 및 복원학’(mean network strength = 0.99)이 가장 낮았으며, 1994년 이전의 연구들은 환경분야와 관련된 연구보다 자원개발 및 탐사와 관련된 연구들이 활발하게 진행되었다는 것으로 해석된다.

1994년 이후의 네트워크의 총 평균 연결 강도(Fig. 6(b))는 5.08으로 1994년 이전 네트워크 총 평균 연결 강도(Fig. 6(a))의 약 2.81배 증가하였으며, 이는 최근에 지질학 분야의 융복합 연구가 활발히 진행된 것으로 판단된다. Fig. 6(b)에서 높은 연관성을 보여주는 연구주제들은 ‘암석학 및 지구화학-광상학’(network strength = 10.62)이었으며, ‘수문학 및 수리지질학-광상학’(network strength = 10.05), ‘토양오염 및 복원학-광상학’(network strength = 9.33) 순으로 높았다. 또한, 가장 낮은 연관성을 보여주는 연구주제들은 ‘암석학 및 지구화학-지구물리 및 물리탐사’(network strength = 3.10)였으며, 위와 같은 연구주제들은 연구주제 간 융합연구가 부족한 실정이기 때문에 이러한 경향을 보이는 것으로 해석된다. 따라서, 1994년 이후의 연구들은 ‘광상학’을 기준으로 학제 간 관점에서 해석하려는 융복합적 연구 동향을 보였으며, 환경문제가 대두됨에 따라 ‘수문학 및 수리지질학’, ‘토양오염 및 복원학’과 같은 연구들이 병행되었다는 것을 확인하였다.

4. 결 론

본 연구는 학술지 “자원환경지질”에서 1968년부터 2020년까지 총 52년간 발행한 논문을 대상으로 토픽모델링을 실시하여 학술지 “자원환경지질”의 역사에 대해 고찰하였다. 토픽모델링의 결과로 학술지 “자원환경지질”의 논문들은 총 8개의 연구주제로 분류할 수 있었다. 또한, 시간에 따른 연구주제별 확률 분포의 변화를 계산하여 연

구 동향을 확인하였으며, 학술지명을 바꾼 1994년을 기준으로 유의미한 변화를 확인하였다. 1994년 이전에는 ‘광상학’과 관련된 자원개발 및 탐사 연구들이 주를 이루었으며, 1994년 이후 지질학기반 환경분야의 연구들이 주를 이루었다. 특히, ‘수문학 및 수리지질학’과 ‘토양오염 및 복원학’과 같은 광산개발 후 수반되는 광산폐수와 광미 등 다양한 오염원을 대상으로 하는 환경분야의 연구들이 활발히 수행되었음을 확인하였다. 다음으로, 연관성 분석을 통해 연구 주제 간 연결 강도를 계산하고 결과를 해석하였다. 1994년 이전의 연구들은 ‘광상학’과 관련된 연구 주제(‘화산학’, ‘암석학 및 지구화학’, ‘점토광물’)들이 가장 연결 강도가 높았으며, 1994년 이후의 연구들은 ‘광상학’과 환경분야의 연구 주제(‘수문학 및 수리지질학’, ‘토양오염 및 복원학’) 간 연결 강도가 높았다. 따라서, 학술지 “자원환경지질”은 ‘광상학’을 주로 이루던 학술지에서 ‘광상학’ 기반 융복합연구를 바탕으로 환경분야까지 범주를 넓혀 광범위한 연구주제를 다루는 것으로 확인되었다.

본 연구에 사용된 LDA는 대용량의 문헌을 대상으로 내재되어 있는 주제들을 전반적으로 확인할 수 있었으나, 주어진 주제들에 대한 명확한 정의를 제공하지 않는 한계점을 나타내었다. 따라서, 토픽모델링을 이용한 모델링 결과를 바탕으로 기존에 수행된 문헌조사 방법과 병행하여 연구를 진행하면 보다 효율적이고 정확한 해석을 도출할 수 있을 것으로 판단된다. 향후 토픽모델링을 이용한 후속 연구 주제의 개수와 분류 방법, 단어들을 바탕으로 연구주제를 추론하는 해석방법 등을 고안하여 타당성과 신뢰도를 높일 수 있는 방법과 그에 따른 학술지의 질적향상을 위한 전략적 계획 수립, 지질학적 연구주제의 분포 변화 예측 등의 후속연구가 필요한 것으로 판단된다.

