

보안 과제(), 일반 과제(✓) / 공개(✓), 비공개()
기상See-At기술개발사업 제3차 연도 최종보고

R&D / 2015-1100

위성-수치모델 기반 태풍발생 탐지/예측 시스템 개발 연구 최종보고서

2018 . 7 . 25

주관연구기관 / 한국해양과학기술원
위탁연구기관 / 한국해양대학교

기 상 청
한국기상산업기술원

제 출 문

기상청장 귀하

'위성-수치모델 기반 태풍발생 탐지/예측 시스템 개발 연구'(연구개발 기간 : 2015. 05. 01 ~ 2018. 04. 30) 과제의 최종보고서 1부를 제출합니다.

2018 . 07 . 25 .

주관연구기관명: 한국해양과학기술원 (대표자) 김웅서 (인)
위탁연구기관명: 한국해양대학교 산학협력단 (대표자) 경갑수 (인)



주관연구기관책임자: 박명숙

협동연구기관책임자: 김형석

기상업무 연구개발사업 처리규정 제28조에 따라 최종보고서 열람에
동의합니다.

보고서 요약서

과제 고유 번호	KMIPA2015-1100	해당 단계 연구 기간	2017.5.1. ~2018.4.30.	단계구분	3/3
연구사업명	중사업명	기상·지진 See-At 기술개발연구			
	세부사업명	기상예보기술			
연구과제명	대과제명	위성-수치모델 기반 태풍발생 탐지/예측 시스템 개발 연구			
	세부과제명	위성-수치모델 기반 태풍발생 탐지/예측 시스템 개발 연구			
연구책임자	박명숙	해당단계 참여 연구원 수	총: 16 명 내부: 5 명 외부: 11 명	해당단계 연구개발비	정부: 200,000천원 기업: 천원 정부 외: 천원 계: 200,000천원
		총 연구기간 참여 연구원 수	총: 30 명 내부: 5 명 외부: 25 명	총 연구개발비	정부: 570,000천원 기업: 천원 정부 외: 천원 계: 570,000천원
연구기관명 및 소속 부서명	한국해양과학기술원 해양위성센터			참여기업명	
국제공동연구	상대국명:			상대국 연구기관명:	
위탁연구	연구기관명: 한국해양대학교			연구책임자: 김형석	
요약				보고서 면수 : 96	
<ul style="list-style-type: none"> □ 전구수치모델/분석장을 이용하여 태풍 씨앗 열대섭동을 자동으로 탐지하는 기술을 개발 □ 탐지된 열대 섭동 주변 대규모 환경 분석 기술을 개발하고, 태풍발생탐지/예측 가능성의 가이드언스를 제공하는 모델을 개발하였음 □ 극궤도 위성 위성자료를 이용하여 태풍발생 과정의 역학 특성을 정량화 하는 인자를 개발하였으며, 기계학습을 이용한 객관적 발생탐지/예측 모델을 개발하고 검증하였음 □ 기계학습 기반 태풍발생 탐지 알고리즘을 개발하였음 					

<p>연구의 목적 및 내용</p>	<ul style="list-style-type: none"> - 태풍의 발생 전 열대섭동의 탐지 및 정량적인 분석을 통해 향후 태풍으로 발생할지 쇠퇴할지의 여부를 객관적으로 결정하는, 태풍발생 탐지/예측 알고리즘을 개발한다. - 전구수치모델/분석장을 이용하여 태풍으로 발달할 가능성이 있는 태풍 씨앗 저기압성 순환인 열대섭동 단계부터 탐지하고 추적하는 알고리즘을 개발한다. - 준실시간 이용 가능한 위성 관측 및 산출물을 이용하여, 태풍발생 시에 바람, 온도, 대류의 공간 특성 및 그 시간변동성을 나타내는 정량화된 태풍발달 예측 인자를 개발한다. - 태풍발생 과정의 역학적, 열적, 수문학적 특성을 나타낼 수 있는 위성관측 기반 인자들을 개발한다. - 위성공간정보의 다양한 기법을 이용하여, 정지궤도위성 적외, 수증기 채널 정보를 이용한 대류 강도/내부 구조와 하층 바람장을 관측할 수 있는 극궤도 마이크로파 위성자료를 효율적으로 결합할 수 있도록 한다. - 통계적 방법과 기계학습기술을 적용하여 위성 기반 태풍발생탐지 알고리즘을 개발하여, 전구수치모델 자료를 융합 이용한 객관화된 태풍발달탐지/예측기술을 개발한다. - 전구수치모델에 의한 대규모-중관 분석과 위성 기반 알고리즘을 효율적 결합을 통한 독자적 태풍발생 탐지/예측 기술을 개발한다. - 전구 수치모델-위성 결합 태풍발생 탐지/예측 시스템의 검증 및 평가를 실시하며, 향후 기상 현업에 활용할 수 있도록 한다. 				
<p>연구개발성과</p>	<ul style="list-style-type: none"> - 전구수치모델/분석장을 이용하여 태풍으로 발달할 가능성이 있는 열대섭동 단계부터 자동으로 탐지하고 추적하는 기술을 개발하였음 - 탐지 된 열대섭동 주변 대규모 환경을 분석하여, 전구모델을 이용한 태풍발생탐지/예측 모델을 구성하였음 [Journal of Climate 1편 게재확정, 기후변화 학회지 1편, S/W 등록 1건] - 준실시간 극궤도 위성 WindSat 산출물(해상풍)을 이용하여 태풍 발생 시 시스템의 대칭성, 조직화, 강도 등을 정량화 하는 인자를 개발하였으며, 기계학습 기반 태풍발생 탐지 알고리즘을 개발하였음 [관련 논문 Remote Sensing of Environment 1편 게재] - 정지궤도 천리안 기상위성 관측을 이용하여 대류의 강도, 조직화 정도를 나타내는 인자를 개발하였음. 특히, GIS의 공간분석 프로그램(Fragstats)을 새롭게 응용하여 태풍 발생과 관련된 구름 인자를 최초로 정량화하였음 [관련 논문 Journal of Geophysical Research 1편 게재] - 태풍 발생 역학에 대한 연구를 수행하였음 [관련논문 Monthly Weather Review 2편 게재] - 통계적인 방법과 기계학습 방법들을 이용하여 위성 인자들과 태풍 발생/비발생을 관련성을 나타내는 객관적 알고리즘들을 개발하였음 - 태풍발생 모델을 구성할 때 일반적인 통계적인 방법(선형회귀방법)에 비해 기계학습 방법을 이용했을 때, 성능 개선되는 점을 연구하였음 - 태풍발생 가능성 여부의 객관적 결정 알고리즘을 개발하고 현업에 사용을 위해 검증하였음 				
<p>연구개발성과의 활용계획(기대효과)</p>	<ul style="list-style-type: none"> - 전구수치 모델장 결과와 다양한 위성자료를 이용한 태풍 발생 탐지, 예측 시스템의 기술을 수립하여, 현업화를 위한 기술을 개발한다. - 향후 태풍 현업 기관인 국가태풍센터와 지속적인 협업을 통해 우리나라 기상현업에서 독자적인 태풍 발생 시스템을 구축하여, 태풍 전 탐지 및 예보시스템이 개선되어 신속하고 효과적인 태풍 현업 업무를 수행할 수 있도록 기여 한다. 				
<p>핵심어(5개 이내)</p>	<p>태풍발생</p>	<p>위성</p>	<p>전구 수치모델</p>	<p>예측</p>	<p>기계학습</p>

< SUMMARY >

<p>Purpose & Contents</p>	<ul style="list-style-type: none"> □ This study purposed to develop satellite–numerical model based tropical cyclone formation detection/prediction technique that can be used for operational forecast at KMA. □ We purposed to develop algorithm to track many tropical disturbances from global models, and analyze large–scale and synoptic conditions and tropical waves. □ From various satellite data, we purposed to develop new predictors to quantify the spatial and temporal variations of wind, temperature, and convection of developing tropical disturbances. □ In particular, several spatial pattern analysis techniques in GIS are newly applied to develop satellite–based tropical cyclone genesis indices. □ Using developed input indices, we purposed to develop objective tropical cyclone formation detection/prediction algorithm through machine learning technology. □ We purposed to perform validation and evaluation of the developed techniques using an independent dataset, and proposed to a model for real–time KMA's operation. 				
<p>Results</p>	<p>Using global model analysis data(NCEP FNL), We have developed an objective tracking technique of detecting many tropical disturbances over the western North Pacific that are possible to develop into tropical cyclone. By analyzing various large–scale factors, we also developed an objective tropical cyclone formation detection technique based on the global model analysis.</p> <p>Through quantitative pattern and intensity analyses of WindSat sea–surface winds, this study has developed numerical indices to quantify the degrees in the symmetry in system cyclonic circulation, organization of strong wind areas, and system intensity. The degree of cyclonic circulation symmetry near the system center is quantified using circular variances, and the degree of strong wind aggregation(heavy rainfall) is defined using a spatial pattern analysis program tool called FRAGSTATS. In addition, the circulation strength and convection are defined based on the areal averages of wind speed and rainfall. An objective TC formation detection model is then developed by applying those indices to machine–learning tree algorithms and linear discriminant analyses(LDA). The comparison between the convective LDA and machine learning based models indicates that application of machine learning techniques can improve hit rate and lead time of tropical cyclone formation detection even using the same input indices. For operation uses, we suggested to consider various tropical cyclone formation models developed in this study, since each of them different aspect (large–scale aspect, system–scale evolution, and mesoscale evolution) of tropical cyclone process.</p>				
<p>Expected Contribution</p>	<p>The developed models will contribute to operational tropical cyclone formation monitoring and prediction in KMA. Establishing the independent, tropical cyclone formation detection and prediction system will contribute to improve tropical cyclone detection and prediction systems at all tropical cyclone life stages that can give more effective warning to public.</p>				
<p>Keywords</p>	<p>Tropical cyclone</p>	<p>Satellite</p>	<p>Numerical model</p>	<p>prediction</p>	<p>Tropical meteorology</p>

<Contents>

Chapter 1. Overview of Research & Development	1
1. Objectives of Research & Development	1
2. Necessities of Research & Development	1
3. Scope of Research & development	4
Chapter 2. Status of International and Domestic Technology Development	5
1. International Technology Development Status	5
가. US JTWC Tropical cyclone warning system	5
나. JMA operational method	6
다. CIMSS Advanced Dvorak Technique(ADT)	8
라. Colorado State University CIRA	8
마. University of Arizona: Deviation Angle Variation(DAV) algorithm	9
2. Domestic Technology Development Status	11
가. KMA National Typhoon Center	11
나. KMA National Meteorological Satellite Center	12
3. Summary of previous techniques	12
Chapter 3. Research & Development Results and Performances	13
1. Detection of tropical disturbances	13
2. Developing TC formation indices	20
가. WindSat-based indices	21
나. COMS/MI-based indices	27
다. Himawari-8/AHI-based indices	34
라. NCEP FNL based indices	41
3. Development of TC occurrence detection model	45
가. Windsat-based model	47
나. COMS/MI-based model	51
다. Himawari-8/AHI-based model	56
라. Global model-based model	59
Chapter 4. Level of Objectives Achievement and Contribution of the Relevant Fields	63
1. Level of objectives achievement	63
가. Performance & method of evaluation	63
나. Evidences of evaluation	64
다. Performance index	70
2. Contribution to the related fields	71
가. Technological aspect	71
나. Economical-industrial aspect	72
Chapter 5. Application plan for Research & Development Results	73
Chapter 6. International Science and Technology Information Derived During Research & Development Processes	78
Chapter 7. Security Level of Research & Development Results	79

Chapter 8. Research Facilities and Equipments enrolled in National Science & Technology Information	80
Chapter 9. Security Actions of Laboratory and other facilities	81
Chapter 10. Representative performances of Research & Development	82
Chapter 11. The others	84
Chapter 12. References	85

〈 목 차 〉

제 1장. 연구개발과제의 개요	1
1. 연구개발 목적	1
2. 연구개발의 필요성	1
3. 연구개발 범위	4
제 2장. 국내외 기술 개발 현황	5
1. 국외 기술 개발	5
가. 미국 JTWC 태풍발생경보 시스템	5
나. 일본 기상청 현업 태풍 탐지 분석 방법-EDA에 기반	6
다. CIMSS Advanced Dvorak Technique(ADT)	8
라. Colorado State University CIRA	8
마. University of Arizona: Deviation Angle Variation(DAV) 알고리즘	9
2. 국내 기술 개발	11
가. 현재 국내 기상청 현업 (국가태풍센터)	11
나. 국가 위성센터 태풍 분석 알고리즘 개발	12
3. 기존 기술의 취약성 및 과제의 목표 설정	12
제 3장. 연구 수행 내용 및 성과	13
1. 태풍 씨앗 열대섭동 탐지 기술 개발	13
2. 태풍발생 인자 개발	20
가. 극궤도 Windsat 해상풍 자료 이용 인자 개발	21
나. 정지궤도 천리안 기상위성(COMS MI) 인자 개발	27
다. 정지궤도 히마와리-8 위성(Himawari-8 AHI) 인자 개발	34
라. 모델장을 이용한 대규모환경 인자 개발	41
3. 태풍발생 탐지/예측 모델 개발	45
가. 극궤도 Windsat 해상풍 자료 기반 태풍발생탐지 모델 개발	47
나. COMS MI 기반 태풍발생탐지 모델 개발	51
다. Himawari-8 AHI 기반 태풍발생탐지 모델 개발	56
라. 전구 모델 기반 태풍발생 탐지/예측 모델 개발	59
제 4장. 목표 달성도 및 관련 분야 기여도	63
1. 목표 달성도	63
가. 연구 개발 성과 및 평가방법	63
나. 평가의 근거	64
다. 성과지표	70
2. 관련 분야 기여도	71
가. 기술적 측면	71
나. 경제적·산업적 측면	72
제 5장. 연구개발성과의 활용계획	73

제 6장. 연구 과정에서 수집한 해외 과학기술 정보	78
제 7장. 연구개발성과의 보안등급	79
제 8장. 국가과학기술종합정보시스템에 등록된 연구시설·장비 현황	80
제 9장. 연구개발과제 수행에 따른 연구실 등의 안전 조치 이행 실적	81
제 10장. 연구개발과제의 대표적 연구 실적	82
제 11장. 기타 사항	84
제 12장. 참고 문헌	85

제 1장. 연구개발과제의 개요

1. 연구개발 목적

- 정지궤도/극궤도 위성 관측 및 산출물(적외, 수증기, 상층온도, 해상풍 등)을 이용한 태풍 발생 탐지 기술 개발
- 전지구 대기 수치모델 예측/분석장을 이용하여 태풍으로 발달할 가능성이 있는 열대 섭동 탐지 및 추적
- 수치모델-위성 결합 시스템 태풍 발생 탐지/예측 시스템 기술 개발, 검증 및 평가.
- 수치모델/위성관측 자료 기반의 태풍발생탐지 기술의 최적화 및 현업 활용.
- 개발될 기술의 수준은 객관적 탐지방에서 세계최고 수준(University of Arizona Deviation Angle Variance 기법)으로 함 : 열대 저압부 발생 이전 혹은 이후 100시간 내 Hit rate 93%, False alarm 22% 수준

2. 연구개발의 필요성

태풍은 우리나라에 큰 인명 피해와 경제적 손실을 초래하는 기상현상이다. 기상청의 과거 기상재해 통계에 따르면(표 1), 2002년 루사, 2003년 매미, 2006년 에위니아, 1999년 올가, 1987년 쉘마, 1995년 제니스 등 태풍 혹은 태풍에 동반되는 집중호우가 피해 순위, 1, 2, 3, 5, 9, 10 위에 기록되고 있다. 태풍은 한번 우리나라에 상륙할 때 막대한 피해를 줄 수 있다.

태풍은 열대 서태평양에서 발생 후 수일 이내에 우리나라에 상륙하여 피해를 줄 수 있다. 최근에도, 2012년 볼라벤과 덴빈은 약 6,365억원, 2012년 산바는 3,657억원의 재산 피해를 낸 바가 있다.

태풍 일생 중 가장 예보가 가장 어려운 단계는 상대적으로 저기압성 순환의 강도가 약한 발생단계이다. 학계 및 선진국의 현업 기관에 태풍발생 시점은 열대저압부(Tropical Depression, TD)가 발생하는 시점이다(표 2). 국내 현업기관인 국가태풍센터에서는 더 강한 35 knots부터 태풍으로 정의한다. 따라서, 본 연구과제에서는 상대적으로 탐지가 더 어려운 열대저압부 생성 시점을 기준으로 알고리즘을 주로 개발하였다. 추가적으로, 최근기간의 현업 검증을 통하여 국내 현업에 이용하는 기준을 이용하여 알고리즘을 조정하였다.

최근, 수치 모델은 슈퍼컴퓨터 성능의 발달로 엄청난 진보가 있었으며, 중규모 모델의 수평 격자는 수 km까지 세밀해졌다. 하지만, 실제 대기 해양의 다양한 규모의 물리과정으로 발생하는 열대섭동에서 태풍으로 변환되는 과정을 수치모델로 정확히 구현하는 것은 한계가 많다. 최근 몇몇 연구들(예, Park et al. 2013)은 중규모 모델이 실제 태풍으로 발달되지 않는 저기압성 순환도 태풍으로 과다 발달(over-deepening)시키는 경향이 있다고 보고하고 있다. 대부분의 전구 모델은 태풍으로 발달하지 않는 경우를 발달로 예보하는 오보(False Alarm)가 많은 것으로 알려져 있다(IWTC-VII).

현재 우리나라 현업기관인 국가태풍센터에서는 경험 있는 예보관이 태풍 발생단계를 정성적/주관적으로 분석하는 작업이 이뤄지고 있다. 태풍발생 가능성을 추정하는 객관적 탐지/예측 알고리즘이 없기 때문에 태풍의 발생 뿐 아니라 전 생애의 탐지가 국외 태풍현업기관들에 많이 의존하고 있는 것이 사실이다. 또한, 숙련된 몇몇의 예보관의 주관에 의존하고 있기 때문에, 예보관이 달라졌을 때 예보 성능이 달라질 수 있다.

표 1. 과거 기상재해(호우, 태풍, 대설) 피해 순위(국민안전처 통계자료, 기상청)

순위	재해기간	주요피해원인	주요피해지역	사 망 · 실 종 (명)	재산피해 (억원)
1	2002년 8.30~9.1	태 풍(루사)	전 국	246	51,479
2	2003년 9.12~9.13	태 풍(매미)	전국(서울, 인천 제외)	131	42,225
3	2006년 7.9~7.29	호 우, 태 풍(에위니아)	전 국	62	18,344
4	1998년 7.31~8.18	호 우	전 국(제주 제외)	324	12,478
5	1999년 7.23~8.4	호 우, 태 풍(올가)	전 국	67	10,490
6	2002년 8.4~8.11	호 우	경기, 강원, 충북, 경북, 경남 등	23	9,181
7	1990년 9. 9~9.12	호 우	서울, 경기, 강원, 충북	163	5,203
8	2004년 3.4~3.5	대 설	충남, 충북, 대전, 경북 등	-	6,734
9	1987년 7.15~7.16	태 풍(셀마)	경남, 강원	345	3,913
10	1995년 8.19~8.30	호 우, 태풍(제니스)	서울, 경기, 충청	65	4,563

표 2. 국제 학계에서 통용되는 태풍 발생의 정의

Title	Definition	Speed	Defining formation
Tropical Disturbance	Distinct area of disturbed weather with no well-defined low-level circulation	< 25	
Tropical Depression	Fairly well-defined low-level circulation (Numbered storm, 01W, 02W, 03W, ...)	25-34	Tropical cyclone formation
Tropical Storm	Named storm	34-63	
Typhoon	With special Super typhoon category ≥ 130 kt	> 63	

본 연구 과제를 통해서 수치-위성을 이용한 한국형 객관적 태풍 발생 탐지/예측 기술을 개발하며 현업화(연계) 기술을 개발하였다. 독자적인 발생 탐지 기술을 구축함으로써, 향후 태풍의 전 단계 강도 탐지 및 예보시스템이 개선되어 신속하고 효과적인 태풍 현업 업무를 수행할 수 있는 기틀을 마련하는 것을 목표로 한다.

3. 연구개발 범위

구분	연도	연구개발의 목표	연구범위
1차년도	2015	위성관측 자료를 이용한 태풍발생 탐지/예측 인자 개발	<ul style="list-style-type: none"> - 준 실시간 이용 가능한 위성 관측 및 산출물 (적외, 수증기, 상층 온도, 해상풍)을 선별 - 선별된 위성관측/산출을 이용하여 태풍발생시 동반되는 순환과 대류의 (시)공간패턴을 잘 표현하는 예측 인자 개발
	2015	위성 인자를 이용한 알고리즘 원형 개발	<ul style="list-style-type: none"> - 통계모형 및 기계학습 기술을 적용하여 객관적 알고리즘 개발 - 경험적 분석에 의한 주관적 방법 개발
	2015	전구 수치 모델을 이용한 열대섭동 탐지	-준 실시간 이용 가능한 전구모델의 저기압성 와도의 열대 섭동을 탐지
2차년도	2016	검증과 평가를 통한 위성 태풍발생 탐지 알고리즘 원형 개선	- 과거 태풍 자료를 이용하여 시스템별 검증으로 태풍발생탐지/예측 시스템의 정량적인 평가
	2016	전구 수치 모델을 이용한 태풍발생 확률 분석	-준실시간 이용 가능한 전구 수치모델들 이용한 대규모-중관 환경 및 열대 파동 등의 태풍발생 호조건 분석
	2016	수치모델-위성 결합 시스템 원형 개발	<ul style="list-style-type: none"> - 전구수치모델의 중간/대규모 환경장 분석결과물, 위성 알고리즘의 결과물, 실제 태풍발생과의 연관성 분석 -어떤 결과물이 실제 태풍발생을 잘 표현하는지 비교하여 평가
3차년도	2017	수치모델과 위성 관측의 효율적 결합 시스템	-시스템 최적화를 위해 전구수치모델 시스템과 위성알고리즘을 효율적 결합 방법을 연구
	2017	검증 및 평가	<ul style="list-style-type: none"> -과거 태풍 자료를 이용한 시스템별 검증으로 태풍 발생탐지/예측 시스템의 정량적 평가 -최종적으로 태풍발생 탐지 및 예측시스템에서 생산된 결과가 표출되어 현업 활용 가능한 체계 구축

제 2장. 국내외 기술 개발 현황

1. 국외 기술 개발

가. 미국 JTWC 태풍발생경보 시스템

미국 Joint Typhoon Warning Center(JTWC)는 자체적으로 태풍발생 가능성이 있는 열대섭동을 추적하고 있으며, 차례대로 시스템 번호(예, 90W Invest, 91W Invest)를 부여한다. 각 열대섭동들에 대해 위성관측, Dvorak Technique, 대기장, 지수 등을 이용하여 앞으로 24시간동안 열대저기압 발생 확률을 예보자가 주관적인 방법으로 판단하고 있다 (Kucas and Darlow, 2012). 태풍발생확률을 이전에는 "Poor", "Fair", "Good" 분류로 제공하였으나, 2010년 이후에는 새로운 분류법인 "low", "medium", "High"를 사용하고 있다. 이 분류의 의미는 다음과 같다.

- "Low": 이 영역의 열대저기압 발생을 감시하고 있으나, 앞으로 24시간 동안 발생하지 않을 것임 [발생가능성: 0%]
- "Medium": 이 영역의 열대저기압 발생을 감시하고 있으며, 열대 저기압 발달할 확률이 증가하고 있음. 그러나, 앞으로 24시간 이후에 발달할 가능성이 많아 보임 [발생가능성 20%]
- "High": 이 영역은 24 시간 이내에 열대저기압으로 발달하거나, 혹은 이미 발달하고 있는 단계. 예보자는 이 단계가 되면 "Tropical Cyclone Formation Alert(TCFA; 태풍 발생 경계)"를 발효한다 [발생가능성: 70%].

JTWC에서 위와 같은 분류의 태풍발생 가능성 예측을 하기 위해서, 예보자는 여러 변수를 분석 한다 (그림 2). 이중 위성관측에서 얻은 변수는 ASCAT, 가시, 적외 영상, 레이더 영상을 통해서 분석된 하층바람장의 중심(Low-level circulation center; LLCC), CIMSS 850 mb 와도장, Dvorak T number, CIMSS 상층 발산 분포, AMSU 대기 온도 등이다.



그림 2. JTWC 태풍발생 탐지 단계

나. 일본 기상청 현업 태풍 탐지 분석 방법-EDA에 기반

일본 기상청도 태풍으로 발달할 가능성이 있는 초기 열대 섭동을 분석하고 있으며, 이를 태풍 현업예보에 적용하고 있다. 일본 기상청은 Dvorak(1984) 분석 방법에 근거하여, 독자적인 EDA 방법을 개발하였고, 1999년부터 실험적으로 태풍예보 현업에 적용하였고 결과를 검증하였다(Tsuchiya et al. 2001). EDA는 열대 해상에 조직화된 대류 시스템(Organized Convective System; OCC)이 존재하는지 여부를 파악하여, OCC의 시간 공간 특성을 주관적으로 분석하여 T number를 결정한다(그림 3). 검증결과를 바탕으로, EDA 방법뿐만 아니라 위성자료, 종관자료, 선박자료, 그리고 수치예보 자료를 사용하여 열대저압부 주변의 분석을 실시한다. 위성분석 담당자가 이미 개발된 위성영상 분석 소프트웨어를 이용하여 위성영상으로 열대저기압이 발달할 가능성이 있는 후보 열대 섭동의 위치와 강도를 분석한다. 이 과정은 다른 관측 자료를 포함하지 않고 오직 위성영상 분석만 실시한다. 분석에 사용되는 영상은 적외, 가시 영상 및 다양한 마이크로파 영상으로 분석시간까지 수신되는 모든 위성 영상자료를 다양하게 사용한다. 분석 시간은 1일 4회이며, 분석 요소는 모두 32개로 조사되었다.

다음단계로, 일기도 분석 담당자는 위성분석자가 분석한 열대 섭동의 위치와 강도를 기본으로 하여, 여러가지 관측 자료(바람, 중심기압, 강풍반경 등)를 분석하여, 최종으로 시스템의 위치와 강도를 결정한다. 최종단계로 주임 예보관이 일기 분석 담당자가 분석한 태풍 정보를 받아서 24시간, 48시간, 72시간 태풍 발생 예보를 한다. 이때, 일본 전구 모델과 ECMWF 모델결과도 참고 자료로 이용된다. 일본 기상청의 방법도 JTWC와 마찬가지로 개발한 위성자료 분석 소프트웨어를 활용하더라도 분석자의 주관적인 판단이 개입된다.

표 3. JTWC에서 태풍발달 판단 시 사용되는 예측 인자 및 사용 자료

태풍 발생 인자	사용 자료
Low-level circulation center symmetry	ASCAT, visible and microwave satellite imagery, radar
850 hPa vorticity (VOR)	CIMSS vorticity product
Dvorak T number	Final T number
Global model development (24-48h)	NOGAPS, GFS, UKMO, ECMWF surface wind fields
MJO-associated OLR anomaly	Australian CAWCR OLR anomaly Hovmoller diagram
Vertical wind shear	CIMSS vertical wind shear
Upper level outflow pattern	CIMSS upper-level feature track winds
Core temperature anomaly	CIMSS AMSU

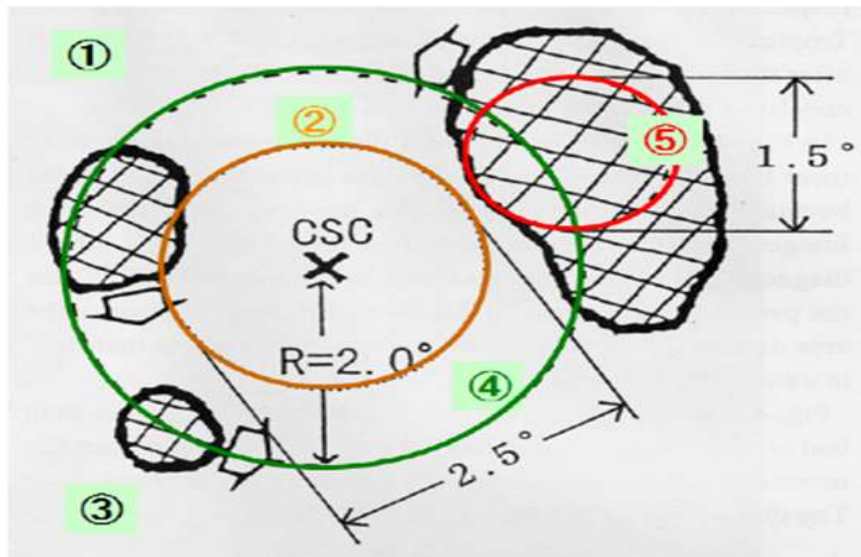


그림 3. EDA의 OCC의 5가지 특징에 따라 T 지수가 결정된다. OCC 지속시간이 12시간 또는 그 이상 지속되어야 한다. CSC는 위도 2.5도 또는 그 규모로 정의된다(주황색 원). CSC는 6시간 또는 그 이상 지속되어야한다. CSC로부터 위도 2도 보다 적은 지역에 발달한 상층 구름대(-31도 또는 그 이하)가 나타난다(초록색 원). dense cloud area는 직경이 위도 1.5도 보다 커야 한다(빨간색 원)

다. CIMSS Advanced Dvorak Technique(ADT)

위에서 언급한 미국과 일본의 현업 태풍발생탐지/예측 기술은 주관적인 Dvorak 분석 기법 등 예보자의 주관적인 결정에 의존하고 있다. 미국 University of Wisconsin-Madison의 CIMSS에서는 정지궤도 위성의 영상을 이용하여 태풍의 강도를 추정하는 Dvorak 분석 기술의 객관화-자동화된 기술인 Advanced Dvorak Technique(ADT)을 개발하였다 (Olander and Velden, 2007). 최근 극궤도 위성 마이크로파 관측을 추가적으로 이용하여, 태풍의 강도의 정확도를 더욱 높이고 있다.

ADT 알고리즘은 객관적인 Dvorak 분석의 개념을 기본으로 하여, 객관화-자동화된 알고리즘이다. 태풍의 강도와 관련된 위성 영상의 패턴을 다양하게 탐지할 수 있도록 여러 가지 인자를 먼저 계산하였다. 예를 들어, 영상에서 태풍의 중심위치를 여러 테스트를 통해서 찾는다. 위성영상을 Log spiral 기법을 활용하여 분석 curved band를 조사하며, 태풍의 눈의 패턴에 따른 eye의 최소 온도, eyewall의 최대온도, eye-eyewall의 온도차 등의 인자들을 산출한다. ADT는 태풍패턴을 분류하는 알고리즘을 회귀방정식(multiple regression equation)을 이용하여, 좀 더 객관적인 위성 관측자료 들의 활용할 수 있도록 하였다.

ADT 알고리즘이 객관화-자동화된 시스템으로 태풍의 강도를 탐지하는 유일한 시스템이나, 태풍 초기 단계에서 정확도가 떨어져서 태풍 발생 단계의 탐지/예측으로 이용될 수 없다(ADT 개발자인 Dr. Chris Velden, 2014년 12월).

라. Colorado State University CIRA

Colorado State University의 Cooperative Institute for Research in the Atmosphere(CIRA)에서는 대서양, 동태평양에서 앞으로 24 시간 혹은 48 시간 이내에 태풍 발생 가능성을 예측하는 시스템을 실시간으로 제공하고 있다 (<http://www.ssd.noaa.gov/PS/TROP/TCFP>). 앞에서 언급한 JTWC, 일본 기상청, CIMMS의

태풍 모니터링 기술은 저기압성 순환을 가진 개개의 열대 섭동 및 저기압에 초점을 두었다. CIRA의 태풍발생 탐지 시스템은 고정된 $5^{\circ} \times 5^{\circ}$ 위·경도 박스 내에서의 태풍 발생 가능성을 예측한다. 이 방법은 전구 수치모델에서 여러 환경 및 대류 조건과 $6.7\mu\text{m}$ 채널 위성 자료를 이용한다. 예측인자는 850-200-hPa 연직바람쉬어, 850 hPa 순환, 연직 대기 불안정도, 850 hPa 발산장, 경도, 지역에서의 육지 면적, 이미 존재하는 태풍과의 거리, 기후학적

해수면온도, 구름 효과가 제거된 수증기채널 밝기 온도, 기후학적 태풍 발생 확률 등이다. 이 예측인자들에서 태풍발생 확률을 추정하기 위해서 다중회귀 기법을 이용하였다.

