Impact of ensemble and hindcast sample size on decadal predictions

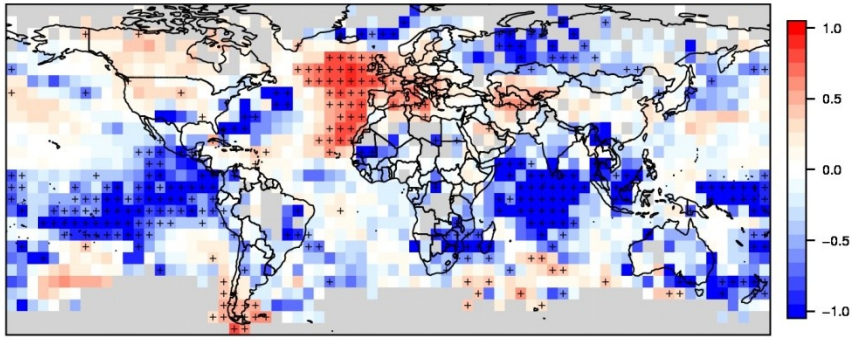
W. A. Müller , F. Sienz, H. Pohlmann

AGCI Workshop Decadal Prediction , Aspen 2015

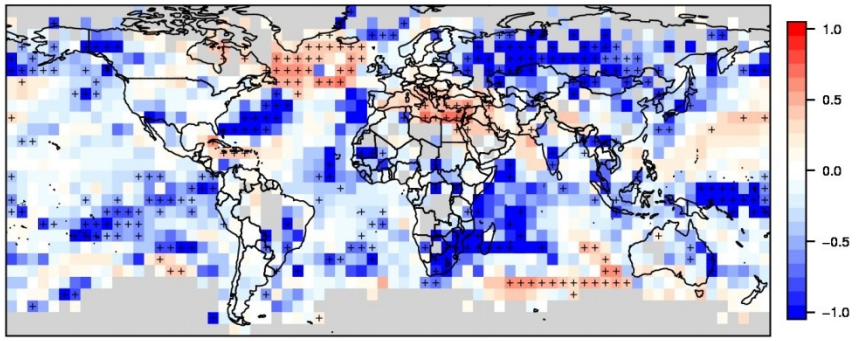
A 3-member example

MPI-ESM-LR, tas, Yr2-5, r(Ini) – r(Unini), 1961-2012

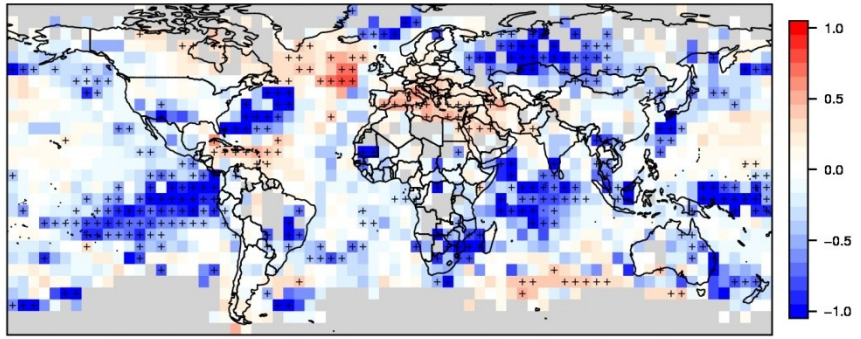
Members 1,2,3



Members 8,9,10



Members 1-10



Outline

- Impact of ensemble size and number of initialization on decadal predictions in a conceptual model
- Impact of ensemble size in the MPI-ESM model
- Decadal predictions from 1901-2010 in the MPI-ESM model

A linear regression model

Data and Region:

North Atlantic SST 20-60N
HadISST

A LM-Model:

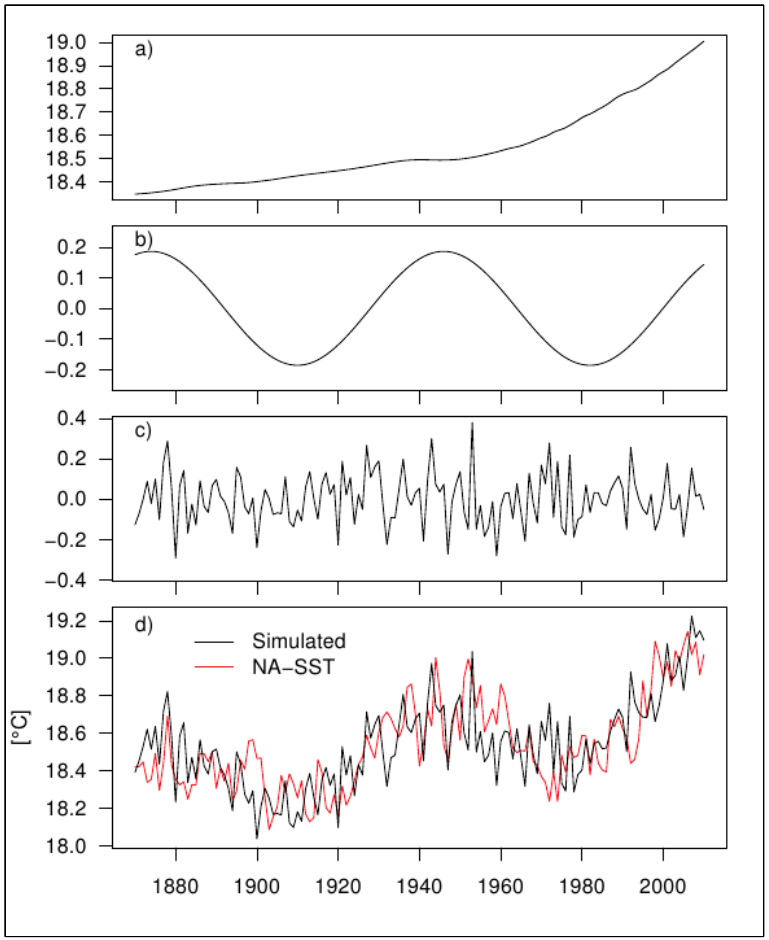
$$y_t = \tilde{y}_t + \varepsilon_t \text{ , with}$$

$$\tilde{y}_t = \alpha_0 + \alpha_1 x_{CO_2,t} + \alpha_2 \sin(2\pi t/P) + \alpha_3 \cos(2\pi t/P)$$

