

# 지구관측데이터 공동활용을 위한 K-GEOSS 시범서비스 개발

정창후\*

\*한국과학기술정보연구원

e-mail:chjeong@kisti.re.kr

## Development of a K-GEOSS Pilot Service for Sharing and Utilizing Global Earth Observation Data

Chang-Hoo Jeong\*

\*Korea Institute of Science and Technology Information

### 요 약

지구관측데이터 공동활용체제로 GEOSS가 제안되면서 한국에서도 K-GEOSS 개발을 통하여 다양한 방식으로 지구관측데이터의 활용·확산에 참여하고 있다. 본 연구에서는 K-GEOSS 시범서비스로 환경 분야의 생물 중 분포 변화 예측 시나리오를 개발하여 미래 생물 종의 서식지 보호를 위한 지구관측데이터의 효과적 사용 사례를 보여준다. 이러한 시범서비스는 K-DMSS 플랫폼을 이용하여 개발하였으며, 데이터 수집·전처리, 모델 학습·평가, 실제 분포 변화 예측까지 모두 자동화하여 수행한다. 최종단계에서 다양한 기후모델 및 기후변화 시나리오에 따른 여러 종류의 예측 결과를 제공함으로써 대체 서식지 보호 및 정책수립을 위한 종합적인 의사결정에 도움을 주도록 하였다. 이와 같은 다양한 종류의 시범서비스 개발 및 지구관측데이터의 공동활용 사례 발굴을 통하여 GEOSS 체계의 정착을 보다 앞당길 수 있다.

### 1. 서론

조정적·포괄적·지속적 지구관측으로 인류에 도움이 되는 결정·행동 실현의 비전을 가지고 지구관측데이터의 공동활용을 추진하는 GEO (Group on Earth Observations)[1]가 창설되고, 비전달성을 위한 통합시스템으로 GEOSS (Global Earth Observation System of Systems)[2]가 제안되면서 참여국들은 GEOSS 통합 시스템을 개발하기 위한 다양한 노력을 경주하고 있다. 이에 보조를 맞추기 위해서 한국에서도 GEOSS로서의 역할을 수행하는 K-GEOSS를 개발하여 지구관측데이터의 활용·확산에 다양한 방식으로 참여하고 있다.

본 연구에서는 지구관측데이터 공동활용을 위한 K-GEOSS 시범서비스 사례로서 환경 분야의 대규모 생물 중 분포 변화 예측 시나리오를 개발하였다. 환경 분야의 생태계 보존을 위해서 많은 연구들이 수행되고 있는데, 그 중에서 미래 기후변화로 인한 생물 종의 분포 변화를 예측하고 보호구역을 보전하는 시나리오는 GEO의 사회편익분야 중 하나인 생태계 분야에서 전 지구 관측 데이터의 효과적 사용을 보여주는 적절한 사례라고 할 수 있다.

### 2. K-GEOSS 시범서비스 개발

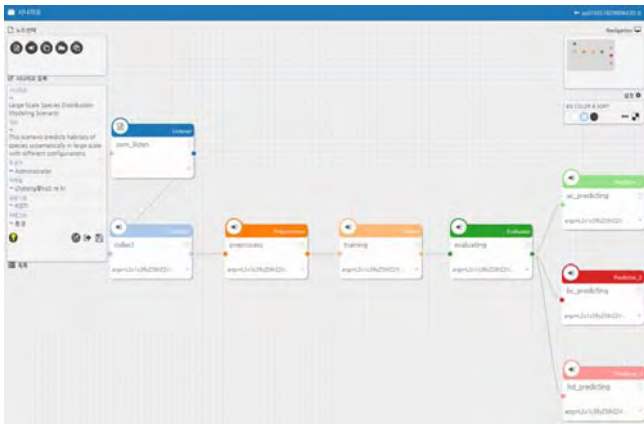
일반적으로 생물 중 분포 모델링에 관련된 연구들은 개인의 연구자가 수작업으로 데이터를 수집하고 분석하는 과정을 거쳐서 최종 적합 서식지를 도출하게 된다. 그러나

K-GEOSS 시범서비스 시나리오에서는 데이터 수집, 전처리, 모델 수행에 이르는 전 과정이 HPC & Big Data 플랫폼 위에서 자동화되어 수행된다. 이러한 자동화된 시나리오 운영은 입력으로 사용되는 생물의 종류가 바뀌거나 기후변화 시나리오의 개정으로 인해 미래 기후변화 데이터의 새로운 버전이 공개되더라도 시스템의 전체 프로세스에 영향을 미치지 않고 적합 서식지를 신속하게 재분석할 수 있는 장점을 가지고 있다.

