

**예측성 향상을 위한 다중모델기반 기후예측기술 개발**  
**Development of multi-model-based climate  
prediction technique for the predictability  
improvement**

**손수진, 김원무, 박경원, 정유란, 정임국**  
**기후사업본부 예측기술과**

**2021.12.**

**APEC기후센터**

## 평문 초록

유사한 기후 특성을 기반으로 하는 지역 기후 전망을 위해서는 지역 맞춤형 기후 예측 기술이 개발 및 활용되어야 합니다. 따라서 이 연구에서는 계절 및 계절 내 시간 규모를 예측 목표로 동아시아와 한반도 지역을 중점 연구지역으로 하여, 수집 가능한 여러 역학 기후 모델을 바탕으로 모델 간 앙상블을 활용하는 기후 예측 기술을 개발하였습니다.

2017년부터 APCC 계절 기후 예측을 지원하던 우리나라 맞춤형 전문가 계절 예측 시스템 ESPreSSO (Expert Seasonal Prediction System for operational Seasonal Outlook)이 한 단계 개선하였습니다. 우리 센터의 최신 다중모델앙상블 구성에 기반한 전문가 가이드와 인공지능의 협력을 통해 예측자료 생산의 민첩성과 예측성을 확보하였고, 평문 예보문과 기후 모식도 등을 통해 예측정보를 쉽게 이해할 수 있도록 접근성이 향상되었습니다. 동아시아 계절 기후 예측을 위해서는 역학 모델의 예측성능을 바탕으로 우수 모델을 선별하여 모델 간 앙상블을 수행하는 방법이 고려되었습니다. 모델 선정의 기준은 기후인자 (엘니뇨-남방진동의 복잡성과 열대 서태양 대류 등)와 그 영향에 대한 모델의 재현 능력을 바탕으로 합니다. 선별 모델 앙상블이 가용 모델 앙상블에 비해 높은 예측성능과 낮은 불확실성을 제시하였습니다. 이러한 모델 선별 앙상블에 적합한 확률 앙상블 방법을 도출 및 적용하여, 계절 예측을 위한 시스템을 구축하고 통합 기후 전망을 위한 표출 시스템을 개발하였습니다. 이 연구를 통해서 개발된 새로운 버전의 ESPreSSO<sup>v3.2</sup>, 동아시아 맞춤형 기후 예측 및 가시화 시스템의 전체 과정은 2021년 동안 시운영되어 그 효용성을 확인, 이에 현업화를 위한 과정이 추진되었습니다.

한편, 정규화 베이스 후처리를 통한 지점 대 지점 간의 선형 후보정의 한계를 극복하고자, 최신 인공지능 신경망 기술을 도입하였습니다. 우리 센터의 다중모델앙상블 계절 기후 예측 자료로부터 의미있는 정보를 추출하여, 동아시아에 맞는 정확한 예측 정보를 제공하기 위해 이미 지 처리에 사용되는 최신 인공신경망 구조인 유-넷을 도입하고, 이를 기후 정보에 맞게 대폭 수정하여 예측 성능이 향상되면서도 안정적인 후보정이 이루어질 수 있도록 하였습니다. 특히 이 과정에서 공간적인 구조를 가지면서 계절별 특성을 가지는 기후 정보로부터 유의미한 정보를 추출하기 위해 다양한 기법을 개발/적용하였습니다. 이러한 맞춤형 인공신경망 후처리 모형을 통해 다중모델앙상블을 후처리한 결과, 동아시아 강수 예측의 경우 기존 다중모델앙상블 확률예측에 비해 예측 적중률이 상승하였고, 계절별/연도별로 안정적인 예측 성능을 보였습니다.

계절내 예측자료에 대한 수요증대와 높은 정확도의 3-4주 예측을 제공하기 위하여, 신경망 기법을 활용하여 계절내 예측의 오차를 줄이는 기법을 개발하였습니다. 신경망 기법은 동아시아 지역의 온도, 강수 등의 다양한 기후변수를 바탕으로 최대온도값을 잘 학습하도록 설계되었습니다. 신경망 기법은 3-4주에서 오차가 줄어든 예측을 생산하였습니다. 한편 연산 속도 개선 및 신경망 모델의 분산 감소를 위한 심층학습 앙상블 학습 기법 개발하였습니다. 심층학습 앙상블은 합성곱 신경망을 통해서 공간 정보를 유지한 채 학습시키고, 이 정보를 기계학습의 새로운 훈련셋으로 전달되도록 구축하였습니다. 3-4주에서 심층학습 앙상블은 단일 모델보다 계절내 예측성을 개선하였습니다. 아울러 샘플데이터의 증강 혹은 변환, 심층학습 알고리즘 개선은 추후 예측성 향상에 기여할 것으로 기대합니다.

# Executive Summary

It is quite essential for valuable and reliable climate outlook at a regional scale based on similar climate variability and characteristics to develop regional-specific climate prediction techniques. Therefore, this research project is set to be carried out for three years from 2019 to 2021 to develop regional-specific leading climate prediction technology. In the first year, regional-specific seasonal multimodel ensemble (MME) climate prediction technology development systems were established. Based on this, regionally customized seasonal and sub-seasonal MME climate prediction technology were developed in 2020. Finally, it has been constructed and operationalized to produce the multi-model-based seasonal climate prediction customized at Korean peninsula and East Asian region as well as it has been further developed to advance the multi-model-based sub-seasonal climate prediction.

First, the Expert Seasonal Prediction System for operational Seasonal Outlook, ESPreSSO, is upgraded to be more agile, accurate, and accessible; and internally operationalized. The updated expert guide is now based on the new APCC multi-model ensemble configuration, which is fed into the artificial intelligence predictor selection system to enhance agility and accuracy. The revised natural language outlook report as well as the conceptual diagram greatly improves the accessibility to the climate outlook data. The new ESPreSSOv3.2 has already passed the internal operationalization screening and waiting to be officially internally operational.

Second, the physical phenomena-based ensemble model reconstruction and probabilistic prediction were performed for the development of multi-model-based seasonal prediction technology for the East Asian region. In the first year, in order to prepare a system for the development of climate prediction technology for regionally tailored seasonal MMEs, the predictive performance of the complexity of El Niño–Southern Oscillation (ENSO) and its climate impacts were diagnosed, and this was used as a factor to improve local climate prediction. In the 2020 year, the variation of precipitation in the western Pacific was discovered as an additional climate factor and applied together with ENSO to develop and evaluate probabilistic MME prediction technology. The probabilistic MME prediction constructed in this way showed more improved predictability in terms of category verification compared to the existing MME based on all available models. In the final year, it has improved through the optimal sampling among all available models and appropriate probabilistic MME methods of calibration and empirical ranking and pooling for the probability density estimation. Finally, the step-wise prediction and visualization systems based on the results has been constructed and experimented for the operationalization.

Third, a SOTA artificial intelligence (AI) techniques are introduced to overcome the limitations of the grid-to-grid post-processing through regularized linear Bayesian approach.

The SOTA U-Net architecture is heavily modified to optimally process climate data, which as spatial characteristics as well as seasonal dependency; it includes the use of variable-specific kernels for convolution, multi-level bypasses, channel and spatial attentions, and dynamic conditional dropouts. The customized AI neural network enhanced the predictability of the East Asian climate outlook as much as 9.5% for precipitation, compared to the current probabilistic multi-model ensemble prediction results.

On the other hand, we developed post correction technique suited to Korean peninsula, by LSTM model based on seasonal prediction data of the Glossea5 and ECMWF models, in the 2019 project. We provided accuracy of highly predictable information of the performance of sub-seasonal data by ConvLSTM method in the 2020 project. We developed ConvLSTM deep learning technology for subseasonal climate model by sensitivity analysis to improve S2S prediction performance on the East Asian region. ConvLSTM model was optimized the parameters and main variables by machine learning and composition of input data. The results were verified with reanalysis data. As a result of the analysis, the accuracy of PCC was low but we confirmed the performance improvement in the RMSE of statistical analysis.

Finally, this study was conducted to develop a deep learning ensemble learning technique to improve the computation speed and reduce the variance of the neural network model as the range of training data input to the deep learning model was expanded to the East Asia region. For the MME-based S2S prediction climate input data for building a learning model, five individual climate models(ECMWF, NCEP, KMA, UKMO, CMA) were combined from the reference forecast frequency(4 times a month on the 1st, 9th, 17th, and 25th) and time range(30 days) during the period of hindcast (1994-2014). As common input variables, climatic variables such as surface average air-temperature, maximum and minimum air-temperature, precipitation, sea level pressure, winds in the east, west, north and south directions at 850hPa pressure surface, and geographic altitude were used. As observation, it was constructed with ERA-Interim data, which are reanalyzed data. The training model based ensemble deep learning(after here, hybrid training model), first, it was trained from a convolutional neural network while maintaining spatial information, and automatically learned this information was constructed as a process of feeding into a new training set of machine learning, for example, random forest. As a result of weekly prediction accuracy analysis of hybrid training model, the improvement of predictability of the hybrid model was still lower than before applying the model to the MME-based S2S, although it slightly improved the uncorrected effect of week 1 that appeared in single machine learning, that is random forest, or single convolutional neural network learning. However, in week 3-4, the predictability of MME based S2S was improved by frequency, indicating that the hybrid training model is contributing to the improvement of S2S prediction than the single model. In next study, improvement of the deep learning ensemble model through continuous deep learning algorithm improvement, such as augmentation or transformation of sample data, optimization of hyperparameter and architecture will contribute to the improvement of MME-based S2S prediction.

## 국문 요약

유사한 기후 특성을 기반으로 하는 지역 규모에서 기후 전망을 위해서는 지역 특화형 선도 기후예측 기술 개발이 이루어져야 한다. 따라서 이 연구과제는 “지역특화형 선도 기후예측 기술 개발”을 목표로 2019년부터 2021년까지 3년 동안 수행되었던 사업으로, 1차년도에는 지역 맞춤형 계절 다중모델앙상블(multimodel ensemble, MME) 및 계절내(sub-seasonal to seasonal, S2S) MME 기후예측 기술 개발 체계가 마련되었다. 2020년에는 이를 기반으로 하여 지역 맞춤형 계절 및 계절내 MME 기후예측 기술을 개발하였다. 이 과제에서는 예측 시간 규모 및 예측 자료의 유용성 등을 고려하여, 예측성 향상을 위한 다중모델 기반 계절 예측 기술개발과 계절내 예측 기발 기술 개발로 구분하여 연구를 진행하였다. 따라서 2021년에는 다중모델 기반 계절 예측 시스템을 구축하여 현업화를 추진하고, 계절내 예측 기술을 고도화하였다. 계절 예측 기술 개발을 위해서 과제는 크게 한반도 맞춤형 지역 예보 생산 및 개선, 물리과정 기반 모델 선별을 통한 동아시아 확률 예측 개선, 동아시아 계절 확률 예측 생산 및 표출 시스템 개발, 인공지능 기반 동아시아 계절 확률 개선 연구를 수행하였다. 또한 다중모델 계절내 예측 기반 기술 개발을 위해서는 인공지능 기반 아키텍처를 활용한 보정 기술 및 시·공간 자료에 적합한 통합 인공지능 기반 보정 기술을 개발하였다.

우리나라 맞춤형 전문가 계절 예측 시스템 ESPreSSO (Expert Seasonal Prediction System for operational Seasonal Outlook)의 민첩성, 예측성, 그리고 접근성을 개선한 ESPreSSOv3.2를 개발하고, 내부 현업화 절차를 진행하였다. 기후 예측 전문가의 경험과 지식이 반영된 전문가 가이드를 우리 센터의 최신 다중모델앙상블 구성에 맞춰 업데이트하고, 전문가 가이드와 인공지능이 협력하여 최종 후보정 예측 자료가 생산되도록 함으로써 민첩성과 예측성을 확보하였다. 또한, 다양한 결과 표출, 향상된 평문 예보문과 기후 예측 현황 모식도 등을 제공하여 예보자가 예측 정보를 쉽게 이해할 수 있도록 접근성을 향상시켰다. 이러한 새로운 버전의 ESPreSSO<sup>v3.2</sup>는 올해 내부 현업화 심의를 통과하여 공식적인 내부 현업 지원을 앞두고 있다.

다중모델 내 샘플링(모델 선별)을 통한 동아시아 지역 맞춤형 계절 확률 예측 연구를 위해서는, 1차 년도에는 지역 맞춤형 계절 MME 기후예측 기술 개발체계를 마련하였다. 우선 엘니뇨-남방진동(El Niño-Southern Oscillation, ENSO)의 복잡성과 그 원격상관의 예측성을 진단하고, 이를 지역 기후예측을 향상시키기 위한 물리 인자로 사용하였다. 2차 년도에는 동아시아 기후 예측을 위해 서태평양 강수 변동을 추가 기후인자로 발굴하여 ENSO와 함께 적용하여 확률론적 MME 예측기술을 개발하고 평가하였다. 이렇게 구성된 확률 MME 예측은 가용한 모든 모델을 바탕으로 한 기존의 MME에 비해서 카테고리 검증 측면에서 예측성이 개선된 결과를 보였다. 따라서 3차년도에는 모델 선별 앙상블에 적합한 확률 요소 및 기법을 선정하여 적용하고 기존의 현업 확률 예측 대비 정확성, 선명도, 정보력에서 그 효용성을 확인하였다. 이러한 예측성 개선 및 추가 정보 제공에 대한 이점을 바탕으로 예측 및 표출 시스템을 구축하고 현업화를 추진하였다.

다음으로 정규화 베이스 후처리를 통한 지점 대 지점 간의 선형 후보정의 한계를 극복하고자, 최신 인공지능 신경망 기술을 도입하였다. 우리 센터의 다중모델앙상블 계절 기후 예측 자

료로부터 의미있는 정보를 추출하여, 동아시아에 맞는 정확한 예측 정보를 제공하기 위해 이미 지 처리에 사용되는 최신 인공신경망 구조인 유-넷을 도입하고, 이를 기후 정보에 맞게 대폭 수정하여 예측 성능이 향상되면서도 안정적인 후보정이 이루어질 수 있도록 하였다. 특히 이 과정에서 공간적인 구조를 가지면서 계절별 특성을 가지는 기후 정보로부터 유의미한 정보를 추출하기 위해 변수(채널)별 합성곱 커널, 채널-공간 어텐션, 동적 조건부 드롭-아웃, 다계층 우회로 등의 다양한 기법을 개발/적용하였다. 이러한 맞춤형 인공신경망 후처리 모형을 통해 다중모델앙상블을 후처리한 결과, 동아시아 강수 예측의 경우 기존 다중모델앙상블 확률예측에 비해 예측 적중률이 약 9.5% 상승하였고, 계절별/연도별로 안정적인 예측 성능을 보이고 있다.

또한 계절내 예측의 오차 저감을 위해 다중모델 기반 계절내 예측 기술이 개발 되었다. 2019년과제에서는 GloSea5와 ECMWF모델 계절내 예측자료 기반으로 LSTM기법에 적용하여 한반도 지역에서 잘맞는 후보정 기법인 딥러닝 기법을 산출하였다. 2020년 과제에서는 한반도 지역의 예측성 높은 정보를 제공하기 위하여 격자기반의 입력자료에 따른 ConvLSTM기법을 적용하여 지점인 한반도 지역의 예측성능을 향상시키는 연구를 수행하였다.

2021년 연구과제에서는 기계학습과 심층학습 복합 활용 알고리즘 개발을 통해 동아시아 계절내(sub-seasonal to seasonal, 이하 S2S) 다중모델앙상블(multi-model ensemble, 이하 MME) 예측 자료의 예측성을 개선하고자 수행되었다. 먼저 다수 기후모델의 S2S 예측 자료를 기반으로 동아시아 및 한반도 영역에 대한 S2S 예측 자료는 과거기후 예측기간(1994-2014년) 동안 기준 예측빈도(1, 9, 17, 25일로 매월 4회) 및 예측 선행시간(30일)을 기준으로부터 5개(ECMWF, NCEP, KMA, UKMO, CMA) 개별 기후모델이 결합되었다. 공통 입력 변수는 지표면 평균기온, 최고기온, 최저기온, 강수, 해면기압, 850hPa 기압면에서의 동서남북 방향의 바람, 지위고도 등의 기후 변수가 이용되었다. 관측값으로는 재분석 자료인 ERA-Interim 자료로 구축되었다. 기계학습 및 심층학습 복합 알고리즘(이하 하이브리드 학습 모델) 개발을 위해서 기계학습에서는 랜덤포레스트를, 심층학습으로는 합성곱 신경망 알고리즘을 채택하였다. 동아시아 지역 계절내 예측의 정확도 향상을 위하여 개별모델과 MME 입력자료에 대한 ConvLSTM 딥러닝 기술개발과 민감도 분석을 실시하였다. 최적화된 파라미터와 머신러닝에서 제공한 주요변수 설정 및 입력자료 구성을 통하여 ConvLSTM 기반 계절내 예측기술을 개발하고, 재분석자료와의 학습결과를 검증하였다.

복합 알고리즘 개발을 위해서는 먼저 합성곱 신경망을 통해서 공간 정보 그대로를 유지한 채 학습시키고, 자동으로 학습된 이 정보를 랜덤포레스트의 새로운 훈련셋으로 feed 되게 하는 과정으로 구축하였다. 구축된 동아시아 MME S2S 예측 자료를 하이브리드 학습 모델로 학습하고 주별 예측 정확도를 분석한 결과, 단일 랜덤포레스트 혹은 단일 합성곱장단기 기억 학습에서 나타났던 week1의 보정 효과 퇴보는 다소 개선되었지만 여전히 보정 전보다 크게 개선되지 않았다. 그러나 week3-4에서 예측 빈도별 개선은 향상되었다. 향후 샘플데이터의 증강 혹은 변환, 하이퍼파라미터 보정 및 아키텍처 보완 등 지속적인 심층학습 알고리즘 개선을 통한 심층학습 복합 모델 개선은 MME 기반 S2S 예측 향상에 기여할 수 있을 것이다. 또한 향후 다양하게 제시되고 있는 고도화된 인공지능 모델을 이용하고, GPU기반의 컴퓨팅의 계산용량이 확충된다면 보다 개선된 예측자료 생산이 가능하며, 이 자료에 기반한 기후감시 및 기후분야 활용에 도움이 될 것이다.

# 목차

평문 초록 .....	i
Executive Summary .....	ii
국문 요약 .....	iv
목차 .....	vi
표 차례 .....	x
그림 차례 .....	xii
1. 서론 .....	1
가. 연구의 배경 .....	1
나. 연구의 목표 .....	2
다. 연구의 내용 및 추진전략 .....	2
라. 연구의 추진체계 .....	3
2. 우리나라 맞춤형 전문가 계절 예측 시스템 .....	6
가. 개요 .....	6
나. 새로운 버전의 ESPreSSOv3.x 개발 .....	6
(1) 민첩성(agility), 예측성(accuracy), 접근성(accessibility)을 향상시킨 ESPreSSOv3 .....	6
다. 민첩성/예측성/접근성이 향상된 ESPreSSOv3의 현업 운영 지원 .....	8
(1) ESPreSSOv3.0 기반의 3 개월 예측 정보 생산 및 제공 (매달) .....	8
(2) 새로운 전문가 가이드가 적용된 ESPreSSOv3.2의 내부 현업화 .....	9
3. 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 개선 및 시스템 구축	10
가. 개요 .....	10
나. 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 .....	10
(1) 앙상블 재구성을 위한 물리현상 .....	10
(2) 물리현상 기반의 앙상블모델 재구성 .....	13
(3) 지역특화 계절예측 기술 개발 .....	15
다. 다양한 기후예측자료의 기후예측성능 검증을 통한 안정성 평가 .....	15
(1) 과거재현 자료의 통계적 안정성 확인 .....	15
(2) 다양한 기후예측자료의 기후예측성능 검증 .....	16
라. 모델 선별을 통한 다중모델기반 기후예측 기술 개발 .....	19

(1) ENSO 유형과 강수예측 성능에 기반 한 앙상블모델 재구성 .....	19
(2) predictability source 및 앙상블 재구성을 바탕으로 한 MME의 한계 파악 .....	21
마. 확률예측 기법 개발 .....	23
바. 동아시아 맞춤형 확률예측시스템 구축 .....	26
사. 동아시아 맞춤형 확률예측을 위한 표출시스템 구축 .....	26
(1) 시스템 및 프로그램의 구성 .....	26
(2) 동아시아 확률예측의 가시화 .....	27
4. 다중모델앙상블의 후처리를 통한 동아시아 예측성 향상 .....	50
가. 개요 .....	50
나. 선형 베이스 접근법에 기반한 다중모델앙상블 격자 자료 후처리 .....	50
(1) 베이스 접근법 .....	50
(2) 정규화를 도입한 베이스 회귀 .....	52
(3) 선형 베이스 후처리 모형을 이용한 보정된 예측 결과 .....	53
다. 맞춤형 인공지능망을 통한 비선형 후처리 .....	55
(1) 기본 아키텍처: 스킵-커넥트 인코더-디코더 유-넷 .....	55
(2) 인코더: 채널별 합성곱 커널, 채널-공간 어텐션 .....	55
(3) 동적 조건부 드롭아웃이 적용된 보틀넥과 고차원 우회로 .....	57
라. 예측성을 향상을 위한 맞춤형 인공지능망의 전이 학습 전략 .....	60
(1) 장기 관측 재분석 자료를 이용한 사전 학습 .....	60
(2) 다중모델앙상블 단순합성 자료와 강수 재분석 자료를 이용한 미세 조정 학습 .....	60
마. 맞춤형 인공지능망 후처리를 통한 동아시아 예측성 향상 .....	61
(1) 맞춤형 인공지능망 후처리 모형을 이용한 예측 결과의 보정 예시 .....	61
(2) 맞춤형 인공지능망 후처리를 통한 예측 성능의 향상과 안정화 .....	61
5. 동아시아 확률예측을 위한 기법 간 상호 비교 .....	64
가. 자료 및 방법 .....	64
나. 검증 결과 .....	64
6. 다중모델 계절내예측 기반기술 개발: 머신러닝 및 딥러닝 복합 활용 기법 개발 .....	68
가. 개요 .....	68
나. 동아시아 지역 계절내 MME 예측 자료 구축 .....	69
다. 머신러닝 및 신경망 모델 결합을 위한 딥러닝 하이브리드 알고리즘 구축 .....	71
(1) 랜덤포레스트 학습 모델 .....	72
(2) 하이브리드 학습 모델 .....	74

(3) 예측성 검증 및 평가 .....	76
라. 동아시아 계절내/MME 예측에 대한 하이브리드 알고리즘 민감도 분석 .....	78
(1) 랜덤포레스트 알고리즘 최적화 .....	78
(2) 합성곱장단기기억 알고리즘 최적화 .....	81
마. 하이브리드 알고리즘 내 동아시아 계절내/MME 예측입력자료의 민감도 분석 .....	81
(1) 입력자료 민감도 분석 .....	82
바. 하이브리드 모델 최적화 .....	84
(1) 개별 기후모델의 LSTM 적용에 따른 예측 성능 평가 .....	84
사. 신경망모델 하이브리드 기반의 동아시아 MME 계절내 예측 기후 성능 평가 기법 구축 및 가시화 .....	87
(1) 평방근제곱근오차 및 편차공간상관성에 따른 예측성 평가 .....	87
(2) 개선 비율 평가 .....	91
(3) 코사인 유사도에 따른 예측성 평가 .....	94
7. 격자기반 계절내 예측 딥러닝 기법 개발 .....	97
가. 개요 .....	97
나. 자료 및 방법 .....	97
(1) 자료 .....	97
(2) 방법 .....	101
다. 결과 .....	109
(1) 개별모델 딥러닝 모델 학습결과 .....	109
(2) 민감도 분석 결과 .....	111
8. 요약 및 결론 .....	116
가. 요약 및 결론 .....	116
(1) 후처리를 통한 우리나라 맞춤형 전문가 계절 예측 시스템의 개발 .....	116
(2) 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 개선 및 시스템 구축 117	
(3) 다중모델앙상블의 후처리를 통한 동아시아 예측성 향상 .....	118
(4) 다중모델 계절내예측 기반기술 개발: 머신러닝 및 딥러닝 복합 활용 기법 개발 .....	118
(5) 격자기반 계절내 예측 딥러닝 기법 개발 .....	119
나. 연구개발결과의 활용 방안 및 기대성과 .....	120
(1) 과학적·기술적 성과 .....	120
(2) 경제적·사회적 성과 .....	120

사사 .....	122
References .....	123

## 표 차례

Table	page
1. Final goal and annual objectives of the research project. ....	3
2. Three versions of hindcast data sets and experimental design. ....	17
3. Several factors for the probability forecast of APCC, NMME, and C3S. ....	24
4. Step-wise probability forecast system developed in the study. ....	26
5. Verification of past forecasts (HSS). ....	43
6. Verification of past forecasts (category matching). ....	44
7. Strategies of Bayesian post-processing with regularization applied in this study. ....	53
8. Information of sub-seasonal to seasonal climate models: used to produce multi-model ensemble sub-seasonal to seasonal prediction data in this study. ....	70
9. Different deep learning models for processing four types of spatial-temporal data: it shows which is learning technique of deep learning adopted in this study (modified from Source: Wang et al. (2019) Deep Learning for Spatio-Temporal Data Mining: A Survey). ....	72
10. Key hyperparameters of random forest training model. ....	73
11. Four scalers information for data transformation in deep learning. ....	82
12. Key hyperparameters of convolutional neural network algorithm in the hybrid model of this study. ....	84
13. Information of four representative tuners for hyperparameter optimization of deep learning algorithms. ....	84
14. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME before applying training models. ....	92
15. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME after applying hybrid training model. ....	92
16. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME after applying random forest (after here, RF in this table) model. ....	92
17. The values of the X-axis and Y-axis of Figure 20, which plot the improvement rates by the hybrid model (i.e. one of post-calibration methods of this study) of the individual climate models and the MME-based subseasonal predicted daily maximum air-temperature: The right of table is the improvement rates of the spatial anomaly pattern correlation coefficient (PCC) calculated by Eq. [1]. The left of table is the improvement rates of the root mean square error (RMSE) calculated by Eq. [2]. A rate with a positive value means ‘improvement by the post-calibration technique’ and is marked in bold. ....	93
18. The values of the X-axis and Y-axis of Figure 20, which plot the improvement rates by the random forest model (i.e. one of post-correction method of this study) of the individual climate models and the MME-based subseasonal predicted daily maximum air-temperature:	

The right of table is the improvement ratio of the spatial anomaly pattern correlation coefficient (PCC) calculated by Eq. [1]. The left of table is the improvement ratio of the root mean square error (RMSE) calculated by Eq. [2]. A rate with a positive value means ‘improvement by the post-calibration technique’ and is marked in bold. ....93

19. Summary of training dataset for ConvLSTM Model. ....101

20. Summary of convolutional layer at ConvLSTM model. ....104

## 그림 차례

Figure	page
1. Research Structure. ....	3
2 The overview of ESPreSSO – Expert Seasonal Prediction System for operational Seasonal Outlook. ....	6
3 An example integrated climate outlook report of ESPreSSOv3. ....	8
4 Examples of operational reports of ESPreSSOv3.0 in 2021. ....	9
5 Operationalization (deliberation process) of ESPreSSOv3.2. ....	9
6. (a) First and (b) second EOF patterns of a monthly rolling 3-month mean SSTAs computed for 25° S–25° N and 140° E–80° W during 1983–2005, based on OISST. The fractional variance explained by each mode and season are provided in the upper right of each panel. The horizontal and vertical black dashed lines denote the equator and International Date Line. ....	12
7. Time series of the combined PCs of (a) the first mode (red solid line for left y-axis) and (b) the second mode (blue dashed line for right y-axis) associated with Figure 2 with all months during January 1982 to December 2005. ....	13
8. Climatological seasonal mean of East Asian and tropical west Pacific rainfall. ....	13
9. Relative signs and distance magnitudes (see the relevant coordinate and scale at bottom, respectively) on a 2-dimensional plane depicting regression coefficients of the seasonal mean anomalous precipitation onto the first two PCs for (a-l) JFM to DJF. ....	14
10. Regression of seasonal mean anomalous East Asian rainfall on tropical Western Pacific rainfall. ....	15
11. Time series (left panels) and their statistical distribution (right panels) of interannual variation of ACC skills averaged during corresponding AMJ–OND and FMA–MJJ seasons to (a) Australasian and East Asian precipitation predictions, based on training (1982–1993) and test (1994–2005) periods. Blue and red circles denote the raw MME based on all-inclusive models and new MME based on best performing models, respectively. Augmented Dickey–Fuller (ADF) tests are applied to test the stationarity of whole time series, and the relevant ADF statistics are provided in parentheses following the labelled targeted data sets of upper left of each panel. The statistical distribution of these time series for the raw and new MMEs are illustrated by box plots and mean values (circles), based on training and test periods. ....	16
12. Time-averaged ACC skills of the two MME mean (upper and middle panels) Australasian (left panels) and East Asian precipitation predictions (right panels) and their difference fraction (%). ....	18
13. Probability distribution functions of regionally aggregated HSS (% , solid) and not EC coverage (% , dashed). ....	20
14. Probability distribution functions of regionally aggregated HSS (%). ....	21
15. Skill variation in response to Nino 3.4. ....	22
16. Skill variation in response to ENSO. ....	22

17. Temporal variations of the proportion of grid points in the above normal category (upper panel) and below normal category (bottom panel). .....	23
18. Aggregated accuracy, resolution, and sharpness in time and space for the raw and debiased probability forecasts of precipitation and temperature. ....	24
19. Aggregated accuracy, reliability, and resolution, in time and space for probability forecasts of precipitation and temperature, based on various combination of PDF estimation and composite. ....	25
20. Aggregated accuracy, reliability, and resolution, in time and space for probability forecasts of precipitation and temperature, based on all available and newly selected models. ....	25
21. Schematic diagram of the probability prediction display system. ....	27
22. Overall PMME precipitation forecast in East Asia. ....	27
23. Overall PMME Temperature forecast in East Asia. ....	29
24. Grid PMME Precipitation forecast. ....	31
25. Grid PMME Temperature forecast. ....	32
26. Forecast of predictors (SST + PC). ....	35
27. Forecast of predictors (TWP Precipitation, left contour; right mesh). ....	36
28. Trends of past forecasts and observations during the recent period. ....	39
29. Relative signs and distance magnitudes (see the relevant coordinate and scale at bottom, respectively) on a 2-dimensional plane depicting regression coefficients of the seasonal mean anomalous precipitation onto the first to PCs for FMA to NDJ. ....	40
30. Regression of seasonal mean anomalous East Asian rainfall on tropical Western Pacific rainfall. ....	41
31. Heidke skill score for East Asia (precipitation). ....	44
32. Heidke skill score for East Asia (temperature). ....	45
33. Forecasted precipitation versus actual (left), forecasted temperature versus actual .....	45
34. Schematics of Bayesian regression. <a href="http://www.indiana.edu/~kruschke/BMLR">http://www.indiana.edu/~kruschke/BMLR</a> . ....	52
35. Distribution of the HSS for hindcast cross-validation of (Left) 1-month lead November, (Middle) 2-month lead December, and (Right) 3-month lead January. ....	54
36. Overview of the customized U-Net architecture with dynamic conditional dropout. ....	55
37. Depth-wise separable convolutional block. Distinct filters by channel convolves separately followed by batch normalization and selu activation. Above operations are performed twice for an individual block. ....	56
38. Operation flow chart for convolutional block attention module. ....	57
39. A newly developed dynamic conditional dropout layer. Auxiliary month information is fed into the network to conditionally control neurons at the bottleneck in a way that each month share significant amount of network, while having limited portion of month-specific pathways. ....	58
40. The overview of the customized artificial neural network developed for the post-processing of MME seasonal climate prediction. ....	59

41. Examples of prediction (left), true observed true category (middle), and observed raw values (right) for selected months (1982-11, 2019-03, 2019-08). .....	61
42. The seasonal and annual hit rates (in %) of APCC-PMMEv1. ....	62
43. The seasonal and annual hit rates of the custom network. ....	62
44. Annual hit rates for APCC-PMMEv1 (red) and the custom network (blue). ....	63
45. Annual hit rates for APCC-PMMEv1 (red) and the custom network (blue). ....	63
46. East Asia probabilistic prediction intercomparison verification results (HSS:upper, RPCC:lower). ....	65
47. Probability prediction regional distribution comparison result (JFM). ....	66
48. Probability prediction regional distribution comparison result (JJA). ....	67
49. The geographic information of the input area in the study. ....	70
50. Machine learning typically follows one of two paths: the most popular methods are in orange color box. ....	71
51. A commonly used class of ensemble algorithms are forests of randomized trees. ....	72
52. The process of the operation in the decision trees of random forest: Each tree in the random forest is built from a sample drawn with replacement from the training set. ....	73
53. The concept of ensemble learning algorithms: A) is describes the process of combining learning and prediction results respectively by multiple learning algorithms. B) describes a model in which multiple learning algorithms are combined to learn and predict(modified from source: Di et al. (2019) An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution). ....	74
54. Diagram of hybrid architecture of this study: (A) work flow for integrating machine learning and deep learning algorithms and (B) simple architecture of convolutional neural network with ensemble classifier. ....	75
55. Cosine similarity measures the similarity between two vectors by calculating the cosine of the angle between the two vectors. Cosine similarity is 1 at theta = 0 and -1 at theta = 180, that means for two overlapping vectors cosine will be the highest (a) and lowest for two exactly opposite vectors (d). ....	78
56. The decision tree structure by the n_estimator(i.e. 5) of the random forest developed in this study. ....	78
57. Optimization of random forest hyperparameters: The result of optimizing the hyperparameters(a: n_estimator, b: min_sample split, c: max depth, d: max_feature, e: min_sample leaf, f: learning curve) of the random forest algorithm developed for the predicted subseasonal daily maximum air-temperature of ECMWF in this study. ....	79
58. Feature Importance of random forest: Feature importance results of random forest calculated by applying the random forest algorithm developed in this study to 5 individual climate models(a: ECMWF, b: GloSea5, c: UKMO, d: CMA, e: NCEP) and MME-based subseasonal daily maximum air-temperature(f). ....	80
59. Comparison of spatial anomaly pattern correlation coefficients of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME before applying	

training models. ....81

60. Histograms and two-dimensional distributions of input data transformed by four transformers for sub-seasonal MME predicted daily maximum air-temperature: The four transformers are (a) minMax scaler, (b) Robust scaler, (c) Standard scaler, (d) maxAbs scaler, respectively. Two-dimensional distribution between the Era-Interim daily maximum air-temperature matched with the input data transformed by the four transformers. ....83

61. Histograms and two-dimensional distributions of input data transformed by four transformers for ECMWF sub-seasonal predicted daily maximum air-temperature: The four transformers are (a) minMax scaler, (b) Robust scaler, (c) Standard scaler, (d) maxAbs scaler, respectively. Two-dimensional distribution between the Era-Interim daily maximum air-temperature matched with the input data transformed by the four transformers. ....83

62. Convolutional short-term memory (ConvLSTM) training results: (a) Loss after training ConvLSTM to sub-seasonal MME predicted daily maximum air-temperature. Losses after learning ConvLSTM of daily maximum air-temperature of ECMWF (b), GloSea5 (c), and UKMO (d), respectively. ....86

63. Comparison of training and test scores of a random forest connected to a hybrid model and a single random forest for the sub-seasonal MME predicted daily maximum air-temperature. ....87

64. Comparison of spatial anomaly pattern correlation coefficients (after here, PCC) by prediction frequency: PCC of predicted daily maximum air-temperature of sub-seasonal MME before applying training models (left). PCC of predicted daily maximum air-temperature of sub-seasonal MME after applying Random Forest model (middle) and the Hybrid model (right).  
88

65. Comparison of the deviation of the spatial anomaly pattern correlation coefficient (after here, PCC) by prediction frequency of predicted daily maximum air-temperature of sub-seasonal MME: The PCC after application by each model (left: Random Forest, right: Hybrid model) is subtracted from the PCC before applying by each model. ....88

66. Comparison of mean, highest, and lowest values of deviations of spatial anomaly pattern correlation coefficients by prediction frequency of predicted daily maximum air-temperature of sub-seasonal MME. ....89

67. Comparison of spatial anomaly pattern correlation coefficients by prediction frequency of predicted daily maximum air-temperature of five individual climate models (from left, ECMWF, NCEP, GloSea5, UKMO, CMA): Spatial correlation coefficients for each prediction frequency before applying models (upper panel), after applying by the hybrid model (center panel) and random forest (lower panel). ....90

68. Comparison of the deviation of the spatial anomaly correlation coefficient by prediction frequency of predicted sub-seasonal daily maximum air-temperature of five individual climate models: Subtracted after application by each each training model (top: Hybrid model, bottom: Random Forest) from before applying by the models. ....91

69. Comparison of improvement rates of post-calibration techniques for the sub-seasonal daily

maximum air-temperature of individual climate models and MME, respectively: The x-axis is the ratio between the spatial anomaly pattern correlation coefficient (after here, PCC) after applying by the training models, for example Random Forest and the Hybrid model and the PCC before applying the models. The y-axis is the ratio between the root mean square error (after here, RMSE) after applying by the training models and the RMSE before applying the models. ....94

70. Comparison of cosine similarity between spatial anomaly pattern correlation coefficients (after here, PCC): Era-Interim's predicted daily maximum air-temperature is defined as observation, that is, true value. Calculation of cosine similarity between the PCC of the Era-Interim's daily maximum air-temperature and the PCC before applying the training models to the daily maximum air-temperature on the sub-seasonal MME (from (a) to (d)). ..95

71. The density of cosine similarity by week in Figure 21 is from a to d. ....95

72. Comparison of cosine similarity between spatial anomaly pattern correlation coefficients (after here, PCC): Calculation of cosine similarity between the PCC of the Era-Interim's daily maximum air-temperature and the PCC before applying the training models to the sub-seasonal daily maximum air-temperature of five individual climate models. Calculation of cosine similarity between the PCC of the daily maximum air-temperature of Era-Interim and the PCC of the sub-seasonal daily maximum air-temperature of five individual climate models applied by the hybrid model and Random Forest. From left, (a) is ECMWF, (b) is NCEP, (c) is Glossea5, (d) is UKMO, (e) is CMA. ....96

73. Glossea5 daily maximum temperature data. ....99

74. ERA-Interim daily maximum temperature data. ....100

75. A diagram of ConvLSTM for Subseasonal model. ....102

76. The structure of parameter of ConvLSTM. ....105

77. Model pattern Extraction and visualization of ConvLSTM ....106

78. GPU monitoring and CUDA activity. ....107

79. Sensitivity analysis of optimizing parameter and configuration of input data. ....108

80. Comparison of the result of convLSTM model with ECMWF model at daily maximum temperature. ....109

81. Comparison of the pattern correlation for (a) convLSTM (b) ECMWF. ....110

82. Comparison of the pattern correlation for 1-4 week of lead-time of Glossea5, NCEP, UKMO model. ....111

83. The sensitivity of analysis for spatial resolution. ....112

84. Comparison of performance of pattern correlation of maximum temperature for 3 week lead-time. ....112

85. Comparison of performance of pattern correlation of maximum temperature for 4 week lead-time. ....113

86. Comparison of map of the pattern correlation from ConvLSTM with MME data. ....114

87. Comparison of pattern correlation at maximum temperature for 3-4 week of lead-time for MME. ....114

88. Comparison of RMSE at maximum temperature for 3-4 week of lead-time for MME. ...115

# 1. 서론

## 가. 연구의 배경

지난 과거 40년 동안, 전 지구 관측 시스템의 획기적인 진보, 지구 시스템의 다양한 요소들에 대한 폭 넓은 이해와 모델링 기술의 향상, 계산 능력 및 컴퓨팅 기술의 혁신을 바탕으로 세계 도처 주요 현업 기관의 날씨 및 기후 예측은 꾸준히 성장하고 진화 해 왔다(Goddard et al. 2001, Guilyard 2009, Wang et al. 2004, Bauer et al. 2015). 이러한 진보에 발 맞추어 APEC (Asia-Pacific Economic Cooperation) 기후센터 (APEC Climate Center, APCC)의 다중모형 앙상블 (Multi-Model Ensemble, MME)이나 World Meteorological Organization (WMO) Lead Center for Long-Range Forecasts MME (LC-LRFMME)와 같은 계절 예측 인프라 (Seasonal Forecast Infrastructure, SFI)는 현업적으로 이미 성숙 단계에 이르렀다. 즉, APCC MME, WMO LC-LRFMME, 북미의 MME (North American Multi-model Ensemble, NMME), 유럽 Copernicus Climate Change Service (C3S)의 MME와 같은 다양한 SFI에서는 주로 전구 관점에서 기후예측 정보를 서비스하고 있으며, 각 SFI는 고유의 특징 및 여건에 따라 특정 지역의 계절 예측 생산 시 활용되고 있다. 또한, 계절 내 (Subseasonal to Seasonal, S2S) 시간규모 (2주-2달) 예측에 대한 수요와 인식의 증가로 최근 폭 넓은 연구 및 방대한 자료 공유 (S2S Prediction Project Database; Vitart et al. 2017)가 이루어지고 있다.

현재 APCC에서는 매월 전 세계의 14개 우수 현업기관 및 연구기관으로부터 진보된 계절 예측 자료를 수집하여 다중모형 앙상블 기법을 활용하여 통합된 계절 예측 자료를 생산하고 있다. 따라서, APCC는 명실상부 최적의 계절 예측 자료를 전 지구 규모에서 제공하고 있으며, 지역, 국가, 지점 규모에서 계절 예측 생산 시 활용 가능한 고품질의 자료를 제공하고 있다. 좀 더 나아가, 태평양 도서국 및 동아시아 지역에 대해서는 해당 지역을 좀 더 확대한 계절 예측 자료를 제공하고 있다(Sohn et al 2018, [clikp.sprep.org:apcc21.or/ser/eastaisa/outlook.do?lang=ko](http://clikp.sprep.org:apcc21.or/ser/eastaisa/outlook.do?lang=ko)).

한편 실질적인 활용성 증대를 위해서, WMO 에서는 유사한 기후 특성을 기반으로 하는 지역 규모에서 기후 전망이나 해석이 이루어져야 함을 권고하고 있다 (<https://public.wmo.int/en/our-mandate/climate/regional-climate-outlook-products>). 또한 다양한 WMO 포럼을 통해서 끊임없이 논의 및 제고되었듯이, 지역 및 국가 규모의 계절 예측을 개발 하기 위해서는 적절한 전구 규모 SFI 활용 전략 수립이 필요하다. 따라서, 지역이나 국가 규모의 예측정보를 생산하기 위한 절차로, WMO에서는 전구 규모의 기후예측정보(즉, 다중모형 또는 MME)를 바탕으로 지역 맞춤형 (Regionally and Nationally Tailored) 예측기술을 개발할 것을 권고하고 있다. 가령, 공식적인 북미 지역의 계절 예측자료 생산 시 NMME는 하나의 권위 있는 원천으로 활용되어 다양한 (통계적인) 예측 도구의 정보들과 통합한다. APCC에서는 동아시아 지역에 대한 기후전망을 제공하고 있으나, 이는 지역 맞춤형으로 후보정 되지 않은 자료이다.

기존 MME 기법은 가능한 모든 자료를 활용하여 결정론적으로 단순 평균하거나(Simple Composite Method, SCM), 앙상블 멤버의 크기에 따라 역으로 가중치를 주어 확률적으로 평균

하는 통상적인 결합방식으로, 중위도 이상 지역에서는 예측성능이 낮다. 일반적으로, MME 기법에서는 개별모델의 역학적 특성뿐만 아니라 모델예측 간 상관관계도 주요한 역할을 하는데, 현행의 MME 기법에서는 그러한 점이 고려되지 않고 있다. 또한, 널리 사용되는 회귀모형과 같은 자료 중심적(data-driven) 방법은 표본수가 충분하지 못한 자료 특성이 있어 과적합(overfitting)의 문제를 내포하고 있다. 아울러, 다중 모델이 예측하는 미래 기후의 경향성을 파악하는 데 집중하여 모형 내/외의 불확실성 등의 부가 정보는 일부만 활용하고 있다.

한편, WMO 선도센터 및 NMME는 최근 S2S 예측자료를 보유 및 제공 중이며 국제적으로 다양한 연구들이 수행되고 있다. 하지만, 그 성능은 단기, 장기 예보에 비해 현저히 떨어진다. 특히 한반도를 비롯한 동아시아 지역의 성능은 다른 지역에 비해 현저히 떨어지고 MME 적용의 효용성도 미비하다. 반면에 최근 각광받고 있는 딥러닝 기술은 다양한 알고리즘의 개발과 디지털 혁명, 컴퓨터의 발달로 갈수록 방대해지는 데이터의 축적에 따라, 다양한 분야에서 복잡한 문제들을 해결하는데 널리 적용되고 있다. 이러한 딥러닝 기술은 예측이 어려운 기후 분야에서도 강수량 등의 예측성능개선에 대한 가능성을 보여 주었다. 일반적인 기계학습 기술을 이용하여 계절 예측 데이터의 정확성을 향상시키려는 연구가 최근 수년 동안 수행되고 있지만, 계절내 예측 자료의 경우 일반 기계학습이나 딥러닝 기술을 활용한 연구는 거의 없다. 따라서 모델 기반 예측기술 개발에 집중하고, 예측성 향상을 위한 계절/계절내 기후예측 기술을 고도화 할 필요가 있다.

MME 계절기후예측의 예측성 근간은 앙상블에 참여하는 개별모델의 성능에 크게 의존하는데, 참여 모델 자체의 성능 개선은 센터 내에서는 수행하기 어려우며 모형 운영 기관 (Model provider)의 모형 향상 계획에 전적으로 의존한다. 또한, 계절기후예측에서는 월별 자료만 제공하는 등 자료 가용성 (data availability)에 대한 한계가 존재한다. 따라서 센터 내 수집된 다중 모델 자료를 바탕으로 기후예측 자료를 분석 및 해석하고, 다양한 통계적 기법을 적용하여 예측성 향상을 위한 계절 및 계절내 기후예측 기술을 개발하고자 한다.

## 나. 연구의 목표

동아시아 (및 한반도) 지역을 위한 맞춤형 기후 예측기술을 개발하기 위한 방안으로, APCC에서 수집 가능한 다중모델 예측 결과를 활용하되 수집된 자료의 시간규모 특징에 따라 크게 두 가지로 나누었다. 즉, 이 과제에서는 동아시아 지역을 중점 연구지역으로 설정하여 지역별 특화된 다중모형 기반 계절/계절내 예측 기술 개발을 통해 계절/계절내 예측성 향상을 목표로 하였다. 따라서 이 과제를 통해

- (1) 한반도 및 동아시아 맞춤형 계절 확률예측 시스템 구축
  - (2) 동아시아 계절내 예측 기후의 예측성 향상을 위한 신기술 하이브리드 기법 구축
- 을 달성하고자 한다.

## 다. 연구의 내용 및 추진전략

이 과제에서는 예측 시간 규모 및 예측자료의 유용성 등을 고려하여, Figure 1과 같이 크게 (1) 예측성 향상을 위한 다중모델 기반 계절 예측 기술개발 및 (2) 다중모델 계절내 예측 기법 기술 개발로 구분하여 연구를 진행하고자 한다. 연구 내용 (1)과 (2)에 관련된 연구 방법 및 결과는 다음 2-5장과 6-7장에서 각각 나타내었다.



Figure 1. Research Structure.

## 라. 연구의 추진체계

이 과제는 5단계 아태사업으로 총 2019년부터 2021까지 3년 동안 수행되는 다단계 사업이다. 5단계 사업의 최종 목표는 “지역특화형 선도 기후예측 기술 개발” 으로 연차별 최종 목표는 Table 1과 같다.

Table 1. Final goal and annual objectives of the research project.

5단계 (2019–2021) 최종목표	
지역특화형 선도 기후예측 기술 개발	
연차별 최종목표 및 현재까지의 달성내용	
1차년도( '19)	지역 맞춤형 계절 다중모델앙상블 및 계절내 MME 기후예측 기술 개발 체계 구축
2차년도( '20)	지역 맞춤형 계절 MME 및 계절내 MME 기후예측 기술 개발
	물리현상 기반의 앙상블모델      물리현상 기반의 앙상블모델 재구성을

	재구성을 통한 지역특화 계절 예측기술 개선 및 예측성 검증	통한 지역특화 계절 예측기술의 예측성능 안정성 평가 및 동아시아 지역의 계절 MME 기후예측을 위한 climate driver 추가 발굴
	베이지 접근법에 의한 다중모델 앙상블 기법 개발	베이지 회귀를 통한 과거 관측 -개별 모델 예측을 적합(fitting)하고 정규화 기법을 적용해 안정성을 향상시킨 후처리 기법을 개발, 과거기후예측 (hindcast) 기간의 삼분위 확률 예측 정보 생산, 기존 기법 및 관측과 결과 비교
	Probability Anomaly Correlation (PAC)를 통한 확률적 기후예측 보정	동아시아 지역 계절 확률 MME 기후예측에 대한 PAC 적용 및 검증
	딥러닝을 활용한 계절내 예측 강수, 최고·최저 기온의 MME 기법 개발 및 계절내 예측의 오차 저감	예측빈도 및 기간이 상이한 기후모델 기반의 통합 MME 기법 구축 및 딥러닝 기술 적용에 따른 MME 기반 계절내 예측자료의 오차 저감
3차년도( '21)	다중모델 기반의 지역특화 계절예측시스템 구축 및 계절내예측 기반기술 기술 고도화	
	다중모델기반 계절예측 기술개발	지역맞춤형 전문가 계절 예측 시스템 개선 및 운영
		기후모델 재구성 및 확률예측 기법 개발, 동아시아 맞춤형 확률예측시스템 구축
		심층학습 인경신공망 모델을 이용한 다중모델 계절예측 후보정 기법 개발 및 예측성이 향상된 기후 예측 자료 생산 기반기술 상호비교
다중모델기반 계절예측 기술개발	머신러닝 및 신경망 모델 결합을 위한 딥러닝 하이브리드 알고리즘 구축, 동아시아 MME 계절내 기후 예측성능 평가	

Figure 1의 연구의 추진 전략에서 나타난 세부 연구와 그에 따른 구체적인 2020년 목표 및 달성내용은 Table 1에 나타내었다. 2020년 dms 전체 사업의 2차년도로 1차년도에 구축된 지역 맞춤형 계절 MME 및 계절내 MME 기후예측 기술 개발 체계를 기반으로 지역 맞춤형 계절 MME 및 계절내 MME 기후예측 기술 개발을 목표로 한다. 과제가 마무리되는 3차년도인 2021

년에는 2차년도에 개발된 계절예측 기술을 바탕으로 다중모델 기반의 지역특화 계절예측시스템을 구축하고, 기반 기술인 계절내 예측 기술을 고도화하고자 한다. 다년간 수행된 세부 연구의 자세한 연구 방법 및 연구 결과는 2장부터 7장까지 기술되었다.

## 2. 우리나라 맞춤형 전문가 계절 예측 시스템

### 가. 개요

이 연구를 통해서 우리나라 맞춤형 전문가 계절 예측 시스템(ESPreSSO)이 고도화하여 민첩성(agility)/정확성(accuracy)/접근성(accessibility)이 크게 향상되었다. 특히, APCC의 최신 다중모델앙상블 계절 기후 예측 시스템의 구성을 반영한 새로운 전문가 가이드(ESPreSSO<sup>v3.2</sup>)를 추가하여, 전문가 가이드와 인공지능이 협력한 최적의 예측 정보를 생성하도록 하였다. 지난 1년간 고도화된 ESPreSSO를 활용한 3 개월 계절 기후 예측을 통해 내부 현업 운영을 지원하였으며, 최신 ESPreSSO<sup>v3.2</sup>의 내부 현업화를 추진하여 ESPreSSO의 계절 기후 예측 정보가 공식적으로 현업 운영에 활용될 수 있도록 조치하였다. 이에, ESPreSSO<sup>v3.2</sup>의 민첩성/정확성/접근성 향상에 대한 구체적인 내용과 내부 현업화 지원에 관한 사항을 정리하고자 한다.

### 나. 새로운 버전의 ESPreSSOv3.x 개발

#### (1) 민첩성(agility), 예측성(accuracy), 접근성(accessibility)을 향상시킨 ESPreSSOv3

이 연구를 통해 기존 ESPreSSO (Expert Seasonal Prediction System for operational Seasonal Outlook)를 새로운 다중모델앙상블 구성에 맞춰 업데이트한 전문가 가이드와 인공지능의 협업을 통해 최적의 예측 결과를 생산할 수 있도록 고도화(Figure 2)하였다.

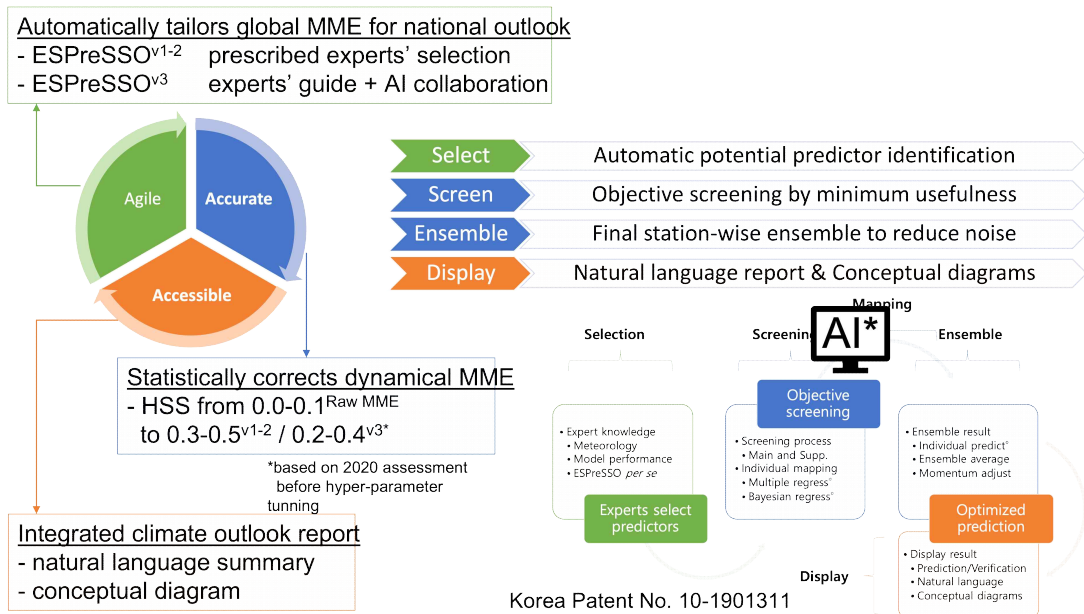


Figure 2 The overview of ESPreSSO - Expert Seasonal Prediction System for operational Seasonal Outlook.

#### (가) 민첩성(agility)의 향상

다중모델앙상블 자료를 기반으로 전문가의 예보 토의를 거쳐 수작업으로 진행하던 최종 계절 기후 예보 생산과정을 미리 작성된 최신 전문가 가이드와 인공지능의 협력으로 대체하여, 새로운 계절 기후 예측값을 활용해 최종 계절 예보를 자동적으로 작성할 수 있게 되었다. 이 과정에서 ESPreSSO의 인공지능 예측 인자 선별 모듈은 전문가 가이드에 따른 잠재 예측 인자 가중치를 고려하는 한편, APCC의 다중모델앙상블 계절 기후 예측 시스템의 예측성 및 재현성, 해당 인자의 위도별 특성, 영향 반경 등을 고려하여 잠재 예측 인자를 추천하고, 이는 최종적으로 이미지 및 라벨 분류를 통한 최종 예측 인자로 선정된다. 이 과정이 자동적으로 수행되어 기존 수작업에 의한 예측 보정에 비해 획기적으로 민첩성이 향상되었다. 또한, 기존 ESPreSSOv1-v2의 전문가의 직접적인 개입을 전문가의 소극적인 가이드와 인공지능의 협력 시스템으로 치환하여, 향후 다중모델앙상블 구성 등 환경 변화에 민첩하게 대처할 수 있도록 구성하였다.

### (나) 예측성(accuracy)의 향상

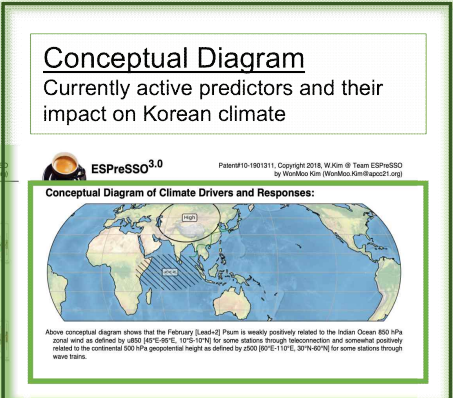
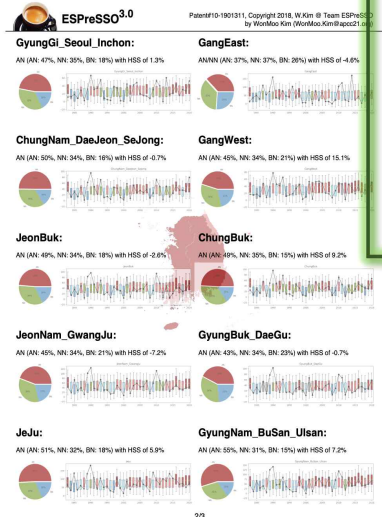
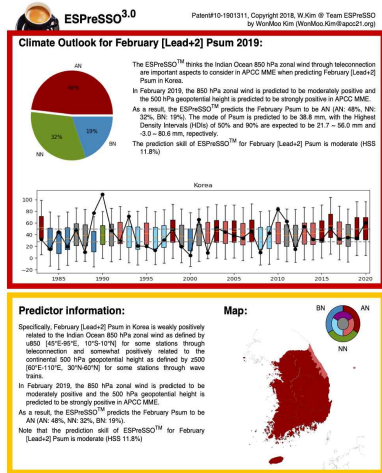
ESPreSSO의 예측 성능을 위해 새로운 다중모델앙상블 구성에 맞춰 갱신된 전문가 가이드와 인공지능 협업에 의해 개별 관측소에 대한 예측 인자를 선정할 수 있도록 제작하였다. 개별 관측소에 대해 선택된 주/부 예측인자를 바탕으로 베이스 회귀를 통한 계절 예측 분포가 생성되며, 이러한 개별 예측을 앙상블하여 예측을 안정화하고 성능이 향상될 수 있도록 조치하였다.

특히, 이번 연구를 통하여 APCC의 최신 다중모델앙상블 계절 기후 예측 시스템의 구성을 반영한 전문가 가이드를 작성(※전문가 가이드 작성: 기후분석과 여새림 박사)하였는데, 새로운 전문가 가이드는 기온 및 강수에 대해 각각 작성되어있으며, 계절별 특성을 반영하여 월별로 분리 작성하였다. 기온의 경우 월별 약 3-4 개의 잠재 예측 인자가, 강수의 경우 월별 약 1-2 개의 잠재 예측 인자가 포함된 최신 전문가 가이드는 APCC의 다중모델앙상블 2020 년 구성을 기반으로 제작되었으나, 예측 특성뿐만 아니라 관측의 특성도 적극 고려하여 향후 다중모델앙상블의 소규모 구성 변화에는 크게 민감하지 않도록 유의하였다.

### (다) 접근성(accessibility)의 향상

ESPreSSO는 삼분위, 시계열, 대푯값, 최고밀도 구간 등 다양한 방식의 예측 및 검증 결과 표출을 통해 다양한 정보를 한눈에 파악할 수 있도록 구성(Figure 3)하였다. 뿐만 아니라, 예보관이 현업 계절 기후 예보에 직접 활용할 수 있도록 주요 예측 인자 및 그 활동 현황, 이들이 우리나라 기후에 영향을 미치는 방식과 강도 등 ESPreSSO의 예측 결과 산출 근거를 자연어 설명을 통해 제공함으로써 블랙박스 모형을 탈피하고 예측 정보 접근성을 확보하였다. 특히, ESPreSSO<sup>v3</sup>에서는 기후 모식도를 제공하여 지도 위에 표현된 주요 예측 인자의 활동과 이들의 영향 방식을 시각적으로 제공함으로써, 예보자가 직관적으로 계절 기후 예측 현황을 파악할 수 있도록 하였다.

**Climate Outlook Summary**  
 (left) Pie chart, (bottom) history, and (right) natural language outlook on prediction (probabilistic/deterministic), predictor activity and impact, predictability assessment



**Verification Information:**

Lead	Cast	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	ALL	13.2	17.1	0.0	15.1	-8.1	2.0	-2.0	7.4	12.8	8.1	16.9	16.9
	HIND	28.0	11.6	-0.0	14.3	-5.4	5.4	10.7	17.0	10.7	29.0	25.0	
2	ALL	5.9	11.8	3.3	7.2					14.9	8.5	8.8	26.4
	HIND	8.9	9.8	3.1						25.0	5.4	17.0	25.0
3	ALL	0.7	5.9	5.9						2.0	-6.8	-6.1	10.8
	HIND	-0.8	5.4	7.1						16.1	9.8	-0.0	-4.5
FORE	ALL	5.0	7.5	2.5	17.5	-5.6	-22.2	-27.8	-11.1	-16.7	-8.3	-8.3	41.7
	HIND	5.0	7.5	2.5	17.5	-5.6	-22.2	-27.8	-11.1	-16.7	-8.3	-8.3	41.7

**Predictor information & Regional outlook**  
 (left) Detailed predictor information  
 (right) regional climate outlook

**Verification information**  
 Heidke Skill Scores (HSS) for cross-validated hindcasts and independent forecasts

Figure 3 An example integrated climate outlook report of ESPreSSO<sup>v3</sup>.

