Tell us your “weather” during a season

1. What happens?
2. Irregularity?
Q: What climate phenomena are you familiar with?
Climate Change

• Changes in our expectation
Warm pool

- Climate Engine: remember “mean” feature
El Niño, ENSO

- A Big Ocean Swing
ENSO impact
Extratropical LFV (SAM, PSA)

Irving and Simmonds (2016)
How they change weather?
Seasonal Prediction (1)
: Introduction/Predictability

Jin Ho Yoo
APEC Climate Center
Overview

• Predictability
• Methods
• Verification + Downscaling
• Operation
Climate prediction
Climate
Climate is what we expect,
Weather is what we get
Climate = Expectation
Climate Change = We need to change our expectation
Climate prediction = Expectation of Expectation

How uncertain!
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

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

\[
\vec{X}(t_0 + \tau) = \vec{X}(t_0) + \int_0^\tau F(\vec{X}(t), a(t)) dt
\]
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
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 30 days later*
3. *Temperature of this room in 30 years later*

**Lead time**($\tau$)
$X(\text{initial}) = 1.$

$X(\text{initial}) = 1.00001$
Predictability

Depends on *what to predict*

**Prediction of**
1. Temperature of Seoul (Korea)
2. Temperature of Jakarta (Indonesia)
3. Temperature of Villa Las Estrellas (Antarctica)

**Location**
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
Climate prediction

How long?

Time mean of weather
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)

Summer mean prec.

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

Primary seasonal weather statistics: seasonal mean

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

\[ X = X_s + X_n \]
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)
Mechanisms of Variability

**Internal**

- **Weather:**
  1. Internal Dynamics of Atmosphere

- **Climate (seasonal-decadal):**
  2. Internal Dynamics of Coupled Ocean-Land-Atmosphere

- **Climate Change:**
  3. Internal Dynamics of Sun-Earth System

**External**

- Boundary Condition of SST, Soil wetness, Snow, Sea ice, etc.
- Solar, Volcanoes
- 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

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

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

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

\[
\frac{\partial \theta}{\partial t} + \bar{u} \cdot \nabla \theta = l
\]
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
Multi Model Ensemble Forecasting

- Run with many models

Which one??
Use all!

Pattern correlation: summer monsoon precip.

\[ P = \sum_{i} a_i F_i \]

0.32 (indv.)
0.44 (MME)
Predictability of Multi Model Ensemble

Correlation skill of a single model

\[ R_i = \frac{x y_i}{\sqrt{V(x)V(y_i)}} \]

Correlation skill of MME

\[ \langle y \rangle = \frac{1}{M} \sum_{i=1}^{M} y_i \]

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

\[ \langle R \rangle = \frac{1}{M} \sum_{i} R_i \]

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

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

\[ \langle r \rangle = \frac{1}{M^2} \sum_{i} \sum_{j} \frac{y_i y_j}{V} \]

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

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

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

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)
APCC MME (TCC)

Rainfall (JJA)

SLP (JJA)
DJF season

Rainfall (JJA)

ROC score
Rainfall (DJF)
APCC operational forecast

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

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

Precipitation for December 2015-February 2016

Anomaly fields are displayed

© APEC Climate Center
Models have biases

Modeled sea surface temperature and bias

SST

Orography

Saurrel et al. (2015)
EOFs of Summer Mean Precipitation

(a) 1st mode of obs. (24.3%)

(b) 2nd mode of obs. (15.7%)

(c) 1st mode of model (23.0%)

(d) 2nd mode of model (12.6%)

(e) 1st mode

(f) 2nd mode
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)
Weakness: overfitting

Consider potential predictability

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

Forcing 1

Forcing 2

Forcing 3

Local Large scale circulation (L)

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
LARGE SCALE PATTERN ASSOCIATED WITH RAINFALL

Local large circulation and Teleconnection

One Point Correlation map with seasonal mean local rainfall with other variables
RELATIONSHIP WITH ENSO

corr. btw. station prec. and ONI
Seasonal Prediction (3) : Evaluation and Downscaling

Jin Ho Yoo
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)
    • $MSE = \frac{1}{N} \sum_i (F_i - O_i)^2$
  – ACC (Anomaly correlation, Pattern), TCC (Temporal correlation)
  – MSSS (Mean Squere 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

