

# Recalibrating and Combining Ensemble Predictions

Michael K. Tippett, Lisa Goddard, Simon Mason (IRI),  
Malaquias Pena Mendez (SAIC-EMC/NCEP/NOAA),  
Huug van den Dool (CPC/NCEP/NWS/NOAA)

Extracting as much useful information  
as possible from model forecasts

# Specifics

- Calibrate model outputs using pattern-based regression.
- Develop PDF from regression statistics.

# Outline

- Are regression forecasts reliable?
- How to pick predictor patterns?
- Flexible Forecast Format

# Are regression forecasts reliable?

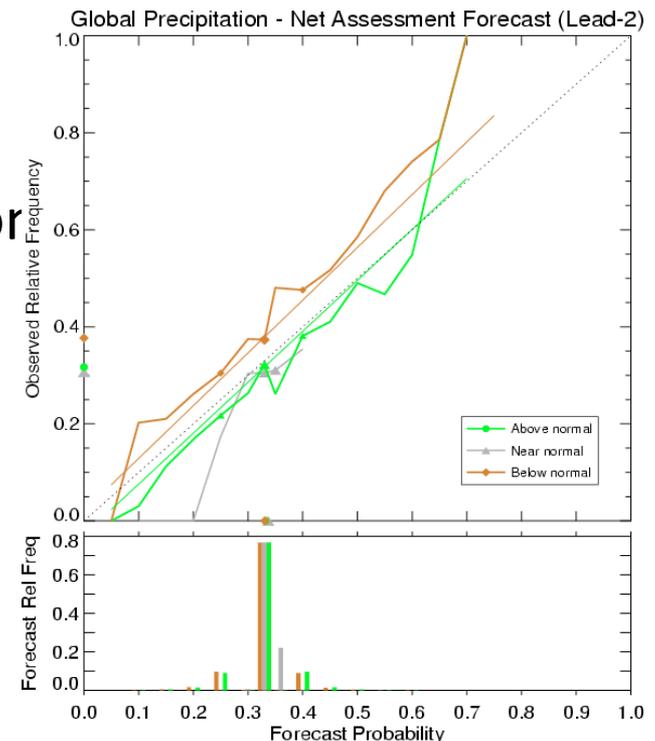
- Model Output Statistics (MOS)
- Predictor = model forecast =  $f$
- Predictand = observation =  $o$
- Regression
  - $E[o | f]$  = expected value of observation given the forecast
    - Perfect model --  $E[o | f] = f$
    - No skill –  $E[o | f] = \text{climatology}$
  - $\text{Var}[o | f]$  = variance of the observation given the forecast
    - Perfect model --  $\text{Var}[o | f] < \text{climatological variance}$
    - No skill –  $\text{Var}[o | f] = \text{climatological variance}$

# Are regression forecasts reliable?

- Linear regression (Gaussian)
  - $E[o | f] = a f + c =$  linear function of  $f$ .
    - Estimate parameters from data.
  - $\text{Var}[o | f] =$  error variance.
    - Estimate from data
- Generalized linear models (e.g., Poisson)
  - Count data

# Are regression forecasts reliable?

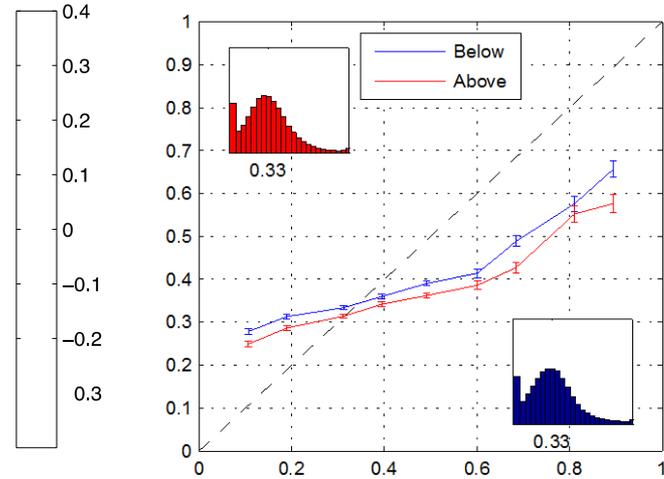
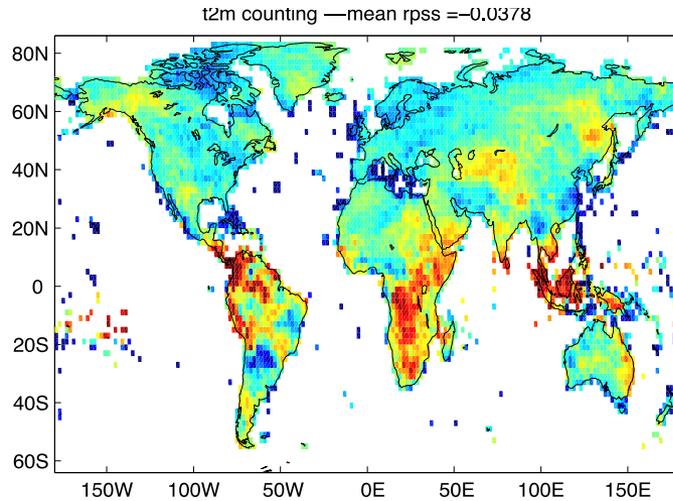
- Reliability diagram (Categorical forecasts)
  - $P$  = forecast probability of an event  $A$  (above)
    - $A = 1$  if the event happens and 0 otherwise
  - Plot  $E[A | P]$  vs.  $P$
  - Reliable if 45 degree line
  - Regression = conditional expectation
  - If  $P$  comes from regression
    - $P = E[A | P]$
    - plotting  $E[A | P]$  vs  $E[A | P]$
    - “guaranteed” reliability



