

바다에서 찾는 **국민의 행복,**
인류에 공헌하는 **해양과학기술**

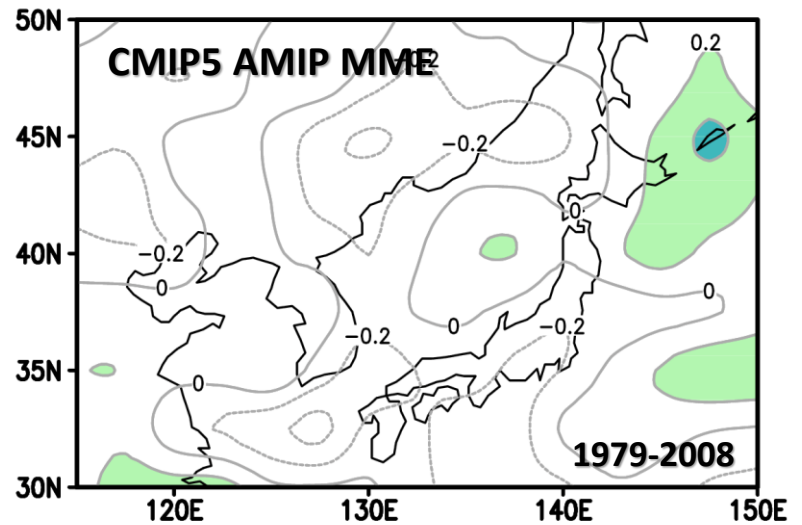
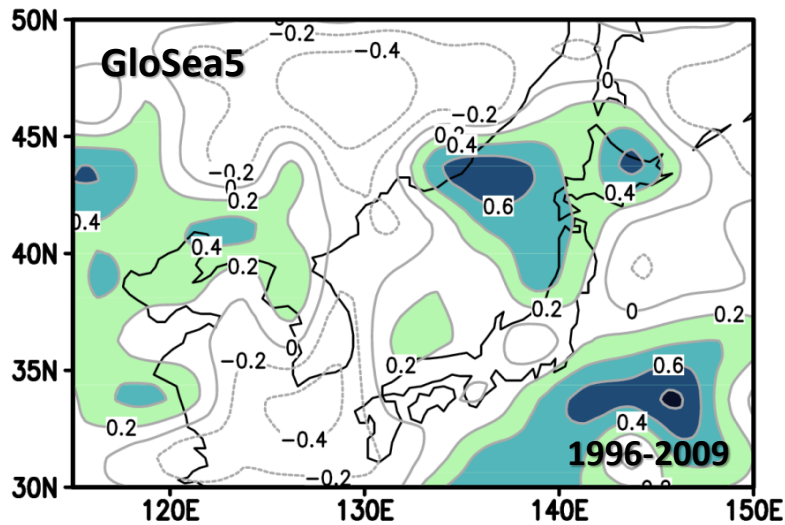
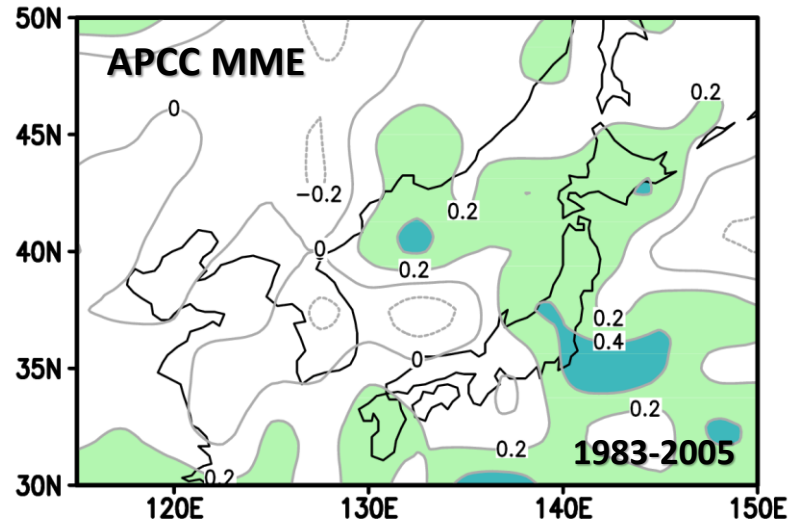
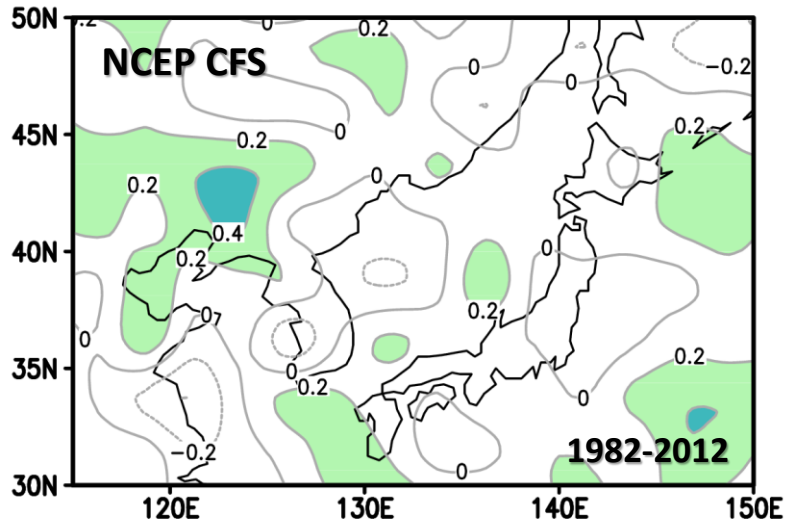
통계 및 역학-통계모형을 이용한 여름철 북동아시아 강수의 계절예측

권민호, 이강진



역학모형을 이용한 여름철 강수량 예측성

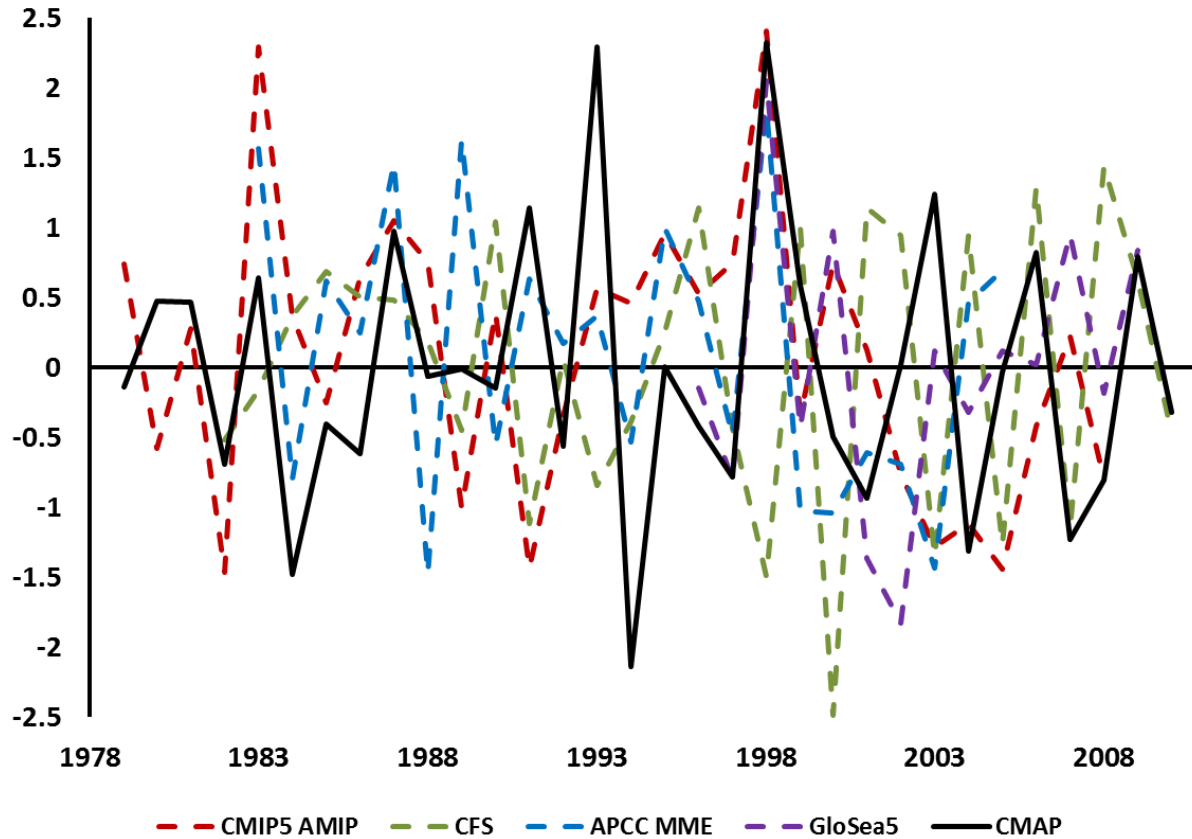
바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



◦ 역학모형을 이용한 여름철 강수량 예측성

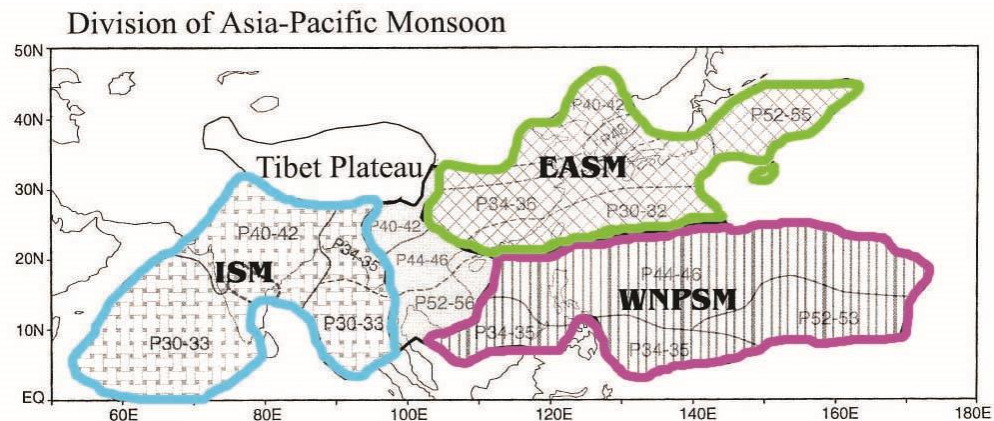
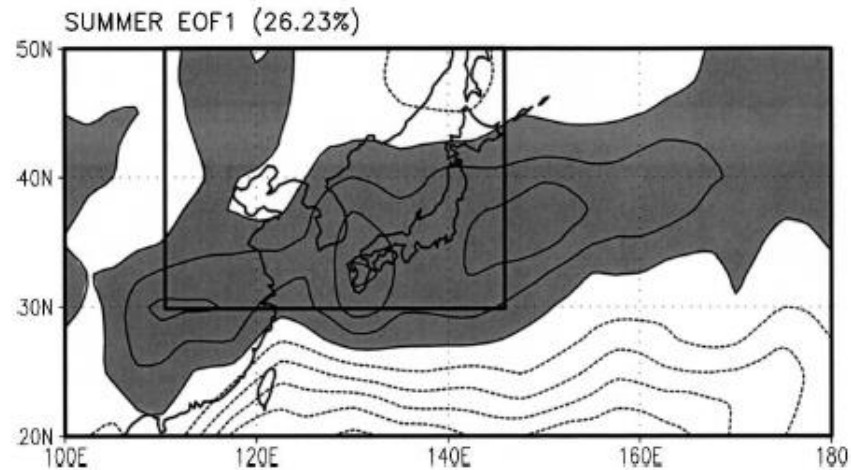
바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

(East Asia Summer Rainfall Anomaly : Pr [115°~150°E, 30°~50°N])



EASRA (East Asia Summer Rainfall Anomaly) :

- ✓ Precipitation anomaly (JJA, 115E-150E, 30N-50N)
← Lee et al. (2005)



Wang and LinHo (2002)

❖ Time-series model

AR → ARIMA → HMM → RNN → LSTM → Transformer (time series)

...

