Funded by the European Union

Statistically downscaled seasonal climate indicators for the agricultural case study

This archive contains statistically downscaled retrospective seasonal climate indicators of agricultural case study in Catalonia. High-resolution seasonal climate indicators relevant to the case study were generated with the aim of evaluating water availability and spring frost events. Essential details on 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: 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.
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-study: The end user for the agricultural case study is Codorníu, a winemaking company based in Catalonia, Spain, known for producing high-quality wines and cava for over 470 years.

Domain: The geographical domain of the study is Catalonia, bounded by latitudes 39°N–44°N and longitudes 1°W–4°E.

Spatial resolution: The horizontal spatial resolution for the downscaled climate information is 5 km.

Predictand: The predictands in this study are minimum temperature, maximum temperature, and precipitation. These variables were used to compute key climate indicators. The Standardized Precipitation Evapotranspiration Index (SPEI) was calculated over accumulation periods of 1, 3, and 6 months (i.e., SPEI1, SPEI3, SPEI6) using minimum and maximum temperature along with precipitation, to provide insights into water availability. Frost Days (FD) were derived from minimum temperature data as an indicator of potential spring frost risk.

Time horizon: This dataset includes retrospective data covering the period from 1993 to 2016. SPEI values are available for each month from April to September, while FD values are provided for the months March, April, and May.

Type of predictions: SPEI provides standardized values of water deficit (Vicente-Serrano et al., 2010). Negative values indicate drier-than-normal conditions for the region, while positive values correspond to wetter-than-normal conditions. As values approach zero, conditions are considered closer to the climatological normal. Traditionally, values below −1 or above +1 are interpreted as indicating considerable drought or wetness events, respectively. Frost Days (FD) represent the number of days within a given month during which the daily minimum temperature falls below 0°C.

Input data

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

Reanalysis data: obtained from the CERRA (Schimanke et al., 2021) dataset, which was used as the reference for statistical downscaling. The product has a spatial resolution of approximately 5 km. Various downscaling experiments were performed using CERRA variables both as predictands (i.e., minimum temperature, maximum temperature, and precipitation) and as predictors. The predictor set included large-scale atmospheric variables such as horizontal wind components at 300, 500, and 850 hPa, temperature, relative humidity, and geopotential height at 850 hPa, as well as mean sea level pressure.

Seasonal predictions data: obtained from the seasonal forecasts produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), specifically the SEAS5v1 system (Johnson et al., 2019). These forecasts are based on ECMWF's Integrated Forecasting System (IFS) and consist of 25 ensemble members for the hindcast dataset. The spatial resolution of the data is approximately 100 km. The same set of variables described in the reanalysis section was utilized from this dataset.

Methodology

As stated above, the primary objective of this study is to produce high-resolution SPEI and FD datasets (Duzenli et al., in prep). To generate these indicators in a skillful way, a two-step sensitivity analysis was conducted using various statistical downscaling methods. The entire methodological workflow is illustrated in Figure 1. The workflow was repeated for each target month and for four different forecast times (1, 2, 3, and 4). For example, the downscaling of a variable evaluated for May was carried out using May values obtained from model runs initialized in February, March, April, and May separately. Downscaling was performed using daily data, which were subsequently aggregated to monthly values prior to the SPEI calculation. For FD, daily minimum temperature values were used directly to count the number of frost days within each month, resulting in a single FD value computed per month. The downscaling procedures were carried out via the R packages CSDownscale (Duzenli et al., 2024) and downscaleR (Bedia et al., 2020).

In the first step, minimum temperature, maximum temperature, and precipitation, which are used as input variables for indicator computation, were individually downscaled using nine different methods. These methods represent various combinations of analogs, bias correction, linear regression, and logistic regression techniques. All approaches in this step followed the Model Output Statistics (MOS) framework, in which the anomaly of the target variable from the GCM served as the predictor, and the corresponding anomaly from the observational dataset was used as the predictand. Results from this phase showed that the analogs method generally outperformed the other approaches. Consequently, the analogs method was selected as the best-performing option in the first step.

