

Research Paper

머신러닝 알고리즘을 이용한 포유류 종 풍부도 매핑 구축 연구

김지영* · 이동근* · 김은섭** · 최지영*** · 전윤호****

서울대학교 조경·지역시스템공학부*, 서울대학교 협동과정 조경학**,
서울대학교 농업생명과학연구원***, 한국환경연구원****

Mapping Mammalian Species Richness Using a Machine Learning Algorithm

Zhiying Jin* · Dongkun Lee* · Eunsub Kim** · Jiyoung Choi*** · Yoonho Jeon****

Department of Landscape Architecture and Rural Systems Engineering, Seoul National University*
Interdisciplinary Program and Life Science, Graduate School of Environmental Studies, Seoul National University**
Research Institute of Agriculture and Life Sciences, Seoul National University***
Korea Environment Institute****

요약: 생물다양성은 환경영향평가 제도의 목표에 중요한 부분으로, 개발대상지 입지 선정, 주변 환경 파악 및 교란으로 인한 생물종 영향 등에서 활용되고 있다. 환경영향평가 분야에서 새로운 기술과 모델을 활용하여 생물다양성을 보다 정확하게 평가하고 예측하는 방안에 대한 연구가 많이 진행되고 있다. 비록 현장, 문헌조사를 통한 데이터를 바탕으로 종 풍부도 지수를 평가하고 있으나, 현장 데이터는 시·공간적으로 미흡하므로 고해상도의 종 풍부도 매핑을 통한 기초자료를 활용함으로써, 모니터링 실효성 문제 해결이 필요하다. 따라서 본 연구에서는 제4차 전국자연환경조사 데이터와 환경변수를 바탕으로 Random forest 모델을 활용하여 종 분포모형을 개발하였다. 해당 모델은 24종의 포유류 종 분포 매핑 결과를 species richness index를 활용하여 100m 해상도의 종 풍부도 매핑 결과를 도출하였다. 연구 결과, 종 분포모형은 평균 0.82의 AUC값으로 우수한 예측 정확도를 보였다. 또한, 전국자연환경조사 데이터와 비교결과, 고 해상도의 종 풍부도 매핑 결과의 종 풍부도 분포는 정규분포의 형태를 가지고 있어 환경영향평가에서의 기초자료로 사용함에 있어 신뢰성이 높다. 본 연구의 분석결과를 추후 도시개발과 사업을 함에 있어 생물다양성 평가, 서식지 보전 등에 새로운 참고자료로 활용될 수 있을 것으로 사료된다.

주요어: 생물다양성, 종분포모형, 포유류, 환경영향평가

Abstract: Biodiversity holds significant importance within the framework of environmental impact assessment, being utilized in site selection for development, understanding the surrounding

First Author: Zhiying Jin, Tel: +82-2-880-4885, E-mail: jzy0305@snu.ac.kr, ORCID: 0009-0005-4760-9758

Corresponding Author: Dongkun Lee, Tel: +82-2-880-4875, E-mail: dklee7@snu.ac.kr, ORCID: 0000-0001-7678-2203

Co-Authors: Eunsub Kim, Tel: +82-2-880-4885, E-mail: mr.solver92@snu.ac.kr, ORCID: 0000-0003-2206-5203

Jiyoung Choi, Tel: +82-2-880-4885, E-mail: jychoi1@snu.ac.kr, ORCID: 0000-0002-2702-1245

Yoonho Jeon, Tel: +82-44-415-7777, E-mail: yhjeon@kei.re.kr, ORCID: 0000-0003-4094-7082

Received: 19 December 2023. Revised: 18 April 2024. Accepted: 22 April 2024.

environment, and assessing the impact on species due to disturbances. The field of environmental impact assessment has seen substantial research exploring new technologies and models to evaluate and predict biodiversity more accurately. While current assessments rely on data from fieldwork and literature surveys to gauge species richness indices, limitations in spatial and temporal coverage underscore the need for high-resolution biodiversity assessments through species richness mapping. In this study, leveraging data from the 4th National Ecosystem Survey and environmental variables, we developed a species distribution model using Random Forest. This model yielded mapping results of 24 mammalian species' distribution, utilizing the species richness index to generate a 100-meter resolution map of species richness. The research findings exhibited a notably high predictive accuracy, with the species distribution model demonstrating an average AUC value of 0.82. In addition, the comparison with National Ecosystem Survey data reveals that the species richness distribution in the high-resolution species richness mapping results conforms to a normal distribution. Hence, it stands as highly reliable foundational data for environmental impact assessment. Such research and analytical outcomes could serve as pivotal new reference materials for future urban development projects, offering insights for biodiversity assessment and habitat preservation endeavors.

Keywords: Biodiversity; Species distribution model; Mammal; Environmental impact assessment

I. 서론

생물다양성은 유전자와 종부터 생태계에 이르기까지 모든 생명체의 다양성을 의미하며, 인류에게 필수적인 서비스를 제공한다. 이미 많은 연구에서 생물다양성의 중요성을 강조하였는데 이는 생물자원의 확보, 생태계 기능의 안정성 보존 등에서 핵심적인 역할을 하며 인류의 생존에 영향을 끼친다(Ives & Carpenter 2007; Jiang & Pu 2009; Cardinale et al. 2012). 그러나 국제사회에서 생물다양성 보전을 위한 노력에도 불구하고 급격한 도시확장과 인간활동으로 발생된 기후변화가 많은 생물종 및 그들의 서식지에 위협을 주면서 생물다양성이 과거에 비해 빠른 속도로 소실되어 가고 있다(Watson et al. 1995, Butchart et al. 2010). 세계경제포럼에서 2023년에 발표한 향후 10년간 전 세계가 맞이하게 될 10대 리스크 중 절반이 환경과 연관되고 그 중 생물다양성 손실과 생태계 붕괴가 4위를 차지하고 있다(WEF 2023). 이는 국제사회에서 생물다양성 손실에 대한 논의가 대두되고 있으며 환경문제를 비롯한 생물다양성 위기가 현재 사회에서 중시되어야 함을 다시 한 번 강조하였다.