❖ Non-time-series model

Analog → RF

MLR → CCA → ANN → SVM → CNN → Transformer (vision)

...

AR : AutoRegressive, ARIMA : AutoRegressive Integrated Moving Average, HMM : Hidden Markov Model, RNN : Recurrent Neural Network, LSTM : Long-Short Term Memory, RF : Random Forest, MLR : Multiple Linear Regression, CCA : Canonical Correlation Analysis, ANN : Artificial Neural Network, SVD : Support Vector Machine, CNN : Convolutional Neural Network

MLR

✓ Possible predictors for East Asian summer monsoon

ENSO : Chang et al. (2000), Wang et al. (2000)

NAM : Yasunari and Seki (1992), Yang et al. (2004)

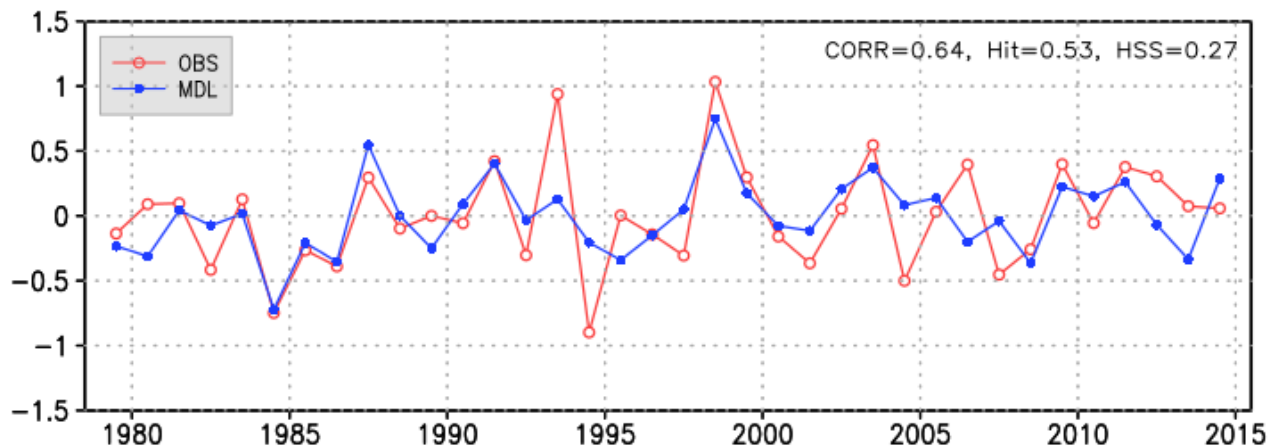
Indian Ocean : Latif et al. (1994), Clark et al. (2000)

Snowcover : Zhang et al. (2004), Yim et al. (2010)

✓ Selected predictors

SST_MAM (45E~65E, 15S~Eq.), H850_MAM (145E~165E, 55N~70N)

Snowcover_MAM (100E~120E, 50N~55N)



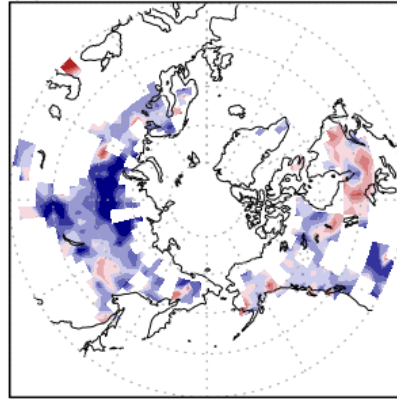
$$HSS(\text{Heidke skill score}) = \frac{\frac{1}{N} \sum n(F_i, O_i) - \frac{1}{N^2} \sum N(F_i)N(O_i)}{1 - \frac{1}{N^2} \sum N(F_i)N(O_i)}$$

0 : no skill

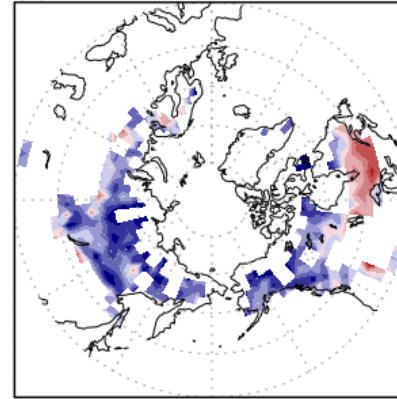
1 : perfect skill

CORR(EASRA,SC_MAM)

(a) 1979-1999



(b) 2000-2021



문제점

- 예측인자와 예측변수간의 상관관계는 통계적 항상성을 만족하지 않음
- 전체기간 예측인자 추출에 의한 통계적 모델의 예측성은 과평가되는 경향을 가짐
- “통계적” 예측인자에 의한 통계모델의 예측성은 불안정을 가짐

개선방안

- 예측인자/예측변수 간 관계의 시간규모에 대한 고려
- 통계적 앙상블(BAGGING)
- 역학적 설명이 가능한 예측인자의 추출
- 역학모형의 활용
- 딥러닝 등의 머신러닝 기법의 적용

