Statistical Seasonal Forecast Methods at the Climate Prediction Center

Peitao Peng

CPC/NCEP/NWS/NOAA, USA

Acknowledgements: H. van den Dool, D. Unger, P. Xie

Outline

- Seasonal forecast methods used in CPC
- Some statistical methods in detail
  - Optimal Climate Normals (OCN)
  - EOF adjusted OCN (EOCN)
  - Constructed Analogue (CA)
  - Forecast tool consolidation

Seasonal Forecast Methods in CPC

1. NCEP Climate Forecast System (CFS), a fully coupled dynamical model;
2. Canonical Correlation Analysis (CCA) (Barnston 1994, He and Barnston 1996);
3. Ensemble CCA (ECCA) (Mo 2004);
4. Screening Multiple Linear Regression (Unger 1996);
5. Markov Model (for ENSO) (Xue et al. 2000).
6. Optimal Climate Normals (OCN) (Huang et al. 1994);
7. EOF adjusted OCN (EOCN) (Peng and van den Dool 2002);
8. Constructed Analogue (van den Dool 1994,2003);
9. Forecast tool consolidation (Unger 1996, Peng et al. 2006)
10. Forecasts from other centers (IRI, CDC, ...)

Why still need statistical methods?

- Dynamical models have demonstrated great success in tropical seasonal to inter-annual predictions (e.g., ENSO), but much less satisfactory for middle and higher latitudes;
- Statistical models can give comparable and even higher skills for some regions and for some variables;
- Statistical models are much more economy than dynamical ones.

Optimal Climate Normals

Rapid change of climate in last 30years has made WMO recommended climate (30-year mean, updated every 10 years) no longer appropriate. The average over last K years may be more representative of current state and a better estimate of the upcoming expected value.

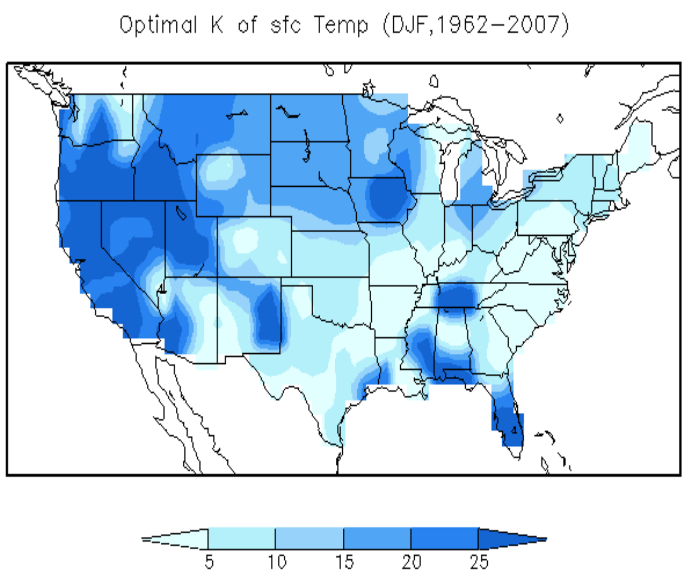
OCN Forecast Method: Taking the average of the most recent K years as the prediction for the coming year (Huang et al 1994).

For station *i* and target year *n*, *T(i,n)* anomaly prediction is

$$\hat{T}_K^f(i,n) = \frac{1}{K}\sum_{j=1}^{K} T(i,n-j) - C_{WMO}(i)$$

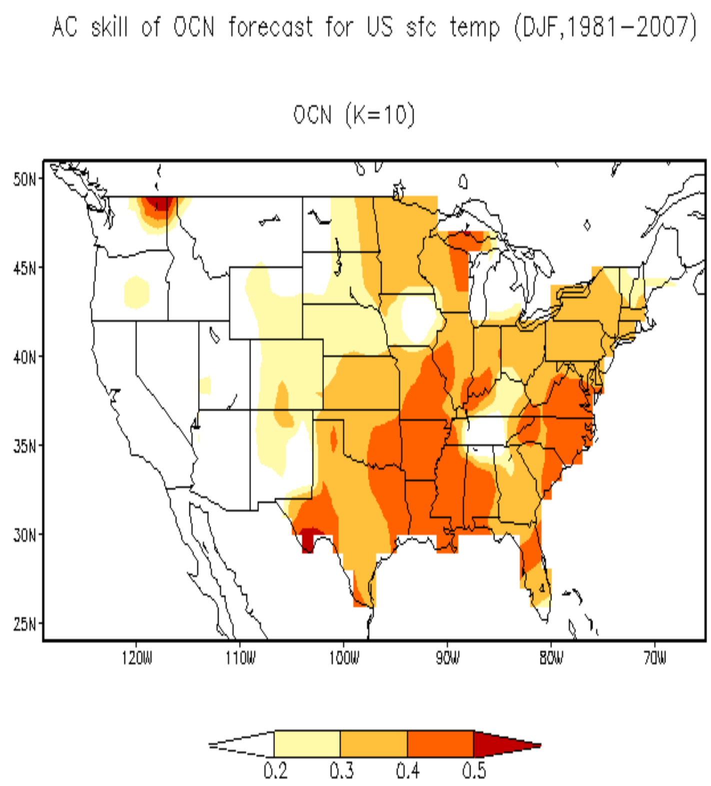
Optimal K is determined by maximizing the correlation skill over the training period:

Optimal K of surface temp (DJF, 1962-2007)



- Abrupt changes in K site by site would lead to a prediction inconsistent in space;
- A single K which makes the skill averaged over space and seasons maximum thus has been used: K~10 yrs

