Representing model uncertainty on seasonal time scales: ENSEMBLES

F. J. Doblas-Reyes

Catalan Institute of Climate Sciences (IC3), Barcelona, Spain

with contributions from, among others, A. Weisheimer, T. Palmer, M. Balmaseda, G. Shutts, M. Leutbecher (ECMWF), J. Murphy, D. Smith, M. McVean (Met Office), and the seasonal contributors to the ENSEMBLES multi-model

The ENSEMBLES project

- ENSEMBLES was a EU-funded project active from 2004 to 2009, with ~70 partners, that worked on the development of an EPS across space and time scales.
- The main objective was to carry out global and regional dynamical simulations across time scales (seasonal, decadal and longer), with the aim of providing high-resolution climate information to specific applications.
- An innovative and comprehensive seasonal-to-decadal (s2d) experiment was one of the main deliverables.

ENSEMBLES s2d experiment

- Model uncertainty is a major source of forecast error. Three approaches to deal with the impact of model uncertainty on forecast error were investigated: multi-model (ECMWF, GloSea, DePreSys, Météo-France, IfM-Kiel, CERFACS, INGV), stochastic physics (ECMWF) and perturbed parameters (DePreSys_PP).
- Hindcasts in two streams:
  - Stream 1: hindcast period 1991-2001, seasonal (7 months, May and November start date), annual (14 months, November start date), 9-member ensembles, ERA40 initialization in most cases; DePreSys (IC and PP ensembles) 10-year runs in every instance.
  - Stream 2: As in Stream 1 but over 1960-2005, with 4 start dates for seasonal hindcasts (Feb, May, Aug and Nov start dates), at least 1 for annual (Nov start date) and at least one 3-member decadal hindcast every 5 years; DePreSys_PP 10-year runs once a year.

ENSEMBLES s2d experiment

partner	Atmospheric model; resolution	Ocean model; resolution	initialization		external forcing	Additional components, comments	references
			atmosphere and land	ocean			
ECMWF	IFS CY31R1; T159/L62	HOPE; 0.3º-1.4º/L29	ERA-40/oper. analysis, atmospheric singular vectors	wind stress perturbations to generate ensemble of ocean reanalyses; SST perturbations at initial time	observed global well-mixed GHGs and sulphate aerosol and A1B from 2000, observed solar activity, no volcanic aerosol nor ozone	Operational Seasonal Forecasting system S3	*Stockdale et al.* (2010)*; Balmaseda et al.* (2008)
	IFS CY33R1; T159/L62	HOPE; 0.3º-1.4º/L29	ERA-40/oper. analysis, atmospheric singular vectors	wind stress perturbations to generate ensemble of ocean reanalyses; SST perturbations at initial time	“	used for the decadal hindcasts only	*Bechtold et al.* (*2008*)
	IFS CY35R3; T159/L62	HOPE; 0.3º-1.4º/L29	ERA-40/oper. analysis, atmospheric singular vectors	wind stress perturbations to generate ensemble of ocean reanalyses; SST perturbations at initial time	“	used with the stochastic physics approach	*Palmer et al.* (*2009*)
UKMO	HadGEM2-A; N96/L38	HadGEM2-O; 0.33º-1º/L20	ERA-40/oper. analysis, anomaly assimilation for soil moisture	wind stress perturbations to generate ensemble of ocean reanalyses; SST perturbations at initial time	observed global well-mixed GHGs, ozone and sulphate aerosol emissions and A1B from 2000, persisted solar activity and volcanic aerosol	fully interactive sea ice module	*Collins et al.* (2008)
	HadAM3; 3.75x2.5°	HadOM	anomaly assimilation of ERA-40/oper. analysis	anomaly assimilation of an ocean reanalysis	“	perturbed-parameter ensemble	*Smith et al.* (2007)
MF	ARPEGE4.6; T63	OPA8.2; 2°/L31	ERA-40/oper. analysis	wind stress, SST and water flux perturbations to generate ensemble of ocean reanalyses	observed global well-mixed GHGs and sulphate aerosol and A1B from 2000, no solar activity nor volcanic aerosol, dynamical ozone	GELATO sea ice model	*Daget et al.* (2009)*; Salas y Melia* (2002)
IFM-GEOMAR	ECHAM5; T63/L31	MPI-OM1; 1.5º/L40	initial condition permutations of three coupled climate simulations from 1950 to 2005 with SSTs restored to observations		observed global well-mixed GHGs, ozone and sulphate aerosol emissions and A1B from 2000, persisted solar activity and volcanic aerosol	-	*Keenlyside et al.* (2005)*; Jungclaus et al.* (2006)
CMCC-INGV	ECHAM5; T63/L19	OPA8.2; 2°/L31	AMIP-type simulations with forced SSTs	wind stress perturbations to generate ensemble of ocean reanalyses, SST perturbations at initial time	observed global well-mixed GHGs and sulphate aerosol and A1B from 2000, no volcanic aerosol nor ozone	dynamical snow-sea ice model and land-surface model	*Weisheimer et al.* (2009); *Alessandri et al.* (2010)

Error and accuracy

- Some definitions:
  - Forecast system: the forecast model (e.g., a coupled ocean/atmosphere model), the initialization (and ensemble generation) method and the statistical model that creates the forecast probabilities (an ensemble is not a probability forecast).
  - Forecast quality: Statistical description of how well the forecasts match the observations; it has multiple attributes.
- Important in probability forecast systems:
  - Systematic error: difference in the statistical properties of the forecast and the reference distribution functions.
  - Accuracy: some sort of distance between forecasts and observations. It is used as an indication of the association, the discrimination or resolution of the forecasts.
  - Reliability: A measure of trustworthiness, not accuracy, that is another component of the systematic error of the forecast system.

