The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

1956-5

Targeted Training Activity: Seasonal Predictability in Tropical Regions to be followed by Workshop on Multi-scale Predictions of the Asian and African Summer Monsoon

4 - 15 August 2008

Status of Seasonal Prediction Current & Predictability.

Kang In Sik

Seoul National University
School of Earth and Environmental Sciences
Division of Atmospheric Sciences
San 56-1 Shillim-dong, Kwanak-ku, Seoul 151-742
REPUBLIC OF KOREA

Current Status of Seasonal Prediction & Predictability

In-Sik Kang

Climate and Environment System Research Center
Seoul National University, Korea

Content

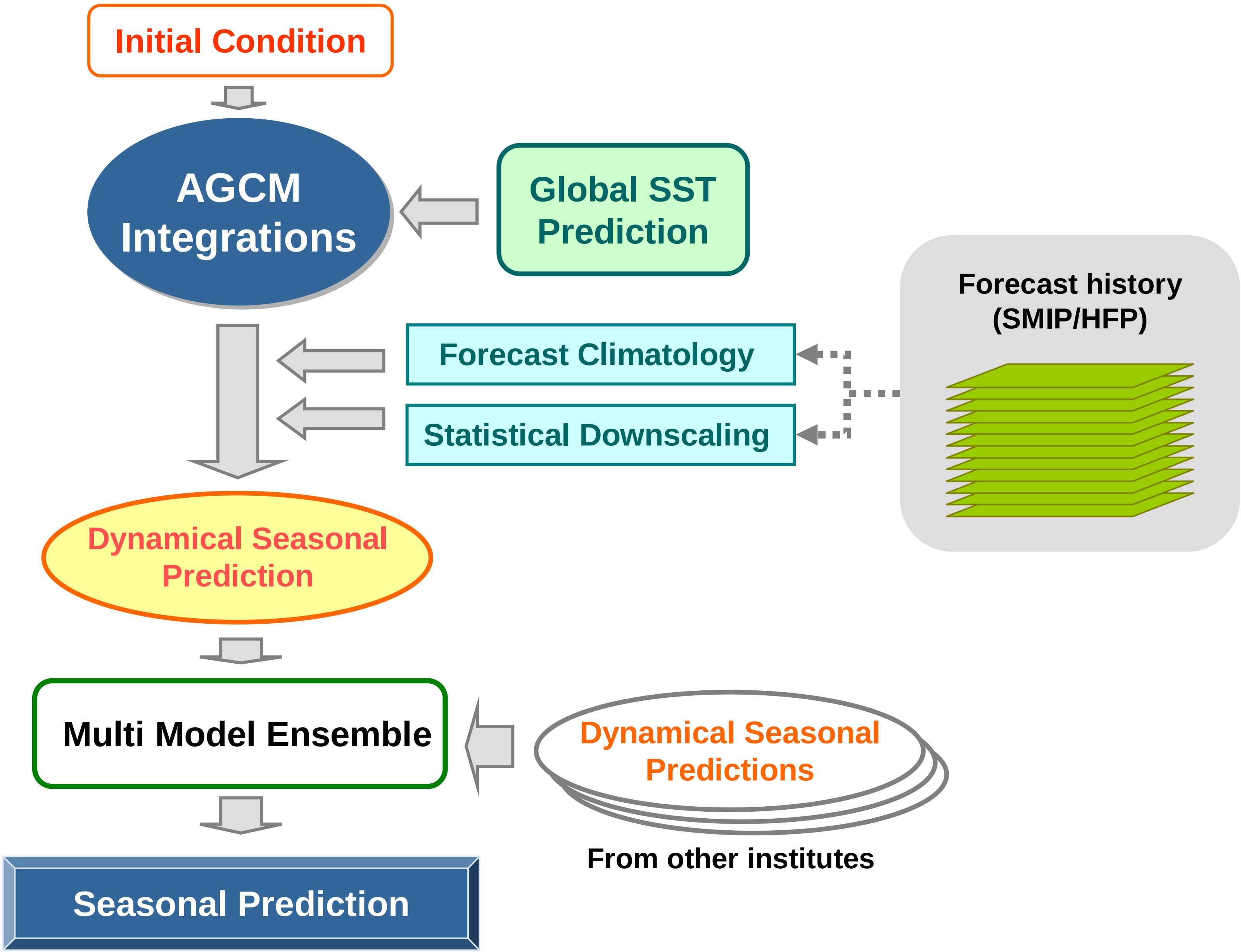
Seasonal Prediction & Predictability

- Potential Predictability
- Real predictability – Tier-2 system
- Coupled model predictability – Tier-1 system
- Multi-model Ensemble

Intraseasonal prediction & predictability

- Statistical models
- Dynamical model
- Statistical + Dynamical Combined

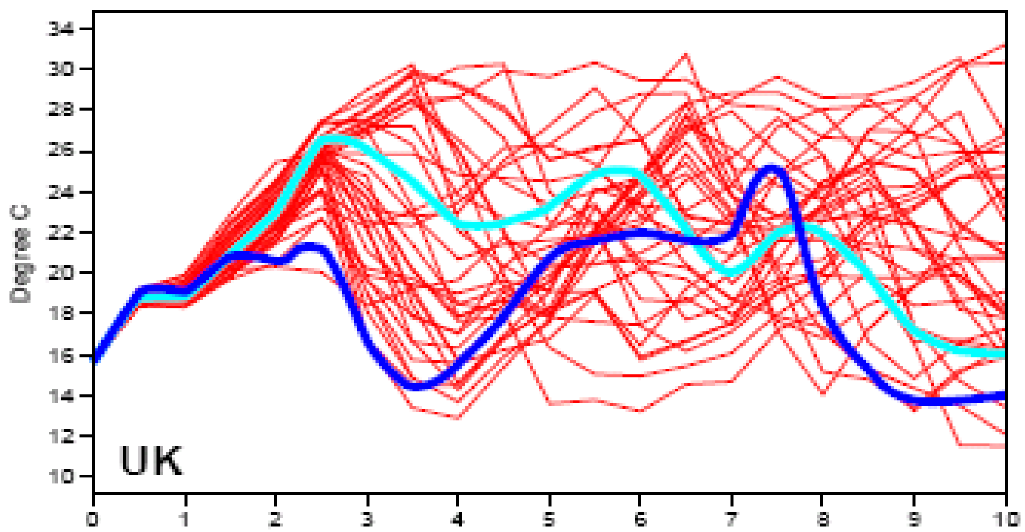
Seasonal Prediction System



Uncertainty of seasonal prediction

The Nature of the Seasonal Prediction

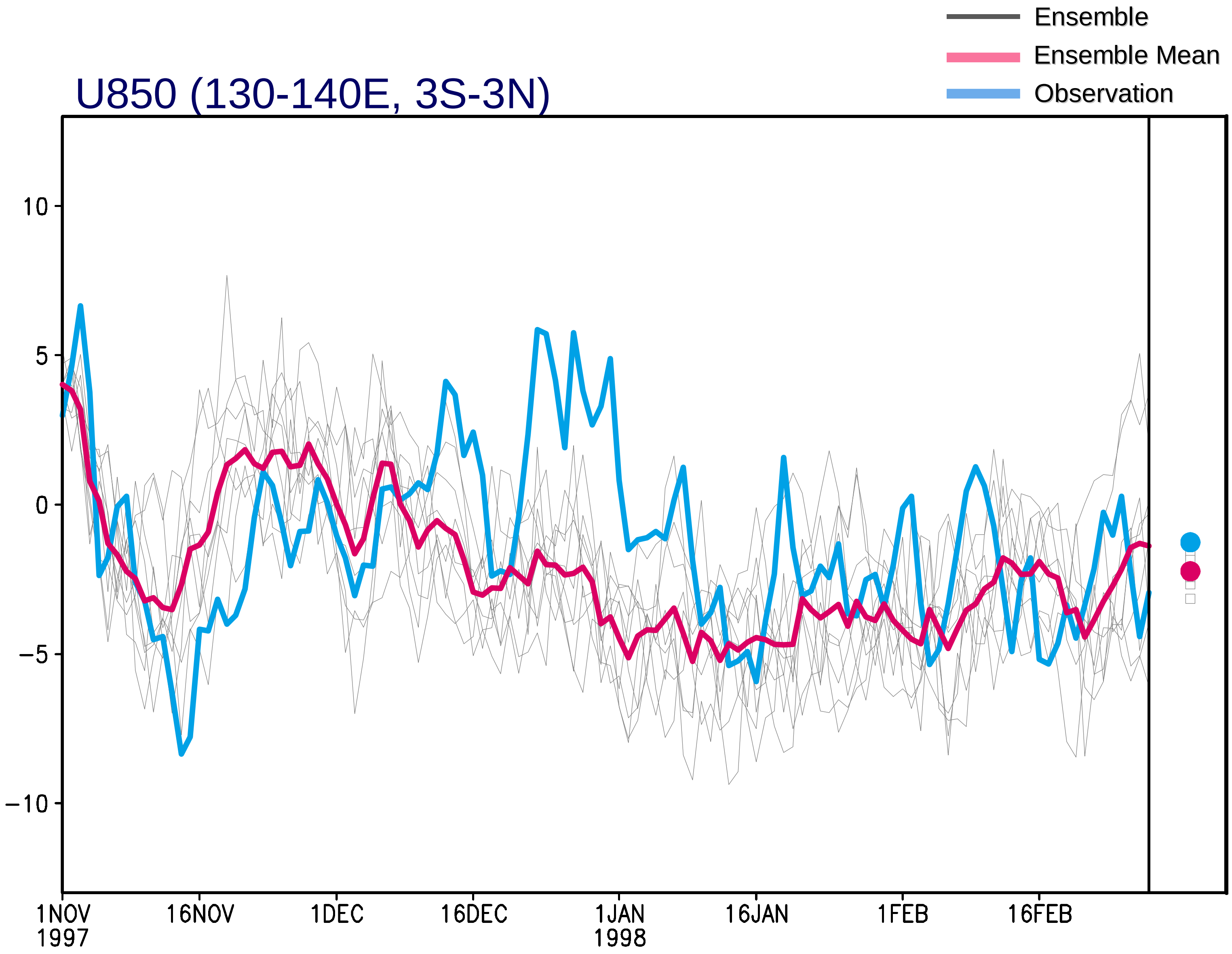
A Realization = Signal + Noise



Ensembles from
Small Initial Perturbation

Ensemble
Ensemble Mean
Observation

Ensembles Forecasts with Small Initial Perturbations



The Nature of the Seasonal Prediction

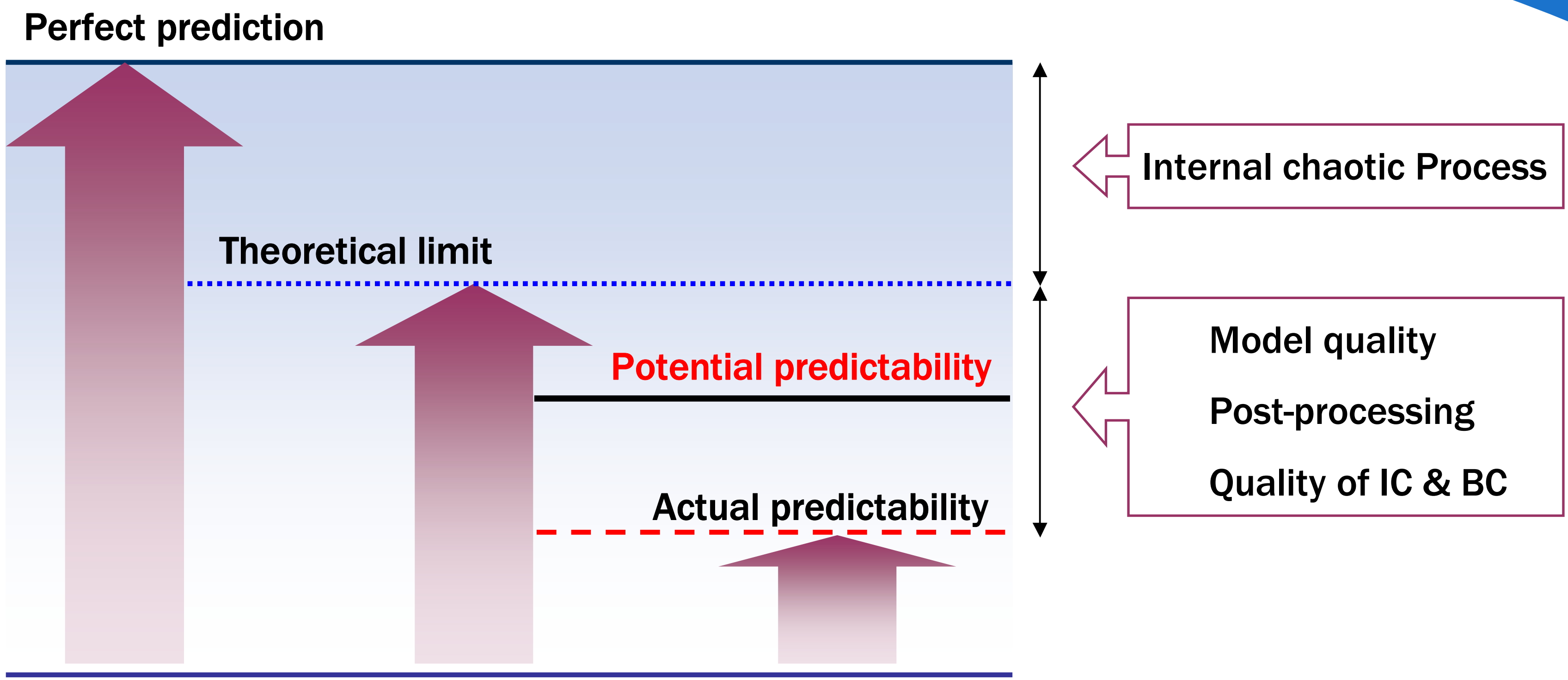
Uncertainty of seasonal prediction

A Realization = True + Error + Noise

- Uncertainty from Initial Condition
- Uncertainty from Model Physics and Dynamics

