Package 'sandwichr'

January 9, 2023

Version 1.0.3

Date 2023-1-8

Title Spatial Interpolation Based on Spatial Stratified Heterogeneity

Author Yue Lin [aut, cre],
Chengdong Xu [aut, ctb],
Jinfeng Wang [aut, ctb]

Maintainer Yue Lin <lin.3326@osu.edu>

Depends R (>= 3.5.0)

Imports sf, ggplot2, gridExtra, geodetector, caret, lwgeom, tools, dplyr

Suggests rmarkdown, knitr, MASS, ggpubr, ape

Description Spatial interpolation is a common practice in social and environmental science. This package enables the implementation of SSH-based spatial interpolation proposed by Wang et al. (2013) <doi:10.1068/a44710>. It provides functions to (1) evaluate stratification schemes, (2) interpolate sampling data over user-defined reporting units, (3) assess interpolation uncertainties, and (4) evaluate overall accuracy using the k-fold cross-validation estimate.

License GPL (>= 2)

URL https://github.com/linyuehzzz/sandwich_spatial_interpolator/tree/master/r/sandwichr

BugReports https://github.com/linyuehzzz/sandwich_spatial_interpolator/issues

LazyData false

RoxygenNote 7.2.2

VignetteBuilder knitr

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2023-01-09 08:10:05 UTC

R topics documented:

`autoplot` *Plot interpolation results*

Description

Plot the estimated mean and standard error for each reporting unit.

Usage

```
## S3 method for class 'sandwich.model'
autoplot(
  object,
  titles = c("Mean", "Standard Error"),
  labels = c("", ""),
  border_color = "darkgray",
  poly_fill_mean = c("white", "red"),
  poly_fill_se = c("white", "blue"),
  ...
)
```

Arguments

`object`	A `sandwich.model` object generated by `sandwich.model`.
`titles`	A list of texts for the titles.
`labels`	A list of texts for the legend labels.
`border_color`	Color for outlines of the polygons.
`poly_fill_mean`	A list of colors for low and high ends of the gradient in the mapping of means.
`poly_fill_se`	A list of colors for low and high ends of the gradient in the mapping of standard errors.
`...`	Ignored.

Value

A list of ggplot objects.

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534.

See Also

sandwich.model

Examples

```
library(ggplot2)
data(sim.data)
sim.sw <- sandwich.model(object=sim.data, sampling.attr="Value", type="shp")
ggplot2::autoplot(object=sim.sw)
```

---

bc.data *Breast cancer incidence in mainland China*

Description

The bc.data dataset consists of two data frames:

- bc.data[[1]]: Breast cancer incidence at 242 sampling units in mainland China, where the SSH stratum (SSHID) and reporting unit (GBCODE) that each sample falls into are specified. The sampling attribute is Incidence.
- bc.data[[2]]: The county-level administrative divisions in mainland China (GBCODE), where the weights of each intersecting stratum (W1 and W2) are specified.

Examples

```
data(bc.data)
```

load.data.shp *Load shapefiles into* sf*s*

Description

Convert shapefiles into a list of sf objects.

Usage

```
load.data.shp(sampling.file,
       ssh.file,
       reporting.file)
```

Arguments

sampling.file	Path to the sampling layer. The sampling layer should be a point sf object that contains an attribute to be interpolated.
ssh.file	Path to the SSH layer. The SSH layer should be a polygon sf object.
reporting.file	Path to the reporting layer. The reporting layer should be a polygon sf object.

Value

A list of sf objects.

Examples

```
sim.sampling.name <- system.file("extdata", "sim.sampling.shp",
                                 package="sandwichr")
sim.ssh.name <- system.file("extdata", "sim.ssh.shp",
                            package="sandwichr")
sim.reporting.name <- system.file("extdata", "sim.reporting.shp",
                                  package="sandwichr")

sim.data <- load.data.shp(sampling.file=sim.sampling.name,
                          ssh.file=sim.ssh.name,
                          reporting.file=sim.reporting.name)
```

load.data.txt *Load text files into data frames*

Description

Convert text files into a list of data frames.

