Package 'SpatialGEV'

June 9, 2024

Title Fit Spatial Generalized Extreme Value Models

Version 1.0.1

Maintainer Meixi Chen <meixi.chen@uwaterloo.ca>

Description
Fit latent variable models with the GEV distribution as the data likelihood and the GEV parameters following latent Gaussian processes. The models in this package are built using the template model builder 'TMB' in R, which has the fast ability to integrate out the latent variables using Laplace approximation. This package allows the users to choose in the fit function which GEV parameter(s) is considered as a spatially varying random effect following a Gaussian process, so the users can fit spatial GEV models with different complexities to their dataset without having to write the models in 'TMB' by themselves. This package also offers methods to sample from both fixed and random effects posteriors as well as the posterior predictive distributions at different spatial locations. Methods for fitting this class of models are described in Chen, Ramezan, and Lysy (2024) <doi:10.48550/arXiv.2110.07051>.

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.5.0)

Imports TMB (>= 1.7.16), mvtnorm, evd, stats, Matrix, methods

LinkingTo TMB, RcppEigen

RoxygenNote 7.2.3

Suggests INLA, testthat, knitr, rmarkdown

Additional_repositories https://inla.r-inla-download.org/R/stable/

VignetteBuilder knitr

NeedsCompilation yes

Author Meixi Chen [aut, cre],
Martin Lysy [aut],
Reza Ramezan [ctb]

Repository CRAN

Date/Publication 2024-06-09 06:20:11 UTC

Contents

`CAsnow` *Gridded monthly total snowfall in Canada from 1987 to 2021.*

Description

Variables containing the monthly total snowfall (in cm) in Canada from 1987 to 2021 and the location information. The data has been gridded and information about the grid size can be found in the paper Fast and Scalable Inference for Spatial Extreme Value Models (arxiv: 2110.07051).

Usage

```
CAsnow
```

Format

A list containing the location information and the observations:

locs A 509x2 matrix with longitude and latitude for each grid cell

n_loc Number of locations

Y A list of length 509 with each element of the list containing the observations at a location

Source

https://climate-change.canada.ca/climate-data/#/monthly-climate-summaries `grid_location` *Grid the locations with fixed cell size*

Description

Grid the locations with fixed cell size

Usage

```
grid_location(
  lon,
  lat,
  sp.resolution = 2,
  lon.range = range(lon),
  lat.range = range(lat)
)
```

Arguments

`lon`	Numeric, `n` longitude values
`lat`	Numeric, `n` latitude values
`sp.resolution`	Numeric, must be a single value that indicates the minimal unit length of a grid cell.
`lon.range`	Optional vector that indicates the range of `lon`. Default is `range(lon)`.
`lat.range`	Optional vector that indicates the range of `lat`. Default is `range(lat)`.

Details

The longitude and latitude of each grid cell are the coordinate of the cell center. For example, if `sp.resolution=1`, then `cell_lon=55.5` and `cell_lat=22.5` correspond to the square whose left boundary is 55, right boundary is 56, upper boundary is 23, and lower boundary is 22.

Value

An `n x 3` data frame containing three variables: `cell_ind` corresponds to unique id for each grid cell, `cell_lon` is the longitude of the grid cell, `cell_lat` is the latitude of the grid cell. Since the output data frame retains the order of the input coordinates, the original coordinate dataset and the output have can be linked one-to-one by the row index.

Examples

```
longitude <- runif(20, -90, 80)
latitude <- runif(20, 40, 60)
grid_locs <- grid_location(longitude, latitude, sp.resolution=0.5)
cbind(longitude, latitude, grid_locs)
```

`kernel_exp` *Exponential covariance function*

Description

Exponential covariance function

Usage

```
kernel_exp(x, sigma, ell, X1 = NULL, X2 = NULL)
```

Arguments

x	Distance measure.
sigma	The scale parameter with the constraint of `sigma` > `0`
ell	The range/lengthscale parameter with the constraint of `ell` > `0`.
X1	A `n1 x 2` matrix containing the coordinates of location set 1. If `x` is not provided, `X1` and `X2` should be provided for calculating their distance.
X2	A `n2 x 2` coordinate matrix.

