

Research Article

농림위성을 위한 기계학습을 활용한 복사전달모델기반 대기보정 모사 알고리즘 개발 및 검증: 식생 지역을 위주로

강유진 ¹⁾ · 김예진 ²⁾ · 임정호 ^{3)*} · 임중빈 ⁴⁾

Machine Learning-Based Atmospheric Correction Based on Radiative Transfer Modeling Using Sentinel-2 MSI Data and Its Validation Focusing on Forest

Yoojin Kang ¹⁾ · Yejin Kim ²⁾ · Jungho Im ^{3)*} · Joongbin Lim ⁴⁾

Abstract: Compact Advanced Satellite 500-4 (CAS500-4) is scheduled to be launched to collect high spatial resolution data focusing on vegetation applications. To achieve this goal, accurate surface reflectance retrieval through atmospheric correction is crucial. Therefore, a machine learning-based atmospheric correction algorithm was developed to simulate atmospheric correction from a radiative transfer model using Sentinel-2 data that have similar spectral characteristics as CAS500-4. The algorithm was then evaluated mainly for forest areas. Utilizing the atmospheric correction parameters extracted from Sentinel-2 and GEOKOMPSAT-2A (GK-2A), the atmospheric correction algorithm was developed based on Random Forest and Light Gradient Boosting Machine (LGBM). Between the two machine learning techniques, LGBM performed better when considering both accuracy and efficiency. Except for one station, the results had a correlation coefficient of more than 0.91 and well-reflected temporal variations of the Normalized Difference Vegetation Index (i.e., vegetation phenology). GK-2A provides Aerosol Optical Depth (AOD) and water vapor, which are essential parameters for atmospheric correction, but additional processing should be required in the future to mitigate the problem caused by their many missing values. This study provided the basis for the atmospheric correction of CAS500-4 by developing a machine learning-based atmospheric correction simulation algorithm.

Keywords: Compact Advanced Satellite 500-4, Light gradient boosting machine, Sentinel-2, GEOKOMPSAT-2A

Received October 3, 2023; Revised October 16, 2023; Accepted October 20, 2023; Published online October 31, 2023

¹⁾ 울산과학기술원 지구환경도시건설공학과 박사후연구원(Postdoctoral Researcher, Department of Civil Urban Earth and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

²⁾ 울산과학기술원 지구환경도시건설공학과 석/박사통합과정생(Combined MS/PhD Student, Department of Civil Urban Earth and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

³⁾ 울산과학기술원 지구환경도시건설공학과 교수(Professor, Department of Civil Urban Earth and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

⁴⁾ 국립산림과학원 국제산림연구과 임업연구소(Researcher, Division of Global Forestry, National Institute of Forest Science, Seoul, Republic of Korea)

* Corresponding author: Jungho Im (ersgis@unist.ac.kr)

Copyright © 2023 by The Korean Society of Remote Sensing. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

요약: 식생에서의 고해상도 정보를 제공하기 위한 목적으로 농림위성이 발사 예정에 있다. 농림위성 자료를 식생 분야에 활용하기 위해서는 정확한 지표면 반사도 추출이 필수적이므로, 이를 위한 대기보정 기술의 개발과 식생에 대한 검증이 선행되어야 한다. 따라서, 본 연구에서는 Sentinel-2를 대체 위성 영상으로 활용하여 복사전달모델 대기보정을 모사하는 기계학습 기반 대기보정 알고리즘을 개발하고, 산림 지역을 위주로 검증을 수행하였다. Sentinel-2 및 GEOKOMPSAT-2A (GK-2A)에서 추출된 대기보정 파라미터를 기반으로 Random Forest와 Light Gradient Boosting Machine (LGBM)을 통하여 대기보정 알고리즘을 개발하고, 산림에 위치한 총 4개 지역에서의 1년 간의 현장 관측 자료를 이용하여 검증하였다. 두 가지 기계학습 기법 중 효율성을 고려했을 때 LGBM이 탁월할 것으로 판단되었으며, 한 관측소를 제외하면 0.91 이상의 상관계수를 보이며 Normalized Difference Vegetation Index를 활용한 연간 식생 활력도의 시계열 변동을 잘 반영할 수 있음을 확인하였다. 대기보정 필수 파라미터인 Aerosol Optical Depth (AOD) 및 water vapor 수급을 위하여 GK-2A 산출물을 활용할 수 있으나, AOD의 지속적인 결측 문제가 지표면 반사도 산출에 치명적일 수 있기 때문에 이를 보완할 필요가 있다고 사료된다. 본 연구는 농림위성을 위한 대기보정 모사 알고리즘을 개발하고, 한계점과 동시에 보완되어야 할 방향에 대해 제시함으로써 추후 농림위성 자료의 정밀 대기보정을 위한 토대로 기여할 수 있을 것으로 기대된다.

주요어: 차세대 중형위성 4호, Light gradient boosting machine, Sentinel-2, 천리안위성 2A호

1. 서론

농지 및 산림에서의 원격 탐사를 위해서 농림위성(차세대 중형위성 4호, CAS500-4)이 발사 예정에 있다. 가시광(Blue, Green, Red), 적색경계(Red edge), 근적외(Near infrared)를 포함한 5개 파장대에 대해 5 m 공간 해상도로 정보를 제공할 수 있어 식생에 대한 고해상도 정보를 제공하는 데 큰 기여를 할 것으로 기대된다(Kwon et al., 2021; Kim et al., 2021). 이는 이상기상 혹은 자연 재해 등에 취약한 농업 및 산림 분야에서의 관측에 있어 아주 유용하다. 위성 영상으로부터 지상에서의 정보를 얻기 위해서는 대기 보정을 통하여 대기에 의한 효과를 제거한 지표 반사도를 취득해야 한다(Zhang et al., 2018). 특히나 식생 지역에서는 여러 파장대에서 취득된 반사도 간의 비율 등을 이용하여 계산된 식생 지수(Normalized Difference Vegetation Index, Normalized Difference Water Index 등)를 통한 식생 관찰이 이루어지고 있으며, 아주 작은 반사도의 차이에도 영향을 크게 받기 때문에 정확한 대기 보정 기술을 적용하는 것이 매우 중요하다(Kang et al., 2021).

대기 보정은 위성 영상에서 관측된 대기 상한 복사휘도 혹은 반사도로부터 복사전달모델을 이용하여 대기 효과를 제거하여 지표에서의 신호를 얻는 과정을 의미한다(Vermote and Kotchenova, 2008). 복사전달모델은 위

성 천정각 및 태양 천정각, 대기 중 에어로졸, 수증기량, 전층 오존 농도 등의 조건을 기반으로 대기 중의 투과, 산란, 흡수 등 다양한 요인을 고려한다. 따라서, 방대한 변수를 고려해 주기 위해서 엄청난 계산량과 시간이 소요된다는 한계점이 있다. 일반적으로 이러한 과정을 간소화하기 위해서 미리 다양한 대기 조건에 따른 복사전달모델 결과를 계산하여 조건표를 구축하고, 이를 위성 관측값에 적용하여 대기보정을 수행한다(Anderson et al., 1999). 이 때, 위성의 스펙이나 대기 보정을 수행하고자 하는 지역의 대기 조건, 대기 보정 파라미터의 민감도 등에 따라 적절한 파라미터 범위 및 간격에 대해서 표를 구축하는 것이 매우 중요하다(Lee et al., 2015). 예를 들어, 대기 보정 파라미터 중 가장 주요한 요소 중 하나인 에어로졸 광학 두께 (Aerosol Optical Depth, AOD)는 매우 촘촘한 간격에 대한 조건표 구축이 요구된다.

뿐만 아니라, 수많은 경우에 대해서 조건표를 구축하더라도 실제 대기에서의 대기 보정 파라미터는 표에 구축되지 않은 값을 가지는 경우가 존재한다. 이는 대기 보정 결과에 불연속성을 야기하며, 지표면 반사도에도 영향을 준다. 따라서, 조건표를 구축함과 동시에 이를 내삽할 수 있는 방안을 마련하는 것은 필수적이다. 데이터 내삽을 위한 다차원 보간 방법으로는 minimum curvature surface, radial basis function 등의 방법이 존재하지만 대기 보정에 요구되는 파라미터의 수가 많고 그

사이에 비선형적인 관계를 감안했을 때, 기계학습을 이용하면 내삽 및 모사를 동시에 잘 수행할 것으로 기대된다(Lee et al., 2020). 기계학습은 수많은 독립 변수와 종속 변수 사이의 비선형적 관계를 풀어내는 데 뛰어난 성능을 보여 원격 탐사 분야에서 널리 활용되는 추세이다(Ghimire et al., 2012; Kim et al., 2014; Mudiyansele et al., 2022). 따라서 본 연구에서는 기계학습을 통한 복사전달모델 기반 대기 보정 모사 알고리즘을 개발하고 적용 가능성을 평가하였다.

기계학습 기반 대기보정 알고리즘에 의해 산출된 결과의 신뢰성을 평가하기 위해, 현장 관측 데이터와의 비교검증은 불가피하다. 대기의 다양한 변수들에 의해 대기 보정 결과가 크게 달라질 수 있으므로, 현장 측정값과의 상관 관계, 편차 등을 꼼꼼히 분석하여 알고리즘의 유효성을 평가해야 한다. 산림 분야에서는 Normalized Difference Vegetation Index (NDVI)와 같은 식생 지수가 산림의 건강성, 수종 분류, 성장 상태 등을 평가하는 중요한 지표로 사용된다(Pettorelli et al., 2005). 반사도는 이런 지수를 산출하는데 있어 기본적인 요소로, 그 정확성은 결과의 신뢰도에 큰 영향을 미친다. 계절별 혹은 다양한 시기에 걸친 현장 관측 데이터와의 지속적인 비

교검증은 알고리즘의 신뢰성을 향상시키는 핵심적인 요소이다(Lim et al., 2022).

