Package: serofoi (via r-universe)

October 29, 2026

Title Bayesian Estimation of the Force of Infection from Serological Data

Type Package

Version 1.0.3

Description Estimating the force of infection from time varying, age varying, or constant serocatalytic models from population based seroprevalence studies using a Bayesian framework, including data simulation functions enabling the generation of serological surveys based on this models. This tool also provides a flexible prior specification syntax for the force of infection and the seroreversion rate, as well as methods to assess model convergence and comparison criteria along with useful visualisation functions.

License MIT + file LICENSE

URL https://github.com/epiverse-trace/serofoi, https://epiverse-trace.github.io/serofoi/

BugReports https://github.com/epiverse-trace/serofoi/issues

Depends R (>= 4.1.0)

Imports bayesplot, checkmate, config, cowplot, dplyr, ggplot2, glue, graphics, loo, expm, methods, purrr, Rcpp (>= 0.12.0), rstan (>= 2.18.1), rstantools (>= 2.3.1), stats, tibble, tidyr

Suggests knitr, rlang, rmarkdown, scales, spelling, testthat (>= 3.0.0)

LinkingTo BH (>= 1.66.0), Rcpp (>= 0.12.0), RcppEigen (>= 0.3.3.3.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), StanHeaders (>= 2.18.0)

VignetteBuilder knitr

Biarch true

Config/Needs/website epiverse-trace/epiversetheme

Config/testthat/edition 3

Encoding UTF-8

Language en-GB

LazyData true

NeedsCompilation yes

RoxygenNote 7.3.2

SystemRequirements GNU make

Author Zulma M. Cucunubá [aut, cre]
(<https://orcid.org/0000-0002-8165-3198>), Nicolás T. Domínguez
[aut] (<https://orcid.org/0009-0002-8484-1298>), Ben Lambert
[aut], Pierre Nouvellet [aut], Geraldine Gómez [ctb]
(<https://orcid.org/0009-0007-8701-0568>), Jaime A.
Pavlich-Mariscal [ctb]
(<https://orcid.org/0000-0002-3892-6680>), Hugo Gruson [ctb]
(<https://orcid.org/0000-0002-4094-1476>), David Santiago
Quevedo [ctb] (<https://orcid.org/0000-0003-1583-4262>), Miguel
Gámez [ctb], Sumali Bajaj [ctb], Everlyn Kamau [ctb], Richard
Creswell [ctb], International Development Research Center
(IDRC) [fnd], Pontificia Universidad Javeriana [cph]

Maintainer Zulma M. Cucunubá <zulma.cucunuba@javeriana.edu.co>

Config/pak/sysreqs make

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

Date/Publication 2025-03-31 10:40:55 UTC

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

RemoteRef HEAD

RemoteSha beb5bc6ffe34a1f391521a92e8193c240d9fa2e5

Contents

serofoi-package 3
add_age_group_to_serosurvey 4
build_stan_data 5
chagas2012 6
chik2015 7
extract_central_estimates 7
fit_seromodel 8
get_foi_index 10
plot_foi_estimates 11
plot_rhats 12
plot_seromodel 13
plot_seroprev_estimates 15
plot_serosurvey 16
plot_summary 17
prepare_serosurvey_for_plot 18
prob_seroprev_age_by_age 19

prob_seroprev_age_time_by_age . . . 19
prob_seroprev_by_age . . . 20
prob_seroprev_gen_by_age . . . 21
prob_seroprev_time_by_age . . . 22
probability_exact_age_varying . . . 23
probability_exact_time_varying . . . 24
set_foi_init . . . 24
set_stan_data_defaults . . . 25
sf_cauchy . . . 26
sf_none . . . 26
sf_normal . . . 27
sf_uniform . . . 27
simulate_serosurvey . . . 28
simulate_serosurvey_age . . . 30
simulate_serosurvey_age_time . . . 31
simulate_serosurvey_general . . . 32
simulate_serosurvey_time . . . 34
summarise_central_estimate . . . 35
summarise_loo_estimate . . . 36
summarise_seromodel . . . 37
veev2012 . . . 39

Index **40**

`serofoi-package` *The 'serofoi' package.*

Description

A DESCRIPTION OF THE PACKAGE

Author(s)

Maintainer: Zulma M. Cucunubá <zulma.cucunuba@javeriana.edu.co> (ORCID)

Authors:

- Nicolás T. Domínguez <ex-ntorres@javeriana.edu.co> (ORCID)
- Ben Lambert
- Pierre Nouvellet

Other contributors:

- Geraldine Gómez (ORCID) [contributor]
- Jaime A. Pavlich-Mariscal (ORCID) [contributor]
- Hugo Gruson (ORCID) [contributor]
- David Santiago Quevedo (ORCID) [contributor]
- Miguel Gámez [contributor]

- Sumali Bajaj [contributor]
- Everlyn Kamau [contributor]
- Richard Creswell [contributor]
- International Development Research Center (IDRC) [funder]
- Pontificia Universidad Javeriana [copyright holder]

References

Stan Development Team (NA). RStan: the R interface to Stan. R package version 2.26.22. https://mc-stan.org

#' @keywords internal

See Also

Useful links:

- https://github.com/epiverse-trace/serofoi
- https://epiverse-trace.github.io/serofoi/
- Report bugs at https://github.com/epiverse-trace/serofoi/issues

---

`add_age_group_to_serosurvey` *Adds age group marker to serosurvey*

Description

Adds age group marker to serosurvey

Usage

```
add_age_group_to_serosurvey(serosurvey)
```

Arguments

`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models)
	`age_min` Floor value of the average between age_min and age_max
	`age_max` The size of the sample
	`n_sample` Number of samples for each age group
	`n_seropositive` Number of positive samples for each age group