9-Month lead DJF Skill of OCN Forecast (K=10)

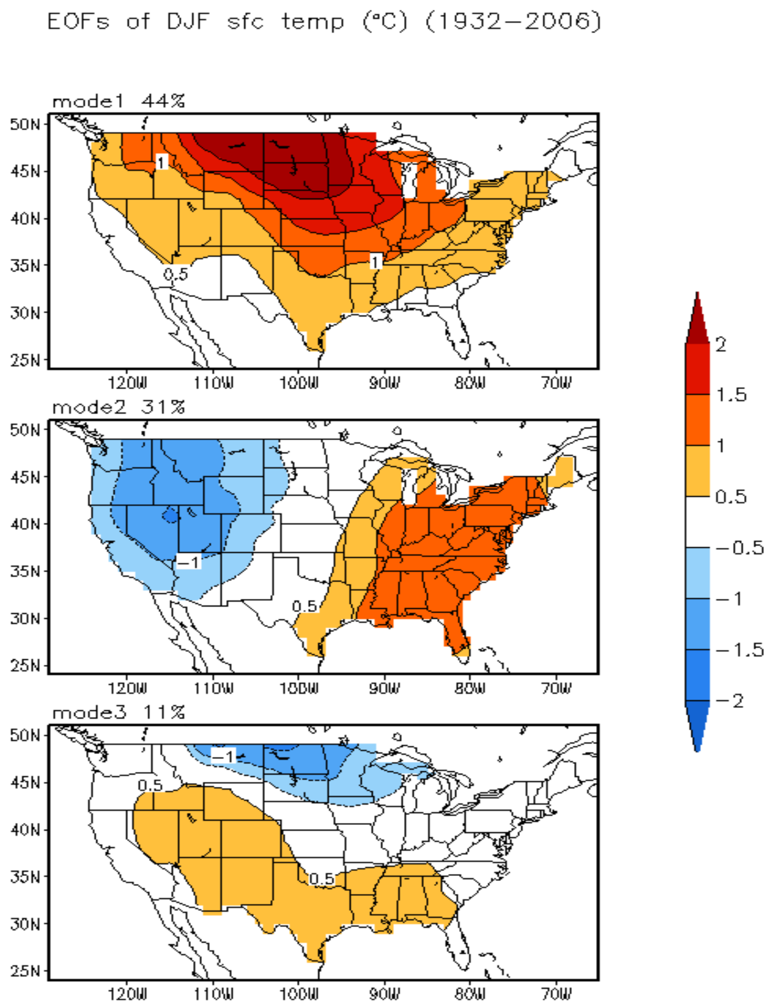


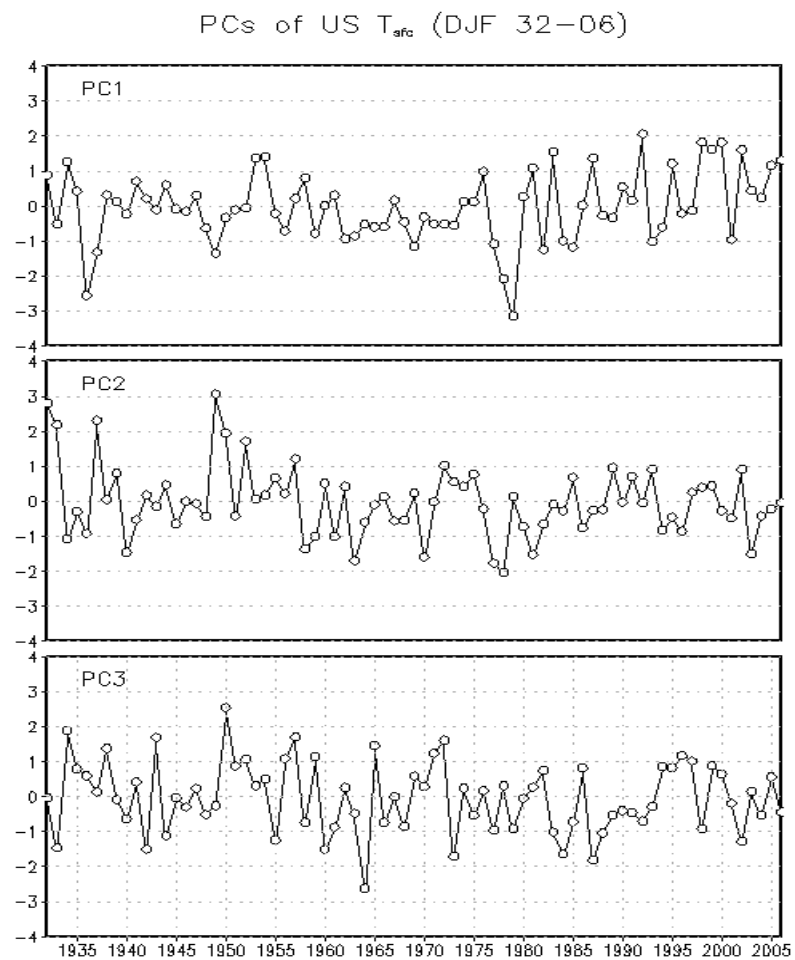
- Decent skills with simple method;
- Caveats:

  a) Geographical changes in K is totally ignored;

  b) Doesn't count multiple timescales of variability.

