

상세화 기법을 응용한 한반도 전문가 계절 예측 시스템 II

Expert Seasonal Prediction System for Seasonal Outlook in Korea
(ESPreSSO-KR) II

김원무 서비스개발팀

여새림, 김유진 기후분석팀

발간사

우리 센터는 국내 및 아태 지역의 지속 가능한 발전을 지원하기 위하여 다중모델 앙상블 기법을 이용한 신뢰도 높은 계절 기후 예측 정보를 생산하고 제공하고 있습니다. 우리 센터에서 운영하는 계절 기후 예측 시스템은 참여 모델의 고도화 및 다변화, 운영 시스템의 안정화, 표출 방식의 다양화 등을 통하여 지속적으로 발전해 왔습니다. 한편으로는 특정 수요자들을 위한 맞춤형 기후 정보 제공을 위한 계절 예측 신뢰도 향상과 맞춤형 콘텐츠 제작을 위하여 노력해왔습니다.

이 연구 보고서 『상세화 기법을 응용한 한반도 전문가 계절 예측 시스템 II』는 우리 센터에서 제공하는 다중모델앙상블 기반의 계절 기후 예측 정보를 우리나라의 기후 예측에 특화하여 예측 성능을 극대화하고자 하는 취지에서 시작되었습니다. 이미 우리 센터에서는 지난 해 『상세화 기법을 응용한 한반도 전문가 계절 예측 시스템』을 발간한 바 있는데, 이를 통하여 개발된 역학-통계, 객관-주관 하이브리드 계절 예측 시스템인 ESPreSSO-KR (Expert Seasonal Prediction System for Seasonal Outlook in Korea)은 지난 2년여 간의 실시간 운영을 통하여 그 높은 예측 성능과 안정성이 검증되었습니다. 이 연구에서는 ESPreSSO-KR의 기능 향상(자동화 및 확률 예보 기능 개발), 시스템 안정화(프로그래밍 언어 변경, 유지 보수 모듈 추가), 사용자 주문 반영(예측 범위 확장, 권역별 예보 기능 추가) 등을 통한 과학적/기술적 시스템 고도화를 이룩하였습니다. 이와 함께 겨울철 한반도 기후 예측 인자에 대한 심도 깊은 연구를 진행하여 학술적 정보도 제공하였습니다. 이 연구를 통하여 개발되고 기능이 향상된 하이브리드 계절 기후 예측 시스템

인 ESPreSSO-KR이 우리나라 계절 예측 성능의 향상에 도움이 되어 기후 예측
정보의 사용자들에게 더욱 신뢰할만하고 유용한 정보가 전달되길 기대합니다.

2018년 3월
APEC 기후센터
원장 권 원 태

ABSTRACT

The Expert Seasonal Prediction System for Seasonal Outlook in Korea (ESPreSSO-KR) v1.0 was developed in 2017 to provide reliable seasonal predictions of monthly mean temperature. It is a hybrid dynamical-statistical system based on seasonal predictions from the APEC Climate Center Multi-Model Ensemble as well as its statistical downscaling techniques and expert knowledge. We have added several new features and sub-modules to the ESPreSSO and proudly present the new ESPreSSOv2.1. It has been completely re-written in Python, a more general language than the original GrADS + FORTRAN. The new ESPreSSOv2.1 provides extended seasonal forecasts with a lead time of up to three-months and sub-regional (provincial) forecasts as requested by the Korea Meteorological Administration. Furthermore, a probabilistic categorical forecast module was developed to generate tercile seasonal forecasts. It also automatically runs in the background, generates new seasonal forecasts, and sends seasonal outlooks to subscribers. We also improved maintenance modules so that an administrator can check the status of ESPreSSOv2.1 at a glance. Each module is self-contained and replaceable so that an administrator can easily switch computational methodologies. We compare the technical details of ESPreSSOv1.0 vs. ESPreSSOv2.1 and provide the practical basis for upgrading ESPreSSO.

목 차

1. INTRODUCTION	1
2. EPreSSO ^{v2.1}	5
2.1 프로그래밍 언어의 변경(포팅)	5
2.2 권역별 예보 기능	6
2.3 확률 삼분위 예보 기능	6
2.4 예측 선행 기간의 확장	7
2.5 자동화 기능 및 인터페이스 개선	8
3. SCIENCE behind the EPreSSO (December)	9
4. RESEARCH RESULTS	12
4.1 EPreSSO ^{v2.1} 의 예측 결과	12
4.2 EPreSSO의 예측 성능 평가 및 검증	18
4.3 프로젝트의 기타 성과물	19
5. CONCLUDING REMARKS	20

■ REFERENCES	21
■ 부록 1. ESPreSSO ^{v2.1} performance	22
■ 부록 2. ESPreSSO ^{v2} core (funcESPreSSO)	23
■ 부록 3. ESPreSSOv2.1 outlook example (2-3-month lead)	30

1. INTRODUCTION

정확하고 믿을만한 계절 예측 정보를 제공하기 위하여, 지난 연구 과제인 [상세화 기법을 응용한 한반도 전문가 계절 예측 시스템: Expert Seasonal Prediction System for Seasonal Outlook in Korea (ESPreSSO-KR)]을 통해 ESPreSSO^{v1.0}을 개발하였다. ESPreSSO는 하이브리드 역학-통계 계절 예측 시스템으로서, APCC의 현업 다중 모형 앙상블(Multi-Model Ensemble, MME) 계절 예측과 기후 예측 전문가의 경험과 지식을 기반으로 우리나라에 최적화된 계절 전망을 생산하는 시스템이다. ESPreSSO의 방법론과 시스템은 대한민국 특허 제 10-1901311호로 등록되었고, 논문(Kim *et al.*, 2018)을 통하여서도 소개되었다. 실제 예측성 측면에서도, ESPreSSO의 계절 전망은 지난 2016-2017년 운용 기간동안 높은 안정성과 적중률(2017년 Heidke Skill Score 기준 계절 예측 내부 검증 결과: 기상청 18.8, APCC MME 37.5, APCC 공식 예보 37.5, SCoPS 25.0, ESPreSSO 50.0)을 보여주었고, 2018년에도 꾸준히 계절 전망을 생산하여 제공(APPENDIX I)하고 있다. 이번 연구를 통해 사용자의 요구를 적극 반영하여 한층 발전된 ESPreSSO^{v2.1}을 개발(APPENDIX II)하였으며, 이 보고서에서는 지난 연구 결과를 바탕으로 새로운 버전에 추가된 기능과 개선 사항에 대한 기술적인 내용을 설명하고자 한다.

ESPreSSO는 전구 수치 기후 모형의 예측값을 기후 예측 전문가가 수치 모형의 특성과 한계를 고려하여 표적 기후에 적합한 하이브리드 역학-통계 계절 예측 시스템이다. 따라서, ESPreSSO가 안정적이고 신뢰할만한 예측 결과를 생산하기 위해서는 다음의 세 가지 기본 가정이 필요하다.

- 첫째, 기반 역학 모형은 어느 정도 예측 성능이 있어야 한다.
- 둘째, 표적 변수는 적절히 균질하며, 한편으로는 적절히 비균질해야 한다.
- 셋째, 전문가는 기후, 모형, 그리고 ESPreSSO에 대한 지식을 갖추어야 한다.

지난 연구 과제인 [상세화 기법을 응용한 한반도 전문가 계절 예측 시스템: Expert Seasonal Prediction System for Seasonal Outlook in Korea (ESPreSSO-KR)]을 통해, ESPreSSO의 기반 역학 예측 시스템인 APCC의 MME는 어느 정도 예측 성능이 있고, 표적 변수인 우리나라 기온은 전 계절에 대하여 적절히 균질/비균질한 특성을 만족

하는 것을 확인(Kim *et al.*, 2017)한 바 있다. 즉, 우리나라 개별 관측소의 기온은 한반도에 영향을 주는 대규모 현상의 영향을 받아 관측 자료의 주성분이 공진성(coherence)을 가지며 상위 수 개(≤ 3)의 성분으로 대부분의 변동성($\geq 90\%$)을 설명할 수 있는 한편, 지형 등의 국지적인 특성도 가지고 있다는 것을 보인 바 있다. 또한, ESPreSSO의 잠재 예측 인자(기반 역학 모형에서 표적 변수의 예측에 사용 가능할 것으로 판단되는, 물리/역학적 근거와 예측성을 가진 인자)를 선택하는 전문가가 기후 현상과 기반 역학 모형의 특성, 그리고 ESPreSSO의 세부 사항에 관한 지식과 경험을 가지고 계절 전망의 예측성을 향상시킬 수 있다는 것을 확인하였다.

새로 개발된 ESPreSSO^{v2.1}은 전체적으로 ESPreSSO^{v1.0}의 구조를 유지하고 있어 모형의 구성 및 구동 단계는 기존의 버전과 동일하다. ESPreSSO의 구성 및 구동 흐름을 되짚어보면 다음과 같다.

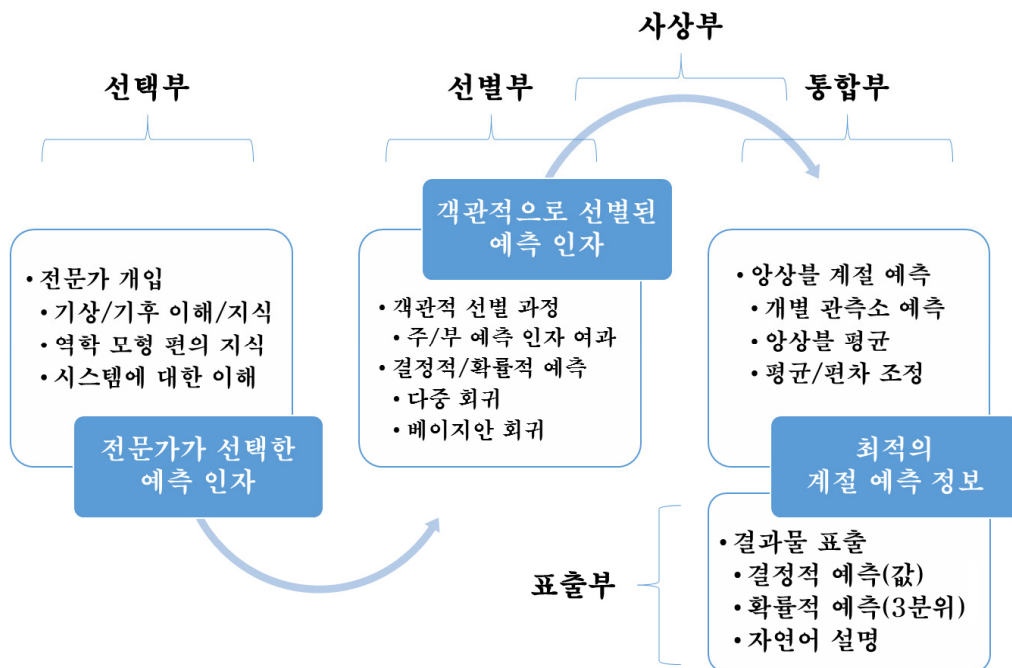


그림 1. ESPreSSO의 구성 및 구동 흐름도, Kim *et al.* (2018)에서 발췌

ESPreSSO^{v2.1}과 이전 버전의 기본 틀은 대체로 동일하나, 전문가 선택부를 제외한 선별부, 사상부, 통합부, 그리고 표출부는 새로운 기능이 추가되었거나 기능이 개선되었다. 특히, ESPreSSO^{v1.0}의 시범 운용을 통하여 높은 예측 성능이 확인된 바, ESPreSSO의 기능에 대한 추가적인 요청을 받을 수 있었는데, 구체적으로는 우리나라 세부 지역에 대한 권역별 예보, 확률 예측 기능, 그리고 변수의 확장(예측 선행 시간 확장 및 강수 예측)에 대한 요구가 있었고, 이러한 요구에 대응하여 ESPreSSO^{v2.1}은 확률 예측 및 1-3 개월 선행 예측뿐만 아니라 우리나라 10개 권역별 상세 예보 기능을 포함할 수 있도록 개발하였다. ESPreSSO^{v2.1}에서 추가된 새로운 기능과 개선 핵심 사항을 간단히 정리하면 다음과 같다.

What's new in ESPreSSO^{v2.1}

1. Completely re-written source code to enhance generality
 - from GrADS + FORTRAN to python3
 - New graphical interface: 'prettified' from GrADS style to python Matplotlib + GIS style
2. Regional anomaly forecast (v2.0)
 - In response to KMA's request, ESPreSSO now provides 10 provincial forecasts as a map
3. Probabilistic forecast (v2.1)
 - Probabilistic tercile forecast based on deterministic MME and fuzzy relationship
4. Improved automation
 - Fully operational at the background without human intervention (optional)
 - > Automatically generates new seasonal prediction when available, and
 - > Automatically sends outlook to subscribers (Sommelier + Baristas + Coffeeholics)
 - Updated KMA interface
5. Extended forecast up to 3-month lead time
 - Up to three-month lead temperature forecast
6. Easier maintenance
 - New interface to check status of potential predictor selection
 - Interim result tables are exported in ASCII form

ESPreSSO^{v2.1}은 기존의 GrADS + FORTRAN의 조합에서 벗어나 상대적으로 다양한 사용자층을 확보하고있는 언어인 python을 기반으로 작성하였으며, 단계별로 모듈화하여 유지/보수 및 기능 개선이 용이하도록 구성하였다. 사용자의 요청에 따라 우리나라 10개 권역별로 상세한 예측값을 제공하고 있으며, 삼분위 확률 예측 기능이 추가되었다. 자동화 방안을 개선하여, ESPreSSO가 새로운 계절 예측 시점이 되면 자동으로 기반 역학 모형의 예측값 생산 여부를 체크하고 새로운 예보를 생산한 후 사용자에게 이메일을 통해서 보고할 수 있도록 하였다. 또한, 예측의 범위를 기존 1개월 선행 예보에서 3개월 선행까지 확대하였다. 그 밖에도, 다양한 응용 분야에서 ESPreSSO의 결과에 쉽게 접근할 수 있도록 내부 인터페이스를 정렬하는 등 사용자 편의성을 획기적으로 개선하였다.

