



[특별세션 14]

극지 GeoAI 데이터 활용 연구

- 한국극지데이터 관리현황
- 극지 빅데이터 플랫폼 활용 사례 조사: 대용량 유전체 데이터 분석
- 극지 위성자료와 빙권정보 수집 및 모니터링 시스템을 통한 GeoAI 활용
- 인공지능을 활용한 얇은 해빙 두께 추정 연구
- K-means clustering 기법을 활용한 해수면온도 자료 기반 북극해 광역 분류 연구

한국극지데이터 관리현황

김고흥^{1,*}, 주동찬¹, 이주한^{2,#}

¹극지연구소 미래기술센터 데이터연구지원실

²극지연구소 미래기술센터

#교신저자: joochan@kopri.re.kr

극지연구소는 남·북극에서 과학 연구를 수행하고 있으며, 다양한 유형의 데이터를 수집하고 있습니다. 극지연구소는 남·북극 지방의 연구 데이터를 수집·관리·공유하기 위해서 한국 극지데이터 센터(Korea Polar Data Center)를 운영하고 있습니다. KPDC 에서는 데이터 관리 시스템 구축과 운영을 하고 있습니다. KPDC 는 극지정보의 효율적인 관리와 유통을 통해 연구자들이 극지정보에 쉽게 접근할 수 있도록 포괄적이고 다양한 극지정보와 데이터를 제공하기 위해 시스템의 관리를 위한 체계적인 도구를 지속적으로 업그레이드하는 것을 목표로 운영하고 있습니다.

극지데이터 관리시스템은 데이터 관리 계획(Data Management Plan), 데이터 수집 및 표준 검사, 데이터 보존, 지도 정보 표출, 데이터의 공개 등의 서비스를 하고 있으며, 극지연구소 큐레이션 시스템과의 데이터 연계 및 외부의 AMD(Antarctic Master Directory), KISTI DataON 시스템, SIOS(Svalbard Integrated Arctic Earth Observing System) 과 연계를 하고 있으며, ArcticPortal.org 와 MOU 체결을 통한 외부기관과의 협력을 모색 중에 있습니다.

극지 빅데이터 플랫폼 활용 사례 조사: 대용량 유전체 데이터 분석

박용길^{1,*}, 이철용^{1,#}, 이주한², 주동찬³

¹한국해양과학기술원 해양빅데이터·AI 센터

²극지연구소 미래기술센터

³극지연구소 데이터연구지원실

#교신저자: cylee82@kiost.ac.kr

극지의 미생물 분석은 극한 환경에서의 생태계와 미생물의 환경 적응 기작, 그리고 유용한 생물소재의 발견과 생분해 과정의 이해 등을 연구하기 위한 중요한 과제이다. 이러한 미생물 군집의 전체 유전체 정보를 분석하면 기존에 알지 못했던 새로운 미생물 종을 발견할 수 있다. 그러나 이러한 분석은 대용량 데이터와 복잡한 분석 과정, 그리고 데이터 구조의 복잡성으로 인해 고성능의 컴퓨팅 자원이 요구된다. 따라서 본 연구에서는 극지 빅데이터 플랫폼을 이용한 미생물 분석 사례를 조사하여 시사점을 도출하도록 하였다.

균유전체학 기반 미생물 분석 방법은 특정 환경에서 채취한 모든 미생물 군집의 유전체를 분석하는 방법으로 이를 통해 미생물 다양성을 탐구하고 생태계 기능을 분석할 수 있으며, 건강과 의학연구 및 산업적 응용에도 활용될 수 있다. 기존의 분석 방법 중 Megahit 는 대규모의 데이터셋을 처리할 수 있고 메모리 사용량을 최소화하는 등의 장점을 가지고 있다. 이는 MPI 방식을 이용하여 클러스터 환경에서 실행될 수 있다. 한편, Apache Spark 를 이용한 균유전체학 분석 시도도 이루어지고 있다. Hadoop 과 Spark 기반의 균유전체학 분석 알고리즘 연구와 함께 클라우드 기반의 분석 플랫폼도 등장하고 있다. 마지막으로, 두 방식을 혼합하여 Spark 가 설치된 서버를 이용하여 Megahit 를 실행하는 방안도 고려할 수 있다.