EOFs and PCs of US DJF Surface Temp (1932-2006)





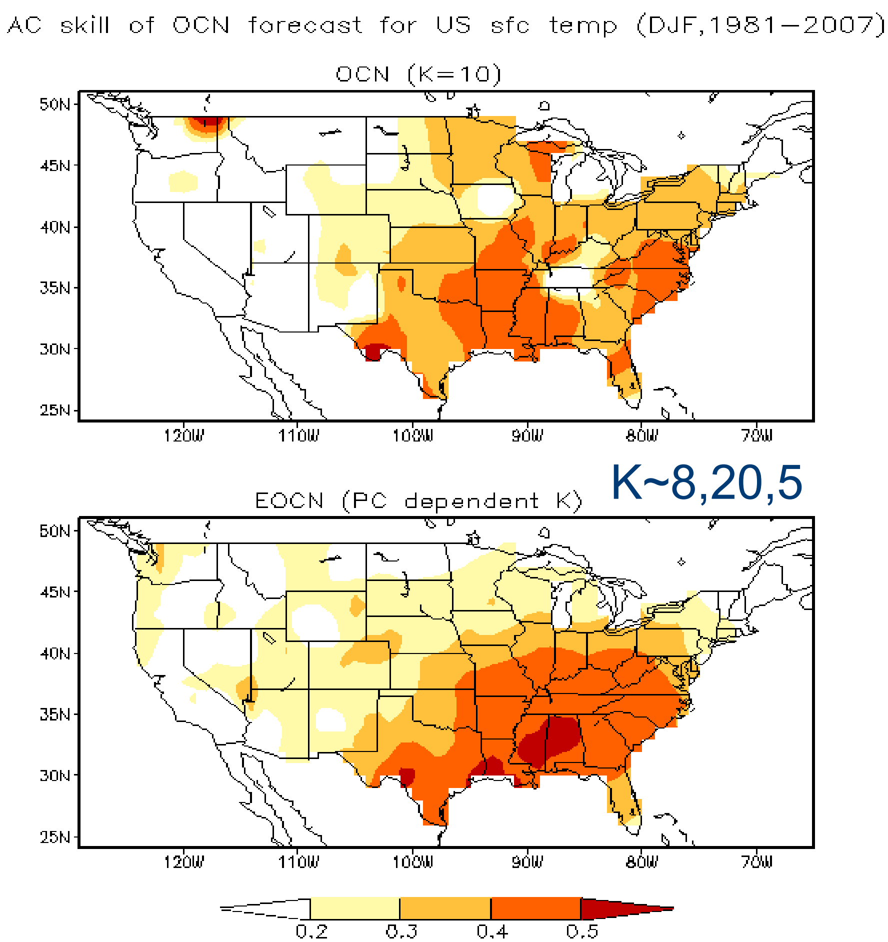
86% variance explained

Timescales are likely PC dependent

EOF Adjusted OCN (EOCN)

- Represent space-time fields of a variable with EOFs and PCs;
- Determine optimal K for each PC;
- Predict coming year PCs with the OCN scheme;
- Synthesize forecast with EOFs and the predicted PCs.

Skill Comparison: OCN vs EOCN



- Modest skill improvement is achieved with the EOF adjustment;
- Caveats:

  a) EOFs may not quite physically based;

  b) PCs not stationary.
- Further improvement is possible by using

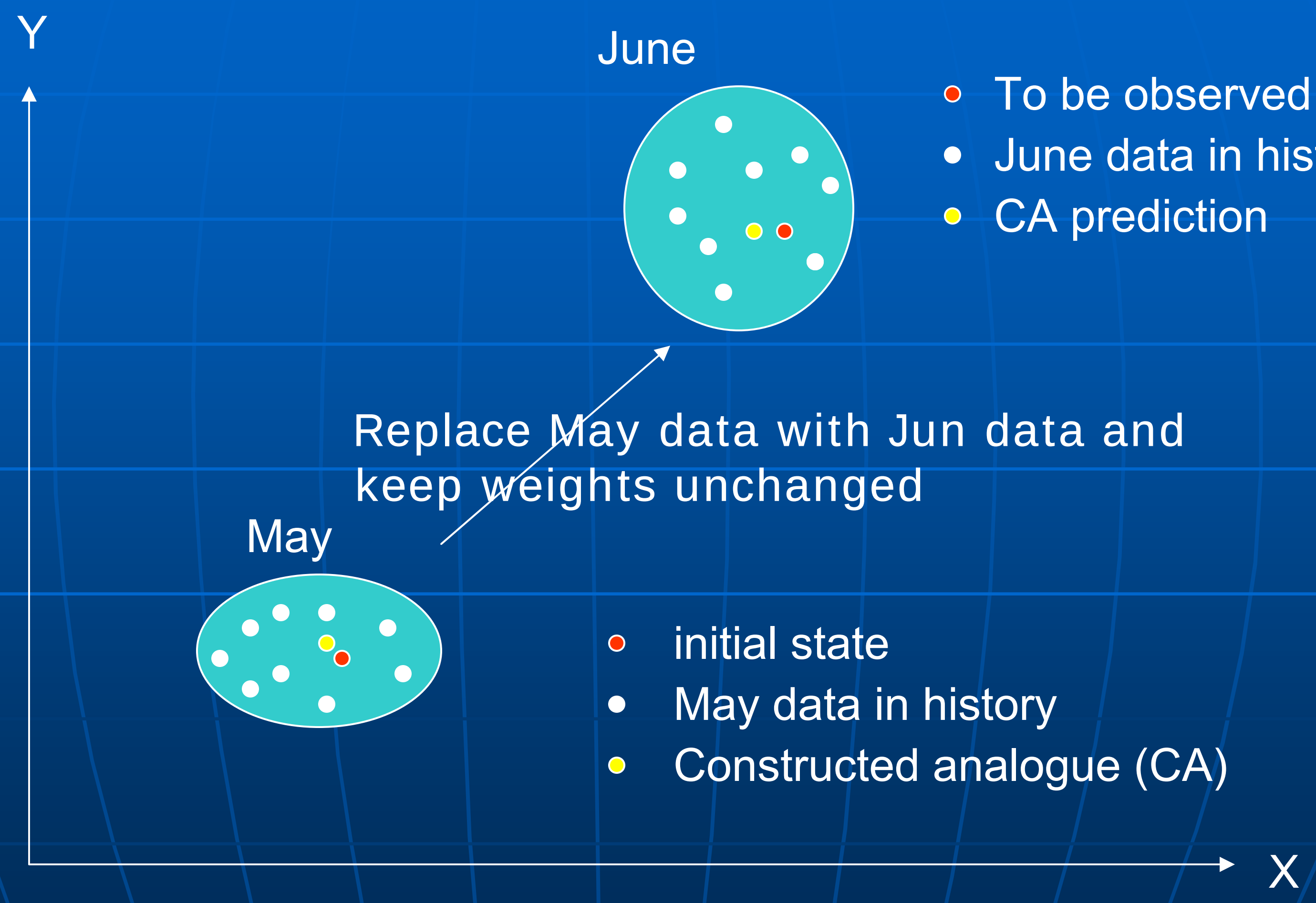
  a) More physically based decomposition;

