

아태지역 15-60일 MME 예측 기술 개발 및 검증 기술 개발

Development of 15-60 Day Multi-Model Ensemble Prediction and
Verification Methods for Asia-Pacific Region

문자연, 이윤영, 오지현, 김가영 예측기술팀

윤순조 기후분석팀

발간사

APEC 기후센터는 지난 14년 간 신뢰도 높은 기후 예측정보를 생산하고 제공하기 위해 예측기술에 대한 개선과 개발을 꾸준히 추진해 오고 있습니다. 특히 아시아 태평양 지역에 다중모델앙상블 계절예측 정보를 생산하여 제공해 온 전문기관으로서 회원국을 대상으로 고품질의 기후정보를 서비스하며 궁극적으로는 아시아 태평양 지역의 기후예측의 품질을 향상시키기 위해 노력하고 있습니다.

계절내 예측은 이상기후로 인해 발생한 재난 대응을 위한 실질적인 많은 정책결정이 내려지는데 유용한 시간규모의 예측 정보로 알려져 있으며, 최근 전 세계적으로 이 시간 규모의 예측기술 개발에 대한 많은 투자와 관심이 집중되고 있습니다. 그러나 현재까지 계절내 예측기술의 발전은 미미할 뿐 아니라 계절내 예측에 대한 이해도 아직 초기단계에 불과한 실정입니다. APCC는 기상청으로부터 위탁 받아 운영하고 있는 WMO 장기예보 선도센터를 통해 WMO S2S(Subseasonal to Seasonal) 프로젝트에서 제공하고 있는 다중 모델 계절내 예측 데이터 베이스를 구축하고 있습니다. 또한 계절내 예측성에 큰 영향을 미치는 것으로 알려져 있는 여름철 계절내 진동의 예측정보도 다중모델에 기반하여 실시간으로 제공하는 등 예측의 불모지로 일컬어지는 계절내 예측에 있어 세계 어느 기관과 견주어도 뒤지지 않을 기술력과 경험을 쌓아오고 있습니다.

이 연구는 아·태지역을 대상으로 15-60일 이전에 발생하는 계절내 기후를 예측하기 위한 기술 개발을 목표로 1단계 첫 번째 과제를 시작하였습니다. 2018년도에는 계절내 예측 분야의 전반적인 예측성 파악을 위해 S2S 자료시스템에 구축되어 있는 역학모델들의 계절내 예측자료를 이용하여 각 모델 및 다중모델앙상블의 일

반적인 계절내 예측성능 및 다양한 조건변화에 따른 민감도 분석이 수행되었습니다. 또한 계절내 예측의 주요 잠재인자로 알려져 있는 Madden and Julian Oscillation(MJO) 및 블로킹 현상의 예측 가능성을 진단하고 각 현상들의 재현 특성을 제시하였습니다.

이 연구보고서가 발간되기까지는 많은 분들의 노력이 있었습니다. 연구에 참여하여 보고서를 집필한 문자연 박사, 오지현 박사, 이윤영 박사, 김가영 연구원, 윤순조 연구원에게 감사드립니다. 보고서의 수준을 제고하도록 논평을 해주신 전남대학교의 정지훈 교수와 이화여자대학교의 유창현 교수께도 감사드립니다. 이 연구보고서가 밑거름이 되어 향후 계절내 예측 관련 연구들이 이어져 예측 능력이 향상되길 바랍니다.

2019년 3월
APEC 기후센터
원장 권 원 태

ABSTRACT

This project was designed to evaluate the current status of subseasonal prediction and predictability and thereby provide results that can contribute to develop suitable prediction systems for a subseasonal time scale. In this study, three models (ECMWF, NCEP, and ECCC) that have common forecast specifications (e.g., initial date and interval, lead time, reforecast periods) are selected to provide a consistent evaluation.

The first aim of this project is to evaluate the general performance of reforecast, real-time forecast, and multi-mode ensembles (MMEs) of 3 models for 11 variables and 12 verification regions. In terms of climatology, the models display a disorganized long-term mean (LTM) pattern for some regions and variables and tend to underestimate interannual variation. Similarities in LTM spatial patterns decrease, although absolute biases increase as lead times increase, in all three models; however, this was expected. Predictability measured by signal to noise ratio (SNR), which is different from actual prediction skill, was the highest for NCEP but lowest in ECCC for the same ensemble size and references. Prediction skill analysis suggests that forecasts are reliable up to 2 or 3 weeks in most regions; the ECMWF has the longest predictability among three models. There is clear dependency between prediction skill and the months, regions, and variables; the features are very similar for reforecasts and real-time forecasts. Interestingly, global subseasonal prediction skill during boreal winter season was shown to be dependent largely on the variability of the ENSO, and the association becomes clearer with longer lead weeks. From the comparison of the four varying MME methods, which have different

combinations of reference period and weightings, we found that the grand ensemble merging MME (MME_g) has a skill that is at least comparable to, and perhaps even better than, the ECMWF model.

The Madden-Julian Oscillation (MJO), one of major sources of subseasonal predictability, and the Ural blocking affecting cold surges in East Asia were investigated to verify the skill and climatic features simulated by the models. According to the verification based on the prediction skill of the MJO, the ECMWF, NCEP, and MME closely resembled observations while the ECCC model underestimated the intensity, eastward evolution speed, and life cycle of the MJO. The prediction skills for the lead time, season, phase, amplitude, and the location of the MJO convection showed contrasting results based on the model and variables used. However, the ECMWF model outperformed in detecting the highest MJO prediction skill.

It is suggested that a modified index for Ural blocking be applied for subseasonal forecasts. Individual models succeeded in capturing Ural blocking with one week of lead time; however, as this increases the skill decreases. Although predicting a blocking event 3-4 weeks ahead is not reliable, the performance can be improved by up to two weeks by applying a simple MME.

목 차

1. 서론	1
2. 자료 및 분석 방법	3
2.1 모형 및 관측 자료	3
2.2 다중모형 양상불 합성	4
2.3 블로킹 및 MJO 정의	6
2.3.1 우랄 블로킹의 정의	6
2.3.2 MJO의 정의	7
2.4 검증 지수 및 검증 지역	7
3. 지역별 변수별 계절내 예측성 평가	11
3.1 세 모형의 기후값 및 예측 잠재력 비교	11
3.2 다중모형 양상불 기법에 따른 계절내 예측성	19
3.2.1 과거재현자료 성능 평가	19
3.2.2 실시간예측자료 성능 평가	29
4. 계절내 규모 기후현상의 재현 특성 및 예측성	38
4.1 MJO	38

4.1.1 MJO의 특징	38
4.1.2 MJO 의 재현 특성 평가	38
4.1.3 MJO 예측성 평가	44
4.2 블로킹	50
4.2.1 관측블로킹의 특징	50
4.2.2 개별모형의 UB 예측성능	52
4.2.3 MME의 UB 예측성능	56
5. 결론	59
5.1 계절내 예측성 평가(계절내 예측의 현주소 파악)	59
5.2. 계절내 기후현상의 재현 특성 및 예측성 제시	60
REFERENCES	63

1. 서론

대기 초기 조건의 영향을 받기에는 긴 2주 이상의 예보로부터 해양 변동성의 영향을 받기에는 짧은 2달 이내의 예보에 해당되는 15-60일(계절내) 예측은 날씨와 기후를 잇는 결정적인 연결고리의 역할을 함에도 불구하고 선행예측이 어려운 영역으로 간주되어 왔다(Brunet et al., 2010; WMO, 2012; Robertson et al., 2015). 소수의 기관에서만 생산하던 계절내 예측 정보가 최근 들어 대부분의 현업기관에서 계절내 예측시스템을 통해 시험적으로 운영되기 시작하고 있으며 Subseasonal to Seasonal (S2S) 예측에 관한 국제간 프로젝트가 2013년부터 착수하여(WMO, 2015) 5년간의 Phase I 단계를 마치고 2018년 11월부터 II 단계가 시작될 예정이다. S2S 프로젝트가 시작되고 2015년 S2S 자료시스템이 구축되면서 계절내 예측성을 향상시키고 계절내 변동을 발생시키는 주요 기후학적인 요인 및 현상을 이해하기 위해 모델링, 예측, 예보검증, 예측인자 발굴 등의 다양한 방면에서의 노력이 이루어지고 있다. 그동안 밝혀진 과학적인 성과를 요약하면 다음과 같다(WMO/WCRP, 2018).

- 7~10개 S2S 참여모형의 MJO 예측성은 선행시간 20일에서 0.5의 bivariate correlation 을 보이며 단 하나의 모형은 30일까지도 유지된다.
- 북대서양에 나타나는 MJO 원격상관은 통계적으로 유의한 시그널이지만 모든 모형에서 비교적 약하게 모의한다.
- MJO 예측성은 성층권의 Quasi-Biennial Oscillation (QBO)가 동풍이 우세한 상태일 경우 몇몇 모형에서 1주일까지 더 연장(개선)된다.
- 2013년 3월 서유럽의 한파는 서태평양의 강한 MJO가 부분적으로 영향을 준 것으로 제안되었다.
- 유럽-대서양 기후지역에 대해 S2S 모형들의 예측성을 평가한 결과, North Atlantic Oscillations (NAO)가 양/음의 위상일 경우 3주까지 예측성이 있고 그 외의 지역은 16일까지 있었다.

계절내 예측 정보는 기상 산업뿐만 아니라 농업, 보건, 수문 등의 응용분야에서도 점차

수요가 늘어나고 있으며 중규모 예측능력의 향상, 예측성의 주요 source의 지속적인 개발로 인해 계절내 예측에 관한 관심이 국제적으로 증가하고 있다. 그러나, 현재까지 밝혀진 과학적인 연구 결과를 포함하여 과거재현자료(reforecast) 및 예측 자료를 활용한 계절내 예측성은 여전히 초기 단계에 있으며 계절내 시간에 적합한 예보 기준, MME 기법, 예측성 검증 등에 대한 체계적인 정립이 되어 있지 않은 상황이다.

APEC 기후센터는 설립이래 MME (Multi Model Ensemble)를 활용한 기후예측 정보 생산 및 서비스를 지속적으로 수행해 오고 있으며 계절예측에 적용할 수 있는 다양한 MME 기법을 개발 및 도입하여(Min et al., 2009; Lee et al., 2011; Lee et al., 2013; Min et al., 2014) 안정적인 현업 계절예보를 하고 있다.

본 연구보고서는 기상-기후간 이음새 없는 예측시스템의 확립을 구현하기 위해 가장 시급하게 해결되어야 하는 계절내 예측을 위한 준비단계로, 계절내 예측자료의 생산조건(생산시기, 간격, 예측기간, 과거재현기간 등)이 동일한 3개의 모형자료를 이용해 계절내 예측성의 다양한 평가를 수행하고 계절내 규모에서 주요한 영향 인자를 선택하여 이들의 특성 및 예측성을 평가하고자 한다. 이로부터 향후 계절내 예측에 적합한 예측기술을 개발하고 계절내 예측자료를 생산함에 있어 근거자료로 활용될 수 있을 것이다.

본 보고서의 2장은 보고서에 사용된 자료와 연구방법, 용어의 정의 등이 작성되었고 3장은 기후값과 예측값을 이용해 다양한 계절내 예측성 평가 결과를 제시한다. 4장은 계절내 예측에 중요한 기후인자로 블로킹과 MJO에 대한 재현능력 및 예측성을 평가하였고 마지막 5장에서 결과를 요약하였다.

2. 자료 및 분석 방법

2.1 모형 및 관측 자료

본 연구에 사용된 계절내 예측자료는 S2S prediction project database (Vitart et al., 2017)에서 수집되고 있는 자료이다. 현재 총 11개의 모형이 S2S 프로젝트를 위해 자료를 제공하고 있으며 본 연구에서는 그 중 세 모형(ECCC, ECMWF, NCEP)의 자료를 활용하여 연구를 수행하였다. 세 모형은 예측날짜와 동일한 날짜에 과거재현자료를 제공하고 있으며, 세 모형 모두 매주 동일한 요일(목요일)에 예측을 수행하므로 본 연구과제에서 우선적으로 분석하는 모형으로 선정되었다. 각 모형들은 저마다의 예측 스케줄과 전산 자원에 따라 모형의 특성이 매우 상이하며(Table 1), 그에 따라 각 모형 간 비교를 위해서는 동일한 기준을 정하여 자료를 표준화할 필요가 있다. 따라서 용이한 비교 분석을 위해 예측자료의 수평 해상도를 모두 $1.5^{\circ} \times 1.5^{\circ}$ 로 변환하였으며, 시간 해상도는 24시간으로 하였다. 또한, 예측 시작날짜가 목요일인 예측 및 과거재현자료만을 사용하여 분석을 수행하였다. 추가적으로 모형의 예측성 등을 검증하기 위한 관측자료는 ERA-Interim 자료를 사용하며, 이때 자료의 수평해상도는 모형자료와 동일하게 수평해상도를 $1.5^{\circ} \times 1.5^{\circ}$ 로, 시간해상도를 24시간으로 변환하여 사용하였다.

Table 1. Description of subseasonal forecast models.

Center	Time Range	Resolution	Ens. Size	Fcst. Frequency	Refcst Period	Refcst Ens. Size	Refcst. Frequency
ECCC	32 days	$0.45^{\circ} \times 0.45^{\circ}$, L40	21	weekly	1995–2014	4	weekly
ECMWF	46 days	Tco639/319 L91	51	2/week	past 20 years	11	2/week
NCEP	44 days	T126 L64	16	daily	1999–2010	4	daily

결정론적 예측을 위한 기후값 계산은 각 모형별로 예측 시작날짜와 동일한 날짜(월, 일)에 해당하는 과거재현자료의 평균으로 구하였다. 아노말리 예측은 이렇게 구해진 각 모형별 기후값을 해당날짜의 예측자료에서 뺀 값으로 구해진다. 확률론적 예측에서는 기

후값 계산에 사용되는 것과 동일한 과거재현자료로 확률분포 값을 계산하게 된다. 다만, 확률론적 예측을 위한 확률분포를 구성함에 있어 일강수량 자료를 사용할 경우 무강수일이 다수 포함되면서 확률분포 상의 삼분위 지점을 구하기 어려운 경우가 발생한다. 따라서 확률론적 예측을 위해서는 주별 평균한 자료를 사용하였다.

2.2 다중모형 앙상블 합성

본 연구에서는 계절내 예측의 방법으로 결정론적 예측과 확률론적 예측을 모두 수행하였다. 결정론적 예측은 각 모형의 예측값에서 모형의 기후값으로 뺀 아노말리 값을 앙상블 합산하는 방법으로, 각 모형의 구조적인 오차를 적절히 줄여주는 효과를 보여준다. 확률론적 예측은 각 모형의 과거재현자료를 확률분포 함수(Probability Distribution Function, PDF로 표기)로 피팅한 후 삼분위 지점을 구하고, 각 모형의 앙상블 예측값이 해당 카테고리에 얼마나 포함되는지 확률을 계산한다. 이때, 확률론적 예측을 위한 PDF를 구성할 시 강수량 변수는 감마분포를 따르는 것으로 가정하며, 나머지 변수는 가우시안 분포를 따르는 것으로 정하였다. 이렇게 구해진 개별 모형의 예측결과를 다중모형 앙상블 예측으로 병합하기 위해 기존 계절 예측에서 사용된 단순 합성법(Simple Composite Method, SCM으로 표기)과 확률론적 다중모형 앙상블(Probabilistic Multi-Model Ensemble, PMME로 표기)이 적용되었다. SCM은 모형별로 계산된 결정론적 예측결과 값을 적절한 가중치를 주어 평균함으로써 결정론적으로 물리적 단위를 갖는 예측값을 제공하는 방법이다. PMME는 각 모형 및 세 카테고리별로 구해진 확률값을 모형간에 적절한 가중치를 주어 평균하는 방법이다.

계절 MME 경험에 따르면, 안타깝게도 모형의 과거 재현 자료의 총 기간은 제 각각이고 여러 모형에 대한 통일된 결과를 얻기 위해 공통 기간을 처리하는 과정에서 그 기간이 더 짧아지게 된다. 따라서 기후값의 안정성 확보가 쉽지 않고 관측과 비교를 위한 기간으로 충분하지 않으며, 일반적으로 통용되는(예를 들어, 과거 30년 기준) 보편적 예보 결과를 보장하지 못하고 그 활용성에 한계를 가져올 수 있다. 이를 극복하는 방안으로 모형 합산 과정의 통일성(공통기간)을 포기하고 전체 가용 자료를 모두 이용하여 기후값 기간을 늘리는 방안에 대한 논의가 센터 내에서 여러 차례 있어 왔다. 계절내 모형 자료 역시

3개 모형만의 공통 기간이 총 12년으로 굉장히 짧아 유사한 문제점을 예상할 수 있다. 또한, MME에 참여한 개별 모형들은 전반적인 성능도 다르고 운용과정에서 특화 여부에 따라 지역적 성능 차이가 크게 날 수 있다. 모형 가중치를 어떻게 주는지에 따라 전구 혹은 각 지역별 예측성능이 달라질 수 있기에, MME 과정에서 개별 모형의 기여도 결정은 매우 중대한 사항이다. 따라서, 본 연구에서는 아노말리를 계산하는 기준 기간을 과거 재현자료에 대한 모형간의 공통기간(1999-2010)과 개별 모형 고유의 기간으로 다변화하고 앙상블 평균시 가중치에 변화를 주는 방식으로, 아래 4개의 다른 방식의 합성기법을 적용하여 그 결과를 비교해 보고자 한다(Table 2).

- MME_c: 기후값 및 PDF 계산 시 공통기간을 사용하여 개별 모형의 결정론적/확률론적 예측을 수행함. 다중모델 앙상블 예측 시 세 모형의 결과를 동일한 가중치로 평균하는 방법을 적용
- MME_e: 개별 모형별로 제공되는 과거재현자료를 모두 사용하여 기후값 및 PDF를 계산하고 각 모형별 예측을 수행. 다중모델 앙상블 예측 시 세 모형의 결과를 동일한 가중치로 평균하는 방법을 적용
- MME_g: MME_c와 마찬가지로 모형 공통기간을 사용하여 기후값 및 PDF 계산. 다중모델 앙상블 예측 시 모든 모형의 개별 앙상블들을 동일 가중치로 합성하는 방법. 즉, 각 모형의 앙상블 개수를 가중치로 주어 세 모형 결과를 합성함
- MME_w: MME_c와 마찬가지로 모형 공통기간을 사용하여 기후값 및 PDF 계산. 앙상블 개수의 제곱근으로 가중치를 주어 세 모형 결과를 합성함

Table 2. Description of four multi-model ensemble variations applied in this study.

Abbr.	Abbr. for DMME/PMME	Reference period	Merging type
MME_c	SCM_c/PMME_c	common (1999–2010)	individual model equal weighted
MME_e	SCM_e/PMME_e	individual model original/extended period	individual model equal weighted
MME_g	SCM_g/PMME_g	common (1999–2010)	grand ensemble merged/ individual model ensemble size weighted
MME_w	SCM_w/PMME_w	common (1999–2010)	individual model root ensemble size weighted

2.3 블로킹 및 MJO 정의

2.3.1 우랄 블로킹의 정의

블로킹을 정의하는 방법은 다양하나, 본 연구에서는 겨울철 우랄 블로킹을 정의하기 위해 500 hPa의 지위고도장을 바탕으로 하는 Cheung et al. (2013, 2015)에서 사용된 우랄 블로킹 탐지 방법을 다소 수정하여 적용하였다. 먼저 아래의 (1)과 (2)를 동시에 만족하는 구역을 찾는다.

$$ZGN(\lambda) = \frac{Z(\lambda, \varphi_N) - Z(\lambda, \varphi_0)}{\varphi_N - \varphi_0} < -10m/deglat \quad (1)$$

$$ZGS(\lambda) = \frac{Z(\lambda, \varphi_0) - Z(\lambda, \varphi_S)}{\varphi_0 - \varphi_S} > 0 \quad (2)$$

where $\lambda \in [0, 357.5]^\circ\text{E}$, $\varphi_N = 80^\circ\text{N} + \Delta$, $\varphi_0 = 60^\circ\text{N} + \Delta$, $\varphi_S = 40^\circ\text{N} + \Delta$, $\Delta = -6^\circ, -3^\circ, 0^\circ, 3^\circ, 6^\circ$ latitude

다음으로 경도방향으로 최소 12° 에 해당하는 영역에 걸쳐 위의 조건 (1)과 (2)를 동시에 만족하는 경도대를 선별한 후 해당 영역의 가장자리를 각각 블로킹의 풍상측과 풍하측으로 선택한다. 위의 과정으로 선별된 블로킹 중 중심부가 $45^\circ\text{--}90^\circ\text{E}$ 에 위치한 경우만을 우랄 블로킹으로 간주한다.

2.3.2 MJO의 정의

본 연구에서 정의한 MJO는 Wheeler and Hendon (2004, 이후 WH)에서 제시한 방법으로 산출된다. 적도 지역(15°S-15°N)에 대해 평균한 OLR과 상하층(200 hPa과 850 hPa) 동서바람의 세가지 변수로 다변수 주성분 분석을 통해 지수를 산출하여 MJO 지수 강도($(PC1^2+PC2^2)^{1/2} \geq 1.0$)가 1.0 이상/미만일 경우를 강한/약한 MJO로 고려하게 된다. 아노말리를 생산하는 자료의 전처리 과정에서 1999-2010의 공통된 기후값과 세번째 harmonics까지 제거한 후 120일 이동평균을 추가로 제거하여 경년변동의 영향을 제외한다. 분석은 총 기간에 대해서와 여름(5~10월) 및 겨울(11월~이듬해 4월)을 6개월씩 구분하여 각각 비교하였다. MJO의 전파는 기존 연구들과 마찬가지로 8개의 위상으로 표출된다.

2.4 검증 지수 및 검증 지역

모든 예측은 1999-2010 기간 기후값에 대한 아노말리값 혹은 삼분위 카테고리를 그에 상응하는 관측카테고리와 비교하여 검증되었다. 기후값 및 결정론적 예보 검증에는 PCC/ACC와 제공근오차(Root Mean Square Error, RMSE)를 확률론적 예보 검증에는 LEPS, SR, ROC, BSS 그리고 HSS 지수를 적용하였다.

○ PCC (Pattern Correlation Coefficient)/ ACC(Anomaly Correlation Coefficient)

PCC는 두 공간분포에서 동일한 위치에 있는 값 간의 상관계수로, 주로 예측공간분포가 관측과 얼마나 유사한 가를 나타내는 지수로 쓰이며 -1~1의 범위에 분포하고 1에 가까울수록 두 패턴이 비슷함을 나타낸다. ACC는 PCC의 특별한 경우인데, 관측과 예측값의 직접적인 상관관계는 계절적 변동으로 인해 오해의 소지가 높은 값을 제공할 수 있으므로 아래 식과 같이 예측과 관측 모두에서 기후값을 제거하여 계산한 상관계수로 예측값을 검증한다.

$$ACC = \frac{\sum_{i=1}^n w_i (f_i - \bar{f})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^n w_i (f_i - \bar{f})^2 \sum_{i=1}^n w_i (o_i - \bar{o})^2}}$$

○ LEPS (Linear Error in Probability Space)

LEPS는 확률예보를 검증하는 지수 중의 하나로, 예측의 누적 확률 분포와 그에 상응하는 관측 확률분포 간의 거리를 정량화한 값으로 아래 식으로 표현된다(Potts et al., 1996).

$$LEPS = [3(1 - |P_f - P_v| + P_f^2 - P_f + P_v^2 - P_v) - 1] \times 100$$

여기서 P_f 는 예측의 누적 확률 분포 함수이고, P_v 는 관측의 누적 분포함수를 나타낸다. LEPS 지수는 -100~100 범위에 분포하고, 100에 가까울수록 좋은 예측이다.