2. ESPreSSO^{v2.1}

2.1 프로그래밍 언어의 변경(포팅)

ESPreSSO^{v1.0}은 대기-해양 과학 분야에서 종종 사용되는 스크립트 언어인 GrADS와 효율적인 계산을 위한 과학 계산 전문 언어인 FORTRAN으로 작성되었다. 따라서 기상/기후 이외의 사용자/관리자가 기후 예측 정보가 생산되는 과정을 이해하고 관리/운영하는 데 어려움이 따랐다. 이를 극복하고자 이번 연구에서는 ESPreSSO의 전문가 선택부를 제외한 모든 부분을 보다 다양한 사용자층을 확보하고 있는 python으로 포팅하는 작업을 수행하였다. 이를 통해 ESPreSSO를 다양한 플랫폼에 쉽게 이식하고 사용자의 요청에 따라 예측 계산 모듈을 간단히 추가할 수 있는 환경이 조성되었으며, 다양한 응용 분야 사용자들에게 보다 친숙한 인터페이스로 기후 예측의 상세한 정보를 제공할 수 있게 되었다. 또한, 관측 자료와 예측 정보는 모두 데이터 클래스를 통하여 관리하여 관리자가 각 단계에서 일어나는 연산 과정을 쉽게 이해할 수 있도록 구성하였다. ESPreSSO^{v2.1}의 내부 모듈과 인터페이스는 [APPENDIX II]에 간략히 정리하였다.

포팅 과정에서 프로그램의 구조와 계산 알고리즘은 대부분 기존 ESPreSSO^{v1.0}와 동일한 방식을 채택하였으나, 세부적인 사항에서 몇 가지 기능 개선이 있었다. 먼저, 예측을 수행하는 회귀 모형을 기존의 결정론적 다중 회귀 모형에서 확률론적 베이지안 능형 회귀로 변경(후술)하였고, 과거 ESPreSSO^{v1.0} 버전에서 오차 보정을 위해 임의적으로 실시했던 평균/편차의 보정 작업을 제거하여 객관성을 제고하였다. 또한, 전문가가 선택한 잠재 예측 인자 정보에서 인자 값을 계산하는 과정에서, ESPreSSO^{v1.0}에서는 GrADS의 영역평균 기능을 이용하였으나 ESPreSSO^{v2.1}에서는 위도 가중치뿐만 아니라 영역의 가장자리 격자가 포함하는 영역의 비율을 제거하여 정확도를 높였다. 또한, 자연어 표시 부분에서 주요 예측 인자가 해당 예측에 유의한 변수로 작용하였는지 표현하는 부분에서 표시 여부를 예측 인자 0.75-표준편차에서 1.0-표준 편차로 상향 조정하여 내용의 변잡함을 줄이는 등 모형의 객관성과 정보 가독성을 높이는 데 주력하였다.

2.2 권역별 예보 기능

시험 운용 기간인 2016-2017년 두 해 동안 ESPreSSO^{v1.0}은 뛰어난 수준의 예측 성능을 보여주어 사용자로부터 새로운 기능에 대한 다양한 요청을 받게 되었다. 그 가운데 우리나라 전체에 대한 예보뿐만 아니라 지역별 예보가 필요하다는 요청이 있었고, 이에 따라 기상청에서 2010년부터 제공하는 권역별 장기/계절 예보에 맞춰 ESPreSSO^{v2.1}에서도 우리나라 10개 권역(서울/인천/경기도, 부산/울산/경상남도, 대구/경상북도, 광주/전라남도, 전라북도, 대전/세종/충청남도, 충청북도, 강원도 영동, 강원도 영서, 제주도)에 대한 권역별 상세 예보를 제공하는 기능을 추가하였다. 지리정보는 강원도 영동과 영서 지방을 나누어 다루기 위하여, 시군구 단위의 행정구역을 병합하여 사용하였다. 권역별 상세 예보는 삼분위 확률 예보값이 지도 형태로 제공된다. 권역별 예보의 생산 방법은 ESPreSSO의 최종 예보 생산 방법과 동일하나, 각 권역에 포함된 관측소별로 개별 예측된 자료만을 앙상블하여 계산한다. 그러나 ESPreSSO의 예측성이 균질/비균질한 자료의 앙상블(통합부)을 통해 얻어지는 바, 권역별 예보의 경우 앙상블 멤버의 숫자가 적어지는 측면이 있어 ESPreSSO의 기본 목적에 맞지 않는 경우가 있으므로 권역별 예측성을 검증하기 전까지는 참고 수준으로만 이용하기를 권장한다.

2.3 확률 삼분위 예보 기능

이전 버전인 ESPreSSO^{v1.0}는 기반 역학 모형에서 계산된 예측 인자와 관측 값을 이용한 선형다중회귀 결정 예측 값을 제공하였다. ESPreSSO^{v2.1}은 기후 예측의 개념에 합치하고 실제 현업 예보에 사용할 수 있도록 확률 예측 값을 삼분위로 제공하도록 설계하였다. 확률 예측의 예측 인자(predictor)는 이전 버전과 동일하여, 각 관측소에서 선별된 주/부 인자를 기반 역학 모형의 단순 합성 자료(Simple Composite Method, SCM)로부터 추출하였다. 이 예측 인자(predictor, X)와 예측할 변수(predictand, y)가 갖는 관계를 베이즈안 능형회귀(Bayesian ridge regression) 분석을 통하여 추정하였다. 일반적인 선형회귀는 예측할 변수와 예측된 값의 잔차 제곱합($\sum(Xw-y)^2$)을 최소화하는 모형 계수(w)를 구하나, 능형회귀에서는 추가적으로 계수의 크기에 패널티를 부여한다($\sum[(Xw-y)^2 + \alpha w^2]$).

베이저안 접근법에서 예측된 변수(y)의 사후 확률 분포는 예측된 값(Xw)에 대해서 가우스 분포를 이루고 있다고 가정($p(y|X, w, \alpha) = N(y|Xw, \alpha)$)하며, 모형 파라미터(w)의 사전 분포는 구형 가우스(spherical Gaussian) 분포($p(w|\lambda) = N(w|0, \lambda^{-1}I_p)$)로 주어진다. 즉, 가우시안(Gaussian) 분포에 대한 켈레 사전(conjugate prior) 분포로 non-informative ($\alpha, \lambda = 10^{-6}$) 감마(Gamma) 분포를 사용하여 주변 로그-우도(marginal log-likelihood)를 극대화하는 방향으로 파라미터를 추정하였다.

실제 예측/검증 기간(2006년 이후)의 모형 파라미터들은 훈련 기간인 1983-2005년의 자료만을 이용하여 추정하였고, 훈련 기간의 예측은 표적 연도를 제외한 one-leave-out 잭나이프 교차 검증 기법을 이용하여 파라미터를 결정하였다. 예측된 분포($\hat{y} \sim N(\mu, \sigma^2)$)로부터 삼분위를 결정하였는데, 삼분위 결정값의 계산을 위해서는 모형의 현업적 안정성을 위하여 훈련 기간의 관측 값을 정규화하여 추정한 삼분위 결정값과 예측 삼분위 결정값을 평균하여 사용하였다. 두 삼분위 결정값 이하, 포함, 이상이 될 확률을 각각 평년 이하(below normal, BN), 평년 비슷(near normal, NN), 평년 이상(above normal, AN)로 구한 후, 5% 단위로 산입하여 최종 삼분위 확률값을 구하였다. 대표 삼분위 예보 범주를 선정하기 위하여 최대 확률값의 범주를 선택하되, 5% 단위로 반올림한 확률 수치가 동일하면 두 범주를 모두 선택(BN/NN, NN/AN, BN/AN)하였고, 각 범주의 확률 차이의 최대값이 5% 이하면 동일 확률(equal change, EQ)로 선언하였다. 예측값이 존재하지 않는 경우 Missing으로 처리하였다. 다중 범주의 선택은 ESPreSSO^{v2.1}의 운용 기간동안 민감도 검사를 거쳐 점진적으로 현업 예보의 기준에 맞춰가야할 것으로 보인다. 한편, 활용성을 위하여 결정론적인 예측 값도 자연어를 통하여 함께 제공하고 있다.

2.4 예측 선행 기간의 확장

ESPreSSO^{v1.0}의 성공적인 1개월 선행 계절 기온 예보를 기반으로 이번 연구에서는 예측 선행 기간을 3개월까지 확장하였다. 사용된 방법론은 1개월 선행 예측과 동일하다. 2, 3개월 선행 예측의 성능은 ESPreSSO^{v2.1}의 예측 결과가 축적된 이후 검증 기간의 결과와 취합하여 월별 추이를 분석할 계획이다.

2.5 자동화 기능 및 인터페이스 개선

ESPreSSO^{v2.1}은 실행시 현재 날짜에 맞는 기반 역학 모형의 자료를 검색하고, 이것이 존재할 경우 최신 계절 예보를 생산해 등록된 구독자들에게 이메일로 결과를 전송한다. 관측 자료 업데이트, 관측/예측 데이터베이스 업데이트, 이메일 수신 기능은 옵션을 통하여 쉽게 제어할 수 있도록 구성하였다. 기반 역학 모형 자료 및 관측 자료의 위치, 훈련/검증 기간의 정의, 변수 리스트, 결과값의 저장 위치 등은 환경 변수 목록을 통하여 관리할 수 있도록 네임리스트 파일로 관리하고 있다. 관리자를 위하여 ESPreSSO^{v1.0}의 GrADS 출력값 및 바이너리 결과값을 ESPreSSO^{v2.1}로 변환하는 변환 프로그램, 자료 업데이트만 빠르게 처리하는 준비 프로그램, 그리고 ESPreSSO 예측 인자의 선정 상태를 표출하는 보고 프로그램도 부차적으로 제공하고 있다.

3. SCIENCE behind the ESPreSSO (December)

ESPreSSO는 전문가의 기후현상과 기반 역학 모형의 특성에 대한 이해를 통해 예측성을 향상시킨다. 실제로 전문가는 잠재예측인자를 선정할 때 실제 기후 현상과 기반 역학 모형의 특성을 고려하는데, 3장에서는 12월 한국 기온과 관계된 기후 현상을 바탕으로 실제 잠재예측인자를 선정하는 과정의 실례를 간단히 소개한다. 자세한 내용은 겨울철의 다른 달과 함께 논문을 통하여 보고할 예정이다.

그림 2는 12월 실제 한반도 기온의 변동성과 관련된 기후 현상이다. 이는 동시상관계수를 그린 것으로, 한반도 기온의 변동성에 영향을 미치는 각 변수의 편차장으로 이해할 수 있다. 특징적으로 보이는 것은, 한반도를 중심으로 한 동서 해면기압장의 차이이다. 12월은 기후학적으로 대륙 고기압이 발달하여 한반도에는 북서풍이 주풍으로 영향을 미치는데, 상관계수 지도(그림 2)에서 한반도 기온이 높은 때는 해면기압장의 동서 차이로 한반도에는 상대적으로 온난, 습윤한 남풍이 유입된다. 이로 인해 한반도 주변 동아시아 전반적으로 같은 기온의 변동성을 보이고, 강수가 상대적으로 많이 발생한다. 또한, 중위도 동아시아 변동성은 열대 지역 해수면 온도(SST) 강제력의 영향을 받는다. 12월 한반도 기온이 엘니뇨-남방진동(ENSO)과 연관되어 있다는 것은 이미 선행연구(Son *et al.*, 2014)에서 분석하였다. ENSO에 따라 아열대 북서태평양 SST가 변동하면, SST forcing은 Gill-type Rossby wave response를 통해 중위도 순환장에 영향을 끼친다. 그림 2에서도 이를 확인할 수 있는데, 열대 태평양에서 나타나는 SST의 동서 대비와 중위도 태평양에서 나타나는 광범위한 고기압성 순환이 그것이다.

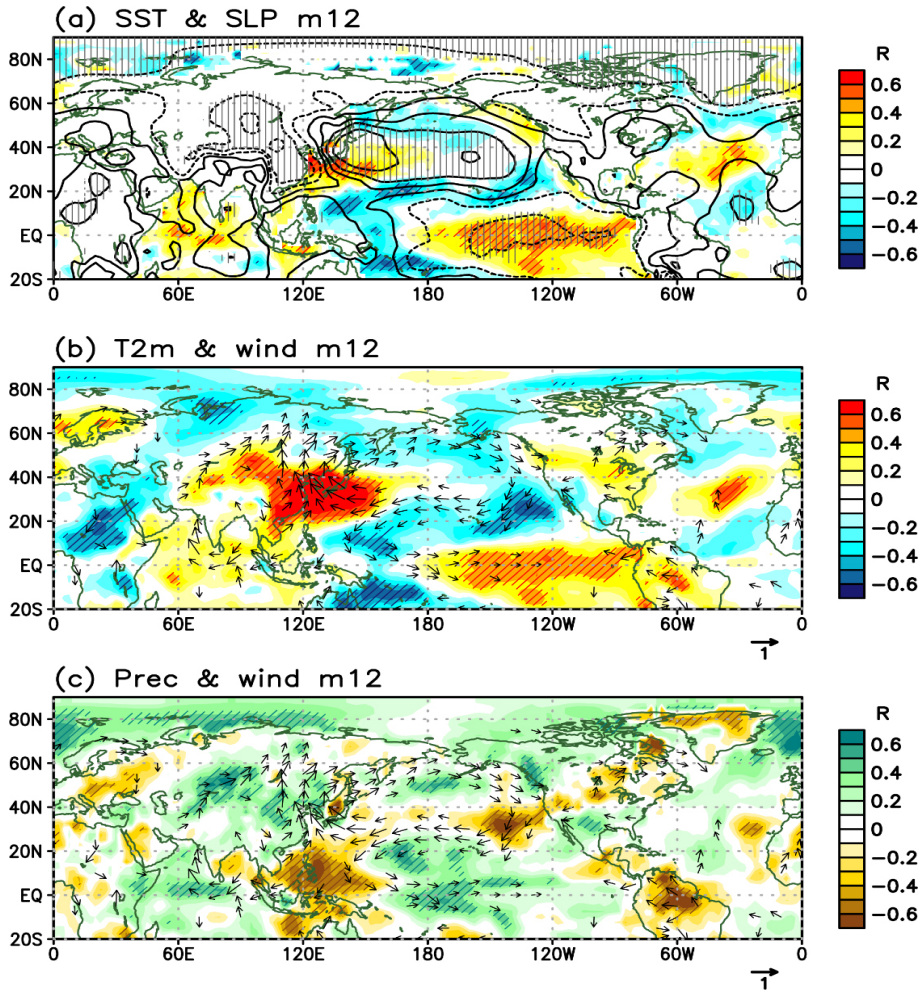


그림 2. 12월 한반도 기온 시계열과 재분석 자료의 (a)해수면 온도(채색), 해면기압(실선), (b)2-m 기온(채색), 850 hPa 바람장(벡터), (c)강수 (채색)의 상관계수. 통계적으로 90% 수준에서 유의미한 값은 사선으로 표시함

한반도 기온 변동성과 관련된 여러 기후 인자는 기반 역학 모형에서도 어느 정도 모의된다. 그림 3은 한반도 12월 기온 변동성과 관련된 APCC MME hindcast의 12월 변동성이다. 이 그림들에서 열대 태평양의 El Niño-like pattern을, 이와 관련하여 아열대 북서태평양(필리핀 해) 대류활동 변동에 의한 강수 감소와 한반도로의 남풍 유입을 확인할 수 있다. 또한 한반도를 중심으로 한 중위도 해면기압장의 동서 대비도 역학 모형에서 잘 모의한다.

