# Assessing representations of model uncertainty in seasonal forecast ensembles

**Antje Weisheimer**
ECMWF and Oxford University, NCAS

Thanks to

**Francisco Doblas-Reyes, Tim Stockdale, Thomas Jung, Glenn Shutts, Hannah Cloke, Florian Pappenberger and Tim Palmer**

# ENSEMBLES
## "ENSEMBLE-based Predictions of Climate Changes and their ImpactS"

### Multi-model ensemble (MME)

- five coupled atmosphere-ocean GCMs for seasonal forecasts developed in Europe
- 9 initial condition ensemble members → 45 members

### Perturbed parameter ensemble (PPE)

- uncertainty in poorly constrained cloud physics and surface parameters in HadCM3
- 8 model versions with simultaneous perturbations to 29 parameters + 1unperturbed member
- no control run available

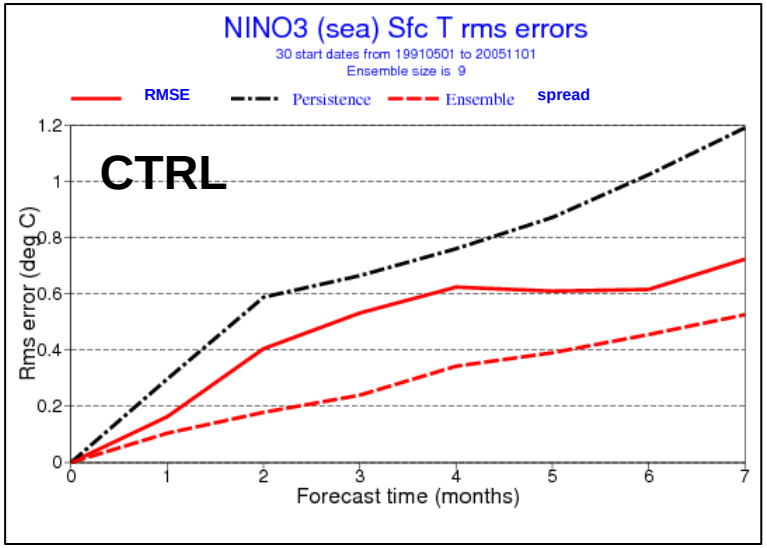
### Stochastic Physics Ensemble (SPE)

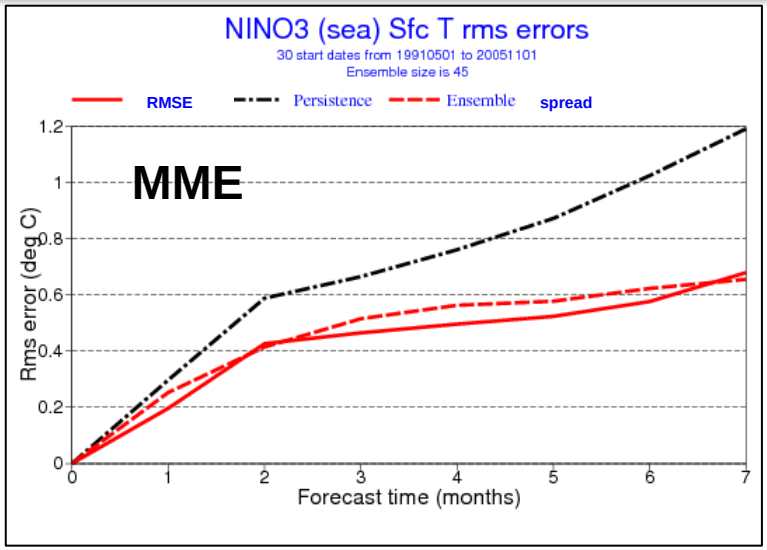
- model uncertainty in coupled ECMWF model due to variability of the unresolved scales
- two-scale perturbed diabatic tendency scheme: $\tau$=6h/30d and L=500km/2500km
- kinetic energy backscatter scheme
- 9 initial condition ensemble members

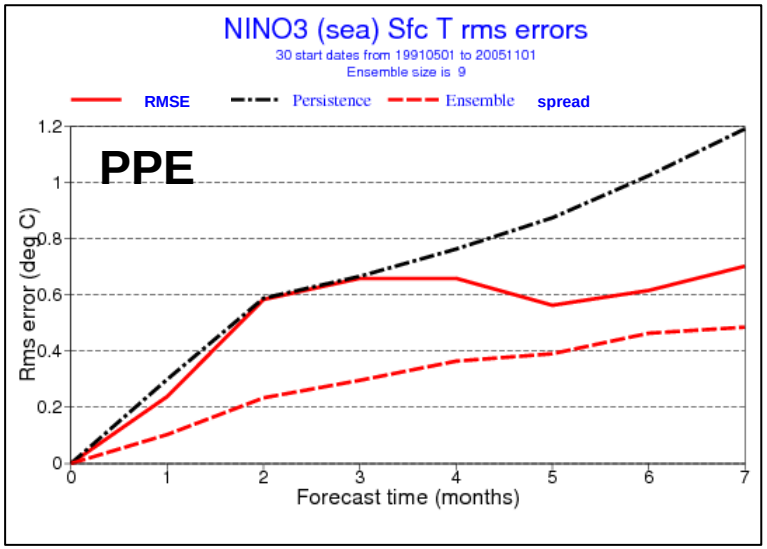
- control run without any stochastic physics (CTRL)

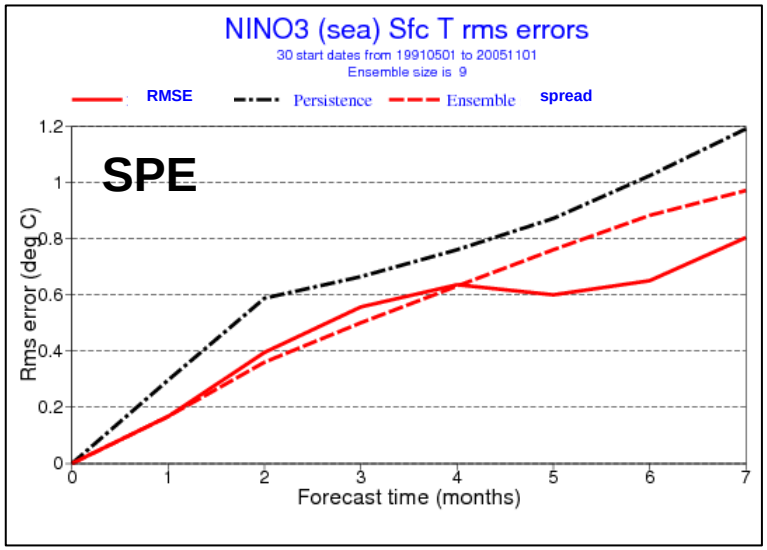
**Coupled seasonal forecasts over the period 1991-2005**
**Initialised on 1st May and 1st November each year**
**Assessment of monthly and seasonal forecast skill**

# Niño3 SST forecasts: spread-skill



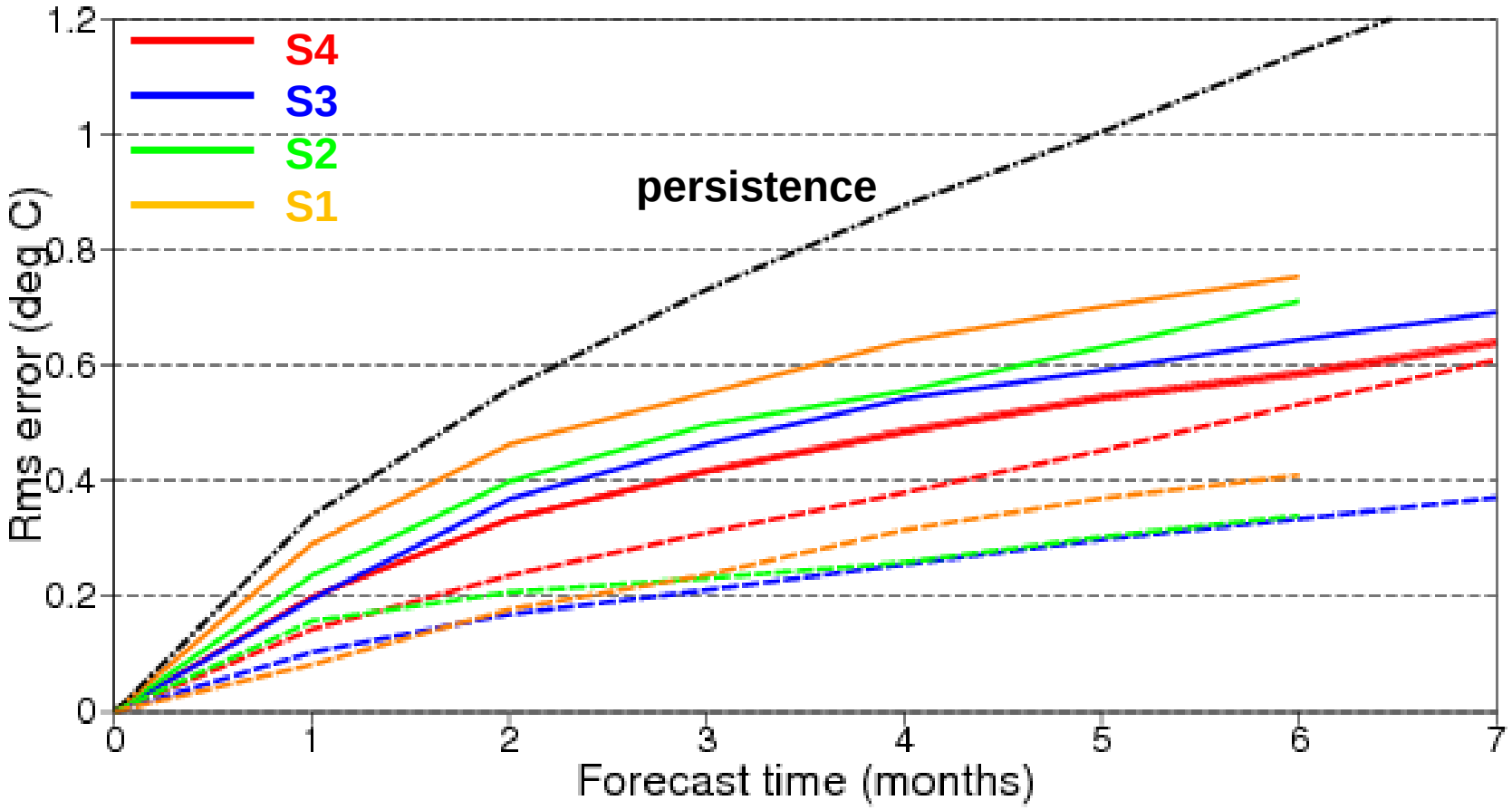




**Weisheimer et al. (2011)**

# Niño3 SST forecasts: spread-skill in operational systems

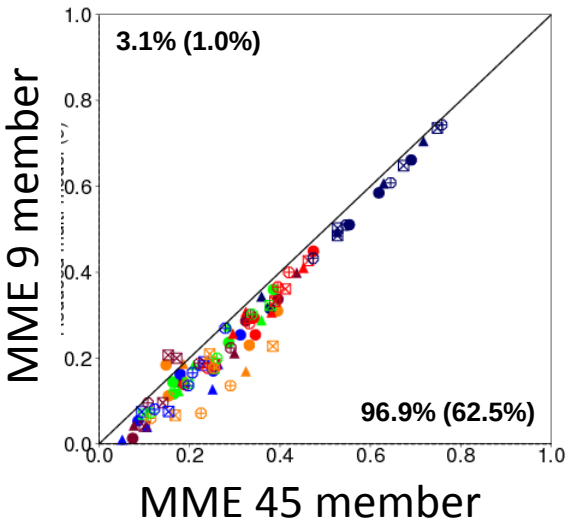
## Performance of the new operational seasonal forecasting system S4 (new stochastic physics) versus previous systems



