

Unveiling research on Baekdudaegan: bibliometric analysis and Latent Dirichlet Allocation topic modelling approaches

Thakur Dhakal, Tae-Su Kim, Shraddha Tiwari, Seong-Hyeon Kim, Anteneh Bongasie, JunYoung Kim, Su-Jin Kim & Gab-Sue Jang

Keywords: Baekdudaegan, biodiversity, Latent Dirichlet Allocation, sustainable development goals, text mining

Abstract

Baekdudaegan is a protected mountain range that spans 1,400 km across the Korean Peninsula and serves as a hub of biodiversity and cultural heritage. This study presents a comprehensive analysis of research insights into this complex mountain ecosystem through bibliometric analysis and Latent Dirichlet Allocation (LDA) topic modelling approaches using metadata from 165 scholarly publications sourced from the Web of Science (WoS) core collection. We observed a discernible upward trend in Baekdudaegan-related literature, with a notable surge in scholarly interest post-2017, primarily from institutions such as the Baekdudaegan National Arboretum and the Korea National Arboretum. According to WoS categories, the examined research was dominated by plant science, ecology and environmental science clusters. Employing the LDA model to the corpus of texts extracted from keywords and abstracts of each observed literature, we identified six principal research topics: biochemical pathways, forest ecosystem dynamics, genomic taxonomy, seedling dynamics, biodiversity conservation and climate change. Finally, we visualized the keyword network and LDA-trained word cloud topics. The findings of this study enhance a deeper understanding of Baekdudaegan Mountains research, offering valuable insights for researchers and policymakers. This study serves as a resource for future investigations and informed decision-making regarding the sustainable management of mountain ecosystems.

Profile

Protected area

Baekdudaegan

Mountains Reserve

Mountain range

Baekdudaegan

mountain range

Country

North and South Korea

Introduction

Mountains serve as significant reservoirs of biodiversity, regulate climate and provide vital resources such as freshwater fuel and other natural materials essential for human survival and ecological balance (Körner 2004; Payne et al. 2017). Research on mountains is increasing and essential for understanding their conservation challenges, ecological dynamics and the impacts of climate change, as well as to provide the status of mountain studies globally (Gunya & Sarmiento 2022; Solár & Solár 2020; Thaler et al. 2021). Much research has been conducted in various disciplines within a variety of approaches and frameworks (Kratzer 2018; Palau & Claramunt-López 2024). Strengthening mountain specific research is necessary to assess sensitivity to global climate scenarios and to ensure sustainable mountain ecosystems (Negi et al. 2014; Schmidt-Vogt et al. 2016). Recent studies have highlighted the importance of expanding literatures, particularly those relating to protected mountain ranges such as Baekdudaegan in East Asia (Chung et al. 2017; Schroth et al. 2009; Theobald et al. 2024).

The Baekdudaegan mountain range (BMR) has significant historical, ecological, cultural and conservation importance in the Korean Peninsula (Choi 2004; Chung et al. 2018a). It is also called the *spine of Korea*, stretching more than 1,400 km across the Korean Pen-

insula, from the northern border with North Korea to the southernmost point of South Korea (Cho & Chun 2015; Min et al. 2019). Baekdudaegan's classification as a biosphere reserve and other conservation efforts have been implemented by governmental and non-governmental organizations (Park & Lee 2014; Hayes 2010). Significant obstacles still exist in some regions despite these initiatives, such as conflicting land use interests, economic constraints and political difficulties (Lee 2019). Previous studies have explored comprehensive strategies to tackle these issues including community involvement and sustainable development methods; strong legislative frameworks are necessary for successful conservation (Cárdenas et al. 2018; Cho et al. 2022; Kim 2011). Scholars have directed their focus towards understanding the significance and characteristics of notable mountains like those of Baekdudaegan in Korea, resulting in a burgeoning of literature offering diverse perspectives (Altaweel et al. 2019; Arroyave et al. 2024).

When one explores the latent structures of existing research, bibliometric studies and topic modelling emerge as effective tools for communicating with scholars (Calvo et al. 2022; Cheng et al. 2018; Davies et al. 2012). Bibliometric studies shed light on research development and its significance by analysing literature quantitatively and focusing on publication patterns and research themes (Cobo et al. 2011; Donthu

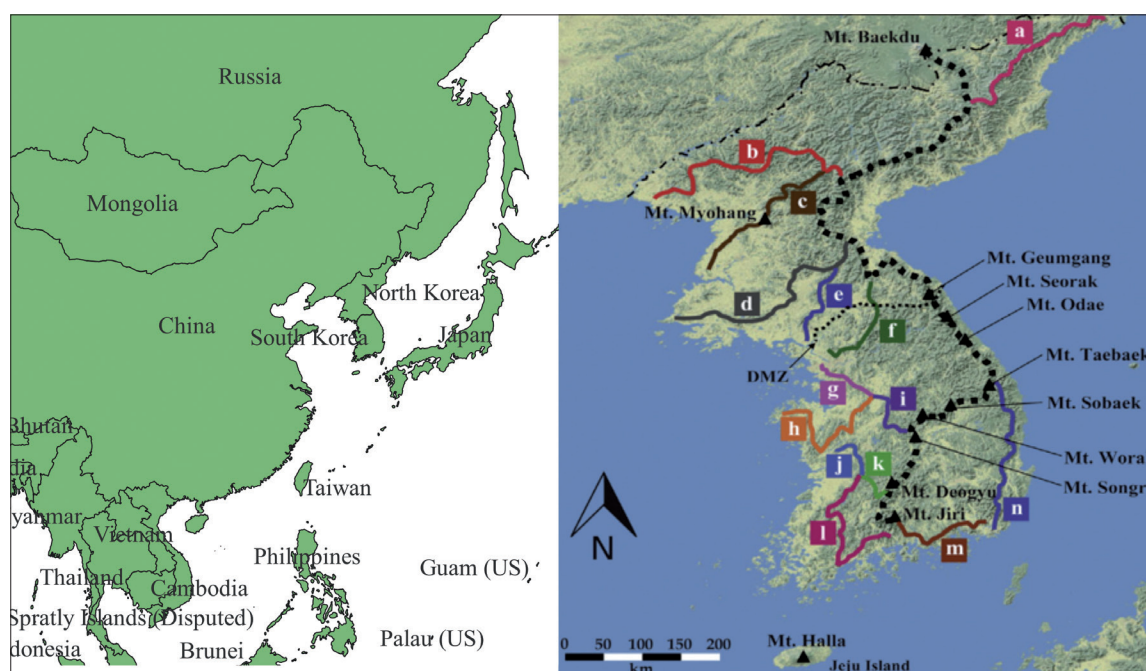


Figure 1 – Location of Korea in the East Asian map and the Baekdudaegan mountain system (where the main ridgeline is represented by a thick black dashed line), which runs north to south along the Korean Peninsula. The thin curved lines in diverse colours represent 14 subsidiary mountain ranges of the Baekdudaegan Mountains (which form one jeonggan and 13 jeongmaeks): a – Jangbaekjeonggan; b – Cheongbukjeongmaek; c – Cheongnamjeongmaek; d – Haeseojeongmaek; e – Imjinbukyesungnamjeongmaek; f – Hanbukjeongmaek; g – Hannamjeongmaek; h – Geumbukjeongmaek; i – Hannamgeumbukjeongmaek; j – Geumnamjeongmaek; k – Geumnambonamjeongmaek; l – Honamjeongmaek; m – Nakanjeongmaek; and n – Nakdongjeongmaek (Chung et al. 2017). Dotted line: DMZ – Korean Demilitarized Zone.

et al. 2021; Gaviria-Marin et al. 2018; Kratzer 2018). Understanding content from a large corpus of documents may be facilitated by means of topic modelling, which uncovers the latent thematic patterns of words (Abdelrazek et al. 2023; Daud et al. 2010). In recent decades, bibliometric and natural language topic modelling techniques have garnered increasing research interest and are gaining in popularity on account of the development of new software programs, the ability to handle large databases and sophisticated machine learning techniques (Ellegaard & Wallin 2015; Grisales et al. 2023; Wang et al. 2023).

A combination of probabilistic natural language-based topic modelling and qualitative bibliometric research methods helps delve into the significance, research trajectories and prospective avenues of scholarly inquiry (Cobelli & Blasi 2024; Jiang et al. 2016). These studies have been conducted in multiple sectors, including ecological dynamics, cultural heritage prevention, climate change, energy conservation and socioeconomic studies (Vlase & Lähdesmäki 2023; Woolston 2023). During our literature search, comprehensive studies on mountain research using bibliometric analysis and topic modelling were limited. Therefore, this paper aims to fill this gap by conducting an in-depth analysis through bibliometrics and topic modelling to analyse research relating to the BMR, offering valuable insights for international dialogues concerning mountain conservation and ecosys-

tem management. The findings and approach enhance understanding and advance knowledge in the field of mountain research.

Materials and methods

Baekdudaegan

Baekdudaegan, designated as the Baekdudaegan Mountains Reserve is located in East Asia and is the ecological axis of the Korean Peninsula. The BMR is critical to Korea's natural heritage and the global significance of its biodiversity. It is an approximately 1,400 km-long central mountain range that extends from Mt Baekdu, North Korea to Mt Jiri, South Korea (Figure 1). There are fourteen subsidiary mountain ranges connected to the BMR (one jeonggan and 13 jeongmaeks), see Figure 1. Baekdudaegan is a treasured representation of Korean identity considering its varied ecosystems, immaculate scenery and rich cultural legacy. It is a key corridor for wildlife, a place of inspiration for artists and a path for spiritual seekers. Baekdudaegan plays a vital role in the country's climate and water systems. It symbolizes Korea's natural heritage and is a focal point for conservation initiatives (Je-Hun & Doo-Hee 2013; Jung 2015; Kim et al. 2023).

The present conservation status of the BMR presents considerable differences between South and North Korea. In South Korea, the Korea Forest Service passed the Baekdudaegan Protection Act in 2003

and designated the Baekdudaegan protected area in 2005 (Cho & Chun 2015). South Korea offers a good degree of protection, as large parts are covered by the network of protected areas (Chung et al. 2017). In contrast, the North Korean part of the Baekdudaegan mountains system is rarely covered by protected areas, whereas deforestation affects large parts, including in the globally recognized Mt Baekdu Biosphere Reserve zone (Tang et al. 2010). The mountain range is rich in species diversity, with an estimated 4,662 vascular plants, 135 bird, 36 mammals and 32 amphibian and reptile species (Chung et al. 2018a). In the South Korean section, there are 379 villages linking directly to BMR (Kim et al. 2023). In this study, we have explored the latent themes contents of research centred on the glacial refugium of the BMR.

