Package 'eSIR'

December 10, 2021

Type Package

Title Extended State-Space SIR Models

Version 0.4.2

Description An implementation of extended state-space SIR models developed by Song Lab at UM school of Public Health. There are several functions available by 1) including a time-varying transmission modifier, 2) adding a time-dependent quarantine compartment, 3) adding a time-dependent antibody-immunization compartment. Wang L. (2020) <doi:10.6339/JDS.202007_18(3).0003>.

License CC BY 4.0

Encoding UTF-8

Imports coda (>= 0.19.3), chron (>= 2.3.54), data.table (>= 1.12.0), ggplot2 (>= 3.2.1), grDevices (>= 3.5.2), graphics (>= 3.5.2), gtools (>= 3.8.1), scales (>= 1.1.0), stats (>= 3.5.2), reshape2 (>= 1.4.3), rjags (>= 4.10), utils (>= 3.5.2)

Depends R (>= 3.5.2)

RoxygenNote 7.1.2

LazyData true

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Lili Wang [aut], Song Lab [aut] (<http://websites.umich.edu/~songlab/>), Paul Egeler [ctb] (Spectrum Health Office of Research and Education), Michael Kleinsasser [cre]

Maintainer Michael Kleinsasser <biostat-cran-manager@umich.edu>

Repository CRAN

Date/Publication 2021-12-10 13:20:02 UTC

R topics documented:

`confirmed` *Confirmed COVID-19 cases*

Description

Confirmed COVID-19 cases in US states

Format

a list with

- Province_State name of the US state
- date ... a column for each date

`death` *Confirmed COVID-19 deaths*

Description

Confirmed COVID-19 deaths in US states

Format

a list with

- Province_State name of the US state
- date ... a column for each date

eSAIR *Extended state-space SIR with a subset of the population showing antibody positivity*

Description

In this function we allow it to characterize time-varying immunization among a subset of the population that have been tested positive in an antibody assessment. We expanded the SIR model by adding a time-varying antibody-positive proportion α_t.

Usage

```
eSAIR(
  Y,
  R,
  alpha0 = NULL,
  change_time = NULL,
  begin_str = "01/13/2020",
  T_fin = 200,
  nchain = 4,
  nadapt = 10000,
  M = 500,
  thn = 10,
  nburnin = 200,
  dic = FALSE,
  death_in_R = 0.02,
  casename = "eSAIR",
  beta0 = 0.2586,
  gamma0 = 0.0821,
  R0 = beta0/gamma0,
  gamma0_sd = 0.1,
  R0_sd = 1,
  file_add = character(0),
  add_death = FALSE,
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  eps = 1e-10
)
```

Arguments

Y	the time series of daily observed infected compartment proportions.
R	the time series of daily observed removed compartment proportions, including death and recovered.

alpha0	a vector of values of the dirac delta function α_t. Each entry denotes the proportion that will be immunized at each change time point. Note that all the entries lie between 0 and 1, its default is NULL.
change_time	the change points over time corresponding to alpha0, to formulate the dirac delta function α_t; its defalt value is NULL.
begin_str	the character of starting time, the default is "01/13/2020".
T_fin	the end of follow-up time after the beginning date begin_str, the default is 200.
nchain	the number of MCMC chains generated by rjags, the default is 4.
nadapt	the iteration number of adaptation in the MCMC. We recommend using at least the default value 1e4 to obtained fully adapted chains.
M	the number of draws in each chain, with no thinning. The default is M=5e2 but suggest using 5e5.
thn	the thinning interval between mixing. The total number of draws thus would become round(M/thn)*nchain. The default is 10.
nburnin	the burn-in period. The default is 2e2 but suggest 2e5.
dic	logical, whether compute the DIC (deviance information criterion) for model selection.
death_in_R	the numeric value of average of cumulative deaths in the removed compartments. The default is 0.4 within Hubei and 0.02 outside Hubei.
casename	the string of the job's name. The default is "eSAIR".
beta0	the hyperparameter of average transmission rate, the default is the one estimated from the SARS first-month outbreak (0.2586).
gamma0	the hyperparameter of average removed rate, the default is the one estimated from the SARS first-month outbreak (0.0821).
R0	the hyperparameter of the mean reproduction number R0. The default is thus the ratio of beta0/gamma0, which can be specified directly.
gamma0_sd	the standard deviation for the prior distrbution of the removed rate γ, the default is 0.1.
R0_sd	the standard deviation for the prior disbution of R0, the default is 1.
file_add	the string to denote the location of saving output files and tables.
add_death	logical, whether add the approximate death curve to the plot, default is false.
save_files	logical, whether to save plots to file.
save_mcmc	logical, whether save (TRUE) all the MCMC outputs or not (FALSE).The output file will be an .RData file named by the $casename$. We include arrays of prevalence values of the three compartments with their matrices of posterior draws up to the last date of the collected data as theta_p[,,1] and afterwards as theta_pp[,,1] for θ_t^S, theta_p[,,2] and theta_pp[,,2] for θ_t^I, and theta_p[,,3] and theta_pp[,,3] for θ_t^R. The posterior draws of the prevalence process of the antibody-immunized compartment can be obtained via thetaA_p and thetaA_pp. Moreover, the input and predicted proportions Y, Y_pp, R and R_pp can also be retrieved. The prevalence and prediceted proportion matrices have rows for MCMC replicates, and columns for days. The MCMC posterior draws of other parameters including beta, gamma, R0, and variance controllers k_p, lambdaY_p, lambdaR_p are also available.

