

GK-2A AMI

Algorithm Theoretical Basis

Document

착빙

개발자 정보

심성문, 울산과학기술원

임정호, 울산과학기술원

박수민, 울산과학기술원

Version 1.1

2019. 11. 15.



국가기상위성센터
National Meteorological Satellite Center

<수정 이력>

문서 버전	수정내용	과학코드 버전	수정일	작성자
0.0	—	0.1.0.0	2016.01.17	심성문, 박수민, 임정호
0.1	전 내용	0.2.0.0	2017.01.15	심성문, 박수민, 임정호
0.2	전 내용	0.3.0.0	2017.09.29	심성문, 박수민, 임정호
0.3	PIREPs 품질관리 추가	0.5.1.0	2018.11.20	심성문, 박수민, 임정호
1.0	AHI-AMI 변환에 따른 세부 내용 수정	0.5.1.0	2019.04.26	심성문, 임정호
1.1	전 내용	0.5.1.0	2019.11.15	심성문, 임정호

<목차>

1. 개요	1
1.1 목적	1
1.2 사용자	1
1.3 내용	1
1.4 관련 문서	2
2. 알고리즘 소개	3
2.1 개요	3
2.2 알고리즘 흐름도	6
2.3 입력자료	6
2.3.1 L1B	6
2.3.2 L2	7
2.3.3 NWP	7
2.3.3 보조 자료	7
2.4 이론적 배경	10
2.4.1 물리적 문제	11
2.4.2 착빙발생 마스크 구축	11
2.4.3 랜덤포레스트	11
2.5 수학적 설명	12
2.5.1 착빙발생 마스크 알고리즘	13
2.5.2 과냉각수적 후처리	14
2.5.3 착빙발생 강도 알고리즘	15
2.5.4 착빙발생 고도 알고리즘	15
2.6. 산출물	15
3. 모의자료 및 검증	16
3.1 모의 입력자료 및 검증자료	16
3.2 검증 방법	16
3.3 검증 결과	16
4. 현업운영 시 고려사항	20
4.1 수치계산 고려사항	20

4.2 프로그래밍 및 절차상의 고려사항	20
4.3 품질 평가 및 진단	20
4.4 예외사항 처리	20
4.5 검증 알고리즘	20
5. 가정 및 제한.....	22
5.1 성능	22
5.2 예상 센서 성능	22
5.3 개선을 위한 사전 계획	22
6. 참고문헌	23

<표 목차>

표 1. AMI와 AHI의 중심파장 및 채널 특성	4
표 2. 착빙 발생 PIREPs 사례에 대한 품질관리	9
표 3. 착빙 미발생 PIREPs 사례에 대한 품질관리	10
표 4. 분류정확도 평가 테이블 구성.....	17
표 5. 검증사례들을 이용한 AMI 랜덤포레스트 모델 평가 결과.....	17

<그림 목차>

그림 1. ELA 산출 영역 ($22.4 - 46.5^{\circ}$ N, $107.3 - 148.1^{\circ}$ E).....	5
그림 2. 착빙 알고리즘 처리 흐름도.....	6
그림 3. (a) CART (Classification and regression trees)와 (b) 랜덤포레스트 (Random forest) 모델의 구성.....	12
그림 4. 랜덤포레스트 AMI 모델 입력 변수들의 상대적 변수중요도. (a) 낮시간 모델, (b) 밤시간 모델.	14
그림 6. 2018.09.16. 08:50-09:10 UTC에 대한 AHI 기반 착빙강도 결과. 미착빙 픽셀은 채널13의 밝기온도로 대체되었음.....	18
그림 7. 2017.01.28. 01:40 UTC 시기의 AHI와 UM768 자료를 이용한 착빙발생 고도 이미지. 청색: 채널13의 밝기온도, 황색: 낮은 고도 착빙, 적색: 높은 고도 착빙...	19

약어표

ABI	Advanced Baseline Imager
AMI	Advanced Meteorological Imager
ATBD	Algorithm Theoretical Basis Document
CART	Classification And Regression Tree
Ch	Channel
COMS	Communication, Ocean and Meteorological Satellite
CPH	Cloud Phase
CTH	Cloud Top Height
CTPS	Cloud Top Properties
DCOEW	Daytime Cloud Optical thickness, Effective radius, Water path
ELA	Extended Local Area
F	False
FA	Flight Altitude
FIT	Flight Icing Threat
GK-2A	Geo-Kompsat-2A
H	Hit
HKO	Hong Kong Observatory
LGT	Light
M	Miss
MI	Meteorological Imager
MOG	Moderate Or Greater
MWIR	Mid-wave infrared
N	Correct negative
NWP	Numerical Weather Prediction
PIREPs	Pilot reports
POD	Probability Of Detection
POFD	Probability Of False Detection
QC	Quality Control
RF	Random Forest
SBAF	Spectral Band Adjustment Factors
SWIR	Shortwave Infrared
SLW	Super-cooled Liquid Water
UM	Unified Model

1. 개요

1.1 목적

작빙(ICING) 알고리즘 기술문서(Algorithm Theoretical Basis Document)는 2018년에 발사된 후속 정지궤도위성인 GK-2A(GeoKOMPSAT-2A)에 탑재될 AMI(Advanced Meteorological Imager)의 산출물들을 이용하여 작빙발생 탐지를 하기 위한 것이다. 작빙 산출물은 작빙 발생지역의 발생강도, 발생고도를 제공한다.

1.2 사용자

이 문서는 알고리즘의 물리적 개념에 대한 이해가 필요하거나 알고리즘의 결과를 활용 및 연구하고자 하는 사용자를 위해 작성되었으며, 특히 항공기 운항 관련 종사자를 위하여 작성되었다. 이는 또한 알고리즘을 유지 보수하거나 수정, 개선하고자 하는 사용자에게도 유용한 정보를 제공한다.

1.3 내용

본 문서는 아래와 같은 목차로 이루어져 있다.