Data

Scholarly data can be obtained from various platforms, including Web of Science (WoS), Scopus, Google Scholar, Semantic Scholar, PubMed and Crossref. WoS is a comprehensive science-indexed database that extensively collects academic journals, conference proceedings and the most accepted peer-reviewed scholarly publications (Falagas et al. 2008; Li et al. 2018). Considering the relevance and impact of publications, this study has examined Baekdudaegan-related research sourced from the WoS Core Collection. A total of 165 scholarly publications retrieved using the search function {ALL="Baekdudaegan"} in the WoS Core Collection, as of 27 December 2023, have been examined for bibliometric analysis and topic modelling in this study (see Supplementary File S1). To examine the transition of dominant latent words and modelled topics, keywords and abstracts of the literature were used without considering general words such as auxiliary verbs, connecting words, stop words and punctuation after transferring all text into lowercase. The research data for 2023 were incomplete. Therefore, we analysed the publication trend and assessed the goodness of fit with R² value (Cameron & Windmeijer 1997), along with overall citations, based on the literature data between 2004 and 2022.

Bibliometrics insights

Bibliometrics offer valuable insights into scholarly communication dynamics and the impact of research across academic disciplines (Ninkov et al. 2022; Pan et al. 2024). Reporting influential articles is essential for recognizing impactful research domains, the sustainable development goals (SDG) alignment of articles enhances the strategic value of academic efforts and WoS categories can effectively navigate and make use of the specific field of scholarly literature. The contributing institution facilitates the potential stakeholders in the research ecosystem, and keywords and their networks enhance an understanding of how different concepts are interconnected (Calvo et al. 2022; Cheng et al. 2018; Davies et al. 2012). Similarly, the number

of authors contributing to a single publication attests to collaborative projects and the authors' network suggests an association between two or more scientists pursuing shared goals (Kratzer 2018; Ortega 2014; Venturini et al. 2024).

We therefore examined the influencing research, SDG alignments, WoS categories, collaborative institutions and authors using the WoS core collection analysis tool. The relationship among keywords that contained at least two publications were visualized using a co-occurrence network graph that was generated using VOS Viewer software version 11.6.18 (Van Eck & Waltman 2018). The keywords and authors networks were graphed with the default configuration of the VOSviewer software (Van Eck & Waltman 2018). The ten most cited articles were summarized with key findings and annual citation rate. We visualized the annual latent words, which examined a collection of five dominant words for each year, and the transitions of the most frequently used 30 words were presented in the RStudio environment (R Core Team 2023).

Topic modelling

Topic modelling is an unsupervised probabilistic text classification technique used to uncover latent textual data frameworks (Griffiths & Steyvers 2004). With advances in research and development (R&D) and robust computational tools, various techniques such as Latent Dirichlet Allocation (LDA) (Blei et al. 2003), non-negative matrix factorization (Da Kuang et al. 2015), correlation (Blei & Lafferty 2005), joint sentiment (Lin & He 2009), bidirectional encoder representations from transformers (Devlin et al. 2019) and models in natural language processing (NLP) have been applied to topic modelling (Abdelrazek et al. 2023). Several researchers have gravitated to LDA, which has been successfully applied to different research clusters (Du et al. 2021; Jelodar et al. 2019; Laureate et al. 2023; Vayansky & Kumar 2020). LDA is highly recommended for analysing latent topics because of its advantages in flexibility and coherence (Stevens et al. 2012; Vayansky & Kumar 2020). LDA is a generative probabilistic model that follows the three-level hierarchical Bayesian rule (Blei et al. 2003).

LDA characterizes documents as distributions over topics, with each topic represented as a collection of words, thereby revealing topic generation within a corpus (Vayansky & Kumar 2020). Using a Dirichlet distribution, LDA initially samples a document-specific multinomial distribution over subjects to construct a document. Words from these topics were sampled repeatedly based on the Bayesian rule (Blei et al. 2003). In LDA, α and β are the corpus-level hyperparameters and are assumed to be sampled once while generating a corpus; θ is the document-level parameter sampled once in a document; and Φ is the topic-level distribution of words. The variables z and w are word-level variables sampled once for each word in each document, as graphically represented in Figure 2. In LDA,

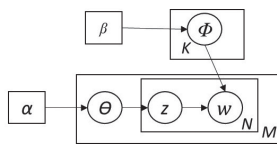


Figure 2 – A graphical representation of the latent Dirichlet allocation (LDA) model. The Figure illustrates the relation between documents, topics and words. The letters α and β are the corpus-level hyperparameters; θ is the document-level parameter sampled once in a document; Φ is the topic-level distribution of words; z and w are word-level variables sampled once for each word in each document; M is the total number of documents; N is the number of words in each document; and K is the predefined number of topics (Blei et al. 2003).

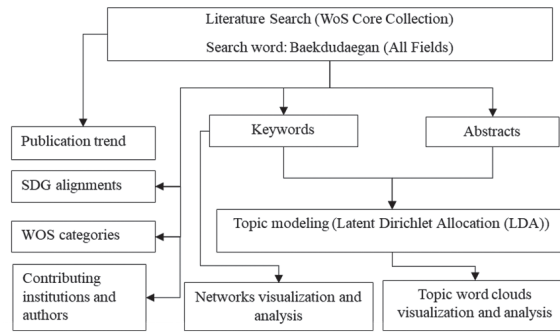


Figure 3 – Workflow summary for bibliometric and topic modelling analysis of Baekdudaegan research indexed in WoS database from 2004 to 2022.

each M document is a mixture of K latent topics and each topic describes a multinomial distribution of N words (Blei et al. 2003). These parameters have been estimated using Gibbs sampling and variational expectation maximization (VEM) approaches, implemented via the *topicmodels* package within the RStudio environment (Grün & Hornik 2011). Gibbs sampling is more accurate and provides more stable topics than the VEM approach (Marani & Baumer 2023; Mardones-Segovia et al. 2022; Zubir et al. 2018). Leveraging this established understanding, the current study employed the Gibbs sampling method for its topic modelling.

It is essential to determine the appropriate number of topics in topic modelling (Arun et al. 2010; Cao et al. 2009). In this study, the possible number of optimal topics (K) was determined by comparing four metrics – Griffiths 2004 (Griffiths & Steyvers 2004), Cao 2009 (Cao et al. 2009), Arun 2010 (Arun et al. 2010) and Deveaud 2014 (Deveaud et al. 2014) – each

offering distinct perspectives on topic selection. Griffiths 2004 assesses topic coherence by measuring the semantic similarity among the top words within each topic, while Cao 2009 focuses on model perplexity, indicating the model's ability to predict unseen data. Arun 2010 strikes a balance between model complexity and coherence by maximizing intra-topic coherence and minimizing inter-topic separation. Deveaud 2014 provides a comprehensive evaluation of topic model quality, considering both coherence and coverage. Our study evaluated these metrics by employing a range of optimization techniques: Griffiths 2004 and Deveaud 2014 were optimized for maximization, while Cao 2009 and Arun 2010 were optimized for minimization. The topic $K=10$ was initially arbitrarily defined, and the scenarios were evaluated for each metric using the *ldatuning* package (Murzintcev & Chaney 2022). Visualizing a larger number of words can lead to overcrowding in the word cloud, making it difficult

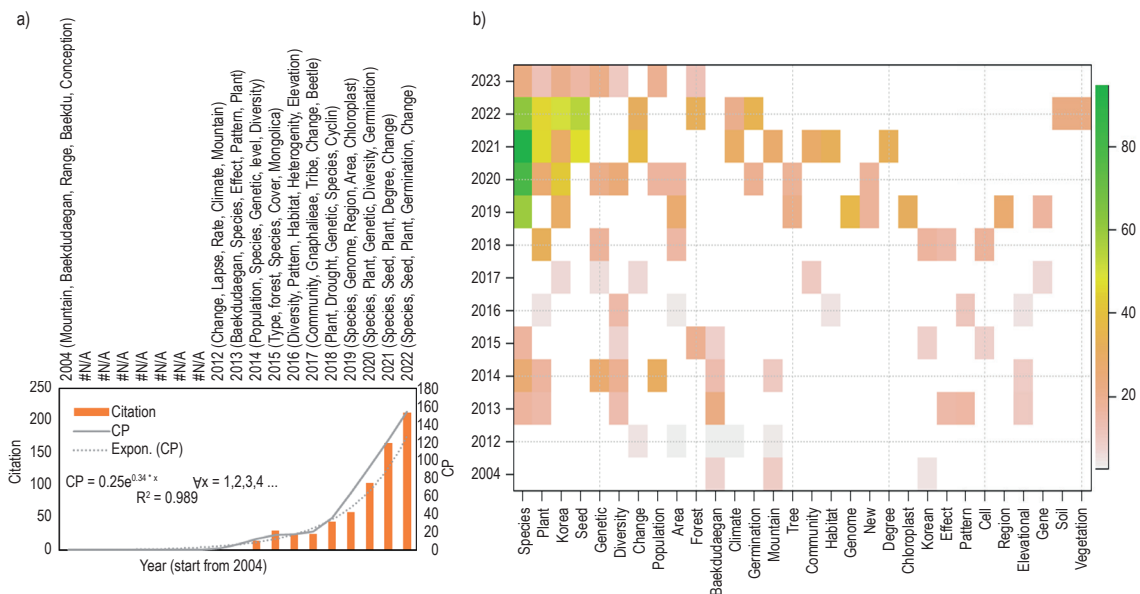


Figure 4 – WoS-indexed Baekdudaegan research contents: a) cumulative publication (CP), citation trend and exponential fit of CP (dominated words in each year in parentheses; #N/A: no record), b) highly dominated 30 words transitions over time (colourbar represents the frequency of words).