Usage

```
load.data.txt(sampling_ssh.file,
       reporting_ssh.file)
```

Arguments

`sampling_ssh.file`
: Path to the file linking sampling and SSH layers. Each row denotes a sampling unit. At least two columns need to be included in this file: value of the sampling attribute and the stratum each sampling unit falling into.

`reporting_ssh.file`
: Path to the file linking reporting and SSH layers. Each row denotes a reporting unit. The weights of each stratum should be specified in the columns.

Value

A list of data frames.

Examples

```
bc.sampling_ssh.name <- system.file("extdata", "bc_sampling_ssh.csv",
                                    package="sandwichr")
bc.reporting_ssh.name <- system.file("extdata", "bc_reporting_ssh.csv",
                                     package="sandwichr")

bc.data <- load.data.txt(sampling_ssh.file=bc.sampling_ssh.name,
                         reporting_ssh.file=bc.reporting_ssh.name)
```

`sandwich.cv` *Perform* k*-fold cross validation*

Description

Perform *k*-fold cross validation to evaluate the overall model accuracy and output the average root mean square error (RMSE).

Usage

```
sandwich.cv(object,
       sampling.attr,
       k=10,
       type="shp",
       ssh.id.col=NULL,
       reporting.id.col=NULL,
       ssh.weights=NULL)
```

Arguments

object	When type="shp", object is a list of three sf objects generated by load.data.shp, including a point sf object used as the sampling layer, a polygon sf object used as the SSH layer, and a polygon sf object used as the SSH layer. When type="txt", object is a list of two data frames generated by load.data.txt, including a file linking sampling and SSH layers and a file linking reporting and SSH layers.
sampling.attr	Text for the name of the attribute to be interpolated in the sampling layer.
k	Number of folds (k > 1). By default, k = 10.
type	Text for the type of input data. type="shp" denotes shapefiles, and type="txt" denotes text files. By default, type="shp".
ssh.id.col	Text for the column that specifies which stratum each sampling unit falls into (see load.data.txt). Set to NULL when type="shp".
reporting.id.col	Text for the column that specifies which reporting unit each sampling unit falls into (see load.data.txt). Set to NULL when type="shp".
ssh.weights	A list that specifies the strata in the SSH layer and their corresponding columns of weights in reporting_ssh.file (see load.data.txt).

Value

A value of the *k*-fold cross validation estimate.

See Also

load.data.shp, load.data.txt

---

sandwich.model *Perform Sandwich model-based mapping*

Description

Estimate the mean and standard error for each reporting unit using SSH-based spatial interpolation.

Usage

```
sandwich.model(object,
       sampling.attr,
       type="shp",
       ssh.id.col=NULL,
       ssh.weights=NULL)
```

Arguments

object	When type="shp", object is a list of three sf objects generated by load.data.shp, including a point sf object used as the sampling layer, a polygon sf object used as the SSH layer, and a polygon sf object used as the SSH layer. When type="txt", object is a list of two data frames generated by load.data.txt, including a file linking sampling and SSH layers and a file linking reporting and SSH layers.
sampling.attr	Text for the name of the attribute to be interpolated in the sampling layer.
type	Text for the type of input data. type="shp" denotes shapefiles, and type="txt" denotes text files. By default, type="shp".
ssh.id.col	Text for the column that specifies which stratum each sampling unit falls into (see load.data.txt). Set to NULL when type="shp".
ssh.weights	A list that specifies the strata in the SSH layer and their corresponding columns of weights in reporting_ssh.file (see load.data.txt).

Value

A sandwich.ci object that contains the estimated mean and standard deviation for each reporting unit.

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534.

See Also

load.data.shp, load.data.txt

Examples

```
data(sim.data)
sim.sw <- sandwich.model(object=sim.data, sampling.attr="Value", type="shp")
```

---

sandwichr — sandwichr*: Spatial Interpolation Based on Spatial Stratified Heterogeneity*

Description

The package sandwichr provides tools to interpolate a spatially stratified heterogeneous population with high accuracy, even when its spatial autocorrelation is weak. These tools allow the calculation of critical summarized statistics such as geodetector q. They also enable the visualization of the interpolated surface as well as the standard error and confidence interval maps.

