GK-2A AMI
Algorithm Theoretical Basis Document

성층권 침투 대류운 탐지 알고리즘

이주현, 울산과학기술원
김미애, 울산과학기술원
임정호, 울산과학기술원

Version 1.1

2019. 11. 25.

국가기상위성센터
National Meteorological Satellite Center

<수정 이력>

문서 버전	수정내용	과학코드 버전	수정일	작성자
0.1	-	0.1.0.0	2016.01.15	김미애, 임정호
0.2	전 내용	0.2.0.0	2017.01.15	김미애, 임정호
0.3	전 내용	0.4.0.0	2017.09.29	이주현, 김미애, 임정호
0.4	- 기계학습 기법 수정에 따른 입력자료 개선 및 보완 - 검증 결과 보완	0.5.0.0	2018.11.11	이주현, 김미애, 임정호
1.0	- AMI 변환에 따른 세부 내용 수정	0.5.0.0	2019.04.15	이주현, 김미애, 임정호
1.1	- GK2A AMI L2 산출물에 대한 결과 추가	0.5.0.0	2019.11.15	이주현, 임정호

<목차>

1. 개요 1

1.1 목적 1

1.2 사용자 1

1.3 내용 1

1.4 관련 문서 1

2. 관측시스템 개요 2

2.1 산출물 2

2.2 센서특성 2

3. 알고리즘 소개 3

3.1 개요 3

3.2 처리 흐름 3

3.3 입력자료 4

3.4 이론적 배경 5

4. 시험자료 및 출력물 11

4.1 모의 입력자료 11

4.2 모의자료 기반 산출결과 11

5. 협업운영시 고려사항 16

5.1 수치계산 고려사항 16

5.2 프로그래밍 절차상의 고려사항 16

5.3 품질평가 및 진단 16

5.4 예외사항 처리 16

5.5 알고리즘 검증 16

6. 가정 및 제한 16

6.1 성능 16

6.2 예상되는 센서성능 17

6.3 산출물 개선을 위한 사전계획 17

7. 참고문헌 18

<표 목차>

표 1. AMI 채널 규격 및 OT 알고리즘에서 사용되는 채널 3
표 2. RANDOM FOREST 기반 성층권 침투 대류운 탐지 모델 구축을 위해 사용된 변수 7
표 3. 랜덤 포레스트 모델 평가 결과 12
표 4. 랜덤 포레스트 기반 상위 10위까지의 변수중요도를 보이는 변수와 각 변수의 특성 13
표 5 GK-2A 성층권 침투 대류운 탐지 알고리즘의 목표정확도. 본 목표정확도는 GOES-16 OT 알고리즘의 목표정확도와 동일. 14
표 6. 성층권 침투 대류운 탐지 예외사항 대처 방안 16

<그림 목차>

그림 1. 성층권 침투 대류운 탐지 대상 지역 (동아시아 확장지역, 22.4 ~ 46.4 °N, 107.3 ~ 148.1 °E) 2
그림 2. 성층권 침투 대류운 탐지 알고리즘 흐름도 4
그림 3. HIMAWARI-8 AHI의 R0.64 가시영상 자료(A)와 AQUA MODIS 0.64 ㎛ 채널의 가시영상 자료(B)를 이용한 성층권 침투 대류운 발생 후보지역 선정 및 밝기온도 기반 자동적 위치 재조정(D)을 통한 성층권 침투 대류운 기준 자료 구축의 예시 8
그림 4. (A) CART(CLASSIFICATION AND REGRESSION TREES)와 (B) 랜덤 포레스트(RF: RANDOM FOREST) 모델의 구성 9
그림 5. BT11.2 적외 채널과 R0.64 가시 채널 기반 후처리 흐름도 11
그림 6. 랜덤 포레스트 모델의 변수 중요도 결과 13
그림 7. 2019년 6월 15일 23:00 UTC 부터 2019년 6월 16일 03:00 UTC 까지의 동아시아 확장지역에 대한 성층권 침투 대류운 탐지 알고리즘의 POD(PROBABILITY OF DETECTION)와 FAR(FALSE ALARM RATE). 알고리즘의 목표 정확도는 POD 25 % 이상, FAR 75 % 이하임. (각 산출물의 시점은 YYYYMMDDHHMM UTC로 표현)) 14
그림 8. 2019년 6월 22일 00:00 UTC, 2019년 8월 8일 01:00 UTC 그리고 2019년 8월 10일 02:00 UTC의 GK2A AMI 영상에 대한 탐지 결과 (노란색 동그라미: 모델 탐지 지역, 노란색 화살표: 잘 탐지된 OT, 파란색 화살표: 미탐지 OT, 빨간색 화살표: 오탐지 지역, 각 영상의 시점은 YYYYMMDDHHMM UTC로 표현) 15

약어표

ABI	Advanced Baseline Imager
AMI	Advanced Meteorological Imager
ATBD	Algorithm Theoretical Basis Document
BT	Brightness temperature
BTD	Brightness Temperature Difference
CART	Classification and Regression Trees
DT	Decision trees
FAR	False Alarm Rate
FOV	Field of View
GK-2A	GEO-KOMPSAT 2A
MODIS	Moderate Resolution Imaging Spectro-radiometer
OT	Overshooting Tops
POD	Probability of Detection
R	Reflectance
RF	Random Forest

1. 개요

1.1 목적

성층권 침투 대류운(OT: Overshooting Top) 탐지 알고리즘 기술문서는 후속 정지궤도위성인 GK-2A에 탑재될 AMI(Advanced Meteorological Imager)의 적외 채널 밝기온도를 이용하여 OT를 탐지하기 위한 알고리즘에 대해 기술한다. 이 문서는 OT 탐지 알고리즘에 필요한 입력 자료와, 이론적인 배경, 현업 운영 시 고려사항, 산출물과 관련된 가정과 한계점을 설명하며 상세한 OT 탐지 알고리즘 절차 및 모의 자료를 이용한 알고리즘 결과가 포함되어있다.

1.2 사용자

이 문서의 대상은 OT 탐지 알고리즘의 물리적 개념에 대한 이해가 필요하거나 OT 탐지 결과를 악기상 자료와 연계하는 등의 연구에 활용하기를 원하는 사용자를 위해 작성되었으며, 알고리즘을 유지 보수하거나 수정, 개선하고자 하는 사용자에게도 유용한 정보를 제공한다.

1.3 내용

이 문서는 아래와 같이 총 5개 부분으로 나뉜다.

- 관측시스템 개요: 알고리즘 산출물에 대한 간략한 설명
- 알고리즘 소개: 알고리즘의 이론적 개념, 입력 및 출력자료를 포함한 OT 탐지 알고리즘에 대한 상세한 설명
- 시험자료 및 출력물: 알고리즘 성능과 산출물 품질의 특성을 결정짓는 테스트 자료에 대해 기술하고 Himawari-8 자료를 이용해 처리된 알고리즘 결과에 대해 기술
- 현업 운영시 고려사항: 수치적인 계산, 프로그래밍과 절차, 품질평가, 진단 및 예외처리와 관련된 이슈에 대한 종합적인 기술
- 가정 및 제한: 알고리즘이 기반하고 있는 이론 관련 모든 가정과 제약사항에 대해 기술 및 토의

1.4 관련 문서

- 성층권 침투 대류운과 관련된 참고문헌
- GOES-16 성층권 침투 대류운 탐지 ATBD

2. 관측시스템 개요

2.1 산출물

OT 탐지 알고리즘의 산출물은 OT의 존재 여부에 대하여 yes/no의 이진 값으로 객체형태로 제공된다. OT 탐지 알고리즘은 동아시아 확장지역(**그림 1**)에 대해 구축되었으며, 탐지된 OT 화소들에 대한 위치정보를 2 km 공간해상도로 제공한다.