**다. 민첩성/예측성/접근성이 향상된 ESPreSSOv3의 현업 운영 지원**

**(1) ESPreSSOv3.0 기반의 3 개월 예측 정보 생산 및 제공 (매달)**

민첩성, 예측성, 그리고 접근성이 향상된 ESPreSSOv3는 2021 년 한 해동안 APCC 내부 예보 토의에 3 개월 예측 정보를 제공함으로써 APCC 내부의 현업 운영을 지원(Figure 4)하였다. 특히, ESPreSSOv3는 다양한 방식으로 표출되는 예측 결과와 기후 예측 현황에 대한 자연어 설명과 더불어 예상되는 주요 예측 인자의 활동과 이들이 우리나라 계절 기후에 미치는 영향이 도식으로 정리된 기후 모식도를 제공하여 접근성을 크게 향상시켰다. 예측 결과의 산출 근거를 다양한 방식으로 제공하여, 블랙박스 모형을 탈피하고 예측 과정과 결과에 대한 예보자의 이해를 증진시키고자 하였다.

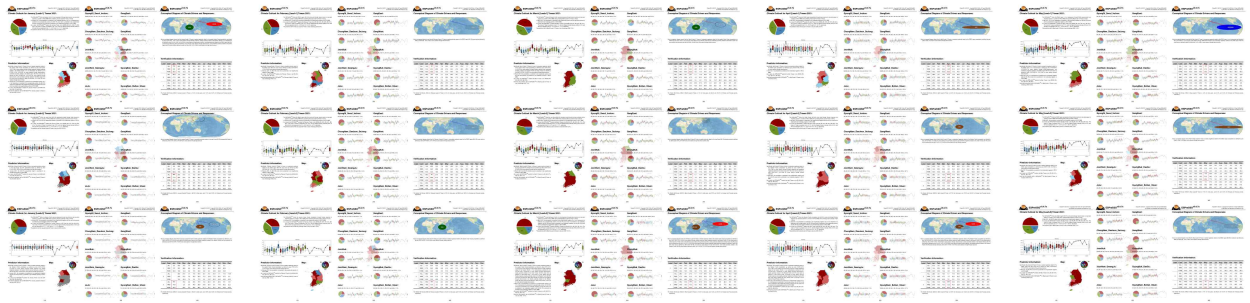


Figure 4 Examples of operational reports of ESPreSSO<sup>v3.0</sup> in 2021.

(2) 새로운 전문가 가이드가 적용된 ESPreSSOv3.2의 내부 현업화

또한, APCC의 최신 다중모델앙상블 구성에 맞춰 갱신한 새로운 전문가 가이드가 적용된 ESPreSSO<sup>v3.2</sup>가 정기적으로 운용될 수 있도록 내부 현업화를 추진(Figure 5)하였고, 현업화 심의에 통과하고 자동화 프로세스를 완료하였다. 이를 통해 ESPreSSO가 APCC 내부의 계절 기후 예측에 공식적으로 활용될 수 있는 기반을 마련하였으며, 연산 자동화를 통해 운영의 효율성을 높이게 되었다.

※ ESPreSSO<sup>v3.2</sup>의 현업화 신청('21.10.13.) 및 심의('21.10.22.)

**현업화 대상: ESPreSSO<sup>v3.2</sup>**

기대효과: 최신 전문가 가이드와 인공지능의 협력을 통한 민첩(Agile)하고, 정확(Accurate)하고, 접근성(Accessible) 높은 우리나라 맞춤형 계절 전망 제공

**ESPreSSO 현업화 기술 심의 완료**

**현업화 기술 심의신청서**

2021. 10. 13. (수석기상관료 김민우)

**현업화 기술 내용**

ESPreSSO<sup>v3.2</sup> (Expert Seasonal Prediction System for operational Seasonal Climate Service)는...

**현업화 신청 사유**

구분	현업화	비현업화
현업화 대상	○	○
현업화 대상	○	○
현업화 대상	○	○

**현업화 필요성**

ESPreSSO<sup>v3.2</sup>는 최신 다중모델앙상블 구성에 맞춰 갱신된 전문가 가이드를 적용하여, 계절 기후 예측의 정확도를 높이고, 연산 자동화를 통해 운영 효율성을 높일 수 있다.

**현업화 방안**

ESPreSSO<sup>v3.2</sup>는 최신 다중모델앙상블 구성에 맞춰 갱신된 전문가 가이드를 적용하여, 계절 기후 예측의 정확도를 높이고, 연산 자동화를 통해 운영 효율성을 높일 수 있다.

**현업화 기대효과**

ESPreSSO<sup>v3.2</sup>는 최신 다중모델앙상블 구성에 맞춰 갱신된 전문가 가이드를 적용하여, 계절 기후 예측의 정확도를 높이고, 연산 자동화를 통해 운영 효율성을 높일 수 있다.

※ 제출된 ESPreSSO<sup>v3.2</sup>의 현업화 기술 심의 신청서

▶ 전문가-인공지능 협력에 기반한 우리나라 맞춤형 계절 전망 생산 시스템인 ESPreSSO<sup>v3.2</sup>

Figure 5 Operationalization (deliberation process) of ESPreSSO<sup>v3.2</sup>.

### 3. 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 개선 및 시스템 구축

#### 가. 개요

1차년도인 2019년에는 “지역 맞춤형 계절 다중모델앙상블 기후예측 기술 개발을 위한 체계가 구축” 되었으며, 이는 물리현상 기반의 앙상블 재구성을 통한 지역특화 계절 예측기술 개발을 통해서 그 체계가 마련되었다. 참여모델의 지역별 예측성능 평가 및 차별화된 MME 구성을 위해서 APCC MME 기후예측에 참여하는 모델의 후보자료(1982-2005)를 바탕으로 SST, 강수량 및 지표면온도의 예측성능이 분석되었다. 엘니뇨-남방진동(El Nino와 Southern Oscillation, ENSO)의 복잡성과 그 원격상관의 예측성능을 진단하고 향상시키기 위한 물리적 인자로 사용되었다. 즉 ENSO 및 그 복잡성에 대한 모델의 재현성이 모델의 충실성(fidelity)을 위해서 사용되었다. 그 결과를 바탕으로 충분한 충실성을 가진 모델을 가지고, 여러 가지 모델 선별 방안으로 MME를 구성하였다. MME 조합은 모델의 충실성을 어떻게 평가하느냐에 따라 달라지는데, ENSO를 바로 평가하거나 그 원격상관의 재현성을 평가하는가에 따라 다르게 구성하였다. 모델 선택의 효과는 계절 및 지역에 따라 달랐다. 전체 연구사업의 2차 년도인 2020년에는 1차 년도에 구축된 기술 개발 체계를 바탕으로, 지역 맞춤형 계절 MME 기후예측 기술을 개발하였다. 따라서 1차 년도에 개발된 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술을 개선하고 그 예측성을 검증하고자 하였다. 3차년도인 당해연도에는 기후모델을 재구성하고 이에 적합한 확률예측 기법을 개발, 이를 바탕으로 최종적으로 동아시아 맞춤형 확률예측 시스템을 구축하였다.

#### 나. 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측 기술

이 과제는 2018년 4월에 과제 컨셉 기획, 같은 해 9월에 과제의 전체 3년 내용이 제안되었다. 따라서 과제 수행 1년차에는 APCC 6개월 계절예측 현업 기준으로 MME version 1에 참여 중인 개별 기후모델을 바탕으로 연구가 진행되었다. 하지만 사용된 모델 자료는 과제 2년차인 2020년 중반에 APCC 6개월 계절예측이 MME version 2로 변경되었다. 따라서 현업 활용성을 제대로 평가하기 위해서 다양한 기후 자료의 기후예측성능 검증을 통한 안정성 평가가 수행 되었다.

##### (1) 앙상블 재구성을 위한 물리현상

1차 년도인 2019년도에는 지역 특화형 선도 기후예측 기술 개발을 위한 지역 맞춤형 계절 다중모델 앙상블 기후예측 기술 개발 체계가 구축되었다. 이 연구에서 개발된 기술 체계는 크

게 다음 두 가지 과정으로 구성된다.

- (1) 참여모델의 지역별 예측 성능 평가,
- (2) 차별화된 MME 구성 방안

참여모델의 일반적 예측 성능 평가 및 특성을 분석하고, 이에 기반을 둔 아태지역 내 지역별 차별화된 MME 구성 적용 가능성을 평가 하였다. 즉 지역 특화형 선도 기후예측 기술 개발을 위해서, 물리현상을 기반으로 앙상블을 재구성하였다. 따라서 이 연구에서는 엘니뇨-남방진동을 새롭게 정의하고, 2차년도에 서태평양 대류활동을 새로운 기후인자로 추가하여 기후영향 (climate impact)을 분석하였다. 이 연구에서 ENSO는 중·동태평양 ENSO의 혼합을 캡처하기 위해서 처음 두 EOF 모드(Figures 6와 7 참조)의 결합으로 정의되었다(Timmermann et al., 2018). 서태평양 대류 활동은 필리핀을 중심으로 하는 지역(Figure 8에서 하위 네모 박스로 표시)의 영향을 나타내며, 동아시아 강수 예측을 위해서  $110^{\circ}$  - $140^{\circ}$  E,  $5^{\circ}$  S- $25^{\circ}$  N의 영역에서 해당 계절 강수 아노말리를 평균하여 사용하였다. 따라서, 이에 대한 모델의 재현성을 바탕으로 모델을 선별하여 지역 맞춤형 MME를 개발하였다.

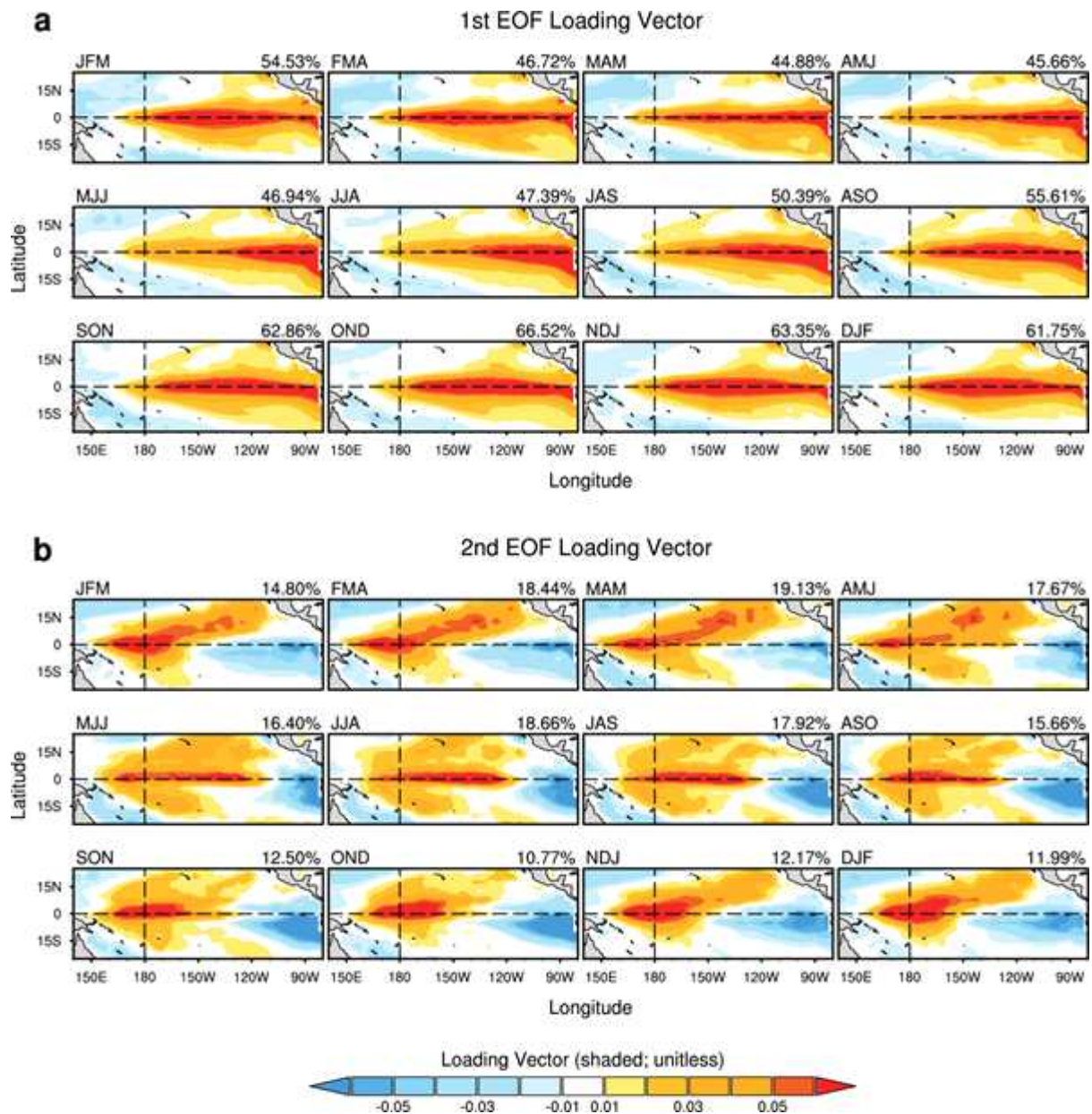


Figure 6. (a) First and (b) second EOF patterns of a monthly rolling 3-month mean SSTAs computed for  $25^{\circ}$  S– $25^{\circ}$  N and  $140^{\circ}$  E– $80^{\circ}$  W during 1983–2005, based on OISST. The fractional variance explained by each mode and season are provided in the upper right of each panel. The horizontal and vertical black dashed lines denote the equator and International Date Line.

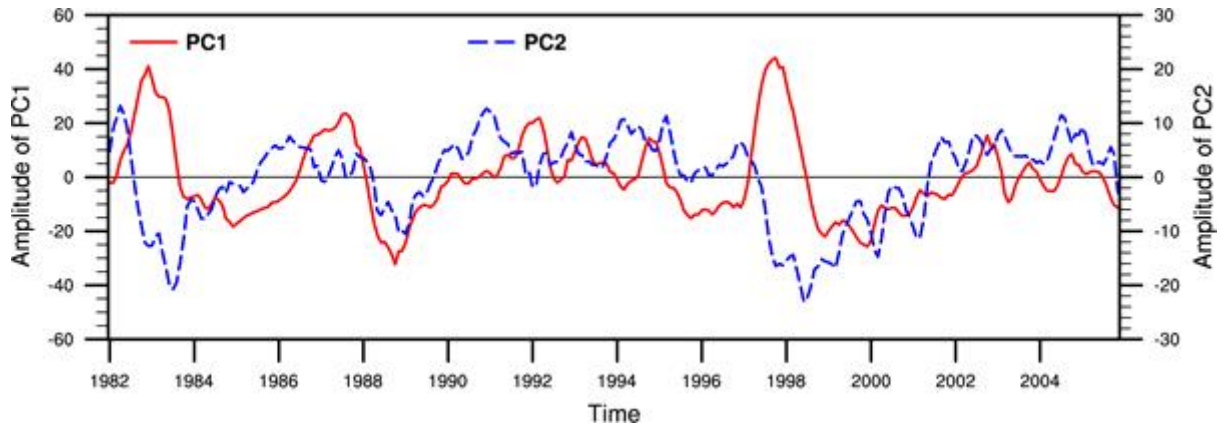


Figure 7. Time series of the combined PCs of (a) the first mode (red solid line for left y-axis) and (b) the second mode (blue dashed line for right y-axis) associated with Figure 2 with all months during January 1982 to December 2005.

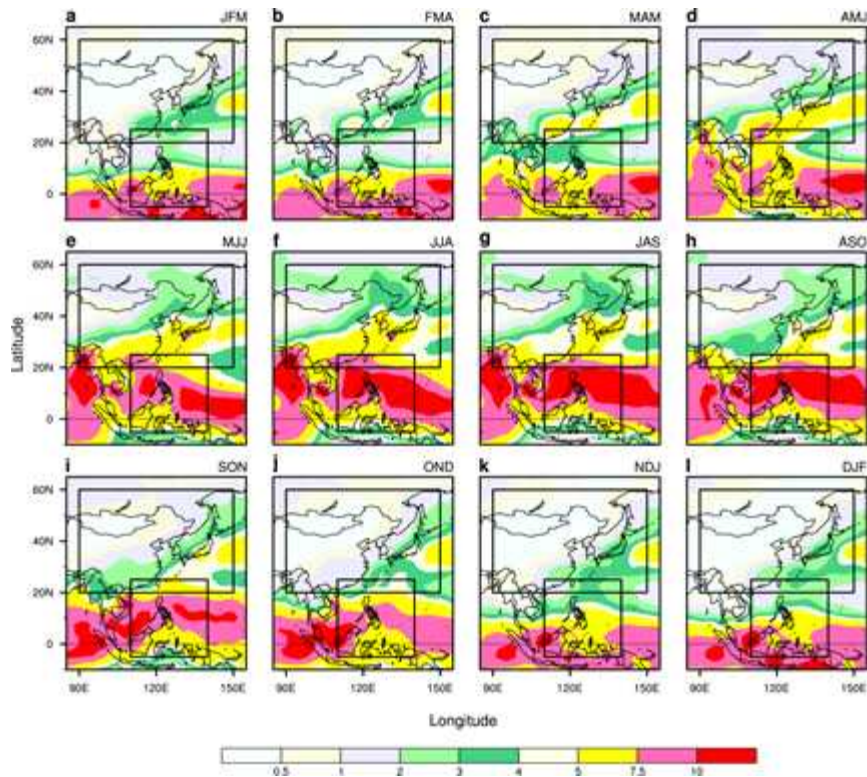


Figure 8. Climatological seasonal mean of East Asian and tropical west Pacific rainfall.

## (2) 물리현상 기반의 앙상블모델 재구성

동아시아 지역 계절예측기술 개발을 위해서 엘니뇨-남방진동의 강수 영향(climate impact)에 기반 하여(Figure 9 참조) 모델을 선별하였다. Figure 10은 열대 서태평양 지역(110° -140° E,

5° S-25° N)에서 영역 평균된 강수 지수에 대한 동아시아 여름 강수 변동성에 대한 회귀 계수를 나타낸다. 이러한 동아시아 여름 강수와 열대 서태평양 강수의 변동성 사이의 관련성은 통계적으로 유의하며, 대부분의 계절에서 95% 신뢰구간을 넘어선다. 즉, ENSO 모드(또는 서태평양 대류 활동)와 동아시아 지역 강수의 관련성에 대한 재현성에 대한 모델의 성능을 평가하고, 이를 새로운 MME의 서브셋을 위해서 성능 좋은 모델을 선별하였다.

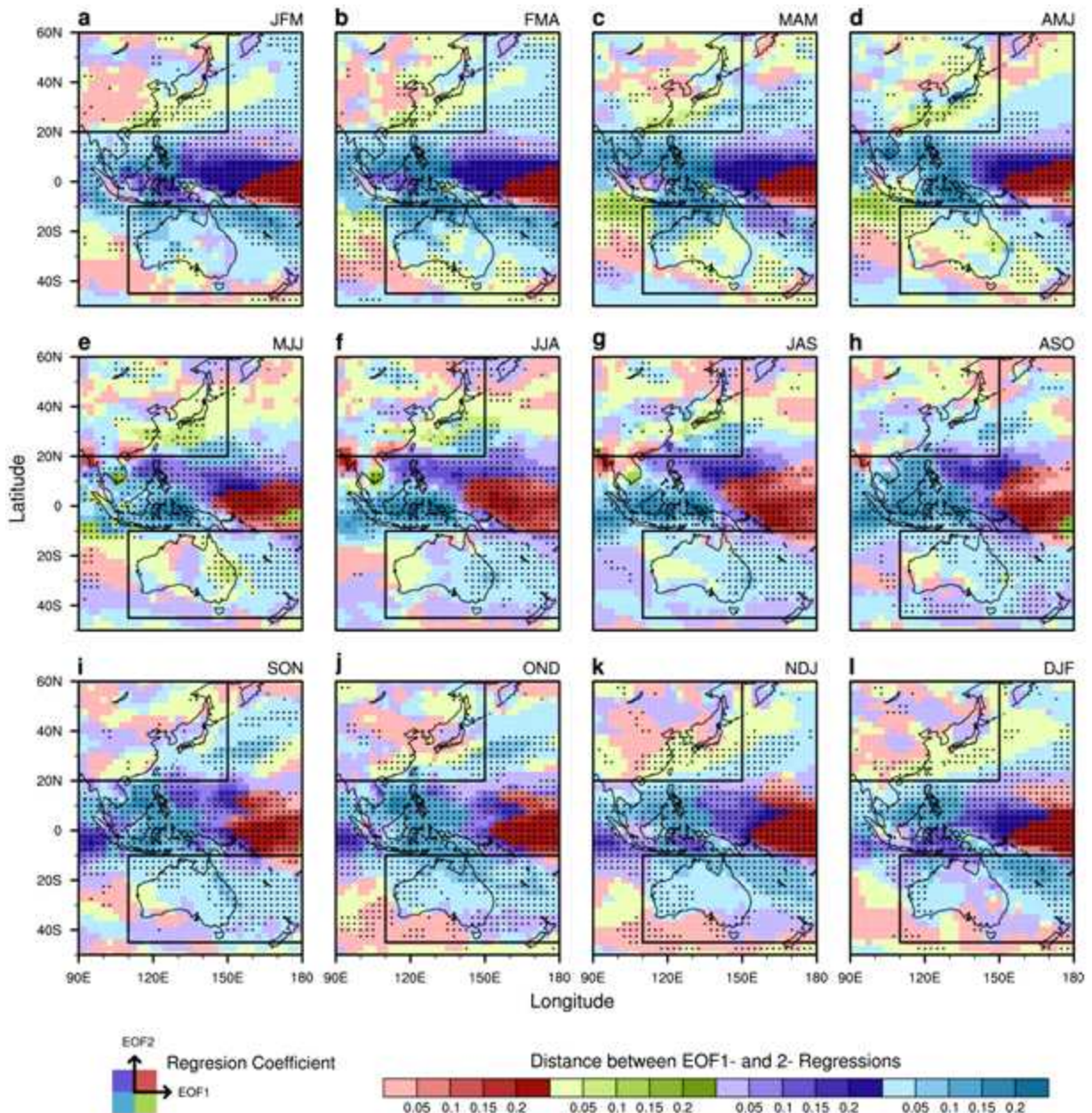


Figure 9. Relative signs and distance magnitudes (see the relevant coordinate and scale at bottom, respectively) on a 2-dimensional plane depicting regression coefficients of the seasonal mean anomalous precipitation onto the first two PCs for (a-l) JFM to DJF.

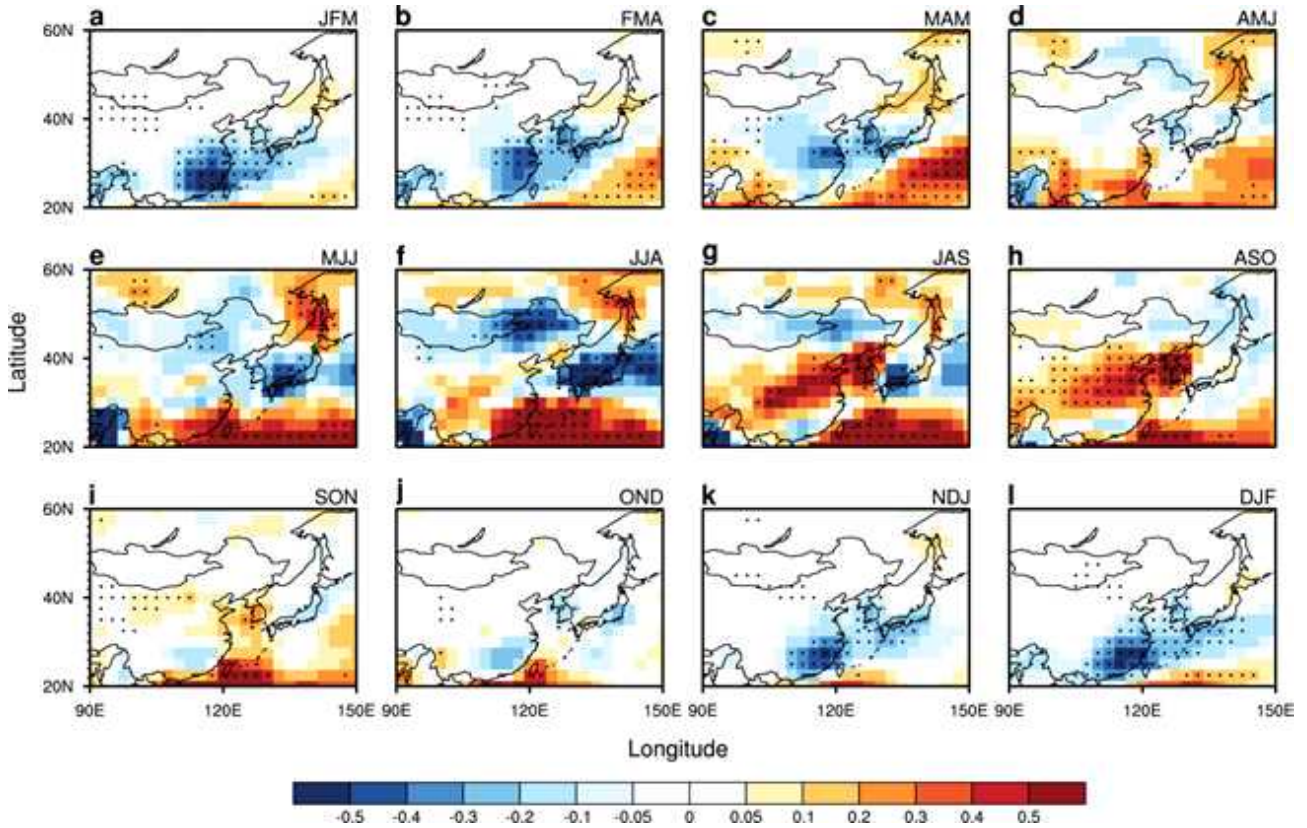


Figure 10. Regression of seasonal mean anomalous East Asian rainfall on tropical Western Pacific rainfall.

### (3) 지역특화 계절예측 기술 개발

다중 참여모델의 일반적인 아태지역 계절예측 성능 평가 및 특성 분석이 되었다. 지역/계절 별로 참여모델의 예측성능을 Anomaly Pattern Correlation Coefficient (ACC), MME Efficiency Index(MEI) 등을 기준으로 평가하고, 평가 결과를 바탕으로 아태지역 내 지역별 차별화된 MME 구성(model combination) 방안을 연구하고 적용 가능성을 평가하였다. 이를 통해 동아시아 지역 특화형 MME 구성을 통한 기후예측의 정확도 향상을 확인하였다.

## 다. 다양한 기후예측자료의 기후예측성능 검증을 통한 안정성 평가

### (1) 과거재현 자료의 통계적 안정성 확인

1차 년도의 연구에서는 ENSO 모드와 지역 강수의 관련성에 대한 재현성을 모델의 성능을 평가하고 새로운 MME의 서브셋을 위해서 성능 좋은 모델을 선별하기 위한 기준으로 사용하였다. 따라서 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술의 예측성능 안정성 평가를 위해 과거예측기간 (1982-2005년)에 대한 예측성능의 정상성(stationarity) 검증을 수행하였다. 과거예측기간 자료를 훈련기간(1982-1993, 12년)과 검증기간(1994-2005, 12년)으로

나누어 분석한 결과, 두 기간 모두 재구성된 앙상블 예측의 예측성능이 우수하고 특이점이 나타나지 않아 예측성이 안정적인 것을 알 수 있다(Figure 11 참조). 두 지역의 강수예측 성능의 전체 시계열은 지역이나 MME에 상관없이 어떠한 추이(trend)도 나타나지 않으며, Augmented Dickey-Fuller (ADF) test를 통한 유의성 검정 수행 결과 1% 유의 수준에서 통계적으로 안정적임을 나타낸다.

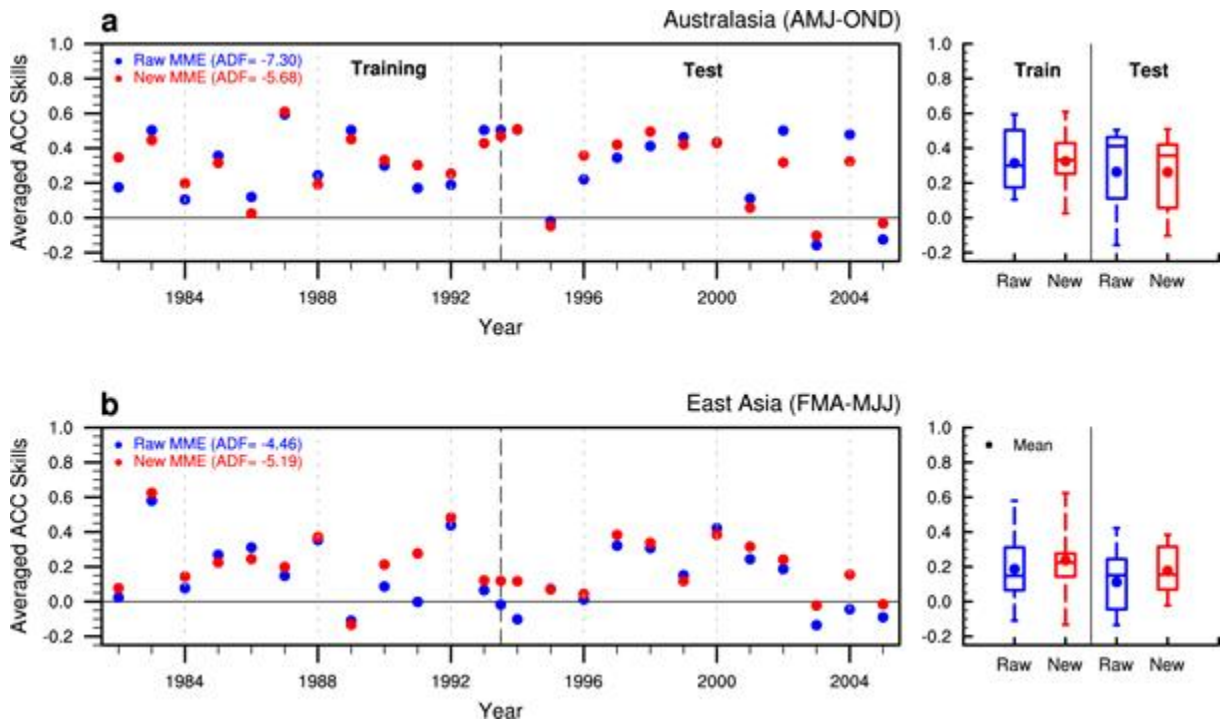


Figure 11. Time series (left panels) and their statistical distribution (right panels) of interannual variation of ACC skills averaged during corresponding AMJ-OND and FMA-MJJ seasons to (a) Australasian and East Asian precipitation predictions, based on training (1982-1993) and test (1994-2005) periods. Blue and red circles denote the raw MME based on all-inclusive models and new MME based on best performing models, respectively. Augmented Dickey-Fuller (ADF) tests are applied to test the stationarity of whole time series, and the relevant ADF statistics are provided in parentheses following the labelled targeted data sets of upper left of each panel. The statistical distribution of these time series for the raw and new MMEs are illustrated by box plots and mean values (circles), based on training and test periods.

## (2) 다양한 기후예측자료의 기후예측성능 검증

물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술의 예측성능 안정성을 평가하기 위해서 다양한 과거기후예측자료 및 실시간 기후예측자료 생산 하였다(Table 17 참조). 이를 바탕으로 예측 성능 검증 및 예측성 향상을 확인하였다.

Table 2. Three versions of hindcast data sets and experimental design.

구성 버전	사용 모델	과거 예측기간	K Fold	실험 설계
V2019	APCC, CMCC, JMA, MSC, NASA, NCEP, PNU, POAMA	1982-2005	23	과거 예측기간 차이
V2019*	상동	1991-2010	19	
V2020	APCC, BCCV2, CMCC_SPS3, GLOSEA5, JMA, MSC_CANSIPSV2, NASA, NCEP, PNU, BOM_ACESS-S1, UKMO	상동	19	사용 모델 차이

위 3.(다).(1)절의 통계 검정을 일반화하기 위해서, cross-validation을 통해 예측을 재평가하였다. 즉 원시 MME와 새 MME의 통계적인 정상성을 확인하고, 뿐만 아니라 임의적으로 생산된 성능, 과적합된 관련성, 임의의 모델 선택을 회피하기 위해서 표준 leave-one-out cross-validation method(Michaelsen, 1987; Jolliffe and Stephenson, 2012)이 아노말리 및 EOF 계산, 재현성 평가, 과거기후예측자료와 관측 사이의 예측성능 검증을 포함하는 전 과정에서 적용되었다. 과거기후 예측기간에 대한 평균 예측력이 교차검정 된 과거기후 예측기간의 예측성능 분포 범위 내에 존재하였다. 따라서 통계 분석 결과가 독립된 자료(independent data)에 대해서 일반화 할 수 있음을 알 수 있었다(Figure 12 참조).

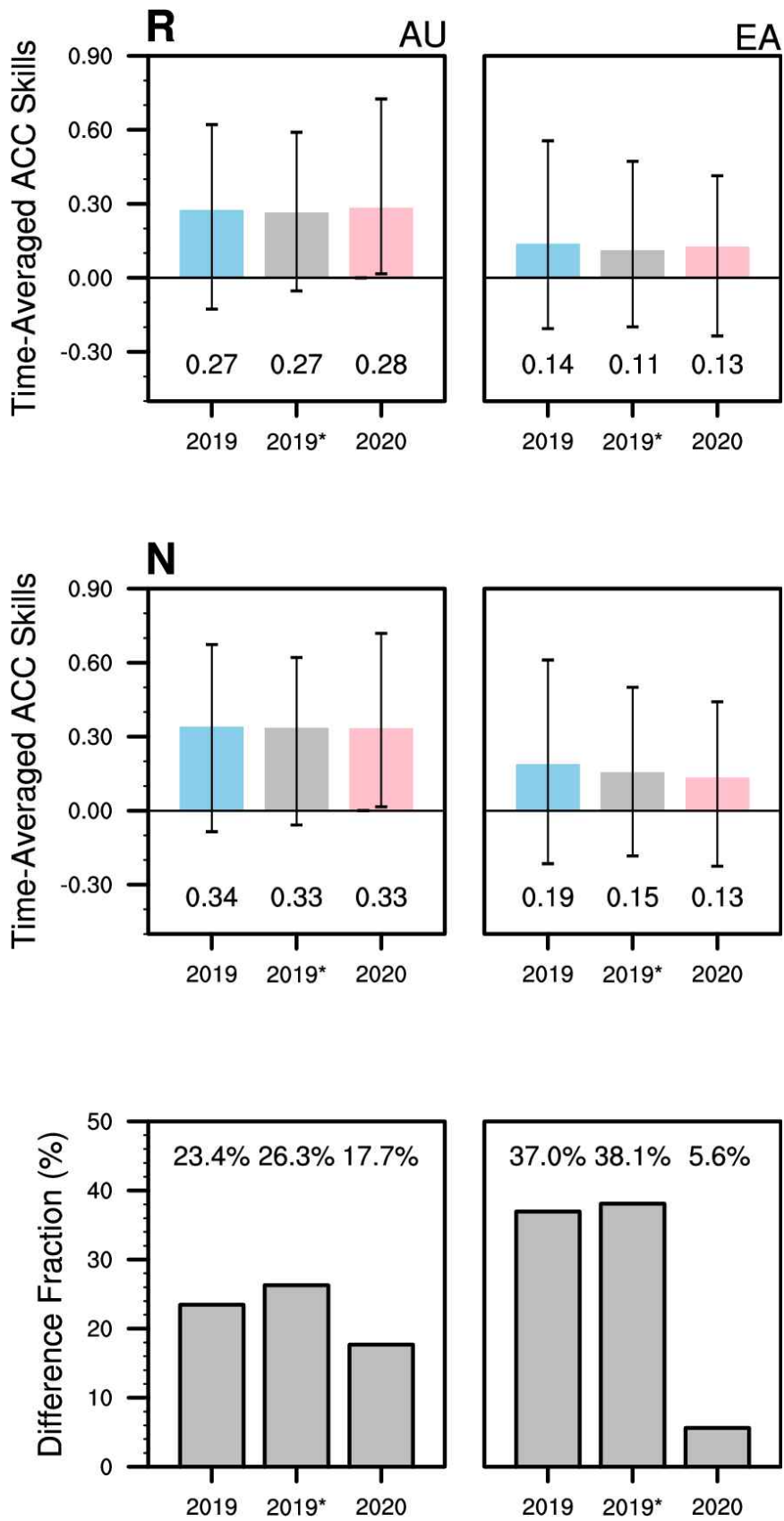


Figure 12. Time-averaged ACC skills of the two MME mean (upper and middle panels) Australasian (left panels) and East Asian precipitation predictions (right panels) and their difference fraction (%).

## 라. 모델 선별을 통한 다중모델기반 기후예측 기술 개발

### (1) ENSO 유형과 강수예측 성능에 기반 한 앙상블모델 재구성

동아시아 강수 변동에 영향을 줄 수 있는 주요 기후 인자인 ENSO와 서태평양 강수 변동을 각각 사용하여 모델 선별을 수행, 앙상블을 재구성하는 경우에 각 기후인자와 동아시아 강수 간 재현성에 대한 모델 분포가 작아 예측성능의 개선이 크지 않음을 알 수 있었다. 이에 Johnson et al. (2014)에서 이용된 방법과 같이 각각의 기후인자에 대해 재구성 두 MME를 합성하여 최종 앙상블을 구성하고자 한다. 아래는 ENSO 유형과 서태평양 강수 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 개선 결과를 일반적인 방법으로 (기후값 대비) 예측성을 검증한 결과(아래 3.라.(1).(가)절 참조)와 앙상블모델 재구성을 통한 MME 예측성능을 이전과 상호 비교한 결과(아래 3.라.(1).(나)절 참조)를 보여준다.

#### (가) 기후값 대비

ENSO 유형 및 열대 태평양 강수의 예측성능을 이용한 다중모델기반 동아시아 기후예측 기술을 개발하였다. Figure 13은 과거재현 자료의 동아시아 강수량에 대해 계산된 HSS와 not EC coverage의 확률밀도 함수를 가용한 모든 모델로 구성된 앙상블(P0, 검은색)과 선별된 모델로 구성된 앙상블(P3, 녹색)에 대해서 나타냈다. P0와 P3를 비교할 때, HSS와 not EC coverage의 품위(performance) 분포가 P0보다 P3에서 모두 개선된 것을 알 수 있다.

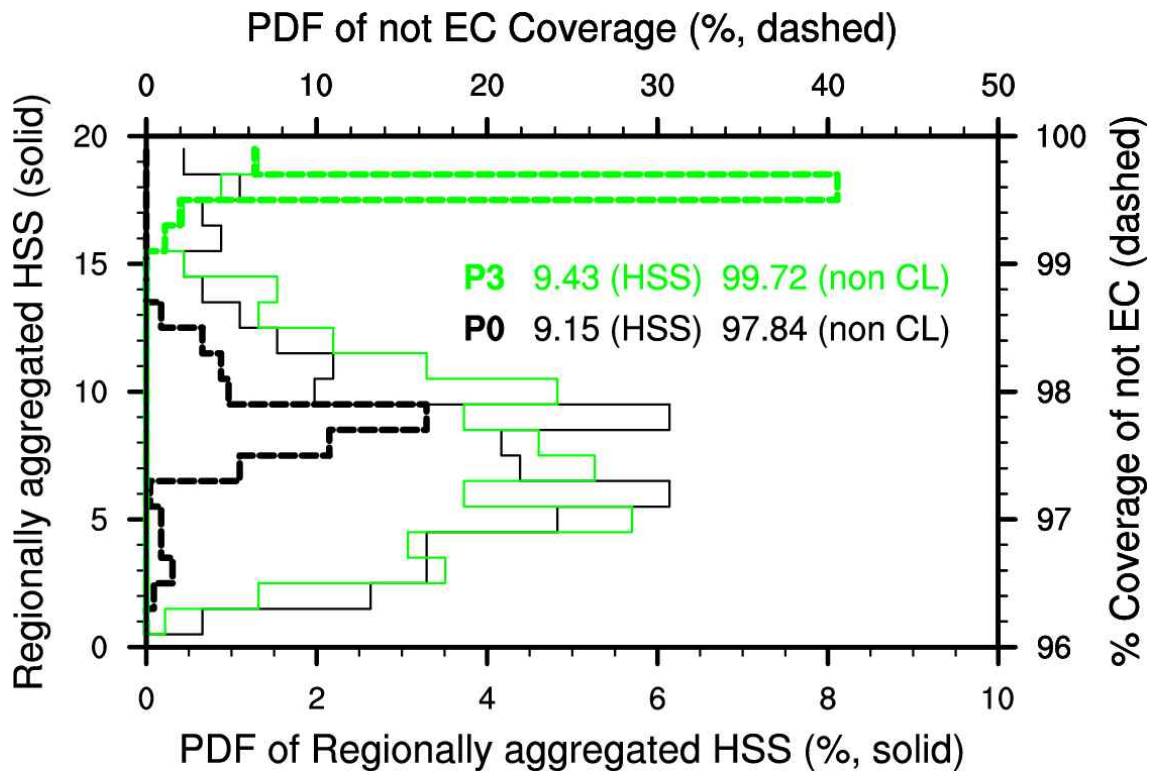


Figure 13. Probability distribution functions of regionally aggregated HSS (%, solid) and not EC coverage (%, dashed).

(나) 표준 예측 시스템 대비

Figure 14은 표준 예측 시스템(여기서는 P0) 대비 계산된 P3의 HSS의 밀도확률함수를 나타낸다. HSS가 0인 경우는 P0와 P3의 성능이 같음을 나타낸다, 음의 값은 P0가 P3보다 성능이 우세한 경우를 나타내며, 반대로 양의 값은 P3가 P0보다 성능이 우세한 경우를 나타낸다. 전체 예측된 값 중 60.53%가 예측성능이 개선된 경우이며, 39.47%가 예측성능이 개선되지 못한 경우를 나타낸다. 평균적으로 P3가 P0에 비해 예측이 개선되었으며, 이는 1,000개의 부트 스트랩을 수행할 경우 95% 신뢰구간에서 통계적으로 유효한 값이다.

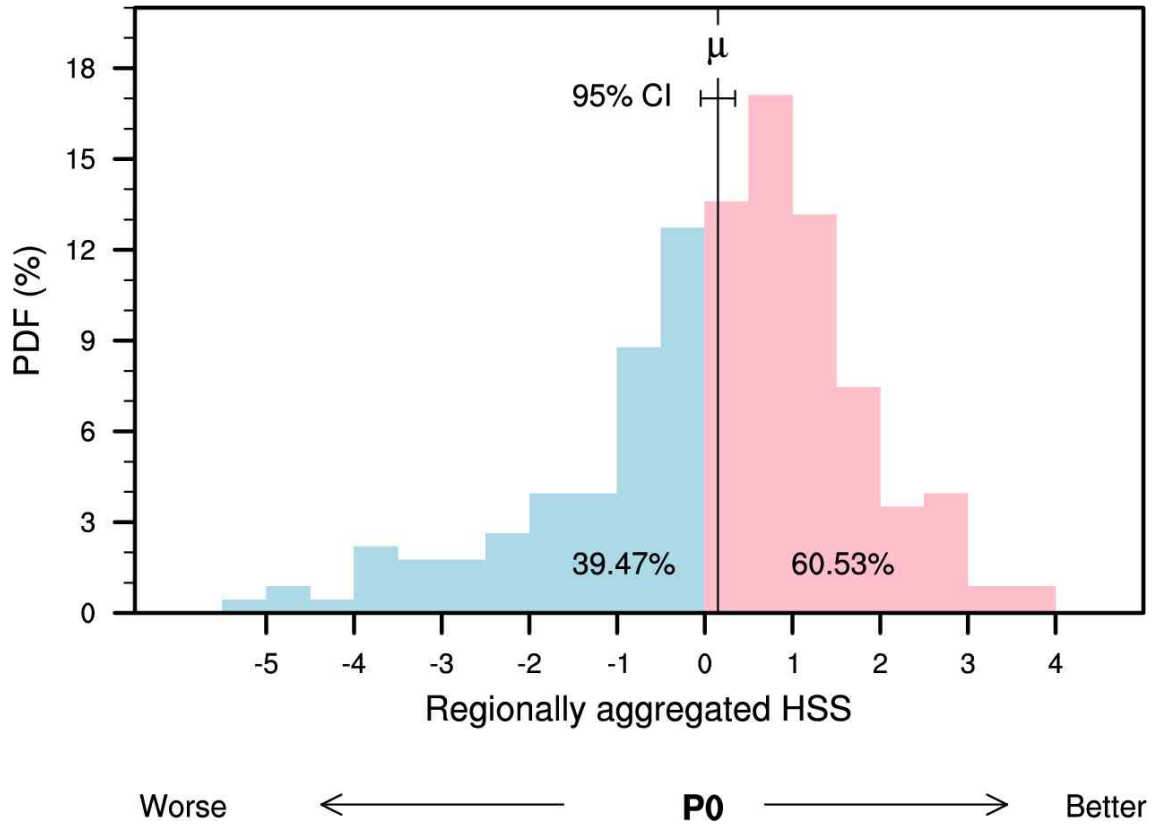


Figure 14. Probability distribution functions of regionally aggregated HSS (%).

선별된 모델을 바탕으로 한 앙상블은 기존의 가용한 모델을 바탕으로 한 앙상블에 비해, 기 후값 대비(Figure 13 참조) 또는 표준 예측 시스템 대비(Figure 14 참조) 향상된 예측 성능을 보여 줌을 알 수 있었다. 따라서 아래 3.라.(2)절에서는 예측성의 원천을 알아보기로 한다.

## (2) predictability source 및 앙상블 재구성을 바탕으로 한 MME의 한계 파악

### (가) ENSO와 관련된 예측성의 시간 변동성 분석

예측성의 원천을 알아보기 위해서, ENSO와 관련된 예측성의 시간 변동성을 분석하였다. Nino 3.4에 대한 HSS의 시간적인 반응을 살펴보면, 선별된 모델을 바탕으로 한 앙상블은 기존의 가용한 모델을 바탕으로 한 앙상블에 비해, ENSO와의 관련성이 더 컸다(Figure 15 참조). 또한 엘니뇨보다는 라니냐이거나 중립일 경우의 예측성능이 훨씬 높았다. ENSO의 복잡성을 기준으로 HSS의 시간적인 반응을 살펴보면, P0보다 P3에서 예측성이 높은 요인은 중동태평양의 warming/cooling보다 그것의 zonal shift가 더 큰 역할을 하는 것을 알 수 있다(Figure 16 참조).

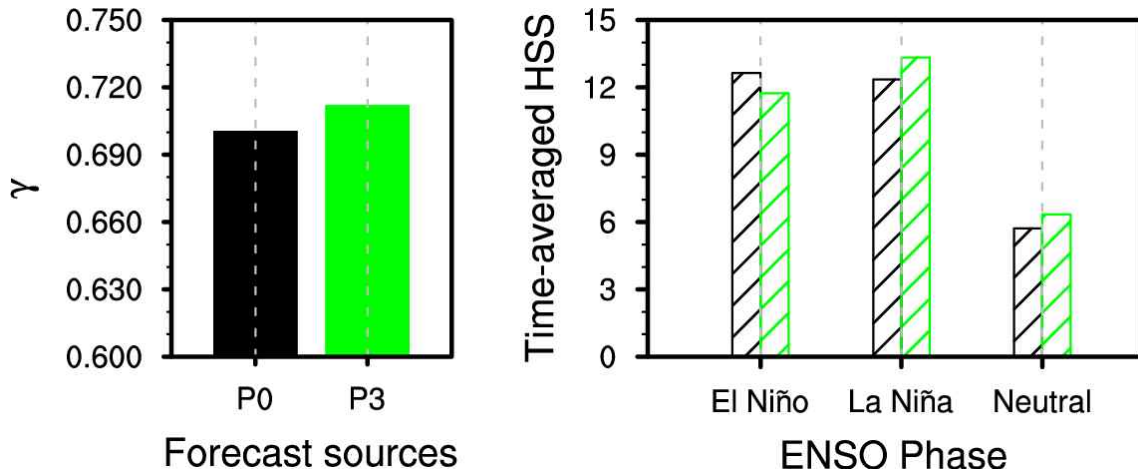


Figure 15. Skill variation in response to Nino 3.4.

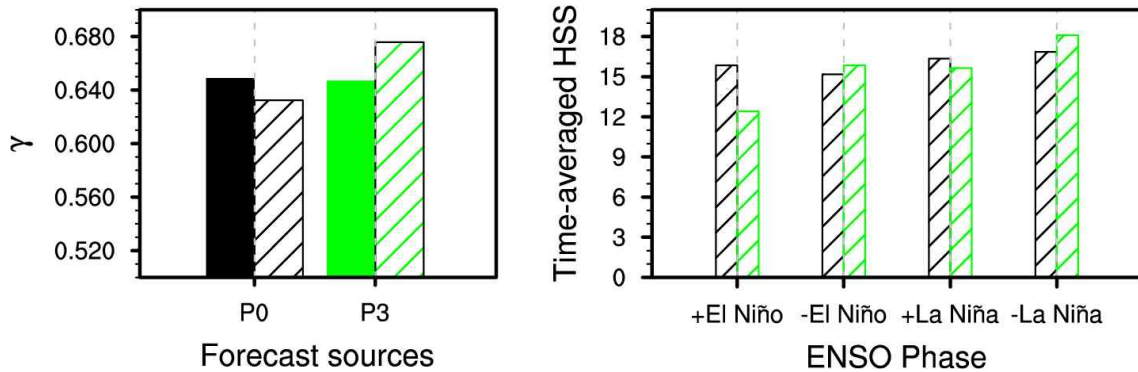


Figure 16. Skill variation in response to ENSO.

#### (나) 앙상블 재구성을 바탕으로 한 MME의 한계 파악

Figure 17은 전체 격자에 대한 각 above와 below 카테고리의 비율(%)의 시간적인 변동성을 나타낸다. HSS는 결국 각 카테고리에 대해서 전체 격자수 중 기후값 대비 예측한 카테고리의 hit 수를 나타내는 것으로, 각 카테고리에 대해서 관측과 같은 비율을 가질수록 높은 예측 성능을 가진 것으로 간주된다. 각 카테고리에 대해서 관측과 좀 더 가깝거나 유사한 시계열을 보일수록 이는 높은 예측성으로 귀결됨을 알 수 있다. 따라서 기존의 가용한 모델을 바탕으로 한 앙상블에 비해, 선별된 모델을 바탕으로 한 앙상블은 관측값과 좀 더 근사한 시계열을 보이면서 결국 예측성능의 향상을 보인다. 하지만, P와 P3의 차이가 크지 않으며, 관측 시계열과 다소 유사성이 떨어지는 것을 알 수 있다. 이는 곧 P0이든 P3이든 모델의 강수 예측성 자체에 의해 제한됨을 알 수 있다. 즉, 두 시스템은 동일 시스템으로 태생적으로 예측성의 한계를 지님을 알 수 있다.

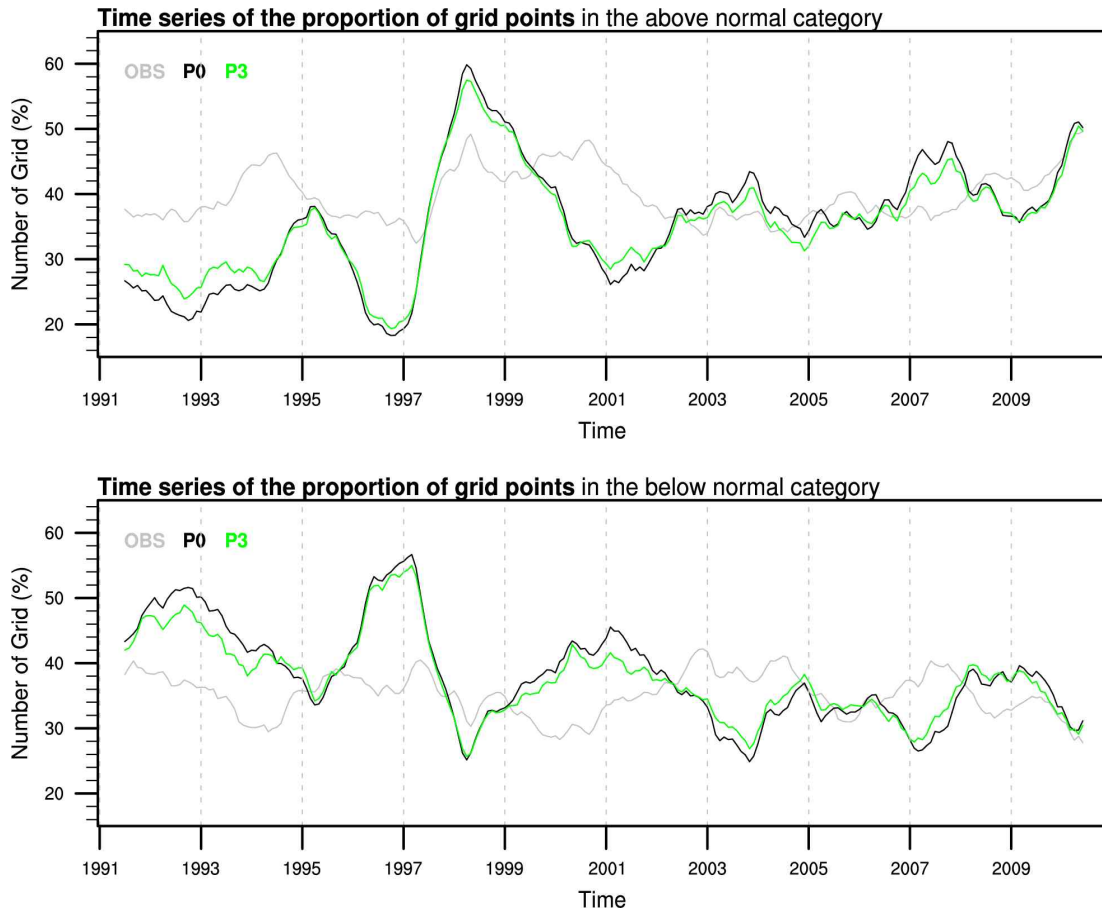


Figure 17. Temporal variations of the proportion of grid points in the above normal category (upper panel) and below normal category (bottom panel).

## 마. 확률예측 기법 개발

동아시아 지역맞춤형 계절기후전망을 위한 물리기반 모델선별 앙상블의 효용성을 다양한 측면에서 확인하였으며, 이에 동아시아 지역 맞춤형 계절 확률 예측에 적합한 계절 확률 예측 기법을 분석하고 반영하고자 한다. 다중모델을 활용한 확률예측의 경우 기관마다 상이한 계절 확률예측기법을 사용하고 있다. 즉 Table 3과 같이 확률예측을 위한 요소 및 조합을 기관마다 다양하게 구성하고 있다. 크게 확률예측을 위한 요소는 보정, 확률추정, 합성의 과정 및 단계를 거치는데, 그 순서와 구성은 다양하게 조합할 수 있다. APCC와 WMO에서 사용하는 확률예측 방법은 어떤 보정도 하지 않은 원시 자료로부터 모수화 방법을 활용하여 개별 모형의 확률 분포를 추정하고 여기에 가중치를 두어 합성한다. 반면에 NMME의 경우 개별 모델별로 기후값을 보정한 후 모수화(또는 비모수화) 방법으로 확률 분포를 추정한 후 전체를 같은 샘플에서 넣어서 확률을 계산한다. 앞절에 보인 결과는 보정한 자료를 비모수화 방법으로 전체 샘플에서 추정한 방법이다. 따라서 이 절에서는 각 단계별로 어떤 방법이 나은지 살펴보고, 특히 모델 선별에 적합한 방법을 찾아서 확률예측을 위한 요소를 구성하고자 한다.

Table 3. Several factors for the probability forecast of APCC, NMME, and C3S.

	APCC (WMO)	NMME	C3S	'2020
보정 (Calibration)	Uncalibrated	Calibrated	-	Calibrated
추정 (Estimation)	Parametric	Parametric Non-parametric	Non-parametric	Non-parametric
합성 (Combination)	Weighted	In model & pooled	pooled	In grand & pooled

먼저 기후학적 평균의 보정의 효능에 대해 살펴보았다. APCC에서 현재 현업으로 활용하고 있는 “모수화+가중치” 기법과 “순위+풀링” 기법에서의 보정의 차이를 살펴보았다. Welch  $t$  검정과 K-S 검정 수행 결과 보정 전과 보정 후의 자료들은 통계적으로 자료 간 서로 독립적이다. 추정 및 합성 방법에 무관하게, 보정은 정확도(Accuracy)와 분해능(Resolution) 측면에서 예측성을 향상, 선명도(Sharpness)를 개선(Figure 18 참조)시킴을 알 수 있었다.

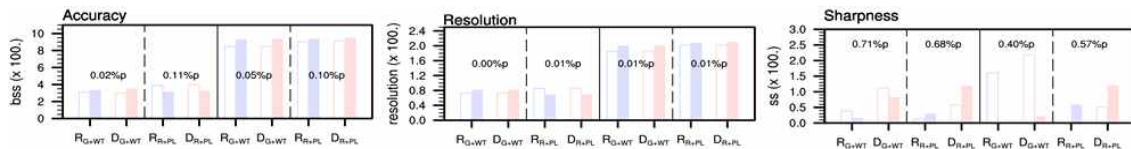


Figure 18. Aggregated accuracy, resolution, and sharpness in time and space for the raw and debiased probability forecasts of precipitation and temperature.

다음으로 추정과 합성에 적합한 방법에 대해서 평가하였다. Figure 19에서 알 수 있듯이, 모수화는 가중치, 순위는 풀링 방법이 높은 예측성을 보였다. 전반적으로 정확성, 신뢰도, 분해능 측면에서는 순위 방법이 높은 예측성을 보였다. 또한 강수와 기온 모두에서 풀링 방법이 높은 예측성을 나타내는 것을 알 수 있었다. 추정과 합성의 효능을 전체 가용 앙상블과 모델 선별 선별 앙상블에 대해서 살펴보면, 강수는 모델 선별 앙상블 후 정확도, 신뢰도, 분해능에서 예측성이 개선, 기온은 모델 선별 앙상블 후 신뢰도에서 예측성이 개선되었다. 모델 선별 앙상블에서는 풀링 방법이 여타 방법에 비해 높은 예측성을 보였다.

따라서 현재 APCC의 전체 가용 앙상블은 “무보정+모수화+가중치” 기법을 바탕으로 확률 예측자료를 생산하고 있지만, 위 분석을 종합적으로 판단하여 모델 선별 앙상블은 “보정+순위+풀링” 기법을 채택하여 확률예측 시스템을 개발 및 구성하였다.

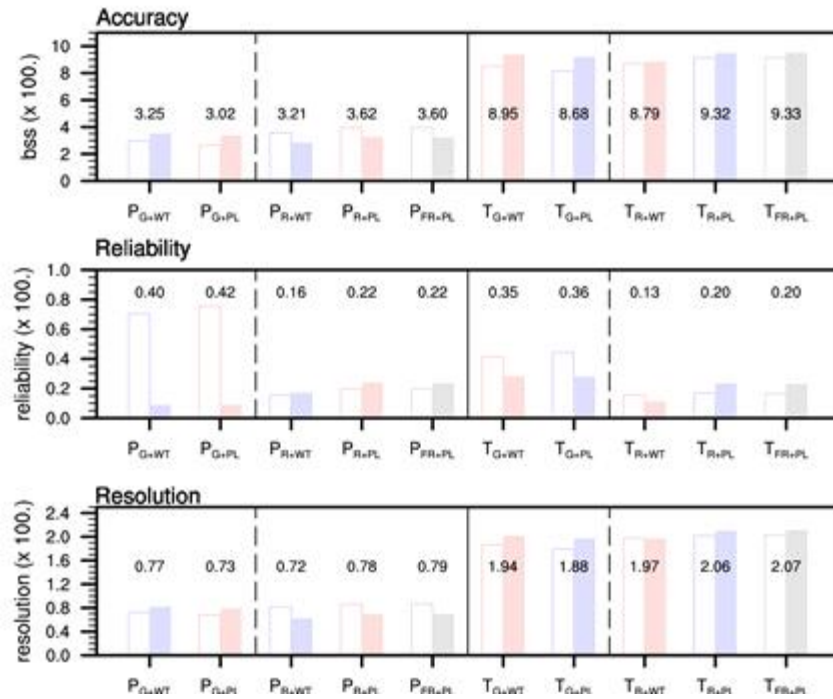


Figure 19. Aggregated accuracy, reliability, and resolution, in time and space for probability forecasts of precipitation and temperature, based on various combination of PDF estimation and composite.

