CLIK Training (16-17 Oct 2018, Hanoi, Vietnam)

Climate Prediction 101

Yun-Young Lee

APCC APEC CLIMATE CENTER

CLIMATE?

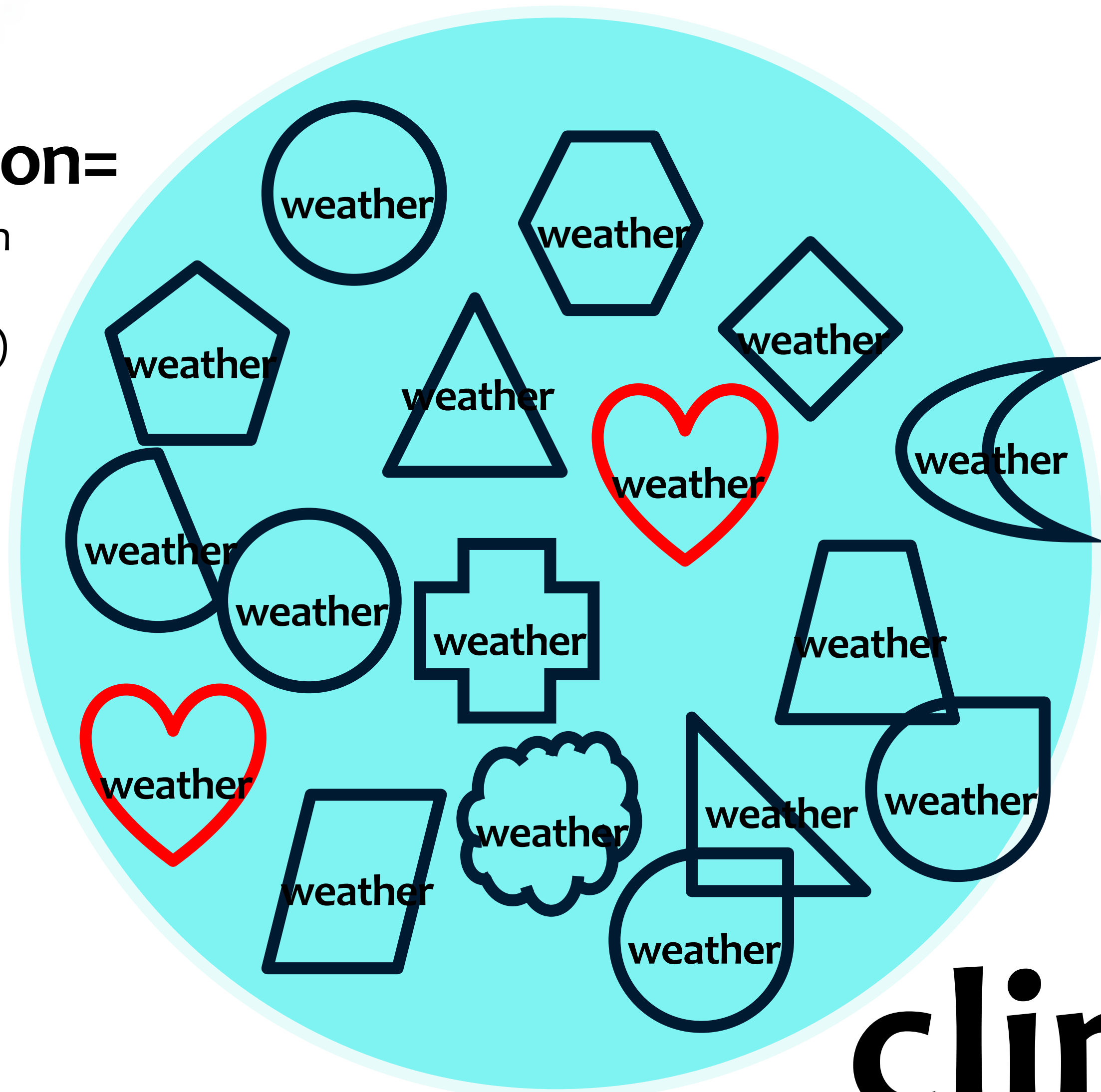
Weather is what we ***get,***

Climate is what we ***expect!***

Weather summary = Climate

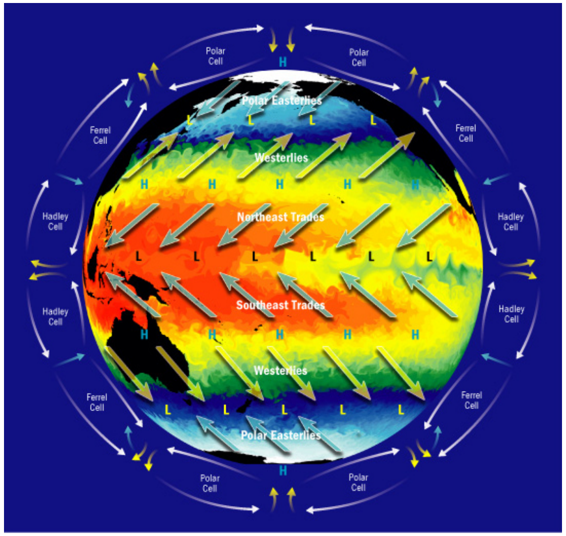
Expectation=

mean condition
of atmosphere
(temp. & Prcp.)



climate

Climate = Expectation



Climate **=** **Expectation**

Change ***Should be changed!***

Climate = Expectation

Prediction of Expectation

How uncertain!

Prediction

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

(in meteorology)

- What : atmospheric state (weather)

- Conditions:
  current state, physical rules, external forcing factors

Atmosphere is dynamical system!

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

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

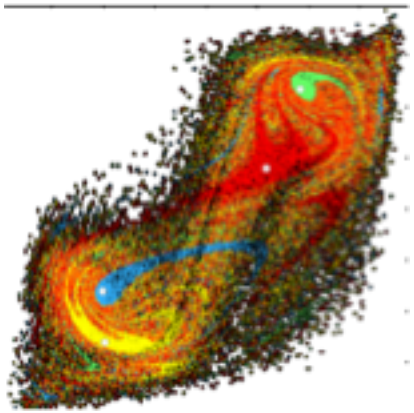
~~Determinism~~

~~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…***

Predictability

Depends on ***what to predict***

Lead time(τ)

- ✓ *Temperature of this room tomorrow*
- ✓ *Temperature of this room in 30days later*
- ✓ *Temperature of this room in 30years later*

Location

- ✓ *Temperature of Busan (Korea)*
- ✓ *Temperature of Ho Chi Minh (Vietnam)*
- ✓ *Temperature of Villa Las Estrellas (Antarctica)*

Predictability

- ✓ *Temperature*
- ✓ *rainfall*
- ✓ *wind speed*

Physical variables

- ✓ *Mean Temperature during a day*
- ✓ *Mean Temperature during a month*
- ✓ *Mean Temperature during a century*

Time-scale of predictand

Signal & Noise

Two scales

- Fast and small scale processes : **noise**
  - Weather by tropical cyclone etc.
- Slow and large processes : **signal**
  - Climate by **ENSO**, ITCZ, monsoon, **MJO** and so forth

- What is the Signal? (How we can "see"?)
  - Tendency of weather that has to be physically caused by slow varying processes

Potential predictability

Matter of
Signal
&
Noise

$$X = Xs + Xn$$

Measured by relative magnitude (variance) of signal and noise

Signal >> Noise : more predictable

Signal << Noise : less predictable

How long
(time domain)?