대규모의 생물 중 분포 변화를 예측하는 K-GEOSS 시범서비스는 K-DMSS(KISTI-Decision Making Support System) 플랫폼을 통하여 이루어진다. K-DMSS는 시나리오 기반으로 각종 데이터를 수집하여 미래 예측 정보를 생산하고, 사회·경제적 영향을 분석하여 과학적인 방법으로 표출할 수 있는 통합 소프트웨어 체계이다[3]. K-DMSS는 고성능컴퓨팅 인프라, 고해상도 모델링·시뮬레이션 기술, 코드 최적·병렬화 기술, 빅데이터 분석 기술, 대용량 데이터 가시화 기술 등을 기반으로 개발되었다[3]. K-DMSS를 활용한 K-GEOSS 시범서비스 시나리오는 그림 1과 같은 워크플로우 형태로 구성되어 운영된다.

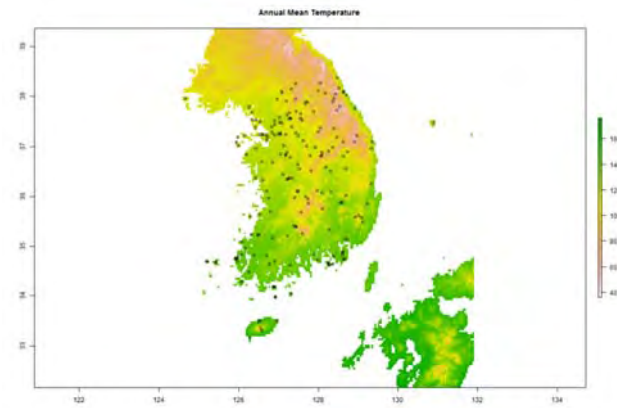
K-DMSS를 구성하는 웹 포털의 워크플로우 설계 화면은 ① 시나리오에 대한 메타데이터 입력 화면(그림 1의 좌측)과 ② 작업 목적(데이터 수집, 데이터 전처리, 모델 수행, 결과 시각화, 결과 후처리 등)에 적합한 노드를 선택하여 작업 흐름을 제어하는 워크플로우 구성 화면(그림 1의 가운데), 그리고 ③ 확인하고자 하는 위치를 선택하여

상세 정보 화면으로 바로 이동할 수 있도록 도와주는 네비게이션 화면(그림 1의 우측 상단)으로 구성되어 있다.



(그림 1) 생물 종 분포 변화 예측 시나리오 설계 화면

그림 1의 시나리오 설계 화면에서 보이는 것과 같이 생물 종 분포 변화 예측 시나리오는 1) Collector: 데이터 수집 단계(생물 종 데이터는 국립생물자원관으로부터, 기후 데이터는 WorldClim으로부터 수집), 2) Preprocessor: 데이터 전처리 단계, 3) Trainer: 모델 학습 단계, 4) Evaluator: 모델 평가 단계, 5) Predictor: 분포 변화 예측 단계로 구성된다. 각 단계가 순차적으로 실행되면서, 최종적으로 GCM (Global Climate Model)과 RCP (Representative Concentration Pathways)에 따른 여러 종류의 예측 결과를 생산한다.



