Package 'CSDownscale'

March 5, 2026

Title Statistical Downscaling of Climate Predictions

Version 0.0.2

Description Statistical downscaling and bias correction of climate predictions.
It includes implementations of commonly used methods such as Analogs,
Linear Regression, Logistic Regression, and Bias Correction techniques,
as well as interpolation functions for regridding and point-based applications.
It facilitates the production of high-resolution and local-scale climate
information from coarse-scale predictions, which is essential for impact analyses.
The package can be applied in a wide range of sectors and studies,
including agriculture, water management, energy, heatwaves, and other
climate-sensitive applications. The package was developed within the framework of
the European Union Horizon Europe projects Impetus4Change (101081555) and AS-
PECT (101081460),
the Wellcome Trust supported HARMONIZE project (224694/Z/21/Z), and the Spanish national project
BOREAS (PID2022-140673OA-I00). Implements the methods described in
Duzenli et al. (2024) <doi:10.5194/egusphere-egu24-19420>.

Depends R (>= 3.6.0)

Imports CSTools, abind, multiApply, nnet, plyr, s2dv, ClimProjDiags,
proxy, easyVerification

License GPL-3

URL https://gitlab.earth.bsc.es/es/csdownscale

BugReports https://gitlab.earth.bsc.es/es/csdownscale/-/issues

SystemRequirements cdo

Encoding UTF-8

RoxygenNote 7.3.1

NeedsCompilation no

Author BSC-CNS [aut, cph],
Jaume Ramon [aut] (ORCID: <https://orcid.org/0000-0003-2818-5206>),
Eren Duzenli [aut] (ORCID: <https://orcid.org/0000-0002-7296-4496>),
Llorenç Lledó [aut] (ORCID: <https://orcid.org/0000-0002-8628-6876>),
Lluís Palma [ctb],

Sara Moreno-Montes [ctb],
Carlos Delgado [ctb],
Nuria Perez-Zanon [ctb] (ORCID:
 <https://orcid.org/0000-0001-8568-3071>),
Javier Corvillo Guerra [ctb],
Raul Marcos [ctb],
Alba Llabrés-Brustenga [ctb],
Verónica Torralba [ctb],
Theertha Kariyathan [ctb],
Victòria Agudetse [cre]

Maintainer Victòria Agudetse <victoria.agudetse@bsc.es>

Contents

Analogs   *Downscaling using Analogs based on large scale fields.*

Description

This function performs a downscaling using Analogs. To compute the analogs given a coarse-scale field, the function looks for days with similar conditions in the historical observations. The analogs function determines the N best analogs based on Euclidian distance, distance correlation, or Spearman's correlation metrics. To downscale a local-scale variable, either the variable itself or another large-scale variable can be utilized as the predictor. In the first scenario, analogs are examined between the observation and model data of the same local-scale variable. In the latter scenario, the function identifies the day in the observation data that closely resembles the large-scale pattern of interest in the model. When it identifies the date of the best analog, the function extracts the corresponding local-scale variable for that day from the observation of the local scale variable. The used local-scale and large-scale variables can be retrieved from independent regions. The input data for the first case must include 'exp' and 'obs,' while in the second case, 'obs,' 'obsL,' and 'exp' are the required input fields. Users can perform the downscaling process over the

subregions that can be identified through the 'region' argument, instead of focusing on the entire area of the loaded data. The search of analogs must be done in the longest dataset posible, but might require high-memory computational resources. This is important since it is necessary to have a good representation of the possible states of the field in the past, and therefore, to get better analogs. The function can also look for analogs within a window of D days, but is the user who has to define that window. Otherwise, the function will look for analogs in the whole dataset. This function is intended to downscale climate prediction data (i.e., sub-seasonal, seasonal and decadal predictions) but can admit climate projections or reanalyses. It does not have constrains of specific region or variables to downscale.

Usage

```
Analogs(
  exp,
  obs,
  exp_lats = NULL,
  exp_lons = NULL,
  obs_lats,
  obs_lons,
  grid_exp,
  obsL = NULL,
  obsL_lats = NULL,
  obsL_lons = NULL,
  nanalogs = 3,
  fun_analog = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  time_dim = "time",
  member_dim = "member",
  metric = "dist",
  region = NULL,
  return_indices = FALSE,
  loocv_window = TRUE,
  ncores = NULL
)
```

Arguments

exp   an array with named dimensions containing the experimental field on the coarse scale for the variable targeted for downscaling (in case obsL is not provided) or for the large-scale variable used as the predictor (if obsL is provided). The object must have, at least, the dimensions latitude, longitude, start date and time. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'. Also, the object can be either hindcast or forecast data. However, if forecast data is provided, the loocv_window parameter should be selected as FALSE.

obs	an array with named dimensions containing the observational field for the variable targeted for downscaling. The object must have, at least, the dimensions latitude, longitude, start date and either time or window. The object is expected to be already subset for the desired region. Optionally, 'obs' can have the dimension 'window', containing the sampled fields into which the function will look for the analogs. Otherwise, the function will look for analogs using all the possible fields contained in obs.
exp_lats	a numeric vector containing the latitude values in 'exp'. Latitudes must range from -90 to 90.
exp_lons	a numeric vector containing the longitude values in 'exp'. Longitudes can range from -180 to 180 or from 0 to 360.
obs_lats	a numeric vector containing the latitude values in 'obs'. Latitudes must range from -90 to 90.
obs_lons	a numeric vector containing the longitude values in 'obs'. Longitudes can range from -180 to 180 or from 0 to 360.
grid_exp	a character vector with a path to an example file of the exp data. It can be either a path to another NetCDF file which to read the target grid from (a single grid must be defined in such file) or a character vector indicating the coarse grid to be passed to CDO, and it must be a grid recognised by CDO.
obsL	an 's2dv_cube' object with named dimensions containing the observational field of the large-scale variable.The object must have, at least, the dimensions latitude, longitude, start date and either time or window. The object is expected to be already subset for the desired region. Optionally, 'obsL' can have the dimension 'window', containing the sampled fields into which the function will look for the analogs. Otherwise, the function will look for analogs using all the possible fields contained in obs.
obsL_lats	a numeric vector containing the latitude values in 'obsL'. Latitudes must range from -90 to 90.
obsL_lons	a numeric vector containing the longitude values in 'obsL'. Longitudes can range from -180 to 180 or from 0 to 360.
nanalogs	an integer indicating the number of analogs to be searched.
fun_analog	a function to be applied over the found analogs. Only these options are valid: "mean", "wmean", "max", "min", "median" or NULL. If set to NULL (default), the function returns the found analogs.
lat_dim	a character vector indicating the latitude dimension name in the element 'data' in exp and obs. Default set to "lat".
lon_dim	a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".
sdate_dim	a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".
time_dim	a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".
member_dim	a character vector indicating the member dimension name in the element 'data' in exp and obs. Default set to "member".