Ensemble Perturbation:

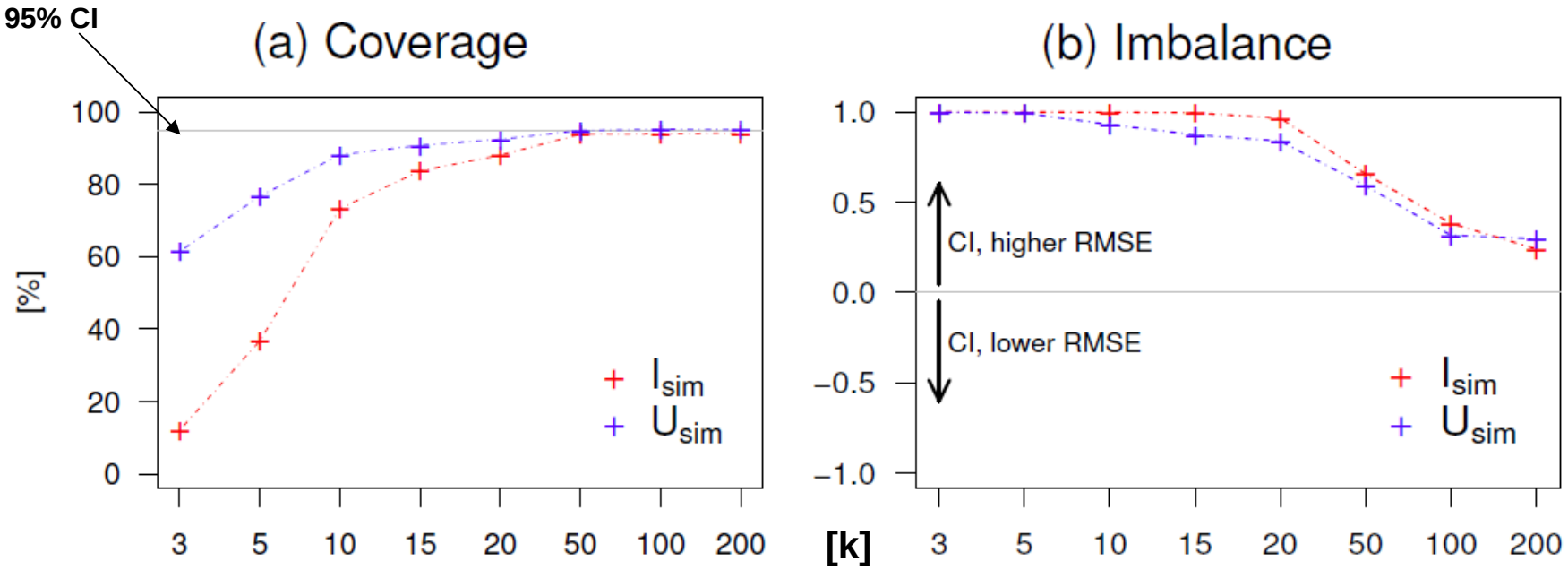
$$I_{sim,k,t} = \tilde{y}_t + \frac{1}{k}\sum_{j=1}^{k} \varepsilon_{j,t}$$

$$U_{sim,k,t} = \frac{1}{k}\sum_{j=1}^{k} (\tilde{y}_{j,t} + \varepsilon_{j,t})$$



Sienz et al., 2015 (submitted)

Impact of Ensemble Size on Test Statistics (e.g. RMSE)

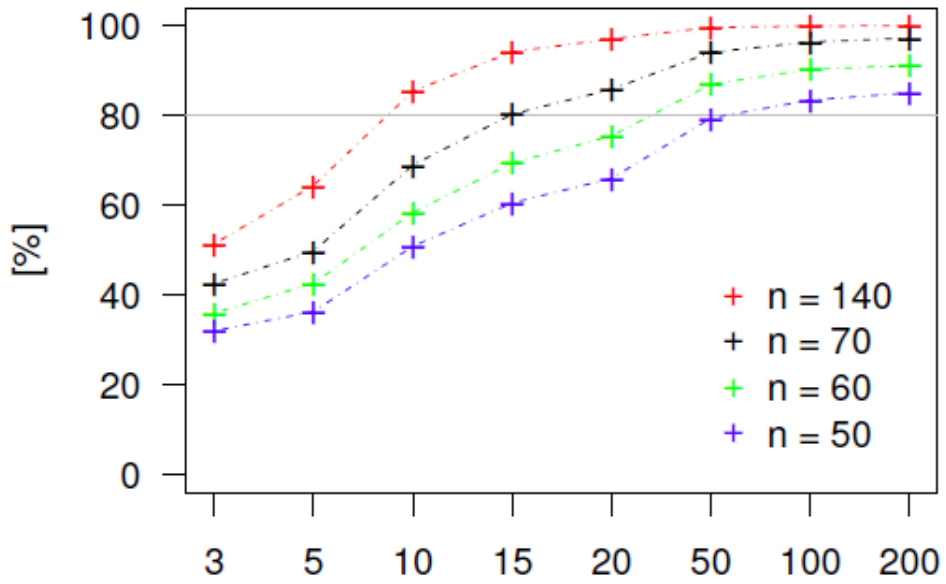


An analysis of confidence interval (CI) and statistical test properties requires the knowledge of the true values.