Details

Let x = dist(x_i, x_j).

```
cov(i,j) = sigma^2*exp(-x/ell)
```

Value

A matrix or a scalar of exponential covariance depending on the type of `x` or whether `X1` and `X2` are used instead.

Examples

```
X1 <- cbind(runif(10, 1, 10), runif(10, 10, 20))
X2 <- cbind(runif(5, 1, 10), runif(5, 10, 20))

kernel_exp(sigma=2, ell=1, X1=X1, X2=X2)

kernel_exp(as.matrix(stats::dist(X1)), sigma=2, ell=1)
```

`kernel_matern` *Matern covariance function*

Description

Matern covariance function

Usage

```
kernel_matern(x, sigma, kappa, nu = 1, X1 = NULL, X2 = NULL)
```

Arguments

x	Distance measure.
sigma	Positive scale parameter.
kappa	Positive inverse range/lengthscale parameter.
nu	Smoothness parameter default to 1.
X1	A `n1 x 2` matrix containing the coordinates of location set 1. If `x` is not provided, `X1` and `X2` should be provided for calculating their distance.
X2	A `n2 x 2` coordinate matrix.

Details

Let x = dist(x_i, x_j).

```
cov(i,j) = sigma^2 * 2^(1-nu)/gamma(nu) * (kappa*x)^nu * K_v(kappa*x)
```

Note that when `nu=0.5`, the Matern kernel corresponds to the absolute exponential kernel.

Value

A matrix or a scalar of Matern covariance depending on the type of `x` or whether `X1` and `X2` are used instead.

Examples

```
X1 <- cbind(runif(10, 1, 10), runif(10, 10, 20))
X2 <- cbind(runif(5, 1, 10), runif(5, 10, 20))

kernel_matern(sigma=2, kappa=1, X1=X1, X2=X2)

kernel_matern(as.matrix(stats::dist(X1)), sigma=2, kappa=1)
```

`matern_pc_prior` *Helper funcion to specify a Penalized Complexity (PC) prior on the Matern hyperparameters*

Description

Helper funcion to specify a Penalized Complexity (PC) prior on the Matern hyperparameters

Usage

```
matern_pc_prior(rho_0, p_rho, sig_0, p_sig)
```

Arguments

`rho_0`	Hyperparameter for PC prior on the range parameter. Must be positive. See details.
`p_rho`	Hyperparameter for PC prior on the range parameter. Must be between 0 and 1. See details.
`sig_0`	Hyperparameter for PC prior on the scale parameter. Must be positive. See details.
`p_sig`	Hyperparameter for PC prior on the scale parameter. Must be between 0 and 1. See details.

Details

The joint prior on `rho` and `sig` achieves

```
P(rho < rho_0) = p_rho,
```

and

```
P(sig > sig_0) = p_sig,
```

where `rho` = `sqrt(8*nu)/kappa`.

Value

A list to provide to the `matern_pc_prior` argument of `spatialGEV_fit`.

References

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to construct priors. Statistical Science.

Examples

```
n_loc <- 20
y <- simulatedData2$y[1:n_loc]
locs <- simulatedData2$locs[1:n_loc,]
fit <- spatialGEV_fit(
  data = y,
  locs = locs,
  random = "abs",
  init_param = list(
    a = rep(0, n_loc),
    log_b = rep(0, n_loc),
    s = rep(-2, n_loc),
    beta_a = 0,
    beta_b = 0,
    beta_s = -2,
    log_sigma_a = 0,
    log_kappa_a = 0,
    log_sigma_b = 0,
    log_kappa_b = 0,
    log_sigma_s = 0,
    log_kappa_s = 0
  ),
  reparam_s = "positive",
  kernel = "matern",
  beta_prior = list(
    beta_a=c(0,100),
    beta_b=c(0,10),
    beta_s=c(0,10)
  ),
  matern_pc_prior = list(
    matern_a=matern_pc_prior(1e5,0.95,5,0.1),
    matern_b=matern_pc_prior(1e5,0.95,3,0.1),
    matern_s=matern_pc_prior(1e2,0.95,1,0.1)
  )
)
```

ONsnow — *Monthly total snowfall in Ontario, Canada from 1987 to 2021.*

Description

A dataset containing the monthly total snowfall (in cm) in Ontario, Canada from 1987 to 2021.