생물다양성의 보존은 지속가능한 개발의 중요한 요소로 환경적으로 건전하여 지속가능한 발전을 구현하

는 핵심정책인 환경영향평가에서 많이 고려되고 있다(Koo & Lee 2012). 생물다양성의 중요성이 전세계적으로 인식되고 생물다양성이 날 따라 감소되고 있는 현황에서 개발 사업을 시행하는 경우 생물다양성에 불리한 영향을 미치는지에 대해서는 환경영향평가에서 생물종에서부터 생태계에 이르기까지 다양한 측면이 고려되어야 한다. 그러나 환경영향평가 중에서도 생물다양성 평가를 포함한 자연생태환경분야는 타 환경 분야에 비해 정량화 하기 어렵다는 점이 있다. 그러므로 생물다양성의 보존과 생물 서식지의 효율적인 보호를 위해서는 환경영향평가 프로세스의 한 부분으로 생물다양성평가를 위한 기초자료가 필요하다.

종 풍부도(Species richness, SR)는 특정한 지리적 범위에서 출현하는 생물종의 총수를 의미하는 것으로, 생물다양성을 평가하는데 가장 간단한 방법으로 널리 사용되고 있다. 종 풍부도는 여러 개 종의 분포를 합산하여 얻을 수 있는데 종 분포모형을 사용하여 종의 서식지분포를 예측하고 결과를 바탕으로 종 풍부도를 얻는 연구가 많이 진행되어 왔다(Myers et al. 2000; Hassall 2012; França & Cabral 2016; McKerrow et al. 2018; Shin et al. 2018; Coll et al. 2019; Li et al. 2020; Hernández-Urcera et al. 2021; Carroll et al. 2022). 국내에서도 종 분포 모형을 이용하여 식물, 포유류 등을 대상으로 한 종 풍부도

분석이 많이 진행되어 왔다(Kwon et al. 2012; Park et al. 2020; Park & Mo 2021). 선행 연구의 주된 초점은 특정 지역 내에서 특정 생물종의 종 다양성 및 풍부도에 관한 분석에 맞춰져 있었다. 이러한 연구들은 지역적, 생물학적 범위를 한정짓고 있어, 종 다양성에 대한 보다 광범위하고 포괄적인 이해를 제한하고 있다. 이런 단일 종의 서식지 분포 예측분석에 비해 전국단위에서 동·식물상을 고려한 생물다양성 분석 연구는 아직 미흡하다. 또한 전체 포유류 분류군을 대상으로 한 종 풍부도의 평가분석에 관한 연구도 많이 진행되지 못했다.

생물다양성 평가를 위한 생물종 데이터를 얻기 위해 한국 환경부는 1986년부터 전국자연환경조사를 시작하였다. 제3차 전국자연환경조사부터는 1:25,000 지형도를 위·경도 2'30" 단위로 나눈 격자를 기본단위로 조사를 진행하였다. 이를 통해 GIS 기반의 종 데이터를 제공하고 한국의 생물다양성을 분석할 수 있었지만 기존 종 출현데이터의 부족과 상대적으로 큰 격자단위로 인해 넓은 범위에서 종 풍부도를 평가하는 것에는 한계가 있었다. 한국에서는 제2차 및 제3차 전국자연환경조사 데이터를 활용하여 전국단위로 포유류의 종 풍부도와 희귀도를 분석한 연구가 진행되었다(Kim et al. 2014). 제2차('97-'05)와 제3차('06-'13)조사에서 총 64종의 포유류가 확인된 것과 비교하면, 제4차 전국자연환경조사에서는 총 54종의 포유류가 발견되었다. 더욱 세밀하고 광범위한 조사를 실시했음에도 불구하고 발견

된 포유류 종수가 줄어든 것으로 나타났다. 이는 10여년 동안 기후변화, 개발사업 등의 영향으로 인해 생물 다양성이 감소되고 있다는 것을 나타낸다. 그러므로 새로운 종 데이터로 다시 한국의 종 풍부도를 분석함으로써 변화하는 환경 조건과 인간의 영향이 생물다양성에 미치는 현재의 영향을 더 정확하게 반영하고자 한다.

본 연구에서는 기존의 생물종 데이터의 한계점을 보완하기 위해 Random Forest(RF)를 활용하여 종 분포모형을 구축하여 제4차 전국자연환경조사 데이터를 기반으로 한국의 포유류 종 분포를 예측함으로써 더욱 정밀하고 해상도가 높은 종 풍부도 지도를 얻고자 한다. 이를 통해 더욱 신뢰성 있는 생물다양성 데이터를 제공하고 환경영향평가중의 자연생태환경분야에 정량화된 시각화 자료를 제공할 수 있다.