Diagram:
- OBS (Observation)
- RMSE (Root Mean Square Error)
- Correlation
- FCST (Forecast)

APCC Climate Center Logo
Probabilistic forecast

• **Brier score (Brier Skill Score)**
  – MSE of prob. forecast
  
  \[ BS = \frac{1}{N} \sum_i (F_i - O_i)^2, \text{ F}=1/0, \text{ O}=1/0 \]

• **BSS (Brier skill score)**
  – \[ 1 - \frac{BS}{BS_c}, \]
  – \( E : \text{ error/penalty}, \ Ec : \text{ error of reference forecast} \)
Probabilistic forecast (Categorical)

- Reliability curve
Probabilistic forecast (Categorical)

- **ROC (Relative Operating Characteristics)**

<table>
<thead>
<tr>
<th>F</th>
<th>O</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td>Hit (H)</td>
<td>False Alarm (F)</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>Miss (M)</td>
<td>Correct Rejection (C)</td>
</tr>
</tbody>
</table>

- **HR (Hit rate)** = $\frac{H}{H+M}$
- **FAR (False Alarm rate)** = $\frac{F}{F+C}$
  - Good forecast: HR↑, FAR ↑
Probabilistic forecast (Categorical)

- ROC (Relative Operating Characteristics)
Probabilistic forecast (Categorical)

- **HSS (Heidke Skill Score)**

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>O</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Hit (H)</td>
<td></td>
<td></td>
<td>False Alarm (F)</td>
</tr>
<tr>
<td>No</td>
<td>Miss (M)</td>
<td></td>
<td></td>
<td>Correct Rejection (C)</td>
</tr>
</tbody>
</table>

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

• \( HSS = (score - score by chance) \)
• \( / (perfect score - score by chance) \)
Forecast economic value

\[ V = \frac{E_{cli} - E_{fore}}{E_{cli} - E_{per}} \]

V=1 : perfect forecast

V=0 : climatological forecast

\[ E_{fore} = (h + f)C + mL \]

\[ E_{per} = hC = \bar{C}C \]

• When the forecast is perfect, \( f = m = 0 \) and \( h = \bar{C} \). Then,

\[ E_{per} = hC = \bar{C}C \]

• When the forecast is climatology. The only one kind of action will be kept.

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

\[ E_{cli} = \min(C, \bar{C}L) \]

\[ V = \frac{\min\left( \frac{C}{L}, \bar{C} \right) - (h + f) \frac{C}{L} - m}{\min\left( \frac{C}{L}, \bar{C} \right) - \frac{C}{L} \bar{C}} \]
Value of Probabilistic forecast (Above normal) : GCMs

(a) Monsoon (40E-160E, 20S-40N)
(b) ENSO (160E-280E, 20S-20N)
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!
DJF season

Rainfall (JJA)

ROC score
Rainfall (DJF)
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?”
• BUT we can “COMPARE” diff. things
Seasonal Prediction (4) : Operation and discussion

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 loose 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) : 40-50 day oscillation (30-60 days ISO)

Eastward moving large scale convective anomaly along the equatorial baroclinic structure

(precipitation anomaly is predominant in Indo-Pacific sector)

It can be a predictability source of extended range forecast in the tropical region.
Madden and Julian (1972)

Like negative EOF 2
Like positive EOF 1
Like positive EOF 2
Like negative EOF 1
(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)
BSISO (Boreal Summer ISO)

The canonical northward propagating component

The AMS pre-monsoon and onset component

BSISO forecast (May-Oct)
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

http://s2sprediction.net
New type of El Nino
El Nino Modoki (Central Pacific El Nino)
Fig. 2  

(a) Composite observed significant SST (°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
So, what are you going to do with CLIK

• Hope you were able to find suitable predictors for your locations
  – Yes : produce forecast
  – No : try more! (It is important to understand the large scale circulation that affects local weather)

• Once you’ve got a forecast, you need to combine them with other informations.
  – It’s an area of “art” at this moment
1. What is the most visible (potentially important) climate fluctuation: La Nina
2. Is the La Nina coming? Is it going to be Strong?
3. How did La Nina change the weather statistics before?
4. What is the forecast from MODELS? How much are they reliable? If it is not reliable, do we have “calibrated” forecast from them??
5. What is your conclusion???
Suggestions?
Questions?

Thanks
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
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
The most important thing you need is, 

Patience