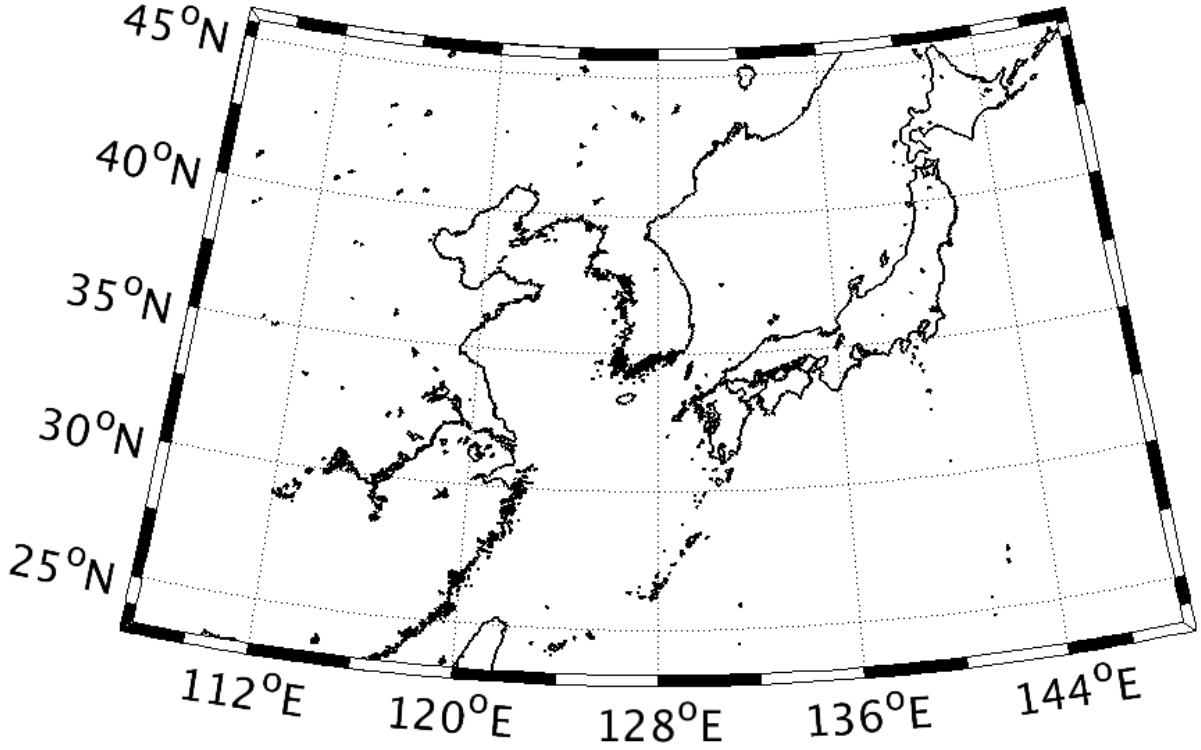

그림 1. 성층권 침투 대류운 탐지 대상 지역 (동아시아 확장지역, 22.4 ~ 46.4 ° N, 107.3 ~ 148.1 ° E)

2.2 센서특성

OT 탐지 알고리즘은 GK-2A에 탑재될 AMI의 BT11.2의 적외선 채널 영상을 사용한다(**표 1**). AMI는 가시, 근적외, 적외 영역을 포함하여 총 16개의 밴드를 가진다. Full-disk 스캔에 소요되는 시간은 10분이며 10개의 적외 채널에 대해서 2 km의 공간해상도를 가진다.

표 1. AMI 채널 규격 및 OT 알고리즘에서 사용되는 채널

채널 종류	범위(㎛)	중심파장(㎛)	FOV(km)	OT 탐지 알고리즘에 사용여부
가시	0.45 – 0.49	0.47	1	
	0.59 – 0.69	0.64	0.5	○
근적외	0.84 – 0.88	0.86	1	
	1.365 – 1.395	1.38	2	
	1.58 – 1.64	1.61	2	
	2.235 – 2.285	2.26	2	
적외	3.80 – 4.00	3.9	2	
	5.77 – 6.60	6.19	2	○
	6.75 – 7.15	6.95	2	
	7.24 – 7.44	7.34	2	
	8.30 – 8.70	8.5	2	○
	9.42 – 9.80	9.61	2	
	10.1 – 10.6	10.35	2	○
	10.8 – 11.6	11.2	2	○
	11.8 – 12.8	12.3	2	○
	13.0 – 13.6	13.3	2	○

3. 알고리즘 소개

3.1 개요

OT 탐지 알고리즘은 AMI로부터 아래의 자료를 이용하여 OT 발생 여부를 산출한다.

- BT(밝기온도, Brightness temperature)6.2, BT8.6, BT10.4, BT11.2, BT12.4, BT13.3 적외선 채널의 영상

OT 탐지 알고리즘이 탐지한 OT 지역으로부터 아래의 산출물을 생산한다.

- 탐지된 OT 화소들에 대한 위치정보

3.2 처리 흐름

OT 탐지 알고리즘은 **그림 2**의 흐름도에 따라 수행된다.

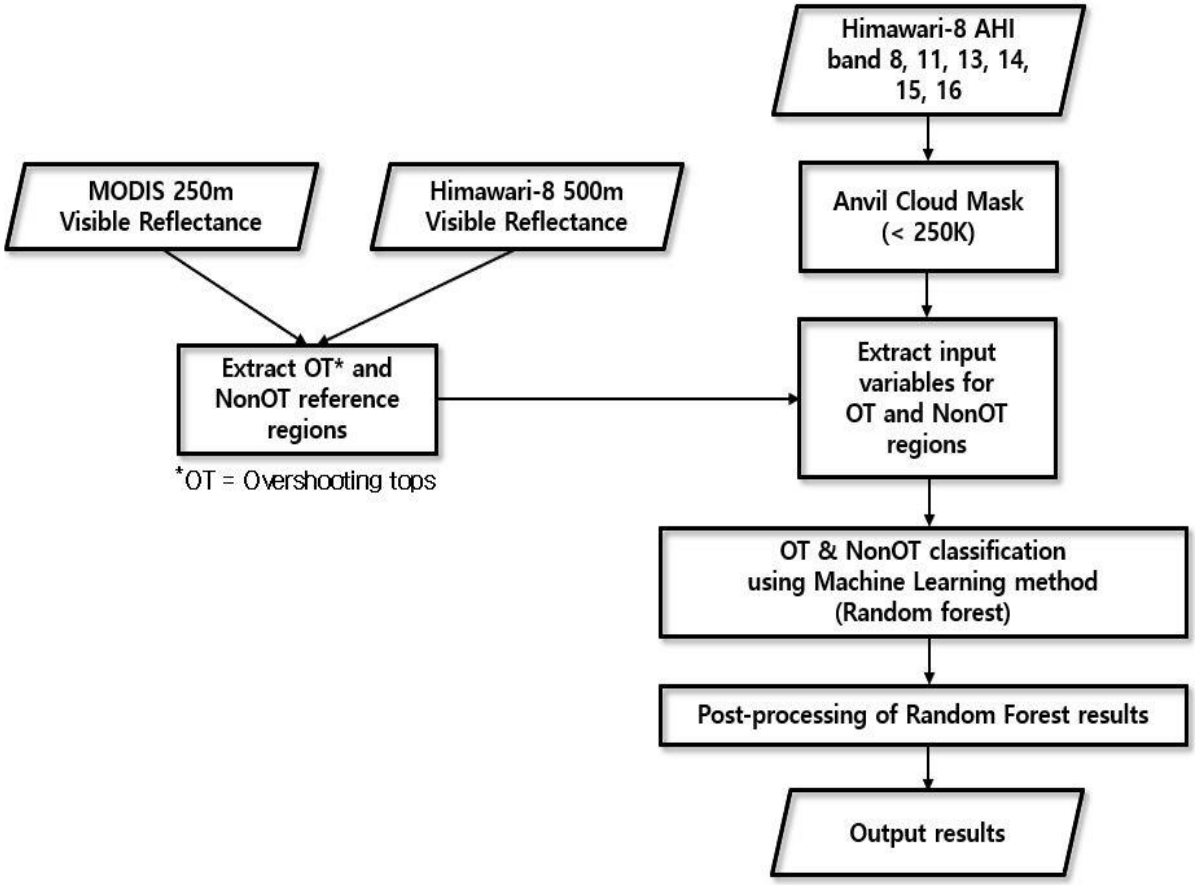

그림 2. 성층권 침투 대류운 탐지 알고리즘 흐름도

3.3 입력자료

3.3.1 주요 센서 자료

OT 탐지 알고리즘에 사용되는 주요 센서 자료는 아래와 같다.