Imperfectness of Model → Systematic Error

Predictability of Seasonal Prediction

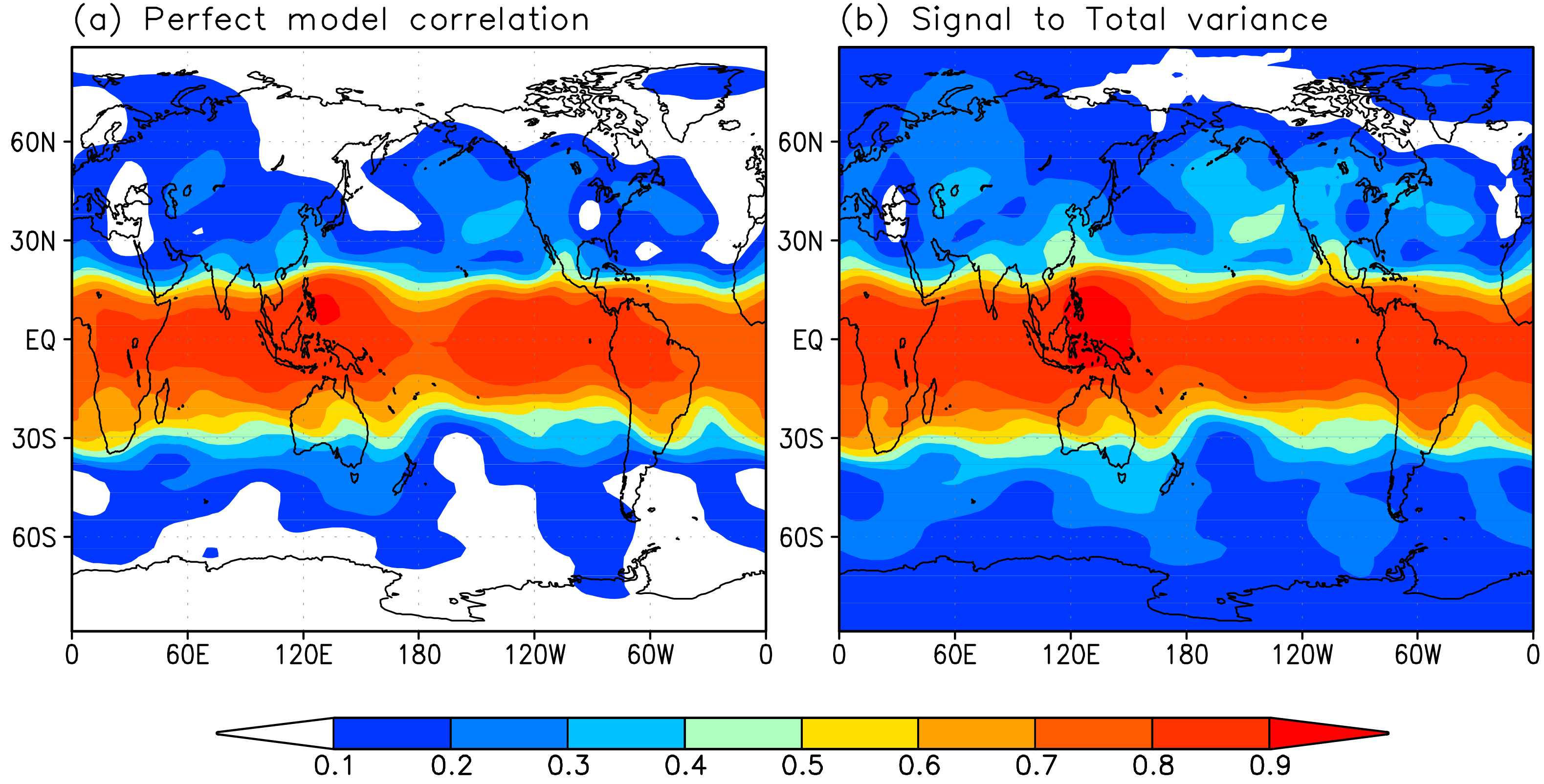


In climate prediction, potential predictability is regarded as the predictability with full information of future boundary condition (e.g., SST). Thus, predictability is varied with similarity between the response of real atmosphere and prediction method to the same BC.

Establish "potentially" possible prediction skill with state-of-art prediction system

Perfect model correlation & Signal to Total variance ratio

Z500 winter (C20C, 100 seasons, 4 members)



Although the 4 member is not enough to estimate Potential predictability precisely, the patterns of 2 metrics are quite similar

Decomposition of climate variables

Climate state variable (X) consists of predictable and unpredictable part.

- Predictable part = *signal* (Xs) : forced variability
- Unpredictable part = *noise* (Xn) : internal variability

$$X = Xs + Xn$$

The dynamical forecast (Y) also have its forced and unforced part.

forecast *signal* (Ys) : forced variability of model

forecast *noise* (Yn) : internal variability of model

$$Y = Ys + Yn$$

The internal variability (noise) is stochastic

If the forecast model is not perfect, Xs≠Ys. (there is a systematic error)

Upper limit of prediction

Maximizing correlation in the presence of error in signal and noise

$$Observation\ :\ x = x_s + x_n$$

$$Forecast : y = y_s + y_n = \alpha x_s + y_e + y_n$$

Noise y_n and Error y_e are not correlated with others.

α : regression coefficient of signal

$$
\begin{aligned}
Cov(x,y) &= (\overline{xy}) = (\overline{(x_s + x_n)(\alpha x_s + y_e + y_n)}) \\
&= (\alpha \overline{x_s^2} + \alpha \overline{x_s x_n} + \overline{x_s y_e} + \overline{x_n y_e} + \overline{x_s y_n} + \overline{x_n y_n}) \\
&= \alpha V(x_s) + \alpha Cov(x_s, x_n) + Cov(x_s, y_e) + Cov(x_n, y_e) + Cov(x_s, y_n) + Cov(x_n, y_n) \\
&= \alpha V(x_s)
\end{aligned}
$$

Correlation between observation (x) and forecast (y)

$$\text{Cor (x,y)} = \frac{\alpha V(x_s)}{V(x)^{1/2}\left[\alpha^2 V(x_s) + V(y_e) + V(y_n)\right]^{1/2}}$$

The correlation coefficient is maximized by removing *$V(y_e)$* and *$V(y_n)$*

➔ The most accurate forecast will be the *SIGNAL* of perfect model.

Strategy of Prediction

The strategy of seasonal prediction is to

obtain "perfect signal" as close as possible.

(i.e. reducing variance of systematic error and variance of noise)

1. Reduction of Noise

- **Averaging large ensemble members**

(if number of ensemble members is infinite, Noise will be zero in the ensemble mean)

2. Correct signal

- **Improving GCM**
- **Statistical post-process (MOS)**
- **Multi-model ensemble**

Maximum prediction skill : potential predictability

$$\text{Cor(x,y)} = \frac{\alpha V(x_s)}{V(x)^{1/2}\left[\alpha^2 V(x_s) + V(y_e) + V(y_n)\right]^{1/2}} = \frac{V(x_s)}{\left[V(x)V(x_s)\right]^{1/2}}$$

$$= \frac{V(x_s)^{1/2}}{V(x)^{1/2}}$$

$$\sqrt{\frac{V_{Signal}}{V_{Total}}} = \sqrt{\frac{\rho}{1+\rho}}, \quad \rho = \frac{V_{Signal}}{V_{Noise}} \equiv SNR$$

$\rho : Signal\ to\ Noise\ Ratio$

$V_{Total} = V_{signal} + V_{noise}$

Maximum prediction skill (=potential predictability of particular predictand) is a function of Signal to Noise Ratio

Forced & Free variance

Forced variance

Climate signals

caused by external forcing (e.g. SST)