The number of times the CI include the true value is the coverage, whereas the imbalance measures the percentage of CI above or below the true value.

For small k: CI not include true values and (here) shift to higher RMSE

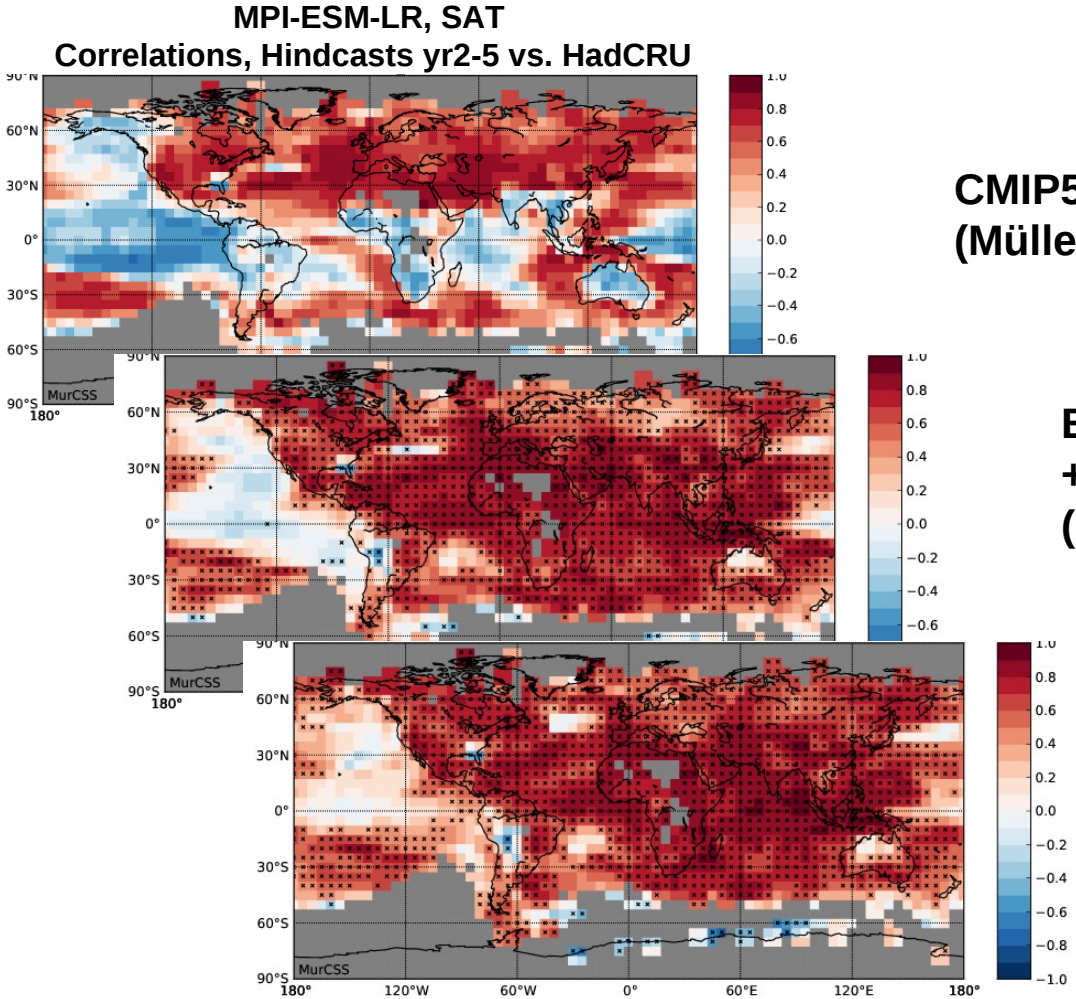
The Power as a function of ensemble Size (k) and number of initialisations (n)



The performance of significance tests is investigated in terms of power. Here: the power to detect a RMSE skill-score of 0.3 as significant greater than zero.

For small k and/or n: Significance test may not be applicable

MiKlip – Towards an operational decadal climate prediction system



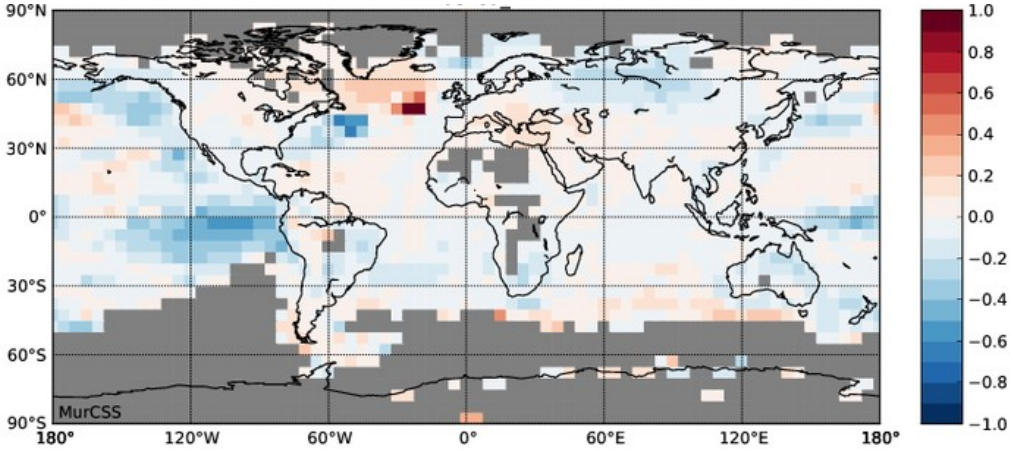
CMIP5, b0, 3 mem
(Müller et al. 2012, GRL)

