Credibility intervals for the Covid19 reproduction number

Gersende Fort
CNRS
Institut de Mathématiques de Toulouse

In collaboration with

- Patrice Abry, CNRS, Laboratoire de Physique de l'ENS Lyon
- Juliette Chevallier, INSA, Institut de Mathématiques de Toulouse
- Barbara Pascal, CNRS, Laboratoire des Sciences du Numérique de Nantes
- Nelly Pustelnik, CNRS, Laboratoire de Physique de l'ENS Lyon

Partly funded by
Fondation Simone et Cino Del Duca, Project OpSiMorE

Outline

- Uncertainty on the reproduction numbers: how to capture it ?
- The statistical model:
  - model
  - high probability intervals
  - a hidden variable model
- Choice of the hyperparameters
- Conclusion

Uncertainty on the reproduction numbers: how to capture it ?

A deterministic criterion

- For a **fixed time horizon** T,
- Observations:
  - $\mathsf{Z}_{1:T} := (\mathsf{Z}_1, \cdots, \mathsf{Z}_T)$,
  - the total infectiousness of infected individuals at time t, for all $t = 1, \cdots, T$: $\Phi_t := \sum_{u=1}^{\tau} \phi_u \mathsf{Z}_{t-u}$.
- positive parameters $\lambda_{\mathsf{R}}, \lambda_{\mathsf{O}}$,

Maximize:

$$(\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \mapsto \mathcal{U}(\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) := \sum_{t=1}^{T} (\mathsf{Z}_t \ln(\mathsf{R}_t \Phi_t + \mathsf{O}_t) - (\mathsf{R}_t \Phi_t + \mathsf{O}_t)) - \lambda_{\mathsf{R}} \|\mathsf{D}\, \mathsf{R}_{1:T}\|_1 - \lambda_{\mathsf{O}} \|\mathsf{O}_{1:T}\|_1$$

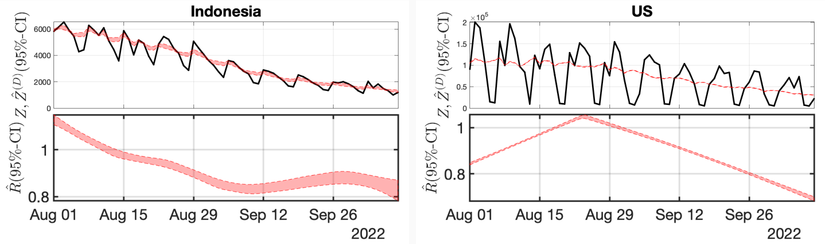
under the constraints

- non negative reproduction numbers: $\mathsf{R}_t \geq 0$
- non negative *intensity*: $\mathsf{R}_t \Phi_t + \mathsf{O}_t \geq 0$

$$\mathcal{D} := \bigcap_{t=1}^{T} \{\mathsf{R}_t \geq 0, \quad \mathsf{R}_t \Phi_t + \mathsf{O}_t \geq 0\}$$

Uncertainty on the R_t's and O_t's

- Set $\theta := (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$
- See θ as a random variable
- *Learn* the distribution of θ given the observations $\mathsf{Z}_1, \cdots, \mathsf{Z}_T$
- Find a $(1-\alpha)\%$ high probability interval for each R_t and each O_t



bottom For each $t \in \{1, \cdots, T\}$, each R_t: a 95% credibility interval.

top, black the observations $\mathsf{Z}_1, \cdots, \mathsf{Z}_T$.

top, red For each $t \in \{1, \cdots, T\}$, each O_t: a 95% credibility interval from which we deduce a credibility interval for a "denoised observation" $\mathsf{Z}_t - \widehat{\mathsf{O}_t}$

The statistical model

The distribution on θ

$\theta := (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$

- The contrast $\mathcal{U}$ to be maximized is a log-density
- The constraint set $\mathcal{D}$ is a support

The distribution of θ:

$$\pi(\theta) \propto \exp(\mathcal{U}(\theta)) \;\; \mathbf{1}_{\mathcal{D}}(\theta)$$

- The normalizing constant is not explicit (no closed form expression)

Remarks:

- A (the) maximizer of π solves

See the talk by Patrice Abry, for the computation

$$\mathrm{argmax}_{\theta \in \mathcal{D}} \;\; \mathcal{U}(\theta)$$

- $\mathcal{U}$ and $\mathcal{D}$ depend on the observations $(\mathsf{Z}_{1-\tau}, \cdots, \mathsf{Z}_T)$.