$$\frac{1}{N-1}\sum_{i=1}^{N}(\bar{X}_i - \bar{\bar{X}})^2 - \alpha$$

$\alpha = \frac{1}{n}\cdot$ **Free variance**

□ **Ensemble mean variation**

Free variance

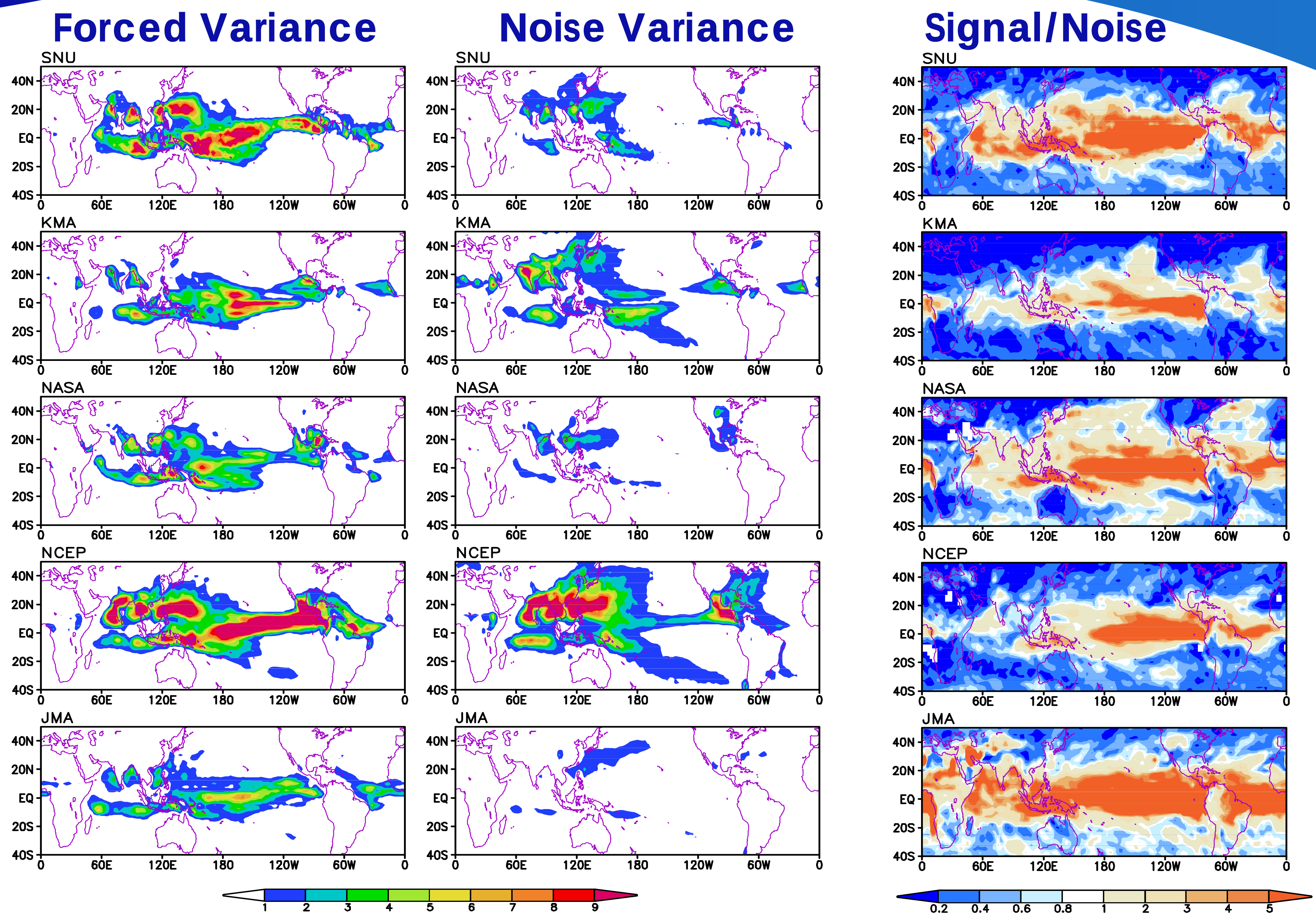
Intrinsic transients

due to natural variability

$$\frac{1}{N(n-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2$$

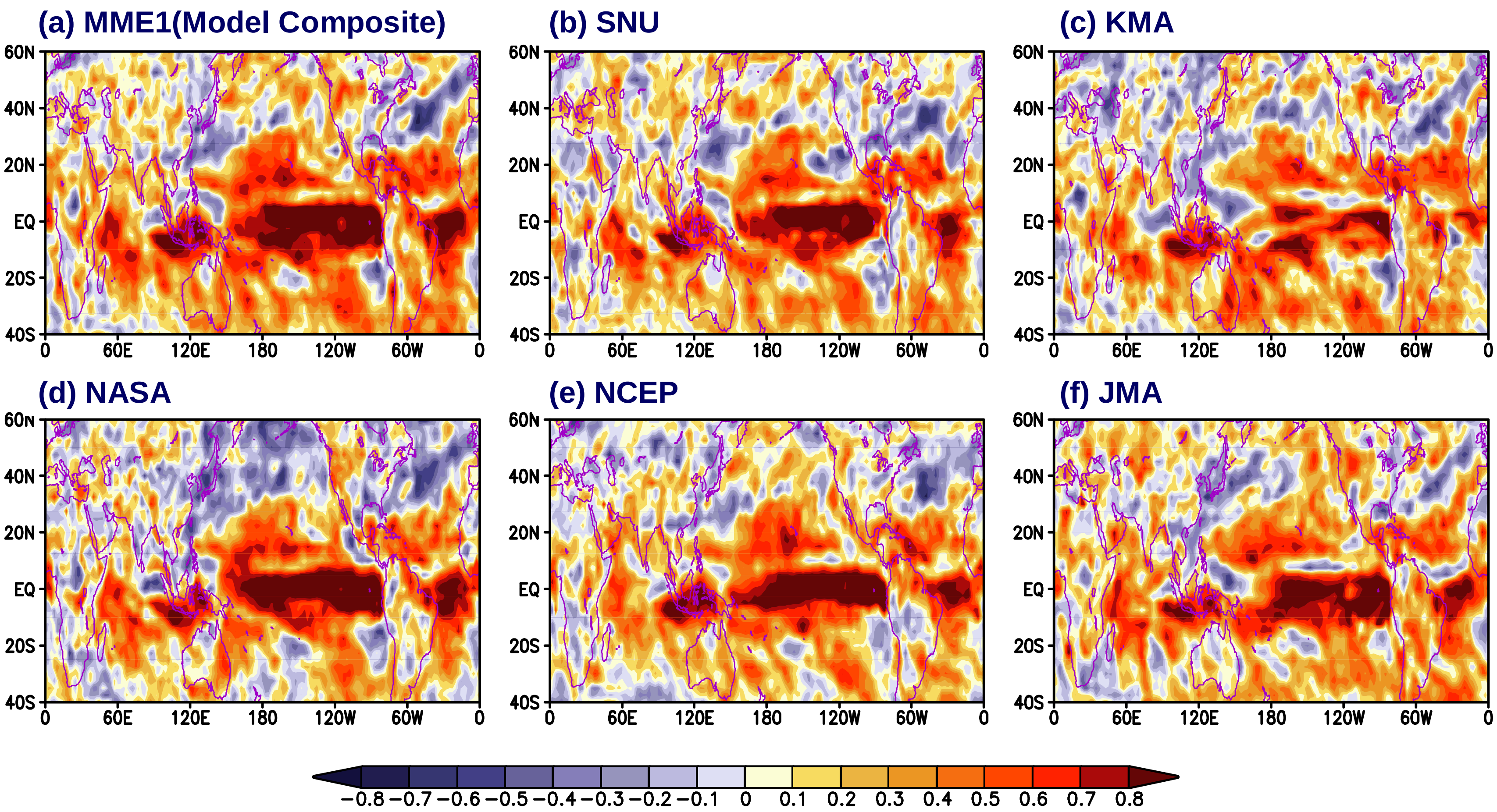
Ensemble spread

Variance analysis of JJA Precipitation Anomalies

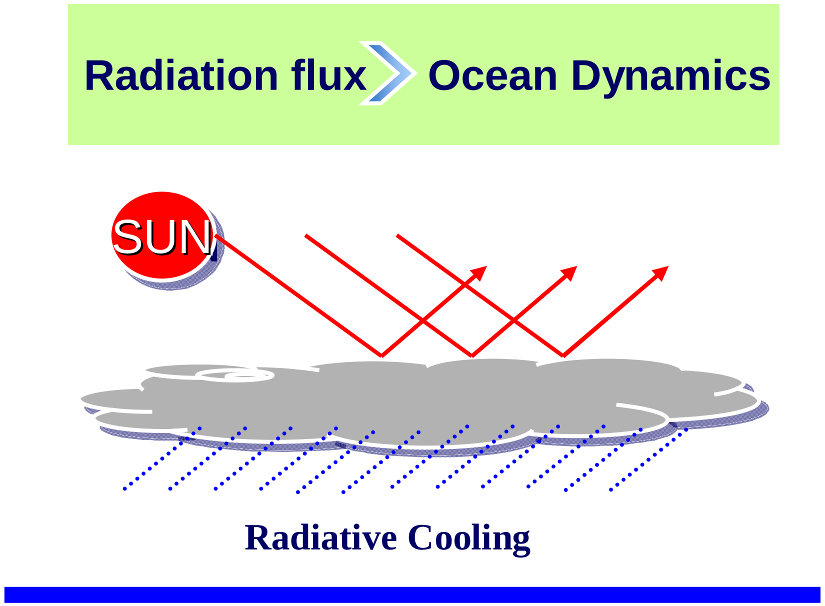


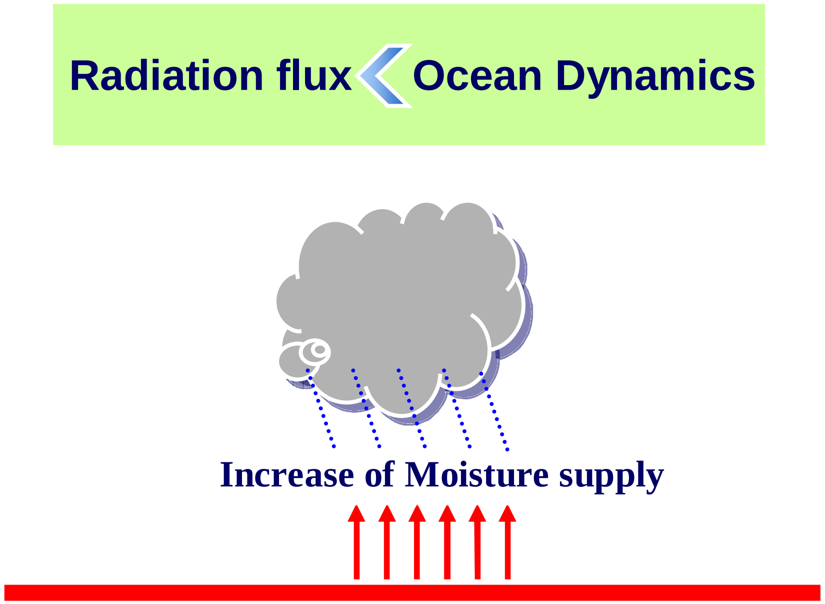
Prediction Skill of JJA Precipitation (21 yr)

Temporal Correlation



Air-sea interaction in the tropical Pacific



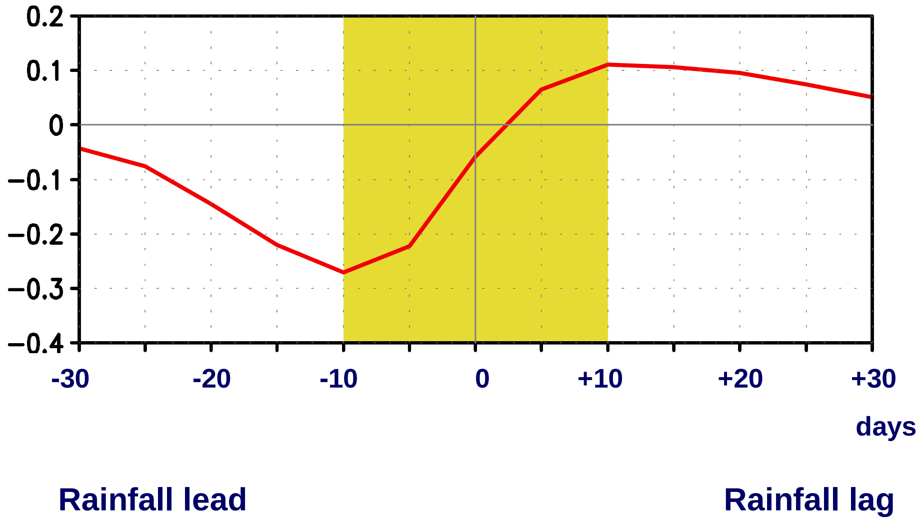


Where radiative flux control the SST…

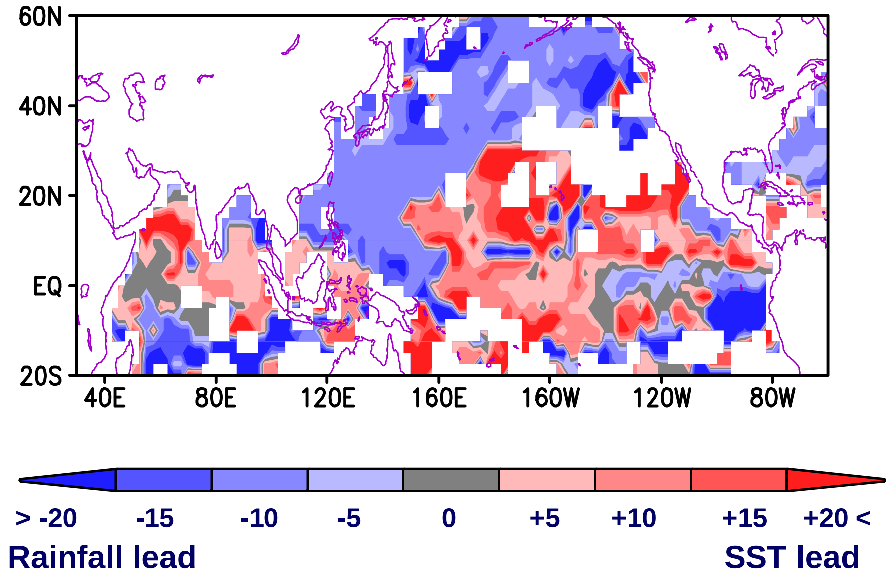
1. Radiative flux would lead the SST anomalies
2. Temporal correlation between PRCP & SST can be a negative sign

Lead-lag correlation between pentad SST and rainfall data for JJA 82-99

Western North Pacific (5-30N, 110-150E)



Lead-lag pentad number

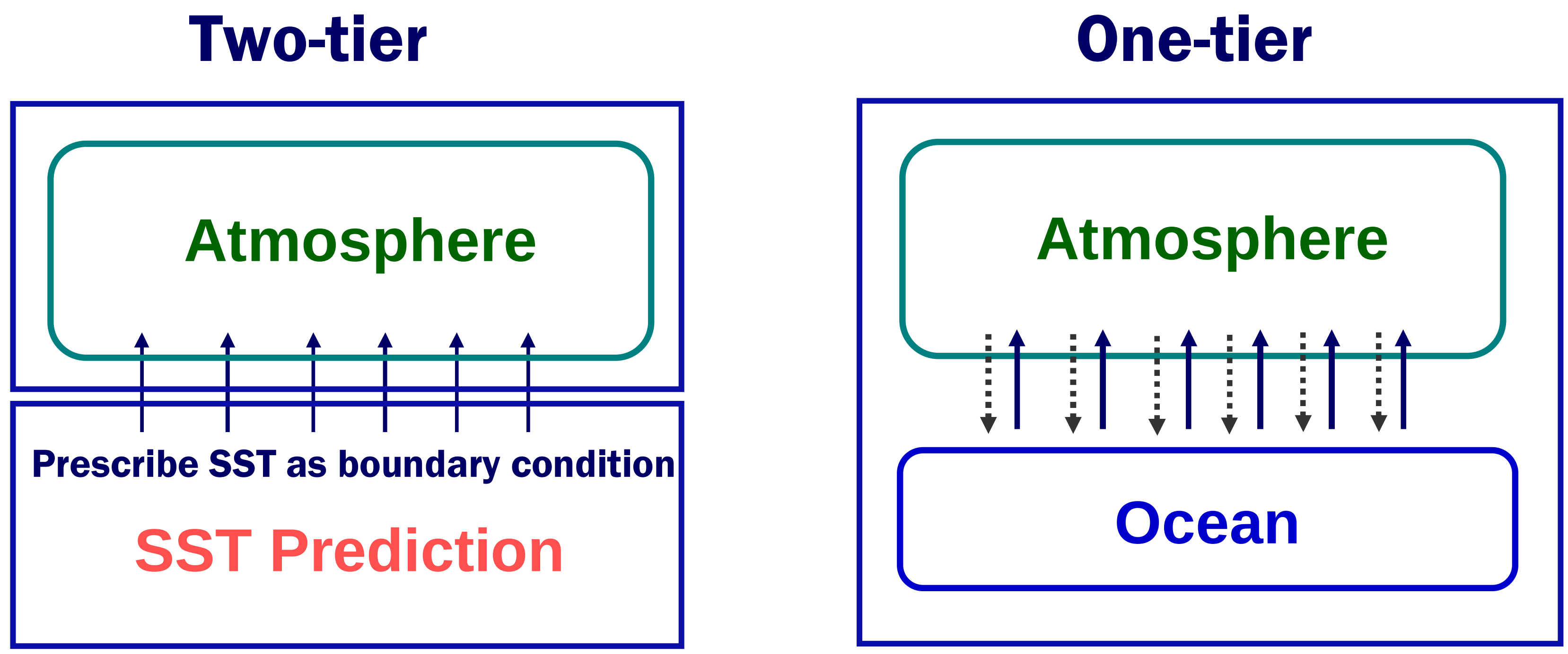


Only more than 95% significance level is shaded

→ **Atmosphere forces the ocean where the correlation coefficients between rainfall and SST show negative.**

Current activities of seasonal prediction

Climate Prediction System

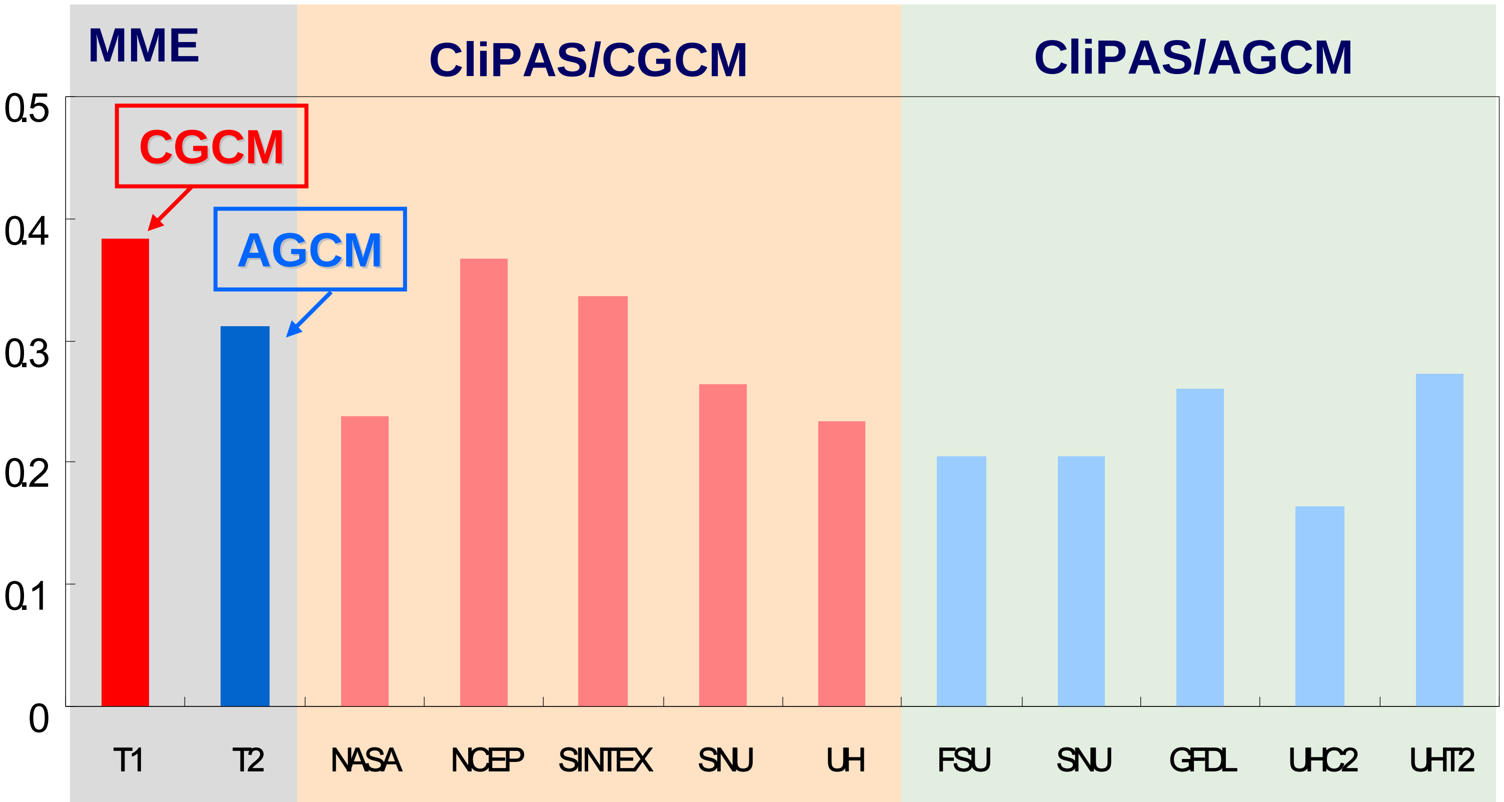


Key **SST prediction skill**

Coupling of atmosphere and ocean process

The state-of-the-art Climate Prediction

Global domain pattern correlation(60S-60N, 0-360)



Strategy of Prediction

The strategy of seasonal prediction is to

obtain "perfect signal" as close as possible.