metric	a character vector to select the analog specification method. Only these options are valid: "dist" (i.e., Euclidian distance), "dcor" (i.e., distance correlation) or "cor" (i.e., Spearman's .correlation). The default metric is "dist".
region	a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function uses the full obs grid as the downscaling region.
return_indices	a logical vector indicating whether to return the indices of the analogs together with the downscaled fields. The indices refer to the position of the element in the vector time * start_date. If 'obs' contain the dimension 'window', it will refer to the position of the element in the dimension 'window'. Default to FALSE.
loocv_window	a logical vector only to be used if 'obs' does not have the dimension 'window'. It indicates whether to apply leave-one-out cross-validation in the creation of the window. It is recommended to be set to TRUE. Default to TRUE.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

A list of three elements. 'data' contains the dowscaled field, 'lat' the downscaled latitudes, and 'lon' the downscaled longitudes. If fun_analog is set to NULL (default), the output array in 'data' also contains the dimension 'analog' with the best analog days.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

Ll. Lledó, <llorenc.lledo@ecmwf.int>

Examples

```
exp <- rnorm(15000)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5, time  = 30)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(27000)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5, time  = 30)
obs_lons <- seq(0,6, 6/14)
obs_lats <- seq(0,6, 6/11)
if (Sys.which("cdo") != "") {
downscaled_field <- Analogs(exp = exp, obs = obs, exp_lats = exp_lats, exp_lons = exp_lons,
                     obs_lats = obs_lats, obs_lons = obs_lons, grid_exp = 'r360x180')
}
```

CST_Analogs   *Downscaling using Analogs based on coarse scale fields.*

Description

This function performs a downscaling using Analogs. To compute the analogs given a coarse-scale field, the function looks for days with similar conditions in the historical observations. The analogs function determines the N best analogs based on Euclidian distance, distance correlation, or Spearman's correlation metrics. To downscale a local-scale variable, either the variable itself or another large-scale variable can be utilized as the predictor. In the first scenario, analogs are examined between the observation and model data of the same local-scale variable. In the latter scenario, the function identifies the day in the observation data that closely resembles the large-scale pattern of interest in the model. When it identifies the date of the best analog, the function extracts the corresponding local-scale variable for that day from the observation of the local scale variable. The used local-scale and large-scale variables can be retrieved from independent regions. The input data for the first case must include 'exp' and 'obs,' while in the second case, 'obs,' 'obsL,' and 'exp' are the required input fields. Users can perform the downscaling process over the subregions that can be identified through the 'region' argument, instead of focusing on the entire area of the loaded data. The search of analogs must be done in the longest dataset posible, but might require high-memory computational resources. This is important since it is necessary to have a good representation of the possible states of the field in the past, and therefore, to get better analogs. The function can also look for analogs within a window of D days, but is the user who has to define that window. Otherwise, the function will look for analogs in the whole dataset. This function is intended to downscale climate prediction data (i.e., sub-seasonal, seasonal and decadal predictions) but can admit climate projections or reanalyses. It does not have constrains of specific region or variables to downscale.

Usage

```
CST_Analogs(
  exp,
  obs,
  obsL = NULL,
  grid_exp,
  nanalogs = 3,
  fun_analog = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  time_dim = "time",
  member_dim = "member",
  metric = "dist",
  region = NULL,
  return_indices = FALSE,
  loocv_window = TRUE,
  ncores = NULL
)
```

Arguments

exp	an 's2dv_cube' object with named dimensions containing the experimental field on the coarse scale for the variable targeted for downscaling (in case obsL is not provided) or for the large-scale variable used as the predictor (if obsL is provided). The object must have, at least, the dimensions latitude, longitude, start date and time. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'. Also, the object can be either hindcast or forecast data. However, if forecast data is provided, the loocv_window parameter should be selected as FALSE.
obs	an 's2dv_cube' object with named dimensions containing the observational field for the variable targeted for downscaling. The object must have, at least, the dimensions latitude, longitude, start date and either time or window. The object is expected to be already subset for the desired region.
obsL	an 's2dv_cube' object with named dimensions containing the observational field of the large-scale variable. The object must have, at least, the dimensions latitude, longitude, start date and either time or window. The object is expected to be already subset for the desired region.
grid_exp	a character vector with a path to an example file of the exp data. It can be either a path to another NetCDF file which to read the target grid from (a single grid must be defined in such file) or a character vector indicating the coarse grid to be passed to CDO, and it must be a grid recognised by CDO.
nanalogs	an integer indicating the number of analogs to be searched
fun_analog	a function to be applied over the found analogs. Only these options are valid: "mean", "wmean", "max", "min", "median" or NULL. If set to NULL (default), the function returns the found analogs.
lat_dim	a character vector indicating the latitude dimension name in the element 'data' in exp and obs. Default set to "lat".
lon_dim	a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".
sdate_dim	a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".
time_dim	a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".
member_dim	a character vector indicating the member dimension name in the element 'data' in exp and obs. Default set to "member".
metric	a character vector to select the analog specification method. Only these options are valid: "dist" (i.e., Euclidian distance), "dcor" (i.e., distance correlation) or "cor" (i.e., Spearman's .correlation). The default metric is "dist".
region	a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function uses the full obs grid as the downscaling region.