- 관측시스템 개요: 알고리즘 산출물에 대한 간단한 설명
- 알고리즘 소개: 알고리즘의 이론적 의미, 입력 및 출력자료를 포함한 작빙 알고리즘에 대한 상세한 설명
- 시험자료 및 출력물: 알고리즘 성능과 산출물의 특성을 결정짓는 테스트자료에 대해 기술하고 AMI 자료를 통해 생성된 알고리즘에 대한 설명
- 현업 운영 시 고려사항: 수치적인 계산, 프로그래밍과 절차, 품질평가, 진단 및 예외처리와 관련된 이슈에 대한 종합적인 기술
- 가정 및 제한: 알고리즘과 관련된 모든 가정과 제약사항에 대해 기술 및 토의관측

1.4 관련 문서

- 대기항공 알고리즘 개발 제안서
- GOES-R(Geostationary Operational Environmental Satellite-R) Flight Icing Threat ATBD
- GOES-R ABI 가이드

2. 알고리즘 소개

2.1 개요

2.1.1 관측시스템 개요

2.1.1.1 산출물 특성

GK-2A 착빙 알고리즘은 AMI 기본 산출물과 구름산출물 정보를 이용하여, 항공기 착빙 발생 환경을 탐지 후, 착빙 발생 유·무를 나타내는 수평적 마스크를 기반으로, 구름분석 산출물을 이용하여 착빙발생 강도, 온·습도 프로파일을 통해 착빙발생 고도를 산출한다. 착빙발생 유·무는 L1B(Level-1B) 자료와 L2(Level-2) 산출물 중 CTPS(Cloud Top Properties) 알고리즘의 구름상(CPH: Cloud Phase)산출물을 입력자료로 하고, 기계학습의 일종인 랜덤포레스트(RF: Random Forest) 방법을 거친 후에 13번 채널의 밝기온도를 이용한 과냉각수적 후처리를 거쳐 산출이 된다. 입력자료로 사용되는 광학채널들의 밤·낮 가용성이 달라, 낮시간(09시-18시 KST)과 밤시간(18시-다음날 09시 KST)을 구분하여 모델을 구동한다. 출력되는 산출물 중 하나인 착빙 발생 강도는 DCOEW(Daytime Cloud Optical thickness, Effective radius, Water path) 알고리즘의 구름수액량 (Liquid Water Path) 산출물을 입력자료로 하여 미착빙(none), 약한 착빙(LGT: Light), 강한착빙(MOG: Moderate Or Greater)로 구분 후 출력이 되며, 착빙 발생 고도는 수치모델의 고도별 온·습도자료를 사용하여 산출을 하는데, UM(Unified Model)1280의 72개의 수직 프로파일 기준으로, 28번째 (약 11km)에서 41번째 (약 6.5km) 구간에 대한 정보를 제공한다. 산출영역은 ELA (Extended Local Area)이고, 공간해상도는 2km x 2km, 시간해상도는 10분으로 산출한다.

2.1.1.2 센서특성

GK-2A에 탑재될 AMI는 현재 COMS(Communication, Ocean, and Meteorological Satellite) 위성의 MI(Meteorological Imager)에 비해 성능이 매우 향상되었다. 밴드 개수는 가시광선(VIS: Visible), 가시근적외선(VNIR: Visible and Near-Infrared), 근적외선(SWIR: Short-Wavelength Infrared), 중적외선(MWIR: Mid-Wavelength Infrared), 열적외선(TIR: Thermal Infrared) 영역을 포함하여 16개를 제공하며, 기존의 full-disk 스캔에 소요되는 시간 또한 15분에서 10분으로 단축되고, 공간해상도는 2배 향상된 2km가 된다. AMI의 채널 특성과, 모의자료로 사용된 Himawari-8 위성의 AHI(Advanced Himawari Imager)의 채널별 중심파장과 물리적 의미는 표 1과 같다.

표 1. AMI와 AHI의 중심파장 및 채널 특성

AMI 채널	중심파장 (μm)		채널 특성
	AMI	AHI	
01(VIS)	0.47	0.46	낮 시간 구름 탐지 및 두께 특성, 육지 에어로졸
02(VIS)	0.511	0.51	낮 시간 구름 탐지 및 두께 특성, 해색
03(VIS)	0.64	0.64	
04(VNIR)	0.856	0.86	낮 시간 구름 탐지 및 두께 특성, 식생지수
05(SWIR)	1.38	—	
06(SWIR)	1.61	1.6 (ch05)	얼음과 물 입자 특성 구분, 적설 및 운상
NIR	—	2.3 (ch06)	
07(SWIR)	3.83	3.9	저층 구름 정보, 야간 안개, 산불 및 바람벡터
08(MWIR)	6.241	6.2	대류권 상층 수증기 구성
09(MWIR)	6.952	7.0	대류권 중층 수증기 구성
10(MWIR)	7.344	7.3	대류권 하층 수증기 구성
11(TIR)	8.592	8.6	운상, 강우강도, 먼지 및 SO ₂ 탐지
12(TIR)	9.625	9.6	
13(TIR)	10.403	10.4	지표면 및 구름 탐지, 운정온도
14(TIR)	11.212	11.2	구름 및 강우강도, 해수면온도
15(TIR)	12.364	12.3	총 가강수량, 화산재 및 해수면온도
16(TIR)	13.31	13.3	공기 온도, 운정 고도 및 운량

2.1.2 알고리즘 개요

본 알고리즘 구축 연구는 후속 정지궤도 위성인 GK-2A의 착빙 산출물 구축을 기반으로 진행되며, 착빙 강도를 미착빙, 약한착빙, 강한착빙 세 가지로 구분 후 예보를 하여 50% 이상의 POD(Probability Of Detection), 40% 이하의 POFD(Probability Of False Detection)를 가지는 것을 목표로 하고 있으며, 이는 미국의 착빙탐지 알고리즘인 FIT (Flight Icing Threat)과 유사하다 (Smith et al., 2010). 출력물은 2km x 2km의 공간해상도를 가지며 10분마다 ELA 영역에서 제공된다 (그림 1).

AMI에서 산출되는 L1B 자료와 구름상 정보를 모델의 입력 자료로 사용한 랜덤

포레스트 기반의 알고리즘과 13번 채널의 밝기온도를 이용한 과냉각수적 필터링을 통하여 착빙발생 마스크를 산출한다. 그리고 해당 착빙발생 마스크 지역의 구름수액량 자료를 사용하여 착빙발생 강도, 수치모델을 통해 고도를 산출한다.

