Seasonal forecasting systems: present and future

Tim Stockdale

European Centre for Medium-range Weather Forecasts

t.stockdale@ecmwf.int

Outline

- Introduction: Why do we use numerical models for seasonal forecasting?

- The practice of numerical seasonal forecasting: SEAS5
  - Model
  - Initial conditions
  - Reforecasts
  - Calibration
  - Forecast assessment
  - Biases
  - Multi-model approach

- Recap

Approaches to seasonal forecasting

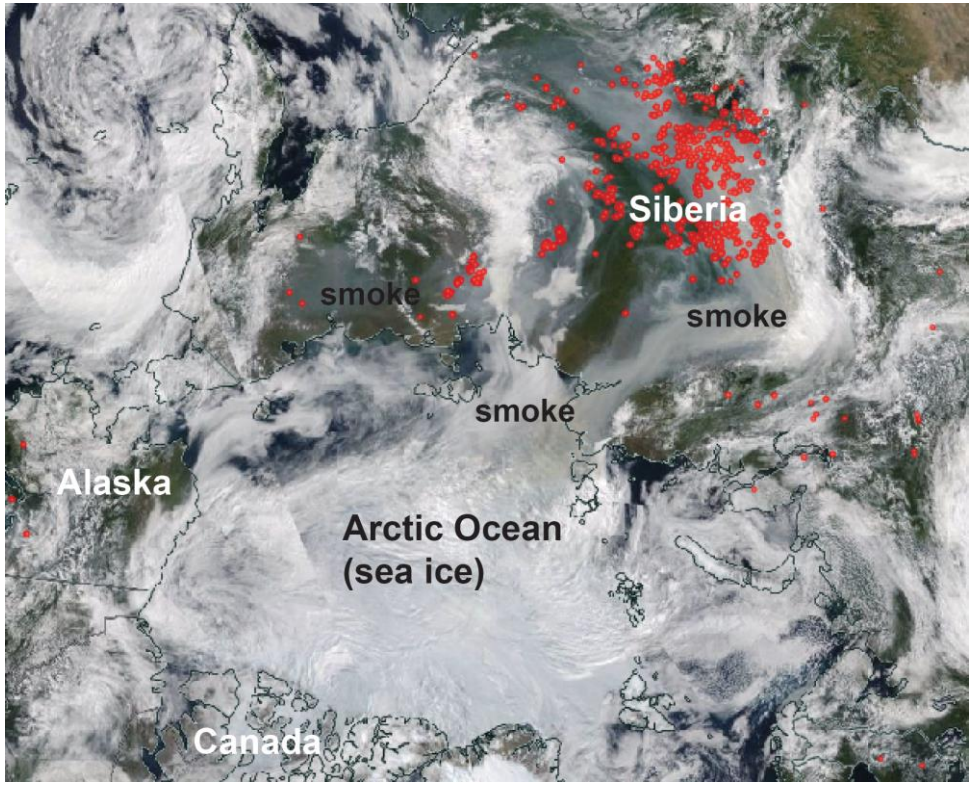
- Empirical or low-order model forecasting
  - Use past observational record and statistical methods
  - Works with reality instead of error-prone numerical models ☺
  - Aim to extract predictable signals, and represent uncertainty in outcome ☺
  - Limited number of past cases: works best when observed variability is dominated by a single source of predictability ☹
  - Limited quality of past data is a problem for some sources of predictability ☹
  - A non-stationary climate is problematic ☹
- Full-physics numerical forecasts
  - Include comprehensive range of sources of predictability ☺
  - Non-linear interactions of oceanic, land and atmospheric sources of predictability ☺
  - Benefit from latest observing systems, use all information about the present state ☺
  - Ensemble forecasts map uncertainty in initial state to uncertainty in outcome ☺
  - Model errors are an issue! ☹

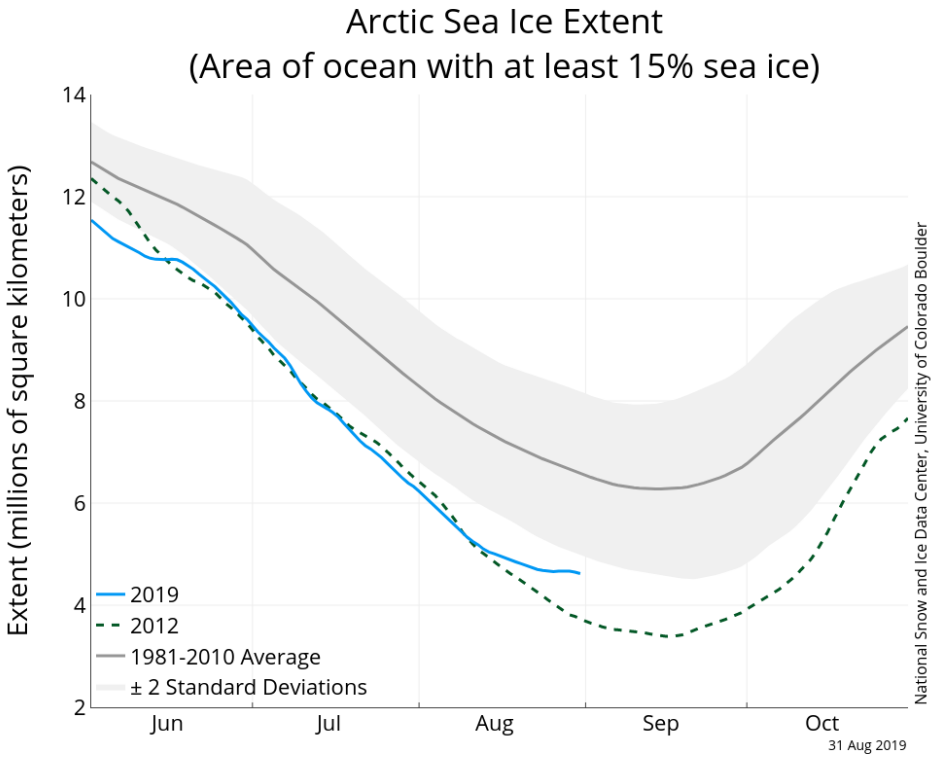
Sources of seasonal predictability

- KNOWN TO BE IMPORTANT:
  - El Nino variability - biggest single signal
  - Other tropical ocean SST - important, but multifarious
  - Climate change - impact is substantial, especially in temperature forecasts, and must be accounted for
  - Local land surface conditions - e.g. soil moisture in spring: dry soil warms up more quickly and is more prone to drought

- OTHER FACTORS:
  - Volcanic eruptions - definitely important for large events, gives global cooling plus sometimes a winter warming in parts of the northern hemisphere
  - Mid-latitude ocean temperatures - Convergence near SST fronts
  - Remote soil moisture/ snow cover - Unclear how large the effects might be
  - Sea ice anomalies - definitely local effects, also weaker remote impacts
  - Dynamic memory of atmosphere - most important for first month
  - Stratospheric influences - polar vortex, QBO, solar cycle, ozone, …
  - Aerosols - natural and anthropogenic
- Unknown or Unexpected - e.g. Arctic fires in summer 2019 ????

Arctic sea-ice, August 2019





Not yet known whether smoke may have had much influence – but is an example of an unexpected event.

Figures from NSIDC, https://nsidc.org/arcticseaicenews/ MODIS image credit: NASA Worldview

Numerical seasonal forecast – what do we need?

- A suitably accurate and comprehensive Earth System Model
- Observational data and an analysis system, to prepare initial conditions

- A re-analysis of the past, ideally several decades or more
- A set of re-forecasts covering these past years, to establish forecast biases (and skill)

- A real-time analysis system, which must produce data compatible with the re-analysis
- A real-time forecast system, with a large enough ensemble to sample evolution uncertainty
- A product generation system, including appropriate calibration

- Verification systems, dissemination systems, quality control systems
- A large computer

SEAS5: a state-of-the-art seasonal forecast system from ECMWF

- Introduced in November 2017

- Model: IFS, WAM, NEMO, LIM
- Analysis system: IFS, OCEAN5
- Re-analyses: ERAI, ORAS5, updated ERAI-Land

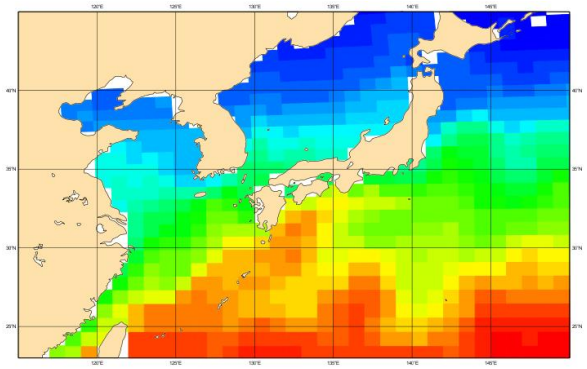
- Re-forecasts
- Product generation and dissemination

SEAS5 forecast model

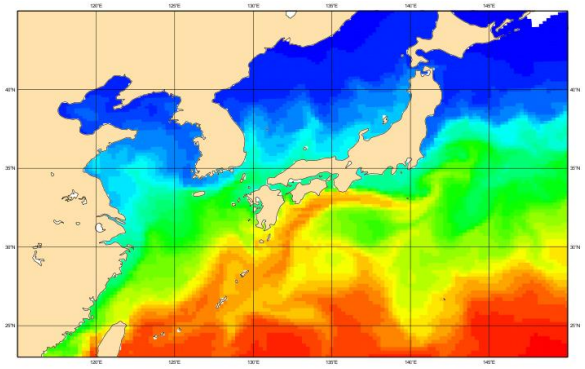
- IFS (atmosphere and land surface)
  - T_{CO}319L91 Cy43r1, 36km grid for physics, full stratosphere
  - All of the physical and dynamic processes of a world-class NWP model
  - Land surface model, multiple soil layers, different soil types, different vegetation types, snow, glaciers
  - Lake model with variable depths, variable mixed layer, surface and bottom temperatures, lake ice
  - Time varying tropospheric sulphate aerosol and stratospheric aerosol from volcanoes
- Wave model
  - Ocean surface waves modify the interaction between ocean and atmosphere. Runs at 0.5 deg resolution.
- NEMO (ocean)
  - Global ocean model, 0.25 deg resolution (eddy permitting), 75 vertical levels
- LIM (sea-ice)
  - Single category ice, solved on same grid as ocean model

SEAS5 – ocean component

- Ocean model resolution upgraded from previous 1x1 deg to 0.25x0.25 deg
- Ocean vertical resolution improved from previous 42 levels to 75 levels
- High ocean resolution is needed to represent ocean eddies, and to better resolve the boundary currents that are important in the ocean, such as the Kuroshio in the Pacific (shown here) and the Gulf Stream in the Atlantic.



NEMO ORCA1 Z42

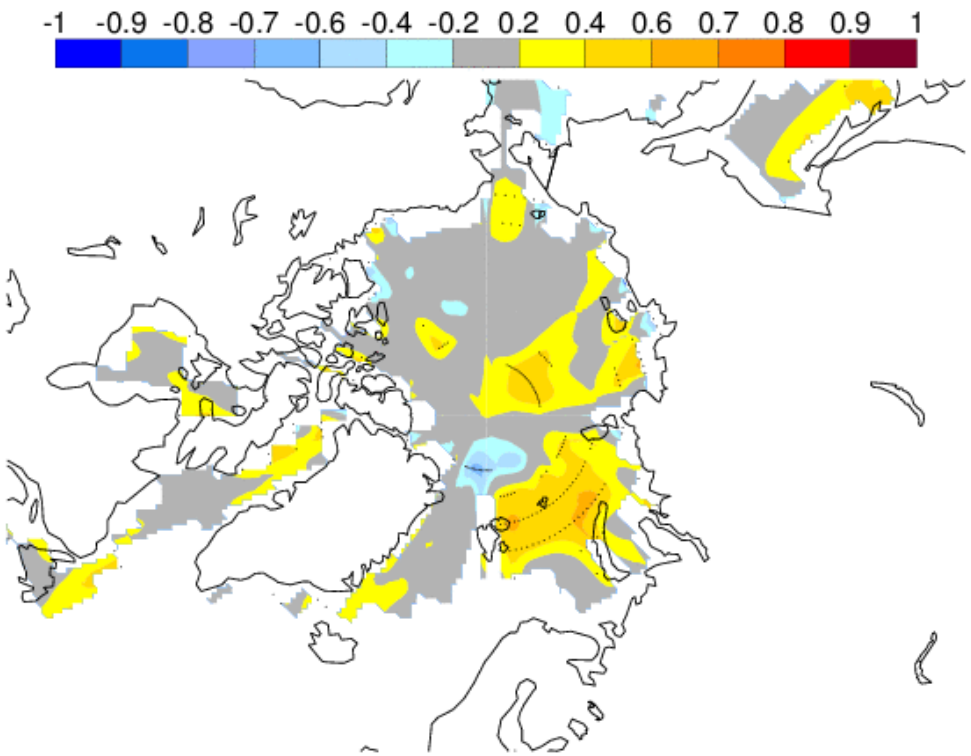


NEMO ORCA025 Z75

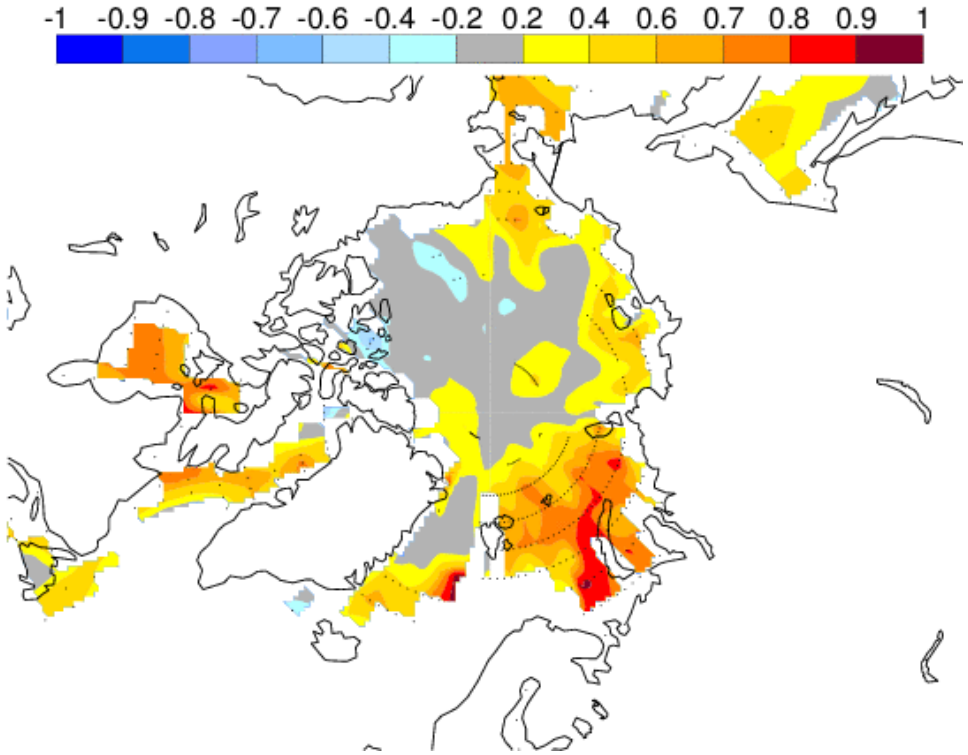
Note: ¼ degree ocean resolution is not enough to resolve eddies or boundary currents properly.

SEAS5 – LIM2 sea-ice model

SEAS4



SEAS5

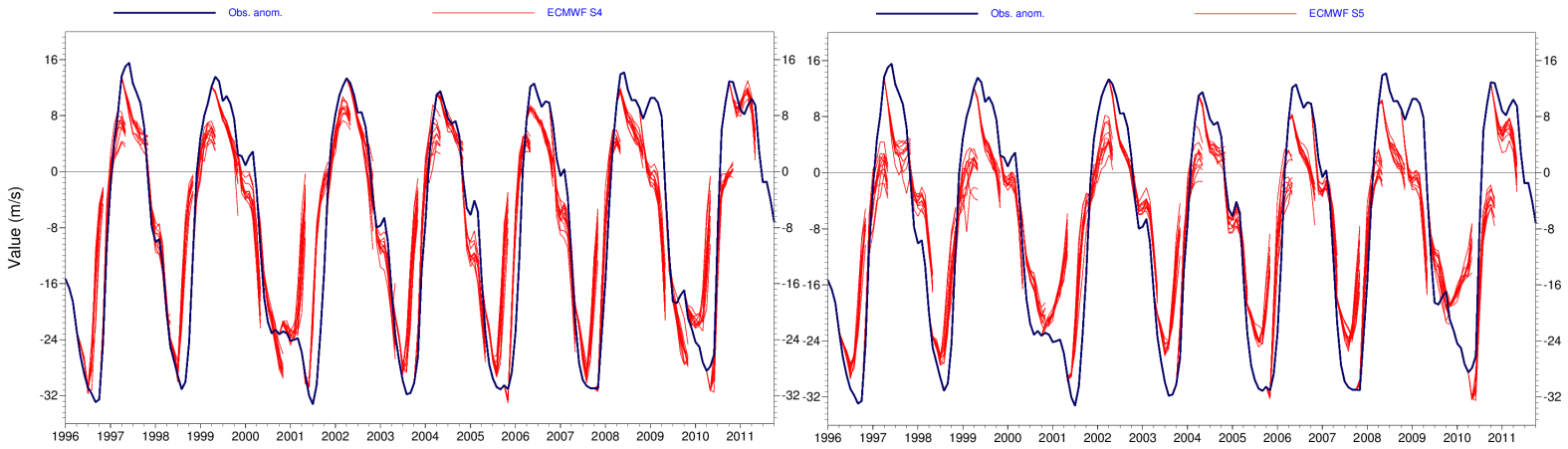


Sea ice cover predictability is improved when we include the interactive sea ice model, illustrated here with correlation scores for predictions of DJF sea ice cover.

Stratosphere - the QBO

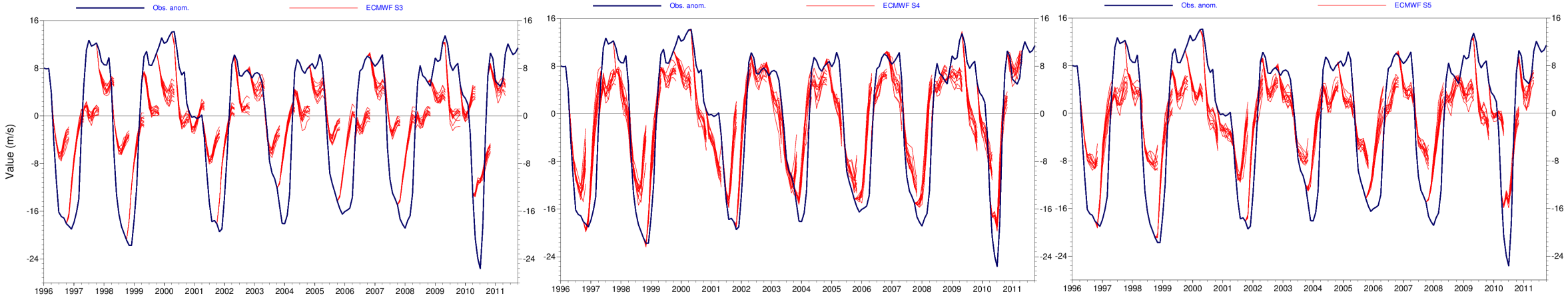
System 4 **SEAS5**

30hPa



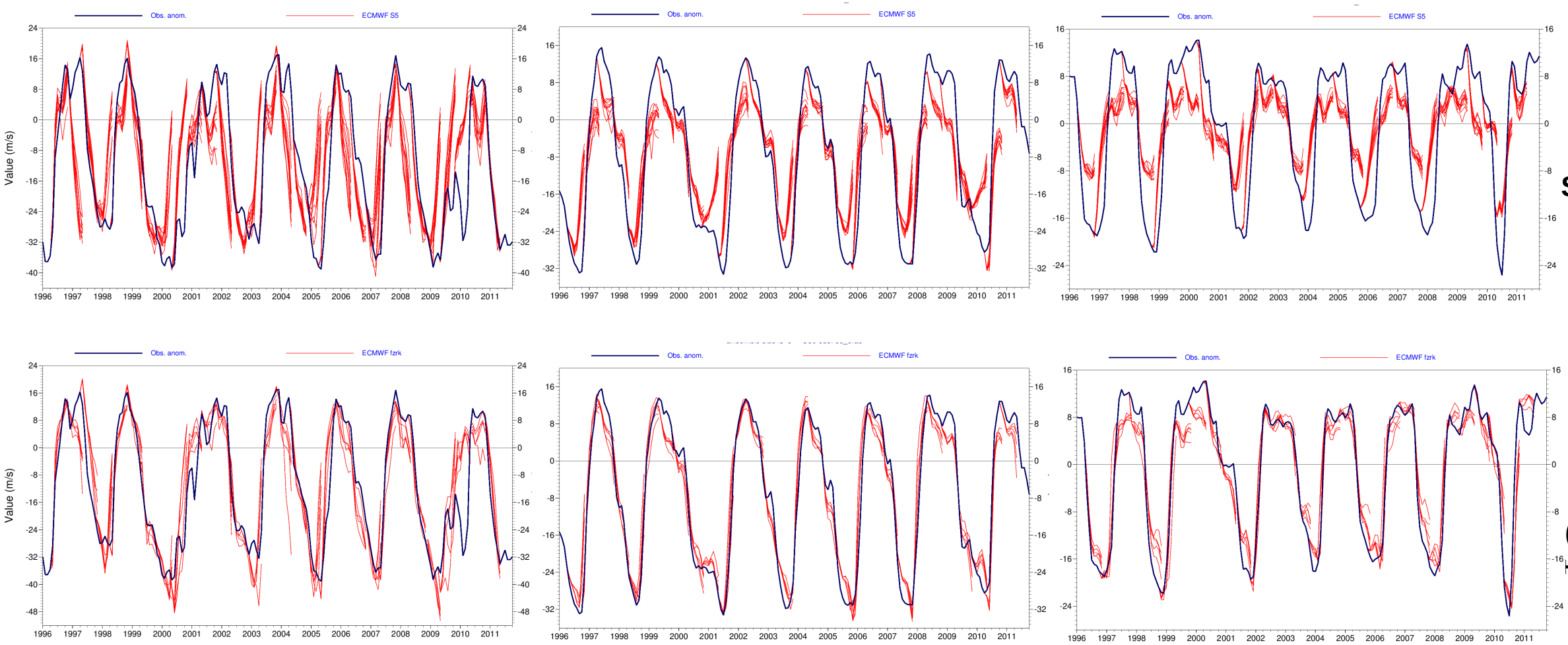
System 3

50hPa



QBO: what can be done …..

