Seasonal Prediction (1)
: Introduction/Predictability

Jin Ho Yoo
APEC Climate Center

APCC APEC CLIMATE CENTER

Overview

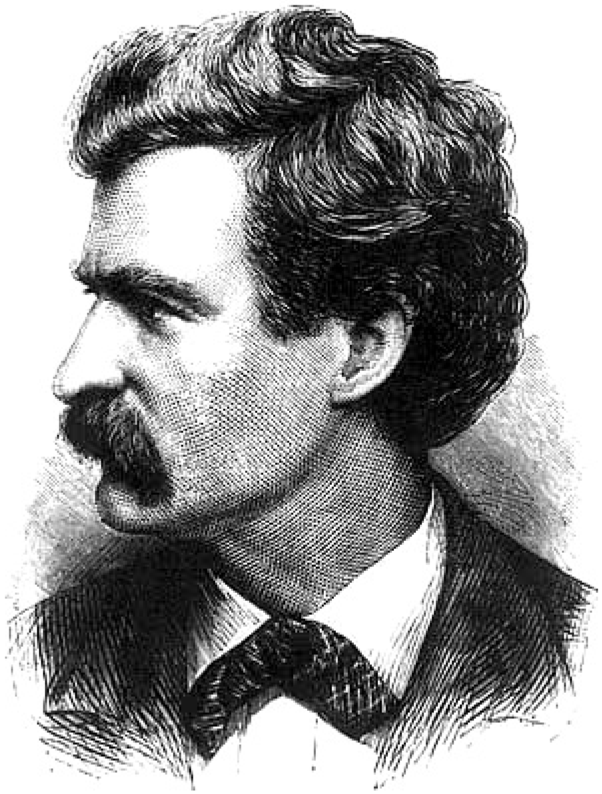
- Predictability
- Methods
- Verification + Downscaling
- Operation

Climate prediction

Climate

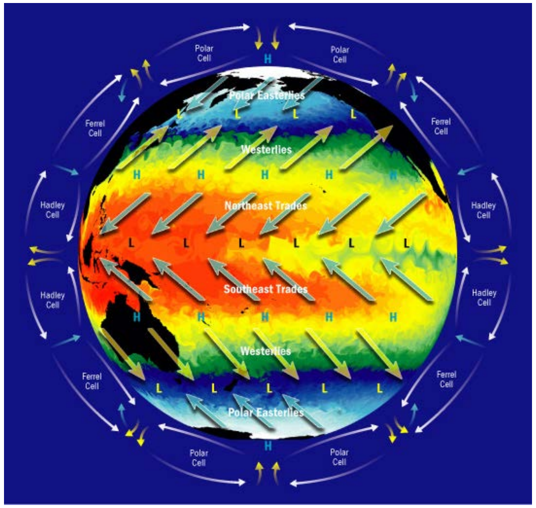
**Climate is
what we expect,**

**Weather is
what we get**



APCC
APEC CLIMATE CENTER

Climate = Expectation



APCC
APEC CLIMATE CENTER

Climate Change = We need to change our Expectation

APCC
APEC CLIMATE CENTER

Climate prediction = Expectation of Expectation

How uncertain!

APCC
APEC CLIMATE CENTER

Prediction

a rigorous, (often quantitative), statement forecasting what will happen under specific conditions

Prediction (in Meteorology)

a rigorous, (often quantitative), statement forecasting what will happen under specific conditions

What : atmospheric state

Conditions??

Atmosphere is dynamical system

$$\frac{d\vec{X}}{dt} = F(\vec{X}, a)$$

$$\vec{X}(t_0 + \tau) = \vec{X}(t_0) + \int_0^{\tau} F(\vec{X}(t), a(t))$$

APCC
APEC CLIMATE CENTER

Prediction (in Meteorology)

a rigorous, (often quantitative), statement forecasting what will happen under specific conditions

What : atmospheric state (weather)

Conditions: Current state, Physical rules, external forcing factors

APCC

APEC CLIMATE CENTER

Determinism

$$\frac{d\vec{X}}{dt} = F(\vec{X}, a)$$

Perfect prediction is possible when we have knowledge of all necessary "conditions"

Chaos

Small difference in the initial state cause huge difference later even in the deterministic nonlinear system.

$$\frac{d\vec{X}}{dt} = F(\vec{X}, a)$$

Our knowledge is never perfect!

→ perfect forecast is impossible

How well we can predict?

“Predictability”

Predictability

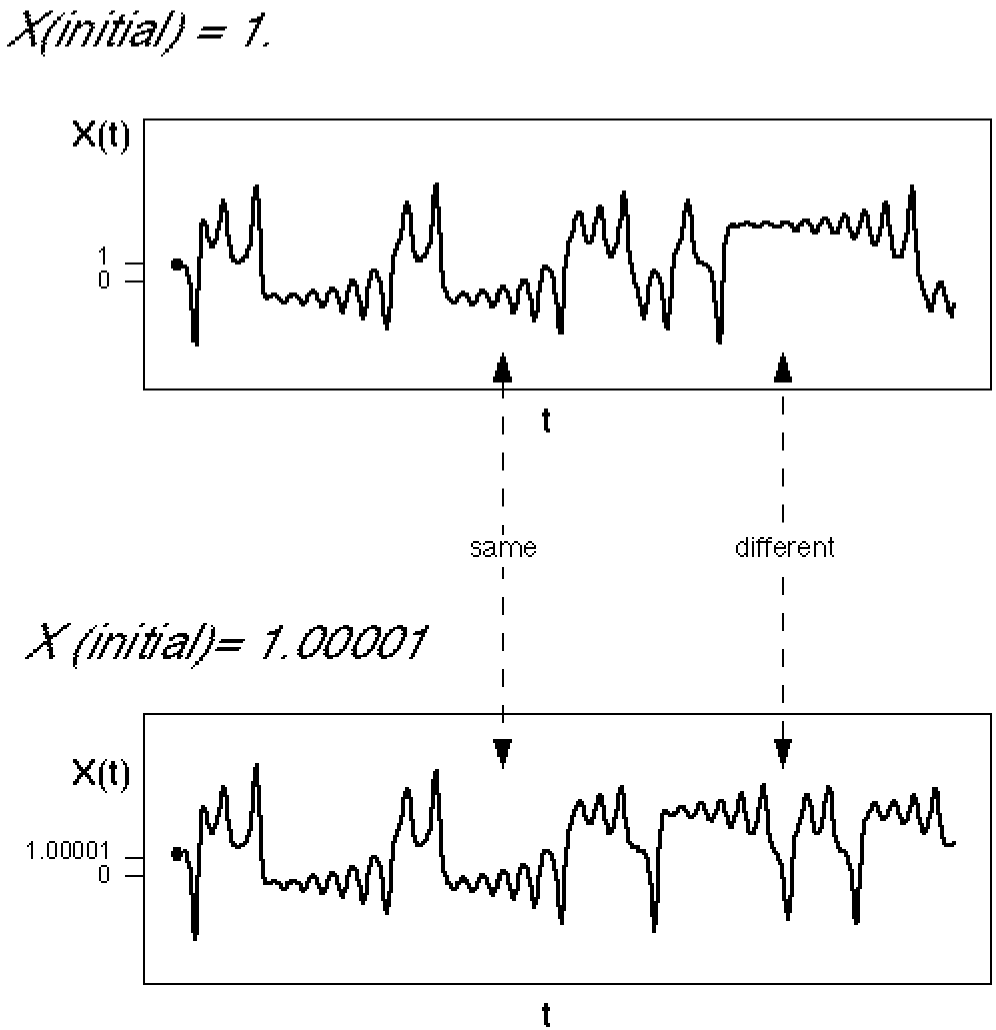
Depends on *what to predict*

Prediction of

1. *Temperature of this room tomorrow*
2. *Temperature of this room in 30days later*
3. *Temperature of this room in 30years later*

Lead time(τ)

APCC
APEC CLIMATE CENTER


X(initial) = 1.
X(t)
1
0
t
same
different
X (initial)= 1.00001
X(t)
1.00001
0
t

Predictability

Depends on *what to predict*

Prediction of

1. *Temperature of Busan*
2. *Temperature of Jakarta*
3. *Temperature of London*

Location

APEC CLIMATE CENTER

Predictability

Depends on *what to predict*

Prediction of

1. *Temperature*
2. *rainfall*
3. *wind speed*

Physical variables

Why Predictability is varying with location/variables

Characteristics of variability is different

- Tropics : weather = local convection (time scale ~ few hours)
- Extratropics : weather = synoptic system (time scale ~ few days)

- Daily rainfall is more chaotic (highly nonlinear) than temperature/pressure

Predictability

Depends on *what to predict*

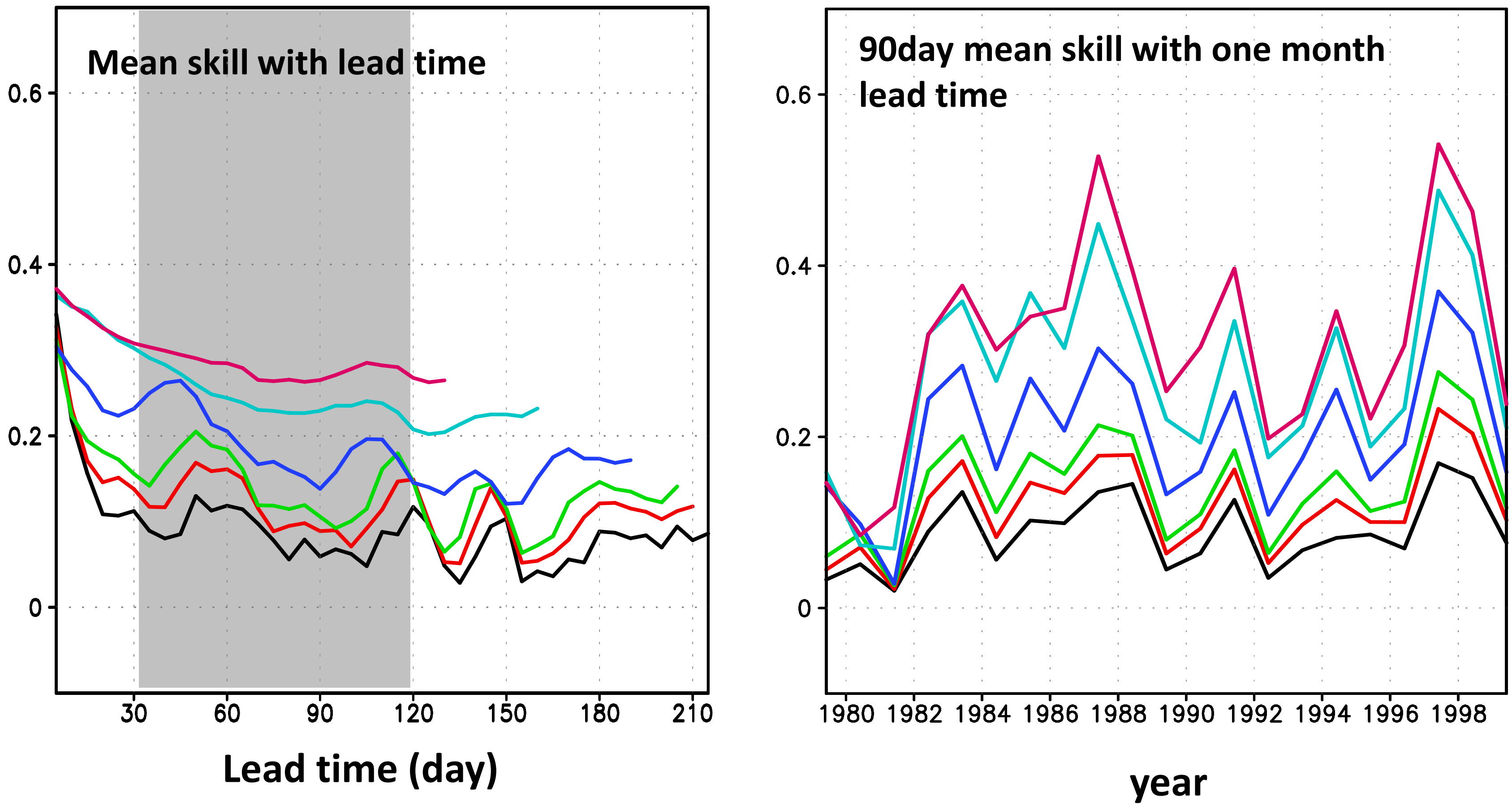
Prediction of

1. *Mean Temperature during a day*
2. *Mean Temperature during a month*
3. *Mean Temperature during a century*

Time scale of predictand

Seasonal mean and Intraseasonal predictability

Global pattern correlation skill of GCPS precipitation forecast (SMIP)



5day, 10day, 15day, 30day, 60day, 90day averaged field

APCC
APEC CLIMATE CENTER

How long?

Time mean of weather

Climate prediction

APEC CLIMATE CENTER

Seasonal forecast

Seasonal Prediction

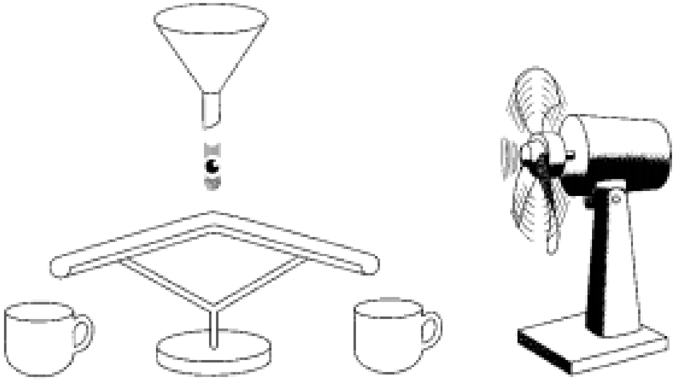
What : state of atmosphere during a season

Condition : Current state, Physical rules, external forcing factor

Lead time ~ 1 month (e.g. DJF forecast at Nov)

History of Short-term (Seasonal) Climate Prediction

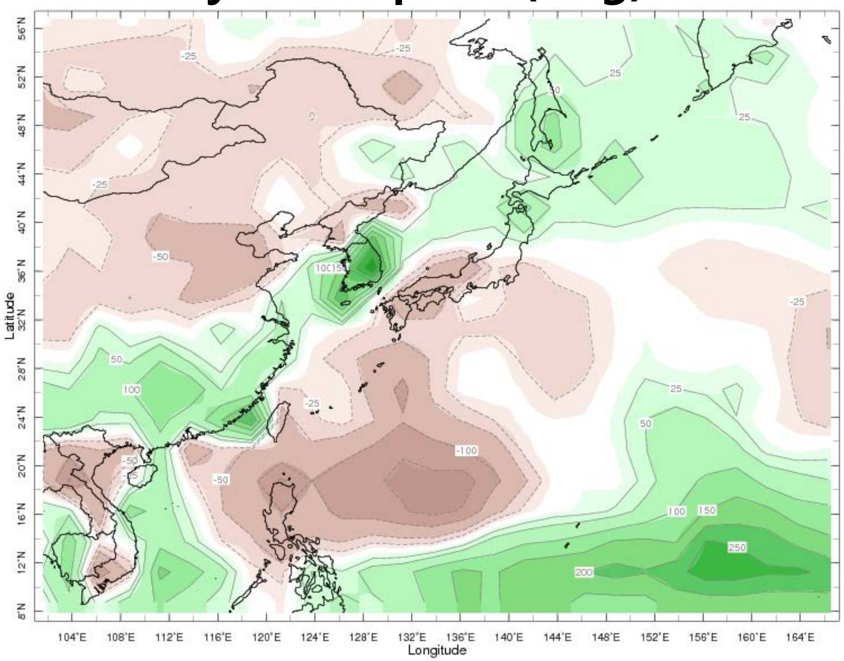
- 1960's : Hypothesis proposed
- 1980's : ENSO prediction + Atm. LFV. (PNA..)
- 1990's : (Experimental) Dyn. Seasonal Fcst.
- 2000's : International collaboration (MIPs)
- 2010's : Operation (GFCS, RCOFs/WMO)



T. Palmer (1998)

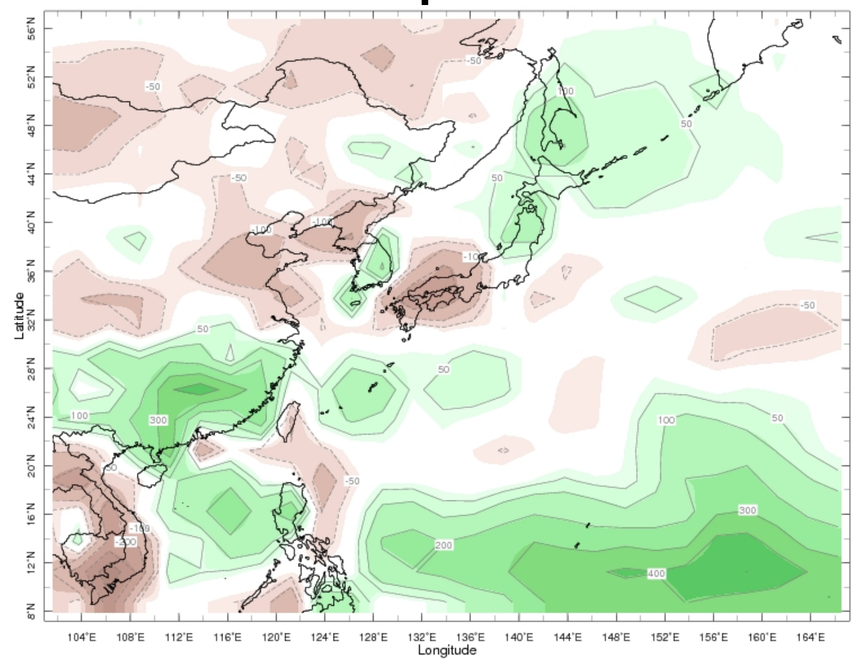
2002 summer rainfall

Monthly mean prec. (Aug)



Aug 2002

Summer mean prec.