Table 3. 2-by-2 contingency table for categorical forecasts.

		Event observed	
		Yes	No
Event forecast	Yes	A	B
	No	C	D

○ Bias

Bias는 이벤트가 관측에 비해 자주 예측되었는지 아니면 덜 예측되었는지를 측정한다. Table 3의 기호를 사용하여 아래와 같이 표현할 수 있다.

$$\text{Bias} = (A+B)/(A+C).$$

1보다 큰 값은 관측에 비해 자주 예측된 것을 의미하며 1보다 작은 것은 관측에 비해 덜 자주 예측된 것을 의미한다.

○ SR (Success Rate)

SR은 카테고리 예보를 검증하는 매우 단순하고 직관적인 지수로서 전체 예측 시도 중 카테고리를 맞춘 비율로 정의되며 지수는 0에서 1의 범위에 분포한다. 삼분위 예보에서 1/3 이하는 무작위로 뽑는 것보다 성능이 낮아 신뢰할 수 없고, 1/3~2/3의 범위는 중간정도, 2/3 이상은 높은 성능으로 판단할 수 있다.

○ ROC (Relative Operating Characteristics)

ROC는 카테고리 예보를 검증하는데 널리 사용되는 지수로서, 오경보율(FAR)에 대비 적중률(HR)의 비로 나타낸다. 이때, 적중률은(HR)은 실제 일어난 사건 중 예보가 맞춘 비율이고 오경보율(FAR)은 일어나지 않은 사건 중 예보가 틀린 비율이다. 적중률이 높을수록 오경보율이 낮을수록, 따라서 ROC가 높을수록 좋은 예보이다. ROC 스코어 값은 HR과 FAR를 잇는 선 아래 면적으로 정의되는데 0~1의 범위에서 분포하고 이때 1은 완벽한 카테고리 예측을 의미한다. ROC 스코어 0.5는 HR과 FAR가 동일한 경우로 유의미한 예보성능이 없다고 판단할 수 있다.

○ HSS (Hiedke Skill Score)

HSS는 랜덤 예측에 비해 예측의 정확도를 평가하는 척도로서, 아래와 같이 구해진다.

$$HSS = 2 \frac{(AD - BC)}{[(A + C)(C + D) + (A + B)(B + D)]}$$

HSS가 1에 가까울수록 정확한 예측을, 0에 가까우면 랜덤 예측과 그 성능이 비슷함을, 0보다 작을수록 랜덤예측보다 못한 정확도를 의미한다.

○ BSS (Brier Skill Score)

Brier score (BS)는 예측값과 관측값의 차이를 측정하며 아래와 같이 나타낸다.

$$BS = \frac{1}{n} \sum_{i=1}^n (p_i - o_i)^2$$

P_i 는 i 번째 예측의 확률이며 관측 O_i 는 i 번째 관측으로서 이벤트가 발생했을 시 1, 발생하지 않았을 때는 0을 갖는다. 따라서 BS의 값이 작을수록 예측값의 적중률이 높은 것을 의미한다. 그러나 발생이 희박한 사례의 경우에도 BS값이 작아지는 경우가 있다. 따라서 본 연구에서는 아래의 식과 같이 기후값에 대비한 BS값을 통하여 예측성능을 평가하고자 한다.

$$BSS = \frac{BS - BS_{clim}}{0 - BS_{clim}} = 1 - \frac{BS}{BS_{clim}}$$

BSS의 경우, $-\infty$ 에서 1까지의 값을 가지며 음의 BSS는 기후값에 의존하여 예측한 결과보다 예측성능이 떨어짐을 나타낸다. BSS값이 1에 가까울수록 예측성능이 좋은 것을 의미한다.

본 연구에서는 계절내 예측의 전반적인 성능을 WMO에서 권고한 12개 지역 도메인 (Table 4) 대하여 제시하였으며, 이를 통해 향후 타 연구와의 비교 등에서 객관성을 확보할 수 있을 것으로 예상된다.

Table 4. Geographic domain of 12 regions verified.

Regions	Latitude	Longitude
Globe	90°S-90°N	0°-360°E
Tropics	20°S-20°N	0°-360°E
S. Extratropics	< 20°S	0°-360°E
N. Extratropics	> 20°N	0°-360°E
East Asia	15°N-60°N	75°E-150°E
South Asia	10°S-35°N	60°-140°E
North America	10°-75°N	190°E-310°E
South America	60°S-10°N	270°E-330°E
Australia (or Australasia)	50°S-0°	110°E-180°E
Australia + S. Pacific	50°S-20°N	110°E-260°E
Russia (or Northern Eurasia)	40°N-80°N	25°E-190°E
East Middle	10°N-45°N	25°E-75°E

3. 지역별 변수별 계절내 예측성 평가

3.1 세 모형의 기후값 및 예측 잠재력 비교

현업 예보 기관 특히 MME에 집중하는 본 센터에서 모형 기후값은 단순히 오차보정을 위해 제거되는 값으로 여겨지는 경향이 있다. 모형이 모사하는 세상이 실제 세상과 차이가 있기에 매우 당연한 과정이지만, 대기 배경장과 변동 성분 사이의 밀접한 역학적 관계 때문에(Franzke et al., 2011; Lau and Boyle, 1987; Luo et al., 2008; Nakamura et al., 1987) 기후값 오류에 대한 진단연구는 모형성능 개선을 위해 필수적이다. 본 연구에서도 계절내 모형 기후값이 얼마나 관측과 유사한지, 선행시간에 따른 변화는 없는지, 변수별 지역별로 분석하였다. 이 때 기후값은 모형 공통기간(1999-2010)을 기준으로 산출되었다. Figure 1의 가로축은 기후값의 오차를 나타낸 것으로, 선행시간이 1일(동그라미)에서 32일(별표)로 길어짐에 따른 추이와 전구/열대/동아시아 지역적 차이를 확인할 수 있다. 초기 선행시점에서 세 모형의 상층 지위고도장이 전반적으로 관측과 유사하고, NCEP을 제외하고는 선행시간이 증가함에도 오차가 크게 변하지 않는다. 하지만 NCEP의 경우 선행시간이 길어짐에 따라 지위고도장을 크게 과소모의하는 것으로 나타난다. 하층 기온의 경우, ECMWF와 ECCC는 초기 선행시점에서는 관측과 유사하나 선행시간이 길어짐에 따라 오차가 증가한다. 하지만 그 경향은 두 모형에서 반대로 나타나는데, 특히 전구에서 ECMWF는 과소모의, ECCC는 과대모의 경향이 뚜렷하다. 두 모형과 달리 NCEP은 전구와 적도에서 특히 초기 선행시간에서 지표온도를 과대모의하고 선행시간이 증가함에 따라 관측에 가깝게 모의하고 있다. 강수의 경우, 초기 선행시점에 세 모형, 세 지역에서 매우 다른 경향을 보이고 선행시간이 길어짐에 따른 뚜렷한 추이를 확인할 수 없다.

계절 예측 모형에서 앙상블 평균 예측의 변동성이 과소모의되는 것은 잘 알려져 있으나, 계절내 예측 모형에서도 이러한 특성이 유지되는지에 대한 답을 얻고자 하였다. 예측 날짜에 대한 경년변동성이 실제 세상과 얼마나 다른지 표준편차값의 오차를 통해 확인할 수 있었다(Figure 1의 y축). 변수/지역/모형에 관계 없이 선행시간이 증가할수록 변동성이 감소하는 경향이 뚜렷하여 예상을 벗어나지 않는다. 하지만 감소폭이 모형/지역별로

다소 차이가 있고, 강수의 경우는 초기 시점에서 동아시아의 변동성이 크게 과장되어 있음을 확인할 수 있다.

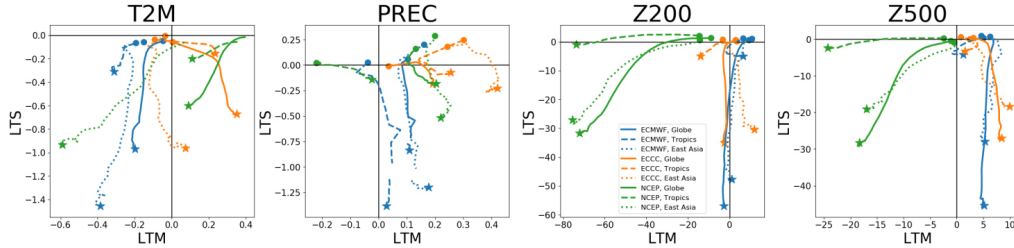


Figure 1. Temporal evolution of biases in long term mean (LTM) and long term standard-deviation (STD) of four variables with increasing lead days from 1 to 32 for three models. Bias is defined as a median value of differences relative to the observed LTM and LTS over Globe, Tropics and East Asia domains during all year round.

기후값 공간분포 유사성 및 절대값 차이를 검증하고자 PCC/RMSE 지수를 적용한 결과는 Figure 2와 같다. 선행시간이 증가함에 따라서 기후값 재현성능이 떨어질 것이라는 기본적인 가정에 기반한 분석으로, 대부분 지역/변수/모형에서 공간분포 유사성이 떨어지고 절대오차가 증가하여 예상에 부합한다. 상층 지위고도장의 경우 기본 패턴에서 남북 방향 경도가 주를 이루고, 따라서 PCC 통계 특성상 전구와 동아시아에서 성능 감소가 거의 나타나지 않는다. 초기 선행시점에서의 성능과 선행시간에 따른 성능감소폭 등을 종합적으로 고려할 때, 기후값 재현성이 가장 좋은 모형은 ECMWF이며 이는 타 모형성능 평가와 일치하는 결과이다. 나머지 두 모형은 우열을 가리기 어려우나 초기 선행시점 성능의 중요도를 고려할 때, ECCO가 NCEP에 비해 다소 관측과 유사한 기후를 창조했다고 판단된다.

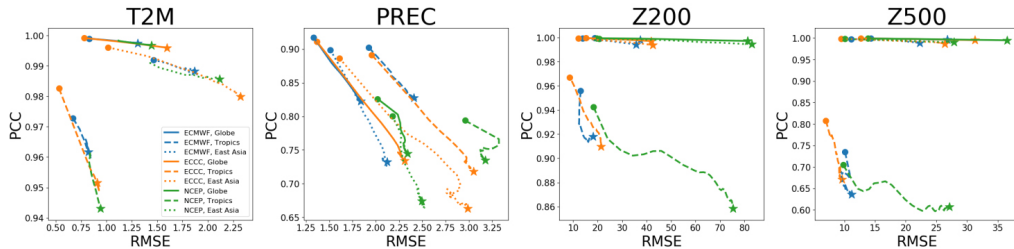


Figure 2. Temporal evolution of PCC and RMSE of long term mean (LTM) pattern of four variables with increasing lead days from 1 to 32 over Globe, Tropics and East Asia domains for three models during all year round.

모형의 잠재적인 성능은 모형 내부 노이즈 대비 시그널의 비율(signal to noise ratio, SNR)로도 판단할 수 있는데(Smith et al., 2014; Smith et al., 2016), 시그널은 앙상블 평균의 경년변동으로 노이즈는 앙상블 스프레드의 기후값으로 정량화된다. 이는 모형 예측성의 최대 한계치에 대한 개념이며, 관측과의 유사성/차이점에 기반한 정확도 검증과는 다르게 받아들여야 한다. 본 연구에서는 매 예측선행시간에서 SNR를 계산하고 이 값이 1보다 감소하는 날짜를 최대 예측일로 정의하였다. Figure 3은 12 지역, 11 변수에 대한 예측일의 12달 분포를 나타낸 것이다. 전반적으로 NCEP의 예측일이 다른 두 모형에 비해서 길고, 이러한 특성은 해양을 많이 포함하는 열대, 남아시아, 남태평양 지역에서 뚜렷하게 나타난다. 반대로 북반구 전체, 북미, 러시아 지역에서는 차이가 크지 않다. 하지만, ECMWF에서도 열대, 남아시아, 남태평양 등 지역에서 상층 지위고도 예측일이 NCEP에 필적할 만큼 길다. ECCC는 남아메리카에서 하층 온도의 예측일이 상대적으로 길다.

전체 SNR 평가에서 시그널과 노이즈가 각각 얼마만큼의 공헌을 하는지를 확인하기 위해서 전구 개별 격자에서의 두 값에 대한 2차원 히스토그램 분석을 수행하였다. 선행시간에 따른 추이를 보기 위해서 7일씩 구분하여 분석하였고, 대표로 하층 기온에 대한 결과를 제시한다(Figure 4). 초기 선행시간에서는 패턴의 대부분이 대각선 위쪽에 분포하지만, 시간이 흐를수록 최대 등치선(음영)이 대각선에 접근하고 4주 선행시간에서는 대부분의 패턴이 대각선 아래쪽에 분포한다. 이는 선행시간이 증가할수록 SNR가 감소하는 모형의 특성을 그대로 드러낸다. ECMWF와 ECCC의 경우, 긴 선행시간으로 갈수록 시그널이 크거나 노이즈가 작은 쪽의 분포가 확연히 줄어들고 최대 등치선(음영)이 대각선에 직각 방향으로 접근하는 특성을 보여, 시그널과 노이즈의 영향이 유사한 수준으로 판단된다. 하지만, NCEP의 경우는 시그널 감소폭에 비해 노이즈 증가폭이 뚜렷하고 최대 등치선(음영)이 대각선에 우측으로 접근하는 양상을 보이며, 4주차에서도 다른 모형에 비해서 작은 노이즈 분포를 보인다. 따라서, NCEP 모형의 최대 예측일이 긴 것은 과소한 앙상블 스프레드 영향으로 분석된다. ECCC는 두 모형과 다르게 초기시점부터 노이즈 값이 상대적으로 크고 선행시간 증가에 따른 증가폭이 크지 않다는 특징이 있다. 앙상블 개수가 NCEP처럼 4개뿐임을 고려하면 매우 이례적인 결과이지만, 그 원인은 확인할 수 없었다. 다만 ECCC가 다른 두 모형과 달리 해양 접합이 없는 대기모형이기 때문에 추후 그 관련성을 점검할 수 있겠다.

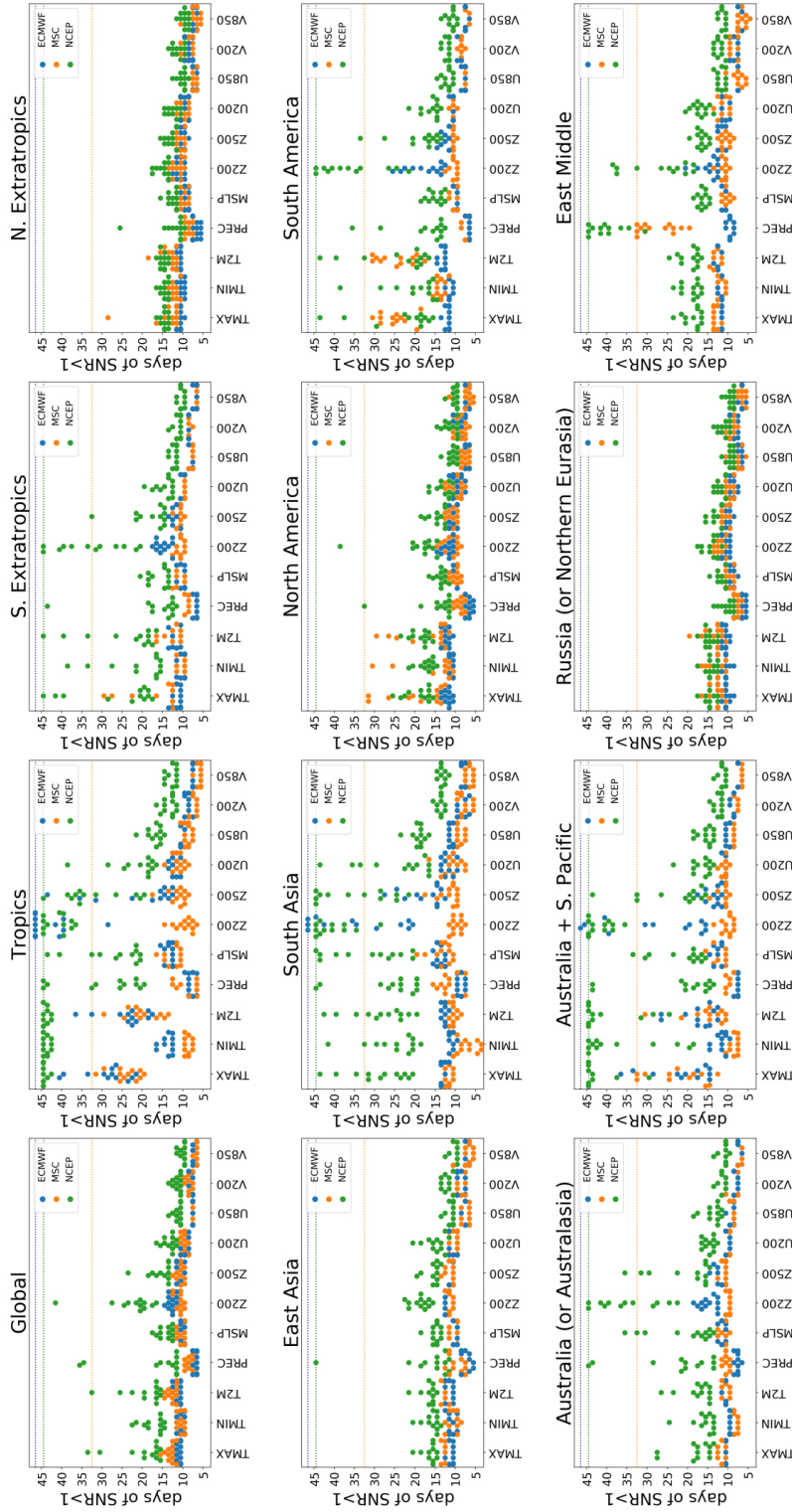


Figure 3. Distribution of predictable days defined as the day when SNR exceeds 1 over 12 regional domains for three models for 11 variables. Individual dots represent the values of 12 months.

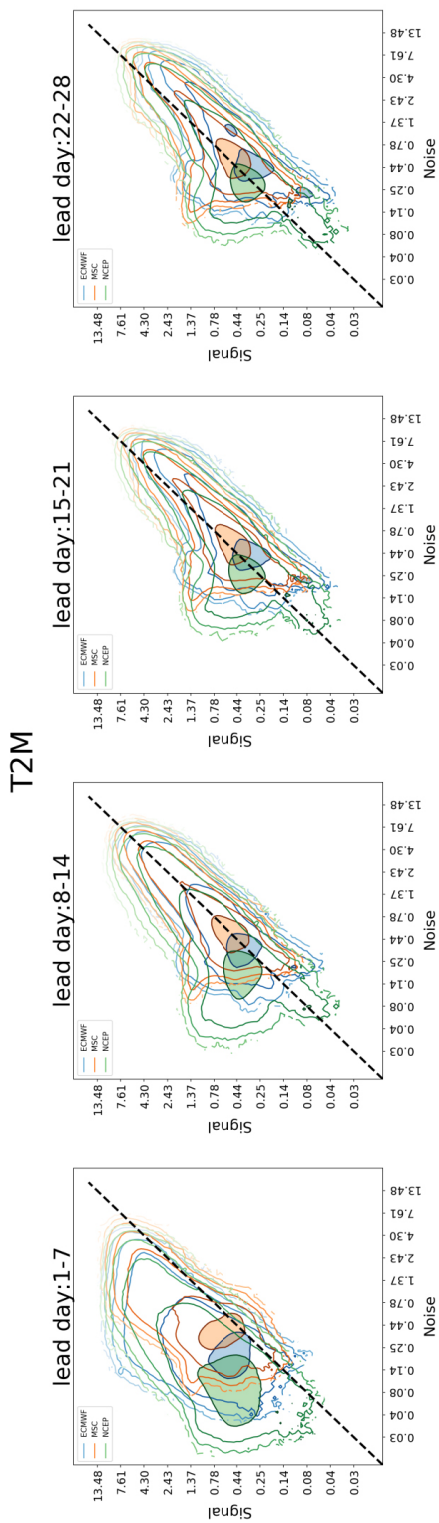


Figure 4. Global 2D histogram of signal and noise of T2M in four lead weeks for three models.

본 연구에서 검증한 세 모형자료는 각 기관의 상황에 따라 각각 다른 과거재현자료 기간과 앙상블로 이루어져 있다. ECMWF와 ECCC는 약 20년 정도의 자료 기간을 갖지만 NCEP은 12년에 불과하다. 또한 강한 97/98년의 엘니노를 포함하지 않아 시그널의 특성이 다른 두 모형과 다를 것이라는 합리적인 의심이 가능하다. 앙상블 사이즈의 경우 ECMWF는 11개여서 어느 정도 앙상블간 스프레드가 형성된다고 볼 수 있지만, 다른 두 모형은 앙상블 사이즈가 4개로 노이즈를 가늠하기에 적합하다고는 볼 수 없다. 하지만, 이러한 분석에 적합한 자료는 존재하지 않기 때문에 기존 자료를 최대한 활용하여 결과를 보이고, 파악된 한계들에 기반한 해석을 하는 것이 최선이라고 판단된다.

앞선 SNR관련 분석(Figure 3과 Figure 4)은 모형 고유의 기간과 앙상블 사이즈에 대한 결과이다. 따라서, NCEP의 SNR이 다른 두 모형에 비해 긴 것은 시그널 분석 기간이 짧고 앙상블 사이즈가 작은 것에 기인한다고 판단된다. 실제 NCEP 과거 재현 자료의 경우, 매일 6시간 간격의 초기조건에 대해서 4번 예측을 수행하고 그것을 앙상블인 것처럼 처리하지만 각각의 초기장에 대한 스프레드 실험은 수행하지 않는다. 이는 NCEP의 노이즈가 값이 전체적으로 작은 것의 주요한 원인으로 파악된다.

위의 SNR 분석에서 모형간 차이에 대한 합리적이고 유의미한 추론은 가능했지만, 모형 고유의 기간과 앙상블 사이즈에 기반한 분석이기에 모형간 SNR 비교는 공정하게 이루어졌다고 볼 수 없다. 0. 기존의 분석에 추가적으로 1. 과거 재현 기간을 1999-2010으로 통일하고 2. 앙상블 사이즈를 4개로 일치시킨 경우에 대한 분석을 수행하고, 세 분석 결과를 방사형 삼각형으로 시각화하였다(Figure 5). NCEP의 경우는 추가 분석에서 조건이 변화하지 않기에 항상 정삼각형 패턴이 나온다. ECCC의 경우는 기간 조건만 조정하기 때문에 항상 이등변 삼각형 패턴이 나온다. Figure 5가 보이듯이, 자료 기간이나 앙상블 사이즈를 통일시켜도 NCEP의 SNR 성능이 가장 길다는 결론은 달라지지 않았다(열대 T2M, Z200 제외). 하지만, 조건을 변화시켰을 때 다른 모형의 SNR 예측일이 어느 정도 예민하게 반응하는 것을 확인할 수 있었다. 분석 기간을 1. 1999-2010으로 통일하면 ECMWF와 ECCC의 예측일이 미세하게 감소하는 모습이 보이고, 이는 그림 5에서 제시된 변수와 지역 전체에서 일관되게 확인된다. 이는 앞서 언급했듯이 97/98 엘니노 해가 제외되면서 시그널(경년변동) 값이 줄어든 영향으로 분석된다. 기간에 더해 2. 앙상블 개수를 4개로 통일하는 실험은 ECMWF 11개의 앙상블 중에서 4개만을 추출하는 모든 경

우(330개 조합)에 대해 SNR를 계산하고 평균값을 취하는 방식으로 수행하였다. 그 결과 매우 뚜렷하게 그리고 일관되게 ECMWF SNR 예측일이 증가한 것을 확인할 수 있다. 이는 노이즈 감소를 의미하고, 샘플 사이즈가 증가(감소)할수록 분산값이 증가(감소)하는 전형적인 통계 특성에 부합하는 결과라고 볼 수 있다. 모형간 비교에서는, 전반적으로 ECCC, ECMWF, NCEP 순서로 높아지는 SNR 예측성이 확인된다(Table 5). 전구에서 네 변수 평균 예측일은 ECCC, ECMWF, NCEP 순으로 각각 10, 14, 19일이다. 동아시아에서는 10, 14, 17일로 큰 차이가 없으나 NCEP의 예측일이 전구에서보다 짧다. 변수별로 살펴보면, 지역에 관계없이 전반적으로 강수의 예측일이 가장 짧고, 반대로 대류권 상부 순환장(Z200)의 예측일이 가장 길게 나타났다. SNR을 기반으로한 예측일 모형의 잠재 최대 성능을 추정하는 값이고 실제 예측성과는 차이를 유념해야 한다. 또한 매우 작은 앙상블 사이즈에 기반한 분석이기에 추후 자료 갱신/확장에 따라 결론이 달라질 가능성이 있음을 명시한다.