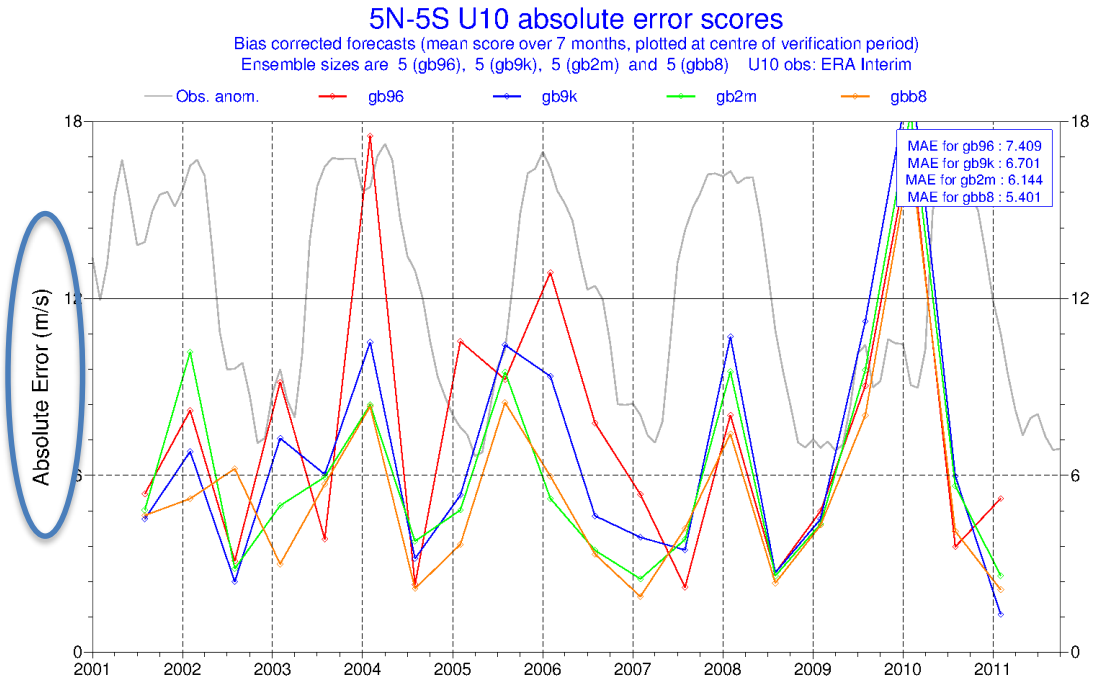
10 hPa **30 hPa** **50hPa**



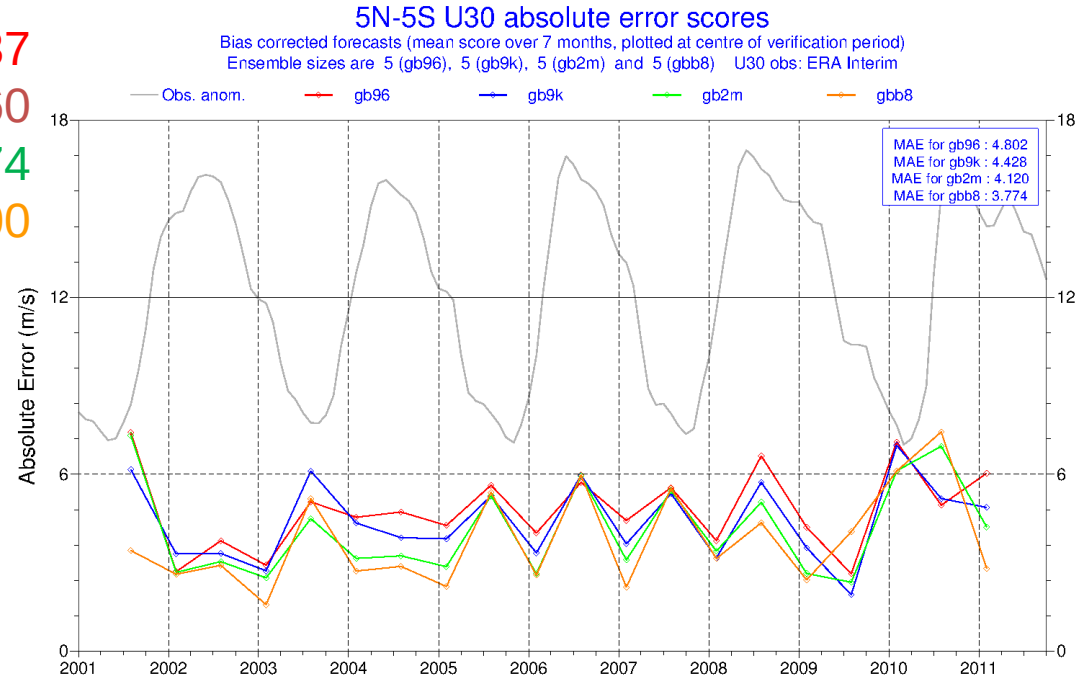
SEAS5

fzrk (L137 -phys)

To get best results for QBO, need resolutions better than L137

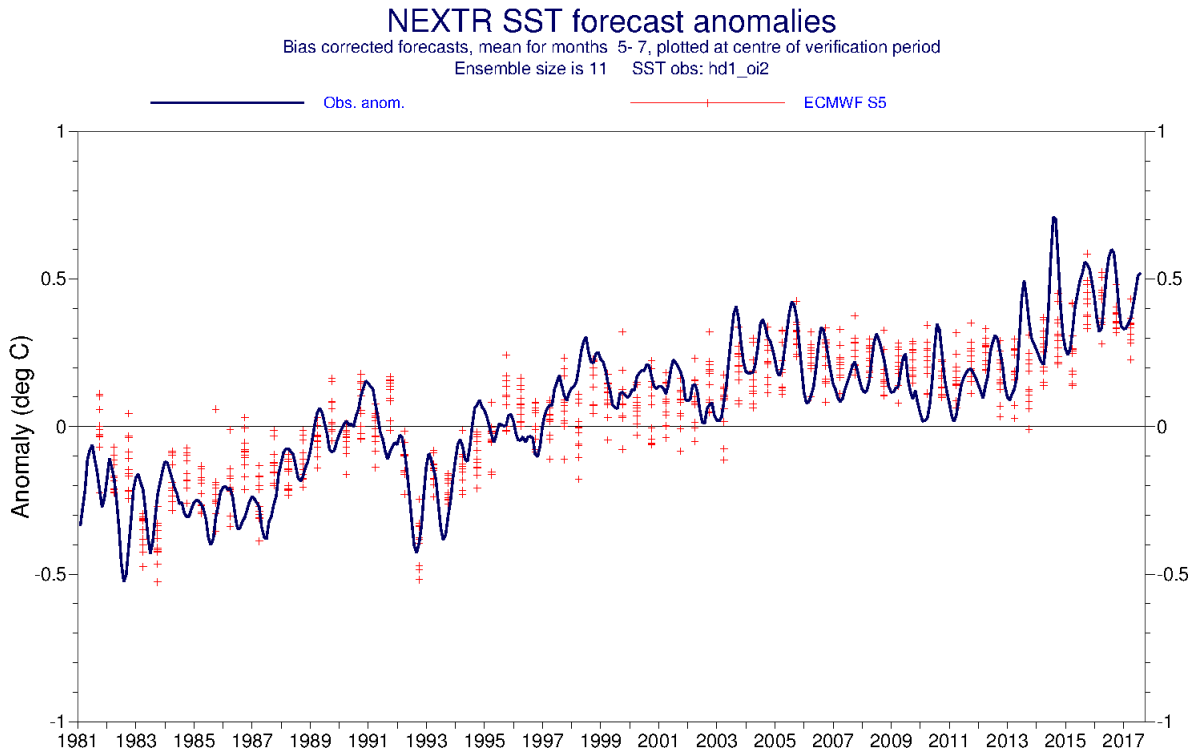


L137
L160
L174
L200



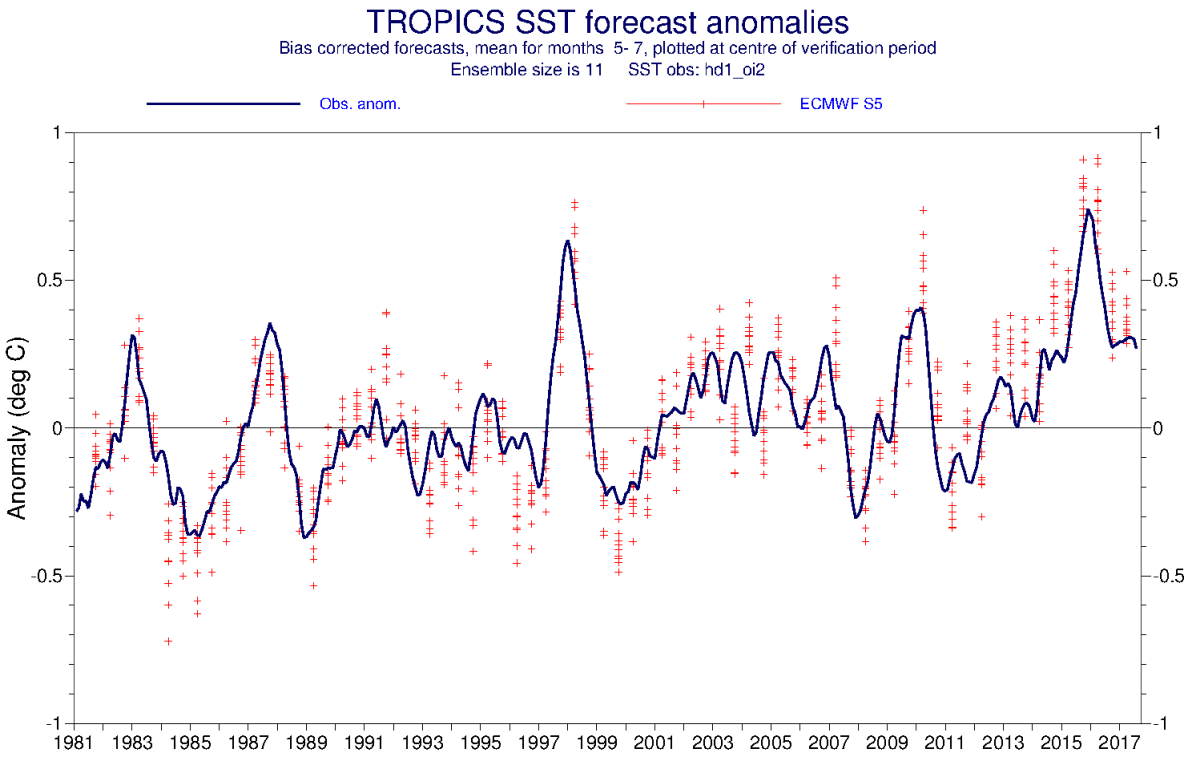
Without interactive ozone (which would enable better results)
Without re-tuning of NOGWD amplitude or other parameters

Trends



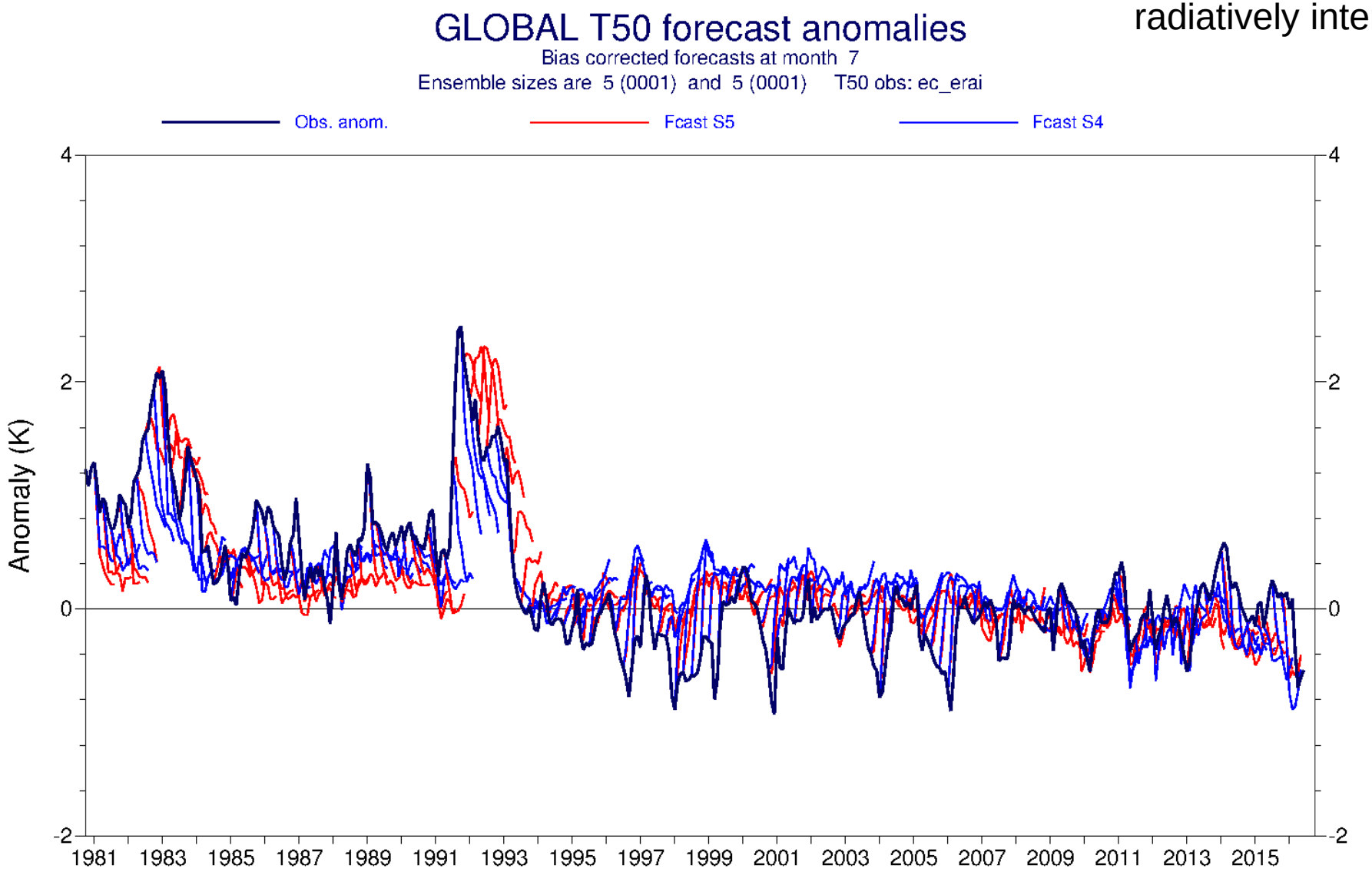
Capturing trends is important. Time-varying CO2 and other factors are important in this.

There is a strong link between seasonal prediction and decadal/ multi-decadal climate prediction.



Stratosphere – volcanic aerosols

SEAS5 does **not** have radiatively interactive ozone

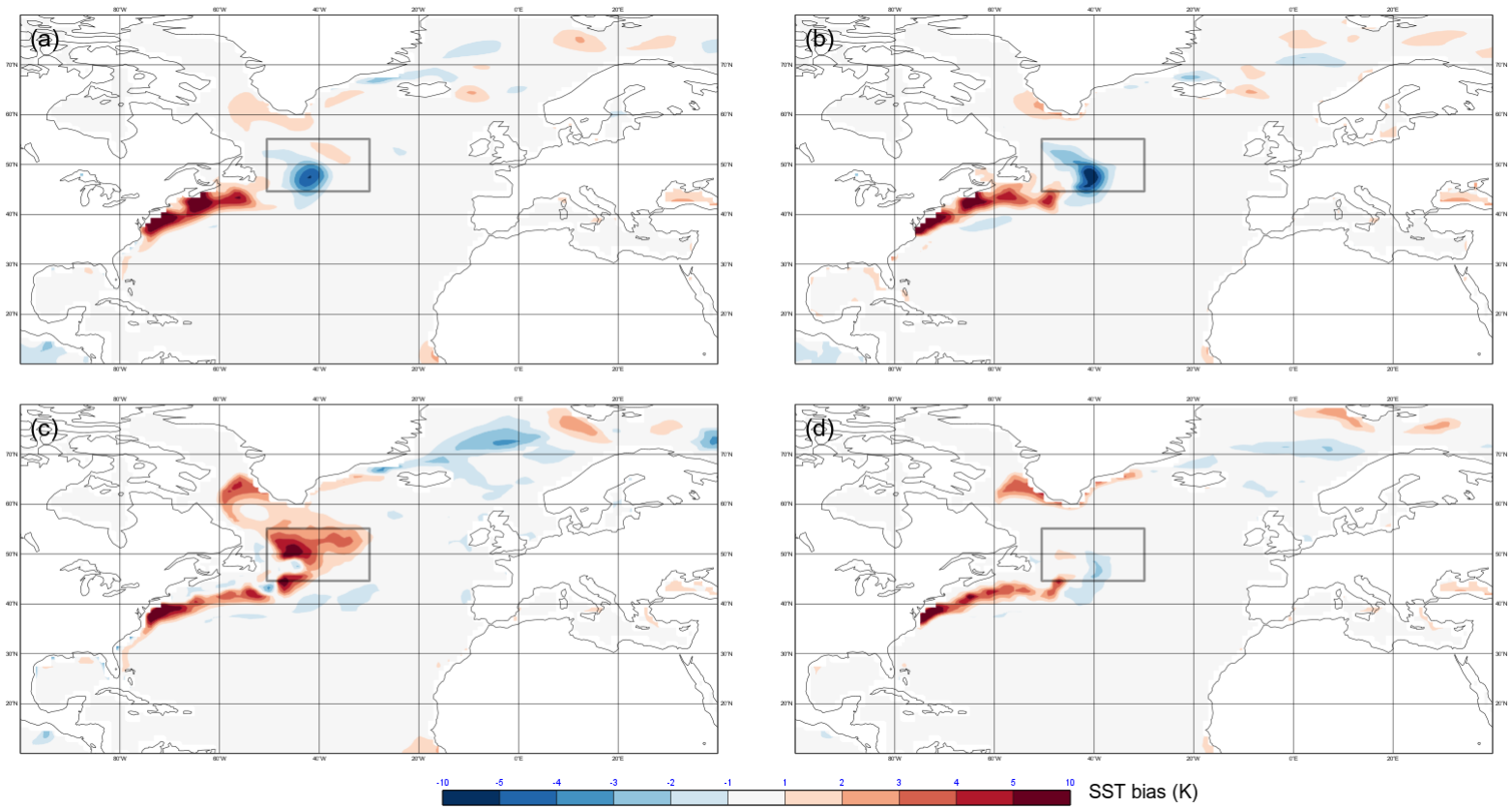


SEAS5 initial conditions

- Issue of consistency
  - Differences between real-time analyses and re-analyses must be minimised
  - Not easy, given evolution of observing system and practical constraints

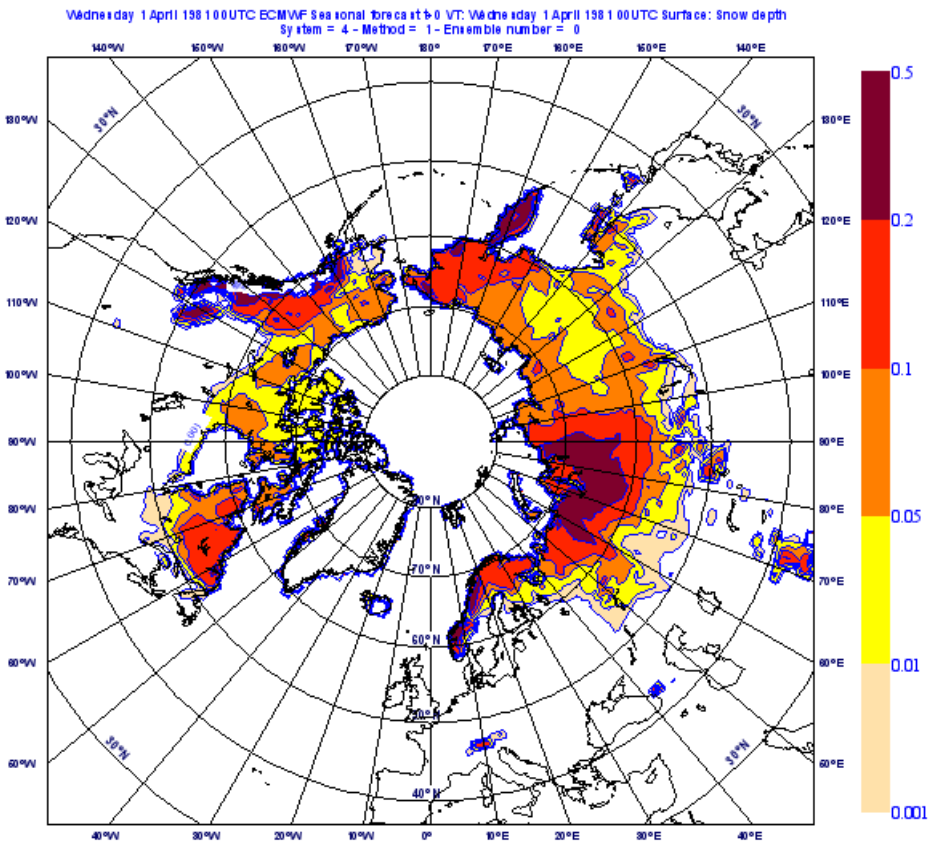
- Issue of uncertainty
  - Attempt to sample uncertainty in initial state
    - SST (50 pertns)
    - ocean sub-surface (5 member analysis)
    - atmosphere (singular vectors, EDA)
    - … but not yet land

