

Estimation of the effective reproduction number in Madagascar

Our objective is to determine the spread of the coronavirus disease (Covid-19) from the number of daily infected cases. We use a model which considers the randomness of infections as a renewal process, in close connection to the Model 2 used in [2].

1/ The reproduction number

Let $I(t-k)$ be the number of daily infected people at time $t - k$, ($t - k < t$, t : date, k : duration), then we can assume that a fraction of the same people at time $t - k$, of size $I(t-k)f(t,k)$, ($0 \leq f(t,k) \leq 1$), will transmit the virus to some individuals at time t . When we consider all cases before the time t (i.e. for all $k \in [1, t-1]$,

then $I_{inf}(t) = \sum_{k=1}^{t-1} I(t-k)f(t,k)$ will be the number of all people who contaminated those infected at time t .

For this setting, we can define the effective reproduction number $R_t = R(t)$ at time t [2] by:

$$R(t) = \frac{I(t)}{I_{inf}(t)} \quad (1)$$

Therefore R_t is the average number of newly infected people at time t and contaminated by one person.

Remarks:

For $R(t) > 1$, the number of infection is fastly growing.

The coefficient $f(t,k)$ is related to the infectiousness profile and is only estimated; in some models [2, 4, 5], this term depends only on the duration k and not explicitly on the time t . It implies that the fraction of those who have transmitted the virus is completely determined when the duration k is known.

Because of the limitations to detect all chains of transmission among the infected population, the way of estimating $f(t,k)$ is by means of data about the serial interval, i.e. the time between the symptom onsets for the infected and his or her transmitter. From such data, one can deduce the serial interval distribution $w(k)$ of the duration k between symptom onsets. Then, one sets: $f(t,k) = w(k)$.

2/ Our model

For models based on the renewal process [8, 7, 5, 4, 2], the purpose is not to compute R_t directly from the equality 1. Following the methods described further, we rather estimate:

I(t) : in our model, $I(t)$ follows a Poisson distribution with mean equals to $R(t)I_{inf}(t) = R(t) \sum_{k=1}^{t-1} I(t-k)w(k)$.

When $R(t)$ and $w(k)$ are known, then one can estimate $I(t)$ using the Poisson distribution.

Rt : We use Bayesian inference to estimate R_t which follows a Gamma distribution in the renewal process model. It is realistic to assume that R_t does not change considerably over a time period $[t, t-k^*+1]$ for some $k^* > 1$. In this case, it is not difficult to show that the shape and the scale of the posterior

Gamma distribution of R_t are respectively equal to $(a + \sum_{k=1}^{t-1} I(t-k))$ and $(1/b + \sum_{s=t-k^*+1}^{t-1} I_{inf}(s))$

where a and b are respectively the shape and the scale of the prior Gamma distribution of R_t . (See Methods).

w(k) : In our model, the serial interval distribution $w(k)$ is taken to follow a Gamma distribution with 6.48 days as mean and 3.83 days as standard deviation [2].

2/ Estimation of $R(t=6/06/20)$: See Figure 1. Each boxplot gives the minimum and maximum values, the 1st (25%), 2nd (median) and 3rd (75%) quartiles of the posterior Gamma distribution of R_t . For the numerical values, we give the mean, the 25% and 75% quartiles.