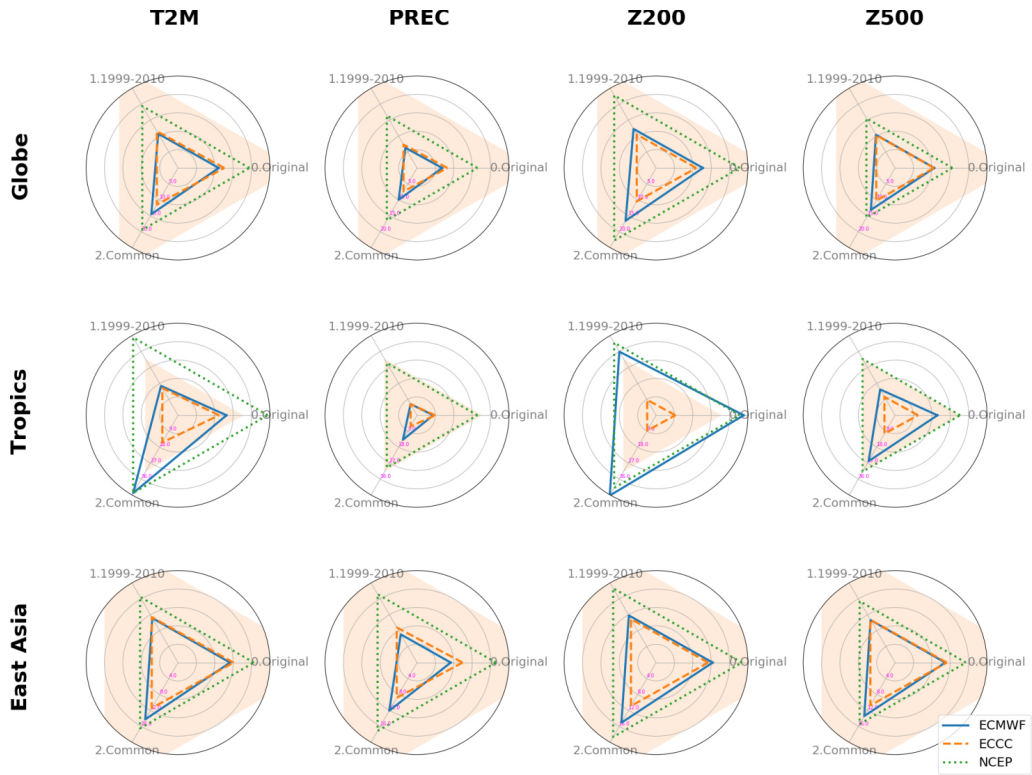


Figure 5. Predictable days derived from three different combinations of reference period and ensemble size for three models in three regions. Calculations are based on “0. Original”: original period and ensemble size, “1. 1999-2010”: common period and original ensemble size, “2. Common”: common period and ensemble size.

Table 5. SNR based predictable days derived from the test of “2. Common”: common period and ensemble size for four variables and three models over two regions.

	Model	T2M	PREC	Z200	Z500	Mean
Globe	ECCC	12	7	11	10	10
	ECMWF	15	10	17	13	14
	NCEP	19	16	23	16	19
E. Asia	ECCC	11	9	11	11	10
	ECMWF	14	12	15	13	14
	NCEP	17	17	19	15	17

3.2 다중모형 앙상블 기법에 따른 계절내 예측성

3.2.1 과거재현자료 성능 평가

a. 일별 예측성

S2S 참여 모형들이 생산하는 2016년 및 2017년의 과거재현자료의 전반적인 일별 예측성을 평가하기 위해 ACC를 통해 예측가능일을 살펴보았다. 예측선행시간이 증가할수록 ACC는 지속적으로 감소하는데, 여기서 예측가능일은 ACC 값이 특정 임계값보다 작아지는 날짜로 정의하였다. 특정 임계값은 임의로 0.3, 0.4, 0.5로 두었고 개별 모형의 모든 변수의 동아시아, 열대, 전구 지역에 대한 예측가능일을 Figure 6에 나타내었다. 임계값이 0.3일 때 대부분 약 2~3주의 예측가능일을 가지고, 0.4 일 때는 대략 1~2주일, 0.5 일 때는 약 1주일 정도의 예측가능일을 보이며, 전반적으로 임계값이 커질 수록 예측가능일이 감소하는 경향이 있다. 모형별로 예측선행시간이 다르므로(Table 1) 모형간 예측가능일을 직접적으로 비교할 수는 없으나, 이를 고려하더라도 대부분의 변수에서 임계값과 상관없이 ECMWF가 열대지역에서 가장 긴 예측가능일을 갖는다. 임계값이 0.4와 0.5일 경우의 PREC 예측가능일은 3개 모형과 3개 도메인에 대해서 0이고, 0.3의 경우 ECMWF, ECCC, NCEP이 열대지역에서의 예측가능일이 각각 6, 3, 1일로 다른 임계값에 비해 증가하나 여전히 일주일이 넘지 않는다. 반면 Z200의 경우에는 임계값이 증가함에도 불구하고 다른 변수에 비해 예측가능일이 크게 줄어들지 않는 모습을 보인다. 특히 ECMWF의 열대지역에 대한 예측가능일은 임계값이 0.5일 경우에도 33일로 유지되고, NCEP은 열대지역에서 21일의 예측가능일을 보여준다.

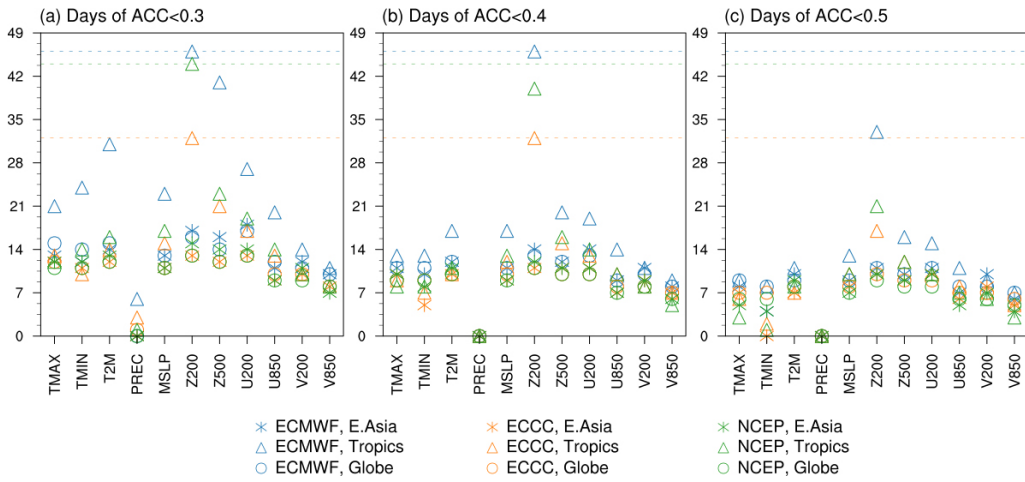


Figure 6. Distribution of predictable days defined as the day when ACC exceeds 0.3, 0.4, or 0.5 over 3 regional domains for three models for 11 variables. Blue, orange and green dashed lines indicate the forecast time range for ECMWF, ECCC and NCEP, respectively.

b. 주별 예측성

본 절에서는 S2S 과거재현자료의 주별 예측성을 살펴보고자 한다. 세 모형의 예측 시작일은 목요일이지만 선행시간 5일인 월요일을 시작으로 7일씩 나눠서 최대 4주까지 주 평균값을 구하고 이를 검증하였다. 여기서는 3개 모형과 더불어 Table 2에서 제시한 4가지 MME기법에 대한 예측성을 함께 비교하였다. Figure 7은 4개 변수에 대해 총 7가지 예측자료의 전구 주별 예측성을 누적하여 나타낸다. 대부분의 예측자료는 변수별로 정도의 차이가 있으나 누적된 예측성능 중 1주 선행시간의 예측성이 가장 큰 부분을 차지하고, 선행시간이 증가할수록 해당 선행시간이 차지하는 예측성의 비율이 줄어든다. 개별 모형 중 ECMWF는 다른 두 모형에 비해 상대적으로 높은 예측성을 4주까지 유지하고 특히 PREC에 대한 예측성은 다른 두 모형에 비해 확연히 높다. 4개 MME 기법의 예측성은 ECMWF와 거의 비슷하거나 높는데, 이 중 SCM_g가 가장 높은 예측성을 갖는다. SCM_g는 앙상블 수에 비례하여 가중치를 주는 기법으로, 개별 모형 중 가장 높은 예측성능과 가장 많은 앙상블 수를 가지는 ECMWF의 영향이 크게 작용한 것으로 판단된다. 또한 앞서 a 절에서 언급하였듯이, PREC의 누적된 예측성능은 다른 변수에 비해 확연히 떨어지고 특히 1주 선행시간에서 다른 변수는 0.5 이상의 예측성을 보이는데 반해 강수

는 1주에서도 0.5에 미치지 못하는 예측성을 보인다. Z200는 4가지 변수 중 가장 높은 누적 예측성능을 보인다.

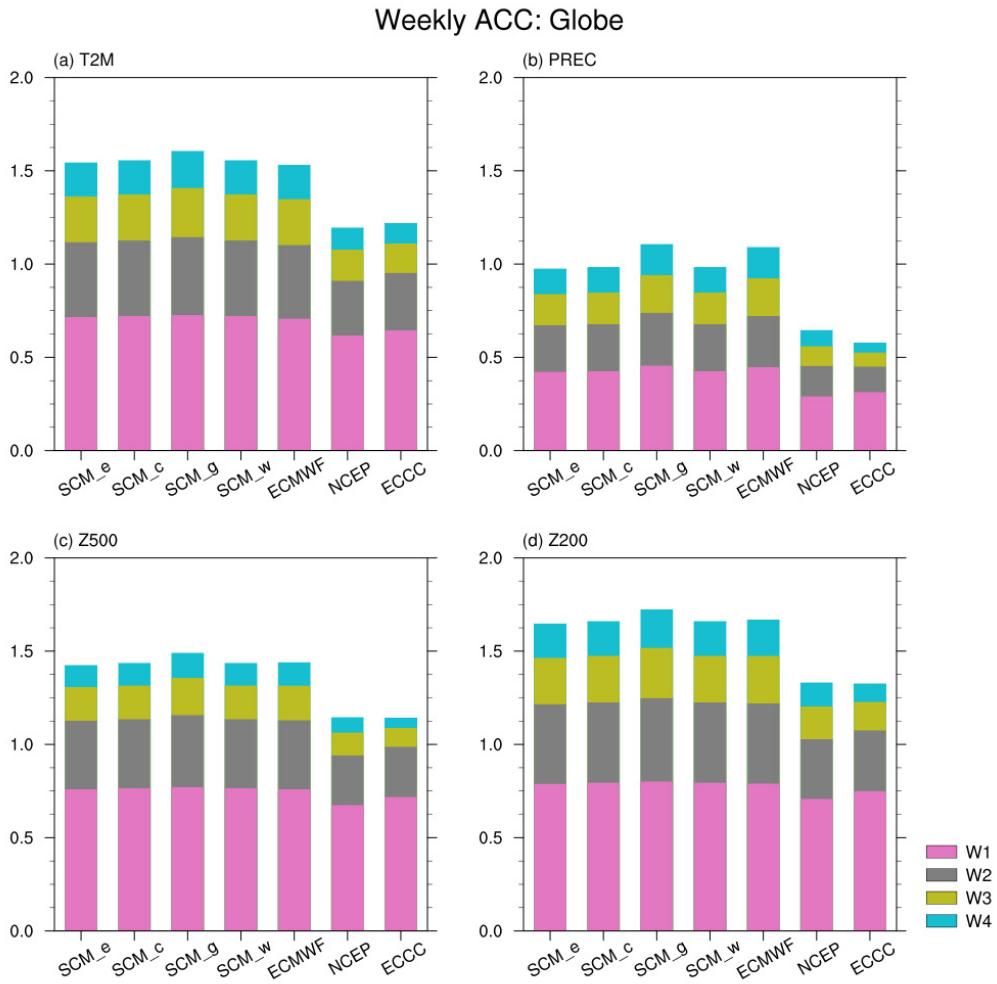


Figure 7. ACC of 4 variables over Globe for 7 simulations (4 SCMs and 3 models) for 4 lead weeks.

7가지 예측자료의 변수별, 지역별 예측성을 주별로 나누어서 Figure 8에 제시하였다. 전반적으로 예측자료들의 변수간 예측스킬의 차이는 있지만 앞서 언급한 바와 같이 1주에서 가장 높고, 같은 변수 내에서는 도메인에 상관없이 서로 비슷한 예측성능을 보이지만 예측 선행시간이 증가할수록 도메인간 예측성능에 차이가 있다. 열대지역의 예측성은

다른 도메인에 비해 변수에 상관없이 4주간 지속되고, 남태평양을 포함하는 호주 지역 또한 열대지역과 유사한 경향을 가진다. 남아시아 지역의 Z500과 Z200 예측성 역시 남태평양을 포함하는 호주 지역과 비슷한 모습을 보인다. PREC 예측성은 역시나 다른 변수에 비해 낮고, 특히 중동지역은 다른 지역에 비해 급격하게 예측성능이 줄어들어 3주부터는 예측성능이 0에 가까워진다. 앞서 언급한 바와 같이, 개별 모형의 관점에서 보면 NCEP과 ECCC의 예측성은 ECMWF보다 모든 도메인, 변수, 선행시간에서 낮고, 4가지 MME 기법의 예측성은 ECMWF를 따라가는 경향이 있다.

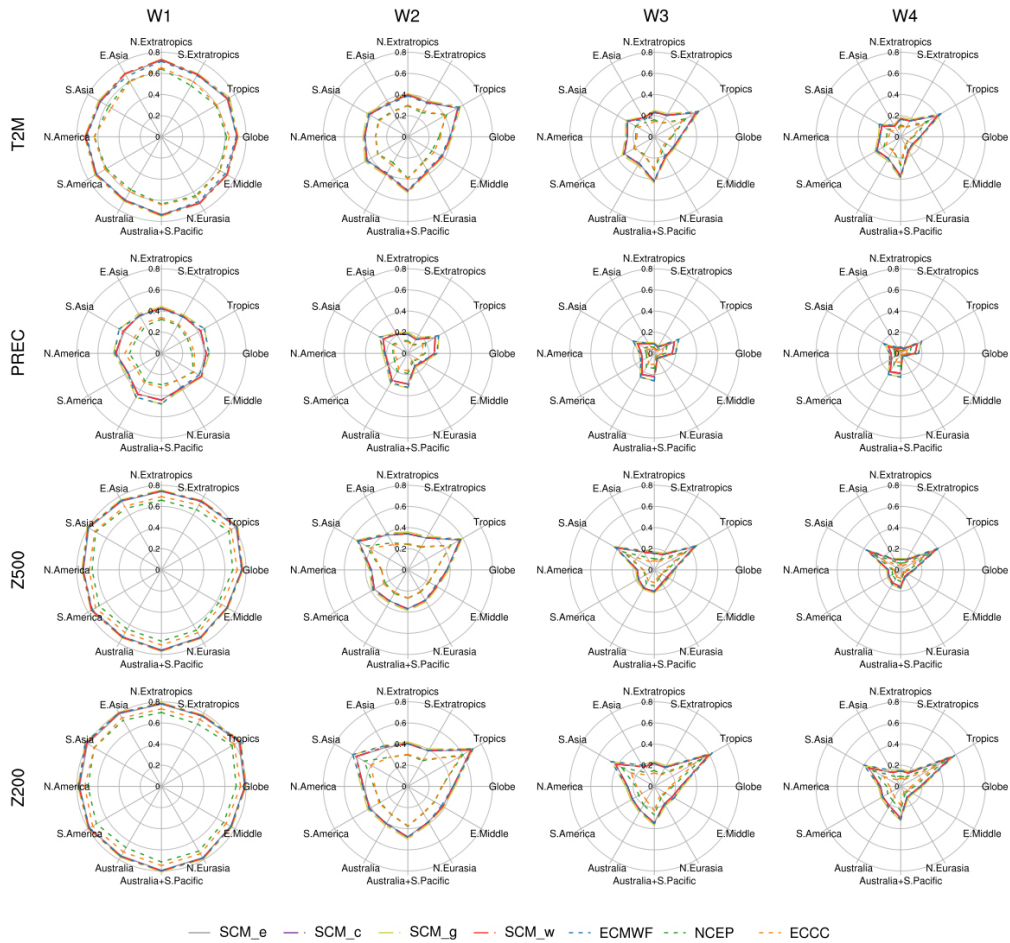


Figure 8. Regional variation of ACC of 4 variables for 7 simulations (4 SCMs and 3 models) for 4 lead weeks.

여러 선행연구에서 계절예측성능의 대부분은 경계조건의 영향을 받고, 특히 El Nino-Southern Oscillation (ENSO)의 활동과 큰 관련이 있다고 제시한 바 있다. 계절 예측의 예측성능은 주로 ENSO 사이클의 계절변동과 연관되어 여름철보다 겨울철에 높은 경향이 있다고 알려져 있다(Wang et al., 2009; Barnston et al., 2010; Min et al., 2014). 이러한 특징이 계절내 예측성능에도 영향을 미치는지 살펴보기 위해 3개의 개별 모형과 SCM_c의 기온에 대해서 ACC와 겨울철 ONI 크기(절대값)의 시계열과 이들 간의 상관계수를 Figure 9에 함께 나타내었다. 여기서 ACC는 겨울에 초기화 된 계절내 예측자료 중 2주 선행시간에 대한 예측성능의 평균값을 연도별로 나타낸 것이다. 개별 모형과 SCM_c의 기온에 대한 ACC 모두 ONI의 크기와 0.73 이상의 높은 상관계수를 가지고 이들간의 차이는 크지 않다. 이는 열대 지역이 ENSO 활동에 직접적으로 영향을 받는 지역이기 때문으로 생각된다. 동아시아 기온에 대해서는 3개 중 2개 모형의 예측성능이 ENSO와의 상관관계가 거의 없음에도 불구하고(ECMWF는 0.51), SCM_c의 동아시아 기온 예측성능은 ENSO와 0.31의 상관계수를 보인다. Figure 10에는 겨울철 강수의 3주 선행시간에 대해서 제시하였는데, 4개 예측자료의 열대지역 강수 예측성능 모두 0.69 이상으로 ENSO와 높은 상관을 보인다. 전구에서도 역시 0.65 이상의 상관계수를 가진다. 한 가지 주목할만한 점은 다른 변수에 비해 강수의 예측성능 자체는 높지 않으나, 강수 예측성능의 경년변동은 ENSO와 크게 관련되어 있다는 것이다. 이를 통해 ENSO는 계절 예측뿐만 아니라 계절내 예측성능에 영향을 주는 주요 소스 중 하나로 판단할 수 있다.

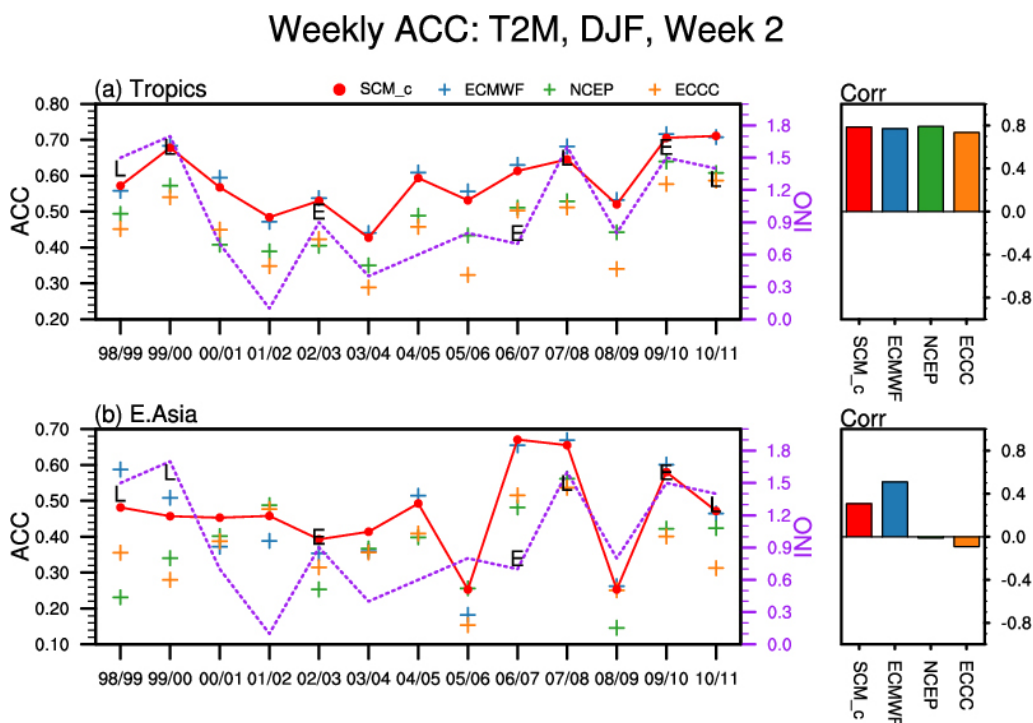


Figure 9. (Left) Timeseries of ACC of T2M for the period of DJF 1998/99–2010/11 for 4 simulations (SCM_c and 3 models) over (a) Tropics and (b) East Asia domains for 2-lead week and ONI index for same period, and (Right) Correlation between ACC timeseries for 4 simulations and ONI index. E and L marks denote El Nino and La Nina year, respectively.