Usage

```
ONsnow
```

Format

A data frame with 63945 rows and 7 variables with each row corresponding to a monthly record at a weather location:

LATITUDE Numeric. Latitude of the weather station

LONGITUDE Numeric. Longitude of the weather station

STATION_NAME Character. Name of the weather station

CLIMATE_IDENTIFIER Character. Unique id of each station

LOCAL_YEAR Integer from 1987 to 2021. Year of the record

LOCAL_MONTH Integer from 1 to 12. Month of the record

TOTAL_SNOWFALL Positive number. Total monthly snowfall at a station in cm

Source

https://climate-change.canada.ca/climate-data/#/monthly-climate-summaries

---

`print.spatialGEVfit` *Print method for spatialGEVfit*

Description

Print method for spatialGEVfit

Usage

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

Arguments

`x`	Model object of class `spatialGEVfit` returned by `spatialGEV_fit`.
`...`	More arguments for `print`.

Value

Information about the fitted model containing number of fixed/random effects, fitting time, convergence information, etc.

print.spatialGEVpred *Print method for spatialGEVpred*

Description

Print method for spatialGEVpred

Usage

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

Arguments

`x`	Object of class `spatialGEVpred` returned by `spatialGEV_predict`.
`...`	Additional arguments for `print`.

Value

Information about the prediction.

print.spatialGEVsam *Print method for spatialGEVsam*

Description

Print method for spatialGEVsam

Usage

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

Arguments

`x`	Object of class `spatialGEVsam` returned by `spatialGEV_sample`.
`...`	Additional arguments for `print`.

Value

Information about the object including dimension and direction to use `summary` on the object.

simulatedData *Simulated dataset 1*

Description

A list of data used for package testing and demos. Both `a` and `logb` are simulated on smooth deterministic surfaces.

Usage

```
simulatedData
```

Format

A list containing the simulation parameters and simulated data on a 20x20 grid:

locs A 400x2 matrix. First column contains longitudes and second contains latitudes

a A length 400 vector. GEV location parameters

logb A length 400 vector. Log-transformed GEV scale parameters

logs A scalar. Log-transformed GEV shape parameter shared across space

y A length 400 list of vectors which are observations simulated at each location

simulatedData2 *Simulated dataset 2*

Description

A list of data used for package testing and demos. `a`, `logb`, `logs` are simulated from respective Gaussian random fields and thus are nonsmooth.

Usage

```
simulatedData2
```

Format

A list containing the simulation parameters and simulated data on a 20x20 grid:

locs A 400x2 matrix. First column contains longitudes and second contains latitudes

a A length 400 vector. GEV location parameters

logb A length 400 vector. Log-transformed GEV scale parameters

logs A length 400 vector. Log-transformed GEV shape parameters

y A length 400 list of vectors which are observations simulated at each location `sim_cond_normal` *Create a helper function to simulate from the conditional normal distribution of new data given old data*

Description

Create a helper function to simulate from the conditional normal distribution of new data given old data

Usage

```
sim_cond_normal(joint.mean, a, locs_new, locs_obs, kernel, ...)
```

Arguments

`joint.mean`	The length `n` mean vector of the MVN distribution. By default mu1 is the first `m` elements of `joint.mean`
`a`	A vector of length `n-m`, the values of mu2 to condition on
`locs_new`	A matrix containing the coordiantes of new locations
`locs_obs`	A matrix containing the coordinates of observed locations
`kernel`	A function (kernel function) that returns a matrix containing the similarity between the two arguments.
`...`	Hyperparameters to pass to the kernel function.

Details

This serves as a helper function for `spatialGEV_predict`. The notations are consistent to the notations on the MVN wikipedia page

Value

A function that takes in one argument `n` as the number of samples to draw from the condition normal distribution of `locs_new` given `locs_obs`: either from `rmvnorm` for MVN or `rnorm` for univariate normal. The old and new data are assumed to follow a joint multivariate normal distribution.

spatialGEV_fit *Fit a GEV-GP model.*

Description

Fit a GEV-GP model.

