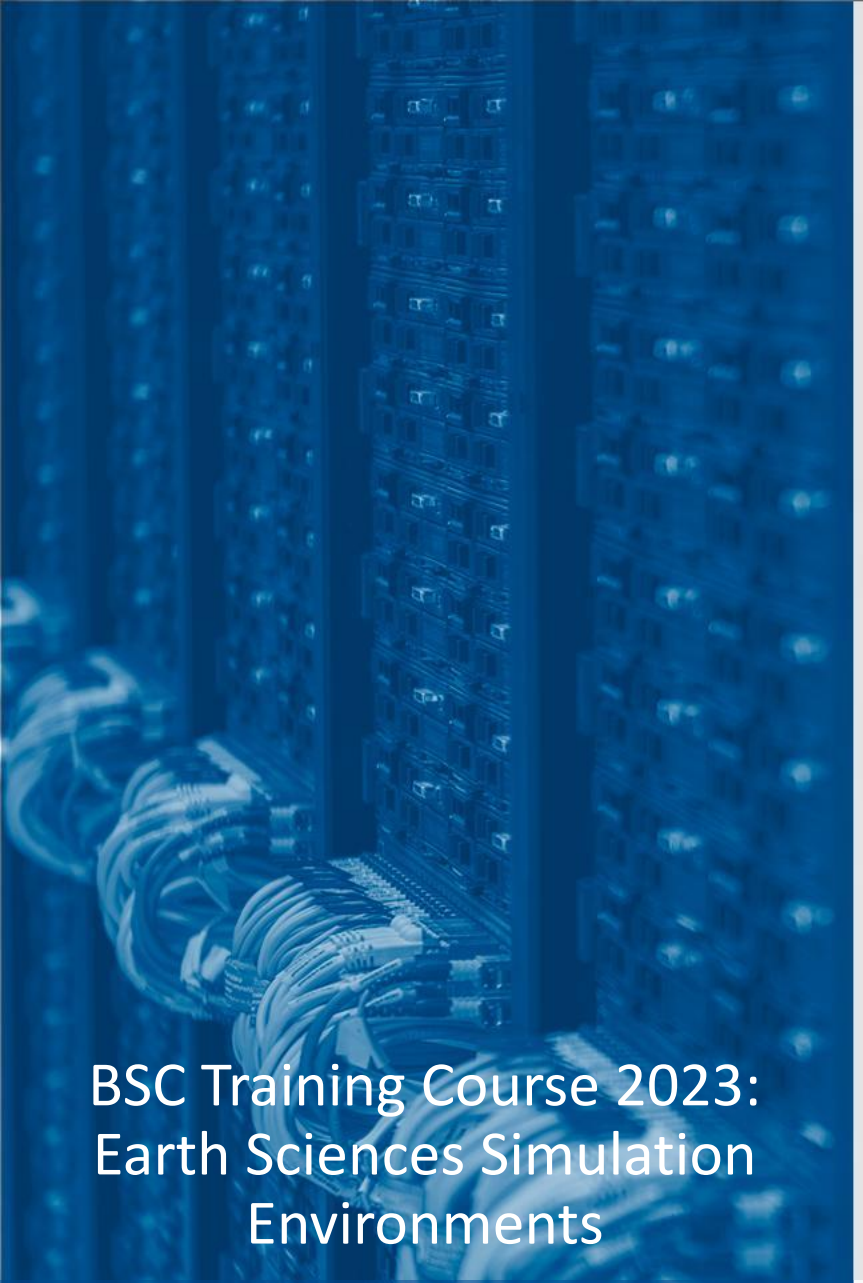


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Climate Forecast Analysis hands-on tutorial: R tools

BSC Training Course 2023: Earth Sciences Simulation Environments

31st October, online

Victòria Agudetse, Eva Rifà, An-Chi Ho

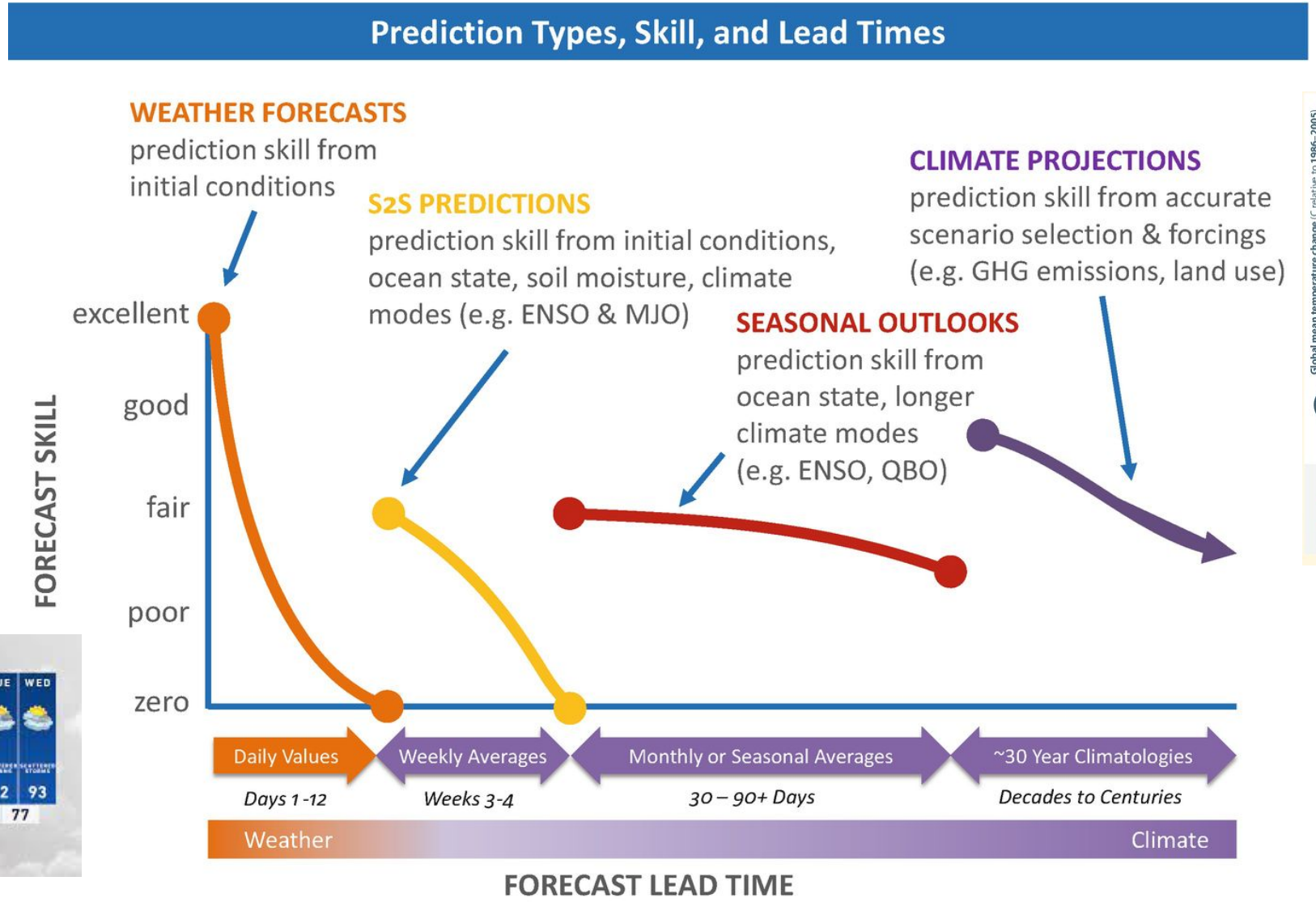
Outline

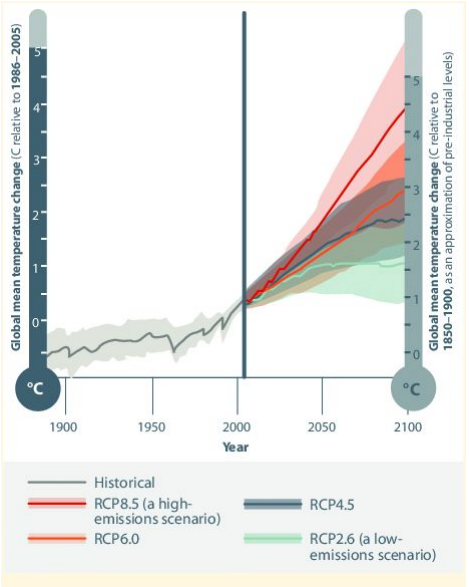
1. Introduction to Climate Forecasts
2. Introduction to the Climate Forecast Analysis Tools
3. Hands-On I: Load data by startR
4. Hands-On II: Data assessment with CSTools and s2dv
5. Hands-On III: SUNSET

1. Introduction to Climate Forecasts

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Climate Forecast: Forecast horizon