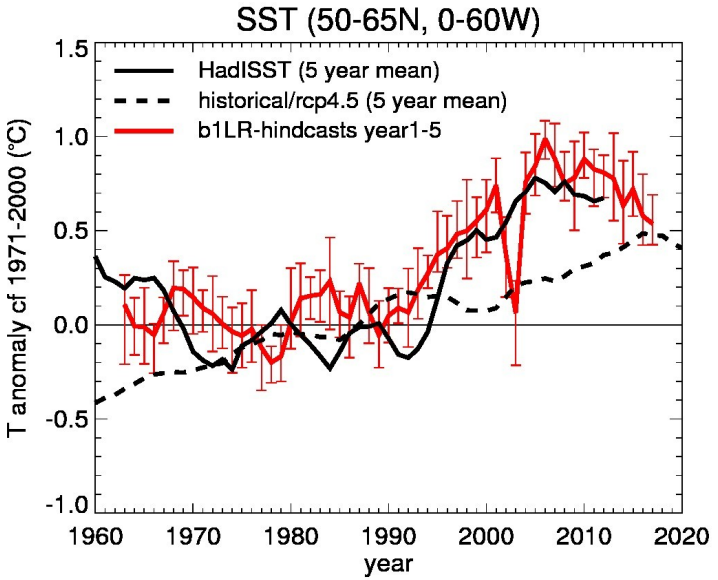
Baseline 1, b1, 10 mem
+ new intitialisation (ORA-S4)
(Pohlmann et al. 2013, GRL)

Prototype, pr, 30 mem
+ new intitialisation
(full field + ORA-S4/GECCO2)

MPI-ESM decadal prediction experiments

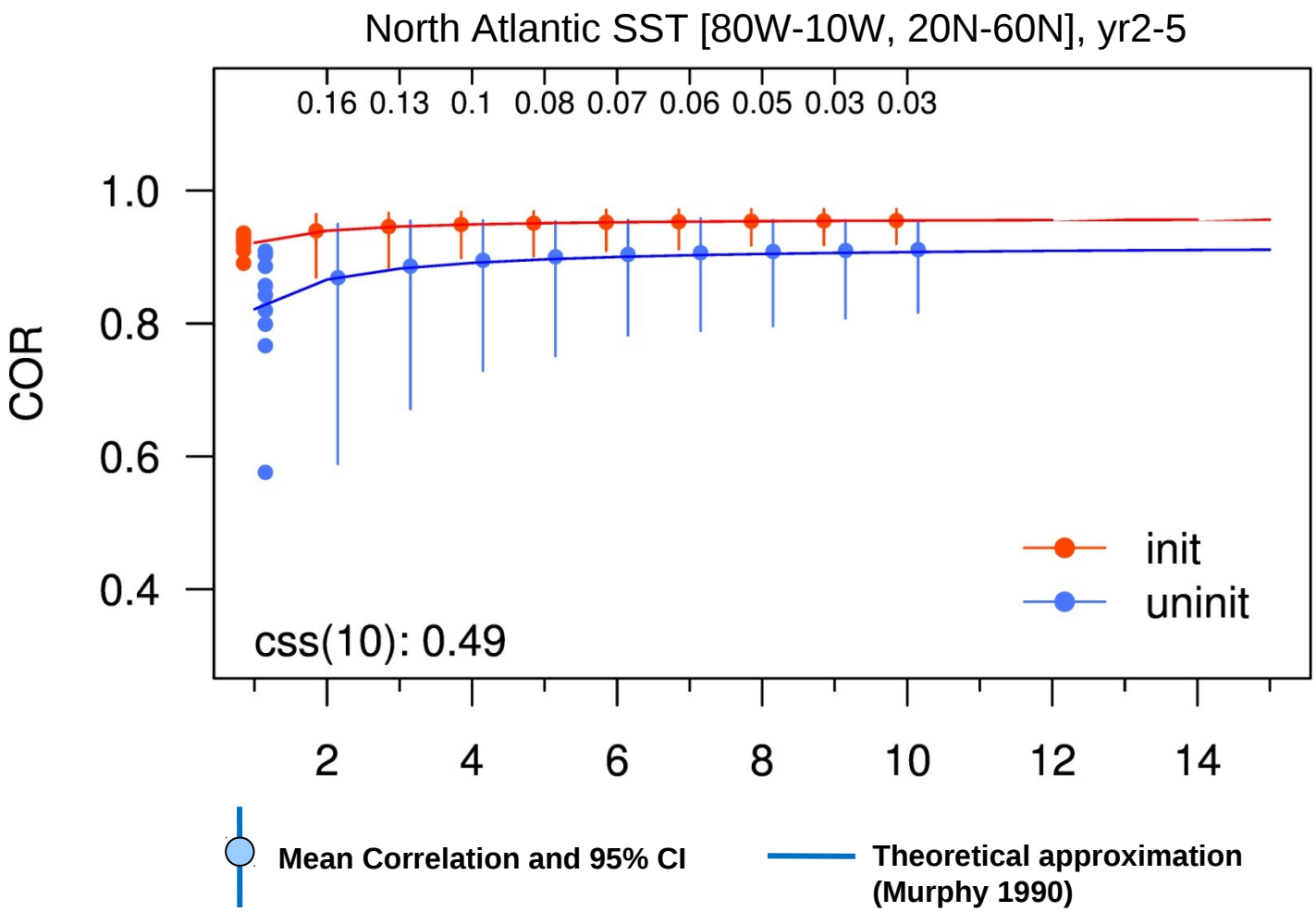
- Low- resolution MPI-ESM-LR (Ech6 T63L47,MPIOM Gr1.5° L40)
- Assimilation
  - Full field atmosphere by ERA40/ERAinterim (vor, div, T, logp)
  - Anomaly ocean by ORA-S4 (T&S)
- Yearly initializations from 1961-2012, 10 members