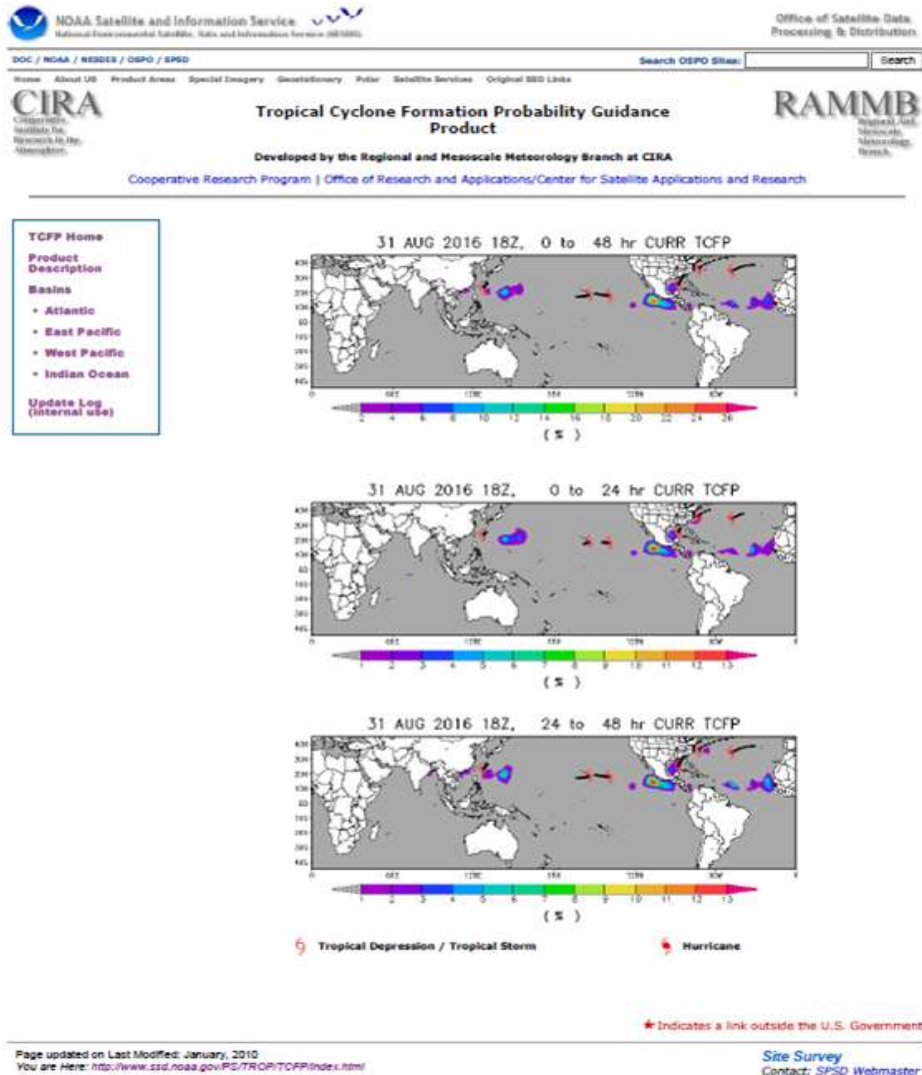


그림 4. CIRA/RAMMB의 태풍발생 가능성 가이드선 산출물

마. University of Arizona: Deviation Angle Variation(DAV) 알고리즘

최근 University of Arizona의 Miguel F. Pineros와 Elizabeth A. Ritchie 교수 연구팀은 정지궤도 위성자료를 이용한 객관적인 방법으로 태풍발생단계를 탐지하려는 연구를 수행하고 있다. 이미 정의된 시스템의 반경주변의 cloud cluster의 조직화(organization)와 비대칭성(Axisymmetry)이 태풍발생과 연관되어 있기 때문에(그림 5), 이를 정량화하여

태풍 발생단계부터 발생할지 여부를 결정하게 한다.

이 선행연구는 Deviation Angle Variation(DAV) 기술을 이용하여 정지궤도 위성의 밝기온도의 공간적 변동성의 통계적 분석을 실시하였다. 이 DAV 기술은 먼저 적외온도 채널의 밝기온도를 직교 좌표계로 변환하여, low-pass filter로 적외 영상 이미지의 noise를 제거한 후, 밝기 온도의 공간 경도 벡터(Gradient vector)를 구한다(그림 6). 그림 6.(c)의 중심처럼 matrix의 높은 값은 라인이 많이 교차하는 것을 의미하고, 이것은 관심 있는 시스템의 중심위치를 찾는 데 도움을 준다. 시스템의 구조가 비대칭성(Axisymmetry)인지 조사하기 위해서, 시스템 중심에서의 radian line을 따라 각 픽셀의 gradient vector의 DAV를 계산하는 것이다.

DAV 알고리즘의 결과는 열대섭동이 열대저압부가 되는 단계인 태풍 발달시간보다 평균 0.6 h 이전에 태풍 발달을 탐지하는 것을 보였다. 이 기술은 특히 발달하는 열대섭동에서 발달하지 않는 열대섭동을 구분해 낼 수 있는 가능성을 보였다. Pineros et al. (2010)은 DAV 알고리즘은 93%의 hit rate와 22%의 낮은 False Alarm rate를 보여 이 방법의 우수성은 강조하였다. 그러나, 탐지 시각은 NHC에서 정의한 TD 발생시간 기준으로 평균 0.6 시간 먼저 탐지하였다. 즉, 93%의 수치는 TD 발달 시점 뿐 아니라 TD 발달 후 (최대 100시간) 탐지까지 포함되어 있어, 실질적으로는 약 53% 만 발달 이전에 탐지를 할 수 있었다.

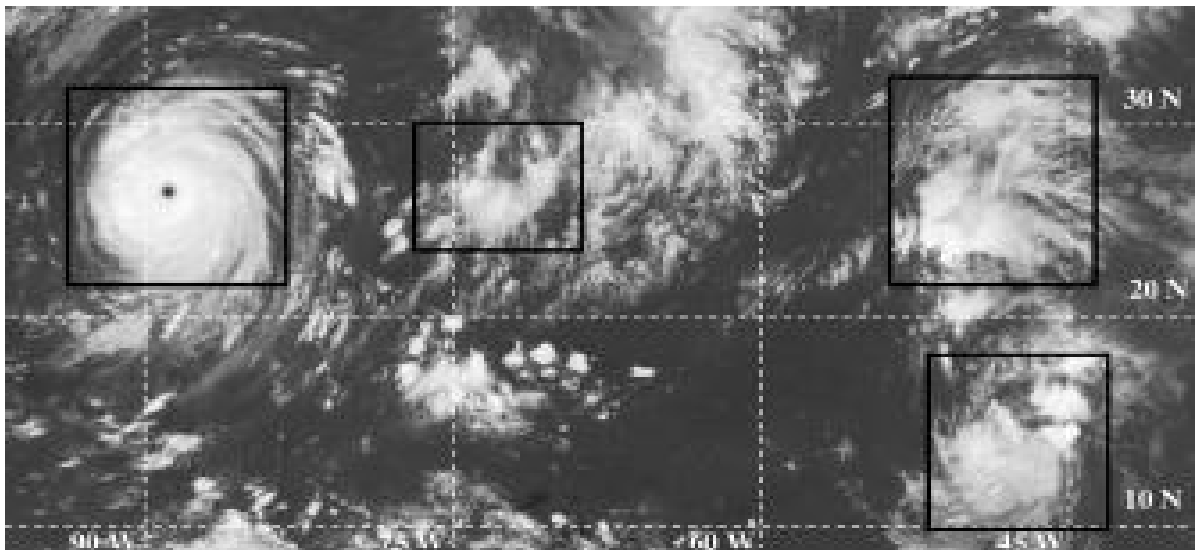


그림 5. 멕시코 만과 서대서양에서의 2005년 8월 28일 1915UTC GOES 적외채널 영상. 제일 좌측 90° W 근처 잘 조직화된 허리케인 Katrina와 덜 조직화된 다른 시스템들

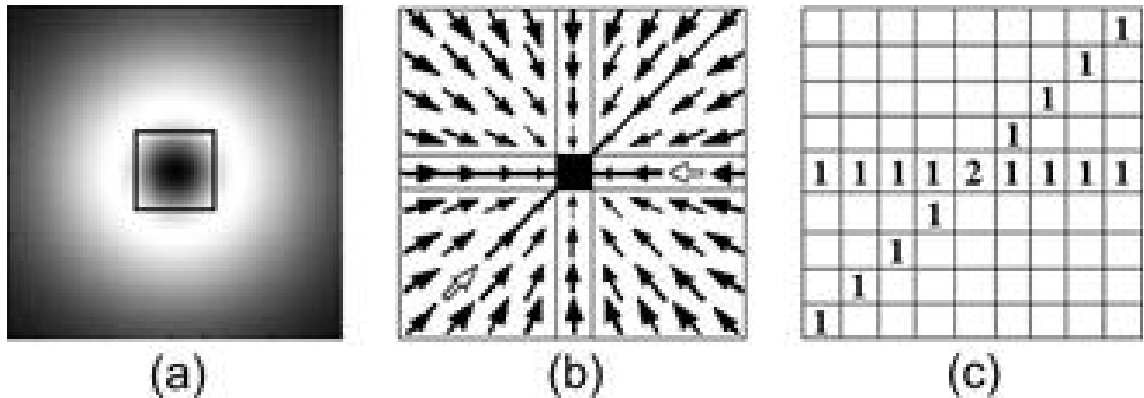


그림 6. University of Arizona 개발된 적외채널 밝기온도의 gradient line density 계산의 예시: (a) 이상적인 와도의 경우 밝기온도의 이미지, (b) gradient vector, (c) density matrix

2. 국내 기술 개발

가. 현재 국내 기상청 현업 (국가태풍센터)

기상청 국가태풍센터는 2011년 열대저압부 분석 시스템을 개발하였으며, 표출 기반을 구축하였다. 2015년 5월부터는 시험적으로 태풍발생 단계 열대저압부(TD) 정보를 제공하고 있다. 제공되는 정보는 현재 시스템의 위치, 24시간 예상 위치, 강도 진행방향, 이동속도 등이다. 예보관들은 열대섭동 단계부터 종관, 위성, 모델 정보를 분석한다. 분석 정보는 850hPa 유선장의 저기압성 흐름, 해수면 온도, Ocean Heat Content, 연직바람 쉬어, 상층 발산장, 계절 내 섭동(Madden Julian Oscillation), ECMWF 예측장 등이다. 분석하는 위성정보는 정지궤도 위성 이미지를 보고 주관적으로 CI & T number를 추정한다. 이 모든 분석 결과를 종합하여 예보관이 태풍발생정보를 제공하게 되는데, 분석 과정은 정성적이며, 최종 태풍발생가능성의 예보 결정과정은 예보관이 주관적으로 결정하고 있다.

현재 태풍발생예보는 여러 표출된 정보들을 이용한 예보관에 판단에 전적으로 의존한다. 또한, 현재 센터에는 태풍발생탐지/예보의 경험이 많은 예보관의 수가 절대적으로 부족하다. 예보관이 바뀌는 경우 예보의 정확도가 바뀔 가능성이 크다. 이런 문제점을 예보관들이 잘 직시하고 있으며, 태풍발생 확률을 객관적으로 결정할 수 있는 알고리즘의 필요성을 제시하였다(2015년 3월 태풍센터 예보관과 인터뷰).

나. 국가 위성센터 태풍 분석 알고리즘 개발

국가 위성센터에는 ADT를 이용한 태풍 최대풍속 및 위치 추정 등 위성영상 분석을 통한 태풍 업무를 수행하고 있다. 그러나, 열대 저압부 발달 이후 시스템에 중점을 주고 있다. 위성 센터에서 직간접 수신한 가용한 위성자료를 이용하여, 위성자료 기반 Subjective/Objective Dvorak 기술을 활용하여 영상을 분석하고 있다. 위성센터에서는 우리나라에 영향을 줄 가능성이 있는 태풍 발생 이후 열대저압부 대한 위성 분석을 실시하며, 결과값은 CI & T number이다. 즉, 태풍 발생 가능성을 산출하는 객관적 알고리즘은 위성센터에 존재하지 않는다(2005년 5월 위성센터 태풍분석 담당자와 이메일 교환).

3. 기존 기술의 취약성 및 과제의 목표 설정

미국 JTWC와 일본 기상청 현업에서 태풍 발생 탐지 및 예보 경보에 사용하는 방법은 예보관의 주관적 결정에 많이 의존하는 기술이다. JTWC의 경우 경험 많은 예보관이 다수 확보되어 있기 때문에, 다소 높은 Hit rate를 보였다. 17개 중 12개의 시스템이 24시간 내 발생확률 높음을 잘 예측하였다(Kucas and Darlow, 2012).

현재 우리나라 기상청 태풍센터도 정성적인 태풍발생분석을 수행하고 있지만, 그 과정이 객관적이지 않고 경험 많은 예보관의 수가 절대적으로 부족하며 예보관들이 기존의 태풍발생분석 시스템에 대한 신뢰성이 부족하였다. 주관적 Dvorak 분석과, 위성의 경우 ASCAT, MTSAT의 정성적인 분석만 수행하고 있으며 정량적인 위성자료 분석은 이뤄지지 않고 있었다. 국가위성센터에서는 TD 발생이후 Dvorak 분석을 실시하고 있으며, 발생 이전 열대 섭동을 분석하여 태풍발생확률을 객관적으로 산출하는 결정 알고리즘은 없었다. 따라서, 본 과제 수행 이전 위성을 이용한 태풍발생탐지 현업에 이용기술 선진국 대비 30% 수준으로 판단되었다.

Pineros et al. (2011)에서는 정지궤도 위성 DAV 방법을 이용한 태풍 발생 탐지 기술 개발을 발표하였다. 이 기술을 이용한 태풍 발생 탐지율은 93% 이며, False Alarm은 약 30%로 그 우수성을 강조하였으나, 실질적으로 태풍 발생 시간 전 Hit rate는 약 50%로 높지 않았다. 따라서, 본 과제 수행 이전 객관화-자동화된 태풍발생탐지 기술 선진국에서 개발 및 현업화하고 있는 단계이나, 우리나라는 객관적 태풍발생탐지 기술은 존재하지 않았다.

제 3장. 연구 수행 내용 및 성과

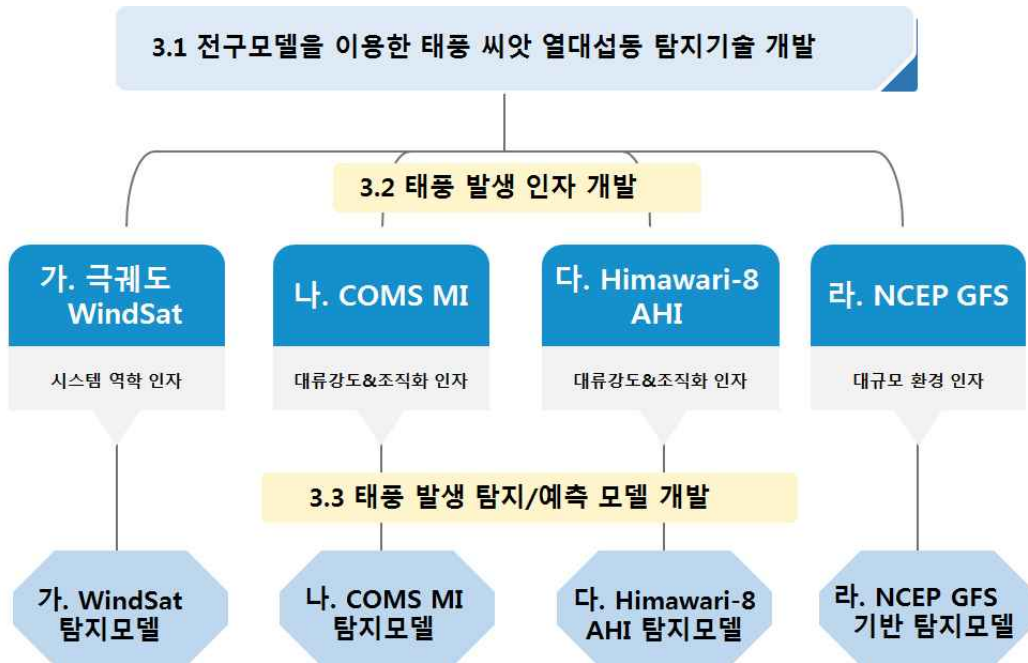


그림 7. 태풍 씨앗 열대섭동 탐지 기술 개발 개요도

본 장에서는 세 가지 단계로 연구 수행 내용에 대해 먼저 서술한다. 전구모델을 이용한 태풍 씨앗 열대섭동 탐지 기술 개발은 3.1절, 태풍 발생 인자 개발에 대해서는 3.2절, 태풍 발생 탐지/예측 모델 개발에 대한 내용은 3.2절에 서술한다. 태풍 발생인자는 태풍의 시스템 역학, 대류강도와 조직화 정도, 대규모 환경 인자들을 모두 고려하였다. 극궤도 WindSat, COMS MI, Himawari-8 AHI, NCEP GFS 기반 태풍발생 인자 개발 결과에 대해서는 3.2절의 가, 나, 다, 라 섹션에서 각각 논의한다. 각 인자를 바탕으로 기계학습 방법을 이용하여 개발된 태풍발생탐지/예측 모델들은 3.3절의 각 섹션에서 논의한다.

1. 태풍 씨앗 열대섭동 탐지 기술 개발

태풍의 발생을 탐지/예측하기 위해서는 태풍 이전의 열대 섭동에 대한 탐지가 필수적이다. 특히 과거 태풍들에 대한 관측 및 분석 자료(best track 자료)에서는 태풍 발생 이전에 대한 정보가 매우 제한적이며 그 자료의 양도 매우 적다. 또한 false alarm을 줄이기 위해서는 과거 태풍으로 발달하지 않는 열대 섭동에 대한 분석이 필수적인데, 이를 위해서는 태풍으로

1) http://rammb.cira.colostate.edu/products/tc_realtime/index.asp

발달하는 열대 섭동과 비슷한 특성을 가지고 있지만 태풍으로 발달하지 않은 열대 섭동 시스템에 대한 정보가 필요하다. Hennon et al. (2013)에서는 위성 관측을 기반으로 열대 구름 군집 자료를 생산하여 열대 섭동에 대한 정보를 생산하기도 하였다. 그러나 이 자료는 위성에서 관측된 구름 자료에 기반하고 있어, 열대 대류의 일 변동성 등에 의해 구름 군집 트랙의 연속성이 떨어지는 경우가 많다. 또한 초기 열대 섭동과 하층와도 및 주변 대기장의 특성을 적절히 반영하지 못하는 단점을 가지고 있다. 본 연구에서는 현업적으로 쉽게 사용 가능하며 균일한 시간 간격으로 연속적인 추적이 가능한 전구 수치모델 분석장을 이용하여 열대지역에서의 태풍으로 발생 가능한 열대섭동을 탐지하는 알고리즘을 개발하였다.

태풍 발달 이전의 열대 섭동은 태풍의 주된 특성인 warm core가 형성되어 있지 않기에, 저기압성 바람의 회전을 기준으로 추적이 되어야 한다. 우선 열대섭동의 특성을 살펴보기 위해 그림 8에서는 2008년 6월 18일 12UTC에 열대 섭동에서 TD(Tropical Disturbance, 열대 저압부)로 성장한 태풍 Fengshen의 열대 섭동의 GFS 분석장의 700hPa 바람장 및 상대와도를 보여주고 있다. 열대 섭동단계에서는 바람장은 원형 형태의 닫힌 저기압 회전의 모습은 매우 약하게 나타나고 보다 넓은 범위에서 열린 저기압 형태로 회전한다. 바람의 회전을 나타내는 지표로 주로 사용되는 것이 상대와도이다. 그림 8에서 음영으로 표시된 상대와도를 살펴보면 태풍으로 성장하였을 때는 상대와도의 중심이 태풍의 회전의 중심과 일치하게 나타난다. 그러나 열대섭동 단계에서는 강한 풍속 지역이 회전의 중심에서 먼 지역에서 나타나므로 상대외도는 회전의 중심보다 바람의 강도가 강한 회전의 가장자리에서 여러 지역에 분산되어 나타나는 경우가 많다. 그리하여 상대와도는 열대섭동의 존재 지역을 확인할 수 있는 중요한 지표로 사용될 수 있으나, 그것만으로 열대 섭동의 중심을 추적하기에는 어려움이 따른다.

본 연구에서는 이러한 문제점을 보완하기 위해 회전하는 벡터의 조직화 정도를 나타내는 지표인 Circular Variance(CV)를 함께 사용하였다. CV는 다음과 같이 정의된다.

$$CV = 1 - \frac{\sqrt{\left(\sum_{i=1}^n \cos\theta_i\right)^2 + \left(\sum_{i=1}^n \sin\theta_i\right)^2}}{n}$$

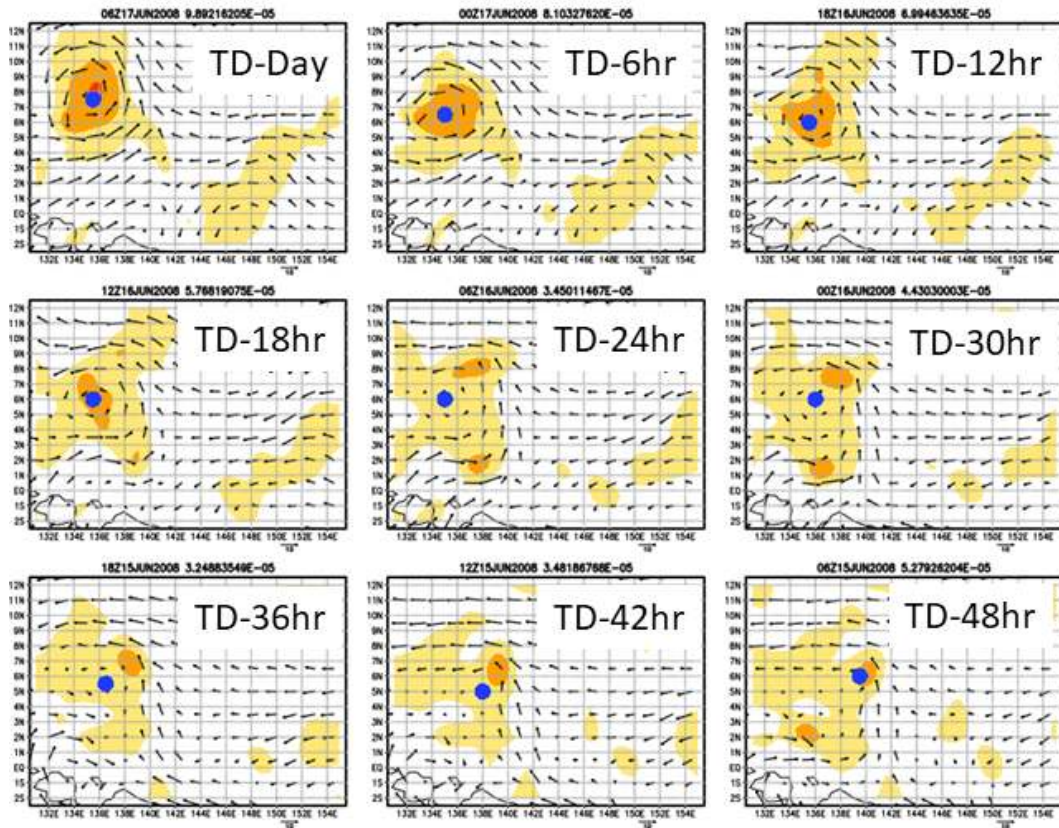


그림 8. 2008년 태풍 Fengshen의 TD 발생이전 열대 섭동, 700-hPa 바람(벡터) 및 상대와도(음영), 파란색 점은 열대 섭동의 중심으로 추정되는 지점을 나타냄

여기서 θ 는 각 지점의 바람 벡터의 방위각을 뜻한다. CV는 0과 1사이의 값을 가지며, 벡터의 방향만을 이용하여 산출되기 때문의 풍속은 CV의 산출에 관여되지 않고, CV 계산 영역에서 바람 벡터의 회전 성분이 강할수록 그 값이 크게 나타난다. 이를 이용하여 모델 분석장의 700hPa 바람장을 이용하여, 각각의 모델 격자를 중심으로 4x4도 범위에서 CV를 산출하였다. 그림 9은 열대 섭동 주변에서 상대와도, CV, 그리고 상대와도와 CV를 곱한 결과를 나타낸 그림이다. 그림에서 볼 수 있듯이 상대와의도 경우에는 회전의 중심보다 바람의 강도가 강한 지역에서 최댓값을 가지며, CV는 바람의 강도와 관계없이 바람 벡터가 회전하는 형태가 나타나는 지역에서 최댓값을 가진다. 그러나 CV의 경우에는 풍속을 고려하지 않으므로 국지적인 바람의 회전이 나타나는 지역이 존재하는 경우 풍속과 관계없이 그 값이 커지는 단점을 가지고 있다. 또한 열대 섭동처럼 열린 저기압 형태의 회전, 비대칭적 회전이 나타나는 경우 그 중심을 정확히 포착하는데 어려움이 발생한다. 이를 보완하기 위해 본 연구에서는 상대와도와 CV를 모두 활용하기 위해 상대와도와 CV를

곱한 값을 이용하여 최종 열대 섭동 추적 알고리즘을 구성하였다. 그림 9에서 상대와도와 CV를 곱한 결과를 보면 열대섭동의 회전의 중심에서 그 값이 최대가 되어 열대 섭동의 중심을 적절히 판단할 수 있음을 확인할 수 있다.

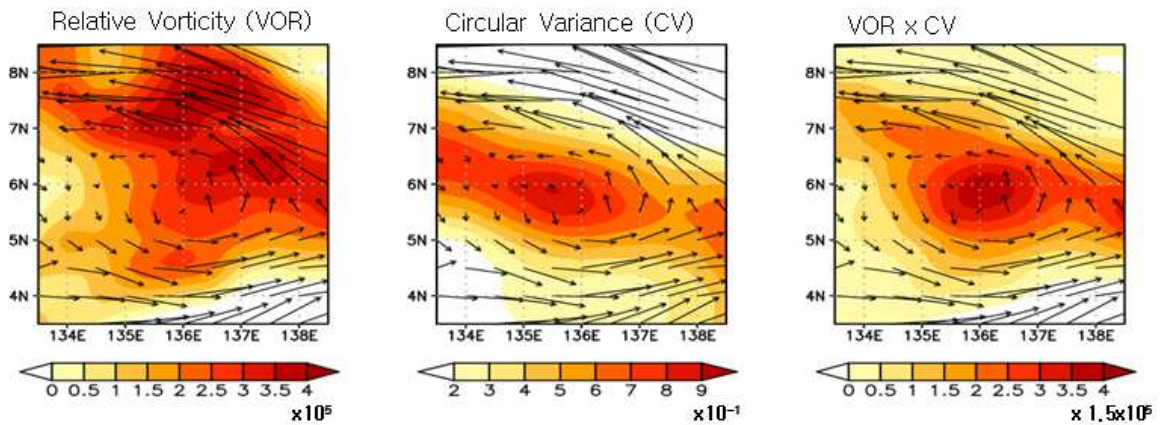


그림 9. 열대섭동 주변의 바람장(벡터)와 회전을 나타내는 지표(음영: 상대와도(좌), CV(중), 및 상대와도와 CV의 곱(우))

알고리즘 개발에 시험적으로 사용된 수치모델 분석장은 미국 NOAA에서 제공하는 GFS 예측 모델 분석장이다. 이 자료는 실시간으로 인터넷을 통해 입수가 가능하여 연구 및 현업 활용에 용이하다. GFS 분석장 자료는 2007-2015년 기간의 6-10월의 북서태평양지역(90E-180E, 0-30N)에서 700hPa 상대와도와 CV 곱(VOR×CV)의 중심을 추적하여 태풍으로 발달한 열대 섭동을 추적하였다. 열대 섭동 추적 알고리즘은 기존의 발생한 태풍의 TD 발생일을 기점으로 열대 섭동을 역추적 하는 알고리즘을 우선적으로 개발하였으며, 이를 비슷한 방법으로 북서태평양 지역에서 발생하는 열대 섭동을 모두 추적하는 알고리즘을 개발하였다.

본 연구에서 개발된 태풍으로 발달한 열대 섭동 역추적 알고리즘은 JTWC Best-Track 자료를 이용하여 TD 발생 지점으로부터 6시간 이내 400km 이내의 700hPa 바람장의 VOR×CV가 주변 반경 2.5도 이내의 지역의 값보다 높은 값을 갖는 지역에서 최댓값을 갖는 지역을 추적하여 연결한다. 만약 VOR×CV 최대값 지역 3도 반경 이내에 $2.5 \times 10^{-5} \text{ s}^{-1}$ 이상의 상대와도가 존재하지 않는 경우에는 열대섭동이 존재하지 않는 것으로 판단하여 추적을 중단한다. 0.5도 이하의 고해상도의 자료에서 상대와도는 하나의 격자에서만 강한 값을 갖는 경우(bull's eye)가 발생하기도 하는데, 이는 열대 섭동을 추적하는데 노이즈로 작용될 수 있다. 이를 해결하기 위해 상대와도를 계산할 때 5도 반경으로 가우시안 필터를

적용하여 앞서 언급한 bull's eye 현상을 제거하였다. 알고리즘에 사용된 여러 기준을 정리하면 다음과 같다.

- VOR×CV 계산 시 상대와도 5도 반경으로 가우시안 필터 적용
- VOR×CV 가 주변 2도 이내지역에서 최댓값을 가질 것
- VOR×CV 최댓값 지점에서 반경 3도 이내에 $2.5 \times 10^{-5} \text{ s}^{-1}$ 이상의 상대와도가 존재할 것
- 열대섭동의 최대 이동 거리는 6시간마다 400km

그림 10는 태풍으로 발달한 열대 섭동을 역추적한 결과 일부를 보여주고 있다. 그림에서 확인할 수 있듯이, 개발된 열대 섭동 역추적 알고리즘은 JTWC Best-Track에 기록된 열대 섭동의 위치와 매우 근접하게 추적하고 있음을 확인할 수 있다. 또한 JTWC Best-Track에 기록된 열대 섭동 보다 더 오랜 시간의 열대섭동의 위치를 추적하고 있음을 확인할 수 있다.

앞선 태풍으로 발달한 열대 섭동을 역추적하는 알고리즘을 바탕으로, 태풍으로 발달하지 않는 열대섭동을 포함하여 북서태평양 지역에서 발생하는 전체 열대 섭동 추적하는 알고리즘을 개발하였다. 전체 열대 섭동을 추적하기 위해서 우선 북서태평양 지역의 700hPa VOR×CV가 반경 2도 이내의 지역의 값보다 높은 값을 갖고, VOR×CV 최댓값 지역 2.5도 반경 이내에 $2.5 \times 10^{-5} \text{ s}^{-1}$ 이상의 상대와도가 존재하는 경우 잠재 열대 섭동으로 정의하여 그 위치를 모두 기록하고, 6시간 이내 400km이내의 잠재 열대 섭동 중 VOR×CV값이 가장 큰 값을 잠재 열대 섭동을 연결하였다. 이렇게 연결된 잠재 열대 섭동이 48시간 이상 연결되는 경우 실제 열대 섭동으로 정의하여 열대 섭동 중심 추적 자료로 저장한다.

2007년-2015년 기간 동안 5월-10월에 GFS 분석장을 기반으로 북서태평양에서 발생한 열대섭동 중 태풍으로 발달한 열대 섭동 167사례 및 비발달 열대 섭동 352사례의 위치를 추적할 수 있었다. 알고리즘을 검증하기 위해서는 과거 열대 섭동 자료가 필요한데 열대 섭동에 대한 객관적인 관측 자료는 부재한 상황이다. 해외 현업 예측기관에서는 태풍으로 발달할 수 있는 열대 섭동에 대해 추적을 실시하여 INVEST 시스템으로 위치정보를 제공하고 있으나 자료는 저장되지 않고 시스템이 태풍으로 발달하거나 소멸하면 자료가 삭제된다. 본 연구에서는 RAMMB/CIRA 의 Current Active Tropical Cyclone 웹페이지¹⁾에서 제공하는 태풍으로 발달 가능성이 있는 열대섭동자료인 INVEST자료를 실시간으로 수집하여 검증하였다. 본 INVEST 자료는 발생 시기에만 공개되고 시스템이 소멸한

경우에는 저장되지 않고 삭제되기 때문에 2015년 이후자료만 수집하여 검증할 수 있었다. 중심 위치에 대한 검증결과 평균 거리 제곱 오차가 0.78도로 사용한 모델의 해상도(0.5도)를 고려하였을 때 본 연구에서 개발한 열대 섭동 위치추적 알고리즘이 우수하게 열대 섭동의 위치를 산출 하고 있음을 알 수 있었다.