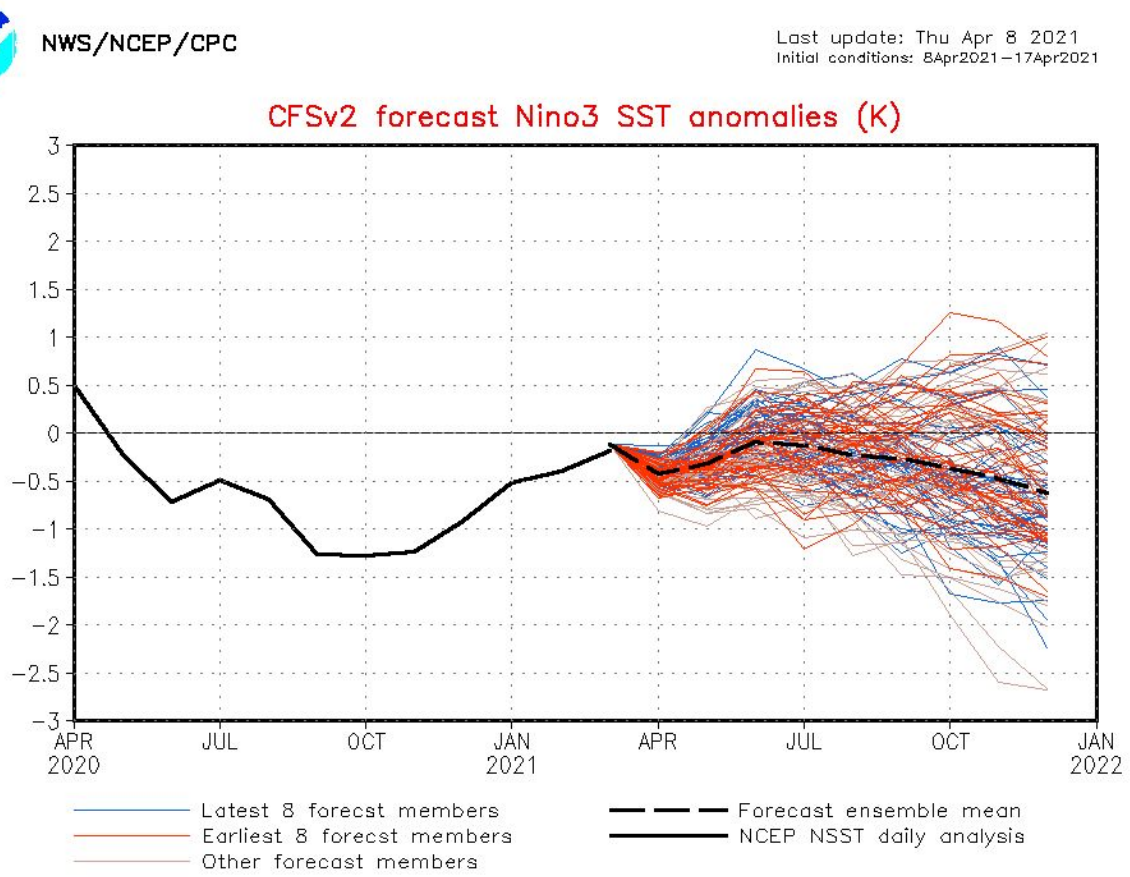
Source: DOI:10.13140/RG.2.2.21145.62564

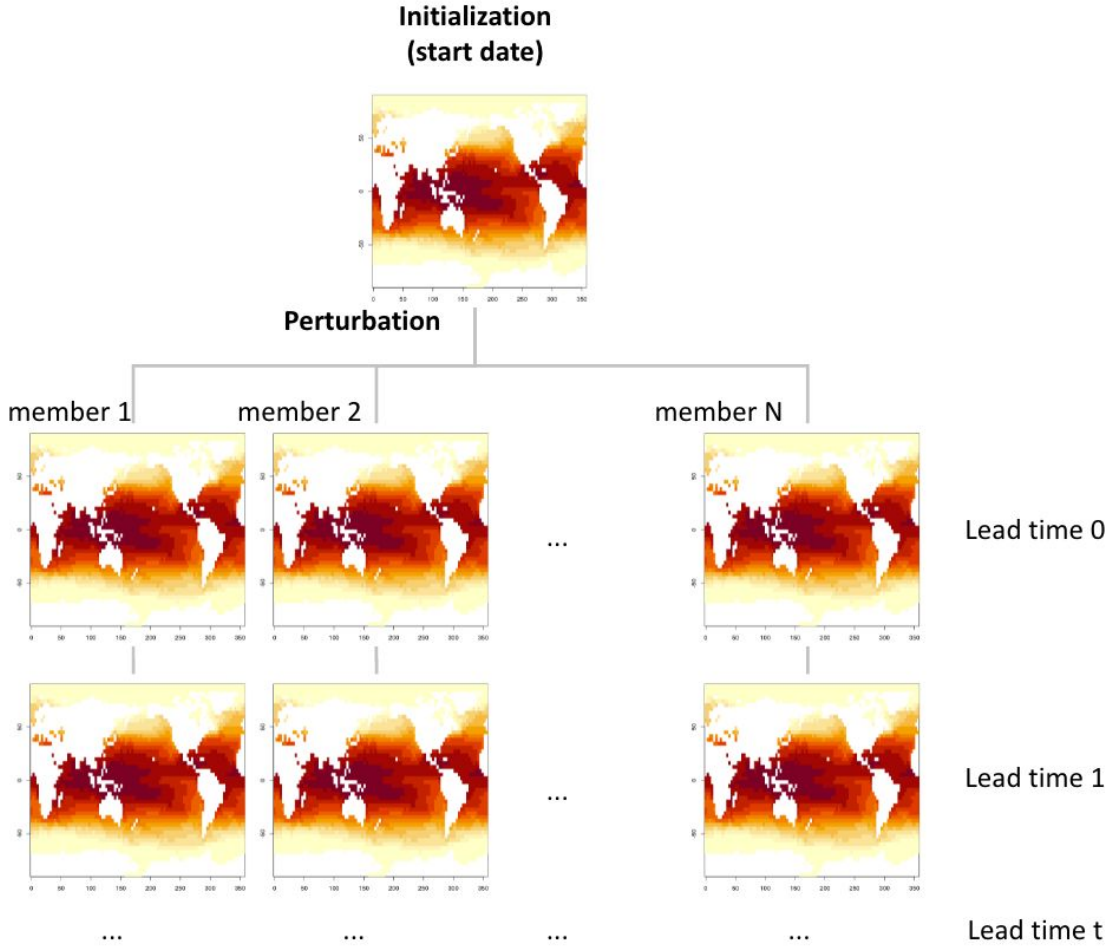
Adapted from iri.columbia.edu/news/qa-subseasonal-prediction-project

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Climate Forecast: Ensemble generation

- Different models represent the equations using different parameterizations
- Perturbations on initial conditions are included to generate a ensemble of simulations



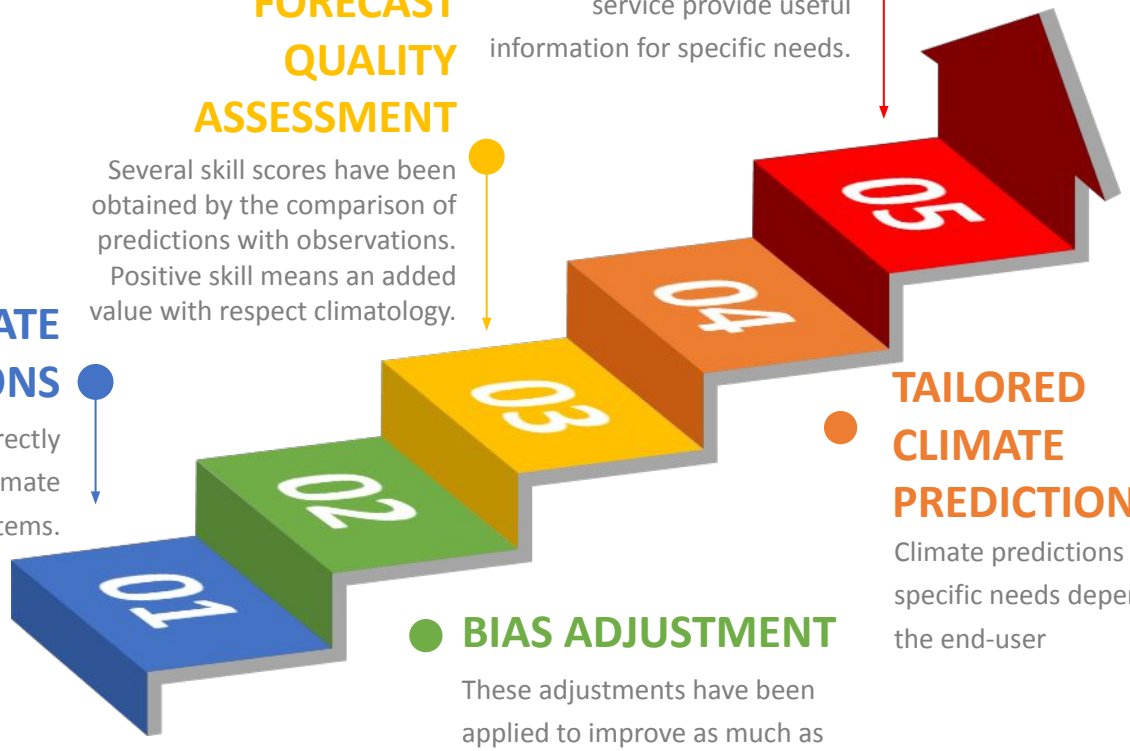


Multi-dimensional array with named dimension:

e.g.: [model = 2, sdates = 30, members = 25, ftime = 7, lat = 90, lon = 360, levels = 10]

Climate Forecast: From Climate data to Climate product

How to turn climate data into useful result or products?



RAW CLIMATE PREDICTIONS

Predictions obtained directly from different climate prediction systems.

BIAS ADJUSTMENT

These adjustments have been applied to improve as much as possible reliability of the climate predictions.

FORECAST QUALITY ASSESSMENT

Several skill scores have been obtained by the comparison of predictions with observations. Positive skill means an added value with respect climatology.

TAILORED CLIMATE PREDICTIONS

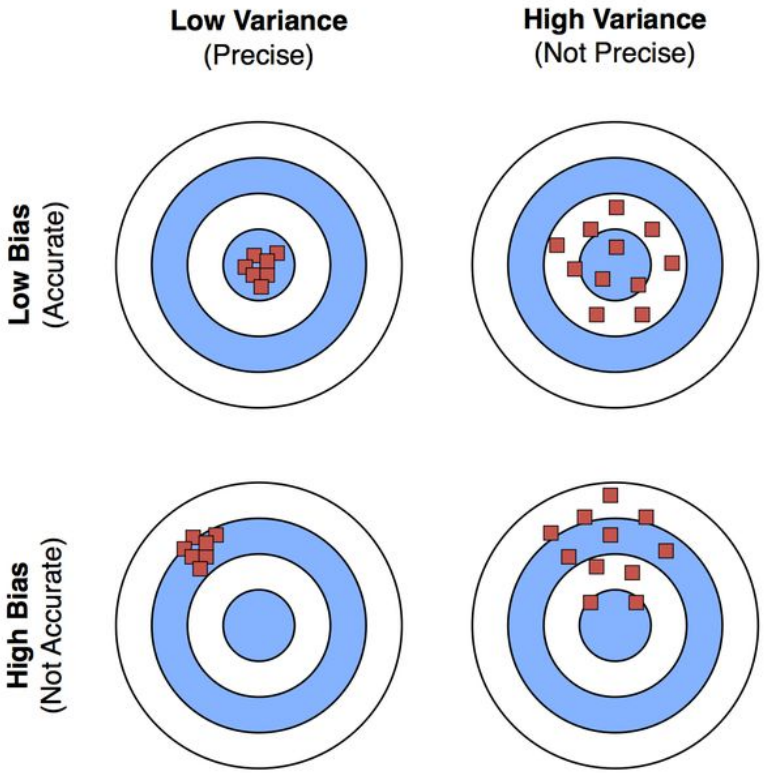
Climate predictions tailored to specific needs depending on the end-user

CLIMATE SERVICE PRODUCT

The final products from the service provide useful information for specific needs.

Climate Forecast: Bias adjustment

- The raw experimental output may have room to be improved by reference data (e.g., observation)
- Many methods can improve the quality of forecast, e.g., bias correction, variance inflation, minimized mean-squared error, etc.



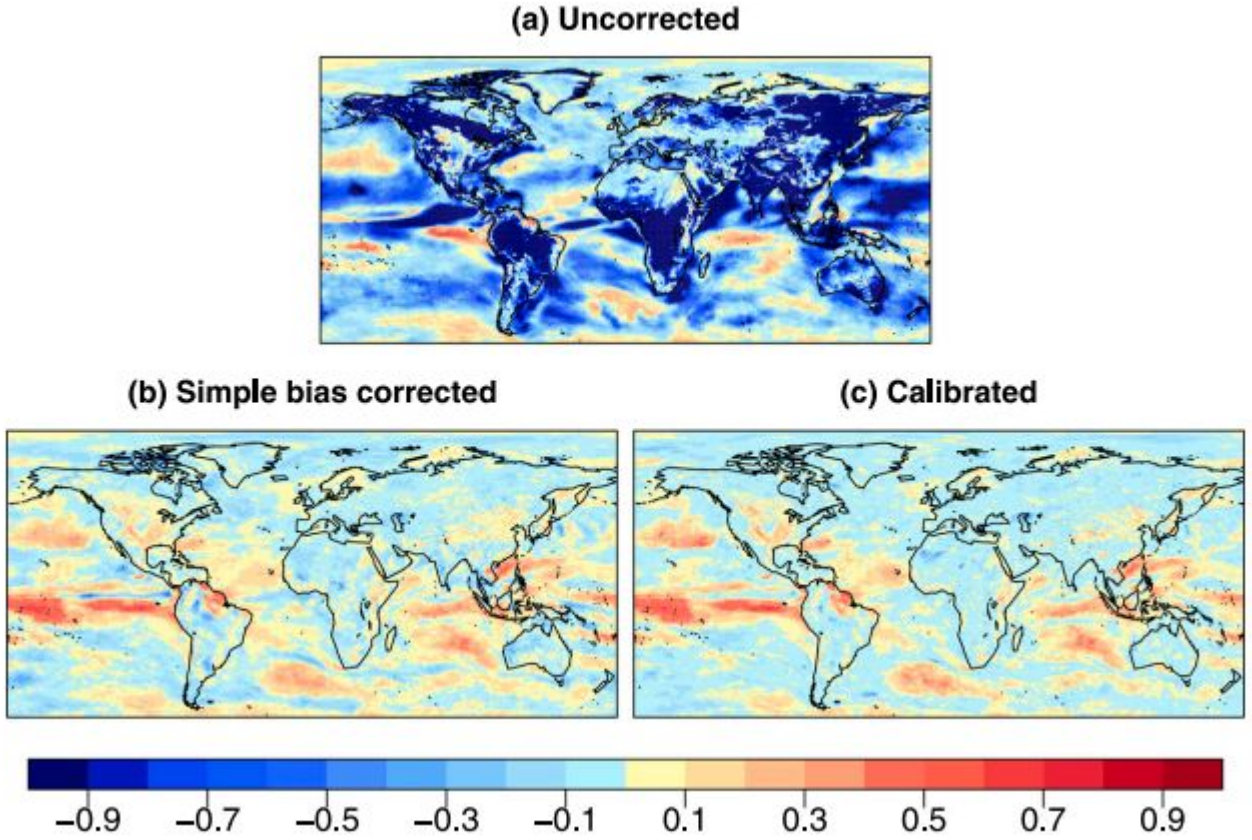


FIG. 4. Fair RPSS for tercile events of 10-m wind speed forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on 1 Nov for the period of 1981–2012.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Torralba, V., Doblas-Reyes, F. J., MacLeod, D., Christel, I., & Davis, M. (2017). Seasonal Climate Prediction: A New Source of Information for the Management of Wind Energy Resources, *Journal of Applied Meteorology and Climatology*, *56*(5), 1231-1247. Retrieved Apr 9, 2021, from https://journals.ametsoc.org/view/journals/apme/56/5/jamc-d-16-0204.1.xml

Climate Forecast Quality assessment

- When issuing a climate product, it is important to provide information about its **reliability**.
- For every initialization, a climate prediction will include two time periods: a forecast and a hindcast. The hindcast is a simulation covering a historical period, which can be compared to the observational information of the same time period to **evaluate the quality of the model**. This is called a Skill Assessment.