save_plot_data	logical, whether save the plotting data or not.
eps	a non-zero controller so that all the input Y and R values would be bounded above 0 (at least eps). Its default value is 1e-10

Value

casename	the predefined casename.
incidence_mean	mean cumulative incidence, the mean prevalence of cumulative confirmed cases at the end of the study.
incidence_ci	2.5%, 50%, and 97.5% quantiles of the incidences.
out_table	summary tables including the posterior mean of the prevalance processes of the 3 states compartments ($\theta_t^S, \theta_t^I, \theta_t^R, \theta_t^H$) at last date of data collected ((t') decided by the lengths of your input data Y and R), and their respective credible inctervals (ci); the respective means and ci's of the reporduction number (R0), removed rate (γ), transmission rate (β).
plot_infection	plot of summarizing and forecasting for the infection compartment, in which the vertial blue line denotes the last date of data collected (t'), the vertial darkgray line denotes the deacceleration point (first turning point) that the posterior mean first-derivative of infection prevalence $\dot{\theta}_t^I$ achieves the maximum, the vertical purple line denotes the second turning point that the posterior mean first-derivative infection proportion $\dot{\theta}_t^I$ equals zero, the darkgray line denotes the posterior mean of the infection prevalence θ_t^I and the red line denotes its posterior median.
plot_removed	plot of summarizing and forecasting for the removed compartment with lines similar to those in the plot_infection. The vertical lines are identical, but the horizontal mean and median correspond to the posterior mean and median of the removed process θ_t^R. An additional line indicates the estimated death prevalence from the input death_in_R.
spaghetti_plot	20 randomly selected MCMC draws of the first-order derivative of the posterior prevalence of infection, namely $\dot{\theta}_t^I$. The black curve is the posterior mean of the derivative, and the vertical lines mark times of turning points corresponding respectively to those shown in plot_infection and plot_removed. Moreover, the 95% credible intervals of these turning points are also highlighted by semi-transparent rectangles.
first_tp_mean	the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots plot_infection and plot_removed, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.
first_tp_mean	the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots plot_infection and plot_removed, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.
first_tp_ci	fwith first_tp_mean, it reports the corresponding credible interval and median.

second_tp_mean the date t at which $\theta_t^I = 0$, calculated as the average of the stationary points of all of posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "pruple" lines in the plots of `plot_infection` and `plot_removed`. The latter indicate stationary t at which the first-order derivative of the averaged posterior of θ_t^I equals zero.

second_tp_ci with `second_tp_mean`, it reports the corresponding credible interval and median.

dic_val the output of `dic.samples()` in `dic.samples`, computing deviance information criterion for model comparison.

gelman_diag_list

Since version 0.3.3, we incorporated Gelman And Rubin's Convergence Diagnostic using `gelman.diag`. We included both the statistics and their upper C.I. limits. Values substantially above 1 indicate lack of convergence. Error messages would be printed as they are. This would be only valid for multiple chains (e.g. nchain > 1). Note that for time dependent processes, we only compute the convergence of the last observation data (`T_prime`), though it shows to be `T_prime+1`, which is due to the day 0 for initialization.

Examples

```
NI_complete <- c(
  41, 41, 41, 45, 62, 131, 200, 270, 375, 444, 549, 729,
  1052, 1423, 2714, 3554, 4903, 5806, 7153, 9074, 11177,
  13522, 16678, 19665, 22112, 24953, 27100, 29631, 31728, 33366
)
RI_complete <- c(
  1, 1, 7, 10, 14, 20, 25, 31, 34, 45, 55, 71, 94,
  121, 152, 213, 252, 345, 417, 561, 650, 811, 1017,
  1261, 1485, 1917, 2260, 2725, 3284, 3754
)
N <- 58.5e6
R <- RI_complete / N
Y <- NI_complete / N - R # Jan13->Feb 11
change_time <- c("02/08/2020")
alpha0 <- c(0.2) # 20% of the susceptible population were found immunized
res.antibody <- eSAIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  alpha0 = alpha0, change_time = change_time,
  casename = "Hubei_antibody", save_files = TRUE, save_mcmc = FALSE,
  M = 5e2, nburnin = 2e2
)
res.antibody$plot_infection


change_time <- c("01/16/2020")
alpha0 <- c(0.2)
NI_complete2 <- c(41, 45)
RI_complete2 <- c(1, 1)
N2 <- 1E3
res3 <- eSAIR(
```

```
  RI_complete2 / N2,
  NI_complete2 / N2,
  begin_str = "01/13/2020",
  T_fin = 4,
  alpha0 = alpha0,
  change_time = change_time,
  dic = FALSE,
  casename = "Hubei_q",
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  M = 50,
  nburnin = 1
)
closeAllConnections()
```

---

qh.eSIR *Extended state-space SIR with quarantine*

Description

Fit an extended state-space SIR model being reduced by in-home hospitalization.