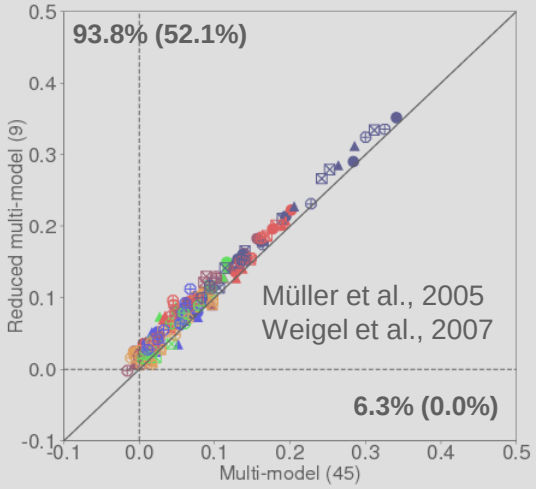

Courtesy Tim Stockdale

# Impact of ensemble size

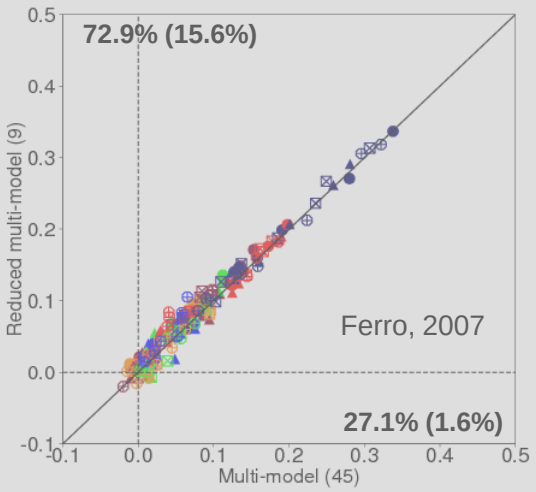
**Anomaly correlation**



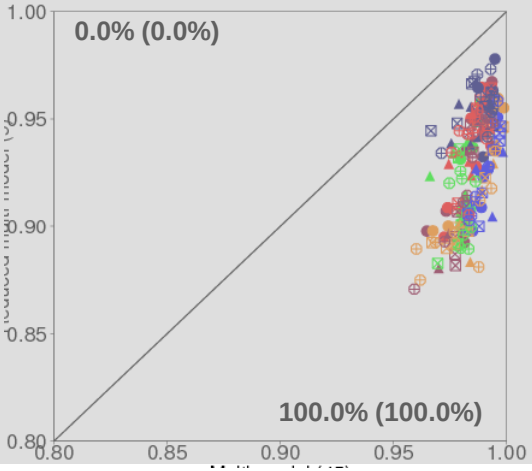

**Debiased Brier Skill Score**



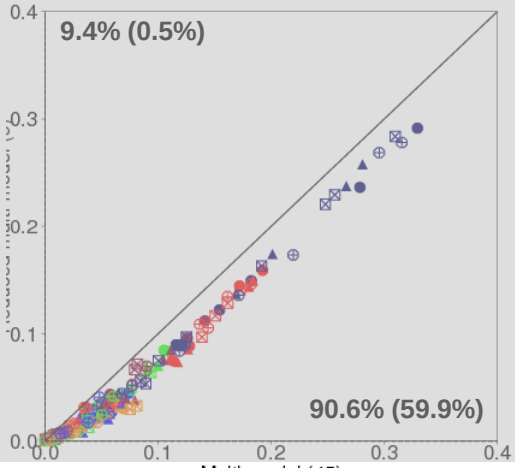

**Infinite size ensemble Brier Skill Score**




**Reliability skill score**

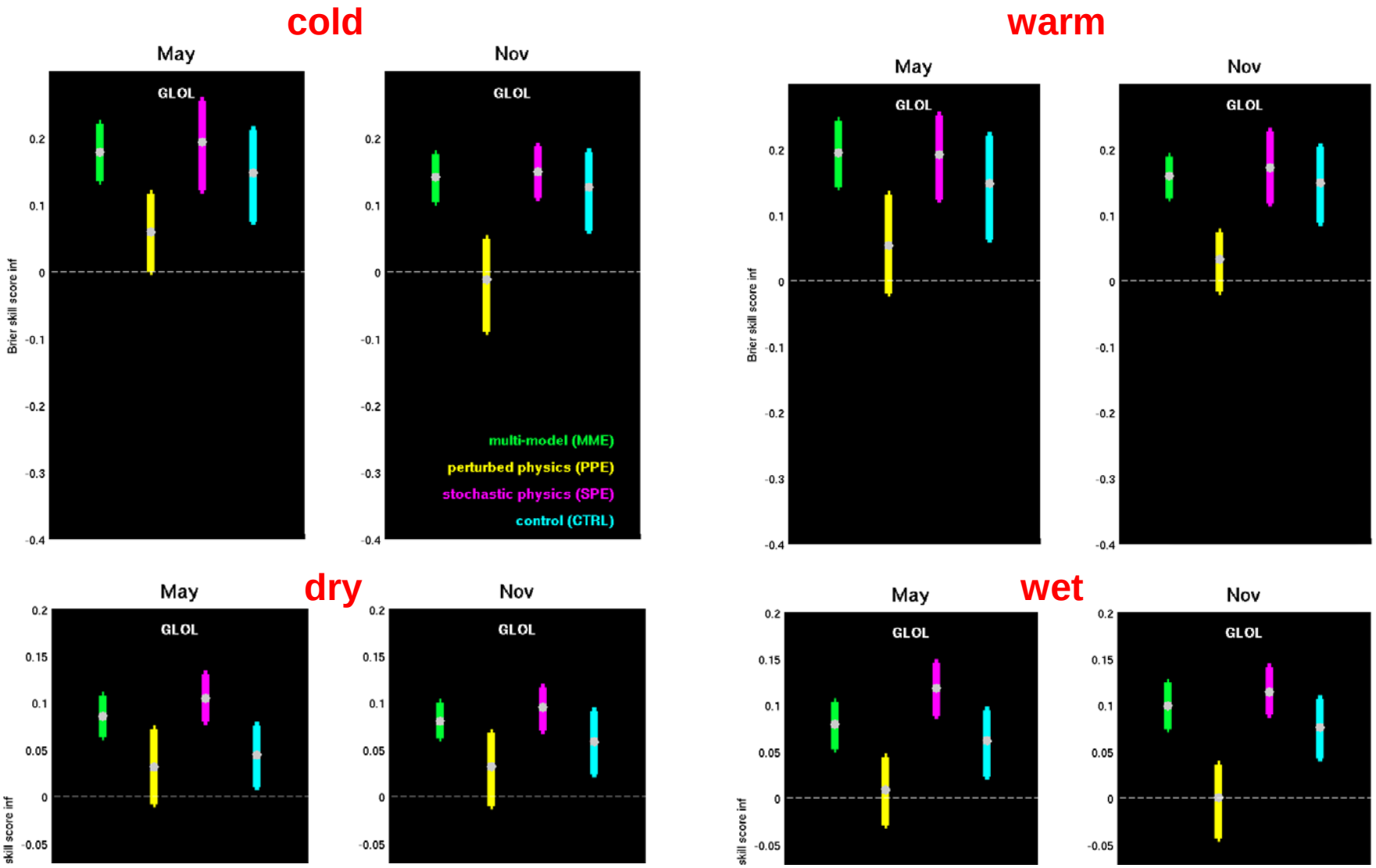



**Resolution skill score**



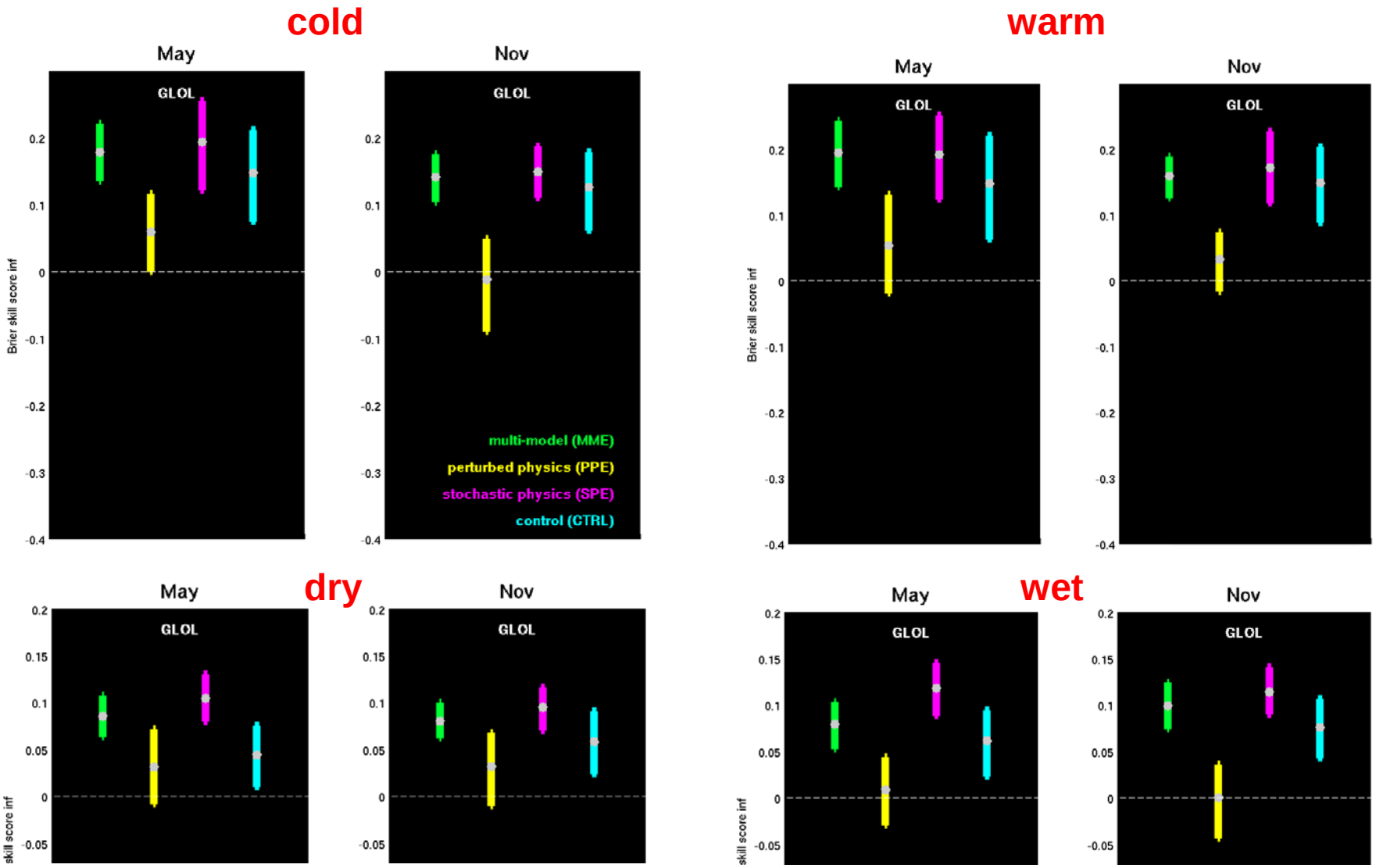

- near-surface temperature (1960-2005) and precipitation (1980-2005)