return_indices	a logical vector indicating whether to return the indices of the analogs together with the downscaled fields. Default to FALSE.
loocv_window	a logical vector only to be used if 'obs' does not have the dimension 'window'. It indicates whether to apply leave-one-out cross-validation in the creation of the window. In this procedure, all data from the corresponding year are excluded (e.g., all days from that year) so that the analogs are selected only from the remaining years. It is recommended to be set to TRUE. Default to TRUE.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

An 's2dv_cube' object. The element 'data' contains the dowscaled field, 'lat' the downscaled latitudes, and 'lon' the downscaled longitudes. If fun_analog is set to NULL (default), the output array in 'data' also contains the dimension 'analog' with the best analog days.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

Examples

```
exp <- rnorm(15000)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5, time  = 30)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(27000)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5, time  = 30)
obs_lons <- seq(0,6, 6/14)
obs_lats <- seq(0,6, 6/11)
exp <- CSTools::s2dv_cube(data = exp, coords = list(lat = exp_lats, lon = exp_lons))
obs <- CSTools::s2dv_cube(data = obs, coords = list(lat = obs_lats, lon = obs_lons))
if (Sys.which("cdo") != "") {
downscaled_field <- CST_Analogs(exp = exp, obs = obs, grid_exp = 'r360x180')
}
```

CST_Intbc   *Downscaling using interpolation and bias adjustment.*

Description

This function performs a downscaling using an interpolation and a later bias adjustment. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
CST_Intbc(
  exp,
  obs,
  exp_cor = NULL,
  target_grid,
  bc_method,
  int_method = NULL,
  points = NULL,
  method_point_interp = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  member_dim = "member",
  time_dim = "time",
  region = NULL,
  ncores = NULL,
  loocv = TRUE,
  ...
)
```

Arguments

exp	an 's2dv object' containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude, longitude, start date and member. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.
obs	an 's2dv object' containing the observational field. The object must have, at least, the dimensions latitude, longitude and start date. The object is expected to be already subset for the desired region.
exp_cor	an optional 's2dv_cube' object with named dimensions containing the seasonal forecast experiment data. If the forecast is provided, it will be downscaled using the hindcast and observations; if not provided, the hindcast will be downscaled instead. The default value is NULL.
target_grid	a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.
bc_method	a character vector indicating the bias adjustment method to be applied after the interpolation. Accepted methods are 'quantile_mapping', 'bias', 'evmos', 'mse_min', 'crps_min', 'rpc-based'. The abbreviations 'qm' can also be used.
int_method	a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2 or older version is required.

points   a list of two elements containing the point latitudes and longitudes of the lo-
                        cations to downscale the model data. The list must contain the two elements
                        named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscal-
                        ing is to a point location, only regular grids are allowed for exp and obs. Only
                        needed if the downscaling is to a point location.

method_point_interp
                        a character vector indicating the interpolation method to interpolate model grid-
                        ded data into the point locations. Accepted methods are "nearest", "bilinear",
                        "9point", "invdist4nn", "NE", "NW", "SE", "SW". Only needed if the down-
                        scaling is to a point location.

lat_dim   a character vector indicating the latitude dimension name in the element 'data'
                        in exp and obs. Default set to "lat".

lon_dim   a character vector indicating the longitude dimension name in the element 'data'
                        in exp and obs. Default set to "lon".

sdate_dim   a character vector indicating the start date dimension name in the element 'data'
                        in exp and obs. Default set to "sdate".

member_dim   a character vector indicating the member dimension name in the element 'data'
                        in exp and obs. Default set to "member".

time_dim   a character vector indicating the time dimension name in the element 'data' in
                        exp and obs. Default set to "time".

region   a numeric vector indicating the borders of the downscaling region. It consists of
                        four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the
                        left border, while lonmax refers to the right border. latmin indicates the lower
                        border, whereas latmax indicates the upper border. If set to NULL (default), the
                        function takes the first and last elements of the latitudes and longitudes in obs.

ncores   an integer indicating the number of cores to use in parallel computation. The
                        default value is NULL.

loocv   a logical indicating whether to apply leave-one-out cross-validation when ap-
                        plying the bias correction. In this procedure, all values from the corresponding
                        year are excluded, so that when building the correction function for a given year,
                        no data from that year are used. Default to TRUE.

...   additional arguments passed to internal methods

Value

An 's2dv' object. The element 'data' contains the dowscaled field, 'lat' the downscaled latitudes,
and 'lon' the downscaled longitudes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(900)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
exp <- CSTools::s2dv_cube(data = exp, coords = list(lat = exp_lats, lon = exp_lons))
obs <- CSTools::s2dv_cube(data = obs, coords = list(lat = obs_lats, lon = obs_lons))
if (Sys.which("cdo") != "") {
res <- CST_Intbc(exp = exp, obs = obs, target_grid = 'r1280x640',
                 bc_method = 'bias', int_method = 'conservative')
}
```

CST_Interpolation	*Regrid or interpolate gridded data to a point location.*

Description

This function interpolates gridded model data from one grid to another (regrid) or interpolates gridded model data to a set of point locations. The gridded model data can be either global or regional. In the latter case, the region is defined by the user. It does not have constrains of specific region or variables to downscale.