Value

serosurvey with additional column specifying age group marker defined as the mean floor between `age_min` and `age_max`

build_stan_data *Builds stan data for sampling depending on the selected model*

Description

Builds stan data for sampling depending on the selected model

Usage

```
build_stan_data(
  serosurvey,
  model_type = "constant",
  foi_prior = sf_uniform(),
  foi_index = NULL,
  is_log_foi = FALSE,
  foi_sigma_rw = sf_none(),
  is_seroreversion = FALSE,
  seroreversion_prior = sf_none()
)
```

Arguments

serosurvey	survey_year Year in which the survey took place (only needed to plot time models) age_min Floor value of the average between age_min and age_max age_max The size of the sample n_sample Number of samples for each age group n_seropositive Number of positive samples for each age group
model_type	Type of the model. Either "constant", "age" or "time"
foi_prior	Force-of-infection distribution specified by means of the helper functions. Currently available options are: **sf_normal** Function to set normal distribution priors **sf_uniform** Function to set uniform distribution priors
foi_index	Integer vector specifying the age-groups for which Force-of-Infection values will be estimated. It can be specified by means of get_foi_index
is_log_foi	Boolean to set logarithmic scale in the FoI
foi_sigma_rw	Prior distribution for the standard deviation of the Force-of-Infection. Currently available options are: **sf_normal** Function to set normal distribution prior. Available for time models in the log-scale **sf_cauchy** Function to set Cauchy distribution prior. Available for time models in regular scale.
is_seroreversion	Boolean specifying whether to include seroreversion rate estimation in the model

`seroreversion_prior`
: seroreversion distribution specified by means of the helper functions. Currently available options are:

  sf_normal Function to set normal distribution priors

  sf_uniform Function to set uniform distribution priors

  sf_none Function to set no prior distribution

Value

List with necessary data for sampling the specified model

`chagas2012` *Chagas seroprevalence data in serofoi*

Description

Datasets that measure the seroprevalence of IgG antibodies against Trypanosoma cruzi infection in rural areas of Colombia corresponding to a serosurvey conducted in 2012 for a rural indigenous community known to have long-term endemic transmission, where some control interventions have taken place over the years.

Usage

```
data(chagas2012)
```

Format

`chagas2012`**:**

A `<data.frame>` with 4 rows and 5 columns:

survey_year Year in which the serosurvey was conducted

n_sample Number of collected samples per age group

n_seropositive Number of positive samples per age group

age_min Age group minimal age

age_max Age group maximal age

Examples

```
data(chagas2012)
```

chik2015 *Chikungunya seroprevalence data in serofoi*

Description

Datasets that measure the seroprevalence of IgG antibodies against the Chikungunya virus conducted in Bahia, Brazil in October-December 2015 by Dias et al. (2018). The survey was conducted immediately after a large Chikungunya epidemic in the area.

Usage

```
data(chik2015)
```

Format

chik2015:

A <data.frame> with 4 rows and 5 columns:

survey_year Year in which the serosurvey was conducted

n_sample Number of collected samples per age group

n_seropositive Number of positive samples per age group

age_min Age group minimal age

age_max Age group maximal age

Examples

```
data(chik2015)
```

extract_central_estimates *Extracts central estimates from stan_fit object for specified parameter*

Description

Extracts central estimates from stan_fit object for specified parameter

Usage

```
extract_central_estimates(
  seromodel,
  serosurvey,
  alpha = 0.05,
  par_name = "foi_vector"
)
```

Arguments

seromodel	stan_fit object obtained from sampling a model with fit_seromodel
serosurvey	survey_year Year in which the survey took place (only needed to plot time models) age_min Floor value of the average between age_min and age_max age_max The size of the sample n_sample Number of samples for each age group n_seropositive Number of positive samples for each age group
alpha	1 - alpha indicates the credibility level to be used
par_name	String specifying the parameter to be extracted from seromodel

Value

A dataframe with the following columns

median Median of the samples computed as the 0.5 quantile

lower Lower quantile alpha

upper Upper quantile 1 - alpha

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
central_estimates <- extract_central_estimates(
  seromodel,
  veev2012,
  par_name = "foi"
)
```

---

fit_seromodel *Runs specified stan model for the Force-of-Infection (FoI)*

Description

Runs specified stan model for the Force-of-Infection (FoI)

Usage

```
fit_seromodel(
  serosurvey,
  model_type = "constant",
  is_log_foi = FALSE,
  foi_prior = sf_normal(),
  foi_sigma_rw = sf_none(),
  foi_index = NULL,
  foi_init = NULL,
```

```
  is_seroreversion = FALSE,
  seroreversion_prior = sf_normal(),
  ...
)
```

Arguments

serosurvey	survey_year Year in which the survey took place (only needed to plot time models) age_min Floor value of the average between age_min and age_max age_max The size of the sample n_sample Number of samples for each age group n_seropositive Number of positive samples for each age group
model_type	Type of the model. Either "constant", "age" or "time"
is_log_foi	Boolean to set logarithmic scale in the FoI
foi_prior	Force-of-infection distribution specified by means of the helper functions. Currently available options are: **sf_normal** Function to set normal distribution priors **sf_uniform** Function to set uniform distribution priors
foi_sigma_rw	Prior distribution for the standard deviation of the Force-of-Infection. Currently available options are: **sf_normal** Function to set normal distribution prior. Available for time models in the log-scale **sf_cauchy** Function to set Cauchy distribution prior. Available for time models in regular scale.
foi_index	Integer vector specifying the age-groups for which Force-of-Infection values will be estimated. It can be specified by means of get_foi_index
foi_init	Initialization function for sampling. If null, default is chosen depending on the foi-scale of the model
is_seroreversion	Boolean specifying whether to include seroreversion rate estimation in the model
seroreversion_prior	seroreversion distribution specified by means of the helper functions. Currently available options are: **sf_normal** Function to set normal distribution priors **sf_uniform** Function to set uniform distribution priors **sf_none** Function to set no prior distribution
...	Additional parameters for rstan

Value

stan_fit object with Force-of-Infection and seroreversion (when applicable) samples

Examples

