Package ‘rflexscan’

March 1, 2023

Type Package
Title The Flexible Spatial Scan Statistic
Version 1.1.0
Date 2023-03-01
Author Takahiro Otani, Kunihiko Takahashi
Maintainer Takahiro Otani <otani@med.nagoya-cu.ac.jp>
Description Functions for the detection of spatial clusters using the flexible spatial scan statistic developed by Tango and Takahashi (2005) <doi:10.1186/1476-072X-4-11>. This package implements a wrapper for the C routine used in the FleXScan 3.1.2 <https://sites.google.com/site/flexscansoftware/home> developed by Takahashi, Yokoyama, and Tango. For details, see Otani et al. (2021) <doi:10.18637/jss.v099.i13>.

URL https://tkhrotn.github.io/rflexscan/
License GPL-3
Depends R (>= 3.1.0)
Imports Rcpp, igraph, sf, grDevices
LinkingTo Rcpp
Suggests knitr, rmarkdown, spdep, spData
RoxygenNote 7.2.3
Encoding UTF-8
NeedsCompilation yes
Repository CRAN
Date/Publication 2023-03-01 08:52:30 UTC

R topics documented:

rflexscan-package ................................................. 2
choropleth ....................................................... 3
plot.rflexscan ................................................... 4
print.rflexscan .................................................. 6
rflexscan-package

Analyze spatial count data using the flexible spatial scan statistic

Description

The rflexscan package provides functions and classes to analyze spatial count data using the flexible spatial scan statistic developed by Tango and Takahashi (2005). This package designed for any of the following interrelated purposes:

1. To evaluate reported spatial disease clusters, to see if they are statistically significant.
2. To test whether a disease is randomly distributed over space.
3. To perform geographical surveillance of disease, to detect areas of significantly high rates.

This package implements a wrapper for the C routine used in the FleXScan 3.1.2 developed by Takahashi, Yokoyama, and Tango.

References


See Also

rflexscan
**Description**

Display choropleth map of detected clusters.

**Usage**

```r
choropleth(
  polygons, 
  fls, 
  col = palette(), 
  region_color = "#F0F0F0", 
  rank = 1:length(fls$cluster), 
  pval = 1, 
  ...
)
```

**Arguments**

- `polygons` A SpatialPolygonsDataFrame.
- `fls` An rflexscan object.
- `col` A vector of colors for each cluster.
- `region_color` Color of regions that are not included in any clusters.
- `rank` An integer vector which specifies ranks of clusters to be displayed.
- `pval` A threshold of P-value. Clusters with P-values of <pval will be displayed.
- `...` Other parameters to be passed to plot function.

**Details**

Clusters are colored using the current palette. Please use `palette` function to specify colors of each cluster. Note that clusters with ranks larger than the number of colors in the palette are not highlighted.

**See Also**

- `rflexscan`

**Examples**

```r
# load sample data (North Carolina SIDS data)
library(sf)
library(spdep)
data("nc.sids")
```
sids.shp <- read_sf(system.file("shapes/sids.shp", package="spData")[1])

# calculate the expected numbers of cases
expected <- nc.sids$BIR74 * sum(nc.sids$SID74) / sum(nc.sids$BIR74)

# run FleXScan
fls <- rflexscan(x = nc.sids$x, y = nc.sids$y,
                 observed = nc.sids$SID74,
                 expected = expected,
                 name = rownames(nc.sids),
                 clustersize = 10,
                 nb = ncCR85.nb)

# display all clusters
choropleth(sids.shp, fls)

# display clusters with rank 1, 2 and 3
choropleth(sids.shp, fls, rank = c(1, 2, 3))

# display clusters of P-value <= 0.05
choropleth(sids.shp, fls, pval = 0.05)

---

**plot.rflexscan**

*Graph plotting of flexscan results*

**Description**

Display detected clusters by a graph representation.