High probability region

$\theta := (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$, $\log \pi(\theta) = \mathcal{U}(\theta)$ **on** $\mathcal{D}$

Our approach:

- High probability intervals for each component of $\theta \rightarrow$ credibility intervals for each R_t and each O_t

- Intricate density π, known up to a normalizing constant
  - Markov chain Monte Carlo sampling: approximate π via *samples* $\theta^{(1)}, \cdots, \theta^{(M)}$ in $\mathcal{D}$,
  - For each component of θ: estimate the quantiles $q_{\alpha/2}$ and $q_{1-\alpha/2}$ of its marginal distribution from these samples
  - Obtain a $(1-\alpha)\%$ credibility interval for each R_t and each O_t

The MCMC sampling ($\star$)

$\theta := (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$,

$$\log \pi(\theta) = \sum_{t=1}^{T} \left(\mathsf{Z}_t \ln(\mathsf{R}_t \Phi_t + \mathsf{O}_t) - (\mathsf{R}_t \Phi_t + \mathsf{O}_t)\right) - \lambda_{\mathsf{R}} \|\mathsf{D}\mathsf{R}_{1:T}\|_1 - \lambda_{\mathsf{O}} \|\mathsf{O}_{1:T}\|_1 \text{ on } \mathcal{D}$$

Goal

- Sample an ergodic *Markov chain* $\theta^{(1)}, \cdots, \theta^{(k)}, \cdots$ having π as unique stationary distribution

Challenges

- $\log \pi(\theta) = -f(\theta) - g(\mathsf{A}\theta)$ sum of two terms
  - a C^1-function f
  - a non-smooth convex function g composed with a matrix A
- a support $\mathcal{D}$

Our approach: Hastings-Metropolis within Gibbs

- proposal: adapt a *Langevin Monte Carlo* dynamics
  - combine *gradient* and *proximal* operators
  - $g(\mathsf{A}\cdot)$ does not have an explicit proximal operator but $g(\cdot)$ has `PGdual`
  - $g(\mathsf{A}\cdot) = \sum_{i=1}^{3} g_i(\mathsf{A}_i \cdot)$ and $g_i(\mathsf{A}_i \cdot)$ has an explicit proximal operator. `PGdec`
- an acceptance-rejection step

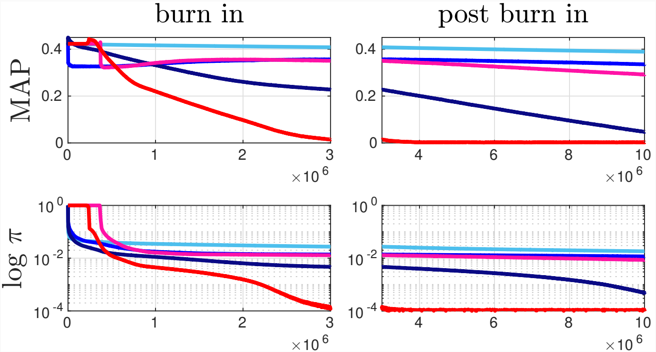


Figure 1: `RW` in light blue, `RW Invert` in blue and `RW Ortho` in dark blue; `PGdual Invert` in pink and `PGdual Ortho` in red. During the burn in period [left] and after [right], evolution of the distance to the MAP along iterations [top] and to $\max \ln \pi$ along iterations [bottom].