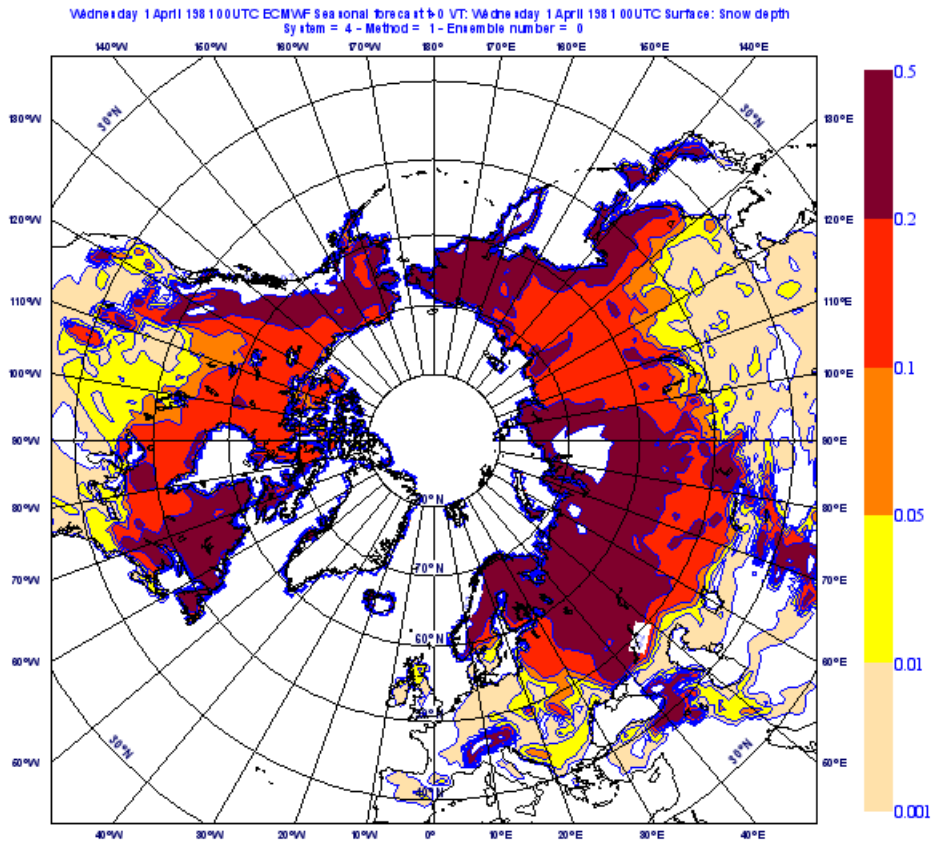
ORAS5: North Atlantic problem



DJF SST bias for November forecasts w.r.t. ERA-Interim for (a), (b) S4 and (c), (d) SEAS5. The bias during the early period 1981-1995 is shown in (a, c), the bias during the late period 2001-2015 is shown in (b, d). From ECMWF Tech Memo 835.

Initial conditions - land surface





Snow depth limits, 1st April

Used as a safeguard to ensure consistency between real-time and re-forecast initial states. Applied to soil moisture, snow, lakes, soil temperatures.

ECMWF SEAS5 forecasts and re-forecasts

- Real time forecasts:
  - **51 member ensemble forecast to 7 months**
  - SST and atmosphere initial perturbations (SV, EDA) added to each member

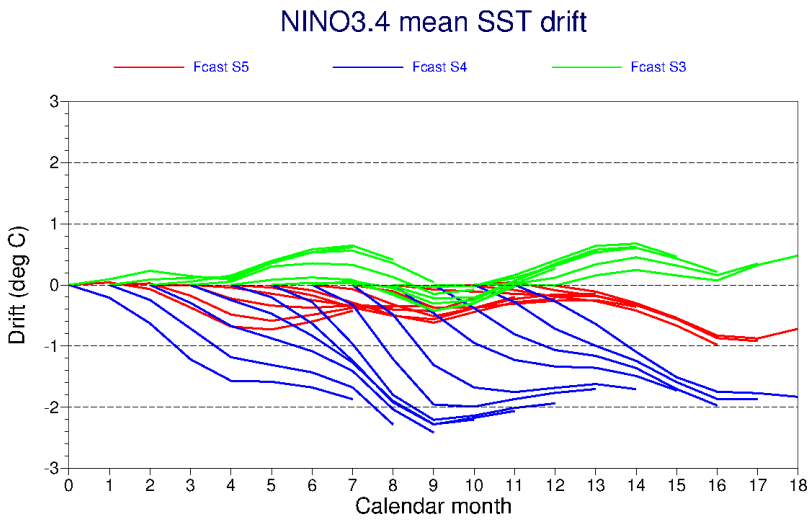
  - **15 member ensemble forecast to 13 months**
  - Designed to give an ‘outlook’ for ENSO
  - Only runs once per quarter (Feb, May, Aug and Nov starts)

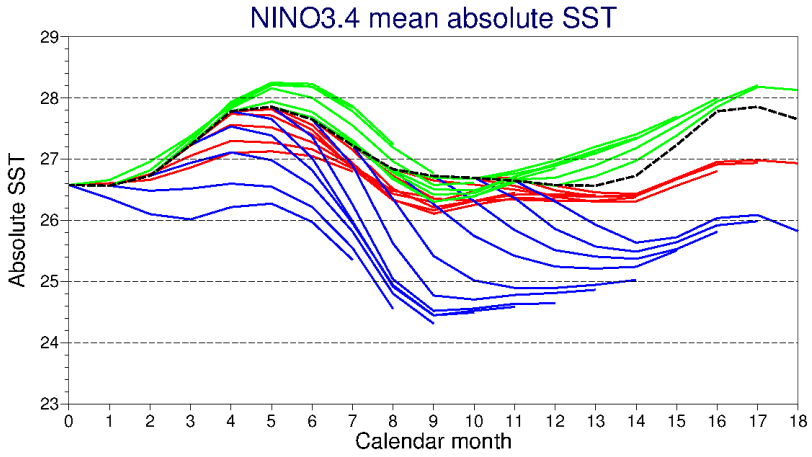
- Re-forecasts from 1981-2016 (**36 years**)
  - **25 member** ensemble every month
  - 15 members extended to 13 months once per quarter

How many re-forecasts?

- Re-forecasts dominate total cost of system
  - SEAS5: 10800 re-forecasts (must be in first year)
    612 real-time integrations (per year)
- Re-forecasts define model climate
  - Need both climate mean and the pdf, latter needs large sample
  - May prefer to use a “recent” period (SEAS5 has 36 years available, but uses only last 24 years for web products)
  - SEAS5 has 600 member climate (25 members * 24 years) for web products, so sampling is basically OK
- Re-forecasts provide information on skill
  - A forecast cannot be used unless we know (or assume) its level of skill
  - Observations have only 1 member, so large ensembles are less helpful than large numbers of cases.
  - Care needed e.g. to estimate skill of 51 member ensemble based on past performance of 25 member ensemble
  - For regions of high signal/noise, SEAS5 gives adequate skill estimates
  - For regions of low signal/noise (eg <= 0.5), need hundreds of years, 36 years available is not enough

Forecast calibration and products

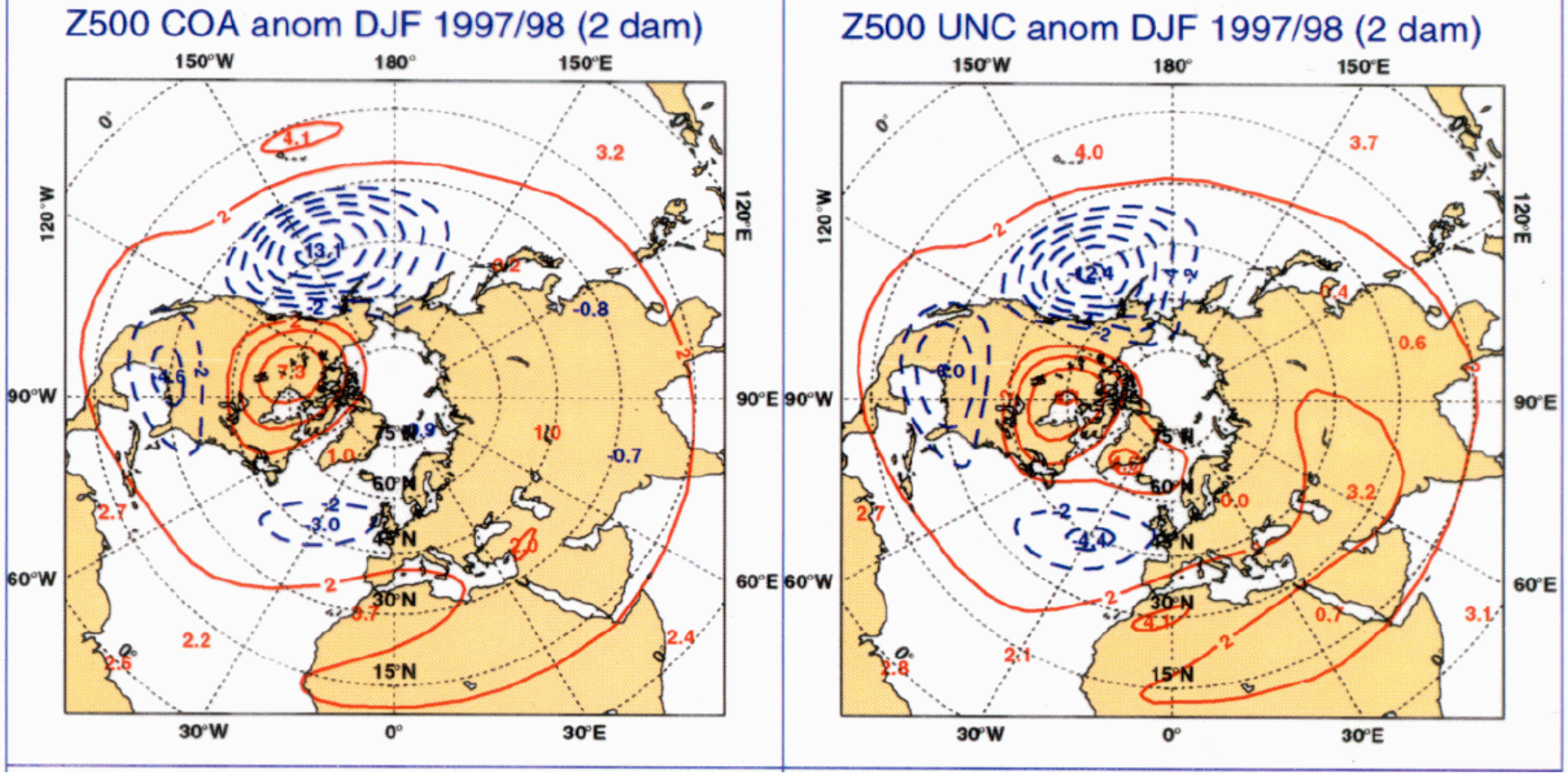




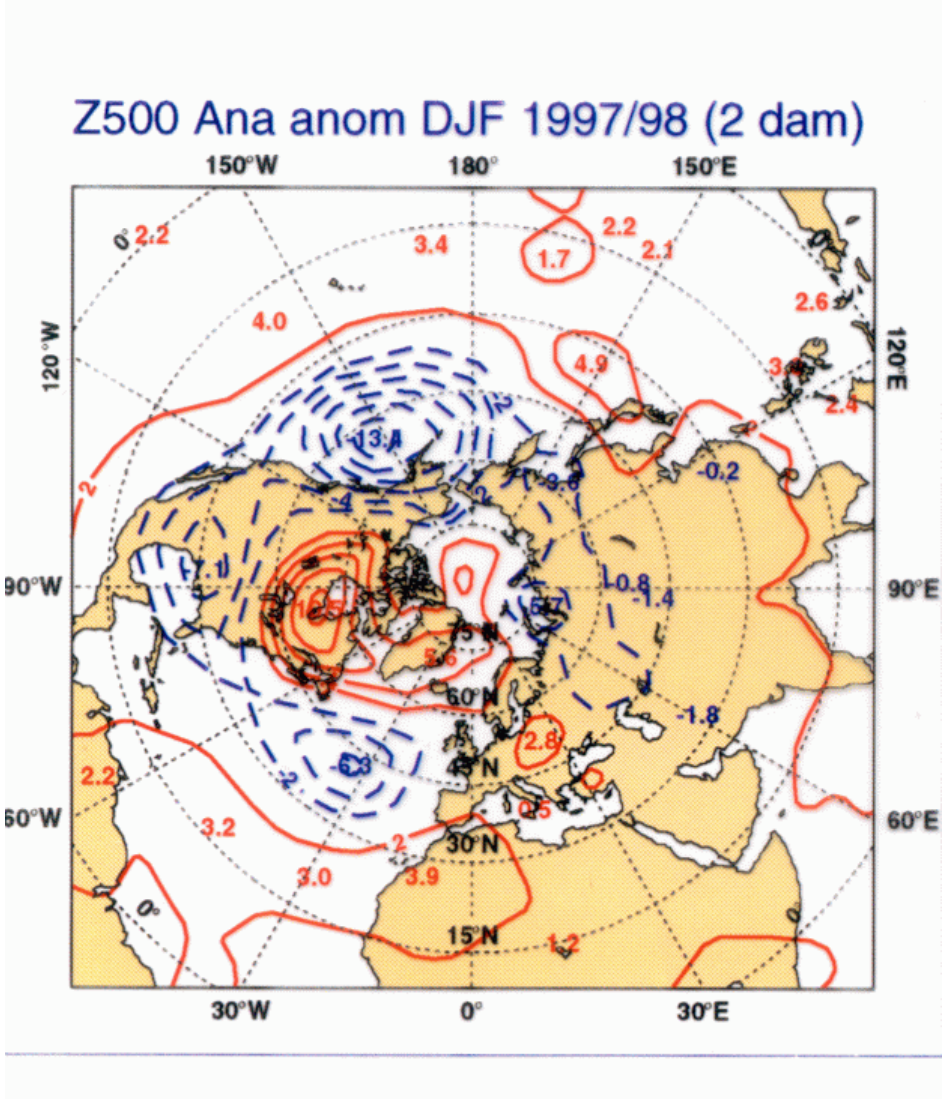
SST bias is a function of lead time and season.

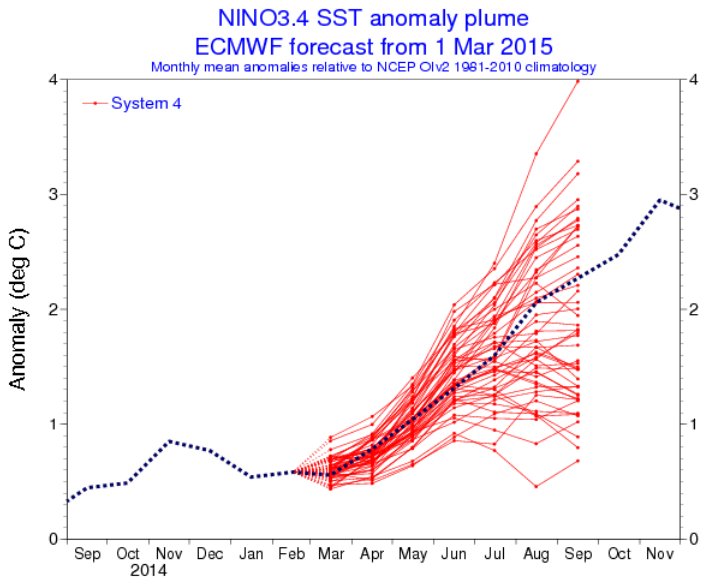
Some systems have less bias, but it is still large enough to require correcting for.

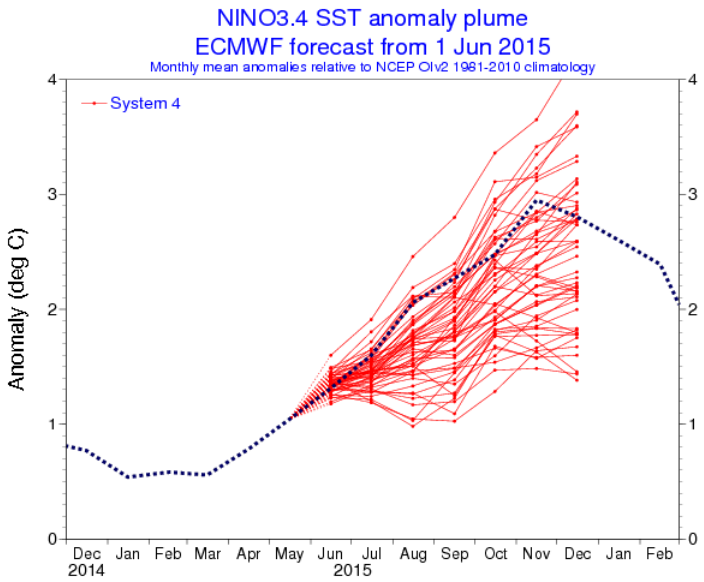
Despite SST bias and other errors, anomalies in the coupled system can be remarkably similar to those obtained using observed (unbiased) SSTs …..

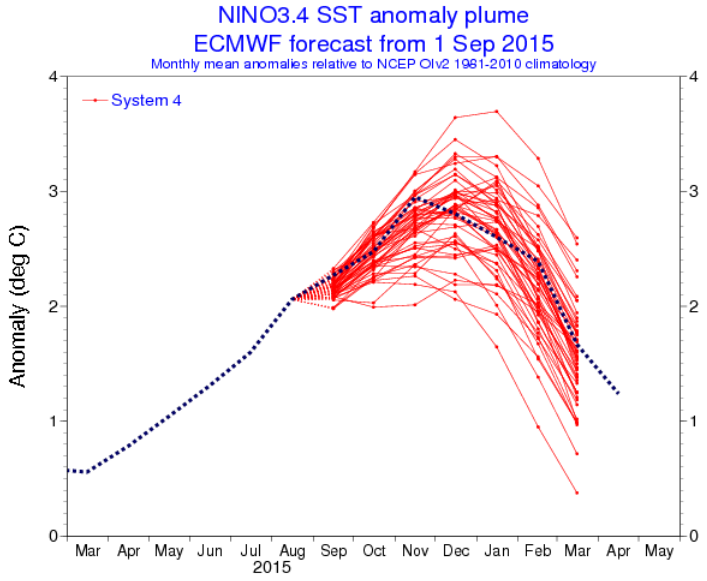


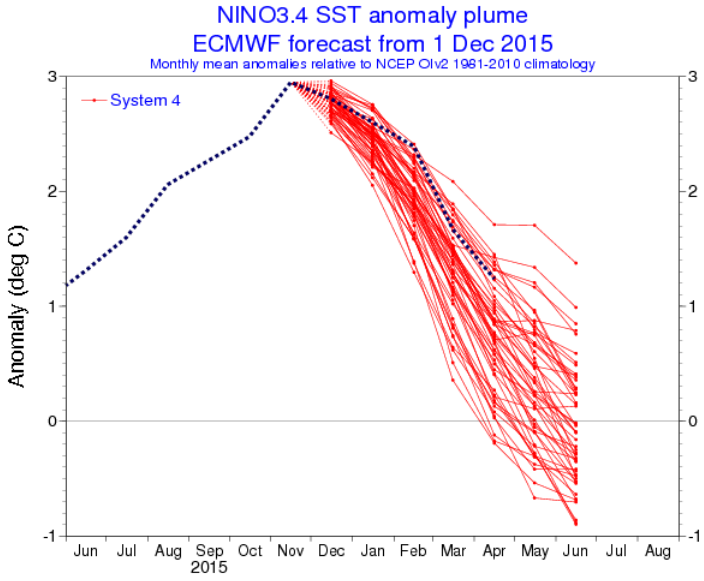
… and can also verify well against observations




NINO3.4 SST anomaly plume
ECMWF forecast from 1 Mar 2015
Monthly mean anomalies relative to NCEP OIv2 1981-2010 climatology
System 4
Anomaly (deg C)
Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov
2014


NINO3.4 SST anomaly plume
ECMWF forecast from 1 Jun 2015
Monthly mean anomalies relative to NCEP OIv2 1981-2010 climatology
System 4
Anomaly (deg C)
Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb
2014
2015


NINO3.4 SST anomaly plume
ECMWF forecast from 1 Sep 2015
Monthly mean anomalies relative to NCEP OIv2 1981-2010 climatology
System 4
Anomaly (deg C)
Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May
2015


NINO3.4 SST anomaly plume
ECMWF forecast from 1 Dec 2015
Monthly mean anomalies relative to NCEP OIv2 1981-2010 climatology
System 4
Anomaly (deg C)
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug
2015

