Sub-seasonal to seasonal forecast Verification

Frédéric Vitart and Laura Ferranti
European Centre for Medium-Range Weather Forecasts

INDEX

1. **Context:** S2S prediction

2. **Issues with S2S verification**

    - Space/Time Averaging
    - Conditional skill
    - Use of re-forecasts for calibration and verification

3. **Verification of weather regime transitions**

4. **Extreme weather verification**

S2S Prediction

Bridging the gap between Climate and weather prediction

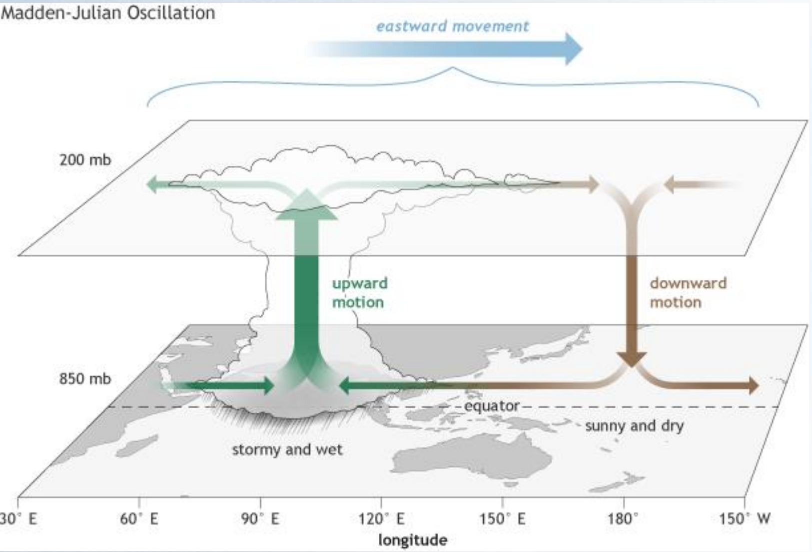
A particularly difficult time range: Is it an atmospheric initial condition problem as medium-range forecasting or is it a boundary condition problem as seasonal forecasting? “Predictability Desert”

Some sources of predictability :

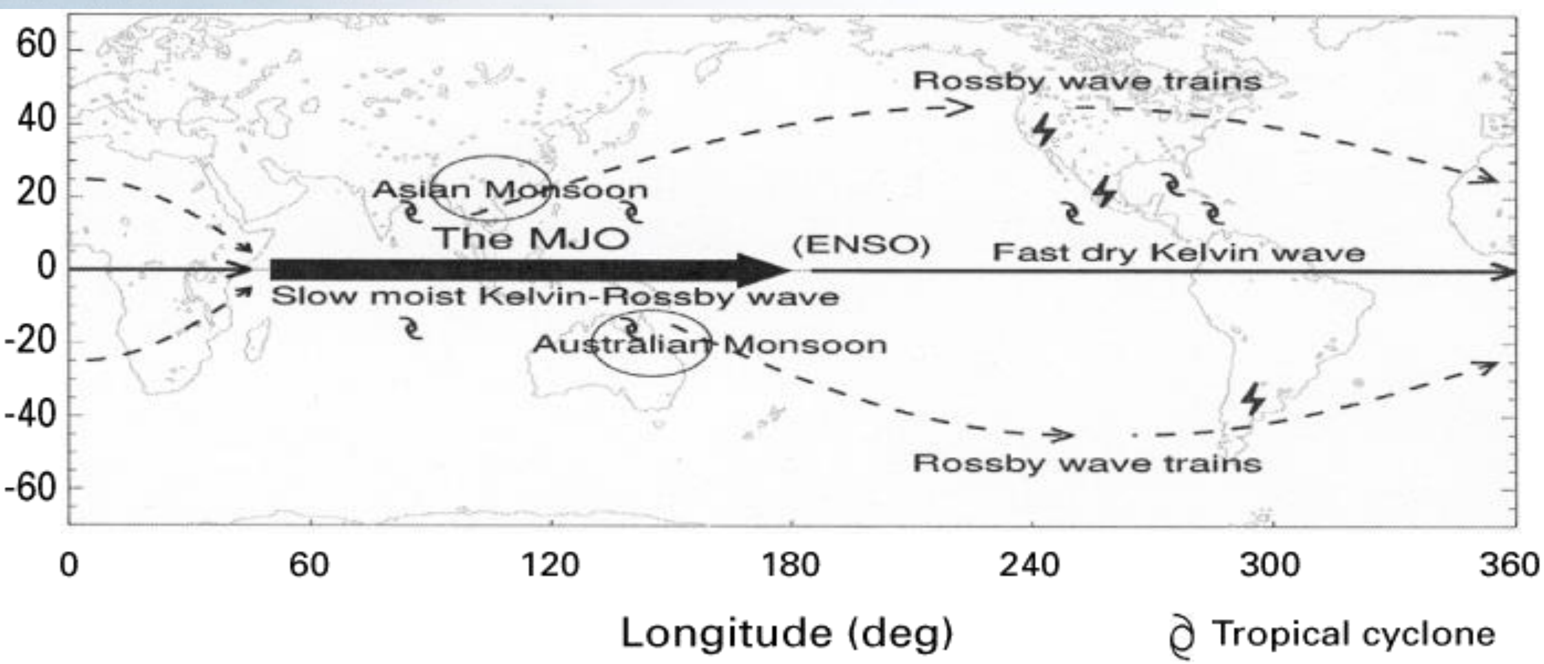
- Madden Julian Oscillation
- ENSO
- Land surface conditions: snow-soil moisture
- Stratospheric variability
- Atmospheric dynamical processes
  (Rossby wave propagations, weather regimes…)
- Sea ice cover –thickness ?

Skill depends on “windows of opportunity”!

Madden-Julian Oscillation and its impacts



The Madden-Julian Oscillation (**MJO**) is the major fluctuation in tropical weather on weekly to monthly timescales. The MJO can be characterised as an eastward moving 'pulse' of cloud and rainfall near the equator that typically recurs every 30 to 60 days.



WWRP/WCRP Sub-seasonal to Seasonal (S2S) Prediction Project

Sub-Projects

- **Teleconnections** *(C. Stan and H. Lin)*
- **Madden-Julian Oscillation** *(D. Waliser and S. Woolnough)*
- **Monsoons** *(H. Hendon)*
- **Africa** *(A. Robertson and R. Graham)*
- **Extremes** *(F. Vitart)*
- **Verification and Products** *(C. Coelho)*

Research Issues

- Predictability
- Teleconnection
- O-A Coupling
- Scale interactions
- Physical processes

Modelling Issues

- Initialisation
- Ensemble generation
- Resolution
- O-A Coupling
- Systematic errors
- Multi-model combination

Needs & Applications

Liaison with SERA (Working Group on Societal and Economic Research Applications)

S2S Database

WWRP/WCRP S2S Database

	Time-range	Resol.	Ens. Size	Freq.	Hcsts	Hcst length	Hcst Freq	Hcst Size
ECMWF	D 0-46	T639/319L91	51	2/week	On the fly	Past 20y	2/weekly	11
UKMO	D 0-60	N216L85	4	daily	On the fly	1993-2015	4/month	3
NCEP	D 0-44	N126L64	4	4/daily	Fix	1999-2010	4/daily	1
ECCC	D 0-32	0.45x0.45 L40	21	weekly	On the fly	1995-2014	weekly	4
BoM	D 0-60	T47L17	33	2/weekly	Fix	1981-2013	6/month	33
JMA	D 0-34	T319L60	25	2/weekly	Fix	1981-2010	3/month	5
KMA	D 0-60	N216L85	4	daily	On the fly	1996-2009	4/month	3
CMA	D 0-45	T106L40	4	daily	Fix	1886-2014	daily	4
CNRM	D 0-32	T255L91	51	weekly	Fix	1993-2014	2/monthly	15
CNR-ISAC	D 0-32	0.75x0.56 L54	40	weekly	Fix	1981-2010	6/month	1
HMCR	D 0-63	1.1x1.4 L28	20	weekly	Fix	1981-2010	weekly	10

s2s.ecmwf.int s2s.cma.cn

Sub-seasonal verification

S2S forecasts are based on ensemble forecasts. Metrics used to verify S2S forecasts include:

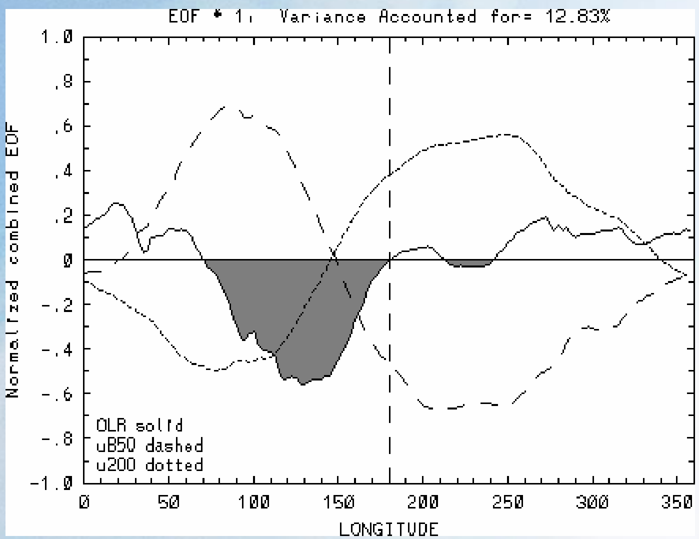
- RMSE/correlations (MJO/ENSO…)
- Reliability diagrams/BS
- RPS
- CRPS
- ROC area
- Potential Economic value

….

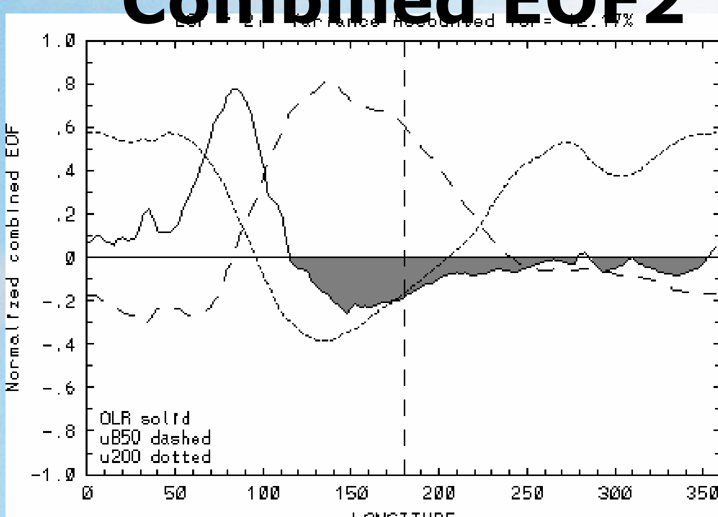
Usually applied on weekly means/monthly means

