

# 생태계 기후변화 영향연구를 위한 생물종 기초조사자료의 통합 및 표준화 방법론 고찰

이경은\* · 홍승범\*\*†

\*국립생태원 생태연구전략실 전임연구원, \*\*국립생태원 기후생태연구실 선임연구원

## Integration and standardization of species survey data sets for climate change impact research for ecosystems

Lee, Kyungeun\* and Hong, Seungbum\*\*†

\*Associate Researcher, Division of Ecological Research Strategy, National Institute of Ecology, Seocheon, Korea

\*\*Senior Researcher, Division of Climate & Ecological Research, National Institute of Ecology, Seocheon, Korea

### ABSTRACT

As the impacts of climate change on ecosystems and human society become increasingly severe, the need for national-level policy responses is growing. However, biodiversity survey data collected by various government agencies in South Korea lack spatial and temporal representativeness at the national scale. This limits the ability to accurately assess the current state of biodiversity and to identify the mechanisms of biodiversity change due to climate change. To address these challenges, this study proposes a strategy for integrating various biodiversity survey datasets to obtain data suitable for analyzing the current state of biodiversity and time-series changes in the context of climate change impacts. The approach involves identifying common elements (such as species names, location, and survey period) and examining standardization methodologies. The key recommendations of this study are as follows: First, standardize species names based on the National Species List to ensure consistency in survey data. Second, improve the precision of location information and correct errors to enhance data reliability. Third, develop a method to derive analysis unit periods for time-series analysis from datasets with irregular survey frequencies. These strategies are expected to strengthen the scientific foundation for ecosystem change research and provide critical baseline data for policy development in response to climate change.

*Key words: Climate Change Impact, Time-series Analysis, Integration of Species Survey Data, Data Standardization, National-level Biodiversity Baseline Data*

## 1. 서론

최근 기후변화가 인간 사회와 생태계에 미치는 영향이 위기로 표현될 정도로 심각성이 점차 고조되고 있으며, 이에 대한 국가 차원의 정책적 대응 요구가 증가하고 있다. 우리나라는 국가기후변화적응대책(Government of Korea, 2010, 2015, 2020)에서 물관리, 생태계, 건강 등 여러 다양한 부문에 대한 기후변화 적응 정책을 추진해 왔으며, 특히 최근 국가기후위기적응강화대책(Government

of Korea, 2022)에서 생태계 부문의 14개 기후변화 리스크를 제시하고, 부정적 영향을 저감하기 위한 정책 개발 및 추진을 강화하고 있다. 생태계 분야 기후변화 적응을 위한 정책의 실효성은 과학적 신뢰도에 기반한 기후변화 영향 분석 정보에 따라 결정되는데, 과거부터 현재까지의 변화상에 대한 충분한 자료 확보와 기후 환경 변화 간의 인과관계 규명이 선행되어야 한다(Cheng et al., 2021). 그러나, 우리나라는 일부 지역 또는 특정 생물종 국한된 연구나(Kim et al., 2014; Kwon et al., 2020; Park et al.,

†Corresponding author : sbhong@nie.re.kr (1210, Geumgang-ro, Maseo-myeon, Seocheon-gun, 33657, Korea. Tel. +82-41-950-5611)

ORCID 이경은 0000-0003-1493-5360

홍승범 0000-0002-1163-7045

2018), 모형을 활용한 간접적인 전국 생물다양성 현황 분석 연구 사례(Kwon et al., 2012; Shin et al., 2018)만 존재할 뿐, 아직 실제 조사자료만을 기반으로 한 전국 단위의 생물다양성 현황 분석과 과거로부터의 시계열 변화에 대한 기후변화 영향 연구가 부족한 상황이다(Lee et al., 2019). 이에 따라 생태계의 기후변화 리스크에 대응하는 정책 수립에 필요한 과학적 근거 정보의 생산에 한계가 발생하고 있다.

우리나라는 환경부, 농림축산식품부, 해양수산부 등 여러 부처와 소속 및 산하기관에 의해 각기 다른 목적과 방법으로 육상생태계의 생물종 분포 기초조사 사업이 추진되어 왔다. 전국 단위의 비교적 체계적이고 표준화된 조사 방식을 적용하여 주기적으로 반복적인 생물 분포 현황을 조사 중인 사업으로는 환경부의 전국자연환경조사가 대표적이다. 하지만 두 가지 한계점으로 인해 국가 수준의 생태 정보로서의 대표성을 갖기에는 부족한 점이 있다. 첫째, 우리나라 생태계의 핵심 보전 지역인 국립공원이 조사 지역에서 제외되어 있다(NIE, 2017)는 점에서, 전국자연환경조사 자료만으로는 국가 수준의 생물다양성 현황을 파악하는 데 한계가 있다(Lee et al., 2023). 국립공원 지역의 생물종 분포 현황 조사는 국립공원공단에서 별도로 추진 중이다(KNPS, 2018). 둘째, 우리나라 생태계의 과거 변화를 분석하기에는 자료수집 기간이 20년 남짓으로 아직 짧다는 점이며, 체계화된 조사 방법론이 적용된 조사자료는 10년 정도에 불과하다(NIE, 2017). 기후변화 영향 연구의 특성상 보통 30년 이상의 관측 및 조사자료가 필요하나, 현재 진행 중인 우리나라 대부분의 생태계 분야 기후변화 관련 연구는 장기간 자료의 시계열 분석이 아닌 생물종이 조사된 지점과 기후 조건 간의 공간적 상관성 분석에 의존하는 경향이 있다(Hong et al., 2022; Kim et al., 2022; Park and Won, 2021; Park et al., 2019). 장기적으로 관찰된 생물종 분포변화 데이터의 부족은 우리나라 생태계 변화에 대한 기후변화 영향의 과학적 이해도 부족과 미래 기후변화 영향 예측의 정확성 문제로 이어진다. 특히, 멸종위기종이나 희귀생물처럼 정책적으로 세밀한 관리와 조치가 필요한 생물종의 경우, 종의 분포가 국소적이고 개체수도 적어 상대적으로 조사된 자료가 부족할 수밖에 없다. 이들에 대한 장기적인 조사자료 확보는 생육 환경과 서식지를 정확히 이해하고 보전 정책의 성공 여부를 판단하는 데 필수적이다.

