

# 해양 생물 종 분포 추정 최근 연구 동향 및 향후 방향: 해양 포유류와 바닷새를 중심으로

## Recent Research Trends and Future Directions on Marine Species Distribution Modeling: Focusing on Marine Mammals and Seabirds

김영준\* · 정시훈\*\* · 성태준\*\* · 임정호\*\*\*

Young Jun Kim · Si Hun Jung · Tae Jun Sung · Jungho Im

**요약** 종 분포 모델은 생물 종의 관측 정보와 관련된 환경 변수 사이의 관계를 분석하여 생물의 출현 가능성 혹은 생물량을 모의하는 모델이다. 각 생물 종들의 기후변화 등에 의한 서식 환경변화에 대한 영향평가와 같은 활용이 가능하지만, 우리나라의 해양 생물 종 분포 추정과 관련된 연구는 아직 미흡한 상황이다. 본 연구는 Scopus 및 Web of Science에 등재된 저널을 기준으로 지난 10년(2011-2020)간 게재된 해양 분야 종 분포 추정 연구 200편에 대한 동향 및 키워드를 분석하고, 최근 5년(2016-2020)간 게재된 해양 포유류 및 바닷새 분포 추정과 관련된 연구 62편에 대한 분석 대상(종), 분석 지역, 분석 기법에 대하여 조사·분석하였다. 해양 종 분포 추정에 대한 연구는 점점 증가하고 있으나, 우리나라 해역에 대한 연구는 없는 것으로 나타났다. 해양 포유류 및 바닷새 분포 추정에 대한 최근 연구는 단일 모델보다는 둘 이상의 모델을 활용한 비교를 수행하며, 기존에 쓰이던 통계모델의 개선된 형태의 모델을 사용하는 것으로 나타났고, 두 가지 이상의 방법을 동시에 활용하는 앙상블(ensemble) 기법도 활용되고 있는 것으로 나타났다. 추후 보다 체계적인 해역이용영향평가를 위해서는 해양 생물 관측 자료의 보완, 다양한 종 분포 추정 모델의 활용 및 앙상블 형태의 접근, 해역이용평가 시 해양포유류 및 바닷새 종 별 고유의 특성을 반영할 수 있는 환경변수에 대한 연구가 필요하다고 판단된다.

**키워드** 종 분포 추정, 해양, 해양 포유류, 바닷새, 리뷰

**Abstract** Species Distribution Models (SDMs) examine the relationships between species occurrence observations and environmental factors. SDMs are widely used to assess the impact of climate change to the habitats of each species. However, SDMs of marine species in Korea have had minimum exploration. The present study reviewed over the 200 articles on marine SDMs published between 2011 and 2020, and specifically examined 62 articles on SDMs for marine mammals or seabirds published between 2016 and 2020 based on the Scopus and Web of Science databases. The number of studies on marine SDMs has been increasing in recent years, but there is none conducted in Korea. Recent studies on SDMs of marine mammals and seabirds tended to focus on comparison of multiple approaches, trying to propose improved models when compared to the existing statistical approaches. Ensemble approaches combining statistical and machine learning models have been adopted in very recent years. In order to establish a more sophisticated marine environmental impact assessment system in Korea, it is crucial to (1) collect a large number of in-situ measurements of marine lives, (2) use sophisticated modeling and ensemble approaches, and (3) comprehensively examine key environmental variables that can reflect the unique characteristics of marine mammals and seabirds in terms of their habitats and behaviors.

**Key words** Species distribution model, marine, marine mammal, seabird, review

### 1. 서론

종 분포 모델(Species Distribution Model)은 생물 종

의 출현·비출현(presence-absence) 정보와 관련된 환경 변수 사이의 관계를 분석하여 대상 종의 지역별 출현 가능성, 혹은 생물량을 모의하는 모델이다. 이러한 종

Received: 2021.12.18, Revised: 2021.12.31, Accepted: 2021.12.31

\* 한국수리학회, 울산과학기술원 도시환경공학과 박사과정, kimyj@unist.ac.kr (주저자)

\*\* 울산과학기술원 도시환경공학과 석박통합과정

\*\*\* 울산과학기술원 도시환경공학과 교수, ersgis@unist.ac.kr (교신저자)

분포 모델을 활용한 연구는 생물의 출현 확률 모의, 기후변화에 의한 영향, 외래종에 대한 분석, 과거 기후에 대한 분포, 생물량 및 군집분포, 서식지 적합도 분석, 분포 추정 방법론에 대한 분석, 계획 및 관리 등 널리 활용되고 있다 (Melo-Merino et al., 2020). 종 분포 모델을 활용하여 생물 종의 분포에 대한 정량적인 예측이 가능하며, 이를 이용해 각 생물 종들의 특정 개발 사업이나 기후변화 등에 의한 서식 환경변화에 대한 영향을 평가하는데 활용할 수 있다(Austin, 2002; Chung et al., 2020). 국내에서도 각종 대규모 개발 사업 등에 앞서 멸종위기 등급 종 및 보호가 필요한 생물군에 대한 분포 분석을 통해 종에 대한 환경 영향 평가를 수행하고 있다. 이러한 종 분포 모형을 활용하여 대상 생물군의 지리적 분포 및 환경 변수에 대한 상관관계를 파악하여 보호구역을 선정하거나 추후 관리 방안을 수립하는데에 활용된다(Franklin, 2009; Thorn et al., 2009; Chung et al., 2020).

하지만, 대규모 해상풍력 단지 건설 사업 및 바다골재 채취사업 등 해양환경에 대한 환경 및 생물 영향평가에 대한 연구는 많이 진행되지 않은 상황이다. 특히, 우리나라 해역에 대한 해양생물 종 분포에 대한 연구가 많이 수행되지 않은 실정이며, 아시아 연안에만 분포하는 상괘이(Neophocaena asiakorialis)의 경우, 해상 풍력기 설치 등에 민감하다고 알려져 있으며(Oh and Yeo, 2019; Oh et al., 2020), 이에 따른 영향평가를 위한 개체수 분포에 대한 연구가 필요한 실정이다. 현재 우리나라는 총 88종의 해양보호생물을 지정·관리하고 있다. 이러한 해양보호생물은 우리나라 고유종으로써 해당 생물군의 학술적 가치가 높으나 개체수가 감소중인 보호가치가 높은 해양생물 중에서 개체 감소에 대한 위협요인을 체계적으로 관리할 필요가 있다고 판단한 생물을 대상으로 지정한다(Kim et al., 2017; Oh et al., 2020). 88종의 해양보호생물 중, 해송(Myriophathes japonica)과 같은 무척추동물(invertebrate)이 36종(41.0%)으로 가장 많이 지정되어 있으며, 이어서 해양포유류(19종, 21.6%) 및 바닷새(16종, 18.2%) 순으로 지정되어 있다. 해양 포유류 및 바닷새의 경우, 활동 반경이 넓고, 관측 빈도와 개체수가 많지 않아 추정과 관련된 연구 및 추정 정확도가 상대적으로 낮은 실정이다. 해역이용계획수립 및 환경 영향평가에 있어 고래 및 돌고래와 같은 해양포유류 및 바닷새 서식지에의 영향은 필수적으로 고려해야 할 중요한 요소이다.

본 연구에서는 지난 10년(2011-2020)동안 게재된 200편의 국내·외 해양 분야의 종 분포 추정 모델을 활용한 연구논문을 조사하여 동향을 분석하고, 이 중 최

근 5년간(2016-2020) 게재된 국내·외 해양포유류 및 바닷새와 관련한 종 분포 모델 활용 연구 62편을 조사·분석하여 최근 연구에서 주로 활용하고 있는 종 분포 추정 방법론에 대하여 비교·분석하였다. 분석 결과를 바탕으로 향후 우리나라 해양 포유류 및 바닷새 분포 추정 연구를 위한 방향을 제시하고, 더 나아가 해역 이용에 대한 영향평가 기법의 고도화를 위한 방안을 제시하였다.

## 2. 연구 방법 및 연구 자료

본 연구에서는 Scopus 및 Web of Science에 등재된 저널을 기준으로 문헌조사를 수행하였으며, 종 분포 모델 기법을 활용하여 해양 포유류 및 바닷새의 분포를 추정하거나 서식지 적합도를 평가한 연구를 대상으로 분석을 수행하였다.

### 2.1 연구 방법 및 자료

국내·외 종 분포 추정 연구 관련 최근 동향을 분석하기 위하여 Scopus 및 Web of Science 기준 최근 10년(2011-2020)간 게재된 논문을 대상으로 문헌조사를 수행하였다. 키워드는 “species distribution models”와 “sea”, “marine”, “ocean”을 조합하여 검색하였으며(Melo-Merino et al., 2020), 중복된 검색 결과는 제외하였다. 이 외에 종 분포 추정 및 서식지 적합도 분석 등과 관련이 없는 검색결과, 종 분포 추정 모델을 활용하지 않은 연구 또한 분석에서 제외하였다. 총 200편의 문헌이 검색되었으며, 연도별 게재 편수와 생물 종과 관련된 주요 키워드 동향을 분석하였다.