Usage

```
spatialGEV_fit(
  data,
  locs,
  random = c("a", "ab", "abs"),
  method = c("laplace", "maxsmooth"),
  init_param,
  reparam_s,
  kernel = c("spde", "matern", "exp"),
  X_a = NULL,
  X_b = NULL,
  X_s = NULL,
  nu = 1,
  s_prior = NULL,
  beta_prior = NULL,
  matern_pc_prior = NULL,
  return_levels = 0,
  get_return_levels_cov = T,
  sp_thres = -1,
  adfun_only = FALSE,
  ignore_random = FALSE,
  silent = FALSE,
  mesh_extra_init = list(a = 0, log_b = -1, s = 0.001),
  get_hessian = TRUE,
  ...
)

spatialGEV_model(
  data,
  locs,
  random = c("a", "ab", "abs"),
  method = c("laplace", "maxsmooth"),
  init_param,
  reparam_s,
  kernel = c("spde", "matern", "exp"),
  X_a = NULL,
  X_b = NULL,
  X_s = NULL,
  nu = 1,
```

```
  s_prior = NULL,
  beta_prior = NULL,
  matern_pc_prior = NULL,
  sp_thres = -1,
  ignore_random = FALSE,
  mesh_extra_init = list(a = 0, log_b = -1, s = 0.001),
  ...
)
```

Arguments

data	If method == "laplace", a list of length n_loc where each element contains the GEV observations at the given spatial location. If method == "maxsmooth" as list with two elements: est, an n_loc x 3 matrix of parameter estimates at each location, and var, a 3 x 3 x n_loc array of corresponding variance estimates.
locs	An n_loc x 2 matrix of longitude and latitude of the corresponding response values.
random	Either "a", "ab", or "abs", where a indicates the location parameter, b indicates the scale parameter, s indicates the shape parameter. This tells the model which GEV parameters are considered as random effects.
method	Either "laplace" or "maxsmooth". Default is "laplace". See details.
init_param	A list of initial parameters. See details.
reparam_s	A flag indicating whether the shape parameter is "zero", "unconstrained", constrained to be "negative", or constrained to be "positive". If model "abs" is used, reparam_s cannot be zero. See details.
kernel	Kernel function for spatial random effects covariance matrix. Can be "exp" (exponential kernel), "matern" (Matern kernel), or "spde" (Matern kernel with SPDE approximation described in Lindgren el al. 2011). To use the SPDE approximation, the user must first install the INLA R package.
X_a	n_loc x r_a design matrix for a, where r-1 is the number of covariates. If not provided, a n_loc x 1 column matrix of 1s is used.
X_b	n_loc x r_b design matrix for log(b). Does not need to be provided if b is fixed.
X_s	n_loc x r_s design matrix for g(s), where g() is a transformation function of s. Does not need to be provided if s is fixed.
nu	Hyperparameter of the Matern kernel. Default is 1.
s_prior	Optional. A length 2 vector where the first element is the mean of the normal prior on s or log(s) and the second is the standard deviation. Default is NULL, meaning a uniform prior is put on s if s is fixed, or a GP prior is applied if s is a random effect.
beta_prior	Optional named list that specifies normal priors on the GP mean function coefficients betas. Each element of the list should be a named length 2 vector in which the first element is mean and second element is sd. E.g. beta_prior=list(beta_a=c(0,100), beta_b=c(0,10), beta_s=c(-2,5)). Default is NULL, which means imposing a noninformative uniform flat prior.

matern_pc_prior
: Optional named list that specifies Penalized complexity priors on the GP Matern covariance hyperparameters sig and rho, where sig = sqrt(sigma) and rho = sqrt(8*nu)/kappa. Names must be matern_a, matern_b, or matern_s. E.g. matern_pc_prior=list(matern_s=matern_pc_prior(100, 0.9, 2, 0.1)). Default is NULL, which means a flat prior. See ?matern_pc_prior for more details.

return_levels
: Optional vector of return-level probabilities. If provided, the posterior mean and standard deviation of the upper-tail GEV quantile at each spatial location for each of these probabilities will be included in the summary output. See ?summary.spatialGEV_fit for details.

get_return_levels_cov
: Default is TRUE if return_levels is specified. Can be turned off for when the number of locations is large so that the high-dimensional covariance matrix for the return levels is not stored.

sp_thres
: Optional. Thresholding value to create sparse covariance matrix. Any distance value greater than or equal to sp_thres will be set to 0. Default is -1, which means not using sparse matrix. Caution: hard thresholding the covariance matrix often results in bad convergence.

