Seasonal Forecasts: how do they work and how we assess their value?

Laura Ferranti
European Centre for Medium-Range Weather Forecasts

Contents:

- Seasonal forecasting with coupled Global Circulation Models
  - Why use GCMs?
  - How we make a forecast
  - Basic calibration – how, why, and what are the problems?
  - Model error

- Operational forecasts: ECMWF System 4
  - System design
  - How good are the El Nino forecasts?
  - How good are the atmospheric forecasts?

- Forecast interpretation
  - Meaning of probabilistic forecasts
  - Outlook

How to make seasonal forecasts:

- Empirical forecasting – **Statistical models**
  - Use past observational record and statistical methods
  - Works with reality instead of error-prone numerical models ➽
  - Limited number of past cases means that it works best when observed variability is dominated by a single source of predictability ➻
  - A non-stationary climate is problematic ➻

- Two-tier forecast systems – **Atmospheric model forced by SST**
  - First predict SST anomalies (ENSO or global; dynamical or statistical)
  - Use ensemble of atmosphere GCMs to predict global response
  - Many people still use regression of a predicted El Nino index on a local variable of interest

- Single-tier GCM forecasts - **Ocean-atmosphere coupled system**
  - Include comprehensive range of sources of predictability ➽
  - Predict joint evolution of SST and atmosphere flow ➽
  - Includes indeterminacy of future SST, important for prob. forecasts ➽
  - Model errors are an issue! ➻

Sources of seasonal predictability

Atmospheric predictability arises from slow variations in lower-boundary forcing

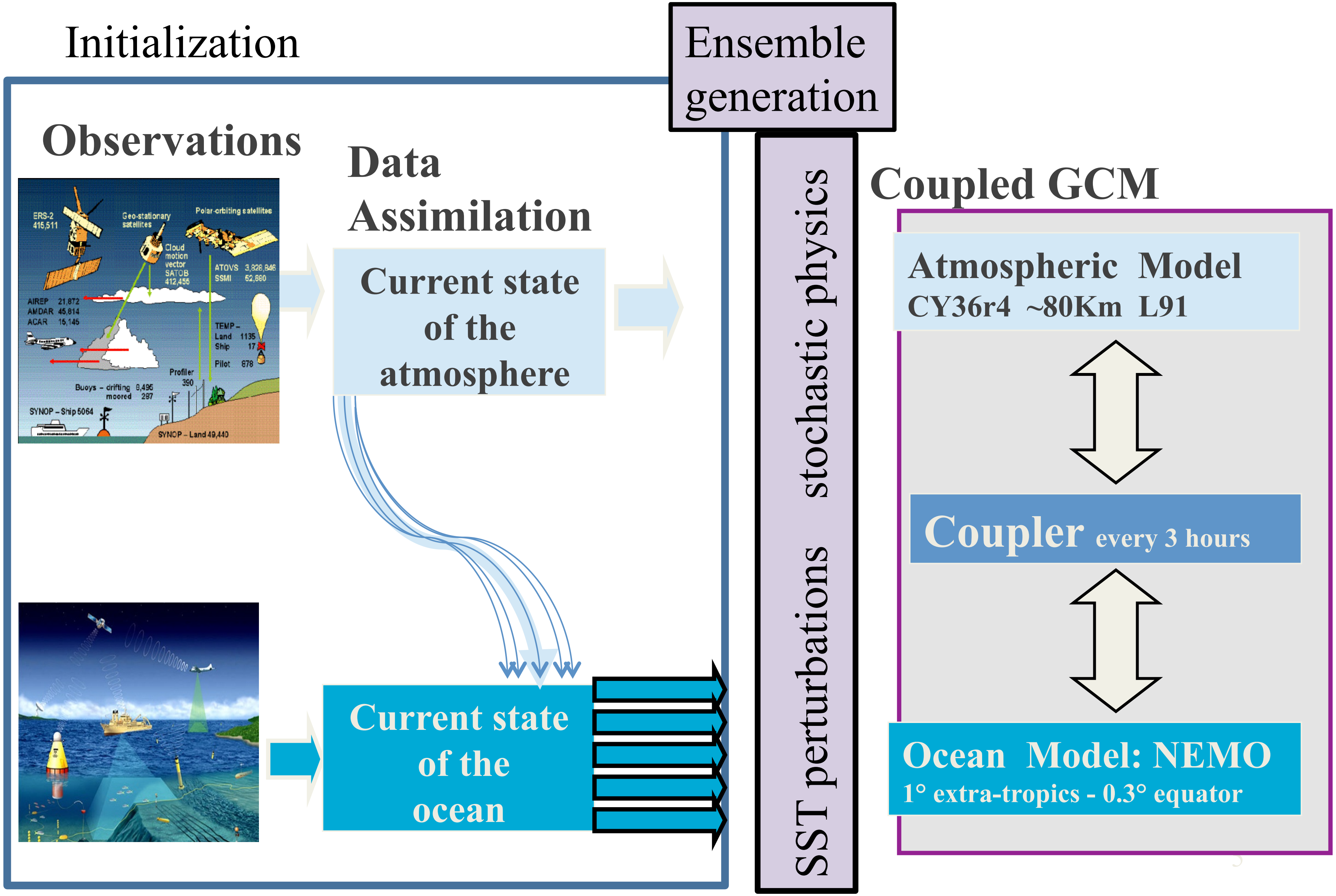
KNOWN TO BE IMPORTANT:

- **El Nino variability** → **biggest single signal**
- **Other tropical ocean SST** - important, but multifarious
- **Climate change** - trends in mid-latitudes
- **Local land surface conditions** - e.g. soil moisture in 2003

OTHER FACTORS:

- Mid-latitude ocean temperatures - always controversial
- Remote soil moisture/snow cover - not yet well established
- Sea ice anomalies - local effects, but remote??
- Dynamic memory of atmosphere - most likely on 1-2 months
- Stratospheric influences - downward propagation of anomalies
- Volcanic eruptions - potentially predictable if contained in initial conditions

ECMWF Seasonal Forecasting System



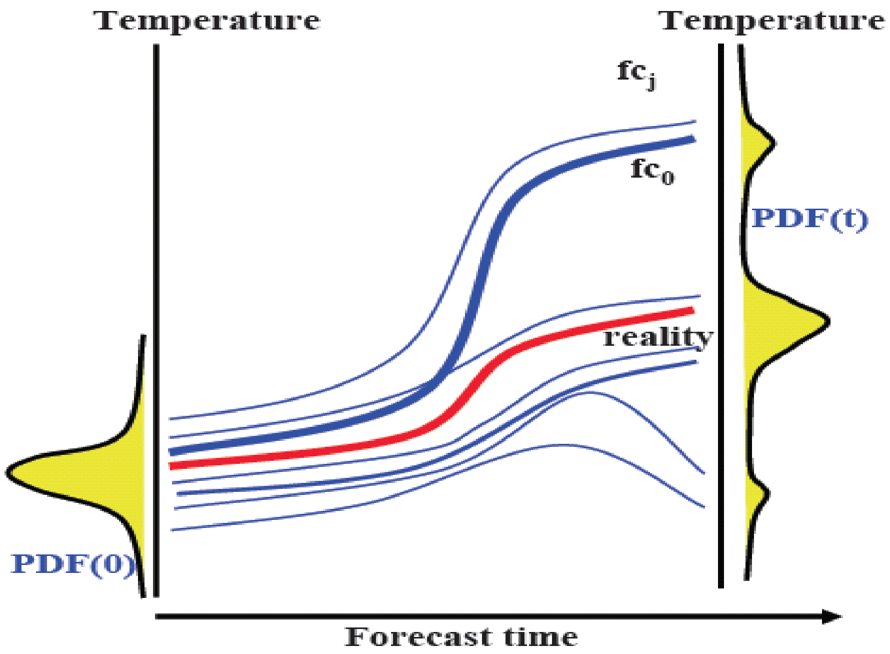
Ensemble forecast:

By sampling the uncertainties at the initial state and running a number of perturbed forecasts it is possible to estimate the probability distribution at future time.

The diversity of these perturbed forecasts (spread) represents the forecast uncertainties.

An ensemble is efficient in predicting forecast uncertainties if the observations is indistinguishable from the ensemble members.

The performance of the ensemble is therefore assessed by the relation between the spread and the mean error for a long term sample.



An ensemble of 51 seasonal forecasts is created every month. Each forecast member is 7 months long.

RMSE and spread in different systems

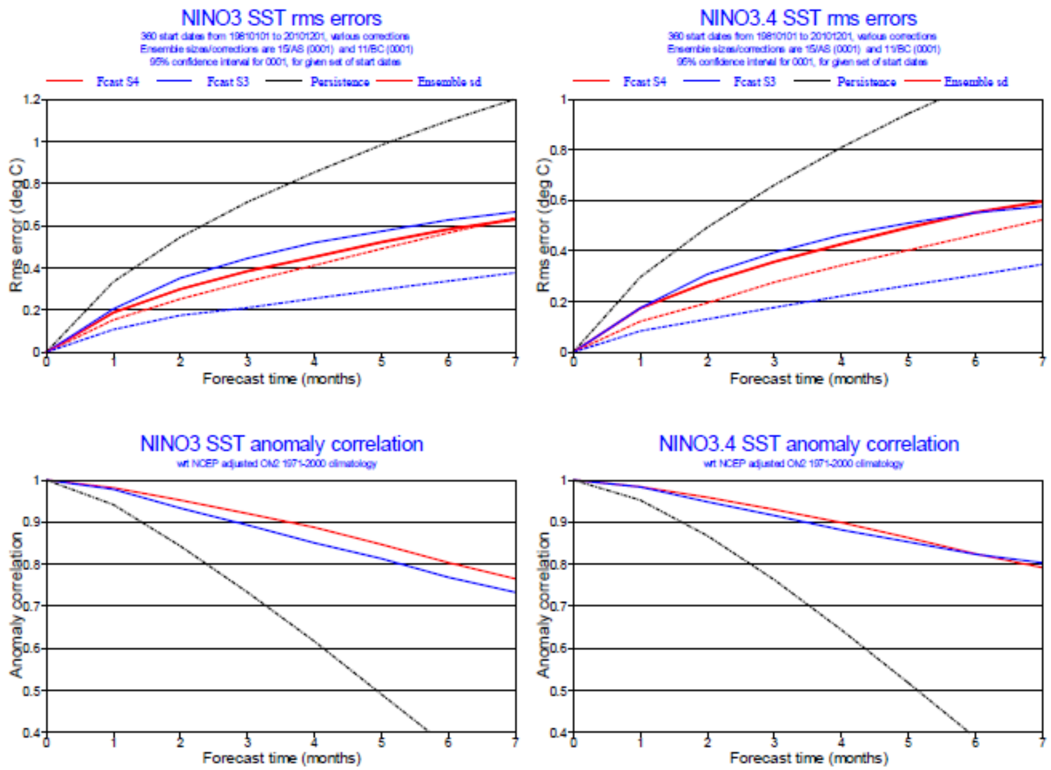


Figure4.1.1. S4 (red) and S3 (blue) NINO3 and NINO3.4 SST scores for the 30 year re-forecast period. S4 has decreased error (solid line) and increased ensemble spread (dashed line).

Rms error of forecasts has been systematically reduced ….

.. but ensemble spread (dashed lines) is still substantially less than actual forecast error.

Calibration - Product generation:

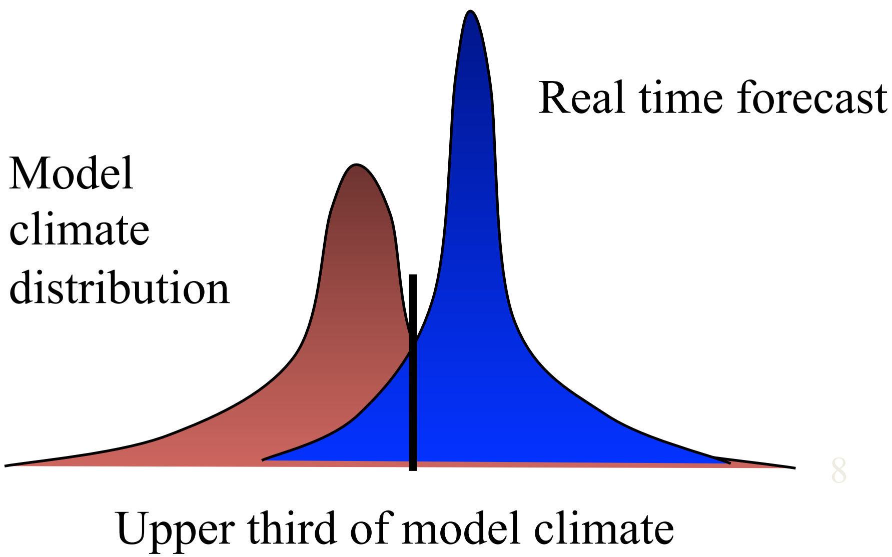
Model drift is a general problem to all GCM. A way to deal with the model drift is to express the forecast with reference to the model climate.