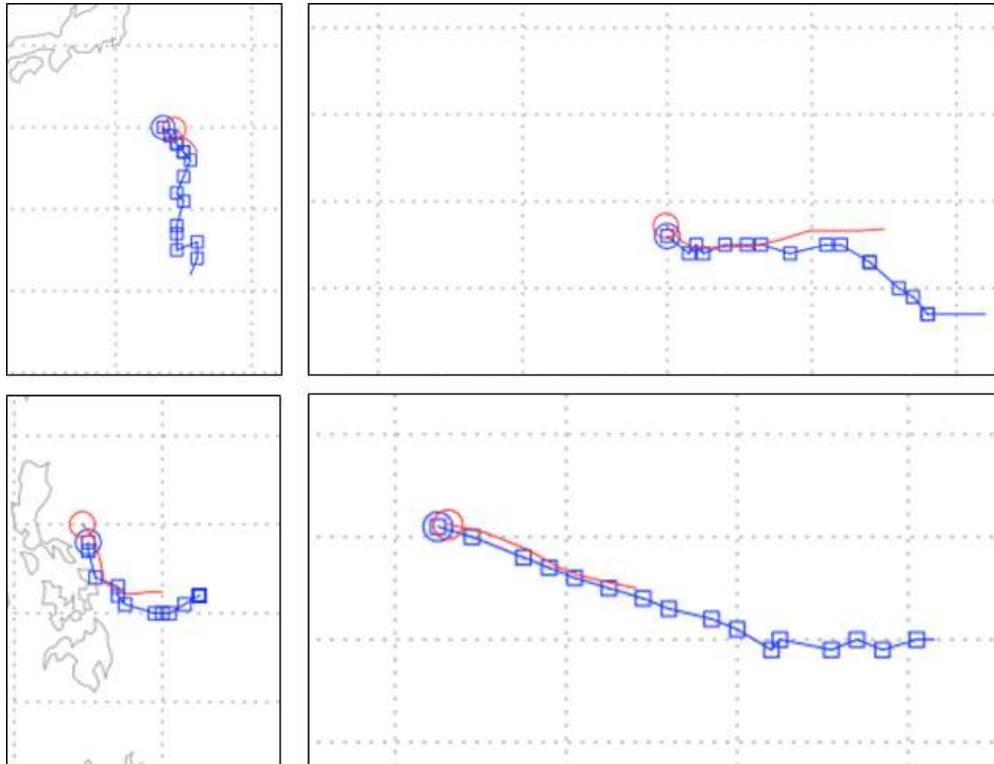


그림 10. 열대 섭동 추적 알고리즘을 활용하여 TD 발생 이전 열대 섭동 추적 결과 (청색)와 JTWC Best-Track에 기록된 열대 섭동 트랙(적색)

실제 태풍 발생 예측에 활용하기 위해서 보다 정확한 분석장을 사용하기 위해 보다 더 많은 관측자료가 자료동화되어 있는 FNL 분석장을 이용하여 열대섭동 자료를 재생산 하였다. FNL 분석장은 1×1도의 수평해상도를 가지고 있으며 6시간의 주기로 생산된다. GFS보다 약 한시간 늦게 제공되지만 태풍 발생 예측을 위한 실시간 자료로 사용될 수 있다. GFS 분석장은 수평해상도 0.5도의 상대적인 고해상도를 가지고 있으나, 상대와도를 활용한 열대 섭동 추적에 있어 한,두개의 격자에서만 강한 상대와도 값을 갖는 경우(bull's eye)가 자주 발생하여, 열대 섭동을 추적하는데 노이즈로 작용되고 있었다. 이를 가우시안 필터를 적용하여 사용하면 . 이를 해결하기 위해 상대와도를 계산할 때 5도 반경으로 가우시안 필터를 적용해보기도 하였으나, 그 경우 상대와도의 값이 다소 낮아져 앞서 설정한 상대와도 기준을 충족하지 못하는 경우가 발생하기도 하였다. 그러나 FNL의 경우 상대적으로

해상도가 낮아 앞서 언급한 bull's eye 현상으로 인한 노이즈가 상대적으로 덜 발생하여 보다 더 안정적으로 열대 섭동을 추적할 수 있었다. 그래서 향후 태풍 발생 모델 개발에서는 FNL 기반으로 생산된 열대섭동 자료를 주로 사용하였다. FNL 분석장을 이용하여 GFS 분석장 기반으로 개발된 알고리즘을 그대로 적용하여 사용하였으며 2007년부터 2017년 총 11년간 5월부터 11월 사이에 48시간 이상 지속된 열대섭동을 추적 생산하였다. 태풍으로 발달한 열대 섭동은 총 243사례, 태풍으로 발달하지 않은 열대섭동은 총 1,068사례를 추적하였다.

그림 11은 FNL 자료에서 추출된 열대섭동의 공간적 활동 분포를 나타내고 있다. 태풍으로 발생한 열대 섭동의 경우 필리핀해 동남쪽 지역에서 가장 많이 발생하는 것으로 나타나며, 태풍으로 발달하지 못하고 소멸한 열대 섭동의 경우 남중국해와 필리핀해 남쪽에서 주로 활동하는 것으로 분석되었다. 특히 남중국해에서는 많은 열대 섭동이 발생하지만 실제 태풍으로 발생하는 경우는 상대적으로 적은 것으로 나타나고 있다. 그리하여 전체 열대섭동 중에 태풍으로 발달하는 열대 섭동의 비율은 필리핀해에서 높게 나타나고 있다. 이를 바탕으로 향후 태풍 발생 모델 개발에서 열대 섭동의 위치 정보도 태풍의 발생 가능성을 탐지하는 데 있어 주된 지표로 사용될 수 있음을 암시해주고 있다.

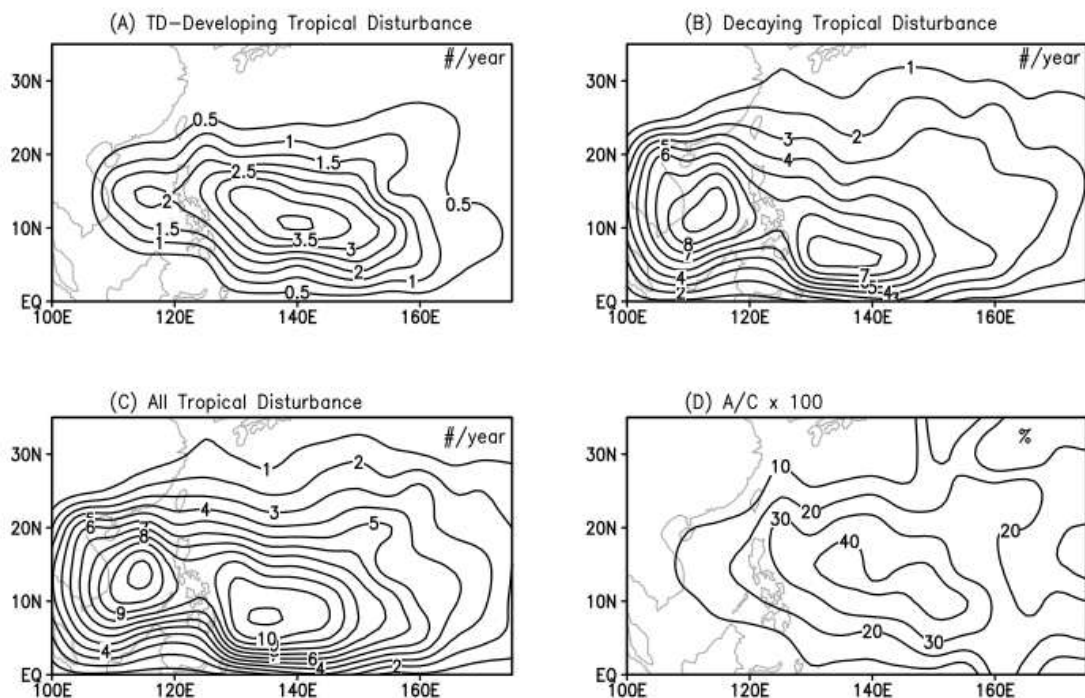


그림 11. 2007-2017년 5-11월 기간에 FNL 분석장 자료에서 탐지된 열대 섭동의 활동 위치 분포 (A: 태풍으로 발달한 열대 섭동, B: 태풍으로 발달하지 않고 소멸한 열대 섭동, C:전체 열대 섭동, D: 전체 열대 섭동 중 태풍으로 발달하는 태풍의 비율)

2. 태풍발생 인자 개발

현상학적으로 태풍의 발생과정은 대기와 해양의 역학, 열역학, 대류과정의 복합적인 변화를 동반한다(그림 12). 태풍은 저기압성 순환의 최대 와도가 하층에 있는 시스템으로(warm core system), 최대 풍속이 기준치(Tropical depression : 20 kts, Tropical storm : 35 kts)를 넘어야 하고, 열역학적으로 warm core system으로 중심에 주변보다 높은 온도를 동반한다. 또한, 최근 학계에서는 태풍으로 성장하는 섭동과 성장하지 않는 섭동의 차이는 중규모 순환 내에서 성장하는 중규모 대류계(mesoscale convective system)의 잠열 방출과 그에 따라 생성되는 저기압성 와도의 구조에 있다고 강조되었다. 고해상도 수치모델링을 통해서 Montomgery et al.(2006)은 태풍이 생성 직전에, 대류성 상승기류 지역에서 다수의 vortical hot tower가 생성되며, 여러 저기압성 와도를 가진 대류 셀 간의 거리가 서로 가까워져서 결국 병합(merging)하면서 up-scale 과정을 겪으면서 태풍으로 발달함을 보였다.

이와 같이 태풍발생 시 다양한 변화를 잘 반영할 수 있는 발생탐지 알고리즘을 개발하기 위해서, 본 연구에서는 정지궤도 위성의 적외, 수증기 채널 자료와 극궤도 마이크로파 관측의 해상도 자료를 이용하여 태풍발생 탐지/예측 인자를 개발하였다.

태풍 발생과 관련된 바람장, 대류, 온도 등 역학-열역학-수문 과정의 강도변화, 경압성, 조직화, 대칭성들을 정량화하였다. 위성정보의 공간적 변동성을 보다 잘 표현하기 위해서 지리정보체계 및 공간 모델링에서 많이 활용되고 있는 다양한 공간지수 개념을 적용하였다. 예를 들어, 원형분산(CV)은 방향성이 있는 바람장의 특성을 잘 나타내기 때문에 태풍 발생 여부 판단에 중요한 변수로 활용할 수 있다. CV 방법을 이용하여 Windsat 자료의 바람패턴을 정량화하는 인자는 이미 개발되었다.

또한 Contagion과 같은 경우는 비슷한 형태의 바람장이 공간적으로 어떻게 분포하고 있는지를 정량화해 주는 지수로, 이 역시 앞으로 태풍으로 발생할지 말지를 판단하는데 도움을 줄 수 있는 변수이다. 이처럼 다양한 공간적인 변수들을 위성자료로부터 추출하여 이를 바탕으로 태풍 발생 여부를 판단하는 모델을 구축하였다. 이 외에도 활용 가능한 공간 변수들은 기상학적으로 균질한 그룹을 하나의 패치라고 할 때 패치 크기와 밀도, 모양 복잡도, 근접도, 대비, 분산 등 다양하게 존재하면 우선 각각에 대해 상관성 분석을 수행하여 태풍 발생 탐지/예측에 효과적인 변수들을 선별하였다. 우선 다양한 위성자료로부터 기상학적 패치들을 추출한 뒤에 이로부터 다양한 공간 변수들을 계산하였고, 여러 분야에서 경관 및 공간분석을 수행하는 데에 많이 사용되는 FRAGSTATS

(<http://www.umass.edu/landeco/research/fragstats/fragstats.html>) 틀을 사용하였다.

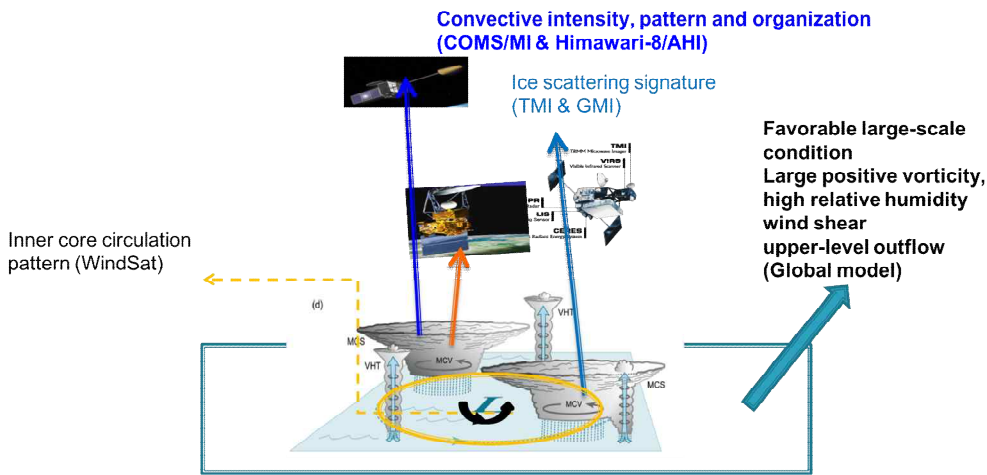


그림 12. 대규모 환경, 시스템 구조, 중규모 대류계의 강도/조직화 등 여러 복잡한 과정을 통해서 태풍의 발생이 결정됨: 태풍발생 탐지를 위해 위성자료와 지구 모델의 융합이 필요한 이유

가. 극궤도 Windsat 해상풍 자료 이용 인자 개발

(a) 자료

극궤도위성의 해상풍 자료를 이용한 태풍 탐지 인자 개발의 장점은 태풍 발생의 역학적 특성을 정량화 할 수 있다는 것이다. 현재의 태풍 발생 정의 방법은 주로 저기압성 순환장의 역학적인 강도와 패턴을 고려한다. 공식적으로, 미국 JTWC 에서는 열대 서태평양에서 닫힌 저기압성 순환 시스템 내의 최대지속풍속(Maximum Sustained Wind; MSW)가 25 knots가 넘을 때 태풍발생을 정의한다. MSW는 JTWC 뿐 아니라 National Hurricane Center 등 다양한 태풍 현업기관에서 이용되는 중요한 메트릭이다.

특별한 태풍 관측 프로젝트가 진행되지 않는 대부분의 경우, 태풍이 생성되는 넓은 바다에서 현장 관측 자료를 얻는 것이 쉽지 않다. 따라서, 역학적인 태풍 변수인 MSW는 매시간 가용한 적외/가시채널 영상을 통한 구름 상부의 특성에 의해서 추정된다. Dvorak 방법은 적외 채널에 기반을 두어 구름 운정온도를 이용하여 대류의 강도, 구름 패턴의 뒤틀림의 정도 등을 고려한다. 정지궤도 위성에 탑재된 적외채널 영상은 매시간 혹은 더 빈번하게 이용 가능하다는 장점이 있다. 그러나, 적외채널 위성 영상은 상층의 구름이 있을 경우, 태풍의 발생 정의에서 가장 중요한 하층 바람장을 탐지에는 민감하지 않다.

마이크로파 원격탐사는 구름과 강수입자의 방출과 산란을 기반으로 이뤄지는데, 연직/수평 편광마이크로파 복사량의 차이는 해양의 표면에 거칠기에 민감함 것으로 알려졌다. 따라서,

마이크로파 위성 영상은 해상의 바람(예를 들어, 10 m 높이 바람)을 산출하는데 이용 될 수 있다(Gaiser et al. 2014). 또한, WindSat과 같은 마이크로파 복사계는 연직, 수평 방향 뿐 아니라, 다양한 방향으로 편광된 복사량을 측정함으로써 바람 방향도 산출할 수 있다. 본 연구는 WindSat 자료를 이용하여 태풍의 발생과 연관되어 있는 역학적인 해상의 바람 패턴, 조직화, 강도를 정량화하는 인자들을 정의한다.

WindSat은 미국 해군 연구소에서 개발한 최초의 수동 마이크로파 복사계로, 코리올리 위성에 탑재되어 2003년부터 현재까지 운용되고 있다. WindSat 센서는 전방/후방 스캔을 수행하기 때문에, Advanced Scatterometer(ASCAT)보다 훨씬 넓은 약 1000 km 정도의 넓은 관측폭을 가지고 있다. 넓은 관측폭으로 인해 태풍으로 발달하는 시스템 규모의 위성 영상을 훨씬 자주 샘플링 할 수 있는 장점이 있다(그림 13).

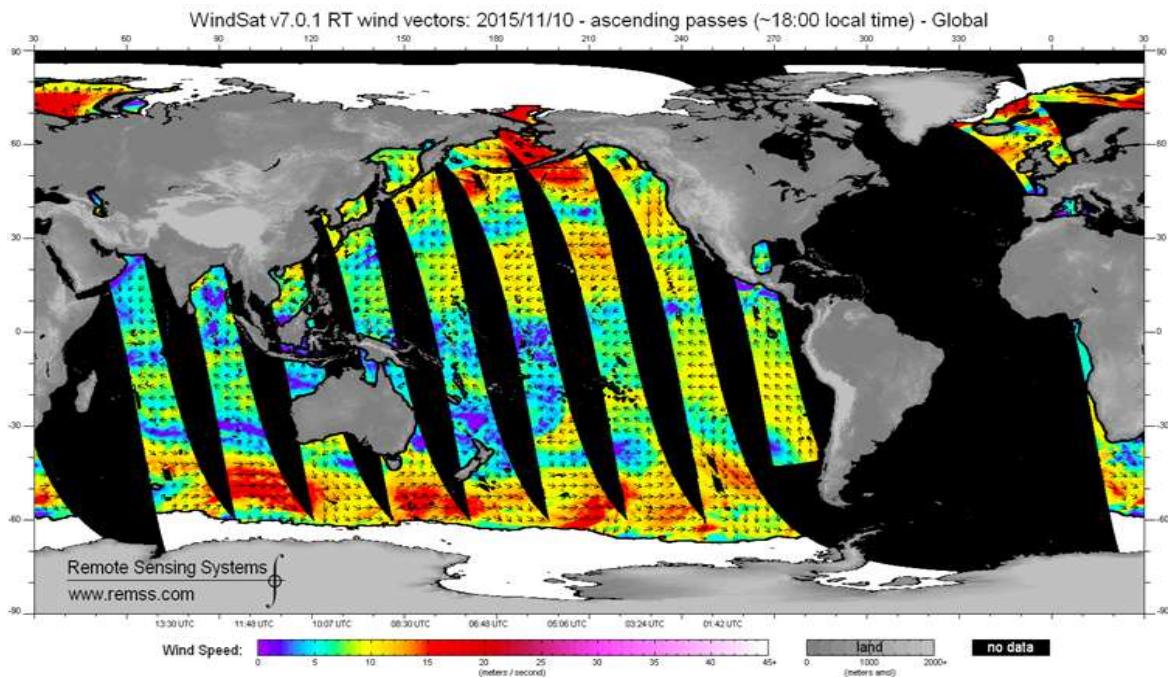


그림 13. 본 연구에서 이용한 전구에서의 WindSat 해상풍 자료: WindSat 위성의 경우 관측 폭(Swath)이 넓어서 태풍 씨앗 시스템 영상을 분석하는데 적합함

WindSat은 6.8, 10.7, 18.7, 23.8, 37.0 GHz 채널에서 관측을 실시하며, 공간 분해능은 37 GHz 채널의 경우 8×13 km 이며, 10.7 GHz의 경우 25×28 km 이다. 본 연구에서는 Remote Sensing System archive (<http://www.remss.com/missions/windsat>)에서 다운 받은 $0.25^\circ \times 0.25^\circ$ WindSat 격자 자료를 이용한다.

본 연구에서는 WindSat 풍속, 풍향, 강수량 위성자료를 Henon et al. (2011)의 Tropical

Cloud Cluster 자료의 위치를 중심으로 상대 좌표로 이동한 2차원 위성 영상을 분석한다. 그림 14에 보이는 것처럼 시스템 중심의 $16^\circ \times 16^\circ$ 위경도 영역의 여러 영상을 구성하여, 이중 위성의 샘플링 면적이 60% 이상인 영상만 분석 샘플로 사용하였다. 본 연구에서는 2005년부터 2009년까지 총 1325개의 WindSat 영상 자료를 구성하였다. 이것은 352개의 발달 시스템과 973개의 비발달 시스템의 영상으로 구성된다. 각 위성자료의 정량적인 영상분석을 통하여, 태풍 발생과 연관된 인자를 추출한다.

(b) 개발된 인자들

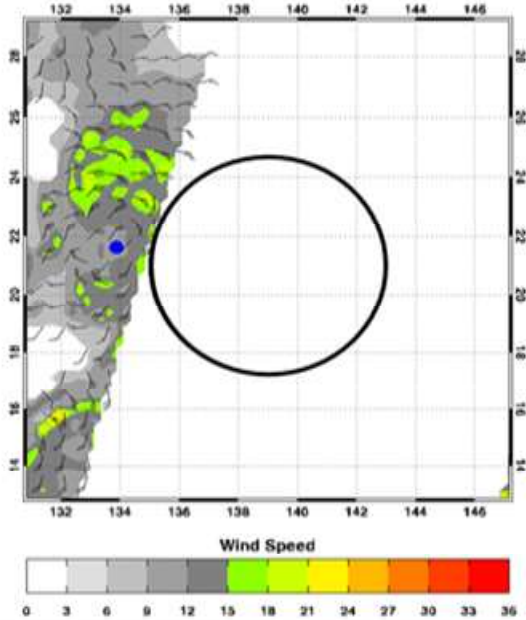
본 연구에서는 WindSat 위성 해상풍 자료를 이용하여 태풍 발생과 관련된 8개의 인자를 개발하였다. 이는 이후 기계학습의 입력 자료로 이용된다. 이 중 5개 인자(wind_ave, wind_cv_fix, wind_cv_mv, wind_ci, wind_pladj)는 시스템의 강도, 바람 순환의 대칭성, 강풍 영역의 조직화 정도를 나타내는 역학 인자이며, 3개 인자(rain_ave, rain_ci, rain_pladj)는 대류의 강도와 강수의 조직화 정도를 나타내는 수문학 인자이다.

발달하는 시스템의 강도를 나타내기 위해서, 시스템 강도 바람장을 WindSat 위성자료를 이용하여 정의하였다. WindSat 평균 풍속 인자(wind_ave)는 지정된 4° 반경의 중심원상에서 가용한 모든 영역의 풍속을 평균한 인자이다. 또한, 본 연구에서는 3.1절에 설명된 원형분산(CV) 개념을 이용하여 발달하는 시스템의 저기압성 순환의 Symmetry(대칭성)를 정의하였다.

본 연구에서는 종관규모 순환의 Symmetry와 중규모 순환의 Symmetry를 동시에 고려하기 위해서 그림 14의 두 박스 영역에서의 원형분산 인자를 정의한다. 발달하는 시스템의 종관 규모의 순환의 Symmetry를 정량화하기 위해서, 고정된 큰 위경도의 $16^\circ \times 16^\circ$ 넓은 박스 영역 내에서의 WindSat 풍향값들을 이용하여 얻어진 원형 분산을 wind_cv_fix로 정의한다. 또한, 시스템의 중규모 순환의 Symmetry를 정량화하기 위해서 $4^\circ \times 4^\circ$ 위경도 박스를 점차적으로 옮겨 가면서 풍향의 원형분산값을 계산하여, 이 중 최대값 wind_cv_mv로 정의한다.

(a) Bad passage

WindSat 2008/08/12/20Z - VONGFONG



(b) Good passage

WindSat 2005/07/20/20Z - BANYAN

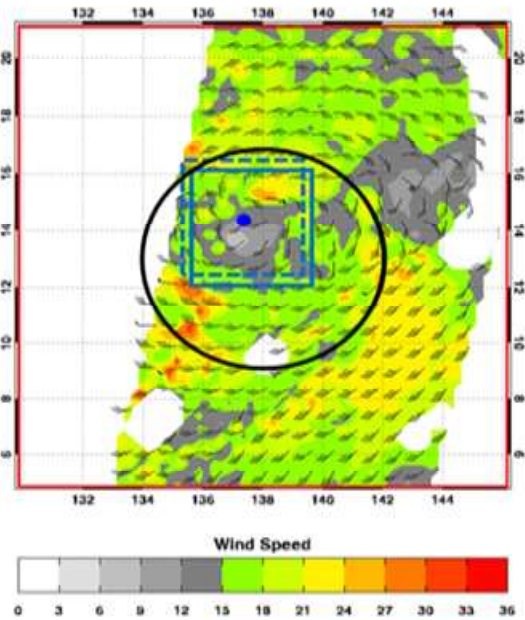


그림 14. 시스템 주변의 극궤도 위성의 WindSat 관측 범위 예시: (a) 좋지 않은 관측 범위, (b) 좋은 관측 범위. 위성 인자 산출을 위한 세 가지 주된 관측 영역: 미리 정의된 시스템 중심에서 4° 반경의 원 (흑색 원), 16° × 16° 크기의 고정된 위경도 박스 영역 (적색 사각형), 크기가 작은 움직이는 4° × 4° 크기의 움직이는 위경도 박스 영역 (청색 사각형) (Park et al. 2016)

표 4. WindSat 기반 태풍발생 인자 추출과 이후 기계학습 구성에 이용된 발달/비발달 샘플 수 (Park et al. 2016)

Period	Dev	Non-Dev	Total
Training (2005-2007)	206	519	725
Hindcast validation (2008-2009)	146	454	600
Total period (2005-2009)	352	973	1325

이론적으로 원형분산은 0에서 1의 값을 가진다. 원형분산이 1에 가까운 값은 영역 내에서 풍향이 다양함을 나타내서, 본 연구에서는 시스템이 잘 조직화되어 있음을 나타낸다. 원형분산이 0에 가까운 값은 풍향이 거의 비슷한 것을 나타내며, 이는 시스템이 조직화 되지 않음을 나타낸다.

그림 15는 실제 여러 발달 단계의 WindSat 영상과 wind_cv_fix, wind_cv_mv 값을 비교하였다. 2006년 태풍 ShanShan의 경우 열대저기압 단계에서 저기압성 순환이 잘 조직화되어 있으며, 정의된 영역에서 바람의 방향이 다양하게 변화하는 것을 확인할 수 있다. wind_cv_fix, wind_cv_mv 값이 거의 1에 가까운 0.982, 0.928로 확인되었다. 2009년 태풍 Morakot이 태풍으로 발생하기 전 섭동 단계일 때는, wind_cv_mv 값은 0.959로 크나, wind_cv_fix 값은 크지 않다. 이는, 발생 전 중규모 규모의 순환은 잘 조직화되었으나, 시스템 규모의 순환은 조직화되지 않음을 의미한다. 2006년 비발달 시스템인 마지막 영상인 경우 바람이 대부분 같은 비슷한 방향으로 서풍 혹은 남서풍으로 불고 있다. 따라서, 두 방향 분산 인자가 모두 낮은 값을 보이는 것을 확인할 수 있다.

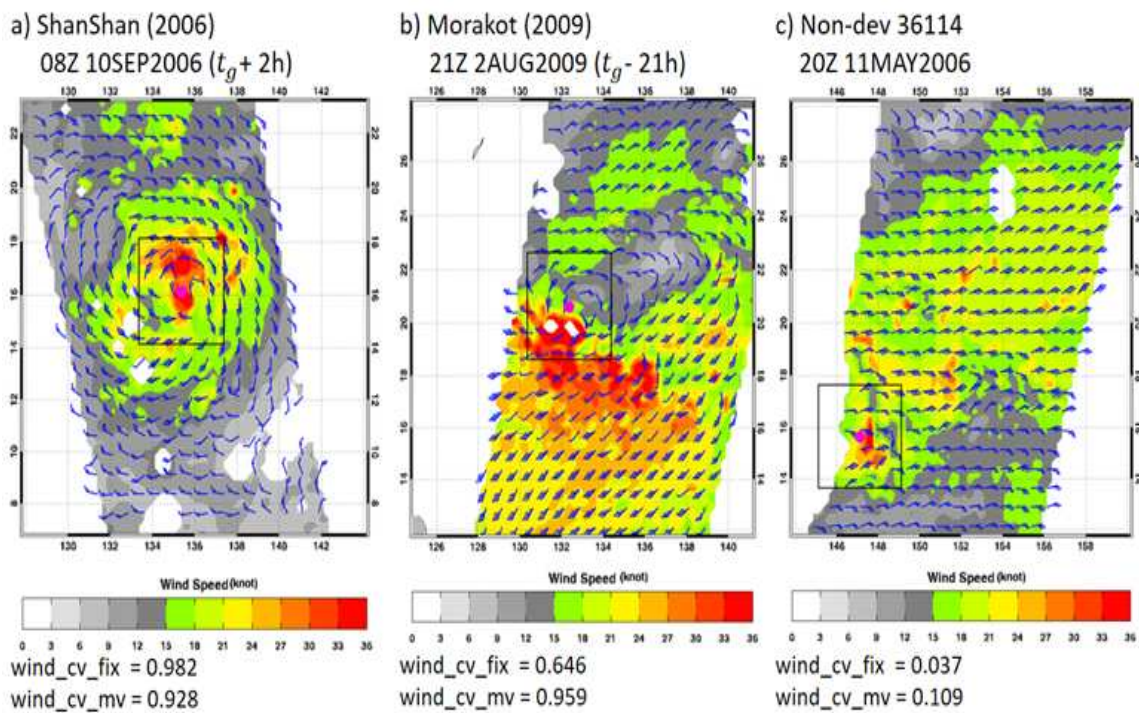


그림 15. 이미 정의된 발달, 비발달 시스템의 중심에 이차원적인 WindSat 바람장 영상에서의 두 개의 원형분산 변수값들: (a) 2006년 태풍 Shanshan의 열대저기압 단계, (b) 2009년 태풍 Morakot이 발생하기 전 섭동 단계, (c) 2006년의 비발달 시스템. JTWC 베스트트랙에서 태풍발생시점(열대저기압 정의 시점)은 t_g 로 정의됨(Park et al. 2016)

태풍으로 발달한 시스템의 경우 강한 양의 와도 영역이 병합(Aggregation)되어 그 면적을 넓혀 간다. 이러한 특성을 위성자료를 이용하여 강풍영역이 병합된 정도를 정량화(인자화)하기 위해서 본 연구에서는 WindSat 풍속 15 m/s 영역을 강한 바람영역으로 정의하며, 그 이상일 경우 1, 아닌 경우를 0으로 지정하여 Binary data의 영상을 구성하였다. 그리고, 처리된 영상을 이용하여 FRAGSTATS 메트릭을 이용하여 PLADJ

인자와 CI 인자를 계산하였다.

PLADJ 인자의 경우 0에서 100의 값을 가지며, CI 인자의 경우 -1에서 1의 값을 가진다. 두 인자의 값이 커질 경우 강풍 영역이 잘 병합된 상태를 뜻하며, 값이 작아질 경우 강풍 영역이 흩어져 있음을 나타낸다.

그림 16는 여러 단계에서 WindSat 영상과 병합인자들의 값을 나타낸다. 2019년 태풍 Nuri의 경우 강풍 영역이 거의 병합되어 넓은 면적을 차지하고 있는데, 이럴 경우 wind_ci의 값은 1에 가깝고 wind_pladj 값은 100에 가깝다. 다음 영상인 태풍 Yagi의 발달 전 단계, 비발달 섭동의 경우 병합의 정도가 약해서 관련 인자들의 값이 작아지는 것을 잘 확인할 수 있다.

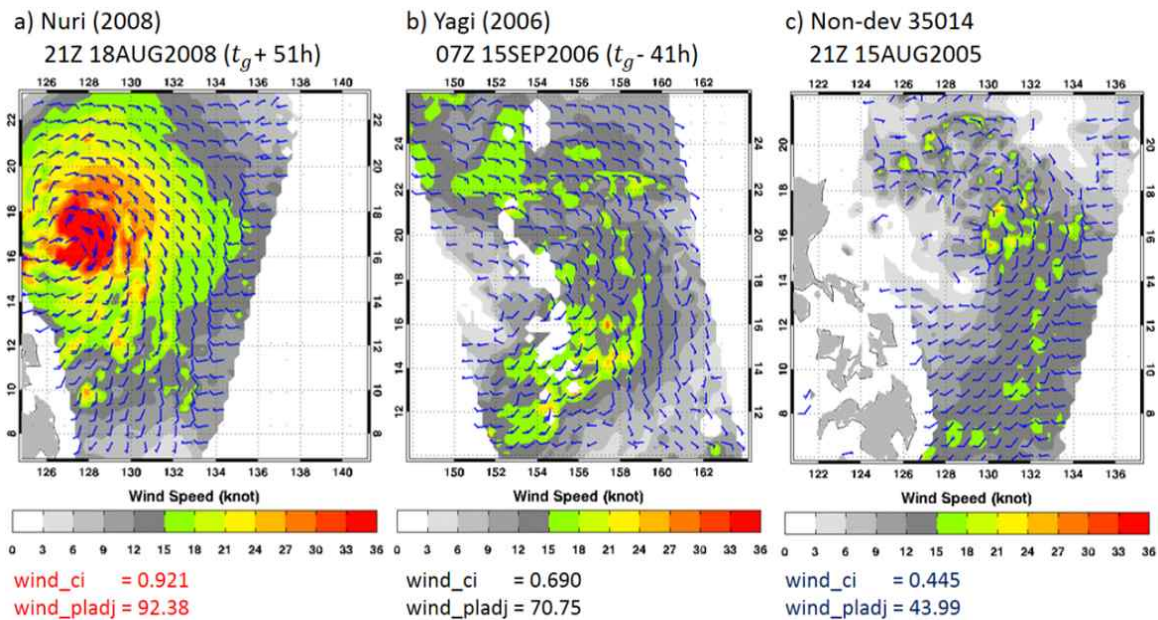


그림 16. 여러 정의된 발달, 비발달 시스템의 중심에서 이차원적 WindSat 바람장 영상과 두 개의 병합정도를 나타내는 인자들(wind_ci, wind_pladj)

그림 17에서는 본 연구를 통해서 개발된 8개의 WindSat 기반 인자들에 대하여 발달과 비발달 시스템의 차이를 박스 플롯(Box-whisker plot)을 통하여 조사하였다. 모든 인자들에 대해서, 중간값은 발달 시스템이 비발달 시스템에 비해 일관적으로 높은 것을 보여주고 있다. 이것은 태풍으로 발생하는 시스템의 경우, WindSat 영상에서 강한 바람 영역과 강한 대류 영역의 정량화를 통해서, 태풍으로 발생하는 경우 더 조직화되며 대칭된 저기압성 순환이 보이며 더 강한 대류가 발달한다는 일반적인 태풍 발생에 대한 역학, 물리 과정과

일치한다. 특히, wind_cv_fix와 wind_ave의 경우 5, 25, 50, 75 percentile의 경우도 일관적으로 발달시스템이 비발달시스템 보다 높은 값을 보인다. 다른 변수들의 경우에는 비발달, 발달시스템 간에 겹치는 영역이 크다. 하지만, wind_cv_fix와 wind_ave 인자의 경우 발달과 비발달시스템이 부분적으로만 겹치며, 두 시스템 분류간의 구별성이 더 크다. 어떤 비발달 시스템의 경우, 어느 정도 강하며, 조직화되어 있거나 강한 대류의 특성을 보이는 것으로 보인다. 이것은 WindSat 단일 방법 태풍을 탐지하는 것은 만족스러운 성능을 보이는데 한계가 있을 수 있음을 보이며, 오탐지율도 다소 있을 수 있음을 의미한다.

