Barcelona Supercomputing Center

- Created in 2005; 350 employees
- Research, develop and manage information technology
- Facilitate scientific progress and its application in society

Earth Sciences Department

- Environmental modelling and forecasting
- Structure: four groups (around 50 people), funded by public and private sources

COMPUTATIONAL EARTH SCIENCES

ATMOSPHERIC COMPOSITION

CLIMATE PREDICTIONS

EARTH SYSTEM SERVICES

Earth System Services

Facilitate technology transfer of state-of-the-art research from local, national to international levels in five areas:

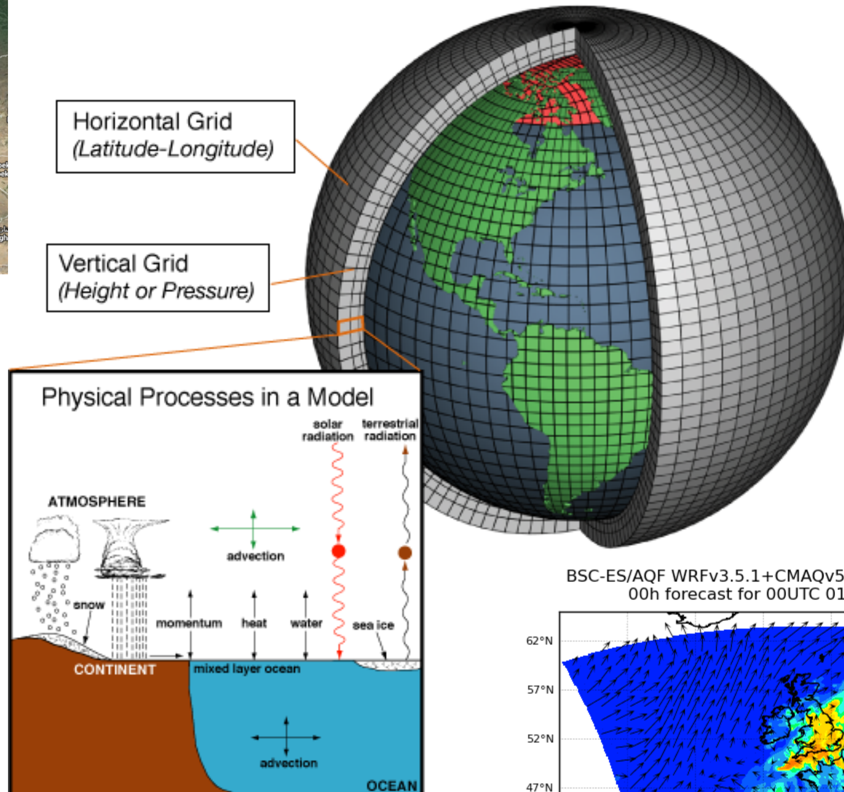
- Air quality assessments

- Mineral Dust modelling

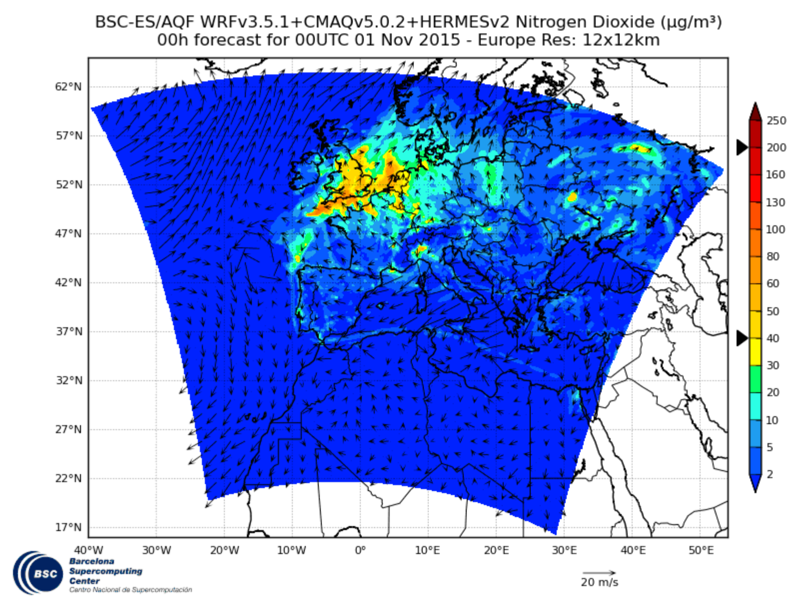
- Weather forecasting
- Climate predictions

- Computational Earth Services

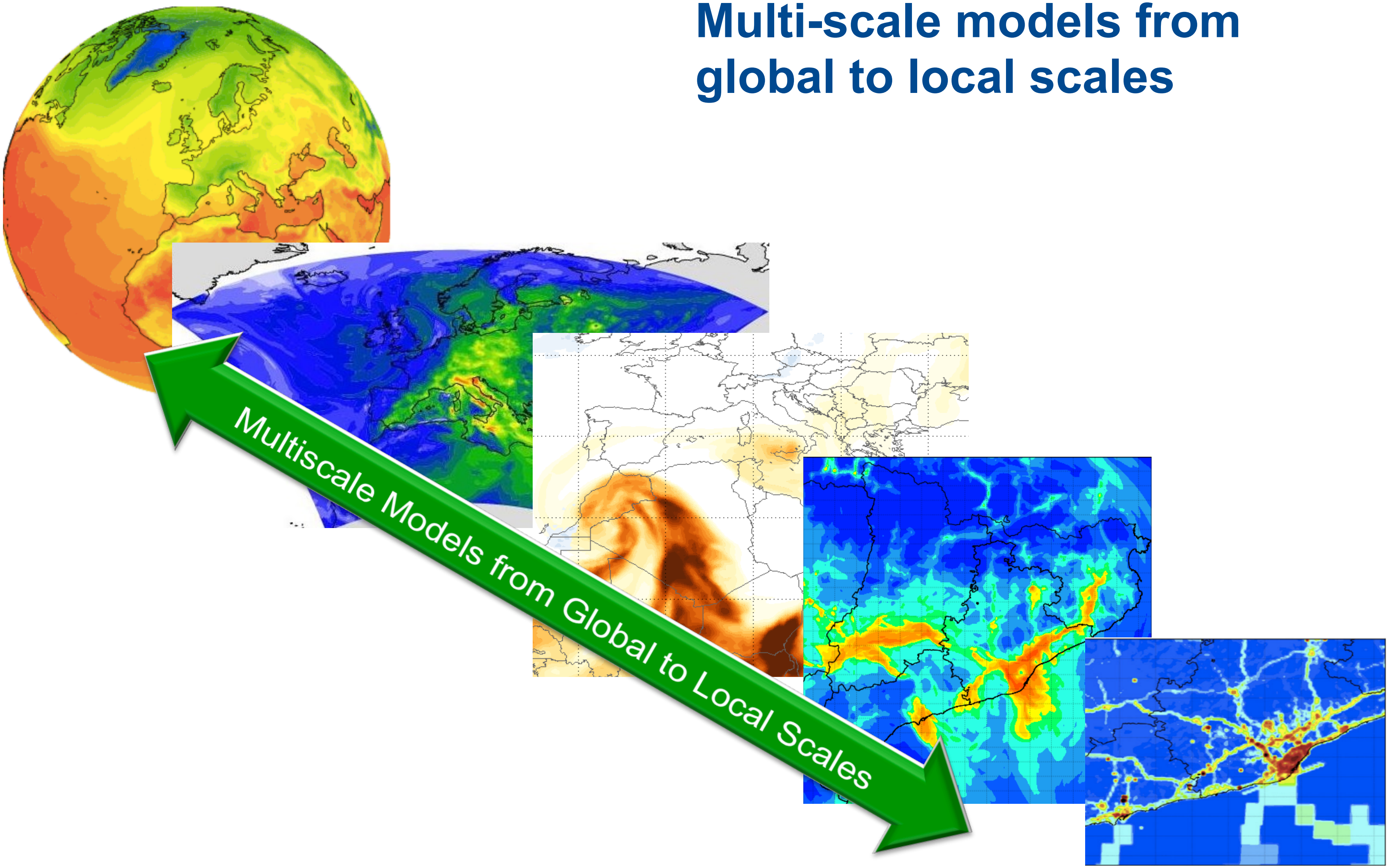
Earth sciences modelling: climate and air quality modelling



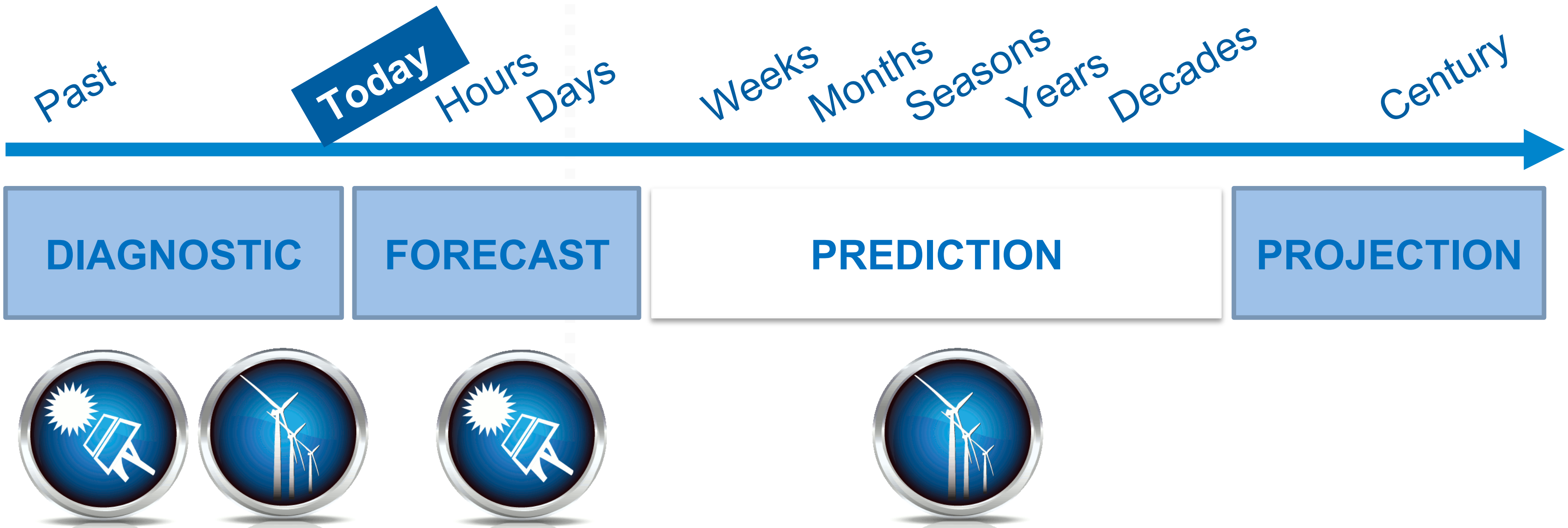
Towards modelling the Earth system



Multi-scale models from global to local scales

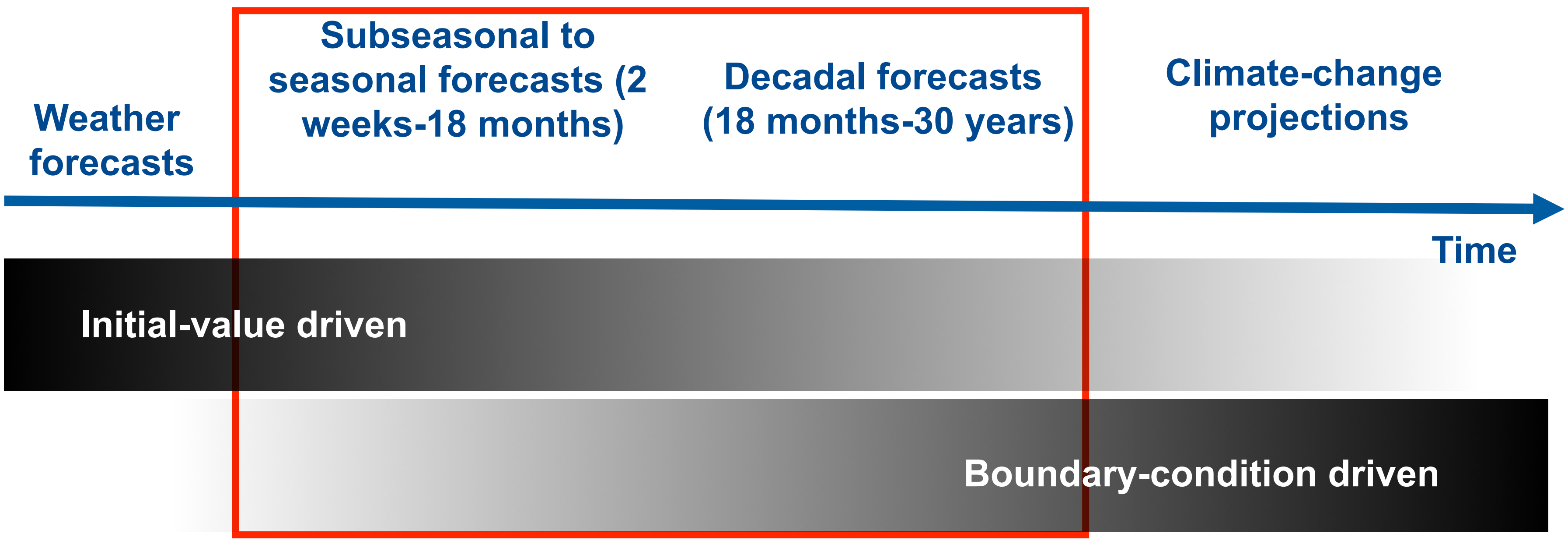


Temporal scales



Climate prediction time scales

Progression from initial-value problems with weather forecasting at one end and multi-decadal to century projections as a forced boundary condition problem at the other, with climate prediction (sub-seasonal, seasonal and decadal) in the middle. Prediction involves initialization and systematic comparison with a simultaneous reference



Adapted from Meehl et al. (2009)

Sources of seasonal predictability

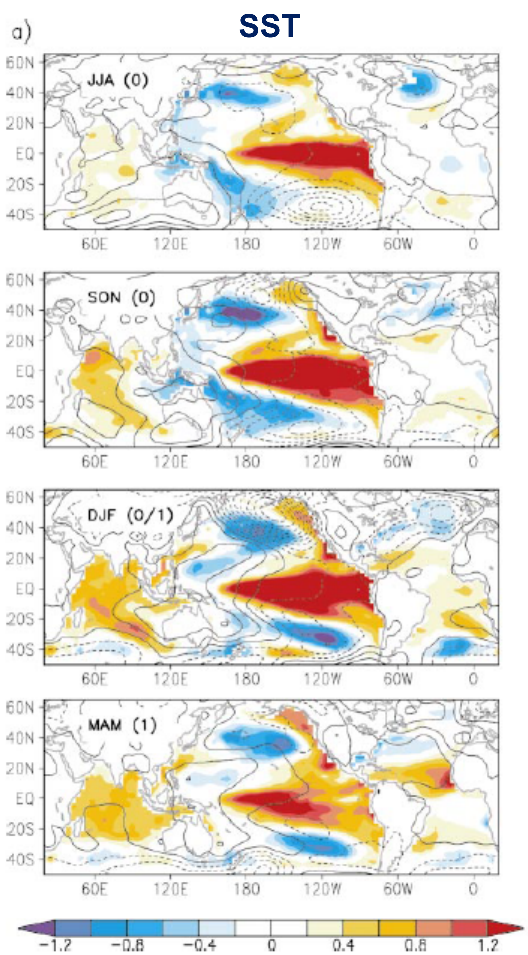
- Important:
  - ENSO (El Niño Southern Oscillation) - biggest single signal
  - Other tropical ocean SST - difficult
  - Climate change - important in mid-latitudes
  - Local land surface conditions - soil moisture, snow
  - Atmospheric composition - difficult