Near-Surface Air Temperature of ECMWF SEAS5 (Interpolation = to system, Aggregation level = score, Cross-validation = anomalies & terciles & crps_clim)

(Ref: ERA5 1993-2016)

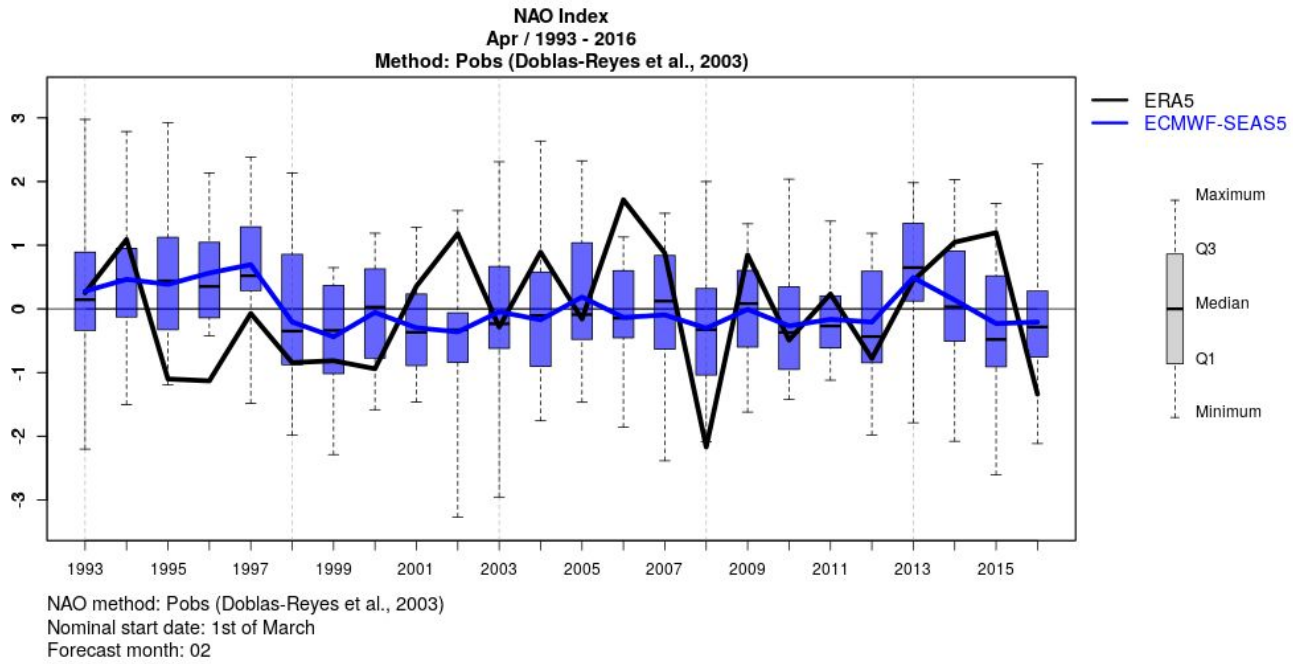
Region	Forecast Month	Mean bias (K) — Start date												Correlation — Start date												RPSS — Start date												CRPSS — Start date												Spread-to-error ratio — Start date											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Extra-tropical NH	1	-0.42	-0.17	-0.10	-0.12	-0.10	-0.08	-0.09	-0.03	0.07	-0.07	-0.24	-0.45	0.69	0.75	0.77	0.73	0.69	0.66	0.66	0.68	0.65	0.66	0.69	0.68	0.28	0.38	0.40	0.34	0.30	0.28	0.26	0.29	0.25	0.27	0.29	0.28	0.27	0.36	0.39	0.34	0.29	0.27	0.28	0.29	0.26	0.27	0.29	0.29	0.96	0.95	0.99	0.93	0.90	0.92	0.93	0.92	0.94	0.95	0.96	0.95
	2	-0.45	-0.20	-0.23	-0.24	-0.18	-0.11	-0.11	0.10	0.22	0.05	-0.37	-0.55	0.29	0.40	0.38	0.34	0.31	0.38	0.40	0.40	0.30	0.25	0.26	0.36	0.03	0.09	0.07	0.06	0.05	0.08	0.09	0.08	0.05	0.00	0.02	0.07	0.06	0.11	0.10	0.09	0.08	0.10	0.12	0.12	0.10	0.05	0.05	0.10	0.98	0.98	0.95	0.94	0.98	1.01	1.00	1.03	0.98	0.96	1.01	1.09
	3	-0.32	-0.34	-0.38	-0.37	-0.25	-0.20	0.04	0.27	0.32	-0.09	-0.49	-0.43	0.29	0.32	0.28	0.26	0.29	0.35	0.37	0.26	0.25	0.25	0.20	0.28	0.03	0.05	0.03	0.02	0.03	0.06	0.07	0.02	0.02	0.01	0.01	0.04	0.07	0.07	0.07	0.06	0.07	0.10	0.11	0.08	0.06	0.06	0.05	0.07	0.97	0.95	0.92	0.99	1.01	1.00	1.03	0.98	0.97	1.01	1.05	1.01
	4	-0.39	-0.46	-0.51	-0.42	-0.34	-0.11	0.17	0.31	0.06	-0.36	-0.42	-0.35	0.25	0.25	0.24	0.26	0.30	0.32	0.30	0.20	0.22	0.23	0.26	0.23	0.01	0.02	0.02	0.01	0.04	0.05	0.04	-0.01	0.00	0.01	0.03	0.01	0.05	0.06	0.06	0.06	0.08	0.09	0.09	0.05	0.06	0.05	0.07	0.05	0.94	0.93	0.99	1.02	0.99	1.05	1.02	0.95	1.00	1.08	1.03	0.97
	5	-0.50	-0.58	-0.54	-0.48	-0.23	-0.06	0.21	0.11	-0.20	-0.39	-0.36	-0.41	0.22	0.26	0.24	0.19	0.33	0.26	0.20	0.22	0.21	0.21	0.23	0.22	-0.01	0.02	0.00	-0.01	0.05	0.02	0.00	0.00	0.01	0.00	0.02	0.00	0.05	0.07	0.05	0.04	0.09	0.08	0.05	0.06	0.06	0.05	0.06	0.05	0.93	1.00	1.02	0.97	1.06	1.03	0.96	1.00	1.07	1.02	0.99	0.94
	6	-0.62	-0.62	-0.60	-0.35	-0.18	-0.01	0.00	-0.19	-0.26	-0.32	-0.43	-0.50	0.25	0.29	0.21	0.26	0.25	0.25	0.21	0.22	0.23	0.22	0.22	0.20	0.01	0.03	-0.01	0.01	0.02	0.02	0.00	0.00	0.02	0.01	0.00	-0.01	0.06	0.08	0.04	0.06	0.08	0.07	0.05	0.06	0.06	0.05	0.05	0.05	1.01	1.05	0.98	1.04	1.03	1.01	1.01	1.07	1.01	0.98	0.96	0.94
Tropics	1	-0.20	-0.21	-0.16	-0.15	-0.11	-0.16	-0.18	-0.18	-0.16	-0.16	-0.21	-0.23	0.81	0.86	0.83	0.83	0.80	0.80	0.78	0.80	0.79	0.79	0.78	0.79	0.45	0.51	0.46	0.44	0.42	0.43	0.40	0.41	0.41	0.41	0.40	0.42	0.43	0.50	0.45	0.42	0.41	0.40	0.39	0.42	0.41	0.42	0.40	0.42	0.96	0.93	0.92	0.95	0.91	0.86	0.88	0.89	0.89	0.92	0.92	0.92
	2	-0.24	-0.20	-0.16	-0.09	-0.10	-0.14	-0.14	-0.15	-0.17	-0.19	-0.24	-0.24	0.66	0.67	0.68	0.63	0.62	0.63	0.65	0.63	0.65	0.59	0.62	0.63	0.27	0.28	0.28	0.25	0.25	0.24	0.27	0.24	0.27	0.23	0.26	0.27	0.26	0.27	0.27	0.24	0.24	0.24	0.27	0.26	0.28	0.24	0.27	0.26	0.98	0.99	1.04	0.98	0.97	0.95	0.97	0.94	0.99	0.99	1.02	1.02
	3	-0.19	-0.18	-0.11	-0.07	-0.09	-0.12	-0.13	-0.18	-0.18	-0.22	-0.22	-0.25	0.60	0.63	0.58	0.56	0.58	0.58	0.59	0.61	0.56	0.55	0.59	0.61	0.21	0.23	0.19	0.20	0.20	0.19	0.21	0.23	0.20	0.21	0.24	0.23	0.21	0.22	0.19	0.20	0.20	0.21	0.22	0.24	0.21	0.23	0.23	0.23	1.00	1.05	0.99	0.99	0.99	0.96	0.95	0.99	1.02	1.02	1.06	0.99
	4	-0.20	-0.15	-0.08	-0.06	-0.08	-0.11	-0.16	-0.20	-0.22	-0.22	-0.23	-0.22	0.59	0.55	0.53	0.56	0.54	0.54	0.59	0.54	0.54	0.53	0.59	0.57	0.20	0.18	0.16	0.18	0.15	0.16	0.20	0.19	0.19	0.19	0.22	0.18	0.19	0.17	0.16	0.19	0.18	0.19	0.22	0.20	0.21	0.20	0.22	0.19	1.06	1.01	0.99	1.02	0.97	0.96	1.01	1.03	1.03	1.03	1.02	1.00
	5	-0.19	-0.13	-0.08	-0.06	-0.10	-0.15	-0.19	-0.25	-0.22	-0.24	-0.21	-0.22	0.52	0.50	0.53	0.52	0.53	0.55	0.51	0.52	0.54	0.53	0.55	0.55	0.14	0.14	0.14	0.13	0.15	0.16	0.15	0.17	0.19	0.17	0.18	0.16	0.14	0.15	0.16	0.16	0.17	0.19	0.17	0.20	0.20	0.18	0.18	0.16	0.99	0.99	1.01	0.99	0.99	1.00	1.02	1.02	1.05	0.99	1.03	1.03
	6	-0.16	-0.12	-0.08	-0.07	-0.14	-0.19	-0.24	-0.24	-0.24	-0.22	-0.22	-0.20	0.46	0.51	0.50	0.50	0.53	0.49	0.49	0.52	0.55	0.51	0.55	0.48	0.11	0.13	0.11	0.13	0.15	0.14	0.15	0.17	0.18	0.14	0.16	0.12	0.11	0.15	0.15	0.15	0.18	0.16	0.17	0.19	0.19	0.15	0.17	0.12	0.98	1.02	0.98	1.00	1.02	1.03	1.02	1.04	1.01	1.00	1.06	0.99
Extra-tropical SH	1	0.07	0.09	0.12	0.05	-0.05	-0.12	0.00	-0.07	-0.03	0.04	-0.07	0.01	0.72	0.74	0.68	0.69	0.68	0.64	0.64	0.68	0.67	0.69	0.72	0.71	0.32	0.36	0.29	0.30	0.29	0.24	0.24	0.27	0.30	0.29	0.34	0.30	0.31	0.34	0.27	0.28	0.29	0.24	0.26	0.27	0.26	0.29	0.33	0.31	0.96	0.99	0.95	1.01	1.05	1.01	0.94	0.98	0.96	1.02	0.98	0.86
	2	0.26	0.31	0.25	0.02	-0.15	0.02	-0.03	-0.01	0.02	-0.11	-0.06	0.10	0.46	0.45	0.40	0.33	0.28	0.34	0.34	0.32	0.40	0.39	0.37	0.45	0.12	0.11	0.09	0.05	0.03	0.06	0.05	0.05	0.09	0.09	0.07	0.11	0.13	0.13	0.10	0.07	0.06	0.07	0.07	0.06	0.10	0.11	0.10	0.12	1.00	1.00	0.99	1.01	1.03	1.01	1.00	0.93	1.03	0.95	0.90	1.01
	3	0.51	0.45	0.17	-0.12	-0.11	-0.09	-0.03	-0.02	-0.17	-0.14	0.03	0.31	0.38	0.37	0.29	0.23	0.27	0.24	0.26	0.31	0.33	0.31	0.38	0.39	0.07	0.07	0.03	0.00	0.03	0.00	0.02	0.04	0.06	0.05	0.08	0.08	0.09	0.09	0.07	0.05	0.06	0.03	0.05	0.08	0.09	0.09	0.09	0.10	1.00	0.98	1.00	1.03	0.99	0.96	0.93	0.99	0.92	0.91	1.05	0.99
	4	0.66	0.37	-0.02	-0.08	-0.19	-0.10	-0.07	-0.21	-0.20	-0.03	0.26	0.57	0.34	0.27	0.24	0.23	0.21	0.19	0.31	0.30	0.33	0.30	0.32	0.34	0.06	0.02	0.01	-0.01	-0.01	-0.03	0.04	0.03	0.05	0.04	0.04	0.05	0.08	0.06	0.05	0.03	0.03	0.02	0.07	0.08	0.10	0.07	0.07	0.08	0.98	0.99	1.03	0.98	0.96	0.91	1.00	0.94	0.93	1.03	0.99	0.97
	5	0.56	0.16	0.00	-0.16	-0.16	-0.10	-0.25	-0.22	-0.06	0.23	0.55	0.76	0.23	0.24	0.22	0.22	0.23	0.27	0.28	0.28	0.29	0.26	0.31	0.28	0.00	0.01	-0.01	-0.01	0.01	0.01	0.03	0.03	0.04	0.01	0.03	0.02	0.05	0.05	0.04	0.05	0.05	0.06	0.07	0.08	0.07	0.06	0.07	0.05	0.96	1.02	0.97	0.98	0.96	0.98	0.93	0.92	1.04	0.97	0.99	0.95
	6	0.32	0.13	-0.12	-0.14	-0.14	-0.26	-0.25	-0.09	0.19	0.53	0.77	0.66	0.20	0.24	0.23	0.24	0.29	0.30	0.26	0.28	0.23	0.27	0.28	0.20	0.00	0.00	0.00	0.01	0.03	0.03	0.02	0.02	0.00	0.02	0.02	-0.01	0.05	0.04	0.04	0.05	0.07	0.08	0.07	0.06	0.05	0.06	0.05	0.03	0.99	0.97	0.99	0.96	1.02	0.94	0.92	1.05	0.99	0.99	0.95	0.96

