Package ‘MBC’

February 5, 2018

Version 0.10-4
Title Multivariate Bias Correction of Climate Model Outputs
Author Alex J. Cannon
Maintainer Alex J. Cannon <alex.cannon@canada.ca>
Description Calibrate and apply multivariate bias correction algorithms
for climate model simulations of multiple climate variables. Three methods
described by Cannon (2016) <doi:10.1175/JCLI-D-15-0679.1> and
Cannon (2018) <doi:10.1007/s00382-017-3580-6> are implemented:
(i) MBC Pearson correlation (MBCp), (ii) MBC rank correlation (MBCr),
and (iii) MBC N-dimensional PDF transform (MBCn).
License GPL-2
Depends R (>= 2.10), Matrix, energy
Encoding UTF-8
NeedsCompilation no
Repository CRAN
Date/Publication 2018-02-05 18:21:30 UTC

R topics documented:

MBC-package ............................................................. 2
cccma ................................................................. 6
escore ................................................................. 7
MBCn ................................................................. 7
MBCp ................................................................. 9
MBCr ................................................................. 10
MRS ................................................................. 12
QDM ................................................................. 13
rot.random ......................................................... 14

Index 15
Description

Calibrate and apply multivariate bias correction algorithms for climate model simulations of multiple climate variables. Three methods are supported: (i) MBC Pearson correlation (MBCp), (ii) MBC rank correlation (MBCr), and (iii) MBC N-dimensional probability density function transform (MBCn). The first two, MBCp and MBCr (Cannon, 2016), match marginal distributions and inter-variable dependence structure. Dependence structure can be measured either by the Pearson correlation (\texttt{mbcp}) or by the Spearman rank correlation (\texttt{mbcr}). The energy distance score (\texttt{escore}) is recommended for model selection. The third, MBCn (Cannon, 2018), which operates on the full multivariate distribution, is more flexible and can be considered to be a multivariate analogue of univariate quantile mapping. All aspects of the observed distribution are transferred to the climate model simulations. In each of the three methods, marginal distributions are corrected by the change-preserving quantile delta mapping (\texttt{qdm}) algorithm (Cannon et al., 2015).

An example application of the three methods using the \texttt{cccma} dataset can be run via:

\texttt{example(MBC, run.dontrun=TRUE)}

Note: because empirical quantiles and their changes are used by \texttt{QDM}, sample sizes of the observed, model calibration, and model projection datasets should be approximately equal.

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>MBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>License:</td>
<td>GPL-2</td>
</tr>
<tr>
<td>LazyLoad:</td>
<td>yes</td>
</tr>
</tbody>
</table>

References


See Also

QDM, MBCp, MBCr, MBCn, score, rot.random, cccma

Examples

```r
## Not run:
data(cccma)
set.seed(1)

# Univariate quantile mapping
qdm.c <- cccma$gcm.c
qdm.p <- cccma$gcm.p
for(i in seq(ncol(cccma$gcm.c))){
  fit.qdm <- QDM(o.c=cccma$rcm.c[,i], m.c=cccma$gcm.c[,i],
               m.p=cccma$gcm.p[i], ratio=cccma$ratio.seq[i],
               trace=cccma$trace[i])
  qdm.c[i] <- fit.qdm$h.ohat.c
  qdm.p[i] <- fit.qdm$h.ohat.p
}

# Multivariate MBCp bias correction
fit.mbcp <- MBCp(o.c=cccma$rcm.c, m.c=cccma$gcm.c,
                 m.p=cccma$gcm.p, ratio.seq=cccma$ratio.seq,
                 trace=cccma$trace)
mbcp.c <- fit.mbcp$h.ohat.c
mbcp.p <- fit.mbcp$h.ohat.p

# Multivariate MBCr bias correction
fit.mbcr <- MBCr(o.c=cccma$rcm.c, m.c=cccma$gcm.c,
                 m.p=cccma$gcm.p, ratio.seq=cccma$ratio.seq,
                 trace=cccma$trace)
mbcr.c <- fit.mbcr$h.ohat.c
mbcr.p <- fit.mbcr$h.ohat.p

# Multivariate MBCn bias correction
fit.mbcn <- MBCn(o.c=cccma$rcm.c, m.c=cccma$gcm.c,
                 m.p=cccma$gcm.p, ratio.seq=cccma$ratio.seq,
                 trace=cccma$trace)
mbcn.c <- fit.mbcn$h.ohat.c
mbcn.p <- fit.mbcn$h.ohat.p
colnames(mbcn.c) <- colnames(mbcn.p) <-
             colnames(cccma$rcm.c)

# Correlation matrices (Pearson and Spearman)
# MBCp
dev.new()
par(mfrow=c(2,2))
plot(c(cor(cccma$rcm.c)), c(cor(qdm.c)), col='black',
     pch=19, xlab='CanRCM4', ylab='CanESM2 MBCp',
     main='Pearson correlation\nMBCp calibration')
abline(0, 1)
```
grid()
points(c(cor(cccma$rcm.c), c(cor(mbcp.c)), col='red')
plot(c(cor(cccma$rcm.p), c(cor(qdm.p)),
    col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcP',
    main='Pearson correlation\nMBcP evaluation')
abline(0, 1)
grid()
points(c(cor(cccma$rcm.p), c(cor(mbcp.p)), col='red')
plot(c(cor(cccma$rcm.c, m='s'), c(cor(qdm.c, m='s'))),
    col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcP',
    main='Spearman correlation\nMBcP calibration')
abline(0, 1)
grid()
points(c(cor(cccma$rcm.c, m='s')), c(cor(mbcp.c, m='s'))),
    col='red')

