

APCC Training workshop on CLIK: Seasonal prediction and its localization

2015.10.6-8

Philippine Atmospheric, Geophysical and Astronomical
Services Administration (PAGASA)



Overview

- **Variability** : “What” do we want predict
- **Predictability** : What is “to predict”
- **Methods** : “How” do we predict
- **Verification** : How to make a “good” prediction
- **Operation** : In reality?

Seasonal Prediction (0) : Climate Variability

Jin Ho Yoo
APEC Climate Center

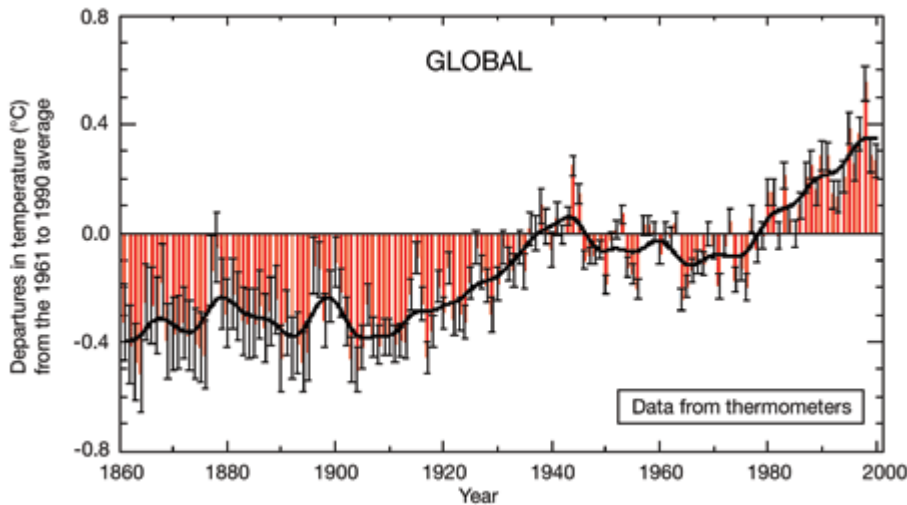


Tell us your “weather” during a season

Q: What climate phenomena are you familiar with?

Climate Change

- Changes in our **expectation**

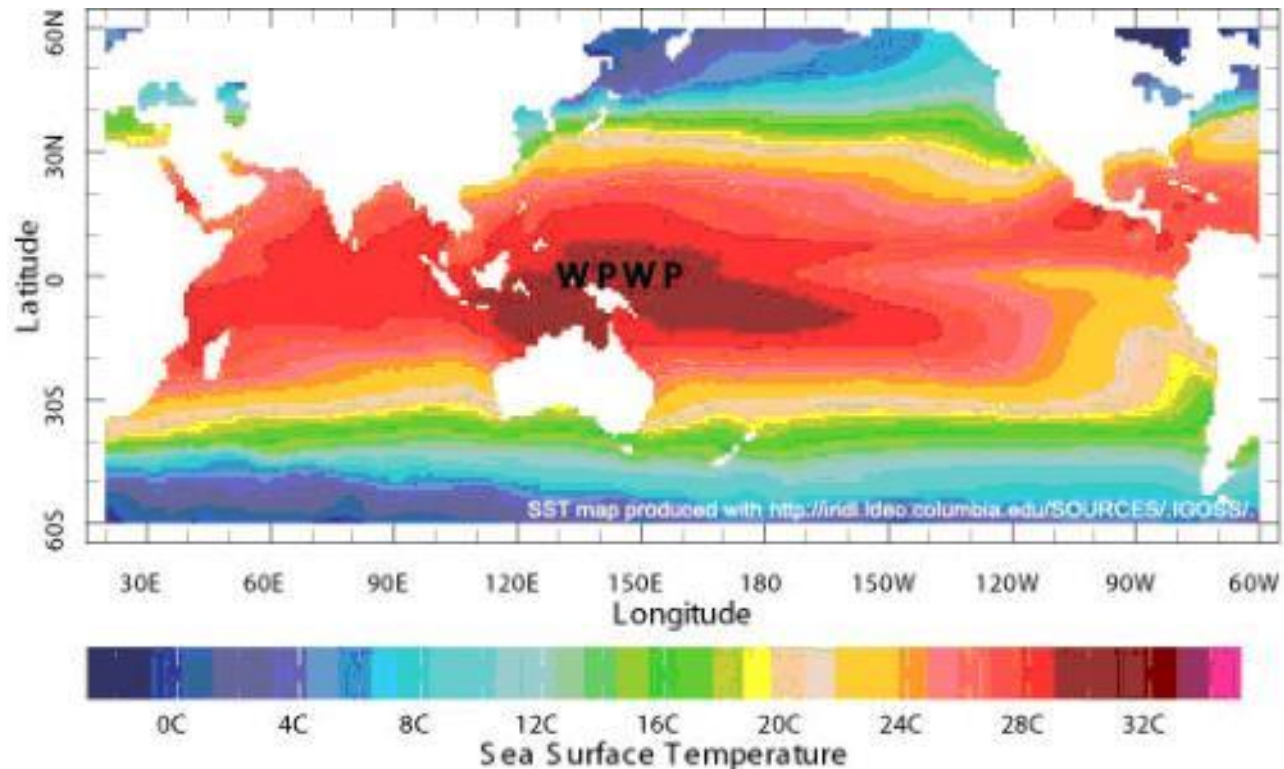


© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com

search ID: mlyn203

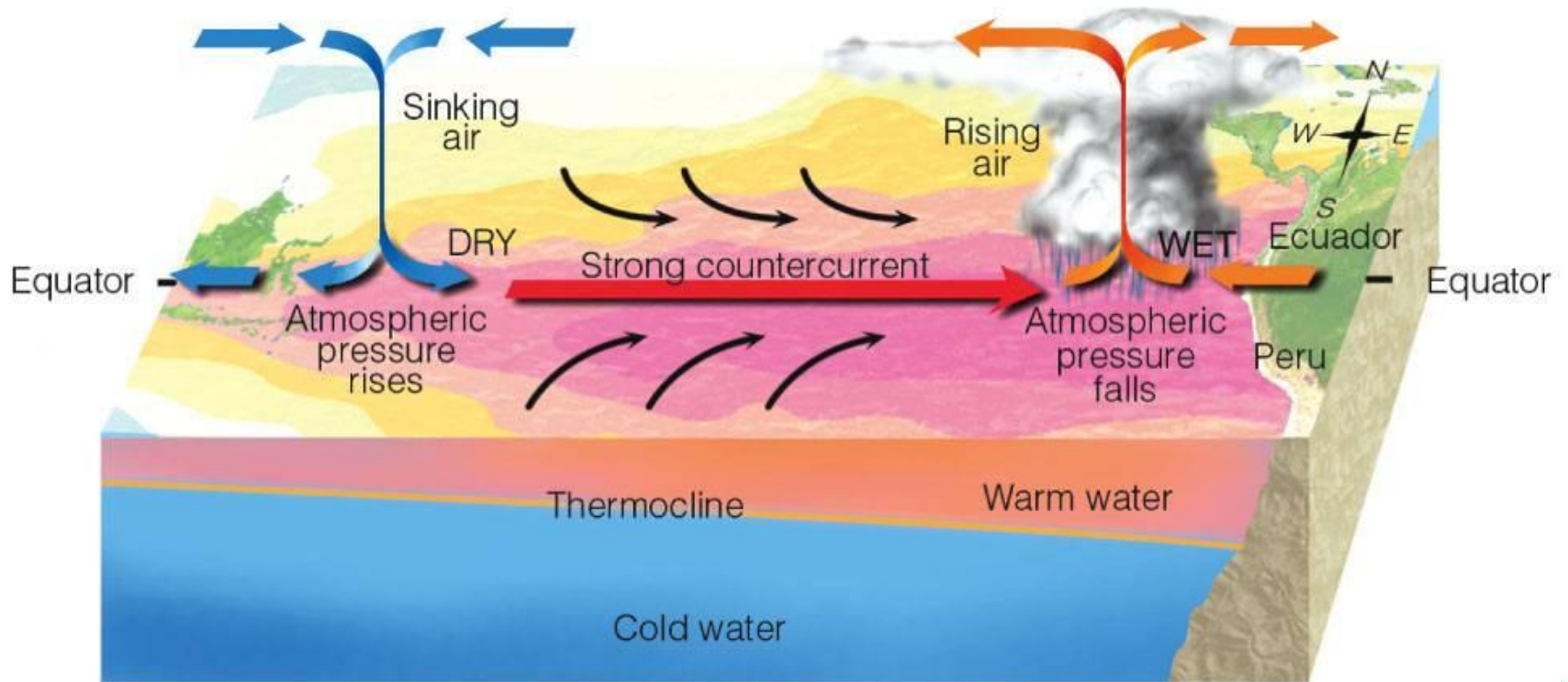
Warm pool

- Climate Engine : remember “mean” feature



El Niño, ENSO

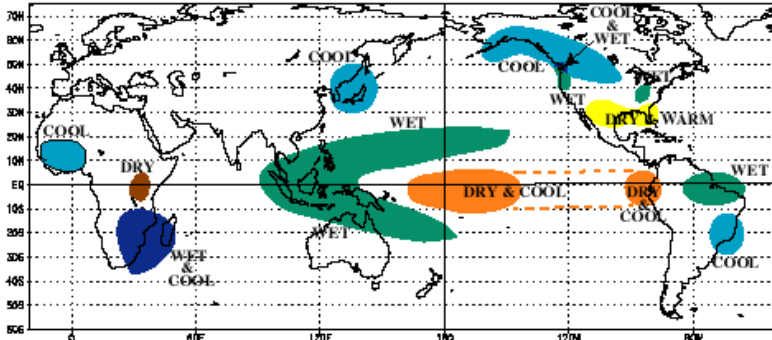
- A Big Ocean Swing



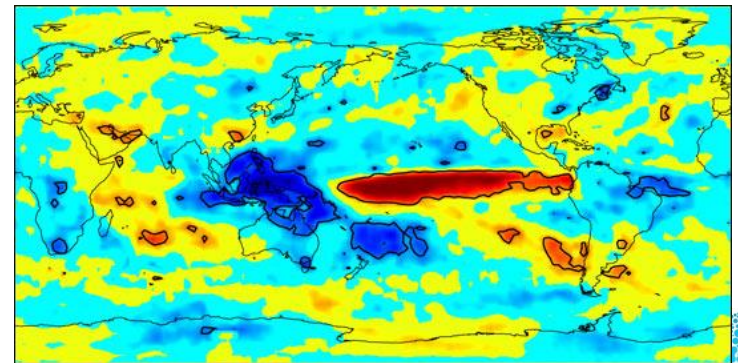
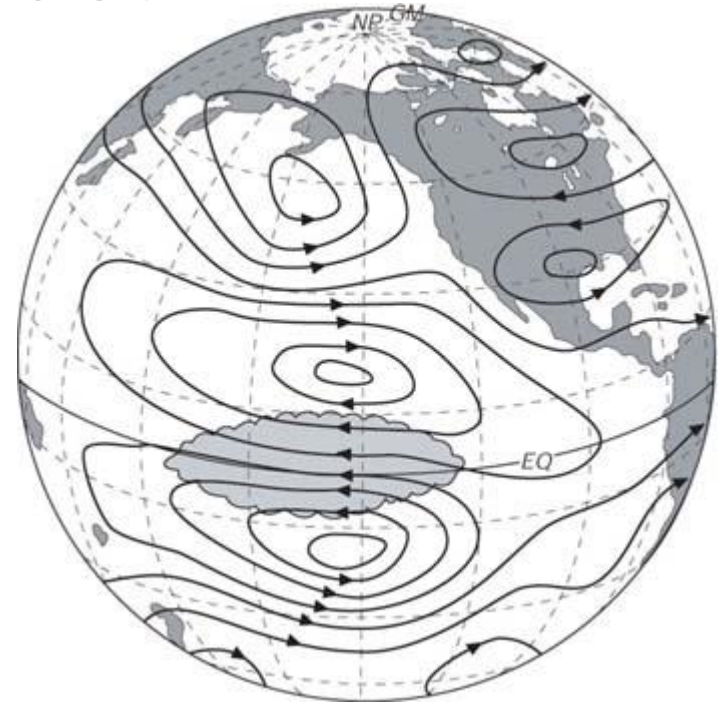
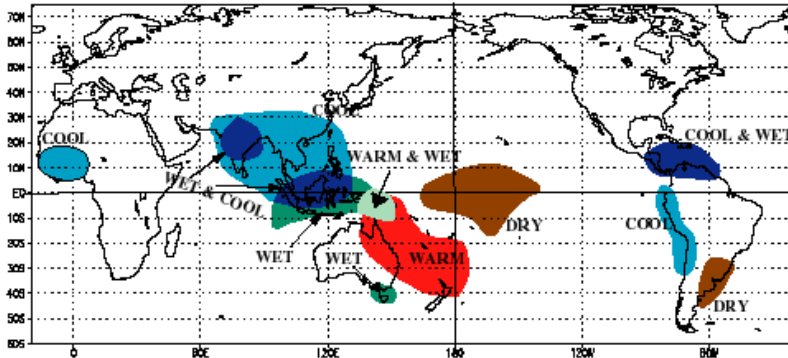
(b) El Niño Conditions

ENSO impact

COLD EPISODE RELATIONSHIPS DECEMBER - FEBRUARY



COLD EPISODE RELATIONSHIPS JUNE - AUGUST



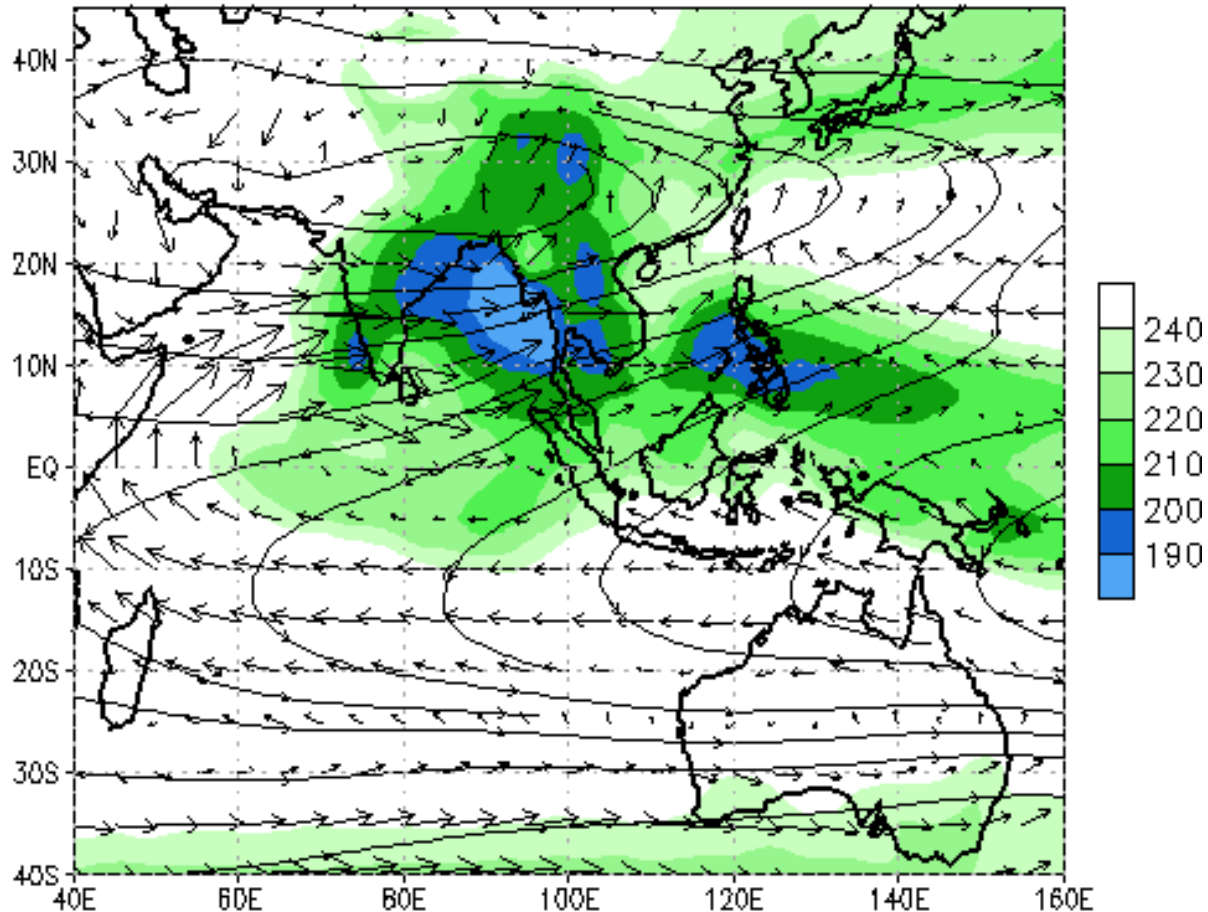
Precipitation (peak El Niño)
 lower than normal normal higher than normal



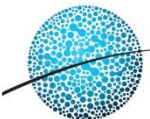
Monsoon

OLR, 200-hPa Streamlines and 850-hPa Wind Clim (1979-1995)

02JUL



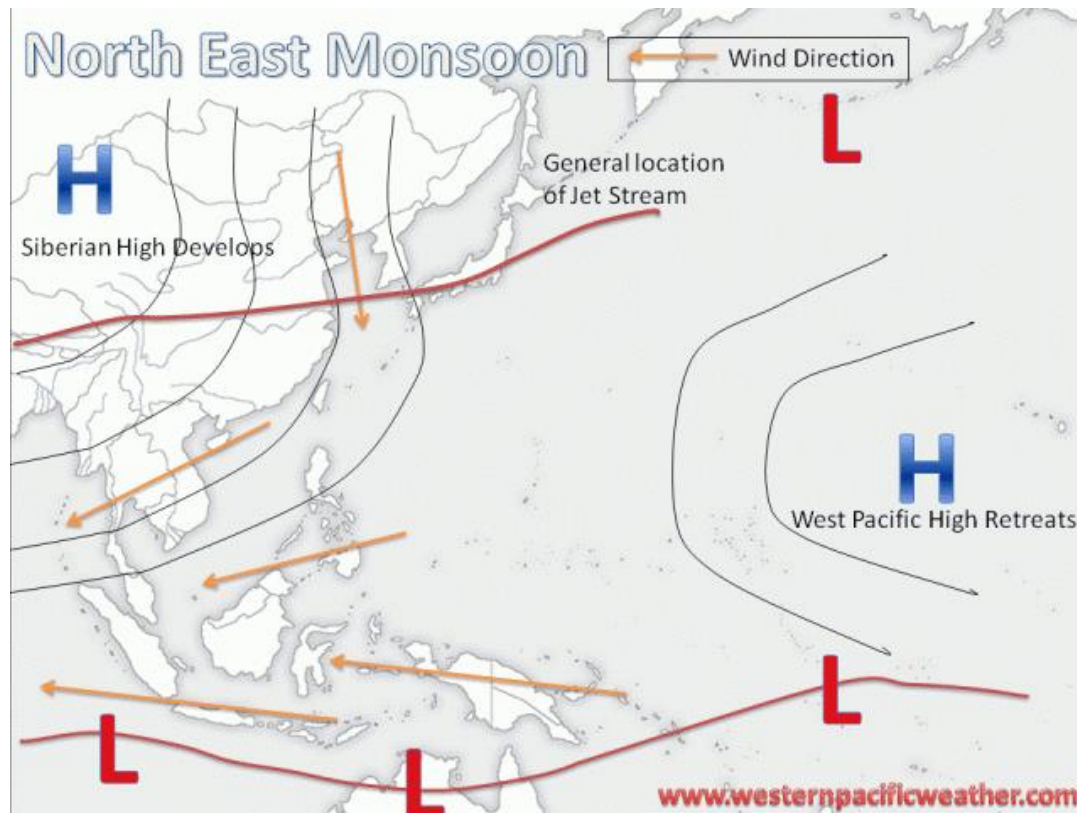
Data Sources: OLR - NESDIS/ORA, Winds - NCEP CDAS/ Reanalysis



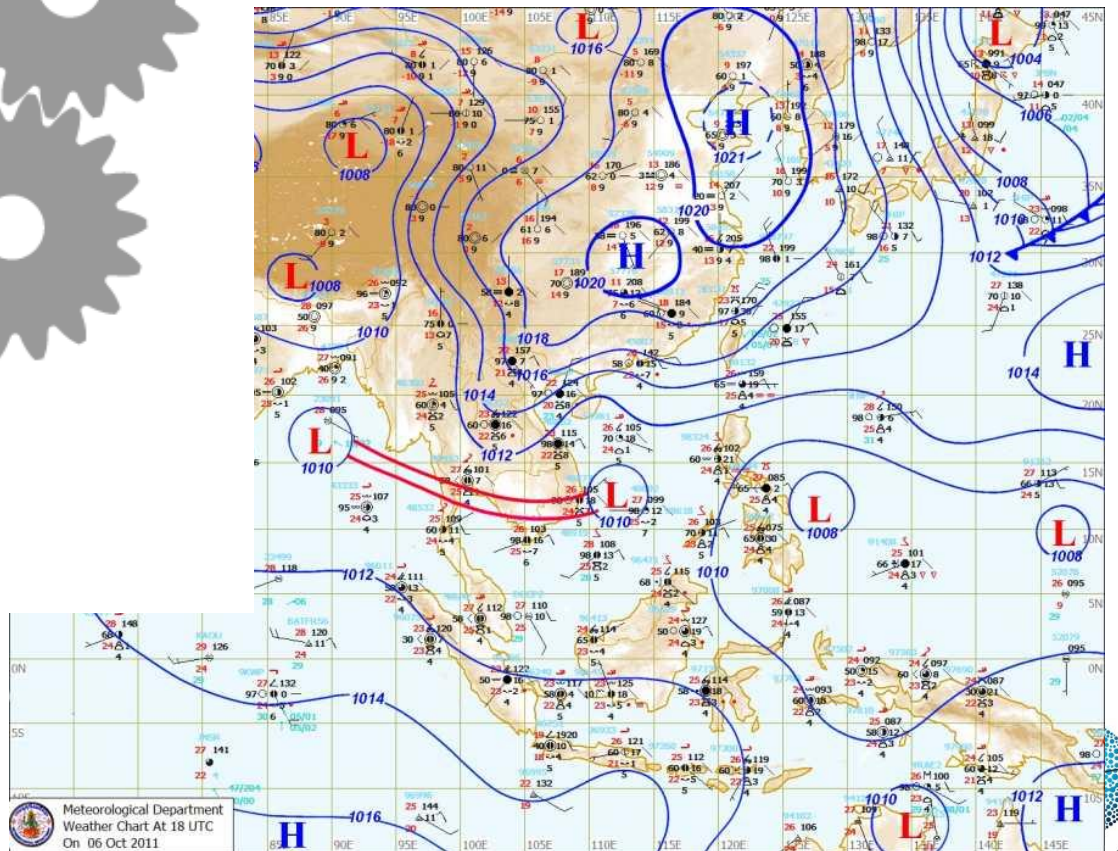
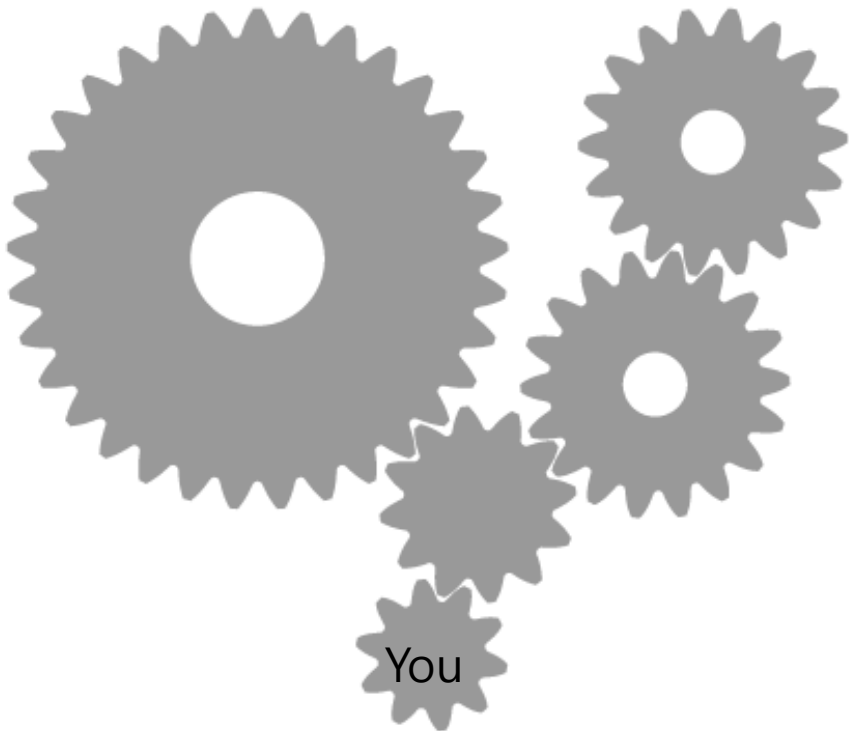
APCC
APEC CLIMATE CENTER

Cold surge

- Outbreak of Siberian cold air



How they change weather?



Seasonal Prediction (1)

Jin Ho Yoo
APEC Climate Center

Overview

- Predictability
- Methods
- Verification
- Operation

Seasonal prediction

- **Target : seasonal weather statistics with a few months lead time**
 - Mostly, seasonal or monthly mean Temp. or Prec.
- **Why we do this? : for planning (serious)**

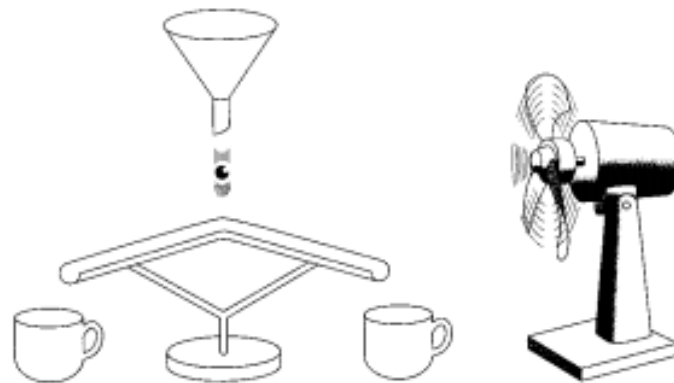
letters to nature

.....
**Forecasting Andean rainfall and
crop yield from the influence of
El Niño on Pleiades visibility**

Benjamin S. Orlove[†], John C. H. Chiang[†] & Mark A. Cane[†]

History of Short-term (Seasonal) Climate Prediction

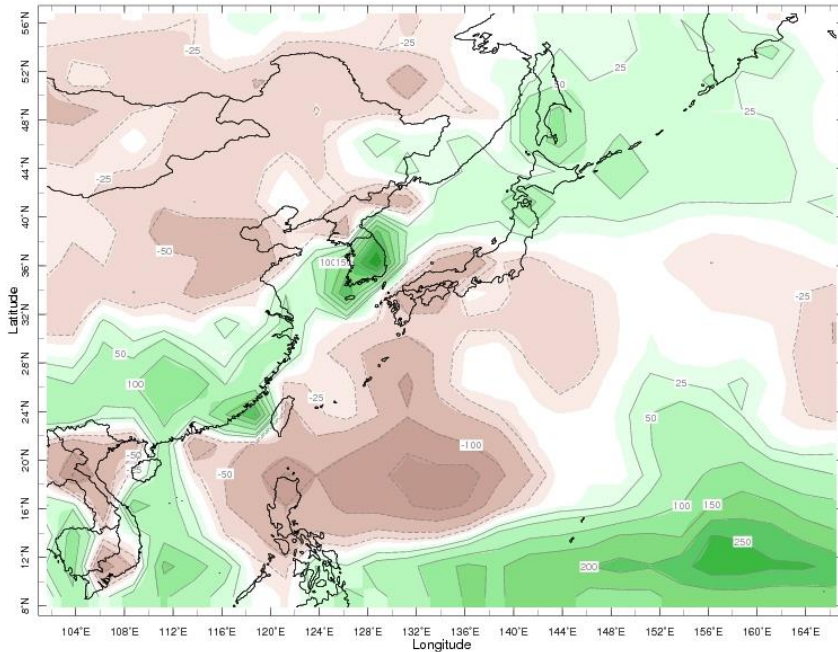
- 1960's : Hypothesis proposed
- 1980's : **ENSO** prediction + Atm. LFV. (PNA..)
- 1990's : (Experimental) Dyn. Seasonal Fcst.
- 2000's : International collaboration (MIPs)
- 2010's : Operation (GFCS, RCOFs/WMO)



T. Palmer (1998)

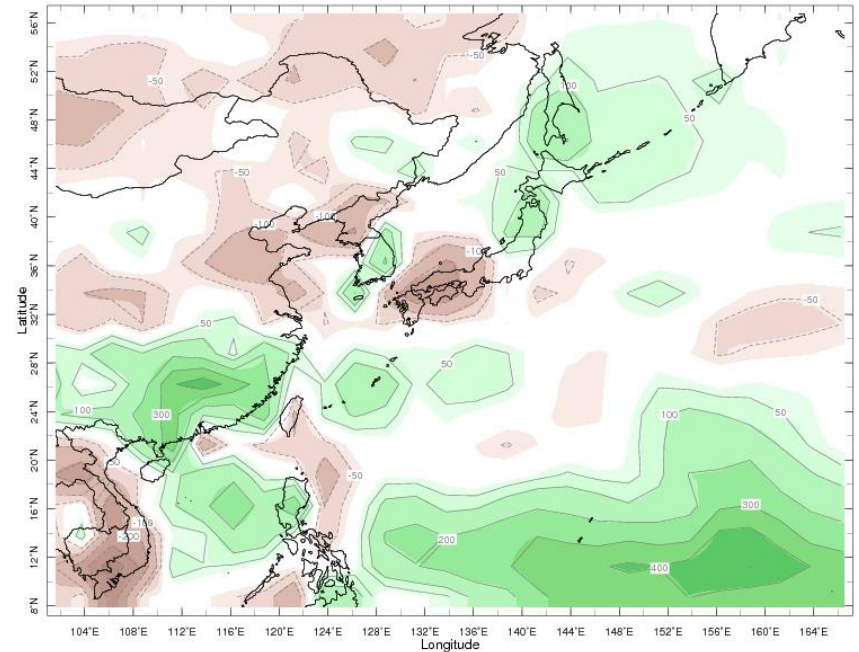
2002 summer rainfall

Monthly mean prec. (Aug)



Aug 2002

Summer mean prec.

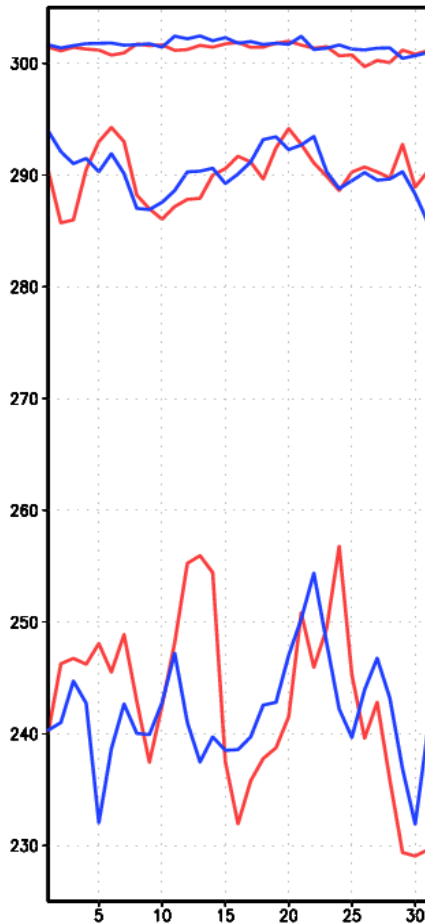


Jun-Aug 2002

Typhoon "RUSA" passed at 8/31 (1000mm a day)

Begin with Questions

How is the seasonal mean determined?

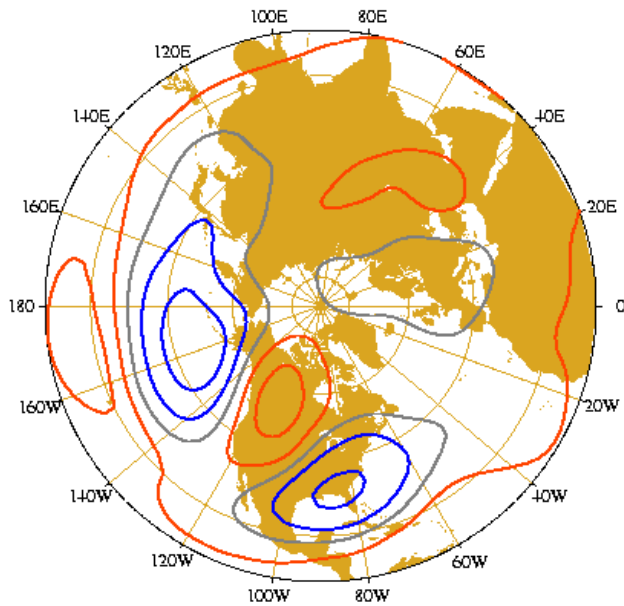


What causes change (variability) of the mean?

