

Package ‘pc’

October 10, 2026

Title Pattern Causality Analysis

Version 0.3

Description Infer causation from observational data through pattern causality analysis (PC), with original algorithm for time series data from Stavroglou et al. (2020) <[doi:10.1073/pnas.1918269117](https://doi.org/10.1073/pnas.1918269117)>, as well as methodological extensions for spatial cross-sectional data introduced by Zhang & Wang (2025) <[doi:10.1080/13658816.2025.2581207](https://doi.org/10.1080/13658816.2025.2581207)>, together with a systematic description proposed in Lyu et al. (2026) <[doi:10.1016/j.compenvurbsys.2026.102435](https://doi.org/10.1016/j.compenvurbsys.2026.102435)>.

License GPL-3

Encoding UTF-8

URL <https://stscl.github.io/pc/>, <https://github.com/stscl/pc>

BugReports <https://github.com/stscl/pc/issues>

Depends R (>= 4.1.0)

LinkingTo Rcpp, RcppThread

Imports ggplot2, methods, sdsfun, sf, stats, terra

Suggests Rcpp, RcppThread, readr, infoxtr, spEDM, tEDM, knitr, rmarkdown

VignetteBuilder knitr

Config/roxygen2/markdown TRUE

Config/roxygen2/version 8.1.0

NeedsCompilation yes

Author Wenbo Lyu [aut, cre, cph] (ORCID:
<<https://orcid.org/0009-0002-6003-3800>>)

Maintainer Wenbo Lyu <lyu.geosocial@gmail.com>

Repository CRAN

Date/Publication 2026-10-10 05:30:02 UTC

Contents

dmi	2
fnn	3
ops	5
pc	8
Index	11

dmi	<i>Delayed Mutual Information</i>
-----	-----------------------------------

Description

Delayed Mutual Information

Usage

```
## S4 method for signature 'data.frame'
dmi(
  data,
  target,
  tau = 1:10,
  pred = NULL,
  k = 3,
  base = exp(1),
  normalize = FALSE,
  threads = length(tau),
  ...
)
```

Arguments

- | | |
|-----------|--|
| data | Observation data. |
| target | Integer of column indice for the target variable. |
| tau | (optional) Step of lag. |
| pred | (optional) Predictions indices. |
| k | (optional) Number of nearest neighbors used for evaluation. |
| base | (optional) Logarithm base of the entropy. |
| normalize | (optional) Whether to normalize MI values. |
| threads | (optional) Number of threads used. |
| ... | Additional arguments to absorb unused inputs in method dispatch. |

Value

A vector.

Note

dmi is only supported for time-series inputs.

References

Fraser, A.M., Swinney, H.L., 1986. Independent coordinates for strange attractors from mutual information. *Physical Review A* 33, 1134–1140.

Kraskov, A., Stogbauer, H., Grassberger, P., 2004. Estimating mutual information. *Physical Review E* 69, 066138.

Examples

```
abun = readr::read_csv(system.file("case/abundance.csv", package = "pc"))
pc::dmi(abun, 2)
```

fnn	<i>False Nearest Neighbors</i>
-----	--------------------------------

Description

False Nearest Neighbors

Usage

```
## S4 method for signature 'data.frame'
fnn(
  data,
  target,
  E = 10,
  k = 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  dist.metric = c("euclidean", "manhattan", "maximum"),
  rt = 10,
  eps = NULL,
  threads = length(E),
  higher.parallel = TRUE,
  ...
)

## S4 method for signature 'sf'
fnn(
  data,
  target,
```

```

    E = 10,
    k = 1,
    tau = 1,
    style = 1,
    lib = NULL,
    pred = NULL,
    dist.metric = c("euclidean", "manhattan", "maximum"),
    rt = 10,
    eps = NULL,
    threads = length(E),
    higher.parallel = TRUE,
    detrend = FALSE,
    nb = NULL,
    ...
)

## S4 method for signature 'SpatRaster'
fnn(
  data,
  target,
  E = 10,
  k = 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  dist.metric = c("euclidean", "manhattan", "maximum"),
  rt = 10,
  eps = NULL,
  threads = length(E),
  higher.parallel = TRUE,
  detrend = FALSE,
  ...
)