```
data(chagas2012)
seromodel <- fit_seromodel(
  serosurvey = chagas2012,
  model_type = "time",
  foi_index = data.frame(
    year = 1935:2011,
    foi_index = c(rep(1, 46), rep(2, 31))
  ),
  iter = 100
)
```

`get_foi_index` *Generates Force-of-Infection indexes for heterogeneous age groups*

Description

Generates a list of integers indexing together the time/age intervals for which FoI values will be estimated in fit_seromodel. The max value in `foi_index` corresponds to the number of FoI values to be estimated when sampling. The serofoi approach to fitting serological data currently supposes that FoI is piecewise-constant across either groups of years or ages, and this function creates a Data Frame that communicates this grouping to the Stan model

Usage

```
get_foi_index(serosurvey, group_size, model_type)
```

Arguments

`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`group_size`	Age groups size
`model_type`	Type of the model. Either "age" or "time"

Value

A Data Frame which describes the grouping of years or ages (dependent on model) into pieces within which the FoI is assumed constant when performing model fitting. A single FoI value will be estimated for ages/years assigned with the same index

Examples

```
data(chagas2012)
foi_index <- get_foi_index(chagas2012, group_size = 25, model_type = "time")
```

`plot_foi_estimates` *Plots Force-of-Infection central estimates*

Description

Plots Force-of-Infection central estimates

Usage

```
plot_foi_estimates(
  seromodel,
  serosurvey,
  alpha = 0.05,
  foi_df = NULL,
  foi_max = NULL,
  size_text = 11,
  plot_constant = FALSE,
  x_axis = NA
)
```

Arguments

`seromodel`	stan_fit object obtained from sampling a model with fit_seromodel
`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`alpha`	1 - alpha indicates the credibility level to be used
`foi_df`	Dataframe with columns `year/age` Year/Age (depending on the model) `foi` Force-of-infection values by year/age
`foi_max`	Max FoI value for plotting
`size_text`	Size of text for plotting (`base_size` in ggplot2)
`plot_constant`	boolean specifying whether to plot single Force-of-Infection estimate and its corresponding rhat value instead of showing this information in the summary. Only relevant when `seromodel@model_name == "constant"`)
`x_axis`	either `"time"` or `"age"`. Specifies time axis values label for constant model additional plots. Only relevant when and `seromodel@model_name == "constant"`

Value

ggplot object with estimated FoI

Examples

```
data(chagas2012)
seromodel <- fit_seromodel(
  serosurvey = chagas2012,
  model_type = "time",
  foi_index = data.frame(
    year = 1935:2011,
    foi_index = c(rep(1, 46), rep(2, 31))
  ),
  iter = 100,
  chains = 2
)
plot_foi_estimates(seromodel, chagas2012)
```

`plot_rhats` *Plot r-hats convergence criteria for the specified model*

Description

Plot r-hats convergence criteria for the specified model

Usage

```
plot_rhats(
  seromodel,
  serosurvey,
  size_text = 11,
  plot_constant = FALSE,
  x_axis = NA
)
```

Arguments

`seromodel`	stan_fit object obtained from sampling a model with fit_seromodel
`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`size_text`	Size of text for plotting (`base_size` in ggplot2)

plot_constant	boolean specifying whether to plot single Force-of-Infection estimate and its corresponding rhat value instead of showing this information in the summary. Only relevant when seromodel@model_name == "constant")
x_axis	either "time" or "age". Specifies time axis values label for constant model additional plots. Only relevant when and seromodel@model_name == "constant"

Value

ggplot object showing the r-hats of the model to be compared with the convergence criteria (horizontal dashed line)

Examples

```
data(chagas2012)
seromodel <- fit_seromodel(
  serosurvey = chagas2012,
  model_type = "time",
  foi_index = data.frame(
    year = 1935:2011,
    foi_index = c(rep(1, 46), rep(2, 31))
  ),
  iter = 100,
  chains = 2
)
plot_rhats(seromodel, chagas2012)
```

---

plot_seromodel *Visualise results of the provided model*

Description

Visualise results of the provided model

Usage

```
plot_seromodel(
  seromodel,
  serosurvey,
  alpha = 0.05,
  bin_serosurvey = FALSE,
  bin_step = 5,
  foi_df = NULL,
  foi_max = NULL,
  loo_estimate_digits = 1,
  central_estimate_digits = 2,
  seroreversion_digits = 2,
  rhat_digits = 2,
```