본 연구에서 WindSat 기반으로 태풍으로 발생하는 시스템의 역학적 특성을 고려하는 인자를 개발하였다. 개발된 인자들은 기계학습인 입력자료로 이용되어, 객관적 태풍발생 탐지/예측 시스템을 개발하는데 이용될 것이다. 이는 3.3.가 절에서 논의할 것이다.

나. 정지궤도 천리안 기상위성(COMS MI) 인자 개발

(a) 자료 및 방법론

극궤도 위성의 경우 하나의 시스템을 지속적으로 관측할 수 없다는 단점이 있다. 이점을 보완하기 위해서, 본 연구에서는 정지궤도 위성자료도 태풍발생 인자들을 개발하였다. 본 섹션에서는 COMS 위성을 이용한 인자개발에 대해서 논의한다. COMS 위성은 기상, 해양관측 및 통신서비스를 수행하는 우리나라 최초의 정지궤도 복합위성으로, 2010년 6월에 적도 상공으로 발사가 되었다. COMS 탑재체 중 기상센서인 Meteorological Imager(MI)는 가시광(Visible), 단파적외(Shortwave Infrared, SWIR), 수증기(Water Vapor, WV), 적외채널 2개 등 총 5개의 채널에서 관측을 15분마다 수행하며, 우리나라 부근의 기상관측을 하고 있다(표 5). COMS/MI의 적외채널/수증기 채널의 공간해상도는 4 km이다.

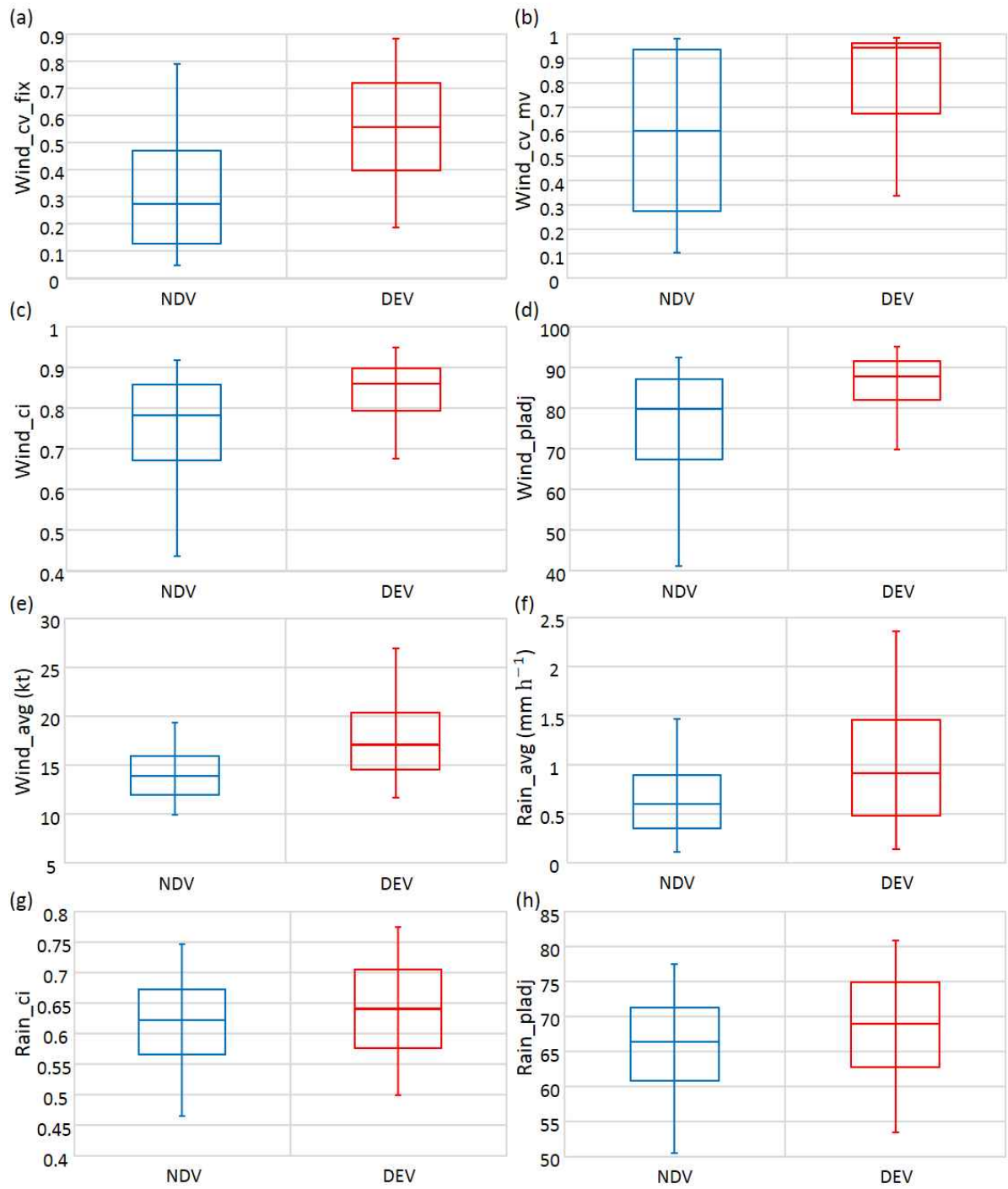


그림 17. 태풍으로 발달하는 시스템(적색 박스)과 비발달 시스템(청색박스)에 대한 개발된 WindSat 위성기반 인자들의 통계적 비교

표 5. COMS 기상센서의 채널, 파장, 공간해상도 및 활용 분야

채널	파장(μm)	공간해상도	활용분야
가시광선 (Vis)	0.67	1km	주간 구름영상, 황사·산불·연무관측, 대기운동벡터
단파적외 (SWIR)	3.7	4km	야간안개 및 하층운, 산불감지, 지면온도 추출
수증기 (WV)	6.7	4km	중상층 대기 수증기량, 상층대기 운동파악
적외1 (IR1)	10.8	4km	구름정보, 해수면온도, 황사관측
적외2 (IR2)	12	4km	구름정보, 해수면온도, 황사관측

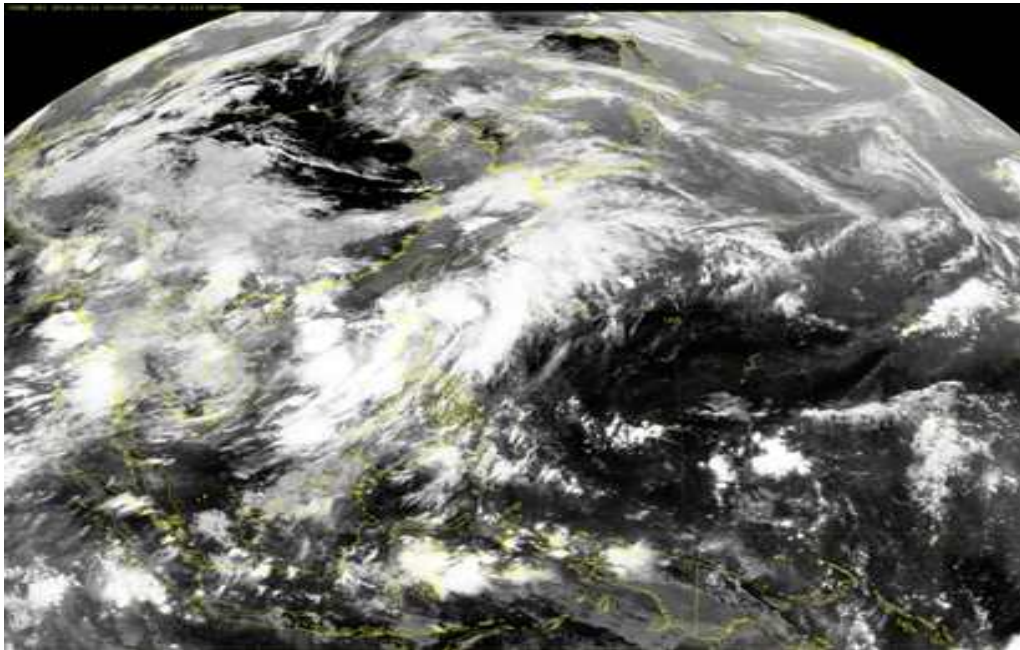


그림 18. 정지궤도 위성(COMS) MI 영상 예시

본 연구개발에는 지속적인 섭동 관측을 위해, 본 연구에서는 가시광선을 제외한 채널들을 이용하여 알고리즘을 구축을 시도하였다. 기계학습 모델들에 입력 변수는 크게 보조변수, 통계변수, 공간변수로 구성하여 입력하였다. 보조변수는 시스템 중심의 위치정보나 현지시간, DOY(day-of-year)로, 통계변수는 섭동시스템 내의 구름들의 각 채널에 대한 밝기온도 평균 값 정보를 뽑아내어 사용하는 인자들로, 공간변수는 시스템 내 구름의 구성 등을 수치화 시켜 사용하는 인자들로 구성되어 있다.

각 트랙은 최적의 품질을 위해, 모든 트랙의 시작 지점에서 24시간까지를 불확실성이 큰 구간으로 평가하여, 제외를 한 후, 나머지 부분에 총 세 개의 라벨을 부여하여 기계학습 모델에 적용을 하였다. 발달하는 트랙의 경우, 강도가 35 knot에 도달하기 직전부터 48시간 이 전까지를 Developing(1) 라벨로, 그 외의 구간은 Continue(2)로 나누었고, 소멸하는 트랙의 경우, 소멸 직전부터 48시간 이 전까지를 Non-developing(0)으로, 그 외의 구간은 Continue(2)로 라벨링을 한 후, 모델을 구축 하였다. 모델의 학습 및 평가를 위해, 샘플 중 2011년에서 2014년까지는 학습 자료로, 2015년 자료는 검증 자료로 사용을 하였다.

(b) 개발된 COMS/MI 위성 인자

본 연구에서는 표 6과 같이 다양한 COMS 기반 위성인자들을 개발 및 고려하였다. 이는 보조변수, 통계변수, 공간변수들을 포함한다. 보조변수는 섭동 중심의 위도, 경도와 함께, 관측 시점의 지역시간과 DOY를 이용한다. 통계변수는 섭동 중심 주변 $3.5^\circ \times 3.5^\circ$ 영역 내에, SWIR과 WV 채널의 평균 밝기온도를 각각 분석하였다. IR1 채널의 경우는, 밝기온도가 270 K보다 낮은 지역(단일채널 구름탐지 기법)을 구름으로 분류 후, 해당 영역의 평균값을 도출해 변수로 사용하였다.

열대섭동은 태풍으로 발생하는 과정에서 구름의 수평 구조는 조직적으로 바뀌며 태풍으로 발달해간다. 이러한 특성을 정량화하기 위해 DAV, IR1_MCI, SHDI_WV 등 다양한 새로운 공간변수를 개발하였다.

정지궤도 위성에서 관측된 구름의 형태를 정량화하기 위해 Pineros et al. [2011]에서 개발된 DAV 인자를 이용하였다. DAV 기술은 정지궤도 위성에서 산출된 구름의 밝기온도의 변화도를 이용하여 중심으로부터 변화도의 제곱 합을 계산해 구름의 대칭도를 계산하는 방법인데, 이 값을 공간변수 인자 중 하나로 사용하였다. 또한, 본 연구에서는 독자적으로 발달할 섭동의 경우, 분산된 작은 중규모 대류계들이 아닌 규모가 큰 대류가 발달하는 것을 확인하여, 이를 정량화 하고자 MCI(monopolized convection index)를 고안하여, 입력자료로 사용하였고, 아래와 같다.

$$MCI = \frac{\text{Top two large EC cluster pixel \#}}{\text{Total EC pixel \#}} \times 100\%$$

IR1 영상 내에서 밝기온도가 240K 보다 낮은 지역을 EC(Extreme Convection)영역으로 구분 후, 영상을 이진 분류(Binary Classification)를 진행한다. 그 후, 가장 영역이 큰 두 개의 패치의 화소 개수가 전체 EC 영역의 화소에서 차지하는 비율을 구하여, 해당 영상의 MCI 값으로 계산한다.

표 6. 본 연구에서 개발된 사용된 COMS 기반 위성인자들

변수명		설명
보조변수	lat	latitude
	lon	longitude
	doy	day-of-year
	local_time	local time
통계변수	swir_mean	average brightness temperature of SWIR
	wv_mean	average brightness temperature of WV
	ir1_mean	average brightness temperature of IR1
공간변수	DAV	deviation angle variance
	ir1_MCI	monopolized convection index of IR1
	ir1wv_MCI	monopolized convection index of IR1-WV
	SHDI_wv	shannon's diversity index of WV
	SHEI_wv	shannon's evenness index of WV
	SHDI_ir1	shannon's diversity index of IR1
	SHEI_ir1	shannon's evenness index of IR1
	SHDI_ir1wv	shannon's diversity index of IR1-WV
SHEI_ir1wv	shannon's evenness index of IR1-WV	

그리고, 영상 내의 온도값 범위를 기준으로 여러 클래스로 구분한 뒤, 당 클래스의 구성비가 어느 정도가 되는지를 인자에 추가 하였다. 이는 WV, IR1, 그리고 IR1-WV영상에 적용이 되었는데, 각 영상의 온도 분포에 대해, 밝기온도의 값이 낮은 20%를 강한 대류(Deep convection), 20-40%를 대류(Convection), 40-60%를 약한 대류(Weak convection) 영역으로 구분하여, 세 클래스의 구성 비율을 SHDI (Shannon's Diversity Index), SHEI (Shannon's evenness index) 분포 통계 변수를 통해 정량화 시켰다 (그림 19). 식은 아래와 같다.

$$SHDI = - \sum_{i=1}^m (P_i \times \ln P_i)$$

$$SHEI = \frac{- \sum_{i=1}^m (P_i \times \ln P_i)}{\ln m}$$

i는 범주 (i=1, 2, 3), P_i 는 범위 안에 i가 차지하는 비율을 나타내고, m은 클래스의 개수를

의미한다.

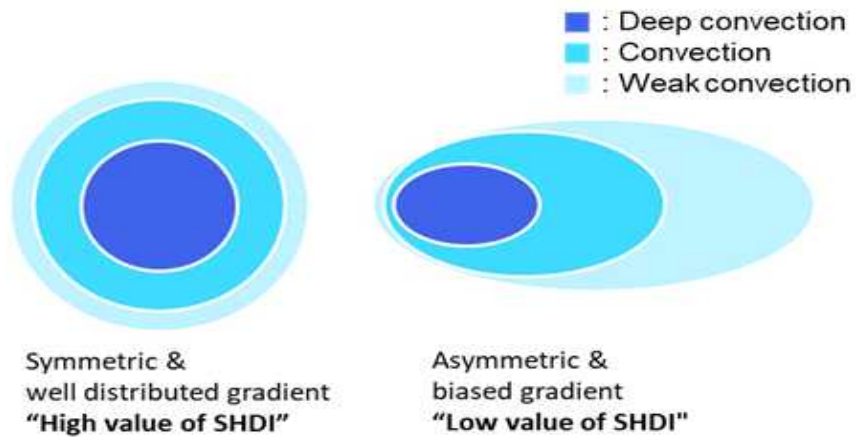


그림 19. SHDI (Shannon's Diversity Index) 도입 배경 개념도

그림20는 여러 발달/비발달 시스템의 COMI MI 영상에서 SHDI와 SHEI 인자들의 예시를 보인다. 발달하고 있는 시스템의 경우 강한 대류의 면적이 넓어지면서 SHDI = 1.07, SHEI = 0.97로 해당 인자들의 값이 커졌다. 비발달 시스템의 위성 영상의 경우 아주 강한 차가운 적외채널 온도의 면적은 아주 작으며, 대류 면적도 넓지 않아 SHDI= 0.19, SHEI = 0.28로 인자 값이 낮았다.

그림 21을 통하여 COMS 기반으로 개발된 태풍발생 인자들의 통계적 분포를 발달/지속/비발달 시스템별로 어떻게 다른지 조사하였다. 분석 결과는 발달 시스템의 경우 DAV의 값이 비발달 시스템과 지속 시스템 보다 훨씬 작은 값을 보여, 태풍으로 발달하기 직전의 시스템이 급격하게 조직화되어 가는 과정을 잘 보여주고 있다. 또한, 발달 시스템의 경우 WV_mean의 값이, 지속과 비발달 시스템보다 훨씬 낮은 값을 보인다. 이는 태풍으로 발달하는 직전 48시간부터 대기 중 수증기량이 증가하는 것이 급작스럽게 증가하는 것을 의미하고 있다. 전체적으로 이 결과에서 주목할 점은 나중에 태풍으로 발달하더라도 48시간 이전의 대류의 조직화 정도와 강도는 비발달 시스템과 다소 유사함을 알 수 있다. 따라서, 이 그림은 본 연구에서 사용되었던, 세 가지 카테고리 Developing, Continuous, Non-developing가 적절하다는 근거 결과가 될 수 있다. COMS/MI 인자들을 이용한 발생 탐지 모델은 3.3.4 절에서 논의한다.

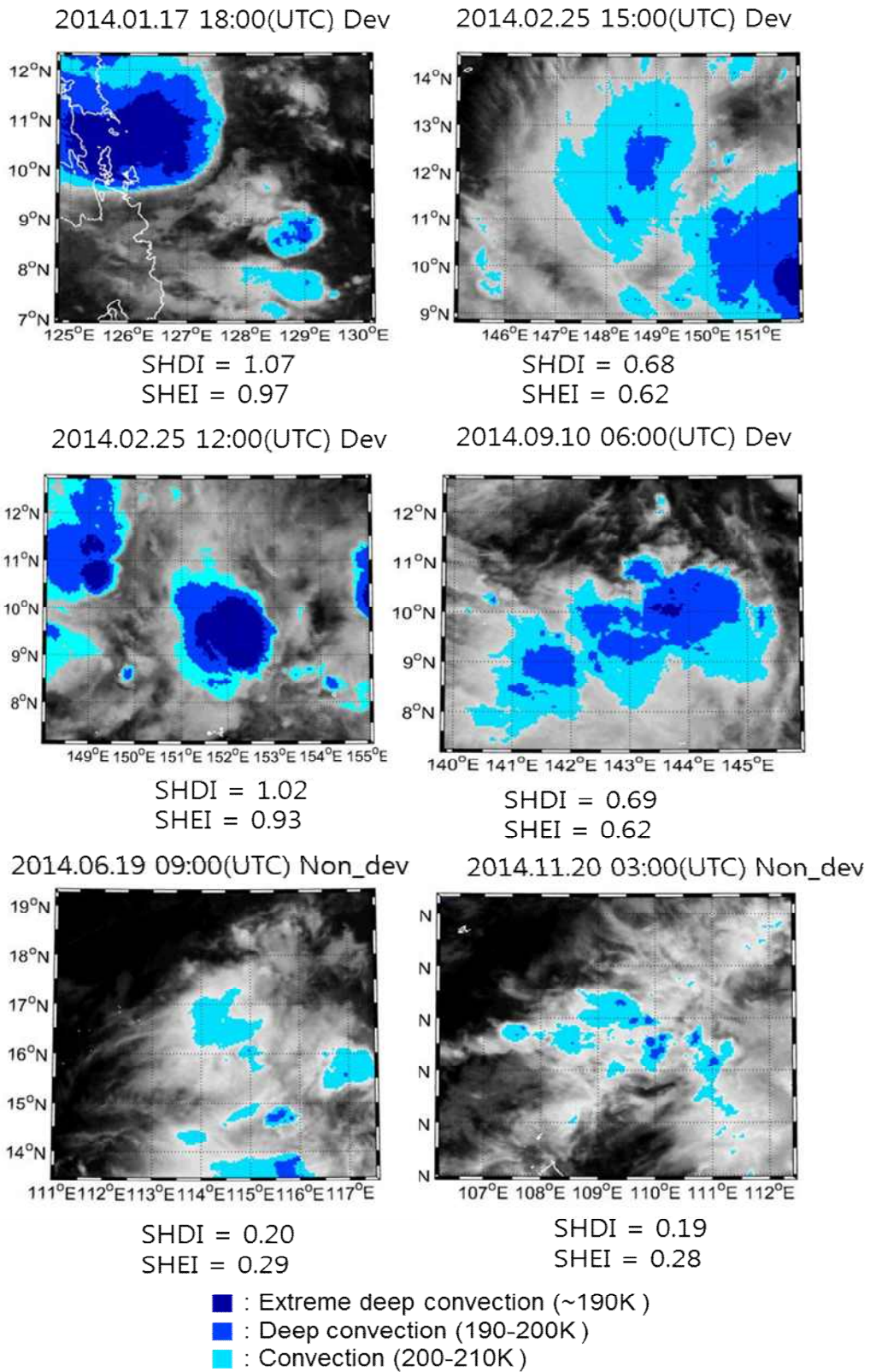


그림 20. 발달/비발달 시스템의 COMS/MI 영상 및 SHDI, SHEI 지수 예시

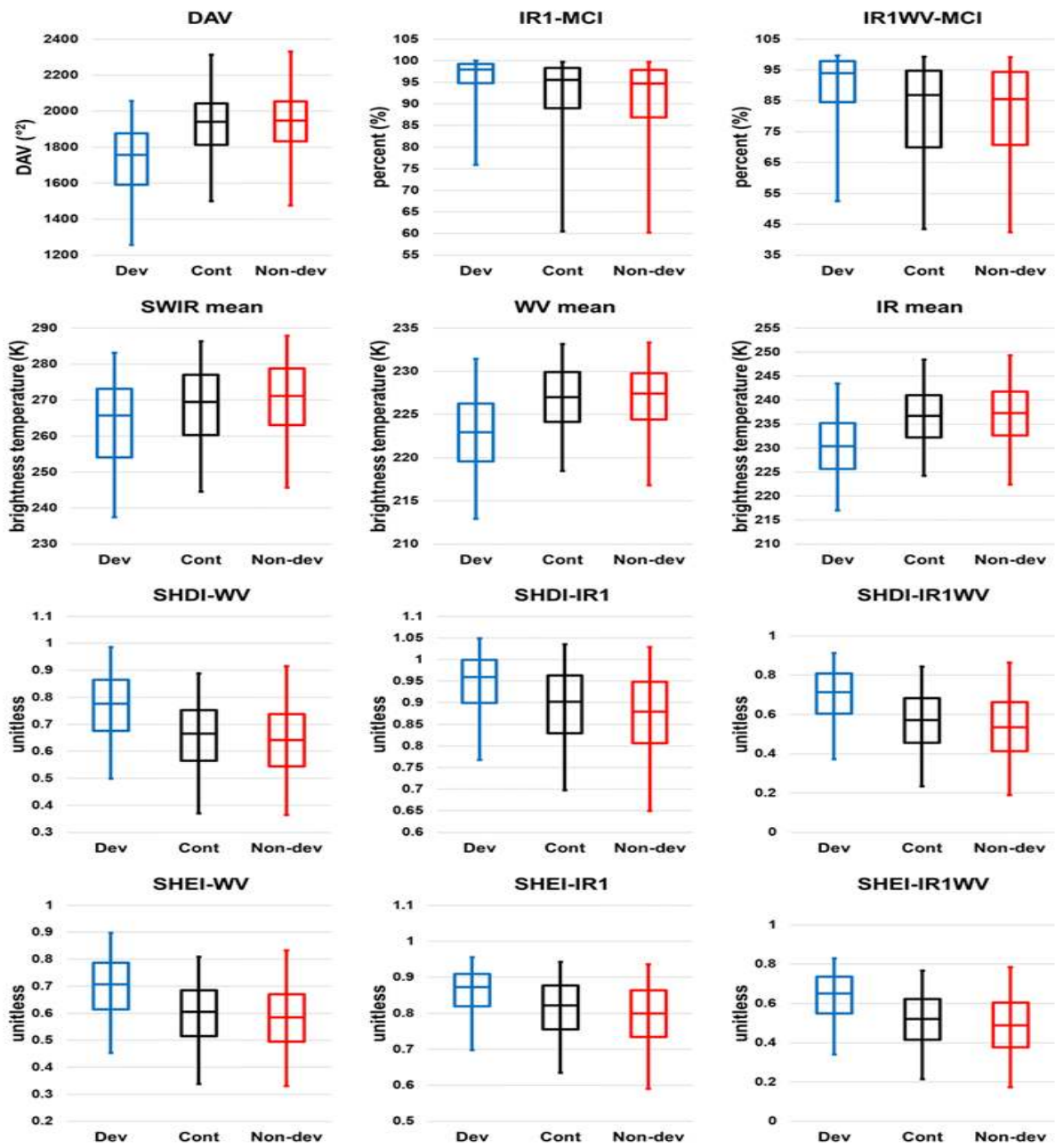


그림 21. 2011년에서 2015년까지 취득된 모든 샘플에 대한 통계 및 공간변수의 클래스별 값 비교 그래프.

다. 정지궤도 히마와리-8 위성(Himawari-8 AHI) 인자 개발

(a) 자료

Himawari-8은 Japan Meteorological Agency(JMA)에서 2014년 10월 7일 새로 발사된

정지 궤도 위성으로 2015년부터 운용되기 시작했다. Himawari-8에는 새로 개발된 imager인 Advanced Himawari Imager(AHI)가 탑재되어 기존의 정지궤도 위성들과 비교하여 더 좋은 분해능을 가지고 있다. COMS와 비교하여 분광 해상도(spectral resolution)의 경우에는 기존 5개 채널(가시광선 채널 1개, 적외선 채널 4개)에 비해 16개 채널(가시광선 채널 3개, 근적외선 채널 3개, 적외선 채널 10개)을 가지고 있다. 공간 분해능의 경우에는 기존 COMS 위성에서 가시광선의 경우 1 km 의 해상도, 적외선 채널에서 4 km 의 해상도를 가지고 있던 것에 비해 Himawari-8 위성의 경우 가시광선 채널의 경우 0.5 km, 근적외선 채널의 경우 1 km, 적외선 채널의 경우 2 km 의 해상도를 가지고 있다. 시간 분해능의 경우에는 COMS 위성은 60 분의 해상도를 가지고 있지만, Himawari-8 위성은 10분 간격의 해상도를 가지고 있어 더 많은 빈도로 관측을 할 수 있게 됐다.

(b) 방법론

본 연구에서는 낮 시간에만 이용 가능한 가시광선, 근적외선 채널을 제외한 적외선 채널 중 8, 10, 11, 14, 15, 16번 채널 자료를 이용하였고, 60분 간격으로 관측 자료를 이용하였다. 정지궤도 위성에서 산출한 대류 관련 정보들을 태풍 발생 예측에 유용한 정보들로 가공하기 위해서 예측 인자를 개발하였다. 기존 연구들에서 주로 이용되던 방법인 Dvorak 기법의 경우 구름의 형태를 분석하여 태풍의 강도를 추정하는 방식이었던 것에 비해 본 연구에서는 대류의 행동의 변화를 태풍 발생의 잠재적인 에너지로 인지하여 특성화 시키는 작업을 통해 예측 변수를 개발하였다.

태풍의 발생을 예측하기 위해 태풍의 특성을 모의하는 인자들을 대류의 강도(intensity), 대칭성(symmetry), 분포(Diversity) 3가지 범례로 분류하여 개발하였다. 대류의 강도의 경우 실제 대류 에너지를 추정하는 것을 목적으로 개발되었다. 대류의 대칭성의 경우에는 기존 Dvorak 기법에서 구름의 모양을 분석하였던 것과 유사하게 구름의 대칭성을 정량화 하였다. 마지막으로 대류의 분포는 강한 대류의 분포를 분석하여 대류의 크기와 집적도를 정량화 하였다.

(c) 개발된 위성 인자

정지궤도 위성 자료는 복사 에너지 값을 제공하기 때문에 Radiance 자료를 밝기온도로 변환하는 작업을 하였다. 근사식(sensor Plank function)의 계산 과정은 아래와 같은 공식을 따르며 Takahashi 와 Center(2017)를 참조하였다.

$$T_b = b_{1i} + B_{2i}T_e + b_{3i}T_e^2$$

$$T_e(B_i) = \frac{\frac{hc}{k}v_i}{\ln\left(\frac{2hc^2v_i^3}{B_i} + 1\right)}$$

먼저 복사(B_i) 값을 이용하여 effective temperature(T_e)를 계산해준다. h 는 플랑크 상수 (Planck coefficient)로 $6.62606957e-34$ [J·s] 로 정의된다. k 는 볼츠만 상수(Boltzmann constant)로 $1.38064880e-23$ [J/K] 로 정의된다. c 는 빛의 속도(speed of light)로 $2.99792458e+08$ [m/s] 로 정의된다. 변환된 effective temperature 값과 상수 값을 이용하여 밝기 온도 값을 계산한다. v_i 는 각 채널의 중심 파수, b_i 는 각 채널의 상수 값으로 표 7과 같이 정의된다.

표 7. Himawari-8 AHI 의 각 채널 별 파수와 band correction coefficients

AHI band	Wavenumber	Band correction coefficients		
	ν (cm-1)	b1	b2	b3
Band8 (6.2 μm)	1609.241	-1.662616	1.003694	-1.732716e-07
Band10 (7.3 μm)	1361.387	-0.06306013	1.000195	-1.069833e-07
Band11 (8.6 μm)	1164.443	-0.1605105	1.000589	-4.019762e-07
Band14 (11.2 μm)	890.741	-0.2530423	1.001233	-1.153788e-06
Band15 (12.4 μm)	809.242	-0.3766459	1.002025	-2.096994e-06
Band16 (13.3 μm)	753.369	-0.09773197	1.000564	-6.266746e-07

변수 개발에 앞서, 구름이 없는 지역을 제거하기 위해 14번 적외선 채널의 밝기온도가 288.15 K 보다 높은 온도의 영역을 제거한 뒤 이용하였다.

강도와 관련된 변수는 단일 채널을 이용한 인자와 다중 채널을 이용한 인자로 구성되어있다. 단일 채널을 이용한 인자는 각 6개의 채널의 밝기 온도를 구름이 없는 지역을 제거한 영역 평균한 값을 계산하였다. 영역의 크기의 경우 태풍 중심을 기준으로 $4^\circ \times 4^\circ$ 정사각형 범위이다. 또한 강한 대류를 모의하기 위해 $1^\circ \times 1^\circ$ 의 영역에서 최솟값을 구하여 이동 윈도우(moving window) 방법으로 계산하여 가장 작은 값을 변수로 개발했다. 다중 채널을 이용한 강도 변수는 서로 다른 두 채널의 차이 값을 영역 평균한 값을 이용하였다. 대표적인 다중 채널을 이용한 분석으로는 Olander and Veldon (2009)의 적외 채널과 수증기 채널의 차이(IR-WV) 값을 이용하여 태풍의 강도를 추정한 연구가 있다. 또한 Lee et al.

(2017)에서 convective initiation을 탐지하기 위해 다중 채널의 차이를 이용하여 대류의 특성들을 변수화한 연구가 있다. 다중 채널의 변수는 다양한 대류의 특성들을 해석 하는데 유용하기 때문에 본 연구에서 다양한 채널의 차이를 정량화 하여 태풍의 중심으로부터 $4^{\circ} \times 4^{\circ}$ 정사각형 범위에서 평균한 값을 이용하여 변수를 개발하였다.