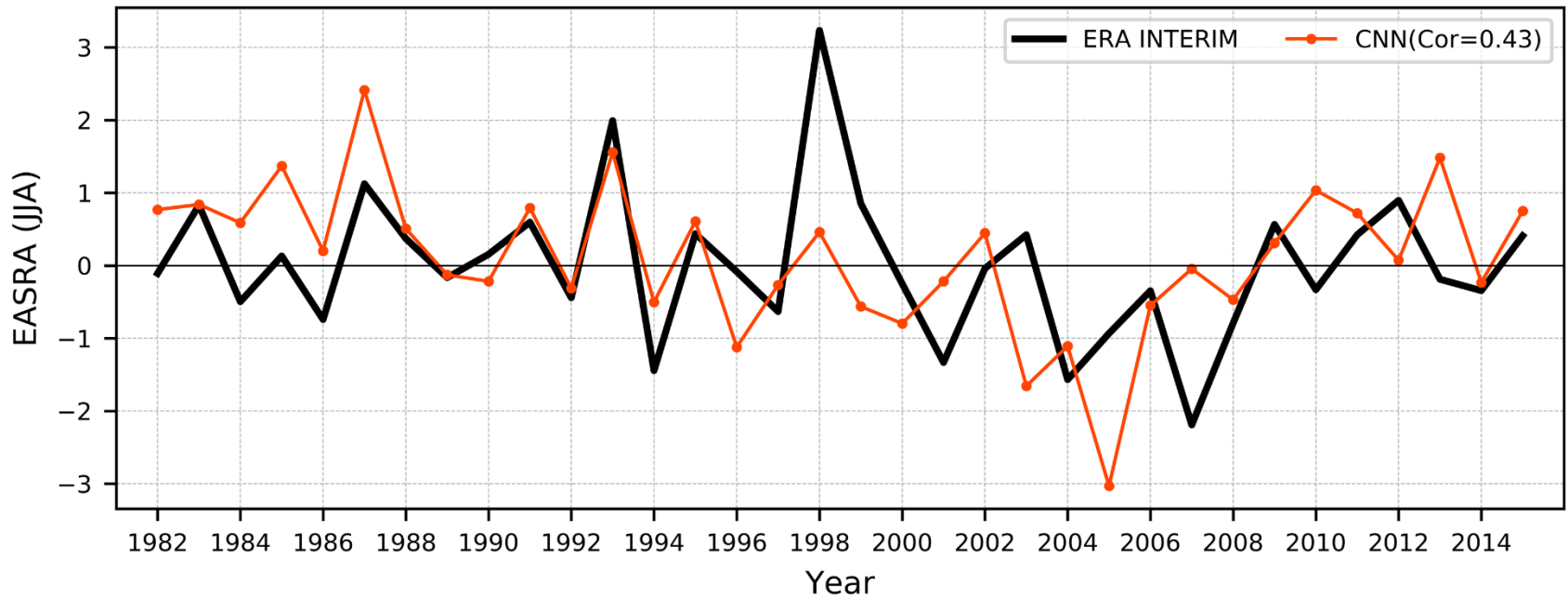
○ CNN을 이용한 여름철 동아시아 강수 예측

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



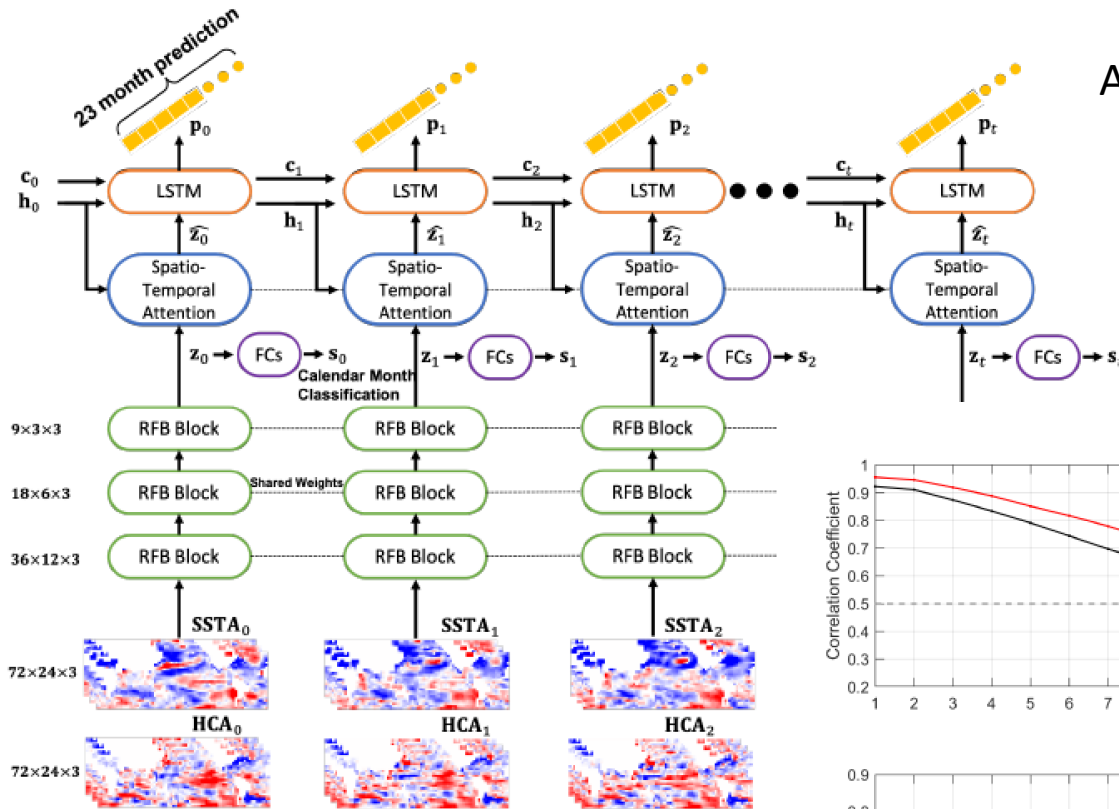
❖ 2D-CNN을 이용한 동아시아 강수 예측

학습 데이터	CMIP5 historical 21개 모형 (3024년)	입력	3, 4, 5월 전지구 해수면 온도 (5°x5°)
검증 데이터	ERA_INTERIM (34년)	출력	6, 7, 8월 평균 EASRA (East Asian Summer Rainfall anomaly, 115°~150°E, 30°~50°N)

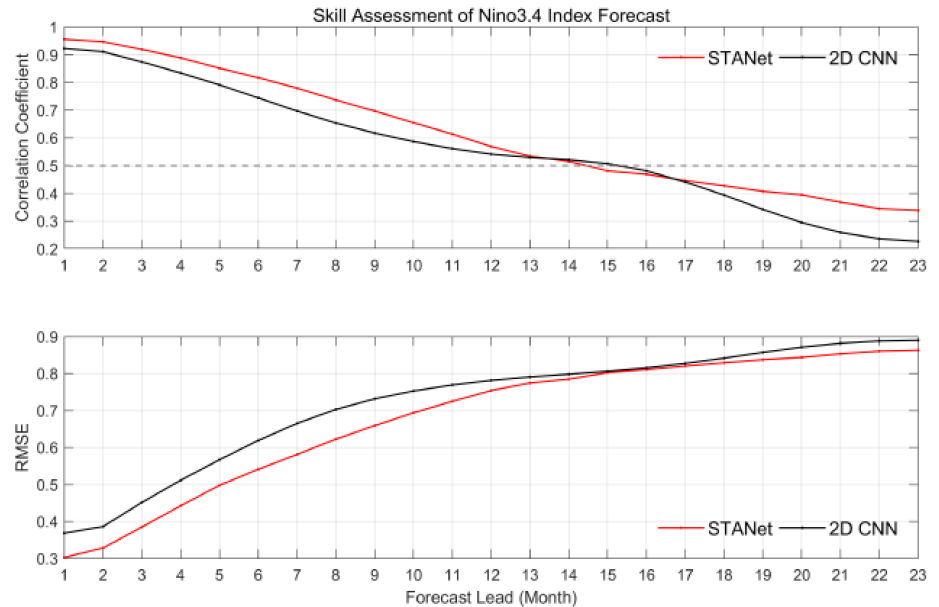


딥러닝을 이용한 엘니뇨 예측

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



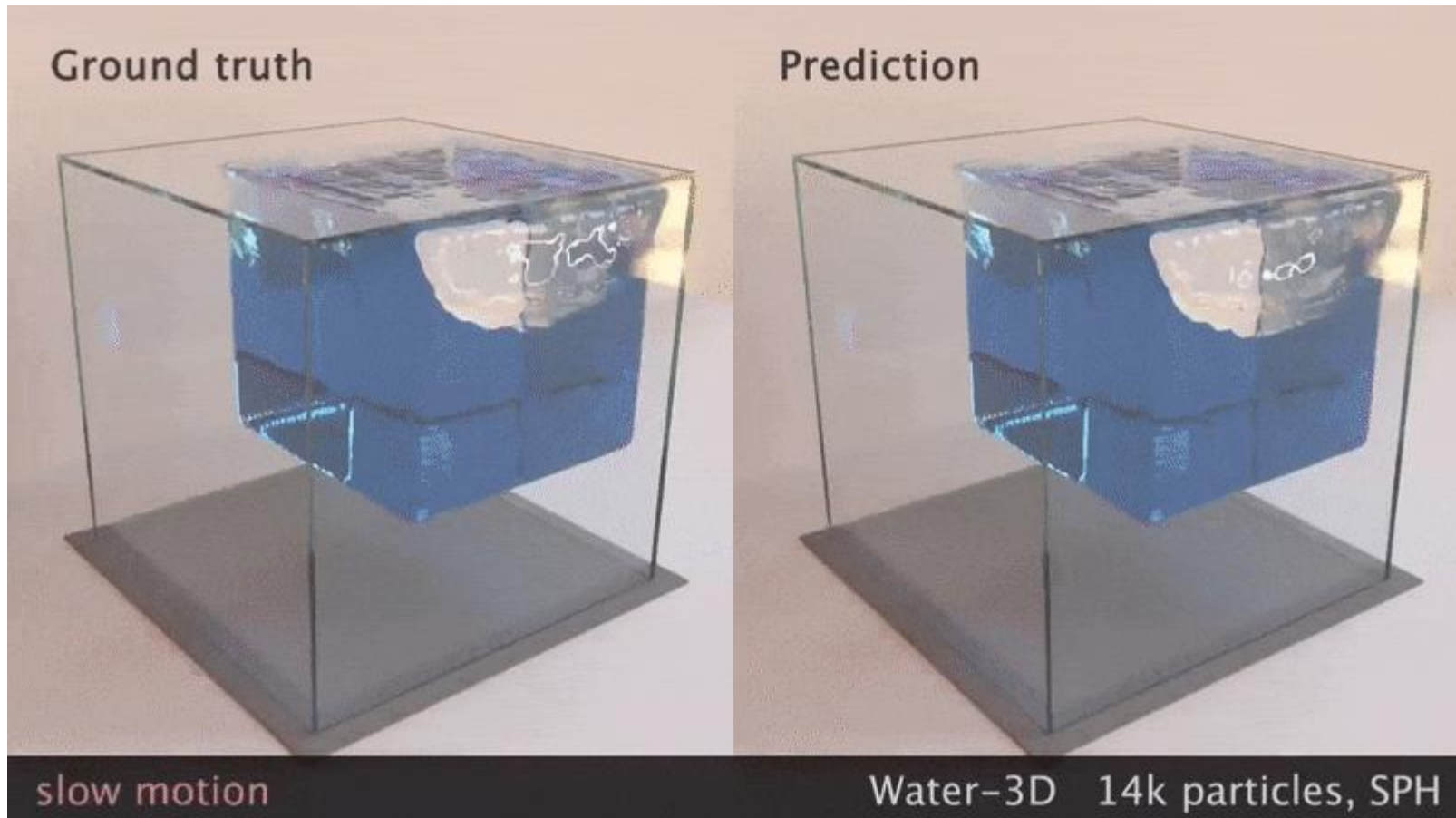
Attention 기반 모델



Structure of Spatiotemporal neural network with attention

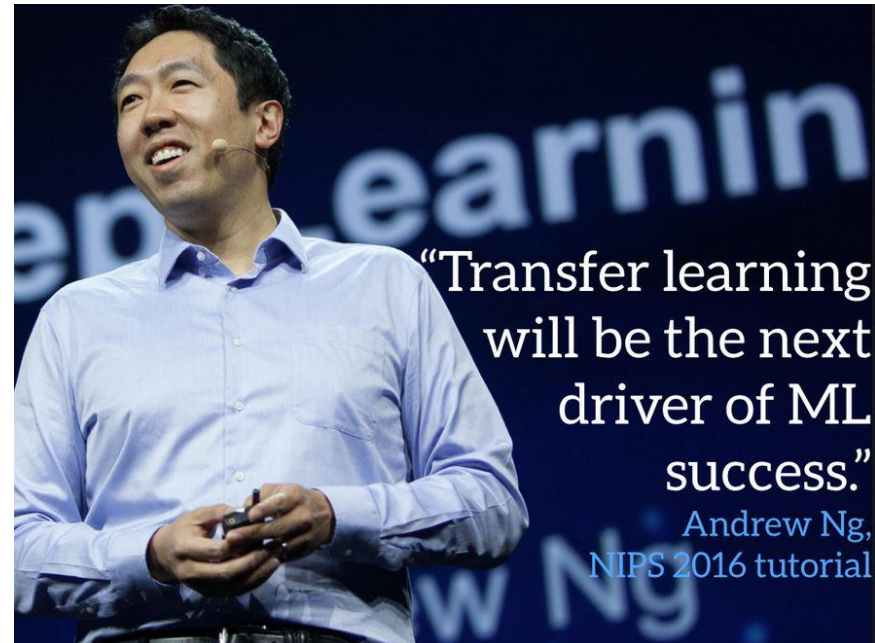
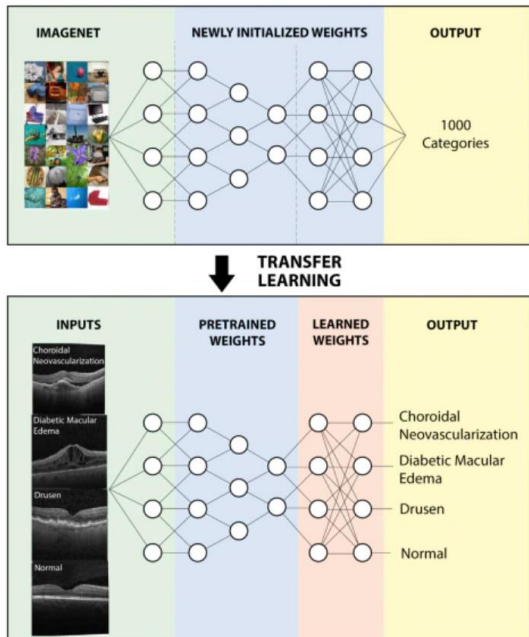
Physics-based deep learning