- By chance?
- By “something”?



PNA debates

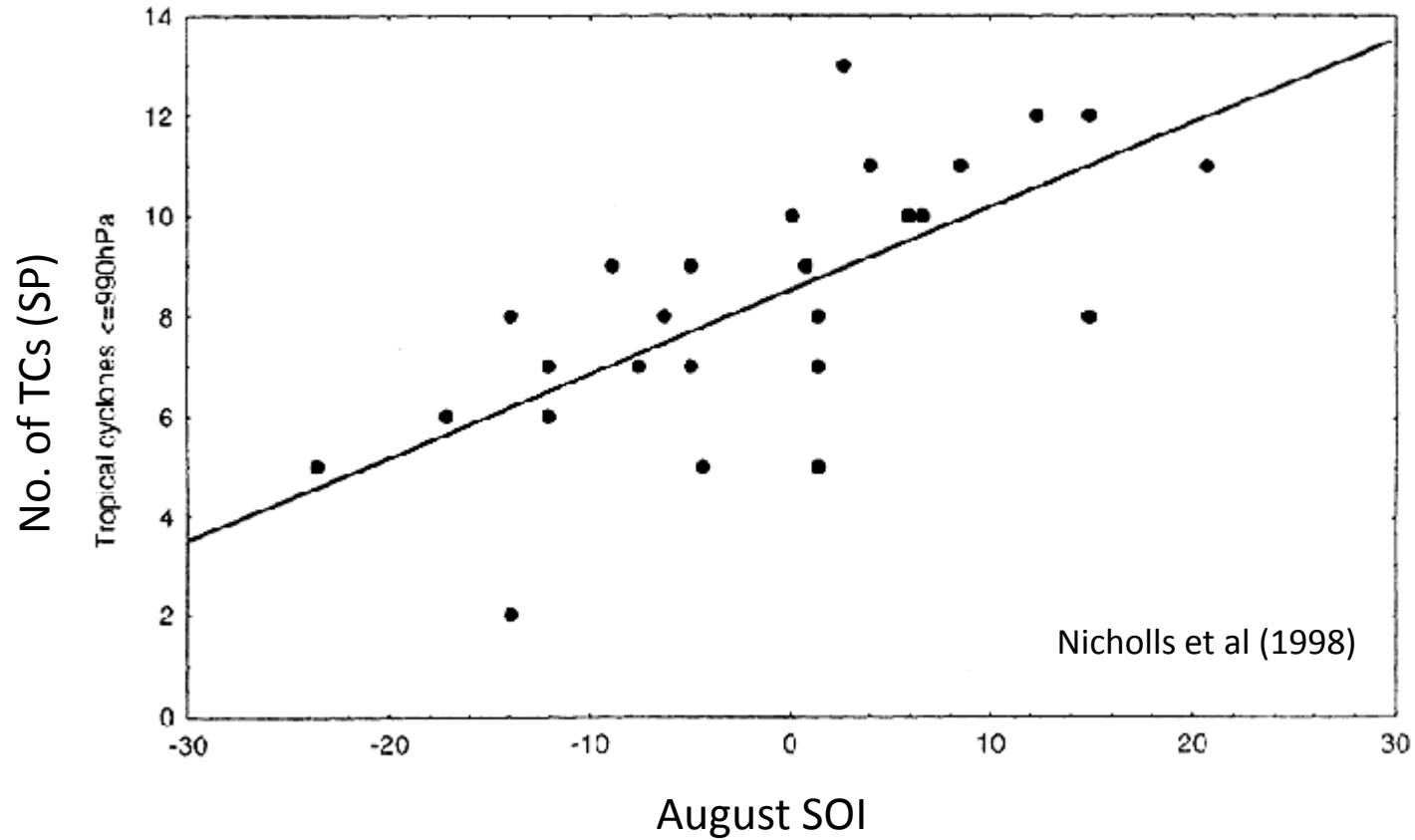


1. Forced by El Nino
2. Atmospheric internal variability (random)

Two scales

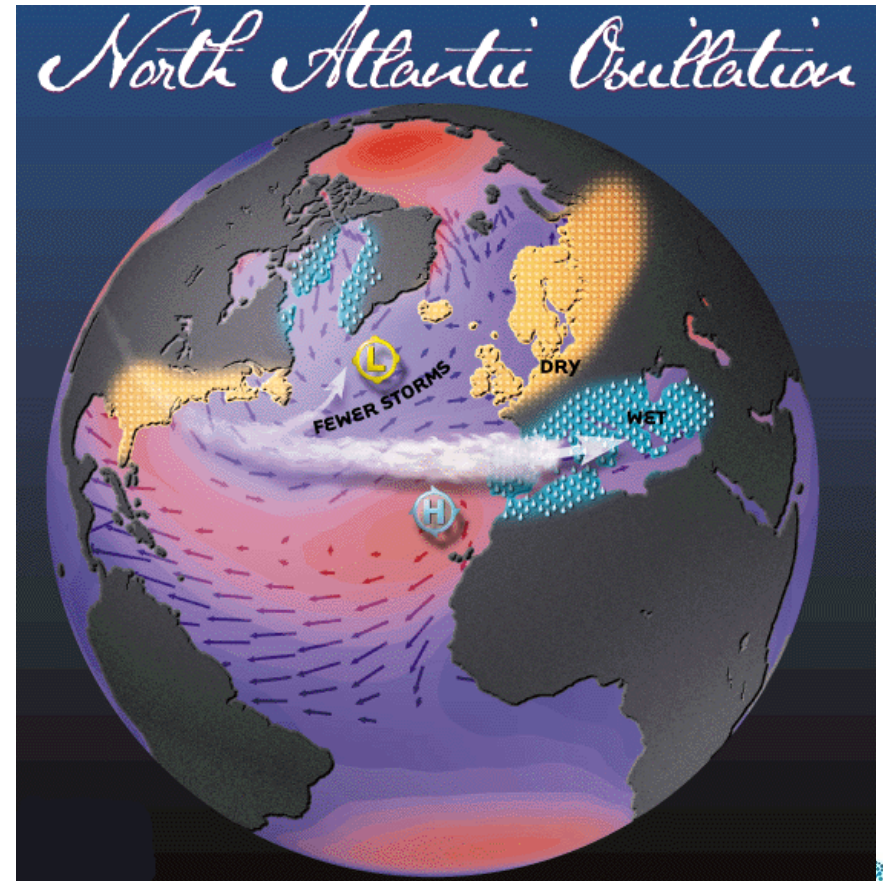
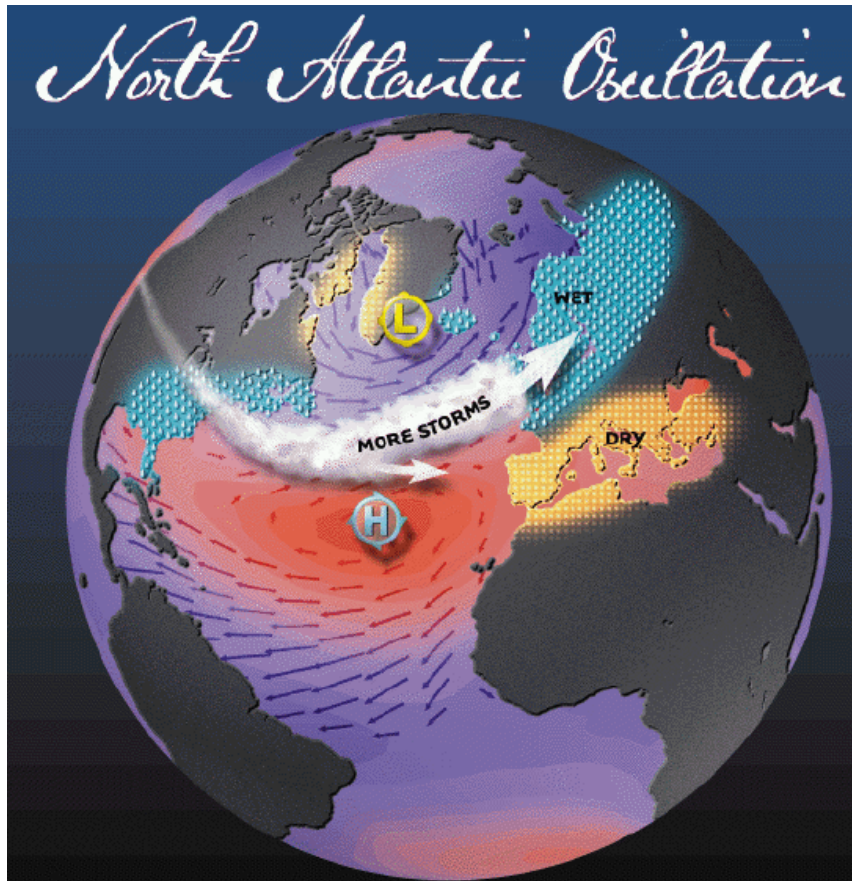
- **Fast and small** scale processes
 - Weather, Tropical cyclone
- **Slow and large** processes
 - Climate, ITCZ, ENSO
- Slow process consist of Fast processes but if it has its own slow dynamics, it can **controls** Fast processes

Two scales

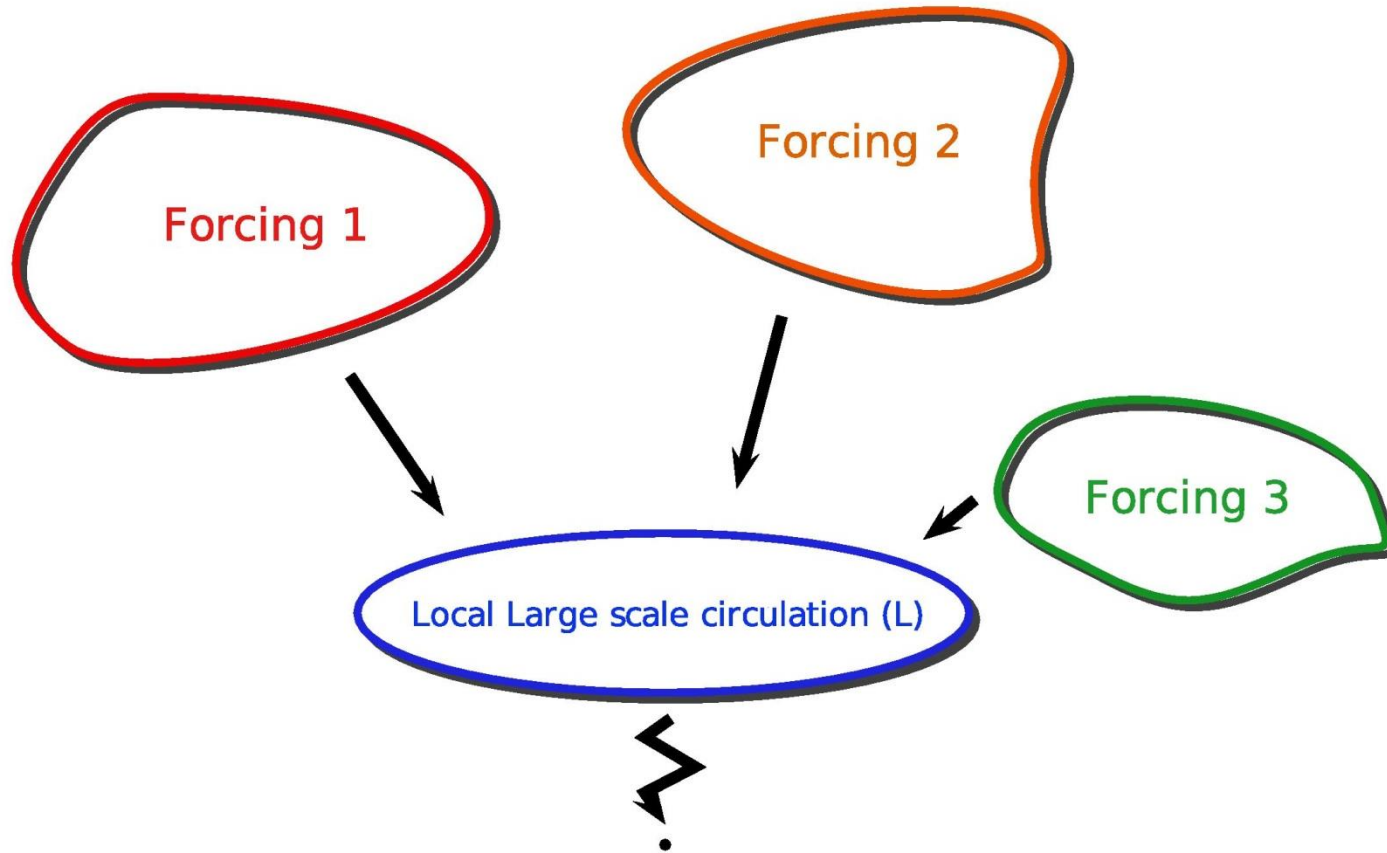


NAO(north Atlantic oscillation)

NAO changes passage of storms (weather system)



Local large scale circulation

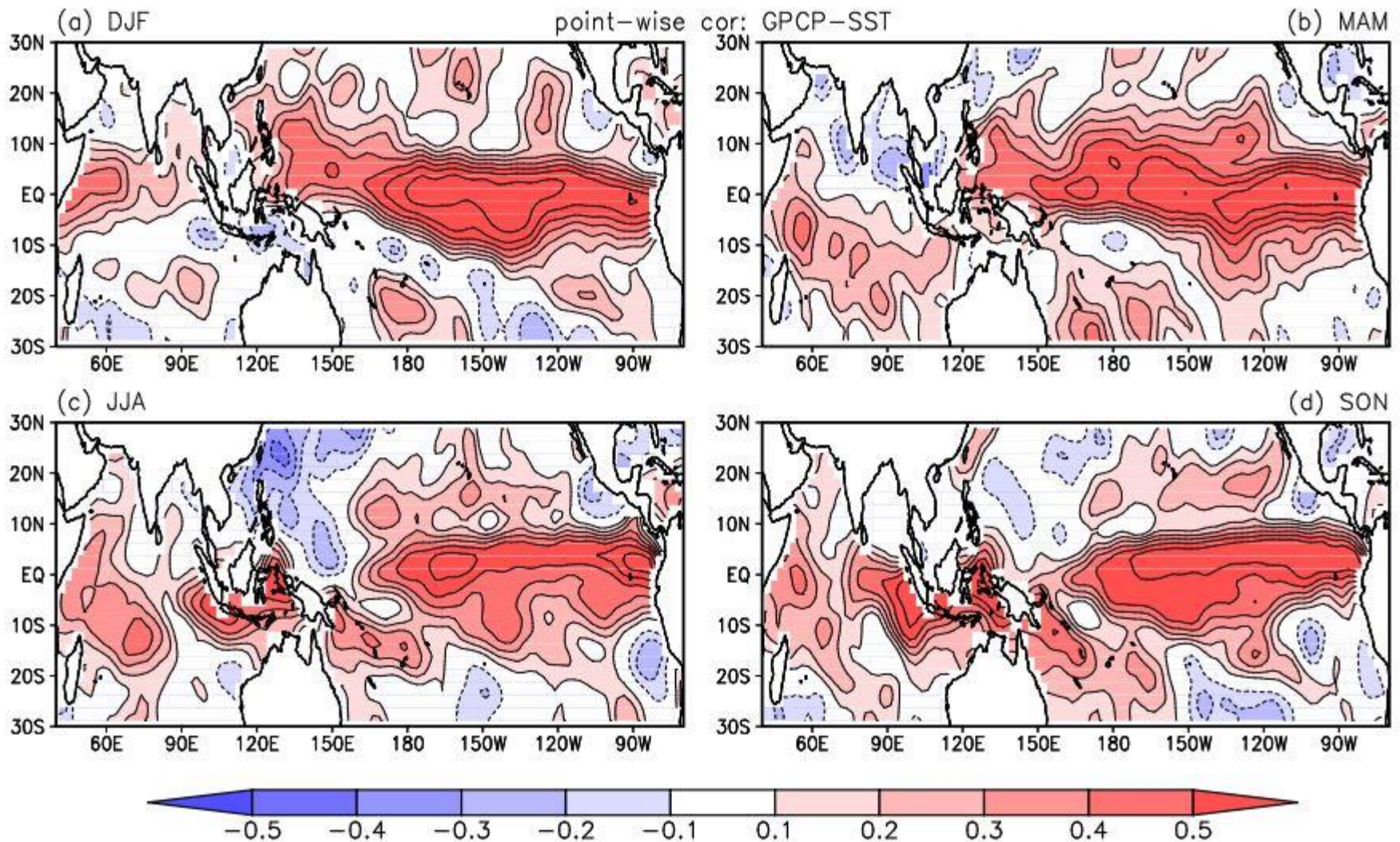


Local weather statistics (Korean summer rainfall)

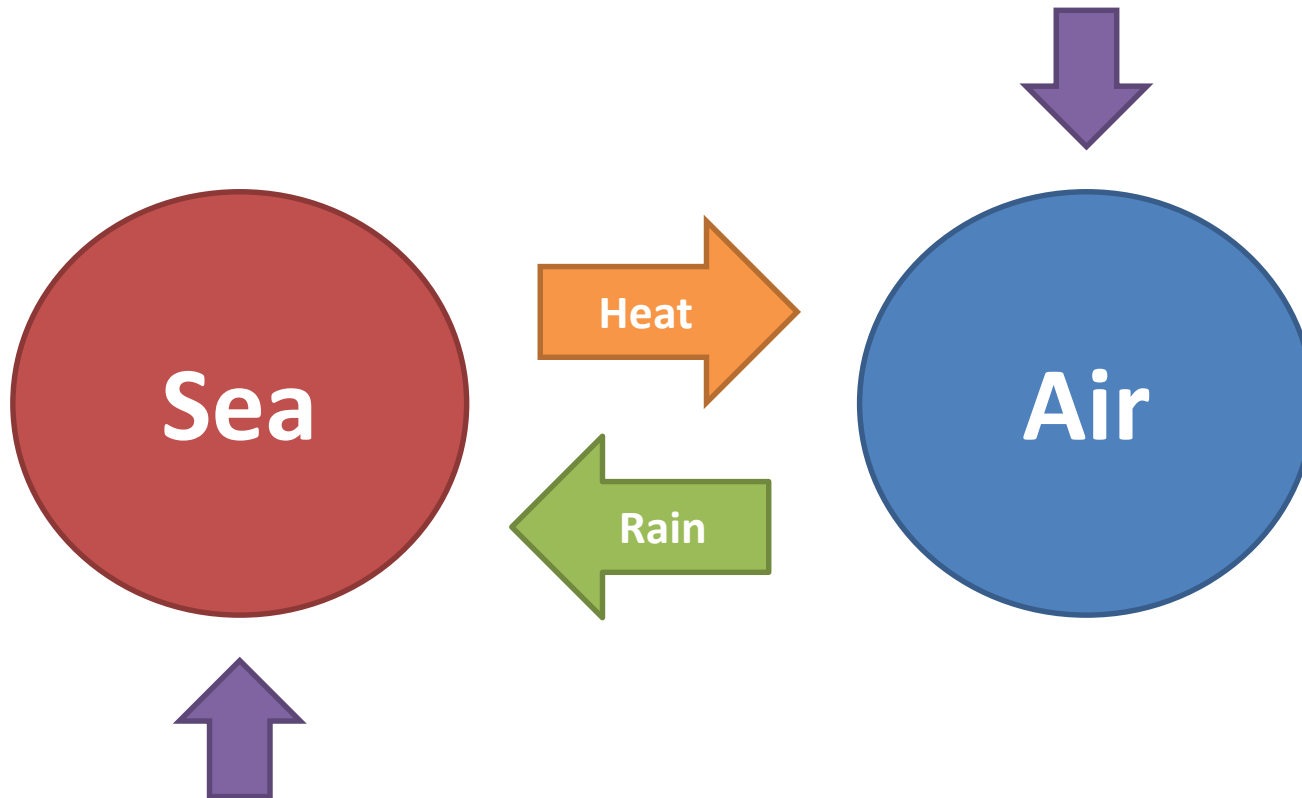
Local large scale circulation

- Local climate (i.e. seasonal mean) is defined by how weather behaved during a season (statistics)
- Therefore, understanding weather behavior is the first step of seasonal forecast (often ignored..)
- In many cases, local large scale pattern that directly affect local weather is visible in seasonal time scale
 - Question is whether we can predict that large scale pattern directly or via teleconnection

Do not miss : air-sea interaction



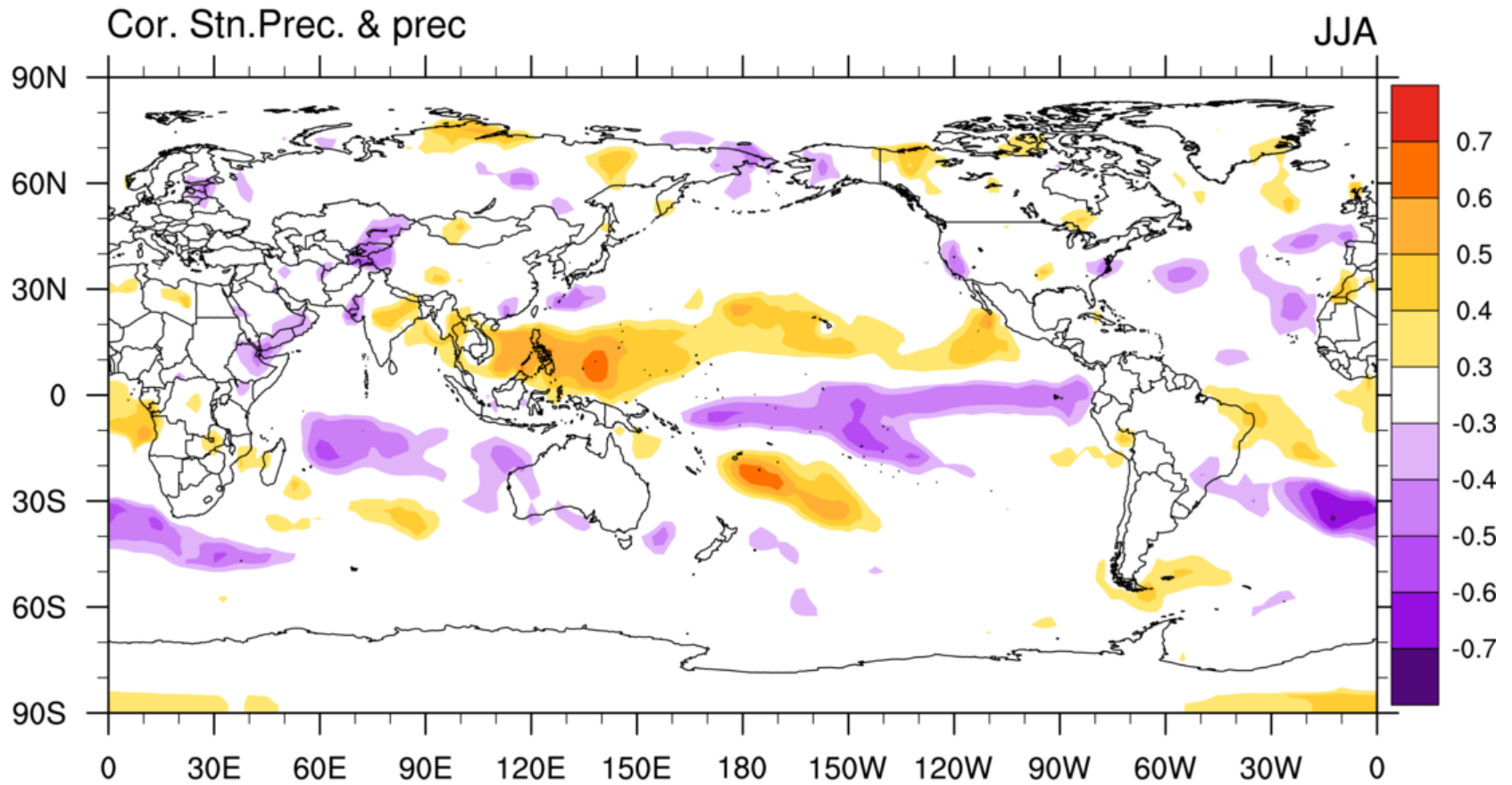
Feedbacks



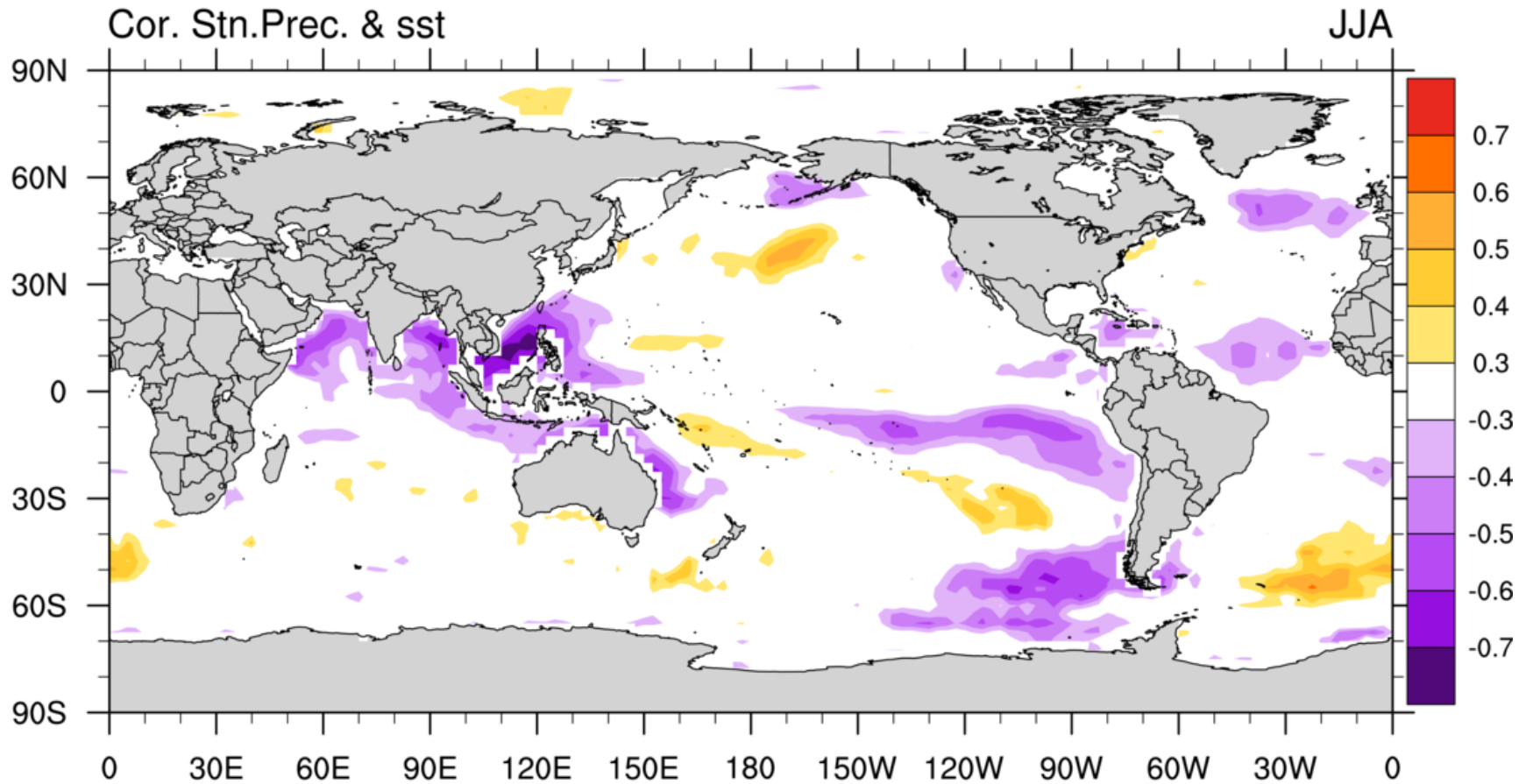
Local large circulation and Teleconnection

LARGE SCALE PATTERN ASSOCIATED WITH RAINFALL

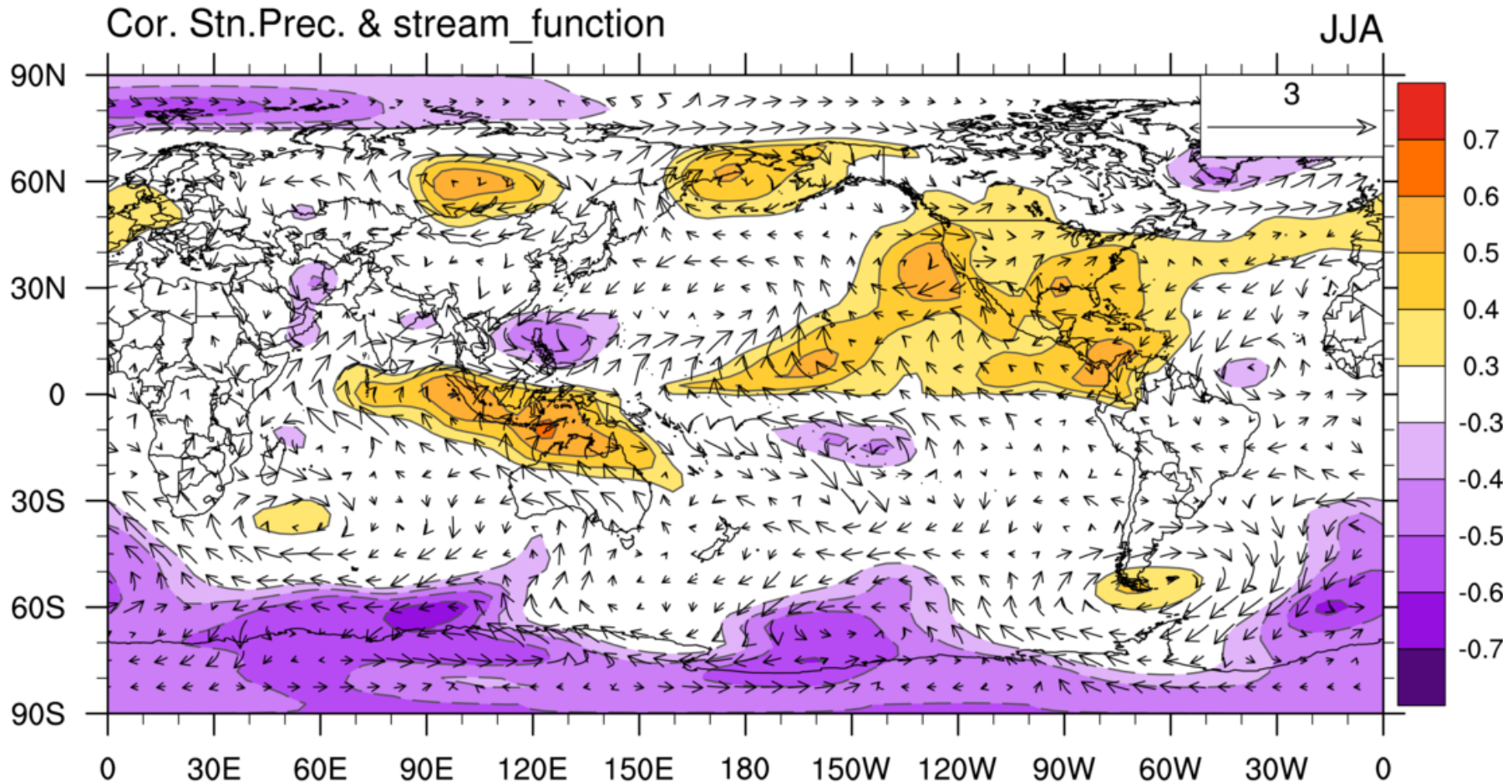
Manila (JJA)