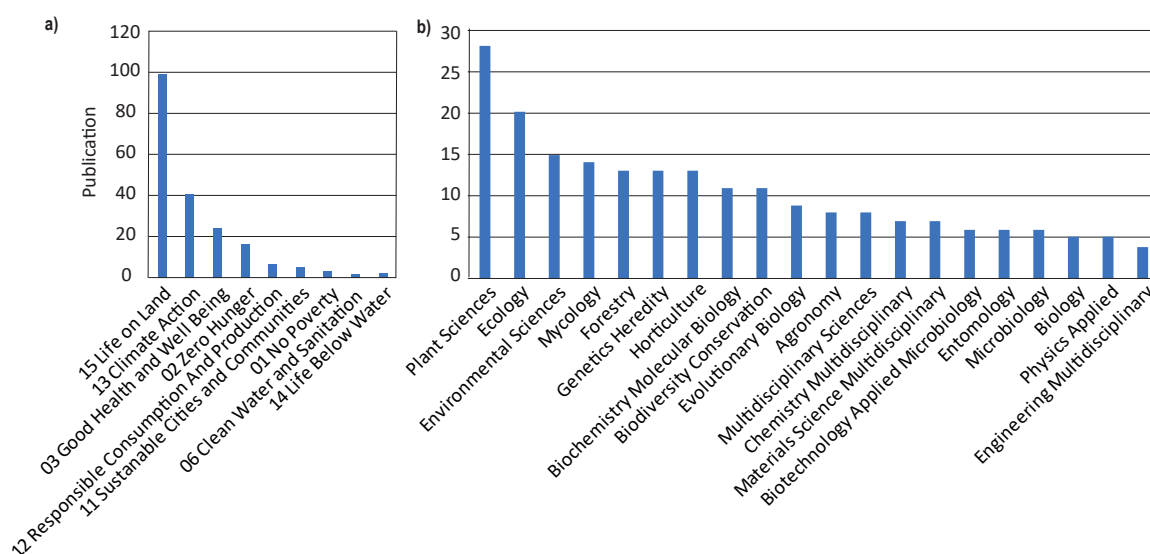


Figure 5 – Frequency plots for research alignment to a) Sustainable Development Goals and b) top 20 Web of Sciences categories.

for viewers to interpret. We therefore extracted the first 100 dominant words for each topic and visualized the word cloud computed in the RStudio environment (Fellows 2022; R Core Team 2023). A detailed summary of research flow is shown in Figure 3.

Results

Publication trend

Choi (2004) was the first report of the 165 publications (Articles: 157; Reviews: 4; Proceeding papers: 1; Editorial material: 2; Meeting abstracts: 1) publications indexed in the WoS database. Remarkably, there was a conspicuous absence of WoS-indexed publications from 2004 to 2012, followed by a notable surge in publications post-2017 (Figure 4a). There is an upward trajectory in cumulative publications, characterized by an exponential growth rate of 0.34 literatures per year with a constant multiplier of 0.25 ($R^2 = 0.989$). Concurrently, citations exhibit a corresponding increase correlating with the rising rate at 5.19 citations per publication. When scoping the influence publication on the topic *Baekdudaegan*, *Altitudinal patterns of plant species richness on the Baekdudaegan Mountains, South Korea: mid-domain effect, area, climate, and Rapoport's rule* (Lee et al. 2013) was observed as the most cited (41 times and 3.42 citations per year) (Table 1). The existing literature on Baekdudaegan mainly focuses on species distribution, climate influences and the glacial refugium characteristics of this mountain region (Chung et al. 2017). The latent words derived from keywords assigned by authors and the abstracts of publications revealed a notable shift in research focus over time. Initially, the discourse was dominated by terms such as *Baekdudaegan* and *mountain*, gradually transitioning to include words like *species*, *seed*, *plant* and *germination* in recent years (as indicated by x-level in parentheses in Figure 4a). Further scrutiny concentrated on the five most frequently employed words, leading to the ex-

traction of the thirty most recurrent terms over time. Notably, terms such as *species*, *plant* and *seed* emerged as dominant themes across various temporal segments, as depicted in the heatmap illustrated in Figure 4b.

We revealed that the research outputs were predominantly aligned with nine specific SDGs that were observed using the SDGs feature within the WoS core collection analysis tool. Notably, a substantial portion of the research corpus (comprising 99 articles) was aligned with the exploration of SDG 15, concerning life on land. This was closely followed by contributions to SDG 13, addressing climate action, and SDG 3, focusing on good health and well-being (Figure 5a). The examined literature ranged from various (a total of 45) WoS categories, with plant science as the dominant WoS category, followed by ecology, environmental science and forestry (see top 20 categories in Figure 5b).

We examined the authors' affiliations, number of authors per article and dominant authors' publications (Figure 6) to identify collaborating institutes and authors. A total of 454 authors from 171 institutions contributed to the examined literature. The Baekdudaegan National Arboretum was the dominant organization, followed by the Korea National Arboretum, Chonnam National University and Korea Forest Research Institute (Figure 6a). Similarly, there were 38 publications contributed by five authors, with the maximum number of authors per article being 17 and an average of eight authors (Figure 6b). The author Kim, YS made the highest number of occurrences ($n=25$), followed by Na, CS ($n=19$) and Kim, DW ($n=18$) (Figure 6c).

Keyword and authors' networks

A total of 399 co-occurring keywords were identified while analysing the VOSviewer framework, where 65 words met the threshold of the minimum number of occurrences (two times) of a keyword visualized in the network graph with thematic clusters (Figure 7a).

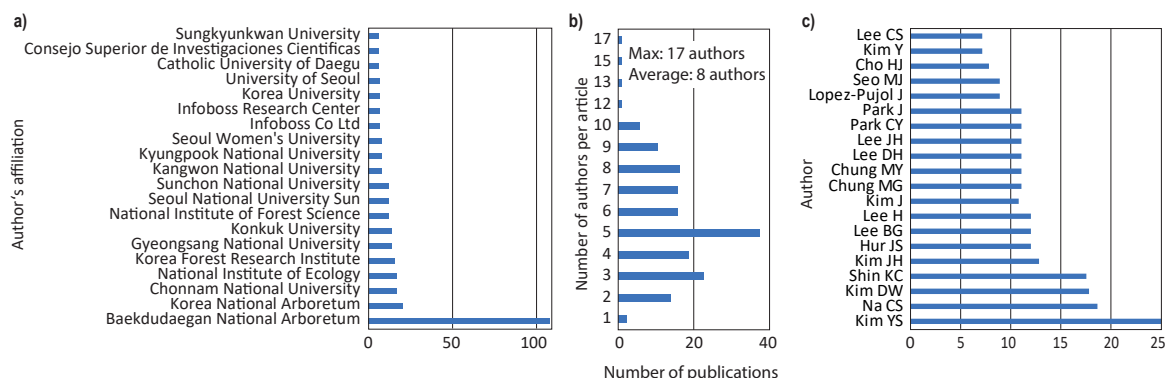


Figure 6 – Publication frequency of a) top 20 dominated institutions b) number of authors per article and c) top 20 authors.

It was observed that the keyword *conservation* was positioned as the central theme of networks linked to the terms *genetic diversity*, *biodiversity*, *diversity*, *phylogeography*, *biosphere reserve* and *Baekdudaegan*. These keywords focus on environmental and ecological research, with a specific emphasis on areas such as the Baekdudaegan Mountains, biodiversity and evolutionary genetics. Similarly, we graphed the co-authorship network (Figure 7b) to visualize the collaborative community, especially working on conservation and biodiversity topics related to the themes linked to the Korean Baekdudaegan environment. We observed that Kim, Yeong-su (Kim Y.S.), Kim, Dae Wook (Kim D.W.), Kim, Jeong Yoon (Kim J.Y.) and Woo H.S. were serving as central nodes and they connected with other authors, forming a collaborative research group. Other researchers such as Kim Jae Woo (Kim J.W.) and Lee, Andosung (Lee A.S.) led to smaller clusters within the network.

Topic visualization

We observed extreme values of the four metrics, Griffiths 2004, Cao 2009, Arun 2010 and Deveaud 2014 on the LDA model (Figure 8). Arun 2010 and Griffiths 2004 demonstrated a better fit when increasing the number of topics (K). Good convergence was observed in Cao 2009 and Deveaud 2014 at $K=6$ that we proceed in LDA modelling for further analysis.

The thematic delineations were examined to facilitate a deeper understanding of the multifaceted dimensions of research within the examined literature. We found six distinct topics, each characterized by dominant terms encapsulating key thematic elements. Word clouds depicting the top 100 dominant words for each topic were visualized (Figure 9). The first topic exhibited a prevalence of terms such as *activity*, *plant*, *cell*, *enzyme* and *production*, indicative of a focus on biochemical processes and productivity within plant systems. In the second topic, prominent terms

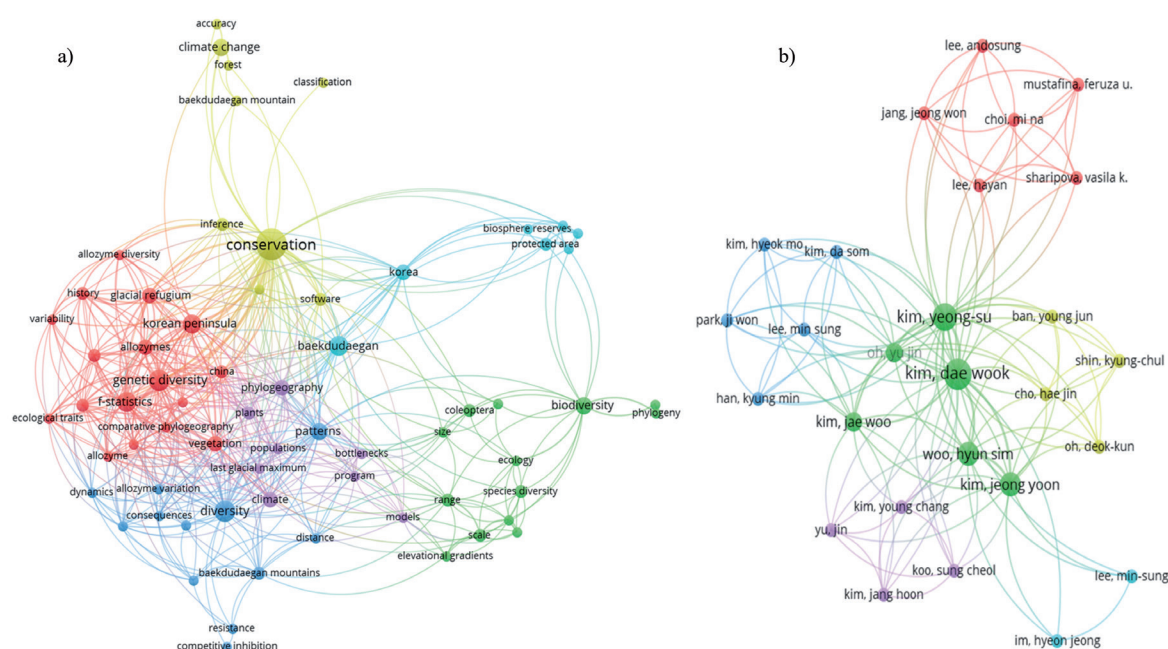


Figure 7 – Network visualization of (a) co-occurrence of keywords and (b) co-authorship relationships (node colors represent the thematic groups, where node size corresponds to the frequency of occurrence (for keywords) or the number of publications (for authors) and link thickness shows the strength of associations, based on co-occurrence frequency).