본 연구에서는 Second Simulation of the Satellite Signal in the Solar Spectrum Vector (6SV) 복사전달모델 기반으로 농림위성의 상세규격과 우리 나라 대기 상황을 고려하여 조건표를 구축하였다. 또한, 시간 및 계산 간소화를 위하여 이를 내삽 및 모사할 수 있는 기계학습 기반 대기보정 모사 알고리즘을 개발하였다. 본 연구에서 개발한 알고리즘은 현재 개발 중인 농림위성과 유사한 분광 해상도의 정보를 제공하는 Sentinel-2를 대체 자료로 활용하였으며, 산림에서의 현장 관측 기반 연간 지표면 반사도 및 식생 지수의 변화 관찰을 통한 비교평가를 수행하였다.

2. 연구 지역 및 자료

2.1. Sentinel-2 위성 자료

Sentinel-2는 유럽 우주국(European Space Agency)에서 운영 중인 위성으로 다중분광센서인 Multi-Spectral Instrument (MSI)를 장착하고 있다(Drusch et al., 2012). Sentinel-2는

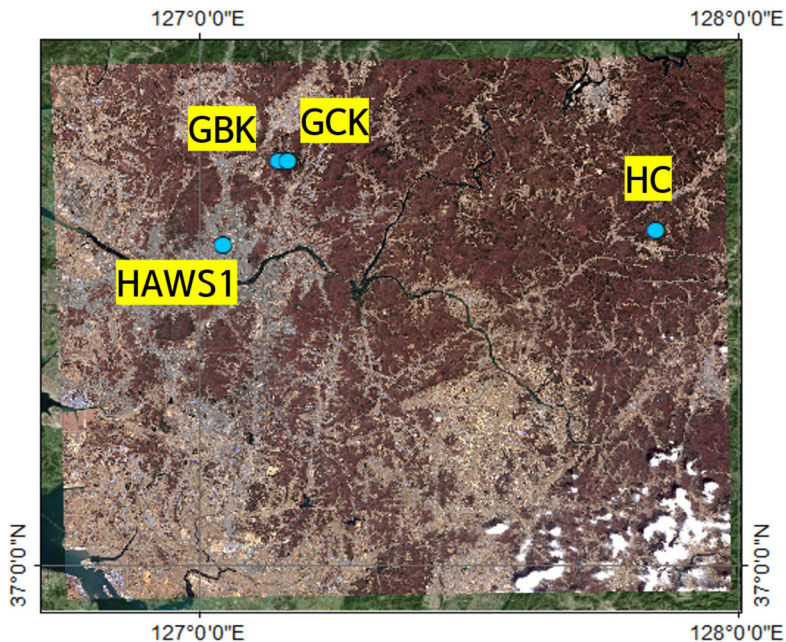


Fig. 1. Sentinel-2 image covering the study area on April 7, 2022. The blue circles represent four in-situ stations that measure bottom of atmosphere reflectance (GBK, GCK, HC, HAWS1).

추후 농림위성에서 제공하게 될 파장 영역과 매우 유사한 파장 정보를 제공하고 있기 때문에 농림위성을 대체할 위성 영상으로 선정하였다. MSI는 3개의 가시광 및 근적외선 영역에 대해서 10 m 공간 해상도로 정보를 제공하고 있으며, 대기 보정이 수행되기 전의 자료는 Level-1C (L1C), Sentinel-2 자체 알고리즘인 Sen2cor 기반으로 대기 보정이 수행된 자료가 Level-2A (L2A)로 정의되어 있다. 개발된 대기 보정 알고리즘의 검증을 위하여 본 연구에서는 L1C 자료를 모델링에 이용하고 그 결과를 L2A와 비교하였다. 본 연구는 산림 현장 관측을 수행하고 있는 4개 플럭스 타워(flux tower)를 포함하는 T52SCG 타일에 대해 수행하였다(Fig. 1).

2.2. 대기보정 파라미터 자료

대기 보정은 대기 중의 정보를 파라미터로 입력 받아 이루어지므로 이를 확보하는 것이 필수적이다. 그 중 가장 주요한 파라미터로는 대기 중의 에어로졸에 관한 정보인 에어로졸 광학 두께(AOD), 수증기량(Water Vapor, WV), 오존(Ozone, O₃), 고도 등이 있다. AOD는 대기 중 산란 효과를, WV는 대기 중 흡수 효과 보정에 사용되는 파라미터로서 매우 중요한 역할을 한다(Kang et al., 2021). Sentinel-2 MSI는 자체적으로 단파적외 파장대를 이용하는 dense dark vegetation method에 기반하여 AOD를 산출하고, atmospheric pre-corrected differential absorption method에 기반하여 WV를 산출한 뒤 대기 보정을 진행한다(Kaufman et al., 1997a; Kaufman et al., 1997b; Schläpfer et al., 1998). 그러나 농림위성은 AOD 및 WV 산출에 필요한 파장 정보를 포함하지 않기 때문에 추후 대기 보정을 위해서는 해당 파라미터를 외부에서 제공할 필요가 있다. 이미지 내에 0의 반사도를 가지는 지점이 존재한다고 가정하고 대기보정을 수행하는 dark object subtraction 방법도 적용 가능하다(Chavez Jr, 1988). 그러나 이는 가정에 적합한 대상이 선정되었는가에 대기보정 결과가 의존하는 한계점이 있으며, dark object subtraction보다 복사전달모델에 의한 대기보정 결과가 더 정확하다는 선행 연구 결과도 존재한다(Lantzanakis et al., 2017). 따라서 본 연구에서는 해당 파라미터를 외부에서 제공하는 방안을 채택하였다.

본 연구에서는 AOD와 WV 파라미터 확보를 위해서 2018년에 발사된 GEO-KOMPSAT-2A (GK-2A) 위성 산

출물을 활용하였다. GK-2A 위성은 AOD는 2분 간격, WV는 10분 간격으로 2 km 공간 해상도로 제공하고 있으므로 대기 보정에 필요한 필수 파라미터를 실시간으로 확보할 수 있다(Jee et al., 2020). 본 연구에서는 Sentinel-2 위성 관측 시간 직전의 30분 동안의 자료를 평균하여 사용하였다.

대기 보정 파라미터 요소의 하나인 고도 자료는 Shuttle Radar Topography Mission (SRTM)에서 제공되는 Digital Elevation Mode (DEM) 자료를 이용하였다. SRTM DEM은 90m 해상도로 제공되며, 유라시아 대륙에서의 오차는 2.54m 수준인 것으로 보고되었다 (Berry et al., 2007). 본 연구에서는 Sentinel-2 MSI 타일 내에 포함된 고도의 평균 값을 대기 보정 파라미터로 활용하였다.

2.3. 현장 관측 자료

대한민국의 지리적 특성을 고려할 때, 국토의 63% 이상을 차지하는 산림 영역의 관측 및 연구는 필수적이다 (Lim et al., 2022). 산림 관측에 필요한 데이터는 국립산림과학원의 협조를 통해 획득하였으며, 주요 관측 데이터는 플럭스 타워를 통한 연속적인 모니터링 결과이다.

플럭스 타워에 설치된 분광센서는 (주)솔단의 SD-500 제품으로, 해당 센서의 기술적 사양은 Table 1에 상세히 기술되어 있으며, 분광 관측 범위는 Table 2에 명시되어 있다. SD-500의 사양을 Sentinel-2 MSI 및 농림위성 Compact Advanced Payload with Wide swath-width (CAP-W) 센서 사양과 비교했을 때, blue 밴드에서의 중심 파장에는 일부 차이가 있으나 그 외의 밴드들은 상당히 유사한 파장대를 공유하고 있었다.