Usage

```
CST_Interpolation(
  exp,
  points = NULL,
  method_remap = NULL,
  target_grid = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  region = NULL,
  method_point_interp = NULL,
  ncores = NULL
)
```

Arguments

exp   s2dv object containing the experimental field on the coarse scale for which the
               downscaling is aimed. The object must have, at least, the dimensions latitude
               and longitude. The field data is expected to be already subset for the desired re-
               gion. Data can be in one or two integrated regions, e.g., crossing the Greenwich
               meridian. To get the correct results in the latter case, the borders of the region
               should be specified in the parameter 'region'. See parameter 'region'.

points	a list of two elements containing the point latitudes and longitudes of the locations to downscale the model data. The list must contain the two elements named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscaling is to a point location, only regular grids are allowed for exp and obs. Only needed if the downscaling is to a point location.
method_remap	a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2 or older version is required.
target_grid	a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.
lat_dim	a character vector indicating the latitude dimension name in the element 'exp' and/or 'points'. Default set to "lat".
lon_dim	a character vector indicating the longitude dimension name in the element 'exp' and/or 'points'. Default set to "lon".
region	a numeric vector indicating the borders of the interpolation region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in exp.
method_point_interp	
	a character vector indicating the interpolation method to interpolate model gridded data into the point locations. Accepted methods are "nearest", "bilinear", "9point", "invdist4nn", "NE", "NW", "SE", "SW".
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

An s2dv object containing the dowscaled field.

Author(s)

J. Ramon, <jaume.ramon@bsc.es>

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5, time  = 1)
lons <- 1:5
lats <- 1:4
exp <- CSTools::s2dv_cube(data = exp, coords = list(lat = lats, lon = lons))
if (Sys.which("cdo") != "") {
res <- CST_Interpolation(exp = exp, method_remap = 'conservative', target_grid = 'r1280x640')
}
```

CST_Intlr   *Downscaling using interpolation and linear regression.*

Description

This function performs a downscaling using an interpolation and a linear regression. Different methodologies that employ linear regressions are available. See parameter 'lr_method' for more information. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
CST_Intlr(
  exp,
  obs,
  exp_cor = NULL,
  lr_method,
  target_grid = NULL,
  points = NULL,
  int_method = NULL,
  method_point_interp = NULL,
  predictors = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  time_dim = "time",
  member_dim = "member",
  large_scale_predictor_dimname = "vars",
  loocv = TRUE,
  region = NULL,
  ncores = NULL
)
```

Arguments

exp   an 's2dv object' containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude, longitude, start date and member. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.

obs   an 's2dv object' containing the observational field. The object must have, at
                        least, the dimensions latitude, longitude and start date. The object is expected to
                        be already subset for the desired region.

exp_cor   an optional 's2dv_cube' object with named dimensions containing the seasonal
                        forecast experiment data. If provided, the forecast will be downscaled using
                        the hindcast and observations; if not, the hindcast will be downscaled instead.
                        The default value is NULL. Since the Intlr function is built separately for each
                        ensemble member, it is not recommended for forecast cases where the mem-
                        ber_dim length of exp_cor differs from that of exp. In such situations, the use
                        of other functions in the package is more appropriate.

lr_method   a character vector indicating the linear regression method to be applied. Ac-
                        cepted methods are 'basic', 'large-scale' and '9nn'. The 'basic' method fits a
                        linear regression using high resolution observations as predictands and the in-
                        terpolated model data as predictor. Then, the regression equation is applied to
                        the interpolated model data to correct the interpolated values. The 'large-scale'
                        method fits a linear regression with large-scale predictors (e.g. teleconnection
                        indices) as predictors and high-resolution observations as predictands. Finally,
                        the '9nn' method uses a linear regression with the nine nearest neighbours as
                        predictors and high-resolution observations as predictands. Instead of construct-
                        ing a regression model using all nine predictors, principal component analysis is
                        applied to the data of neighboring grids to reduce the dimension of the predic-
                        tors. The linear regression model is then built using the principal components
                        that explain 95% of the variance. The '9nn' method does not require a pre-
                        interpolation process.

target_grid   a character vector indicating the target grid to be passed to CDO. It must be a
                        grid recognised by CDO or a NetCDF file.

points   a list of two elements containing the point latitudes and longitudes of the lo-
                        cations to downscale the model data. The list must contain the two elements
                        named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscal-
                        ing is to a point location, only regular grids are allowed for exp and obs. Only
                        needed if the downscaling is to a point location.

int_method   a character vector indicating the regridding method to be passed to CDORemap.
                        Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be
                        used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2
                        or older version is required.

method_point_interp
                        a character vector indicating the interpolation method to interpolate model grid-
                        ded data into the point locations. Accepted methods are "nearest", "bilinear",
                        "9point", "invdist4nn", "NE", "NW", "SE", "SW".

predictors   an array with large-scale data to be used in the 'large-scale' method. Only
                        needed if the linear regression method is set to 'large-scale'. It must have, at
                        least the dimension start date and another dimension whose name has to be
                        specified in the parameter 'large_scale_predictor_dimname'. It should contain
                        as many elements as the number of large-scale predictors.

lat_dim   a character vector indicating the latitude dimension name in the element 'data'
                        in exp and obs. Default set to "lat".

lon_dim	a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".
sdate_dim	a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".
time_dim	a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".
member_dim	a character vector indicating the member dimension name in the element 'data' in exp and obs. Default set to "member".
large_scale_predictor_dimname	
	a character vector indicating the name of the dimension in 'predictors' that contain the predictor variables. See parameter 'predictors'.
loocv	a logical indicating whether to apply leave-one-out cross-validation when generating the linear regressions. In this procedure, all values from the corresponding year are excluded, so that when building the regression model for a given year, none of that year's data are used. Default to TRUE.
region	a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in obs.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