환경부의 전국자연환경조사 사례처럼 하나의 조사 사업만으로 국가 차원의 공간적 대표성을 갖는 생물다양성 현황 정보를 도출하는 것이 어려우며, 기후변화 영향연구

에 필요한 시계열 데이터 또한 충분하지 않다. 이러한 한계를 해결하기 위한 대안으로 각 부처와 기관별로 산재해 있는 다양한 기초 조사자료를 수집하고 통합하는 방안을 고려할 수 있다. 예를 들어, 가장 오랜 기간 축적된 데이터를 보유하고 있는 국립생물자원관의 생물종 표본자료(Bae et al., 2021)와 공간적으로 균등한 방식의 전국 단위 생물조사가 이루어지는 전국자연환경조사 자료, 그리고 핵심 생태계 지역의 생물 조사가 수행 중인 국립공원공단의 자연자원조사 자료를 통합함으로써 전국 및 지역 단위의 생물다양성 현황 정보와 생물종의 시계열 변화에 대한 기초 데이터를 확보할 수 있을 것이다.

여러 기관에서 다양한 방식으로 수집된 자료를 통합할 때, 자료 간 품질의 불균형, 시간적 불연속성, 특정 지역에 대한 데이터의 편중, 데이터의 중복 등의 여러 문제가 발생할 수 있으므로, 이러한 문제를 최소화하기 위한 데이터 표준화가 매우 중요하다. Darwin Core는 국제적인 생물 정보 공유를 목적으로 개발된 표준 체계로, 전 세계적인 생물다양성 정보 구축에 유용한 도구이다(Wieczorek et al., 2012). 그러나 우리나라는 자체적으로 관리하는 국가 표준 종목록인 ‘국가생물종목록’을 보유하고 있어, 이를 국가 수준의 통합된 생물종 데이터 표준화 방법으로 Darwin Core를 바로 채택하는 것은 현재로서는 어려움이 있다. 그 이유는 표준화된 생물종 학명과 국가생물종목록의 학명이 일치하지 않는 종이 많아, Darwin Core 표준 체계를 준용할 경우 상당한 데이터가 유실될 가능성이 크기 때문이다. 따라서, 국제 기준에 부합하는 생물다양성 데이터베이스를 구축하기 위해, 먼저 국내 부처의 생물다양성 자료를 표준화할 방안을 마련하는 것이 시급하다. 이에 Hong et al. (2023)은 환경부 소속 및 산하 주요 기관들을 대상으로 관속식물종 조사자료를 수집하여 시범적으로 통합데이터를 구축하고, 전국 식물종 분포 현황과 과거 시계열 분석을 시도한 바 있으며, 그 과정에서 기후변화 영향연구 차원의 자료 통합 과정을 소개하였다.

본 연구에서는 관속식물종 조사자료를 표준화한 선행연구 경험을 바탕으로(Hong et al., 2023), 우리나라 육상생태계의 모든 분류군에 대하여 생물다양성 현황과 과거 변화, 그리고 기후변화 영향 예측을 위해 필요한 기초 생물종 조사자료를 통합하고 표준화하는 방법을 고찰하고자 한다.

## 2. 국내 생물종 기초조사 자료

본 연구에서 참고한 자료는 환경부 주요 소속 및 산하

Table 1. Summary of species survey projects under the Ministry of Environment in Korea

Institutes	Projects	Spatial Coverage	Temporal Coverage	Legal Basis	References
National Institute of Biological Resources (NIBR)	· Project of Native Species Exploration · Species Specimen Collection	Nationwide	1885 ~ 2021	-	Bae et al. (2021)
Korea National Park Service (KNPS)	· Natural Resource Survey of National Parks	National Parks	2000 ~ 2018	National Parks Act (Article 36)	KNPS (2018) KNPS (2019)
National Institute of Ecology (NIE)	· National Survey for Ecosystems in Korea	Nationwide (except National Parks)	1997 ~ 2019	Natural Environmental Conservation Act (Article 30)	NIE (2018) NIE (2023a)
	· Survey of Inland Wetlands	Inland Wetlands	2016 ~ 2021	Wetland Conservation Act (Article 4)	NIE (2020) NIE (2022)
	· Survey of Backdudaegan Protected Areas	Backdudaegan Areas	2015 ~ 2019	Natural Environmental Conservation Act (Article 31)	NIE (2021)
	· Survey of Endangered Species	Nationwide	1997 ~ 2021	Wildlife Protection and Management Legislation (Article 6)	NIE (2023b)

기관들이 수행한 생물종 기초조사 6개 사업(국립공원공단의 자연자원조사, 국립생태원의 전국자연환경조사(2~4차), 백두대간 정밀 조사, 내륙습지(기초, 정밀) 조사)과 수장고의 표본 정보(국립생물자원관의 표본자료)를 통해 수집된 것들이다(Table 1).

이 중 가장 장기간에 걸쳐 축적된 자료는 국립생물자원관의 생물종 표본자료로, 100년 이상의 데이터를 보유하고 있다. 그러나 우리나라 전국 단위의 생물종 조사는 2000년 이후부터 본격화되었기 때문에, 그 이전 시기의 데이터는 전체 데이터 수의 10% 정도 수준으로 종별 조사 표본 수가 매우 부족하여 전국적인 생물종 분포 현황과 생물다양성 변화 정보를 도출하는 데 한계가 있다 (Hong et al., 2023). 다만, 일부 특정 종에 한해서는 50년 이상의 장기적인 분포변화 정보를 도출할 수 있다.

국립생물자원관의 자료를 제외한 다른 조사사업들은 각 사업 목적에 맞는 법적 근거(자연환경보전법, 자연공원법 등)에 따라 주기적으로 추진하게 되어 있다. 동일한 사업 내에서 일관된 조사 방법론과 조사 지역이 유지되고, 장기간에 걸쳐 사업이 지속될 경우 신뢰성 높은 시계열 변화 정보를 도출할 수 있게 된다. 그러나 대부분의 기초 조사 사업은 그 추진 기간이 비교적 짧거나 초기 조사자료의 체계화가 이루어지지 않아, 기후변화 영향연구에 활용할 수 있는 데이터의 기간은 대략 20년 정도에 불과하다(Table 1). 예를 들어, 전국자연환경조사사업의 경우 전

국을 격자화하여 조사 방법론을 체계화한 것은 3차 사업부터로, 그 이전에 수집된 자료는 시계열 분석 활용에 한계가 있다.