-0.56 -0.45 -0.34 -0.23 -0.11 0 0.11 0.23 0.34 0.45 0.56 | -1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1 | -0.5 -0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5 | -0.5 -0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5 | 0 0.6 0.7 0.8 0.9 1 1.2 1.4 1.6 1.8 2

‘Scorecard’ plot showcasing different prediction skill metrics (credit: Nadia Milders)

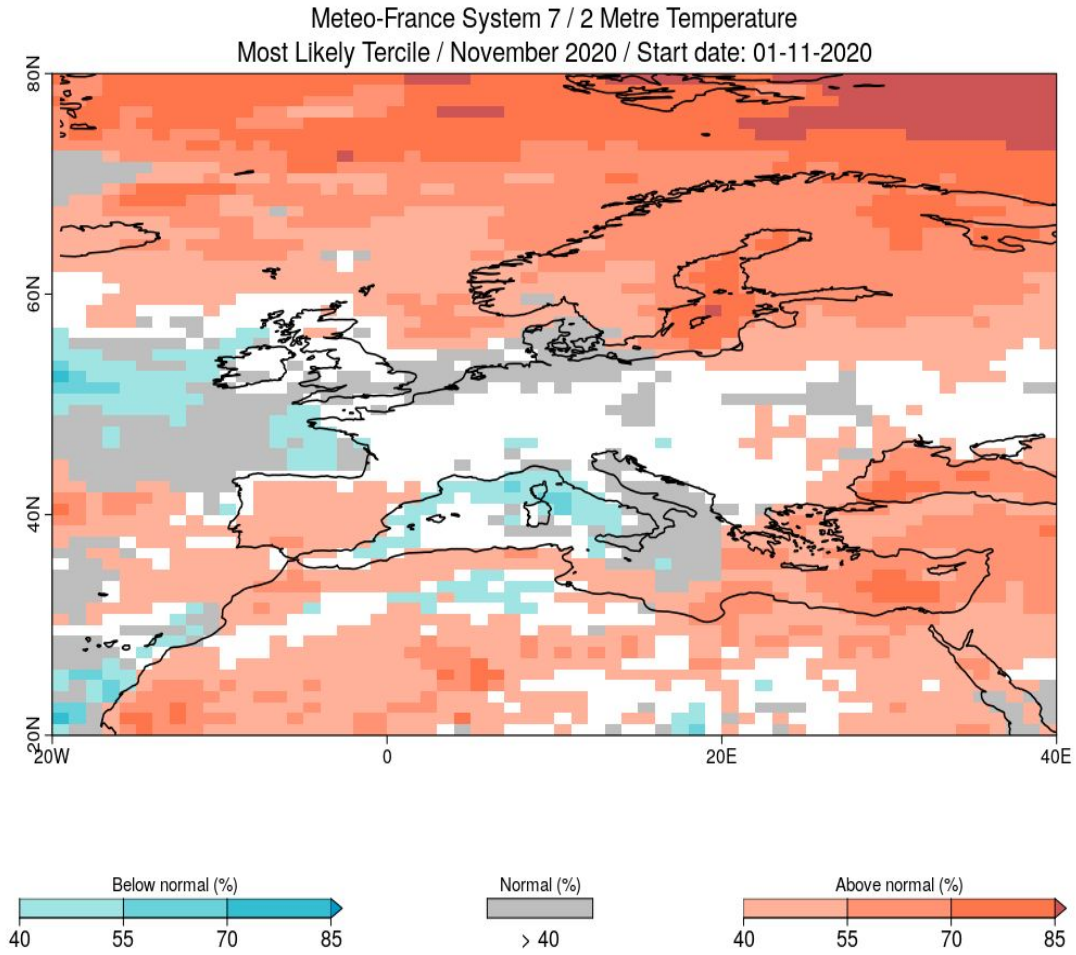
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Tailored climate prediction



NAO Index time series box plot

(credit: Núria Pérez-Zanón)

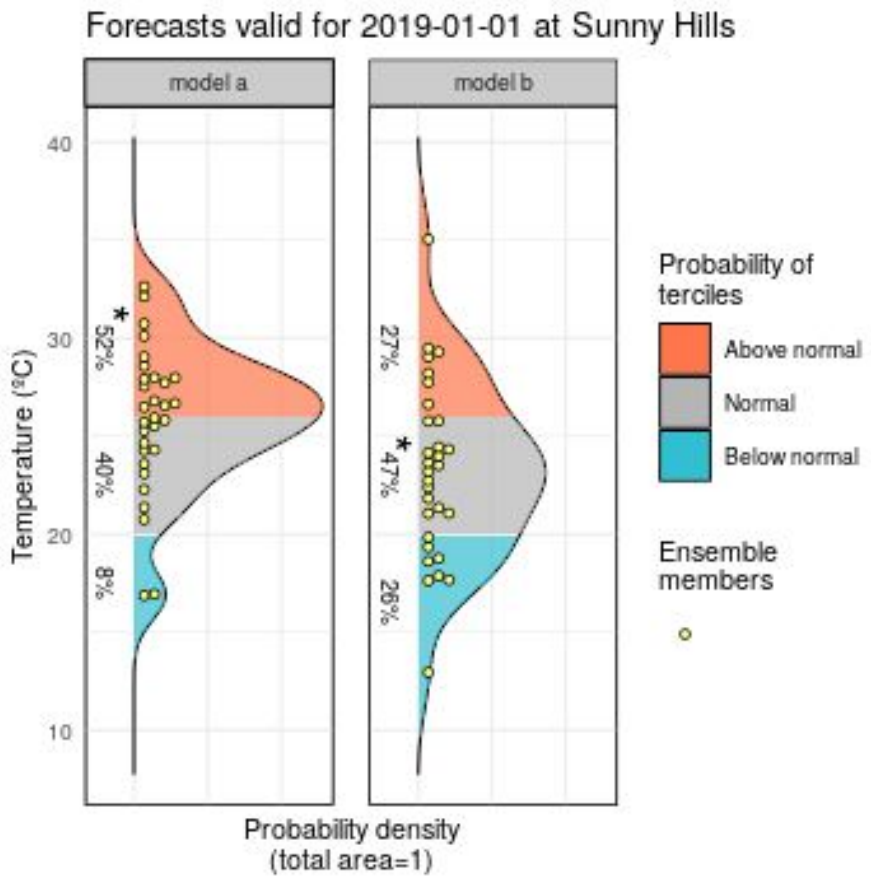


Map of the Most Likely Terciles

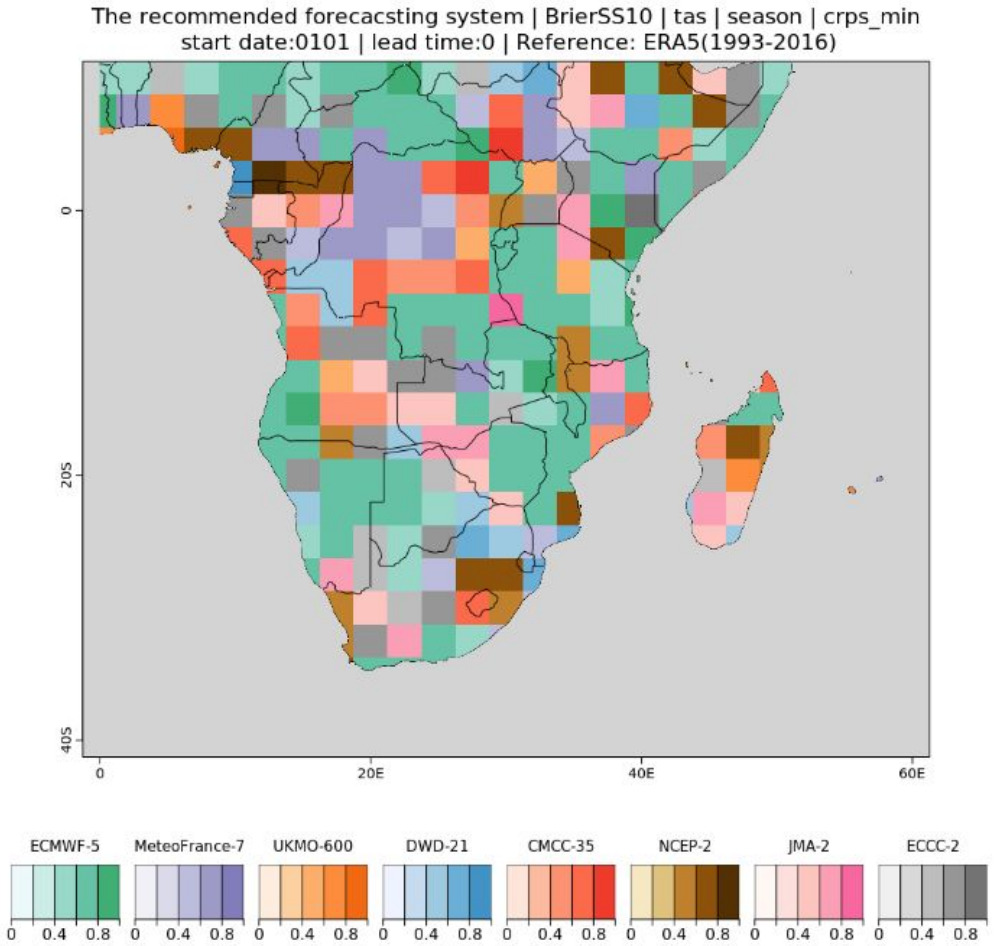
(by `CSTools::PlotMostLikelyQuantileMap`)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Tailored climate prediction