```
  size_text = 11,
  plot_constant = FALSE,
  x_axis = NA
)
```

Arguments

seromodel	stan_fit object obtained from sampling a model with fit_seromodel
serosurvey	survey_year Year in which the survey took place (only needed to plot time models) age_min Floor value of the average between age_min and age_max age_max The size of the sample n_sample Number of samples for each age group n_seropositive Number of positive samples for each age group
alpha	1 - alpha indicates the credibility level to be used
bin_serosurvey	If TRUE, serodata is binned by means of prepare_bin_serosurvey. Otherwise, age groups are kept as originally input.
bin_step	Integer specifying the age groups bin size to be used when bin_serosurvey is set to TRUE.
foi_df	Dataframe with columns year/age Year/Age (depending on the model) foi Force-of-infection values by year/age
foi_max	Max FoI value for plotting
loo_estimate_digits	Number of loo estimate digits
central_estimate_digits	Number of central estimate digits
seroreversion_digits	Number of seroreversion rate digits
rhat_digits	Number of rhat estimate digits
size_text	Size of text for plotting (base_size in ggplot2)
plot_constant	boolean specifying whether to plot single Force-of-Infection estimate and its corresponding rhat value instead of showing this information in the summary. Only relevant when seromodel@model_name == "constant")
x_axis	either "time" or "age". Specifies time axis values label for constant model additional plots. Only relevant when and seromodel@model_name == "constant"

Value

seromodel summary plot

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
plot_seromodel(seromodel, veev2012)
```

`plot_seroprev_estimates` *Plot seroprevalence estimates on top of the serosurvey*

Description

Plot seroprevalence estimates on top of the serosurvey

Usage

```
plot_seroprev_estimates(
  seromodel,
  serosurvey,
  alpha = 0.05,
  size_text = 11,
  bin_serosurvey = FALSE,
  bin_step = 5
)
```

Arguments

`seromodel`	stan_fit object obtained from sampling a model with fit_seromodel
`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`alpha`	1 - alpha indicates the credibility level to be used
`size_text`	Size of text for plotting (`base_size` in ggplot2)
`bin_serosurvey`	If `TRUE`, `serodata` is binned by means of `prepare_bin_serosurvey`. Otherwise, age groups are kept as originally input.
`bin_step`	Integer specifying the age groups bin size to be used when `bin_serosurvey` is set to `TRUE`.

Value

ggplot object with seroprevalence estimates and serosurveys plots

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
plot_seroprev_estimates(seromodel, veev2012)
```

`plot_serosurvey` *Plots seroprevalence from the given serosurvey*

Description

Plots seroprevalence from the given serosurvey

Usage

```
plot_serosurvey(
  serosurvey,
  size_text = 11,
  bin_serosurvey = FALSE,
  bin_step = 5
)
```

Arguments

`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`size_text`	Size of text for plotting (`base_size` in ggplot2)
`bin_serosurvey`	If `TRUE`, `serodata` is binned by means of `prepare_bin_serosurvey`. Otherwise, age groups are kept as originally input.
`bin_step`	Integer specifying the age groups bin size to be used when `bin_serosurvey` is set to `TRUE`.

Value

ggplot object with seroprevalence plot

Examples

```
# Chikungunya example serosurvey
data(chik2015)
plot_serosurvey(chik2015)

# VEEV example serosurvey
data(veev2012)
plot_serosurvey(veev2012)
```

`plot_summary` *Plots model summary*

Description

Plots model summary

Usage

```
plot_summary(
  seromodel,
  serosurvey,
  loo_estimate_digits = 1,
  central_estimate_digits = 2,
  rhat_digits = 2,
  size_text = 11,
  plot_constant = FALSE
)
```

Arguments

`seromodel`	stan_fit object obtained from sampling a model with fit_seromodel
`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`loo_estimate_digits`	Number of loo estimate digits
`central_estimate_digits`	Number of central estimate digits
`rhat_digits`	Number of rhat estimate digits
`size_text`	Size of text for plotting (`base_size` in ggplot2)
`plot_constant`	boolean specifying whether to plot single Force-of-Infection estimate and its corresponding rhat value instead of showing this information in the summary. Only relevant when `seromodel@model_name == "constant"`)

Value

ggplot object with a summary of the specified model

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
plot_summary(seromodel, veev2012)
```

`prepare_serosurvey_for_plot` *Prepares serosurvey for plotting*

Description

Adds seroprevalence values with corresponding binomial confidence interval

Usage

```
prepare_serosurvey_for_plot(serosurvey, alpha = 0.05)
```

Arguments

`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`alpha`	1 - alpha indicates the confidence level to be used

Value

serosurvey with additional columns:

seroprev Seroprevalence computed as the proportion of positive cases `n_seropositive` in the number of samples `n_sample` for each age group

seroprev_lower Lower limit of the binomial confidence interval of `seroprev`

seroprev_upper Upper limit of the binomial confidence interval of `seroprev`

prob_seroprev_age_by_age

Generate probabilities of seropositivity by age based on an age-varying Force-of-Infection (FoI) model.

Description

This function calculates the probabilities of seropositivity by age based on an age-varying FoI model. It takes into account the FoI and the rate of seroreversion.

Usage

```
prob_seroprev_age_by_age(foi, seroreversion_rate)
```

Arguments

foi	A dataframe containing the FoI values for different ages. It should have two columns: 'age' and 'foi'.
seroreversion_rate	A non-negative numeric value representing the rate of seroreversion.

Value

A dataframe with columns 'age' and 'seropositivity'.

prob_seroprev_age_time_by_age

Generate probabilities of seropositivity by age based on an age-and-time varying Force-of-Infection (FoI) model.

Description

This function calculates the probabilities of seropositivity by age based on an age-and-time-varying FoI model. It takes into account the FoI and the rate of seroreversion.

Usage

```
prob_seroprev_age_time_by_age(foi, seroreversion_rate)
```

Arguments

foi	A dataframe containing the FoI values for different ages. It should have three columns: 'year', 'age' and 'foi'.
seroreversion_rate	A non-negative numeric value representing the rate of seroreversion.

Value

A dataframe with columns 'age' and 'seropositivity'.

`prob_seroprev_by_age` *Generate probabilities of seropositivity by age based on model choice.*

Description

This function generates seropositivity probabilities based on either a time-varying Force-of-Infection (FoI) model, an age-varying FoI model, or an age-and-time-varying FoI model. In all cases, it is possible to optionally include seroreversion.