adfun_only
: Only output the ADfun constructed using TMB? If TRUE, model fitting is not performed and only a TMB tamplate adfun is returned (along with the created mesh if kernel is "spde"). This can be used when the user would like to use a different optimizer other than the default nlminb. E.g., call optim(adfun$par, adfun$fn, adfun$gr) for optimization.

ignore_random
: Ignore random effect? If TRUE, spatial random effects are not integrated out in the model. This can be helpful for checking the marginal likelihood.

silent
: Do not show tracing information?

mesh_extra_init
: A named list of scalars. Used when the SPDE kernel is used. The list provides the initial values for a, log(b), and s on the extra triangles created in the mesh. The default is list(a=1, log_b=0, s=0.001).

get_hessian
: Default to TRUE so that spatialGEV_sample() can be used for sampling from the Normal approximated posterior with the inverse Hessian as the Normal covariance.

...
: Arguments to pass to INLA::inla.mesh.2d(). See details ?inla.mesh.2d() and Section 2.1 of Lindgren & Rue (2015) JSS paper. This is used specifically for when kernel="spde", in which case a mesh needs to be constructed on the spatial domain. When no arguments are passed to inla.mesh.2d(), a default argument is max.edge=2, which simply specifies the largest allowed triangle edge length. It is strongly suggested that the user should specify these arguments if they would like to use the SPDE kernel. Please make sure INLA package is installed before using the SPDE approximation.

Details

This function adopts Laplace approximation using TMB model to integrate out the random effects.

Specifying method="laplace" means integrating out the random effects u in the joint likelihood via the Laplace approximation: $p_{\mathrm{LA}}(y \mid \theta) \approx \int p(y, u \mid \theta)\, \mathrm{d}u$. Then the random effects posterior is constructed via a Normal approximation centered at the Laplace-approximated marginal likelihood mode with the covariance being the quadrature of it. If method="maxsmooth", the inference is carried out in two steps. First, the user provide the MLEs and variance estimates of a, b and s at each location to data, which is known as the max step. The max-step estimates are denoted as $\hat{u}$, and the likelihood function at each location is approximated by a Normal distribution at $\mathcal{N}(\hat{u}, \widehat{Var}(u))$. Second, the Laplace approximation is used to integrate out the random effects in the joint likelihood $p_{\mathrm{LA}}(\hat{u} \mid \theta) \approx \int p(\hat{u}, u \mid \theta)\, \mathrm{d}u$, followed by a Normal approximation at mode and quadrature of the approximated marginal likelihood $p_{\mathrm{LA}}(\hat{u} \mid \theta)$. This is known as the smooth step.

The random effects are assumed to follow Gaussian processes with mean 0 and covariance matrix defined by the chosen kernel function. E.g., using the exponential kernel function:

```
cov(i,j) = sigma*exp(-|x_i - x_j|/ell)
```

When specifying the initial parameters to be passed to init_param, care must be taken to count the number of parameters. Described below is how to specify init_param under different settings of random and kernel. Note that the order of the parameters must match the descriptions below (initial values specified below such as 0 and 1 are only examples).

- random = "a", kernel = "exp": a should be a vector and the rest are scalars. log_sigma_a and log_ell_a are hyperparameters in the exponential kernel for the Gaussian process describing the spatial variation of a.

```
init_param = list(a = rep(1,n_locations), log_b = 0, s = 1,
                  beta_a = rep(0, n_covariates),
                  log_sigma_a = 0, log_ell_a = 0)
```

Note that even if reparam_s=="zero", an initial value for s still must be provided, even though in this case the value does not matter anymore.

- random = "ab", kernel = "exp": When b is considered a random effect, its corresponding GP hyperparameters log_sigma_b and log_ell_b need to be specified.

```
init_param = list(a = rep(1,n_locations),
                  log_b = rep(0,n_locations), s=1,
                  beta_a = rep(0, n_covariates), beta_b = rep(0, n_covariates),
                  log_sigma_a = 0,log_ell_a = 0,
                  log_sigma_b = 0,log_ell_b = 0).
```

- random = "abs", kernel = "exp":

```
init_param = list(a = rep(1,n_locations),
                  log_b = rep(0,n_locations),
                  s = rep(0,n_locations),
                  beta_a = rep(0, n_covariates),
                  beta_b = rep(0, n_covariates),
                  beta_s = rep(0, n_covariates),
                  log_sigma_a = 0,log_ell_a = 0,
                  log_sigma_b = 0,log_ell_b = 0).
                  log_sigma_s = 0,log_ell_s = 0).
```