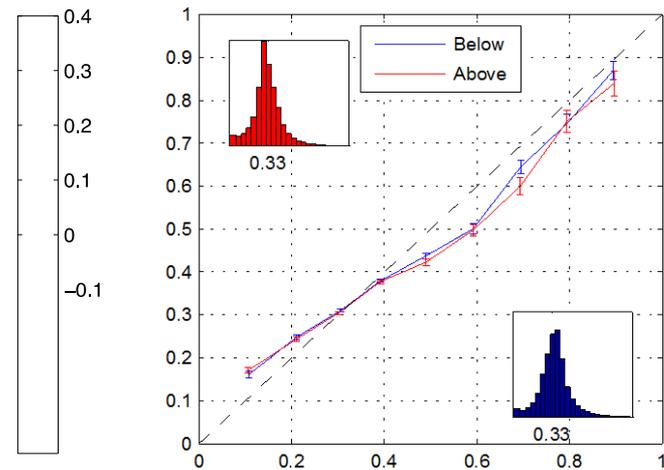
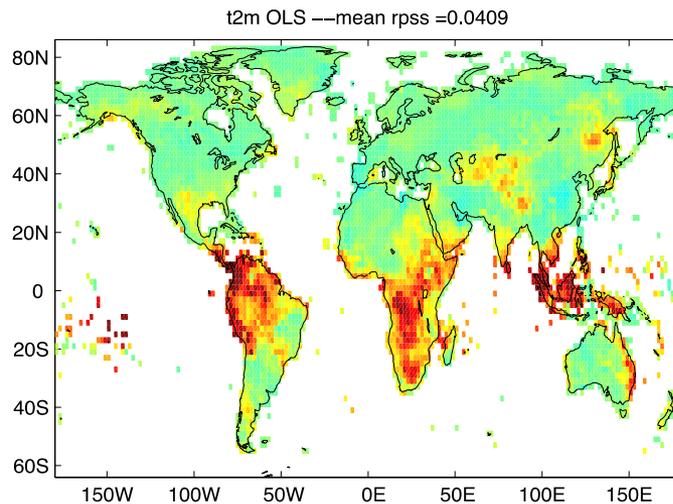
# Are regression forecasts reliable in practice?

DJF t2m  
Nov 1  
start

Ensemble  
Frequency

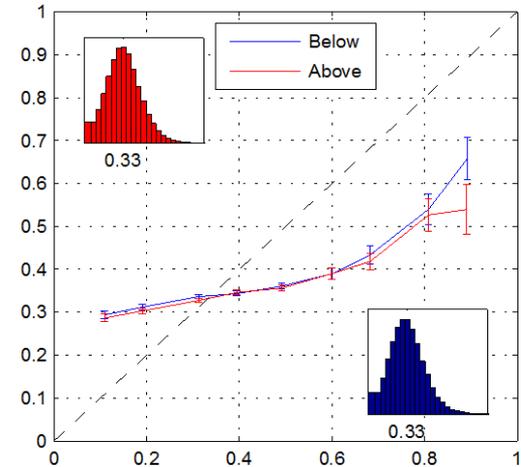
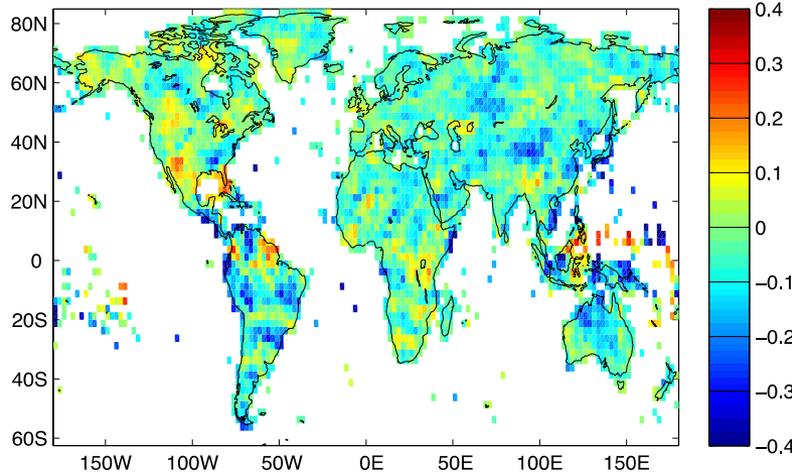


Linear  
regression

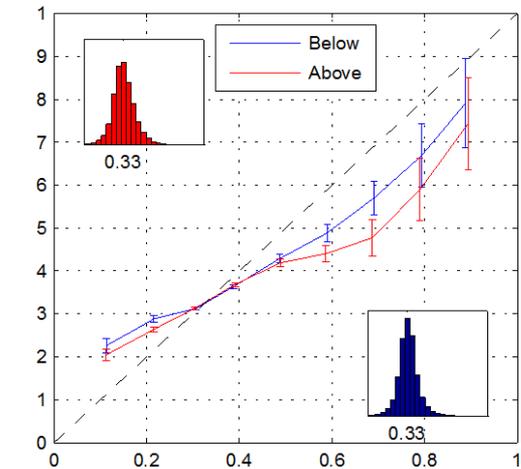
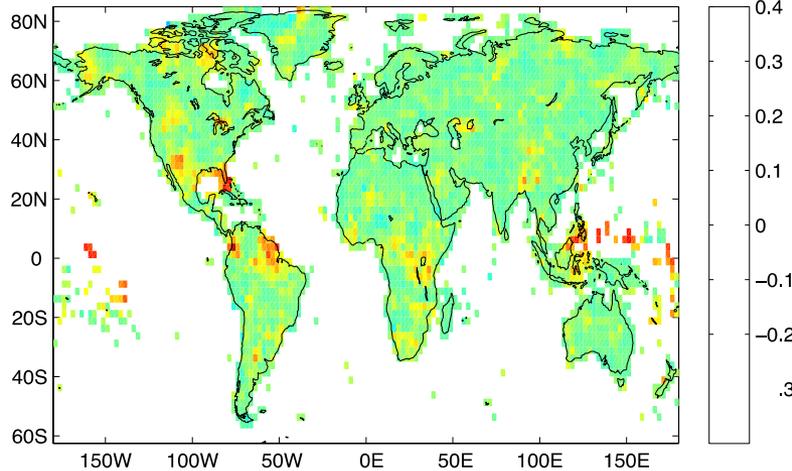