(i.e. reducing variance of systematic error and variance of noise)

1. Reduction of Noise

- Averaging large ensemble members

(if number of ensemble members is infinite, Noise will be zero in the ensemble mean)

2. Correct signal

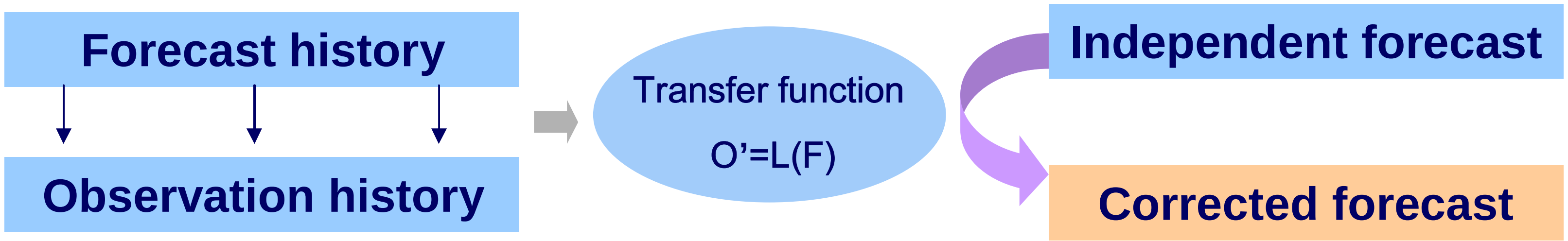
- Improving GCM
- Statistical post-process (MOS)
- Multi-model ensemble

Error correction

Multi model ensemble prediction

Noise dynamics

Correcting signal : Statistical Post process



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

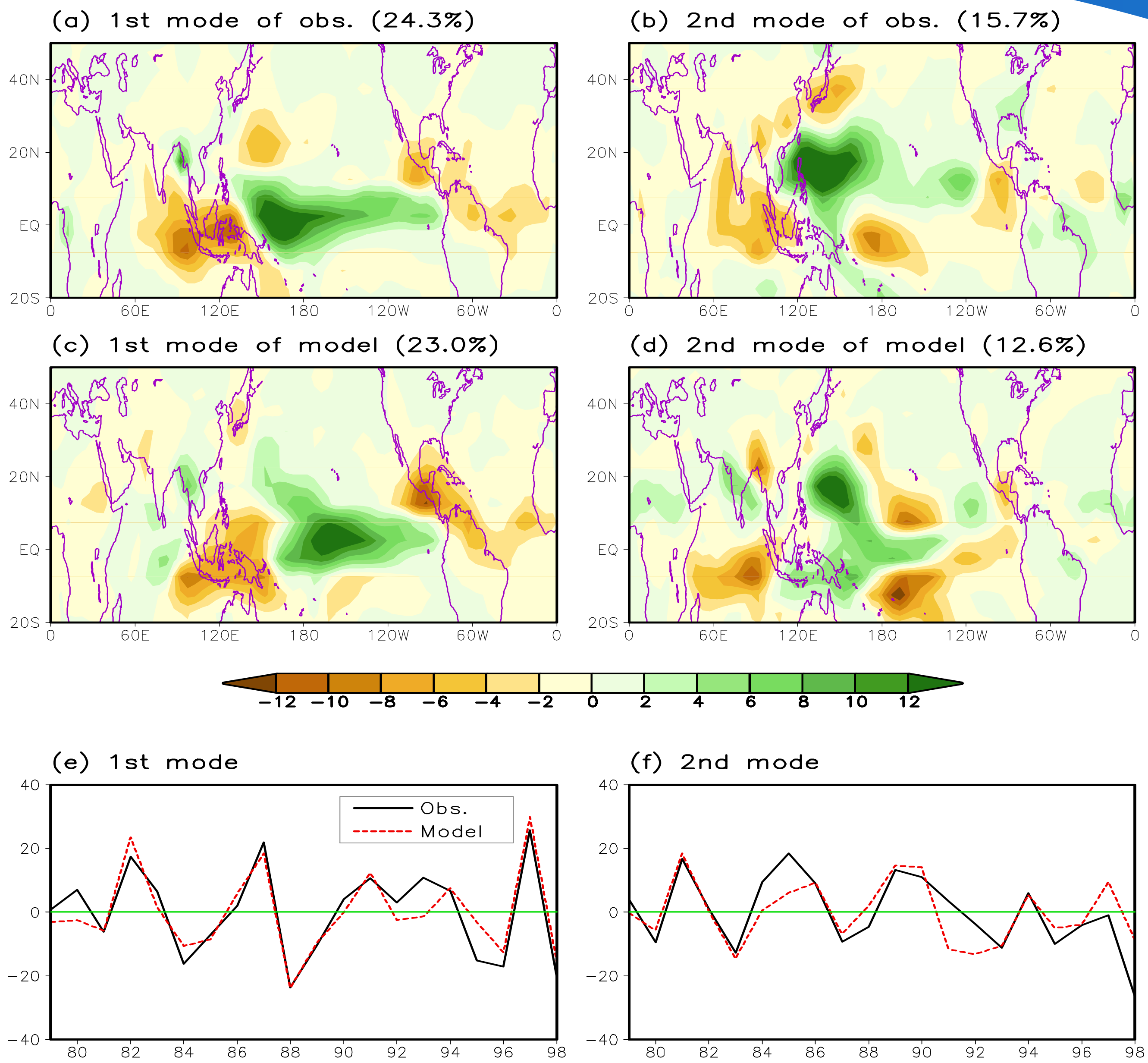
If statistics is not stationary, post-process will not work in independent forecast

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

Regarding actual constraints, available large ensemble forecast with well-tuned post process will be an appropriate strategy of seasonal forecast.

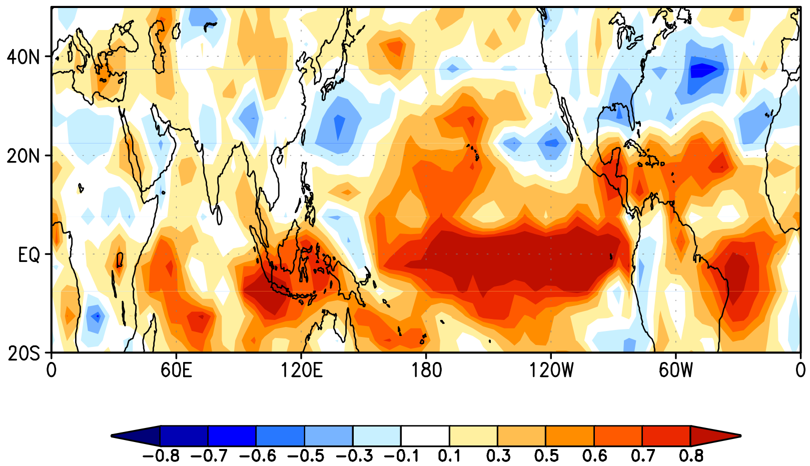
➔ Statistically optimized multi model ensemble prediction

EOF of Summer Mean Precipitation

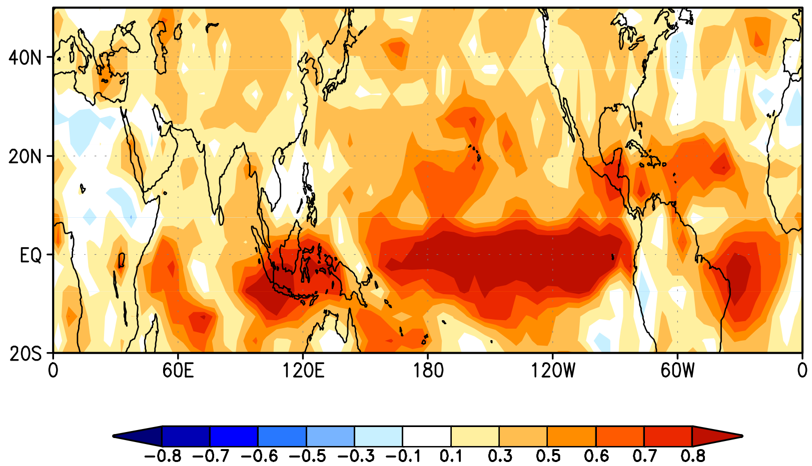


Correlation and Forecast Skill Score

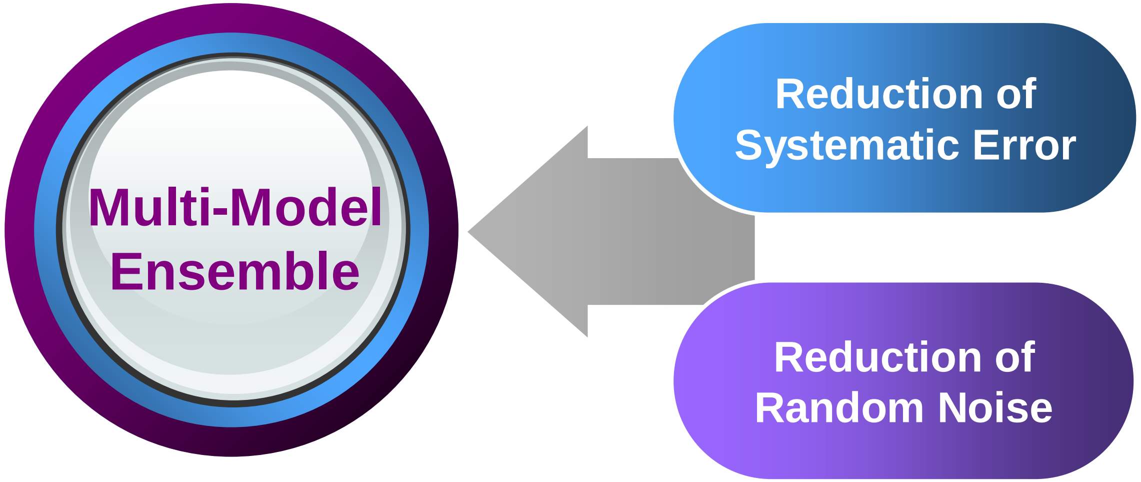
Before Bias Correction



After Bias Correction



Multi-model Ensemble Prediction



Cancellation of errors

: Multi-model

More samples

: Ensembles

Benefits of Multi Model Ensemble

$$x_i = y + e_i + \varepsilon_i$$

Forecast = True + Error + Noise

E_M : Error variance of Multi-Model Ensemble $E_M = \frac{1}{M}\left[\overline{V}(e) + \frac{2}{M}\sum_{ij} COV(e_i, e_j) + \overline{V}(\varepsilon)\right]$

E_S : Error variance of Single Model $E_s = \overline{V}(e) + \overline{V}(\varepsilon)$

[Example] *Two Models (M=2)*

Error variance of Multi-Model Ensemble $E_M = \frac{1}{2}\left[\overline{V}(e) + COV(e_1, e_2) + \overline{V}(\varepsilon)\right]$

Error variance of Single Model $E_s = \overline{V}(e) + \overline{V}(\varepsilon)$

Mean Square Error $E = E_M - E_s = \underbrace{\boxed{-\frac{1}{2}\overline{V}(e) + \frac{1}{2}COV(e_1, e_2)}}_{\text{Reduction of Systematic Error}} \underbrace{\boxed{-\frac{1}{2}\overline{V}(\varepsilon)}}_{\text{Reduction of Random Noise}}$

- **For Improvement ➔ E < 0**
- **Where Models are independent each other ➔** $COV(e_i, e_j) = 0$