Model climate are estimated from a set of past forecasts (reforecast)

- 30 years of forecasts (1981-2010), all of which use a 15 member ensemble.
- Model climate has both a mean and a distribution, allowing us to estimate e.g. tercile boundaries.
- Model climate is a function of start date and forecast lead time.
- EXCEPTION:	Nino SST indices are bias corrected to absolute values, and anomalies are displayed w.r.t. a 1971-2000 climate. Nino SST indices have been scaled so that also variance the model variance matches the observed variance in the calibration period

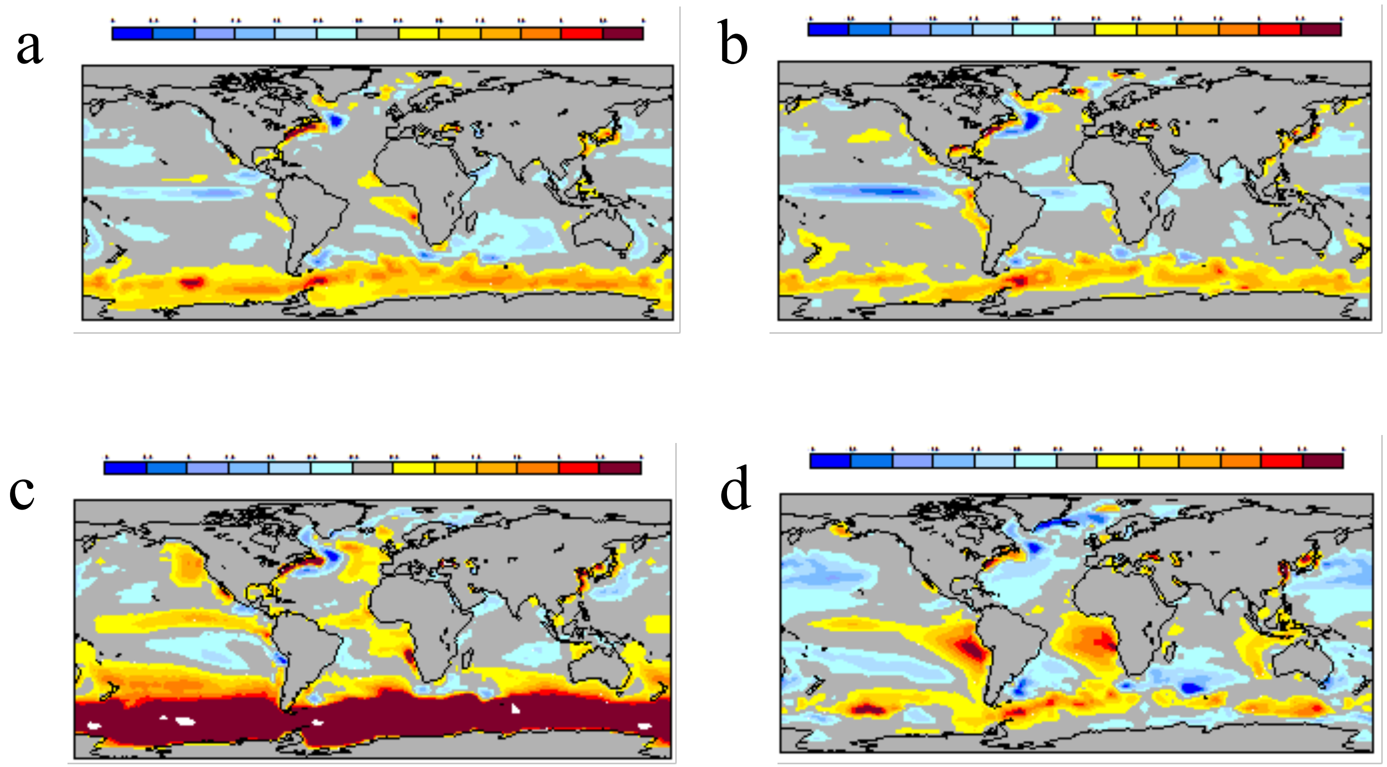
Implicit assumption of linearity.

Seasonal Forecasts are made relative to model climate

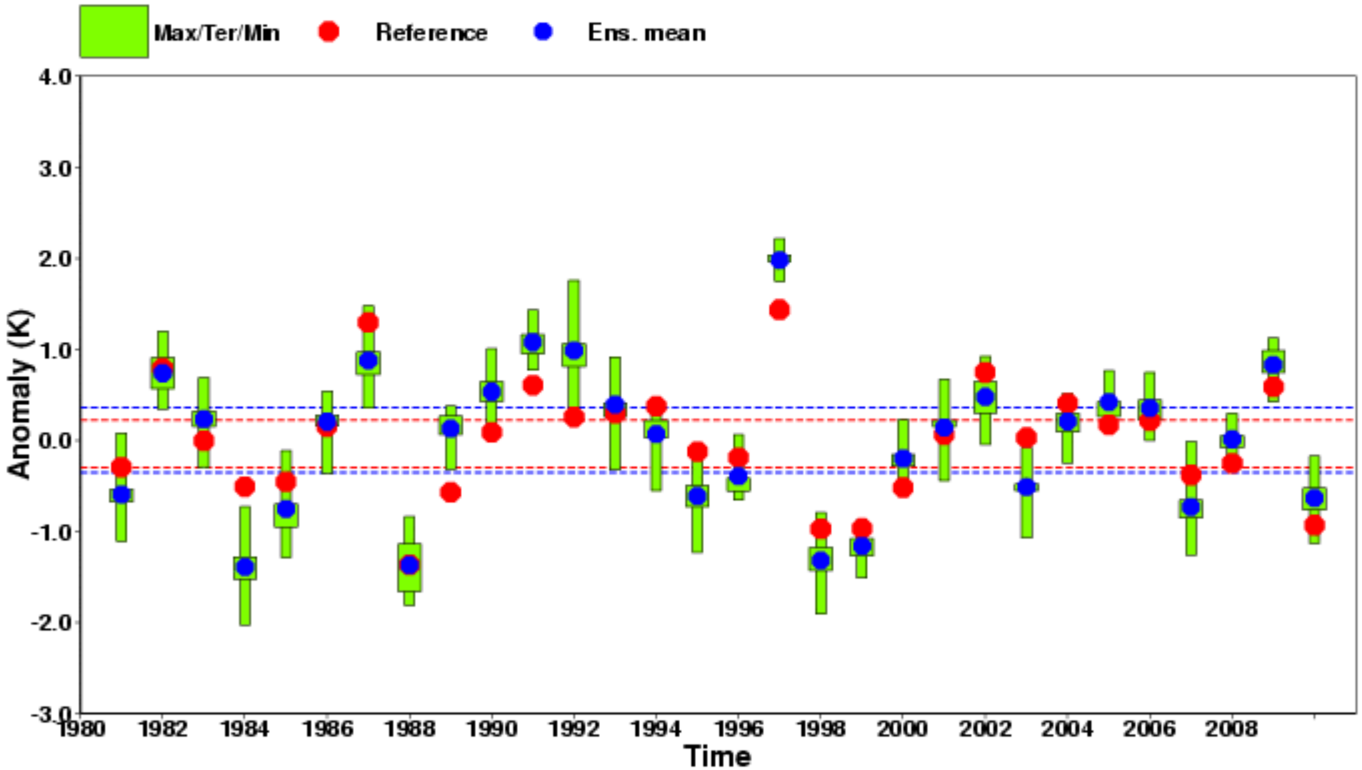


SST Biases for DJF

Biases from 4 independent coupled systems included in the EUROSIP multi-model (1996-2009)



Nino 3.4 SST anomalies:



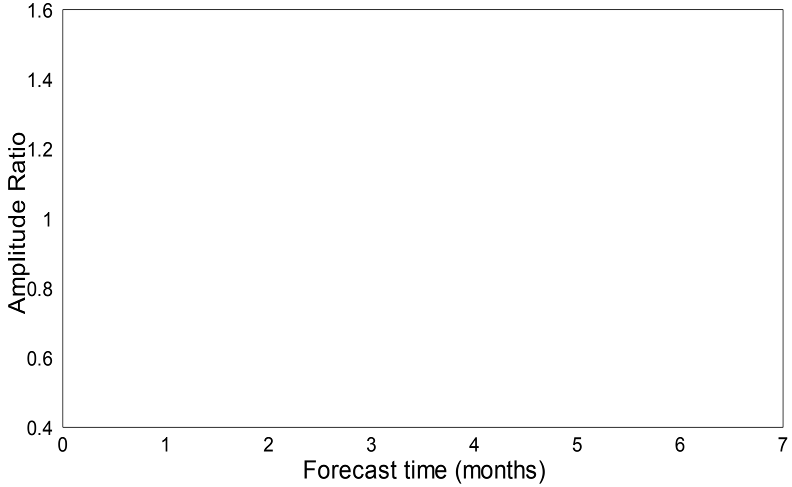
NINO 3.4	Correlation	Sd m/Sd obs
System 4	0.88	1.30

Nino plumes: variance scaling

- Model Nino SST anomalies in S4 have **too large amplitude**
- Problem is especially acute in boreal spring and early summer (model bias of "permanent La Nina" does not allow spring relaxation physics to apply; this was something S3 did very well)
- We plot the "Nino plumes" corrected for both mean **and** variance, instead of just the mean.
- This is done by scaling the model anomalies so that the model variance matches the observed variance in the calibration period
- We use the same approach (cross-validated) when calculating scores
- This affects the *plotting*, not the data
- The spatial maps are not affected: the tercile and quintile probability maps are already implicitly standardized wrt model variance
- Also can be used for other models, eg in multi-model system

Calibration of ENSO SST indices

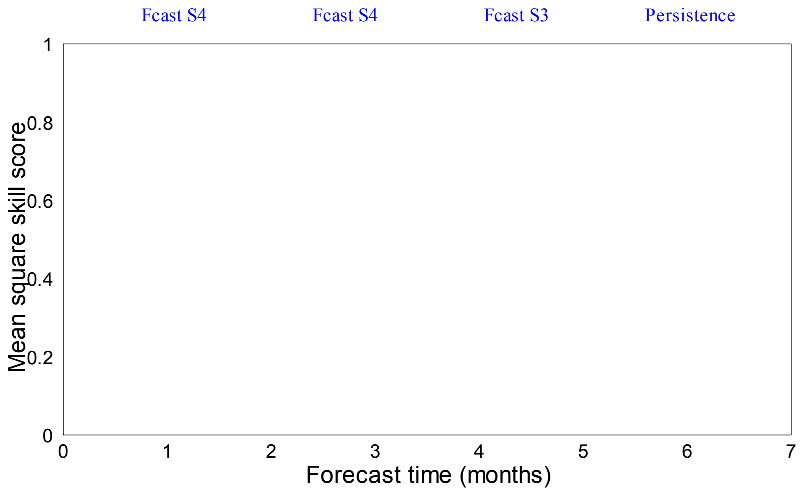
NINO3 SST anomaly amplitude ratio



NINO3 SST mean square skill scores

150 start dates from 19910201 to 20081101, various corrections

Ensemble sizes are 15 (0001), 11 (0001) and 11 (0001)



S4 non calibrated

S4 calibrated

S3

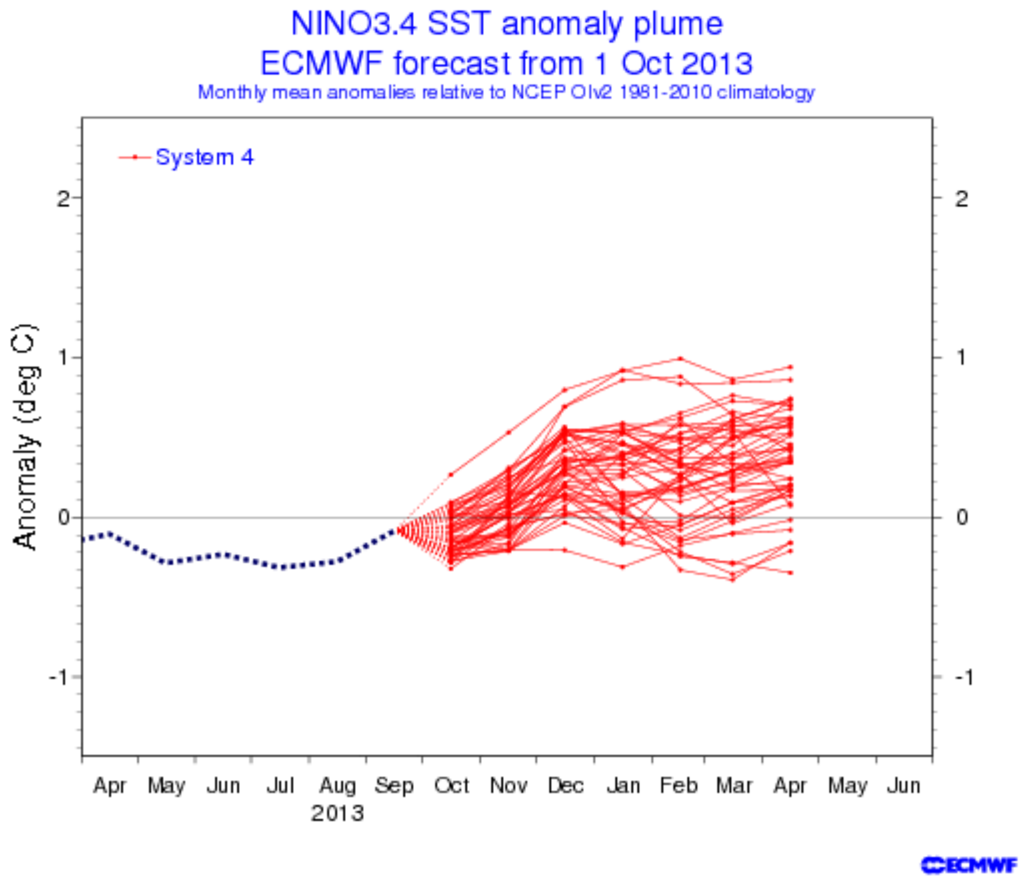
Seasonal Forecast Products :

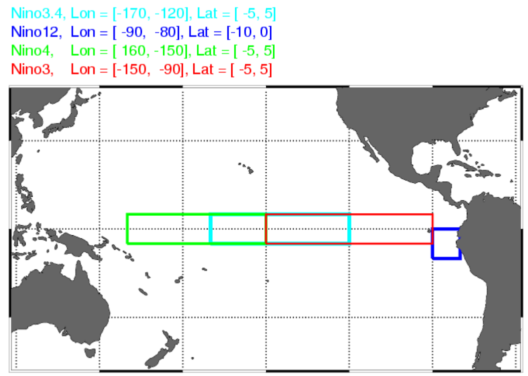
Seasonal outlook (graphic products):

(up to 7 months ahead)

