Funded by the European Union

ASPECT FACILITATING SEAMLESS CLIMATE ADAPTATION

Barcelona Supercomputing Center Centro Nacional de Supercomputación

# Statistically downscaled decadal climate variables for the agricultural, pensions sector and British Red Cross case studies

This archive contains statistically downscaled retrospective decadal climate predictions of anomalies of near-surface mean temperature (TAS) and precipitation (PR) for the next case studies: the agricultural sector in Catalonia, and the pensions sector and disaster response in the UK. High-resolution decadal climate variables relevant to each case study were generated: PR in the UK for the pensions sector, TAS in the UK for disaster response, and both variables in Catalonia for the agricultural case study. Essential details about the archived data are provided below.

## Scope and rationale

The data was produced in the framework of the European project ASPECT (Adaptation-oriented Seamless Predictions of European ClimaTe) and contributes towards ASPECT's agricultural case study, pensions sector case study and disaster response case study: more information on the project and its case studies can be found at the ASPECT website https://www.aspect-project.eu/. The agricultural case study is led and coordinated by the Barcelona Supercomputing Center, while the pensions sector and disaster response case studies are led and coordinated by the Met Office.

This work is part of a coordinated effort to demonstrate the value of statistically-downscaled climate information for a number of societal sectors. Please refer to ASPECT's dedicated infosheet for information on ASPECT's statistical downscaling approach.

## Context and definitions

*Case-studies:*

- The end user for the agricultural case study is a winemaking company based in the Spanish region of Catalonia.
- The end user for the pensions sector case study is the range of actors within the pensions sector from the UK.
- The end user for the disaster response case study is the British Red Cross.

*Domain*: the geographical domain is Western Europe. This domain covers latitudes from 34.5ºN to 58.8ºN and longitudes from 12ºW to 10ºE.
*Spatial resolution:* the target horizontal spatial resolution for the downscaled climate information is of 0.1º (approximately 10 km).
*Predictand:* the predictand is the own variable for each one.
*Time horizon:* Retrospective predictions are issued for the standard 1960-2013 verification period with a forecast period of 1 to 5 years.
*Type of predictions:* The predictions are for anomalies of the variables TAS and PR.

**Input data**

The archived data is computed based on reanalysis and decadal prediction.

***Reanalysis data:*** obtained from the ERA5-Land reanalysis (Muñoz-Sabater et al., 2021) dataset, which was used as the observation-based reference for statistical downscaling. The product has a spatial resolution of approximately 10 km. Depending on the downscaling method, monthly values (transformed to yearly means) or daily values of TAS and PR have been used. In addition, daily mean sea-level pressure (PSL) and 500 hPa geopotential height (zg500), and monthly teleconnection indices were used as predictors in some methods. The teleconnection indices used are: North Atlantic Oscillation (NAO; Wanner et al., 2001), Atlantic Multidecadal Variability (AMV; Trenberth & Shea, 2006), South Pacific Ocean Dipole (SPOD; Huang et al., 2006), and Trans-Polar Index (TPI; Pittock, 1980).

***Decadal predictions data:*** obtained from 13 decadal forecast systems contributing to the first component of the Decadal Climate Prediction Project (DCPP-A; Boer et al., 2016), as part of phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016). Details on the DCPP members, spatial resolutions, and references of the 13 decadal forecast systems can be found in Table 1 of Delgado-Torres et al. (2022). The spatial resolution varies across models, ranging from 0.7°×0.7° in EC-Earth to 2.8°×2.8° in CanESM5. Depending on the method, either monthly values (converted to yearly means) or daily values of TAS and PR were used. In addition, monthly teleconnection indices were employed as predictors in some methods.

**Methodology**

As stated above, the primary objective of this study is to produce high-resolution datasets of TAS and PR anomalies and to build a multi-model ensemble with the results obtained from each dataset. The anomalies are calculated relative to the 1980–2010 reference period. The different downscaling methods are applied independently to each dataset, and the multi-model ensemble is constructed afterwards. Multiple downscaling methods have been tested in order to identify the one that performs best for each variable (Moreno-Montes et al. in prep).