  b) New techniques to deal with non-stationary.

Constructed Analogues (CA)

- Why constructed?
  - Natural analogues are highly unlikely to occur in high degree-of-freedom processes;
- CA method
  - Linearly combine past observed anomaly patterns such that the combination is as close as desired to the initial state, then carry forward in time with weights persisting;

Constructed Analogues (concept)



Constructed Analogues (*formulas*)

Use the weighted average of historical data to approximate current data (IC):

$$X^{IC}(s,t_0) \approx X^{CA}(s,t_0) = \sum_{t=1}^{T} \alpha(t) X(s,t)$$

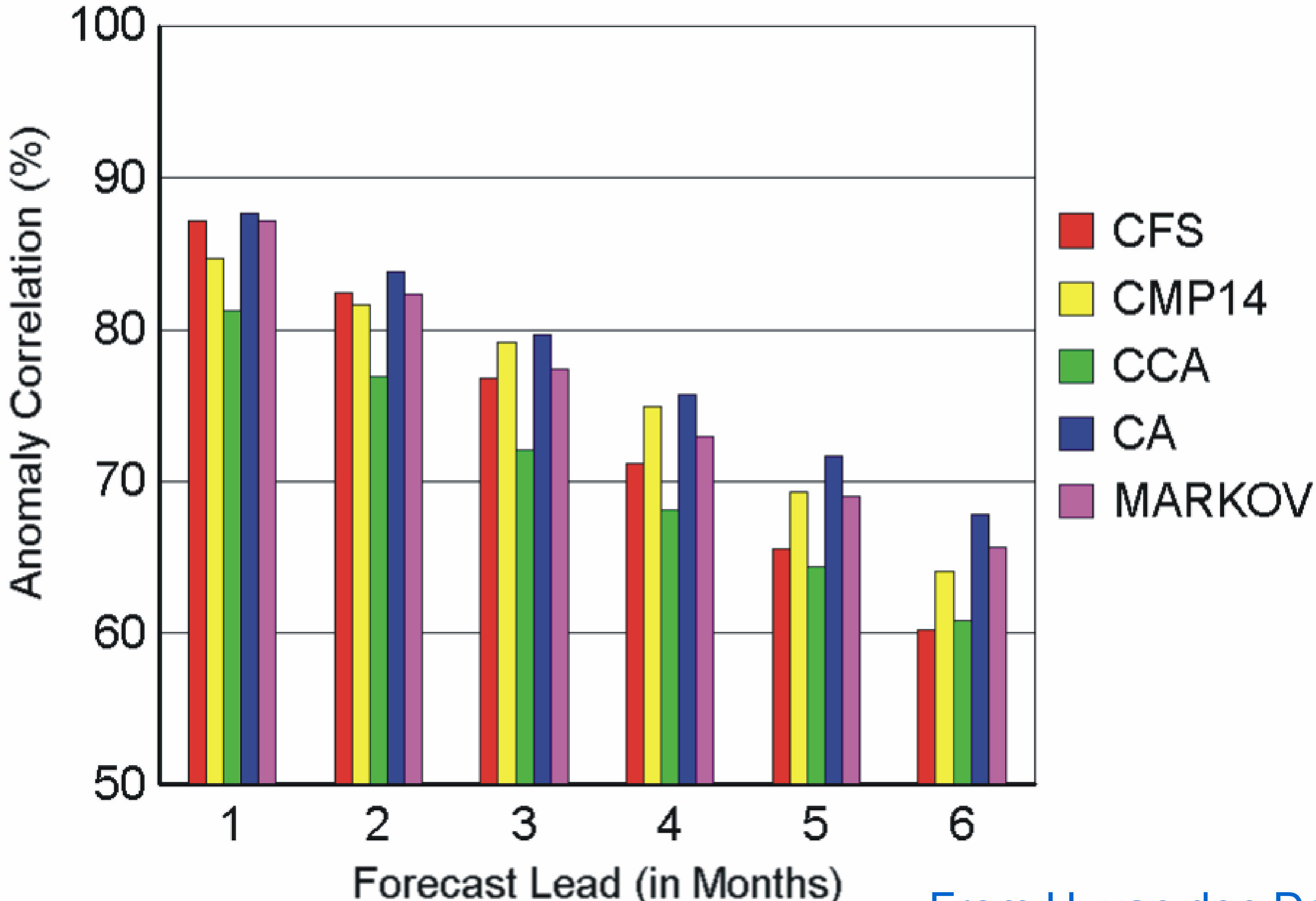
The weights are obtained by minimizing the error:

$$R = \sum_{s=1}^{S} (X^{ic}(s,t_0) - X^{CA}(s,t_0))^2$$

Construct forecast by using the same weights for the “future” data:

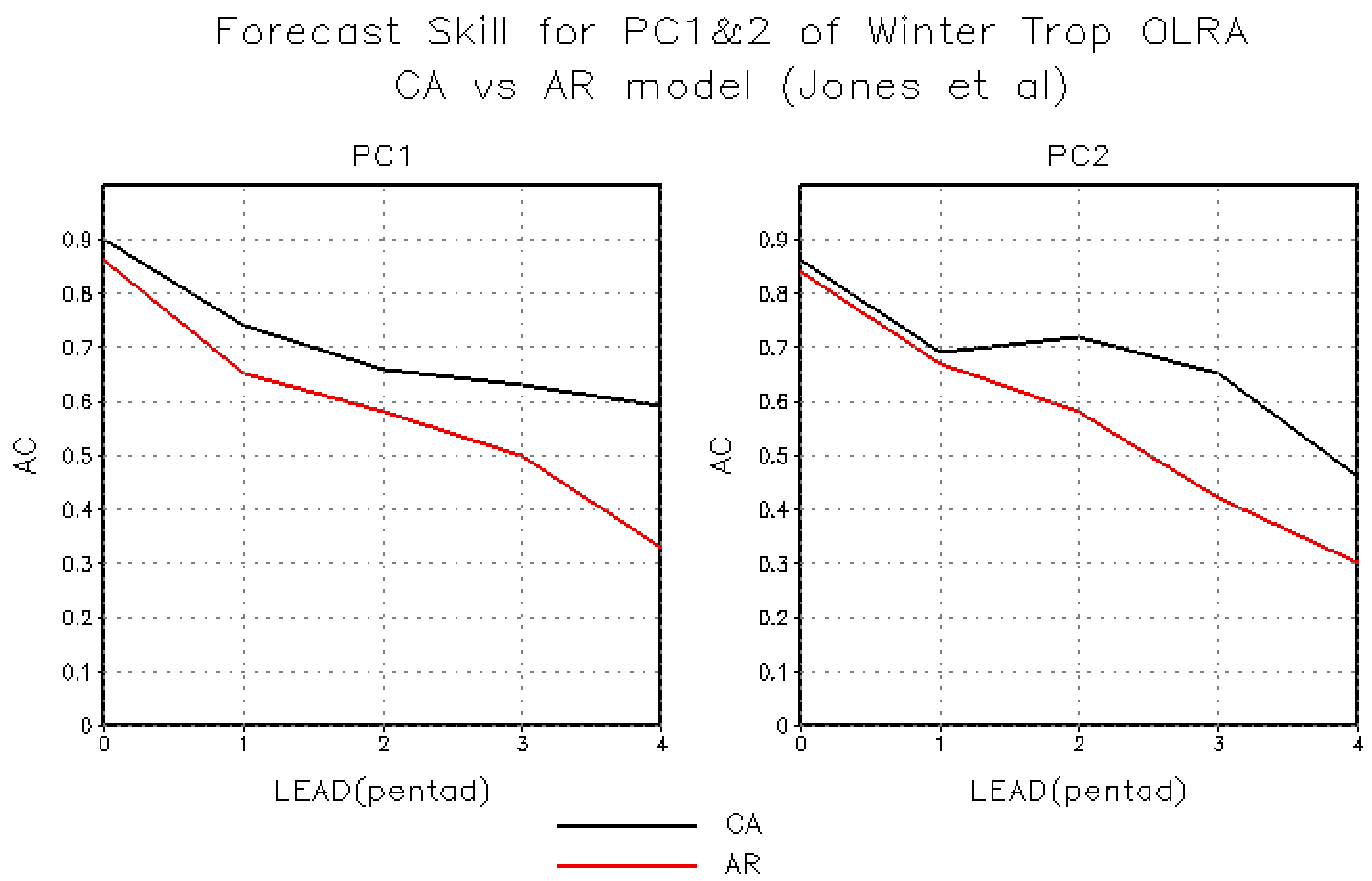
$$X^{F}(s,t_0 + \Delta t) = \sum_{t}^{T} \alpha(t) X(s,t + \Delta t)$$

Skill in SST Anomaly Prediction
Nino-3.4 (DJF 81/82 to DJF 03/04)



From H. van den Dool

CA vs AR for MJO index forecast



For MJO forecast, CA can be better than AR

Forecast Tool Consolidation (an Experiment)

Because different tools have different physical basis or are based on different statistical schemes, their skill scores can be quite different in geographical distribution and magnitude. Thus it is necessary to optimally combine predictions from various forecast tools.

Recent relevant works:

Krishnamurti et al 1999, 2000; Kharin and Zwiers 2002; Peng et al. 2002; and others.