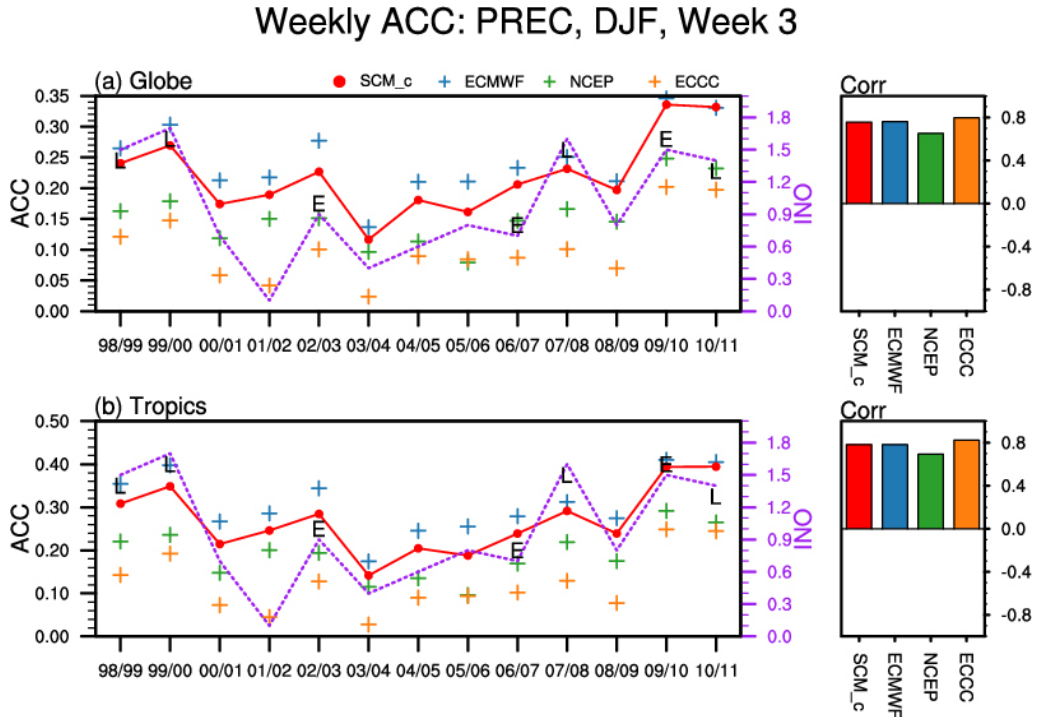


Figure 10. (Left) Timeseries of ACC of PREC for the period of DJF 1998/99–2010/11 for 4 simulations (SCM_c and 3 models) over (a) Globe and (b) Tropics domains for 3-lead week and ONI index for same period, and (Right) Correlation between ACC timeseries for 4 simulations and ONI index. E and L marks denote El Nino and La Niña year, respectively.

12 지역에 대한 4개 변수의 1~4주 선행시간의 겨울철 초기화된 SCM_c의 ACC와 ONI의 상관계수를 Figure 11에 요약하여 나타내었다. 여기서 선행시간은 위쪽 삼각형 으로부터 시작하여 반시계 방향으로 진행될수록 증가하며 위, 왼쪽, 아래, 오른쪽의 삼각 형이 각각 1, 2, 3, 4주 선행시간의 상관계수이다. 지역별로 살펴보면, 열대지역에서는 변수와 상관없이 예측성과 ONI와의 상관성이 가장 높게 나타나는데, 이는 앞서 언급한 바와 같이 열대 지역이 ENSO의 영향을 직접적으로 받는 지역이기 때문으로 생각된다. 또한 전체적으로 선행시간이 1주에서 4주로 증가할수록 상관계수가 커지는 경향이 있고, 이러한 특징은 한반도가 포함된 동아시아 지역에서도 잘 나타난다. 이를 통해, 계절내 예보 내에서 짧은 선행시간 예측은 중기예보의 특성이 강하고 긴 선행시간으로 갈수록 장기예보의 특성이 주요해지는 것으로 볼 수 있다. 모형 예측에서, 선행시간이 길어질수 록 대기 초기 조건에 대한 반응의 불확실성이 증가하고 따라서 예측성능이 감소하는 것은

자명하다. 예를 들어, 계절내 예측성의 원천이라고 알려진 MJO의 경우 메모리가 상대적으로 짧기 때문에 긴 선행 시간에서는 그 영향이 사라지게 된다. 초기 인자들의 영향이 사라진 대신, 메모리가 긴 ENSO의 영향은 여전히 살아 남아 그 상관성이 커지는 것처럼 보일 수도 있다. 하지만, 긴 선행 시간 예측성에서 상관관계 값이 매우 높기 때문에, ENSO 상태에 따른 예측 정보의 조건적 활용이 기대된다. 전구와 비교할 때 중동 지역은 선행시간에 따라 다른 양상을 보이는데, T2M의 경우 대부분의 지역에서 해양배경장과 예측성능의 관련성이 시간에 따라 커지는 것에 반해 중동 지역에서는 T2M 예측성능과의 관련성이 시간에 따라 감소하는 경향을 보인다.

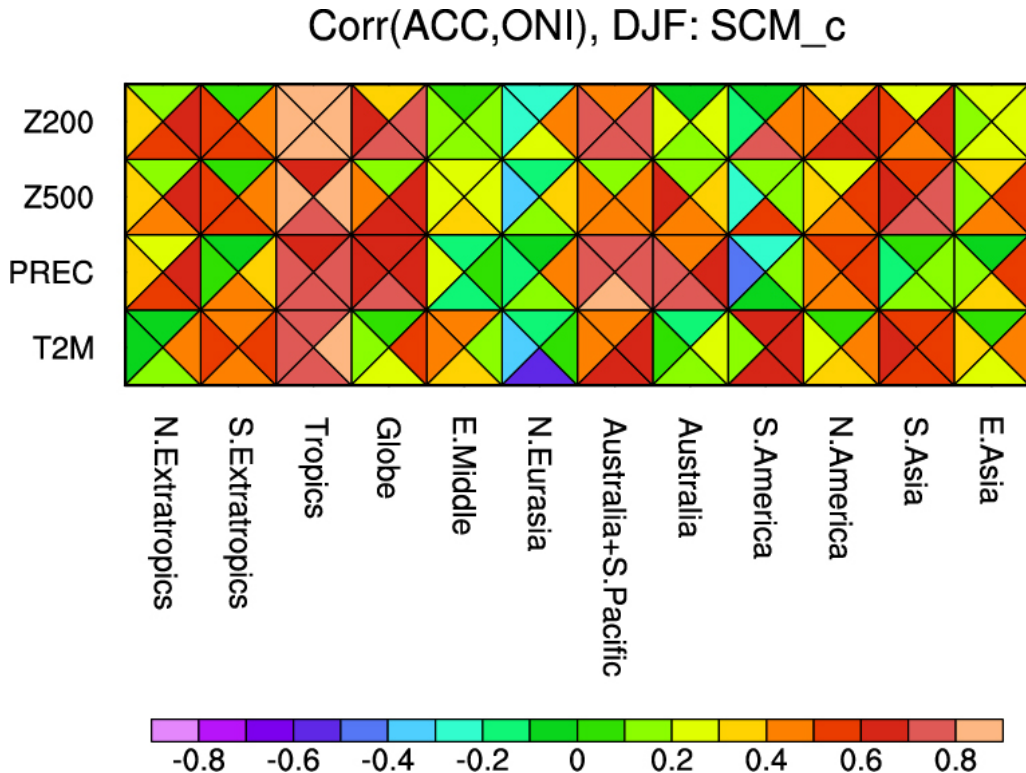


Figure 11. Correlation between ACC timeseries of SCM_c for 4 variables over 12 domains and ONI index for Correlation the period of DJF season for 4 lead weeks. The top, left, bottom and right triangles represent for 1-lead, 2-lead, 3-lead and 4-lead week, respectively.

전구와 중동지역에 대한 상관계수를 자세하게 살펴보고자 SCM_c와 함께 개별 모형들의 상관계수를 Figure 12에 제시하였다. 전구는 적도 지역을 포함하고 있기 때문에 Figure 11에서 보여준 적도의 상관성을 가지고, 선행시간이 증가함에 따라 전체적으로 ENSO와 ACC간의 상관성이 증가하는 경향을 보인다. 모든 변수에서 NCEP이 다른 모형이나 SCM_c에 비해 상관계수가 낮지만, 다른 모형과 비슷한 수준으로 판단되고 모형간의 편차도 크지 않다(Figure 12 (a)-(d)). 항상 뜨거운 기온이 유지되는 사막기후의 특성을 가진 중동 지역 T2M 예측성능의 경우, 1주 선행시간에서 초기조건에 크게 예민하지 않고 해양경계조건에 영향을 강하게 받는 것으로 보인다. 물리적 거리를 고려하면 중동지역과 태평양 경년변동 사이에 역학적 관계 때문에 강한 의존성을 보인다고 해석하는 것은 무리이다. 계절 예측 모형 예측성은 태평양 ENSO에 기인한다고 알려져 있고, 강한 시그널이 존재할 경우에는 중위도 대기 반응까지도 신뢰할 만한 수준의 예측성이 보인다. 계절내 예측에서도 강한 ENSO 시그널이 경계조건으로 존재할 때 깊은 대륙 중동지역의 예측성이 유의미하게 높아지는 것으로 해석된다. 선행시간이 증가할수록 경계조건의 영향이 줄어드는 것으로 보이는데, 이는 항상 덥고 공간적으로 단조로운 기온분포를 보이는 중동 지역에서는 긴 선행시간에서도 통계의 특성상 인위적으로 일정 정도의 ACC 값이 매해 유지되고 따라서 유의미한 경년변동 자체가 나타날 수 없기 때문으로 판단된다. 나머지 3개 변수들을 살펴보면, 중동 계절내 예측성의 ENSO 의존성은 전 선행시간에서 미비하고 모형들간의 일관성도 존재하지 않음을 확인할 수 있다. 도메인에 적도가 포함되지 않기 때문에 적도 보다는 적도와 극 사이의 온도차 같은 중위도 자체 역학이 예측성에 주는 영향이 클 것으로 생각된다.

예측성이 매우 높고 ENSO 주요 활동 무대인 적도와 그 외 영역에 대한 차이를 살펴보기 위해, 대표로 열대와 북반구 아적도(북위 20도 이북) 두 지역의 ACC 예측성능과 ONI 간의 상관계수를 Figure 13에 나타내었다. 전반적으로, 열대에서는 높은 상관성이 1주부터 4주까지 유지되어 선행시간에 상관없이 ENSO에 영향을 크게 받는 반면, 북반구 아적도에서는 1주에서는 ENSO보다는 MJO 등의 영향을 더 크게 받다가 선행시간이 증가할수록 다른 강제력이나 대기초기장 보다 ENSO의 영향이 커지는 것으로 보인다. 또한, T2M을 이외의 변수에서 3-4주차에서의 ENSO 영향은 두 지역에서 큰 차이를 확인할 수 없었다. 따라서, ENSO의 활동에 따라 3~4주의 아적도 지역 예측성능이 높아질 것으

로 기대할 수 있고, 이는 계절내 예측자료의 활용에 좋은 정보가 될 수 있을 것으로 생각된다. 서로 유사한 패턴을 보이는 T2M, Z500, Z200와 달리, PREC의 경우는 그 추이가 독특하다. 열대에서는 ENSO와의 상관성이 1주에 다소 낮은 값에서 시작하는데, 이는 MJO 직접 강수의 영향을 크게 받기 때문으로 풀이된다. 반대로, 북반구 아적도 1주에서는 상관성이 다른 변수들에 비해 다소 높다. 결론적으로, 열대와 북반구 아적도의 차이가 다른 변수들에 비해 크지 않다. 강수가 해수면 온도와 직접적으로 연관된 물리과정의 결과물을 것을 감안하면 이러한 관련성의 추이를 이해할 수 있으며, 추후 이와 관련된 추가 분석이 필요할 것으로 판단된다.

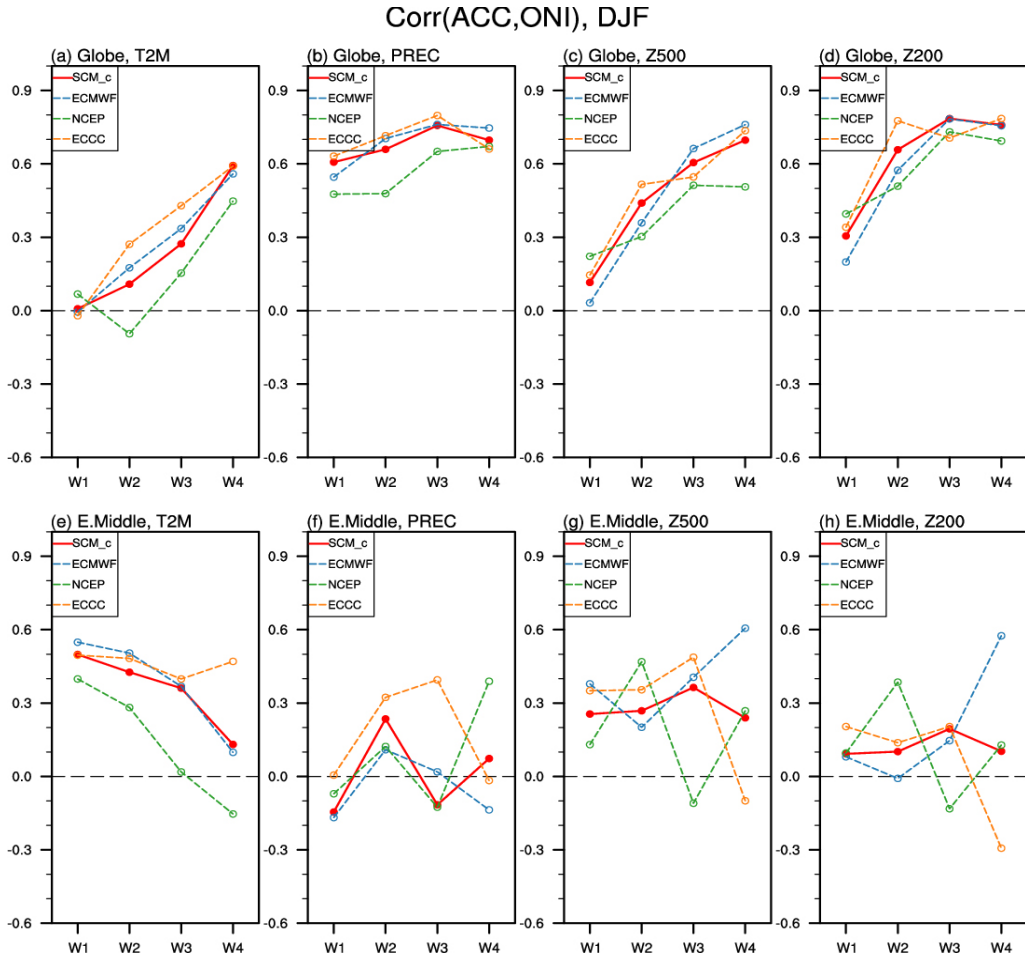


Figure 12. Temporal variation of correlation between ACC timeseries of SCM_c and 3 models for 4 variables over Globe and East correlation domains and ONI index for the period of DJF season.

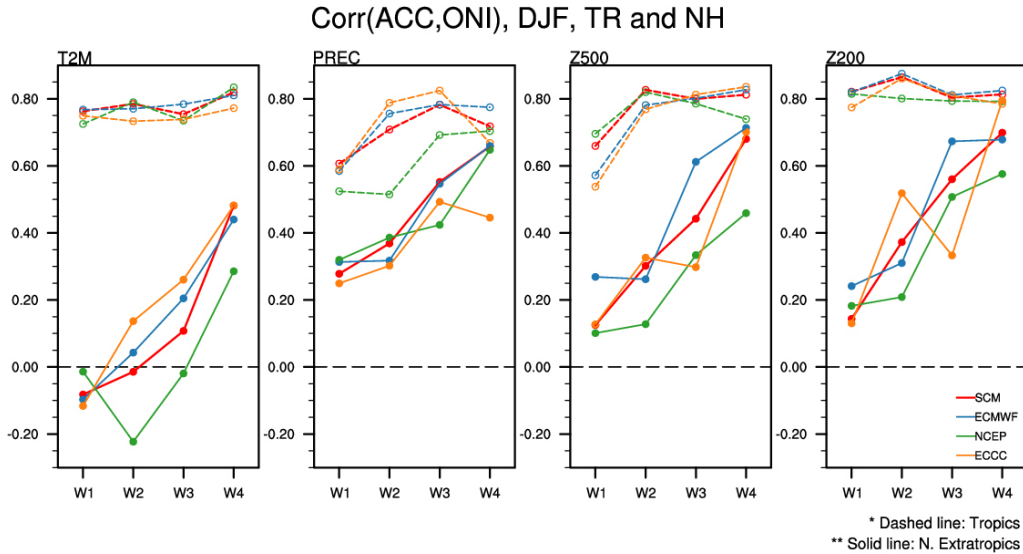


Figure 13. Temporal variation of correlation between ACC timeseries of SCM_c and 3 models for 4 variables over Tropics (dashed line) and N. Extratropics (solid line) domains and ONI index for the period of DJF season.

3.2.2 실시간예측자료 성능 평가

2016/17 두 해의 S2S 실시간 주평균 예측에 대한 3개의 개별 모형과 4개 MME의 전반적인 결정론적/확률론적 예보성능을 Figure 14에 요약 제시하였다. 결정론적 예보에 대한 ACC 검증 결과(상단 1열), ECCO와 NCEP을 제외한 ECMWF 와 4개 SCM은 유사한 수준의 예측성을 보이고 예측 선행시간이 증가할수록 성능이 급격히 감소하는 등, 앞선 과거재현자료의 성능평가와 일관된 모습이다. 변수별로는 강수를 제외하고 하층 기온, 상층 지위고도장의 예측성이 높게 나타났다. 이는 변수의 고유 특성에 기인한다. 강수는 지역적으로 나타나는 현상인데 반해 다른 세 변수는 넓은 공간해상도를 갖고 특히 상층으로 갈수록 점점 평탄해지는 특성이 있기 때문에 높은 성능을 나타낼 수 밖에 없다. 2016년의 성능이 2017년에 비해서 높은 것을 알 수 있는데, 앞서 언급한대로 강한 ENSO 시그널이 존재한 해(2016)에 계절내 예측 성능도 높은 것으로 해석할 수 있겠다.

LEPS/ROC/SR 로 검증한 확률론적 예보 성능도 위의 결정론적 예보 성능의 특징과 크게 다르지 않지만(Figure 14, 상단 2-4열), 일부에서 차이점도 존재한다. 예를 들어, 확률예보의 경우는 500 hPa 지위고도장이 하층기온에 비해 미세하게 성능이 높다.

LEPS 검증 결과에서는 ECMWF 와 ECCG가 뚜렷하게 최고성능 최저성능 모형으로 구분이 되는데, ROC/SR에서 아주 미세하게 MME_g가 최대 성능을 보이는 것과 대조적이다. 이는 검증지수 각각이 예보를 바라보는 시점이 다르기 때문이다. ROC/SR은 카테고리를 맞추는 정도를 성능의 기준으로 보지만 LEPS는 특정 카테고리를 얼마나 확신을 가지고 예측했는지를 평가한다. ECMWF가 가장 높은 확률로 정확한 카테고리를 예보하기 때문에 PMME에서 확률을 조합하는 것으로는 LEPS 값이 떨어질 수 밖에 없음을 보여준다. 또한 NCEP는 PMME에 필적할 LEPS 성능을 보이는데, 카테고리를 맞추는 능력은 ECCG와 크게 다르지 않지만 카테고리를 맞춘 예보에서는 상대적으로 매우 강한 확신(높은 확률)을 보였기 때문으로 해석 가능하다.

전반적인 모형별, 선행시간별, 변수별 그리고 연별 상대적 예측성을 확인하기 위해 최대 예측 스کیل에 대한 비율을 도식화 하였다(Figure 14, 하단 1-4열). ACC/ROC/SR에서의 예측 민감도는 일관된 모습을 보인다. MME_g가 미미하게나마 최고 성능을 보이고, 선행시간과 변수에 대해 성능이 줄어드는 비율도 유사하다. 하지만 LEPS에서는 ECMWF 모형 성능이 뚜렷하게 높고, 2주차 예보 성능 하락이 상대적으로 크며, 강수량의 성능이 다른 변수들에 비해 매우 낮은 것으로 나온다. 이러한 결과는 LEPS 검증 지수의 특징을 여실히 드러낸다. LEPS는 “카테고리를 높은 확률로 맞추는 것”으로 정확한 예측을 정의하기 때문에, 그 결과값의 분포가 넓고 예보간의 차이가 뚜렷한 매우 정직한 검증 지수라고 볼 수 있겠다. 하지만 여러 모형의 확률을 조합하는 MME에 적용하면, 그 결과는 가장 좋은 하나의 개별 모형보다 높아질 가능성은 수학적으로 크지 않다. 이는 MME 성능 평가의 목적과도 맞지 않기에, 적어도 계절내 MME 예보에는 LEPS 검증지수 적용은 적절하지 않은 것으로 판단된다. 추후, 충분히 많은 개별 모형을 합성하는 연구를 수행하여 LEPS 검증 지수의 특성 및 한계를 좀 더 객관적으로 평가해 보고자 한다.

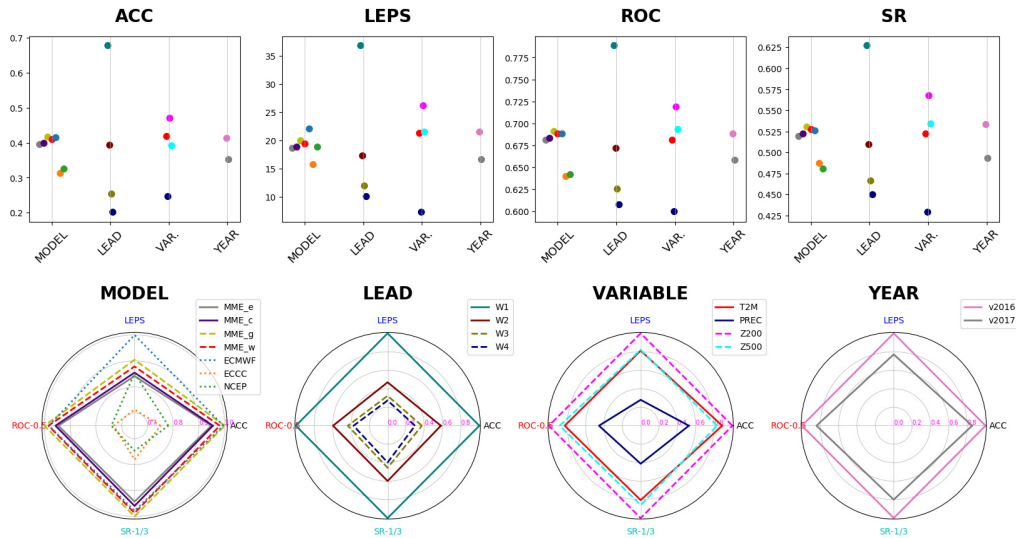


Figure 14. (Upper) Prediction skill dependent on different simulations, lead weeks, variables, and reforecast versions verified from four metrics: ACC, LEPS, ROC, and SR. (Lower) Fractional prediction skill relative to the top performer.

과거 재현 자료 검증에서처럼, 선행시간과 지역별 예측성능의 민감도를 분석하고 하층 기온에 대한 결과를 대표로 제시하였다(Figure 15). 앞선 결과에서와 마찬가지로 ACC는 결정론적 예보를, LEPS/ROC/SR는 확률론적 예보를 평가한 결과이다. 예보 첫 주에는 크고 원에 가까운 형태를 보이는데, 이는 예측성능이 높고 지역적으로 차이가 없음을 의미한다. 하지만 선행 시간이 증가할수록 원이 찌그러진 형태로 줄어들어서 성능저하가 지역별로 다르게 나타나는 것을 확인할 수 있다. 열대지역과 호주를 포함한 남태평양은 긴 선행시간에도 모든 검증지수에서 다소 높은 예측성능을 유지한다. 북미대륙과 남아시아의 경우는 전반적으로 성능이 유지되는 것으로 보이지만 지수별로 차이가 있다. 많은 면적이 대륙으로 구성된 동아시아, 중동, 북유라시아는 모든 지수에서 일관되게 성능 저하가 뚜렷한데, 중위도 대륙에서 예측성이 낮은 기후 모형의 특성이 계절내 규모에서도 일관됨을 확인할 수 있다. 모형별로 비교해보면, ECMWF와 4개 MME의 선들이 뭉쳐서 나타나 그 성능이 유사함을 알 수 있다. 이들의 성능이 다른 두 개별 모형에 비해 뚜렷하게 높는데, 이는 지역에 관계없이 일관되고 선행시간이 길어져도 그 상대적인 차이는 유지되는 것으로 보인다. 하지만 LEPS 검증에서는 ECMWF의 성능이 지역에 크게 상관없이 뚜렷이 높고, 이는 앞선 그림의 결과와 일관된다.