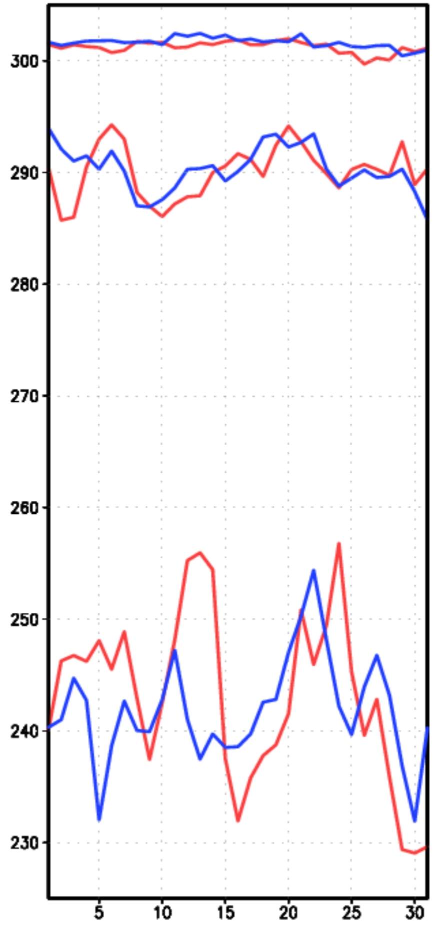


Jun-Aug 2002

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

Seasonal forecast

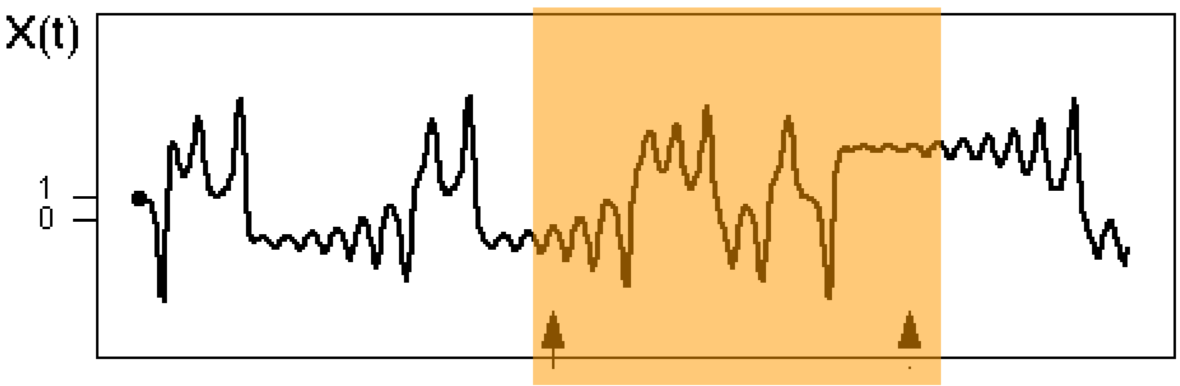
How is the seasonal mean determined?

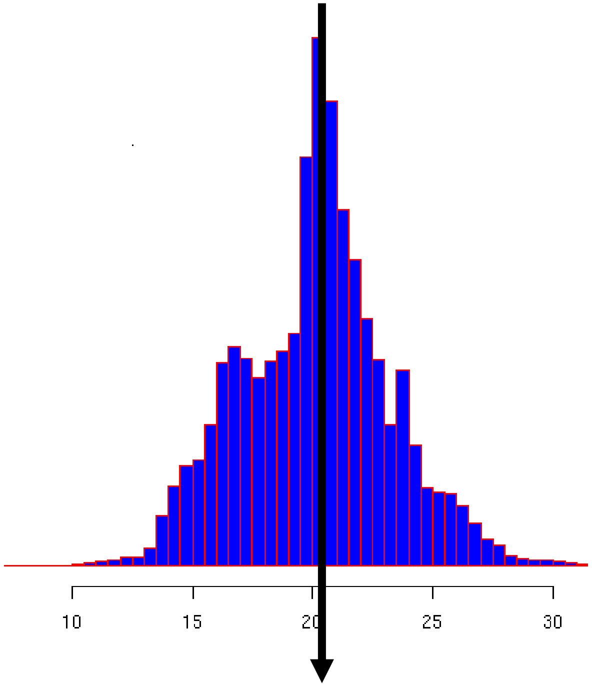


What causes change (variability) of the mean?

- **By chance?**
- **By "something"?**

Weather statistics

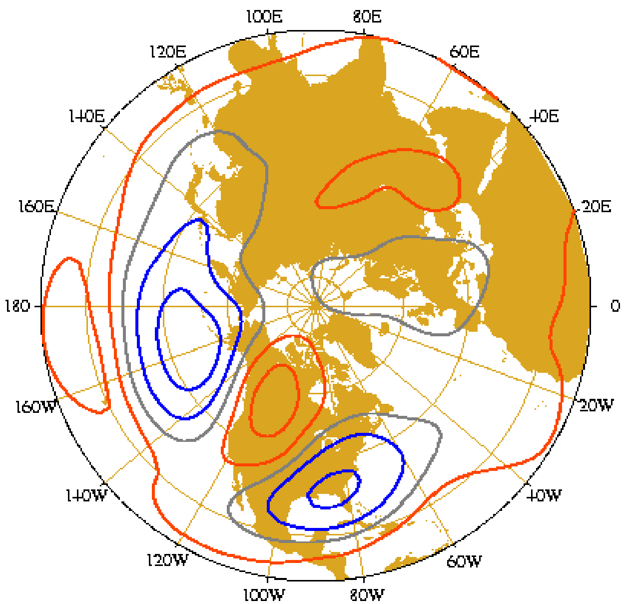




Seasonal mean

Primary seasonal weather statistics : seasonal mean

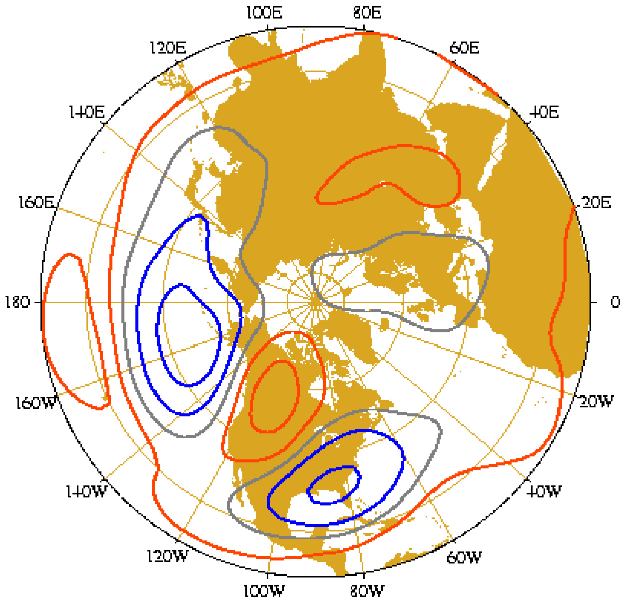
PNA debates



1. Forced by El Nino

2. Atmospheric internal variability (random)

PNA debates



1. Forced by El Nino
   Predictable (signal)

2. Atmospheric internal
   variability (random)
   Unpredictable (noise)

Matter
of
Signal
&
Noise

$$X = Xs + Xn$$

APCC
APEC CLIMATE CENTER

Potential predictability

Measured by relative magnitude (variance) of signal and noise

Signal >> Noise : more predictable

Signal << Noise : less predictable

Signal in Seasonal prediction

- What is the Signal? (How we can "see"?)
  - Tendency of weather that has be physically caused by slow varying processes
- What derives the Signal?
  - External forcing (or interaction)
  - Slow varying processes (ENSO)

APCC
APEC CLIMATE CENTER

Mechanisms of Variability

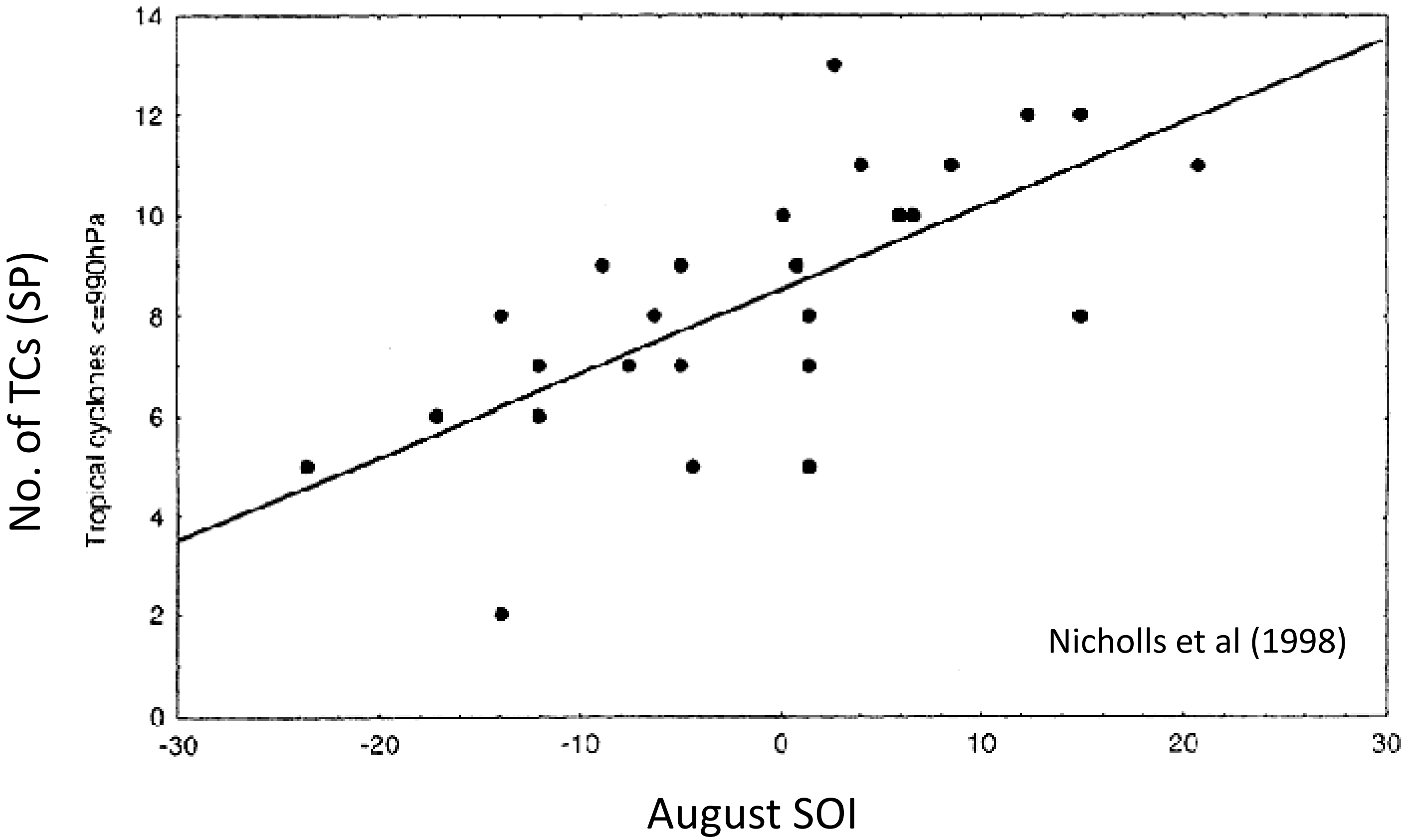
	Internal	External
Weather:	1. Internal Dynamics of Atmosphere	• Boundary Condition of SST, Soil wetness, Snow, Sea ice, etc.
Climate: (seasonal-decadal)	2. Internal Dynamics of Coupled Ocean-Land-Atmospshere	• Solar, Volcanoes
Climate Change:	3. Internal Dynamics of Sun-Earth System	• Human effects: (Greenhouse gases, land use changes)

From J. Shukla (2007)

Two scales

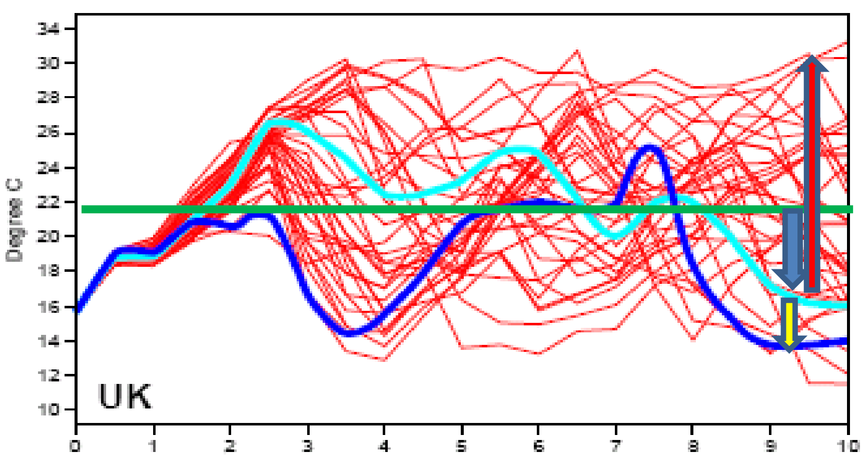
- Fast and small scale processes : noise
  - Weather, Tropical cyclone
- Slow and large processes : signal
  - Climate, ITCZ, ENSO

Two scales


No. of TCs (SP)
Tropical cyclones <=990hPa
0
2
4
6
8
10
12
14
-30
-20
-10
0
10
20
30
Nicholls et al (1998)
August SOI

Predictability

- Relative ratio between signal and noise
- BUT we don't know actual signal
  - Estimation of potential predictability by models
  - Ensemble prediction

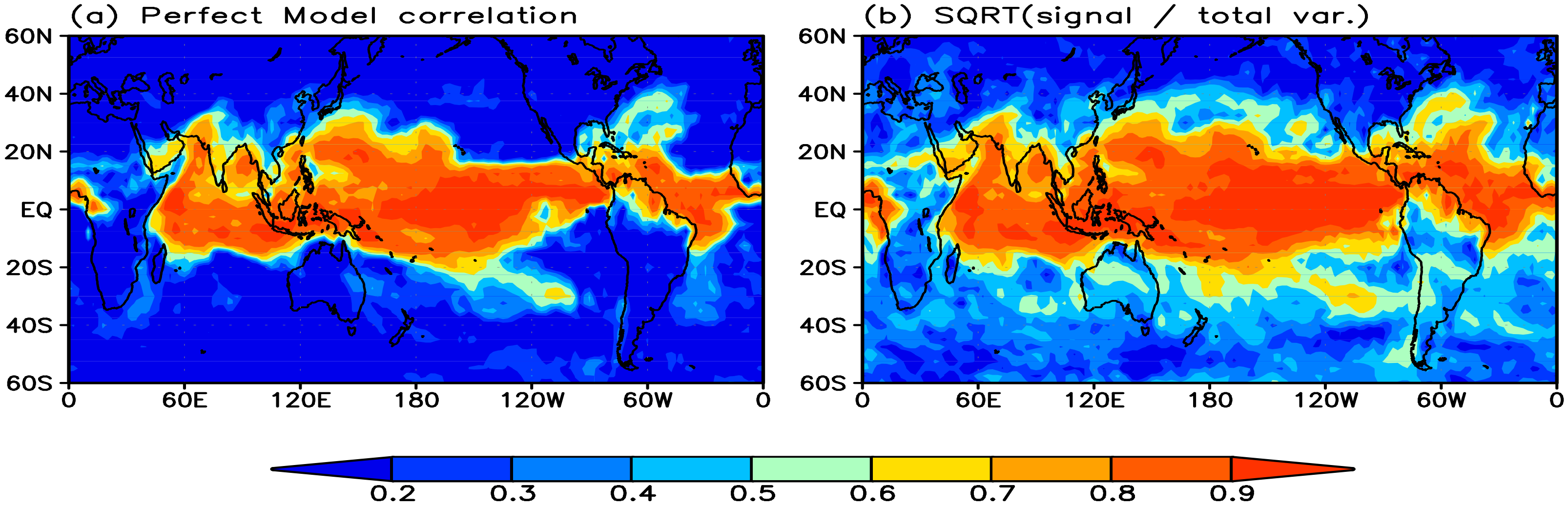


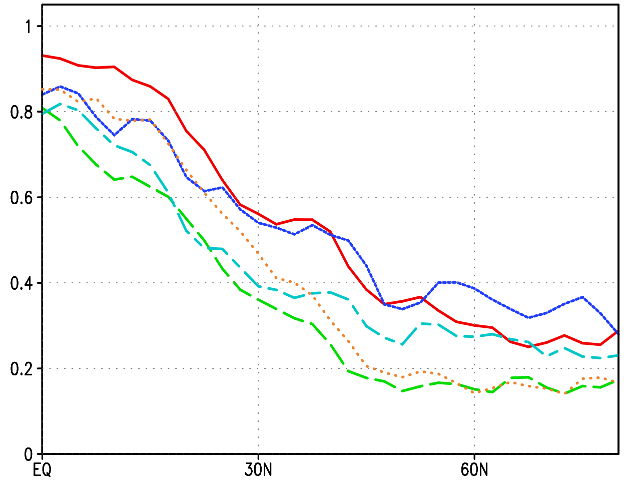
$X = X_s + X_n$

X_s : ensemble mean

X_n : deviation from ensemble mean

Estimated potential predictability of rainfall





Potential predictability

- Estimated limit of the predictability given prediction methods (model)
  - Depends on nature itself as well as prediction model
  - We cannot change the nature but model is our product
  - Potential predictability may be able to be improved (or not) if our model is improved

Seasonal Prediction (2)
: Methods

Jin Ho Yoo
APEC Climate Center

APCC 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?