In the second step, rather than relying solely on the target variable as the predictor, additional large-scale variables, known to influence the variability of local conditions, were also incorporated. The aim was to investigate whether these extended approaches provide added value compared to the basic Analogs-MOS method. The specific predictors used for each target variable are listed in Table 1. For temperature-related variables, pressure-based fields were selected as predictors, while for precipitation, variables associated with adiabatic cooling and moisture advection were used.

When using multiple predictors, each variable was first standardized individually. PCA was then performed on the combined dataset of standardized predictors, and the components explaining 95% of the total variance were retained as predictors. These large-scale predictors were used to search for the most similar historical time step. After identifying the closest match based on these large-scale fields, the corresponding local-scale value from the observational dataset was extracted and assigned as the downscaled result for the target time step.

Additionally, a downscaling approach based on weather regimes (WRs) was applied (Olmo and Bettolli, 2022; Duzenli et al., 2025). WRs were calculated separately for each month using mean sea level pressure fields over the North Atlantic. Four distinct regimes were defined per month, and each time step of both the model outputs and observational data was classified accordingly prior to the analog search. The search was then restricted to observational time steps belonging to the same WR category as the corresponding model time step. This approach introduces a pre-filtering step based on large-scale circulation patterns, ensuring that analogs are selected only from periods influenced by similar atmospheric conditions.

According to the results, the optimal downscaling method differed for each target variable. In general, MOS-based approaches performed better for temperature-related variables, whereas Perfect Prognosis (PP)-based methods, which use observed large-scale variables as predictors, provided more skillful outputs compared to MOS for the downscaling of precipitation.

For FD calculation, minimum temperature values that were downscaled using the best-performing method for that variable (Analogs-MOS) were used. For SPEI, all three variables were downscaled using a common predictor set, specifically the one used for precipitation in Table 1. This ensured that the same days were selected as analogs for the corresponding time steps of each variable, thereby maintaining physical consistency. For example, using a common predictor set can prevent situations where a downscaled maximum temperature might be lower than the corresponding minimum temperature on a given day, or where unrealistic relationships might arise between temperature and precipitation values. The reason for using the predictor combination that yielded the best performance for precipitation is that precipitation was identified as the most influential variable in determining SPEI skill.

In cases where the SPEI scale was longer than the forecast time, observational data were incorporated. For example, when calculating SPEI3 for May based on model outputs initialized in April, downscaled climate predictions were utilized for April and May, while the values for March were taken from the reference dataset, namely CERRA. In addition, potential evapotranspiration (PET) used in the SPEI calculation was estimated using the Hargreaves equation (Hargreaves and Allen, 2003). The water deficit (precipitation minus PET) is standardized through a three-parameter shifted log-logistic probability distribution. The parameters of this distribution are estimated using the unbiased probability weighted moments approach, as described by Beguería et al. (2014) and Stagge et al. (2015). All the processes explained in the methodology section were implemented using a leave-one-year-out cross validation approach to avoid potential overfitting.