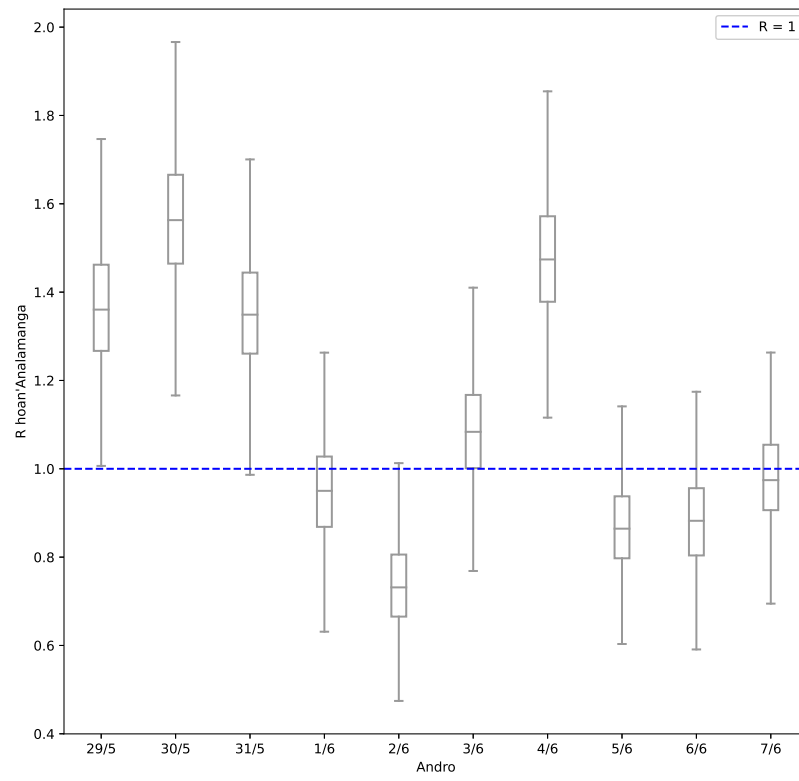


Figure 1: Data taken from 3/05/20 to 7/06/20. $R(t=6/06) = 0.88220$ (75%-quartile: 0,95091 ; 25%-quartile: 0,80891) for a and b such that the prior and the posterior distribution of R_t are equal.

Discussions and limitations of the model:

1/ The estimation of R_t depends on the prior parameters a and b . For $t = 6/06/20$, this dependence is shown by the figure 2. The red region corresponds to values of R_t which does not change significantly with rapport to a and b . In the blue region, a and b are both usually less than 1, and R_t is lower than its values in the red region.

For example, $R(t=6/06) = 0,78333$ (75%-quartile: 0,84466 ; 25%-quartile: 0,71788) for $a = b = 0,1$. And $R(t=6/06) = 0,94372$ (75%-quartile: 1,01521 ; 25%-quartile: 0,86759) for $a = b = 5$.

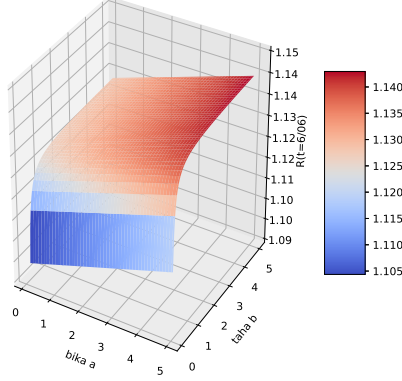


Figure 2: Variation of $R(t=6/06/20)$ according to parameters a and b .

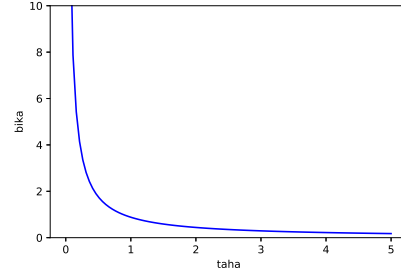


Figure 3: Curve of points (a, b) for which the prior and the posterior distributions of R_t are equal when $t = 6/06/20$.

2/ It is better to choose values for a and b such that the prior and the posterior distributions of R_t are equal. Indeed, both probabilities can be used to estimate R_t . In such cases, the values (a, b) which equalize the two distributions are situated on a particular curve in the space of possible parameter values.

3/ The time-windows $[t, t-k^*-1]$ are chosen for a value of k^* such that 1) R_t is globally stable during the time period and 2) the posterior density probability $\Pr(\text{daily infected} = I(t), \dots, I(t-k^*-1) | R_t)$ given R_t is maximum so that the model fits closely to the data [4, 2].

5/ The estimation of R_t depends on the duration of the period associated to the collected data. This duration have to be taken into account when one computes the optimal duration k^* . In fact, it should be largely greater than k^* .