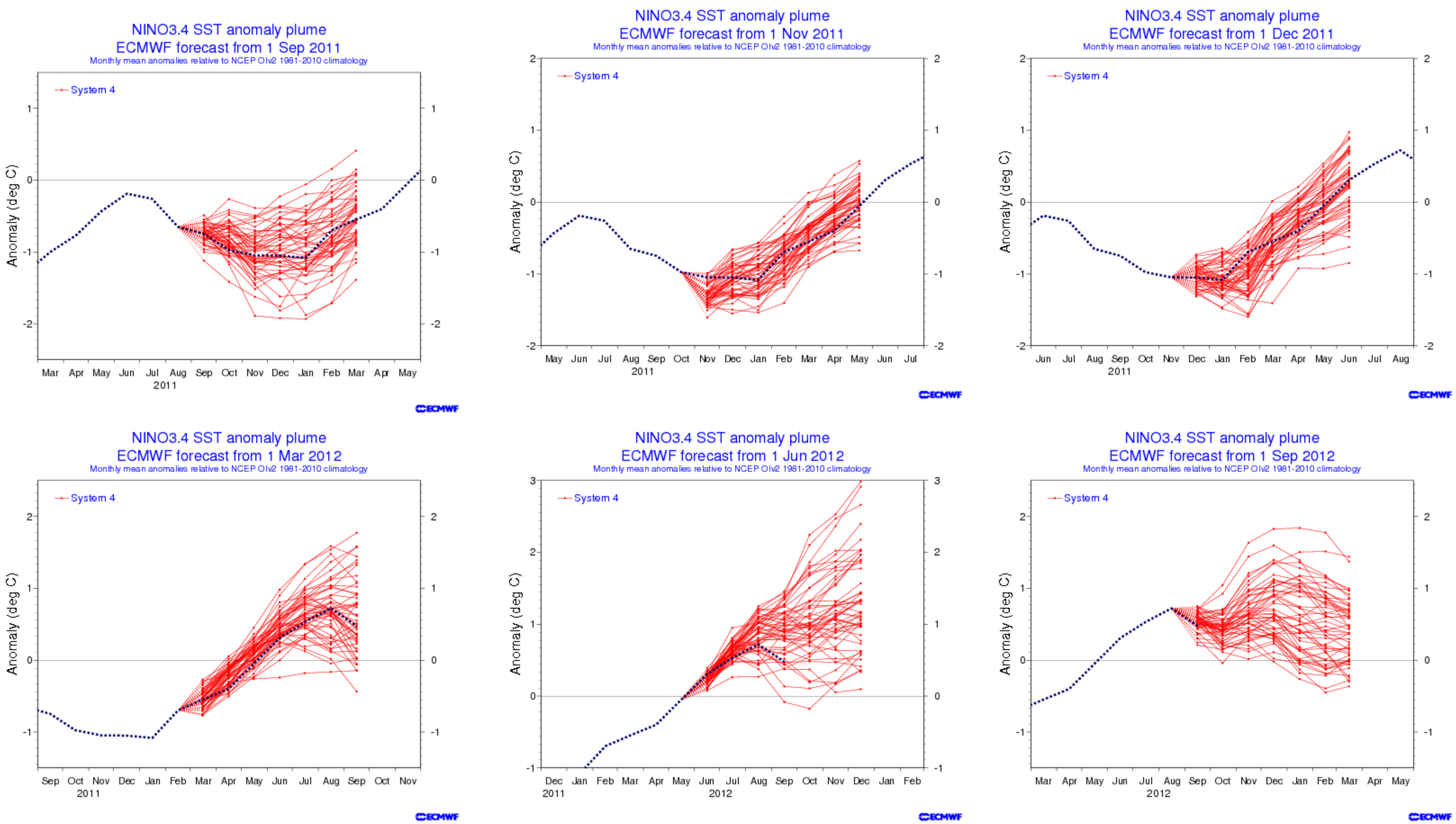
- Forecasts for Nino3, Nino3.4 and Nino4
- Spatial plots (ens.mean anomaly, terciles ..)
- Climagrams (similar to Epsgrams)
- Tropical storms

Nino 3.4 predictions:

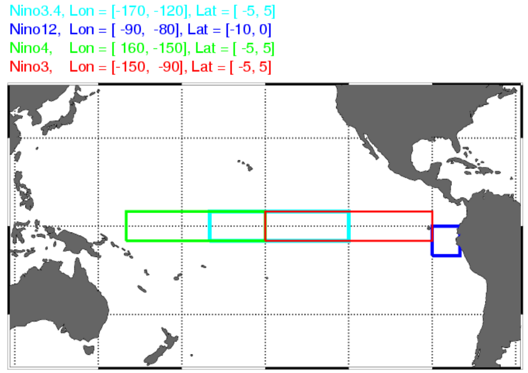




Seasonal forecasts
sea surface temperature, tropical Pacific



SST deterministic scores:



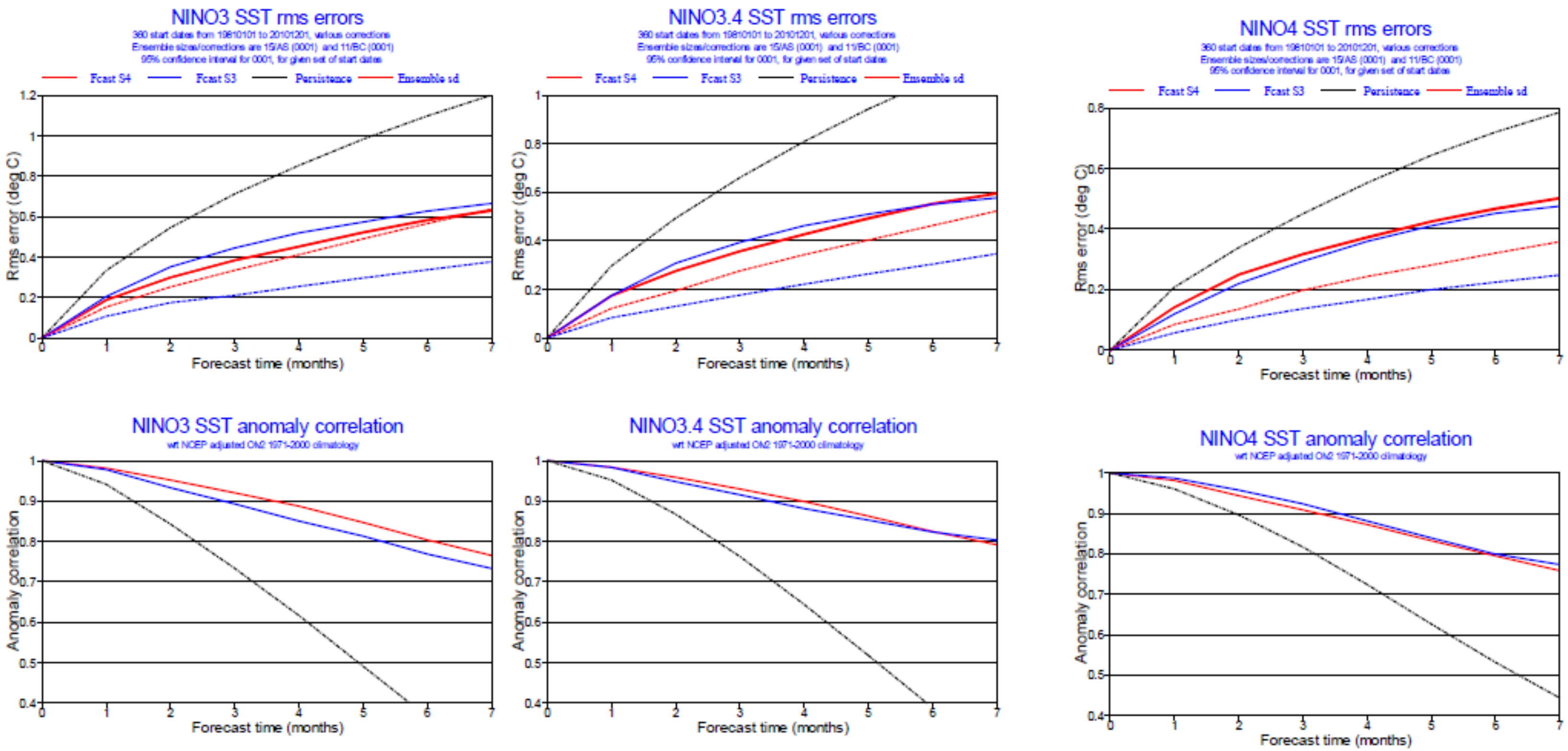
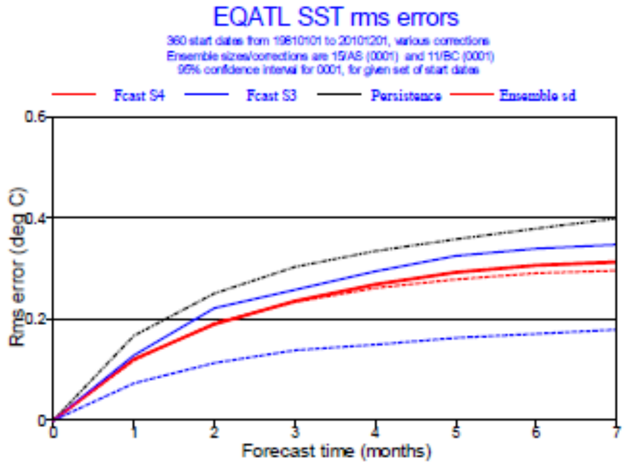
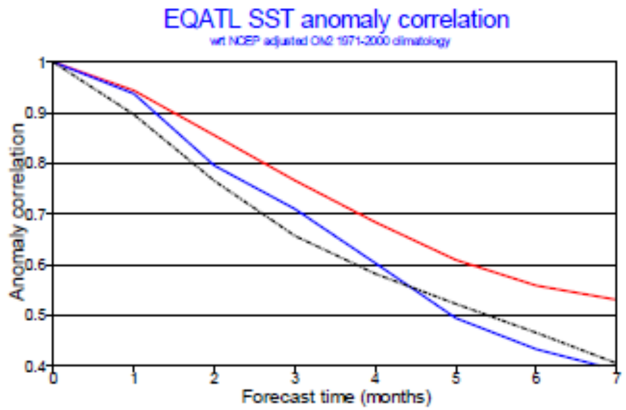


Figure4.1.1. S4 (red) and S3 (blue) NINO3 and NINO3.4 SST scores for the 30 year re-forecast period. S4 has decreased error (solid line) and increased ensemble spread (dashed line).

SST deterministic scores: Eq. Atlantic





torial Atlantic SST.

Seasonal Forecast Products :

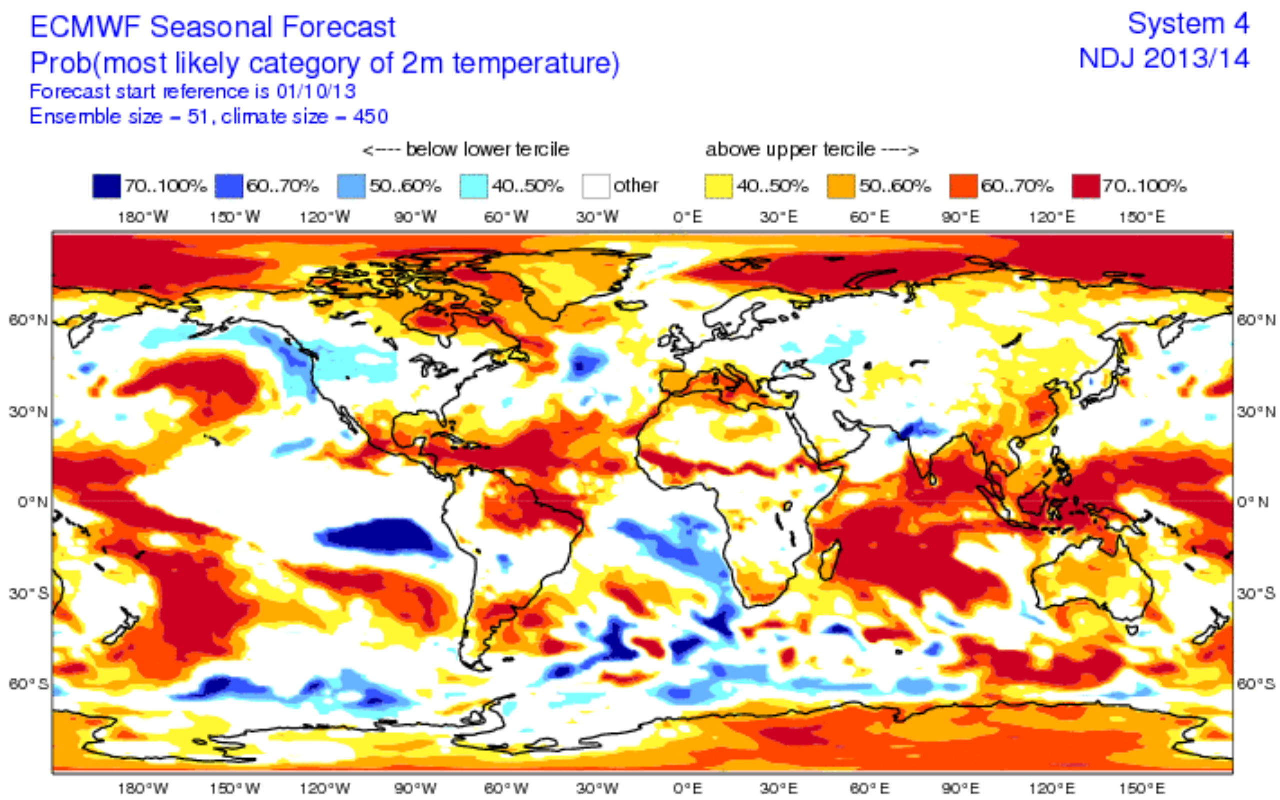
Seasonal outlook (graphic products):

(up to 7 months ahead)

- Forecasts for Nino3, Nino3.4 and Nino4
- Spatial plots (ens.mean anomaly, terciles ..)
- Climagrams (similar to Epsgrams)
- Tropical storms

Seasonal forecast charts :

Spatial maps representing the seasonal forecast in terms of model probabilities stratified by terciles.