Usage

```
qh.eSIR(
  Y,
  R,
  phi0 = NULL,
  change_time = NULL,
  begin_str = "01/13/2020",
  T_fin = 200,
  nchain = 4,
  nadapt = 10000,
  M = 500,
  thn = 10,
  nburnin = 200,
  dic = FALSE,
  death_in_R = 0.02,
  casename = "qh.eSIR",
  beta0 = 0.2586,
  gamma0 = 0.0821,
  R0 = beta0/gamma0,
  gamma0_sd = 0.1,
  R0_sd = 1,
  file_add = character(0),
  add_death = FALSE,
```

```
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  eps = 1e-10
)
```

Arguments

Y	the time series of daily observed infected compartment proportions.
R	the time series of daily observed removed compartment proportions, including death and recovered.
phi0	a vector of values of the dirac delta function ϕ_t. Each entry denotes the proportion that will be qurantined at each change time point. Note that all the entries lie between 0 and 1, its default is NULL.
change_time	the change points over time corresponding to phi0, to formulate the dirac delta function ϕ_t; its defalt value is NULL.
begin_str	the character of starting time, the default is "01/13/2020".
T_fin	the end of follow-up time after the beginning date begin_str, the default is 200.
nchain	the number of MCMC chains generated by rjags, the default is 4.
nadapt	the iteration number of adaptation in the MCMC. We recommend using at least the default value 1e4 to obtained fully adapted chains.
M	the number of draws in each chain, with no thinning. The default is M=5e2 but suggest using 5e5.
thn	the thinning interval between mixing. The total number of draws thus would become round(M/thn)*nchain. The default is 10.
nburnin	the burn-in period. The default is 2e2 but suggest 2e5.
dic	logical, whether compute the DIC (deviance information criterion) for model selection.
death_in_R	the numeric value of average of cumulative deaths in the removed compartments. The default is 0.4 within Hubei and 0.02 outside Hubei.
casename	the string of the job's name. The default is "qh.eSIR".
beta0	the hyperparameter of average transmission rate, the default is the one estimated from the SARS first-month outbreak (0.2586).
gamma0	the hyperparameter of average removed rate, the default is the one estimated from the SARS first-month outbreak (0.0821).
R0	the hyperparameter of the mean reproduction number R0. The default is thus the ratio of beta0/gamma0, which can be specified directly.
gamma0_sd	the standard deviation for the prior distrbution of the removed rate γ, the default is 0.1.
R0_sd	the standard deviation for the prior disbution of R0, the default is 1.
file_add	the string to denote the location of saving output files and tables.
add_death	logical, whether add the approximate death curve to the plot, default is false.

`save_files` logical, whether to save plots to file.

`save_mcmc` logical, whether save (`TRUE`) all the MCMC outputs or not (`FALSE`).The output file will be an `.RData` file named by the $casename$. We include arrays of prevalence values of the three compartments with their matrices of posterior draws up to the last date of the collected data as `theta_p[,,1]` and afterwards as `theta_pp[,,1]` for θ_t^S, `theta_p[,,2]` and `theta_pp[,,2]` for θ_t^I, and `theta_p[,,3]` and `theta_pp[,,3]` for θ_t^R. The posterior draws of the prevalence process of the quarantine compartment can be obtained via `thetaQ_p` and `thetaQ_pp`. Moreover, the input and predicted proportions `Y`, `Y_pp`, `R` and `R_pp` can also be retrieved. The prevalence and prediceted proportion matrices have rows for MCMC replicates, and columns for days. The MCMC posterior draws of other parameters including `beta`, `gamma`, `R0`, and variance controllers `k_p`, `lambdaY_p`, `lambdaR_p` are also available.

`save_plot_data` logical, whether save the plotting data or not.

`eps` a non-zero controller so that all the input `Y` and `R` values would be bounded above 0 (at least `eps`). Its default value is 1e-10

Details

In this function we allow it to characterize time-varying proportions of susceptible due to government-enforced stringent in-home isolation. We expanded the SIR model by adding a quarantine compartment with a time-varying rate of quarantine ϕ_t, the chance of a susceptible person being willing to take in-home isolation at time t.

Value

`casename` the predefined `casename`.

`incidence_mean` mean cumulative incidence, the mean prevalence of cumulative confirmed cases at the end of the study.

`incidence_ci` 2.5%, 50%, and 97.5% quantiles of the incidences.

`out_table` summary tables including the posterior mean of the prevalence processes of the 3 states compartments ($\theta_t^S, \theta_t^I, \theta_t^R, \theta_t^H$) at last date of data collected ((t') decided by the lengths of your input data `Y` and `R`), and their respective credible intervals (ci); the respective means and ci's of the reproduction number (R0), removed rate (γ), transmission rate (β).