Time mean of
weather

Subseasonal

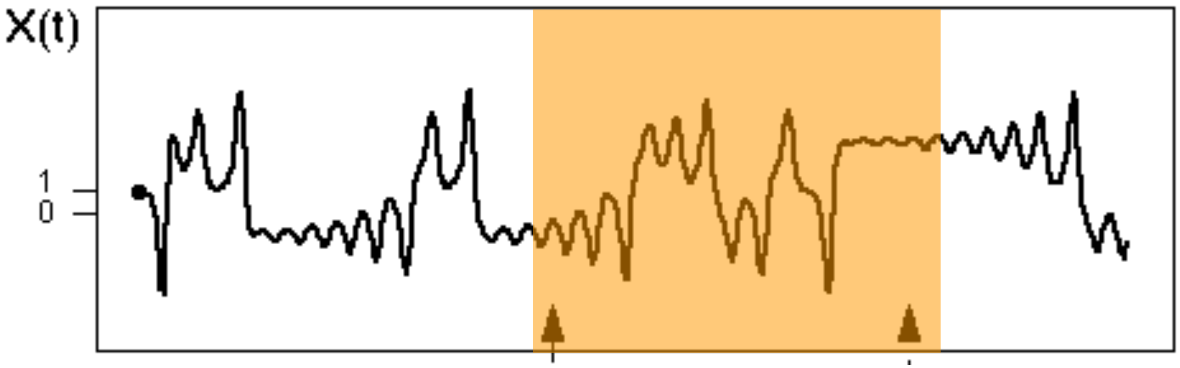
Seasonal

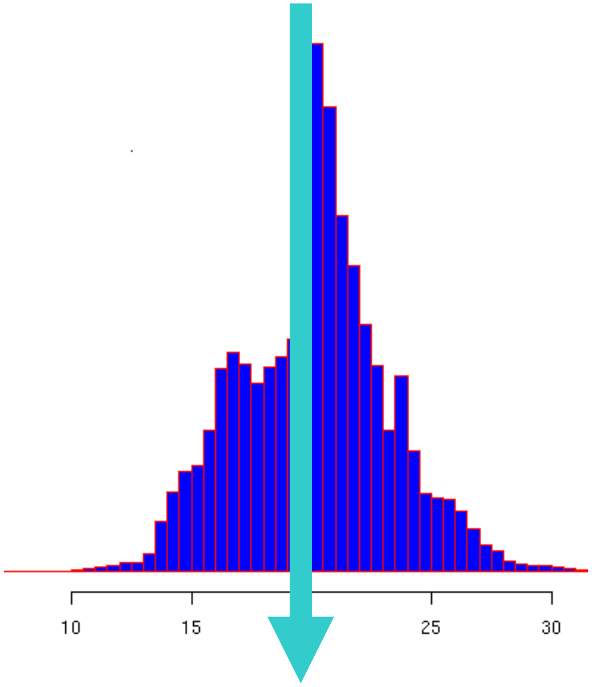
Climate

prediction

forecast

Weather statistics



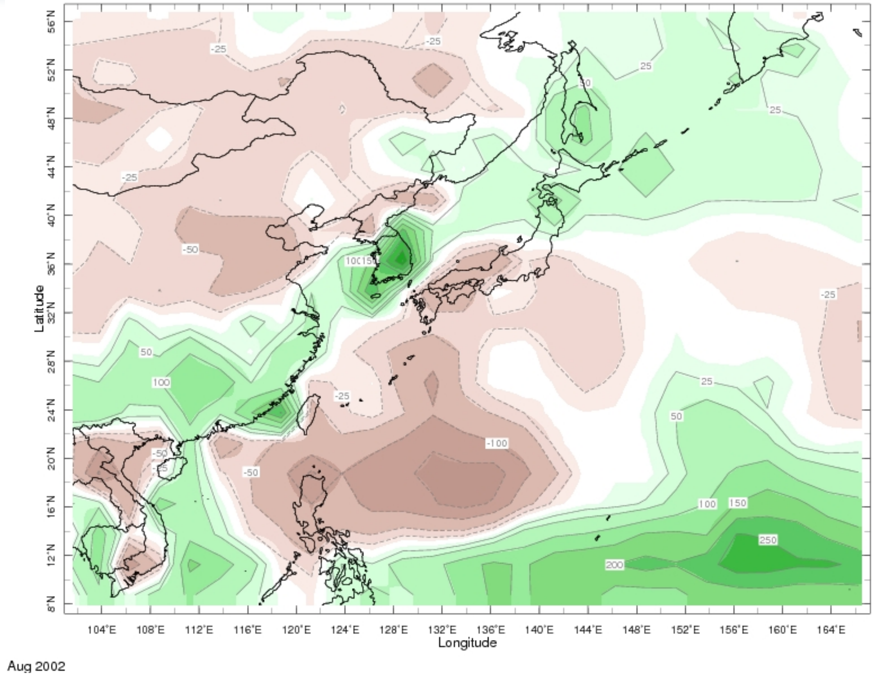


Primary seasonal weather statistics :
seasonal **mean**

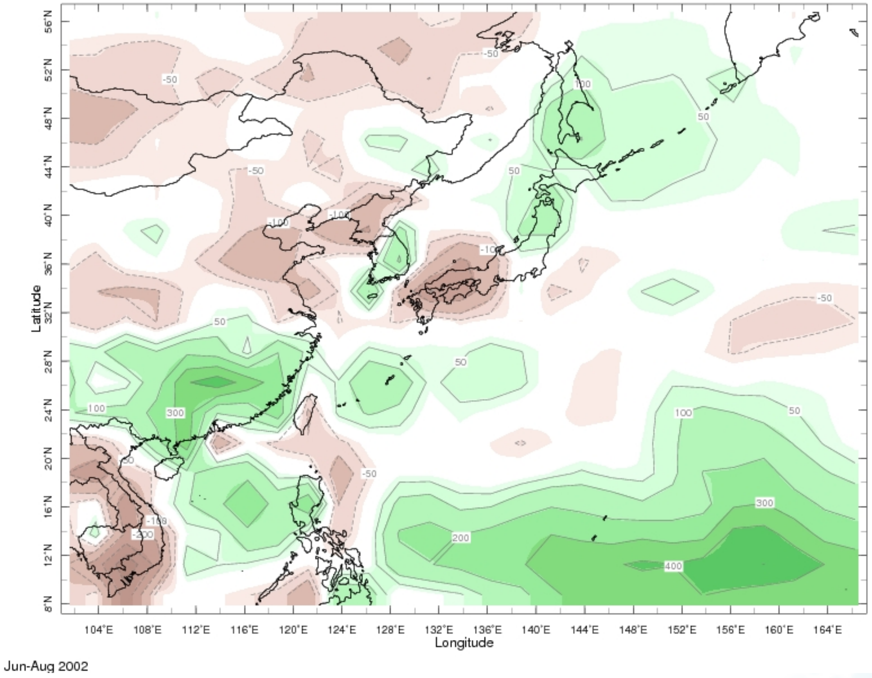
Seasonal
mean

2002 summer (JJA) rainfall

Monthly prec. Anomaly (Aug)



Summer mean prec. anomaly



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

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

Pros/Cons

Statistical/Empirical

- Simple and cheap
- Based on (observation) data
- Short observing history. Do we have enough?
- Unprecedented events predictable ????

Dynamical

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

Statistical forecasting

(0) Climatology

- Baseline of seasonal forecasting
- "Nothing particular, Sir."

$$x(t+1) = x$$

✓ Rainfall amount will be similar to 30 year average.

✓ I don't know? (33%:33%:33%) or Near Normal?

(1) Persistence

- Assume that future will be same as it is now.

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

✓ Often Close to people's expectation.

✓ Effective when the autocorrelation is large (e.g. ENSO forecast)

(2) Regression

- The most popular method and many variations

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

x : predictand (e.g. rainfall at a station)

y : predictor (e.g. NINO3.4 SST)

Regression based forecast

Question #1 : Predictor selection

- How to define **predictor (y)**?
- By definition, predictor should cause some changes in variation of predictand.

Question #2 : appropriate Function

- How to define **a** and **b**?
- *Your choice*: linear, nonlinear, single, multi… .complex ones are not necessarily better.

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

Predictand : my mood in the morning
Predictor? Relationship?

$x'(t+1) = ay(t) + b$
$x'(t+1) = a_1 y_1(t) + a_2 y_2(t) b$
$x_1'(t+1) + x_2'(t+1) = a_1 y_1(t) + a_2 y_2(t) b$

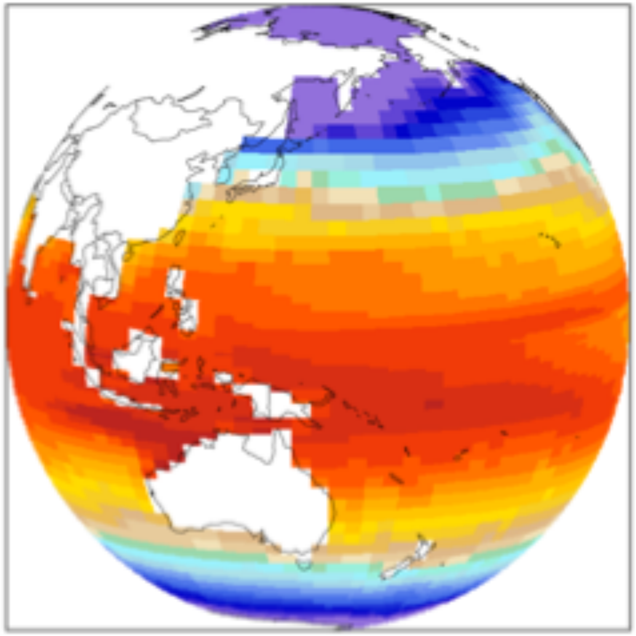
One to One : often not very satisfactory
One to Multi : easy to overfit (lie)
Multi to Multi : looks nice but often produce nothing practical

✓ If they give similar results, **the simpler is the better!**

✓ 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

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)

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

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

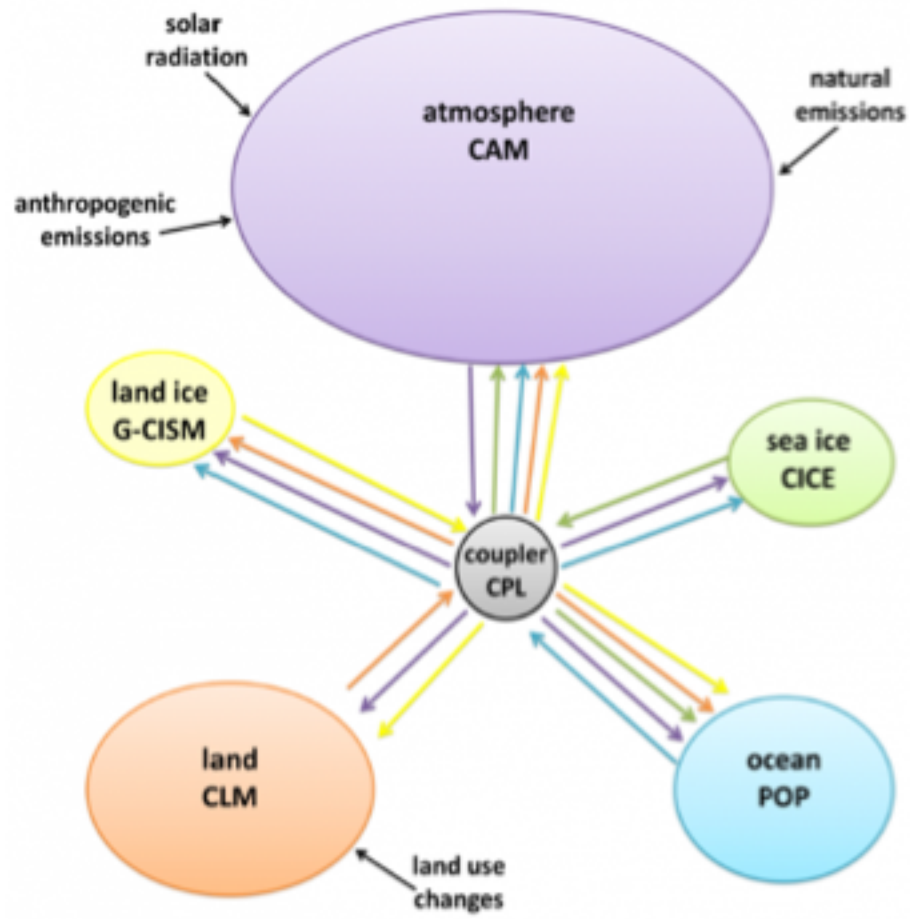
$$\frac{\partial \theta}{\partial t} + \vec{u} \cdot \vec{\nabla}\theta = l$$

$$\vdots$$

```
//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
```

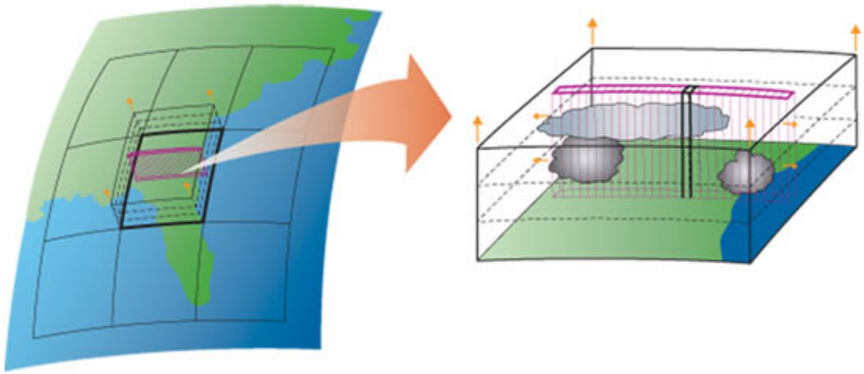
GCMs

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



Numerical modeling

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



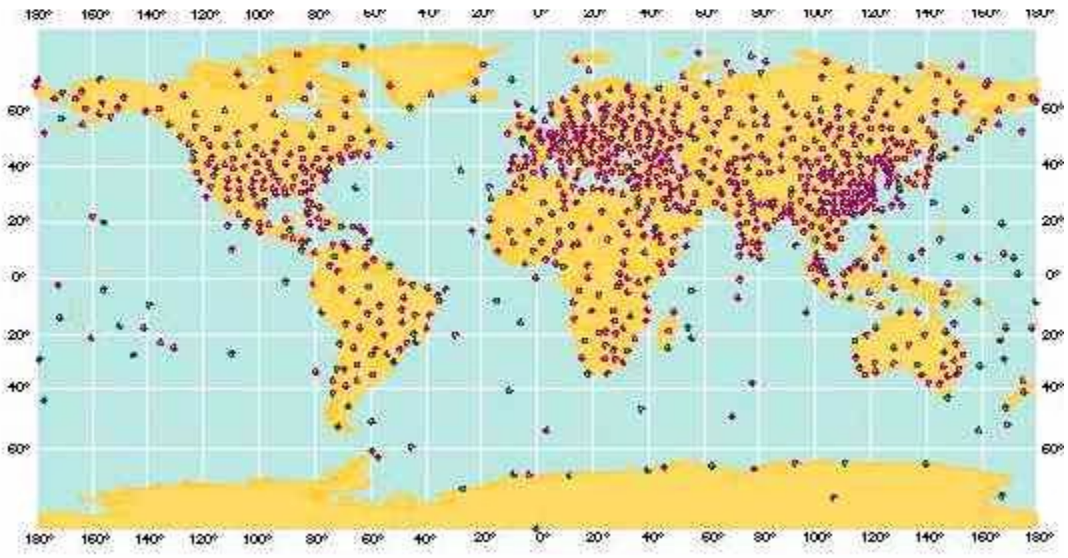
  - Unknown processes, tunable parameters
  - Initialization (for forecasting)