Wheeler and Hendon MJO Index

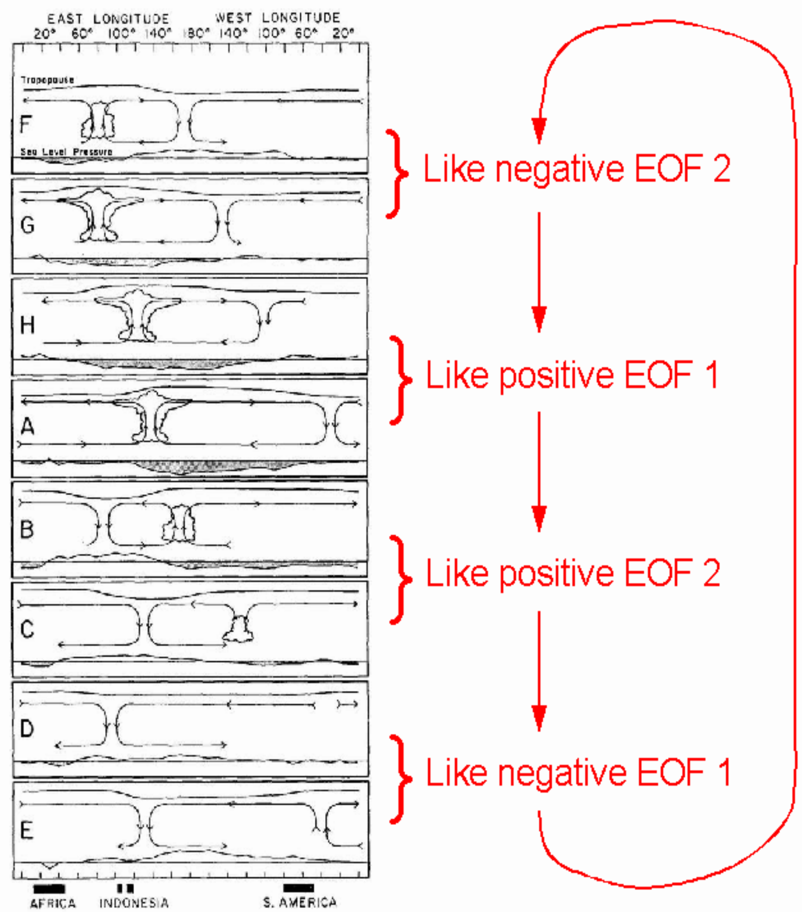
Combined EOF1



Combined EOF2

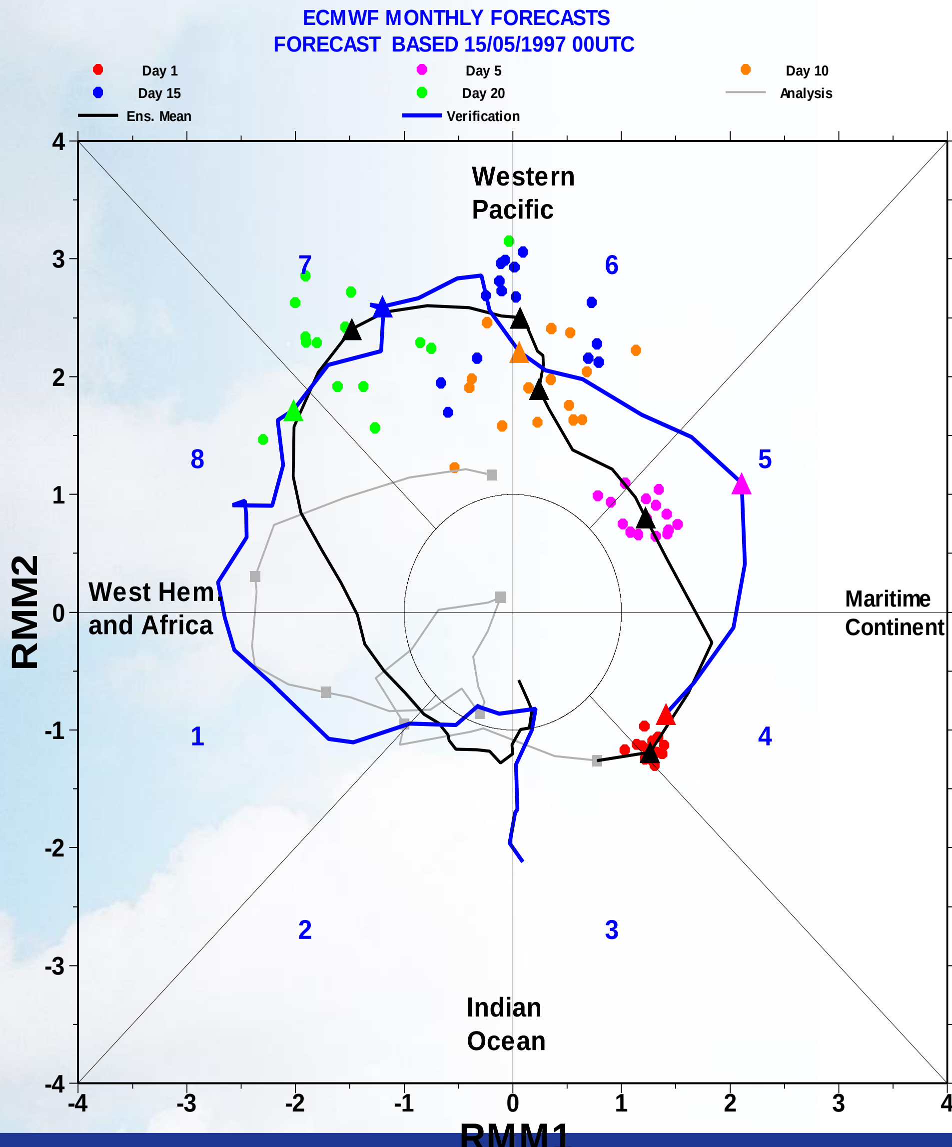


Madden and Julian's (1972) schematic

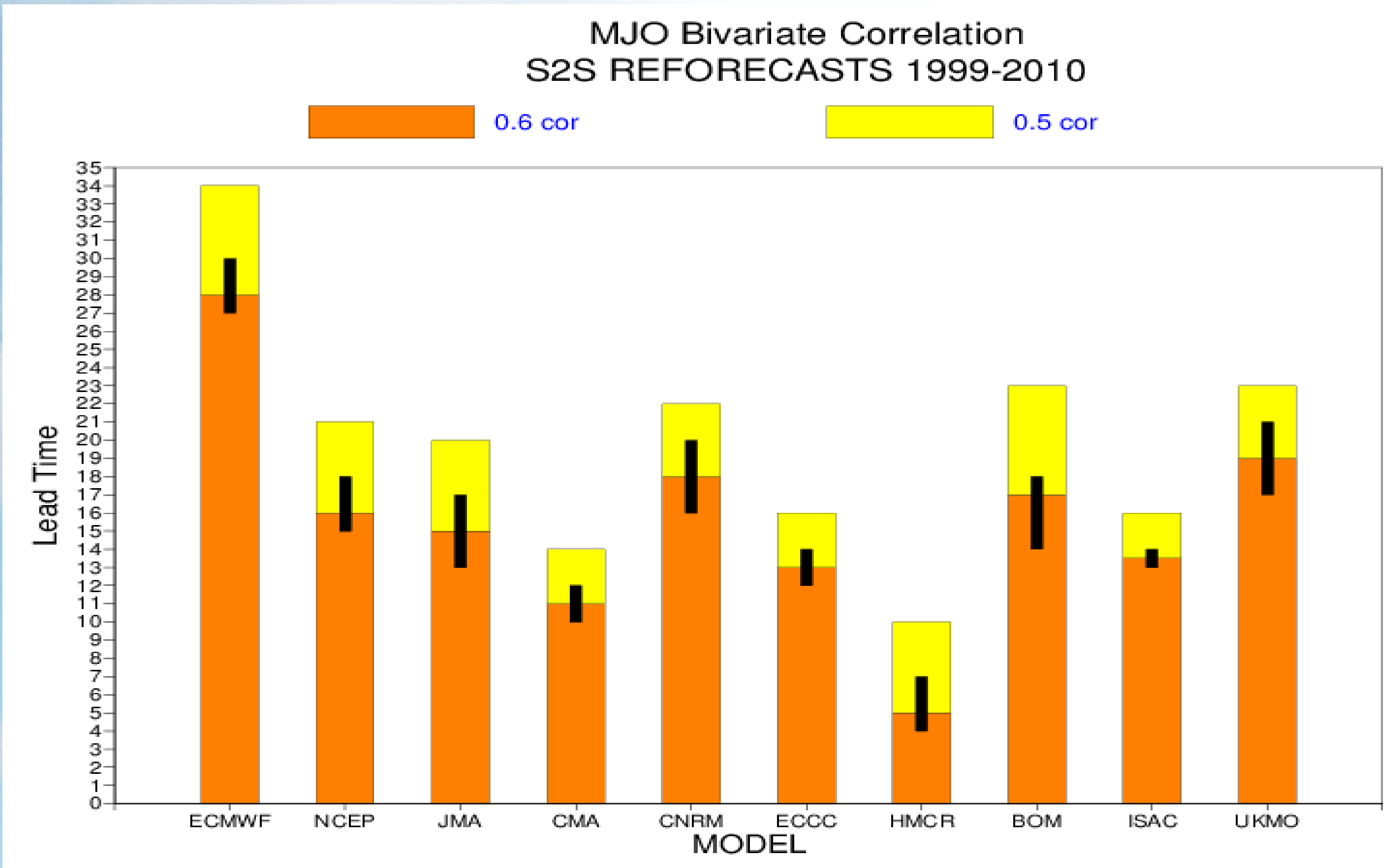


From Wheeler and Hendon, BMRC

MJO FORECAST



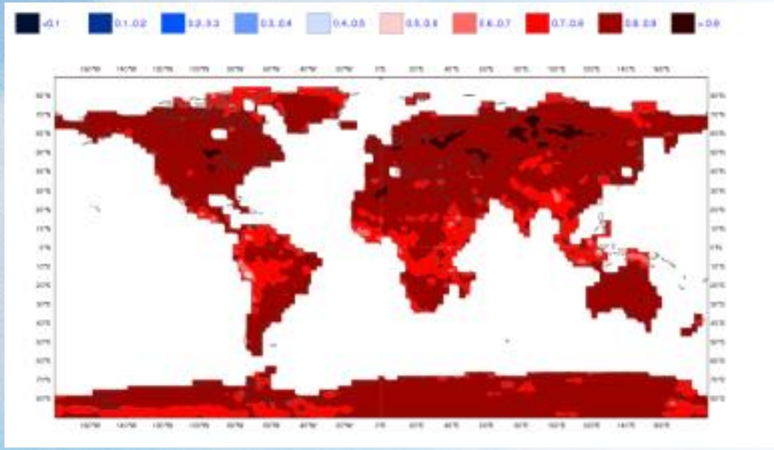
Bivariate Correlation with ERA Interim – Ensemble Mean 1999-2010 re-forecasts



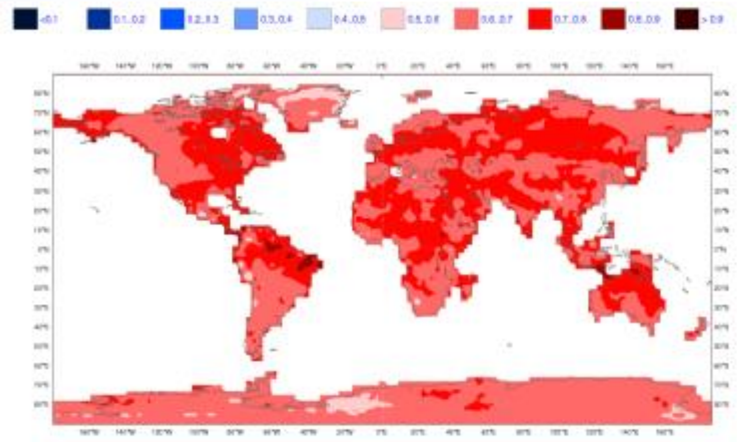
Skill of the ECMWF Extended-range forecasts

ROC area: 2-meter temperature in the upper tercile

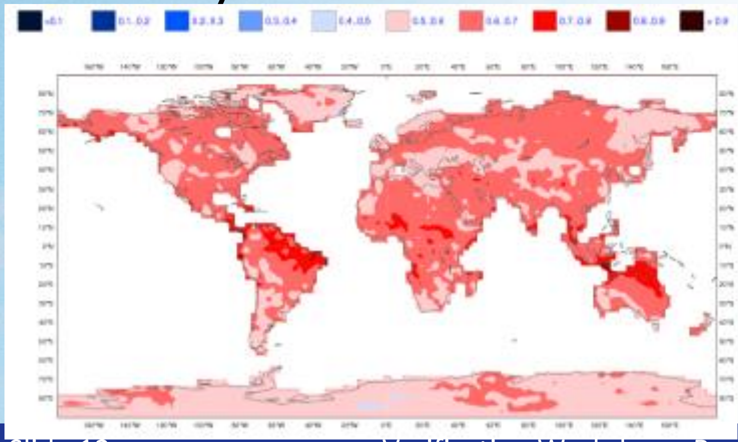
Day 5-11



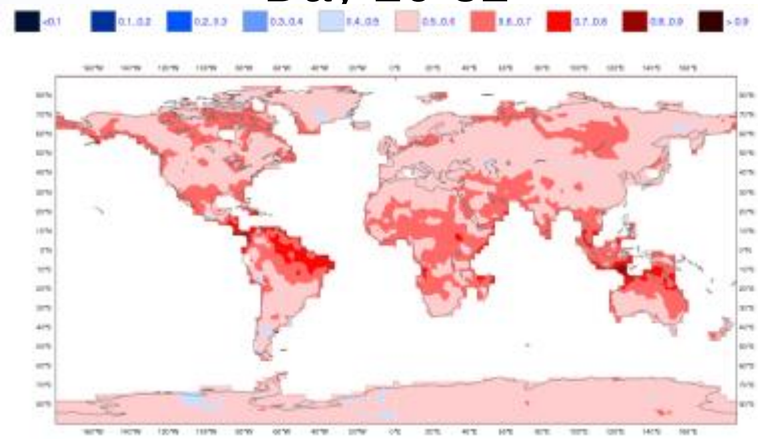
Day 12-18



Day 19-25



Day 26-32



S2S verification

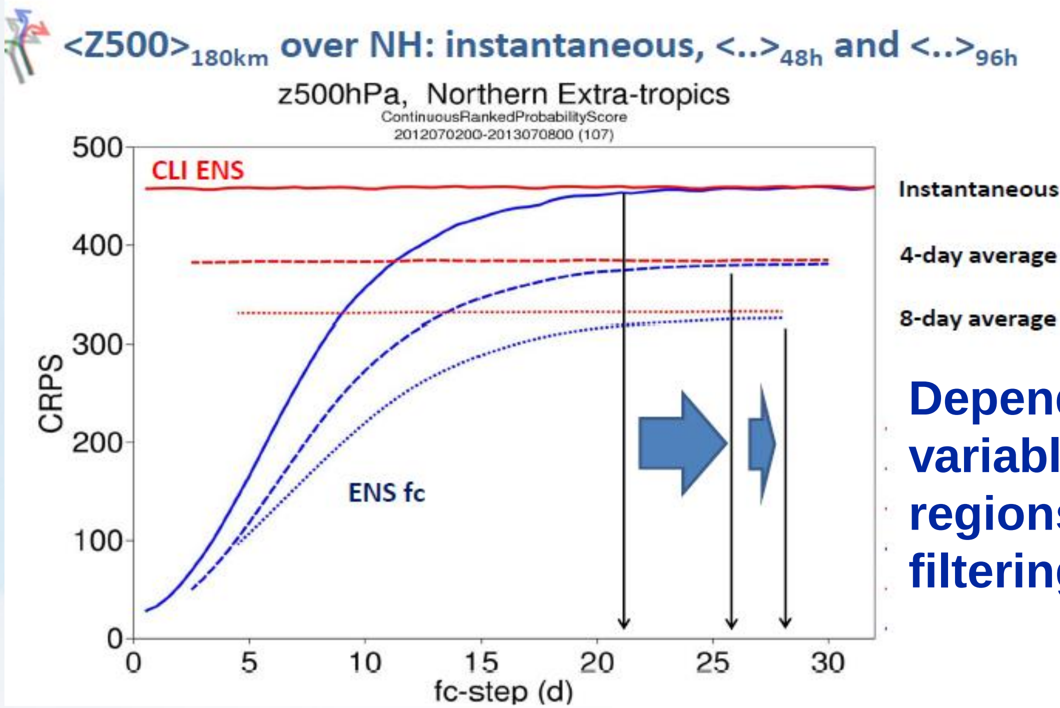
Important challenges with S2S verification:

- Extended-range forecasts have very little skill to predict the day to day variability of the weather. There is a need to verify S2S forecasts over longer time period and larger domains. What is the optimum space/time filtering?

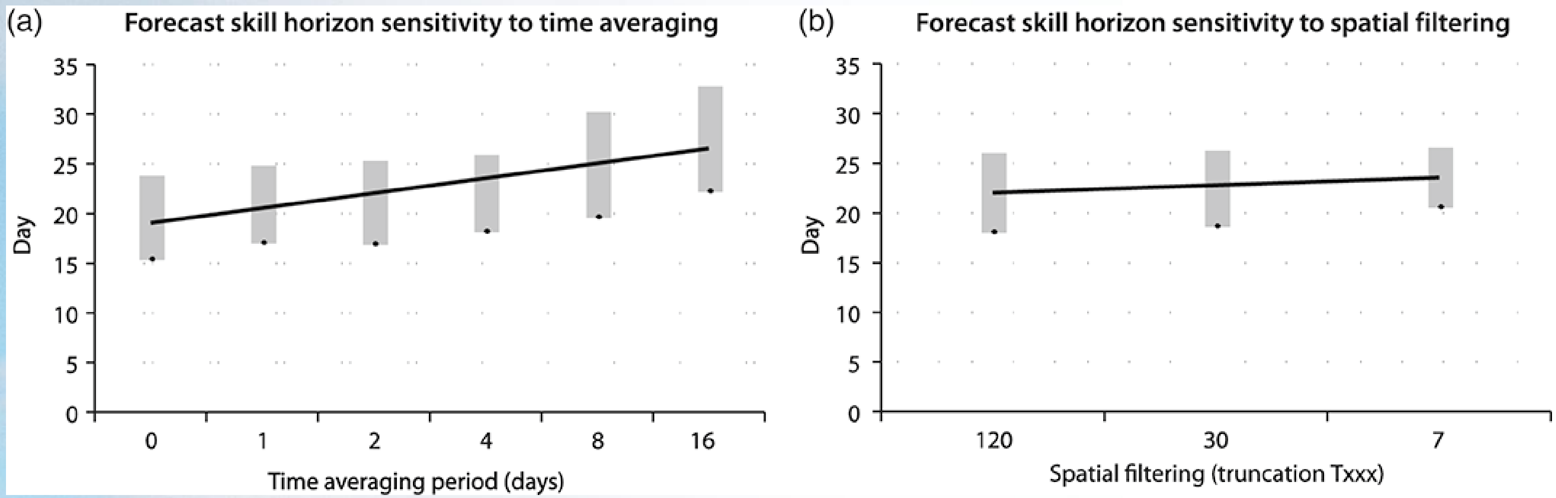
- Forecast skill is very flow dependent. Need for conditional verification on MJO, ENSO, NAO, IOD and SAM phases as well as on particular weather regimes

- Models drift quickly towards there own climatology. Calibration is necessary. Operational centres produce re-forecasts to calibrate real-time S2S forecasts and also for skill assessment.

The predictability limit is the time when the forecast error crosses a certain threshold. As threshold, m - 2 σ was used, where m is the average climatological error.