`plot_infection` plot of summarizing and forecasting for the infection compartment, in which the vertical blue line denotes the last date of data collected (t'), the vertical darkgray line denotes the deacceleration point (first turning point) that the posterior mean first-derivative of infection prevalence $\dot{\theta}_t^I$ achieves the maximum, the vertical purple line denotes the second turning point that the posterior mean first-derivative infection proportion $\dot{\theta}_t^I$ equals zero, the darkgray line denotes the posterior mean of the infection prevalence θ_t^I and the red line denotes its posterior median.

`plot_removed` plot of summarizing and forecasting for the removed compartment with lines similar to those in the `plot_infection`. The vertical lines are identical, but the horizontal mean and median correspond to the posterior mean and median of the

removed process θ_t^R. An additional line indicates the estimated death prevalence from the input `death_in_R`.

`spaghetti_plot` 20 randomly selected MCMC draws of the first-order derivative of the posterior prevalence of infection, namely $\dot{\theta}_t^I$. The black curve is the posterior mean of the derivative, and the vertical lines mark times of turning points corresponding respectively to those shown in `plot_infection` and `plot_removed`. Moreover, the 95% credible intervals of these turning points are also highlighted by semi-transparent rectangles.

`first_tp_mean` the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots `plot_infection` and `plot_removed`, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.

`first_tp_mean` the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots `plot_infection` and `plot_removed`, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.

`first_tp_ci` fwith `first_tp_mean`, it reports the corresponding credible interval and median.

`second_tp_mean` the date t at which $\theta_t^I = 0$, calculated as the average of the stationary points of all of posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "pruple" lines in the plots of `plot_infection` and `plot_removed`. The latter indicate stationary t at which the first-order derivative of the averaged posterior of θ_t^I equals zero.

`second_tp_ci` with `second_tp_mean`, it reports the corresponding credible interval and median.

`dic_val` the output of `dic.samples()` in `dic.samples`, computing deviance information criterion for model comparison.

`gelman_diag_list`

Since version 0.3.3, we incorporated Gelman And Rubin's Convergence Diagnostic using `gelman.diag`. We included both the statistics and their upper C.I. limits. Values substantially above 1 indicate lack of convergence. Error messages would be printed as they are. This would be only valid for multiple chains (e.g. nchain > 1). Note that for time dependent processes, we only compute the convergence of the last observation data (`T_prime`), though it shows to be `T_prime+1`, which is due to the day 0 for initialization.

Examples

```
NI_complete <- c(
  41, 41, 41, 45, 62, 131, 200, 270, 375, 444, 549, 729,
  1052, 1423, 2714, 3554, 4903, 5806, 7153, 9074, 11177,
  13522, 16678, 19665, 22112, 24953, 27100, 29631, 31728, 33366
)
RI_complete <- c(
  1, 1, 7, 10, 14, 20, 25, 31, 34, 45, 55, 71, 94, 121, 152, 213,
```

```
  252, 345, 417, 561, 650, 811, 1017, 1261, 1485, 1917, 2260,
  2725, 3284, 3754
)
N <- 58.5e6
R <- RI_complete / N
Y <- NI_complete / N - R # Jan13->Feb 11

change_time <- c("01/23/2020", "02/04/2020", "02/08/2020")
phi0 <- c(0.1, 0.4, 0.4)
res.q <- qh.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  phi0 = phi0, change_time = change_time,
  casename = "Hubei_q", save_files = TRUE, save_mcmc = FALSE,
  M = 5e2, nburnin = 2e2
)
res.q$plot_infection
# res.q$plot_removed

res.noq <- qh.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, casename = "Hubei_noq",
  M = 5e2, nburnin = 2e2
)
res.noq$plot_infection


change_time <- c("01/16/2020")
phi0 <- c(0.1)
NI_complete2 <- c(41, 45)
RI_complete2 <- c(1, 1)
N2 <- 1E3
res2 <- qh.eSIR(
  RI_complete2 / N2,
  NI_complete2 / N2,
  begin_str = "01/13/2020",
  T_fin = 4,
  phi0 = phi0,
  change_time = change_time,
  dic = FALSE,
  casename = "Hubei_q",
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  M = 50,
  nburnin = 1
)
closeAllConnections()
```

`recovered` *Confirmed COVID-19 recovered*

Description

Confirmed COVID-19 recovered in US states

Format

a list with

- Province_State name of the US state
- date ... a column for each date

`tvt.eSIR` *Fit extended state-space SIR model with time-varying transmission rates*

Description

Fit extended state-space SIR model with pre-specified changes in the transmission rate, either stepwise or continuous, accommodating time-varying quarantine protocols.