included *forest*, *species*, *vegetation*, *area* and *soil*, suggesting an emphasis on ecosystem dynamics and landscape-level interactions within forested environments. The third topic was characterized by terms such as *species*, *genome*, *region*, *gene* and *taxon*, suggesting a concentration on the genomic and taxonomic aspects of biodiversity. Conversely, the fourth topic featured terms like *seed*, *germination*, *temperature*, *treatment* and *dormancy*, indicating a focus on the dynamics and mechanisms governing seed behaviour under varying environmental conditions. The fifth topic highlighted terms such as *genetic*, *population*, *species*, *Baekdudaegan* and *diversity*, suggesting a thematic exploration of genetic diversity and population dynamics, particularly within mountainous regions. Lastly, the sixth topic was marked by words such as *species*, *change*, *plant*, *climate* and *pattern*, indicating an investigation into the impacts of climate change on plant communities and ecological patterns. Through careful examination of the word distribution within each topic, we synthesized the overarching themes into six categories: bioactive plant components, forest ecosystem restoration, genomic diversity, seed germination dynamics, mountain biodiversity and climate-induced changes.

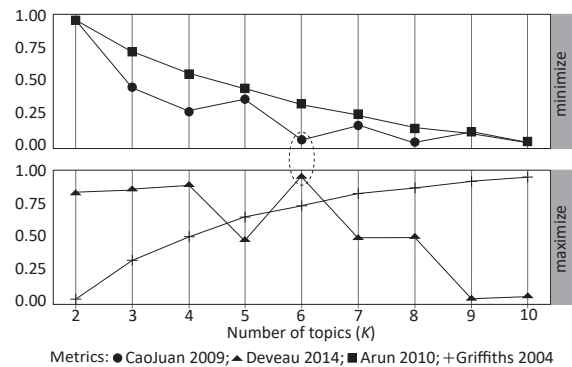


Figure 8 – Estimation of the optimal number of topics (six topics) based on LDA model quality using four measures: Griffiths (2004), CaoJuan (2009), Arun (2010), and Deveau (2014). The top panel shows metrics (CaoJuan (2009) and Arun (2010)) that should be minimized, as lower values indicate better topic separation. The bottom panel presents metrics (Griffiths (2004) and Deveau (2014)) that should be maximized, as higher values reflect improved topic coherence and model performance. Based on these measures, six topics were selected as the optimal number.

Table 1 – Top ten most-cited research and their findings linking to Baekdudaegan indexed in WoS until 27 December 2023. TC: Total citations; C/Y: Citations per year

Title (Type)	Key findings	TC (C/Y)
Altitudinal Patterns of Plant Species Richness in the Baekdudaegan Mountains, South Korea: Mid-domain Effect, Area, Climate and Rapoport's rule (Article) (Lee et al. 2013)	Hump-shaped species richness pattern in the Baekdudaegan mountains where plant richness is influenced by mid-domain effect (MDE) and climatic factors.	41 (3.42)
The role of the Baekdudaegan (Korean Peninsula) as a major glacial refugium for plant species: A priority for conservation. (Review) (Chung et al. 2017).	The Baekdudaegan mountains hypothetically served as a glacial refugium and have genetic diversity of plant populations. They are also regarded as the holy mountain system of Korea.	40 (5)
Environmental drivers of patterns of plant diversity along a wide environmental gradient in Korean temperate forests. (Article) (Lee & Chun 2016).	Plant species distribution in the Baekdudaegan mountains follows diverse elevational patterns, with regional area, habitat heterogeneity and climate being significant influencing factors.	26 (2.89)
Incorporating differences between genetic diversity of trees and herbaceous plants in conservation strategies. (Article) (Chung et al. 2020)	Along the Baekdudaegan mountain network, forest trees exhibit higher genetic diversity at both population and species levels compared to annuals and herbaceous perennials and herbaceous plant populations are low and vulnerable to the loss of genetic diversity.	23 (4.6)
The Korean Baekdudaegan mountains: A glacial refugium and a biodiversity hotspot that needs to be conserved. (Article) (Chung et al. 2018a).	The Baekdudaegan, a biodiversity hotspot with cultural and spiritual significance, requires concerted conservation and cooperation efforts for its sustainable management.	23 (3.29)
Large effective population sizes and high levels of gene flow between subpopulations of <i>Lilium cernuum</i> (Liliaceae). (Article) (Chung & Chung 2014).	In the Baekdudaegan mountain region, there are high levels of within-population genetic variation and low between-population genetic differentiation. Cold-adapted species are found in large refugial areas.	16 (1.45)
Were the main mountain ranges in the Korean Peninsula a glacial refugium for plants? Insights from the congeneric pair <i>Lilium cernuum</i> – <i>Lilium amabile</i> . (Article) (Chung et al. 2014).	The Baekdudaegan mountain range serves as glacial refugia for boreal plant species, providing stable habitats and supporting large population sizes.	15 (1.36)
Potential distribution of amphibians with different habitat characteristics in response to climate change in South Korea. (Article) (H.W. Kim et al. 2021)	Amphibian species face heightened vulnerability to climate change, anticipating significant declines and potential extinctions. Their habitats are projected to concentrate along the Baekdudaegan mountain range and in the southeastern region, demanding urgent conservation action.	14 (3.5)
Comparison of genetic variation between northern and southern populations of <i>Lilium cernuum</i> (Liliaceae): Implications for Pleistocene refugia. (Article) (Chung et al. 2018b).	There is genetic diversity of boreal or temperate plant species in the southern glacial refugium in the Baekdudaegan mountains and in the northern area (Northeast China and the Russian Far East).	14 (2)
Driven to the edge: Species distribution modeling of a clawed salamander (Hynobiidae: <i>Onychodactylus koreanus</i>) predicts range shifts and drastic decrease of suitable habitats in response to climate change. (Article) (Y. Shin et al. 2021)	Climate change significantly impacts amphibian diversity, resulting in a northeastward shift of their habitat suitability towards high-elevation areas along the central ridge of the Baekdudaegan range, as it serves as a climatic refugium for the species.	12 (3)

Overall, research linked to the BMR is an increasing trend, largely due to its critical role in conservation and ecosystem management. The research is predominantly aligned with SDG 15, particularly in the field of plant science, and is largely conducted by the Baekdudaegan National Arboretum. The unsupervised topics extracted from the LDA model also revealed a strong emphasis on conservation, biodiversity and ecosystem services. Given the growing importance of the BMR and similar protected mountain areas, it is crucial to continue research and systematically to monitor their conservation status.

Discussion

Mountains are special places to be protected for sustainable conservation (Kollmair et al. 2005). They serve as habitats for species worldwide, emphasizing the importance of effective management strategies and monitoring such indicators as conservation status, threats, socioeconomic context and social valuation and perception (Chung et al. 2017; Schmeller et al. 2022; Shin 2004). Representing approximately a quarter of the world's land area across 139 countries, mountains are integral landscapes of the planet, necessitating research to understand their systems and current threats (Kohler et al. 2012; Ng 2022; Schmeller et al. 2022). The inclusion of a dedicated mountain chapter in UN Agenda 21 (United Nations 1992, Chapter 13: Managing Fragile Ecosystems: Sustainable Mountain Development) catalysed significant progress in mountain research, leading to the establishment of global mountain institutions such as the Global Moun-

tain Biodiversity Assessment Network, the Mountain Research Initiative, the International Centre for Integrated Mountain Development and the Mountain Partnership (Bhandary 2015; Grumbine & Xu 2021; Price et al. 2021). These institutions facilitate interdisciplinary research, knowledge exchange with local communities and awareness-raising efforts regarding mountain ecosystems and their services at both local and global levels. As global and local interest in mountain research continues to rise, literature centred on mountainous regions has experienced a notable surge (Gurgiser et al. 2022; Ng 2022). Research linked to local mountains highlights indigenous conservation practices, traditional ecological knowledge and community-based management strategies that support sustainable conservation efforts (Farina & Mullet 2024; Garrard 2012; Palau & Claramunt-López 2024; Xu et al. 2023). Acknowledging the assistance of bibliometric studies and LDA modelling to identify major research concerns and guide future investigations, the current study examined the insights from 165 research papers linked to the Korean BMR obtained from the WoS core collection database.

The BMR in South Korea occupies a significant position in ecological research because of its designation as a biodiversity hotspot and its role as a crucial glacial refugium refuge for plant species (Chung et al. 2018a; Shin et al. 2021). In recent years, an increasing number of studies have focused on various facets of this mountain range, including changes in forest landscapes, village characteristics, plant species diversity and the impacts of climate change on species distribution (Choi et al. 2006; Dhakal et al. 2023; Kim et al.