ECMWF Seasonal Forecast

Mean 2m temperature anomaly

Forecast start is 01/09/17, climate period is 1993-2016

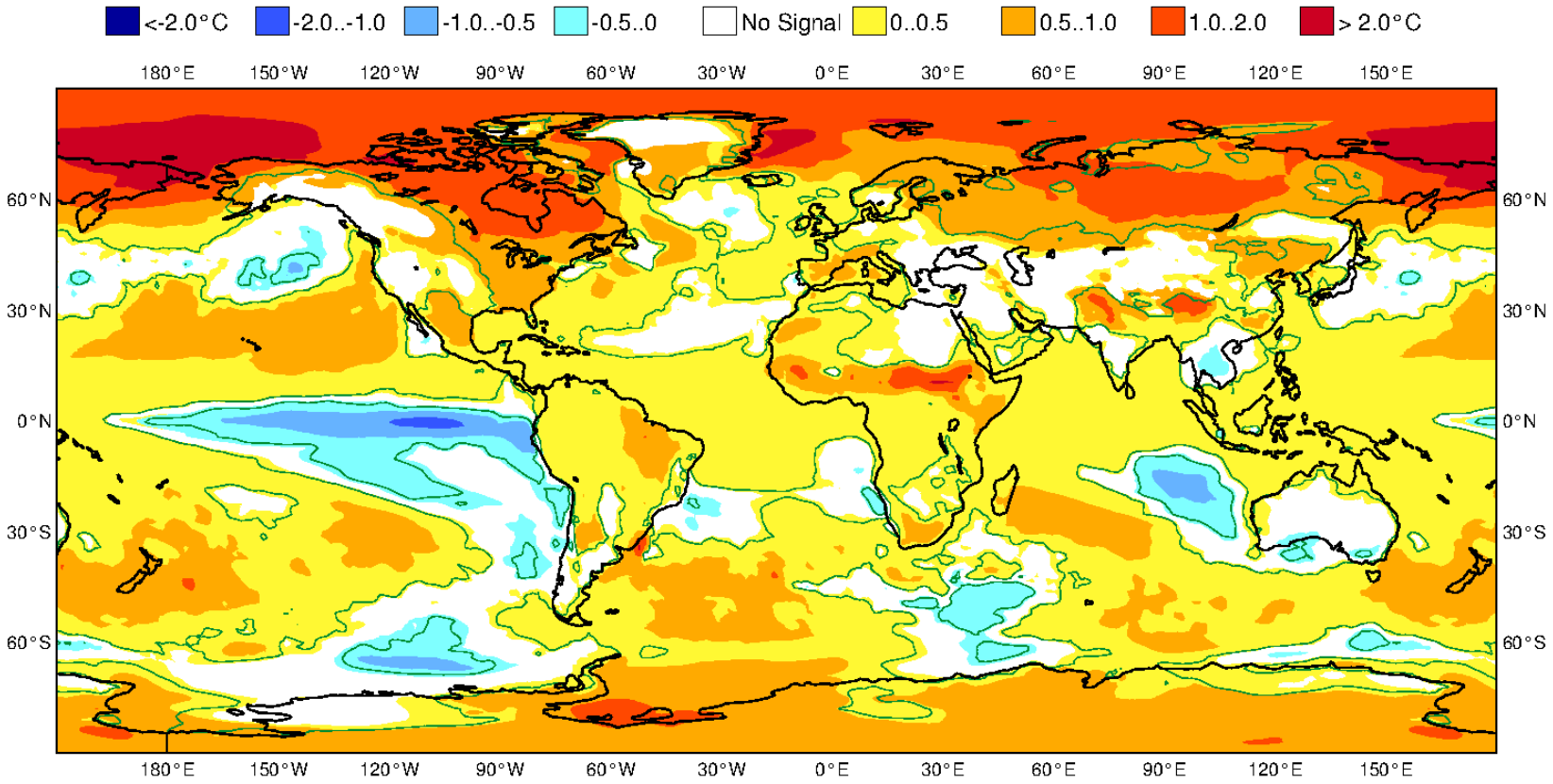
Ensemble size = 51, climate size = 600

System 5

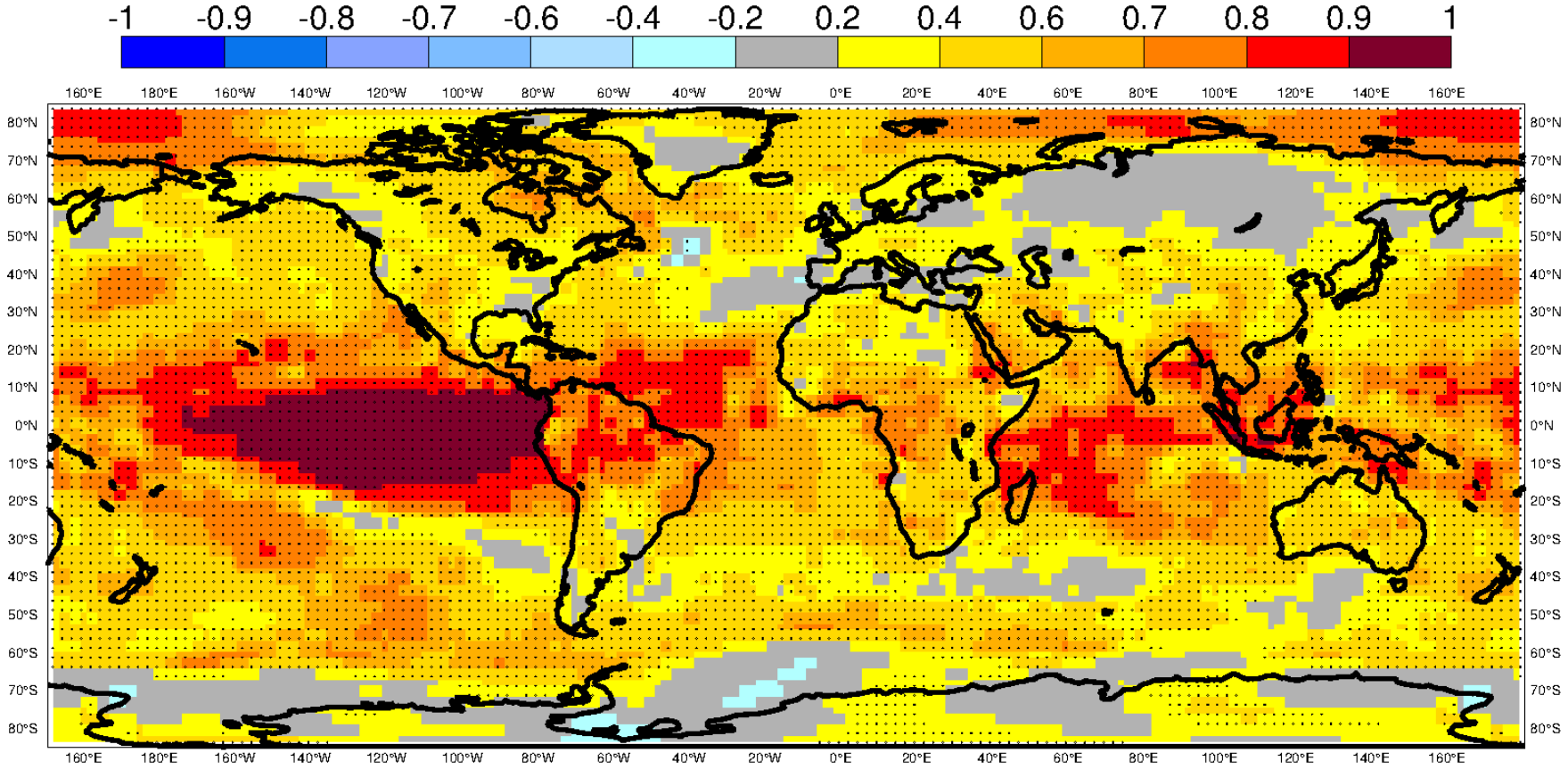
OND 2017

Shaded areas significant at 10% level

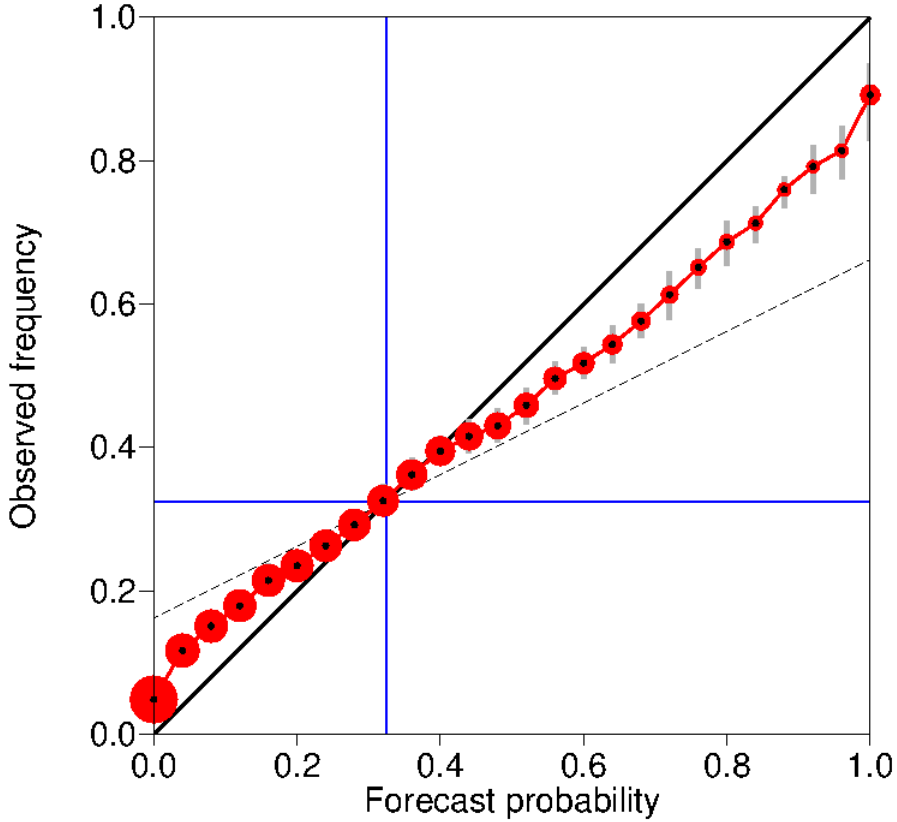
Solid contour at 1% level



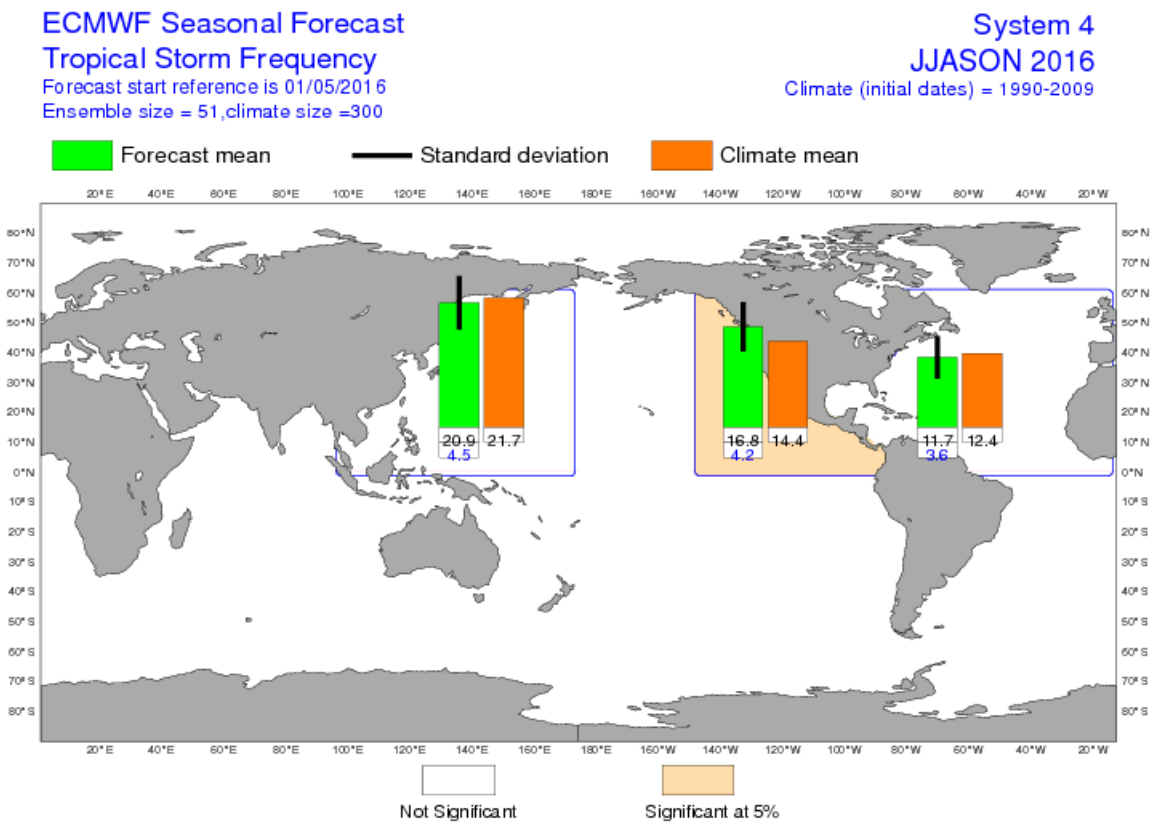
Anomaly Correlation Coefficient for 0001 with 25 ensemble members
Near-surface air temperature
Hindcast period 1981-2016 with start in September average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (1000 samples)

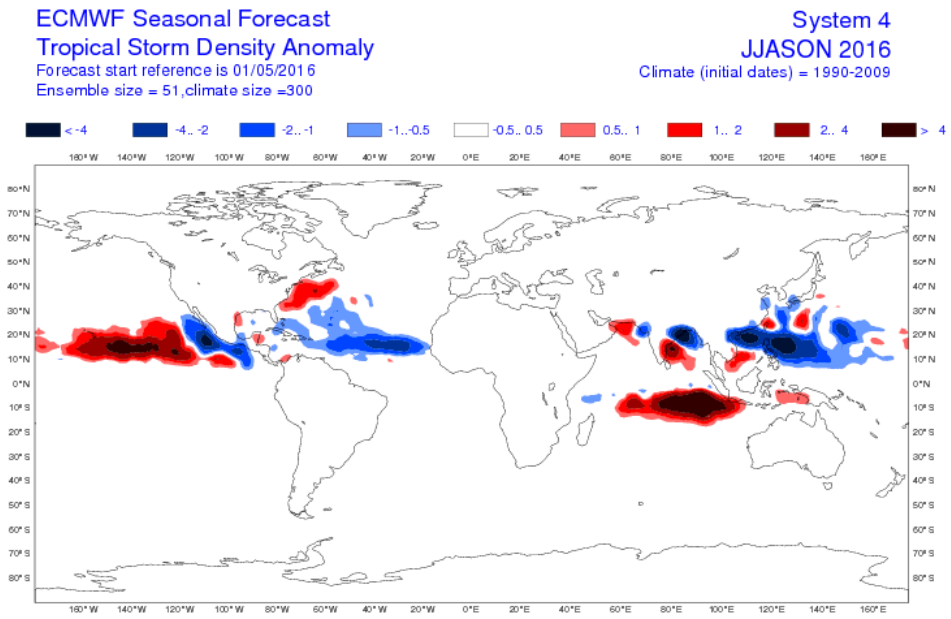


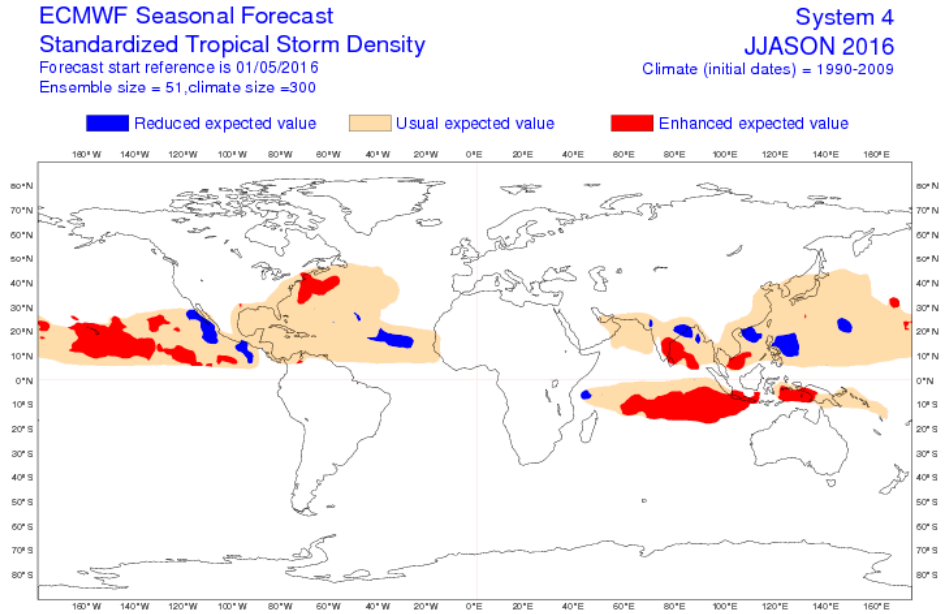
Reliability diagram for 0001 with 25 ensemble members
Near-surface air temperature anomalies above the upper tercile
Accumulated over global (land and sea points)
Hindcast period 1981-2016 with start in September average over months 2 to 4
Skill scores and 95% conf. intervals (1000 samples)
Brier skill score: 0.175 (0.134, 0.215)
Reliability skill score: 0.982 (0.976, 0.987)
Resolution skill score: 0.193 (0.155, 0.231)



Tropical storm forecasts



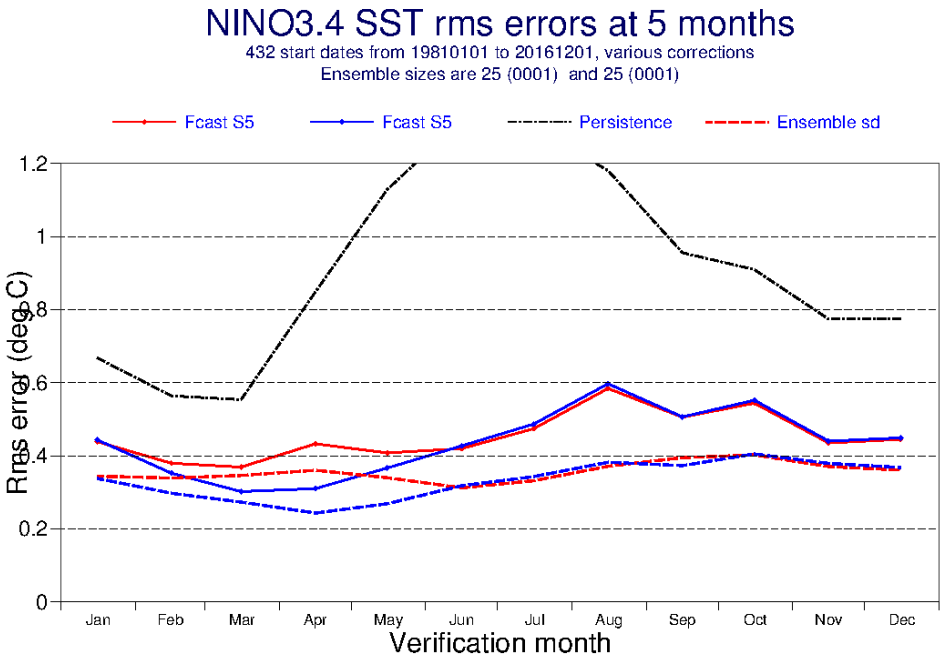


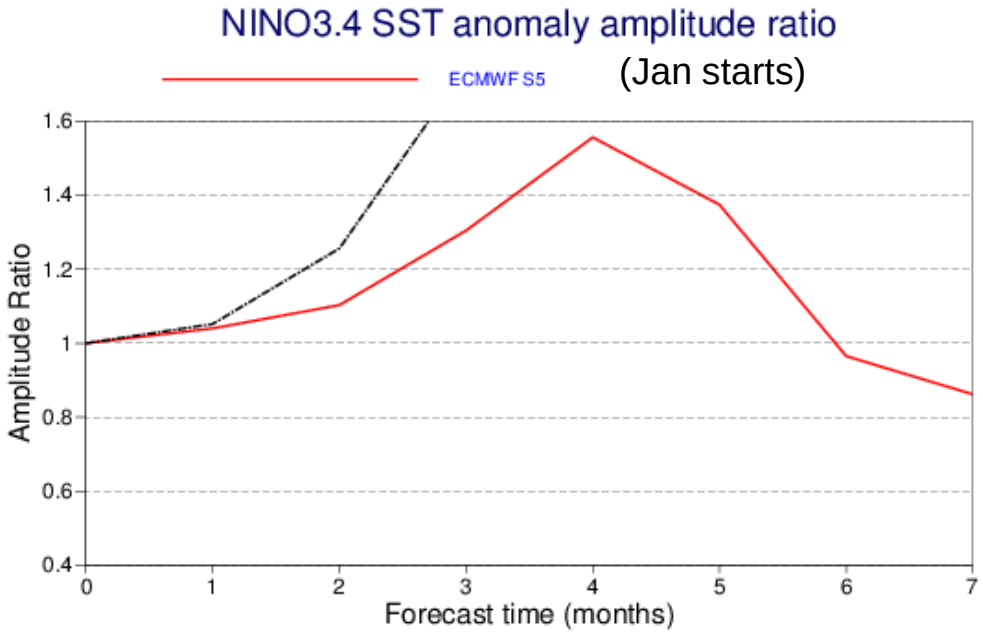


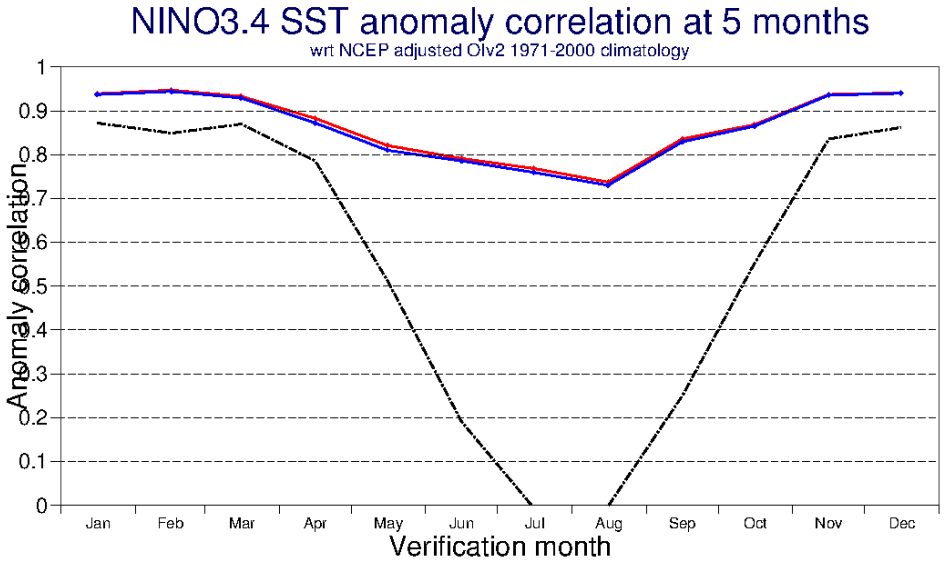
Variance adjustment

This very simple calibration scales the forecast climatological variance to match the observed climatological variance. The scaling is seasonally dependent. This calibration can substantially improve forecast products (and their verification scores). This calibration was used for our previous system, but was turned off in SEAS5.

