Multimodel Ensemble forecasts

Calibrated methods

Michael K. Tippett

International Research Institute for Climate and Society
The Earth Institute, Columbia University

ERFS Climate Predictability Tool Training Workshop
May 4-9, 2009

Ideas

- Extract useful information from ensembles
- Simple ways of combining ensembles from different models

Methods that depend on model performance.

Calibration methods

- Forecast probabilities are constructed taking into account model performance.
- Historical observations and hindcasts are necessary.
- Correct some aspects of the ensemble information and use them to construct probabilities.

Independent calibration

First consider methods that use equal weighting across models or individual model performance, independent of the behavior of other models.

Spread correction

Given the mean of the forecast pdf,
construct a distribution about that mean based on past model performance.

- The forecast mean may come one of the methods discussed previously.
- Goal is to treat problems with ensemble spread.

Historical ensemble

Ideally, the verifying observation is indistinguishable from an ensemble member.

Historical ensemble = mean + (Past obs - past means)

Number of historical ensemble member = number of hindcasts.

Historical ensemble is constant from year to year.
Varies according to start and lead.

Historical Gaussian distribution

Forecast distribution is a Gaussian distribution with mean specified.

Variance estimate from (Past obs - past means)

Not optimal if the mean is poorly calibrated.

Forecast variance may be greater than the climatological variance.

A forecast with zero mean would indicate enhanced probability not of the normal category but of the above and below categories.

Historical Gaussian distribution (2)

A “signal-noise” decomposition.

First compute the correlation r between the mean and observations.

The signal-to-noise ratio S is

$$S = \frac{r^2}{1 - r^2} = \frac{\text{signal variance}}{\text{noise variance}}$$

Noise variance = (variance of forecast mean) / S

Historical Gaussian distribution (2)

Normalize forecast mean so that variance of forecast mean = 1

Forecast distribution is Gaussian with normalized forecast mean and variance $1/S$

Climatology distribution has mean zero and variance equal to the sum of the signal and noise variances.

In the limit that $r \to 0$, the forecast distribution equals the climatological distribution. No predictability.

For $r > 0$ the forecast distribution is strictly narrower than the climatological distribution.

Nonhomogeneous Gaussian regression

Gneiting et al. 2005, Wilks 2006

In the historical Gaussian distribution, the forecast variance is constant from year to year. Depends only on start and lead.

In NGR, forecast variance is a constant plus term that is proportional to the ensemble variance.

Constant of proportionality is found based on past performance. Choose constant to optimize the skill of past forecasts.

For instance, optimize the continuous ranked probability skill score.

Separate regression

Perform regressions between each model and observations.
Average the separate regressions.
Model the uncertainty using a Gaussian distribution.
Equivalent to multiple linear regression with the assumption that the models are independent.
Avoids assigning negative weights to model with positive skill.

Categorical probability calibration

Generalized linear model

Develop a regression between forecast mean and event occurrence (binary).

Generalized linear model.
“Probit”

Categorical probability calibration

Bayesian optimal weighting

Rajagopalan et al 2002, Robertson et al 2004

Weights are assigned to ensemble probabilities and climatological probabilities.

Weights are chosen to optimized the log-likelihood of the observations. Proportional to average "ignorance".

Resulting probabilities are between the ensemble probabilities and climatological probabilities.

Cannot sharpen probabilities.

Joint calibration

Multimodel ensemble methods where the calibration of a model to observations depends on the behavior of other models.

Advantage

- Incorporates information about all the possible relations between observations and models.

Disadvantage

- Incorporates information about all the possible relations between observations and models.
- Many parameters to estimate from little data!
- Over-fitting.

“Super-ensembling”

Multiple regression between the ensemble means of each model and observations.

A forecast distribution about the result of the MLR can be constructed from past performance.

“Super-ensembling”

Problem 1: collinearity. Predictors are highly correlated.

Data is well fit, but regression coefficients are highly sensitive to the data and tend to have large errors.

Poor predictions!

Problem2: Overfitting. Poor predictions!

Fixes

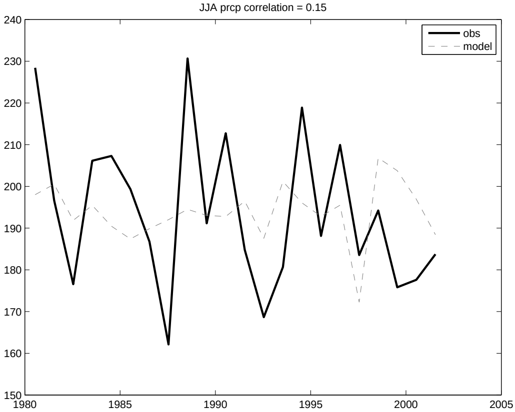
Reduce the number of predictors.