랜덤포레스트는 수 백 개의 CARTs(Classification And Regression Trees)로 구성되어 있는데, 이는 단일 CART 기반 모델의 과적합 문제를 해결하기 위해 구성된 알고리즘이다. 착빙발생 강도 산출은 착빙 마스크 지역에 대해 구름수액량을 단일 입력 자료로 사용하여 낮시간의 약한착빙과 강한착빙 강도를 산출하고, 착빙발생 고도 산출 알고리즘은 수치모델의 고도별 온·습도자료를 사용하여 고도별 착빙발생위치를 산출하는데, UM1280의 72개의 수직 프로파일 기준으로, 28번째(약 11km)에서 41번째(약 6.5km) 구간에 대한 정보를 제공한다.

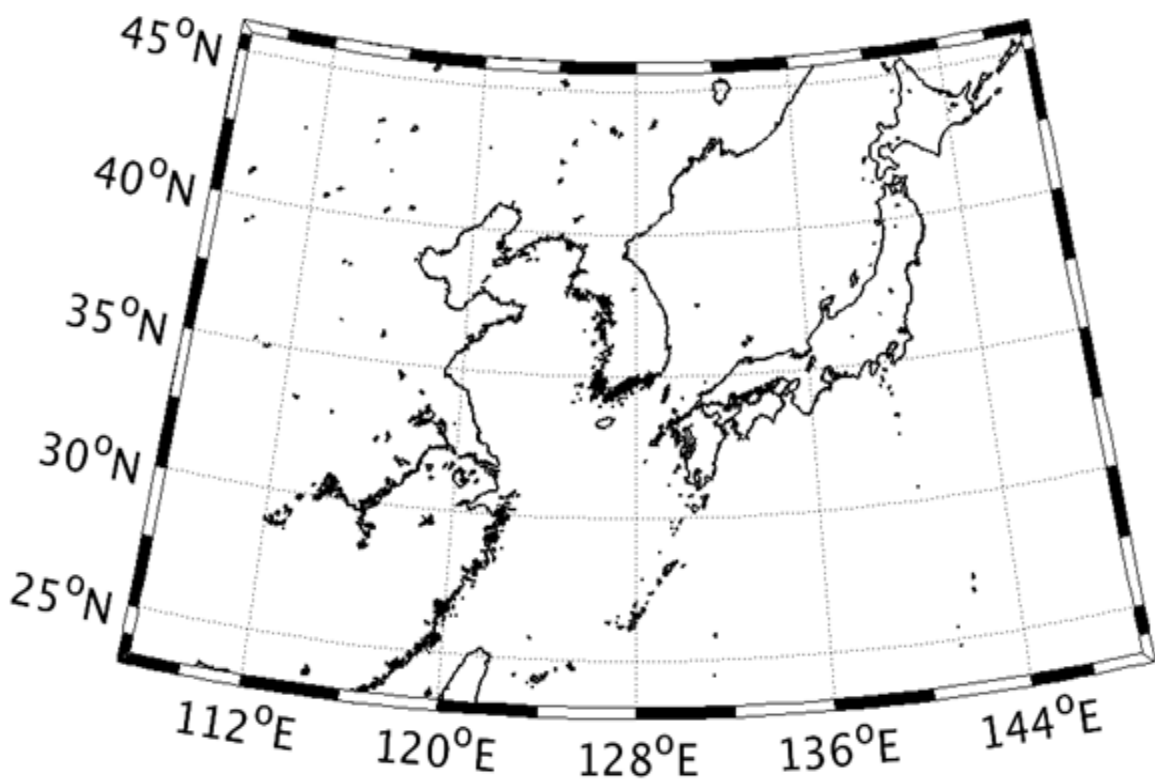


그림 1. ELA 산출 영역 (22.4 - 46.5° N, 107.3 - 148.1 ° E)

2.2 알고리즘 흐름도

전체 알고리즘의 흐름은 그림 2와 같다. 착빙 알고리즘은 낮/밤에 따라 다른 모듈이 구동이 된다.