APCC
APEC CLIMATE CENTER

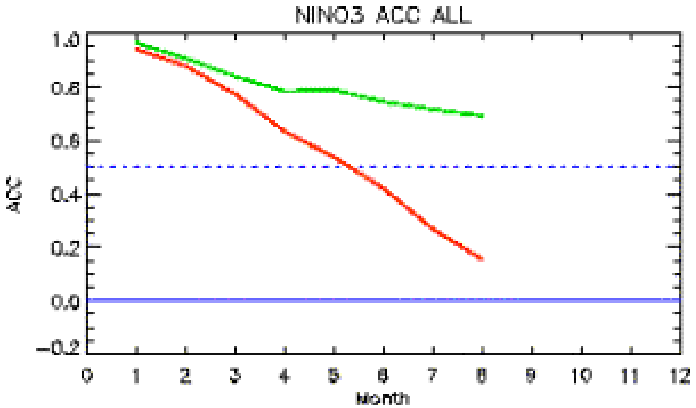
Statistical forecasting

- (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%)

APCC
APEC CLIMATE CENTER

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)



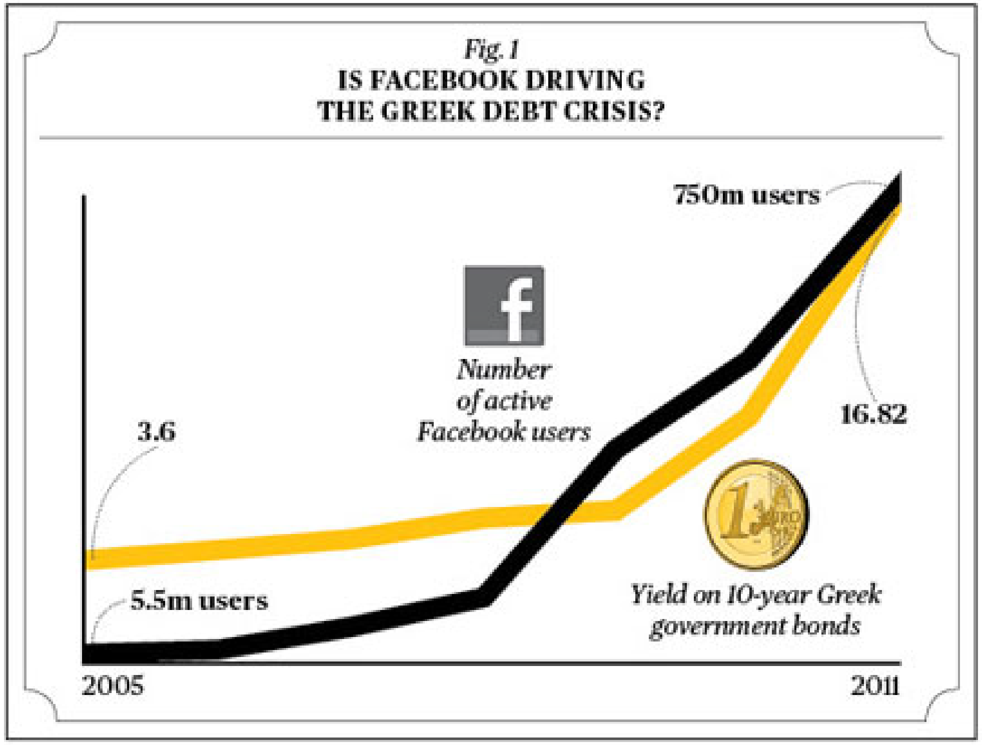
Statistical forecasting

- (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)

APCC
APEC CLIMATE CENTER

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?

APCC
APEC CLIMATE CENTER

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

APCC
APEC CLIMATE CENTER

Regression based forecast

- Question #2 : appropriate Function

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

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

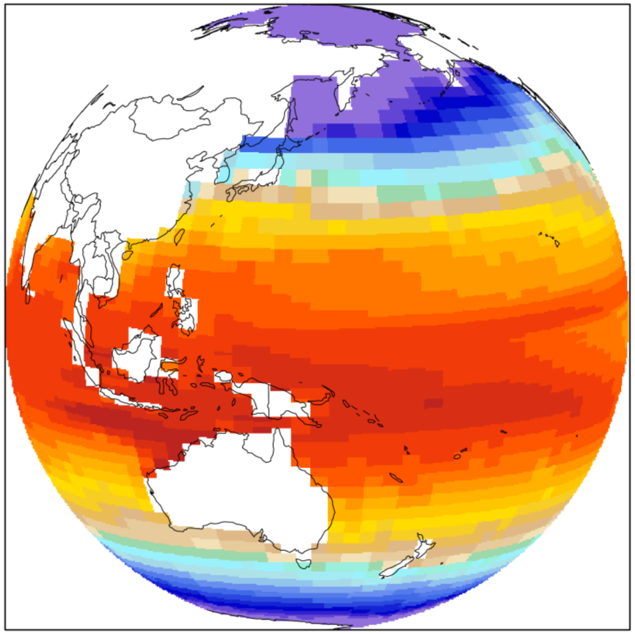
$$x_1'(t+1)+x_2'(t+1) = a_1y_1(t) + a_2y_2(t)b$$

- 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

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)

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

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

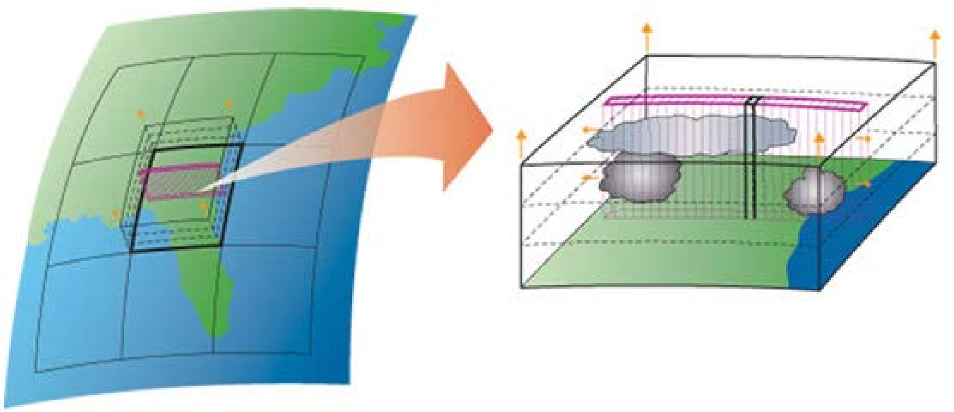
$$\frac{\partial \theta}{\partial t} + \vec{u} \cdot \vec{\nabla} \theta = l$$

$$\vdots$$

```
//Belehradek functions:
//
//     DTstage(T+9.11)^-2.05
//
    MinTime=pow(T[j]+9.11, -2.05);//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 4oC.
//
//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; //Overrride temp dependent mortality
```

Numerical modeling

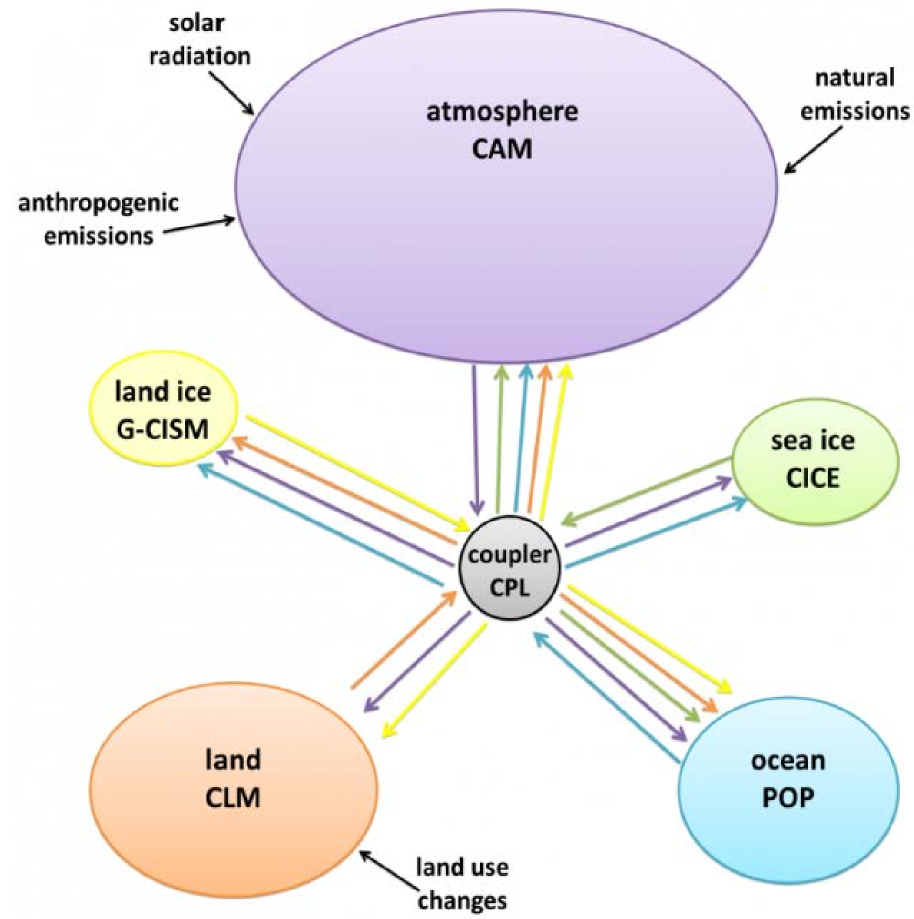
- Issue
  - Digitization (physical variable is continuous, but computer needs digitization”
    - Resolution, subgrid-scale parameterization



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

GCMs

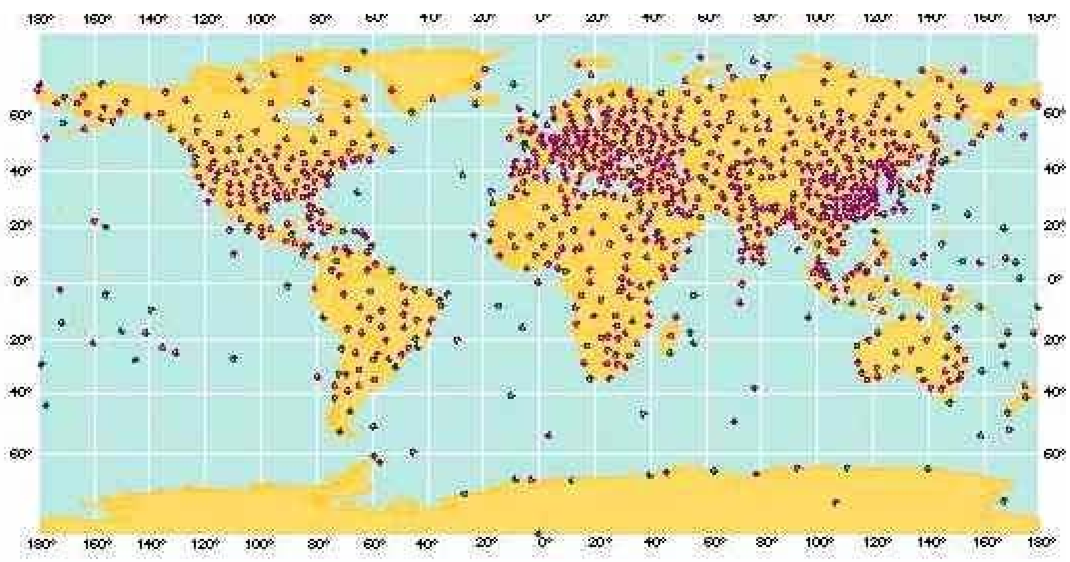
- Coupled GCM
  - Atmosphere
  - Ocean
  - Sea-Ice
  - Land surface
  - Chemistry
  - Biosphere

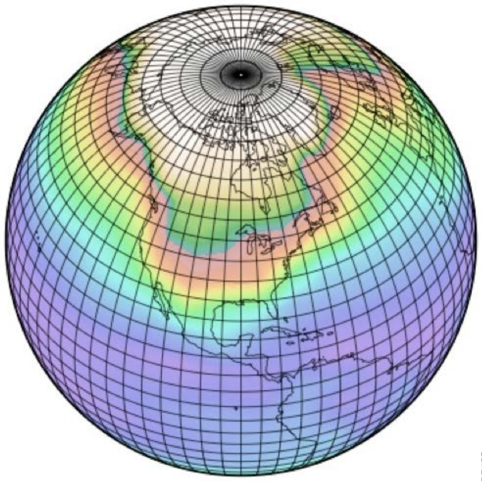


Initialization

Estimating Current status of climate system

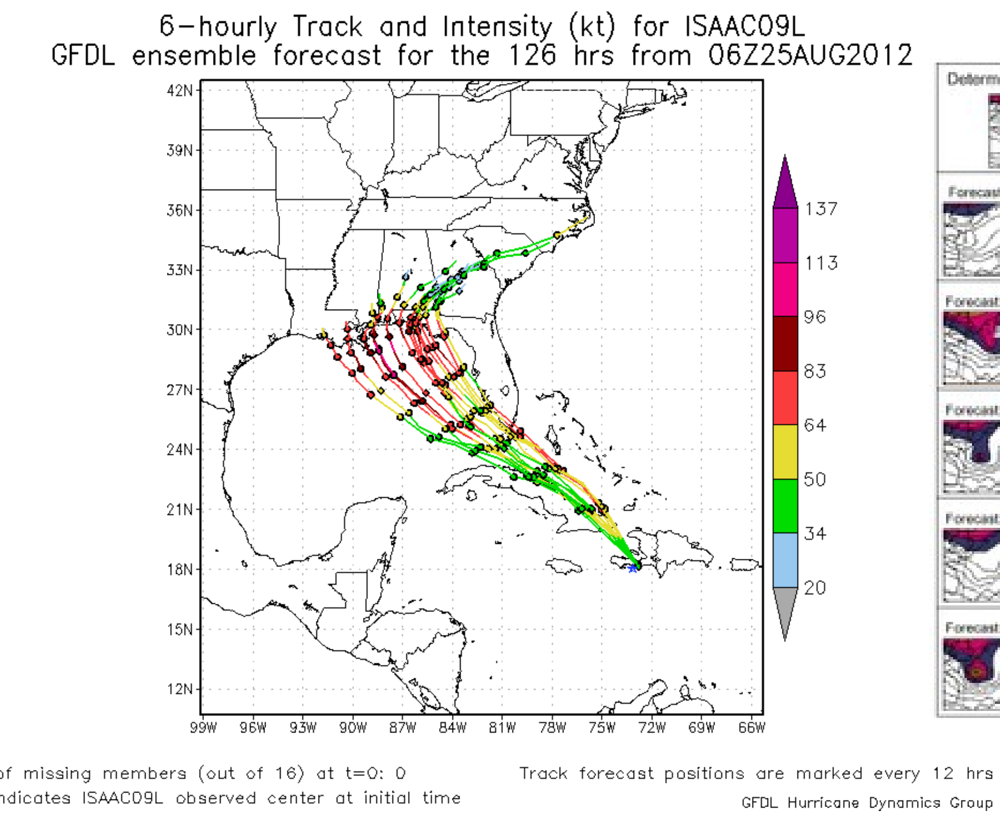
- Preparing the beginning climate state of GCM with available observation
  - Balance between Wrong GCM vs Wrong OBS.
  - Balance between components (Atm, Ocn)