본 연구에서는 산림 영역의 검증에 중점을 두고 T52SCG 타일에 위치한 4개 관측 지점에 대한 심층 분석을 수행하였다. 이 네 지점은 광릉 활엽수림 플럭스

Table 1. Specifications of the Soldan SD-500 sensor

Detector Element	Light Emitting Diode, Photodiode
Detector List	Red, Green, Blue, Near Infrared
Operating Temperature Range	-35°C ~ 50°C
Operating Humidity Range	0-100% RH
Cosine Response	±5% (Zenith angle within 70°)
Field of view angle	180°

Table 2. Central wavelength and bandwidth of SD-500, Sentinel-2 MSI, and CAS500-4 CAP-W sensors

Band	SD-500		Sentinel-2 MSI		CAS500-4 CAP-W	
	Central Wavelength (nm)	Bandwidth (nm)	Central Wavelength (nm)	Bandwidth (nm)	Central Wavelength (nm)	Bandwidth (nm)
Blue	412	59	490	65	490	65
Green	560	42	560	35	560	35
Red	656	44	665	30	665	30
Red-edge	-	-	705, 740, 783	20, 20, 10	705	15
Near infrared	850	100	842	115	842	115

Table 3. Height above canopy and selected days in the study period for each in-situ station

In-situ measurement	GBK	GCK	HC	HAWS1
Height above canopy	20 m	5 m	10 m	5 m
Date	2022.01.07	2022.01.07	2022.01.07	2022.01.07
	2022.01.27	2022.01.12	2022.01.27	2022.01.17
	2022.02.06	2022.01.17	2022.02.06	2022.02.06
	2022.02.11	2022.01.27	2022.02.11	2022.02.11
	2022.02.16	2022.02.06	2022.02.21	2022.02.16
	2022.03.03	2022.02.11	2022.03.03	2022.02.21
	2022.03.08	2022.02.16	2022.03.08	2022.03.03
	2022.03.28	2022.03.03	2022.04.02	2022.03.08
	2022.04.02	2022.03.08	2022.04.07	2022.03.28
	2022.04.07	2022.03.28	2022.04.17	2022.04.02
	2022.04.17	2022.04.02	2022.04.27	2022.04.07
	2022.04.27	2022.04.07	2022.05.02	2022.04.17
	2022.05.02	2022.04.17	2022.05.12	2022.04.27
	2022.05.07	2022.04.27	2022.05.17	2022.05.02
	2022.05.12	2022.05.07	2022.06.01	2022.05.07
	2022.05.17	2022.05.12	2022.06.11	2022.05.12
	2022.05.27	2022.05.17	2022.09.24	2022.05.17
	2022.06.01	2022.05.27	2022.09.29	2022.05.22
	2022.06.11	2022.06.01	2022.10.19	2022.05.27
	2022.09.19	2022.06.11	2022.12.23	2022.06.01
	2022.09.24	2022.09.19	2022.12.28	2022.06.11
	2022.10.19	2022.09.24	-	2022.06.21
	2022.10.29	2022.10.19	-	2022.07.01
	2022.12.08	2022.10.29	-	2022.09.09
	2022.12.13	2022.12.08	-	2022.09.19
	2022.12.18	2022.12.18	-	2022.09.24
	2022.12.28	2022.12.28	-	2022.10.19
-	-	-	2022.10.29	
-	-	-	2022.12.08	
-	-	-	2022.12.13	

타워(GBK), 광릉 침엽수림 플렉스 타워(GCK), 홍천 플렉스 타워(HC) 그리고 국립산림과학원 플렉스 타워(HAWS1)로 구성되어 있다. 2022년 동안의 Sentinel-2 영

상 자료 중 구름의 영향을 크게 받지 않는 일자를 추출하여, 네 관측 지점에 대해 총 105일 동안의 샘플 데이터를 구축하였다(Table 3). 지역별로 분류하면 GBK 및

GCK 지역에서는 각각 27개의 샘플, HC에서는 21개의 샘플, 그리고 HAWS1에서는 30개의 샘플을 확보할 수 있었다.

3. 연구 방법

3.1. 복사전달모델 기반 대기보정 및 조건표 구축

대기 보정은 위성에서 관측된 값에서부터 대기 중의 산란, 흡수 등에 의한 효과를 제거하여 지표 반사도를 계산할 수 있는 복사전달모델을 통해 이루어진다. 대기 상한 복사 휘도 값이 주어졌을 때 각각 투과, 산란, 반사에 대한 항인 대기 보정 계수 x_a, x_b, x_c 를 이용하면 아래 식을 통하여 지표면 반사도로 변환할 수 있다(식 1). 그러나 매 이미지에 대해 복사전달모델을 실행하는 것은 방대한 시간과 계산량을 필요로 하므로 주로 사전에 구축된 조건표를 활용하여 보정 작업을 수행한다. 조건표는 대기 보정 파라미터의 효과를 미리 계산하고 정리한 것으로, 복사전달모델을 실행하지 않고도 즉시 대기 보정을 수행할 수 있는 장점이 있다(Vermote and Kotchenova, 2008). 본 연구에서는 6SV 복사전달모델을 사용하였다. 6SV는 6S 복사전달모델에서 편광 효과를 계산할 수 있도록 개선된 모델로서 successive orders of scattering 방법을 통하여 복사전달 방정식을 풀어가며, 0.3-4 μm 파장 범위를 2.5 nm 분광해상도 단위로 구분하여 모의할 수 있다(Vermote et al., 1997). 본 연구에서는 Python에서 6SV를 구현할 수 있도록 구축된 Py6S 라이브러리를 활용하였다(Wilson, 2013).

$$BOA\ reflectance = \frac{x_a \times TOA\ radiance - x_b}{1 + x_c \times (x_a \times TOA\ radiance - x_b)} \quad (1)$$

식(1)에서 Bottom of Atmosphere (BOA) reflectance는 지표면 반사도를, Top of Atmosphere (TOA) radiance는 대기상한 복사휘도를 의미한다. x_a (the inverse of the transmittance)는 대기 중의 투과에 관한 항으로, 대기보정 파라미터 중에서는 AOD, SZA, VZA, WV에 의한 영향을 크게 받는 항이다. x_b (the scattering term of the atmosphere)는 산란에 관한 항으로, x_a 와 마찬가지로 AOD, SZA, VZA에 영향을 많이 받는다. x_c (spherical albedo, atmospheric reflectance for isotropic light)는 AOD

Table 4. Atmospheric parameter range, interval, and number of cases used for the look-up-table

Paramete	Range	Interval	No. of case
SZA	0–70°	5°	15
VZA	0–30°	10°	4
RAA	0–180°	20°	10
WV	0–4 g/cm ²	1 g/cm ²	5
O3	300–350 DU	50 DU	2
AOD	0, 0.02, 0.04, 0.06, 0.08, 0.1, 0.12, 0.14, 0.16, 0.18, 0.20, 0.25, 0.3, 0.35, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0		24
DEM	0–1800 m	300 m	7

에 의한 영향력만을 주로 받는 것으로 알려져 있다(Lee et al., 2020).

본 연구에서는 태양 천정각(solar zenith angle, SZA), 위성 천정각(view zenith angle, VZA), 상대 방위각(relative azimuth angle, RAA), WV, O₃, AOD, DEM을 포함한 7개 대기보정 파라미터에 대해서 대기 보정 계수인 x_a, x_b, x_c 를 복사전달모델을 통하여 산출하고, 이를 조건표로 구축하였다(Table 4). Py6S에 적용되는 aerosol type은 본 연구의 대상 지역이 육상에 존재하는 식생 지역임을 고려하여 미리 정의된 ‘continental’로 설정하였다. 조건표 구축에 필요한 대기 보정 파라미터의 범위 및 간격은 Sentinel-2 위성과 우리나라 대기 환경을 잘 반영할 수 있는 KOMPSAT-3 위성 영상 대기 보정 조건표를 참고하였으며, 각 파라미터마다 계산되는 경우의 수는 Table 4에 나타내었다(Jung et al., 2020). AOD와 같이 대기 보정에 미치는 영향이 매우 큰 경우에는 타 파라미터에 비하여 많은 경우에 대해 계산을 수행하였으며, 발생 빈도를 고려하여 낮은 AOD에 대해서는 더 촘촘한 간격으로 조건표를 구축하였다(Lee et al., 2015). 따라서 조건표는 1,008,000건의 경우의 수에 대한 정보를 담게 되며, 각 밴드별로 구축하여 총 5개의 조건표가 구축되었다.

3.2. 복사전달모델 모사 기계학습 모델

조건표는 대기 보정 파라미터마다 특정 범위 내에서 일정한 값들의 조합으로 구축되었다. 그러나 현실에서의 대기 조건은 매우 다양하기 때문에 표에 구축된 것과 같이 특정 대기 보정 파라미터 값에 해당하지 않는 상황이 일반적이다. 따라서, 표에 제시되지 않은 대기

Table 5. Default settings of aerosol optical depth and water vapor for each month

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
AOD	0.1	0.1	0.2	0.2	0.3	0.4	0.4	0.3	0.2	0.2	0.1	0.1
WV [g/cm ²]	0.5	0.5	1	1	1	2	2	2	1	1	1	0.5

보정 파라미터에 해당할 경우에는 조건표로부터 내삽 기법을 통해 지표 반사도를 추출할 수 있도록 내삽 방법을 제안하였다. 대기 보정 변수와 대기 보정 계수 사이의 비선형적인 관계가 있기 때문에, 기계학습이 빠르고 높은 정확도를 가질 것으로 예상되었으므로 조건표 내삽 기법으로 두 가지 기계학습 방법을 비교 분석하였다.

본 연구에서 시험한 기계학습 기법은 의사 결정 나무에 기반을 둔 Random Forest (RF)와 Light Gradient Boosting Machine (LGBM)이다. RF는 무작위로 선택된 다수의 의사 결정 나무를 기반으로 하는 앙상블 학습 기법으로, 다수의 의사 결정 나무에서 얻은 예측 결과를 평균하여 회귀 문제의 최종 결과를 생성한다(Belgiu and Drăguț, 2016). RF의 주요한 특징은 의사 결정 나무들 사이의 독립성을 보장하기 위한 무작위성이다. 훈련 데이터에서 무작위로 일부를 추출하여 부트스트랩과 집계 (bagging)를 통해 사용하고, 독립 변수도 일부를 무작위로 선택하여 의사 결정 나무를 형성한다(Abdi, 2020).