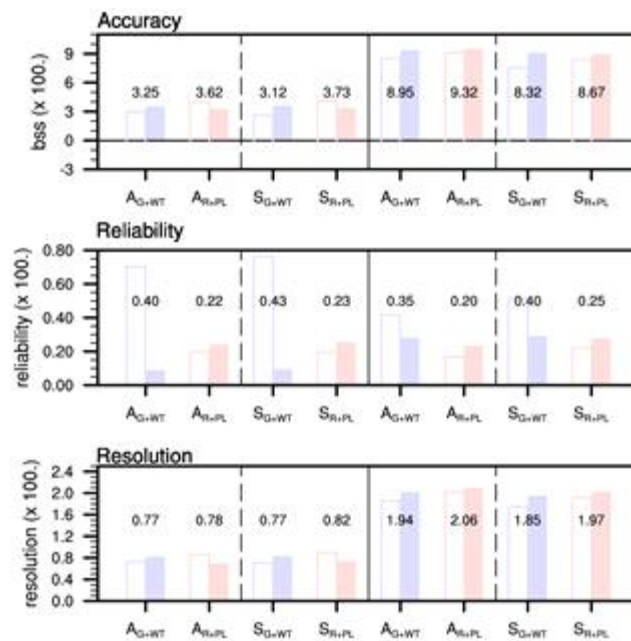


Figure 20. Aggregated accuracy, reliability, and resolution, in time and space for probability forecasts of precipitation and temperature, based on all available and newly selected models.

## 바. 동아시아 맞춤형 확률예측시스템 구축

위의 결과들을 종합적으로 반영하여 물리 기반의 모델 선별 앙상블은 “보정+순위+폴링” 기법을 채택하여 확률예측 시스템을 개발 및 구성하였다. 단 시스템은 자동화를 위하여 step-wise 확률예측 시스템으로 개발하였다(Table 4 참조). 즉 안정적인 현업을 위해서 각각의 소스 코드는 각각의 예측 및 과거기후예측을 수행할 때 어떤 수정도 하지 않도록 모듈화하여 간 단계를 설계 및 구성하였다. 시스템은 0번부터 5번까지 총 6개의 단계로 구성되었으며, 짝수(0, 2, 4) 단계는 과거기후예측 자료를 다루는 reference 추정, 모델 선별, 기후학적 삼분위 추정의 작업을 수행한다. 반면에 홀수(1, 3, 5) 단계는 실제 기후예측을 위한 단계로 바이어스 보정, 기후인자의 결정론적 예측, 강수 및 기온의 확률예측을 위한 작업을 수행한다. 여기에서 2, 3, 5 단계의 결과물은 다음 절의 표출 시스템과 연계되도록 설계 되었다.

Table 4. Step-wise probability forecast system developed in the study.

Step 0	1	2	3	4	5
<b>Hindcast Preprocessing and Bias Correction</b>	<b>Forecast Preprocessing and Bias Correction</b>	<b>Skill Assessment of Regression Field and Model Selection</b>	<b>Forecast of Climate Driver</b>	<b>Climatological Tercile Estimation and Model Combination</b>	<b>Probability Prediction</b>
Data Retrieval Seasonal Mean Ens Mean Ens spread Climatology Bias Corrected	Data Retrieval Seasonal Mean Ens Mean Ens Spread Bias Corrected	ENSO Definition TWP Definition Regression Skill NMME of SST	NMME of SST ENSO Idx ENSO Complexity NMME of Prec	Climatological Tercile	Calculation of Observed category Calculation of Probability Combination of Probability
Reference Creation	Bias-Removed Calibration	Model Selection	Deterministic Prediction	Reference PDF Estimation	Probabilistic Prediction

## 사. 동아시아 맞춤형 확률예측을 위한 표출시스템 구축

### (1) 시스템 및 프로그램의 구성

가시화 시스템에 사용된 언어는 다중 언어 사용을 통한 복잡한 이해관계를 해결하고자 python version 3만을 사용하여 생산할 수 있도록 구성되었다. APCC의 현업시스템과의 연계를 위해 HPC의 python/3.7-anaconda-res를 사용하여 모두 생산할 수 있도록 최적화하였다. python 언어를 사용하기 위해서는 여러 부가적인 python 패키지가 요구되는데, 현 시스템에서는 일반적으로 많이 사용되는 pandas, numpy, statistics, math, xarray, datetime, matplotlib, seaborn, os, scipy, basemap, cartopy, sklearn 패키지를 사용하였다. 동아시아 지역맞춤형 표출시스템은 “1. 동아시아에 대한 기후전망 2. 기후전망에 이용된 예측인자의 전망 3. 최근 동안의 지난 예보 및 관측의 추이 4. 예측인자와 동아시아 기후의 관련성 5. 최근 동안 지난 예보의 검증 6. 과거 재현기후 예보의 검증” 에 대한 13개의 python 코드와 PDF 작성을 위한 한 개의 python 코드 그리고 이를 총괄하는 코드로 구성되어있다. 총괄하는 코드를 통해 효율적인 관리 시스템을 구성하고자 하였다. 총괄 python 코드의 실행을 통해 입력한 해당 월에 생산된 지역맞춤형

예측정보 및 관측자료를 이용하여 동아시아 예측정보 및 검증정보가 일괄 생성되도록 설계되었다. Figure 1은 확률예측 가시화 시스템에 대한 구성을 나타낸다.

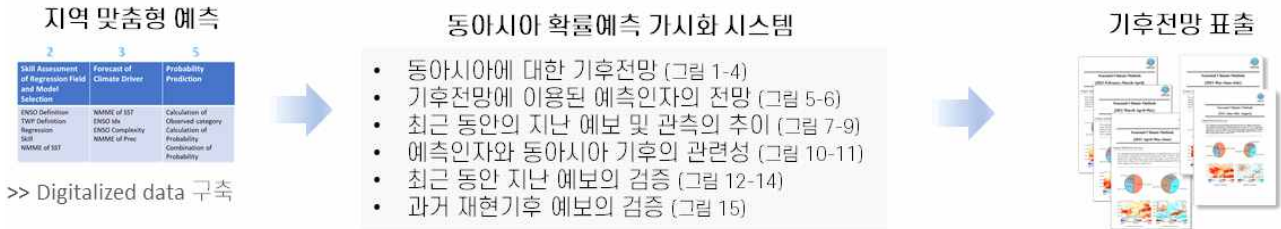
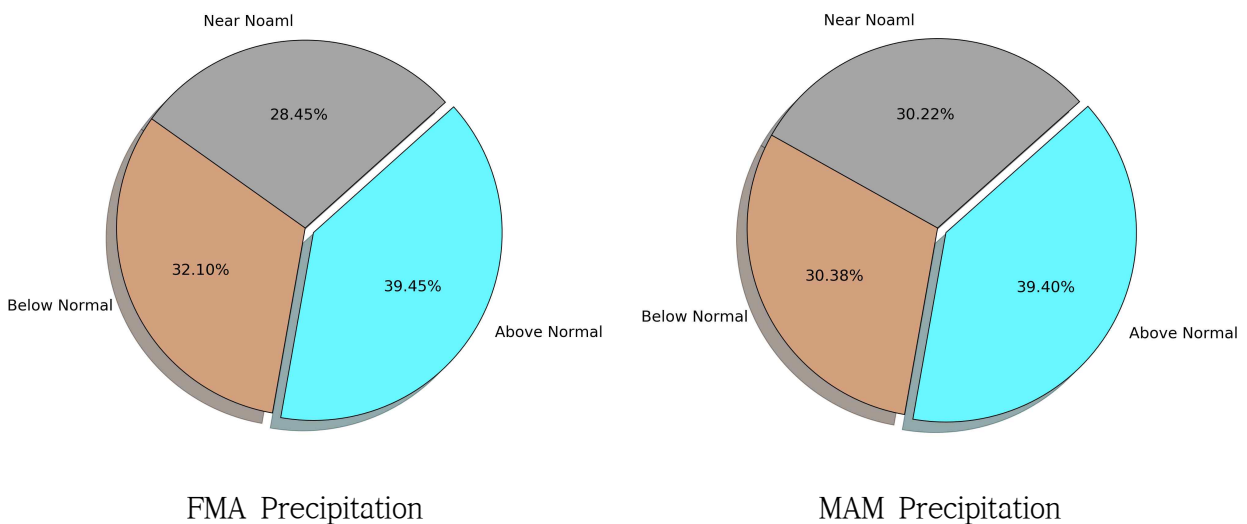


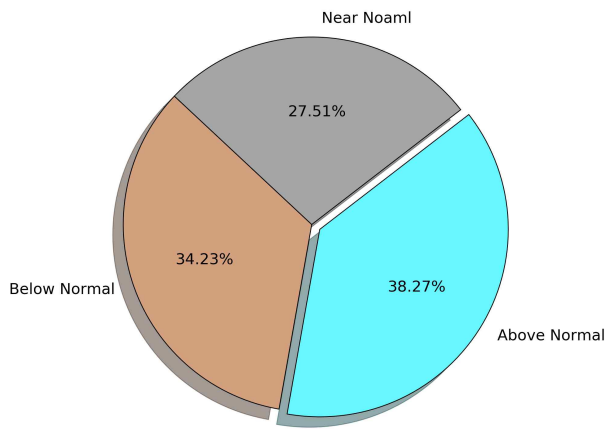
Figure 21. Schematic diagram of the probability prediction display system.

## (2) 동아시아 확률예측의 가시화

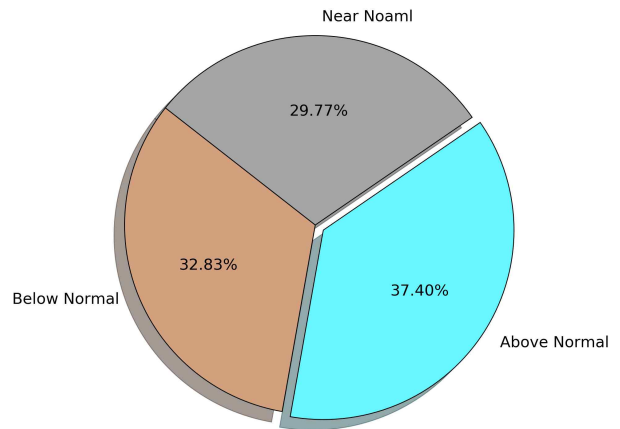
### (가) 동아시아에 대한 기후 전망

동아시아에 대한 기후 전망은 PMME를 통한 지역 전체의 확률분포와 동아시아 전체의 확률을 통해 표출한다. 지역 전체 PMME 전망은 Pie Chart를 통해 정보를 제공하고, 격자형 PMME 전망을 통해 동아시아 지역 전반에 대한 정보를 제공하였다. 동아시아에 대한 기후 전망의 정보는 강수와 기온에 대한 정보를 매일 제공하도록 설계하였다. Figure 22은 2021년 FMA 부터 NDJ 까지의 동아시아 지역 전체 PMME 강수량의 전망을 나타내고 Figure 23은 같은 기간의 동아시아 지역 전체 PMME 기온의 전망을 나타낸다. Figure 24은 같은 기간의 격자형 PMME 강수 전망에 대해 나타내며, Figure 25은 같은 기간의 PMME 기온 전망에 대해 나타낸다.

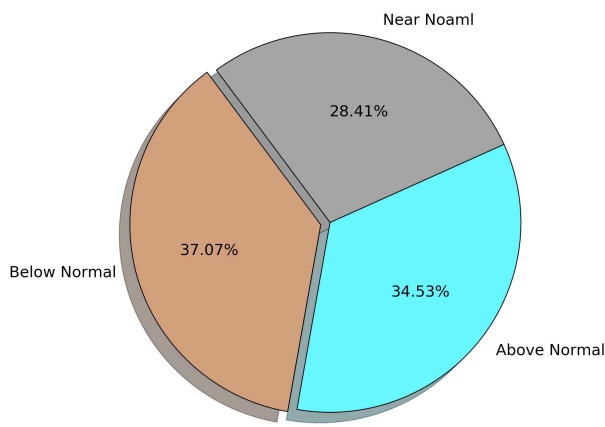




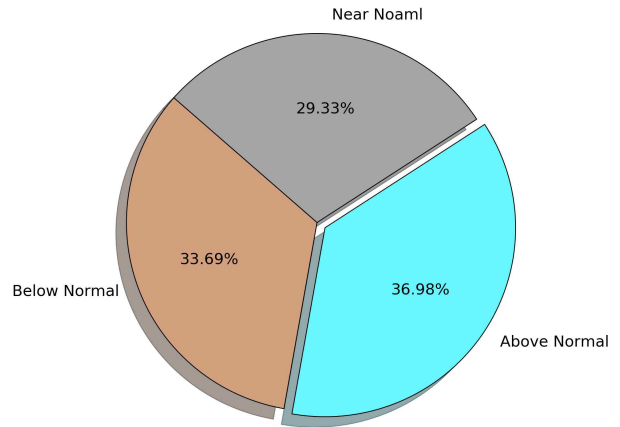
AMJ Precipitation



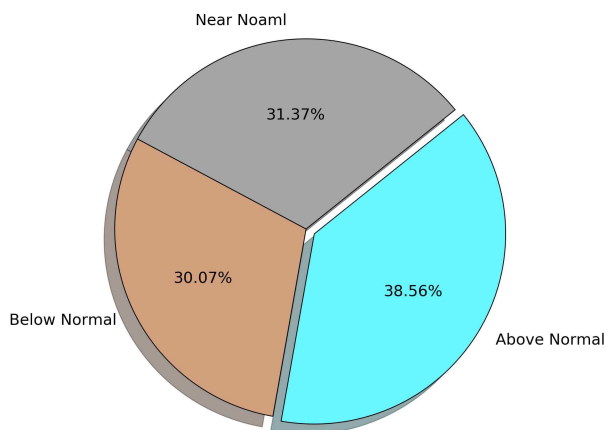
MJJ Precipitation



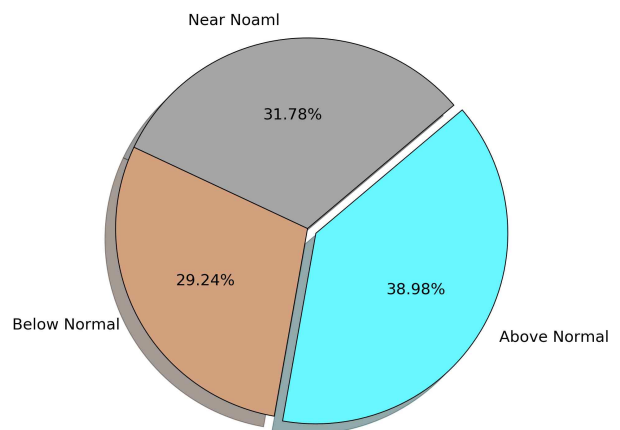
JJA Precipitation



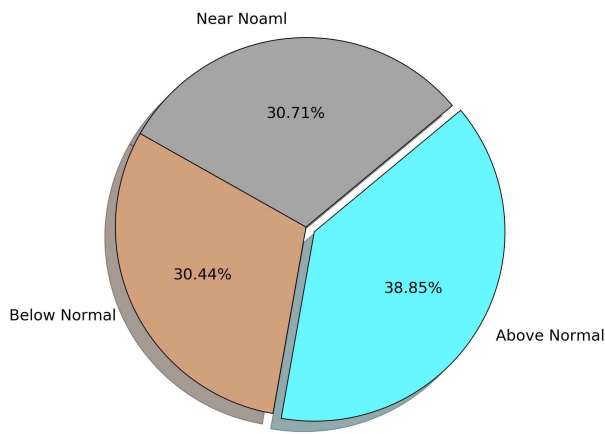
JAS Precipitation



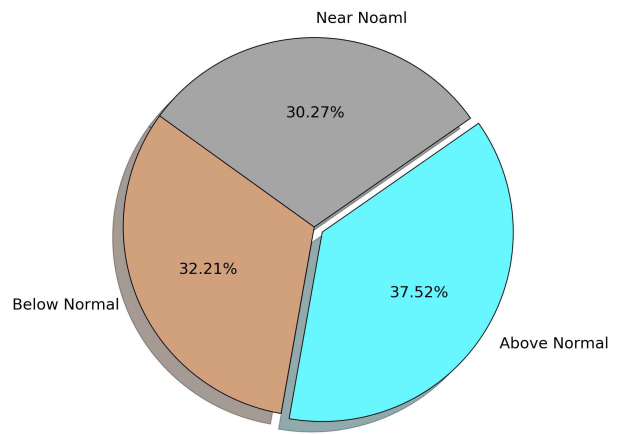
ASO Precipitation



SON Precipitation

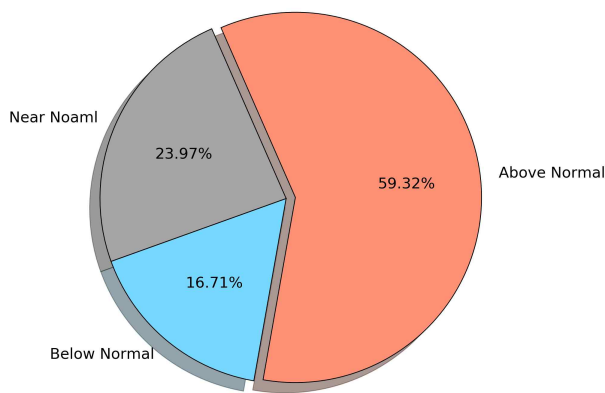


OND Precipitation

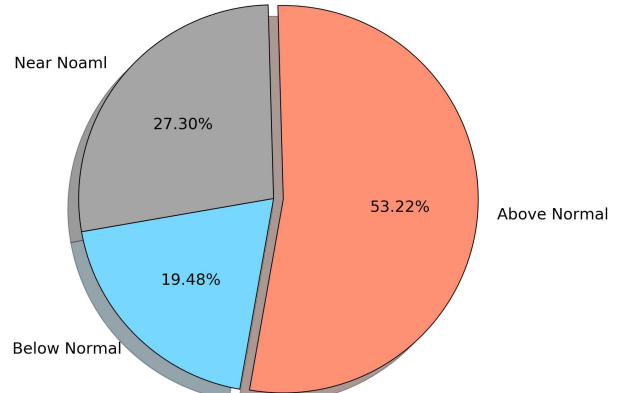


NDJ Precipitation

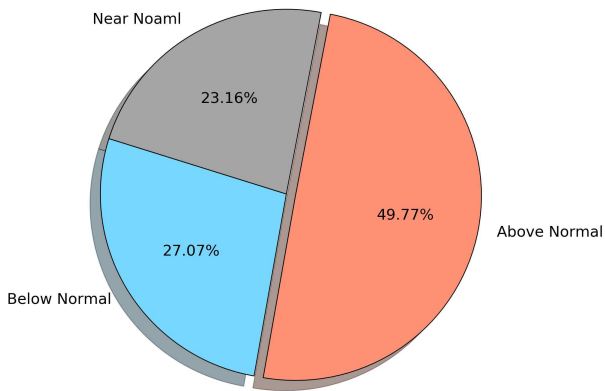
Figure 22. Overall PMME precipitation forecast in East Asia.



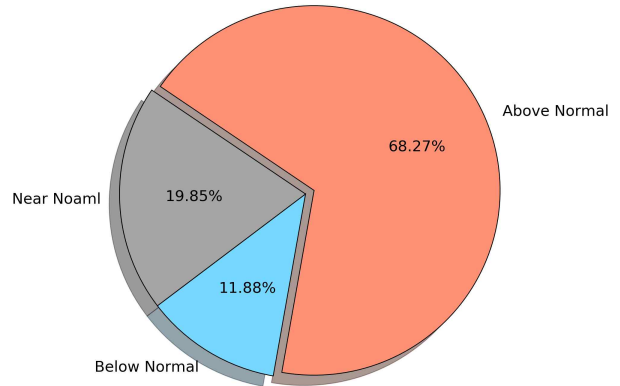
FMA Temperature



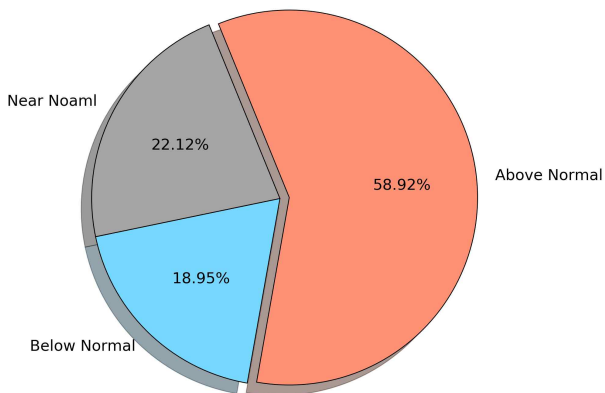
MAM Temperature



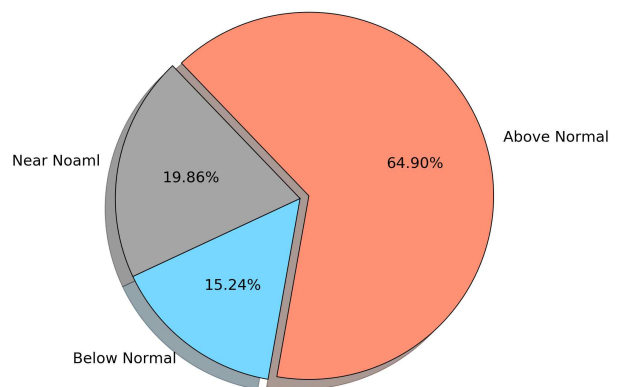
AMJ Temperature



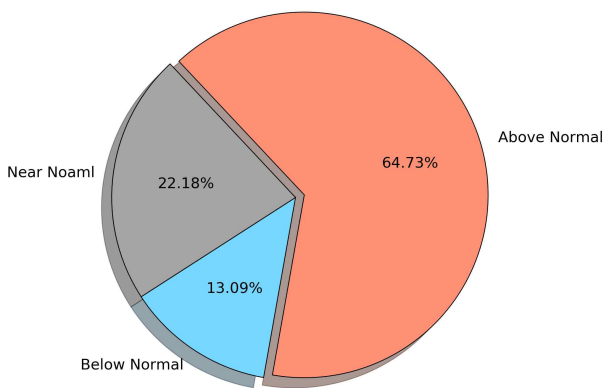
MJJ Temperature



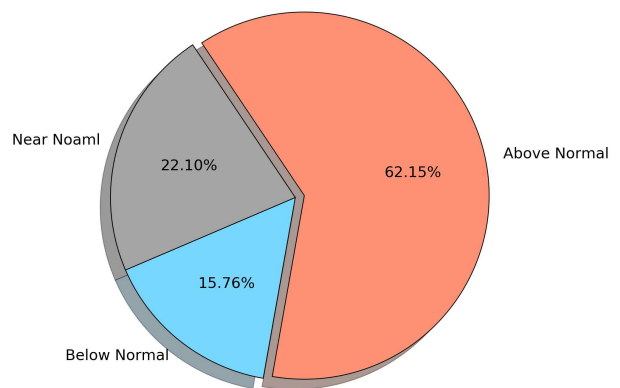
JJA Temperature



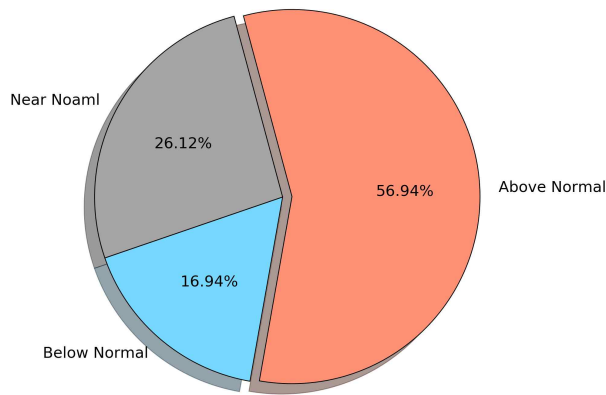
JAS Temperature



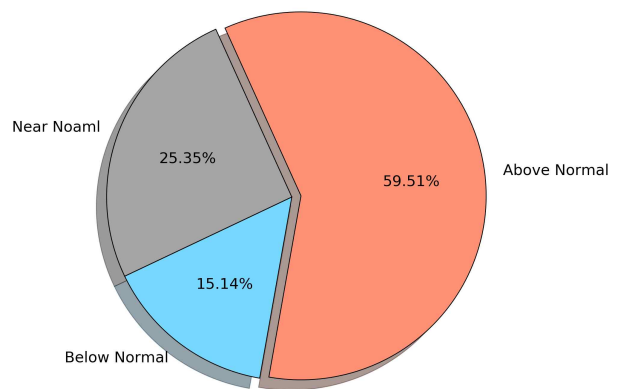
ASO Precipitation



SON Precipitation

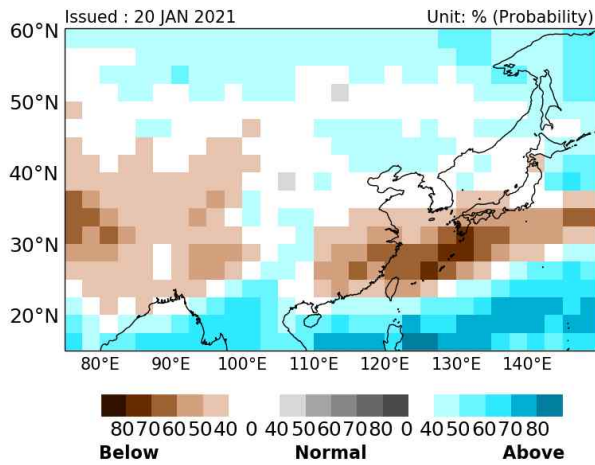


OND Precipitation

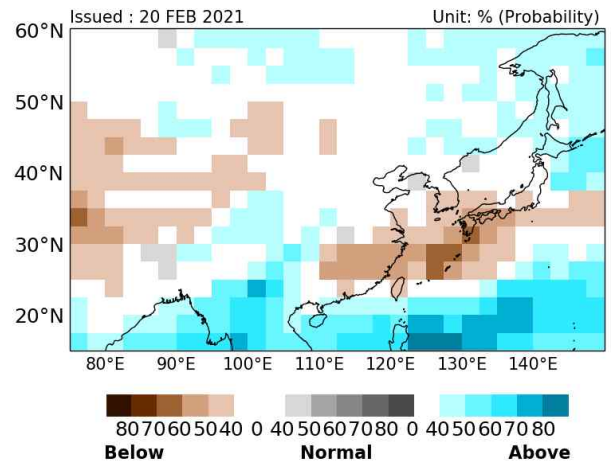


NDJ Precipitation

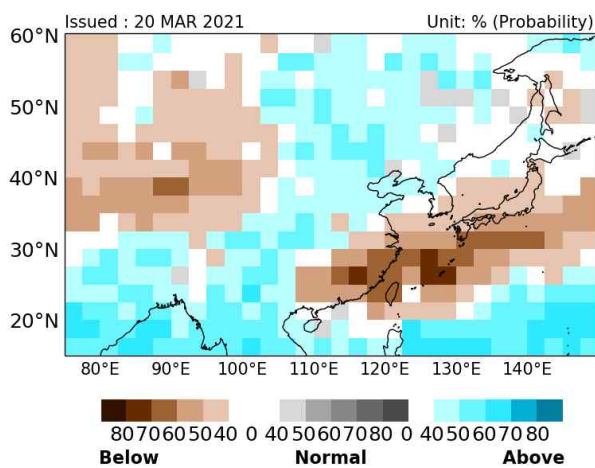
Figure 23. Overall PMME Temperature forecast in East Asia.



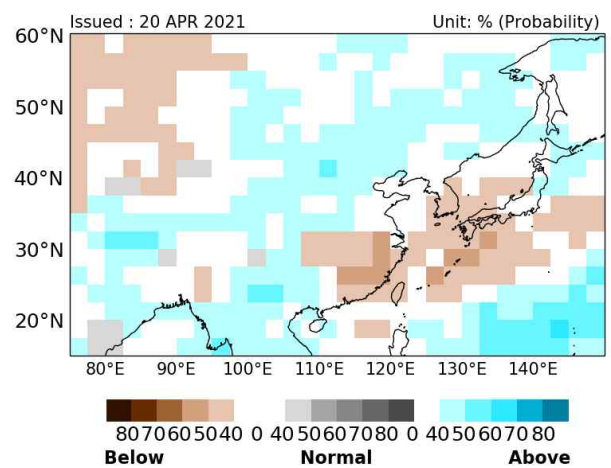
FMA Precipitation



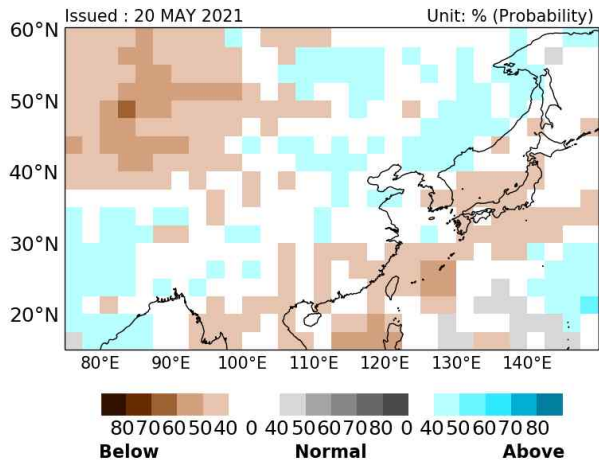
MAM Precipitation



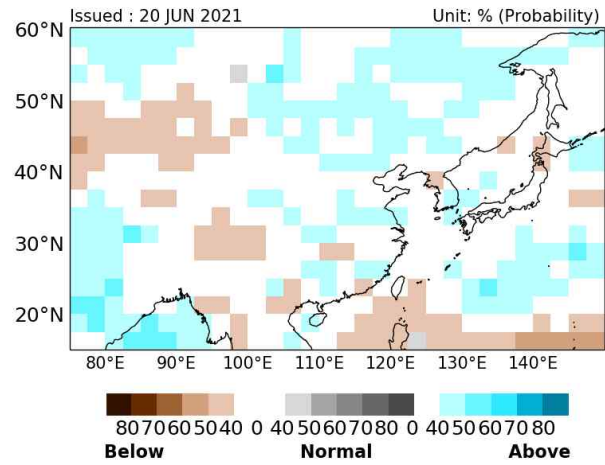
AMJ Precipitation



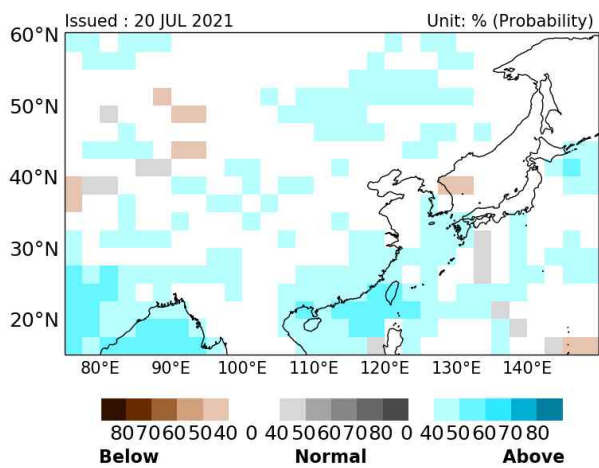
MJJ Precipitation



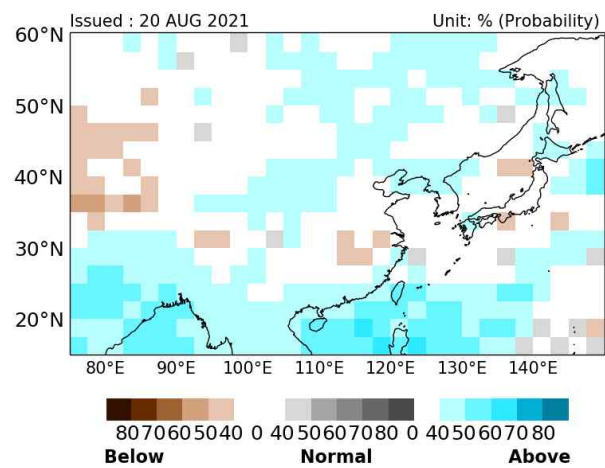
JJA Precipitation



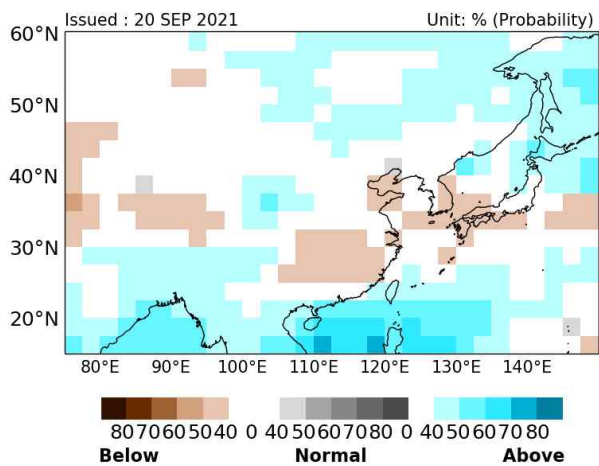
JAS Precipitation



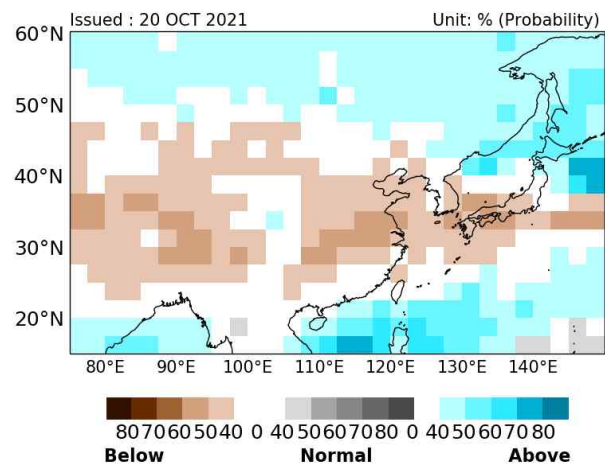
ASO Precipitation



SON Precipitation

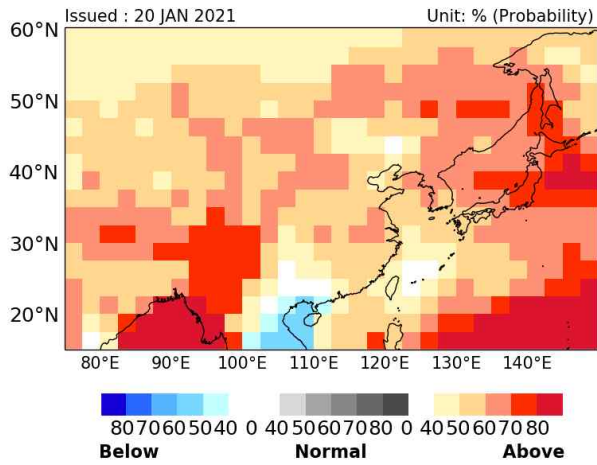


OND Precipitation

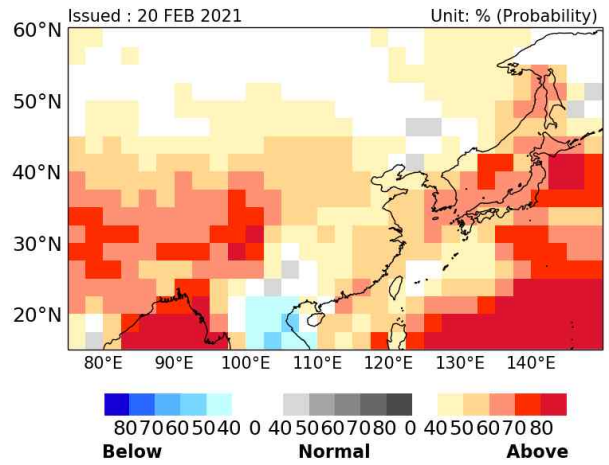


NDJ Precipitation

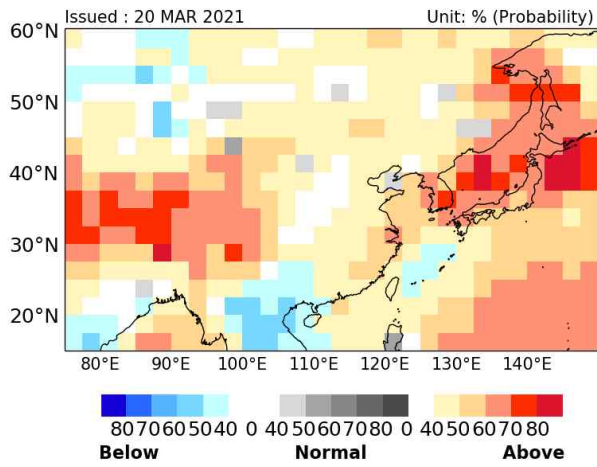
Figure 24. Grid PMME Precipitation forecast.



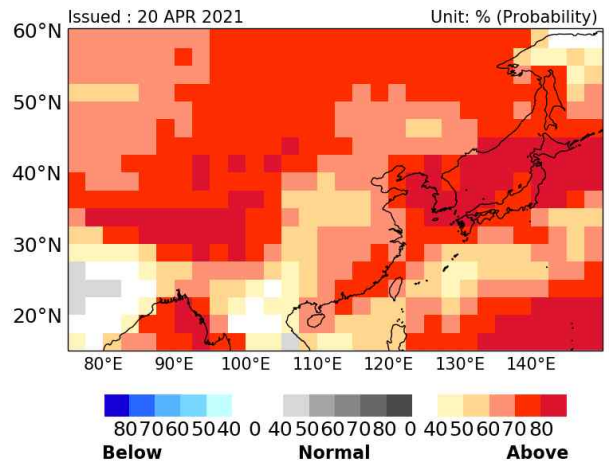
FMA Temperature



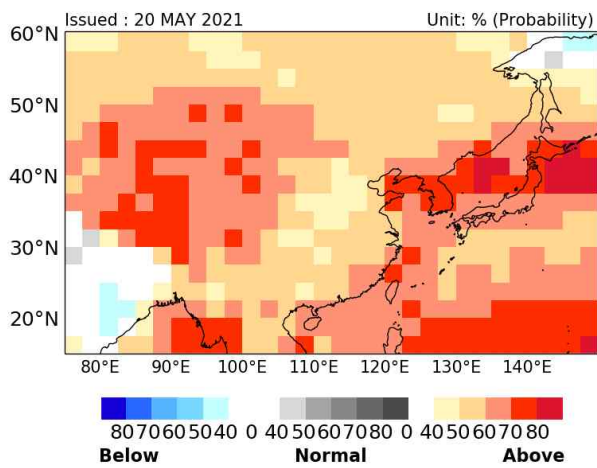
MAM Temperature



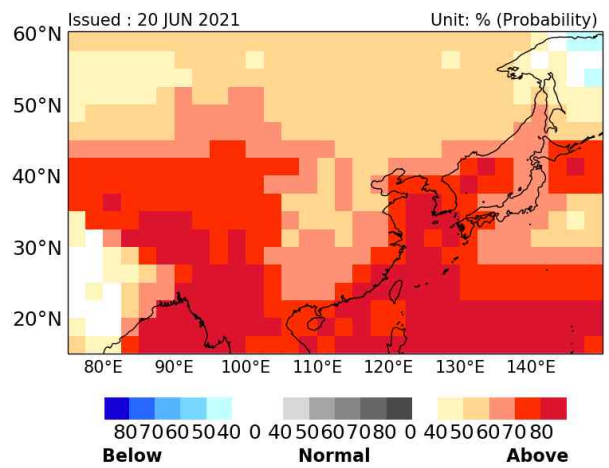
AMJ Temperature



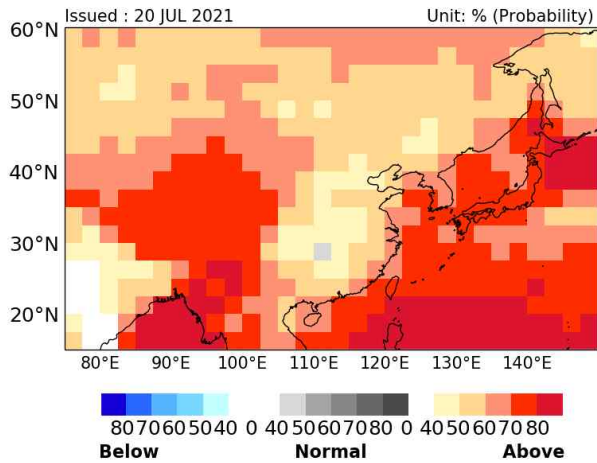
MJJ Temperature



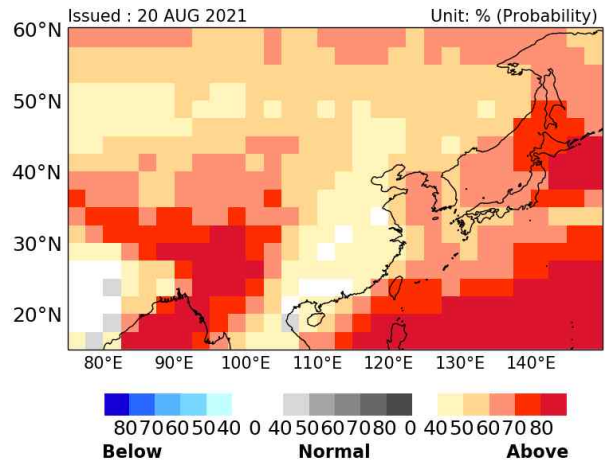
JJA Temperature



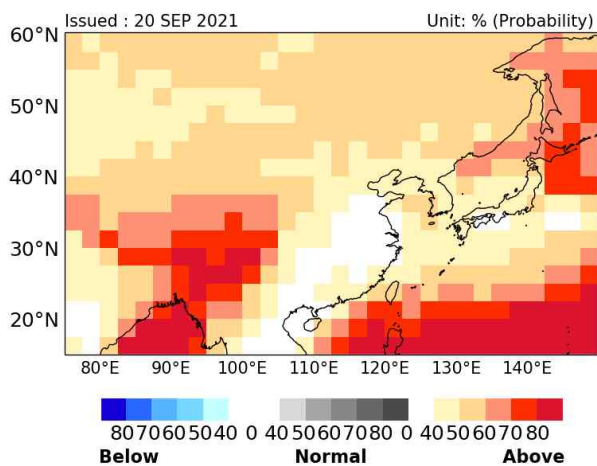
JAS Temperature



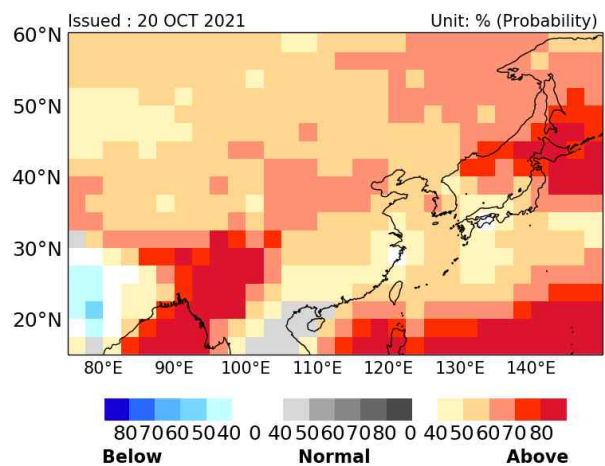
ASO Precipitation



SON Precipitation



OND Precipitation



NDJ Precipitation

Figure 25. Grid PMME Temperature forecast.

### (나) 동아시아에 대한 기후 전망

기후 전망에 이용된 예측인자의 전망은 매 계절별로 동아시아 강수 변동과 영향 기후인자인 ENSO/서태평양 강수 변동 간 물리적 관계를 분석한 결과를 토대로 이를 표출하도록 설계하였다. 기후 전망에 이용된 예측인자의 전망으로 ENSO 지역의 SST (Sea surface temperature) 와 EOF (Empirical orthogonal functions)의 PC (Principal component)와, TWP Precipitation 정보를 매월 제공한다. Figure 26은 2021년 FMA부터 NDJ까지의 ENSO 지역의 SST와 PC 정보를 나타내고, Figure 27은 TWP Precipitation 정보를 나타내며 Contour 형태와 Mesh 형태 모두를 제공한다.

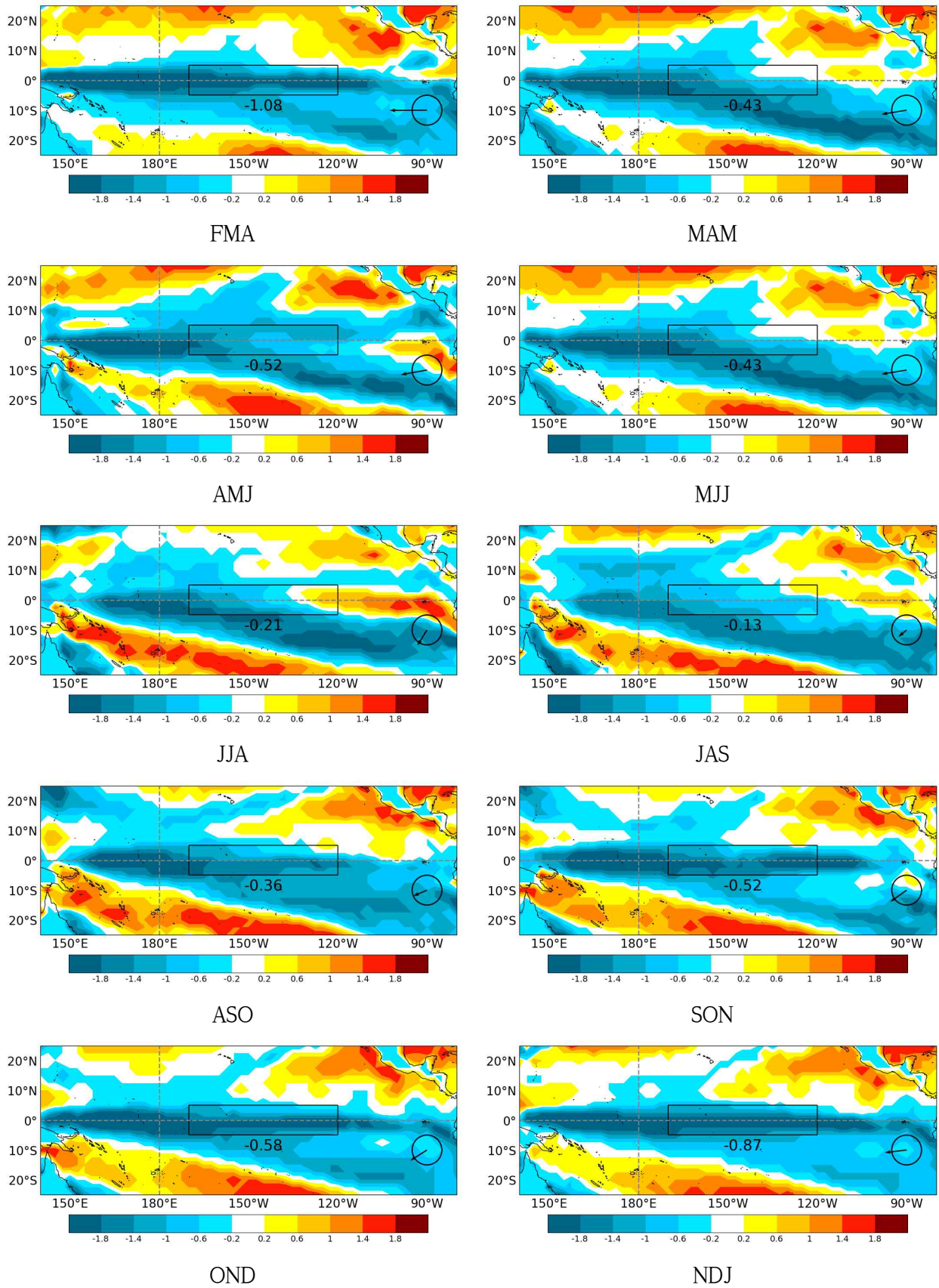
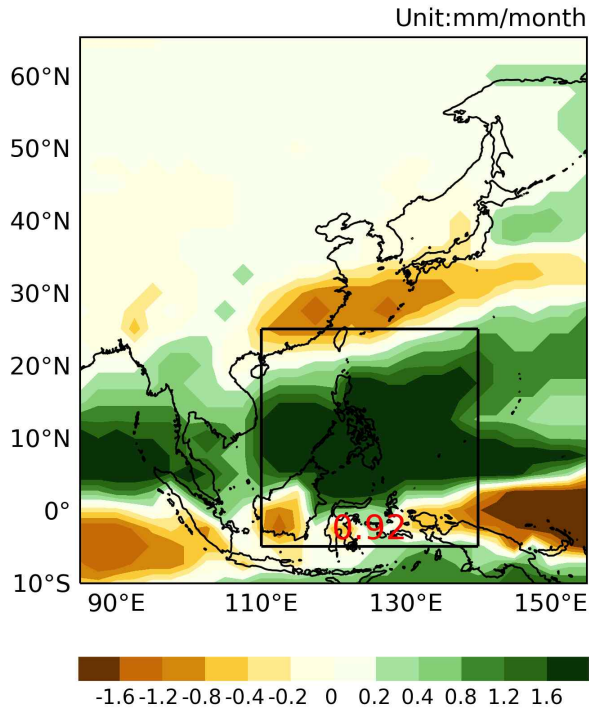
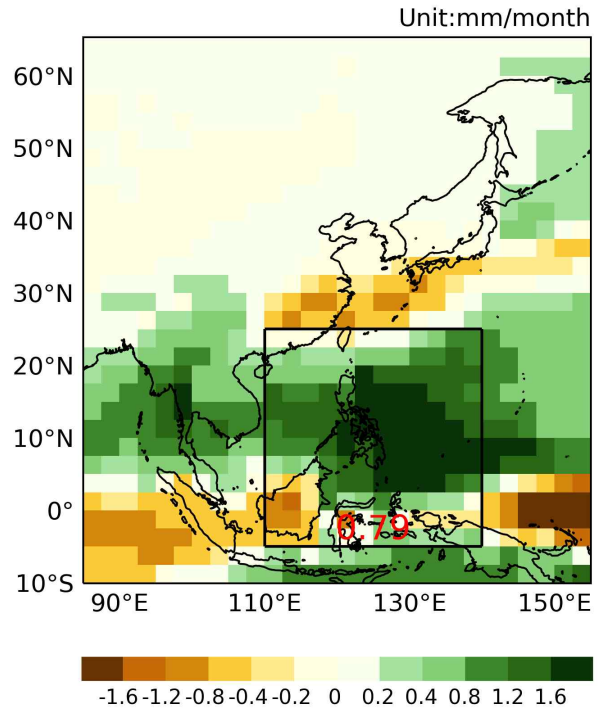


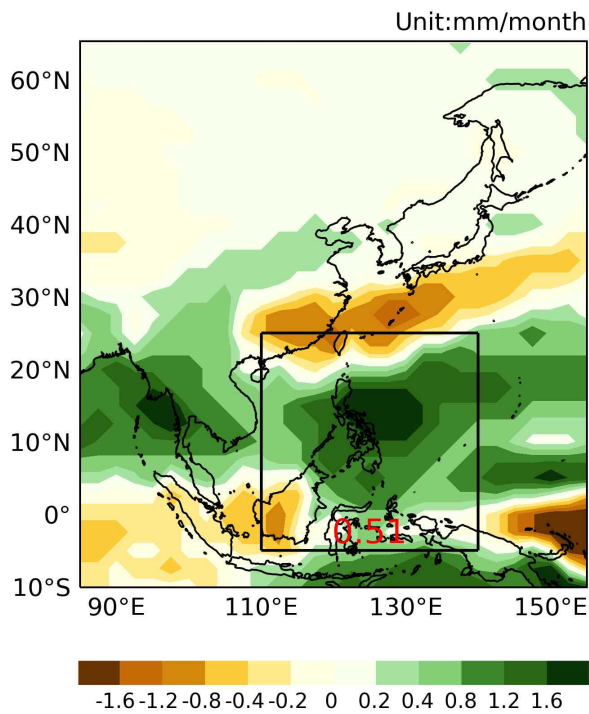
Figure 26. Forecast of predictors (SST + PC).



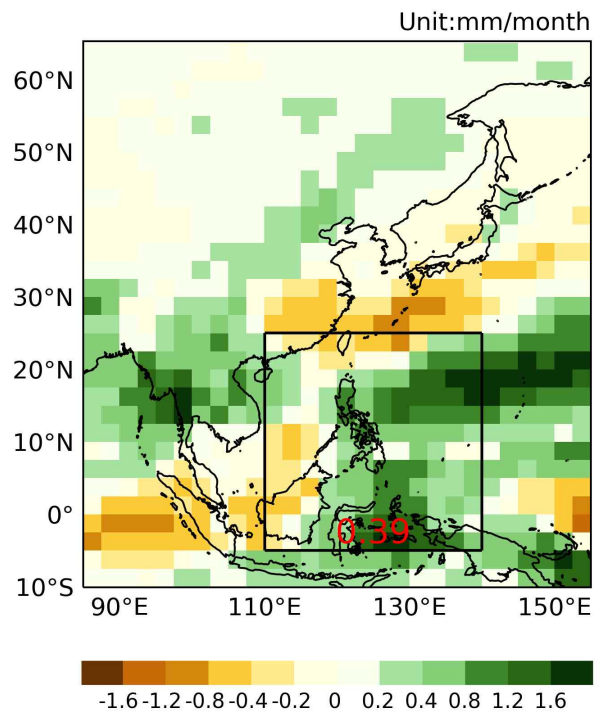
FMA



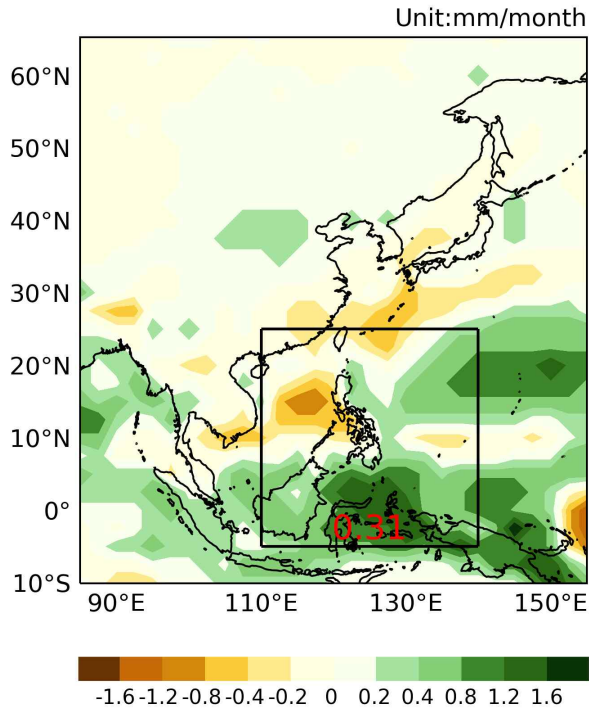
MAM



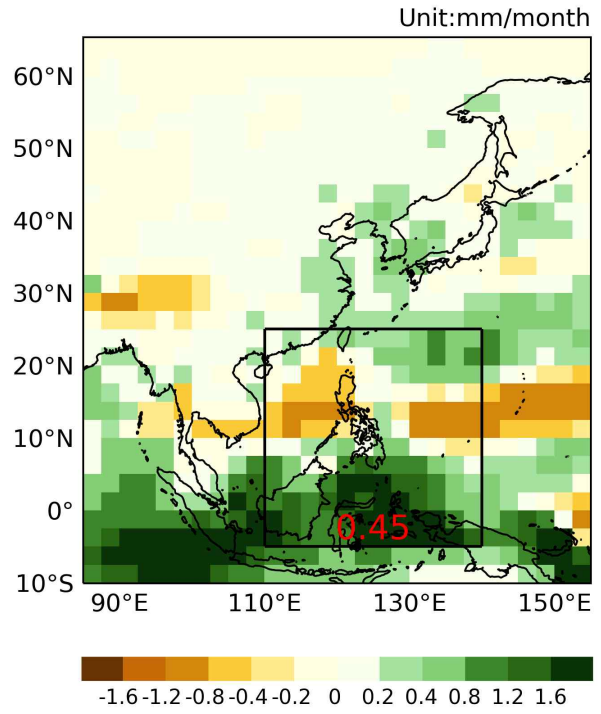
AMJ



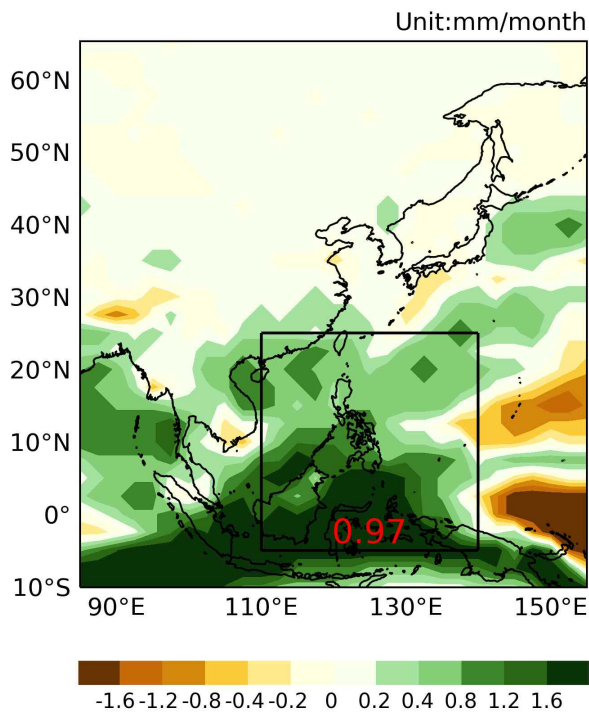
MJJ



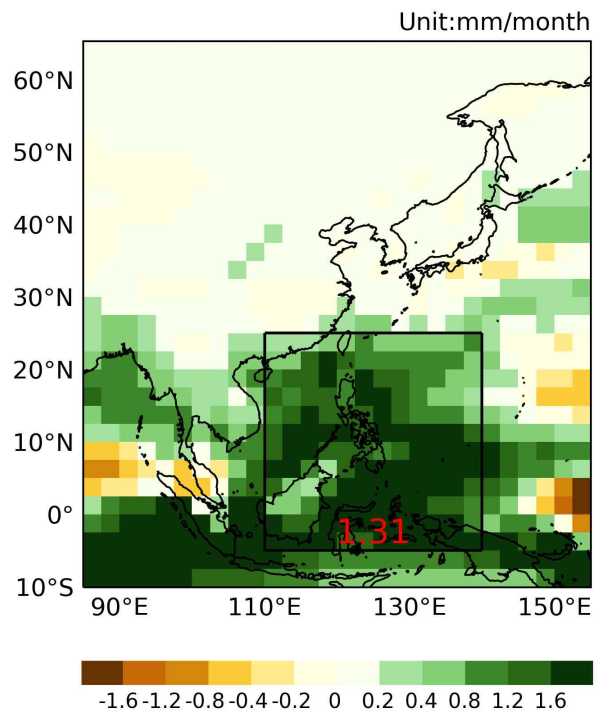
JJA



JAS



ASO



SON

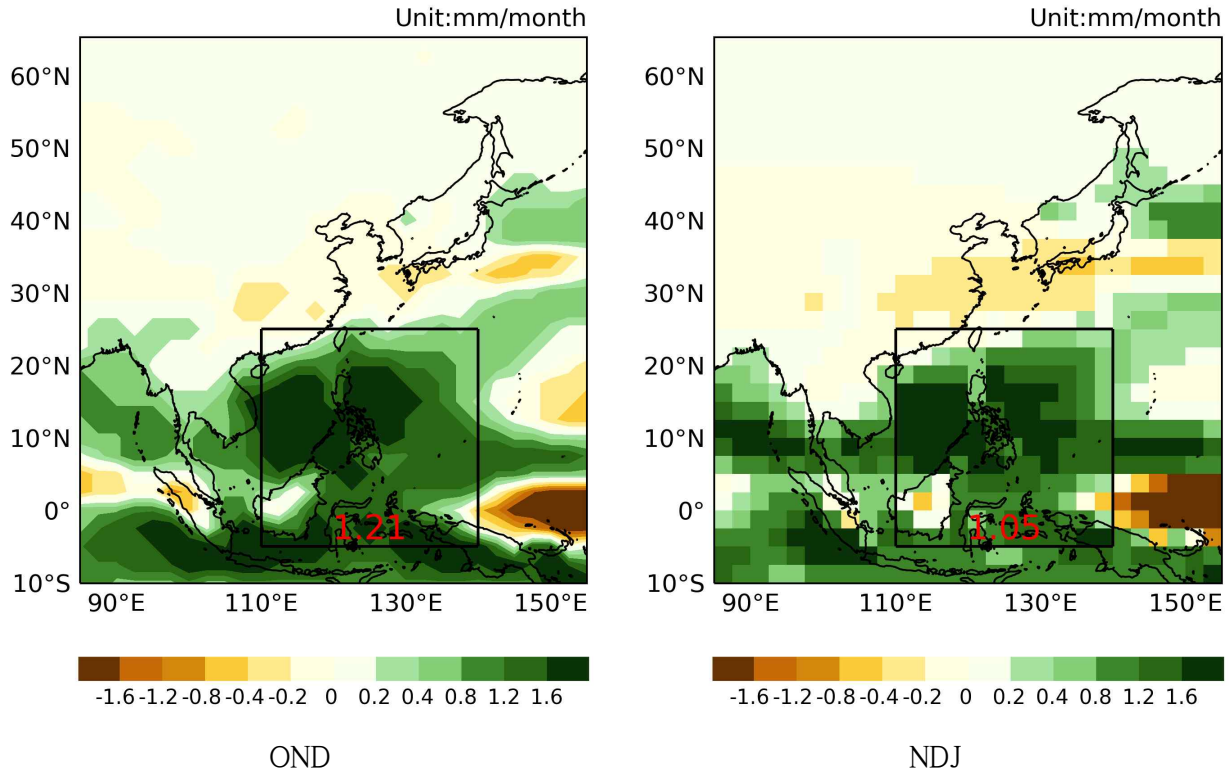


Figure 27. Forecast of predictors (TWP Precipitation, left contour; right mesh).