- six large-scale regions (tropical/NH/SH land/ocean)
- Feb, May, Aug and Nov start dates
- lead times 2-4 and 5-7 months
- lower and upper tercile events

# Global land areas  Brier Skill Score ∞  1st month



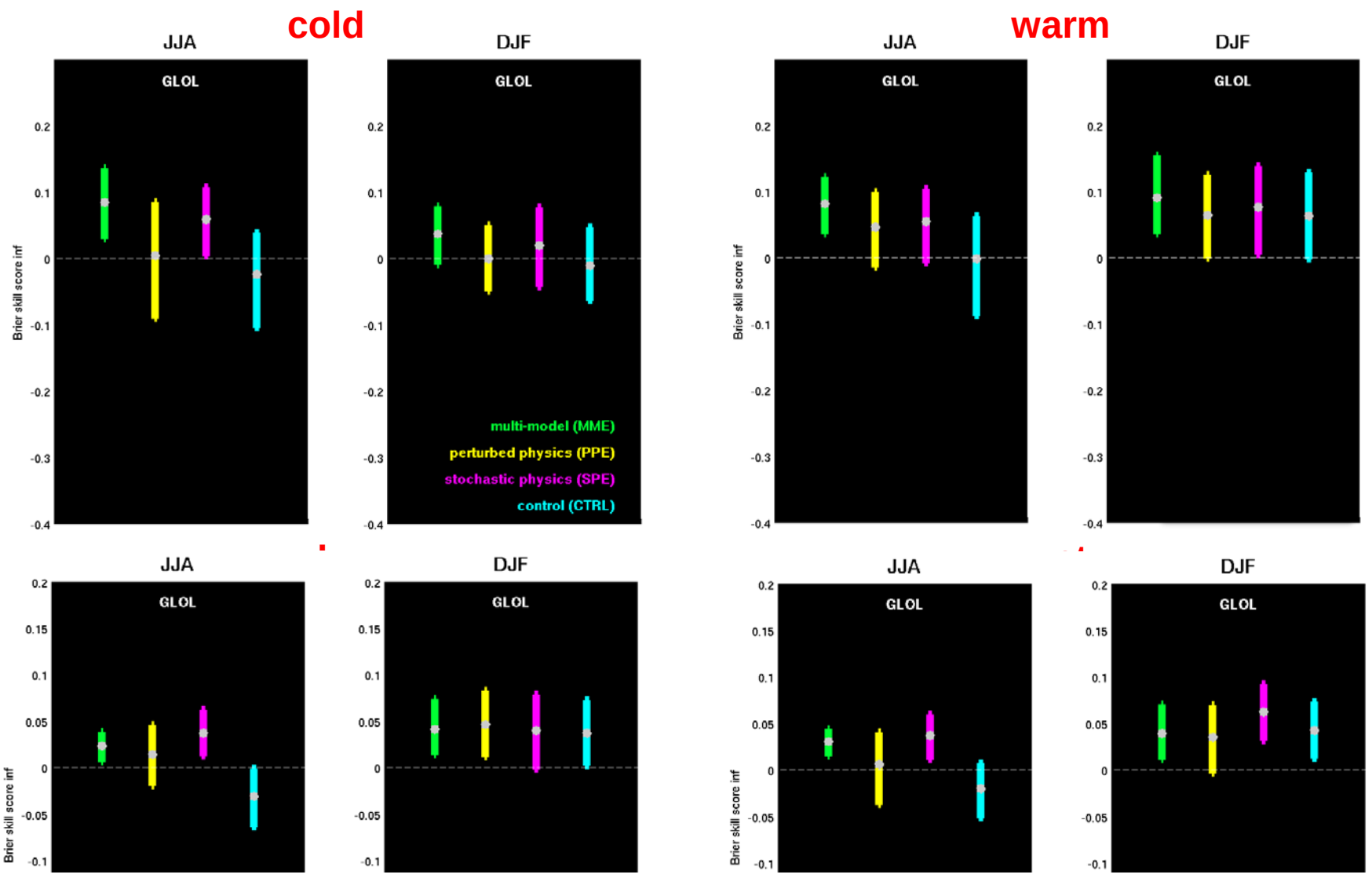

Weisheimer et al. (2011)

# Global land areas  Brier Skill Score ∞  1st month




Weisheimer et al. (2011)

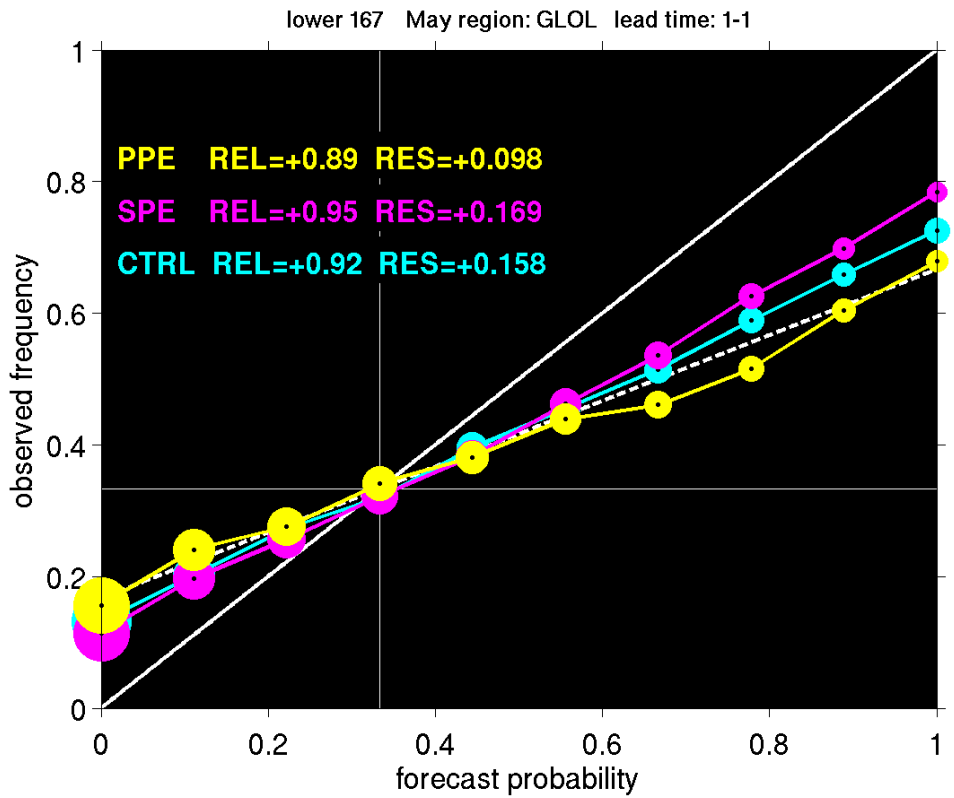
# Global land areas  Brier Skill Score ∞  months 2-4



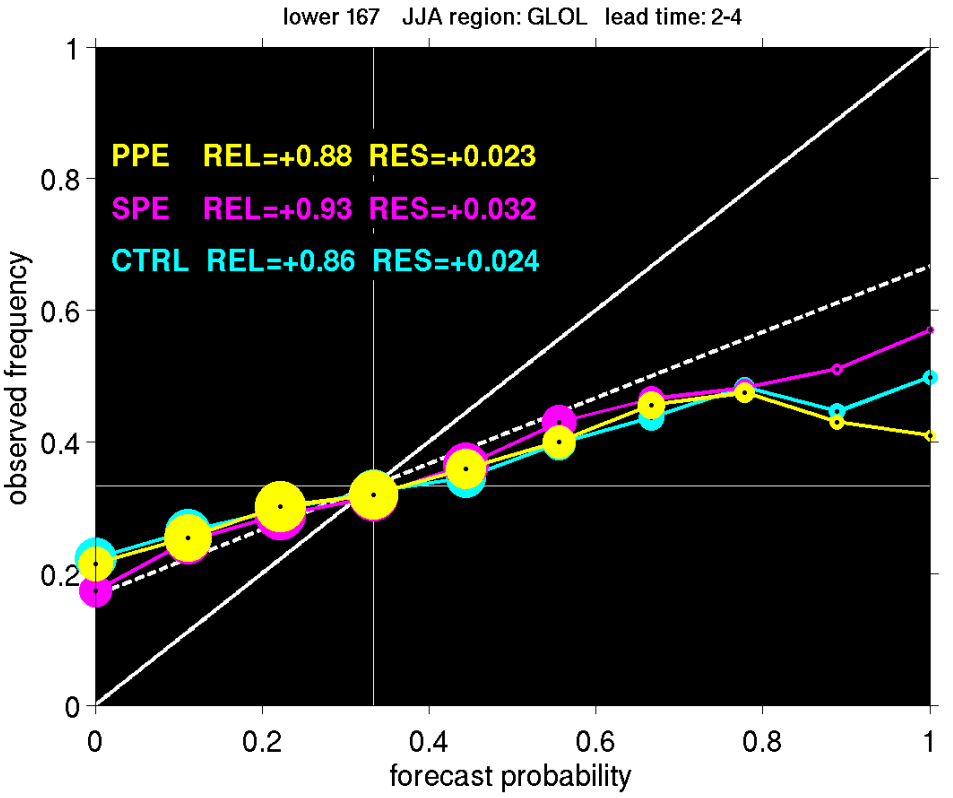
# Global land areas reliability diagrams