Characteristics of each MME method

MME1

$$P = \frac{1}{M}\sum_i F_i$$

- simple composite
- equal weighting

MME2

$$P = \sum_i a_i F_i$$

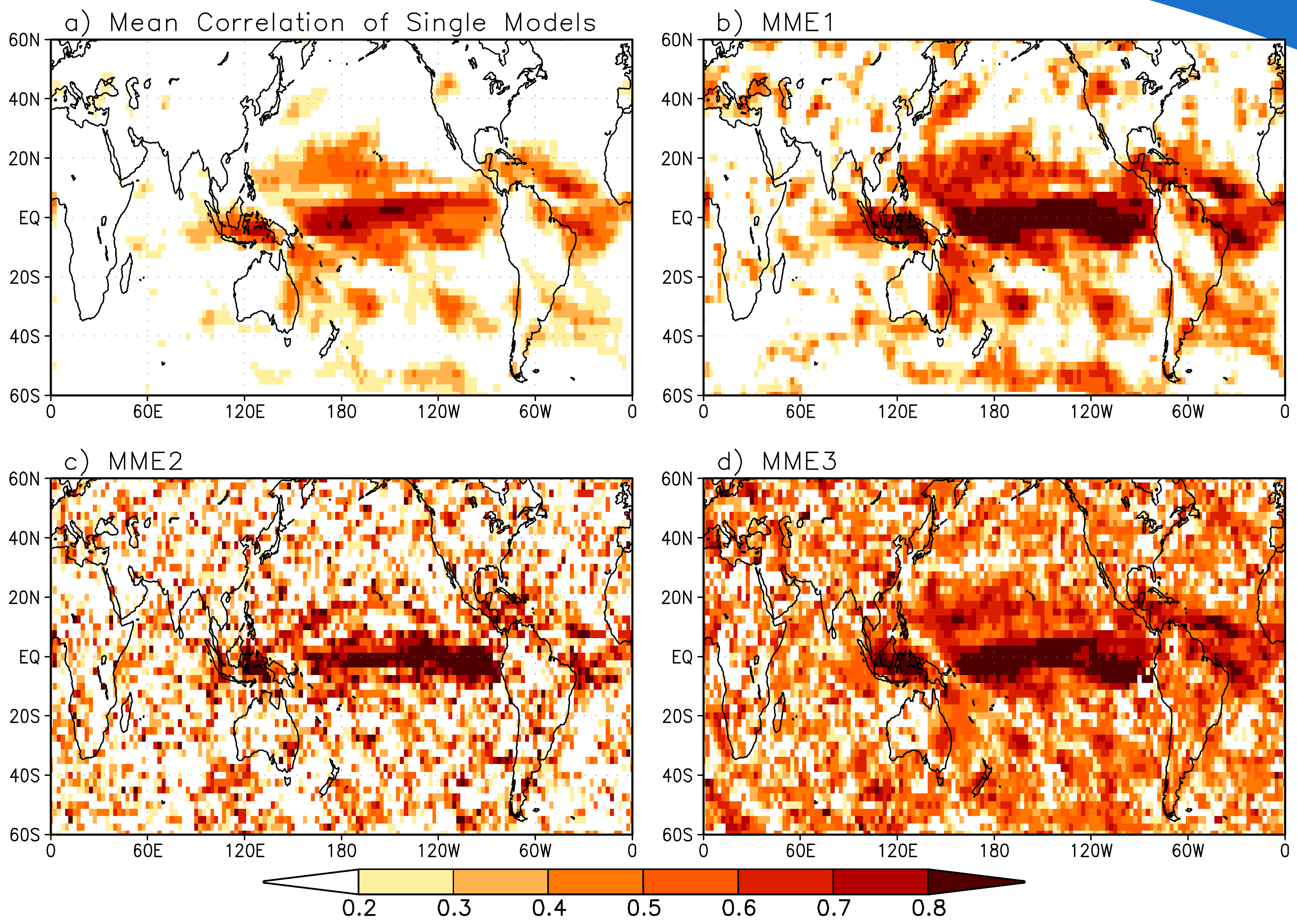
- superensemble
- Weighted Ensemble

MME3

$$P = \frac{1}{M}\sum_i \hat{F}_i$$

- simple composite after correction

Correlation Skill of MME



Combined and calibrated predictions of intraseasonal variation with dynamical and statistical methods

Targeted Training Activity, Aug 2008

What should we predict?

Previous studies

Different predictands

Statistical ISV prediction

Studies	Statistical Models	Predictand
Waliser et al. (1999)	SVD	Filtered OLR,U200
Lo and Hendon (2000)	EOF and regression	OLR, stream function
Mo (2001)	SSA and regression	Filtered OLR
Goswami and Xavier (2003)	EOF and regression	Rainfall
Jones et al. (2004)	EOF and regression	Filtered OLR, U200, U850
Webster and Hoyos (2004)	Wavelet and regression	Rainfall, River Discharge
Jiang et al. (2008)	Regression	RMM index, OLR, U200, U850

Dynamical ISV prediction

Studies	Dynamical Models	Predictand
Chen and Alpert (1990)	NMC/NCEP DERF (DERF- Dynamical Extended Range Forecast)	30-90d filtered OLR,U200
Lau and Chang (1992)		OLR, stream function
Jones et al. (2000)		Filtered OLR, U200
Seo et al. (2005)		OLR, U200, U850
Vitart et al. 07	ECMWF-MFS	RMM index

What should we predict?

Previous studies

Statistical ISV prediction

EOF, regression, wavelet, SSA, …

Forecast skill : 15 - 25 days

Dynamical ISV prediction

DERF-based model

Forecast skill : 7-10 days

Different predictands

Fair and rigorous reassessment is needed in real-time prediction framework

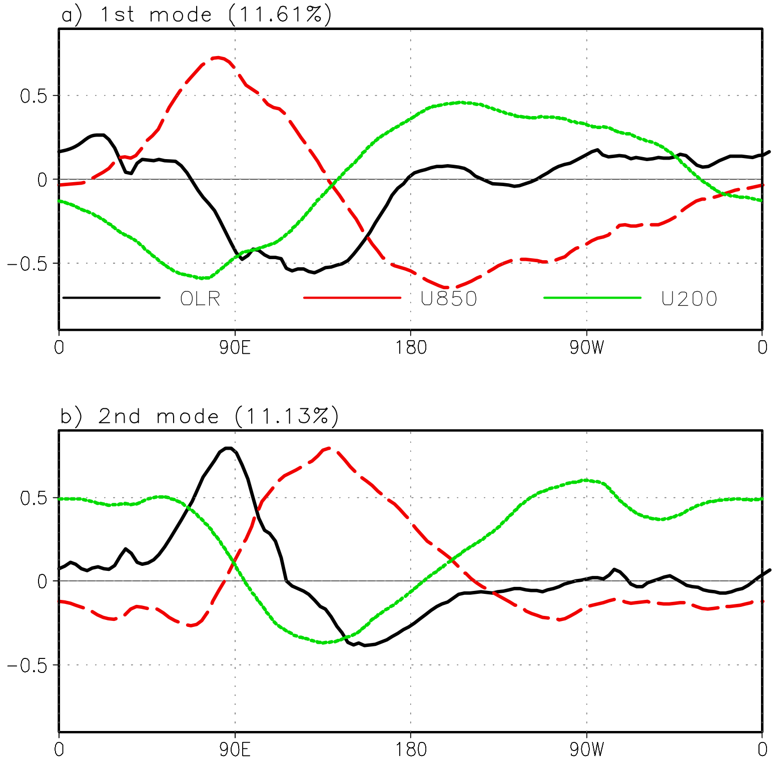
What should we predict?

Real-time Multivariate MJO index (RMM):

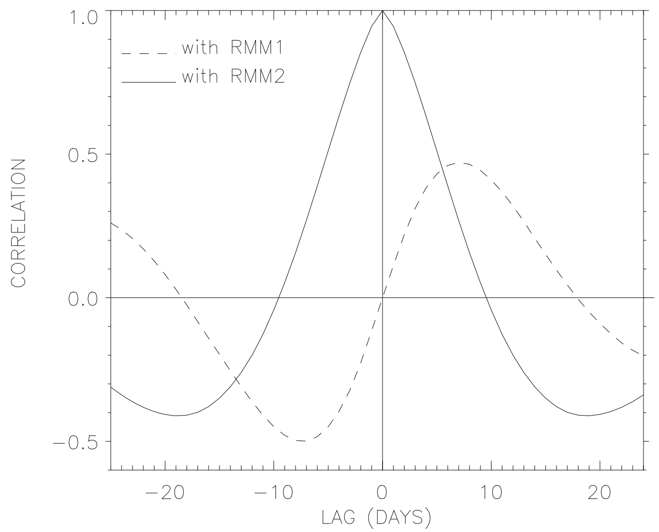
The PCs of combined EOFs (Equatorially averaged OLR, U850, U200)

(Wheeler and Hendon 04)

Combined EOF



Lag correlation: RMM1



1. Annual cycle removed;

2. Interannual variability (ENSO) removed:

- Regression pattern of each variable against the NINO3.4 index

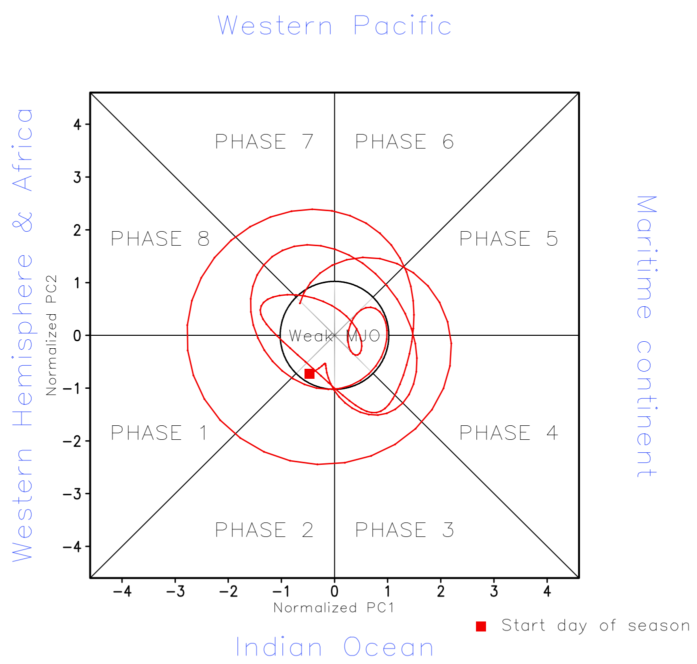
- Mean of previous 120 days

What should we predict?

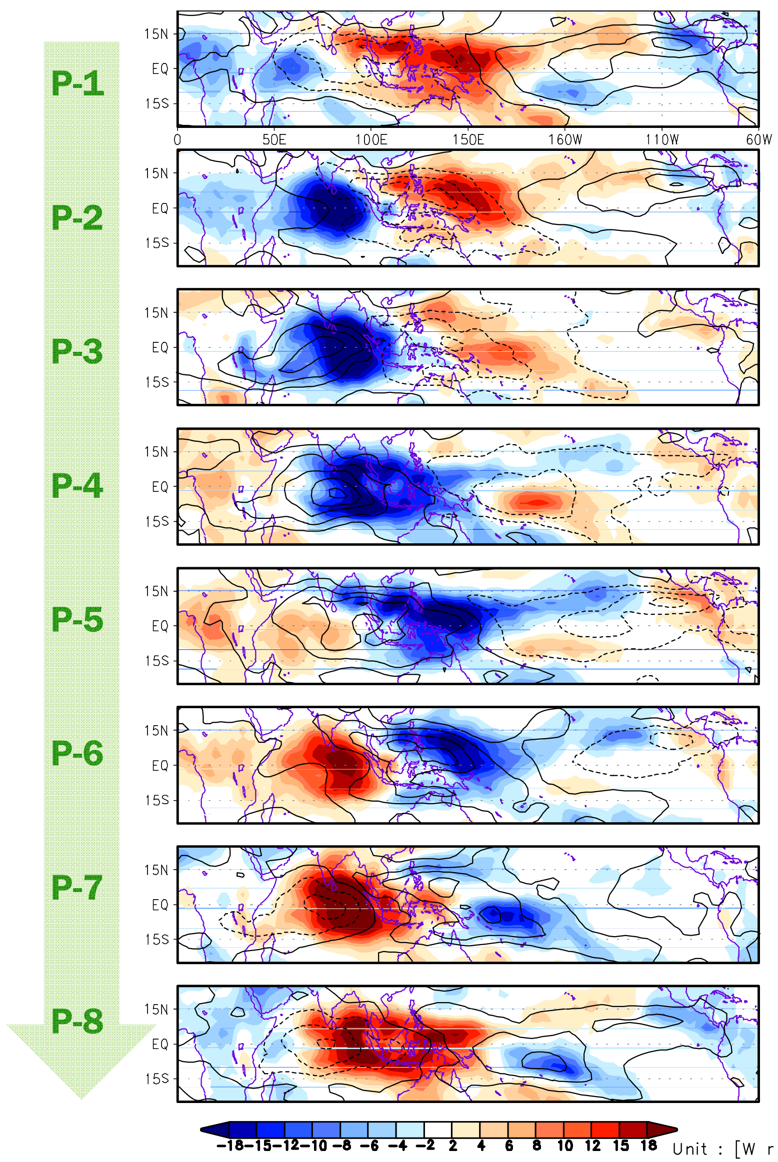
Advantages of RMM index

1. Avoid the typical Filtering problem in real-time use
2. Convenient for application (monitoring and prediction): Reduction of parameters
3. Represent the MJO in individual phase

Phase diagram (RMM1, RMM2): 1979 Jan-Dec



Composite: OLR & U850



Statistical prediction

- Multi regression model
- Wavelet based model
- SSA based model

Statistical model

Predictand: RMM index

Multi-regression — **Wavelet** — **SSA**

Multi-regression

Prediction of RMMs (regression)

$$PC(t_0+\tau)=\sum_{p=1}^{m}\sum_{j=1}^{\lambda}B_{pj}(\tau)PC_p(t_0-j+1)$$

$$j = j^{th}\ day\ earlier\ from\ t_0$$

$$\tau = lead\ time$$

$$\lambda = lag$$

$$m = PCs$$

$$B_{pj} = lag-regression\ coeff$$

Wavelet

Wavelet analysis

↓

Prediction of bands (regression)

↓

Reconstruction

$$W_n(s)=\sum_{n'=0}^{N-1}x_{n'}\left(\frac{\delta t}{s}\right)^{1/2}\psi_o\left[\frac{(n'-n)\delta t}{s}\right]$$