II. 연구방법

본 연구는 크게 2단계로 진행되었다(Figure 1). 첫 단계에서는 포유류 대상을 선정한 후 대상종들의 서식지 분포에 영향을 주는 환경변수를 구축하고 종 분포모형중의 Random forest 모형을 사용하여 종들의 서식 분포를 예측하였다. 다음 단계에서는 모든 대상종들의 서식예측 분포도를 합산하여 전국단위의 포유류 종 풍부도를 도출하고 분석하였다.

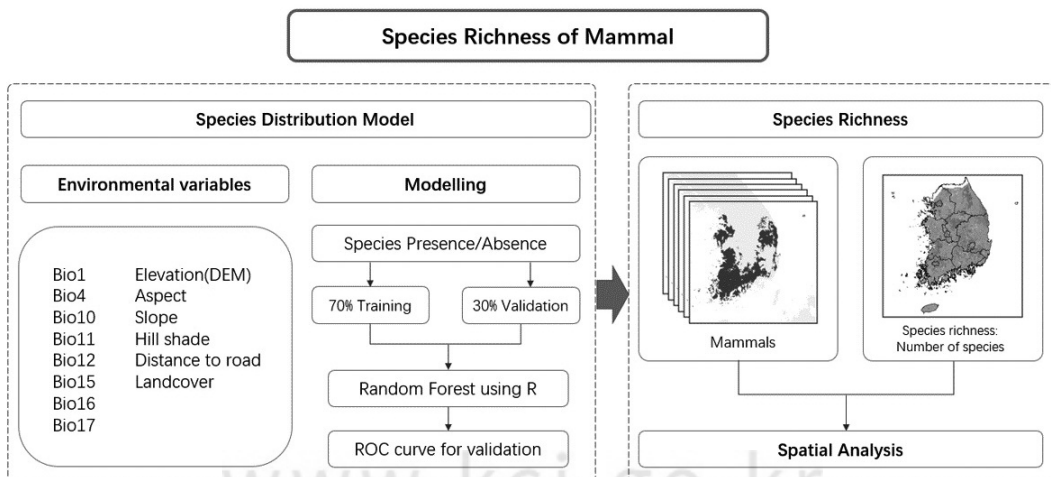


Figure 1. Research flow

Table 1. Target species

| Species | Count | Species | Count |
|----------------------------------|-------|----------------------------------|-------|
| <i>Hydropotes inermis</i> | 12195 | <i>Martes flavigula</i> | 755 |
| <i>Prionailurus bengalensis</i> | 4410 | <i>Rattus norvegicus</i> | 141 |
| <i>Mogera wogura</i> | 4173 | <i>Naemorhedus caudatus</i> | 128 |
| <i>Sus scrofa Linnaeus</i> | 3842 | <i>Pteromys volans</i> | 100 |
| <i>Nyctereutes procyonoides</i> | 3141 | <i>Pipistrellus abramus</i> | 80 |
| <i>Meles leucurus</i> | 2616 | <i>Apodemus peninsulae</i> | 74 |
| <i>Sciurus vulgaris Linnaeus</i> | 2113 | <i>Crociodura lasiura</i> | 70 |
| <i>Eutamias sibiricus</i> | 1942 | <i>Myotis petax</i> | 59 |
| <i>Capreolus pygargus</i> | 1323 | <i>Micromys minutus</i> | 57 |
| <i>Mustela sibirica</i> | 1118 | <i>Erinaceus amurensis</i> | 46 |
| <i>Lepus coreanus</i> | 969 | <i>Rhinolophus ferrumequinum</i> | 41 |
| <i>Apodemus agrarius</i> | 767 | <i>Crociodura shantungensis</i> | 33 |

1. 종 분포모형을 위한 대상종 선정

본 연구는 전국을 대상으로 제4차 전국자연환경조사('14-'18) 포유류 위치자료를 활용하였다. 생태계 구성요소 중 포유류는 소속 생태계내의 최상위 소비자로서 해당 생태계의 안정성 여부를 판단할 수 있는 근거로 된다. 제4차 전국자연환경조사에서는 총 54종의 포유류 데이터가 기록되었다. 종분포모형 구축을 위해 공간통계 분석을 위한 최소 데이터 수가 필요하며, 일반적으로 최소 7개 이상의 데이터가 요구된다(Flancklin 2009). 본 연구는 전국 범위의 포유류 종 풍부도 지도 작성을 위한 종 분포 예측의 정확성과 신뢰성을 보장하

기 위해 위치자료 수가 30지점 이상인 24종을 대상으로 하였다(Table 1). 그 중 산양은 멸종위기 야생생물 I급 종이고 삿, 담비, 하늘다람쥐는 멸종위기 야생생물 II급 종이다.

2. 환경변수 구축

기존의 서식지 관련 연구 및 대상종과 관련한 연구를 참고하여 기후(기온, 강수), 지형(향, 표고, 지형기복), 거리(도로로부터의 거리), 토지피복 등의 변수들을 100m 공간해상도로 구축하였다(Table 2). 기후자료는 Worldclim에서 제공하는 기후자료인 BIOCLIM자료를

Table 2. Environmental variables

| Variables | Description | Data source |
|-------------|-------------------------------------|--|
| Bio1 | Annual mean air temperature | WorldClim data website |
| Bio4 | Temperature seasonality | |
| Bio10 | Mean temperature of warmest quarter | |
| Bio11 | Mean temperature of coldest quarter | |
| Bio12 | Annual precipitation | |
| Bio15 | Precipitation of the driest quarter | |
| Bio16 | Precipitation of the driest quarter | |
| Bio17 | Precipitation of wettest quarter | |
| Aspect | Aspect | |
| Slope | Slope | |
| Hillshade | Hillshade | |
| DEM | Elevation | |
| Dis_to_road | Distance to road | Environmental Geographic Information Service |
| Landcover | Landcover | |

