Package ‘CDFt’

October 12, 2022

Version 1.2

Date 2021-02-16

Title Downscaling and Bias Correction via Non-Parametric CDF-Transform

Author Mathieu Vrac <mathieu.vrac@lsce.ipsl.fr> and Paul-Antoine Michelangeli <pam@climpact.com>

Maintainer Mathieu Vrac <mathieu.vrac@lsce.ipsl.fr>

Depends R (>= 1.8.0), stats

Description

Statistical downscaling and bias correction (model output statistics) method based on cumulative distribution functions (CDF) transformation. See Michelangeli, Vrac, Loukos (2009) Probabilistic downscaling approaches: Application to wind cumulative distribution functions. Geophysical Research Letters, 36, L11708, <doi:10.1029/2009GL038401>. ; and Vrac, Drobinski, Merlo, Herrmann, Lavaysse, Li, Somot (2012) Dynamical and statistical downscaling of the French Mediterranean climate: uncertainty assessment. Nat. Hazards Earth Syst. Sci., 12, 2769-2784, www.nat-hazards-earth-syst-sci.net/12/2769/2012/, <doi:10.5194/nhess-12-2769-2012>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2021-02-18 12:30:06 UTC

R topics documented:

CDFt *Downscaling or bias correction of CDF through CDF-transformation*

Description

Downscales (or corrects the model outputs) cumulative distribution function (CDF) of a climate variable from large- to local-scale by applying a equivalent of proportionality transformation : i.e., based on a CDF representing a variable at large scale (i.e., low spatial resolution) and the equivalent CDF at a local scale (e.g., modeled at a weather station), this method finds a mathematical transformation allowing to go from the large- to the local-scale CDF. Hence, when a new large-scale CDF is given, a new local-scale CDF is downscaled based on this transformation.

Usage

```
CDFt(ObsRp, DataGp, DataGf, npas=1000, dev=2)
```

Arguments

`ObsRp`	Observed time series of the variable (e.g., temperature) at the local scale to be used for estimation of the calibration local-scale CDF.
`DataGp`	Large-scale time series to be used for estimation of the calibration large-scale CDF.
`DataGf`	Large-scale time series to be used for estimation of the large-scale CDF to be downscaled.
`npas`	Number of "cuts" for which quantiles will be empirically estimated (Default is 1000).
`dev`	Coefficient of development (of the difference between the mean of the large-scale historical data and the mean of the large-scale future data to be downscaled). This development is used to extend range of data on which the quantiles will be calculated for the CDF to be downscaled (Default is 2).

Details

For details on the mathematical formulation of the transformation used to translate the large-scale CDF to the local-scale one, see the reference below. Note that in this R package, the large-scale data (i.e., DataGp and DataGf) are automatically transformed to have the same mean as the ObsRp time series. This avoid to get out of the range of applicability of the CDFt method. However, the large-scale output CDFs have their initial mean (i.e., not centered).

P.-A. Michelangeli, M. Vrac, H. Loukos. "Probabilistic downscaling approaches: Application to wind cumulative distribution functions", Geophys. Res. Lett., doi:10.1029/2009GL038401, 2009.

Value

A message is returned if the "dev" parameter is too small to capture the whole range of the downscaled CDF. Otherwise, CDFt returns a list with components

DS	Downscaled time series generated by "Quantile-matching" method performed between large-scale CDF to be downscaled, and the local-scale downscaled CDF. Note that the length of this array is equal to the length of DataGf
x	an array containing values of the variable (e.g., temperature) where the downscaled (and other) CDF has been estimated.
FRp	an array containing the values of the local-scale CDF used for calibration, evaluated at the points in x.
FGp	an array containing the values of the large-scale CDF used for calibration, evaluated at the points in x.
FGf	an array containing the values of the large-scale CDF used for downscalingn, evaluated at the points in x.
FRf	an array containing the values of the downscaled CDF evaluated at the points in x.

Author(s)

M. Vrac (mathieu.vrac@lsce.ipsl.fr) and P.-A. Michelangeli (pam@climpact.com)

See Also

CramerVonMisesTwoSamples,KolmogorovSmirnov

Examples

```
## Example

### Generation of example data
O <- rnorm(2100,mean=0,sd=1)
Gp <- rnorm(300,mean=3,sd=1)
Gf <- rnorm(300,mean=4,sd=1)

### Call of the CDFt method
CT <- CDFt(O,Gp,Gf)

x <- CT$x
FGp <- CT$FGp
FGf <- CT$FGf
FRp <- CT$FRp
FRf <- CT$FRf
ds <- CT$DS

### Plot the results
par(mfrow=c(1,2))

plot(x, FGp,type="l",lty=2,ylim=c(0,1),xlab="x",ylab="CDF(x)")
lines(x,FGf,type="l",lty=2,col=2)
lines(x,FRp,type="l")
lines(x,FRf,type="l",col=2)

plot(Gf,ds,xlab="Large-scale data", ylab="Downscaled data")
```

CramerVonMisesTwoSamples
Computation of the two-sample Cramer-von Mises statistics

Description

This function computes the two-sample Cramer-von Mises statistics U.

Usage

```
CramerVonMisesTwoSamples(S1, S2)
```

Arguments

S1	Vector containing the sample 1 from which CDF1 will be estimated.
S2	Vector containing the sample 2 from which CDF2 will be estimated.

Details

CDF1 and CDF2 are estimated empirically to compute the two-sample Cramer-von Mises statistics.

Value

U: The value of the Cramer-von Mises statistics.

Author(s)

P.-A. Michelangeli (pam@climpact.com) and M. Vrac (mathieu.vrac@lsce.ipsl.fr)

References

T.W. Anderson "On the distribution of the Two-sample Cramer-von Mises criterion". The Annals of Mathematical Statistics, 33 (3), 1148-1159 (1962).

P.-A. Michelangeli, M. Vrac, H. Loukos. "Probabilistic downscaling approaches: Application to wind cumulative distribution functions", Geophys. Res. Lett., doi:10.1029/2009GL038401, 2009.

See Also

KolmogorovSmirnov,CDFt

Examples

```
# generate random values (for the example)
S1 = rnorm(100)
S2 = rnorm(100)
CramerVonMisesTwoSamples(S1, S2)
```

`KolmogorovSmirnov` *Computation of the Kolmogorov-Smirnov statistics*

Description

This function computes the Kolmogorov-Smirnov statistics (KS).

Usage

```
KolmogorovSmirnov(S1, S2)
```

Arguments

`S1`	Vector containing the sample 1 from which CDF1 will be estimated.
`S2`	Vector containing the sample 2 from which CDF2 will be estimated.

Details

CDF1 and CDF2 are estimated empirically to compute the Kolmogorov-Smirnov statistics.

Value

Returns the value of the Kolmogorov-Smirnov statistics.

Author(s)

P.-A. Michelangeli (pam@climpact.com) and M. Vrac (mathieu.vrac@lsce.ipsl.fr)

References

D.A. Darling. "The Kolmogorov-Smirnov, Cramer-von Mises tests", Ann. Math. Statist., 28 (4), 823-838 (1957).

P.-A. Michelangeli, M. Vrac, H. Loukos. "Probabilistic downscaling approaches: Application to wind cumulative distribution functions", Geophys. Res. Lett., doi:10.1029/2009GL038401, 2009.

See Also

`CramerVonMisesTwoSamples`,`CDFt`

Examples

```
# generate random values (for the example)
S1 = rnorm(100)
S2 = rnorm(100)
KolmogorovSmirnov(S1, S2)
```

Index