LGBM은 의사 결정 나무에 기반을 두는 점이 RF와 동일하지만, bagging이 아닌 boosting 방법으로 모델을 훈련하는 점이 가장 큰 특징이다. LGBM은 트리의 균형을 맞추지 않고, 손실값이 가장 큰 리프 노드를 연속적으로 분할하여 깊고 불균형한 트리를 생성하는 리프 기반 분할 방식을 채택하고 있다(Fan et al., 2019). 이를 통해 전통적인 트리 기반 분할 방식보다 예측 오류 손실을 최소화할 수 있다. 특히 대량의 데이터를 다루는 데 뛰어난 성능을 보이고 있다.

본 연구에서 기계학습 기반 내삽 모델은 조건표 구축에 활용된 7개 대기 보정 파라미터를 독립 변수로, 대기 보정 계수인 x_a, x_b, x_c 를 각각 종속 변수로 활용하여 개발되었다. 각 대기 보정 계수를 추정할 수 있는 모델을 개별적으로 구축하였기 때문에 5개 위성 밴드와 3개 대기 보정 계수에 대해서 총 15개 모델이 개발되었다.

3.3. 대기보정 결과 산출 방법

대기 보정에 필요한 7개 파라미터 중 SZA, VZA,

RAA는 위성 관측 자체에서 취득이 가능한 값으로 실시간으로 얻을 수 있다. 이외에 외부에서 취득이 필요한 값으로는 AOD, WV, O₃, 고도 정보가 있다. 따라서 외부에서 취득이 필요한 파라미터 수급 방안에 따라서 두 가지 방법으로 대기보정을 수행하였다. 첫 번째 방법은 L2A 자료 산출에 사용되고 있는 대기 보정 파라미터 값을 그대로 사용하는 것으로, 이를 통해 동일 대기 보정 파라미터 조건에서 L2A 지표면 반사도와 비교를 수행하였다. 두 번째 방법은 향후 실제 농림위성에 적용할 수 있도록 설계하였다. 선행 연구에 제시된 파라미터 민감도 분석 결과를 토대로 DEM은 SRTM으로부터 취득된 값을 활용하였으며 Sentinel-2 타일의 평균 값, O₃은 300 DU 고정 값으로 사용하였다(Kang et al., 2021). AOD와 WV는 GK-2A 위성 산출물로부터 취득된 값을 활용하였으며, 별도의 전처리 과정을 거쳐 대기 보정에 활용하였다(Fig. 2).

GK-2A로부터 취득된 AOD 및 WV 값은 구름과 자체 quality flag 정보를 활용하여 유의미한 값들만을 활용하였다. 이후 Sentinel-2 위성 관측 시간 이전의 30분 동안의 자료를 모두 평균하고, 그 이후에도 만약 결측치가 존재할 경우에는 이중선형 보간법을 이용하여 메꾸어 주었다. 간혹 구름이 한반도 전반에 머물고 있거나 자료 수급의 문제로 대기 보정 파라미터가 제공되지 않는 경우가 존재할 수 있으므로, 본 연구에서는 우리나라

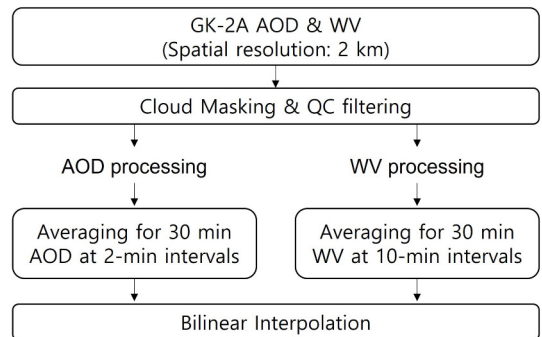


Fig. 2. Processing flow to derive atmospheric correction parameters from GK2A data.

라의 대기 조건을 고려하여 AOD 및 WV의 월별 기본 값을 구축하고 이를 이용할 수 있도록 하였다(Table 5). 월별 기본 값은 동아시아 지역 내에서의 AOD 및 WV에 대한 선행 연구를 기반으로 설정되었다(Tan et al., 2022; Kang et al., 2021).

3.4. 평가 방법

3.4.1. 복사전달모델 모사 알고리즘 평가 방법

복사전달모델 기반의 조건표를 내삽하고 모사할 수 있는 모델을 개발하기 위하여 조건표를 훈련 자료로 사용하고, 조건표에 구축된 조건과 다른 범위 및 간격을 통하여 별도의 검증표를 구축하였다(Table 6). 검증표는 기존에 구축된 조건표와 전혀 겹치지 않는 대기 보정 파라미터들의 조합으로 이루어져 있기 때문에 개발된 모델의 내삽 성능을 평가하기에 적합하며, 총 588,000 경우의 수에 대하여 구축되었다. 검증 지표로는 결정계수 (coefficient of determination, R^2), 평균 제곱근 오차(Root Mean Squared Error, RMSE), 평균 편향 오차(Mean Bias Error, MBE), 평균절대비오차(Mean Absolute Percentage Error, MAPE)를 사용하였다(식 2-5).

Table 6. Atmospheric parameter range, interval, and number of cases used for evaluation

Parameter	Range	Interval	No. of case
SZA	3.5–75°	5°	15
VZA	5–40°	10°	4
RAA	10–200°	20°	10
WV	0.5–5 g/cm ²	1 g/cm ²	5
O3	290–350 DU	50 DU	2
AOD	0.015, 0.075, 0.135, 0.195, 0.23, 0.5, 0.7, 0.9, 1.3, 1.8, 2.3, 2.8, 3.5, 4.5		14
DEM	150–2100 m	300 m	7

$$R^2 = \frac{\sum_{i=1}^n (y_i - x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (3)$$

$$MBE = \frac{\sum_{i=1}^n (x_i - y_i)}{n} \quad (4)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{y_i} \right| \quad (5)$$

식에서 n 은 전체 자료의 개수, x_i 와 y_i 는 각각 모델 예측 값과 실측 값을 의미하며, \bar{x} 와 \bar{y} 는 모델 예측 값과 실측 값들의 평균값을 의미한다.

3.4.2. 현장 관측 대비 정확도 평가 방법

본 연구는 Table 3에서 제시된 날들에 대해 세 가지 다른 대기보정 알고리즘, 즉 Sentinel-2 L2A, Sentinel-2에서 추출된 대기 보정 파라미터를 활용한 기계학습 대기 보정 결과 그리고 GK-2A 대기 보정 파라미터를 활용한 기계학습 대기보정 결과를 기반으로 현장 관측 반사도 값을 추출하고, 이를 현장에서 직접 관측된 데이터와 비교하였다.

대기 보정된 결과는 위치 오차를 감안하여 플렉스 타워 좌표를 중심으로 주변 3 by 3 격자(30 × 30 m 범위) 내의 데이터를 평균하여 현장 관측값과 비교하였다. 현장 관측값은 Sentinel-2 위성이 해당 지역을 촬영하는 시간대를 고려하여 오전 11시부터 11시 20분 사이의 데이터를 평균하여 분석에 활용하였다.

본 연구에서는 대기 보정의 정확도를 평가하기 위하여 두 가지 접근법을 채택하였다. 첫째로, 밴드별 반사도와 식생 지수를 대상으로 통계적 검증 지표를 활용하여 평가하였다. 상관계수(Correlation coefficient, R), RMSE, MBE 등의 지표를 이용하여 대기 보정 결과와 현장에서의 관측치 사이의 상관성과 편차를 정량적으로 분석하였다. 본 연구에서 중점적으로 살펴본 식생 지수는 NDVI로, 이는 Red 밴드와 NIR 밴드를 활용한 식(6)를 통해 도출되었다. NDVI의 값은 일반적으로 -1에서 1 사이로 나타나며, 식생의 건강 상태 및 농업 생산량을 간접적으로 추정하는 데에 주요한 역할을 한다.

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (6)$$

식에서의 ρ 는 각 밴드에 대한 반사도를 나타낸다.

둘째로, 시계열 데이터 기반의 다시기 검증을 통해 평가하였다. 시기별로 대기 보정 결과와 현장에서의 관측치 간의 편차에서 나타나는 특성을 분석하였으며, 현장에서 관측된 NDVI 값의 패턴이 위성 영상 기반 NDVI에서 얼마나 잘 재현되는지에 대한 유사성도 함께 검토하였다.