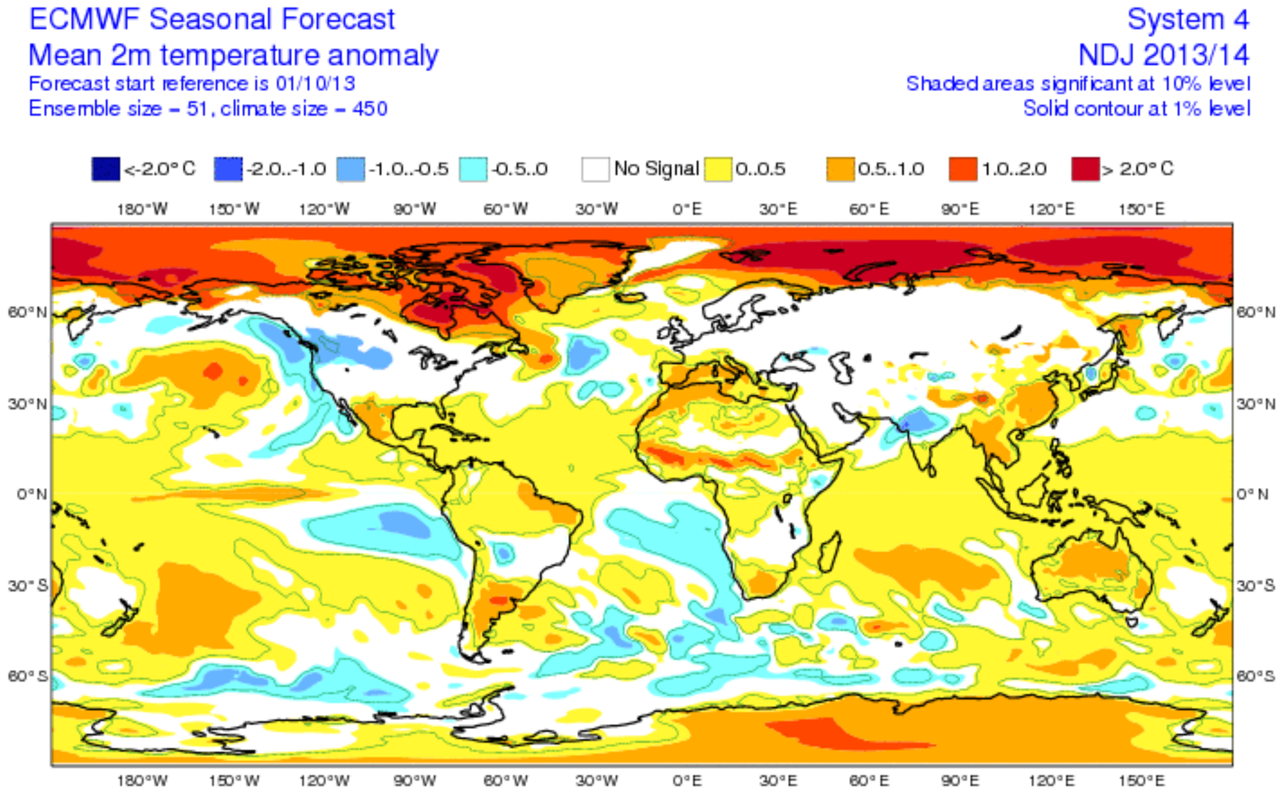
Available parameters are:

- 2m Temperature
- Mean sea level pressure
- Precipitation
- Sea surface temperature
- 850 hPa temperature
- 500 hPa geopotential

Forecast is made available on the 8th of each month.

Seasonal forecast charts :

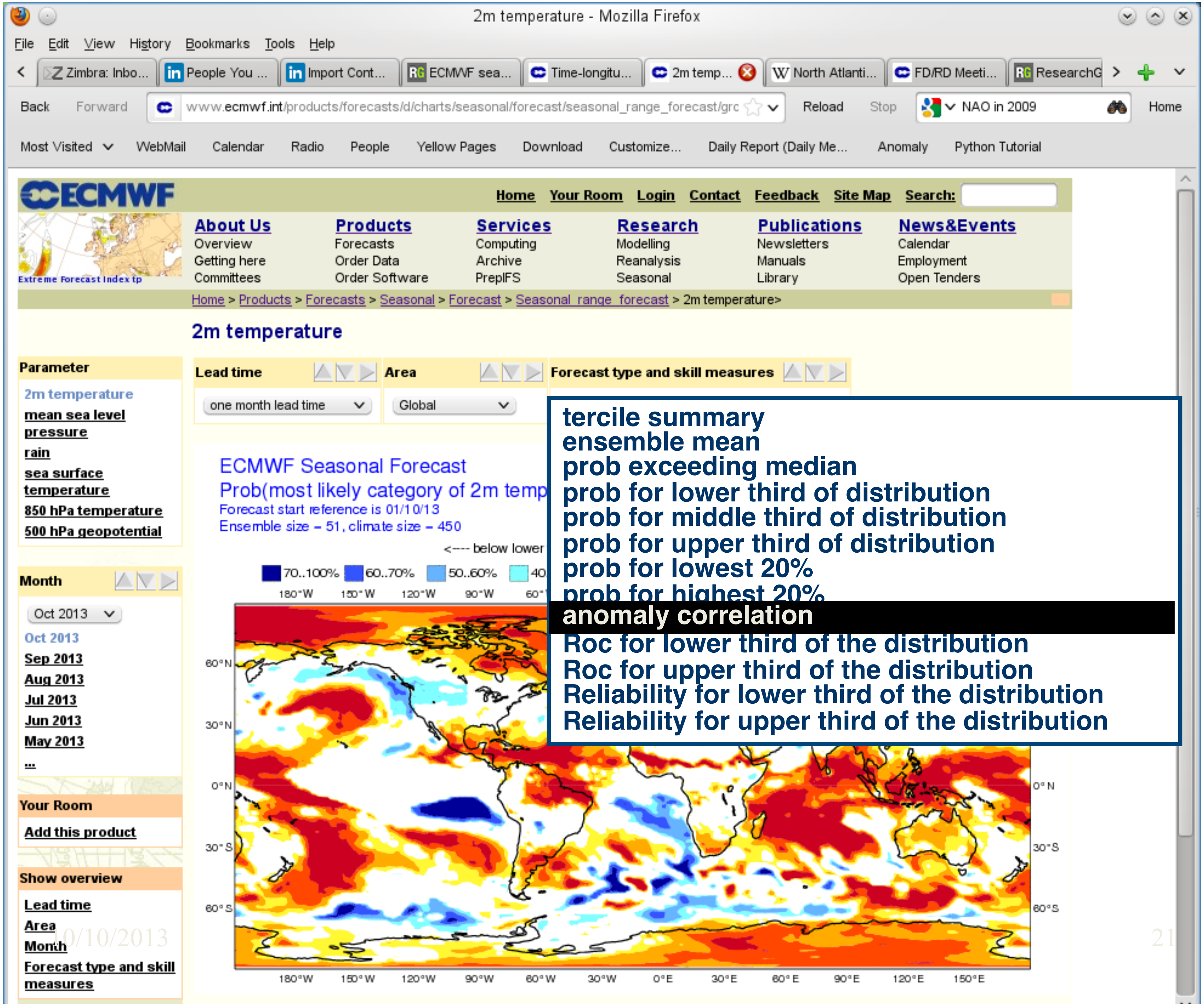
Spatial maps representing the seasonal forecast in terms of model probabilities stratified by terciles.



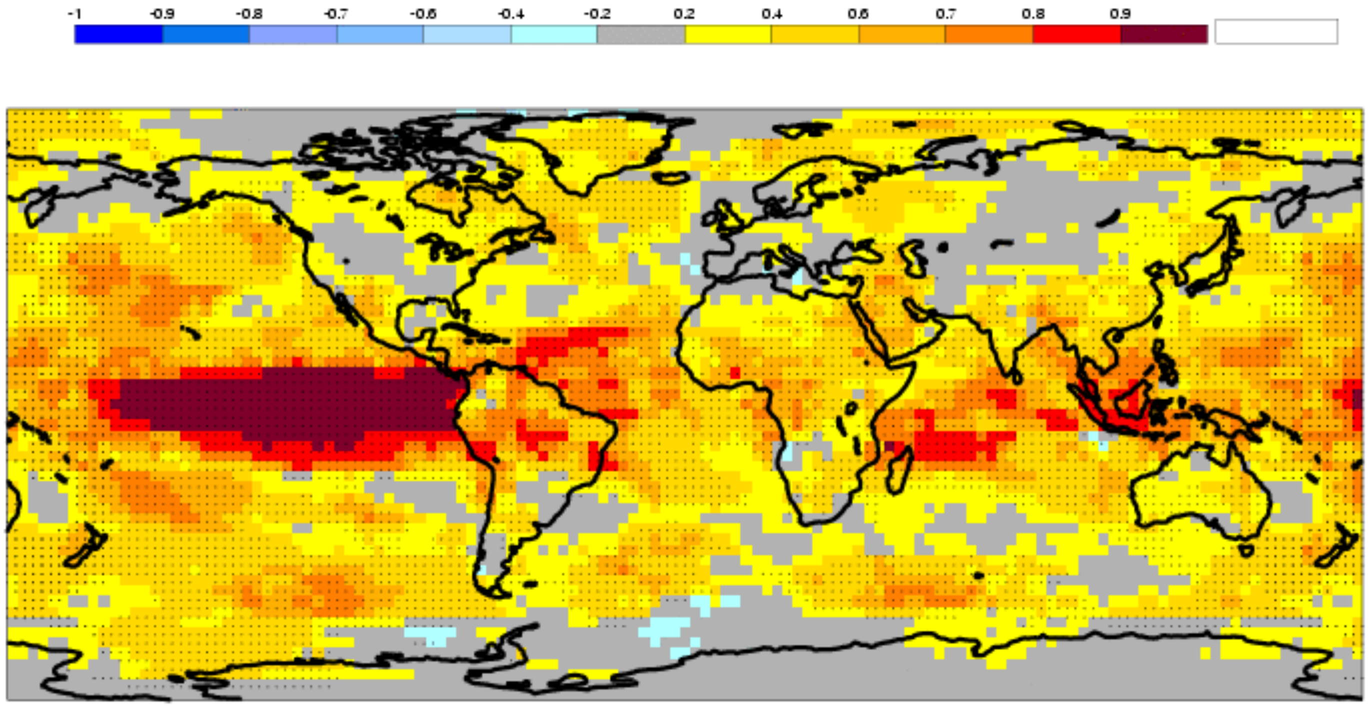
Available parameters are:

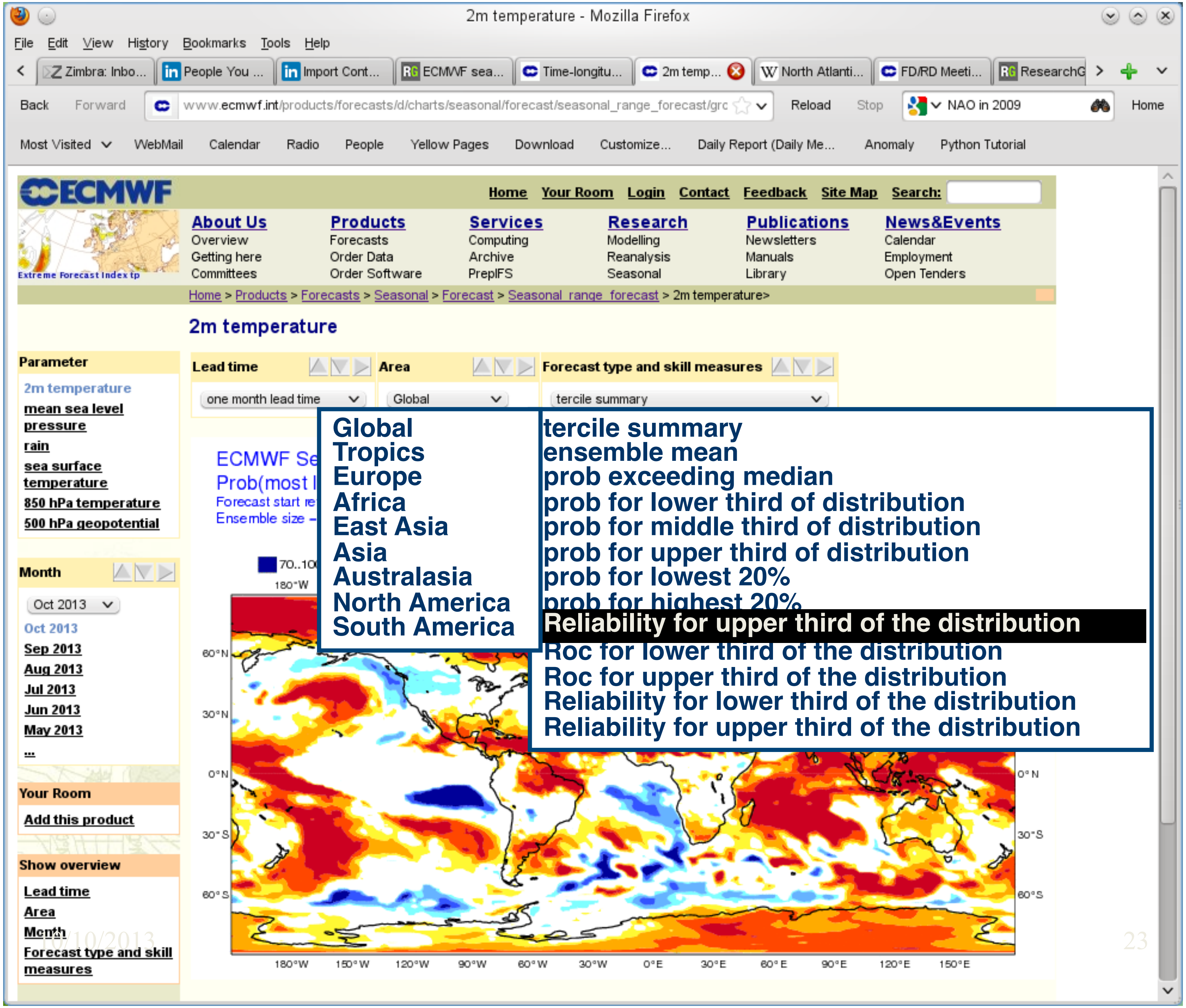
- 2m Temperature
- Mean sea level pressure
- Precipitation
- Sea surface temperature
- 850 hPa temperature
- 500 hPa geopotential

Forecast is made available on the 8th of each month.