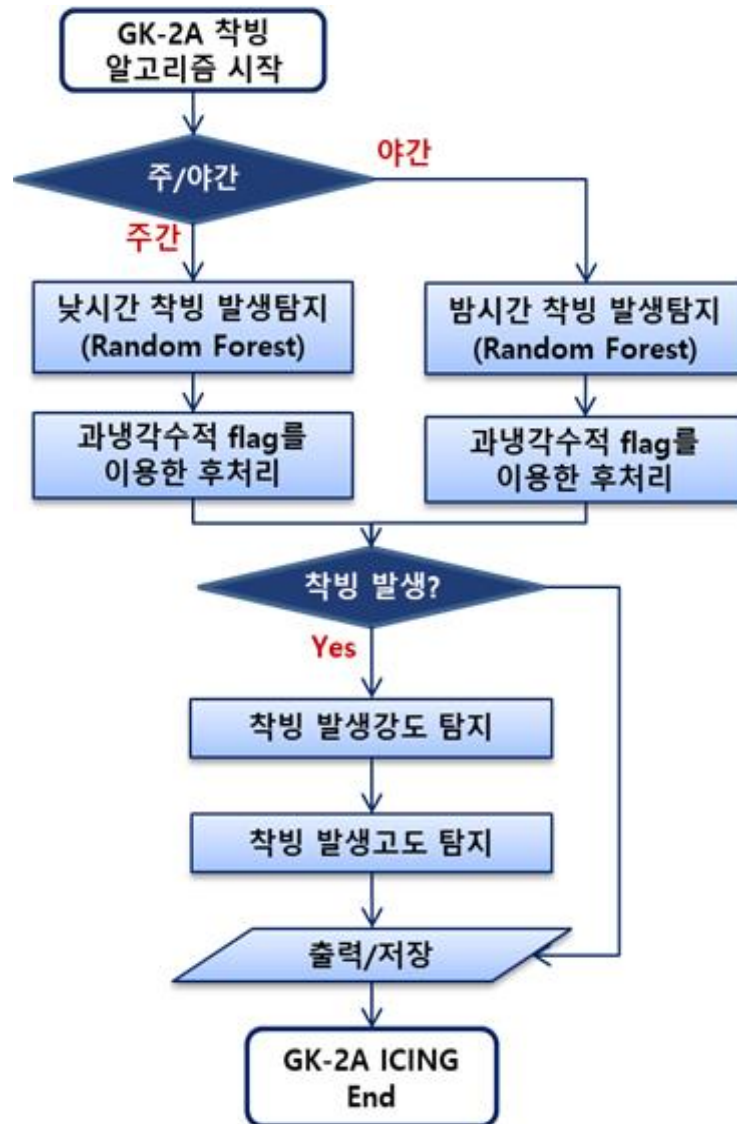


그림 2. 착빙 알고리즘 처리 흐름도

2.3 입력자료

2.3.1 L1B

낮시간 착빙 탐지 모듈에서는 채널 1, 2, 4, 6(AHI기준 Ch5)의 반사도(Reflectance)와 7, 8, 9, 10, 11, 13, 14, 15, 16의 밝기온도(Brightness Temperature)를, 밤시간 착빙 탐지 모듈에서는 채널 7, 8, 9, 10, 11, 13, 14, 15, 16의 밝기온도를 입력자료로 사용한다. 픽셀 기반으로 구동이 되므로, 입력 자료의 해상도는 2km x 2km로 일치 후 사용한다.

2.3.2 L2

2.3.2.1 PIREPs 품질관리

PIREPs 품질관리에서는 CTPS 알고리즘의 구름상, 운정고도를 입력자료로 사용한다.

2.3.2.2 착빙발생마스크

착빙발생 마스크에서는 CTPS 알고리즘의 구름상을 입력자료로 사용한다.

2.3.2.3 착빙발생 강도

착빙발생 강도에서는 DCOEW 알고리즘의 구름수액량을 입력자료로 사용한다.

2.3.3 NWP

착빙발생 고도 모델에서는 수치예보모델인 UM1280의 예측장 자료를 10분 간격, 2km x 2km 해상도로 일치를 시킨 후, 위성관측 공간에 맞추어 사용을 한다.

2.3.3 보조 자료

2.3.3.1 PIREPs

PIREPs(Pilot REPorts)는 항공기 운항 중에 조종사가 항공기의 위치, 상태 및 운항중인 비행 환경에 대하여 기록을 한 보고이다. 관측 당시의 위치 좌표, 고도와 함께 비행기 동체에 발생한 착빙에 대해서도 기록이 되는데, 이는 기존 착빙 발생 관련 연구에서도 널리 사용되어지고 있는 자료이다.

본 연구에서는 모의자료의 운영 기간을 토대로, 2015년 07월 01일 부터 2017년 12월 31일까지의 PIREPs 자료를 항공 기상청과 HKO(HongKong Observatory)에 의뢰하여 취득 후 사용하고 있다. 해당 기간에서는 21개의 착빙 발생 데이터가 취득이 되었다. PIREPs의 유의 범위에 대한 기존 연구를 바탕으로, 해당 착빙 발생 지점 주변 20km와 착빙 미발생 지점 주변 15km 반경 지역까지를, 발생 시간 전·후 5분 이 내에 발생 지점을 착빙 발생 구역이라고 판단하고, 해당 지역의 모의자료 픽셀 값들을 추출하여 연구를 진행하였다.

항공기 조종사의 육안 판단에 의해 기록되는 PIREPs의 특성상, 실제 착빙이 발생한 위치 및 시간과 일치하지 않을 가능성이 있다. 이에 대한 불확실성을 줄이고자 착빙 발생 및 미발생 검증자료들에 대해 품질관리를 실시하여, 이를 통과한 사례들만을 신뢰도 높은 검증 자료로 사용하고자 한다 (Sim et al., 2018). 품질관리 방법은 표 2, 3와 같다.

표 2. 착빙 발생 PIREPs 사례에 대한 품질관리

룰	내용	방법	플래그
룰1	운정고도보다 높은 고도에서의 비행	$FA > CTH + err$	부적합
룰2	얼음운정상구름의 운정고도 근처에서의 비행	$FA \sim CTH + err$ $CPH = ICE$	부적합
룰3	두꺼운 얼음운정상 구름의 내부고도에서의 비행	$FA < CTH + e$ $R01 > 30\%$	적합
룰4	얇은 얼음운정상 구름보다 낮은 고도에서의 비행	$FA < CTH - e$ $R01 < 30\%$	부적합
룰5	따뜻한 운정온도의 구름 근처 및 아래 고도에서의 비행	채널 13의 밝기온도 $> 270K$	부적합
룰6	매우 차가운 운정온도 구름 근처 및 아래 고도에서의 비행	채널 13의 밝기온도 $< 228K$	부적합
최종품질평가		착빙발생 버퍼영역 중 적합의 비율이 20% 이상일 경우	고품질
FA: Flight Altitude; 비행고도 (m) CTH: Cloud Top Height; 운정고도 (km) CPH: Cloud Phase; 운상 err: error; 오차 (RMSE x 0.5) R01: 채널 1번의 반사도 (%)			

표 3. 착빙 미발생 PIREPs 사례에 대한 품질관리

룰	내용	방법	플래그
룰1	운정고도보다 높은 고도에서의 비행	$FA > CTH + err$	부적합
룰2	과냉각수적운정상 구름의 운정고도 근처 및 아래 고도에서의 비행	$FA \sim CTH + err$ $CPH = SLW$	부적합
최종품질평가		착빙발생 버퍼영역 중 부적합의 비율이 80% 이하일 경우	고품질
FA: Flight Altitude; 비행고도 (m) CTH: Cloud Top Height; 운정도고 (km) CPH: Cloud Phase; 운상 err: error; 오차 (RMSE x 0.5) SLW: Super-cooled Liquid Water; 과냉각수적			

2.3.3.2 기상항공기

2018년 1월부터 국립기상과학원에서는 기상항공기를 주 2회 띄워, 구름의 물성 측정을 하고 있다. 온도, 습도, 풍향, 풍속을 측정하는 항공자료조사시스템 (AIMMS-20)을 비롯하여, 구름응결핵 입자크기분포를 측정하는 구름응결핵계수기 (CCN-200), 에어로졸 입자크기분포를 측정하는 광학입자계수기 (Sky-OPC), 구름입자 크기분포와 형태를 측정하는 구름입자측정기 (CCP), 강수입자의 크기분포와 형태를 측정하는 강수입자 측정기 (PIP), 액체수함량과 총수함량을 측정하는 수함량측정기 (WCM2000), 그리고 결빙여부를 관측하는 결빙센서가 있다.