Usage

```
prob_seroprev_by_age(model, foi, seroreversion_rate = 0)
```

Arguments

`model` A string specifying the model type which can be either '"age"', '"time"', '"age-time"'.

`foi` A dataframe containing the FoI values. For time-varying models the columns should be:

year Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of the FoI by year

For age-varying models the columns should be:.

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

foi Corresponding values of the FoI by age

For age-and-time-varying models the columns should be:

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

time Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of FoI by age and year

`seroreversion_rate` A non-negative value determining the rate of seroreversion (per year). Default is 0.

Value

A dataframe with columns 'age' and 'seropositivity'.

Examples

```
prob_seroprev_by_age(
  model = "age",
  foi = data.frame(
    age = 1:80,
    foi = rep(0.01, 80)
  )
)
```

`prob_seroprev_gen_by_age` *Generate probabilities of seropositivity by age based on a general Force-of-Infection (FoI) model.*

Description

This function calculates the probabilities of seropositivity by age based on an abstract model of the serocatalytic system.

Usage

```
prob_seroprev_gen_by_age(
  construct_A_fun,
  calculate_seroprev_fun,
  initial_conditions,
  max_age,
  ...
)
```

Arguments

`construct_A_fun`	A function that constructs a matrix that defines the multiplier term in the linear ODE system.
`calculate_seroprev_fun`	A function which takes the state vector and returns the seropositive fraction.
`initial_conditions`	The initial state vector proportions for each birth cohort.
`max_age`	The maximum age to simulate seropositivity for.
`...`	Additional parameters for `construct_A_fun`

Value

A dataframe with columns 'age' and 'seropositivity'.

Examples

```
# define age- and time-specific multipliers
foi_df_time <- data.frame(
  year = seq(1946, 2025, 1),
  foi = c(rep(0, 40), rep(1, 40))
)

foi_df_age <- data.frame(
  age = 1:80,
  foi = 2 * dlnorm(1:80, meanlog = 3.5, sdlog = 0.5)
)

u <- foi_df_age$foi
v <- foi_df_time$foi

# function to construct A matrix for one piece
construct_A <- function(t, tau, u, v) {
  u_bar <- u[t - tau]
  v_bar <- v[t]

  A <- diag(-1, ncol = 12, nrow = 12)
  A[row(A) == (col(A) + 1)] <- 1
  A[1, 1] <- -u_bar * v_bar
  A[2, 1] <- u_bar * v_bar
  A[12, 12] <- 0

  A
}

# determines the sum of seropositive compartments of those still alive
calculate_seropositivity_fn <- function(Y) {
  sum(Y[2:11]) / (1 - Y[12])
}

# initial conditions in 12D state vector
initial_conditions <- rep(0, 12)
initial_conditions[1] <- 1

# calculate probability
seropositive_hiv <- prob_seroprev_gen_by_age(
  construct_A,
  calculate_seropositivity_fn,
  initial_conditions,
  max_age = 80,
  u,
  v
)
```

`prob_seroprev_time_by_age` *Generate probabilities of seropositivity by age based on a time-varying Force-of-Infection (FoI) model.*

Description

This function calculates the probabilities of seropositivity by age based on a time-varying FoI model. It takes into account the FoI and the rate of seroreversion.

Usage

```
prob_seroprev_time_by_age(foi, seroreversion_rate)
```

Arguments

`foi`	A dataframe containing the FoI values for different years. It should have two columns: 'year' and 'foi'.
`seroreversion_rate`	A non-negative numeric value representing the rate of seroreversion.

Value

A dataframe with columns 'age' and 'seropositivity'.

`probability_exact_age_varying` *Computes the probability of being seropositive when Forces-of-Infection (FoIs) vary by age*

Description

Computes the probability of being seropositive when Forces-of-Infection (FoIs) vary by age

Usage

```
probability_exact_age_varying(ages, fois, seroreversion_rate = 0)
```

Arguments

`ages`	Integer indicating the ages of the exposed cohorts
`fois`	Numeric atomic vector corresponding to the age-varying Force-of-Infection to simulate from
`seroreversion_rate`	Non-negative seroreversion rate. Default is 0.

Value

vector of probabilities of being seropositive for age-varying FoI including seroreversion (ordered from youngest to oldest individuals)

`probability_exact_time_varying` *Computes the probability of being seropositive when Forces-of-Infection (FoIs) vary by time*

Description

Computes the probability of being seropositive when Forces-of-Infection (FoIs) vary by time

Usage

```
probability_exact_time_varying(years, fois, seroreversion_rate = 0)
```

Arguments

`years`	Integer indicating the years covering the birth ages of the sample
`fois`	Numeric atomic vector corresponding to the age-varying FoI to simulate from
`seroreversion_rate`	Non-negative seroreversion rate. Default is 0.