Probability distribution of ensemble forecast
(by `CSTools::PlotForecastPDF.R`)



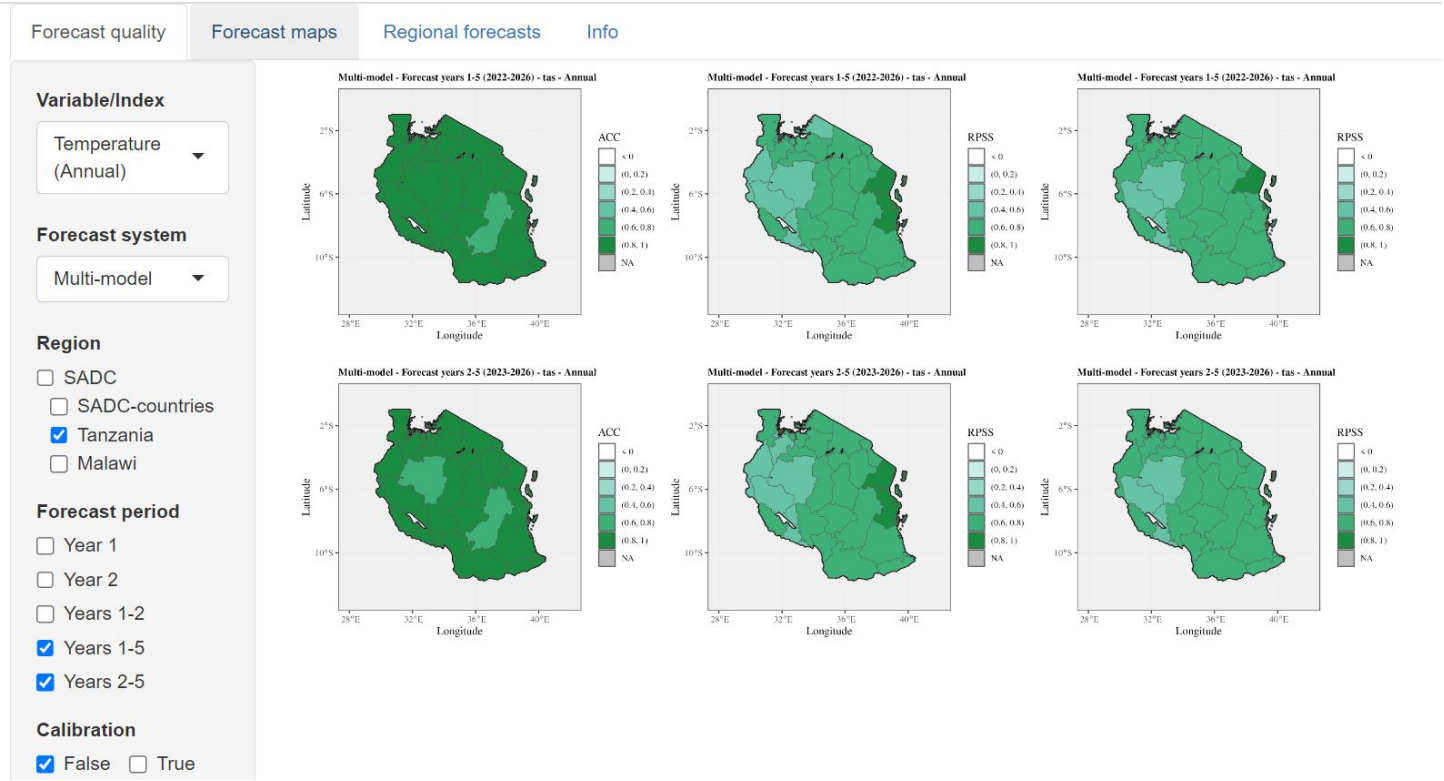
Recommended forecast system based on Brier skill score
(credit: Chihchung Chou)

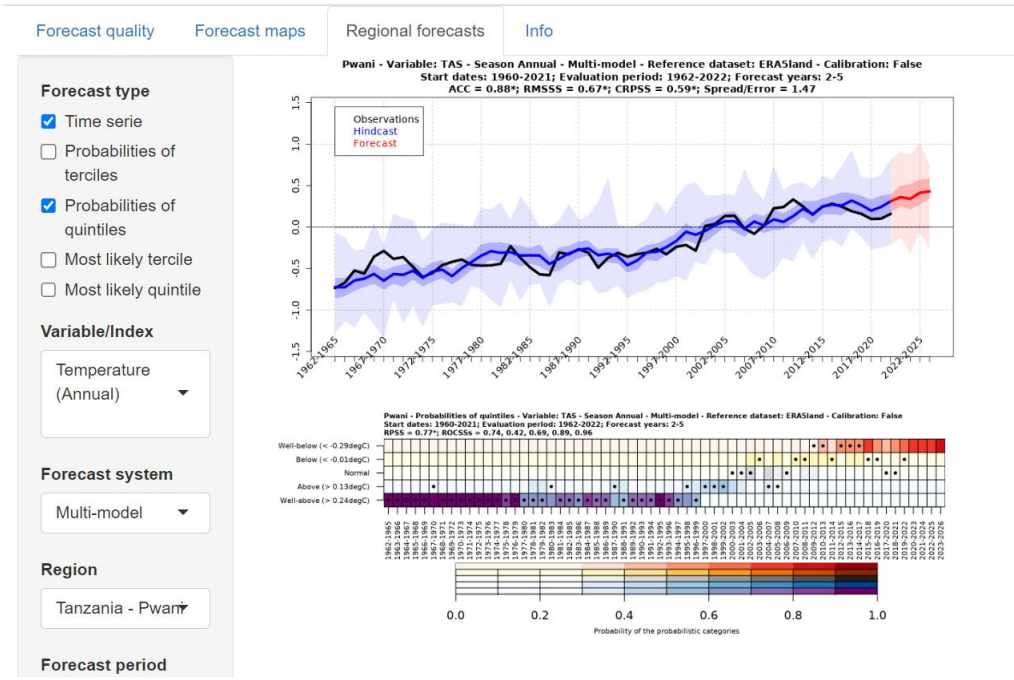
Climate service product: Shiny app

Shiny is an R package that makes it easy to build **interactive web apps** straight from R & Python.

Example: https://earth.bsc.es/shiny/cdelgado_FOCUS-Africa-casestudy/ (credit: Carlos Delgado)

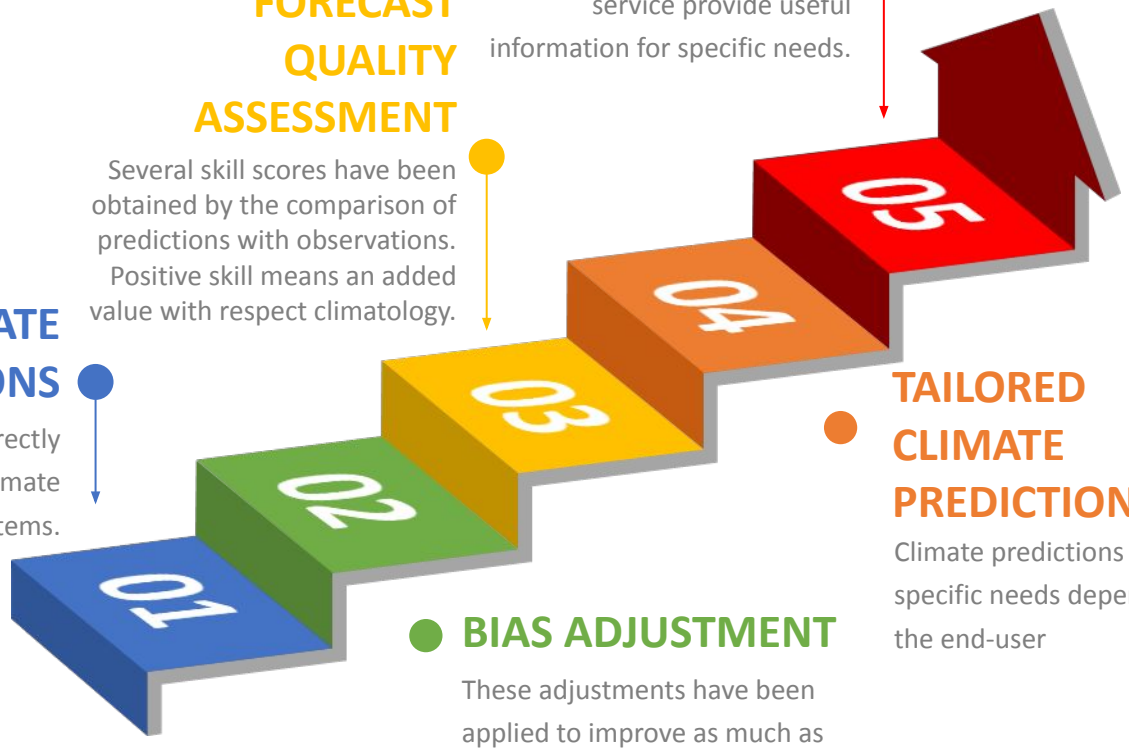
- `user: focus-africa-casestudy`
- `password: climateservices`





Climate Forecast: From Climate data to Climate product

How to turn climate data into useful result or products?



RAW CLIMATE PREDICTIONS

Predictions obtained directly from different climate prediction systems.

BIAS ADJUSTMENT

These adjustments have been applied to improve as much as possible reliability of the climate predictions.

FORECAST QUALITY ASSESSMENT

Several skill scores have been obtained by the comparison of predictions with observations. Positive skill means an added value with respect climatology.

TAILORED CLIMATE PREDICTIONS

Climate predictions tailored to specific needs depending on the end-user

CLIMATE SERVICE PRODUCT

The final products from the service provide useful information for specific needs.

We need TOOLs for data processing.