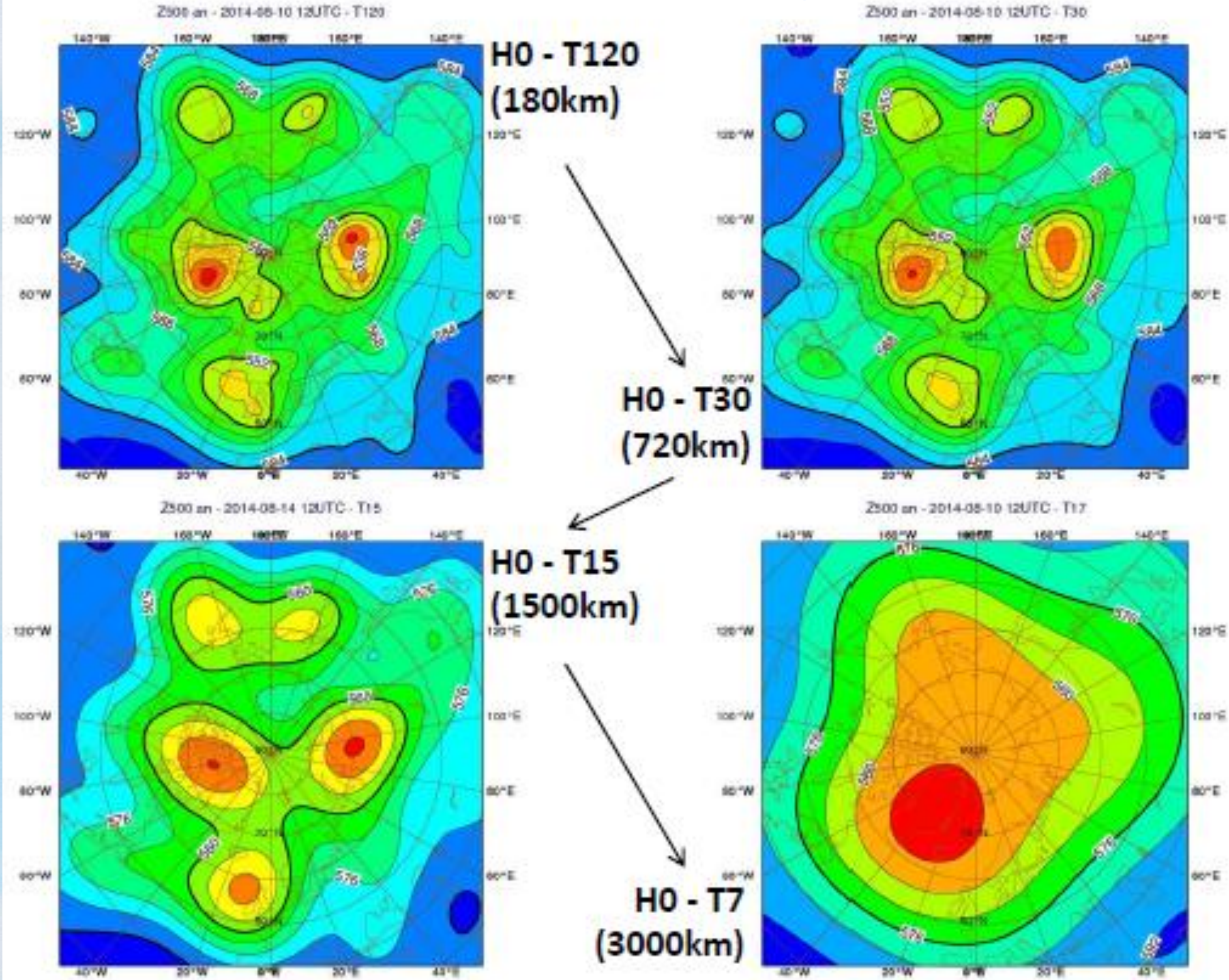


Depends on variables, regions, spatial filtering

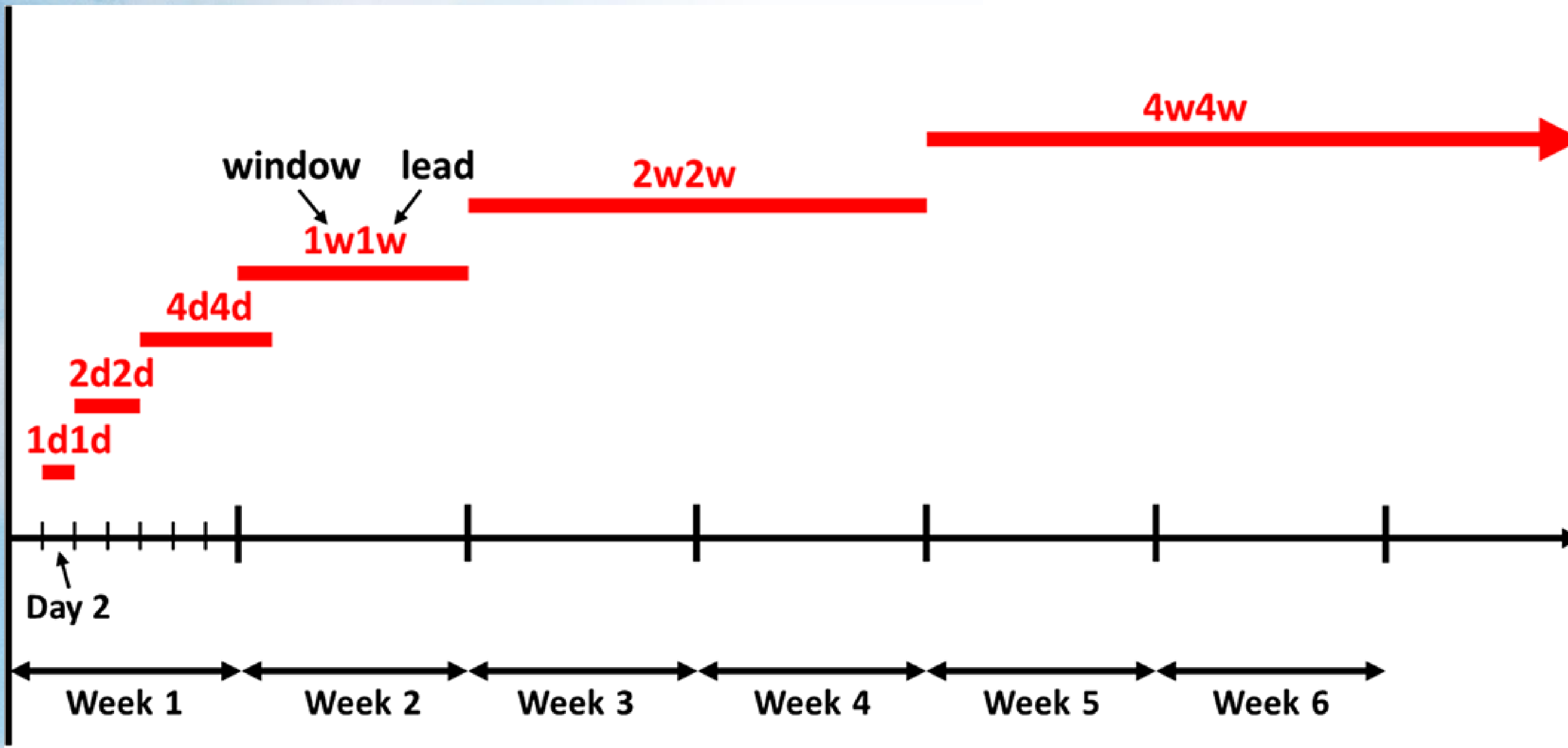


(Z500, T850, U950, V850) and three regions (NH, SH, TR). *Buizza and Leutbecher, 2015*