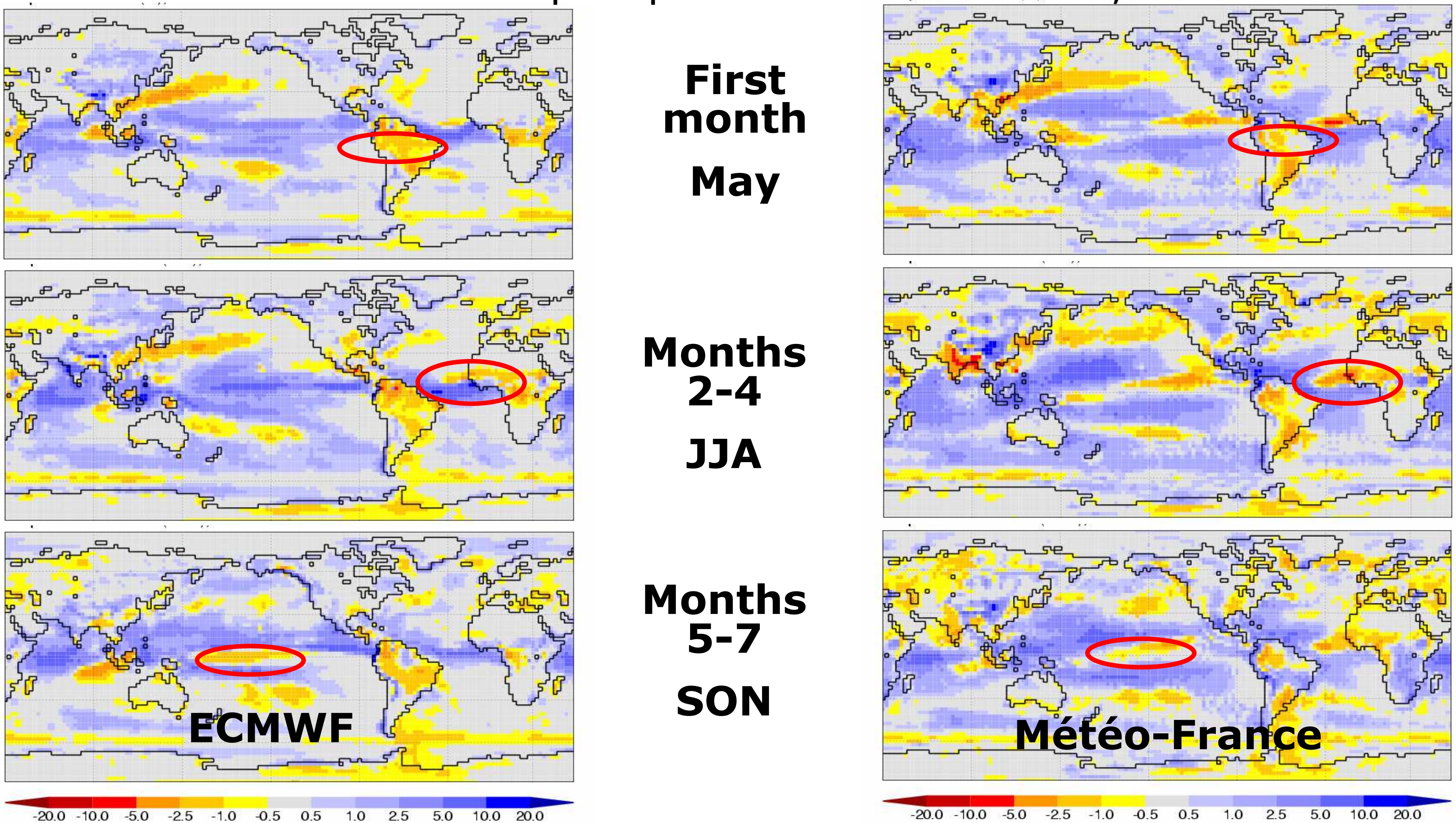
Dealing with systematic errors

- Model drift is comparable in size to the predicted signal
  - Both in ocean and atmosphere fields.
- Predictions are made *relative* to past model integrations
  - Anomalies computed in one-year-out cross-validation.
  - Model climate estimated from all available years and all ensemble members, performed separately for each single-model or model version.
  - Model climate is a function of start date and lead time.
- Implicit assumption of linearity
- Main systematic errors in dynamical climate forecasts:
  - Differences between the model climatological pdf (computed for a lead time from all start dates and ensemble members) and the reference climatological pdf (for the corresponding times of the reference dataset): systematic errors in mean and variability.
  - Conditional biases in the forecast pdf: errors in conditional probabilities implying that probability forecasts are not trustworthy. This type of systematic error is best assessed using the reliability diagram.