- Other factors:
  - Volcanic eruptions - important for large events
  - Mid-latitude ocean temperatures - still somewhat controversial
  - Remote soil moisture/snow cover - not well established
  - Sea-ice anomalies - at least local effects
  - Stratospheric influences - various possibilities
  - Remote tropical atmospheric teleconnections

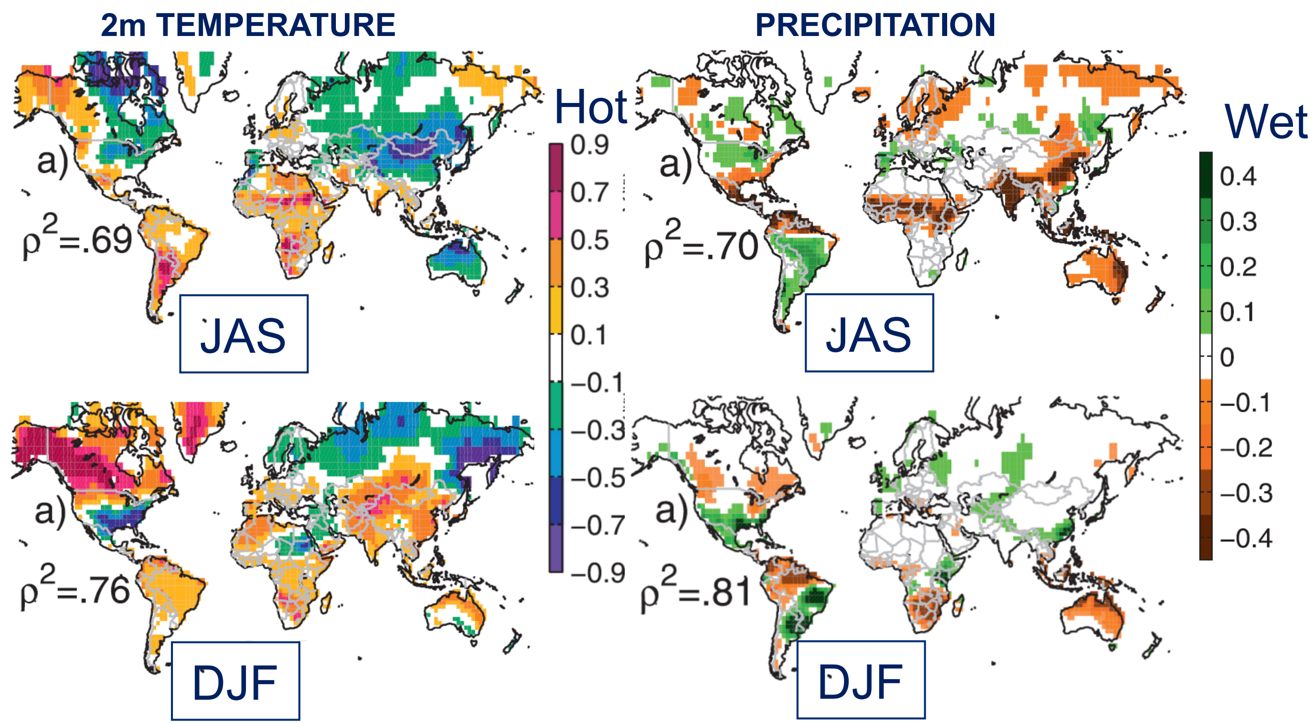
- Unknown or Unexpected

Sources of seasonal predictability

ENSO is the most important source of predictability at seasonal timescales (see e.g. Doblas-Reyes et al. 2013)



ENSO teleconnections: regressions onto Niño3.4



(Alexander et al. 2002; Yang and DelSole 2012)

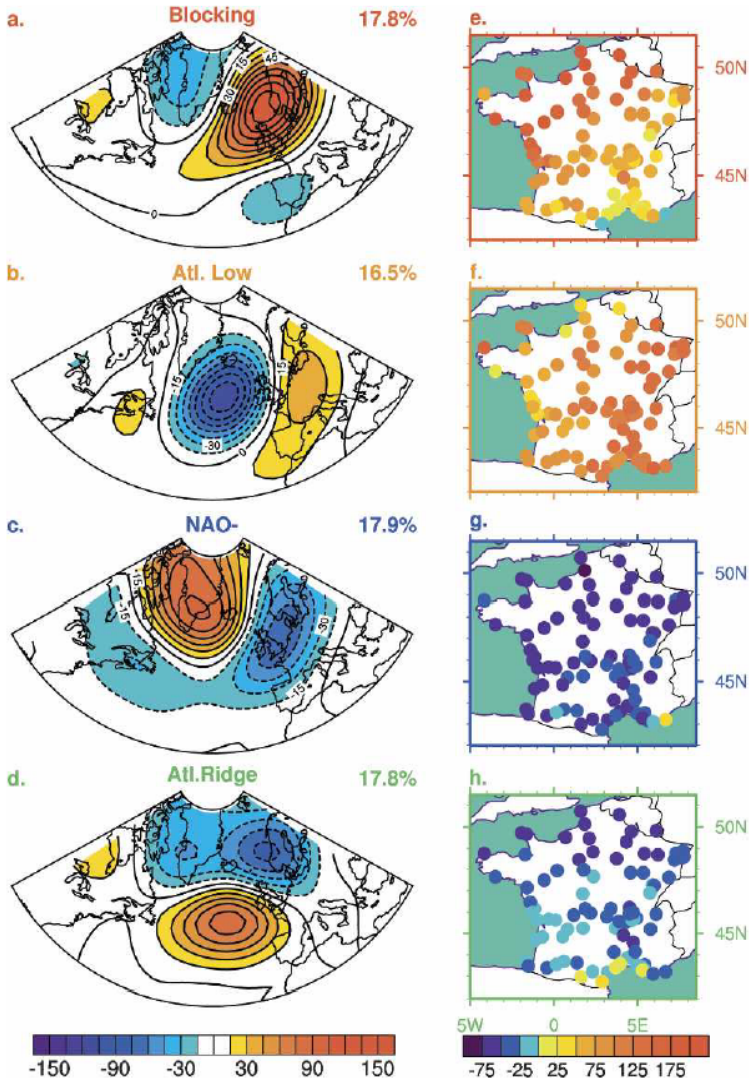
Methods of seasonal forecasting

- Empirical forecasting
  - Use past observational record and statistical methods
  - Works with reality instead of error-prone numerical models
  - Limited number of past cases
  - A non-stationary climate is problematic
  - Can be used as a benchmark

- Coupled GCM forecasts
  - Include comprehensive range of sources of predictability
  - Predict joint evolution of ocean and atmosphere flow
  - Includes a large range of physical processes
  - Includes uncertainty sources, important for prob. Forecasts
  - Systematic model error is an issue!

An example: Weather types

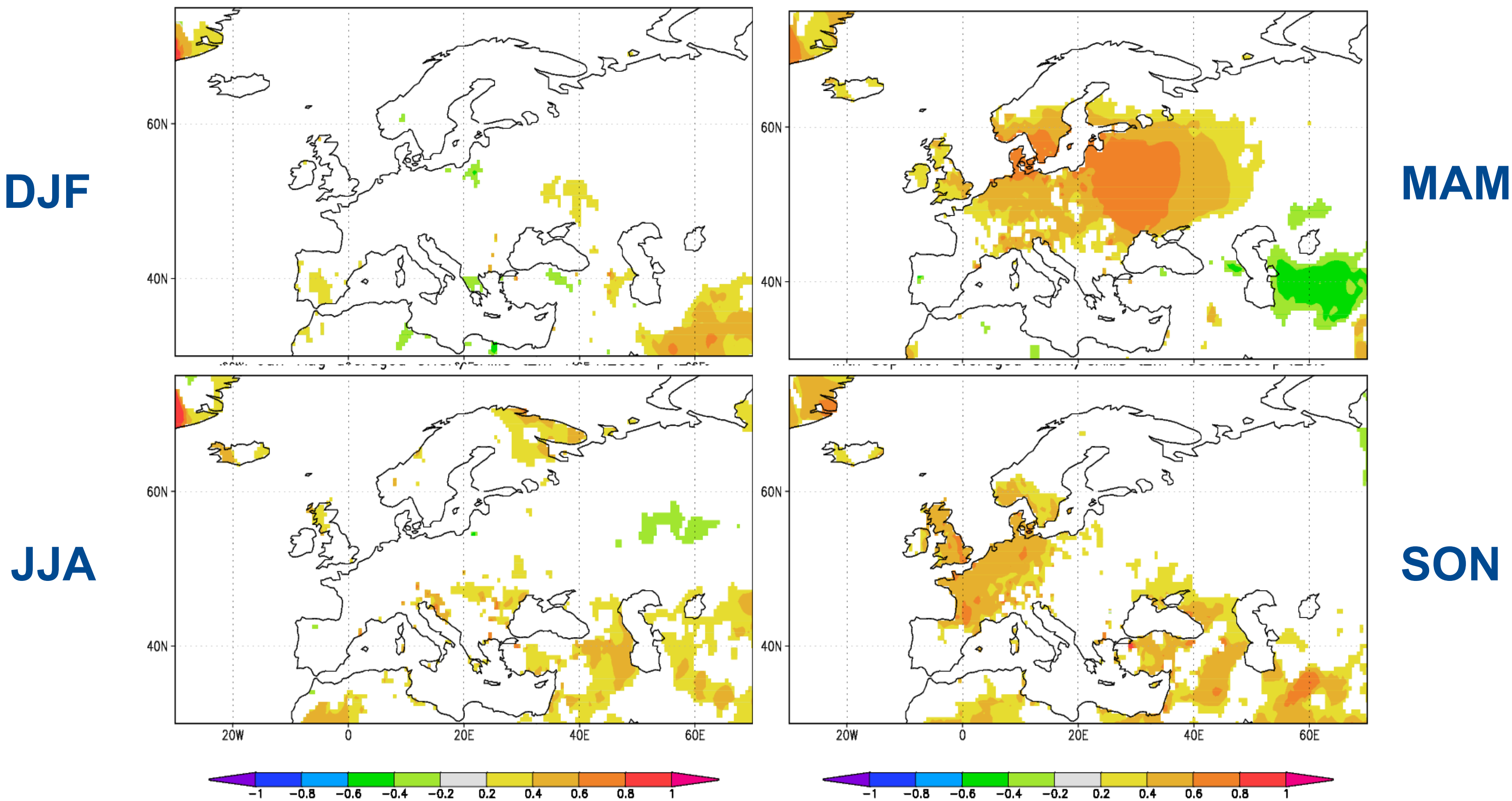
Z500 summer weather types and frequency change (%) of warm days



Cassou et al. (2005)

Temperature skill: persistence

Correlation of GHCN temperature of one-month lead anomaly persistence over 1981-2005. Only values statistically significant with 80% confidence are plotted.

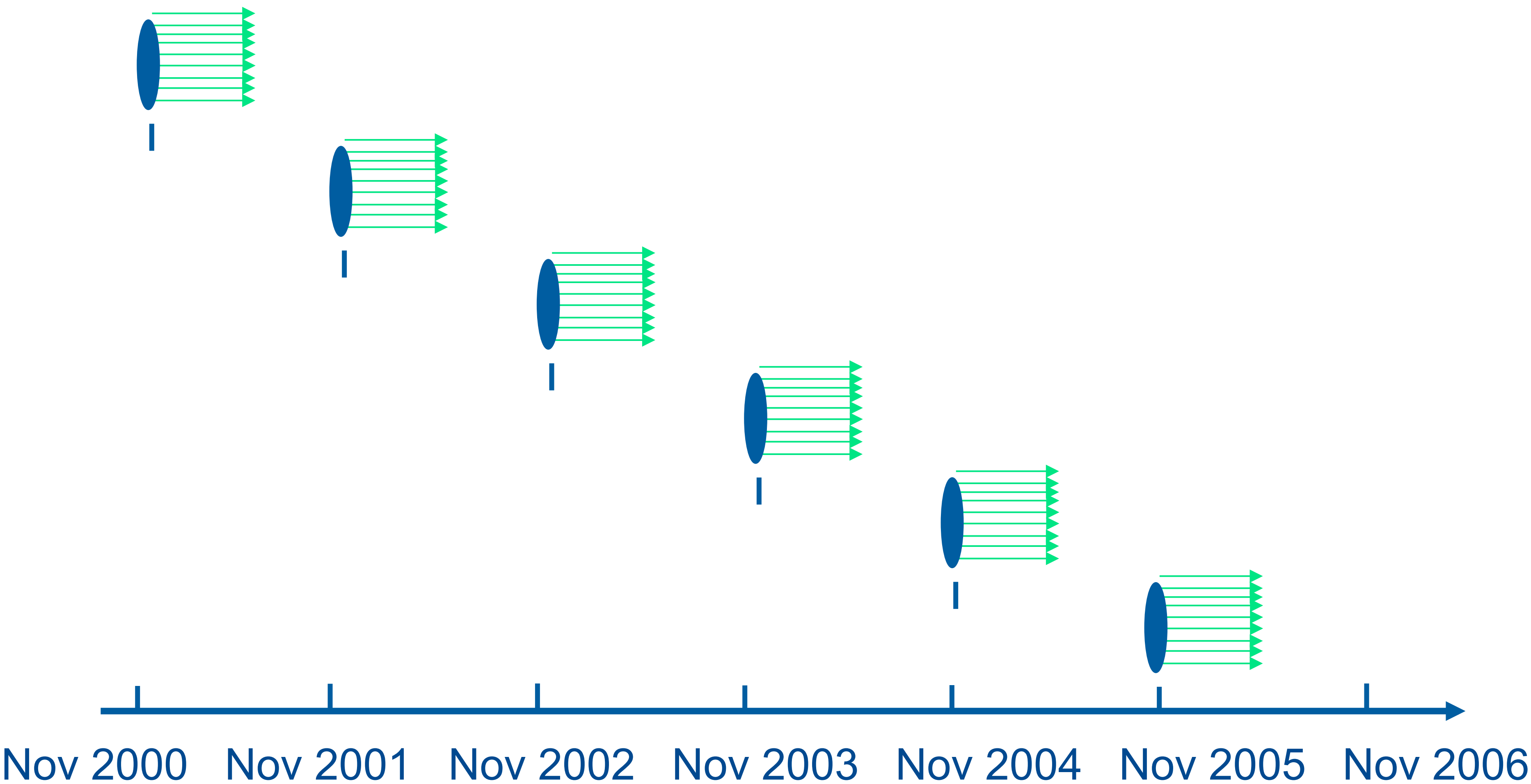


To produce dynamical forecasts

- Build a coupled model
- Prepare initial conditions
- Initialize coupled system
  - The aim is to start the system close to reality. Accurate SST is particularly important, plus ocean sub-surface.
- Run an ensemble forecast
  - Explicitly generate an ensemble on the e.g. 1st of each month, with perturbations to represent the uncertainty in the initial conditions; run forecasts for several months.
- Produce probability forecasts from the ensemble
- Apply calibration and combination if significant improvement is found, for which hindcasts are required