Spatial Filtering



Seamless prediction and verification



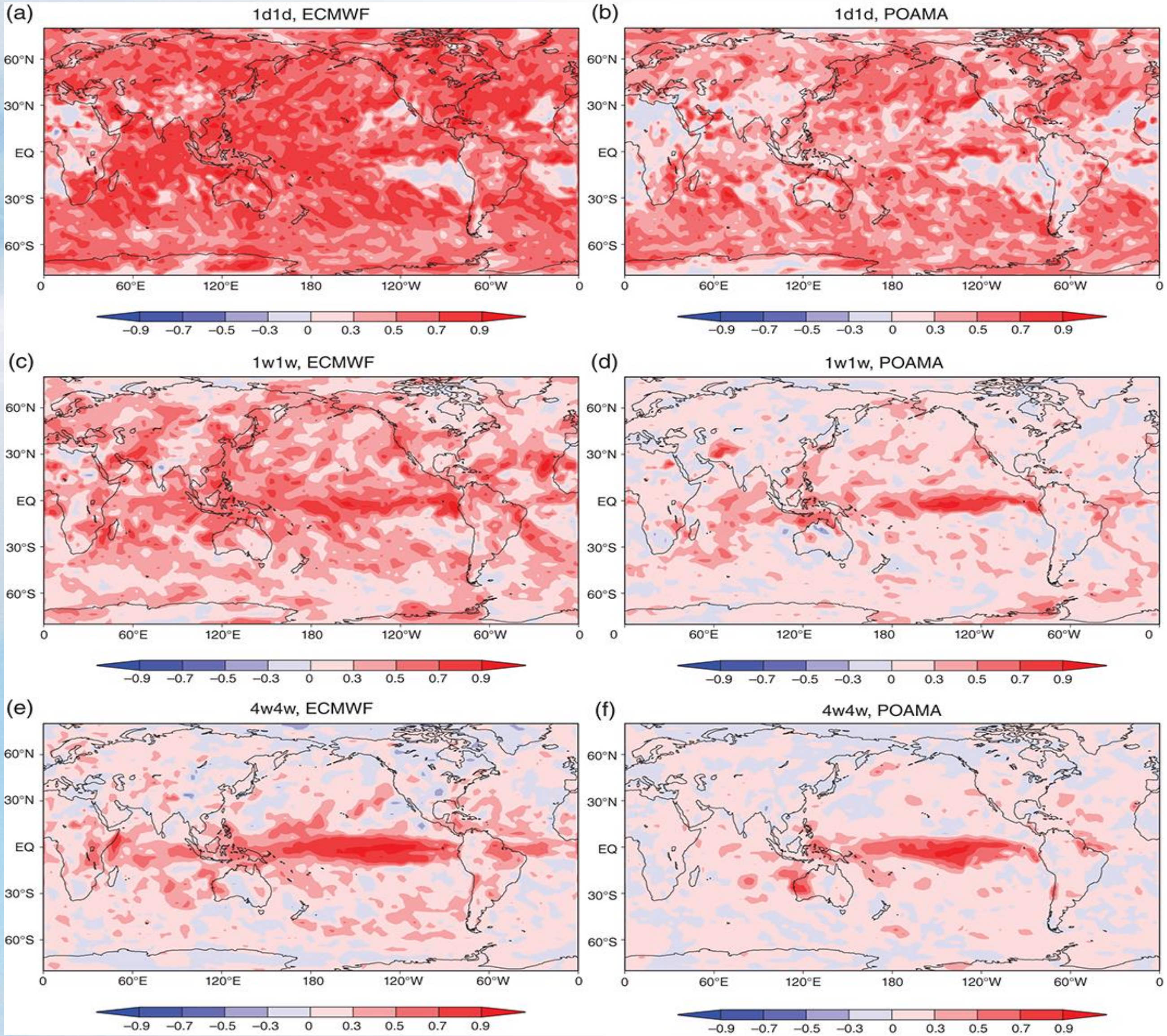
Wheeler et al, 2016

Example of seamless Verification

1d1d

1w1w

4w4w



Short range

Medium range

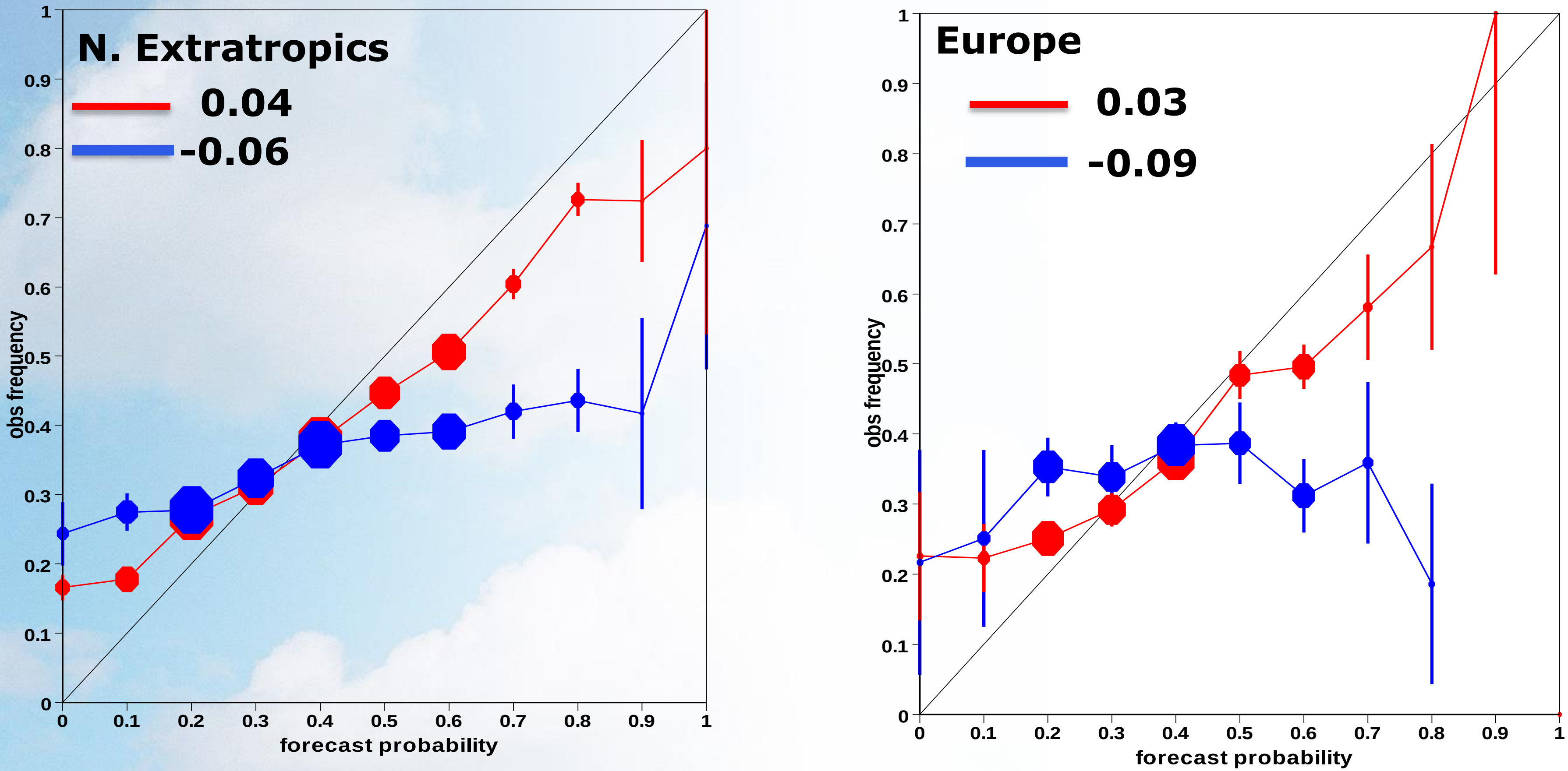
Extended range

Maps of CORa actual skill for precipitation

Wheeler et al, 2016

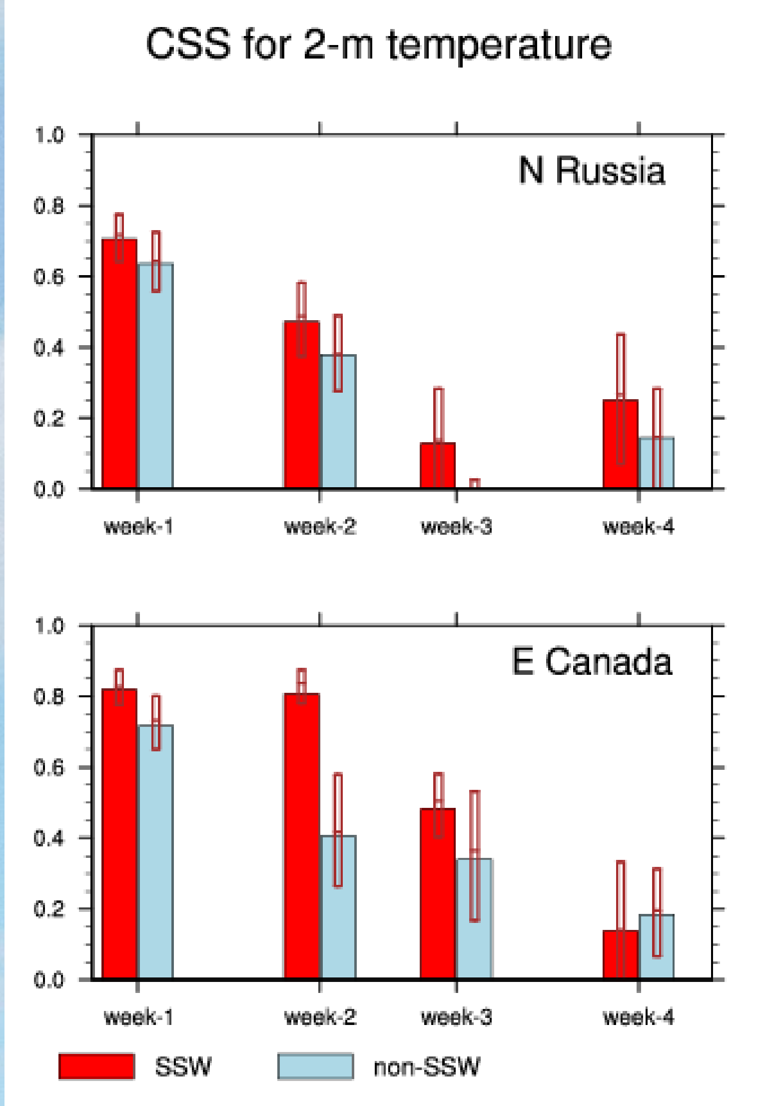
Impact of MJO on S2S skill scores

Reliability Diagram
Probability of 2-m temperature in the upper tercile
Day 19-25



MJO in IC **NO MJO in IC** *Vitart and Molteni, 2010*

Impact of SSWs on forecast skill scores



From Tripathi et al. (2015)

Conditional verification is useful:

- Better understanding the contribution of climate drivers in the model
- For users to have more/less confidence in a forecast a priori.

This type of verification needs adequate samples (including re-forecasts) to allow sub-setting of the data to provide meaningful verification.

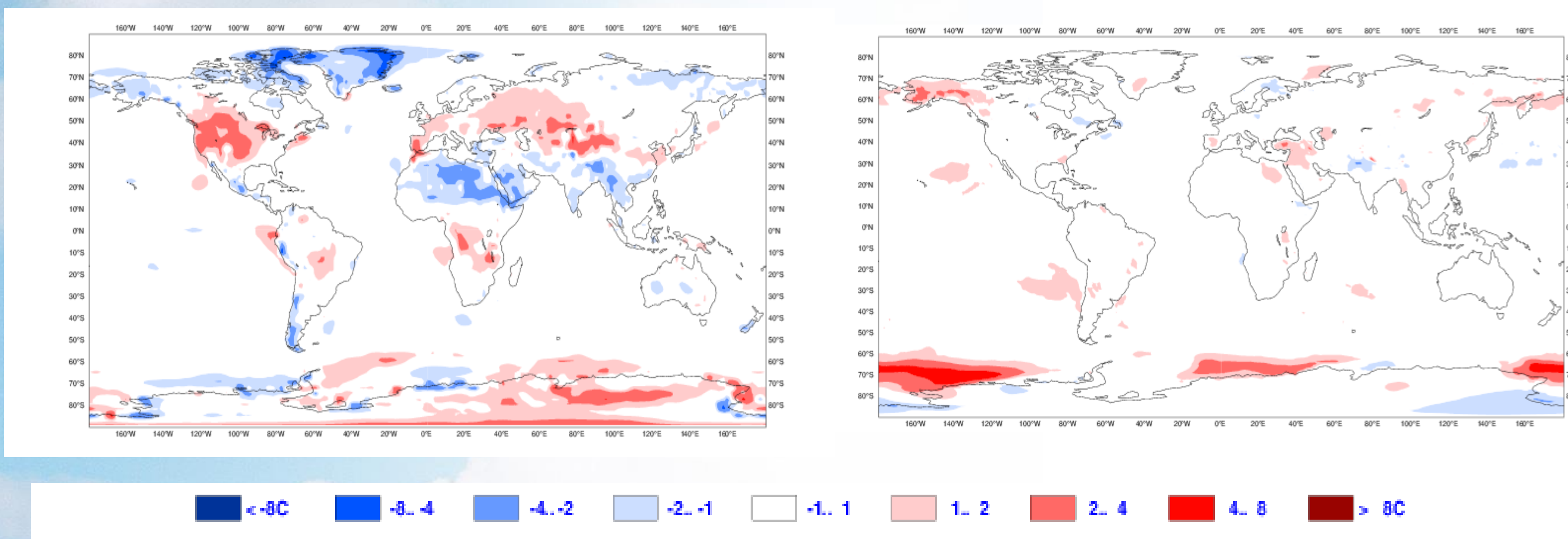
Need to calibrate extended-range forecasts

2m-temp forecast day 26-32

1st August start dates

Model Bias (1996-2015)

Forecast anomalies

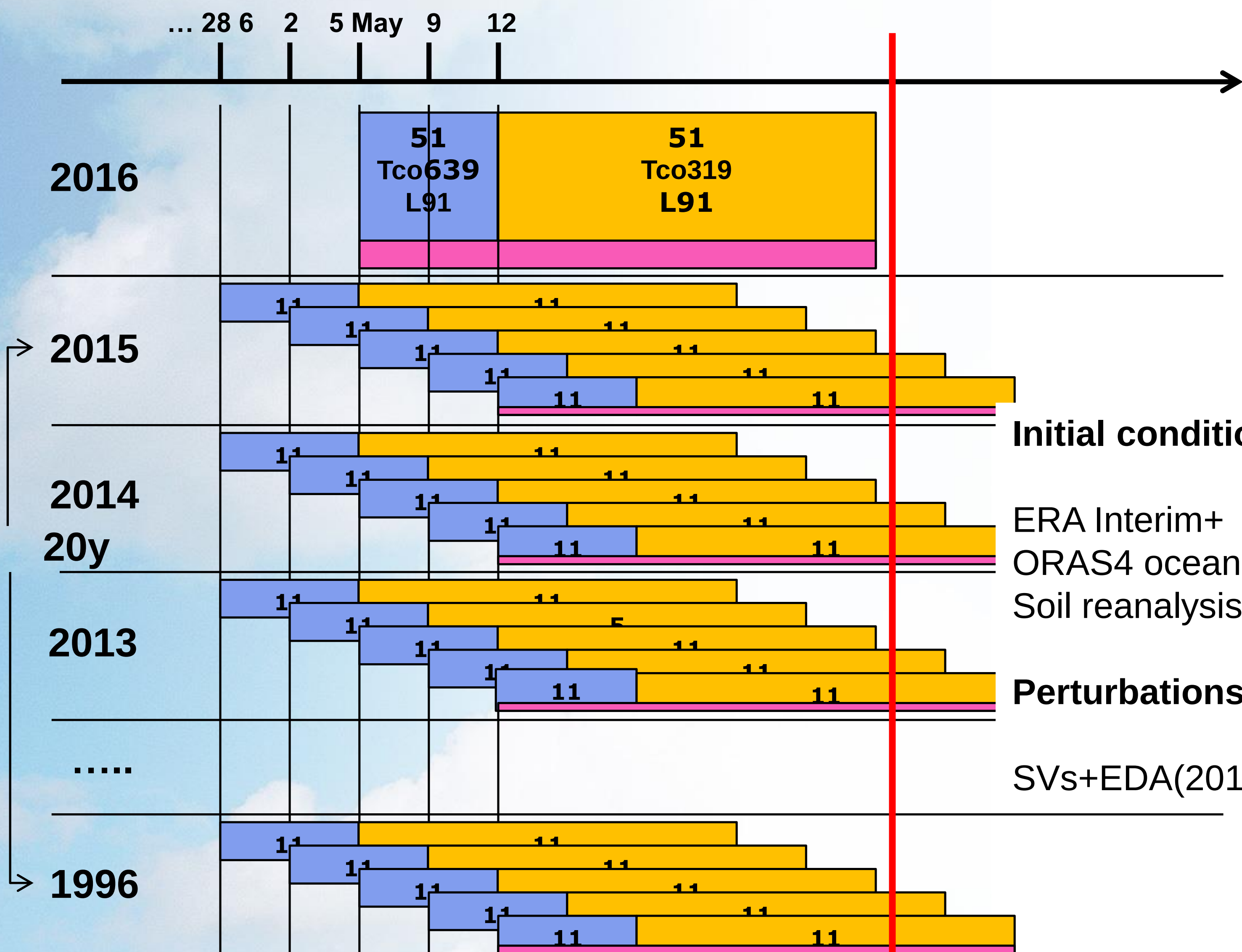


Biases (eg 2mT as shown here) can have a magnitude larger than the anomalies we want to predict

WWRP/WCRP S2S Database

	Time-range	Resol.	Ens. Size	Freq.	Hcsts	Hcst length	Hcst Freq	Hcst Size
ECMWF	D 0-46	T639/319L91	51	2/week	On the fly	Past 20y	2/weekly	11
UKMO	D 0-60	N216L85	4	daily	On the fly	1993-2015	4/month	3
NCEP	D 0-44	N126L64	4	4/daily	Fix	1999-2010	4/daily	1
ECCC	D 0-32	0.45x0.45 L40	21	weekly	On the fly	1995-2014	weekly	4
BoM	D 0-60	T47L17	33	2/weekly	Fix	1981-2013	6/month	33
JMA	D 0-34	T319L60	25	2/weekly	Fix	1981-2010	3/month	5
KMA	D 0-60	N216L85	4	daily	On the fly	1996-2009	4/month	3
CMA	D 0-45	T106L40	4	daily	Fix	1986-2014	daily	4
CNRM	D 0-32	T255L91	51	weekly	Fix	1993-2014	2/monthly	15
CNR-ISAC	D 0-32	0.75x0.56 L54	40	weekly	Fix	1981-2010	6/month	1
HMCR	D 0-63	1.1x1.4 L28	20	weekly	Fix	1981-2010	weekly	10

The ECMWF ENS re-forecast suite to estimate the M-climate



Initial conditions:

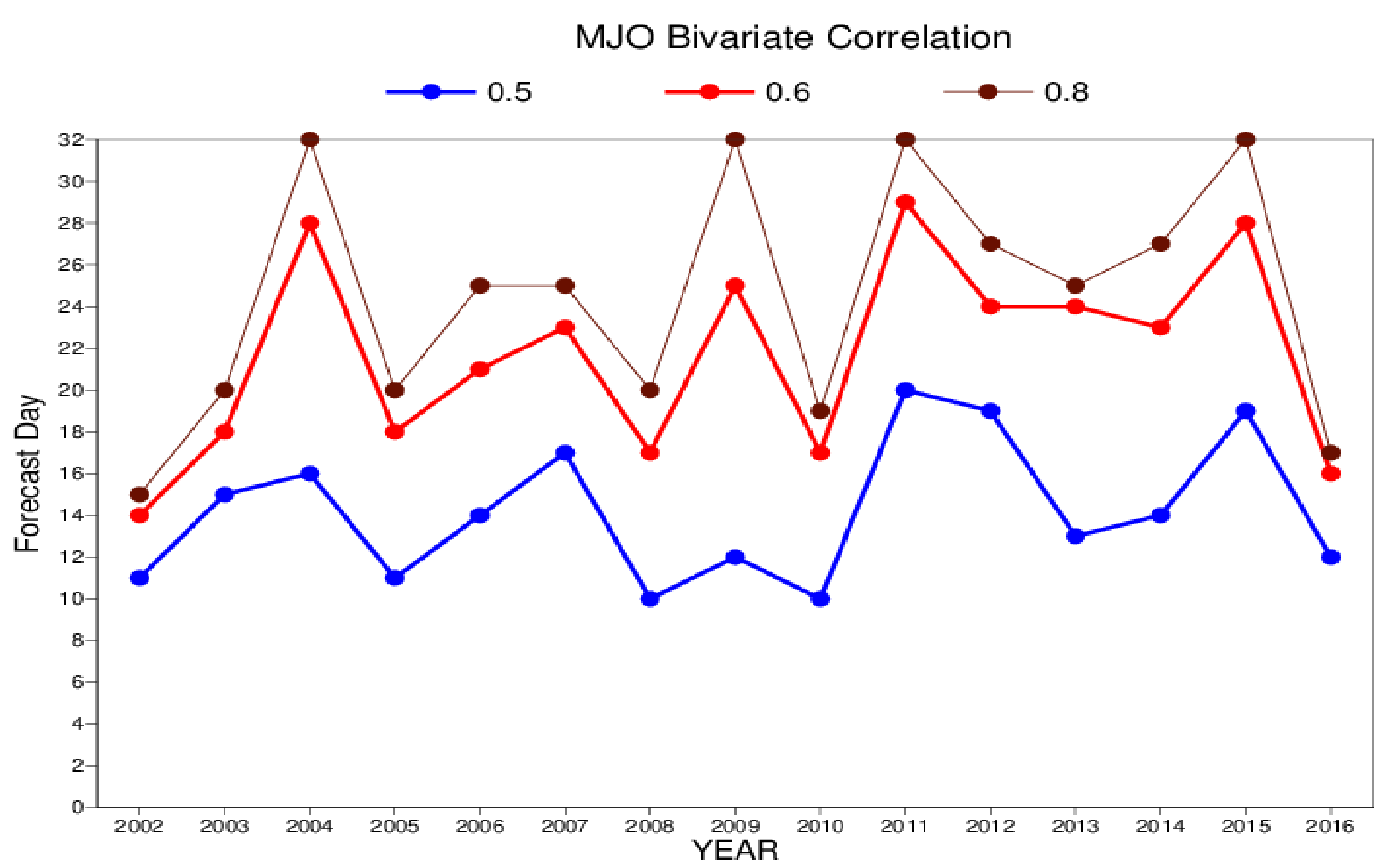
ERA Interim+
ORAS4 ocean Ics+
Soil reanalysis

Perturbations:

SVs+EDA(2016)+SPPT+SKEB

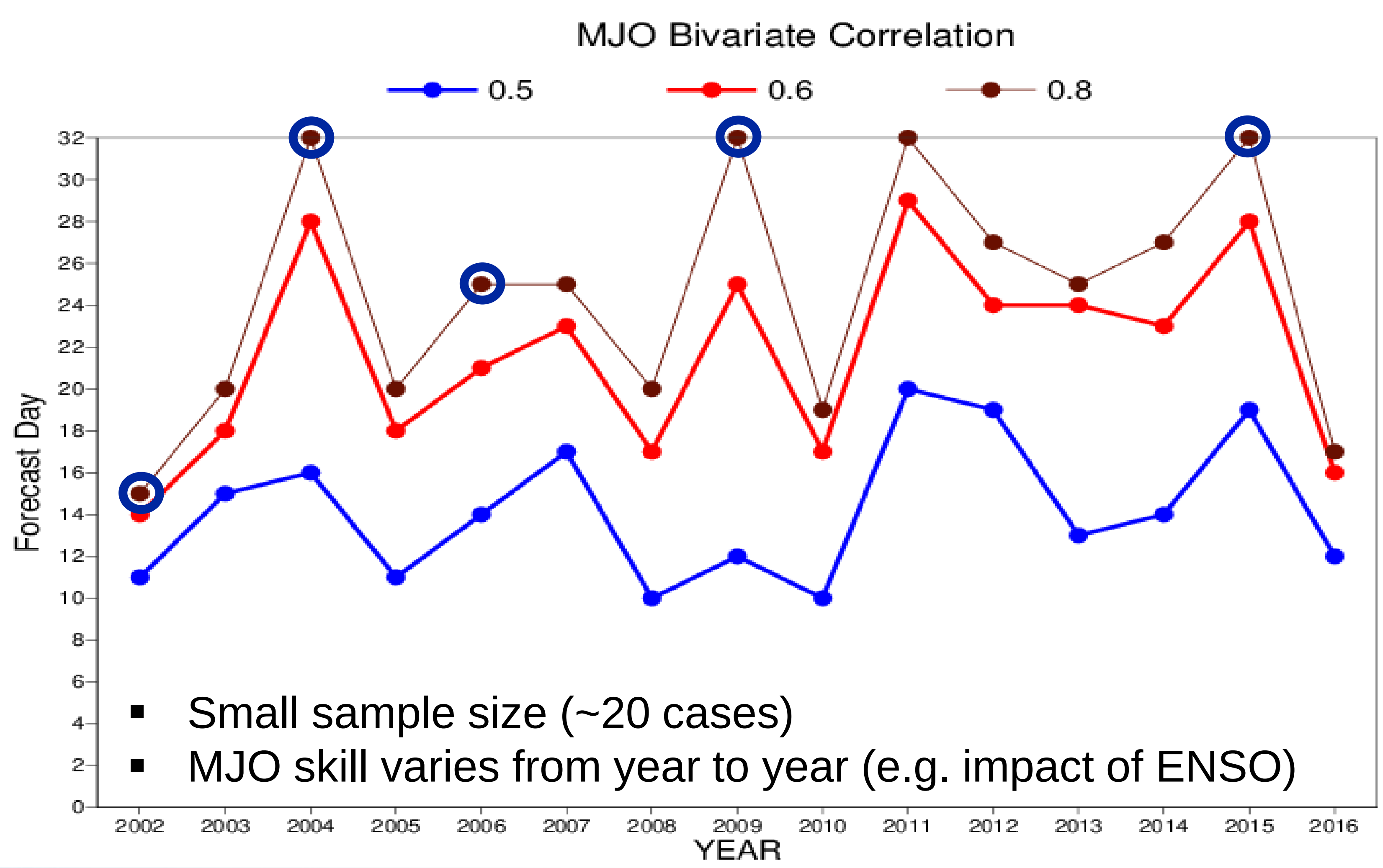
MJO WH index bivariate correlation

ECMWF Real-time forecasts - NDJFM 2002-2016



MJO WH index bivariate correlation

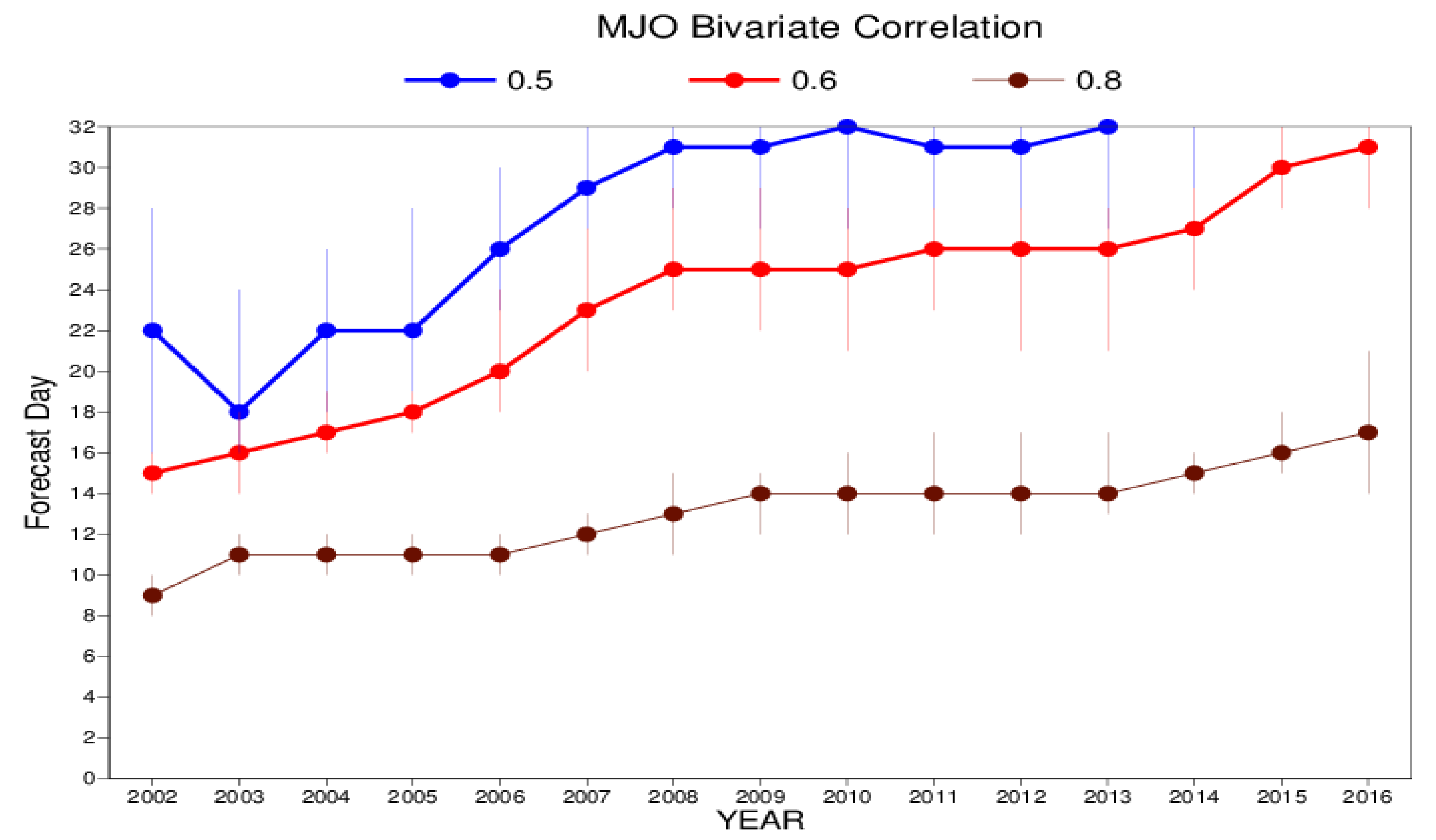
ECMWF Real-time forecasts - NDJFM 2002-2016



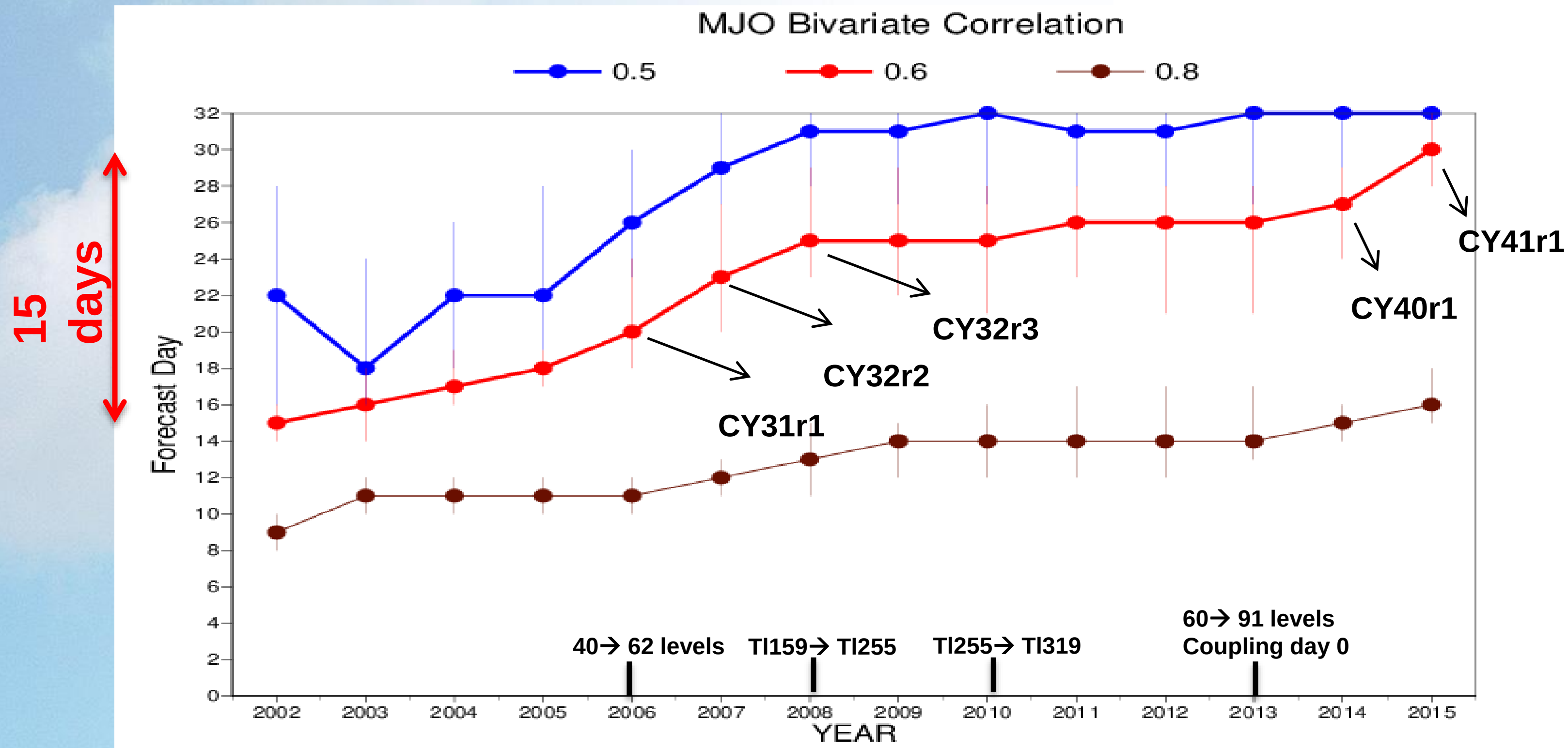
- Small sample size (~20 cases)
- MJO skill varies from year to year (e.g. impact of ENSO)

MJO WH index bivariate correlation

Re-forecasts - common period 1995-2001



MJO WH index bivariate correlation

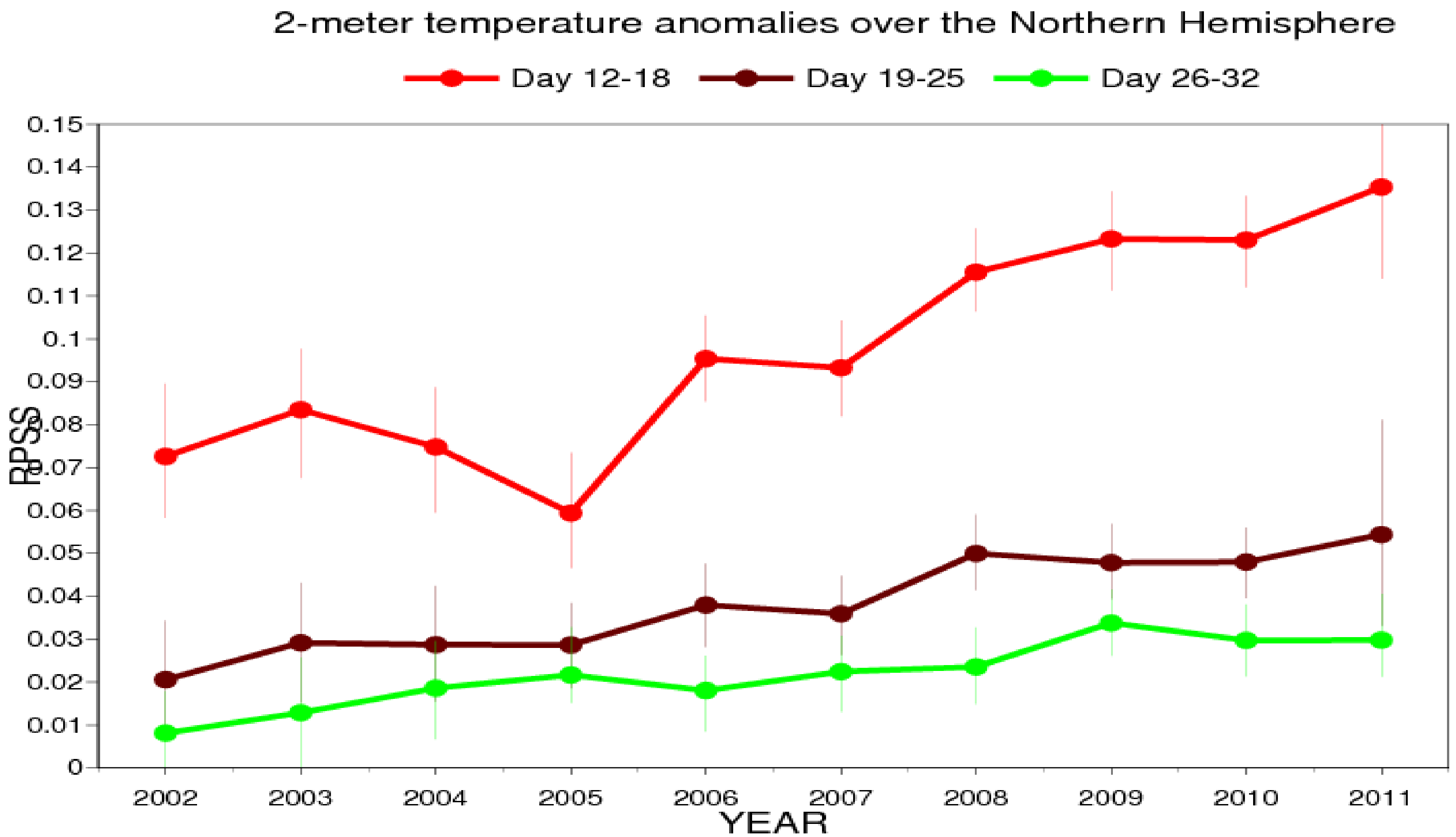


CY31R1: Parameterisation of ice supersaturation
CY32R2: McRAD (radiation scheme)
CY32R3: Changes in convective scheme (Bechtold at al. 2008)
CY40R1: Improved diurnal cycle of precipitation
CY41R1: revised organized convective detrainment and the revised convective momentum transport. …