#### (다) 동아시아에 대한 기후 전망

최근 동안의 지난 예보 및 관측의 추이에서는 기후 전망에 이용된 예측인자의 전망은 통해 생산된 결과를 해당 월까지의 누적 결과를 표출하도록 설계하였다. 가시화 시스템을 구동하는 목표연도와 월을 기준으로 과거 11개월까지 총 12개월의 추이를 확인할 수 있다. 추가적으로 관측자료의 추이를 함께 제공함으로써, 지난 예보에 대한 추이와 관측에 대한 추이를 함께 제공하고자 하였다. Figure 28은 ENSO 지역의 1년간의 PC정보를 관측자료와 함께 제공하기 위한 결과와, TWP Precipitation의 1년간의 정보를 관측자료와 함께 제공하기 위한 결과를 2021년 NDJ를 기준으로 나타낸다.

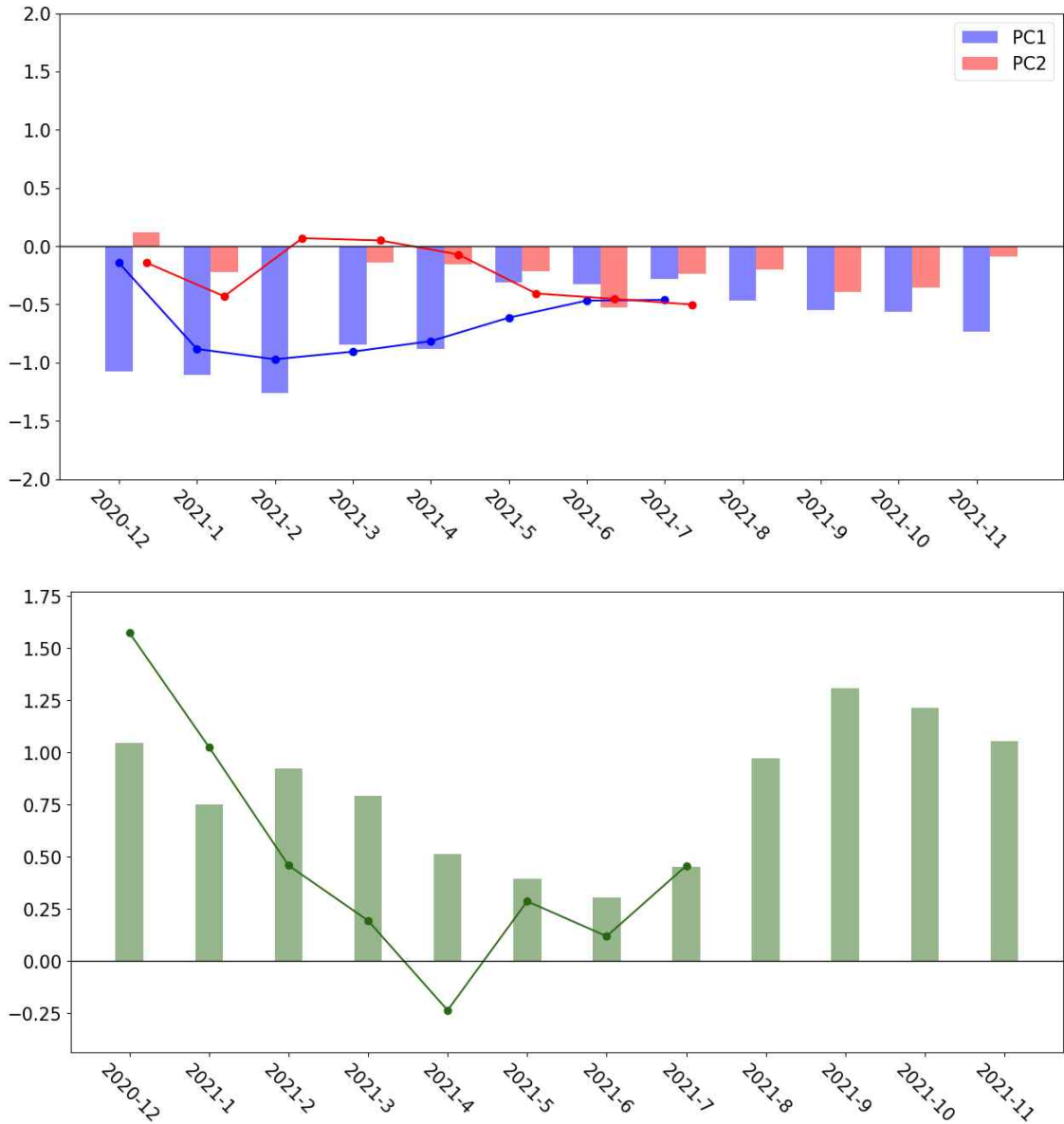
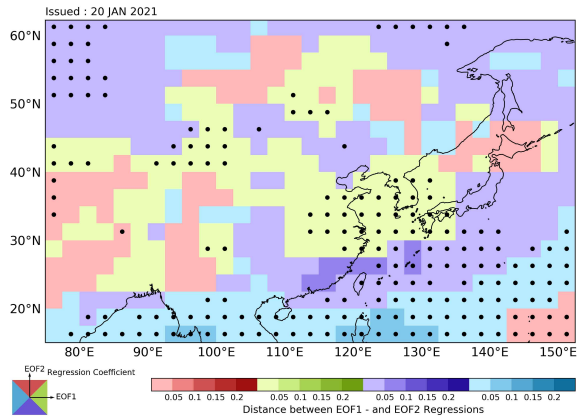


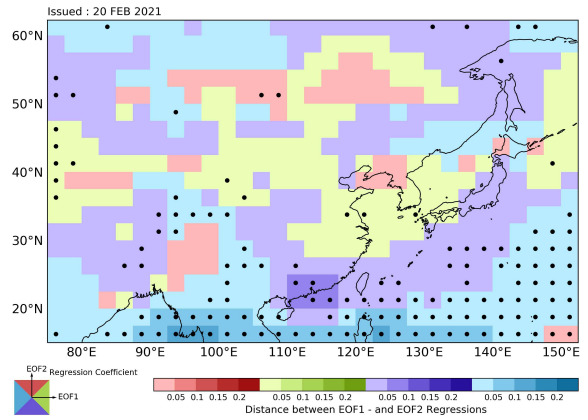
Figure 28. Trends of past forecasts and observations during the recent period.

(라) 예측인자와 동아시아 기후의 관련성

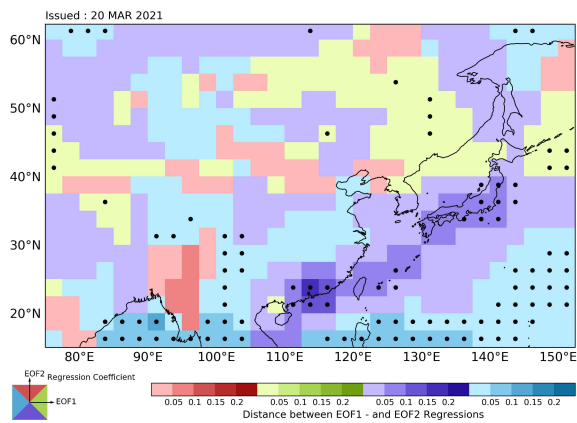
예측인자와 동아시아 기후의 관련성과 관련된 결과로는 두 가지를 제안하였다. 첫 번째는 물리현상 기반의 앙상블모델 재구성으로, 새 MME를 바탕으로 예측성능 및 계절 간 변동의 특징이 분석된 결과와 ENSO의 non-mature phase에서의 예측성능 및 요인을 분석한 결과를 표출한다. 두 번째는 열대 서태평양 지역에서 영역 평균된 강수 지수에 대한 동아시아 여름 강수 변동성에 대한 회귀 계수를 나타내는 열대 서태평양 강수와 동아시아 강수의 관련성 결과를 제시하였다. Figure 29은 2021년 FMA부터 NDJ까지의 물리현상 기반의 앙상블모델 재구성 결과이고, Figure 30은 같은 기간의 열대 서태평양 강수와 동아시아 강수의 관련성 결과이다.



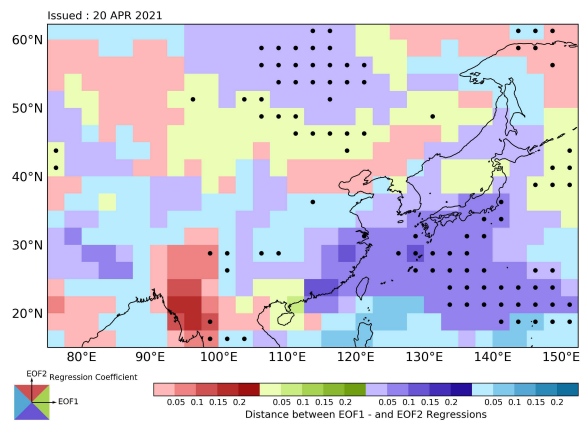
FMA



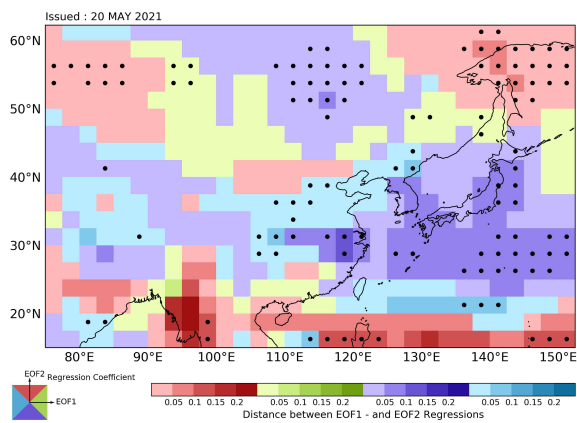
MAM



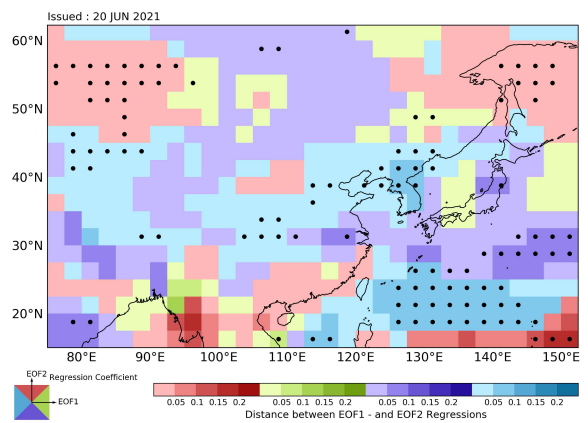
AMJ



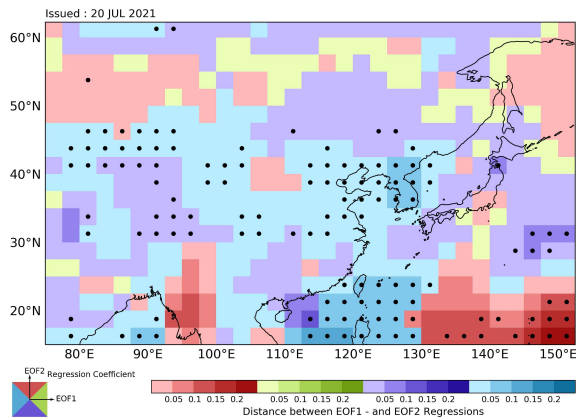
MJJ



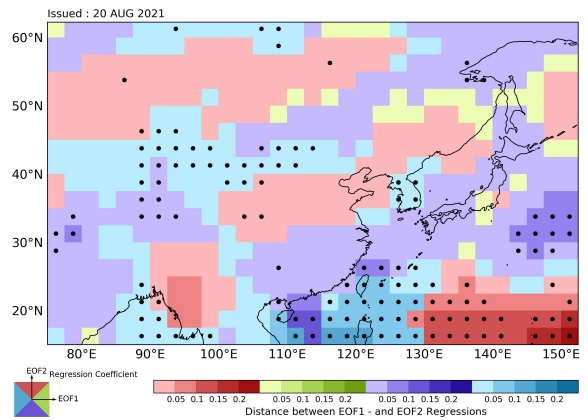
JJA



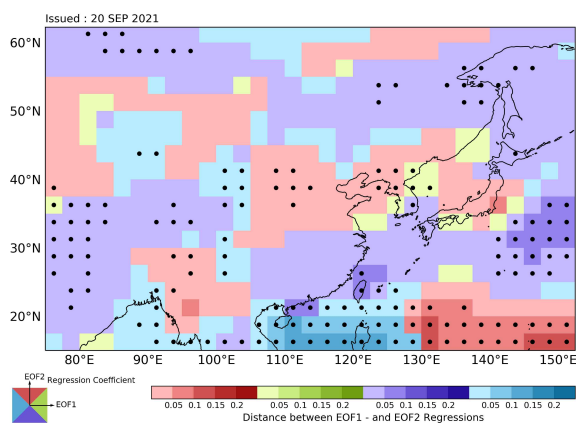
JAS



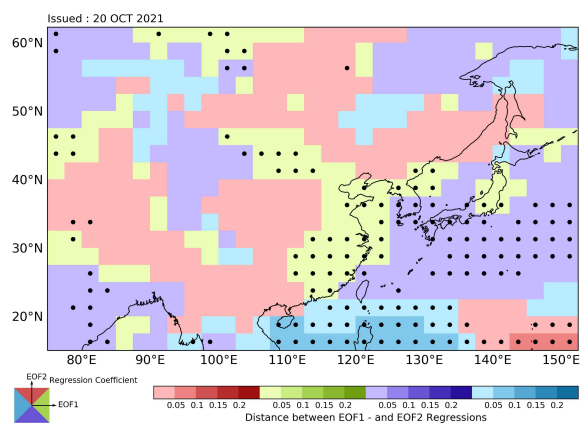
ASO



SON

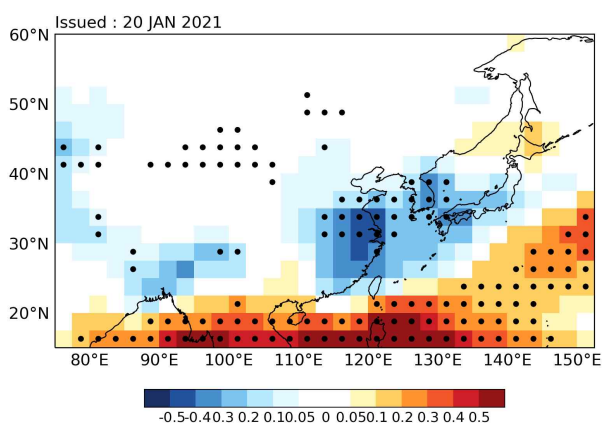


OND

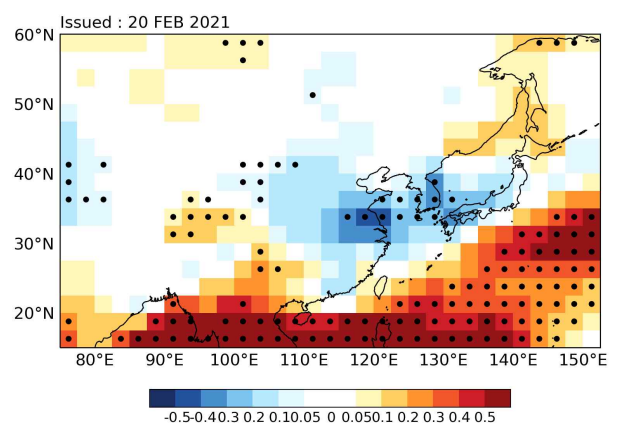


NDJ

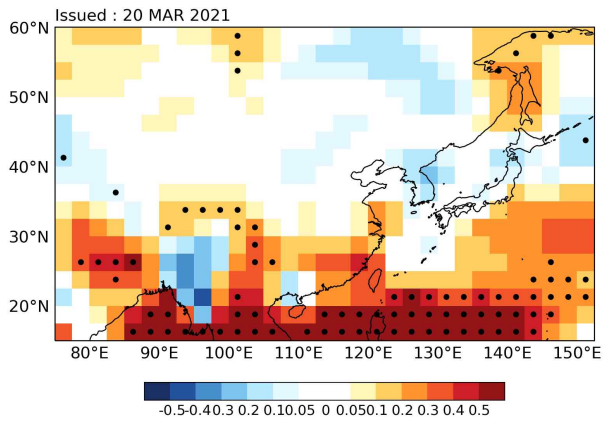
Figure 29. Relative signs and distance magnitudes (see the relevant coordinate and scale at bottom, respectively) on a 2-dimensional plane depicting regression coefficients of the seasonal mean anomalous precipitation onto the first to PCs for FMA to NDJ.



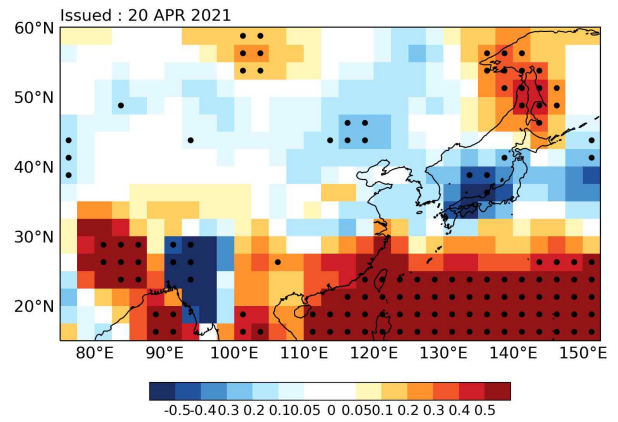
FMA



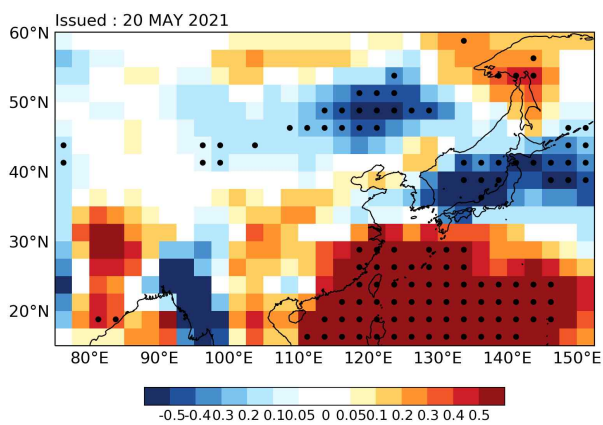
MAM



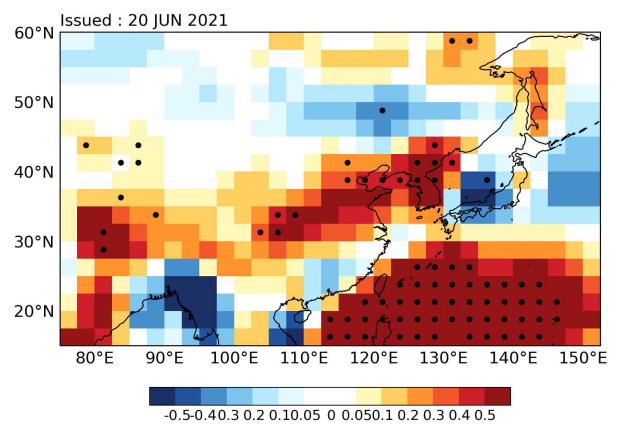
AMJ



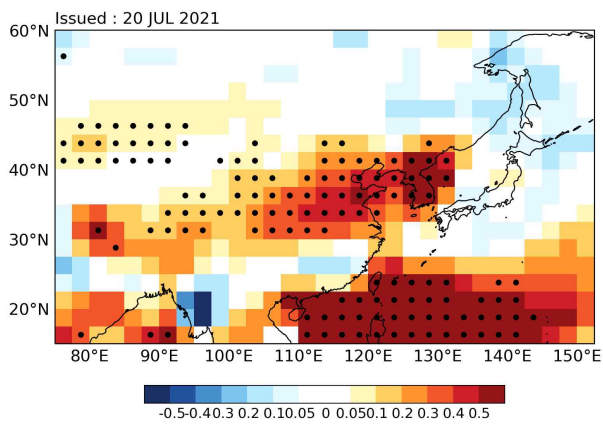
MJJ



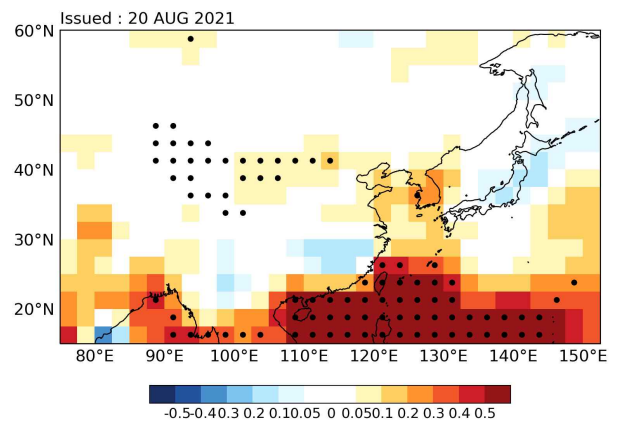
JJA



JAS



ASO



SON

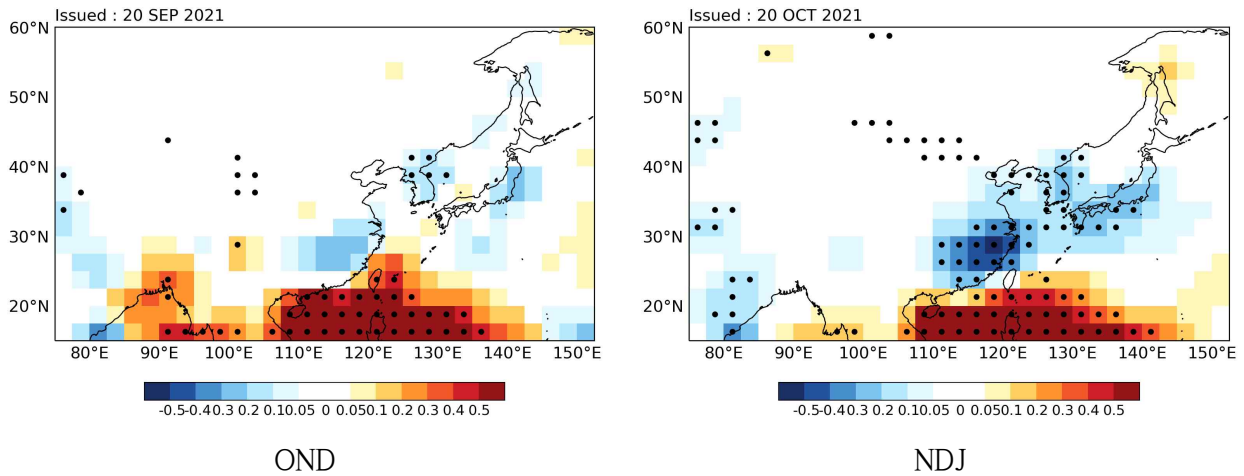


Figure 30. Regression of seasonal mean anomalous East Asian rainfall on tropical Western Pacific rainfall.

#### (마) 최근 동안의 지난 예보의 검증

동아시아에 대한 과거 재현기후 예보의 검증은 현재까지의 HSS 결과와, 관측자료와 비교하여 카테고리 일치 비교 결과를 함께 제시한다. HSS의 결과와 관측자료와 카테고리 일치 비교의 결과는 목적 연도와 월을 기준으로 과거의 모든 결과를 제시하여 과거 재현기후 예보의 검증을 한눈에 할 수 있도록 설계하였다. 또한 동아시아에 대한 HSS 지도를 매일 제공함으로써, 전망에 대한 예측성능을 함께 보여준다. 강수와 기온의 확률 예측의 결과가 실제 관측의 값을 비교하여 제공하는 결과를 통해 해당 예보가 정확성을 가지는지도 확인하고자 하였다. Table 5은 지난 예보의 검증으로 2021년도 NDJ를 기준으로 한 HSS의 누적 결과를 보여준다. Table 6은 같은 기간을 기준으로 한 관측자료와의 카테고리 일치 결과에 대한 정리를 보여준다. 해당 카테고리 일치의 결과가 일치한다면 초록색 불일치의 경우 붉은색으로 표출하였다. Figure 31는 2021년도 FMA부터 NDJ까지의 동아시아 지역의 강수 확률예보의 HSS 결과는 나타내고, Figure 32는 같은 기간의 동아시아 지역의 기온 확률예보의 HSS 결과를 나타낸다. Figure 33은 2021년 FMA부터 JAS까지의 강수와 기온에 대해 예측과 실제의 비교 결과를 나타낸다.

Table 5. Verification of past forecasts (HSS).

	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
2021	24.1	45.2	9.8	17.2	25.4	11.4	21.6					
2020	33.8	19.5	15.7	0.2	22.8	22.1	25.9	41.4	27.4	31.2	9.3	13.2

#### Precipitation

	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
2021	36.1	37.4	17.2	-2.4	13.4	9.3	35.8					
2020	67.1	45.0	33.0	1.7	5.5	11.1	26.9	23.3	40.4	22.8	87.5	86.8

#### Temperature

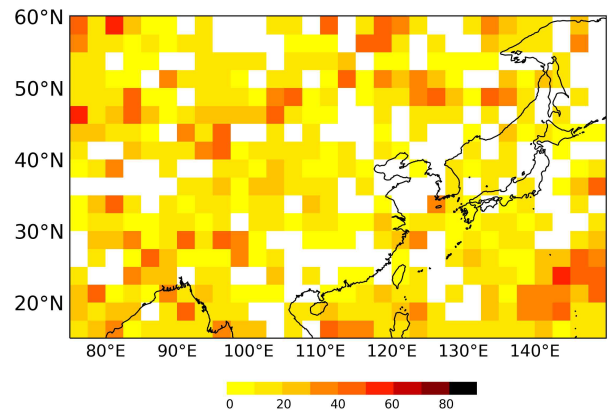
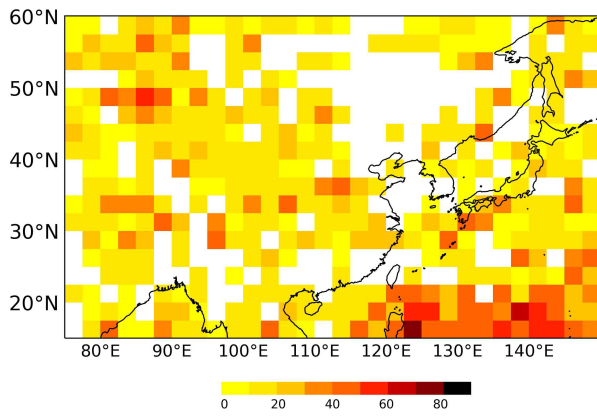
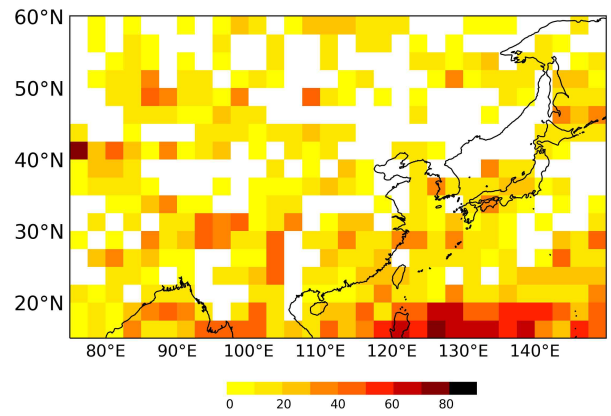
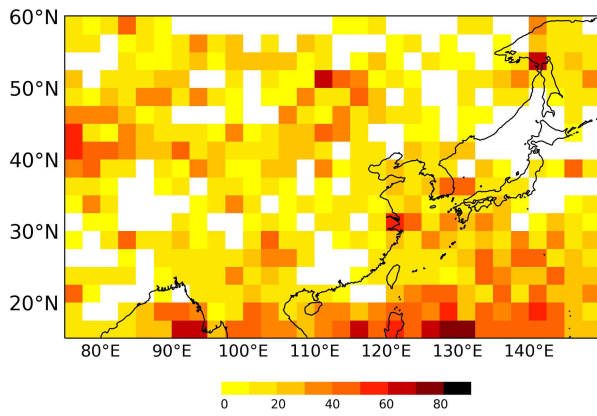
Table 6. Verification of past forecasts (category matching).

	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
2021	+/+	+/+	+/+	+/+	+/+	+/+	+/+					
2020	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+

Precipitation

	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
2021	+/+	+/+	+/+	-/+	+/+	+/+	+/+					
2020	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+

Temperature



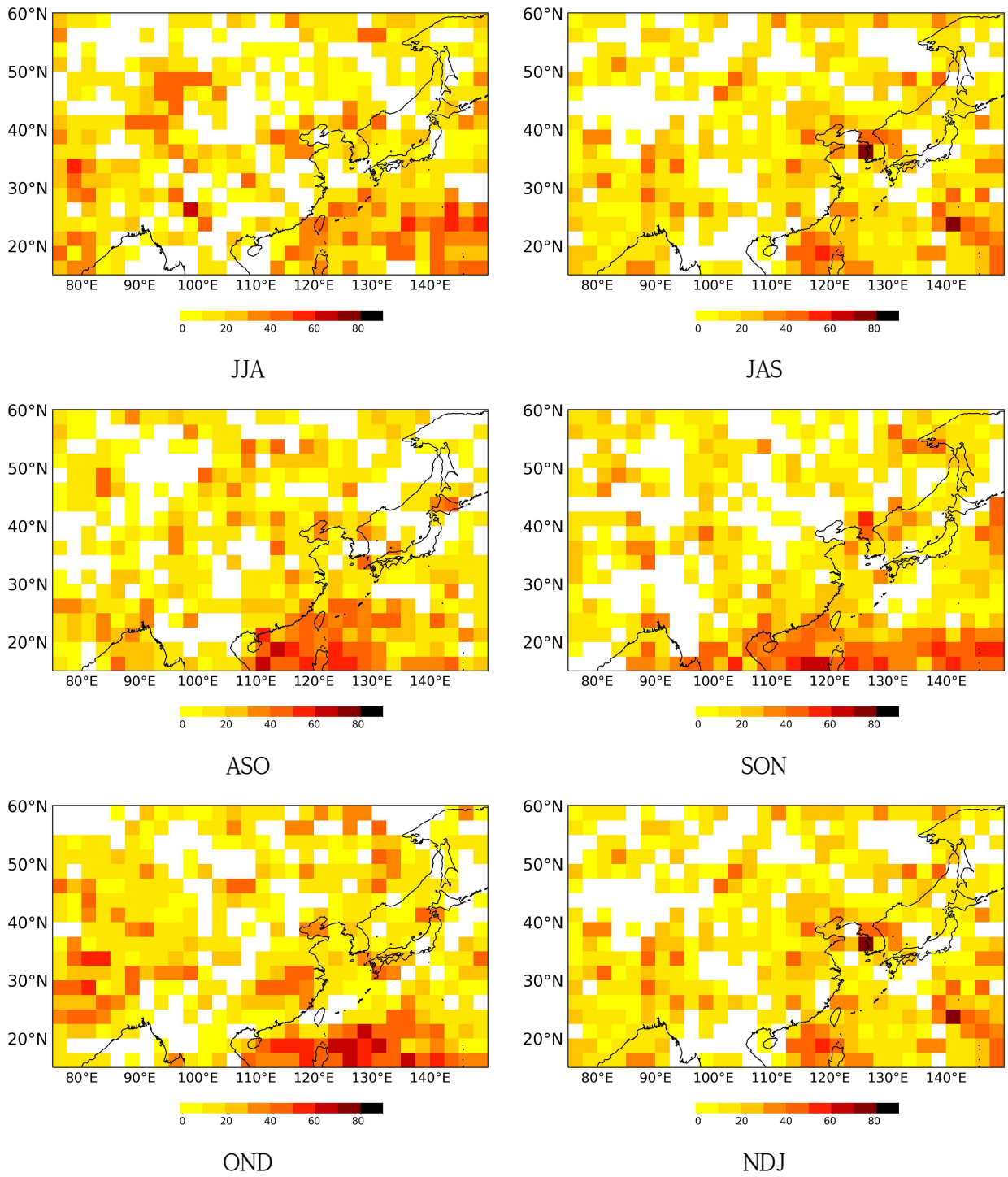
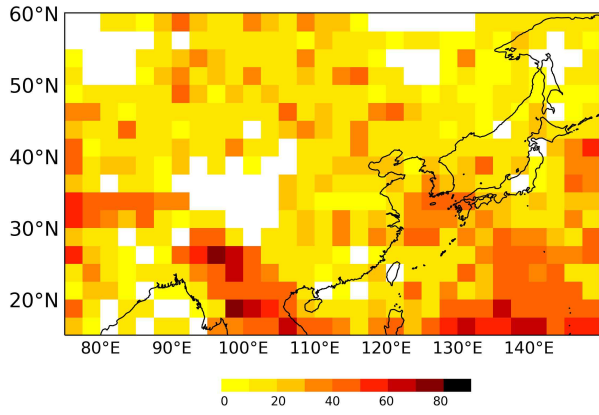
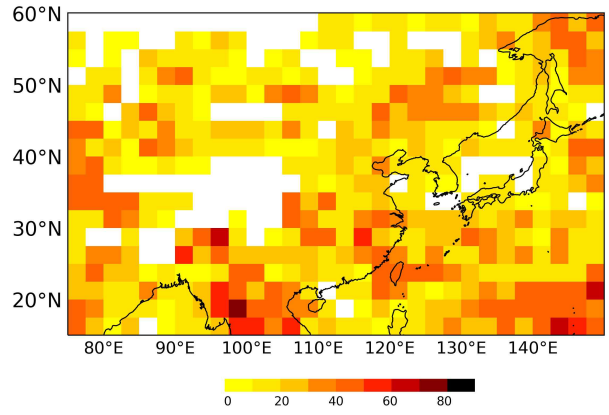


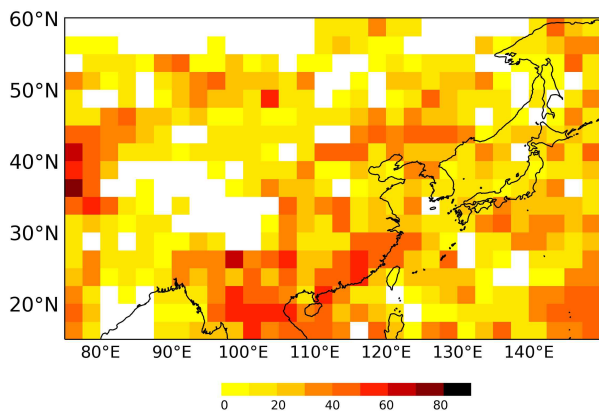
Figure 31. Heidke skill score for East Asia (precipitation).



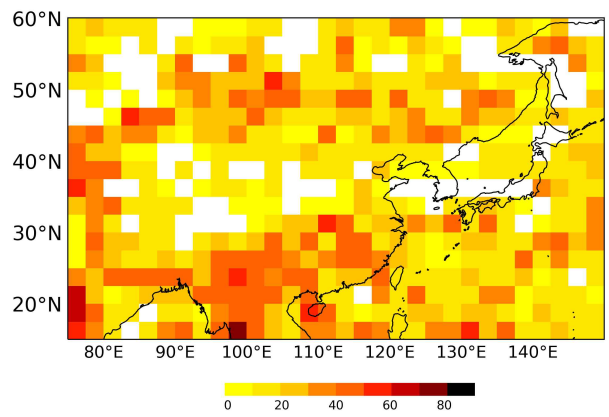
FMA



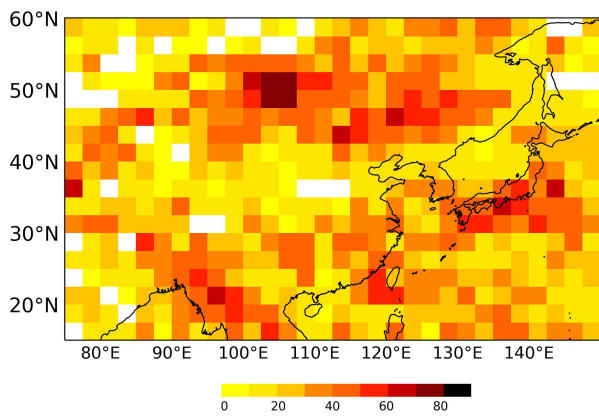
MAM



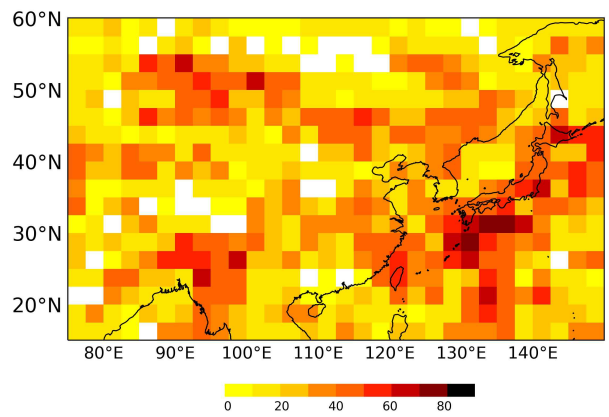
AMJ



MJJ



JJA



JAS

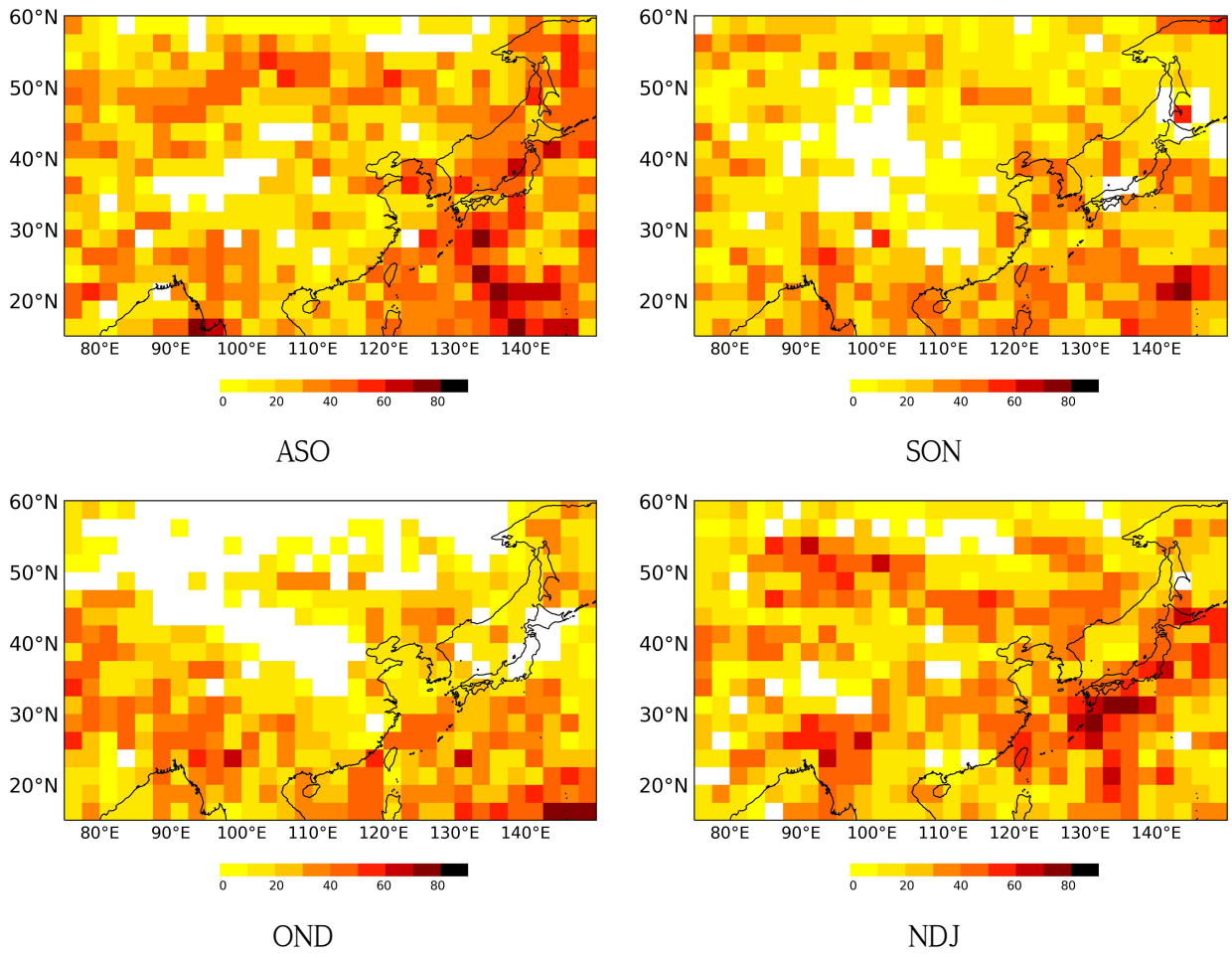
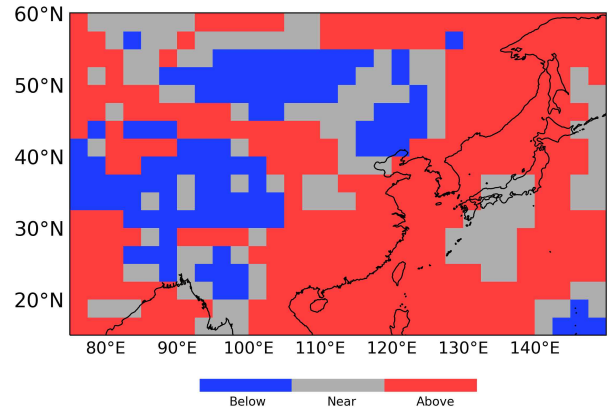
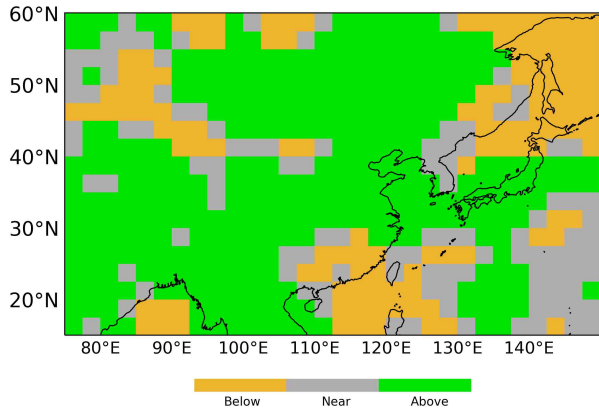
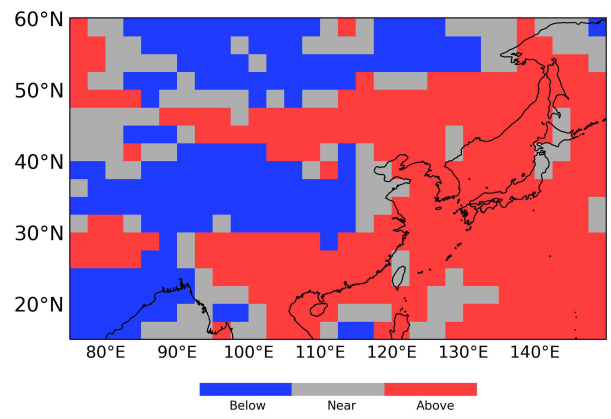
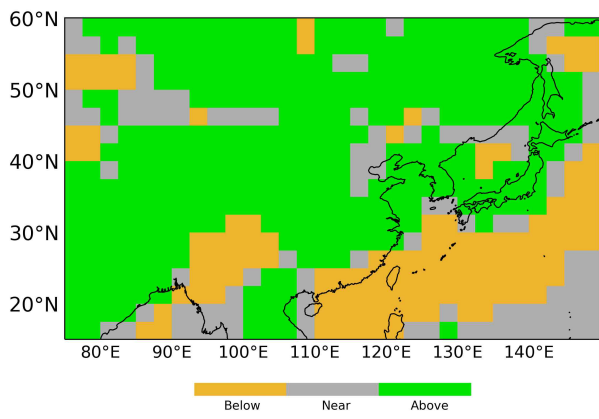


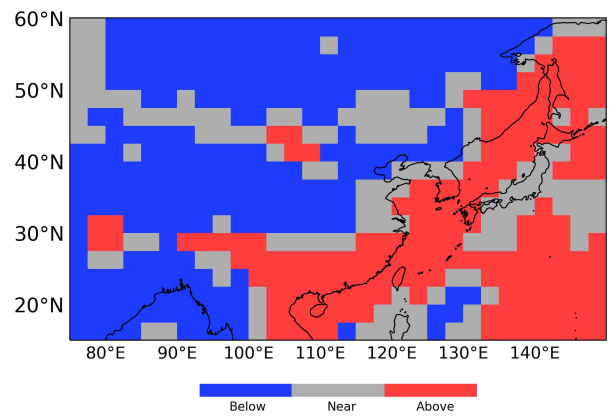
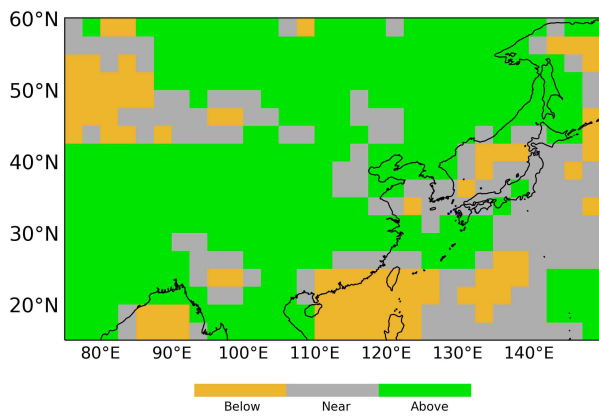
Figure 32. Heidke skill score for East Asia (temperature).



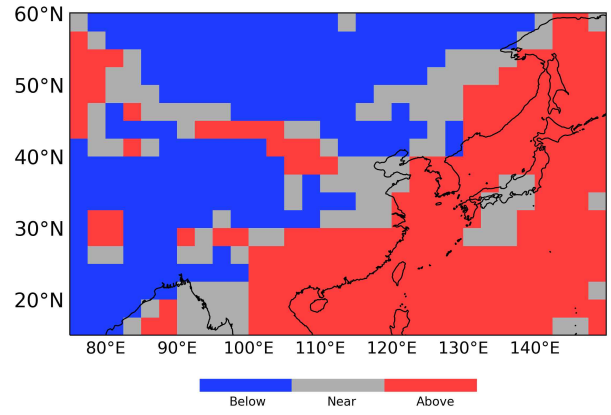
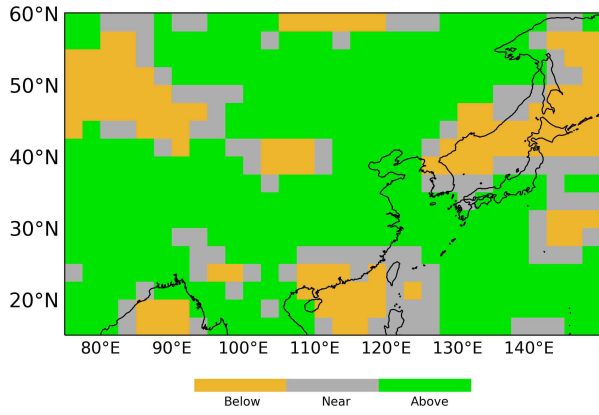
FMA



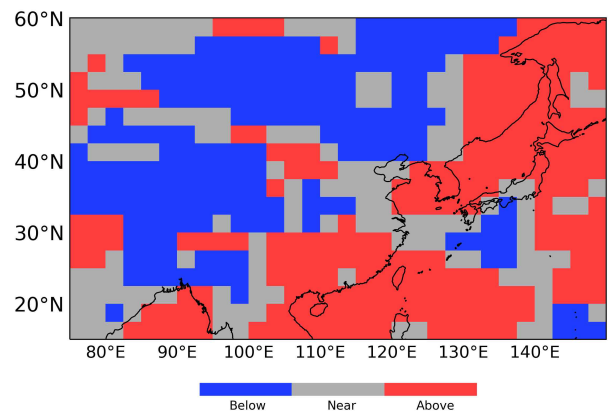
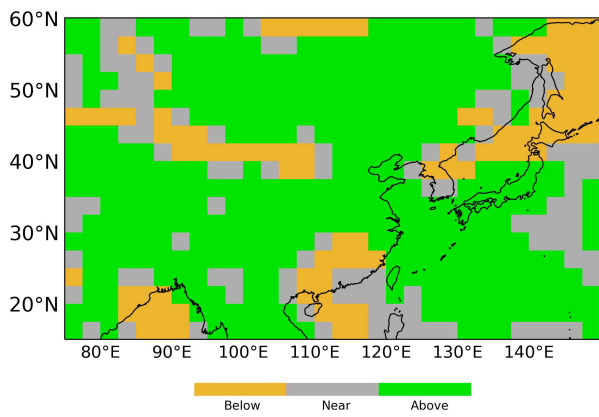
MAM



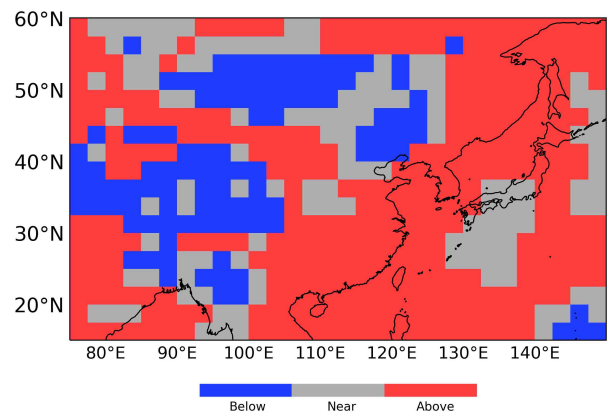
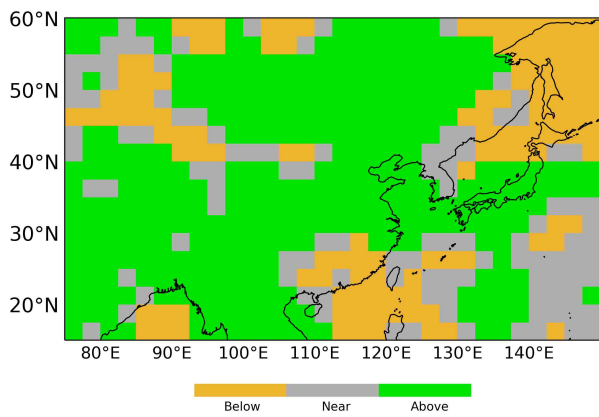
AMJ



MJJ



JJA



JAS

Figure 33. Forecasted precipitation versus actual (left), forecasted temperature versus actual (right).

## 4. 다중모델앙상블의 후처리를 통한 동아시아 예측성 향상 가. 개요

이 연구에서는 다중모델앙상블(MME; Multi-Model Ensemble) 계절 기후 예측 자료의 후처리를 통해 보다 예측성이 향상된 보정 계절 예측 자료를 생산할 수 있는 기술을 개발하고, 이를 통해 예측성 향상을 꾀한다. 이를 위하여 베이지 확률론에 기반한 조건부 확률 갱신 접근법(Bayesian Approach)과 최신 인공지능 기술(AI; Artificial Intelligence)을 도입하여 동아시아 지역에 최적화된 예측 정보를 생산하고자 한다. 특히, 최근 컴퓨터 영상 처리 및 모델링 분야에서는 디지털화된 빅데이터(Big Data)의 누적과 연산 능력의 발전, 그리고 관련 알고리즘의 개발에 힘입어 인공신경망(Artificial Neural Networks) 기반의 인공지능 기술이 획기적인 발전을 이루고 있다. 이 연구에서는 컴퓨터 비전 및 영상 예측 분야의 최신(SOTA; State-Of-The-Art) 아키텍처인 유-넷(U-Net)을 기반으로, 기후 정보에 적합한 다양한 예측 성능 향상 기작을 추가한 맞춤형 후처리 아키텍처를 개발하고, 이를 동아시아 계절 기후 예측에 맞게 학습시켜, 기존 다중모델앙상블 확률 예측 대비 약 9.5%의 예측 성능 향상을 이루었다. 이에, 맞춤형 아키텍처의 구성, 각 파트별 개발 내용, 예측 성능을 향상시키기 위한 사전 학습 전략 및 미세 조정을 위한 최종 학습 방법과 함께 예측 결과를 소개하고 기존 확률 예측과 비교하고자 한다.

### 나. 선형 베이지 접근법에 기반한 다중모델앙상블 격자 자료 후처리

신뢰도 높은 계절 예측을 위하여 우리 센터에서는 다중모델앙상블(Multi-Model Ensemble; MME) 기법을 활용하여 여러 예측 센터에서 생산한 역학 모델의 예측 결과를 종합한 최적의 예보를 생산하고 있다. 이러한 다중모델앙상블 기법을 통해 앙상블에 참여하는 개별 모델의 바이어스가 제거되어 보다 안정적이고 신뢰도 높은 계절 예측이 가능한 것으로 알려져있다. 우리 센터에서는 객관적인 예측 정보 생산을 위하여 주로 결정론적 예측값은 단순합성법을 사용하고 확률론적 예측값은 가우스 방법을 활용하고 있다. 그러나, 이러한 자료 정합성 중심의 접근 방법은 역학 모델이 모의하는 기후 시스템과 실제 기후 시스템간의 괴리에 대한 고려가 부족하여, 예측성이 중요한 현업 기후 계절 전망 측면에서는 아쉬움이 남아있다. 따라서, 이 연구에서는 자료 정합성 중심의 다중모델앙상블 기법을 확장하여 현업 예측성 향상에 초점을 맞추고, 이를 위해 정규화 기법이 적용된 베이지 회귀를 이용하여 과거 관측 자료와 다중 모델의 예측 자료를 적합하고자 하였다.

#### (1) 베이지 접근법

기존의 빈도론적 접근법에서 회귀는 다음과 같이 표현된다.

$$f(\mathbf{X}) = \beta_0 + \sum_{j=1}^p \mathbf{X}_j \beta_j + \epsilon = \beta^T \mathbf{X} + \epsilon$$

여기서  $\beta$  는 회귀 계수를 의미하며  $\epsilon \sim N(0, \sigma^2)$  는 원점 주위에 표준편차  $\sigma$  로 분포하고 있는 오차를 나타낸다. 일반적인 최소곱제법에서는 잔차제곱항의 합을 최소화하는 방법으로 회귀 계수의 초평면을 적합하며, 이 경우 최대 우도 추정법의 회귀 계수는 다음과 같이 나타낼 수 있다.

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

한편, 베イズ 추론에서는 수집된 정보(E)를 바탕으로 기존의 확률(p(H))을 갱신(p(H|E))해 나간다. 이를 수식으로 표현하면 다음과 같다.

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} = \frac{P(D|\theta)P(\theta)}{\int P(D|\theta)P(\theta)d\theta}$$

베イズ 회귀를 통해 다음과 같은 예측값을 업데이트하게 된다.

$$\mathbf{y} \sim \mathcal{N}(\beta^T \mathbf{X}, \sigma^2 \mathbf{I})$$

이 연구에서는 다중모델앙상블에 참여하는 개별 모델의 예측값을 관측에 적합하는 과정에서 베イズ 접근법을 활용하였다. 적합의 대상이 되는 관측 자료로 ECMWF (European Centre for Medium-Range Weather Forecasts)의 재분석 자료인 ERA5를 사용하였고, 예측 변수는 우리 센터에서 운용하고 있는 다중모델앙상블에 참여하는 역학 모델 예측 자료의 개별 평균값을 이용하였다. 기본적인 자료의 정합성을 유지하기 위해, 관측 재분석 자료의 변수에 해당하는 개별 모델의 예측 변수만을 사용하였으며, 개별 모델의 격자값으로부터 관측 재분석 자료의 격자값을 예측하는 격자-격자 적합을 수행하였다. 이러한 자료를 이용하여 각 격자의 개별 모델 예측 격자값(x)으로부터 베イズ 회귀를 통해 예측 격자값을 예측하게 되며, 이를 수식으로 표현하면 위와 같다. 예측 값의 사전 확률 분포는 가우스 분포를 가정하였으며, 이는 원점 주위에 매우 넓게 분포된 무정보적 가우스 분포를 사용하였다. 이를 기반으로 각 격자별 개별 예측 모델의 평균값과 관측 재분석 자료의 회귀를 통해 최종 예측값의 위치 및 모양 파라미터를 추정하였다.

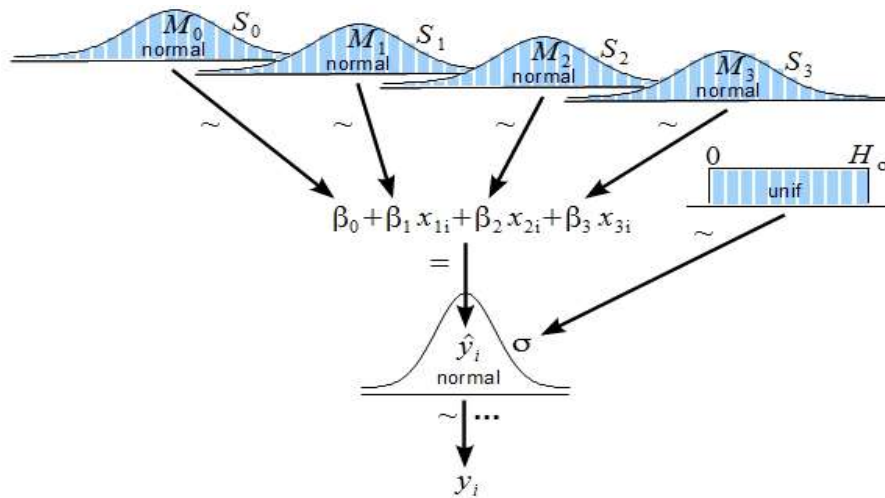
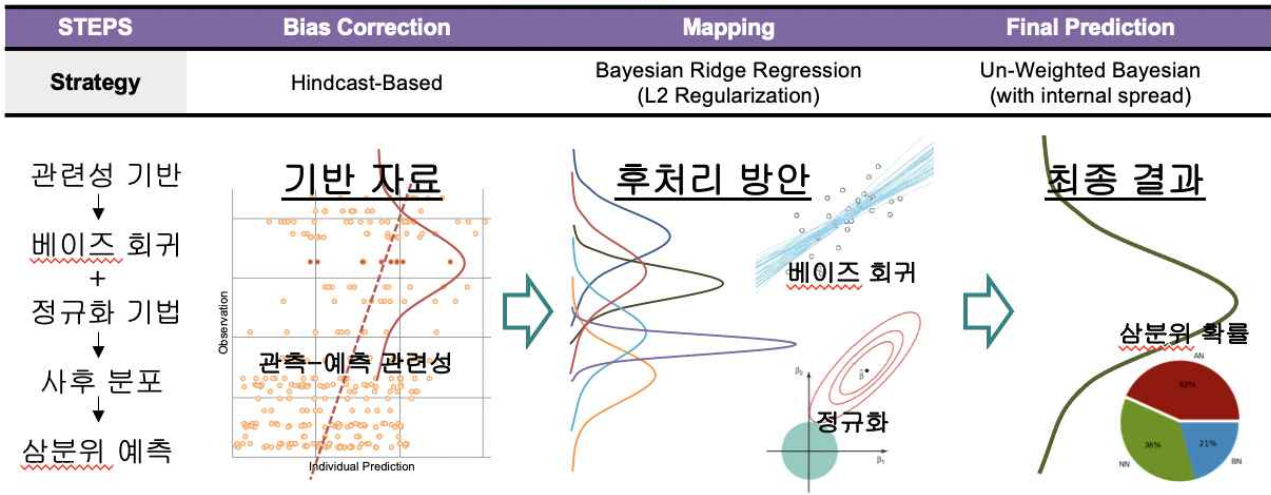


Figure 34. Schematics of Bayesian regression.  
<http://www.indiana.edu/~kruschke/BMLR>.