SEAS5 verification includes the amplitude ratio, which should be used *a posteriori* to interpret the Nino plumes. This is important for forecasts of March, April and May.







Magics 2.34.0 (64 bit) - caliban - net - Tue May 1 17:24:21 2018

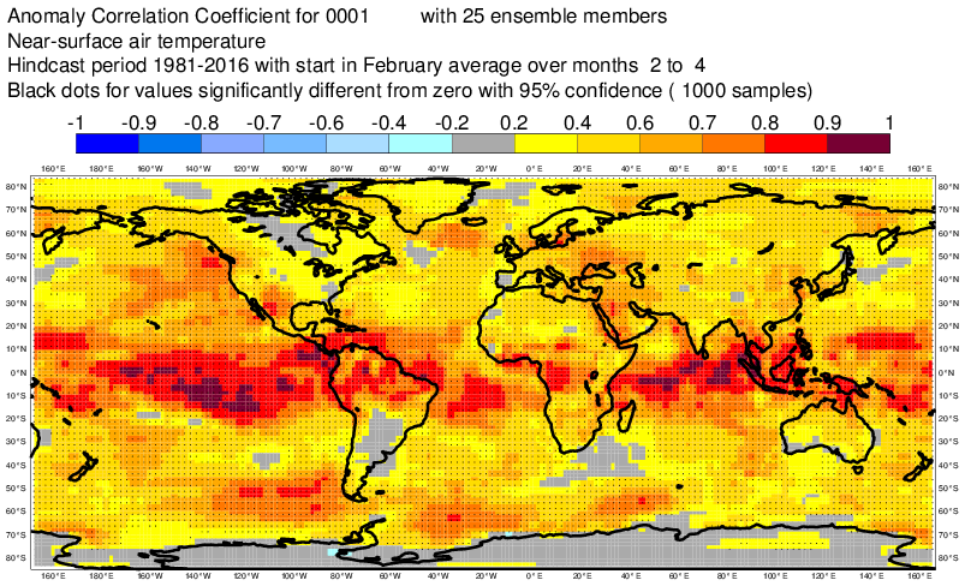
The challenge of assessment

- How good are the forecasts?

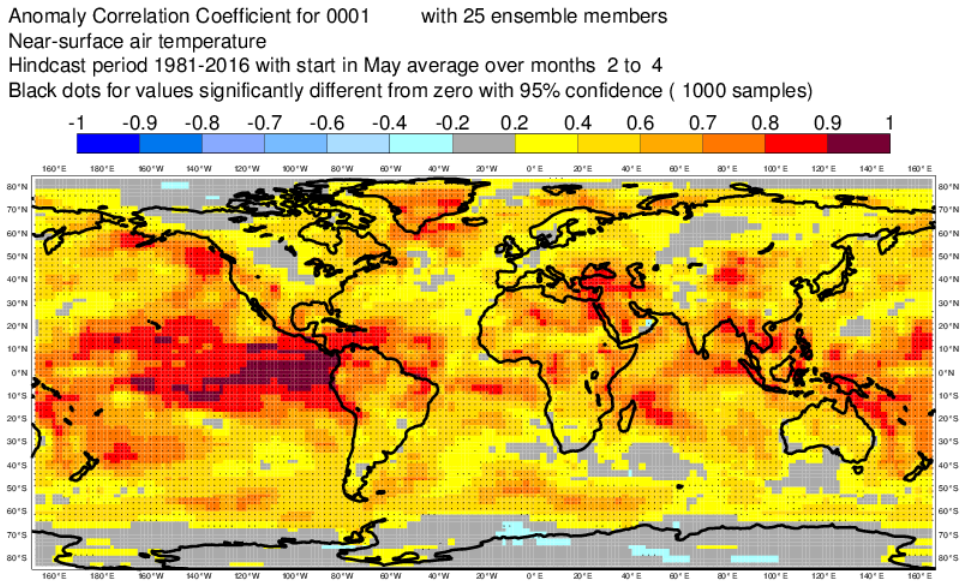
- Is a new system better or worse than the previous version?

Skill can vary a lot with season ……

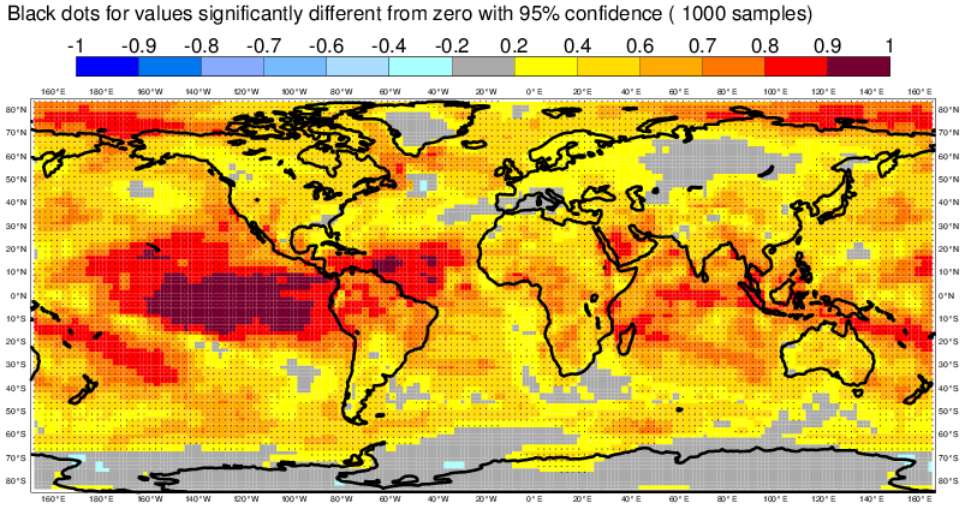
MAM



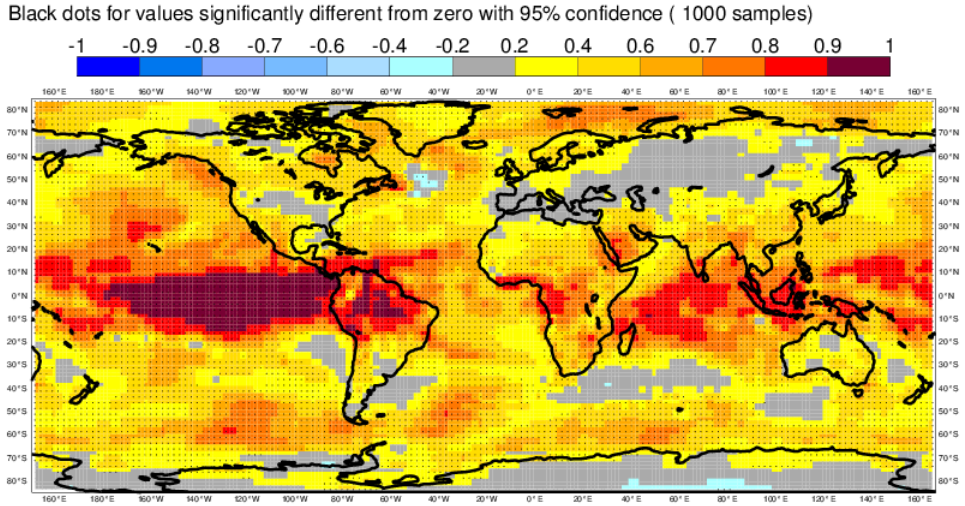
JJA



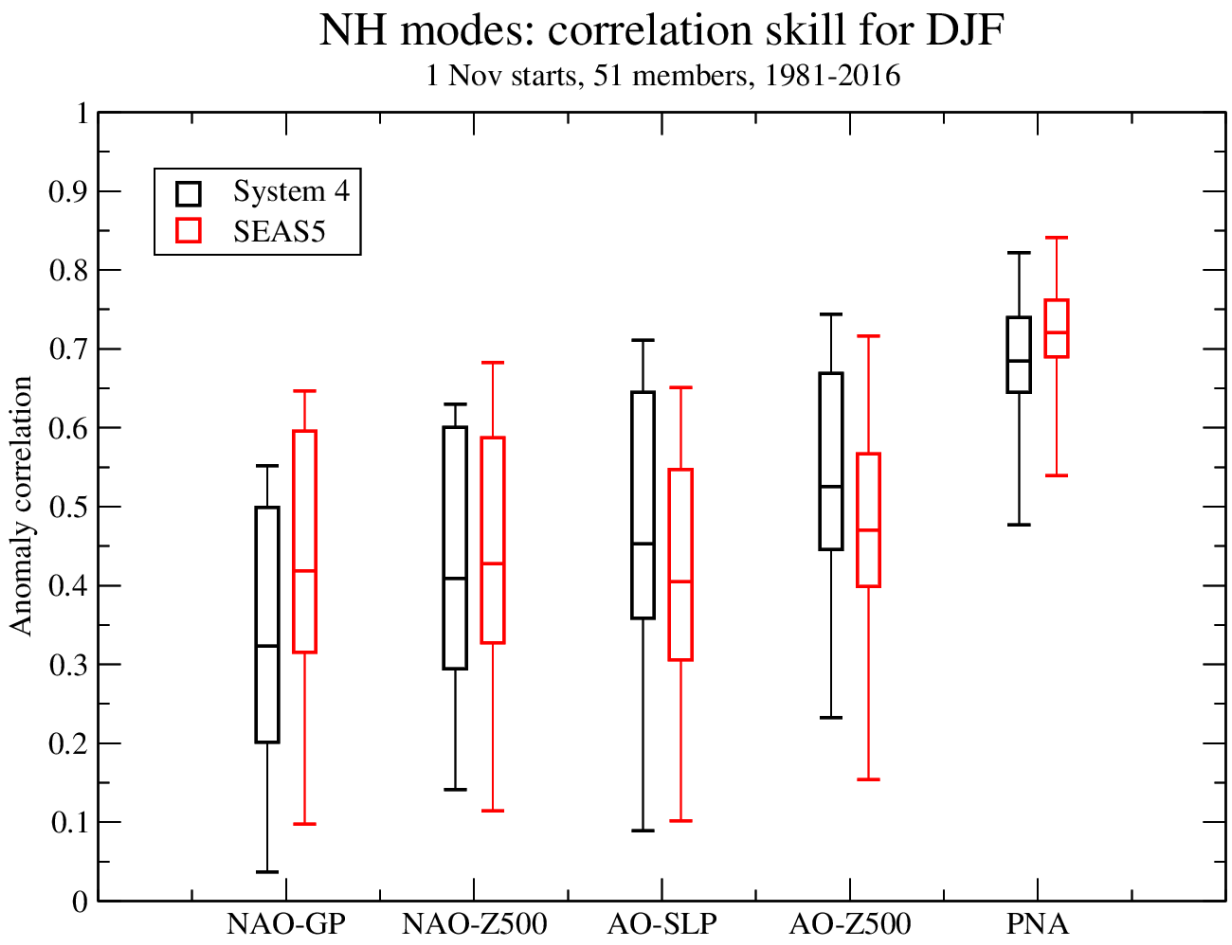
SON



DJF

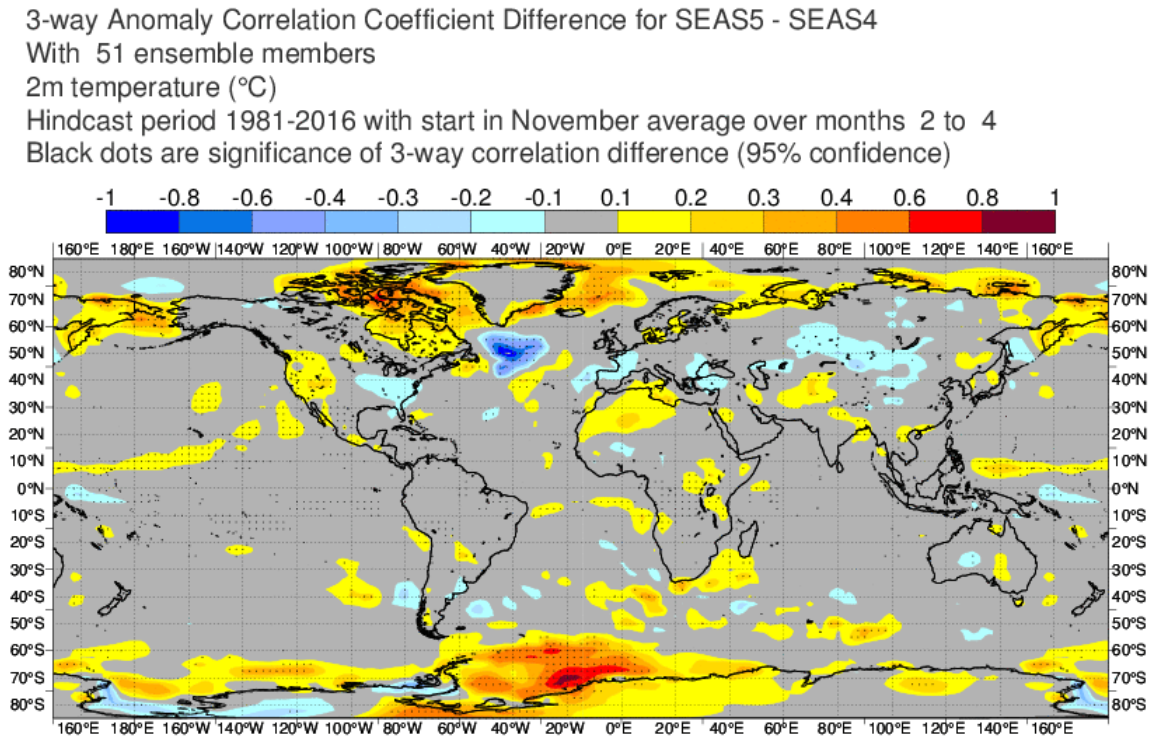


NAO, AO and PNA

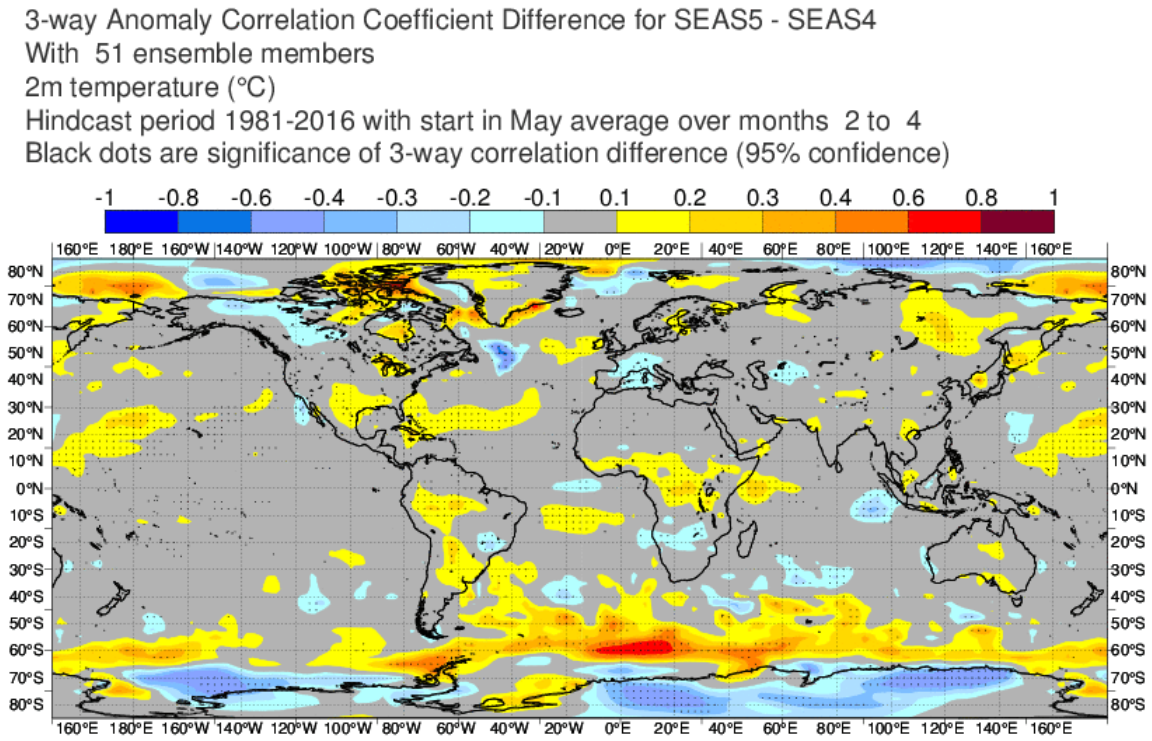


Box = 95% interval, bootstrapping on ensemble size

Whiskers = 95% interval, bootstrapping on sampled years


3-way Anomaly Correlation Coefficient Difference for SEAS5 - SEAS4
With 51 ensemble members
2m temperature (°C)
Hindcast period 1981-2016 with start in November average over months 2 to 4
Black dots are significance of 3-way correlation difference (95% confidence)
-1 -0.8 -0.6 -0.4 -0.3 -0.2 -0.1 0.1 0.2 0.3 0.4 0.6 0.8 1

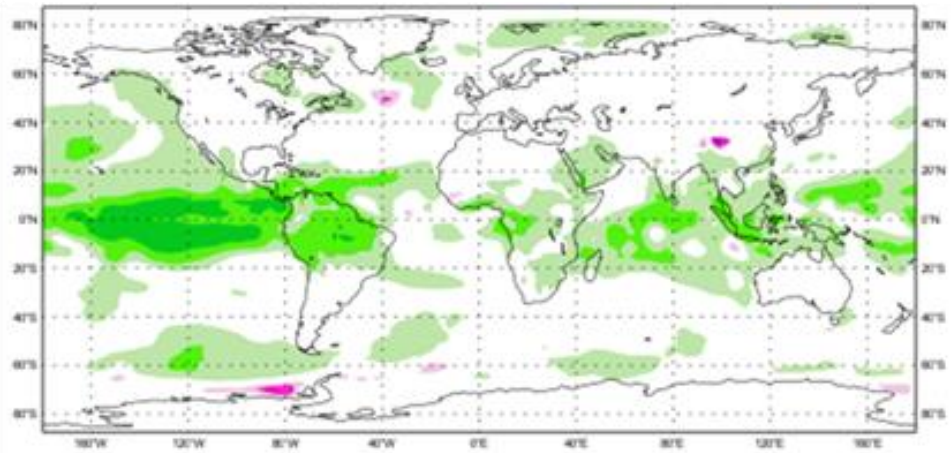
DJF

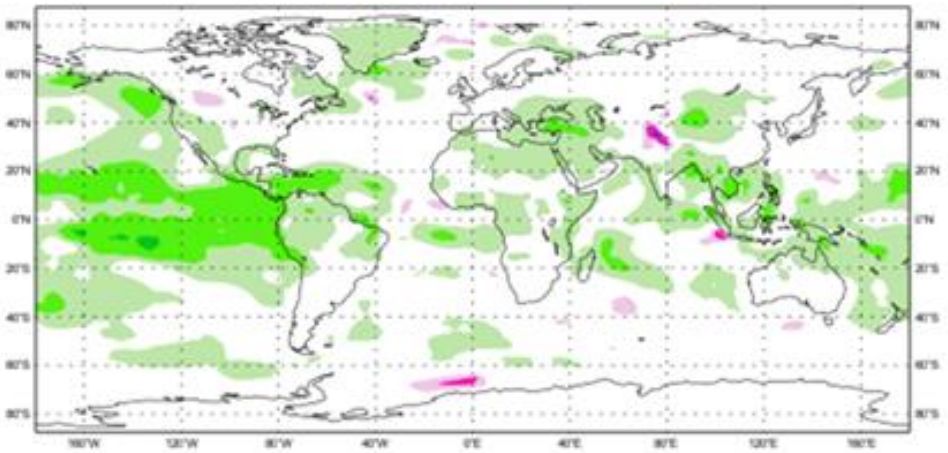

3-way Anomaly Correlation Coefficient Difference for SEAS5 - SEAS4
With 51 ensemble members
2m temperature (°C)
Hindcast period 1981-2016 with start in May average over months 2 to 4
Black dots are significance of 3-way correlation difference (95% confidence)
-1 -0.8 -0.6 -0.4 -0.3 -0.2 -0.1 0.1 0.2 0.3 0.4 0.6 0.8 1