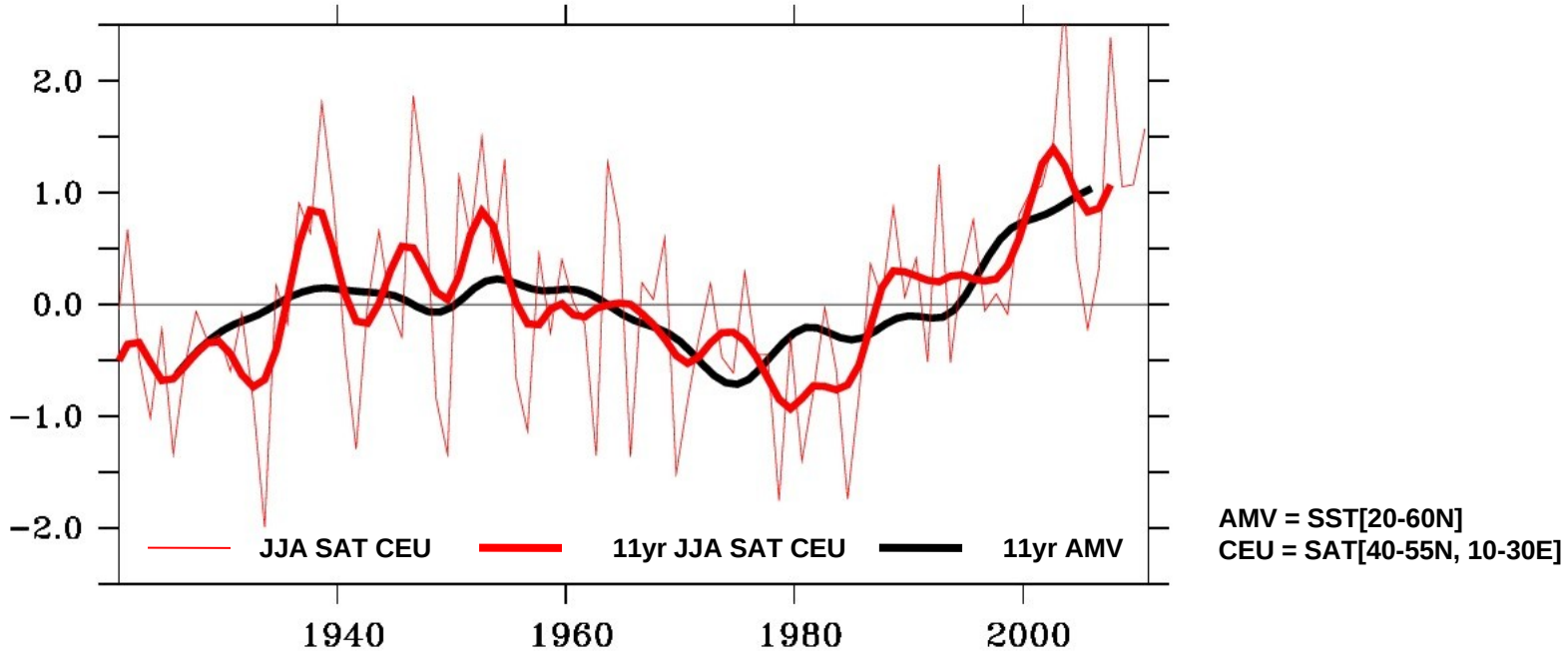


Ensemble Size & Forecast Skill

correlation and 95% confidence interval



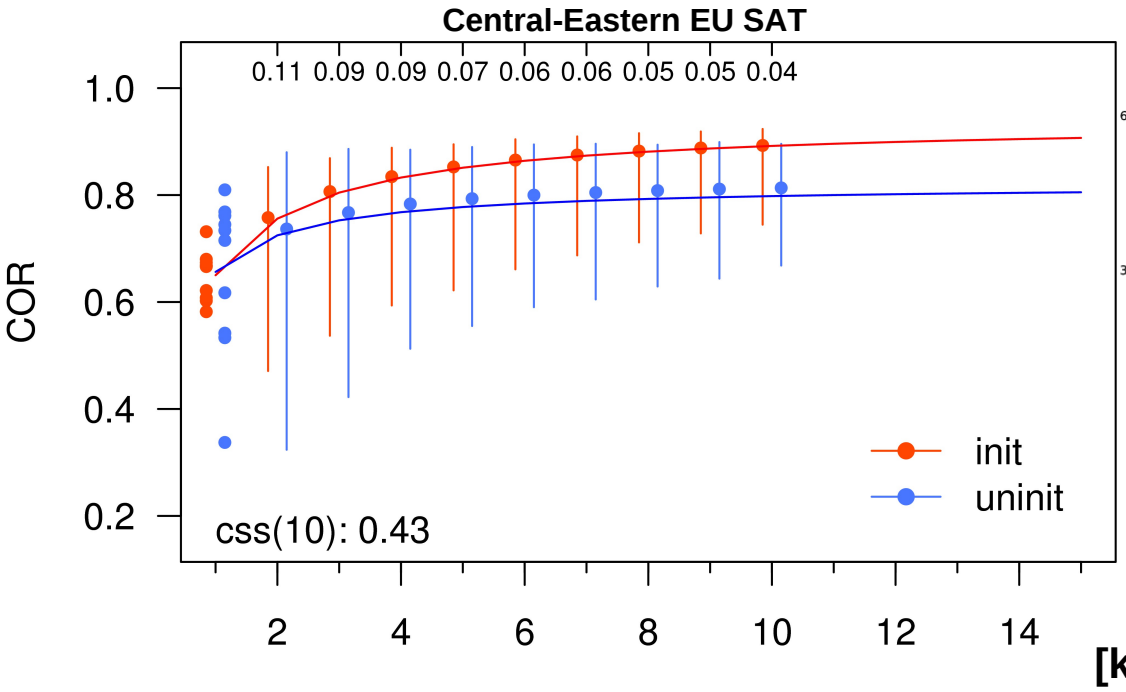
Multi-decadal variations of European summer climate, here 20CR

