

# Package ‘sandwichr’

February 8, 2022

**Version** 1.0.1

**Date** 2022-01-25

**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

**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](https://github.com/linyuehzzz/sandwich_spatial_interpolator/tree/master/r/sandwichr)

**BugReports** [https://github.com/linyuehzzz/sandwich\\_spatial\\_interpolator/issues](https://github.com/linyuehzzz/sandwich_spatial_interpolator/issues)

**LazyData** false

**RoxygenNote** 7.1.2

**VignetteBuilder** knitr

**Encoding** UTF-8

## R topics documented:

autoplot . . . . .	2
bc.data . . . . .	3
hs.data . . . . .	4
load.data.shp . . . . .	4
load.data.txt . . . . .	5
sandwich.ci . . . . .	6
sandwich.cv . . . . .	6
sandwich.model . . . . .	7

sandwichr . . . . .	8
ssh.data.shp . . . . .	9
ssh.data.txt . . . . .	9
ssh.test . . . . .	10
summary . . . . .	11

<b>Index</b>	<b>13</b>
--------------	-----------

---

autoplot	<i>Plot interpolation results</i>
----------	-----------------------------------

---

**Description**

Plot the estimated mean, standard error, and confidence interval 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"),
  ...
)

## S3 method for class 'sandwich.ci'
autoplot(
  object,
  titles = c("Confidence Interval (Lower Bound)", "Confidence Interval (Upper Bound)"),
  labels = c("", ""),
  border_color = "darkgray",
  poly_fill_ci = c("white", "red"),
  ...
)
```

**Arguments**

object	A sandwich.ci object generated by <a href="#">sandwich.ci</a> , or a sandwich.model object generated by <a href="#">sandwich.model</a> .
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.
poly_fill_ci	A list of colors for low and high ends of the gradient in the mapping of confidence intervals.

**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](#), [sandwich.ci](#)

**Examples**

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

hs.sw.ci <- sandwich.ci(object=hs.sw, level=.95)
ggplot2::autoplot(object=hs.sw.ci)
```

---

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)
```

---

hs.data	<i>Heshun human population</i>
---------	--------------------------------

---

### Description

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

- `hs.data[[1]]`: Human population of 167 sampling villages within the Heshun region, China, which serves as the sampling layer. The sampling attribute is `Population`.
- `hs.data[[2]]`: A zonation map with 5 strata that serves as a candidate SSH layer.
- `hs.data[[3]]`: The administrative division of 10 townships in Heshun county that served as the reporting layer.

### References

Li, L., Wang, J., & Wu, J. (2012). A spatial model to predict the incidence of neural tube defects. *BMC Public Health*, 12(1), 951.

### Examples

```
data(hs.data)
```

---

load.data.shp	<i>Load shapefiles into sf</i>
---------------	--------------------------------

---

### Description

Convert shapefiles into a list of `sf` objects.

### Usage

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

### Arguments

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

### Value

A list of `sf` objects.

**Examples**

```

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

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

```

---

load.data.txt	<i>Load text files into data frames</i>
---------------	---

---

**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.ci	<i>Calculate confidence interval for the mean</i>
-------------	---

---

### Description

Calculate the confidence interval for the mean of each reporting unit.

### Usage

```
sandwich.ci(object,
             level=.95)
```

### Arguments

object	A sandwich.model object generated by <a href="#">sandwich.model</a> .
level	The confidence level required. By default, level = .95.

### Value

A sandwich.ci object that contains the lower and upper bounds of confidence intervals.

### See Also

[autoplot](#), [sandwich.model](#)

### Examples

```
data(hs.data)
hs.sw <- sandwich.model(object=hs.data, sampling.attr="Population", type="shp")
hs.sw.ci <- sandwich.ci(object=hs.sw, level=.95)
```

---

sandwich.cv	<i>Perform k-fold cross validation</i>
-------------	--

---

### 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 <a href="#">load.data.shp</a> , 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 <a href="#">load.data.txt</a> , 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 <a href="#">load.data.txt</a> ). Set to NULL when type="shp".
reporting.id.col	Text for the column that specifies which reporting unit each sampling unit falls into (see <a href="#">load.data.txt</a> ). 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 <a href="#">load.data.txt</a> ).

**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 <a href="#">load.data.shp</a> , 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 <a href="#">load.data.txt</a> , 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 <a href="#">load.data.txt</a> ). 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 <a href="#">load.data.txt</a> ).

### 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(hs.data)
hs.sw <- sandwich.model(object=hs.data, sampling.attr="Population", 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.



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(hs.data)
hs.ssh2.name <- system.file("extdata", "hs.ssh2.shapefile.shp",
                           package="sandwichr")
hs.ssh2 <- read_sf(dsn=dirname(hs.ssh2.name),
                  layer=file_path_sans_ext(basename(hs.ssh2.name)))
hs.join <- ssh.data.shp(object=hs.data[[1]], ssh.lyr=hs.data[[2]], ssh.id="STR_1")
hs.join <- ssh.data.shp(object=hs.join, ssh.lyr=hs.ssh2, ssh.id="STR_2")
```

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(hs.data)
hs.ssh2.name <- system.file("extdata", "hs.ssh2.shapefile.shp",
                           package="sandwichr")
hs.ssh2 <- read_sf(dsn=dirname(hs.ssh2.name),
                  layer=file_path_sans_ext(basename(hs.ssh2.name)))
hs.join <- ssh.data.shp(object=hs.data[[1]], ssh.lyr=hs.data[[2]], ssh.id="STR_1")
hs.join <- ssh.data.shp(object=hs.join, ssh.lyr=hs.ssh2, ssh.id="STR_2")
ssh.test(object=hs.join, y="Population", x=c("STR_1", "STR_2"), test="factor")
```

---

summary

*Summarize interpolation results*

---

## Description

Summarize the estimated mean, standard error, and confidence interval.

## Usage

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

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

## Arguments

object	A <code>sandwich.ci</code> object generated by <a href="#">sandwich.ci</a> , or a <code>sandwich.model</code> object generated by <a href="#">sandwich.model</a> .
...	Ignored.

## Value

Summarized statistics for the estimated mean, standard error, and confidence interval.

## See Also

[sandwich.model](#), [sandwich.ci](#)

**Examples**

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

```
hs.sw.ci <- sandwich.ci(object=hs.sw, level=.95)
summary(object=hs.sw.ci)
```

# Index

autoplot, [2](#), [6](#)

bc.data, [3](#)

hs.data, [4](#)

load.data.shp, [4](#), [7](#), [8](#)

load.data.txt, [5](#), [7](#), [8](#), [10](#)

sandwich.ci, [2](#), [3](#), [6](#), [11](#)

sandwich.cv, [6](#)

sandwich.model, [2](#), [3](#), [6](#), [7](#), [11](#)

sandwichr, [8](#)

ssh.data.shp, [9](#), [10](#), [11](#)

ssh.data.txt, [9](#), [10](#), [11](#)

ssh.test, [10](#)

summary, [11](#)