## cold events

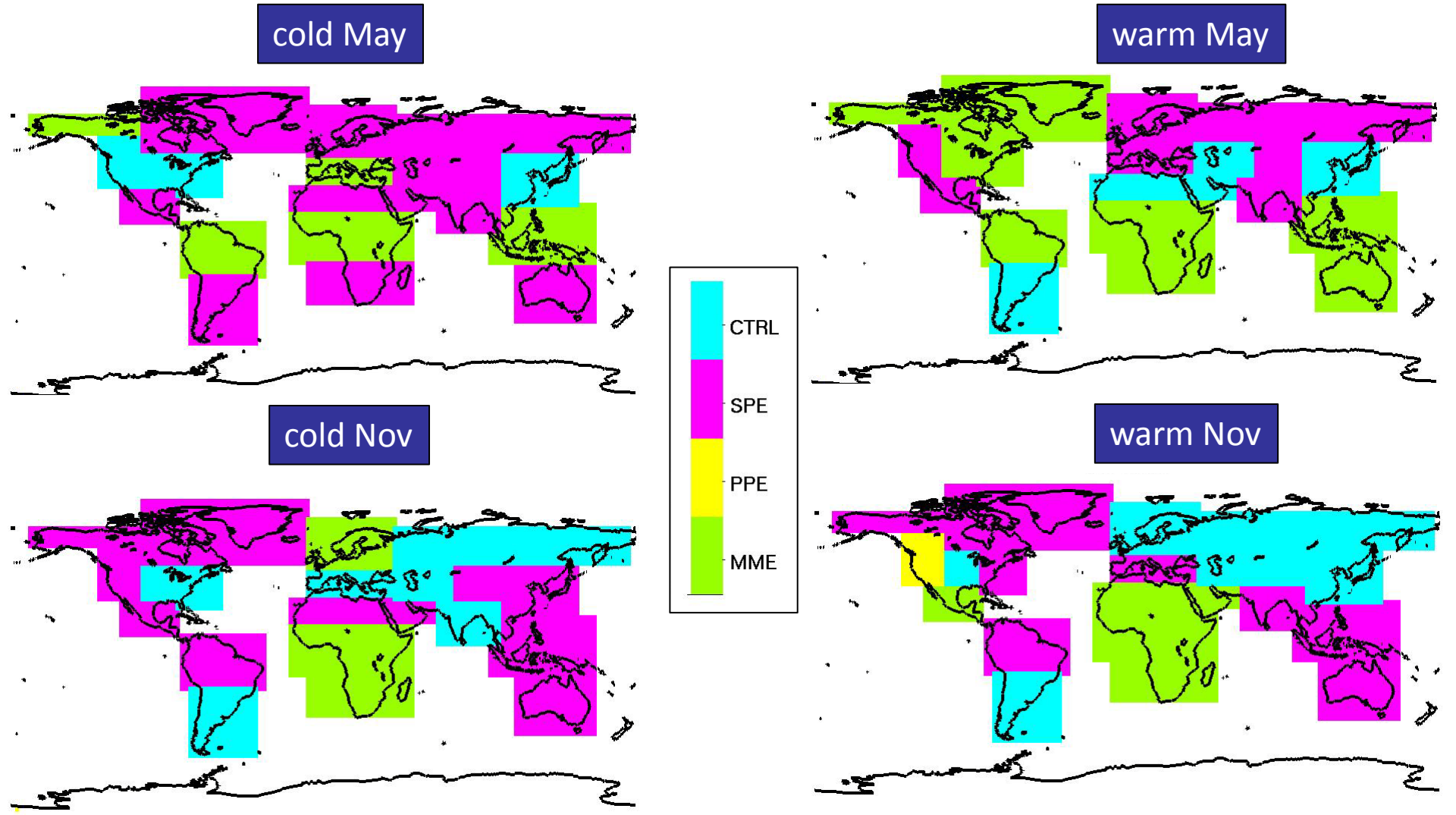
### May



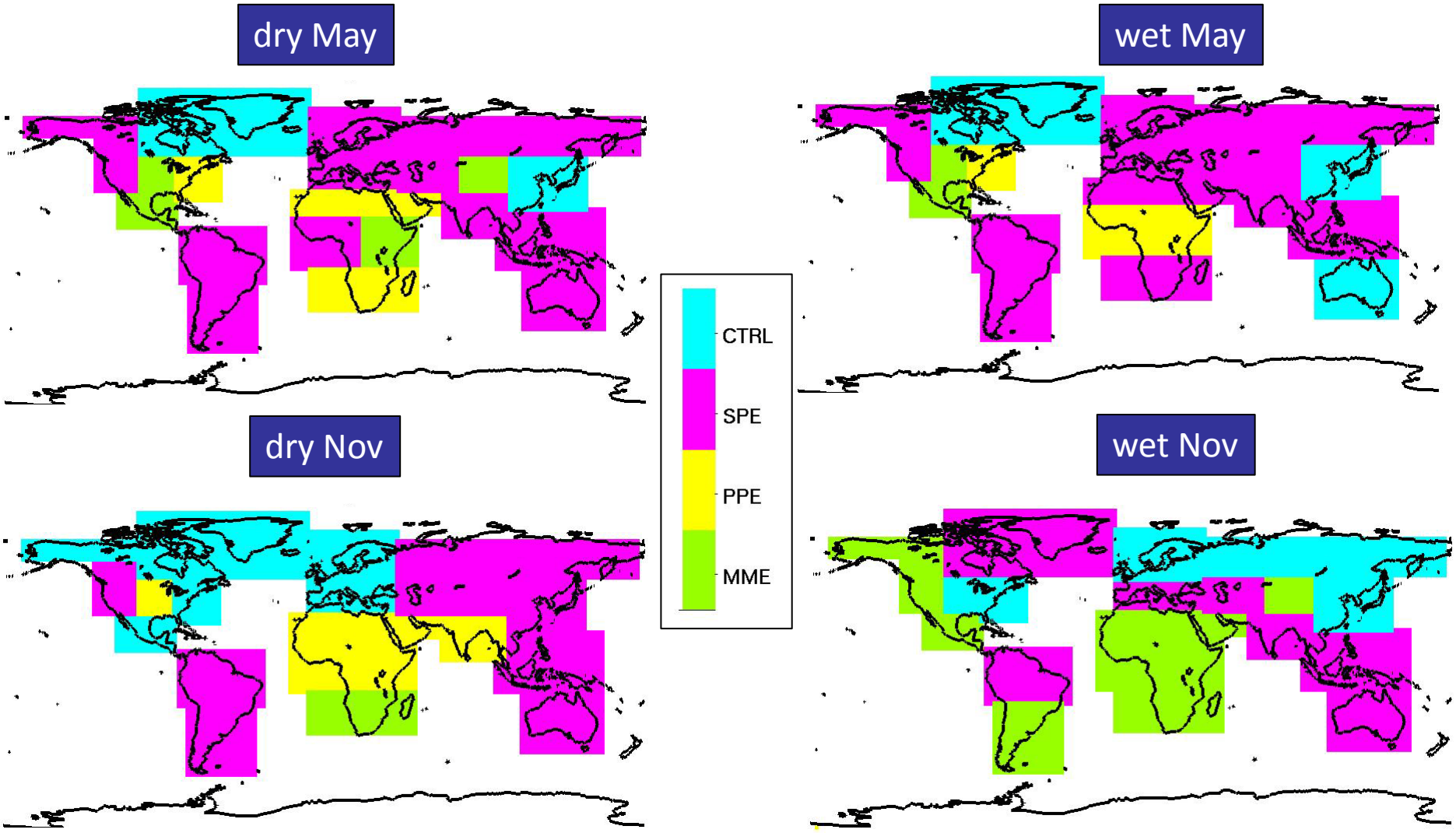

### JJA



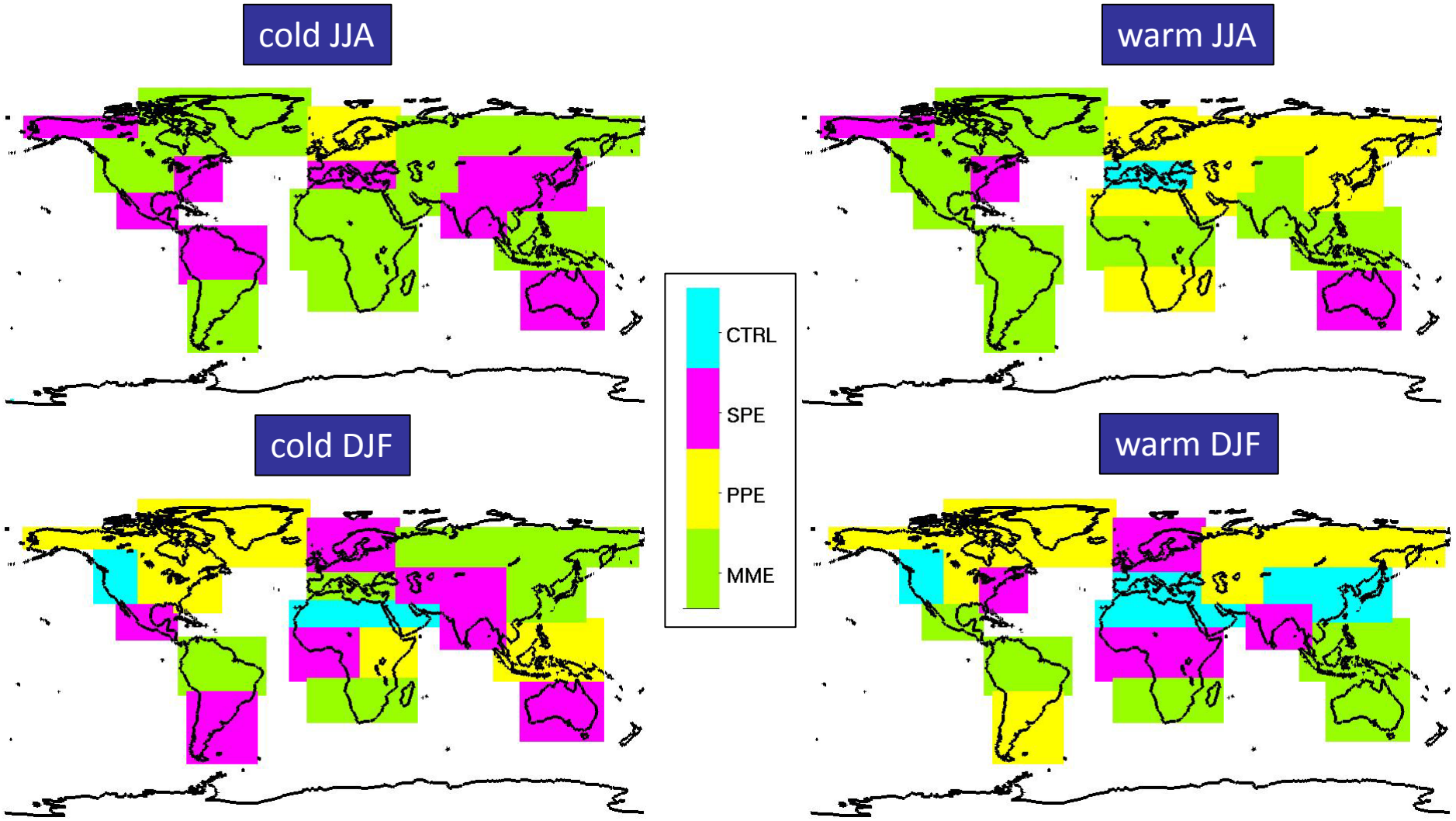
# temperature highest BSS ∞ 1st month



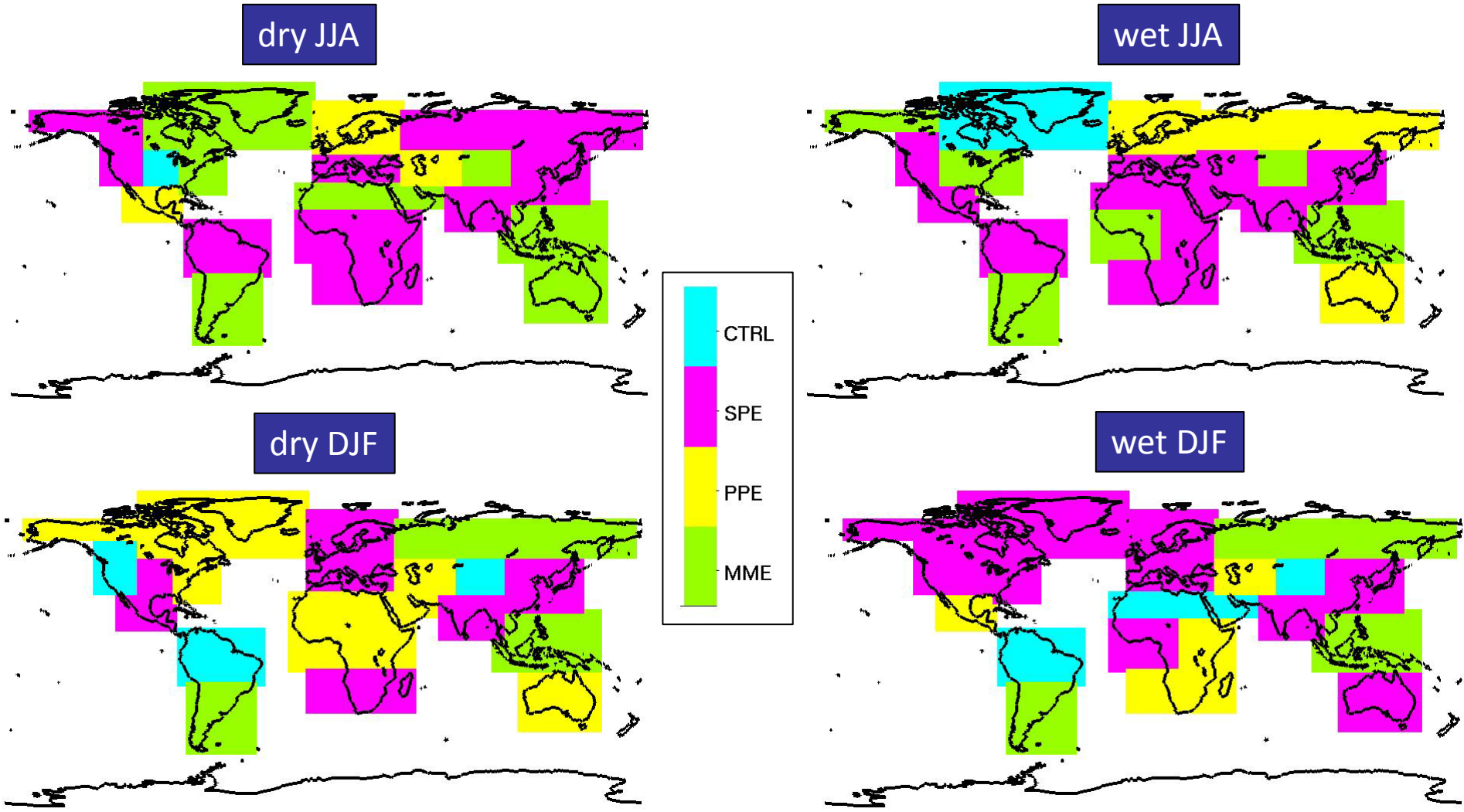
# precipitation highest BSS ∞ 1st month



# temperature highest BSS ∞ 2-4 months



# precipitation highest BSS ∞ 2-4 months

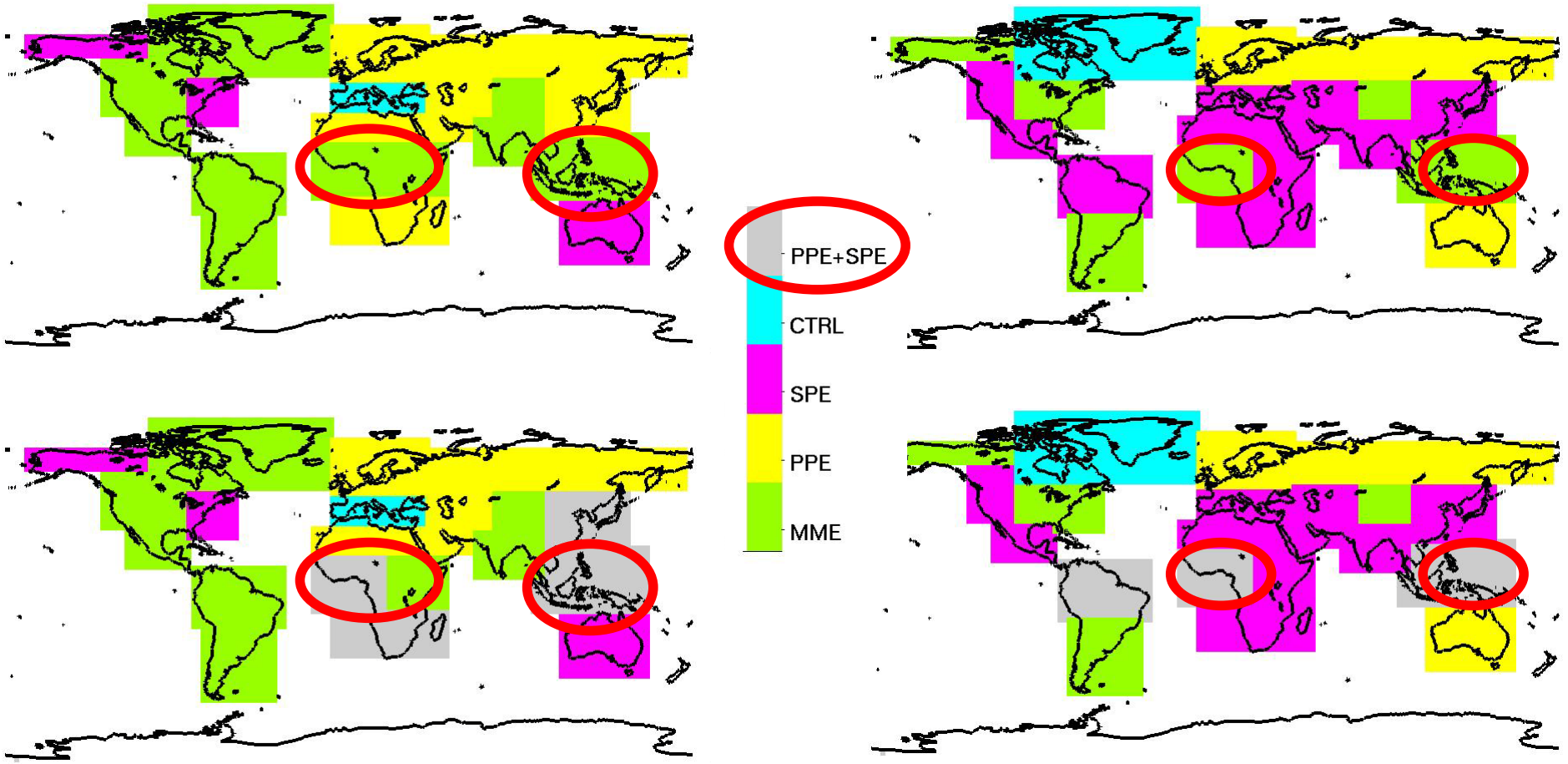


# combined PPE & SPE

2-4 months

warm JJA