Value

vector of probabilities of being seropositive for age-varying FoI including seroreversion (ordered from youngest to oldest individuals)

`set_foi_init` *Sets initialization function for sampling*

Description

Sets initialization function for sampling

Usage

```
set_foi_init(foi_init, is_log_foi, foi_index)
```

Arguments

`foi_init`	Initialization function for sampling. If null, default is chosen depending on the foi-scale of the model
`is_log_foi`	Boolean to set logarithmic scale in the FoI
`foi_index`	Integer vector specifying the age-groups for which Force-of-Infection values will be estimated. It can be specified by means of get_foi_index

Value

Function specifying initialization vector for the Force-of-Infection

Examples

```
data(chagas2012)
foi_index <- get_foi_index(chagas2012, group_size = 5, model_type = "age")
foi_init <- set_foi_init(
  foi_init = NULL,
  is_log_foi = FALSE,
  foi_index = foi_index
)
```

`set_stan_data_defaults` *Set stan data defaults for sampling*

Description

Set stan data defaults for sampling

Usage

```
set_stan_data_defaults(stan_data, is_log_foi = FALSE, is_seroreversion = FALSE)
```

Arguments

`stan_data`	List to be passed to rstan
`is_log_foi`	Boolean to set logarithmic scale in the FoI
`is_seroreversion`	Boolean specifying whether to include seroreversion rate estimation in the model

Value

List with default values of stan data for sampling

sf_cauchy *Sets Cauchy distribution parameters for sampling*

Description

Sets Cauchy distribution parameters for sampling

Usage

```
sf_cauchy(location = 0, scale = 1)
```

Arguments

location	Location of the Cauchy distribution
scale	Scale of the Cauchy distribution

Value

List with specified statistics and name of the distribution

Examples

```
my_prior <- sf_cauchy()
```

sf_none *Sets empty prior distribution*

Description

Sets empty prior distribution

Usage

```
sf_none()
```

Value

List with the name of the empty distribution

sf_normal *Sets normal distribution parameters for sampling*

Description

Sets normal distribution parameters for sampling

Usage

```
sf_normal(mean = 0, sd = 1)
```

Arguments

mean	Mean of the normal distribution
sd	Standard deviation of the normal distribution

Value

List with specified statistics and name of the model

Examples

```
my_prior <- sf_normal()
```

sf_uniform *Sets uniform distribution parameters for sampling*

Description

Sets uniform distribution parameters for sampling

Usage

```
sf_uniform(min = 0, max = 10)
```

Arguments

min	Minimum value of the random variable of the uniform distribution
max	Maximum value of the random variable of the uniform distribution

Value

List with specified statistics and name of the model

Examples

```
my_prior <- sf_uniform()
```

`simulate_serosurvey` *Simulate serosurvey data based on various Force-of-Infection (FoI) models.*

Description

This function generates binned serosurvey data based on either a time-varying FoI model, an age-varying FoI model, or an age-and-time-varying FoI model. In all cases, it is possible to optionally include seroreversion. This function allows construction of serosurveys with binned age groups, and it generates uncertainty in the distribution of a sample size within an age bin through multinomial sampling.

Usage

```
simulate_serosurvey(model, foi, survey_features, seroreversion_rate = 0)
```

Arguments

`model` A string specifying the model type which can be either '"age"', '"time"', '"age-time"'.

`foi` A dataframe containing the FoI values. For time-varying models the columns should be:

year Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of the FoI by year

For age-varying models the columns should be:.

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

foi Corresponding values of the FoI by age

For age-and-time-varying models the columns should be:

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

time Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of FoI by age and year

`survey_features` A dataframe containing information about the binned age groups and sample sizes for each. It should contain columns:

age_min Left limits of the age groups to be considered in the serosurvey

age_max Right limits of the age groups to be considered in the serosurvey

n_sample Number of samples by age group

The resulting age intervals are closed to the left `[` and open to the right `)`.

`seroreversion_rate` A non-negative value determining the rate of seroreversion (per year). Default is 0.

Value

A dataframe with simulated serosurvey data, including age group information, overall sample sizes, the number of seropositive individuals, and other survey features.

Examples

```
# time-varying model
foi_df <- data.frame(
  year = seq(1990, 2009, 1),
  foi = rnorm(20, 0.1, 0.01)
)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey(
model = "time",
foi = foi_df,
survey_features = survey_features)

# age-varying model
foi_df <- data.frame(
  age = seq(1, 20, 1),
  foi = rnorm(20, 0.1, 0.01)
)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey(
model = "age",
foi = foi_df,
survey_features = survey_features)

# age-and-time varying model
foi_df <- expand.grid(
  year = seq(1990, 2009, 1),
  age = seq(1, 20, 1)
)
foi_df$foi <- rnorm(20 * 20, 0.1, 0.01)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey(
model = "age-time",
foi = foi_df,
survey_features = survey_features)
```

`simulate_serosurvey_age` *Simulate serosurvey data based on an age-varying Force-of-Infection (FoI) model.*

Description

This function generates binned serosurvey data based on an age-varying FoI model, optionally including seroreversion. This function allows construction of serosurveys with binned age groups, and it generates uncertainty in the distribution of a sample size within an age bin through multinomial sampling.

Usage

```
simulate_serosurvey_age(foi, survey_features, seroreversion_rate = 0)
```

Arguments

`foi` A dataframe containing the FoI values. For time-varying models the columns should be:

year Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of the FoI by year

For age-varying models the columns should be:.

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

foi Corresponding values of the FoI by age

For age-and-time-varying models the columns should be:

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

time Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of FoI by age and year

`survey_features` A dataframe containing information about the binned age groups and sample sizes for each. It should contain columns:

age_min Left limits of the age groups to be considered in the serosurvey

age_max Right limits of the age groups to be considered in the serosurvey

n_sample Number of samples by age group

The resulting age intervals are closed to the left `[` and open to the right `)`.

`seroreversion_rate` A non-negative value determining the rate of seroreversion (per year). Default is 0.

Value

A dataframe with simulated serosurvey data, including age group information, overall sample sizes, the number of seropositive individuals, and other survey features.

Examples

```
# specify FOIs for each year
foi_df <- data.frame(
  age = seq(1, 20, 1),
  foi = rnorm(20, 0.1, 0.01)
)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey_age(
foi_df, survey_features)
```

---

simulate_serosurvey_age_time

Simulate serosurvey data based on an age-and-time-varying Force-of-Infection (FoI) model.