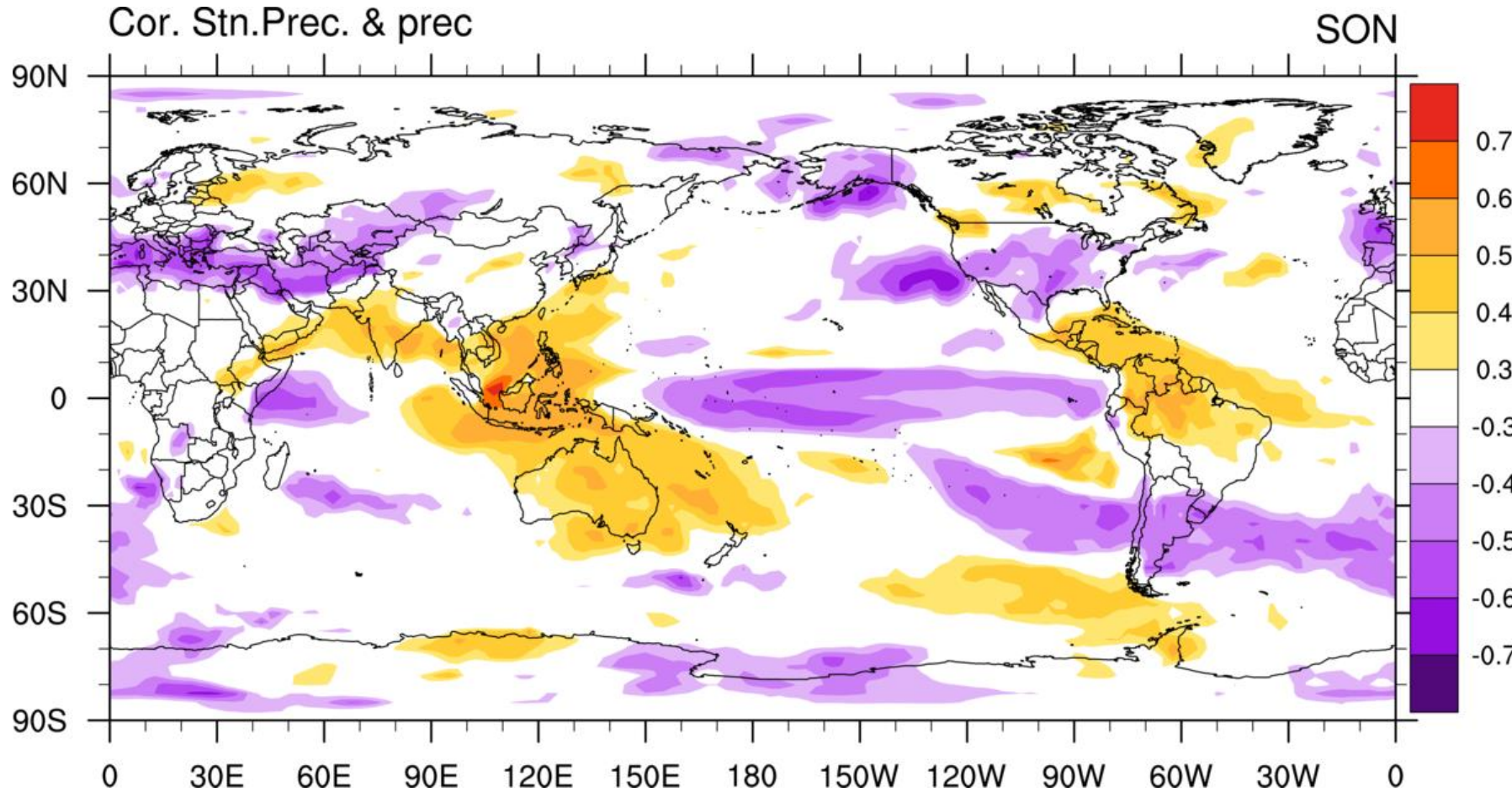
Manila (JJA)



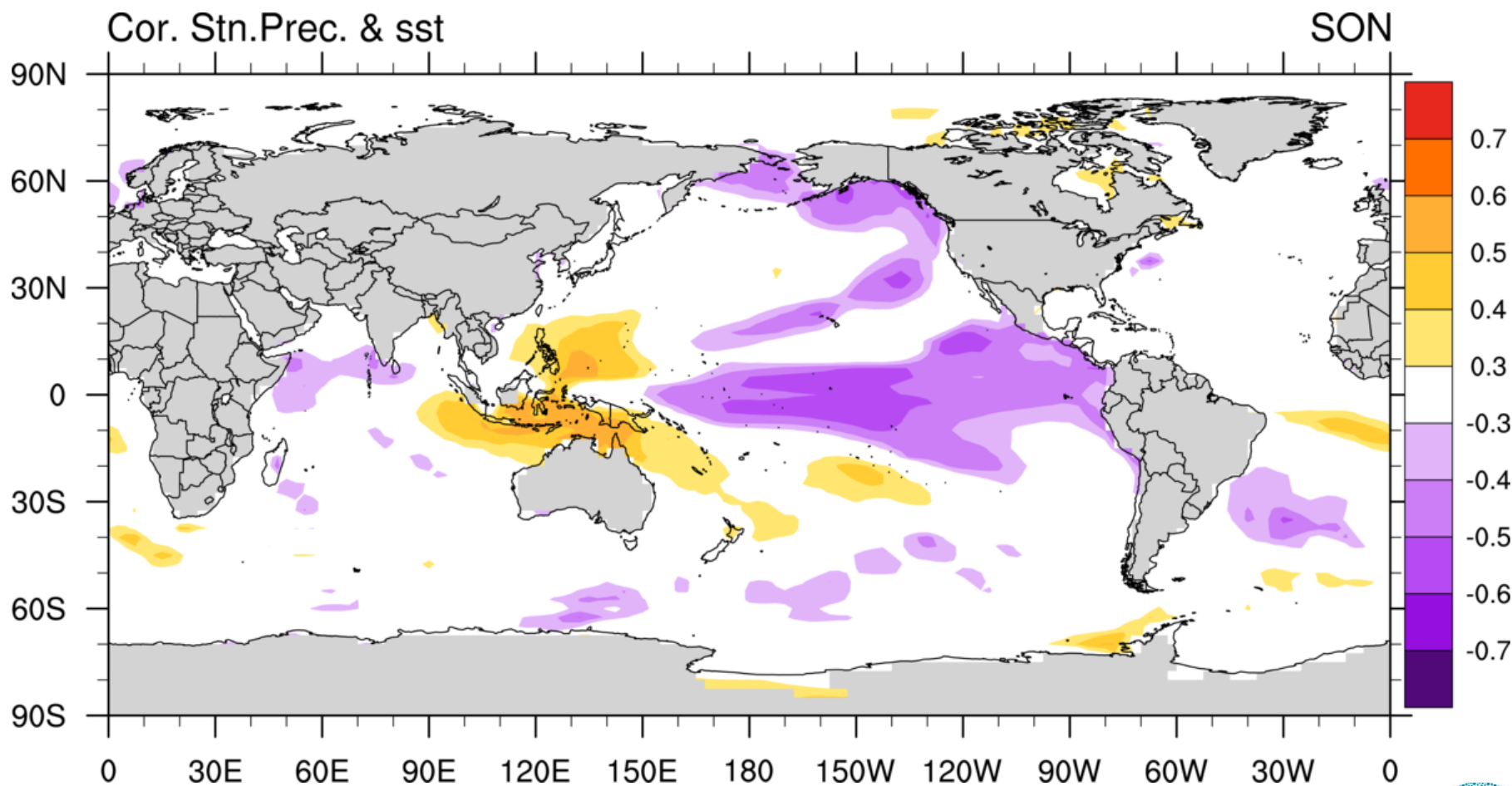
Manila (JJA)



Manila (SON)



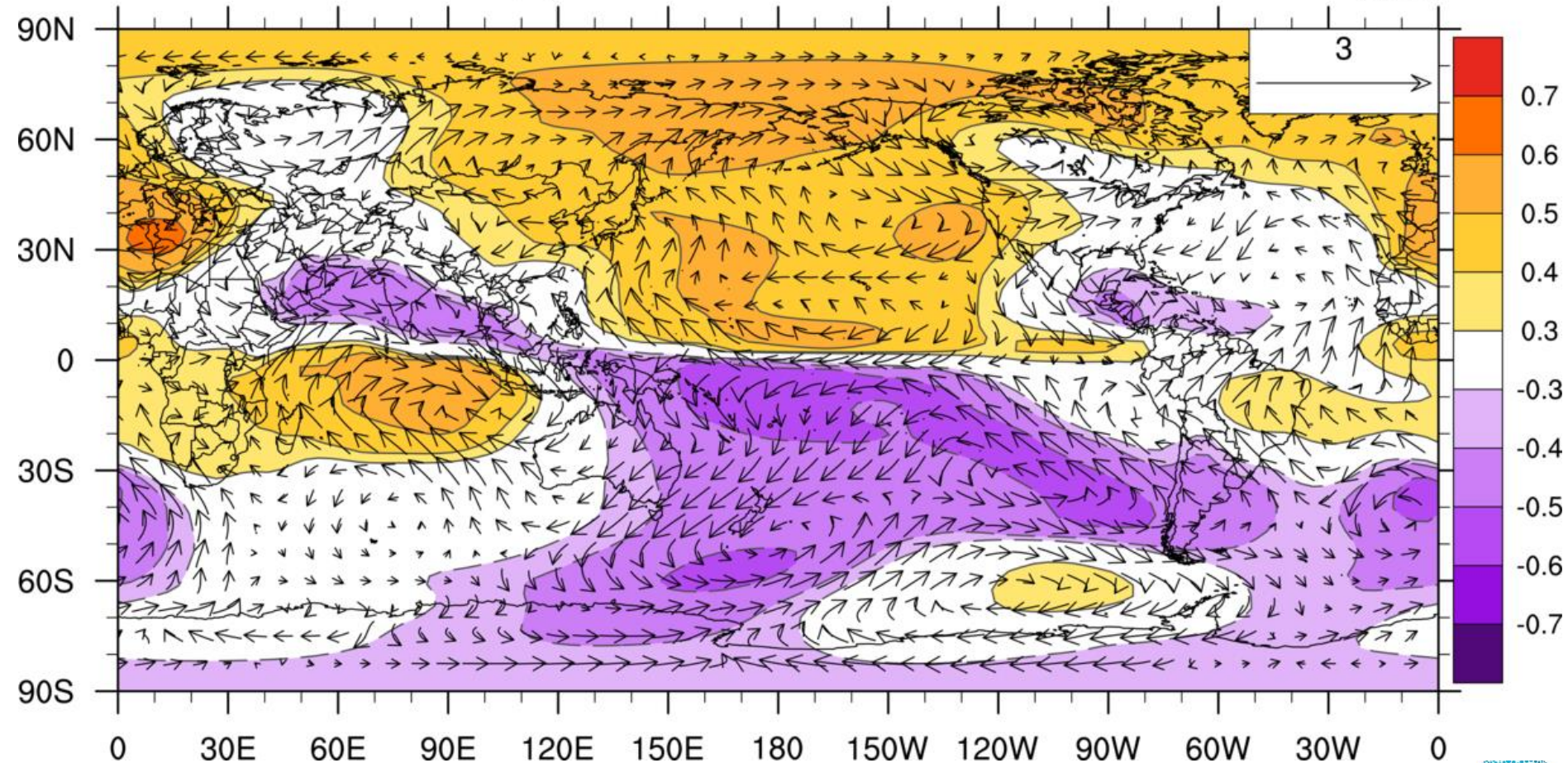
Manila (SON)



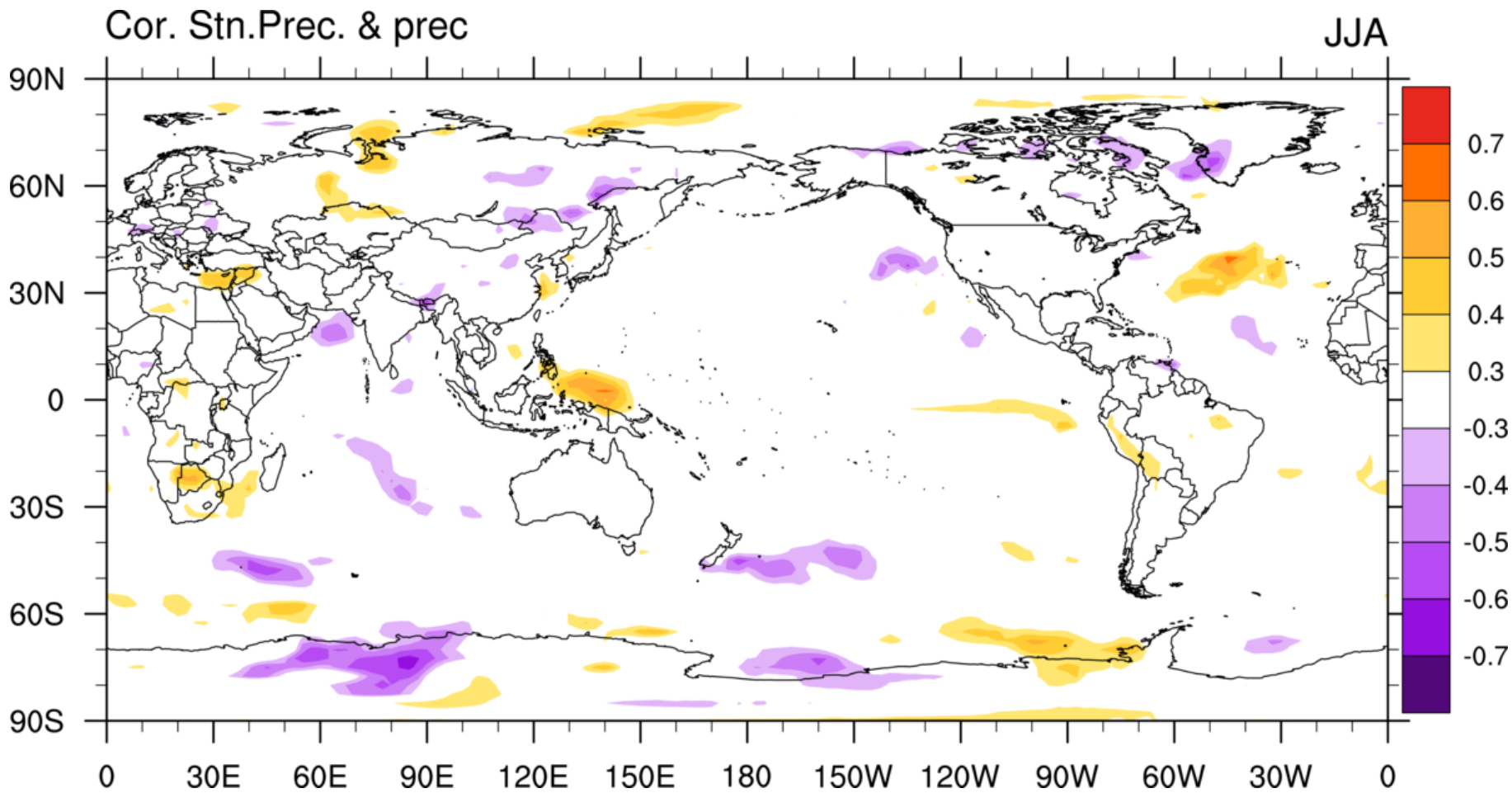
Manila (SON)

Cor. Stn.T2M & stream_function

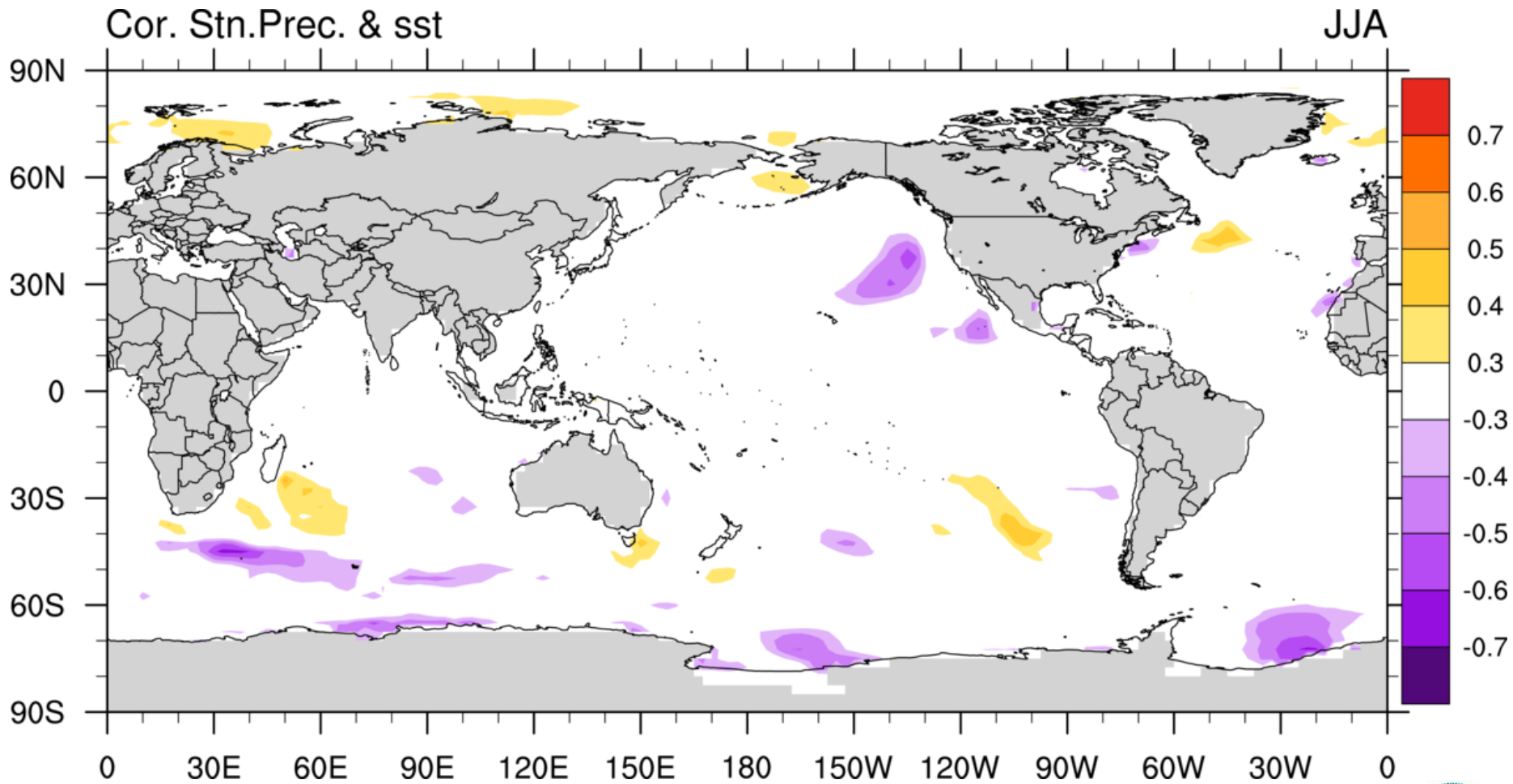
SON



Davao (JJA)



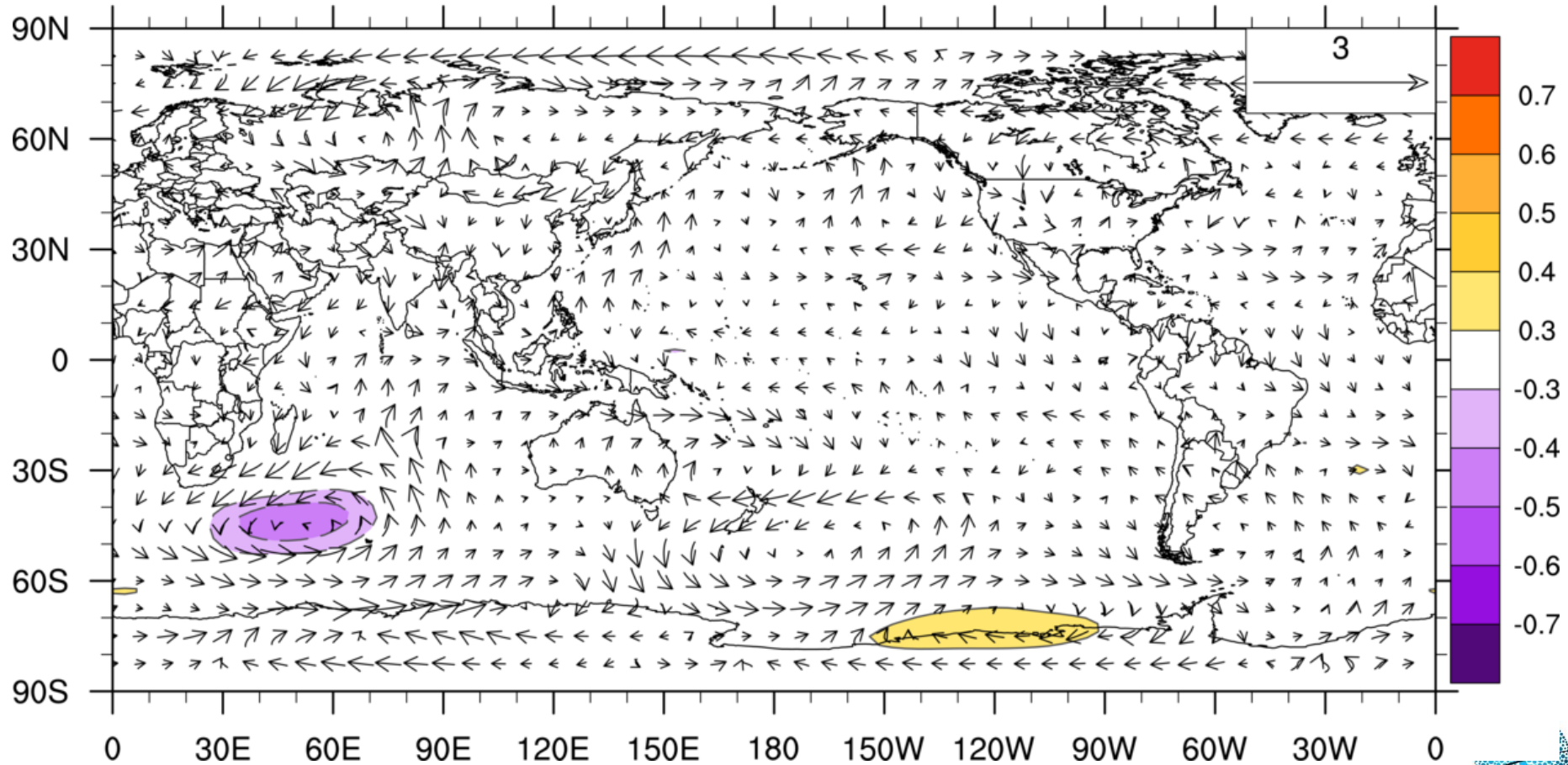
Davao (JJA)



Davao (JJA)

Cor. Stn.T2M & stream_function

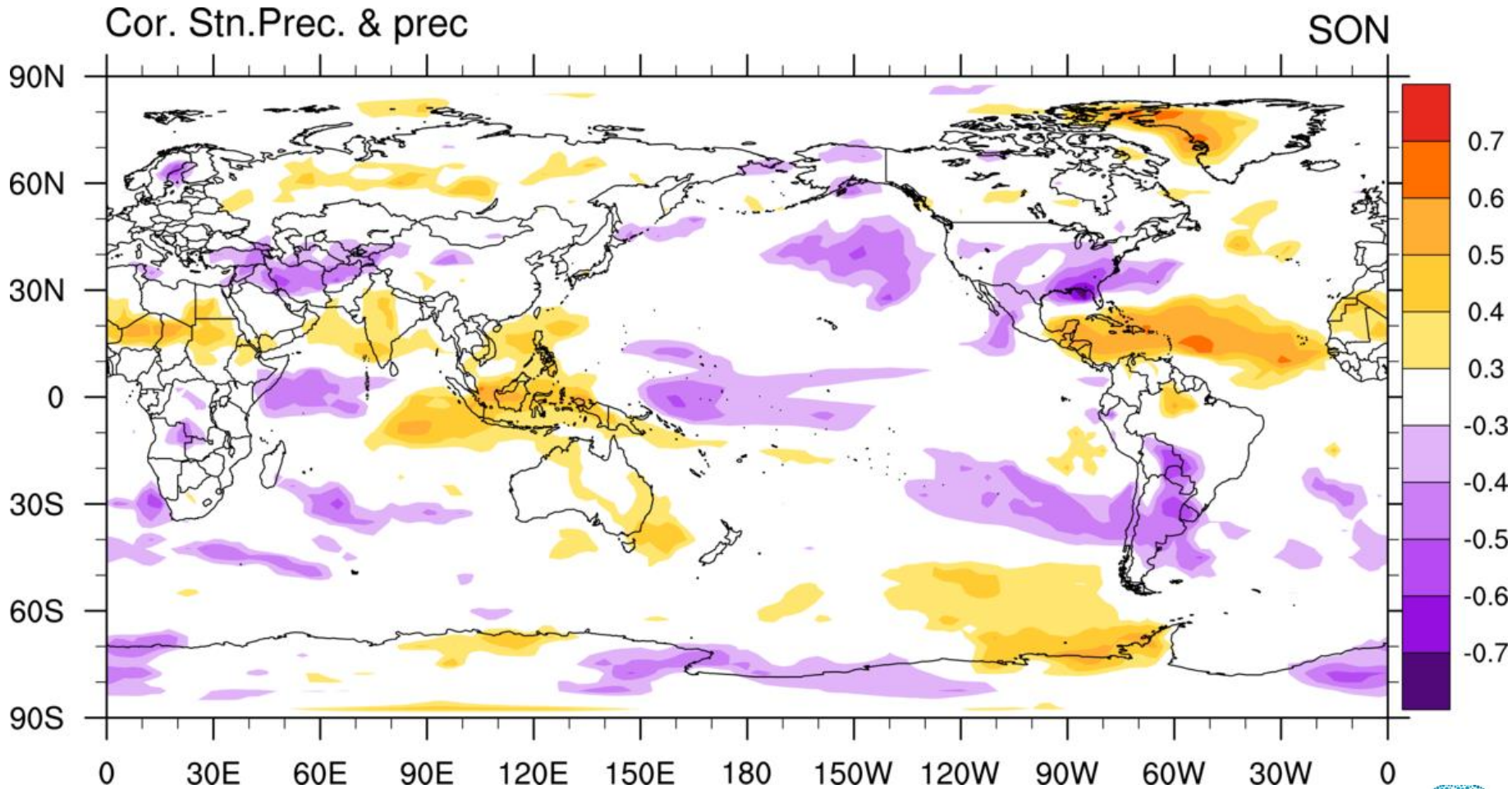
JJA



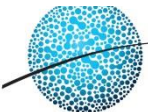
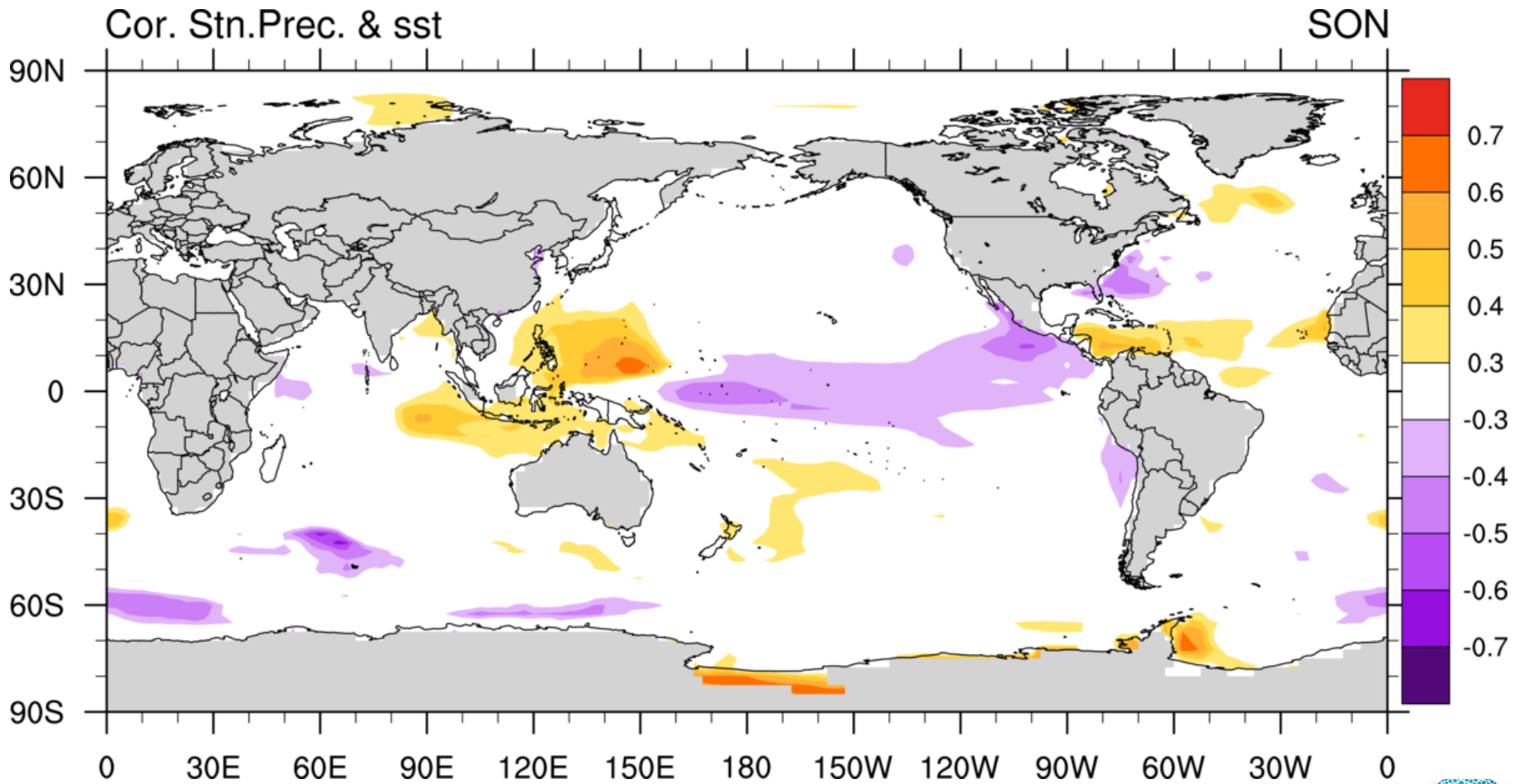
Davao City enjoys a mild tropical climate. Compared with other parts of the Philippines in which there is a distinct hot and wet season. **It enjoys the privilege of a climate where the days are always sunny and mild followed by nights of rain.** The city is outside the typhoon belt and **lacks major seasonal variations.** The predominant wind direction is northward from the Davao Gulf where the cooler air of the sea replaces the warm air mass over the city. A surrounding mountain chains protect the city effectively from strong winds.

NOTHING to PREDICT?