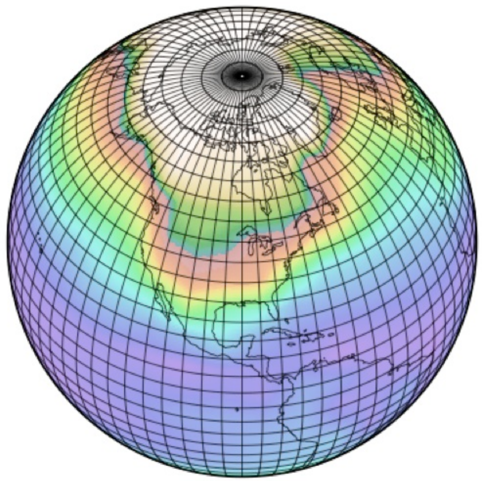
NOT the same with NATURE
NO perfect GCM

Initialization

Estimating Current status of climate system

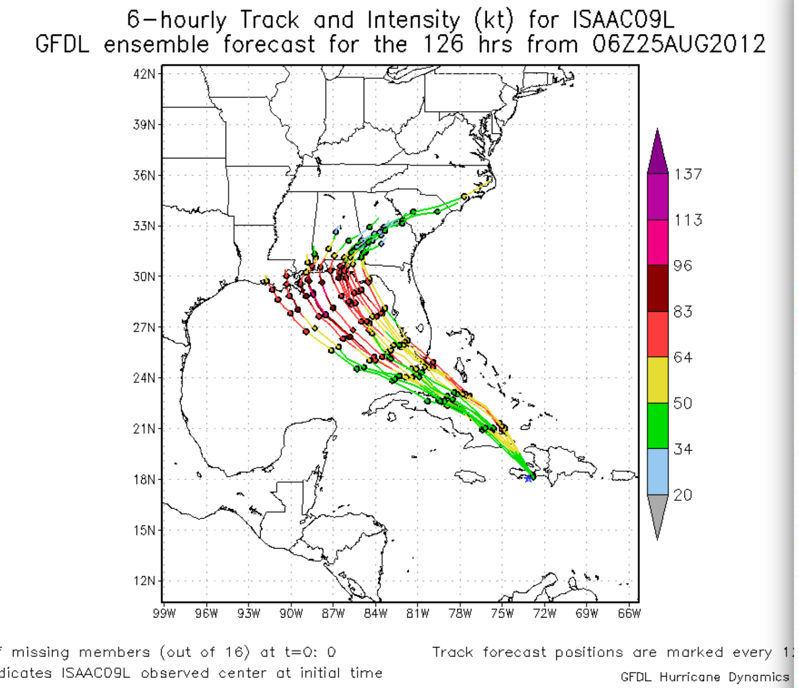
- Preparing the beginning climate state of GCM with available observation
  - Balance between Wrong GCM vs Wrong OBS.
  - Balance between components (Atm, Ocn)

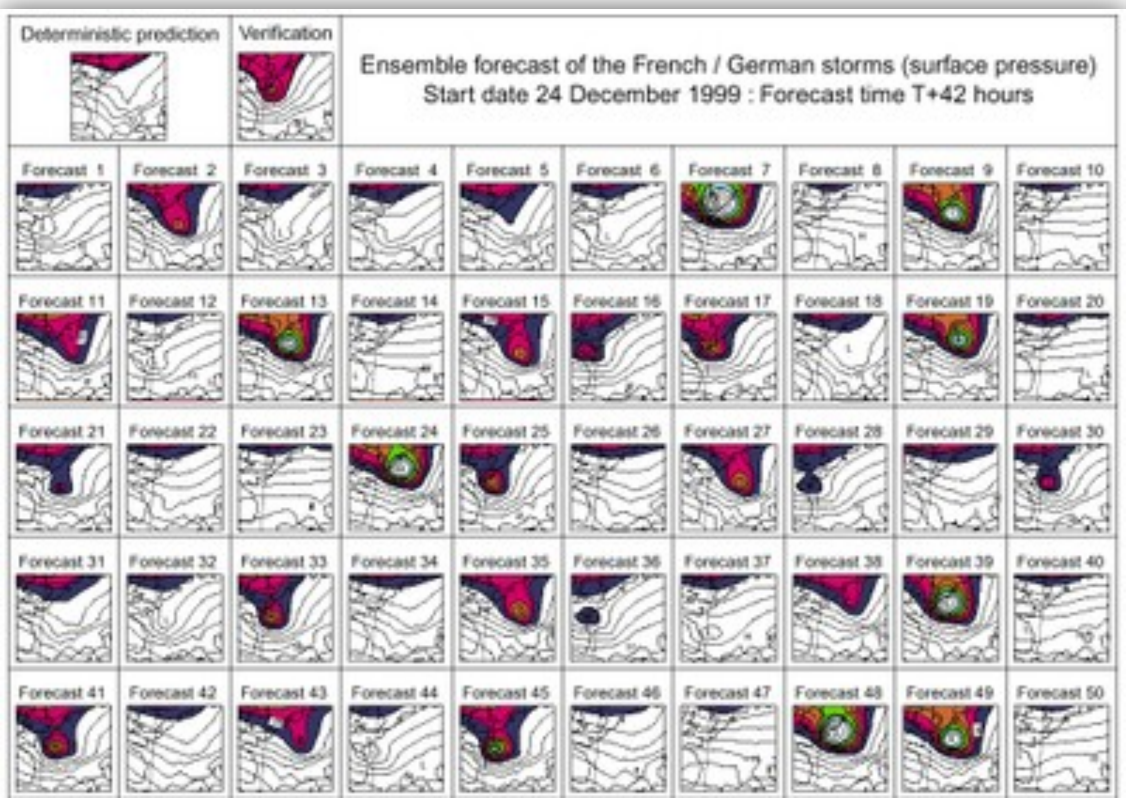




Ensemble Forecasting

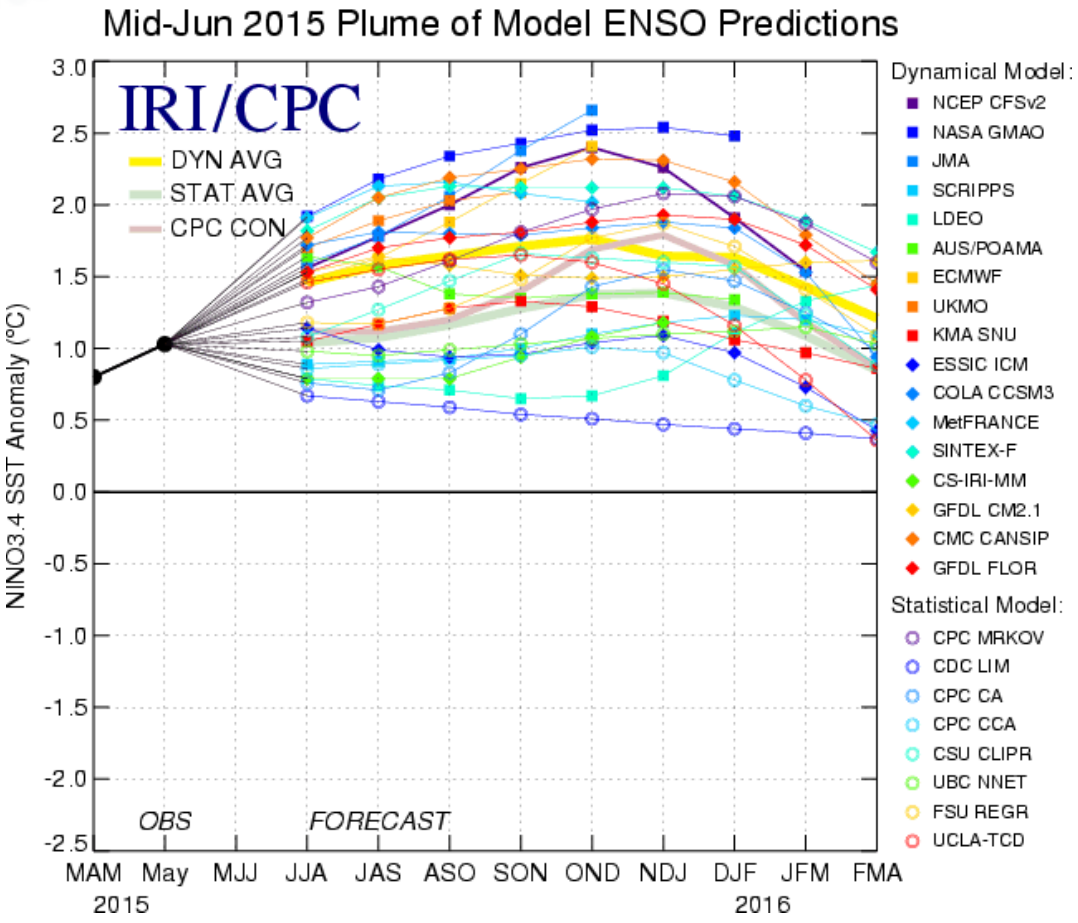
- Run many times
  - Starts from slightly different initial conditions