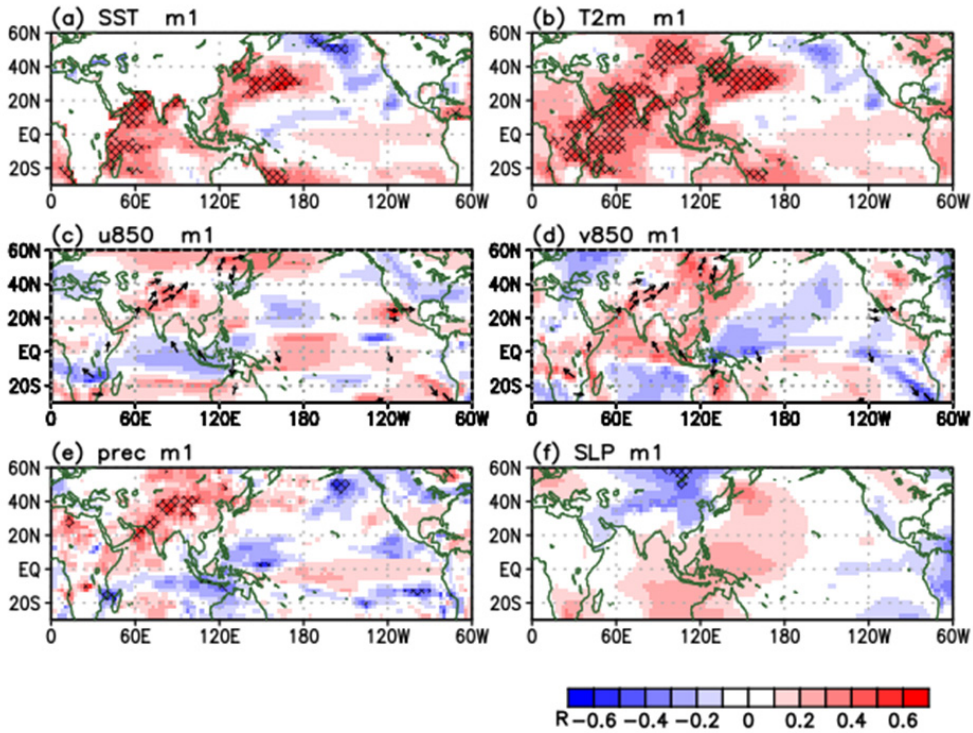


그림 3. 12월 한반도 기온 시계열과 APCC-MME hindcast 자료의 1983-2005 기간동안 (a)해수면 온도, (b)2-m 기온, (c)850 hPa 동-서 방향 바람, (d)850 hPa 남-북 방향 바람, (e)강수, (f)해면기압의 상관계수. 이는 1개월 선행예측 결과로 11월에 발간된 12월 예측장임. 통계적으로 90% 수준에서 유의미한 값은 사선으로 표시함

따라서 12월 실제 기후 현상과 기반 역학 모형의 특성을 고려하여 한반도로 유입되는 남풍, 필리핀해의 대류활동, 그리고 중국 대륙의 해면기압장을 해당 월의 잠재 예측 인자로 선택하였다(표 1).

표 1. 12월 우리나라 기온 예측을 위한 MME 예측장에서 선정된 예측인자 요약

varName	lon1	lon2	lat1	lat2	description
v850	120°E	152.5°E	22.5°N	50°N	in-situ meridional wind
prec	112.5°E	145°E	5°N	25°N	Philippines Sea convection
slp	75°E	125°E	27.5°N	57.5°N	sea level pressure in China(-)

4. RESEARCH RESULTS

4.1 EPreSSO^{v2.1}의 예측 결과

EPreSSO^{v2.1}의 지난 2018년 10월까지 수행한 11월까지의 1개월 선행 예측 결과는 다음 첨부된 그림4와 같다. 추가적인 2-3개월 선행 기온 계절 예측 결과는 [APPENDIX III]에 정리하였다. EPreSSO^{v2.1}는 먼저 우리나라 평균 기온의 확률 예측 결과를 박스-위 스커 시계열로 보여준다. 삼분위 예측 범주의 기준값은 안정성을 위하여 훈련 기간의 관측 삼분위 값과 해당 기간의 예측(hindcast) 삼분위 값의 평균을 사용하였다. 삼분위 기준값은 기상청에서 사용하는 $\pm 0.43\sigma$ 기준과 합치하도록 훈련 기간의 관측/예측 자료가 가우스 분포를 가진다고 가정하고 정규 분포의 삼분위 값을 사용하였다. 박스에는 최대 확률을 가지는 범주에 해당하는 색을 채색함으로써, 해당 시기의 EPreSSO 예측 결과를 쉽게 파악할 수 있도록 하였다. BN, NN, AN (below normal, near normal, above normal)의 세 가지 범주는 각각 파란색 회색, 빨간색으로 채색하였다. 한편, 두 범주가 경쟁하여 우세를 판별하기 힘든 경우, 즉, 두 범주의 차이가 5% 이내인 경우, BN/NN, NN/AN, BN/AN과 같이 두 범주를 모두 표시하고 각각 하늘색, 분홍색, 보라색으로 표시하였다. 세 가지 범주에서 우세한 경우가 없는 경우, 즉, 세 가지 범주의 차이가 최대 5% 이내인 경우, EQ (equal chance)로써 흰 색으로 채색하였다. 검증 기간에서 예측 기간으로 넘어가는 시기에 APCC의 MME 자료 중 누락된 변수들이 있어 안정적인 예측 값이 생산되지 않는 경우가 있는데, 이 경우는 Missing 처리하여 표시하지 않는다. EPreSSO^{v2.1}에서는 확률 예측뿐만 아니라, 권역별 상세 예보를 지도 형태로 제시한다. 권역별 예보의 표시는 확률 예측 결과의 색깔과 동일하게 각 지역을 채색하여 제시하고 있다. 이와 함께 해당 월의 주요 예측 인자와 예측 당시에 의미있게 작용한 인자를 자연어로 제공하여, 현업 예보에 도움이 되도록 하였다. 현재 EPreSSO^{v2.1} 버전에서는 최종 예보문에 삼분위 예측값의 범주 예측 검증이 포함되지 않고 상관계수만을 제시하였는데, 범주 예측 검증 지수(Heidke Skill Score 등)를 포함하는 것은 차년도 검증 진행 과정에서 추가할 예정이다.

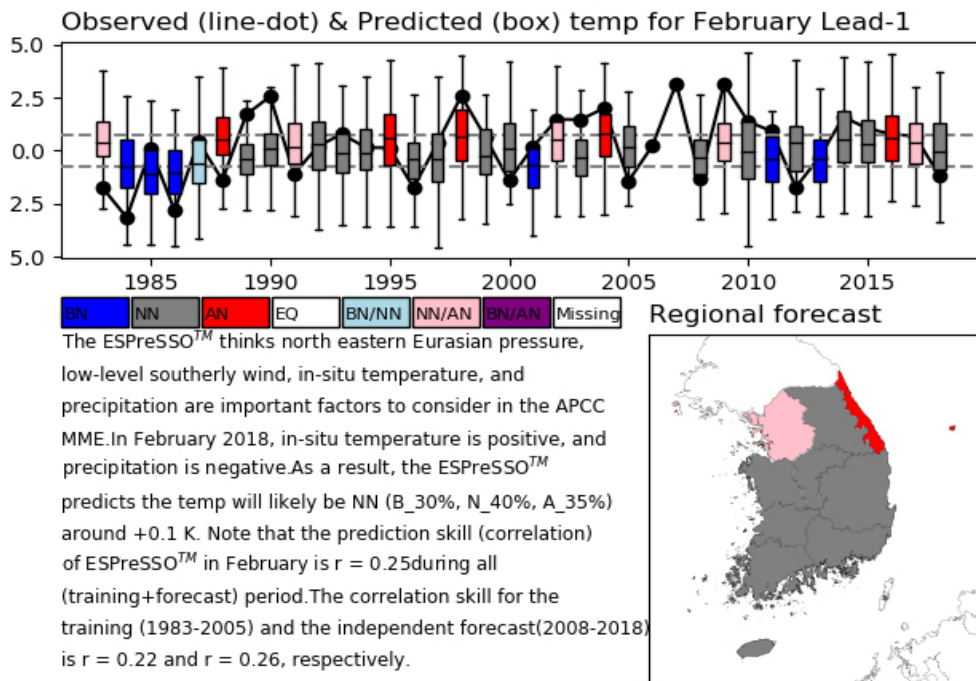
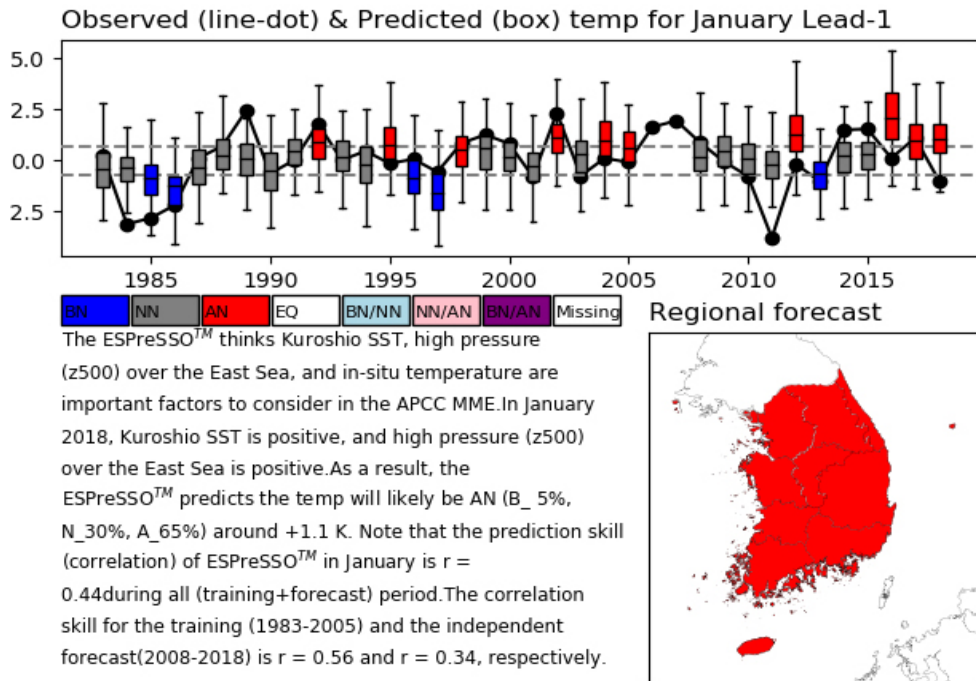


그림 4. ESPreSSo를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조) (계속)