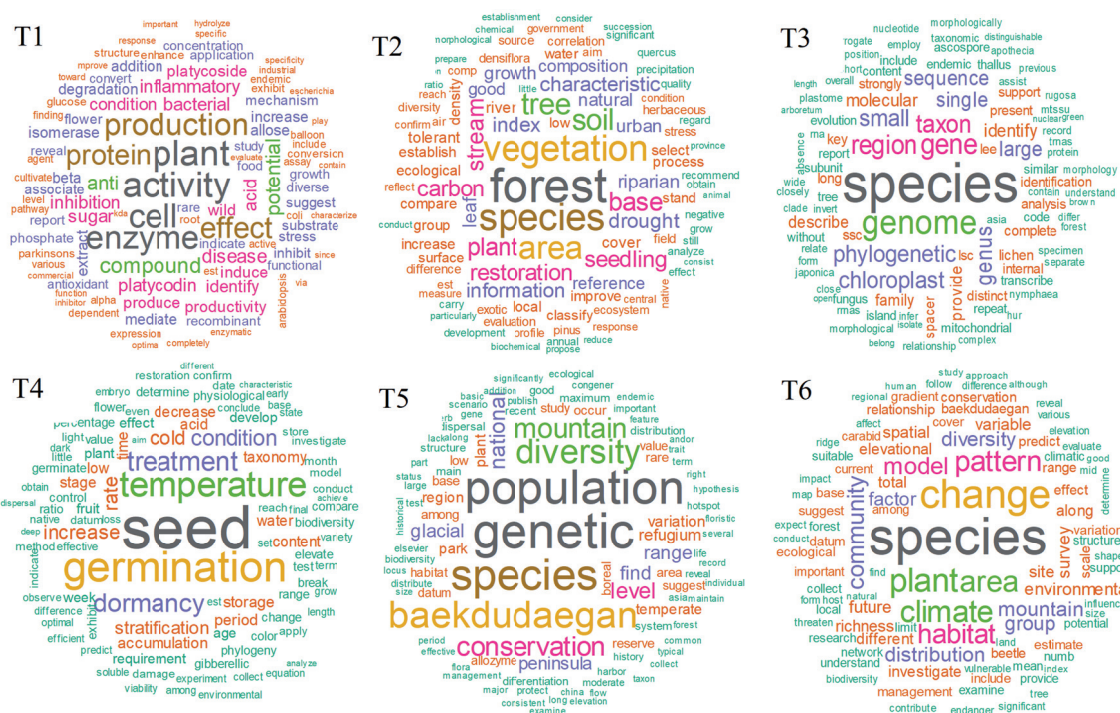


Figure 9 – Visualization of the 100 dominant words within each topic (T1 to T6). The size of each word represents its dominance based on frequency, whereas the colours are used to enhance readability and differentiate words visually.

2021; Kim et al. 2023; Park et al. 2022, 2017; Shin et al. 2021). We also observed that the most influential highly influenced literature focused on the species distribution, genetic diversity and glacial refugium characteristics of the BMR and the annual changes in year wise dominant thematic words within the literature highlighted the genetics, species, population and seeds as current focal points.

Mountain ecosystem research is increasing globally (Gurgiser et al. 2022; Liu et al. 2022). We also observed that Baekdudaegan-linked research has exhibited an increasing pattern, but one that appears to be in the early stages of research diffusion (Rogers 1983). The publication of Baekdudaegan-linked research followed an exponential trajectory with a rate of 0.34 WoS indexed publications in a year. This increasing trend could be of growing interest in investigating the ecological, cultural and socioeconomic dimensions of the study region. Forest and mountain-related research is aligned with multiple disciplines of SDGs (Baumgartner 2019; Timko et al. 2018). We observed that the alignment of the examined literature was dominated by SDG 15 with a particular focus on biodiversity conservation and ecosystem management (UNDP 2021). By aligning research efforts with the broader framework of the SDGs, Baekdudaegan-linked research can contribute significantly to sustainable development outcomes, while fostering interdisciplinary collaboration and knowledge exchange (Bettencourt & Kaur 2011; Zimmermann et al. 2018).

We observed the research landscape linked to the BMR, dominant latent words, SDG alignments, research categories, contributing authors, institutions and LDA topics. This would support collaborative effort on local as well as global mountain studies. Similar WoS categories such as plant science and ecology were the leading clusters within Baekdudaegan studies as highlighted in the previous mountain-related research (Gurgiser et al. 2022; Liu et al. 2021; Mengist et al. 2020). Understanding research-contributing institutions supports resource and knowledge sharing (Mufungizi et al. 2023; Shen et al. 2023; Subramanyam 1983). The Baekdudaegan National Arboretum is the predominant contributing institution that has conducted numerous projects focused on in- and ex-situ conservation, and the use of biological resources (Baekdudaegan National Arboretum 2024). In addition, we employed visualization techniques to elucidate keyword networks, facilitating the identification of interdisciplinary keywords and revealing link-strength distributions (Van Eck & Waltman 2018).

The LDA model was fitted to the keywords and abstracts, suggesting that six topics provided insights into the current research trends in the forest dynamics of the BMR. The first LDA-modelled topic comprised words related to biochemical pathways and their impact on industries and health supplements. The second topic included complex forest ecosystems and myriad indices defining ecological balance related to compre-

hensive evaluations and strategies for forest conservation. The third topic consisted of genomic taxonomy and molecular identification across species, using molecular identification techniques related to deciphering evolutionary relationships and preserving diverse genetic lineages. The fourth topic covered seed physiology, germination dynamics and environmental factors related to agricultural practices, ecosystem restoration and conservation. The fifth topic covered mountain biodiversity conservation, genetic diversity and population genetics related to sustaining ecological balance in mountainous landscapes. The sixth topic comprised the species response to climate change and adoption, focusing on habitat distribution which is inclined to stress the urgency of setting strategies to mitigate the threats of ongoing environmental challenges.

Based on the distribution of dominant words in LDA, we have proposed six potential broad topics: 1) bioactive plant components, biochemical significance and disease; 2) forest ecosystem restoration by contrasting urban-natural dynamics; 3) genomic taxonomy for molecular insights and evolutionary significance; 4) seed germination dynamics and environmental factors; 5) mountain biodiversity and conservation strategies; and 6) climate-induced changes and species resilience that contribute to ecosystem services, ecological policy research and SDGs (Altaweel et al. 2019; Liu et al. 2022; Mengist et al. 2020). The WoS categorization of the analysed literature predominantly aligned with SDG 15. Additionally, the results from the LDA model reveal that topics primarily associated with SDG 15 often intersect with other SDGs, depending on their emphasis on aspects such as health, climate, ecosystems and sustainability. This convergence between the WoS report and LDA findings validates the consistent association of the topics with the SDGs, highlighting the importance of BMR research for sustainable biodiversity conservation and land management.

Although the current study explores research insights gained from the publication trends and thematic structure of BMR linked studies, it has several limitations. The inherent limitations of bibliometric and LDA topic modelling methods is their inclination towards data-driven research outputs rather than specific content. This study relied only on the WoS database, missing valuable information from other sources such as Google Scholar, Scopus and the Korean Science Index (KSCI). Although there are biases, bibliometric studies explore various insights such as authors, institutional and document contents (Donthu et al. 2021). We examined publication trends, article keyword networks, contributing institutions and the 2030 agenda for the SDG alignment, as suggested in the WoS database. In-depth research on science mapping, research performance and network analysis of all scholarly sources can influence the results. Furthermore, the study employs LDA modelling, defining six topics concerning current research trends that may

oversimplify complex phenomena. Comparing different metrics and topic modelling approaches, such as non-negative matrix factorization (NMF) and latent semantic analysis (LSA), could achieve a more robust and nuanced examination of scientific content related to the Baekdudaegan mountain network.

This is a case study involving the Korean Peninsula. We have identified some noticeable areas for future research linking mountain ecosystems such as climate change impact and adaptation, biodiversity and conservation, ecosystem services, forest dynamics and restoration ecology, genetic diversity and evolution, anthropological impact and conflict mitigation and ecological monitoring and data generation that can enhance understanding and management (Blatter et al. 2017; Canedoli et al. 2024; Shen et al. 2023). The increasing BMR research centred on biodiversity and nature has contributed to South Korea achieving the seventh position on the 2023 Nature Index (NI). NI is examined based on research publications in 82 natural science journals selected by an independent group of researchers (Woolston 2023). Upgrading the national NI rank would be a future goal for researchers in this region. The Man and Biosphere Program (MAB) of UNESCO has established fundamental principles and methodologies for proficiently governing the Biosphere Reserve, necessitating ecological research and conservation endeavours to uphold biodiversity and sustainable practices (Batisse 1982). In the keywords network of this study, terms such as *protected area*, *Biosphere Reserve*, and *Korea* are cohesively clustered, yet challenges persist in the incorporation of the BMR into Biosphere Reserve regions (Lee 2019). Examining social issues and contributing factors will be the next research path. Continuous research and development within the Baekdudaegan networks hold promise for achieving diverse objectives, serving as a crucial pathway towards enhanced ecological understanding, conservation efficacy and sustainable resource management. Future study could also be bibliometrically comparative, employing topic modelling on other locally and globally protected, as well as non-protected, mountain ranges in the world.

Conclusions

Mountain research has focused on the unique species and ecological characteristics of different mountain ranges worldwide. This study has explored insights into the research surrounding the BMR in the Korean peninsula applying bibliometric analysis and LDA modelling to the metadata of 165 WoS-indexed articles. We examined research trends and their alignment with WoS categories and the 2030 agenda for UN SDGs, co-occurrence networks, collaborating institutions and LDA topics. Overall, the results of the hybrid bibliometric analysis and topic modelling approaches of this study benefit local ecosystems and open avenues for studies across global mountain

networks. Beyond academia, this study encourages co-operation and knowledge exchange among researchers and policymakers to address challenging environmental issues and to promote sustainable management practices.

Funding

This research was supported by grants from the National Institute of Forest Science (Grant No. FE0100-2023-01-2023), the Republic of Korea, and the National Research Foundation of Korea (NRF-2022R1A6A3A01086781) funded by the Government of the Republic of Korea.

Data availability

Raw data will be available at <https://Figureshare.com/s/f832b3cc6f17ce84207c> and as Supplementary File S1.