## (2) 정규화를 도입한 베이지 회귀

베이지 회귀를 통한 추정에서 입력 자료는 다중모델앙상블에 참여하는 개별 모델의 예측 평균값이나, 간혹 개별 예측의 바이어스 등으로 인하여 실제 예측에 기여하지 못하거나 부정적인 영향을 미치는 경우가 존재한다. 이는 추후 격자-격자 모형을 벗어나 모델 결과 통계량 (Model Output Statistics; MOS)를 이용하는 방법에서도 나타날 수 있는 현상으로, 회귀 모형이 실제 관련성에 적합되기보다는 자료 자체의 특성에 적합되는 경우를 의미한다. 이러한 자료 과적합을 방지하기 위하여, 이 연구에서는 정규화 기법을 도입하였다. 정규화 기법은 회귀 계수에 패널티를 부여함으로써 특정 회귀 계수가 과도하게 성장하는 것을 막고 회귀에 영향을 미치지 못하는 변수의 가중치를 제거하여 회귀 모형이 안정적으로 파라미터를 추정할 수 있도록 돕는다. 이 연구에서는 유클리드 공간에서의 거리 제곱에 해당하는 l2-norm 을 이용한 능형 회귀를 도입하여 위와 같은 자료 과적합 현상을 방지하고 안정적인 예측이 수행되도록 구성하였다. 이 연구에서 사용한 후처리 방법론과 자료 처리 방안을 간단히 모식화하면 아래와 같다.

Table 7. Strategies of Bayesian post-processing with regularization applied in this study.



### (3) 선형 베이지스 후처리 모형을 이용한 보정된 예측 결과

하인드캐스트 기간에 leave-one-out 방식으로 계산된 예측 결과는 전체적으로 정규화 베이지스 7 ~ 9 월 1 개월 선행 예측의 특성과 10, 11 월 1 개월 선행 예측의 특성이 조금 다른 것으로 나타난다. 즉, 10, 11 월의 경우 예측 분산이 과소 추정되어 대부분의 예측 범주가 니어-노멀에 머무는 문제가 있는 것으로 나타났으며, 이는 추후 연구를 통해 보완해야 할 사항으로 보인다. 많은 경우, 기반 MME 예측과 정규화 베이지스 후처리를 통한 예측이 비슷한 범주값을 예측하는 것으로 보이나, 일부 관측 자료에 적합된 정규화 베이지스 후처리는 기반 MME 예측과 상이한 결과를 보이기도 한다. 이러한 경향성은 대륙 혹은 해양 등 경계면의 특성이나 위도에 관계 없이 공히 관측되고 있으나, 전체적으로 해양 지역에서의 후처리 효과가 탁월한 것으로 판명되었다. 자료에 민감하게 반응하는 대륙에서의 후처리는 예측성에 큰 영향을 주지 못하는 것으로 나타난다. 하인드캐스트 기간에서 교차 검증한 정규화 베이지스 후처리 결과와 기존 APCC PMME의 예측 성능 검토를 위해, 범주형 예측 자료 검증에 주로 사용되는 Heidke Skill Score를 계산하여 상호 비교하였다. 그러나 짧은 자료 기간으로 인하여 대부분의 경우 후처리 결과가 기존 예측값에 비해 향상되지 않은 것을 확인할 수 있다. 다만, 해양의 영향이 강해지는 겨울철로 진입할 수록 해양 지역에서 후처리 효과는 긍정적인 역할을 기대할 수 있는 것으로 보인다.

하인드캐스트 기간의 자료를 바탕으로 구축된 정규화 베이지스 후처리 모형을 이용하여, 실시간 예측도 수행하였다. 그러나, 예측 분산이 과소모의되는 경향이 커서 각 범주별로 정확한 범주 구분이 이루어지지 않는 문제가 발생하고 있다. 이는 향후 후처리 모형 구성을 발전시키는 데 고려해야 할 것으로 보인다. 기반 APCC-PMME 예측에서는 북서태평양 북동인도양 등 해양 지역에서의 고온이 예측되는 반면 후처리 이후 이러한 고온 예측은 줄어들고 중국 동남쪽으로부터 우리나라를 통과하는 밴드 지역의 평년값 예측이 두드러지게 올라가는 모습을 보인다. 예측 성능은 평균적으로는 향상되는 것으로 보이나, 공간적인 산포가 매우 크게 벌어지고 있음을 확인할 수 있다. 따라서, 적은 자료를 이용해 안정적으로 적합을 수행하고 최종 예측 값을 추

정하는 방식으로 추가 개발이 진행되어야 할 것으로 보인다.

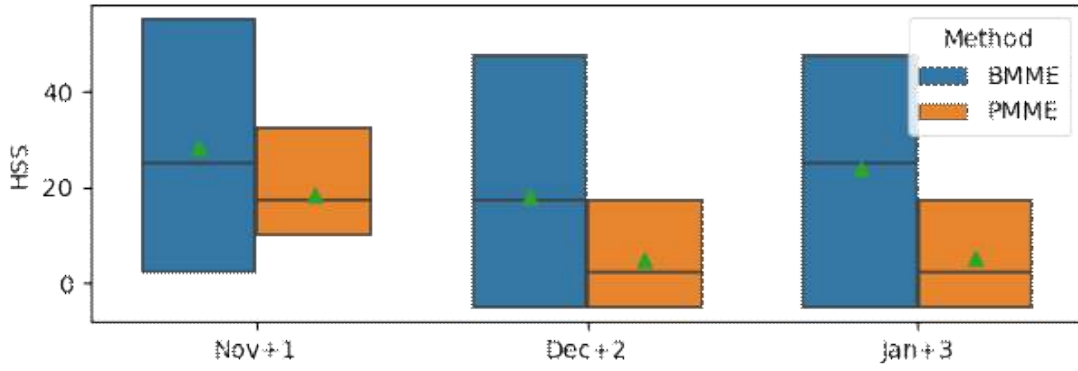


Figure 35. Distribution of the HSS for hindcast cross-validation of (Left) 1-month lead November, (Middle) 2-month lead December, and (Right) 3-month lead January.

## 다. 맞춤형 인공지능망을 통한 비선형 후처리

### (1) 기본 아키텍처: 스킵-커넥트 인코더-디코더 유-넷

이 연구에서는 공간적인 특성이 중요한 다중모델앙상블의 계절 기후 예측 정보로부터 의미 있는 패턴을 추출하기 위해 컴퓨터 비전 이미지 처리 분야의 최신 아키텍처인 인코더-디코더 네트워크와 이의 변형인 유-넷(U-Net; Ronneberger et al., 2015)을 기반으로 맞춤형 후처리 아키텍처를 구성(Figure 36)하였다. 다중모델앙상블 계절 기후 예측 결과의 공간적인 정보를 활용하기 위해 합성곱 신경망(Convolutional Neural Networks)를 이용한 인코더를 구성하였고, 이로부터 저차원의 예측 인자 정보를 추출하게 된다. 추출된 저차원의 정보는 이 연구를 통해 독자적으로 개발된 동적 조건부 드롭아웃 연결층을 통과하며 해당 시기에 맞는 네트워크를 통과하고, 이는 다시 전치 합성곱 신경망을 지나며 동아시아 지역에 맞는 후처리 예측값으로 복원되도록 하였다. 이 과정에서 채널별 합성곱 커널의 사용, 채널-공간 어텐션 기법의 적용, 동적 조건부 드롭아웃의 개발, 그리고 완전합성곱-특성-직접-우회로를 활용함으로써 예측력을 향상시키고자 하였다.

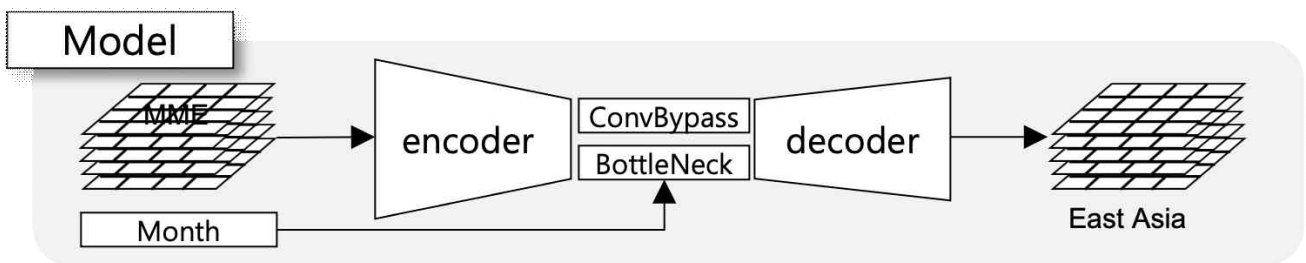


Figure 36. Overview of the customized U-Net architecture with dynamic conditional dropout.

위 구성도(Figure 36)은 이 연구에서 개발한 동적 조건부 드롭아웃이 포함된 맞춤형 유-넷 아키텍처의 전체 구조를 나타낸다. 주 입력 자료로 APCC의 결정론적 다중모델앙상블인 단순합성 계절 예측이 변수별로 입력되고, 동적 드롭아웃을 위한 부차적인 입력 자료로 월 정보가 정수형으로 전달된다. 전자의 예측 공간 정보는 인코더를 통해 저차원 정보를 추출하는 데 사용되는 한편, 스킵 커넥션을 통해 디코더로 특성이 전달된다. 추출된 저차원 정보는 부차적 입력 정보인 월별 마스크가 적용된 동적 조건부 연결층을 통과하고, 이는 다시 디코더를 통해 공간 정보로 복원되는 구조이다.

### (2) 인코더: 채널별 합성곱 커널, 채널-공간 어텐션

#### (가) 채널별 합성곱 커널

이 연구에서 사용된 맞춤형 후처리 아키텍처의 앞쪽은 공간적 의미를 포함하는 다중모델앙

상블 계절 기후 예측 정보로부터 예측을 위한 저차원 정보를 추출하는 합성곱 신경망 기반의 인코더 모듈이다. 일반적인 컴퓨터 비전/이미지 처리 분야의 입력 정보는 대개 RGB 채널의 사진 혹은 영상 자료로서 채널간의 특성이 크게 다르지 않지만, 기후 정보의 경우 변수(온도, 기온, 기압 등)간의 공간적인 특성이 매우 상이하다. 이러한 변수간의 물리/역학적인 특성을 적절히 반영하기 위하여 이 연구에서는 Figure 37와 같이 모든 채널을 동일한 커널을 사용하는 일반적인 합성곱 신경망이 아닌 채널별로 다른 커널을 적용한 뎀스-와이즈 합성곱(Depthwise Convolution)을 적용하였다. 또한 연산 효율성을 고려하여 분해 합성곱을 수행하는 뎀스-와이즈 분해 합성곱 신경망 층을 구성하였고, 셀루(selu)를 이용한 비선형 활성화 이전 배치 정규화(Batch Normalization)을 수행함으로써 모형이 안정적으로 학습을 수행할 수 있도록 조치하였다. 이러한 채널별 분해 합성곱 활성화 신경망층을 2 회 연속 배치하여 공간 정보를 추출할 수 있도록 합성곱 블록을 구성하였다.

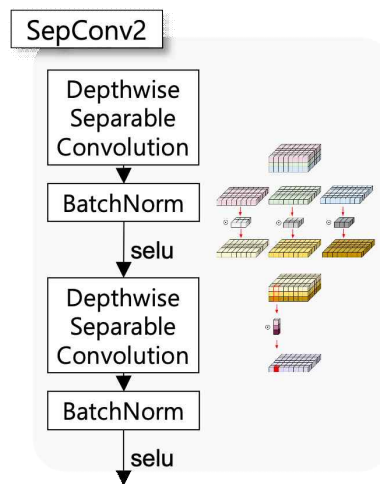


Figure 37. Depth-wise separable convolutional block. Distinct filters by channel convolves separately followed by batch normalization and selu activation. Above operations are performed twice for an individual block.

### (나) 채널-공간 어텐션

합성곱 블록을 통과한 정보는 이후 뒤따르는 합성곱 블록 어텐션 모듈(Figure 38; CBAM; Convolutional Block Attention Module)을 통과한다. 이 연구에서 적용한 어텐션 기법은 선행 연구(Woo *et al.*, 2018)에서 제시한 채널 어텐션과 공간 어텐션을 독립적으로 수행한 후 이어붙이는 CBAM 기법을 적용하였다. 채널 어텐션은 전역 평균 풀링 및 최대 풀링 개별적인 완전 연결층으로 2 회 넘겨받아 합산한 후 최종적으로 시그모이드 활성화를 수행하고, 공간 어텐션은

채널별 평균 풀링과 최대 풀링을 이어붙여 합성곱을 수행한 후 배치 정규화 후 시그모이드 활성화를 수행한다. 채널 어텐션 이후 공간 어텐션까지 수행된 정보는 CBAM 블록으로 엮어 모형이 집중해야할 채널 정보와 공간 정보를 학습할 수 있도록 구성하였다.

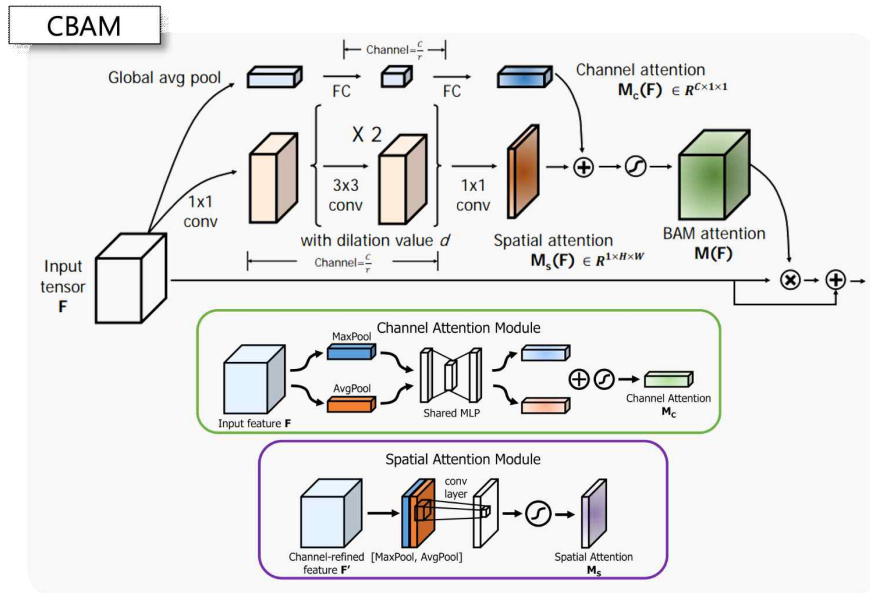


Figure 38. Operation flow chart for convolutional block attention module.

### (3) 동적 조건부 드롭아웃이 적용된 보틀넥과 고차원 우회로

#### (가) 동적 조건부 드롭아웃

기후 현상은 동일한 물리/역학에 지배받는 기후 시스템의 결과인 한편, 계절별로 상이한 경계값에 따라 월별로 독특한 특색을 나타내기도 한다. 따라서, 전체적인 기후 현상의 특성은 유지하되, 일부 월별 특색을 고려할 수 있다면 단일 모형을 이용해 다양한 계절을 안정적으로 예측할 수 있을 것이다. 또한, 월별 특색을 고려할 때, 멀리 떨어진 월에 비해 인접 월과는 비슷한 성질을 가지는 경우를 고려한다면, 부드러운 계절 전이를 표현할 수 있게 된다. 기후 자료의 계절 전이 특성을 고려하기 위해, 이 연구에서는 새로 개발한 동적 조건부 드롭아웃(Figure 39)을 도입하였다.

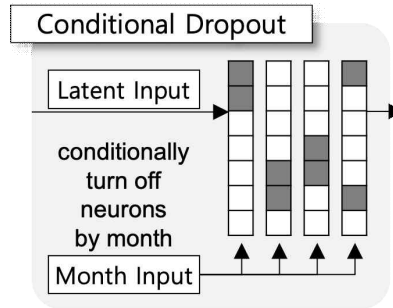


Figure 39. A newly developed dynamic conditional dropout layer. Auxiliary month information is fed into the network to conditionally control neurons at the bottleneck in a way that each month share significant amount of network, while having limited portion of month-specific pathways.

무작위로 섞인 입력 자료에 대응하기 위해 부차적 월 정보에 따라 동적인 연산을 수행하도록 개발하였으며, 전체 보틀넥 뉴런(N) 중 1/12의 정수배에 해당하는 뉴런( $\text{int}(N/12)$ )이 영향을 받고, 나머지 뉴런( $N \bmod 12$ )은 항상 작동하도록 조치하였다. 월에 따라 3/12에 해당하는 뉴런이 소거되며, 인접 월끼리는 2/12의 소거 뉴런을 공유하도록 제작함으로써, 월별 특색을 고려하는 한편, 비슷한 계절내 연관성은 일부 유지할 수 있도록 개발하였다.

#### (나) 고차원 우회로

예측성의 향상을 위하여 다중 우회로(multi-level bypass)를 제공하여 정보가 계층적으로 전달될 수 있도록 모형을 개발하였다. 기존 유-넷의 형태와 마찬가지로 저차원 보틀넥 경로 외에 완전 합성곱 우회로를 제공하여 저차원에서도 일부 공간 정보가 유지될 수 있도록 구성하였다. 또한, 중간층에서 특성 우회로(feature bypass)를 통해 디코더로 중층의 정보를 전달하고, 입력층의 기후 정보 자료를 간단한 평균 풀링만 거친 후 디코더를 우회하여 최종 재귀 합성곱 층으로 전달함으로써 예측 성능의 향상을 꾀했다.

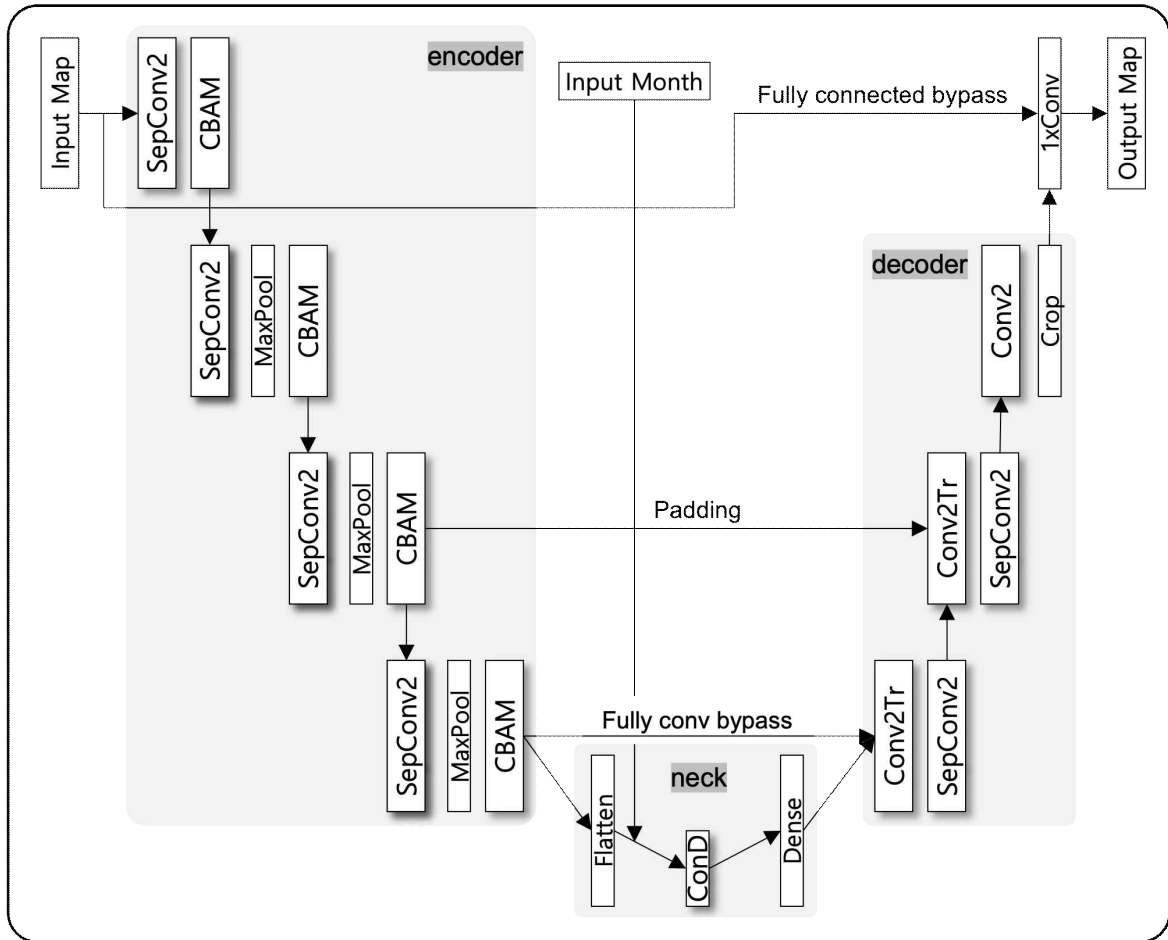


Figure 40. The overview of the customized artificial neural network developed for the post-processing of MME seasonal climate prediction.

위에서 열거한 맞춤형 인공지능망 모형의 전체 구성은 Figure 40와 같다. 다중모델앙상블 결정론적 단순 합성곱 계절 기후 예측 결과와 함께 해당 월 정보가 입력으로 전달되어, 최종적으로 관측 재분석 자료에 적합된 동아시아 맞춤형 삼분위 후보정 계절 기후 예측 결과가 생산된다.

## 라. 예측성을 향상을 위한 맞춤형 인공지능망의 전이 학습 전략

### (1) 장기 관측 재분석 자료를 이용한 사전 학습

APCC의 다중모델앙상블 계절 기후 예측 자료는 가용 기간이 길지 않아(버전에 따라 20-30년 내외), 많은 수의 모수를 추정해야하는 인공지능망 학습에는 적절치 않다. 따라서, 최대한 가용 기후 정보를 확보하기 위해 한 모형을 통해 모든 월을 처리할 수 있는 동적 조건부 드롭아웃 기술 등이 개발/적용되었다. 또한, 비교적 긴 기간을 포함하는 관측 재분석 자료의 활용을 시도하였다. 이 연구에서는 1948년부터 2020년까지의 NCEP/NCAR Reanalysis 1 (이후 NCEP1) 자료를 이용하여 상기 맞춤형 인공지능망 모형을 사전 학습시켜 예측 성능을 확보하였다. 사전 학습을 위한 입력 변수는 다중모델앙상블 계절 기후 예측 자료의 변수와 일치할 수 있도록 850 hPa 기온(air), 바람(uwnd, vwnd), 그리고 500 hPa 지위고도(hgt)를 이용하였으며, 출력 변수는 동일한 재분석 자료의 강수량(prate) 추정값을 사용하였다. 사전 학습의 검증 을 위해서 임의의 5 개 해(최종 모형의 학습을 위해서는 1982, 1992, 2002, 2012, 2019 년을 사용함)를 검증 세트로 사용하였으며, 이들은 모형의 학습에는 사용되지 않는다. 이를 제외한 학습 세트로부터 모멘트(평균과 표준편차)를 계산하고, 이를 기준으로 아노말리를 계산하였다. 검증 자료의 아노말리 계산에도 동일한 학습 기반의 모멘트가 사용되었다. 그 밖에도, 학습 속도를 향상시키기 위한 약간의 최적화 작업이 수행되었고, 학습률을 조정해가며 모형이 안정화될 때까지 반복적으로 학습을 수행하였다.

사전 학습을 통해 검증 자료를 예측한 결과 삼분위 예측에서 평균적으로 약 45% 정도의 정답률을 기록하였다. 임의 추출시 성능이 약 33%, 그리고 APCC의 다중모델앙상블 확률 예측의 정답률이 약 38% 인 것을 감안하면, 사전 학습을 통해 모형이 기후 정보를 잘 학습한 것으로 판단할 수 있다. 특히, 동적 조건부 드롭아웃을 적용함으로써, 하나의 모형을 활용하여 전계절을 학습시키는 다중 작업 처리 학습(multi-task learning)이 가능하였고, 이를 통해 월별로 안정적인 예측을 수행할 수 있었다.

### (2) 다중모델앙상블 단순합성 자료와 강수 재분석 자료를 이용한 미세 조정 학습

사전 학습이 완료된 맞춤형 인공지능망 모형의 동적 조건부 드롭아웃 보틀넥을 제외한 모든 학습 모수를 동결하고 APCC의 다중모델앙상블 결정론적 단순합성 계절 기후 예측 자료를 입력으로 사용하고 CPC Merged Analysis of Precipitation (이하 CMAP)을 삼분위 범주로 환산한 강수 자료를 출력으로 사용한 미세 조정 학습을 수행하였다. 미세 조정 학습에서는 월별로 다른 모수를 학습하여, 계절에 특화된 예측 보정을 수행할 수 있도록 조치하였다. 따라서, 기본 맞춤형 인공지능망 모형의 구조는 계절과 상관없이 동일하지만, 학습된 모형의 모수는 월별로 다른 형태를 가지게 된다. 입력 자료는 APCC-MME version 1의 u850, v850, t850과 z500이며, 출력 자료는 CMAP의 precip이 된다. 미세 조정을 위한 학습 자료는 1991년부터 시작하며, 안정적인 모형의 검증을 위하여 leave-one-out 교차 검증을 수행하였다.

## 마. 맞춤형 인공지능망 후처리를 통한 동아시아 예측성 향상

### (1) 맞춤형 인공지능망 후처리 모형을 이용한 예측 결과의 보정 예시

이 연구를 통하여 개발된 맞춤형 인공지능망 후처리 모형을 장기 관측 재분석 자료를 바탕으로 사전 학습시키고 APCC의 다중모델앙상블 계절 기후 예측 결과 및 CMAP 강수 추정 자료를 통해 미세조정 학습시킨 후 leave-one-out 교차 검증을 통해 예측 정보를 생산하였다. Figure 13과 같이 보정된 예측 결과는 삼분위 확률값(좌)으로 표현되며, 이는 CMAP 강수(우)의 삼분위 범주(중)와 공간적으로 매우 유사한 패턴을 보여, 맞춤형 인공지능망 후처리 모형이 적절한 보정을 수행하고 있는 것으로 나타난다.

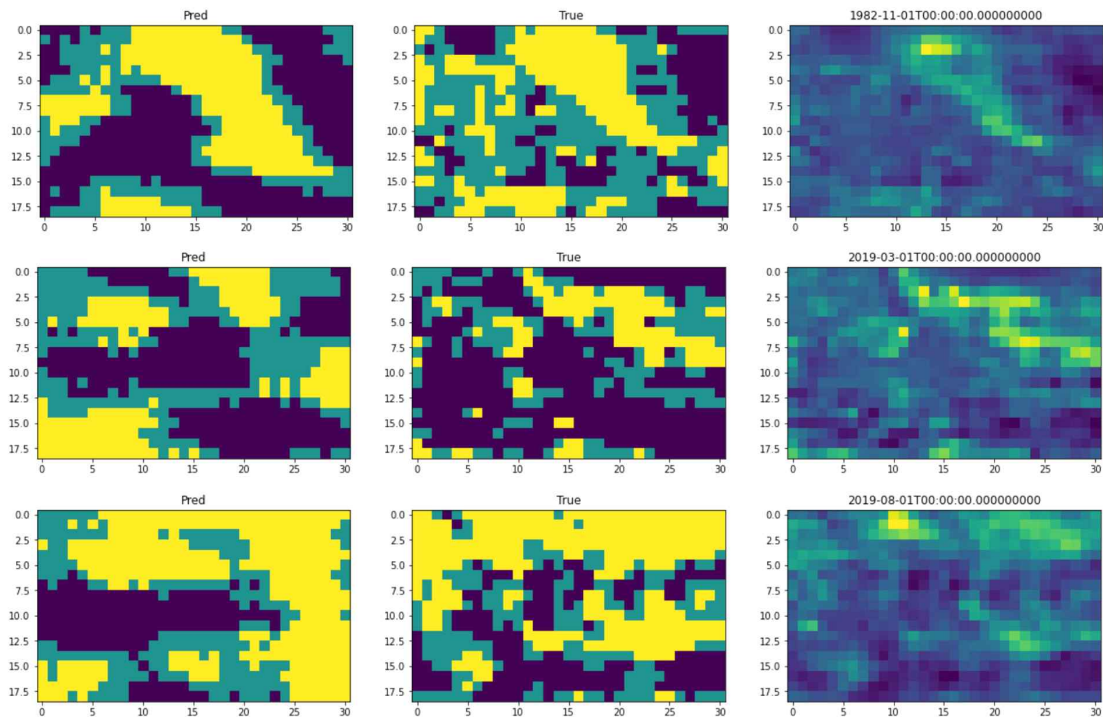


Figure 41. Examples of prediction (left), true observed true category (middle), and observed raw values (right) for selected months (1982-11, 2019-03, 2019-08).

### (2) 맞춤형 인공지능망 후처리를 통한 예측 성능의 향상과 안정화

기존 APCC의 다중모델앙상블 가우시안 확률 예측 결과(Figure 42; 이후 APCC-PMMEv1)와 비교하여 후보정 효과(Figure 43)를 살펴보았다.

(가) 기존 APCC-PMMEv1의 예측 성능

APCC-PMMEv1의 범주 적중률이 평균 38.7% 정도로 나타났으며, 엘니뇨의 활동이 강한 해(1997, 1998, 2010 등)의 예측 성능은 비교적 양호한 한 편, 비-엘니뇨해의 예측 성능은 떨어지는 모습을 보인다. 계절적으로도 예측성 장벽(predictability barrier)로 알려져있는 봄-여름 전이 기간에 적중률이 약간 하락하는 경향이 관찰된다. 많은 경우 적중률이 40%를 넘지 못하나, 강한 엘니뇨의 발달/전성기에는 50%를 넘는 적중률을 보이기도 한다.

ave(HR)	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	Annual	
1991	41.77%	39.73%	37.69%	36.33%	36.33%	35.99%	43.12%	31.75%	35.31%	37.01%	41.09%	37.83%		
1992	37.18%	43.12%	44.31%	41.09%	32.94%	36.84%	33.11%	38.20%	41.94%	33.45%	23.26%	26.83%	36.02%	
1993	37.18%	40.24%	44.14%	44.99%	45.50%	45.50%	41.94%	38.88%	35.31%	39.90%	27.33%	26.98%	38.24%	
1994	29.88%	30.73%	33.11%	33.62%	28.01%	38.20%	41.60%	40.58%	48.22%	49.41%	43.63%	36.84%	37.82%	
1995	33.62%	51.95%	43.80%	39.90%	31.75%	35.82%	35.14%	30.90%	35.65%	43.97%	50.25%	57.56%	40.86%	
1996	43.29%	43.63%	36.50%	38.20%	36.16%	39.39%	28.88%	36.33%	43.46%	37.01%	34.97%	43.63%	38.46%	
1997	44.31%	46.01%	39.73%	36.84%	32.26%	40.24%	37.52%	43.97%	38.37%	37.86%	48.73%	47.37%	41.10%	
1998	42.95%	47.54%	46.52%	46.18%	51.10%	39.73%	48.73%	37.52%	41.60%	51.10%	55.25%	64.18%	47.72%	
1999	48.22%	46.01%	40.24%	41.94%	40.07%	30.73%	34.63%	36.16%	39.56%	43.46%	37.35%	38.88%	39.77%	
2000	36.33%	41.09%	46.87%	36.67%	38.88%	35.99%	33.78%	27.67%	28.55%	23.60%	33.62%	34.63%	34.68%	
2001	37.18%	33.45%	39.39%	41.26%	40.07%	44.14%	42.28%	40.07%	48.56%	33.79%	35.48%	44.14%	39.98%	
2002	42.44%	33.28%	40.92%	36.33%	31.24%	37.35%	37.18%	33.45%	41.43%	45.50%	42.44%	41.43%	38.58%	
2003	33.62%	29.71%	33.28%	27.16%	35.48%	42.78%	37.01%	40.41%	31.07%	31.07%	31.24%	34.80%	33.97%	
2004	38.38%	38.38%	44.55%	36.33%	39.39%	43.12%	31.58%	35.65%	41.94%	41.60%	44.99%	41.09%	38.26%	
2005	40.41%	38.71%	43.97%	37.35%	32.09%	31.82%	31.58%	31.24%	35.99%	37.69%	27.84%	30.03%	34.90%	
2006	31.58%	32.77%	35.48%	29.03%	26.99%	39.05%	40.24%	35.48%	40.07%	38.37%	31.58%	32.43%	34.42%	
2007	35.48%	31.41%	32.43%	39.39%	35.65%	32.09%	30.90%	43.29%	45.16%	46.86%	42.61%	40.07%	37.95%	
2008	31.58%	48.73%	42.95%	42.16%	45.67%	40.07%	42.28%	39.22%	34.13%	36.67%	38.88%	41.94%	40.41%	39.70%
2009	39.22%	33.28%	41.60%	36.16%	39.39%	43.12%	37.52%	43.80%	39.90%	41.26%	51.27%	41.94%	39.32%	
2010	53.65%	39.05%	53.14%	43.97%	35.14%	30.56%	43.97%	53.14%	44.48%	46.69%	44.31%	42.11%	44.19%	
Seasonal	38.25%	39.14%	41.08%	38.01%	36.13%	37.88%	37.14%	38.20%	39.47%	39.39%	39.27%	40.32%	38.69%	

Figure 42. The seasonal and annual hit rates (in %) of APCC-PMMEv1.

ave(HR)	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	Annual
1991	39.90%	41.94%	38.03%	39.56%	37.52%	44.31%	42.28%	41.26%	39.73%	39.90%	40.41%	40.44%	
1992	48.82%	43.97%	46.35%	42.44%	45.67%	40.07%	42.28%	39.22%	43.97%	43.97%	37.86%	43.48%	43.48%
1993	43.63%	41.77%	49.07%	46.52%	45.84%	41.60%	42.44%	38.71%	37.86%	37.01%	39.22%	39.90%	41.96%
1994	43.80%	38.88%	39.56%	39.39%	46.86%	42.44%	45.33%	43.29%	42.95%	40.75%	45.67%	39.05%	42.33%
1995	42.44%	42.11%	41.09%	48.22%	43.46%	42.61%	41.43%	39.39%	40.75%	39.90%	45.50%	58.74%	43.80%
1996	38.54%	43.46%	41.60%	44.65%	37.69%	40.07%	36.84%	39.73%	38.54%	37.86%	43.46%	46.33%	40.73%
1997	39.73%	46.69%	44.31%	37.35%	43.46%	43.12%	40.41%	50.25%	47.88%	36.16%	44.99%	46.35%	43.39%
1998	42.95%	50.93%	47.20%	48.05%	44.99%	44.82%	41.60%	44.99%	42.11%	41.94%	45.16%	45.33%	45.01%
1999	45.33%	46.35%	41.09%	36.84%	41.26%	40.07%	38.88%	40.07%	43.29%	37.01%	42.11%	41.94%	41.19%
2000	41.94%	46.35%	41.60%	42.11%	39.05%	39.05%	43.97%	43.97%	40.41%	43.46%	48.38%	43.63%	42.61%
2001	46.18%	50.75%	40.58%	43.12%	42.61%	43.97%	44.48%	41.43%	40.41%	42.28%	42.11%	41.77%	43.31%
2002	45.50%	46.69%	45.16%	45.33%	39.73%	37.86%	39.39%	41.94%	39.05%	43.97%	46.01%	37.86%	42.37%
2003	45.33%	43.63%	41.77%	43.63%	46.01%	43.63%	39.90%	45.33%	43.46%	40.07%	37.18%	39.73%	42.47%
2004	44.63%	43.46%	42.95%	40.07%	42.69%	40.07%	36.84%	39.73%	40.41%	37.18%	39.90%	42.78%	40.92%
2005	40.75%	37.86%	44.31%	40.75%	44.48%	44.99%	38.54%	37.01%	40.07%	39.39%	37.69%	40.92%	40.56%
2006	41.94%	41.60%	37.01%	35.65%	38.03%	41.09%	34.97%	45.16%	35.82%	47.88%	40.41%	44.48%	40.34%
2007	44.31%	45.84%	42.78%	43.46%	40.41%	38.54%	38.10%	42.28%	42.26%	53.99%	43.97%	47.20%	44.68%
2008	40.07%	44.14%	41.09%	41.77%	43.29%	38.03%	39.73%	44.99%	46.26%	49.75%	42.95%	42.11%	42.87%
2009	38.37%	39.22%	41.77%	41.26%	39.39%	45.84%	38.71%	37.35%	39.39%	36.84%	48.22%	48.38%	41.23%
2010	41.43%	45.16%	45.33%	40.07%	42.61%	39.22%	41.09%	44.14%	42.11%	57.39%	47.37%	40.07%	43.83%
Seasonal	42.93%	44.10%	42.84%	42.09%	42.52%	41.23%	40.94%	42.06%	41.31%	42.17%	42.66%	43.73%	42.38%

Figure 43. The seasonal and annual hit rates of the custom network.

(나) 맞춤형 인공지능망을 이용한 후처리에 따른 예측 성능

Figure 43에서 나타난 후처리 적중률은 대부분의 경우 40%를 상회하며, 평균 42.4%의 범주 적중률을 기록하고 있다. 이는 기존 APCC-PMMEv1의 예측 적중률인 38.7% 대비 약 9.5%의 평균 예측이 상승한 것이다. 이러한 예측 성능은 연도별(Figure 44)로나 계절별(Figure 45)로 나누어 살펴보았을 때에도 비교적 안정적으로 유지되고 있는 것으로 보인다. 특히, 엘니뇨-남방진동의 활동이 크지 않은 해의 예측 성능이 기존 APCC-PMMEv1과 같이 크게 하락하지 않고 꾸준히 유지되는 것을 확인할 수 있다. 다만 강한 엘니뇨 발달/전성기에는 기존 APCC-PMMEv1의 예측 성능이 약간 높은 것을 확인할 수 있다. 흥미롭게도 일반적인 계절 기후 예측 모델이 가지고 있는 특징인 예측성 장벽이 맞춤형 인공지능망 후처리 모형에서는 발견되지 않는다는 점(Figure 45)이다. 이 연구를 통해 개발한 후처리 모형은 엘니뇨-남방진동의 위상전이가 나타

나는 봄-여름철 예측 적중률이 다른 계절과 마찬가지로 꾸준히 유지되는 안정적인 모습을 보인다.

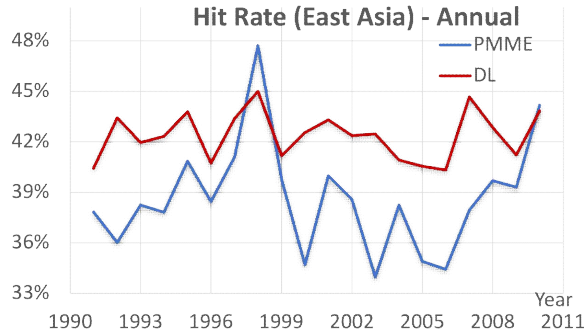


Figure 44. Annual hit rates for APCC-PMMEv1 (red) and the custom network (blue).

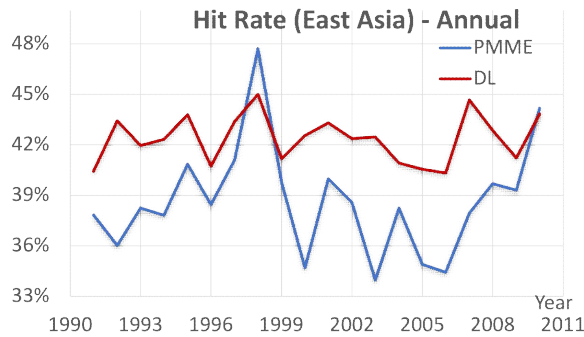


Figure 45. Annual hit rates for APCC-PMMEv1 (red) and the custom network (blue).

전체적으로 기존 APCC-PMMEv1 대비 약 9.5%의 향상된 예측 성능인 평균 42.4%의 적중률과 연도별/계절별 예측 성능의 부침이 크지 않은 맞춤형 인공지능기후 후처리 모형은 실제 안정적이고 정확한 계절 기후 예측에 중요한 기여를 할 수 있다.

## 5. 동아시아 확률예측을 위한 기법 간 상호 비교

### 가. 자료 및 방법

동아시아 맞춤형 확률예측 검증 시스템은 APCC의 PMME와 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술의 결과와 딥러닝을 활용한 동아시아 PMME 예측결과를 각각 상호비교하기 위해 설계되었다. 상호비교 검증 시스템은 강수량을 대상으로 작성되었고, 강수량 관측자료는 APCC의 현업시스템과 동일하게 CAMS-OPI 강수 자료를 활용하여 검증 시스템을 설계하였다. 확률 예측의 검증은 HSS와 RPSS를 활용하여 검증 시스템을 구축하였으며, Hindcast의 기간은 마찬가지로 APCC의 현업시스템과 동일하게 1991년부터 2010년까지로 정의하였다. HSS는 관측값에 대해 정확한 강우강도 자료와 우연히 맞은 강우강도 자료의 차이를 평가하기 위해 사용하는 일반화된 정확성 계산기법에 속한다. 3개의 tercile 기반 범주에 대한 확률 예측은 다음과 같은 식 [1]과 같다.

$$Heidke\ skill\ score = 100 \left( \frac{Hits - Expected}{Total - Expected} \right) \quad [1]$$

RPSS (Ranked Probability Skill Score)는 예측의 범주형 확률과 관측의 범주형 확률 간의 누적 제공 오차를 측정하는 방법으로, 관찰된 범주형 확률은 관찰된 범주에서 100%이고 다른 모든 범주에는 0%로 정의한다. RPSS 산출을 위해서는 다음의 식[2], 식[3]과 같다.

$$RPS = \sum_{cat=1}^{Ncat} (Pcum_{F(cat)} - Pcum_{O(cat)})^2 \quad [2]$$

$$RPSS = 1 - \frac{RPS_{forecast}}{RPS_{reference}} \quad [3]$$

여기서 Ncat의 경우 3분위 tercile의 경우 3이 된다. “cum”은 합산이 먼저 1에 대해 수행한 다음 1과 2에 대해 그리고 1과 2 그리고 3에 대해 수행됨을 의미한다. RPS가 높을수록 예측은 좋지 않음을 의미하고, RPS가 0일 경우 관측과 예측이 같음을 의미한다. 실제 예측에 대한 RPS가 참조 예측에 대한 RPS보다 작은 경우 RPSS > 0 이다.

### 나. 검증 결과

Figure 46은 동아시아 확률 예측에 대한 상호비교 결과로 HSS와 RPSS에 대해 JFM부터 DJF까지의 12개월 결과를 보여준다. HSS를 기준으로 본 경우 동아시아 지역을 대상으로 한 APCC의 확률 예측 결과보다 물리현상 기반의 지역 특화 계절 예측 기술의 결과와 딥러닝을 활용한 동아시아 확률 예측 결과가 높은 성능을 보여준다. 딥러닝이 전반적으로 가장 우수한 성능을 보여주는 것으로 볼 수 있다. 하지만 RPSS에서는 딥러닝이 다른 두 방법에 비해 가장 낮은 성능을 보여주는 것을 알 수 있다. 이는 딥러닝을 활용한 동아시아 확률 예측의 결과의 특징이

Above, Near, Below를 나타내는 확률값이 매우 강하게 나타남이기 때문인데, 이를 Figure 47을 통해 알 수 있다. Figure 47은 1992년, 1998년, 1999년, 2000년, 2008년의 JFM을 기준으로 각 3가지 방법의 확률 예측 결과를 나타낸다. Figure 48은 같은 년도의 JJA을 기준으로 각 3가지 방법의 확률 예측 결과를 나타낸다. 해당 그림에서 알 수 있듯이 딥러닝을 활용한 확률 예측 결과가 해당 normal을 강하게 나타내는 것을 확인할 수 있다. RPSS는 여러 주요 요인에 의해 낮아질 수 있는데, 그 중 평균 확률의 편향이 딥러닝을 통한 예측 결과에서 RPSS를 낮게 나타내는 요인이 된 것으로 사료된다. 이는 딥러닝을 계절 예측 분야에 적용함에 있어서 지속적인 연구가 필요함을 시사한다고 사료된다.

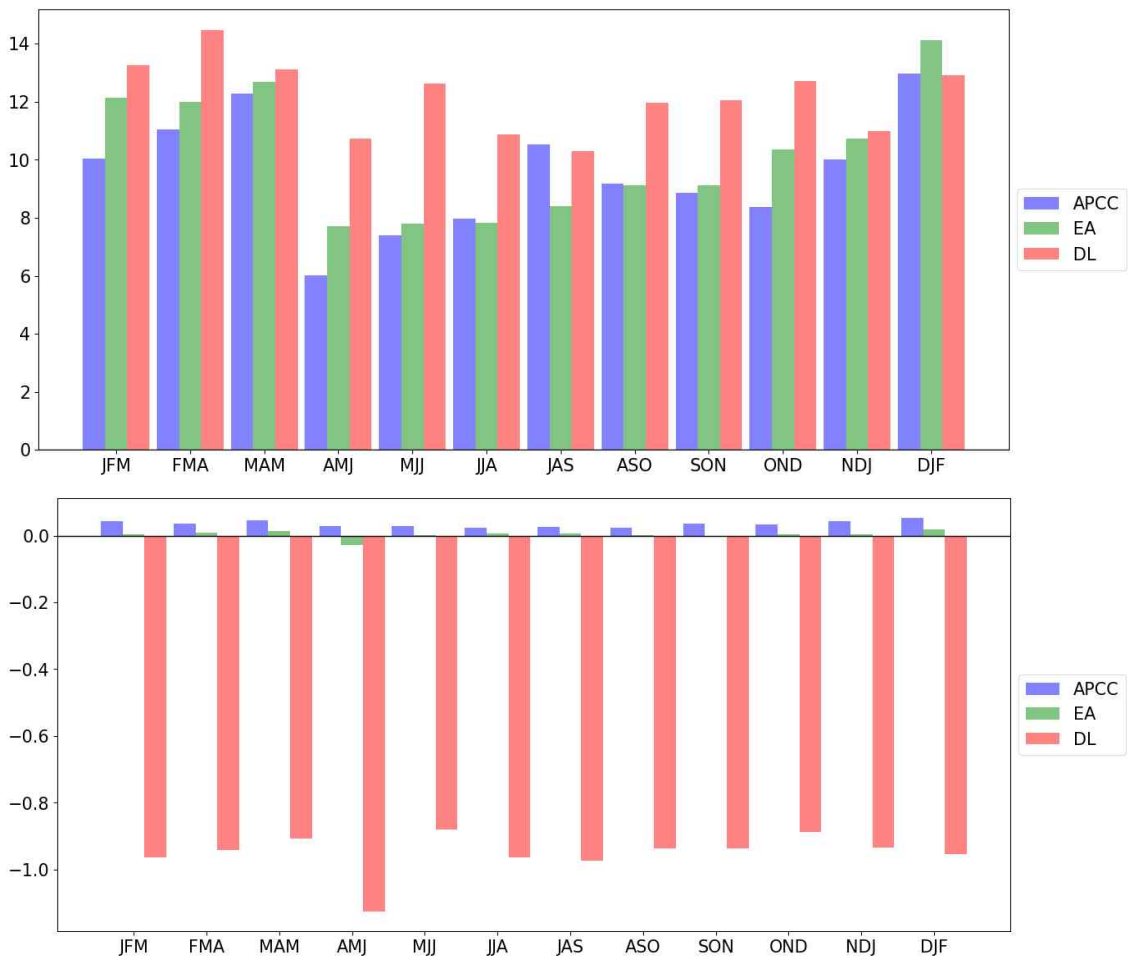


Figure 46. East Asia probabilistic prediction intercomparison verification results (HSS:upper, RPCC:lower).

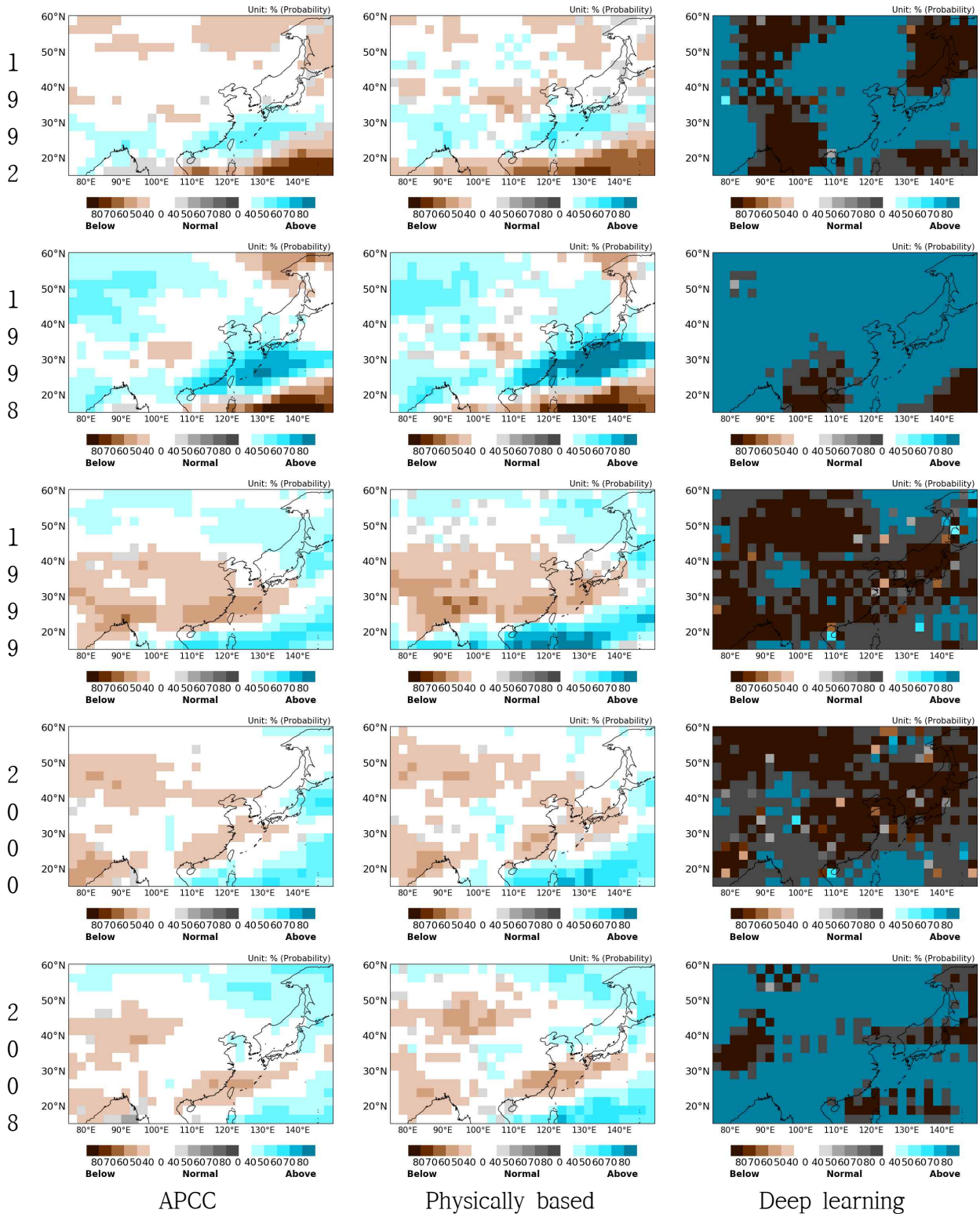


Figure 47. Probability prediction regional distribution comparison result (JFM).

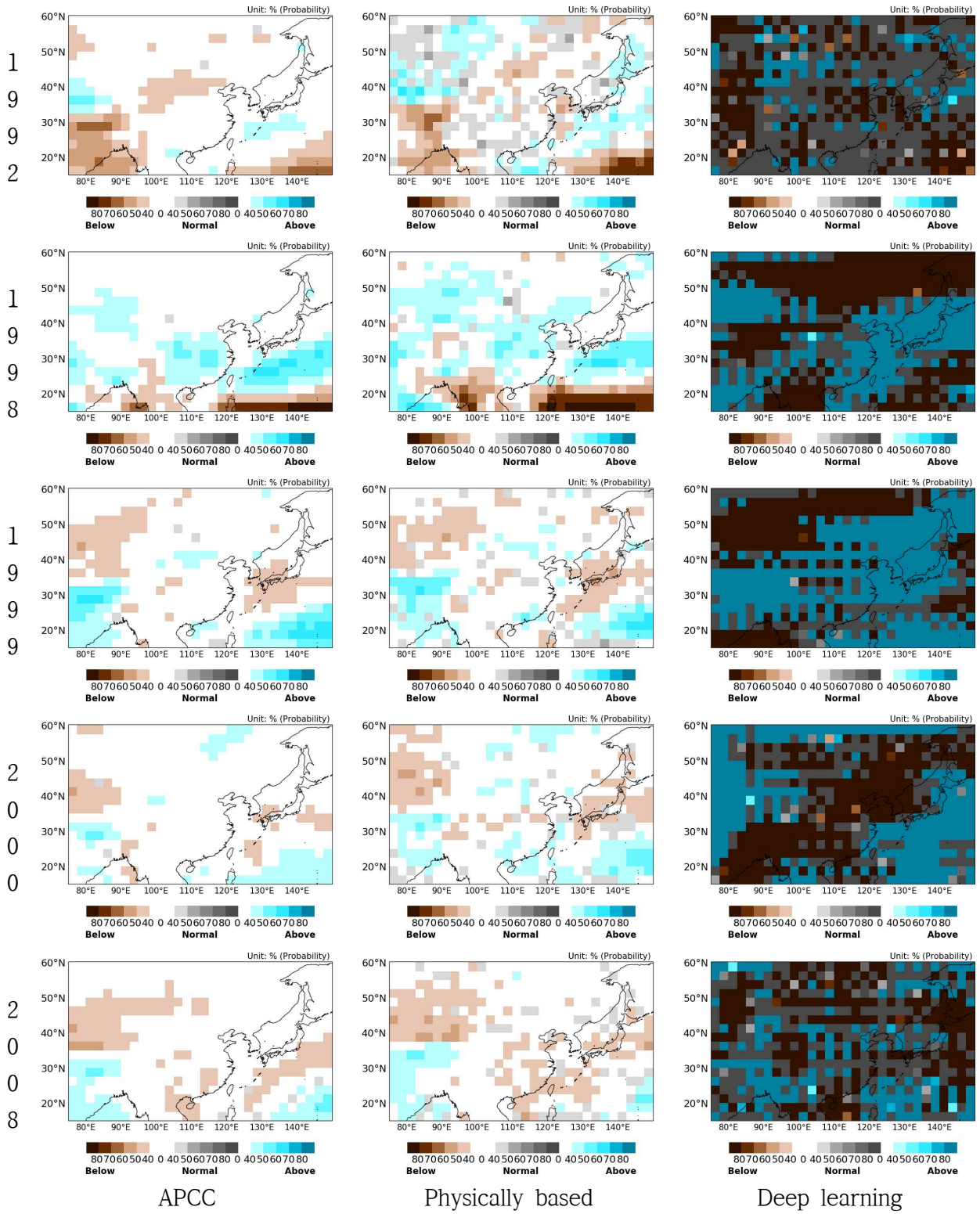


Figure 48. Probability prediction regional distribution comparison result (JJA).

## 6. 다중모델 계절내예측 기반기술 개발: 머신러닝 및 딥러닝 복합 활용 기법 개발

### 가. 개요

계절 및 계절내 예측 개선을 위한 딥러닝 기법 도입 및 연구는 지난 몇 년간 폭발적으로 증가해왔다(Chattopadhyay et al., 2020; Hernández et al., 2016; LeCun et al., 2015; Miao et al., 2019; Rasp et al., 2018; Reichstein et al., 2019). 이에 본 연구도 2019년부터 계절내 예측 개선을 위한 기반 기술 개발로 딥러닝 기술 적용 연구를 수행하기 시작하였다. 먼저 한반도 지점에 대해 2개 개별 기후모델(ECMWF 및 GloSea5)의 계절내(Sub-seasonal to seasonal, 이하 S2S) 예측 시계열 자료에 장단기기억훈련법(Long-Short Term Memory, 이하 LSTM)을 적용하여, 목표 선행예측기간(week3-4) 내 자료의 예측성을 향상시킬 수 있음을 확인하였다. 또한 여러 기후모델을 한번에 고려하는 다중모델앙상블(multi-model ensemble, 이하 MME) 기법은 기후 예측의 신뢰성을 증가시킨다고 알려져 있다(Park et al., 2018; Vitart and Robertson, 2018; Wang et al., 2020). 따라서, 참여 개별 기후모델의 확대(UKMO, CMA, MSC, NCEP 추가)를 통해 MME 기반 S2S 예측 자료를 구축하고, LSTM 적용을 통해 3-4주 내 MME S2S 예측 기후의 예측 오차를 개선하였다. 또한 LSTM 피쳐엔지니어링 민감도 분석을 통해 S2S 자료의 예측 정확도를 개선하고, 격자 기반의 학습 알고리즘 개발을 위해 합성곱 신경망 모델(Convolutional Long-Short Term Memory, 이하 ConvLSTM) 프레임워크를 도출하였다(Shi et al., 2015).

그러나, 온전한 한반도의 예측 기후 정보를 신경망 모델에 입력하기 위해서는 대상 지역이 동아시아로 확장된 기후 변수를 비롯한 기후 인자까지 고려한 MME 기반 S2S 예측 기후 자료의 개선이 필요하였다. 또한 격자 기반의 ConvLSTM 프레임워크가 도출되었기는 하지만 모델에 입력되는 훈련 자료의 범위가 동아시아로 확장된 만큼 신경망 모델의 높은 분산을 줄이고 예측 성능을 강화시킬 수 있는 신경망 모델의 개발 및 개선이 요구되었다. 신경망 모델의 높은 분산을 줄이는 방법으로 단일 신경망 모델 대신 여러 신경망 모델로부터 학습시켜 그 예측 값을 결합하는 앙상블 학습 기법이 활용되고 있다(Thongsuwan et al., 2021; Wang et al., 2019).

따라서, 본 연구는 입력자료 개선을 위해 한반도를 포함한 동아시아 지역 기반의 개별 기후 모델 및 MME 기반 S2S 예측 자료를 생산하고, 딥러닝 모델의 분산, 예를 들면 불확실성을 줄이기 위해 기계학습 및 신경망 알고리즘 결합을 통해 단일 모델 적용에서 나타날 수 있는 분산 혹은 불확실성을 줄일 수 있도록 하이브리드 알고리즘을 구축하였다. 또한, 하이브리드 알고리즘내 동아시아 개별 기후모델 및 MME 계절내 예측 입력자료의 민감도를 분석하고, 신경망 알고리즘의 주요 하이퍼파라미터 선정을 통해 하이브리드 모델을 최적화하였다. 생산된 동아시아 지역 MME 기반 S2S 예측 자료에 하이브리드 모델을 적용하여 그 예측 성능을 평가하였다.

## 나. 동아시아 지역 계절내 MME 예측 자료 구축

동아시아 지역 계절내(Sub-seasonal to seasonal, 이하 S2S) MME(multi-model ensemble, 이하 MME) 예측 자료를 구축하기 위한 개별 기후모델은 Table 8과 같다. 도메인은 경도 60에서 150, 위도 -21에서 51의 1.5도로 y축(경도)으로 61개, x축(위도)로 49의 격자를 가진 입력자료로 확장되었다(Figure 49). 따라서 기후모델의 앙상블 변수가 급격히 증가됨으로써 발생하는 컴퓨팅 계산 문제를 해결하기 위해, 선례 연구와 같이 기후모델의 앙상블(control과 perturbation)과 상관없이 기후 변수들이 평균되도록 처리하였다. 예를 들면, ECMWF 기후모델의 경우 control에 대한 17개 기후 변수와 perturbed에 대한 17개 기후 변수는 과거 재현에 대한 11개 앙상블에 따라 총 187개로 구성될 뿐만 아니라, 이것이 전체 도메인 격자(187x49x61)로 구성되기 때문에, read/open 하는데 어려움이 발생하였다. 따라서 입출력 문제 해결을 위해 11개의 앙상블에 상관없이 17개 기후 인자별로 평균하였다. 마찬가지로 NCEP, UKMO, CMA 기후 모델도 과거 재현 앙상블이 각각 4, 7, 4개로 앙상블에 상관없이 기후 인자별로 평균되도록 처리하였다. 반면, KMA(이하 GloSea5) 기후모델은 앙상블이 3개임에도 선례 연구와의 일치성으로 평균하지 않고, 앙상블에 따라 변수가 각각 출력되도록 처리하였다.

구축된 개별 기후모델의 S2S 일최고기온 예측 자료로부터 동아시아 지역 MME 기반 S2S 훈련 자료를 생산하였다. 5개 개별 기후모델(ECMWF, NCEP, GloSea5, UKMO, CMA)의 과거 재현 기간 전체가 포함되는 범위는 1999년에서 2010년으로 12년에 불과하기 때문에 이 기간을 MME 기간으로 설정할 수 없었다. 선례 연구에서 기계학습 및 심층학습을 위한 훈련 및 검증 기간 전체를 포함하는 기간에 대해 20년을 적정 기간으로 채택한 바가 있어, 본 연구에서도 훈련 및 검증의 전체 적정 기간을 20년으로 설정하여, 5개 개별 기후모델이 적절하게 배치될 수 있는 기간으로 1994년에서 2014년을 선정하였다. 또한 5개 개별 기후모델의 공통 기후변수는 총 12개 (TMAX, TMIN, T2M, PREC, U50, U200, U850, V200, V850, W200, Z200, Z500)로 선정되었으며, 공통 예측 선행시간(time range)은 30일로 설정하였다. 또한 5개 개별 기후모델의 공통 예측 빈도(hindcast frequency)의 기준은 매월 4회로 1일, 9일, 17일 25일로 선정하였다. 관측 자료 즉 훈련 학습에서 레이블 및 참값으로는 입력해야 하는 정보는 재분석(Era-Interim) 자료로 구축되었다.

