

# Package ‘EpiInvert’

August 9, 2022

**Version** 0.2.0

**Type** Package

**Title** Estimation of Time Varying Reproduction Numbers and Restored Incidence Curves Using Variational Techniques

**Maintainer** Luis Alvarez <lalvarez@ulpgc.es>

**Description** Estimation, by inverting a renewal equation, of time-varying reproduction numbers and restored incidence curves with festive days and weekly biases corrected as described in Alvarez et al. (2021) <[doi:10.1073/pnas.2105112118](https://doi.org/10.1073/pnas.2105112118)> and Alvarez et al. (2022) <[doi:10.3390/biology11040540](https://doi.org/10.3390/biology11040540)>. EpiInvert can manage daily incidence data and weekly aggregated incidence data. This version of the package also includes EpiInvertForecast, a learning method for the short time forecast of the restored incidence curve.

**URL** <https://github.com/lalvarezmat/EpiInvert>

**License** GPL (>= 2)

**Encoding** UTF-8

**LazyData** true

**Imports** Rcpp (>= 1.0.8.3)

**LinkingTo** Rcpp

**Suggests** knitr, ggplot2, grid, dplyr, testthat (>= 3.0.0), rmarkdown

**Config/testthat/edition** 3

**Depends** R (>= 2.10)

**LazyLoad** yes

**RoxygenNote** 7.1.2

**NeedsCompilation** yes

**Author** Luis Alvarez [aut, cre] (<<https://orcid.org/0000-0002-6953-9587>>),  
Jean-David Morel [ctb] (<<https://orcid.org/0000-0002-7122-9924>>),  
Jean-Michel Morel [ctb] (<<https://orcid.org/0000-0002-6108-897X>>)

**Repository** CRAN

**Date/Publication** 2022-08-09 21:50:08 UTC

## R topics documented:

EpiInvert	2
EpiInvertForecast	5
EpiInvertForecast_plot	7
EpiInvert_plot	8
festives	9
incidence	9
incidence_weekly_aggregated	10
restored_incidence_database	11
select_params	11
si_distr_data	12
<b>Index</b>	<b>13</b>

---

EpiInvert	<i>EpiInvert estimates the reproduction number <math>R_t</math> and a restored incidence curve from the original daily incidence curve and the serial interval distribution. EpiInvert also corrects the festive and weekly biases present in the registered daily incidence.</i>
-----------	---

---

### Description

EpiInvert estimates the reproduction number  $R_t$  and a restored incidence curve from the original daily incidence curve and the serial interval distribution. EpiInvert also corrects the festive and weekly biases present in the registered daily incidence.

### Usage

```
EpiInvert(
  incid,
  last_incidence_date,
  festive_days = rep("1000-01-01", 2),
  config = EpiInvert::select_params()
)
```

### Arguments

incid	The original daily incidence curve (a numeric vector).
last_incidence_date	The date of the last value of the incidence curve in the format YYYY-MM-DD. EpiInvert does not allow missing values. On days when a country does not report data, a zero must be registered as the value associated with the incidence of that day.
festive_days	The festive or anomalous dates in the format YYYY-MM-DD (a character vector). In these dates we "a priori" expect that the incidence value is perturbed.
config	An object of class estimate_R_config, as returned by function select_params. The element of config are :

- `si_distr`: a numeric vector with the distribution of the serial interval (the default value is an empty vector). If this vector is empty, the serial interval is estimated using a parametric shifted log-normal.
- `shift_si_distr`: shift of the above user provided serial interval. This shift can be negative, which means that secondary cases can show symptoms before the primary cases (the default value is 0).
- `max_time_interval`: Maximum number of days used by EpiInvert (the default value is 150, which means that EpiInvert uses the last 150 days. The computational cost strongly depends on this value).
- `mean_si`: mean of the parametric shifted log-normal to approximate the serial interval (in the case the above `si_distr` vector is empty), (the default value is 12.267893).
- `sd_si`: standard deviation of the parametric shifted log-normal to approximate the serial interval (in the case the above `si_distr` vector is empty), (the default value is 5.667547).
- `shift_si=`: shift of the parametric shifted log-normal to approximate the serial interval (in the case the above `si_distr` vector is empty), (the default value is -5).
- `Rt_regularization_weight`: regularization weight for  $R_t$  in the variational model used by EpiInvert (the default value is 5).
- `seasonality_regularization_weight`: regularization weight for the weekly bias correction factors in the variational model used by EpiInvert (the default value is 5).
- `incidence_weekly_aggregated`: a boolean value which determines if we use weekly aggregated incidence. In such case every week a single data is stored with the accumulated incidence in the last 7 days (the default value is FALSE).