2m temperature - Mozilla Firefox
File Edit View History Bookmarks Tools Help
www.ecmwf.int/products/forecasts/d/charts/seasonal/forecast/seasonal_range_forecast/grc
ECMWF
Home Your Room Login Contact Feedback Site Map Search:
About Us
Overview
Getting here
Committees
Products
Forecasts
Order Data
Order Software
Services
Computing
Archive
PrepIFS
Research
Modelling
Reanalysis
Seasonal
Publications
Newsletters
Manuals
Library
News&Events
Calendar
Employment
Open Tenders
Home > Products > Forecasts > Seasonal > Forecast > Seasonal_range_forecast > 2m temperature>
2m temperature
Parameter
2m temperature
mean sea level pressure
rain
sea surface temperature
850 hPa temperature
500 hPa geopotential
Lead time
one month lead time
Area
Global
Forecast type and skill measures
Month
Oct 2013
Sep 2013
Aug 2013
Jul 2013
Jun 2013
May 2013
Your Room
Add this product
Show overview
Lead time
Area
Month
Forecast type and skill measures
ECMWF Seasonal Forecast
Prob(most likely category of 2m temp
Forecast start reference is 01/10/13
Ensemble size = 51, climate size = 450
tercile summary
ensemble mean
prob exceeding median
prob for lower third of distribution
prob for middle third of distribution
prob for upper third of distribution
prob for lowest 20%
prob for highest 20%
anomaly correlation
Roc for lower third of the distribution
Roc for upper third of the distribution
Reliability for lower third of the distribution
Reliability for upper third of the distribution

Anomaly Correlation Coefficient for ECMWF with 15 ensemble members
Near-surface air temperature
Hindcast period 1981-2010 with start in October average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)





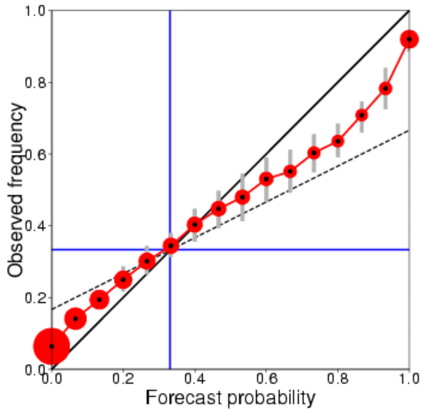
2m temperature - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Zimbra: Inbo... People You ... Import Cont... ECMWF sea... Time-longitu... 2m temp... North Atlanti... FD/RD Meeti... ResearchG
Back Forward www.ecmwf.int/products/forecasts/d/charts/seasonal/forecast/seasonal_range_forecast/grc Reload Stop NAO in 2009 Home
Most Visited WebMail Calendar Radio People Yellow Pages Download Customize... Daily Report (Daily Me... Anomaly Python Tutorial
ECMWF
Home Your Room Login Contact Feedback Site Map Search:
About Us Overview Getting here Committees
Products Forecasts Order Data Order Software
Services Computing Archive PrepIFS
Research Modelling Reanalysis Seasonal
Publications Newsletters Manuals Library
News&Events Calendar Employment Open Tenders
Extreme Forecast Index tp
Home > Products > Forecasts > Seasonal > Forecast > Seasonal_range_forecast > 2m temperature>
2m temperature
Parameter
2m temperature
mean sea level pressure
rain
sea surface temperature
850 hPa temperature
500 hPa geopotential
Lead time
one month lead time
Area
Global
Forecast type and skill measures
tercile summary
Global
Tropics
Europe
Africa
East Asia
Asia
Australasia
North America
South America
tercile summary
ensemble mean
prob exceeding median
prob for lower third of distribution
prob for middle third of distribution
prob for upper third of distribution
prob for lowest 20%
prob for highest 20%
Reliability for upper third of the distribution
Roc for lower third of the distribution
Roc for upper third of the distribution
Reliability for lower third of the distribution
Reliability for upper third of the distribution
ECMWF Se
Prob(most l
Forecast start re
Ensemble size -
70..10
Month
Oct 2013
Oct 2013
Sep 2013
Aug 2013
Jul 2013
Jun 2013
May 2013
...
Your Room
Add this product
Show overview
Lead time
Area
Month
Forecast type and skill measures
60°N 30°N 0°N 30°S 60°S
180°W 150°W 120°W 90°W 60°W 30°W 0°E 30°E 60°E 90°E 120°E 150°E

Forecast skill is different from region to region

DJF 2m temp > upper tercile

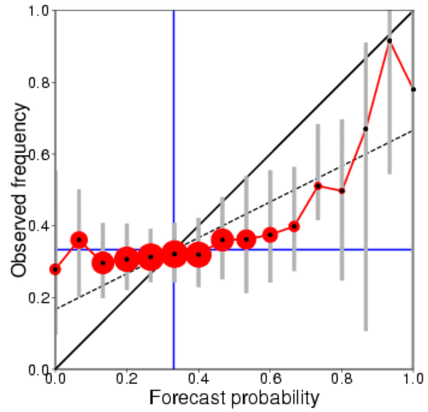
Tropics

Reliability diagram for ECMWF with 15 ensemble members
Near-surface air temperature anomalies above the upper tercile
Accumulated over tropical band (land and sea points)
Hindcast period 1981-2010 with start in November average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: 0.279 (0.197, 0.351)
Reliability skill score: 0.973 (0.956, 0.982)
Resolution skill score: 0.306 (0.231, 0.372)



Europe

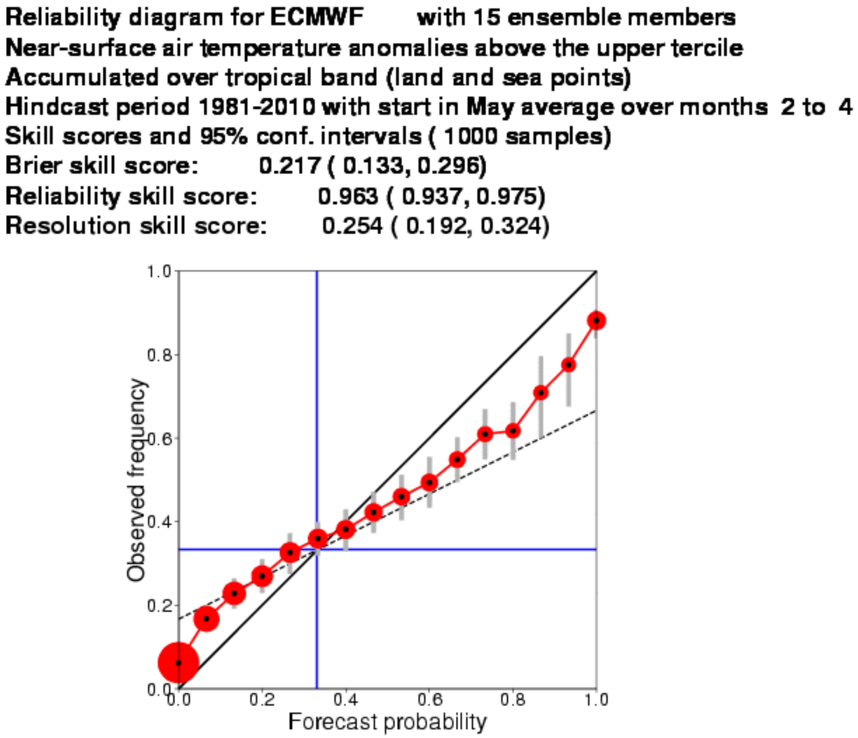
Reliability diagram for ECMWF with 15 ensemble members
Near-surface air temperature anomalies above the upper tercile
Accumulated over Europe (land and sea points)
Hindcast period 1981-2010 with start in November average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: -0.081 (-0.191, 0.011)
Reliability skill score: 0.908 (0.790, 0.965)
Resolution skill score: 0.011 (0.006, 0.053)



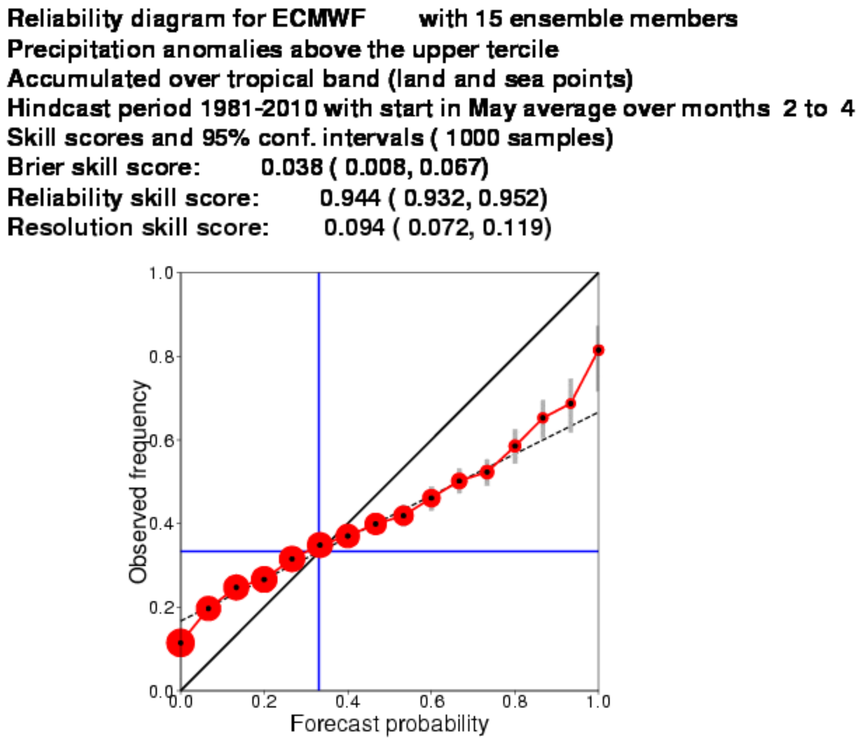
Forecast skill is different for different parameters

JJA Tropics

2m temp > upper tercile



precip>upper tercile

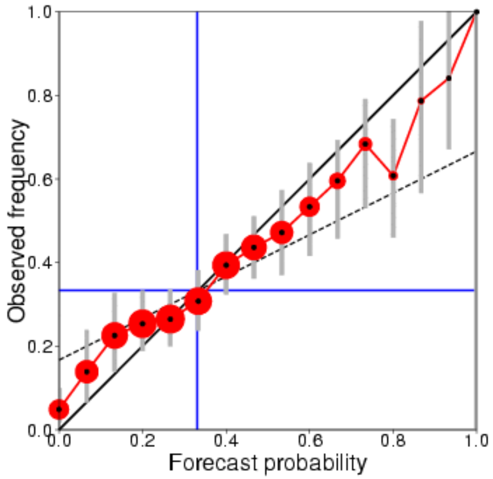


Forecast skill is different for different seasons

2m temp > upper tercile over Europe

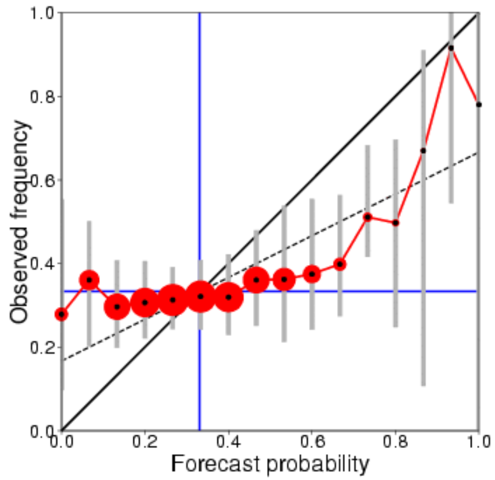
JJA

Reliability diagram for ECMWF with 15 ensemble members
Near-surface air temperature anomalies above the upper tercile
Accumulated over Europe (land and sea points)
Hindcast period 1981-2010 with start in May average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: 0.092 (0.007, 0.162)
Reliability skill score: 0.986 (0.950, 0.994)
Resolution skill score: 0.106 (0.056, 0.173)



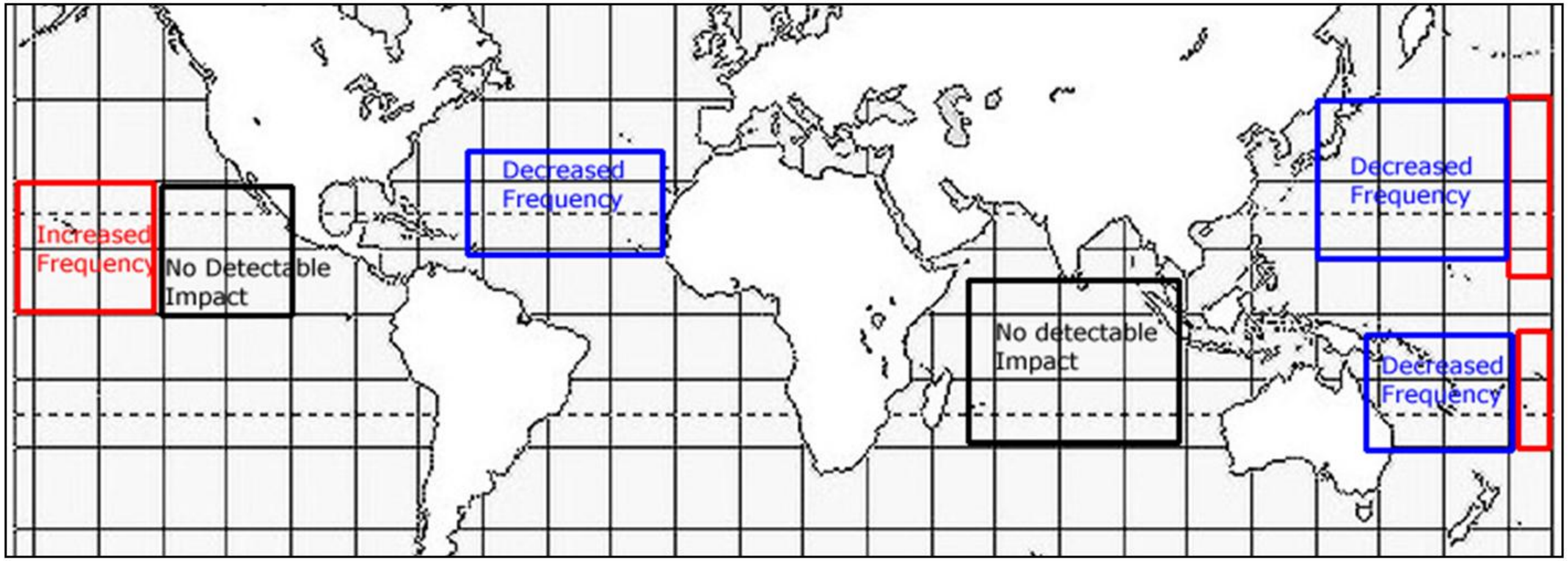
DJF

Reliability diagram for ECMWF with 15 ensemble members
Near-surface air temperature anomalies above the upper tercile
Accumulated over Europe (land and sea points)
Hindcast period 1981-2010 with start in November average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: -0.081 (-0.191, 0.011)
Reliability skill score: 0.908 (0.790, 0.965)
Resolution skill score: 0.011 (0.006, 0.053)