사 사

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업의 연구결과로 수행되었음”(2021-0-01533)

References

- Ahn, J.P. and Lee, M.H. (2018) Sorption Efficiency of the Bamboo Charcoal to Remove the Cesium in the Contaminated Water System. *Econ. Environ. Geol.*, v.51, p.87-97. doi: 10.9719/EEG2018.51.2.87.
- Bak, J.W. and Kim, D.H. (2015) Mineral Medicine Described in the Oriental Medicine Book and Mineral Medicines Applicable to Atopic Dermatitis Treatment. *Econ. Environ. Geol.*, v.48 p.231-

240. doi: 10.9719/eeg.2015.48.3.231.
- Ben, A.K. and Yvette, M. (2000) Environmental Contamination and Bioavailability of Toxic Element around the Daduk Mine Area. Korea. *Econ. Environ. Geol.*, v.33, p.273-282.
- Blei, D.M. (2012) Introduction to Probabilistic Topic Models. *Commun. ACM.*, v.55, p.77-84. doi: 10.1145/2133806.2133826.
- Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet allocation. *J. Mach. Learn. Res.*, v.3, p.993-1022. doi: 10.1016/b978-0-12-411519-4.00006-9
- Chi, J.M. (1975) A Study on the Genesis of Fluorite Deposits of South Korea. *Econ. Environ. Geol.*, v.8, p.25-56.
- Choi, S.G., Chi, S.J. and Park, S.W. (1988) Gold-Silver Mineralization of the Au-Ag Deposits at Yeongdong District, Chung-cheongbuk-Do. *Econ. Environ. Geol.*, v.21, p.367-380.
- Chon, H.T., Cheong, Y.W. and Kim, I.J. (1991) Oxygen and Sulfur Isotope Composition, and Genesis of Some Pyrophyllite Deposits Distributed in the Kyeongnam and Cheonnam Provinces. *Econ. Environ. Geol.*, v.24, p.97-105.
- Gatti, C.J., Brooks, J.D. and Nurre, S.G. (2015) A Historical Analysis of the Field of OR/MS using Topic Models. arXiv: 1510.05154.
- Griffiths, T.L. and Steyvers, M. (2004) Finding scientific topics. *proc. Natl. Acad. Sci.*, v.101, p.5228-5253. doi: 10.1073/pnas.0307752101.
- Hamm, S.Y., Cheong, J.Y., Kim, H.S., Hahn, J.S. and Cha, Y.H. (2005) Groundwater Flow Modeling in a Riverbank Filtration Area, Deasan-Myeon, Changwon City. *Econ. Environ. Geol.*, v.38, p.67-78.
- Han, J.K. and Choi, S.H., (2012) Ore Geology of Skarn Ore Bodies in the Kasihan Area, East Java, Indonesia. *Econ. Environ. Geol.*, v.45, p.1-8. doi: 10.9719/EEG2012.45.1.001.
- Heo, C.H. and Kim, S.Y. (2018) Analysis of Research Trends in the Journal of Economic and Environmental Geology: 1968-2017. *Econ. Environ. Geol.*, v.51, p.249-277. doi: 10.9719/EEG2018.51.3.249
- Hsu, H. and Lachenbruch, P.A. (2005) Paired t test. *Encycl. Biostat.*, v.6.
- Hwang, J.S., Min, K.D., Choi, C. and Yu, S.H. (2003) Gravity Survey Around the Palgongsan Granitic Body and Its Vicinity. *Econ. Environ. Geol.*, v.36, p.305-312.
- Hwang, S.K. (1993) One-cyclic Volcanic Processes at Udo Crater, Korea. *Econ. Environ. Geol.*, v.26, p.55-65.
- Hwang, S.K., Jo, I.H. and Yi, K. (2017) SHRIMP U-Pb Dating and Volcanic Processes of the Volcanic Rocks in the Guamsan Caldera, Cheongsong, Korea. *Econ. Environ. Geol.*, v.50, p.467-476. doi: 10.9719/EEG2017.50.6.467.
- Hyun, S.G., Woo, N.C., Shin, W.S. and Hamm, S.Y. (2006) Characteristics of Groundwater Quality in a Riverbank Filtration Area. *Econ. Environ. Geol.*, v.39, p.151-162.