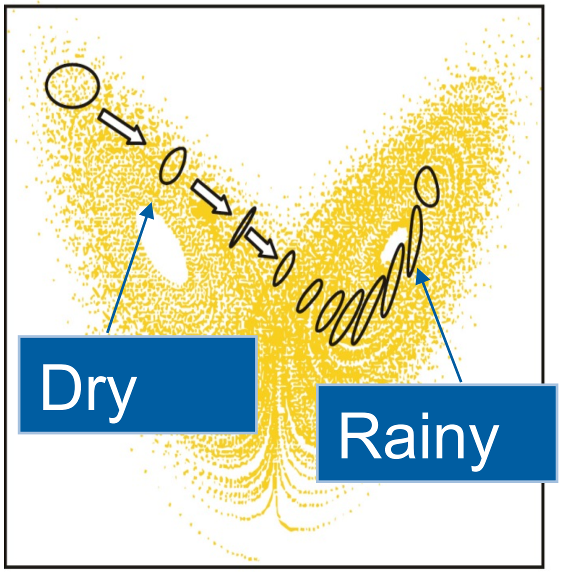
Ensemble predictions

Ensemble forecast system



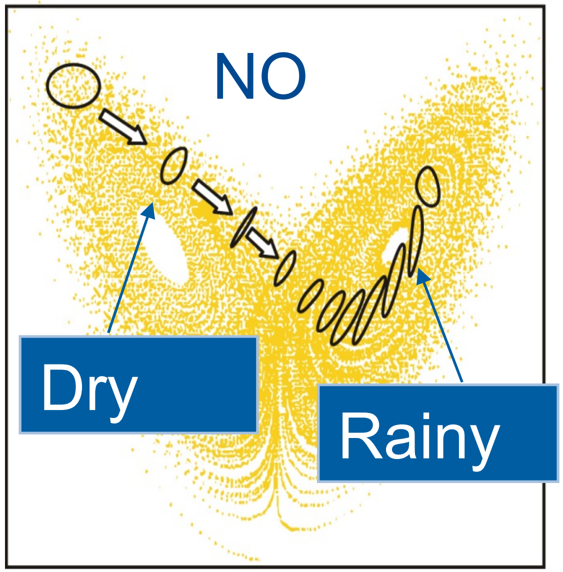
Why running several forecasts

A farmer is planning to spray a crop tomorrow



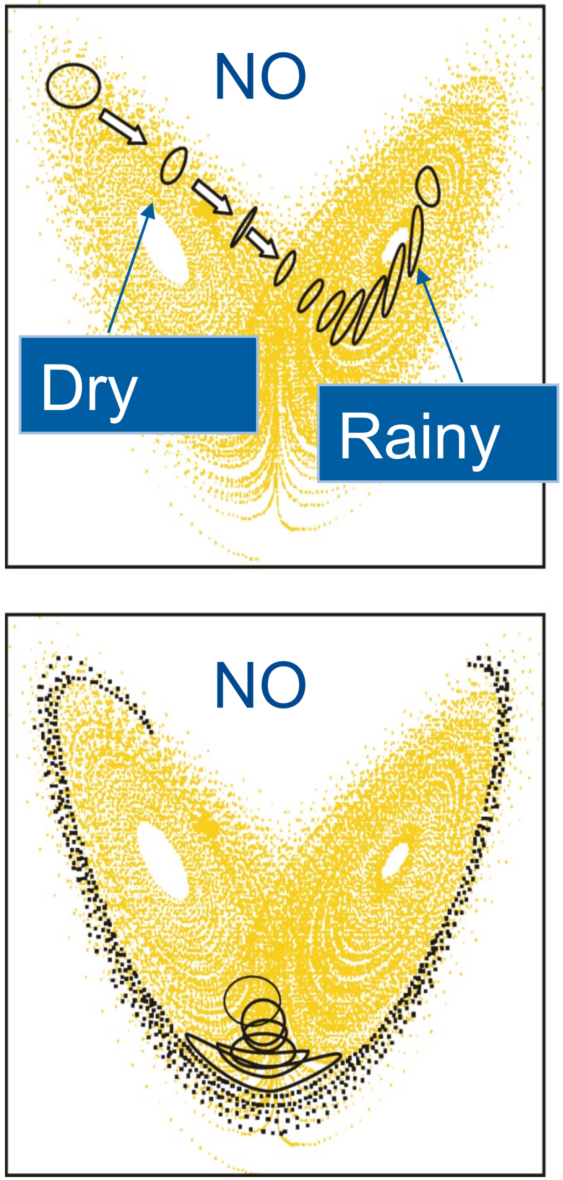
Why running several forecasts

A farmer is planning to spray a crop tomorrow



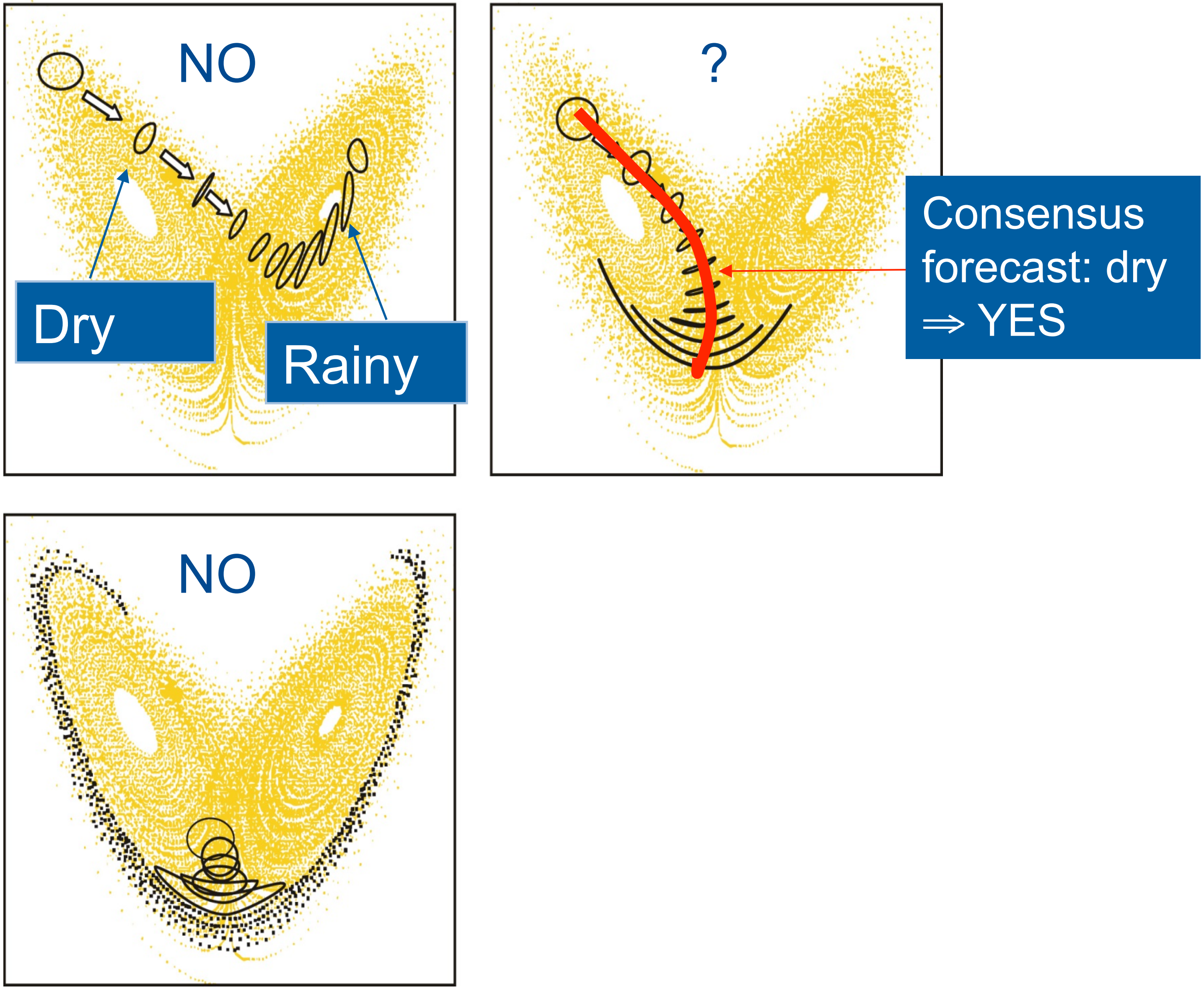
Why running several forecasts

A farmer is planning to spray a crop tomorrow

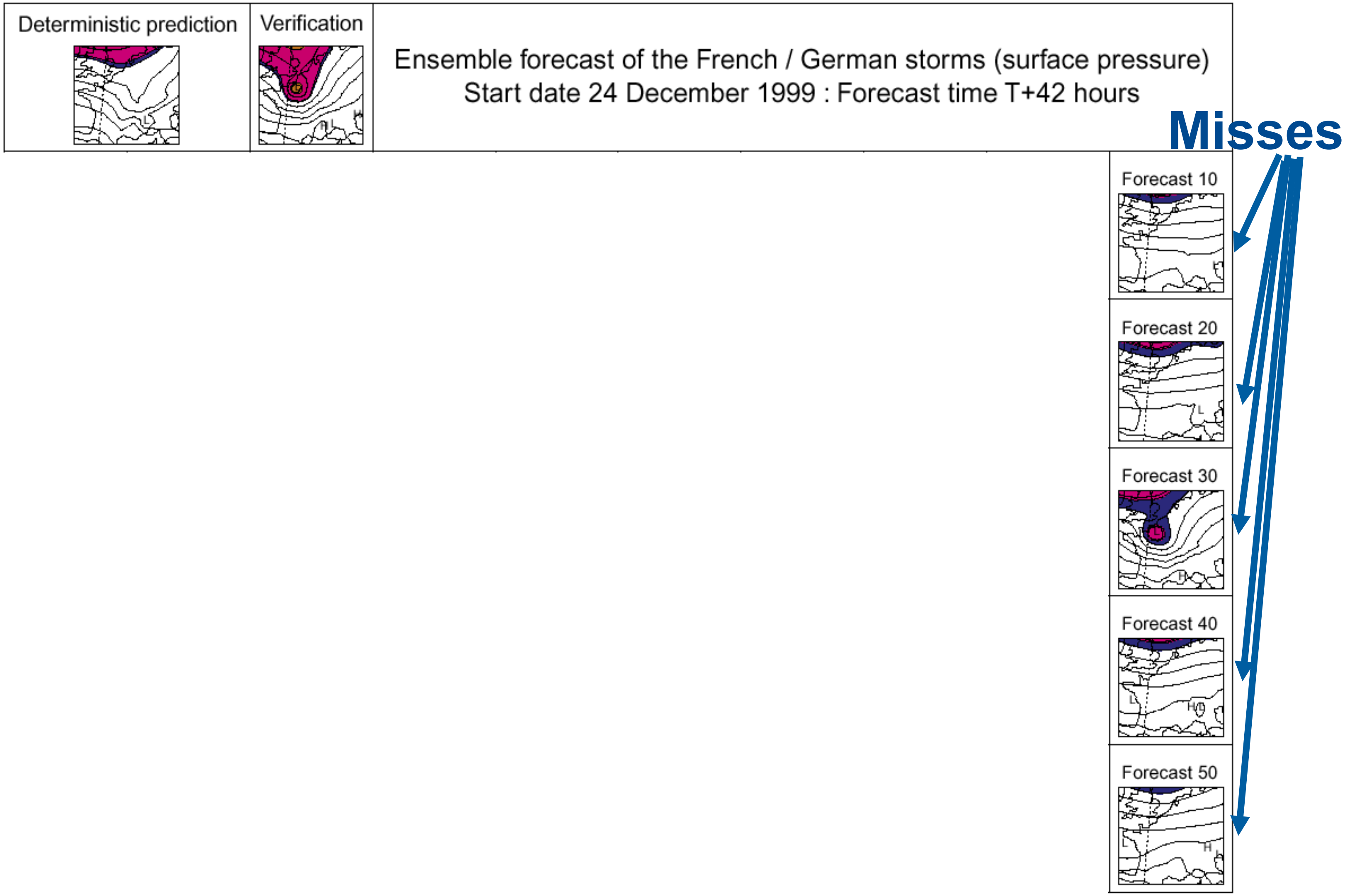


Why running several forecasts

A farmer is planning to spray a crop tomorrow

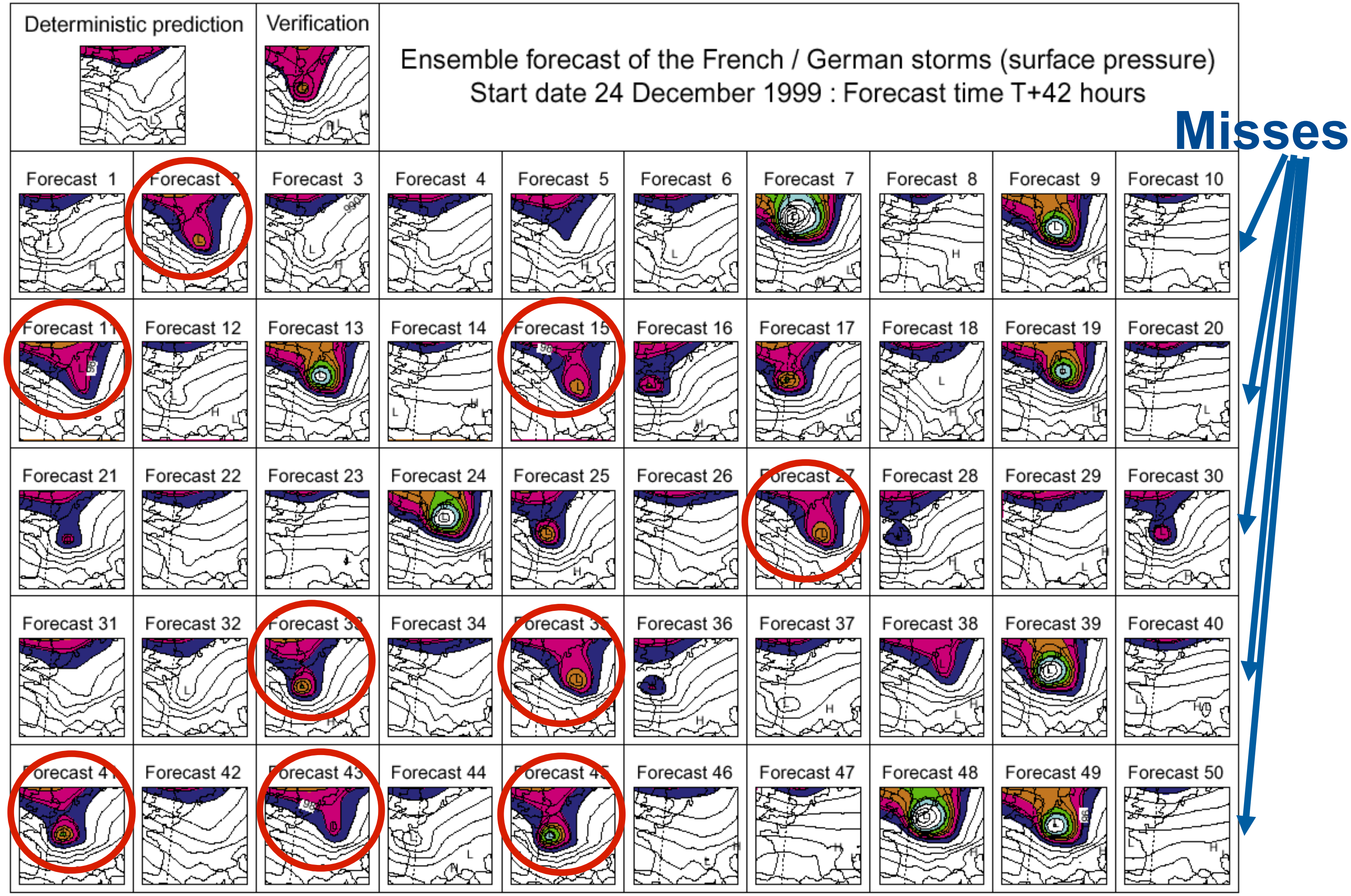


ECMWF forecasts (D+42) for the storm Lothar



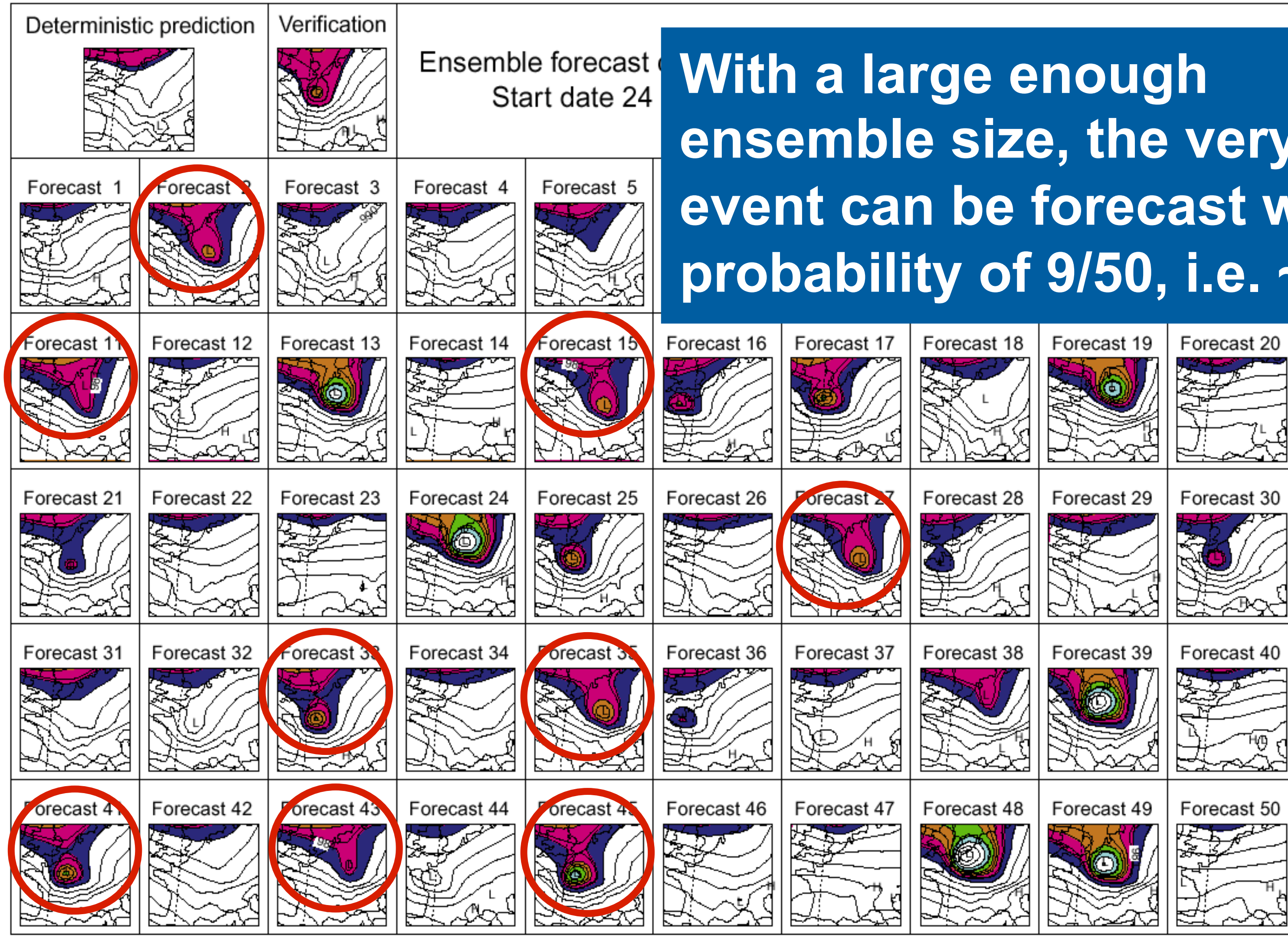
How many members: ensemble size

ECMWF forecasts (D+42) for the storm Lothar



How many members: ensemble size

ECMWF forecasts (D+42) for the storm Lothar



With a large enough ensemble size, the very rare event can be forecast with a probability of 9/50, i.e. ~20%

ECMWF System 4: main features

- **Operational forecasts**
  - 51 member ensemble from 1st day of the month
  - released on the 8th
  - 7-month integration

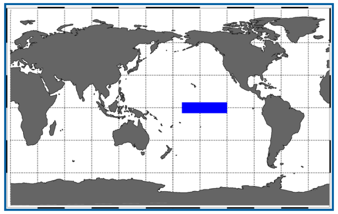
- **Re-forecast set**