Author(s)

Yue Lin, Chengdong Xu, and Jinfeng Wang

References

Wang, J. F., Haining, R., Liu, T. J., Li, L. F., & Jiang, C. S. (2013). Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10), 2515-2534.

`sim.data` *Simulated data*

Description

The `sim.data` dataset consists of three `sf` objects:

- `sim.data[[1]]`: 41 sampling units. The sampling attribute is `Value`.
- `sim.data[[2]]`: A zonation map with 4 strata that serves as a candidate SSH layer.
- `sim.data[[3]]`: 7 reporting units.

Examples

```
data(sim.data)
```

`ssh.data.shp` *Prepare shapefile data for Geodetector*

Description

Join a candidate SSH layer to an existing layer for the geographical detector model.

Usage

```
ssh.data.shp(object,
       ssh.lyr,
       ssh.id,
       ssh.name=NULL)
```

Arguments

`object`	A point `sf` object to be joint. Its geometry should be consistent with the sampling layer, but it may have additional attributes from the SSH layer(s).
`ssh.lyr`	A polygon `sf` object used as the candidate SSH layer. Its stratum ID will be linked to `object`.
`ssh.id`	Text for the field name of stratum ID in the SSH layer.
`ssh.name`	Text for the new field name assigned to the stratum ID in the output object. If `NULL`, the attribute name will be the same as that in the old SSH layer.

Value

An sf object with the sampling units and the Strata they fall into.

Examples

```
library(sf)
library(tools)
data(sim.data)
sim.join <- ssh.data.shp(object=sim.data[[1]], ssh.lyr=sim.data[[2]], ssh.id="X")
head(sim.join)
```

---

ssh.data.txt *Prepare text data for Geodetector*

Description

Prepare the input from text files for the geographical detector model.

Usage

```
ssh.data.txt(object)
```

Arguments

object	A list of two data frames generated by load.data.txt, including a file linking sampling and SSH layers linking reporting and SSH layers and a file.

Value

A list of two data frames.

Examples

```
data(bc.data)
bc.join <- ssh.data.txt(object=bc.data)
```

`ssh.test` *Evaluate stratification*

Description

Calculate the factor detector q-statistic and the interaction detector in the geographical detector model described by Wang et al. (2010). The q-statistic measures the SSH of the sampling attribute in terms of a given stratification, which can be used for the selection of an SSH layer for Sandwich model-based mapping. The interactive effects indicate whether a combination of two stratifications enhances the SSH of the sampling attribute.

Usage

```
ssh.test(object,
       y,
       x,
       test="factor",
       type="shp")
```

Arguments

`object`	An object generated by `ssh.data.shp` or `ssh.data.txt`.
`y`	Text for the name of the explained variable (sampling attribute) in `object`.
`x`	Text for the name(s) of the explanatory variable(s) (stratification(s)) in `object`.
`test`	Text for the type of test. `test="factor"` denotes the factor detector, and `test="interaction"` denotes the interaction detector. By default, `test="factor"`.
`type`	Text for the type of input data. `type="shp"` denotes shapefiles, and `type="txt"` denotes text files. By default, `type="shp"`.

Value

A value of the q-statistic or the combined q-statistic.

References

Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X., & Zheng, X. Y. (2010). Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*, 24(1), 107-127.

See Also

`ssh.data.shp`, `ssh.data.txt`

Examples

```
library(sf)
library(tools)
data(sim.data)
sim.join <- ssh.data.shp(object=sim.data[[1]], ssh.lyr=sim.data[[2]], ssh.id="X")
head(sim.join)
ssh.test(object=sim.join, y="Value", x=c("X"), test="factor")
```

---

summary *Summarize interpolation results*

Description

Summarize the estimated mean and standard error.

Usage

```
## S3 method for class 'sandwich.model'
summary(object, ...)
```

Arguments

object	A sandwich.model object generated by sandwich.model.
...	Ignored.

Value

Summarized statistics for the estimated mean and standard error.

See Also

sandwich.model

Examples

```
data(sim.data)
sim.sw <- sandwich.model(object=sim.data, sampling.attr="Value", type="shp")
summary(object=sim.sw)
```

Index