- BT6.2, BT8.6, BT10.4, BT11.2, BT12.4, BT13.3 적외선 채널의 검보정된 영상
- R(반사도, Reflectance)0.64 파장대의 검보정 된 가시 영상

3.3.2 보조자료

OT 탐지 알고리즘 수행을 위해 필요한 보조 자료는 다음과 같다. 이 보조 자료들은 OT 탐지 알고리즘 내에 직접적으로 사용되는 것이 아닌 OT 탐지 모델을 구축할 때 훈련자료를 위한 OT 발생 사례 구축을 위해 사용된다.

- MODIS(Moderate-resolution imaging spectroradiometer)의 가시 채널 반사도 영상 자료
- Himawari-8 AHI(Advanced Himawari Imager)의 가시 채널 반사도 영상 자료

■ Himawari-8 AHI 11.2 ㎛ 채널 밝기온도 자료

성층권 침투 대류운의 대표적인 특징에 해당하는 대류운 표면의 공간적인 특성을 고려하기 위하여 가시영상이 주요 보조 자료로 이용된다. MODIS 와 AHI 의 R0.64 영상을 이용하여 OT 후보지를 설정하며 AHI BT11.2 자료를 기반으로 주변의 밝기온도를 확인하여 OT 후보지역이 주변보다 밝기온도가 낮은 지 확인한다. 낮지 않을 경우 밝기온도를 기반으로 한 위치 재조정 과정이 필요하다.

3.4 이론적 배경

OT 탐지 알고리즘은 대류권계면을 뚫고 성층권 하부로 강한 상승기류를 가진 대류운의 정상을 탐지하는 과정이다. 적외선 채널 영상에서의 OT는 주변의 상대적으로 더 따뜻한 모루운에 비해서 매우 온도가 낮은 밝기 온도를 가지는 지역으로 보이며, 가시 채널 영상에서는 평평한 모루운 위로 돔 형태로 불룩하게 솟은 모습을 보이며 태양천정각이 높으면 그 주변으로 그림자가 생긴다. 또한 OT는 주변 모루운에 비해서 울퉁불퉁한 질감을 가진다(**그림 3**). 강한 상승기류를 가지는 대류운에 의해서 OT가 발생하게 되면 대류권에서 성층권 하부로 수증기가 투입되게 된다. 이론적으로 고도가 높아질수록 온도가 높아지는 성층권 하부에서 위성의 수증기 채널 영상에서는 밝기온도가 높게 나오는 반면 같은 위치에서 적외 채널 영상은 온도가 낮다. 이러한 수증기와 적외 채널의 양수의 차를 이용하여 OT를 탐지하는 기법인 WV-IRW BTD(Water Vapor-InfraRed Window channel Brightness Temperature Differences) 방법이 소개되었고 AVHRR, MODIS, SEVIRI, GOES 등의 다양한 위성영상을 활용하여 많은 연구가 수행되어왔다(Bedka 등, 2010; Setvak 등, 2007). 강한 적란운의 상승기류에 의해 성층권 하부로 대류권의 수증기가 침투하게 되면 수증기 밴드에서의 밝기 온도가 적외선 밴드에 비해 더 높은 값을 가진다. 고도가 증가할수록 온도가 높아지는 성층권 하부에서 수증기는 복사에너지를 방출하여 수증기 채널에서는 적외 채널보다 더 높은 온도를 보인다. 이 차이를 이용하여 성층권 침투 대류운을 탐지하는 방식이 WV-IRW BTD 방법이다. 하지만 이 방법은 종종 흔히 관측되는 OT의 크기보다 상당히 과대 탐지되며 또한 많은 수의 false OT가 탐지되는 한계점을 가진다. 이 기법의 전제는 다음과 같다.

- 성층권 하부에서는 고도에 따라 대기 온도가 증가함.
- 적란운의 상승기류에 의해 수증기가 구름 꼭대기를 넘어서 성층권 하부로 들어감.
- 이 수증기는 따뜻한 성층권 온도에서 방출하는 반면 적외선 채널에서의 방출은 차가운 구름 꼭대기로부터 비롯됨.
- 따뜻한 수증기 밝기온도와 차가운 적외선 채널의 밝기온도의 플러스의 차이가 OT 발생지를 나타냄.

하지만 WV-IRW BTD의 밝기온도 차에 대한 경계값은 센서 자료의 공간해상도, 대류운 상승기류의 강도, 성층권에서 수증기의 지속시간 등에 따라서 달라질 수 있으므로 이를 특정한 값으로 사용하게 되면 정확한 OT를 탐지하는데 한계를 가진다(Setvak 등, 2007; Bedka 등, 2010). BTD의 경계값에 따라서 OT 탐지 결과가 다르게 산출되며 실제 OT 발생 지역에 비해서 과대 탐지를 하는 경향이 있다(**그림 3**). 이에 OT의 반경과 주변 모루운과의 밝기온도 차이 등의 OT의 특성들을 이용한 객관적 OT 탐지 알고

리즘인 IRW-texture 방법이 소개 되었다(Bedka 등, 2010; Bedka, 2011; Bedka 등, 2012). 이 방법은 기존 WV-IRW BTD 방법의 OT 크기의 과대 산정, 수증기의 공간 분포, 입력 센서 자료의 공간 및 분광 해상도로부터의 한계점들을 극복하였다. 하지만 IRW-texture 방법은 OT의 크기와 모루운과 OT 간의 밝기온도 차이 등에 대해서 특정 임계값을 이용하고 있기 때문에 규모가 작은 모루운 내에서 발생하는 OT에 대해서는 탐지가 안 되는 등 다양한 OT 발생 사례를 탐지하는데 있어 한계를 가진다. 따라서 다양한 OT 발생 사례들과 함께 기계학습 기법을 적용하여 OT의 특성에 대한 정보를 추출하고 다양한 조건에서 발생하는 OT들을 탐지할 수 있는 알고리즘이 개발될 필요가 있다. 5차년도 까지 총 네 가지 기계학습 기법 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), SVM(Support Vector Machines)과 Himawari-8 AHI 위성 영상으로부터 다양한 입력 변수들을 이용하여 OT와 non-OT 지역 분류를 통해 OT 탐지를 수행하고 분석하였으며, 최종적으로 타 알고리즘보다 좋은 탐지성능을 가지며, 현업에서의 활용이 용이한 RF 모델을 기반으로 알고리즘을 구성하였다.

3.4.1 입력자료 및 기준자료 구축

OT 탐지 모델은 한반도와 중국 동부, 일본을 포함하는 동아시아 확장지역을 대상으로 구축되었다. RF 기반 OT 탐지 모델을 구축하기 위해 Himawari-8 BT6.4, BT8.6, BT10.4, BT11.2, BT12.4, BT13.3 영상으로부터 총 32가지의 입력 변수들을 추출하였다(**표 2**).