또한 200편의 선행연구 중, 최근 5년(2016-2020) 사이에 게재된 해양 포유류 및 바닷새 분포 추정과 관련된 연구논문을 기준으로 해양포유류 및 바닷새 종 별 분석 빈도, 분석 지역, 분포 추정 기법에 대한 동향을 분석하였다. 본 연구에서는 선행연구들과 같이 해양에 서식하는 고래, 바다코끼리 등과 같은 포유류를 해양 포유류로 정의하였고, 해안 지역에 군락을 이루는 등 해양환경에 적응한 갈매기, 바다쇠오리 등과 같은 조류 종으로 정의하였다. Scopus 및 Web of Science 등재지 기준 최근 5년간 해양포유류 및 바닷새 분포 추정 관련 연구는 총 62건으로, 저자, 제목, 출판년도, 등재지널, 분석 대상(종), 분석 지역, 분석 기법에 대하여 비교·분석하였다.

마지막으로 최근 연구 동향 분석을 통하여 향후 우리나라 해역의 해양생물 종 분포 연구의 발전 방향 및 해양환경영향평가 기법 개선을 위한 방안을 제시하였다.

## 2.2 종 분포 추정 모델

기존 널리 활용되고 있는 종 분포 추정 모델의 종류는 크게 통계 기반의 모델과 기계학습 기반의 모델로 구분할 수 있다. 최근엔 기계 학습 기법을 활용한 접근이 보다 보편화되고 있는 추세이며, 기존의 통계 기반의 모델에 대하여 변수간의 공선성(collinearity)에 덜 민감하다는 장점이 있다(Fox et al., 2017). 또한, 기존 통계 및 기계학습 기반의 단일 모델이 보여주는 불확실성에 대한 한계점을 해결하기 위하여 두 가지 이상의 기법을 함께 사용하는 앙상블(ensemble) 등의 다중모형을 적용하는 연구가 증가하고 있다(Kwon, 2014). 이때, 앙상블 모델에 들어간 각각의 추정 모델의 추정 정확도에 비례한 가중치를 적용하여 앙상블을 수행한다(Hunt et al., 2020; Marmion et al., 2009).

### 2.2.1 통계 기반 추정 모델

Generalised Linear Models(GLM)은 선형 모델을 연결 함수를 통해 종속 변수의 분포를 대변하도록 모의한다(Nelder et al., 1972). 일반화 선형 모델은 세가지 성분으로 랜덤 성분(Random component), 체계적 성분(Systematic component), 그리고 연결 함수(Link function)로 구성되어 있다. 첫 번째로, 랜덤 성분은 반응 변수(종속 변수)의 확률 분포를 나타낸다. 두 번째로, 체계적 성분은 설명 변수(독립 변수)들이 결합된 형태를 말한다. 세 번째로, 연결함수는 랜덤 성분과 체계 성분을 연결해주는 함수이다. GLM은 정량적, 그리고 정성적 자료 모두에 적용이 가능하다. 특히, 반응 변수와 설명 변수가 비정규성(비선형성)일 경우 기존의 선형 모델보다 우수하게 모의한다.

Generalised Additive Models(GAM)은 앞서 언급한 GLM과 전체적인 식과 개념 구성은 유사하다. 하지만, 각 변수 간의 비선형적 관계를 모의하기 위해서 GLM과 달리, 각 설명 변수의 비선형 함수를 허용하여 일반적인 선형 회귀모델에서 확장된 가법적(additive) 형태를 가진다(Hastie et al., 2017). 비선형 함수를 활용함에 따라 반응 변수를 보다 우수하게 모의할 수 있다. 하지만, 독립 및 종속변수의 관계가 가법적 이어야 한다는 가정이 요구되며, 많은 설명변수가 있는 조건에서는 각 변수간의 상호작용이 모호해지는 단점을 가진다.

Hierarchical Bayesian model(HBM)은 계층적 베이저안 분포를 가정하여 종 분포를 추정하는 기법으로, 서식지 적합성과 종 탐지 가능성 등 다양한 계층적 과정의 결과로 관찰 정보(관찰 여부 및 개체수)와 관련 매개변수와의 관계를 해석할 수 있다. 이러한 베이저

안 접근 방식은 매개변수간 구조적 계층이 존재하거나 누락된 자료가 포함되는 것과 같은 복잡한 모델을 할 때 특히 적합하다고 알려져 있다(Rivera et al., 2019).

Multi-variate Adaptive Regression Splines (MARS)는 변수선택방법(subset selection method)을 이용해 다양한 입력변수에 적합한 스플라인(spline) 기저 함수를 찾는 기법으로, 입력변수가 많은 다차원의 회귀 분석에 적합하다(Friedman, 1991). MARS는 각 입력 변수에 대한 관측값들을 매듭점(knot point)으로 한 여러 개의 조각인 선형 스플라인들로 도출한다. 다수의 기저 함수들의 곱을 활용해 오차의 제곱합을 최소화시키는 방향으로 전진단계선형회귀(forward stepwise linear regression)방법을 활용하여 불필요한 기저함수쌍은 제거하고 추론에 효율적인 기저함수쌍을 남겨 기저 함수의 계수를 추정한다. 단계적, 조각적인 연산을 통해서 비선형적인 입력자료 추론의 정확도를 향상시켰다. MARS의 학습법은 Classification And Regression Tree(CART)와 유사성을 보인다. CART와 다른 방식으로 추론하는 것으로 보이지만 CART는 추론을 위한 나무를 성장시킨다는 점에서 학습방식의 유사함과 높은 정확도를 보인다. 기계학습 기법의 정확도와 비교할 수 있는 정도의 우수한 추론 기법이다.

Flexible Discriminant Analysis(FDA)은 선형 회귀 모델의 혼합을 기반으로 하는 분류 모델이다(Hastie et al., 1994). 입력 자료의 특성에 따라 적절한 선형 분리를 위해 최적 가중치를 도출한다. 앞서 도출된 가중치를 다중 가법 회귀모델에 활용하여 자료의 특성을 대표할 수 있는 판별 면(기준)을 생산한다(Hallgren et al., 2019; Freidman et al., 2017). FDA는 예측 변수의 수가 많을 때, 우수한 각 변수들 사이의 상호관계 분석, 효율적인 연산, 그리고 이상치에 대한 영향을 적게 받는 다양한 장점을 갖는다. 하지만, 모델의 가중치 연산을 위한 적합한 파라미터 산정의 어려움, 과적합 위험성, 그리고 다른 선형 회귀모델 복잡한 해석으로 인한 최적의 결과 모의에 어려움이 존재한다.

Ecological Niche Factor Analysis(ENFA)는 관측(presence)자료만으로 대규모의 종 분포를 추정하는데 활용되는 다변수 분석 모델로써, 기존 통계기법에서 널리 쓰이는 요소 분석(factor analysis)와 같이 다양한 종 분포와 관련된 환경변수를 몇 가지의 주성분(principal component)로 해석하는데 초점을 맞추고 있으며, 유일한 차이점은 각 주성분은 생태학적 의미를 갖는다는 것에 있다. 야생동물 관리, 서식지 평가, 서식지 예측 등 다양한 분야에서 폭넓게 적용되고 있다(Ouyang and Liu, 2008).

Getis-ord는 특정 현상에 대한 공간적 분포를 모의하

기 위한 통계적 기법으로 특정 공간에 집중 또는 분산되어있는 정도를 모의하는 기법이다(Ord and Getis, 1995; Getis and Ord, 2010). Getis-ord은 전역적 및 국지적 군집패턴 모두 구분하여 모델링을 할 수 있다(Lee and Seung, 2016). 전역적 군집분석은 General G 또는 Moran's I 통계량을 이용하며, 국지적 군집분석은 Getis-ord Gi, Local Indicator of Spatial Association (LISA)를 활용한다. 대표적으로 가장 널리 활용되어지는 Getis-ord Gi는 z-score와 p-value를 활용하여 인접한 영역에서 유사한 특징을 가진 변수들 간의 상관관계가 높거나 낮은 지역에 대한 공간적 군집분석을 한다.