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



Sanchez-Gonzalez et al. (2020)

전이학습(transfer learning)

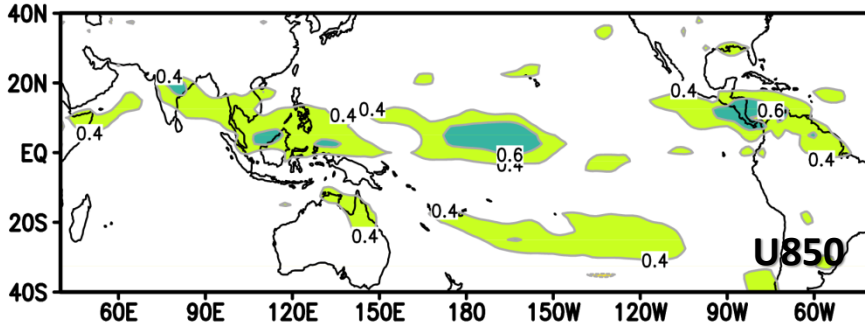


- 계절예측을 위한 기후자료의 부족한 샘플수 → 전이학습
- 딥러닝 기술(예, transfer model)의 빠른 발전 → 관련 전문가와 협업
- 기후예측에 패턴인식 딥러닝 기술 적용 → 물리기반 딥러닝 모델 개발

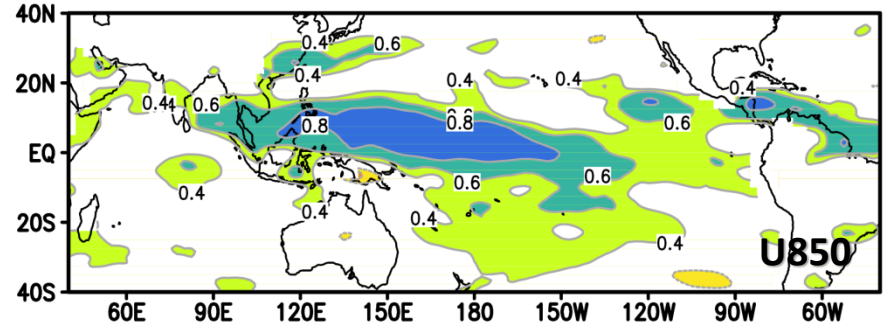
역학모형을 이용한 열대 지역 예측성

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

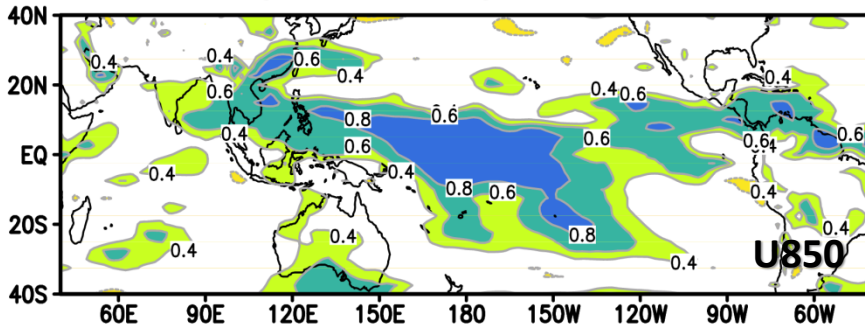
NCEP CFS (1982-2012)



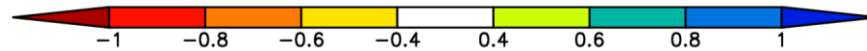
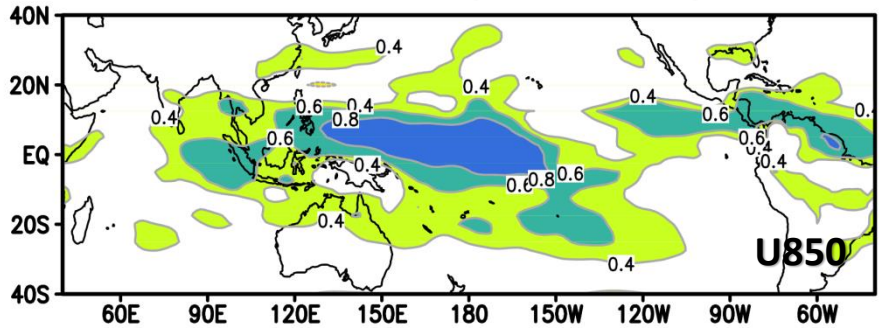
APCC MME (1983-2005)



GloSea5 (1996-2009)



CMIP5 AMIP MME (1979-2008)



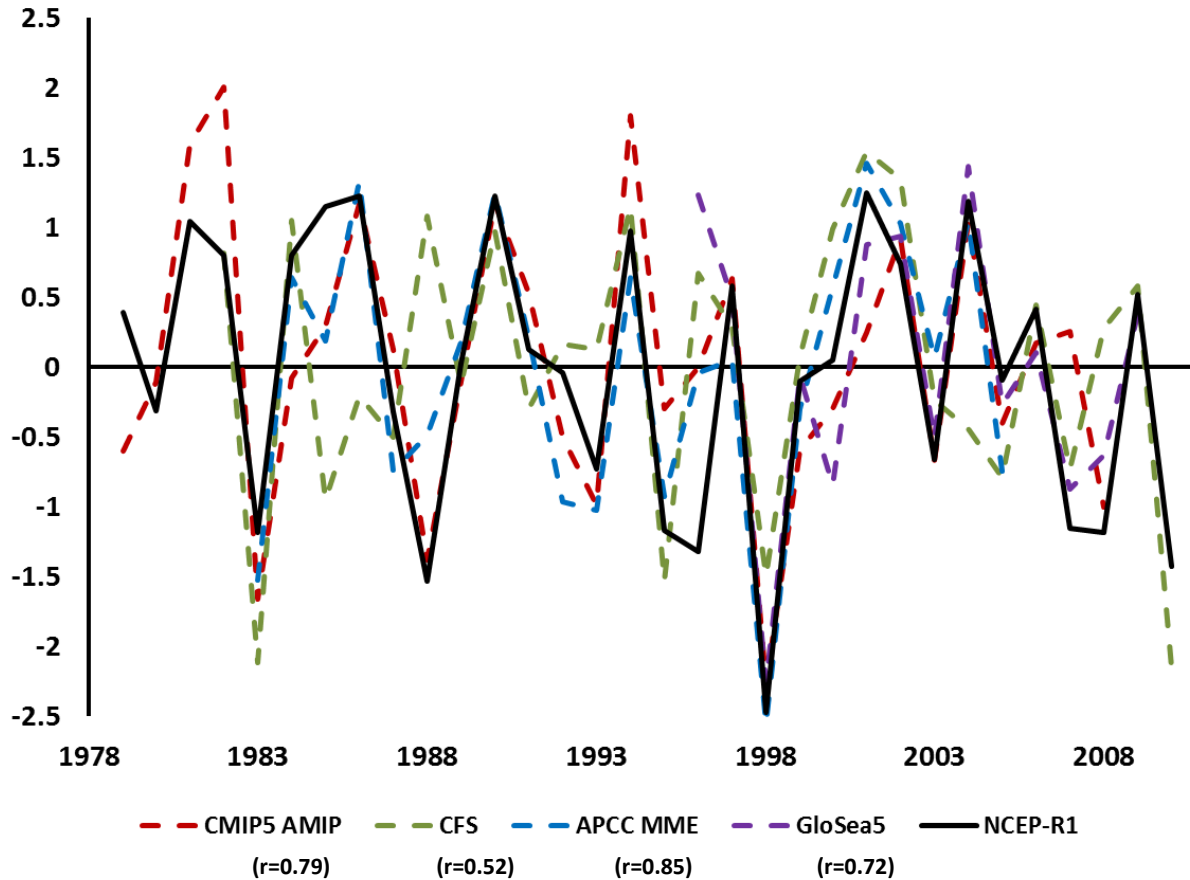
Prediction skills of JJA-mean U850 initialized on 1 May

◦ 북서태평양 여름몬순의 예측성

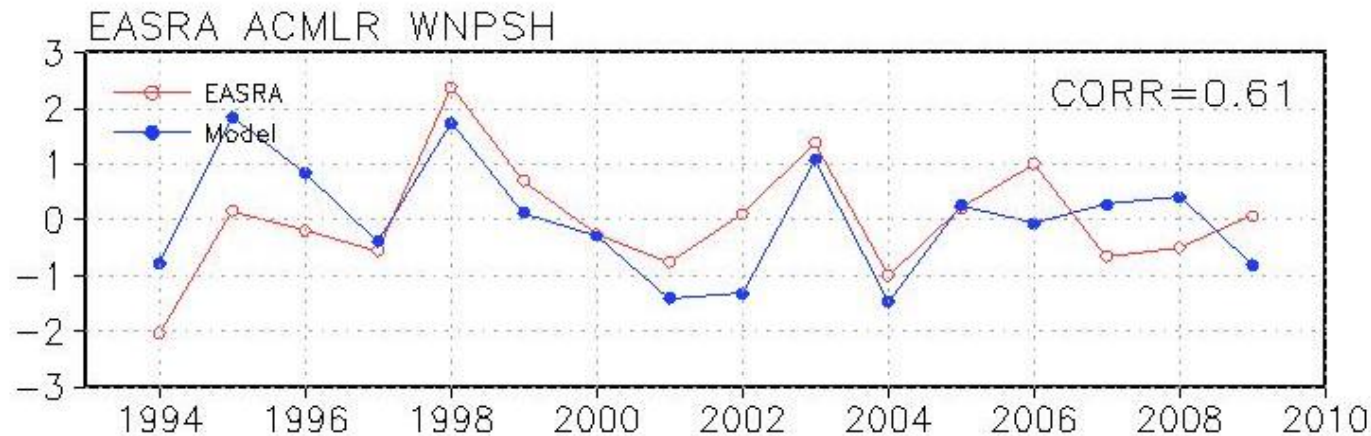
바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



(Western North Pacific Monsoon Index : U850 [100°~130°E, 5°~15°N] - [110°~140°E, 20°~30°N])



WNPSH의 예측성
 GFDL(r=0.80)
 ABOM(r=0.76)
 Xiang et al. (2013)