```

Arguments

<code>data</code>	Observation data.
<code>target</code>	Integer of column indice for the target variable.
<code>E</code>	(optional) Embedding dimensions.
<code>k</code>	(optional) Number of nearest neighbors used for evaluation.
<code>tau</code>	(optional) Step of lag.
<code>style</code>	(optional) Embedding style (0 includes current state, 1 excludes it).
<code>lib</code>	(optional) Libraries indices.
<code>pred</code>	(optional) Predictions indices.
<code>dist.metric</code>	(optional) Distance measure to be used.

rt (optional) Relative distance threshold.
 eps (optional) Absolute distance threshold.
 threads (optional) Number of threads used.
 higher.parallel (optional) Whether to use a higher level of parallelism.
 ... Additional arguments to absorb unused inputs in method dispatch.
 detrend (optional) Whether to remove the linear trend.
 nb (optional) Neighbours list.

Value

A vector.

References

Kennel, M.B., Brown, R., Abarbanel, H.D.I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A* 45, 3403–3411.

Examples

```
crash = sf::read_sf(system.file("case/crash.gpkg", package = "pc"))
pc::fnn(crash, 1, threads = 1)
```

 ops

Optimal Parameter Search for Pattern Causality

Description

Optimal Parameter Search for Pattern Causality

Usage

```
## S4 method for signature 'data.frame'
ops(
  data,
  source,
  target,
  E = 3:5,
  k = E,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  maximize = c("dark", "positive", "negative"),
  dist.metric = c("euclidean", "manhattan", "maximum"),
```

```

    zero.tolerance = max(k),
    relative = TRUE,
    weighted = TRUE,
    threads = length(E),
    higher.parallel = TRUE,
    h = 0,
    ...
)

## S4 method for signature 'sf'
ops(
  data,
  source,
  target,
  E = 3:5,
  k = E + 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  maximize = c("dark", "positive", "negative"),
  dist.metric = c("euclidean", "manhattan", "maximum"),
  zero.tolerance = max(k),
  relative = TRUE,
  weighted = TRUE,
  threads = length(E),
  higher.parallel = TRUE,
  detrend = FALSE,
  nb = NULL,
  ...
)

## S4 method for signature 'SpatRaster'
ops(
  data,
  source,
  target,
  E = 3:5,
  k = E + 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  maximize = c("positive", "negative", "dark"),
  dist.metric = c("euclidean", "manhattan", "maximum"),
  zero.tolerance = max(k),
  relative = TRUE,
  weighted = TRUE,

```

```

    threads = length(E),
    higher.parallel = TRUE,
    detrend = FALSE,
    ...
  )

```

Arguments

<code>data</code>	Observation data.
<code>source</code>	Integer of column indice for the source variable.
<code>target</code>	Integer of column indice for the target variable.
<code>E</code>	(optional) Embedding dimensions.
<code>k</code>	(optional) Number of nearest neighbors used for projection.
<code>tau</code>	(optional) Step of lag.
<code>style</code>	(optional) Embedding style (0 includes current state, 1 excludes it).
<code>lib</code>	(optional) Libraries indices.
<code>pred</code>	(optional) Predictions indices.
<code>maximize</code>	(optional) Causality metric to maximize: one of "positive", "negative", or "dark".
<code>dist.metric</code>	(optional) Distance measure to be used.
<code>zero.tolerance</code>	(optional) Maximum number of zeros tolerated in signature space.
<code>relative</code>	(optional) Whether to calculate relative changes in embedding.
<code>weighted</code>	(optional) Whether to weight causal strength.
<code>threads</code>	(optional) Number of threads used.
<code>higher.parallel</code>	(optional) Whether to use a higher level of parallelism.
<code>h</code>	(optional) Prediction horizon.
<code>...</code>	Additional arguments to absorb unused inputs in method dispatch.
<code>detrend</code>	(optional) Whether to remove the linear trend.
<code>nb</code>	(optional) Neighbours list.

Value

A list.

`param` The selected optimal parameter combination.

`causality` A data.frame containing results across parameter settings.