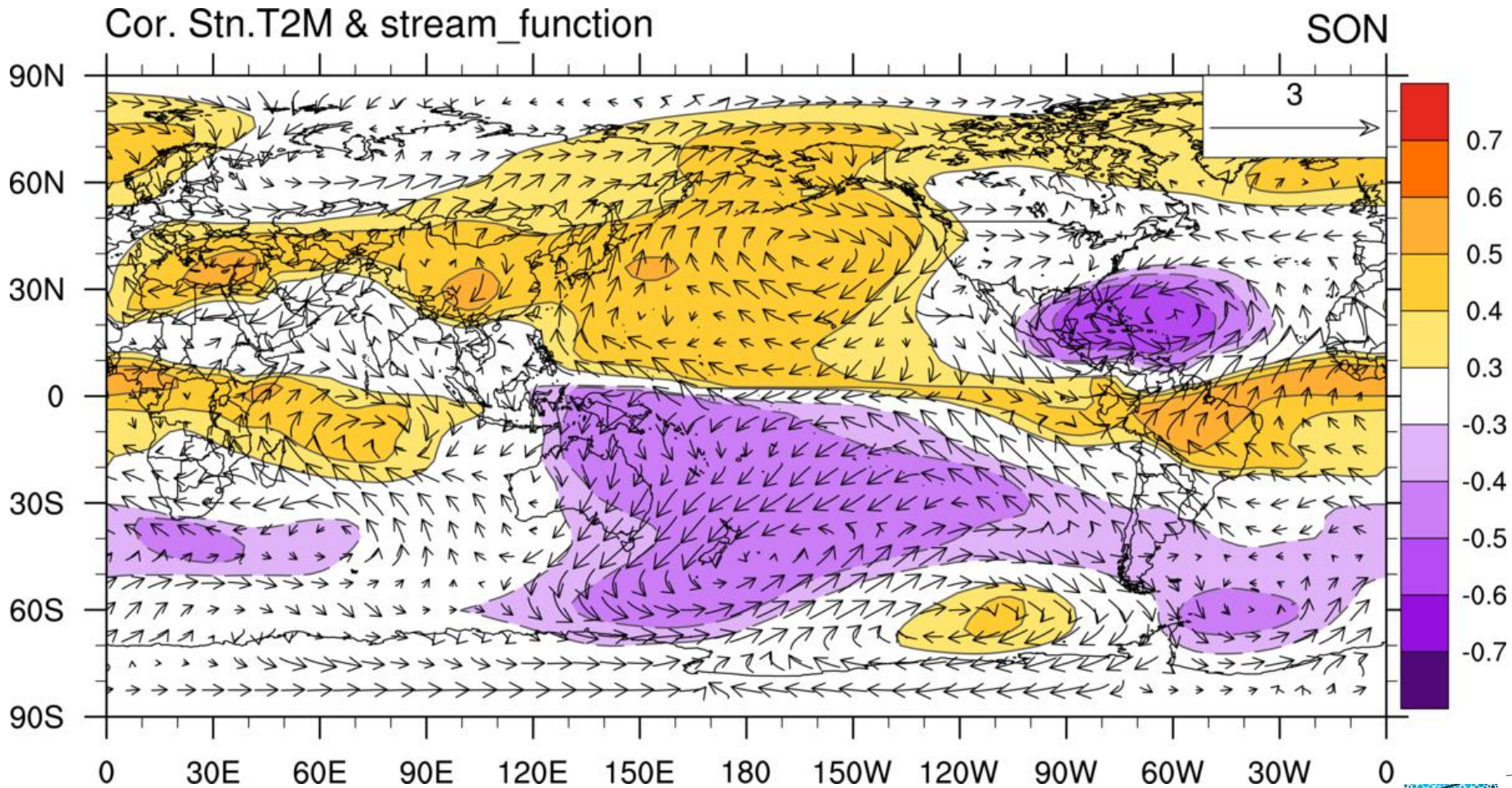
Davao (SON)



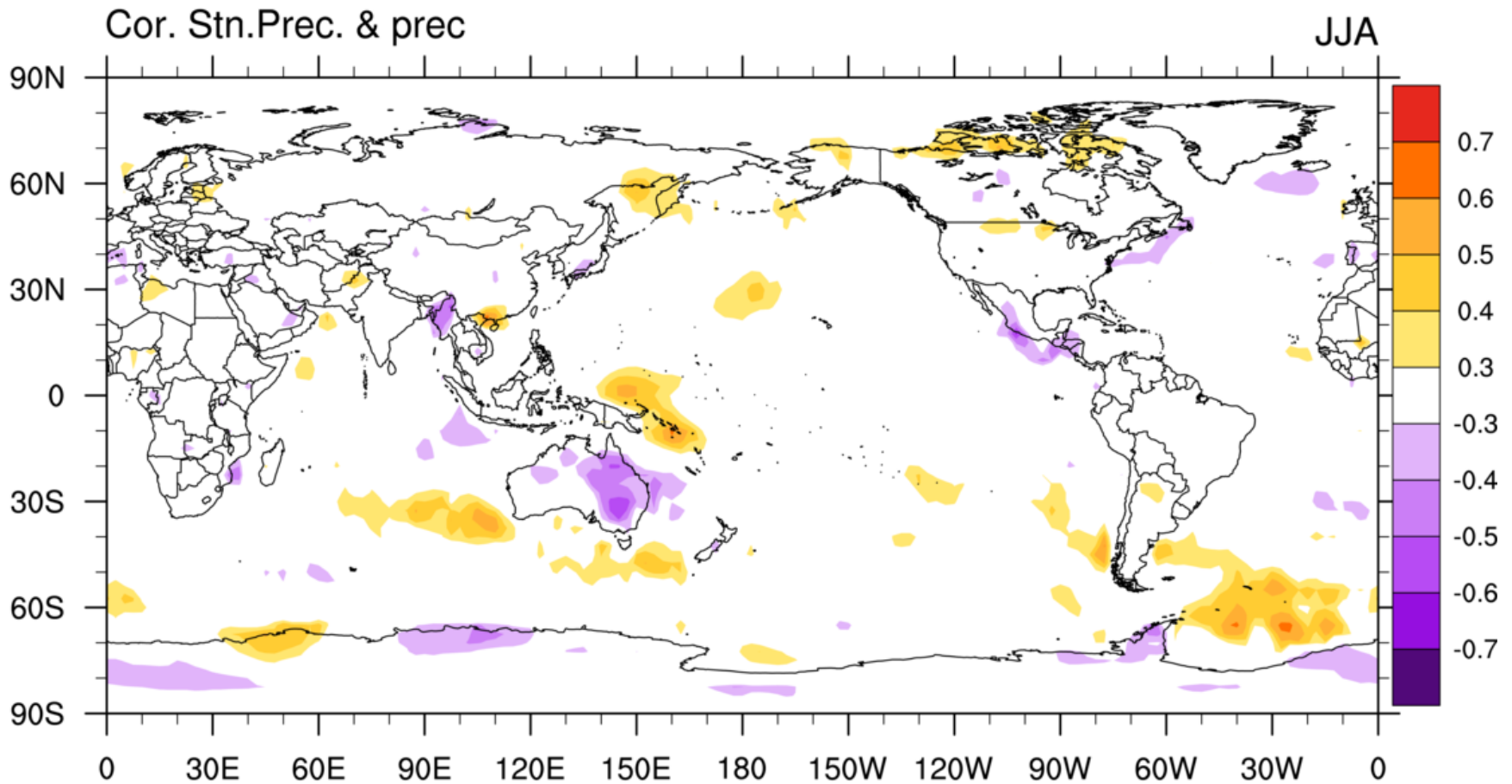
Davao (SON)



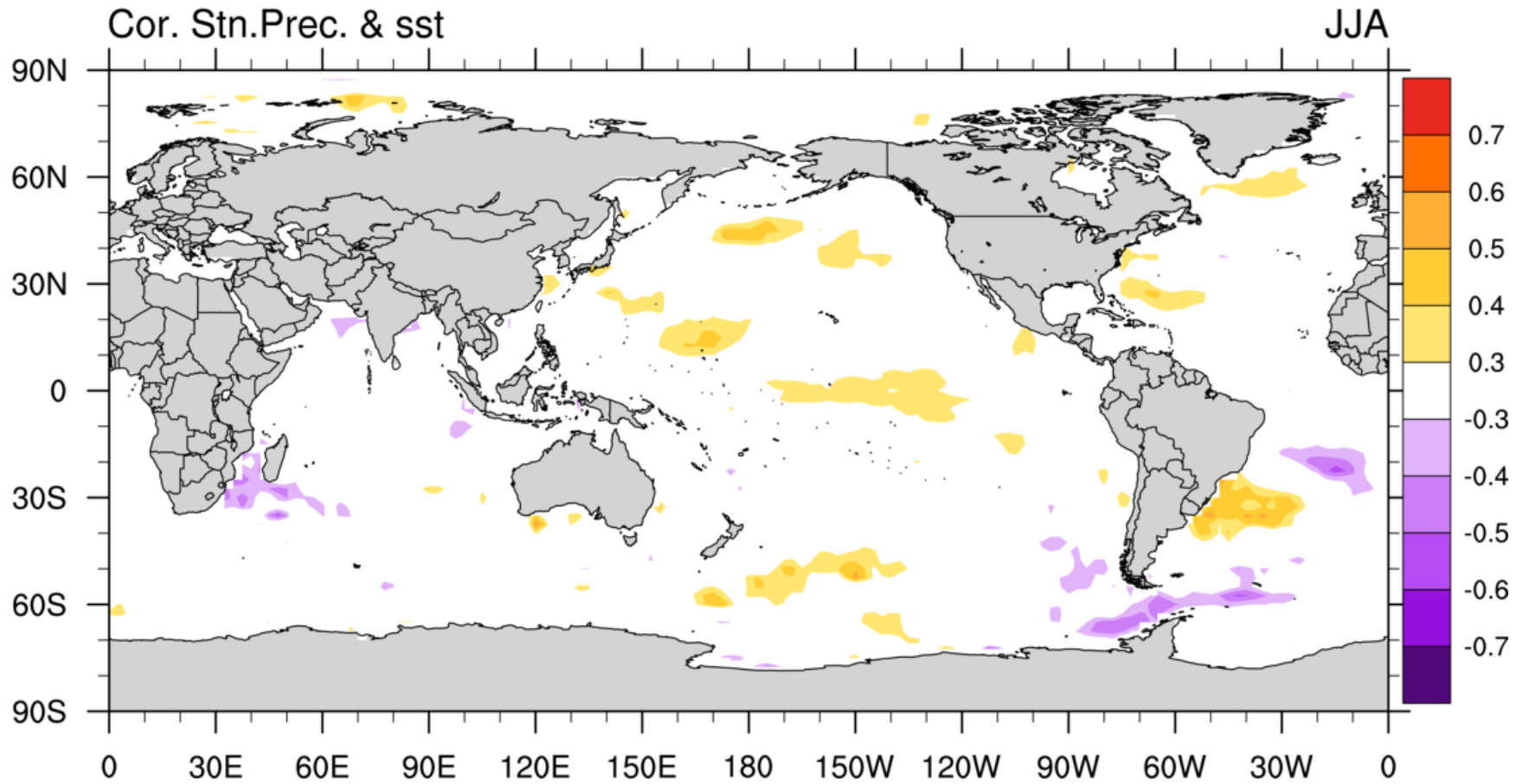
Davao (JJA)



Hanoi (JJA)



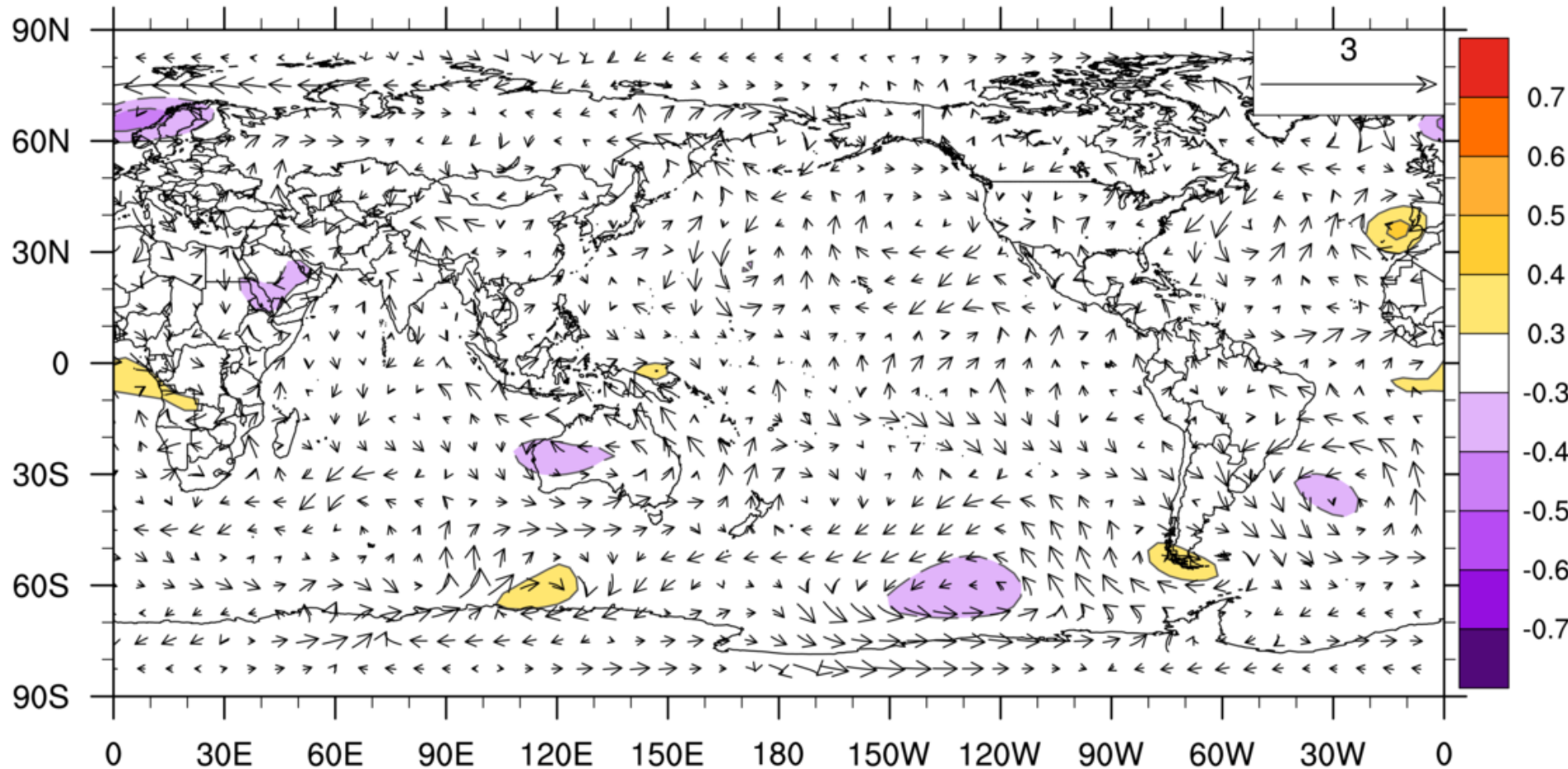
Hanoi (JJA)



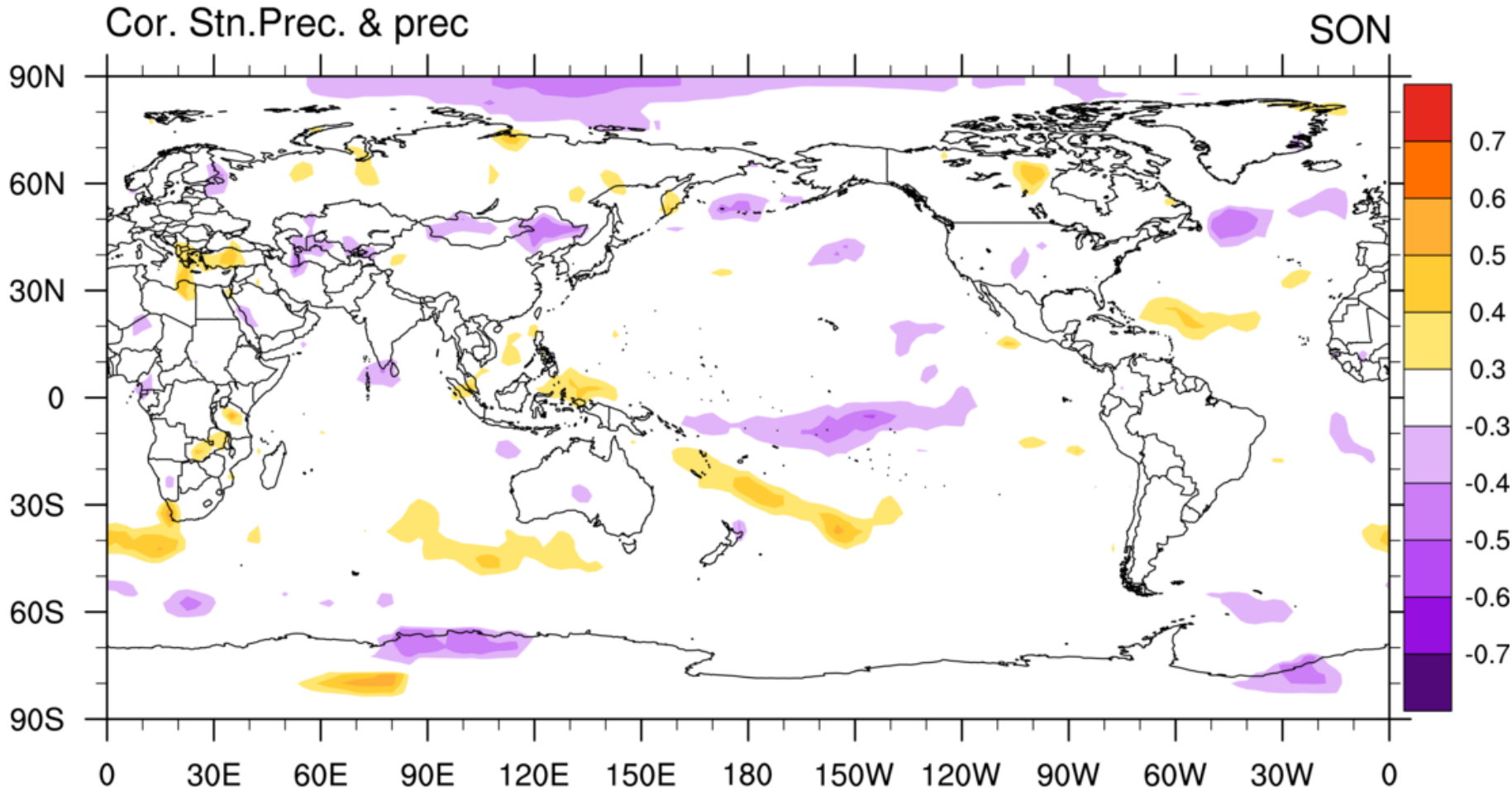
Hanoi (JJA)

Cor. Stn. Prec. & stream_function

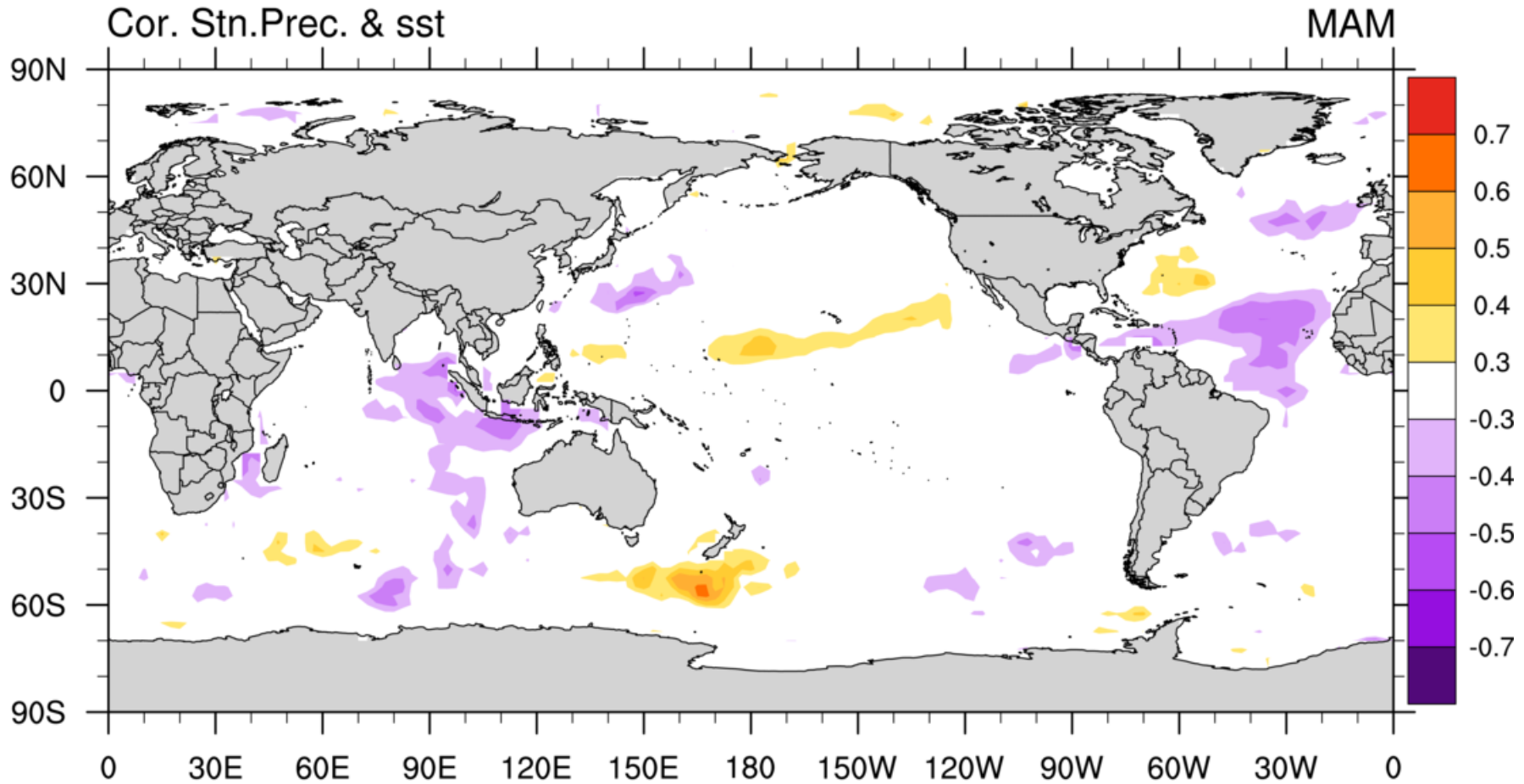
JJA



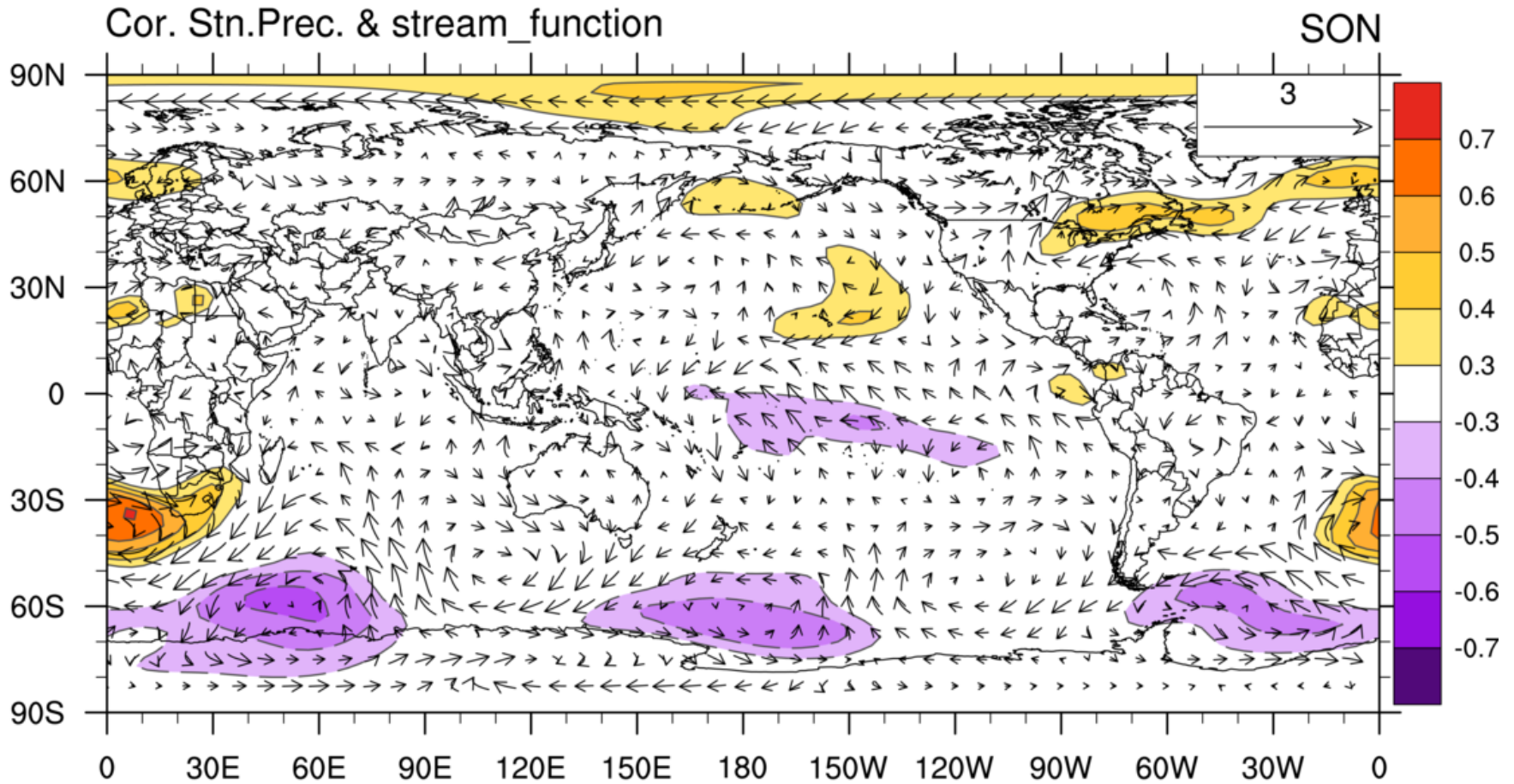
Hanoi (SON)



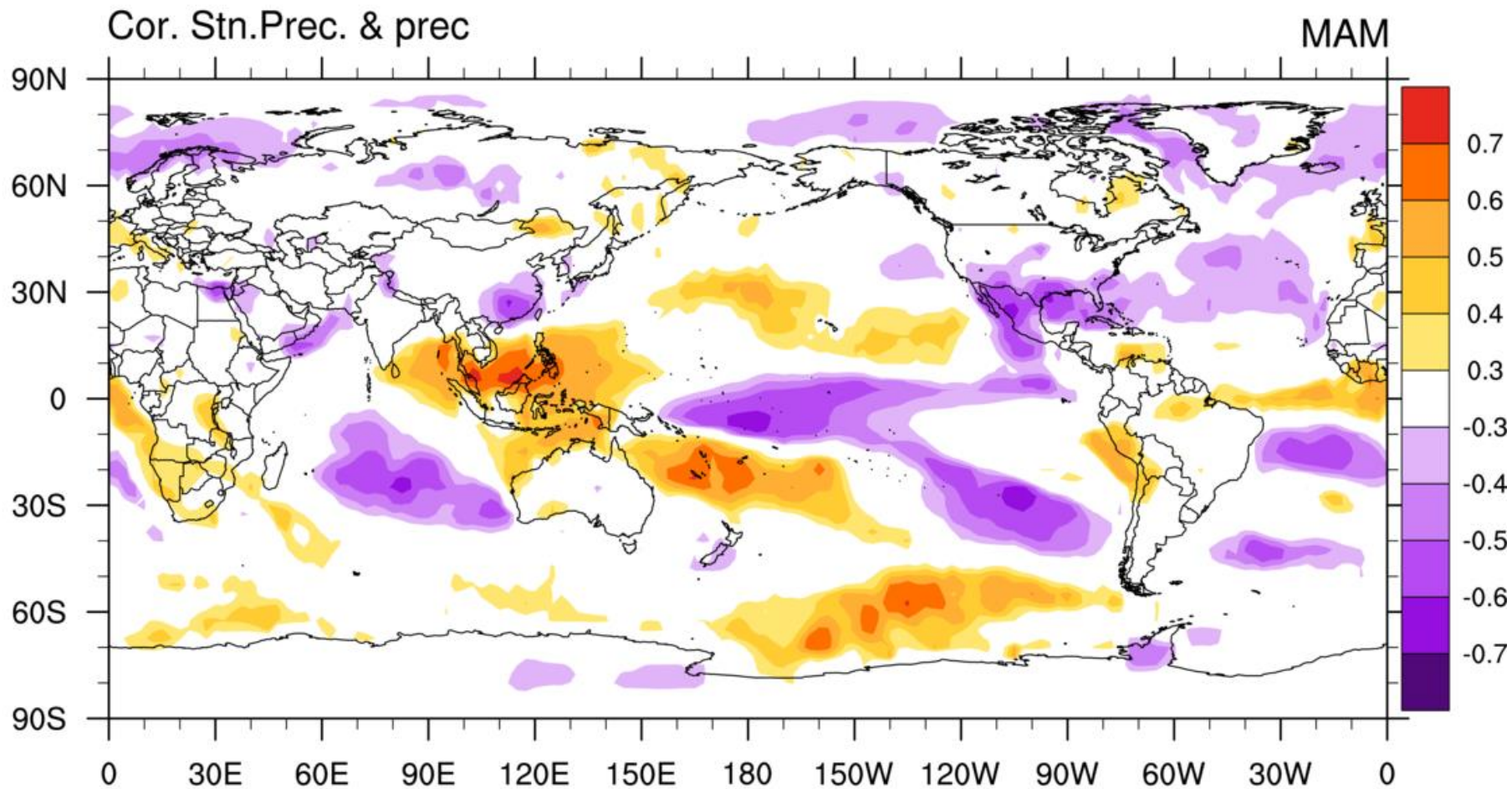
Hanoi (SON)



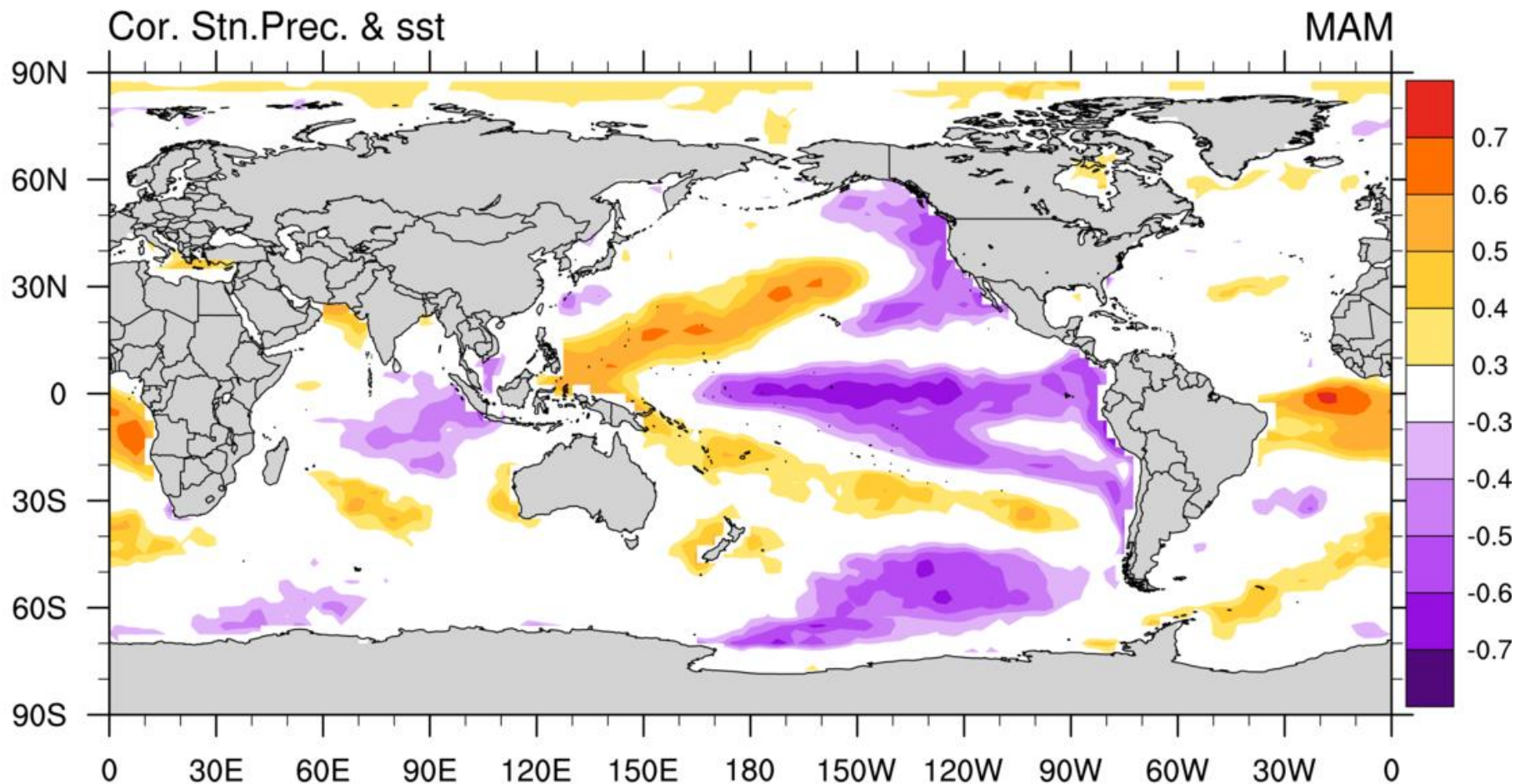
Hanoi (SON)



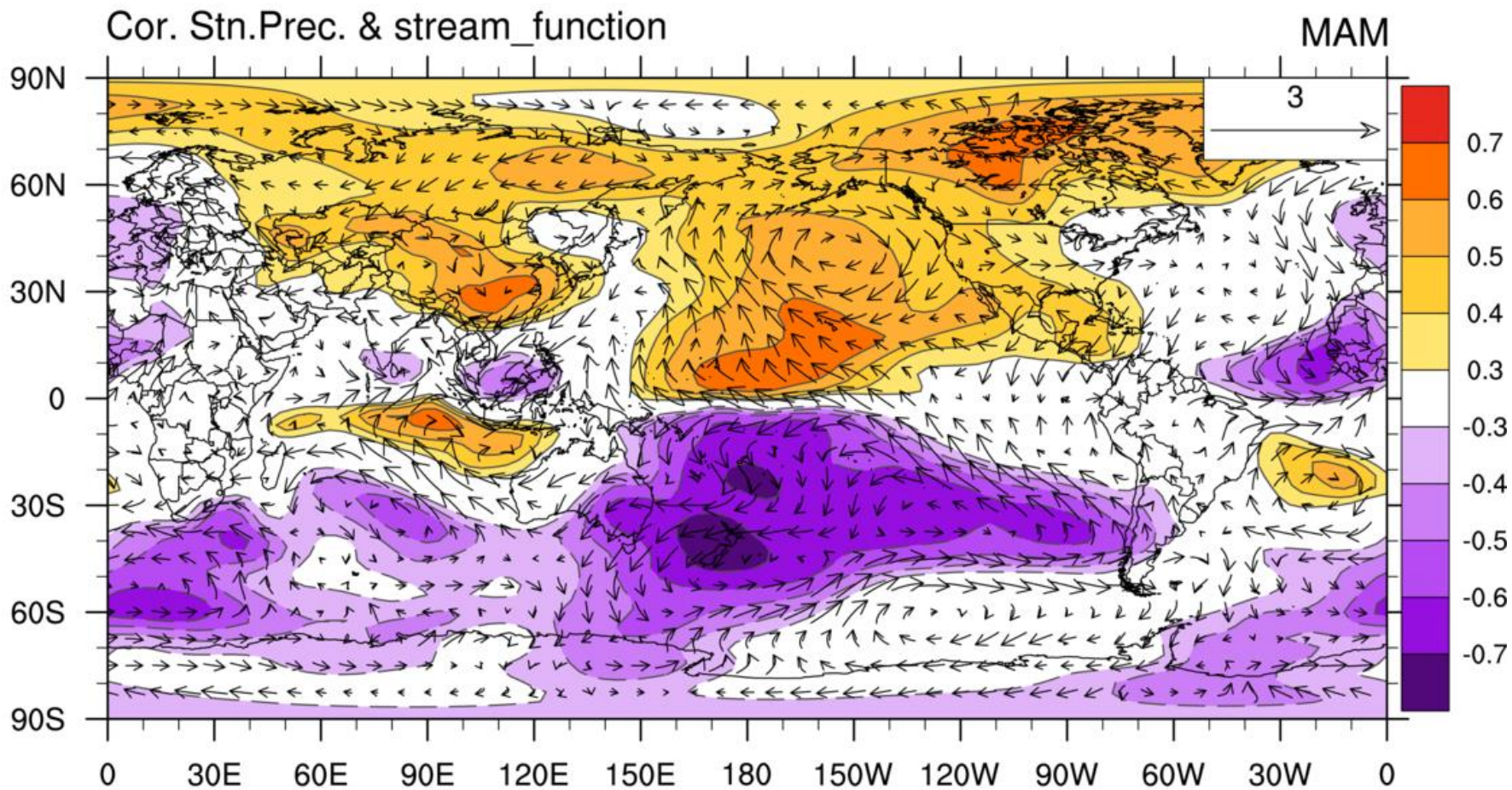
Bangkok (MAM)



Bangkok (MAM)



Bangkok (MAM)



Predictability

- More structured pattern : higher potential
- But actual predictability depends on model's ability to predict the pattern
 - How do we predict?