Table 7. Evaluation results of the RF atmospheric correction simulation model

Band number	Atmospheric coefficient	R ²	RMSE	MAPE (%)	MBE
Band 2	xa	0.93	0.0142	14.8	-0.0028
	xb	0.95	0.9629	14.5	-0.2532
	xc	0.99	0.0055	2.4	0.0003
Band 3	xa	0.92	0.0114	14.3	-0.0021
	xb	0.95	0.6143	14.5	-0.1609
	xc	0.99	0.0063	3.1	-0.0001
Band 4	xa	0.92	0.0083	13.7	-0.0015
	xb	0.95	0.3543	14.5	-0.0921
	xc	0.99	0.0069	4.0	-0.0003
Band 5	xa	0.92	0.0079	12.9	-0.0014
	xb	0.95	0.3015	14.4	-0.0781
	xc	0.99	0.0070	4.3	-0.0004
Band 8	xa	0.92	0.0078	10.9	-0.0014
	xb	0.95	0.1950	14.4	-0.0493
	xc	0.99	0.0066	4.9	-0.0006

Table 8. Evaluation results of the LGBM atmospheric correction simulation model

Band number	Atmospheric coefficient	R ²	RMSE	MAPE (%)	MBE
Band 2	xa	0.90	0.0162	17.9	0.0021
	xb	0.93	1.1275	22.7	0.1629
	xc	0.99	0.0068	2.6	-0.0004
Band 3	xa	0.90	0.0110	16.1	0.0010
	xb	0.93	0.5994	21.1	0.0614
	xc	0.99	0.0078	3.5	-0.0006
Band 4	xa	0.91	0.0076	15.0	0.0005
	xb	0.93	0.3225	24.0	0.0222
	xc	0.99	0.0085	4.7	-0.0008
Band 5	xa	0.91	0.0071	14.3	0.0004
	xb	0.93	0.2666	25.4	0.0154
	xc	0.99	0.0085	5.1	-0.0008
Band 8	xa	0.91	0.0069	12.4	0.0003
	xb	0.93	0.1647	28.3	0.0050
	xc	0.99	0.0079	5.8	-0.0009

4. 연구 결과 및 분석

4.1. 기계학습 모델 자체 평가

4.1.1. 기계학습 기법 간의 성능 비교

RF 및 LGBM 기계학습 기법을 통하여 밴드별 대기 보정계수를 추정하는 모델의 정확도 평가 결과는 각각 Tables 7, 8에 나타내었다. RF와 LGBM 모두 R²가 0.90 이상으로 매우 좋은 성능을 보여주었다. 3개의 대기보

정계수의 절대적인 범위는 각각 xa는 0-0.4, xb는 0-30, xc는 0.1-0.3 수준임을 고려했을 때, RMSE, MAPE, MBE가 모두 낮게 나타났으므로 기계학습 모델의 대기보정 계수 추정의 정확도는 매우 높다고 간주할 수 있다. 밴드별로 정확도에 나타나는 경향성은 눈에 띄지 않았으며, 대기보정 계수별로 평가 결과를 비교했을 때는 xa < xb < xc의 순서대로 정확도가 더 높게 나타났다. 모든 밴드에 대해서 xc 추정 모델의 R²는 0.99로 매우 높았으며,

MAPE도 6% 이하로 매우 낮았다. 두 기법이 가지는 성능은 각 밴드별, 대기보정 계수별로 순위가 서로 다르게 나타났기 때문에 특정 기법이 항상 뛰어나다고 말할 수는 없었다. 그러나 공통적인 특징은 RF는 LGBM에 비해서 MBE가 약 2배에서 10배가량 크게 나타났다는 점이다. 따라서 일반적인 성능이나 산출 속도의 효율성을 고려했을 때 RF보다 LGBM 기법이 복사전달모델 기반 대기보정 알고리즘을 모사하는 데 더 탁월하다고 판단되었다.

4.1.2. 최고 기계학습 기법의 현장 관측 대비 평가 결과

Table 9에서는 현장 관측소에서 얻은 데이터를 기반으로 Sentinel-2 L2A, Sentinel-2에서 추출된 대기 보정 파라미터를 활용한 기계학습 대기보정 결과 그리고 GK-2A 대기 보정 파라미터를 활용한 기계학습 대기보정 결과의 정확도를 비교하였으며, 기계학습 기법으로는 LGBM 결과를 활용하였다. 광릉 활엽수 지역(GBK)에서의 세 비교 모델의 정확도는 R 값이 0.73-1.00, RMSE는 0.007-0.068로써, 네 관측소 중에서 가장 높은 일치도를 보였다. 또한, Figs. 3(a-c)의 NDVI 비교 그래프에서도 해당 지역의 알고리즘 성능이 높은 일치도를 나타냈

다. 반면에 광릉 침엽수림(GCK) 지역은 상대적으로 가장 낮은 상관 관계를 보였다. 해당 지역의 전나무림 특성 상 연중 밴드별 반사도가 더 낮고, 시계열 NDVI 변동폭이 크지 않아 선형 상관성이 낮게 나타났다는 것이 Lim et al. (2022) 연구 결과와 일치하였다. 홍천 지역(HC)에서의 분석 결과, Sentinel-2 L2A와 Sentinel-2 파라미터 기반 대기보정 결과는 R 값이 최소 0.53에서 최고 0.99로 현장 관측값과 상당히 높은 선형 관계를 나타내었다. 그러나 GK-2A 파라미터 기반 대기보정 결과는 상대적으로 높은 RMSE 값을 보이며 성능 저하가 관찰되었다. 국립산림과학원(HAWS1) 지역에서는 모든 밴드가 전반적으로 현장 관측치와 높은 상관성을 나타냈으나, 대기 보정된 NDVI가 현장관측 NDVI에 비해 과대모의되는 경향이 관찰되었다. Fig. 3(j, k, l)에 따르면, 현장 관측 기반의 NDVI가 상대적으로 낮을 때, 대기보정된 NDVI는 더 높은 값을 나타내었다. 이러한 현상은 식생 활력도가 약한 시기(1~3월, 11~12월)에 Band 2, 3, 4의 반사도 값이 현장 관측값에 비해 낮게 추정되었지만, Band 8은 현장에서의 관측값과 거의 일치하게 추정되었기 때문에 발생한 것으로 해석된다. 결과적으로 Sentinel-2 L2A가 현장 관측값과 가장 높은 상관관계를 보였고, 이

Table 9. Comparison of the evaluation results against in-situ measurements by atmospheric correction algorithms (i.e., Sentinel-2 L2A, Sentinel-2 parameters based LGBM, and GK-2A parameters based LGBM)

In-situ	Band number	N	Sentinel-2 L2A		Sentinel-2 parameters based LGBM		GK-2A parameters based LGBM	
			R	RMSE	R	RMSE	R	RMSE
GBK	Band 2	27	0.99	0.008	0.98	0.011	0.73	0.039
	Band 3		0.99	0.008	0.97	0.011	0.84	0.025
	Band 4		0.99	0.007	0.98	0.010	0.95	0.015
	Band 8		1.00	0.018	0.95	0.068	0.98	0.034
GCK	Band 2	27	0.96	0.007	0.75	0.011	0.56	0.047
	Band 3		0.97	0.010	0.87	0.008	0.68	0.021
	Band 4		0.87	0.011	0.69	0.012	0.61	0.012
	Band 8		0.97	0.051	0.96	0.042	0.96	0.042
HC	Band 2	21	0.94	0.018	0.84	0.029	0.45	0.063
	Band 3		0.90	0.021	0.88	0.031	0.33	0.056
	Band 4		0.93	0.023	0.91	0.033	0.89	0.040
	Band 8		0.95	0.033	0.93	0.059	0.94	0.059
HAWS1	Band 2	30	0.98	0.007	0.95	0.009	0.53	0.038
	Band 3		0.96	0.007	0.95	0.011	0.68	0.026
	Band 4		0.96	0.014	0.94	0.016	0.79	0.025
	Band 8		0.99	0.016	0.98	0.041	0.98	0.043

어서 Sentinel-2 파라미터 기반 LGBM 대기보정 그리고 GK-2A 파라미터 기반 LGBM 대기보정 순서로 상관관계가 높았다.

Sentinel-2 parameter를 사용했음에도 LGBM 모델의 결

과가 L2A 결과와 다른 원인에는 복사전달모델의 차이, LUT 구축 간격 및 범위의 차이, LGBM 모델 성능에 의한 차이 등이 있다. 본 연구에서는 6SV 복사전달모델을 기반으로 대기보정을 수행하였으나 Sentinel-2 지표면