사용하였으며, 19개의 BIOCLIM 자료 중 PCA(Principal Components Analysis)분석을 통하여 전체 변수를 대표할 수 있는 8개의 변수(Bio1, Bio4, Bio10, Bio11, Bio12, Bio15, Bio16, Bio17)를 선택하였다(Seo et al., 2009). DEM (Digital Elevation Model)을 통해 지형기복, 경사도, 향, 수분지수를 구축하였다. 도로로부터의 거리는 국토교통부의 자료를 사용하였다. 자료구축은 ArcMAP 10.8.2에서 수행하였다.

3. 머신러닝 알고리즘을 이용한 종 분포모형

종 분포모형(Species Distribution Models, SDMs)은 기존의 종 분포와 환경 변수 간의 통계적 상관성에 의존하고 있다. 이는 관찰된 종의 출현을 특정 지점의 환경 변수 값과 연관시킴으로써 전체지역의 종 분포를 예측할 수 있다(Li & Wang 2013).

종의 분포를 예측하기 위해 현재 다양한 모형들이 사용되고 있다(Seo et al., 2008). 국내에서는 Maxent (Maximum entropy method) 모형, GLM (Generalized linear models)모형, GAM (Generalized additive model) 모형 등이 포유류, 식물 등의 종 분포를 예측분석 하는데 많이 사용되었다(Kim et al. 2014; Lee et al. 2016). GLM모형과 GAM모형은 회귀분석에서 자주 사용되는 모형으로 모형 복잡성이 상대적으로 낮고 쉽게 사용할 수 있다는 장점이 있지만 이런 모형은 작은 연구 범위에 더 잘 적용되고 여러가지 환경변수와의 복잡한 관계를 예측함에 있어서는 모형 정확도가 낮을 수 있다(Aertsen et al. 2010).

Random forest는 여러 개의 의사결정나무(Decision tree)를 결합한 모델로 하나의 의사결정나무보다 더 우수한 예측 성능을 갖는다(Pavlov 2019). 일부 의사결정 나무가 잘못된 예측을 진행하더라도 여러 의사결정 나무 모델의 예측 결과를 종합함으로써 정확한 예측이 가능한 것에 있다(Bae & Kim 2021). 본 연구에서는 제4차 전국자연환경조사에서 제공되는 종 출현 자료와 비출현 자료를 기반으로 Random forest 모델을 통해 전국 범위에서 각 대상종의 종 분포를 예측하였다. 비출현 자료는 ArcMAP 10.8.2에서 각 생물종의 출현자료와 같은 수량으로 랜덤으로 생성하였다. 종 분포모형의 구축은 R 4.3.1을 이용하였다. 모형예측의 정확도는

ROC (Receiver Operating Characteristic)의 AUC (Area Under Cover) 값을 통해 측정하였다.

4. 종 풍부도 지수 평가

종 분포모형의 결과는 종분포확률로 나타나는데, 이 결과를 출현/비출현 지역으로 예측하기 위한 임계값(threshold)이 필요하다. 본 연구에서는 출현지역의 예측률과 비출현지역의 예측률의 합이 최대가 되는 값을 기준으로 하여 0(비출현)과 1(출현)을 갖는 이항형 자료를 구축하였다(Hu & Jiang 2011; Kim et al. 2014). 0과 1로 결과가 도출된 각 종의 분포예측결과를 합산하여 종 풍부도 지도를 생성하였다.

$$SR = \sum_{i=1}^n S_i$$

n = 대상지역의 총 종수

S_i = 대상종이 출현하는 격자(1의 값을 가진 격자)

마지막으로 본 연구에서 구축된 고해상도의 종 풍부도 매핑 결과와 전국자연환경조사 데이터를 통해 구축된 매핑 결과간 비교를 통해 해당 연구 결과에 대해 검토하였다.

III. 연구결과 및 고찰

1. 모형예측 정확도

포유류 24종에 대한 종분포모형을 구축하였고 각각의 모형 정확도는 ROC (Receiver Operating Characteristic)의 AUC (Area Under Cover) 값을 통해 측정하였다 (Table 3). 전체 대상종의 분포예측 결과 AUC값이 최대로 나타난 종은 우수리박쥐(0.97)이고 최소로 나타난 종은 우수리맛쥐(0.679)이다. 전체 대상종의 평균AUC 값은 0.82로서 우수한 모형 정확도를 보였다.