Seasonal Prediction (2)

Jin Ho Yoo
APEC Climate Center

Methods

- Statistical (Empirical)
 - Use observed relationship of climate system to predict future
 - Linear
- Dynamical
 - Based on “physical law” of climate system and expect to mimic “the memory”
 - Nonlinear

Which one is better?

Statistical

- Simple and cheap
- Based on data
- Data is real thing but do we have enough?

Dynamical

- Complex and expensive
- Based on Law
- Is our understanding accurate?

Statistical forecasting

- (0) Climatology

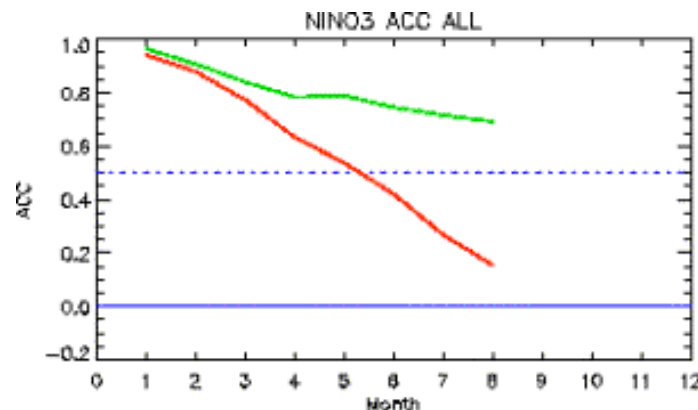
$$x(t + 1) = \bar{x}$$

- Baseline of seasonal forecasting
- “Nothing particular, Sir.”
- Deterministic forecast
 - Rainfall amount will be similar to 30year average

- Probabilistic forecast
 - Near normal ?
 - I don't know? (33%:33%:33%)

Statistical forecasting

- (1) Persistence $x'(t + 1) = x'(t)$
 - Assume that future will be same as it is now
 - ANOMALY !
 - Often Close to people's expectation
 - Effective when the autocorrelation is large
 - Often used for ENSO forecast (Nino3.4)

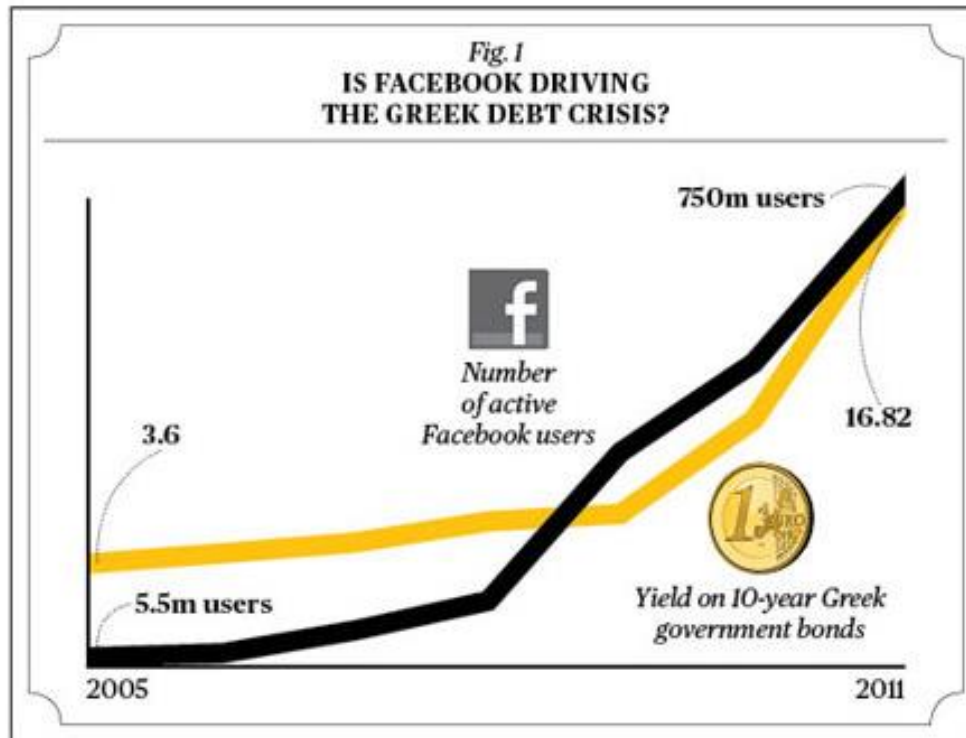


Statistical forecasting

- (2) Regression $x'(t + 1) = ay(t) + b$
 - The most popular method and many variations
 - x : predictand (e.g. rainfall at a station)
 - y : predictor (e.g. NINO3.4 SST)

Predict yield of Greek bonds with Facebook users

- Is it appropriate?



If yes, why?

If not, why?

From *business week*

Regression based forecast

- Question #1 $x'(t + 1) = ay(t) + b$
 - How to define predictor (y)?
 - By definition, predictor should cause some changes in variation of predictand
 - Predictand : my mood in the morning
 - Predictor?

Regression based forecast

- Question #2

$$x'(t + 1) = ay(t) + b$$

- How to define **a** and **b**?
- your choice. Linear, nonlinear, single, multi....
 - Complex one is not necessarily better.
- Predictand : my mood in the morning
- Predictor :
- a , b?

Regression based forecast

- Question #1 : Predictor selection
 - Should be based on Physical relationship between predictors and predictands
 - Predictor cannot be tiny signal in the seasonal forecast
 - Keep “doubt” on the possibility of selection by chance
 - Selected predictor should be validated with separate data

Regression based forecast

- Question #2 : appropriate Function

$$x'(t + 1) = ay(t) + b$$

$$x'(t + 1) = a_1y_1(t) + a_2y_2(t)b$$

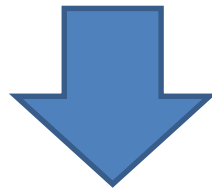
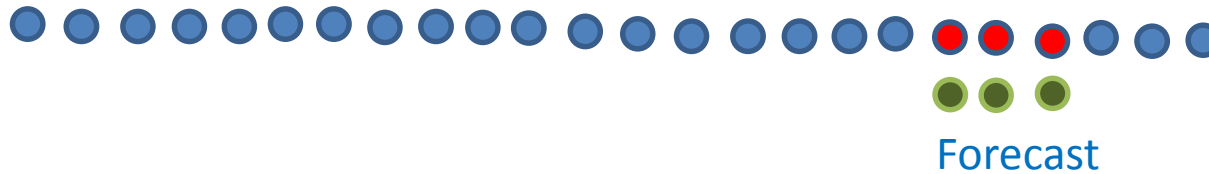
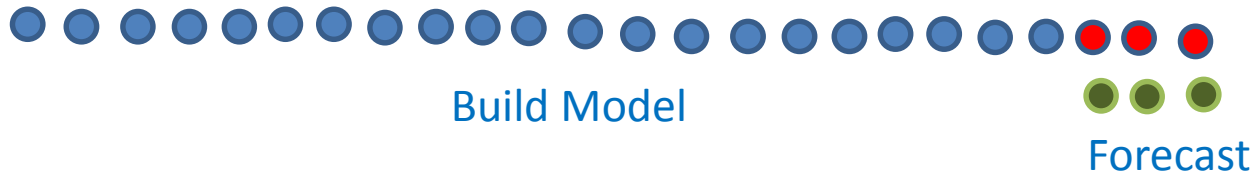
$$x_1'(t + 1) + x_2'(t + 1) = a_1y_1(t) + a_2y_2(t)b$$

- One to One : often not very satisfactory, limited cases
- One to Multi : easy to overfit (lie)
- Multi to Multi : looks nice but often produce nothing practical
- If they gives similar result, the simpler is the better

Cross-validation

- Climate data record is short (~3-40years)
 - EX) We have only a few El Nino events.
 - It's better to consider as many event as possible
 - Then, how to verify? (How to estimate skill of developed method?) for new cases?
 - Devide data into **Training** and **Validation** period

Cross-validation



- Exact value of model parameters can change

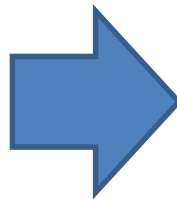
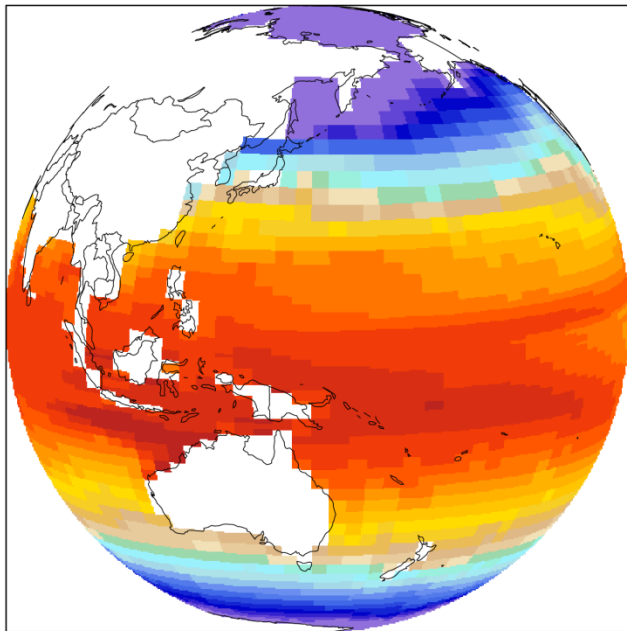
- NEVER use “omitted” data in ANY CASE of model development

The most important thing

- Physical understanding of;
 1. What weather event/system consists of your seasonal climate (LOCAL, predictand)
 2. What external (slow varying factor) controls the weather system (GLOBAL, predictor)

Dynamical forecast

- Use GCM : Global Climate Model
 - It used to be called “General Circulation Model”



Dynamical forecast

- Governing Equations
 - Written as computer program code (NWP)

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \nabla \Phi - 2\Omega \times \mathbf{u} - \frac{1}{\rho} \nabla p + \mathcal{F}$$

$$\frac{\partial \rho}{\partial t} + \bar{\nabla}(\rho \bar{\mathbf{u}}) = 0 \quad \Leftrightarrow \quad \frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{u}$$

$$\frac{\partial \theta}{\partial t} + \bar{\mathbf{u}} \cdot \bar{\nabla} \theta = l$$



```
//Behradok functions:
//
//   DTstage(T+9.11)^-2.85
//
MinTime=pow(T[j]+9.11, -2.85);//Minimum time to advance to stage (in days)
for(k=0;k<numLifeStage;k++)
{
  MaxRate[k]=MinTime*DTstage[k];
  MaxRate[k]=MaxRate[k]*ToSecs;//Convert to seconds
  MaxRate[k]=1.0/MaxRate[k];//Convert to rate
}

//Parameters for Ivlev functions controlling food dependence
//
//   R=a[1-exp[-b*(food-c)]]--development rate (days^-1)
//
// But, idea is that temp sets max growth rate, and food tells us how close
// we get to the max. In this sense, a=1 (Campbell figured an absolute
// rate, we're essentially normalizing his rates by rate at 40C.
//
//b=[ones(1,6)*params.bnaup,ones(1,6)*params.bcop];

for(k=0;k<6;k++)
  Rfood[k]=[1.-exp[-(F[j]-c)*params.bnaup]];
for(k=6;k<12;k++)
  Rfood[k]=[1.-exp[-(F[j]-c)*params.bcop]];

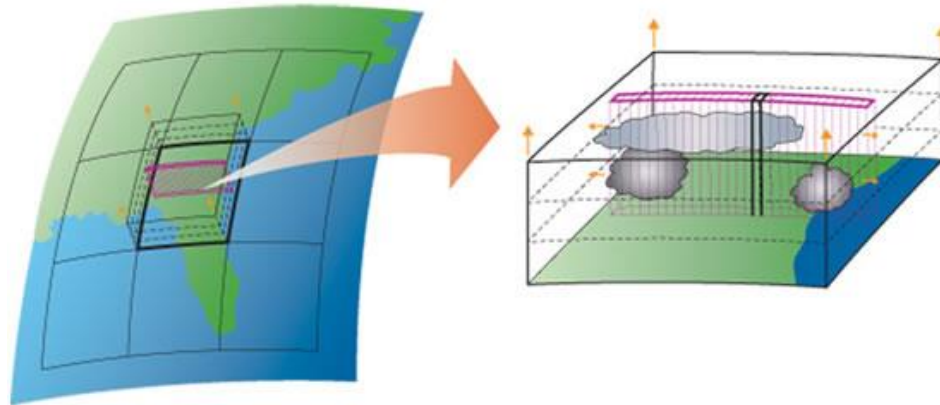
//Multiply Rfood by MaxRate to get the actual rate.
for(k=0;k<12;k++)
  R[k]=MaxRate[k]*Rfood[k];

R[12]=0.;//adults don't molt

//M[k]=mortality rate for stage k at node j
//
gammaT=gamma0*(1.-gamma0)*pow(T[j]/Tc,z);
//gammaT=0.1; //Override temp dependent mortality
```

Numerical modeling

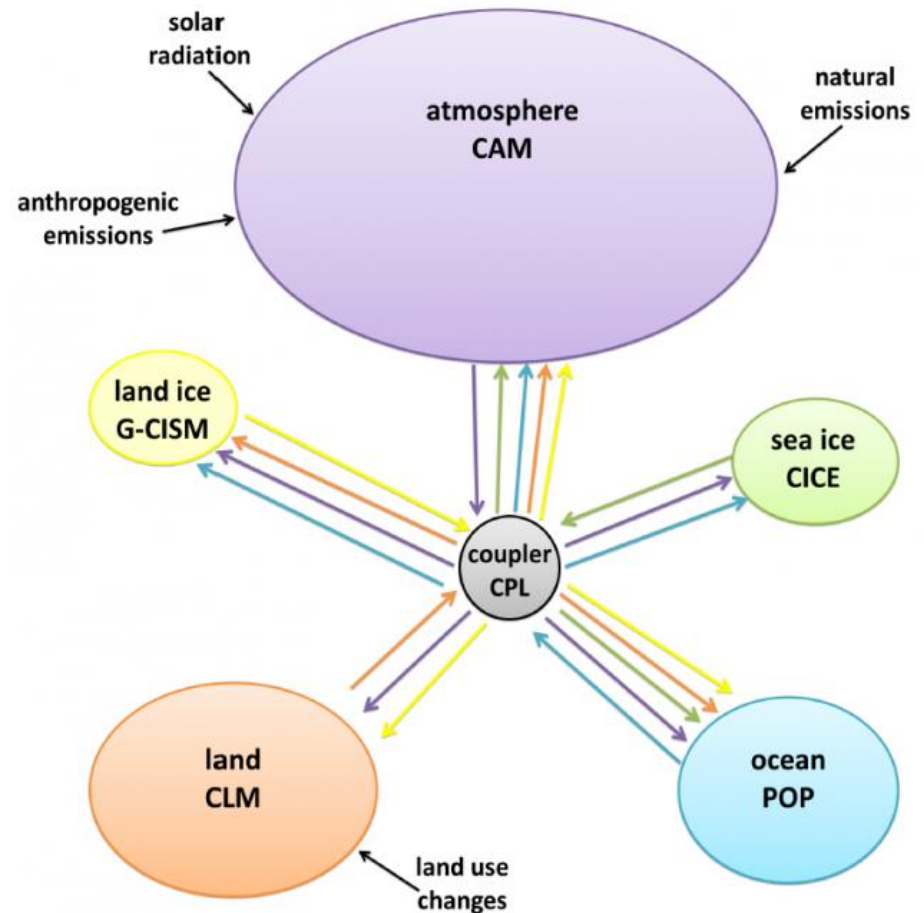
- Issue
 - Digitization (physical variable is continuous, but computer needs digitization”
 - Resolution, subgrid-scale parameterization



- Unknown processes, tunable parameters
- Initialization (for forecasting)

GCMs

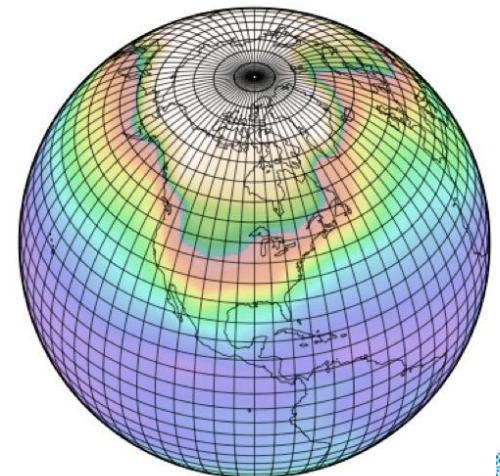
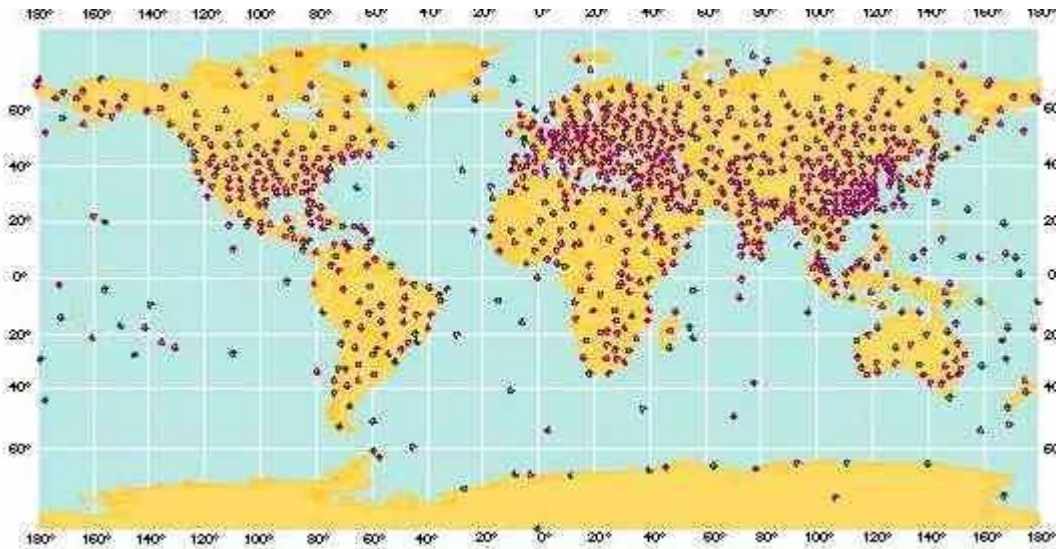
- Coupled GCM
 - Atmosphere
 - Ocean
 - Sea-Ice
 - Land surface
 - Chemistry
 - Biosphere



Initialization

Estimating Current status of climate system

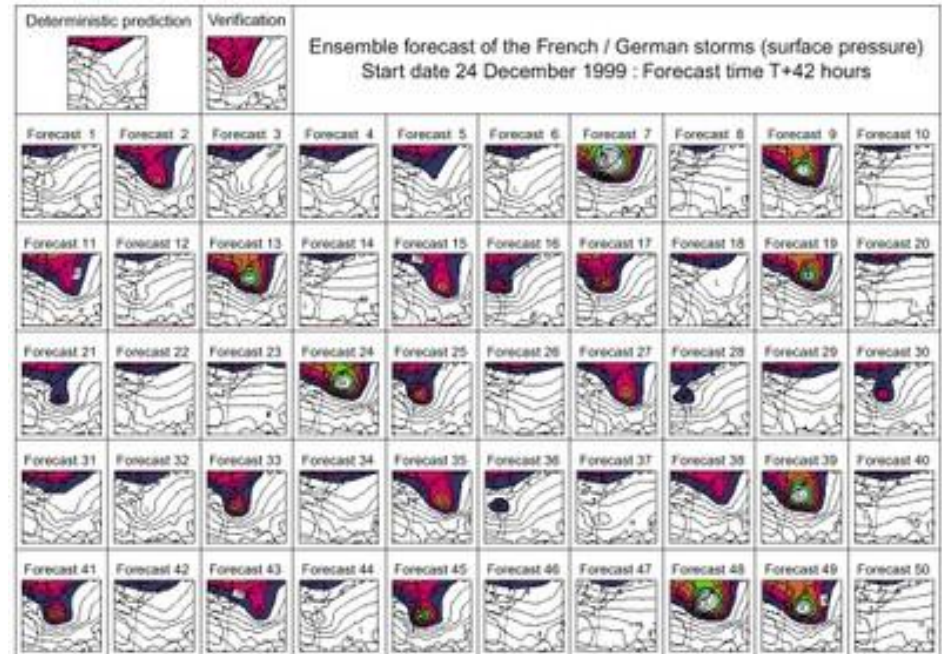
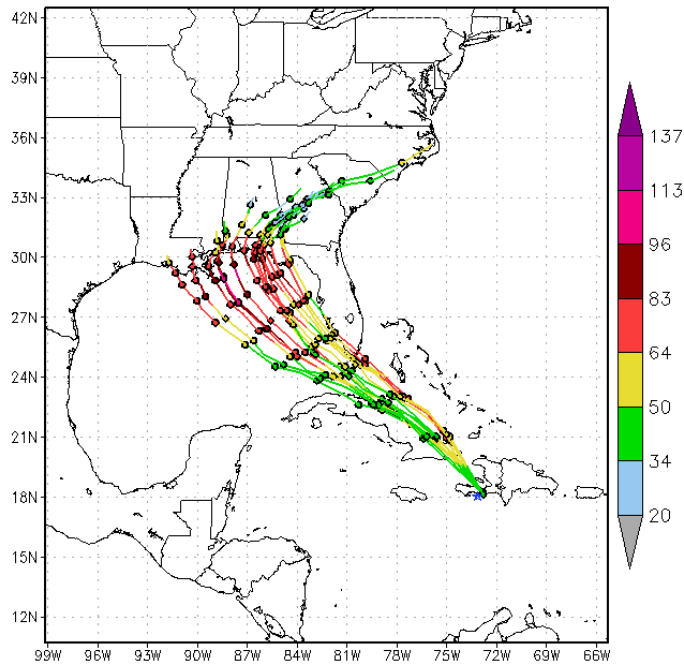
- Preparing the beginning climate state of GCM with available observation
 - Balance between Wrong GCM vs Wrong OBS.
 - Balance between components (Atm, Ocn)



Ensemble Forecasting

- Run many times
 - Starts from slightly different initial conditions

6-hourly Track and Intensity (kt) for ISAAC09L
 GFDL ensemble forecast for the 126 hrs from 06Z25AUG2012



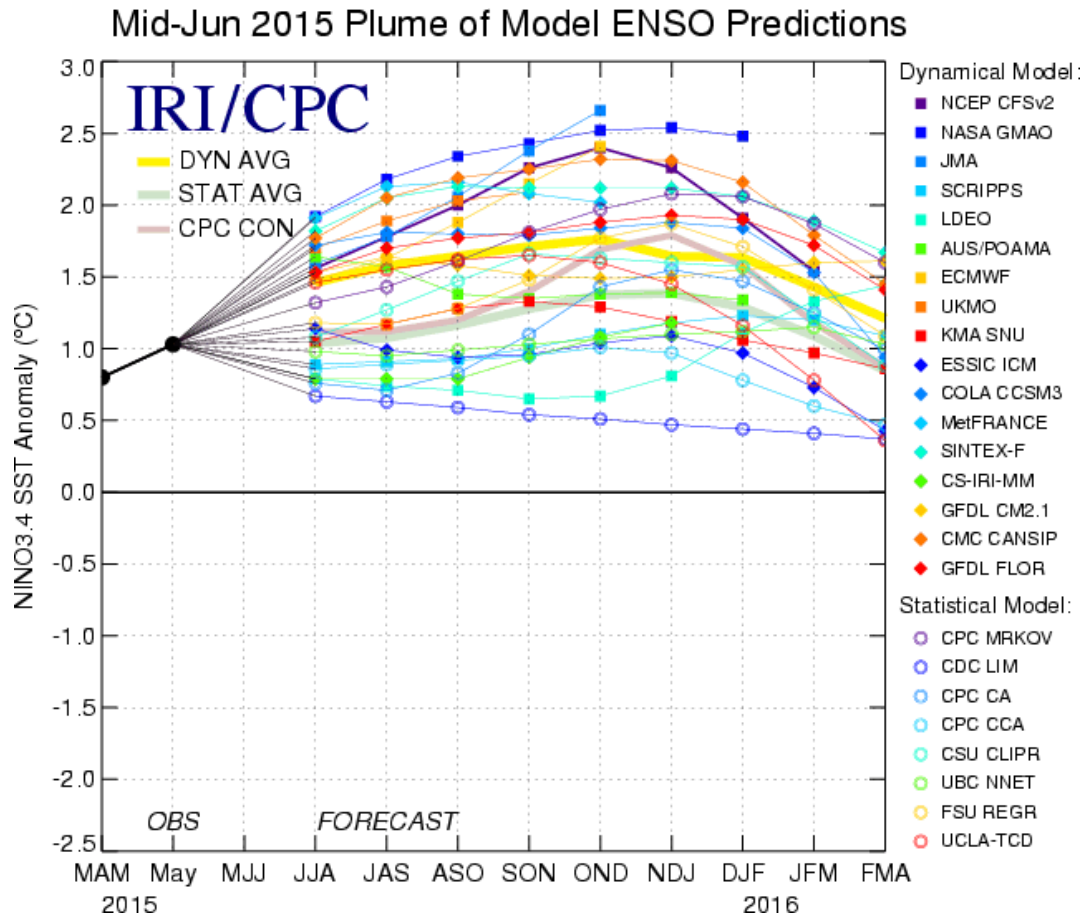
of missing members (out of 16) at t=0: 0
 indicates ISAAC09L observed center at initial time

Track forecast positions are marked every 12 hrs

GFDL Hurricane Dynamics Group

Multi Model Ensemble Forecasting

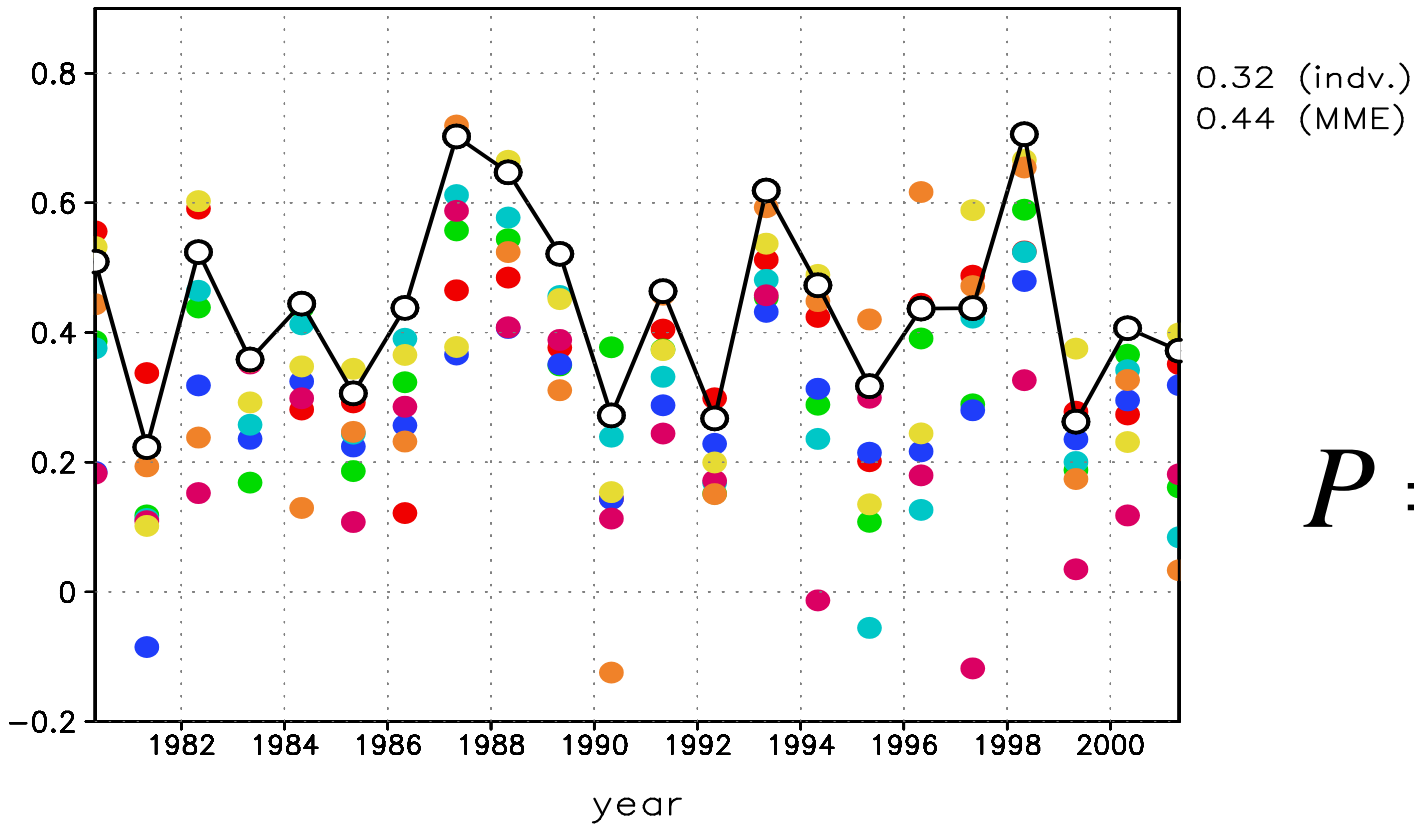
- Run with many models



Which one??

Use all!