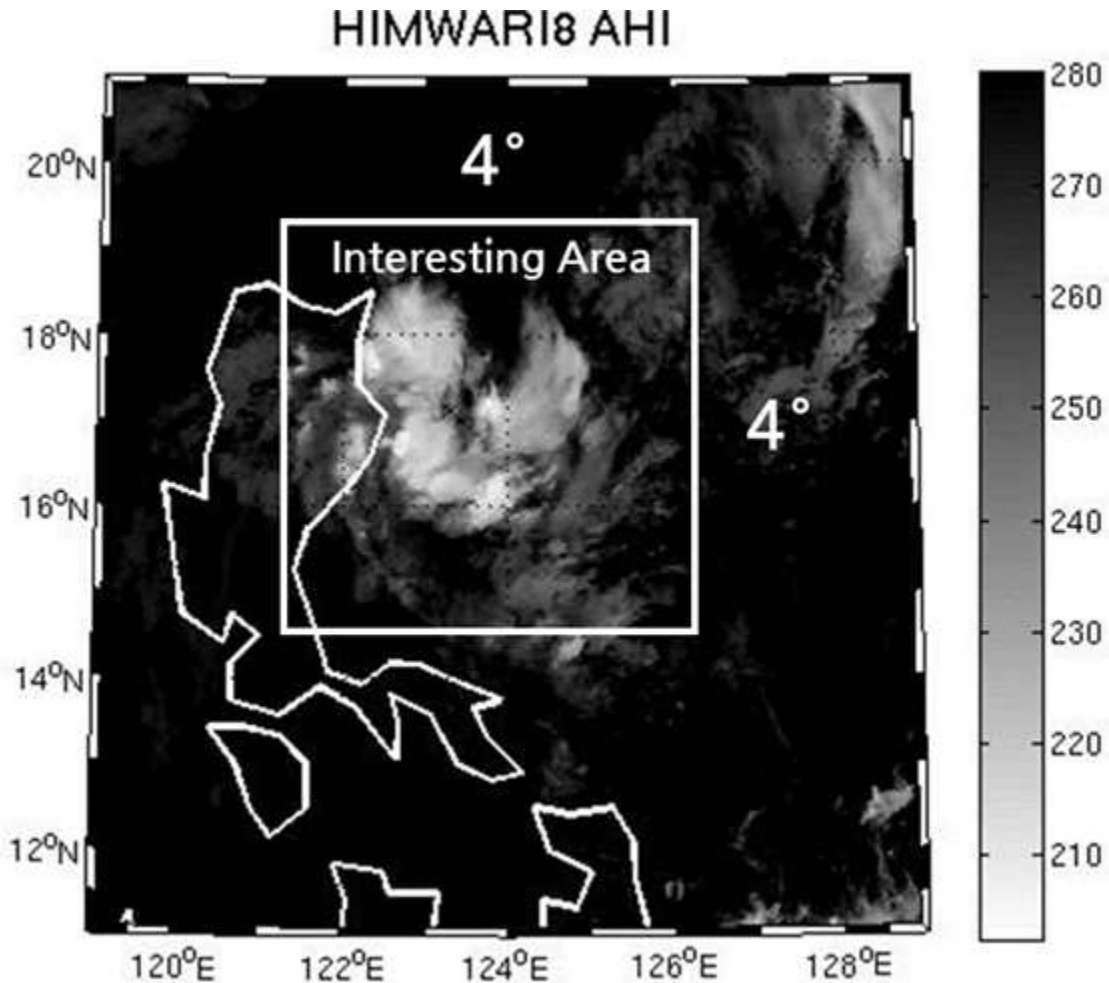


그림 22. Himawari-8 AHI IR 14번 채널이미지 예시. 강도와 관련된 변수들은 태풍 중심으로부터 $4^{\circ} \times 4^{\circ}$ 정사각형 범위 (interesting area)에서 산출되었다.

태풍이 발생하는 과정에서 구름의 겉보기 형태는 조직적으로 바뀌며 강한 태풍으로 발달해간다. 정지궤도 위성에서 관측된 구름의 형태를 정량화하기 위해 Pineros et al. (2011)에서 개발된 DAV 기술을 이용하였다. DAV 기술은 정지궤도 위성에서 산출된 구름의 밝기온도의 변화도를 이용하여 중심으로부터 변화도의 제곱 합을 계산해 구름의 대칭도를 계산하는 방법이다. 본 연구에서는 10번 수증기 채널과 14번 적외 채널의 값을 이용하여 DAV 산출 값을 변수화 하였다. DAV 계산 과정에서 고해상도의 적외 채널값이 오차를 만들어 실제 계산 과정에서는 해상도를 10 km 로 낮춘 뒤 로우 패스 필터(low pass

filter)를 적용하였다. 각 그리드마다 이동 윈도우 방식으로 DAV를 계산하여 만들어진 새로운 DAV 지도에서 가장 낮은 값을 가지는 위치를 대류의 중심으로 가정하여 변수로 이용하였다.

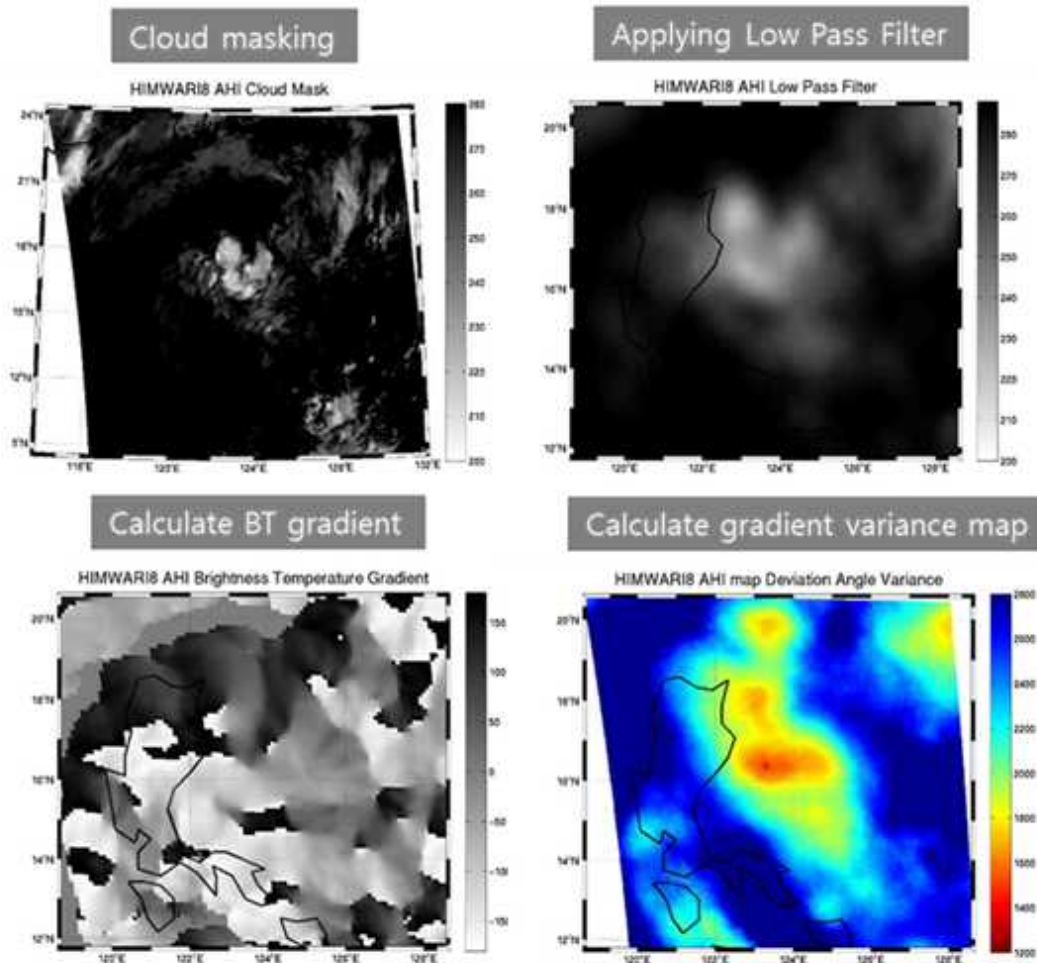


그림 23. Himawari-8 AHI IR 14번 채널에서 DAV 기술 적용 예시. 왼쪽 상단은 구름을 제거한 그림이고, 다음 과정으로 10 km 해상도로 낮춘 뒤 로우 패스 필터를 적용한 그림이다 (오른쪽 상단). DAV 계산을 위해 각 그리드 별로 밝기온도의 기울기를 계산하였고 (왼쪽 하단), 기울기 값의 제곱 합을 하여 DAV를 계산한 맵이 오른쪽 하단에 나타내었다.

구름의 형태만큼 구름의 크기와 모인 정도(집적도)가 대류 운동의 크기를 예측하는데 중요한 인자로 이용될 것이라는 가정을 하여 대류의 분포 변수를 개발하였다. 대류의 분포를 정량화하기에 앞서 14번 적외 채널을 이용하여 중심으로부터 $4^{\circ} \times 4^{\circ}$ 정사각형 범위 대류의 크기를 세 가지 범위로 나누었다. 240 K 보다 작은 값을 갖는 지역을 매우 강한 대류 영역, 240 K 이상 265 K 미만인 지역을 강한 대류의 영역, 265 K 이상 280 K 미만인 지역을

대류 영역으로 나누었다. 각 임계값은 태풍의 주변 영역의 밝기 온도 값을 누적 분포 함수로 만들어 임의로 나누어 준 값이다.

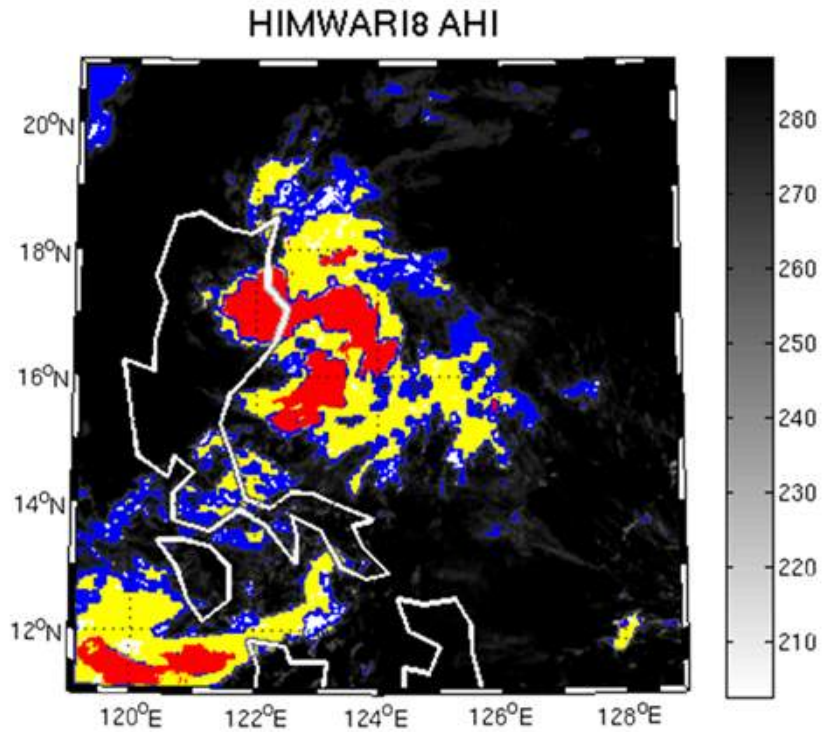


그림 24. Himawari-8 AHI IR 14번 채널의 밝기온도 값에 따른 분류 이미지. 파란색 영역은 대류 지역으로 265 K 이상 280 K이하의 온도 범위를, 노란색 영역은 강한 대류 지역으로 240 K 이상 265 K 미만의 온도 범위를, 빨간색 영역은 매우 강한 대류 지역으로 240 K 미만의 온도 범위를 나타낸다.

세 개의 범주에 따라 공간 분포를 계산하기 위해 Shannon's Diversity Index(SHDI), Simpson's Diversity Index(SIDI), Simpson Index(SIMI) 분포 통계 변수를 이용하였다. 식은 아래와 같다.

$$SHDI = - \sum_{i=1}^m (p_i * \ln P_i)$$

$$SIDI = 1 - \sum_{i=1}^m p_i^2$$

$$SIMI = \frac{1}{\sum_{i=1}^m P_i^2}$$

i 는 범주($i=1, 2, 3$), P_i 는 범위 안에 i 가 차지하는 비율을 나타낸다. SHDI의 범위는 0부터 무한대를 가진다. SHDI는 균등한 분포를 가질수록 값이 크고 한 범례의 비율이 높을수록 낮은 값을 가지는 경향을 보인다. SIDI는 0에서 1 사이의 값을 가진다. SIMI는 1에서 3 사이의 값을 갖는다($m=3$). SIDI와 SIMI의 값이 커질수록 다양성이 높아지고, 값이 작아질수록 한 종류의 비율이 높은 것을 의미한다.

대류의 가장 큰 특성 중 하나인 일 변동(diurnal variation)은 태풍 발달을 탐지하는데 잘못된 예측하게 영향을 미친다. 일 변동의 영향을 줄이기 위해 24시간 이동 평균법(running average)을 생산된 모든 변수에 적용하였다. 본 연구에서 개발된 위성 인자와 간단한 설명을 아래 표에 정리하였다. 이동 평균법을 각 변수들에 적용한 변수는 “TR_변수”로 명명하였다. Himawari-8/AHI를 이용한 발생 탐지 알고리즘은 3.3.4 절에서 논의한다.

표 8. Himawari-8 AHI 산출 위성 인자 리스트 및 설명

No.	위성 인자	범례	설명
#1	M08	Intensity	Upper-level water vapor
#2	M10	Intensity	Lower-level water vapor
#3	M11	Intensity	Cloud phase
#4	M14	Intensity	Cloud height
#5	M15	Intensity	Cloud height
#6	M16	Intensity	Cloud height and amount
#7	D(08-14)	Intensity	WV-IR
#8	D(08-16)	Intensity	WV-IR
#9	D(10-14)	Intensity	WV-IR
#10	D(10-16)	Intensity	Cloud top height
#11	D(16-14)	Intensity	Cloud top glaciation
#12	D(15-14)	Intensity	Cloud top glaciation
#13	D(11-14)	Intensity	Ice cloud detection
#14	D(08-10)	Intensity	Water vapor difference between top and bottom
#15	Min08	Intensity	Upper-level water vapor in strong convective region
#16	Min10	Intensity	Lower-level water vapor in strong convective region
#17	Min14	Intensity	Cloud height in strong convective region
#18	Min16	Intensity	Cloud height in strong convective region
#19	DAV10	Symmetry	Degree of symmetry of water vapor cluster
#20	DAV14	Symmetry	Degree of symmetry of cloud cluster
#21	SHDI14	Diversity	Proportional abundance of each patch type multiplied by that proportion
#22	SIDI14	Diversity	Measure of diversity for patch type
#23	SIMI14	Diversity	Dominance of patch

라. 모델장을 이용한 대규모환경 인자 개발

열대 섭동이 태풍으로 발달하거나 소멸하는 과정에서 주변 대규모 환경 인자가 어떠한 형태를 띠고 있는지 살펴보기 위해 열대 섭동 주변의 환경 요소 중 850hPa의 상대와도, 200hPa의 발산장, 850hPa 수분 발산장, 700hPa 상대 습도, 10m 상공 풍속, 850hPa-200hPa 연직바람시어, 그리고 500hPa의 연직 기압 풍속장의 시간에 따른 변화를 태풍으로 발달한 열대 섭동과 그렇지 못한 열대 섭동의 값을 비교 분석하였다.

850hPa의 상대와도의 경우, 열대 섭동 발생 후 태풍으로 발달한 경우와 소멸한 경우 모두 태풍으로 발달한 열대 섭동의 상대와도 값이 크게 나타났으며 시간에 상대와도가 크게 증가하는 추세도 더 크게 나타났다(그림 25). 그리고 태풍 발달 전 48시간 시점부터 태풍으로 발달한 열대 섭동의 상대와도 값이 크게 증가하였다. 태풍으로 발생하는 열대섭동의 경우 48시간 이전의 하층 와도 3.5×10^5 이상, 평균 하층 와도 증가율 $2 \times 10^5 \text{ s}^{-1}/\text{day}$ 이상의 값을 가지고 있다. 그러므로 하층 와도의 절대적인 크기와 하층 와도의 증가 경향성 모두 태풍 발생 탐지에 주요 환경인자로 사용될 수 있다.

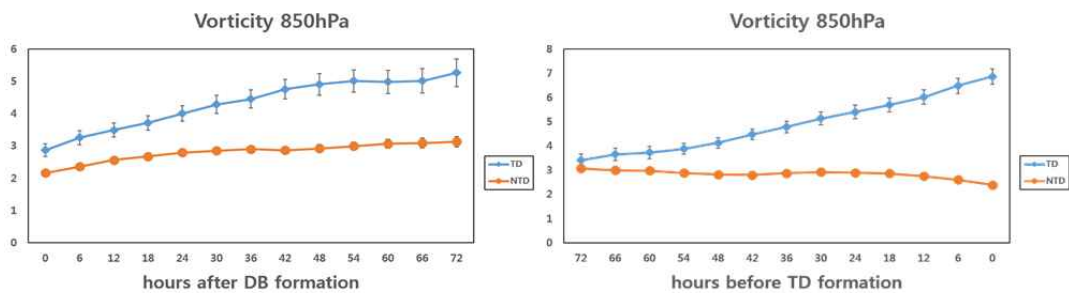


그림 25. 열대 섭동 발생 이후(좌), 열대 저압부로 발달 또는 소멸하기 이전(우) 72시간 동안의 열대 섭동 주변 850hPa 상대와도의 변화, (청색: 태풍으로 발달하는 열대 섭동, 적색: 소멸하는 열대섭동)

상층 발산은 태풍의 구조를 형성하기 위한 태풍 발달의 주요한 환경인자 중에 하나이다. 열대 섭동 주변의 200hPa의 발산장은 태풍으로 발달하는 열대섭동의 경우 태풍으로 발달하지 않는 열대 섭동보다 매우 큰 값을 유지하였다(그림 26). 태풍으로 발달한 경우 시간적 변화는 미미하였으나, 소멸한 열대 섭동의 경우에는 상층 발산이 감소하였다. 상층 발산장의 경우 절대적인 값의 크기가 태풍의 발달을 탐지하는데 주요한 인자로 사용될 수 있으며, 경향성 또한 태풍 발달 탐지의 보조 인자로 사용될 수 있을 것이다.

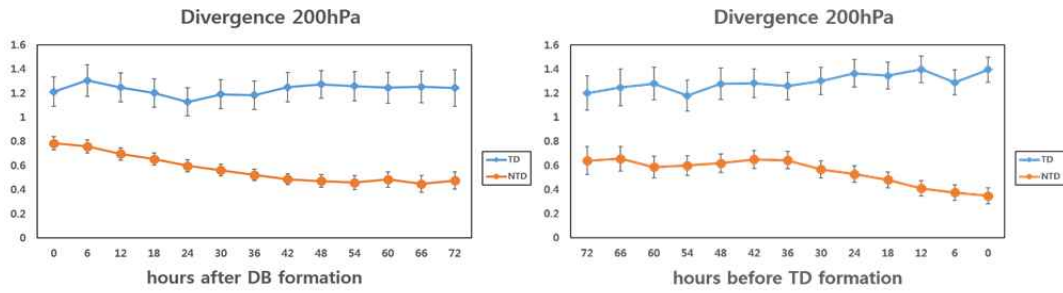


그림 26. 그림 23과 같음, 단 200hPa 발산장

하층 수분 발산장은 열대섭동 주변에서 음수의 값을 가지고 있어 수분 수렴이 이루어지는 것으로 나타나고 있다. 태풍으로 발달한 열대 섭동의 값이 소멸한 열대 섭동보다 수분 수렴이 훨씬 크게 나타나고 있다(그림 27). 열대 섭동 발생 후 72시간의 경우 두가지 경우 모두 수분 수렴이 증가(수분 발산 감소)하는 추세였으나 태풍으로 발달한 열대 섭동의 값이 시간이 지남에 따라 더 크게 수분 수렴이 증가하였다. 그리고 태풍 발달 72시간 전의 경우 태풍으로 발달한 열대 섭동은 수분 수렴이 크게 증가하였으나 소멸한 열대 섭동의 경우 거의 수분 수렴은 거의 일정한 값을 가졌다.

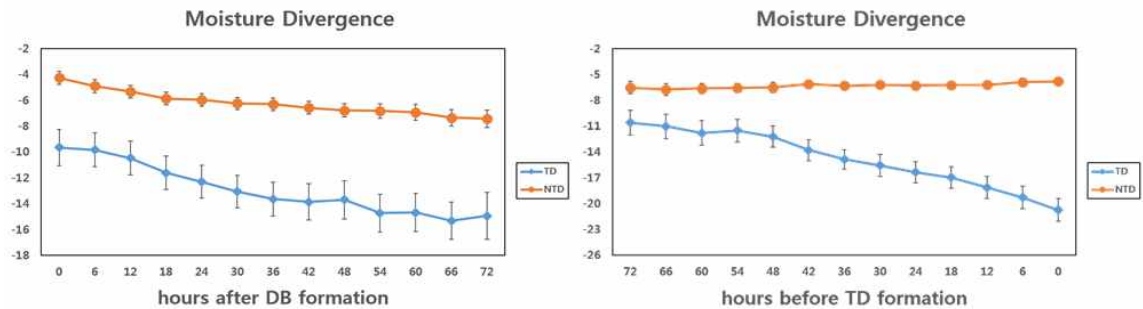


그림 27. 그림 23과 같음, 단 하층 수분 발산장

열대섭동 주변 상대습도는 태풍으로 발달한 열대 섭동의 값이 소멸한 열대 섭동에 비해 큰 값을 나타냈다(그림 28). 시간적 변화는 미미하므로 시간적인 추세보다 절대적인 상대습도의 값이 태풍 발달을 판단하는 주요 인자로 사용될 수 있다.

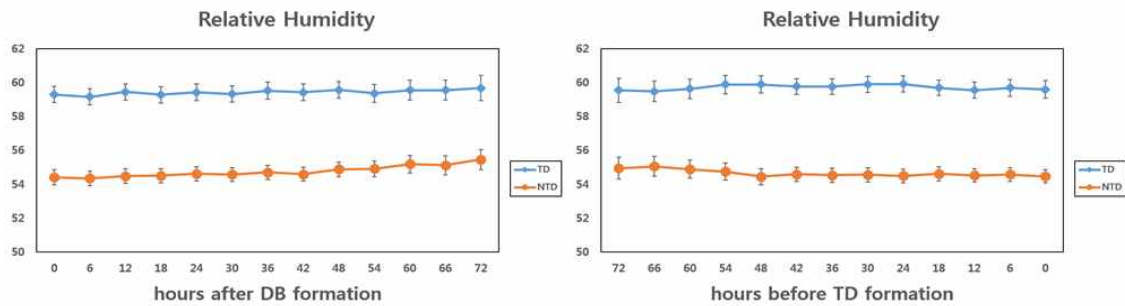


그림 28. 그림 23과 같음, 단 상대습도

지면 풍속의 경우 태풍으로 발달한 열대 섭동은 시간이 지남에 따라 크게 증가하는 반면 소멸한 열대 섭동은 거의 일정한 값을 나타냈다(그림 29). 태풍은 지면 풍속이 강한 열대 저기압으로 정의되기 때문에 태풍으로 발달하는 열대 섭동에서 지면 풍속이 강해지는 것은 당연한 결과이다. 이러한 지면 풍속의 증가 경향은 태풍 발달을 탐지하는데 보조적인 인자로 사용될 수 있을 것이다.

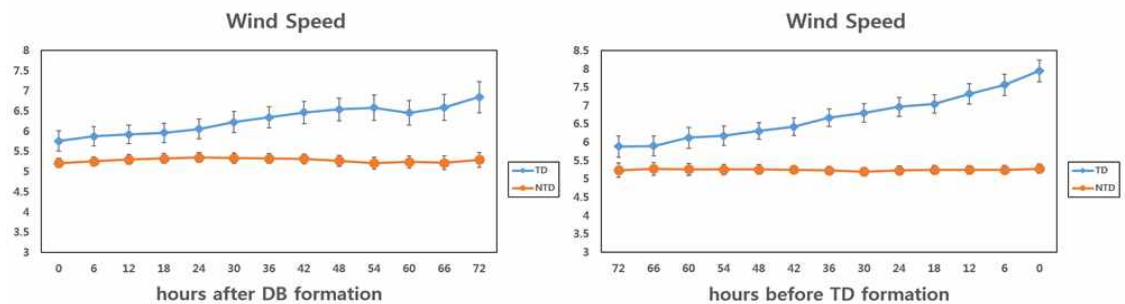


그림 29. 그림 23과 같음, 단 풍속

태풍은 약한 연직바람시어에서 발생하는데, 연직바람시어가 약해야 태풍의 구조가 수직적으로 형성되어 잠열이 태풍에 효율적으로 전달될 수 있는 환경장을 형성하게 된다. 태풍으로 발달한 열대 섭동이 소멸한 열대 섭동에 비해 연직바람시어의 값은 낮게 나타났다(그림 30). 시간적 변화는 소멸하는 열대섭동의 경우 열대 섭동 발생이후 감소하는 경향성을 보이나 소멸하기 전에는 다시 증가하는 것으로 나타나고 있다. 태풍으로 발생하는 열대 섭동도 태풍으로 발달 이전에 연직시어가 증가하는 경향성을 나타내고 있다. 이러한

연직시어의 경향성은 알려진 연직시어와 태풍과의 관련성과는 다르게 나타나고 있어 연직시어를 태풍 발달 환경인자로 사용하는 경우 변화 경향성 보다 절대적인 연직시어의 값의 차이를 이용할 수 있을 것이다.

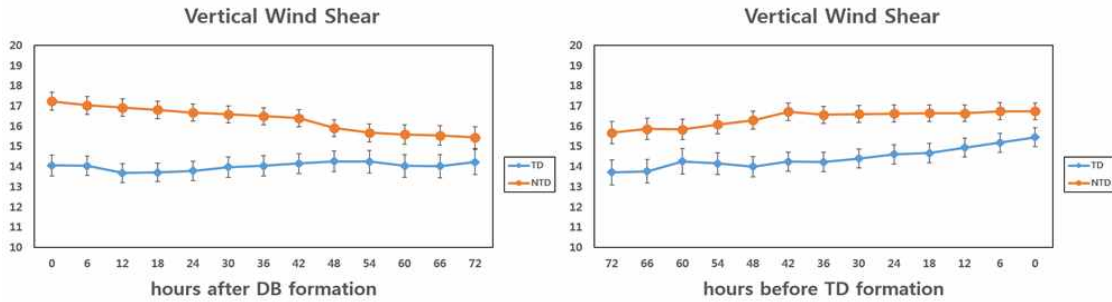


그림 30. 그림 23과 같음, 단 연직바람시어

연직 기압 풍속은 음의 값이 클수록 상승이 강함을 나타낸다. 열대 섭동에서의 상승 기류는 태풍으로 발달한 열대 섭동이 소멸한 열대 섭동이 그 절댓값이 매우 크다 (그림 31).

또한 소멸한 열대 섭동은 상승기류기 지속적으로 약화되고 있음을 알 수 있다. 열대 섭동에서의 상승기류의 가도는 태풍으로 발달하는 열대 섭동과 소멸하는 열대 섭동 간에 확연한 차이를 보이고 있어 태풍 발달을 판단하는 주요한 환경 인자로 사용될 수 있을 것이다.

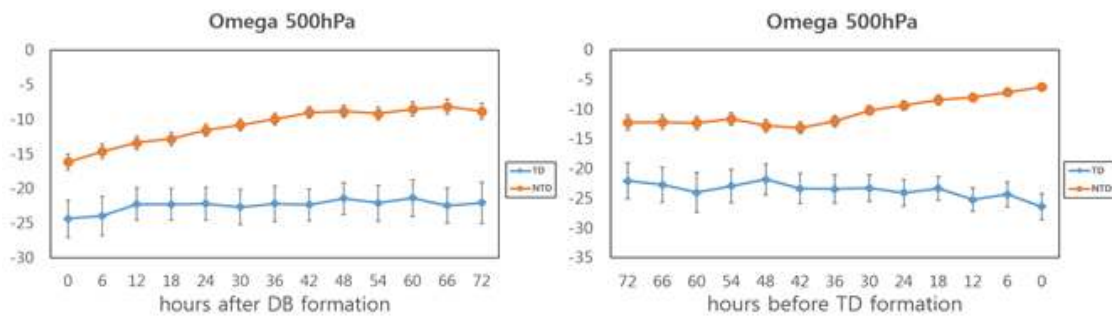


그림 31. 그림 23과 같음, 단 연직 기압 풍속

3. 태풍발생 탐지/예측 모델 개발

본 연구에서는 위성자료 및 수치모델 자료에서 추출한 다양한 공간 변수들을 활용하여 태풍 발생 여부를 판단하는 모델을 구축하기 위해 기본적인 통계기법과 보다 발전된 기계학습 기법이 이용되었다(그림 32).

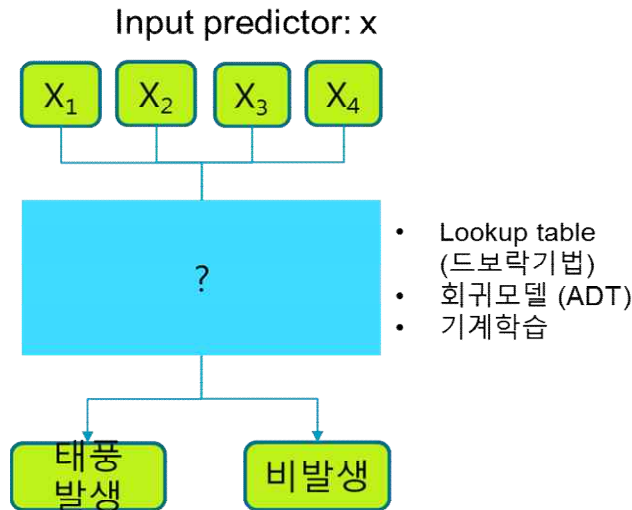


그림 32. 위성 자료 및 수치모델 자료에서 추출한 태풍발생 인자들과 태풍발생, 비발생의 관련성을 찾기 위해서는, 여러 가지 통계적인 기법이나 기계학습 등이 이용된다.

연구에서 활용할 기계학습 기법은 결정나무(decision trees), random forest(RF), support vector machines, 인공신경망(artificial neural networks) 등이다(그림 33). 결정나무는 일반적으로 하나의 샘플 그룹을 서로 이질적이며 각각은 균질한 두 그룹으로 나누는 반복적인 이진 분할 기법을 기반으로 하고 있으며 이질성을 정량화 하는데 Information Gain, Gain Ratio, Gini 지수 등 다양한 형태의 지수를 사용한다.

본 연구에서는 RuleQuest Research, Inc.에서 개발한 See5.0 결정나무 모듈을 활용하고자 한다. RF는 널리 활용되고 있는 결정나무의 한 종류인 Classification and Regression Trees(CART)를 기반으로 하고 있으며 결정나무의 단점으로 알려져 있는 과모의(overfitting), 훈련자료에 대한 높은 민감도를 완화시키기 위해 두 가지 무작위 기법을 도입하였다.

RF는 무수히 많은 CART를 만들어 내서 그 결과를 통합하는 형태를 띠고 있으며, 각각의 CART를 생성할 때 동일한 훈련자료를 사용하는 것이 아니라 무작위로 추출한 일부분만을 사용하고, 또한, 각 CART의 노드에서 전체 입력 변수를 사용하는 것이 아니라 무작위로 추출한 일부분의 변수만을 사용한다. 이를 통해 독립적인 결정나무가 무수히 많이 생성되고, 다양한 결과를 앙상블 형태로 통합함으로써 그 정확도를 보다 높일 수 있는 장점이 있다.

Support vector machines은 자료를 가장 효과적으로 구분하는 초평면(hyperplane)을 결정하는 알고리즘으로 적은 훈련자료로도 탁월한 효과를 보여 주는 장점이 있다. 일반적으로 자료가 가지고 있는 차원(dimension)에서는 그 구분이 쉽지 않기 때문에 보통 고차원으로의 변형을 통하여 작업을 수행하게 되는데 이는 Kernel function이라고 한다. Support vector machines에서는 Kernel function의 선택이 매우 중요하며 적절한 Kernel function을 선택했다 하더라도 그와 관련된 다양한 파라미터들을 어떻게 최적화하는지도 결과에 상당한 영향을 미친다.