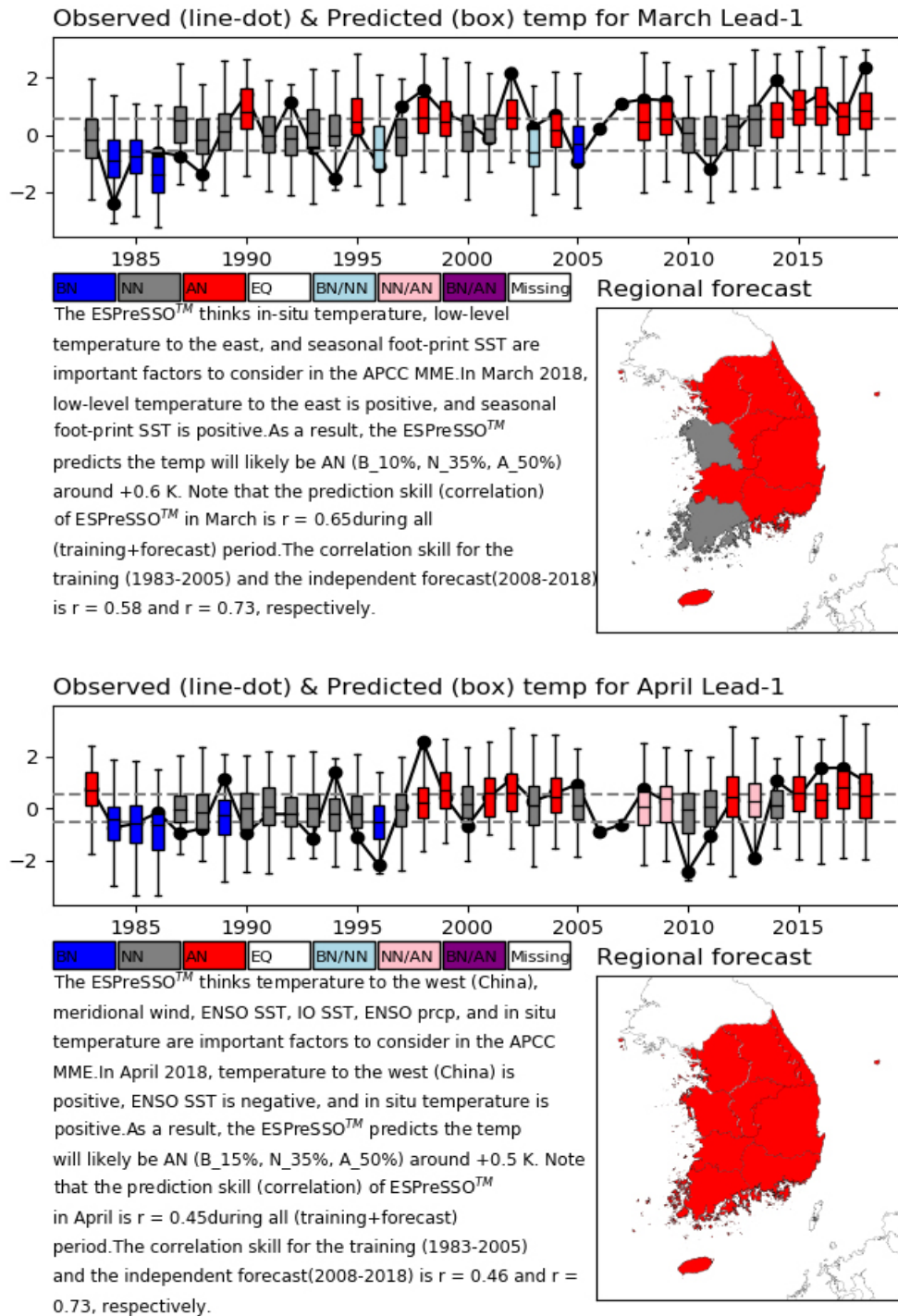


그림 4. ESPreSSo를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조) (계속)

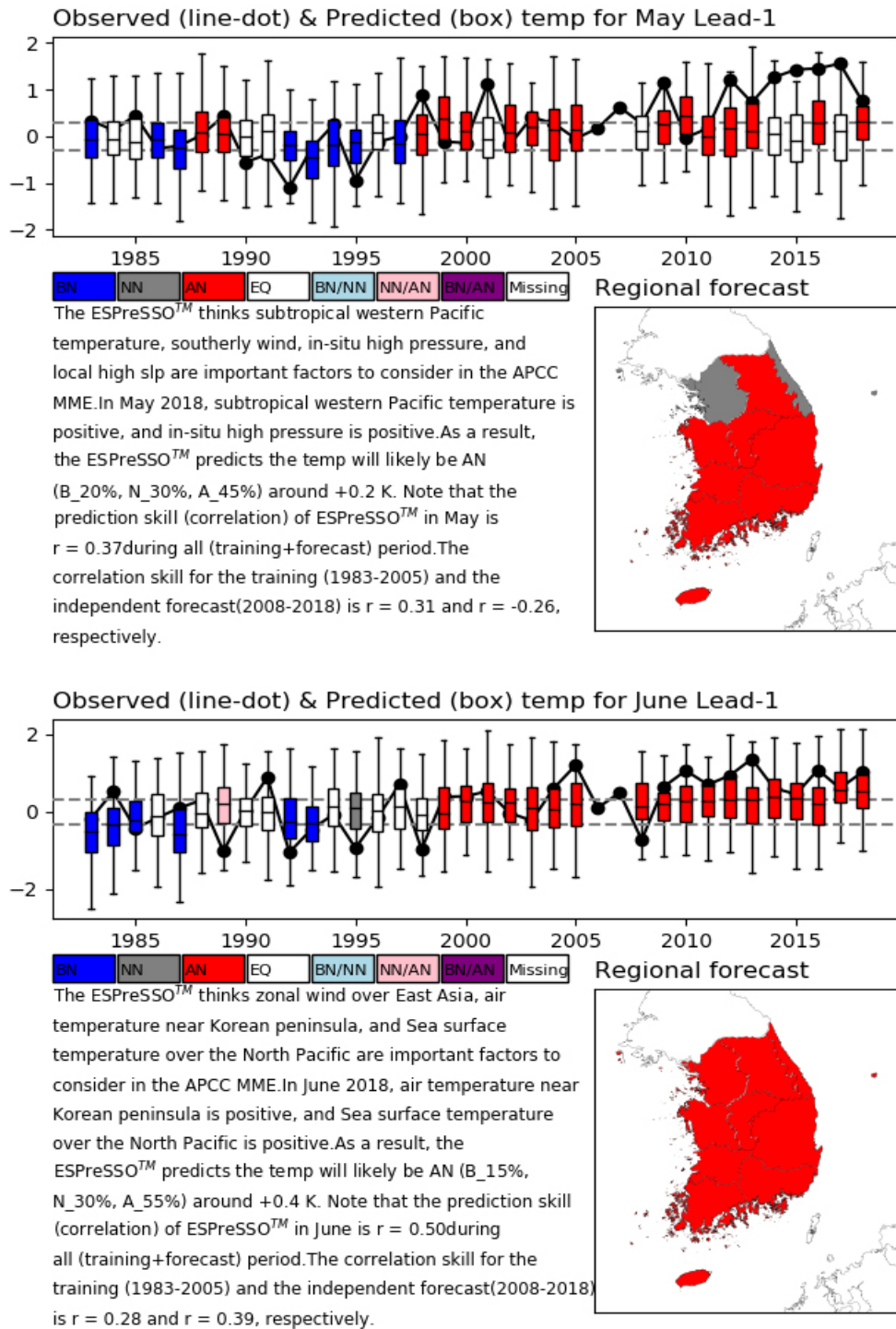


그림 4. ESPreSSo를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조) (계속)

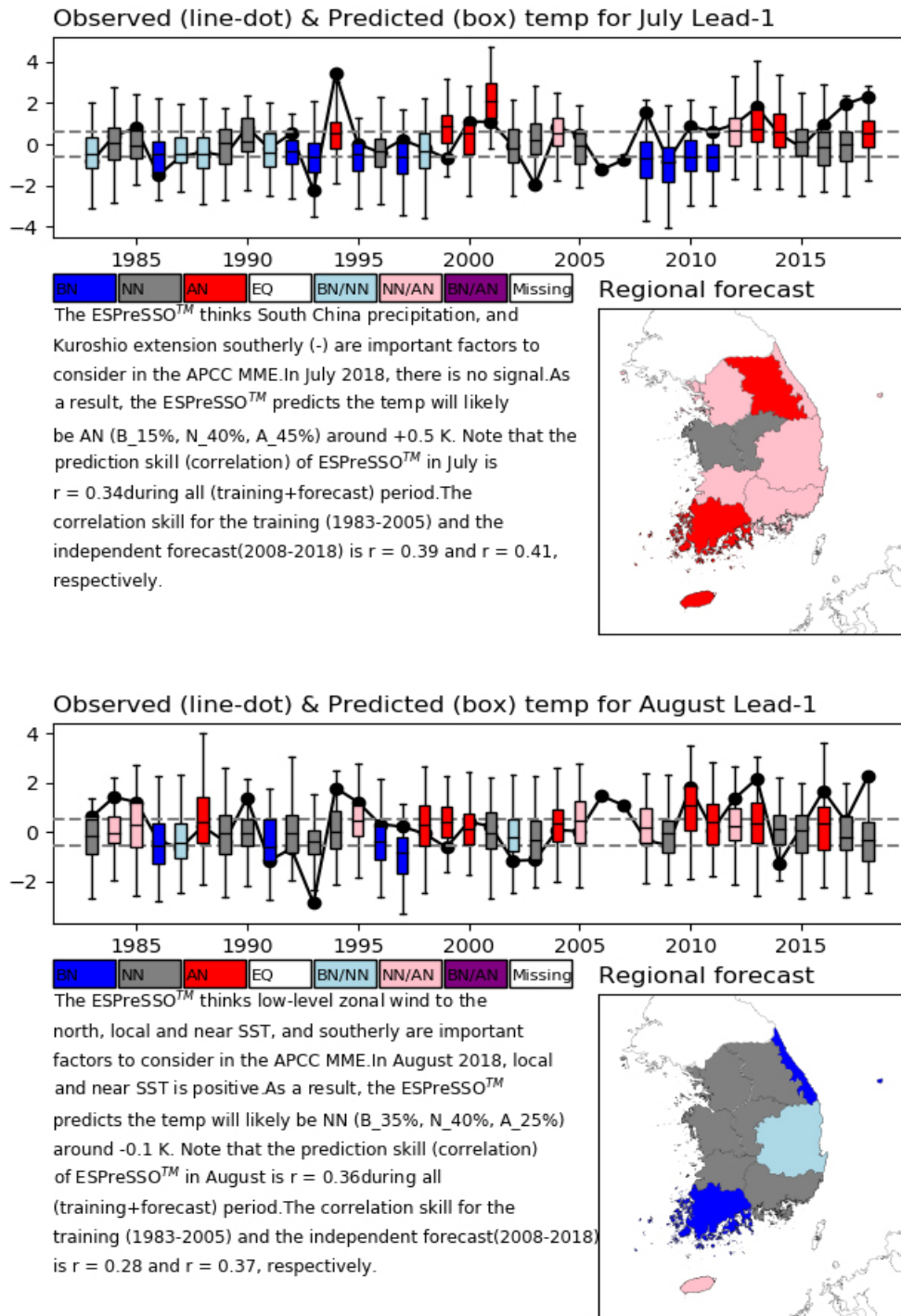


그림 4. ESPreSSo를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조) (계속)

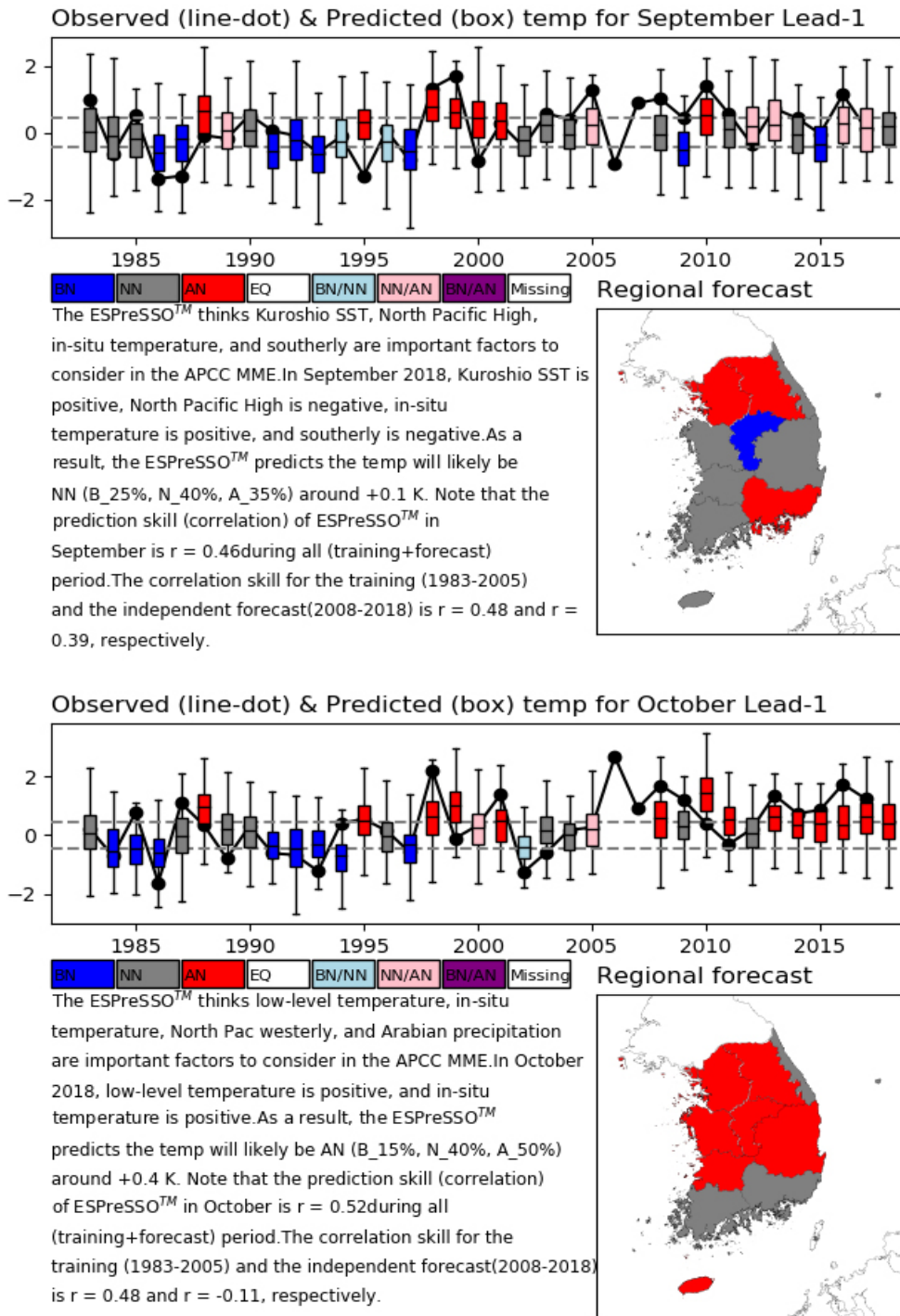


그림 4. ESPreSSo를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조) (계속)