A list with two s2dv_cube objects, exp and obs, each with elements 'data' containing the downscaled field, 'coords' containing the coordinate information, 'dims' describing the dimension structure, and 'attrs' containing the associated attributes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(900)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
exp <- CSTools::s2dv_cube(data = exp, coords = list(lat = exp_lats, lon = exp_lons))
obs <- CSTools::s2dv_cube(data = obs, coords = list(lat = obs_lats, lon = obs_lons))
if (Sys.which("cdo") != "") {
res <- CST_Intlr(exp = exp, obs = obs, target_grid = 'r1280x640',
```

```
                    lr_method = 'basic', int_method = 'conservative')
    }
```

CST_LogisticReg   *Downscaling using interpolation and logistic regression.*

Description

This function performs a downscaling using an interpolation and a logistic regression. See `multinom` for further details. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
CST_LogisticReg(
  exp,
  obs,
  exp_cor = NULL,
  target_grid,
  int_method = NULL,
  log_reg_method = "ens_mean",
  probs_cat = c(1/3, 2/3),
  return_most_likely_cat = FALSE,
  points = NULL,
  method_point_interp = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  member_dim = "member",
  time_dim = "time",
  region = NULL,
  loocv = TRUE,
  ncores = NULL
)
```

Arguments

exp   an 's2dv object' with named dimensions containing the experimental field on
                 the coarse scale for which the downscaling is aimed. The object must have, at
                 least, the dimensions latitude, longitude, start date and member. The object is
                 expected to be already subset for the desired region. Data can be in one or two
                 integrated regions, e.g., crossing the Greenwich meridian. To get the correct
                 results in the latter case, the borders of the region should be specified in the
                 parameter 'region'. See parameter 'region'.

obs   an 's2dv object' with named dimensions containing the observational field. The
                         object must have, at least, the dimensions latitude, longitude and start date. The
                         object is expected to be already subset for the desired region.

exp_cor   an optional array with named dimensions containing the seasonal forecast exper-
                         iment data. If the forecast is provided, it will be downscaled using the hindcast
                         and observations; if not provided, the hindcast will be downscaled instead. The
                         default value is NULL.

target_grid   a character vector indicating the target grid to be passed to CDO. It must be a
                         grid recognised by CDO or a NetCDF file.

int_method   a character vector indicating the regridding method to be passed to CDORemap.
                         Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be
                         used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2
                         or older version is required.

log_reg_method   a character vector indicating the logistic regression method to be used. Accepted
                         methods are "ens_mean", "ens_mean_sd", "sorted_members". "ens_mean" uses
                         the ensemble mean anomalies as predictors in the logistic regression, "ens_mean_sd"
                         uses the ensemble mean anomalies and the ensemble spread (computed as the
                         standard deviation of all the members) as predictors in the logistic regression,
                         and "sorted_members" considers all the members ordered decreasingly as pre-
                         dictors in the logistic regression. Default method is "ens_mean".

probs_cat   a numeric vector indicating the percentile thresholds separating the climatolog-
                         ical distribution into different classes (categories). Default to c(1/3, 2/3). See
                         convert2prob.

return_most_likely_cat
                         if TRUE, the function returns the most likely category. If FALSE, the function
                         returns the probabilities for each category. Default to FALSE.

points   a list of two elements containing the point latitudes and longitudes of the lo-
                         cations to downscale the model data. The list must contain the two elements
                         named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscal-
                         ing is to a point location, only regular grids are allowed for exp and obs. Only
                         needed if the downscaling is to a point location.

method_point_interp
                         a character vector indicating the interpolation method to interpolate model grid-
                         ded data into the point locations. Accepted methods are "nearest", "bilinear",
                         "9point", "invdist4nn", "NE", "NW", "SE", "SW". Only needed if the down-
                         scaling is to a point location.

lat_dim   a character vector indicating the latitude dimension name in the element 'data'
                         in exp and obs. Default set to "lat".

lon_dim   a character vector indicating the longitude dimension name in the element 'data'
                         in exp and obs. Default set to "lon".

sdate_dim   a character vector indicating the start date dimension name in the element 'data'
                         in exp and obs. Default set to "sdate".

member_dim   a character vector indicating the member dimension name in the element 'data'
                         in exp and obs. Default set to "member".

time_dim	a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".
region	a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in obs.
loocv	a logical vector indicating whether to perform leave-one-out cross-validation in the fitting of the logistic regression. In this procedure, all values from the corresponding year are excluded, so that when fitting the model for a given year, none of that year's data is used. Default to TRUE.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

A list with two s2dv_cube objects, exp and obs, each with elements 'data' containing the down-scaled data, that could be either in the form of probabilities for each category or the most likely category. 'coords' containing the coordinate information, 'dims' describing the dimension structure, and 'attrs' containing the associated attributes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

See Also

multinom

Examples

```
exp <- rnorm(1500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 15)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(2700)
dim(obs) <- c(lat = 12, lon = 15, sdate = 15)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
exp <- CSTools::s2dv_cube(data = exp, coords = list(lat = exp_lats, lon = exp_lons))
obs <- CSTools::s2dv_cube(data = obs, coords = list(lat = obs_lats, lon = obs_lons))
if (Sys.which("cdo") != "") {
res <- CST_LogisticReg(exp = exp, obs = obs, int_method = 'bil', target_grid = 'r1280x640',
                       probs_cat = c(1/3, 2/3))
}
```