- PCA on model predictors. Retain components that explain the most variances. Components are uncorrelated.

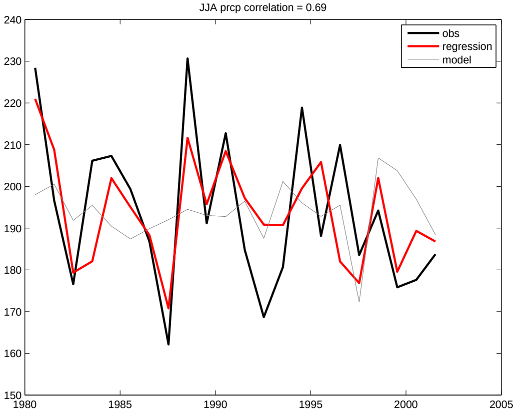
Ridge regression. Bayesian interpretation (Delsole 2004).
Prior is:

- model weighting is $1/N$.
- equal model weighting

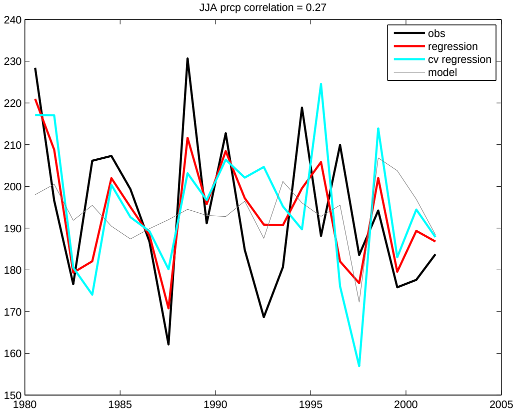
Example: precip



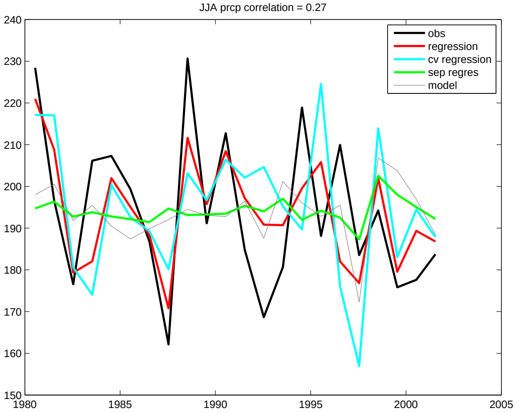
Example: precip, MLR in sample



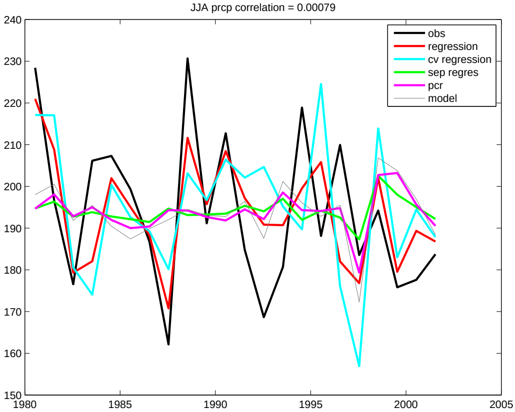
Example: precip MLR cross-validated



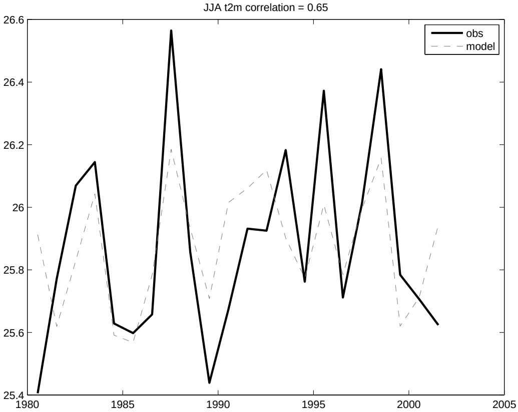
Example: precip separate regressions



Example: precip PCR

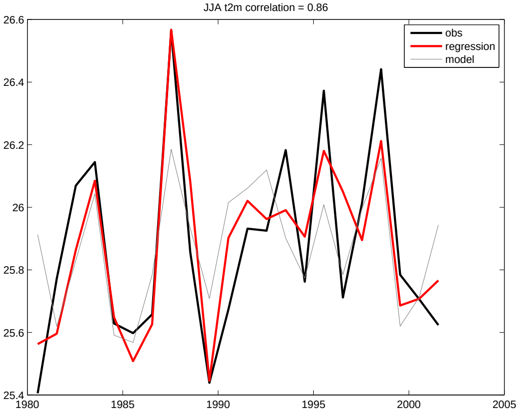


Example: t2m

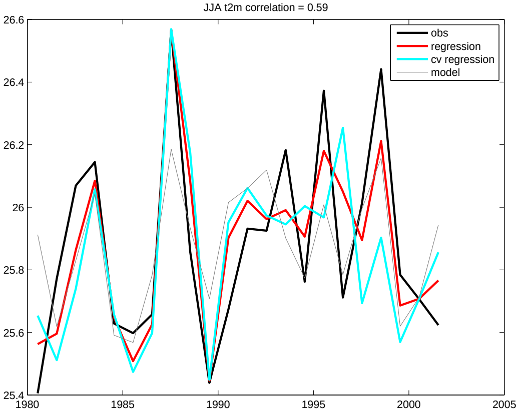


JJA t2m correlation = 0.65
obs
model
26.6
26.4
26.2
26
25.8
25.6
25.4
1980
1985
1990
1995
2000
2005

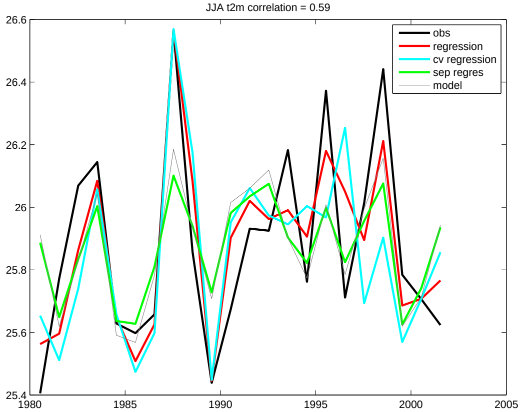
Example: t2m, MLR in sample