# MBcR
dev.new()
par(mfrow=c(2, 2))
plot(c(cor(cccma$rcm.c), c(cor(qdm.c)), col='black',
    pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcR',
    main='Pearson correlation\nMBcR calibration')
abline(0, 1)
grid()
points(c(cor(cccma$rcm.c), c(cor(mbc.p)), col='blue')
plot(c(cor(cccma$rcm.p), c(cor(qdm.p)),
    col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcR',
    main='Pearson correlation\nMBcR evaluation')
abline(0, 1)
grid()
points(c(cor(cccma$rcm.p), c(cor(mbc.p)), col='blue')
plot(c(cor(cccma$rcm.c, m='s')), c(cor(qdm.c, m='s'))),
    col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcR',
    main='Spearman correlation\nMBcR calibration')
abline(0, 1)
grid()
points(c(cor(cccma$rcm.c, m='s')), c(cor(mbc.p, m='s'))),
    col='blue')
plot(c(cor(cccma$rcm.p, m='s')), c(cor(qdm.p, m='s'))),
    col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
    xlab='CanRCM4', ylab='CanESM2 MBcR',
    main='Spearman correlation\nMBcR calibration')
abline(0, 1)
grid()
MBC-package

```r
xlab='CanRCM4', ylab='CanESM2 MBCr',
main='Spearman correlation\nMBCr evaluation'
abline(0, 1)
grid()
points(c(cor(cccma$rcm.p, m='s')), c(cor(mbcr.p, m='s')),
col='blue')

# MBCn
dev.new()
par(mfrow=c(2, 2))
plot(c(cor(cccma$rcm.c)), c(cor(qdm.c)), col='black',
pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
xlab='CanRCM4', ylab='CanESM2 MBCn',
main='Pearson correlation\nMBCn calibration'
abline(0, 1)
grid()
points(c(cor(cccma$rcm.c)), c(cor(mbcn.c)), col='orange')
plot(c(cor(cccma$rcm.p)), c(cor(qdm.p)),
col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
xlab='CanRCM4', ylab='CanESM2 MBCn',
main='Pearson correlation\nMBCn evaluation'
abline(0, 1)
grid()
points(c(cor(cccma$rcm.p)), c(cor(mbcn.p)), col='orange')
plot(c(cor(cccma$rcm.c, m='s')), c(cor(qdm.c, m='s')),
col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
xlab='CanRCM4', ylab='CanESM2 MBCn',
main='Spearman correlation\nMBCn calibration'
abline(0, 1)
grid()
points(c(cor(cccma$rcm.c, m='s')), c(cor(mbcn.c, m='s')),
col='orange')
plot(c(cor(cccma$rcm.p, m='s')), c(cor(qdm.p, m='s')),
col='black', pch=19, xlim=c(-1, 1), ylim=c(-1, 1),
xlab='CanRCM4', ylab='CanESM2 MBCn',
main='Spearman correlation\nMBCn evaluation'
abline(0, 1)
grid()
points(c(cor(cccma$rcm.p, m='s')), c(cor(mbcn.p, m='s')),
col='orange')

# Pairwise scatterplots
dev.new()
pairs(cccm$gcm.c, main='CanESM2 calibration', col='#0000001A')
dev.new()
pairs(cccm$rcm.c, main='CanRCM4 calibration', col='#0000001A')
dev.new()
pairs(qdm.c, main='QDM calibration', col='#0000001A')
dev.new()
pairs(mbcp.c, main='MBCp calibration', col='#FF00001A')
dev.new()
pairs(mbcr.c, main='MBCr calibration', col='#0000FF1A')
dev.new()
```
Sample CanESM2 and CanRCM4 data

**Description**

Sample CanESM2 (T63 grid) and CanRCM4 (0.22-deg grid) data (122.5 deg W, 50 deg N).