### 2.2.2 기계학습 기반 추정 모델

트리 기반(Tree-based) 기법은 의사결정나무(Decision tree)를 기반으로 하는 가장 널리 사용되는 기계학습 모델이다. 생물 종 분포 추정 분야에서는 기본적인 Classification and Regression Trees(CART) 및 Classification Tree Analysis (CTA) 부터, 이를 기반으로 하는 앙상블(Ensemble) 기법인 Random Forest(RF)와 Gradient Boosting Machine(GBM) 등 다양한 트리 기반 모델을 활용한 연구가 수행되고 있다. CART는 독립변수와 종속변수 사이의 관계를 학습하는 모델로서 의사결정나무를 사용한다(Loh, 2011). 의사결정나무는 특정 기준에 따라 데이터를 이진 분류하는 과정을 반복하여 분류, 또는 회귀 문제를 해결할 수 있는 지도학습 모델이다. 이때, 초기 분류 기준을 루트 노드, 중간 분류 기준을 자식 노드, 가장 마지막 노드를 리프 노드라고 한다. CART는 리프 노드가 가장 균일한 상태로 분류되는 것, 즉 엔트로피가 가장 낮아지는 방향으로 학습한다. CTA는 CART와 같은 원리로 학습하나, 학습 대상이 분류의 형태를 가진다는 차이가 있다. RF는 CART를 기반으로, 다수의 의사결정나무를 학습하는 배깅(Bagging) 기반의 앙상블 모델이다. 부트스트랩(Bootstrap) 기법을 통해 무작위로 선택된 훈련자료와 입력변수에 대해 훈련된 기초 분류기들을 결합하여 개별 결정 나무의 과적합(Overfitting) 문제를 방지할 수 있다(Breiman, 2001). 생성된 개별 의사결정나무들은 분류의 경우 다수결, 회귀의 경우 평균을 통해 결합된다. GBM은 의사결정나무를 사용하는 앙상블 중 부스팅(Boosting) 기반의 모델이다. 분석용 데이터 관측 값의 가중치가 동일한 상태에서 시작되어, 형성된 초기 분류기에 의해 오 분류된 관측 값은 다음 관측 값에 높은 가중치를 주게 되지만, 정 분류된 관측 값은 낮은 가중치를 주는 과정을 반복함으로써 최종 분류기를 형성한다(Ayyadevara, 2018). 이를 통

해 모델의 편향과 분산을 줄일 수 있지만 과적합의 위험이 있다. 여기서 CART와 boosting 기법을 결합한 Boosted Regression Trees (BRT) 기법도 널리 쓰이고 있다.

Support Vector Machine(SVM)은 서포트 벡터(Support vector)와 초 평면(Hyper-plane)이 주요 개념인 기계학습 기법이며, 간결한 알고리즘과 뛰어난 성능을 장점으로 많은 종 분포 연구에 활용되었다. SVM은 주어진 데이터 집합이 투영(Projection)된 공간에서 새로운 데이터가 어느 카테고리에 속할지 판단하기 위해 가장 큰 폭을 가진 선형 경계를 찾는다(Noble, 2006). 이때 해당 결정 경계를 초평면, 경계와 가장 가까이 있는 데이터 샘플들을 서포트 벡터라고 하며, 서포트 벡터와 초평면 사이의 거리를 마진(Margin), 데이터 샘플이 마진 내에 포함되거나 반대쪽에 있는 것을 마진 오류(Margin error)라고 한다. SVM 분류는 마진 오류를 최소화하면서 가장 큰 마진을 가지는 초평면을 찾는 것을 목표로 한다(Noble, 2006). 반대로 회귀 문제에서는 마진 내에 데이터 집합을 최대한으로 포함하는 최소한의 마진을 가지는 초평면을 찾는다(Awad & Khanna, 2015). SVM은 선형 경계를 통해 데이터를 분류하므로, 비선형 분류를 위해 데이터를 더 높은 차원으로 투영하는 과정을 거치게 되며, 이때 연산량을 줄이기 위해 커널 트릭(Kernel trick)을 사용한다(Suthaharan, 2016). SVM은 선형(Linear), 가우시안(Gaussian), 방사형 기저 함수(Radial Basis Function) 등의 커널 함수를 사용하며, 커널 함수의 종류와 파라미터에 따라 성능이 크게 달라질 수 있다(Min & Lee, 2005; Song et al, 2008).

Artificial Neural Network(ANN)은 인간의 신경망 구조와 원리를 모방한 기계학습 알고리즘으로, Deep Neural Network, Convolutional Neural Network 등 다양한 딥러닝 모델의 기초가 되는 기법이다. 인간의 뇌에서 뉴런들이 자극을 받고, 그 자극을 처리하여 일정 임계치(Threshold)를 넘으면 결과 신호를 발산하는 과정에서 착안하였으며, 수학적 뉴런에서 자극은 입력 값, 처리 과정은 가중치(Weight)를 이용한 연산, 임계치는 활성화 함수(Activation function), 발산하는 신호는 출력 값에 해당한다(Jain et al, 1996). ANN 구조는 개별적인 수학적 뉴런이 모여 형성된 입력 층, 은닉 층, 출력 층으로 구성되며, 각 층은 가중치로 연결되어 있다. ANN의 입력 층에서 데이터를 받고, 은닉 층에서 연산을 수행한 뒤 출력 층에서 데이터를 출력한다. 역전파 학습법은 ANN의 가장 대표적인 학습 방법이며, 입력자료가 입력 층에 제공되고 그에 따라 출력 값이 계산되는 전향전파 단계와, 오류를 줄이기 위하여

연결 가중치를 조정하는 후향전파 단계를 반복하며 최적의 가중치를 찾아낸다(Goh, 1995). ANN의 장점은 연산부에 해당하는 은닉층과 뉴런의 개수를 유연하게 조절할 수 있다는 점이지만, 너무 단순한 구조는 학습력을 감소시키고, 너무 복잡한 구조는 과적합 문제를 발생시킬 수 있다(Moghim & Bras, 2017)

Maximum Entropy(Maxent) 모델은 출현 자료만을 가지고 생물 종의 분포를 추정할 수 있는 기계학습 기반의 종 분포모델이다. Maxent 모델은 생물종의 서식지 적합도를 연구 대상지 전체에 대한 확률 분포로 추정한다(Phillips & Dudik, 2008). 이때, 최대 엔트로피 접근법을 기반으로 엔트로피가 극대화되는 방향으로 종 분포를 추정하며, 종 분포와 상관성이 있는 환경변수를 추가함에 따라 해당 엔트로피에 제약을 거는 식으로 작동한다(Lee et al, 2010). Maxent 모델은 종속 변수로 종의 출현자료만을 요구하기 때문에 비 출현 자료를 획득하기 어려운 경우에, 제한된 자료로 보다 예측력이 높은 결과를 도출할 수 있다(Phillips et al., 2004). 이러한 장점으로 Maxent 모델은 많은 종 분포 추정 연구에 활용되고 있으며, 다양한 상황에서 높은 추정 성능을 보이고 있다.

### 3. 연구 결과

#### 3.1 국내/외 해양분야 종 분포 추정 연구 동향

최근 10년간 출판된 해양 분야 종 분포 추정 관련 국내/외 선행 연구는 총 200편으로 나타났다. 지난 10년간 관련 연구가 10년간 꾸준히 증가하고 있는 것으로 나타났으며(Fig. 1), 우리나라 해역을 대상으로 수행된 연구는 없는 것으로 나타났다.

조사된 문헌 200편의 생물 종과 관련된 키워드를 분석한 결과(Fig. 2), 해양 분야 종별 연구 동향은 바닷새(seabird)가 16편으로 가장 많은 것으로 나타났고, 뒤

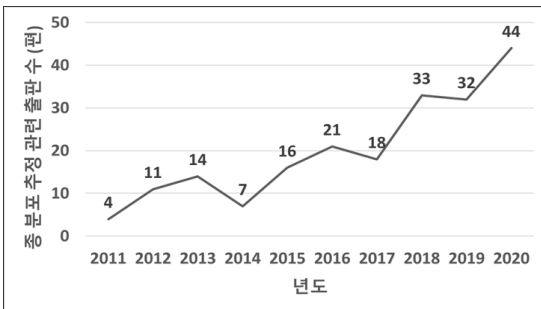


Figure 1. Trends in the number of publications per year regarding species distribution models applied to marine species (2011–2020)

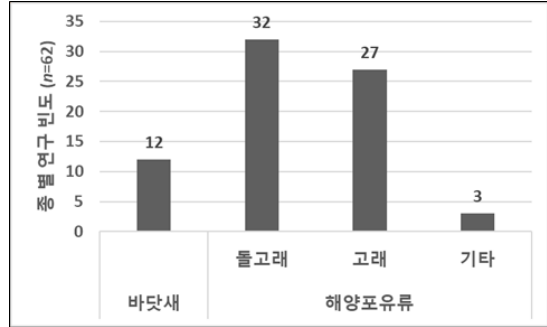


Figure 2. Tendency in the target species of distribution models for marine mammals and seabirds (2016–2020)

를 이어서 산호(coral reef)와 고래류(cetacea)에 대한 연구가 각각 15편과 13편 순으로 나타났다. 그 뒤로 물고기(fishes) 8편 및 저서생물(benthos) 6편의 순으로 나타났다. 바닷새는 ‘seabird’와 ‘bird’ 키워드 빈도를 합한 결과이며, 산호 또한 ‘coral’과 ‘coral reef’ 키워드 빈도를 합한 결과이다. 총 200편의 분석대상에 비하여 분류 결과가 상대적으로 작게 나타났는데, 이는 생물 종에 대한 키워드만 분석한 결과로, 다른 주요 키워드는 종 자체보다는 ‘species distribution model’ (118건), ‘climate change’ (58건), ‘ecological modeling’ (49건) 등 모델링과 기후환경에 초점이 맞춰져 있는 것으로 나타났다.