전반적으로 가장 높은 성능을 보인 MME_g에 대해서 4개 변수의 예측성능 추이를 선행시간과 지역별로 나타낸 결과는 Figure 16과 같다. 상층 지위고도의 남아시아 성능은 하층 기온에 비해서 매우 높고, 열대지역 성능에도 필적한다. 최신 S2S 예측 연구 결과들은 MJO 예측력이 4주까지도 유지된다고 밝히고 있다(Vitart et al., 2010; Reshied et al., 2011; Zhang et al., 2013; Kim et al., 2014). 남아시아 지역은 열대 인도양과 태평양을 순환하는 MJO와 그 원격상관에 밀접하게 반응하는 지역이기 때문에, 상층 지위고도장의 높은 예측성을 MJO 성과와 연결 지을 수 있겠다. 강수는 계절내 규모에서 매우 중요한 변수이지만 안타깝게도 3-4주 예측에서는 모형이나 지역에 관계없이 예측성이 형편없다. 하지만 중동지역에서는 확률예보 성능이 매우 높은 것으로 의외의 결과를 보이는데, 일반적으로 중위도 내륙의 예보성능이 낮은 것과 반대이다. 이는 중동지역이 일년 내내 매우 건조하고 강수가 드문 사막 기후이기 때문에 나타나는 인위적 값으로 판단된다. 매우 적은 유효 강수 샘플로부터 감마 분포에 의해서 얻어진 삼분위값은 0보다 크기 때문에 강우가 없는 대부분의 경우에 BN를 예보하게 되고 따라서 확률/카테고리 예보의 성능은 높아질 수 밖에 없다. 이를 통해, 일부 지역이나 변수에 대해서는 검증 지수를 숫자 그대로 해석하는 것을 경계해야 함을 배울 수 있다. 변수간의 성능을 비교하면, 전반적으로는 강수, 하층 기온, 500 hPa, 200 hPa 지위고도 순으로 높다. 한반도를 포함하는 동아시아 지역이 그러한데, 상층 지위고도의 예측성능이 상대적으로 긴 선행시간에서 유지되기 때문에 통계적 후처리를 통한 맞춤형 서비스를 계획한다면 상층변수를 통해 예측성능 향상의 기회를 엿볼 수 있을 것이다. 하지만, 일부 지역/선행시간에서는 변수별 예측력의 순위가 바뀌기도 한다. 북미 지역이 대표적인데, 2주차부터 하층 기온의 성능이 상층 지위고도 성능을 넘어서는 것으로 파악된다.

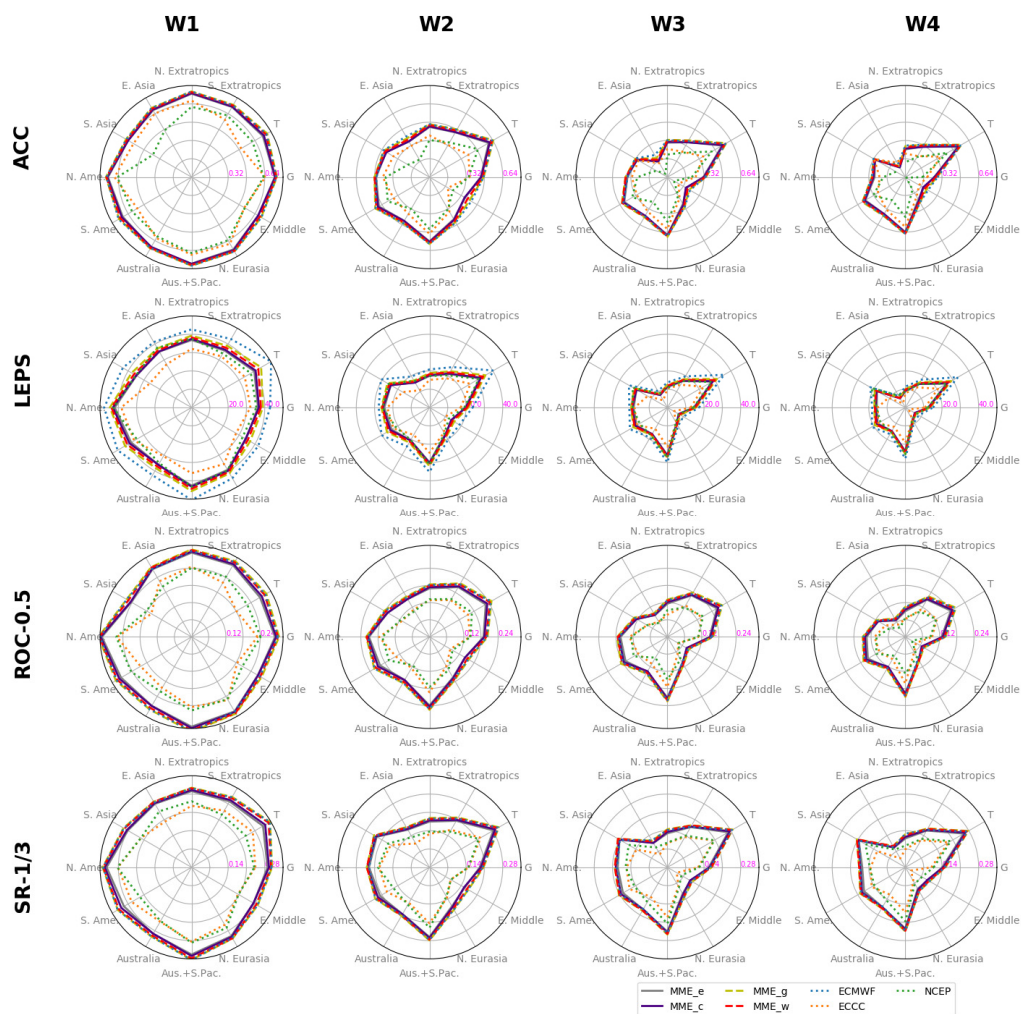


Figure 15. Regional variation of prediction skill of T2M for 7 simulations (4 MMEs and 3 models) verified from 4 metrics for 4 lead weeks.

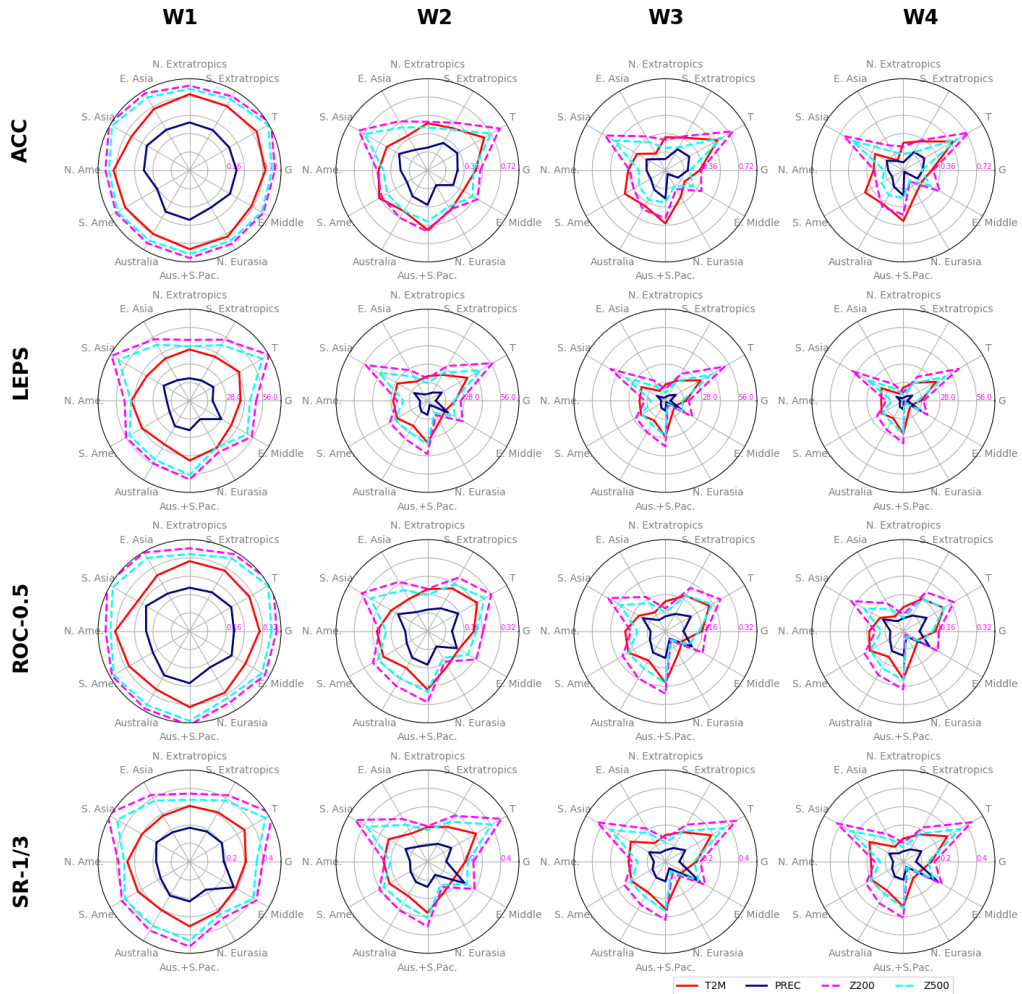


Figure 16. Regional variation of prediction skill of MME_g for 4 variables (T2M, PREC, Z200 and Z500) verified from 4 metrics for 4 lead weeks.

앞선 분석은 각 검증지수의 평균값을 비교함으로써 계절내 예보에 다중모형 합성법이 효과가 있는지, 가장 적절한 합성방법은 무엇인지에 대한 답을 제시한다. 모형 개수가 3개로 적고 다른 모형 대비 성능이 월등한 ECMWF 모형을 합성했다는 점에서 분석 결과를 일반화하는데 한계가 있지만, 모든 양상분을 같은 비중으로 합성하는 MME_g의 방법에서 ECMWF 대비 성능 향상을 확인할 수 있었다. 하지만 지역별 변수별로 차이가 있고, 절대적인 향상폭은 크지 않았다. MME의 효과를 확인하기 위해 추가적으로 7개 예측(3개 모형, 4개 MME)의 검증지수 값을 비교하여 각각의 예측이 최고 성능을 보인 횟수를 선행

주수, 변수, 검증 지수, 그리고 지역에 대하여 막대그래프로 나타냈다(Figure 17). MME_g나 ECMWF가 최고 순위 빈도가 가장 높고, 둘 중에서 MME_g의 비율이 좀 더 높다. NCEP은 일부 최고 성능을 보이는 경우가 있지만 ECCC는 거의 없다. MME 방법 중에서 세 모형을 동일 가중치를 주어 합성하는 MME는 기후값 기준 기간에 관계없이 최고 성능을 보인 횟수가 매우 드물다. 이는 ECMWF를 상대적으로 낮은 성능의 두 모형과 같은 비중으로 합성했기 때문에 나타난 성능 저하이다. 같은 맥락으로, MME_w에 대비 MME_g의 빈도가 훨씬 높은 것은 ECMWF에 좀 더 많은 가중치가 부여됐기 때문으로 해석할 수 있다.

좀 더 세분화해서 살펴보면, 예측 주수가 길어질수록(특히 3주) MME_g와 ECMWF의 빈도차가 증가하는 경향이 있다. 따라서, 선행 시간이 길수록 다중모형 합성법의 효과가 높다고 판단된다. 변수별로 살펴보면, MME_g와 ECMWF 빈도차가 하층 기온에서 가장 크고, 따라서 MME 효과가 가장 큰 것으로 해석된다. 반대로 200 hPa 지위고도에서는 ECMWF가 MME_g에 비해 최고 순위 빈도가 높다. 검증 지수별로 살펴보면, 결정론적 예보를 검증(ACC)에서는 MME_g와 ECMWF의 빈도차는 크지 않은데 반해 확률론적 예보 검증(ROC/SR)에서는 상대적으로 그 차이가 매우 크다. 하지만 LEPS에서는 ECMWF 빈도가 매우 높고 대신 모든 MME들의 빈도는 거의 0이다. 이는, 앞서 언급했듯이 검증 지수의 고유 특성 때문으로 LEPS가 MME의 효과를 계량하는 지수로서는 부적합하다는 점을 다시 한번 확인할 수 있다. 검증 지역별로 살펴보면, MME_g와 ECMWF의 빈도차가 거의 없는 남반구 아적도에 비해 북반구 아적도에서는 그 차이가 매우 커서 MME 효과를 기대해 볼 수 있겠다. 반대로 남아시아와 중동지역에서는 ECMWF 빈도가 가장 높고, 후반 선행시점(3-4주)에서는 동아시아 지역도 마찬가지로의 결과를 보인다(그림 생략). 따라서, 일부 지역에서는 MME 효과가 미비한 것으로 판단된다.

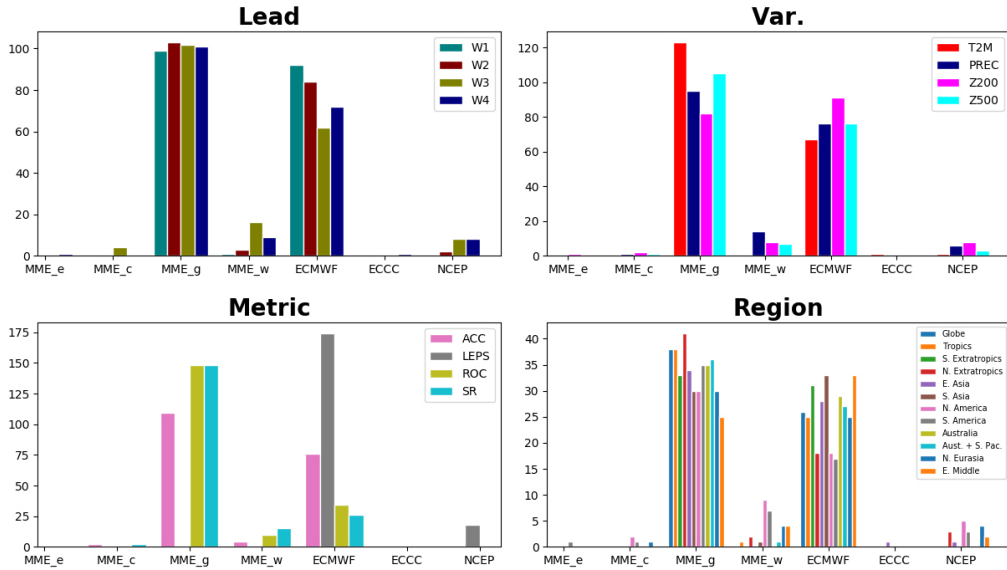


Figure 17. Frequency of 7 simulations counted as a top ranker in terms of lead weeks, variable, metrics, and regions.

앞선 분석들은 각기 다른 모형들을 통합하여 예측성을 높이려는 MME 접근법이 계절 내 예측에서도 변수, 시기, 지역에 따라 일정 정도 효과가 있음을 보여 매우 의미 있는 결과를 도출하였다. 하지만 모형 개수가 3개로 제한되었다는 것과 그 중 뚜렷한 최고 성능모형인 ECMWF를 포함했다는 점이 한계로 지적될 수 있다. 특히, MME 적용 후에 예측성능에 있어 ECMWF 대비 괄목할만한 개선을 보이지 않았다는 점은 MME 적용의 효율에 대해 시사하는 바가 적지 않다. ECMWF의 성능이 높은 것은 초기 2주까지 매우 높은 대기 해상도(16 km)로 모의를 수행하고, 앙상블의 개수가 상대적으로 많으며, 다양한 형태의 앙상블 수행과 자료동화 기법이 적용되기 때문으로 알려져 있다. ECMWF의 성능을 크게 뛰어넘기 위해서는 MME 과정에서 불확실성의 감소가 두드러지게 나타나야 하는데, NCEP과 ECCC가 섞이는 과정에서는 안타깝게도 그 효과가 미비한 것으로 볼 수 있다. 이는 앞선 SNR 분석에서 NCEP의 SNR가 세 모형 중에 가장 높게 나타났기에 매우 예상과 다른 결과이다. SNR은 개별 앙상블 모형의 가능성을 가늠하는 중요한 잣대 중 하나이지만, 예측성능과 항상 비례하는 것은 아니다. 이는 앙상블 간의 스프레드가 모델이 가리키는 시그널을 확신할 수 없을 만큼 넓지 않다는 개념(precision)이며, 실제 정확도(accuracy)와는 구분해서 생각해야 한다. 만약 특정 모형 시그널이 지속적인 bias

를 가지고 있다면 높은 SNR 값에도 불구하고 예측성능은 낮고, 따라서 다른 모형과 섞이는 과정에서 오히려 불확실성이 증가해 그 효과가 미비할 수도 있다. 향후 MME 성능 개선을 위해서는 적절한 오차 보정법이 수반되어야 할 것으로 사료된다. 아울러, 다양한 계절내 모형들을 추가하여 모형간 분석을 통해 SNR과 예측성능과의 관계를 좀 더 일반화하고 관련 요인을 파악할 수 있을 것으로 생각된다.

4. 계절내 규모 기후현상의 재현 특성 및 예측성

4.1 MJO

4.1.1 MJO의 특징

적도 계절안진동의 가장 뚜렷한 모드로 알려진 MJO는 Wheeler and Hendon (2004) 이 개발한 real-time multivariate (RMM; OLR, 850 and 200 hPa zonal winds) index를 통해 현업기관에서 연중 예측정보를 제공하는 주요 기후인자 중의 하나이다. 4.1장에서는 이 현상이 S2S 모형에서 어떻게 모의되며 어느 정도의 예측성이 있는지를 가늠해보고자 서론에서 언급된 바와 같이 공통기간 및 조건에 해당하는 3개의 모형과 이들의 단순평균을 통한 MME까지 총 4가지 경우를 관측과 비교하며 분석하였다.

4.1.2 MJO 의 재현 특성 평가

각 모형에서 MJO를 어느 정도 적절하게 모의하는지 평가해 보기 위해 모형별 그리고 MME를 선행시간 별로 나누어 관측과 비교하였다. Figure 18, Figure 19, Figure 20은 MJO 전파모습을 선행시간 1주 시, 전체기간, 여름(MJJASO), 겨울(NDJFMA)에 대해 각각 제시한 것이다. 관측 (a) 에서 나타나는 모습은 이미 이전의 수많은 연구에서 제시된 바 있다. 주시해야 할 점은 총 기간, 여름, 겨울로 나누어 모형별 MJO의 특징이 어떻게 재현되고 계절 변화의 특성도 모의하는 지에 대한 것이다. 먼저 관측에서 적도 인도양에서부터 따뜻한 해양을 따라 대류구역(음의 OLR)이 동진하며 중태평양에 이르러 점차 소멸하는 모습은 모든 기간, 모형에서 관측과 일치하게 모의되고 있다. 3개의 모형을 단순 평균한 MME의 경우 관측과 아주 유사한 전파속도와 대류구역의 위치가 뚜렷하며 강도 면에서는 관측보다 약하게 모의되고 있는데 이는 아노말리장의 합성을 통한 상쇄작용에 의한 결과인 것으로 보인다. 각 모형별 세부적인 특징을 보게되면 ECCO의 경우 가장 약한 강도와 서태평양에서의 대류활동의 약화로 인한 동진 축소를 지적할 수 있고 ECMWF와 NCEP의 경우 대류활동의 강도가 관측과 유사한 크기로 나타남과 동시에 대

류역의 강한 동진모습이 보이며 두 모형 모두 우수한 모의를 하고 있는 것으로 판단된다. 이들의 활동을 여름과 겨울철로 6개월씩 구분하여 보게되면 MJJASO 여름철에는 동진뿐만 아니라 북동진하는 구조가 발생하며 여름철 북반구 몬순의 활동과도 관련이 되는데 관측과 MME를 비교하였을 때 이러한 복합적인 구조의 전파형태가 잘 모의되고 있다. 개별 모형 특징은 ECCC가 MJO 관련 대류활동이 약하게 모의되긴 하지만 전체적으로 MJO 위상별 대류구역의 위치, 전파 특성 등은 모든 모형에서 잘 모의하고 있다. 겨울철은 MJO가 가장 활발한 시기로 그림에서 보듯이 강한 강도의 대류활동이 동진하며 이러한 구조가 MME를 비롯하여 개별 모형에서 잘 나타나고 있다. ECCC는 전체 위상에 대해 다소 MJO 대류활동이 약하게 모의되고 ECMWF와 NCEP은 계절에 상관없이 대류구역의 위도별 위치 및 위상속도 등이 관측과 유사하게 표현되고 있음을 볼 수 있다. 다만 여기서 제시한 것은 가장 초기의 1주 선행시간에서의 특징이기 때문에 예측성의 향상과는 연결지어 생각하기 어렵지만 MJO 자체의 특징을 우수하게 모의하는 것도 매우 중요하다.

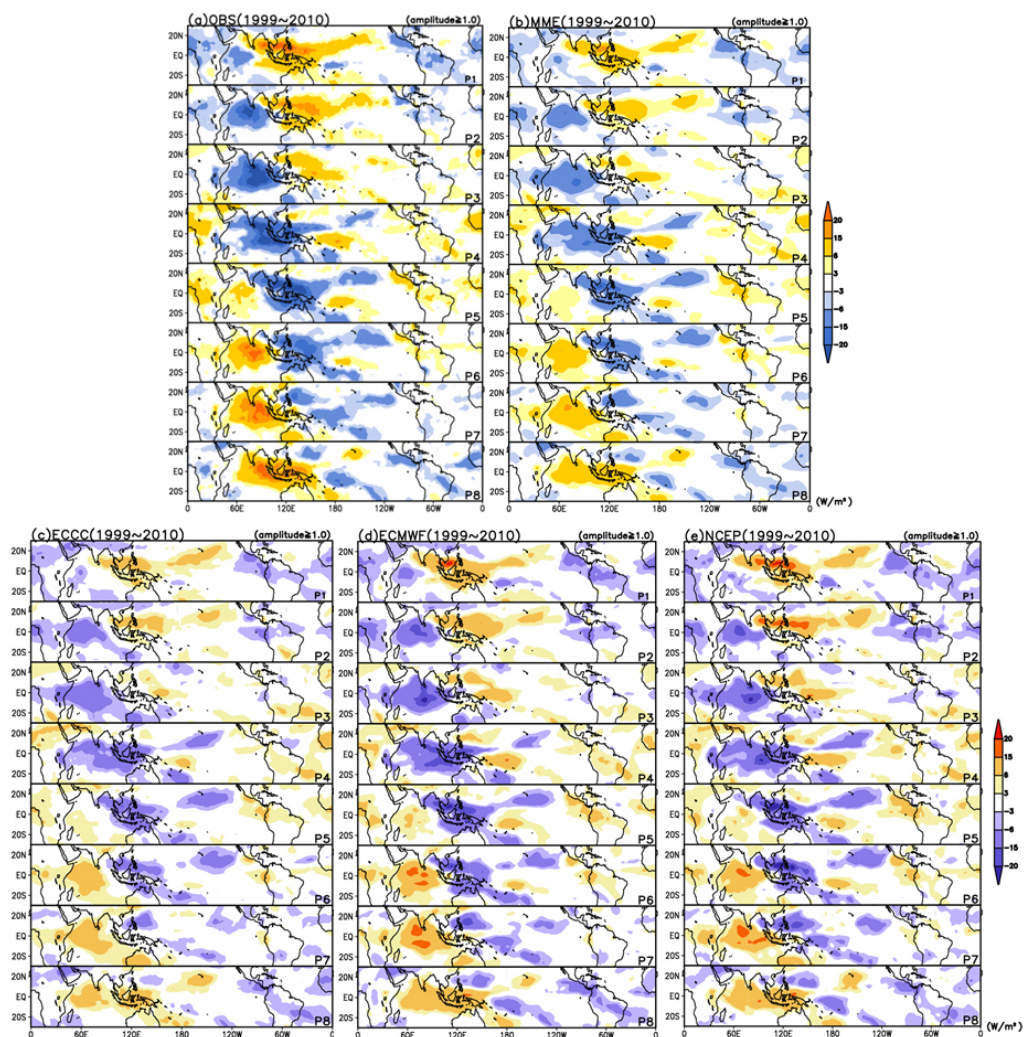


Figure 18. The composite maps of intraseasonal anomalies of OLR (negative anomaly depicts active convection, unit: W/m^2) from the (a) observation (b) MME (c) ECCC (d) ECMWF (e) NCEP reforecast during 1999~2010 at lead week1, with reference to WH (2004) when the amplitude of MJO index is greater than 1.0.