INSTITUT CATALÀ DE CIÈNCIES DEL CLIMA

Systematic errors in seasonal forecasts

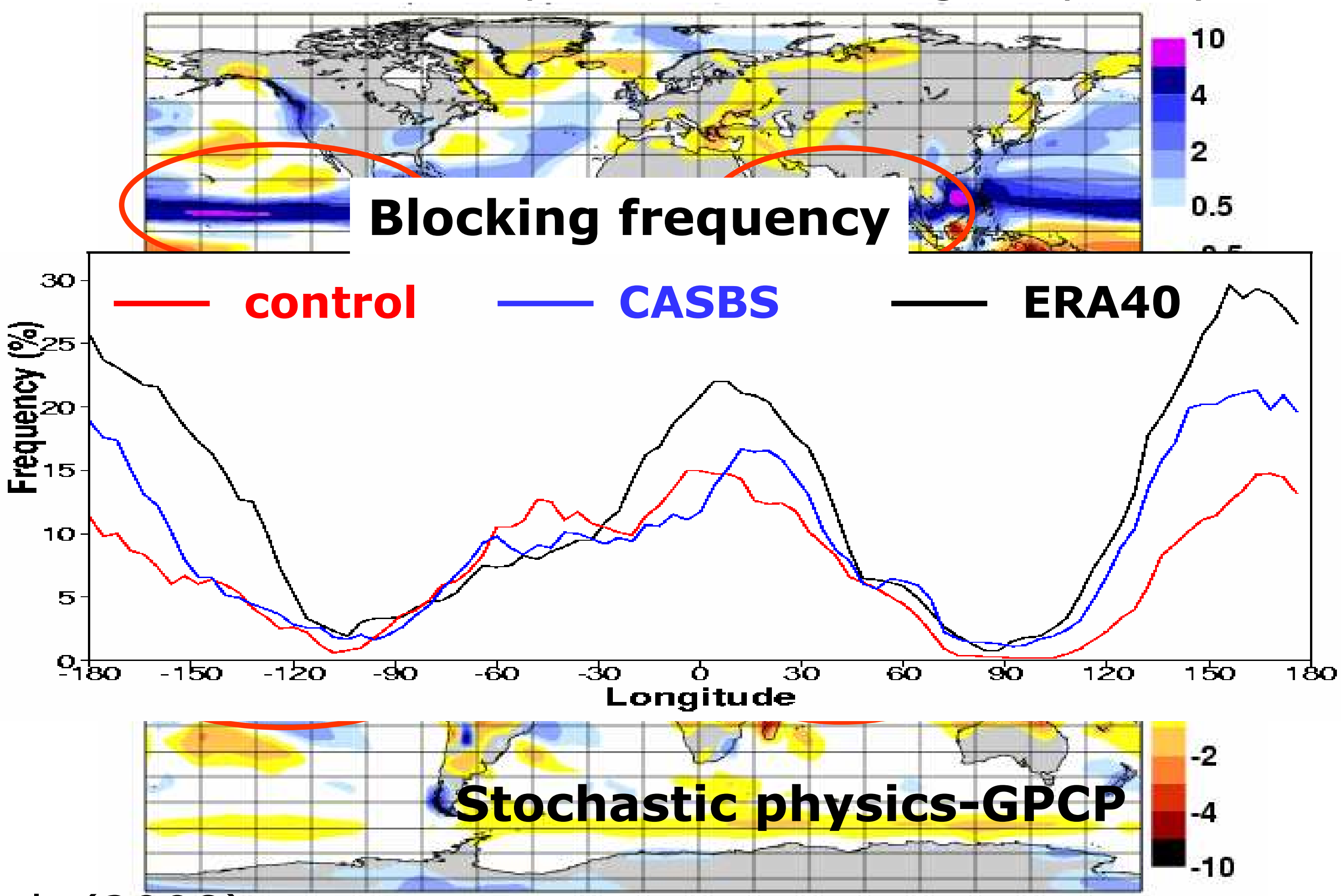
ENSEMBLES Stream 2 precip. mean bias wrt GPCP, 1979-2005



Error reduction: stochastic physics

Precipitation bias (DJF, 1-month lead, 1991-2001, CY29R2)

CASBS reduces the tropical and blocking frequency biases

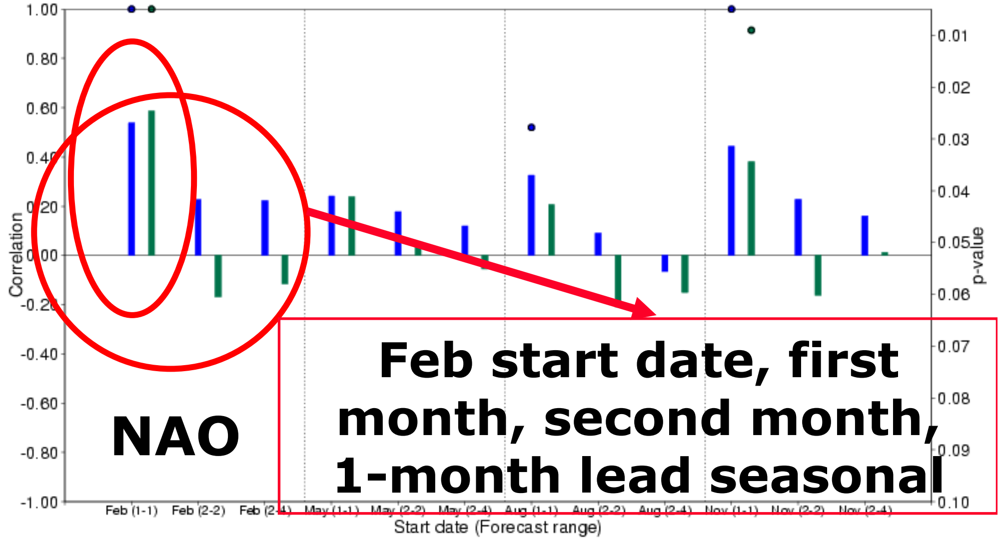


Berner et al. (2008)

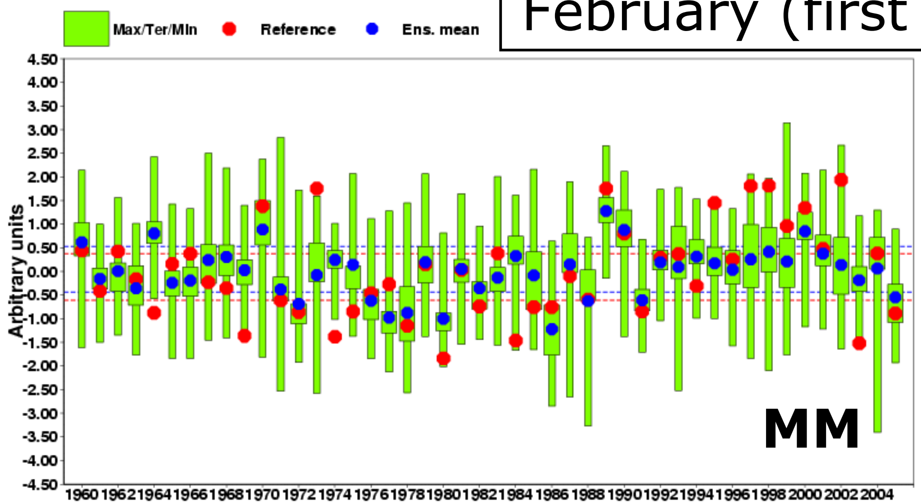
MM vs PP ensembles

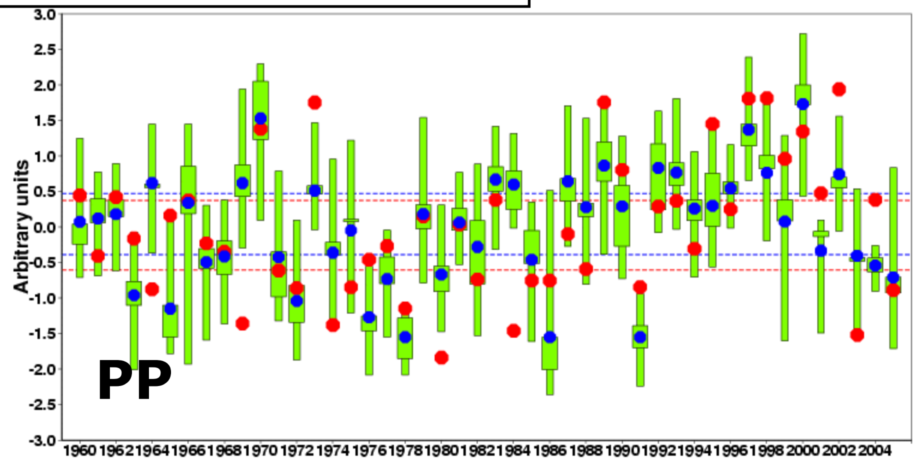
Ensemble-mean anomaly correlation (1960-2005): **MM** (45) **PP** (9)