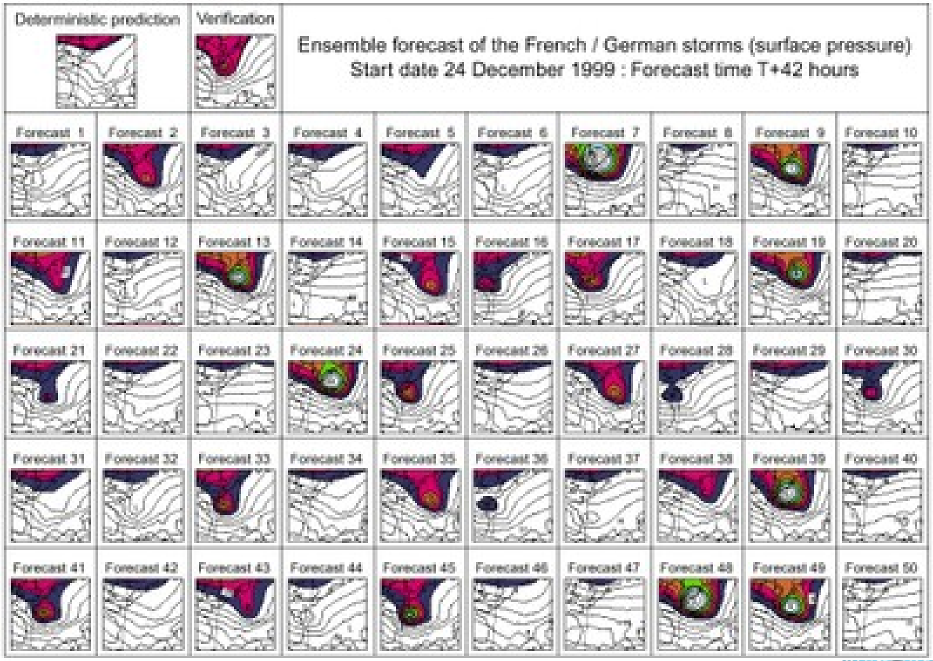




Ensemble Forecasting

- Run many times
  - Starts from slightly different initial conditions

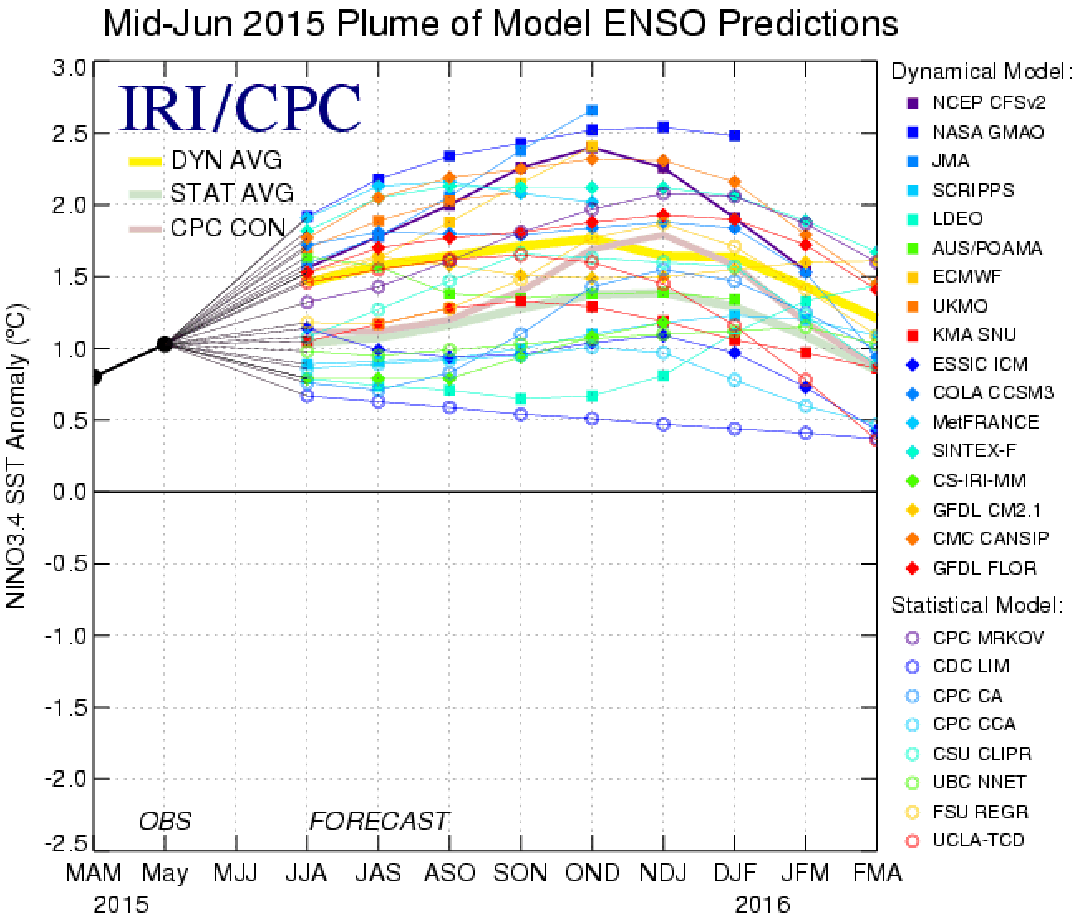




APCC
APEC CLIMATE CENTER

Multi Model Ensemble Forecasting

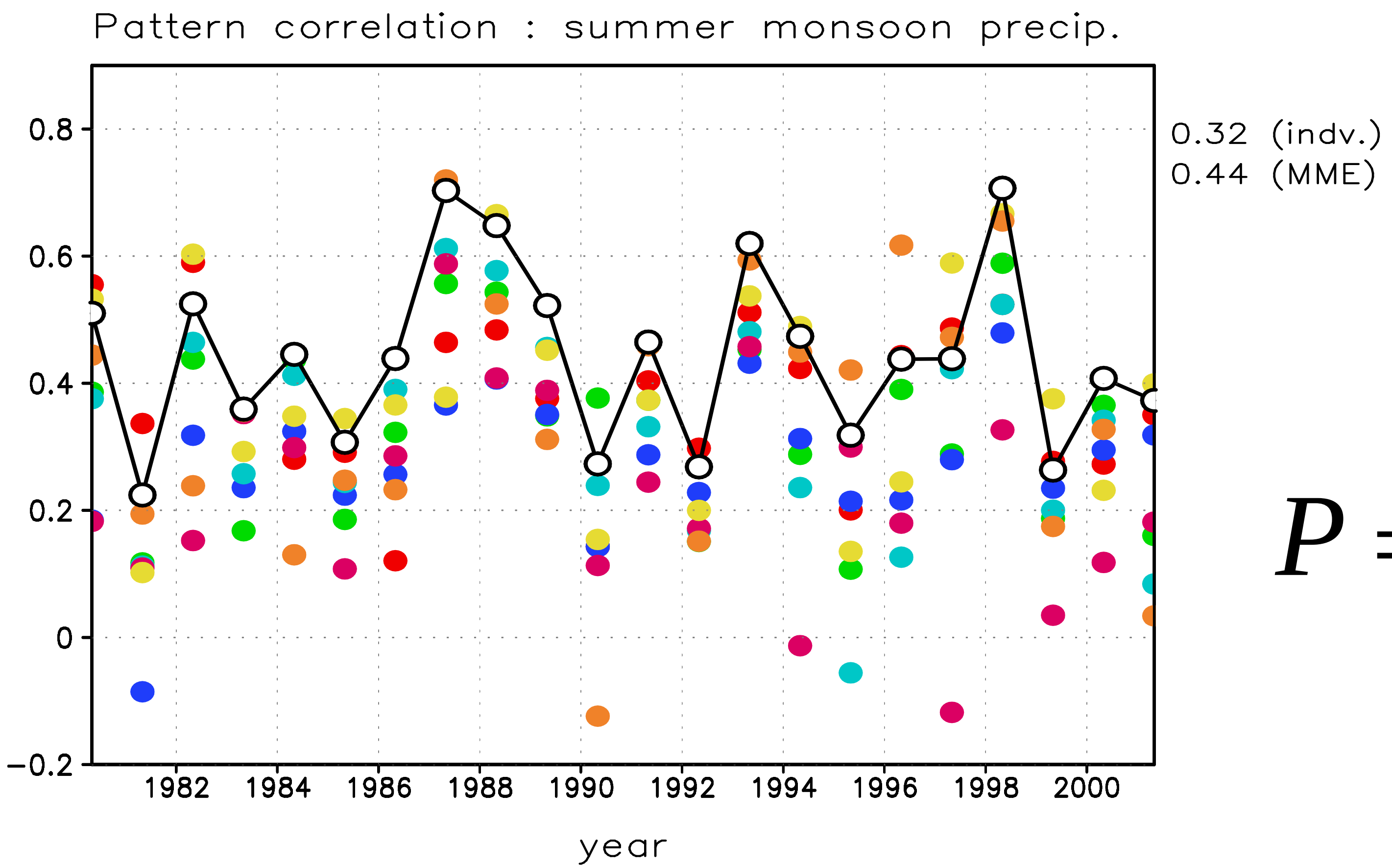
- Run with many models



Which one??

APCC
APEC CLIMATE CENTER

Use all!



$$P = \sum_i a_i F_i$$

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

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

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

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

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

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

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

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

$$Observation : x = x_s + x_n$$

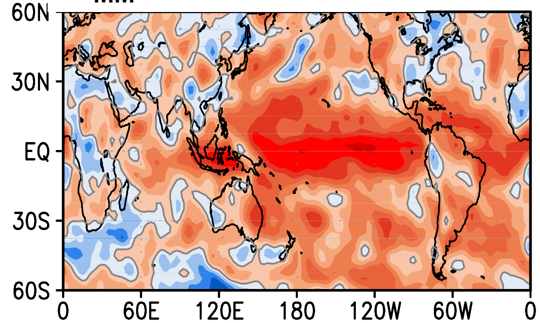
$$Forecast : y = y_s + y_n = x_s + e + y_n$$

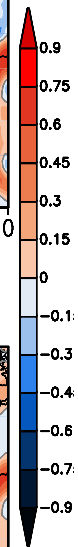
APEC CLIMATE CENTER

Temporal correlation skill (SUMMER MEAN PRCP)

Multi-model ensemble correlation skill

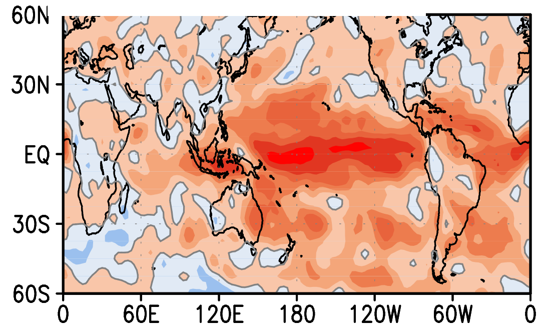
R_{MM}





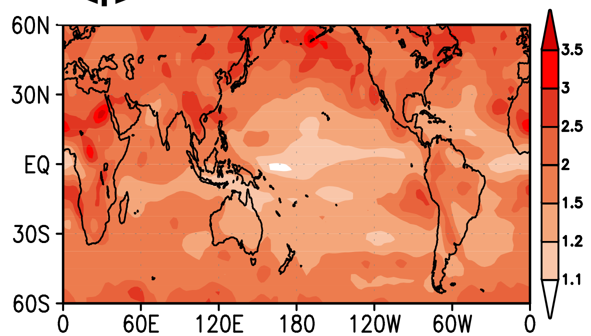
Mean correlation skill of individual models

<R>



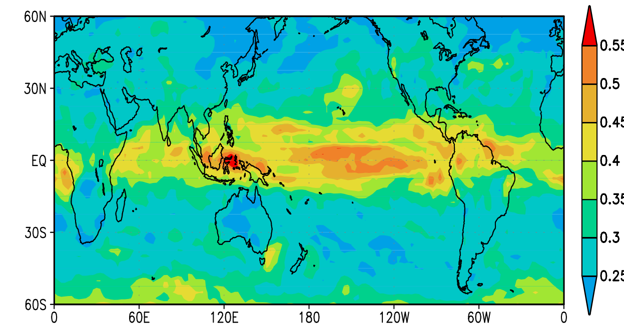
Inflation factor of correlation skill by multi-model ensemble

$\langle r\rangle^{-1/2}$



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

Contribution of systematic error (conditional) cancellation

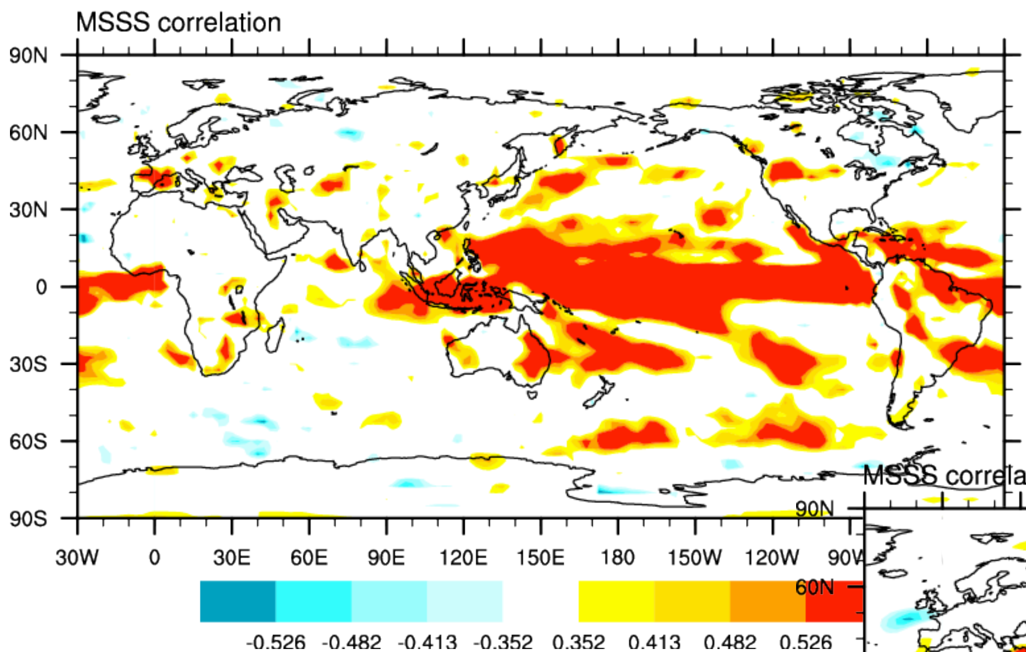


$$\boxed{R_{MM} = \frac{\langle R\rangle}{\sqrt{V(\langle y\rangle)}} = \frac{\langle R\rangle}{\sqrt{\langle r\rangle}}}$$

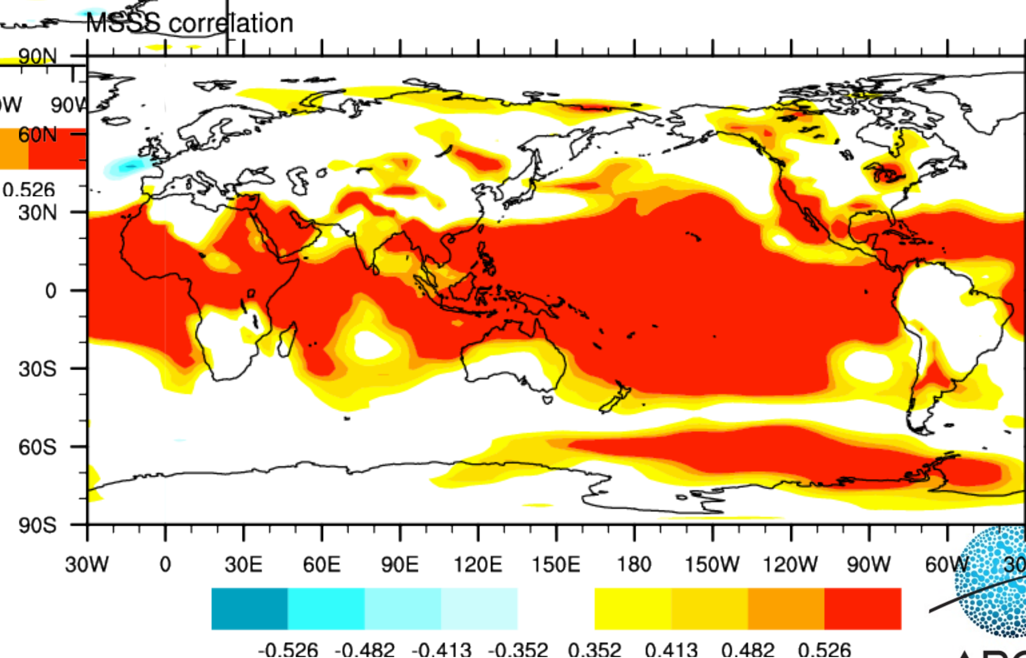
Independent and good models : Best forecast result (on average)

APCC
APEC CLIMATE CENTER

APCC MME (TCC)



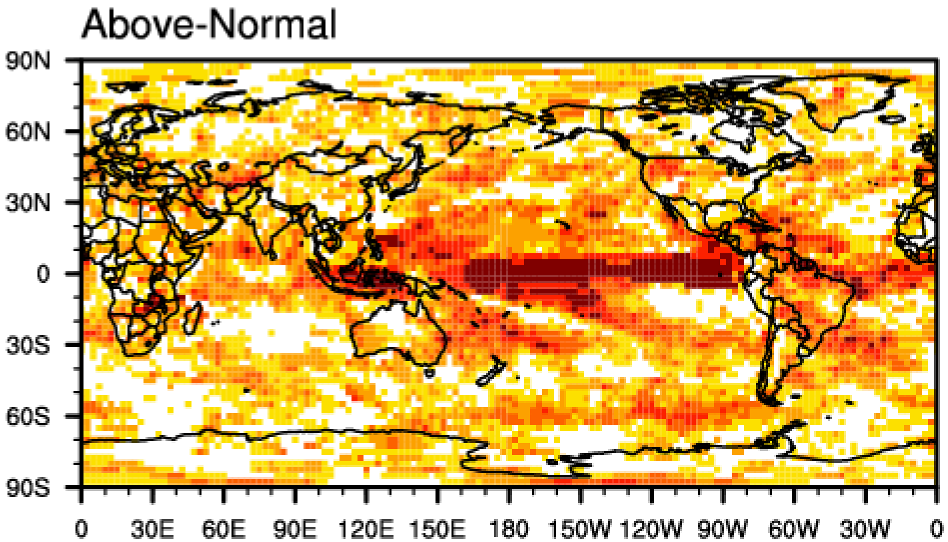
Rainfall (JJA)