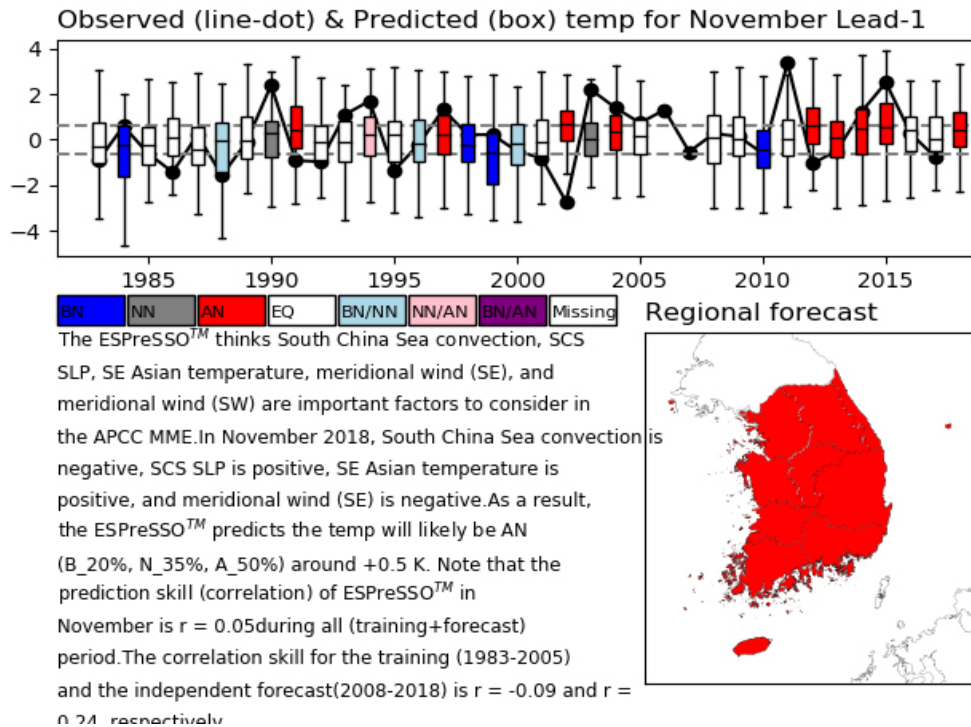


그림 4. ESPreSSO를 통한 2018년 1월부터 11월까지 1개월 선행 시간에서 예측한 우리나라 기온 예측 결과(자세한 내용은 본문 참조)

4.2 ESPreSSO의 예측 성능 평가 및 검증

2018년 10월까지 ESPreSSO^{v2.1}의 예측 결과와 관측 값을 비교해보면 [APPENDIX I]과 같다. ESPreSSO^{v2.1}은 개발이 완료되었으나, 과거 자료에 대한 hindcast 결과 등을 재생산하여 결과를 검증하고 민감도를 검증하지 않은 상태이다. 따라서, 자동으로 발신되는 ESPreSSO의 계절 예측 정보에는 ESPreSSO^{v1.0}의 결과가 첨부되어 제공(현업 예보를 위해서는 결과가 검증된 ESPreSSO^{v1.0}을 따를 것을 추천)하고 있다. 대부분의 경우 두 버전의 ESPreSSO는 비슷한 결과를 보이고 있으나, 세부적인 결과 값에는 차이가 있는 경우가 관찰되었다. 예를 들어, ESPreSSO^{v1.0}에서는 지난 2월을 평년보다 추울 것으로 예상한 반면 ESPreSSO^{v2.1}에서는 평년과 비슷한 기온을 예측하여 차이를 보였다. 일반적으로 ESPreSSO^{v2.1}의 예측 최빈값이 결정론적 예측인 ESPreSSO^{v1.0}의 예측 값에 비해 평균값으로 치우치는 경향이 있는데, 이는 능형회귀의 특성이라 할 수 있겠다. 그러나

종종 두 버전에서 상이한 예측 값을 보이는 경우가 있는데, 이는 2.1장에서 언급한 잠재 예측인자 풀에서 예측 인자를 선정하는 방식의 차이에서 기인한 것으로 보인다. 다만, 상세한 부분은 차년도 연구에서 검증을 진행하는 과정이 필요하다.

4.3 프로젝트의 기타 성과물

상세화 기법을 응용한 한반도 전문가 계절 예측 시스템인 ESPreSSO^{v2.1}은 이메일을 통하여 확률 예보 및 권역별 전망 등을 제공하여 계절 기후 예측 능력 향상에 직접적으로 기여하고 있다. 이전 버전과 마찬가지로 전문가의 지식이 데이터베이스화되어 접근이 용이해졌으며, 전문가의 인적 조합에 영향을 받지 않는 안정적인 예보가 가능하게 되었다. 범용 프로그램 언어로 포팅하여 관리 및 운영의 편의성을 제고하였으며, 다양한 응용 분야의 잠재적 수요에 대응하여 다양한 방식으로 중간 결과값을 보고할 수 있는 기능을 구축하였다.

기본적인 성과물 이외에도, ESPreSSO의 방법론은 대한민국 특허로 등록(특허번호 10-1901311)였고, 기술적인 내용과 예측성에 관한 내용을 학술 논문을 통하여 학계에 보고(Kim *et al.*, 2018)하였다.

특허 제10-1901311호:

다중모형 앙상블과 상세화 기법을 응용한 계절 예측 시스템 및 방법

특허권자: 재단법인 에이팩 기후센터

발명자: 김원무, 김유진, 여새림

학술 논문:

Kim, W., S.-R. Yeo, and Y. Kim (2018), Development of the expert seasonal prediction system: An application for the seasonal outlook in Korea, *Asia-Pac. J. Atmos. Sci.*, **54**(4): 563-573, 10.1007/s13143-018-0052-9.

5. CONCLUDING REMARKS

계속 과제인 [상세화 기법을 응용한 한반도 전문가 계절 예측 시스템 II]는 ESPreSSO의 기술적 개발을 완료하는 것과 예측 선행 시간을 확장하는 것에 초점이 맞추어져 진행하였다. 연구 내용에 기술적인 부분이 많아 보고서에서는 심도 있게 다루지 않았으나, 일부 기술적인 내용과 과학적 배경에 대해서는 특허 문서와 논문을 통해서도 접근이 가능하다. 한편, ESPreSSO^{v2.1}은 기존 계획을 벗어나는 추가적인 요구를 반영하게되어 개발 기간이 늘어나 결과 값에 대한 자세한 검증이 부족한 것으로 판단된다. 차년도 과제에서는 강수 예측 시스템 개발뿐만 아니라, 현업 시스템의 운영에 초점을 맞춰 ESPreSSO가 현업적으로 유용하게 활용될 수 있도록 검증 지표를 개발/제공할 계획이다.

REFERENCES

- Son, H.-Y., J.-Y. Park, J.-S. Kug, J. Yoo, and C.-H. Kim (2014): Winter precipitation variability over Korean Peninsula associated with ENSO, *Clim. Dynam.*, **42**: 3171–3186.
- Kim W., S.-R. Yeo, and Y. Kim (2018): Development of the expert seasonal prediction system: An application of the seasonal outlook in Korea, *Asia-Pac. J. Atmos. Sci.* Online Published
- Yeo S.-R., S.-W. Yeh, Y. Won, H.-S. Jo, and W. Kim (2017): Distinct mechanisms of Korean summer temperature variability during early and late summer, *J. Geophys. Res. Atmos.*, **122**.
- Yeo S.-R., S.-W. Yeh, Y. Kim, and S.-Y. Yim (2018): Monthly climate variation over Korea in relation to the two types of ENSO evolution. *Int. J. Climatol.*, **38**: 811–824.
- Expert seasonal prediction system (ESPreSSO) Patent#**10-1901311**.

Appendix 2 EPreSSO^{v2} core (funcEPreSSO)

EPreSSO (version 2.0)

EPreSSO-KR: Expert Seasonal Prediction System for Seasonal Outlook in Korea.

The EPreSSO is a hybrid tailored seasonal forecast system. It is based on the dynamic-model based Multi-Model Ensemble (MME) forecast provided by APCC, as well as the experts knowledge of climate system and MME forecast system.

Assumptions:

1. The base models should have some level of predictability.
 - Systematic biases, however, are acceptable and overcome by EPreSSO.
 - Operational MME set is assumed to be stable.
2. The target variables should be homogeneous and inhomogeneous at the same time.
 - Being homogeneous implies the target variable is physically and dynamically governed by similar process.
 - Being inhomogeneous introduces diversity in the potential predictor selection.
3. The experts should meet three criteria: climate system, MME system, and EPreSSO
 - Climate system - understand physical/dynamical processes that govern

target climate.

- MME system - have experience in operational characteristics of MME and understand the limitation.
- ESPreSSO *per se* - understand the ESPreSSO work flow, esp. how the predictors are handled.

Processes and modules:

1. Selection module

- Experts select potential predictor sets using their knowledge and experience.

2. Screening module:

- ESPreSSO objectively screens potential predictors based on minimum usefulness.

3. Mapping module:

- Make regression-based statistical mapping (deterministic + probabilistic).

4. Ensemble module:

- Logical ensemble to produce a stable forecast.

5. Display module:

- Exports ESPreSSO results to subscribers including natural language outlook.

References:

1. ESPreSSO:

- Kim W., S.-R. Yeo, and Y. Kim (2018): Development of the expert seasonal prediction system: An application of the seasonal outlook in Korea, Asia-Pac. J. Atmos. Sci.

2. Dynamics:

- Yeo S.-R., S.-W. Yeh, Y. Won, H.-S. Jo, and W. Kim (2017): Distinct mechanisms of Korean summer temperature variability during early and late summer, J. Geophys. Res. Atmos., 122.
- Yeo S.-R., S.-W. Yeh, Y. Kim, and S.-Y. Yim (2018): Monthly climate variation over Korea in relation to the two types of ENSO evolution. Int. J. Climatol., 38: 811-824.

3. Technical reports:

- Kim W., S.-R. Yeo, and Y. Kim (2017 and 2018): Expert Seasonal Prediction System for Seasonal Outlook in Korea (ESPreSSO-KR), APCC.

4. Patent:

- Expert seasonal prediction system (ESPreSSO) Patent#10-1901311.

The ESPreSSO comes with absolutely no guarantee. Please contact the author for more details.

Modules

calendar	sklearn.linear_model	numpy	matplotlib.pyplot
csv	matplotlib	posixpath	shapefile
envESPreSSO	netCDF4	pandas	textwrap

Classes

[builtins.object](#)
[dataESP](#)
[dataObs](#)
[dataPrd](#)
[html](#)
[infoESP](#)

class **dataESP**([builtins.object](#))

The [dataESP](#) class stores ESPreSSO data as a pandas DataFrame.

Attributes:

mark : index name
 data : pandas DataFrame
 file : base data file in CSV format
 tarVar : target variable (str)
 tarMon : target month (int)
 tarLead : target lead (int)

Methods defined here:

__init__(self, info, tarVar, tarMon=4, tarLead=0, mark='Year', file=None)
 Initialize [dataESP](#) from CSV file if exists.

Updates:

all attributes

col(self, col)

Get list of column data from column value.

get(self, row, col)

Get data value from index row and column values.

row(self, row)

Get list of row data from index row value.

save(self, file)

Save data to CSV file.

class **dataObs**([dataESP](#))

Subclass of [dataESP](#) – for observation data.

Method resolution order:

[dataObs](#)
[dataESP](#)
[builtins.object](#)

Methods defined here:

down(self, info, years=None)

Search missing points in the existing data, and download KMA data from their website.

Updates:

file_dataKMA

load(self, info, und=None)

Search missing points in the existing data, and fill'em from KMA data files.

Updates:

data

Methods inherited from [dataESP](#):

__init__(self, info, tarVar, tarMon=4, tarLead=0, mark='Year', file=None)

Initialize [dataESP](#) from CSV file if exists.

Updates:

all attributes

col(self, col)

Get list of column data from column value.

get(self, row, col)

Get data value from index row and column values.

row(self, row)

Get list of row data from index row value.

save(self, file)

Save data to CSV file.

class **dataPrd**([dataESP](#))

Subclass of [dataESP](#) – for prediction data.

Method resolution order:

[dataPrd](#)

[dataESP](#)

[builtins.object](#)

Methods defined here:

load(self, info, other)

Search missing points in the existing data, and fill'em from the new ESPreSSO results.

Updates:

data

Methods inherited from [dataESP](#):

__init__(self, info, tarVar, tarMon=4, tarLead=0, mark='Year', file=None)

Initialize [dataESP](#) from CSV file if exists.

Updates:

all attributes

col(self, col)

Get list of column data from column value.

get(self, row, col)

Get data value from index row and column values.

row(self, row)

Get list of row data from index row value.

save(self, file)

Save data to CSV file.

class **html**([builtins.object](#))

HTML interface.

Methods defined here:

parse_html(self, table)

parse_url(self, url)

class **infoESP**([builtins.object](#))

The [infoESP](#) class stores ESPreSSO information.

Attributes:

file : file name

mark : index if not assigned <> data.columns[0]

data : an indexed DataFrame <> pd.read_csv(file).set_index(mark)

cols : list of column names <> data.columns.tolist()

rows : list of index names <> data.index.tolist()

Methods defined here:

__init__(self, file, mark=None)

Initialize [infoESP](#) from CSV file.

Updates:

all attributes

col(self, col)

Get list of column data from column value.

get(self, row, col)

Get data value from index row and column values.

row(self, row)

Get list of row data from index row value.

save(self, file=None)

Save data to CSV file.

Functions

aave(data, lat)

Calculate area-weighted average.

```

addDomain(address, domain)
    Add domain after address@ if necessary.
get_Region(dataStn, file='/data16/WonMoo.Kim/DATA/ESPreSSO/infoGeo.csv')
    Generate regional averages.
get_dataMME(tarVar, tarMon, tarLead, tarExp, bYear=None, bLon=None, bLat=None,
und=None, tooBig=900000000.0)
    Get MME prediction data from corresponding experiment files.
get_dataPrd(info, stnID, tarMon, tarLead)
    Generate predictor timeseries from the potential predictor information & MMM
    data.
get_dataWhy(info, tarYear)
    Get important potential predictors and find current status.
moments(data, ref=0)
    Calculate moments.
plt_mapKorea(result, vmax=None)
    Plot regional forecast on a map.
plt_talk(dict, tarVar, tarYear, tarMon, tarLead, prd)
    Plot natural language outlook.
plt_timeSeries(Obs, Prd)
    Plot timeseries of observations and predictions.
report(txt, pp, suf="")
    Report the progress of ESPreSSO modules by %.
sendMail(sub=None, pre="", mid="", suf="", files=None, to=[])
    Send ESPreSSO outlook to subscribers.
skill_Cor(daObs, daPrd, year=None)
    Calculate correlation skill.