Table 8. Information of sub-seasonal to seasonal climate models: used to produce multi-model ensemble sub-seasonal to seasonal prediction data in this study.

Center		Hindcast Ens. size	Hindcast length	Hindcast frequency	Time range	Ensemble
KMA	Seoul	3	1991-2010	4/month (1,9,17,25)	60	MME (1994-2014)
UKMO	Exeter	7	1993-2015	4/month (1,9,17,25)	60	
CMA	Beijing	4	1994-2014	daily	60	
ECMWF	ECMWF	11	past 20 years	2/week (Mon, Thu)	46	
NCEP	Washington	4	1999-2010	daily	44	

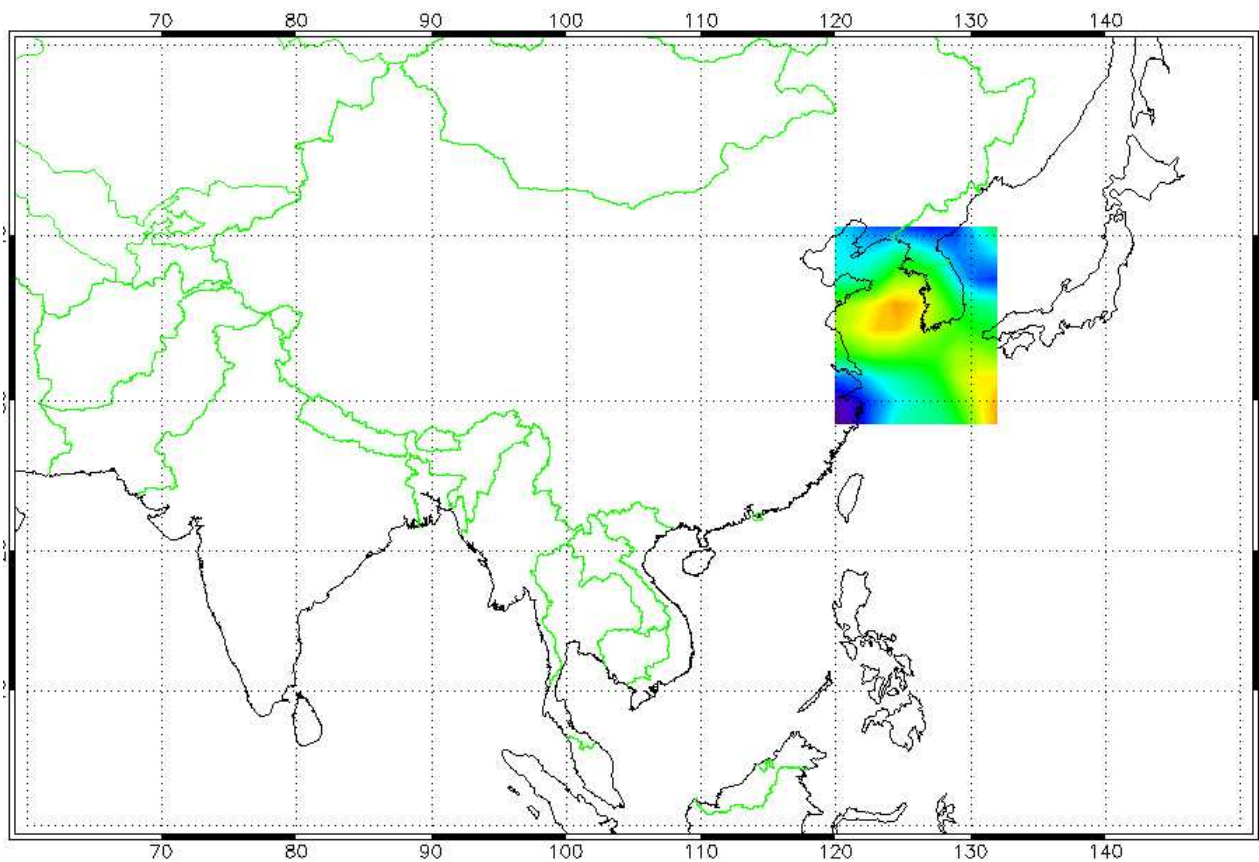


Figure 49. The geographic information of the input area in the study.

## 다. 머신러닝 및 신경망 모델 결합을 위한 딥러닝 하이브리드 알고리즘 구축

Figure 50은 기계학습 및 심층학습에서 많이 적용되는 대표 훈련 기법들을 도식화한 것이다 (<https://www.weareworldquant.com/en/thought-leadership/welcome-to-the-machine/>). 본 연구에서도 기계학습의 대표적인 훈련 기법으로 랜덤포레스트(random forest)를 선정하였다. Table 9는 심층학습에서 자료 형태에 따른 다양하게 적용되는 딥러닝 알고리즘들을 비교한 것이다 (Wang et al., 2019). 본 연구에서도 적용 자료 형태가 포인트 기반(point based)의 시계열(time series) 자료에서 격자 기반(spatial based)의 시계열로 확장하면서 장단기기억(long-short term memory, 이하 LSTM) 훈련 기법에서 합성곱 신경망모델(Convolutional Neural Network 및 ConvLSTM 등)로 알고리즘을 선정하였다.

### METHOD ACTING

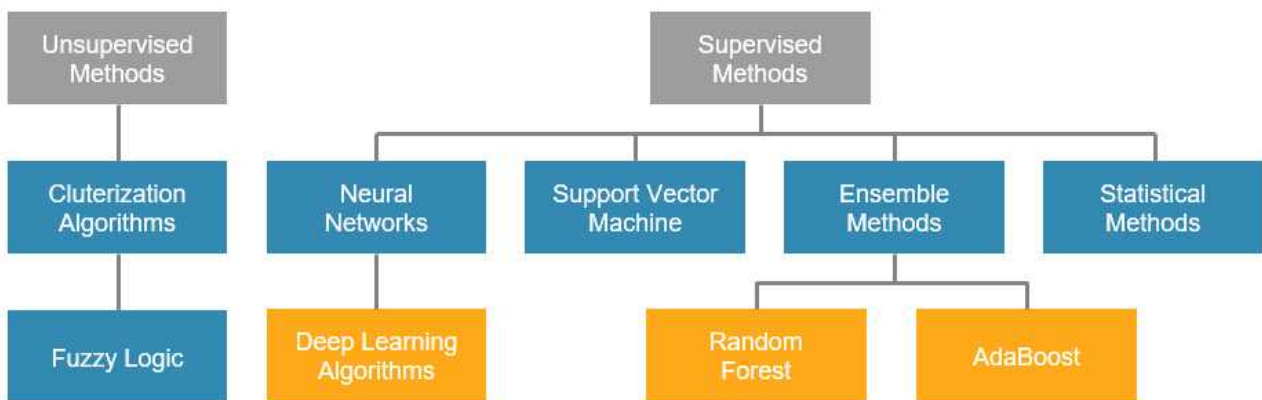


Figure 50. Machine learning typically follows one of two paths: the most popular methods are in orange color box.

Table 9. Different deep learning models for processing four types of spatial-temporal data: it shows which is learning technique of deep learning adopted in this study (modified from Source: Wang et al. (2019) Deep Learning for Spatio-Temporal Data Mining: A Survey).

Deep Learning Model	Data type	
	Time series	Spatial maps
RNN (LSTM, GRU)	●	●
Seq2Seq (Sequence to Sequence)	●	
AE (autoencoder/decoder)	●	
CNN		●
ConvLSTM		●
Hybrid	●	●

### (1) 랜덤포레스트 학습 모델

랜덤포레스트 기계학습 기법으로 대표적인 앙상블 모델이며, Gradient Boosting 기법과 같이 기본 요소는 결정 나무(이하 decision tree)이다(<https://www.datasources.ai/en/data-science-article/s/random-forests-for-complete-beginners>). 또한 잘 작동되는 tree를 찾아가면서 과적합(이하 overfitting)된 양을 줄여가는 기법으로 많이 알려져 있다(Figure 51 & 52).

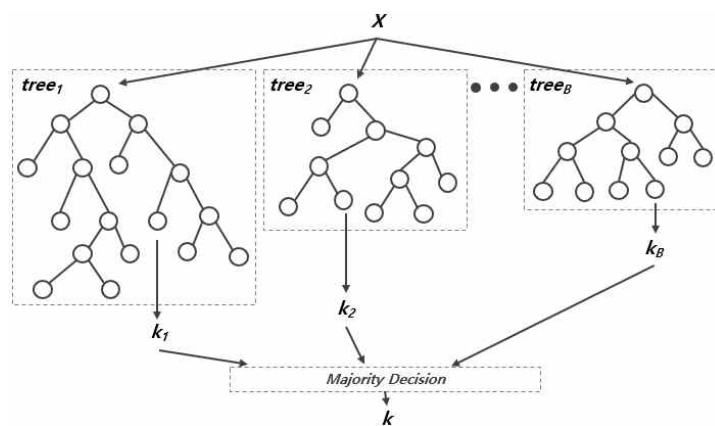


Figure 51. A commonly used class of ensemble algorithms are forests of randomized trees.

랜덤포레스트의 기본 요소는 decision tree로써, 이들은 조금씩 다른 여러 decision tree의 묶음으로, 각각의 tree 묶음들은 대부분 예측을 잘 할 수 있지만, 입력 데이터에 따라 overfitting하는 경향이 만들어지는 tree 들이 존재할 수 있는데, 이 tree들을 평균하면서 overfitting된 양을 줄여갈 수 있게 한다(Breiman et al.,1984; Misra and Wu, 2020). Figure 4에서 굵게 표시한 선(bold lines)들이 예측을 잘하면서 overfitting을 줄여가는 방향을 도식화하는

과정을 표현한 것이다.

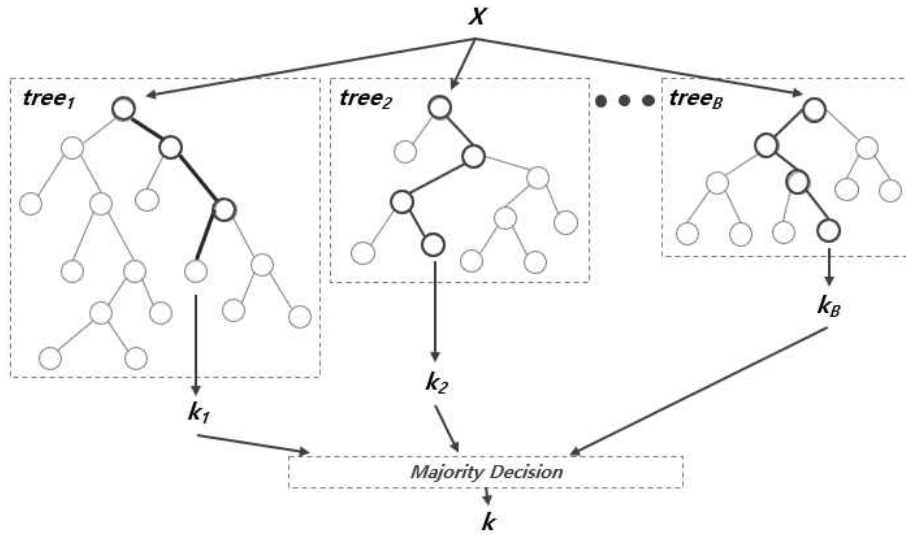


Figure 52. The process of the operation in the decision trees of random forest: Each tree in the random forest is built from a sample drawn with replacement from the training set.

랜덤포레스트에서 직접 결정해 줘야 하는 주요 초매개모수(이하, hyperparameter)는 tree를 몇 개 생성할지를 결정하는 n\_estimators, tree의 최대 깊이(max\_depth), tree의 split과 leaf(min\_sample\_split & min\_sample\_leaf) 등으로 알려져 있다(Table 10). 이러한 hyperparameters는 파이선에서 제공하는 내장함수(built-in function)인 GridSearch 함수를 이용하여 결정될 수 있다. Hyperparameters의 범위를 설정한 후, 내장함수 GridSearch 함수를 통해 랜덤포레스트의 hyperparameters를 도출하였다. min\_sample\_split과 min\_sample\_leaf는 tree가 전개해 나가는 방향을 결정하는데, 특히 min\_sample\_leaf의 node는 예측 성능의 개선 확률을 결정하는 중요한 역할을 한다.

Table 10. Key hyperparameters of random forest training model.

Parameter	Description
n_estimator	구축하려는 tree의 수
max_depth	tree의 최대 깊이를 결정
max_features	최상의 분할을 찾을 때 고려해야 할 특성의 수
min_sample_split	임의로 분리되어야 할 노드(leaf)의 최소 수
min_sample_leaf	마지막 노드(leaf)에서의 최소 샘플 수

랜덤포레스트가 많은 장점을 가진 기계학습임에도 불구하고 블랙박스 모형이기 때문에, 설명 변수와 반응 변수의 설명력을 확보하기 어렵다는 단점이 있다. 그래서, 이를 어느 정도 해결하기 위해 변수 중요도(feature importance)라는 척도를 통해 어느 변수가 예측 성능에 중요한 역할을 하는지를 설명한다. Tree가 어떻게 작동하는지를 요약하기 위한 변수 중요도를 계산하는 방법에는 3개가 있는데, 첫 번째는 MDI(mean decrease in impurity) importance, 두 번째는 permutation importance, 3번째는 drop column importance가 있다(Louppe et al., 2013). 본 연구에서는 MDI importance를 계산하였다(Eq. [4]). MDI importance는 가장 대표적인 변수 중요도로서, 역시 파이선에서 scikit-learn의 default로 내장되어있는 함수이다.

$$\Delta i(t) = i(t) - \frac{N_{il}}{N_t} i(t_l) - \frac{N_{ir}}{N_t} i(t_r) \quad [4]$$

Eq. [4]은 각 변수가 split될 때 impurity 감소분의 평균을 중요도로 정의한다. Eq. [4]에서  $i(t)$ 는  $t$ 노드의 impurity(entropy, gini index, variance) 이고,  $N_t$ 는  $t$ 노드의 관측치 개수를 나타낸다. Importance score(dimensionless)의 0은 전혀 사용되지 않았음을, 1은 완벽하게 예측에 사용되었음을 나타낸다. 그러나 이 feature importance score가 낮다고 해서 이 특성이 유용하지 않다는 뜻은 아니다.

## (2) 하이브리드 학습 모델

일반적으로 현실 세계의 정보량은 모델이 특정 문제의 핵심 기능을 발견하는 것을 어렵게 만든다. 이러한 이유로 일반적인 방법은 수동 기능의 피쳐 엔지니어링(feature engineering)에 의존하며, 이는 외부 환경이 일부 기능을 편향시키고 부적절한 기능 추가 또는 삭제로 이어질 수 있기 때문에 최적의 해결책에 도움이 되지 않을 수 있다.

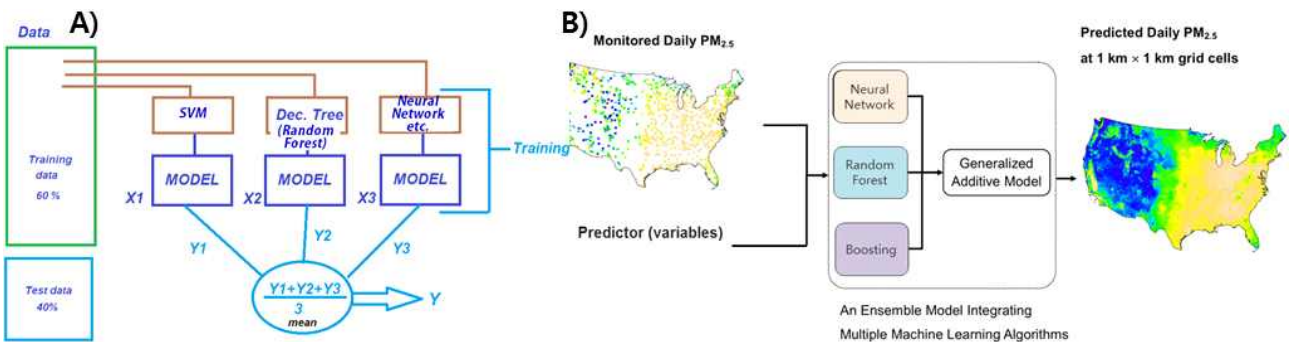


Figure 53. The concept of ensemble learning algorithms: A) is describes the process of combining learning and prediction results respectively by multiple learning algorithms. B) describes a model in which multiple learning algorithms are combined to learn and predict(modified from source: Di et al. (2019) An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution).

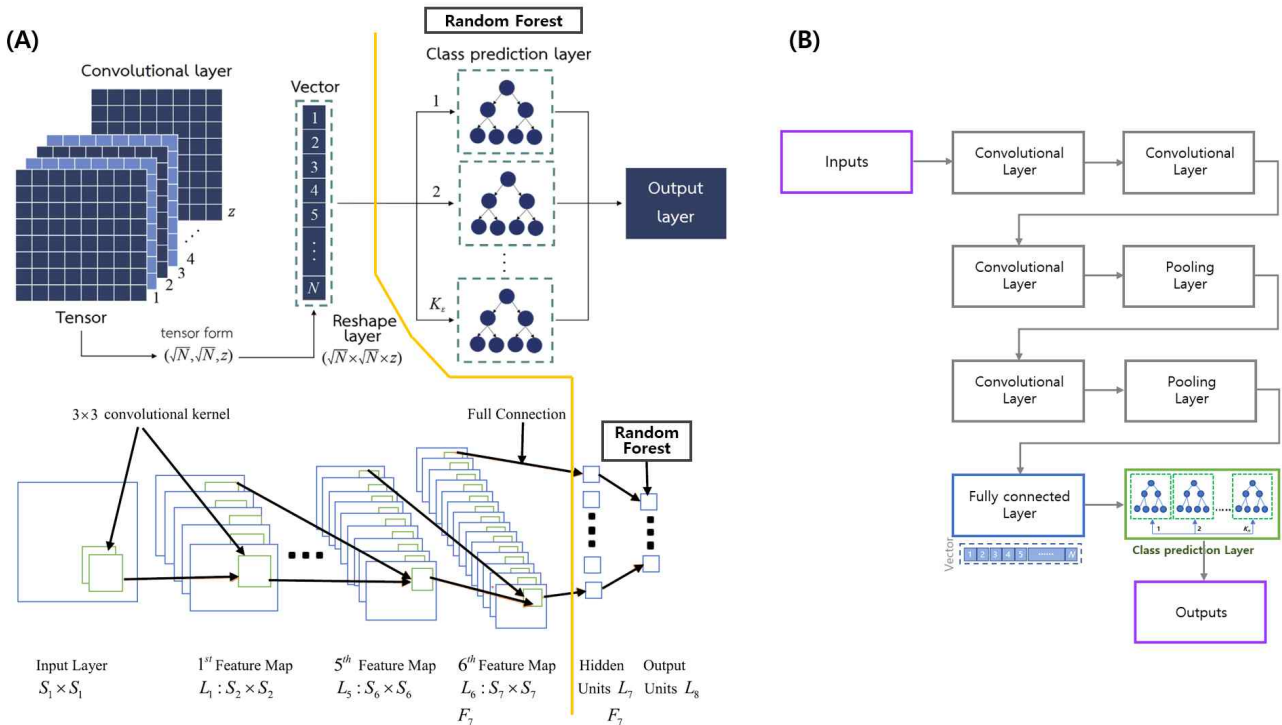


Figure 54. Diagram of hybrid architecture of this study: (A) work flow for integrating machine learning and deep learning algorithms and (B) simple architecture of convolutional neural network with ensemble classifier.

Figure

반면, 자동 기능 학습은 이 문제를 해결할 수 있는 것으로 알려져 있는데, 원시 데이터에서 주요 특징 혹은 주요 기능 등을 자동으로 검색하게 한다. 따라서 효과적인 모델의 중요성과 완전한 모델을 찾기 위한 자동 기능 학습 기능을 잘 활용할 필요가 있다. 그러나 기존의 머신러닝 모델과 딥러닝을 포함한 대부분의 모델은 여전히 단일 모델만 훈련에 사용한다.

Figure 53는 본 연구에서 여러 참고문헌(Di et al., 2019)을 기반으로 계획한 초기 앙상블 학습 알고리즘에 대한 예시이다. Figure 53의 A는 여러 학습 알고리즘에 의해 학습된 후 각각의 예측 결과를 결합하는 과정, 다시 말해서 여러 학습 알고리즘에 의해 학습 및 예측된 결과를 각각 결합하는 과정을 설명한다. Figure 53의 B는 여러 학습 알고리즘이 결합되고, 그 결합된 알고리즘에 의해 학습 및 예측되는 모델을 설명한 것이다.

Figure 54은 본 연구에서 최종 활용한 앙상블 학습 모델(이하 하이브리드 모델)을 설명하는 모식도이다. Figure 54의 A에서 오렌지선을 기준으로 좌측은 하이브리드 결합을 위한 기본 합성곱 장단기 기억의 아키텍처를 설명한다. 합성곱 알고리즘(ConvLSTM)의 역할은 공간/시간적 중요한 특성을 추출하고, 이것을 완전연결 레이어에 담는다. Figure 54의 A에서 오렌지선 우측은 완전연결 레이어의 정보를 랜덤포레스트 알고리즘에 입력하여 결정 트리로부터 최종 예측하는 아키텍처를 설명한다.

Figure 54의 B는 Thongsuwan et al.(2021)과 Weldegebriel et al.(2020)의 CNNXGBoost 모델

을 수정하여 본 연구에서 하이브리드 아키텍처를 간략하게 설명한 것이다. 합성곱 신경망은 모델에 대한 사전 지식을 수용하기 위해 이미지의 기하학을 탐색하도록 설계된 특별한 종류의 인공신경망으로, 합성곱 신경망은 이미지 데이터 세트에서 더 깊은 기능을 학습하기 위해 대부분 계층적 아키텍처로 구축된다. 합성곱2D(Conv2D) 레이어 3개와 같은 차원(2D)의 maxpooling 레이어 및 완전연결(fully-connected, dense) 레이어로 연결하였으며, 깊이(depth) 및 층별로 다른 kernel size(2, 3, 5 등)를 통해 연산량 조절이 가능하도록 depthwise convolutional 네트워크로 구성하였다. Feature 등 완전연결 레이어의 정보가 랜덤포레스트(regression) 모델로 입력되도록 구현되었다.

### (3) 예측성 검증 및 평가

#### (가) 평방근제곱근오차 및 편차공간상관계수

훈련 기법에 따른 선행예측시간(lead time range) 전체 예측 기후 변수의 예측빈도별(frequency), 선행주수별 및 여름철 예측성 평가는 평방근제곱근오차(Root Mean Square Error, 이하 RMSE)와 편차공간상관계수(anomaly spatial Pattern Correlation Coefficient, 이하 PCC)를 계산하여 사용된 딥러닝 훈련 기법 및 훈련 기법간에 대한 예측 정확성 및 변동성과 상관성 등을 평가하였다(Eq. 5 & 6).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{1,i} - x_{2,i})^2}{n}} \quad [5]$$

여기서  $x_{1,i}$ 는  $i$ 일짜의 학습되지 않은 개별 기후모델 혹은 MME 기반 S2S 예측기후 자료이며,  $x_{2,i}$ 는  $i$ 일짜에서 딥러닝에 의해 학습된 개별 기후모델 혹은 MME 기반 S2S의 예측기후 자료이다.

$$PCC_i = \frac{\sum_{j=1}^N (x_{i,j} - \bar{x}_i) \times (y_j - \bar{y})}{\sqrt{\sum_{j=1}^N (x_{i,j} - \bar{x}_i)^2} \times \sqrt{\sum_{j=1}^N (y_j - \bar{y})^2}} \quad [6]$$

여기서  $x_i$ 는 학습 전 혹은 학습 후 각각의 개별 기후모델 및 MME 기반의 S2S 예측기후 자료의 예측 날짜  $i$ 일짜의 값이고,  $\bar{x}_i$ 는 훈련 기간 동안 개별 기후모델 혹은 MME 기반 계절내 예측기후 자료의 평균, 즉 모델의 기후 평균값을 나타낸다.  $y_i$ 는  $i$ 일짜에서의 관측 자료,  $\bar{y}_i$ 는 훈련 기간 동안 관측 자료의 평균, 즉 기후값을 나타낸다.  $j$ 는 선행예측시간(lead time range),  $N$ 은 예측 개수를 각각 나타낸다.

## (나) 개선 비율

개별 기후모델의 학습 모델 적용 전의 S2S 예측의 RMSE 및 PCC가 학습 모델 적용 후에 의해 얼마만큼 개선 및 변화되었는지 비교하기 위해 개선 비율(importance ratio)을 계산하였다 (Lang et al., 2020). Y축은 Eq. [7]으로부터 계산된 개별 기후모델의 RMSE이름 놓고, X축에는 Eq. [8]로부터 계산된 PCC를 분포시킨다. Eq. [7]에서 0의 의미는 기준 모델이 완벽하다고 가정할 때 완벽 모델의 RMSE는 0이다. Eq. [8]에서 1은 역시 기준 모델이 완벽하다고 가정하면 완벽 모델의 편차 공간상관계수는 1이다라는 가정으로 개선 비율이 평가된다.

$$RMSE_{improv.} = \frac{MSE_{DL\ or\ ML} - MSE_{s2s}}{0 - MSE_{s2s}} \quad [7]$$

여기서  $RMSE_{DL\ or\ ML}$ 에서  $DL\ or\ ML$ 의 의미는  $ML$ 은 랜덤포레스트 학습 모델에 따른 RMSE 값이고,  $DL$ 은 하이브리드 학습 모델에 따른 RMSE 값이다.  $RMSE_{s2s}$ 는 학습 모델이 적용되지 않은 개별 기후모델 혹은 MME 기반 S2S의 RMSE 값이다.

$$PCC_{improv.} = \frac{PCC_{DL\ or\ ML} - PCC_{s2s}}{1 - PCC_{s2s}} \quad [8]$$

여기서  $PCC_{DL\ or\ ML}$ 에서도 역시  $DL\ or\ ML$ 의 의미는  $ML$ 은 랜덤포레스트 학습 모델에 따른 PCC 값이고,  $DL$ 은 하이브리드 학습 모델에 따른 PCC 값이다.  $PCC_{s2s}$ 는 학습 모델이 적용되지 않은 다시 말해서 보정되지 않은 S2S의 PCC 값이다.

## (다) 코사인 유사도

본 연구에서는 학습 효과에 대한 척도로 코사인 유사도를 계산하였다(Eq. [9] & Figure 78). 코사인 유사도는 Data Science에서 유사성 검증을 위해 가장 많이 활용되는 척도로 두 벡터간의 코사인 각도를 계산한다(<https://towardsdatascience.com/cosine-similarity-how-does-it-measure-the-similarity-maths-behind-and-usage-in-python-50ad30aad7db>). 코사인 유사도가 높을수록 서로 유사성이 높다는 의미로 간주된다(Han et al., 2012; He et al., 2020). 따라서, 본 연구에서는 학습 기법이 적용되지 않은 S2S와 관측값간의 코사인 유사도와 학습 기법이 적용된 S2S와 관측값간의 코사인 유사도를 비교하여 학습 효과의 변화를 분석하였다.

$$\cos(\theta) = \frac{\langle \hat{y}, y^* \rangle}{\|\hat{y}\|_2 \|y^*\|_2} \quad [9]$$

Eq. [9]에서,  $\hat{y}$ 은 예측값이고,  $y^*$ 에는 참값 혹은 label이 입력되는데, 여기서는  $y^*$ 이

observation, 즉 Era-Interim 값이 된다.

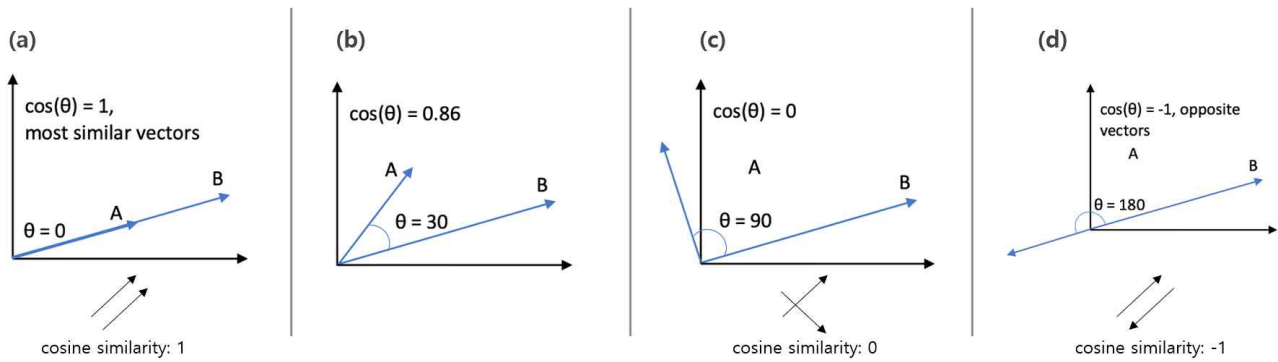


Figure 55. Cosine similarity measures the similarity between two vectors by calculating the cosine of the angle between the two vectors. Cosine similarity is 1 at theta = 0 and -1 at theta = 180, that means for two overlapping vectors cosine will be the highest (a) and lowest for two exactly opposite vectors (d).

## 라. 동아시아 계절내/MME 예측에 대한 하이브리드 알고리즘 민감도 분석

### (1) 랜덤포레스트 알고리즘 최적화

랜덤포레스트 알고리즘을 개별 기후모델 및 다중모델앙상블(multi-model ensemble, 이하 MME) 기반 계절내(Sub-seasonal to seasonal, 이하 S2S) 예측 자료에 적용하여 랜덤포레스트 알고리즘을 최적화하기 위해서는 먼저 초매개모수(이하 hyperparameter)의 최적화가 필요하다.

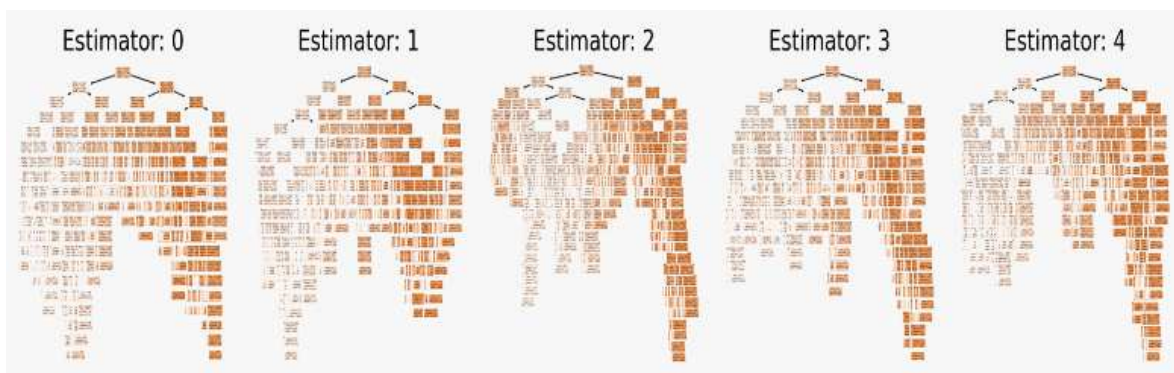


Figure 56. The decision tree structure by the n\_estimator(i.e. 5) of the random forest developed in this study.

Figure 56은 hyperparameter  $n\_estimator$ 별 tree 구조를 테스트한 것이다. 또한 Figure 57는 GridSearch 함수를 통해 본 연구에서 ECMWF의 S2S 예측 자료에 대한 랜덤포레스트 학습 모델의 최적화(optimization)된 hyperparameters 결과이다. 오렌지선은 훈련 자료에서, 파란선은 검증 자료에서 각각 hyperparameter에 의해 도출될 수 있는 정밀도와 재현율에 대한 상대적 기여도를 나타내며 이 지수의 범위는 0에서 1이다.  $n\_estimator$ 와  $max\_feature$ 는 각각 10과 3에서 score가 증가한 후 estimator 및  $max\_feature$ 가 증가해도 변화가 나타나지 않았다.  $max\_depth$ 의 경우, 6 이후부터 오렌지선과 파란선이 벌어져서 과적합(이하 overfitting)이 될 수 있음을 시사했다.  $min\_sample\_leaf$ 의 경우, 6 이후부터 오렌지색 선과 파란색 선이 좁아진 후 더 이상의 변화가 나타나지 않았다. 이와 같이 hyperparameter의 score 값이 0과 1의 범위에서 적히도록 출력되면서 최상의 조합을 도출하였다.

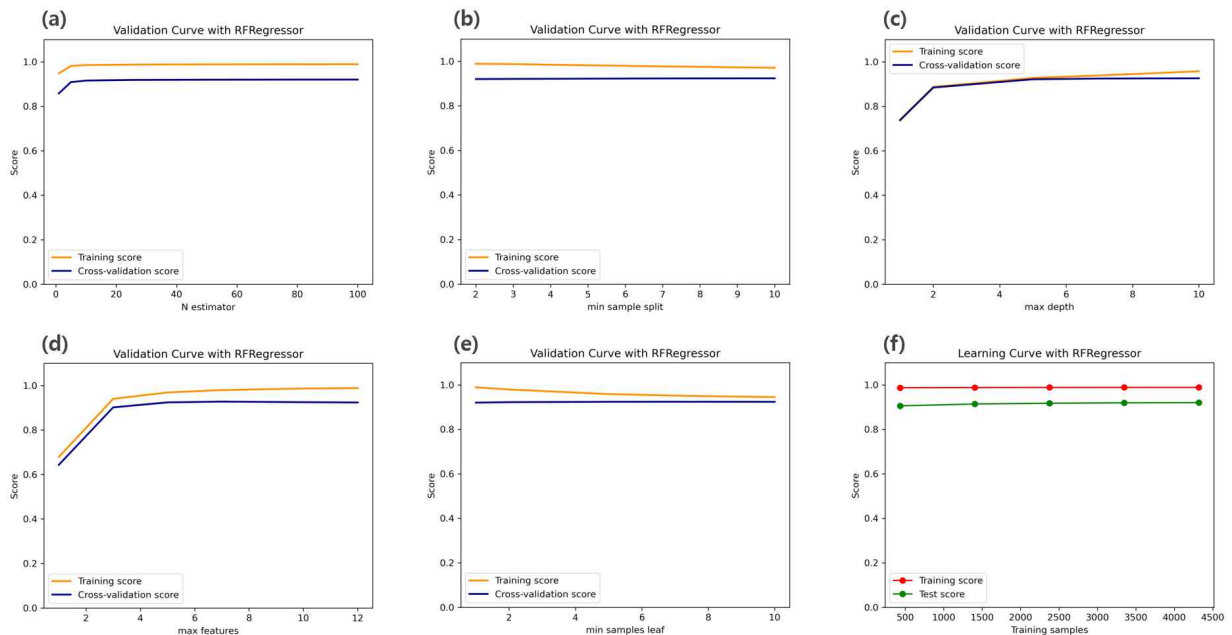


Figure 57. Optimization of random forest hyperparameters: The result of optimizing the hyperparameters(a:  $n\_estimator$ , b:  $min\_sample\_split$ , c:  $max\_depth$ , d:  $max\_feature$ , e:  $min\_sample\_leaf$ , f: learning curve) of the random forest algorithm developed for the predicted subseasonal daily maximum air-temperature of ECMWF in this study.

또한, 입력 변수(반응 변수)와 예측 변수(설명 변수)간의 랜덤포레스트 알고리즘에 대한 예측 성능 설명을 확보하기 위해 변수 중요도(Feature Importance Score)를 계산하였다. 변수 중요도의 0은 전혀 사용되지 않았음을, 1은 완벽하게 예측에 사용되었음을 나타낸다. 그러나 변수 중요도가 낮다고 이 특성이 유용하지 않다는 뜻은 아니다. Figure 58의 a에서 f는 랜덤포레스트 알고리즘을 5개 개별 기후모델 및 MME 기반의 S2S 일 최고기온에 적용하여 변수 중요도를 계산한 결과이다. 0과 1사이에서 절대적인 값의 차이는 있지만 NCEP과 UKMO를 제외하고,

대기 이동에 따른 바람(V850 및 U850) 인자가 목표 변수인 일 최고기온 예측에 영향을 미치는 것으로 나타났다. 그러나, UKMO와 NCEP은 바람 인자보다는 복사(OLR) 및 고도(Z200 등) 인자들이 일 최고기온 예측에 더 영향을 미치는 것으로 나타났으며(Figure 58의 c와 e), MME S2S의 경우는 더 복잡하게 나타났다(Figure 58의 f). 또한, 5개 개별 기후모델의 S2S에 대한 랜덤 포레스트 모델의 변수 중요도에서는 예측 변수(설명 변수)에 대해 영향을 미치는 입력 변수(반응 변수)에 대한 설명이 뚜렷하게 나타난 반면, MME S2S는 우세 인자가 뚜렷하게 나타나지는 않았다.

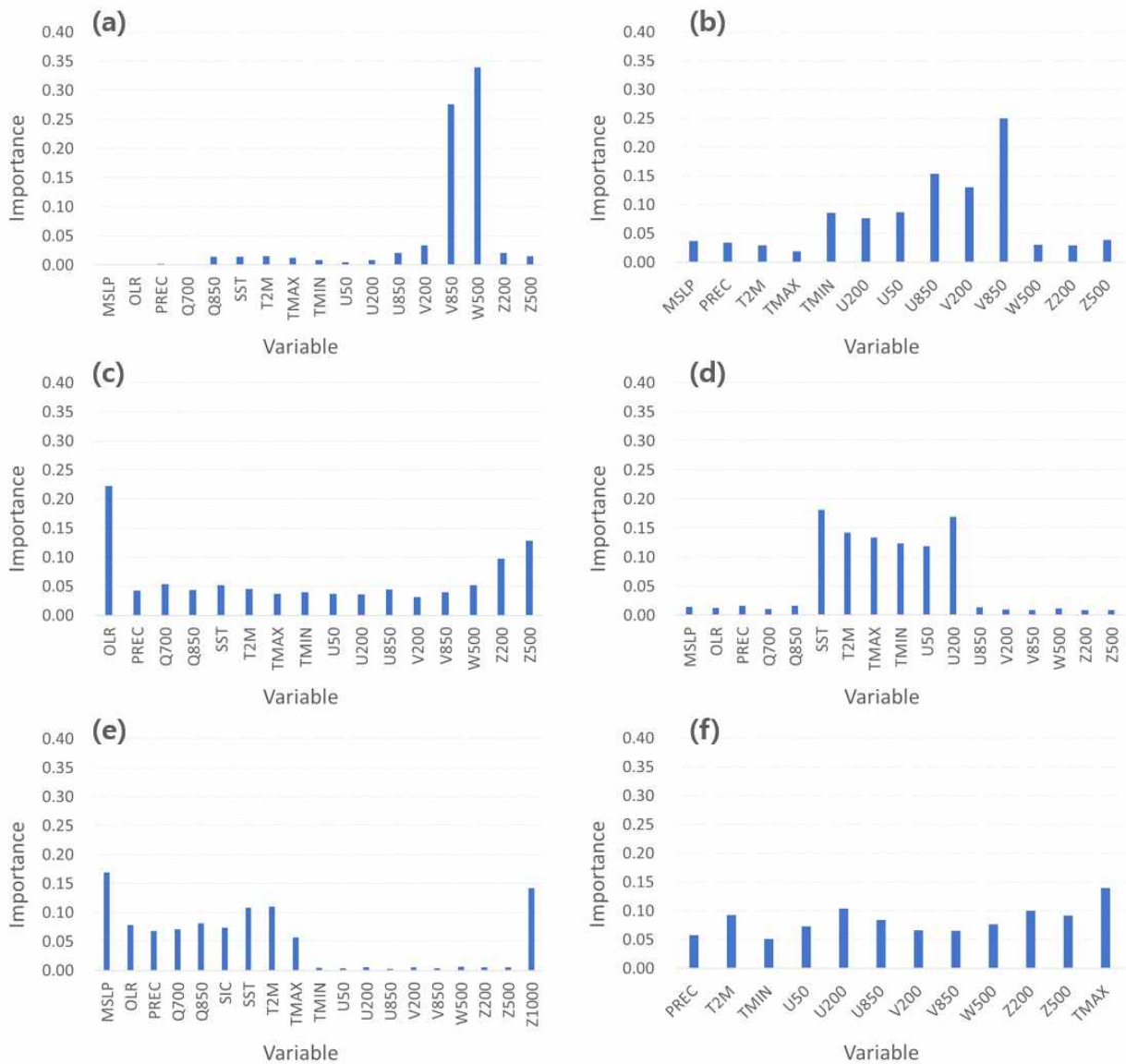


Figure 58. Feature Importance of random forest: Feature importance results of random forest calculated by applying the random forest algorithm developed in this study to 5 individual climate models(a: ECMWF, b: GloSea5, c: UKMO, d: CMA, e: NCEP) and MME-based subseasonal daily maximum air-temperature(f).

## (2) 합성곱장단기기의 알고리즘 최적화

합성곱 장단기기억(Convolutional Long-Short Term Memory, 이하 ConvLSTM) 알고리즘을 개별 기후모델 및 MME 기반 S2S 예측 자료에 적용하여 ConvLSTM 알고리즘을 최적화하기 위해서는 신경망 구조, 입력 자료 변형 및 hyperparameter의 결정과 최적화가 필요하다. 따라서, 기존 ConvLSTM의 신경망 모델을 선정하기 위해 민감도 분석을 하였다. 먼저, 기존 ConvLSTM의 신경망 선정하기 위해, base 신경망 이외 다른 레이어 구조를 가진 3개의 신경망을 구성하였다. 첫 번째는 base 신경망에서 1개 합성곱 레이어를 추가하여 연장된 신경망으로 구성하였다. 두 번째는 첫 번째 구조에서 batch normalization 레이어를 혼합한 구조이다. 세 번째는 두 번째에 1개 합성곱 레이어를 추가하고 dropout이 혼합된 신경망으로 구성하였다. Base 신경망을 포함하는 4개의 신경망에 대해 Randomsearch의 내장 함수(tuner.get\_best\_models)를 통해 기존 신경망을 선정하기 위한 민감도 분석 결과, base 신경망에 1개 합성곱 레이어가 추가된 구조가 기존 ConvLSTM의 신경망으로 선정되었으며 이것을 하이브리드 모델의 버전 1(v1)로 명명하였다.

### 마. 하이브리드 알고리즘 내 동아시아 계절내/MME 예측입력자료의 민감도 분석

먼저 Eq. [8]로부터 개별 기후모델 및 MME 기반 S2S의 편차공간상관계수(anomaly spatial pattern correlation coefficient, 이하 PCC)를 계산하여 S2S 예측 자료가 잘 구축되었는지 확인하였다. Week1에서 ECMWF를 제외한 나머지 개별 기후모델의 PCC보다 높게 나타났다(Figure 59). Week2 이후부터는 MME의 PCC가 나머지 개별 기후모델의 PCC와 비슷하였다.

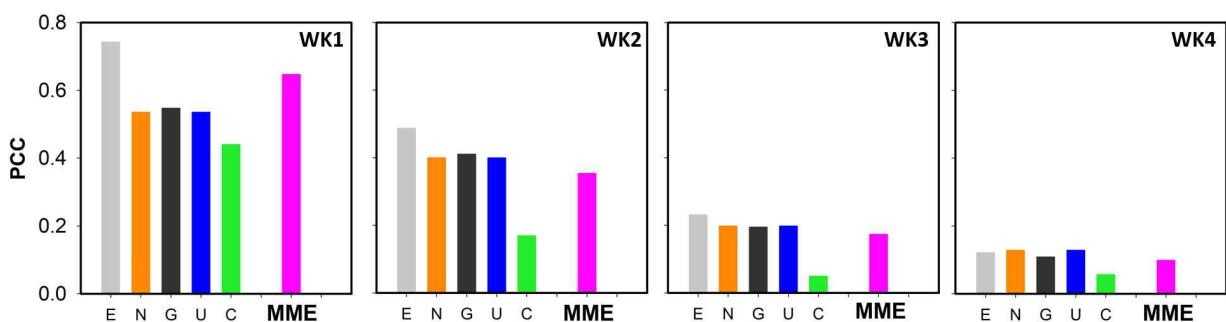


Figure 59. Comparison of spatial anomaly pattern correlation coefficients of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME before applying training models.

## (1) 입력자료 민감도 분석

앞서 언급한 바와 같이, 신경망 모델의 입력 자료 변형(transformation)에 따라서도 신경망 모델의 예측 성능이 영향을 받기 때문에 입력 자료의 변형에 대한 민감도 분석이 필요하다. 따라서 하이브리드 모델(v1)에 결합된 기준 ConvLSTM의 신경망 모델의 개별 기후모델 및 MME S2S 예측 입력 자료 변형 방법에 따른 민감도 분석을 하였다. 기계학습과 심층학습에서 입력 자료의 변형을 위해 많이 사용되는 대표 변환기로는 MinMax Scaler, Robust Scaler, Standard Scaler, MaxAbs Scaler 등이 있다(Table 11).

Table 11. Four scalers information for data transformation in deep learning.

Scaler	Description
MinMax	최대 및 최소값이 각각 1과 0이 되도록 스케일링
Robust	중앙값(median)과 IQR(interquartile range)을 사용. 아웃라이어의 영향을 최소화
Standard	기본 스케일. 평균 및 표준편차 사용
MaxAbs	최대절대값과 0이 각각 1과 0이 되도록 스케일링

Figure 60는 4개의 변환기에 의해 변형된 Era-Interim의 일 최고기온 자료와 MME S2S 예측 일 최고기온 자료의 히스토그램 분포와 2차원 분포이다. a와 e는 MinMax Scaler, b와 f는 Robust Scaler, c와 g는 Standard Scaler, d와 h는 MaxAbs Scaler에 의한 결과이다. 각각의 변환기에 따른 기준 ConvLSTM의 신경망 모델의 정확도 분석 결과, MaxAbs Scaler에 의한 정확도가 가장 높은 것으로 나타났다. Figure 61은 ECMWF S2S 예측 일 최고기온 자료에 대한 4개 변환 결과이다. 기준 하이브리드 모델에서 4개 변환기의 훈련자료 변환에 따른 민감도 분석 결과, ECMWF의 경우 기준 하이브리드 모델의 네트워크에서 Robust Scaler로 선정하였고, MME의 경우 MaxAbs Scaler가 선정되었다.

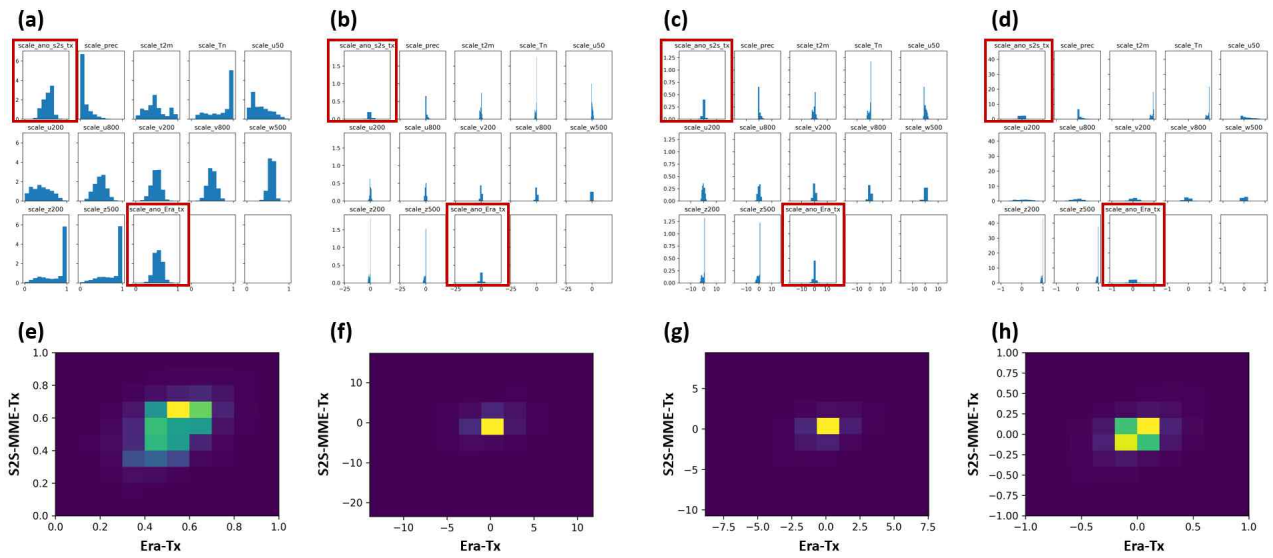


Figure 60. Histograms and two-dimensional distributions of input data transformed by four transformers for sub-seasonal MME predicted daily maximum air-temperature: The four transformers are (a) minMax scaler, (b) Robust scaler, (c) Standard scaler, (d) maxAbs scaler, respectively. Two-dimensional distribution between the Era-Interim daily maximum air-temperature matched with the input data transformed by the four transformers.

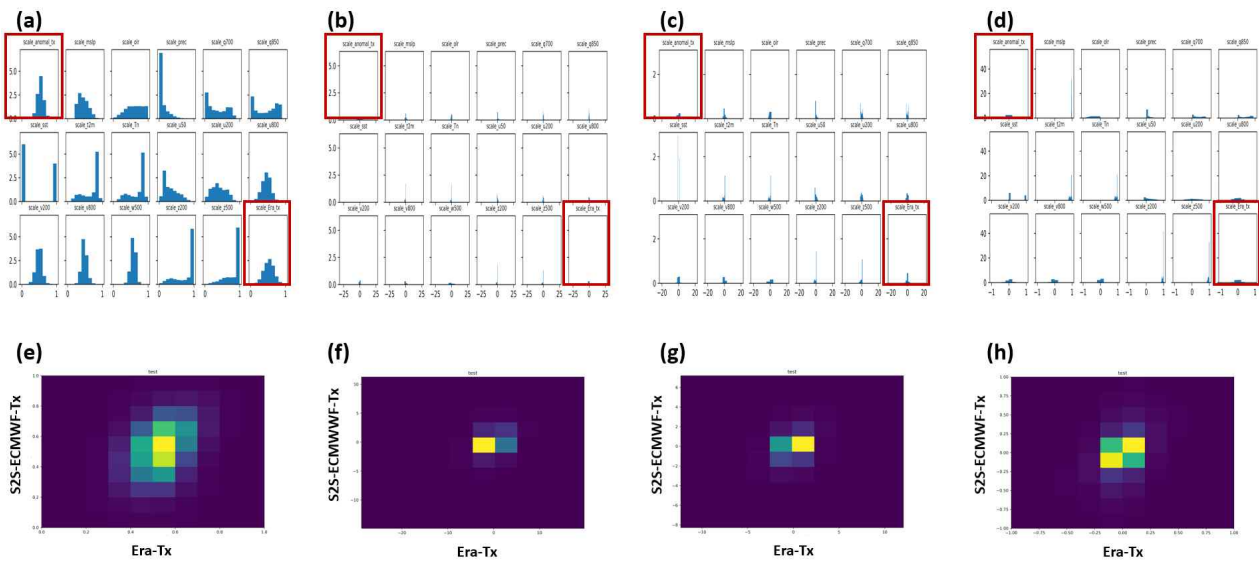


Figure 61. Histograms and two-dimensional distributions of input data transformed by four transformers for ECMWF sub-seasonal predicted daily maximum air-temperature: The four transformers are (a) minMax scaler, (b) Robust scaler, (c) Standard scaler, (d) maxAbs scaler, respectively. Two-dimensional distribution between the Era-Interim daily maximum air-temperature matched with the input data transformed by the four transformers.

## 바. 하이브리드 모델 최적화

### (1) 개별 기후모델의 LSTM 적용에 따른 예측 성능 평가

랜덤포레스트 알고리즘의 최적화에서 언급한 바와 같이, 합성곱 알고리즘에 관여하는 hyperparameter의 결정 및 최적화 역시 필요하다. 신경망 모델의 hyperparameter에는 batch size, epoch, learning rate, activation function, kernel size, dense units 등이 있다(Table 12). 이러한 hyperparameter에 대한 객관적 선정을 위해 tensorflow의 라이브러리 keras Tuner 등을 도입하여, 본 연구의 기준 ConvLSTM의 신경망 모델의 hyperparameter의 최적화가 필요하다. Hyperparameter를 객관적으로 선정하기 위한 도구를 제공한다(Table 13).

Table 12. Key hyperparameters of convolutional neural network algorithm in the hybrid model of this study.

Hyperparameter	Description
batch size	한 번의 반복(i.e. epoch)에서 사용된 훈련 예제의 수
epoch	전체 학습 데이터 세트의 처리 수
learning rate	특정 뉴런에 대한 최적화에 사용되는 오류 기여도의 비율
activation function	네트워크가 데이터의 복잡한 패턴을 잘 학습할 수 있도록 인공신경망에 추가하는 함수
kernel size	필터 마스크의 너비 x 높이를 의미

Table 13. Information of four representative tuners for hyperparameter optimization of deep learning algorithms.

Tuner	Description
BaysianOptimization	블랙박스 기능을 최적화하는 튜너로 최적화를 위해 Bayes 규칙을 사용
RandomSearch	주어진 범위 내에서 랜덤(임의)으로 값을 검색하여 선택 튜닝 시간을 줄이기 위해 적은 epoch로 많은 모델의 훈련을 시도하고 훈련 결과가 가장 좋은 모델에서 경쟁을 시도하기 때문에 시간이 많이 걸리는 단점
Hyperband	교차 검증된 하이퍼파라미터 검색을 사용

대표적인 Keras 튜너에는 BaysianOptimization Tuner, RandomSearch Tuner, Hyperband Tuner 등이 있는데, Tuner 별로 다소 다른 hyperparameter가 도출되는 문제가 발생된다. 그래서 기준 ConvLSTM의 신경망 모델의 hyperparameter를 어떤 Tuner로 최적화할지에 대한 결정이 필요하였다. 따라서 객관적 튜너 선정을 위해 먼저 자원 및 계산 속도 측면에서 max trials

옵션을 통해 연산 속도 조절이 가능한 RandomSearch Tuner와 BayesianOptimization Tuner를 1차 선정하여 테스트하였다. 그 결과, 최종 BayesianOptimization을 Tuner가 선정되었다. BayesianOptimization Tuner에 의한 ECMWF의 경우, 기존 ConvLSTM의 신경망 모델에서 batch size는 32, epoch는 75, dense unit은 128, activation function는 relu로, learning rate는 0.001로 선정되었다.

하이브리드 모델(v1)에서 선정된 hyperparameter 기반의 기존 ConvLSTM의 신경망 모델로 MME 기반 S2S 예측 일 최고기온을 훈련한 결과가 Figure 62의 a이고, b에서 d는 각각 ECMWF, UKMO, GloSea5 결과이다. 파랑색이 훈련 자료에서 평가된 평균제곱오차(Mean Square Error, MSE)를 나타내고, 오렌지색은 검증 자료를 나타내는데, 검증 자료에서 미세한 차이이지만, 더 균일한 값으로 하이브리드 모델에 연결된 기존 ConvLSTM의 신경망 모델의 퍼포먼스가 이루어진 것을 알 수 있다.

마찬가지로 Figure 63는 하이브리드 모델(v1)에서 랜덤포레스트 모델로 MME 기반 S2S 예측 일 최고기온을 훈련한 결과이다. x축은 training score를 놓고, y축은 test score를 놓은 것이다. 파란색은 단일 랜덤포레스트, 오렌지색은 하이브리드 모델 내의 랜덤포레스트의 학습 결과로, 하이브리드 모델 내의 랜덤포레스트의 학습이 더 잘 이루어졌으며, 이것은 신경망 모델에 의한 feature(형상 혹은 패턴)의 특성 추출의 영향으로 향후, 단일 랜덤포레스트에 특성 추출(예, 주 성분) 부분을 추가했을 경우와 비교 분석을 통해서 설명 가능할 것으로 판단된다.

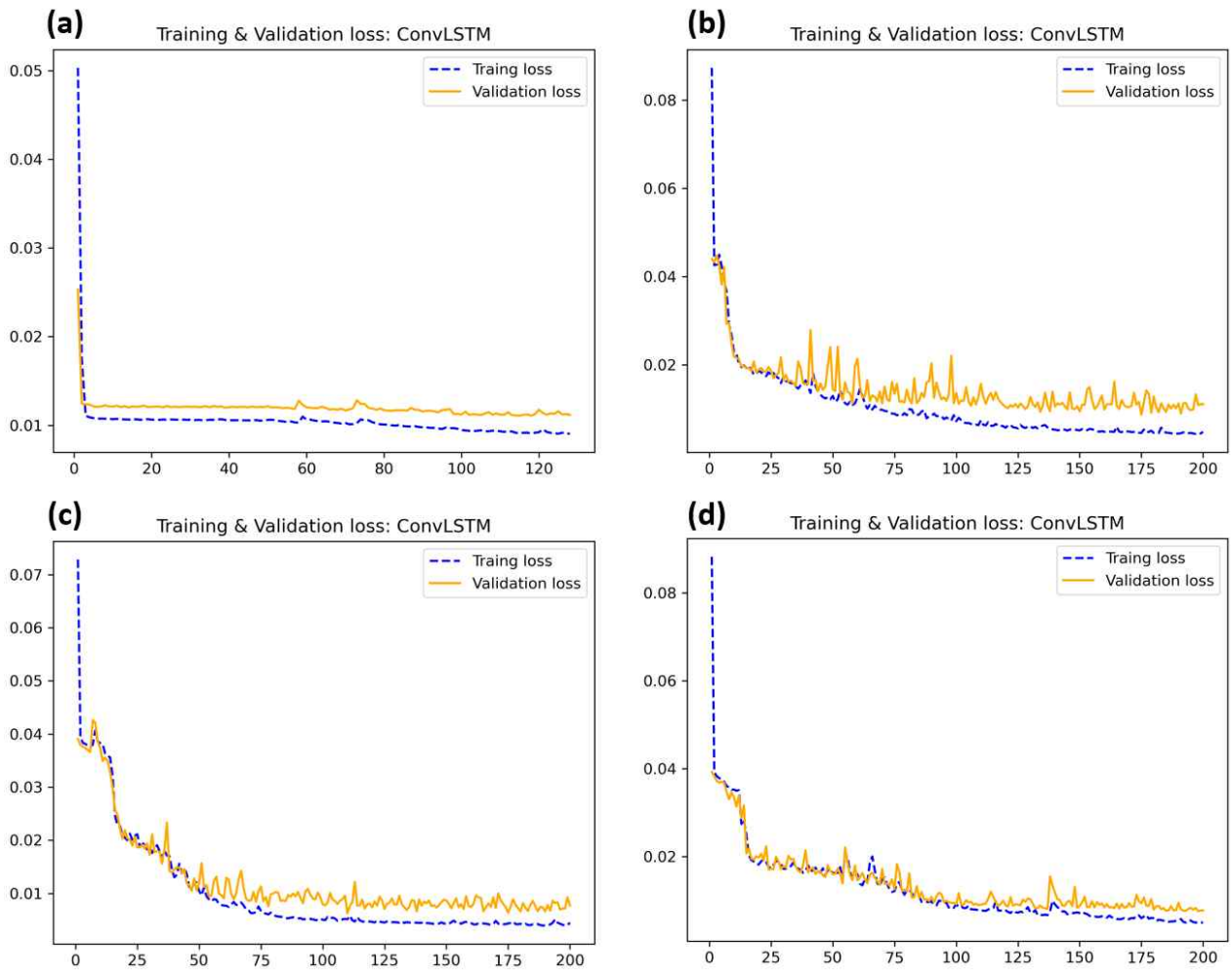


Figure 62. Convolutional short-term memory (ConvLSTM) training results: (a) Loss after training ConvLSTM to sub-seasonal MME predicted daily maximum air-temperature. Losses after learning ConvLSTM of daily maximum air-temperature of ECMWF (b), GloSea5 (c), and UKMO (d), respectively.