### 3. 기후변화 영향연구를 위한 생물종 기초조사자료의 통합 및 표준화 방법론 고찰

#### 3.1. 자료 통합을 위한 공통 요소 정보 추출

자료 통합에 있어서 가장 기본이 되는 것은 각 자료가 담고 있는 여러 정보 요소 중 표준화가 가능한 공통 요소를 추출하는 것이며, 그 공통 요소들은 본 연구의 목적상 전국 단위 생물다양성 현황과 기후변화 영향 분석을 가능케 할 수 있어야 한다. 본 연구에서 수집한 자료들로부터 제공되는 정보의 종류와 정보별 제공 또는 표출 형식은 Table 2와 같으며, 모든 자료가 공통으로 제공하고 있는 정보 요소는 종명, 조사 위치(또는 종의 출현 위치), 조사 시기 등 3가지 정도이다. 이 중 종명과 종의 출현 위치 정보는 대부분의 생물종 분포 현황을 주제로 한 선행 연구들이 주로 활용하는 기본 요소들에 해당하며(Hong et al., 2022; Kim et al., 2014; Shin et al., 2018), 기후변화 영향 연구에 필요한 종별 시계열 변화 데이터의 구축을 위해 조사 시기 정보가 필수적으로 추가되는 요소가 된다.

그 외 종별 개체수 조사 정보, 생물종 출현 지역의 지리

Table 2. Summary of available information and recorded formats from survey data sources

Name of Data	Name of Species	Population info.	Format of Coordinate	Format of Survey Time	Address	Others
Species Specimen Collection (NIBR)	●	×	DMS	yyyy-mm-dd	●	-
Survey data of National Parks (KNPS)	●	⓪	Decimal	yyyy-mm-dd	●	National Park Names
The 2nd National Survey Data (NIE)	●	×	Decimal	yyyymmdd	×	Altitude, Moisture Conditions, Representative Mountain, etc.
The 3rd National Survey Data (NIE)	●	⓪	DMS	yyyy-mm-dd	×	
The 4th National Survey Data (NIE)	●	⓪	Decimal	yyyy-mm-dd	×	
Survey Data of Inland Wetlands (NIE)	●	×	Decimal	yyyy-mm-dd	●	Inland Wetland Naems
Survey Data of Backdudaegan (NIE)	●	⓪	Decimal	yyyy-mm-dd	●	Altitude, Moisture Conditions, etc
Survey Data of Endangered Species (NIE)	●	⓪	DMS	yyyy-mm-dd	●	Name of Survey Porject

[Information Availability] ● Full ⓪ Partial × No-info

[Format of Coordinate] Decimal : real number format, DMS : Divided into Degree(°), Minutes('), and Seconds(") of latitude & longitude info.

[Format of Survey Time] yyyy : 4-digit integer of year info., mm : 2-digit integer of month info., dd : 2-digit integer of date info.

적, 물리 환경적 특성 정보 등은 자료 통합을 위한 공통 요소로 고려하는 것은 어렵다. 종별 개체수는 Shannon’s diversity Index (Shannon, 1948) 같은 세밀한 생물다양성 정보 도출을 위해 중요한 요소지만, 기본 조사 항목에 포함되지 않은 사업들이 있어 현재의 통합 요소로 정하는 것은 어렵다. 특히, 국립생물자원관의 자료는 동일 지역, 동일 시점에 여러 개의 기록이 존재하나 이는 종의 개체 수 정보로 보기 어렵기 때문에 자료 통합 시, 단일 자료로 처리될 필요가 있다. 또한, 개체수 정보를 수집하는 사업 간에도 조사 방법이 일치하지 않아 자료 통합 시 불확실성을 높일 수 있다. 이에 본 연구에서 개체수 정보는 자료 통합 요소에서 제외하였으며, 이에 통합데이터의 활용성은 종풍부도 수준의 생물다양성 분석 정도로 제한된다.

기후변화 영향 연구의 핵심 요소인 생물종 관찰 지점의 위치정보는 조사에 사용된 장비의 정밀도에 따라 정보의 정확도가 결정된다. 2000년대 이후 대부분 조사 사업에서는 인공위성 기술을 활용한 정밀한 GPS 장비를 사용한 것으로 추정되지만, 조사자가 위치정보를 야장에 기록하는 과정에서 발생할 수 있는 실수는 데이터의 불확실성을 증가시킬 수 있다(Hong et al., 2023). 특히, 2000년대 이전의 기록은 정밀한 좌표 기록 장비가 널리 사용되지 않았던 시기의 자료로, 활용 시 주의가 필요하다.

고도, 토양 특성 등의 생물종 출현 지역, 즉 서식지의 환경적 특성 정보들은 사업별로 그 기록방식이 매우 다르고 대부분 표준화되지 않아 통합을 위한 공통 요소로서는 부적합하다. 그러나 생물종의 정확한 출현 위치 정보와

출현 시기 정보가 확보되었다고 가정하면, 생물종의 서식 환경에 대한 정보는 지형도, 토양도, 토지피복도 등의 공간 정보와 기온 및 강수와 같은 기후 요소와의 분석을 통해 간접적으로 추출할 수 있다.

### 3.2. 생물종 기초조사자료의 통합을 위한 공통 요소들의 표준화 방법

#### 3.2.1. 종명 표준화 및 보정

우리나라에 자생하는 생물종 기록은 「생물다양성 보전 및 이용에 관한 법률」 제10조와 시행령 제2조 및 제3조에 따라 국가생물다양성위원회(위원장: 환경부 차관) 산하에 국가생물종목록 실무위원회(위원장: 국립생물자원관장)를 두고 국립생물자원관에서 구축하고 있다.

2023년 기준, 총 60,010종이 등록되어 있으며, 각 생물종의 학명과 분류체계 등의 최신 정보를 제공하고 있다. 이를 통해 각 기관의 데이터를 하나로 통합하여 국가적으로 관리하는 것이 가능해졌다(NIBR, 2024). 따라서 국가 수준에서 두 개의 종목록이 존재하는 관속식물류, 곤충류, 균류, 지의류를 제외한 나머지 종들은 국가생물종목록을 기준으로 데이터를 통합하면 된다. 다만 국가생물종목록은 주기적으로 갱신되므로, 변경 이력은 남겨두고 목록이 최신화될 때마다 전체 자료를 다시 한번 검토해야 한다. 이는 조사 당시에는 두 개의 종이 분리되어 있었는데 최신 목록에서는 한 종으로 합쳐지는 경우 또는 분류학적 위치가 변경되어 속명이 바뀌는 경우 등이 있기 때문이다.

