Package: geoaddSAE2 (via r-universe)

October 15, 2026

Type Package

Title Geoadditive Small Area Estimation for Area-Level Model

Version 0.1.2

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>. Geoadditive models combine nonlinear covariate effects and spatial variation as described by Kammann and Wand (2003) <doi:10.1111/1467-9876.00385>, while their application to small area estimation is discussed by Pusponegoro et al. (2019) <doi: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, sae, spdep, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

Depends R (>= 3.5.0)

Imports mgcv, stats

Config/roxygen2/version 8.1.0

NeedsCompilation no

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

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

Repository https://amlyhrzky.r-universe.dev

Date/Publication 2026-10-15 19:20:02 UTC

RemoteUrl https://github.com/cran/geoaddSAE2

RemoteRef HEAD

RemoteSha 2dc1b8632fb6d986573bc53fc1606cd0c2b28820

Contents

geosae 2
print.geosae 4
simulated_sae 5
summary.geosae 6

Index **7**

`geosae` *Fit an Area-Level Geoadditive Small Area Estimation Model*

Description

Fits area-level Geoadditive Small Area Estimation (SAE) models, an extension of the Fay-Herriot model that incorporates linear covariate effects, nonlinear covariate effects represented by penalized splines, and spatial variation represented by a bivariate thin plate regression spline.

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, known sampling variances, covariates, and spatial coordinates.
`formula`	A model formula giving the response variable on the left-hand side and covariates to be modelled linearly on the right-hand side. Use `y ~ 1` if no linear covariates are used.
`vardir`	Unquoted name of the column containing the known sampling variances of the direct estimates.
`nonlinear`	Character vector containing the names of covariates to be modelled nonlinearly using P-splines.
`spatial`	Character vector of length two containing the names of the spatial coordinate variables, for example `c("lat", "lon")`.
`knots`	Optional basis dimension for the nonlinear P-spline terms. If `NULL`, the basis dimension is selected automatically.
`spatial_k`	Optional basis dimension for the spatial thin plate regression spline. If `NULL`, the default from **mgcv** is used.
`method`	Smoothing-parameter estimation method passed to `mgcv::gam`. Currently, `"REML"` is used by default.
`compare`	Logical. If `TRUE`, Fay-Herriot and Spatial Fay-Herriot comparison models are also fitted. These comparison models require the optional **sae** package.
`proxmat`	Optional spatial proximity matrix used by the Spatial Fay-Herriot comparison model.
`proxmat_k`	Number of nearest neighbours used when constructing the spatial proximity matrix automatically. Default is 4.
`B`	Number of parametric bootstrap replications used to estimate the MSE. Default is 100.
`seed`	Optional integer seed for reproducibility.

Details

Model parameters and smoothing parameters are estimated by Restricted Maximum Likelihood (REML) within a linear mixed model framework using **mgcv**. The Mean Squared Error (MSE) of the Geoadditive SAE estimates is estimated using parametric bootstrap.

When `compare = TRUE`, the function also fits the classical Fay-Herriot and Spatial Fay-Herriot models using implementations provided by the optional **sae** package. The **sae** package is not required when `compare = FALSE`.

Value

An object of class `"geosae"`. The returned list contains area-level estimates, model diagnostics, fixed-effect parameter estimates, model comparisons when requested, and the underlying fitted model objects.

Examples

```
data(simulated_sae)

fit <- geosae(
  data      = simulated_sae,
  formula   = y ~ x1 + x2,
  vardir    = vardir,
  nonlinear = "x3",
  spatial   = c("lat", "lon"),
  B         = 10,
  seed      = 1
)

fit
summary(fit)
head(fit$estimation)


## Not run:
fit_cmp <- geosae(
  data      = simulated_sae,
  formula   = y ~ x1 + x2,
  vardir    = vardir,
  nonlinear = "x3",
  spatial   = c("lat", "lon"),
  compare   = TRUE,
  B         = 100,
  seed      = 1
)

fit_cmp$comparison

## End(Not run)
```

`print.geosae` *Print method for geosae object*

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` *Simulated area-level small area estimation data*

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: $x1\ Bernoulli(0.5)$ and $x2\ U(1, 5)$. A nonlinear auxiliary variable $x3\ U(1, 5)$ 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 $g(x3) = x3^2$.
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 $theta = \alpha + \beta_1 x1 + \beta_2 x2 + g(x3) + h(lat, lon)$.
8. The number of sampled units in each area is generated from $n_i\ U10, 50$. The sampling error variance is then defined as $vardir = \sigma_e^2/n_i$, where $\sigma_e^2 = 1$.
9. The direct estimator is generated according to the sampling model $y = theta + e$, where $e\ N(0, vardir)$.
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.

vardir 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

`object`	An object of class `geosae`
`...`	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, 5

geosae, 2

print.geosae, 4

simulated_sae, 5
summary.geosae, 6