표 2. Random Forest 기반 성층권 침투 대류운 탐지 모델 구축을 위해 사용된 변수

Variable	Meaning	Characteristics
bt (BT11.2 영상에서의 후보 위치 픽셀 값)	Brightness temperature of center of OT	Physical
sw62_112 (BT6.2와 BT11.2의 밝기온도차)	Water vapor content in the upper atmosphere with longwave infrared	
sw62_104 (BT6.2와 BT10.4의 밝기온도차)	Water vapor content in the upper atmosphere with clear longwave infrared	
sw86_112 (BT8.6와 BT11.2의 밝기온도차)	Water vapor content in the middle−upper atmosphere	
sw124_104 (BT12.4와 BT10.4의 밝기온도차)	IR window with low−impact of water vapor	
sw124_112 (BT12.4와 BT11.2의 밝기온도차)	Cirrus detection	
std3x3, 5x5, 7x7, 9x9, 11x11 at BT11.2, BT13.3 (sub window 내의 밝기온도의 표준편차)	Roughness of cloud top	Spatial
diff3x3, 5x5, 7x7, 9x9, 11x11 at BT11.2, BT13.3 (sub window 내의 중심픽셀과 주변픽셀 평균간의 밝기온도 차이)	Lumpy texture of cloud top	
Correaltion ratios between 5x5 sub window and ideal OT shape (R0.64 가시영상기반 41x41 sub window와 OT 모형간의 2차원 상관비)	Spatial correlation ratio with ideal OT shape	
latitude	Correction to the difference of value according to the variation of the zenith angle	

성층권 침투 대류운 탐지 알고리즘 구축 위해 모의자료인 Himawari-8 AHI 자료가 이용 가능한 2015년 7월부터 시작하여 2017년 12월까지 성층권 침투 대류운 발생 사례 구축을 시행하였다. 성층권 침투 대류운 발생 사례는 **그림 3**와 같이 MODIS 혹은 Himawari-8의 가시영상을 기반으로 대류운 내에서 돔 모양으로 튀어나온 부분이 관찰되며, 주변에 모루운이 발생한 경우를 기준으로 육안판독을 통해 후보지가 구축되었다. 사례 구축의 객관화를 위해 구축된 후보지(**그림 3** (d)의 노란색 동그라미)의 밝기온도가 주변 3*3 픽셀 내에서 가장 낮은 온도가 아닐 경우 가장 낮은 온도를 가지는 곳(**그림 3** (d)의 빨간색 동그라미)으로 성층권 침투 대류운의 위치를 조정하여 사례를 구축하였다. 이렇게 구축된 성층권 침투 대류운 발생 탐지 사례는 2015년 7월부터 2017년 12월에 대해 총 1311 개의 샘플을 취득하였다. 이 중 80%는 모델의 훈련자료로 사용하였고 나머지 20%는 구축된 모델 평가에 이용되었다. 알고리즘 0.5.0.0 버전은 Himawari-8 AHI 영상 기반의 샘플만을 이용하여 모델을 구축하였다. 향후 GK2A AMI의 영상이 충분히 확보된 후 모의자료를 AMI 기반 데이터로 치환하여, AMI에 최적화된 알고리즘을 제공할 예정이며, 더 좋은 탐지 성능을 보일 것으로 기대한다.

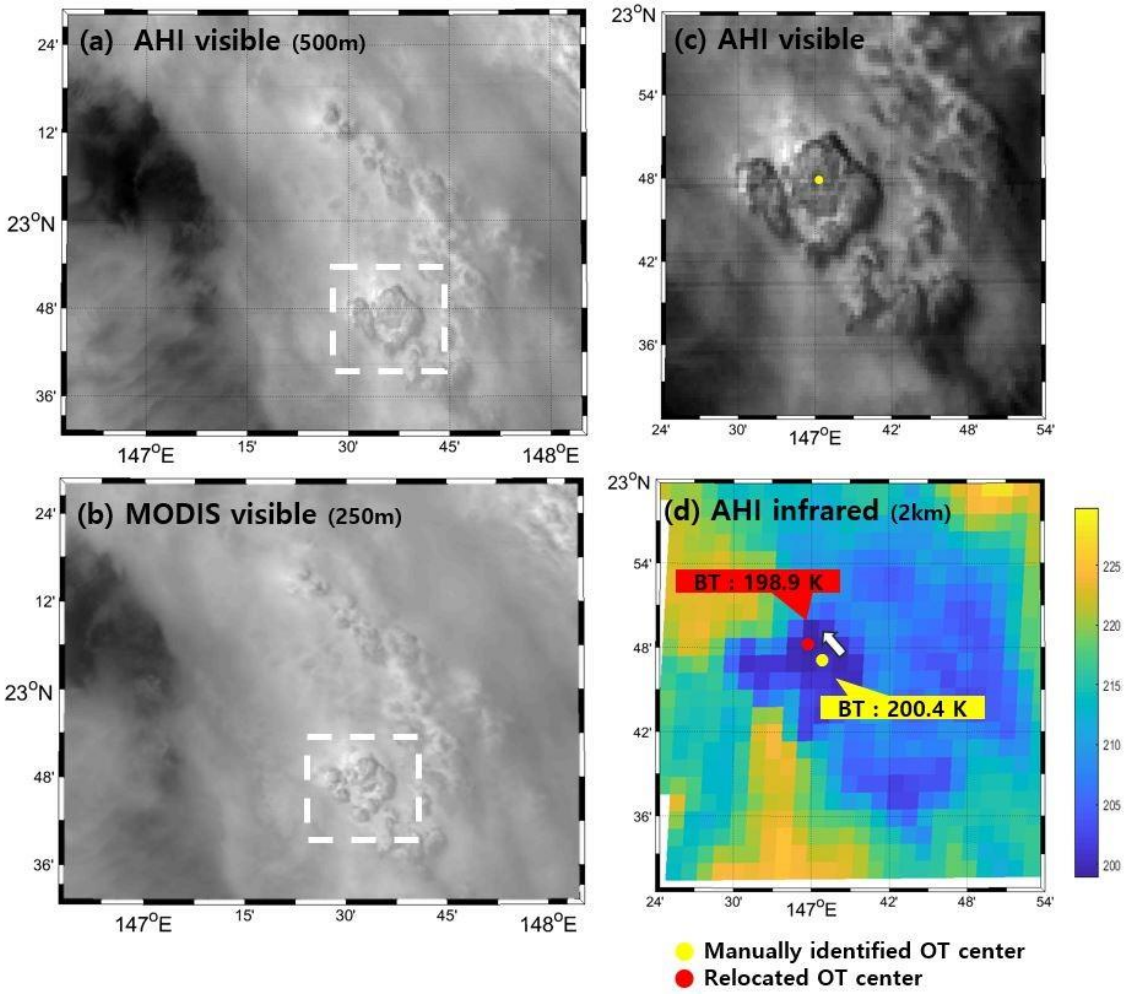

그림 3. Himawari-8 AHI의 R0.64 가시영상 자료(a)와 Aqua MODIS 0.64 ㎛ 채널의

가시영상 자료(b)를 이용한 성층권 침투 대류운 발생 후보지역 선정 및 밝기온도 기반 자동적 위치 재조정(d)을 통한 성층권 침투 대류운 기준 자료 구축의 예시

3.4.2 기계학습 기반 성층권 침투 대류운 탐지 모델 구축

본 알고리즘에서는 기존 연구(Han et al., 2015; Kim et al., 2017)를 통해 대류운 초기 탐지 및 성층권 침투 대류운 탐지에 유의미하다는 것이 증명된 랜덤포레스트 모델(Random forest, RF)을 이용하였다. 랜덤포레스트 알고리즘은 걸정나무 기반 알고리즘인 CART(Clssification and regression trees)를 확률적으로 접근한 알고리즘으로(Breinman et al., 2001), CART는 훈련자료를 기반으로 하여 타겟 변수를 위한 규칙으로 구성된 결정나무를 구성하는 알고리즘이다. RF 알고리즘은 변수 추출과 훈련자료 추출 과정에 있어 무작위 과정(randomization)을 거쳐 하나의 결정나무가 아닌 수백~수천개 이상의 결정나무를 구성한다. 이를 기반으로 과반투표(majority voting) 혹은 사용자 지정 알고리즘을 통해 결과를 예측하는 알고리즘이다. **그림 4**는 RF 알고리즘의 구성을 보여준다.