사례 분석 결과, 균유전체학 분석을 위해서는 데이터 어셈블리 작업에 최적화된 효율성을 우선한다면 Megahit 를 MPI 와 함께 사용하는 것이 더 적합하다는 결론을 도출하였다. 유연한 데이터 처리 및 분석 워크플로우가 필요하고, 대규모 데이터 처리 및 분석 작업을 통합적으로 관리하기 위해서는 Spark 기반의 균유전체학 분석을 활용할 수 있다.

사사

이 논문은 2024 년도 해양수산부 재원으로 해양수산과학기술진흥원의 지원을 받아 수행된 연구임(20210629, 극한지 개발 및 탐사용 협동이동체 시스템 기술 개발 사업)

참고문헌

1. Wall, D.P., Kudtarkar, P., Fusaro, V.A. *et al.* Cloud computing for comparative genomics. *BMC Bioinformatics* 11, 259 (2010). doi: 10.1186/1471-2105-11-259.
2. Wiewiórka MS, Messina A, Pacholewska A, Maffioletti S, Gawrysiak P, Okoniewski MJ. SparkSeq: fast, scalable and cloud-ready tool for the interactive genomic data analysis with nucleotide precision. *Bioinformatics*. 30(18):2652-3 (2014). doi: 10.1093/bioinformatics/btu343.

극지 위성자료와 빙권정보 수집 및 모니터링 시스템을 통한 GeoAI 활용

이성재^{1,*}, 김현철^{1,#}

¹극지연구소 원격탐사빙권정보센터

#교신저자: kimhc@kopri.re.kr

수많은 위성이 지구를 관측하게 되면서 위성정보를 어떻게 효율적으로 활용것인지에 대한 중요성이 부각되고 있다. 특히 접근의 제약이 따르는 극지방의 환경, 지질, 해양 등 빙권 연구를 위해서는 위성자료의 활용이 매우 중요한 역할을 하고 있다. 이러한 자료의 체계적인 관리를 위해서 극지연구소 원격탐사빙권정보센터에서는 극지연구에 필요한 위성 데이터의 수집 및 DB 화를 통하여 GeoAI 연구에 활용하고자 하였다. 또한 다양한 연구를 통해 생성된 산출물은 'STAR Archive System'을 통하여 공유될 예정이다. 해당 시스템은 웹기반으로 개발되었으며, 주요 기능으로는 3 차원 공간정보 상에서 위성자료 또는 빙권정보를 가시화 및 메타데이터를 확인할 수 있다. 기존에는 위성을 운영하는 기관에 따라 각각 제공되었다면, 본 연구에서 구축한 시스템은 하나의 사이트에서 극지연구에 필요한 자료를 효율적으로 검색하고 활용할 수 있다. 한국항공우주연구원(KARI)의 아리랑위성을 포함한 Sentinel(ESA), Suomi-NPP(NASA), Landsat(USGS) 등 다양한 극지연구에 활용할 수 있는 자료를 구축하고 있다. 향후에는 극지연구과제를 통해 개발된 산출물인 해빙두께, 해빙농도, 해빙이동장 등 해빙정보를 올려 여러 연구자와 공유할 예정이다. 또한 북극과 남극 해양연구활동을 책임지고 있는 쇄빙연구선 아라온호의 운항 지원을 위한 위성자료 및 해빙정보를 자동으로 제공하여 항해에 필요한 항로결정의 기초자료로 활용할 수 있도록 지원하고 있다. 극지 원격탐사 연구에 다양한 역할을 하고 있는 STAR Archive System 을 통하여 연구자의 극지 위성정보 활용을 촉진시키고, 다양한 활용분야 개척에 도움이 될 것으로 기대하고 있다.

사사

본 연구는 한국해양과학기술원 부설 극지연구소의 쇄빙연구선 위성 지원(PE24825) 및 북극 빙권변화 정량분석을 위한 원격탐사연구(PE24040) 연구과제의 지원을 받아 수행되었습니다.