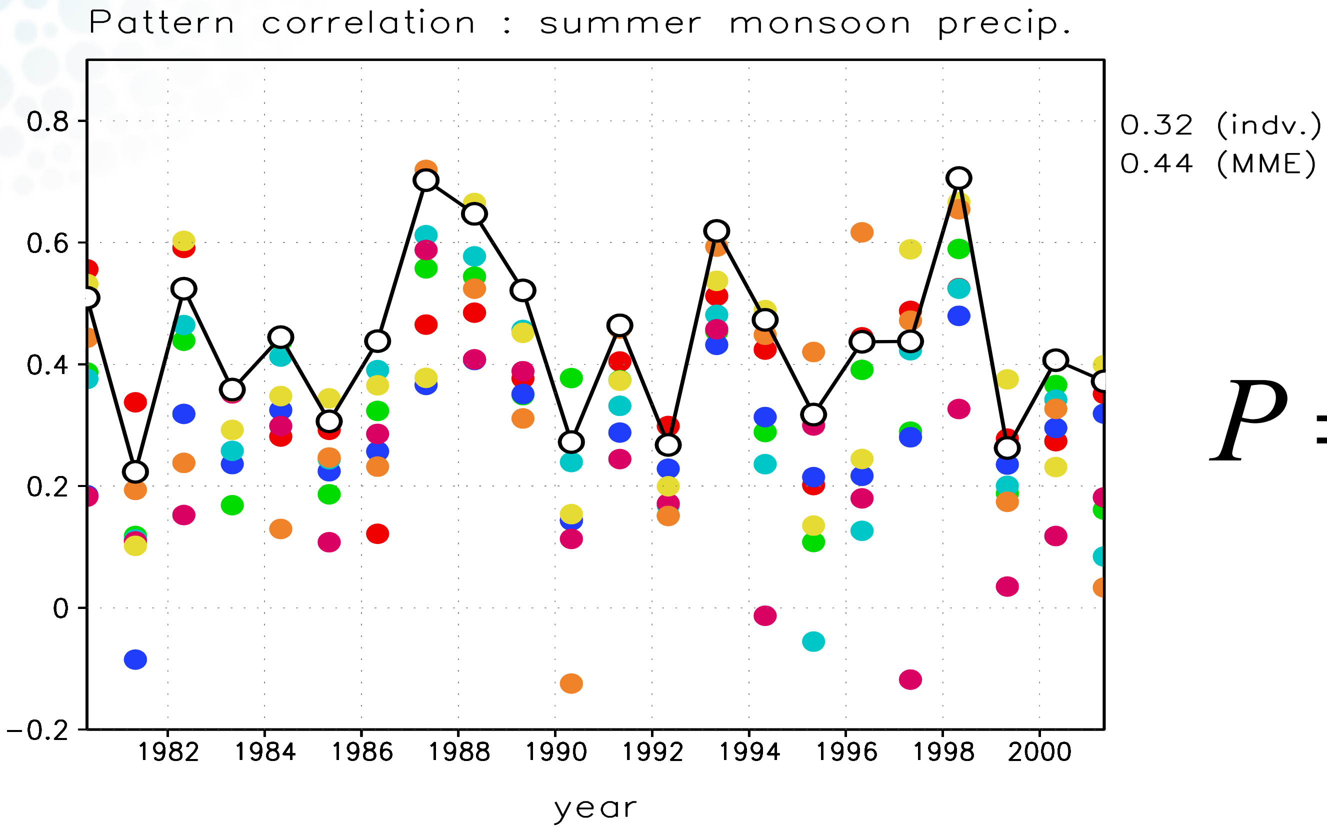
Multi Model Ensemble Forecasting

- Run with many models



Which one??

Use all!



$$P = \sum_i a_i F_i$$

A way to **reduce prediction uncertainties** generated when any single model is selected (to avoid extremely wrong forecast), in the end, to **improve overall prediction skill!**

APCC MME Climate Prediction

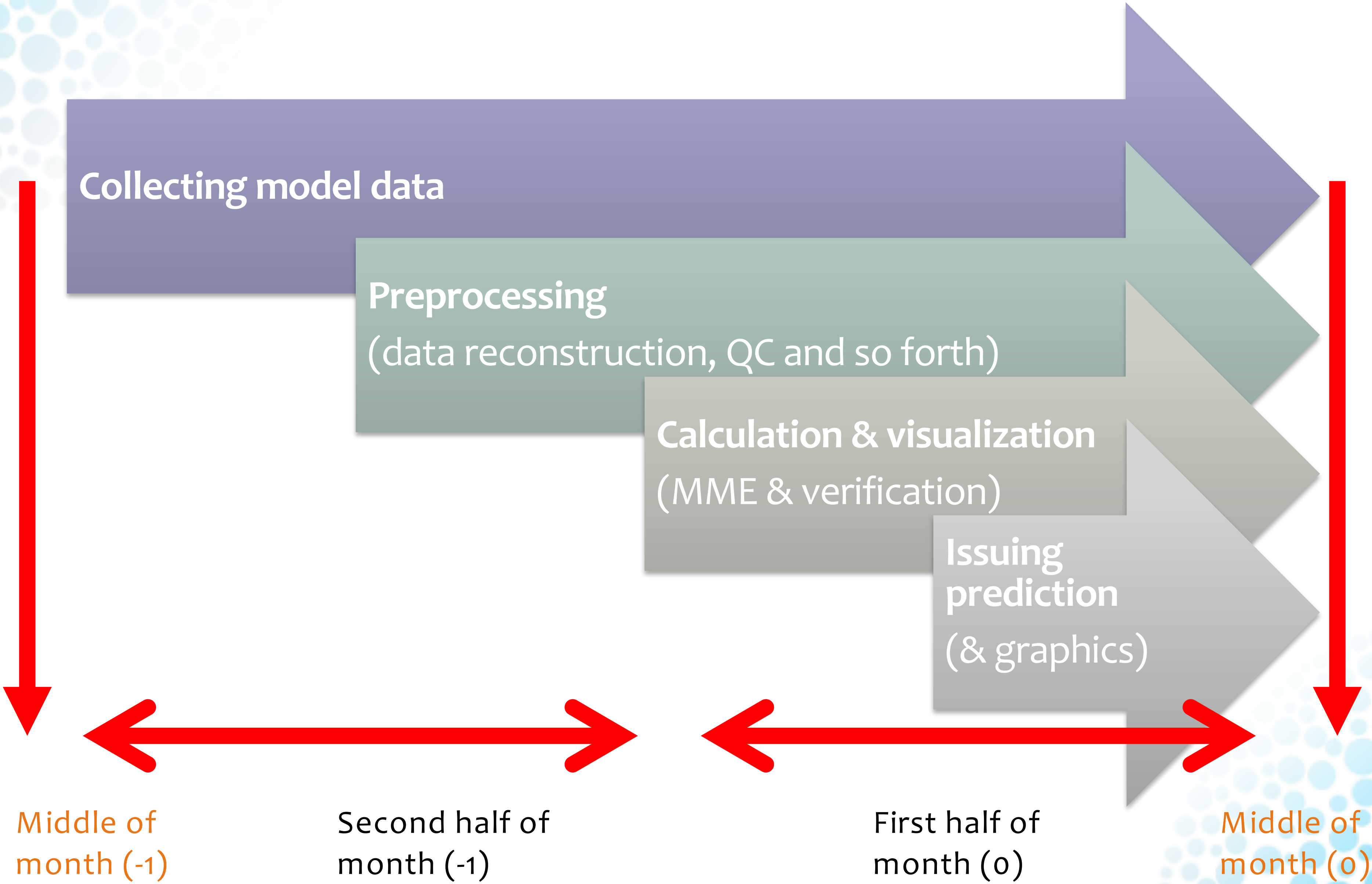
Collection of Dynamic ensemble seasonal prediction data from NMHS and research institutes (16 operations/institutions from 10 countries)

List of available models

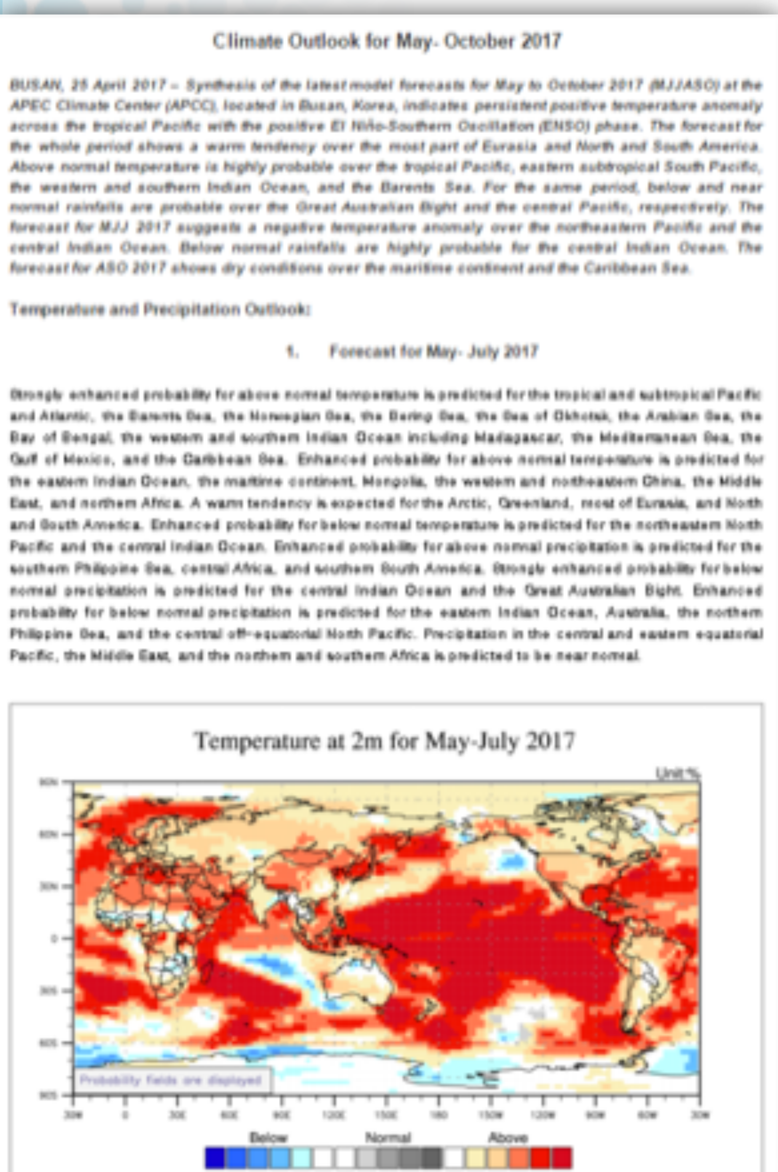
(as of March 2018)

Nation	Organization	Acronym
Australia	Bureau of Meteorology	BoM
Canada	Meteorological Service of Canada	MSC
Chinese Taipei	Central Weather Bureau	CWB
Italy	Centro Euro-Mediterraneo sui Cambiamenti Climatici	CMCC
Japan	Japan Meteorological Agency	JMA
Korea	APEC (Asia-Pacific Economic Cooperation) Climate Center	APCC
	Pusan National University	PNU
USA	National Centers for Environmental Prediction, NOAA	NCEP
	National Aeronautics and Space Administration	NASA

APCC operational forecast



APCC Climate Information



Climate Outlook for May- October 2017

BUSAN, 25 April 2017 – Synthesis of the latest model forecasts for May to October 2017 (MJJASO) at the APEC Climate Center (APCC), located in Busan, Korea, indicates persistent positive temperature anomaly across the tropical Pacific with the positive El Niño-Southern Oscillation (ENSO) phase. The forecast for the whole period shows a warm tendency over the most part of Eurasia and North and South America. Above normal temperature is highly probable over the tropical Pacific, eastern subtropical South Pacific, the western and southern Indian Ocean, and the Barents Sea. For the same period, below and near normal rainfalls are probable over the Great Australian Bight and the central Pacific, respectively. The forecast for MJJ 2017 suggests a negative temperature anomaly over the northeastern Pacific and the central Indian Ocean. Below normal rainfalls are highly probable for the central Indian Ocean. The forecast for ASO 2017 shows dry conditions over the maritime continent and the Caribbean Sea.

Temperature and Precipitation Outlook:

1. Forecast for May- July 2017

Strongly enhanced probability for above normal temperature is predicted for the tropical and subtropical Pacific and Atlantic, the Barents Sea, the Norwegian Sea, the Bering Sea, the Sea of Okhotsk, the Arabian Sea, the Bay of Bengal, the western and southern Indian Ocean including Madagascar, the Mediterranean Sea, the Gulf of Mexico, and the Caribbean Sea. Enhanced probability for above normal temperature is predicted for the eastern Indian Ocean, the maritime continent, Mongolia, the western and northeastern China, the Middle East, and northern Africa. A warm tendency is expected for the Arctic, Greenland, most of Eurasia, and North and South America. Enhanced probability for below normal temperature is predicted for the northeastern North Pacific and the central Indian Ocean. Enhanced probability for above normal precipitation is predicted for the southern Philippine Sea, central Africa, and southern South America. Strongly enhanced probability for below normal precipitation is predicted for the central Indian Ocean and the Great Australian Bight. Enhanced probability for below normal precipitation is predicted for the eastern Indian Ocean, Australia, the northern Philippine Sea, and the central off-equatorial North Pacific. Precipitation in the central and eastern equatorial Pacific, the Middle East, and the northern and southern Africa is predicted to be near normal.