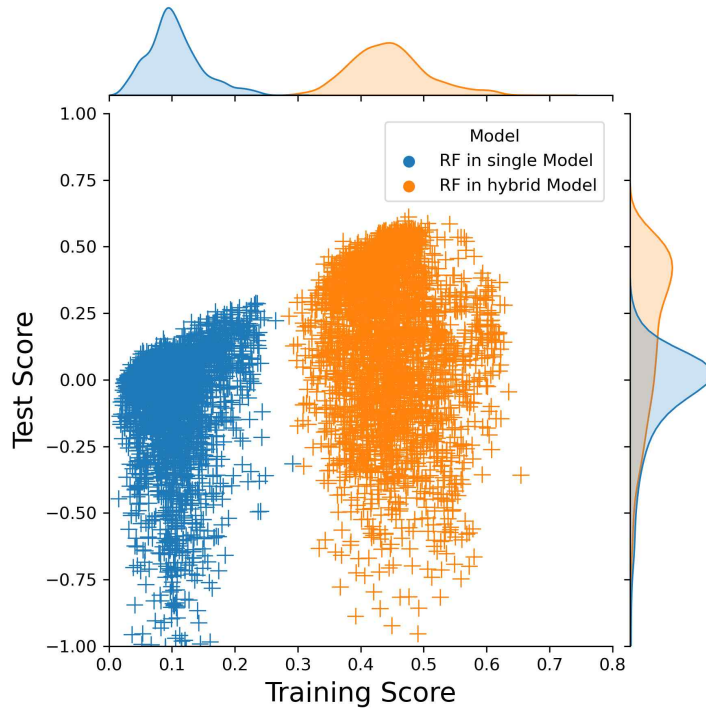


Figure 63. Comparison of training and test scores of a random forest connected to a hybrid model and a single random forest for the sub-seasonal MME predicted daily maximum air-temperature.

## 사. 신경망모델 하이브리드 기반의 동아시아 MME 계절내 예측 기후 성능 평가 기법 구축 및 가시화

### (1) 평방근제곱근오차 및 편차공간상관성에 따른 예측성 평가

Figure 64에서 MME 기반 S2S 예측 일 최고기온에 대해 학습 모델을 적용하지 않은 경우 (좌), 단일 랜덤포레스트 모델을 적용한 경우(가운데) 및 하이브리드 학습 모델을 적용한 경우 (우)의 계산된 공간상관계수(pattern correlation coefficient, 이하 PCC)를 heatmap으로 도출한 것이다. 예측빈도(frequency)는 48회로 단일 랜덤포레스트 학습 모델을 적용한 경우의 PCC가 적용하지 않았을 때보다 낮은 값들로 분포하고 있어 랜덤포레스트 학습 효과가 낮음을 직관적으로 판단할 수 있다.

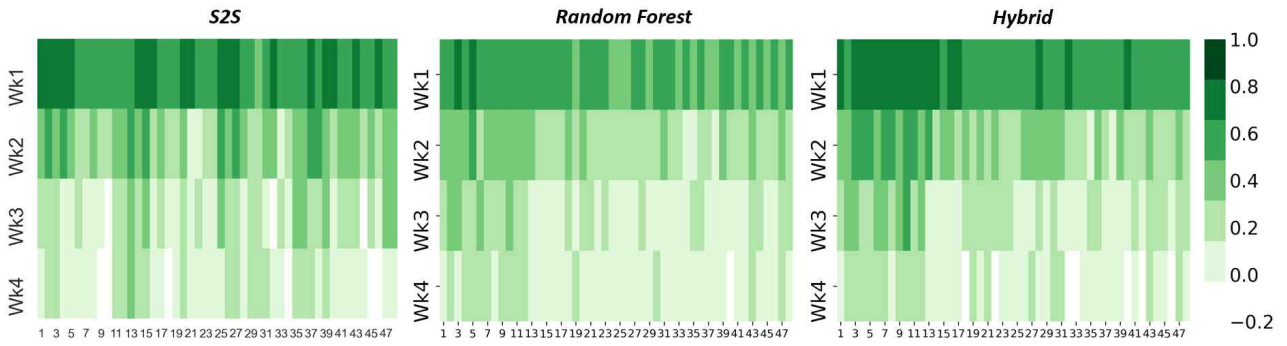


Figure 64. Comparison of spatial anomaly pattern correlation coefficients (after here, PCC) by prediction frequency: PCC of predicted daily maximum air-temperature of sub-seasonal MME before applying training models (left). PCC of predicted daily maximum air-temperature of sub-seasonal MME after applying Random Forest model (middle) and the Hybrid model (right).

Figure 65은 PCC의 편차(difference) heatmap으로 왼쪽은 단일 랜덤포레스트 학습 모델이 적용된 경우의 PCC에서 학습 모델이 적용되지 않은 경우의 PCC를 뺀 편차, 오른쪽은 하이브리드 학습 모델이 적용된 경우의 PCC에서 학습 모델이 적용되지 않은 경우의 PCC를 뺀 편차 heatmap이다. 예측 빈도별 ‘학습 모델에 의한 후보정 효과가 좋아짐’은 양수(빨간색)이어야 하는데, 하이브리드 모델의 보정에서 Week1-2에서도 보정 효과가 반영되었고, Week3-4에서는 보정 효과가 좋은 예측 빈도일이 증가하고 있다.

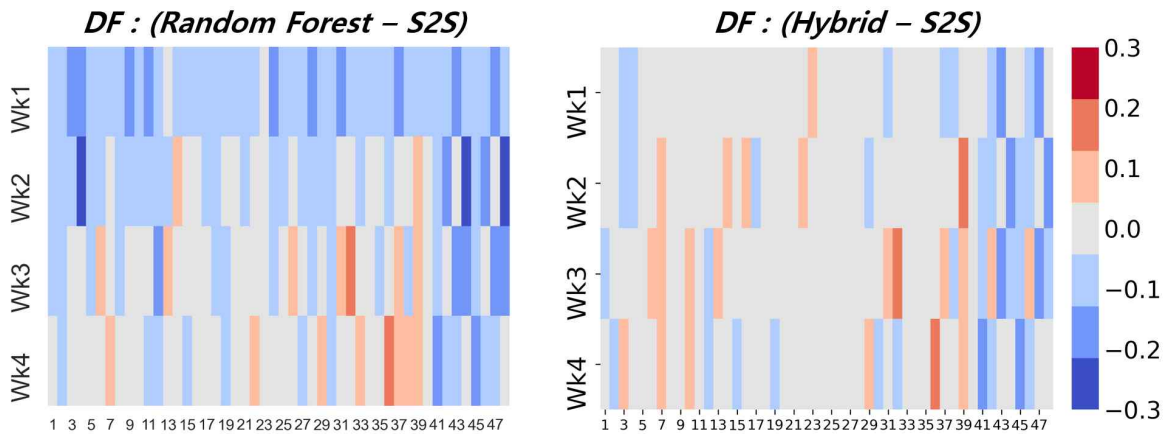


Figure 65. Comparison of the deviation of the spatial anomaly pattern correlation coefficient (after here, PCC) by prediction frequency of predicted daily maximum air-temperature of sub-seasonal MME: The PCC after application by each model (left: Random Forest, right: Hybrid model) is subtracted from the PCC before applying by each model.

Figure 66은 예측 빈도별 PCC 편차의 평균, 최고 및 최저값을 비교한 것이다(Weyn et al.,

2021). 보정 전에 대한 하이브리드 보정의 PCC 편차의 최고, 최저뿐만 아니라 평균이 오른쪽으로, 즉 양의 방향으로 이동되는 것으로 나타났다(오른쪽: 편차가 양수). 하이브리드 학습 모델에 의한 보정이 S2S 예측 개선에 영향을 미친다고 판단된다.

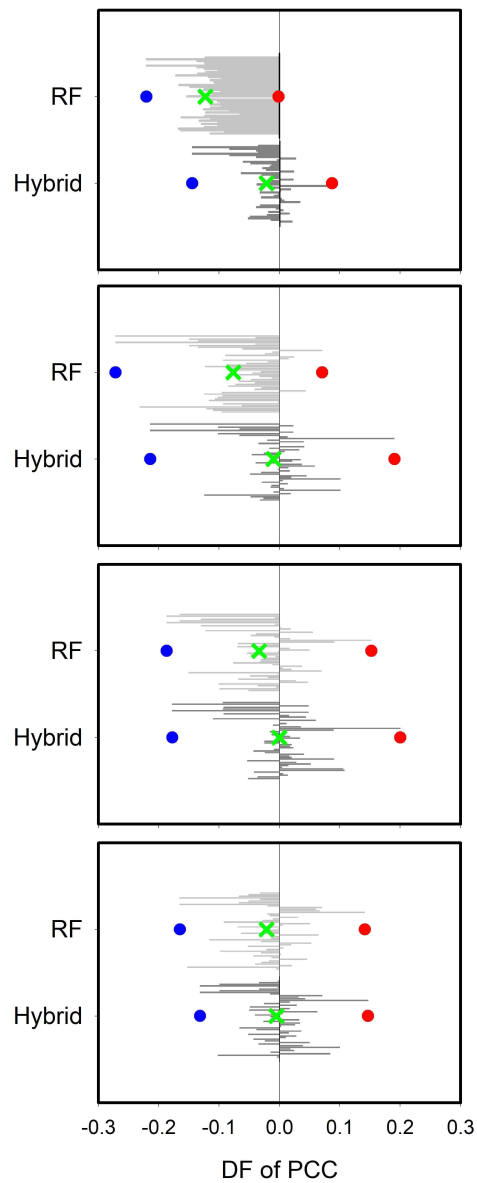


Figure 66. Comparison of mean, highest, and lowest values of deviations of spatial anomaly pattern correlation coefficients by prediction frequency of predicted daily maximum air-temperature of sub-seasonal MME.

Figure 67는 MME에 참여한 개별 기후모델들의 PCC 계산 결과이다. Figure 19에서 학습 모

델이 적용되지 않은 경우(위 패널), 하이브리드 학습 모델이 적용된 경우(가운데 패널), 단일 랜덤포레스트 학습 모델이 적용된 경우(아래 패널)에 대해 계산된 PCC의 heatmap이다. MME 기반 S2S에 대한 PCC heatmap 결과와는 다소 다르게 개별 기후모델 S2S에 대한 PCC heatmap으로는 모델 간의 차이를 직관적으로 보이지 않았다.

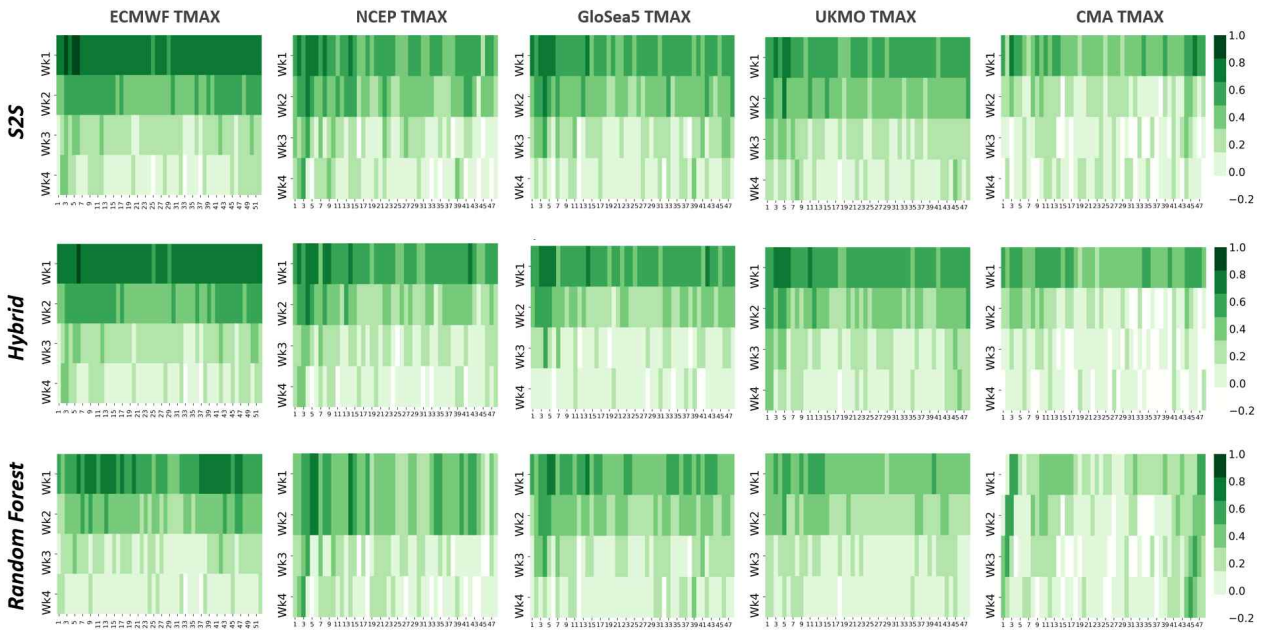


Figure 67. Comparison of spatial anomaly pattern correlation coefficients by prediction frequency of predicted daily maximum air-temperature of five individual climate models (from left, ECMWF, NCEP, GloSea5, UKMO, CMA): Spatial correlation coefficients for each prediction frequency before applying models (upper panel), after applying by the hybrid model (center panel) and random forest (lower panel).

개별 기후모델의 S2S에 대한 PCC 편차를 비교한 것은 Figure 68이다. 특히 사항은 예측성이 낮다고 평가된 기후 모델(예, NCEP이나 CMA)의 경우 Week3-4에서 향상이 나타났다. 또한, 랜덤포레스트 보정에서는 보이지 않았던 Week1에서의 향상이 나타나며, Week3-4에서의 개선 빈도도 증가하는 것으로 나타났다(빨간색 증가).

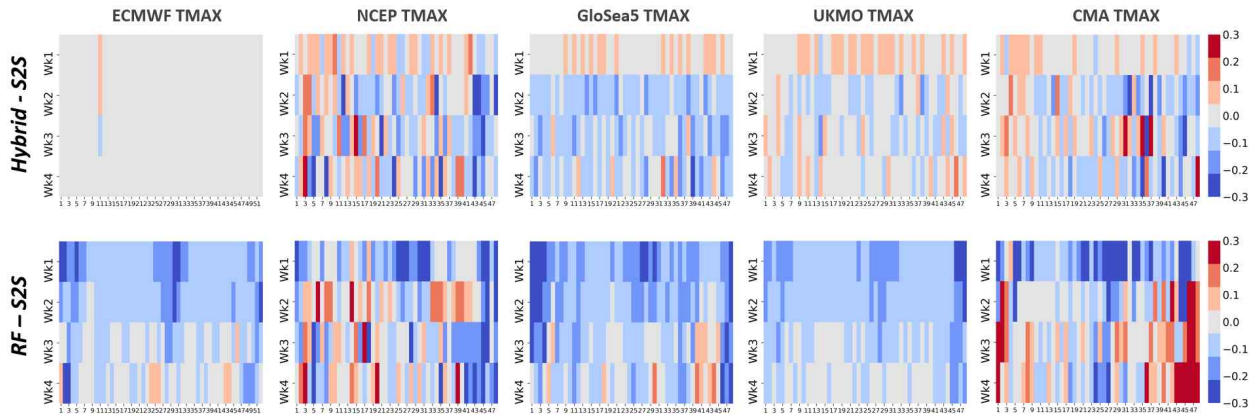


Figure 68. Comparison of the deviation of the spatial anomaly correlation coefficient by prediction frequency of predicted sub-seasonal daily maximum air-temperature of five individual climate models: Subtracted after application by each each training model (top: Hybrid model, bottom: Random Forest) from before applying by the models.

## (2) 개선 비율 평가

Table 14에서 Table 16은 Figure 64와 67에서 5개 개별 기후모델 및 MME 기반 S2S에 대한 PCC와 평방근제곱근오차(Root Mean Square Error, 이하 RMSE)로써 개선 비율 계산 결과의 이해를 돕고자 한 것이다. 학습 모델 보정 기술에 대한 개선 비율을 비교하였다(Table 17 및 Table 18과 Figure 69). 이것은 백분율보다는 비율 개념(=in proportion rather than percentage terms)으로, 어떤 기후모델의 S2S 예측 자료에 대해 학습 모델에 의한 보정 전 RMSE 및 PCC가 학습 모델 보정 기법에 의해 얼마만큼 개선 혹은 변화되었는가를 비교하기 위해 계산한 것이다. 계산한 결과, Week1에서 후보정 기법에 따른 보정 효과는 미흡하였으며 오히려 퇴보되는 문제가 발생하였다. 또한 Week2-4에서 예측성이 낮은 기후모델(CMA, GloSea5, NCEP 등) 및 MME에서 후보정에 의한 RMSE 개선은 보이지만 PCC의 개선은 미흡함이 나타났다. 향후 Week1-2에서 개선율이 네거티브(퇴보)인 이유 분석 필요하며, 학습 모델에 의한 후보정 기술이 RMSE 개선에서는 비교적 일관성이 보이지만, Correlation 개선이 미흡한 문제, 즉 학습 알고리즘의 안정성 및 일관성에 대한 한계에 대해 향후 원인 분석 및 극복이 필요할 것으로 판단된다.

Table 14. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME before applying training models.

S2S	Pattern Correlation Coefficient						Root Mean Square Error					
	MME	ECMWF	NCEP	GloSea5	UKMO	CMA	MME	ECMWF	NCEP	GloSea5	UKMO	CMA
W1	0.65	0.74	0.54	0.55	0.54	0.44	1.53	0.86	1.88	1.16	1.39	2.15
W2	0.35	0.49	0.4	0.41	0.4	0.17	1.88	1.1	2.1	1.37	1.32	2.89
W3	0.17	0.23	0.2	0.19	0.2	0.05	2.0	1.23	2.31	1.47	1.42	2.98
W4	0.1	0.12	0.13	0.11	0.13	0.06	2.04	1.26	2.36	1.5	1.44	3.02

Table 15. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME after applying hybrid training model.

Hybrid	Pattern Correlation Coefficient						Root Mean Square Error					
	MME	ECMWF	NCEP	GloSea5	UKMO	CMA	MME	ECMWF	NCEP	GloSea5	UKMO	CMA
W1	0.63	0.74	<b>0.56</b>	<b>0.58</b>	<b>0.57</b>	<b>0.45</b>	1.8	1.18	2.0	1.29	<b>1.28</b>	<b>1.96</b>
W2	0.34	0.49	0.36	0.33	0.37	0.12	<b>1.84</b>	1.2	<b>2.04</b>	<b>1.31</b>	<b>1.27</b>	<b>2</b>
W3	0.17	0.23	0.16	0.12	0.17	<b>0.06</b>	<b>1.87</b>	<b>1.22</b>	<b>2.06</b>	<b>1.32</b>	<b>1.29</b>	<b>2.01</b>
W4	0.09	0.12	0.09	0.05	<b>0.13</b>	0.02	<b>1.87</b>	<b>1.22</b>	<b>2.07</b>	<b>1.32</b>	<b>1.29</b>	<b>2.02</b>

Table 16. Spatial anomaly pattern correlation coefficients (right) and root mean square error (left) of predicted daily maximum air-temperature within 5 individual climate models for sub-seasonal and MME after applying random forest (after here, RF in this table) model.

RF	Pattern Correlation Coefficient						Root Mean Square Error					
	MME	ECMWF	NCEP	GloSea5	UKMO	CMA	MME	ECMWF	NCEP	GloSea5	UKMO	CMA
W1	0.52	0.62	0.41	0.48	0.42	0.26	1.85	1.11	<b>1.86</b>	1.27	1.49	2.2
W2	0.28	0.39	<b>0.41</b>	0.36	0.31	0.15	1.98	1.22	<b>2.01</b>	1.4	1.42	<b>2.43</b>
W3	0.14	0.18	0.14	0.17	0.14	<b>0.1</b>	2.06	1.3	<b>2.18</b>	1.48	1.51	<b>2.43</b>
W4	0.08	0.09	0.09	0.1	0.08	<b>0.1</b>	2.08	1.33	<b>2.22</b>	1.5	1.53	<b>2.47</b>

Table 17. The values of the X-axis and Y-axis of Figure 20, which plot the improvement rates by the hybrid model (i.e. one of post-calibration methods of this study) of the individual climate models and the MME-based subseasonal predicted daily maximum air-temperature: The right of table is the improvement rates of the spatial anomaly pattern correlation coefficient (PCC) calculated by Eq. [1]. The left of table is the improvement rates of the root mean square error (RMSE) calculated by Eq. [2]. A rate with a positive value means ‘improvement by the post-calibration technique’ and is marked in bold.

Hybrid	Pattern Correlation Coefficient Improvement Ratio						Root Mean Square Error Improvement Ratio					
	MME	ECMWF	NCEP	GloSea5	UKMO	CMA	MME	ECMWF	NCEP	GloSea5	UKMO	CMA
W1	-0.06	-0.02	<b>0.05</b>	<b>0.06</b>	<b>0.07</b>	<b>0.02</b>	-0.18	-0.37	-0.07	-0.11	<b>0.08</b>	<b>0.09</b>
W2	-0.02	<b>0</b>	-0.07	-0.14	-0.06	-0.06	<b>0.02</b>	-0.09	<b>0.03</b>	<b>0.04</b>	<b>0.04</b>	<b>0.31</b>
W3	<b>0</b>	-0.01	-0.05	-0.09	-0.03	<b>0.01</b>	<b>0.07</b>	<b>0.01</b>	<b>0.11</b>	<b>0.1</b>	<b>0.09</b>	<b>0.33</b>
W4	-0.01	<b>0</b>	-0.05	-0.06	<b>0</b>	-0.04	<b>0.08</b>	<b>0.03</b>	<b>0.12</b>	<b>0.12</b>	<b>0.1</b>	<b>0.33</b>

Table 18. The values of the X-axis and Y-axis of Figure 20, which plot the improvement rates by the random forest model (i.e. one of post-correction method of this study) of the individual climate models and the MME-based subseasonal predicted daily maximum air-temperature: The right of table is the improvement ratio of the spatial anomaly pattern correlation coefficient (PCC) calculated by Eq. [1]. The left of table is the improvement ratio of the root mean square error (RMSE) calculated by Eq. [2]. A rate with a positive value means ‘improvement by the post-calibration technique’ and is marked in bold.

RF	Pattern Correlation Coefficient Improvement Ratio						Root Mean Square Error Improvement Ratio					
	MME	ECMWF	NCEP	GloSea5	UKMO	CMA	MME	ECMWF	NCEP	GloSea5	UKMO	CMA
W1	-0.35	-0.47	-0.27	-0.14	-0.26	-0.32	-0.21	-0.29	<b>0.01</b>	-0.09	-0.07	-0.02
W2	-0.12	-0.19	<b>0.01</b>	-0.09	-0.16	-0.02	-0.05	-0.11	<b>0.04</b>	-0.03	-0.08	<b>0.16</b>
W3	-0.04	-0.07	-0.07	-0.03	-0.08	<b>0.05</b>	-0.03	-0.06	<b>0.06</b>	-0.01	-0.06	<b>0.18</b>
W4	-0.02	-0.04	-0.04	-0.01	-0.05	<b>0.05</b>	-0.02	-0.05	<b>0.06</b>	0	-0.06	<b>0.18</b>

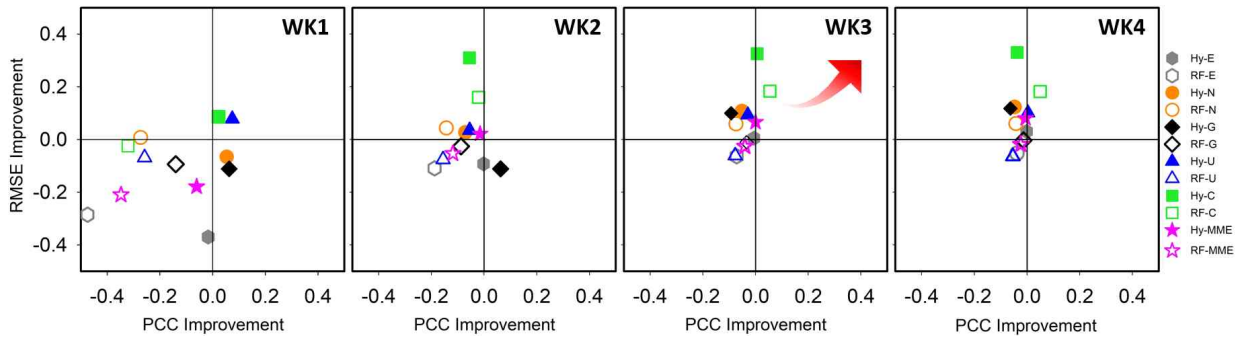


Figure 69. Comparison of improvement rates of post-calibration techniques for the sub-seasonal daily maximum air-temperature of individual climate models and MME, respectively: The x-axis is the ratio between the spatial anomaly pattern correlation coefficient (after here, PCC) after applying by the training models, for example Random Forest and the Hybrid model and the PCC before applying the models. The y-axis is the ratio between the root mean square error (after here, RMSE) after applying by the training models and the RMSE before applying the models.

### (3) 코사인 유사도에 따른 예측성 평가

앞서 언급한 바와 같이, S2S 예측 기후 자료에 기계학습 및 심층학습 모델 후보정 기술 적용을 통해 관측값, 즉 Era-Interim의 기후를 얼마나 잘 재현 혹은 유사 정도(예, 얼마나 잘 모사할 수 있는가)를 비교 분석하기 위해 본 연구에서 코사인 유사도(cosine similarity) 분석을 도입하였다. Figure 70는 MME S2S 예측 기후 자료에 학습 모델을 적용하지 않은 경우와 적용한 경우에 대한 코사인 유사도 분석 결과이다. ‘코사인 유사도가 높을수록 서로 유사성이 높다’는 의미로 초록색이 많을수록 유사성이 높은 것이다. Week3에서 높은 코사인 유사도 값, 즉 초록색이 증가가 나타남으로써 하이브리드 보정이 보정 전의 예측 일 최고기온을 잘 모의하고 있음을 알 수 있다. Week4에서는 하이브리드 보정에 대한 유사성이 다소 낮아지지만, density plot을 보면 높은 코사인 유사도가 하이브리드 보정에서 유지되고 있는 것을 알 수 있다(Figure 71 from (a) to (d)).

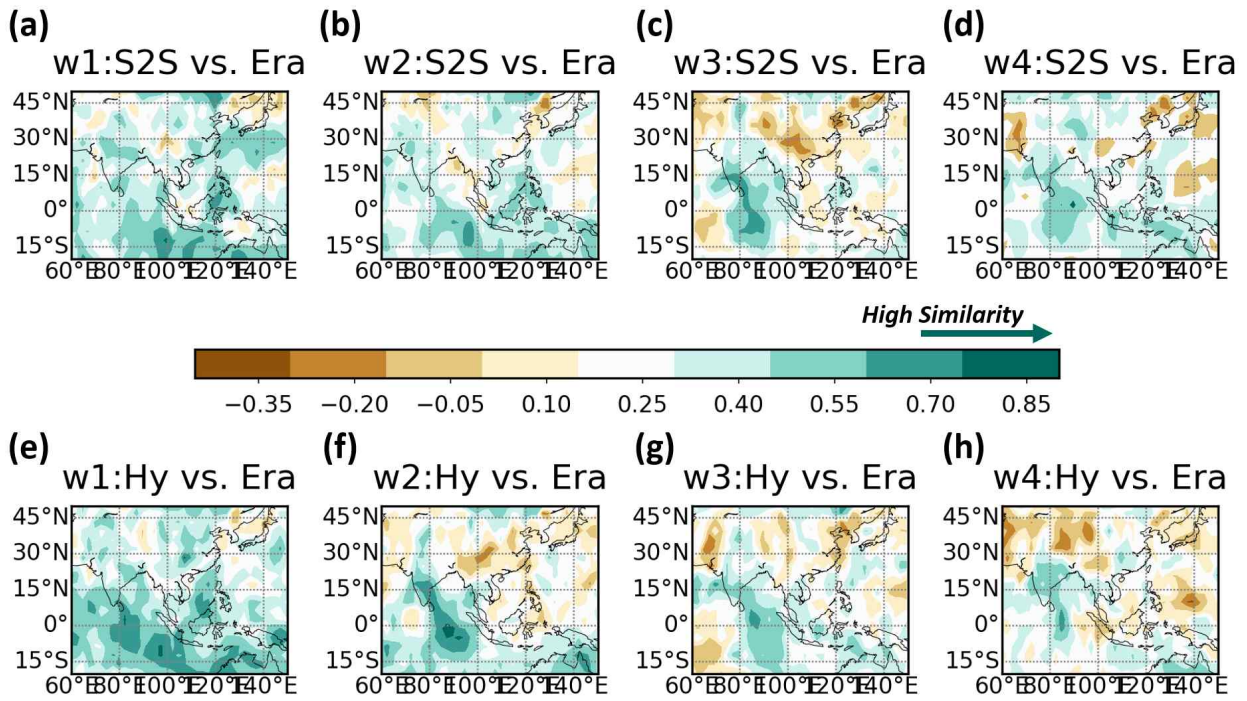


Figure 70. Comparison of cosine similarity between spatial anomaly pattern correlation coefficients (after here, PCC): Era-Interim's predicted daily maximum air-temperature is defined as observation, that is, true value. Calculation of cosine similarity between the PCC of the Era-Interim's daily maximum air-temperature and the PCC before applying the training models to the daily maximum air-temperature on the sub-seasonal MME (from (a) to (d)).

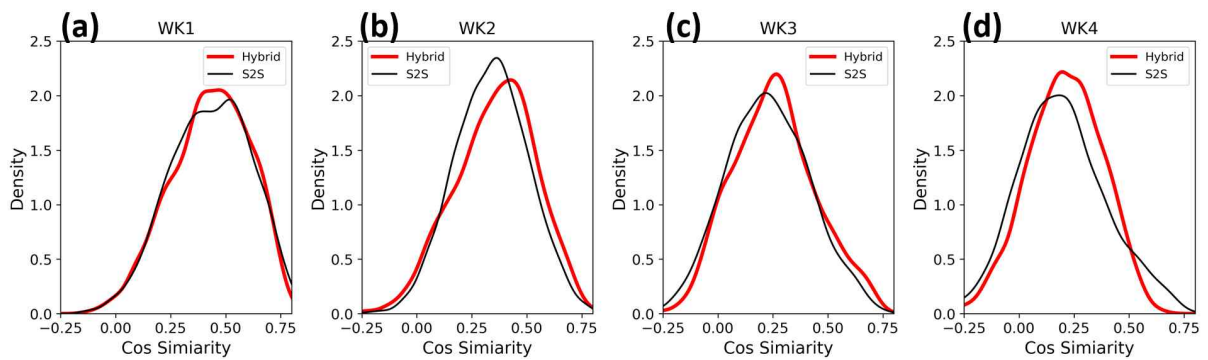


Figure 71. The density of cosine similarity by week in Figure 21 is from a to d.

MME에 참여한 5개 개별 기후모델의 코사인 유사도 분석 결과는 Figure 72이다. 개별 기후모델의 코사인 유사도는 week3-4만 plot을 했는데, 앞서 언급바와 같이 낮은 예측성의 개별 기후모델(예, NCEP이나 CMA)의 하이브리드 후보정에서 관측값과의 높은 코사인 유사성을 나타내고, 낮은 유사성은 감소하는 것을 알 수 있다. 코사인 유사도 역시 양의 값(초록색)이 많아지는 것이 목표임을 감안하면, week3과 4에서 RF 보정에 따른 코사인 유사도가 점차 양의 값으

로 많아지는 것을 볼 수 있음. 따라서 후보정 처리가 week3-4의 예측 개선에 기여한다고 판단 된다.

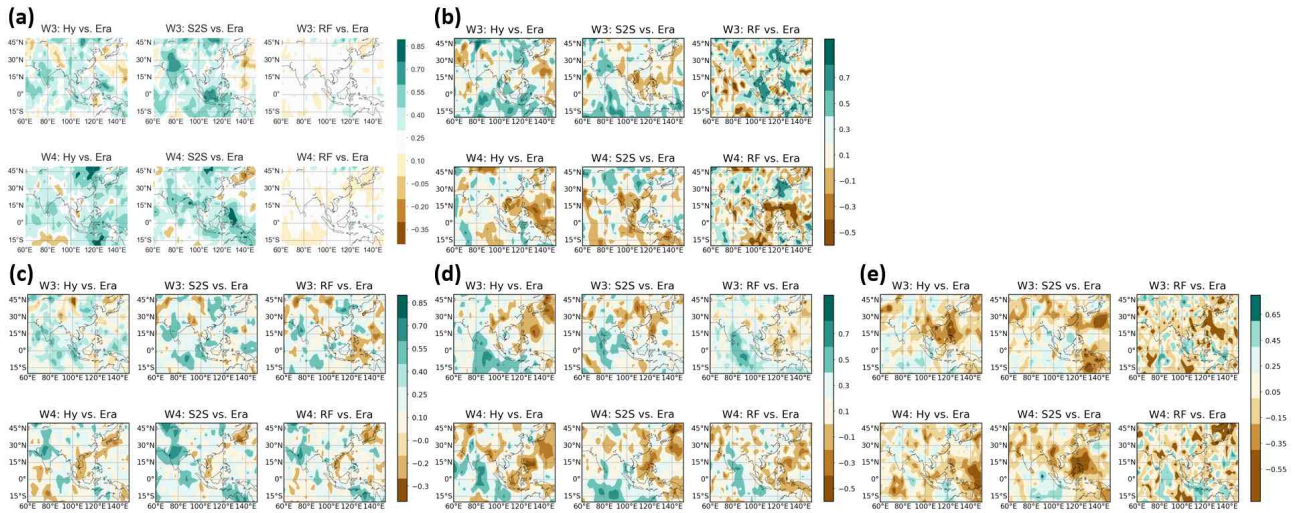


Figure 72. Comparison of cosine similarity between spatial anomaly pattern correlation coefficients (after here, PCC): Calculation of cosine similarity between the PCC of the Era-Interim's daily maximum air-temperature and the PCC before applying the training models to the sub-seasonal daily maximum air-temperature of five individual climate models. Calculation of cosine similarity between the PCC of the daily maximum air-temperature of Era-Interim and the PCC of the sub-seasonal daily maximum air-temperature of five individual climate models applied by the hybrid model and Random Forest. From left, (a) is ECMWF, (b) is NCEP, (c) is Glossea5, (d) is UKMO, (e) is CMA.

## 7. 격자기반 계절내 예측 딥러닝 기법 개발

### 가. 개요

최근 단기예보와 중장기 예보를 잇는 3~4주의 계절내 예측자료는 이상기후 감시의 조기대응 및 현업 예측활용에 있어 갈수록 중요성이 증대되고 있다. 따라서 WMO(World Meteorological Organization)는 2013년부터 S2S(Subseasonal to Seasonal) 프로젝트 연구가 진행중에 있으며, 기상청 또한 GloSea5(Global Seasonal Forecasting Model Version 5)모델을 이용하여 계절내 예측시스템을 운영중에 있다. 하지만 한반도 및 동아시아 지역은 지면과 상호작용 등 다양한 변수로 계절내 예측에 어려움이 많은 것이 사실이다.

전통적인 기후모델의 기법과 다르게 최근 신경망 기법을 이용하여 예측기술을 향상시키고자 하는 연구가 활발하게 진행되고 있으며, 기존의 다양한 방법에서 발생할 수 있는 오차를 최대한 줄이기 위한 딥러닝 기술이 기후분야에 적용되고 있는 실정이다.

다양한 딥러닝 기술에서 시계열 예측에 유용한 순환신경망(Recurrent Neural Network, RNN)기법은 연속적인 문장 번역 등에 가장 많이 이용되면서 기후예측에 많이 이용되기 시작하였다. 하지만 장기간의 예측에서 초기의 입력정보들이 시간에 지날수록 사라지고 입력과 출력 데이터의 기간이 길어질수록 연관 관계가 줄어드는 장기 의존성(Long-Term Dependency)문제가 발생하기 때문에 이러한 장기의존성의 문제를 해결하기 위하여 LSTM(Long Short Term Memory) 기법이 발전하게 되었다.

2019년과제에서는 계절내 예측모형인 GloSea5와 ECMWF모델을 입력기반으로 LSTM기법에 적용하여 한반도 지역에서 3~4주에 잘맞는 딥러닝 기법을 산출하였다. 2020년 과제에서는 한반도 지역의 예측성 높은 정보를 제공하기 위하여 격자기반의 입력자료에 따른 ConvLSTM기법을 적용하여 지점인 한반도 지역의 예측성능을 향상시키는 연구를 수행하였다.

2021년 연구과제에서는 APCC의 서비스 지역인 동아시아지역의 기후정보 수요증가에 따라 격자기반 기후모델 입력자료에 따른 동아시아지역 최대온도를 산출할 수 있는 기법 연구를 수행하였다. 머신러닝 기법과 딥러닝 기법을 결합한 하이브리드 모델을 개발하고 다양한 민감도 분석을 실시하여, 최종적으로 최적화된 신경망 기법을 개발하고자 한다.

### 나. 자료 및 방법

#### (1) 자료

##### (가) 계절내 개별 기후모델자료

계절내 기후 예측자료는 WMO(World Meteorological Organization)의 계절내 예측 프로젝트 기반하여 다양한 기후모델 자료(ECMWF, GloSea5, NCEP, MSC, CMA, UKMO)를 제공하고 있다.

한국 기상청에서 운용하고 있는 Glosea5(Global Seasonal Forecasting System version5; MacLachlan et al 2015, Scaife et al 2014)는 Met Office 기후모델인 HadGEM3의 패밀리 형태의 모델로 고해상도 앙상블 모델로서 대기-지표-해양-해빙 결합모델로 대기모형은 Unified Model(UM)을 기본모델로 지면모델은 UK Land Environment Simulator(JULES; Best et al., 2011), 해양은 Nucleus for European Modeling of the Ocean(NEMO; Madec, 2008), 해빙모델은 The Los Alamos Sea Ice Mode(CICE; Hunke and Lipscomb, 2010)을 모델로 구성되어 있다. Glosea5는 N216공간해상도로 대략 위도 0.8°, 경도 0.5°의 해상도를 가지고 있으며, 1991년부터 2010년까지 자료를 전처리 하였다. (Figure 73). NCEP 계절내 예측자료를 1991년부터 2010년까지 12년간의 데이터에서 자료를 평균하여 MSLP-OLR-PREC-Q700-Q850-SIC-SST-T2M-TMAX-TMIN-U50-U200-U850-V200-V850-W500-Z200-Z500의 변수를 배열(12,43,61,49,19)의 입력자료로 변환하여 DB 구축완료하였다. UKMO는 1993년부터 2015년까지 23년의 자료를 앙상블 평균하여 OLR-PREC-Q700-Q850-SST-T2M-TMAX-TMIN-U50-U200-U850-V200-V850-W500-Z200-Z500의 변수 자료를 이용하여 리드타임 60일과 위도 -21~50, 경도 60-150에 대한 배열(23,60,61,49,16)으로 입력자료를 구성하여 DB구축하여 입력자료호 활용하였다. CMA자료는 2001년부터 2014년까지 14년간의 데이터를 평균하여 MSLP-OLR-PREC-Q700-Q850-SST-T2M-TMAX-TMIN-U50-U200-U850-V200-V850-W500-Z200-Z500의 변수를 배열(14,59,61,49,17)의 입력자료로 변환하여 전처리하였다. CMA자료는 Daily 자료로 Glosea5와 같은 리드타임에 맞추어 자료를 DB구축하였다. MSC자료는 1995년부터 2017년의 23년치의 자료를 다운받았으나 1995-1997년, 2015-2017년의 자료가 없는 상황으로 최종자료는 1994년부터 2014년까지 DB구축하였다. 변수는 MSLP-OLR-PREC-Q700-Q850-SST-T2M-TMAX-TMIN-U50-U200-U850-V200-V850-W500-Z200-Z500이다. ECMWF자료는 1998년부터 2017년까지 자료를 구축하였으며, MSLP-OLR-PREC-Q700-Q850-SST-T2M-TMAX-TMIN-U50-U200-U850-V200-V850-W500-Z200-Z500의 변수를 전처리 하였다.

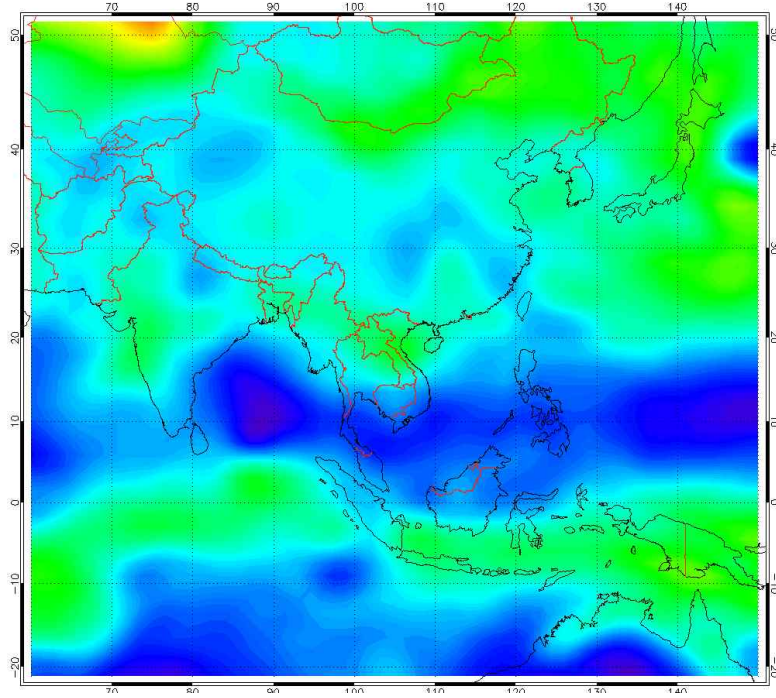


Figure 73. GloSea5 daily maximum temperature data.

#### (나) ERA-Interim

ERA-Interim는 대기모델 ERA-40에서 발전되어 제공하는 재분석자료로 1979년부터 2018년도까지 자료를 제공하고 있다. ERA-Interim자료는 다양한 변수와 격자화된 시공간 해상도의 자료를 제공하고 있으며, 모델과 관측자료를 결합한 재분석자료이다. 제공되는 자료는 1979년 1월부터 2019년 12월까지 자료를 수집하였고, 자료의 포맷은 netCDF, GRIB형태로 제공된다. 시간 해상도는 sub-daily, daily, monthly자료를 제공하며 공간적 해상도는  $0.75^{\circ} \times 0.75^{\circ}$  으로 해양 및 지표로 관측자료는 지상관측자료, 위성자료, 위성 radiance 및 모델 자료로 이루어져 있다(Figure 74). 재분석자료 ERA-Interim는  $1^{\circ}$  의 해상도로 전처리 하였다.

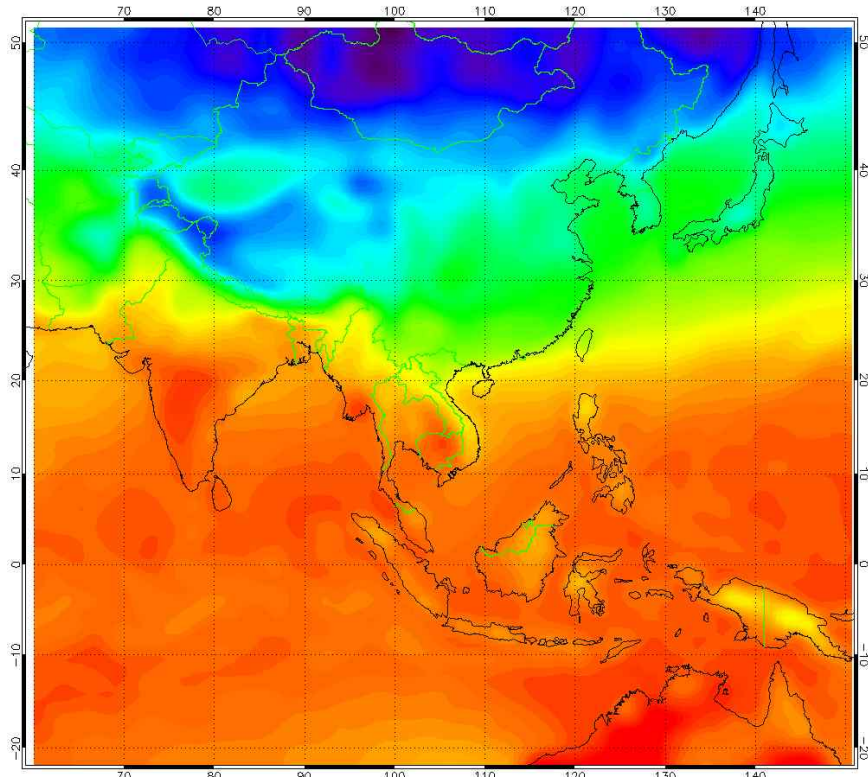


Figure 74. ERA-Interim daily maximum temperature data.

재분석자료인 ERA-Interim자료는 ECMWF, GloSea5, NCEP, MSC, CMA, UKMO의 개별모델과 MME 자료 기간에 맞추어 배열(61,49)의 형태로 리드타임 60일 자료를 구축하였고, 필요에 따라 개별모델의 리드타임에 맞추어 기간을 정해 전처리 하였다. 아시아 지역의 타겟 변수는 최대온도 값을 정하였고, 향후 최저온도, 강수에 대해서도 자료를 구축하여 학습할 예정이다.

ConvLSTM 모델의 입력자료는 개별모델 및 MME에 따라 틀리며, 같은 release date를 가지는 GloSea5, UKMO, CMA는 각각 1991-2010, 1993-2015, 2001-2014년까지 데이터로 각각의 모델은 일자료로 다양한 변수와 각기 다른 앙상블 개수를 입력자료로 활용하며, release date에 따라 월에 1,9,17,24일의 총 4개의 데이터셋으로 1년 48개의 총 데이터에서 한번 release date 60일 예측자료가 더해진다. 학습해야 할 변수는 최대 온도 자료이다. ECMWF는 11개의 앙상블 개수 자료와 1998년부터 2017년까지 자료를 구축하였고, NCEP자료는 1999년부터 2010년까지 4개의 앙상블 개수 자료를 평균하여 자료를 구축하였다.(Table 19).

Table 19. Summary of training dataset for ConvLSTM Model.

	Observed		S2S model			
	ERA-Interim	GloSea5	UKMO	NCEP	ECMWF	CMA
Period	1990-2018	1991-2010	1993-2015	1999-2010	1998-2017	2001-2014
Time scale	daily	daily	daily	daily	daily	daily
Time range	-	0-60	0-60	0-44	0-46	0-60
Ens. size		3	7	4	11	4
ConvLSTM input	release date(60 day)					
ConvLSTM output	daily Tmax					

## (2) 방법

### (가) 딥러닝 방법

한반도 지역외에 아시아 태평양 지역의 격자기반의 계절내 예측자료를 바탕으로 정확한 예측자료를 생산하기 위하여 ConvLSTM을 이용하여 예측모델을 개발하였다(Xingian Shi et al, 2015). ConvLSTM은 RNN의 하나의 방법으로 RNN은 연속되는 데이터 값을 딥러닝 모델 연속적으로 출력되는 값이 다시 입력 값에 다시 포함되는 방법으로 RNN딥러닝 모델은 시계열데이터에 좋은 결과를 제시하지만 장기간에 걸쳐 발생하는 예측값을 잘 예측하지 못하는 문제점이 발견되었는데 특정시점의 오차의 출력 기울기는 출력층에서 역전파해 온 출력값의 기울기와 이후 시점에서 거슬러 온 출력 기울기 합의 합이 되는데, 몇 개의 레이어를 걸치면서 오차를 저파시키면 기울기가 소실 또는 발산하는 문제가 있어, 특정시점의 정보가 일정 시간이 지난 시점에까지 가중치가 전달되지 않는데 이러한 문제점을 해결하기 위해 LSTM (Long Short-Term Memory) 알고리즘이 연구되어 장기간의 시계열 자료를 예측하는 연구에 많이 이용되었다.

한반도 지역의 한 지점의 시계열 자료의 다음값을 예측할 때 LSTM방법내에서 일반적으로 이용하는 FC(Fully-Connected)-LSTM을 이용하여 다양한 기후변수를 1차원 벡터로 입력하고 최종적으로 1차원 벡터 대상 변수 값인 TMAX, TMIN, PREC을 산출하는 방법은 이전 연구를 통하여 개발된 방법으로 계절내 예측모델을 이용하는데 유용한 방법중의 하나이다. 이러한 결과를 바탕으로 동아시아와 같은 지역에 격자기반의 계절내 예측의 정확도를 향상하기 위하여 한반도 한지점에서 사용한 1차원 벡터의 LSTM을 바로 사용하기엔 무리가 있어 2차원 벡터를 기반으로 시계열 자료 예측에 유리한 ConvLSTM을 선택하여 연구에 적용하였다. 2차원 배열을 가진 ECMWF, Glosea5, CMA, MSC, NECP, UKMO의 개별모델과 개별모델을 기반으로 MME한 자료를 입력자료로 데이터셋을 구성하고 각각의 모델에 맞는 데이터에 맞춰 관측자료 ERA-Interim자료를 전처리 하여 훈련 데이터셋을 구성하였다. 2차원 Glosea5 데이터는 빅데이터로 데이터 차원이 크게 증가하고 또한 공간적 상관성이 있는 데이터이므로 피쳐간의 특성을

고려한 CNN을 이용하는 것이 향후 정확도 개선에 합리적 방법이라고 판단된다. 본 연구에서는 LRCN(Donahue, 2017)에 기반하여 CNN으로 각각의 GloSea5 feature 특징 벡터를 추출하고 이를 기반으로 LSTM을 적용하였다(Figure 75).

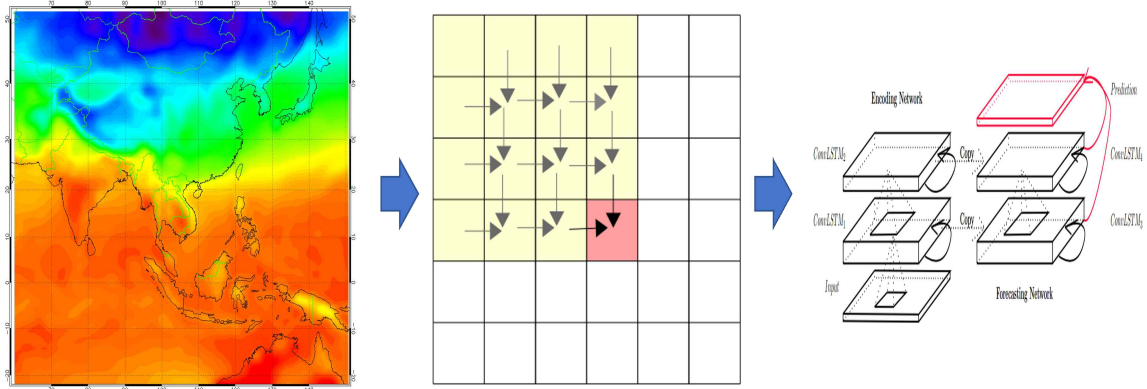


Figure 75. A diagram of ConvLSTM for Subseasonal model.

$$i_t = \sigma(W_{\xi}^* X_t + W_{hi}^* H_{t-1} + W_{ci} \odot C_{t-1} + b_i)$$

$$f_t = \sigma(W_{xf}^* X_t + W_{hf}^* H_{t-1} + W_{cf} \odot C_{t-1} + b_f)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_{xc}^* X_t + W_{hc}^* H_{t-1} + b_c)$$

$$o_t = \sigma(W_{xo}^* X_t + W_{ho}^* H_{t-1} + W_{co} \odot C_t + b_o)$$

$$H_t = o_t \odot \tanh(c_t)$$

ConvLSTM는 아시아 태평양 지역의 계절내 예측자료  $x_t$ 가 2차원의 ECMWF, GloSea5, MSC, CMA, NCEP, UKMO, MME 각자 모델의 격자 입력자료 피쳐가 들어가고 ConvLSTM의 입력, 망각, 출력 게이트  $i, f, o$ 와 셀 입력  $X$ , 셀 출력  $C$ , 셀 상태  $H$ 는 모두 3차원 텐서로 리드타임 3,4 주 3차원 벡터로 예측값을 산출하도록 설계하였다.

ConvLSTM은 개별모델과 MME모델은 MSLP, PREC, T2M, TMAX, TMIN, U200, U850, V200, V850, W500, Z200, Z500의 각기 다른 변수값을 Conv2D필터를 통해서 피쳐 특징을 뽑아내고 다시 MaxPooling을 통해서 차원을 줄이고, MaxPooling후에 다시 Conv2D를 실행하고 다시 Maxpooling을 실시하여 출력은 ERA-Interim과 맞도록 최종적으로 각각의 픽셀값의 시계열자료를 산출하는 FC-ConvLSTM 아닌 3차원 벡터를 한번에 리드타임 3-4주 자료가 산출이 되도록 Conv3D함수를 이용하여 한번에 출력하도록 ConvLSTM아키텍처를 설계하였다(Table 20).

```

from keras.models import Sequential

from keras.layers import MaxPooling2D, Dropout, Batchnormalization, Conv3D

from keras.layers import MinMaxScaler

from keras.layers import ConvLSTM2D

```

케라스에서 변수의 각기 다른 최대 최소값을 상대적으로 일정하게 맞춰 질수 있도록 정규화하기 위한 MinMaxScaler와 학습에 필요한 convLSTM2D, Conv3D 라이브러리로 활용하기 위해서 함수를 호출하고 개별모델자료와 ERA-Interim자료를 anomaly값으로 변환후 타겟변수인 Tmax에 대하여 훈련셋을 총 sample수에서 70%를 훈련셋으로 30%를 테스트셋으로 나누어 자료를 나누고 학습을 진행하였다.

```

with tf.device('/device:gpu:0'):

model = Sequential()

model.add(ConvLSTM2D(filters=64, kernel_size=(3, 3),input_shape=(2, lonnum, latnum, ffx),
padding='same', return_sequences=True, activation='tanh', recurrent_activation=
'hard_sigmoid', kernel_initializer='glorot_uniform', unit_forget_bias=True, dropout=0.3,
recurrent_dropout=0.3))

model.add(BatchNormalization())

model.add(ConvLSTM2D(filters=64, kernel_size=(3, 3),input_shape=(2, lonnum, latnum, ffx),
padding='same', return_sequences=True, activation='tanh', recurrent_activation=
'hard_sigmoid', kernel_initializer='glorot_uniform', unit_forget_bias=True, dropout=0.3,
recurrent_dropout=0.3))

model.add(BatchNormalization())

model.add(Conv3D(filters=1, kernel_size=(1, 1, 1),
padding='same', data_format='channels_last'))

model.summary()

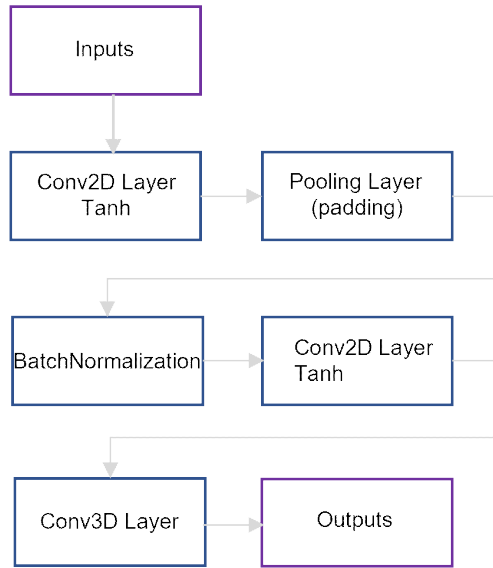
```

Table 20. Summary of convolutional layer at ConvLSTM model.

Layer (type)	Output Shape	Param #
conv_lst_m2d_560 (ConvLSTM2D)	(None, 2, 61, 49, 64)	159232
batch_normalization_560 (Batch Normalization)	(None, 2, 61, 49, 64)	256
conv_lst_m2d_561 (ConvLSTM2D)	(None, 2, 61, 49, 64)	295168
batch_normalization_561 (Batch Normalization)	(None, 2, 61, 49, 64)	256
conv_lst_m2d_562 (ConvLSTM2D)	(None, 2, 61, 49, 64)	295168
batch_normalization_562 (Batch Normalization)	(None, 2, 61, 49, 64)	256
conv_lst_m2d_563 (ConvLSTM2D)	(None, 2, 61, 49, 64)	295168
batch_normalization_563 (Batch Normalization)	(None, 2, 61, 49, 64)	256
conv3d_140 (Conv3D)	(None, 2, 61, 49, 1)	65

ConvLSTM모델은 입력자료의 배열이 아시아 태평양 지역의  $61 \times 49$ 의 배열과 평균 13개의 입력변수로 48주 각 주별로 계산하기 때문에 계산시간이 오래 걸리는 단점이 발생하였고, 이를 해결하기 위하여 배치사이즈를 크기 잡아 계산하였다. 레이어는 64를 4개로 이어붙여 적용하였고, kernel\_size는  $3 \times 3$ 으로 일반적으로 많이 적용하는 크기로 적용하였다(Figure 76). optimizer 함수는 'adam' 으로 손실함수는 mean square error (mse)함수로 선택하여 학습을 진행하였다. 대용량의 학습데이터에서 오는 과적합을 줄이기 위하여 batch normalization을 적용하였으며, padding='same', return\_sequences=True, activation='tanh', recurrent\_activation='hard\_sigmoid', kernel\_initializer='glorot\_uniform', unit\_forget\_bias=True, dropout=0.4, recurrent\_dropout=0.3으로 가장 많이 사용하는 옵션을 선택하여 학습하였다. epochs는 EarlyStopping을 적용하여 정확도가 나오면 멈출수 있도록 셋팅하였으며, 수동으로 300-1500회를 변화를 주면서 정확도 개선을 위한 민감도 분석을 실시하였다(Figure 77).

### ConvLSTM Architecture



#### Hyper-parameters

<i>Filters</i>	<i>=28</i>
<i>Layer</i>	<i>Dropout = 0.4</i>
	<i>Kerner sizes=3, 3</i>
<i>Training</i>	<i>Optimizer: Tanh</i>
	<i>Epoch = 300</i>
	<i>Batch size = 8</i>

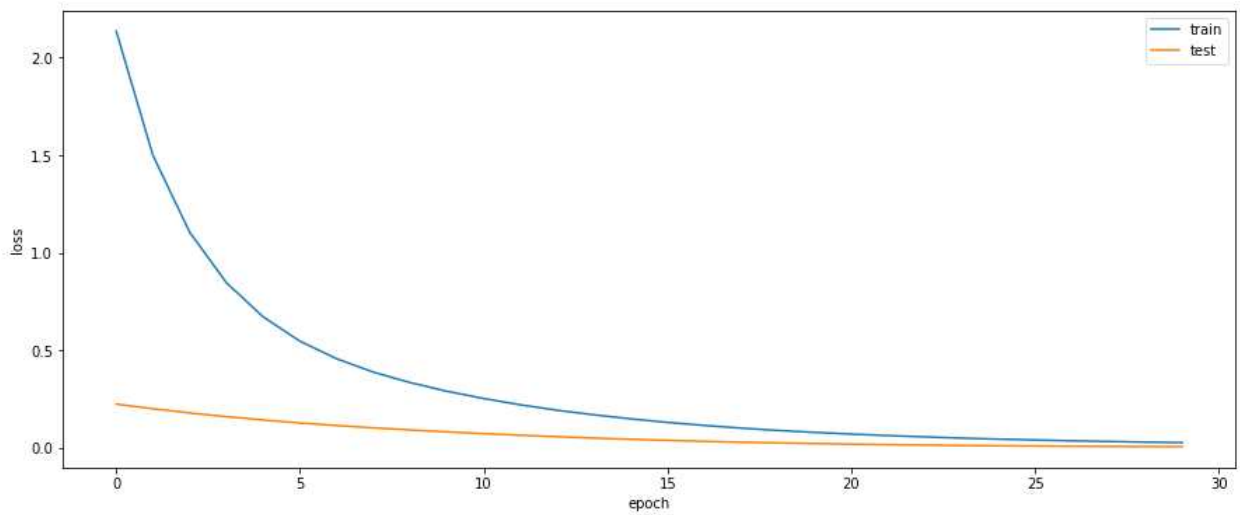


Figure 76. The structure of parameter of ConvLSTM.