랜덤 포레스트 모델은 알고리즘을 구성하는 변수들의 변수중요도를 제공한다. 최적화 알고리즘을 구성하기 위한 무작위 과정을 반복 할 때, 각 변수를 leave-one-out 방식으로 치환해 가며 결정나무 알고리즘을 구성하고, 결정나무의 결과를 검증하는 방식으로 각 변수의 변수중요도를 결정한다. 이 때, 각 변수에 대해 중요도 점수(importance score)가 각각 기록되며, 각 중요도 점수는 표준 오류(Standard error)로 나누어 준다. 이 변수는 "Mean Decrease Accuracy"로 제공된다. 또한 랜덤포레스트 알고리즘은 "Mean Decrease Gini"지수도 제공하는데, 이는 랜덤포레스트 모델을 구성하는 각 결정나무가 변수에 대한 규칙 찾아갈 때 규칙을 통과함으로 인해 줄어드는 엔트로피값을 정량화 하여 나타낸 지수이다. 랜덤포레스트 모델에서는 "Mean Decrease Accuracy"지수와 "Mean Decrease Gini"지수가 높을수록 높은 변수중요도를 가진다.

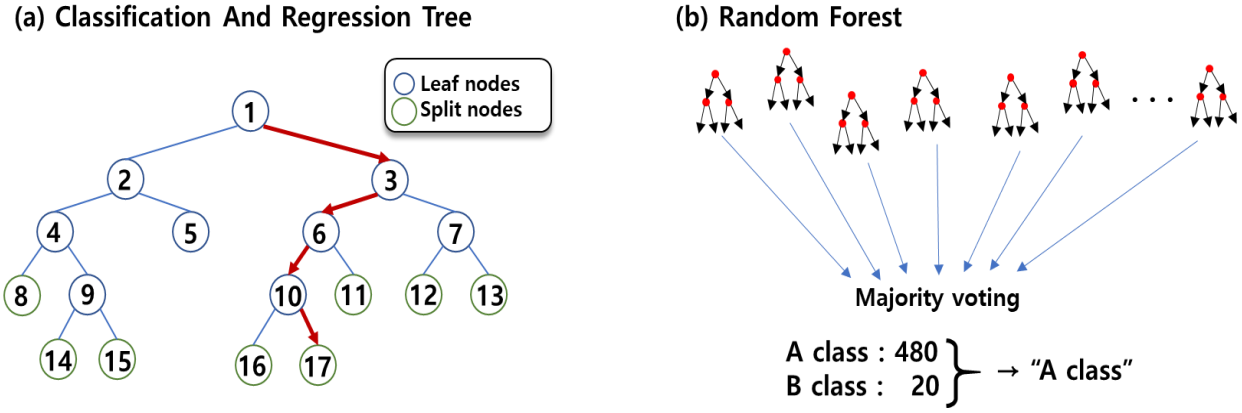

그림 4. (a) CART(Classification and regression trees)와 (b) 랜덤 포레스트(RF: Random forest) 모델의 구성

3.4.3 BT11.2 적외 채널 및 R0.64 가시채널 기반 후처리

일반적인 성층권 침투 대류운에서는 앞서 시행한 기계학습 기반 성층권 침투 대류운 탐지 모델에서 기존 연구 결과에 비해 좋은 분류성능을 보이지만, 권운이나 발생 초기의 대류운과 같은 성층권 침투 대류운과 비슷한 물리적, 공간적 특성을 가지는 지역에 대해서는 오탐지율이 높은 경향을 보인다. 이를 해소하기 위해 본 알고리즘에서는 기계학습 기반으로 탐지된 성층권 침투 대류운 후보지역에 대해 후처리를 시행하여, 결과적으로 실질적인 성층권 침투 대류운의 공간적특성과 유사한 지역을 최종 산출물로 제공한다. **그림 5**는 후처리 과정을 나타낸다. 후처리 과정은 BT11.2 적외 채널을 기반으로 한 후처리와 R0.64 가시채널을 이용한 후처리로 구성되며, 정상적인 가시채널 영상 이용이 불가능한 밤 시간대(1100 ~ 1950 UTC)의 경우 적외 채널 기반 후처리만 적용된다. 각 후처리는 아래와 같이 시행된다.

- BT11.2 적외 채널 기반 후처리: 성층권 침투 대류운 후보지를 기준으로 한 7x7, 9x9, 11x11 sub window 내에서 밝기온도가 낮은 지역의 분포가 성층권 침투 대류운의 공간적 특성을 따르는 지 판별하기 위한 단계이다. sub window 내에서 밝기온도가 낮은 지역의 X 좌표와 Y좌표의 표준편차 차이의 절대값을 정량화하여 이용하고, 이 값이 낮을수록 성층권 침투 대류운의 특성을 따른다는 가정을 기반으로 한다.
- R0.64 가시 채널 기반 후처리: 표면 거칠기를 정량화하여 판별하기 위한 단계이며, 앞서 적외 채널 기반 후처리에서 성층권 침투 대류운의 특성과 유사하다고 판단되어 걸러진 지역에 대해 적용된다. 가시영상에서 후보지역을 기반으로 41x41 sub window를 생성한 후, 영상의 구배 크기(Gradient magnitude)를 계산하여 상위 30%의 구배 크기의 평균값이 클수록 성층권 침투 대류운의 특성을 따른다는 가정을 기반으로 한다.

후처리 각 과정은 임계 값 기반 알고리즘으로 구성되며, 임계 값은 시험적 결정 방(trial-and-error method)을 통해 지정하였다. GK-2A에 탑재된 알고리즘에서는 후처리에 이용되는 각 임계 값이 매개변수로 지정되어 있어, 향후 경험적 분석을 통해 변경 가능하다.