References

- Abdelrazek, A., Y. Eid, E. Gawish, W. Medhat & A. Hassan 2023. Topic modeling algorithms and applications: A survey. *Information Systems* 112: 102131. <https://doi.org/10.1016/J.IS.2022.102131>
- Altaweel, M., C. Bone & J. Abrams 2019. Documents as data: A content analysis and topic modeling approach for analyzing responses to ecological disturbances. *Ecological Informatics* 51: 82–95. <https://doi.org/10.1016/j.ecoinf.2019.02.014>
- Arroyave, F.J., J. Jenkins, S. Shackelton, B. Jackson & A.M. Petersen 2024. Research alignment in the U.S. national park system: Impact of transformative science policy on the supply and demand for scientific knowledge for protected area management. *Journal of Environmental Management* 357: 120699. <https://doi.org/10.1016/J.JENVMAN.2024.120699>
- Arun, R., V. Suresh, C.E.V. Madhavan & M.N. Murty 2010. On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. *Lecture Notes in Computer Science* 6118: 391–402. https://doi.org/10.1007/978-3-642-13657-3_43
- Baekdudaegan National Arboretum 2024. An Overview of Research. Available at: <https://www.bdna.or.kr/eng/contents/content?menuId=M0021&upMenuId=M0003&cs0> (accessed 19/06/2024).
- Batisse, M. 1982. The biosphere reserve: A tool for environmental conservation and management. *Environmental Conservation* 9: 101–111. <https://doi.org/10.1017/S0376892900019937>
- Baumgartner, R.J. 2019. Sustainable development goals and the forest sector – A complex relationship. *Forests* 10(2): 152. <https://doi.org/10.3390/F10020152>
- Bettencourt, L.M.A. & J. Kaur 2011. Evolution and structure of sustainability science. *PNAS* 108(49): 19540–19545. <https://doi.org/10.1073/PNAS.1102712108>

- Bhandary, R.R. 2015. Coalition strategies in the climate negotiations: an analysis of mountain-related coalitions. *International Environmental Agreements: Politics, Law and Economics* 17: 173–190. <https://doi.org/10.1007/S10784-015-9313-6>
- Blattert, C., R. Lemm, O.Thees, M.J. Lexer & M. Hanewinkel 2017. Management of ecosystem services in mountain forests: Review of indicators and value functions for model based multi-criteria decision analysis. *Ecological Indicators* 79: 391–409. <https://doi.org/10.1016/j.ecolind.2017.04.025>
- Blei, D.M. & J.D. Lafferty 2005. Correlated topic models. In: Weiss, Y., B. Schölkopf & J. Platt (eds.), *Advances in Neural Information Processing Systems*: 147–154.
- Blei, D.M., A.Y. Ng & M.I. Jordan 2003. Latent Dirichlet allocation. *The Journal of Machine Learning Research* 3: 993–1022.
- Calvo, D.C., H.J. Luna, J.A. Arango, C.I. Torres & B.E. Rittmann 2022. Determining global trends in syn-gas fermentation research through a bibliometric analysis. *Journal of Environmental Management* 307: 114522. <https://doi.org/10.1016/J.JENVMAN.2022.114522>
- Cameron, A.C. & F.A.G. Windmeijer 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* 77: 329–342. [https://doi.org/10.1016/S0304-4076\(96\)01818-0](https://doi.org/10.1016/S0304-4076(96)01818-0)
- Canedoli, C., N. Rota, I.N. Vogiatzakis, A. Zanchi, M. Drius, H. Nagendra & E. Padoa-Schioppa 2024. Review of indicators for mountain ecosystem services: Are the most frequently used also the best? *Ecological Indicators* 166: 112310. <https://doi.org/10.1016/J.ECOLIND.2024.112310>
- Cao, J., T. Xia, J. Li, Y. Zhang & S. Tang 2009. A density-based method for adaptive LDA model selection. *Neurocomputing* 72: 1775–1781. <https://doi.org/10.1016/J.NEUCOM.2008.06.011>
- Cárdenas, M.R., M. Clüsener-Godt, G. Köck & V. Van Rysseghem 2018. *Man and the Biosphere Programme: Biennial activity report 2016-2017*. UNESCO Man and Biosphere Programme, Lammerhuber, Baden (Austria). <https://doi.org/https://unesdoc.unesco.org/ark:/48223/pf0000266210>
- Cheng, X., C. Shuai, J. Liu, J. Wang, Y. Liu, W. Li & J. Shuai 2018. Topic modelling of ecology, environment and poverty nexus: An integrated framework. *Agriculture, Ecosystems & Environment* 267: 1–14. <https://doi.org/10.1016/J.AGEE.2018.07.022>
- Cho, K.-H., H.-J. Kim, M.-K. Jang, J.-H. Ryu, D. Yoo, H.-E. Kang & J.-Y. Park 2022. Detection of African swine fever at an abattoir in South Korea, 2020. *Veterinary Sciences* 9: 150. <https://doi.org/10.3390/vetsci9040150>
- Cho, W. & B.K. Chun 2015. *Restoration of the Baekdudaegan mountains in the Republic of Korea*. Unasyuva 66.
- Choi, K.H., J.S. Kim, Y.S. Kim, M.A. Yoo & T.S. Chon 2006. Pattern detection of movement behaviors in genotype variation of *Drosophila melanogaster* by using self-organizing map. *Ecological Informatics* 1: 219–228. <https://doi.org/10.1016/J.ECOINF.2005.12.002>
- Choi, Y.-K. 2004. Baekdudaegan, the central axis of the Korean Peninsular: The path toward management strategies regarding to its concepts. In: Hong, S.-K., J.A. Lee, B.-S. Ihm, A. Farina, Y. Son, E.-S. Kim & J.C. Choe (eds.), *Ecological Issues in a Changing World*: 355–383. https://doi.org/10.1007/978-1-4020-2689-8_22
- Chung, M.Y. & M.G. Chung 2014. Large effective population sizes and high levels of gene flow between subpopulations of *Lilium cernuum* (Liliaceae). *Biochemical Systematics and Ecology* 54: 354–361. <https://doi.org/10.1016/J.BSE.2014.03.009>
- Chung, M.Y., M.G. Chung, J. López-Pujol, M.X. Ren, Z.Y. Zhang & S.J. Park 2014. Were the main mountain ranges in the Korean peninsula a glacial refugium for plants? Insights from the congeneric pair *Lilium cernuum* – *Lilium amabile*. *Biochemical Systematics and Ecology* 53: 36–45. <https://doi.org/10.1016/J.BSE.2013.12.019>
- Chung, M.Y., J. López-Pujol & M.G. Chung 2017. The role of the Baekdudaegan (Korean peninsula) as a major glacial refugium for plant species: A priority for conservation. *Biological Conservation* 206: 236–248. <https://doi.org/10.1016/J.BIOCON.2016.11.040>
- Chung, M.Y., S. Son, S. Herrando-Moraira, C.Q. Tang, M. Maki, Y.D. Kim, J. López-Pujol, J.L. Hamrick & M.G. Chung 2020. Incorporating differences between genetic diversity of trees and herbaceous plants in conservation strategies. *Conservation Biology* 34: 1142–1151. <https://doi.org/10.1111/COBI.13467>
- Chung, M.Y., S. Son, S., G.U. Suh, S. Herrando-Moraira, C.H. Lee, J. López-Pujol & M.G. Chung 2018a. The Korean Baekdudaegan Mountains: A glacial refugium and a biodiversity hotspot that needs to be conserved. *Frontiers in Genetics* 9: 395213. <https://doi.org/10.3389/fgene.2018.00489>
- Chung, M.Y., S.H. Vu, J. López-Pujol, S. Herrando-Moraira, S. Son, G.U. Suh, H.T.Q. Le & M.G. Chung 2018b. Comparison of genetic variation between northern and southern populations of *Lilium cernuum* (Liliaceae): Implications for Pleistocene refugia. *PLOS One* 13. <https://doi.org/10.1371/JOURNAL.PONE.0190520>
- Cobelli, N. & S. Blasi 2024. Combining topic modeling and bibliometric analysis to understand the evolution of technological innovation adoption in the healthcare industry. *European Journal of Innovation Management* 27: 127–149. <https://doi.org/10.1108/EJIM-06-2023-0497>
- Cobo, M.J., A.G. López-Herrera, E. Herrera-Viedma & F. Herrera 2011. Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology* 62: 1382–1402. <https://doi.org/10.1002/asi.21525>
- Da Kuang, D., J. Choo & H. Park 2015. Nonnegative matrix factorization for interactive topic modeling and document clustering. In: Kuang, D. J. Choo & H. Park

- (eds.), *Partitional clustering algorithms*: 215–243. https://doi.org/10.1007/978-3-319-09259-1_7
- Daud, A., J. Li, L. Zhou & F. Muhammad 2010. Knowledge discovery through directed probabilistic topic models: A survey. *Frontiers of Computer Science in China* 4: 280–301. <https://doi.org/10.1007/S11704-009-0062-Y>
- Davies, N.B., J.R. Krebs & S.A. West 2012. An introduction to behavioural ecology, 4th edn. Oxford.
- Deveaud, R., E. SanJuan & P. Bellot 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. *HAL open science*. <https://doi.org/10.3166/dn.17.1.61-84>
- Devlin, J., M.W. Chang, K. Lee & K. Toutanova 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Burstein, J., C. Doran & T. Solorio (eds.), *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies – Proceedings of the Conference*. Association for Computational Linguistics (ACL): 4171–4186.
- Dhakal, T., K.H. Cho, S.J. Kim & M.S. Beon 2023. Modeling decline of mountain range forest using survival analysis. *Frontiers in Forests and Global Change* 6: 1183509. <https://doi.org/10.3389/FFGC.2023.1183509>
- Donthu, N., S. Kumar, D. Mukherjee, N. Pandey & W.M. Lim 2021. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* 133: 285–296. <https://doi.org/10.1016/J.JBUSRES.2021.04.070>
- Du, H., Y. Guo, Z. Lin, Y. Qiu & X. Xiao 2021. Effects of the joint prevention and control of atmospheric pollution policy on air pollutants – A quantitative analysis of Chinese policy texts. *Journal of Environmental Management* 300: 113721. <https://doi.org/10.1016/J.JENVMAN.2021.113721>
- Ellegaard, O. & J.A. Wallin 2015. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* 105: 1809–1831. <https://doi.org/10.1007/S11192-015-1645-Z>
- Falagas, M.E., E.I. Pitsouni, G.A. Malietzis & G. Pappas 2008. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal* 22: 338–342. <https://doi.org/10.1096/FJ.07-9492LSF>
- Farina, A. & T.C. Mullet 2024. Sonotope patterns within a mountain beech forest of Northern Italy: a methodological and empirical approach. *Frontiers in Ecology and Evolution* 12. <https://doi.org/10.3389/fevo.2024.1341760>
- Fellows, I. 2022. Word Clouds. Available at: <https://cran.r-project.org/web/packages/wordcloud/wordcloud.pdf> (Accessed 13/03/2025)
- Garrard, R. 2012. The effects of climate change on people in the Andes: Melting glaciers means drought, tension and conflict for Peru's mountain people in and around Huascarán National Park. *eco.mont - Journal on protected mountain areas research and management* 4: 35–40. <https://doi.org/10.1553/eco.mont-4-1s35>
- Gaviria-Marin, M., J.M. Merigo & S. Popa 2018. Twenty years of the Journal of Knowledge Management: a bibliometric analysis. *Journal of Knowledge Management* 22: 1655–1687. <https://doi.org/10.1108/JKM-10-2017-0497>
- Griffiths, T.L. & M. Steyvers 2004. Finding scientific topics. *PNAS* 101: 5228–5235. <https://doi.org/10.1073/PNAS.0307752101>
- Grisales, A.M., S. Robledo & M. Zuluaga 2023. Topic modeling: Perspectives from a literature review. *IEEE Access* 11: 4066–4078. <https://doi.org/10.1109/ACCESS.2022.3232939>
- Grumbine, R.E. & J. Xu 2021. Mountain futures: Pursuing innovative adaptations in coupled social–ecological systems. *Frontiers in Ecology and Environment* 19(6): 342–348. <https://doi.org/10.1002/fee.2345>
- Grün, B. & K. Hornik 2011. Topicmodels: An R package for fitting topic models. *Journal of Statistical Software* 40: 1–30. <https://doi.org/10.18637/jss.v040.i13>
- Gunya, A. & F. Sarmiento 2022. Report on the International Symposium of Mountain Studies, part of the 34th International Geographical Congress, Istanbul (16–20 August 2021). *eco.mont - Journal on protected mountain areas research and management* 14: 47–48. <https://doi.org/10.1553/ECO.MONT-14-1S47>
- Gurgiser, W., M.F. Price, I.F. Juen, C. Körner, M. Bahn, B. Gerns, M. Meyer, K. Nicolussi, U. Tappeiner & S. Mayr 2022. Rising slopes – Bibliometrics of mountain research 1900–2019. *PLOS One* 17. <https://doi.org/10.1371/JOURNAL.PONE.0273421>
- Hayes, P. 2010. Sustainable security in the Korean peninsula: Envisioning a Northeast Asian biodiversity corridor. *The Korean Journal of International Studies* 12: 1–14. <https://doi.org/10.14731/KJIS.2010.12.8.2.197>
- Je-Hun, R. & W. Doo-Hee 2013. The modern production of multiple meanings of the Baekdudaegan mountain system. *Korea Journal* 53: 103–132. <https://doi.org/10.25024/kj.2013.53.3.103>
- Jelodar, H., Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li & L. Zhao 2019. Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications* 78: 15169–15211. <https://doi.org/10.1007/S11042-018-6894-4>
- Jiang, H.C., M.S. Qiang & P. Lin 2016. Finding academic concerns of the Three Gorges Project based on a topic modeling approach. *Ecological Indicators* 60: 693–701. <https://doi.org/10.1016/J.ECOLIND.2015.08.007>
- Jung, H.-J. 2015. *Landscape as heritage: Towards a conservation framework for scenic sites in Korea* (PhD thesis). University of Sheffield.
- Kim, H.W., P. Adhikari, M.H. Chang & C. Seo 2021. Potential distribution of amphibians with different habitat characteristics in response to climate change in South Korea. *Animals* 11(8): 2185. <https://doi.org/10.3390/ANI11082185>