- Jeon, J.H., Kim, I.S., Lee, M.H. and Jang, Y.Y. (2006) Study of Solidification by Using Portland and MSG (micro silica grouting) Cements for Metal Mine Tailing Treatment. *Econ. Environ. Geol.*, v.39, p.699-710.
- Kang, I.M. (2012) About short-stacking effect of illite-smectite mixed layers. *Econ. Environ. Geol.*, v.45, p.71-78. doi: 10.9719/EEG2012.45.2.071.
- Kim, C. B. and Kim, Y.J. (1990) Geochronology and Petrochemistry of Foliated Granites between Damyang and Jinan. *Econ. Environ. Geol.*, v.23, p.233-244.
- Kim, C.H., Park, C.H., Ko, Y.T., Jung, E.Y., Kwak, J.Y., Yoo, S.H. and Min, K.D. (2007) The study of Formation for Dokdo seamounts at the northeastern part of the Ulleung basin using gravity and magnetic data. *Econ. Environ. Geol.*, v.40, p.153-170.
- Kim, H.G., Hwang, J.K. and Hwang, J.S. (2020) Topic Modeling of The Journal of the Wind Engineering Institute of Korea using LSA and LDA. *J. Wind Eng. Inst. Korea* v.24, p.113-120. doi: 10.37109/weik.2020.24.3.113
- Kim, I.J. (1992) Alteration zoning, mineral assemblage and geochemistry of the hydrothermal clay deposits related to Cretaceous felsic magmatism in the Haenam area, southwest Korea. *Econ. Environ. Geol.*, v.25, p.397-416.
- Kim, K.H. (1986) Isotope Geochemistry of Uranium Ore Deposits in Okcheon Metamorphic Belt, South Korea. *Econ. Environ. Geol.*, v.19, p.163-173.
- Kim, K.H., Kim, O.J., Min, K.D. and Lee, Y.S. (1984) Structural, paleomagnetic and petrological studies of the Chugaryeong Rift Valley. *Econ. Environ. Geol.*, v.17, p.215-230.
- Kim, O.J. (1973) The stratigraphy and geologic structure of the metamorphic complex in the northwestern area of the Kyonggi Massif. *Econ. Environ. Geol.*, v.6, p.201-216.
- Kim, S.G. and Bae, H.S. (2006) Investigation of post-seismic sites using local seismic tomography in the Korean peninsula. *Econ. Environ. Geol.*, v.39, p.111-128.
- Kim, S.K., Jun, M.S. and Jeon, J.S. (2006) Recent Research for the Seismic Activities and Crustal Velocity Structure. *Econ. Environ. Geol.*, v.39, p.369-384.
- Kim, S.O. and Park, M.E. (2015) Standardization Studies for the Oriental Mineral Medicine. *Econ. Environ. Geol.*, v.48, p.187-197. doi: 10.9719/eeg.2015.48.3.187
- Kim, S.Y., Lee, J.W. and Ahn, E.Y. (2004) A Suggestion for Improvement in Quality through Bibliometrics of Research Articles published in Domestic Journals in Earth and Environmental Sciences. *J. Geol. Soc. Korea.*, v.40, p.103-111.
- Kim, S.W. (1971) Effects of Geological Conditions on the Geomorphological Development of the Southwestern Coastal Regions of Korea. *Econ. Environ. Geol.*, v.4, p.11-18.
- Kim, Y.J., Park, Y.S. and Kang, S.W. (1994) The Study on Geochronology and Petrogenesis of Foliated Granites in the Honam Shear Zone, South Korea. *Econ. Environ. Geol.*, v.27, p.247-261.
- Kim, Y., Oh, J.M., Jung, H.Y., Lee, S.Y. and Roh, Y. (2014) Characterization of microbial diversity of metal-reducing bacteria enriched from groundwater and reduction/biomineralization of iron and manganese. *Econ. Environ. Geol.*, v.47, p.431-439. doi: 10.9719/EEG2014.47.4.431.
- Kim, Y.K. and Lee, D.S. (1983) Petrology of alkali volcanic rocks in northern part of Ulrung Island. *Econ. Environ. Geol.*, v.16, p.19-36.
- Ko, K.S., Kim, Y., Koh, D.C., Lee, K.S., Lee, S.G., Kang, C.H.,