↓

$$R(t_0+\tau)=\sum_{p=1}^{m}\sum_{j=1}^{\lambda}B_{pj}(\tau)R_p(t_0-j+1)$$

↓

$$x_n=\frac{\delta j\,\delta t^{1/2}}{C_\delta\psi_0(0)}\sum_{j=0}^{J}\frac{\Re\{W_n(s_j)\}}{s_j^{1/2}}$$

SSA

SSA

↓

Prediction of PCs (regression)

↓

Reconstruction

$$PC_k(t+\tau)=\sum_{p=1}^{m}\sum_{j=1}^{\lambda}B_{pj}(\tau)PC(t+\tau-j)$$

↓

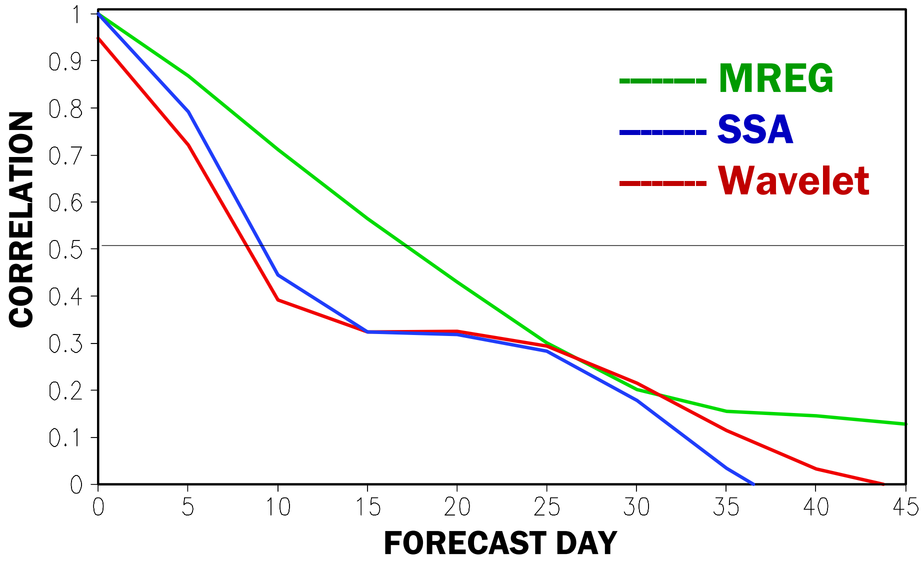
$$RC_k(t+\tau)=\frac{1}{M}\sum_{s=1}^{M}PC_k(t-s+1)EV_k(s)$$

Statistical model

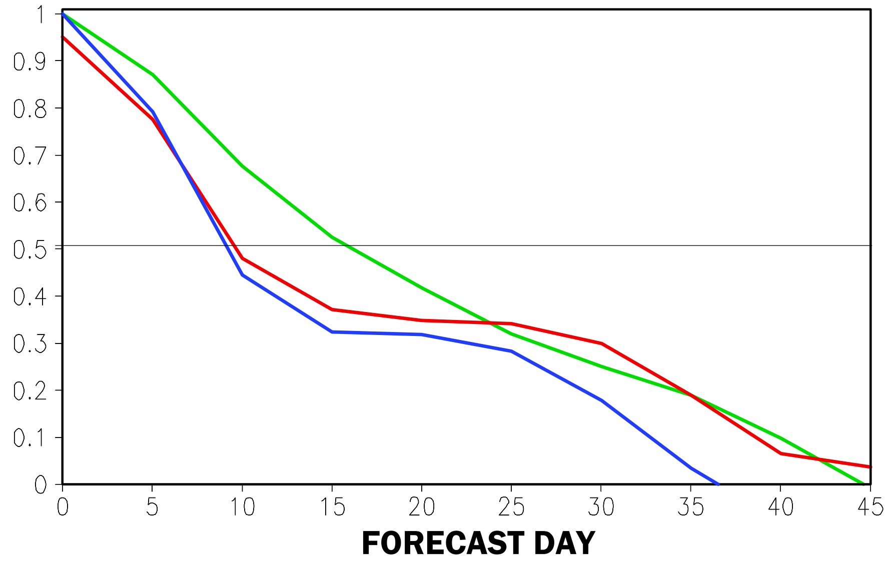
Multi-regression — **Wavelet** — **SSA**

Correlation 0.5 at (day)		
	RMM1	RMM2
MREG	16-17	15-16
Wavelet	7-8	9-10
SSA	8-9	9-10

RMM1



RMM2



Statistical model

Multi-regression: Downscaling

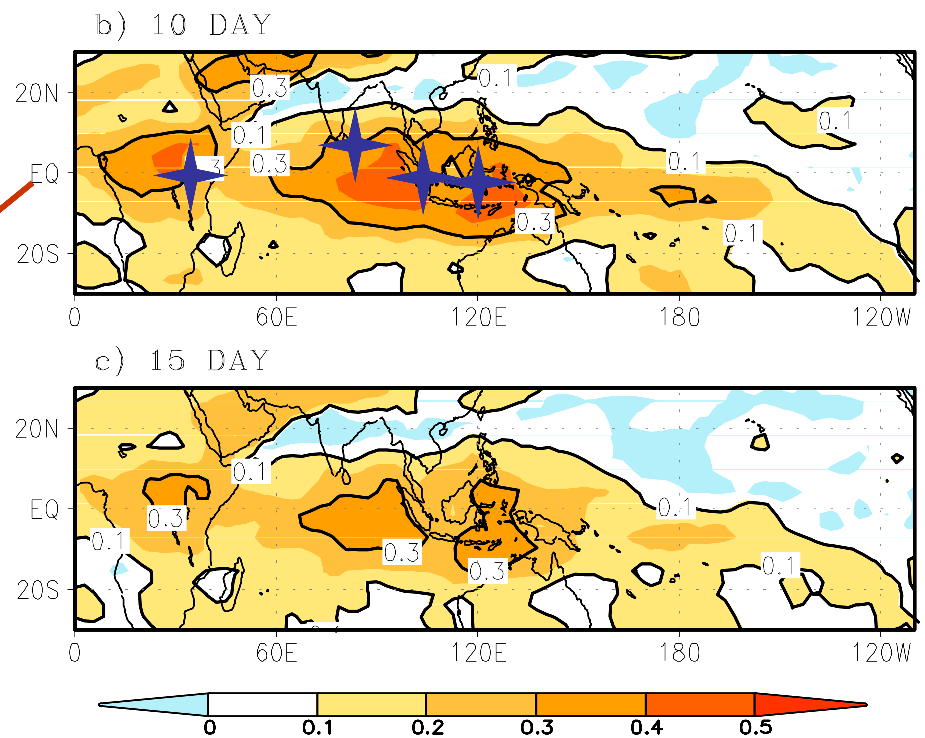
Downscaling to grids

$$X(t) = \beta_1 RMM_1(t) + \beta_2 RMM_2(t)$$

Regression coefficients can be obtained from historical data

Kenya (30E, EQ)
Sri-Lanka (80E, 5N)
Singapore(105E, EQ)
Indonesia(120E, EQ)

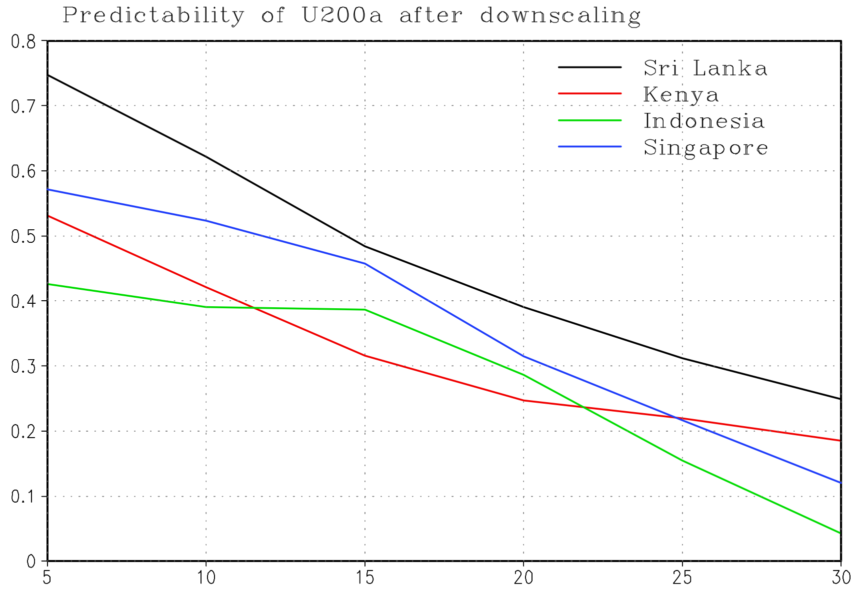
Predictability of downscaling results: unfiltered-OLRa



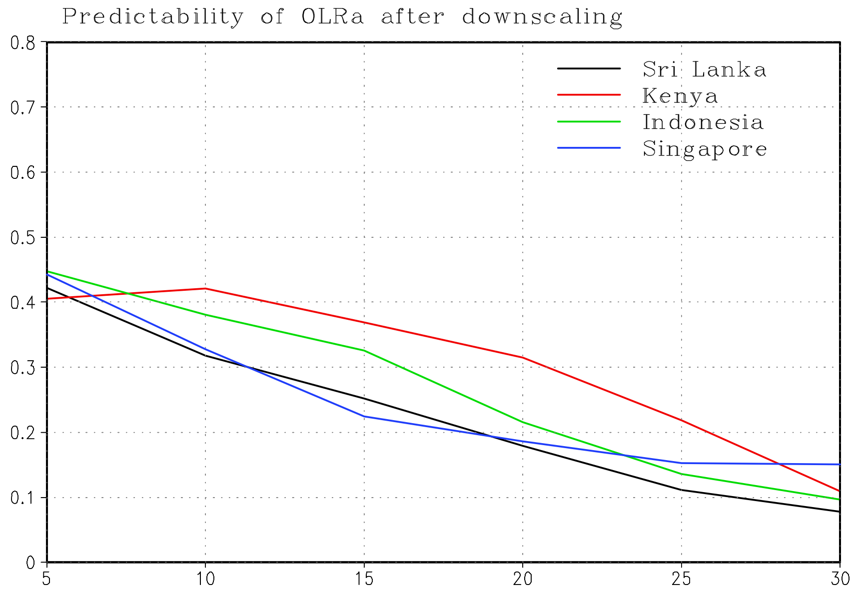
Statistical model

Multi-regression: Downscaling

Unfiltered U200 anomaly



Unfiltered OLR anomaly



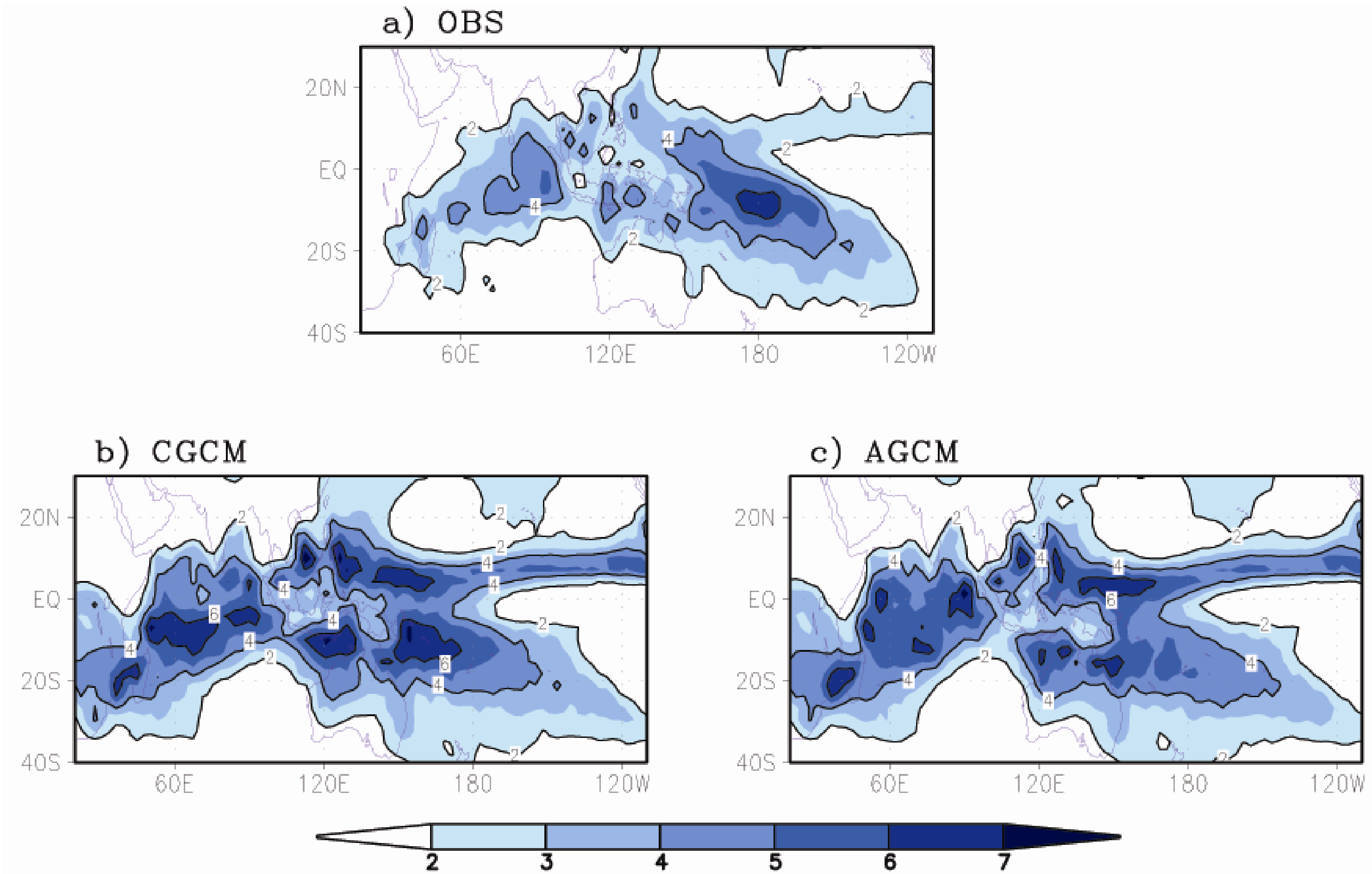
Dynamical prediction

- Simulation Performance
- Optimal Experimental Design
- Dynamical Predictability