- pr: precipitation (mm day⁻¹)
- tas: average surface temperature (deg. C)
- dtr: diurnal temperature range (deg. C)
- sfcWind: surface wind speed (m s⁻¹)
- ps: surface pressure (ps)
- huss: surface specific humidity (kg kg⁻¹)
- rsds: surface downwelling shortwave radiation (W m⁻²)
- rlds: surface downwelling longwave radiation (W m⁻²)

**Value**

a list of with elements consisting of:

- `gcm.c`: matrix of CanESM2 variables for the calibration period.
- `gcm.p`: matrix of CanESM2 variables for the validation period.
- `rcm.c`: matrix of CanRCM4 variables for the calibration period.
- `rcm.p`: matrix of CanRCM4 variables for the validation period.
- `ratio.seq`: vector of logical values indicating if samples are of a ratio quantity.
- `trace`: numeric values indicating trace thresholds for each ratio quantity.
escore

**Energy distance score**

**Description**

Calculate the energy distance score measuring the statistical discrepancy between samples \( x \) and \( y \) from two multivariate distributions.

**Usage**

```r
escore(x, y, scale.x=FALSE, n.cases=NULL, alpha=1, method='cluster')
```

**Arguments**

- **x** numeric matrix.
- **y** numeric matrix.
- **scale.x** logical indicating whether data should be standardized based on \( x \).
- **n.cases** the number of sub-sampled cases; \( \text{NULL} \) uses all data.
- **alpha** distance exponent in \((0,2]\)
- **method** method used to weight the statistics

**References**


mbcn

**Multivariate bias correction (N-pdft)**

**Description**

Multivariate bias correction that matches the multivariate distribution using QDM and the N-dimensional probability density function transform (N-pdft) following Cannon (2018).

**Usage**

```r
mbcn(o.c, m.c, m.p, iter=30, ratio.seq=rep(FALSE, ncol(o.c)),
     trace=0.05, trace.calc=0.5*trace, jitter.factor=0,
     n.tau=NULL, ratio.max=2, ratio.max.trace=10*trace,
     ties='first', qmap.precalc=FALSE, rot.seq=NULL,
     silent=FALSE, n.escore=0, return.all=FALSE, subsample=NULL,
     pp.type=7)
```
Arguments

- `o.c` matrix of observed samples during the calibration period.
- `m.c` matrix of model outputs during the calibration period.
- `m.p` matrix of model outputs during the projected period.
- `iter` maximum number of algorithm iterations.
- `ratio.seq` vector of logical values indicating if samples are of a ratio quantity (e.g., precipitation).
- `trace` numeric values indicating thresholds below which values of a ratio quantity (e.g., `ratio=TRUE`) should be considered exact zeros.
- `trace.calc` numeric values of thresholds used internally when handling of exact zeros; defaults to one half of `trace`.
- `jitter.factor` optional strength of jittering to be applied when quantities are quantized.
- `n.tau` number of quantiles used in the quantile mapping; `NULL` equals the length of the `m.p` series.
- `ratio.max` numeric values indicating the maximum proportional changes allowed for ratio quantities below the `ratio.max.trace` threshold.
- `ratio.max.trace` numeric values of trace thresholds used to constrain the proportional change in ratio quantities to `ratio.max`; defaults to ten times `trace`.
- `ties` method used to handle ties when calculating ordinal ranks.
- `qmap.precalc` logical value indicating if `m.c` and `m.p` are outputs from QDM.
- `rot.seq` use a supplied list of random rotation matrices. `NULL` generates on the fly.
- `silent` logical value indicating if algorithm progress should be reported.
- `n.escore` number of cases used to calculate the energy distance when monitoring convergence.
- `return.all` logical value indicating whether results from all iterations are returned.
- `subsample` use `subsample` draws of size `n.tau` to calculate initial empirical quantiles; if `NULL`, calculate normally.
- `pp.type` type of plotting position used in quantile.

Value

- a list of with elements consisting of:
  - `mhat.c` matrix of bias corrected `m.c` values for the calibration period.
  - `mhat.p` matrix of bias corrected `m.p` values for the projection period.

References


See Also

QDM, MBcp, MBcr, MRS, escore, rot.random

**MBcp**

Multivariate bias correction (Pearson correlation)

**Description**

Multivariate bias correction that matches marginal distributions using QDM and the Pearson correlation dependence structure following Cannon (2016).