Intbc	*Downscaling using interpolation and bias adjustment.*

Description

This function performs a downscaling using an interpolation and a later bias adjustment. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
Intbc(
  exp,
  obs,
  exp_cor = NULL,
  exp_lats,
  exp_lons,
  obs_lats,
  obs_lons,
  target_grid,
  bc_method,
  int_method = NULL,
  points = NULL,
  method_point_interp = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  time_dim = "time",
  member_dim = "member",
  source_file_exp = NULL,
  source_file_obs = NULL,
  region = NULL,
  ncores = NULL,
  loocv = TRUE,
  ...
)
```

Arguments

exp	an array with named dimensions containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude, longitude, start date and member. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in

the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.

obs   an array with named dimensions containing the observational field. The object must have, at least, the dimensions latitude, longitude and start date. The object is expected to be already subset for the desired region.

exp_cor   an optional array with named dimensions containing the seasonal forecast experiment data. If the forecast is provided, it will be downscaled using the hindcast and observations; if not provided, the hindcast will be downscaled instead. The default value is NULL.

exp_lats   a numeric vector containing the latitude values in 'exp'. Latitudes must range from -90 to 90.

exp_lons   a numeric vector containing the longitude values in 'exp'. Longitudes can range from -180 to 180 or from 0 to 360.

obs_lats   a numeric vector containing the latitude values in 'obs'. Latitudes must range from -90 to 90.

obs_lons   a numeric vector containing the longitude values in 'obs'. Longitudes can range from -180 to 180 or from 0 to 360.

target_grid   a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.

bc_method   a character vector indicating the bias adjustment method to be applied after the interpolation. Accepted methods are 'quantile_mapping', 'bias', 'evmos', 'mse_min', 'crps_min', 'rpc-based'. The abbreviations 'qm' can also be used.

int_method   a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2 or older version is required.

points   a list of two elements containing the point latitudes and longitudes of the locations to downscale the model data. The list must contain the two elements named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscaling is to a point location, only regular grids are allowed for exp and obs. Only needed if the downscaling is to a point location.

method_point_interp
                       a character vector indicating the interpolation method to interpolate model gridded data into the point locations. Accepted methods are "nearest", "bilinear", "9point", "invdist4nn", "NE", "NW", "SE", "SW". Only needed if the downscaling is to a point location.

lat_dim   a character vector indicating the latitude dimension name in the element 'data' in exp and obs. Default set to "lat".

lon_dim   a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".

sdate_dim   a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".

time_dim   a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".

member_dim   a character vector indicating the member dimension name in the element 'data'
                    in exp and obs. Default set to "member".

source_file_exp

                    a character vector with a path to an example file of the exp data. Only needed if
                    the downscaling is to a point location.

source_file_obs

                    a character vector with a path to an example file of the obs data. Only needed if
                    the downscaling is to a point location.

region   a numeric vector indicating the borders of the downscaling region. It consists of
                    four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the
                    left border, while lonmax refers to the right border. latmin indicates the lower
                    border, whereas latmax indicates the upper border. If set to NULL (default), the
                    function takes the first and last elements of the latitudes and longitudes in obs.

ncores   an integer indicating the number of cores to use in parallel computation. The
                    default value is NULL.

loocv   a logical indicating whether to apply leave-one-out cross-validation when ap-
                    plying the bias correction. In this procedure, all values from the corresponding
                    year are excluded, so that when building the correction function for a given year,
                    no data from that year are used. Default to TRUE.

...   additional arguments passed to internal methods

Value

An list of three elements. 'data' contains the dowscaled field, 'lat' the downscaled latitudes, and
'lon' the downscaled longitudes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

See Also

[BiasCorrection](#)

[Calibration](#)

[QuantileMapping](#)

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(900)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
```

```
if (Sys.which("cdo") != "") {
res <- Intbc(exp = exp, obs = obs, exp_lats = exp_lats, exp_lons = exp_lons,
             obs_lats = obs_lats, obs_lons = obs_lons, target_grid = 'r1280x640',
             bc_method = 'bias', int_method = 'conservative')
}
```

Interpolation	*Regrid or interpolate gridded data to a point location.*

Description

This function interpolates gridded model data from one grid to another (regrid) or interpolates gridded model data to a set of point locations. The gridded model data can be either global or regional. In the latter case, the region is defined by the user. It does not have constrains of specific region or variables to downscale.

Usage

```
Interpolation(
  exp,
  lats,
  lons,
  points = NULL,
  source_file = NULL,
  method_remap = NULL,
  target_grid = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  region = NULL,
  method_point_interp = NULL,
  ncores = NULL
)
```

Arguments

exp	an array with named dimensions containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude and longitude. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.
lats	a numeric vector containing the latitude values. Latitudes must range from -90 to 90.
lons	a numeric vector containing the longitude values. Longitudes can range from -180 to 180 or from 0 to 360.

points	a list of two elements containing the point latitudes and longitudes of the locations to downscale the model data. The list must contain the two elements named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscaling is to a point location, only regular grids are allowed for exp and obs. Only needed if the downscaling is to a point location.
source_file	a character vector with a path to an example file of the exp data. Only needed if the downscaling is to a point location.
method_remap	a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be used, CDO_1.9.8 or newer version is required.
target_grid	a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.
lat_dim	a character vector indicating the latitude dimension name in the element 'exp' and/or 'points'. Default set to "lat".
lon_dim	a character vector indicating the longitude dimension name in the element 'exp' and/or 'points'. Default set to "lon".
region	a numeric vector indicating the borders of the interpolation region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in exp.
method_point_interp	
	a character vector indicating the interpolation method to interpolate model gridded data into the point locations. Accepted methods are "nearest", "bilinear", "9point", "invdist4nn", "NE", "NW", "SE", "SW". Only needed if the downscaling is to a point location.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

An list of three elements. 'data' contains the dowscaled field, 'lat' the downscaled latitudes, and 'lon' the downscaled longitudes.