Niño3.4



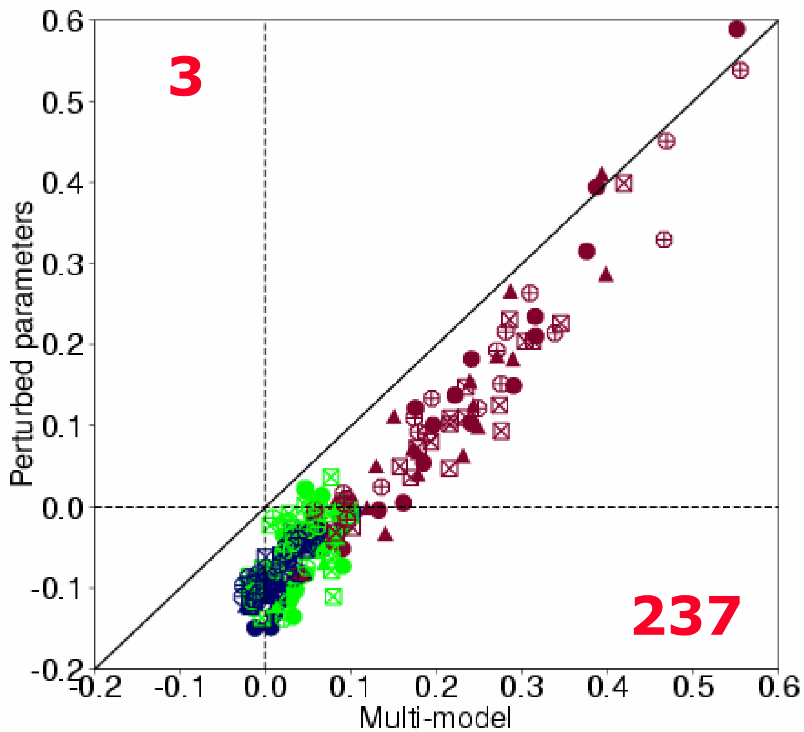
February (first month) NAO predictions

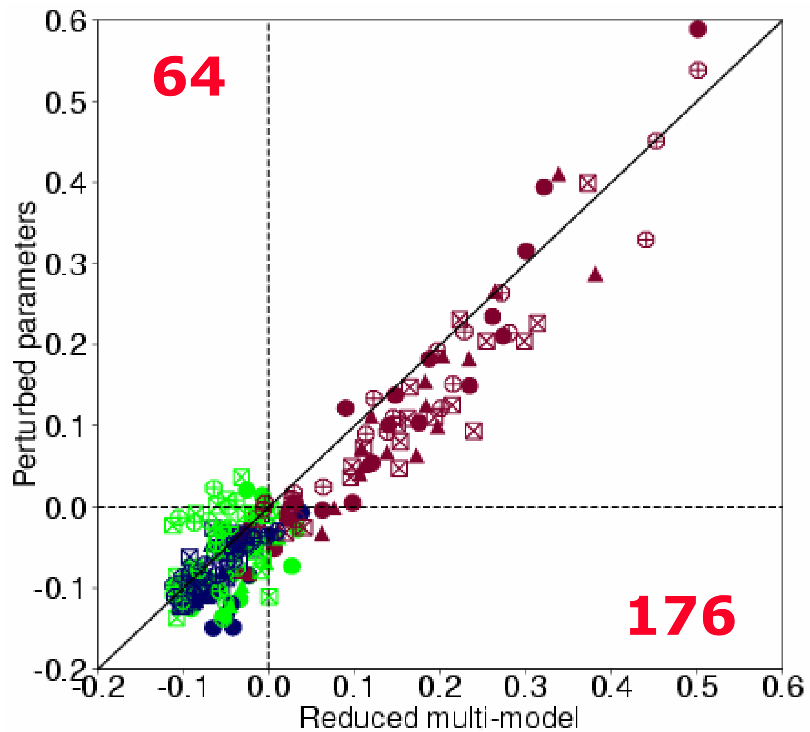




Probabilistic predictions

Brier skill score for several regions (Northern Hemisphere, Tropics, Southern Hemisphere), events (anomalies above/below the upper/lower tercile), lead times (2-4, 5-7 months), start dates (Feb, May, Aug and Nov) and variables (near-surface temperature, precipitation, Z500, T850 and MSLP) computed over the period 1960-2005. The inset numbers indicate the number of cases where a system is superior.

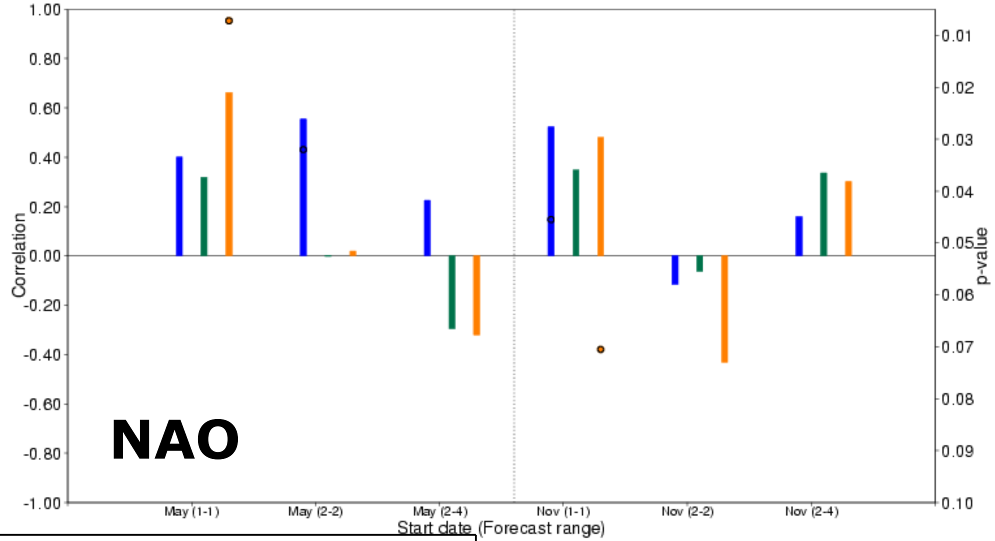




MM vs PP vs SP ensembles

Ensemble-mean anomaly correlation (1991-2005): **MM** (45) **PP** (9) **SP** (9)