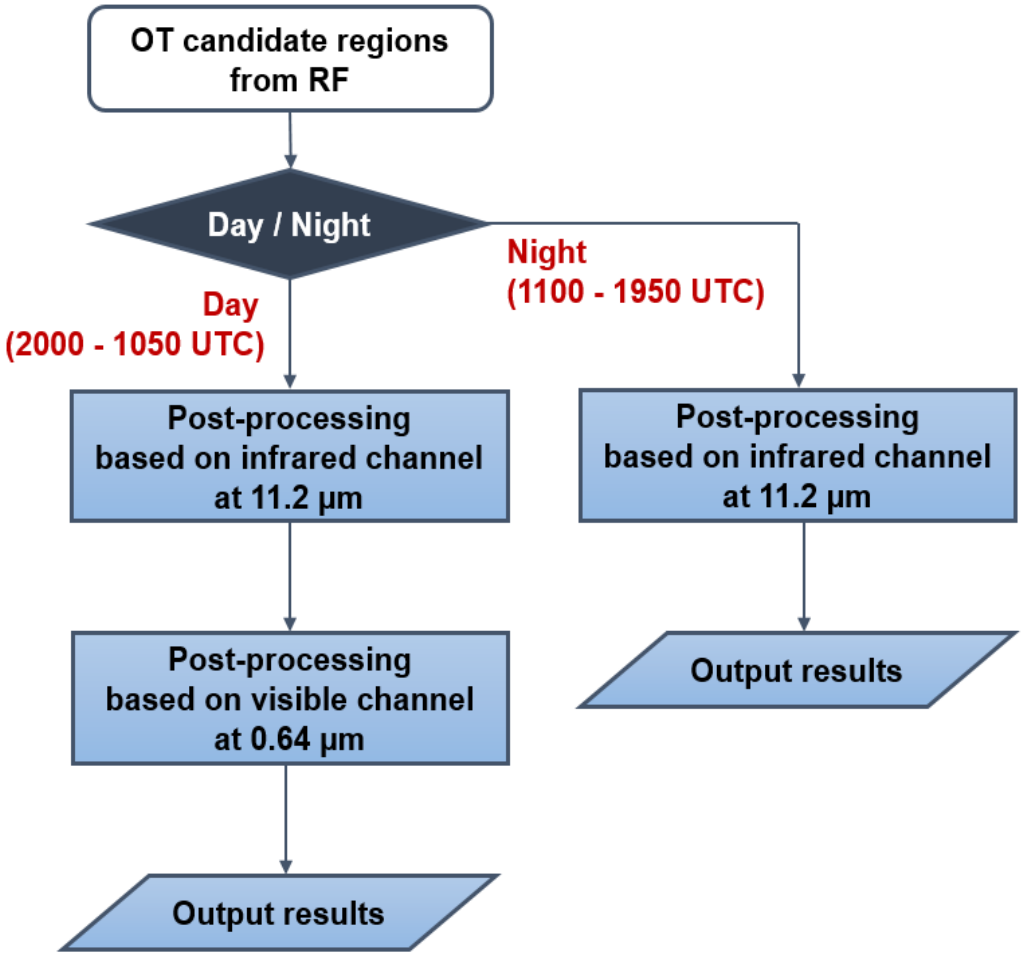

그림 5. BT11.2 적외 채널과 R0.64 가시 채널 기반 후처리 흐름도

4. 시험자료 및 출력물

4.1 모의 입력자료

GK-2A 위성의 영사기인 AMI와 가장 유사한 채널을 보유한 Himawari-8에 탑재된 AHI(Advanced Himawari Imager)로부터의 적외선 채널 영상은 OT 탐지 알고리즘의 주요 모의 자료로 사용된다. AHI를 탑재한 Himawari-8 위성은 동아시아와 서태평양 지역에 대한 기상 현상을 관측하기 위한 기상 관측 위성이다. 총 16개의 밴드를 가지며 그 기능과 사양이 기존의 MTSAT(Multi-functional Transport Satellite) 위성 시리즈들에 탑재된 영사기보다 시간해상도와 정확도 면에서 향상되었다. 본 연구에서는 AHI의 band 8, 11, 13, 14, 15, 16의 BT6.2, BT8.6, BT10.4, BT11.2, BT12.4, BT13.3 영상을 이용하여 다양한 입력 변수들을 추출하여 OT 탐지를 수행하였다.

4.2 모의자료 기반 산출결과

4.2.1 동아시아 확장지역의 Himawari-8 영상을 이용한 성층권 침투 대류운 탐지 결과

4.2.1.1 성층권 침투 대류운 탐지 모델 평가 및 변수 중요도 결과

표 3 는 테스트 자료에 대한 모델의 성능 평가 결과를 보여준다. 본 알고리즘에서는 1차적인 성층권 침투 대류운 후보지역 탐지를 위해 랜덤포레스트 기법을 이용하였다. 랜덤포레스트는 트리 기반의 모델인 결정나무 모델(Decision trees)를 기반으로 확률적으로 접근하여 과 적합 문제를 해소한 기계학습 기법 중 하나이며, 본 모델에서는 전체 평가 정확도가 91.29 %로 유효한 성능 평가 결과를 보였다. 또한 RF 모델에서는 성층권 침투 대류운 픽셀 중 약 23 %의 오탐지율을 보였고, 반면 성층권 침투 대류운이 아닌 지역에 대해서는 단 2 %만이 오분류 되었다. 본 기계학습 모델을 이용하여 성층권 침투 대류운의 후보지를 결정하는데 유효한 성능을 보인다고 판단하였다. **그림 6**에서 확인할 수 있듯이 RF 모델에서는 대기 상층부의 수증기 함량을 의미하는 sw124_104 변수가 가장 중요한 변수중요도를 보였다. 뒤이어 sw62_104, bt, sw86_112 등의 물리적 변수들과 더불어 성층권 침투 대류운의 공간적 특성을 대변하는 diff3x3, mws5_corr5, mws5_corr3 등의 공간적 변수들이 균형적으로 상위의 변수중요도를 보이는 것으로 확인되었다. 이에 본 모델이 구축되는 데 성층권 침투 대류운의 물리적, 공간적 특성이 골고루 고려되었다는 것을 확인할 수 있다 (**표 4**).

표 3. 랜덤 포레스트 모델 평가 결과

Reference / Classified as	OT	Non-OT	Sum	User's accuracy
OT	199	13	212	93.87%
Non-OT	62	587	649	90.45%
Sum	261	600	861	
Producer's accuracy	76.25%	97.83%		
Overall accuracy	91.29%			
Kappa coefficient	0.7823			