Table 3. Comparison between the national species list and the national standard plant & insect list

Based on the KME list	Vascular Plants	Insects	Examples	
			Korean Names	Scientific Names
Cases when both the Korean names and scientific names match	2,453	14,203		
Cases when the scientific names match only	120	204	꼭두선이 vs. 꼭두서니	<i>Rubia argyi</i> (H. Lév. & Vaniot) H. Hara ex Lauener & D.K. Ferguson
			밀드리좁벌 vs. 밀들이벌	<i>Leucospis japonica</i> Walker, 1871
Cases when the Korean names match only	1,855	4,086	장대나물	<i>Arabis glabra</i> (L.) Bernh. vs. <i>Turritis glabra</i> (L.)_synonym
			줄물방개	<i>Hydaticus (Prodaticus) satoi satoi</i> Wewalka, 1975 vs. <i>Hydaticus (Prodaticus) satoi</i> Wewalka, 1975
Cases when both the Korean names and scientific names mismatch	73	143	사탕무우 vs. 사탕무	<i>Beta vulgaris</i> var. <i>altissima</i> Döll vs <i>Beta vulgaris</i> var. <i>saccharifera</i> Alef._synonym
			작은복숭아명나방 vs. 소나무명나방	<i>Conogethes pinicolalis</i> Inoue and Yamanaka, 2006 vs. <i>Conogethes pinicolalis</i> Inoue & Yamanaka, 2006
Not-exist species in the KFS list	140	2,074	땃꽃	<i>Halenia corniculata</i> (L.) Cornaz
Total	4,641	20,710		

앞서 언급한 관속식물류, 곤충류, 균류 및 지의류는 국가 수준에서 두 개의 종목록이 존재하는데, 하나는 국가 생물종목록이며, 다른 하나는 국립수목원이 정리한 목록이다. 국립수목원의 국가표준식물목록과 국가표준지의류 목록은 「수목원 조성 및 진흥에 관한 법률」 제3조에 의해 국립수목원장을 위원장으로 하는 국가수목유전자원목록 심의회를 통해 구축되었다. 이 위원회는 국립수목원과 식물분류학회가 공동으로 조직하여 목록작성에 필요한 모든 사항을 검토하고 결정한다. 국가표준곤충목록은 한국 곤충명집(ESK and KSAE, 1994)과 한국곤충총목록(Baek et al., 2010)을 바탕으로, 국립수목원에서 진행한 연구 결과를 반영하여 정리되었다. 국가표준버섯목록은 사단법인 한국균학회가 발간한 한국의 버섯목록(KSM, 2013)과 국립생물자원관의 국가생물종목록집(NIBR, 2015)을 기반으로 작성되었다(KNA, 2024). 이 중 많은 조사 사업의 대상 분류군인 관속식물류와 곤충류 종목록을 비교해 보면, 두 목록 간에는 종수와 증명에서 차이가 있다. 이는 데이터 통합 과정에서 중요한 고려 사항이 된다. 국가생물종목록에서는 식물이 4,641종, 곤충이 20,710종으로 기록되어 있지만, 국가표준식물목록과 국가표준곤충목록에서는 각각 4,320종과 18,256종이 기록되어 있다. 두 목록 간 증명과 학명이 일치하는 경우는 식물 2,453종, 곤충 14,203종으로 각각 53%, 69%에 해당한다. 국명은 같으나 학명이 다른

경우도 많아, 식물은 1,855종(40%), 곤충은 4,086종(20%) 이 학명이 달랐다. 대부분의 차이는 명명자 표기의 차이 또는 국가생물종목록에서 이명이나 원기재명으로 표기된 것이 국가표준식물목록에서는 정명으로 쓰였기 때문이다. 예를 들어, 팔꽃나무의 학명은 *Daphne genkwa* Siebold & Zucc.이나 국가표준식물목록에서는 *Wikstroemia genkwa* (Siebold & Zucc.) Domke로 표기된다. 두 목록이 국가 수준에서 동시에 존재하는 이유에 대해서는 여기서 다루지 않겠지만, 현재 연구자들은 두 목록 중 하나에 맞춰 종목록을 작성하고 있다. 조사 사업마다 정해진 기준 종목록이 있으나, 연구자에 따라 더 친숙한 목록, 도감이 있으므로 자연스럽게 그 목록을 기준으로 작성하게 된다. 이에 따라 일부 종은 같은 종임에도 불구하고 서로 다른 국명이 쓰인 자료가 발생하게 된다. 예를 들어 국가생물종목록에서는 꼭두선이(*Rubia argyi* (H. Lév. & Vaniot) H. Hara ex Lauener & D.K. Ferguson)라고 쓰지만, 국가표준식물목록에서는 꼭두서니(*Rubia argyi* (H. Lév. & Vaniot) H. Hara ex Lauener & D.K. Ferguson)가 정명이다(Table 3). 사업 담당자가 연구자들이 제출한 데이터를 모두 검토해서 정리를 해야 하지만, 현실은 아직도 두 목록의 증명이 섞인 자료가 나오고 있다.

식물과 곤충에 대한 증명과 학명의 차이는 데이터 통합 과정에서 데이터 손실 및 왜곡을 초래할 수 있다. 전국자

연환경조사 4차 식물상 조사자료로 비교해 보면 원본 데이터 수는 194,466개였으나, 국가생물종목록과 종명이 일치하지 않는 데이터를 일괄 제거하여 제공하고 있는 국립생태원의 에코뱅크(국제생태정보종합은행) 데이터는 192,995개로 1,471개가 손실되었다. 그러나 두 종목록을 모두 검토하여 Table 3처럼 같은 종이지만 종명이 다를 경우 국가생물종목록 기준으로 보정했을 때는 데이터 수가 193,823개로 손실된 데이터의 절반 이상이 복구되었다. 식물은 생태계의 근간이므로 이에 대한 조사와 정리가 명확히 되어야 다른 생물들과의 상호작용 연구에서도 높은 정확성을 보장할 수 있다. 특히 곤충의 경우, 국명이 정해지지 않은 종(1,062종)이 많아 새로운 이름이 부여되면 과거 자료에서도 이를 반영할 수 있도록 해야 한다. 또한 기존의 국명 또는 학명이 이명이 되거나, 다른 종명이 정명으로 바뀌는 경우, 혹은 연구 결과에 따라 목록상 다른 종이 같

은 종으로 판명될 때마다 기존 자료의 종명도 변경되도록 변경 이력을 남겨야 한다. 목록에 없어서 제외된 자료들도 후에 종이 목록에 추가될 경우 자료로 사용할 수 있도록 보관해야 한다. 시계열 데이터를 확보하기 위해 과거 자료를 최신 종목록 기준으로 통합할 때는 이명, 원기재명, 국명 및 학명 변경을 반드시 고려해 데이터 손실을 최소화해야 한다.