  - 30 years, start dates from 1 Jan 1981 to 1 Dec 2010
  - (51) 15-member ensembles, 7-month integrations
  - 13-month extension from 1st Feb/May/Aug/Nov

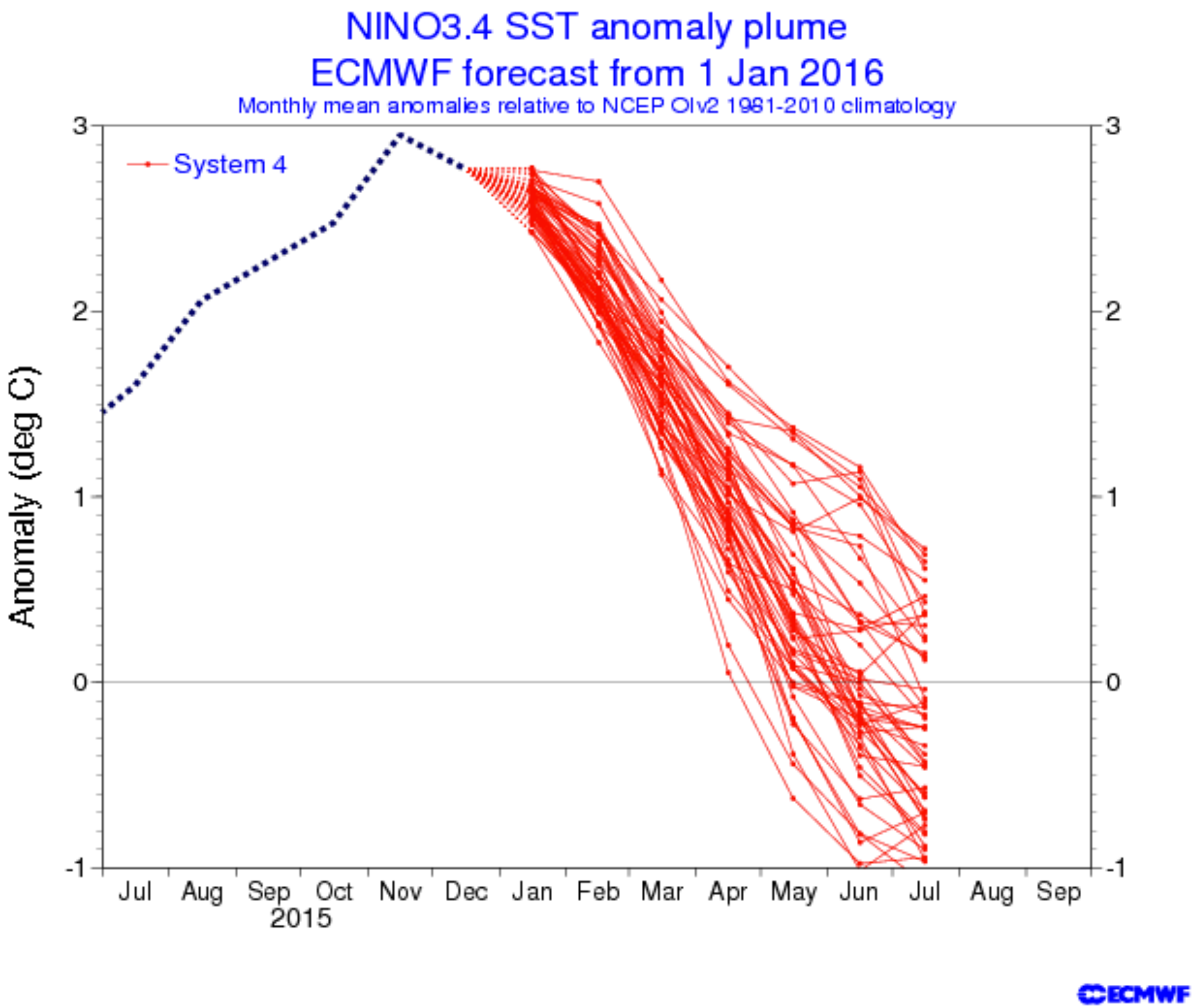
A Re-forecast is a database of historical "forecasts" performed with the same model: allow us to statistically assess model skill and to calibrate the forecasts

There are systematic errors …

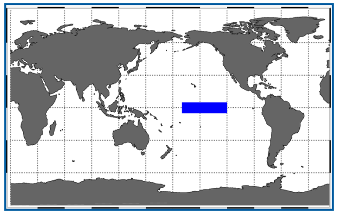
- **Model drift is typically comparable to signal**
  - Both SST and atmosphere fields
- **Forecasts are made *relative* to past model integrations**
  - Model climate estimated from 30 years of forecasts (1981-2010), all of which use a 15 (51) member ensemble
  - Model climate has both a mean and a distribution, allowing us to estimate eg tercile boundaries
  - Model climate is a function of start date and forecast lead time
- **Implicit assumption of linearity**
  - We implicitly assume that a shift in the model forecast relative to the model climate corresponds to the expected shift in a true forecast relative to the true climate, despite differences between model and true climate
  - Most of the time, the assumption seems to work pretty well. But not always