2. Introduction to Climate Forecast Analysis Tools

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

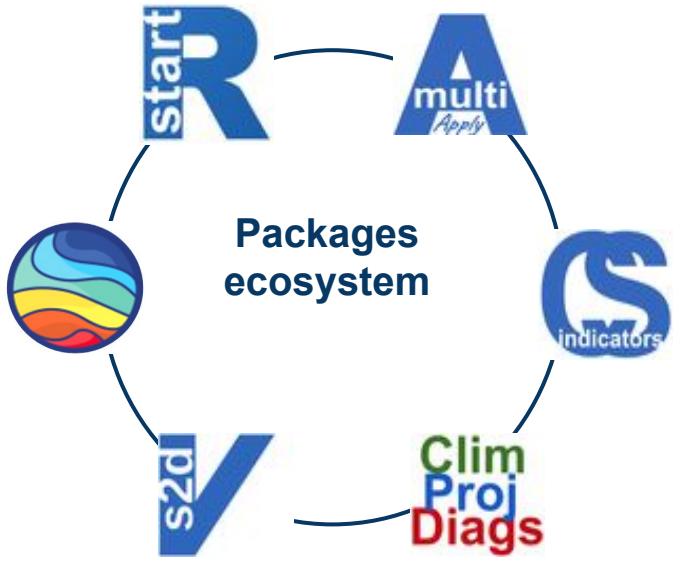
R programming language

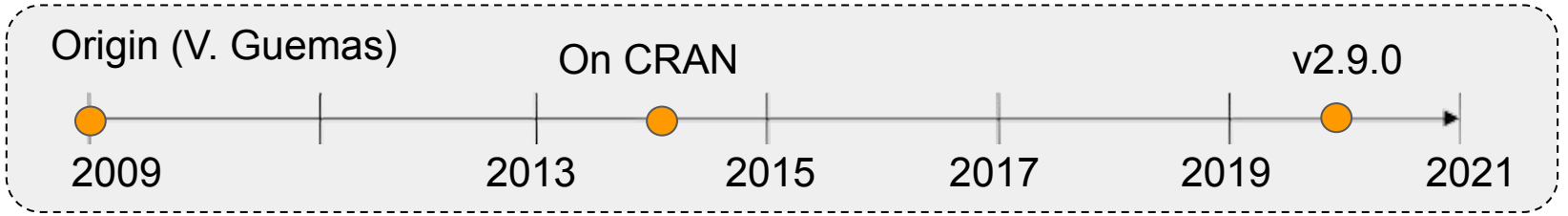
R is a generic programming language, especially features in a strong framework for statistical computing and graphics.

- Free software (under the GNU GPL license.) You can install R from https://cran.r-project.org/
- Provides a wide variety of statistical techniques (linear and non-linear modelling, classical statistical tests, classification and simulation...)
- Well developed plotting tools (e.g., ggplot2)
- Once R is installed, the base R packages are installed along. But you can further install other packages. See the list of all the available packages on CRAN https://cran.r-project.org/web/packages/available_packages_by_name.html

Climate Forecast Analysis Tools: Early version

- ★ Methods developed by the department to assess the quality of the forecast were gathered in s2dverfication R package
- ★ Researchers could easily share their methods and replicate colleagues analysis on their own data
- ★ Common needs detected





***New packages and tools** for different needs are being **developed continuously**.*

Climate Forecast Analysis Tools: Table of packages

★ Functions are split into packages depending on their objective or project requirement

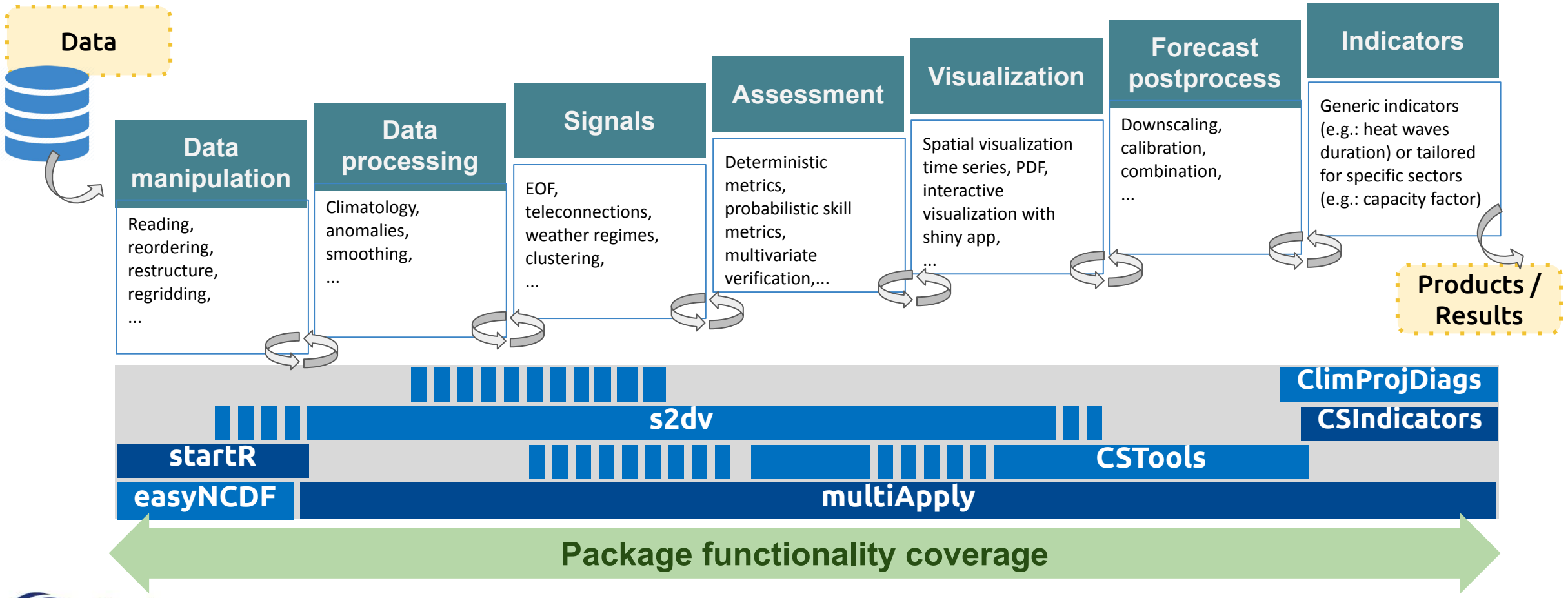
★ Functions from different packages can be used together to perform analyses or obtain climate service products

	Package name	Short description	Link to CRAN and GitLab
Data loading and manipulation	**easyNCDF**	Read/write netCDF files into/from multidimensional R array	CRAN.R-project.org/package=easyNCDF earth.bsc.es/gitlab/es/easyNCDF
	startR	Data retrieval and processing tools	CRAN.R-project.org/package=startR earth.bsc.es/gitlab/es/startR
	multiApply	Apply functions to multiple multidimensional arrays or vectors allowing parallel computation	CRAN.R-project.org/package=multiApply earth.bsc.es/gitlab/ces/multiApply
Analysis and processing	**s2dv**	Functions for Forecast Verification and visualization	CRAN.R-project.org/package=s2dv earth.bsc.es/gitlab/es/s2dv
	CSTools	Methods for forecast calibration, statistical and stochastic downscaling, optimal forecast combination and tools to obtain tailored products.	CRAN.R-project.org/package=CSTools earth.bsc.es/gitlab/external/cstools
Climate indicators	**CSIndicators**	Sectorial Indicators for Climate Service	CRAN.R-project.org/package=CSIndicators earth.bsc.es/gitlab/es/csindicators
	ClimProjDiags	Climate extreme indices, evaluation of the agreement between models, weight and combination functions.	CRAN.R-project.org/package=ClimProjDiags earth.bsc.es/gitlab/es/ClimProjDiags

Climate Forecast Data Analysis Procedure and Tools

★ The package ecosystem aims to cover the whole data analysis cycle.

★ The tools can be used interchangeably, depending on the needs.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

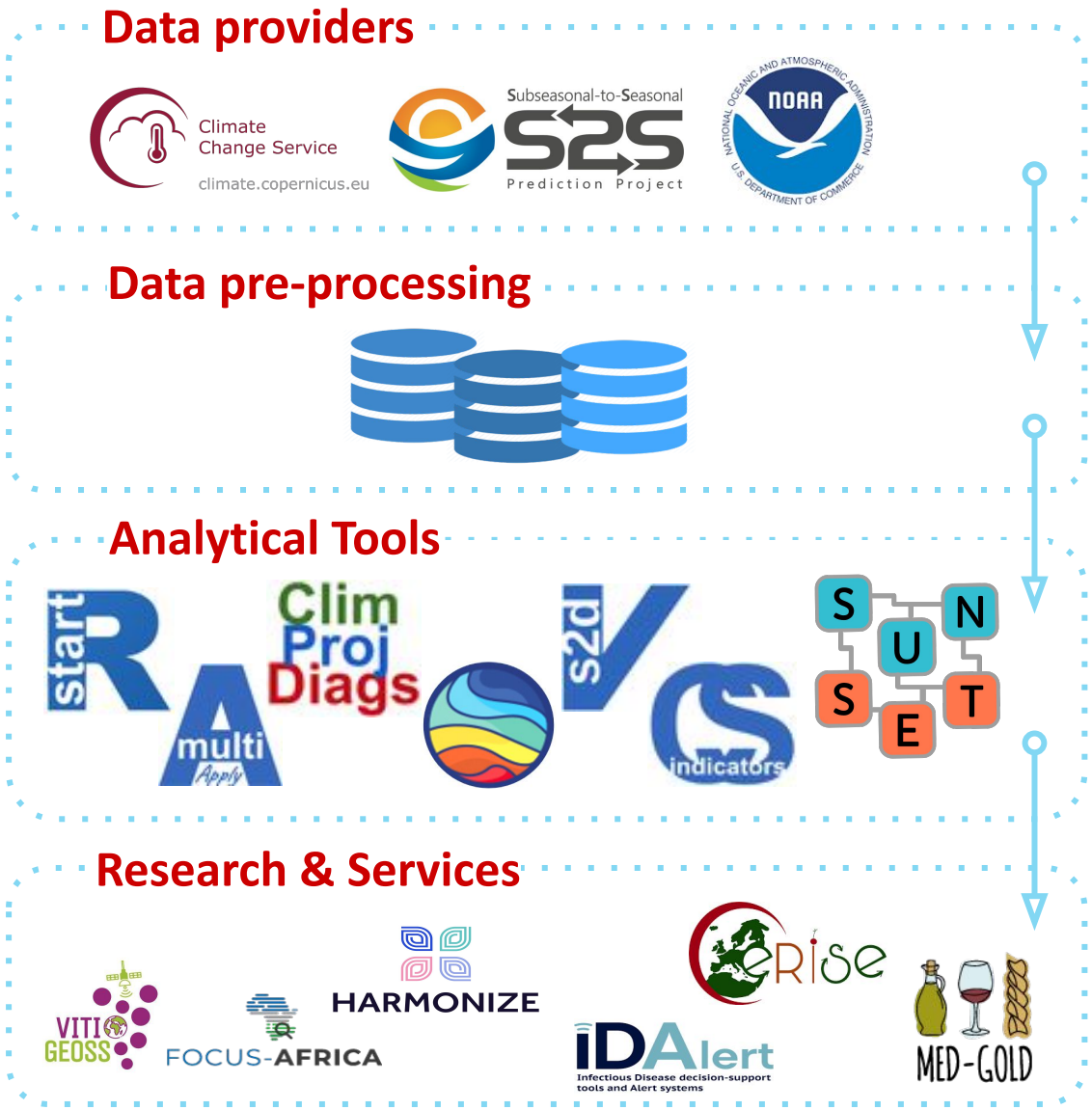
Important Features in Our Tools

LOADING

- Input data format: netCDF
- Different datasets under convention can be loaded
  - ➢ Data loading flexibility required

ANALYZING

- Accepted R object type by functions: Named multi-dimensional array mainly
- Different forecast horizons and frequency to be analyzed
  - ➢ Function flexibility required
- Multiple-core/node and parallel computation on HPCs
  - ➢ use package “multiApply”



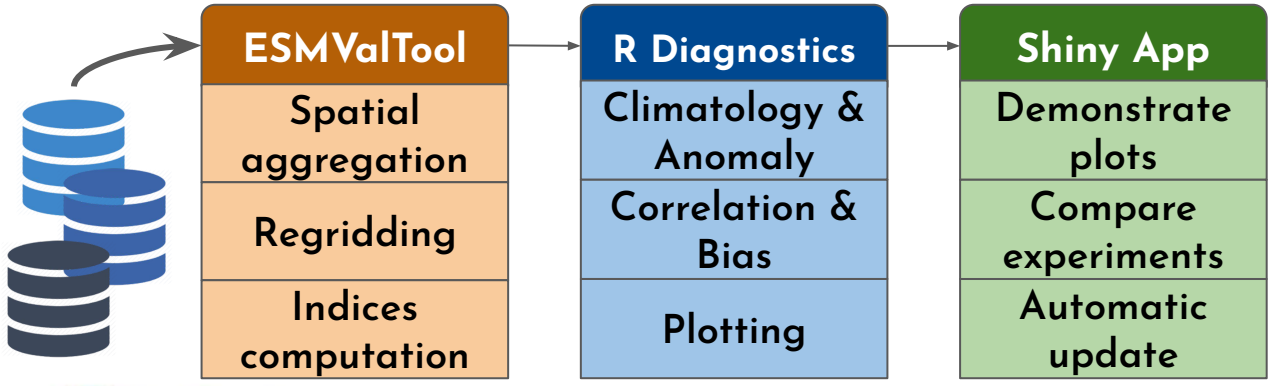
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Climate Forecast Analysis Tools: Integration

The packages are individual, but they can be integrated into a comprehensive data analysis workflow.