생물조사 자료를 분석하기 전에 발생할 수 있는 오류는 크게 두 가지이다. 첫 번째는 현장에서 수기로 수집한 내용을 전산화하는 과정에서 발생하는 오류이며, 두 번째는 데이터를 2차 가공할 때 발생하는 오류이다. 데이터가 전산화된 이후에는 해당 파일만 주고받으며 분석에 사용되므로 초기 입력 단계에서 오류가 발생하지 않도록 하는 것이 매우 중요하다. 숫자데이터의 오류는 데이터 유효성 검증을 통해 수정할 수 있지만, 국명이나 학명 같은 문자

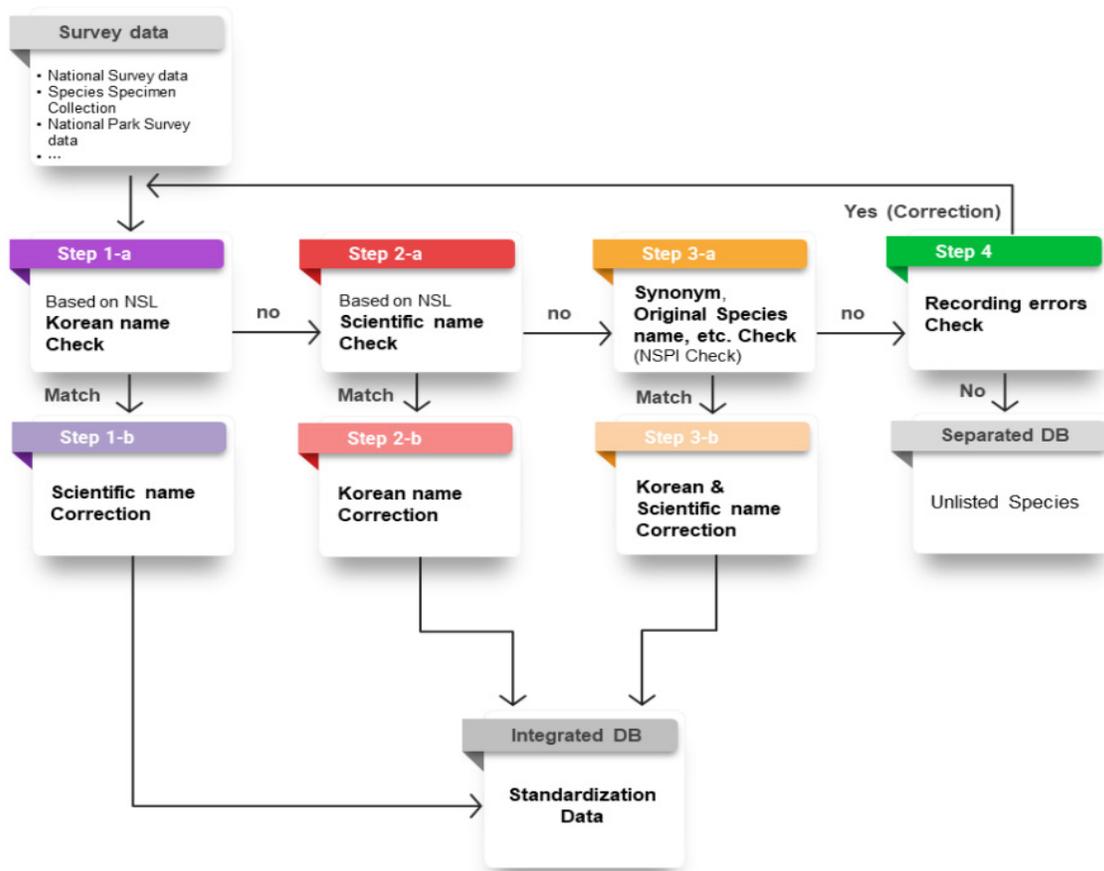


Fig. 1. Procedure of species name standardization for data integration. NSL: National Species List, NSPI: National Standard Plant & Insect List

데이터의 오류는 찾아내기 어렵고 시간이 많이 소요된다. 이러한 오류를 줄이기 위해 데이터를 전산화할 때 기준이 되는 목록과 다른 종명이 입력되지 않도록 하거나, Global Biodiversity Information Facility (GBIF) species matching tool (GBIF, 2024)처럼 오류가 발생하면 자동으로 표시되도록 하는 시스템 기능이 필요하다. 하지만 모든 조사자료가 하나의 기준으로 통일되어 입력되기 위해서는 이를 통제하는 중앙기관과 시스템이 필요하다. 또한, 시계열 데이터를 확보하기 위해 과거 자료까지 모두 통합하려면 오류가 발생한 데이터를 하나씩 검토해야 한다. 예를 들어 졸참나무(*Quercus serrata*)가 졸참나무로 되어 있거나, 개차즈기의 학명은 *Amethystea caerulea*인데 자료에 *Amethystea coerulea*로 기재된 경우, 이는 오기로 볼 수 있다. 하지만 데이터상으로는 오류로 인식되어 다른 종으로 처리되거나 이상치로 삭제될 가능성이 있어, 분석에 사용되는 데이터 수가 줄어들 수 있다. 따라서 종명이나 학명에 오류가 발생하더라도 무조건 제거하는 것은 지양하고, 오류 데이터에 대한 별도의 검토가 필요하다.

본 연구에서 제안하는 증명 표준화는 국가생물종목록을 기준으로 하며, 국명을 우선순위로 두는 방법이다(Fig. 1). 이는 조사 현장에서 최초 기록 시 국명을 사용하기 때문이다. 국가생물종목록과 조사자료의 국명과 학명이 일치하면 해당 자료는 바로 통합 데이터베이스(DB)로 넘어가게 되고, 학명이 일치하지 않을 때는 학명을 국가생물종목록의 학명으로 수정하여 통합 DB로 이동시킨다(Fig. 1, Step one). 나머지 자료들은 국가생물종목록의 학명과의 일치 여부를 점검한다. 일치할 경우 국명을 국가생물종목록으로 수정하여 통합 DB로 이동시킨다(Fig. 1, Step two). 학명과의 일치 여부를 점검하는 단계에서는 동일 종으로 판단할 수 있는 부분 일치 기준(예: 학명 글자의 80% 이상 일치)을 설정하는 방법도 고려할 수 있다. 이는 학명 오기로 발생한 불일치 자료를 찾을 수 있기 때문이다.