- random = "abs", kernel = "matern" or "spde": When the Matern or SPDE kernel is used, hyperparameters for the GP kernel are `log_sigma_a/b/s` and `log_kappa_a/b/s` for each spatial random effect.

```
init_param = list(a = rep(1,n_locations),
                  log_b = rep(0,n_locations),
                  s = rep(0,n_locations),
                  beta_a = rep(0, n_covariates),
                  beta_b = rep(0, n_covariates),
                  beta_s = rep(0, n_covariates),
                  log_sigma_a = 0,log_kappa_a = 0,
                  log_sigma_b = 0,log_kappa_b = 0).
                  log_sigma_s = 0,log_kappa_s = 0).
```

`raparam_s` allows the user to reparametrize the GEV shape parameter `s`. For example,

- if the data is believed to be right-skewed and lower bounded, this means `s>0` and one should use `reparam_s = "positive"`;
- if the data is believed to be left-skewed and upper bounded, this means `s<0` and one should use `reparam_s="negative"`.
- When `reparam_s = "zero"`, the data likelihood is a Gumbel distribution. In this case the data has no upper nor lower bound. Finally, specify `reparam_s = "unconstrained"` if no sign constraint should be imposed on `s`.

Note that when reparam_s = "negative" or "postive", the initial value of `s` in `init_param` should be that of log(|s|).

When the SPDE kernel is used, a mesh on the spatial domain is created using `INLA::inla.mesh.2d()`, which extends the spatial domain by adding additional triangles in the mesh to avoid boundary effects in estimation. As a result, the number of `a` and `b` will be greater than the number of locations due to these additional triangles: each of them also has their own `a` and `b` values. Therefore, the fit function will return a vector `meshidxloc` to indicate the positions of the observed coordinates in the random effects vector.

Value

If `adfun_only=TRUE`, this function outputs a list returned by `TMB::MakeADFun()`. This list contains components `par, fn, gr` and can be passed to an R optimizer. If `adfun_only=FALSE`, this function outputs an object of class `spatialGEVfit`, a list

- An adfun object
- A fit object given by calling `nlminb()` on the adfun
- An object of class `sdreport` from TMB which contains the point estimates, standard error, and precision matrix for the fixed and random effects
- Other helpful information about the model: kernel, data coordinates matrix, and optionally the created mesh if ‘kernel="spde" (See details).

`spatialGEV_model()` is used internally by `spatialGEV_fit()` to parse its inputs. It returns a list with elements `data`, `parameters`, `random`, and `map` to be passed to `TMB::MakeADFun()`. If `kernel == "spde"`, the list also contains an element `mesh`.