Pattern correlation : summer monsoon precip.



$$P = \sum_i a_i F_i$$

Predictability of Multi Model Ensemble

Correlation skill of a single model

$$R_i = \frac{\overline{xy_i}}{\sqrt{V(x)V(y_i)}}$$

Correlation skill of MME

$$\langle y \rangle = 1/M \sum_{i=1}^M y_i$$

$$R_{MM} = \frac{x \langle y \rangle}{\sqrt{V(x)V(\langle y \rangle)}} = \frac{1}{M} \sum_{i=1}^M \left(R_i \sqrt{\frac{V(y_i)}{V(\langle y \rangle)}} \right) = \langle R \rangle \sqrt{\frac{\langle V(y) \rangle}{V(\langle y \rangle)}}$$

$$\langle R \rangle = \frac{1}{M} \sum_i R_i$$

$$V(\langle y \rangle) = \langle V_{Single} \rangle - \frac{M-1}{M} \langle V(y_n) \rangle - \frac{M-1}{M} \langle (V(e) - C(e)) \rangle$$

$$R_{MM} = \frac{\langle R \rangle}{\sqrt{V(\langle y \rangle)}} = \frac{\langle R \rangle}{\sqrt{\langle r \rangle}}$$

$$\langle r \rangle = \frac{1}{M^2} \sum_i \sum_j \frac{\overline{y_i y_j}}{V}$$

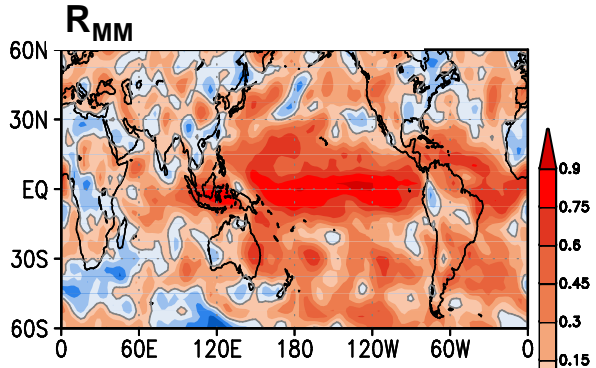
$$E_{MM} = \langle V_{Single} \rangle (1 + \langle r \rangle - 2 \langle R \rangle)$$

Observation : $x = x_s + x_n$
 Forecast : $y = y_s + y_n = x_s + e + y_n$

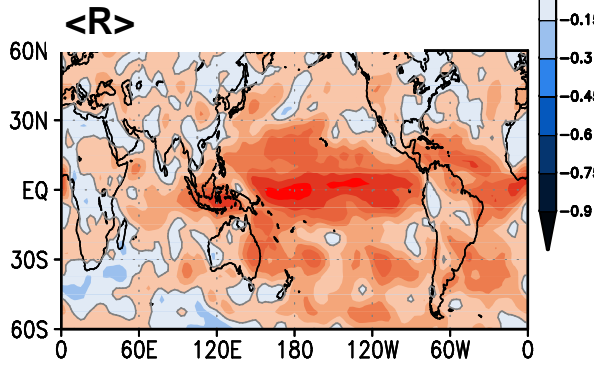


Temporal correlation skill (SUMMER MEAN PRCP)

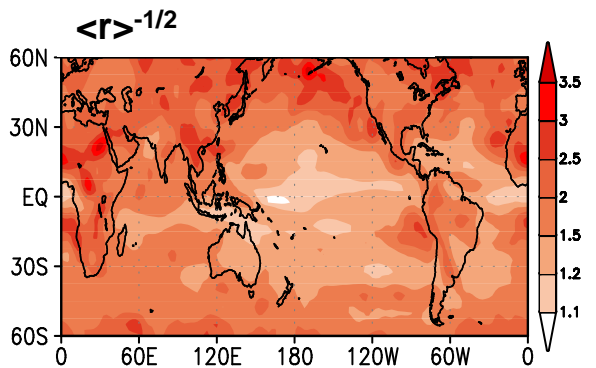
Multi-model ensemble correlation skill



Mean correlation skill of individual models

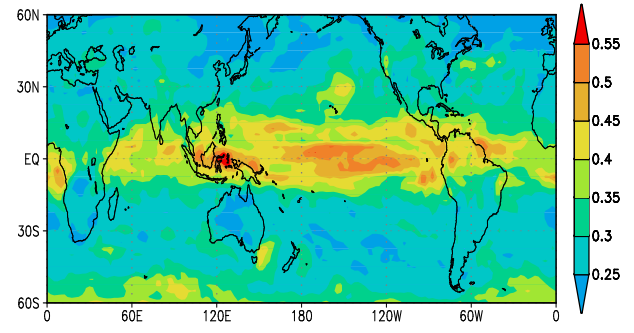


Inflation factor of correlation skill by multi-model ensemble



$$V(\langle y \rangle) = V_{Single} - \frac{M-1}{M} \langle V(y_n) \rangle - \frac{M-1}{M} \langle (V(e) - C(e)) \rangle$$

Contribution of systematic error (conditional) cancellation



$$R_{MM} = \frac{\langle R \rangle}{\sqrt{V(\langle y \rangle)}} = \frac{\langle R \rangle}{\sqrt{\langle r \rangle}}$$

Independent and good models : Best forecast result (on average)

Beauty of Democracy

- Independent and Rational individuals :
 - Best decision for society (in a long run)



Seasonal Prediction (3)

Jin Ho Yoo
APEC Climate Center

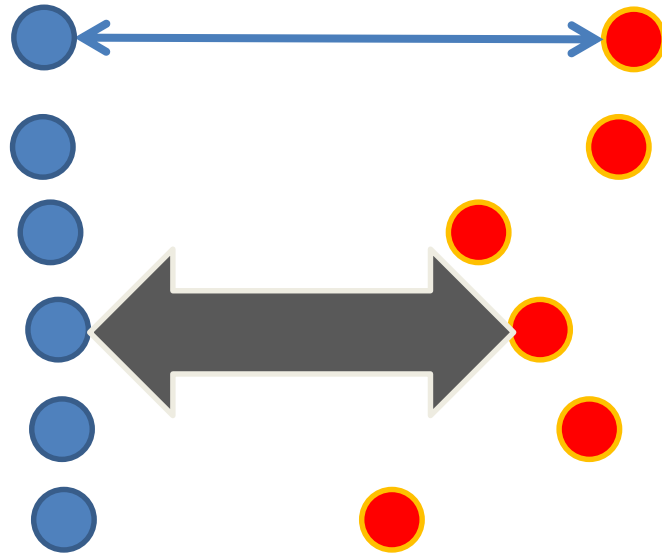
Which one?

- Evaluation of forecast : verification



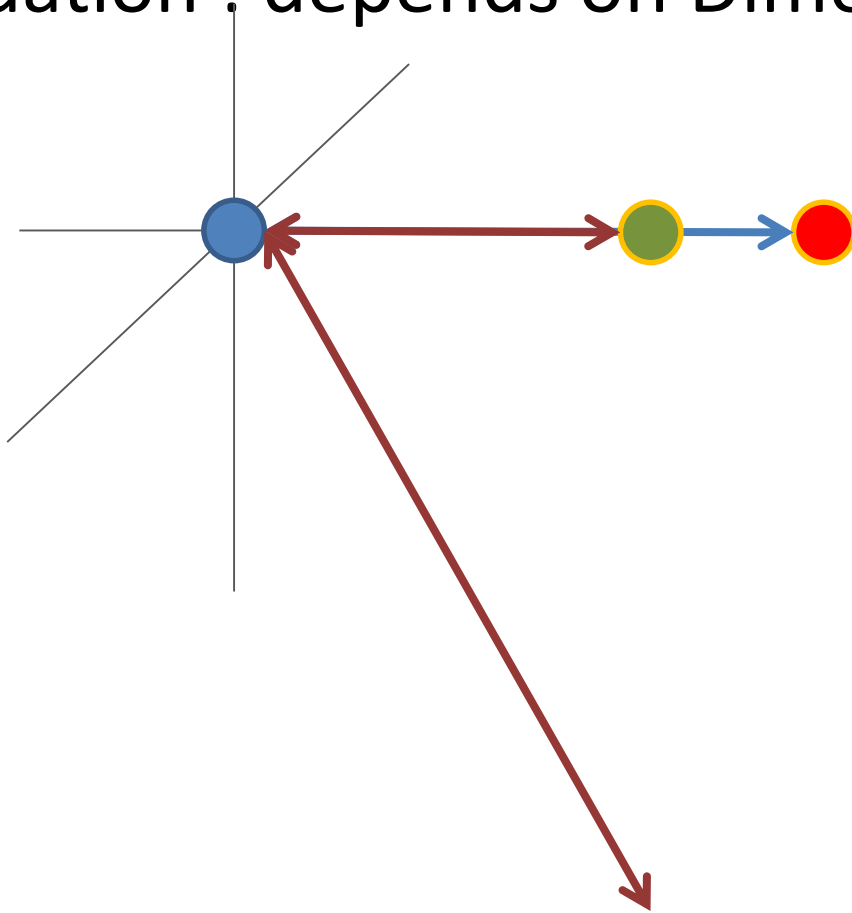
Verification

- Evaluation : measure of closeness



Verification

- Evaluation : depends on Dimension/Viewpoint



Deterministic forecast

- Various measures

- MSE (Mean Square Error), RMSE (Root MSE)

- $MSE = \frac{1}{N} \sum_i (F_i - O_i)^2$

- ACC (Anomaly correlation, Pattern), TCC (Temporal correlation)

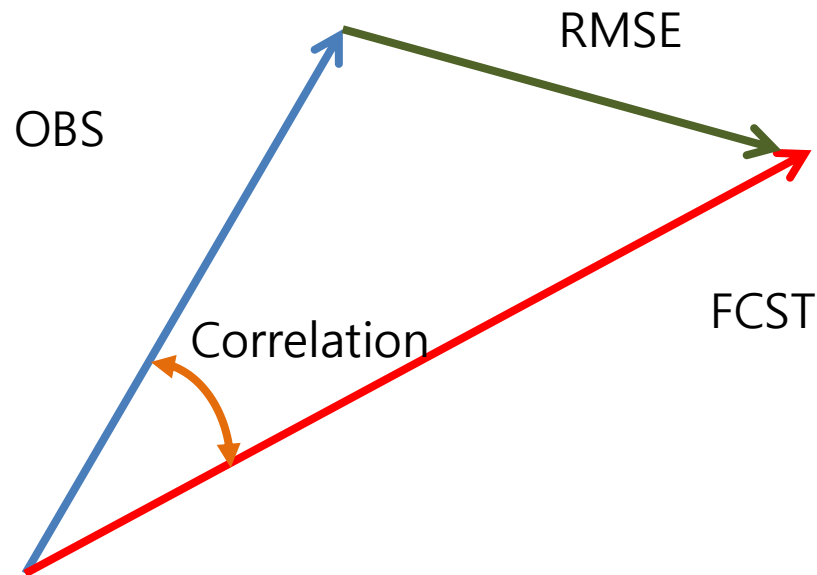
- MSSS (Mean Square Skill Score)

- Conventional form of skill score

- $1 - \frac{E}{E_c}$, E : error/penalty, E_c : error of reference forecast

Verification

- Evaluation : depends on Dimension/Viewpoint

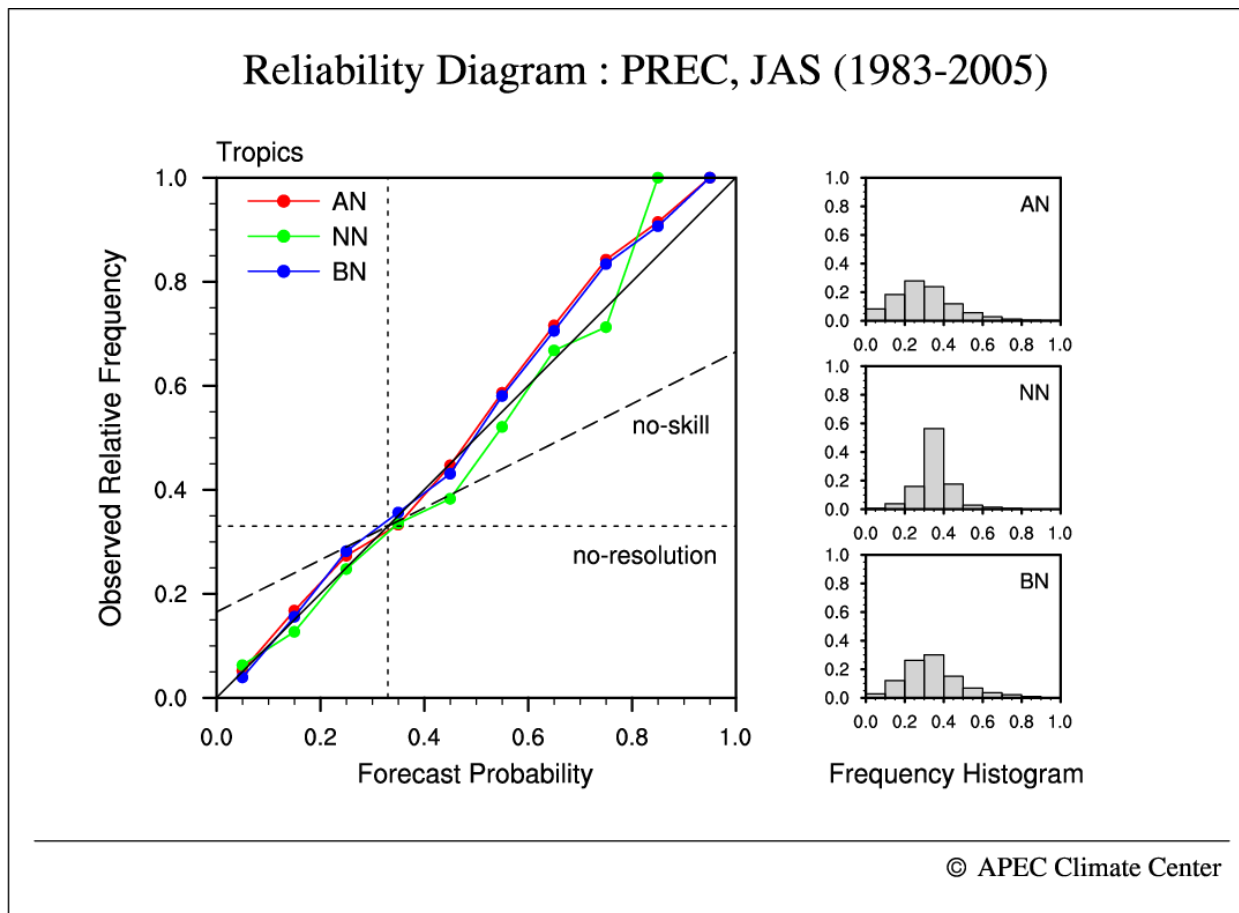


Probabilistic forecast

- Brier score (Brier Skill Score)
 - MSE of prob. forecast
 - $BS = \frac{1}{N} \sum_i (F_i - O_i)^2$, $F=1/0$, $O=1/0$
 - BSS (Brier skill score)
 - $1 - \frac{BS}{BS_c}$,
 - E : error/penalty, E_c : error of reference forecast

Probabilistic forecast (Categorical)

- Reliability curve



Probabilistic forecast (Categorical)

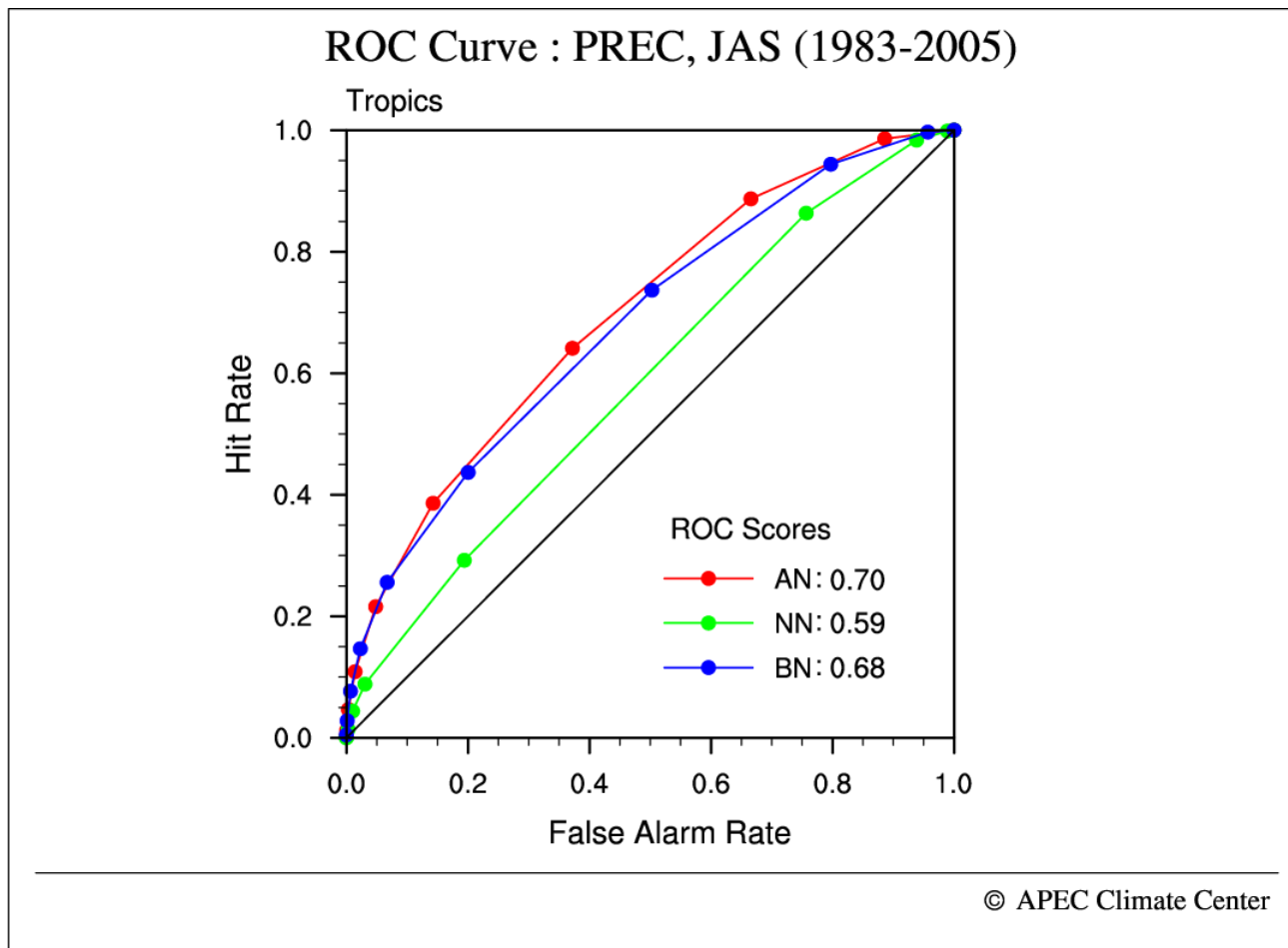
- ROC (Relative Operating Characteristics)

F	O	Yes	No
Yes		Hit (H)	False Alarm (F)
No		Miss (M)	Correct Rejection (C)

- HR (Hit rate) = $H/(H+M)$
- FAR (False Alarm rate) = $F/(F+C)$
 - Good forecast : $HR \uparrow, FAR \uparrow$

Probabilistic forecast (Categorical)

- ROC (Relative Operating Characteristics)



Probabilistic forecast (Categorical)

- HSS (Heidke Skill Score)

F \ O	Yes	No
Yes	Hit (H)	False Alarm (F)
No	Miss (M)	Correct Rejection (C)

- $$HSS = \frac{(score - score\ by\ chance)}{(perfect\ score - score\ by\ chance)}$$

$$\frac{\{(h+c)/n - [(h+f)(h+m) + (f+c)(m+c)]/n^2\}}{\{1 - [(h+f)(h+m) + (f+c)(m+c)]/n^2\}}$$

Forecast economic value

$$V = \frac{E_{cli} - E_{fore}}{E_{cli} - E_{per}}$$

V=1 : perfect forecast

V=0 : climatological forecast

E_{fore} : Expected expense of forecast

E_{per} : Expected expense of perfect forecast

E_{cli} : Expected expense of climatological forecast

- When the forecast is **perfect**, $f = m = 0$. and $h = \bar{o}$. Then, $E_{per} = hC = \bar{o}C$
- When the forecast is **climatology**. The only one kind of action will be kept.

If Yes : $E = (h+f)C = C$, otherwise $E = mL = \bar{o}L$. If decision maker is rational, he/she will choose action of low expense. Thus, $E_{cli} = \min(C, \bar{o}L)$

		Observation (real event)	
		Yes	No
Forecast (action)	Yes	Hit (h) Cost	False alarm (f) Cost
	No	Miss (m) Loss	Correct rejection (c) 0

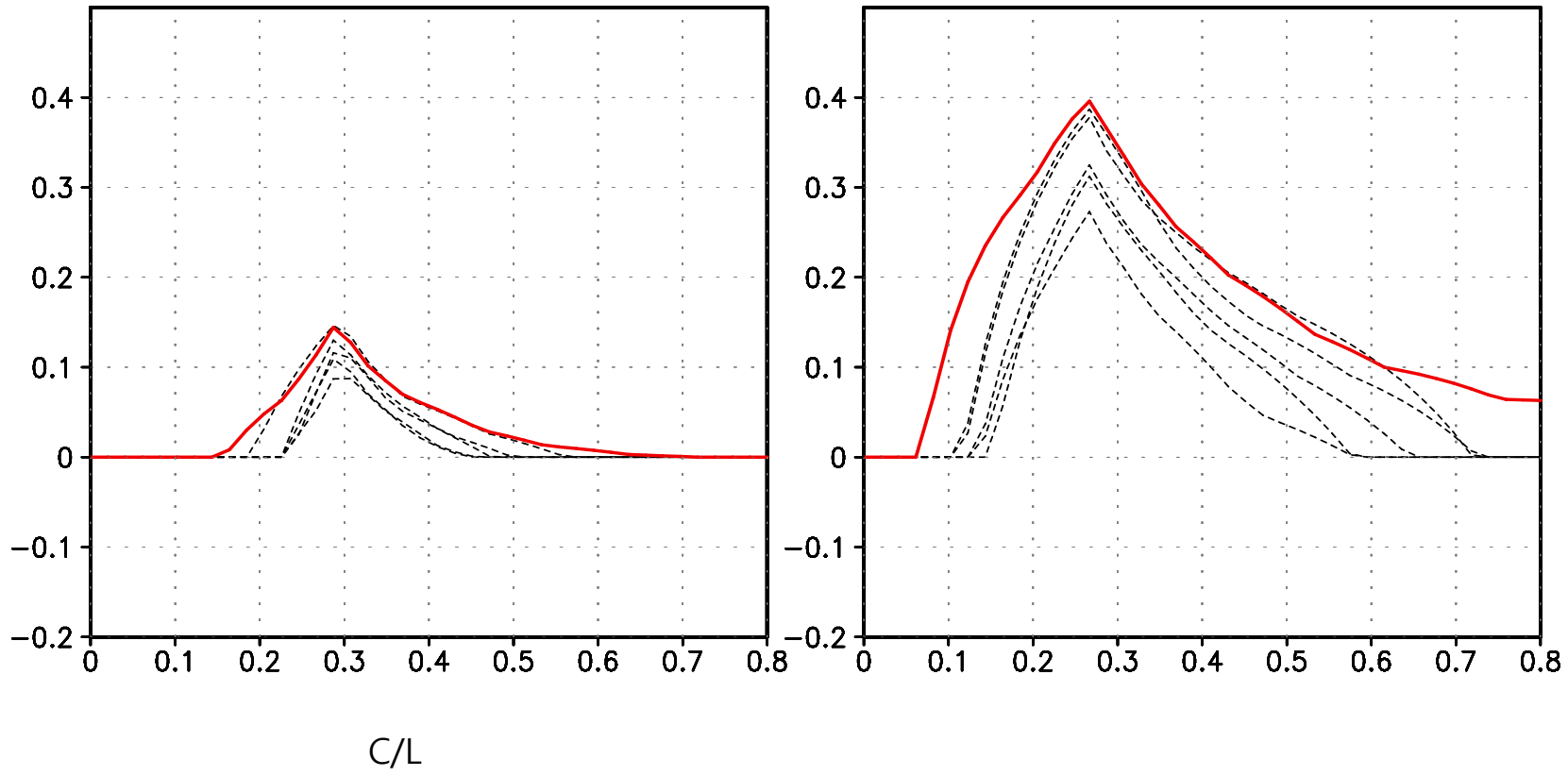
$$E_{fore} = (h + f)C + mL$$

$$V = \frac{\min(\frac{C}{L}, \bar{o}) - (h + f)\frac{C}{L} - m}{\min(\frac{C}{L}, \bar{o}) - \frac{C}{L}\bar{o}}$$

Value of Probabilistic forecast (Above normal) : GCMs

(a) Monsoon(40E-160E,20S~40N)

(b) ENSO (160E-280E,20S~20N)



----- Single model
—— MME

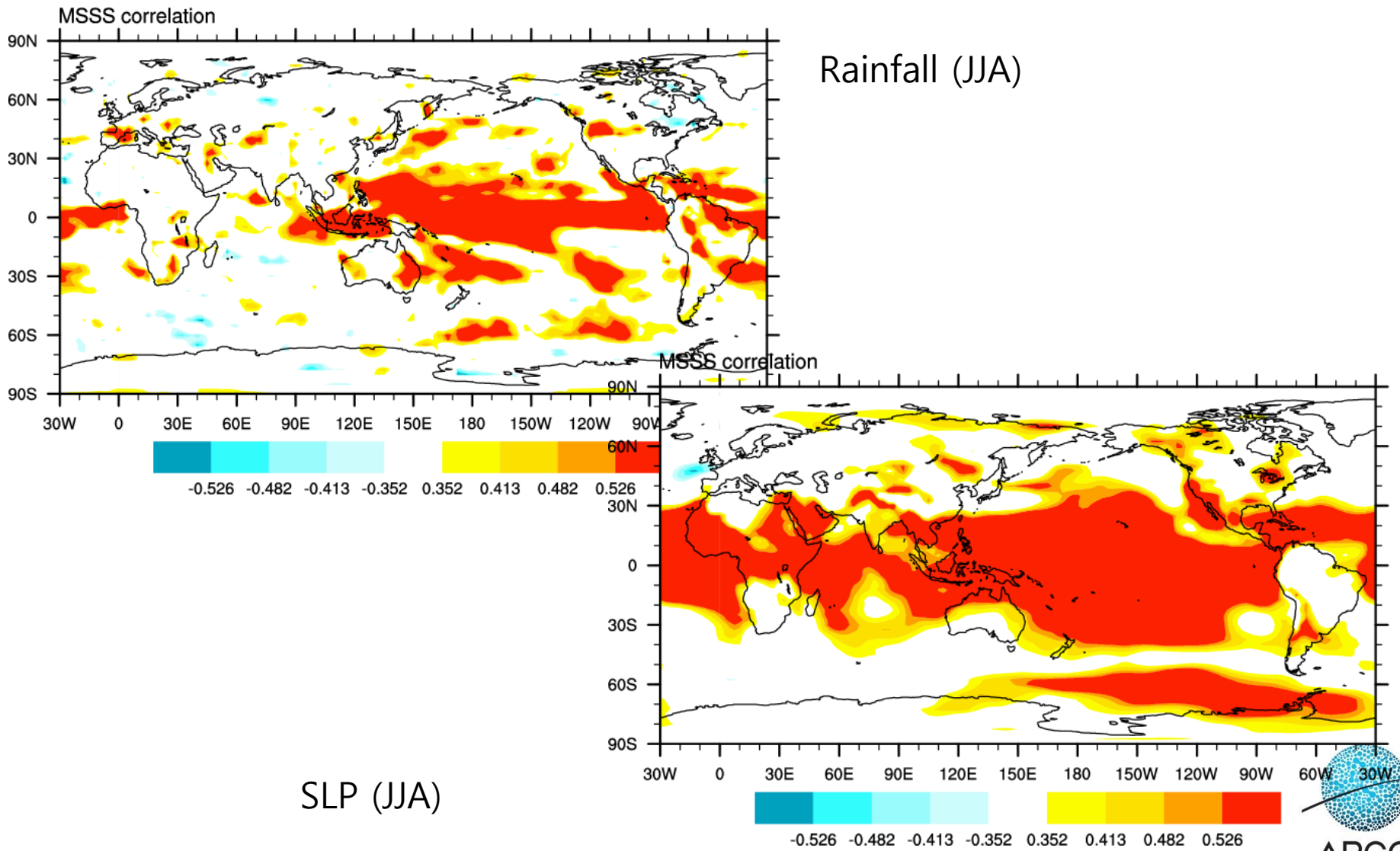


Forecast verification

- There are numerous ways
- Can be chosen by “what” do you want to see
- If not clear, use popular one.
- Difficulties in “translating” meteorological skill score into Public wording.

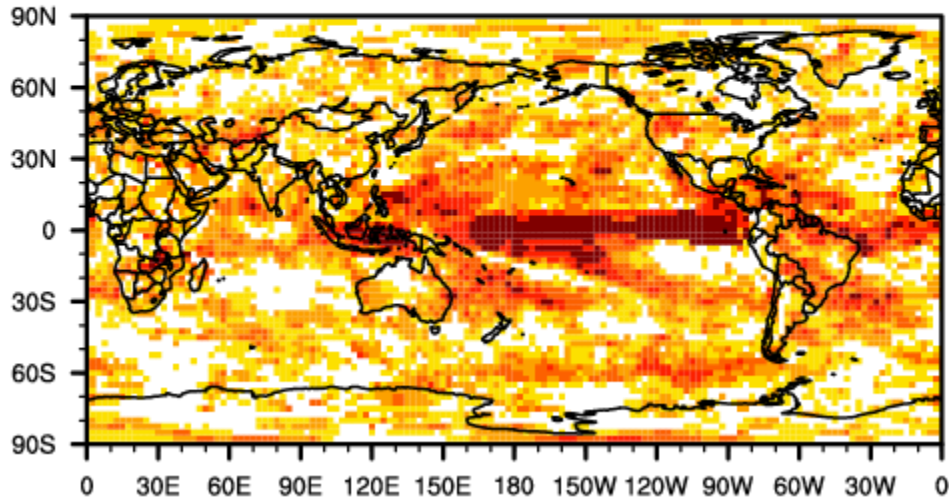
- Let’s see some results!

APCC (TCC)

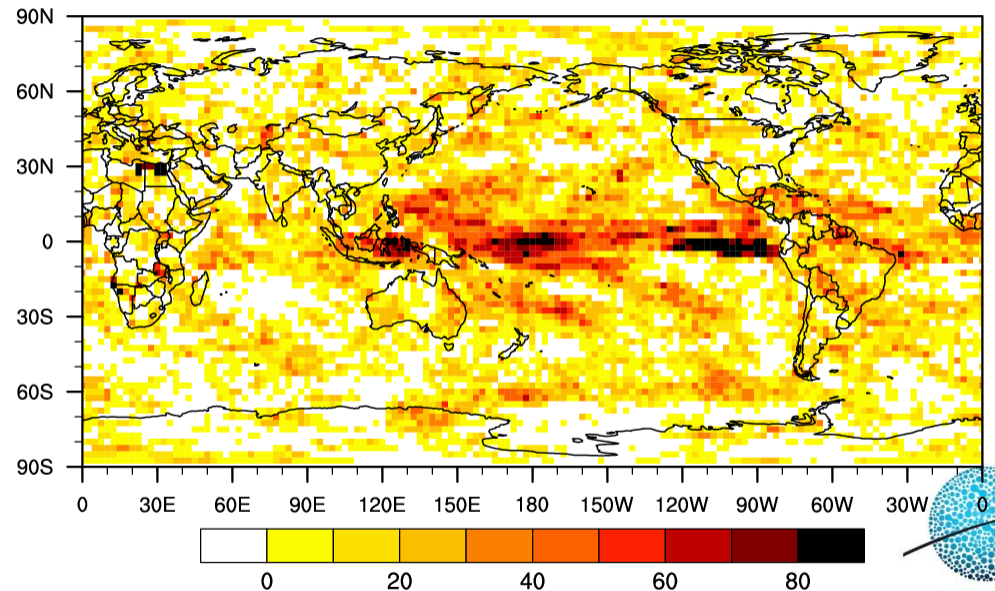


ROC Score : PREC, JJA (1983-2005)

Above-Normal



Heidke Skill Score : PREC, JJA (1983-2005)



R = 0.5, how good it is?

- Explaining 25% variance (R^2)
- A single verification score cannot tell everything.
- Multi aspect evaluation is necessary
- User oriented verification would be useful
- “this man can run fast, how good he is?”

Seasonal Prediction (4)

Jin Ho Yoo
APEC Climate Center

What we do?

- Collecting data and information
- Combine them
- Make a draft (preliminary decision)
- Consultation (discussion)
- Issue!