#### 3.2 해양포유류 및 바닷새 분포 관련 연구 동향

해양포유류와 바닷새 분포 추정과 관련된 총 62건의 선행연구에 대한 분석 결과(Fig. 3), 중복을 포함하여 해양포유류와 관련된 연구가 62건, 바닷새와 관련된 연구가 12건으로, 해양 포유류에 대하여 보다 많은 연구가 진행되고 있는 것으로 나타났다. 해양포유류 중 돌고래와 관련된 연구가 32건, 고래와 관련된 연구가

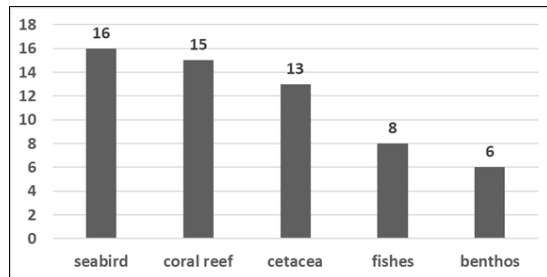


Figure 3. Trends of keywords of publications regarding species distribution models applied to marine species (2011–2020)

27건으로 총 62건의 해양포유류 관련 연구 중 59건(약 95%)이 돌고래 및 고래와 관련된 연구인 것으로 나타났다. 기타 해양포유류 3건은 각각 북극곰(polar bear), 바다사자(sea lion), 매너티로(manatee), 극지 해양에서 주로 서식하는 해양포유류와 관련된 연구에서 조사되었다.

해양포유류와 바닷새 분포 추정과 관련된 총 62건의 선행연구가 수행된 분석 대상 지역에 대한 분석 결과 (Fig. 4), 전 세계 해역을 모두 아우르는 연구는 단 1건인 것으로 나타났으며, 그 외 연구는 모두 지역적인 분석을 수행한 것으로 나타났다. 지역적인 분석을 수행한 선행연구의 경우, 중복을 포함하여 멕시코만을 포함하는 대서양에 대한 연구가 25건으로 가장 활발히 수행되고 있는 것으로 나타났다. 그 뒤 태평양에 대한 분석이 15건, 지중해 지역에 대한 분석이 12건, 북극과 남극을 포함하는 극지해역에 대한 분석이 6건, 이 외 기타 해역에 대한 분석이 5건으로 조사되었다. 기타 해역의 경우 특정 해역으로 구분하기 힘든 해역으로, 페르시아 만(Hemami et al., 2018), 호주남부에 위치한 Coffin Bay (Passadore et al., 2018) 등으로 나타났으며, 우리나라 해역을 포함한 연구는 없는 것으로 나타났다.

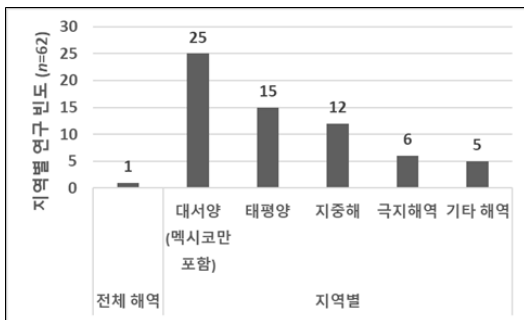


Figure 4. Tendency in the use of species distribution models for marine mammals and seabirds (2016-2020)

### 3.3 해양포유류 및 바닷새 분포 추정 기법 현황

해양포유류와 바닷새 분포 추정과 관련된 총 62건의 선행연구에 활용된 분포 추정 기법에 대한 분석 결과 (Fig. 5), 크게 통계기법과 기계학습기법, 그리고 둘 이상의 추정기법을 동시에 활용하는 앙상블 기법을 활용한 연구로 구분할 수 있다.

중복을 포함하여 통계기법은 총 52번 활용되었으며, GAM, GLM, BHM, MARS, FDA, ENFA, Geties-Ord와 같은 7가지 기법이 활용되고 있는 것으로 나타났

다. 그 중 GAM이 29건(약 56%)으로 가장 많은 빈도를 보였고, 뒤이어 GLM(10건, 약 19%), BHM(6건, 약 12%)인 것으로 나타났다. 중복을 포함하여 기계학습 기법의 경우, 총 44번 활용되었으며, MaxEnt, RF, GBM, BRT, CTA, SVM, ANN와 같이 7가지 기법이 활용되고 있는 것으로 나타났다. 그 중 MaxEnt기법이 19건(약 43%)으로 가장 많이 활용되는 것으로 나타났고, 뒤이어 RF(9건, 약 20%)와 GBM(5건, 약 11%) 순으로 많이 활용되고 있는 것으로 나타났다. Derville et al. (2018)는 기계학습(MaxEnt, BRT, SVM)과 통계기법(GLM, GAM) 총 다섯 가지의 종 분포 추정 알고리즘을 활용하여 동일한 환경변수를 입력하였을 때 뉴칼레도니아 해역의 흑등고래(Megaptera novaeangliae)에 대한 분포 추정 성능을 정량적으로 비교하였다. 해당 연구에서 세 가지 기계학습 기법 모두 통계기법 모델보다 높은 추정 정확도를 보였으나, SVM의 경우, 해석이 어렵다는 점과 BRT 모델의 과적합에 대한 한계점을 제시하였다(Derville et al., 2018).

앙상블 기반의 접근을 한 6편의 선행연구의 경우, 해양포유류 (Hunt et al., 2020; Passadore et al., 2018; Purdon et al., 2020; Zanardo et al., 2017) 뿐만 아니라 바닷새(Fox et al., 2017; Pereira et al., 2018) 분포 추정에도 높은 단일 추정 모형보다 높은 추정능력을 보였다. 특히, Fox et al. (2017)은 같은 앙상블 모델을 활용한 다른 5개의 연구와 다르게 상대적인 출현율(probability of presence) 대신 직접적인 바닷새의 생물량(population density; birds/km<sup>2</sup>)을 추정하였으며, 앙상블 기법이 생물량 추정에도 높은 정확도를 보이는 것을 밝혔다.

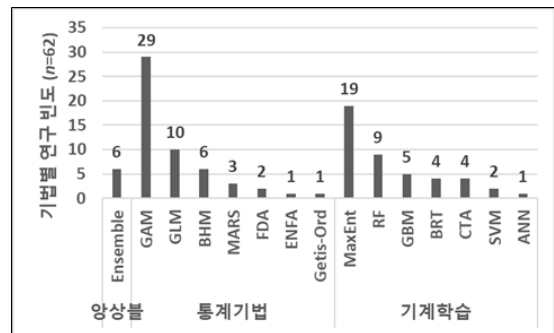


Figure 5. Tendency in the use of species distribution model algorithms for marine mammals and seabirds (2016-2020)

최근 동향의 경우, 단일 모델이 아닌 두 가지 이상의 모델을 활용하여 결과를 비교하는 연구(comparative

study)가 최근 5년간 각각 11, 21, 23, 30, 44%로 점차 증가하고 있는 것으로 나타났다. 같은 종에 대한 분포를 추정할 때, 입력자료의 형태(only presence, presence/absence 등) 및 환경변수의 조합에 따라 두 가지 이상의 독립된 모델의 추정 능력을 비교하여 보다 신뢰할 수 있는 모델을 개발하고 있는 추세이다. Becker et al. (2020)의 경우, 7종의 돌고래와 고래류의 분포 추정을 위하여 통계기반인 GAM과 기계학습 기반인 BRT 모델을 동시에 활용하여 추정 결과를 정성/정량적으로 비교·분석하였다.

또한, 단일 모델만을 활용한 연구의 경우, 기존 널리 활용되던 통계모델이 개선된 형태의 모델을 활용하고 있는 것으로 나타났다. 가장 널리 활용되어온 GAM과 GLM의 경우, zero-inflated GAM (Hemami et al., 2018), GAM-GEE (Bonizzoni et al., 2019), mixed GAM (Thorne et al., 2019), GEE-GLM (Waggitt et al., 2020) 등의 형태로 기존의 단점을 조금씩 보완한 모델을 활용하고 있는 것으로 나타났다.