Dynamical model

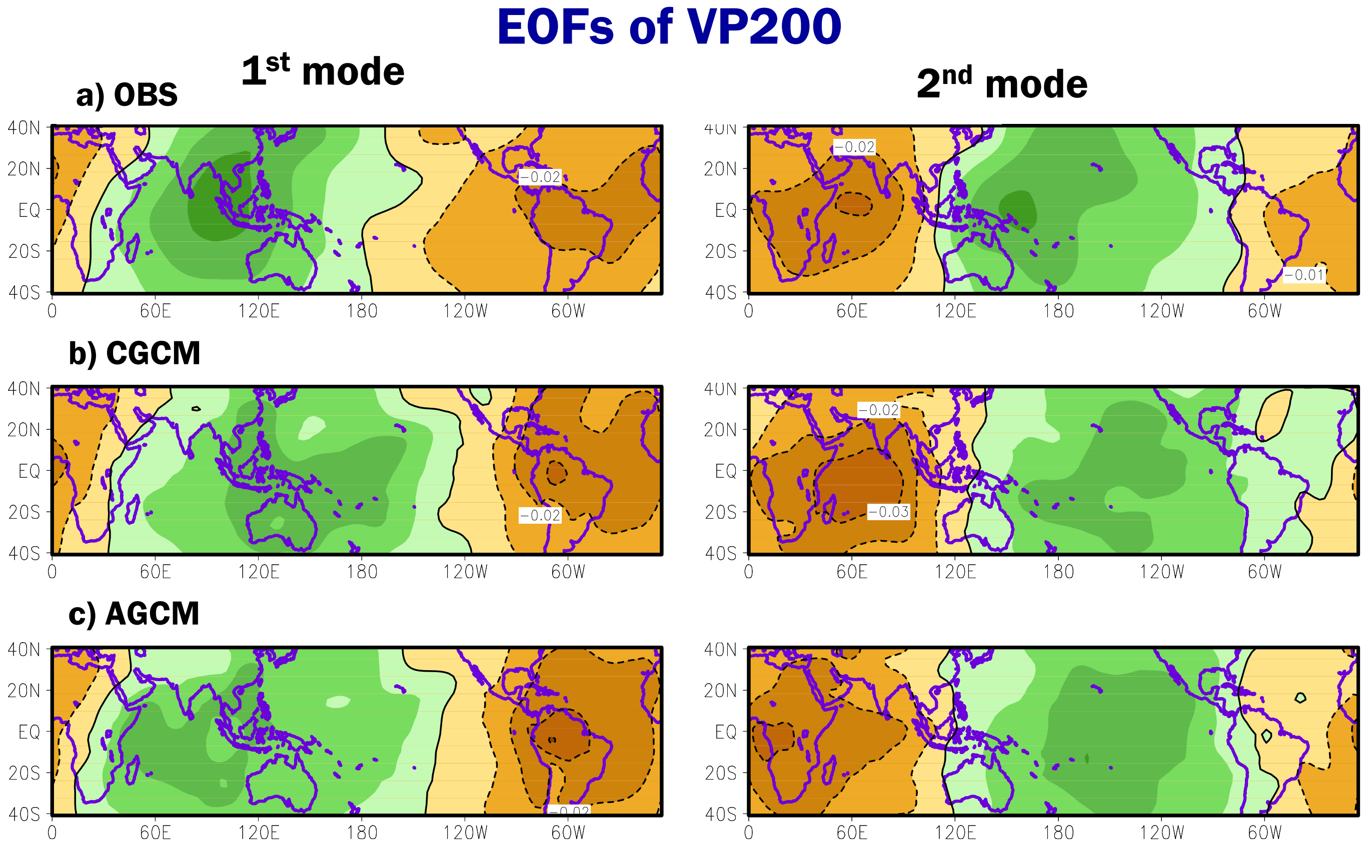
MJO simulation: Variability

Standard deviation of 20-70 filtered PRCP (1-30 day FCST)



Dynamical model

MJO simulation: Propagation

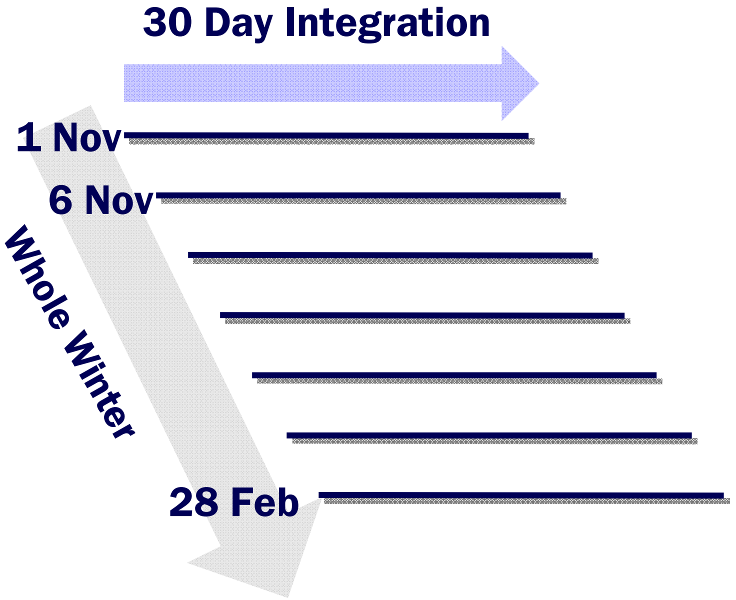


The observed two leading EOFs

- Eastward propagation mode
- Highly correlated between PC1 and PC2
- Two modes Explains more than half of the total variance

Dynamical model: Experimental design

Serial integration through all phases of MJO life cycle



Serial run > Seasonal prediction

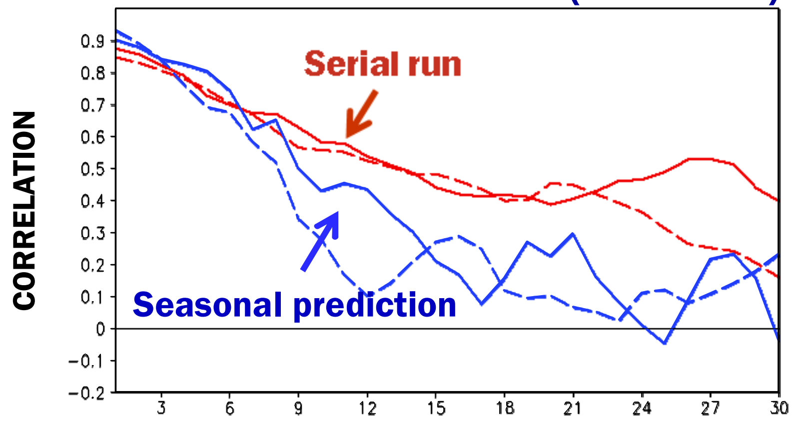
- Plenty of prediction samples
- Include whole initial phases

Does seasonal prediction work for MJO prediction?

Serial run with SNU GCM

EXP	Period	Total 30-day forecasts	Using 1-CPU
AGCM Long-term	27-year (79-05)	621	4 month
CGCM	8-year (98-05)	184	2 month

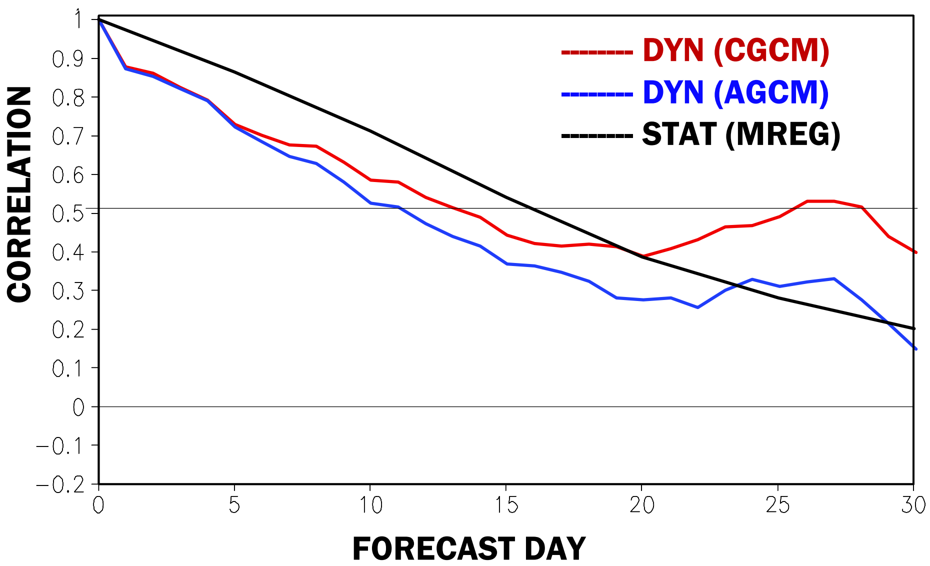
Forecast skill : RMM1 and 2 (SNU CGCM)



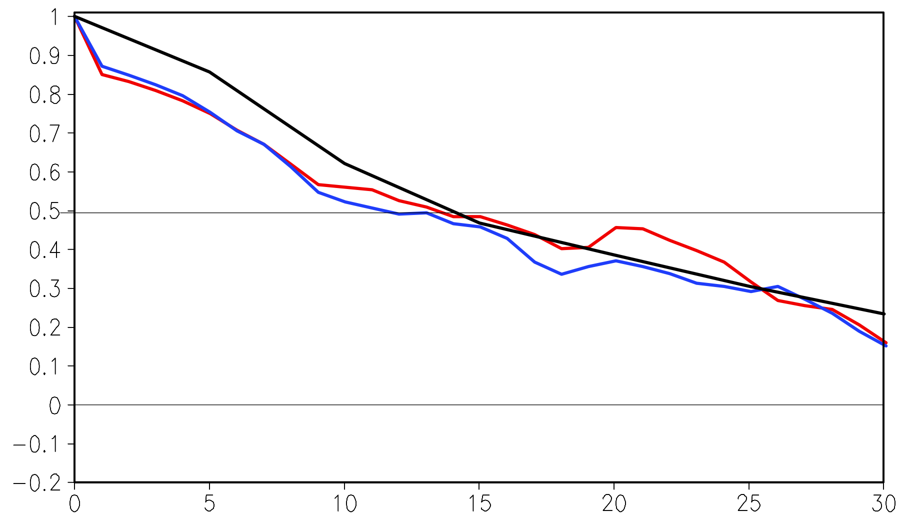
Statistical & Dynamical

Statistical vs. Dynamical prediction

Forecast skill: RMM1



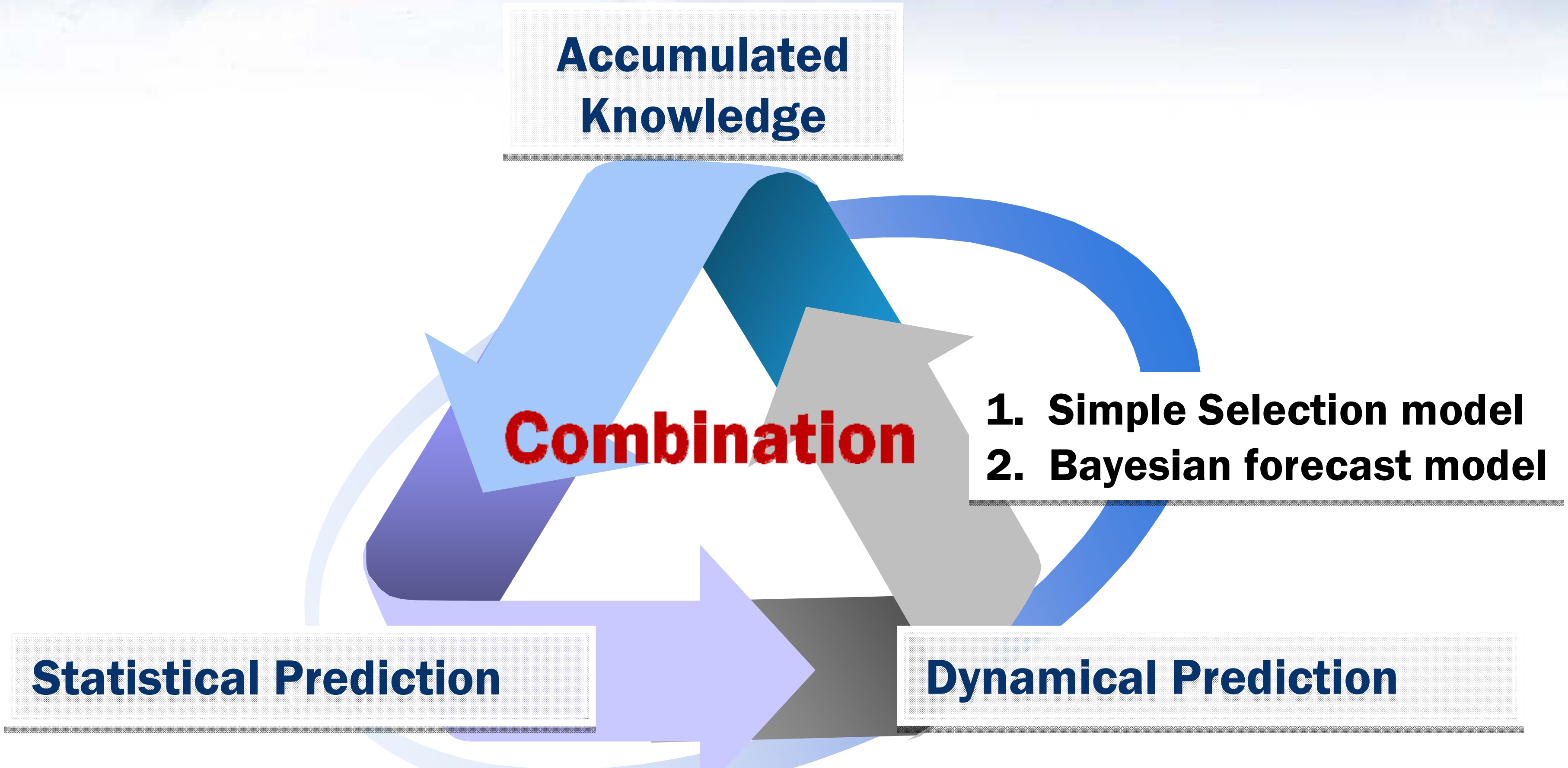
Forecast skill: RMM2



Correlation 0.5 at (day)		
	RMM1	RMM2
DYN (CGCM)	13~14	13~14
DYN (AGCM)	11~12	11~12
STAT (MREG)	15~16	11~12