EXAMPLE 1 s2d Experiment Monitoring Tool

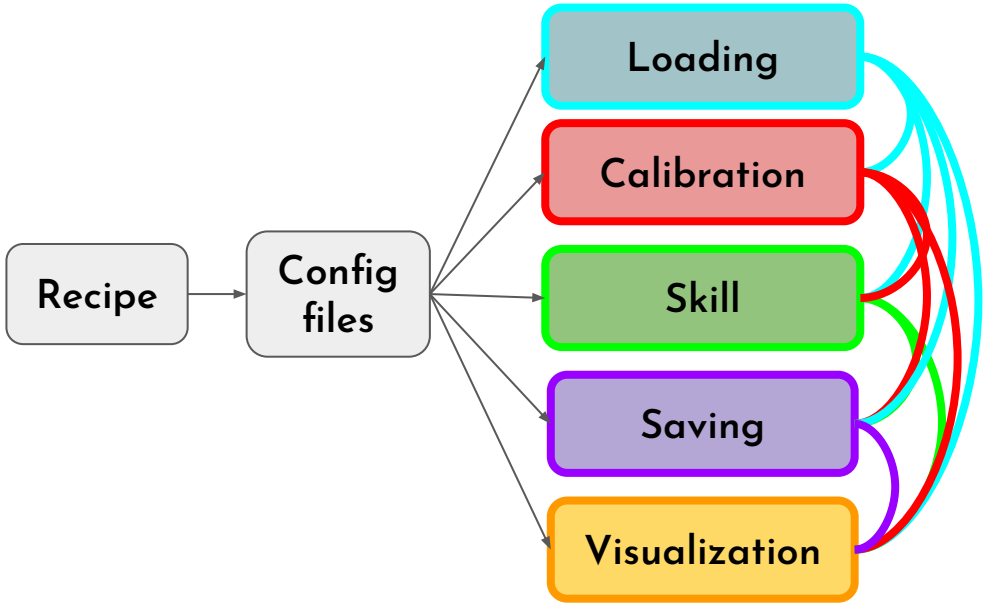
A workflow to monitor the seasonal/decadal predictions while they run. It consists of three parts:
(1) Raw data pre-processing by ESMValTool
(2) Diagnostics computed and plotted by our R packages
(3) Shiny app to show the plots



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXAMPLE 2 SUNSET: Climate Services Suite

This suite intakes a user-defined **recipe** and relies on **configuration** files. It has a **modularized** structure providing flexibility, so that parts of the workflow can be skipped or reordered.



Hands-On Preparation

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Prepare the environment on VM and MN4

1. If you use Windows, remember to open **Xming** first for showing the plots on the fly later.

   (download here)
2. log in VM: `ssh -XY patc{xx}@bsceshandson01.bsc.es`
3. ssh to mn4: `ssh -XY mn4` (passwordless)
4. Require resources:

   ```
   salloc -t 02:00:00 -n 1 -c 16 -J patc_test --x11 --qos=training
   ```
5. Load the required modules:

   ```
   module load gcc/7.2.0 pcre2 intel R/4.1.2 mkl impi CDO/1.8.2
netcdf/4.4.1.1
   ```
6. Open R: `R`

3. Hands-On I: Load data by startR

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

startR features

- ★ An R package for data retrieval and processing into multi-dimensional array
- ★ Accepted data format: netCDF
- ★ Apply multiApply paradigm, which provides flexibility in multi-dimensional data processing and multiple-core usage
- ★ Pre-processing: data regridding and reordering/reshaping/renaming dimensions along with data loading
- ★ Automatically chunk the data and dispatch on HPCs for parallel data-processing and combine chunks back together in the end
- ★ Well-preserved metadata from netCDF files
- ★ Use ecFlow or Autosubmit for job dispatch and monitoring on HPCs

Hands-on 1 overview

★ Goal: Use the package startR to load and pre-processing data

Tools needed:

- R, NCO
- startR package (v2.3.0)
- lubridate

Follow the Markdown file:

https://earth.bsc.es/gitlab/es/startR/-/blob/master/inst/doc/tutorial/PATC2023/hands on_1-data-loading.md

(Check the answer if needed)

NOTE: Don't close the session when you finish the exercise! We will use the data in the next hands-on.

4. Hands-On II: Data assessment with CSTools and s2dv

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

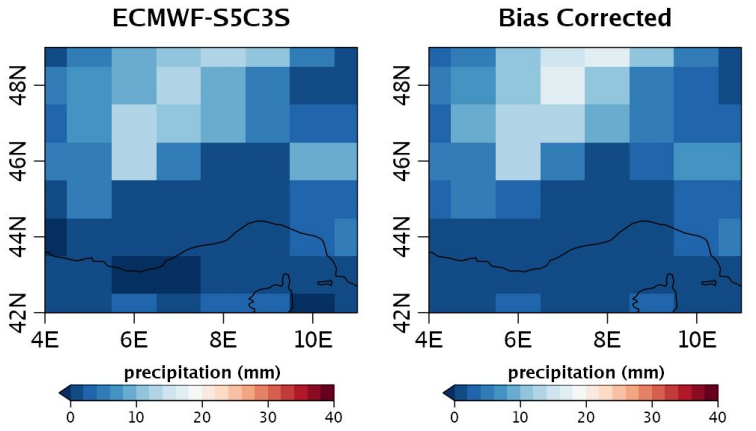
CSTools package

★ Climate forecast data can be postprocessed to obtain relevant information for the end-users.

Basic functions	Correction	Downscaling	Evaluation	Plotting	Classification
CST_Start CST_SaveExp CST_SplitDim CST_MergeDims …	CST_BEI_Weighting CST_BiasCorrection CST_QuantileMapping CST_Calibration …	CST_Analogs CST_RainFARM CST_AdamontAnalog CST_RFTemp …	CST_MultiMetric CST_MultivarRMSE	PlotCombinedMap PlotForecastPDF PlotPDFsOLE PlotWeeklyClim …	CST_MultiEOF CST_WeatherRegimes CST_RegimsAssign CST_EnsClustering …

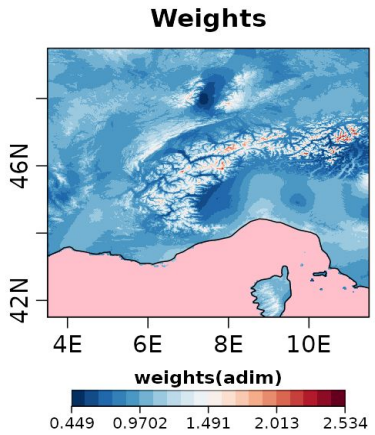
Load the data

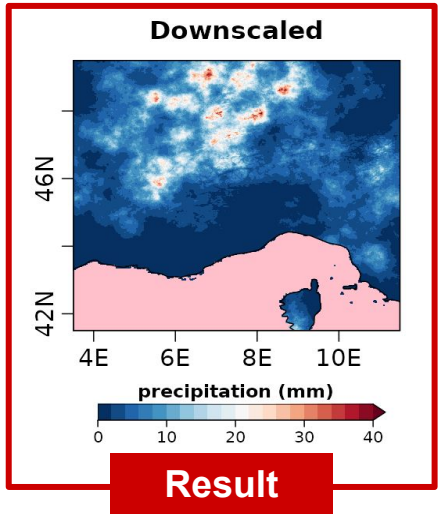
Step 1



Step 2

Step 3





Barcelona Supercomputing Center
Centro Nacional de Supercomputación

- GitLab repo: https://earth.bsc.es/gitlab/external/cstools
- On CRAN: https://cran.r-project.org/web/packages/CSTools/index.html

s2dv package

★ This package is designed to compare experimental and observational datasets. It covers from data retrieval, data post-processing, skill scores against observation, to visualization.

Verification

Compute deterministic and probabilistic scores and skill scores and correlation with reliability indicators such as p-values and confidence intervals.

- **Functions:** NAO, RMSSS, RMS, BrierScore, RPS, RPSS, CRPSS, MSE, RPS, RPSS, …

Statistics

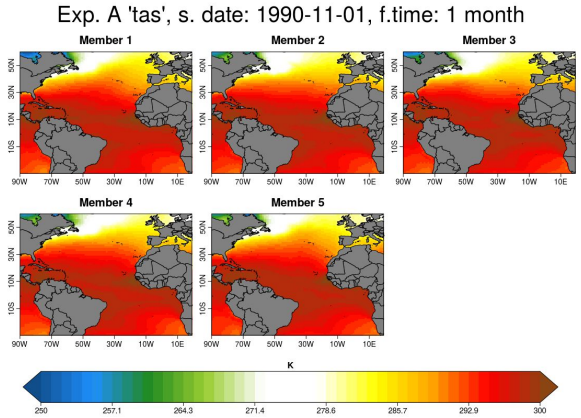
Functions that are commonly used in the forecast verification process to modify, describe or generate fields, a field being a series of modelled or observed measurements for a certain physic variable of interest.

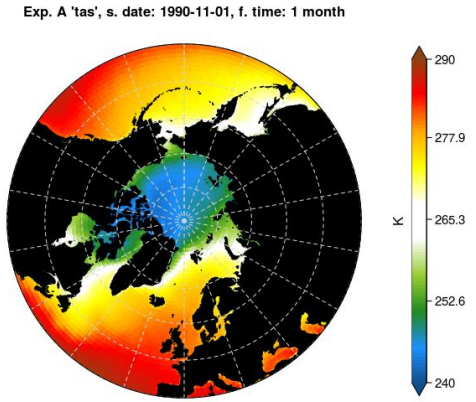
- **Functions:** Clim, Ano, Trend, Consist_Trend, Regression, Eno, Trend, ProbBins, …

Visualisation

Plotting functions are also provided to plot the results obtained from any of the modules above.

- **Functions:** PlotClim, PlotAno, PlotACC, PlotEquiMap, PlotStereoMap, AnimateMap, PlotLayout, …





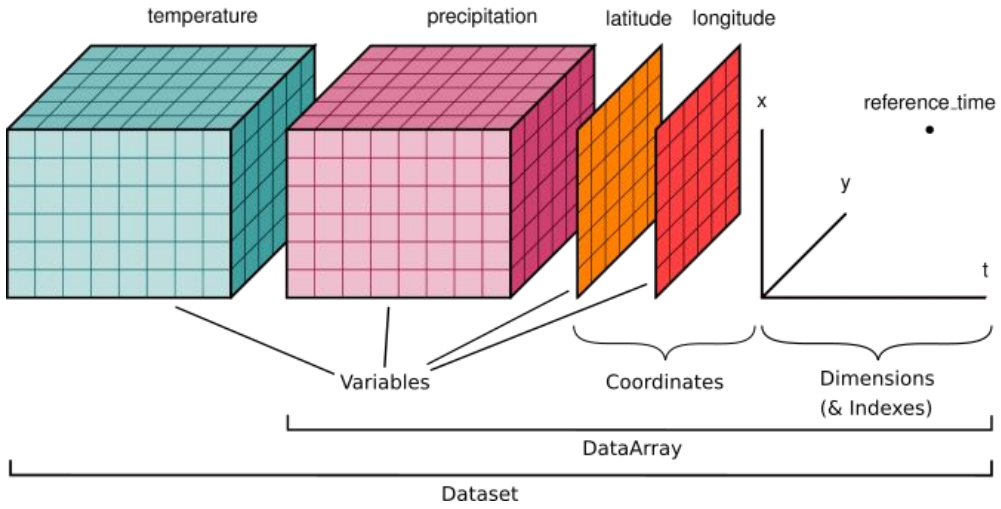
- GitLab repo: https://earth.bsc.es/gitlab/es/s2dv
- On CRAN: https://cran.r-project.org/web/packages/s2dv/

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

s2dv_cube object

★ **s2dv_cube** is an object to store ordered multidimensional array with named dimensions, specific coordinates and stored metadata.