#### 4. 논의 및 제언

해양 종 분포 모델을 활용한 국내외 연구 동향을 조사한 결과(2011-2020), 우리나라의 해역을 대상으로 분석한 연구는 없는 것으로 나타났다. 이로부터 다음과 같은 논의 및 제언이 가능하다고 보이며, 추후 우리나라의 해양 환경 자체에 대한 이해도와 각종 해양 이용 사업에 의한 영향평가 및 기후변화 영향에 대한 이해 확장을 위해서 해결해야 할 것이다.

##### 4.1 우리나라 해역의 해양생물 관측 자료의 보완

우리나라 해역의 해양생물 종 분포 추정 연구의 활성화를 위하여 보다 시·공간적으로 체계적인 관측 시스템의 구축이 필요하다. 해양포유류 및 바닷새 분포 추정과 관련한 해외 연구 동향 조사 결과(2016-2020), 대다수의 연구들이 관측선, 항공기, 육안 관측 등을 활용한 국가(지자체) 자체 관측 자료 혹은 공개되어있는 다양한 출처의 관측 자료를 융합 활용하고 있는 것으로 나타났다.

우리나라의 경우, 국립수산물과학원에서 고래연구센터 운영을 통해 목시관측을 시행하고 있고, 국립공원 연구원에서 허베이 스피리트호 유류유출 사고와 관련하여 유류에 의한 장기간의 변화 경향을 조사하기 위하여 태안해안국립공원 등 특정도서지역에 대하여 해양포유류에 대한 목시관측이 이루어지고 있으나, 해당 기관들에서 우리나라 해역 전체를 관할하기에는 한계가 있는 실정이다. 현재 고래 조사 또한 일부 지역을

제외하면 선박을 활용한 체계적인 목시관측보다는 어업활동 등에 의해 혼획된 개체수를 활용한 간접 추정 (Lee et al., 2018)이 이루어지고 있는 실정으로, 정량적인 관측자료와 GIS 등을 활용한 체계적인 해양포유류 및 바닷새의 종 분포 추정 모의가 어렵다. 바닷새 조사 또한, 16종의 바닷새가 보호종으로 지정되어있는 반면, 국립공원연구원의 조류연구소에서 신안, 태안, 거제를 기점으로 동북아시아를 이동하는 철새에 대한 국지적인 조사만이 체계적으로 이루어지고 있는 실정이며, Kim et al. (2014)에서 분석한 우리나라 서해 연안 봄철 바닷새 분포 조사 결과 이외에는 뚜렷한 생체량 등에 대한 정보가 부족한 실정이다(Kim et al., 2014). 이러한 자료의 부족을 해결하기 위해서는 보다 다양한 국가 혹은 기관과의 협력을 통하여 시공간적으로 체계적인 관측이 이루어질 필요가 있다.

단일기관에서 확보하기 힘들다는 한계를 극복하기 위하여 Roberts et al. (2016)의 경우, 다음과 같이 다양한 기관의 협업을 통하여 대서양지역의 미국 일부 해역 및 멕시코만의 고래류의 생물량을 추정하였다: NOAA NMFS Northeast Fisheries Science Center (NEFSC), New Jersey Department of Environmental Protection (NJDEP), NOAA NMFS Southeast Fisheries Science Center (SEFSC), University of North Carolina Wilmington (UNCW), Virginia Aquarium & Marine Science Center (VAMSC).

관측 자료량의 부족을 해결할 수 있는 또 하나의 방안으로는 Derville et al. (2017)과 같이 시민 자원봉사자들에 의하여 관측된 해양 포유류 자료를 활용하는 방법도 있다. 해당 연구에서 활용한 시민 참여 자료의 경우, 관측한 정확한 GPS위치 정보와 더불어 다음과 같이 세가지 조건 중 하나가 충족될 때 활용이 가능하다고 판단하였다: (1) 종의 구분이 가능한 수준의 사진 자료가 함께 제공될 때, (2) 시민이 해당 생물군의 종 구분이 가능한 수준의 지식을 가지고 있을 때, (3) 제보자가 관측한 생물군의 종 구분을 위한 상세한 묘사가 함께 제공될 때. Kim et al. (2018)의 경우, 소셜 미디어 정보를 활용하여 제주도 해역의 남방큰돌고래 (*Tursiops aduncus*)의 분포를 간접적으로 추정하고자 하였으며, 이와 같이 우리나라도 체계적인 시민 참여형 관측 시스템을 구축한다면 보다 많은 관측 자료를 확보할 수 있을 것으로 기대된다.

##### 4.2 환경변수를 이용한 종 분포 추정 모델의 활용도 증진

보다 신뢰할 수 있고 상대적으로 불확실성이 작은 해양 생물 종 분포 추정 모델 개발을 위하여 고도화된

중 분포 추정 기법 활용해야 한다. 기존 국내 중 분포 추정 연구의 경우, 육상 분야에서는 통계, 기계학습 및 앙상블을 활용한 연구가 진행되고 있으나(Chung et al., 2020), 해양 분야에는 많이 진행되지 않았다.

국외에서는 해양포유류 및 바닷새 분포 추정을 위하여 다양한 통계, 기계학습 기법 및 앙상블을 활용한 연구가 활발히 진행되고 있는 반면(Fig. 5), 우리나라의 경우, 그렇지 못한 실정이다. Lee et al. (2017)은 수산과학원에서 관측한 밍크고래(*Balaenoptera acutorostrata*) 개체 수 자료(2000-2012; 총 195건)와 표층 클로로필 농도와 일차생산성에 대한 Moderate Resolution Imaging Spectroradiometer (MODIS) 위성의 산출물의 상관관계를 활용하여 우리나라 동해 지역 밍크고래 분포를 간접적으로 추정하였다. 해당 연구에서 개체수와 환경 변수가  $R^2$  0.9 이상의 높은 상관관계를 보이지만, 같은 밍크고래 분포를 추정하기 위하여 클로로필 농도 뿐만 아니라, 해수면 온도, 바람, 해류, 다양한 지형학, 생물학적 환경 변수를 고려하여 중 분포 추정 모델을 도출한 선행연구(Roberts et al., 2016; Tobena et al., 2016)와 비교하였을 때, 상대적으로 불확실성이 크다. 아직까지 우리나라가 확보하고 있는 해양포유류 및 바닷새에 대한 샘플 수가 적어 통계학적인 모델링에 한계점이 있을 수 있지만, 기계학습 및 앙상블 기법을 활용한다면 보다 고도화된 중 분포 추정 모형의 개발이 가능 할 것으로 보인다.

#### 4.3 해양포유류 및 바닷새의 특성을 고려한 환경 변수 개발

마지막으로, 우리나라 해역이용에 따른 환경 및 생물 영향평가를 위하여 해양포유류 및 바닷새의 특성을 고려할 수 있는 환경변수에 대한 추가적인 연구가 필요하다.

앞서 소개하였듯, 우리나라에서는 해양포유류 19종과 바닷새 16종을 포함하여 총 88종의 해양보호생물을 선정하여 보호활동을 장려하고 있다. 하지만, 해상 풍력단지 개발 등과 같은 해역이용에 따른 생물종에 대한 영향과 관련한 연구가 많이 부족한 실정이다(Oh et al., 2020). 특히, 해역이용에 따른 수중소음의 발생이 상괭이, 남방큰돌고래를 비롯한 우리나라 주요 해양포유류에 미치는 영향 및 이에 따른 수중소음 영향 기준에 대한 기준이 정립되지 않은 실정이며(Sohn et al., 2015), 우리나라에 서식하거나 우리나라 해역을 거쳐 이동하는 바닷새의 이동패턴(migration pattern)과 기초적인 공간분포에 대한 분석 또한 제대로 이루어지지 않은 실정이다(Oh et al., 2020). 반면 국외에서는 수중소음에 따른 해양생태계 영향에 대한 평가

(Bailey et al., 2010; Southall et al., 2008) 및 바닷새와 해상풍력건설에 따른 바닷새의 민감도 평가(Best and Halpin, 2019; Bradbury et al., 2014)가 적극적으로 이루어지고 있다.

이와 같은 해역이용영향평가에 대한 기술 개발 및 보안을 통하여 해양생물의 중 분포 추정 모델을 고도화할 수 있을 것으로 기대되며, 더 나아가 추후 우리나라 해역이용에 따른 생물 영향평가에 중요한 지표로도 활용할 수 있을 것으로 기대된다.

## 5. 요약 및 결론

본 연구에서는 지난 10년(2011-2020)간 국내/외 해양 생물에 대한 중 분포 추정 모델을 활용한 연구의 동향 및 주요 키워드를 분석하고, 최근 5년간(2016-2020) 게재된 해양포유류 및 바닷새에 대한 분포 추정 연구의 연구 대상 및 분석 지역, 중 분포 추정 기법에 대한 현황을 분석하였다. 분석 결과, 국외에서는 다양한 중 분포 추정 기법을 활용하여 다양한 해양 생물 종에 대한 연구를 진행하고 있는 반면, 우리나라 해역에 대한 해양포유류 및 바닷새를 비롯한 중 분포 추정 연구가 많이 미흡한 것으로 나타났다. 본 연구에서는 연구 동향 분석 결과를 통하여 우리나라 해역에 대한 해양 생물 관측자료의 보완, 체계적인 중 분포 추정 모델의 활용, 해양포유류 및 바닷새의 특성을 고려할 수 있는 환경 변수의 보완/개발의 필요성을 제시하였다. 본 연구에서 제시한 논의사항을 통해 추후 우리나라 해역에 대한 생물 중 분포 연구의 발전과 해역이용에 따른 환경영향평가 기법의 개발 및 보완이 기대된다.