Usage

```
tvt.eSIR(
  Y,
  R,
  pi0 = NULL,
  change_time = NULL,
  exponential = FALSE,
  lambda0 = NULL,
  begin_str = "01/13/2020",
  T_fin = 200,
  nchain = 4,
  nadapt = 10000,
  M = 500,
  thn = 10,
  nburnin = 200,
  dic = FALSE,
  death_in_R = 0.02,
  beta0 = 0.2586,
  gamma0 = 0.0821,
  R0 = beta0/gamma0,
  gamma0_sd = 0.1,
  R0_sd = 1,
  casename = "tvt.eSIR",
  file_add = character(0),
  add_death = FALSE,
  save_files = FALSE,
  save_mcmc = FALSE,
```

```
  save_plot_data = FALSE,
  eps = 1e-10
)
```

Arguments

Y	the time series of daily observed infected compartment proportions.
R	the time series of daily observed removed compartment proportions, including death and recovered.
pi0	the time-dependent transmission rate modifier $\pi(t)$ between 0 and 1.
change_time	the change points over time for step function pi, defalt value is NULL.
exponential	logical, whether $\pi(t)$ is exponential $\exp(-\lambda_0 t)$ or not; the default is FALSE.
lambda0	the rate of decline in the exponential survival function in $\exp(-\lambda_0 t)$.
begin_str	the character of starting time, the default is "01/13/2020".
T_fin	the end of follow-up time after the beginning date begin_str, the default is 200.
nchain	the number of MCMC chains generated by rjags, the default is 4.
nadapt	the iteration number of adaptation in the MCMC. We recommend using at least the default value 1e4 to obtained fully adapted chains.
M	the number of draws in each chain, with no thinning. The default is M=5e2 but suggest using 5e5.
thn	the thinning interval between mixing. The total number of draws thus would become round(M/thn)*nchain. The default is 10.
nburnin	the burn-in period. The default is 2e2 but suggest 2e5.
dic	logical, whether compute the DIC (deviance information criterion) for model selection.
death_in_R	the numeric value of average of cumulative deaths in the removed compartments. The default is 0.4 within Hubei and 0.02 outside Hubei.
beta0	the hyperparameter of average transmission rate, the default is the one estimated from the SARS first-month outbreak (0.2586).
gamma0	the hyperparameter of average removed rate, the default is the one estimated from the SARS first-month outbreak (0.0821).
R0	the hyperparameter of the mean reproduction number R0. The default is thus the ratio of beta0/gamma0, which can be specified directly.
gamma0_sd	the standard deviation for the prior distrbution of the removed rate γ, the default is 0.1.
R0_sd	the standard deviation for the prior disbution of R0, the default is 1.
casename	the string of the job's name. The default is "tvt.eSIR".
file_add	the string to denote the location of saving output files and tables.
add_death	logical, whether add the approximate death curve to the plot, default is false.
save_files	logical, whether to save plots to file.

save_mcmc — logical, whether save (TRUE) all the MCMC outputs or not (FALSE).The output file will be an .RData file named by the $casename$. We include arrays of prevalence values of the three compartments with their matrices of posterior draws up to the last date of the collected data as theta_p[,,1] and afterwards as theta_pp[,,1] for θ_t^S, theta_p[,,2] and theta_pp[,,2] for θ_t^I, and theta_p[,,3] and theta_pp[,,3] for θ_t^R. Moreover, the input and predicted proportions Y, Y_pp, R and R_pp can also be retrieved. The prevalence and predicted proportion matrices have rows for MCMC replicates, and columns for days. The MCMC posterior draws of other parameters including beta_p, gamma_p, R0_p, and variance controllers k_p, lambdaY_p, lambdaR_p are also available.

save_plot_data — logical, whether save the plotting data or not.

eps — a non-zero controller so that all the input Y and R values would be bounded above 0 (at least eps). Its default value is 1e-10

Details

We fit a state-space model with extended SIR, in which a time-varying transmission rate modifier $\pi(t)$ (between 0 and 1) is introduced to model. This allows us to accommodate quarantine protocol changes and ignored resources of hospitalization. The form of reducing rate may be a step-function with jumps at times of big policy changes or a smooth exponential survival function $\exp(-\lambda_0 t)$. The parameters of the function and change points, if any, should be predefined.

Value

casename — the predefined casename.

incidence_mean — mean cumulative incidence, the mean prevalence of cumulative confirmed cases at the end of the study.

incidence_ci — 2.5%, 50%, and 97.5% quantiles of the incidences.

out_table — summary tables including the posterior mean of the prevalance processes of the 3 states compartments ($\theta_t^S, \theta_t^I, \theta_t^R$) at last date of data collected ((t') decided by the lengths of your input data Y and R), and their respective credible inctervals (ci); the respective means and ci's of the reporduction number (R0), removed rate (γ), transmission rate (β).

plot_infection — plot of summarizing and forecasting for the infection compartment, in which the vertial blue line denotes the last date of data collected (t'), the vertial darkgray line denotes the deacceleration point (first turning point) that the posterior mean first-derivative of infection prevalence $\dot{\theta}_t^I$ achieves the maximum, the vertical purple line denotes the second turning point that the posterior mean first-derivative infection proportion $\dot{\theta}_t^I$ equals zero, the darkgray line denotes the posterior mean of the infection prevalence θ_t^I and the red line denotes its posterior median.

plot_removed — plot of summarizing and forecasting for the removed compartment with lines similar to those in the plot_infection. The vertical lines are identical, but the horizontal mean and median correspond to the posterior mean and median of the removed process θ_t^R. An additional line indicates the estimated death prevalence from the input death_in_R.