남은 자료는 이명, 원기재명 등을 확인하며, 국립수목원의 국가표준생물목록도 참고할 수 있다(Fig. 1, Step three). 위에서 언급한 것처럼 국가표준생물목록의 정명이 국가생물종목록에서는 이명으로 처리한 경우가 있기 때문이다. 또한 국가생물종목록은 지속적으로 최신화되므로, 이명 관리도 증명 표준화에서 중요한 부분이다. 조사자료의 종명이 이명 등과 일치하면 국가생물종목록의 정명으로 수정하여 통합DB로 이동시킨다. 통합DB로 보내고 남은 자료는 띄어쓰기, 오기 등 단순 오류를 점검한다(Fig. 1, Step Four). 이 단계에서는 해당 분류군 전문가의

수작업이 필요하나, 남은 자료가 많지 않기 때문에 체크가 가능하다. 오류를 수정한 자료는 다시 첫 단계(Step one)로 돌아가 위의 단계를 다시 거친다. 최종적으로 통합에서 제외된 자료들은 후에 국가생물종목록의 최신화 시 포함될 가능성을 고려해 별도의 데이터베이스로 관리할 것을 제안한다.

### 3.2.2. 좌표체계의 표준화 및 보정

위치 정보의 표기 방법은 Table 1에서 알 수 있듯이 주로 DMS(도분초)와 Decimal(실수) 두 가지 형식으로 기록된다. DMS 형식에서는 ‘도’와 ‘분’이 정수 형태로, ‘초’는 소수점 둘째 자리까지의 실수(floating number) 형식으로 기록되는 것이 일반적이다. 따라서 경도와 위도, 두 개의 기본적인 위치 정보는 결과적으로 네 개의 정수와 두 개의 실수 형태로 저장된다. 반면, Decimal 방식은 두 개의 실수만으로 저장되므로, 데이터 저장의 용량 관리와 데이터 처리 효율성 면에서 유리하다. 단, 위치 정보의 정밀도 측면에서 Decimal 방식 자료를 표준화할 때는 경위도 수치의 소수점 이하 자릿수를 적절히 설정할 필요가 있다.

본 연구에서 다루는 조사 자료의 통합 목적이 기후변화 연구라는 점과, 현재 우리나라의 기후변화 자료의 공간적 해상도가 약 1 km임을 고려할 때, 현시점에서 생물종 위치 좌표의 정밀도는 0.001°(중위도에서 약 100 m 거리)에서 0.01°(중위도에서 약 1 km 거리) 정도면 충분할 것으로 생각한다. 다만, 생물종 출현 위치의 다양한 물리·환경적 요소들(기온, 강수 등의 기후 요소와 고도, 경사도, 향, 토양 특성 등)이 생물종의 서식 및 생육 특성 설명에 매우 중요하다는 점을 고려하여, 최종적으로는 공간 해상도 1 m 수준에 해당하는 소수점 이하 6자리까지의 정밀도를 표준화 방식으로 제안한다.

국립생물자원관이 보유한 생물종 표본 정보 중 50년 이상 된 데이터들은 당시 현장 장비의 부족으로 인해 위치 좌표 정보가 없는 경우가 많았다. 고품질 데이터 확보를 위해 이러한 데이터는 보통 제거되지만, 데이터의 불확실성을 고려하여 거시적 관점에서 기후변화 연구에 활용할 수 있도록, 주어진 행정 주소 구역(최소 시군구 수준) 내 임의의 위치 정보(예: 중심 좌표)로 대체하는 방안을 고려할 수 있다(Hong et al., 2023). 특히 위도별 기온 구배가 뚜렷한 경우, 생물종의 위치 정보가 다소 부정확하더라도 위와 같은 보정 방식을 통해 생물종 표본 수가 증가하면 기후변화 영향 연구에 유리할 수 있다. 예를 들어, 국립생물자원관의 생물종 표본 정보는 100년 이상의 시계열 테

이터를 포함하고 있어, 이를 최대한 활용하면 특정 생물종의 과거 서식지나 범위 변화에 대한 분석이 가능할 수 있다. 체계화된 우리나라 생물종 조사 사업의 역사가 짧다는 점을 고려할 때, 최대한 많은 시계열 데이터 확보는 통합자료 구축에 매우 중요한 요소이다. 물론, 이러한 보정된 자료를 포함한 통합된 생물종 자료를 분석에 활용하기 위해 데이터의 불확실성 관리에 대한 주의가 필요하다.

### 3.2.3. 조사시기 정보의 표준화

Hong et al. (2023)은 조사 자료의 연도 정보만을 활용하여 생물다양성의 시계열 분석을 시도하였다. 이는 통합에 활용된 자료의 조사 사업마다 조사 주기가 상이하고, 동일한 조사 사업 내에서도 분류군별로 연중 조사 시점이 다양하여 일 단위 수준의 자료를 활용한 분석에는 한계가 있기 때문이다. 물론, 현재 자료 수집 상황을 감안했을 때, 연 단위 수준의 정보만으로도 장기적인 생물종 분포 변화의 경향성 분석은 가능하다. 그러나 앞으로 각 기관에서 충분한 기간의 조사 사업이 추진되고 충분한 조사 자료가 확보될 것을 고려하여, 생물종의 계절적 특성을 반영한 심도 있는 기후변화 영향 연구를 위해 월 단위 및 일 단위의 정보까지 통합 정보에 포함될 필요가 있다.

채집 시기 정보의 표준화는 비교적 단순하며, 일반적으로 어떤 형태로든 활용에 큰 무리가 없다. 하지만 전산화할 때 날짜 표기를 연, 월, 일로 하고, 하나의 필드에 모두 입력하는 것이 아니라 필드별로 분리하여 입력하면, 연도별, 계절별, 월별 등 다양한 방식으로 데이터를 추출할 수 있어 데이터 사용자의 활용성이 크게 높아질 것이다.

### 3.3. 통합자료의 시계열 분석을 위한 단위 기간 도출 방법 고찰

3.2장에서 논의한 최소한의 공통 요소 기반 자료 표준화 방법은 특정 시점 또는 일정한 시간적 범위 내에서 종 풍부도 수준의 생물다양성 현황을 분석하는 데 유용하다. 그러나 이 방법은 생물다양성의 시계열적 변화를 분석하는 데 여전히 한계가 있다. 이는 조사 사업 간의 주기 차이뿐만 아니라, 동일한 사업 내에서도 조사자 개인 간의 노력 차이 등으로 특정 시기에 조사자료의 편중 현상이 발생할 수 있기 때문이다. 이러한 문제로 인해 주어진 조사 시기 정보를 그대로 시계열 분석에 활용할 경우, 결과의 불확실성이 발생할 수 있다. 따라서, 생물종 다양성의 시계열 분석을 위한 적절한 단위 기간 선정 방법이 고려

되어야 한다.