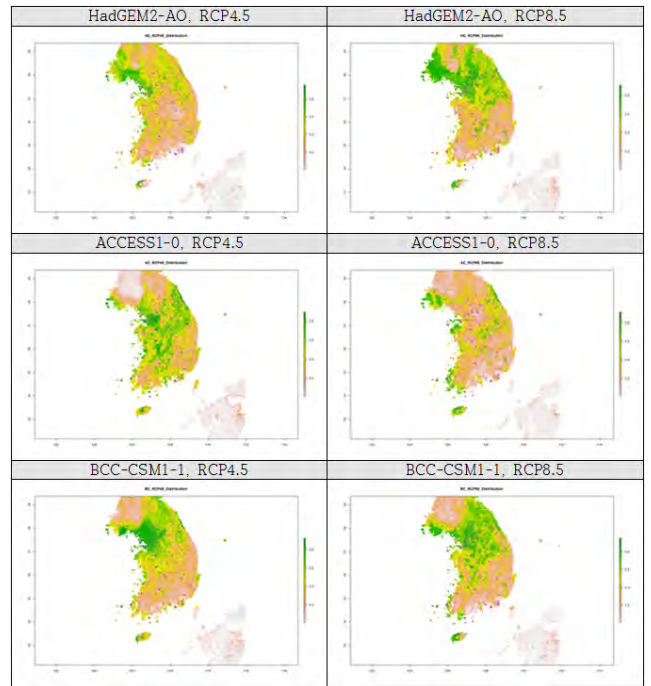
(그림 2) 산왕거미 분포도 및 한국의 연간 평균 기온

그림 2는 생물 종 분포 변화 예측을 위해서 본 시나리오의 사례 연구로 사용한 산왕거미(Araneus ventricosus)의 현재 분포와 산왕거미가 서식하고 있는 대한민국의 연간 평균 기온 정보를 보여준다.

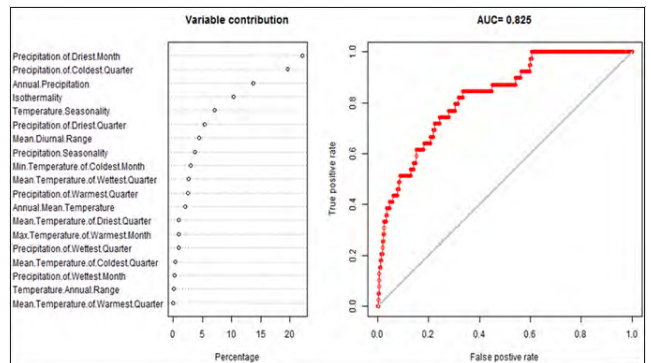
그림 3은 다양한 미래 조건 하에서 산왕거미가 어디에 분포할 것인지를 예측한 결과를 보여준다. 여기에서 예측된 결과를 바탕으로 해당 종의 대체 서식지를 마련하거나 보존하는 정책을 수립할 수 있다.

그림 4는 산왕거미의 미래 분포를 예측한 모델의 변수 중요도와 유효성을 보여준다. 건조한 달의 강수량이 중요한 변수로 작용하고 있다는 것과 80% 이상의 AUC 값을

가지고 모델 예측이 유효하게 동작하고 있다는 것을 확인할 수 있다.



(그림 3) GCM과 RCP에 따른 산왕거미 분포 예측 결과



(그림 4) 변수 중요도와 모델 유효성

### 3. 결론 및 고찰

본 연구에서는 지구관측데이터 공동활용을 위한 K-GEOS 시범서비스로서 생물 종의 미래 서식지 분포 변화를 예측하고 대체 서식지를 보존하는 시나리오에 대해서 설명하였다. 이러한 시나리오는 K-DMSS라는 HPC & Big Data 플랫폼 상에서 자동화되어 수행되기 때문에 실용성 및 확장성 측면에서 많은 장점을 가지고 있다. 향후 이러한 활용사례를 더욱 발굴하고 고도화함으로써 GEOS 체계의 정착을 보다 앞당길 수 있다.

### 참고문헌

[1] GEO, <http://www.earthobservations.org/index2.php>  
 [2] GEOS, <https://www.earthobservations.org/geos.php>  
 [3] 조민수, "KISTI 재난대응 의사결정지원시스템(K-DMSS) 소개", 기상기술정책지, Vol. 10, No. 1, pp.56-70, 2017.