# Are regression forecasts reliable in practice?

DJF prcp  
Nov 1



Ensemble  
Frequency

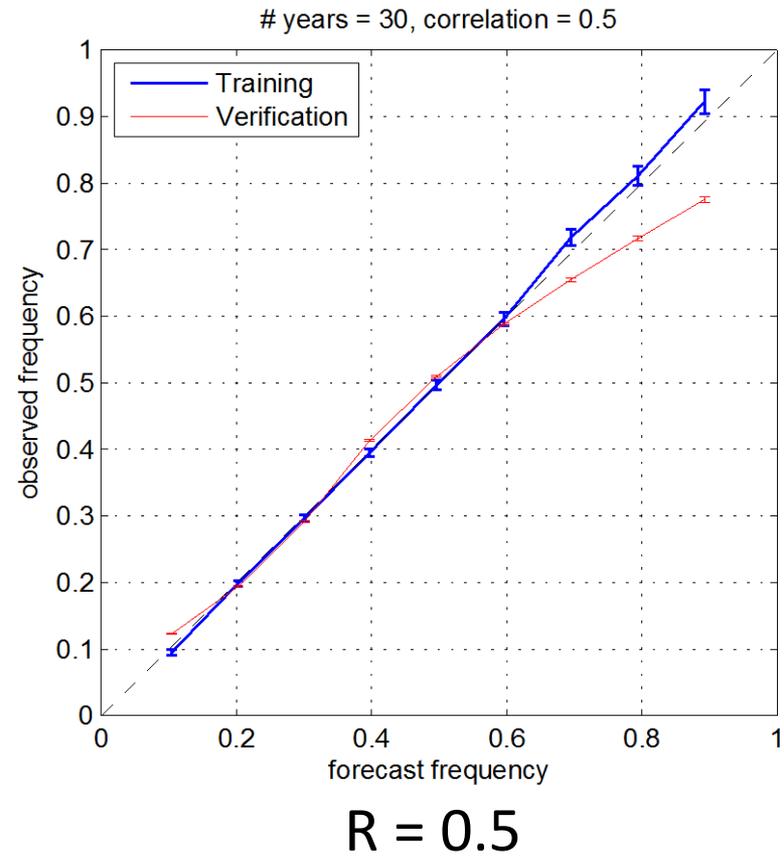
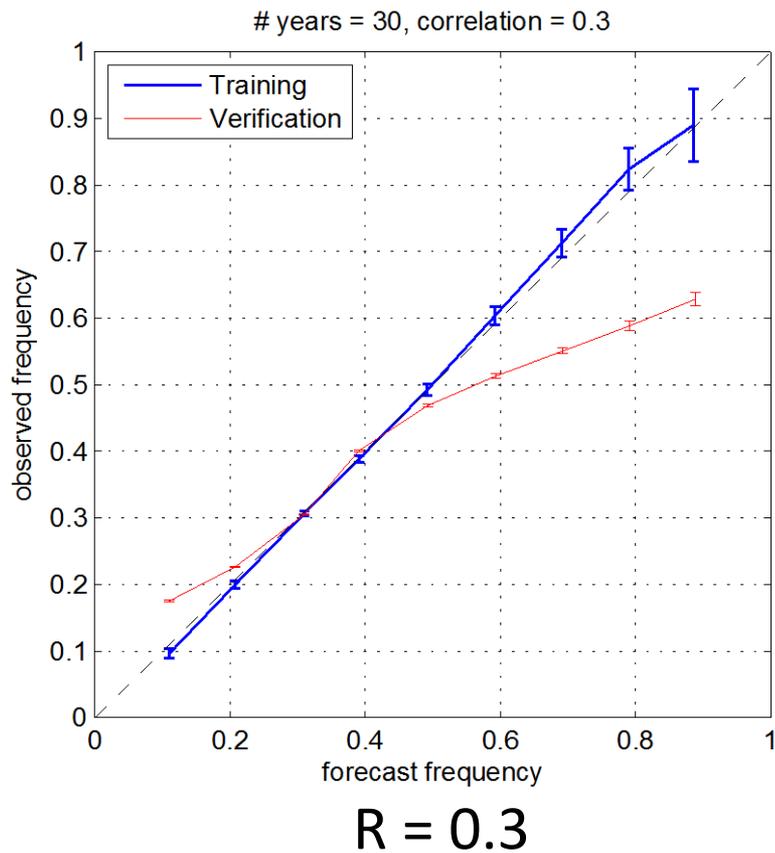


Linear  
regression

# What is the gap between theory and practice?

- Reliability exposes issues AC does not.
  - Correct signal strength
  - Correct uncertainty
- What is the source of over-confidence?
- Linear regression assumptions?
  - Gaussian distributions?
  - Spread too small?

# Synthetic data



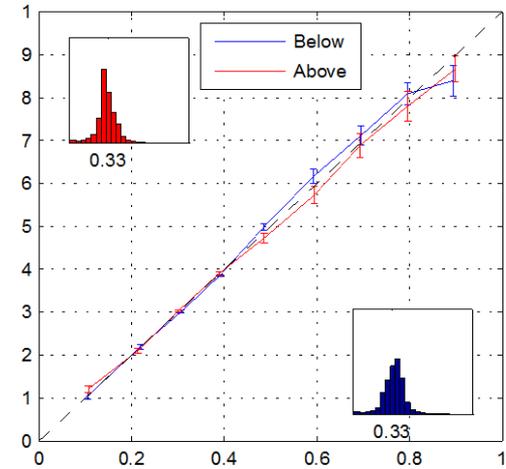
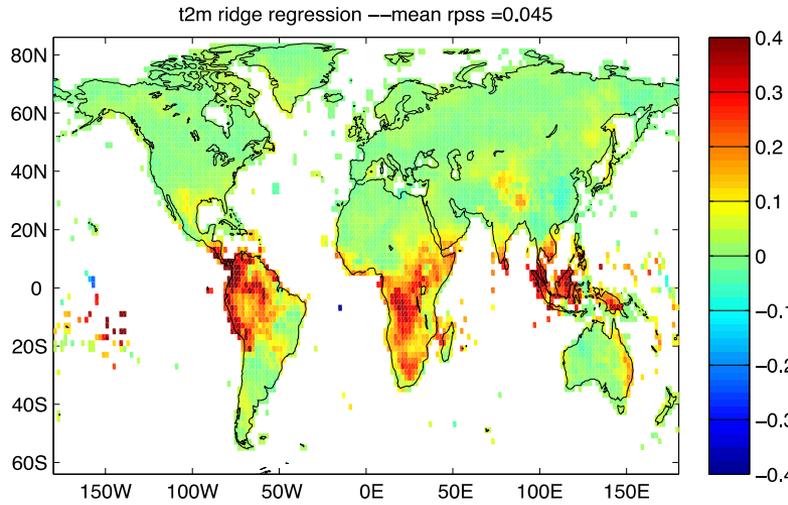
Training / Verification