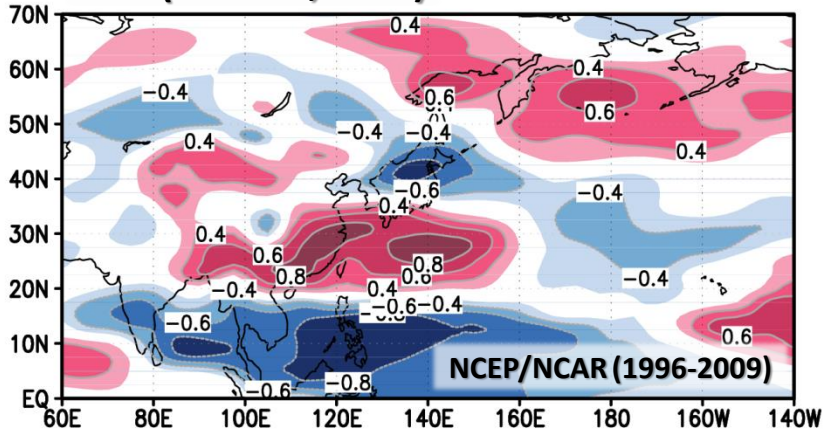


- ❖ 예측인자 : 북서태평양 고기압 지수(WNPSH) ← GloSea5
- ❖ 예측변수 : EASRA(East Asian Summer Rainfall Anomaly)
Precipitation anomaly (JJA, 115E-150E, 30N-50N)

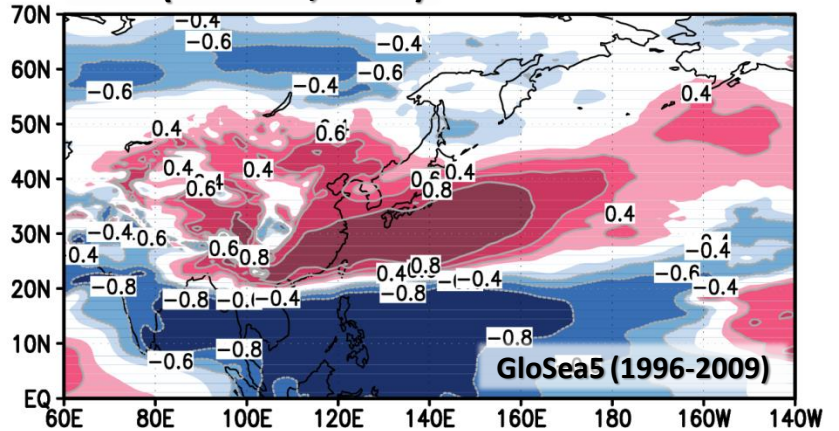
북서태평양 원격상관의 예측성

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

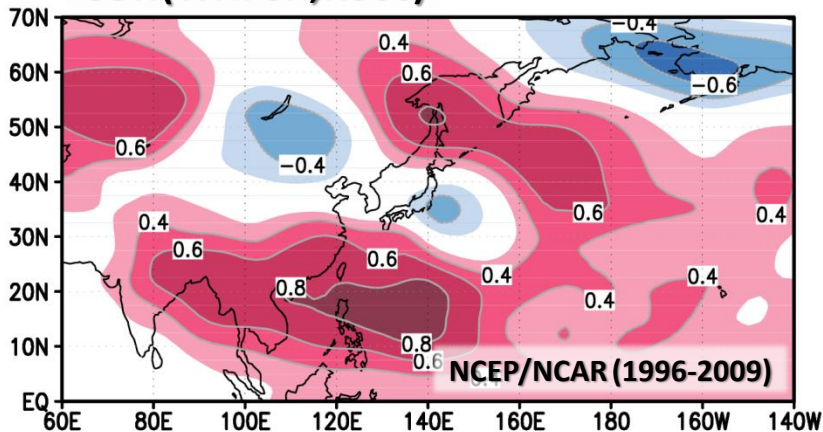
COR (WNPSH,U850)



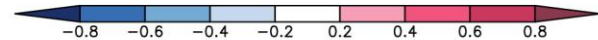
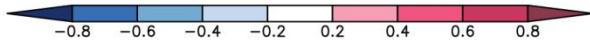
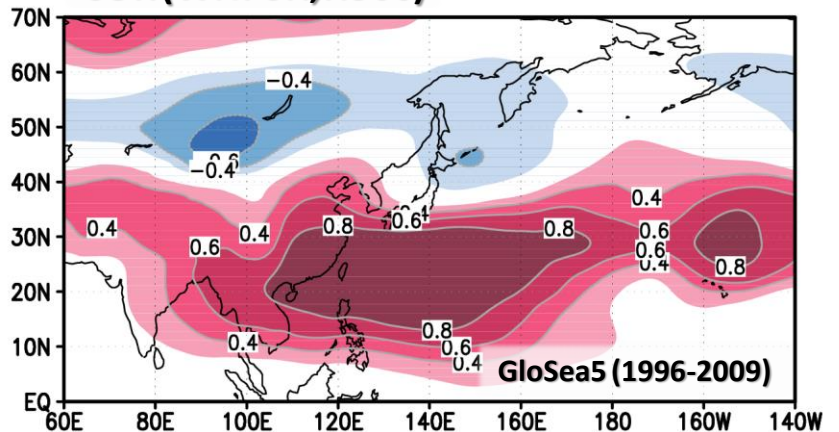
COR (WNPSH,U850)



COR (WNPSH,H500)



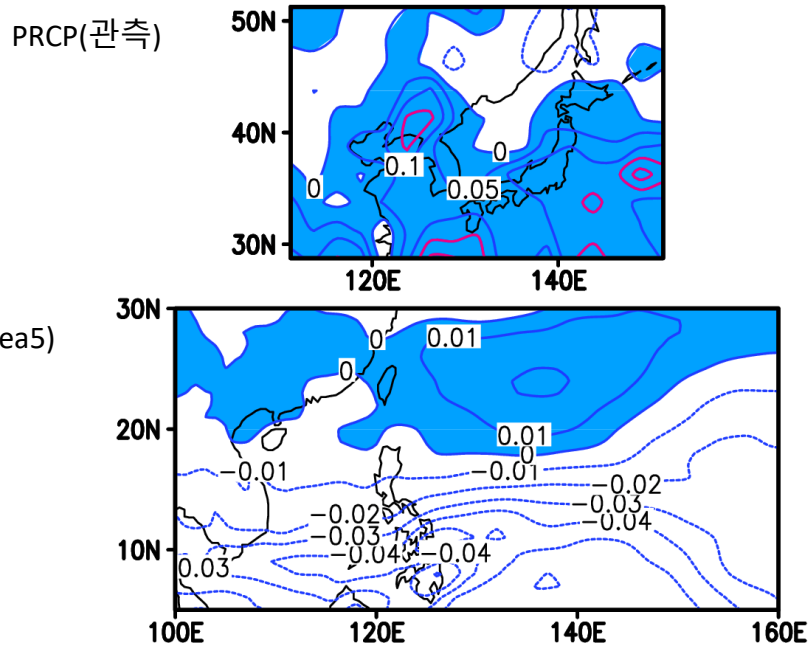
COR (WNPSH,H500)



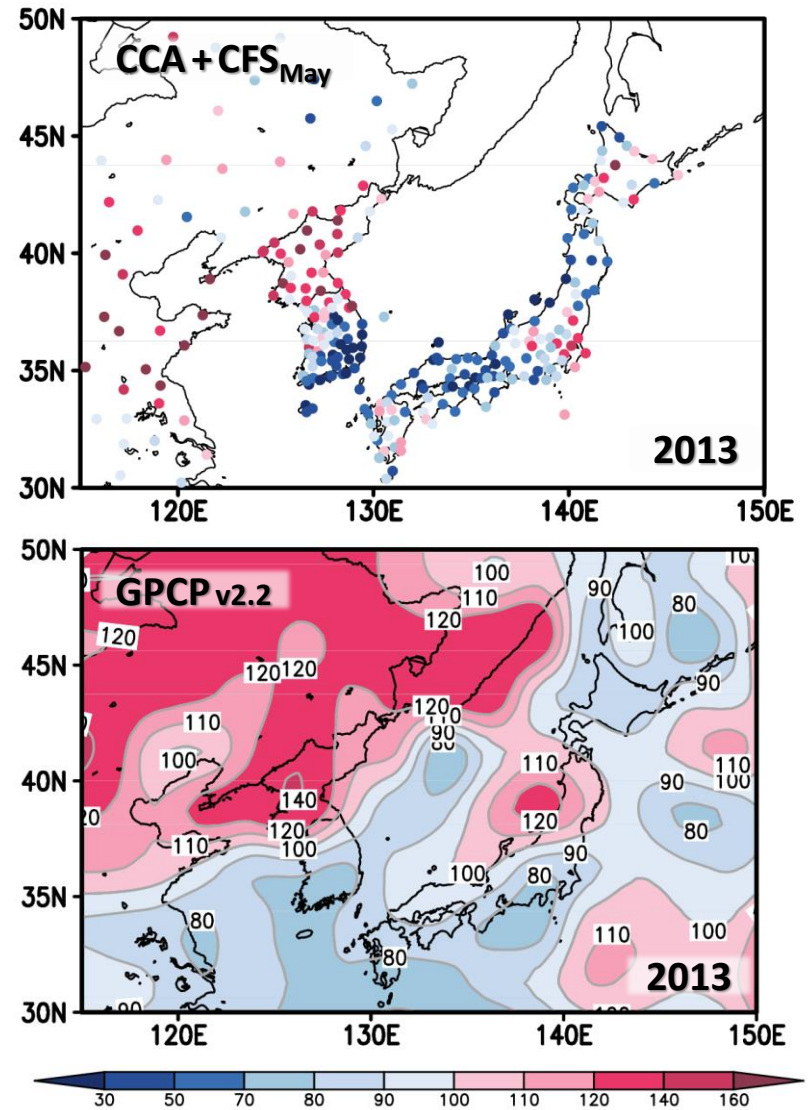
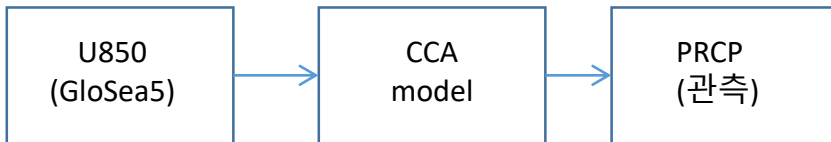
역학-통계모형을 이용한 계절예측

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

❖ CCA (Canonical Correlation Analysis) model



U850(GloSea5)