Member models for the experiment

- NCEP Climate Forecast System (CFS)
  A coupled dynamical model used for climate forecast in NCEP;

- Canonical Correlation Analysis (CCA)
  Predictors: global SST, Z700, surface air temp (T2m) or precipitation rate (Prate)
  Predictants: T2m or Prate

- Screening Multiple Linear Regression (SMLR)
  Predictors: global SST, Z700, soil moisture,T2m or Prate
  Predictants: T2m or Prate

Data

- Model data:

  JFM mean T2m from one-month lead hindcasts by CFS, CCA, and SMLR for the period of 1982-2006.
- Verification: NCDC climate division data.
- All the data are interpolated to 102 US climate divisions and normalized with its standard deviation at each division.

Consolidation Techniques

General formulation:

$$F = \sum_{i=1}^{I} w_i f_i \qquad I = 3$$

Where, w_i is the weight assigned to the forecast by i^{th} member model f_i .

1. Equally weighted ensemble (EW)

$$w_i = 1 / I$$

2. Ridging Regression (RR)

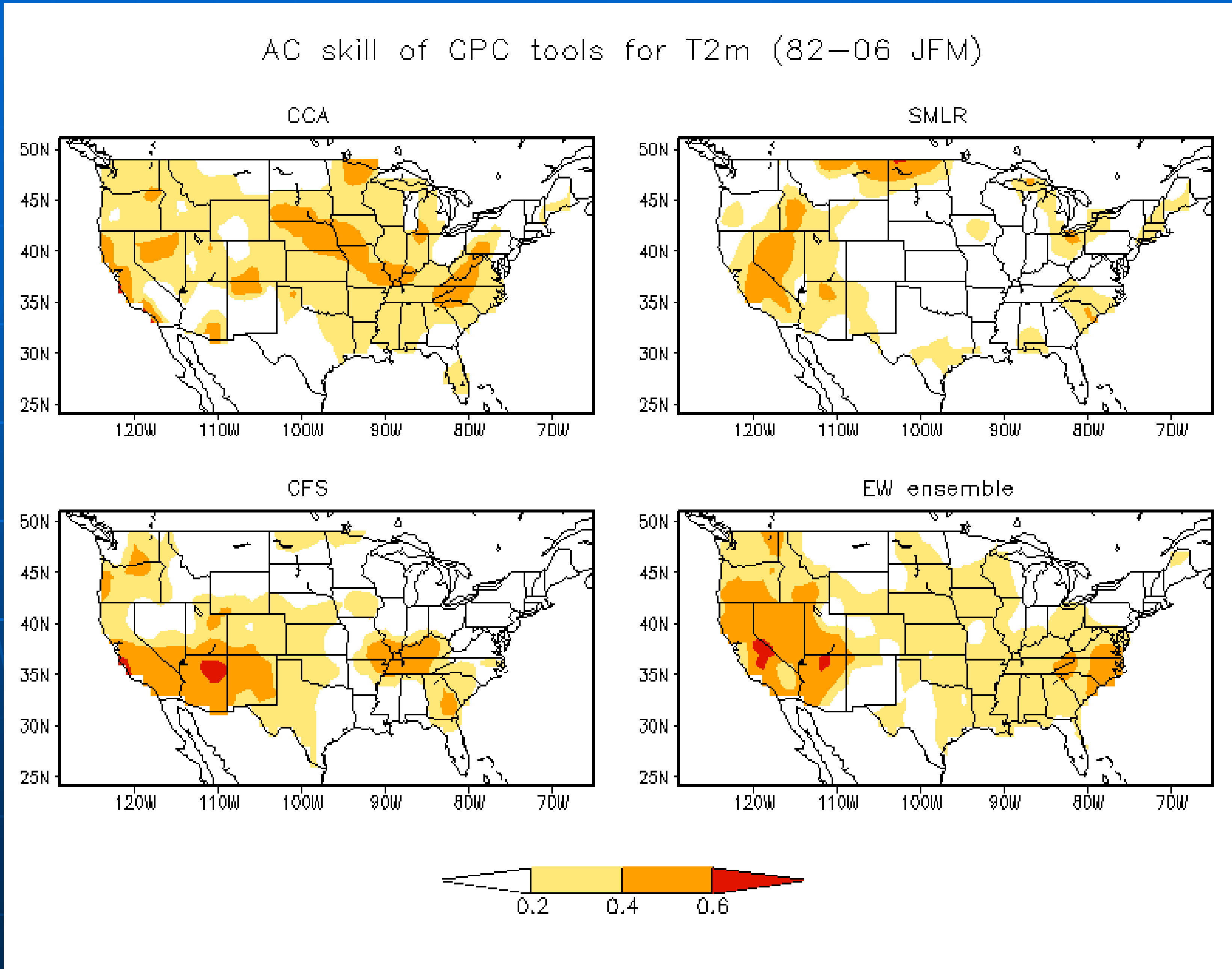
   "optimal" weights are determined by minimizing RMS errors of the consolidated forecast over the training period. The singularity problem of the regression matrix caused by the "co-linearity" of model data is eliminated with the ridging technique.