결빙센서는 니켈합금 센서의 공명주파수의 변화를 관측하는 방법으로, 센서에 결빙이 발생했을 경우 진동수가 떨어지는 원리에 기반하여 착빙 여부를 판단한다. 국립기상과학원에서는 400 Hz 이상 진동수 감소가 관측 시에, 착빙이 발생했다고 판단을 한다.

2.4 이론적 배경

항공기 착빙 지역을 탐지하는 것은 과냉각수적(SLW: Super-cooled Liquid Water)이 풍부한 위치를 찾는 것이다. 이론적으로 수동적 위성센서를 통해서 구름의 특

성들을 파악 후 과냉각수적의 분포 위치를 찾을 수 있다.

2.4.1 물리적 문제

일반적으로 구름에서는 빙결점 이하의 온도에서 물방울이 존재하며, 이를 과냉각수적이라고 한다. 과냉각수적이 항공기와 같은 딱딱한 물체에 부딪히면, 얼음피막이 형성되는데, 이를 착빙이라 한다. 착빙은 항공기 조종에 문제를 일으켜 항공기 사고의 위험을 높인다. 이런 착빙으로 인한 사고를 줄이기 위해 착빙이 잘 생기지 않게 만드는 코팅 및 De-icing 장비를 사용하기도 하지만, 심한 착빙이 발생하면 제 기능을 하지 못한다. 따라서 착빙 발생 가능성이 높은 지역을 미리 파악하고 피해가는 것이 중요하다. 과냉각수적 생성환경은 구름의 특성과 관련이 있으므로 GK-2A AMI 산출물로부터 얻은 구름 특성을 이용하여 착빙 발생 가능성을 알 수 있다.

2.4.2 착빙발생 마스크 구축

해당 모델 입력 자료의 물리적 의미는 표 1와 같다.

채널 1, 2, 4번은 가시광선 영역에 해당한다. 두꺼운 구름들은 가시광선 투과가 힘들어, 이 영역에서 높은 반사도를 가진다. 즉 구름의 두께는 1, 2, 4번 채널의 반사도 값에 영향을 미친다. 따라서 가시광선 영역의 반사도 값을 통해 구름의 수직적 특성, 특히 상층부분에서의 특성을 알 수 있다.

채널 6(AHI기준 채널 5번), 7번은 근적외선 영역에 해당한다. 근적외선이 입자 크기에 영향을 받으므로 구름 내 물방울 입자와 얼음 입자의 구성 정도를 알 수 있으며, 저층 구름에 민감하다. 따라서 구름의 구성 입자의 상태를 반영하여, 착빙 발생 환경을 알 수가 있고, 저층구름 특성 여부 정보를 알 수 있다.

채널 8, 9, 10번은 수증기 탐지 영역에 해당한다. 수증기 탐지 영역은 구름 내 고도별 수분 분포와 밀접한 상관성이 있다. 따라서 구름 내 수분의 수직적 분포에 관한 정보를 얻을 수 있다.

채널 11, 13, 14, 15, 16번은 열적외선 영역에 해당한다. 열적외선 영역을 통해 구름 상층부의 온도 및 고도, 가강수량 등의 특성을 알 수 있다.

2.4.3 랜덤포레스트

본 알고리즘에서는 기존 연구(Sim et al., 2018)를 통해 착빙 탐지에 유의미하다는 것이 증명된 랜덤포레스트 모델을 이용하였다. 랜덤포레스트 모델은 결정나무 기반

알고리즘인 CARTs를 확률적으로 접근한 알고리즘으로 (Breinman et al., 2001), CARTs는 훈련자료를 기반으로 하여 타겟 변수를 위한 규칙으로 구성된 결정나무를 구성하는 알고리즘이다. 랜덤포레스트 모델은 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 3은 CART와 랜덤포레스트 모델의 구성을 보여준다.

랜덤 포레스트 모델은 알고리즘을 구성하는 변수들의 변수중요도를 제공한다. 최적화 알고리즘을 구성하기 위한 무작위 과정을 반복할 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 통과함으로써 인해 줄어드는 엔트로피값을 정량화 하여 나타낸 지수이다. 랜덤포레스트 모델에서는 "Mean Decrease Accuracy"지수와 "Mean Decrease Gini"지수가 높을수록 높은 변수중요도를 가진다

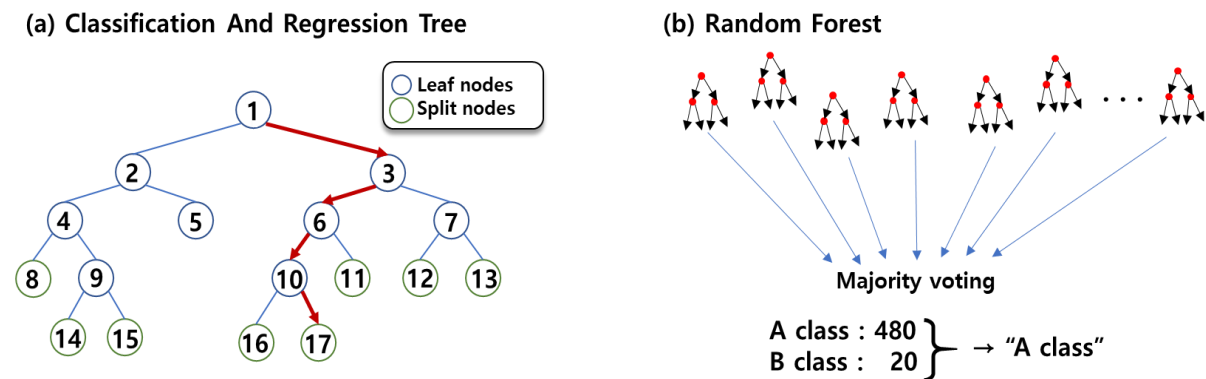


그림 3. (a) CART (Classification and regression trees)와 (b) 랜덤포레스트 (Random forest) 모델의 구성

2.5 수학적 설명

착빙발생은 과냉각수적의 존재여부와 그 분포구성에 의해 결정된다. 따라서 본 알고리즘은 과냉각수적에 관한 정보를 AMI 산출물들을 통해서 알아내었다.