- Kim, K., E. Ha, C. Lee & K. Ahn 2021. Analysis of the forest landscape changes in Mt. Jiri Portion of Baekdudaegan. *Journal of the Korean Cartographic Association* 21: 39–52. <https://doi.org/10.16879/jkca.2021.21.1.039>
- Kim, S.M. 2011. Study on application of IUCN Management Category System on Baekdudaegan Protected Area. *Journal of Korean Society of Forest Science* 100: 494–503. <https://doi.org/10.14578/JKFS.2011.100.3.23>
- Kim, T.S., T. Dhakal, S.H. Kim, J.H. Lee, S.J. Kim & G.S. Jang 2023. Examining village characteristics for forest management using self- and geographic self-organizing maps: A case from the BMR network in Korea. *Ecological Indicators* 148: 110070. <https://doi.org/10.1016/j.ecolind.2023.110070>
- Kohler, T., J. Pratt, B. Dabarbeiux, J. Balsiger, G. Rudaz & D. Maselli 2012. *Sustainable mountain development green economy and institutions. From Rio 1992 to Rio 2021 and beyond. Final Draft for Rio 2012*. <https://doi.org/10.7892/boris.17634>
- Kollmair, M., G.S. Gurung, K. Hurni & D. Maselli 2005. Mountains: Special places to be protected? An analysis of worldwide nature conservation efforts in mountains. *International Journal of Biodiversity Science & Management* 1: 181–189. <https://doi.org/10.1080/17451590509618091>
- Körner, C. 2004. Mountain biodiversity, its causes and function. *Ambio* 33(13): 11–17. <https://doi.org/10.1007/0044-7447-33.sp13.11>
- Kratzer, A. 2018. Biosphere reserves research: A bibliometric analysis. *eco.mont - Journal on protected mountain areas research and management* 10(2): 36–49. <https://doi.org/10.1553/eco.mont-10-2s36>
- Laureate, C.D.P., W. Buntine & H. Linger 2023. A systematic review of the use of topic models for short text social media analysis. *Artificial Intelligence Review* 56: 14223–14255. <https://doi.org/10.1007/S10462-023-10471-X>
- Lee, C.B. & J.H. Chun 2016. Environmental drivers of patterns of plant diversity along a wide environmental gradient in Korean temperate forests. *Forests* 7(1): 19. <https://doi.org/10.3390/F7010019>
- Lee, C.B., J.H. Chun, H.K. Song & H.J. Cho 2013. Altitudinal patterns of plant species richness on the Baekdudaegan Mountains, South Korea: Mid-domain effect, area, climate, and Rapoport's rule. *Ecological Research* 28(1): 67–79. <https://doi.org/10.1007/S11284-012-1001-1>
- Lee, J.-H. 2019. Analyzing local opposition to biosphere reserve creation through semantic network analysis: The case of Baekdu mountain range, Korea. *Land Use Policy* 82: 61–69. <https://doi.org/10.1016/j.landusepol.2018.11.055>
- Li, K., J. Rollins & E. Yan 2018. Web of Science use in published research and review papers 1997–2017: A selective, dynamic, cross-domain, content-based analysis. *Scientometrics* 115: 1–20. <https://doi.org/10.1007/S11192-017-2622-5/TABLES/8>
- Lin, C. & Y. He 2009. Joint sentiment/topic model for sentiment analysis. *Proceedings of the 18th ACM Conference on Information and Knowledge Management*: 375–384. <https://doi.org/10.1145/1645953.1646003>
- Liu, W., J. Zheng, Z. Wang, R. Li & T. Wu 2021. A bibliometric review of ecological research on the Qinghai–Tibet Plateau, 1990–2019. *Ecological Informatics* 64: 101337. <https://doi.org/10.1016/j.ecoinf.2021.101337>
- Liu, W., Z. Wang, R. Li & T. Wu 2022. A bibliometric analysis of mountain ecosystem services, 2000–2019. *Environmental Science and Pollution Research* 29: 16633–16652. <https://doi.org/10.1007/s11356-021-16766-2>
- Marani, H.A. & E.P.S. Baumer 2023. A review of stability in topic modeling: Metrics for assessing and techniques for improving stability. *ACM Computing Surveys* 56(5): 1–32. <https://doi.org/10.1145/3623269>
- Mardones-Segovia, C., H.J. Choi, M. Hong, J.M. Wheeler & A.S. Cohen 2022. Comparison of estimation algorithms for latent Dirichlet allocation. *Springer Proceedings in Mathematics & Statistics* 393: 27–37. https://doi.org/10.1007/978-3-031-04572-1_3
- Mengist, W., T. Soromessa & G. Legese 2020. Ecosystem services research in mountainous regions: A systematic literature review on current knowledge and research gaps. *Science of The Total Environment* 702: 134581. <https://doi.org/10.1016/j.scitotenv.2019.134581>
- Min, K.H., Y. Son & K. Furuya 2019. Categorizing types of transition areas in biosphere reserves: A case study of the BMRs in South Korea. *International Review for Spatial Planning and Sustainable Development* 7: 83–100. https://doi.org/10.14246/IRSPSD.7.1_83
- Mufungizi, A.A., W. Musakwa & N. Chanza 2023. Shifting ecosystems, past, current, and emerging trends: A bibliometric analysis and systematic review of literature. *Ecological Indicators* 156: 111175. <https://doi.org/10.1016/J.ECOLIND.2023.111175>
- Murzintcev N. & N. Chaney 2022. *ldatuning: Tuning of the latent Dirichlet allocation models parameters*. CRAN.
- Negi, G.C.S., R.S. Rawal, S. Sharma, K. Kumar & P.P. Dhyan 2014. Need for strengthening mountain-specific research. *Current Science* 106(5): 659–661.
- Ng, S.L. 2022. Bibliometric analysis of literature on mountain tourism in Scopus. *Journal of Outdoor Recreation and Tourism* 40: 100587. <https://doi.org/10.1016/J.JORT.2022.100587>
- Ninkov, A., J.R. Frank & L.A. Maggio 2022. Bibliometrics: Methods for studying academic publishing. *Perspectives on Medical Education* 11(3): 173–176. <https://doi.org/10.1007/s40037-021-00695-4>
- Ortega, J.L. 2014. Influence of co-authorship networks in the research impact: Ego network analyses from Microsoft Academic Search. *Journal of Informetrics* 8(3): 728–737. <https://doi.org/10.1016/j.joi.2014.07.001>
- Palau, A.Z. & B. Claramunt-López 2024. Mountain research for sustainability: Where are we and where