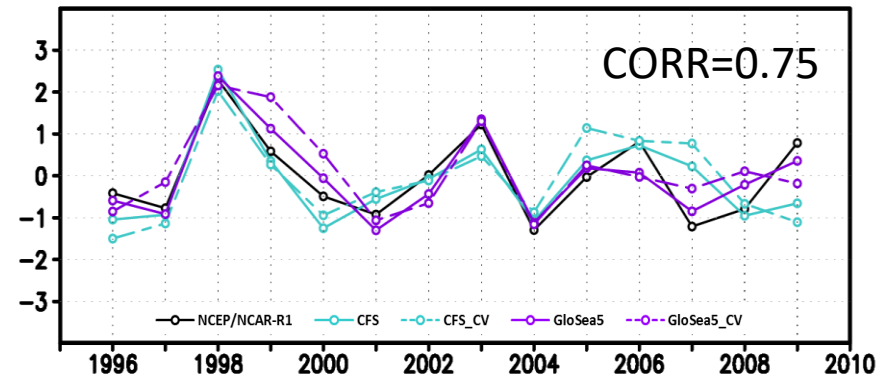
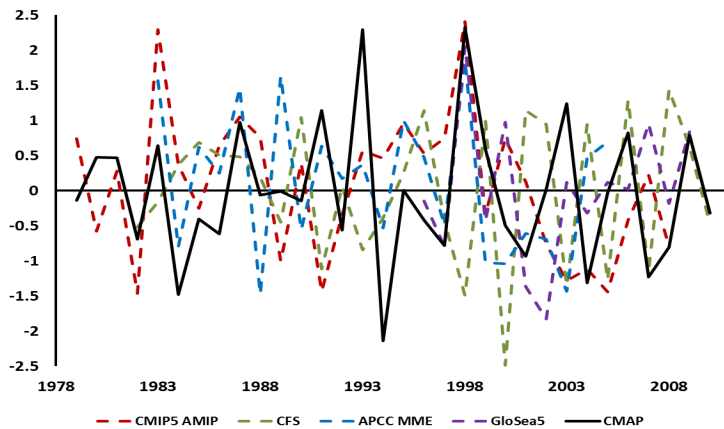
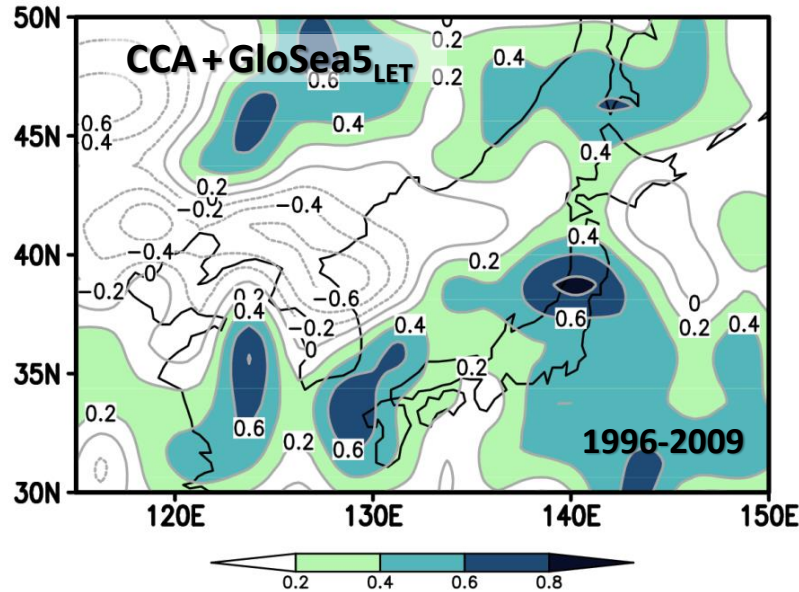
2013

2013

CCA (Canonical Correlation Analysis) model

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

예측성



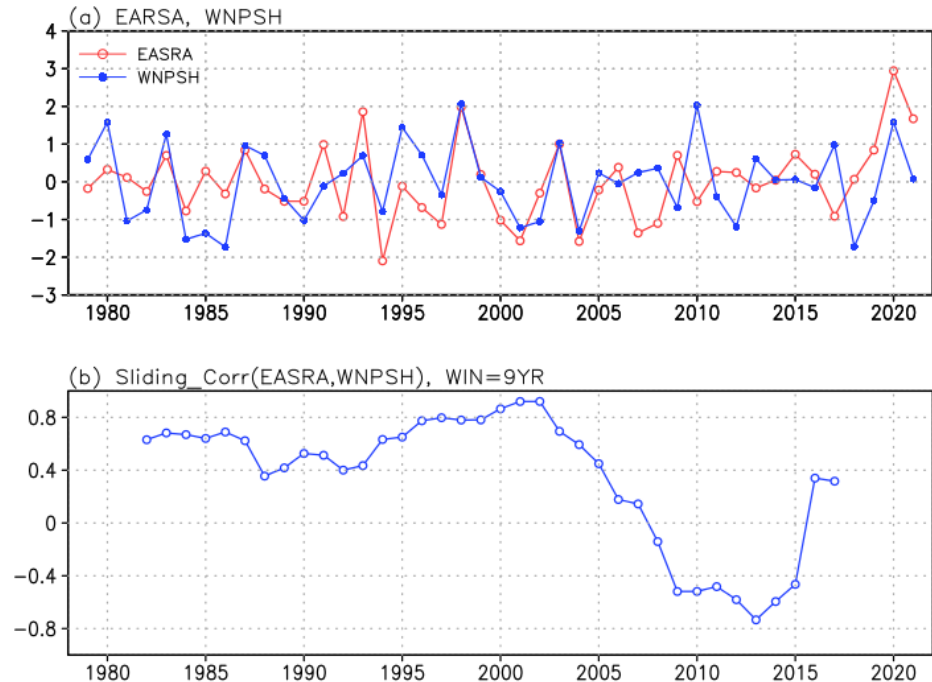
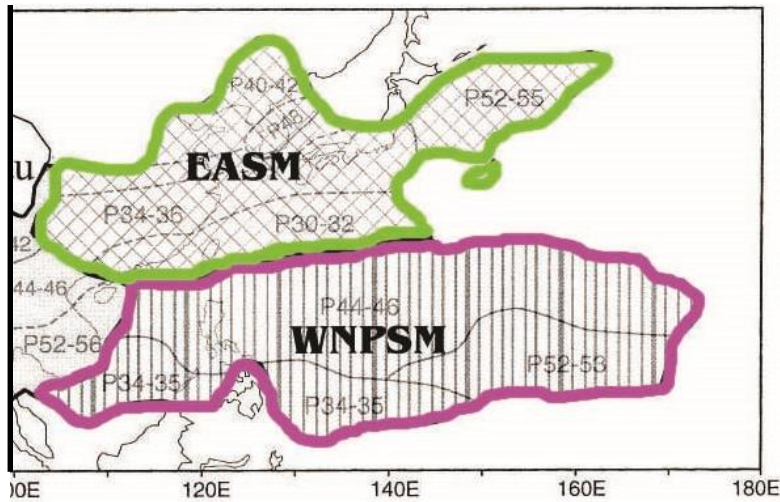
Lee and Kwon (2015)

동아시아여름몬순과 북서태평양여름몬순

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

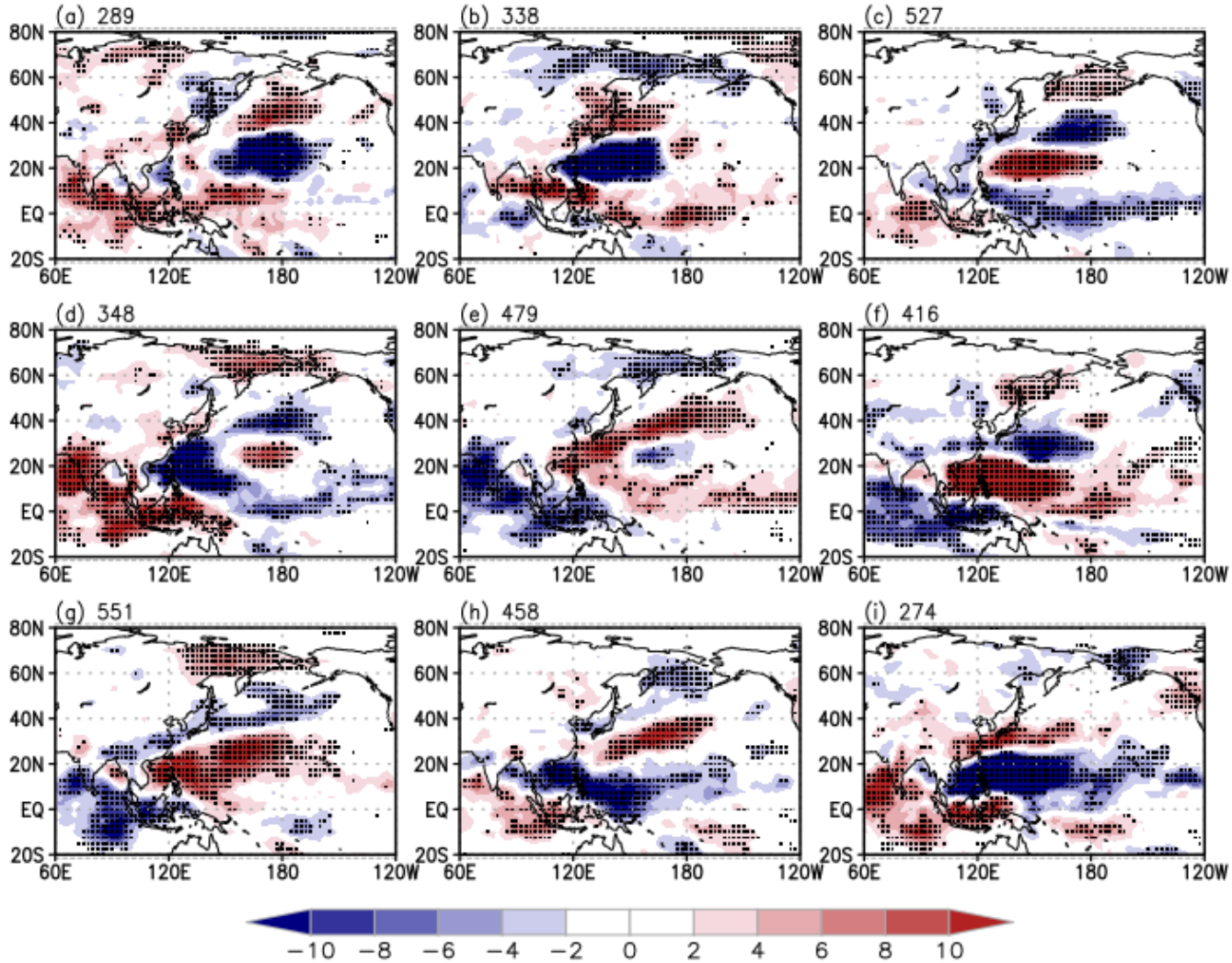
- East Asian Summer Rainfall Anomaly (EASRA) Lee et al. (2005)
- Western North Pacific Subtropical High Index (WNPSH) Wang et al. (2013)