국내에서는 대상종 중 멸종위기 야생생물II급 종인 삿과 담비의 서식지분석과 분포에 관한 연구가 많이 진행해왔다(Kim et al. 2012; Jin et al. 2015; Lim et al. 2017). 본 연구에서는 종 분포모형 예측결과로 삿과 담비의 결과를 예시하였다(Figure 2). 삿과 담비의 분포예측 결과 AUC값은 각기 0.8과 0.88로서 높은 모형 정확도를 보였다. 삿의 경우 한국 남부의 산림지대에 많이 출현할 것으로 나타나고 담비의 경우 고도와 토지피복 유형이

Table 3. AUC of species distribution model of mammals

| Species | AUC | Species | AUC |
|----------------------------------|-------|----------------------------------|-------|
| <i>Hydropotes inermis</i> | 0.818 | <i>Martes flavigula</i> | 0.880 |
| <i>Prionailurus bengalensis</i> | 0.800 | <i>Rattus norvegicus</i> | 0.792 |
| <i>Mogera wogura</i> | 0.744 | <i>Naemorhedus caudatus</i> | 0.965 |
| <i>Sus scrofa Linnaeus</i> | 0.810 | <i>Pteromys volans</i> | 0.910 |
| <i>Nyctereutes procyonoides</i> | 0.783 | <i>Pipistrellus abramus</i> | 0.943 |
| <i>Meles leucurus</i> | 0.857 | <i>Apodemus peninsulae</i> | 0.680 |
| <i>Sciurus vulgaris Linnaeus</i> | 0.778 | <i>Crocodyra lasiura</i> | 0.679 |
| <i>Eutamias sibiricus</i> | 0.794 | <i>Myotis petax</i> | 0.970 |
| <i>Capreolus pygargus</i> | 0.858 | <i>Micromys minutus</i> | 0.905 |
| <i>Mustela sibirica</i> | 0.712 | <i>Erinaceus amurensis</i> | 0.769 |
| <i>Lepus coreanus</i> | 0.798 | <i>Rhinolophus ferrumequinum</i> | 0.900 |
| <i>Apodemus agrarius</i> | 0.761 | <i>Crocodyra shantungensis</i> | 0.758 |

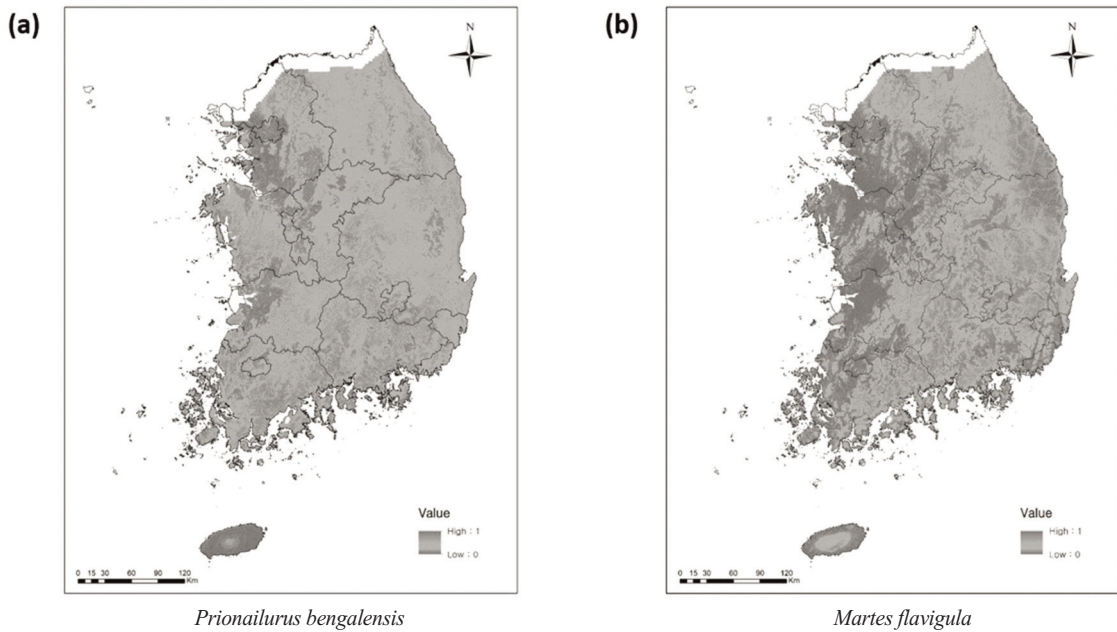


Figure 2. Species distribution of target species

종 분포에 대한 영향이 컸으며 한국 동부의 산맥지대에 많이 출현할 것으로 나타낸다.

2. 전국 포유류 종 풍부도

각 대상종의 분포예측결과를 합산하여 100m 공간해상도의 종 풍부도를 도출하였다(Figure 3). 전체 종을 대상으로 종 풍부도를 합산한 결과 0-23까지의 범주를 가지고 있다. 각 종 풍부도의 면적을 통계한 결과 5종

이상의 생물종이 출현할 것으로 예측된 지역은 전체 면적의 68.27%를 차지하고 15종 이상의 생물종이 출현할 것으로 예측된 지역은 전체면적의 4.36%를 차지한다.

선행연구에 의하면 포유류의 서식지 분포는 주로 고도, 지형기복 등 지형적 요인의 영향을 받는 것으로 나타났다(Kim et al. 2012). 본 연구에서 종 풍부도에 영향을 주는 환경변수를 분석한 결과 고도, 도로로부터의 거리와 가장 건조한 분기의 강수량이 종 풍부도에