★ Its “methods” are **CST** prefix functions from **CSTools** and **CSIndicators**




```
's2dv_cube'
Data         [294.97520446, 295.9965820, 296.99915313, 296.87461853, ... ]
Dimensions   (syear = 24, ensemble = 25, time = 2, lat = 61, lon = 61)
Coordinates
 * syear : 19931101, 19941101, 19951101, 19961101, 19971101, ...
   ensemble : 1, 2, 3, 4, 5, 6, ...
   time : 1, 2
 * lat : 20, 21, 22, 23, 24, 25, 26, 27, ...
 * lon : -20, -19, -18, -17, -16, -15, ...
Attributes
   Dates  : 1993-11-30 1994-11-30 1995-11-30 1996-11-30 1997-11-30 ...
   varName  : tas
   metadata :
      time
        units : hours since 1993-11-01 00:00:00
      tas
        units : K
        long name : 2 metre temperature
   Datasets  : dat1
   when  : 2023-10-21 17:40:59
   source_files  : /exp/meteofrance/system7c3s/monthly_mean/tas_f6h/ ...
   load_parameters  :
       ( dat1 )  : dat = dat1, var = tas, syear = 19931101 ...
```

Functions used

★ **CSTools::CST_Calibration** - Used to correct biases as well as dispersion errors of the model. This function contain different methods of member-by-member bias correction.

★ **CSTools::CST_Anomaly** - Compute deviations from the average conditions. This function computes the anomalies relative to a climatology computed along the selected dimension.

★ **s2dv::RPSS** - Is the skill score based on the Ranked Probability Score. It can be used to assess whether a forecast presents an improvement or worsening with respect to a reference.

- The RPSS ranges between minus infinite and 1. If the **RPSS > 0**, the forecast has higher skill than the reference, if **RPSS < 0** means that it has a lower skill.

Hands-on 2 overview

★ **Goal:** Use the package CSTools and s2dv to perform the hindcast quality assessment

Tools needed:

- R
- CSTools package v5.1.0
- s2dv package v2.0.0

Follow the Markdown file:

https://earth.bsc.es/gitlab/external/cstools/-/blob/doc-bsc_training_2023/inst/doc/tutorial/PATC2023/handson_2-data-assesment.md

(Check the answer if needed)

5. Hands-On III: SUNSET

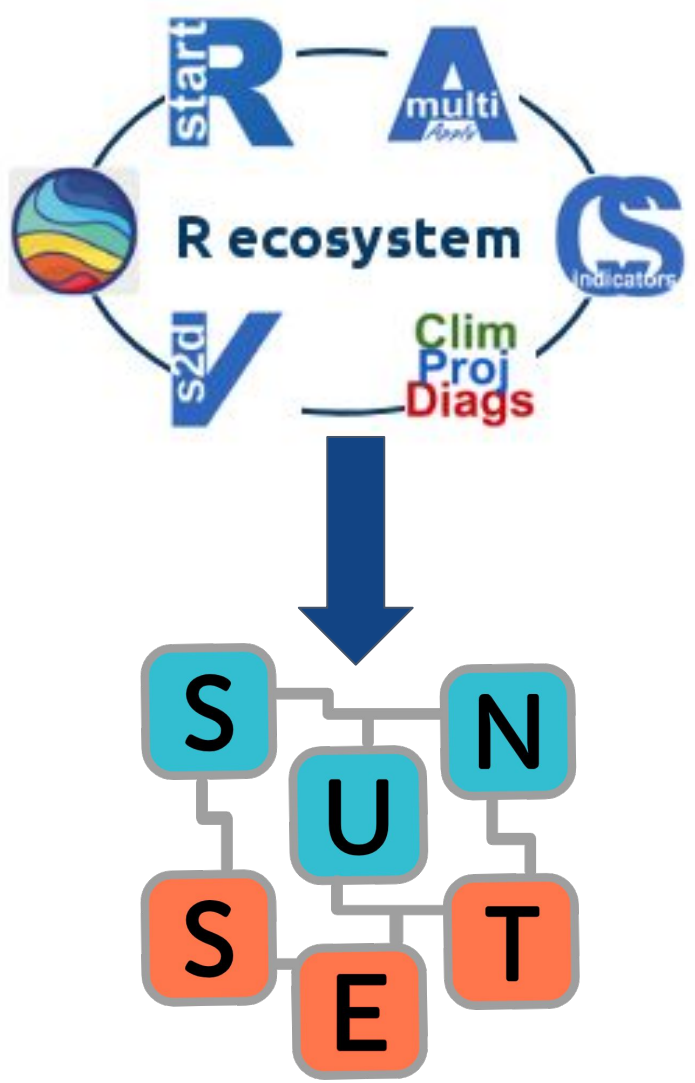
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Introduction to SUNSET

SUNSET, or " SUbseasoNal to decadal climate forecast post-processIng and asSEssmenT suite", is an **R-based tool for forecast post-processing and skill assessment workflows** that takes advantage of the R ecosystem in the Earth Sciences department.

The purpose of this tool is to foster collaboration among people working on forecast verification analysis within the department, to facilitate code reusability and a common framework.

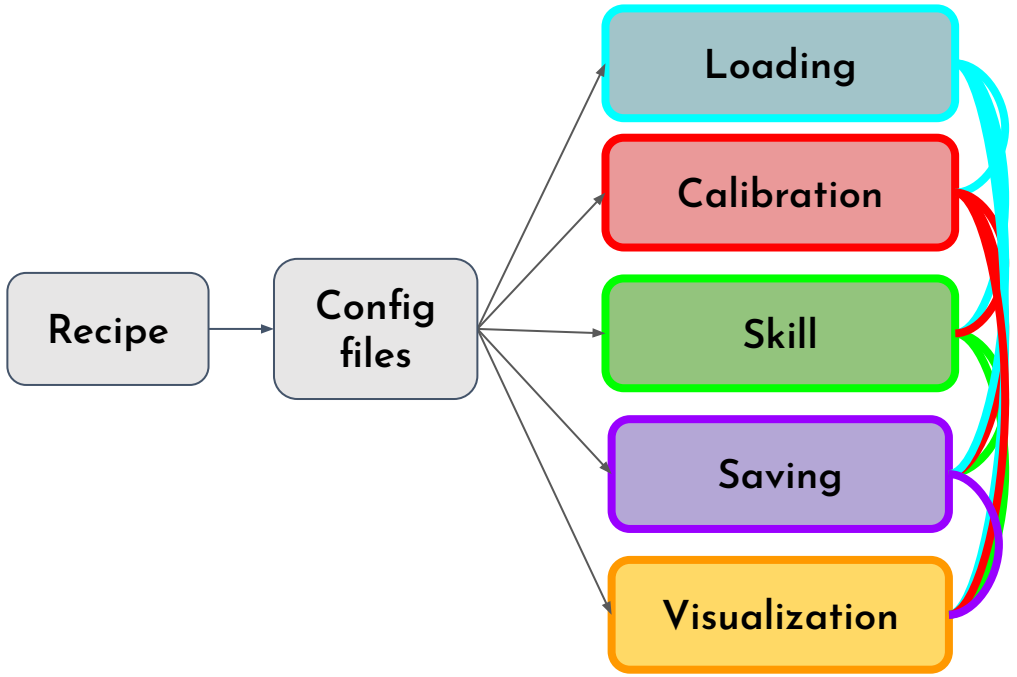
It **integrates** startR, CSTools, s2dv, multiApply… and many other packages in our R ecosystem.



Using SUNSET: Modular structure

SUNSET intakes a user-defined **recipe** with the information of the data to be loaded, and relies on configuration files and **modules**. Each module performs a different task: Loading, Downscaling, Calibration, Anomalies, computing Indices or Indicators, Skill Assessment, Visualization… The sequence of tasks in an analysis is called the **workflow** and is defined in an R script by the user.

The modularized structure provides flexibility, so that parts of the workflow can be skipped or executed at different points.



How do we run SUNSET?

SUNSET is designed for Climate Services applications and therefore can be run in two ways:

1. **Parallelizing workflows:** Creating a recipe with one or multiple datasets that require the same analysis. When launching SUNSET, the recipe is split into individual verifications that are run in parallel on a computer cluster, with the option to use Autosubmit as a workflow manager.
2. **Interactively on the terminal:** Create an individual recipe directly and run the functions on the terminal (for small tests and individual one-time analysis).

For the hands-on tutorial, we will focus on the second approach, which allows us to observe every step of the workflow.

Method 1. Splitting the recipe and parallelizing workflows

USER

RECIPE

**Specifies the assessment to carry on a model output. For example:
2 variables (e.g. temperature at surface and mean sea level pressure) and
2 different initialization dates (e.g. 1st of May and 1st of November)
for 1 region (e.g. Europe).**

AUTOSUBMIT

Atomic recipe 1	Atomic recipe 2	Atomic recipe 3	Atomic recipe 4
Assessment of 1st of **May** initialization day of the air **temperature** at surface	Assessment of 1st of **May** initialization day of the mean sea level **pressure**	Assessment of 1st of **November** initialization day of the air **temperature** at surface	Assessment of 1st of **November** initialization day of the mean sea level **pressure**
Load	Load	Load	Load
Calibration	Calibration	Calibration	Calibration
Anomalies	Anomalies	Anomalies	Anomalies
Skill scores	Skill scores	Skill scores	Skill scores
Probability computation	Probability computation	Probability computation	Probability computation
Spatial visualization	Spatial visualization	Spatial visualization	Spatial visualization

Save logs and outputs

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Method 2. Running an individual workflow interactively

USER

RECIPE

Specifies the assessment to carry on a model output for two variables (e.g. temperature at surface and mean sea level pressure) and 2 different initialization dates (e.g. 1st of May and 1st of November) on global scale.

AUTOSUBMIT

Atomic recipe 1

Assessment of 1st of **May** initialization day of the air **temperature** at surface

- Load
- Calibration
- Anomalies
- Skill scores
- Probability computation
- Spatial visualization

Atomic recipe 2

Assessment of 1st of **May** initialization day of the mean sea level **pressure**

- Load
- Calibration
- Anomalies
- Skill scores
- Probability computation
- Spatial visualization

Atomic recipe 3

Assessment of 1st of **November** initialization day of the air **temperature** at surface

- Load
- Calibration
- Anomalies
- Skill scores
- Probability computation
- Spatial visualization

Atomic recipe 4

Assessment of 1st of **November** initialization day of the mean sea level **pressure**

- Load
- Calibration
- Anomalies
- Skill scores
- Probability computation
- Spatial visualization

Save logs and outputs

Hands-on 3 overview

★ Goal: Create a SUNSET recipe and use the SUNSET functions to perform a climate forecast analysis

Tools needed:

- R
- R packages: yaml, startR, CSTools, s2dv, docopt, log4r, multiApply, easyNCDF, abind, lubridate, PCICt, RColorBrewer, stringr

Steps

1. Copy and edit the SUNSET recipe template
2. Load the data and run the workflow
3. Review the output files and plots

SUNSET Hands-on tutorial

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thank you

If you have any question, feel free to contact us on GitLab or through email.

- Victòria Agudetse (victoria.agudetse@bsc.es)
- Eva Rifà (eva.rifarovira@bsc.es)
- An-Chi Ho (an.ho@bsc.es)

[Acknowledgements]
All the current and previous Data and Diagnostic Team and Climate Services Team members' contribution on the presentation slides and tools