Temperature at 2m for May-July 2017

Seasonal outlook (global)

ENSO & IOD prediction

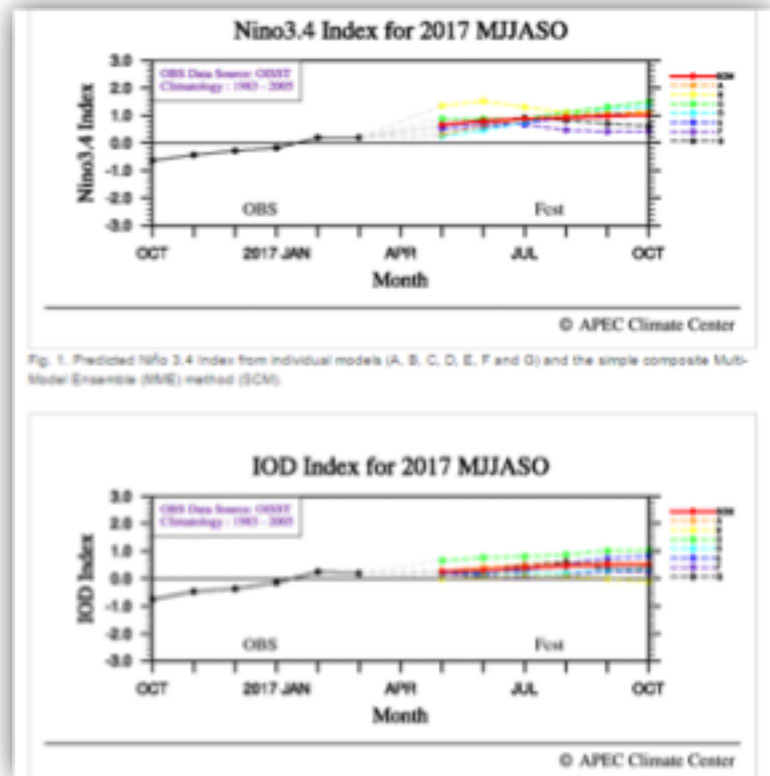
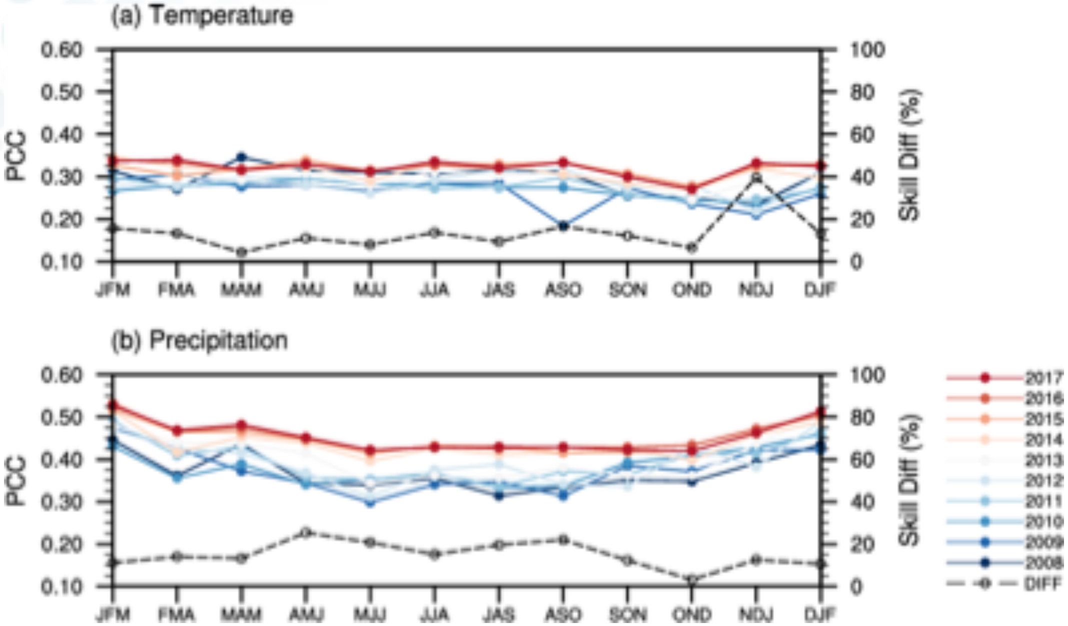


Fig. 1. Predicted Niño 3.4 Index from individual models (A, B, C, D, E, F and G) and the simple composite Multi-Model Ensemble (MME) method (SCM).

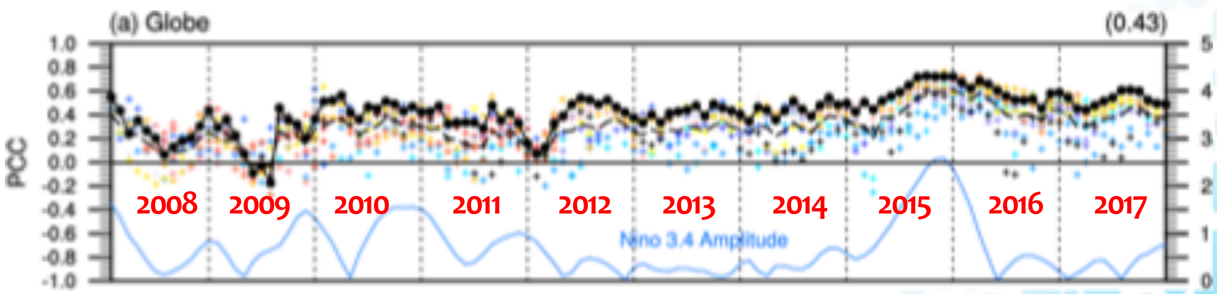
APCC MME skill improvement

Hindcast Skill: ACC



By YM Min (APCC)

RT Forecast Skill: T850, ACC



CLIK Training (16-17 Oct 2018, Hanoi, Vietnam)

Evaluation of Seasonal Prediction

Yun-Young Lee

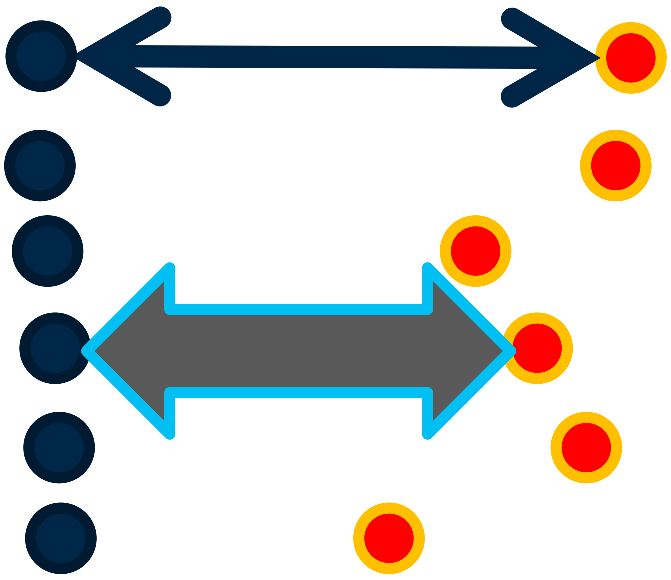
APCC APEC CLIMATE CENTER

How GOOD?

- Evaluation of forecast : verification

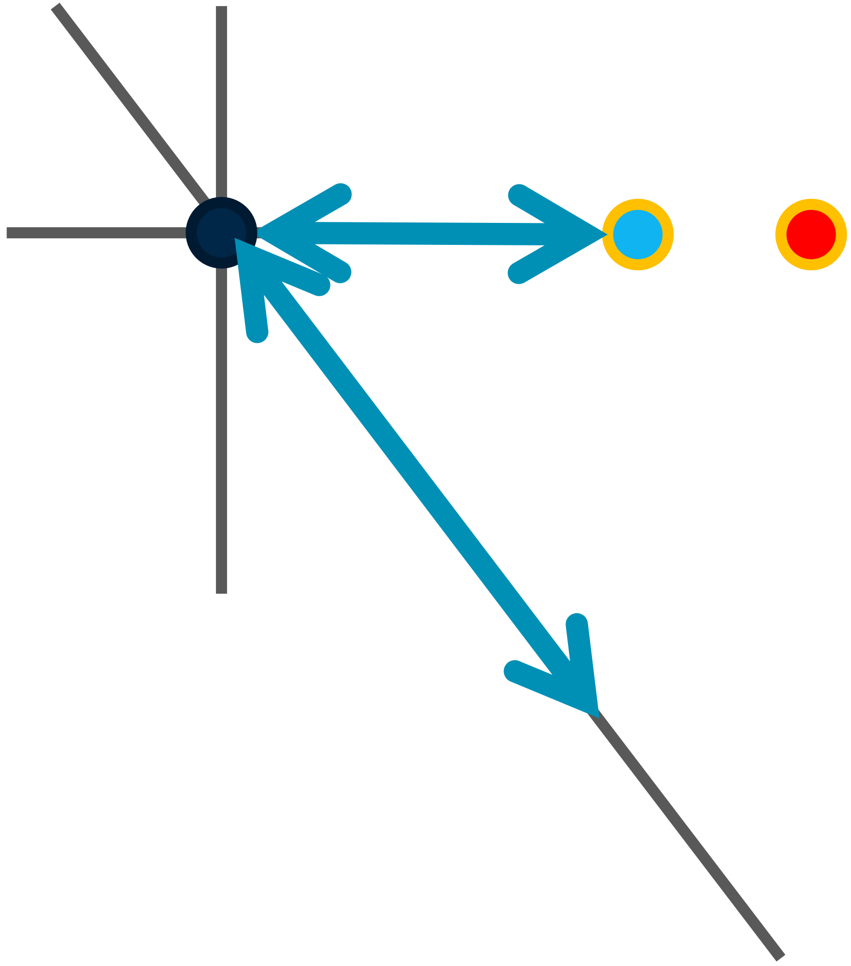
Verification

- Evaluation : measure of **closeness**



Verification

- Evaluation : depends on **Dimension/Viewpoint**



Deterministic forecast

- Various measures
  - **MSE** (Mean Square Error), RMSE (Root MSE)

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

skill score = $\frac{SCORE_{forecast} - SCORE_{reference}}{SCORE_{perfect\ forecst} - SCORE_{reference}}$

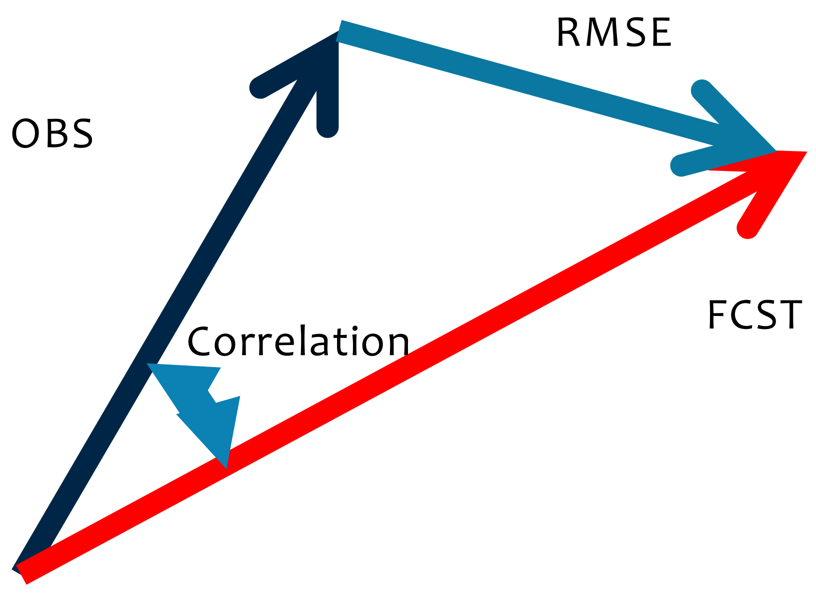
,Which is designed to give an answer for the question "What is the relative improvement of the forecast over some reference forecast?