2.5.1 착빙발생 마스크 알고리즘

해당 모델은 해당 픽셀의 착빙 발생 여부를 AMI 추출 값들을 랜덤포레스트 모델을 통해 분류해냈다. 랜덤포레스트 모델 학습 시, 입력 변수들을 하나씩 제외하며 에러 변화율을 측정하여 상대적 변수 중요도를 파악할 수 있는데, 본 알고리즘에 사용된 변수들의 상대적 중요도는 **그림 4** 와 같다. 두 모델 모두 운정의 상태를 나타내는 구름상과 상층 수증기량을 나타내는 8번 채널에서 높은 변수중요도를 가지고 있다. 밤모델의 경우, 저층구름의 정보를 나타내는 7번 채널과 운정온도를 나타내는 13번 채널의 중요도가 높았다. 구름의 상태, 수증기량, 온도는 SLW의 발생환경에 매우 중요한 요소이고, SLW는 착빙 발생을 결정짓는 요인으로, 당 모델은 착빙 발생에 중요한 인자들의 변수중요도가 높게 나타남을 알 수 있다.

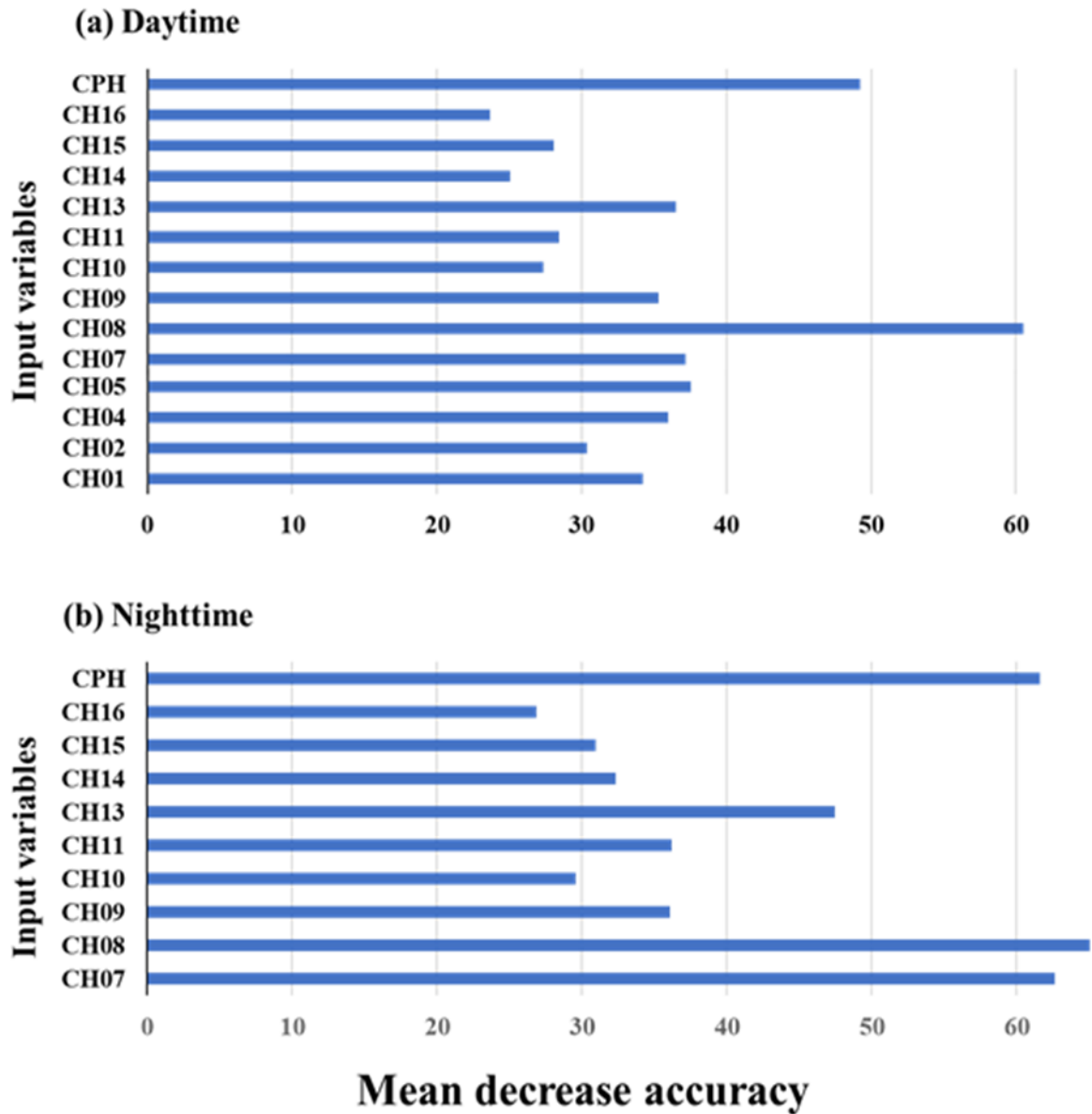


그림 4. 랜덤포레스트 AMI 모델 입력 변수들의 상대적 변수중요도. (a) 낮시간 모델, (b) 밤시간 모델.

2.5.2 과냉각수적 후처리

랜덤포레스트를 통해 분류된 착빙 픽셀 중, 13번 채널의 밝기온도 값이 270K보다 높은 화소들은 과냉각 수적이 존재할 수 없는 지역으로 판단하여, 미착빙으로 추가 분류하였다 (Sim et al., 2018).

2.5.3 착빙발생 강도 알고리즘

착빙발생 마스크 알고리즘을 통해 생성된 착빙 발생 지역에 해당하는 화소 중, 구름수액량의 값이 488 gm-2 보다 높을 경우 강한착빙, 그렇지 않을 경우 약한착빙으로 분류를 하였다 (Smith et al., 2012).

2.5.4 착빙발생 고도 알고리즘

착빙발생 마스크 알고리즘을 통해 생성된 착빙 발생 지역에 해당하는 화소 중, 수치모델기반 온·습도 프로파일의 분포가 온도의 경우 -35°C 이상 0°C 이하이고, 상대습도 경우 60% 이상인 고도에서 착빙이 발생한다고 구분을 하였다 (Bernstein et al., 2005).