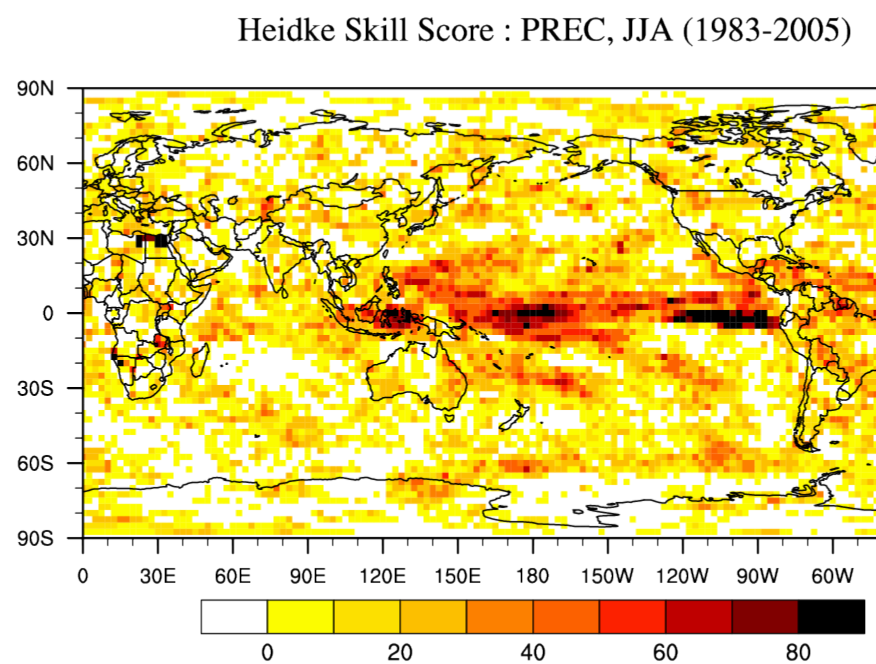


SLP (JJA)

APCC
APEC CLIMATE CENTER

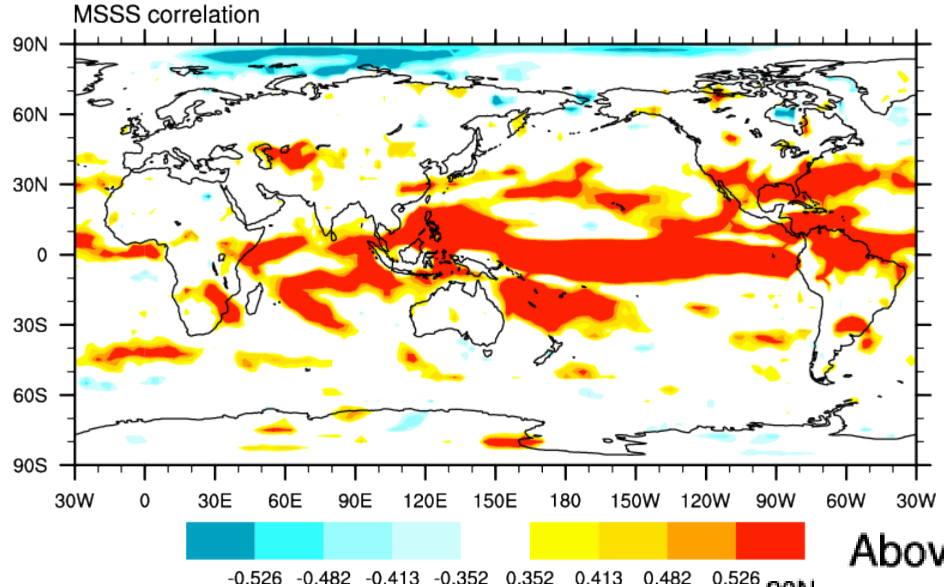
ROC Score : PREC, JJA (1983-2005)





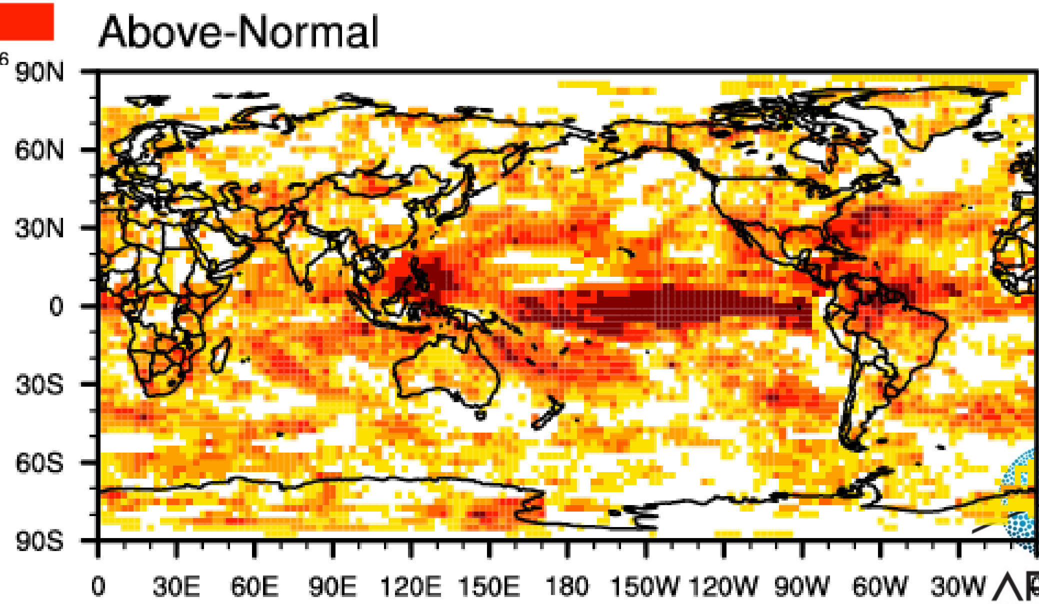
APCC
APEC CLIMATE CENTER

DJF season



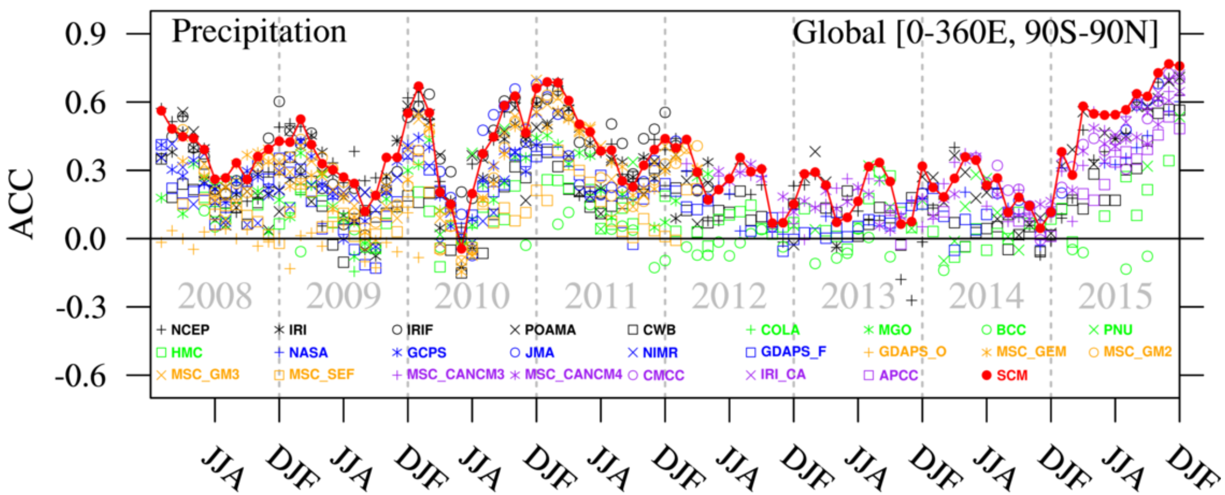
Rainfall (JJA)

ROC score
Rainfall (DJF)



APCC
APEC CLIMATE CENTER

APCC operational forecast



	APCC	WMOLC	ECMWF	NCEP	UKMO	JMA
AN	0.569457	0.541897	0.535531	0.52996	0.528975	0.531497
NN	0.520962	0.521424	0.537661	0.519823	0.524022	0.514656
BN	0.567702	0.533777	0.516511	0.535767	0.516994	0.534244

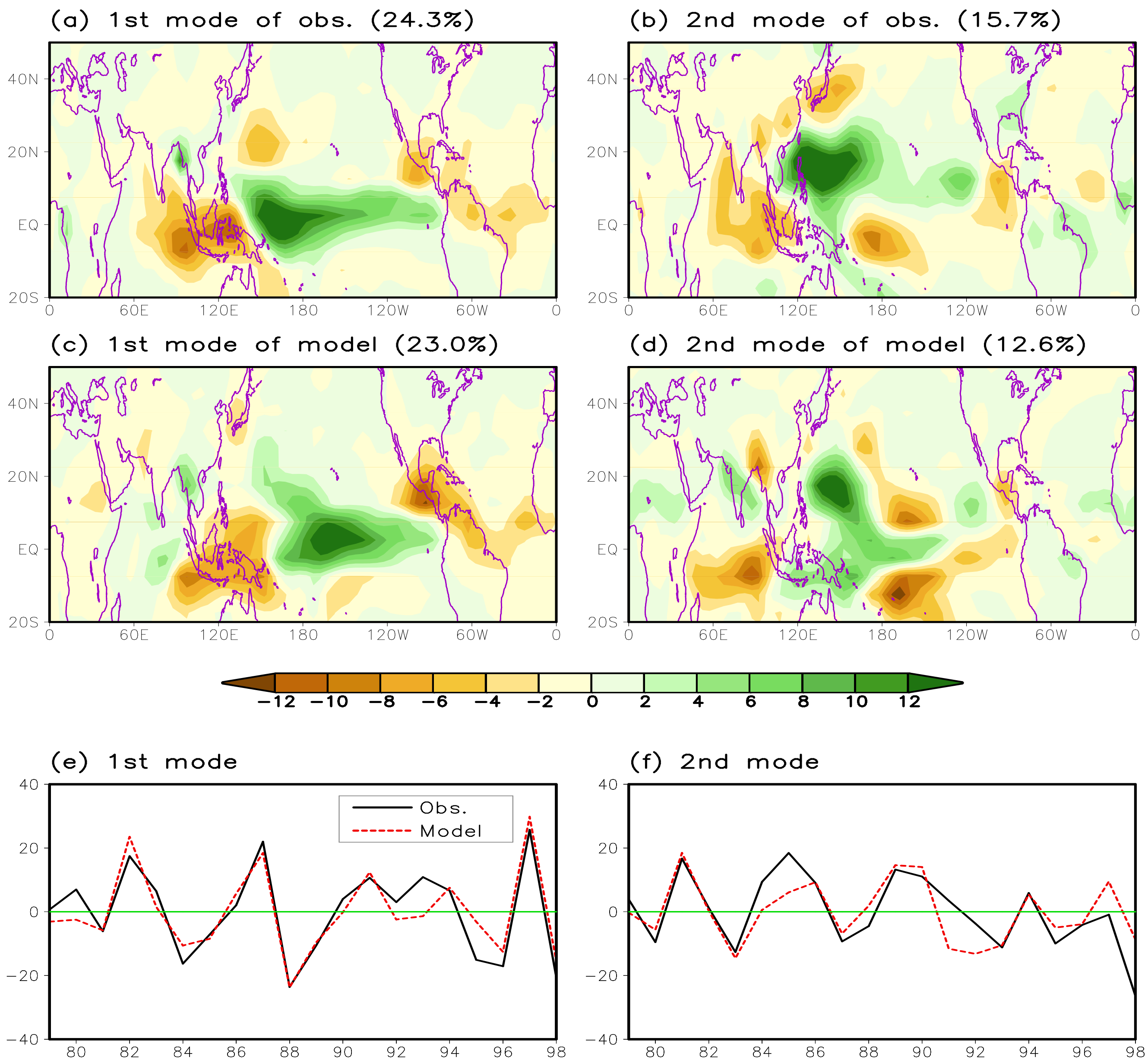
Realtime rainfall forcast for last 4 years (12-15)
ROC score : Perfect = 1, Meaningless(no skill) =0.5,

APCC
APEC CLIMATE CENTER

Even with MME,

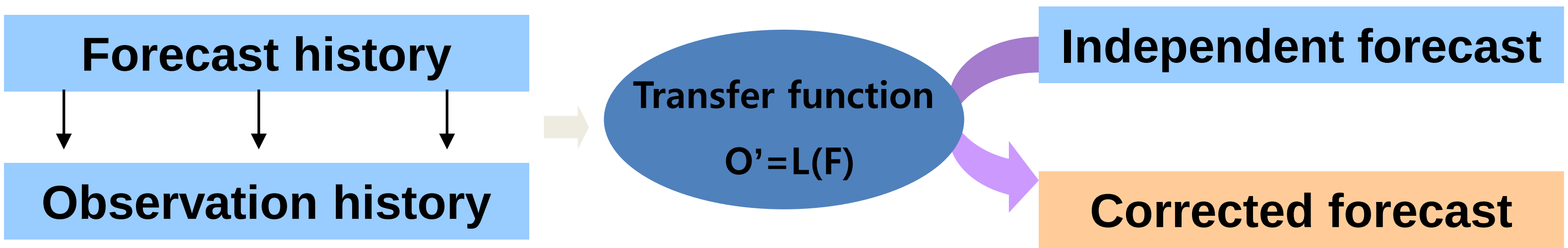
- Still many region in the world, predictability is low
- Any room for further improvement?
  - Post process

EOFs of Summer Mean Precipitation



APCC
APEC CLIMATE CENTER

Statstical downscaling : 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)

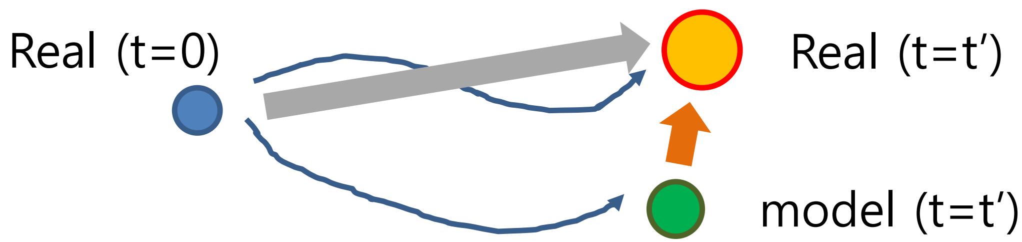
APCC
APEC CLIMATE CENTER

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**

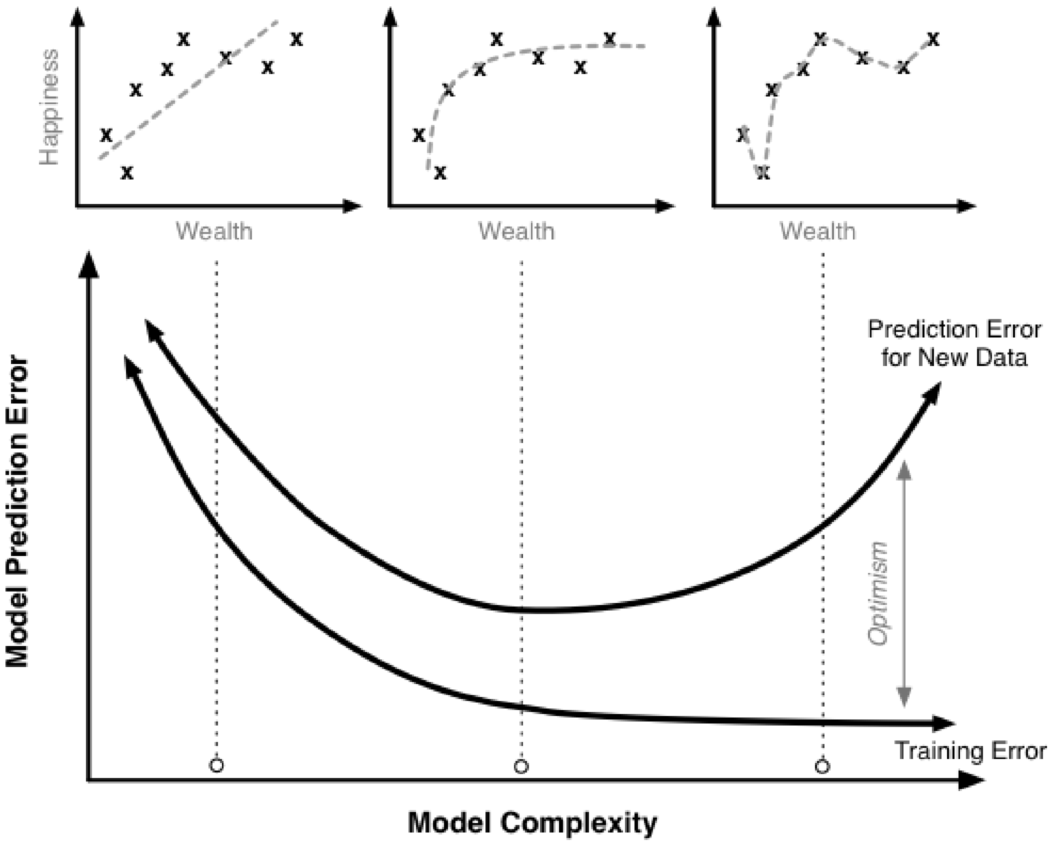
$$\sum_j b_j y_j = \sum_i a_i x_i + \epsilon$$