아시아-태평양 여름몬순



OLR 합성도(북서태평양 클러스터 분석)

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

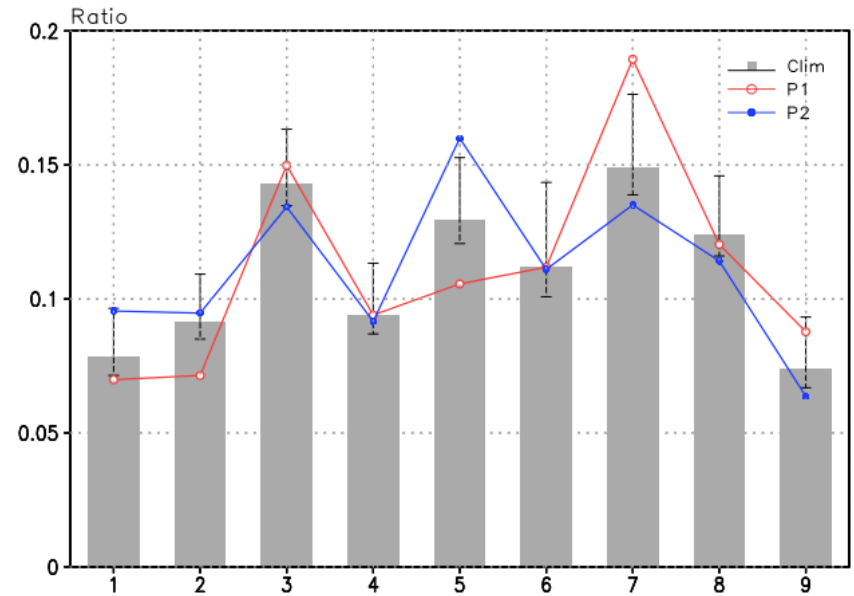
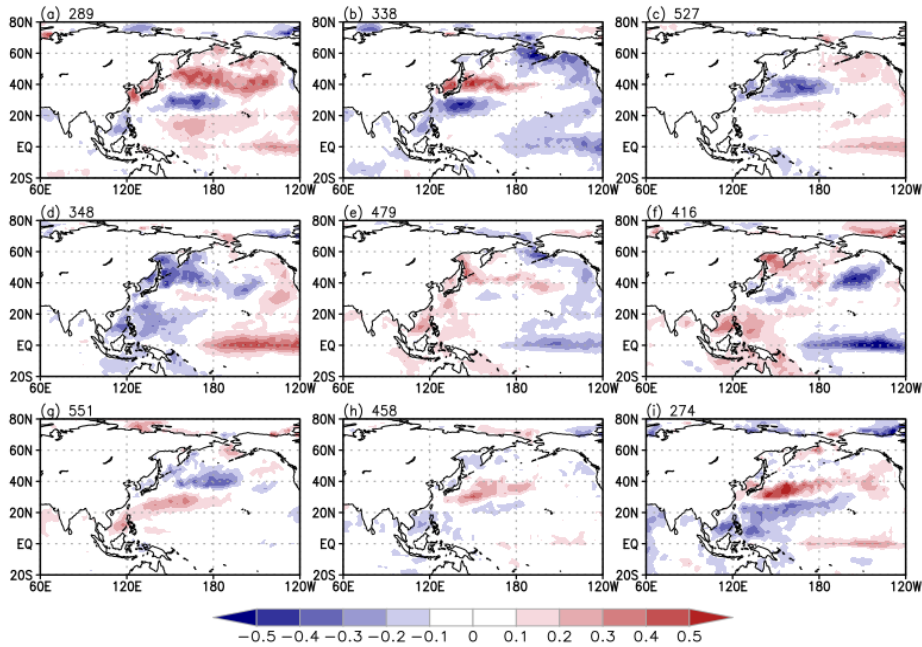


Stippled area : 95% significant

클러스터들의 확률분포

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

SSTA SOM



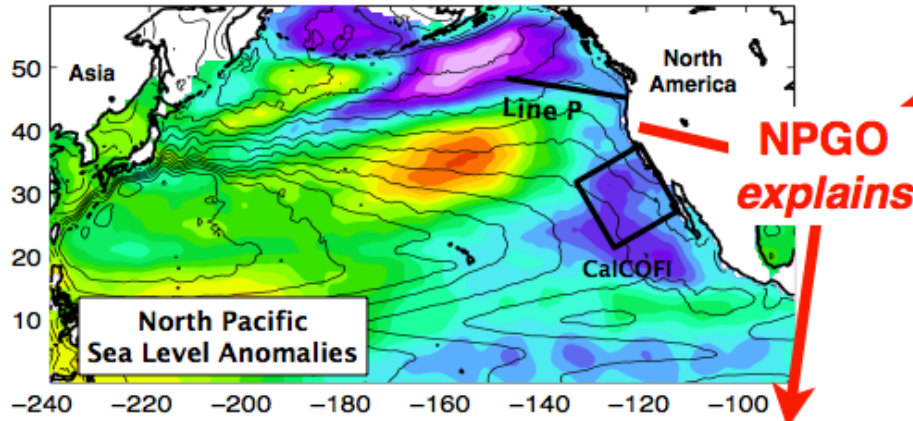
P1 : 1991-2004

P2 : 2005-2018

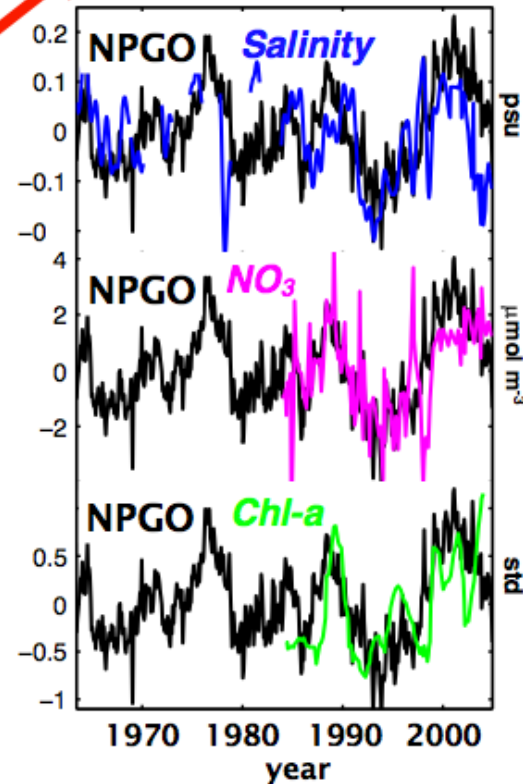
○ NPGO (North Pacific Gyre Oscillation)

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

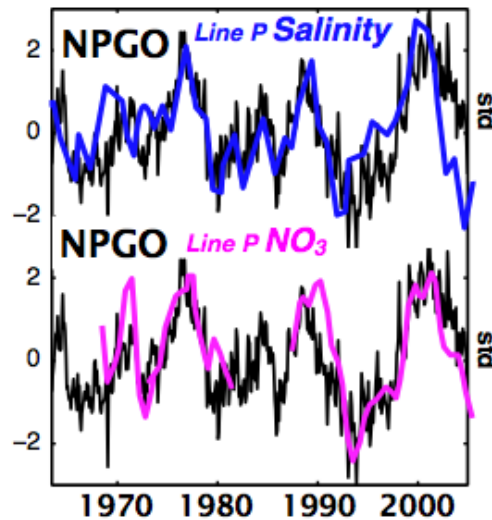
North Pacific Gyre Oscillation (NPGO)



California Current CalCOFI Observations



Gulf of Alaska Line P Observations

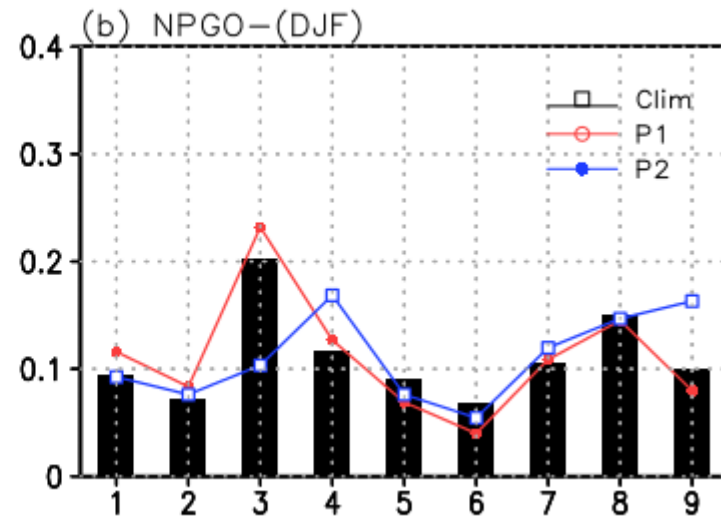
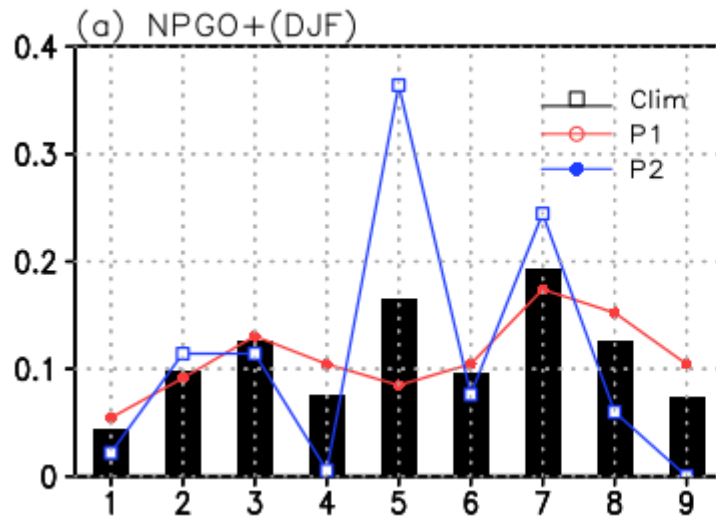
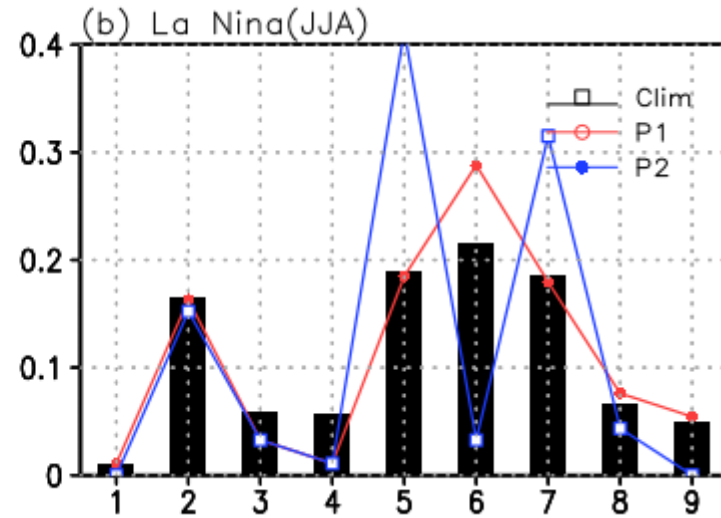
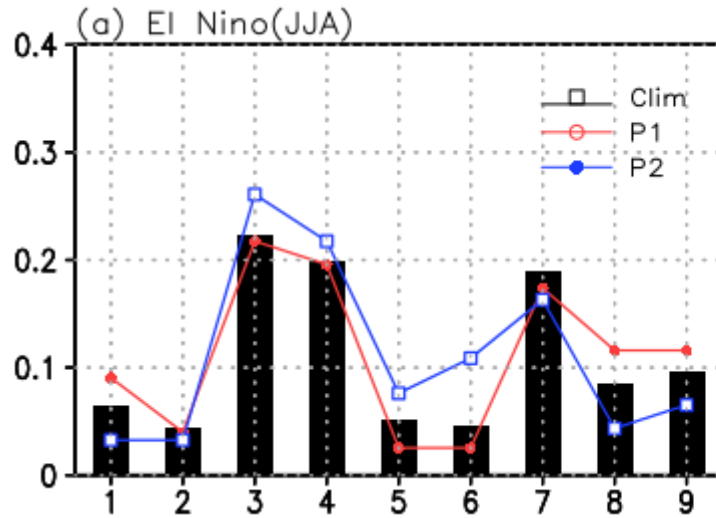