Figure 3. Species richness of mammals in Korea

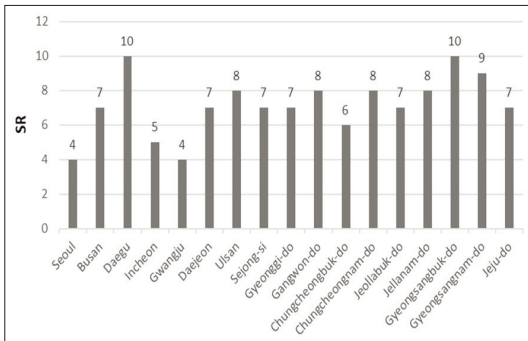


Figure 4. The average species richness of each region

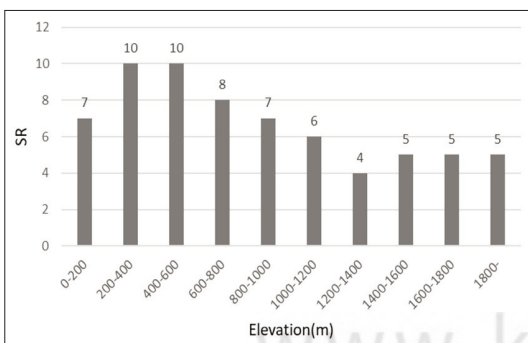


Figure 5. The average species richness of each elevation

중요한 영향을 준다고 나타났다. 이것은 백두대간을 중심으로 주변의 높은 산맥과 지형기복이 심한 지대에 종 풍부도가 높게 나타난 것에서 검증할 수 있다. 특히 경기도와 강원도의 인접지대와 강원도와 경상북도 인접지대의 태백산, 응봉산, 청량산 등 산맥에서의 종 풍부도가 높게 나타났다.

지역차원으로 보면 평균적으로 대구시의 종 풍부도가 가장 높고 서울시와 광주시의 종 풍부도가 가장 낮게 나타났다(Figure 4). 종 풍부도와 고도 사이의 관계를 보면 고도가 1200m이하의 지역에 생물종이 더 많이 출현되고 특히 200-600m 고도의 종 풍부도가 상대적으로 높게 나타났다(Figure 5).

3. 고 해상도 매핑 결과와 전국자연환경조사 데이터간 비교

본 연구에서는 제4차 전국자연환경조사 자료를 이용하여 고해상도 종 분포모형의 결과와 원 자료로 생성된 매핑 결과를 비교하였다(Figure 6). 제4차 조사의 생물종 데이터로 생성된 종 풍부도 지도는 아래와 같은 한계점이 있다. 첫째, 조사의 격자 단위가 상대적으로 크기에 세부적인 범위에서의 종 풍부도 정보를 얻을 수 없다. 둘째, 인위적인 현장조사이기 때문에 사람이 접근하기 어려운 지역의 조사에는 제한적이고 종 출현 데이터가 부족하다. 종 분포모형을 활용한 전국 단위의 고 해상도 종 풍부도 매핑은 이런 한계점을 일부 보완할 수 있었다.

토지피복 데이터에서 산림지역내 종 풍부도 분포를 비교했을 때, 전국자연환경조사 데이터를 통해 구축된 매핑 자료는 0-18의 보다 좁은 종 풍부도의 범주를 나타냈고 산림지역 내 종 풍부도 값이 0인 격자가 전체 격자수의 39%를 차지하였다(Figure 7). 이는 종 출현 데이터만 사용했을 때 인위적인 영향을 많이 받았기에 전국범위에서의 정확한 종 풍부도를 얻기에는 미흡하다는 것을 알 수 있다. 한편 고 해상도의 종 풍부도 매핑은 종 풍부도의 범주가 더 넓고 각 종 풍부도 값이 정규분포의 형태를 가지고 있어 보다 높은 신뢰성을 가지고 있다.

본 연구에서 구축된 종 풍부도 매핑 결과는 해상도가 높으므로, 환경영향평가 동물상 부문에서 현황조사 및 사업으로 인한 영향예측 단계에서 정보를 제공할 수