```

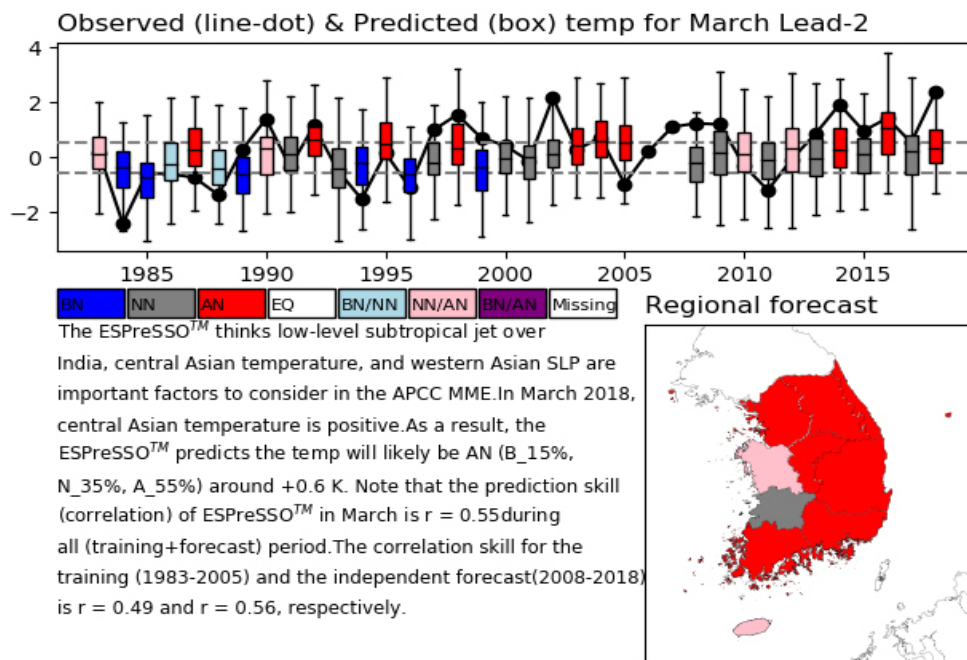
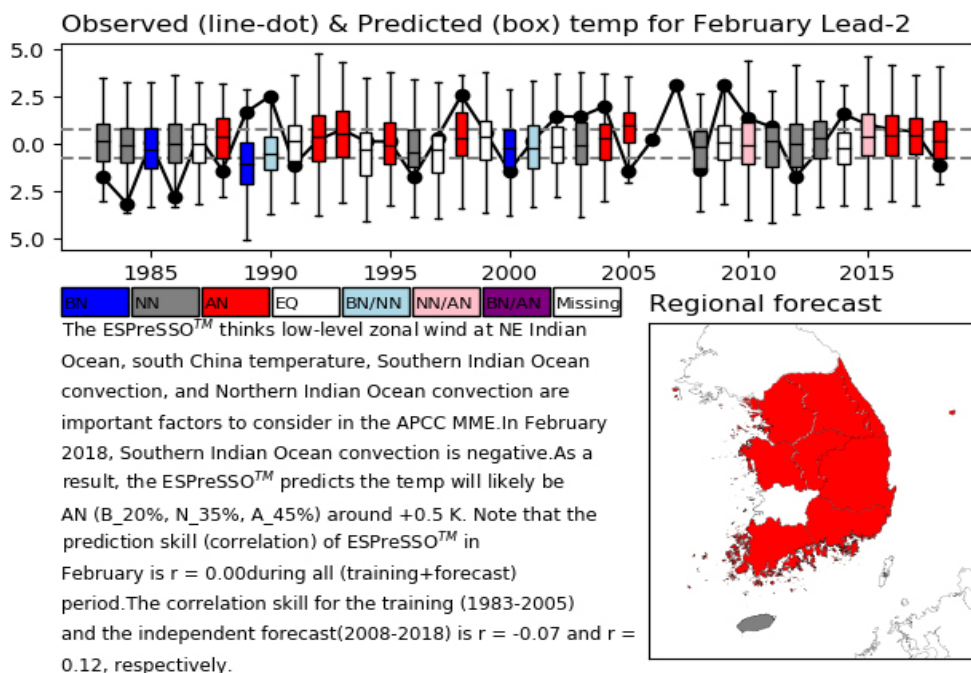
Data

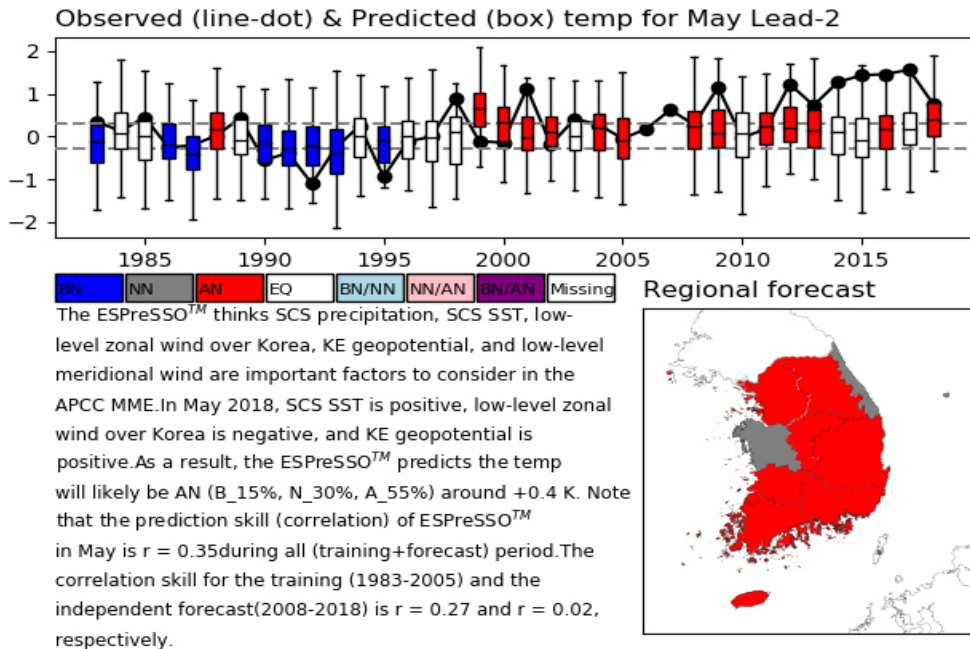
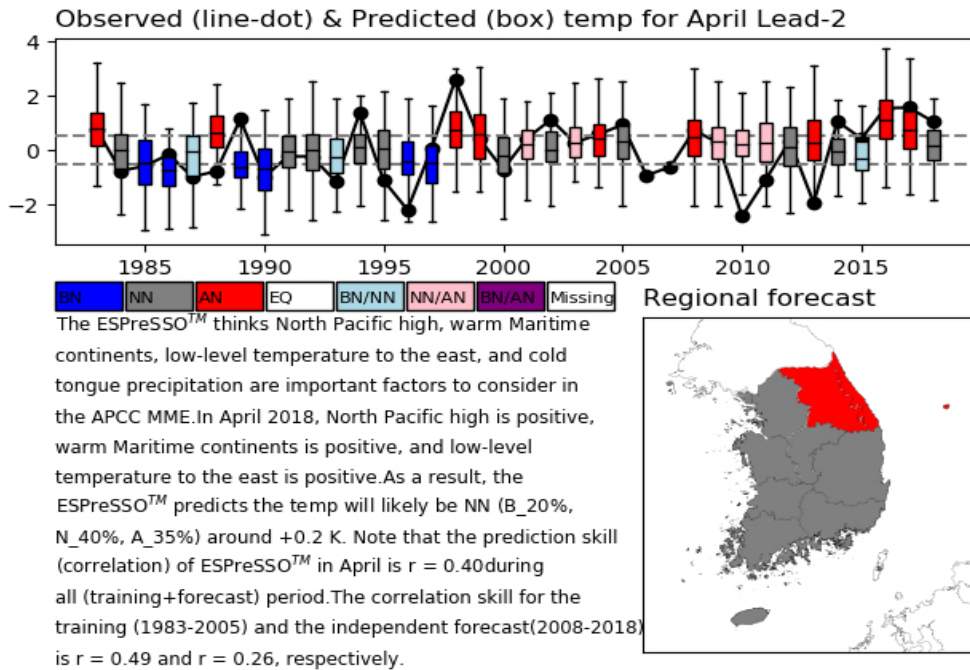
```

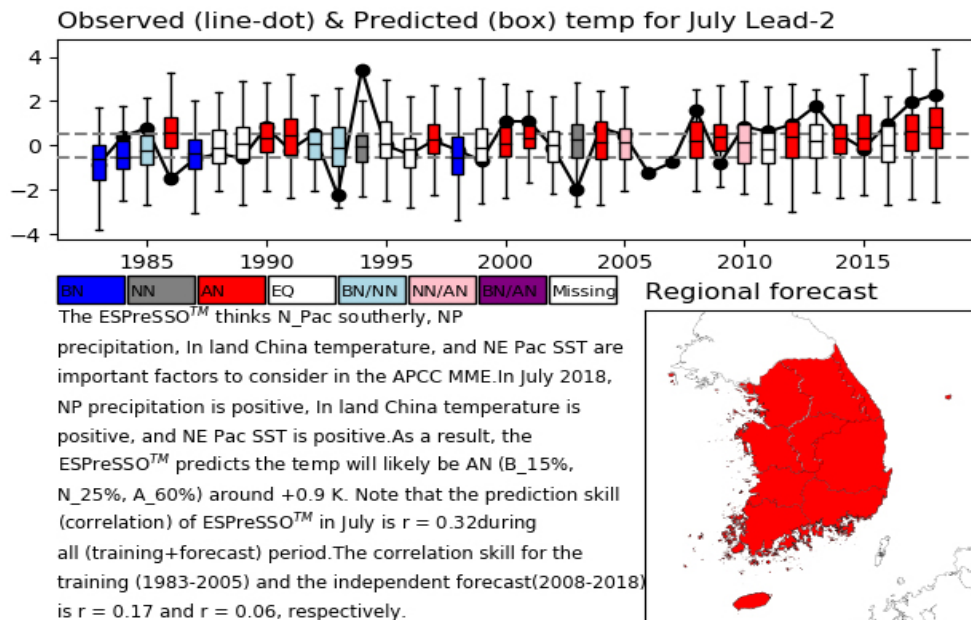
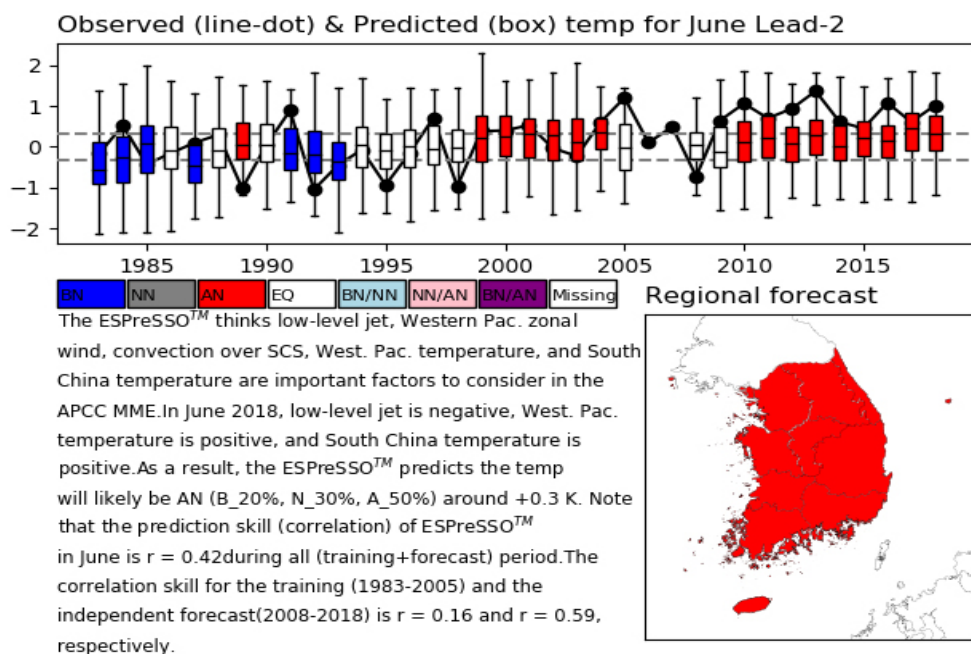
__contact__ = 'WonMoo.Kim@apcc21.org'
__copyright__ = 'Patent#10-1901311, Copyright 2018, Team ESPreSSO'
__note__ = 'This version of ESPreSSO was developed just with... - WonMoo 2018-04-04'
__updated__ = '2018-09-03'
now = datetime.datetime(2018, 4, 26, 14, 6, 49, 999852)

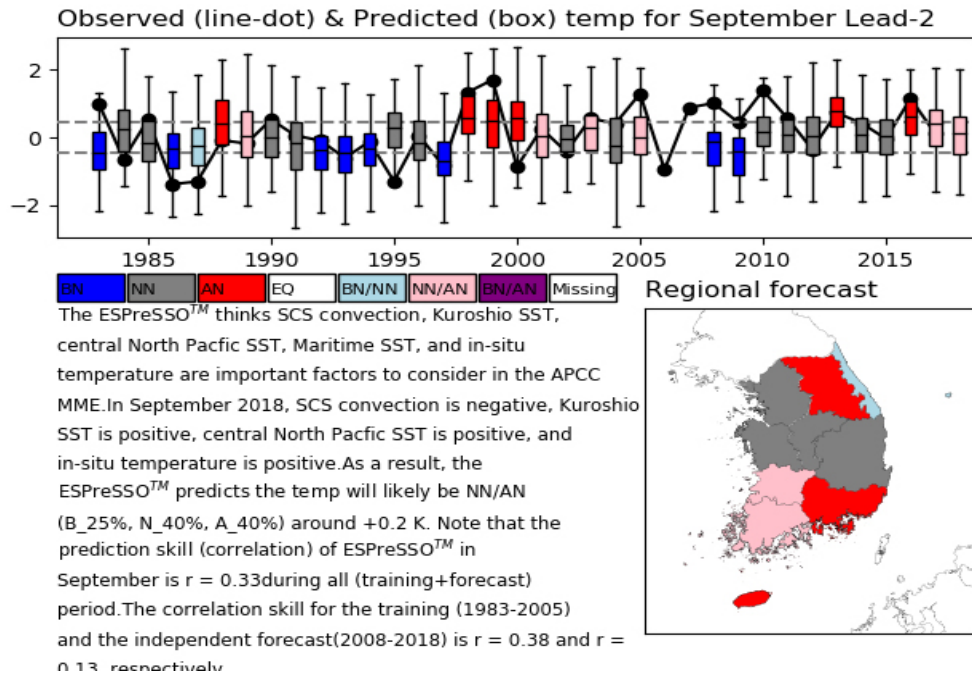
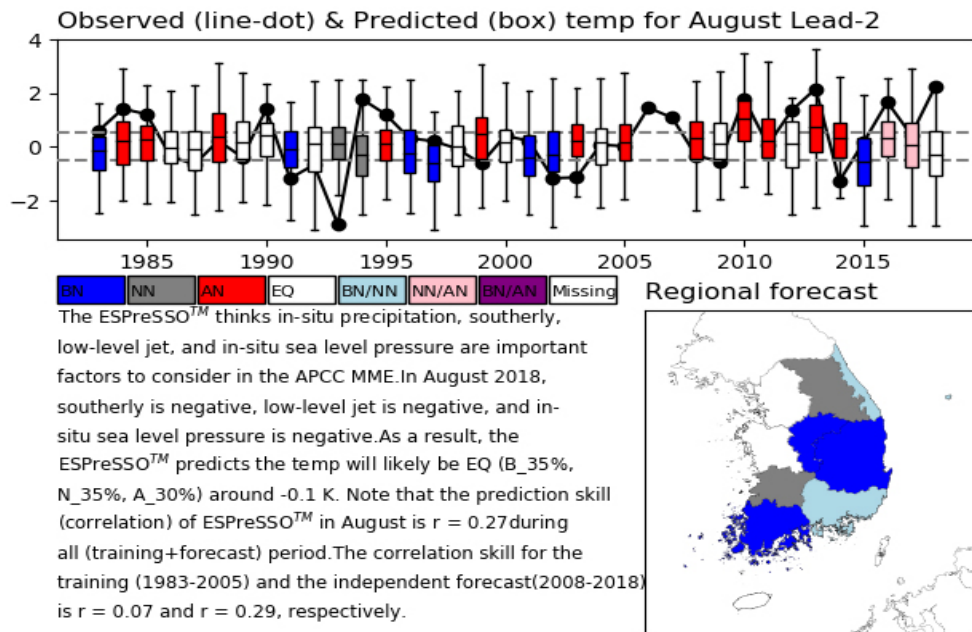
```