Examples

```

crash = sf::read_sf(system.file("case/crash.gpkg", package = "pc"))
pc::ops(crash, 2, 1, E = 3:10, maximize = "positive", threads = 1)

```

pc

Pattern Causality

Description

Pattern Causality

Usage

```
## S4 method for signature 'data.frame'
pc(
  data,
  source,
  target,
  libsizes = NULL,
  E = 3,
  k = E,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  boot = 99,
  replace = FALSE,
  seed = 42L,
  dist.metric = c("euclidean", "manhattan", "maximum"),
  zero.tolerance = max(k),
  relative = TRUE,
  weighted = TRUE,
  threads = length(libsizes),
  higher.parallel = TRUE,
  verbose = TRUE,
  h = 0,
  ...
)

## S4 method for signature 'sf'
pc(
  data,
  source,
  target,
  libsizes = NULL,
  E = 3,
  k = E + 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
```

```

boot = 99,
replace = FALSE,
seed = 42L,
dist.metric = c("euclidean", "manhattan", "maximum"),
zero.tolerance = max(k),
relative = TRUE,
weighted = TRUE,
threads = length(libsizes),
higher.parallel = TRUE,
verbose = TRUE,
detrend = FALSE,
nb = NULL,
...
)

## S4 method for signature 'SpatRaster'
pc(
  data,
  source,
  target,
  libsizes = NULL,
  E = 3,
  k = E + 1,
  tau = 1,
  style = 1,
  lib = NULL,
  pred = NULL,
  boot = 99,
  replace = FALSE,
  seed = 42L,
  dist.metric = c("euclidean", "manhattan", "maximum"),
  zero.tolerance = max(k),
  relative = TRUE,
  weighted = TRUE,
  threads = length(libsizes),
  higher.parallel = TRUE,
  verbose = TRUE,
  detrend = FALSE,
  ...
)

```

Arguments

data	Observation data.
source	Integer of column indice for the source variable.
target	Integer of column indice for the target variable.
libsizes	(optional) Number of observations used.
E	(optional) Embedding dimensions.

<code>k</code>	(optional) Number of nearest neighbors used for projection.
<code>tau</code>	(optional) Step of lag.
<code>style</code>	(optional) Embedding style (0 includes current state, 1 excludes it).
<code>lib</code>	(optional) Libraries indices.
<code>pred</code>	(optional) Predictions indices.
<code>boot</code>	(optional) Number of bootstraps to perform.
<code>replace</code>	(optional) Should sampling be with replacement?
<code>seed</code>	(optional) Random seed.
<code>dist.metric</code>	(optional) Distance measure to be used.
<code>zero.tolerance</code>	(optional) Maximum number of zeros tolerated in signature space.
<code>relative</code>	(optional) Whether to calculate relative changes in embedding.
<code>weighted</code>	(optional) Whether to weight causal strength.
<code>threads</code>	(optional) Number of threads used.
<code>higher.parallel</code>	(optional) Whether to use a higher level of parallelism.
<code>verbose</code>	(optional) Whether to show the progress bar.
<code>h</code>	(optional) Prediction horizon.
<code>...</code>	Additional arguments to absorb unused inputs in method dispatch.
<code>detrend</code>	(optional) Whether to remove the linear trend.
<code>nb</code>	(optional) Neighbours list.

Value

A list.

causality A data.frame of causality results. When `libsizes` is `NULL`, it contains per-sample causality estimates; otherwise, it contains causality results evaluated across different library sizes.

summary A data.frame summarizing overall causality metrics. Only returned when `libsizes` is `NULL`.

References

Stavroglou, S.K., Pantelous, A.A., Stanley, H.E., Zuev, K.M., 2020. Unveiling causal interactions in complex systems. *Proceedings of the National Academy of Sciences* 117, 7599–7605.

Examples

```
crash = sf::read_sf(system.file("case/crash.gpkg", package = "pc"))
p1 = pc::pc(crash, 2, 1, E = 3, k = 7, threads = 1)
print(p1)
plot(p1)

# convergence diagnostics
p2 = pc::pc(crash, 2, 1, libsizes = seq(10,172,40), E = 3, k = 7, threads = 1)
print(p2)
plot(p2)
```

Index

`dmi`, [2](#)
`dmi`, `data.frame`-method (`dmi`), [2](#)

`fnn`, [3](#)
`fnn`, `data.frame`-method (`fnn`), [3](#)
`fnn`, `sf`-method (`fnn`), [3](#)
`fnn`, `SpatRaster`-method (`fnn`), [3](#)

`ops`, [5](#)
`ops`, `data.frame`-method (`ops`), [5](#)
`ops`, `sf`-method (`ops`), [5](#)
`ops`, `SpatRaster`-method (`ops`), [5](#)

`pc`, [8](#)
`pc`, `data.frame`-method (`pc`), [8](#)
`pc`, `sf`-method (`pc`), [8](#)
`pc`, `SpatRaster`-method (`pc`), [8](#)