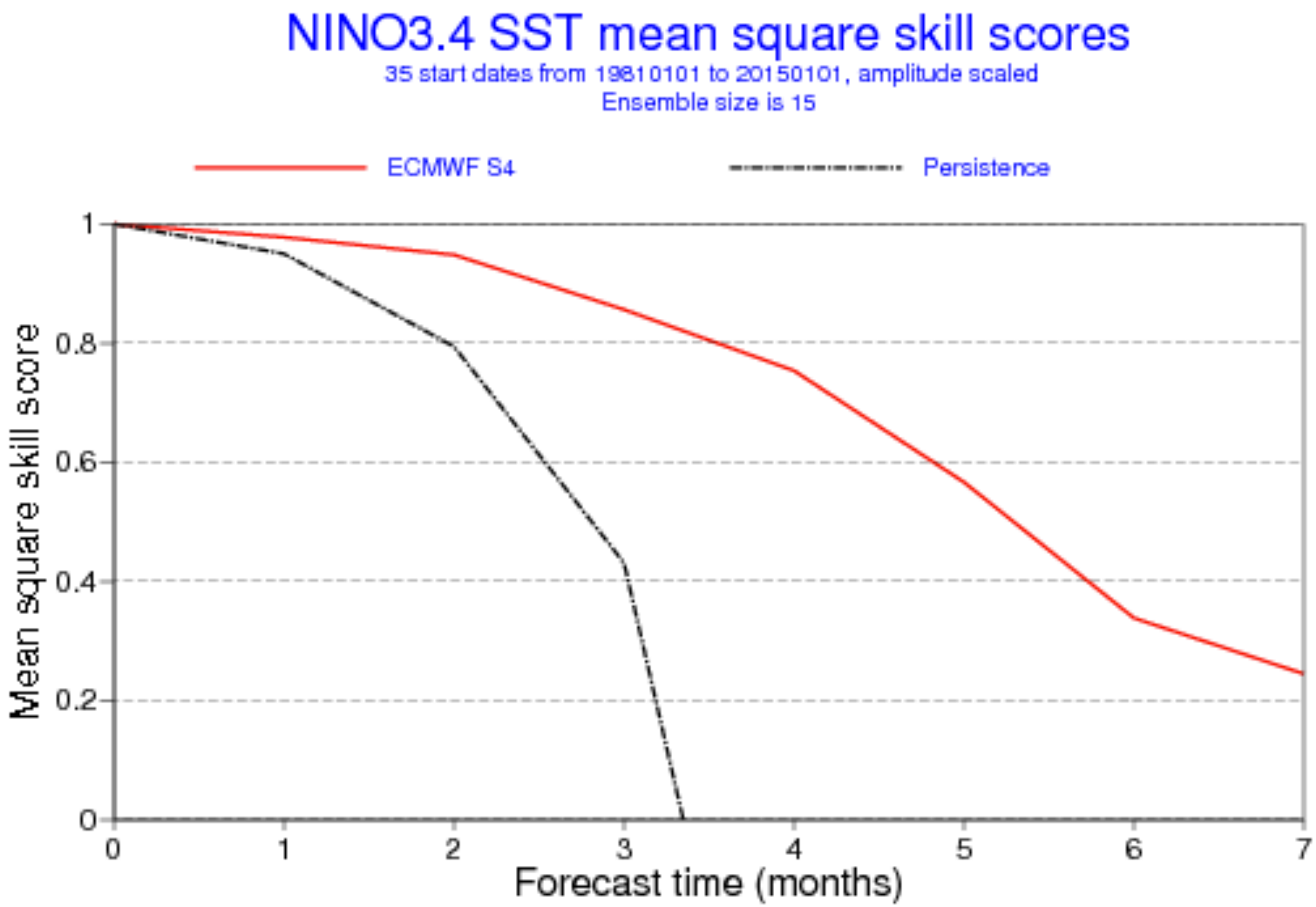
ENSO ensemble predictions





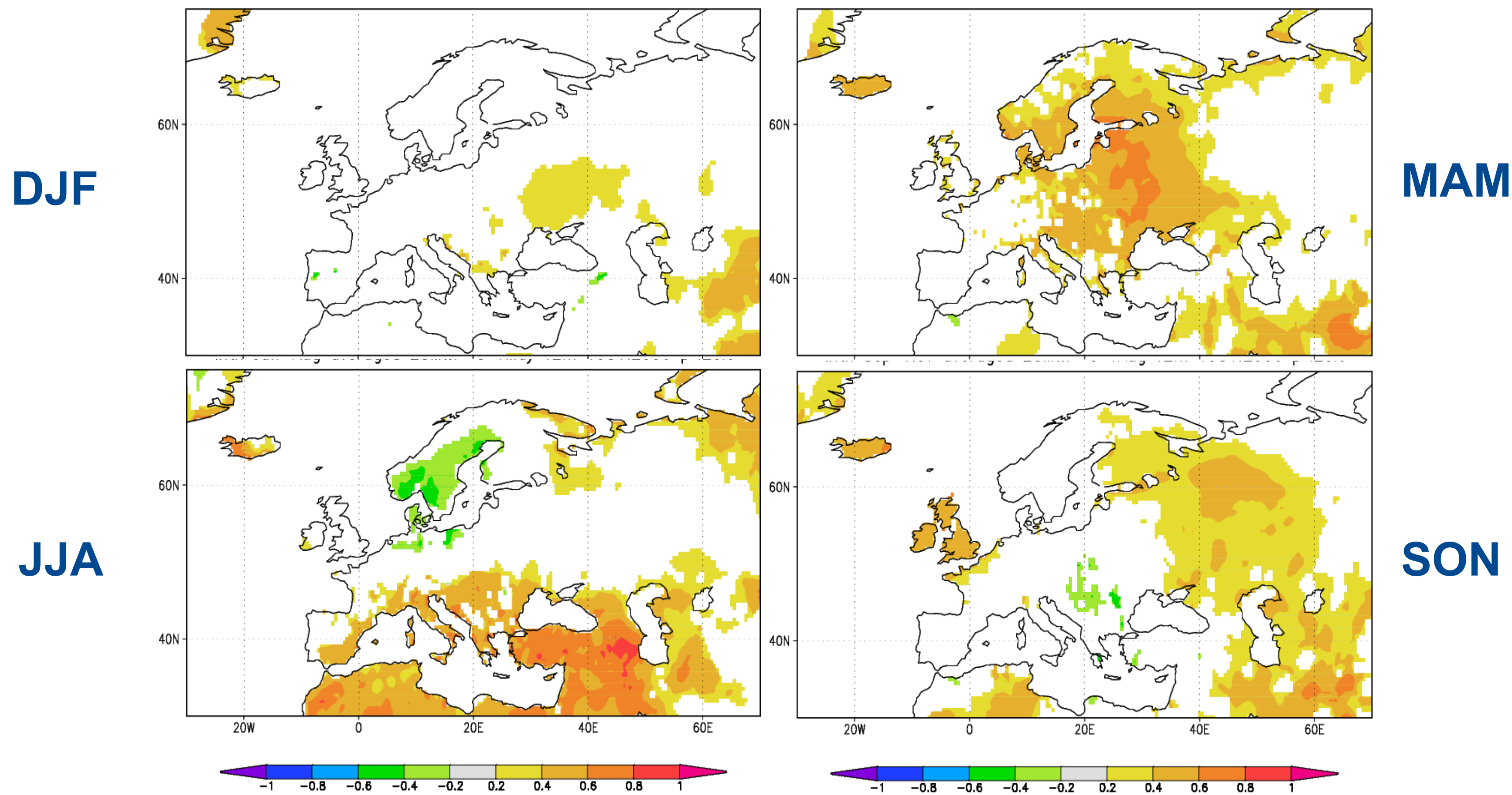
ENSO ensemble predictions





Regional skill: System 4

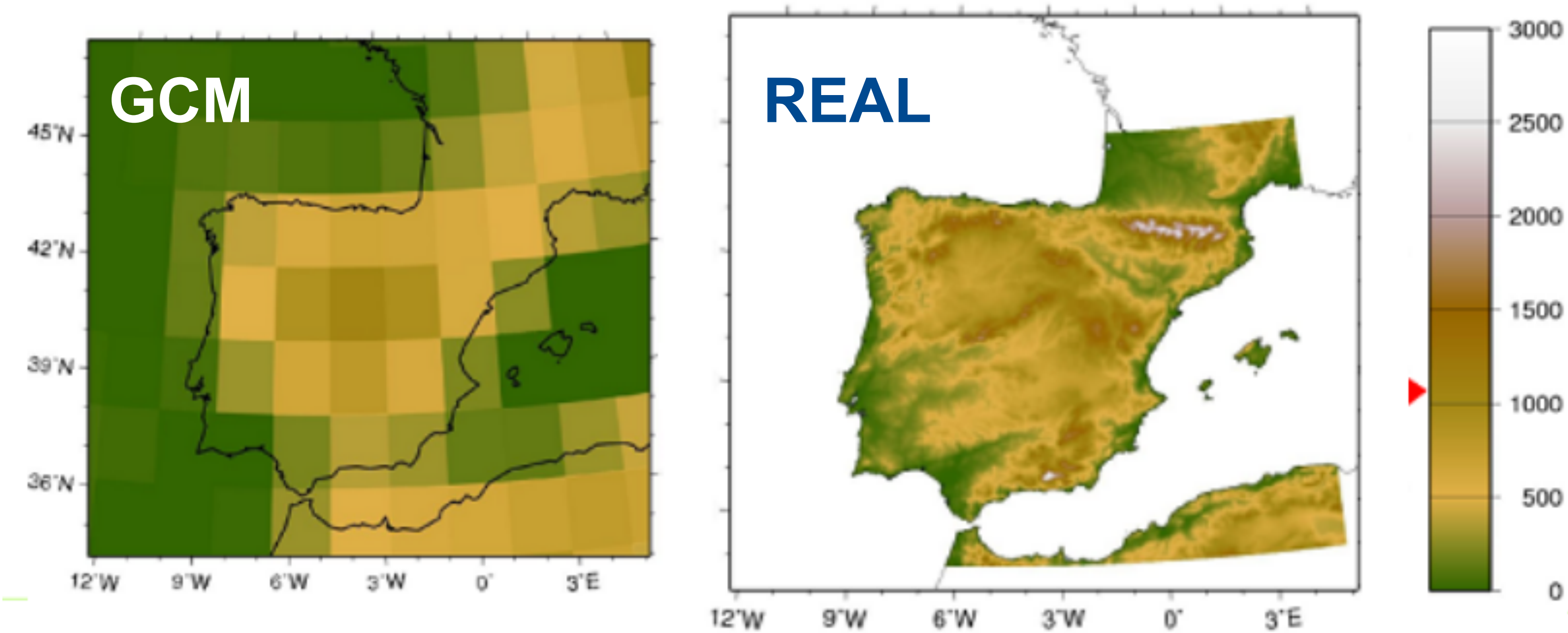
Correlation of System 4 seasonal forecasts (lead times: 2-4) of temperature wrt GHCN over 1981-2010. Only values statistically significant with 80% confidence are plotted.



Testing seasonal forecast for MARS

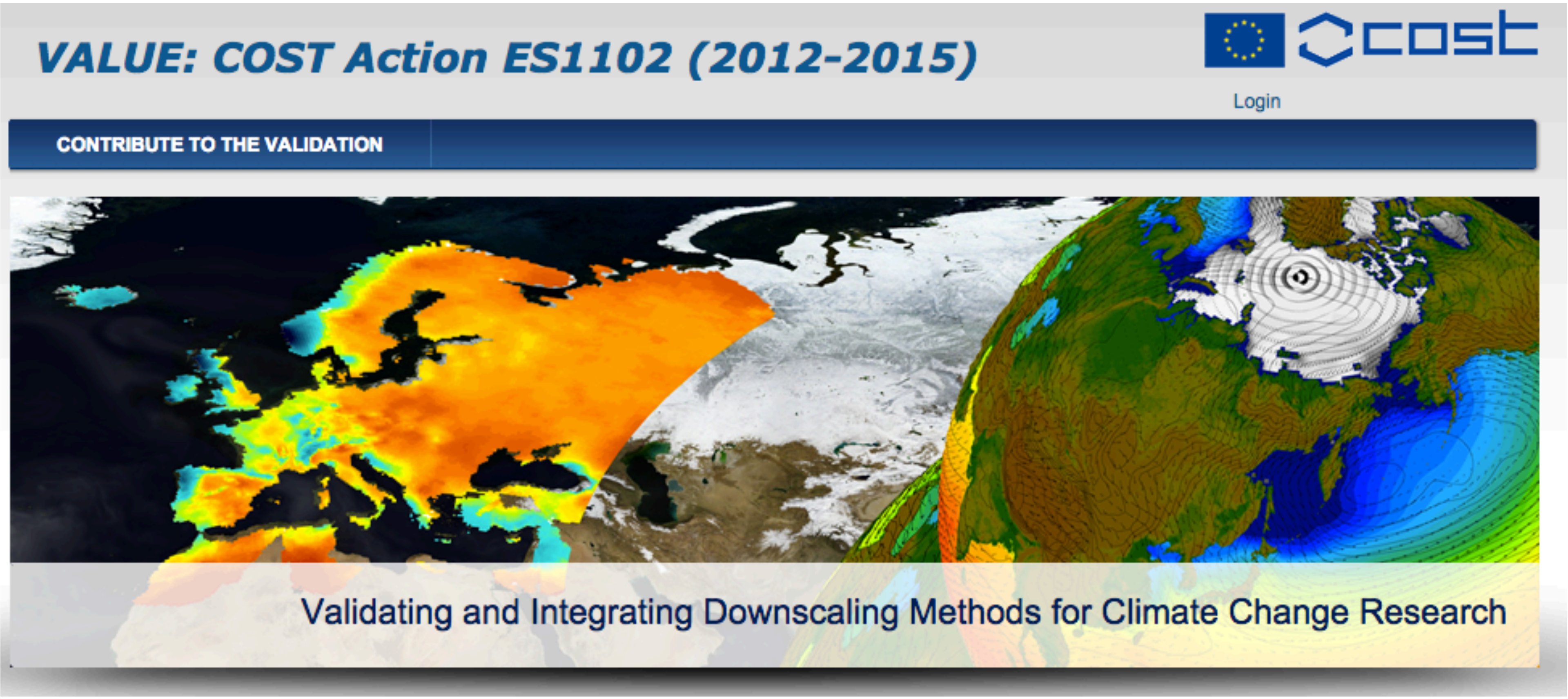
- Exploit existing seasonal forecasts: exploring how the MARS Crop Yield Forecasting System (MCYFS) could ingest the seasonal forecast after bias adjustment and downscaling (when required) for a future operational use
- The deliverables planned are:
  - D1. Review of current methodologies (month 2)
  - D2. Evaluation of the forecast quality of the S4 reforecasts (months 3-5)
  - M1-3: Implementation of a test environment for 2016 using April, May and June start dates (Milestones in month 4, 5 and 6)
  - D3. Description of the test environment for 2016 (month 6)

- GCM models have bias
- GCM grid scale mismatches decision scales



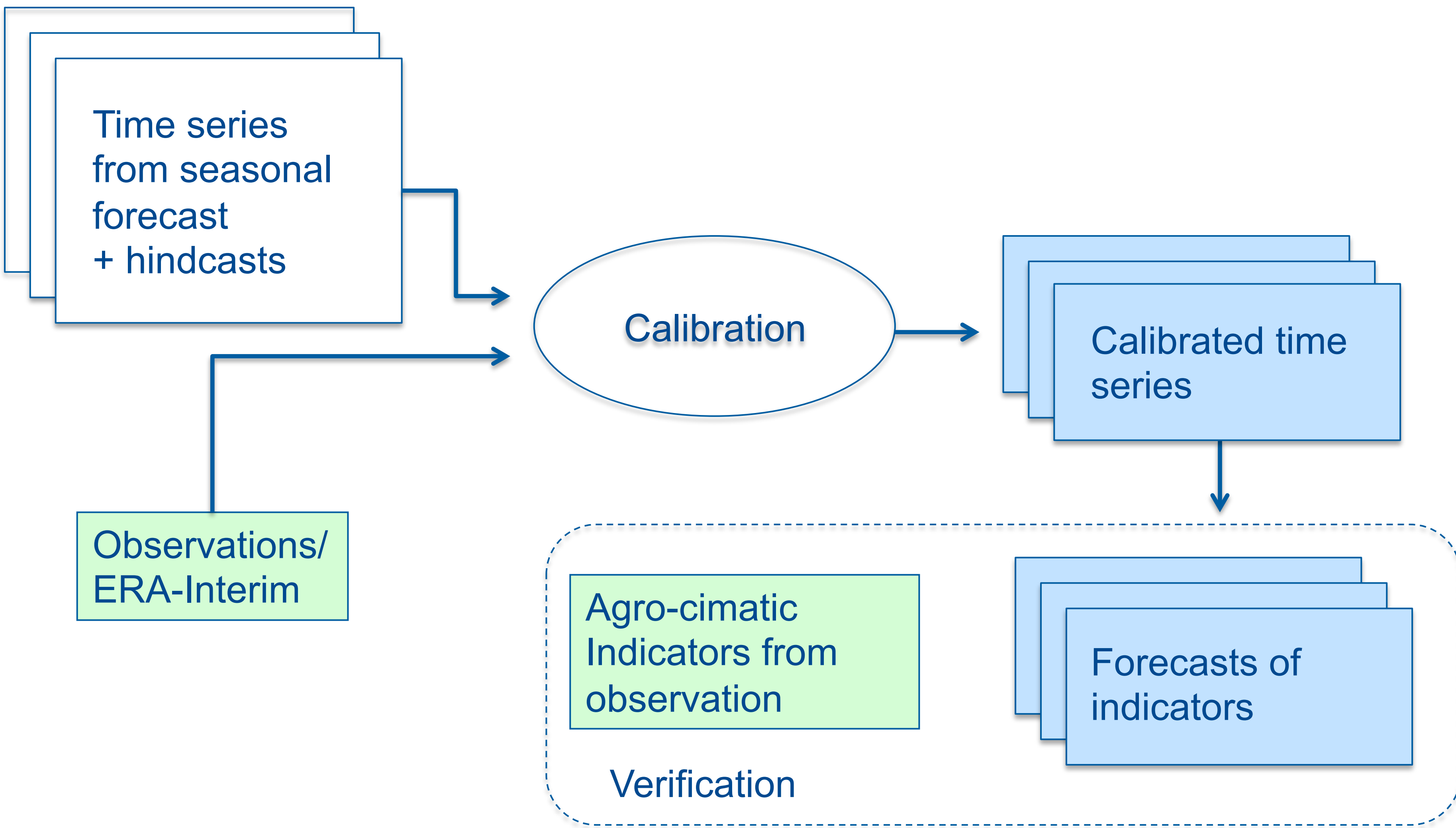
Bias correction/calibration

We are using the statistical correction methods evaluated in the framework of the COST Action VALUE