Niño3.4

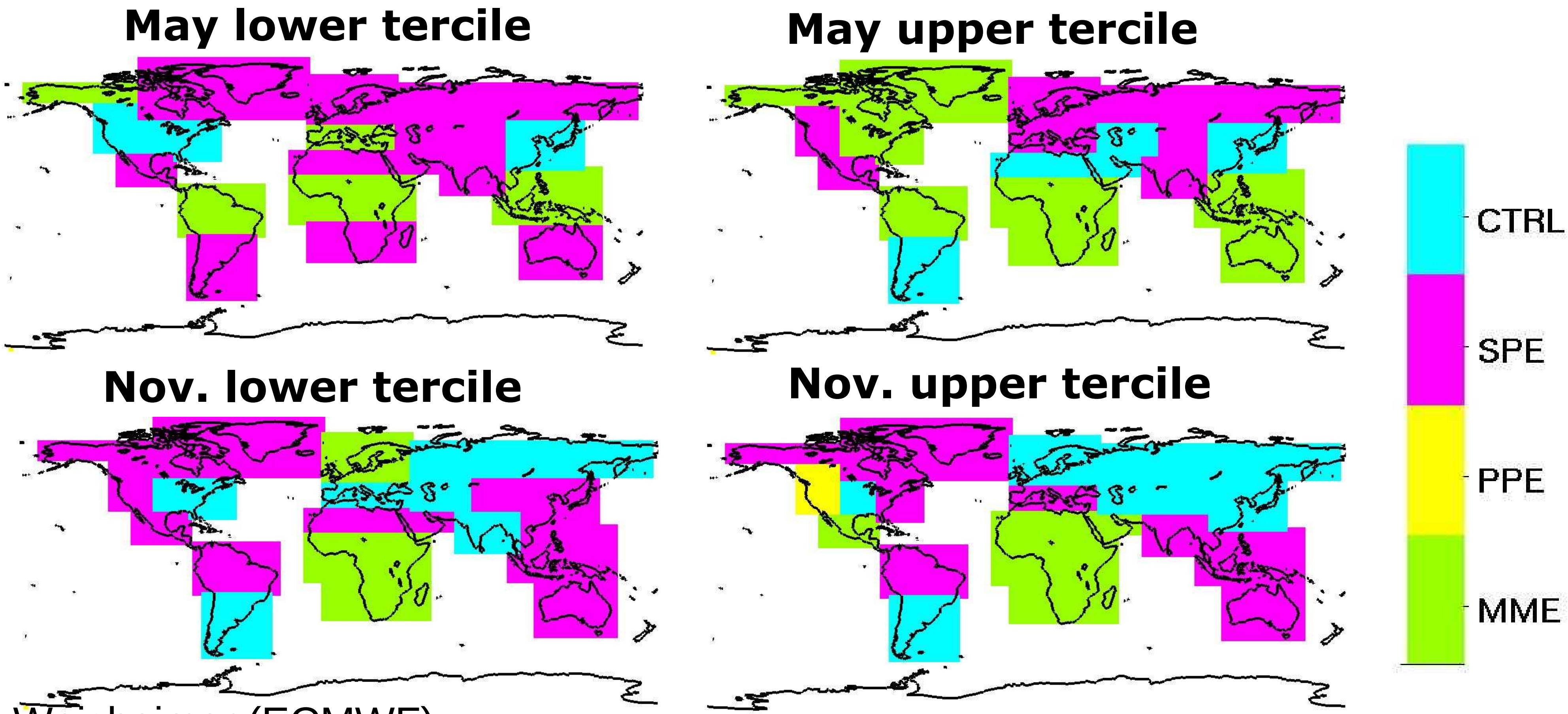


NAO

PNA

Method comparison: first month

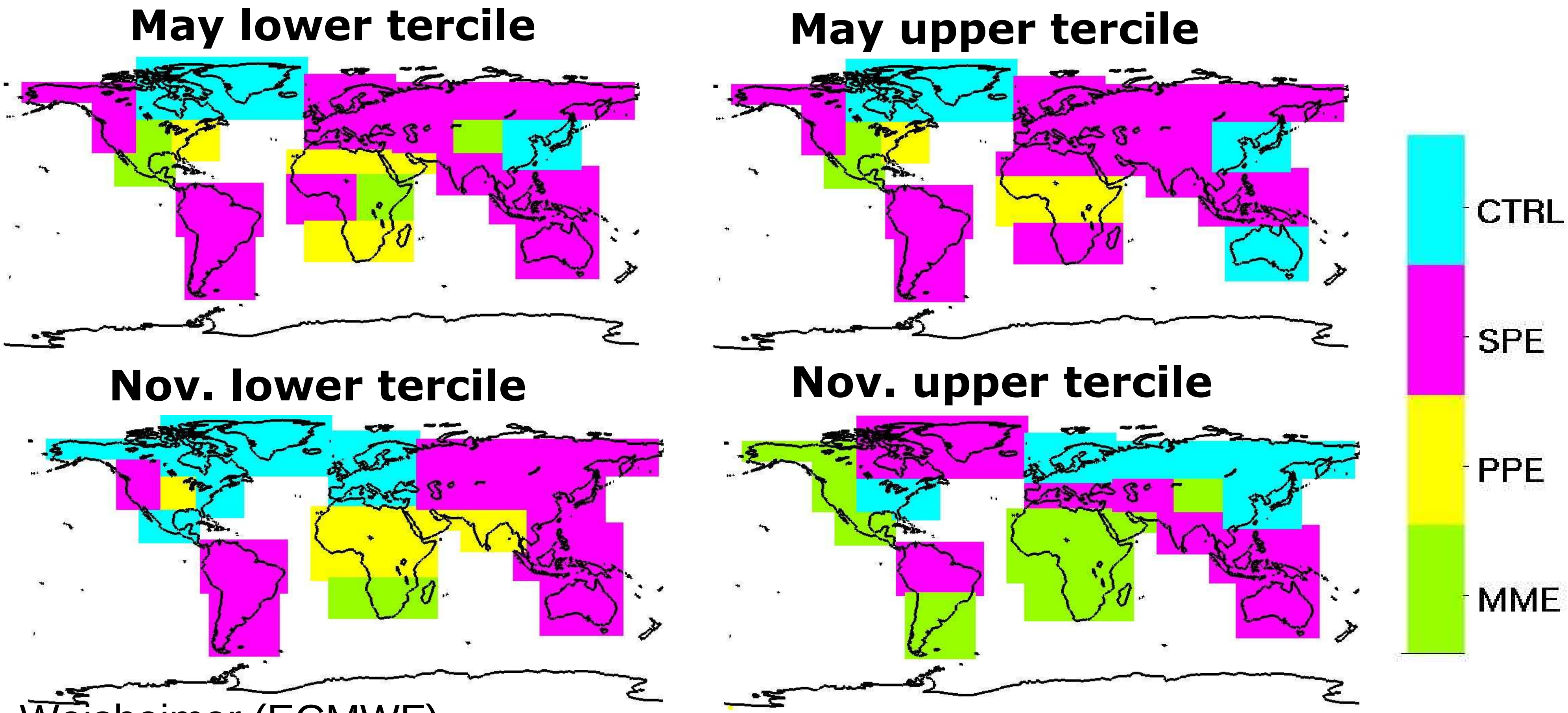
Forecast system with the largest BSS(∞) for temperature tercile events wrt ERA40/ERAInt of the first month predictions, 1991-2005