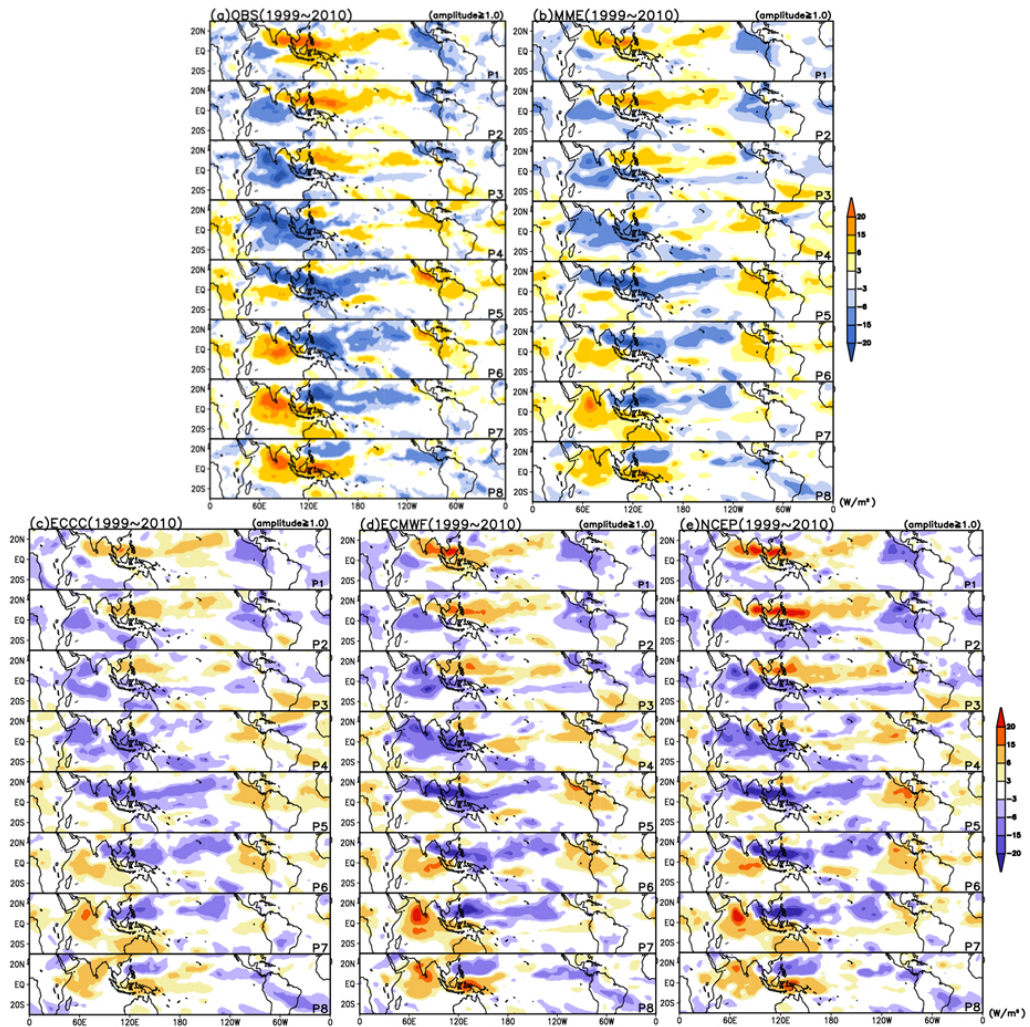


Figure 19. Same as in Figure 18 but for the period of May to October during 1999~2010.

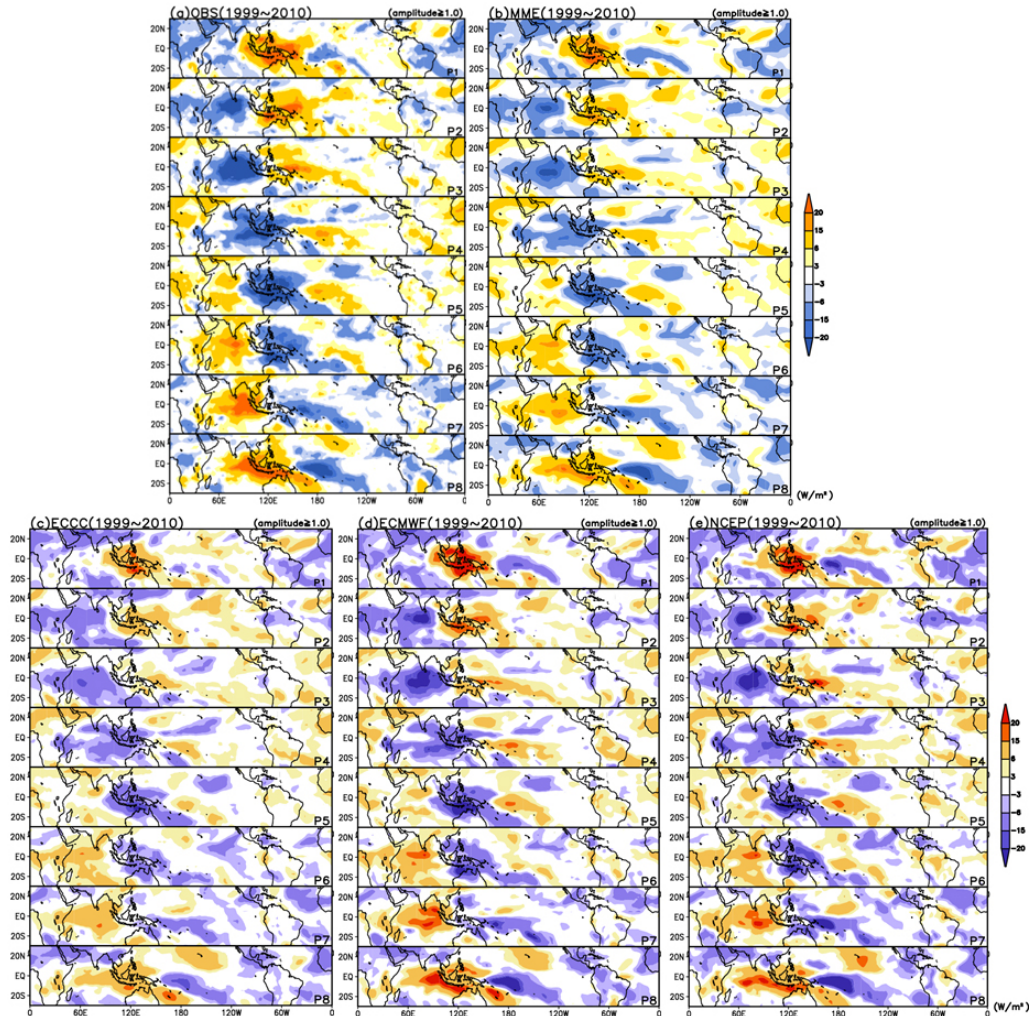


Figure 20. Same as in Figure 18 but for the period of November to April during 1999-2010.

예측 선행시간에 따라 각 모형별 MJO 재현 특성이 어떠한 차이를 보이는 지 알아보기 위해 적도지역(15°S-15°N) 평균 동서구조의 위상시간에 따른 분포를 Figure 21에 제시하였다. 앞에서 나타난 특징처럼 ECCC의 경우 세 모형간의 비교에서는 가장 약한 대류 활동이 나타나며 선행시간 2주에서부터 서태평양에서의 시그널은 아주 약하며 3-4주에서는 거의 찾아보기 힘들다. ECMWF는 모형들 중 가장 우수한 예측가능성을 보여주는데 3주 선행시간까지는 강도 또한 유지되며 MJO의 동진이 중태평양까지 잘 모의된다. 4주에서는 서태평양에서 시그널이 약화되지만 여전히 동진의 전파특성은 모의한다. NCEP

은 선행시간이 길어져도 동진하는 구조를 유지하고 있지만 선행시간이 길어질수록 MJO의 서태평양 대류영역의 표출이 다소 지연되어 나타나는 특징을 보인다. MME의 경우 우수한 개별 모형에 비해 강도는 약하지만 4주까지 동진하는 구조는 잘 모의되고 ECMWF에 비해서는 예측성이 떨어지는 것으로 판단된다.

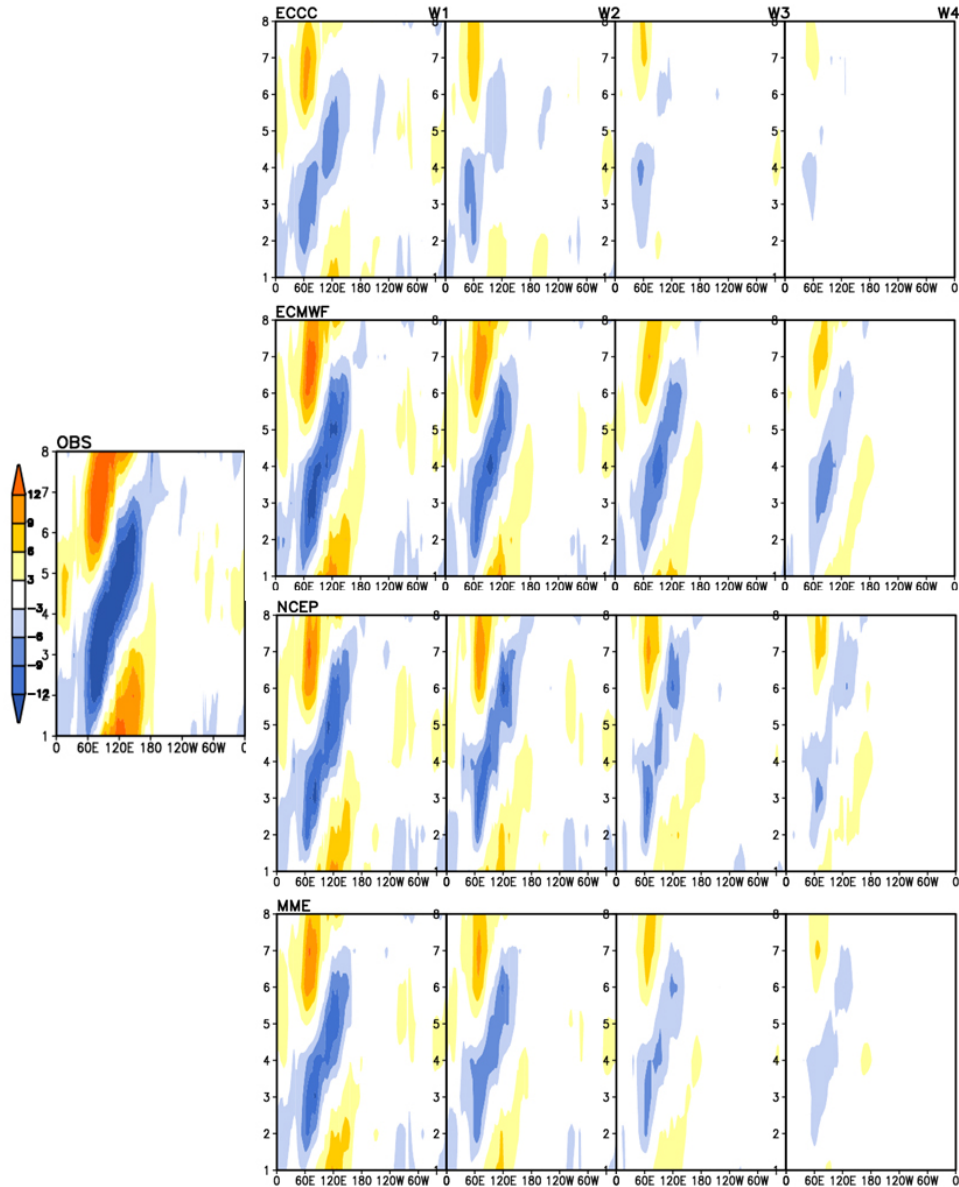


Figure 21. Longitude-MJO phase hovmoller diagrams of latitudinal mean (15°S–15°N) OLR anomalies from observation, ECCC, ECMWF, NCEP, and MME during all year round from 1999~2010.

4.1.3 MJO 예측성 평가

현업 예보를 하고 있는 몇몇 기관에서 과거재현자료를 이용하여 MJO를 가장 우수하게 예측한 경우 그 한계가 10-34일(겨울의 경우 14-42일)로 알려져 있지만 실제로 예측하게 될 경우 기대보다 하향될 수 있다(Waliser et al., 2003; Kim et al., 2014; Kim et al., 2018). MJO는 그 자체의 특성뿐만 아니라 원거리 영향력에서도 중요한 기후 인자이다. 적도에서 대규모로 발생하는 수평-연직 대류활동은 그 지역적인 위치에 따라 원격 상관패턴을 통해 전지구에 영향을 미친다.

MJO에 의한 대류활동과 이와 관련된 중위도 순환장의 선행시간별, 모형별 예측성을 보기 위해 Figure 22에 각 위상별 MJO OLR과 상층 지위고도의 공간상관값(PCC)을 구해서 전체기간, 여름, 겨울철로 나타내었다. 시계방향으로 모형별, MME의 결과를 각각 선행시간별로 볼 수 있다. 대류활동의 경우 가장 뚜렷한 특징으로 ECCO의 현저히 낮은 예측성을 들 수 있는데 선행시간 3주까지는 서서히 떨어지다가 4주에는 급격히 떨어져 전체기간의 경우 0.2까지 도달한다. 여름철이 겨울철에 비해 예측성이 더 높게 유지되는 것을 볼 수 있고 위상별 성능의 차이는 계절별로 상이하다. ECMWF의 경우 앞의 그림에서도 보았듯이 1주부터 4주까지 모든 기간에 대해 큰 감소 없이 0.6 이상의 값을 유지한다. 겨울철보다 여름철이 더 안정적으로 유지되는 것을 볼 수 있다. NCEP의 경우 1-2주는 유사한 경향을 보이다가 3주에 떨어지며 4주에는 급격하게 낮아지는 경향을 보인다. 여름철이 겨울보다 더 낮게 감소한다. MME를 하게 되면 특정 모형들에서 나타난 선행시간별 급격한 성능 감소가 없으며 1-4주간 비슷한 값으로 유지되는 것을 볼 수 있다. 그러나 ECMWF를 넘어서진 못하며 위상별 차이가 여름에 겨울보다 크게 나타난다.

MJO와 관련된 중위도 상층 고도장(200 hPa)의 결과를 보게되면 대류활동(OLR)보다는 위상별 차이도 크고 계절별 차이도 보인다. 전체기간에서 보게되면 모든 모형 및 MME에서 2주까지는 위상2를 제외하고 0.5 이상의 값을 유지하다가 3-4주에는 더 감소하는 경향을 보이며 NCEP의 경우 4주에서의 감소폭이 특정 위상에서 강하게 발생한다. 위상2에서의 성능 감소는 거의 모든 모형과 계절에서 발생하는 것으로 이에 대한 원인 분석은 추후 수행해야할 필요가 있겠다. 계절별로 나누어 보면 여름철이 겨울에 비해 위상별 성능 차이가 작으며 ECMWF의 경우 3주까지는 0.5를 유지하지만 4주에서는 몇몇

위상에서 0에 가까운 값으로 떨어진다. 겨울의 경우 더 큰 변동성을 보여 높을 경우는 여름철보다 우세하지만 낮은 경우는 더 낮은 예측성을 보인다.

계절내 예측성은 MJO의 강도에 따라서 영향을 받을 수 있다. 강한 대류활동으로 인해 유도되는 지역적 강한 순환구조는 그 주변장에 추가적인 에너지 전달이 가능하게 되기 때문인데 이러한 대류활동의 강도에 따른 예측성의 차이를 Figure 23에서 알아보았다. 모든 경우 MJO가 강할 때 관측과의 유사도가 높아지는 것을 알 수 있고 예상대로 선행시간에 따른 예측성의 감소가 뚜렷하며 강도에 상관없이 감소비율은 유사하게 나타난다. ECMWF에서 가장 높은 값을 보여주며 4주에서의 성능이 다른 모형들의 2주에서의 성능보다 높게 나타난다. 앞에서 제시된 결과와 유사하게 MME의 경우 다른 두 모형에 비해서는 높고 ECMWF보다는 낮다.

MJO에 의한 상층 순환장이 어느 정도 변환될 수 있는 지 알아보기 위해 WH(2004)의 RMM 지수가 설명하는 상층 순환장의 분산에 기여하는 비율을 퍼센트로 구해 관측과 MME에 대해 여름과 겨울로 구분하여 Figure 24와 Figure 25에 제시하였다. MJO의 영향이 가장 크게 설명되는 구역은 여름에는 아프리카/서인도양과 서-중태평양, 그리고 동태평양-남미-대서양으로 이어지는 열대지역이며 중위도로 올라가면 북위 25-30도가 한계지점으로 나타난다. MME의 경우도 관측과 아주 유사하게 나타나는데 선행시간에 따라 구역이 감소하고 비율도 감소되는 경향이 보인다. 겨울은 여름에 비해 동서의 차이가 크게 나타난다. 인도-남아시아-서태평양까지는 높은 영향력이 나타나지만 중-동태평양과 북미지역은 아주 낮으며 이는 그 구역에서부터 MJO의 쇠퇴가 발생하기 때문이기도 하다. 여름에 비해 보다 중위도 근처까지 발생하며 선행시기별 감소도 느리게 나타난다.

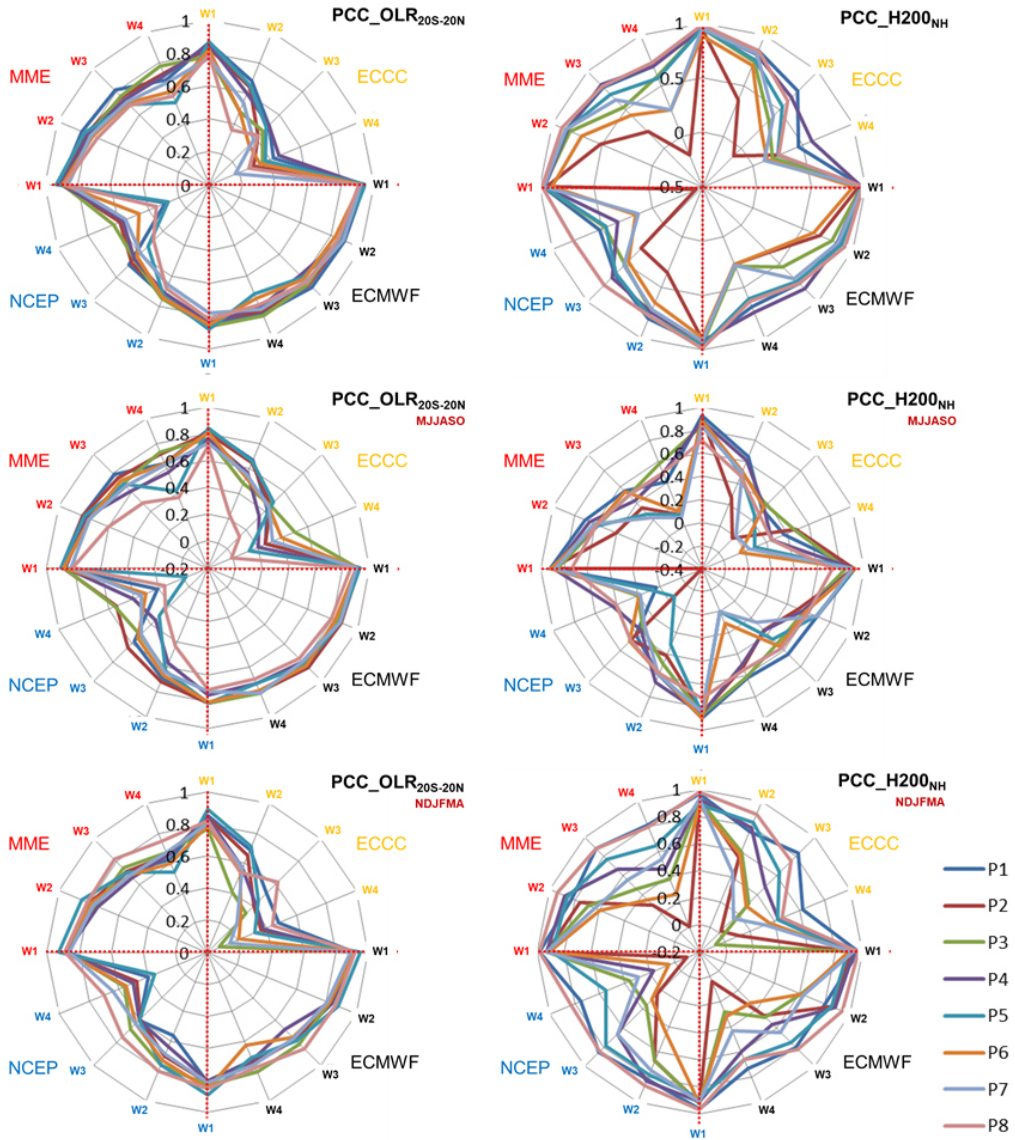


Figure 22. Prediction skills of anomalous OLR and geopotential height at 200 hPa simulated by ECCC, ECMWF, NCEP, and MME over the tropics (20°S–20°N for OLR) and northern extratropics (20°–90°N, H200) as function of the lead time by calculating spatial correlation coefficient with the observed pattern.