## Details

EpiInvert estimates the reproduction number  $R_t$  and a restored incidence curve by inverting the renewal equation :

$$i(t) = \sum_k i(t-k)R(t-k)\Phi(k)$$

using a variational formulation. The theoretical foundations of the method can be found in references [1] and [2].

## Value

an object of class `estimate_EpiInvert`, given by a list with components:

- `i_original`: the original daily incidence curve. In the case of weekly aggregated incidence, we initialize the original curve assigning each day of the week  $1/7$  of the weekly aggregated value.
- `i_festive`: the incidence after correction of the festive days bias.

- `i_bias_free`: the incidence after correction of the festive and weekly biases.
- `i_restored`: the restored incidence obtained using the renewal equation.
- `Rt`: the reproduction number  $R_t$  obtained by inverting the renewal equation.
- `Rt_CI95`: 95% confidence interval radius for the value of  $R_t$  taking into account the variation of  $R_t$  when more days are added to the estimation.
- `seasonality`: the weekly bias correction factors.
- `dates`: a vector of dates corresponding to the incidence curve.
- `festive`: boolean associated to each incidence value to check if it has been considered as a festive or anomalous day.
- `epsilon`: normalized error curve obtained as  $(i\_bias\_free - i\_restored) / i\_restored^a$ .
- `power_a`: the power,  $a$ , which appears in the above expression of the normalized error.
- `si_distr`: values of the distribution of the serial interval used in the EpiInvert estimation.
- `shift_si_distr`: shift of the above distribution of the serial interval `si_distr`.

### Author(s)

Luis Alvarez <lalvarez@ulpgc.es>

### References

- [1] Alvarez, L.; Colom, M.; Morel, J.D.; Morel, J.M. Computing the daily reproduction number of COVID-19 by inverting the renewal equation using a variational technique. Proc. Natl. Acad. Sci. USA, 2021.
- [2] Alvarez, Luis, Jean-David Morel, and Jean-Michel Morel. "Modeling COVID-19 Incidence by the Renewal Equation after Removal of Administrative Bias and Noise" Biology 11, no. 4: 540. 2022.
- [3] Ritchie, H. et al. Coronavirus Pandemic (COVID-19), OurWorldInData.org. Available online: <https://ourworldindata.org/coronavirus-source-data> (accessed on 5 May 2022).

### Examples

```
## load data on COVID-19 daily incidence up to 2022-05-05 for France,
## and Germany (taken from the official government data) and for UK and
## the USA taken from reference [3]
data(incidence)

## EpiInvert execution for USA with no festive days specification
## using the incidence 70 days in the past
res <- EpiInvert(incidence$USA,
"2022-05-05",
"1000-01-01",
select_params(list(max_time_interval = 70))
)

## Plot of the results
EpiInvert_plot(res)
```

```

## load data of festive days for France, Germany, UK and the USA
data(festives)

## EpiInvert execution for France with festive days specification using
## 70 days in the past
res <- EpiInvert(incidence$FRA,"2022-05-05",festives$FRA,
                select_params(list(max_time_interval = 70)))

## Plot of the incidence between "2022-04-01" and "2022-05-01"
EpiInvert_plot(res,"incid","2022-04-01","2022-05-01")

## load data of a serial interval
data("si_distr_data")

## EpiInvert execution for Germany using the uploaded serial interval shifted
## -2 days, using the incidence 70 days in the past
res <- EpiInvert(incidence$DEU,"2022-05-05",festives$DEU,
                EpiInvert::select_params(list(si_distr = si_distr_data,
                shift_si_distr=-2,max_time_interval = 70)))

## Plot of the serial interval used (including the shift)
EpiInvert_plot(res,"SI")

## EpiInvert execution for UK changing the default values of the parametric
## serial interval (using a shifted log-normal) using 70 days in the past
res <- EpiInvert(incidence$UK,"2022-05-05",festives$UK,
                EpiInvert::select_params(list(mean_si = 11,sd_si=6,shift_si=-1,max_time_interval = 70))
                )

## Plot of the reproduction number Rt including an empiric 95% confidence
## interval of the variation of Rt. To calculate Rt on each day t, EpiInvert
## uses the past days (t'<=t) and the future days (t'>t) when available.
## Therefore, the EpiInvert estimate of Rt varies when there are more days
## available. This confidence interval reflects the expected variation of Rt
## as a function of the number of days after t available.
EpiInvert_plot(res,"R")

```

---

EpiInvertForecast

*EpiInvertForecast computes a 28-day forecast of the restored incidence curve including confidence interval radius using the percentiles 50,75,90 and 95. using the weekly seasonality, from the forecasted restored incidence curve we also estimate a 28-day forecast of the original incidence curve.*

---

## Description

EpiInvertForecast computes a 28-day forecast of the restored incidence curve including confidence interval radius using the percentiles 50,75,90 and 95. using the weekly seasonality, from the forecasted restored incidence curve we also estimate a 28-day forecast of the original incidence curve.