Examples

```
library(SpatialGEV)
n_loc <- 20
a <- simulatedData$a[1:n_loc]
logb <- simulatedData$logb[1:n_loc]
logs <- simulatedData$logs[1:n_loc]
y <- simulatedData$y[1:n_loc]
locs <- simulatedData$locs[1:n_loc,]
# No covariates are included, only intercept is included.
fit <- spatialGEV_fit(
  data = y,
  locs = locs,
  random = "ab",
  init_param = list(
    a = rep(0, n_loc),
    log_b = rep(0, n_loc),
    s = 0,
    beta_a = 0,
    beta_b = 0,
    log_sigma_a = 0,
    log_kappa_a = 0,
    log_sigma_b = 0,
    log_kappa_b = 0
  ),
  reparam_s = "positive",
  kernel = "matern",
  X_a = matrix(1, nrow=n_loc, ncol=1),
  X_b = matrix(1, nrow=n_loc, ncol=1),
  silent = TRUE
)
print(fit)

# To use a different optimizer other than the default `nlminb()`, create
# an object ready to be passed to optimizer functions using `adfun_only=TRUE`
obj <- spatialGEV_fit(
  data = y,
  locs = locs, random = "ab",
  init_param = list(
    a = rep(0, n_loc),
    log_b = rep(0, n_loc),
    s = 0,
    beta_a = 0,
    beta_b = 0,
    log_sigma_a = 0,
    log_kappa_a = 0,
    log_sigma_b = 0,
    log_kappa_b = 0
  ),
  reparam_s = "positive",
  kernel = "matern",
  X_a = matrix(1, nrow=n_loc, ncol=1),
  X_b = matrix(1, nrow=n_loc, ncol=1),
```

```
  adfun_only = TRUE
)
fit <- optim(obj$par, obj$fn, obj$gr)


# Using the SPDE kernel (SPDE approximation to the Matern kernel)
# Make sure the INLA package is installed before using `kernel="spde"`
## Not run:
library(INLA)
n_loc <- 20
y <- simulatedData2$y[1:n_loc]
locs <- simulatedData2$locs[1:n_loc,]
fit_spde <- spatialGEV_fit(
  data = y,
  locs = locs,
  random = "abs",
  init_param = list(
    a = rep(0, n_loc),
    log_b = rep(0, n_loc),
    s = rep(-2, n_loc),
    beta_a = 0,
    beta_b = 0,
    beta_s = -2,
    log_sigma_a = 0,
    log_kappa_a = 0,
    log_sigma_b = 0,
    log_kappa_b = 0,
    log_sigma_s = 0,
    log_kappa_s = 0
  ),
  reparam_s = "positive",
  kernel = "spde",
  beta_prior = list(
    beta_a=c(0,100),
    beta_b=c(0,10),
    beta_s=c(0,10)
  ),
  matern_pc_prior = list(
    matern_a=matern_pc_prior(1e5,0.95,5,0.1),
    matern_b=matern_pc_prior(1e5,0.95,3,0.1),
    matern_s=matern_pc_prior(1e2,0.95,1,0.1)
  )
)
plot(fit_spde$mesh) # Plot the mesh
points(locs[,1], locs[,2], col="red", pch=16) # Plot the locations

## End(Not run)
```

`spatialGEV_predict` *Draw from the posterior predictive distributions at new locations based on a fitted GEV-GP model*

Description

Draw from the posterior predictive distributions at new locations based on a fitted GEV-GP model

Usage

```
spatialGEV_predict(
  model,
  locs_new,
  n_draw,
  type = "response",
  X_a_new = NULL,
  X_b_new = NULL,
  X_s_new = NULL,
  parameter_draws = NULL
)
```

Arguments

`model`	A fitted spatial GEV model object of class `spatialGEVfit`
`locs_new`	A `n_test x 2` matrix containing the coordinates of the new locations
`n_draw`	Number of draws from the posterior predictive distribution
`type`	A character string: "response" or "parameters". The former returns draws from the posterior predictive distribution, and the latter returns parameter draws (all on original scale).
`X_a_new`	`n_test x r1` design matrix for a at the new locations. If not provided, the default is a column matrix of all 1s.
`X_b_new`	`n_test x r2` design matrix for log(b) at the new locations
`X_s_new`	`n_test x r2` design matrix for (possibly transformed) s at the new locations
`parameter_draws`	Optional. A `n_draw x n_parameter` matrix, or an object that is of class 'spatialGEVsam'. If `spatialGEV_sample()` has already been called, the output matrix of parameter draws can be supplied here to avoid doing sampling of parameters again. Make sure the number of rows of `parameter_draws` is the same as `n_draw`.

Value

An object of class `spatialGEVpred`, which is a list of the following components:

- An `n_draw x n_test` matrix `pred_y_draws` containing the draws from the posterior predictive distributions at `n_test` new locations
- An `n_test x 2` matrix `locs_new` containing the coordinates of the test data
- An `n_train x 2` matrix `locs_obs` containing the coordinates of the observed data