`spaghetti_plot` 20 randomly selected MCMC draws of the first-order derivative of the posterior prevalence of infection, namely $\dot{\theta}_t^I$. The black curve is the posterior mean of the derivative, and the vertical lines mark times of turning points corresponding respectively to those shown in `plot_infection` and `plot_removed`. Moreover, the 95% credible intervals of these turning points are also highlighted by semi-transparent rectangles.

`first_tp_mean` the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots `plot_infection` and `plot_removed`, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.

`first_tp_ci` fwith `first_tp_mean`, it reports the corresponding credible interval and median.

`second_tp_mean` the date t at which $\theta_t^I = 0$, calculated as the average of the stationary points of all of posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "pruple" lines in the plots of `plot_infection` and `plot_removed`. The latter indicate stationary t at which the first-order derivative of the averaged posterior of θ_t^I equals zero.

`second_tp_ci` with `second_tp_mean`, it reports the corresponding credible interval and median.

`dic_val` the output of `dic.samples()` in `dic.samples`, computing deviance information criterion for model comparison.

`gelman_diag_list`

Since version 0.3.3, we incorporated Gelman And Rubin's Convergence Diagnostic using `gelman.diag`. We included both the statistics and their upper C.I. limits. Values substantially above 1 indicate lack of convergence. Error messages would be printed as they are. This would be only valid for multiple chains (e.g. nchain > 1). Note that for time dependent processes, we only compute the convergence of the last observation data (`T_prime`), though it shows to be `T_prime+1`, which is due to the day 0 for initialization.

Examples

```
NI_complete <- c(
  41, 41, 41, 45, 62, 131, 200, 270, 375, 444, 549, 729,
  1052, 1423, 2714, 3554, 4903, 5806, 7153, 9074, 11177,
  13522, 16678, 19665, 22112, 24953, 27100, 29631, 31728, 33366
)
RI_complete <- c(
  1, 1, 7, 10, 14, 20, 25, 31, 34, 45, 55, 71, 94, 121, 152, 213,
  252, 345, 417, 561, 650, 811, 1017, 1261, 1485, 1917, 2260,
  2725, 3284, 3754
)
N <- 58.5e6
R <- RI_complete / N
Y <- NI_complete / N - R # Jan13->Feb 11
### Step function of pi(t)
change_time <- c("01/23/2020", "02/04/2020", "02/08/2020")
```

```
pi0 <- c(1.0, 0.9, 0.5, 0.1)
res.step <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, pi0 = pi0, change_time = change_time, dic = TRUE,
  casename = "Hubei_step", save_files = TRUE,
  save_mcmc = FALSE, M = 5e2, nburnin = 2e2
)
res.step$plot_infection
res.step$plot_removed
res.step$dic_val

### continuous exponential function of pi(t)
res.exp <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, exponential = TRUE, dic = FALSE, lambda0 = 0.05,
  casename = "Hubei_exp", save_files = FALSE, save_mcmc = FALSE,
  M = 5e2, nburnin = 2e2
)
res.exp$plot_infection
# res.exp$plot_removed

### without pi(t), the standard state-space SIR model without intervention
res.nopi <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, casename = "Hubei_nopi", save_files = FALSE,
  M = 5e2, nburnin = 2e2
)
res.nopi$plot_infection
# res.nopi$plot_removed


change_time <- c("01/18/2020")
pi0<- c(1.0, 0.9)
NI_complete2 <- c(41, 45, 62, 131)
RI_complete2 <- c(1, 1, 7, 10)
N2 <- 1E3
res1 <- tvt.eSIR(
  RI_complete2 / N2,
  (NI_complete2 - RI_complete2) / N2,
  begin_str = "01/10/2020",
  T_fin =10,
  pi0 = pi0,
  change_time = change_time,
  dic = FALSE,
  casename = "Hubei_step",
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  M = 50,
  nburnin = 1
)
closeAllConnections()
```

`tvt.eSIR2` *Fit extended state-space SIR model with time-varying transmission rates*

Description

Fit extended state-space SIR model with pre-specified changes in the transmission rate, either stepwise or continuous, accommodating time-varying quarantine protocols.