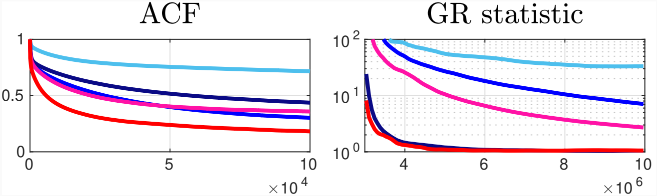


Figure 2: `RW` in light blue, `RW Invert` in blue and `RW Ortho` in dark blue; `PGdual Invert` in pink and `PGdual Ortho` in red. [left] Mean absolute value of the ACF vs the first 10^5 lags. [right] The GR statistic vs iterations.

A hidden variable model ?

$\theta := (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$, $\log \pi(\theta) = \mathcal{U}(\theta)$ **on** $\mathcal{D}$

- Observations: $\mathsf{Z}_t \in \mathbb{N}$, for $t = 1, \cdots, T$.
- Hidden variables: $\theta = (\mathsf{R}_{1:T}, \mathsf{O}_{1:T}) \in \mathbb{R}^{2T}$

A dynamical model:

- a time-evolution of the hidden processes
- a model for the observations, conditionally to the hidden processes

Remark:

$$\text{argmax}_{\text{constraints}}\, \mathcal{U}(\mathsf{R}_{1:T}, \mathsf{O}_{1:T})$$

is the computation of a Maximum a Posteriori (MAP).

A model for the hidden variables

- From the log-density of the *joint distribution*, equal to up to additive constants

$$\ln \pi(\theta) = \sum_{t=1}^{T} \left(\mathsf{Z}_t \ln(\mathsf{R}_t \Phi_t + \mathsf{O}_t) - (\mathsf{R}_t \Phi_t + \mathsf{O}_t)\right) \; -\lambda_{\mathsf{R}} \|\mathsf{D}\mathsf{R}_{1:T}\|_1 - \lambda_{\mathsf{O}} \|\mathsf{O}_{1:T}\|_1$$

on $\mathcal{D} := \bigcap_{t=1}^{T} \{\mathsf{R}_t \geq 0, \mathsf{R}_t \Phi_t + \mathsf{O}_t \geq 0\}$

- Hidden processes

  - Conditionally to $\mathsf{R}_1, \mathsf{R}_2$,

$$\mathsf{R}_t | \mathrm{past}_{t-1} \quad \sim \quad r \mapsto \frac{\lambda_{\mathsf{R}}}{2} \exp(-\lambda_{\mathsf{R}} |r - 2\mathsf{R}_{t-1} + \mathsf{R}_{t-2}|)$$

  - $\mathsf{O}_1, \cdots, \mathsf{O}_T$ are i.i.d. with distribution $u \mapsto (\lambda_{\mathsf{O}}/2) \exp(-\lambda_{\mathsf{O}} |u|)$
  - $(\mathsf{R}_3, \cdots, \mathsf{R}_T)$ and $(\mathsf{O}_3, \cdots, \mathsf{O}_T)$ are independent

A model for the observations

- From the log-density of the *joint distribution*, equal to up to additive constants

$$\ln \pi(\theta) = \boxed{\sum_{t=1}^{T} \left(\mathsf{Z}_t \ln(\mathsf{R}_t \Phi_t + \mathsf{O}_t) - (\mathsf{R}_t \Phi_t + \mathsf{O}_t)\right)} - \lambda_{\mathsf{R}} \|\mathsf{D}\mathsf{R}_{1:T}\|_1 - \lambda_{\mathsf{O}} \|\mathsf{O}_{1:T}\|_1$$

$$\boxed{\text{on } \mathcal{D} := \bigcap_{t=1}^{T} \{\mathsf{R}_t \geq 0, \mathsf{R}_t \Phi_t + \mathsf{O}_t \geq 0\}}$$

- The likelihood
  - if $(\mathsf{R}_t, \mathsf{O}_t)$ satisfies the constraints:

$$\mathsf{Z}_t | \mathrm{past}_{t-1} \sim \text{Poisson}\left(\mathsf{R}_t \Phi_t + \mathsf{O}_t\right) \qquad t \geq 3$$