인공신경망은 인간의 신경망의 기작을 본 따 만든 알고리즘의 형태이다. 뉴런(neuron)이라고 불리는 개체가 망의 형태로 구성되어 있는데, 일반적으로 많이 활용되고 있는 인공신경망은 다중 레이어 전진 역전파(multi-layer feed-forward backpropagation) 신경망이다. 각 뉴런은 입력값을 받아 들여 선형 및 비선형 변환을 통해 하나의 출력값을 만들어 내고 이 출력값은 다음 레이어의 뉴런에 입력값으로 사용되어 진다. 인공신경망은 에러를 낮추는 방향이나 에러의 도함수 기울기를 낮추는 방향으로 반복적으로 수행되며 다양한 파라미터가 사용되기 때문에 망의 구조나 파라미터들을 적절히 결정하는 것이 매우 중요하다. 이상에서 기술한 기계학습 기법에 초점을 맞추어 다양한 위성자료에서 추출한 공간변수들을 바탕으로 태풍 발달 여부를 모의하였다.

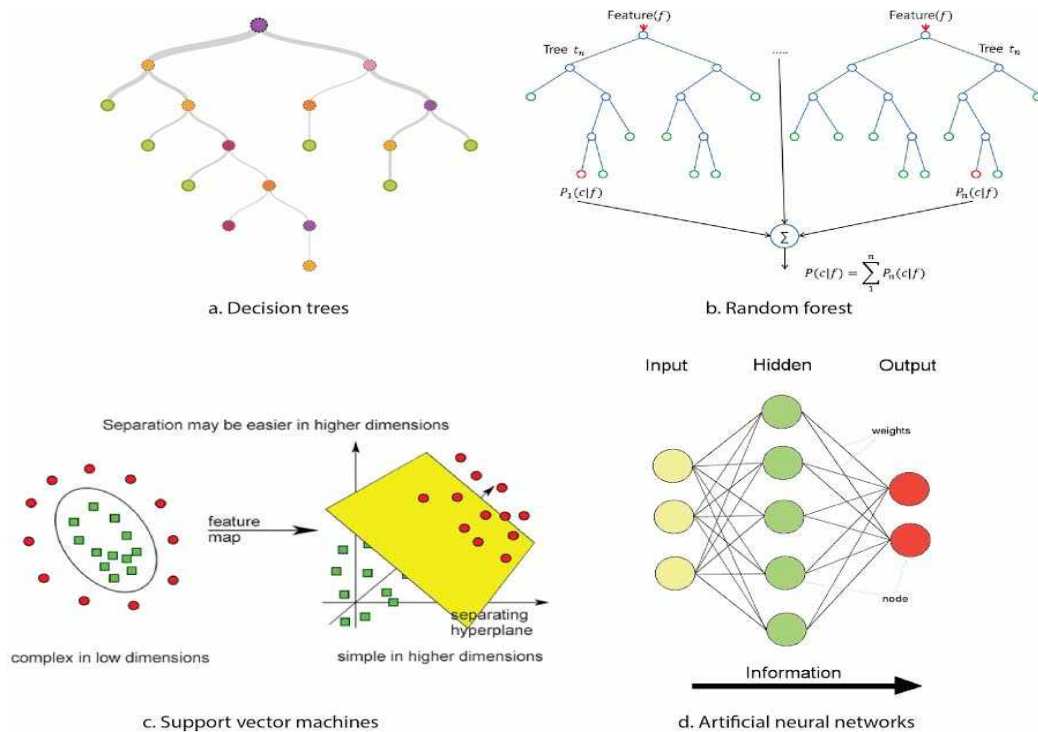


그림 33. 본 연구에 사용될 4가지 기계학습 모식도

가. 극례도 Windsat 해상풍 자료 기반 태풍발생탐지 모델 개발

□ Decision tree 기반 WindSat 태풍발생탐지 모델

3.2.가 절에서 언급된 WindSat 기반 태풍발생탐지 인자들을 이용하여, 결정나무 기법을 이용한 태풍발생탐지 모델을 개발하였다. 개발된 8개의 WindSat 위성 인자를 입력자료로 C5.0을 이용하여, 기계학습의 훈련(Training)과 검증(Validation)을 실시하였다. 그 결과는 8개 인자 중 역확인자인 wind_ave, wind_cv_fix, wind_pladj 만 결정나무 기반 알고리즘에서 이용하며, WindSat 강수 관련 인자들은 이용하지 않는 규칙이 결정되었다. WindSat 기반 결정나무 모델에서 입력변수의 중요도가 표 9에 정리되었다. 이 결과는 태풍발생과 비발생을 탐지하는데 wind_ave와 wind_ave_fix가 가장 중요한 입력 인자로 판단한다.

C5.0은 결정나무를 통해서 일련에 규칙들을 생성하며, 결정나무의 장점들은 이러한 규칙들이 직관적이며 물리적인 해석이 용이하다는 것이다. 개발된 모델에서는 6개의 규칙들이 생성되었다.

표 10에서 제시되는 규칙 5에 의하면, 시스템 내의 바람이 충분히 강하다면(wind_ave >

16.39 kts), 어느 정도의 최소한의 Symmetry만 만족해도 태풍으로 발생할 수 있음을 나타낸다. 이 규칙은 예보관이나 분석자의 정성적인 이해, 태풍 발생의 경우 충분히 강한 바람과 대칭적인 저기압성 순환이 지상 근처에 존재한다는 통상적인 관점과 일치한다. 다만 이 연구에서 제시하는 점은 태풍 발생의 정성적인 기준을 제시한다는 것이다. 규칙 4에 의하면, 바람이 강하더라도($wind_ave > 16.39$ kts) 저기압성 순환의 대칭성이 높지 않다면($wind_cv_fix \leq 0.289$) 태풍으로 발생하지 않는다. 태풍 발생/비발생 판단 시 불확실성이 커지는 영역은, 중간 정도의 바람 강도이다. 이 경우는, 순환의 대칭성이 커도 추가적인 변수인 $wind_pladj$ 의 조건을 만족하지 않으면 비발달로 판단하나, 정확도는 0.734로 다른 규칙 조건 보다 다소 떨어진다.

표 9. 결정나무 모델에 의해 결정된 WindSat 입력 변수의 중요도 (Park et al. 2016)

Order	Predictor	Attribute usage
1	$wind_ave$	90%
2	$wind_cv_fix$	89%
3	$wind_pladj$	64%

표 10. WindSat 기반 태풍발생탐지 결정나무 방법에 의한 규칙들 (Park et al. 2016)

Rule	Descriptions	Classification	Accuracies
1	If $wind_pladj \leq 65.7895$, then a disturbance will not develop.	Non-dev	0.952
2	If $wind_ave \leq 11.49$, then a disturbance will not develop.	Non-dev	0.950.
3	If $wind_ave \leq 16.39$ and $wind_cv_fix \leq 0.352$, a disturbance will not develop.	Non-dev	0.905
4	If $wind_cv_fix \leq 0.289$, then a disturbance will not develop.	Non-dev	0.864
5	If $wind_ave > 16.39$ and $wind_cv_fix > 0.289$, then a disturbance will develop into a TC.	Dev	0.933
6	If $wind_ave > 11.49$ and $wind_cv_fix > 0.352$ and $wind_pladj > 65.7895$, then a disturbance will develop into a TC.	Dev	0.734

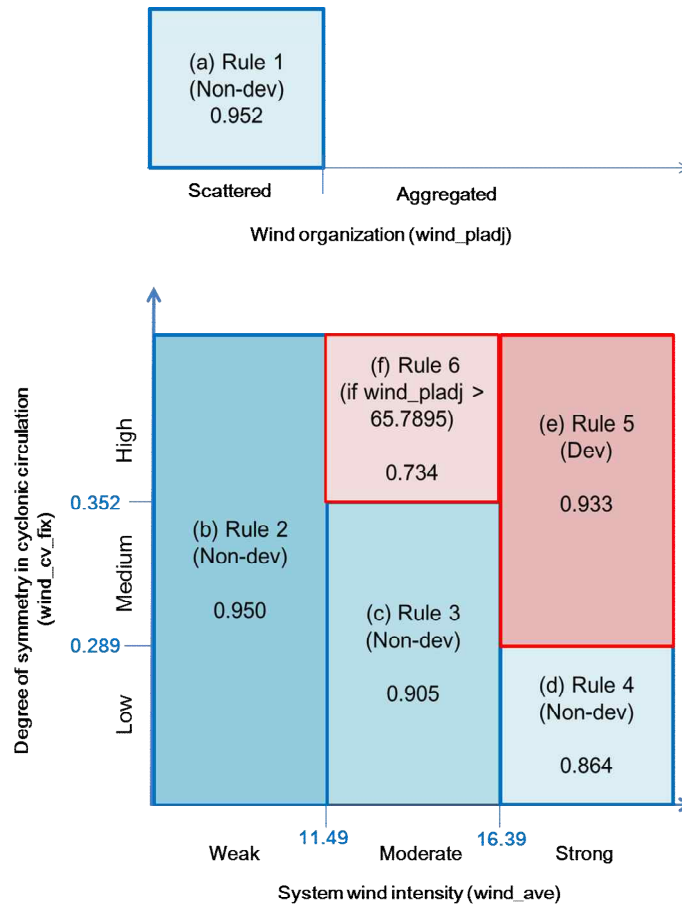


그림 34. Decision Tree(결정나무) 방법을 이용한 WindSat 자료 기반 태풍발생탐지 모델 개념도 (Park et al. 2016)

표 11. 검증 정확도 평가용 분류 표

		Reference	
		Develop	Non-develop
Model classified	Develop	Hit (H)	False (F)
	Non-develop	Miss (M)	Correct negative (N)

개발된 태풍발생 탐지 기술의 검증은 기계학습 Training에 사용되지 않은 2008-2009년 JTWC Best track 자료를 이용하여 수행하였다. 검증을 위해 평가 지수로는 탐지율 뿐 아니라 오탐지율도 고려하여야 한다. 탐지율(POD)와 오탐지율(POFD)은 다음과 같이 정의된다.

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

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

$$\text{OA (overall accuracy)} = (N+H)/(N+F+M+H) \times 100\%$$

표 11에서 나타나는 것처럼 H, M, F, N은 각각 Hit, Miss, False, Correct Negative를 나타낸다. 탐지율(POD)은 실제 발생 태풍 중 모델에서 태풍 발생하는 비율을 계산한 값이다. 탐지율의 범위는 0-100으로, 값이 높을수록 태풍 발생 예측을 잘하는 모델이다. 오탐지율은 태풍으로 발달하지 않는 열대 섭동 중 개발된 기계학습 모델에서 태풍으로 예측한 비율을 계산한 값이다. 오탐지율의 범위도 0-100으로, 값이 낮을수록 열대 섭동을 태풍과 잘 구분하는 모델이고, 100에 가까울수록 태풍이 아닌 섭동을 태풍으로 잘못된 탐지를 하는 모델이다.

표 12. 2008-2009년 검증기간 동안의 Hindcast validation에 의한 Contingency Table (Park et al. 2016)

Validation		Observed		Sum of classification
		DEV	non-DEV	
Classification	DEV	41	72	113
	non-DEV	2	181	183
Sum of observation		43	253	296
Overall Accuracy		75.00%		
Hit Rate		95.35%		
False Alarm Rate		28.46%		

2008-2009년 과거 경우 검증 결과 기계학습을 이용한 태풍 태풍발생탐지 기술의 탐지율 95.35%, 오탐지율 28.46%의 성능을 보였다(표 12). 검증 기간 동안 WindSat 위성이 관측할 수 있었던, 약 43개의 태풍 발달 시스템 중, 41개를 정확하게 태풍 발생으로 탐지하였다. Wood et al.(2015)에 의하면 정지궤도 위성을 이용한 DAV 방법의 경우 탐지율이 96.9%, 오탐지율이 25.6%였다. 본 연구결과의 탐지율과 오탐지율 등의 결과는 선행연구의 정지궤도 위성 기반 결과와 상응할 만한 좋은 성능을 보였다.

개발된 모델의 Lead time은 실제 모델이 태풍 발달 시점으로부터 얼마나 빠르게 예측이 가능한지에 대한 시간이다. 양의 값을 가질수록 태풍 발달 시점보다 빠르게 예측을 할 수 있는 것이고, 음의 값을 가질수록 실제 태풍 발달보다 늦게 태풍발생을 결정한다는

의미이다. 그림 34는 WindSat에서 관측한 첫 번째 관측 시간의 히스토그램을 JTWC 태풍발생 시간에 비교한 상대적인 시간에 따라 나타내었다. 극궤도 위성에서 첫 번째 관측이 이뤄진 이후, 곧 개발된 모델에 의해 태풍으로 탐지되는 경향을 찾았다. 본 연구에서 태풍발생이 탐지되는 시간은 평균적으로 태풍으로 발생하기 약 1.5 h 이전으로 계산되었다. 17 개의 지연된 태풍탐지의 경우 그 중 12개의 경우는 지연된 관측으로 인해 발생되었다. 따라서, 현재 이 섹션에서 개발된 알고리즘은 TCC자료를 이용하여 태풍발생 이전 선행시간이 길지 않다. 그러나, 실제 현업에서 본 연구과제에서 개발된 태풍씨앗섭동 데이터 베이스를 이용할 경우 탐지 시간이 더 빨라질 수 있다.

□ 선형분별법과 기계학습(DT, RF, SVM)간 비교 검증 연구

본 연구에서는 동일한 WindSat 위성인자를 이용하여 다른 통계적방법과 기계학습 방법을 이용하여 태풍발생탐지 예측 모델을 개발하였다. 즉, 세 가지 기계학습 방법인 Decision Tree, Random Forest, Support Vector Machine 기술과 선형분별법(Linear Discriminant Analysis, LDA)을 적용한 4가지 알고리즘의 성능을 비교검증을 실시하였다.

전반적으로, 개발된 기계학습 기반 모델들과 LDA 기반 모델과의 성능 비교를 수행했을 때, 기계학습 적용 모델의 Hit rate이 94-96 %로, LDA 모델의 Hit rate인 63% 보다 훨씬 높게 나타났다. 그러나 False Alarm rate의 경우 LDA가 더 낮은 것으로 나타났다. 태풍탐지 선행시간을 분석했을 때, 같은 샘플을 이용하더라도, LDA를 이용하는 것 보다 기계학습을 이용했을 때 탐지를 더 빨리 할 수 있는 것으로 나타났다. 따라서 WindSat 인자를 이용한 태풍발생탐지 모델 개발의 경우 선형분별법 보다 기계학습 방법을 이용하는 것이 성능이 더 우수한 것으로 판단되었다.

나. COMS MI 기반 태풍발생탐지 모델 개발

COMS MI 기반 태풍발생탐지 모델 개발에서는, 2011년 4월부터 2015년 12월까지의 태풍섭동자료를 이용하여 개발을 하였다. 본 연구과제를 통하여 얻어진 NCEP FNL 기반 태풍 씨앗 섭동 자료를 이용하였으며, 발달 섭동과 소멸 섭동을 구분하여 6시간별 중심 위치 및 강도를 예측해 만들었다. 이 때 만들어진 트랙 자료는 정지궤도 위성 자료에 바로 적용하기에는 시공간 해상도가 무척 낮아, 정지궤도 위성 정보를 이용해 중심을 재조정하는 것이 필요하다. 이를 위해 시간적으로는 선형 내삽을 한 뒤, Pineros et al.(2011)에서 제시한 DAV 기술을 도입하여, 고해상도의 섭동 중심 정보를 취득하였다. 그리고 앞서 제시한 보조, 통계, 공간 변수들을 추출한 뒤, DT와 RF 모델에 적용하여, 정확도 평가를 진행하였다. 2011년에서 2014년까지의 샘플은 학습자료로, 2015년 샘플은 검증자료로 사용하여 모델 구축

및 평가를 진행하였다.

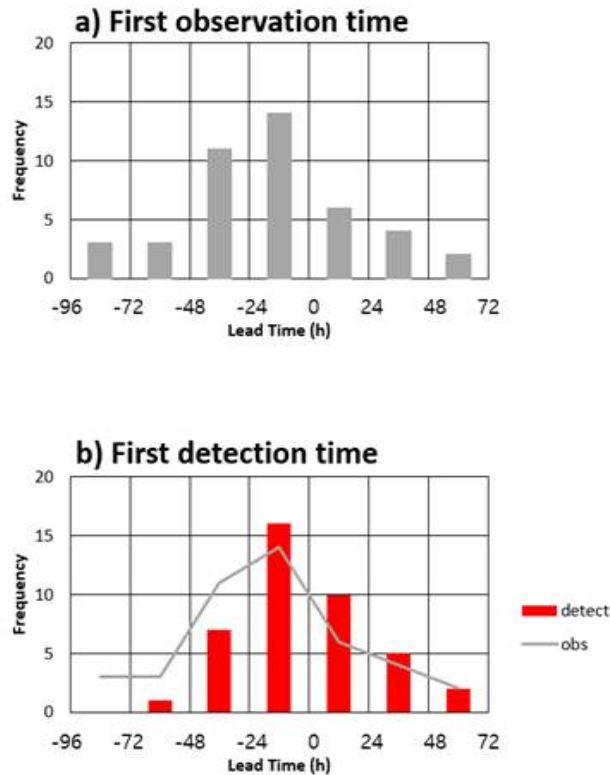


그림 35. (a) WindSat에서 관측한 첫 번째 관측 시간의 히스토그램을 JTWC 태풍발생 시간에 비교한 상대적인 시간에 따라 나타냄, (b) 발달시스템의 경우 처음으로 태풍으로 탐지된 시간의 히스토그램 (Park et al. 2016)

모델 분류 결과, DT와 RF 모델 모두 90%를 초과하는 높은 탐지율을 보였으나, DT의 경우 59%, RF의 경우 39%정도의 높은 오탐지율을 나타내었다(표 13). 즉, 이 검증 결과는 각 시스템 트랙의 발달 또는 소멸 과정 중에 발달한다는 신호가 한 번이라도 발생하였을 경우, 당 트랙 전부를 발달 신호가 잡히는 트랙이라고 간주하여 평가한 결과이다. 이 기준을 조정할 경우 탐지율과 오탐지율이 변동될 수 있다.

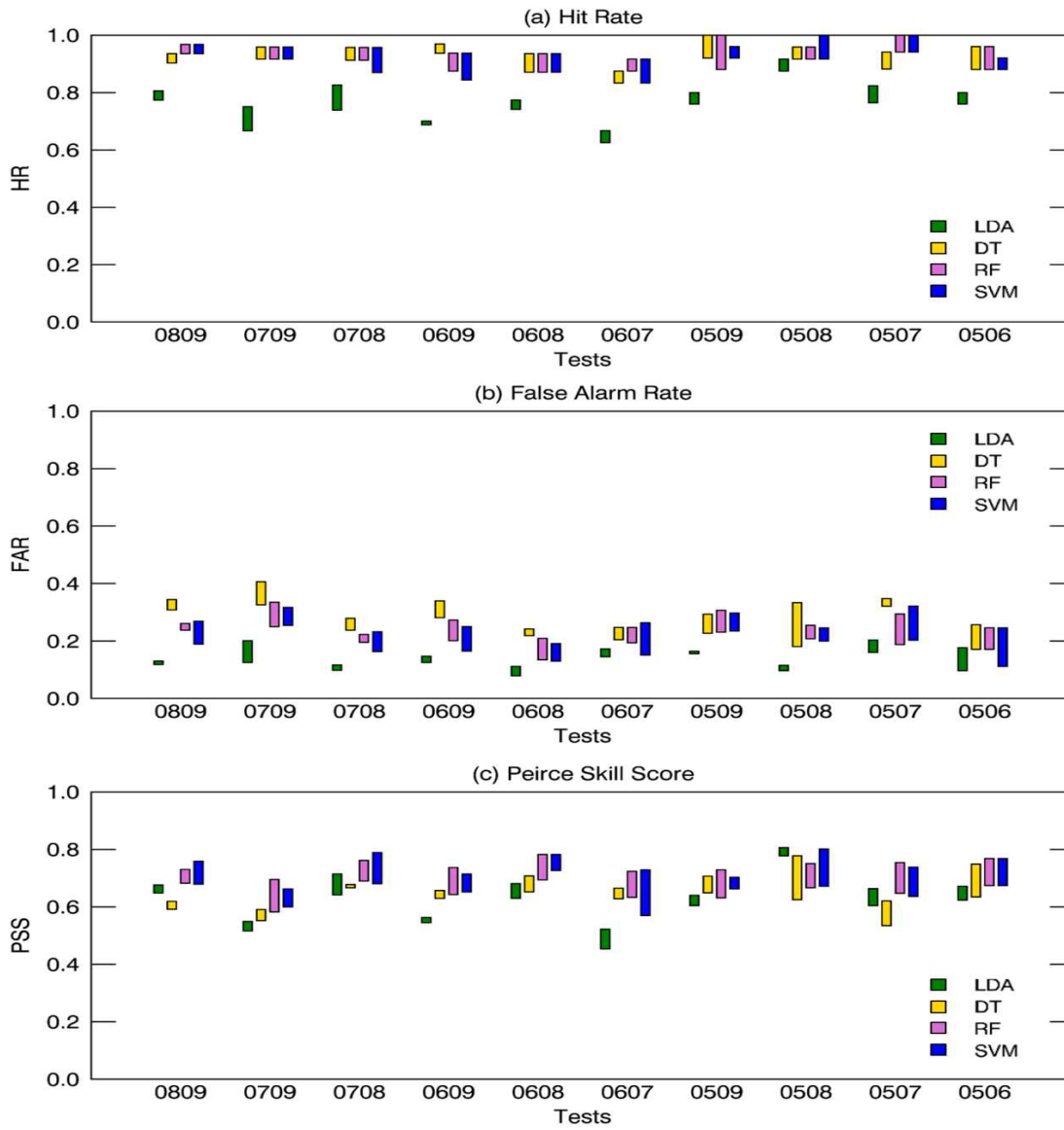


그림 36. LDA, DT, RF, SVM 방법의 성능 비교를 교차 검증을 통해 실시: (a) Hit rate, (b) False Alarm Rate, (c) Peirce skill score. 아래 숫자는 검증을 실시한 기간의 년도를 의미함, 기계학습의 시험자료와 검증자료를 연도별로 교차하면서, 검증을 실시함

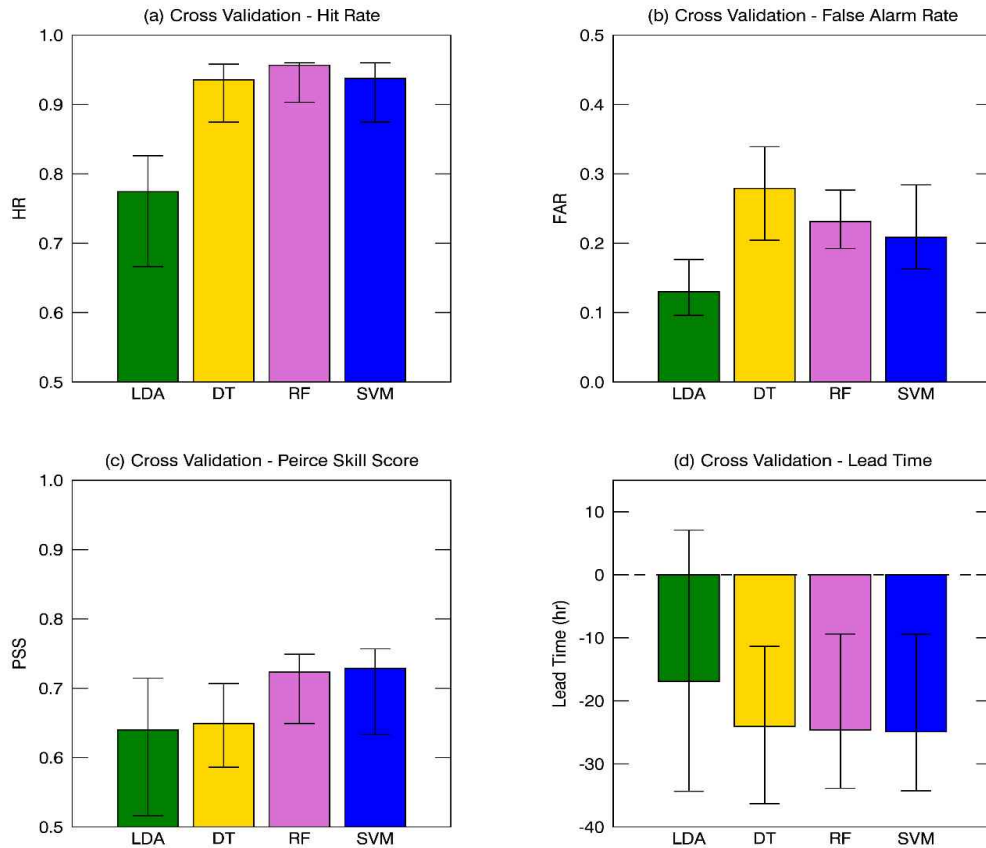


그림 37. 교차검증을 통한 LDA, DT, RF, SVM 방법의 (a) Hit rate (POD), (b) False Alarm Rate, (c) Peirce Skill Score, (d) Lead Time

표 13. COMS/MI 기반 태풍발생탐지 모델의 검증 성능

(a) 35knot 기준 DT single_Hit		Original	
		Develop	Non-develop
Prediction	Develop	16	33
	Non-develop	0	23
POD = 100 % POFD = 59.0 % OA = 54.2 %			
(b) 35knot 기준 RF single_Hit		Original	
		Develop	Non-develop
Prediction	Develop	15	22
	Non-develop	1	34
POD = 93.8 % POFD = 39.3 % OA = 68.1 %			

그림 38에 보이는 것처럼, 개발된 태풍발생탐지/예측 모델의 정확도와 함께 변수 중요도 평가를 진행하였다. DT의 경우, See5.0에서 variable usage라는 변수 중요도 정보를 제공한다. Variable usage는 생성된 DT 알고리즘에서 각 변수가 최종 결정까지의 사용되는 빈도를 이용해 순위를 매기는 것이다. RF의 경우 R-gui의 RF 패키지를 사용하였고, 여기에서는 Mean decrease accuracy 라는 변수 중요도를 제공한다. Mean decrease accuracy는 각 변수를 제거하고 모델을 구축했을 시, 정확도가 얼마나 떨어지느냐를 평가한 척도이다.

각 모델에 대해 변수중요도 평가 결과, 두 모델 모두 DAV, lon, lat, DOY의 중요도가 높게 나온 것을 알 수 있었다. DAV는 선행연구인 Wood et al.(2015)에서 단일변수와 단일 임계값만으로 태풍 발생 전조 탐지를 위해서 사용된 바 있다. 그리고 위치, 날짜 등의 보조변수의 경우, 쇠퇴하는 섭동의 쇠퇴 지점과 발달하는 섭동의 발달 지점이 통계적으로 많이 떨어져 있고 (발달-육지 근처; 쇠퇴-원해), 발달하는 섭동이 빈번하게 발생하는 계절적 시기가 구분이 되어있어 중요도가 높게 나타난 것으로 보인다. 공간변수 중에서는 SWIR_mean과 wv_mean의 기여도가 높게 나타났다. 태풍 발생 직전 대류의 조직화에 의해 면적이 넓어지기 때문에, SWIR_mean은 발생과 비발생의 경우 차이가 클 수 있다. 또한, 태풍 발생전 반복적인 대류활동에 의해 대기중의 수증기량이 증가하여 wv_mean이 태풍 발생 시에 중요한 변수가 될 수 있다.

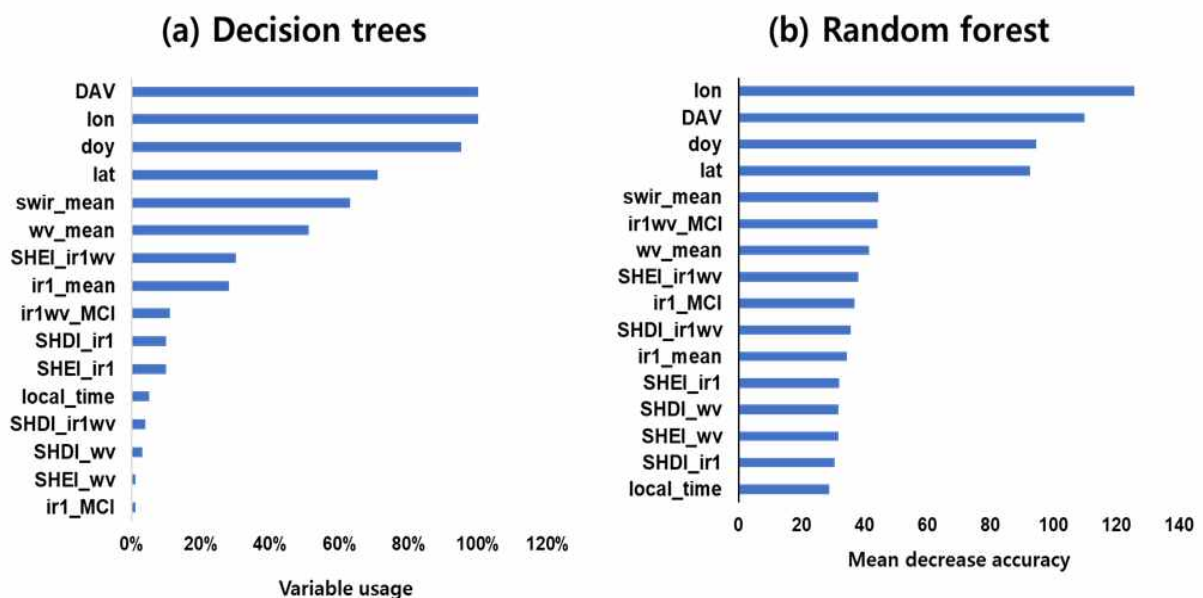


그림 38. DT와 RF 태풍탐지/예측 모델의 변수중요도 비교

다. Himawari-8 AHI 기반 태풍발생탐지 모델 개발

Himawari-8 AHI에서 산출된 변수들을 기계학습에 적용하여 태풍발생탐지 모델을 개발하였다. 본 연구과제의 3.1절에서 설명된, 열대 섭동 트랙을 이용할 경우 태풍 발달 72시간 전부터 발달 섭동의 위치 정보가 가용하다. Himawari-8 위성의 운용이 2015년 7월에 시작했기 때문에, 샘플의 부족을 보완하기 위해 여러 가지 샘플링 기법을 적용하였다. 제한적인 샘플 수를 이용하여 샘플에 대한 의존도를 낮추기 위해 k-fold cross validation 기법을 이용하였다. 본 연구에서는 3 folding을 하여 모델 검증을 수행하였다. 기계학습의 경우 모델의 학습 과정에서 샘플 비율이 결과에 영향을 준다. 최적의 샘플 비율을 얻기 위해 random sampling을 하여 민감도 테스트를 실시해 발달 (developing): 비발달(non-developing) 샘플의 비가 1:2 일 때 가장 좋은 훈련 결과를 나타내는 것을 확인하였다.

본 연구에서는 앞에서 사용하였던 세 가지 기계학습 방법들이 태풍발생 탐지 모델 개발에 사용되었다. 모델의 샘플은 Himawari-8 AHI에서 산출된 변수(predictor)와 범주(target)를 포함하고 있다. 본 연구는 발달, 비발달 하는 대류를 예측하기 위해 이진 분류(binary classification) 기반으로 개발되었다. 각 모델은 공통적으로 훈련(training)과 검증(validating) 과정으로 구성된다. 샘플을 통해 훈련된 기계학습 모델들은 검증 과정에서 결과가 알려지지 않은 샘플을 예측하여 실제 결과와 비교를 통해 모델 검증이 된다.

정지궤도 위성 관측의 경우 1시간 간격의 샘플이 이용되기 때문에 검증 결과를 다시 각 태풍별로 수집하여 “탐지” 혹은 “미탐지”를 결정하는 과정이 필요하다. 대류의 경우 일변동이 있기 때문에 검증의 변동성이 보인다. 태풍의 발달 결정을 위해 24시간 내에 매 시간 태풍발생 탐지 횟수에 대한 민감도 테스트를 실시하였다. 하루 내에 태풍발생 탐지 횟수가 낮을수록 태풍의 예측은 높아지지만 반대로 태풍이 아닌 열대 섭동에 대한 예측 성능이 낮아지고, 탐지 횟수가 높을수록 태풍의 예측이 낮아지게 된다. 그 결과 개발된 모델에서, 하루 10회 이상 태풍발생 탐지가 되었을 때를 해당 시스템의 태풍 발생 탐지 정의하였다. 이 기준을 기반으로 개발된 모델들의 성능 비교를 통해 모델 간 비교를 실시하였다.