인공지능을 활용한 얇은 해빙 두께 추정 연구

박진구^{1,*}, 김정훈², 김현철^{1,#}

¹극지연구소 원격탐사빙권정보센터

²극지연구소 생명과학연구본부

#교신저자: kimhc@kopri.re.kr

남극의 로스해는 전 지구 열염 순환의 핵심 요소인 남극심층수(Anantarctic Bottom Water) 형성에 중요한 역할을 한다. 남극심층수의 약 25%는 로스해에서 기여하며, 폴리냐(polynya)에서의 해빙 형성이 중요한 메커니즘이다. 기후 변화에 따른 남극심층수의 변화를 파악하기 위해서는 장기적인 해빙생산력 자료를 구축하는 것이 필수적이며, 이를 위해 위성 자료를 활용한 연구가 가장 유효한 방법이다. 일반적으로 위성 기반 해빙생산력을 추정하기 위해서는 얇은 해빙 두께 (thin ice thickness)가 필요하다. 현재 대기의 영향을 받지 않고 지속적인 얇은 해빙 두께 자료를 구축하기 위해서는 수동마이크로파(passive microwave) 자료가 사용되는 것이 일반적이며, 이러한 추정을 위해 열적외 기반 해빙 두께와 수동마이크로파의 편광비 간의 관계를 통해 추정될 수 있다. 하지만 열적외 위성 기반의 얇은 해빙 두께 추정치는 대기의 영향을 받기 때문에 시공간적으로 연속적인 자료 구축이 힘들며, 기존 관계식들은 두 자료 간의 비선형적인 관계를 반영하지 않으며 최대 0.2m 까지의 얇은 해빙 두께만 산출할 수 있다. 이 밖에도 내재되어 있는 다양한 문제를 해결하기 위해, 본 연구는 인공지능 기반의 엔드 투 엔드 접근법을 통해 얇은 해빙 두께를 보다 정확하게 추정하고자 한다. 이를 통해 수동 마이크로파 자료와 열 유속 모델을 통합하여 해빙 생산의 장기 시계열을 구축할 수 있으며, 궁극적으로 폴리냐 역학과 남극저층수에 대한 최근 변동성과 추세를 파악할 수 있다. 본 연구는 해빙 분포와 전 지구 해양학적 변화에 대한 중요한 통찰을 제공하며, 특히 남극 연안 폴리냐에서의 고염분 대륙붕수 형성과 이에 따른 전 지구 열염 순환에 미치는 영향을 이해하는 데 기여할 것이다.

사사

이 연구는 해양수산부의 재원으로 해양수산과학기술진흥원의 지원을 받아 수행되었습니다(과제번호 : KIMST 20220547).

K-means clustering 기법을 활용한 해수면온도 자료 기반 북극해 광역 분류 연구

윤주은^{1,*}, 김현철^{1,#}

¹극지연구소 원격탐사빙권정보센터

#교신저자: kimhc@kopri.re.kr

북극해는 해빙, 육지로부터 유입되는 담수, 태평양과 대서양에서 유입되는 해수 등 여러 환경 요인들이 복잡하게 얽혀 작용한다. 이에 따라 최근 심각해지는 지구온난화 현상에 대한 북극 해양 환경의 반응을 이해하기 위해서는 해양의 물리·생지화학적 특성을 기반으로 북극해를 광역 분류하고, 각 광역별 시·공간적 변동 특성을 파악하는 것이 필수적으로 요구된다. 본 연구는 대기-해양의 상호작용, 해양 환경의 변화 특성을 대표할 수 있는 인자인 해수면온도(Sea Surface Temperature)자료를 수집하여 분석에 활용하였다. 사용된 해수면온도자료는 1982년부터 2023년까지의 NOAA OISST v2.1 (Optimum Interpolation Sea Surface Temperature) 자료로 AVHRR 위성 관측 자료와 현장관측자료(부이, 아고, 선박)를 0.25°×0.25° 격자 간격으로 일일 합성한 자료이다. 해수면온도자료는 NSIDC(National Snow & Ice Data Center)에서 제공하는 마스크(Mask) 정보를 활용하여 육상의 영향을 제거한 후 개방구역(Open Water)자료만 추출하여 사용하였다. 북극해에서 대표적으로 나타나는 해수면온도의 시·공간적 변동성을 기반으로 북극해 광역을 분류하기 위해 획득된 해수면온도 자료는 통계함수인 경험적 직교함수(Empirical Orthogonal Function, EOF)법에 적용하여 주성분을 분석하였다. 본 연구에서는 경험적으로 누적 기여율이 85% 이상이 되는 주성분값에 대해 K-평균(K-means Clustering) 알고리즘을 적용하여 북극해 광역을 분류하고, 각 광역별 해수면온도의 시·공간적 변동패턴을 비교·분석하여 북극해의 광역별 지구온난화의 영향을 이해하고자 한다.

사사

북극 빙권변화 정량 분석을 위한 원격탐사 연구(PE24040152)