통합된 식물 종 조사자료를 활용하여 전국 단위의 식물 종 풍부도의 시계열 변화를 분석할 수 있는 적절한 단위 기간을 도출하고자 종 누적 곡선을 응용한 연구가 있다 (Hong et al., 2023). 이 연구에서는 여러 단위 기간(예: 1년~20년)을 설정하여 취합된 조사자료로 시계열 데이터를 구성한 뒤, 각 단위 기간별로 추출된 종 누적 곡선을 통계적으로 비교하며 가장 적절한 단위 기간을 찾는 방식을 사용하였다. 이때, 단위 기간이 길어질수록 분석할 수 있는 시계열 데이터의 수가 줄어드는 문제를 해결하기 위해, 단위 기간별 조사자료를 이동 누계 방식으로 추출하였는데, 우리나라처럼 조사자료의 수집 기간이 비교적 짧은 경우 활용 가능한 방법이다. 연구 결과, 통합된 식물종 자료로 전국 단위의 식물 종 풍부도 현황을 도출하는 데는 성공적이었으나, 종 풍부도의 시계열적 변화를 분석하기에는 여전히 충분하지 않았다(Hong et al., 2023). 하지만, 각 기관의 조사 사업이 지속되고 조사자료의 통합 확대가 이루어질 것으로 기대되는 만큼, 위의 단위 기간 추출 방법을 통해 전국 단위의 종 다양성 시계열 분석도 가능할 것으로 본다.

## 4. 결론

본 연구에서는 부처와 기관별로 흩어져 있는 우리나라 생물종 기초조사 자료의 통합을 위한 표준화 방법론에 대해 논의하였다. 생물종 기초조사 자료의 통합은 생태 부문에서 기후변화 대응에 필요한 기초과학 정보를 생산하기 위한 기반이며, 전국 단위의 생물종 분포 현황과 장기적인 시계열 변화 정보를 구축하기 위함이다. 이를 위해 본 연구는 분산된 생물종 기초조사 자료의 통합을 위한 세 가지 공통 요소 정보를 선정하고, 각 요소의 표준화 방향을 제안하였다.

첫 번째로 논의된 종 명칭의 표준화는 국가 수준의 통합된 생물종 기초조사 자료를 지속적으로 최신화하고 관리하는 데 있어 매우 중요한 부분이다. 본 연구에서 사용된 증명 표준화의 기준은 새로운 종 명칭의 추가나 기존 종 명칭의 변경 등이 지속적으로 최신화되는 환경부의 국가생물종목록이다. 이에 따라 이미 구축된 통합 자료도 종 목록이 최신화될 때마다 검토가 필요하다. 이를 위해 자료의 이명 관리(변경 전 증명)도 함께 요구되는데, 이러한 이명 관리는 환경부와 산림청의 종 목록 간 불일치를 해결하는 데에도 필요한 부분이다. 궁극적으로는 국가 차

원에서 통일된 생물종 명칭을 사용하고 전문가 간 이견을 최소화하기 위해 환경부와 산림청에서 각각 독립적으로 운영 중인 관련 위원회를 일원화할 필요가 있다.

두 번째, 생물종의 위치 정보와 조사 시기 정보의 표준화는 비교적 단순한 과정이나, 위치 정보 표준화 과정에서 좌표 표기의 오류를 보정하는 단계는 잘못된 입력으로 인해 유실될 수 있는 정보를 최소화하는 데 도움이 된다. 특히, 위치 정보의 오류 점검 과정에서 함께 제공되는 행정구역 주소 정보를 활용하는 방법은 통합 자료의 신뢰도를 높이는 데 기여할 것이다. 더 나아가, 위치 정보의 보정이 불가능할 경우 주소 정보를 활용한 위치 정보의 대체 방법은 거시적 관점에서 생물 분포 변화 경향 분석에 유용하게 활용될 것으로 기대된다.

마지막으로, 본 연구에서 제안한 분석 기간의 표준화 방법은 통합된 자료를 활용하여 전국 생물다양성 분포의 시계열 분석을 수행하기 위한 합리적인 접근 방식으로, 자료별로 불규칙한 조사 주기의 문제를 해결하는 데 기여할 수 있다. Hong et al. (2023)은, 이 방법을 식물종 통합 자료에 시범 적용하여 전국 단위의 이론적인 식물종 풍부도 현황 정보를 도출하였으나, 향후 전국자연환경조사 5차 자료나 다른 기관의 지속적인 기초조사 자료가 추가되면 국가 수준의 생물다양성 시계열 변화 정보를 도출할 수 있을 것으로 기대된다. 또한, 이 방법은 행정구역이나 도엽 단위와 같은 더 작은 공간 규모에 적용할 경우, 지역별 생물종 기초조사 사업의 실패를 진단하고, 조사 미비 지역의 조사 사업 개선 방향을 도출하는 데에도 유용하게 활용될 것으로 기대된다.

최근 환경부는 국가 수준의 생태 정보 관리와 기후위기 대응을 위한 과학 기반 정책 지원 정보 생산하기 위해 기후변화와 생태 정보를 통합·관리하는 시스템 개발 사업을 추진하고 있다(Ock et al., 2023). 이 사업의 핵심 목표 중 하나는 각 기관에 산재해 있는 생물종 기초조사 자료를 통합하는 것이며, 본 연구에서 제시한 자료 표준화 과정을 적용하면 국가 수준의 대표성을 갖춘 통합된 생물종 분포 자료를 구축할 수 있을 것이다. 본 연구에서 제안하는 표준화 방법은 종별 위치 정보를 기본적으로 포함하고 있어 생물종의 서식 환경과 관련된 다양한 공간 정보를 연계하기에 매우 용이하다. 이는 점차 공간적으로 고도화되어 제공되는 기후변화 예측 정보, 위성 관측 정보, 그리고 각종 환경 변화 정보와 연계하여, 기후변화뿐만 아니라 다양한 환경 변화에 따른 생태계의 변화를 분석하고, 정책적으로 활용 활용할 수 있는 정보를 생산하는 데 유

용하게 쓰일 것이다.

## 사사

본 연구는 2024년 국립생태원 「생태계의 기후변화 리스크에 대응한 적응역량 강화 연구 ('24)」(NIE-고유연구-2024-35)를 위한 연구비 일부 지원으로 수행되었으며, 이에 감사드립니다.

