

Package: geoaddSAE2 (via r-universe)

October 5, 2026

Type Package

Title Geoadditive Small Area Estimation for Area-Level Model

Version 0.1.1

Description Fits area-level geoadditive small area estimation (SAE) models by extending the Fay-Herriot area-level model with linear, nonlinear, and spatial effects. The Fay-Herriot model is described by Fay and Herriot (1979) [<doi:10.1080/01621459.1979.10482505>](https://doi.org/10.1080/01621459.1979.10482505). Geoadditive models combine nonlinear covariate effects and spatial variation as described by Kammann and Wand (2003) [<doi:10.1111/1467-9876.00385>](https://doi.org/10.1111/1467-9876.00385), while their application to small area estimation is discussed by Puspongoro et al. (2019) [<doi:10.21108/JDSA.2019.2.15>](https://doi.org/10.21108/JDSA.2019.2.15). Nonlinear covariate effects are represented using penalized splines, while spatial effects are represented using a smooth function of geographic coordinates. Models are estimated using restricted maximum likelihood (REML), and mean squared error (MSE) is estimated using a parametric bootstrap. The package also provides comparisons with the Fay-Herriot and spatial Fay-Herriot (SFH) models.

License GPL (>= 3)

Encoding UTF-8

LazyData true

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

Depends R (>= 3.5.0)

Imports mgcv, sae, stats

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Amalyah Rizky Khairunnisa Sutarto [aut, cre], Novi Hidayat Puspongoro [aut, ths]

Maintainer Amalyah Rizky Khairunnisa Sutarto <222212489@stis.ac.id>

Config/pak/sysreqs cmake make
Repository <https://amlyhrzky.r-universe.dev>
Date/Publication 2026-10-05 07:30:02 UTC
RemoteUrl <https://github.com/cran/geoaddSAE2>
RemoteRef HEAD
RemoteSha bbd634ce97162e34a018e1166b1eb13276c96055

Contents

geosae	2
print.geosae	4
simulated_sae	4
summary.geosae	5
Index	7

geosae	<i>Fit an area-level Geoadditive Small Area Estimation model</i>
--------	--

Description

Fits area-level Geoadditive Small Area Estimation (SAE) models: a extension of the Fay-Herriot model that represents nonlinear covariate effects with and spatial variation, estimated by REML within a linear mixed model framework (via **mgcv**), with Mean Squared Error obtained by parametric bootstrap. The package also lets users compare the geoadditive model against the classical Fay-Herriot and Spatial Fay-Herriot models, both of which are provided by, and fitted with, the existing **sae** package.

Usage

```
geosae(  
  data,  
  formula,  
  vardir,  
  nonlinear,  
  spatial,  
  knots = NULL,  
  spatial_k = NULL,  
  method = c("REML"),  
  compare = FALSE,  
  proxmat = NULL,  
  proxmat_k = 4,  
  B = 100,  
  seed = NULL  
)
```

Arguments

data	A data frame containing the direct estimates, the known sampling variances, the covariates, and (if used) spatial coordinates.
formula	A model formula $y \sim x_1 + x_2$ giving the response (direct estimator) on the left-hand side and the <i>*linear*</i> covariates on the right-hand side. Use $y \sim 1$ if there are no linear covariates.
vardir	Unquoted name of the column in data holding the known sampling variances of the direct estimator (e.g. <code>vardir = vardir</code>).
nonlinear	Character vector of covariate names to be modelled nonlinearly with a P-spline.
spatial	Character vector of length 2 giving the names of the two spatial coordinate columns (e.g. <code>c("lat", "lon")</code>).
knots	Optional. Controls the basis dimension (number of knots) of each nonlinear P-spline term. NULL (default) uses an automatic rule of thumb following Ruppert (2002).
spatial_k	Optional basis dimension for the spatial thin-plate smooth. NULL uses mgcv 's default.
method	Smoothing-parameter selection method passed to <code>mgcv::gam</code> (and to the sae comparison models when <code>compare = TRUE</code>). Default "REML".
compare	Logical. If FALSE (default), only the geoadditive model is fitted. If TRUE, comparison models are fitted and a comparison table (<code>model mse rmse</code>) is returned.
proxmat	Optional spatial proximity matrix for the Spatial Fay-Herriot model. The matrix should have one row and one column for each area, with zero diagonal and row-standardized spatial weights. If NULL and <code>compare = TRUE</code> with <code>spatial</code> specified, the matrix is constructed automatically using k-nearest neighbours based on the spatial coordinates.
proxmat_k	Number of nearest neighbours used to construct the automatic spatial proximity matrix when <code>proxmat = NULL</code> . Default is 4.
B	Number of parametric bootstrap replicates. Default 100.
seed	Optional integer seed for the bootstrap, for reproducibility.