Depending on the nature of the chosen predictors, two different statistical downscaling approaches were applied: perfect prognosis (PP) and model output statistics (MOS). PP approaches assume a 'perfect' relationship between the large-scale observations or reanalysis data (predictors) and local-scale observations (predictands). MOS methods, however, acknowledge the imperfections of the model and aim to correct systematic biases while increasing resolution. They establish a statistical relationship between predictors and predictands using simulated predictors and high-resolution observations or reanalysis data as predictands. A number of methods from these two approaches were used:

- Interpolation: Five interpolation methods are compared: two conservative variants (first-order and second-order), bilinear, bicubic, and nearest neighbour.
- Interpolation plus calibration: This MOS approach involves combining the initial interpolations with different calibration techniques. The calibrations applied are: bias and variance adjustment (EVMOS-cal; Vannitsem, 2009), Mean Square Error minimisation (MSE-min; Doblas-Reyes et al., 2005), Continuous Ranked Probability Score minimisation (CRPS-min; Gneiting et al., 2005), and ratio of predictable

components-based calibration (RPC-based; Eade et al., 2014). Besides, different versions of the quantile mapping method (QM; Panofsky and Brier, 1968) are applied to precipitation.

- Linear or multilinear regressions: These methods establish statistical relationships between large-scale climate predictors and high-resolution observational data. The regressions are built using model outputs, observational data, or teleconnection indices (from both observations and models). The teleconnection indices used are: NAO, AMV, SPOD and TPI. Depending on the type of predictor used, the regressions follow either a PP or MOS approach.
- Analogs: This method, which uses daily data, searches for past fields similar to the one being predicted within a reference large-scale historical dataset at a coarser resolution. The most similar analog day is selected based on Spearman correlation, and the corresponding high-resolution reference field from that day is retrieved as the prediction. The search of the most similar analog day in the large-scale dataset can be done using the same variable as predictor (MOS) or another variable such as PSL and zg500 (PP).

The downscaling procedures were carried out using a leave-nine-year-out cross-validation to avoid potential overfitting. All the methods are available in the R package CSDownscale (Duzenli et al., 2024).

**Forecast evaluation**

To evaluate the performance of each method, the Pearson's Anomaly Correlation Coefficient (ACC; Wilks, 2010) was computed between the downscaled results and the ERA5-Land reanalysis data at each grid point. The ACC ranges from −1, indicating a perfect inverse correlation, to 1, indicating a perfect correlation. To assess the statistical significance of the results, a one-sided t-test was performed to determine whether the ACC values differ significantly from zero at the 95% confidence level. The effective number of degrees of freedom has been used to account for the time series autocorrelation during the significance testing following von Storch and Zwiers (1999).

Based on this metric, some methods were able to preserve or even improve the original skill of the multi-model predictions for each variable. Considering all the region, for TAS, bicubic interpolation achieved the highest ACC values, whereas for PR, the analogs method using PSL as a predictor outperformed all other methods. As an example, Figure 1 illustrates the ACC results for the multi-model of PR predictions using analogs-PSL method. The percentage of significant points is indicated in brackets.