## 감사의 글

본 연구는 2021년 해양수산부 재원으로 해양수산과학기술진흥원(과학기술기반 해역이용영향평가 기술개발, 20210427)의 지원을 받아 수행하였습니다.

## 참고문헌

1. Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling*, 157(2-3), 101-118.
2. Awad, M., & Khanna, R. (2015). Support vector regression. In *Efficient learning machines* (pp. 67-80). Apress, Berkeley, CA.



3. Ayyadevara, V. K. (2018). Gradient boosting machine. In *Pro machine learning algorithms* (pp. 117-134). Apress, Berkeley, CA.
4. Azzolin, M., Arcangeli, A., Cipriano, G., Crosti, R., Maglietta, R., Pietroluongo, G., ... & Carlucci, R. (2020). Spatial distribution modelling of striped dolphin (*Stenella coeruleoalba*) at different geographical scales within the EU Adriatic and Ionian Sea Region, central-eastern Mediterranean Sea. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30(6), 1194-1207.
5. Bailey, H., Senior, B., Simmons, D., Rusin, J., Picken, G., & Thompson, P. M. (2010). Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Marine Pollution Bulletin*, 60(6), 888-897.
6. Baines, M., & Weir, C. R. (2020). Predicting suitable coastal habitat for sea whales, southern right whales and dolphins around the Falkland Islands. *Plos One*, 15(12), e0244068.
7. Barragán-Barrera, D. C., Do Amaral, K. B., Chávez-Carreño, P. A., Fariás-Curtidor, N., Lancheros-Neva, R., Botero-Acosta, N., ... & Palacios, D. M. (2019). Ecological niche modeling of three species of *Stenella* dolphins in the Caribbean Basin, with application to the Seaflower Biosphere Reserve. *Frontiers in Marine Science*, 6, 10.
8. Becker, E. A., Carretta, J. V., Forney, K. A., Barlow, J., Brodie, S., Hoopes, R., ... & Hazen, E. L. (2020). Performance evaluation of cetacean species distribution models developed using generalized additive models and boosted regression trees. *Ecology and Evolution*, 10(12), 5759-5784.
9. Becker, E. A., Forney, K. A., Redfern, J. V., Barlow, J., Jacox, M. G., Roberts, J. J., & Palacios, D. M. (2019). Predicting cetacean abundance and distribution in a changing climate. *Diversity and Distributions*, 25(4), 626-643.
10. Becker, E. A., Forney, K. A., Thayre, B. J., Debich, A. J., Campbell, G. S., Whitaker, K., ... & Hildebrand, J. A. (2017). Habitat-based density models for three cetacean species off Southern California illustrate pronounced seasonal differences. *Frontiers in Marine Science*, 4, 121.
11. Bedriñana-Romano, L., Huckle-Gaete, R., Viddi, F. A., Morales, J., Williams, R., Ashe, E., ... & Ruiz, J. (2018). Integrating multiple data sources for assessing blue whale abundance and distribution in Chilean Northern Patagonia. *Diversity and Distributions*, 24(7), 991-1004.
12. Best, B. D., & Halpin, P. N. (2019). Minimizing wildlife impacts for offshore wind energy development: Winning tradeoffs for seabirds in space and cetaceans in time. *Plos one*, 14(5), e0215722.
13. Bonizzoni, S., Furey, N. B., Santostasi, N. L., Eddy, L., Valavanis, V. D., & Bearzi, G. (2019). Modelling dolphin distribution within an Important Marine Mammal Area in Greece to support spatial management planning. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(10), 1665-1680.
14. Bradbury, G., Trinder, M., Furness, B., Banks, A. N., Caldow, R. W., & Hume, D. (2014). Mapping seabird sensitivity to offshore wind farms. *Plos One*, 9(9), e106366.
15. Breen, P., Brown, S., Reid, D., & Rogan, E. (2016). Modelling cetacean distribution and mapping overlap with fisheries in the northeast Atlantic. *Ocean & Coastal Management*, 134, 140-149.
16. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
17. Cañadas, A., De Soto, N. A., Aissi, M., Arcangeli, A., Azzolin, M., B-Nagy, A., ... & Roger, T. (2018). The challenge of habitat modelling for threatened low density species using heterogeneous data: the case of Cuvier's beaked whales in the Mediterranean. *Ecological Indicators*, 85, 128-136.
18. Carlucci, R., Cipriano, G., Paoli, C., Ricci, P., Fanizza, C., Capezzuto, F., & Vassallo, P. (2018). Random Forest population modelling of striped and common-bottlenose dolphins in the Gulf of Taranto (Northern Ionian Sea, Central-eastern Mediterranean Sea). *Estuarine, Coastal and Shelf Science*, 204, 177-192.
19. Carlucci, R., Fanizza, C., Cipriano, G., Paoli, C., Russo, T., & Vassallo, P. (2016). Modeling the spatial distribution of the striped dolphin (*Stenella coeruleoalba*) and common bottlenose dolphin

- (*Tursiops truncatus*) in the Gulf of Taranto (Northern Ionian Sea, Central-eastern Mediterranean Sea). *Ecological indicators*, 69, 707-721.
20. Chung, H. I., Choi, Y., Ryu, J., & Jeon, S. W. (2020). Accuracy Evaluation of Potential Habitat Distribution in *Pinus thunbergii* using a Species Distribution Model: Verification of the Ensemble Methodology. *The Korean Society of Climate Change Research*, 11(1), 37-51.
  21. Cleasby, I. R., Owen, E., Wilson, L., Wakefield, E. D., O'Connell, P., & Bolton, M. (2020). Identifying important at-sea areas for seabirds using species distribution models and hotspot mapping. *Biological Conservation*, 241, 108375.
  22. Correia, A. M., Gil, Á., Valente, R., Rosso, M., Pierce, G. J., & Sousa-Pinto, I. (2019). Distribution and habitat modelling of common dolphins (*Delphinus delphis*) in the eastern North Atlantic. *Journal of the Marine Biological Association of the United Kingdom*, 99(6), 1443-1457.
  23. Cursach, J. A., Arriagada, A., Rau, J. R., Ojeda, J., Bizama, G., & Becerra, A. (2019). Predicting the potential distribution of the endemic seabird *Pelecanus thagus* in the Humboldt Current Large Marine Ecosystem under different climate change scenarios. *PeerJ*, 7, e7642.
  24. de Rivera, O. R., Blangiardo, M., López-Quílez, A., & Martín-Sanz, I. (2019). Species distribution modelling through Bayesian hierarchical approach. *Theoretical Ecology*, 12(1), 49-59.
  25. Derville, S., Torres, L. G., Iovan, C., & Garrigue, C. (2018). Finding the right fit: Comparative cetacean distribution models using multiple data sources and statistical approaches. *Diversity and Distributions*, 24(11), 1657-1673.
  26. Dwyer, S. L., Pawley, M. D., Clement, D. M., & Stockin, K. A. (2020). Modelling habitat use suggests static spatial exclusion zones are a non-optimal management tool for a highly mobile marine mammal. *Marine Biology*, 167(5), 1-20.
  27. Fernandez, M., Yesson, C., Gannier, A., Miller, P. I., & Azevedo, J. M. N. (2018). A matter of timing: how temporal scale selection influences cetacean ecological niche modelling. *Marine Ecology Progress Series*, 595, 217-231.
  28. Fiedler, P. C., Redfern, J. V., Forney, K. A., Palacios, D. M., Sheredy, C., Rasmussen, K., ... & Ballance, L. T. (2018). Prediction of large whale distributions: a comparison of presence-absence and presence-only modeling techniques. *Frontiers in Marine Science*, 5, 419.
  29. Fox, C. H., Huettmann, F. H., Harvey, G. K. A., Morgan, K. H., Robinson, J., Williams, R., & Paquet, P. C. (2017). Predictions from machine learning ensembles: marine bird distribution and density on Canada's Pacific coast. *Marine Ecology Progress Series*, 566, 199-216.
  30. Friedman, J. (1991). Multivariate adaptive regression splines (with discussion), *Annals of Statistics*, 19, 1-141.
  31. Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction*. Springer open.
  32. Getis, A., & Ord, J. K. (2010). The analysis of spatial association by use of distance statistics. In *Perspectives on spatial data analysis* (pp. 127-145). Springer, Berlin, Heidelberg.
  33. Giannoulaki, M., Markoglou, E., Valavanis, V. D., Alexiadou, P., Cucknell, A., & Frantzis, A. (2017). Linking small pelagic fish and cetacean distribution to model suitable habitat for coastal dolphin species, *Delphinus delphis* and *Tursiops truncatus*, in the Greek Seas (Eastern Mediterranean). *Aquatic Conservation: Marine and Freshwater Ecosystems*, 27(2), 436-451.
  34. Goh, A. T. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143-151.
  35. Gomez, C., Lawson, J., Kouwenberg, A. L., Moors-Murphy, H., Buren, A., Fuentes-Yaco, C., ... & Wimmer, T. (2017). Predicted distribution of whales at risk: identifying priority areas to enhance cetacean monitoring in the Northwest Atlantic Ocean. *Endangered Species Research*, 32, 437-458.
  36. Goyert, H. F., Gardner, B., Sollmann, R., Veit, R. R., Gilbert, A. T., Connelly, E. E., & Williams, K. A. (2016). Predicting the offshore distribution and abundance of marine birds with a hierarchical community distance sampling model. *Ecological Applications*, 26(6), 1797-1815.