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

APCC
APEC CLIMATE CENTER

Weakness : overfitting

- **Consider potential predictability**



from Scott Fortmann-Roe 2012

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

APCC
APEC CLIMATE CENTER

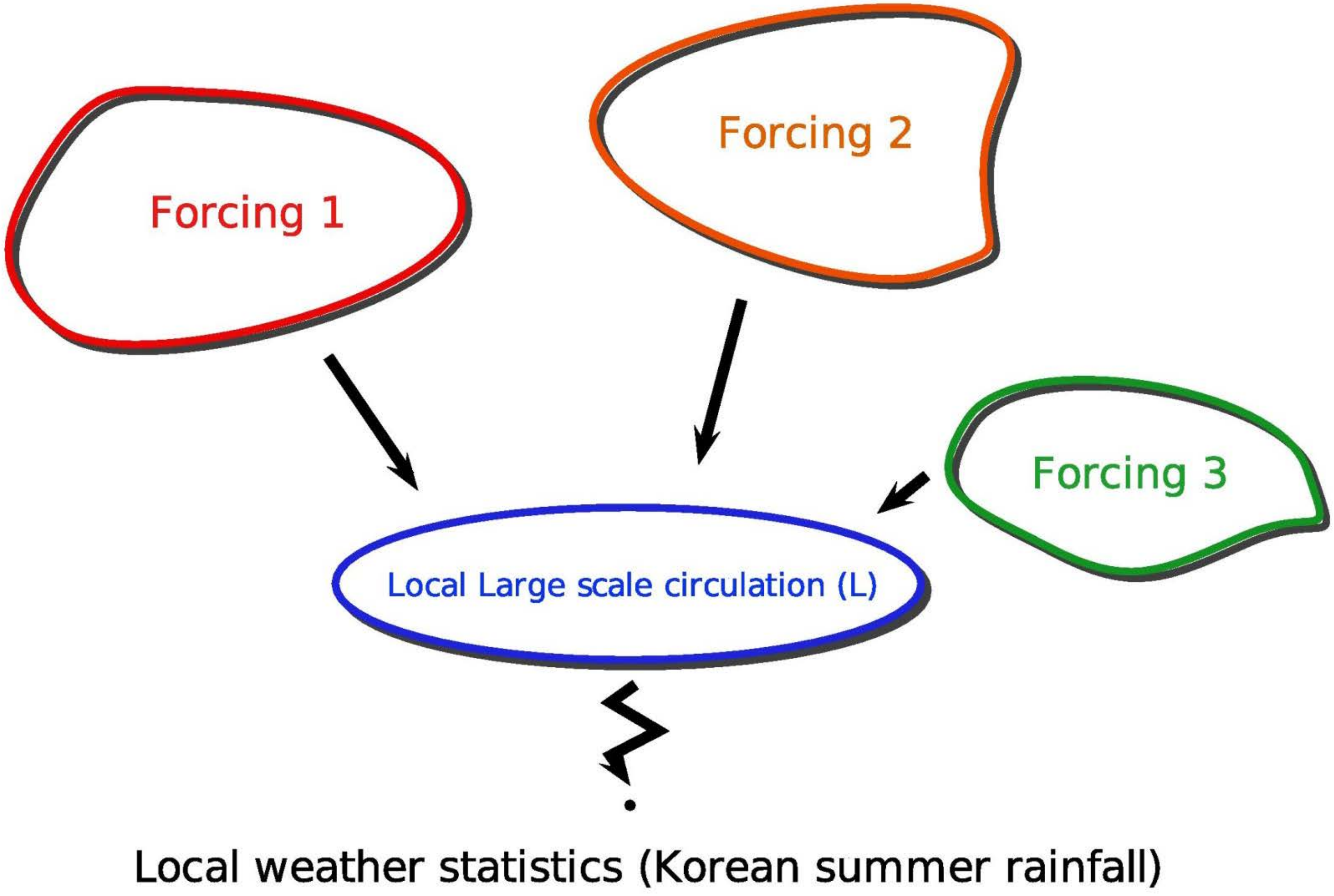
Downscaling (post-process) should be based on

- 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

Local large scale circulation



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

APCC
APEC CLIMATE CENTER

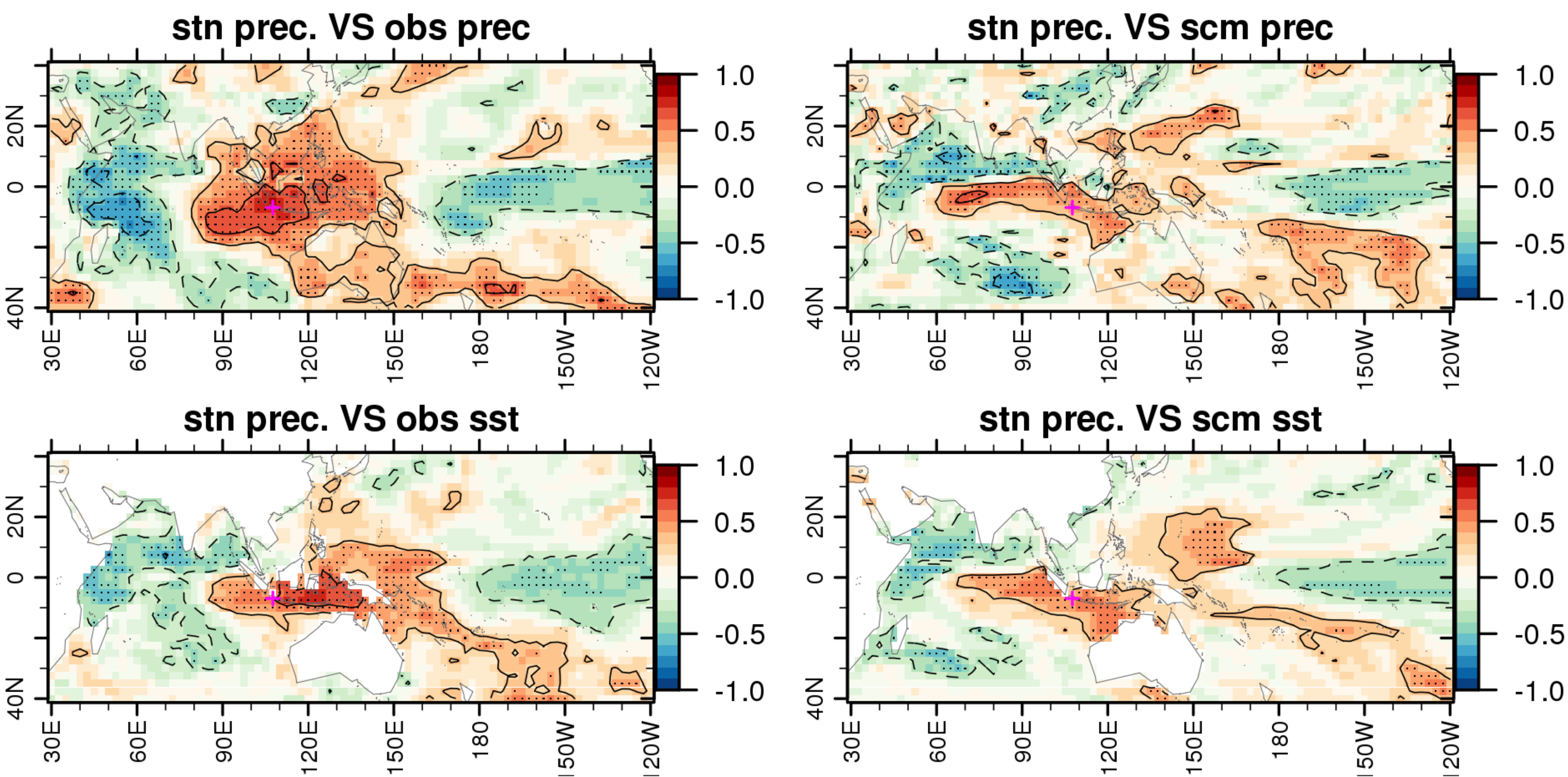
LARGE SCALE PATTERN ASSOCIATED WITH RAINFALL

Local large circulation and Teleconnection

One Point Correlation map with seasonal mean local rainfall(APRODITE or CRU data) with other variables

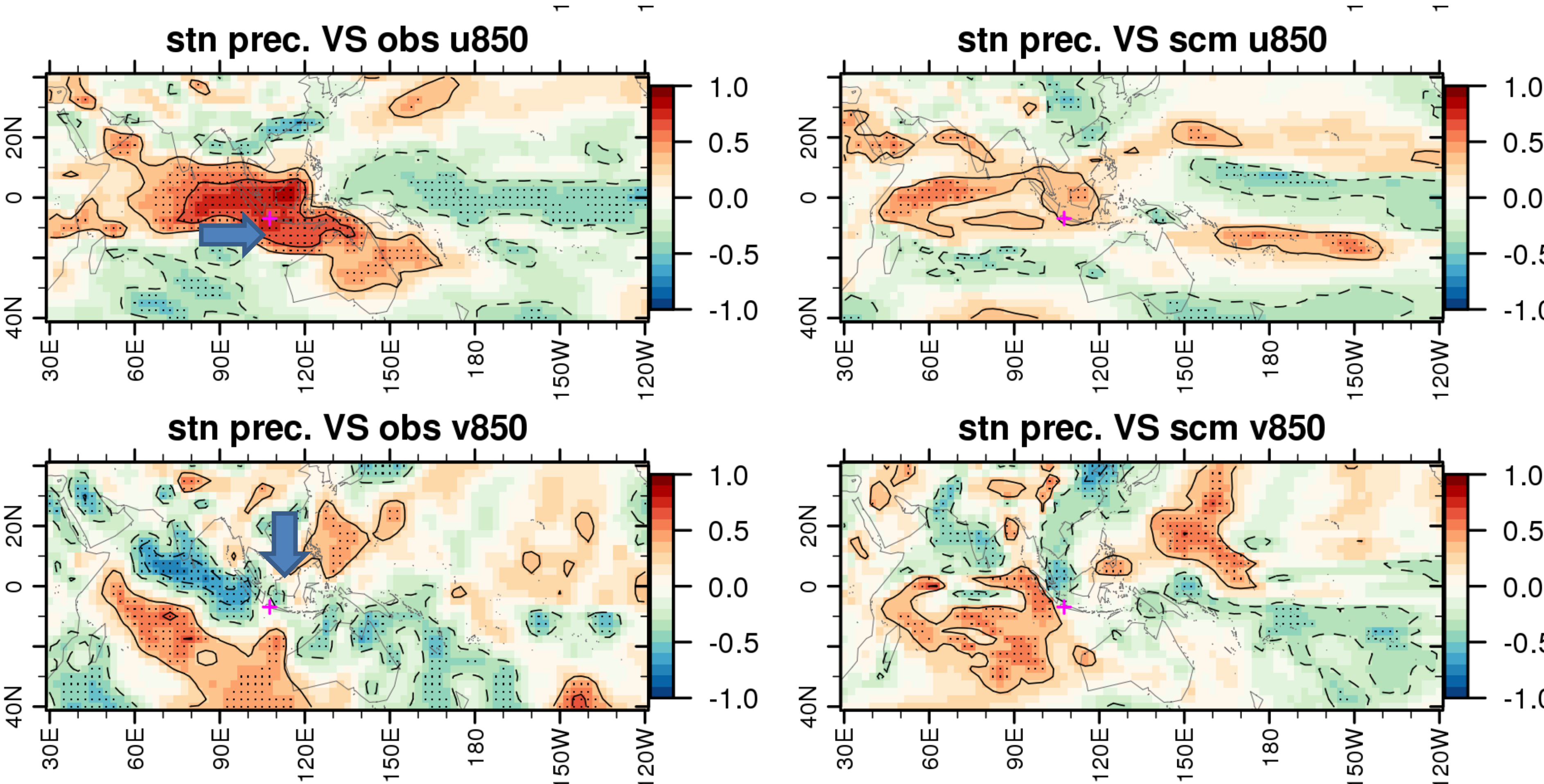
APCC
APEC CLIMATE CENTER

SON [Jakarta]



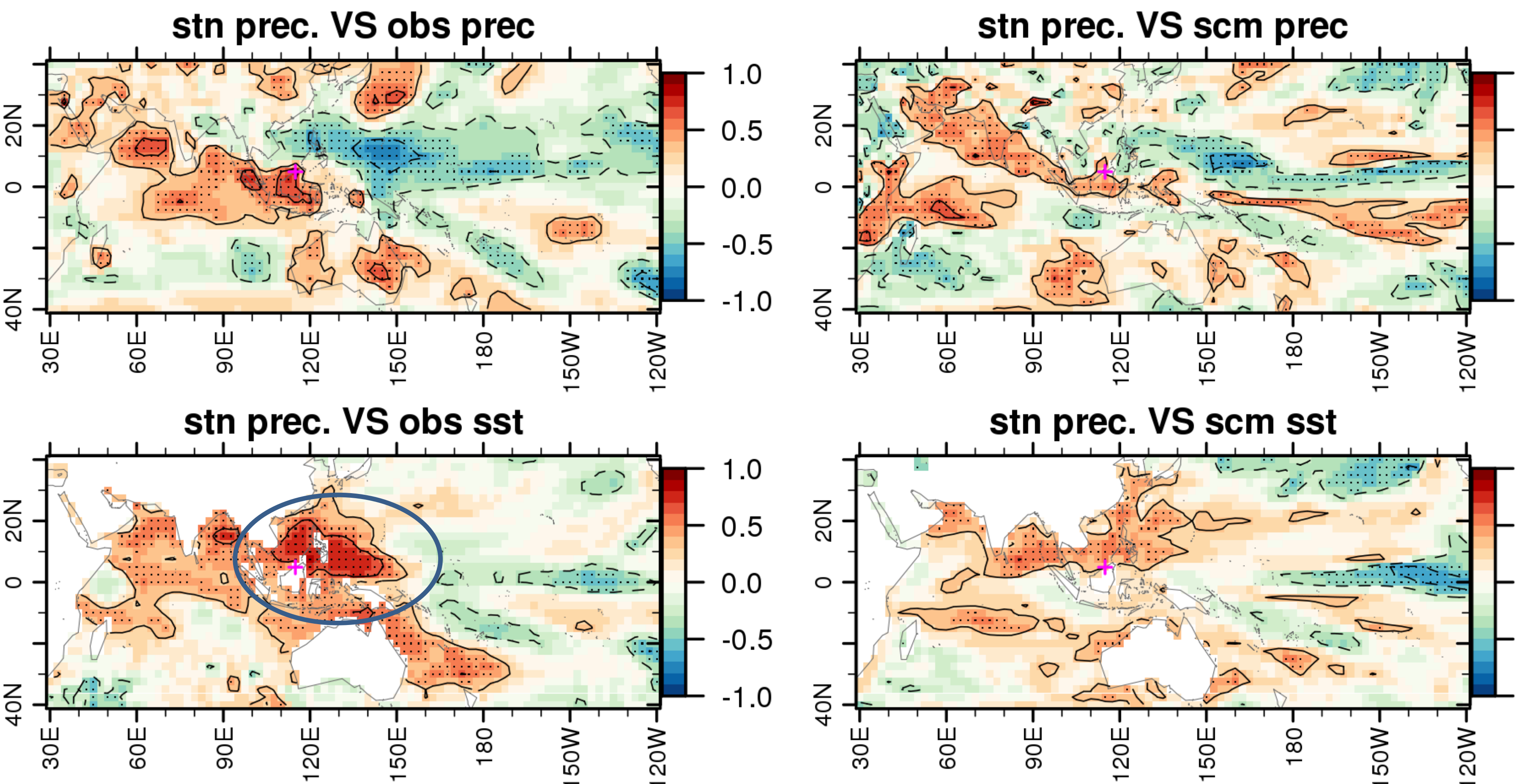
La Nina (and IOD) signature

SON(Jakarta)



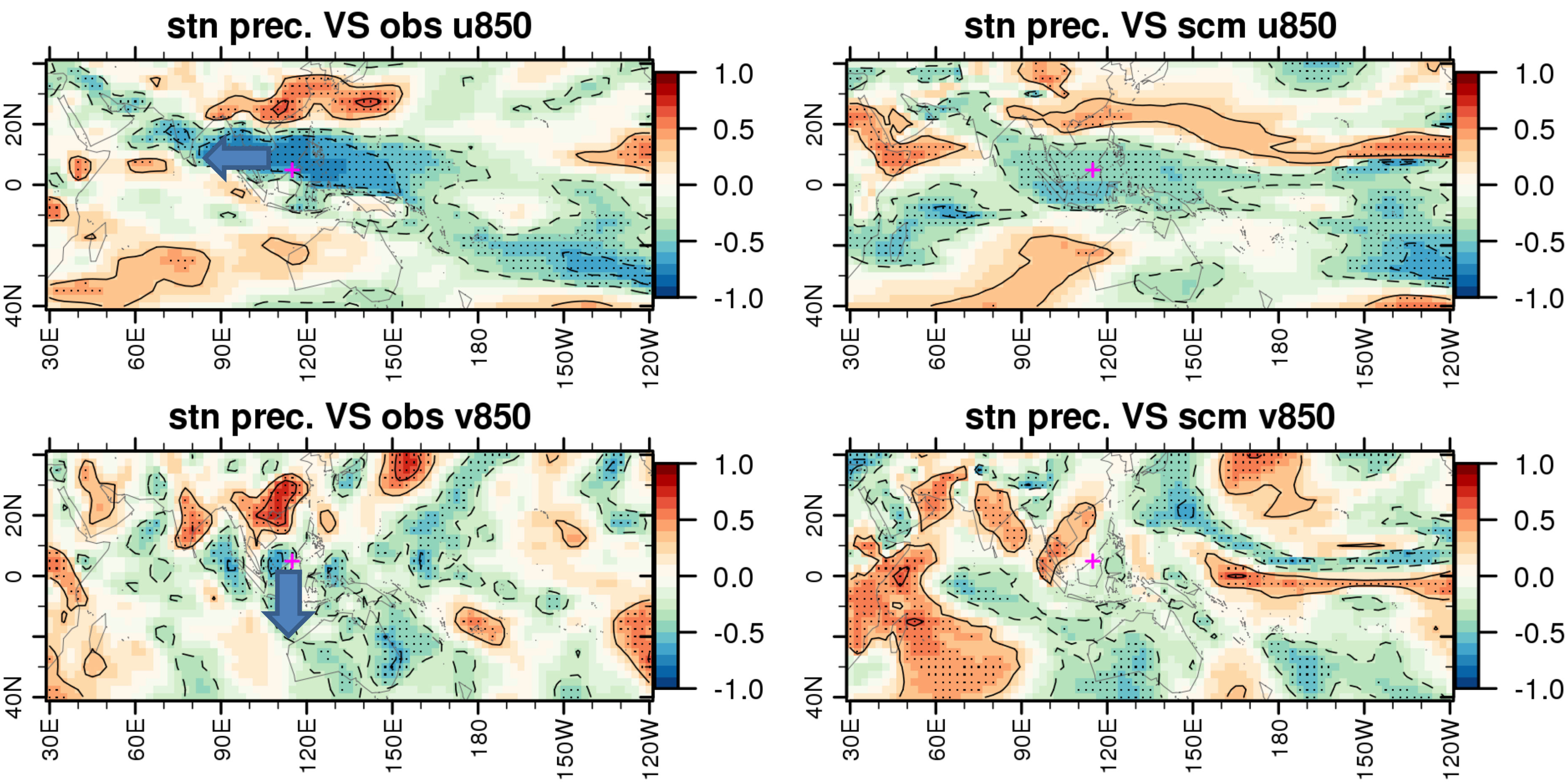
stn prec. VS obs u850
stn prec. VS scm u850
stn prec. VS obs v850
stn prec. VS scm v850
20N
0
40N
30E
60E
90E
120E
150E
180
150W
120W
1.0
0.5
0.0
-0.5
-1.0