Appendix 3 EPreSSO^{v2.1} outlook example (2-3-month lead)

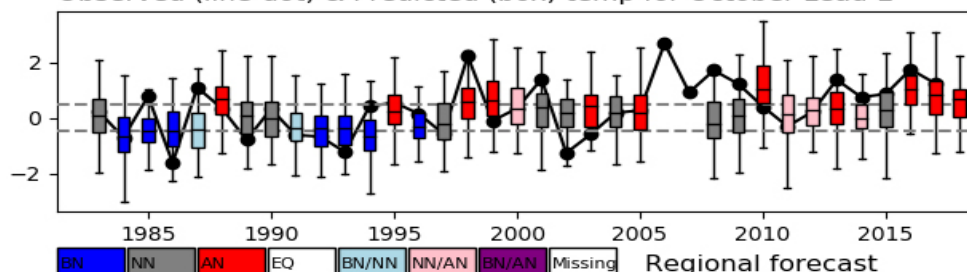








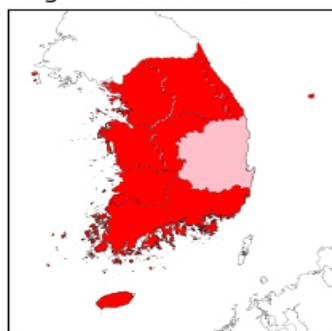
Observed (line-dot) & Predicted (box) temp for October Lead-2



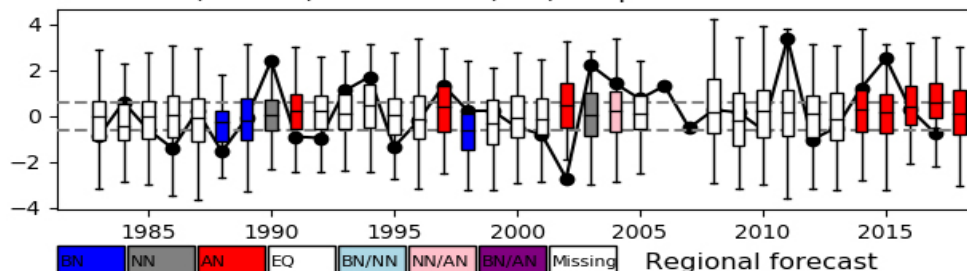
The ESPreSSOTM thinks SCS T850, KE SST, in-situ

temperature, and SCS t2m are important factors to consider in the APCC MME. In October 2018, there is no signal. As a result, the ESPreSSOTM predicts the temp will likely be AN (B_15%, N_40%, A_45%) around +0.4 K. Note that the prediction skill (correlation) of ESPreSSOTM in October is $r = 0.41$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = 0.34$ and $r = 0.02$, respectively.

Regional forecast



Observed (line-dot) & Predicted (box) temp for November Lead-2



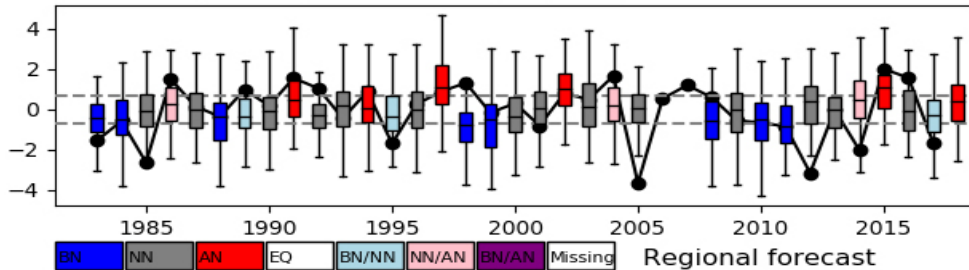
The ESPreSSOTM thinks Maritime SST, temperature in

China, and meridional wind to China are important factors to consider in the APCC MME. In November 2018, there is no signal. As a result, the ESPreSSOTM predicts the temp will likely be AN (B_25%, N_35%, A_40%) around +0.3 K. Note that the prediction skill (correlation) of ESPreSSOTM in November is $r = 0.01$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = -0.08$ and $r = 0.05$, respectively.

Regional forecast

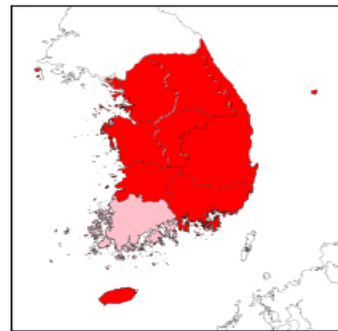


Observed (line-dot) & Predicted (box) temp for December Lead-2

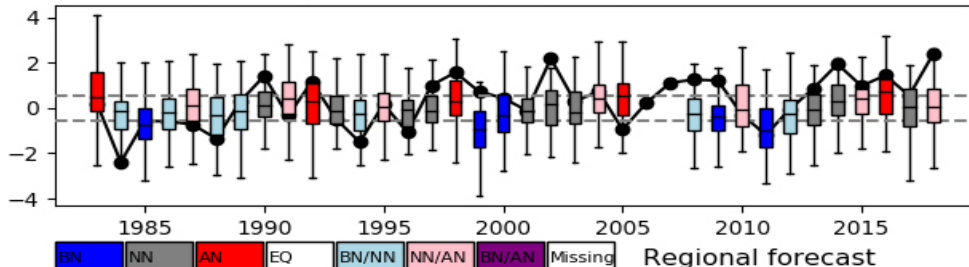


The ESPreSSOTM thinks eq. low-level zonal wind, W. Pac precipitation, low-level meridional wind at Philippines Sea, and SE Asian temperature are important factors to consider in the APCC MME. In December 2018, there is no signal. As a result, the ESPreSSOTM predicts the temp will likely be AN (B_20%, N_35%, A_45%) around +0.5 K. Note that the prediction skill (correlation) of ESPreSSOTM in December is $r = 0.26$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = 0.29$ and $r = 0.19$, respectively.

Regional forecast

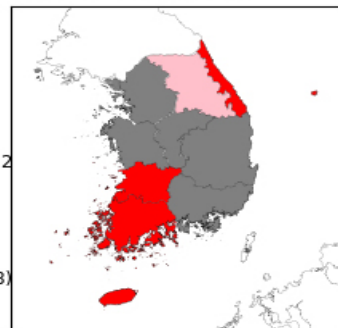


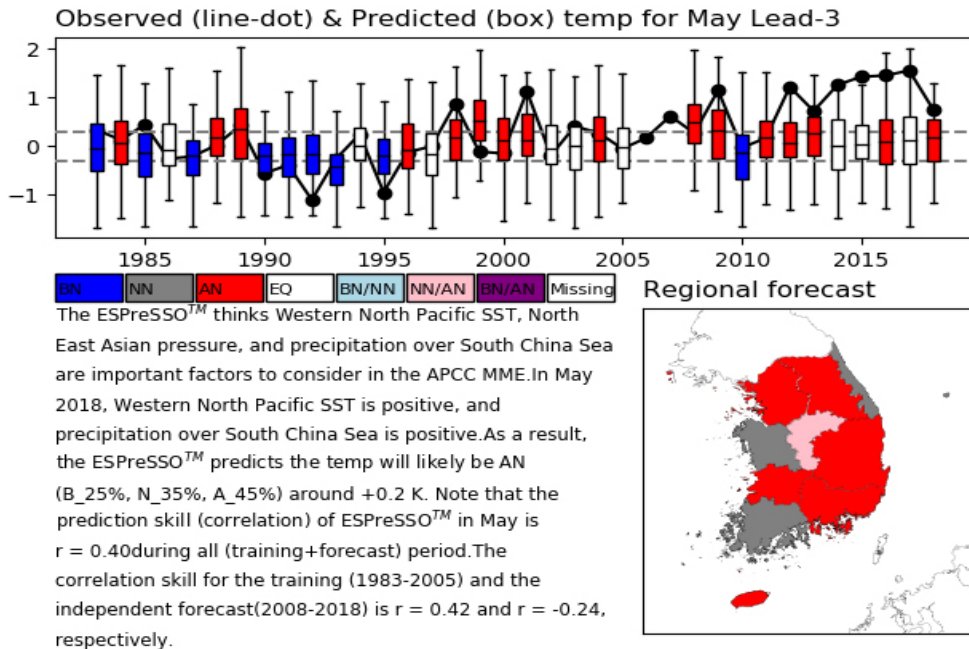
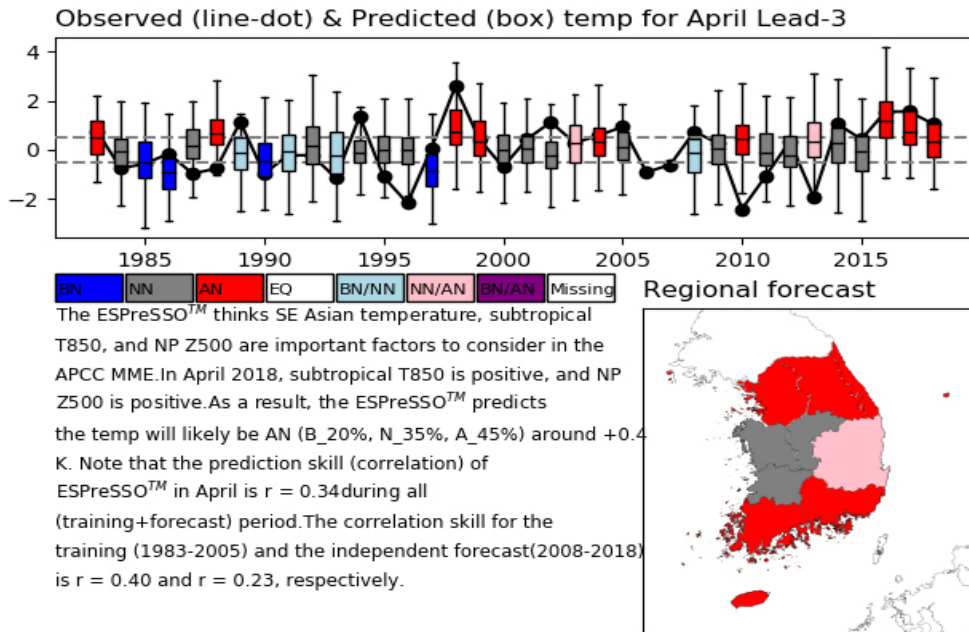
Observed (line-dot) & Predicted (box) temp for March Lead-3