**Usage**

```
EpiInvertForecast(EpiInvert_result, restored_incidence_database)
```

**Arguments**

`EpiInvert_result`

output list of the EpiInvert execution including, in particular, the restored incidence curve and the seasonality.

`restored_incidence_database`

a database including 27,418 samples of different restored incidence curves computed by EpiInvert using real data. Each restored incidence curve includes the last 56 values of the sequence. That is this database can be viewed as a matrix of size 27,418 X 56

**Details**

EpiInvertForecast estimates a forecast of the restored incidence curve using a weighted average of 27,418 restored incidence curves previously estimated by EpiInvert and stored in the database "restored\_incidence\_database". The weight, in the average computation, of each restored incidence curve of the database depends on the similarity between the current curve in the last 28 days and the first 28 days of the database curve. Each database curve contains 56 days. The first 28 days are used for comparison with the current curve and the last 28 days are used for forecasting.

**Value**

a list with components:

- `dates`: a vector of dates corresponding to the forecast days.
- `i_restored_forecast`: a numeric vector with the forecast of the restored incidence curve for the next 28 days.
- `i_original_forecast`: a numeric vector with the forecast of the original incidence curve for the next 28 days.
- `i_restored_forecast_CI50`: radius of an empiric confidence interval, with percentile 50, for the restored incidence forecast following the number of days passed since the current day.
- `i_restored_forecast_CI75`: radius of an empiric confidence interval, with percentile 75, for the restored incidence forecast following the number of days passed since the current day.
- `i_restored_forecast_CI90`: radius of an empiric confidence interval, with percentile 90, for the restored incidence forecast following the number of days passed since the current day.
- `i_restored_forecast_CI95`: radius of an empiric confidence interval, with percentile 95, for the restored incidence forecast following the number of days passed since the current day.

**Author(s)**

Luis Alvarez <lalvarez@ulpgc.es>

## References

- [1] Alvarez, L.; Colom, M.; Morel, J.D.; Morel, J.M. Computing the daily reproduction number of COVID-19 by inverting the renewal equation using a variational technique. Proc. Natl. Acad. Sci. USA, 2021.
- [2] Alvarez, Luis, Jean-David Morel, and Jean-Michel Morel. "Modeling COVID-19 Incidence by the Renewal Equation after Removal of Administrative Bias and Noise" Biology 11, no. 4: 540. 2022.
- [3] Ritchie, H. et al. Coronavirus Pandemic (COVID-19), OurWorldInData.org. Available online: <https://ourworldindata.org/coronavirus-source-data> (accessed on 5 May 2022).
- [4] Alvarez, Luis, Jean-David Morel, and Jean-Michel Morel. EpiInvertForecast Available online: <https://ctim.ulpgc.es/covid19/EpiInvertForecastPaper.html>

## Examples

```
## load data on COVID-19 daily incidence up to 2022-05-05 for France,  
## and Germany (taken from the official government data) and for UK and  
## the USA taken from reference [3]  
data(incidence)  
  
## load of the database of restored incidence curves.  
data("restored_incidence_database")  
  
## EpiInvert execution for USA with no festive days specification  
## using the incidence 90 days in the past  
res <- EpiInvert(incidence$USA,  
"2022-05-05",  
"1000-01-01",  
select_params(list(max_time_interval = 90))  
)  
  
## EpiInvertForecast execution using the EpiInvert results obtained by USA  
forecast <- EpiInvertForecast(res,restored_incidence_database)
```

---

EpiInvertForecast\_plot

EpiInvertForecast\_plot *plot the restored incidence forecast*

---

## Description

EpiInvertForecast\_plot plot the restored incidence forecast

## Usage

```
EpiInvertForecast_plot(EpiInvert_results, Forecast, what = "CI95")
```

**Arguments**

EpiInvert_results	the list returned by the EpiInvert execution
Forecast	the list returned by the EpiInvertForecast execution
what	one of the following drawing options: <ul style="list-style-type: none"> <li>• CI50: use confidence interval for the restored incidence forecast with percentil 50.</li> <li>• CI75: use confidence interval for the restored incidence forecast with percentil 75.</li> <li>• CI90: use confidence interval for the restored incidence forecast with percentil 90.</li> <li>• CI95: use confidence interval for the restored incidence forecast with percentil 95.</li> </ul>