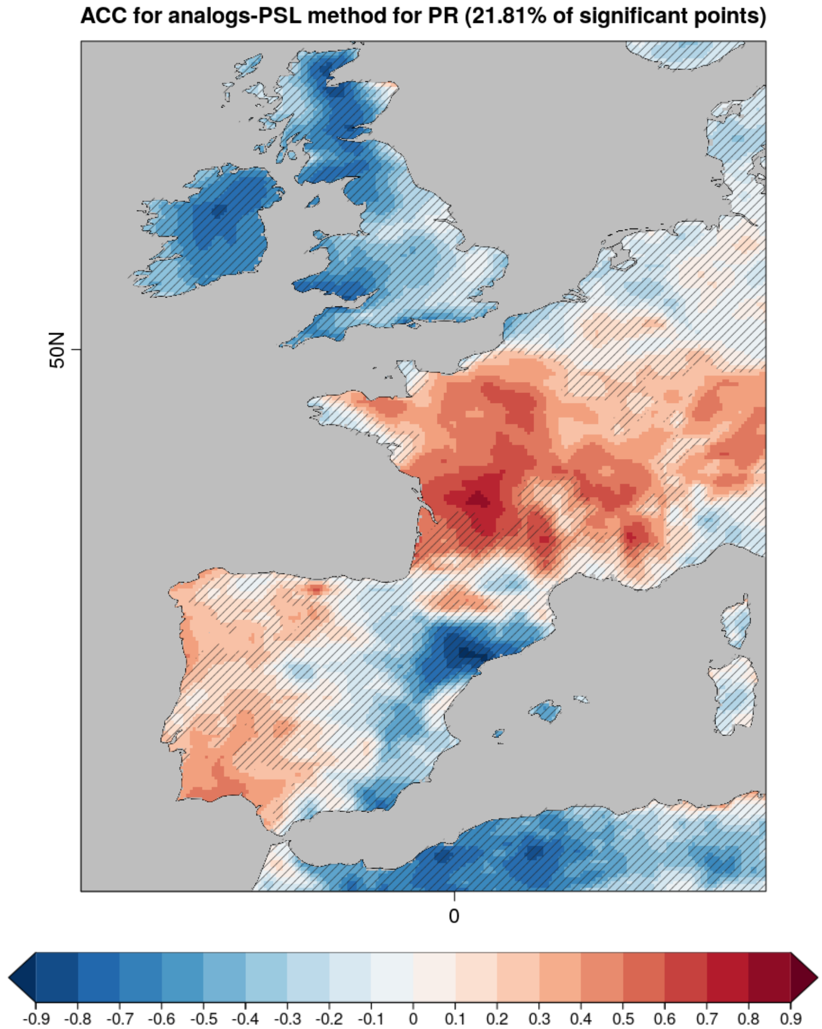


**Figure 1:** *ACC obtained with the multi-model predictions of PR using analogs-PSL method. Hatched regions indicate non-significant ACC value*

**Output data format**

The archive includes data for both variables across multiple start dates from the 13 decadal forecast systems. For each variable, results are shown for the method that yielded the highest ACC values: bicubic interpolation for TAS and the analogs method using PSL as a predictor for PR. Due to the different nature of the methods, the naming conventions also vary accordingly.

For the interpolation method, the mean of the forecast years 1-5 is computed before applying the downscaling technique. As a result, one output file is generated per start date, containing the downscaling predictions averaged over the next 5 years. Accordingly, the naming convention for TAS files is as follows:

i) dcpp_tas_model_member_ssyear_forecastperiod.nc where syear indicates each one of the start years (1960-2013) and the forecast period ranges from 1961-1996 to 2014-2018.

The analogs method, however, calculates the downscaling results separately for each forecast year, resulting in a different file for each one of the 5 forecast years. The naming convention for PR files is:

i) dcpp_pr_model_member_ssyear_forecastyear.nc

where syear ranges again from 1960 to 2013 but the forecast year corresponds to one year.

Figure 1 illustrates the data structure of one archived file. Each file contains a 221 by 301 spatial grid in latitude and longitude.

```
smoreno@bscearth378:/esarchive/scratch/smoreno/aspect_outputs/outputs_downscaling_nc> ncdump -h dcpp_tas_CanESM5_r8i1
p2f1_s1966_1967-1972.nc
netcdf dcpp_tas_CanESM5_r8i1p2f1_s1966_1967-1972 {
dimensions:
        longitude = 221 ;
        latitude = 301 ;
variables:
        double tas(latitude, longitude) ;
                tas:variable = "tas" ;
                tas:model = "CanESM5" ;
                tas:member_name = "r8i1p2f1" ;
                tas:reference_period = "1981-2010" ;
                tas:grid = "r3600x1801" ;
                tas:original_grid = "2.8ºx2.8º" ;
                tas:sdates = 1966 ;
                tas:target_period = "1967-1972" ;
                tas:forecast_period = "years1-5" ;
                tas:standard_name = "air_temperature_anomaly" ;
                tas:long_name = "Near-Surface Air Temperature anomaly" ;
                tas:comment = "near-surface (usually, 2 meter) air temperature anomaly" ;
                tas:units = "K" ;
        double latitude(latitude) ;
                latitude:units = "degrees_north" ;
        double longitude(longitude) ;
                longitude:units = "degrees_east" ;

// global attributes:
                :Downscaling\ method = "Bicubic Interpolation" ;
                :Project = "ASPECT" ;
                :Institution = "Barcelona Supercomputing Center (BSC)" ;
                :Activity = "tas calculation from ERA5land dataset" ;
}
```