JJA [Bandar_Seri_Begawan]

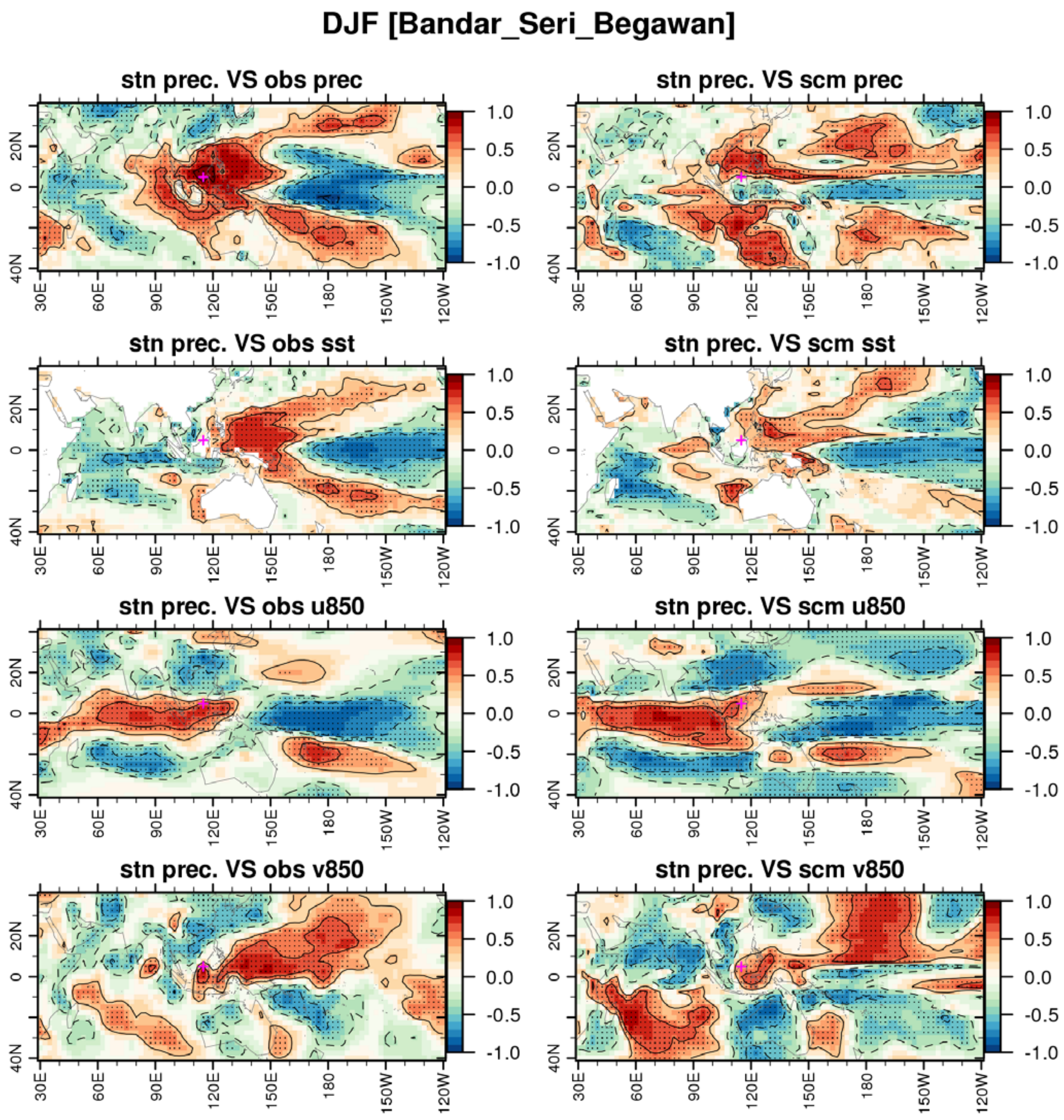


APCC
APEC CLIMATE CENTER

JJA(Bandar_Seri_Begawan)

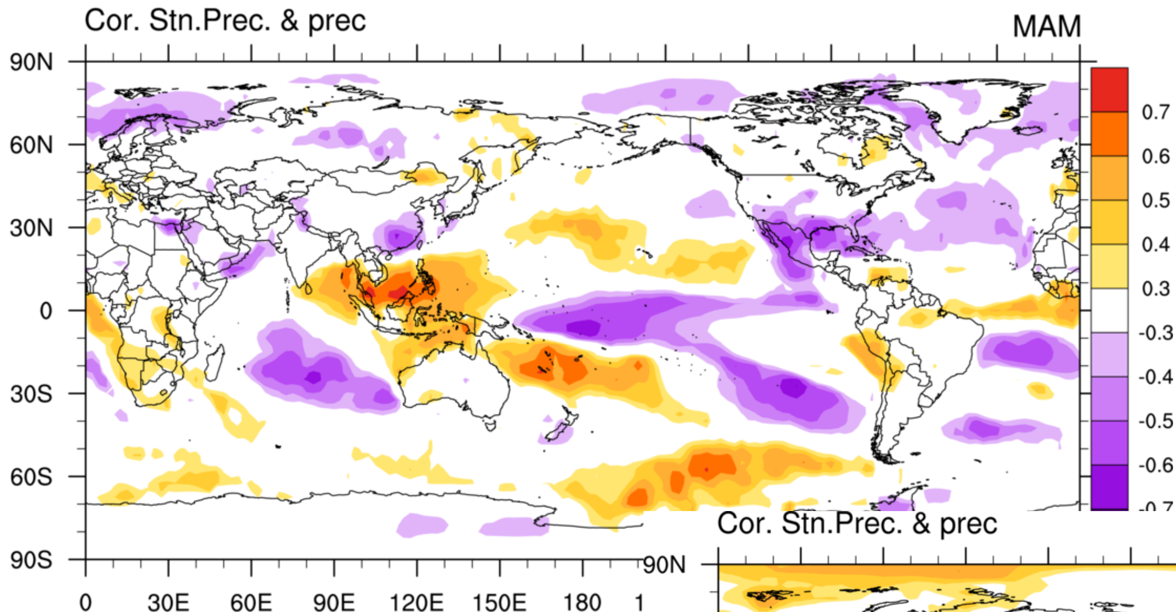


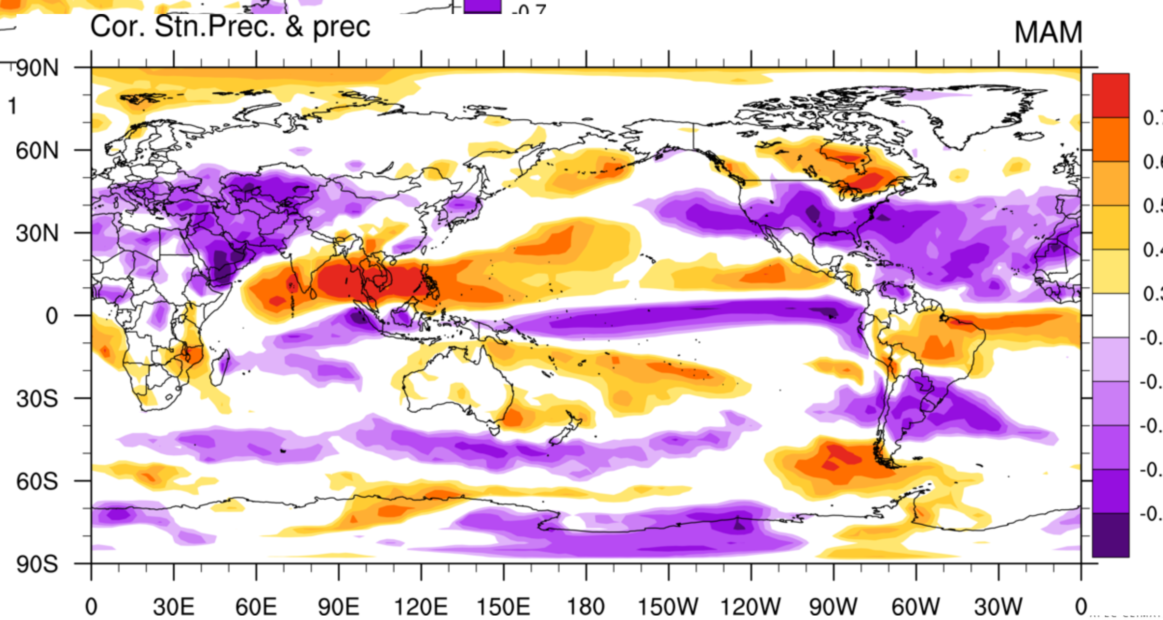
APCC
APEC CLIMATE CENTER


DJF [Bandar_Seri_Begawan]
stn prec. VS obs prec
stn prec. VS scm prec
stn prec. VS obs sst
stn prec. VS scm sst
stn prec. VS obs u850
stn prec. VS scm u850
stn prec. VS obs v850
stn prec. VS scm v850
1.0
0.5
0.0
-0.5
-1.0
40N
0
20N
30E
60E
90E
120E
150E
180
150W
120W

La Nina!

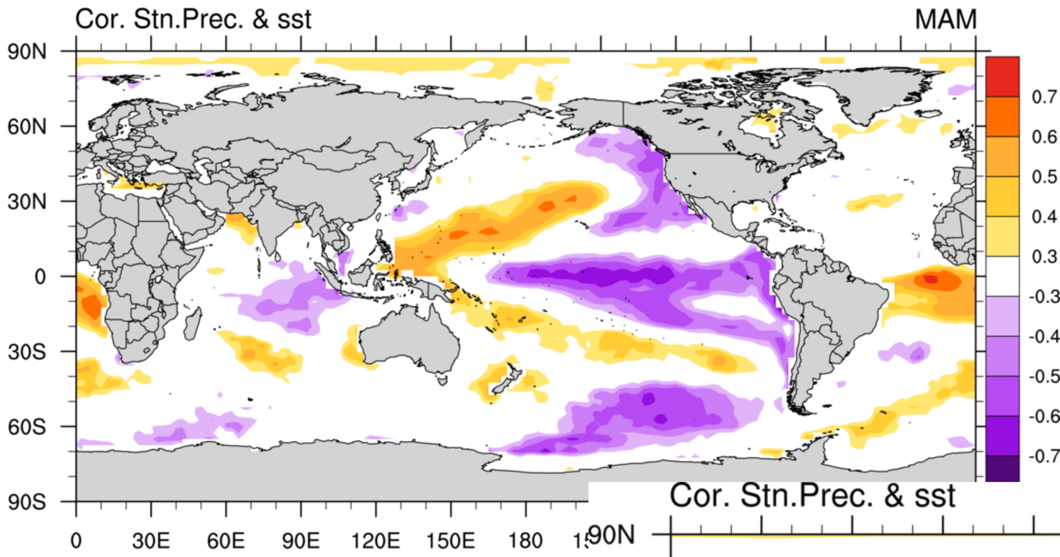
Bangkok (MAM)

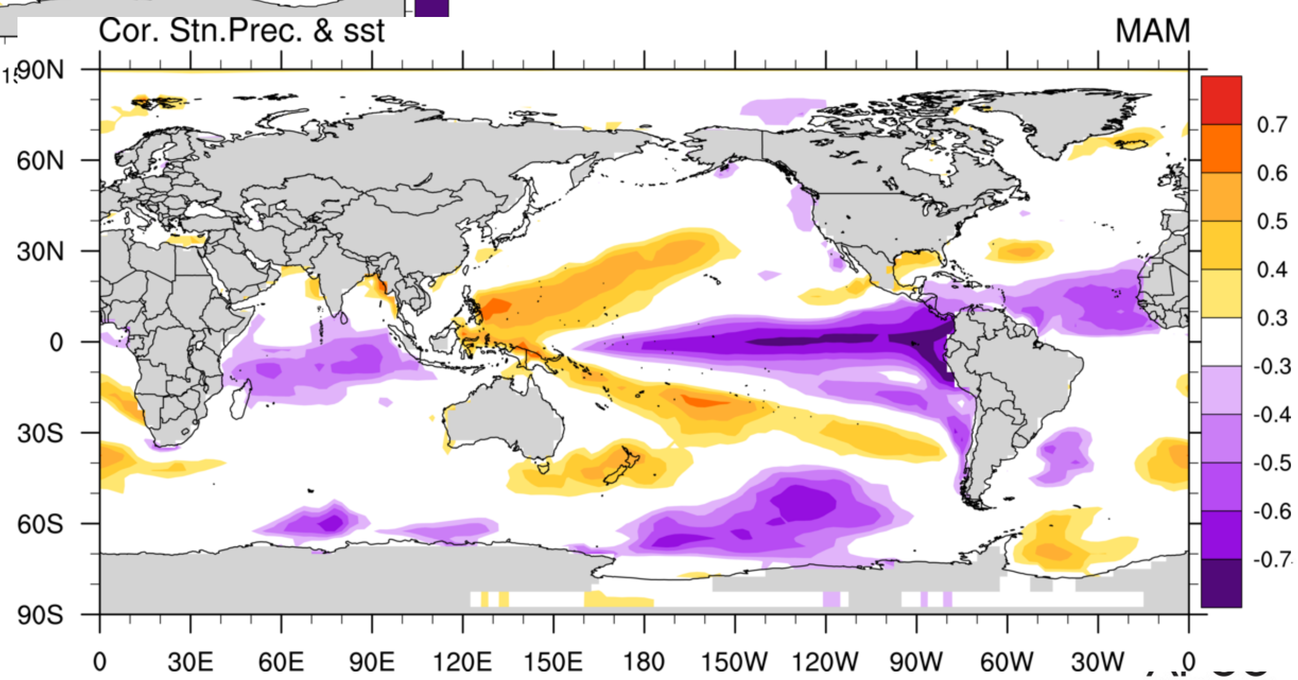




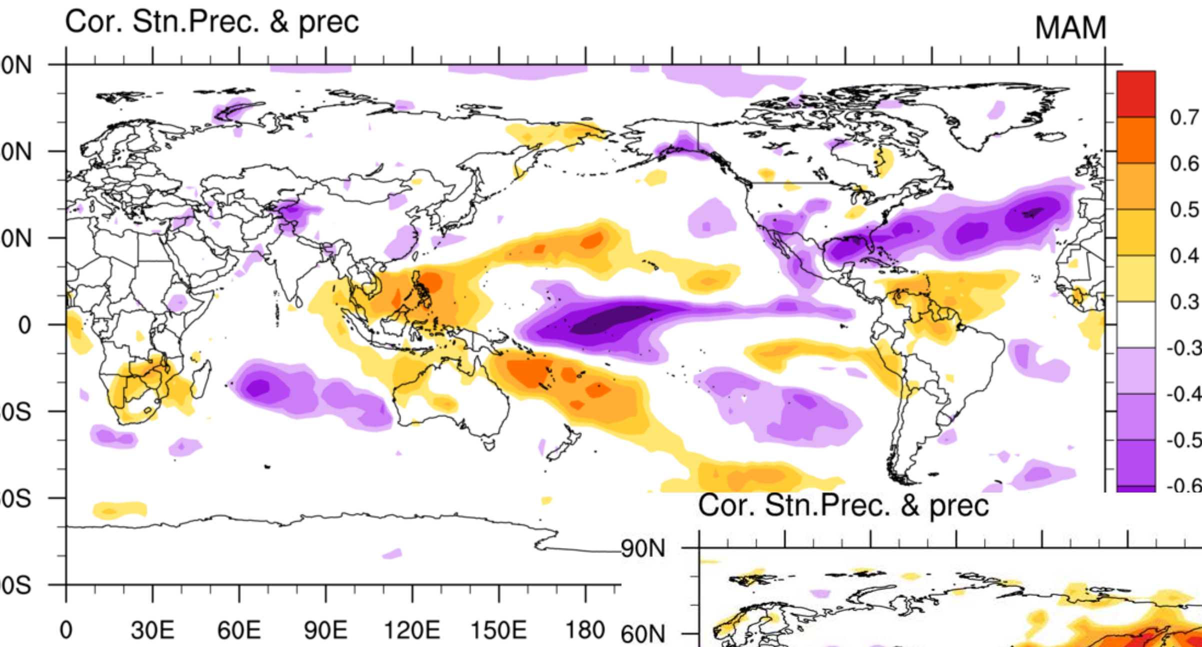
Model

Bangkok (MAM)