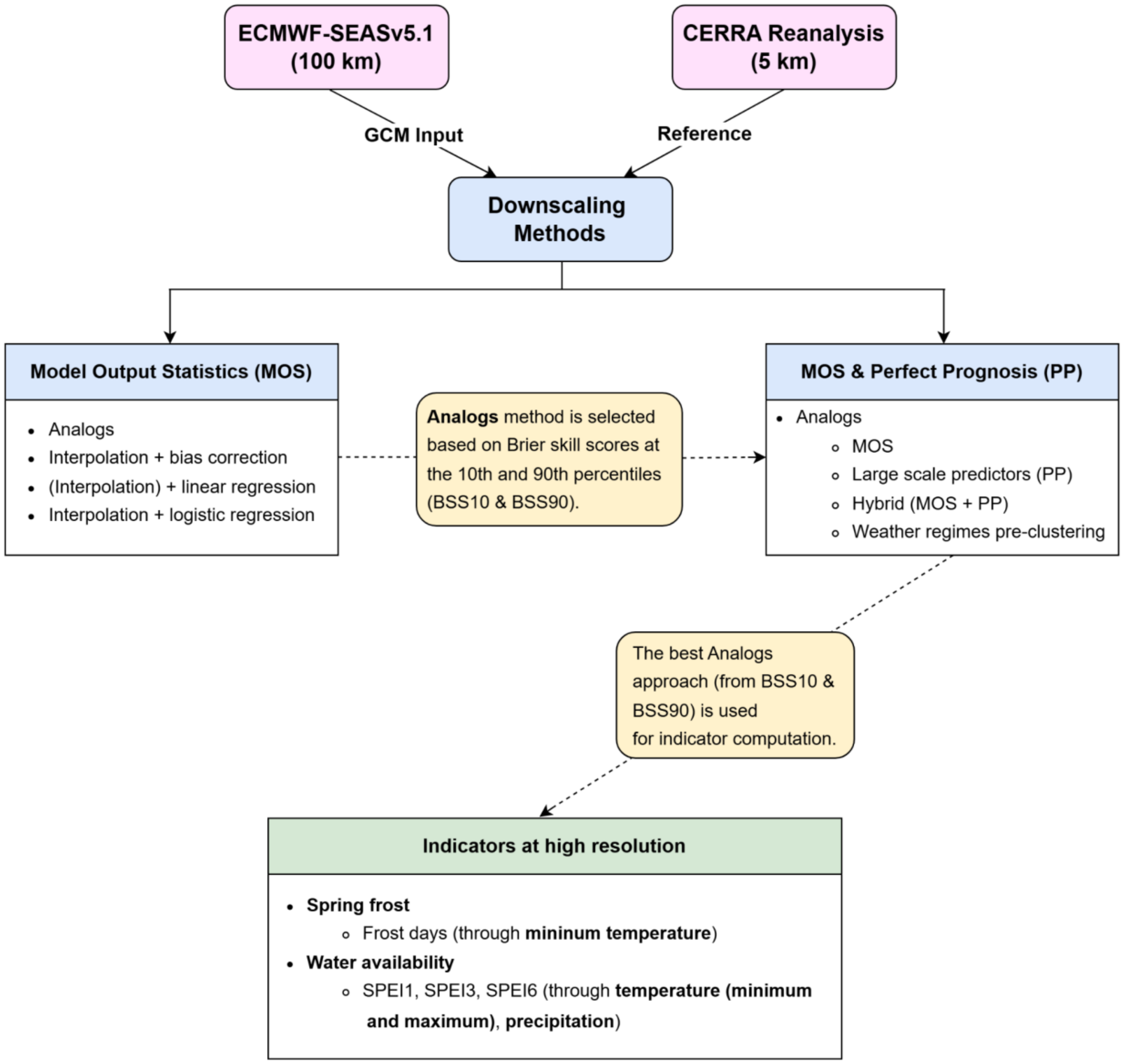

Figure 1: *Summary of the downscaling methodology*

Table 1. Large-scale predictors with respect to the predictands. **ta850**: temperature at 850 hpa pressure level; **g500**: geopotential height at 500 hpa pressure level; **g850**: geopotential height at 850 hpa pressure level; **ua300**: zonal wind at 300 hpa; **ua500**: zonal wind at 500 hpa; **ua850**: zonal wind at 850 hpa; **hur500**: relative humidity at 500 hpa pressure level; **hur850**: relative humidity at 850 hpa pressure level; **slp**: mean sea level pressure

Predictand	Large scale predictors
Temperature (min, mean, max, dew-point)	ta850, g850, slp
Precipitation	ua300, ua500, ua850, g500, g850, hur500, hur850, slp

Forecast evaluation

In the sensitivity analysis part, the Brier Skill Score at the 10th and 90th percentiles (BSS10 and BSS90) were utilized to evaluate the models' ability to predict lower and upper extreme events, respectively. The primary goal of this step is to identify the methods most effective at capturing extreme values in the atmospheric variables, as these variables are directly used to compute impact-relevant indicators. Once the indicators were calculated, their overall predictive performance was further evaluated by incorporating the Ranked Probability Skill Score (RPSS), which measures the accuracy of probabilistic forecasts across ordered categories. In this study, RPSS was applied to tercile-based indicator forecasts to complement the evaluation based on BSS10 and BSS90, thereby enabling a more comprehensive assessment of the high-resolution indicators under both general and extreme conditions.