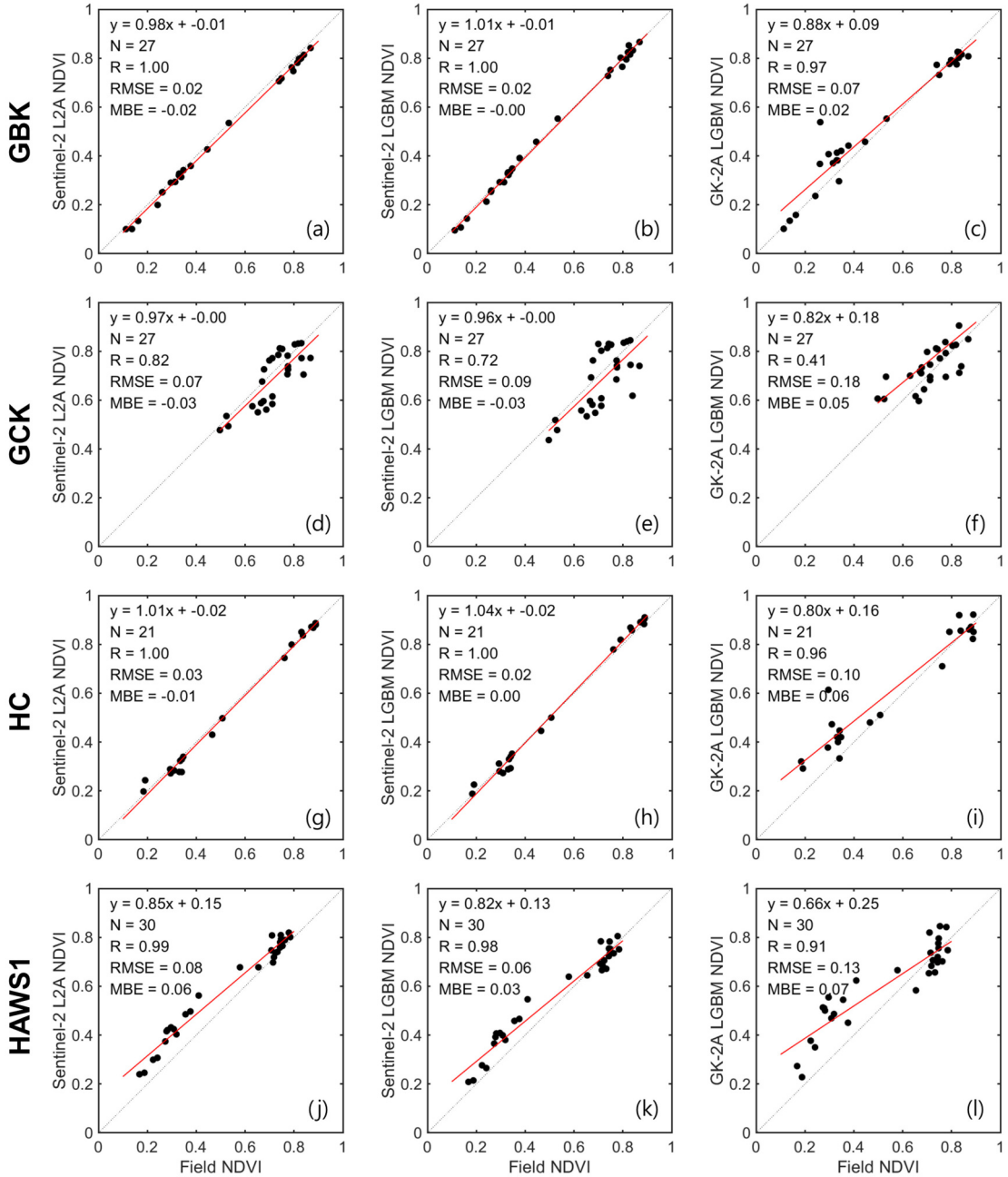


Fig. 3. Comparative assessment of three atmospheric correction algorithms (i.e., L2A, Sentinel-2 parameter-based LGBM, and GK-2A parameter-based LGBM) using in-situ measurements. (a)–(c) Gwangneung broadleaved forest (GBK), (d)–(f) Gwangneung coniferous forest (GCK), (g)–(i) Hongcheon (HC), and (j)–(l) National Institute of Forest Science (HAWS1).

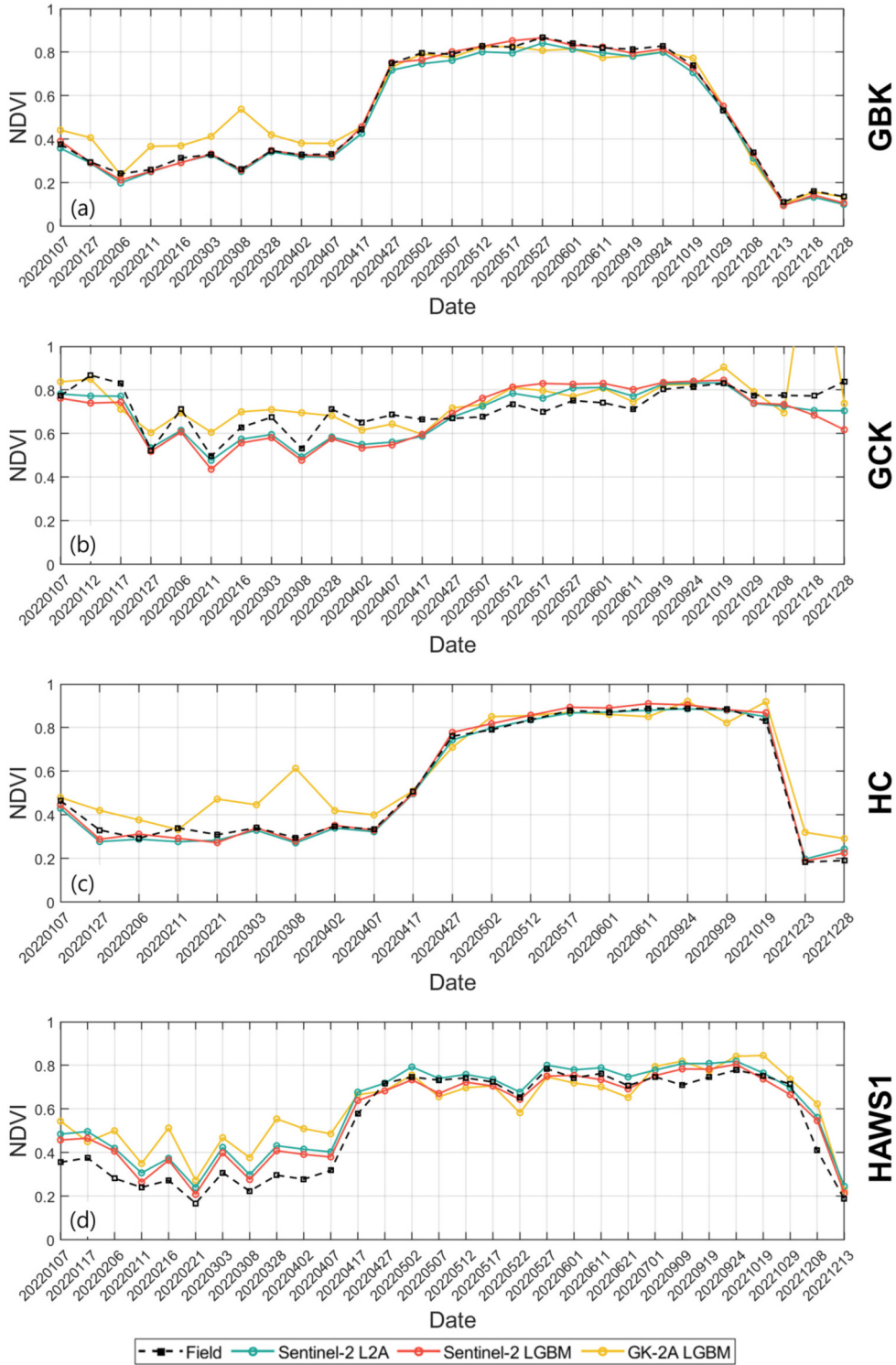


Fig. 4. NDVI time series at (a) Gwangneung broadleaved forest (GBK), (b) Gwangneung coniferous forest (GCK), (c) Hongcheon (HC), and (d) National Institute of Forest Science (HAWS1)) for in-situ measurements and three atmospheric correction algorithms (i.e., L2A, Sentinel-2 parameter-based LGBM, and GK-2A parameter-based LGBM).

반사도를 산출하는 Sen2Cor는 libRadtran을 기반으로 대기보정을 수행하므로 근본적으로 동일한 대기보정 파라미터 조건에서도 산출되는 지표면 반사도는 같을 수 없다(Main-Knorn et al., 2017). 뿐만 아니라, 본 연구에서 LUT를 구축할 때에 설정한 대기보정 파라미터의 범위 및 간격이 Sentinel-2 대기보정을 위하여 구축된 조건과 상이하며, LGBM 모델이 높은 성능을 보장하기는 하지만 복사전달모델 결과를 100% 구현할 수 없으므로 이에 의한 차이도 존재한다.

4.2. 시계열 평가

Fig. 4에는 관측소별 NDVI 시계열 패턴이 제시되어 있다. 광릉 활엽수림(GBK), 홍천(HC), 국립산림과학원(HAWS1) 지역은 활엽수림과 혼효림의 특성상 봄 및 여름 기간 동안 식생의 활성화와 관련하여 NDVI 지수가 약 0.8 정도로 상승하는 반면, 가을과 겨울에는 식생의 비활성화로 인해 약 0.2까지 NDVI 지수의 감소가 관찰되었다. 그러나 광릉 침엽수림(GCK) 지역에서는 혼효림에 비해 상대적으로 0.5-0.8 범위의 일정한 NDVI 분

포를 보여주었다. 특히, 가을에도 잎을 떨어뜨리지 않는 침엽수의 특성에 따라 연간 높은 NDVI 값을 유지하는 것이 확인되었다(Son et al., 2021).

대기 보정 알고리즘을 통해 얻어진 NDVI 값은 전반적으로 현장 관측값에 잘 부합하는 경향을 나타냈다. 그러나 GK-2A 파라미터를 기반으로 한 대기보정 결과는 일부 이상치를 포함하는 것으로 확인되었다. 특히, Fig. 4(a) 광릉 활엽수림과 (c) 홍천 지역에서 2022년 3월 8일에 현장 관측값과의 차이가 두드러졌다. Fig. 5를 기준으로 볼 때, 해당 일자의 GK-2A AOD 값은 Sentinel-2 AOD 값에 비해 지나치게 높았던 부분이 대기보정 과정에서 반사도 산출에 오차를 초래한 것으로 추측된다. 반면에, Sentinel-2 파라미터를 활용한 LGBM 기반 대기보정 기법은 L2A 결과와 상당히 유사한 정확도를 나타냈다. 이러한 결과는 결측 영향을 최소화할 수 있는 고품질의 파라미터를 활용하게 되면 기계학습 기반의 조건표 생성 방식이 복잡한 복사전달모델과 상응하는 성능을 가질 수 있음을 시사하며, 자세한 분석은 이미지 평가 섹션에서 함께 수행하였다.