그림 39은 Himawari-8 AHI에서 산출된 변수들의 RF 훈련 과정에서 계산된 변수중요도를 나타낸다. 가장 변수중요도가 높은 변수는 구름의 조직화를 나타내는 DAV14로, 앞에서 개발된 COM/MI의 결과와 일관되게 나타났다. 그 다음으로 중요한 변수들은 대류의 강도와 관련된 인자들이다. 두 번째로 중요한 변수인 11번과 14번 채널의 차이에 대한 변수인 D(11-14)는 WV-IR을 계산한 변수로, 대류권계면 위로 높게 발달하는 Overshooting

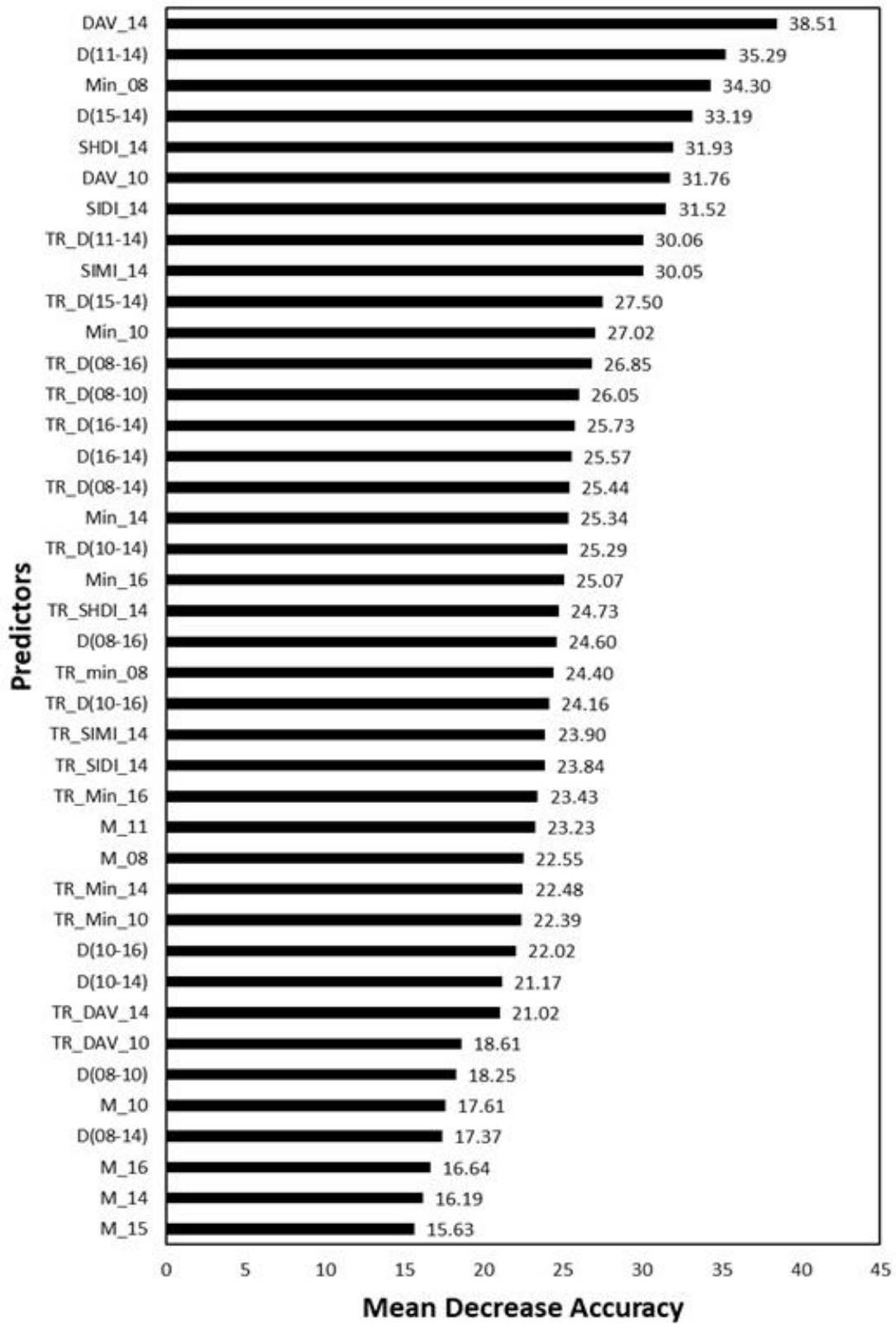


그림 39. Himawari-8 AHI에서 산출된 변수들의 훈련 과정에서 계산된 변수중요도

Convective Cell에서는 양의 값을 보인다. 최근 연구에서 태풍 발생 전 높게 발달하는 대류셀에 의한 양의 와도(Vortical hot tower)의 역할이 강조된 바 있다(Montgomery et al. 2006). 본 연구에서 개발된 위성인자와 기계학습 방법이 태풍의 중요한 역학과정을 잘 고려하여, 태풍발생을 탐지하고 있음을 나타낸다.

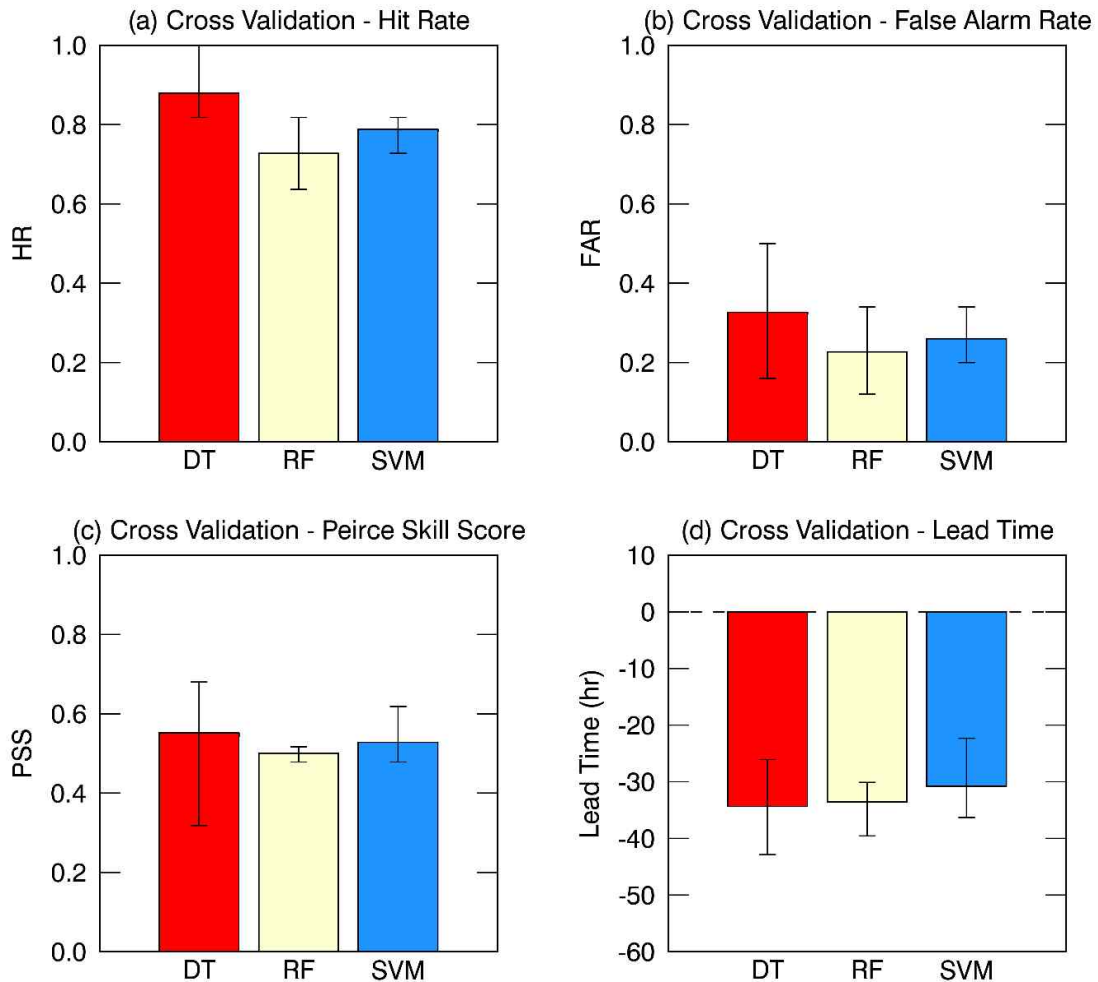


그림 40. Himawari-8 AHI에서 산출된 변수를 이용하여 의사 결정 나무 (red), 랜덤 포레스트 (ivory), 서포트 벡터 머신 (blue) 기반의 태풍발달 탐지 모델의 (a) 적중률, (b) 오경보율, (c) PSS, (d) 리드 타임 비교 결과

그림 40는 세가지 기계학습(DT, RF, SVM)을 이용하여 개발된 Himawari-8/AHI 기반 태풍발생탐지/예측 모델을 성능을 비교하였다. 탐지율의 경우 세 모델에서 평균 80%의 성능을 보인다. DT가 가장 좋은 성능(87%)을 보이고, RF(73%)가 가장 낮은 성능을 보인다.

오탐지율(False Alarm Rate)의 경우 평균 27%의 성능을 보인다. 가장 오탐지율이 낮은 모델은 RF(23%)이고, 가장 높은 모델은 DT(33%)이다. 탐지율과 오탐지율의 결과는 탐지율이 높을수록 오보율도 같이 높아지는 특성을 보이고 있어, 종합적인 성능을 보기위해 PSS를 확인했다. 평균적인 PSS 는 53%이고, 세 가지 다른 기계학습 방법에서 거의 유사하지만 가장 높은 모델은 DT(54%) 로 RF(50%) 보다 좋은 성능을 보이고 있다. 본 연구의 결과는 일반적으로 기계학습 모델 성능 비교 연구에서 일반적으로 RF가 DT에 비해 더 나은 성능을 보이는 것과 반대되는 결과를 보였다.

라. 전구 모델 기반 태풍발생 탐지/예측 모델 개발

본 연구에서는 NCEP FNL 분석장에서 추출된 2007년부터 2015년까지 5월부터 11월까지 열대 섭동 자료를 이용하여 태풍의 발생을 탐지/예측하는 모델을 개발하였다. 태풍 발달을 판단하기 위해 사용된 환경인자는 앞서 살펴보았던 하층 상대와도, 수분 수렴(발산), 상층 발산, 상대 습도, 연직 바람 시어, 하층 수렴, 중층 연직 바람의 환경인자를 사용하였으며, 각각의 인자의 절댓값 및 48시간 동안의 변화 경향성을 모두 잠재인자로 설정하여 모델을 구성하고 검증결과를 평가하여 최적의 예측을 위한 환경 인자를 선정하였다. 환경인자를 계산하는데 있어 태풍 주변 방위각으로 평균 방법을 사용하여 모델의 해상도와 관계없이 환경인자가 계산될 수 있도록 하였다. 또한 열대섭동의 경의 시공간적인 위치도 태풍으로 발달하는데 중요한 요인으로 작용하기에 열대섭동의 발생일(Julian day), 위도, 경도 위치 정보 또한 태풍 발달 예측 인자로 포함하였다.

태풍 발생을 예측하기 위해 기계학습의 한 방법인 RF를 사용하였다. RF는 두 가지 단계로 진행되는데 모델이 예측을 하기 전 입력된 자료를 토대로 decision tree를 구성하는 학습단계와 학습단계에서 구성된 decision tree를 활용해 예측을 하고 예측 결과의 정확도를 검증하는 검증단계로 진행된다. 열대섭동 자료 중 2007년부터 2013년까지의 자료를 학습단계에 사용하였고, 2014년부터 2015년까지의 자료를 검증단계에 사용하였다.



그림 41. RF 수행을 위한 열대 섭동 발달 과정 구분

발달 단계에 따라 열대 섭동을 구분하기 위해 종료 시점에서 48시간 이전까지의 자료를

열대 섭동이 유지되는 상태(1)로 지정하였고, 48시간 이후부터 종료 시점까지 태풍으로 발달하는 경우(2)와 열대 섭동이 소멸하는 경우(0)로 구분하였다 (그림 41). 이렇게 구분하게 되면 실제 현업 예측에서는 예측시점에서 열대 섭동이 48시간 이내 열대 섭동 상태로 유지되는지(1), 소멸하는지(0), 태풍으로 발달하는지(2)로 예측이 가능하다. 현업 예측에서는 현재 열대 섭동의 자료를 입력한 예측결과가 태풍 발달(2)로 산출되면 ‘48시간 이내 태풍으로 발달할 가능성이 매우 높음’, 열대섭동 유지(1)로 산출되면 ‘48시간 이내 태풍 발달 가능성 존재, 지속적인 관찰 필요’, 열대섭동 소멸(0)로 산출되면 ‘48시간 이내 태풍으로 발달할 가능성 매우 낮음’ 으로 판단을 내릴 수 있을 것이다.

앞서 설명한 환경인자와 시공간 정보를 이용하여 RF 모델을 구성하여 2007년부터 2014년까지의 열대 섭동관련 인자를 학습하였다. 그 결과 예측인자들이 예측 결과에 영향을 미치는 중요도에 대해 확인할 수 있었다(표 14). Mean Decrease Accuracy와 Mean Decrease Gini는 그 값이 클수록 모델에서 주로 사용되는 인자를 나타내고 있다. 사용한 인자들 중 상대와도가 가장 큰 중요도를 나타냈으며, 공간 정보, 상층 발상장의 경향성, 시간 정보 순으로 모델 예측에 많은 영향을 미친다는 것을 확인할 수 있었다.

RF 기계학습은 과거 자료를 학습하여 모델을 구성하는데 소멸하는 열대 섭동의 개수가 태풍으로 발생하는 태풍보다 4배 이상 많기 때문에 발달하지 않는 열대 섭동이 더 많이 학습되어 실제 예측에서도 태풍으로 발달하지 않는 예측으로 편향성을 가질 수 있는 우려가 존재하다. 그래서 이를 방지하고 예측 정확도 향상을 위해서 학습단계에서 사용한 소멸한 열대 섭동의 개수를 태풍으로 발달한 열대 섭동의 개수에 비해 1, 2, 3, 4배로 제한하여 모델을 학습하여 예측을 수행하였고 각각의 예측 결과를 비교했다. RF 예측 결과 학습단계에서 태풍으로 발달하지 않은 열대 섭동의 개수를 제한하여 실행하였을 때 아래와 같은 검증 결과가 도출되었다 (표 15). Hit Rate는 태풍으로 발달한 열대 섭동을 RF가 태풍으로 발달했다고 예측한 경우의 정확도를 계산하였고, False Alarm은 태풍으로 발달하지 못한 열대 섭동을 RF가 태풍으로 발달하였다고 예측한 경우의 정확도를 계산한 것이다. 발달하지 않은 열대 섭동의 개수를 제한할수록 False Alarm Rate과 Hit Rate이 모두 증가하는 것으로 나타나고 있다. 즉 발달하지 않은 열대 섭동을 많이 학습할수록 Hit Rate은 떨어지나 발달하지 않는 열대섭동을 태풍으로 발달 할 것으로 잘못 예보하는 경우는 줄어들게 된다.

모든 열대섭동을 사용하여 학습한 결과에서 Hit Rate 은 0.878로 33개의 태풍 중 29개의

발달을 적절히 예측하였다. False Alarm Rate의 경우는 0.086으로 매우 낮은 값을 보이고 있지만 실제 소멸한 22개의 열대 섭동을 태풍으로 발달하는 것으로 예측하였다. 낮은 False Alarm Rate을 보이고 있으나 실제 발달하지 않은 열대 섭동을 태풍으로 발달했다고 예측한 열대 섭동의 개수가 다소 많으므로 현업 활용시 이에 대한 고려가 필요할 것이다.

표 14. 전구수치모델 기반 RF 결과 각 인자들에 대한 중요도

Index	Mean Decrease Accuracy	Mean Decrease Gini
Julian day	63.3	262.6
Latitude	76.3	283.2
Longitude	78.3	300.6
Relative vorticity	97.2	415.4
Moisture divergence	29.9	170.0
Trend of moisture divergence	60.8	255.3
Divergence	21.1	147.2
Trend of divergence	68.7	303.5
Relative humidity	49.0	214.1
Wind speed	31.3	159.1
Vertical wind shear	46.3	218.1
Convergence	18.4	144.6
Omega	26.7	157.1

표 15. Random forest(RF) 결과를 이용해 계산한 예측 정확도. NTD/TD는 소멸한 열대 섭동의 개수를 제한한 비율, TD(OBS)와 DB(OBS)는 실제 태풍으로 발달한 열대 섭동의 개수와 소멸한 열대 섭동의 개수, 그리고 HIT(RF)는 RF 모델이 태풍으로 발달하는 태풍을 예측한 횟수, False Alarm(RF)는 태풍으로 발달하지 않은 열대 섭동을 태풍으로 발달한다고 잘못 예측한 횟수

NTD/TD	TD (OBS)	HIT (RF)	DB (OBS)	False Alarm (RF)	Hit Rate	False Alarm Rate
1	33	33	254	81	1.0	0.318
2	33	32	254	63	0.969	0.248
3	33	32	254	44	0.969	0.173
4	33	31	254	29	0.939	0.114
No limit	33	29	254	22	0.878	0.086

제 4장. 목표 달성도 및 관련 분야 기여도

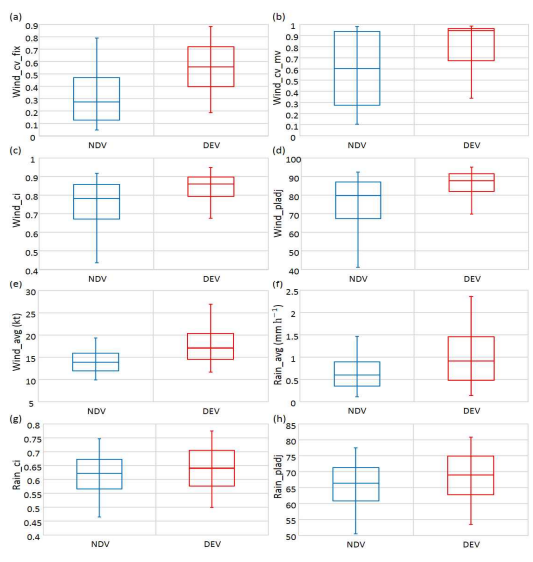
1. 목표 달성도

가. 연구 개발 성과 및 평가방법

연도	세부 연구목표	세부 연구 개발 내용	평가목표 및 착안점
1차 연도	위성관측 자료를 이용한 태풍발생 탐지/예측 인자 개발	<ul style="list-style-type: none"> WindSat 해상풍 자료를 이용하여 태풍 발달 직전 강풍 영역 조직 정도, 중심 부근 시스템 강도, 순환의 대칭성을 정량화하는 인자를 개발 TRMM TMI 자료를 이용하여 태풍 발달 과정에 변화하는 중규모 대류계의 면적과 강도를 정량화 하는 인자를 개발 정지궤도 위성의 적외, 수증기 채널과 채널 간 차이 등 대류의 패턴과 강도를 정량화 하는 인자 개발 	1. 향후 현업에서 이용가능하며, 태풍발생과정을 잘 나타내는 정량화된 인자를 개발하였는지?
	위성 인자를 이용한 알고리즘 원형 개발	<ul style="list-style-type: none"> 새로운 위성 인자를 이용 태풍발생/비발생을 탐지하는 객관적 태풍발생탐지/예측 알고리즘 원형 개발 LDA와 기계학습 방법(DT, RF, SVM)들을 활용하여 객관화된 알고리즘을 개발 	2. 예측인자와 태풍발생의 관계를 나타내는 알고리즘 원형(순서도)를 개발하였는가?
	전구 수치 모델을 이용한 열대섭동 탐지	<ul style="list-style-type: none"> GFS 수치모델 분석장을 이용하여 열대서 태평양에서 태풍으로 발달할 가능성이 있는 열대섭동들을 추적 (발달 및 비발달 시스템 포함) 	3. 태풍으로 발생 가능한 저기압성 와도를 가진 열대섭동이 탐지되었는가?
2차 연도	검증과 평가를 통한 위성 태풍발생 탐지 알고리즘 원형 개선	<ul style="list-style-type: none"> 기 개발된 극궤도 위성 기반 태풍발생 탐지/예측 모델을 계속적으로 검증하고 개선함 정지궤도 위성을 이용한 태풍발생 인자들을 정의하여 정지궤도 위성 기반 태풍발생 탐지 모델을 개발함 	4. 기 개발된 태풍발생 탐지/예측 모델을 계속적으로 검증하고 개선하였는지?
	전구 수치 모델을 이용한 및 태풍발생확률	<ul style="list-style-type: none"> 전구 수치모델 분석장에서 대규모 환경변수를 정의하고, 이를 이용한 객관적인 태풍발생 탐지/예측 모델을 개발 	5. 태풍발생에 중요한 대규모 인자들이 도출되었으며, 전구수치모델 기반

	분석		태풍발생탐지/예측모델을 개발하였는가?
	위성-수치모델 결합 시스템 원형 개발	<ul style="list-style-type: none"> 위성정보와 전구수치모델 자료의 정보를 효율적으로 결합하는 방법을 연구 위성-수치모델 융합 태풍발생 탐지/예측 모델이 개발 	6. 위성과 전구수치모델 자료를 융합한 태풍발생 탐지/예측 모델의 원형 개발되었는지?
3차	수치모델과 위성 관측의 효율적 결합 시스템 개선	<ul style="list-style-type: none"> 시스템 최적화를 위해 전구수치모델 시스템과 위성알고리즘을 효율적 결합 방법을 연구 	7. 모델과 위성관측을 결합하기 위해서 적절한 방법이 사용되었는가?
	검증 및 평가	<ul style="list-style-type: none"> 과거 태풍 자료를 이용한 시스템별 검증으로 태풍 발생탐지/예측 시스템의 정량적 평가 	8. 검증과 평가가 정량적으로 이뤄졌는지?
	최종결과 활용을 위한 체계 구축	<ul style="list-style-type: none"> 최종적으로 태풍발생 탐지 및 예측시스템에서 생산된 결과가 표출되어 현업 활용 가능한 체계 구축 	9. 최종결과를 활용한 체계가 구축되었는가?

나. 평가의 근거

작업점/평가	평가의 근거
1. 향후 현업에서 이용가능하며, 태풍발생과정을 잘 나타내는 정량화된 인자를 개발하였는지?	<ul style="list-style-type: none"> 개발된 태풍 발생 인자의 통계- 아래 발생 섭동과 비발생 섭동이 경우 WindSat 해상풍의 강도 및 조직화에서 차이를 보임. 개발된 정량화된 위성 인자들이 태풍 발생 과정을 잘 나타낸다고 할 수 있음  <p>그림 A. 태풍으로 발달하는 시스템(적색 박스)과 비발달 시스템(청색박스)에 대한 개발된 WindSat 위성 기반 태풍발생 인자들의 통계적 비교 (Park et al. 2016)</p>

- 각 시스템별 인자의 타임시리즈: 개발된 인자들이 태풍으로 발달직전 증가하는 경향이 보임. 태풍의 발생-강화 과정을 잘 나타내고 있음.

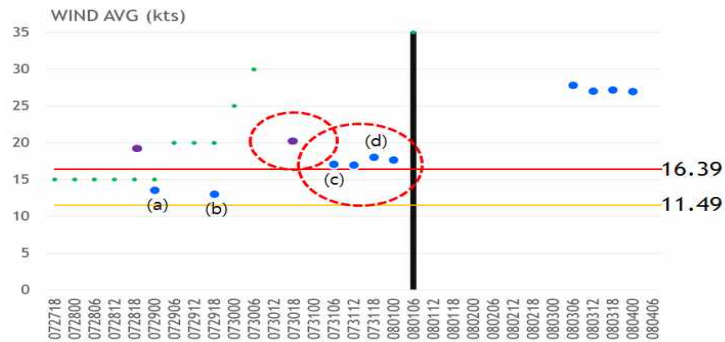


그림 B. 해상풍 자료를 이용한 시스템의 강도 인자(wind_ave)의 시간 변동성: 2015년 태풍 Shouldelor 발달 시

- 새로운 위성 인자를 이용 태풍발생/비발생을 탐지하기 위한 기계학습을 이용한 객관적 알고리즘 원형 개발

2. 예측인자와 태풍발생의 관계를 나타내는 알고리즘 원형(순서도)를 개발하였는가?

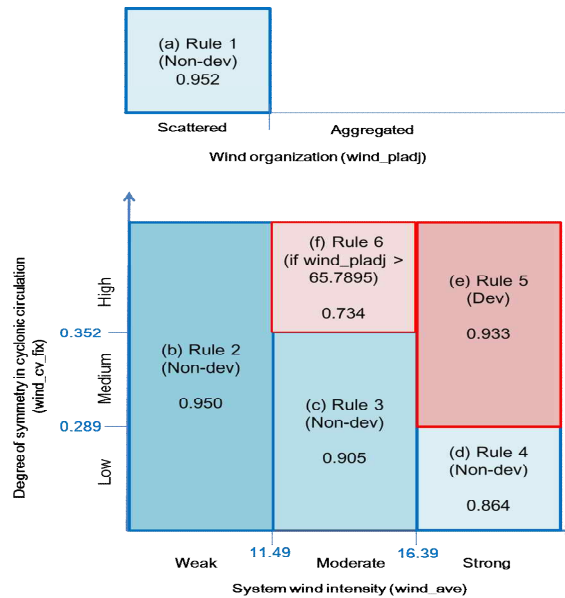


그림 C. 해상풍 자료를 이용한 인자들(wind_ave, wind_cf_vix, wind_pladj)을 이용하여 태풍 발생 및 비발생을 탐지하는 규칙 및 경계값 (Park et al., 2016)

3) 태풍으로 발생 가능한 저기압성 와도를 가진 열대섭동이 탐지되었는가?

상대와도와 원형분산(circular variance)를 결합한 태풍 발생 이전 열대섭동 탐지 알고리즘 개발 (그림 D)- 프로그램 등록 1건

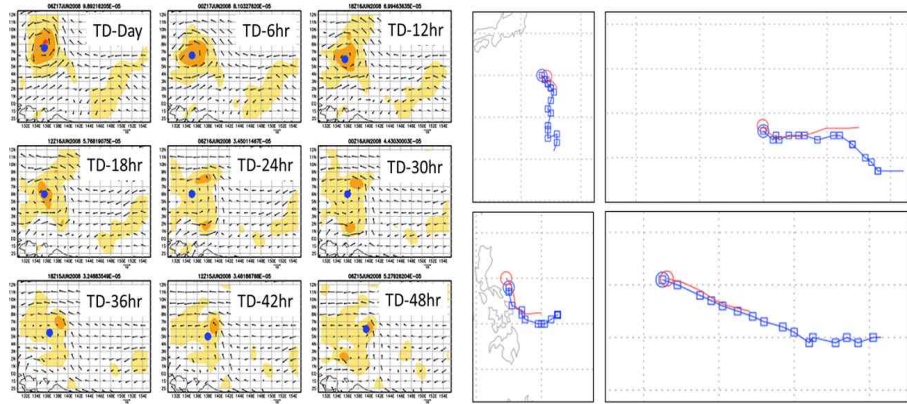


그림 D. (좌) GFS 모델 분석장을 이용한 태풍 발생 이전의 열대 섭동의 위치 추정 알고리즘 결과 예시 (shading:850-hPa relative vorticity, vector : 850-hPa wind), (우) 2015년 INVEST 시스템(붉은색)과 개발된 알고리즘에서 추적된 시스템 (파란색)의 경로비교 결과 예시 - K-fold Cross validation 통하여 기 개발된 위성기반 기반 알고리즘을 기계학습/통계 방법별 계속적으로 검증하였음

4) 기 개발된 태풍발생 탐지/예측 모델을 계속적으로 검증하고 개선하였는지?

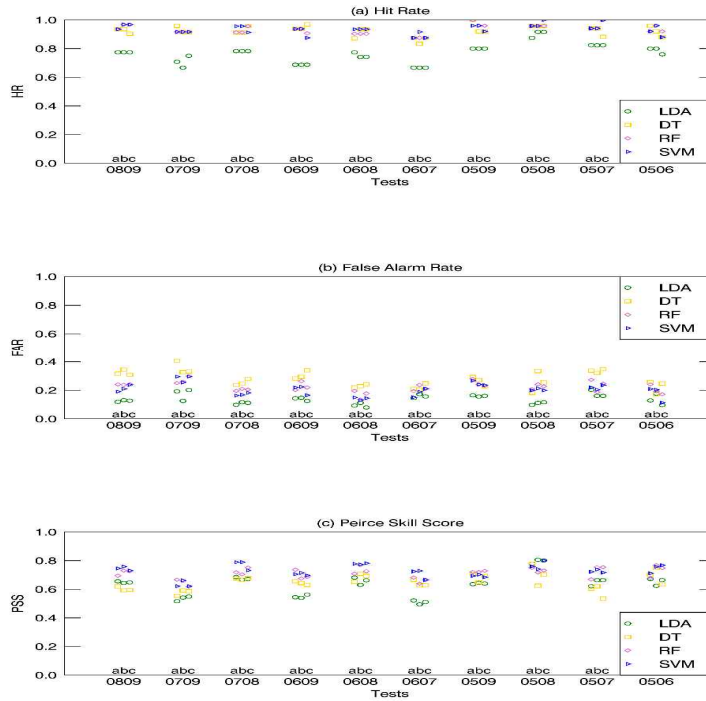


그림 E. 세 가지 기계학습(DT, RF, SVM)과 선형회귀법(LDA)를 이용한 WindSat 기반 태풍발생 탐지 모델을 여러 교차 검증한 결과: (a) Hit rate, (b) False Alarm Rate, (c) Peirce Skill Score

5) 태풍발생에 중요한 대규모

- 전구 수치모델 기반 대규모 인자들이 도출되어, 기계학습의 입력자료로 이용되었음

인자들이
도출되었으며,
전구수치모델 기반
태풍발생탐지/예측
모델을
개발하였는가?

Indices	Variables	Spatial/Temporal info.	Misc.
jdy	Julian Day	DB Day	
lat	Latitude	DB Center	
lon	Longitude	DB Center	
vor5	850-hPa Vorticity	5 degree from DB Center	
vor5_tr	850-hPa Vorticity Trend	5 degree from DB Center, (Max.) 48 hour	
div5	200-hPa Divergence	5 degree from DB Center	Radial
div5_tr	200-hPa Divergence Trend	5 degree from DB Center, (Max.) 48 hour	average of the
vws3_8	Vertical Wind Shear	3-8 degree from DB center	azimuthally
rh5	Relative Humidity	5 degree from DB Center	averaged
rh5_tr	Relative Humidity Trend	5 degree from DB Center, (Max.) 48 hour	values

표 A. 전구수치모델 태풍발생인자 입력 변수 리스트

변수중요도

Index	DB소멸	DB상태유지	DB발달	MeanDecreaseAccuracy	MeanDecreaseGini
jdy	56.2	70.8	42.4	85.8	417.6
lat	55.4	70.0	47.2	87.7	365.9
lon	50.0	73.7	42.8	84.8	380.5
vor5	87.2	62.3	115.3	111.3	626.1
vor5_tr	62.3	59.8	52.0	85.4	508.0
div5	35.4	42.0	41.6	60.3	367.1
div5_tr	32.0	48.2	18.5	59.6	344.8
vws3_8	45.4	67.4	29.6	80.9	404.9
rh5	35.2	53.3	23.5	60.0	389.5
rh5_tr	39.9	49.8	27.7	61.5	366.6

표 B. 개발된 태풍발생탐지/예측 모델에서 인자별 중요도

6) 위성과
전구수치모델
자료를 융합한
태풍발생 탐지/예측
모델의 원형
개발되었는지?

- 태풍발생 48시간 이전과 열대섭동 소멸 48시간 이전에는 발달과 소멸 시스템의 차이가 명백하지 않다고 판단하여, Medium으로 분류하여 알고리즘 산출/검증 Match-up 자료를 구성

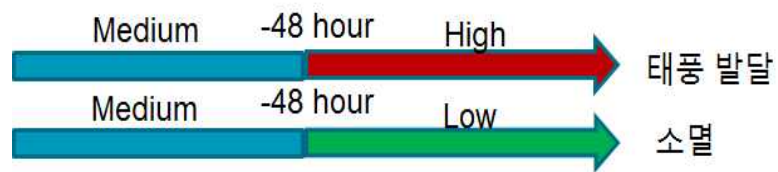


그림 F. 위성과 전구수치모델 융합 태풍발생 탐지/예측 시스템의 최종산출물: 앞으로 48시간 이내에 태풍으로 발달할 확률을 3단계(High, Medium, Low)로 구분함

- 위성-수치모델링 자료 융합을 위해, 전구수치모델 인자와 위성자료 인자를 입력자료로 준비하여 기계학습을 구성하였음

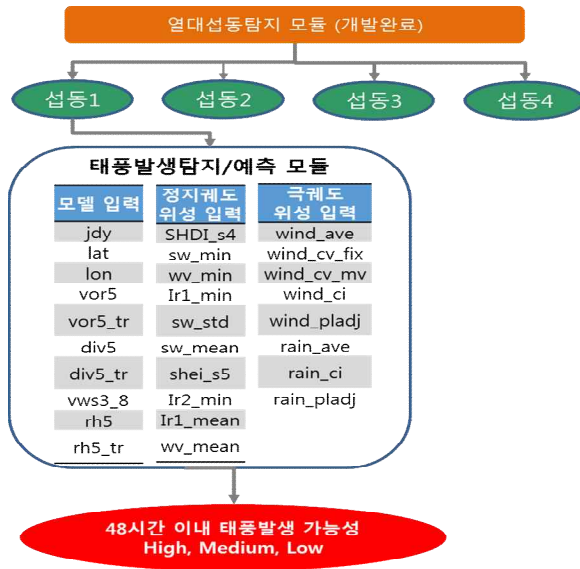


그림 G. 위성-전구 수치모델 자료 융합한 태풍발생 탐지/예측 모델의 원형

	개발된 모델의 의한 48h 이내 태풍으로 발생할 확률		
	발달(High)	지속(Medium)	쇠퇴(Low)
발달 시스템	7	0	1
비발달 시스템	7	0	24

그림 I. 위성-수치모델 융합 태풍발생탐지/예측시스템의 Contingency Table

10. 모델과 위성관측을 결합하기 위해서 적절한 방법이 사용되었는가?