# Why over-confident?

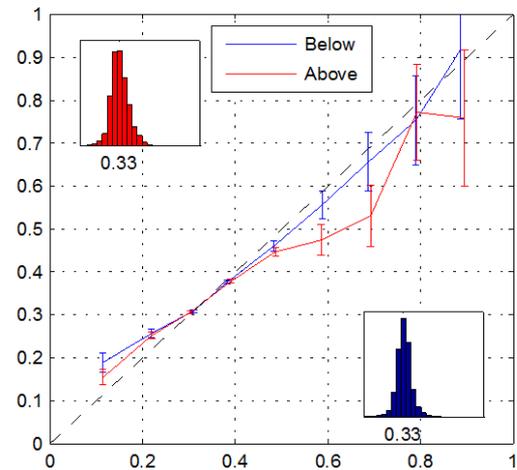
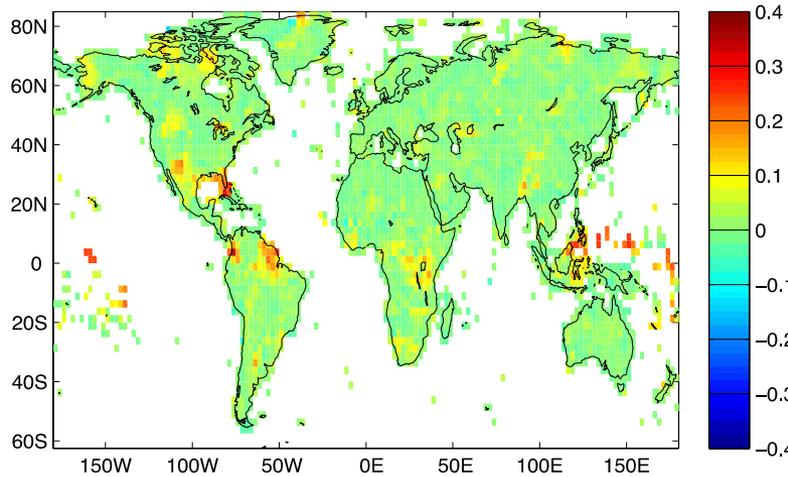
- Spread too small? No.
- Signal too strong? Yes.
  - $MOS = a f + c$ 
    - a estimated from data
    - a = “true” a + sampling error
    - $\text{mean}(\text{sampling error}) = 0$
  - Correct signal variance =  $a^2 \text{Var}(f)$
  - Signal variance in practice
    - $(a^2 + \text{sampling variance}) \text{Var}(f) > \text{true sig. var.}$

# Shrinkage

Ridge



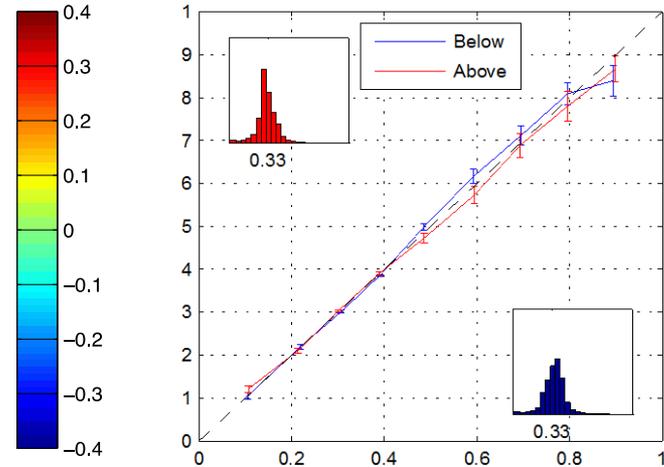
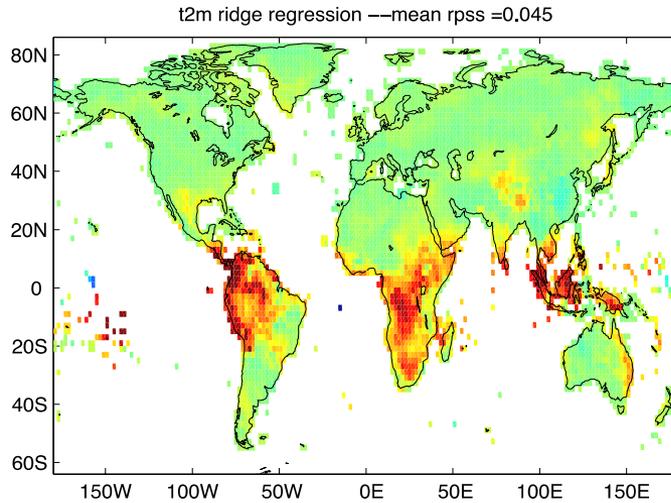
Lasso



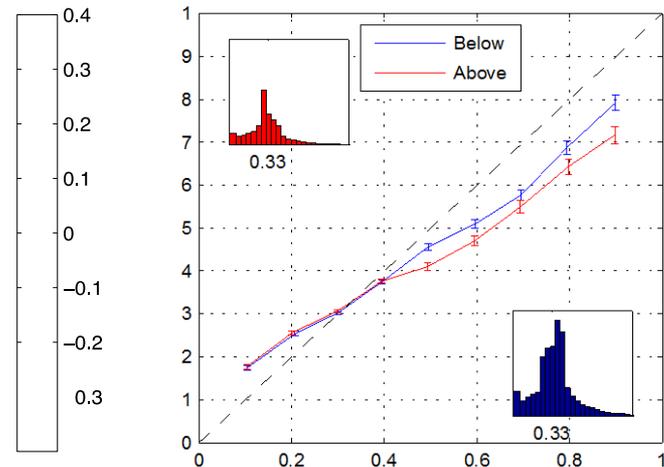
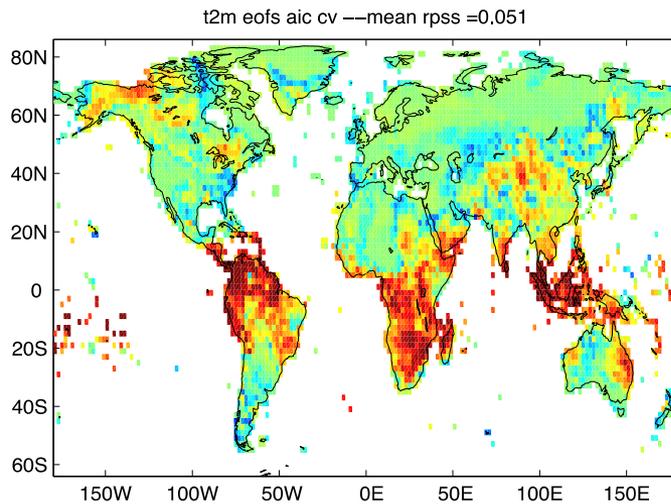
How to select predictor patterns?