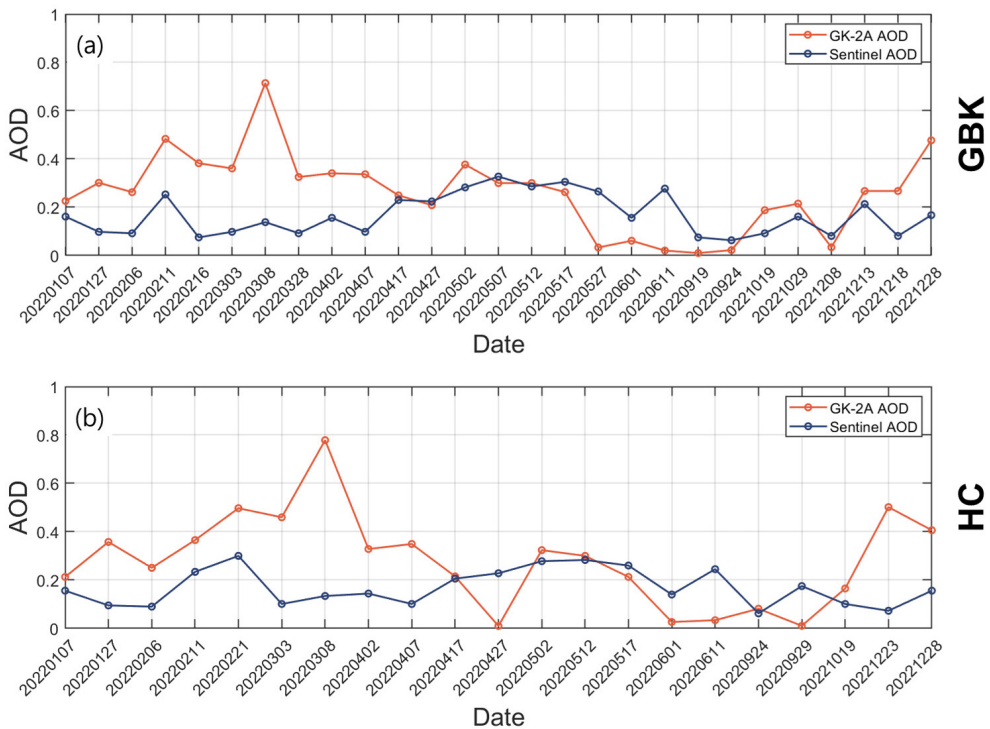


Fig. 5. AOD time series of Sentinel-2 and GK-2A at (a) Gwangneung broadleaved forest (GBK) and (b) Hongcheon (HC).

4.3. 이미지 평가

시계열 평가 시 일부 관측소에서 Sentinel-2 L2A 값과 LGBM 기반 대기보정 모사 알고리즘의 결과의 차이가 가장 컸던 2022년 3월 8일자 영상 분석을 통하여 상기 두 알고리즘의 차이 발생 원인과 개선 방향을 살펴보았다. Blue band에서의 차이가 극명하게 나타났으므로, blue band를 활용하여 이미지 간 비교 평가를 수행하였다. 당시 관측소 일대의 LGBM blue band 반사도는 L2A 반사도에 비해서 전반적으로 낮게 나타난 반면에 AOD는 이와 반대의 경향을 보여주었다(Fig. 6). L2A AOD는 범위가 0.01-0.02 수준에 머물렀으나, LGBM 대기보정에 활용된 GK-2A AOD는 비교적 높은 값(0.1-0.7)을 가지고 있었다(Figs. 6a, b). 일반적으로 대기보정 결과에 가장 영향이 큰 대기보정 파라미터로 AOD가 꼽히며, 특히 파장이 짧은 blue band에 대한 영향이 더욱 큰 것으로 알려져 있다(Kang et al., 2021). AOD가 높으면 위성에서 관측된 신호 중 대기에 의한 비중이 높기 때문에 지표면 반사도가 낮게 산출된다. 따라서 L2A에 비해서 LGBM의 반사도가 낮게 나타난 이유는 AOD에 가장 크게 영

향을 받았을 것으로 보이며, L2A에 비하여 LGBM의 반사도가 특히 낮게 나타난 우측 상단부에서 GK-2A AOD가 매우 높았던 것을 통해서도 확인할 수 있다.

GK-2A AOD가 높게 나타난 이유는 누적 AOD 자료를 사용했음에도 불구하고 quality flag로 인해 발생한 결측값 때문이다. Fig. 6에 의하면 현장 관측소 위치 주변으로 GK-2A AOD에 결측이 존재하였다. 최초로 결측된 부분들이 보간을 수행하면서 높은 AOD 값을 가지도록 처리됨으로써 관측소 위치에서는 높은 AOD를 활용한 대기보정이 수행되었다. 본 연구에서는 AOD 및 WV에 결측이 존재할 경우 이를 보간하여 대기보정에 사용함으로써 대기보정의 산출 범위를 최대한으로 확보하고자 하였다. 그러나 Fig. 6와 같이 결측 범위가 넓은 경우에는 보간에 의한 AOD에 대한 신뢰도가 떨어져 대기보정의 정확도가 저하되는 한계점이 있다. 위 분석을 통하여 GK-2A AOD의 결측이 존재할 경우 대기보정의 정확도에 미치는 영향을 알아보았으며, 결측이 존재하지 않을 경우에는 안정적인 결과를 도출할 수 있는 것으로 판단되었다. 따라서, 본 연구에서는 GK-2A AOD를 통한 대기보정이 가능하지만, 결측 문제를 해

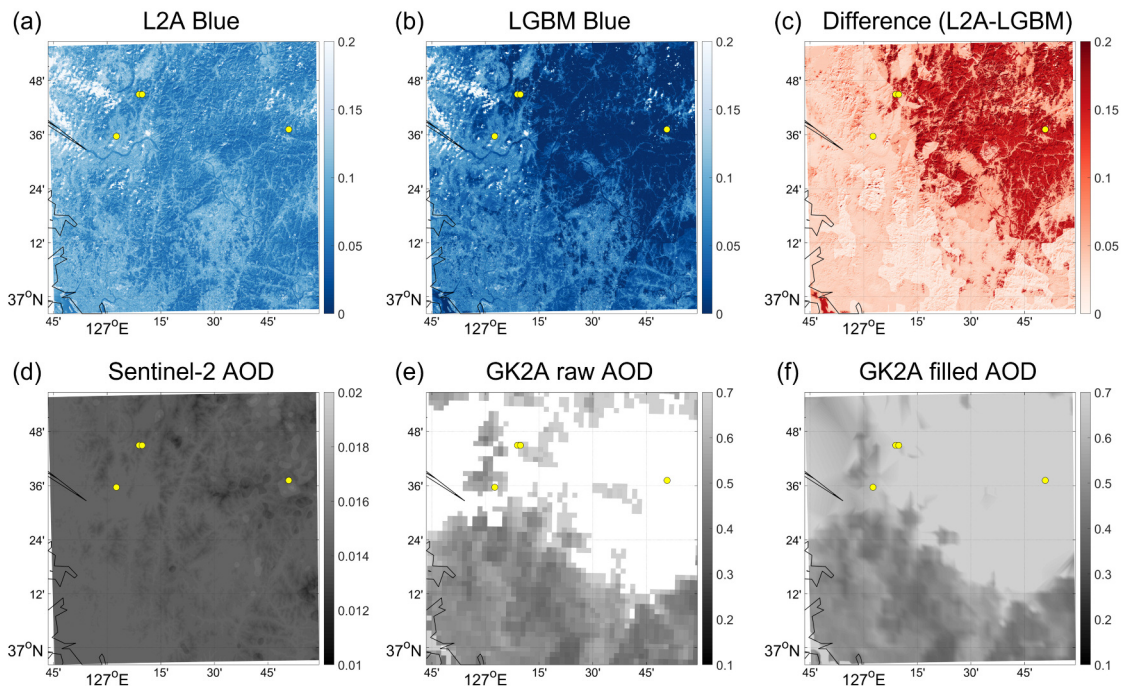


Fig. 6. Spatial distribution of surface reflectance (a-b), difference between L2A and GK-2A parameter-based LGBM (c), sentinel-2 AOD (d), GK-2A raw AOD (e), and GK-2A AOD after preprocessing (f) on March 8th, 2022.

결하기 위하여 추후 여러 위성으로부터 산출된 융합 AOD 혹은 향상된 AOD 보간 방법 등을 활용할 경우 대기보정 알고리즘 성능이 향상될 것으로 기대된다.

사사

본 연구는 농촌진흥청 연구사업(과제번호: PJ01623 42021)의 지원을 받아 수행되었습니다.

5. 결론

본 연구에서는 Sentinel-2 위성 영상을 이용하여 복사전달모델 기반 대기보정을 모사하는 기계학습 기반 대기보정 알고리즘을 개발하고 적용 가능성을 평가하였다. RF와 LGBM 두 가지 기계학습 기법을 적용하였으며, 두 기법 모두 대기보정을 수행하기에 충분한 성능을 보여주었으나 산출 속도를 고려했을 때 LGBM이 효율적인 모델이라고 판단되었다. GK-2A 대기보정 파라미터를 사용한 LGBM 기반 대기보정 모사 알고리즘의 현장 관측대비 NDVI의 정확도는 R: 0.72-1, RMSE: 0.02-0.09, MBE: -0.03-0.03으로써, 활용수립을 포함한 혼합림에서 나타나는 연중 식생 활력도 변화를 잘 모의할 수 있었다.