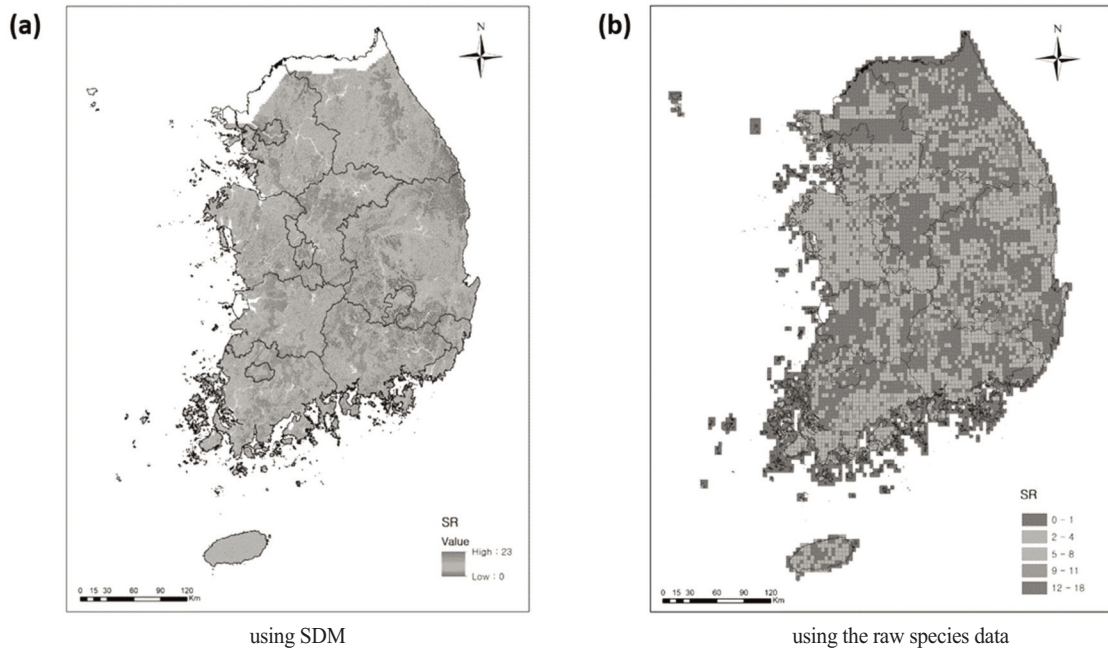


Figure 6. Species richness of mammals in Korea using different data

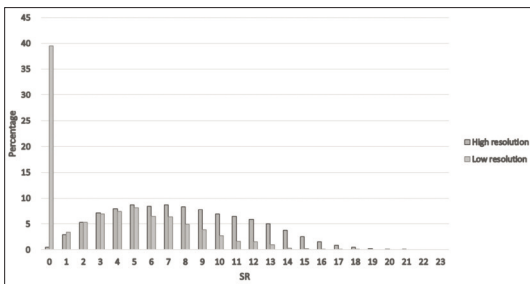


Figure 7. Frequency distributions of species richness

있다. 특히 개발 사업 면적이 작은 경우, 전국자연환경 조사 해상도에서의 데이터는 활용하기 어려우므로, 해당 연구의 매핑 결과가 중요한 기초자료로 활용될 수 있을 것으로 판단된다. 고 해상도의 종 풍부도 매핑은 전국자연환경조사의 공간적 해상도 한계점을 보완하고 신뢰성 있는 기초자료로 사용할 수 있는데 그 의의가 있다. 하지만 종 분포모형을 사용함에 있어 각 종에 대한 환경변수를 구축하지 않은 점과 데이터수가 부족하여 해당 연구에 활용되지 못한 부분은 본 연구에서의 한계점으로 나타난다. 이는 추후 종 별로 부동한 환경변수를 구축하고, 데이터 수가 부족한 생물종에 대한 추가적인 모델링 연구가 필요하다.

IV. 결론

본 연구는 생물다양성의 분포 패턴을 조사하기 위해 제4차 전국자연 환경 조사 데이터를 활용하여 생태계 핵심 종 중 포유류를 대상으로 Random forest 모델을 활용하여 종 풍부도를 분석하였다. 종 풍부도 매핑은 100m 해상도에서 구축되었으며, AUC 값이 0.82로 높게 나타났다. 또한 고도, 도로로부터의 거리, 강수량이 종 풍부도에 중요한 환경 변수로 확인되었다. 특히 백두대간을 중심으로 주변의 고지대와 도로에서 멀리 떨어진 지역에서 종 풍부도가 높게 나타났다. 하지만 본 연구에서는 모델링의 신뢰도를 위해 종 출현 수가 30 미만인 종은 제외되었다. 이로 인해 정확한 생물종의 분포를 반영하는 데 일부 한계가 있으나, 환경영향평가 동물상에서 사업으로 인한 예측부문에서 기초자료로 활용될 것으로 사료된다.

매핑 결과의 신뢰도는 기존에 활용중인 전국자연환경조사 데이터를 기반으로 구축된 종 풍부도 매핑 결과와의 비교를 통해 검토할 수 있다. 산림지역의 종 풍부도 분포를 비교했을 때, 전국자연환경조사 데이터는 0인 값이 많이 나타났지만 고 해상도의 종 풍부도 매

핑 결과의 종 풍부도 분포는 정규분포의 형태를 가지고 있어 환경영향평가 동물상 부문에서 현황조사 및 사업으로 인한 영향예측 단계에서 중요한 참고자료로 사용될 수 있다.

해당 연구는 머신러닝 알고리즘을 이용하여 고해상도 종 분포 매핑 구축을 통해 환경영향평가 생물다양성 평가에 정보를 제공할 수 있다. 또한 소규모 개발 프로젝트에서 매핑 결과를 기초자료로 활용할 수 있다는 점에도 의의가 있다. 향후 환경영향평가 모니터링 및 영향평가의 실효성을 제고하기 위해서는 종 별 환경변수 구축과 데이터 수가 부족한 생물종에 대한 모델링 연구가 지속적으로 수행되어야 한다.

사사

본 결과물은 환경부의 재원으로 한국환경산업기술원의 ICT기반 환경영향평가 의사결정 지원 기술개발 사업(2021003360002)의 지원을 받아 연구되었습니다.