# Selecting predictor patterns

Ridge

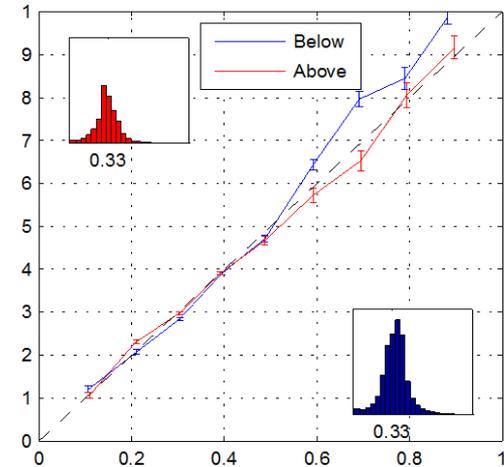
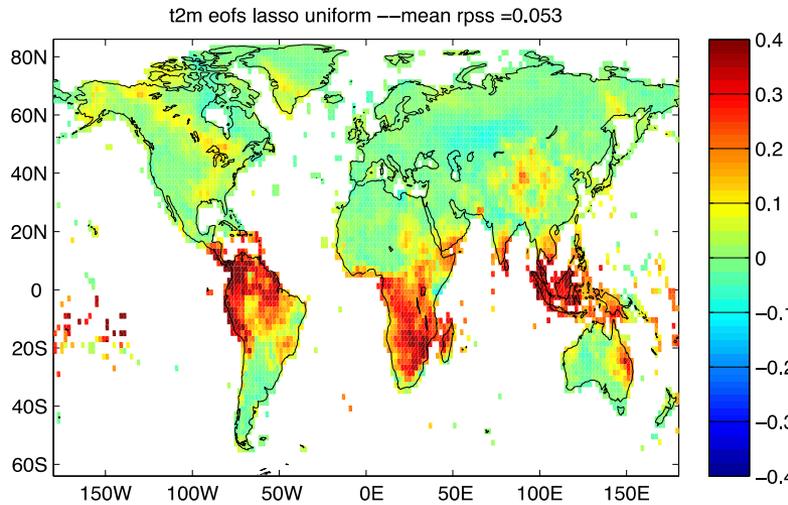


EOFs  
# by AIC

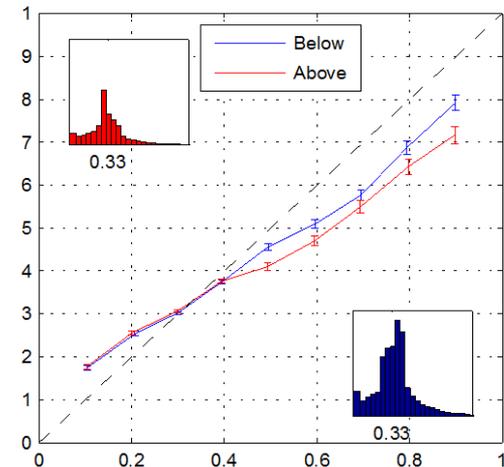
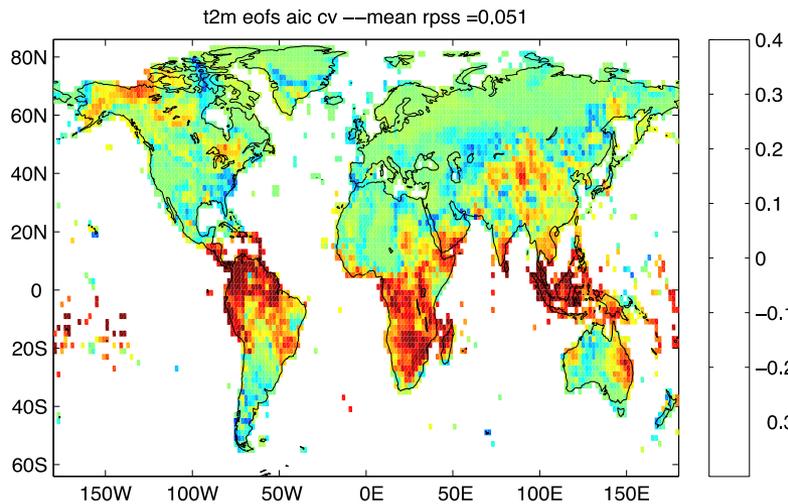