  - otherwise

$$\mathsf{Z}_t | \mathrm{past}_{t-1} \sim \text{distribution that makes } \mathsf{Z}_t \in \mathbb{N} \text{ impossible}$$

Conclusion: a hidden Markov model

- With this descripion of a dynamical model

  $\pi(\theta) \propto \exp(-\mathcal{U}(\theta))$ on $\mathcal{D}$
  is proportional to an posteriori distribution of $(\mathsf{R}_{3:T}, \mathsf{O}_{3:T})$

- Corollary: a sequential model for the evolution of $(\mathsf{R}_t, \mathsf{O}_t, \mathsf{Z}_t)$
- Remark: the model depends on two parameters $\lambda_\mathsf{R}, \lambda_\mathsf{O}$

Choice of the hyperparameters

Hyperparameters: $\lambda_{\mathsf{R}}, \lambda_{\mathsf{O}}$

$$\log \pi(\theta) = \sum_{t=1}^{T} \left(\mathsf{Z}_t \ln(\mathsf{R}_t \Phi_t + \mathsf{O}_t) - (\mathsf{R}_t \Phi_t + \mathsf{O}_t)\right) - \lambda_{\mathsf{R}} \|\mathsf{D}\mathsf{R}_{1:T}\|_1 - \lambda_{\mathsf{O}} \|\mathsf{O}_{1:T}\|_1 \qquad \theta \in \mathcal{D}$$

- either fixed by the experts
- or estimated from the data $(\mathsf{Z}_1, \cdots, \mathsf{Z}_T)$

Our approach:

- We defined a joint distribution of the complete data $\mathsf{Z}_{3:T}, \mathsf{R}_{3:T}, \mathsf{O}_{3:T}$ conditionally to $\mathsf{R}_1, \mathsf{R}_2, \mathsf{Z}_{1-\tau}, \cdots, \mathsf{Z}_2$
- The distribution of the observations is the *marginal distribution*

$$\int \mathcal{L}\left(\mathsf{Z}_{3:T}, \mathsf{R}_{3:T}, \mathsf{O}_{3:T}; \lambda_{\mathsf{R}}, \lambda_{\mathsf{O}}\right) d\mathsf{R}_{3:T}\ d\mathsf{O}_{3:T}$$

- Strategy: choose $(\lambda_{\mathsf{R}}, \lambda_{\mathsf{O}})$ maximizing this likelihood

Solving the optimization problem

$$\operatorname{argmax}_{\lambda_{\mathsf{R}}>0,\lambda_{\mathsf{O}}>0} \int \mathcal{L}\left(\mathsf{Z}_{3:T}, \mathsf{R}_{3:T}, \mathsf{O}_{3:T}; \lambda_{\mathsf{R}}, \lambda_{\mathsf{O}}\right)\ \mathsf{d}\mathsf{R}_{3:T}\ \mathsf{d}\mathsf{O}_{3:T}$$

- The integral is not explicit
- The gradient is not explicit

Our approach:

- A stochastic optimization method among the "Majorization-Minimization" methods
- via a Stochastic Expectation-Maximization algorithm
- named the Stochastic Approximation Expectation Maximization (SAEM, 1999) algorithm

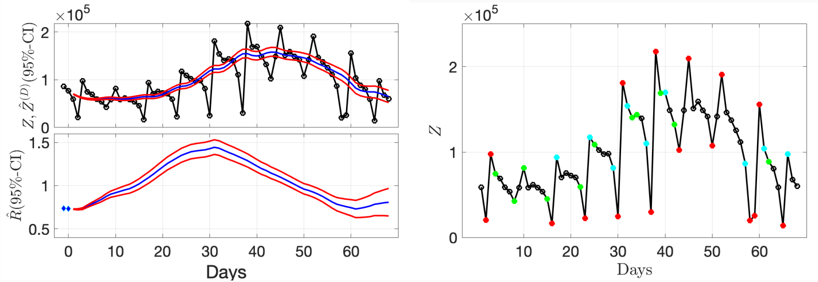


Figure 3: Left: Credibility interval estimates of R_t (bottom plot, in red), with initial values R_0, R_{-1} shown as blue diamonds; and of denoised counts $Z_t^{(D)}$ (top plot, in red), superimposed to counts Z_t (black dotted line). In blue, the estimate of the median.
Right: Counts Z_t marked in red, cyan, green and black, when the a posteriori probability that $|O_t|$ is large, is respectively in $[0.9, 1]$, in $[0.8, 0.9]$, in $[0.7, 0.8]$, and less than 0.7.