Output data format

The archive includes data for four indicators (i.e., SPEI1, SPEI3, SPEI6, and FD) across multiple start months and forecast times. These data are available for both the downscaled seasonal climate predictions from ECMWF-SEAS5.1 and the CERRA reference dataset. The files follow the naming conventions outlined below:

i-) dwn_ecmwfseas5-1_ind_smX_ftY.nc: Represents the downscaled ECMWF-SEAS5.1 product.
ii-) obs_cerra_ind_mX.nc: Represents the reference indicator computed from the CERRA dataset.

In both filenames, the product name (ecmwfseas5-1 or cerra) is included. The "ind" refers to one of the four indicators: SPEI1, SPEI3, SPEI6, or FD. For the downscaled product, data are provided for four different forecast times. Accordingly, smX denotes the start month X, while ftY denotes the forecast time Y.

For the reference data, start month and forecast time are not applicable; therefore, a simplified format mX is used to denote the target month X. Thus, the naming convention for the SPEI3 data for May, derived from a forecast initialized in February, would be as follows:

i-) dwn_ecmwfseas5-1_spei3_sm2_ft4.nc (downscaled forecast),
ii-) obs_cerra_spei3_m5.nc (reference).

Figure 2 illustrates the data structure of the archived files. Each file in the archive contains 23 or 24 time steps, corresponding to the periods 1993 to 2016 or 1994 to 2016, and a 100 by 100 spatial grid in latitude and longitude. Additionally, the downscaled ECMWF-SEAS5.1 files include an ensemble dimension with 25 members, which is not present in the reference-based product.

```
[eduzenli@bsceshub06 nc_outputs]$ ncdump -h dwn_ecmwfseas5-1_SPEI3_sm4_ft2.nc
netcdf dwn_ecmwfseas5-1_SPEI3_sm4_ft2 {
dimensions:
        time = 24 ;
        lat = 100 ;
        lon = 100 ;
        ens_mem = 25 ;

variables:
        long time(time) ;
                time:units = "Monthly SPEI3 data for May during the period 1993 - 2016" ;
                time:long_name = "time" ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:long_name = "latitude" ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:long_name = "longitude" ;
        long ens_mem(ens_mem) ;
                ens_mem:long_name = "ensemble member" ;
        float spei3(ens_mem, lon, lat, time) ;
                spei3:units = "unitless" ;
                spei3:_FillValue = 1.e+20f ;
                spei3:long_name = "Standardized Precipitation-Evapotranspiration Index over a 3-month
accumulation period" ;
                spei3:Projection = "Regular latitude-longitude grid" ;

// global attributes:
                :Project = "ASPECT" ;
                :Institution = "Barcelona Supercomputing Center (BSC)" ;
                :Activity = "Statistical Downscaling of ECMWF-SEAS5.1 seasonal predictions" ;
                :Start month of the downscaled predictions = "April. Since the SPEI scale (i.e., 3 mon
ths) exceeds the forecast range (i.e., 2 months), CERRA data is used to supplement the months outside
the forecast period." ;
                :Potential evapotranspiration (PET) = "Estimated using Hargreaves equation (Hargreaves
 and Allen, 2003)" ;
                :Standardization of the water deficit (Precipitation - PET) = "Done using the three-pa
rameter shifted log-logistic probability distribution, in which the parameters are computed using the
unbiased probability weighted moments method (Beguería et al. 2014; Stagge et al. 2015)" ;
                :Reference 1 = "Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014).
Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspir
ation models, tools, datasets and drought monitoring. International journal of climatology, 34(10), 30
01-3023. https://doi.org/10.1002/joc.3887" ;
                :Reference 2 = "Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Ha
rgreaves evapotranspiration equation. Journal of irrigation and drainage engineering, 129(1), 53-63. h
ttps://doi.org/10.1061/(ASCE)0733-9437(2003)129:1(53)" ;
                :Reference 3 = "Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., & S
tahl, K. (2015). Candidate distributions for climatological drought indices (SPI and SPEI). Internatio
nal Journal of Climatology, 35(13), 4027-4040. https://doi.org/10.1002/joc.4267" ;
}
```