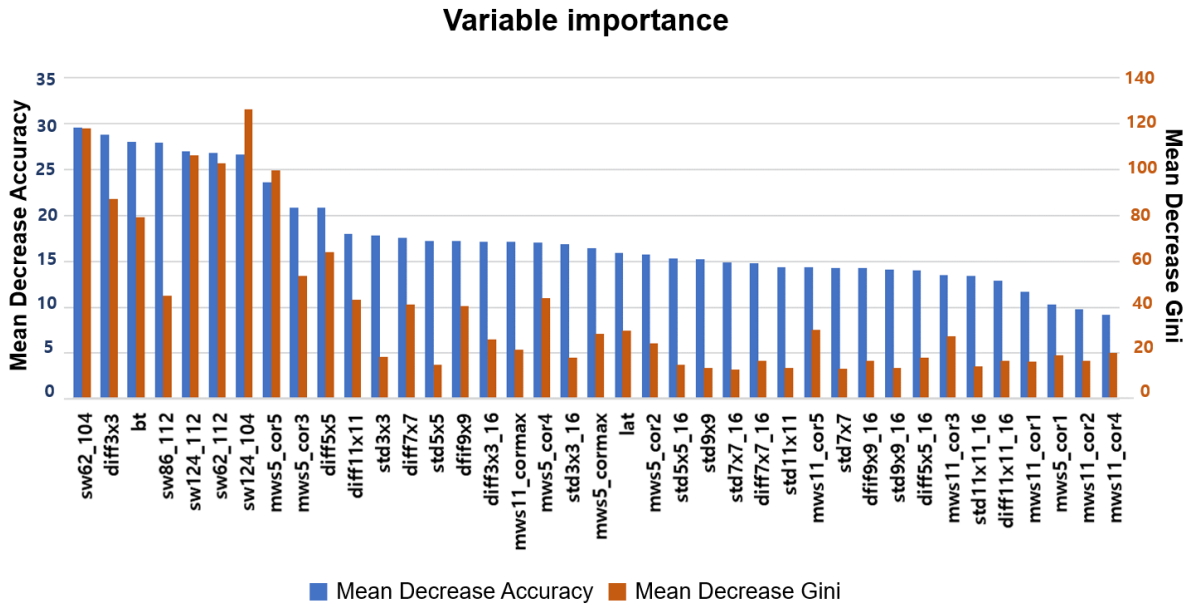

그림 6. 랜덤 포레스트 모델의 변수 중요도 결과

표 4. 랜덤 포레스트 기반 상위 10위까지의 변수중요도를 보이는 변수와 각 변수의 특성

변수 중요도 순서	변수	변수의 특성	
1	sw124_104	수증기의 영향을 배재한 적외선 투과창(IR window)의 값	Physical
2	sw62_104	대기 상층부의 수증기 함유량	Physical
3	diff3x3	3x3 영역에서의 중심과 주변 밝기온도의 차이	Spatial
4	bt	BT11.2	Physical
5	sw86_112	대기 중층부의 수증기 함유량	Physical
6	sw124_112	권운 영역	Physical
7	sw62_112	대기 상층부의 수증기의 함유량	Physical
8	mws5_cor5	후보지의 5x5 영역과 이상적인 OT와의 평면 공간적 상관성	Spatial
9	mws5_cor3	후보지의 5x5 영역과 이상적인 OT와의 평면 공간적 상관성	Spatial
10	diff5x5	5x5 영역에서의 중심과 주변 밝기온도의 차이	Spatial

4.2.1.2 성층권 침투 대류운 발생 분류 매핑 및 정량적 평가 결과

구축된 성층권 침투 대류운 탐지 알고리즘(랜덤포레스트 + 후처리)은 모델 구축이나 평가에 사용되지 않은 2019년 6월 15일에서 16일까지의 GK2A AMI 동아시아 확

장지역 영상들에 대해서 적용되어 POD(Probability Of Detection)와 FAR(False Alarm Ratio)를 계산하였다(**그림 7**). 영상에 따라서 모델의 결과는 POD와 FAR가 각각 83%에서 33%의 가장 낮은 정확도를 가지는 경우부터 92%의 POD와 29% 이내의 FAR를 가지는 좋은 결과를 보여주기도 한다. 본 알고리즘은 해당 기간에 대해서 평균적으로 86.2%의 POD와 27.4%의 FAR를 보여준다. 본 결과는 알고리즘의 목표 정확도인 POD 25% 이상, FAR 75% 이하 (**표 5**)를 만족하는 결과를 보이며, 4차 년도 모델인 트리융합 (Aggregated decision trees, ADT)기반 성층권 침투 탐지 모델보다 나은 성능을 보여준다. **그림 8**는 2019년 6월 22일 00:00 UTC, 2019년 8월 8일 01:00 UTC 그리고 2019년 8월 10일 02:00 UTC의 GK2A AMI 기반 성층권 침투 대류운 탐지 결과를 보여준다. 성층권 침투 대류운의 발생 초기부터 소멸 시기까지의 연속적인 객체 탐지가 이루어지는 것을 확인 할 수 있다.

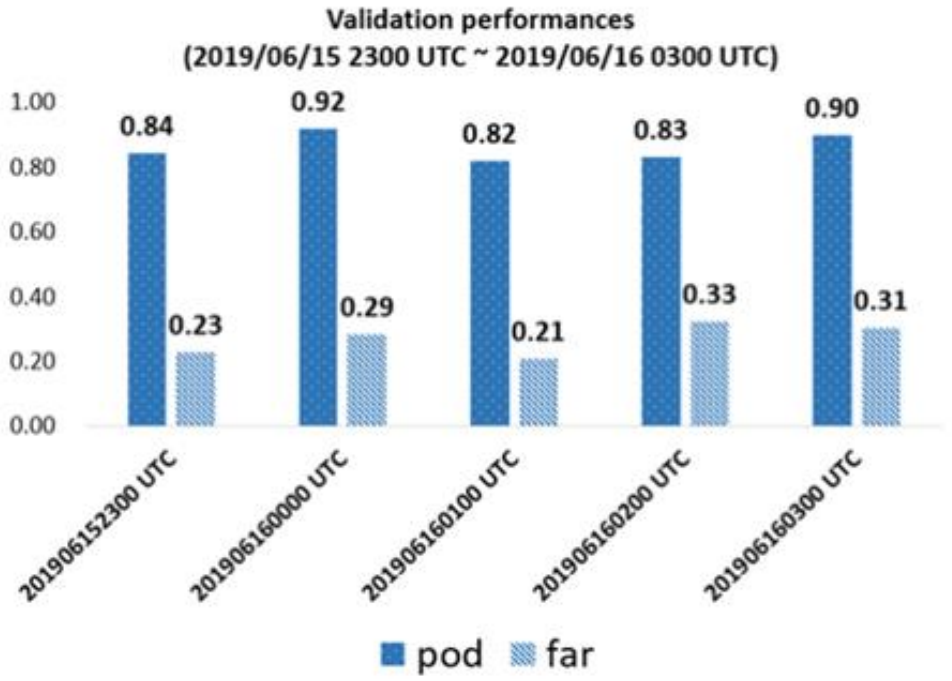

그림 7. 2019년 6월 15일 23:00 UTC 부터 2019년 6월 16일 03:00 UTC 까지의 동아시아 확장지역에 대한 성층권 침투 대류운 탐지 알고리즘의 POD(Probability of detection)와 FAR(False alarm rate). 알고리즘의 목표 정확도는 POD 25 % 이상, FAR 75 % 이하임. (각 산출물의 시점은 YYYYMMDDhhmm UTC로 표현))

표 5 GK-2A 성층권 침투 대류운 탐지 알고리즘의 목표정확도. 본 목표정확도는 GOES-16 OT 알고리즘의 목표정확도와 동일.