**Usage**

```r
MBcp(o.c, m.c, m.p, iter=20, cor.thresh=1e-4, ratio.seq=rep(FALSE, ncol(o.c)), trace=0.05, trace.calc=0.5*trace, jitter.factor=0, n.tau=NULL, ratio.max=2, ratio.max.trace=10*trace, ties='first', qmap.precalc=FALSE, silent=FALSE, subsample=NULL, pp.type=W)
```

**Arguments**

- `o.c`: matrix of observed samples during the calibration period.
- `m.c`: matrix of model outputs during the calibration period.
- `m.p`: matrix of model outputs during the projected period.
- `iter`: maximum number of algorithm iterations.
- `cor.thresh`: if greater than zero, a threshold indicating the change in magnitude of Pearson correlations required for convergence.
- `ratio.seq`: vector of logical values indicating if samples are of a ratio quantity (e.g., precipitation).
- `trace`: numeric values indicating thresholds below which values of a ratio quantity (e.g., ratio=TRUE) should be considered exact zeros.
- `trace.calc`: numeric values of thresholds used internally when handling of exact zeros; defaults to one half of `trace`.
<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>jitter.factor</td>
<td>optional strength of jittering to be applied when quantities are quantized.</td>
</tr>
<tr>
<td>n.tau</td>
<td>number of quantiles used in the quantile mapping; NULL equals the length of the m.p series.</td>
</tr>
<tr>
<td>ratio.max</td>
<td>numeric values indicating the maximum proportional changes allowed for ratio quantities below the ratio.max.trace threshold.</td>
</tr>
<tr>
<td>ratio.max.trace</td>
<td>numeric values of trace thresholds used to constrain the proportional change in ratio quantities to ratio.max; defaults to ten times trace.</td>
</tr>
<tr>
<td>ties</td>
<td>method used to handle ties when calculating ordinal ranks.</td>
</tr>
<tr>
<td>qmap.precalc</td>
<td>logical value indicating if m.c and m.p are outputs from QDM.</td>
</tr>
<tr>
<td>silent</td>
<td>logical value indicating if algorithm progress should be reported.</td>
</tr>
<tr>
<td>subsample</td>
<td>use subsample draws of size n.tau to calculate initial empirical quantiles; if NULL, calculate normally.</td>
</tr>
<tr>
<td>pp.type</td>
<td>type of plotting position used in quantile.</td>
</tr>
</tbody>
</table>

**Value**

a list of with elements consisting of:

- mhat.c: matrix of bias corrected m.c values for the calibration period.
- mhat.p: matrix of bias corrected m.p values for the projection period.

**References**


**See Also**

QDM, MBCr, MRS, MBCn escore

---

**Description**

Multivariate bias correction that matches marginal distributions using QDM and the Spearman rank correlation dependence structure following Cannon (2016).
Usage

MBCr(o.c, m.c, m.p, iter=20, cor.thresh=1e-4,
ratio.seq=rep(FALSE, ncol(o.c)), trace=0.05,
trace.calc=0.5*trace, jitter.factor=0, n.tau=NULL,
ratio.max=2, ratio.max.trace=10*trace, ties='first',
qmap.precalc=FALSE, silent=FALSE, subsample=NULL,
pp.type=W)

Arguments

- **o.c**: matrix of observed samples during the calibration period.
- **m.c**: matrix of model outputs during the calibration period.
- **m.p**: matrix of model outputs during the projected period.
- **iter**: maximum number of algorithm iterations.
- **cor.thresh**: if greater than zero, a threshold indicating the change in magnitude of Spearman rank correlations required for convergence.
- **ratio.seq**: vector of logical values indicating if samples are of a ratio quantity (e.g., precipitation).
- **trace**: numeric values indicating thresholds below which values of a ratio quantity (e.g., ratio\_true) should be considered exact zeros.
- **trace.calc**: numeric values of thresholds used internally when handling of exact zeros; defaults to one half of trace.
- **jitter.factor**: optional strength of jittering to be applied when quantities are quantized.
- **n.tau**: number of quantiles used in the quantile mapping; NULL equals the length of the m.p series.
- **ratio.max**: numeric values indicating the maximum proportional changes allowed for ratio quantities below the ratio.max.trace threshold.
- **ratio.max.trace**: numeric values of trace thresholds used to constrain the proportional change in ratio quantities to ratio.max; defaults to ten times trace.
- **ties**: method used to handle ties when calculating ordinal ranks.
- **qmap.precalc**: logical value indicating if m.c and m.p are outputs from QDM.
- **silent**: logical value indicating if algorithm progress should be reported.
- **subsample**: use subsample draws of size n.tau to calculate empirical quantiles; if NULL, calculate normally.
- **pp.type**: type of plotting position used in quantile.