2.6. 산출물

2km x 2km 공간해상도와 10분 간격의 시간해상도로 착빙 발생 강도에 대한 결과를 강한착빙과 약한착빙 강도로 산출을 하고 (ICING 산출물), 착빙발생 고도결과는 72개의 수직 프로파일 기준으로, 28번째 (약 11km)에서 41번째 (약 6.5km) 구간에 대한 정보를 이진으로 변환 후 제공한다 (ICING_PLEV 산출물). 이 때, 착빙으로 예측이 되는 픽셀 중 착빙고도가 산출이 안되는 픽셀은 저품질로 평가가 되어, 'Bad (2)' 플래그가 부여가 되고, 그 외의 픽셀은 밤낮 여부에 따라 'Day-time cloud (0)' 과 'Night-time cloud (1)' 가 부여된다 (DQF_ICING 산출물)

3. 모의자료 및 검증

3.1 모의 입력자료 및 검증자료

본 연구는 GK-2A AMI 의 모의자료인 Himawari-8 AHI 를 SBAF (Spectral Band Adjustment Factors)를 통해 AMI 로 변환한 자료와 AMI 자료를 복합적으로 이용하여 모의하였다 (Chander et al., 2012).

품질관리를 통과한 PIREPs 에서 착빙발생 16 개, 미발생 샘플 48 개와, 기상항공기를 통해 취득된 13 개의 착빙발생 사례와 17 개의 미발생 사례를 모두 사용하여, 총 29 개의 착빙발생 사례와 65 개의 미발생 사례가 검증자료로 사용되었다.

3.2 검증 방법

총 29 개의 착빙발생 사례와 65 개의 미발생 사례 중 26 개의 착빙발생 사례와 56 개의 미발생 사례를 랜덤포레스트 학습자료로, 3 개의 착빙발생 사례와 9 개의 미발생 사례를 검증자료로 구분하여 사용하였다.

3.3 검증 결과

분류검증 결과, POD는 66.6%, POFD는 22.2%를 나타내었다 (표 5,6). 그림 6 은 2018.09.16. 08:50-09:10 UTC 시기에 대한 착빙발생 강도 결과 이미지이다. 08:50에는 낮시간 모듈이, 09:00과 09:10에는 밤시간 모듈이 사용되었는데, 세 영상간의 착빙예측 분포가 연속적으로 연결되는 것을 확인할 수 있다. 그림 7은 2017.01.28. 01:40 UTC 때의 착빙발생 고도 결과를 나타낸 그림이다.

표 4. 분류정확도 평가 테이블 구성

Modeled	Reference (observed)	
	No	Yes
No	Correct negative (N)	Miss (M)
Yes	False (F)	Hit (H)

probability of detection (POD)= $H/(H+M) \times 100\%$;

probability of false detection ($POFD$)= $F/(N+F) \times 100\%$

표 5. 검증사례들을 이용한 AMI 랜덤포레스트 모델 평가 결과

Modeled	Reference (observed)	
	No	Yes
No	7	1
Yes	2	2

POD = 66.6%

POFD = 22.2%

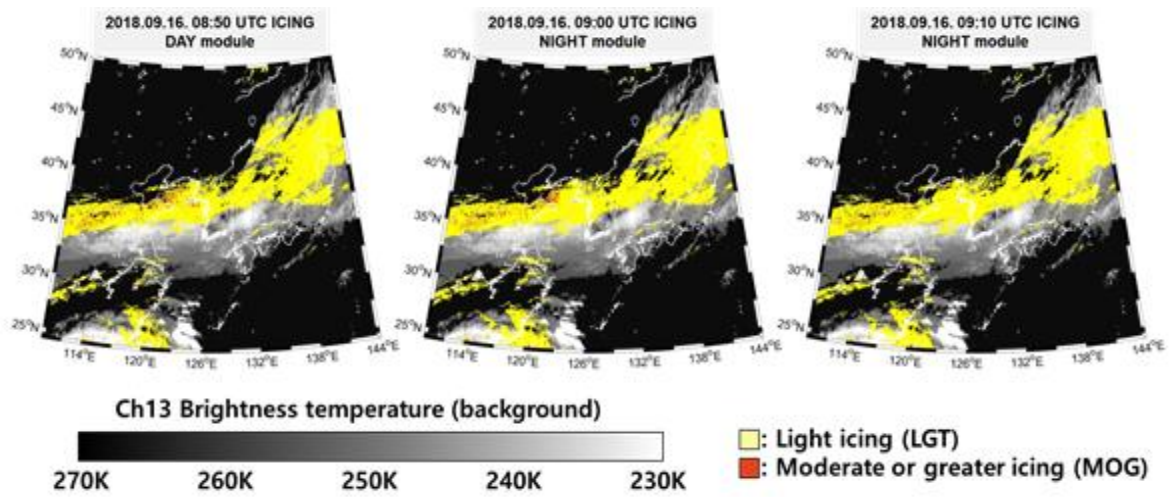


그림 5. 2018.09.16. 08:50-09:10 UTC에 대한 AHI 기반 착빙강도 결과. 미착빙 픽셀은 채널13의 밝기온도로 대체되었음.