## References

- Bae YJ, Cho K, Min G-S, Kim B-J, Hyun J-O, Lee JH, Lee HB, Yoon J-H, Hwang JM, Yum JH. 2021. Review of the Korean indigenous species investigation project (2006-2020) by the National Institute of Biological Resources under the Ministry of Environment, Republic of Korea (in Korean with English abstract). *Korean J Environ Biol* 39(1): 119-135. doi: 10.11626/KJEB.2021.39.1.119
- Baek MK, Hwang JM, Jeong KS, Kim TW, Kim MC, Lee YJ, Cho YB, Park SW, Lee HS, Ku DS, Jeong JC, Kim KG, Choi DS, Shin EH, Hwang JH, Lee JS, Kim SS, Bae YS. 2010. Checklist of Korean insects. Seoul: Nature & Ecology.
- Chang CS, Kwon SY, Kim H. 2021. Status and quality analysis on the biodiversity data of East Asian vascular plants mobilized through the Global Biodiversity Information Facility (GBIF). *Journal of Korean Society of Forest Science* 110(2): 179-188. doi: 10.14578/jkfs.2021.110.2.179
- ESK (The Entomological Society of Korea), KSAE (Korean Society of Applied Entomology), Editors. 1994. Check list of insects from Korea. Seoul: Konkuk University Press.
- GBIF (Global Biodiversity Information Facility). 2024. GBIF species matching tool. [accessed 2024 Aug 9]. <https://www.gbif.org/tools/species-lookup>
- Government of Korea. 2010. Korea national climate change adaptation plan (2011-2015). Sejong, Korea: Ministry of Environment.
- Government of Korea. 2015. The 2nd Korea national

- climate change adaptation plan (2015-2020). Sejong, Korea: Ministry of Environment.
- Government of Korea. 2020. The 3th Korea national climate change adaptation plan (2021-2025). Sejong, Korea: Ministry of Environment.
- Government of Korea. 2022. Measures to strengthen adaptation to the national climate crisis. Sejong, Korea: Ministry of Environment.
- Hong S, Jang I, Kim D, Kim S, Park HS, Lee K. 2022. Predicting potential habitat changes of two invasive alien fish species with climate change at a regional scale. *Sustainability* 14(10): 6093. doi: 10.3390/su14106093
- Hong S, Oh J, Cha J, Lee K. 2023. Plant species richness in Korea utilizing integrated biological survey data (in Korean with English abstract). *Korean J Ecol Environ* 56(4): 363-374. doi: 10.11614/KSL.2023.56.4.363
- Kim J, Kwon H, Seo C, Kim M. 2014. A nationwide analysis of mammalian biodiversity hotspots in South Korea (in Korean with English abstract). *J Environ Impact Assess* 23(6): 453-465. doi: 10.14249/eia.2014.23.6.453
- Kim WM, Kim C, Cho J, Hur J, Song W. 2022. Prediction of *Acer pictum* subsp. *mono* distribution using bioclimatic predictor based on SSP scenario detailed data (in Korean with English abstract). *Ecol Resil Infrastruct* 9(3): 163-173. doi: 10.17820/eri.2022.9.3.163
- KNA (Korea National Arboretum). 2024. National species knowledge information system. [accessed 2024 May 23]. <http://www.nature.go.kr>
- KNPS (Korea National Park Service). 2018. Study of systematization of the 4th natural resources survey. Wonju, Korea: Korea National Park Research Institute.
- KNPS (Korea National Park Service). 2019. Development of information analysis system of the national park resources survey data. Wonju, Korea: Korea National Park Research Institute.
- KSM (Korean Society of Mycology), Editors. 2013. List of mushrooms in Korea. Seoul: Author.
- Kwon HS, Lee YK, Yoo SH, Kim DW, Kim JS. 2020. Distribution patterns of biodiversity hotspot using bird data from the 3rd National Ecosystem Survey in South Korea. *J Korean Env Res Tech* 23(3): 81-89. doi: 10.13087/kosert.2020.23.3.81
- Kwon HS, Seo CW, Park CH. 2012. Development of species distribution models and evaluation of species richness in Jirisan region (in Korean with English abstract). *J Korean Soc Geospat Inf Sci* 20(3): 11-18.
- Lee HS, Sagong H, Joo YJ, Jeong SK. 2019. Development of ecoinformatics biodiversity assessment tools. Sejong, Korea: Korea Environment Institute. KEI Research Report 2019-20.
- Lee K, Kim D, Cha J, Hong, S. 2023. Fine-scale species distribution modeling of *Abies koreana* across a subalpine zone in South Korea for in Situ species conservation. *Sustainability* 15(11): 8964. doi: 10.3390/su15118964
- NIBR (National Institute of Biological Resource). 2015. National list of species\_Basidiomycota. Incheon, Korea: Author.
- NIBR (National Institute of Biological Resource). 2024. Biodiversity of Korean peninsula. [accessed 2024 May 20]. <https://species.nibr.go.kr>
- NIE (National Institute of Ecology). 2017. National ecosystem survey in Korea for 30 years: 1986-2015. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2018. The 4th national ecosystem survey in Korea. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2020. Guidelines for inland wetland survey. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2021. Comprehensive report on the ecosystem survey of the Baekdudaegan protected area. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2022. Status of inland wetland in Korea. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2023a. The 5th national ecosystem survey in Korea. Seochon, Korea: Author.
- NIE (National Institute of Ecology). 2023b. Guideline for

- national survey of endangered species. Seocheon, Korea: Author.
- Ock G, No D, Lee H, Park E-J. 2023. Strategy and case analysis for eco-climate integrated information system (in Korean with English abstract). *J Clim Change Res* 14(6-2): 981-988. doi: 10.15531/KSCCR.2023.14.6.981
- Park CY, Won MY. 2021. Impact of climate change on urban bird species richness and the importance of urban green spaces (in Korean with English abstract). *J Clim Change Res* 12(5-1): 371-381. doi: 10.15531/KSCCR.2021.12.5.371
- Park H, Park D-Y, Choi S-K. 2018. Changes in avian distribution in the Republic of Korea -A response to climatic and environmental condition change-. *J Energy Clim Change Educ* 8(2): 235-253. doi: 10.22368/ksecce.2018.8.235
- Park SU, Koo KA, Kong WS. 2019. Climate-related range shifts of climate-sensitive biological indicator species in the Korean Peninsula: A role of dispersal capacity (in Korean with English abstract). *J Clim Change Res* 10(3): 185-198. doi: 10.15531/KSCCR.2019.10.3.185
- Shannon CE. 1948. A mathematical theory of communication. *Bell Syst Tech J* 27(3): 379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Shin MS, Seo C, Lee M, Kim JY, Jeon JY, Adhikari P, Hong SB. 2018. Prediction of potential species richness of plants adaptable to climate change in the Korean Peninsula (in Korean with English abstract). *J Environ Impact Assess* 27(6): 562-581. doi: 10.14249/eia.2018.27.6.562
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D. 2012. Darwin Core: An evolving community-developed biodiversity data standard. *PLoS ONE* 7(1): e29715. doi: 10.1371/journal.pone.0029715