Usage

```
tvt.eSIR2(
  Y,
  R,
  pi0 = NULL,
  change_time = NULL,
  exponential = FALSE,
  lambda0 = NULL,
  begin_str = "01/13/2020",
  T_fin = 200,
  nchain = 4,
  nadapt = 10000,
  M = 500,
  thn = 10,
  nburnin = 200,
  dic = FALSE,
  death_in_R = 0.02,
  beta0 = 0.2586,
  gamma0 = 0.0821,
  R0 = beta0/gamma0,
  gamma0_sd = 0.1,
  R0_sd = 1,
  casename = "tvt.eSIR2",
  file_add = character(0),
  add_death = FALSE,
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  eps = 1e-10,
  time_unit = 1
)
```

Arguments

`Y`	the time series of daily observed infected compartment proportions.
`R`	the time series of daily observed removed compartment proportions, including death and recovered.

pi0	the time-dependent transmission rate modifier $\pi(t)$ between 0 and 1.
change_time	the change points over time for step function pi, defalt value is NULL.
exponential	logical, whether $\pi(t)$ is exponential $\exp(-\lambda_0 t)$ or not; the default is FALSE.
lambda0	the rate of decline in the exponential survival function in $\exp(-\lambda_0 t)$.
begin_str	the character of starting time, the default is "01/13/2020".
T_fin	the end of follow-up time after the beginning date begin_str, the default is 200. This value must be longer than the length of your input observed time series data.
nchain	the number of MCMC chains generated by rjags, the default is 4.
nadapt	the iteration number of adaptation in the MCMC. We recommend using at least the default value 1e4 to obtained fully adapted chains.
M	the number of draws in each chain, with no thinning. The default is M=5e2 but suggest using 5e5.
thn	the thinning interval between mixing. The total number of draws thus would become round(M/thn)*nchain. The default is 10.
nburnin	the burn-in period. The default is 2e2 but suggest 2e5.
dic	logical, whether compute the DIC (deviance information criterion) for model selection.
death_in_R	the numeric value of average of cumulative deaths in the removed compartments. The default is 0.4 within Hubei and 0.02 outside Hubei.
beta0	the hyperparameter of average transmission rate, the default is the one estimated from the SARS first-month outbreak (0.2586).
gamma0	the hyperparameter of average removed rate, the default is the one estimated from the SARS first-month outbreak (0.0821).
R0	the hyperparameter of the mean reproduction number R0. The default is thus the ratio of beta0/gamma0, which can be specified directly.
gamma0_sd	the standard deviation for the prior distrbution of the removed rate γ, the default is 0.1.
R0_sd	the standard deviation for the prior disbution of R0, the default is 1.
casename	the string of the job's name. The default is "tvt.eSIR2".
file_add	the string to denote the location of saving output files and tables.
add_death	logical, whether add the approximate death curve to the plot, default is false.
save_files	logical, whether to save plots to file.
save_mcmc	logical, whether save (TRUE) all the MCMC outputs or not (FALSE).The output file will be an .RData file named by the $casename$. We include arrays of prevalence values of the three compartments with their matrices of posterior draws up to the last date of the collected data as theta_p[,,1] and afterwards as theta_pp[,,1] for θ_t^S, theta_p[,,2] and theta_pp[,,2] for θ_t^I, and theta_p[,,3] and theta_pp[,,3] for θ_t^R. Moreover, the input and predicted proportions Y, Y_pp, R and R_pp can also be retrieved. The prevalence and predicted proportion matrices have rows for MCMC replicates, and columns for days. The MCMC posterior draws of other parameters including beta_p, gamma_p, R0_p, and variance controllers k_p, lambdaY_p, lambdaR_p are also available.

save_plot_data	logical, whether save the plotting data or not.
eps	a non-zero controller so that all the input Y and R values would be bounded above 0 (at least eps). Its default value is 1e-10
time_unit	numeric, newly added argument, which can be changed to an integer (ceiling) of the input to indicate the time unit for each data point. The default is one-day, i.e., we let each input time series data correspond to each day. If this value is set to be 7, each data point represents one-week's aggregated data.

Details

We fit a state-space model with extended SIR, in which a time-varying transmission rate modifier $\pi(t)$ (between 0 and 1) is introduced to the model. This allows us to accommodate quarantine protocol changes and ignored resources of hospitalization. The form of reducing rate may be a step -function with jumps at times of big policy changes or a smooth exponential survival function $\exp(-\lambda_0 t)$. The parameters of the function and change points, if any, should be predefined.

Value

casename	the predefined casename.
incidence_mean	mean cumulative incidence, the mean prevalence of cumulative confirmed cases at the end of the study.
incidence_ci	2.5%, 50%, and 97.5% quantiles of the incidences.
out_table	summary tables including the posterior mean of the prevalence processes of the 3 states compartments ($\theta_t^S, \theta_t^I, \theta_t^R$) at last date of data collected ((t') decided by the lengths of your input data Y and R), and their respective credible intervals (ci); the respective means and ci's of the reproduction number (R0), removed rate (γ), transmission rate (β).
plot_infection	plot of summarizing and forecasting for the infection compartment, in which the vertical blue line denotes the last date of data collected (t'), the vertial dark-gray line denotes the deacceleration point (first turning point) that the posterior mean first-derivative of infection prevalence $\dot{\theta}_t^I$ achieves the maximum, the vertical purple line denotes the second turning point that the posterior mean first-derivative infection proportion $\dot{\theta}_t^I$ equals zero, the darkgray line denotes the posterior mean of the infection prevalence θ_t^I and the red line denotes its posterior median.
plot_removed	plot of summarizing and forecasting for the removed compartment with lines similar to those in the plot_infection. The vertical lines are identical, but the horizontal mean and median correspond to the posterior mean and median of the removed process θ_t^R. An additional line indicates the estimated death prevalence from the input death_in_R.
spaghetti_plot	20 randomly selected MCMC draws of the first-order derivative of the posterior prevalence of infection, namely $\dot{\theta}_t^I$. The black curve is the posterior mean of the derivative, and the vertical lines mark times of turning points corresponding respectively to those shown in plot_infection and plot_removed. Moreover, the 95% credible intervals of these turning points are also highlighted by semi-transparent rectangles.