# Selecting predictor patterns

EOFs+  
lasso



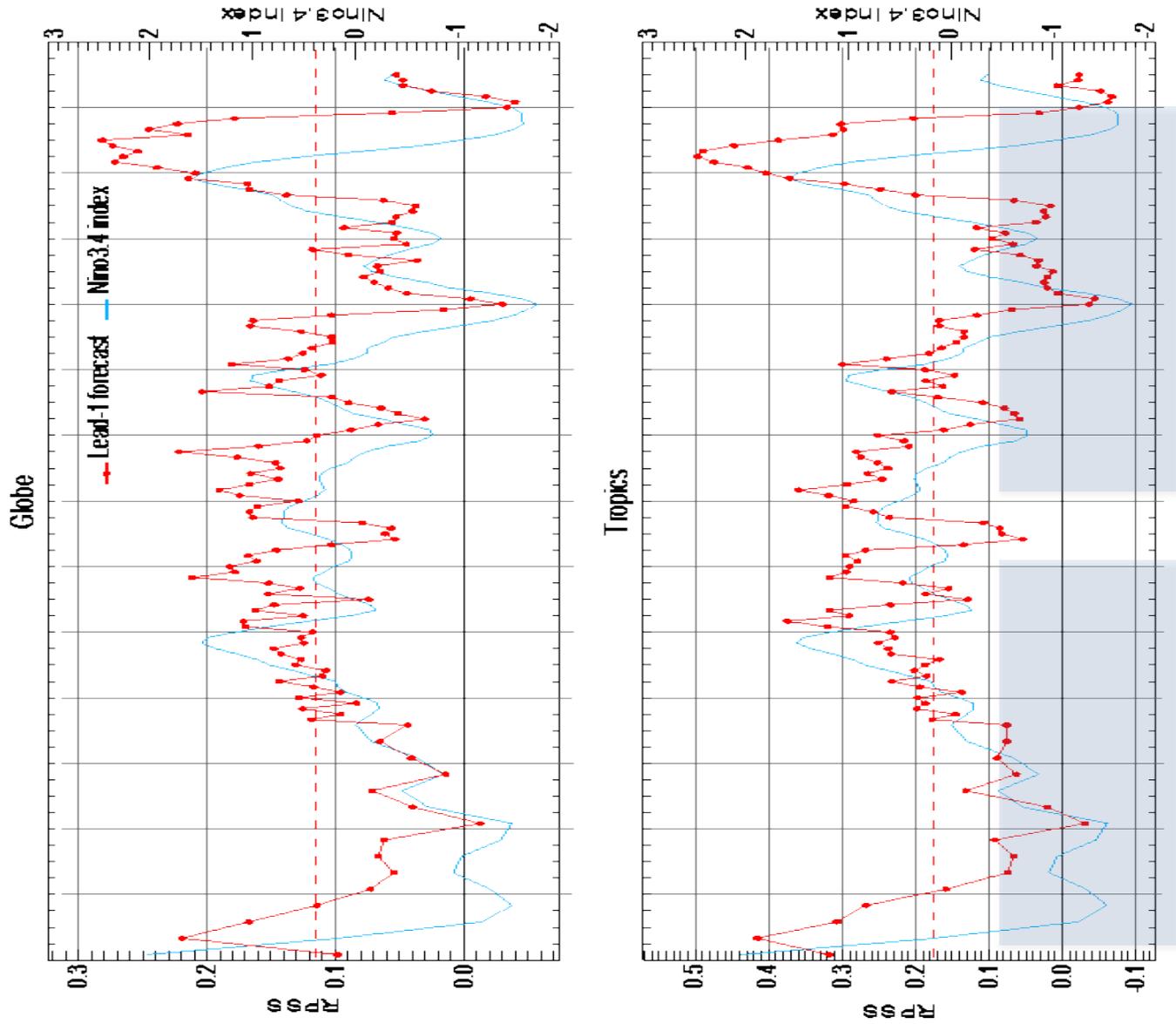
EOFs  
# by AIC



# IRI Forecast Methodology

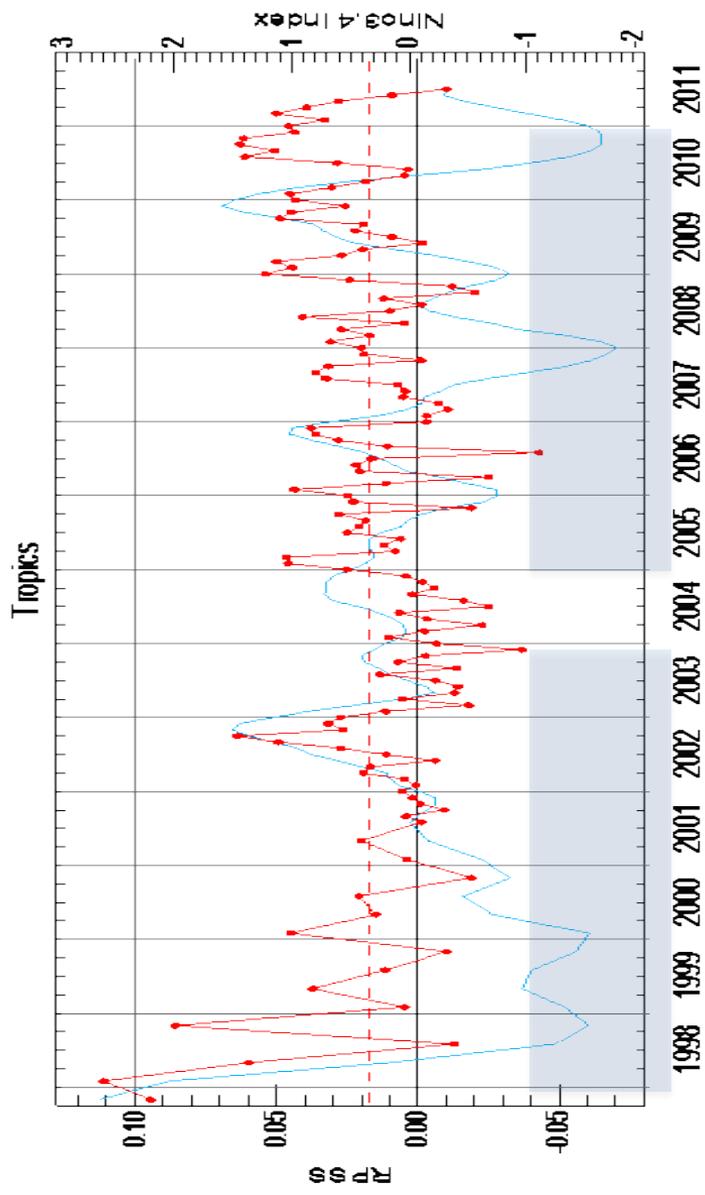
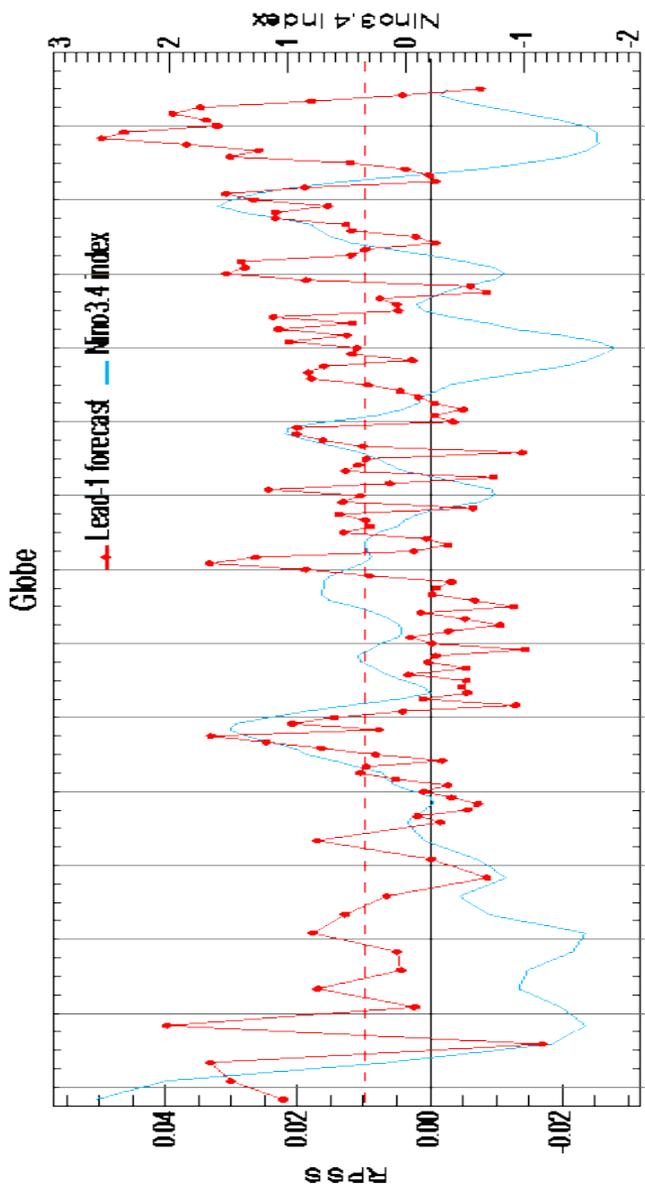
- Atmospheric GCMs forced with forecast SST scenarios
  - Mean of CFSv1, CA and LDEO
  - Positive and negative scenarios based on historical error
- Pattern-based correction of individual model ensemble means.
  - Regression based on historical model runs
    - Forecast SST (CA)
    - Observed SST
  - Spread estimate from historical forecasts with forecast SST.
- Equal weighting of corrected models
- Forecast probabilities
  - Gaussian distribution for temperature
  - Transformed Gaussian for precipitation