Details

Optionally (`compare = TRUE`), the classical Fay-Herriot and Spatial Fay-Herriot models are also fitted – using the existing, well-tested implementations in the **sae** package (`sae::mseFH`, `sae::eblupSFH`, `sae::mseSFH`) rather than re-implemented here – so that all three models can be compared side by side.

Value

An object of class "geosae", a list containing the call, settings, estimation results, diagnostics, fixed-effect parameters, model comparisons (if requested), and underlying fitted model objects.

print.geosae	<i>Print method for geosae object</i>
--------------	---------------------------------------

Description

Print method for geosae object

Usage

```
## S3 method for class 'geosae'
print(x, digits = 4, ...)
```

Arguments

x	An object of class geosae
digits	Number of decimal digits to print. Default 4.
...	Additional arguments passed to other methods.

Value

Invisibly returns the input object of class "geosae". The method prints the main model settings and area-level estimation results to the console.

simulated_sae	<i>Simulated area-level small area estimation data</i>
---------------	--

Description

A synthetic area-level small area estimation dataset used in package examples, demonstrations, and tests.

The data are generated using the following steps:

1. The population consists of 100 small areas, and a seed is set to ensure reproducibility of the simulated data.
2. Two-dimensional spatial coordinates are generated for each area, where latitude is generated from $U(-11, 6)$ and longitude is generated from $U(95, 141)$.
3. Two linear auxiliary variables are generated for each area: $x_1 \text{ Bernoulli}(0.9)$ and $x_2 U(1, 5)$. A nonlinear auxiliary variable $x_3 U(1, 5)$ is also generated.
4. The model parameters are set deterministically as $\alpha = 0.5$, $\beta_1 = 1$, and $\beta_2 = 1$.
5. The nonlinear covariate effect is generated using $g(x_3) = x_3^2$.
6. A continuous spatial effect is generated from the spatial coordinates using the smooth function $h(lat, lon) = \sin(lat) + \cos(lon)$.
7. The true small area parameter is calculated as $theta = \alpha + \beta_1 x_1 + \beta_2 x_2 + g(x_3) + h(lat, lon)$.

8. The number of sampled units in each area is generated from $n_i \sim U(10, 50)$. The sampling error variance is then defined as $var_{dir} = \sigma_e^2/n_i$, where $\sigma_e^2 = 1$.
9. The direct estimator is generated according to the sampling model $y = \theta + e$, where $e \sim N(0, var_{dir})$.
10. The area identifier, direct estimator, auxiliary variables, spatial coordinates, sampling variance, and true small area parameter are combined into a data frame called `simulated_sae`.

Usage

```
simulated_sae
```

Format

A data frame with 100 rows and 9 columns:

area Area identifier.

y Direct estimator (response).

x1 Binary auxiliary variable with a linear effect on the response.

x2 Continuous auxiliary variable with a linear effect on the response.

x3 Continuous auxiliary variable with a nonlinear effect on the response.

lat Latitude-like spatial coordinate.

lon Longitude-like spatial coordinate.

var_dir Known sampling variance of the direct estimator.

theta True small area parameter used to generate the direct estimator.

Examples

```
data(simulated_sae)
head(simulated_sae)
```

```
summary.geosae
```

Summary method for geosae object

Description

Summary method for geosae object

Usage

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

Arguments

<code>object</code>	An object of class <code>geosae</code>
<code>...</code>	Additional arguments passed to other methods.

Value

Invisibly returns the input object of class "geosae". The method prints model diagnostics, fixed-effect parameter estimates, smooth-term information, MSE and RMSE summaries, and model comparison results when available.

Index

* **datasets**

simulated_sae, [4](#)

geosae, [2](#)

print.geosae, [4](#)

simulated_sae, [4](#)

summary.geosae, [5](#)