- Seong, H.J. and Park, W.B. (2005) Hydrogeochemical Characterization of Groundwater in Jeju Island using Principal Component Analysis and Geostatistics. *Econ. Environ. Geol.*, v.38, p.435-450.
- Koh, G.W. and Park, J.B. (2010a) The Study on geology and volcanism in Jeju Island (II): Petrochemistry and $^{40}\text{Ar}/^{39}\text{Ar}$ absolute ages of the volcanic rocks in Gapado-Marado, Jeju Island. *Econ. Environ. Geol.*, v.43, p.53-66.
- Koh, G.W. and Park, J.B. (2010b) The study on geology and volcanism in Jeju Island (III): Early lava effusion records in Jeju Island on the basis of $^{40}\text{Ar}/^{39}\text{Ar}$ absolute ages of lava samples. *Econ. Environ. Geol.*, v.43, p.163-176.
- Koh, G.W., Park, J.B. and Park, Y.S. (2008) The Study on Geology and Volcanism in Jeju Island (I): Petrochemistry and $^{40}\text{Ar}/^{39}\text{Ar}$ Absolute ages of the Subsurface Volcanic Rock Cores from Boreholes in the Eastern Lowland of Jeju Island. *Econ. Environ. Geol.*, v.41, p.93-113.
- Kwon, Y.K. (2012) Sequence Stratigraphy of the Yeongweol Group (Cambrian-Ordovician), Taebaeksan Basin, Korea: Paleogeographic Implications. *Econ. Environ. Geol.*, v.45, p.317-333. doi: 10.9719/EEG2012.45.3.317.
- Kwon, Y.K., Kwon, Y.J., Yeo, J.M. and Lee, C.Y. (2019) Basin Evolution of the Taebaeksan Basin during the Early Paleozoic. *Econ. Environ. Geol.*, v.52, p.427-448. doi: 10.9719/EEG2019.52.5.427.
- Lee, C.H., Lee, H.K. and Kim, K.W. (1997) Petrochemistry and environmental geochemistry of shale and coal from the Daedong Supergroup, Chungnam Coal Field, Korea. *Econ. Environ. Geol.*, v.30, p.417-431.
- Lee, C.H., Lee, H.K., Lee, J.C. and Jeon, S.R. (2001) Environmental geochemistry and contamination assessment of the Tohyun mine creek, Korea. *Econ. Environ. Geol.*, v.34, p.471-483.
- Lee, H.J., Kim, E.J. and Moon, D.H. (2011) Identification of advanced argillic-altered rocks of the haenam area, using by aster spectral analysis. *Econ. Environ. Geol.*, v.44, p. 463-474. doi: 10.9719/EEG2011.44.6.463.
- Lee, H.J. and Lee, M.H. (2012) Investigation of the rice plant transfer and the leaching characteristics of copper and lead for the stabilization process with a pilot scale test. *Econ. Environ. Geol.*, v.45, p.255-264. doi: 10.9719/EEG2012.45.3.255.
- Lee, H.K. and Lee, C.H. (1997) Geochemical origin, behavior and enrichment of environmental toxic elements in coaly metapelite from the Deokpyeong area, Korea. *Econ. Environ. Geol.* v.30, p.553-566.
- Lee, H.K., Moon, H.S., Min, K.D., Kim, I.S., Yun, H. and Itaya, T. (1992) Paleomagnetism, stratigraphy and geologic structure of the Tertiary Pohang and Changgi basins; K-Ar ages for the volcanic rocks. *Econ. Environ. Geol.*, v.25, p.337-349.
- Lee, J.S. (2014) Geochemical Characteristics of the Sub-alkaline Basalt in the Udo Island, Jeju. *Econ. Environ. Geol.*, v.47, p.601-610. doi: 10.9719/EEG2014.47.6.601.
- Lee, J.Y. (2020) Boosting interest and roles of geology as pure and applied sciences. *J. Geol. Soc. Korea.*, v.56 p.1-2. doi: 10.14770/jgsk.2020.56.1.1
- Lee, J.S. and Chon, H.T. (2004) Human risk assessment of toxic heavy metals around abandoned metal mine sites. *Econ. Environ. Geol.*, v.37, p.73-86.
- Lee, S.Y., Roh, Y. and Jeong, J.T. (2014) Changes of the Oxidation/Reduction Potential of Groundwater by the Biogeochemical Activity of Indigenous Bacteria. *Econ. Environ. Geol.*, v.47, p.61-69.
- Min, K.D. and Cho, K.E. (1995) Gravity Survey of the Subsurface Geology and Geologic Structure between Samcheog and Taebaek Area. *Econ. Environ. Geol.*, v.28, p.79-88.
- Na, C.K., Lee, I.S. and Chung, J.I. (1997) Petrogenetic Study on the Foliated Granitoids in the Chonju and the Sunchang area (II)-In the Light of Sr and Nd Isotopic Properties. *Econ. Environ. Geol.*, v.30, p.249-262.
- Na, K.C., Kim, H.S., Lee, D.J. and Lee, S.H. (1982) Comparative studies between Chungju and Seosan groups. *Econ. Environ. Geol.*, v.15, p.177-188.
- Oh, C.W., Yu, Y.H., Lee, P.K. and Lee, Y.U. (2003) The effects of pH change in extraction solution on the heavy metals extraction from soil and controversial points for partial extraction in Korean Standard Method. *Econ. Environ. Geol.*, v.36, p.159-170.
- Oh, J.Y. and Suh, M.C. (2005) Site Characteristics Around the Gongsansung Circular Pond in Gongju Based on the Seismic Methods. *Econ. Environ. Geol.*, v.38, p.623-631.
- Park, Y.S. (1997) Nd Model Age and Nd Isotopic Evidence of Granitoid Rocks in the Gwangju-Naju Area, Korea. *Econ. Environ. Geol.*, v.30, p.153-161.
- Porteous, I., Newman, D., Ihler, A., Asuncion, A., Smyth, P. and Welling, M. (2008) Fast collapsed gibbs sampling for latent dirichlet allocation. the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p.569-577. doi: 10.1145/1401890.1401960
- Sang, K.N. and Chung, W.W. (1990) Research Trend of Geology and Mineral Resources in Korea. *J. Geol. Soc. Korea*, v.26, p.500-512.
- Shin, D. (2020) A comparative study of domestic and international research trends of mathematics education through topic modeling. *Math. Educ.*, v.59, p.63-80. doi: 10.7468/mathedu.2020.59
- Shin, I.H. and Kagami, H. (1996) Rb-Sr Isochron Ages, Sr and Nd Isotopic Compositions of Granophyre in the Haenam-Wando Areas, Korea. *Econ. Environ. Geol.*, v.29, p.151-158.
- Son, M. and Kim, I.S. (1994) Geological Structures and Evolution of the Tertiary Chŏngja Basin, Southeastern Margin of the Korean Peninsula. *Econ. Environ. Geol.*, v.27, p.65-80.
- Sun, L. and Yin, Y. (2017) Discovering themes and trends in transportation research using topic modeling. *Transp. Res. Part C Emerg. Technol.*, v.77, p.49-66. doi: 10.1016/j.trc.2017.01.013
- Um, J.G., Kang, T. and Hwang, J.Y. (2005) Rock Slope Stability Investigations Conducted on the Road Cut in Samrangjin-Miryang Area. *Econ. Environ. Geol.*, v.38, p.305-317.
- Wee, S.H. and Kim, S.K. (2008) The Site Effect of the Broadband Seismic Stations in Korea. *Econ. Environ. Geol.*, v.41, p.225-242.
- Woo, Y.K. (1988) Genesis of Iron Ore Deposits in the south-eastern Part of Gyeongnam Province, Korea. *Econ. Environ. Geol.*, v.21, p.45-56.