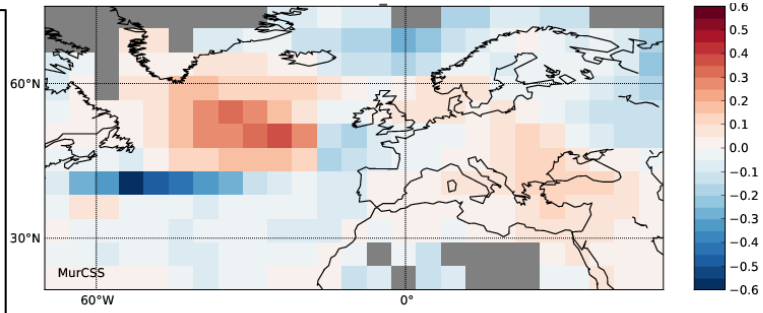


Central European SAT exhibits a pronounced multidecadal variability. During the latest positive phase the well-known heat waves occur, such as 2003, 2007, 2010.

There has been a recognition that North Atlantic MDV has an influence on EU summer (e.g. Sutton and Hodson, 2005). Suggested mechanism by geostrophic adjustment to diabatic heating anomalies (Ghosh et al. 2015, submitted)

Prediction of European multi-years summer means





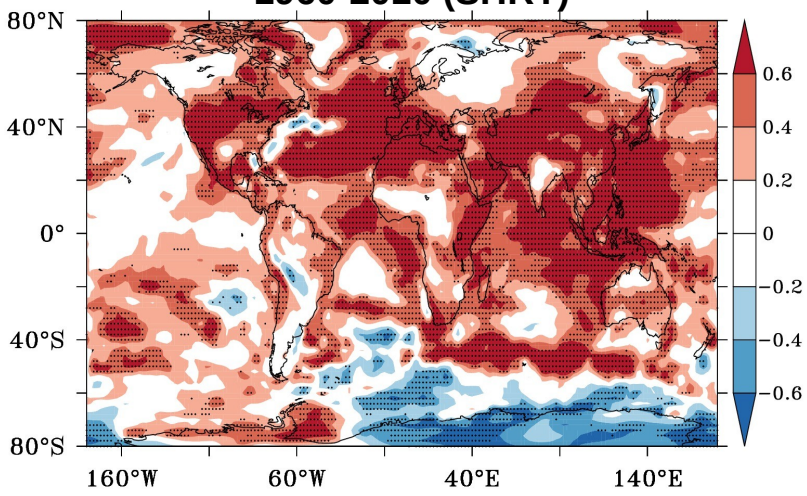
Forecast skill in multi-year summer in central-eastern Europe (Müller et al., 2012 GRL)

Strong contribution by the trend. For large ensemble size forecast skill from initializations becomes significantly different from historicals (Sienz et al., 2015)

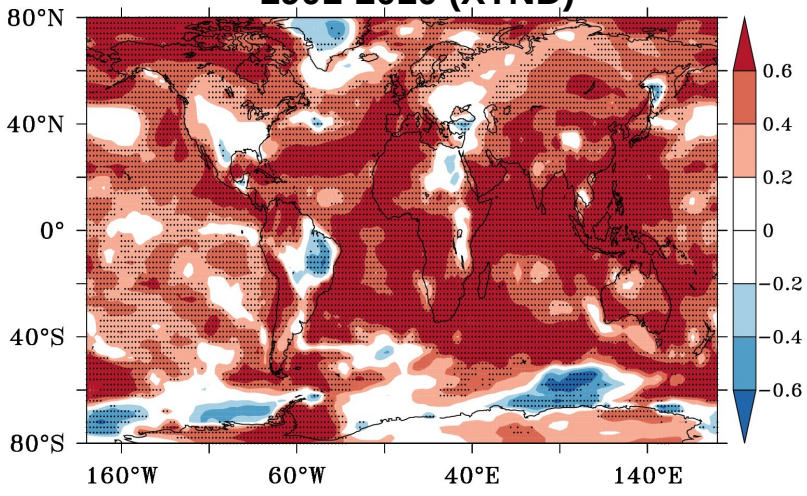
Extended decadal prediction experiments

- MPIOM (1.5°) forced with an ensemble of the 20th century reanalysis for the period 1871-2010, 3 members (Müller et al. 2014 ClimDyn)
- Assimilation with MPI-ESM
  - From 3 forced ocean experiments, only 3D temp. and salinity are assimilated, no atmospheric assimilation
- Yearly initializations from 1901-2010, 3 members (Müller et al. 2014 GRL)

1960-2010 (SHRT)

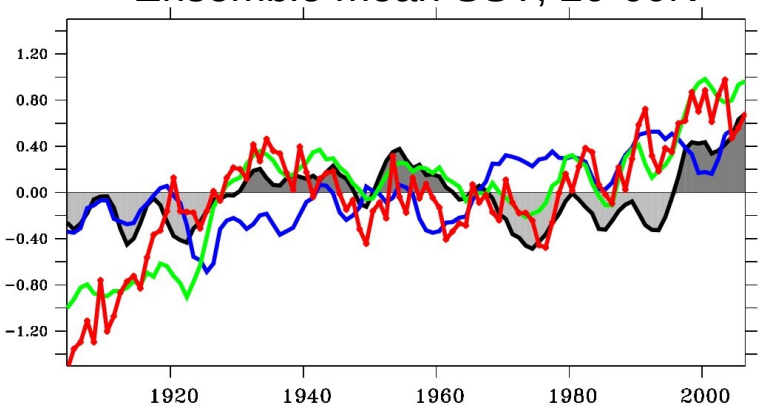


1901-2010 (XTND)