The NPGO index measures changes in the North Pacific gyres circulation and explains key physical-biological ocean variables

Di Lorenzo et al., 2008

북서태평양 패턴과 NPGO

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



❖ 예측인자

- 역학모형에서 예측된 여름철 평균 U850 편차(100E-180, Eq-30N)
- 한반도 기후와 관련 있는 패턴 추출 → SOM 패턴 관련 시계열 추출

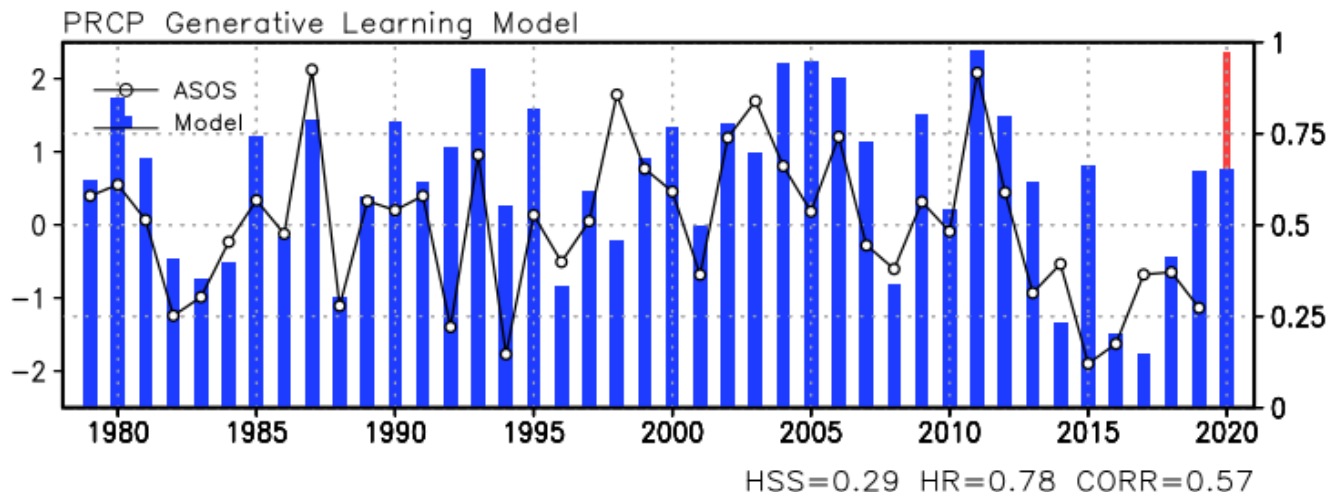
$$x' = x \cdot |\tanh(\tau y)|$$

❖ 생성학습모델

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} = \frac{p(x|y)p(y)}{\sum_i p(x|y_i)p(y_i)}$$

← 학습을 통하여 likelihood function
을 최대로 만드는 매개변수 추정

$$l(\phi, \mu_y, C) = \log \prod_i p(x_i, y_i; \phi, \mu_y, C)$$



- ✓ 북동아시아 지역 여름철 강수량의 통계적 계절예측은 모형의 검증을 고려하지 않고, 현업적인 측면에서 최근 기간에 추출된 예측인자만 활용하는 것이 유리함
- ✓ ENSO, PMM 등 기후변동성은 계절내 시간 규모로 더 좁은 영역에 뚜렷한 영향을 미치므로 계절내규모의 통계적 모델이 효율적임
- ✓ 발전된 딥러닝 기법이 기후예측에 일반적으로 약간의 예측성 향상에 도움을 주지만 주어진 기후예측 문제에서 개선된 딥러닝 기법이 더 높은 예측성을 담보하지는 않음
- ✓ 샘플수가 부족한 기후예측에서 간단한 머신러닝 기법으로도 높은 예측성의 확보가 가능함
- ✓ 전이학습의 광범위한 적용은 적은 샘플수를 극복할 수 있는 방안이나 그 실효성 여부는 검토해야 함
- ✓ 물리과정을 학습한 딥러닝모델이라도 하더라도 역학모형의 문제해결을 보장하지는 않음(예, Lyapunov stability problem)

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술

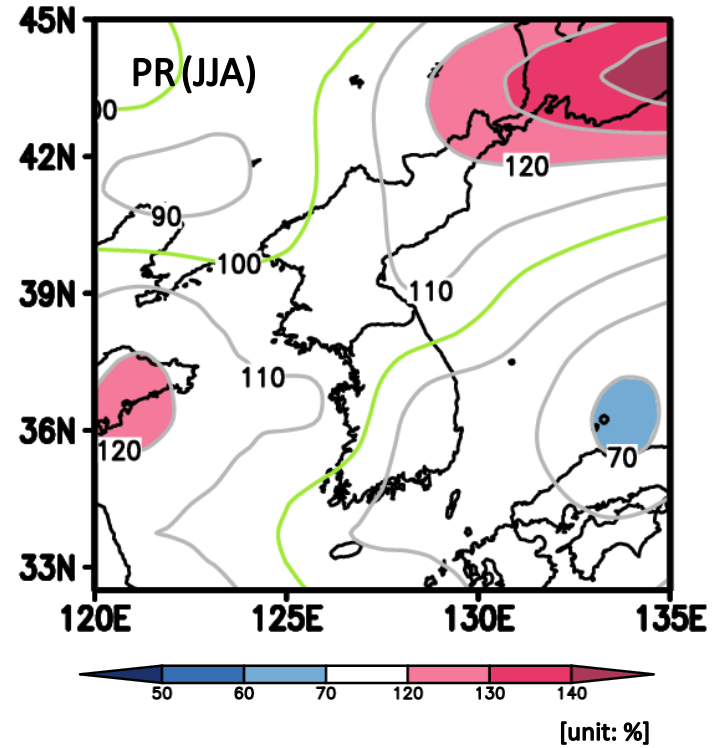
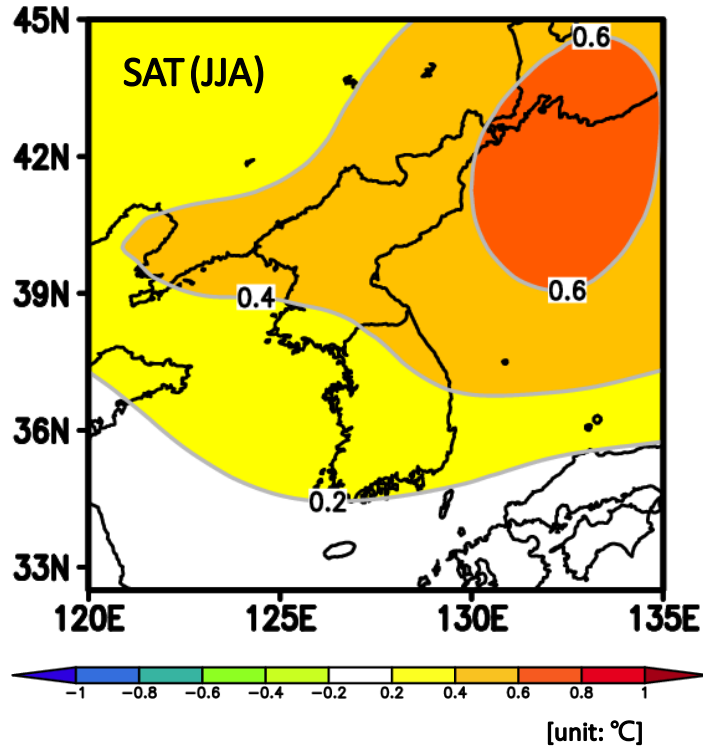


감사합니다

T H A N K Y O U

2022년 여름철 기후 편차 전망

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



◦ SOM 패턴을 이용한 역학통계모형

바다에서 찾는 국민의 행복,
인류에 공헌하는 해양과학기술



❖ 예측인자

- 역학모형에서 예측된 여름철 평균 U850 편차(100E-180, Eq-30N)
- 한반도 기후와 관련 있는 패턴 추출 → SOM 패턴 관련 시계열 추출

$$x' = x \cdot |\tanh(\tau y)|$$

❖ 생성학습모델

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} = \frac{p(x|y)p(y)}{\sum_i p(x|y_i)p(y_i)}$$

← 학습을 통하여 likelihood function
을 최대로 만드는 매개변수 추정

$$l(\phi, \mu_y, C) = \log \prod_i p(x_i, y_i; \phi, \mu_y, C)$$

$$p(y) = \phi^y (1 - \phi)^{1-y}$$

$$p(x|y) = \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp\left(-\frac{1}{2} (x - \mu_y) C^{-1} (x - \mu_y)\right)$$

❖ 가정

$y \sim B(1, \phi)$ 베르누이 확률분포 $x|y \sim N(\mu, C)$ 다변량 정규분포