- Comparable skill

Combination and Calibration

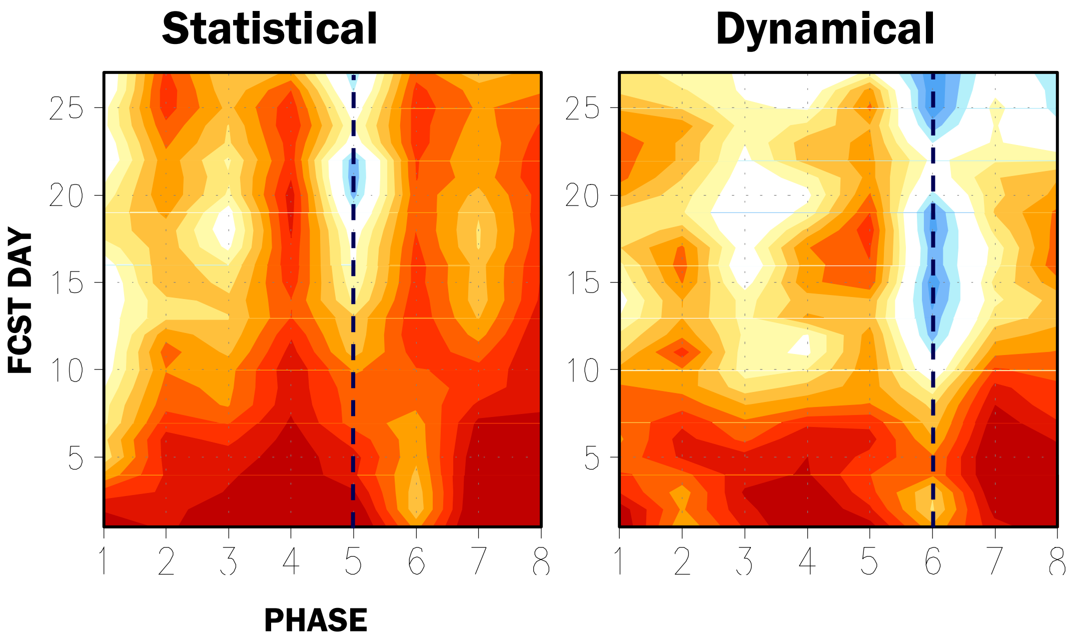


- Comparable predictability
- Independent predictions

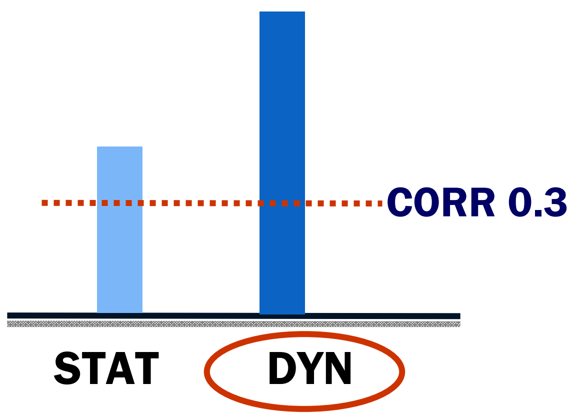
Combination: Selection model

Forecast skill of RMM1

Strong MJO



Selection process



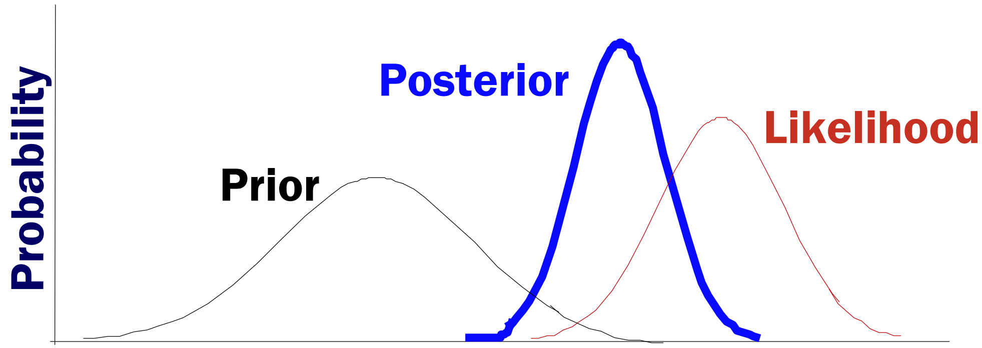
- More than 0.3: Better prediction
- Lesser than 0.3: Persistence

Combination: Bayesian forecast

Bayes' theorem

To construct a reliable data with combination of existing knowledge

$$\overbrace{p(S_f \mid D_f)}^{\text{Posterior}} = \frac{\overbrace{p(D_f \mid S_f)}^{\text{Likelihood}}\,\overbrace{p(S_f)}^{\text{Prior}}}{p(D_f)}$$



→ **Prior PDF is updated by likelihood function to get the less uncertain posterior PDF**

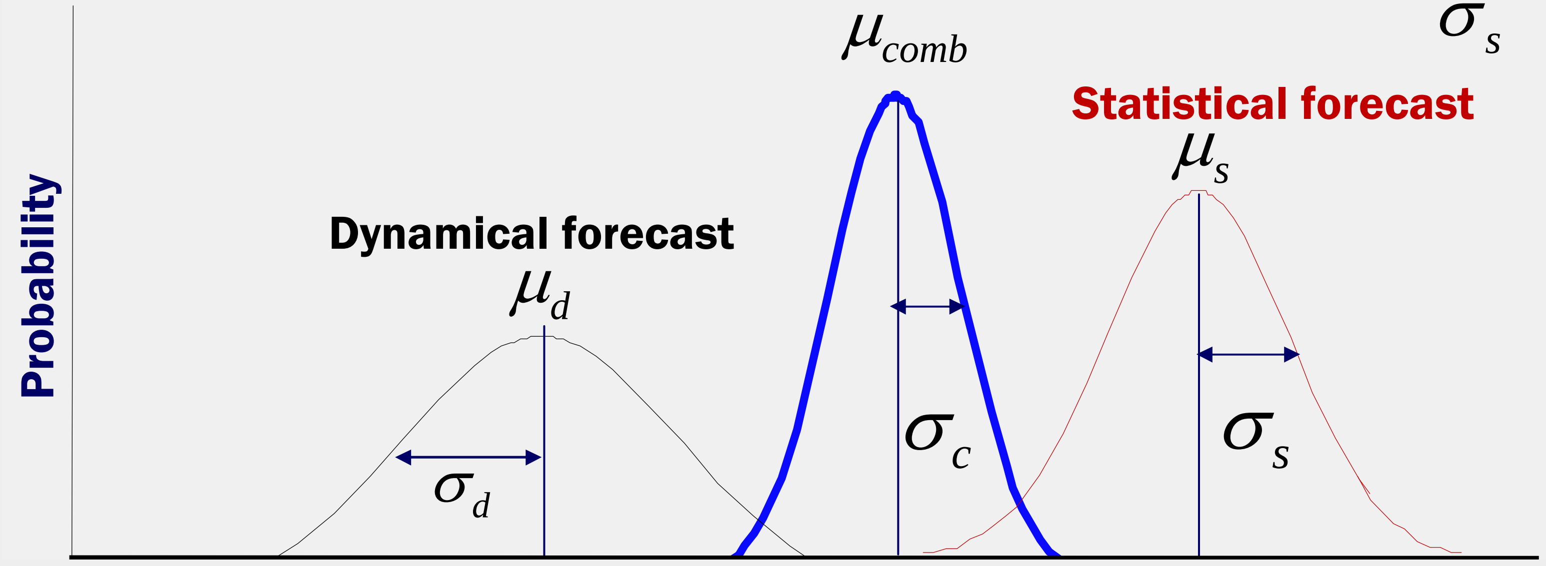
- Choice of the Prior: Statistical forecast (MREG)
- Modeling of the likelihood function:
  Linear regression of past dynamical prediction and on past observation
- Determination of the posterior

Combination: Bayesian forecast

Minimize the forecast error

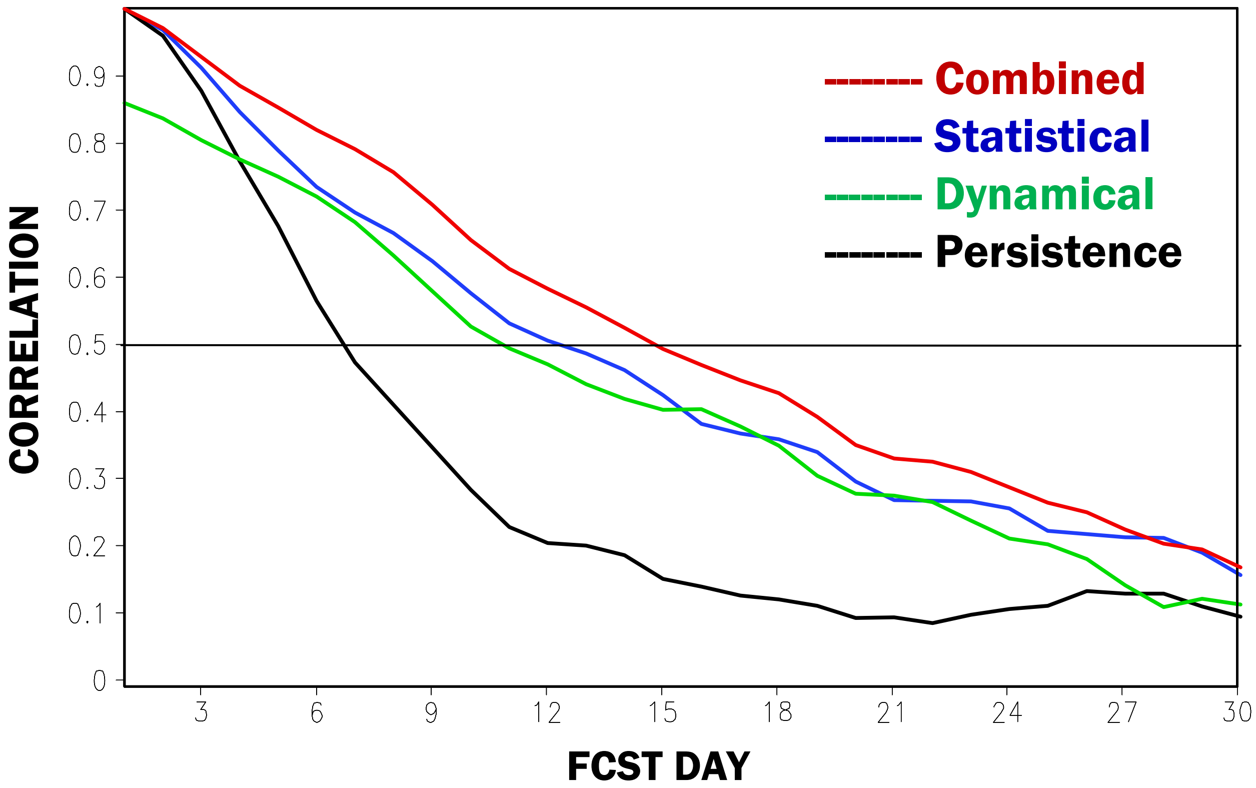
$$\Psi_{comb} = (1-K)\cdot\Psi_{stat} + K\cdot\Psi_{dyn}$$

Combined forecast→ $\mu_{comb} = \dfrac{\sigma_d^{\,2}\mu_s + \sigma_s^{\,2}\mu_d}{\sigma_s^{\,2} + \sigma_d^{\,2}}$



Combination: Bayesian forecast

Forecast skill of RMM1



Correlation 0.5 at (day)	
Combined	14~15
Statistical	12~13
Dynamical	11~12
Persistence	7~8

Improvement of forecast skill through combination by Bayesian forecast model
Improvement of forecast skill through combination by Bayesian forecast model

Other Important Issues

1. Initialization
2. Model improvement
   - Physical parameterization
   - High resolution modeling
3. Subseasonal (MJO) prediction

Initialization process of various institutes

Nudging 3DVAR OI method	Simple method Weighing function is constant
	Used in ECMWF, NCEP, IRI, GFDL, MRI, SNU, BMRC
4DVAR	Forecast errors in the assimilation window is minimized Tangent linear operator & adjoint operator is needed
Ensemble Kalman Filter	Evolving individual members with full model : Advantage for ensemble prediction
	Tested in GFDL

✓ Advantages
✓ Drawbacks

Initialization

Model improvement

Physical Parameterization
High Resolution modeling

Subseasonal prediction

Physical parameterization

Problem | Solution | Solution | Problem

Simplified Arakawa-Schubert cumulus convection scheme

- Loose convection criteria suppress the well developed large-scale eastward waves
- Minimum entrainment constraint
- Relative Humidity Criterion
- Unrealistic precipitation occurs over the warm but dry region

Impact of Cumulus Parameter

Relaxed Arakawa-Schubert scheme

- Absence of cloud microphysics
- Prognostic clouds & Cloud microphysics (McRAS)

Large-scale condensation scheme

Influence of Cloud-Radiation Interaction

- Cloud-radiation interaction suppress the eastward waves
- Layer-cloud precipitation time scale

Physical parameterization