wet JJA

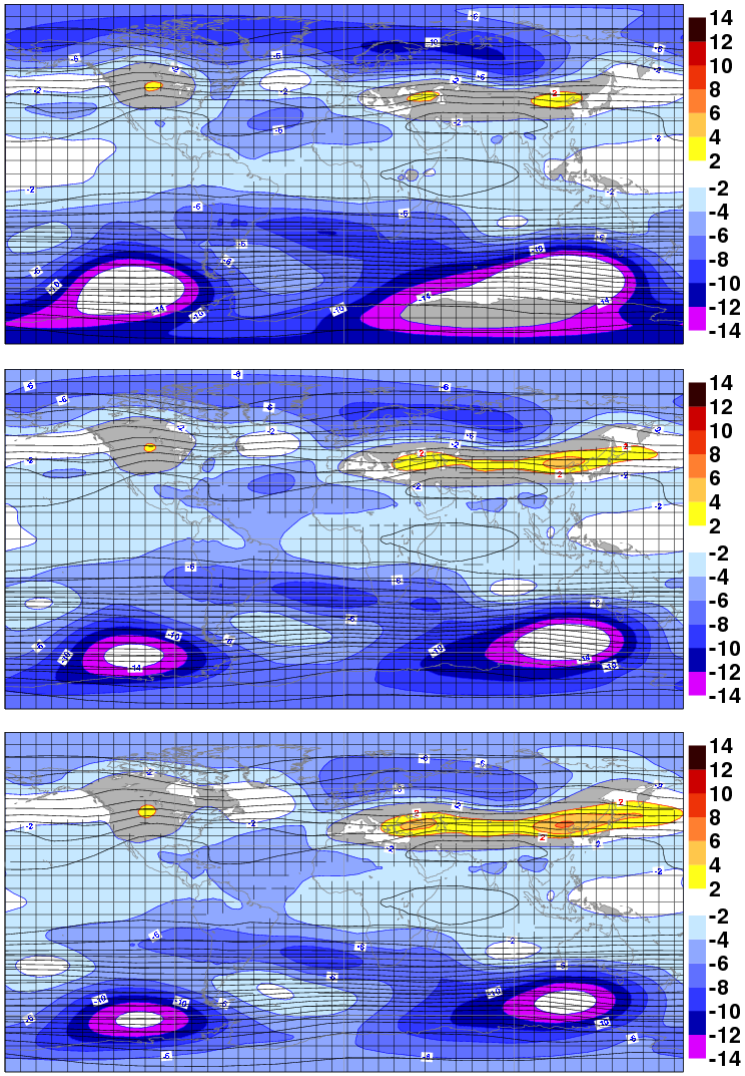


# Outlook: vorticity confinement

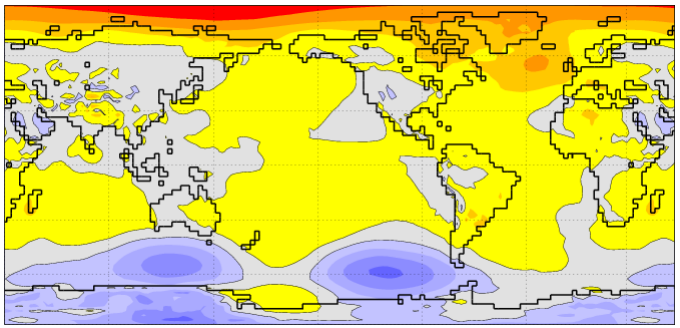
**Z200 in JJA**

**model bias**

**MSLP in JJA**

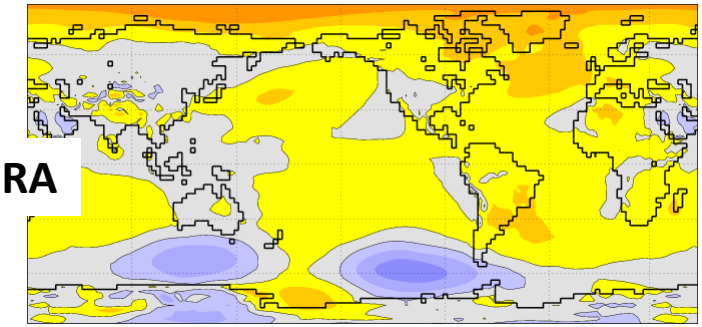



**Ctrl - ERA**



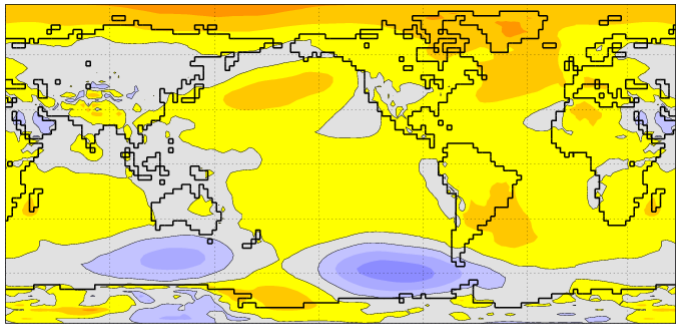



**determVC - ERA**



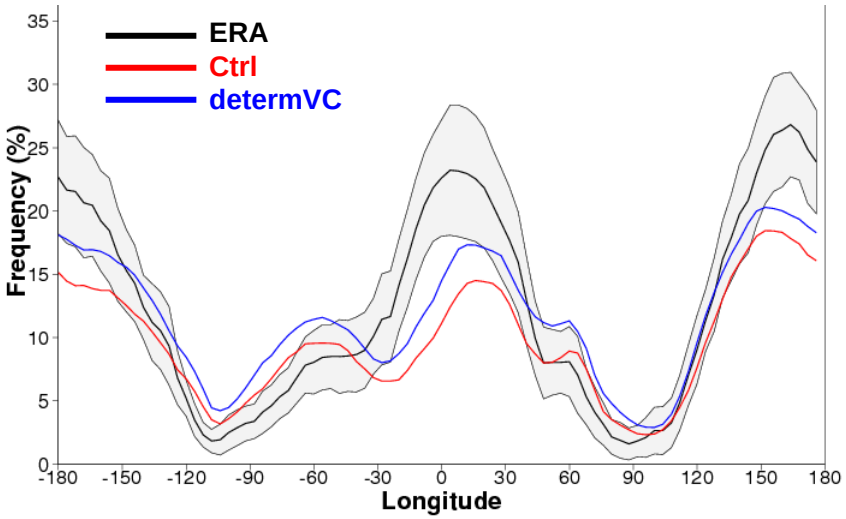


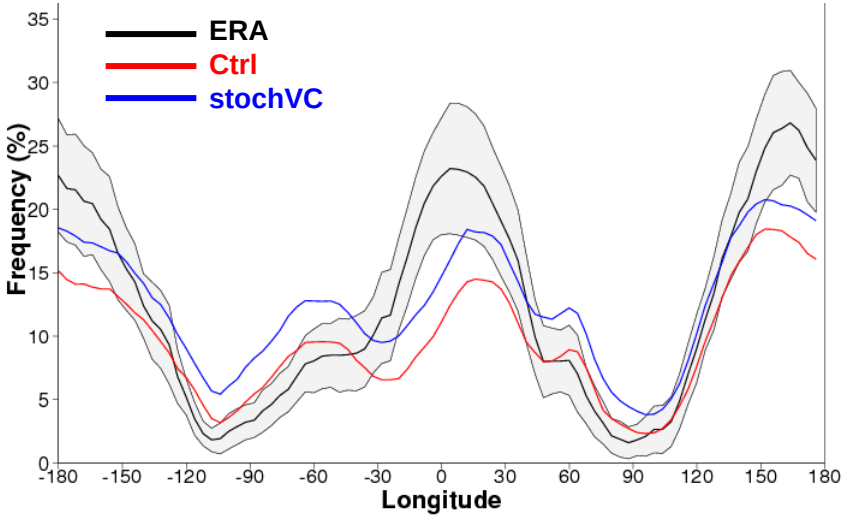

**stochVC - ERA**



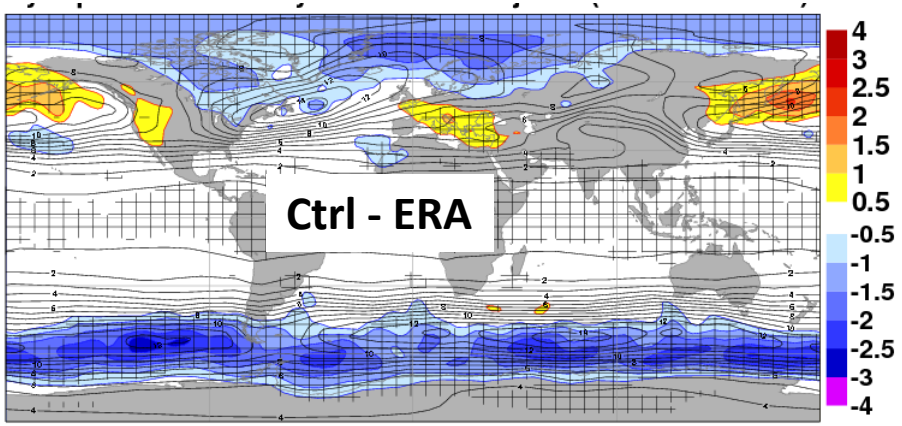