**Value**

a plot with the last 28 days of the original and restored incidence curves and a 28-day forecast of the same curves. It also includes a shaded area with an empiric confidence interval of the restored incidence forecast estimation

---

EpiInvert_plot	<i>EpiInvert_plot plots the results obtained by EpiInvert</i>
----------------	---

---

**Description**

EpiInvert\_plot plots the results obtained by EpiInvert

**Usage**

```
EpiInvert_plot(
  x,
  what = "all",
  date_start = "1000-01-01",
  date_end = "3000-01-01"
)
```

**Arguments**

x	an object of class estimate_EpiInvert.
what	one of the following drawing options: <ul style="list-style-type: none"> <li>• all: a plot combining the main EpiInvert results.</li> <li>• R: a plot of the reproduction number Rt estimation.</li> <li>• incid: a plot combining the obtained incidence curves.</li> <li>• SI: the serial interval used in the EpiInvert estimation.</li> </ul>
date_start	the start date to plot
date_end	the final date to plot



**Value**

a plot.

**See Also**

[EpiInvert](#)

---

festives	<i>A dataset containing festive days in France, Germany, UK and the USA</i>
----------	---

---

**Description**

A dataset containing festive days in France, Germany, UK and the USA

**Usage**

festives

**Format**

A list with 4 variables:

**FRA** festive day of France

**DEU** festive day of Germany

**UK** festive day of UK

**USA** festive day of USA

---

incidence	<i>A dataset containing daily incidence of COVID-19 for France, Germany, UK and the USA</i>
-----------	---

---

**Description**

A dataset containing daily incidence of COVID-19 for France, Germany, UK and the USA

**Usage**

incidence

**Format**

A data frame with 5 variables:

**date** date of the incidence.

**FRA** incidence of France.

**DEU** incidence of Germany.

**UK** incidence of UK.

**USA** incidence of USA.

An updated version of this dataset can be found at <https://www.ctim.es/covid19/incidence.csv>

**Source**

<https://github.com/owid/covid-19-data/tree/master/public/data>

<https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolu>

[https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4/page/page\\_1/](https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4/page/page_1/)

---

incidence\_weekly\_aggregated

*A dataset containing weekly aggregated incidence of COVID-19 for France, Germany, UK and the USA*

---

**Description**

A dataset containing weekly aggregated incidence of COVID-19 for France, Germany, UK and the USA

**Usage**

incidence\_weekly\_aggregated

**Format**

A data frame with 5 variables:

**date** date of the weekly aggregated incidence.

**FRA** weekly aggregated incidence of France.

**DEU** weekly aggregated incidence of Germany.

**UK** weekly aggregated incidence of UK.

**USA** weekly aggregated incidence of USA.

An updated version of this dataset can be found at <https://www.ctim.es/covid19/incidence.csv>

**Source**

<https://github.com/owid/covid-19-data/tree/master/public/data>

<https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolu>

[https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4/page/page\\_1/](https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4/page/page_1/)

---

restored\_incidence\_database

*A dataset including 27,418 samples of different restored incidence curves computed by EpiInvert using real data. Each restored incidence curve includes the last 56 values of the sequence.*

---

**Description**

A dataset including 27,418 samples of different restored incidence curves computed by EpiInvert using real data. Each restored incidence curve includes the last 56 values of the sequence.

**Usage**

restored\_incidence\_database

**Format**

A 27,418 X 56 numeric matrix

---

select\_params

*select\_params function to select EpiInvert parameters*

---

**Description**

select\_params function to select EpiInvert parameters

**Usage**

```
select_params(x = "")
```

**Arguments**

x                    a list with elements of the class estimate\_R\_config

**Value**

an object of class estimate\_R\_config.

**See Also**

[EpiInvert](#)

---

si_distr_data	<i>A dataset containing the values of a serial interval</i>
---------------	---

---

**Description**

A dataset containing the values of a serial interval

**Usage**

si\_distr\_data

**Format**

A numeric vector with 1 variable representing the serial interval

# Index

## \* datasets

- festives, [9](#)
- incidence, [9](#)
- incidence\_weekly\_aggregated, [10](#)
- restored\_incidence\_database, [11](#)
- si\_distr\_data, [12](#)

- EpiInvert, [2](#), [9](#), [11](#)
- EpiInvert\_plot, [8](#)
- EpiInvertForecast, [5](#)
- EpiInvertForecast\_plot, [7](#)

- festives, [9](#)

- incidence, [9](#)
- incidence\_weekly\_aggregated, [10](#)

- restored\_incidence\_database, [11](#)

- select\_params, [11](#)
- si\_distr\_data, [12](#)