Value

a list of with elements consisting of:

- **mhat.c**: matrix of bias corrected m.c values for the calibration period.
- **mhat.p**: matrix of bias corrected m.p values for the projection period.
References


See Also

QDM, MBCp, MRS, MBCn escore

---

**MRS**

*Multivariate linear rescaling using Cholesky decomposition*

**Description**

Multivariate linear bias correction based on Cholesky decomposition of the covariance matrix following Scheuer and Stoller (1962) and Bürger et al. (2011). Bias correction matches the multivariate mean and covariance structure.

**Usage**

```r
MRS(o.c, m.c, m.p, o.c.chol=NULL, o.p.chol=NULL, m.c.chol=NULL, m.p.chol=NULL)
```

**Arguments**

- `o.c`: matrix of observed samples during the calibration period.
- `m.c`: matrix of model outputs during the calibration period.
- `m.p`: matrix of model outputs during the projected period.
- `o.c.chol`: precalculated Cholesky decomposition of the `o.c` covariance matrix; `NULL` calculates internally.
- `o.p.chol`: precalculated Cholesky decomposition of the target `o.p` covariance matrix; `NULL` defaults to `o.c.chol`.
- `m.c.chol`: precalculated Cholesky decomposition of the `m.c` covariance matrix; `NULL` calculates internally.
- `m.p.chol`: precalculated Cholesky decomposition of the `m.p` covariance matrix; `NULL` calculates internally.

**Value**

A list of with elements consisting of:

- `mhat.c`: matrix of bias corrected `m.c` values for the calibration period.
- `mhat.p`: matrix of bias corrected `m.p` values for the projection period.
References


See Also

MBCp, MBCr

qdm

Univariate bias correction via quantile delta mapping

Description

Univariate bias correction based on the quantile delta mapping QDM version of the quantile mapping algorithm from Cannon et al. (2015). QDM constrains model-projected changes in quantiles to be preserved following bias correction by quantile mapping.

Usage

qdm(oNc, mNc, mNp, ratio=False, trace=0.05, trace.calc=0.5*trace,
    jitter.factor=0, n.tau=NULL, ratio.max=2,
    ratio.max.trace=10*trace, ECBC=FALSE, ties='first',
    subsample=NULL, pp.type=7)

Arguments

- o.Nc: vector of observed samples during the calibration period.
- m.Nc: vector of model outputs during the calibration period.
- m.Np: vector of model outputs during the projected period.
- ratio: logical value indicating if samples are of a ratio quantity (e.g., precipitation).
- trace: numeric value indicating the threshold below which values of a ratio quantity (e.g., ratio=TRUE) should be considered exact zeros.
- trace.calc: numeric value of a threshold used internally when handling of exact zeros; defaults to one half of trace.
- jitter.factor: optional strength of jittering to be applied when quantities are quantized.
- n.tau: number of quantiles used in the quantile mapping; NULL equals the length of the m.p series.
- ratio.max: numeric value indicating the maximum proportional change allowed for ratio quantities below the ratio.max.trace threshold.
- ratio.max.trace: numeric value of a trace threshold used to constrain the proportional change in ratio quantities to ratio.max; defaults to ten times trace.
rot.random

ECBC logical value indicating whether mhat.p outputs should be ordered according to o.c ranks, i.e., as in the empirical copula-bias correction (ECBC) algorithm.
ties method used to handle ties when calculating ordinal ranks.
subsample use subsample draws of size n.tau to calculate empirical quantiles; if NULL, calculate normally.
pp.type type of plotting position used in quantile.

Value

a list of with elements consisting of:

mhat.c vector of bias corrected m.c values for the calibration period.
mhat.p vector of bias corrected m.p values for the projection period.

References


See Also

MBCp, MBCr, MRS, escore

Description

Generate a k-dimensional random orthogonal rotation matrix.

Usage

rot.random(k)

Arguments

k the number of dimensions.

References

Index

*Topic package
  MBC-package, 2

cccma, 3, 6

escore, 2, 3, 7, 9, 10, 12, 14

MBC (MBC-package), 2
MBC-package, 2
MBCn, 2, 3, 7, 10, 12
MBCp, 2, 3, 9, 12–14
MBCr, 2, 3, 9, 10, 10, 13, 14
MRS, 9, 10, 12, 12, 14

QDM, 2, 3, 7, 9, 10, 12, 13

rot.random, 3, 9, 14