JJA

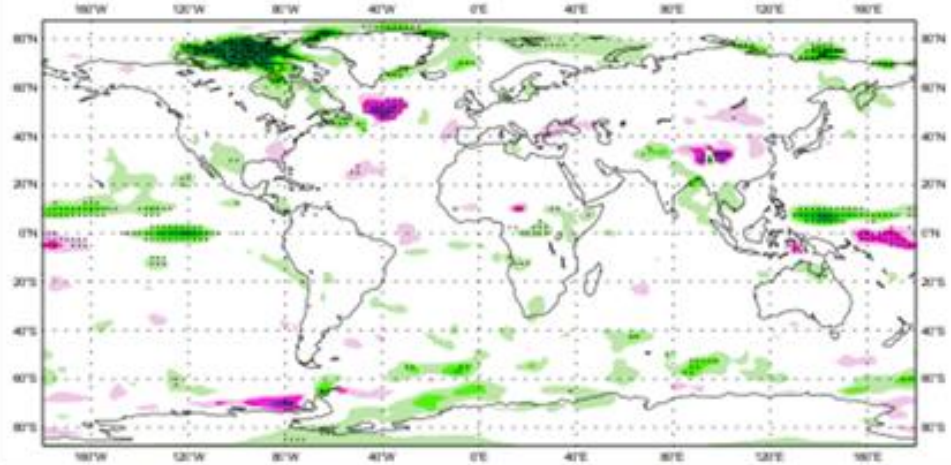
DJF T2m. CRPSS SEAS5 months 2-4

JJA T2m. CRPSS SEAS5 months 2-4

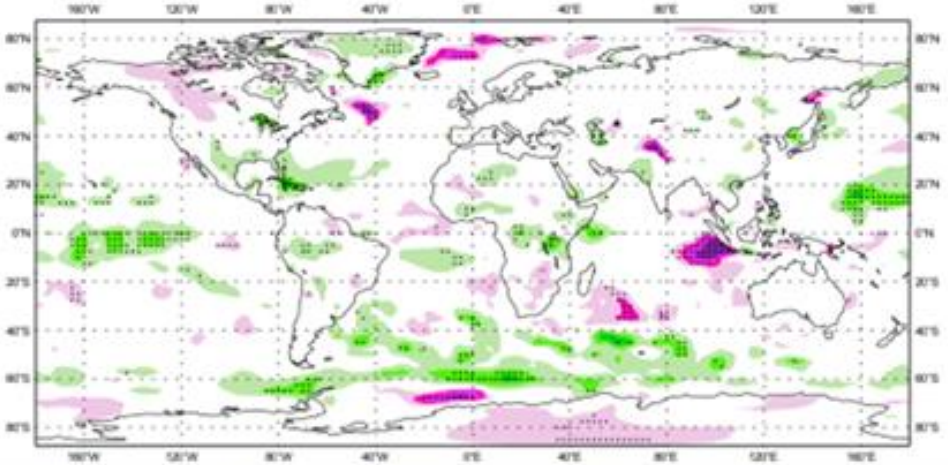




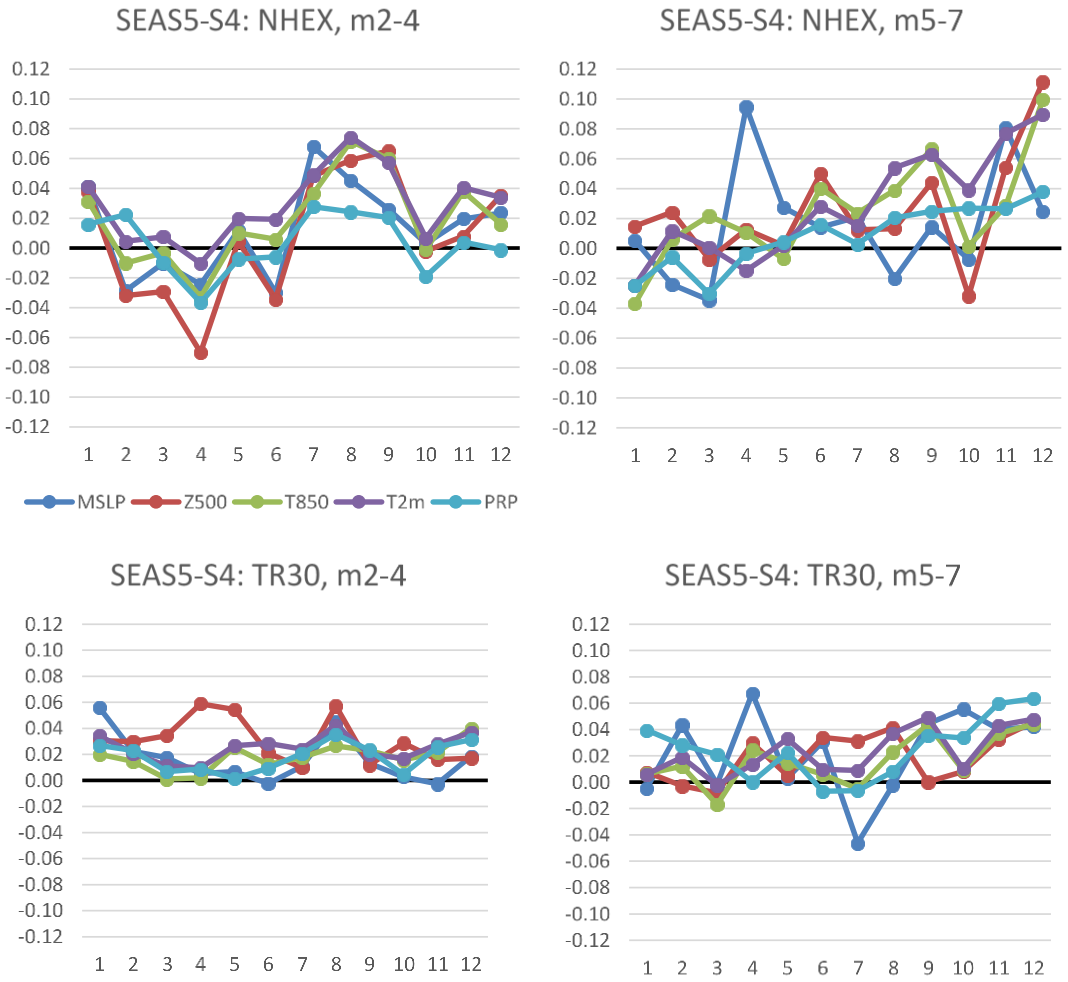
CRPSS SEAS5 – CRPSS S4

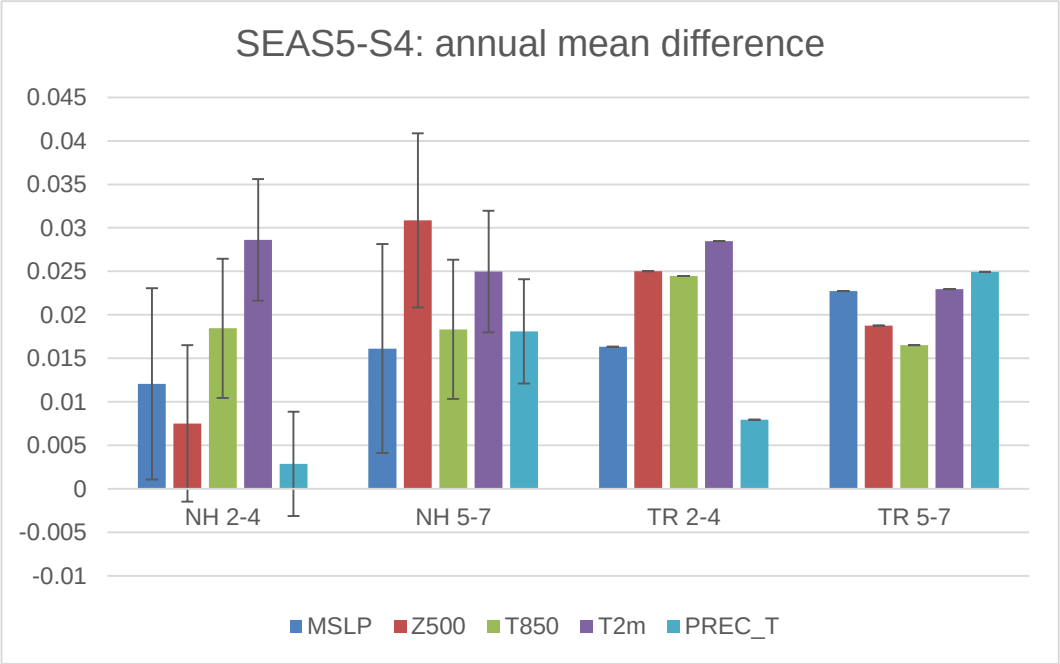


CRPSS SEAS5 – CRPSS S4



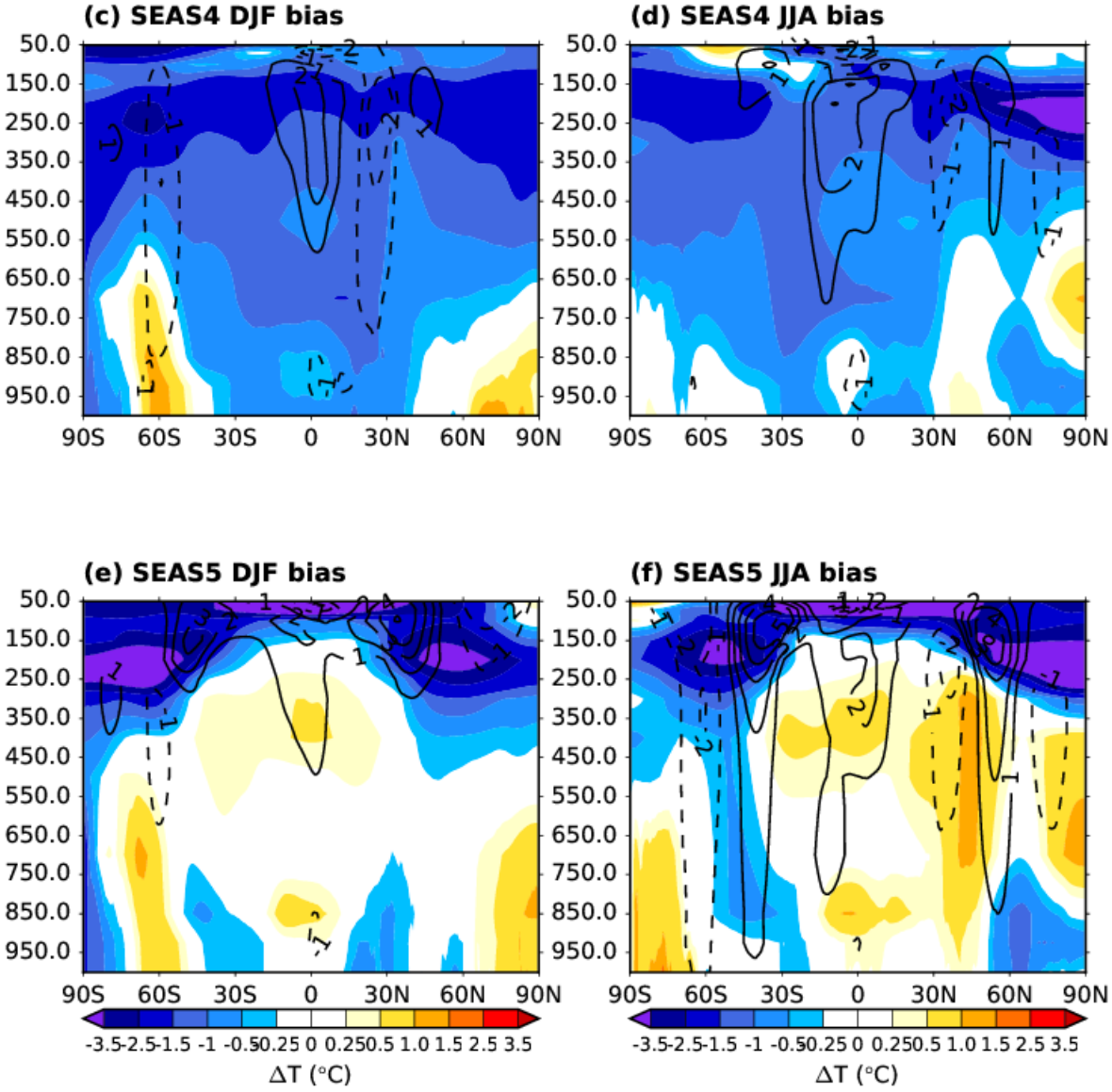
Aggregating scores over all start months





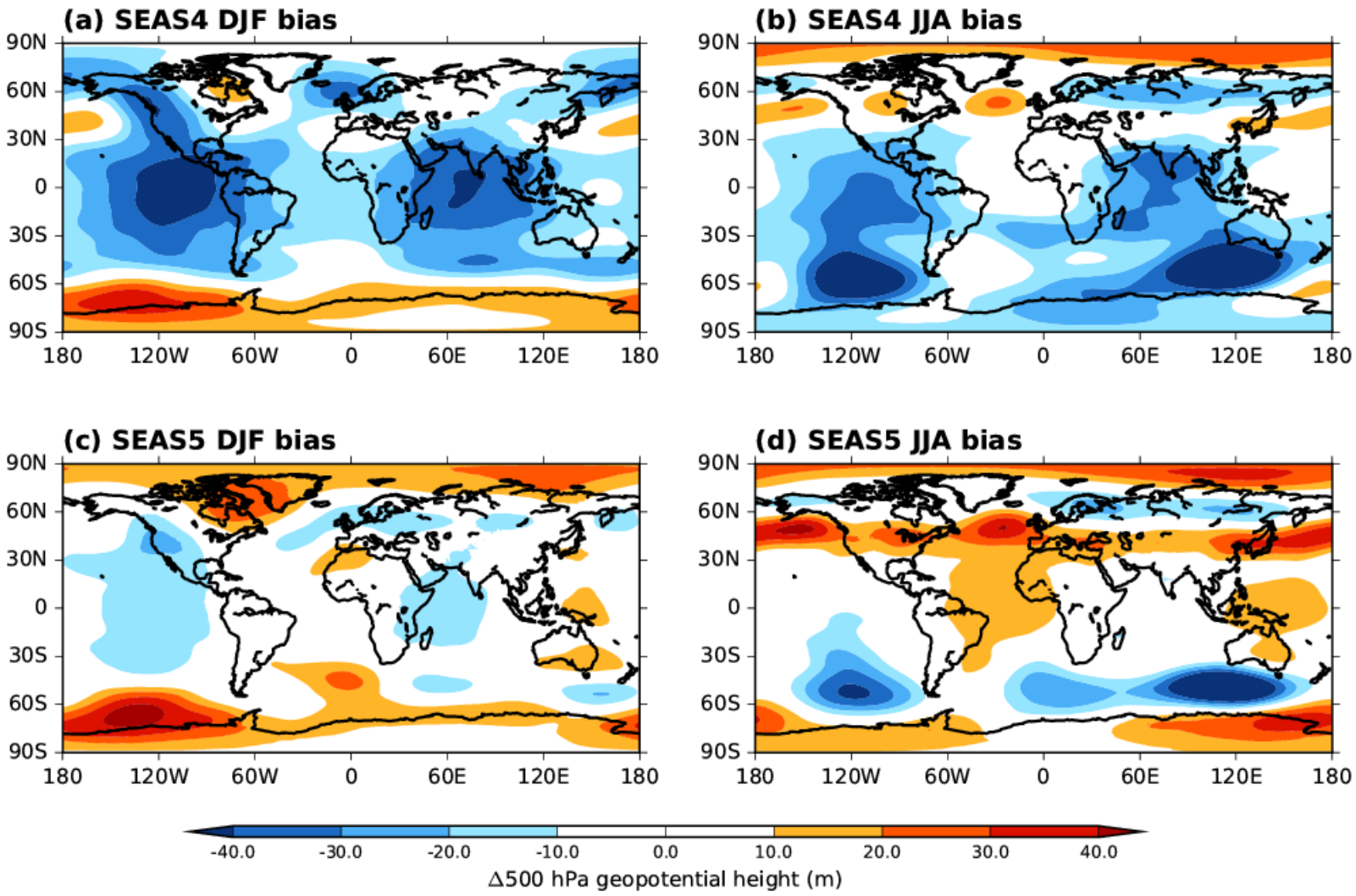
Model biases and performance as a function of model version

Zonal mean biases: T and u



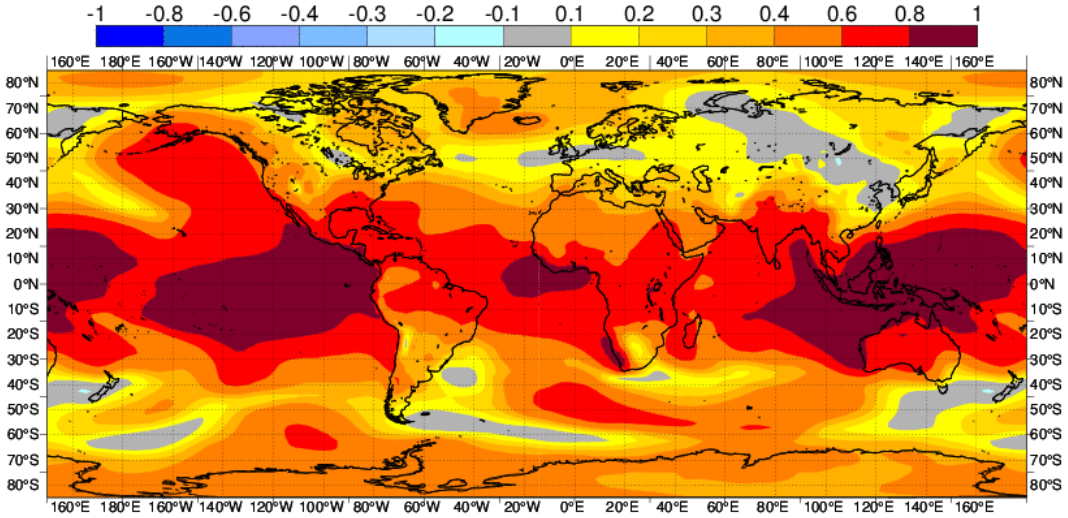
Figures from S. Johnson

Z500 bias

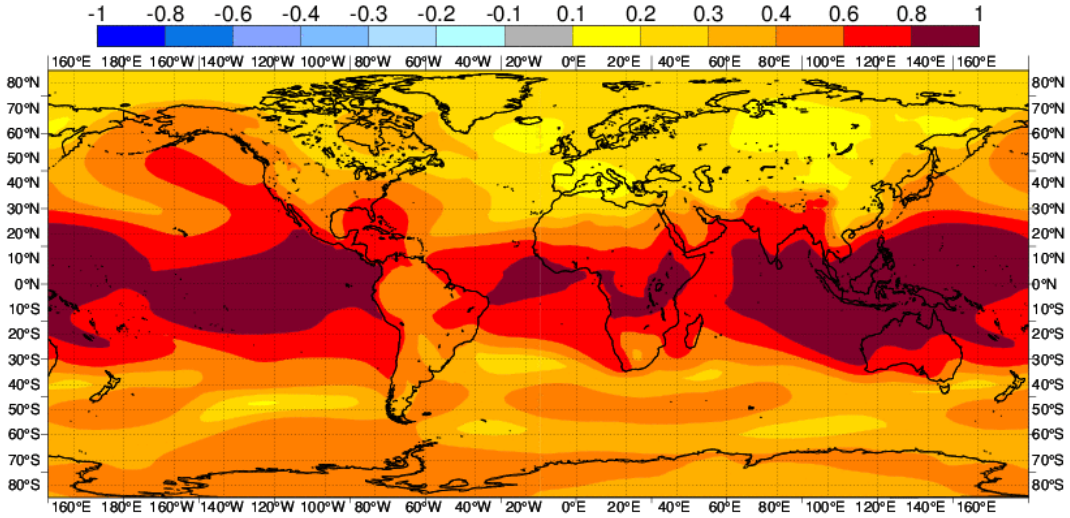


Figures from S. Johnson

Anomaly Correlation Coefficient for ECMWF S5 with 51 ensemble members
Mean sea level pressure (hPa)
Hindcast period 1981-2016 with start in November average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (200 samples)

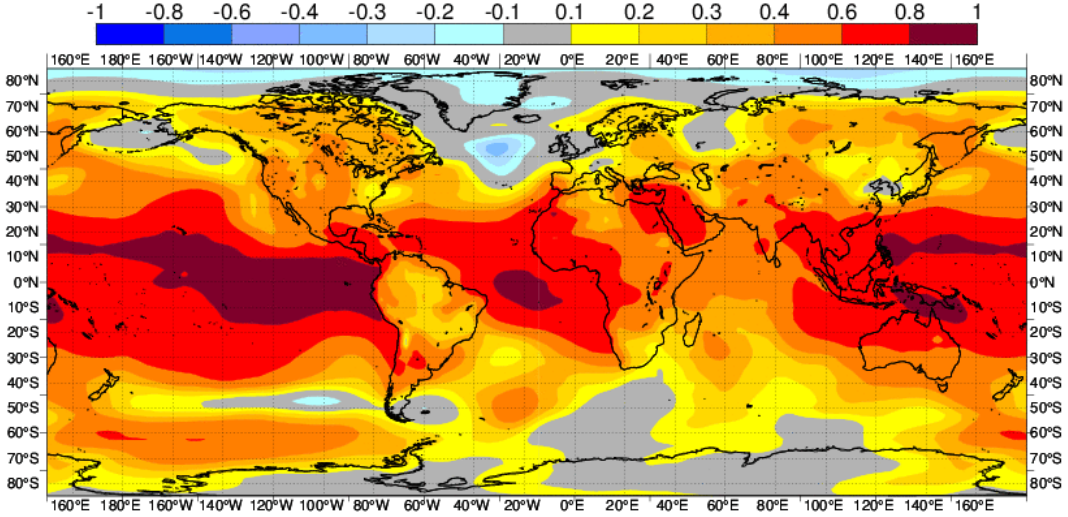


Perfect-model Anomaly Correlation Coefficient for ECMWF S5 with 51 ensemble members
Mean sea level pressure (hPa)
Hindcast period 1981-2016 with start in November average over months 2 to 4
Black dots where perfect model assumption is violated with 95% confidence (200 samples)