37. Grüss, A., Drexler, M. D., Ainsworth, C. H., Roberts, J. J., Carmichael, R. H., Putman, N. F., ... & Love, M. S. (2018). Improving the spatial allocation of marine mammal and sea turtle biomasses in spatially explicit ecosystem models. *Marine Ecology Progress Series*, 602, 255-274.
38. Hallgren, W., Santana, F., Low-Choy, S., Zhao, Y., & Mackey, B. (2019). Species distribution models can be highly sensitive to algorithm configuration. *Ecological Modelling*, 408, 108719.
39. Harvey, G. K., Nelson, T. A., Fox, C. H., & Paquet, P. C. (2017). Quantifying marine mammal hotspots in British Columbia, Canada. *Ecosphere*, 8(7), e01884.
40. Hastie, T. J., & Tibshirani, R. J. (2017). *Generalized additive models*. Routledge.
41. Hastie, T., Tibshirani, R., & Buja, A. (1994). Flexible discriminant analysis by optimal scoring. *Journal of the American Statistical Association*, 89(428), 1255-1270.
42. Hemami, M. R., Ahmadi, M., Sadegh-Saba, M., & Moosavi, S. M. H. (2018). Population estimate and distribution pattern of Indian Ocean humpback dolphin (*Sousa plumbea*) in an industrialised bay, northwestern Persian Gulf. *Ecological Indicators*, 89, 631-638.
43. Herr, H., Kelly, N., Dorschel, B., Huntemann, M., Kock, K. H., Lehnert, L. S., ... & Scheidat, M. (2019). Aerial surveys for Antarctic minke whales (*Balaenoptera bonaerensis*) reveal sea ice dependent distribution patterns. *Ecology and evolution*, 9(10), 5664-5682.
44. Hunt, T. N., Allen, S. J., Bejder, L., & Parra, G. J. (2020). Identifying priority habitat for conservation and management of Australian humpback dolphins within a marine protected area. *Scientific Reports*, 10(1), 1-14.
45. Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44.
46. Kanaji, Y., & Gerrodette, T. (2020). Estimating abundance of Risso's dolphins using a hierarchical Bayesian habitat model: A framework for monitoring stocks of animals inhabiting a dynamic ocean environment. *Deep Sea Research Part II: Topical Studies in Oceanography*, 175, 104699.
47. Kim, H. W., Kim, Y. H., An, Y., Park, K. J., & An D. (2014). Seabird Distribution Patterns by Strip Transect in the Yellow Sea in Spring. *Korean Journal of Fisheries and Aquatic Sciences*, 47(6), 973-977.
48. Kim, H. W., Lee, D., & Sohn, H. (2018). Distribution Status of Indo-Pacific Bottlenose Dolphin *Tursiops aduncus* in the Jeju Island Based on Social Media Data. *Korean Journal of Fisheries and Aquatic Sciences*, 51(5), 600-605.
49. Kim, J., Lim, S. Y., & Yoo, S. H. (2017). Measuring the economic benefits of designating Baegnyeong Island in Korea as a marine protected area. *International Journal of Sustainable Development & World Ecology*, 24(3), 205-213.
50. Koo, K. A., Park, S. U., & Seo, C. (2017). Effects of climate change on the climatic niches of warm-adapted evergreen plants: expansion or contraction?. *Forests*, 8(12), 500.
51. Kwon, H. S. (2014, December 31). Applying Ensemble Model for Identifying Uncertainty in the Species Distribution Models. *Journal of Korean Society for Geospatial Information System*. The Korean Society for Geospatial Information Systems. <https://doi.org/10.7319/kogsis.2014.22.4.047>
52. La Manna, G., Ronchetti, F., Sarà, G., Ruiu, A., & Ceccherelli, G. (2020). Common Bottlenose Dolphin Protection and Sustainable Boating: Species Distribution Modeling for Effective Coastal Planning. *Frontiers in Marine Science*, 7, 955.
53. Lee, D. K., & Kim, H. G. (2010). Habitat potential evaluation using Maxent model-Focused on riparian distance, stream order and land use. *Journal of the Korean Society of Environmental Restoration Technology*, 13(6), 161-172.
54. Lee, D., An, Y. R., Park, K. J., Kim, H. W., Lee, D., Joo, H. T., ... & Lee, S. H. (2017). Spatial distribution of common Minke whale (*Balaenoptera acutorostrata*) as an indication of a biological hotspot in the East Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 143, 91-99.
55. Lee, J., & Seong, S. (2016). An Investigation into Supply Characteristics and Spatial Clustering Pattern of Office Buildings in Seoul : Major

- Office Buildings between 2003 and 2012. *Journal of Korea Planning Association*-Vol, 51(3), 83-96.
56. Lee, S., Choi, S., Kim, J. H., Kim, H. W., & Sohn, H. (2018). Characteristics of the Cetacean Bycatch in Korean Coastal Waters from 2011 to 2017. *Korean Journal of Fisheries and Aquatic Sciences*, 51(6), 704-713.
  57. Llapapasca, M. A., Pacheco, A. S., Fiedler, P., Goya, E., Ledesma, J., Peña, C., & Vásquez, L. (2018). Modeling the potential habitats of dusky, commons and bottlenose dolphins in the Humboldt Current System off Peru: The influence of non-El Niño vs. El Niño 1997-98 conditions and potential prey availability. *Progress in Oceanography*, 168, 169-181.
  58. Loh, W. Y. (2011). Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1), 14-23.
  59. Lombal, A. J., O'dwyer, J. E., Friesen, V., Woehler, E. J., & Burridge, C. P. (2020). Identifying mechanisms of genetic differentiation among populations in vagile species: historical factors dominate genetic differentiation in seabirds. *Biological Reviews*, 95(3), 625-651.
  60. Mäkinen, J., & Vanhatalo, J. (2018). Hierarchical Bayesian model reveals the distributional shifts of Arctic marine mammals. *Diversity and Distributions*, 24(10), 1381-1394.
  61. Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15(1), 59-69.
  62. Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4), 603-614.
  63. Moghim, S., & Bras, R. L. (2017). Bias correction of climate modeled temperature and precipitation using artificial neural networks. *Journal of Hydrometeorology*, 18(7), 1867-1884.
  64. Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3), 370-384.
  65. Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.
  65. Oh, H. T. and Yeo, M. Y. (2019). Status and improvement of marine environmental impact assessment for offshore windfarm project, *Journal of The Korean Society Fisheries And Sciences Education*, 31(5), 1470-1481.
  67. Oh, H. T., Yeo, M. Y., Jung, H. E., & Shim, J. M. (2020). Status and Improvement of Environmental Impacts Assessment on the Marine Endangered Species around the Coastal Area of Offshore Wind Energy - Case Study of the Marine Mammals and Sea Birds -. *Journal of The Korean Society Fisheries And Sciences Education*, 32(6), 1428-1444.
  68. Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27(4), 286-306.
  69. Ouyang, Z., & Liu, J. (2008). Application of ecological-niche factor analysis in habitat assessment of giant pandas. *Acta Ecologica Sinica*, 28(2).
  70. Pace, D. S., Arcangeli, A., Mussi, B., Vivaldi, C., Ledon, C., Lagorio, S., ... & Ardizzone, G. (2018). Habitat suitability modeling in different sperm whale social groups. *The Journal of Wildlife Management*, 82(5), 1062-1073.
  71. Panigada, S., Lauriano, G., Donovan, G., Pierantonio, N., Cañadas, A., Vázquez, J. A., & Burt, L. (2017). Estimating cetacean density and abundance in the Central and Western Mediterranean Sea through aerial surveys: implications for management. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141, 41-58.
  72. Paradell, O. G., López, B. D., & Methion, S. (2019). Modelling common dolphin (*Delphinus delphis*) coastal distribution and habitat use: Insights for conservation. *Ocean & Coastal Management*, 179, 104836.
  73. Passadore, C., Möller, L. M., Diaz-Aguirre, F., & Parra, G. J. (2018). Modelling dolphin distribution to inform future spatial conservation decisions in a marine protected area. *Scientific Reports*, 8(1), 1-14.
  74. Pendleton, D. E., Holmes, E. E., Redfern, J., &