  - **MSSS** (Mean Square Skill Score)
    - Conventional form of "skill score"
    - $1 - \frac{MSE}{MSE_c}$, MSE : error/penalty, MSEc : error of climatology forecast
  - **ACC**(Anomaly correlation, Pattern), TCC (Temporal correlation)

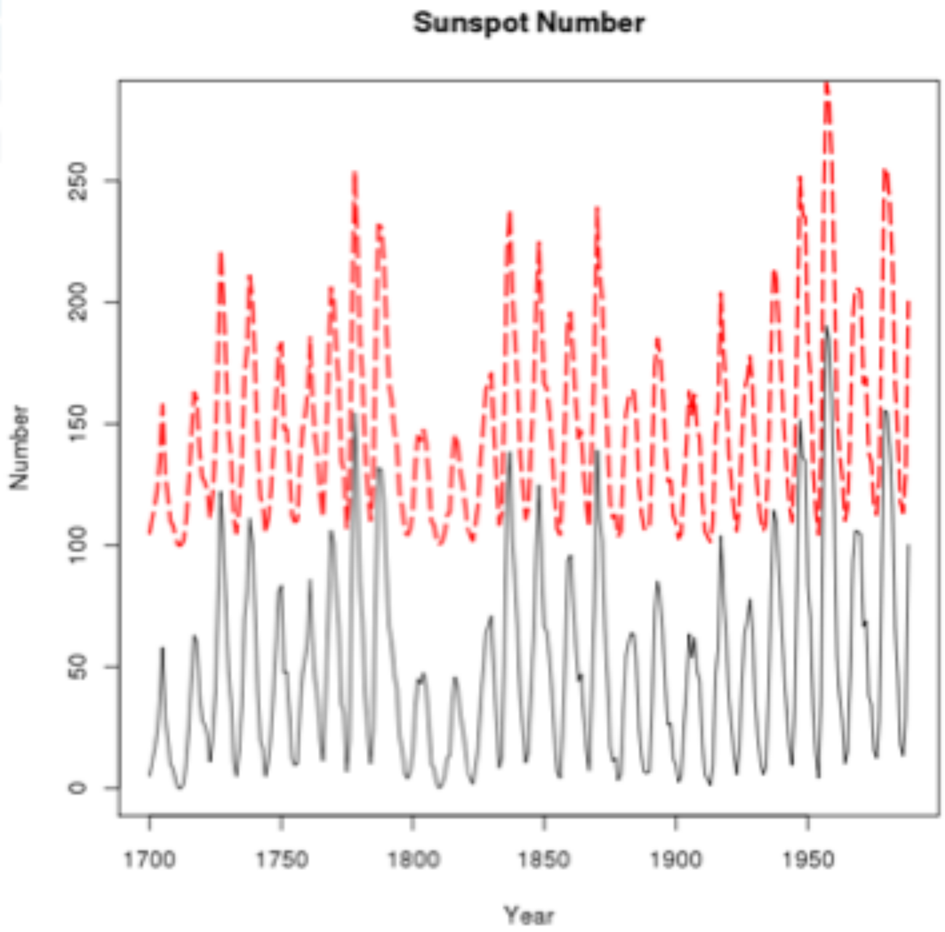
$$ACC = \frac{\sum_{i=1}^{N} w_i(f_i - f)(o_i - o)}{\sqrt{\sum_{i=1}^{N} w_i(f_i - f)^2 \sum_{i=1}^{N} w_i(o_i - o)^2}}$$

Verification

- Evaluation : depends on **Dimension/Viewpoint**



Examples



F(t)
G(t) = F(t) + 100

RMSE?

But, Correlation?

Probabilistic forecast

- Brier score (BS)
  - **MSE of prob. forecast**

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

    F=probability(forecast),

    O=1/0 (actual outcome of instance)

    Eg. Binary events such as **“rain” or “no rain”**

    Range: 0 to 1, Perfect score = 0

    Intuitively, Higher score = Better skill?

skill score = $\frac{SCORE_{forecast} - SCORE_{reference}}{SCORE_{perfect\ forecst} - SCORE_{reference}}$

,Which is designed to give an answer for the question “What is the relative improvement of the forecast over some reference forecast?

- Brier **Skill Score (BSS)**

  BSS = 1 $-\frac{\boldsymbol{BS}}{\boldsymbol{BS_c}}$

  BS : error/penalty,

  BS_c: BS of climatology forecast

  Range: -infinity to 1

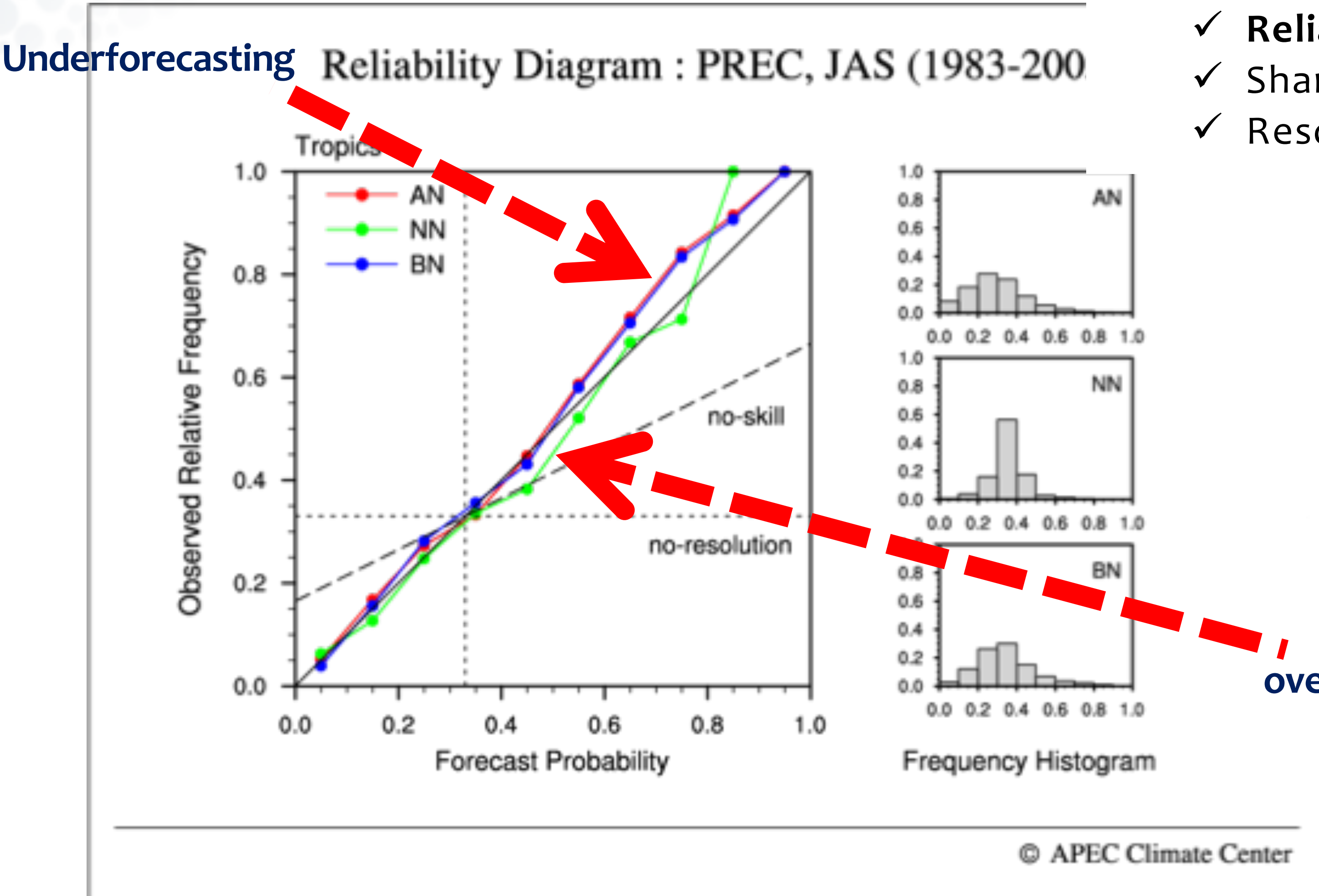
  Perfect score = 1

Probabilistic forecast (Categorical)

- Reliability curve

Accurate probability forecast system

- ✓ **Reliability**
- ✓ Sharpness
- ✓ Resolution



Probabilistic forecast (Categorical)

“rain” “no rain”