	POD (Probability of detection)	FAR (False alarm rate)
목표정확도	25 % 이상	75 % 이하

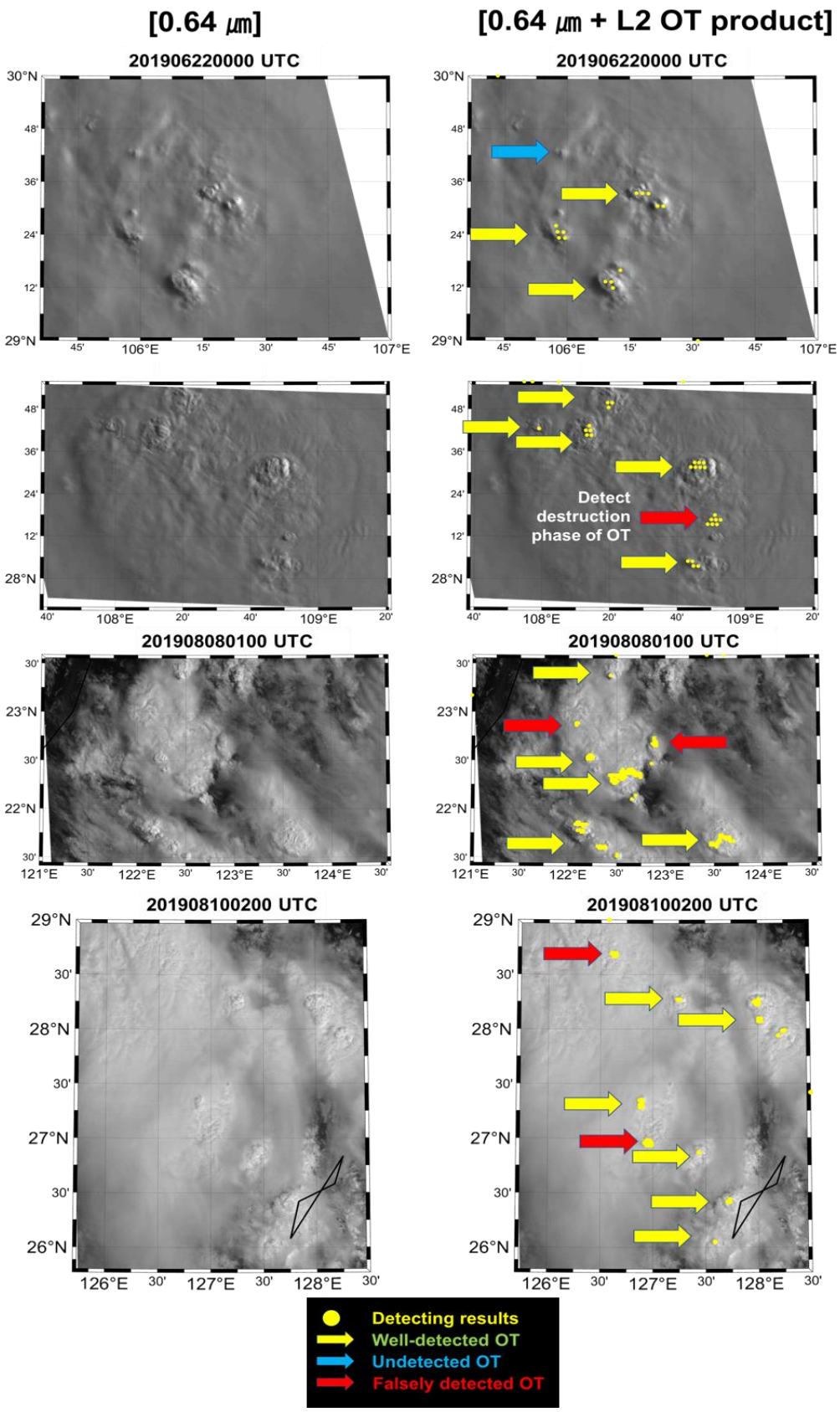

그림 8. 2019년 6월 22일 00:00 UTC, 2019년 8월 8일 01:00 UTC 그리고 2019년 8월 10일 02:00 UTC의 GK2A AMI 영상에 대한 탐지 결과 (노란색 동그라미: 모델 탐지 지

역, 노란색 화살표: 잘 탐지된 OT, 파란색 화살표: 미탐지 OT, 빨간색 화살표: 오탐지 지역, 각 영상의 시점은 YYYYMMDDhhmm UTC로 표현)

5. 협업운영시 고려사항

5.1 수치계산 고려사항

해당사항 없음.

5.2 프로그래밍 절차상의 고려사항

해당사항 없음.

5.3 품질평가 및 진단

GK-2A AMI 센서의 가시영상채널이 가용하지 않을 경우 후처리 과정이 완전하지 않으므로 알고리즘 성능에 영향을 준다.

5.4 예외사항 처리

성층권 침투 대류운 탐지의 예외사항 대처 방안은 **표 6**에 요약 되어있다. 보조자료에 대한 대응책으로는 원하는 지역에 대한 MODIS 자료가 가용하지 않을 경우, Himawari-8 AHI 가시채널 기반 육안 판독과 적외 채널 기반 밝기온도 확인을 통하여 구축된 OT 발생 사례를 기준 혹은 검증 자료로 이용하고자 한다. 또한 AMI의 가시채널의 영상이 가용하지 않을 경우 적외 채널 기반 후처리만 적용되며, 이 경우 알고리즘의 성능에 영향을 준다.

표 6. 성층권 침투 대류운 탐지 예외사항 대처 방안

위험요소 분류	위험 요소
보조자료	• 원하는 지역의 MODIS 자료가 가용하지 않는 경우
보조자료	• 가시채널이 가용하지 않을 경우 알고리즘 성능에 영향을 줌

5.5 알고리즘 검증

5.3 와 동일.

6. 가정 및 제한

6.1 성능

OT 탐지 알고리즘의 성능을 추정하기 위해서 처리 시스템은 다중 화소들을 한

번에 처리가 가능하며 알고리즘은 화소들의 공간적인 통계 값과 표면 거칠기 정보를 추출할 수 있다는 가정을 포함한다.

6.2 예상되는 센서성능

AMI 센서가 현재 사양을 만족하지 못하여 적외 채널 영상이 고해상도로 제공되지 못하면 하나의 적외 채널 영상만을 이용해서 수행되는 OT 탐지 알고리즘의 결과에 영향을 미칠 수 있으므로 OT 탐지 알고리즘에 필요한 적외 채널 영상이 이용 불가능한 경우에 대비하여 다른 적외 채널 영상을 사용할 수 있을 것이다.

6.3 산출물 개선을 위한 사전계획

OT 탐지 알고리즘 개선을 지속하면서 아래의 안건에 집중하여 산출물을 개선시키고자 한다.

6.3.1 성층권 침투 대류운 발생 사례 구축

OT 탐지 모델은 기계학습 기법을 기반으로 만들어지기 때문에 다양한 OT 발생 사례에 대해서 많은 훈련 자료를 구축하는 것이 중요하다. 따라서 OT 기준 자료 구축에 사용 될 수 있는 MODIS와 Himawari-8 AHI 센서의 영상을 이용하여 동아시아 지역에 대한 OT 발생 사례를 지속적으로 구축하여 알고리즘을 개선할 예정이다.

6.3.2 후처리 알고리즘 임계 값 조정

후처리 알고리즘에 이용되는 공간변수에 대한 임계 값은 trial-and-error 방법을 통해 경험적으로 추출한 값으로, 향후 알고리즘 운영 환경 및 대기환경 변화에 따라 임계 값을 조정할 수 있다.

7. 참고문헌

Bedka, K., Brunner, J., Dworak, R., Feltz, W., Otkin, J., & Greenwald, T. (2010). Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients. Journal of Applied Meteorology and Climatology, 49(2), 181-202.

Bedka, K. M. (2011). Overshooting cloud top detections using MSG SEVIRI Infrared brightness temperatures and their relationship to severe weather over Europe. Atmospheric Research, 99(2), 175-189.

Bedka, K. M., Dworak, R., Brunner, J., & Feltz, W. (2012). Validation of satellite-based objective overshooting cloud-top detection methods using CloudSat cloud profiling radar observations. Journal of Applied Meteorology and Climatology, 51(10), 1811-1822.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46.

Han, H., Lee, S., Im, J., Kim, M., Lee, M. I., Ahn, M. H., & Chung, S. R. (2015). Detection of convective initiation using Meteorological Imager onboard Communication, Ocean, and Meteorological Satellite based on machine learning approaches. Remote Sensing, 7(7), 9184-9204.

Ho, T. K. (1995, August). Random decision forests. In Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on (Vol. 1, pp. 278-282). IEEE.

Inoue, T. (1987). A cloud type classification with NOAA 7 split-window measurements. Journal of Geophysical Research: Atmospheres, 92(D4), 3991-4000.

Kim, M., Im, J., Park, H., Park, S., Lee, M. I., & Ahn, M. H. (2017). Detection of Tropical Overshooting Cloud Tops Using Himawari-8 Imagery. Remote Sensing, 9(7), 685.

Mikuš, P., & Mahović, N. S. (2013). Satellite-based overshooting top detection methods and an analysis of correlated weather conditions. Atmospheric research, 123, 268-280.

Schmetz, J., Tjemkes, S. A., Gube, M., & Van de Berg, L. (1997). Monitoring deep convection and convective overshooting with METEOSAT. Advances in Space

Research, 19(3), 433-441.

Setvak, M., Rabin, R. M., & Wang, P. K. (2007). Contribution of the MODIS instrument to observations of deep convective storms and stratospheric moisture detection in GOES and MSG imagery. Atmospheric research, 83(2), 505-518.

Walker, S. H., & Duncan, D. B. (1967). Estimation of the probability of an event as a function of several independent variables. Biometrika, 54(1-2), 167-179.