Improvements in MJO Prediction mostly due to changes in convective parameterization

Performance of the extended-range Forecasts

2-metre temperature RPSS over Northern Extratropics



Vitart, 2014

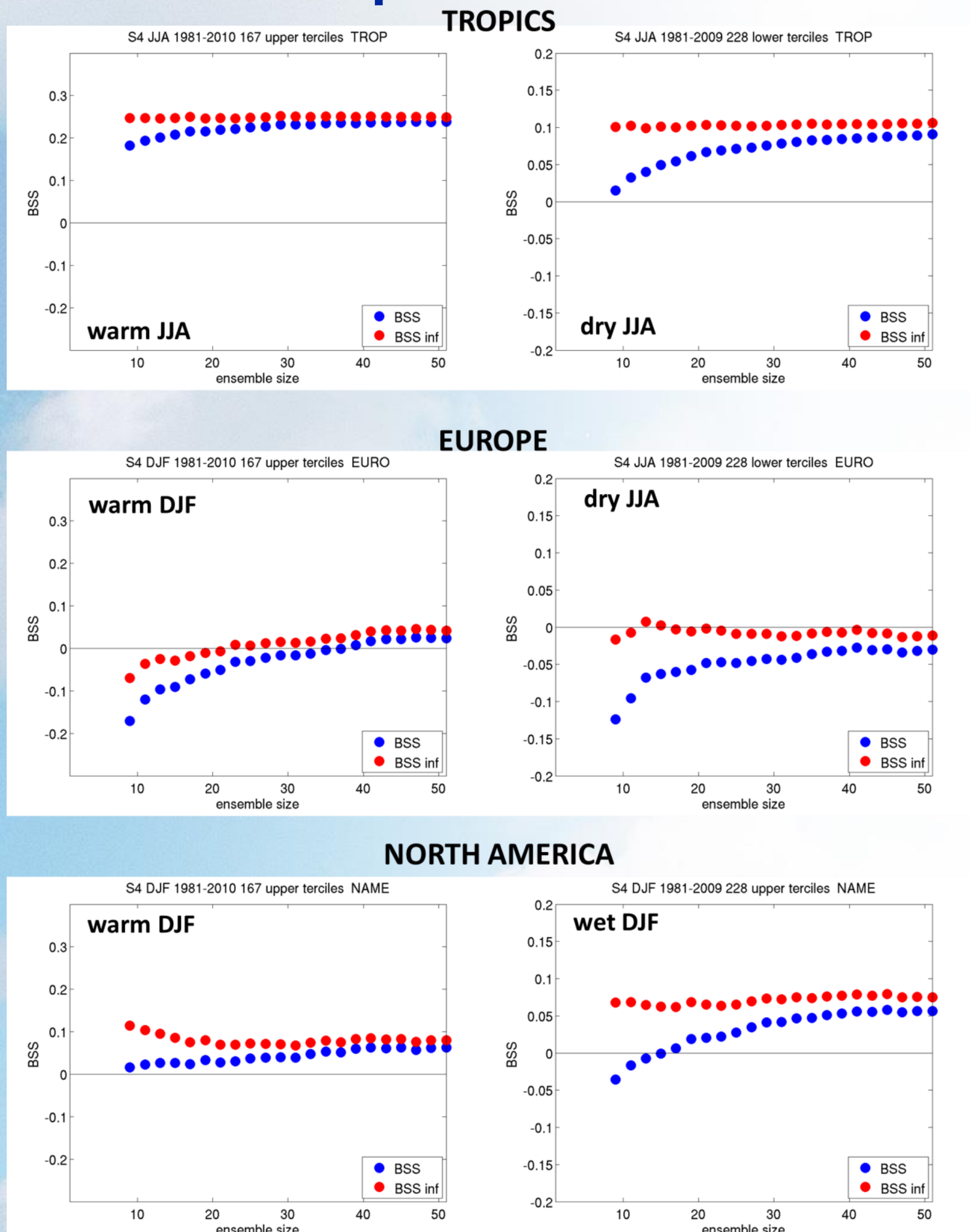
Issues with re-forecast verification

- Re-forecast initialization (re-analyses) is different from the initialization of real-time forecasts (operational analyses)

- Re-forecast ensemble size is often small (typically 5 members) compared to real-time forecasts. Skill is likely to be underestimated.

- Number of re-forecast years is generally too small for properly sampling events like ENSO.

Impact of re-forecast ensemble Size



Brier Skill Score (BSS) with climatology as a reference.

Two versions of the BSS are used:

- **the uncorrected BSS as estimated from the ensemble directly.**

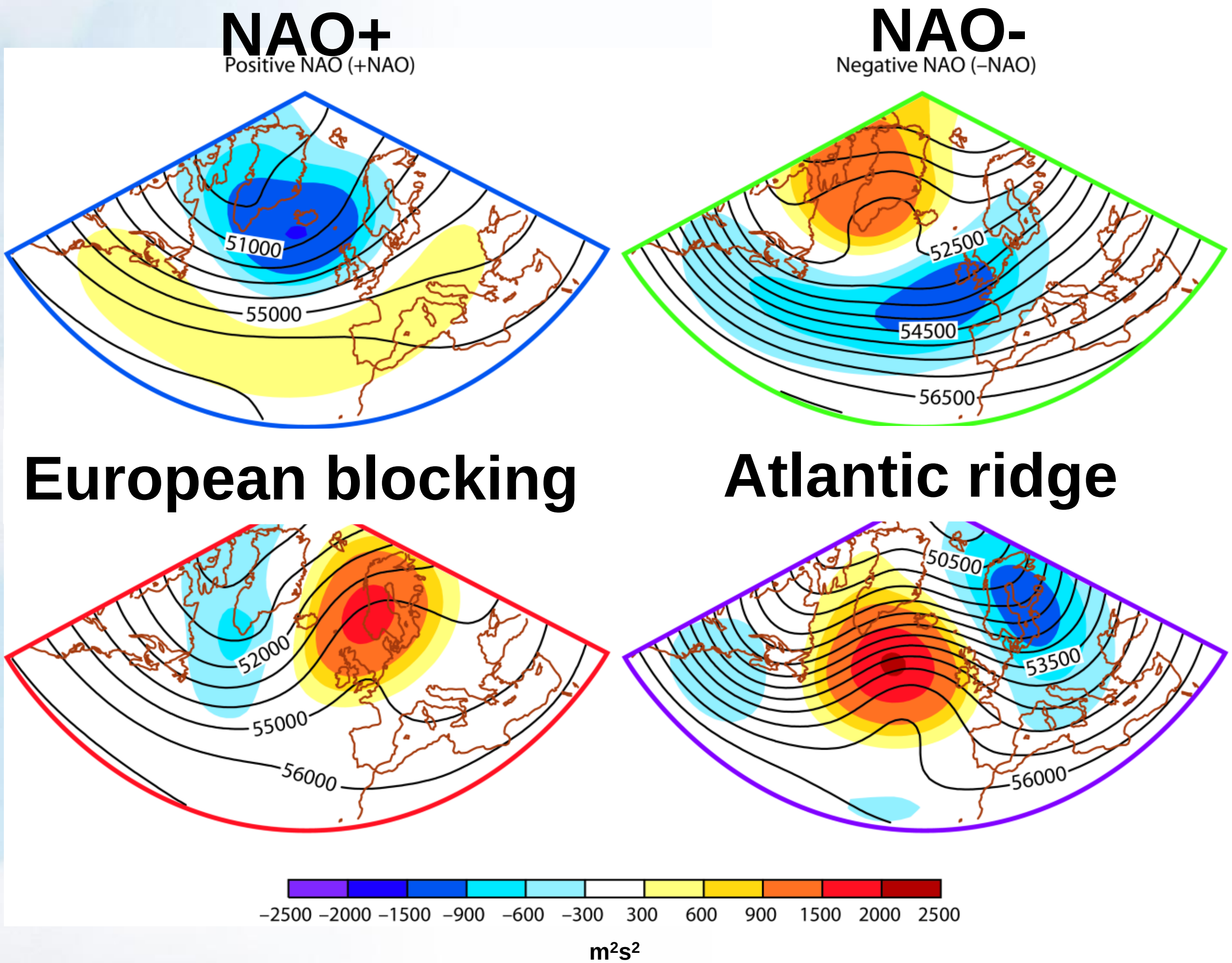
- **an analytical correction of the BSS extrapolating towards a hypothetical infinite ensemble size (Ferro, 2007) .**

Ferranti, Corti, Weisheimer

Verification of Weather regime transition (L. Ferranti)

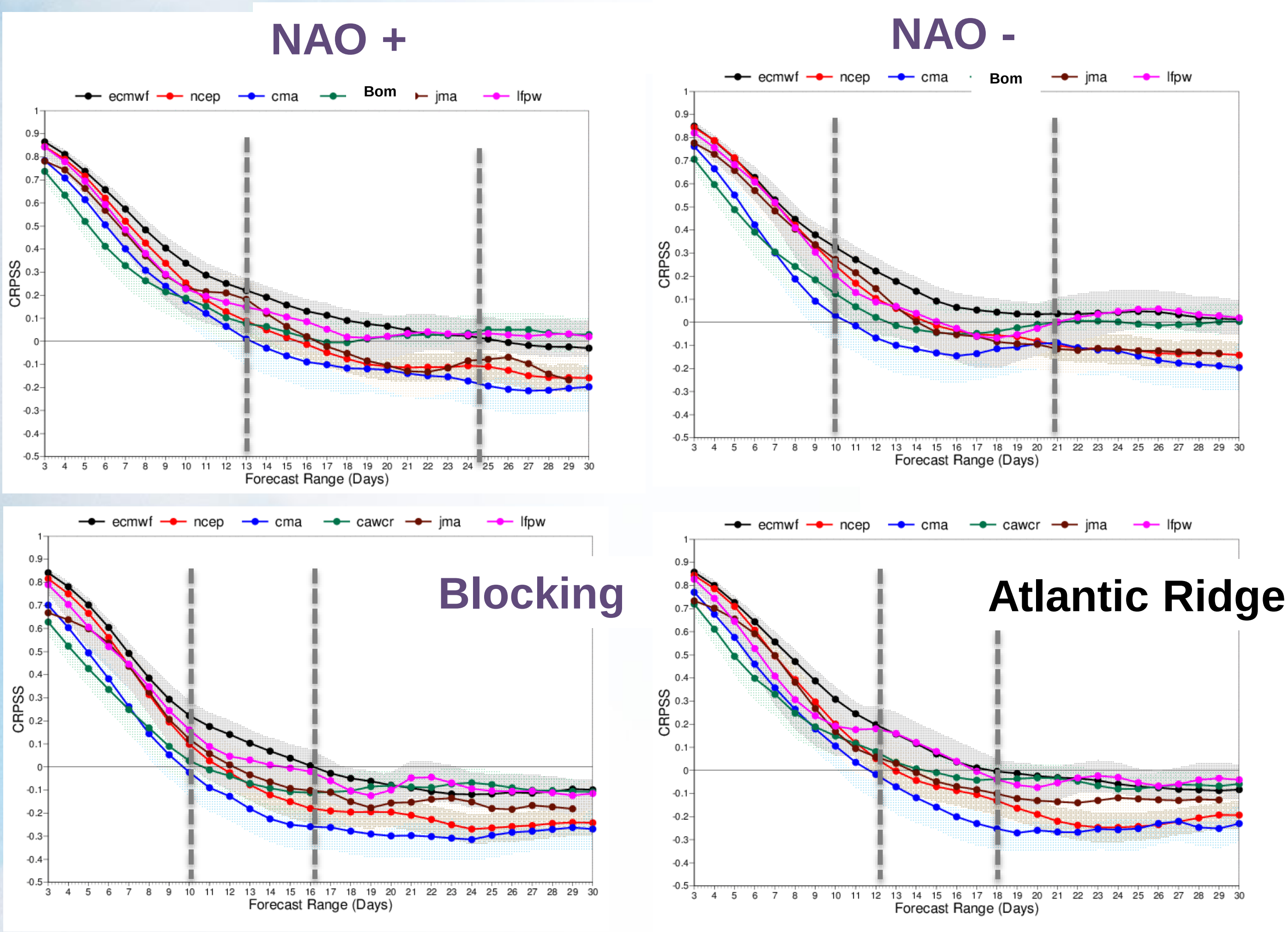
Regimes based on clustering of daily anomalies for 29 cold seasons (1980-2008)

500 hPa geopotential



'k means' clustering applied to EOF pre-filtered data (retaining 80% of variance)

Predicting skill (CRPSS) associated with the Euro-Atlantic Regimes:

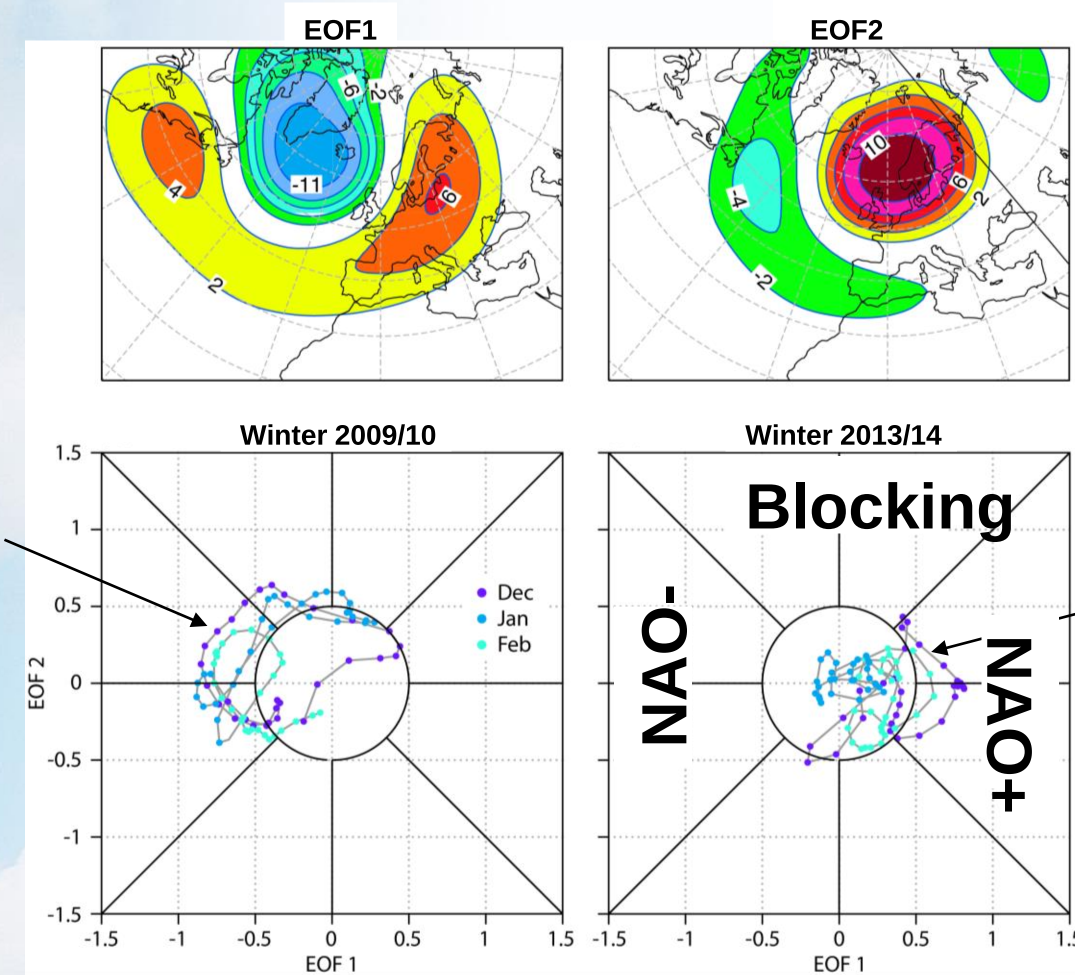


How can we evaluate the model ability in predicting regimes transitions?