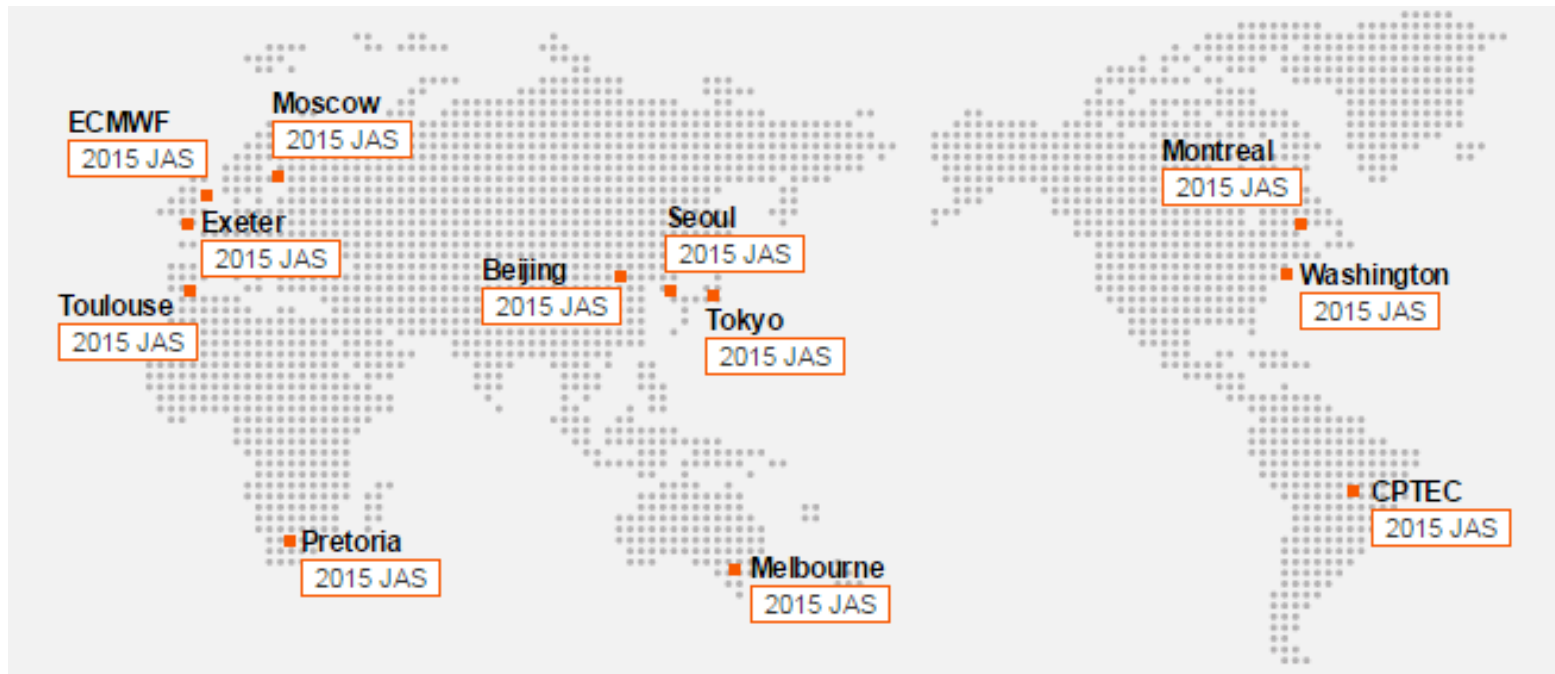
Current observation (monitoring)

- ENSO
 - WMO El Nino Update
- IOD
- ISO
 - CPC MJO page, APCC BSISO page

Why we monitor (analyze) current climate state?

Global forecast Information

- Dynamical Seasonal Prediction
 - GPCs, [WMO LC_LRF](#), [APCC](#), IRI, NMME



www.wmolc.org : only open to WMO members

Monitoring & Forecast information

- More maps are not always helpful unless they are **DIGESTED** properly
- It is known that **Multi Model Ensemble** tends to produce better forecast than a single model but it can lose regional details (maybe because of this, general skill is high)
- At best, all the information is merely explain large scale feature

Combining information

- If you can trust one thing, that is enough
- If you have different information with similar reliability, trust both
 - Are they **Independent**?
- If you can distinguish good and better information (but they are different), combine them with weight
- If you don't have any idea on the reliability, treat them similarly (they are all 'state-of-art' information)

Cautions

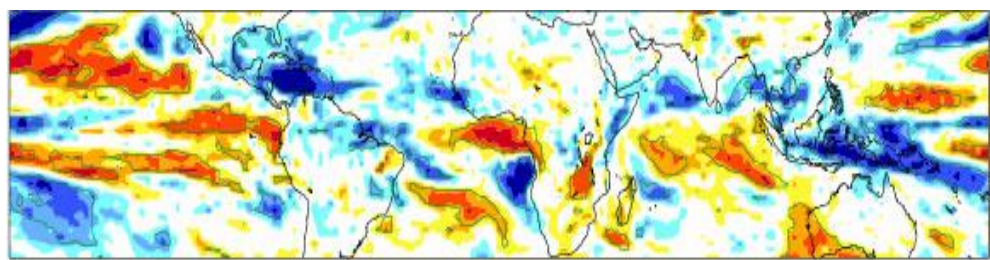
- How reliable our evaluation is?
- Even if you trust them, they can be wrong.
 - One reason to issue “probabilistic forecast”

2006 JJA mean Rainfall forecast

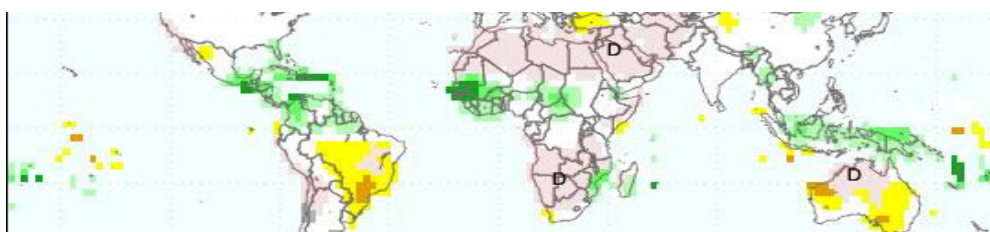
Warm colors : **dry**

Cool colors : **wet**

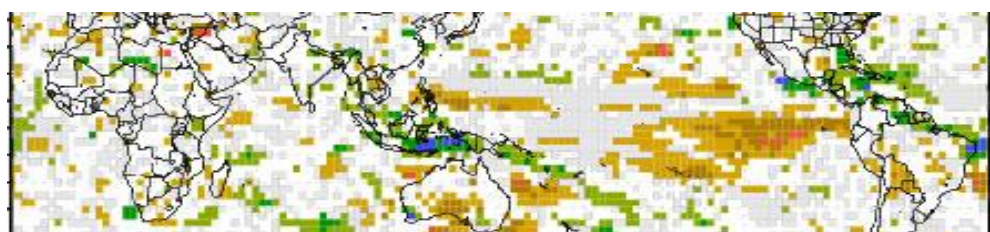
ECMWF



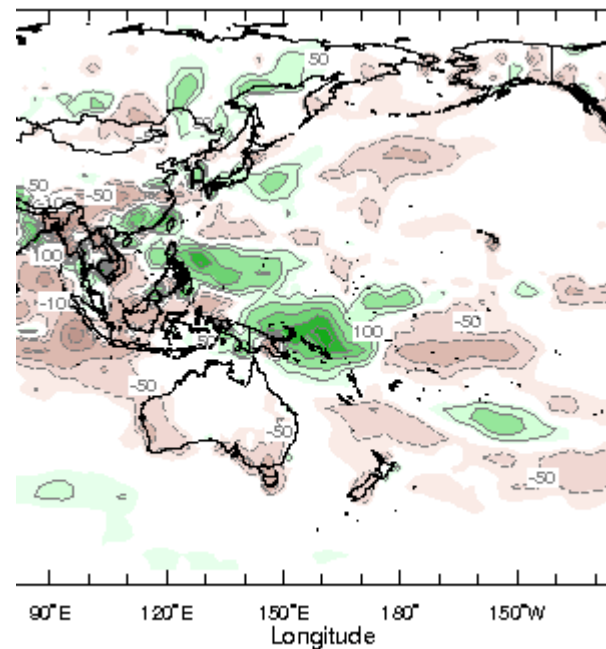
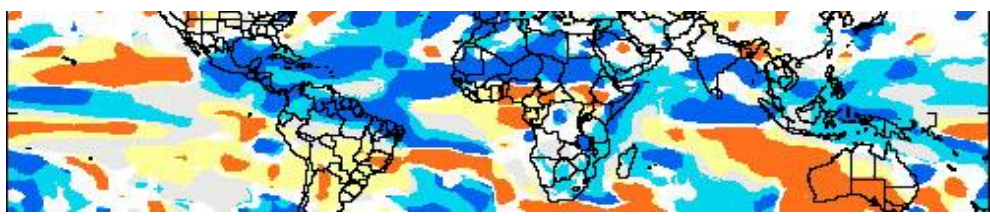
IRI



JMA



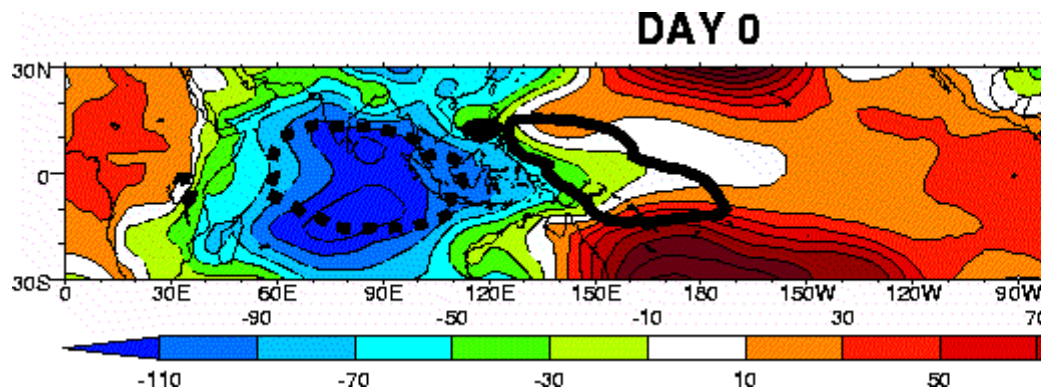
UKMO



A few more...

- Subseasonal information (MJO...)
- A new type of El Nino (El Nino Modoki)
- Way forward

Madden-Julian Oscillation

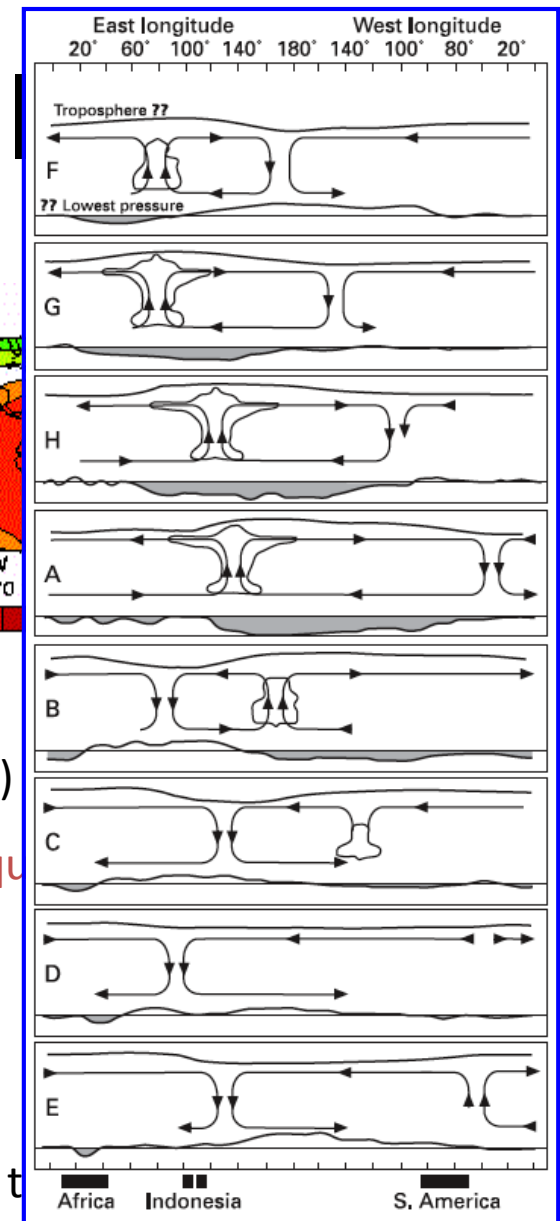


Madden & Julian (1971) : 40-50day oscillation (30-60days ISO)

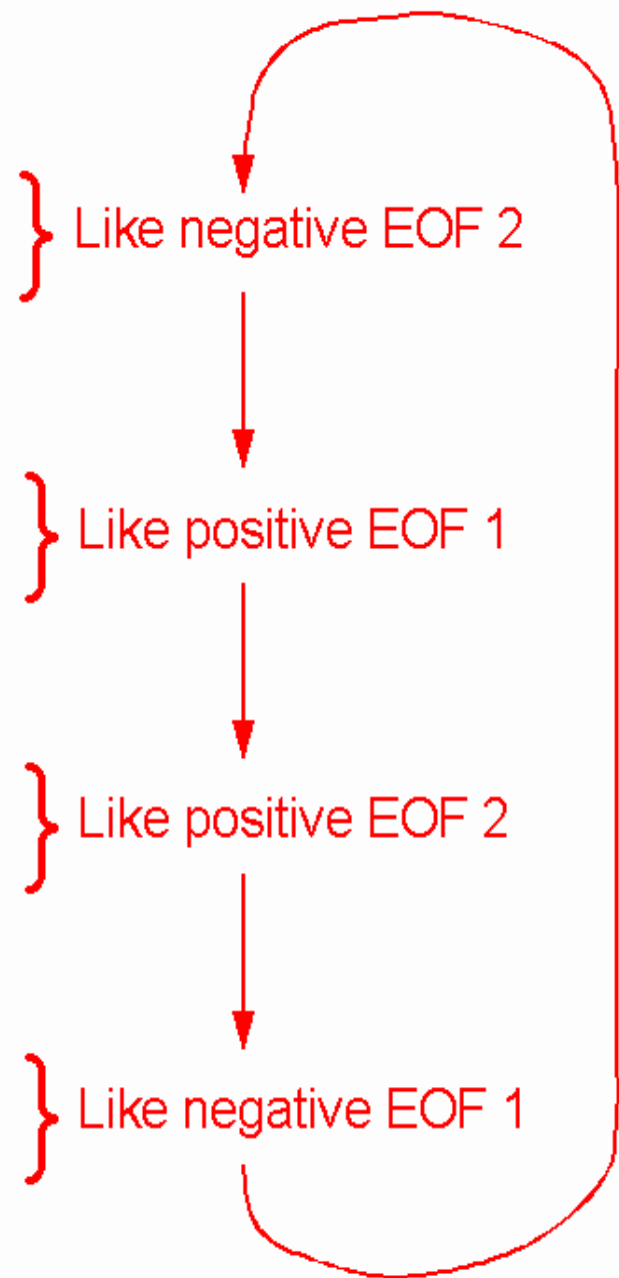
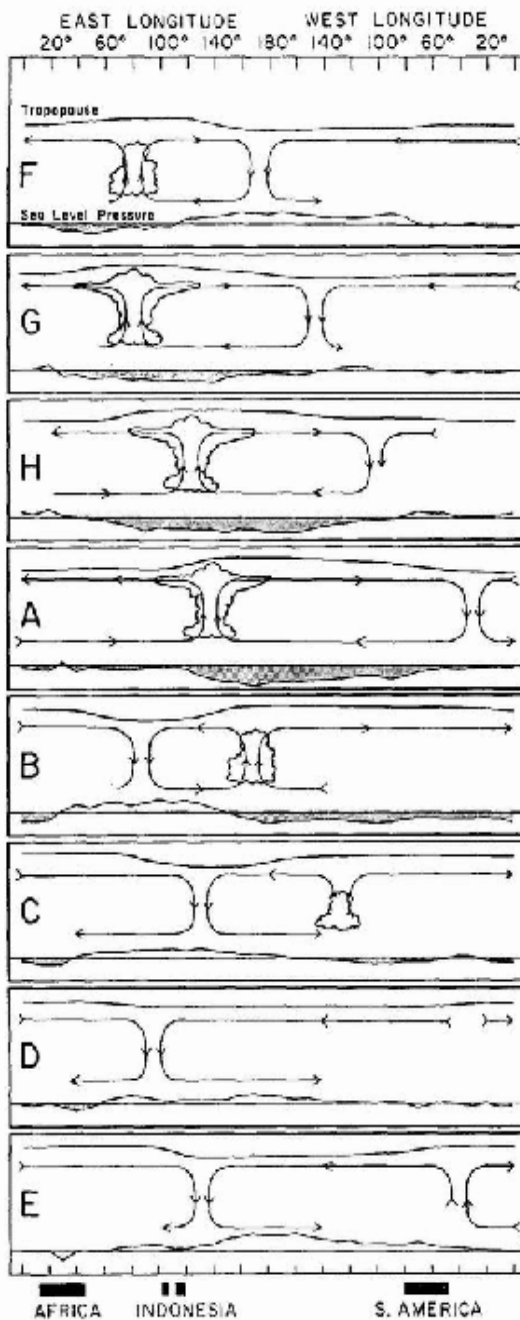
Eastward moving large scale convective anomaly along the equator
baroclinic structure

(precipitation anomaly is predominant in Indo-Pacific sector)

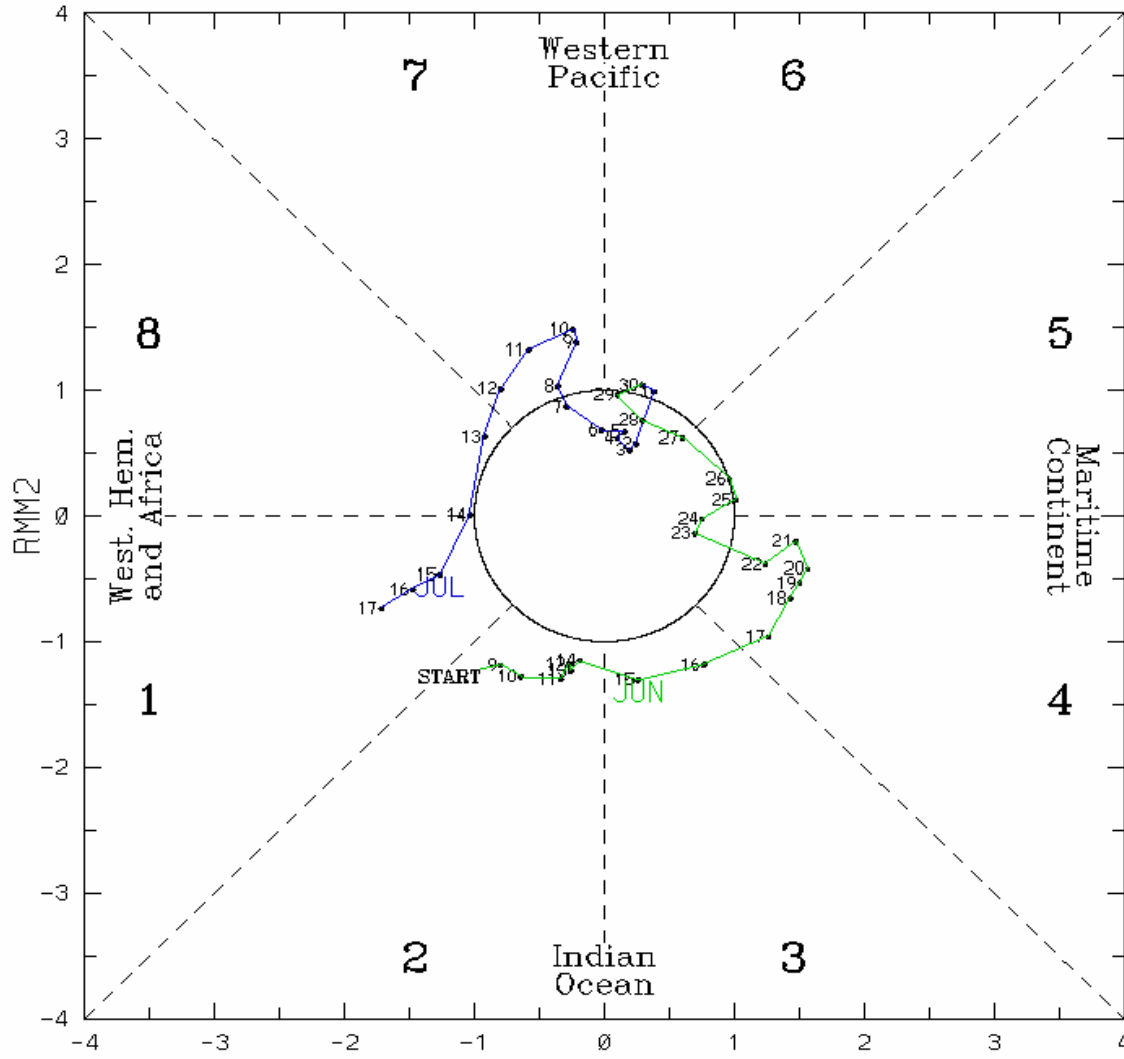
It can be a predictability source of extended range forecast in t



Madden and Julian (1972)



(RMM1,RMM2) phase space for 8-Jun-2007 to 17-Jul-2007

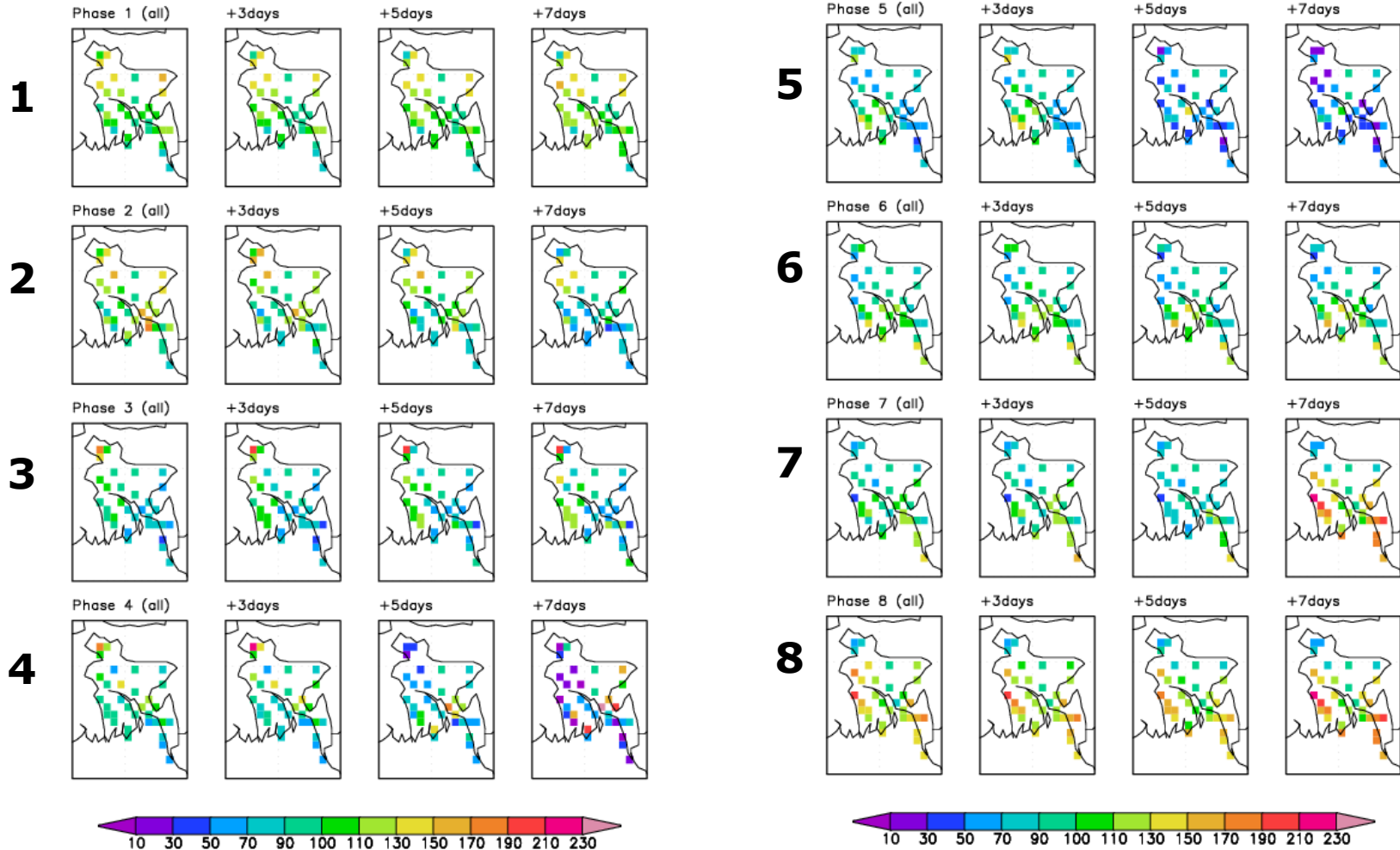


Labelled dots for each day.
Blue line is for Jul, green line is for Jun.

Wheeler and Hendon (2004)
BMRC Climate Forecasting

MJO and Bangladesh rainfall (% of climatology)

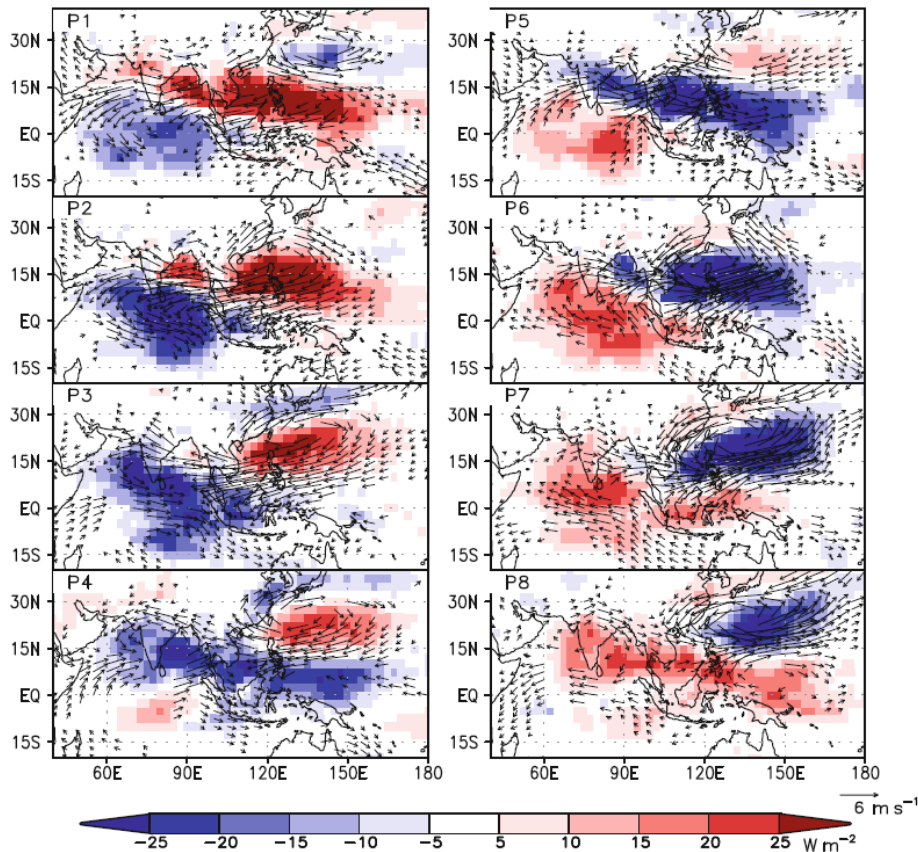
MJO duration



BSISO (Boreal Summer ISO)

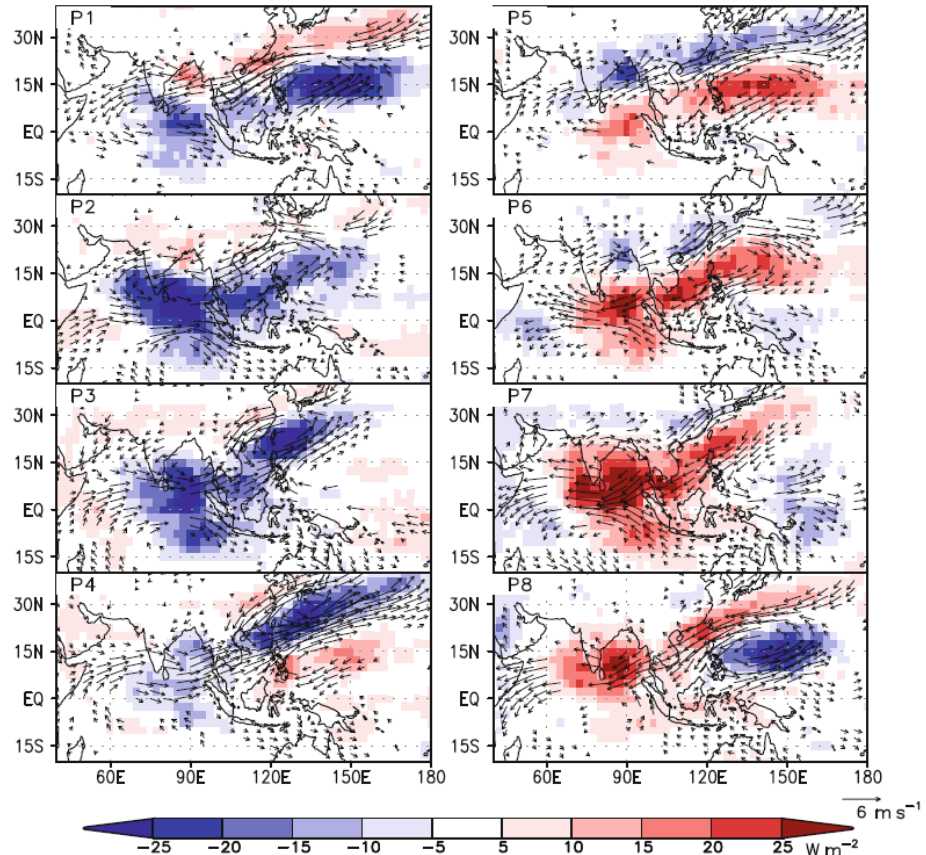
The canonical northward propagating component

BSISO1



The AMS pre-monsoon and onset component

BSISO2



Lee, J.-Y., B. Wang, M. C. Wheeler, X. Fu, D.E. Waliser, and I.-S. Kang, 2013: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Clim. Dyn.*, 40, 493-509.

BSISO forecast (May-Oct)

- Service
- 6-month Forecast
- Past Forecast
- BSISO Forecasts**
- Forecasts
- State of our climate
- CLIK
- TRACE

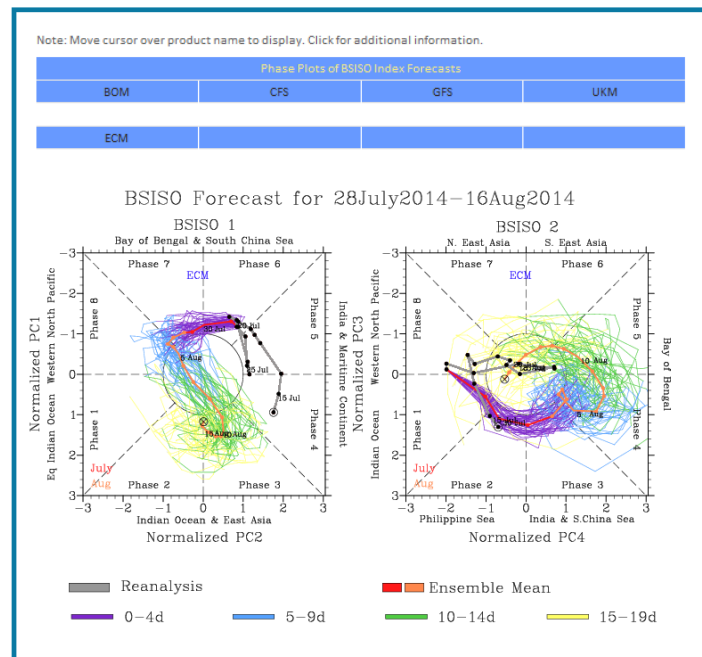
Forecasts

Welcome to the Boreal Summer Intraseasonal Oscillation (BSISO) forecast website. The BSISO forecast activity has been initiated in 2013 with the goal of improving our ability to understand and forecast the BSISO based on numerical models in cooperation with the CAS/WCRP Working Group on Numerical Experimentation (WGNE) Madden-Julian Oscillation (MJO) Task Force, and hosted at the APEC Climate Center (APCC). This website will be updated as additional models become available and verification statistics and various ways of displaying forecast information generated. Below are links to the BSISO monitoring website and the MJO model forecasts

BSISO Realtime Monitoring
Operational Realtime Dynamical Model MJO Forecasts

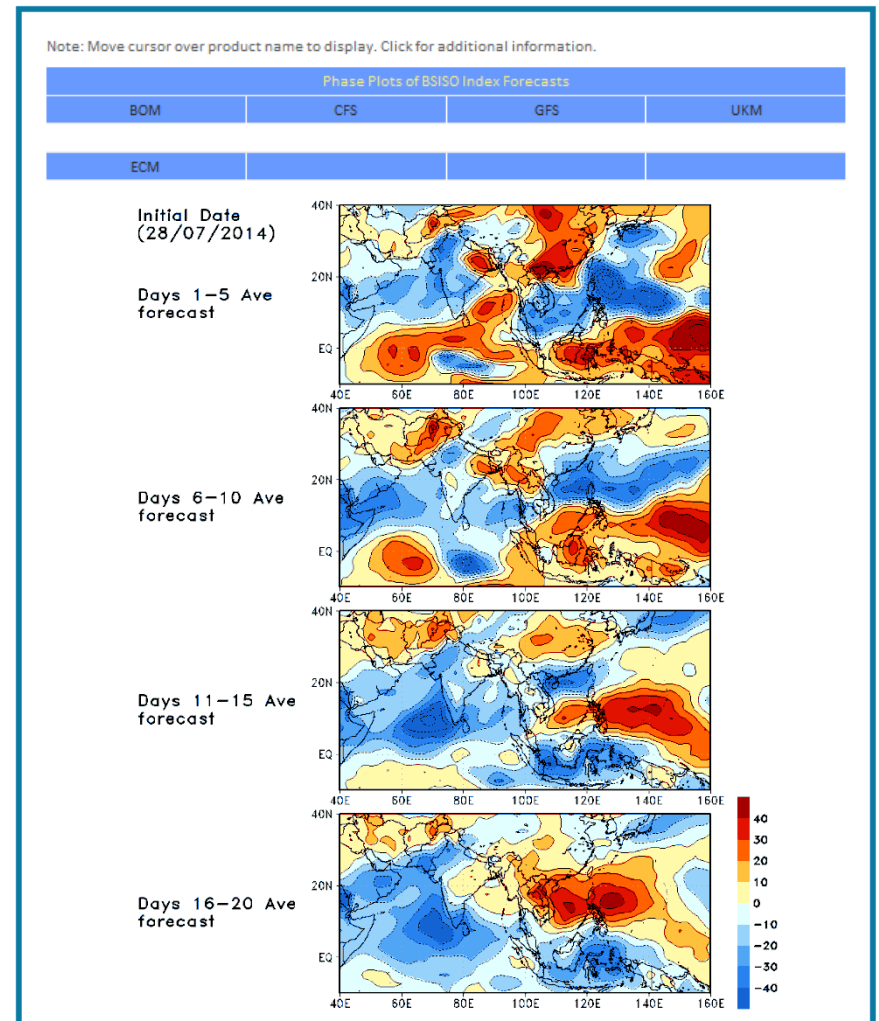
Dynamical Model BSISO Forecasts

A key for the label headings in the figure box is provided below.