`first_tp_mean` the date t at which $\ddot{\theta}_t^I = 0$, calculated as the average of the time points with maximum posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "darkgreen" lines in the two plots `plot_infection` and `plot_removed`, which indicate the stationary point such that the first-order derivative of the averaged posterior of θ_t^I reaches its maximum.

`first_tp_ci` fwith `first_tp_mean`, it reports the corresponding credible interval and median.

`second_tp_mean` the date t at which $\theta_t^I = 0$, calculated as the average of the stationary points of all of posterior first-order derivatives $\dot{\theta}_t^I$; this value may be slightly different from the one labeled by the "pruple" lines in the plots of `plot_infection` and `plot_removed`. The latter indicate stationary t at which the first-order derivative of the averaged posterior of θ_t^I equals zero.

`second_tp_ci` with `second_tp_mean`, it reports the corresponding credible interval and median.

`dic_val` the output of `dic.samples()` in `dic.samples`, computing deviance information criterion for model comparison.

`gelman_diag_list`

Since version 0.3.3, we incorporated Gelman And Rubin's Convergence Diagnostic using `gelman.diag`. We included both the statistics and their upper C.I. limits. Values substantially above 1 indicate lack of convergence. Error messages would be printed as they are. This would be only valid for multiple chains (e.g. nchain > 1). Note that for time dependent processes, we only compute the convergence of the last observation data (`T_prime`), though it shows to be `T_prime+1`, which is due to the day 0 for initialization.

Examples

```
NI_complete <- c(
  41, 41, 41, 45, 62, 131, 200, 270, 375, 444, 549, 729,
  1052, 1423, 2714, 3554, 4903, 5806, 7153, 9074, 11177,
  13522, 16678, 19665, 22112, 24953, 27100, 29631, 31728, 33366
)
RI_complete <- c(
  1, 1, 7, 10, 14, 20, 25, 31, 34, 45, 55, 71, 94, 121, 152, 213,
  252, 345, 417, 561, 650, 811, 1017, 1261, 1485, 1917, 2260,
  2725, 3284, 3754
)
N <- 58.5e6
R <- RI_complete / N
Y <- NI_complete / N - R # Jan13->Feb 11
### Step function of pi(t)
change_time <- c("01/23/2020", "02/04/2020", "02/08/2020")
pi0 <- c(1.0, 0.9, 0.5, 0.1)
res.step <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, pi0 = pi0, change_time = change_time, dic = TRUE,
  casename = "Hubei_step", save_files = TRUE,
  save_mcmc = FALSE, M = 5e2, nburnin = 2e2
)
```

```
res.step$plot_infection
res.step$plot_removed
res.step$dic_val

### continuous exponential function of pi(t)
res.exp <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, exponential = TRUE, dic = FALSE, lambda0 = 0.05,
  casename = "Hubei_exp", save_files = FALSE, save_mcmc = FALSE,
  M = 5e2, nburnin = 2e2
)
res.exp$plot_infection
# res.exp$plot_removed

### without pi(t), the standard state-space SIR model without intervention
res.nopi <- tvt.eSIR(Y, R,
  begin_str = "01/13/2020", death_in_R = 0.4,
  T_fin = 200, casename = "Hubei_nopi", save_files = FALSE,
  M = 5e2, nburnin = 2e2
)
res.nopi$plot_infection
# res.nopi$plot_removed


change_time <- c("01/18/2020")
pi0<- c(1.0, 0.9)
NI_complete2 <- c(41, 45, 62, 131)
RI_complete2 <- c(1, 1, 7, 10)
N2 <- 1E3
res1 <- tvt.eSIR(
  Y = RI_complete2 / N2,
  R = (NI_complete2 - RI_complete2) / N2,
  begin_str = "01/10/2020",
  T_fin =10,
  pi0 = pi0,
  change_time = change_time,
  dic = FALSE,
  casename = "Hubei_step",
  save_files = FALSE,
  save_mcmc = FALSE,
  save_plot_data = FALSE,
  M = 50,
  nburnin = 1
)
closeAllConnections()
```

USA_state_N *Population of US states*

Description

Data frame with the populations of each state.

Format

a list with

- Province_State State as a character string
- N Population of the state

Index