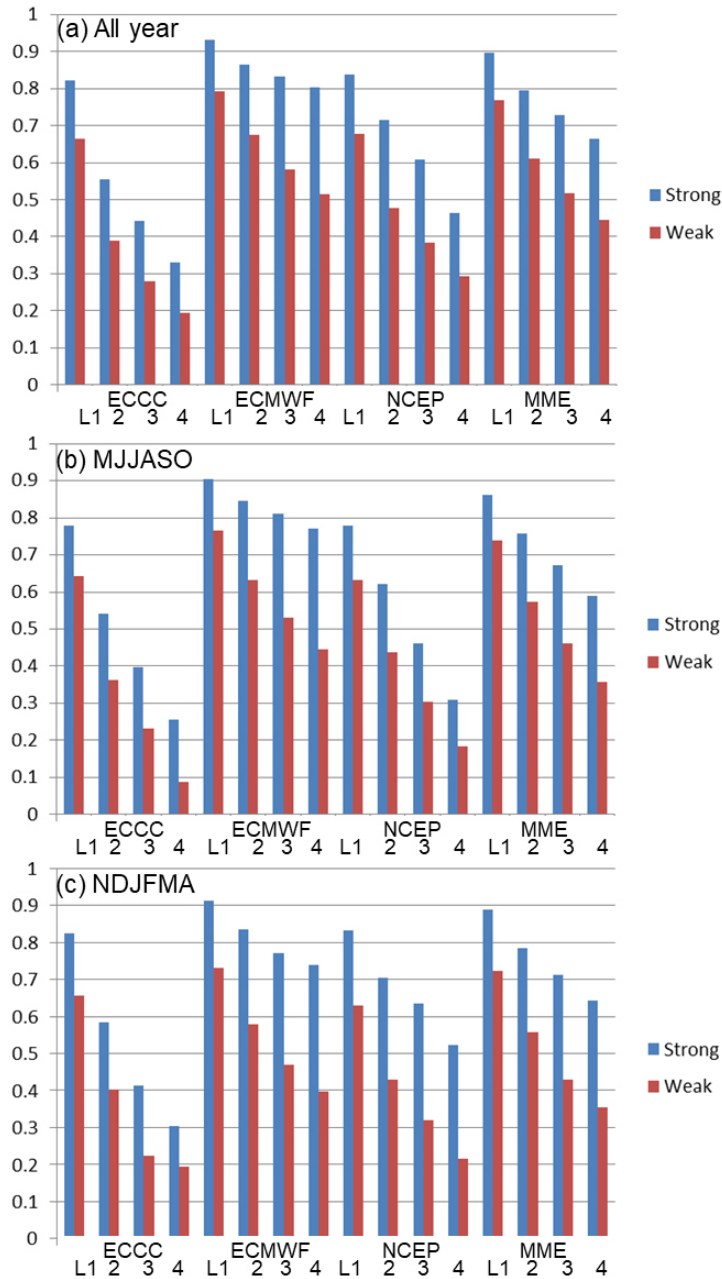


Figure 23. Prediction skills of anomalous OLR simulated by ECCC, ECMWF, NCEP, and MME over the tropics (20°S–20°N for OLR) as function of the lead time by calculating spatial correlation coefficient with the observed pattern. Blue/red bar indicates the strong/weak (amplitude greater/less than 1.0) MJO.

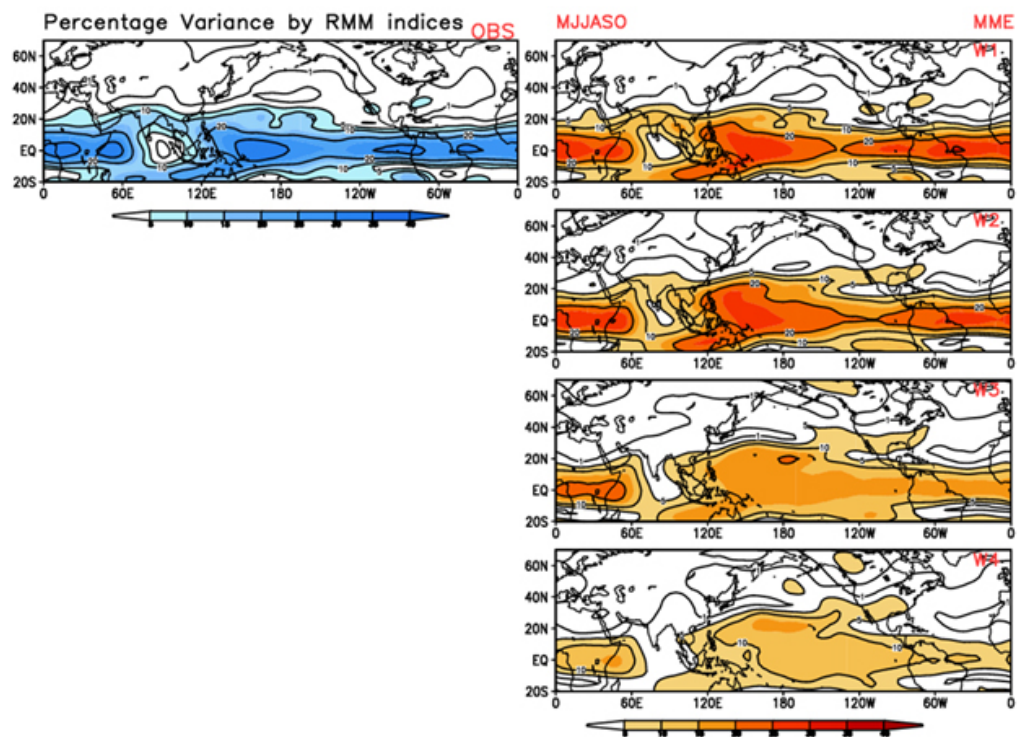


Figure 24. The percentage variance of geopotential height at 200 hPa anomalies accounted for by the RMM index during MJJASO from the observation (left) and the MME during lead time 1 to 4. The percentage variance is defined as $(corr.(RMM1,GPH200)^2 + corr.(RMM2,GPH200)^2) * 100$ (Kim et al., 2018).

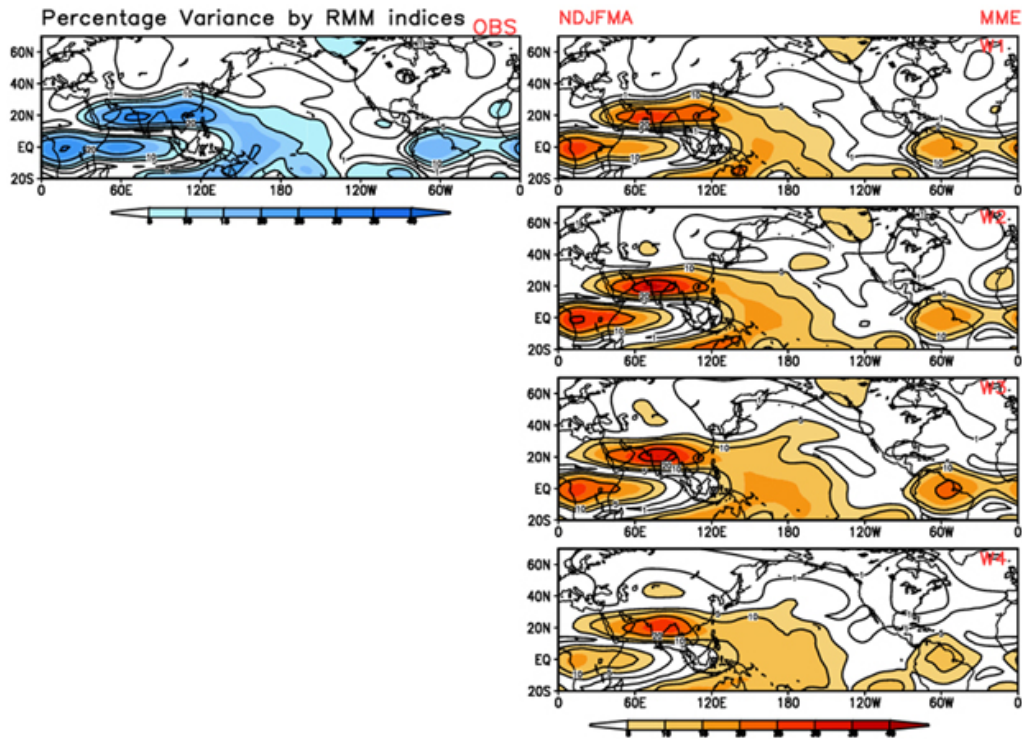


Figure 25. Same as in Figure 24 but for NDJFMA.

4.2 블로킹

4.2.1 관측블로킹의 특징

앞서 소개된 우랄블로킹 정의에 의해 탐지된 블로킹 이벤트 발생시 500 hPa의 지위 고도장의 수평분포는 Figure 26(a)와 같다. 500 hPa의 기압능이 40°-90°E사이에 선명하게 존재하며, 지위고도 양의 아노말리의 중심이 우랄산맥 근처인 60°N, 60°E에 위치한다. 이 고기압 동쪽 가장자리에 북풍이 강화되어 시베리아 지역에 찬 공기가 쌓이게 되고, 결국 동아시아 겨울철 한파를 유발하게 된다(Cheung et al., 2012). Figure 26(b)와 같이, 우랄 블로킹 블로킹 이벤트 발생시 블로킹 고기압 풍하측에 위치한 동아시아 기온의 음의 아노말리가 우세한 것을 확인할 수 있다.

이렇게 탐지된 우랄 블로킹은 Figure 27의 지연-선행 합성도에서 보인 바와 같이 수일에 걸쳐 지속된다.

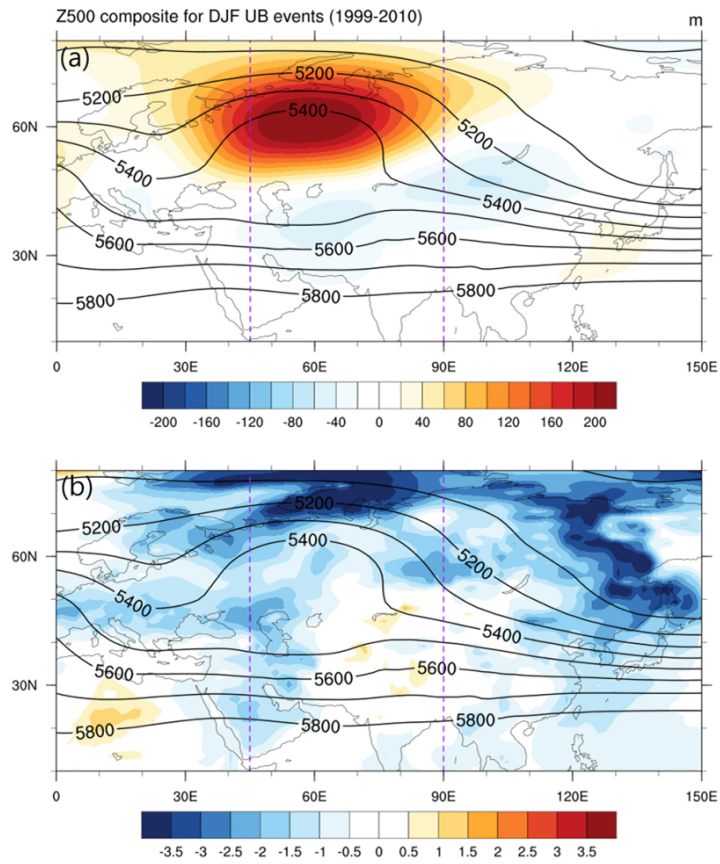


Figure 26. Composite of Z500 ((a): contour), Z500 anomalies ((a): shaded), and T2m anomalies ((b): shaded) for UB events during DJF 1999-2010.

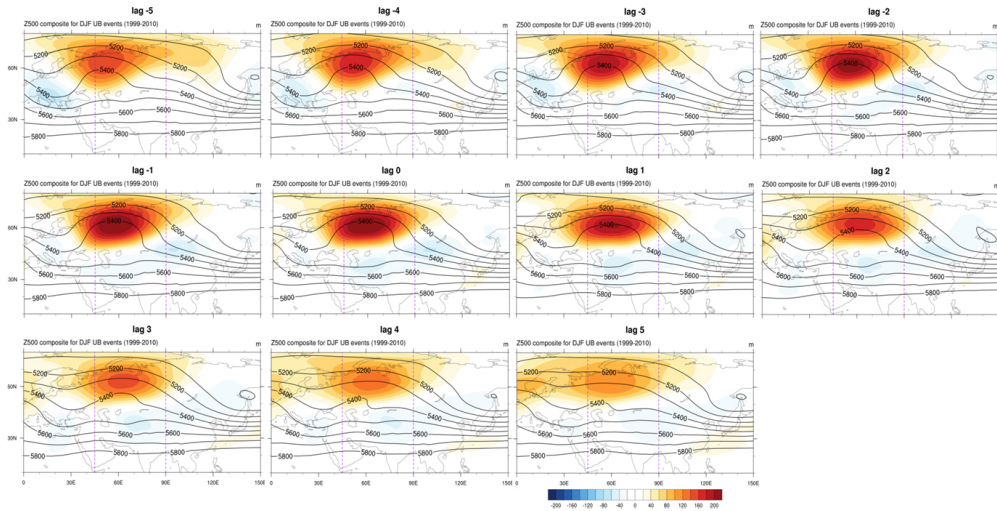


Figure 27. Lead-lag composite of Z500 (contour) and Z500 anomalies (shaded) for UB events during DJF 1999-2010.

4.2.2 개별모형의 UB 예측성능

앞 장에서 계절내 예측자료의 변수별 예측성능을 평가한 결과, 상대적으로 Z500이나 Z200과 같은 상층변수들의 예측성이 2m 기온이나 강수에 비해 더 나은 것으로 나타났다. 따라서 상층변수를 활용하여 유용한 예측정보를 이끌어내는 것이 계절내 예측자료의 활용성을 최대로 끌어올리는데 도움이 될 것이다. 이러한 측면에서 Z500을 사용하는 블로킹 지수의 예측은 계절내 예측에서 시도해볼 만한 과업이라 할 수 있겠다. 본 절에서는 세 개의 예측 모형에서 우랄 블로킹을 각각 잘 예측하는지 평가해보고자 한다. 본 보고서에는 2017년 예측장에 해당하는 과거 재현자료의 결과만을 수록하고자 한다. Figure 28은 예측선행시간에 따라 각 모델 별 전체 앙상블 멤버 중 몇 개의 앙상블 멤버가 겨울철 우랄블로킹 발생을 예측하고 있는가에 대한 비율을 나타낸다. 각 모델 별로 과거재현 자료 기간과 전체 앙상블 멤버의 수가 다른 것에 유의한다. 세 모형의 개별 앙상블 멤버들이 모두 1주전 예측에서는 관측에서 나타난 우랄 블로킹 발생과 유사한 양상을 보이나, 예측 선행시점이 길어질수록 실제 블로킹이 관측된 날에 블로킹 발생을 예측하는 앙상블 멤버 수가 줄어 들 뿐 아니라, 모델 간의 예측성능 차이를 보인다. 예를 들어, ECMWF 경우 3, 4주 전 예측에 겨울철 모든 기간 동안 앙상블 중 한개라도 블로킹이 발생한다고 예측

하는 경향이 두드러진다. 반면에 NCEP의 경우 3, 4주 전 블로킹을 예측하는 앙상블 멤버 수가 현저히 감소한다.

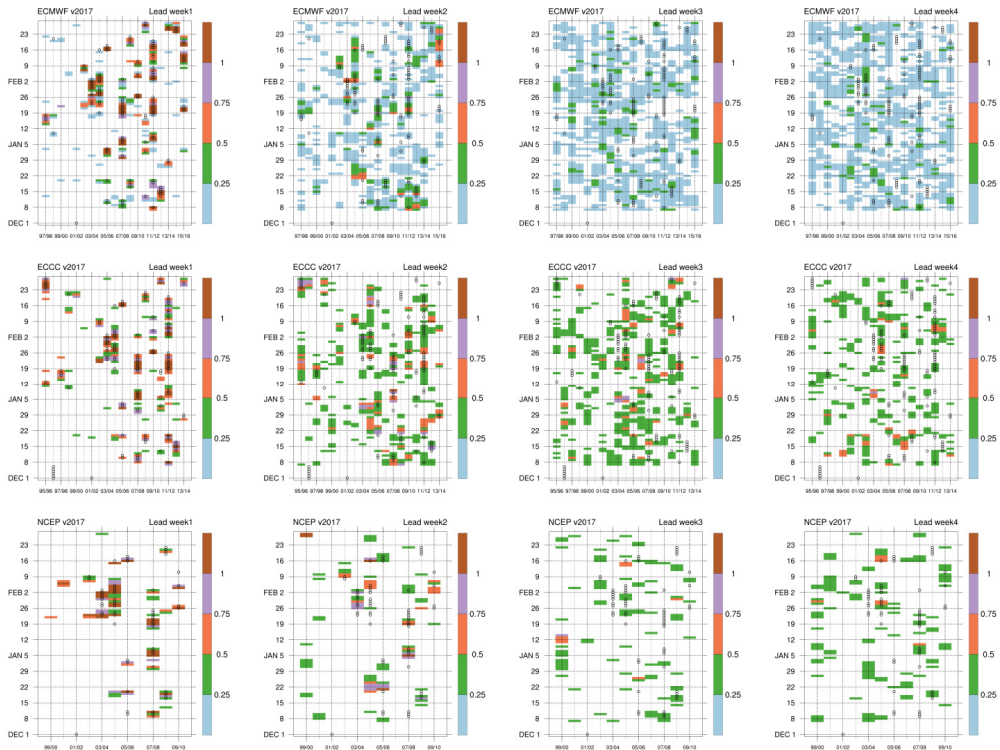


Figure 28. Ratio of number of ensemble members predicting DJF UB to total number of ensemble members of ECMWF (top), ECCO (middle), and NCEP (bottom) with different lead week. Marker indicates observed UB event. Note prediction years (x-axis) differ among models.

각 모델의 블로킹 예측성능을 정량적으로 평가하기 위해 예측 선행 시간 별 BSS를 비교하였다(Figure 29). 각 예측모형의 블로킹 발생 확률을 해당 예측 일자에 블로킹 발생을 예측하는 앙상블 멤버의 수를 전체 앙상블 멤버의 수로 나눠주는 것으로 얻어진다. 또한 각 모델 별로 과거 재현자료 기간이 다르므로 BSS 또한 모형별로 각기 다른 기간에 대해 구해졌음을 유의한다. BS의 비교값으로 1981-2010년간의 겨울철 블로킹의 기후학적 발생률을 사용하였다. Figure 29에서 1주전 예측의 경우 전체 모형의 예측이 블로킹 발생의 30년 기후값을 이용한 것 보다 우수하고, ECMWF, ECCO의 성능이 NCEP에 비해 좋은 것을 확인할 수 있다. 그러나 예측선행시점이 증가함에 따라 모형의 예측성능이

급격히 떨어지는 것을 확인할 수 있다. ECMWF의 경우 2주전 예측의 BSS값이 양의 값을 갖지만 매우 작다. 나머지 모형들은 모두 2주전 예측부터 BSS값이 음수값을 갖는 것을 확인할 수 있으며 이는 블로킹의 기후값에 의존한 예측성능에 모형 성능이 따라가지 못함을 의미한다.

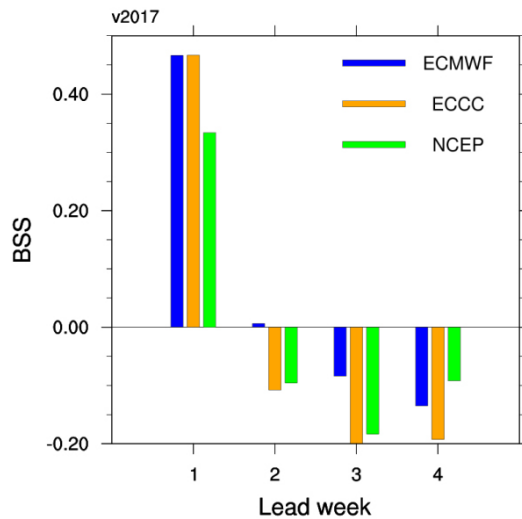


Figure 29. BSS for probabilistic UB forecast of each model.

블로킹 발생을 1, 그렇지 않은 경우를 0으로 취급하는 결정론적 예측방법을 취할 경우 모형 별 블로킹 발생을 예측하는 앙상블 멤버의 수에 따른 예측성능을 알아보하고자 한다. 예측성 평가에 있어 다양한 측면을 고려해야 하므로 Figure 30과 같이 4개의 평가 항목에 따라 그 예측성능을 평가해보고자 한다. 다시 말해, Figure 30은 각 예측선행시점에 최종적으로 블로킹 발생으로 간주되기 위해 고려될 예측 앙상블 개수를 달리했을 때 모형 별로 4개의 예측성 평가 지수에서 어떤 차이를 나타내는지 보인 것이다. Bias의 경우, 양의 값은 관측치보다 빈번한 블로킹 발생을 예측한 것을 의미하며 음의 값은 관측치보다 블로킹 발생이 덜 빈번할 것으로 예측함을 의미한다. 예를 들어 ECMWF는 예측선행시점에 관계없이 전체 12개의 앙상블 멤버 중 1개라도 블로킹 발생을 예측했을 때 최종 블로킹 발생을 1로 간주할 경우 Bias 값이 모두 양의 값을 나타낸다. 이는 관측보다 과도하게 자주 블로킹이 발생한다고 예측하는 것을 의미한다. 만약 전체 11개의 앙상블 멤버 모두

가 블로킹이 발생한다고 예측해야만 최종 블로킹 발생을 1로 간주할 경우 예측선행시점에 관계없이 모두 음의 값을 보이므로 이는 관측에 비해 덜 빈번하게 블로킹이 발생한다고 예측함을 의미한다. HSS의 경우 1에 가까울수록 예측성능이 우수한데, 1주전 예측의 경우 ECMWF는 블로킹 발생을 만족하는 양상블의 개수가 11개(100%)인 경우를 제외하고 모두 0.5 이상의 값을 보이거나, 예측선행시점이 증가함에 따라 그 값이 급속히 줄어드는 것을 확인할 수 있다. Figure 30의 False Alarm Ratio (FAR)는 Table 3의 기호를 이용하면 $B/(A+B)$ 로 표현이 되며 전체 이벤트 예측 중에 오경보율을 나타낸 것이다. FAR이 높을수록 부정확한 예측이 많음을 의미하는데, 그림에서 볼 수 있다시피 ECMWF의 경우, 1주 전 예측은 다른 예측선행 시점에 비해 FAR은 낮지만 양상블 개수를 적게 고려할수록 FAR이 높아지는 경향을 보인다. HR (Hit Rate)는 전체 블로킹 예측 중에 실제로 적중한 경우를 나타내며 $A/(A+C)$ 로 표현된다. ECMWF의 경우 1주전 예측의 적중률이 가장 적중률이 높은 경향을 보이며 선행시간이 증가할수록 그 값이 감소한다. 양상블 개수가 적을수록 HR이 높은 것을 알 수 있다. 세 모형 모두 3-4주 전 예측성능이 좋지 못하므로 2주전 예측성능을 극대화하는 방안을 모색하는 것이 현재 예측기술 수준에서 현실적으로 실현가능한 목표치라는 결론을 잠정적으로 내릴 수 있다. 따라서 Bias가 0에 가까울수록, HSS가 1에 가까울수록, FAR이 작을수록, HR이 높을수록 우수한 예측성능을 의미하므로 종합적으로 판단을 내렸을 때, 결정론적 예측의 경우 ECCO나 NCEP의 경우 4개의 양상블 중 1개의 양상블이라도 블로킹 발생을 예측했을 때, ECMWF의 경우 11개의 양상블 중 25%에 해당하는 2-3개의 양상블이 블로킹을 예측했을 때 블로킹 발생으로 간주하는 것이 적절한 것으로 보여진다.