References

- Aertsen W, Kint V, Van Orshoven J, Özkan K, Muys B. 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological Modelling* 221, 1119-30.
- Bae J, Kim S. 2021. Predictions of COVID-19 in Korea Using Machine Learning Models. *Journal of the Korean Institute of Industrial Engineers*, 47(3), 272-279, 10.7232/JKIE.2021.47.3.272. [Korean Literature]
- Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA, Baillie JEM, Bomhard B, Brown C, Bruno J, Carpenter KE, Carr GM, Chanson J, Chenery AM, Csirke J, Davidson NC, Dentener F, Foster M, Galli A, ... Watson R. 2010. Global Biodiversity: Indicators of Recent Declines. *Science*, 328(5982), 1164-1168.
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava DS, Naeem S. 2012. Biodiversity loss and its impact on humanity. *Nature*, 486(7401), Article 7401.
- Carroll KA, Farwell LS, Pidgeon AM, Razenkova E, Gudex-Cross D, Helmers DP, Lewinska KE, Elsen PR, Radeloff VC. 2022. Mapping breeding bird species richness at management-relevant resolutions across the United States. *Ecological Applications*, 32(6), e2624.
- Chung H, Choi Y, Ryu J, Jeon S. 2020. Accuracy Evaluation of Potential Habitat Distribution in *Pinus thunbergii* using a Species Distribution Model: Verification of the Ensemble Methodology. *Journal of Climate Change Research*, 11(1), 37-51. [Korean Literature]
- Coll M, Pennino MG, Steenbeek J, Sole J, Bellido JM. 2019. Predicting marine species distributions: Complementarity of food-web and Bayesian hierarchical modelling approaches. *Ecological Modelling*, 405, 86-101.
- França S, Cabral HN. 2016. Predicting fish species distribution in estuaries: Influence of species' ecology in model accuracy. *Estuarine, Coastal and Shelf Science*, 180, 11-20.
- Franklin J. 2009. *Mapping species distributions spatial inference and prediction*, Cambridge: Cambridge University Press.
- Hassall C. 2012. Predicting the distributions of under-recorded Odonata using species distribution models. *Insect Conservation and Diversity*, 5(3), 192-201.
- Hernández-Urcera J, Murillo FJ, Regueira M, Cabanellas-Reboredo M, Planas M. 2021. Preferential habitats prediction in syngnathids using species distribution models. *Marine Environmental Research*, 172,

- 105488.
- Hu J, Jiang Z. 2011. Climate Change Hastens the Conservation Urgency of an Endangered Ungulate, PLOS ONE, 6(8), e22873.
- Ives AR, Carpenter SR. 2007. Stability and Diversity of Ecosystems. Science, 317(5834), 58-62.
- Jiang L, Pu Z. 2009. Different Effects of Species Diversity on Temporal Stability in Single-Trophic and Multitrophic Communities. The American Naturalist, 174(5), 651-659.
- Jin LS, Kim J, Park Y-C. 2015. Analysis of habitat characteristics of leopard cat (*Prionailurus bengalensis*) in Odaesan National Park. Journal of Agriculture & Life Science, 49(3), 99-111. [Korean Literature]
- Kim J, Kwon H, Seo C, Kim M. 2014. A nationwide analysis of mammalian biodiversity hotspots in South Korea. Journal of Environmental Impact Assessment, 23(6), 453-465. [Korean Literature]
- Kim J, Seo C, Kwon H, Ryu J, Kim M. 2012. A Study on the Species Distribution Modeling using National Ecosystem Survey Data. Journal of Environmental Impact Assessment, 21(4), 593-607. [Korean Literature]
- Koo M, Lee D. 2012. A Study on the National and International Research Trend of Biodiversity Assessment method and Its Application of Environmental Impact Assessment. Journal of Environmental Impact Assessment, 21(1), 119-132. [Korean Literature]
- Kwon H, Seo C, Park C. 2012. Development of Species Distribution Models and Evaluation of Species Richness in Jirisan region. Journal of Korean Society for Geospatial Information Science, 20(3), 11-18. [Korean Literature]
- Lee S, Cho KH, Lee W. 2016. Prediction of Potential Distributions of Two Invasive Alien Plants, *Paspalum distichum* and *Ambrosia artemisiifolia*, Using Species Distribution Model in Korean Peninsula. Ecology and Resilient Infrastructure, 3(3), 189-200. [Korean Literature]
- Li J, Fan G, He Y. 2020. Predicting the current and future distribution of three *Coptis* herbs in China under climate change conditions, using the MaxEnt model and chemical analysis. Science of The Total Environment, 698, 134141.
- Li X, Wang Y. 2013. Applying various algorithms for species distribution modelling. Integrative Zoology, 8(2), 124-135.
- Lim C, Lee C, Jung S, Park Y. 2017. A Study on the Trail Mangement in National Park Using Habitat Suitability Assessment: A Case Study of Yellow-throated Marten Habitats in Mt. Mudeung National Park. Journal of the Korea Society of Environmental Restoration Technology, 20(4), 63-75. [Korean Literature]
- McKerrow AJ, Tarr NM, Rubino MJ, Williams SG. 2018. Patterns of species richness hotspots and estimates of their protection are sensitive to spatial resolution. Diversity and Distributions, 24(10), 1464-1477.
- Park C, Mo Y. 2021. Impact of Climate Change on Urban Bird Species Richness and the Importance of Urban Green Spaces. Journal of Climate Change Research, 12(5-1), 371-381. [Korean Literature]
- Pavlov YL. 2019. Random Forests, 1-122.
- Seo C, Park Y, Choi Y. 2008. Comparison of Species Distribution Models According to Location Data. Journal of the Korean society for geospatial information system, 16(4), 59-64. [Korean Literature]
- Seo C, Thorne JH, Hannah L, Thuiller W. 2009. Scale effects in species distribution models: implications for conservation planning under climate change, Biology Letters, 5(1), 39-43. [Korean Literature]
- Shin M-S, Seo C, Lee M, Kim J-Y, Jeon J-Y, Adhikari P, Hong S-B. 2018. Prediction of Potential Species

Richness of Plants Adaptable to Climate Change in the Korean Peninsula. *Journal of Environmental Impact Assessment*, 27(6), 562-581. [Korean Literature]

Watson RT, Heywood VH, Baste I, Dias B, Gámez R, Janetos T, Reid W, Ruark G. 1995. Global

Biodiversity Assessment, Summary for Policy-Makers, Cambridge University Press, Cambridge (published for the United Nations Environment Programme).

WEF. “Global Risk 2023” Report.