**Time Series of RPSS for Temperature (Lead-1)**



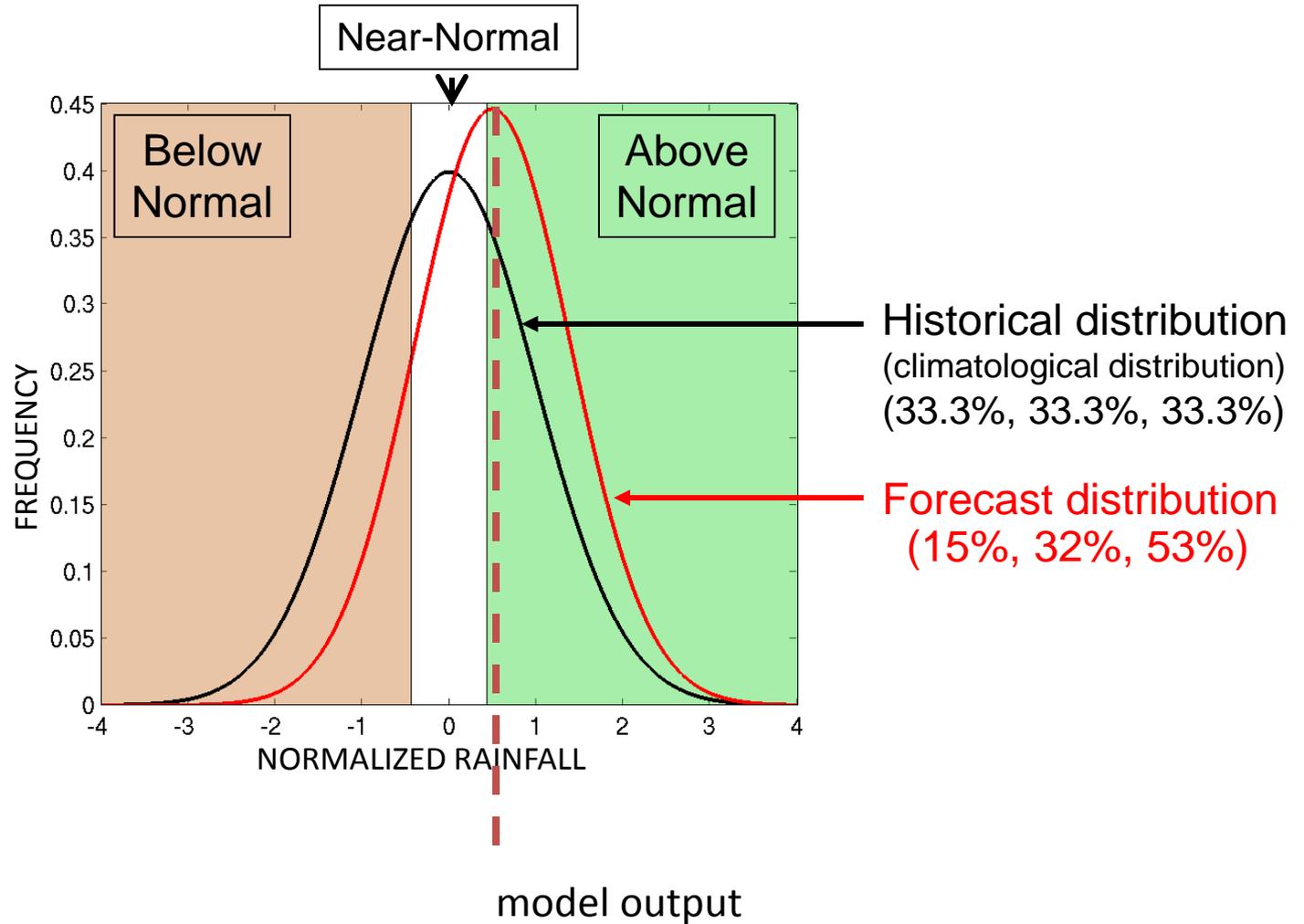
1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011