Conclusion

The contributions of the OptiMoCSI team

Modelization:

- $(\mathsf{R}_1, \cdots, \mathsf{R}_T)$ and $(\mathsf{O}_1, \cdots, \mathsf{O}_T)$ seen as random variables
- Definition of a probability distribution
- Credibility intervals via Monte Carlo sampling
- Data-based estimation of hyperparameters.

Computational statistics:

- New MCMC samplers
- A Stochastic EM algorithm for the estimation of the hyperparameters

Bibliography from the OptiMoCSI team

From a deterministic approach to a statistical modelization

P. Abry, G. Fort, B. Pascal and N. Pustelnik. Temporal Evolution of the Covid19 pandemic reproduction number: Estimations from Proximal optimization to Monte Carlo sampling. HAL-03565440. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, United Kingdom* 2022, pp. 167-170

Markov Chain Monte Carlo sampling

H. Artigas, B. Pascal, G. Fort, P. Abry and N. Pustelnik. Credibility Interval Design for Covid19 Reproduction Number from nonsmooth Langevin-type Monte Carlo sampling. HAL-03371837. *2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia*, 2022, pp. 2196-2200.

P. Abry, G. Fort, B. Pascal and N. Pustelnik. Credibility intervals for the reproduction number of the Covid-19 pandemic using Proximal Lanvevin samplers. October 2022. HAL-03902144 Accepted for publication in *EUSIPCO 2023*.

G. Fort, B. Pascal, P. Abry and N. Pustelnik. Covid19 Reproduction Number: Credibility Intervals by Blockwise Proximal Monte Carlo Samplers. HAL-03611079. *IEEE Trans. Signal Processing*, 71:888-900, 2023.

Stochastic Approximation EM

P. Abry, J. Chevallier, G. Fort and B. Pascal. Pandemic Intensity Estimation From Stochastic Approximation-Based Algorithms. Accepted to *CAMSAP 2023*, HAL-04174245v1.