Description

This function generates binned serosurvey data based on an age-and-time-varying FoI model, optionally including seroreversion. This function allows construction of serosurveys with binned age groups, and it generates uncertainty in the distribution of a sample size within an age bin through multinomial sampling.

Usage

```
simulate_serosurvey_age_time(foi, survey_features, seroreversion_rate = 0)
```

Arguments

`foi` A dataframe containing the FoI values. For time-varying models the columns should be:

year Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of the FoI by year

For age-varying models the columns should be:.

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

foi Corresponding values of the FoI by age

For age-and-time-varying models the columns should be:

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

time Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of FoI by age and year

`survey_features`

A dataframe containing information about the binned age groups and sample sizes for each. It should contain columns:

age_min Left limits of the age groups to be considered in the serosurvey

age_max Right limits of the age groups to be considered in the serosurvey

n_sample Number of samples by age group

The resulting age intervals are closed to the left `[` and open to the right `)`.

`seroreversion_rate`

A non-negative value determining the rate of seroreversion (per year). Default is 0.

Value

A dataframe with simulated serosurvey data, including age group information, overall sample sizes, the number of seropositive individuals, and other survey features.

Examples

```
# specify FOIs for each year
foi_df <- expand.grid(
  year = seq(1990, 2009, 1),
  age = seq(1, 20, 1)
)
foi_df$foi <- rnorm(20 * 20, 0.1, 0.01)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey_age_time(
foi_df, survey_features)
```

`simulate_serosurvey_general` *Simulate serosurvey data based on general serocatalytic model.*

Description

This simulation method assumes only that the model system can be written as a piecewise-linear ordinary differential equation system.

Usage

```
simulate_serosurvey_general(
  construct_A_fun,
  calculate_seroprev_fun,
  initial_conditions,
  survey_features,
  ...
)
```

Arguments

`construct_A_fun`	A function that constructs a matrix that defines the multiplier term in the linear ODE system.
`calculate_seroprev_fun`	A function which takes the state vector and returns the seropositive fraction.
`initial_conditions`	The initial state vector proportions for each birth cohort.
`survey_features`	A dataframe containing information about the binned age groups and sample sizes for each. It should contain columns: **age_min** Left limits of the age groups to be considered in the serosurvey **age_max** Right limits of the age groups to be considered in the serosurvey **n_sample** Number of samples by age group The resulting age intervals are closed to the left `[` and open to the right `)`.
`...`	Additional parameters for `construct_A_fun`

Value

A dataframe with simulated serosurvey data, including age group information, overall sample sizes, the number of seropositive individuals, and other survey features.

Examples

```
foi_df_time <- data.frame(
  year = seq(1946, 2025, 1),
  foi = c(rep(0, 40), rep(1, 40))
)

foi_df_age <- data.frame(
  age = 1:80,
  foi = 2 * dlnorm(1:80, meanlog = 3.5, sdlog = 0.5)
)

# generate age and time dependent FoI from multipliers
foi_age_time <- expand.grid(
  year = foi_df_time$year,
  age = foi_df_age$age
```

```
) |>
  dplyr::left_join(foi_df_age, by = "age") |>
  dplyr::rename(foi_age = foi) |>
  dplyr::left_join(foi_df_time, by = "year") |>
  dplyr::rename(foi_time = foi) |>
  dplyr::mutate(foi = foi_age * foi_time) |>
  dplyr::select(-c("foi_age", "foi_time"))

# create survey features for simulating
max_age <- 80
n_sample <- 50
survey_features <- data.frame(
  age_min = seq(1, max_age, 5),
  age_max = seq(5, max_age, 5)) |>
  dplyr::mutate(n_sample = rep(n_sample, length(age_min))
  )

# simulate survey from age and time FoI
serosurvey <- simulate_serosurvey(
  model = "age-time",
  foi = foi_age_time,
  survey_features = survey_features
)
```

`simulate_serosurvey_time` *Simulate serosurvey data based on a time-varying Force-of-Infection (FoI) model.*

Description

This function generates binned serosurvey data based on a time-varying FoI model, optionally including seroreversion. This function allows construction of serosurveys with binned age groups, and it generates uncertainty in the distribution of a sample size within an age bin through multinomial sampling.

Usage

```
simulate_serosurvey_time(foi, survey_features, seroreversion_rate = 0)
```

Arguments

`foi`	A dataframe containing the FoI values. For time-varying models the columns should be: **year** Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey **foi** Corresponding values of the FoI by year For age-varying models the columns should be:.

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

foi Corresponding values of the FoI by age

For age-and-time-varying models the columns should be:

age Ages starting at 1 and up to the age of the oldest person in the serosurvey

time Calendar years starting at the birth year of the oldest person and up to the time of the serosurvey

foi Corresponding values of FoI by age and year

`survey_features`

A dataframe containing information about the binned age groups and sample sizes for each. It should contain columns:

age_min Left limits of the age groups to be considered in the serosurvey

age_max Right limits of the age groups to be considered in the serosurvey

n_sample Number of samples by age group

The resulting age intervals are closed to the left `[` and open to the right `)`.

`seroreversion_rate`

A non-negative value determining the rate of seroreversion (per year). Default is 0.

Value

A dataframe with simulated serosurvey data, including age group information, overall sample sizes, the number of seropositive individuals, and other survey features.