Spatial OLR Anomalies

A key for the label headings in the figure box is provided below.



S2S project

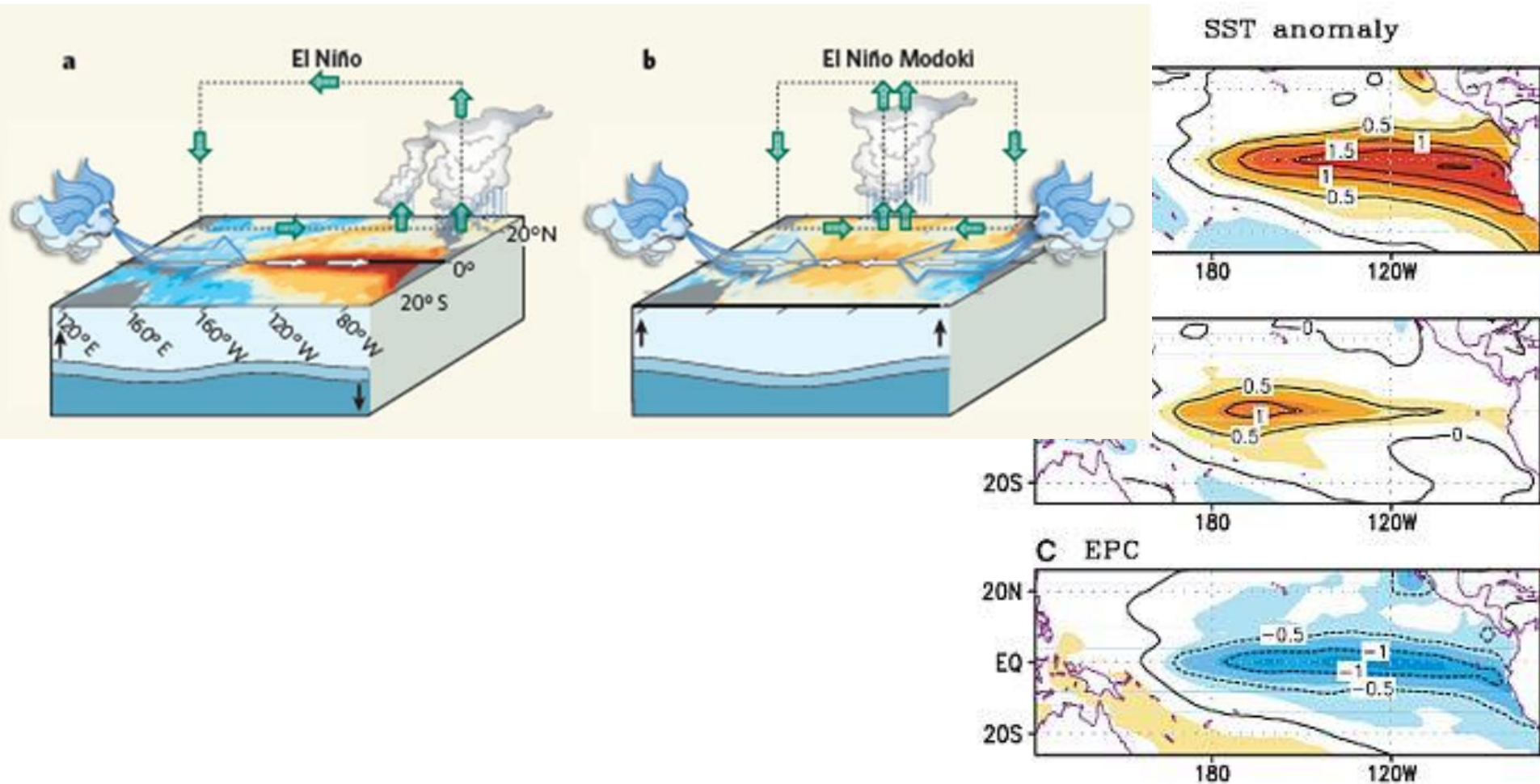
Subseasonal to seasonal (15-60days)

Objectives

- 1.To improve forecast skill and understanding on the subseasonal to seasonal timescale with special emphasis on high-impact weather events
- 2.To promote the initiative's uptake by operational centres and exploitation by the applications community
- 3.To capitalize on the expertise of the weather and climate research communities to address issues of importance to the Global Framework for Climate Services

New type of El Niño

El Niño Modoki (Central Pacific El Niño)



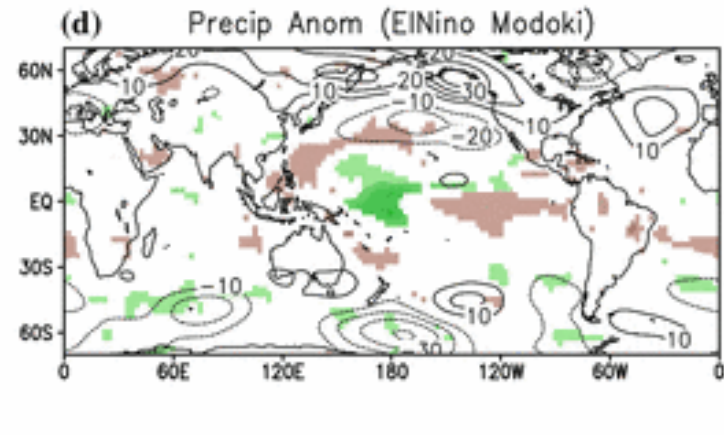
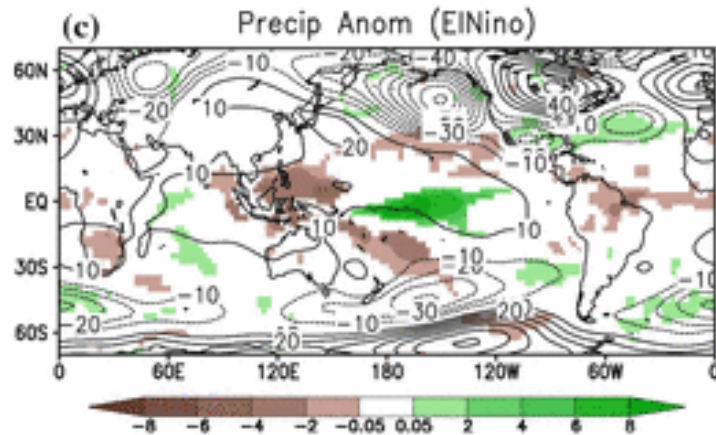
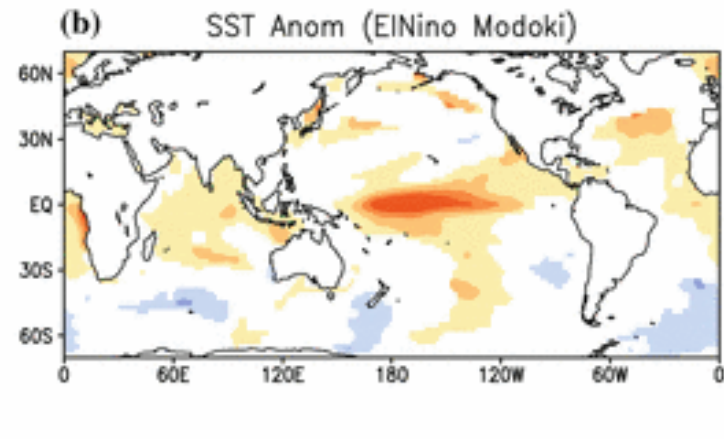
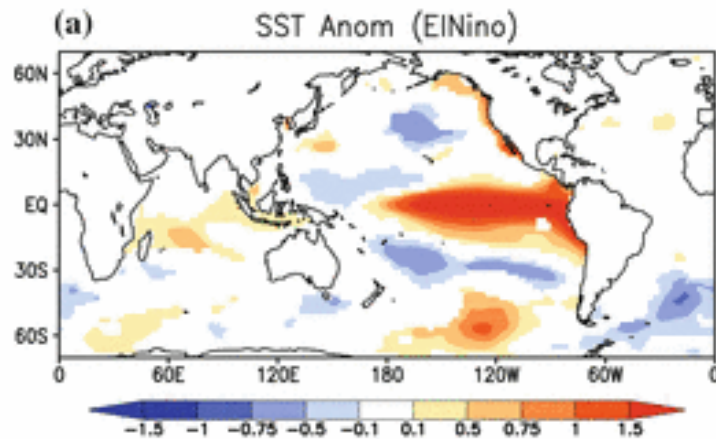
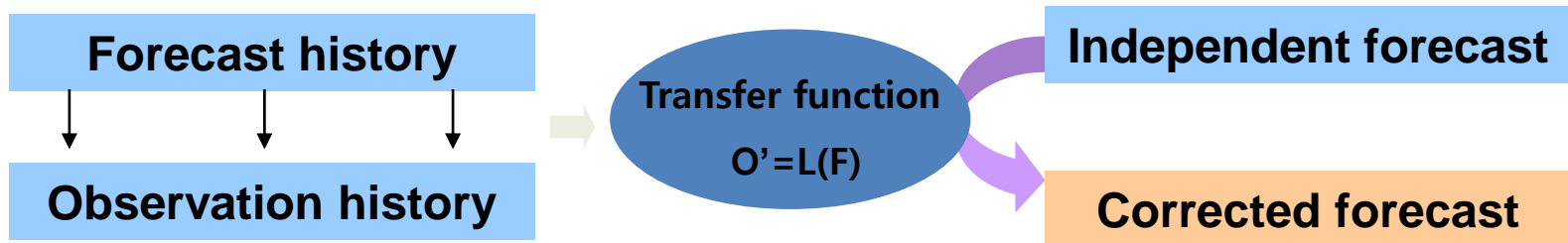


Fig. 2 a Composite observed significant SST ($^{\circ}\text{C}$) anomalies during El Niño years. b Same as a but for El Niño Modoki years. c Composite observed significant precipitation (mm/day; shaded)

anomalies and 500 hPa Geopotential Height (m; contours) anomalies during El Niño Years. d Same as c but for El Niño Modoki years. All the shaded values are significant at 90% using t test

Post-processing : CLIK



There are many approaches in post-process, All of them share similar assumption. :
Statistics between forecast and observation is stationary

If statistics is not stationary, post-process will not work in independent forecast

Thus, statistical stability is a rule of thumb in the statistical post-process (avoiding overfitting)

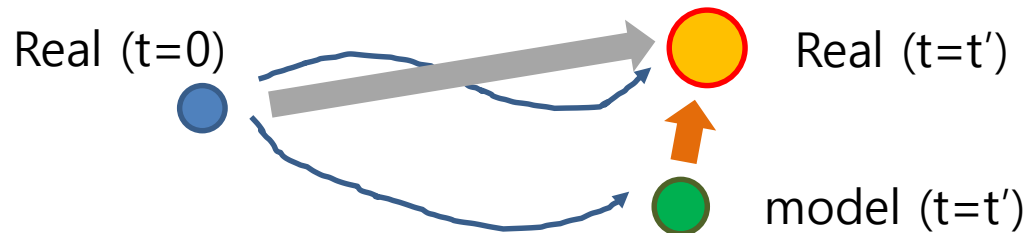
Approach

■ Statistical forecasting based on past forecast

$y(s, t)$: observation

$x(s, t)$: forecast

$$y'(t) = f(x(t), \alpha), \alpha = g(x(1 : t - 1), y(1 : t - 1))$$

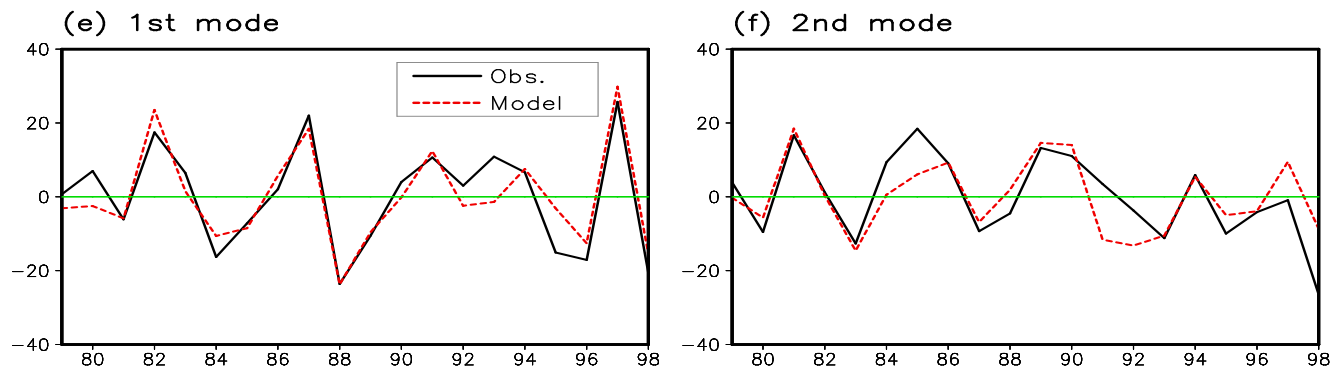
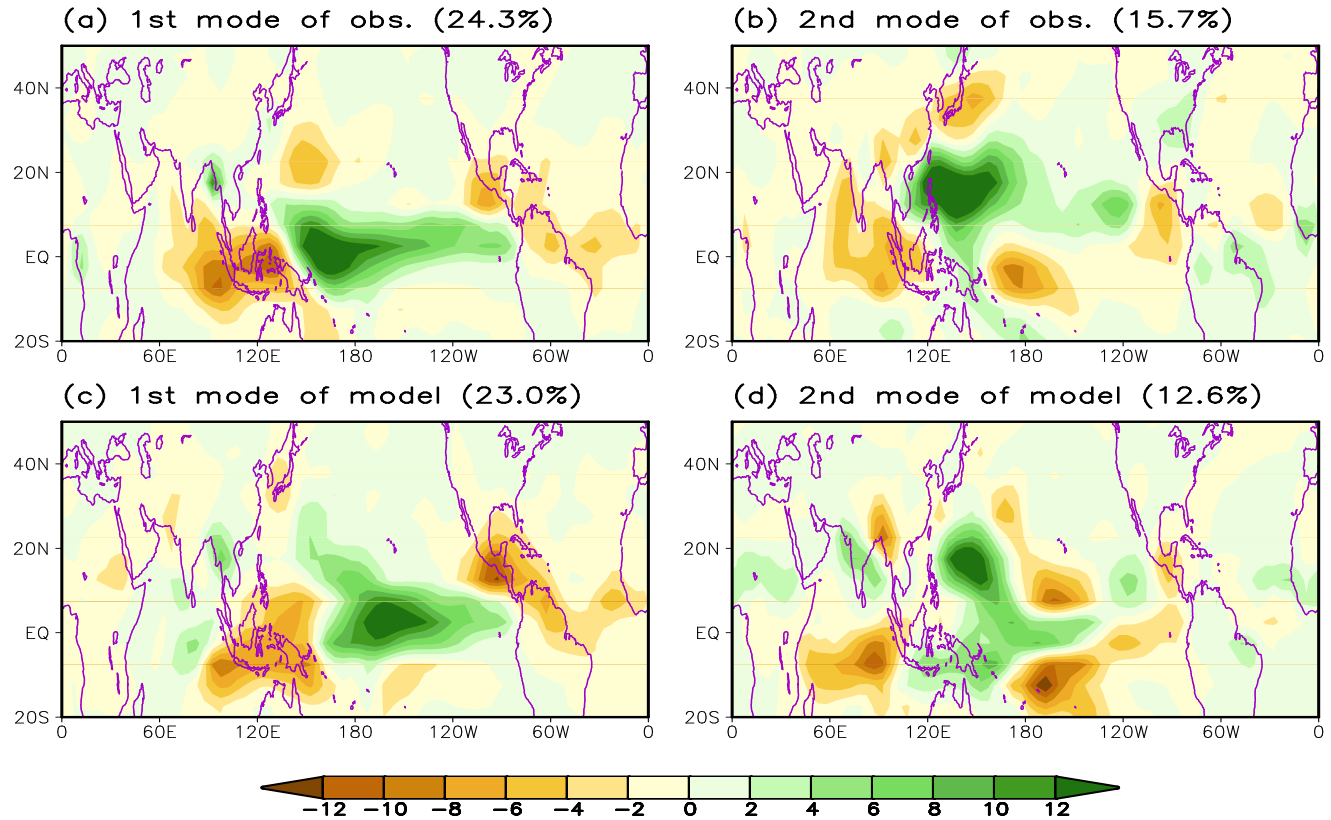


■ The most common way : Regression

$$\sum_j b_j y_j = \sum_i a_i x_i + \epsilon$$

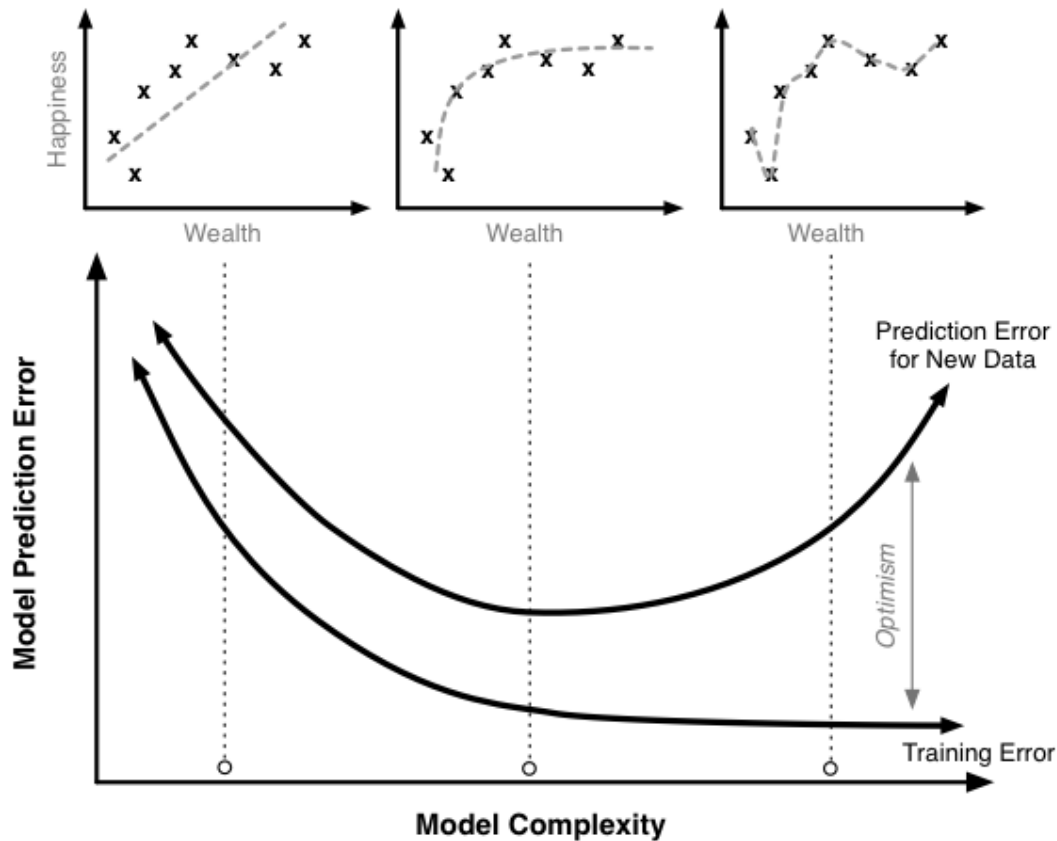
- If $i \& j = 1$: Linear regression
- $i > 1, j = 1$: Multiple regression
- $i \& j > 1$: CCA, SVD, etc

EOF of Summer Mean Precipitation



Weakness : overfitting

■ Consider potential predictability

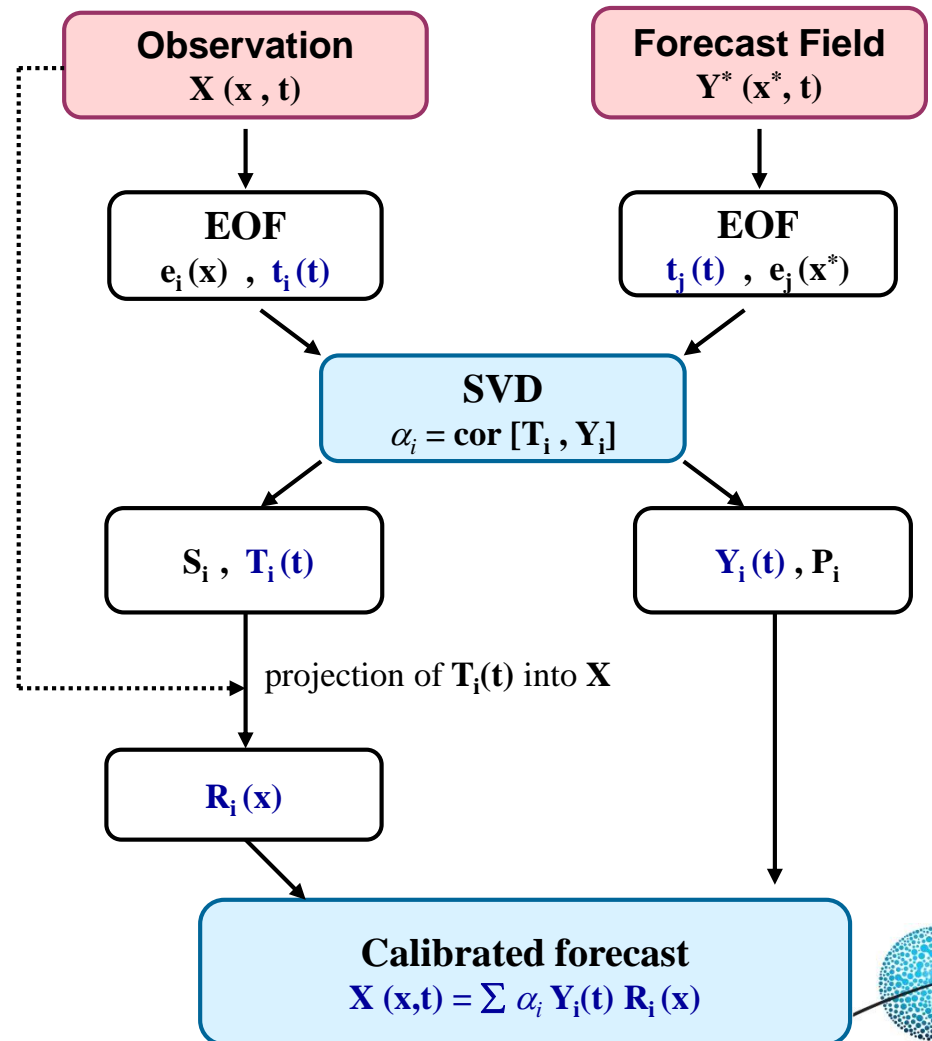
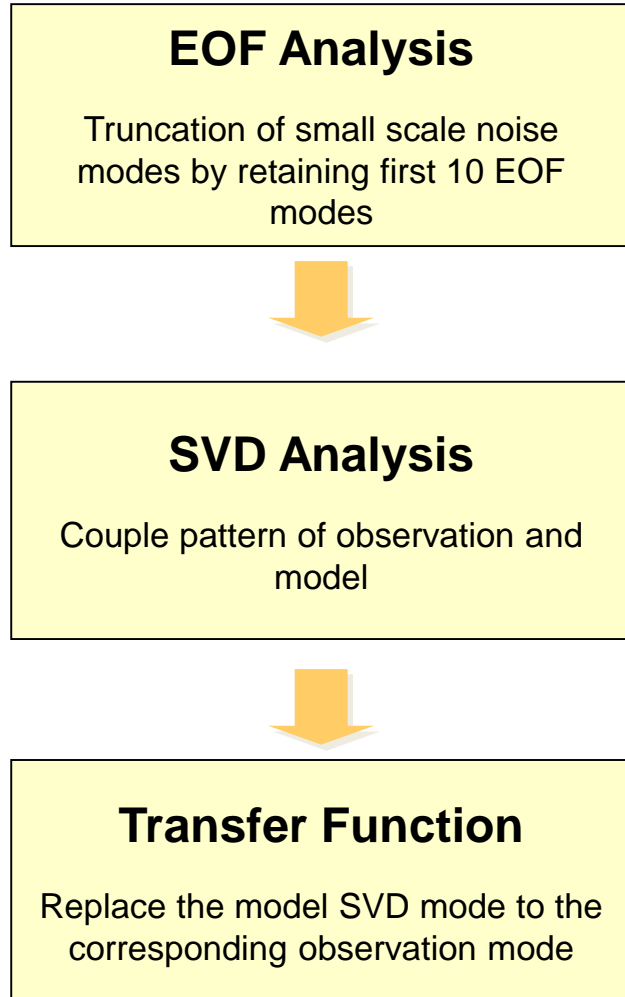


If model output is fitted to the unpredictable noise : Overfitting.
What if we remove “noise” in the observation?

from Scott Fortmann-Roe 2012

Anomaly Bias Correction (example)

Procedure of Anomaly Bias Correction



The most important thing

- Physical understanding of;
 1. What weather event/system consists of your seasonal climate (LOCAL, predictand)
 2. What external (slow varying factor) controls the weather system (GLOBAL, predictor)

And, whether model is able to predict 1 or 2

Thank you

jhyoo@apcc21.org



CLIK

<http://clik.apcc21.org>



CLIK : online prediction tool

- For those **who wants to play with model data**
- To allow **user manipulation of multi model prediction** in producing his/her own forecast
- To provide **statistical downscaling** capability using multi model prediction
 - Prediction : Diff. combinations
 - Downscaling : model to station matching

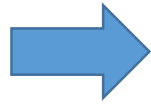
Downscaling in CLIK

- Use “observed” large scale pattern (X) associated with climate variability at stations
- X needs to be predicted by GCMs to some degree
 - X becomes predictor (user selected area)
- CLIK **does not provide any prior information** for selection of predictor (to avoid overfitting)
 - Basic knowledge on Local large scale circulation and associated global teleconnection is necessary

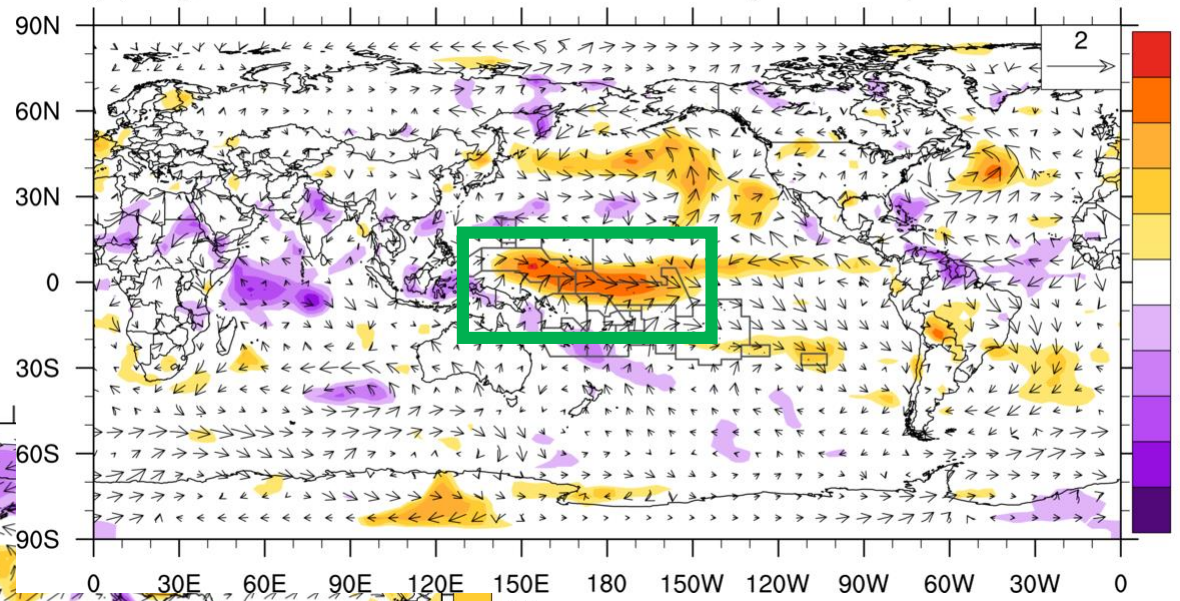
Predictor selection

Meaningful pattern? (hopeful)
: significance score

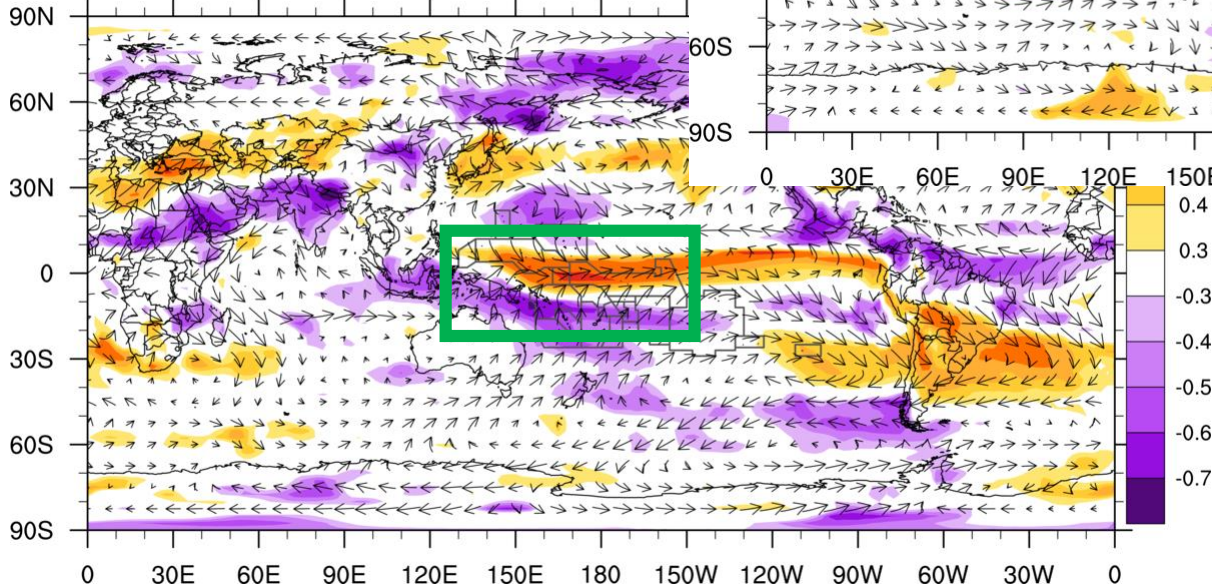
Station data



(a) Reg. between Obs. & Station 91348 [JJA - Precipitation & Wind]



(a) Reg. between MME & Station 91348



Consistency between obs. and
GCMs (good)
: pattern score

The most important thing you need is,

Patience



How to deal with many “hopeful” results?

- Combine (average) them
- Pick up one (a few) with highest corr.
- Automatic? : In the future..