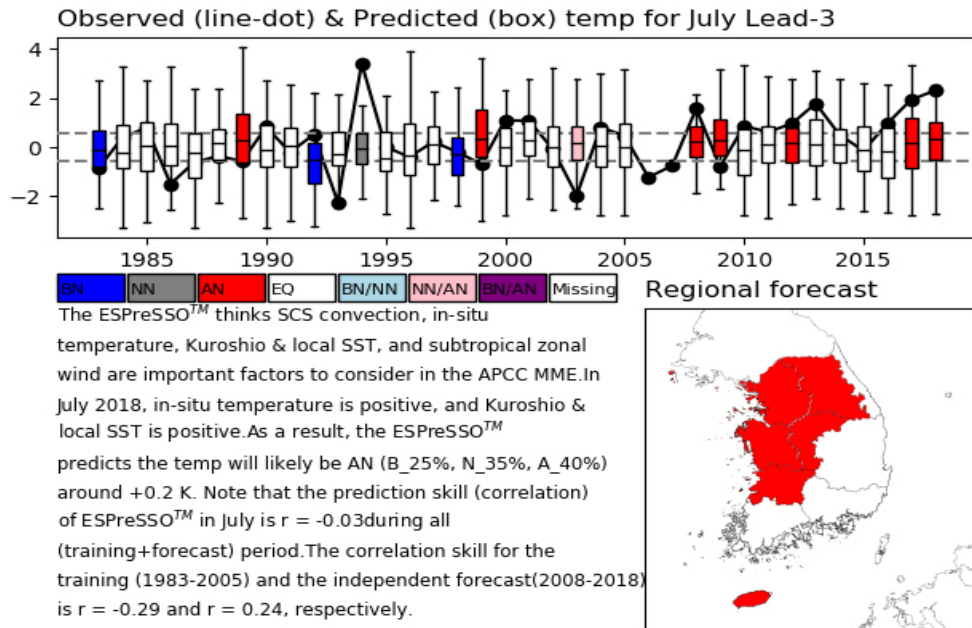
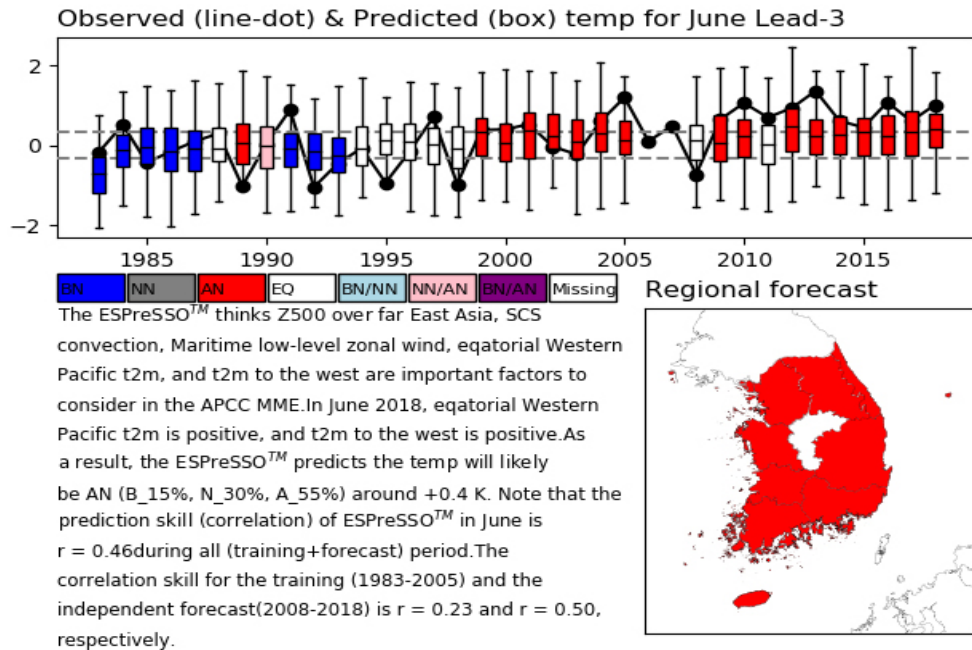


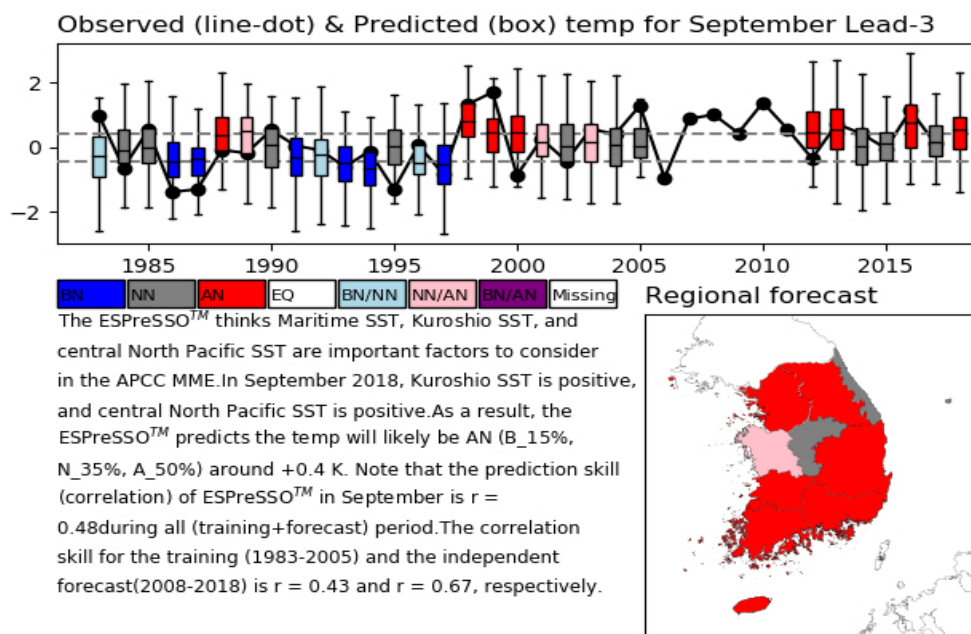
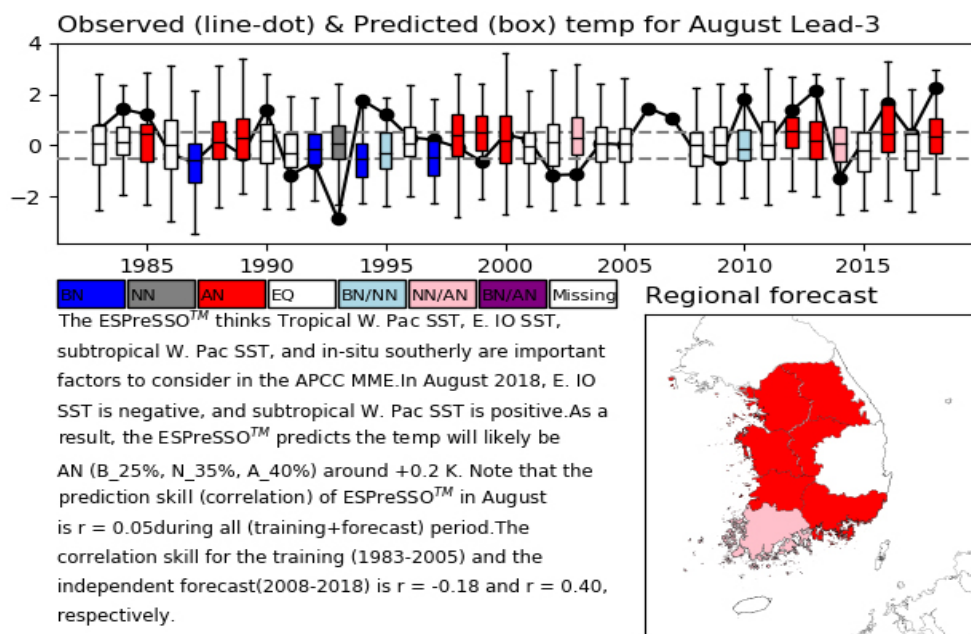
The ESPreSSOTM thinks maritime precipitation, eastern EQ Pacific precipitation, EQ low-level zonal wind W. Pacific, and in-situ temperature are important factors to consider in the APCC MME. In March 2018, in-situ temperature is positive. As a result, the ESPreSSOTM predicts the temp will likely be NN/AN (B_25%, N_40%, A_40%) around +0.2 K. Note that the prediction skill (correlation) of ESPreSSOTM in March is $r = 0.30$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = 0.26$ and $r = 0.49$, respectively.

Regional forecast

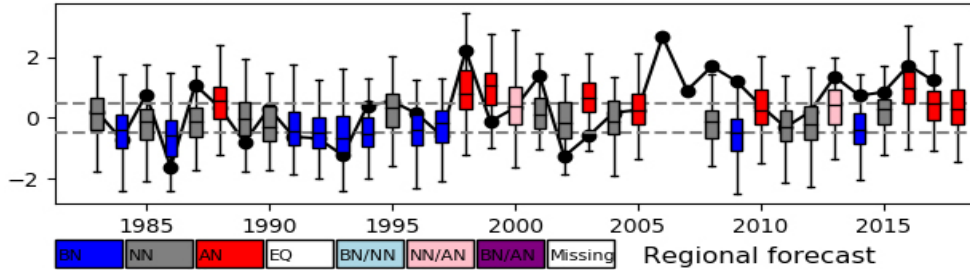








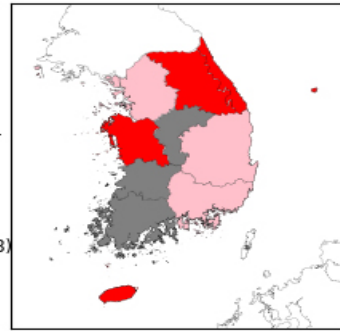
Observed (line-dot) & Predicted (box) temp for October Lead-3



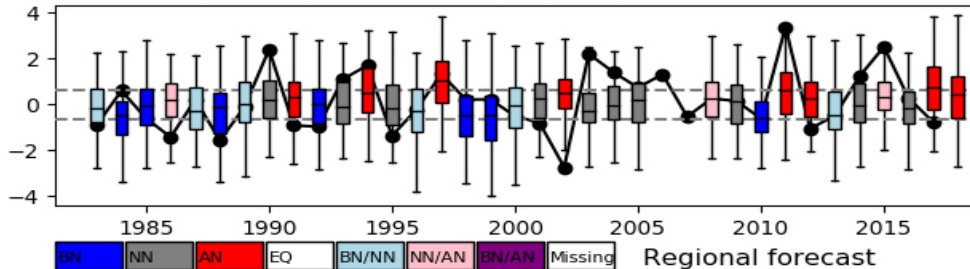
The ESPreSSOTM thinks local SST, local temperature,

China temperature, and near-by precipitation are important factors to consider in the APCC MME. In October 2018, there is no signal. As a result, the ESPreSSOTM predicts the temp will likely be AN (B_15%, N_40%, A_45%) around +0.3 K. Note that the prediction skill (correlation) of ESPreSSOTM in October is $r = 0.48$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = 0.51$ and $r = 0.40$, respectively.

Regional forecast



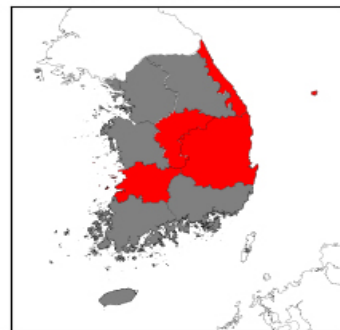
Observed (line-dot) & Predicted (box) temp for November Lead-3

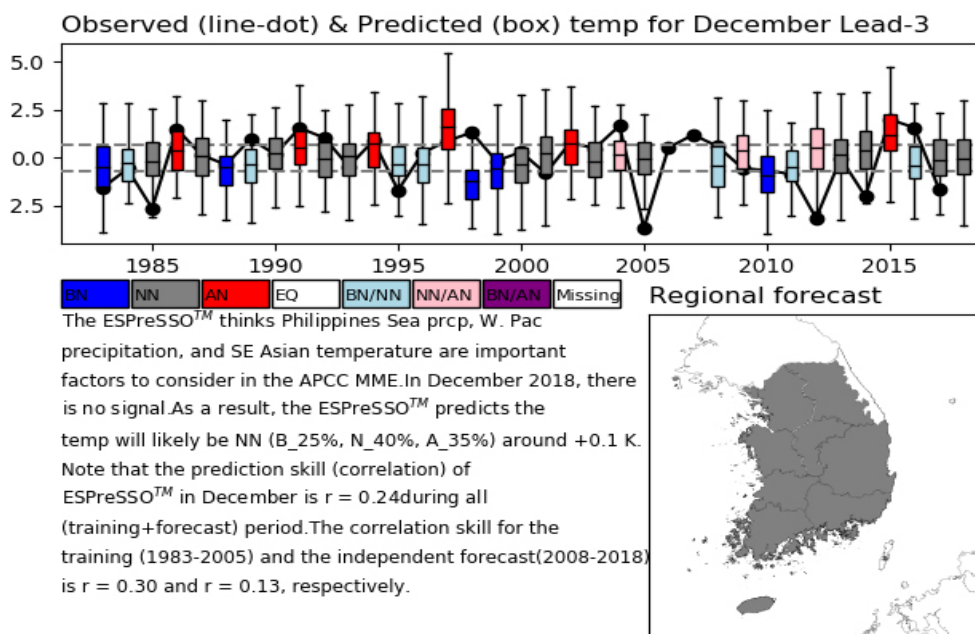


The ESPreSSOTM thinks SE Indian Ocean SST, western

Indian Ocean precipitation, NE Indian Ocean prcp, and SE Indian Ocean prcp are important factors to consider in the APCC MME. In November 2018, there is no signal. As a result, the ESPreSSOTM predicts the temp will likely be AN (B_25%, N_35%, A_40%) around +0.2 K. Note that the prediction skill (correlation) of ESPreSSOTM in November is $r = 0.33$ during all (training+forecast) period. The correlation skill for the training (1983-2005) and the independent forecast (2008-2018) is $r = 0.23$ and $r = 0.48$, respectively.

Regional forecast





【연구자】

김원무 선임연구원

여새림 선임연구원

김유진 선임연구원

연구보고서 2018-10

상세화 기법을 응용한 한반도 전문가 계절 예측 시스템 II

I S B N 979-11-5698-242-5

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.