Examples

```
# specify FOIs for each year
foi_df <- data.frame(
  year = seq(1990, 2009, 1),
  foi = rnorm(20, 0.1, 0.01)
)
survey_features <- data.frame(
  age_min = c(1, 3, 15),
  age_max = c(2, 14, 20),
  n_sample = c(1000, 2000, 1500))
serosurvey <- simulate_serosurvey_time(
foi_df, survey_features)
```

`summarise_central_estimate` *Summarise central estimate*

Description

Summarise central estimate

Usage

```
summarise_central_estimate(
  seromodel,
  serosurvey,
  alpha,
  par_name = "seroreversion_rate",
  central_estimate_digits = 2
)
```

Arguments

seromodel	stan_fit object obtained from sampling a model with fit_seromodel
serosurvey	survey_year Year in which the survey took place (only needed to plot time models) age_min Floor value of the average between age_min and age_max age_max The size of the sample n_sample Number of samples for each age group n_seropositive Number of positive samples for each age group
alpha	1 - alpha indicates the credibility level to be used
par_name	String specifying the parameter to be extracted from seromodel
central_estimate_digits	Number of central estimate digits

Value

Text summarising specified central estimate

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
summarise_central_estimate(
  seromodel,
  veev2012,
  alpha = 0.05,
  par_name = "foi"
  )
```

---

summarise_loo_estimate

Extract specified loo estimate

Description

Extract specified loo estimate

Usage

```
summarise_loo_estimate(
  seromodel,
  par_loo_estimate = "elpd_loo",
  loo_estimate_digits = 2
)
```

Arguments

seromodel	stan_fit object obtained from sampling a model with fit_seromodel
par_loo_estimate	Name of the loo estimate to be extracted. Available options are: "elpd_loo" Expected log pointwise predictive density "p_loo" Effective number of parameters "looic" Leave-one-out cross-validation information criteria For additional information refer to loo.
loo_estimate_digits	Number of loo estimate digits

Value

Text summarising specified loo estimate

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
summarise_loo_estimate(seromodel)
```

---

summarise_seromodel *Summarise specified model*

Description

Summarise specified model

Usage

```
summarise_seromodel(
  seromodel,
  serosurvey,
  alpha = 0.05,
  par_loo_estimate = "elpd_loo",
  loo_estimate_digits = 1,
  central_estimate_digits = 2,
  rhat_digits = 2
)
```

Arguments

`seromodel`	stan_fit object obtained from sampling a model with fit_seromodel
`serosurvey`	`survey_year` Year in which the survey took place (only needed to plot time models) `age_min` Floor value of the average between age_min and age_max `age_max` The size of the sample `n_sample` Number of samples for each age group `n_seropositive` Number of positive samples for each age group
`alpha`	1 - alpha indicates the credibility level to be used
`par_loo_estimate`	Name of the loo estimate to be extracted. Available options are: `"elpd_loo"` Expected log pointwise predictive density `"p_loo"` Effective number of parameters `"looic"` Leave-one-out cross-validation information criteria For additional information refer to loo.
`loo_estimate_digits`	Number of loo estimate digits
`central_estimate_digits`	Number of central estimate digits
`rhat_digits`	Number of rhat estimate digits

Value

A list summarising the specified model

`model_name` Name of the model

`elpd` elpd and its standard deviation

`foi` Estimated foi with credible interval (for 'constant' model)

`foi_rhat` foi rhat value (for 'constant' model)

`seroreversion_rate` Estimated seroreversion rate

`seroreversion_rate_rhat` Seroreversion rate rhat value

Examples

```
data(veev2012)
seromodel <- fit_seromodel(veev2012, iter = 100)
summarise_seromodel(seromodel, veev2012)
```

`veev2012` *Venezuelan Equine Encephalitis Virus (VEEV) seroprevalence data in serofoi*

Description

Datasets that measure the seroprevalence of IgG antibodies against VEEV in a rural village in Panamá in 2012 [Carrera2020].

Usage

```
data(veev2012)
```

Format

`veev2012`**:**

A `<data.frame>` with 4 rows and 5 columns:

survey_year Year in which the serosurvey was conducted

n_sample Number of collected samples per age group

n_seropositive Number of positive samples per age group

age_min Age group minimal age

age_max Age group maximal age

Examples

```
data(veev2012)
```

Index

∗ **datasets**
    chagas2012, 6
    chik2015, 7
    veev2012, 39

add_age_group_to_serosurvey, 4

build_stan_data, 5

chagas2012, 6
chik2015, 7

extract_central_estimates, 7

fit_seromodel, 8, *8*, *10–12*, *14, 15*, *17*, *36–38*

get_foi_index, *5*, *9*, 10, *24*
ggplot2, *11, 12*, *14–17*

loo, *37, 38*

plot_foi_estimates, 11
plot_rhats, 12
plot_seromodel, 13
plot_seroprev_estimates, 15
plot_serosurvey, 16
plot_summary, 17
prepare_serosurvey_for_plot, 18
prob_seroprev_age_by_age, 19
prob_seroprev_age_time_by_age, 19
prob_seroprev_by_age, 20
prob_seroprev_gen_by_age, 21
prob_seroprev_time_by_age, 22
probability_exact_age_varying, 23
probability_exact_time_varying, 24

rstan, *9*, *25*

serofoi *(*serofoi-package*)*, 3
serofoi-package, 3
set_foi_init, 24
set_stan_data_defaults, 25
sf_cauchy, *5*, *9*, 26
sf_none, *6*, *9*, 26
sf_normal, *5, 6*, *9*, 27
sf_uniform, *5, 6*, *9*, 27
simulate_serosurvey, 28
simulate_serosurvey_age, 30
simulate_serosurvey_age_time, 31
simulate_serosurvey_general, 32
simulate_serosurvey_time, 34
summarise_central_estimate, 35
summarise_loo_estimate, 36
summarise_seromodel, 37

veev2012, 39