Examples

```
set.seed(123)
library(SpatialGEV)
n_loc <- 20
a <- simulatedData$a[1:n_loc]
logb <- simulatedData$logb[1:n_loc]
logs <- simulatedData$logs[1:n_loc]
y <- simulatedData$y[1:n_loc]
locs <- simulatedData$locs[1:n_loc,]
n_test <- 5
test_ind <- sample(1:n_loc, n_test)

# Obtain coordinate matrices and data lists
locs_test <- locs[test_ind,]
y_test <- y[test_ind]
locs_train <- locs[-test_ind,]
y_train <- y[-test_ind]

# Fit the GEV-GP model to the training set
train_fit <- spatialGEV_fit(
  data = y_train,
  locs = locs_train,
  random = "ab",
  init_param = list(
    beta_a = mean(a),
    beta_b = mean(logb),
    a = rep(0, n_loc-n_test),
    log_b = rep(0, n_loc-n_test),
    s = 0,
    log_sigma_a = 1,
    log_kappa_a = -2,
    log_sigma_b = 1,
    log_kappa_b = -2
  ),
  reparam_s = "positive",
  kernel = "matern",
  silent = TRUE
)

pred <- spatialGEV_predict(
  model = train_fit,
  locs_new = locs_test,
  n_draw = 100
)
summary(pred)
```

`spatialGEV_sample` *Get posterior parameter draws from a fitted GEV-GP model.*

Description

Get posterior parameter draws from a fitted GEV-GP model.

Usage

```
spatialGEV_sample(model, n_draw, observation = FALSE, loc_ind = NULL)
```

Arguments

`model`	A fitted spatial GEV model object of class `spatialGEVfit`
`n_draw`	Number of draws from the posterior distribution
`observation`	whether to draw from the posterior distribution of the GEV observation?
`loc_ind`	A vector of location indices to sample from. Default is all locations.

Value

An object of class `spatialGEVsam`, which is a list with the following elements:

`parameter_draws` A matrix of joint posterior draws for the hyperparameters and the random effects at the `loc_ind` locations.

`y_draws` If `observation == TRUE`, a matrix of corresponding draws from the posterior predictive GEV distribution at the `loc_ind` locations.

Examples

```
library(SpatialGEV)
n_loc <- 20
a <- simulatedData$a[1:n_loc]
logb <- simulatedData$logb[1:n_loc]
logs <- simulatedData$logs[1:n_loc]
y <- simulatedData$y[1:n_loc]
locs <- simulatedData$locs[1:n_loc,]
beta_a <- mean(a); beta_b <- mean(logb)
fit <- spatialGEV_fit(
  data = y,
  locs = locs,
  random = "ab",
  init_param = list(
    beta_a = beta_a,
    beta_b = beta_b,
    a = rep(0, n_loc),
    log_b = rep(0, n_loc),
    s = 0,
    log_sigma_a = 0,
    log_kappa_a = 0,
    log_sigma_b = 0,
    log_kappa_b = 0
  ),
  reparam_s = "positive",
  kernel = "spde",
```

```
  silent = TRUE
)

loc_ind <- sample(n_loc, 5)
sam <- spatialGEV_sample(model=fit, n_draw=100,
                         observation=TRUE, loc_ind=loc_ind)

print(sam)
summary(sam)
```

---

summary.spatialGEVfit *Summary method for spatialGEVfit*

Description

Summary method for spatialGEVfit

Usage

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

Arguments

`object`	Object of class `spatialGEVfit` returned by `spatialGEV_fit`.
`...`	Additional arguments for `summary`. Not used.

Value

Point estimates and standard errors of fixed effects, random effects, and the return levels (if specified in `spatialGEV_fit()`) returned by TMB.

---

summary.spatialGEVpred *Summary method for spatialGEVpred*

Description

Summary method for spatialGEVpred

Usage

```
## S3 method for class 'spatialGEVpred'
summary(object, q = c(0.025, 0.25, 0.5, 0.75, 0.975), ...)
```

Arguments

`object`	Object of class `spatialGEVpred` returned by `spatialGEV_predict`.
`q`	A vector of quantile values used to summarize the samples. Default is `c(0.025, 0.25, 0.5, 0.75, 0.975)`.
`...`	Additional arguments for `summary`.

Value

Summary statistics of the posterior predictive samples.

`summary.spatialGEVsam` *Summary method for spatialGEVsam*

Description

Summary method for spatialGEVsam

Usage

```
## S3 method for class 'spatialGEVsam'
summary(object, q = c(0.025, 0.25, 0.5, 0.75, 0.975), ...)
```

Arguments

`object`	Object of class `spatialGEVsam` returned by `spatialGEV_sample`.
`q`	A vector of quantile values used to summarize the samples. Default is `c(0.025, 0.25, 0.5, 0.75, 0.975)`.
`...`	Additional arguments for `summary`. Not used.

Value

Summary statistics of the posterior samples.

Index