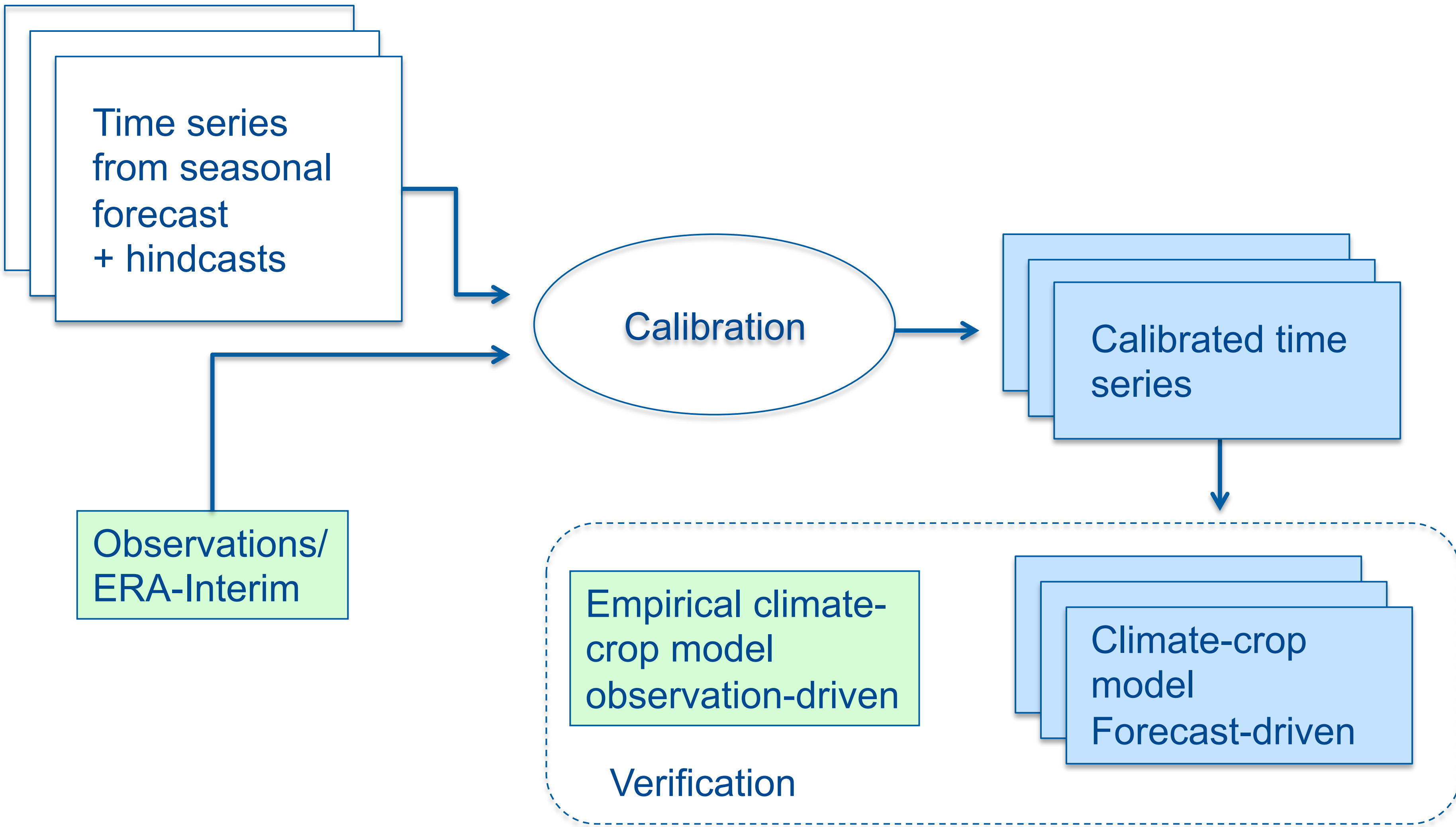


http://www.value-cost.eu/
Maraun et al. 2015

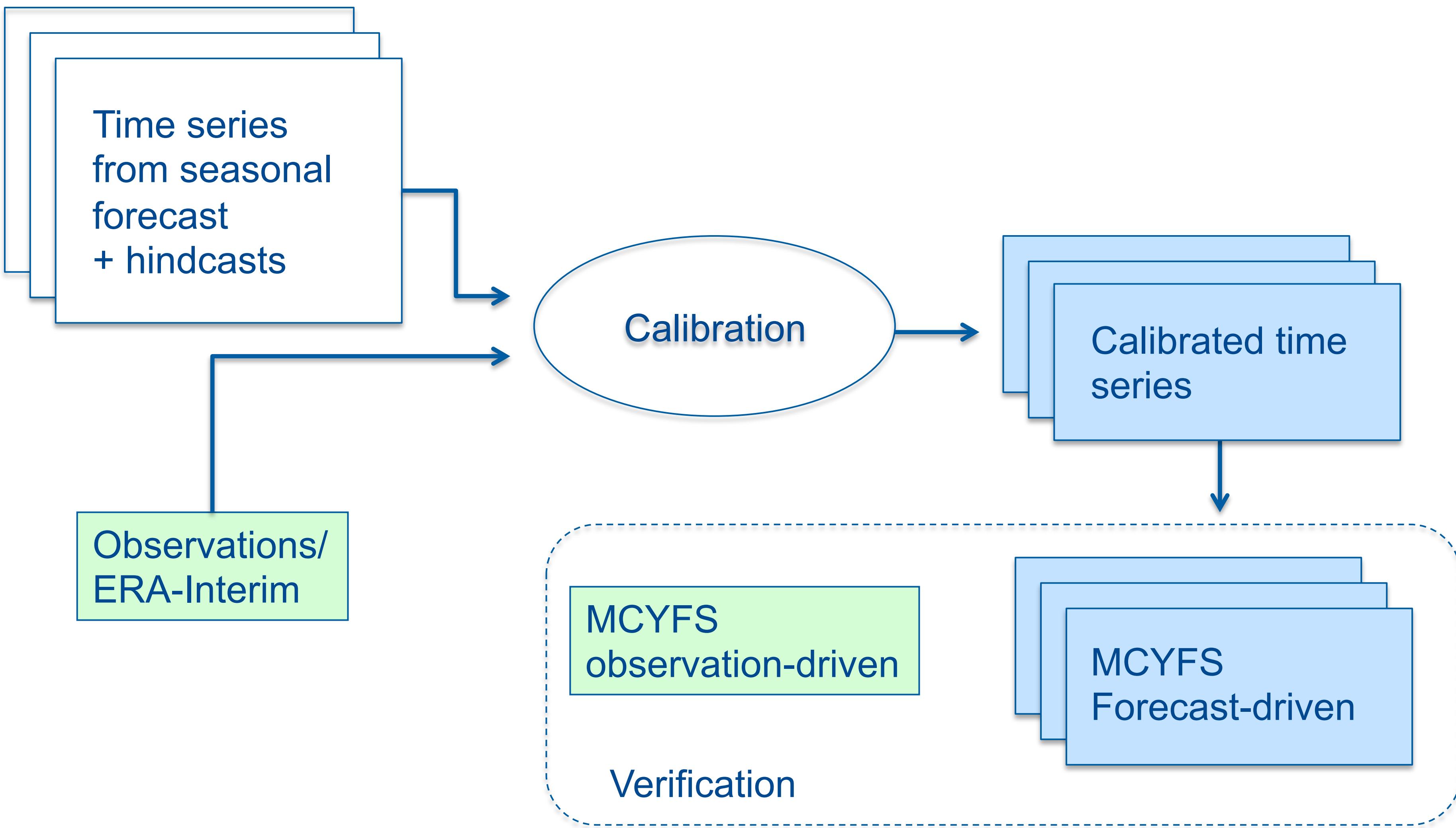
Analysis scheme



Analysis scheme

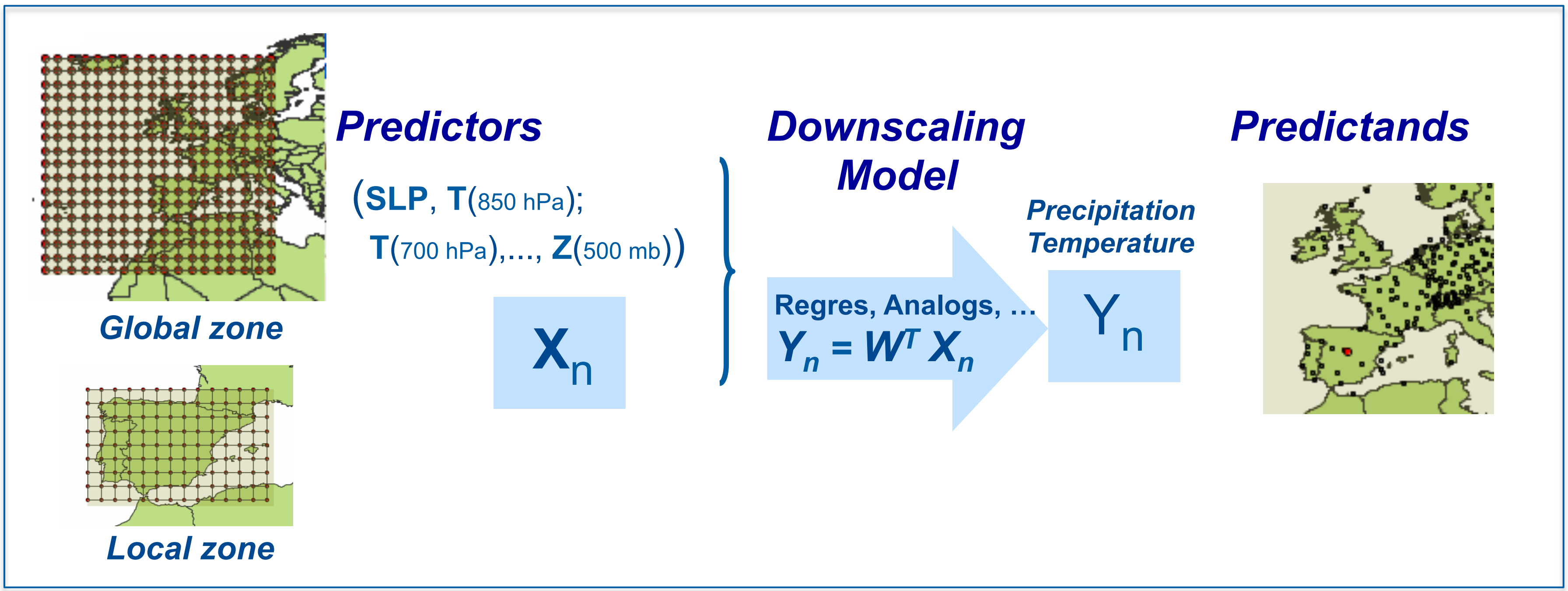


Analysis scheme



Perfect prognosis approach:

- In the training phase the statistical model is calibrated using observational data for both the predictands and predictors (e.g. reanalysis data)
- Typical techniques: transfer functions, analogs, weather typing, weather generators, etc. (Maraun et al. 2010)

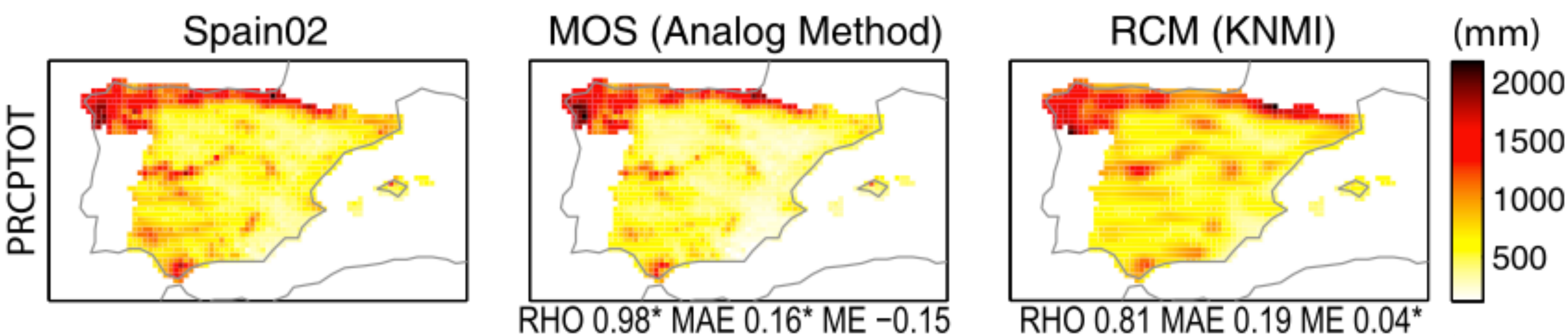


Source: José M. Gutiérrez, University of Cantabria

Bias correction/downscaling

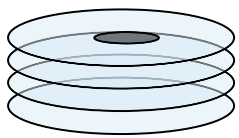
MOS (Model Output Statistics) approach:

- The predictors are taken from the same model for both the training and downscaling phases (e.g. Eden and Widmann 2014)
- Typical techniques:
  - Linear-scaling: Tadj=Traw-(<Tmod>-<Tobs>)
  - Quantile mapping (empirical or parametric)
  - Linear regression, analog methods, etc.

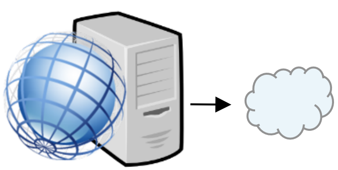


Source: Turco et al. 2011

S2dverification is an R package to verify seasonal to decadal forecasts by comparing experiments with observational data. It allows analyzing data available either locally or remotely. It can also be used online as the model runs. Available from CRAN.



LOCAL STORAGE



ESGF NODE or OPeNDAP SERVER

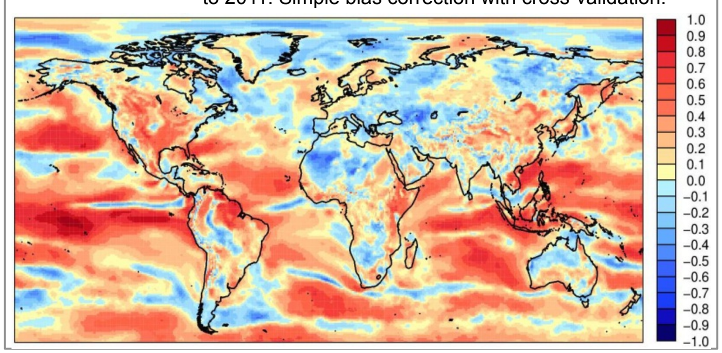
- Supports datasets stored locally or in ESGF (OPeNDAP) servers.
- Exploits multi-core capabilities
- Collects observational and experimental datasets stored in multiple conventions:
  - NetCDF3, NetCDF4
  - File per member, file per starting date, single file, …
  - Supports specific folder and file naming conventions.