일부 현장 관측 자료 대비 차이가 큰 경우는 대기보정에 활용된 AOD에 의해 크게 좌우되는 것으로 분석되었다. 농림위성의 경우 주요 파라미터인 AOD를 자체 제공하기 어렵기 때문에 외부에서 제공하기 위한 방안으로 GK-2A AOD를 활용하였으나, 광역에 대해서 지속적으로 결측이 존재할 경우 대기보정 결과에 치명적일 것으로 예상된다. 따라서, 대기보정 된 지표면 반사도의 신뢰성을 확보하기 위해서는 대기보정 파라미터 수급에 따른 quality flag를 제공하는 방안이 마련될 필요가 있다고 사료된다. 더 나아가, 지표면 반사도 제공 범위를 최대한으로 확보하기 위해서는 추후 여러 위성으로부터 산출된 융합 AOD 혹은 향상된 AOD 보간 방법 등을 개발할 필요가 있다.

본 연구는 Sentinel-2 위성 영상을 대체로 활용하여 기계학습 기반 대기보정 모델을 개발하고, 현재의 한계점과 향후 더 정확한 지표면 반사도 산출을 위하여 고려해야 할 사항에 대해 제시하였다. 추후 농림위성 발사 후 본 연구를 토대로 하여 추가 분석을 통해 더욱 정밀한 대기보정이 수행될 수 있을 것으로 기대된다.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

Abdi, A. M., 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, 57(1), 1–20. <https://doi.org/10.1080/15481603.2019.1650447>

Anderson, G. P., Pukall, B., Allred, C. L., Jeong, L. S., Hoke, M., Chetwynd, J. H. et al., 1999. FLAASH and MODTRAN4: state-of-the-art atmospheric correction for hyperspectral data. In *Proceeding of the 1999 IEEE Aerospace Conference (Cat. No.99TH8403)*, Snowmass, CO, USA, Mar. 7, pp. 177–181. <https://doi.org/10.1109/AERO.1999.792088>

Belgiu, M., and Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjrs.2016.01.011>

Berry, P. A. M., Garlick, J. D., and Smith, R. G., 2007. Near-global validation of the SRTM DEM using satellite radar altimetry. *Remote Sensing of Environment*, 106(1), 17–27. <https://doi.org/https://doi.org/10.1016/j.rse.2006.07.011>

Chavez Jr, P. S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*,

- 24(3), 459–479. [https://doi.org/10.1016/0034-4257\(88\)90019-3](https://doi.org/10.1016/0034-4257(88)90019-3)
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F. et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>
- Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X., and Zeng, W., 2019. Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agricultural Water Management*, 225, 105758. <https://doi.org/10.1016/j.agwat.2019.105758>
- Ghimire, B., Rogan, J., Galiano, V. R., Panday, P., and Neeti, N., 2012. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing*, 49(5), 623–643. <https://doi.org/10.2747/1548-1603.49.5.623>
- Jee, J. B., Lee, K. T., Lee, K. H., and Zo, I. S., 2020. Development of GK-2A AMI aerosol detection algorithm in the East-Asia region using Himawari-8 AHI data. *Asia-Pacific Journal of Atmospheric Sciences*, 56, 207–223. <https://doi.org/10.1007/s13143-019-00156-3>
- Jung, D., Jin, D., Seong, N. H., Lee, K. S., Seo, M., Choi, S. et al., 2020. Sensitivity analysis of surface reflectance retrieved from 6SV LUT for each channel of KOMPSAT-3/3A. *Korean Journal of Remote Sensing*, 36(5–1), 785–791. <https://doi.org/10.7780/kjrs.2020.36.5.1.11>
- Kang, Y. J., Cho, D. J., Han, D. H., Im, J. H., Im, J. B., Oh, K. H. et al., 2021. Sensitivity analysis for CAS500-4 atmospheric correction using simulated images and suggestion of the use of geostationary satellite-based atmospheric parameters. *Korean Journal of Remote Sensing*, 37(5–1), 1029–1042. <https://doi.org/10.7780/kjrs.2021.37.5.1.16>
- Kaufman, Y., Tanré, D., Gordon, H., Nakajima, T., Lenoble, J., Frouin, R. et al., 1997a. Passive remote sensing of tropospheric aerosol and atmospheric correction for the aerosol effect. *Journal of Geophysical Research: Atmospheres*, 102(D14), 16815–16830. <https://doi.org/10.1029/97JD01496>
- Kaufman, Y., Tanré, D., Remer, L. A., Vermote, E., Chu, A., and Holben, B., 1997b. Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *Journal of Geophysical Research: Atmospheres*, 102(D14), 17051–17067. <https://doi.org/10.1029/96JD03988>
- Kim, J. U., Yun, J. H., Lee, K. W., Joo, M. W., Lim, T. K., Choi, D. J. et al., 2021. Development of electro-optical subsystem for CAS500-4 CAP-W. In *Proceeding of the Korean Society for Aeronautical and Space Sciences Fall Conference*, Jeju, Republic of Korea, Nov. 11–19, pp. 465–466.
- Kim, Y. H., Im, J., Ha, H. K., Choi, J. K., and Ha, S., 2014. Machine learning approaches to coastal water quality monitoring using GOCI satellite data. *GIScience & Remote Sensing*, 51(2), 158–174. <https://doi.org/10.1080/15481603.2014.900983>
- Kwon, S. K., Kim, K. M., and Im, J. B., 2021. A study on pre-evaluation of tree species classification possibility of CAS500-4 using RapidEye satellite imageries. *Korean Journal of Remote Sensing*, 37(2), 291–304. <https://doi.org/10.7780/kjrs.2021.37.2.9>
- Lantzanakis, G., Mitraaka, Z., and Chrysoulakis, N., 2017. Comparison of physically and image based atmospheric correction methods for Sentinel-2 satellite imagery. In: Karacostas, T., Bais, A., Nastos, P. (eds.), *Perspectives on atmospheric sciences*, Springer, pp. 255–261. https://doi.org/10.1007/978-3-319-35095-0_36
- Lee, C. S., Yeom, J. M., Lee, H. L., Kim, J. J., and Han, K. S., 2015. Sensitivity analysis of 6S-based look-up table for surface reflectance retrieval.

- Asia-Pacific Journal of Atmospheric Sciences*, 51, 91–101. <https://doi.org/10.1007/s13143-015-0062-9>
- Lee, K. S., Lee, C. S., Seo, M., Choi, S., Seong, N. H., Jin, D. et al., 2020. Improvements of 6S look-up-table based surface reflectance employing minimum curvature surface method. *Asia-Pacific Journal of Atmospheric Sciences*, 56, 235–248. <https://doi.org/10.1007/s13143-019-00164-3>
- Lim, J. B., Cha, S. E., Won, M. S., Kim, J., Park, J. H., Ryu, Y. R. et al., 2022. Design of calibration and validation area for forestry vegetation index from CAS500-4. *Korean Journal of Remote Sensing*, 38(3), 311–326. <https://doi.org/10.7780/kjrs.2022.38.3.7>
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., and Gascon, F., 2017. Sen2Cor for sentinel-2. In *Proceeding of the SPIE Conference on Image and Signal Processing for Remote Sensing 23*, Warsaw, Poland, Sept. 11–14, pp. 37–48. <https://doi.org/10.1117/12.2278218>
- Mudiyansele, S., Abd-Elrahman, A., Wilkinson, B., and Lecours, V., 2022. Satellite-derived bathymetry using machine learning and optimal Sentinel-2 imagery in South-West Florida coastal waters. *GIScience & Remote Sensing*, 59(1), 1143–1158. <https://doi.org/10.1080/15481603.2022.2100597>
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., and Stenseth, N. C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9), 503–510. <https://doi.org/10.1016/j.tree.2005.05.011>
- Schläpfer, D., Borel, C. C., Keller, J., and Itten, K. I., 1998. Atmospheric precorrected differential absorption technique to retrieve columnar water vapor. *Remote Sensing of Environment*, 65(3), 353–366. [https://doi.org/https://doi.org/10.1016/S0034-4257\(98\)00044-3](https://doi.org/https://doi.org/10.1016/S0034-4257(98)00044-3)
- Son, M. B., Chung, J. H., Lee, Y. G., and Kim, S. J., 2021. A comparative analysis of vegetation and agricultural monitoring of Terra MODIS and Sentinel-2 NDVIs. *Journal of The Korean Society of Agricultural Engineers*, 63(6), 101–115. <https://doi.org/10.5389/KSAE.2021.63.6.101>
- Tan, Y., Wang, Q., and Zhang, Z., 2022. Assessing spatiotemporal variations of AOD in Japan based on Himawari-8 L3 V31 aerosol products: Validations and applications. *Atmospheric Pollution Research*, 13(6), 101439. <https://doi.org/https://doi.org/10.1016/j.apr.2022.101439>
- Vermote, E. F., and Kotchenova, S., 2008. Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research: Atmospheres*, 113(D23). <https://doi.org/https://doi.org/10.1029/2007JD009662>
- Vermote, E. F., Tanré, D., Deuze, J. L., Herman, M., and Morcette, J. J., 1997. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3), 675–686. <https://doi.org/10.1109/36.581987>
- Wilson, R. T., 2013. Py6S: A Python interface to the 6S radiative transfer model. *Computers & Geosciences*, 51(2), 166–171. <https://doi.org/10.1016/j.cageo.2012.08.002>
- Zhang, Z., He, G., Zhang, X., Long, T., Wang, G., and Wang, M., 2018. A coupled atmospheric and topographic correction algorithm for remotely sensed satellite imagery over mountainous terrain. *GIScience & Remote Sensing*, 55(3), 400–416. <https://doi.org/10.1080/15481603.2017.1382066>