Figure 2: *Data structure of downscaled product*

Figure 3 shows the spatial distribution of RPSS values for SPEI3 and precipitation in May, based on forecasts from models initialized in March and April. For the model initialized in April, SPEI3 predictions appear considerably more skillful than those for precipitation. This is expected, as the precipitation, minimum temperature, and maximum temperature values used to calculate SPEI3 for May come from the model for April and May, but from observations for March. However, even when all input variables for SPEI3 are sourced from the March-initialized model, SPEI3 still outperforms precipitation in terms of skill. This result highlights a situation in which the standardized combination of precipitation with minimum and maximum temperatures provides greater skill than precipitation alone when evaluated through RPSS.

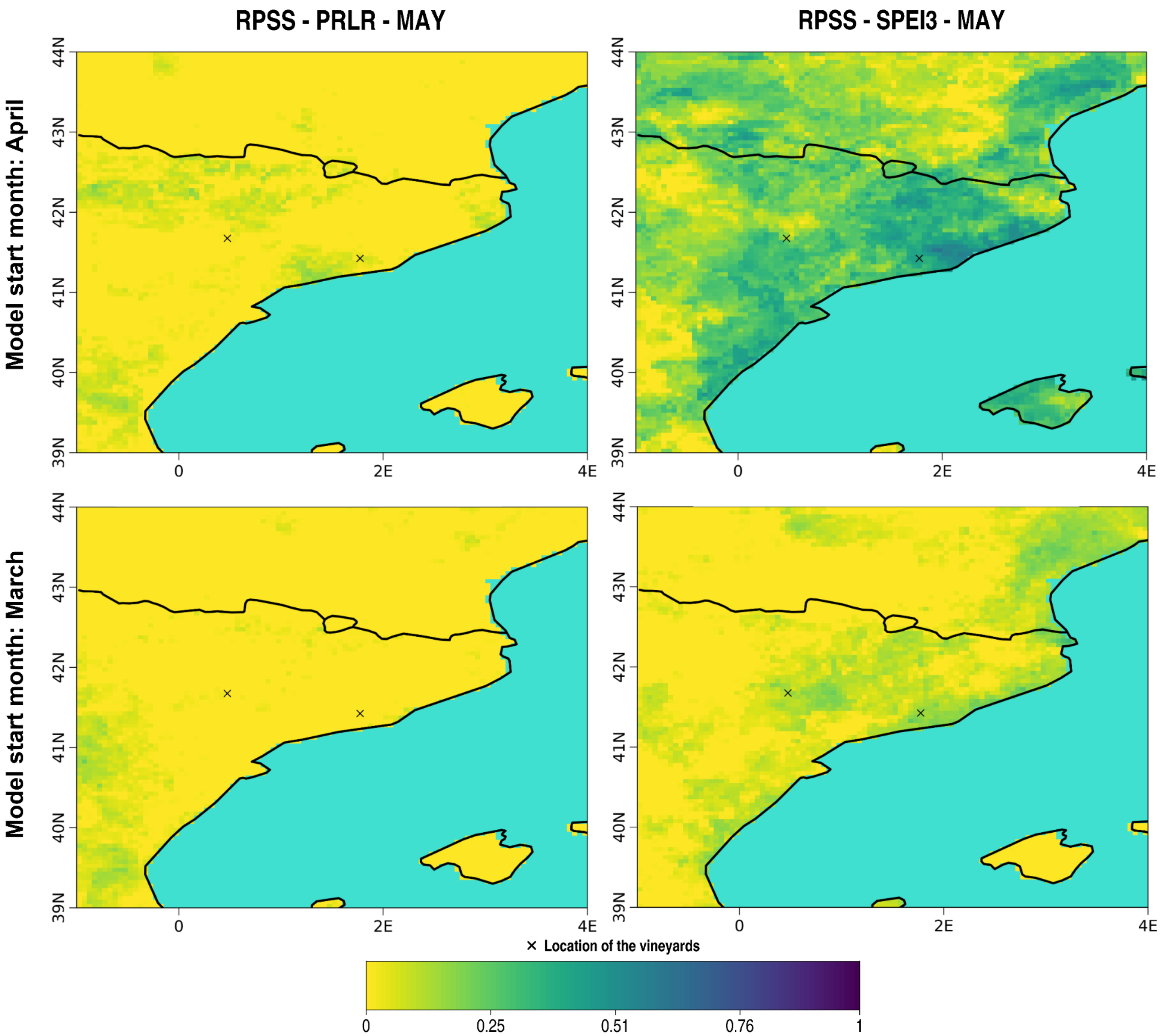

Figure 3: *Spatial distribution of RPSS values for precipitation (PRLR) and SPEI3 in May, based on predictions from models initialized in March and April.*

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

Bedia, J., Baño-Medina, J., Legasa, M. N., Iturbide, M., Manzanas, R., Herrera, S., ... & Gutiérrez, J. M. (2020). Statistical downscaling with the downscaleR package (v3. 1.0): Contribution to the VALUE intercomparison experiment. Geoscientific Model Development, 13(3), 1711-1735. DOI: 10.5194/gmd-13-1711-2020

Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. International journal of climatology, 34(10), 3001-3023. DOI: 10.1002/joc.3887

Duzenli, E., Ramon, J., 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. DOI: 10.5194/egusphere-egu24-19420.