Trajectories in phase space (*c.f.* MJO propagation)

- ±EOF1 and +EOF2 represent quite well ±NAO and BL
- Trajectories in phase space summarise regime evolution
- Unlike MJO, no preferred direction

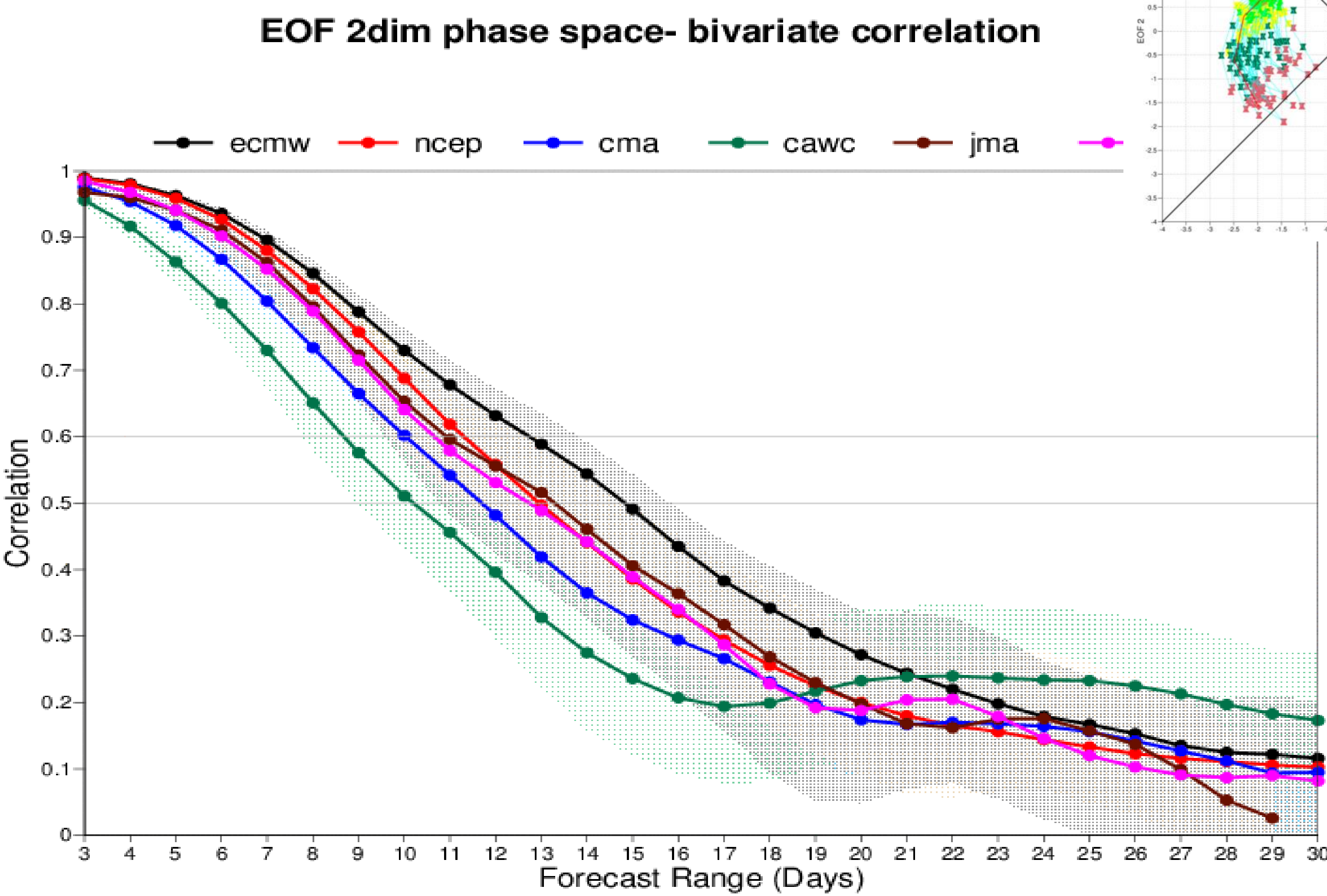
BL: record-breaking cold temperatures over Europe

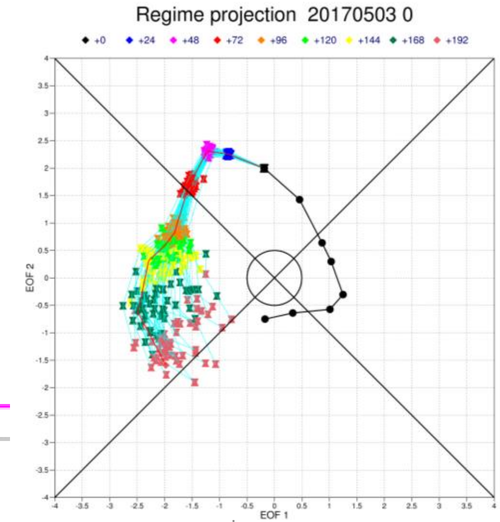


+NAO: exceptional storminess, but mild temperatures over Europe

Based on 5-day running means

Regime transitions:





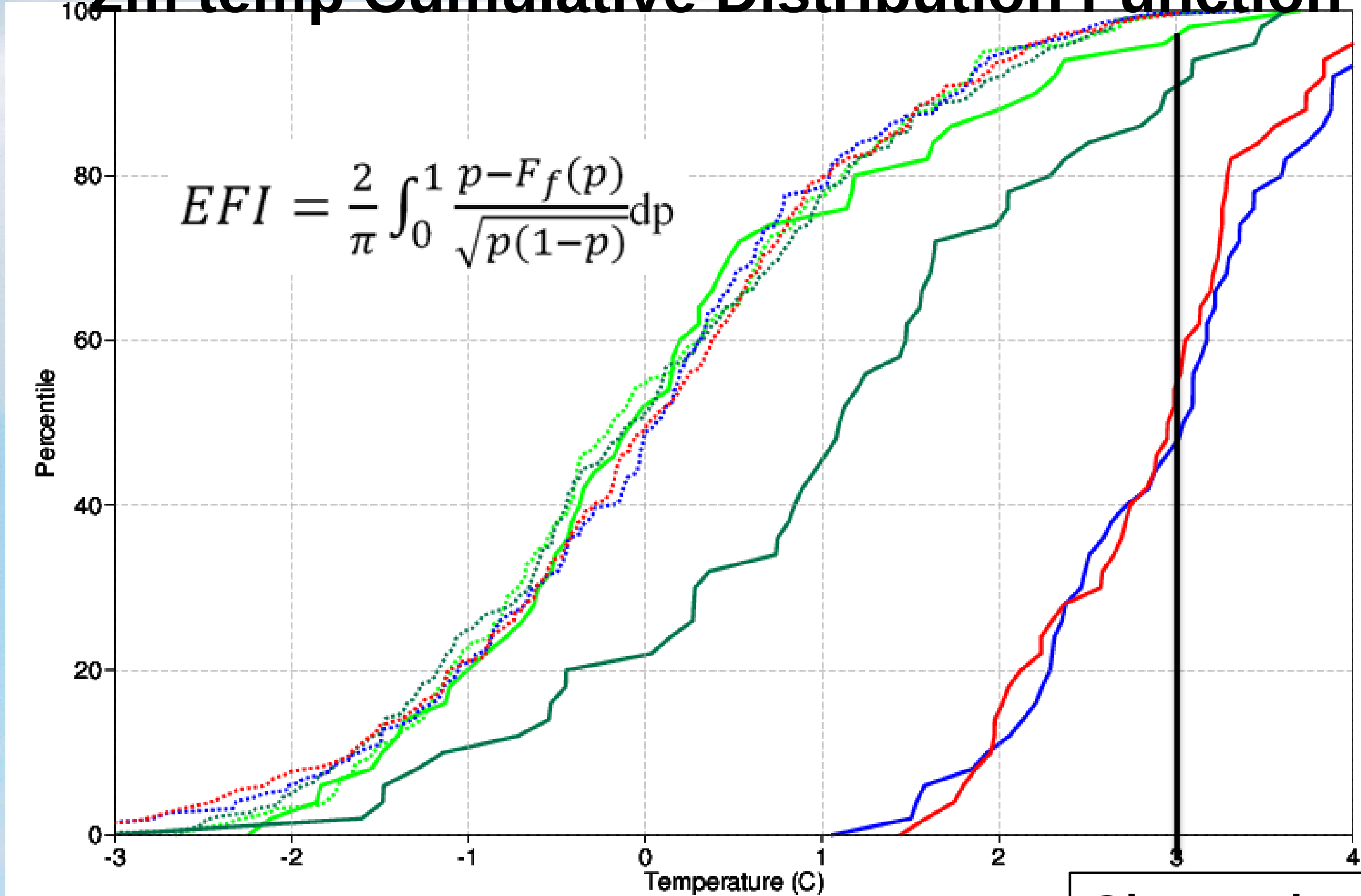
Extreme event verification

Extreme Forecast Index (EFI)

ensemble predictions for 29 June - 5 July 2015

- - - - - Climate — 15 June 2015 (15-21d) — 18 June 2015 (12-18d) — 22 June 2015 (8-14d) — 25 June 2015 (5-11d)

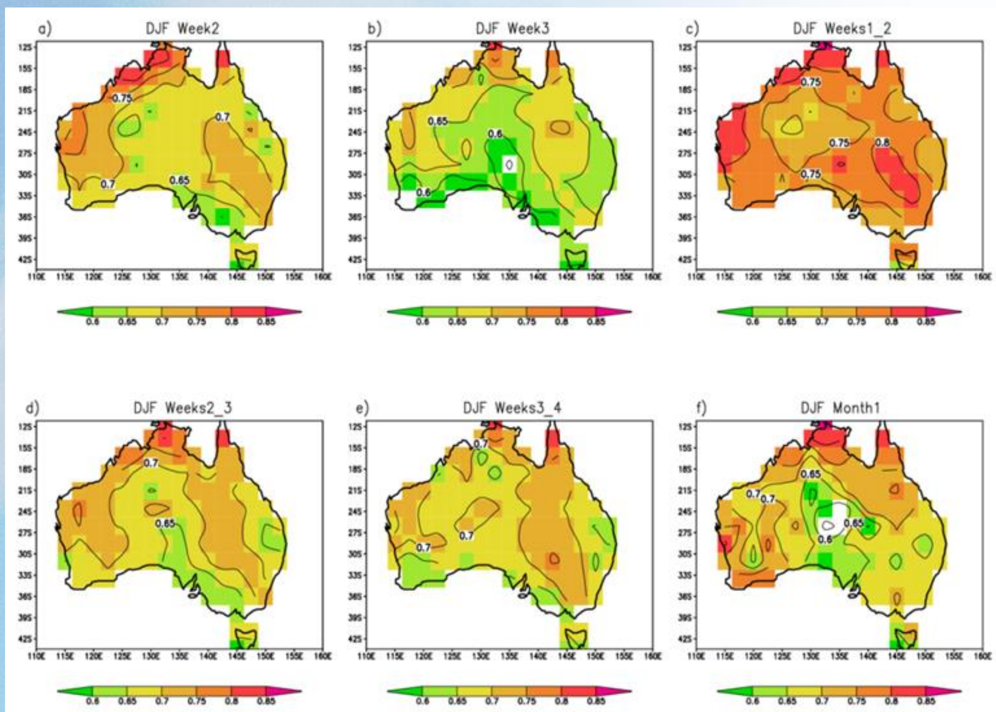
2m temp Cumulative Distribution Function



Observed anomaly

Example of heat wave verification

(Excess Heat Factor index > 0)



For a heatwave to be present, the temperature averaged over three consecutive days has to be greater than the climatological 95th percentile (T95) of daily mean temperature for a given region.

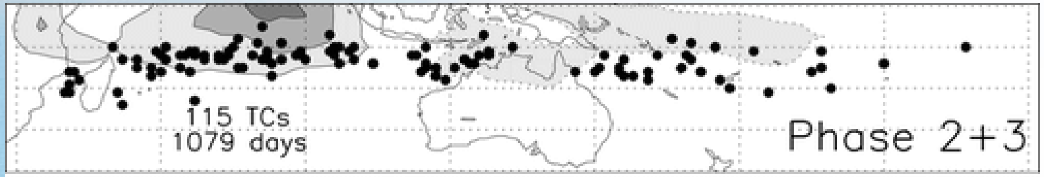
ROC area of the probability of the occurrence of a heatwave (Excess Heat Factor index > 0) [From Hudson and Marshall 2016)

S2S prediction/verification of tropical cyclones

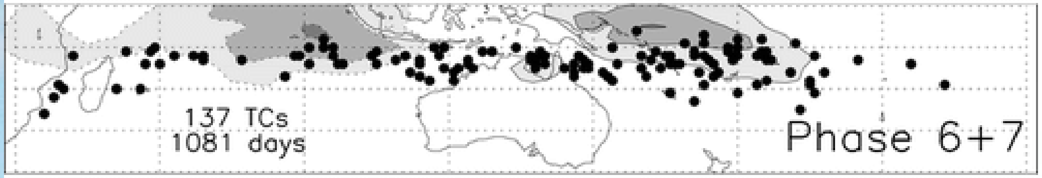
Prediction of more/less TC activity over a sufficiently large area and time window.

Justification: TC genesis is strongly modulated by various models of variability: ENSO, MJO, IOD…. Which can be predicted by models weeks in advance.