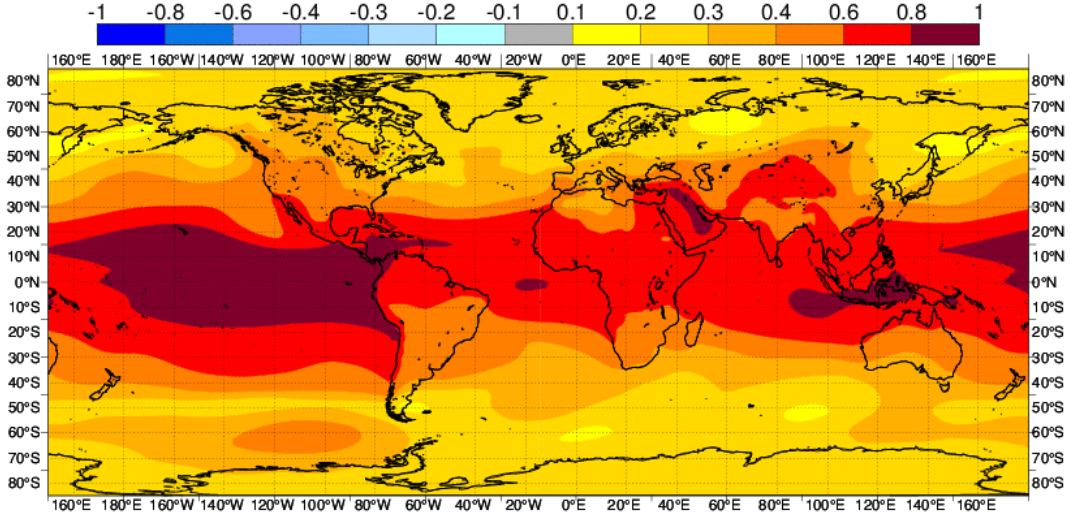


DJF

Hindcast period 1981-2016 with start in May average over months 2 to 4
Black dots for values significantly different from zero with 95% confidence (200 samples)



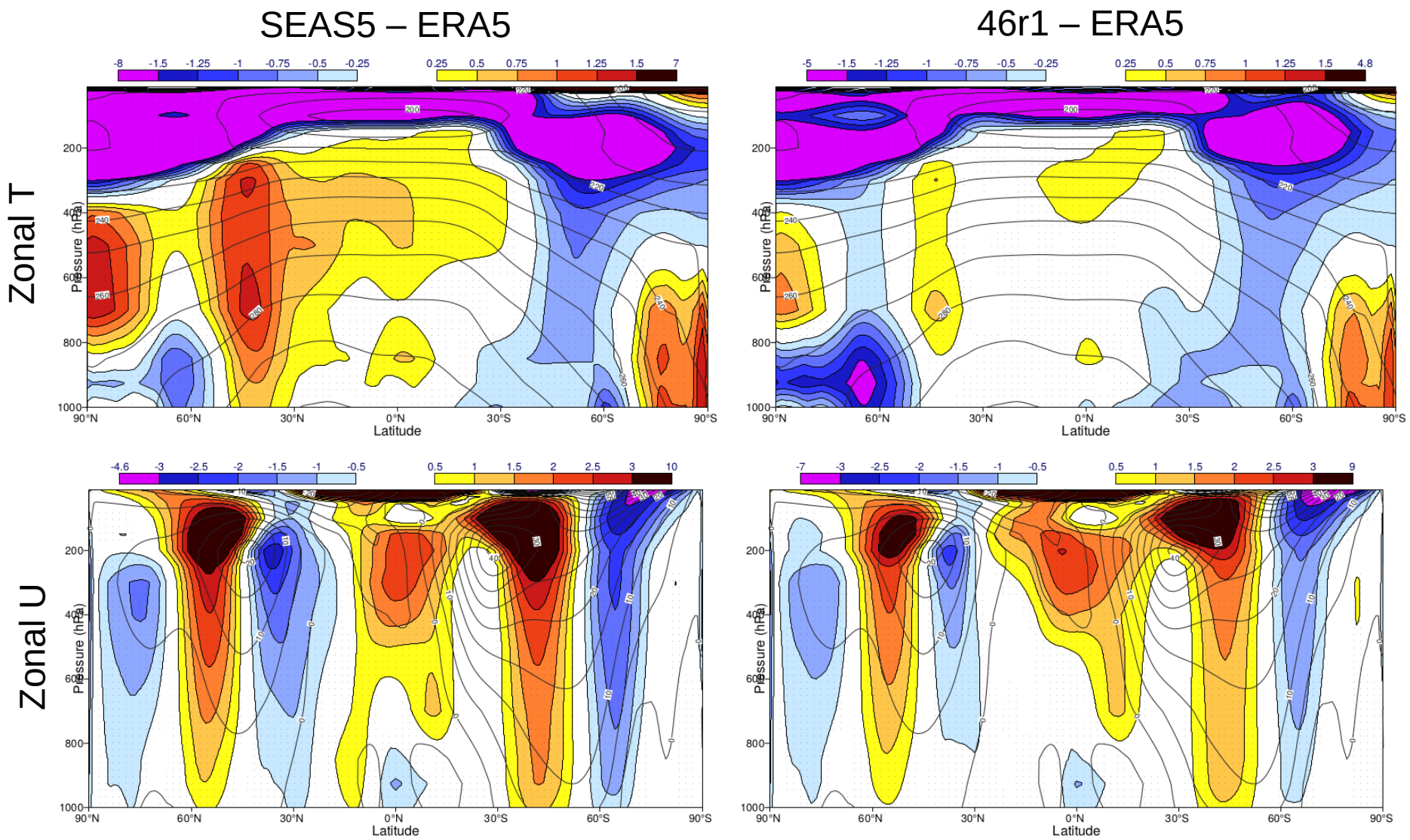
Hindcast period 1981-2016 with start in May average over months 2 to 4
Black dots where perfect model assumption is violated with 95% confidence (200 samples)



JJA

Recent improvements in summer jet bias (46r1)

Zonal wind bias reduced in magnitude, but stubbornly persists, perhaps reflecting the influence of the stratospheric temperature biases.

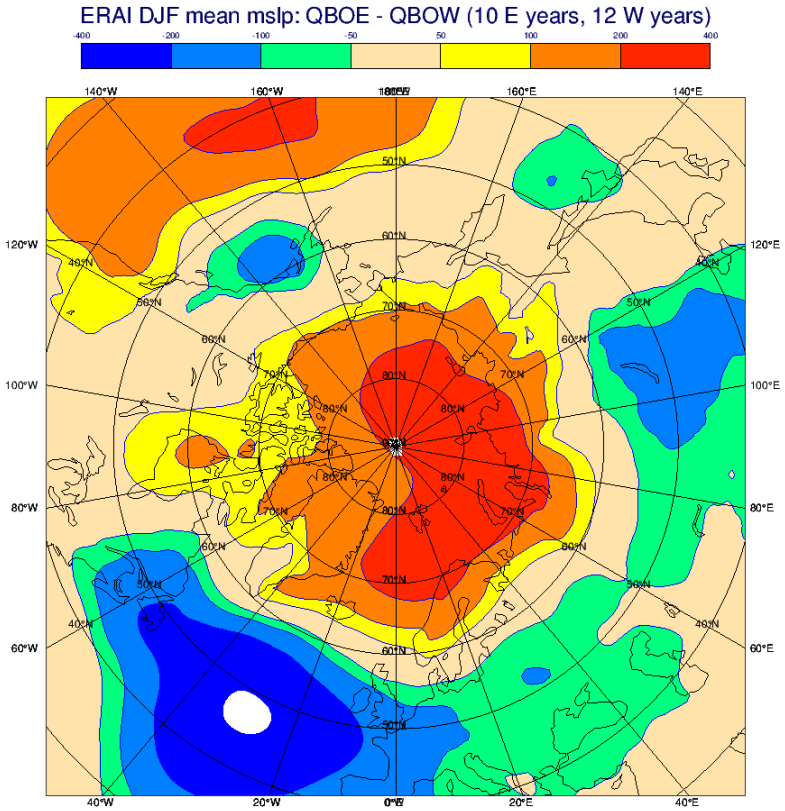


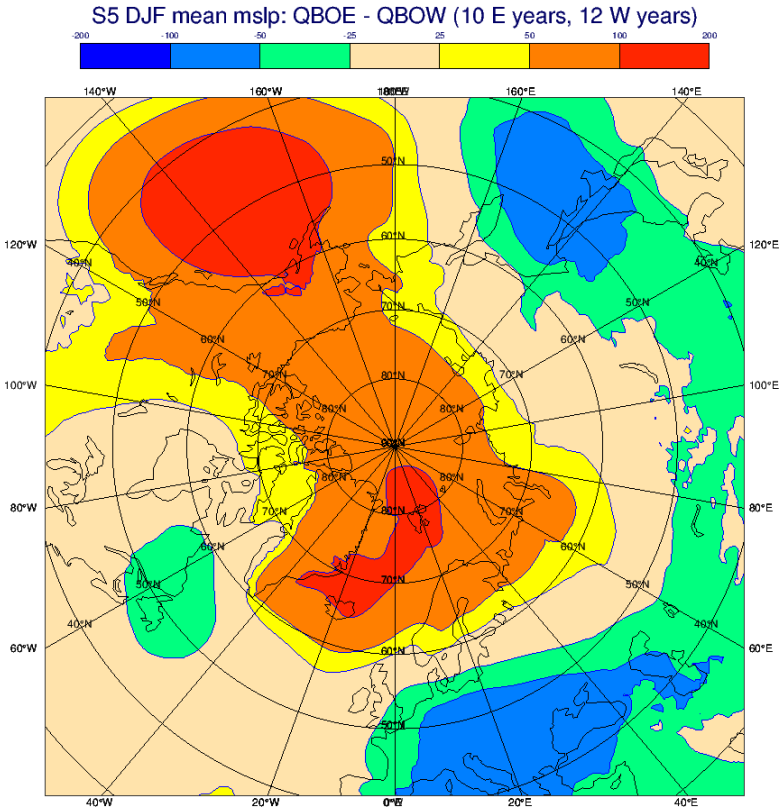
Figures from S. Johnson

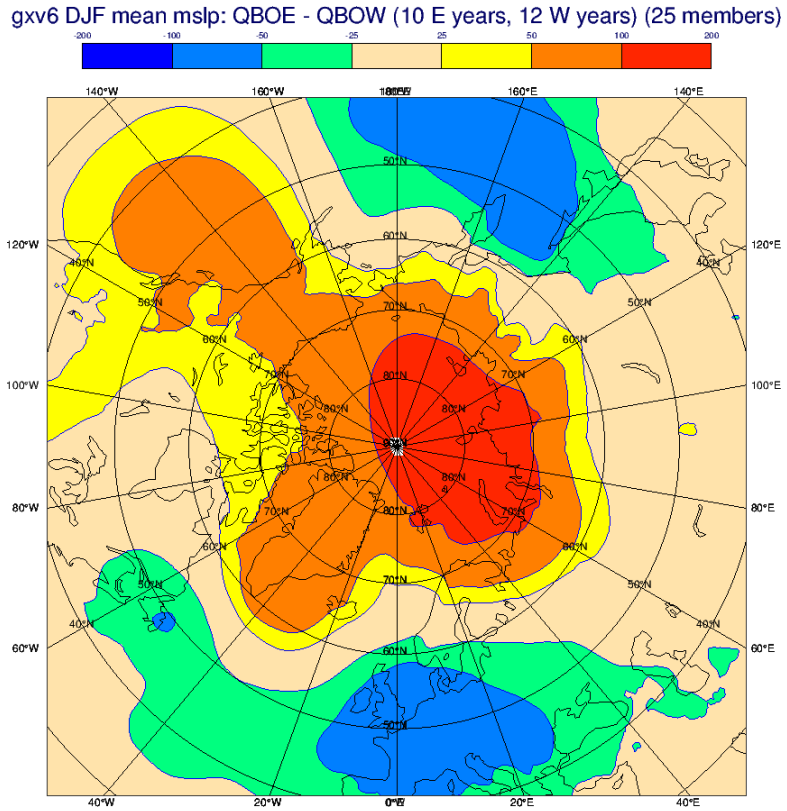
QBO teleconnections – NH winter MSLP

SEAS5

SEAS5 L137







Contour intervals: double

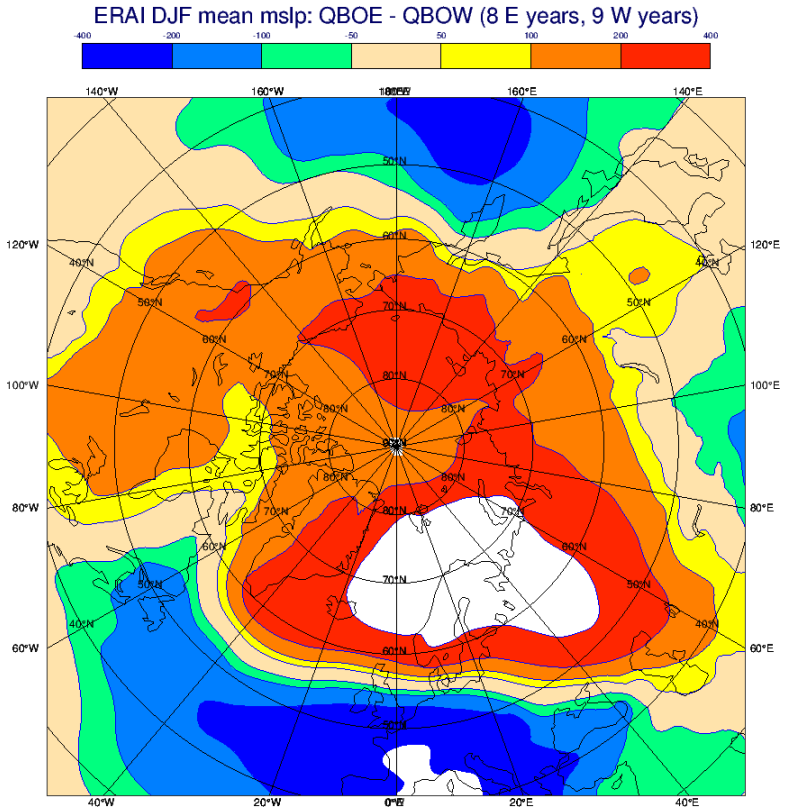
<- Contour intervals: single ->

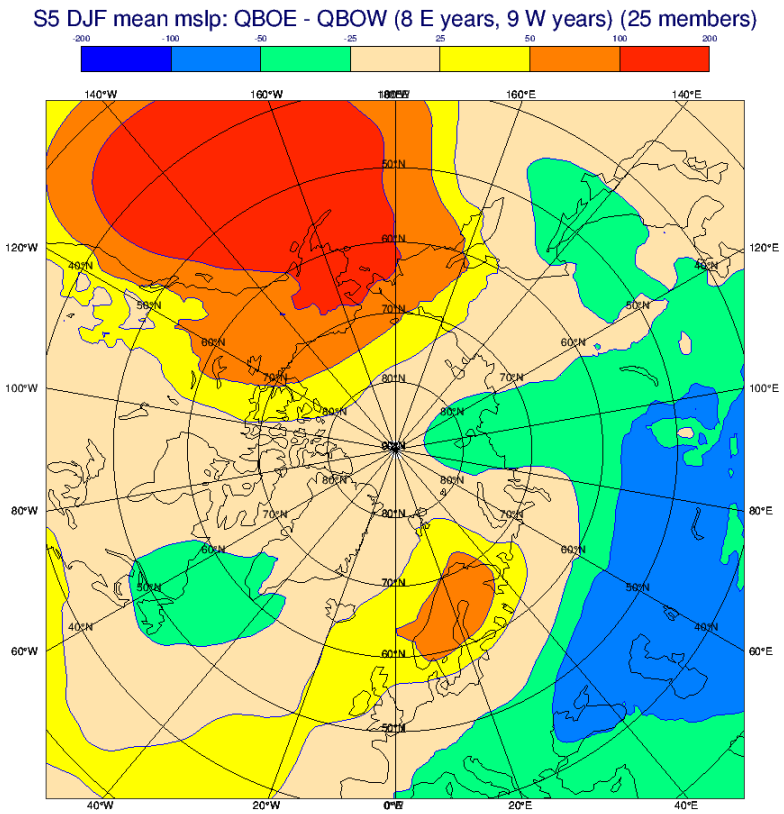
QBO composite years for 1981-2005, following Boer and Hamilton (2008). Contour interval is 0.5 hPa for ERAI, 0.25 hPa for model. Model composites based on 25 member ensemble.

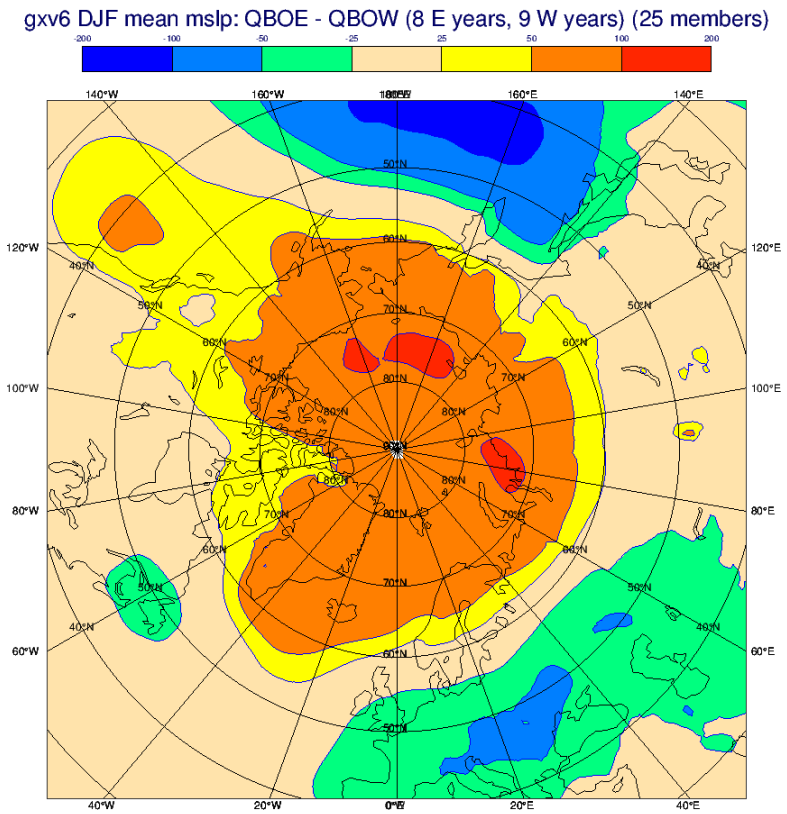
QBO teleconnections – NH winter MSLP – different QBO definition

SEAS5

SEAS5 L137







Contour intervals: double *<- Contour intervals: single ->*

QBO composite years for 1981-2016, following Martin Andrews, defined to maximize NAO teleconnection, equivalent to ~20 hPa level, approx 90 degree phase shift to previous definition. Contour interval is 0.5 hPa for ERAI, 0.25 hPa for model. Model composites based on 25 member ensemble. Note slightly smaller sample despite longer period.