BASIC STATISTICS

SCORES
Correlation, RMSSS, CRPS, ...

PLOTS

Anomaly Correlation Coefficient. 10M Wind Speed ECMWF S4 1 month lead with start dates once a year on first of November and Era-Interim in DJF from 1981 to 2011. Simple bias correction with cross-validation.

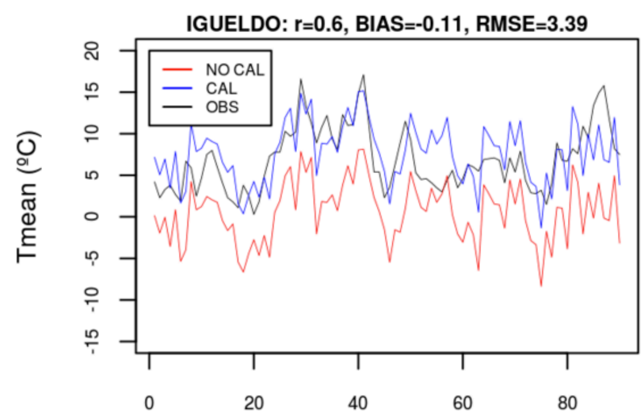


N. Manubens (IC3)

Common tools for data analysis

downscaleR is an R package for climate data access and analysis, with a special focus on empirical-statistical downscaling of daily data. It is developed by SantanderMetGroup and has been supported by the SPECS project.

- Accessing and loading (local and remote) data with package Standard data transformation (Time aggregation, Subsetting
- Regridding and interpolation)
- Principal Components, Clustering and Weather Types
- Visualizing and validating seasonal forecasts
- Statistical Downscaling Methods:
  - Bias Correction methods
  - Analog method
  - Linear regression
  - Generalized Linear Regression (GLM)



Summary: Our open fronts

Work on initialisation: initial conditions for all components (better ocean and sea ice), better ensemble generation, etc.

Model improvement: leverage knowledge and resources from modelling at other time scales, aim for a drift reduction, More efficient codes and adequate computing resources

Calibration and combination: empirical prediction (better use of current benchmarks), local knowledge

Forecast quality assessment: scores closer to user, reliability as main target, process-based verification www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you!

For further information please contact:

info-services-es@bsc.es
marco.turco@bsc.es
francisco.doblas-reyes@bsc.es

www.bsc.es

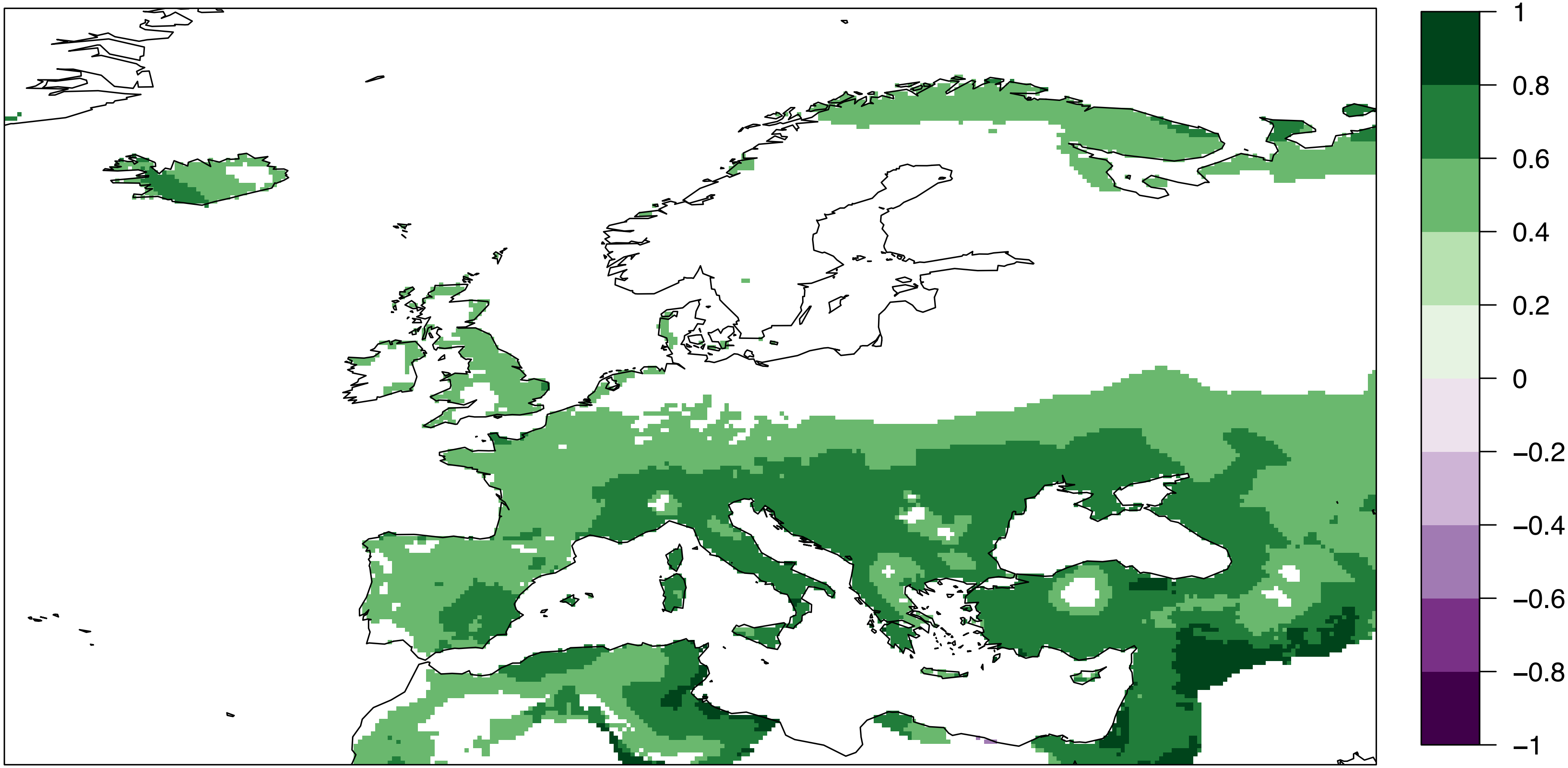
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

EXTRA SLIDES

Preliminary evaluation

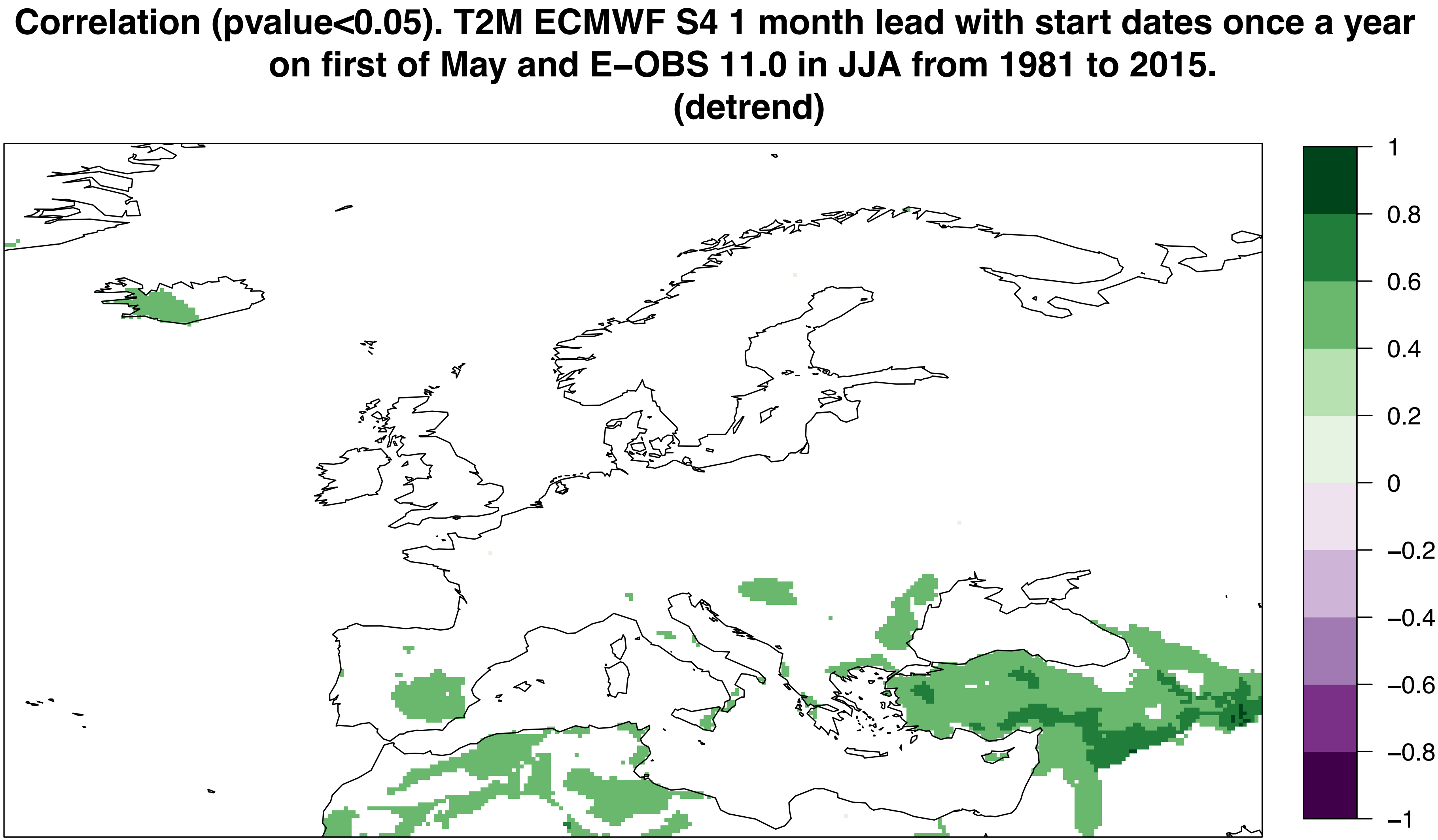
Correlation (pvalue<0.05). T2M ECMWF S4 1 month lead with start dates once a year on first of May and E–OBS 11.0 in JJA from 1981 to 2015. (raw data)