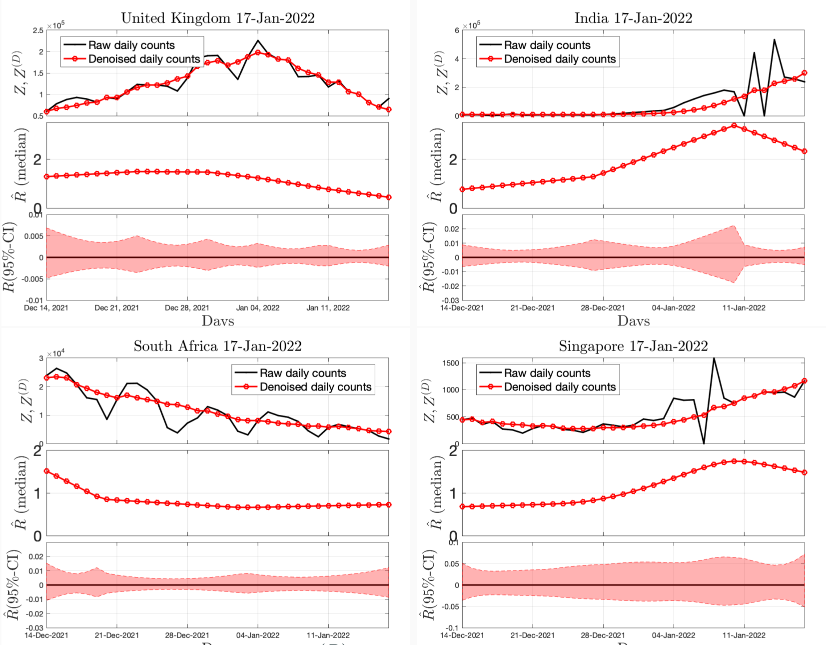


Figure 4: Top: Raw (Z, black) and denoised ($\widehat{Z}^{(D)}$, red) daily new infection counts. Middle: a posteriori median (50%-quantile) estimate for ω. Bottom: 95%-credibility interval estimate for ω, reported as the plots of the 97.5% and 2.5%-quantiles, after subtraction of the 50%-quantile.

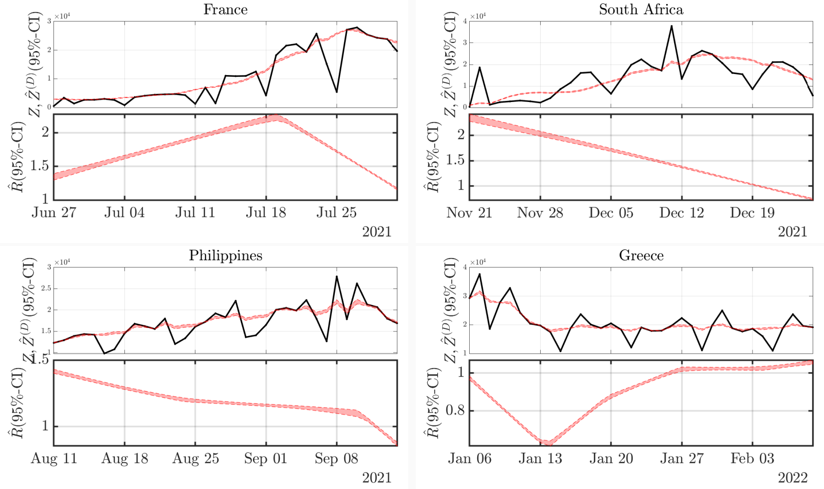
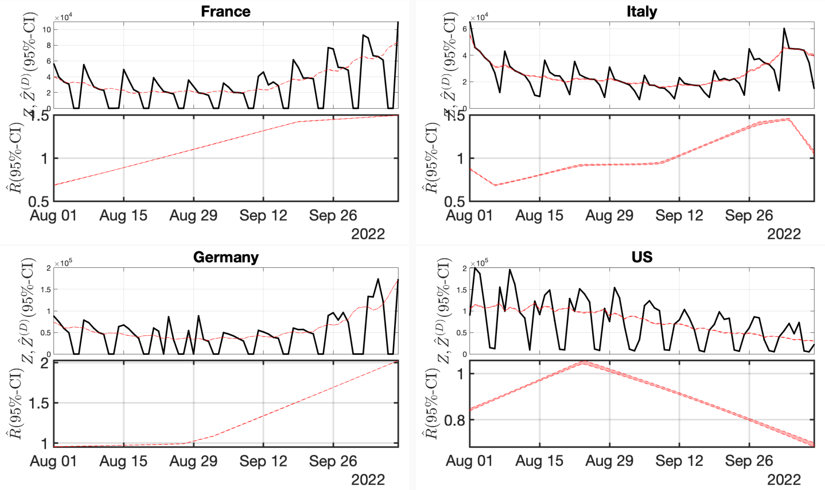
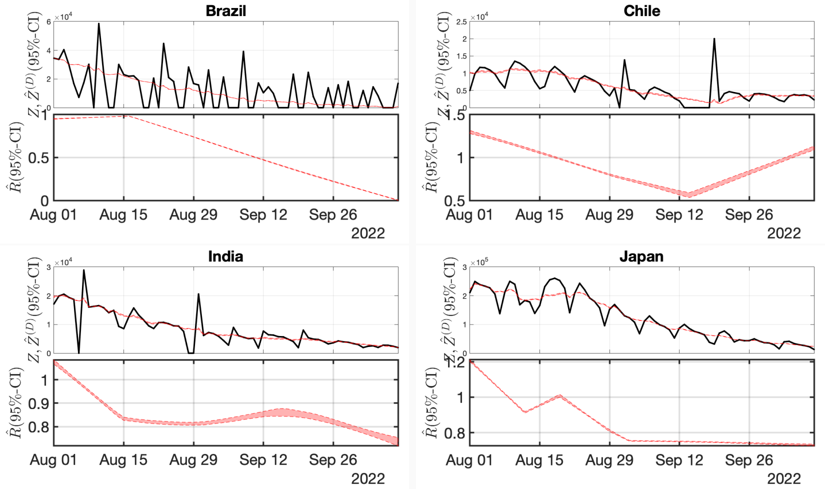