Correct forecast

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

- ***HR (Hit rate)*** *= H/(H+M)*

  0 to 1, perfect score = 1

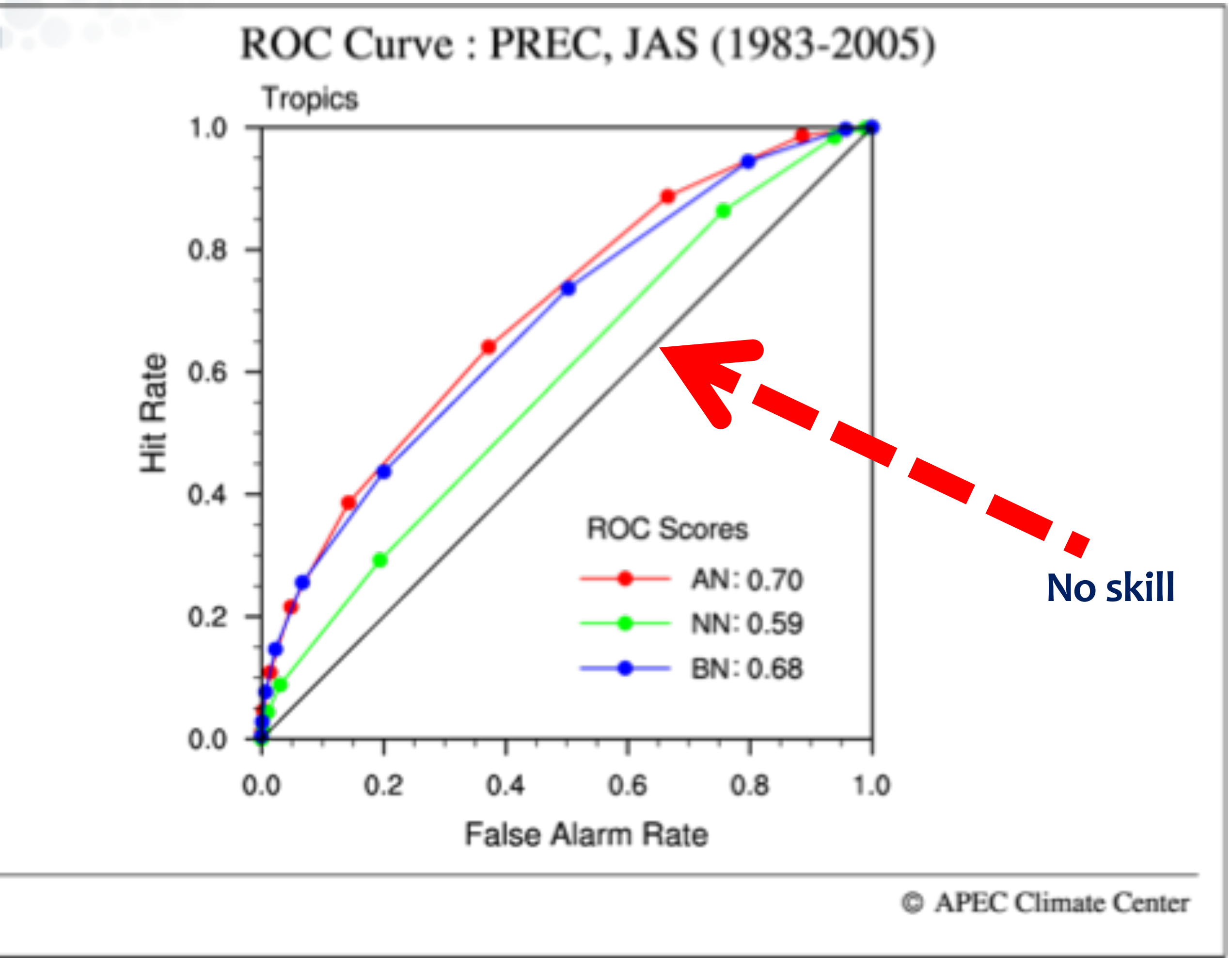
- ***FAR (False Alarm rate)*** *= F/(F+C)*

  0 to 1, perfect score = 0

Good forecast : ***HR↑, FAR ↓***

Probabilistic forecast (Categorical)

- ROC (Relative Operating Characteristics)

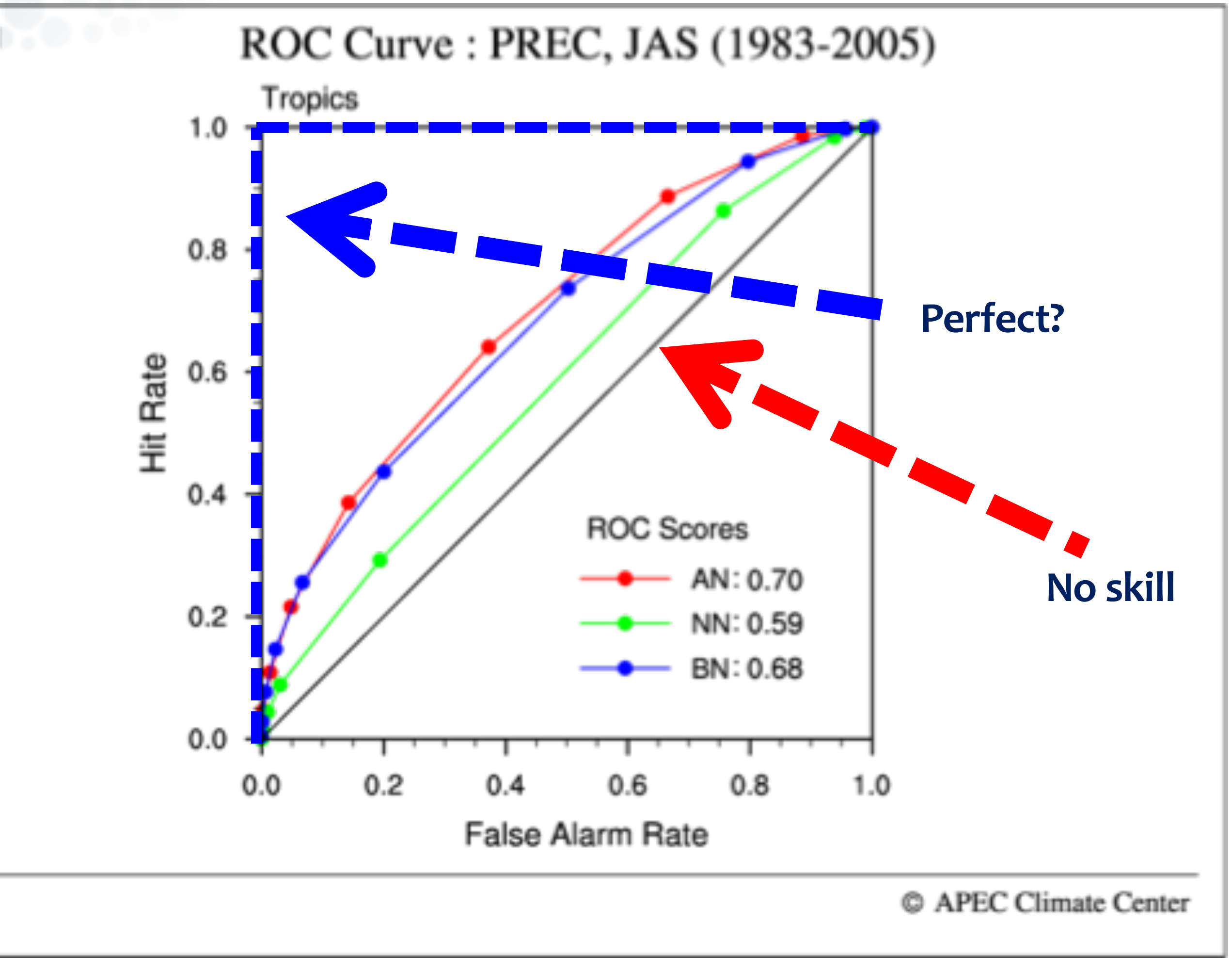


ROC score
=Area Under ROC curve

Range: 0 to 1
perfect: ROC score = 1
no skill: ROC score = 0.5
(no added value)

Probabilistic forecast (Categorical)

- ROC (Relative Operating Characteristics)

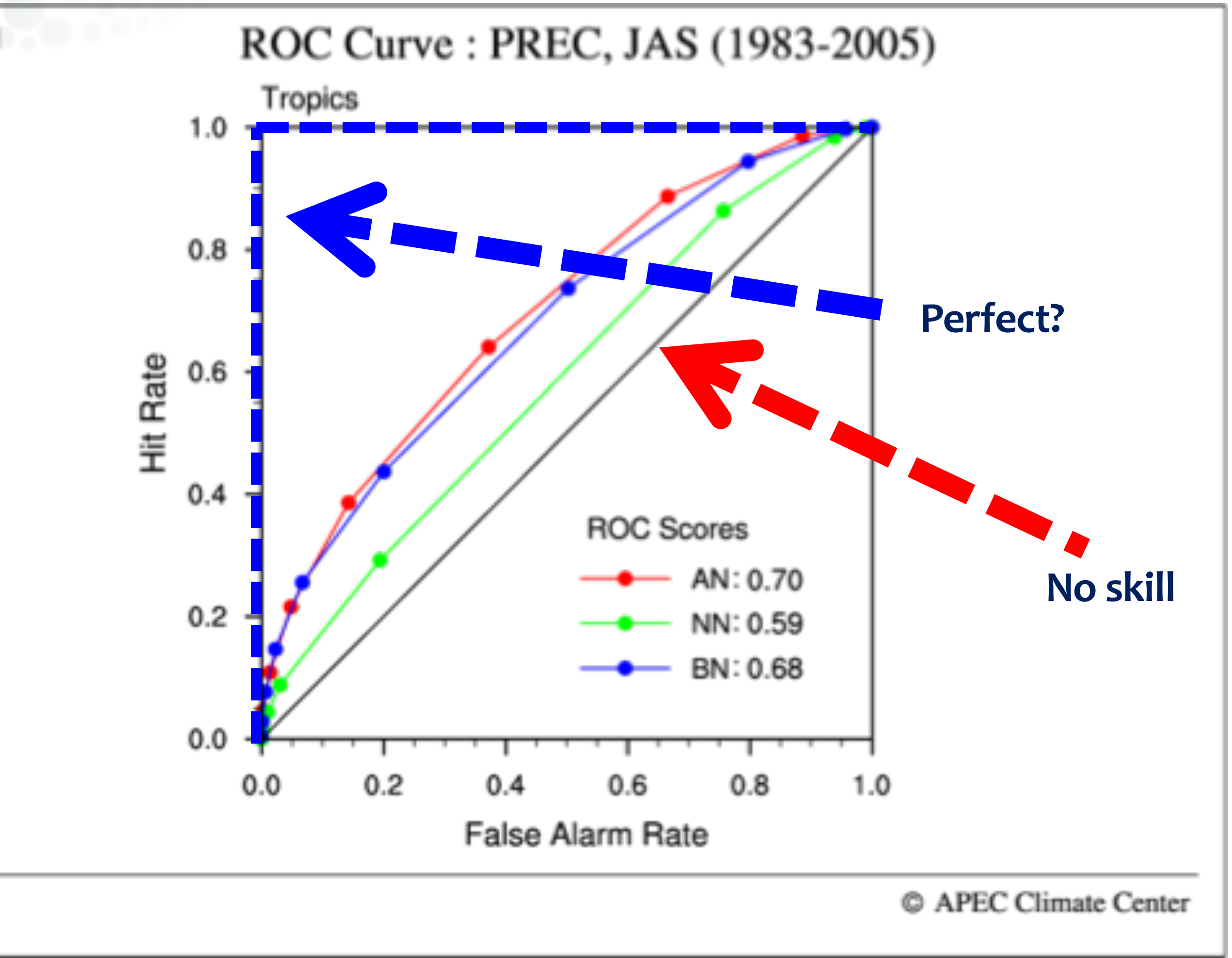


ROC score
=Area of ROC curve

Range: 0 to 1
perfect: ROC score = 1
no skill: ROC score = 0.5
(no added value)

Probabilistic forecast (Categorical)

- ROC (Relative Operating Characteristics)



- ✓ Biased forecast with high ROC score
- ✓ A measure of ***potential usefulness***

ROC score
=Area of ROC curve

Range: 0 to 1
perfect: ROC score = 1
no skill: ROC score = 0.5
(no added value)

Probabilistic forecast (Categorical)

▪ HSS (Heidke **Skill Score**)

*Giving the answer for the question **"What was the accuracy of the forecast in predicting the correct category, relative to that of random chance?"***

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

$$HSS = \frac{SCORE_{forecast} - SCORE_{by\ chance}}{SCORE_{perfect\ forecst} - SCORE_{by\ chance}}$$

$$= \frac{\left\{\frac{(H+C)}{n} - \frac{[(H+F)(H+M)+(F+C)(M+C)]}{n^2}\right\}}{\left\{1 - \frac{[(H+F)(H+M)+(F+C)(M+C)]}{n^2}\right\}}$$

*Range: - infinity to 1, **0=no skill,** 1=perfect skill*

Money Problem!

Drought expected!