- 모든 위성과 모델 인자를 입력자료로 한 단일한 알고리즘을 고려할 경우, 모델인자의 의존성이 커져서 태풍발생의 다양한 특성을 알 수 없음. 따라서, 태풍 발생 시에 대규모적 관점(전구수치모델 기반), 시스템 조직화(WindSat 위성 기반), 대류의 강도 및 조직화(정지궤도 기반) 등 다양한 측면을 모니터링 할 수 있도록 중간산출물을 모니터링 하며 각각의 알고리즘의 태풍발생 탐지/예측 결과를 산출하여 정보를 제공하도록 결합하였음

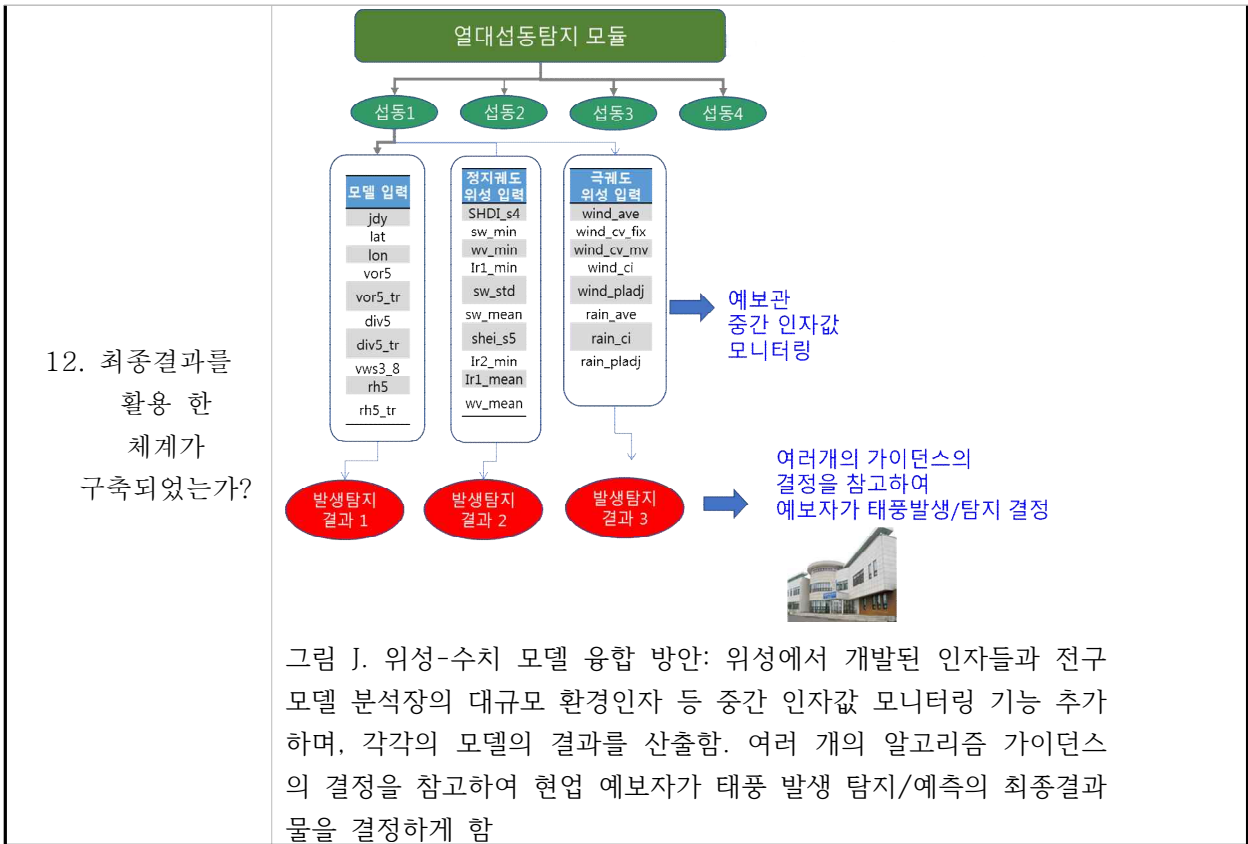
변수명	모델(M)/위성(S)	Mean Decrease Accuracy
mvot	M	119.2832
mvws	M	106.1533
jul_day		103.8612
lat		91.80863
lon		83.98509
mdvt	M	79.11046
mrht	M	73.19908
local_time		59.31836
sw_mean	S	56.7057
DAV	S	55.58434
SHDL_wv	S	44.41981
SHDL_IR1	S	39.37296
ir1_mean	S	38.96147
wv_mean	S	35.97598

표 B. 융합 방안 1에 따라 모든 위성인자와 수치모델인자를 입력자료로 이용하여 하나의 모델로 개발한 변수 중요도

11. 검증과 평가가 정량적으로 이뤄졌는지?

<RF Track based>		Classified		
		Dev	Cont	Non-dev
Original	Dev	7	0	0
	Non-dev	6	0	24

표 C. 위성-수치모델 융합 모델의 검증 결과 (POD = 100 %, FAR = 20%)



다. 성과지표

지표 유형	개념 및 유형분류의 예		비고
투입 지표	○ 연구 개발과정에서 사용된 투입물: 총 연구비 총액 570,000 천원, 투입된 박사급 인력 수 5인		-
과정 지표	○ 과제 수행기간 3년		
산출 지표	양	○ SCI 논문 5편 출판, 비 SCI 논문 1편 출판, S/W 등록 4건	
산출 지표	질	○ 질 높은 논문 성과를 산출하였음.. 과제 성과로 포함된 SCI 논문 5편은 원격탐사, 기상, 기후 분야에서 저명한 인용지수(IF)가 높은 저널 들임 - Remote Sensing of Environment 1편 (IF: 6.265) - Monthly Weather Review 2편 (IF: 3.043) - Journal of Climate 1편 (IF: 4.850) - Journal of Geophysical Research 1편 (IF: 3.450) ○ 전문가의 정성적 평가결과 (미국 Colorado State University, Michale Bell 교수, 2016년 2월 자문) - “태풍발생탐지를 위해서 기계학습 분야와 위성관측을 이용하는 A급 기술”	질적 성과지표
결과 지표	○ 연구 개발 전 객관적인 태풍 발생 탐지 기술 국내에 없었음 -> 2차년도 개발된 위성-전구수치모델 자료 융합 태풍 발생 탐지 기술의		

	<p>최신결과에 따르면 Hit rate 100 %, False alarm rate 23 %</p> <ul style="list-style-type: none"> ○ 주관적 기술에 바탕으로 정성적 분석, 타 현업 기관에 의존적 이었던 태풍 현업 업무를 개선시킬 수 있는 객관적기술, 정량적 분석 방법을 개발. 독자적이고 선도적인 기술 개발 ○ 위성을 활용한 태풍 분석 및 탐지 인력 국내 인프라 구축 및 국제 협력 연구 추진
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

2. 관련 분야 기여도

가. 기술적 측면

- 본 연구에서는 기존 일본과 미국의 방법을 기반을 두지만, 새로운 접근법으로 독창적으로 기술을 개발함. 과학적 분석과 원리 기반을 둔 기술을 개발, 수준 높은 논문들을 출판하여 우리나라의 태풍 연구의 수준을 높일 것으로 기대

- 태풍 발달 시에(TD 단계) 강한 태풍 TY혹은 TS로 발달할 가능성에 대한 정보를 제공, 우리나라에 영향을 줄 수 있는 강한 태풍에 대한 대비를 신속하게 할 수 있음. 수치모델에서 이런 태풍 발생과 강화가 모의에 대한 신뢰도가 높지 않음. 예보자가 태풍 정보를 주는데 중요한 정보가 될 것임

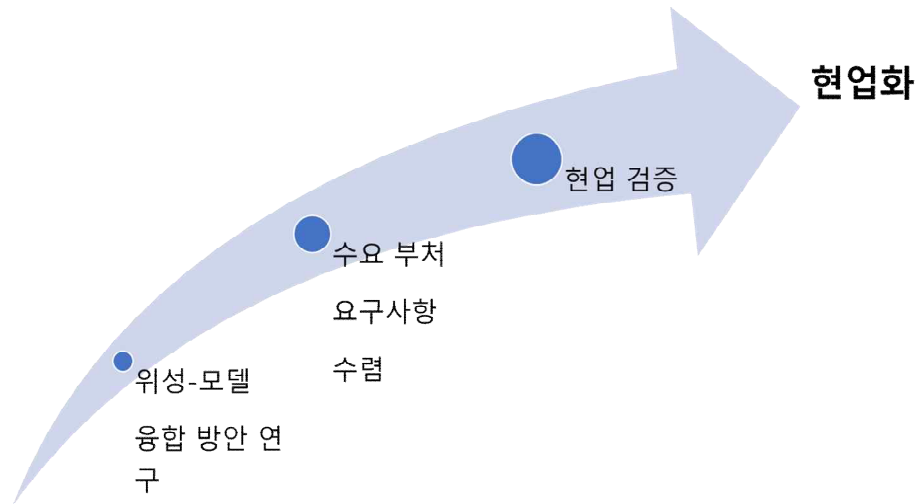
- 다른 나라에 태풍센터에 의존하지 않고, 독자적인 태풍발생 탐지를 할 수 있는 기술을 개발함

성과 목표	기술적 기여
위성관측 자료를 이용한 태풍발생 탐지/예측 인자 개발	현존하는 Dvorak 방법은 적외 채널을 이용한 구름 패턴의 정량적 분석 기술이라면, 개발된 기술은 위성 마이크로라 채널을 이용한 하층 순환 패턴과 강도를 정량화 하는 기술을 개발. 태풍 시스템 자체를 보는 가장 직접적인 방법이며 진취적인 성과라 생각됨.
위성 인자를 이용한 알고리즘 원형 개발	현재 현업에서 발생 단계의 Dvorak 분석은 주관적이므로, 분석자에 따라 결과가 다름. 본 연구를 통해서 객관적인 위성 인자를 이용한 알고리즘 원형이 개발되어, 성능의 차이가 없고 숙련자가 아니라도 이용가능 함. 또한 지금까지 이 분야에서는 선형분별법, 선형회귀법 등을 사용하여 태풍발달 여부나 가능성 정보를 생산하였는데, 개발된 기술은 인공지능 개념의 첨단 기술을 이용하여 알고리즘을 원형을 개발.
전구 수치 모델을 이용한 열대섭동 탐지	대부분의 태풍 현업기관에서는 주관적인 방법으로 열대섭동을 탐지하며, 태풍으로 발생할 가능성이 있는 Invest 정보를 준 실시간 생산함. 개발된 열대섭동 탐지 기술은 객관적이고 자동화된 기술로, 주관적인 분석이 필요하지 않음. 태풍 현업 인력이 충분하지 않은 우리나라에 현실이 적합한 기술이라 평가됨.

나. 경제적·산업적 측면

- 우리나라 기상청의 태풍 모니터링 능력의 향상으로, 신속하게 태풍 재해 경보시스템을 구축 경제적, 산업적 피해를 최소화
- 위성을 이용한 태풍 분석 인력 개발. 앞으로 위성자료를 이용하여 태풍을 분석하는 전문 인력에 대한 수요가 커질 것으로 예상됨. 본 과제를 통해서 대학원생들이 연구를 수행함으로써, 태풍과 위성자료를 잘 아는 전문 인력이 키워질 것으로 생각됨
- 본 과제에서 태풍발생과 연관된 정량적인 인자나 태풍발생탐지/예측 알고리즘이 충분히 검증되고 그 과학적 우수성이 논문, 특허 등으로 검증된다면, 국외로 수출이 가능함

제 5장. 연구개발성과의 활용계획



연구 개발 기술

그림 42. 본 연구 사업에서 수행한 연구 개발 기술에서 현업화 기술로 가기 위한 여러 단계들

3장의 1, 2, 3 절에 서술한대로, 본 연구과제에서 과학적 방법에 기반한 연구 기술을 개발하였고, 개발된 기술에 대한 타당성 및 수준을 국제 학계의 리뷰 후 논문 출판으로 검증하였다. 그러나, 본 과제에서 연구 목적으로 개발된 기술이 현업화되기 위해서는, 현업 수요기관에 맞는 현업화 기술 개발이 필요하다.

그림 42은 본 연구 사업에서 수행한 연구 개발 기술에서 현업화 기술로 가기 위한 여러 단계들 도식화 하였다. 먼저 앞 절에서 기술한 각각 위성, 수치 기반으로 개발된 기술들이 어떻게 융합될지에 대한 방안을 연구하였다. 또한, 개발된 기술의 연구 결과와 현업화 방향성에 대한 현업 수요처의 의견을 조사하여 수렴 반영 하였다. 또한, 앞 절에서 서술한 여러 태풍발생탐지 기술들은 대부분의 경우 국제 학계 기준의 태풍 발생 탐지인 열대저압부 탐지를 기준으로 개발되었다. 현업과 비슷한 상태에서 검증을 실시하기 위해서 국내 태풍현업에서 사용하는 기준으로 조정하여 현업검증을 실시하였다.

본 절에서는 전구수치모델 시스템과 위성 알고리즘의 효율적 결합 방법에 대한 조사를 서술한다. 융합 방법에 대해서 두 가지 안이 제시되었다.

(방법 1) 위성에서 추출한 태풍발생인자들과 전구모델에서 추출한 인자들을 모두

입력자료로 하여 하나의 태풍발생탐지/모델을 구성하는 안(그림 42)

(방법 2) 위성에서 개발된 인자들과 전구모델분석장의 대규모 환경인자 등 중간 인자값 모니터링 기능 추가하며, 각각의 모델의 결과를 산출함. 여러 가이드선의 결정을 참고하여 현업예보자가 태풍 발생 탐지/예측의 최종결과물을 결정하는 안(그림 44)

방법 1에 따라 COMS 위성 인자와 NCEP GFS 인자들을 입력자료로 이용하여 하나의 기계학습(RF) 모델을 개발하였다. 표 17은 태풍발생탐지/예측 모델의 변수 중요도를 나타낸다. 이때 국내 기상현업을 위해 기상청 기준 풍속 17 m/s인 태풍을 탐지하는 것으로 목표로 하였다. 흥미로운 것은 모델과 위성에서 추출한 인자를 동시에 입력한 방법 1에서는 전구모델 인자의 중요도가 COMS 위성 인자보다 훨씬 크게 나왔다. 이는 단일한 탐지 알고리즘을 적용할 경우 태풍발생탐지 결정 과정에서 대규모 환경 인자가 중요한 역할을 한다는 것이다. 현업 예보관들과 단일 알고리즘이 대규모 과정에 크게 의존하는 이유에 대해 논의하였다. 하나의 가정은 모델분석장에는 이미 위성관측 결과가 어느 정도 반영되어 있기 때문에, 모델 분석장이 중요한 역할을 했을 가능성이 있다.

예보자들은 기계학습에 의한 객관적인 최종 결과를 제시하는 것 보다, 현재 존재하고 있는 태풍발생에 대한 물리적 이해를 돕는 것이 중요하다는 의견을 제시하였다. 따라서, 현업화 기술에서는 그림 53과 같이 위성에서 개발된 인자들과 전구모델분석장 대규모 환경인자 등 중간인자값 모니터링 기능을 추가하였다. 또한, 각각의 별도 알고리즘에서 나오는 태풍발생탐지/예측 결과는 현재의 태풍 발생 상태에서 대규모 환경, 시스템 순환, 대류 강도 및 패턴 등 어떤 요소가 발생의 호조건인지에 대한 정보를 제공할 수 있다. 예를 들어, 대류 강도나 패턴을 고려한 정지궤도 위성 적외채널을 이용한 발생탐지는 “발생”으로 산출되었으나, 대규모장을 고려한 모델기반 발생탐지 결과는 “비발생”으로 산출되었을 때를 가정할 수 있다. 이런 경우, 현재 태풍내의 시스템의 발생에 가까우나, 주변환경이 호조건이 아니므로 “비발생”될 가능성도 있다고 해석할 수 있다. 따라서, 본 연구에서 제공하는 현업화 기술은 그림 44과 같은 형태로 제공될 것이다.

표 16. 융합 방안 1에 따라 모든 위성인자와 수치모델인자를 입력자료로 이용하여 하나의 모델로 개발한 변수 중요도

변수명	모델(M)/위성(S)	Mean Decrease Accuracy
mvot	M	119.2832
mvws	M	106.1533
jul_day		103.8612
lat		91.80863
lon		83.98509
mdvt	M	79.11046
mrht	M	73.19908
local_time		59.31836
sw_mean	S	56.7057
DAV	S	55.58434
SHDI_wv	S	44.41981
SHDI_IR1	S	39.37296
ir1_mean	S	38.96147
wv_mean	S	35.97598

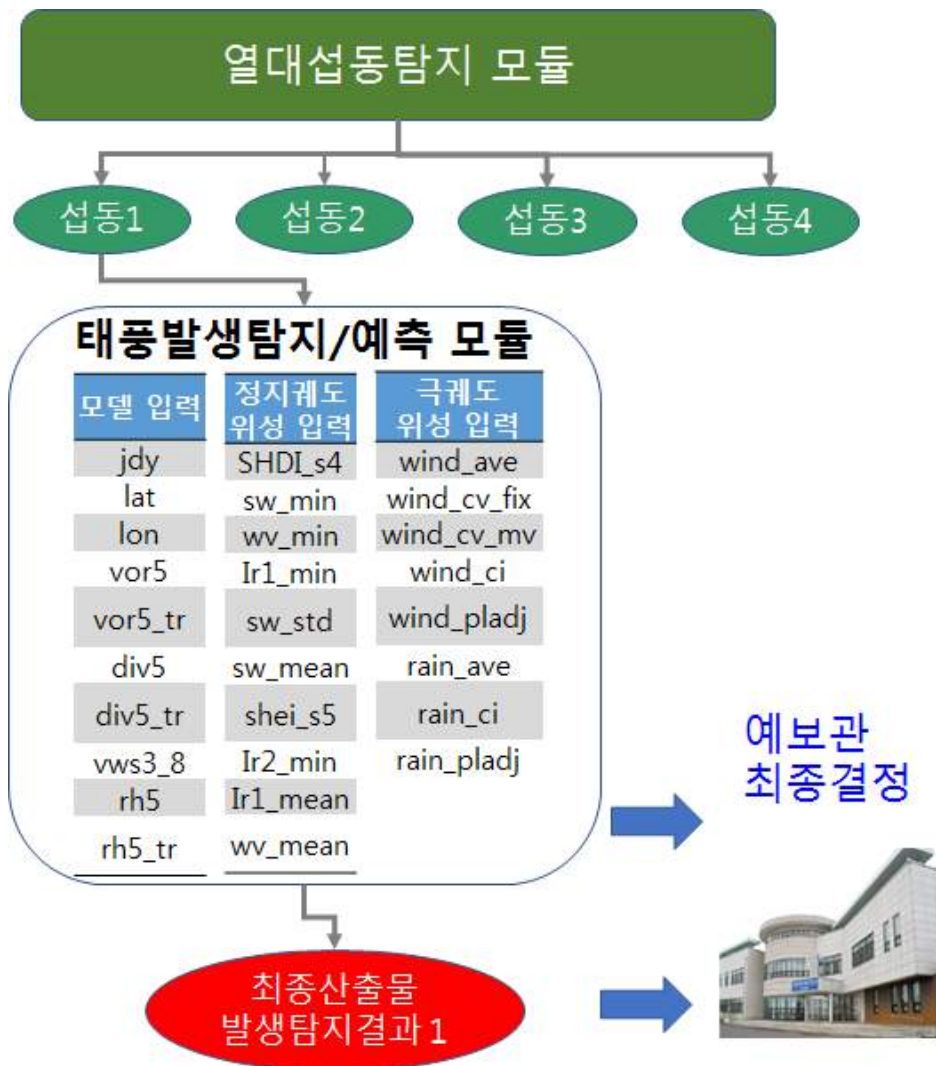


그림 43. 위성-수치 모델 융합 방안 1: 위성에서 개발된 인자와 전구모델에서 개발된 대규모 인자를 모두 고려하여 하나의 태풍발생탐지/예측 현업화 기술로 개발

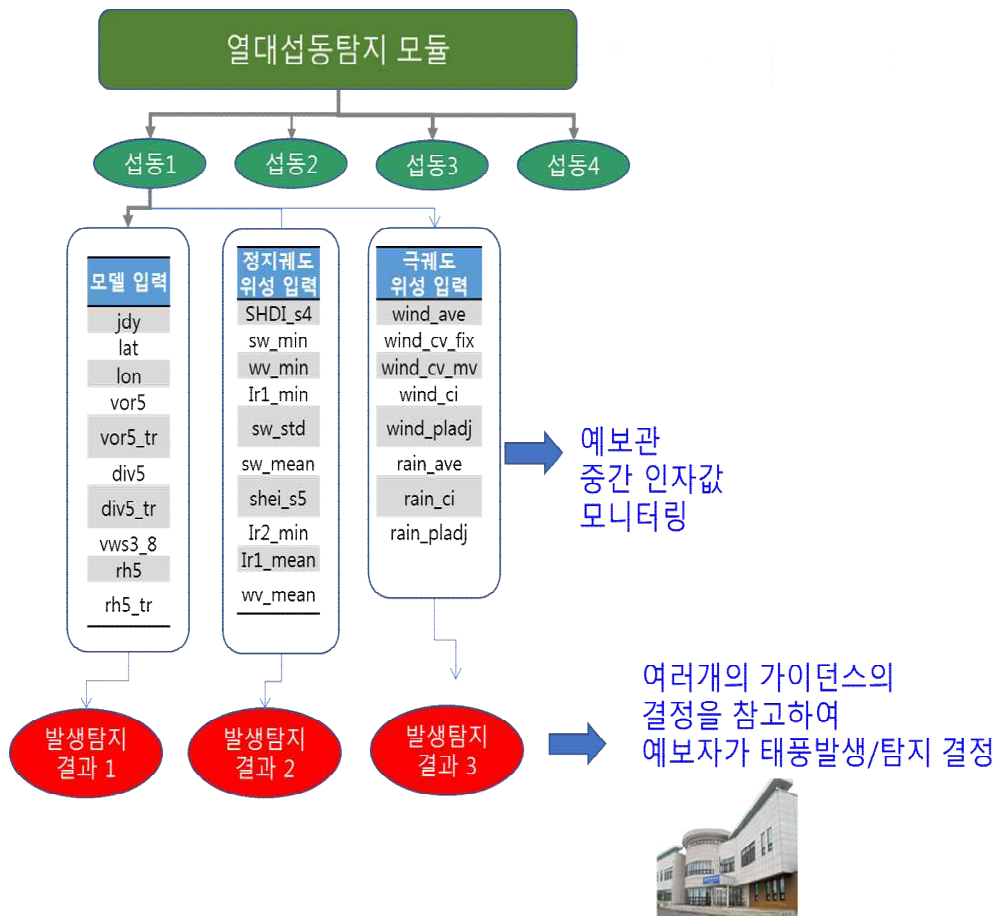


그림 44. 위성-수치 모델 융합 방안 2: 위성에서 개발된 인자들과 전구 모델 분석장의 대규모 환경인자 등 중간 인자값 모니터링 기능 추가하며, 각각의 모델의 결과를 산출함. 여러 개의 알고리즘 가이드언스의 결정을 참고하여 현업 예보자가 태풍 발생 탐지/예측의 최종결과물을 결정하게 함

제 6장. 연구 과정에서 수집한 해외 과학기술 정보

제 7장. 연구개발성과의 보안등급

제 8장. 국가과학기술종합정보시스템에 등록된 연구시설·장비 현황

해당없음

구입 기관	연구시설/ 연구장비명	규격 (모델명)	수량	구입 연월일	구입 가격 (천원)	구입처 (전화번호)	비고 (설치 장소)	NTIS장비 등록 번호

제 9장. 연구개발과제 수행에 따른 연구실 등의 안전 조치 이행 실적

○ 기술적 위험 요소 분석

컴퓨터를 사용하는 연구실로서 전기 안전관련 위험 요소 존재

○ 안전 관리 대책

- 연구실 안전점검 정기적 실시: 연구실의 기능 및 안전을 유지 관리하기 위하여 안전점검지침에 따라 연구실에 관한 안전점검을 정기적으로 실시함.
- 연구 내용 및 결과물 안전 확보 : 정기적으로 인원 및 시설 보안 항목, 문서보안 항목 그리고 정보보안 항목의 보안점검을 실시하여 연구 내용 및 결과물의 안전을 확보하고 있음. 참여 연구원들을 대상으로 연구 결과의 안전한 관리를 위한 안전교육 실시
- 연구실 안전 확보 계획 : 참여 연구원들이 안전관련 각종 법규, 규정 및 지침을 준수하도록 하며, 요구되는 안전교육 및 훈련 실시, 연구실의 잠재되어 있는 위험성 발견 및 위험물질과 각종 실험장비 등 사용에 따른 안전수칙이 잘 이행될 수 있도록 지도점검 및 교육

제 10장. 연구개발과제의 대표적 연구 실적

번호	구분 (논문/특허/기타)	논문명/특허명/ 기타	소속 기관명	역할	논문 게재지/ 특허 등록 국가	영향력 지수	논문 게재일 /특허 등록일	사사 여부 (단독 사사 또는 중복 사사)	특기 사항 (SCI 여부/인용 횟수 등)
	논문	Detection of tropical cyclone genesis via quantitative satellite ocean-surface wind pattern and intensity analyses using decision trees	울산과학기술원	주저자, 공동저자	Remote Sensing of Environment	6.265	2016.09.15	단독	SCI/ 8회
	논문	열대 해수면 온도 분포와 북서태평양 태풍의 계절적 활동 시작일변동 사이의 관련성	한국해양대학교	주저자, 공동저자	한국기후변화학회지	KCI IF : 0.9 (자기인용제외 : 0.73)	2017.03.15	단독	-
	논문	Comments on “Multiscale Structure and Evolution of Hurricane Earl (2010) during Rapid Intensification”	University of Colorado	제 2저자	Monthly Weather Review	3.043	2017.04.04	단독	SCI/ 1회
	논문	Land-Based Convection Effects on Formation of Tropical Cyclone Mekkhala (2008)	울산과학기술원	주저자, 공동저자	Monthly Weather Review	3.043	2017.03.14	단독	SCI

논문	Multiday evolution of convective bursts during western North Pacific tropical cyclone development and nondevelopment using geostationary satellite measurements	서울대학교	공동저자	Journal of geophysical research	3.318	2017.02.14	중복	SCI/1회
논문	Variation of the Tropical Cyclone Season Start in the Western North Pacific	한국해양대학교	주저자, 공동저자	Journal of Climate	4.161	2017.04.12	단독	SCI/4회
S/W	머신러닝 기반의 WindSat 위성자료를 이용한 태풍 발달 탐지 모델	한국해양과학기술원		대한민국		2018.04.06	단독	-
S/W	머신러닝 기반의 히마와리 위성자료를 이용한 태풍 발달 탐지 모델	한국해양과학기술원		대한민국		2018.04.06	단독	-
S/W	천리안기상위성(COMS MI) 기반 태풍전조탐지 알고리즘	한국해양과학기술원		대한민국		2018.04.06	단독	-
S/W	모델 분석장 기반 랜덤포레스트 태풍 발생 예측 및 평가 프로그램	한국해양과학기술원		대한민국		2018.05.02	단독	-

제 11장. 기타 사항

제 12장. 참고 문헌

- Chang, M., M.-S. Park, C.-H. Ho, M.-H. Ahn (2015). Comparison of deep convective burst between developing and nondeveloping tropical disturbances from infrared and water vapor channels of geostationary satellite, will be submitted to Monthly Weather Review
- Dvorak, V. F. (1984). Tropical cyclone intensity analysis using satellite data (Vol. 11). US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service.
- Gaiser, P. W., St Germain, K. M., Twarog, E. M., Poe, G. A., Purdy, W., Richardson, D., ... & Cleveland, J. (2004). The WindSat spaceborne polarimetric microwave radiometer: Sensor description and early orbit performance. *IEEE Transactions on Geoscience and Remote Sensing*, 42(11), 2347–2361.
- Hennon, C. C., Helms, C. N., Knapp, K. R., & Bowen, A. R. (2011). An objective algorithm for detecting and tracking tropical cloud clusters: Implications for tropical cyclogenesis prediction. *Journal of Atmospheric and Oceanic Technology*, 28(8), 1007–1018.
- Hennon, C. C., Papin, P. P., Zarzar, C. M., Michael, J. R., Caudill, J. A., Douglas, C. R., ... & Scales, M. A. (2013). Tropical cloud cluster climatology, variability, and genesis productivity. *Journal of Climate*, 26(10), 3046–3066.
- ITWC-VII, Seventh International Workshop on Tropical Cyclones, WMO/WWRP, La Réunion, FR, Nov (2010).
- Kucas, M. E., & Darlow, J. W. (2012). A subjective method for assessing tropical cyclogenesis at the Joint Typhoon Warning Center. *Tropical Cyclone Research and Review*, 1, 325–330.
- Lee, M. I., Schubert, S. D., & Kim, D. (2011). Representation of tropical storms in the northwestern pacific by the Modern-Era Retrospective analysis for research and applications. *Asia-Pacific Journal of Atmospheric Sciences*, 47(3), 245–253.
- Lee, S., Han, H., Im, J., Jang, E., & Lee, M. I. (2017). Detection of deterministic and probabilistic convection initiation using Himawari-8 Advanced Himawari Imager data. *Atmospheric Measurement Techniques*, 10(5), 1859.
- Montgomery, M. T., Nicholls, M. E., Cram, T. A., & Saunders, A. B. (2006). A vortical hot tower route to tropical cyclogenesis. *Journal of the atmospheric sciences*, 63(1), 355–386.
-

-
- Olander, T. L., & Velden, C. S. (2007). The advanced Dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. *Weather and Forecasting*, 22(2), 287–298.
- Olander, T. L., & Velden, C. S. (2009). Tropical cyclone convection and intensity analysis using differenced infrared and water vapor imagery. *Weather and Forecasting*, 24(6), 1558–1572.
- Park, M. S., Choi, Y. S., Ho, C. H., Sui, C. H., Park, S. K., & Ahn, M. H. (2007). Regional cloud characteristics over the tropical northwestern Pacific as revealed by Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar and TRMM Microwave Imager. *Journal of Geophysical Research: Atmospheres*, 112(D5).
- Park, M. S., Penny, A. B., Elsberry, R. L., Billings, B. J., & Doyle, J. D. (2013). Latent heating and cooling rates in developing and nondeveloping tropical disturbances during TCS-08: Radar-equivalent retrievals from mesoscale numerical models and ELDORA. *Journal of the Atmospheric Sciences*, 70(1), 37–55.
- Park, M.-S., (2014). Mesoscale convective systems at tropical cyclone formation time: Quantitative analyses of satellite microwave polarized corrected temperature images 31th Conference on Hurricanes and Tropical Meteorology, San Diego, CA, March 2014.
- Park, M. S., Kim, M., Lee, M. I., Im, J., & Park, S. (2016). Detection of tropical cyclone genesis via quantitative satellite ocean surface wind pattern and intensity analyses using decision trees. *Remote sensing of environment*, 183, 205–214.
- Piñeros, M. F., Ritchie, E. A., & Tyo, J. S. (2010). Detecting tropical cyclone genesis from remotely sensed infrared image data. *IEEE Geoscience and Remote sensing letters*, 7(4), 826–830.
- Piñeros, M. F., Ritchie, E. A., & Tyo, J. S. (2011). Estimating tropical cyclone intensity from infrared image data. *Weather and Forecasting*, 26(5), 690–698
- Schumacher, A. B., DeMaria, M., & Knaff, J. A. (2009). Objective estimation of the 24-h probability of tropical cyclone formation. *Weather and Forecasting*, 24(2), 456–471.
- Tsuchiya, A., Mikawa, T. and Kikuchi, A. (2001): Method of distinguishing between early stage cloud systems that develop into tropical storms and ones that do not, *Geophysical Magazine Series 2*, Vol.4, Nos.1–4, 49–59.
-

Wood, K. M., & Ritchie, E. A. (2015). A definition for rapid weakening of North Atlantic and eastern North Pacific tropical cyclones. *Geophysical Research Letters*, 42(22).

김민상, 박명숙, & 이명인. (2014). Windsat 위성 해상풍 자료를 이용한 태풍발생탐지 및 예측인자 개발. 한국기상학회 학술대회 논문집, 777-778.