**Usage**

```r
## S3 method for class 'rflexscan'
plot(
x,
     rank = 1:length(x$cluster),
     pval = 1,
     vertexsize = max(x$input$coordinates[, 1]) - min(x$input$coordinates[, 1]),
     xlab = colnames(x$input$coordinates)[1],
     ylab = colnames(x$input$coordinates)[2],
     xlim = c(min(x$input$coordinates[, 1]), max(x$input$coordinates[, 1])),
     ylim = c(min(x$input$coordinates[, 2]), max(x$input$coordinates[, 2])),
     col = palette(),
     frame_color = "gray40",
     vertex_color = "white",
     ...
    )
```
Arguments

- **x**: An `rflexscan` object.
- **rank**: An integer vector which specifies ranks of clusters to be displayed.
- **pval**: A threshold of P-value. Clusters with P-values of < `pval` will be displayed.
- **vertexsize**: Size of vertex of the graph.
- **xlab**: A label of the x axis.
- **ylab**: A label of the y axis.
- **xlim**: The x limits of the plot.
- **ylim**: The y limits of the plot.
- **col**: A vector of colors for each cluster.
- **frame_color**: Color of frames in the graph.
- **vertex_color**: Fill color of vertices that are not included in any clusters.
- **...**: Other parameters to be passed to `plot.igraph` function.

Details

Clusters are colored using the current palette. Please use `palette` function to specify colors of each cluster. Note that clusters with ranks larger than the number of colors in the palette are not highlighted.

See Also

- `rflexscan`

Examples

```r
# load sample data (North Carolina SIDS data)
library(spdep)
data("nc.sids")

# calculate the expected numbers of cases
expected <- nc.sids$BIR74 * sum(nc.sids$SID74) / sum(nc.sids$BIR74)

# run FleXScan
fls <- rflexscan(x = nc.sids$x, y = nc.sids$y,
                  observed = nc.sids$SID74,
                  expected = expected,
                  name = rownames(nc.sids),
                  clustersize = 10,
                  nb = ncCR85.nb)

# display all clusters
plot(fls)

# display clusters with rank 1, 2 and 3
plot(fls, rank = c(1, 2, 3))
```
print.rflexscanCluster

# display clusters of P-value <= 0.05
plot(fls, pval = 0.05)

print.rflexscan

Print rflexscan object

Description
Print method for the rflexscan object.

Usage

## S3 method for class 'rflexscan'
print(x, ...)

Arguments

x An rflexscan object to be printed.
...
Ignored.

See Also
rflexscan

print.rflexscanCluster

Print rflexscanCluster object

Description
Print method for the rflexscanCluster object.

Usage

## S3 method for class 'rflexscanCluster'
print(x, ...)

Arguments

x An rflexscanCluster object to be printed.
...
Ignored.
print.summary.rflexscan

Print summary of flexscan results

Description
Print summary of flexscan results to the terminal.

Usage
## S3 method for class 'summary.rflexscan'
print(x, ...)

Arguments
x An summary.rflexscan object to be printed.
... Ignored.

See Also
rflexscan, summary.rflexscan

rflexscan
Detect spatial disease clusters using the flexible/circular scan statistic

Description
This function analyzes spatial count data using the flexible spatial scan statistic developed by Tango and Takahashi (2005) or Kulldorff’s circular spatial scan statistic (1997), and detect spatial disease clusters.

Usage
rflexscan(
  x,
  y,
  lat,
  lon,
  name,
  observed,
  expected,
  population,
  nb,
  clustersize = 15,
  radius = 6370,
Arguments