ENSO teleconnection

Tropical cyclone activity

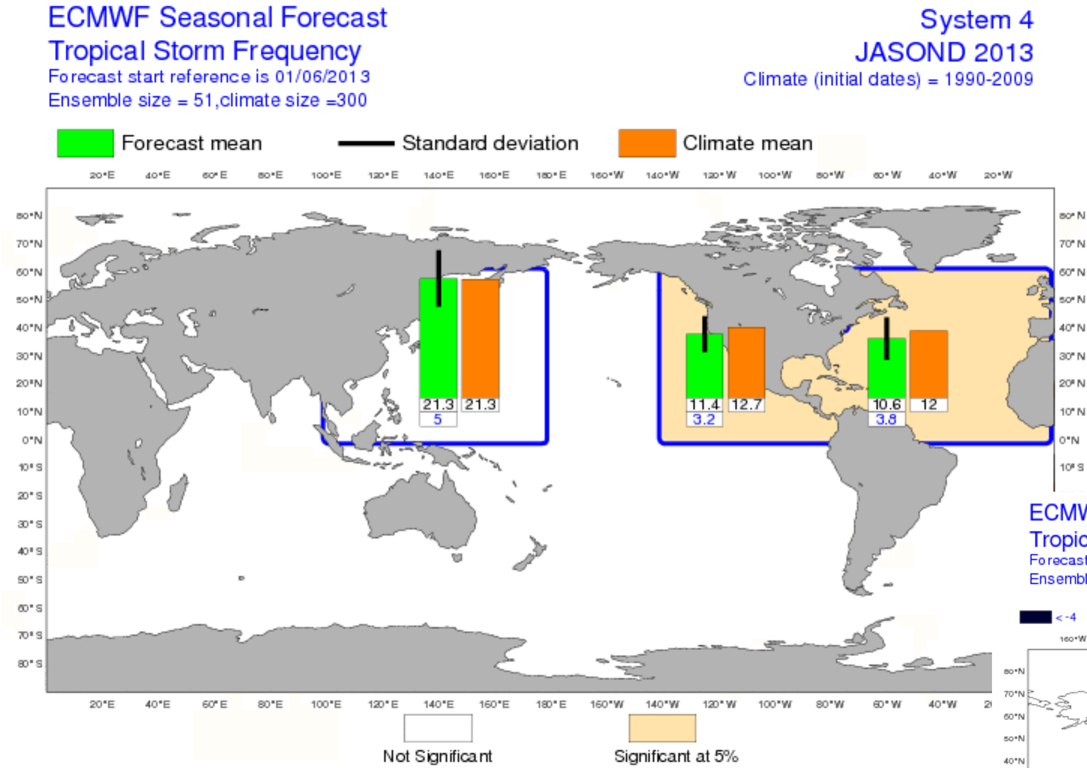


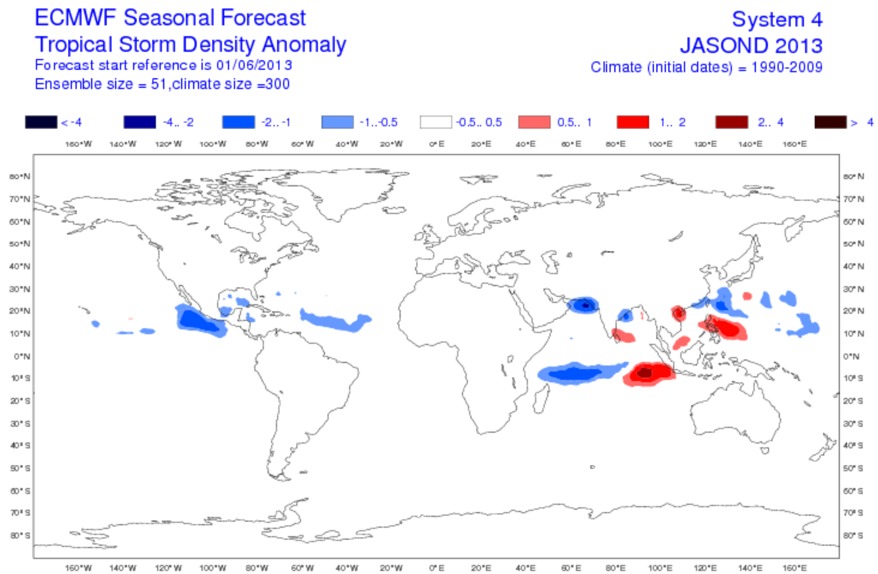
In El Nino year

- **Reduced** frequency :
  - Atlantic: Upper-level westerlies increase, increased vertical wind shear
  - Australia: Convection shifts east, monsoon trough weakens
  - Northwest Pacific (west of 160°E): Monsoon trough shifts away from area
- **Increased** frequency :
  - NW Pacific (east of 160E): Monsoon trough shifts into this area
  - NE Pacific (near Hawaii): Increased convection due to warmer SSTs

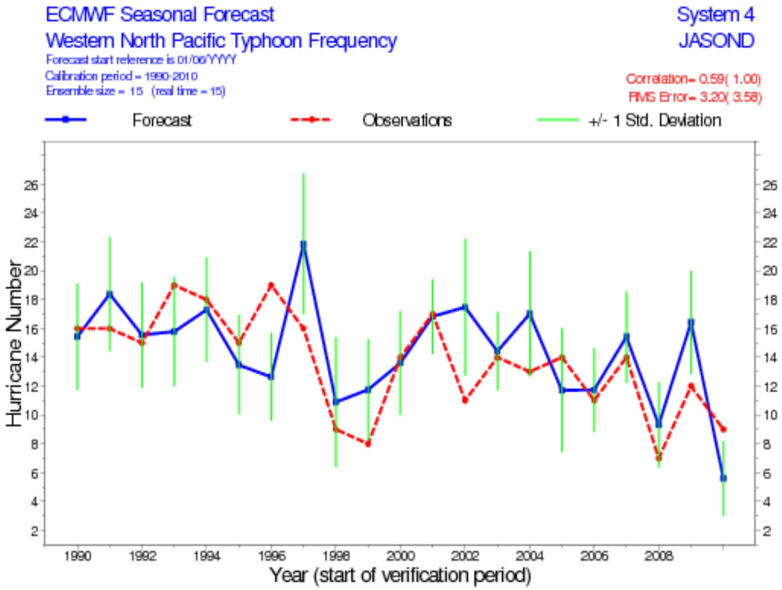
(Gray 1984, Chen et al. 1998; Chan 2000; Chia and Ropelewski 2002; Wang and Chan 2002; Wu et al. 2004; Kim et al. 2005; Camargo et al. 2007, …)

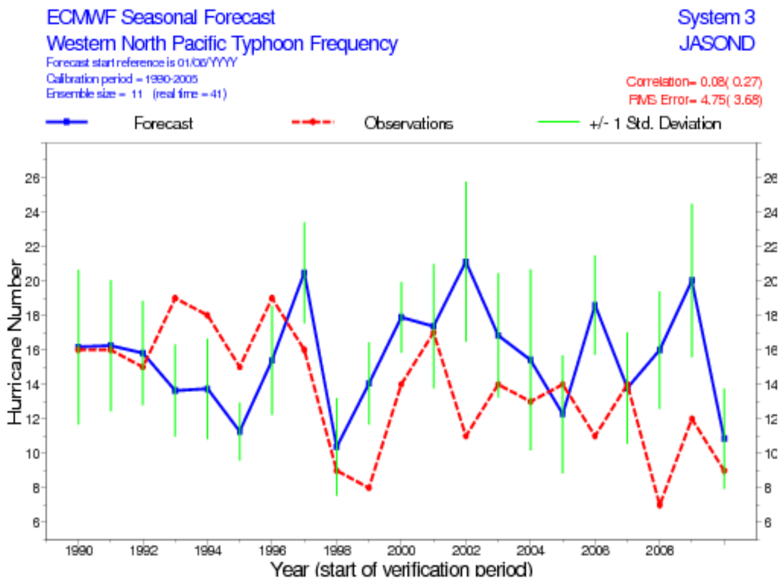
Tropical storm forecasts





Prediction of tropical cyclone frequency: NW Pacific



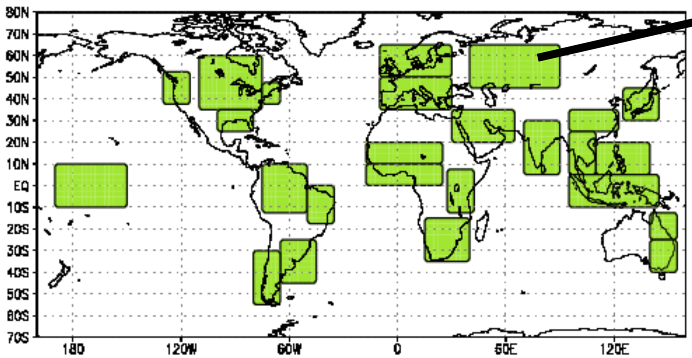


System 4
vs. ERA-Int.

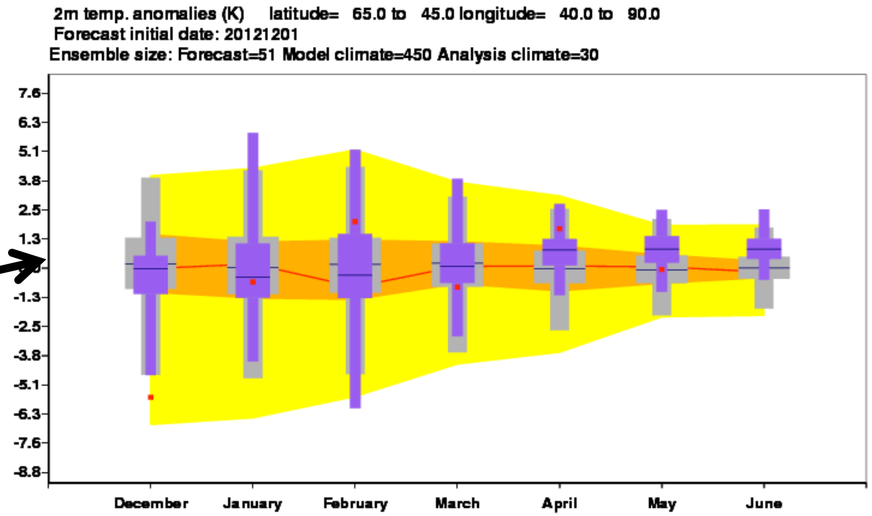
July-Dec.
1990-2010

System 3
vs. ERA-Int.

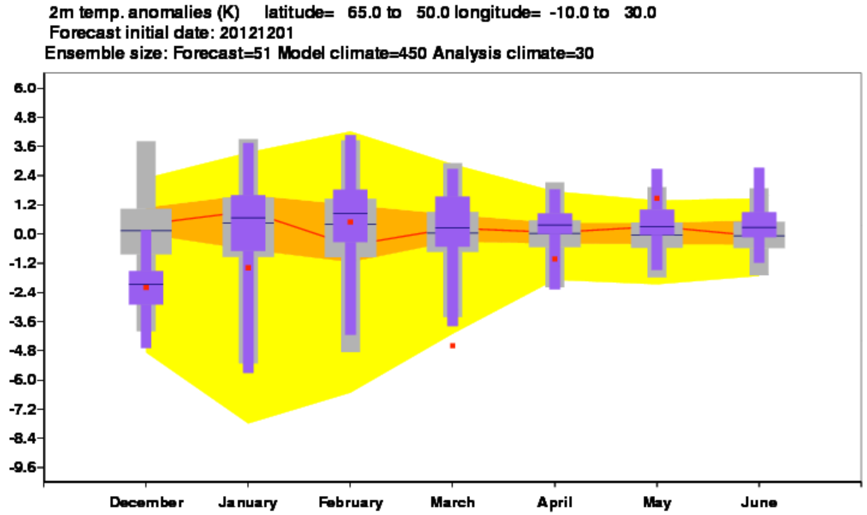
Climagrams:

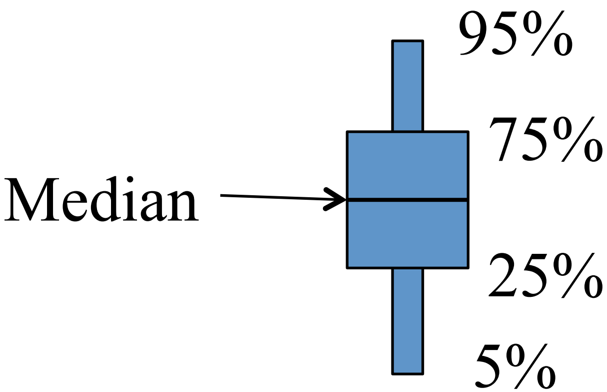


Central Asia

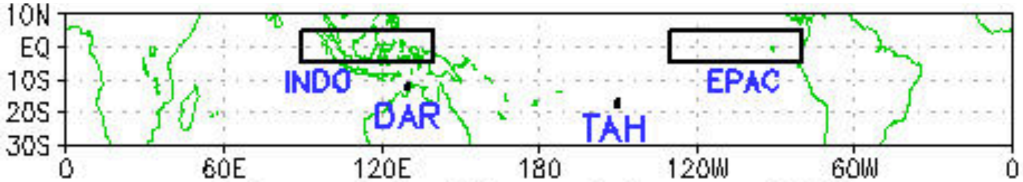


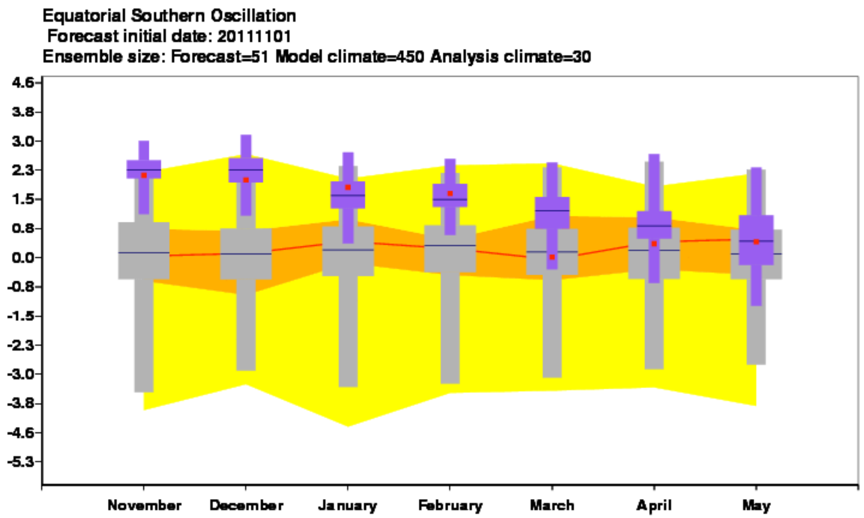
Northern Europe



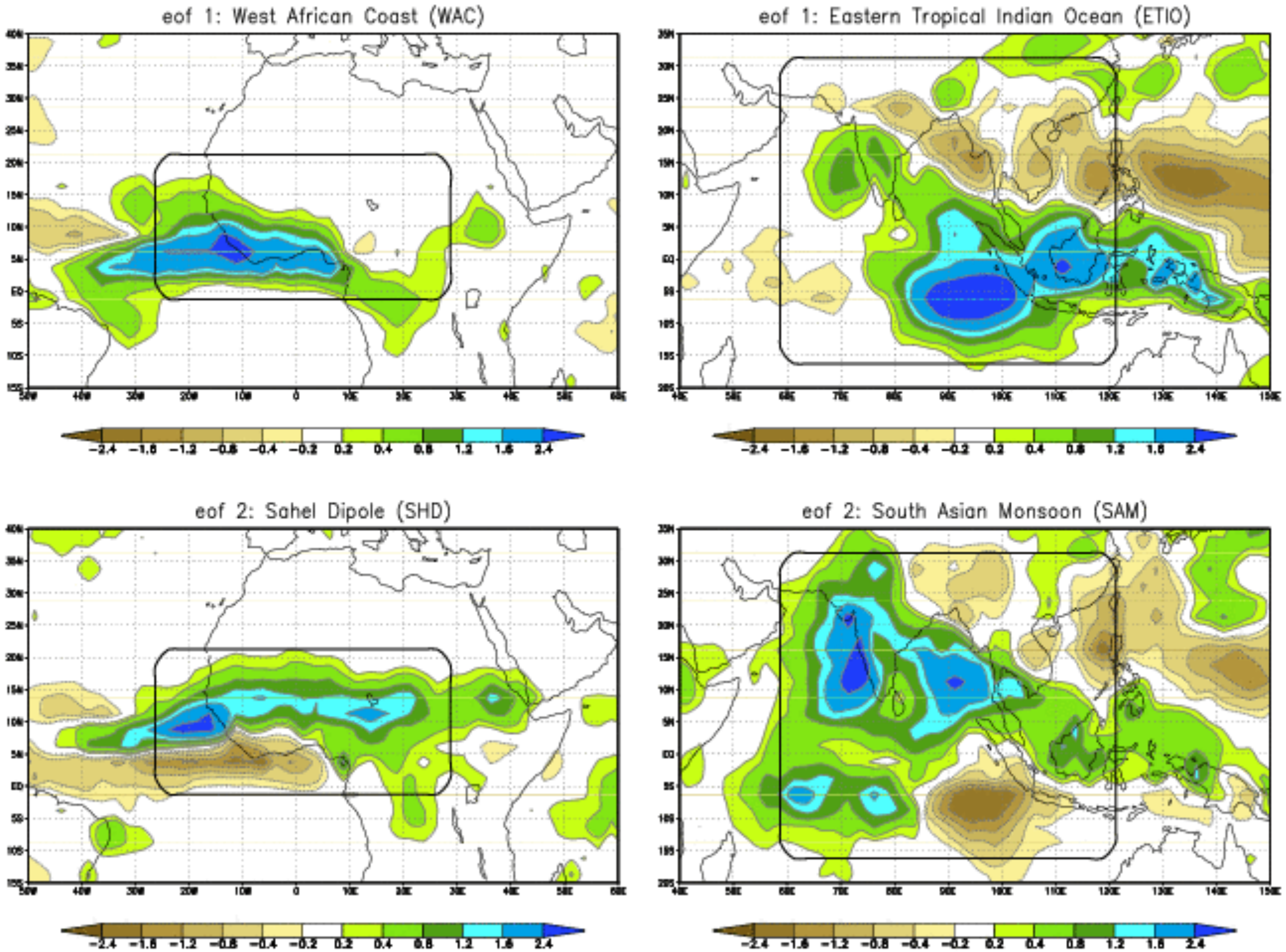


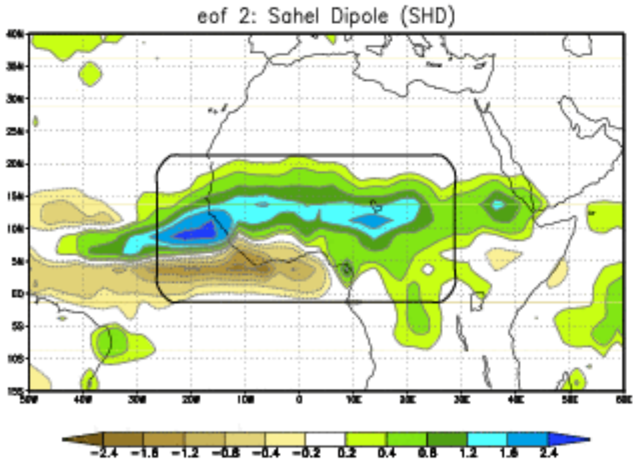
Teleconnections Indices, SOI:



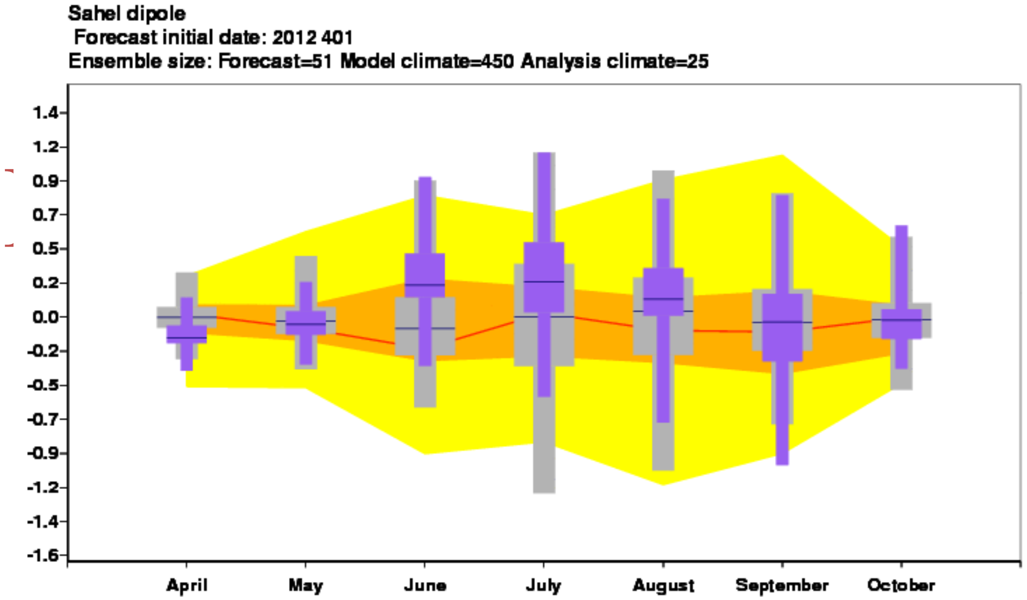


Monsoon Indices:

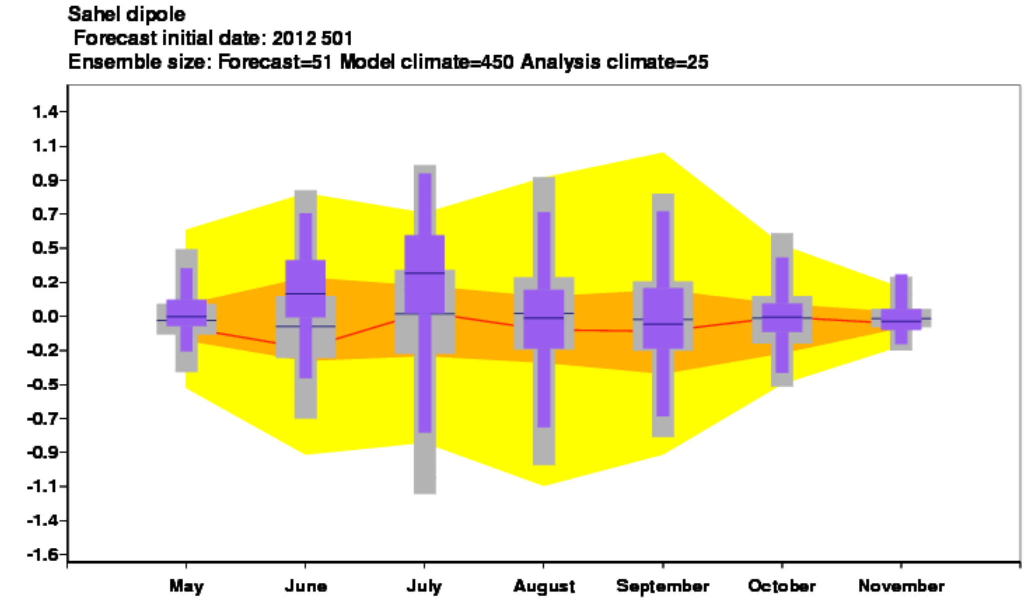




eof 2: Sahel Dipole (SHD)
-2.4 -1.6 -1.2 -0.8 -0.4 -0.2 0.2 0.4 0.8 1.2 1.6 2.4



Sahel dipole
Forecast initial date: 2012 401
Ensemble size: Forecast=51 Model climate=450 Analysis climate=25
1.4
1.2
0.9
0.7
0.5
0.2
0.0
-0.2
-0.5
-0.7
-0.9
-1.2
-1.4
-1.6
April
May
June
July
August
September
October



Sahel dipole
Forecast initial date: 2012 501
Ensemble size: Forecast=51 Model climate=450 Analysis climate=25
1.4
1.1
0.9
0.7
0.5
0.2
0.0
-0.2
-0.5
-0.7
-0.9
-1.1
-1.4
-1.6
May
June
July
August
September
October
November

Seasonal forecast skill

- **The verification of ensemble forecasts requires a sufficient number of verification samples and involves the application of probabilistic skill metrics. The measure of skill conforms to a common standard defined by the WMO. (WMO SVS for LRF)**

- **Seasonal forecasts show high prediction skill in the tropics, particularly the ENSO region.**

- **Predictability is low in the extra-tropics. Over Europe, seasonal forecasts are at best only slightly better than climatology.**

- **The skill depends strongly on the season, location and parameter.**

Model error and forecast interpretation

- Model error is large
  - It dominates El Nino forecast errors
  - Mean state and variability errors are very significant
  - Errors cannot be easily fixed

- Products typically account for sampling error only
  - **Do not take model probabilities as true probabilities**

- Estimating forecast skill can be difficult
  - In mid-latitudes, data is insufficient to produce robust skill estimates
  - Makes interpretation of forecasts tough

- In the end we need "good" models
  - (Multi-model ensembles are small, and only partially span the space of model errors)

Multi-model approach:

- Multi-model approach is an alternative way to sample model errors.

- The combination of several independent models widens the ensemble spread by sampling model errors.

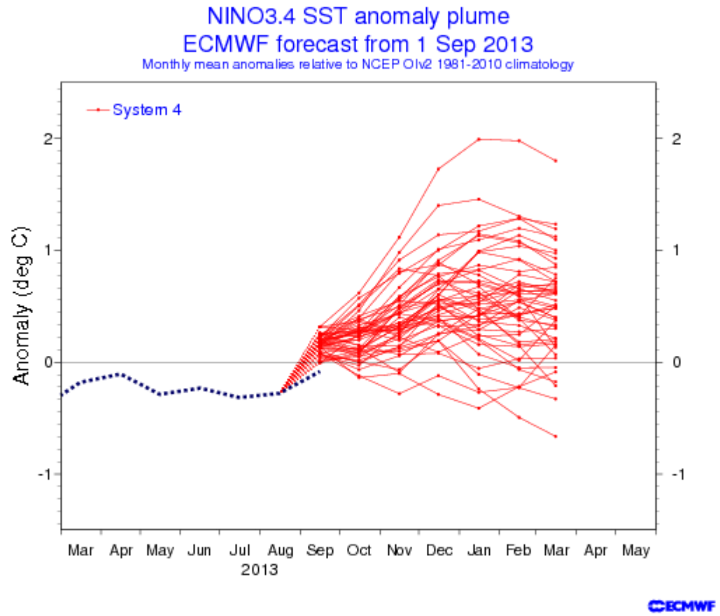
- The multi-model forecast can better represent the full range of uncertainties. Its spread can represents better the unpredictable noise so that the multi-model forecast is more reliable .