Prediction of North Atlantic SST

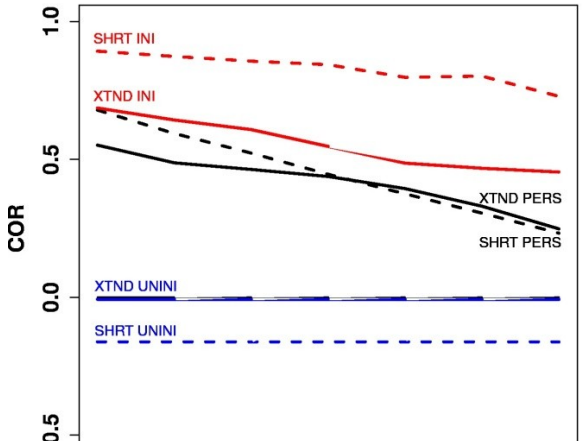
Ensemble mean SST, 20-60N



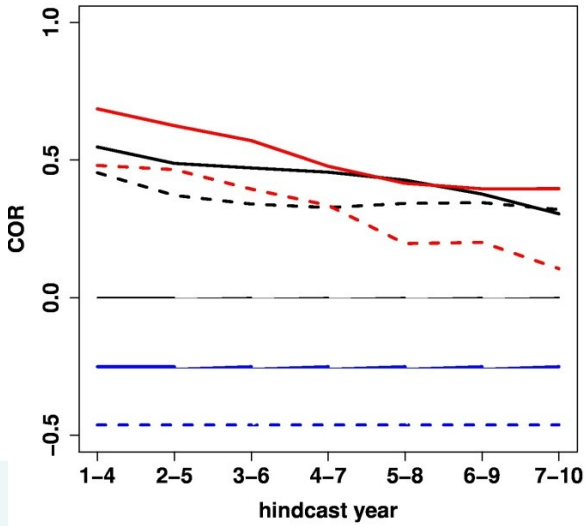
20cr
Uninitialised
Initialised yr2-5
Assimilation

- **Hindcasts before 1950s are possible**
- **Skills become more robust for XTND**

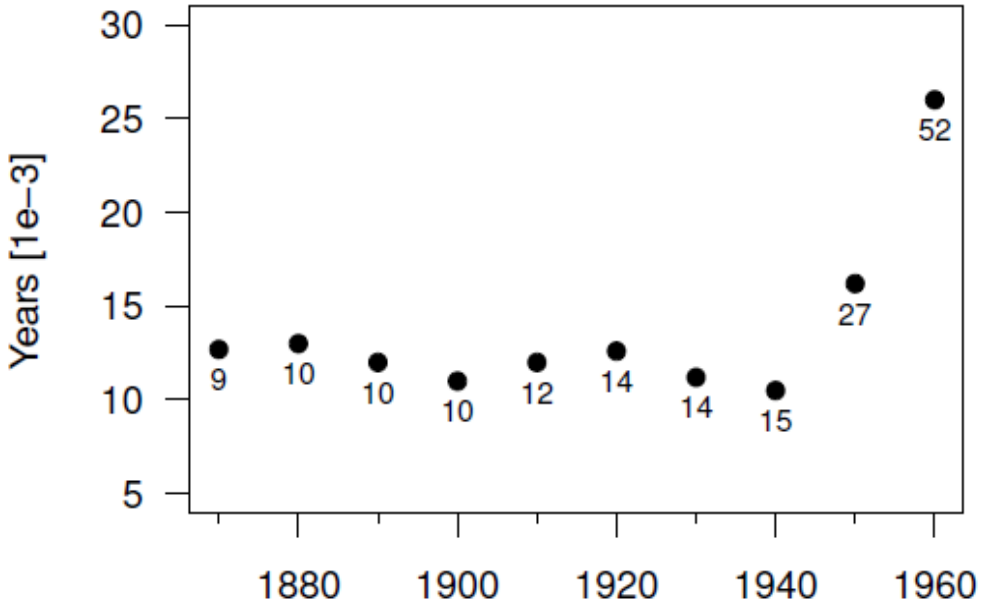
Ensemble mean SST, 40-60N
With trend included