**Figure 2:** *Screenshot of the "ncdump -h file.nc" output showing an example of the data structure of downscaled predictions*

**Acknowledgements**

This research was supported by the ASPECT project, which has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081460. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

**References**

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F. J., Danabasoglu, G., Kirtman, B., ... & Wyser, K. (2016). The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. Geoscientific Model Development, 9(10), 3751–3777. DOI: 10.5194/gmd-9-3751-2016

Delgado-Torres, C., Gualdi, S., Bilbao, R., Cortesi, N., Caron, L.-P., Bellprat, O., ... & Doblas-Reyes, F. J. (2022). Multi-model forecast quality assessment of CMIP6 decadal predictions. Journal of Climate, 35(13), 4363–4382. DOI: 10.1175/JCLI-D-21-0811.1

Doblas-Reyes, F. J., Andreu-Burillo, I., Chikamoto, Y., García-Serrano, J., Guemas, V., Kimoto, M., ... & Mochizuki, T. (2013). Initialized near-term regional climate change prediction. Nature Communications, 4, 1715. DOI: 10.1038/ncomms2704

Duzenli, E., Gamon, J. R., Llabres, A., Torralba, V., Palma Garcia, L., Moreno Montes, S., ... & Marcos, R. (2024, April). Climate Services Downscale (CSDownscale): A statistical downscaling tool for (sub) seasonal to decadal climate predictions. In EGU General Assembly Conference Abstracts (p. 19420). DOI: 10.5194/egusphere-egu24-19420

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., & Robinson, N. (2014). Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?. Geophysical research letters, 41(15), 5620-5628. DOI: 10.1002/2014GL061146

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958. DOI: 10.5194/gmd-9-1937-2016

Gneiting, T., Raftery, A. E., Westveld, A. H., & Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Monthly Weather Review, 133(5), 1098–1118. DOI: 10.1175/MWR2904.1

Huang, B., & Shukla, J. (2006). Interannual SST variability in the southern subtropical and extra-tropical ocean. Technical Report 223, Center for Ocean-Land-Atmosphere Studies.

Moreno-Montes, S., Delgado-Torres, C., Duzenli, E., Pérez-Zanón, N., & Soret, A. . Comparative analysis of statistical downscaling methods in a multimodel framework for decadal climate predictions over Western Europe (in preparation)

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., ... & Zsoter, E. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9), 4349–4383. DOI: 10.5194/essd-13-4349-2021

Panofsky, H. A., & Brier, G. W. (1968). Some applications of statistics to meteorology. *(No Title)*.

Pittock, A. B. (1980). Patterns of climatic variations in Argentina and Chile – I. Precipitation, 1931–60. Monthly Weather Review, 108(9), 1347–1361. DOI: 10.1175/1520-0493(1980)108<1347:POCVIA>2.0.CO;2

von Storch, H., & Zwiers, F. W. (1999). Statistical analysis in climate research. Cambridge University Press. DOI: 10.1017/CBO9780511612336

Trenberth, K. E., & Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. Geophysical Research Letters, 33(12), L12704. DOI: 10.1029/2006GL026894

Vannitsem, S. (2009). A unified linear model output statistics scheme for both deterministic and ensemble forecasts. Quarterly Journal of the Royal Meteorological Society, 135(643), 1801–1815. DOI: 10.1002/qj.498

Wanner, H., Brönnimann, S., Casty, C., Gyalistras, D., Luterbacher, J., Schmutz, C., Stephenson, D. B., & Xoplaki, E. (2001). North Atlantic Oscillation – Concepts and studies. Surveys in Geophysics, 22, 321–381. DOI: 10.1023/A:1014217317898

Wilks, D. S. (2011). Statistical methods in the atmospheric sciences (Vol. 100). Academic press.