- Zhang, J. (2020). Using modelled prey to predict the distribution of a highly mobile marine mammal. *Diversity and Distributions*, 26(11), 1612-1626.
75. Pennino, M. G., Mérigot, B., Fonseca, V. P., Monni, V., & Rotta, A. (2017). Habitat modeling for cetacean management: Spatial distribution in the southern Pelagos Sanctuary (Mediterranean Sea). *Deep Sea Research Part II: Topical Studies in Oceanography*, 141, 203-211.
76. Pereira, J. M., Krüger, L., Oliveira, N., Meirinho, A., Silva, A., Ramos, J. A., & Paiva, V. H. (2018). Using a multi-model ensemble forecasting approach to identify key marine protected areas for seabirds in the Portuguese coast. *Ocean & Coastal Management*, 153, 98-107.
77. Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.
78. Phillips, S. J., Dudík, M., & Schapire, R. E. (2004, July). A maximum entropy approach to species distribution modeling. In *Proceedings of the twenty-first international conference on Machine learning* (p. 83).
79. Pitchford, J. L., Howard, V. A., Shelley, J. K., Serafin, B. J., Coleman, A. T., & Solangi, M. (2016). Predictive spatial modelling of seasonal bottlenose dolphin (*Tursiops truncatus*) distributions in the Mississippi Sound. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 26(2), 289-306.
80. Purdon, J., Shabangu, F. W., Yemane, D., Pienaar, M., Somers, M. J., & Findlay, K. (2020). Species distribution modelling of Bryde's whales, humpback whales, southern right whales, and sperm whales in the southern African region to inform their conservation in expanding economies. *PeerJ*, 8, e9997.
81. Redfern, J. V., Moore, T. J., Fiedler, P. C., de Vos, A., Brownell Jr, R. L., Forney, K. A., ... & Ballance, L. T. (2017). Predicting cetacean distributions in data-poor marine ecosystems. *Diversity and Distributions*, 23(4), 394-408.
82. Roberts, J. J., Best, B. D., Mannocci, L., Fujioka, E. I., Halpin, P. N., Palka, D. L., ... & Lockhart, G. G. (2016). Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. *Scientific Reports*, 6(1), 1-12.
83. Rockwood, R. C., Elliott, M. L., Saenz, B., Nur, N., & Jahncke, J. (2020). Modeling predator and prey hotspots: Management implications of baleen whale co-occurrence with krill in Central California. *Plos One*, 15(7), e0235603.
84. Santora, J. A., Eisner, L. B., Kuletz, K. J., Ladd, C., Renner, M., & Hunt Jr, G. L. (2018). Biogeography of seabirds within a high-latitude ecosystem: Use of a data-assimilative ocean model to assess impacts of mesoscale oceanography. *Journal of Marine Systems*, 178, 38-51.
85. Silber, G. K., Lettrich, M. D., Thomas, P. O., Baker, J. D., Baumgartner, M., Becker, E. A., ... & Waples, R. S. (2017). Projecting marine mammal distribution in a changing climate. *Frontiers in Marine Science*, 4, 413.
86. Sohn HS, An DH and Kim HW(2015). A Study on the Legal Frame to Manage Anthropogenic Underwater Noise for Marine Mammal Protection in Korean Waters. *Korean Journal of Maritime Policy*, 30(2): 165~188.
87. Sohn, H., Park, K. J., An, Y. R., Choi, S. G., Kim, Z. G., Kim, H. W., An, D. H., Lee, Y. R., & Park, T. (2012). Distribution of Whales and Dolphins in Korean Waters Based on a Sighting Survey from 2000 to 2010. *Korean Journal of Fisheries and Aquatic Sciences*, 45(5), 486-492.
88. Song, H., Ding, Z., Guo, C., Li, Z., & Xia, H. (2008, December). Research on combination kernel function of support vector machine. In *2008 International conference on computer science and software engineering* (Vol. 1, pp. 838-841). IEEE.
89. Southall, B. L., Bowles, A. E., Ellison, W. T., Finneran, J. J., Gentry, R. L., Greene Jr, C. R., ... & Tyack, P. L. (2008). Marine mammal noise-exposure criteria: initial scientific recommendations. *Bioacoustics*, 17(1-3), 273-275.
90. Stephenson, F., Goetz, K., Sharp, B. R., Mouton, T. L., Beets, F. L., Roberts, J., ... & Lundquist, C. J. (2020). Modelling the spatial distribution of cetaceans in New Zealand waters. *Diversity and Distributions*, 26(4), 495-516.
91. Studwell, A. J., Hines, E., Elliott, M. L., Howar,

- J., Holzman, B., Nur, N., & Jahncke, J. (2017). Modeling nonresident seabird foraging distributions to inform ocean zoning in Central California. *Plos One*, 12(1), e0169517.
92. Suthaharan, S. (2016). Support vector machine. In *Machine learning models and algorithms for big data classification* (pp. 207-235). Springer, Boston, MA.
93. Tardin, R. H., Chun, Y., Simão, S. M., & Alves, M. A. S. (2017). Modeling habitat use by Bryde's whale *Balaenoptera edeni* off southeastern Brazil. *Marine Ecology Progress Series*, 576, 89-103.
94. Tardin, R. H., Maciel, I. S., Espécie, M. A., Melo-Santos, G., Simao, S. M., & Alves, M. A. S. (2020). Modelling habitat use by the Guiana dolphin, *Sotalia guianensis*, in south-eastern Brazil: Effects of environmental and anthropogenic variables, and the adequacy of current management measures. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30(4), 775-786.
95. Thorne, L. H., Baird, R. W., Webster, D. L., Stepanuk, J. E., & Read, A. J. (2019). Predicting fisheries bycatch: A case study and field test for pilot whales in a pelagic longline fishery. *Diversity and Distributions*, 25(6), 909-923.
96. Tobeña, M., Prieto, R., Machete, M., & Silva, M. A. (2016). Modeling the potential distribution and richness of cetaceans in the Azores from fisheries observer program data. *Frontiers in Marine Science*, 3, 202.
97. Vassallo, P., Marini, C., Paoli, C., Bellingeri, M., Dhermain, F., Nuti, S., ... & Gnone, G. (2020). Species-specific distribution model may be not enough: The case study of bottlenose dolphin (*Tursiops truncatus*) habitat distribution in Pelagos Sanctuary. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30(8), 1689-1701.
98. Viddi, F. A., Harcourt, R. G., & Huccke-Gaete, R. (2016). Identifying key habitats for the conservation of Chilean dolphins in the fjords of southern Chile. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 26(3), 506-516.
99. Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., ... & Ridoux, V. (2019). Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: Large-scale modelling of deep-diving cetacean habitats. *Global Ecology and Biogeography*, 28(3), 300-314.
100. Virgili, A., Lambert, C., Pettex, E., Dorémus, G., Van Canneyt, O., & Ridoux, V. (2017). Predicting seasonal variations in coastal seabird habitats in the English Channel and the Bay of Biscay. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141, 212-223.
101. Wakefield, E. D., Owen, E., Baer, J., Carroll, M. J., Daunt, F., Dodd, S. G., ... & Bolton, M. (2017). Breeding density, fine-scale tracking, and large-scale modeling reveal the regional distribution of four seabird species. *Ecological Applications*, 27(7), 2074-2091.
102. Warwick-Evans, V. C., Atkinson, P. W., Robinson, L. A., & Green, J. A. (2016). Predictive modelling to identify near-shore, fine-scale seabird distributions during the breeding season. *Plos One*, 11(3), e0150592.
103. Wiest, W. A., Correll, M. D., Marcot, B. G., Olsen, B. J., Elphick, C. S., Hodgman, T. P., ... & Shriver, W. G. (2019). Estimates of tidal-marsh bird densities using Bayesian networks. *The Journal of Wildlife Management*, 83(1), 109-120.
104. Zanardo, N., Parra, G. J., Passadore, C., & Möller, L. M. (2017). Ensemble modelling of southern Australian bottlenose dolphin *Tursiops* sp. distribution reveals important habitats and their potential ecological function. *Marine Ecology Progress Series*, 569, 253-266.
105. Zerbini, A. N., Friday, N. A., Palacios, D. M., Waite, J. M., Ressler, P. H., Rone, B. K., ... & Clapham, P. J. (2016). Baleen whale abundance and distribution in relation to environmental variables and prey density in the Eastern Bering Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 312-330.