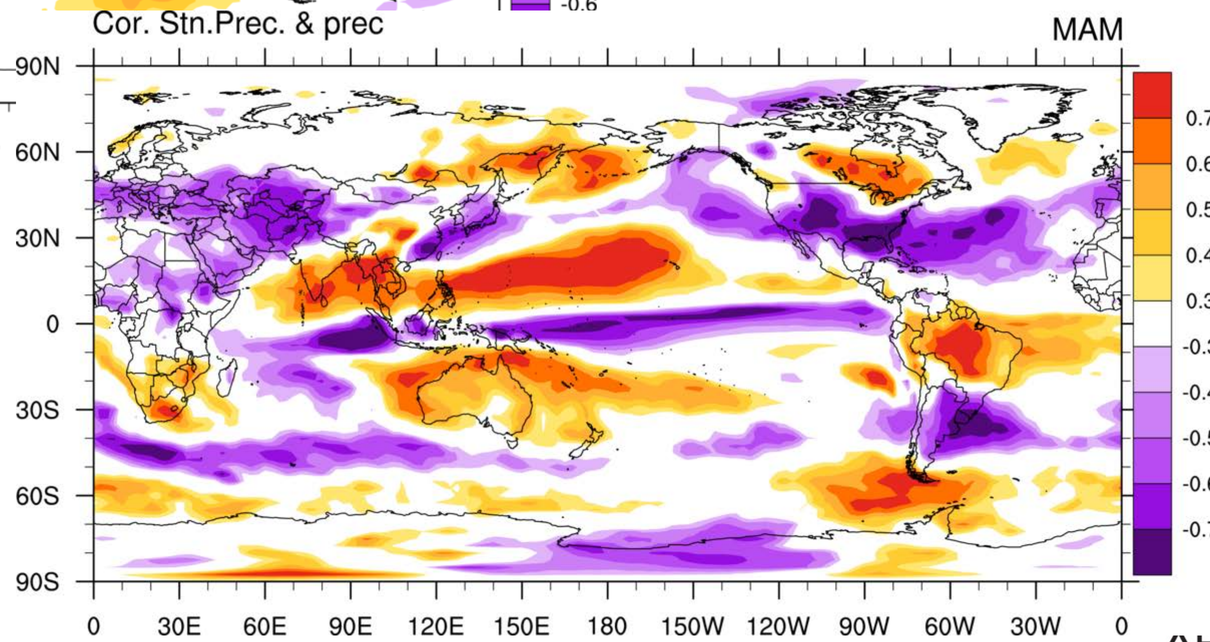

Cor. Stn.Prec. & sst
MAM
90N
60N
30N
0
30S
60S
90S
0
30E
60E
90E
120E
150E
180
0.7
0.6
0.5
0.4
0.3
-0.3
-0.4
-0.5
-0.6
-0.7


Cor. Stn.Prec. & sst
MAM
90N
60N
30N
0
30S
60S
90S
0
30E
60E
90E
120E
150E
180
150W
120W
90W
60W
30W
0
0.7
0.6
0.5
0.4
0.3
-0.3
-0.4
-0.5
-0.6
APEC CLIMATE CENTER

Hochiminh

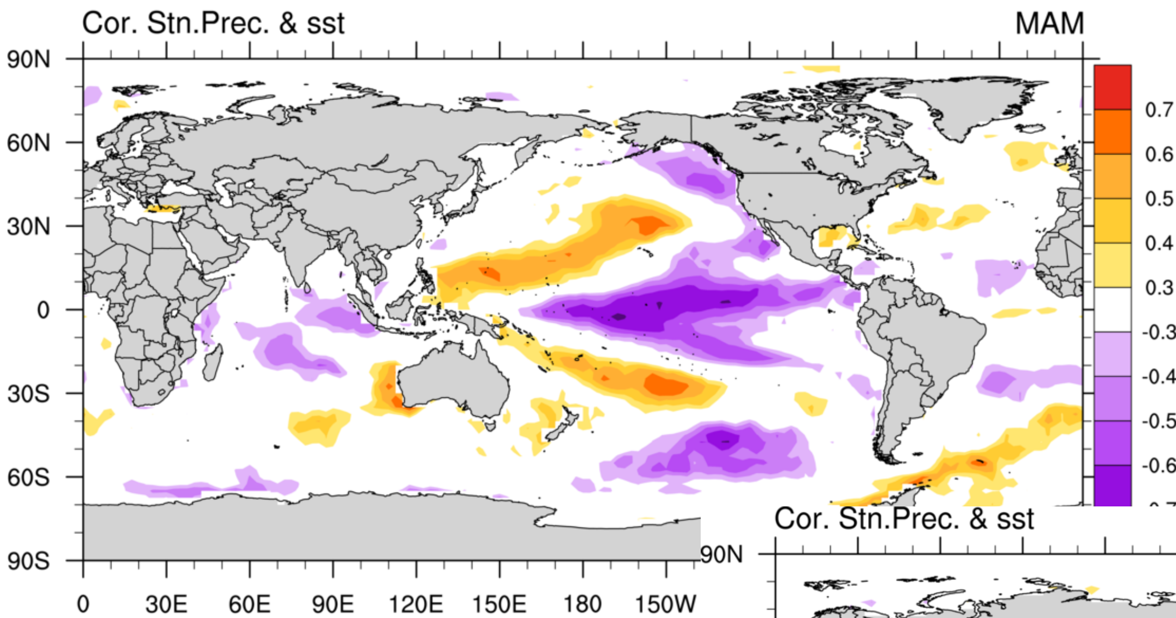


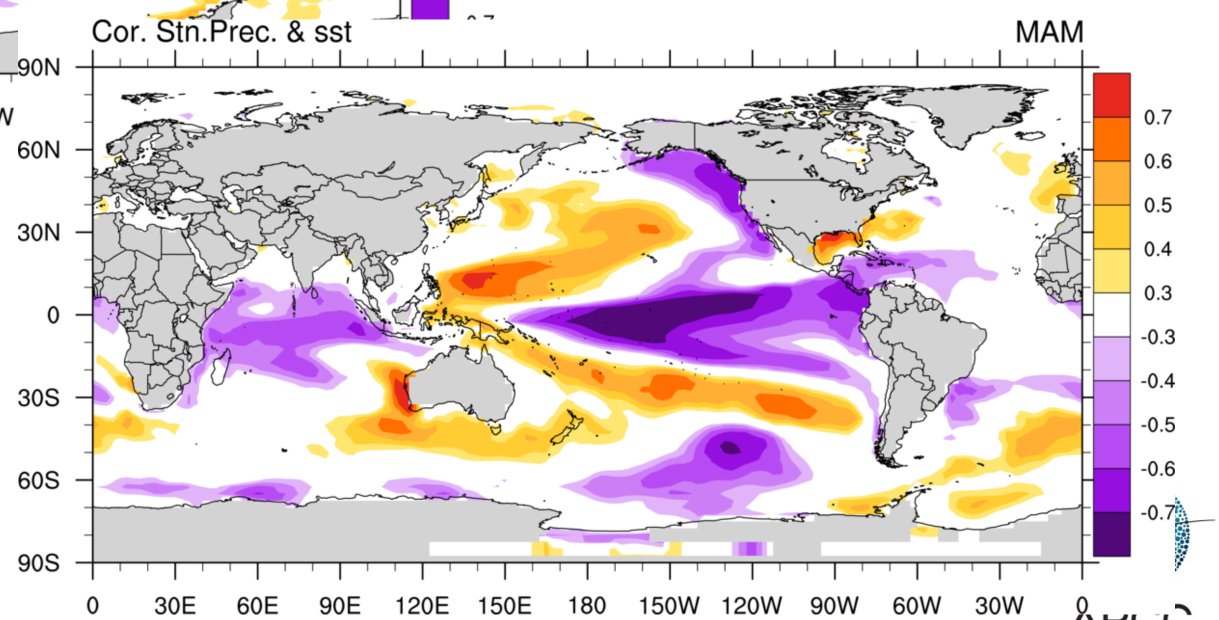
Model



APCC
APEC CLIMATE CENTER

Hochiminh

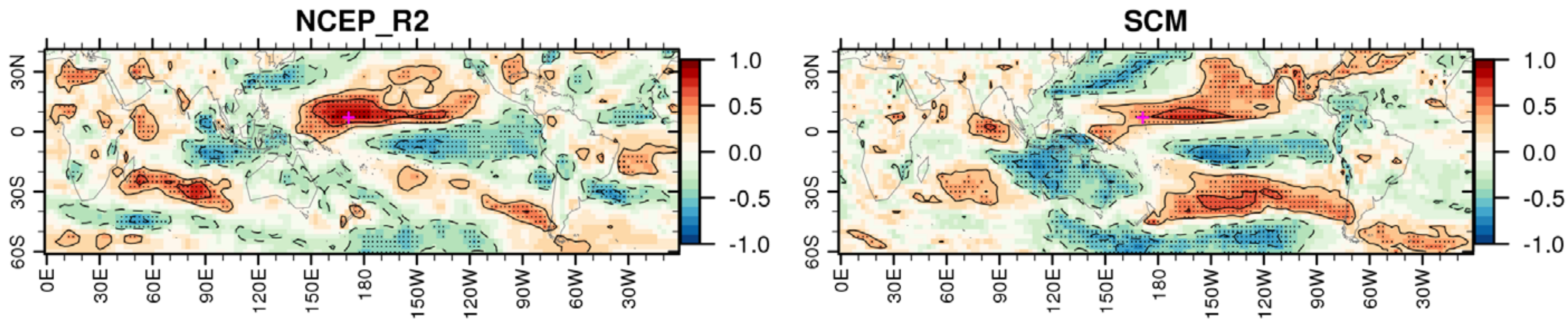




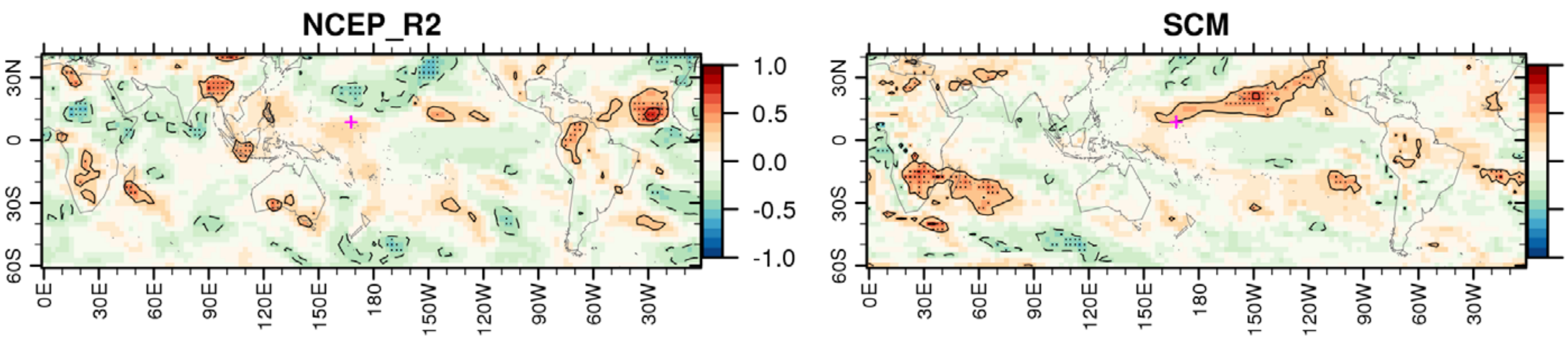
APCC
APEC CLIMATE CENTER

Majuro

prec, AMJ [Marshall_Islands, 91376]



prec, OND [Marshall_Islands, 91366]



Note that,

- YOU should have an “Guess field” that is associated with your seasonal mean climate variability (“positive SST over certain region causes more rainfall at our station”)

- Model should be able to mimic that physical relationship even with some error

- CLIK will work if you can find a predictor satisfying above two thing

APCC
APEC CLIMATE CENTER

Downscaling in CLIK

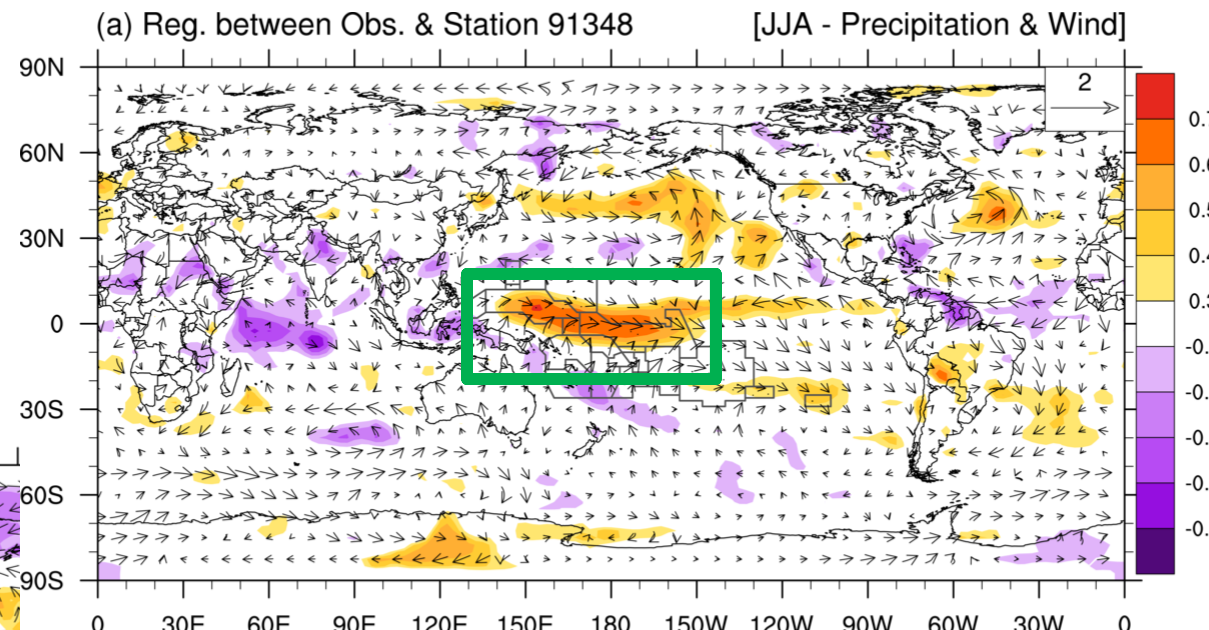
- Use "observed" large scale pattern (X) associated with climate variability at stations
- X needs to be predicted by GCMs to some degree
  - X becomes predictor (user selected area)

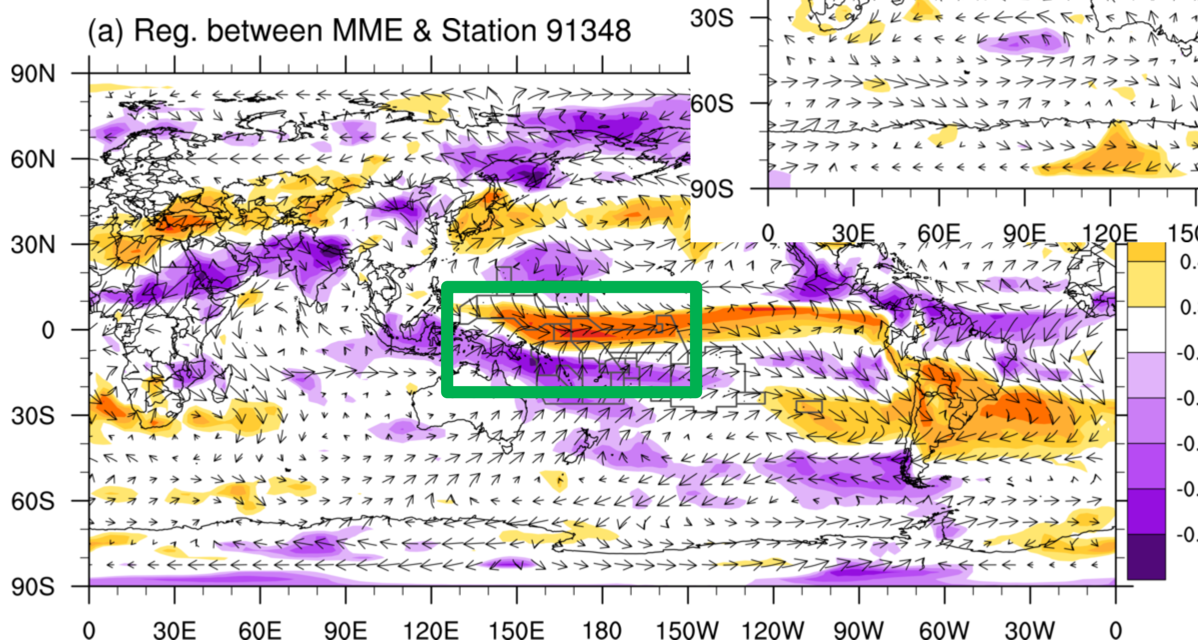
- CLIK does not provide any prior information for selection of predictor (to avoid overfitting)
  - Basic knowledge on Local large scale circulation and associated global teleconnection is necessary

Predictor selection

Meaningful pattern? (hopeful)
: significance score

Station data





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

APCC
APEC CLIMATE CENTER

The most important thing you need is,

Patience

APCC
APEC CLIMATE CENTER