Duzenli, E., Ramon, J., Torralba, V., Pickard, S., Bojovic, D., Trascasa-Castro, P., and Muñoz, Á. G. (2025, May. Assessing the added value of statistical downscaling to the predictive skill of global subseasonal temperature forecasts during the Paris 2024 Olympics. In EGU General Assembly Conference Abstracts. DOI: 10.5194/egusphere-egu25-18397.

Duzenli E., Torralba V., Delgado-Torres C., et al. Statistical downscaling of seasonal climate indicators for an agricultural case study: High-Resolution estimates of water availability and spring frost for vineyard applications (in preparation)

Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Hargreaves evapotranspiration equation. Journal of irrigation and drainage engineering, 129(1), 53-63. DOI: 10.1061/(ASCE)0733-9437(2003)129:1(53).

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., ... & Monge-Sanz, B. M. (2019). SEAS5: the new ECMWF seasonal forecast system. Geoscientific Model Development, 12(3), 1087-1117. DOI: 10.5194/gmd-12-1087-2019.

Olmo, M. E., and M. L. Bettolli, 2022: Statistical downscaling of daily precipitation over southeastern South America: Assessing the performance in extreme events. Int. J. Climatol., 42, 2, 1283–1302, DOI: 10.1002/joc.7303.

Schimanke S., and Coauthors, 2021: CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), Accessed on: 02-08-2024, DOI: 10.24381/cds.622a565a.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., & Stahl, K. (2015). Candidate distributions for climatological drought indices (SPI and SPEI). International Journal of Climatology, 35(13), 4027-4040. DOI: 10.1002/joc.4267.

Vicente-Serrano S.M., Santiago Beguería, Juan I. López-Moreno, (2010). A Multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index - SPEI. Journal of Climate 23: 1696-1718, DOI: 10.1175/2009JCLI2909.1.