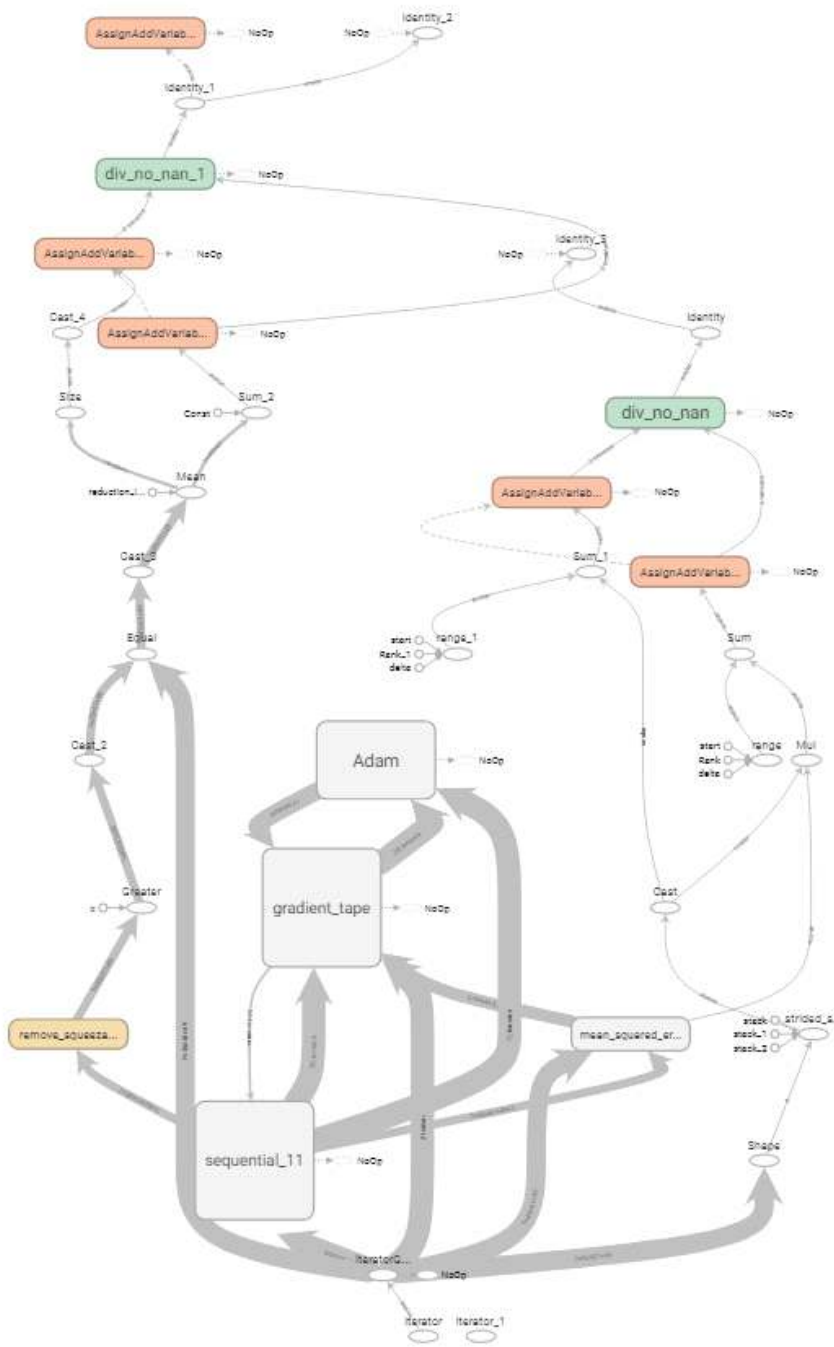


Figure 77. Model pattern Extraction and visualization of ConvLSTM

## (나) 계산방법

ConvLSTM의 입력자료는 5차원의 대용량의 자료와 4개의 레이어를 48주로 합성곱으로 계산하게 되면 PC의 CPU로 계산하기에 어려운 문제점이 있다. 딥러닝 모델의 계산시간을 효과적으로 단축하기 위해서는 GPU의 병렬컴퓨팅이 학습하는데 매우 효과적으로 계산시간을 단축할 수 있다. 현재 GPU 컴퓨팅의 서버를 구축중에 있기 때문에 최근에 발표된 Nvidia GTX3090을 기반으로 학습하였다(Figure 78). Nvidiacolab은 구글 클라우드에서 무료로 가상서버를 지원하기 때문에 가상서버를 통하여 부분 학습하였다. 구글 드라이브와 연동되기 때문에 훈련셋 자료는 구글 드라이브에 업로드하여 colab과 연동하면 colab에서 구글드라이브를 하드처럼 인식하여 사용가능하다. colab은 지속적으로 12시간 사용이 가능하기 때문에, 12시간보다 더 자원이 필요할 경우에는 연산자원에 대한 업그레이드가 필요하다.

NVIDIA-SMI 497.09			Driver Version: 497.09			CUDA Version: 11.5		
GPU	Name	TCC/WDDM	Bus-Id	Disp.A	Volatile	Uncorr.	ECC	
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute M.	MIG M.	
0	NVIDIA GeForce ...	WDDM	00000000:08:00:0	On	1%	Default	N/A	
0%	49C	P8	34W / 370W	1973MiB / 24576MiB			N/A	
Processes:								
GPU	GI ID	CI ID	PID	Type	Process name	GPU Memory Usage		
0	N/A	N/A	1776	C+G	Insufficient Permissions	N/A		
0	N/A	N/A	5704	C+G	...ge#Application#msedge.exe	N/A		
0	N/A	N/A	9996	C+G	C:#Windows#explorer.exe	N/A		
0	N/A	N/A	14644	C+G	...IPanel#SystemSettings.exe	N/A		
0	N/A	N/A	14744	C+G	...txyew#MiniSearchHost.exe	N/A		
0	N/A	N/A	23732	C+G	...20#Office110#Bin#hwp.exe	N/A		
0	N/A	N/A	25020	C+G	...artMenuExperienceHost.exe	N/A		
0	N/A	N/A	31052	C+G	...me#Application#chrome.exe	N/A		
0	N/A	N/A	45152	C+G	...n1h2txyew#SearchHost.exe	N/A		
0	N/A	N/A	45644	C+G	...cw5n1h2txyew#LockApp.exe	N/A		
0	N/A	N/A	46684	C+G	...TeamViewer#TeamViewer.exe	N/A		
0	N/A	N/A	52836	C+G	...y#ShellExperienceHost.exe	N/A		
0	N/A	N/A	55232	C+G	...oot#Office16#POWERPNT.EXE	N/A		

Figure 78. GPU monitoring and CUDA activity.

## (라) 민감도 분석

예측의 정확도를 향상시키기 위해서 다양한 민감도 분석을 위하여 MSLP, PREC, T2M, TMAX, TMIN, U200, U850, V200, V850, W500, Z200, Z500의 변수 중에서 가장 중요한 T2M, TMX, TMIN, U50, U200 5개의 변수를 머신러닝 기법을 이용하여 선택하고, 격자크기에 따른 정확도 변화여부와 MME를 통한 정확도 개선과 데이터를 결합하였을때의 정확도 개선여부를 알아보기 위하여 다양한 민감도 분석을 실시하여 정확도를 개선하고자 하였다. 더불어 ConvLSTM의 딥러닝 모델의 배치사이즈 및 dropout같은 파라미터 최적화를 통하여 민감도 분석을 실시하고 개선된 모델을 개발하고자 하였다(Figure 79).

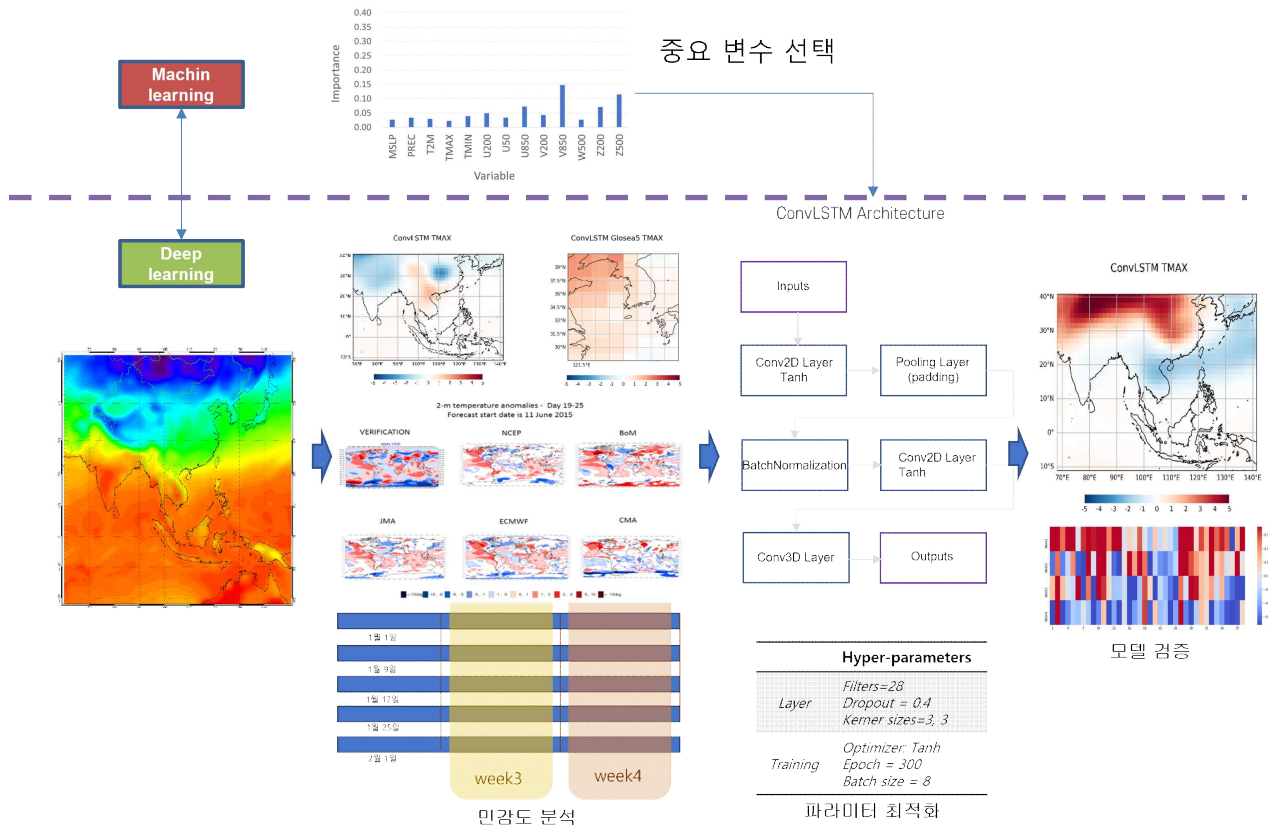


Figure 79. Sensitivity analysis of optimizing parameter and configuration of input data.

### (라) 검증

RMSE와 PCC를 사용하여 다양한 민감도 분석의 결과 예측 모델의 성능의 평가를 위한 통계적 방법으로 사용하였다.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (X_t - \hat{X}_t)^2}{n}}$$

$X_t$ 는 딥러닝모델을 이용한 예측된 값과  $\hat{X}_t$ 는 ERA-intrim을 이용한 RMSE값과 계절내 예측모델, MME과 ERA-intrim과의 RMSE값을 상호비교하여 정확도 검증을 비교하였다.

$$PCC(t) = \frac{\frac{1}{n} \sum_{y=1}^n (s'_{yt} - \bar{s}_t)(o_{yt} - \bar{o}_t)}{\sqrt{\frac{1}{n} \sum_{y=1}^n (s'_{yt} - \bar{s}_t)^2} \sqrt{\frac{1}{n} \sum_{y=1}^n (o_{yt} - \bar{o}_t)^2}}$$

PCC는 위의 식으로 계산할 수 있으며 S와 O는 각각 예측값과 관측값을 의미한다.

## 다. 결과

### (1) 개별모델 딥러닝 모델 학습결과

ConvLSTM의 정확도 검증을 하기 위하여 ECMWF와 ERA-Interim의 최대온도값의 anomaly 자료를 각각의 52주 45일 리드타임에 대해서 학습을 진행하여 PCC값을 산출하였다. Figure 80에서 볼 수 있듯이 첫째주 셋째주의 ConvLSTM 학습결과와 관측자료인 ERA-Interim자료, 관측자료인 ECMWF예측자료를 비교해보면 전체적으로 패턴은 비슷한 형태를 보이나 학습한 결과는 smoothing되어 나타나는데 데이터가 smoothing되는 현상은 sample 개수의 부족으로 인하여 충분히 학습되지 못한 결과일 것으로 판단된다. 52주 각각의 ConvLSTM의 예측된 값과 ECMWF값의 ERA-Interim의 PCC값을 상호비교하면 모델 성능이 거의 비슷하거나 조금 떨어지는 결과값을 보인다(Figure 81). Sample수가 거의 20년치의 훈련셋으로 다양한 Tmax의 패턴을 학습하기에는 데이터수가 떨어지는 것으로 보인다.

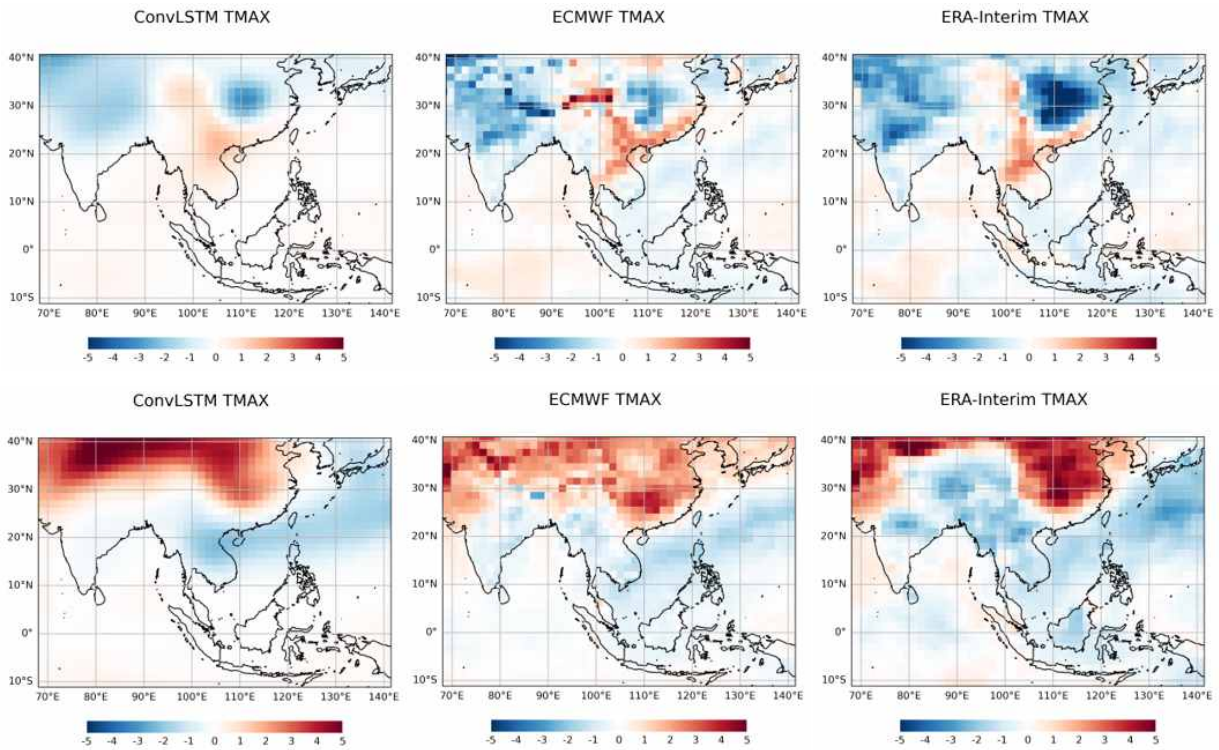
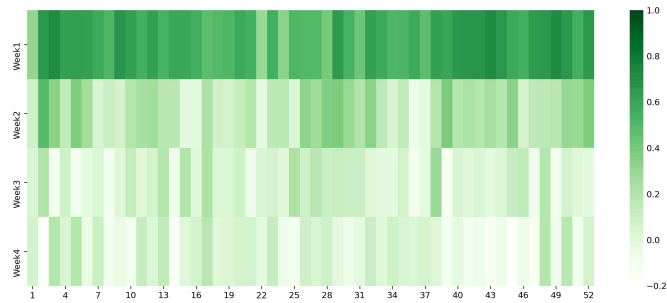
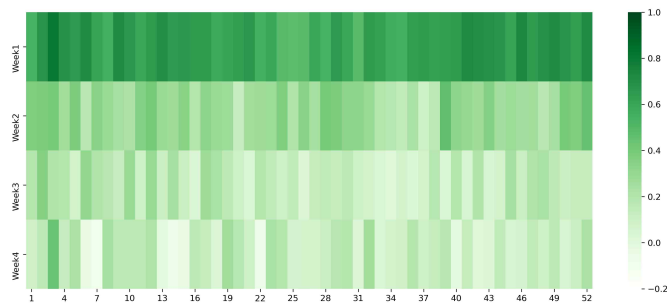


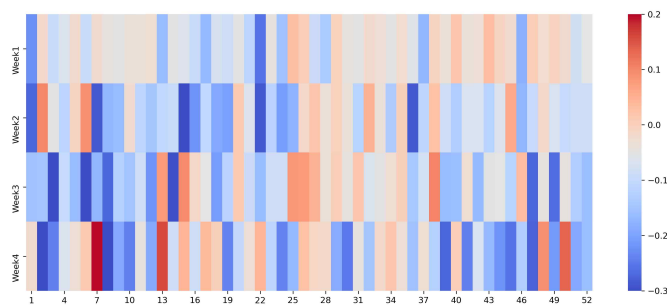
Figure 80. Comparison of the result of convLSTM model with ECMWF model at daily maximum temperature.



(a)



(b)



(c)

Figure 81. Comparison of the pattern correlation for (a) convLSTM (b) ECMWF.

ECMWF이외의 개별모델 GloSea5, NCEP, UKMO의 ConvLSTM의 3-4주의 정확도 검증을 하기 위하여 ECMWF와 같이 GloSea5, NCEP, UKMO와 ERA-Interim의 최대온도값의 anomaly 자료를 ConvLSTM아키텍처에 맞춰 학습을 진행하고 PCC를 이용한 검증 결과 GloSea5의 경우 딥러닝 모델 학습결과는 ECMWF에서 보여지는 결과와 비슷하게 나타난다. NCEP과 UKMO의 학습결과도 NCEP, UKMO의 원 자료에 비해서 3,4주에 대해서 높은 성능을 보이는 week도 있지

만 현저하게 성능이 떨어지는 결과도 보인다. 이는 데이터 sample수가 적어 데이터에서 학습된 패턴이 보이면 좋은 결과를 보이지만 그렇지 못한 경우에는 아예 학습이 되지 않는 것을 볼 수 있다. 52주 모든 학습을 한번에 진행하기 때문에 학습모델이 학습을 진행할때마다 달리 나타난다고 볼 수 있다(Figure 82).

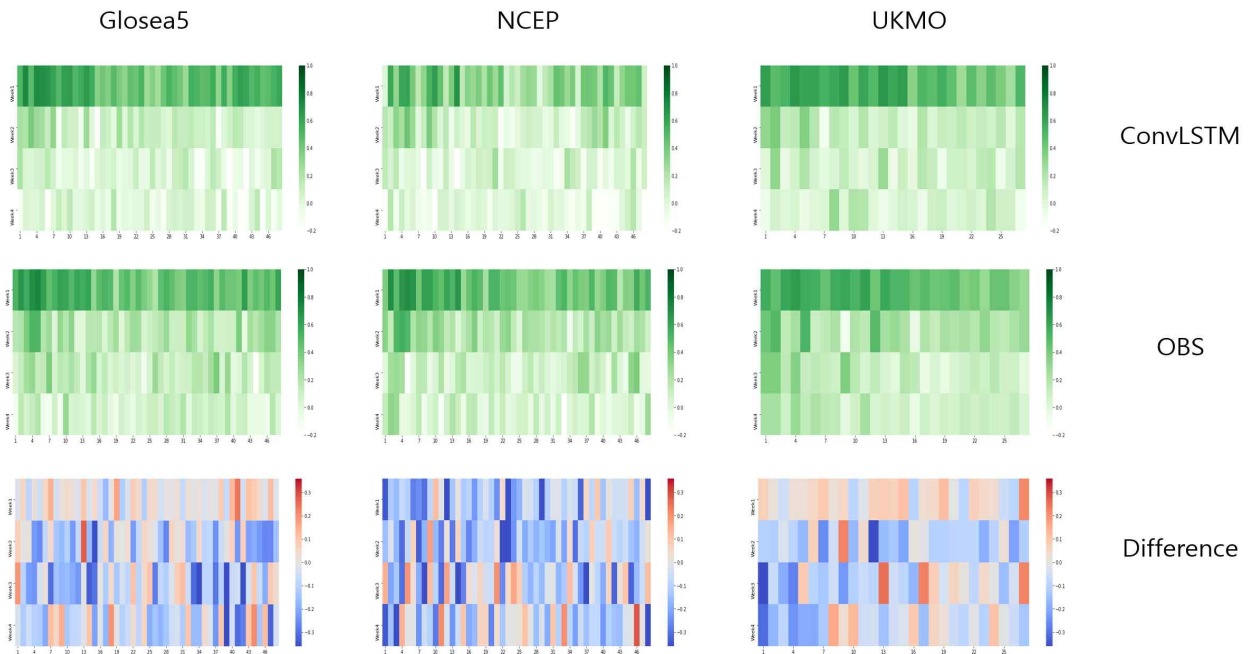


Figure 82. Comparison of the pattern correlation for 1-4 week of lead-time of GloSea5, NCEP, UKMO model.

## (2) 민감도 분석 결과

개별모델의 ConvLSTM의 격자기반 학습된 예측결과는 3-4주에 대해서 개별 기후모델보다 거의 비슷하거나 조금 낮은 결과를 보이고 있기 때문에  $61 \times 49$ 의 격자의 크기의 의한 낮은 정확도를 보이고 있기 때문에,  $61 \times 49$ 의 격자에 변화를 주는 민감도 분석을 실시하였다. 격자를 한반도 주변으로 한정하고  $9 \times 9$  배열로 격자 크기를 줄이면 정확도 향상이 되는지 ConvLSTM 아키텍처를 고정하고 학습을 진행하였다. 모델 입력은 머신러닝에서 제시한 중요변수 5개를 입력자료로 관측자료 ERA-Interim자료를  $9 \times 9$ 로 입력자료와 같이 전처리 하여 70%, 30%의 훈련셋과 테스트셋으로 나누어서 학습하였다. Convolution Layer를 겹겹이 쌓고 학습한 결과는 격자크기에는 크게 정확도 차이가 없는 것을 확인할 수 있었다(Figure 83).

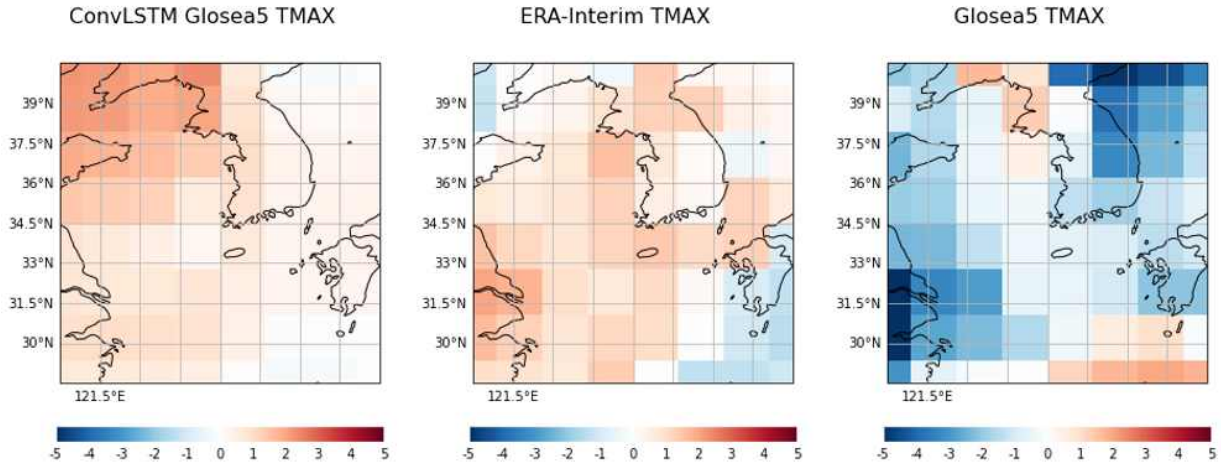


Figure 83. The sensitivity of analysis for spatial resolution.

sample 자료의 부족으로 발생하는 과적합을 보정하기 위하여 23년치의 UKMO자료를 기반으로 최대 5주의 자료를 확대하여 입력자료로 사용하였다. 데이터는 기준일(day)을 기점으로 day-1, day-2, day-3, day-4의 데이터를 합쳐서 sample 년도를 최대 115년치의 입력자료를 구성하고 훈련셋과 테스트셋으로 나누어 학습을 수행하였다. 머신러닝에서 산출한 중요변수 T2M, TMAX, TMIN, U50, U200의 입력자료를 선택하고 학습하였다. 리드타임 3주 평균 학습결과 평균 PCC는 0.31, UKMO는 0.322이며(Figure 84), 리드타임 4주차에서는 학습결과 PCC는 0.312, UKMO는 0.311로 거의 유사한 결과를 보였으나 PCC 통계값에서는 성능이 더 향상되지 못하는 결과를 나타내었다. (Figure 85).

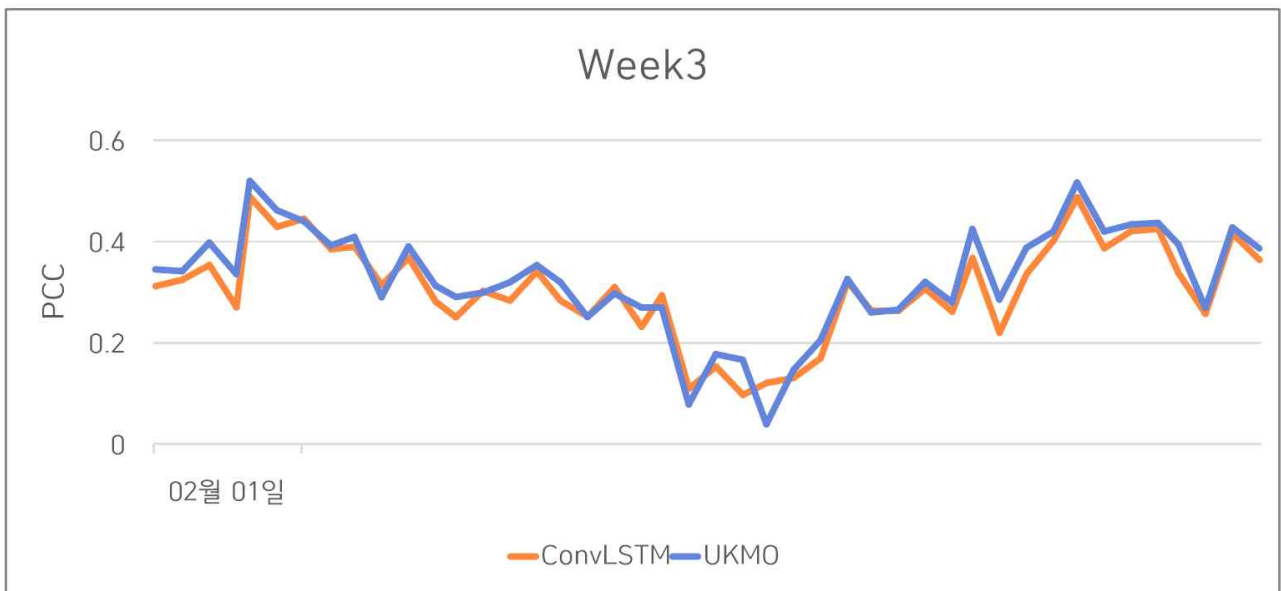


Figure 84. Comparison of performance of pattern correlation of maximum temperature for 3 week lead-time.

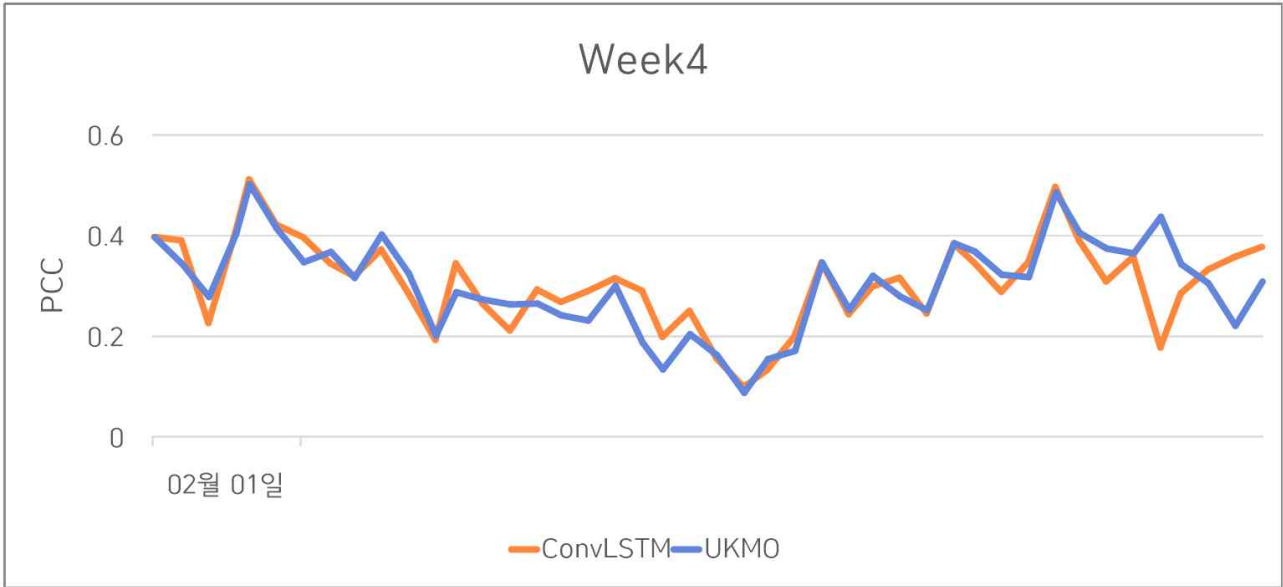


Figure 85. Comparison of performance of pattern correlation of maximum temperature for 4 week lead-time.

UKMO의 개별모델과 동일한 딥러닝 모델 파라미터 최적화 조건에서 GloSea5, ECMWF, NCEP, UKMO, CMA등의 자료를 앙상블로 결합한 MME자료를 이용하여 학습결과를 분석하였다. UKMO와 같은 기계학습에서 제시된 중요변수 5개를 입력자료를 이용하여 구성된 자료를 이용하여 예측성능 평가하였다. 다양한 epochs, batch size따른 민감도 분석을 통하여 최적화된 파라미터를 기반으로 학습한 결과는 Figure 86에서 알 수 있듯이 관측자료인 ERA-Interim자료의 공간패턴을 잘 모의하는 것을 알수 있었다.

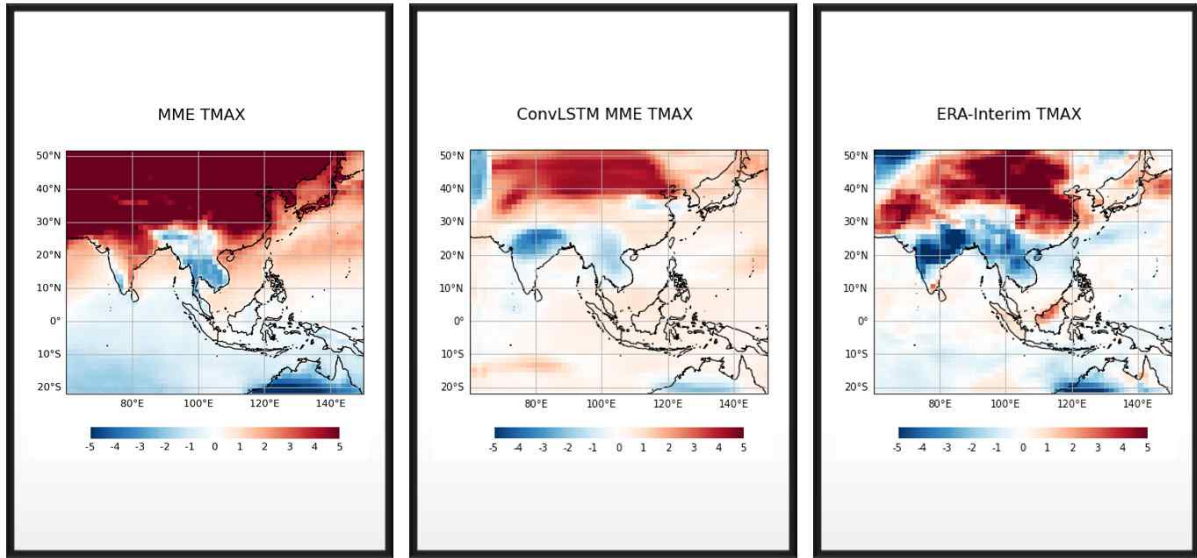


Figure 86. Comparison of map of the pattern correlation from ConvLSTM with MME data.

최종적으로 MME의 PCC 학습한 결과는 3-4주차에서 각각의 주별 패턴은 비슷한 경향을 나타내기는 하지만 개별모델과 같이 더 정확도 높은 성능은 보여주지 못하는 것을 알 수 있었다 (Figure 87). 한편, RMSE의 결과는 Figure 88에서 보이듯이 MME자료보다 학습된 자료에서 더 좋은 성능을 보여주고 있다. 적은 양의 sample 데이터는 공간상관성을 충분히 나타내기 어려운 자료이지만 각각의 픽셀에 대한 최적화 딥러닝 모델의 결과값은 RMSE를 줄여주는 역할을 한다. 향후 sample양의 확대와 최근 많이 이용되고 있는 딥러닝 모델들의 채택으로 현재 제한적인 모델의 결과가 향상될 것으로 본다.

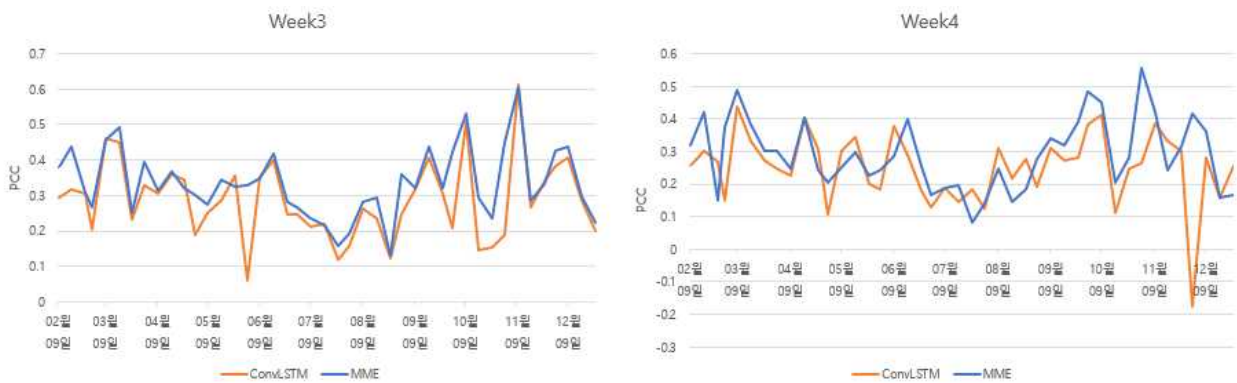


Figure 87. Comparison of pattern correlation at maximum temperature for 3-4 week of lead-time for MME.

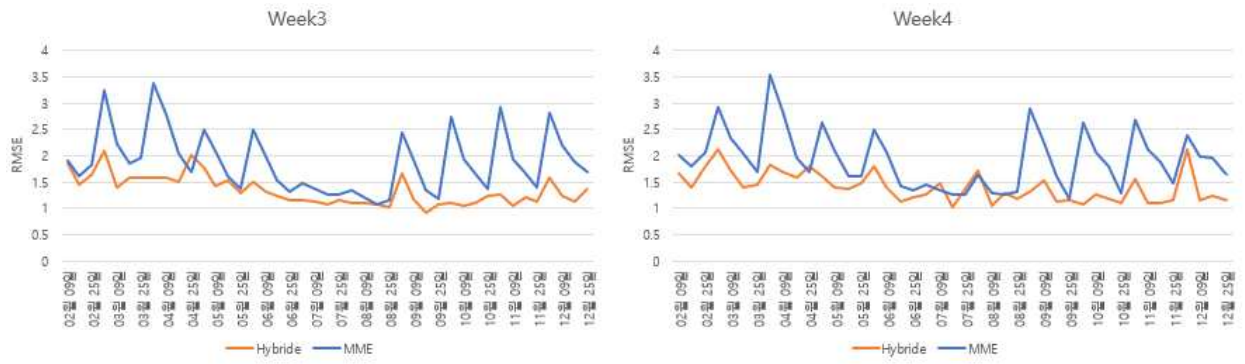


Figure 88. Comparison of RMSE at maximum temperature for 3-4 week of lead-time for MME.

## 8. 요약 및 결론

### 가. 요약 및 결론

유사한 기후 특성을 기반으로 하는 지역 규모에서 기후 전망이나 해석을 위해서는 지역 특화형 선도 기후예측 기술 개발이 이루어져야 한다. MME 계절기후예측의 예측성 근간은 앙상블에 참여하는 개별모델의 성능에 크게 의존하는데, 참여 모델 자체의 성능 개선은 센터 내에서는 수행하기 어려우며 모형 운영 기관의 모델 향상 계획에 전적으로 의존한다. 또한, 계절기후예측에서는 월별 자료만 제공하는 등 자료 가용성에 대한 한계가 존재한다. 따라서 센터 내 수집된 다중모델 자료를 바탕으로 기후예측 자료를 분석 및 해석하고, 다양한 통계적 기법을 적용하여 예측성 향상을 위한 계절 및 계절내 기후예측 기술을 개발하고자 하였다. 이 과정은 “지역특화형 선도 기후예측 기술 개발”을 최종목표로 하는 5단계 아태사업으로 총 2019년부터 2021까지 3년 동안 수행되었다. 1차년도에 구축된 지역 맞춤형 계절 MME 및 계절내 MME 기후예측 기술 개발 체계를 기반으로, 2차년도인 2020년에는 동아시아 및 한반도 지역을 중점 연구지역으로 설정하고 지역 특화된 다중모형 기반 계절 MME 및 계절내 MME 기후예측 기술을 개발하였다. 이를 기반으로 3차년도에는 지역 맞춤형 계절 예측 시스템을 구축하고 사용자에게 확률예측 기반의 통합기후전망을 제공하고자 연구결과의 현업화를 추진하였다. 또한 인공지능 기술을 고도화하여 목표 예측 지역을 동아시아로 확장하고 계절내 예측 기술을 개선하였다. 따라서 본 연구는 다음 4가지 중점 사항을 통해서 계절/계절내 예측성 향상을 수립하고자 하였다.

- (1) 후처리를 통한 우리나라 맞춤형 전문가 계절 예측 시스템의 개발,
- (2) 샘플링을 통한 동아시아 계절 확률 예측 생산 및 표출 시스템 개발,
- (3) 인공지능 기반 계절 확률 예측 개선,
- (4) 인공지능 기반 계절내 예측 개선

이 과제에서는 예측 시간 규모 및 예측자료의 유용성 등을 고려하여, (1) 예측성 향상을 위한 다중모델 기반 계절 예측 기술개발 및 (2) 다중모델 계절내 예측 기반 기술 개발로 구분하여 연구를 진행하고자 하였다. (1)을 위해서는 (가) 후처리를 통한 우리나라 맞춤형 전문가 계절 예측 시스템의 개발, (나) 샘플링을 통한 동아시아 계절 확률 예측 생산 및 표출 시스템 개발, (다) 인공지능 기반 계절 확률 예측 개선, (2)을 위해서는 인공지능 기반 아키텍처를 활용한 보정 및 시·공간 자료에 적합한 통합 인공지능 기반 보정 연구를 수행하였다.

#### (1) 후처리를 통한 우리나라 맞춤형 전문가 계절 예측 시스템의 개발

새로운 전문가 가이드를 적용한 ESPreSSO<sup>v3.2</sup>의 개발과 내부 현업화를 통해, 우리나라에도 민첩성/예측성/접근성이 크게 향상된 맞춤형 계절 기후 예측 정보를 제공하게 되었다. 이 연구를 통하여, 기존 전문가의 직접적인 개입을 통해 수행되었던 다중모델앙상블 자료의 후보정 작업을 전문가의 소극적인 개입(전문가 가이드)과 인공지능의 협력을 통한 예측 인자의 선별로 대체하여 환경 변화에 능동적으로 대응할 수 있는 민첩성을 크게 향상시켰다. 또한 APCC의 최

신 다중모델앙상블 구성을 반영한 새로운 전문가 가이드를 구축하여 예측성을 끌어올렸다. 한편으로는 다양한 방식의 예측 정보 제공, 자연어를 통한 예측 현황 설명과 더불어 계절 기후 예측 현황 모식도를 제공하여 예보자가 ESPreSSO가 수행한 계절 기후 예측의 근거를 쉽게 파악할 수 있도록 하여 접근성이 향상되었다.

## (2) 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술 개선 및 시스템 구축

1차 년도인 2019년에는 “지역 맞춤형 계절 다중모델앙상블 기후예측 기술 개발을 위한 체계가 구축” 되었으며, 이는 물리현상 기반의 앙상블 재구성을 통한 지역특화 계절 예측기술 개발을 통해서 그 체계가 마련되었다. 참여모델의 지역별 예측성능 평가 및 차별화된 MME 구성을 위해서 APCC MME 기후예측에 참여하는 모델의 후보자료(1982-2005)를 바탕으로 SST, 강수량 및 지표면온도의 예측성능이 분석되었다. 엘니뇨-남방진동의 복잡성과 그 원격상관의 예측성능을 진단하고, 지역 기후예측을 향상시키기 위한 물리적 인자로 사용하였다. 모델 선택의 효과는 계절 및 지역에 따라 달랐다. 물리현상 기반의 앙상블 재구성을 통해 Non-mature phase of ENSO 계절 동안 동아시아-호주 몬순지역 강수량의 계절 예측성능이 향상됨을 발견하였다.

2020년에는 전체 연구사업의 2차 년도로 1차 년도에 구축된 기술 개발 체계를 바탕으로, 지역 맞춤형 계절 MME 기후예측 기술을 개발하였다. 따라서 1차 년도에 개발된 물리현상 기반의 앙상블모델 재구성을 통한 지역특화 계절 예측기술을 개선하고 그 예측성을 검증하였다. 과제 수행 1년차에는 APCC 6개월 계절예측 현업 기준으로 MME version 1에 참여 중인 개별 기후모델을 바탕으로 연구가 진행되었다. 과제 2년차 중반에 현업 운영 중인 APCC 6개월 계절예측이 공식적으로 MME version 2로 변경되었다. 따라서 현업 활용성을 제대로 평가하기 위해서 다양한 기후 자료의 기후예측성능 검증을 통한 안정성 평가가 수행 되었다. 또한 1차 년도에는 모델의 충실성을 판단하기 기후 인자로 ENSO만을 고려하고 여러 대상 지역에 적용하여 결정론적 MME 예측기술을 개발하고 평가하였으나, 2차 년도에는 대상 지역을 동아시아로 특화하고 기후 인자를 추가로 발굴하여 ENSO와 함께 적용하여 확률론적 MME 예측기술을 개발하고 평가하였다. 2020년에는는 동아시아 강수예측을 좀 더 향상시키기 위해서, 동아시아 강수 변동성과의 물리적 관련성 및 모델 내 재현성을 바탕으로 열대태평양 강수 변동을 추가 기후인자로 발굴하였다. ENSO 및 서태평양 강수를 기준으로 각각 선별된 모델을 바탕으로 MME를 구성, 이를 합성한 새로운 확률 MME 예측 시스템을 개발 하였다. 이렇게 구성된 확률 MME 예측은 가용한 모든 모델을 바탕으로 한 기존의 MME에 비해서 카테고리 검증 측면에서 좀 더 예측성이 개선된 결과를 보였다. 또한 예측 검증의 기준을 기후값으로 하든 표준 예측 시스템 (기존의 MME)으로 하든 예측성능이 개선된 결과를 보여 줌으로써, 물리현상을 바탕으로 한 앙상블 재구성이 지역 특화 계절 예측 성능 향상을 위한 방안임을 증명하였다.

따라서 2021년에는 이런 연구 결과를 바탕으로 다중모델의 샘플링에 좀 더 적합한 확률예측 기법을 연구하고 기존의 방법과는 달리 3가지 관점에서 확률예측 기법을 개선하였다.

1. 모델의 기후학적 평균을 제거하여 모델의 계통적인 평균 오차를 보정하였다. 이는 PDF

추정 및 모델 간 합성 방법에 상관 없이 개선된 결과를 보였다.

2. 기존의 모수화 방법에서 탈피하여 경험적으로 순위를 매겨 확률밀도함수를 추정하는 방법을 사용하였다. 확률밀도함수를 추정하는 기후학적 기간이 짧은 APCC의 계절 예측 자료에 유용한 방법으로 보이며, 특히나 짧은 자료 기간을 사용할 때 모수화 추정과 관련된 문제로부터 자유로울 수 있다.
3. 확률 합성에 관해서는 개별 모델의 확률 산정 후 가중치를 주어 합산하는 과정에서 탈피하여, 선별된 전체 자료를 같은 샘플로 두어 확률밀도함수를 추정하여 가중치를 고려하지 않아도 되는 장점이 있다.

물리 인자 기반의 모델 선별을 통한 샘플링으로부터 확률 추정, 예측 결과의 검증까지 일련의 과정을 예측 시스템으로 구축하였다. 또한 이러한 결과의 통합 기후 전망을 제공하기 위하여 가시화 시스템을 개발하고 예측 시스템과 연계하였다.

### (3) 다중모델앙상블의 후처리를 통한 동아시아 예측성 향상

이 연구를 통해 개발된 맞춤형 인공지능망 후처리 모형을 통해 동아시아 지역의 계절 기후 예측 성능을 동아시아 강수 예측 범주 적중률 42.4% 수준(기존 성능 대비 9.5% 향상)으로 끌어올리는 한편, 연도별/계절별로 안정적인 예측 정보를 제공할 수 있게 되었다. 특히 이미지 처리 및 관계 예측에 사용되는 최신 기법과 아키텍처를 도입하고, 이를 계절 기후 예측에 특화시킨 맞춤형 인공지능망 후처리 모형을 구성하였다. 이 과정에서 채널별 커널의 도입, 채널-공간 어텐션 모듈의 사용, 동적 조건부 드롭-아웃의 개발, 그리고 다양한 차원의 우회로 신설을 통해 계절 기후 예측 성능을 향상시키고자 하였다. 또한, 맞춤형 인공지능망 후처리 모형의 학습 단계에서도 장기 관측 재분석 자료를 통해 사전 전이 학습을 수행함으로써 예측성을 확보하였고, APCC의 다중모델앙상블 계절 기후 예측 정보와 관측 강수 자료를 이용해 미세 조정 학습을 수행하여 최종적으로 약 9.5% 정도 향상된 범주 예측의 정확성을 확보하였다.

### (4) 다중모델 계절내예측 기반기술 개발: 머신러닝 및 딥러닝 복합 활용 기법 개발

여러 기후모델을 한 번에 고려하는 다중모델앙상블(multi-model ensemble, 이하 MME) 기법은 기후 예측의 신뢰성을 증가시킨다고 알려져 있다. 이에 본 연구는 가용할 수 있는 동아시아 지역 개별 기후모델 5개(ECMWF, NCEP, KMA, UKMO, CMA)로부터 MME 기반의 계절내(Sub-seasonal to seasonal, 이하 S2S) 예측 자료를 생산하고 기계학습 및 심층학습의 훈련 자료로 변환하였다. 또한, 모델에 입력되는 훈련 자료의 범위가 동아시아로 확장된 만큼 신경망 모델의 높은 분산을 줄이고 예측 성능을 강화시킬 수 있는 앙상블 학습 기법(이하 하이브리드 모델)을 개발하였다. 또한 구축된 동아시아 지역 MME 기반 S2S 예측 자료에 하이브리드 모델의 적용 및 예측 성능의 평가가 수행되었다. 구축된 동아시아 지역 MME 기반의 S2S 예측 자료에 개발된 하이브리드 모델을 적용하여 주별 예측 정확도를 평방근제곱근오차(root mean square error, 이하 RMSE)와 공간상관계수(pattern correlation coefficient, 이하 PCC)로 분석한 결과,

단일 기계학습 즉 랜덤포레스트(random forest) 혹은 단일 합성곱 신경망(convolutional neural network) 학습의 week1에서 예측 빈도별 낮은 PCC가 대부분 개선되었지만 여전히 보정 전보다 낮은 PCC가 분포하였다. 그러나 week3-4에서 예측 빈도별 높은 PCC가 뚜렷하게 증가함으로써 하이브리드 모델이 단일 모델보다 MME 기반 S2S 예측 자료의 예측성 개선에 기여하고 있음이 확인되었다. 동아시아 지역 MME 기반 S2S 예측 자료의 예측 성능 향상을 위한 하이브리드 모델 활용 기술에 대한 결과와 이 연구를 통한 교훈을 다음과 같이 요약하였다.

- (1) 후보정 기술의 예측 성능 향상 한계: 오차는 개선되나, 편차공간상관에서 1-2주 미개선 및 3-4주의 낮은 예측 개선 등 후보정 기법에 의한 예측 성능 향상의 한계 및 원인 등이 탐색되었다.
- (2) 개선 미흡의 문제 및 원인에 대한 고찰: 입력 변수의 결정보다는 네트워크 및 아키텍처의 구조, 혹은 전처리를 통한 데이터의 변형, 혹은 훈련 기법의 feature 특징 추출 등에 더 영향을 받는 것으로 추측된다. 또한, S2S 예측 훈련 자료에 적합한 전처리 혹은 훈련 기법, 아키텍처 등의 선정에 대한 레퍼런스의 부재 등으로 다양한 신경망 훈련 기법을 모두 테스트하기에 시간 및 비용 등이 부족하다.
- (3) 향후, S2S 예측 자료 개선에 최적화된 훈련 기법과 모델 구현 및 기술의 안정성 및 정확성 확보 필요: 아키텍처 및 네트워크 구조, 데이터의 변형, 하이퍼파라미터 등에 대한 민감도 분석 및 최적화를 통해 선례 연구에서 부딪친 문제점을 극복하고 개선 방향을 제시할 수 있을 것으로 기대한다.

동아시아 지역 MME S2S 예측 훈련 자료의 처리 과정은 기후학적 대기 이동이 고려된 한반도 및 동아시아 지역 MME S2S 예측 기후 정보를 생산하고, 머신러닝과 딥러닝 복합 활용에 기반한 하이브리드 모델의 파이프라인은 다양한 인공지능 학습에 기반한 MME S2S 예측 성능 개선을 위한 기술로 S2S 예측 자료를 다양한 기법으로 보정할 수 있는 S2S 예측 기술 다원화에 기여하였다. 그러나 1주에서 예측 성능이 퇴보되거나 2-4주에서 예측 성능이 낮아지는 등 모델의 문제 및 한계 등이 탐색된 만큼 향후 S2S 예측에 최적화된 아키텍처의 개발, 최적의 데이터의 변형 및 하이퍼파라미터 선정 등을 통해 모델의 안정성 및 정확성이 확보된 보정 기술의 개선이 필요하다.

## (5) 격자기반 계절내 예측 딥러닝 기법 개발

동아시아 지역의 3-4주의 계절내 예측자료 생산을 위하여 격자기반 딥러닝 프레임 ConvLSTM 구축과 정확도 높은 예측자료를 위한 민감도 분석을 실시하였다. 계절내 예측자료를 산출하기 위하여 입력자료는 ECMWF, GloSea5, NCEP, CMA, UKMO의 개별모델과 MME를 구축하여 각각의 개별모델의 활용가능한 기간의 자료와 변수를 이용하여 ConvLSTM 딥러닝 모델을 이용하여 훈련셋을 구축하고 재분석자료 ERA-Interim을 전처리하였다.

ConvLSTM 아키텍처는 4개의 Convolutional Layer를 연계하고 Conv3D 함수를 이용하여 입력자료가 3차원 벡터로 출력이 될수 있도록 모델을 구성하였다. 타겟변수인 최대온도에 잘 맞

을 수 있도록 정규화, dropout, batch size을 최적화하였다. 정확도 향상을 위하여 격자변화, 개별모델, MME 등 입력자료의 구성과 자료확장, 기계학습을 통한 변수선택 등의 민감도 분석을 실시하였다. 최적화된 모델과 민감도 분석을 통한 3-4주의 학습자료의 분석결과는 PCC는 원자료보다 비슷하거나, 성능개선이 되지 않았지만 RMSE는 모든 딥러닝 모델에서 성능이 개선되는 것을 확인할수 있었다.

계절내 기후모델과 이에 기반한 MME자료에 대한 아시아 지역에 대한 격자기반 ConvLSTM 딥러닝 프레임워크의 파라미터와 민감도 분석을 실시하였고, RMSE의 통계적 분석에서 성능개선을 확인하였다. 하지만 공간패턴의 정확도는 아직 개선이 없었으며, 이를 해결하기 위해서는 입력자료의 확장과 S2S에 보다 적합한 인공지능 모델의 개발이 필요하다. 또한 최근 다양하게 제시되고 있는 시계열 분석에 적합한 모델을 발전시키고, GPU기반의 컴퓨팅의 계산용량이 확충된다면 보다 개선된 예측자료 생산이 가능하며, 이 자료에 기반한 기후감시 및 기후분야 활용에 도움이 될 것이다.

## 나. 연구개발결과의 활용 방안 및 기대성과

이 연구는 2019년부터 2021까지 수행되는 다년도 과제로, “지역 맞춤형 계절 MME 및 계절내 MME 기후예측 기술 개발”을 완료 후 이를 발판으로 “다중모델 기반의 지역특화 계절 예측시스템 구축 및 계절내예측 기반기술 고도화”를 마무리하고, 최종적으로 “지역특화형 선도 기후예측 기술 개발”을 하였다. 이 연구개발결과는 크게 다음의 과학적·기술적 측면과 경제적·사회적 측면으로 활용 가능하다.

### (1) 과학적·기술적 성과

1. 한반도 및 동아시아 지역에 특화된 MME 기반 기후예측 기술 개발 및 예측 향상에 기여, 신뢰도 높은 예측자료 제공
2. 한반도 및 동아시아 내 3-4주 계절내 예측오차 저감 기술 개발 및 방향 제시
3. 계절내 예측 기술과 딥러닝 기술 융합에 기여

### (2) 경제적·사회적 성과

1. 세계기상기구의 계절예측 실효성 제고 방안인 지역 규모 계절 예측 생산 및 활용에 대한 권고사항 준수
2. 1개월(4주) 전망에 보조 자료로써 제공
3. 인공지능 예보에 계절내 예측 기술 참여 및 확대
4. 기후 빅데이터 활용 및 딥러닝 기술 확대를 통한 그린뉴딜 정책의 선도적 이행에 기여

따라서 궁극적으로 이러한 선도적 기술 개발을 통해서 동아시아 계절 및 계절내 기후예측과 그 향상에 기여할 것으로 기대한다.

## 사사

본 연구는 APCC의 지원을 통해 수행되었습니다.

## References

- Bauer P, Thopre A, Brunet G (2015) The quiet revolution of numerical weather prediction. *Nature* 52:47-55
- Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J., 1984. *Classification And Regression Trees* (1st ed.). Routledge. <https://doi.org/10.1201/9781315139470>
- Chattopadhyay, A., Hassanzadeh, P., Pasha, S., 2020. Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data. *Sci. Rep.* 10, 1317. <https://doi.org/10.1038/s41598-020-57897-9>.
- Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M.B., Choirat, C., Koutrakis, P., Lyapustin, A., Wang, Y., Mickley, L.J., Schwartz, J., 2019. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* 130, 104909. <https://doi.org/https://doi.org/10.1016/j.envint.2019.104909>.
- Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA (2001) Current approaches to seasonal to interannual climate predictions. *Int J Climatol* 21:1111-1152. <https://doi.org/10.1002/joc.636>
- Guilyardi E, Wittenberg A, Fedorov A, Collins M, Wang C, Capotondi A, von Oldenborgh GJ, Stockdale T (2009) Understanding El Niño in ocean-atmospheric general circulation models: progress and challenges. *Bull Am Meteorol Soc* 90:325-340
- Han, J., Kamber, M., Pei, J., 2012. 2 - Getting to Know Your Data, in: Han, J., Kamber, M., Pei, J. (Eds.), *Data Mining (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, Boston, pp. 39-82. <https://doi.org/https://doi.org/10.1016/B978-0-12-381479-1.00002-2>.
- He, S., Li, X., DeSole, T., Ravikumar, P., Banerjee, A., 2020. Sub-Seasonal Climate Forecasting via Machine Learning: Challenges, Analysis, and Advances. *arXiv* 1-24.
- Hernández, E., Sanchez-Anguix, V., Julian, V., Palanca, J., Duque, N., 2016. Rainfall Prediction: A Deep Learning Approach, in: Martínez-Álvarez, F., Troncoso, A., Quintián, H., Corchado, E. (Eds.), *Hybrid Artificial Intelligent Systems*. Springer International Publishing, Cham, pp. 151-162.
- Johnson, N. C., D. C. Collins, S. B. Feldstein, M. L. L'heureux, and E. E. Riddle, 2014: Skillful wintertime North American temperature forecasts out to 4 weeks based on the state of ENSO and the MJO. *Weather Forecast.*, 29, 23-38.

<https://doi.org/10.1175/WAF-D-13-00102.1>.

- Jolliffe, I. T., and D. B. Stephenson, Eds., 2011: Forecast Verification: A Practitioner's Guide in Atmospheric Science. 2nd Edition. John Wiley & Sons, 292 pp.
- Lang, Y., Luo, L., Ye, A., Duan, Q., 2020. Do CFSv2 seasonal forecasts help improve the forecast of meteorological drought over Mainland China? *Water (Switzerland)* 12, 1–14. <https://doi.org/10.3390/w12072010>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>
- Louppe, G., Wehenkel, L., Suter, A., Geurts, P., 2013. Understanding variable importances in Forests of randomized trees. *Adv. Neural Inf. Process. Syst.* 1–9.
- Miao, Q., Pan, B., Wang, H., Hsu, K., Sorooshian, S., 2019. Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network. *Water (Switzerland)* 11. <https://doi.org/10.3390/w11050977>
- Michaelsen, J., 1987: Cross-Validation in Statistical Climate Forecast Models. *J. Clim. Appl. Meteorol.*, 26, 1589–1600, [https://doi.org/10.1175/1520-0450\(1987\)026<1589:cviscf>2.0.co;2](https://doi.org/10.1175/1520-0450(1987)026<1589:cviscf>2.0.co;2).
- Misra, S., Wu, Y., 2020. Chapter 10 - Machine learning assisted segmentation of scanning electron microscopy images of organic-rich shales with feature extraction and feature ranking, in: Misra, S., Li, H., He, J. (Eds.), *Machine Learning for Subsurface Characterization*. Gulf Professional Publishing, pp. 289–314. <https://doi.org/10.1016/B978-0-12-817736-5.00010-7>
- Park, J., Kang, H.-S., Hyun, Y.-K., Nakazawa, T., 2018. Predictability of the Arctic Sea Ice Extent from S2S Multi Model Ensemble 28, 15–24. <https://doi.org/10.14191/Atmos.2018.28.1.015>
- Rasp, S., Pritchard, M.S., Gentile, P., 2018. Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci.* 115, 9684. <https://doi.org/10.1073/pnas.1810286115>.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. *Nature* 566, 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234–241). Springer, Cham.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., Woo, W., 2015. Convolutional LSTM

Network: A Machine Learning Approach for Precipitation Nowcasting.

- Sohn, SJ, WM Kim, JH Yoo, YY Lee, SM Oh, BR Kim, H Lee, S Kim, S Seuseu, and N Pelesikoti (2018) The Republic of Korea–Pacific Islands Climate Prediction Services Project, *Bulletin of the American Meteorological Society (BAMS)*, 99, 253–257, DOI: 10.1175/BAMS-D-17-0075.1.
- Thongsuwan, S., Jaiyen, S., Padcharoen, A., Agarwal, P., 2021. ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost. *Nucl. Eng. Technol.* 53, 522–531. <https://doi.org/https://doi.org/10.1016/j.net.2020.04.008>
- Timmermann, A., and Coauthors, 2018: El Nino&#8211;Southern Oscillation complexity. *Nature*, 559, 535&#8211;545, <https://doi.org/10.1038/s41586-018-0252-6>.
- Vitart, F., and Coauthors (2017) The Subseasonal To Seasonal (S2S) Prediction Project Database. *Bulletin of American Meteorological Society*, 98, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>
- Vitart, F., Robertson, A.W., 2018. The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Clim. Atmos. Sci.* 1, 3. <https://doi.org/10.1038/s41612-018-0013-0>
- Wang B, Lee JY, Kang IS et al (2009) Advance and prospectus of seasonal prediction: assessment of the APCC/ClipAS 14-model ensemble retrospective seasonal prediction (1984–2004). *Clim Dyn* 33:93–1117
- Wang, S., Cao, J., Yu, P.S., 2019. Deep learning for spatio-temporal data mining: a survey. *arXiv* 1–21. <https://doi.org/10.1109/tkde.2020.3025580>
- Wang, Y., Ren, H.-L., Zhou, F., Fu, J.-X., Chen, Q.-L., Wu, J., Jie, W.-H., Zhang, P.-Q., 2020. Multi-Model Ensemble Sub-Seasonal Forecasting of Precipitation over the Maritime Continent in Boreal Summer. *Atmosphere (Basel)*. 11, 157–172.
- Weldegebriel, H.T., Liu, H., Haq, A.U., Bugingo, E., Zhang, D., 2020. A New Hybrid Convolutional Neural Network and eXtreme Gradient Boosting Classifier for Recognizing Handwritten Ethiopian Characters. *IEEE Access* 8, 17804–17818. <https://doi.org/10.1109/ACCESS.2019.2960161>.
- Weyn, J.A., Durran, D.R., Caruana, R., Cresswell-Clay, N., 2021. Sub-Seasonal Forecasting With a Large Ensemble of Deep-Learning Weather Prediction Models. *J. Adv. Model. Earth Syst.* 13. <https://doi.org/10.1029/2021MS002502>.
- Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In&#160;Proceedings of the European conference on computer vision

(ECCV)&#160;(pp. 3-19).