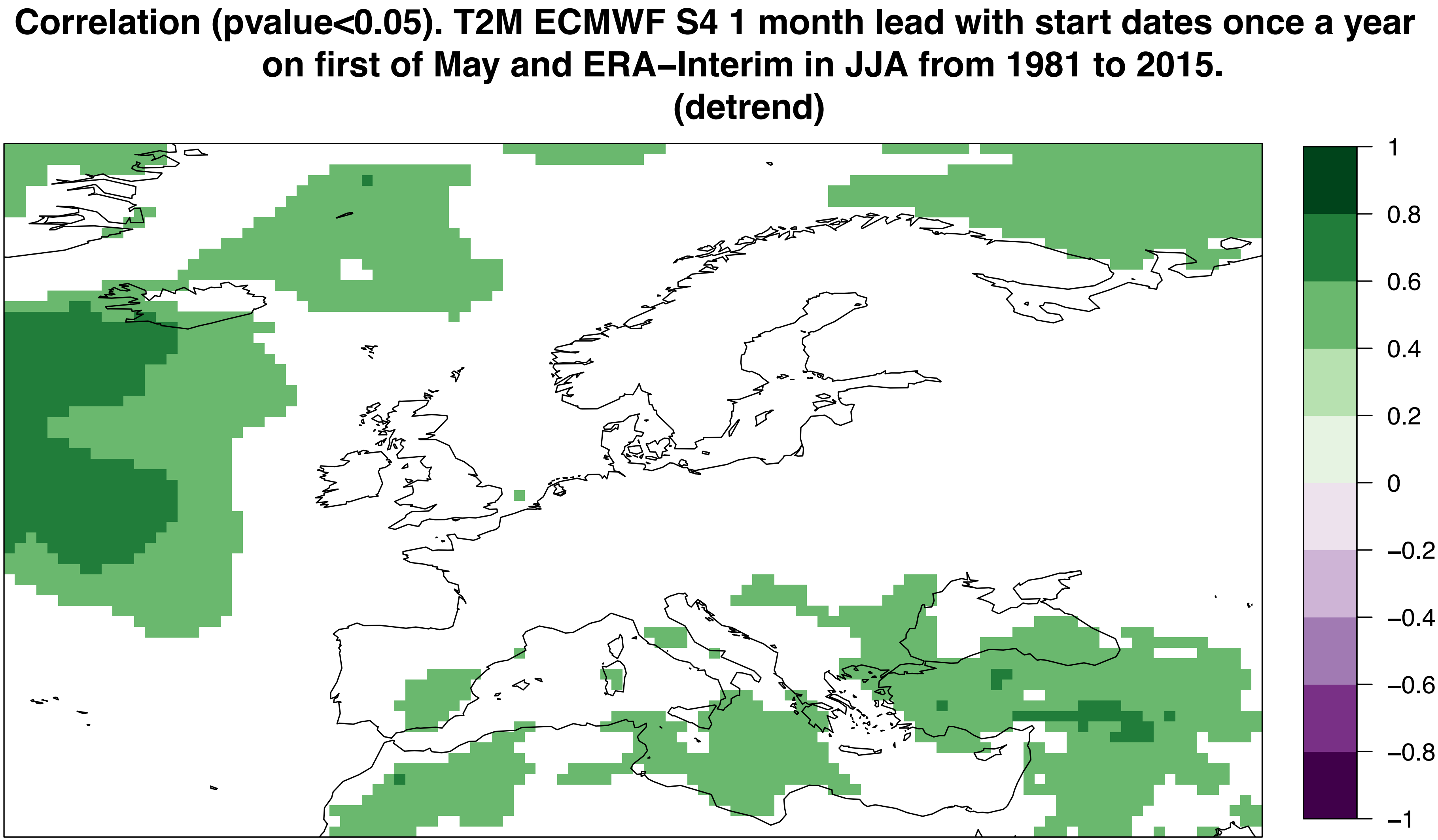
Preliminary evaluation

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

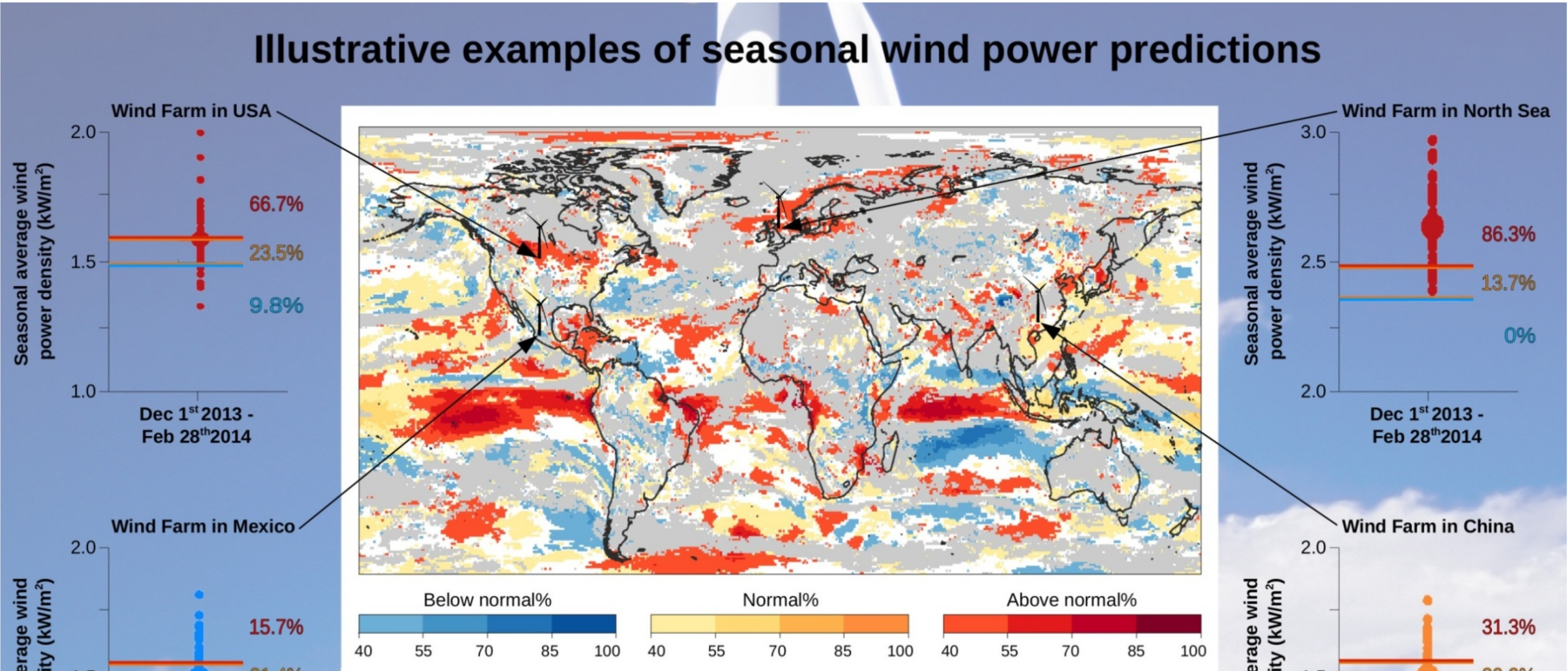
EXCELENCIA SEVERO OCHOA



Preliminary evaluation



Example: wind power predictions



Wind power prediction for December 1st 2013 - February 28th 2014, issued on November 1st 2013.

The most likely wind power category (**below normal**, **normal** or **above normal**), and its percentage probability to occur is shown. “Normal” represents the average of the past 30 years.

White areas demonstrate where the probability is <40% and approximately equal for all three categories.

Grey areas show where the climate prediction model does not improve upon the standard and current approach, which projects past climate data into the future.

How can we predict climate for the coming season if we cannot predict the weather next week?

Weather forecasts

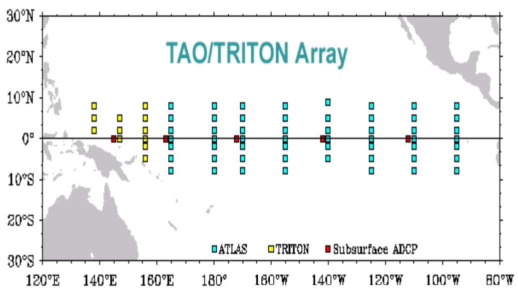
The forecasts are based in the initial conditions of the **atmosphere**, which is highly variable and develops a chaotic behaviour after a few days

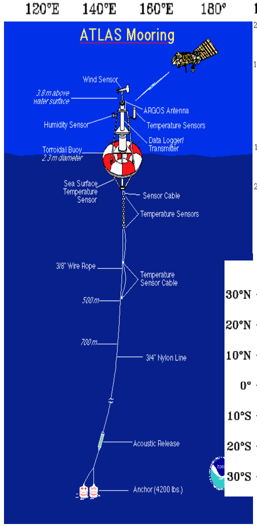
Climate predictions

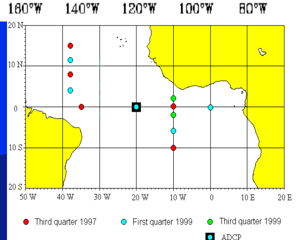
The predictions are based in the initial conditions of the **sea surface temperature**, **snow cover** or **sea ice**, which have a slow evolution that can range from few months to years.

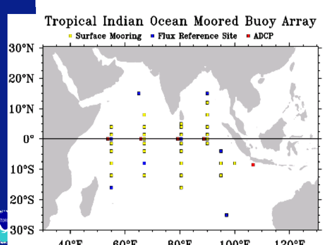
Real-time ocean observations

Moorings

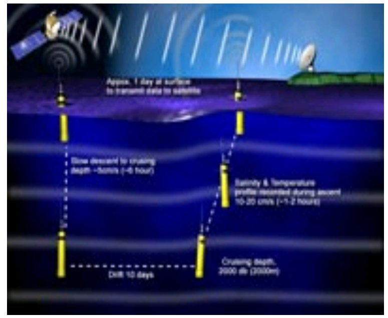


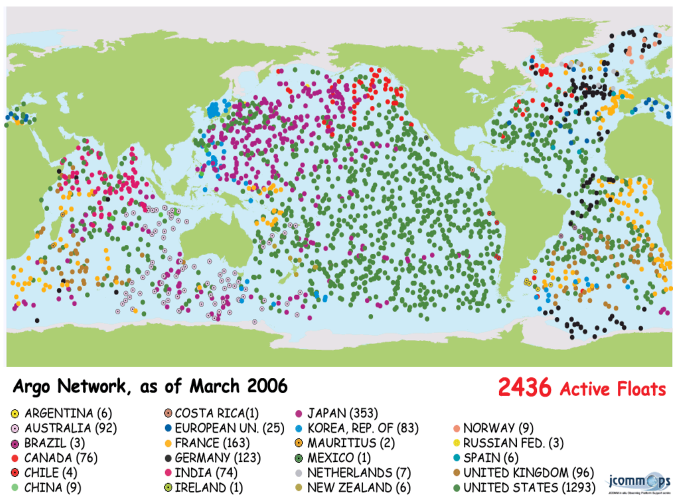






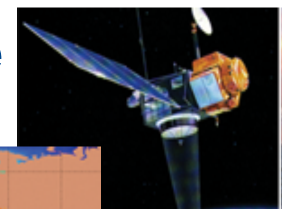
ARGO floats

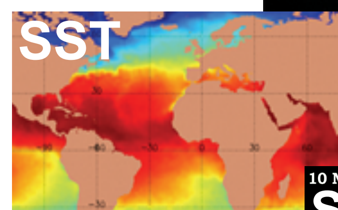


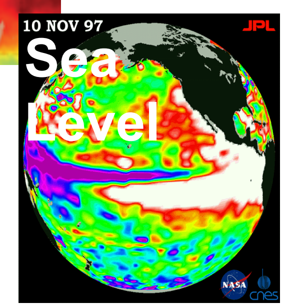


XBT (eXpendable BathiThermograph)

Satellite

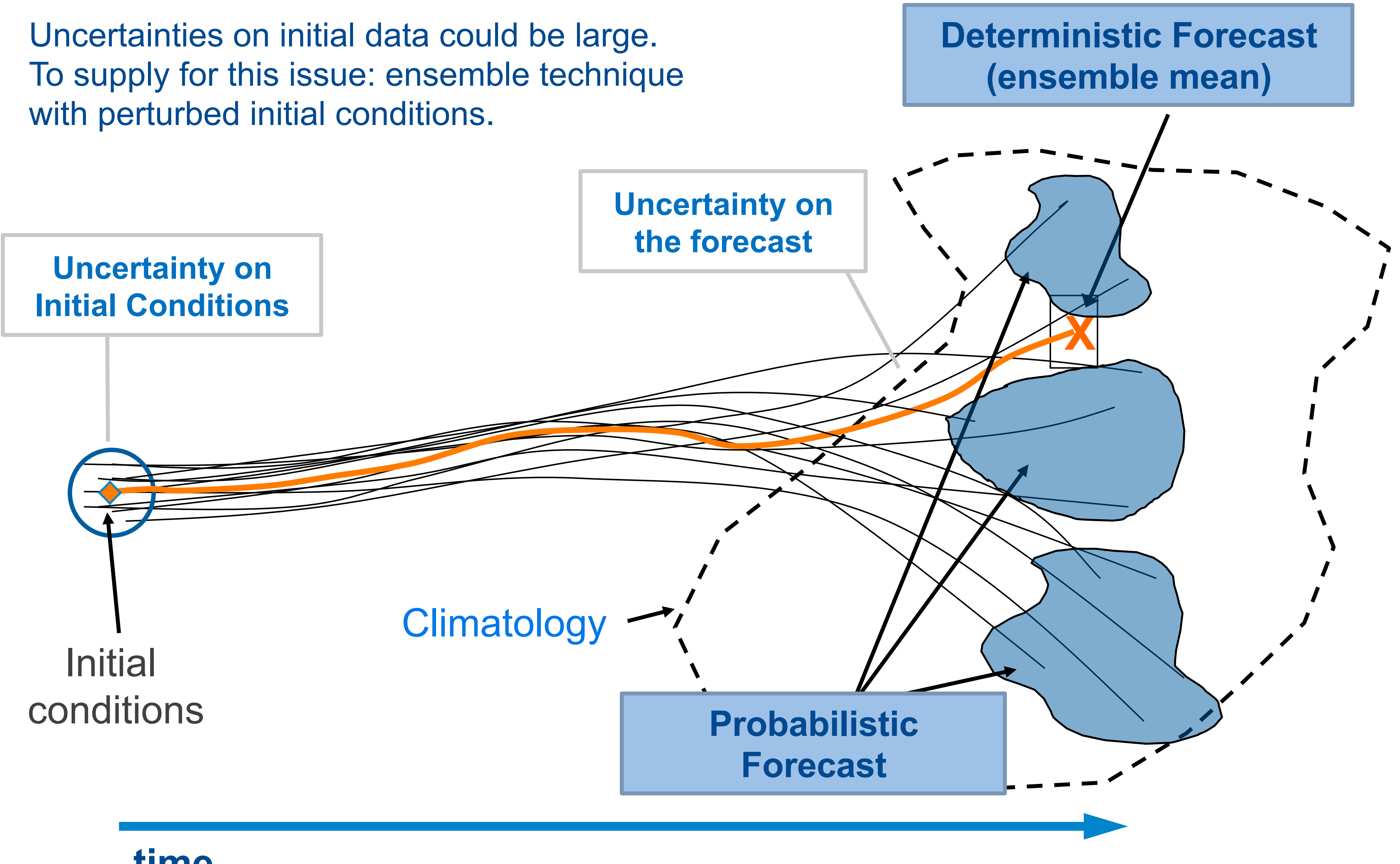






Climate predictions and predictability

Uncertainties on initial data could be large.
To supply for this issue: ensemble technique with perturbed initial conditions.

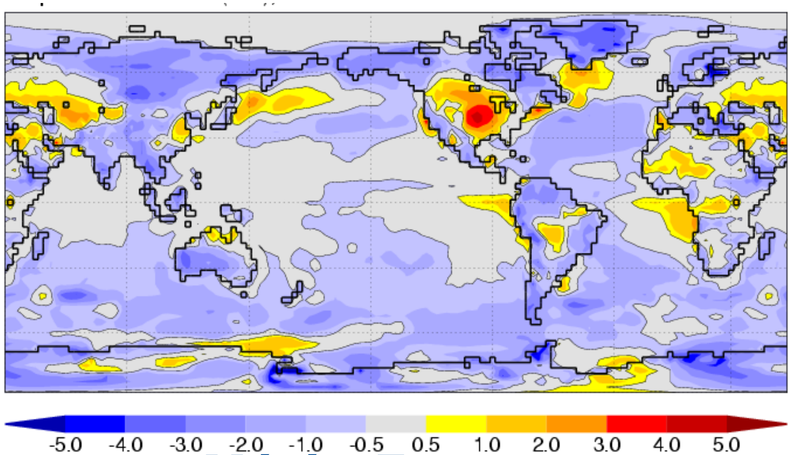


Source: S.Gualdi readapted from Trzaska (http://portal.iri.columbia.edu)

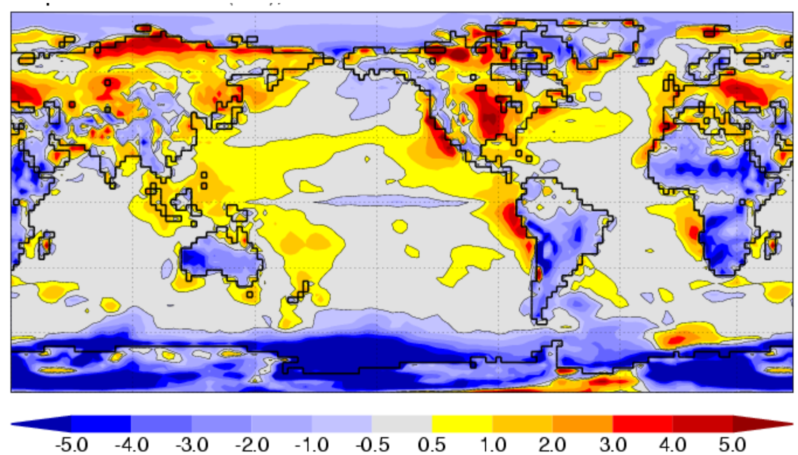
Mean error

Mean biases (JJA 2mT over 1993-2005) are often comparable in magnitude to the anomalies which we seek to predict

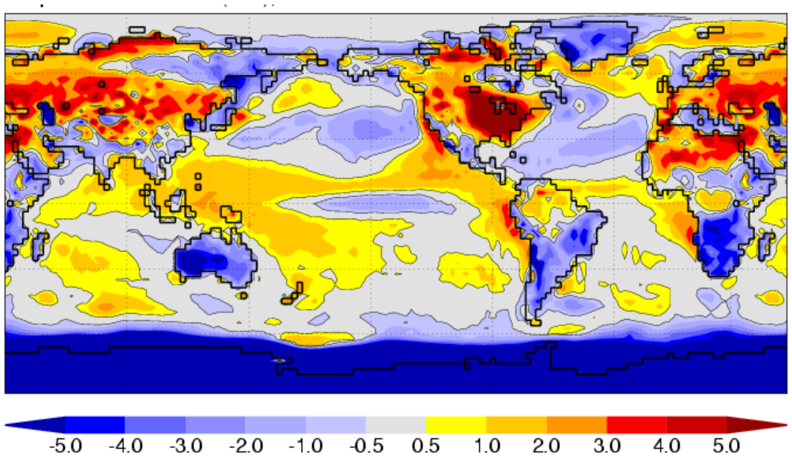
ECMWF



Met Office



Météo-France



CFS