A. Weisheimer (ECMWF)

Method comparison: first month

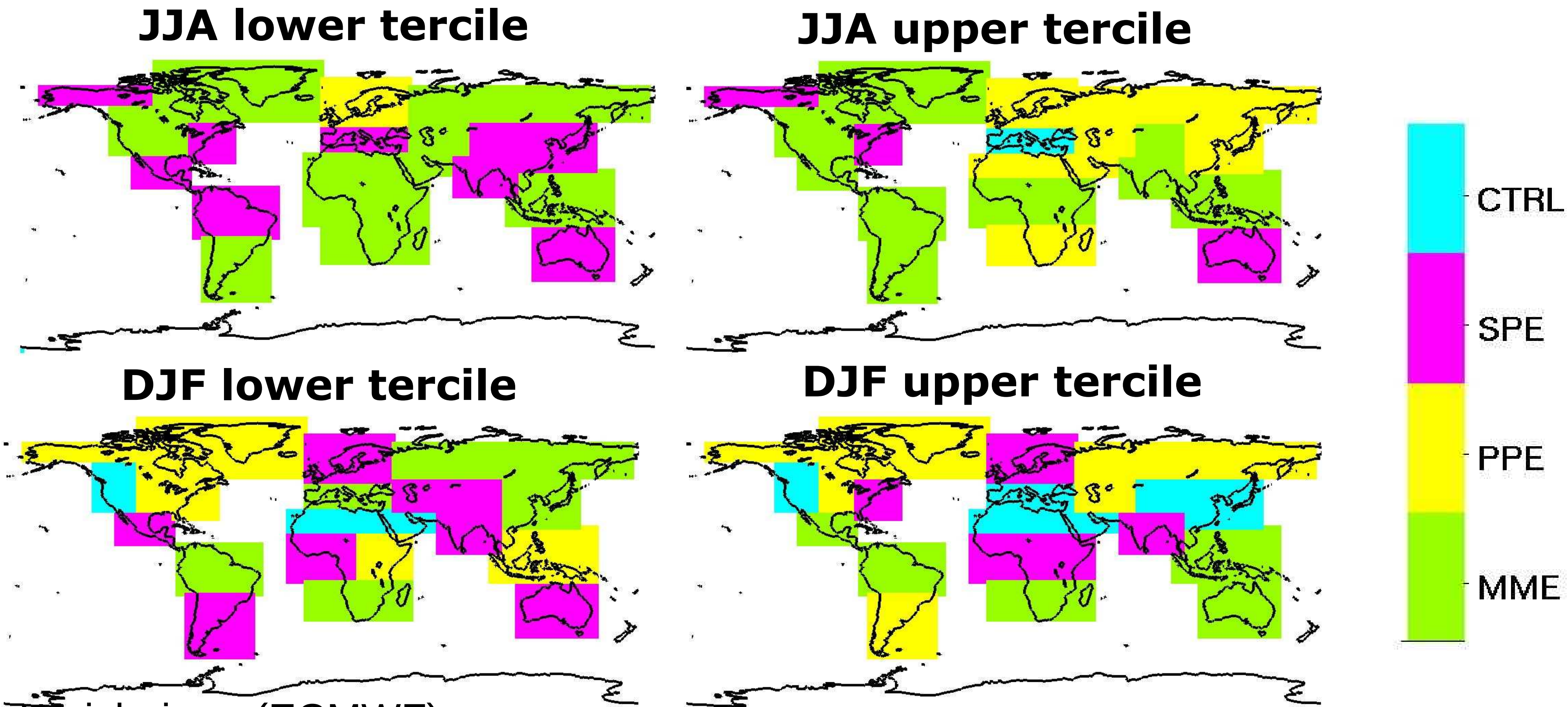
Forecast system with the largest BSS(∞) for precipitation tercile events wrt GPCP of the first month predictions, 1991-2005