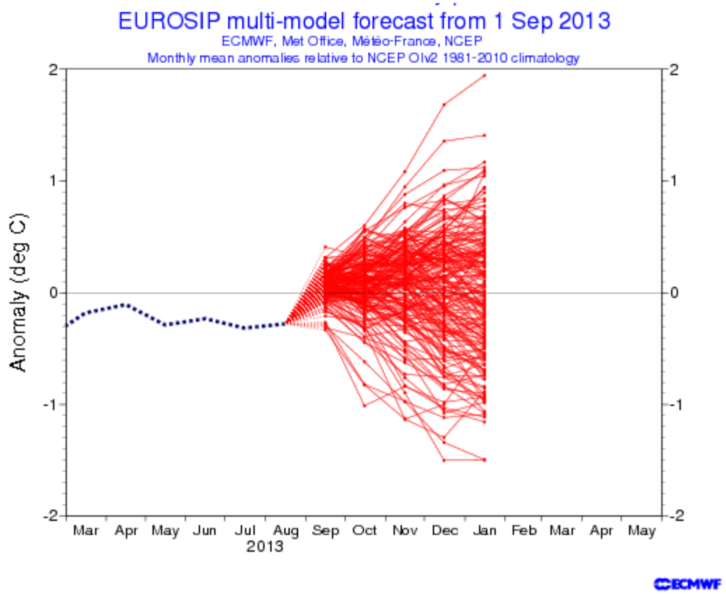
EUROSIP multi-model system:

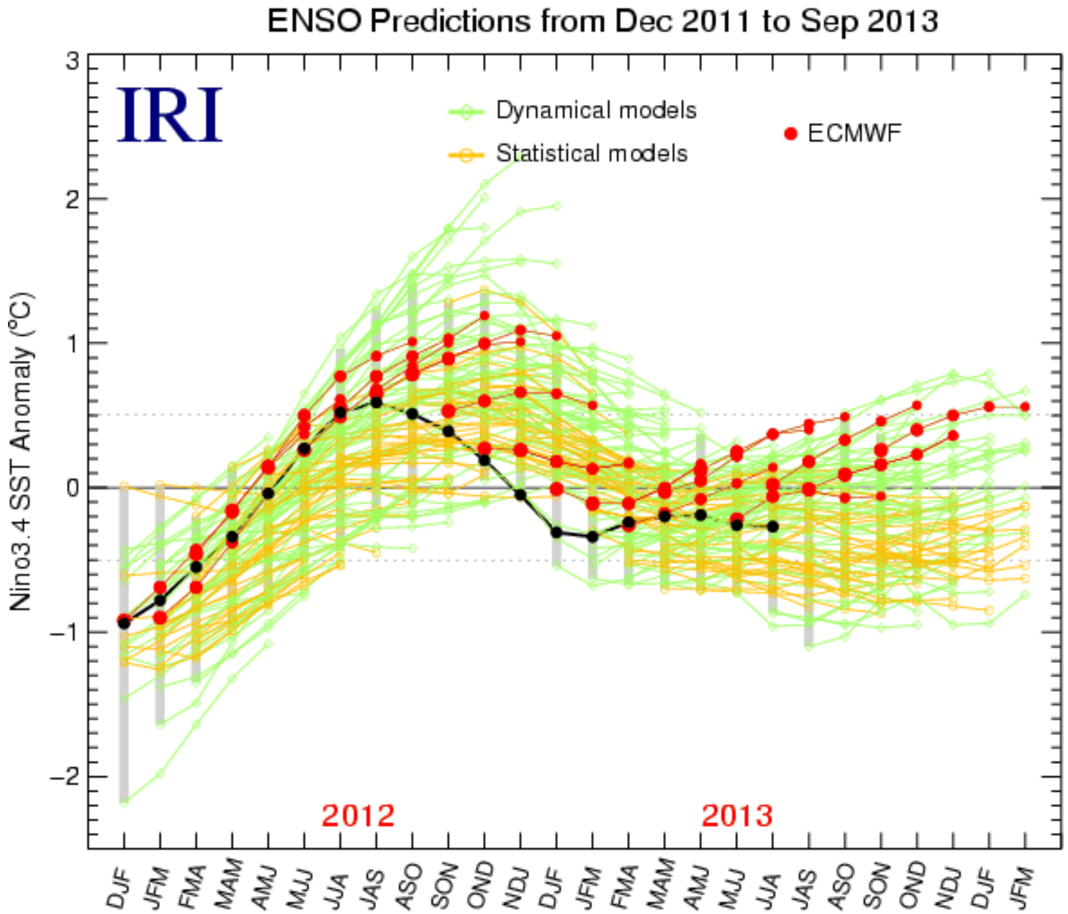
4 Coupled Systems: ECMWF, Météo France, Met Office, Ncep

- **Ensemble generation for the 4 systems is different**
- **Development of multi-model products is ongoing**
- **EUROSIP products <u>are</u> available to WMO users**

Nino 3.4 outlook



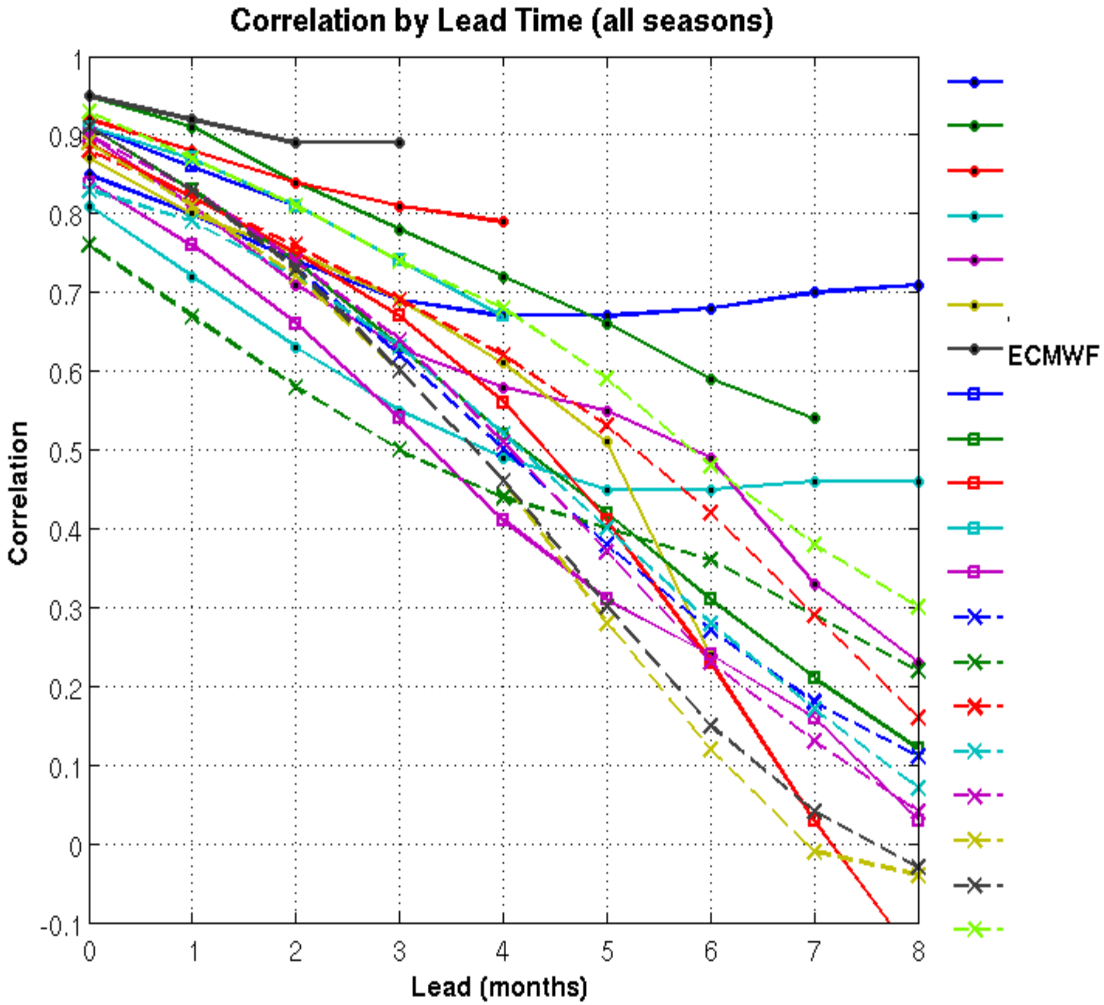




ENS

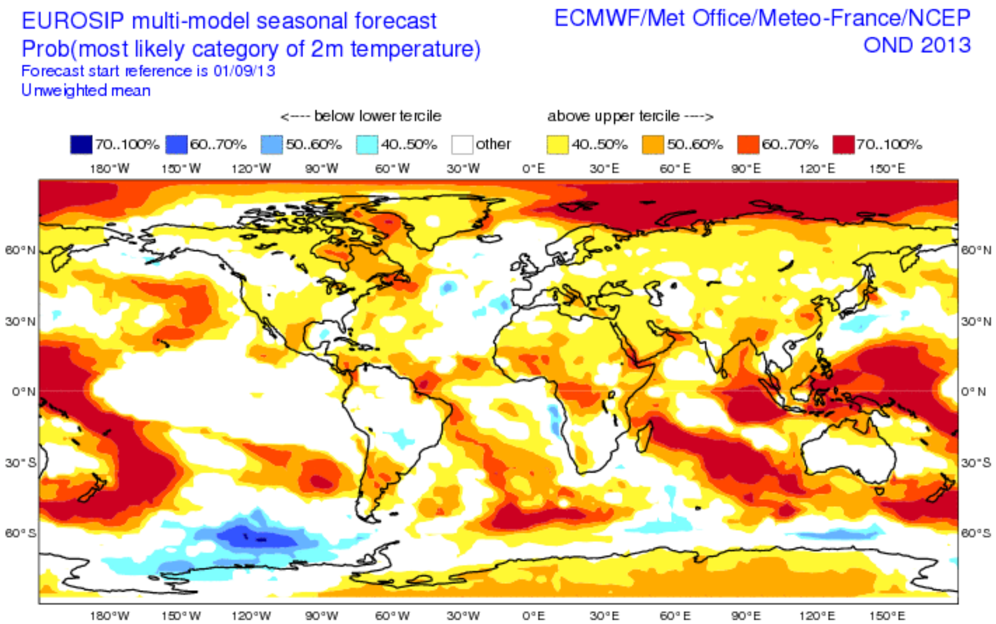
NINO3.4
Anomaly
Correlation

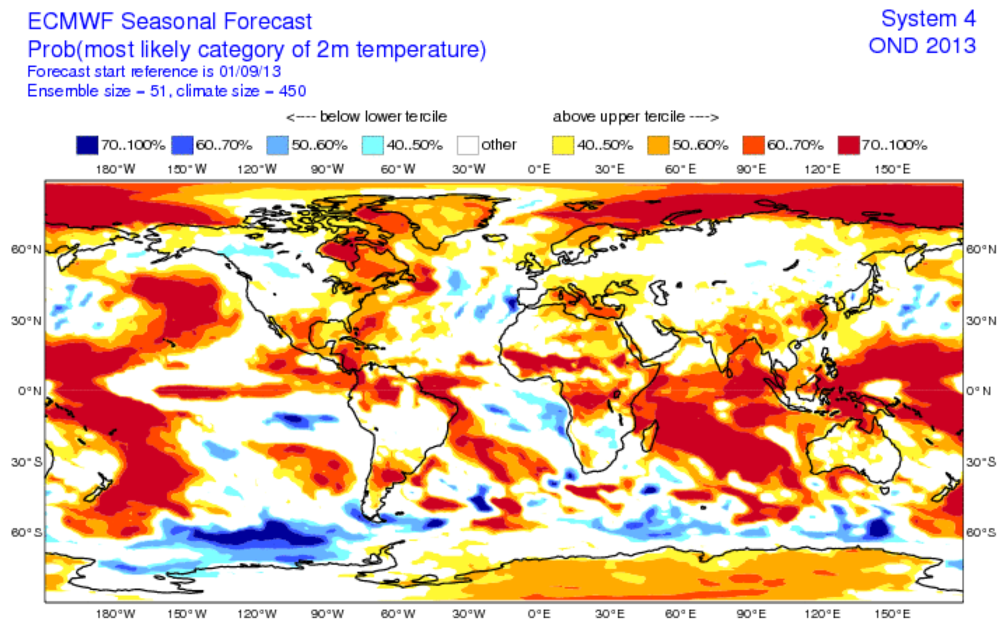
3-month
running means



From: Barnston et al. 2011: Skill of Real-time Seasonal ENSO Model Predictions during 2002-2011—Is Our Capability Increasing? BAMS 2012

EUROSIP multi-model prediction:





Summary

- The current operational seasonal forecast system provides a set graphic products to the WMO users.

- The ECMWF seasonal forecast is a good system for El Nino predictions.

- Seasonal forecast predictions, particularly over mid-latitudes, should be used in combination with some estimate of the forecast skill. **Various skill estimates conformed to the WMO-SVS for LRF are published on the ECMWF web site.**

- Multi-model approach: a way to deal with model error (model calibration) and to enhance forecast reliability. **EUROsip Multi Model products are available to the WMO users**

- ECMWF Technical memorandum N.656 available at http://www.ecmwf.int/publications/library/do/references/list/14 describes the upgraded seasonal forecast system and its major features in comparison with the previous operational system.

- Plenty of scope for improving model forecasts
  - Nino SST forecasts still significantly worse than predictability limits
  - Model errors still obvious in many cases, some processes poorly treated
  - Ocean initial conditions probably OK in Pacific for reasonable number of years, recently improved elsewhere by ARGO
- Model output -> use of forecast
  - Calibration and presentation of forecast information
  - Potential for multi-model ensembles
  - Integration with decision making
- Timescale for improvements
  - Optimist: in 10 years, we'll have much better models, pretty reliable forecasts, confidence in our ability to handle climate variations
  - Pessimist: in 10 years, modelling will still be a hard problem, and progress will largely be down to improved calibration

Summary (2)

- **Seasonal fc. System-4 (S4):** IFS-NEMO coupled model, 3-D var. ocean data assimilation (NEMOVAR), higher atmos. spatial resolution than S3, larger ensemble size, extended re-forecast set.
- **Model biases**: much reduced extra-tropical biases, too strong trade winds and cold SST bias in the equatorial Pacific. ENSO SST variability is over-estimated.
- **SST forecast skill**: similar to S3 in the NINO regions (better in NINO3, slightly worse in NINO4), increased in the tropical and sub-trop. Atlantic.
- **Skill for atmospheric variables:** spatial averages of ensemble-mean scores are consistently higher than in S3 (NH summer better than winter).
- **Tropical atmospheric variability:** more realistic patterns of rainfall variability, better simulation of the interannual and decadal variation in tropical cyclone frequency.
- **Reliability:** the enhanced internal variability and better match between spread and error lead to more reliable seasonal forecasts w.r.t. S3 in both tropical and extra-tropical regions.

Operational seasonal forecasts

- Real time forecasts since 1997
  - "System 1" initially made public as "experimental" in Dec 1997
  - System 2 started running in August 2001, released in early 2002
  - System 3 started running in Sept 2006, operational in March 2007
  - System 4 started running in July 2011, operational in November 2011

- Burst mode ensemble forecast
  - Initial conditions are valid for 0Z on the 1st of a month
  - Forecasts are usually complete by late on the 2nd.
  - Forecast and product release date is 12Z on the 8th.

- Range of operational products
  - Moderately extensive set of graphical products on web
  - Raw data in MARS
  - Formal dissemination of real time forecast data