MJO Phase 2-3

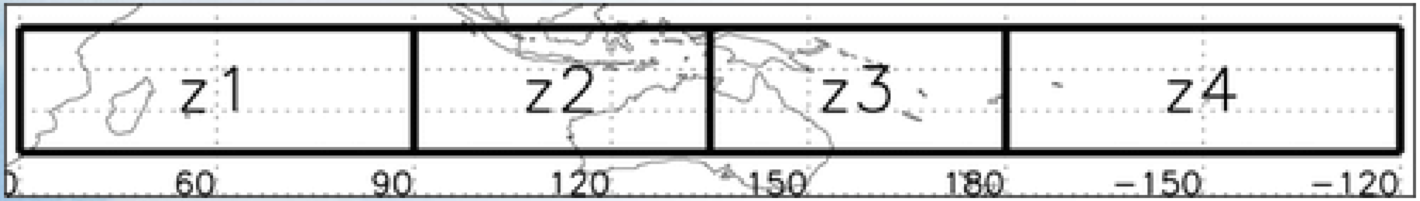


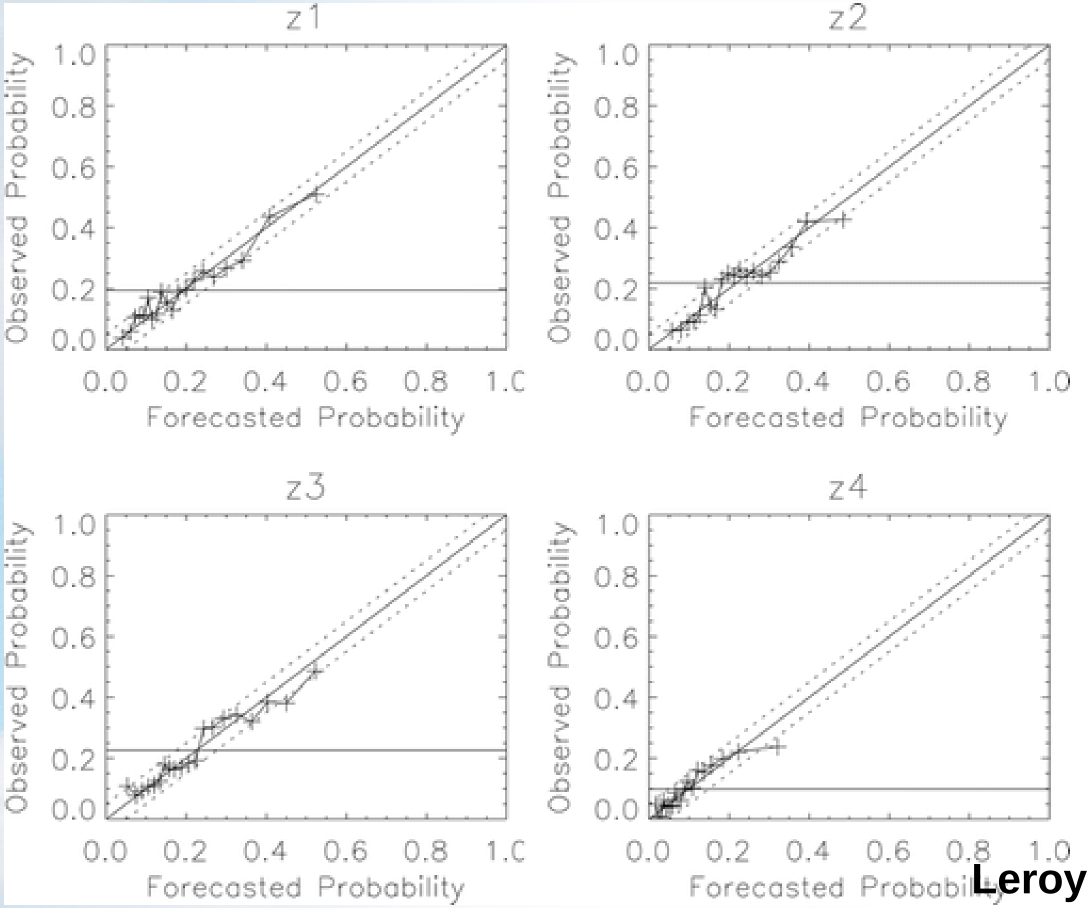
MJO Phase 6-7



Leroy and Wheeler 2008

Verification of statistical TC genesis forecast

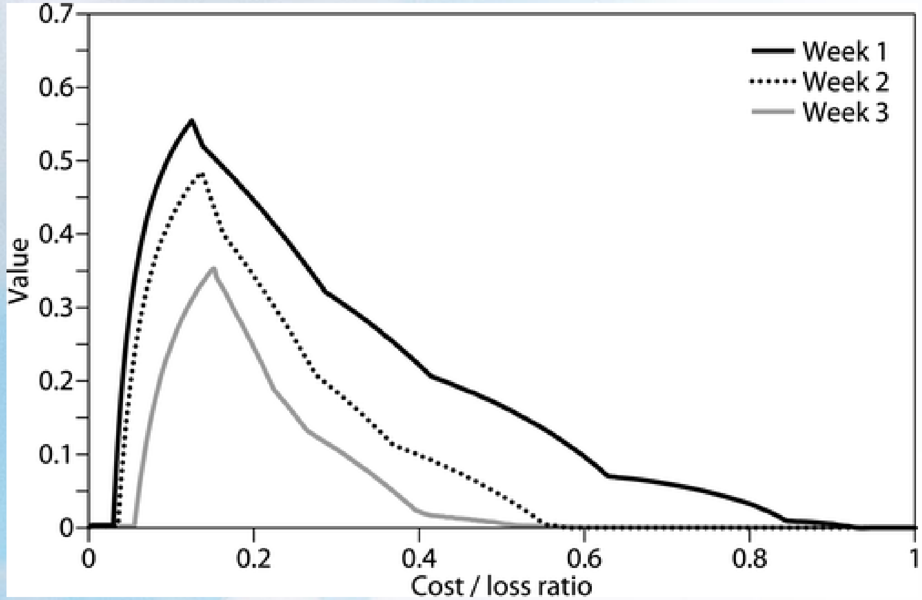


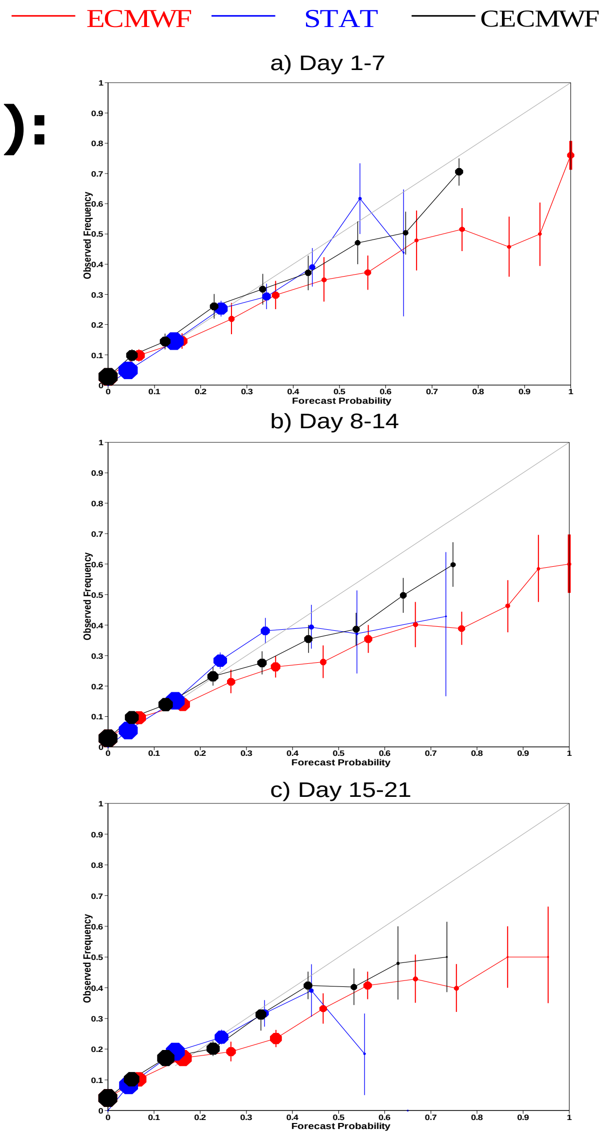


Leroy and Wheeler 2008

Verification over the Southern Hemisphere as in Leroy et al (2008):

Probability of a TC occurrence during **a weekly period over 20x15 degree domains**





Vitart, Leroy and Wheeler, MWR 2010

Conclusions

- S2S prediction and verification is still in an early stage

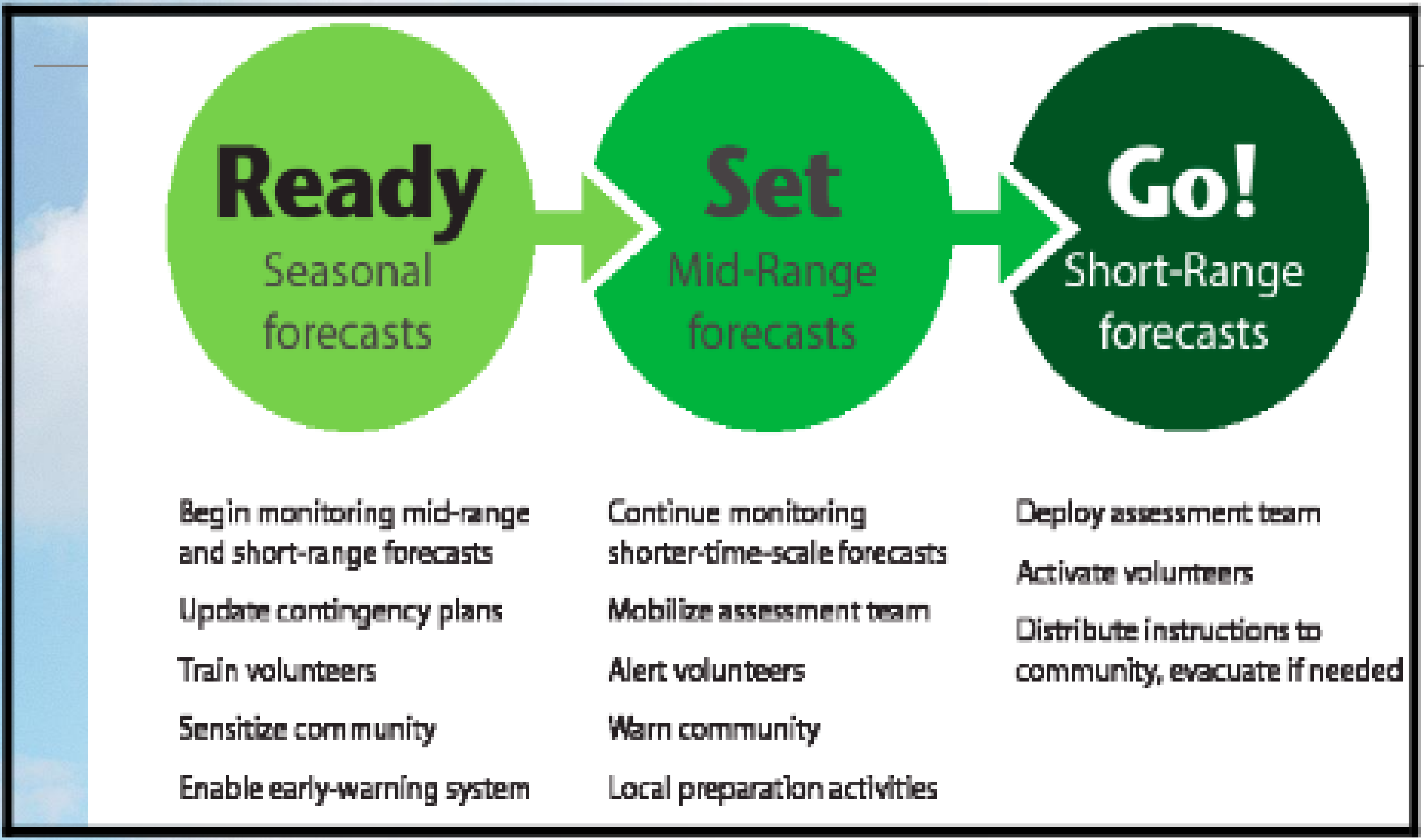
- Re-forecasts needed for model calibration and skill assessment

- A main challenge for S2S verification is computing forecast probabilities under limited (small) ensemble sizes and sometimes relatively small number of re-forecast years in the hindcasts.

- How can we best address verification in a seamless manner, for comparing forecasts across timescales?

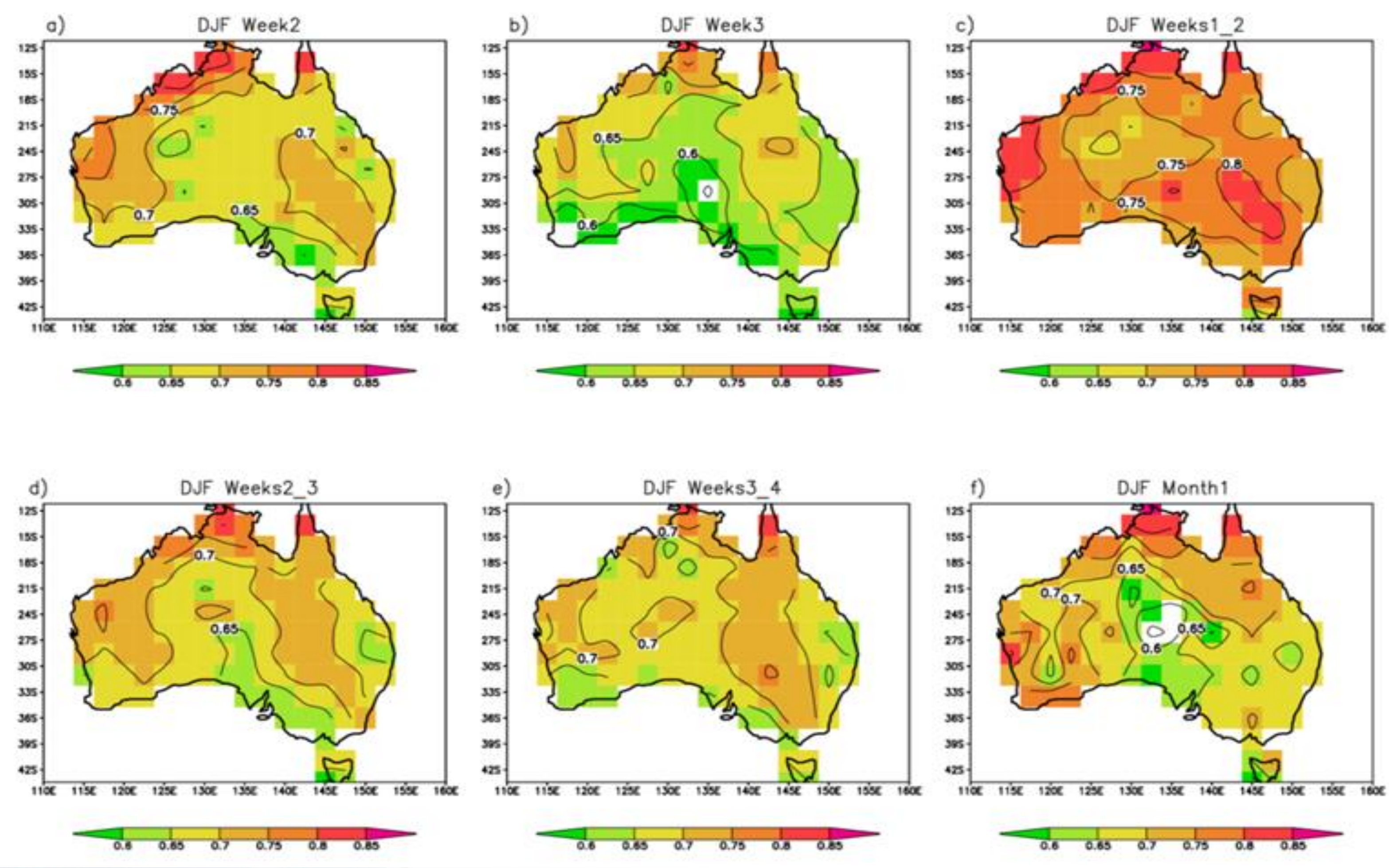
- The S2S database represents an important resource for inter-comparison of S2S forecasting systems and evaluation of the benefits of the multi-model ensemble approach.

Opportunity to use information on *multiple* time scales



Red Cross - IRI example

Example of heat wave verification



ROC area of the probability of the occurrence of a heatwave (Excess Heat Factor index > 0) [From Hudson and Marshall 2016)