Preparing (reservoir)	Drought happens?	Net cost
Yes (cost) 	No (No money loss because of drought)	
No (no cost)	Yes (money loss) 	

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

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

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

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

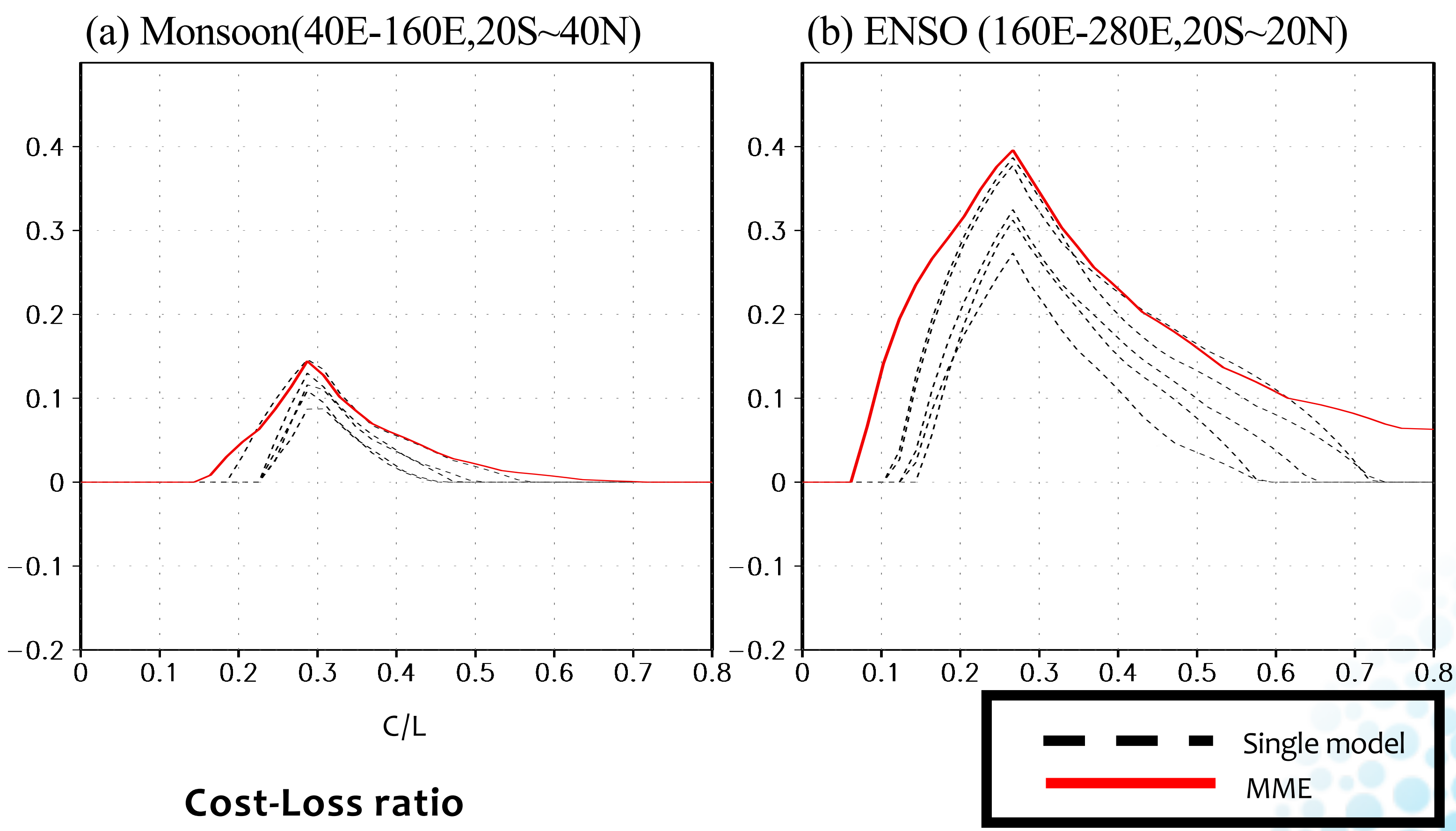
If Yes : E → C, otherwise E → ōL.

Decision: action of low expense. Thus, $\boxed{E_{cli} = \min(C, \bar{o}L)}$

$$V = \frac{\min(C, oL) - (h+f)C - mL}{\min(C, oL) - oC} = \frac{\min\left(\frac{C}{L}, o\right) - (h+f)\frac{C}{L} - m}{\min\left(\frac{C}{L}, o\right) - o\frac{C}{L}}$$

Value of Probabilistic forecast (Above normal) : GCMs

$$V = \frac{\min(\frac{C}{L},\bar{o}) - (h+f)\frac{C}{L} - m}{\min(\frac{C}{L},\bar{o}) - \frac{C}{L}\bar{o}} = \frac{\min(\frac{C}{L},\bar{o}) - f(1-\bar{o})\frac{C}{L} + h\bar{o}(1-\frac{C}{L}) - \bar{o}}{\min(\frac{C}{L},\bar{o}) - \frac{C}{L}\bar{o}}$$



Forecast Verification

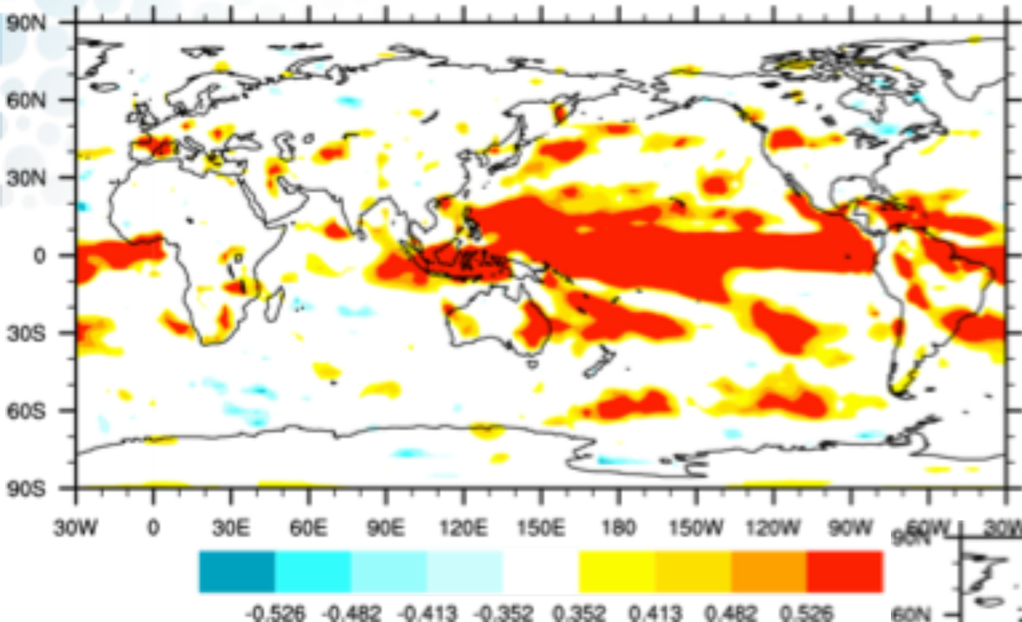
- Multi aspect evaluation is necessary.
- A single verification score (e.g. R=0.5, explaining 25% variance) cannot tell everything.

- User oriented verification would be useful.
- If not clear, use popular one.

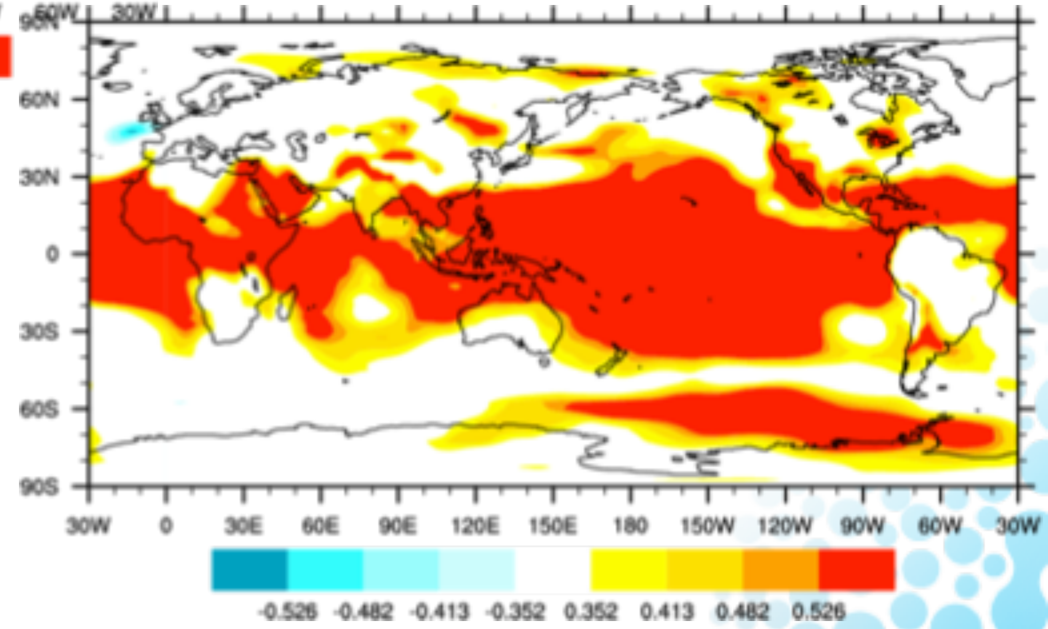
- Difficulties in "translating" meteorological skill score into Public wording.

- Let's see some results!!!

APCC MME (TCC)

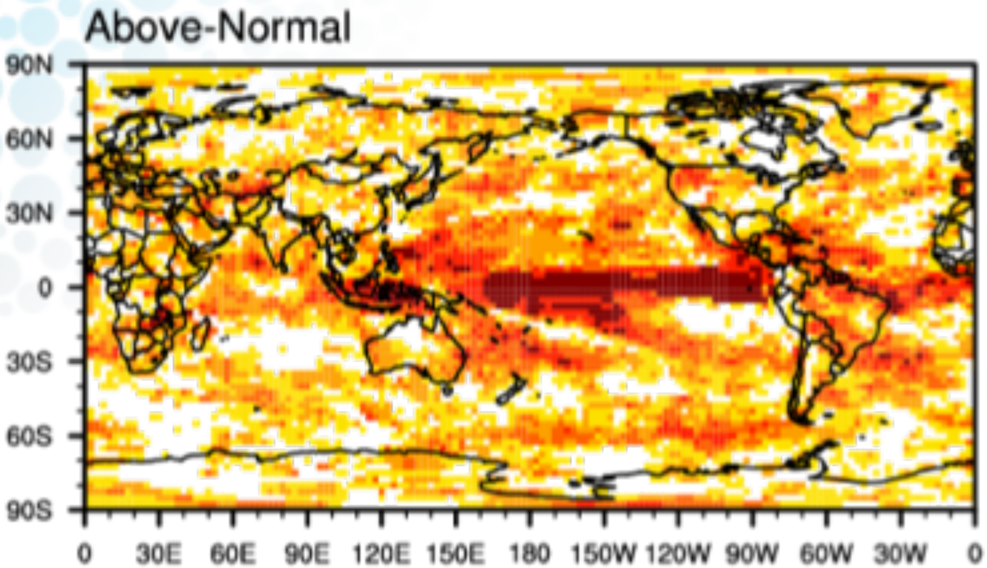


Rainfall (JJA)



SLP (JJA)

ROC Score : PREC, JJA (1983-2005)



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