A. Weisheimer (ECMWF)

Method comparison: seasonal

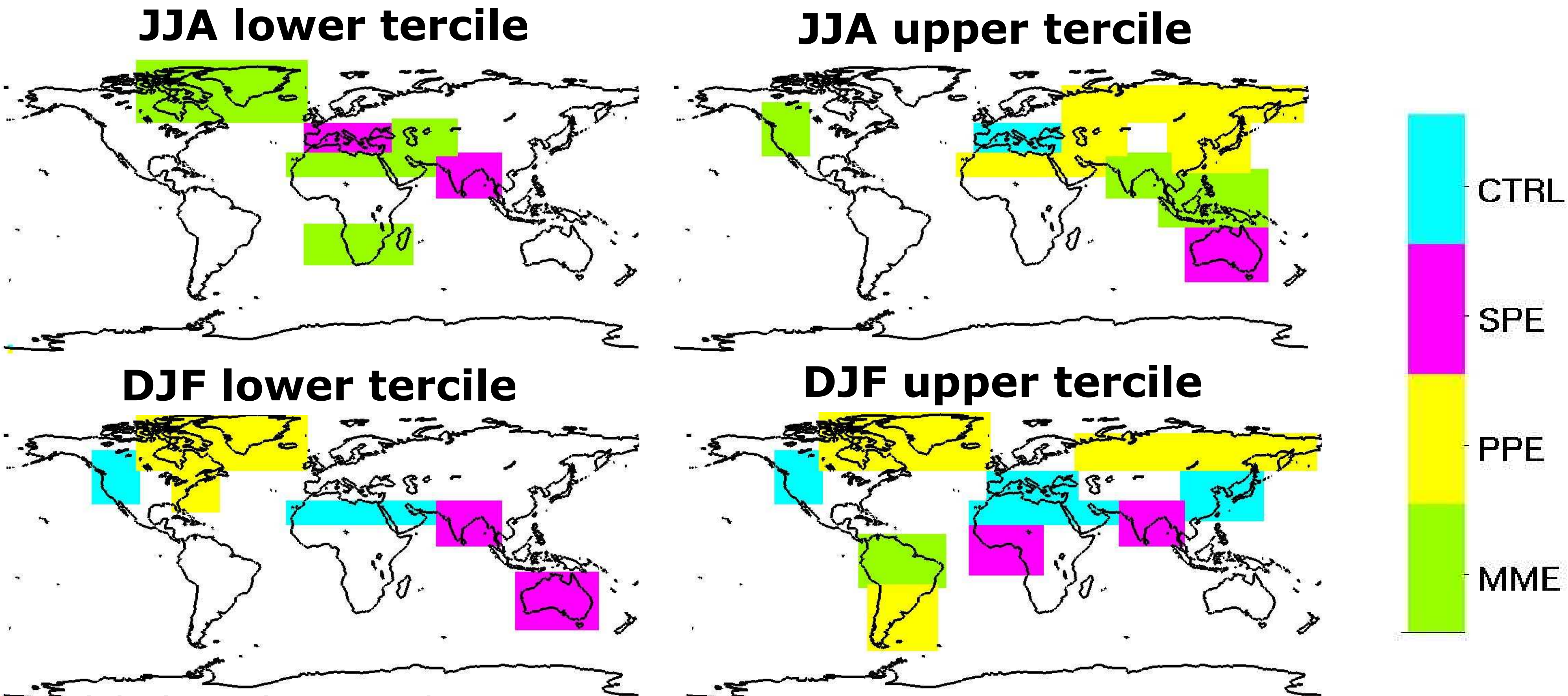
Forecast system with the largest BSS(∞) for temperature tercile events wrt ERA40/ERAInt of one-month lead seasonal predictions, 1991-2005



A. Weisheimer (ECMWF)

Method comparison: seasonal

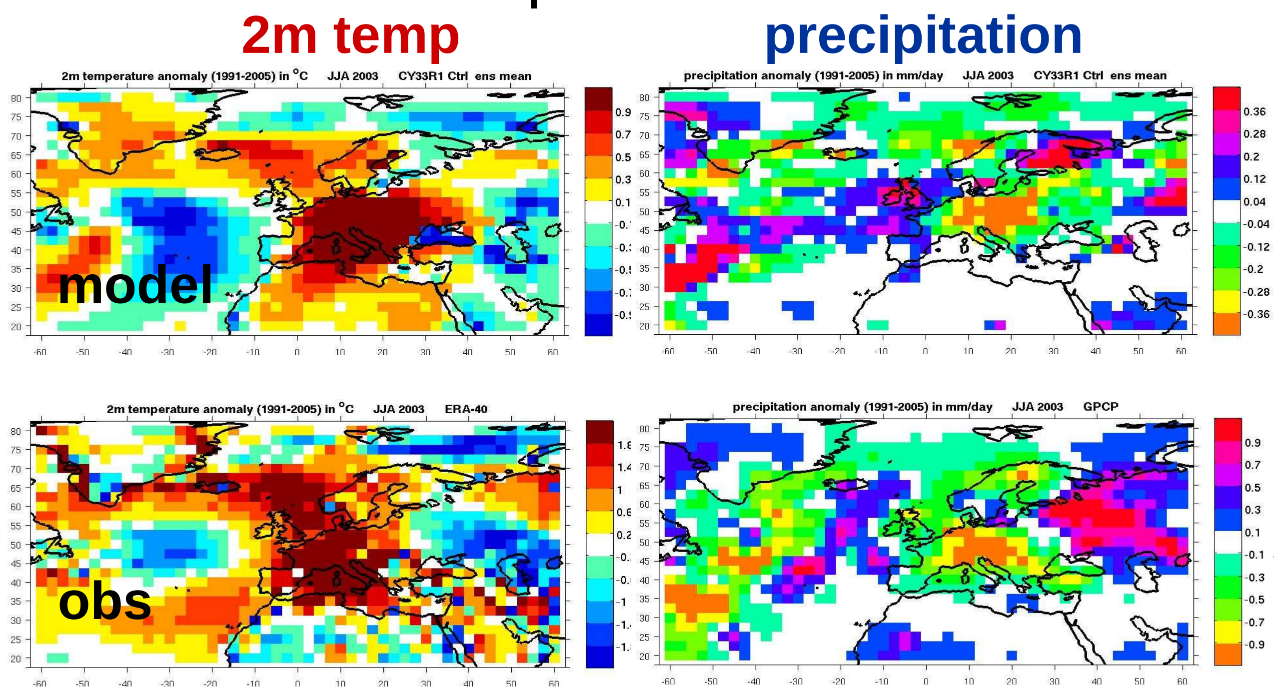
Forecast system with the largest statistically significantly positive BSS(∞) for temperature tercile events wrt ERA40/ERAInt of one-month lead seasonal predictions, 1991-2005