Time Series of RPSS for Precipitation (Lead-1)



# Constructing forecast probabilities beyond tercile categories

# Shifted Gaussian



# Flexible format map room



Data Library

Flexible  
Forecasts

Temperature

Temperature

Africa

Asia Indonesia

Australia

Central  
America

Europe

Global

Middle East

North America

South America

help@iri

Printable Page

## IRI Seasonal Temperature *Flexible* Forecast

Forecast issued: May 2002

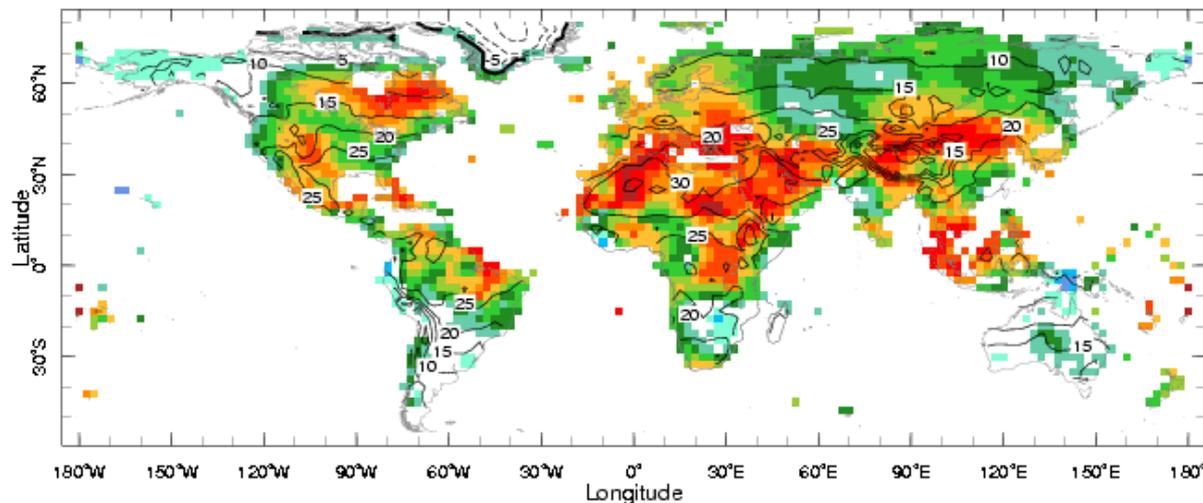
Target Season:  JJA  JAS  ASO  SON

Probability of  exceedance  non-exceedance

Percentile threshold: 50 %-ile

Climatology: 1970 - 1999

80N

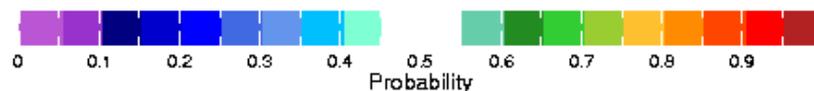


55S

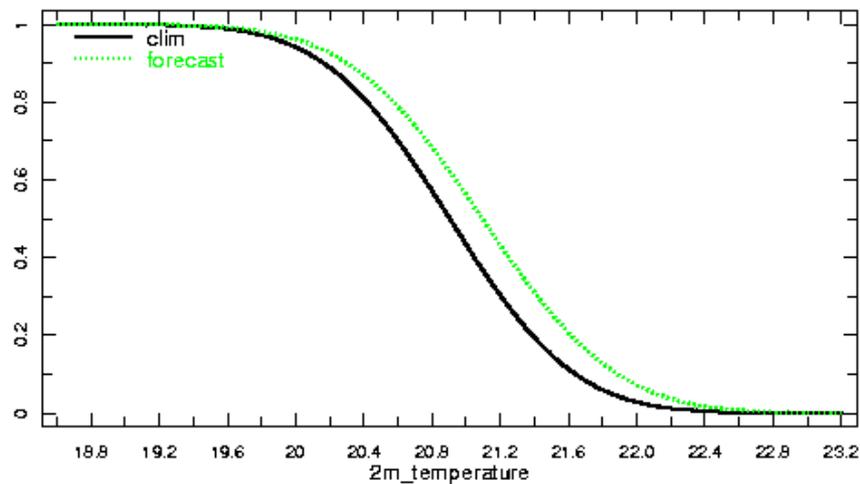
JAS IRI Seasonal Temperature Forecast issued May 2002

-181.25

181.25

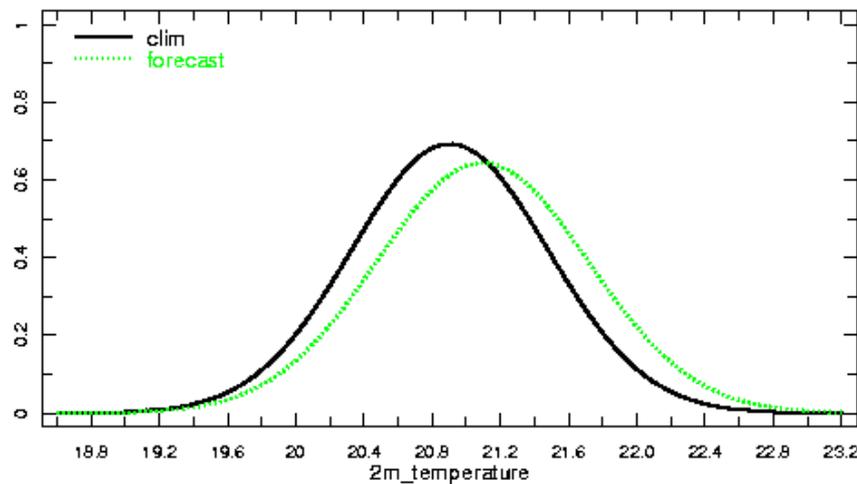


# Flexible format map room



Longitude 122.5W Latitude 40N lead 3.5 months start 0000 1 May

Probability of Exceedance



Longitude 122.5W Latitude 40N lead 3.5 months start 0000 1 May

Probability Distribution

# Summary

- Shrinkage improves reliability of regression forecasts
- Penalty methods can be used to choose EOFs
- Forecast pdfs can be constructed from regression models.
- Complete forecast information provided.