- to go? *Sustainability Science* 19: 1693–1707. <https://doi.org/10.1007/S11625-024-01530-5>
- Pan, B., F. Zhang, X. Zhu, L. Huang, Y. Wu, J.Q. Tang & N.X. Feng 2024. Global trends and hotspots evolution in soil microplastic pollution research: A bibliometric analysis based on the Web of Science. *Ecological Indicators* 161: 111974. <https://doi.org/10.1016/J.ECOLIND.2024.111974>
- Park, C., J. Byeon, J. Park, S. Lee, B. Park, J. Kim & H. Woo 2022. Impact of climate change on species distribution of *Picea jezoensis* Carrière in Baekdudaegan using ensemble modeling. *Sensors and Materials* 34: 4629–4637. <https://doi.org/10.18494/SAM4163>
- Park, M.S. & H. Lee 2014. Forest policy and law for sustainability within the Korean peninsula. *Sustainability* 6(8): 5162–5186. <https://doi.org/10.3390/su6085162>
- Park, Y., J. Kim, T. Jang, H. Chae & Y. Takami 2017. Local climate mediates spatial and temporal variation in carabid beetle communities in three forests in Mount Odaesan, Korea. *Ecological Entomology* 42(2): 184–194. <https://doi.org/10.1111/een.12370>
- Payne, D., E.M. Spehn, M. Snethlage & M. Fischer 2017. Opportunities for research on mountain biodiversity under global change. *Current Opinion in Environmental Sustainability* 29: 40–47. <https://doi.org/10.1016/J.COSUST.2017.11.001>
- Price, M.F., T. Schaaf & M.R.C. Tomažič 2021. Mountain research in UNESCO's Man and the Biosphere Programme: The first five decades. *eco.mont - Journal on protected mountain areas research and management* 13: 93–101. <https://doi.org/10.1553/ECO.MONT-13-SIS93>
- R Core Team 2023. *RStudio | Open source & professional software for data science teams* - RStudio. RStudio.
- Rogers, E.M. 1983. *Diffusion of innovations*. New York.
- Schmeller, D.S., D. Urbach, K. Bates, J. Catalan, D. Cogălniceanu, M.C. Fisher, J. Friesen, L. Füreder, V. Gaube, M. Haver, D. Jacobsen, G. Le Roux, Y.P. Lin, A. Loyau, O. Machate, A. Mayer, I. Palomo, C. Plutzer, H. Sentenac, R. Sommaruga, R. Tiberti & W.J. Ripple 2022. Scientists' warning of threats to mountains. *Science of The Total Environment* 853: 158611. <https://doi.org/10.1016/J.SCITOTENV.2022.158611>
- Schmidt-Vogt, D., M. Foggin & C. Hergarten 2016. Strengthening mountain societies in Central Asia in a context of multidimensional change. *Mountain Research and Development* 36(3): 380–383. <https://doi.org/10.1659/MRD-JOURNAL-D-16-00101.1>
- Schroth, G., P. Laderach, J. Dempewolf, S. Philpott, J. Hagggar, H. Eakin, T. Castillejos, J.G. Moreno, L.S. Pinto, R. Hernandez, A. Eitzinger & J. Ramirez-Villegas 2009. Towards a climate change adaptation strategy for coffee communities and ecosystems in the Sierra Madre de Chiapas, Mexico. *Mitigation and Adaptation Strategies for Global Change* 14: 605–625. <https://doi.org/10.1007/s11027-009-9186-5>
- Shen, Z., Y. Tian, Y. Yao, W. Jiang, J. Dong, X. Huang, X. Wu, T.H. Farooq & W. Yan 2023. Ecological restoration research progress and prospects: A bibliometric analysis. *Ecological Indicators* 155: 110968. <https://doi.org/10.1016/j.ecolind.2023.110968>
- Shin, J.H. 2004. Management area and management strategy of Baekdudaegan. *Korean Journal of Environment and Ecology* 18: 197–204.
- Shin, S., J.H. Kim, J.H. Dang, I.S. Seo & B.Y. Lee 2021. Elevational distribution ranges of vascular plant species in the BMR, South Korea. *Journal of Ecology and Environment* 45: 7. <https://doi.org/10.1186/s41610-021-00182-1>
- Shin, Y., M.S. Min & A. Borzée 2021. Driven to the edge: Species distribution modeling of a clawed salamander (Hynobiidae: *Onychodactylus koreanus*) predicts range shifts and drastic decrease of suitable habitats in response to climate change. *Ecology and Evolution* 11(21): 14669–14688. <https://doi.org/10.1002/ECE3.8155>
- Solár, J. & V. Solár 2020. Land-cover change in the tatra mountains, with a particular focus on vegetation. *eco.mont - Journal on protected mountain areas research and management* 12(1): 15–26. <https://doi.org/10.1553/eco.mont-12-1s15>
- Stevens, K., P. Kegelmeyer, D. Andrzejewski & D. Buttlar 2012. Exploring topic coherence over many models and many topics. In: *EMNLP-CoNLL 2012 - 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Proceedings of the Conference*: 952–961. Association for Computational Linguistics.
- Subramanyam, K. 1983. Bibliometric studies of research collaboration: A review. *Journal of Information Science* 6(1): 33–38. <https://doi.org/10.1177/016555158300600105>
- Tang, L., G. Shao, Z. Piao, L. Dai, M.A. Jenkins, S. Wang, G. Wu, J. Wu & J. Zhao 2010. Forest degradation deepens around and within protected areas in East Asia. *Biological Conservation* 143(5): 1295–1298. <https://doi.org/10.1016/J.BIOCON.2010.01.024>
- Thaler, T., C. Clar, L. Junger & R. Nordbeck 2021. Opportunities and challenges for transdisciplinary research in flood risk management: Some critical reflections and lessons learnt for research on sustainability. *eco.mont - Journal on protected mountain areas research and management* 13: 42–47. <https://doi.org/10.1553/0X003C9DA2>
- Theobald, D.M., A.L. Jacob, P.R. Elsen, E.A. Beever, L. Ehlers & J. Hilty 2024. Evaluating ecosystem protection and fragmentation of the world's major mountain regions. *Conservation Biology* 38(3): e14240. <https://doi.org/10.1111/cobi.14240>
- Timko, J., P. Le Billon, H. Zerriffi, J. Honey-Rosés, I. de la Roche, C. Gaston, T.C. Sunderland & R.A. Kozak 2018. A policy nexus approach to forests and the SDGs: tradeoffs and synergies. *Current Opinion in Environmental Sustainability* 34: 7–12. <https://doi.org/10.1016/J.COSUST.2018.06.004>

UNDP 2021. What are the Sustainable Development Goals?. United Nations Development Programme. Available at: <https://www.undp.org/united-states/sustainable-development-goals> (accessed: 29/07/2022).

Van Eck, N.J. & L. Waltman 2018. *VOSviewer manual*. Leiden.

Vayansky, I. & S.A.P. Kumar 2020. A review of topic modeling methods. *Information Systems* 94: 101582. <https://doi.org/10.1016/J.IS.2020.101582>

Venturini, S. & S. Sikdar, F. Rinaldi, F. Tudisco & S. Fortunato 2024. Collaboration and topic switches in science. *Scientific Reports* 14: 1258. <https://doi.org/10.1038/s41598-024-51606-6>

Vlase, I. & T. Lähdesmäki 2023. A bibliometric analysis of cultural heritage research in the humanities: The Web of Science as a tool of knowledge management. *Humanities and Social Sciences Communications* 10: 84. <https://doi.org/10.1057/s41599-023-01582-5>

Wang, F.Y., J. Yang, X. Wang, J. Li & Q.L. Han 2023. Chat with ChatGPT on Industry 5.0: Learning and Decision-Making for Intelligent Industries. *IEEE/CAA Journal of Automatica Sinica* 10(4): 831–834. <https://doi.org/10.1109/JAS.2023.123552>

Woolston, C. 2023. Nature Index Annual Tables 2023: China tops natural-science table. Nature: Epub ahead of print. <https://doi.org/10.1038/d41586-023-01868-3>

Xu, J., P. Yan, Z. Liu, H. Qin & A. Jiang 2023. Avitourism as an aspect of sustainable mountain development: a case study from Southern China. *eco.mont - Journal on protected mountain areas research and management* 15(1): 38–45. <https://doi.org/10.1553/eco.mont-15-1s38>

Zimmermann, A.B., S.W. von Dach, S.L. Mathez-Stiefel, D. Molden & T. Breu 2018. A brief history of peer-reviewed mountain journals: How platforms for knowledge relevant to sustainable mountain development emerged. *eco.mont - Journal on protected mountain areas research and management* 10(2): 84–87. <https://doi.org/10.1553/eco.mont-10-2s84>

Zubir, W.M.A.M., I. Abdul Aziz, J. Jaafar & M.H. Hasan 2018. Inference algorithms in latent Dirichlet allocation for semantic classification. *Advances in Intelligent Systems and Computing* 662: 173–184. https://doi.org/10.1007/978-3-319-67621-0_16

Authors

Thakur Dhakal^{1†}

is an Assistant Professor at the Department of Life Sciences, Yeungnam University, specializing in AI, machine learning and system dynamics. His work focuses on environment management, climate change, energy sustainability and resource optimization. E-mail: thakur_dhakal2003@yu.ac.kr

Tae-Su Kim^{1†}

is a postdoctoral researcher at the Department of Life Sciences, Yeungnam University. His research

focuses on forest ecology, spatial modeling and environmental monitoring. His research interests include climate change, habitat protection and forest management. E-mail: taesunim@ynu.ac.kr

Shraddha Tiwari^{1†}

is an Assistant professor at the Department of Life Sciences, Yeungnam University. She is a veterinary professional and aims to contribute to the fields of environmental ecology, veterinary science and life sciences. She is passionate about advancing veterinary science and improving animal welfare and environmental health. E-mail: shraddha_tiwari@yu.ac.kr

Seong-Hyeon Kim¹

is a PhD fellow at the Department of Life Sciences, Yeungnam University. His research interests include habitat and climate change modelling for sustainable ecosystems. E-mail: seong5465@ynu.ac.kr

Anteneh Bongasie¹

is a PhD fellow at the Department of Life Sciences, Yeungnam University. His research interests are in climate change modelling and sustainable ecosystem management. E-mail: antenehwork55@gmail.com

Jun Young Kim¹

is a PhD fellow at the Department of Life Sciences, Yeungnam University. His research interests include machine learning and deep learning for wildlife conservation and management. E-mail: kjy3702@ynu.ac.kr

Su-Jin Kim^{2*}

is a senior researcher at the Forest Ecology Division, National Institute of Forest Science. He contributes to the systematization of forest spatial management and conducts research on forest conservation and restoration by monitoring forest ecosystems and developing forest management technologies. E-mail: forestryhydro@korea.kr

Gab-Sue Jang^{1*}

is a senior Professor at the Department of Life Sciences, Yeungnam University. His research focuses on ecology, spatial modeling and environmental monitoring. He has contributed to interdisciplinary studies, collaborating in fields such as precision agriculture, landscape ecology, ecological modelling and environmental engineering. E-mail: sunside@ynu.ac.kr

¹ Department of Life Sciences, Yeungnam University, Gyeongsan 38541, Republic of Korea

² Division of Forest Ecology, National Institute of Forest Science, 02455, Republic of Korea

[†] These authors contributed equally to this work.

* Corresponding author