



## Climate Services Downscale (CSDownscale): A statistical downscaling tool for (sub)seasonal to decadal climate predictions

**Eren Duzenli**<sup>1</sup>, Jaume Ramon Gamon<sup>1</sup>, Alba Llabres<sup>1</sup>, Verónica Torralba<sup>1</sup>, Lluís Palma Garcia<sup>1</sup>, Sara Moreno Montes<sup>1</sup>, Carlos Delgado-Torres<sup>1</sup>, Nuria Perez-Zanon<sup>1</sup>, Javier Corvillo Guerra<sup>1</sup>, and Raul Marcos<sup>1,2</sup>

<sup>1</sup>Barcelona Supercomputing Center, Earth Sciences Department, Barcelona, Spain

<sup>2</sup>Department of Applied Physics, University of Barcelona, Barcelona, Spain

Statistical downscaling is a technique that allows to obtain high-resolution climate information from the coarse-resolution Global Climate Model (GCM) outputs through the long-term relationship between the GCM output and a reference dataset such as in-situ observations. The key benefit of employing statistical downscaling (SD) methods over the dynamical approaches is their significantly less computational costs. The cost-effectiveness of these methods enables the processing of large hindcasts, including multi-model systems with numerous ensemble members, which is highly relevant for the users. Thus, a comprehensive tool that allows users to apply state-of-the-art statistical downscaling methods on climate variables is crucial. CSDownscale is a new generation R package that has been developed to statistically downscale subseasonal to seasonal to decadal climate predictions in the context of climate services, including its use in operational applications. The tool produces a downscaled field or time series using several bias correction, regression (i.e., linear and logistic) and analogs methods. Additionally, the package contains various interpolation methods such as nearest neighbor or bilinear approaches, which are used for regridding purposes. Users can easily combine these with bias correction and regression methods to perform downscaling. When applying these combined operations, the GCM data is initially interpolated to the resolution of the reference dataset, then the selected bias correction or regression method is implemented on the regridded data. However, the package also incorporates a method that infers the high-resolution values using a multi-linear regression with the four closest coarse-scale grid points, which skips the step of interpolation. Furthermore, the CSDownscale package includes an analogs based method, which looks for fields with similar conditions to the one being predicted and returns the high-resolution outcome of past conditions that are most similar, a certain number of similar fields or a combination of them. Finally, the CSDownscale package allows for the GCM data to be downscaled to either a reference grid space or a specific point location. All the methods are conceived to be done in cross-validation, that is, by excluding data from the specific time step being post-processed to avoid overfitting and, consequently, the overestimation of the actual prediction skill.