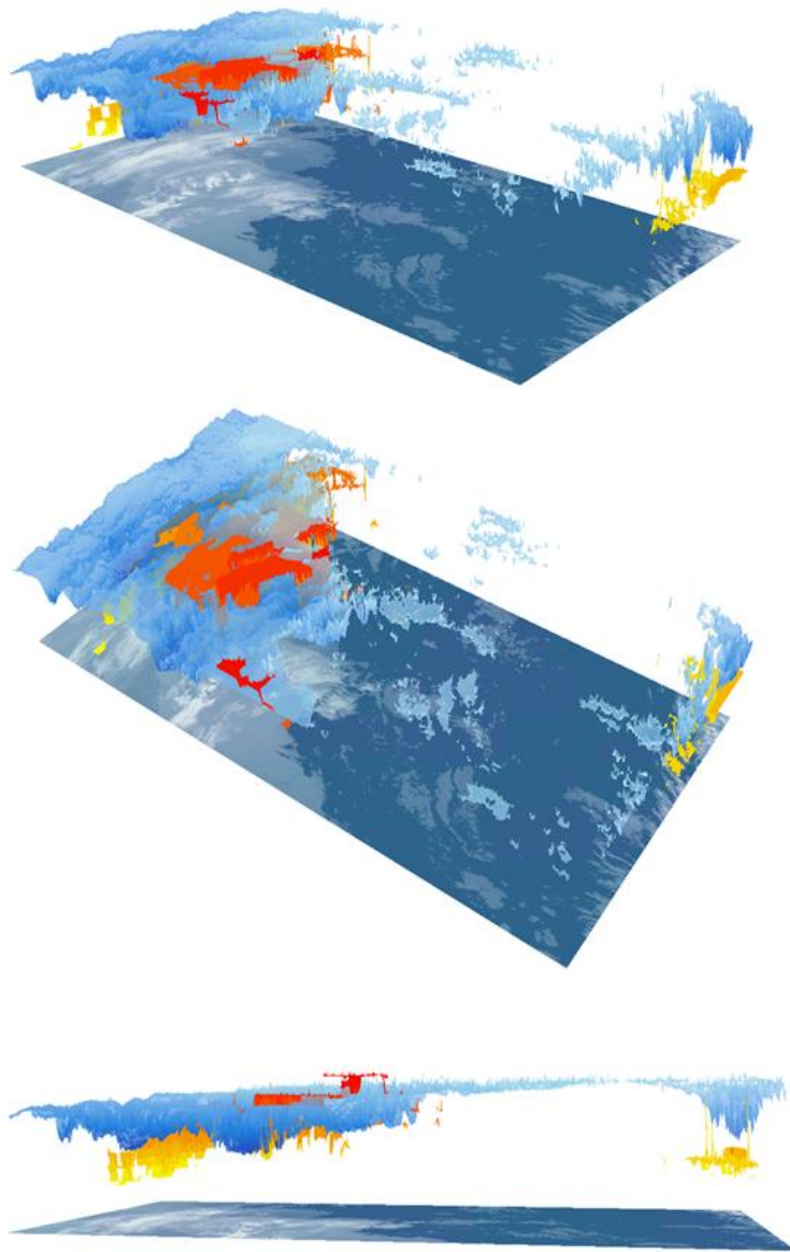


그림 6. 2017.01.28. 01:40 UTC 시기의 AHI와 UM768 자료를 이용한 착빙발생 고도 이미지. 청색: 채널13의 밝기온도, 황색: 낮은 고도 착빙, 적색: 높은 고도 착빙

4. 현업운영 시 고려사항

4.1 수치계산 고려사항

착빙 알고리즘 구동 중 유효숫자의 오류로 인해 오류가 생길 수 있고, 이를 방지하기 위해 프로그램 상의 수치 유효숫자를 구체적으로 지정한다. 또한, 입력 자료의 정확도 및 손상 정도에 민감할 수 있어, L1B자료를 비롯한 L2자료, 수치예보자료의 품질 확인이 선행되어야 한다.

4.2 프로그래밍 및 절차상의 고려사항

착빙 알고리즘은 화소 기반으로 구동하며, 입력자료들의 크기가 프로그래밍 된 크기와 맞지 않을 시, 구동에 영향을 끼치게 된다. 또한 구동에 필요한 낮과 밤의 랜덤 포레스트 모델이 지정된 경로에 위치해 있음을 확인한다.

4.3 품질 평가 및 진단

주기적으로 정확도 평가를 수행하여 알고리즘이 목표 정확도를 잘 달성하고 있는지에 대한 평가를 진행한다.

4.4 예외사항 처리

착빙 알고리즘은 입력 자료의 품질과 유효성 그리고 각 픽셀의 위치정보나 기하학 정보를 평가 후 진행해야 한다. 입력 자료에 문제가 발생 시, 당 알고리즘의 구동이 불가하거나 낮은 신뢰도를 갖게 될 수 있다.

4.5 검증 알고리즘

검증자료인 PIREPs를 기반으로 진행이 되며, 착빙이 발생한 사례가 주기적으로

취득되지 않을 수 있다. 또한 품질관리를 거치더라도 PIREPs의 불확실성이 존재할 수 있다.

5. 가정 및 제한

5.1 성능

이 문서에 있는 착빙탐지 알고리즘을 수행하기 위해서는 아래와 같은 가정이 필요하다.

- 착빙 알고리즘에 사용되는 입력변수들의 품질은 매우 우수하다.
- 샘플로 구축된 착빙데이터는 공간적으로는 20km 반경, 시간적으로는 ± 5 분으로 착빙이 고르게 발생한 경우이다.

5.2 예상 센서 성능

GK-2A AMI의 모의자료인 Himawari-8 AHI를 SBAF를 통해 AMI로 변환한 자료와 AMI 자료를 복합적으로 이용하여 알고리즘 모델을 구축하였다. 이 때 변환식이 완벽하지 않아, 구축된 임계모델의 성능이 AMI로만 구축한 모델보다 낮을 수 있다. 추후 AMI 기반의 샘플을 다량 취득 후, 이만을 이용한 모델 학습을 진행하도록 한다.

5.3 개선을 위한 사전 계획

현재 착빙발생에 대한 샘플이 매우 부족하여, 입력변수 최적화가 제대로 되지 않았다. 따라서 착빙발생에 대한 샘플을 국외의 착빙발생 PIREPs도 취득을 하여 샘플을 늘린 후, 입력변수의 최적화 및 독립된 데이터를 이용한 검증 역시 시행할 예정이다.

6. 참고문헌

- Bernstein, B. C., McDonough, F., Politovich, M. K., Brown, B. G., Ratvasky, T. P., Miller, D. R., ... & Cunning, G. (2005). Current icing potential: Algorithm description and comparison with aircraft observations. *Journal of Applied Meteorology*, 44(7), 969–986.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Sim, S., Im, J., Park, S., Park, H., Ahn, M. H., & Chan, P. W. (2018). Icing Detection over East Asia from Geostationary Satellite Data Using Machine Learning Approaches. *Remote Sensing*, 10(4), 631.
- Smith Jr, W. L.; Patrick Minnis.; Cecilia Fleegeer. Algorithm Theoretical Basis Document: Flight Icing Threat; NOAA NESDIS Center for Satellite Applications and Research; College Park, MD, USA, 2010.
- Smith Jr, W. L., Minnis, P., Fleegeer, C., Spangenberg, D., Palikonda, R., & Nguyen, L. (2012). Determining the flight icing threat to aircraft with single-layer cloud parameters derived from operational satellite data. *Journal of Applied Meteorology and Climatology*, 51(10), 1794–1810.