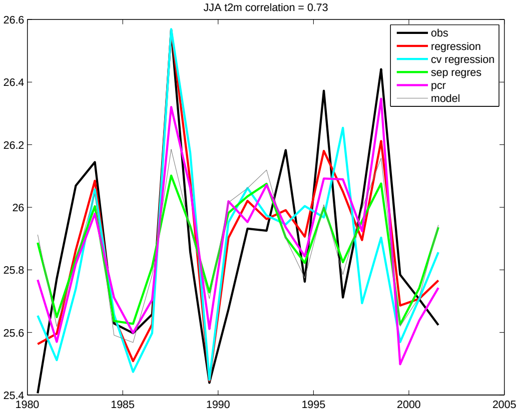
Example: t2m MLR cross-validated



Example: t2m separate regressions



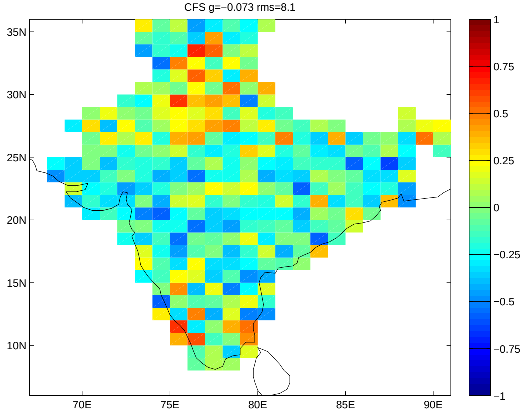
Example: t2m PCR



Summary

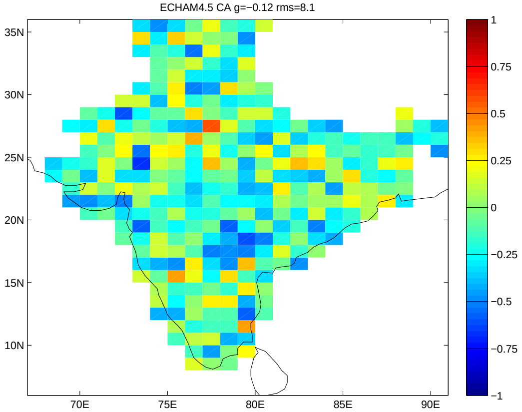
- Independent calibration
  - Base spread on historical performance rather than ensemble.
  - Correct models separately and combine equally.
- Joint calibration
  - MLR – collinearity, overfitting.
  - separate regression
  - Ridge regression
  - PCR (across models)
- Be honest! Cannot pick predictors using the complete data set. Need independent data.

CFS MOS



CFS g=−0.073 rms=8.1
35N
30N
25N
20N
15N
10N
70E
75E
80E
85E
90E
1
0.75
0.5
0.25
0
−0.25
−0.5
−0.75
−1

ECHAM CA MOS



ECHAM4.5 CA g=−0.12 rms=8.1
35N
30N
25N
20N
15N
10N
70E
75E
80E
85E
90E
1
0.75
0.5
0.25
0
−0.25
−0.5
−0.75
−1

ECHAM GML MOS