- Xiong, H., Cheng, Y., Zhao, W. and Liu, J. (2019) Analyzing scientific research topics in manufacturing field using a topic model. *Comput. Ind. Eng.*, v.135, p.333-347. doi: 10.1016/j.cie.2019.06.010
- Yang, D.Y. and Chi, J.M. (1986) Grain Size and Texture of Silver Minerals from Duk-Eum Ore Deposits. *Econ. Environ. Geol.*, v.19, p.227-237.
- Yang, S.I., Kang, D.H., Kim, T.Y., Chung, S.Y. and Kim, M.C. (2008) Characteristics of Groundwater Quality in Bedrock and Tailing Dumps at the Abandoned Dalcheon Mine Area. *Econ. Environ. Geol.*, v.41, p.47-56.
- Yin, J., Liu, C., Park, J.H., Shao, X., Yang, H., Xu, H. and Wang, J. (2013) The Geochemical and Zircon Trace Element Characteristics of A-type Granitoids in Boziguoer, Baicheng County, Xinjiang. *Econ. Environ. Geol.*, v.46, p.179-198. doi: 10.9719/eeg.2013.46.2.179
- Yoo, B.C., Lee, H.K. and Choi, S.G. (2002) Stable isotope, Fluid Inclusion and Mineralogical Studies of the Samkwang Gold-Silver Deposits, Republic of Korea. *Econ. Environ. Geol.*, v.35, p.299-316.
- Yun, S. (1983) Skarn-ore associations and phase equilibria in the Yeonhwa-Keodo Mines, Korea. *Econ. Environ. Geol.*, v.16, p.1-10.
- Yun, S.H. and Hwang, I.H. (1988) Volcano-Stratigraphy and Petrology of the Volcanic Mass in the Koheung Peninsula, South Cheolla Province, Korea. *Econ. Environ. Geol.*, v.21, p.335-348.