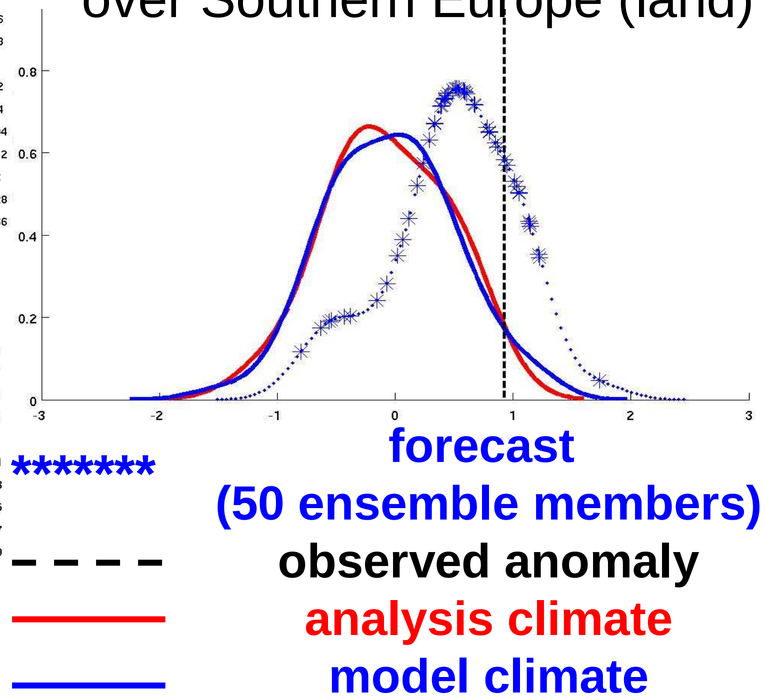
A. Weisheimer (ECMWF)

Do not forget improving the system

Seasonal prediction with ECMWF system improved wrt System 3 (changes in radiation, soil scheme and convection) for summer 2003 with start date on the 1st of May. Anomalies wrt period 1991-2005.



2m temperature over Southern Europe (land)

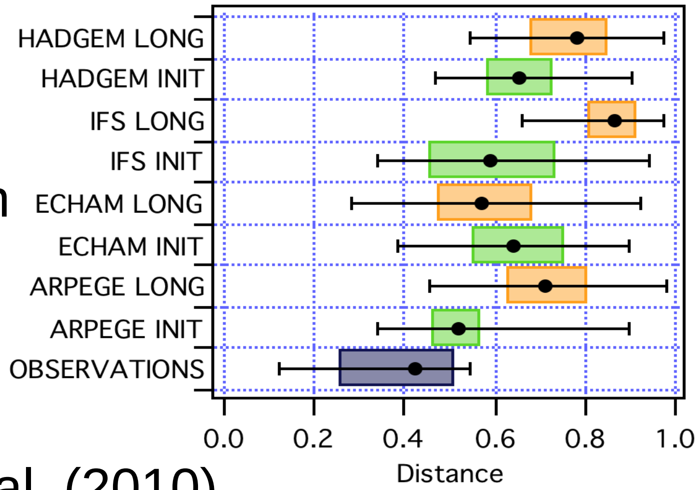


Weisheimer et al. (2011)

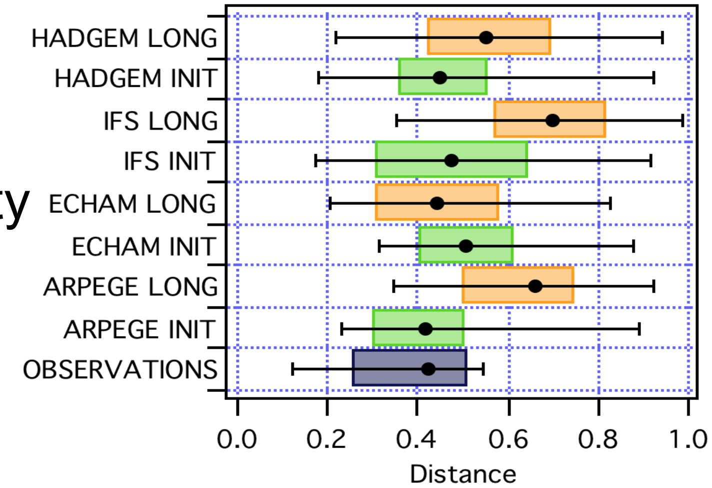
Seamless for ISV systematic errors

- Two metrics for the summer tropical ISV are defined based on the spatio-temporal propagation of precipitation IS perturbations. Realism measures the distance to mean observed pattern and reproducibility to mean model pattern.
- Results are valid for May, but with different lead times:
  - Init - first month of the integration (Seasonal runs)
  - Long - seventh month of the integration (Annual runs)

Realism



Reproducibility



Xavier et al. (2010)

Summary

- Substantial systematic error, including lack of reliability, is still a fundamental problem in dynamical forecasting.
- Comprehensive forecast quality measures (including standard error estimates) are needed in forecast system comparisons.
- Stochastic physics schemes can reduce systematic error without affecting forecast quality (unlike post-processing).
- At seasonal scale, PP and SP ensembles are competitive with MM ensembles, with gains both in accuracy and reliability.
- The current comparison of the methods does not allow for the possibility to unequivocally single out the best one.
- A new type of collaborative efforts that employs the full range of climate information sources acknowledging the seamlessness of climate variability and change has been explored in ENSEMBLES.

Just a bit more

- Impact of appropriately representing GHG, aerosol (anthropogenic and natural, including volcanoes), ozone, … concentrations in the hindcasts.
- New protocols recently developed for data dissemination and experiment documentation (ESGF, Metafor, and all the technology developed for CMIP5).
- Common verification tools in the short-, medium-range, subseasonal, seasonal and decadal communities: dealing with observational uncertainty, development of methods to assess forecast quality for extremes, …
- Robust forecast calibration and combination.