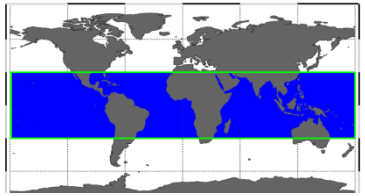
Alternative approach to model errors: multi-model

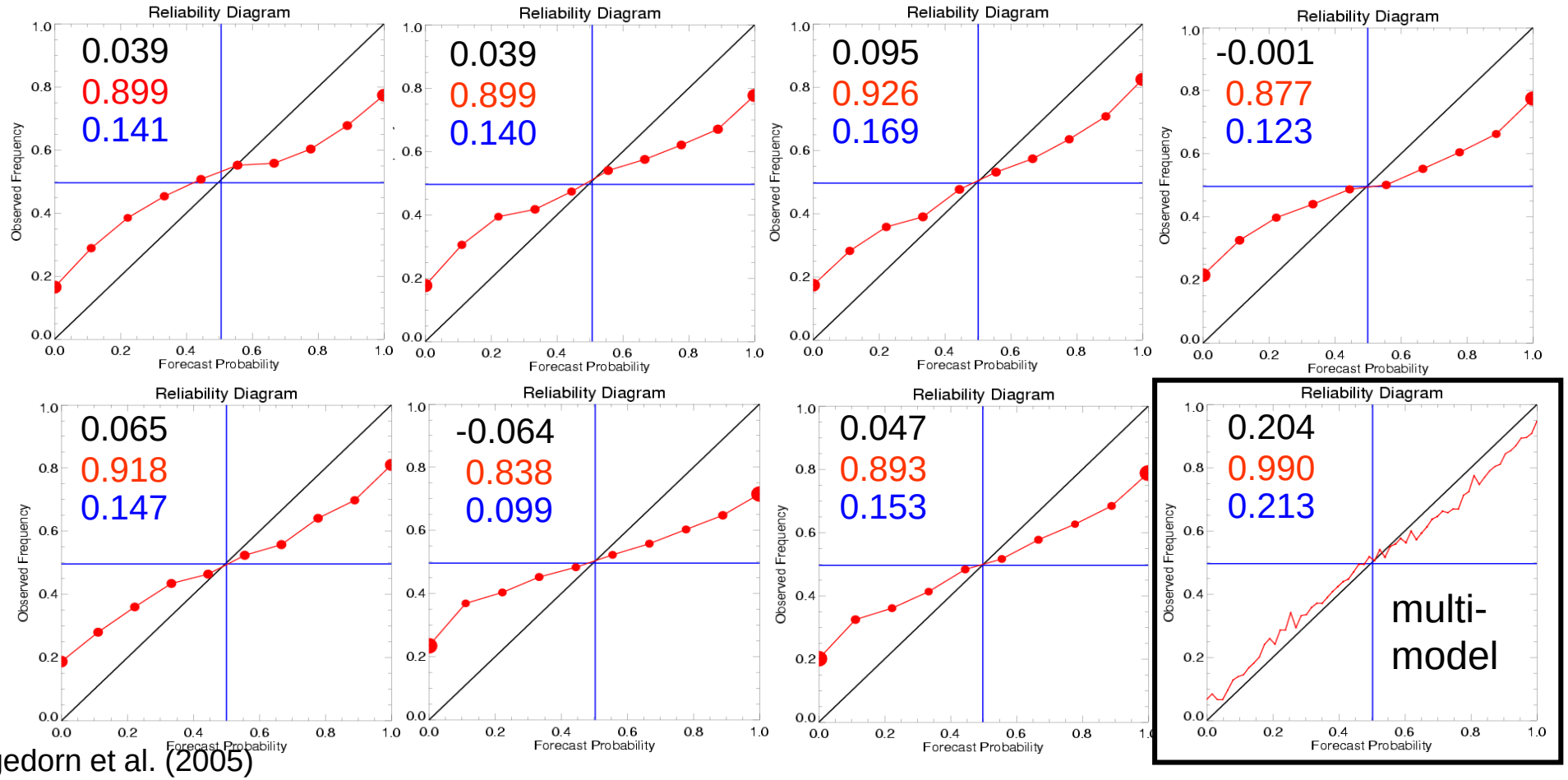
DEMETER: multi-model vs single-model

BSS
Rel-Sc
Res-Sc

Reliability diagrams (T2m > 0)

1-month lead, start date May, 1980 - 2001



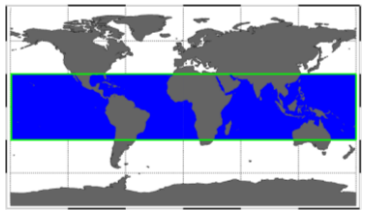


Hagedorn et al. (2005)

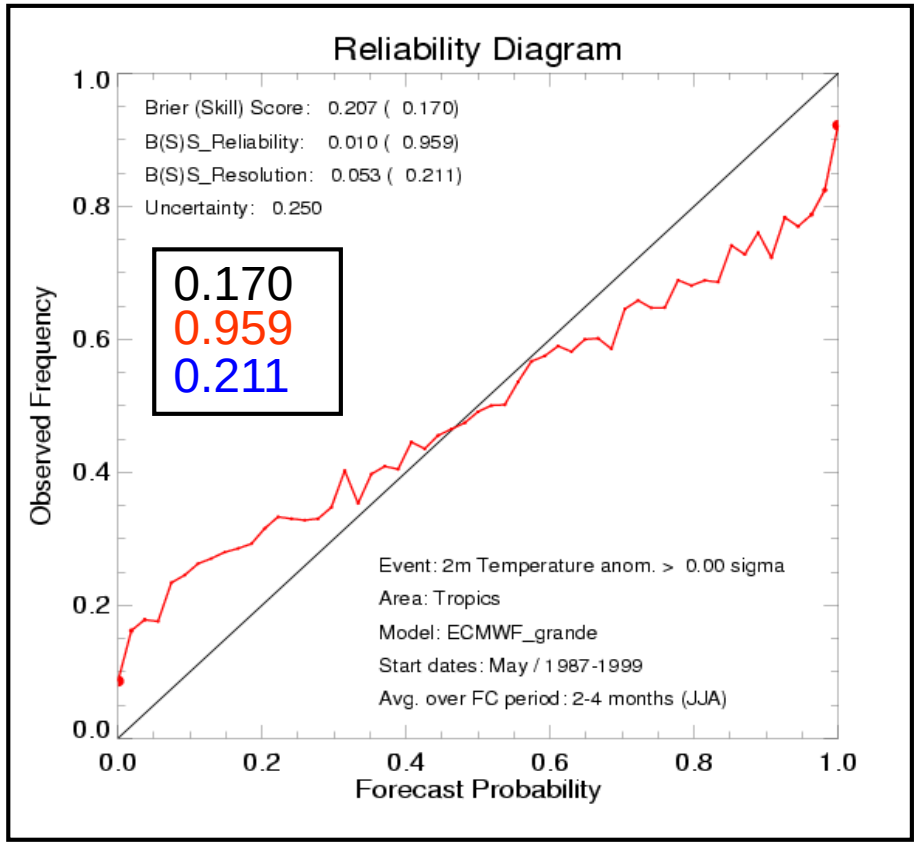
DEMETER: impact of ensemble size

BSS
Rel-Sc
Res-Sc

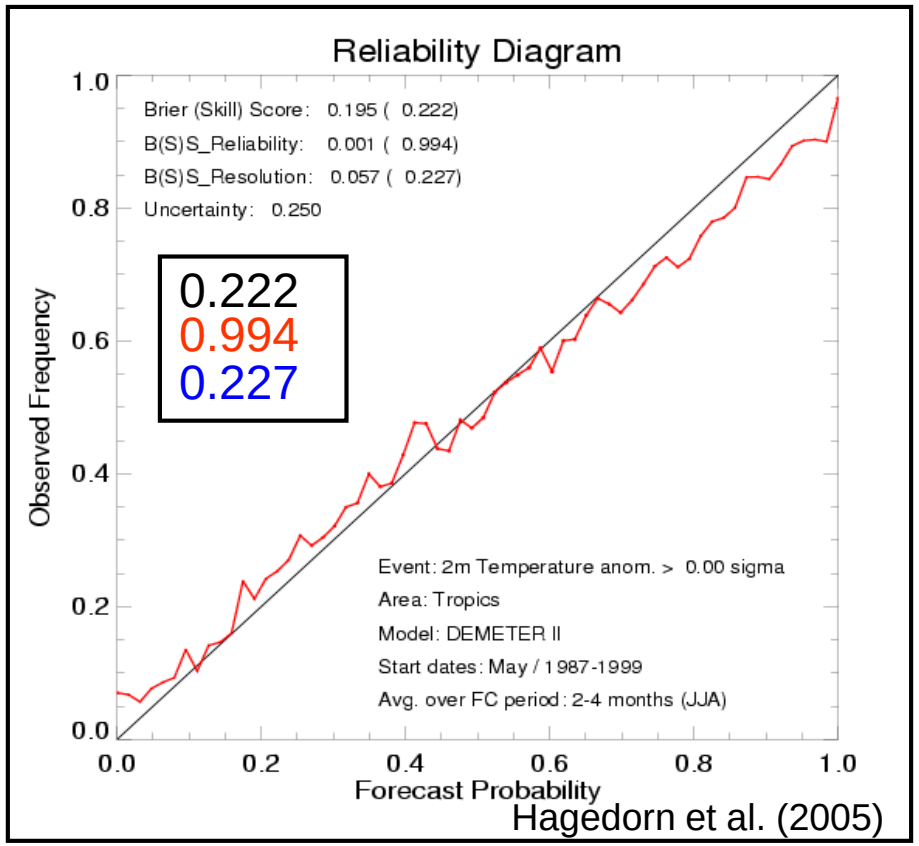
Reliability diagrams (T2m > 0)
1-month lead, start date May, 1987 - 1999



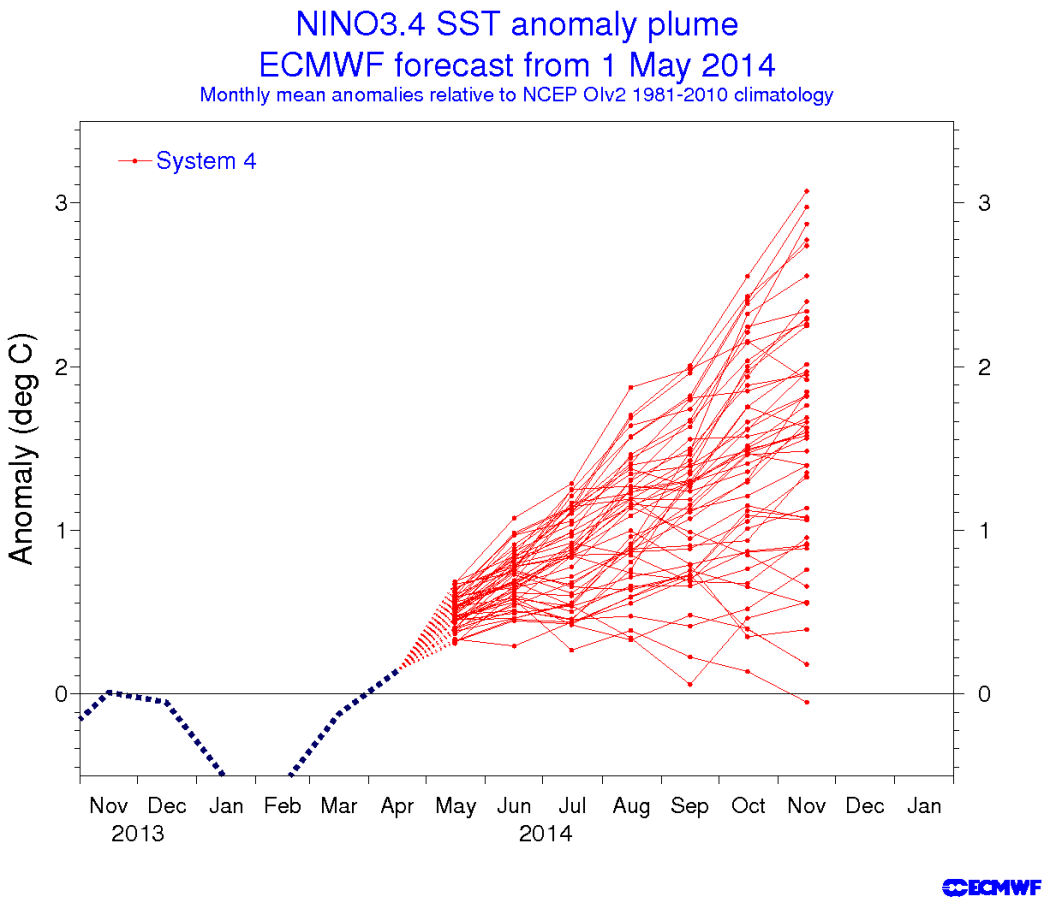
single-model [54 members]

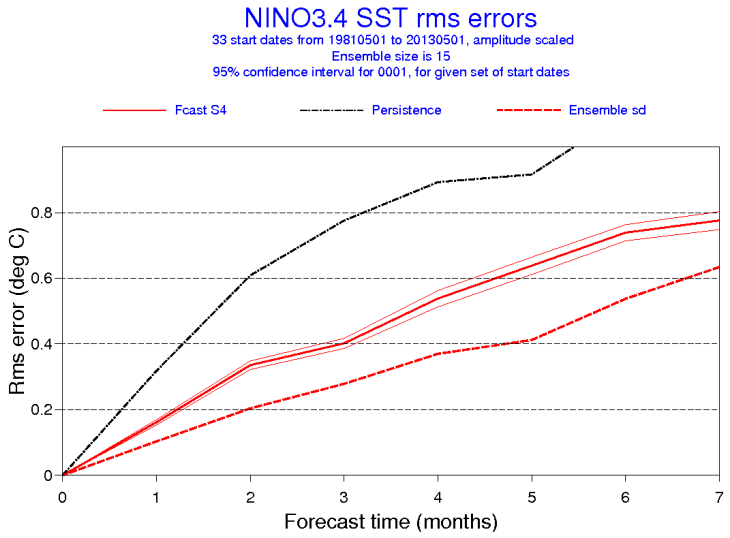


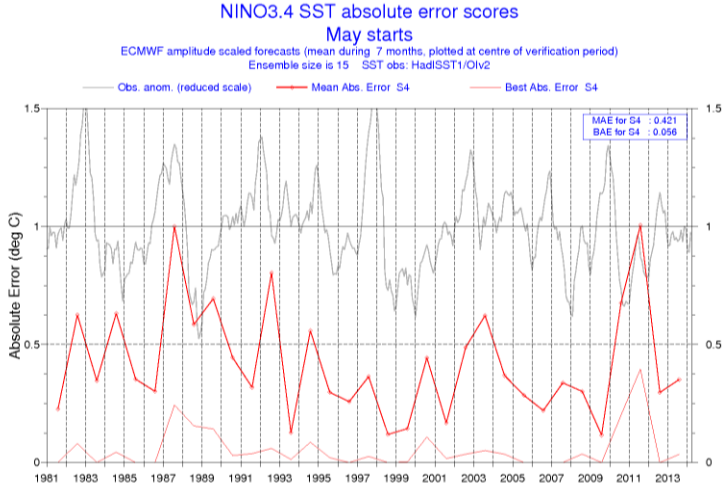
multi-model [54 members]



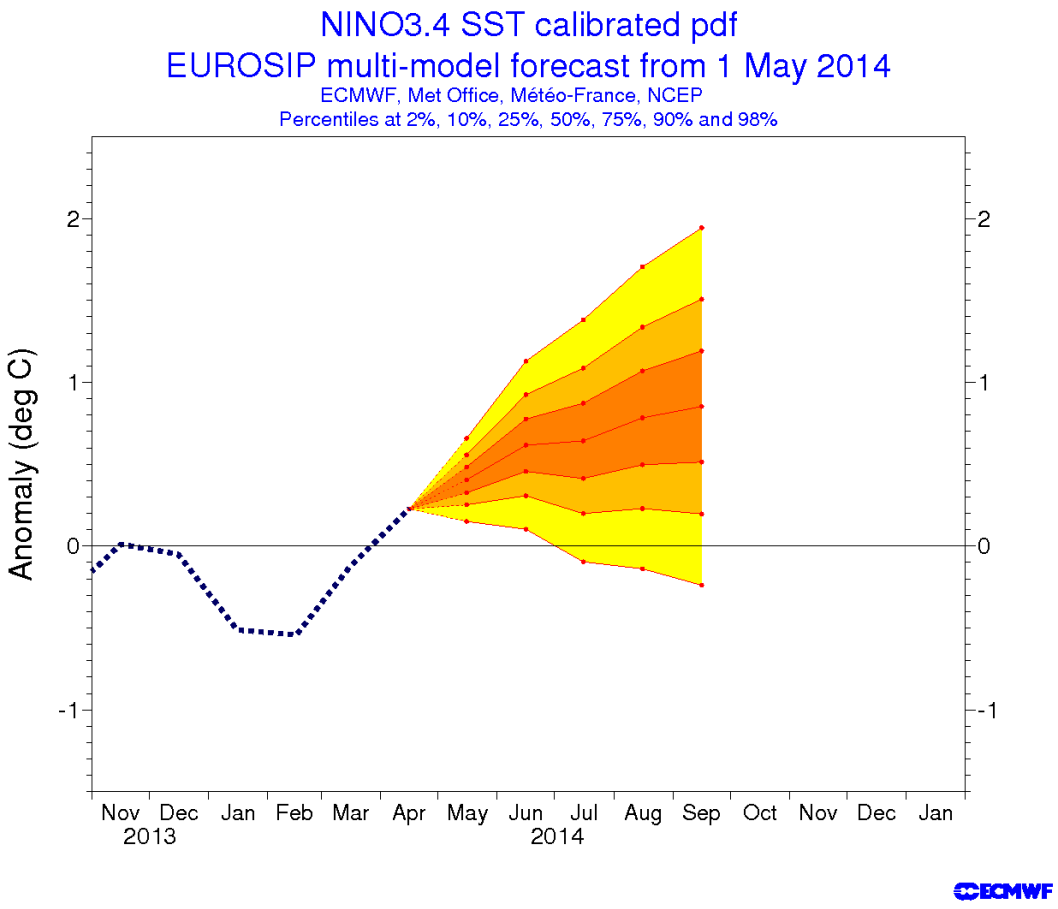
Hagedorn et al. (2005)

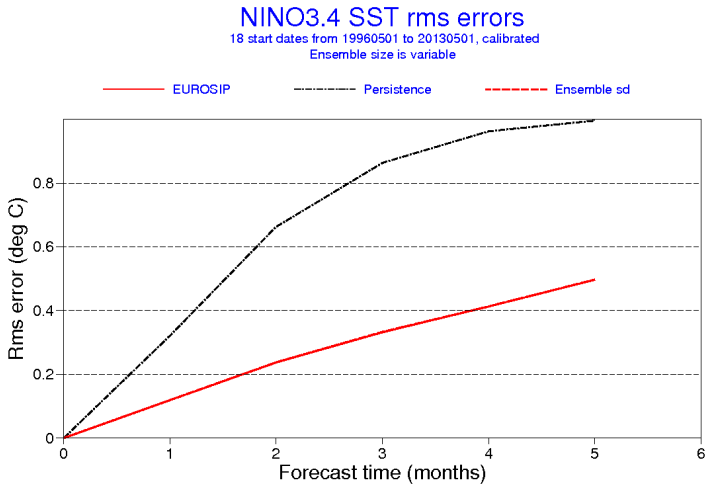


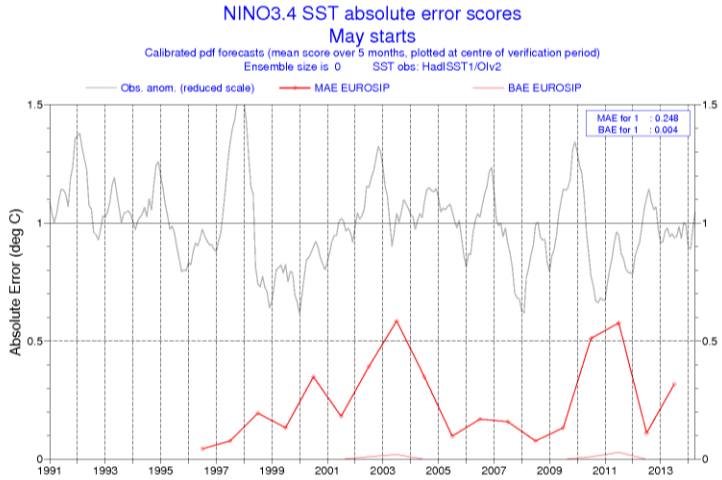




Past performance







Past performance

Summary

- Comprehensive numerical seasonal forecasting is a major undertaking
- Systems are relatively well developed, reliability is moderate, calibration is necessary
- Care is needed to ensure forecast information is properly interpreted and used sensibly

- Forecasting models are fairly realistic in many ways, but remaining errors are enough to substantially impact forecast skill and reliability, even after calibration
- Creating consistent initial conditions for past and present is a challenge, due in particular to the lack of observational data in the past. Observing systems are better now, but still need some improvements.
- Limited predictability and limited past data prevent us being sure about the skill levels of today's forecast systems, and assessment of improvements in forecast skill is difficult.
- Multi-model ensembles are helpful, **but** they only partially span the space of model errors.
- In the end, the only way to achieve high reliability is to build trustworthy models

Postscript: Why *do* we use full-physics numerical models for seasonal forecasting?

- To capture complexity of the real world
  - Multiple and variable sources of predictability
  - Non-linearity of interactions
  - BUT: to add value, predictability gain from these complexities must be greater than loss due to model errors. Is this the case?
- Epistemological: use initial state as a source of knowledge about the future
  - In principle, we can observe initial state and calculate future from laws of physics
  - Knowledge of past “initial states” is limited, and the past cannot be re-observed. Even if a linear, process-based approach to prediction could in principle capture all useful predictability, the evidence-base from the past is limiting and estimates of forecast uncertainty will be unreliable.
  - “Correction” of model forecast errors based on past performance is limited. To go beyond this, we need model biases to be so small that non-linear interactions are negligible. This means very small!
- To deal with a non-stationary climate

References and further reading

SEAS5 forecasts on www.ecmwf.int/en/forecasts/charts and https://climate.copernicus.eu/seasonal-forecasts

ECMWF Seasonal Forecast User Guide

SPECS fact sheets http://www.specs-fp7.eu/Fact%20sheets on seasonal forecasting

Boer, G.J. and K. Hamilton, 2008: QBO influence on extratropical predictive skill. Clim. Dyn. 31:987–1000. doi: 10.1007/s00382-008-0379-5

Hagedorn, R., F.J. Doblas-Reyes and T.N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept. *Tellus*, 57A, 219-233.

Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T., and Vitart, F, 2011: The new ECMWF seasonal forecast system (System 4), ECMWF Tech. Memo 656., DOI:10.21957/4nery093i

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H., and Monge-Sanz, B., 2018: SEAS5: The new ECMWF seasonal forecast system, Geosci. Model Dev. DOI: 10.5194/gmd-2018-228

Stockdale, T., Alonso-Balmaseda, M, Johnson, S, Ferranti, L, Molteni, F, Magnusson, L, Tietsche, S, Vitart, F, Decremer, D, Weisheimer, A, Roberts, CD, Balsamo, G, Keeley, S, Mogensen, K, Zuo, H, Mayer, M, and Monge-Sanz, BM, 2018: SEAS5 and the future evolution of the long-range forecast system. ECMWF Tech Memo 835, DOI: 10.21957/z3e92di7y