detrended



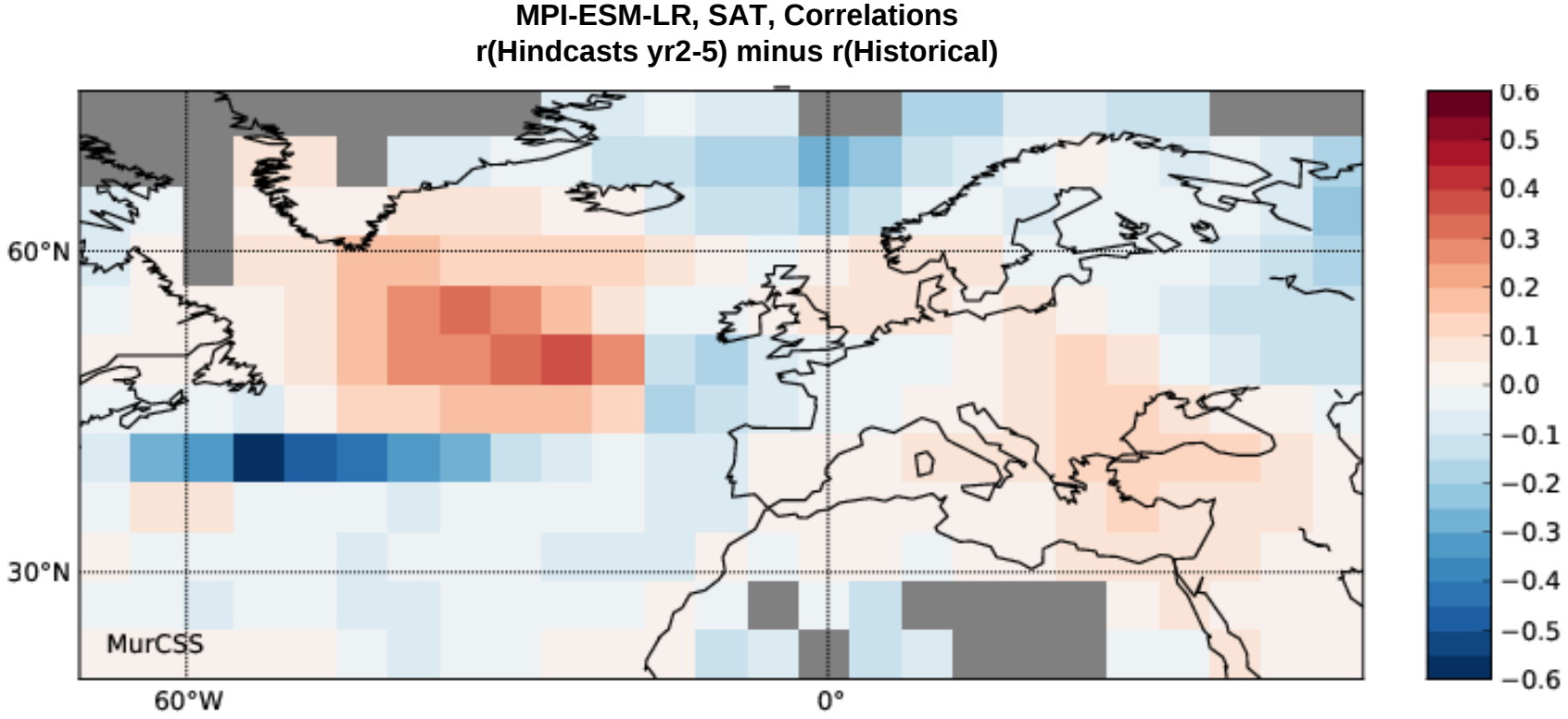
Summary



Number of required model years to achieve a power of 0.8, dependent on hindcast years and ensemble size in a conceptual model for North Atlantic SST.

- Low ensemble and hindcast sample size may result in biased skill properties and misleading significance estimates
- Relevance for prediction system development and comparisons
- Given costs of hindcast setup, backward extension of initializations most effective

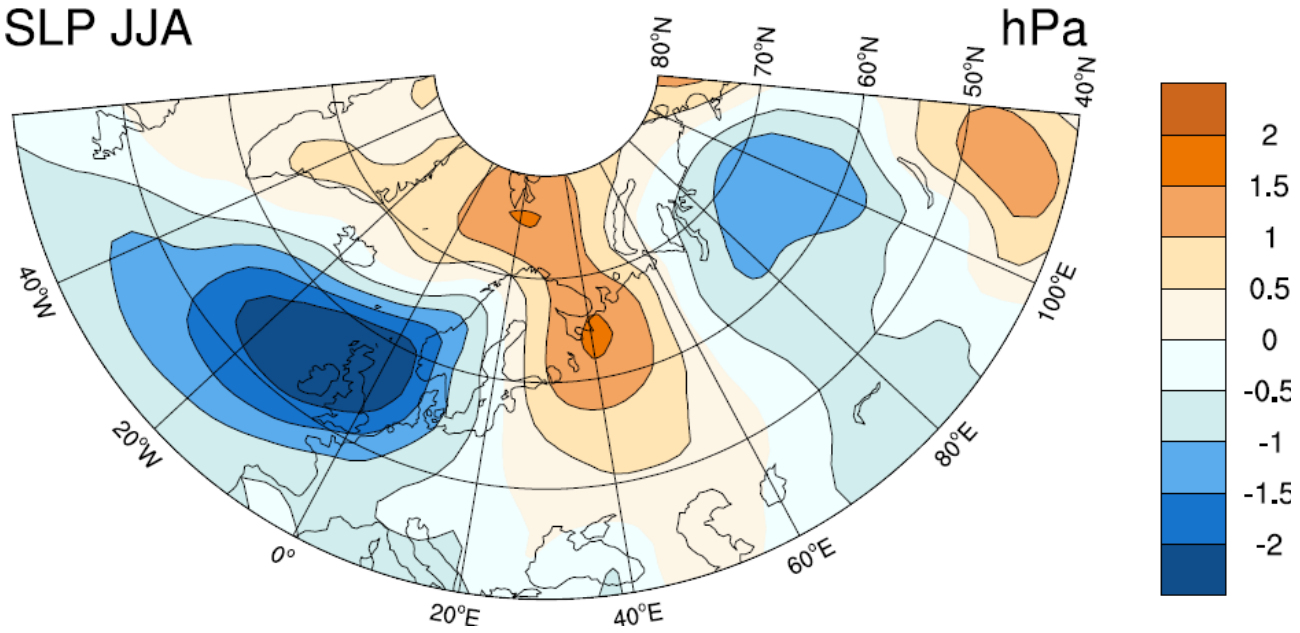
Prediction of European multi-years summer means



Forecast skill in multi-year summer in central-eastern Europe (Müller et al., 2012 GRL)

A Quasi-geostrophic response

Composites of SLP, 20th century reanalysis



A low surface pressure emerge eastward of the positive heat fluxes and a wavetrain further downstream. The response is weakly baroclinic (Kushnir et al 1995), and not related to NAO

At its eastern part of the low pressure, temperature is geostrophically advected. The subsequent positive anomaly favor a blocking-like effect, with local radiative and possible land-surface influence and thus alter central EU temperature variations.

Ghosh et al., 2015 (in prep.)