# Outlook: vorticity confinement

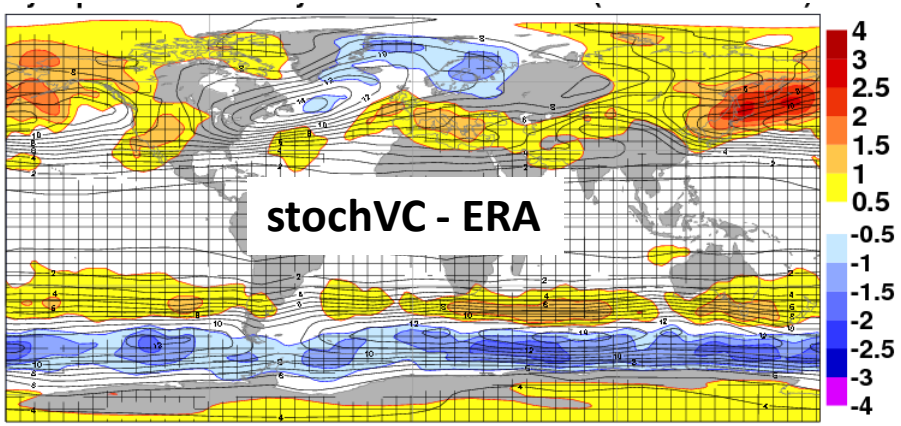
## blocking frequency




## DJF Z500 synoptic activity

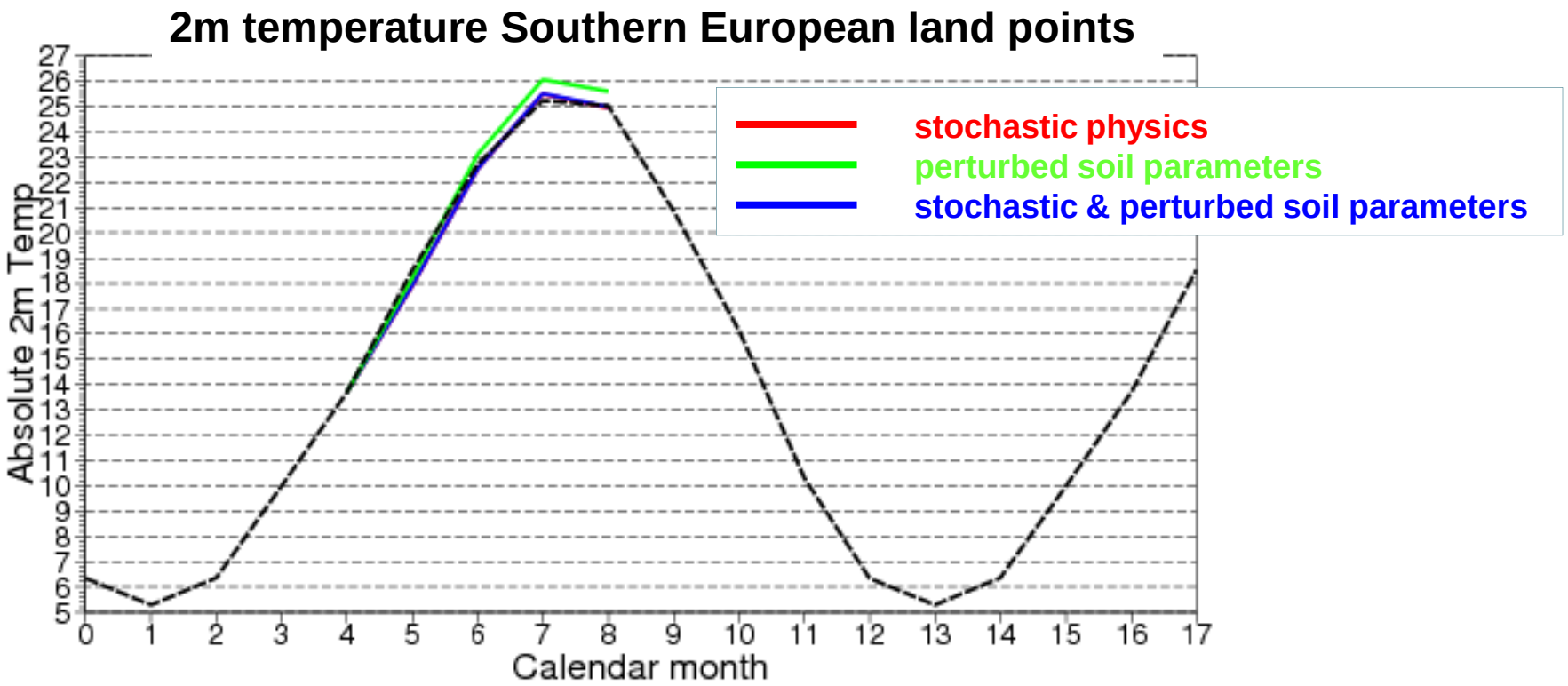


# Outlook: Land surface parameter uncertainty

**seasonal forecast ensemble:**
- **T159, 25 members**
- **1989-2008**
- **May start dates (→ JJA)**

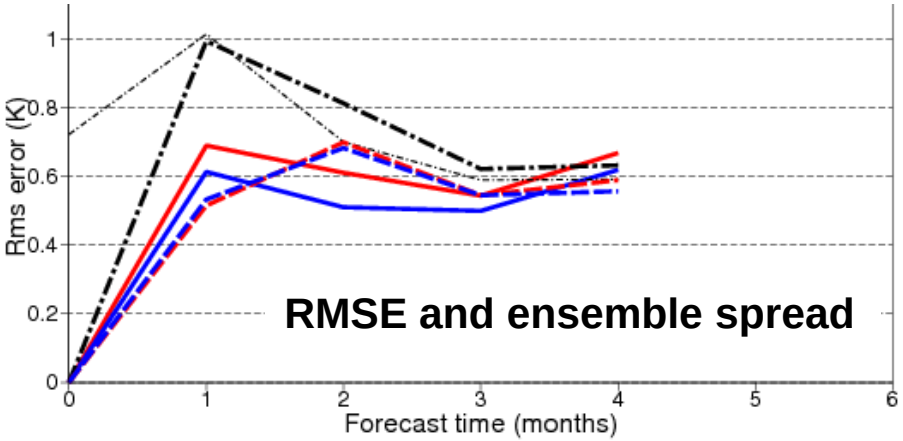
Perturbed land-surface parameters (+/- 20%):
- Hydraulic conductivity