Figure 5: For different countries and different time periods. For each country, observed counts $\mathbf{Z}$ (black solid lines) and 95% credibility interval denoised counts $\mathbf{Z}^{(D)}$ (red pipe) [top] ; 95% credibility interval estimates for R [Bottom].


France
Italy
Germany
US
$Z, \hat{Z}^{(D)}$(95%-CI)
$\hat{R}$(95%-CI)
Aug 01
Aug 15
Aug 29
Sep 12
Sep 26
2022


Brazil
Chile
India
Japan
$Z, \hat{Z}^{(D)}$(95%-CI)
$\hat{R}$(95%-CI)
Aug 01
Aug 15
Aug 29
Sep 12
Sep 26
2022

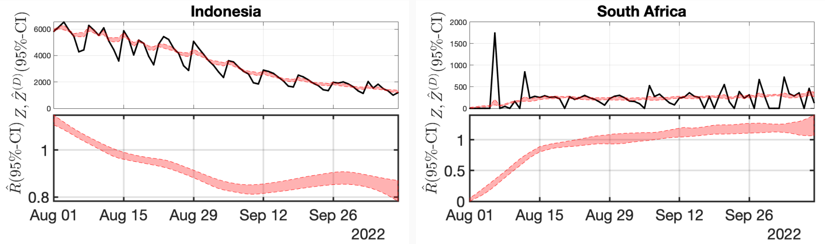
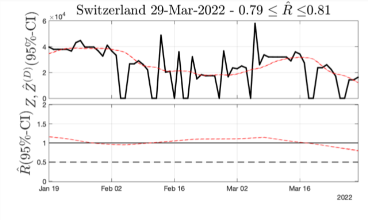
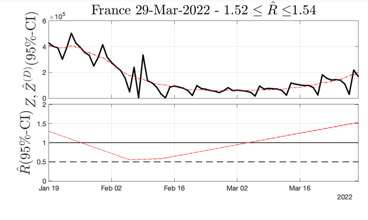
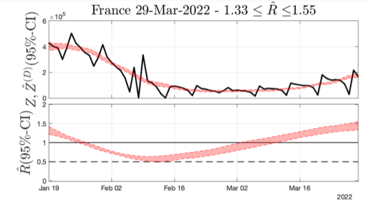


Figure 6: Top: Observed counts (black) and 95% CIs estimations of actual counts of new infections (red); Bottom: 95% CIs estimations of the R_t's.


Switzerland 29-Mar-2022 - 0.79 ≤ $\hat{R}$ ≤0.81
$\hat{R}$(95%-CI) Z, $\hat{Z}^{(D)}$(95%-CI)
×10^4
Jan 19
Feb 02
Feb 16
Mar 02
Mar 16
2022


France 29-Mar-2022 - 1.52 ≤ $\hat{R}$ ≤1.54
$\hat{R}$(95%-CI) Z, $\hat{Z}^{(D)}$(95%-CI)
×10^5
Jan 19
Feb 02
Feb 16
Mar 02
Mar 16
2022


France 29-Mar-2022 - 1.33 ≤ $\hat{R}$ ≤1.55
$\hat{R}$(95%-CI) Z, $\hat{Z}^{(D)}$(95%-CI)
×10^5
Jan 19
Feb 02
Feb 16
Mar 02
Mar 16
2022