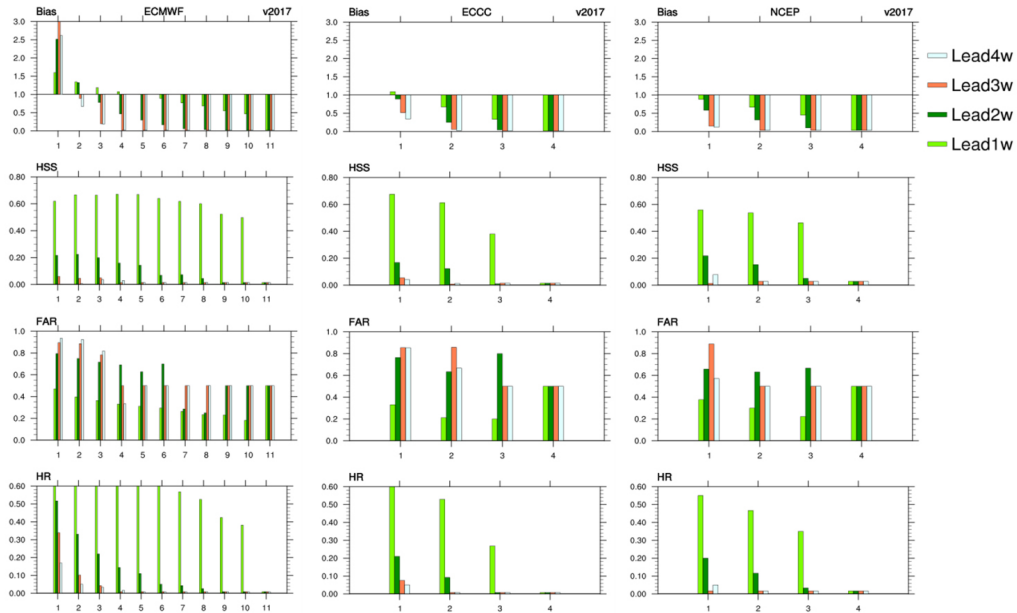


Figure 30. Skill score considering different number of ensembles for a UB event to occur.

4.2.3 MME의 UB 예측성능

이번 절에서는 세 모형의 과거재현자료가 공통적으로 존재하는 1999-2010년에 MME에 의한 우랄 블로킹의 예측성능을 평가하고자 한다. Figure 31은 세 개 모형의 앙상블 멤버를 모두 합친 것 중에서 우랄 블로킹의 발생을 예측한 멤버의 수의 비율을 예측 선행시간에 따라 나타낸 것이다. 앞서 개별 모형에서 모았던 것과 유사하게 예측 선행 시간이 길어질 수록 관측과 다른 예측의 양상을 띄며, 4주 전 예측의 경우는 겨울철 거의 대부분의 날에서 전체 앙상블 멤버의 25% 이상의 앙상블 멤버가 블로킹이 발생할 것으로 예측하고 있다. 보다 정량적인 비교를 위하여 예측 선행시점에 따른 BSS를 계산 하였다(Figure 32). 1999-2010년의 공통기간에 대해 3 모형을 비교한 결과 1주전 예측에서는 ECCO가 세 모형 중 가장 좋은 성능을 보였다. 예측 선행 시점이 증가함에 따라 BSS는 급격하게 감소하지만 2주전 예측에서 3 모형 중 ECMWF가 그나마 나은 성능을 보이는 것으로 나타났다. MME의 블로킹 예측성능을 비교해 본 결과 특히 2주 전 예측에서 그 성능이 가장 우수한 것으로 나타났다.

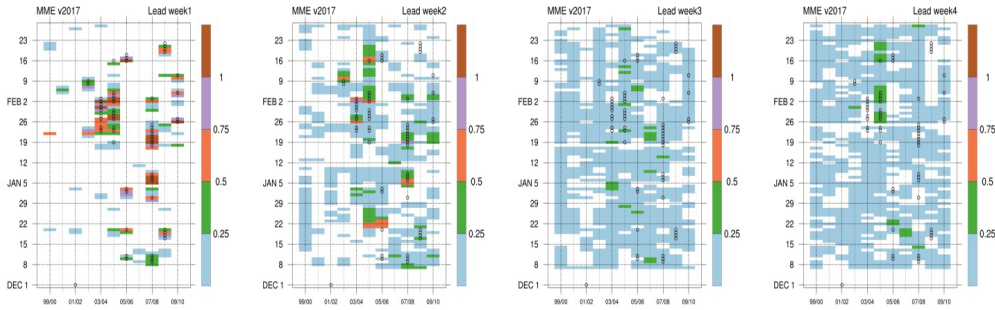


Figure 31. Ratio of number of ensemble members predicting DJF UB to total number of ensemble members of MME from 1999/2000 to 2009/2010 with different lead week. Marker indicates observed UB event.

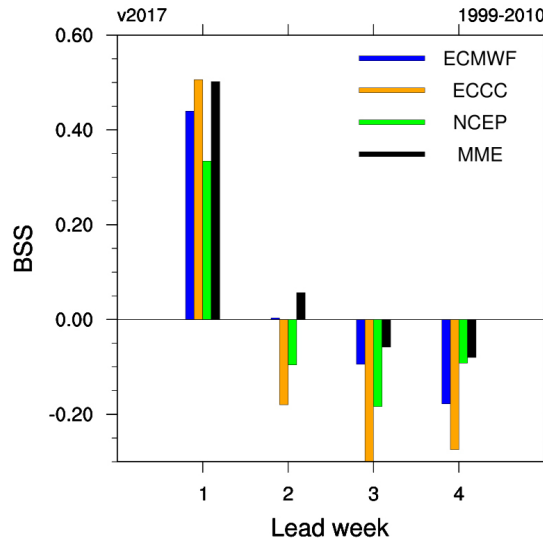


Figure 32. BSS for probabilistic UB forecast of each model and MME.

MME 예측의 수평 분포를 살펴보기 위해 관측 우랄 블로킹 발생시 500 hPa 지위고도장의 패턴을 나타내었다. Figure 33은 관측에서 우랄 블로킹이 발생 시, 해당 날짜로부터 예측 선행시간 4주까지에 3개 모형 모든 앙상블 멤버의 지위고도장을 합성하여 나타낸 것이다. MME의 경우, 1주전 예측에서 우랄 산맥 근처에서 선명하게 보이던 기압능이 예측선행시간의 증가와 함께 점차 사라지면서 블로킹 패턴이 소멸되는 것이 보인다. 이와 함께 예측 선행 시간이 증가함에 따라 전체 모형 앙상블 스프레드가 북위 60도를 기준으로 증가하는 것을 확인할 수 있다(Figure 34).

MME UB composite

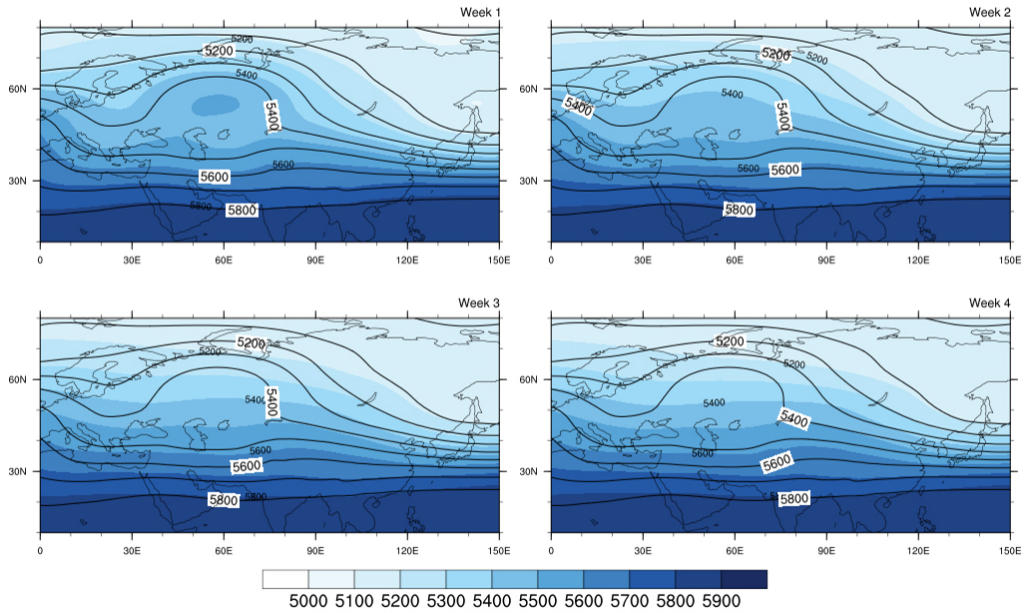


Figure 33. Z500 composite for DJF UB events from 1999/2000 to 2009/2010 (Contour: observation, Shaded: MME).

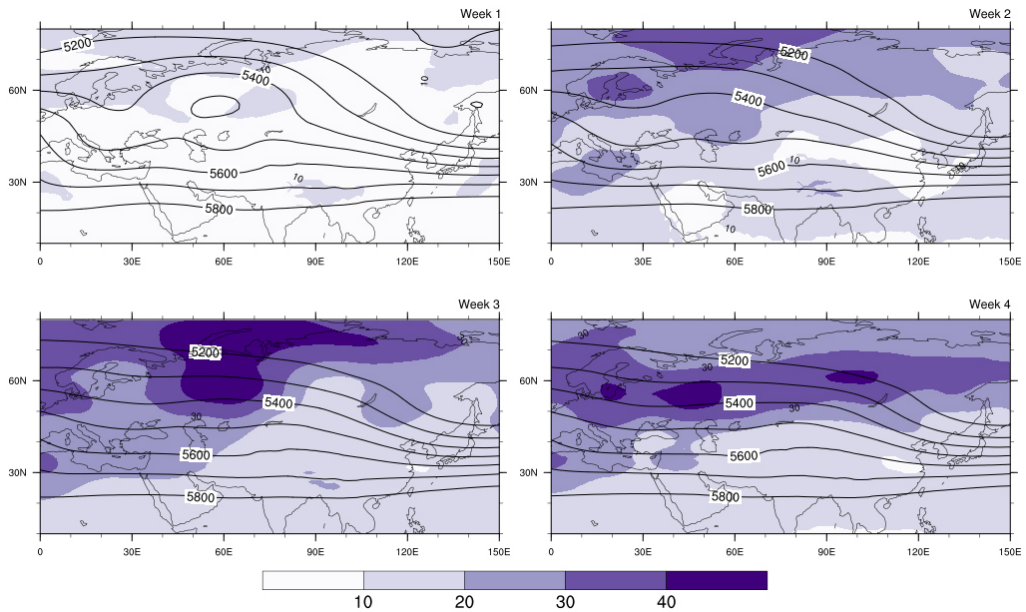


Figure 34. Multi-model ensemble spread (shaded) and composite of Z500 (contour) for DJF UB events at different lead weeks.

5. 결론

APEC 기후센터는 국제협력을 바탕으로 수집 및 가공된 다중모델앙상블 기후예측 정보를 제공하며 다양한 기법의 MME 개발을 통해 기후예측기술의 선진화를 위한 노력을 꾸준히 해오고 있다. 최근, 이음새 없는 예보를 위한 국제적인 관심이 쏟아지고 기상산업 및 응용분야에서의 수요가 증가하고 있는 바, 예측성의 불모지로 알려진 계절내 예측 분야에 핵심기술을 확보하고 국제적인 흐름에 적극 참여하기 위한 자체적인 역량 강화가 필요하게 되었다.

따라서, 본 과제는 시작단계로 향후 계절내 규모에 적합한 예측기술 및 검증기술을 개발함에 있어 필요사항 등이 무엇인지 여러 기준들을 제시하고 계절내 예측의 현주소를 객관적으로 평가하는 것이 목표이다. 이 보고서는 아래의 순서로 작성되었고 간략한 결과를 함께 제시하고자 한다.

5.1 계절내 예측성 평가(계절내 예측의 현주소 파악)

○ 계절내 기후 모의 능력 검증

- 선행시간이 길어짐에 따라 공간분포 유사성이 감소하고 절대오차가 증가하는 등 일관되게 성능 감소가 나타나지만 모형별/지역별 격차가 크다. 경년변동은 관측에 비해 작고 선행시간이 길어짐에 따라 감소하는 경향이 뚜렷하다.
- SNR 값은 기준 기간과 앙상블 개수에 민감하나, 조건을 통일시켰을 때 그 값이 NCEP, ECMWF, ECCC 순으로 줄어들었다. NCEP과 ECCC의 경우 앙상블 개수가 처음부터 동일하나 앙상블간의 스프레드는 매우 다른데, 모형 자체, 접합 구조, 앙상블 처방 등이 상이한 것 때문으로 추측된다.

○ 과거재현자료 예측성 평가

- 계절내 예측 과거재현자료는 대략 2~3주의 예측가능일을 보이고, 모형별로 선행시간이 다르므로 직접적으로 비교할 수는 없지만 ECMWF가 전체적으로 가장 긴

예측가능일을 보인다. Z200은 다른 변수에 비하여 임계값이 증가함에도 예측가능일이 크게 줄어들지 않는 특징이 있다.

- 대부분 지역에서의 계절내 예측성능의 경년변동은 ENSO와 크게 연관되어 있고, 특히 선행시간이 증가할 수록 ENSO에 대한 예측성능 의존도가 커지며 중기예보에서 장기예보의 특성이 나타나는 것으로 판단된다. 그러나, 중동 지역의 계절내 예측성은 ENSO에 그다지 의존하지 않고 적도와 극 사이의 온도차 같은 중위도 자체 역학이 예측성에 영향을 주는 것으로 생각된다.

○ 실시간 예측자료의 예측성 평가

- 모형/선행시간/변수/예측해에 따른 성능 변화를 확인하였다.
- 선행 주수가 증가함에 따라 성능이 급격히 감소하고 엘니뇨 해인 2016년이 성능이 2017년에 비해 높으며 강수 예측성능이 다른 변수에 비해 매우 낮다.
- 다중모형 합성법을 처방했을 때(특히 MME_g) 최소 ECMWF 수준의 성능을 보여 그 효과를 확인할 수 있다.
- 선행 주수가 길거나, 하층 기온의 경우, 또는 북반구 아적도 지역에서 MME_g가 ECMWF 성능을 넘어서는 경우가 훨씬 많았다. 하지만 특정 검증 지수(LEPS)에서는 MME 처방으로도 ECMWF 성능을 넘지 못하는데, 목적에 따라 적절한 검증 지수 선택이 중요함을 알 수 있다.
- 긴 선행시간에서 성능은 지역별로 큰 차이가 있다. 대륙을 포함하는 동아시아, 북러시아, 중동 지역 등 북반구 아적도 지역의 성능은 변수에 관계없이 해양을 포함하는 지역에 비해 아주 낮은 예측성을 보였다. 강수량이 거의 없는 중동지역에서는 통계 특성 때문에 카테고리 예측성능이 높은 것처럼 인위적인 결과가 도출되며, 따라서 검증 대상에 대한 이해를 바탕으로 검증 지수를 적절하게 해석하는 것이 중요하다.

5.2. 계절내 기후현상의 재현 특성 및 예측성 제시

계절내 예측에서 중요하게 고려되어야 할 여러 기후인자들 중 두가지, MJO와 블로킹의 재현 특성과 예측성을 평가하였다. MJO는 계절내 규모에서 가장 중요하게 고려되고

있는 예측인자의 하나이며 이미 MJO 예측은 현업에서 제공되고 기후예측에 활용 되고 있다. 블로킹의 경우, 계절내 예측자료의 변수별 예측성능을 평가해 보면 Z500이나 Z200과 같은 상층 변수들의 예측성이 2m 기온이나 강수에 비해 상대적으로 우수하게 나타나 이러한 상층 변수를 활용하여 유용한 예측정보를 이끌어 내는 것이 계절내 예측자료의 활용성을 최대화 시키는데 도움이 될 것으로 판단되어 선택하였다.

○ MJO

- 3개의 개별 모형 및 MME가 재현하는 MJO 는ECCC 모형이 MJO의 강도, 동진 속도 및 쇄퇴 지역 등에서 가장 과소모의 하였지만 나머지 두 모형(ECMWF와 NCEP)은 관측과 아주 유사하게 나타났다. MME는 강도가 다소 약화되었지만 전반적인 MJO의 전파특성을 잘 표현하였다.
- MJO와 관련된 적도 대류활동과 중위도 순환장의 선행시간별, 모형별 예측성을 본 결과, 대류활동의 경우 ECCC의 현저히 낮은 예측성과 NCEP 모형과 함께 3-4주에서의 예측성의 급격한 감소가 나타났고 ECMWF와 MME는 선행시간의 증가에 따른 예측성의 감소가 거의 없었고 ECMWF가 가장 우수한 예측성을 유지하였다. 위상별 성능의 차이는 여름철이 NCEP과 MME에서 겨울에 비해 크게 나타났다. 중위도 순환장의 경우 위상별 성능의 차이도 적도 대류활동에 비해 훨씬 크고 겨울이 여름에 비해 위상간의 차이가 더 크게 발생하였다.
- MJO의 강도에 따른 예측성 비교 결과, MJO가 강할 때 관측과의 유사도가 높아 졌고 강도에 상관없이 선행시간에 따른 예측성의 감소는 유사했다. 모형별 비교에서는 ECMWF가 모든 경우에서 가장 높은 값을 보여주었고 4주에서의 성능이 다른 모형 및 MME의 2주의 경우보다 높게 나타났다.
- MME의 경우 각 모형들의 단순평균에 의한 결과로 강도면에서 약화된 특징이 있었지만 ECCC와 NCEP에서 보여졌던 약점들이 상쇄되는 영향이 나타났고 우수한 ECMWF를 능가하진 못했다.

○ 블로킹

- 블로킹 지수는 40° - 90° E 내 각 경도에 대해 Z500의 남북경도를 구하여 블로킹을 판별한다. 아노말리를 정의할 필요가 없는 점이 장점이다. S2S 예측자료와 같이 자료 기간이 일정하지 않고 계절내 예측 자료의 아노말리를 구하는 방법이 정형화 되지 않은 상황에서 사용하기 적합하다.
- 개별 모형의 경우 모두 1주 전 예측에서는 관측에서 나타난 우랄 블로킹 발생을 잘 예측하였지만, 예측 선행시간이 길어질수록 예측성능이 저하되었다.
- 다중 모형 앙상블을 통해 2주전 예측의 성능을 다소 개선시킬 수 있었다.
- 블로킹의 계절내 예측은 현재로서 3-4주전 예측은 불가능해 보이지만, 2주 전 예측은 어느 정도 활용이 가능할 것으로 보인다.
계절내 시간규모에 적합한 블로킹의 주별 예측 정보의 가공을 더욱 정교화하는 노력이 필요하다.

REFERENCES

- Andrew W. Robertson, Arun Kumar, Malaquias Pena, and Frederic Vitart, 2015: Improving and Promoting Subseasonal to Seasonal Prediction. *BAMS*, 96, ES49–ES53.
- Barnston, A. G., L. Shuhua, S. J. Mason, D. G. DeWitt, L. Goddard, and X. Gong (2010), Verification of the first 11 years of IRI's seasonal climate forecast, *J. Appl. Meteorol. Climatol.*, 49, 493–520.
- Brunet et al., 2010, Collaboration of the Weather and Climate Communities to Advance Subseasonal-to-Seasonal Prediction, *BAMS*, DOI :10.1175/2010BAMS 3013.1.
- Franzke, C., S. B. Feldstein, and S. Lee, 2011: Synoptic analysis of the Pacific-North American teleconnection pattern. *Quarterly Journal of the Royal Meteorological Society*, 137, 329–346.
- Kim, H.-M., P. J. Webster, V. E. Toma, and D. Kim, 2014: Predictability and prediction skill of the MJO in two operational forecasting systems. *J. Climate*, 27, 5364–5378, doi:10.1175/JCLI-D-13-00480.1.
- Lau, K. M., and J. S. Boyle, 1987: Tropical and Extratropical Forcing of the Large-Scale Circulation: A Diagnostic Study. *Monthly Weather Review*, 115, 400–428.
- Lee DY, K Ashok, and JB Ahn (2011): Toward enhancement of prediction skills of multimodel ensemble seasonal prediction: A climate filter concept, *J Geophys Res*, 116, D06116, doi:10.1029/ 2010JD014610
- Lee DY, JB Ahn, and K Ashok (2013): Improvement of multi-model ensemble seasonal prediction skills over East Asian summer monsoon region using a climate filter concept, *J. Appl. Meteorol. Climatol.*, 52, 1127–1138.
- Luo, D., T. Gong, and L. Zhong, 2008: Dynamical Relationship between the Phase of North Atlantic Oscillations and the Meridional Excursion of a Preexisting Jet: An Analytical Study. *Journal of the Atmospheric Sciences*, 65, 1838–1858.
- Min YM, VN Kryjov, and CK Park (2009): A probabilistic multimodel ensemble approach to seasonal prediction, *Weather Forecasting*, 24, 812–828, doi:10.1175/2008WAF2222140.1
- Min YM, VN Kryjov, and JH Oh (2011): Probabilistic interpretation of regression-based downscaled seasonal ensemble predictions with the estimation of uncertainty, *J Geophys Res*, 116, D08101, doi:10.1029/2010JD015284
- Min YM, VN Kryjov, and SM Oh (2014): Assessment of APCC multimodel ensemble prediction

- in seasonal climate forecasting: Retrospective (1983–2003) and real-time forecasts (2008–2013), *J Geophys Res*, 119, 12,132–12,150, doi:10.1002/2014JD022230
- Nakamura, H., M. Tanaka, and J. M. Wallace, 1987: Horizontal Structure and Energetics of Northern Hemisphere Wintertime Teleconnection Patterns. *Journal of the Atmospheric Sciences*, 44, 3377–3391.
- Potts, J.M., C.K. Folland, I.T. Jolliffe, and D. Sexton, 1996: Revised “LEPS” Scores for Assessing Climate Model Simulations and Long-Range Forecasts. *J. Climate*, 9, 34–53, [https://doi.org/10.1175/1520-0442\(1996\)009<0034:RSFACM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<0034:RSFACM>2.0.CO;2)
- Rashid, H. A., H. H. Hendon, M. C. Wheeler, and O. Alves, 2011: Prediction of the Madden-Julian oscillation with the POAMA dynamical prediction system. *Climate Dyn.*, 36, 649–661, doi:10.1007/s00382-010-0754-x.
- Smith, D. M., A. A. Scaife, R. Eade, and J. R. Knight (2014), Seasonal to decadal prediction of the winter North Atlantic Oscillation: emerging capability and future prospects, *Quarterly Journal of the Royal Meteorological Society*.
- Smith, D. M., A. A. Scaife, R. Eade, and J. R. Knight (2016), Seasonal to decadal prediction of the winter North Atlantic Oscillation: emerging capability and future prospects, *Quarterly Journal of the Royal Meteorological Society*, 142(695), 611–617.
- Vitart, F., A. Leroy, and M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, 138, 3671–3682, doi:10.1175/2010MWR3343.1.
- Vitart, F., and Coauthors (2016), The Subseasonal to Seasonal Prediction (S2S) project database. *Bull. Amer. Meteor. Soc.*, 98(1), 163–173.
- Wang, B., et al. (2009), Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004), *Clim. Dyn.*, 33,93–117, doi:10.1007/s00382-008-0460-0.
- WMO, 2012, *WMO Bulletin*, 61 (2), 48pp.
- WMO Publication, 2015: *Seamless Prediction of the Earth System: from minutes to months*.
- Zhang, C., J. Gottschalck, E. D. Maloney, M. Moncrieff, F. Vitart, D. E. Waliser, B. Wang, and M. C. Wheeler, 2013: Cracking the MJO nut. *Geophys. Res. Lett.*, 40, 1223–1230, doi:10.1002/grl.50244.

【연구자】

문자연 선임연구원

이윤영 선임연구원

오지현 선임연구원

김가영 연구원

윤순조 연구원

연구보고서 2018-08

아태지역 15-60일 MME 예측 기술 개발 및 검증 기술 개발

I S B N 979-11-5698-240-1

발 행 2019년 3월

발 행 인 권 원 태

발 행 처 APEC기후센터
부산시 해운대구 센텀7로 12

제작·인쇄 경성문화사

이 보고서는 APEC기후센터 홈페이지(<http://www.apcc21.org/>)에서 이용하실 수 있습니다.
또한 보고서에 실린 내용은 출처를 명시하면 자유롭게 인용할 수 있습니다. 단, 무단 전재 및
복제를 금합니다.