Author(s)

J. Ramon, <jaume.ramon@bsc.es>

Ll. Lledó, <llorenc.lledo@ecmwf.int>

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5, time  = 1)
lons <- 1:5
lats <- 1:4
if (Sys.which("cdo") != "") {
res <- Interpolation(exp = exp, lats = lats, lons = lons,
```

```
                        method_remap = 'conservative', target_grid = 'r1280x640')
}
```

Intlr   *Downscaling using interpolation and linear regression.*

Description

This function performs a downscaling using an interpolation and a linear regression. Different methodologies that employ linear regressions are available. See parameter 'lr_method' for more information. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
Intlr(
  exp,
  obs,
  exp_cor = NULL,
  exp_lats,
  exp_lons,
  obs_lats,
  obs_lons,
  lr_method,
  target_grid = NULL,
  points = NULL,
  int_method = NULL,
  method_point_interp = NULL,
  source_file_exp = NULL,
  source_file_obs = NULL,
  predictors = NULL,
  lat_dim = "lat",
  lon_dim = "lon",
  sdate_dim = "sdate",
  time_dim = "time",
  member_dim = "member",
  region = NULL,
  large_scale_predictor_dimname = "vars",
  loocv = TRUE,
  ncores = NULL
)
```

Arguments

exp	an array with named dimensions containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude, longitude and start date. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.
obs	an array with named dimensions containing the observational field. The object must have, at least, the dimensions latitude, longitude and start date. The object is expected to be already subset for the desired region.
exp_cor	an optional 's2dv_cube' object with named dimensions containing the seasonal forecast experiment data. If provided, the forecast will be downscaled using the hindcast and observations; if not, the hindcast will be downscaled instead. The default value is NULL. Since the Intlr function is built separately for each ensemble member, it is not recommended for forecast cases where the member_dim length of exp_cor differs from that of exp. In such situations, the use of other functions in the package is more appropriate.
exp_lats	a numeric vector containing the latitude values in 'exp'. Latitudes must range from -90 to 90.
exp_lons	a numeric vector containing the longitude values in 'exp'. Longitudes can range from -180 to 180 or from 0 to 360.
obs_lats	a numeric vector containing the latitude values in 'obs'. Latitudes must range from -90 to 90.
obs_lons	a numeric vector containing the longitude values in 'obs'. Longitudes can range from -180 to 180 or from 0 to 360.
lr_method	a character vector indicating the linear regression method to be applied. Accepted methods are 'basic', 'large-scale' and '9nn'. The 'basic' method fits a linear regression using high resolution observations as predictands and the interpolated model data as predictor. Then, the regression equation is applied to the interpolated model data to correct the interpolated values. The 'large-scale' method fits a linear regression with large-scale predictors (e.g. teleconnection indices) as predictors and high-resolution observations as predictands. Finally, the '9nn' method uses a linear regression with the nine nearest neighbours as predictors and high-resolution observations as predictands. Instead of constructing a regression model using all nine predictors, principal component analysis is applied to the data of neighboring grids to reduce the dimension of the predictors. The linear regression model is then built using the principal components that explain 95% of the variance. The '9nn' method does not require a pre-interpolation process.
target_grid	a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.
points	a list of two elements containing the point latitudes and longitudes of the locations to downscale the model data. The list must contain the two elements

named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscaling is to a point location, only regular grids are allowed for exp and obs. Only needed if the downscaling is to a point location.

int_method   a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2 or older version is required.

method_point_interp

a character vector indicating the interpolation method to interpolate model gridded data into the point locations. Accepted methods are "nearest", "bilinear", "9point", "invdist4nn", "NE", "NW", "SE", "SW".

source_file_exp

a character vector with a path to an example file of the exp data. Only needed if the downscaling is to a point location.

source_file_obs

a character vector with a path to an example file of the obs data. Only needed if the downscaling is to a point location.

predictors   an array with large-scale data to be used in the 'large-scale' method. Only needed if the linear regression method is set to 'large-scale'. It must have, at least the dimension start date and another dimension whose name has to be specified in the parameter 'large_scale_predictor_dimname'. It should contain as many elements as the number of large-scale predictors.

lat_dim   a character vector indicating the latitude dimension name in the element 'data' in exp and obs. Default set to "lat".

lon_dim   a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".

sdate_dim   a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".

time_dim   a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".

member_dim   a character vector indicating the member dimension name in the element 'data' in exp and obs. Default set to "member".

region   a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in obs.

large_scale_predictor_dimname

a character vector indicating the name of the dimension in 'predictors' that contain the predictor variables. See parameter 'predictors'.

loocv   a logical indicating whether to apply leave-one-out cross-validation when generating the linear regressions. In this procedure, all values from the corresponding year are excluded, so that when building the regression model for a given year, none of that year's data are used. Default to TRUE.

ncores   an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

A list of three elements. 'data' contains the dowscaled field, 'lat' the downscaled latitudes, and 'lon' the downscaled longitudes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

Examples

```
exp <- rnorm(500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 5)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(900)
dim(obs) <- c(lat = 12, lon = 15, sdate = 5)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
if (Sys.which("cdo") != "") {
res <- Intlr(exp = exp, obs = obs, exp_lats = exp_lats, exp_lons = exp_lons,
             obs_lats = obs_lats, obs_lons = obs_lons, target_grid = 'r1280x640',
             lr_method = 'basic', int_method = 'conservative')
}
```

LogisticReg   *Downscaling using interpolation and logistic regression.*

Description

This function performs a downscaling using an interpolation and a logistic regression. See multinom for further details. It is recommended that the observations are passed already in the target grid. Otherwise, the function will also perform an interpolation of the observed field into the target grid. The coarse scale and observation data can be either global or regional. In the latter case, the region is defined by the user. In principle, the coarse and observation data are intended to be of the same variable, although different variables can also be admitted.

