Lecture 4.
The Basics of Downscaling Methodology

Yun-Young Lee
7 Sep 2016
Why downscaling is required?
1. Low skill of dynamical seasonal forecast...

- near surface temperature...

TCC (NNR2 vs SCM), t2m
1. Low skill of dynamical seasonal forecast...

- **Rainfall?**
- **tropical region only?**
- **South America?**

**TCC (NNR2 vs SCM), prec**
1. Low skill of dynamical seasonal forecast...

- In particular stations...
1. Low skill of dynamical seasonal forecast...

- *In particular stations...*
Peru & Chile

- **Pacific Ocean** to the west and the **Andes mountains** to the east running parallel to the Pacific Ocean
- Three regions: the costa (coast), the sierra (highlands), and the selva (jungle).
- Ocean climate vs Mountain climate?
2. Climate Locality

- Complicate climate of Peru

Peru map of Köppen climate classification

Can 2.5 by 2.5 grid pixel simulate steep terrain effect therefore locality of station climate? Not really.
2. Climate Locality

How different is the climate between 2 adjacent stations in the east of Arequipa? But, they are in one grid! OMG!!!
2. Climate Locality

How different is the climate between 12 adjacent stations near Arequipa? But, they are in one grid! OMG!!!
2. Climate Locality: annual rainfall
What is downscaling?
Downscaling
Average face (1)

Yoo-Jin

Yun-Young

Inja

Hyun-Rok

Jin-Ho

Koreans
Average face (2)
Average face (3)
In seasonal prediction, how to draw the information of a particular “point” from COARSE grid model data?

We need additional information to downscale!

How to overcome?
- Resolve the (dynamical/physical) process determining point value
- Find a relationship between the point and large scale field
What is dynamical downscaling?

• Simply, it is running RCM.
  BC from GCM, IC → solving dynamic equations!
Statistical downscaling

- Based on statistical relationship between precipitation/temperature at particular stations and adjacent/remote upper level atmospheric fields

→ Developing simple downscaling model
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).
Accuracy of Large Scale Information?

Simulated Large scale information is correct? **NOPE!!!**

Downscaling = Statistical Post-process
Why statistical downscaling rather than dynamical downscaling?
Dynamic downscaling is costing...

**Dynamical downscaling**

- **High performance computing system**
  - computation to solve dynamic equations

- **Fast Network system**
  - boundary condition from GCM

- **Data Storage**
  - archiving high-resolution spatial outputs

- **Well-educated human resources**
  - operation
  - maintenance
  - Interpretation of prediction
Statistical downscaling forecast

✓ No HPC needed
✓ No huge data storage needed
✓ Less dependent on internet speed
✓ Easy to interpret (simple but based on dynamical relationship)
Statistical downscaling forecast 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
Regression concept!

Scatter plot & least square fitting!

\[ Y = a^*x + b \]

\( x \): predictor (ONI)
\( y \): predictand (rainfall)
Station to LSMP relationship

sibayo [FMA]

Domain candidates: ①②③: monopole
④: dipole

+ area: ESP, part of SPCZ (domain ③)
SST↑ → PREC. ↑

- area: equatorial central to eastern Pacific (domain ①②)
SST↓ → PREC. ↑
Weakness: overfitting

Consider potential predictability

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

from Scott Fortmann-Roe 2012
CLIK downscaling

A way to localize existing coarse climate information

CLIK downscaling is mainly based on station to Large Scale Meteorological Field (LSMF) relationship. \((Y = a*X + b)\) By utilizing the simulated LSMF \((X, \text{predictor})\), CLIK estimates seasonal mean precipitation/temperature \((Y, \text{predictand})\) at specific station.

**Dynamical fcst.**

- Simulated LSMF from Individual models

**Statistical dwnscl.**

- Station to LSMF regression relationship

- Hypothesis: The station to LSMF relationship is well replicated in individual models

- Predictor: LSMF
- Predictand: Prec. at specific station

A kind of hybrid system for point-wise seasonal forecast
Predict yield of Greek bonds with number of Facebook users

Is it appropriate?

If yes, why?

If not, why?

From business week
Predict global average temperature with Carbon dioxide concentration

Is it appropriate?

Image Source: images.iop.org

If yes, why?
If not, why?
The most important things...

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

   ➔ Finding predictors (large scale meteorological patterns (circulations) associated with local prcp/temp of your station)

2. whether GCM (MME) is able to reproduce those patterns/relationship?

   ➔ **Applicability** of downscaling
Example of Domain Selection
Maritime Continent: Indonesia (1)

Weakening of Mean monsoon flow (southwesterly)

-> Increase SST + Rainfall
Maritime Continent: Indonesia (2)

SON [Jakarta]

La Nina (and negative IOD) signature
West South America: Sibayo

sibayo [FMA]

NNR2 prec

SCM prec

NNR2 sst

SCM sst

La Nina signature
CLIK downscaling procedure

- **Screening process**
  - Significant test
  - Pattern correlation

- **set Var. & Domain**
  - correlation map
  - regional climate understanding

- Downscaling from regression *procedure* over predictor domain

- The most important thing, *Test! Test! Test!*

- Pass

- Fail

- **Success**
Thank you
2. Climate Locality: seasonal cycle
Our final goal

: Predicting next season climate at particular station point (your city, town, village)

✓ What we have
  : Prcp/temp station data (past) & coarse GCM data (past & future)

✓ What we know
  : past relationship (station point –large scale field)
Accuracy of Large Scale Information?

Simulated Large scale information is correct? NOPE!!

Under this circumstance,

Statistical downscaling can be understood as a statistical “forecast” based on “forecasted” large scale information.
Uncertainty in climate prediction

Increase Greenhouse effect due to CO2 increase

Image source: http://www.slideshare.net/ShrutiPithadia/ppt-on-global-warming
Maritime Continent: East Timor

Difficult but Some interesting signal at equatorial Central Pacific (El Nino Modoki?)
Central Pacific: Fiji

Fiji, 91680, Nadi Airport [JFM]

Prec. Anom. down. from v850

Prec. Anom. down. from s1p

Prec. Anom. down. from v850

Prec. Anom. down. from s1p

corr. coef. = 0.82

corr. coef. = 0.72
Should remind!

1. Based on *past* relationship. It doesn’t guarantee that it works in present/future.

2. Can be *subjective* (dependent on *human*)
   - Which variable is the best?
   - Which domain is the base for predictor?

*Stationary statistics* & *physical understanding* should be based.