- **x**: A vector of X-coordinates.
- **y**: A vector of Y-coordinates.
- **lat** (DEPRECATED): A vector of latitude.
- **lon** (DEPRECATED): A vector of longitude.
- **name**: A vector of names of each area.
- **observed**: A vector with the observed number of disease cases.
- **expected**: A vector with the expected number of disease cases under the null hypothesis. This is used on "Poisson" model.
- **population**: A vector with the background population at risk in each area. This is used on "Binomial" model.
- **nb**: A neighbors list or an adjacency matrix.
- **clustersize**: The number of maximum spatial cluster size to scan, i.e., the maximum number of regions included in the detected cluster.
- **radius**: Radius of Earth to calculate a distance between two sets of latitude and longitude. It is approximately 6370 km in Japan. This parameter is used when lat and lon are specified. This is DEPRECATED. The distance calculated using this parameter is not accurate. This feature is implemented to maintain compatibility with FleXScan. It is recommended to transform latitude and longitude onto the Cartesian coordinate system beforehand and use the x and y parameters that are projected coordinates.
- **stattype**: Statistic type to be used (case-insensitive).
  - "ORIGINAL": the likelihood ratio statistic by Kulldorff and Nagarwalla (1995)
  - "RESTRICTED": the restricted likelihood ratio statistic by Tango (2008), with a preset parameter ralpha for restriction
- **scanmethod**: Scanning method to be used (case-insensitive).
  - "FLEXIBLE": flexible scan statistic by Tango and Takahashi (2005)
  - "CIRCULAR": circular scan statistic by Kulldorff (1997)
- **ralpha**: Threshold parameter of the middle p-value for the restricted likelihood ratio statistic.
- **simcount**: The number of Monte Carlo replications to calculate a p-value for statistical test.
raitype

The type of random number for Monte Carlo simulation (case-insensitive).
"MULTINOMIAL" Total number of cases in whole area is fixed. It can be chosen in either Poisson or Binomial model.
"POISSON" Total number of cases is not fixed. It can be chosen in Poisson model.

comments

Comments for the analysis which will be written in summary.

verbose

Print progress messages.

secondary

The number of secondary clusters to be enumerated. If NULL is specified (default), the search for secondary clusters is stopped when the Monte Carlo p-value reaches 1.

clustertype

Type of cluster to be scanned.
"HOT" Hot-spot clusters with elevated risk.
"COLD" Cold-spot clusters with reduced risk.
"BOTH" Hot- and cold-spot clusters simultaneously.

Details

Centroid coordinates for each region should be specified EITHER by Cartesian coordinates using arguments x and y or by latitudes and longitudes using arguments lat and lon. Note that lat and lon are DEPRECATED due to accuracy issues. This feature is implemented to maintain compatibility with FleXScan software. We recommend to transform latitude and longitude onto the Cartesian coordinate system beforehand (using spTransform function in sp package, for example) and use the x and y parameters that are projected coordinates.

Value

An rflexscan object which contains analysis results and specified parameters.

References


See Also

summary.rflexscan, plot.rflexscan, choropleth
Examples

# load sample data (North Carolina SIDS data)
library(spdep)
data("nc.sids")

# calculate the expected numbers of cases
expected <- nc.sids$BIR74 * sum(nc.sids$SID74) / sum(nc.sids$BIR74)

# run FleXScan
fls <- rflexscan(x = nc.sids$x, y = nc.sids$y,
  observed = nc.sids$SID74,
  expected = expected,
  name = rownames(nc.sids),
  clustersize = 10,
  nb = ncCR85.nb)

# print rflexscan object
print(fls)

# print properties of the most likely cluster
print(fls$cluster[[1]])

# print summary to the terminal
summary(fls)

# plot graph
plot(fls, col = palette())
labs <- 1:length(fls$cluster)
legend("bottomleft", legend = labs, col = palette(), lty = 1)

---

runFleXScan  

Run main routine of FleXScan.

Description

Run main routine of FleXScan.

Usage

runFleXScan(setting, case_mat, coord_mat, adj_mat)

Arguments

setting  
A list of parameter setting.

case_mat  
A matrix of case counts.

coord_mat  
A matrix of coordinates.

adj_mat  
A matrix of neighbourhood relationships.
**summary.rflexscan**

*Summary method for rflexscan objects.*

**Usage**

```r
## S3 method for class 'rflexscan'
summary(object, ...)
```

**Arguments**

- `object` An rflexscan object to be summarized.
- `...` Ignored.

**See Also**

- `rflexscan`
Index

choropleth, 3, 9
palette, 3, 5
plot.igraph, 5
plot.rflexscan, 4, 9
print.rflexscan, 6
print.rflexscanCluster, 6
print.summary.rflexscan, 7

rflexscan, 2, 3, 5–7, 7, 11
rflexscan-package, 2
runFleXScan, 10

summary.rflexscan, 7, 9, 11