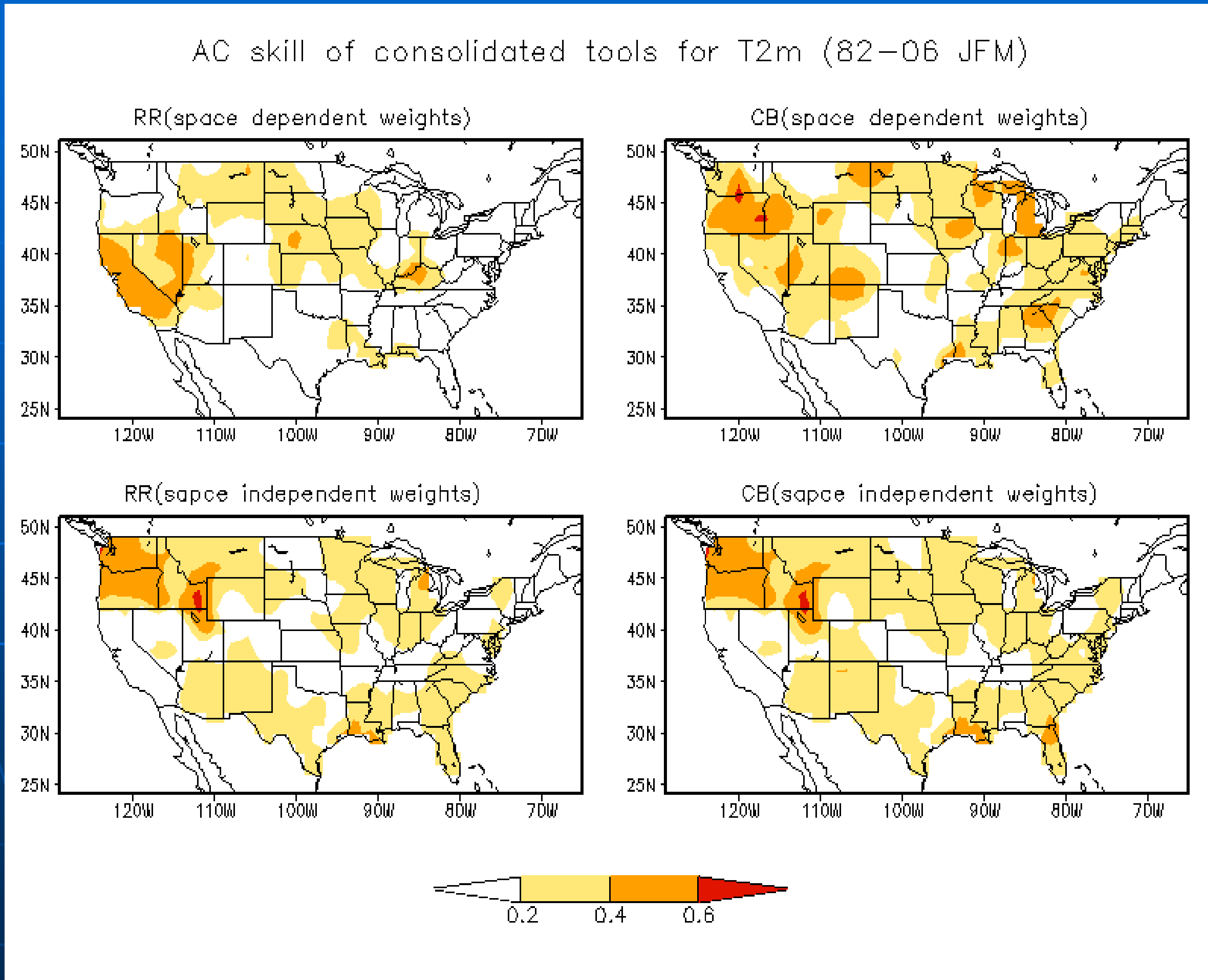
In order for the training data to be independent, cross validation is applied in weight calculations.

3. Cross-validated best model (CB)

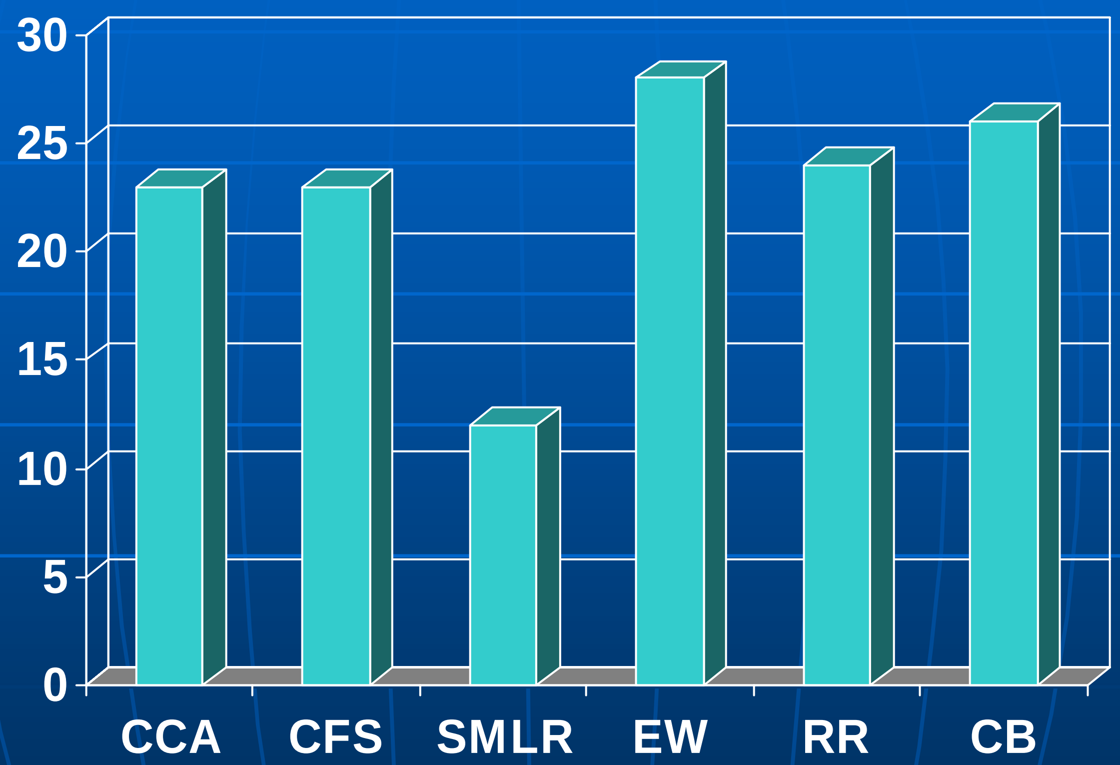
The weights are based on the performance of member models in the history (in terms of cross validation). The best model receives weight 1, while others get weight 0.