- Curve shape parameter of soil moisture characteristic (van Genuchten $\alpha$)

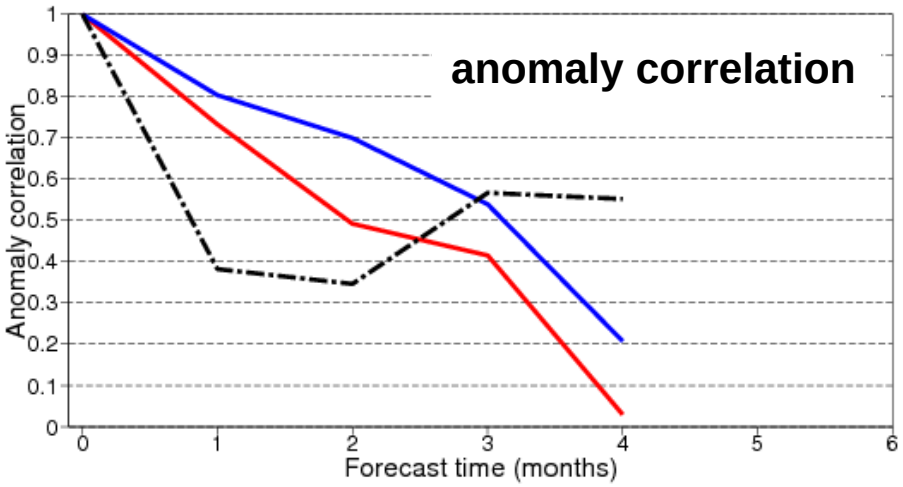


# Outlook: Land surface parameter uncertainty

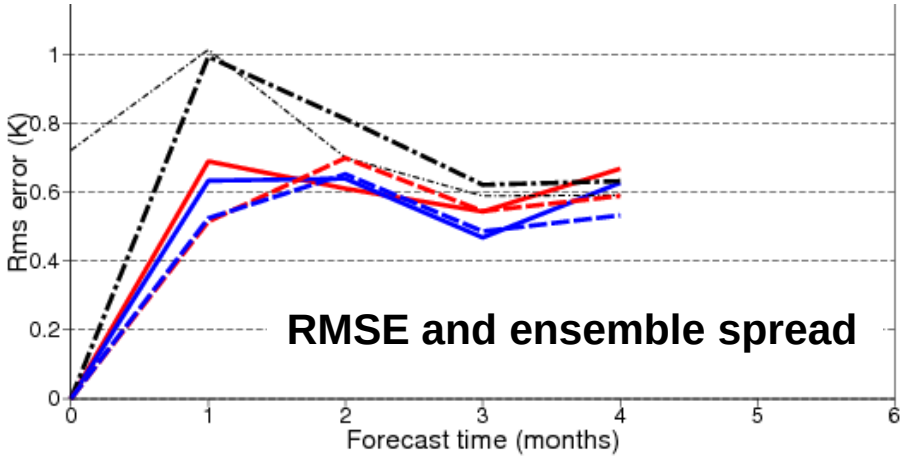
## 2m temperature Southern European land points

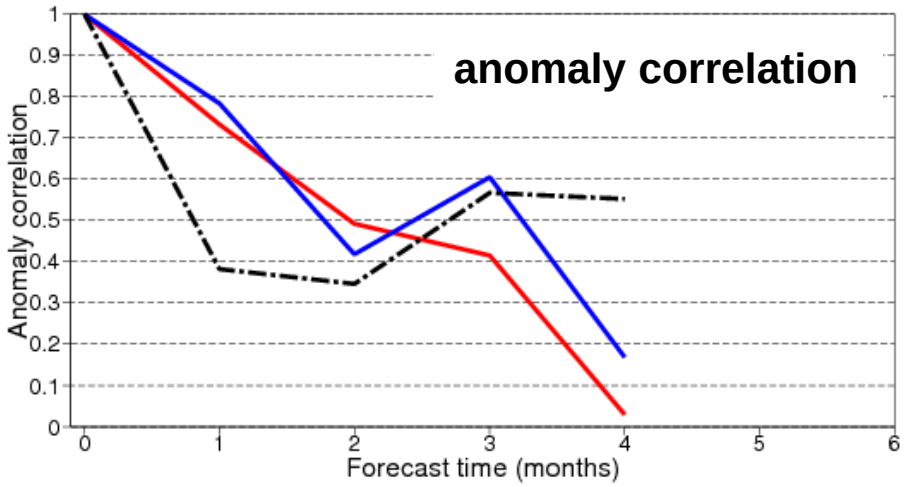
control (stoch phys only)
+ perturbed soil parameters




only stochastic physics
only perturbed soil parameters



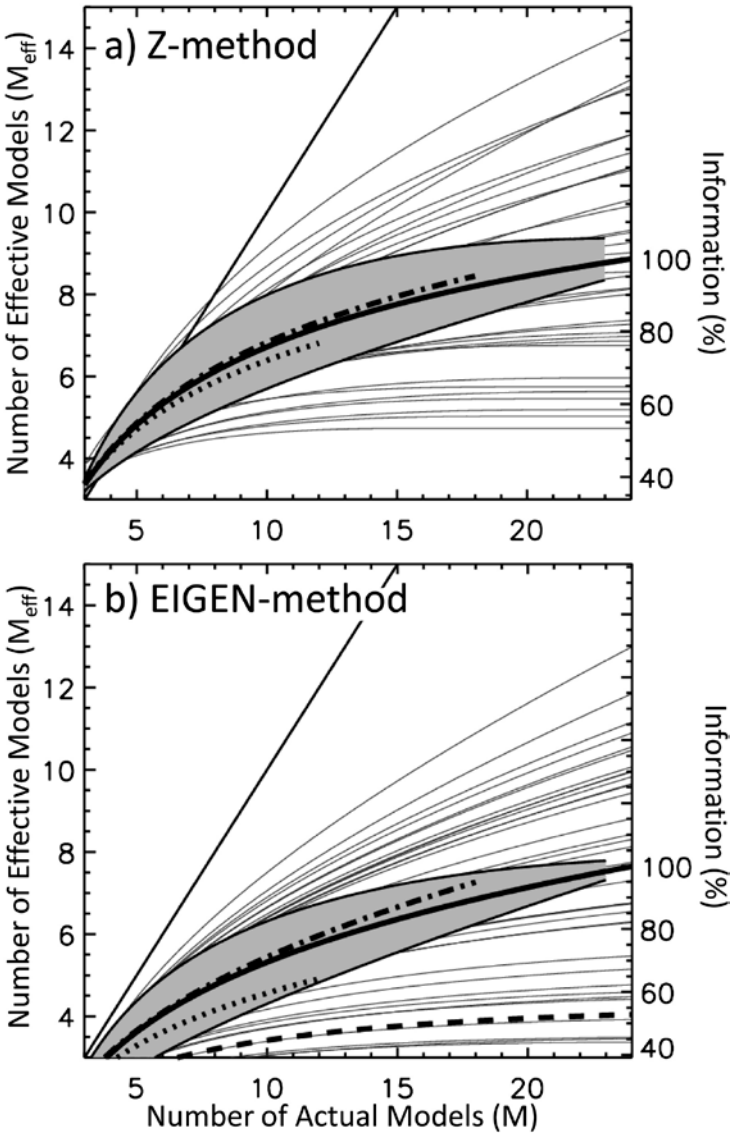
# Effective size of multi-model ensembles