Usage

```
LogisticReg(
  exp,
  obs,
  exp_cor = NULL,
  exp_lats,
  exp_lons,
  obs_lats,
  obs_lons,
```

```
    target_grid,
    int_method = NULL,
    log_reg_method = "ens_mean",
    probs_cat = c(1/3, 2/3),
    return_most_likely_cat = FALSE,
    points = NULL,
    method_point_interp = NULL,
    lat_dim = "lat",
    lon_dim = "lon",
    sdate_dim = "sdate",
    member_dim = "member",
    time_dim = "time",
    source_file_exp = NULL,
    source_file_obs = NULL,
    region = NULL,
    loocv = TRUE,
    ncores = NULL
)
```

Arguments

exp	an array with named dimensions containing the experimental field on the coarse scale for which the downscaling is aimed. The object must have, at least, the dimensions latitude, longitude, start date and member. The object is expected to be already subset for the desired region. Data can be in one or two integrated regions, e.g., crossing the Greenwich meridian. To get the correct results in the latter case, the borders of the region should be specified in the parameter 'region'. See parameter 'region'.
obs	an array with named dimensions containing the observational field. The object must have, at least, the dimensions latitude, longitude and start date. The object is expected to be already subset for the desired region.
exp_cor	an optional array with named dimensions containing the seasonal forecast experiment data. If the forecast is provided, it will be downscaled using the hindcast and observations; if not provided, the hindcast will be downscaled instead. The default value is NULL.
exp_lats	a numeric vector containing the latitude values in 'exp'. Latitudes must range from -90 to 90.
exp_lons	a numeric vector containing the longitude values in 'exp'. Longitudes can range from -180 to 180 or from 0 to 360.
obs_lats	a numeric vector containing the latitude values in 'obs'. Latitudes must range from -90 to 90.
obs_lons	a numeric vector containing the longitude values in 'obs'. Longitudes can range from -180 to 180 or from 0 to 360.
target_grid	a character vector indicating the target grid to be passed to CDO. It must be a grid recognised by CDO or a NetCDF file.
int_method	a character vector indicating the regridding method to be passed to CDORemap. Accepted methods are "con", "bil", "bic", "nn", "con2". If "nn" method is to be

used, CDO_1.9.8 or newer version is required. For method "con2", CDO_2.2.2 or older version is required.

log_reg_method   a character vector indicating the logistic regression method to be used. Accepted methods are "ens_mean", "ens_mean_sd", "sorted_members". "ens_mean" uses the ensemble mean anomalies as predictors in the logistic regression, "ens_mean_sd" uses the ensemble mean anomalies and the ensemble spread (computed as the standard deviation of all the members) as predictors in the logistic regression, and "sorted_members" considers all the members ordered decreasingly as predictors in the logistic regression. Default method is "ens_mean".

probs_cat   a numeric vector indicating the percentile thresholds separating the climatological distribution into different classes (categories). Default to c(1/3, 2/3). See convert2prob.

return_most_likely_cat

if TRUE, the function returns the most likely category. If FALSE, the function returns the probabilities for each category. Default to FALSE.

points   a list of two elements containing the point latitudes and longitudes of the locations to downscale the model data. The list must contain the two elements named as indicated in the parameters 'lat_dim' and 'lon_dim'. If the downscaling is to a point location, only regular grids are allowed for exp and obs. Only needed if the downscaling is to a point location.

method_point_interp

a character vector indicating the interpolation method to interpolate model gridded data into the point locations. Accepted methods are "nearest", "bilinear", "9point", "invdist4nn", "NE", "NW", "SE", "SW". Only needed if the downscaling is to a point location.

lat_dim   a character vector indicating the latitude dimension name in the element 'data' in exp and obs. Default set to "lat".

lon_dim   a character vector indicating the longitude dimension name in the element 'data' in exp and obs. Default set to "lon".

sdate_dim   a character vector indicating the start date dimension name in the element 'data' in exp and obs. Default set to "sdate".

member_dim   a character vector indicating the member dimension name in the element 'data' in exp and obs. Default set to "member".

time_dim   a character vector indicating the time dimension name in the element 'data' in exp and obs. Default set to "time".

source_file_exp

a character vector with a path to an example file of the exp data. Only needed if the downscaling is to a point location.

source_file_obs

a character vector with a path to an example file of the obs data. Only needed if the downscaling is to a point location.

region   a numeric vector indicating the borders of the downscaling region. It consists of four elements in this order: lonmin, lonmax, latmin, latmax. lonmin refers to the left border, while lonmax refers to the right border. latmin indicates the lower border, whereas latmax indicates the upper border. If set to NULL (default), the function takes the first and last elements of the latitudes and longitudes in obs.

loocv	a logical vector indicating whether to perform leave-one-out cross-validation in the fitting of the logistic regression. In this procedure, all values from the corresponding year are excluded, so that when fitting the model for a given year, none of that year's data is used. Default to TRUE.
ncores	an integer indicating the number of cores to use in parallel computation. The default value is NULL.

Value

A list of three elements. 'data' contains the dowscaled data, that could be either in the form of probabilities for each category or the most likely category. 'lat' contains the downscaled latitudes, and 'lon' the downscaled longitudes.

Author(s)

J. Ramon, <jaumeramong@gmail.com>

E. Duzenli, <eren.duzenli@bsc.es>

See Also

[multinom](#)

Examples

```
exp <- rnorm(1500)
dim(exp) <- c(member = 5, lat = 4, lon = 5, sdate = 15)
exp_lons <- 1:5
exp_lats <- 1:4
obs <- rnorm(2700)
dim(obs) <- c(lat = 12, lon = 15, sdate = 15)
obs_lons <- seq(1,5, 4/14)
obs_lats <- seq(1,4, 3/11)
if (Sys.which("cdo") != "") {
res <- LogisticReg(exp = exp, obs = obs, exp_lats = exp_lats, exp_lons = exp_lons,
                   obs_lats = obs_lats, obs_lons = obs_lons, int_method = 'bil',
                   target_grid = 'r1280x640', probs_cat = c(1/3, 2/3))
}
```

Index