In techniques RR and CB, weights are determined either for individual climate divisions (space-dependent weights) or for whole 102 climate divisions (space-independent weights). In the latter case, weights are expected to be more stable than in the former case, owning to the bigger sample size.


AC skill of CPC tools for T2m (82–06 JFM)
CCA
SMLR
CFS
EW ensemble
50N
45N
40N
35N
30N
25N
120W
110W
100W
90W
80W
70W
0.2
0.4
0.6

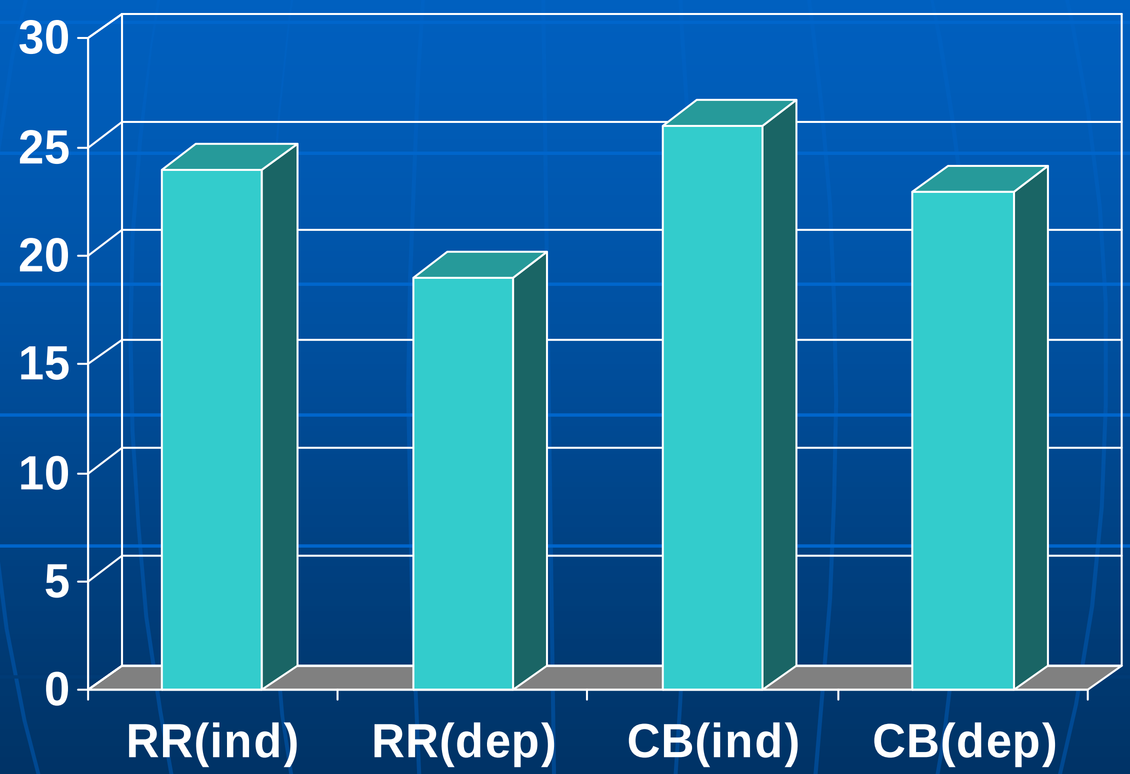

AC skill of consolidated tools for T2m (82-06 JFM)
RR(space dependent weights)
CB(space dependent weights)
RR(sapce independent weights)
CB(sapce independent weights)
50N
45N
40N
35N
30N
25N
120W
110W
100W
90W
80W
70W
0.2
0.4
0.6

AC skill (%) of 82-06 JFM T2m for US 102 CD (CV3)
weights in RR and CB are space-independent



1. Consolidated forecasts are better than individual models;
2. More sophisticated schemes (RR & CB) are hard to beat the simplest one (EW).

AC Skill Comparison for RR and CB space-dependent weights vs space-independent weights



1. Space-independent weights give higher skill than space-dependent weights;
2 why? Sample size problem?

Weights in RR and their variability for whole102 CDs

	Space independent weights		Space Dependent weights	
	mean	σ/mean	mean	σ /mean
CCA	.17	23%	.23	64%
CFS	.16	16%	.22	41%
SMLR	.06	52%	.14	67%

For the limited length of training data (~20 years), EW gives the highest skill. The reason is that the training data are too short to generate stable weights for RR and CB.

Summary

1. Seasonal forecast skill of statistical models is comparable and even higher than dynamical models in many circumstances;
2. Because of the difficulty for dynamical models to deal with the complexity of the nature, statistical models will keep useful to foreseeable future.
3. Good statistical models should be physically based; Statistical models and dynamical models need to compensate each other.
4. In our experiment, sophisticated consolidation schemes (RR and CB) are hard to beat the simplest one (EW), the reason is that the training data (~ 20 years long) are too short to generate stable weights.