24 CMIP-3 climate models for the 20th century, 35 quantities

$M_{eff}$ for individual quantities:
3 … 15

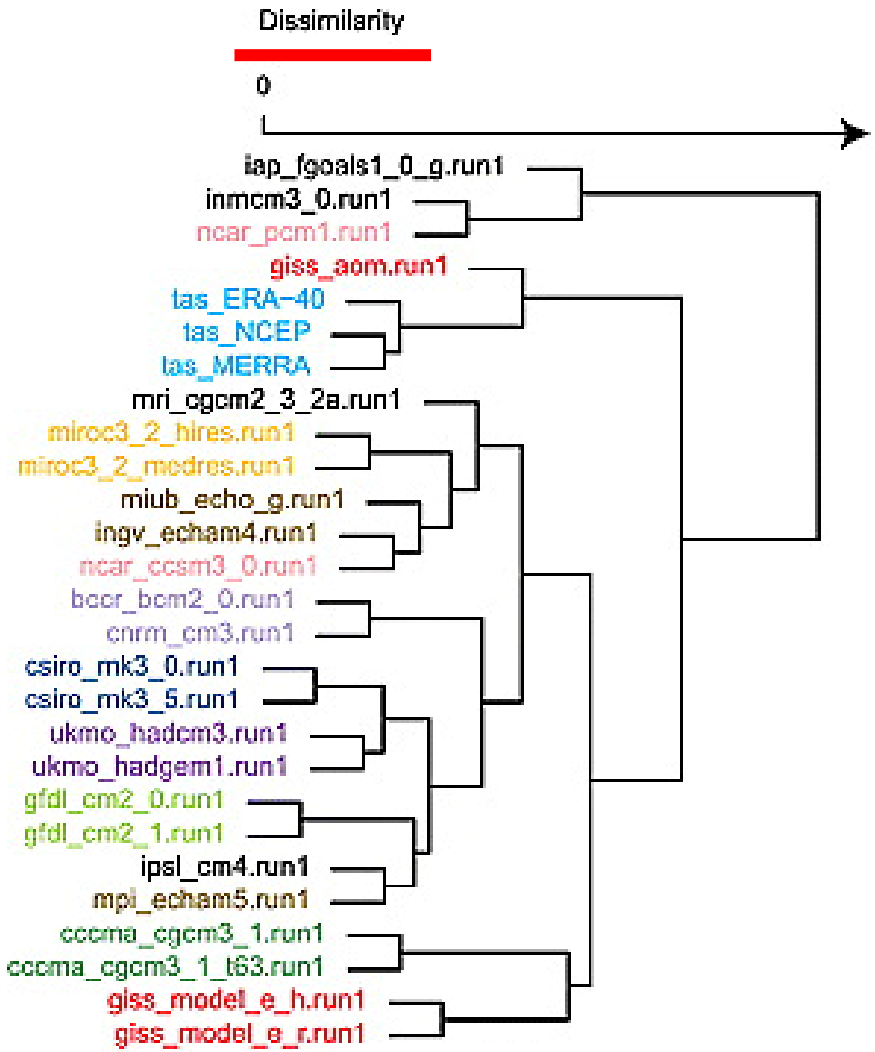
$M_{eff}$ averaged over all quantities:
7.5 … 9

Current interpretations of MME may lead to overly confident predictions

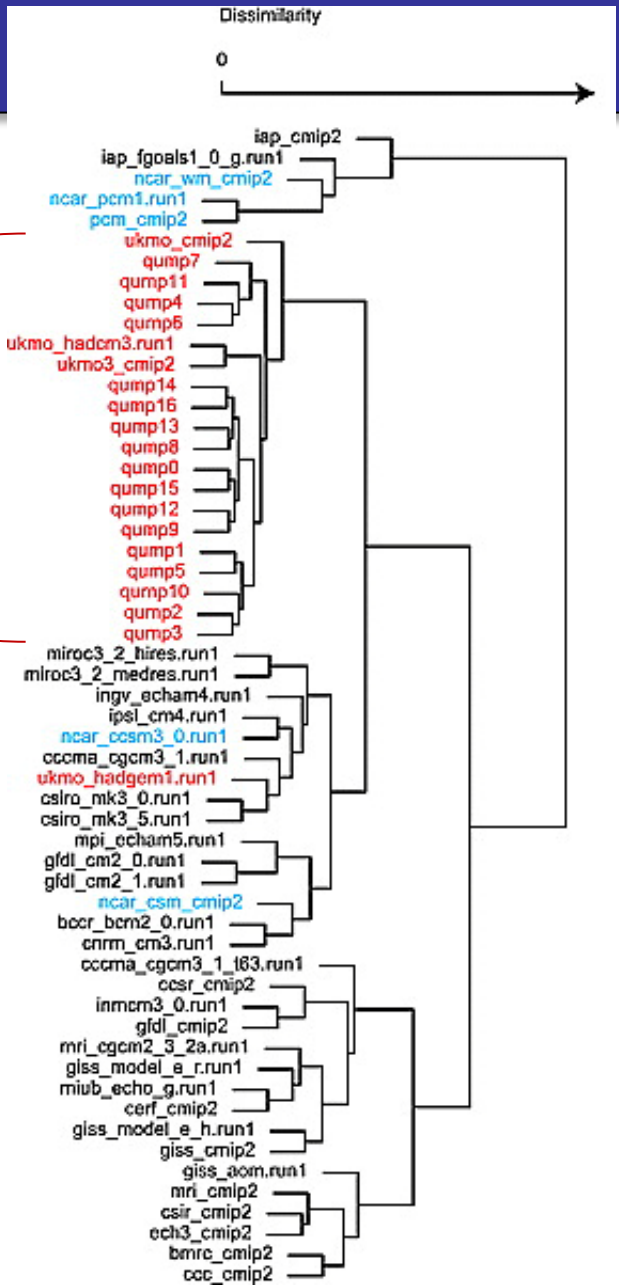



**Pennell and Reichler (2011)**

# Model similarity in MMEs and PPEs




PPE




**Masson and Knutti (2011)**

# Conclusions

- MME performs on average better than any single model ensemble due to reduced overconfidence
- **ENSO: MME very good, SPE improved over CTRL, PPE rather poor**
- **Monthly forecasts: SPE globally most skilful for most land temperature and precipitation events, regional variations**
- **Seasonal forecasts: MME (SPE) on average most skilful for temperature (precipitation) over land, regional variations**
- **PPE: no control ensemble available → quality of base model?**
- **SPE becomes competitive to MME and should also be included in uncertainty estimates of climate predictions/projections**
- **Combination of PPE and SPE has potential to improve skill further beyond the MME**
- Some promising test results with vorticity confinement (mean flow, blocking, synoptic activity)
- Preliminary test on including uncertainty in land surface parameter perturbations (improved forecast skill over southern Europe summer temperatures)

# References


- Doblas-Reyes, F.J., A. Weisheimer, M. Déqué, N. Keenlyside, M. McVean, J.M. Murphy, P. Rogel, D. Smith and T.N. Palmer (2009): Addressing model uncertainty in seasonal and annual dynamical seasonal forecasts. *Quart. J. R. Meteorol. Soc*, 135, 1538-1559.
- Doblas-Reyes, F.J., A. Weisheimer, T.N. Palmer, J.M. Murphy, and D. Smith (2010): Forecast quality assessment of the ENSEMBLES seasonal-to-decadal Stream 2 hindcasts. *ECMWF Tech. Memo.,* 621, 45pp.
- Ferro, C. (2007): Comparing probabilistic forecasting systems with the Brier Score. *Wea. Forecasting*, 22, 1076- 1088.
- Masson, D. and R. Knutti (2011): Climate model genealogy, *Geophys. Res. Lett.*, 38, L08703, doi:10.1029/2011GL046864.
- Müller, W.A., C. Appenzeller, F. Doblas-Reyes and M. A. Liniger, 2005: A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. *J. Climate*, 18, 1513-1523.
- Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer and A. Weisheimer (2009). Stochastic parametrization and model uncertainty. *ECMWF Tech. Memo.,* 598, 42pp.
- Pennel, C. and T. Reichler (2011): On the Effective Number of Climate Models, *J. Climate*, 24, 2358–2367.
- Weigel, A.P., M.A. Liniger and C. Appenzeller, 2007. The discrete Brier and ranked probability skill scores. *Mon. Weather Rev.,* 135, 118-124.
- Weisheimer, A., F.J. Doblas-Reyes, T.N. Palmer, A. Alessandri, A. Arribas, M. Deque, N. Keenlyside, M. MacVean, A. Navarra and P. Rogel (2009). ENSEMBLES - a new multi-model ensemble for seasonal-to-annual predictions: Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.,* 36, L21711, doi:10.1029/2009GL040896.
- Weisheimer, A., F. Doblas-Reyes and T. Palmer (2010): Model uncertainty in seasonal to decadal forecasting – insight from the ENSEMBLES project. *ECMWF Newsletter*, 122, 21-26.
- Weisheimer, A., T.N. Palmer and F. Doblas-Reyes (2011): Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.,* submitted.