6/ The estimation of R_t depends on the the optimal value k^* . However, one can compare the estimated R_t with the one obtained directly from the equality 1.

7/ Because we have taken $w(k)$ with mean equal to 6.5 days, i.e. because we admit that there are about 6.5 days between the symptom onsets for the infected and his or her transmitter, our model cannot give an estimation of R_t for the first 7 days of the outbreak.

Weekly evaluation:

The numbers of daily infections, collected during a week W , can be analyzed so that one can estimate a global reproduction number associated to the week W . In our model, this weekly reproduction is chosen so that for each $t \in W$, the average value of $I(t)$, given by the product $R(t) \sum_{k=1}^{t-1} I(t-k)w(k)$, is close to the

observed data.

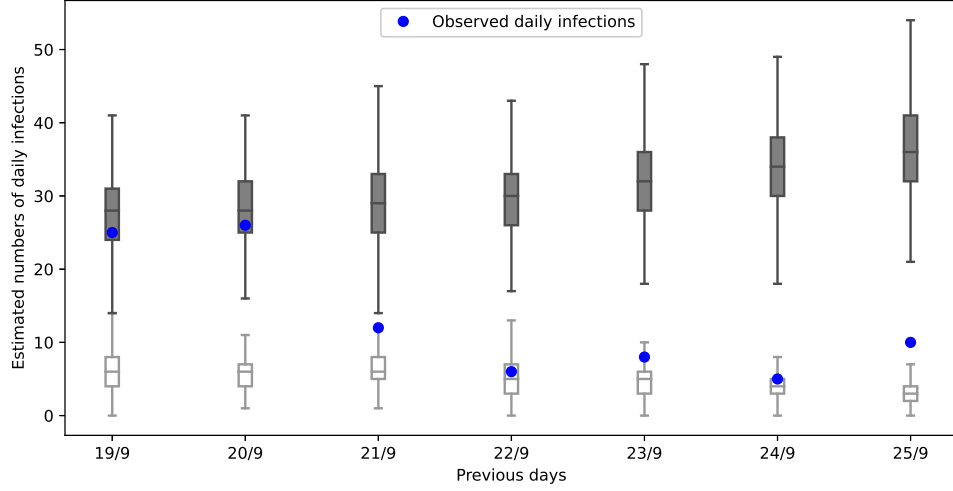


Figure 4: Week evaluation for the period of 19-25/09/20 in **Analamanga**. Transparent boxplots correspond to estimated average values of $I(t)$ when the reproductive number is constantly equal to $R(14/09) = 0,33014$ (75%-kartil: 0,37450 ; 25%-kartil: 0,28034) during the period of 19-25/09/20.

Colored boxplots correspond to estimated average values of $I(t)$ when the reproductive number is constantly equal to $R(19/09) = 1,40746$ (75%-kartil: 1,52539 ; 25%-kartil: 1,28102) during the period of 19-25/09/20.

4/ Methods

A renewal process describes random events K_1, K_2, \dots which occur independently between different random time intervals, k_1, k_2, \dots . These events (eg. time intervals between consecutive infections) are independently and identically distributed probability distribution.

For a renewal process K_1, K_2, \dots following an exponential distribution with rate r , one can prove that:

1/ the random variable $TK_i = \sum_{j=1}^i K_j$, ($i = 1, 2, \dots$), which gives the time of apparition of the event K_i ($i = 1, 2, \dots$), follows a Gamma distribution with the shape i and rate r ;

2/ the random variable $IK(k)$ which computes the number of all events K_1, K_2, \dots happening before the time k , follows a Poisson distribution with mean rk .

Ny fitsipika fanombanana Bayesianana dia mikendry ny hanombana taha iray miainga amin'ny antontan'isa ananana. Izany fanombanana izany dia mifototra amin'ny famaritana ny herijika manaraka fepetra (conditional probability):

$$\begin{aligned} Pr.pdf((antontan'isa = I) \cap (taha = r)) &= Pr.pdf(antontan'isa = I | taha = r) Pr.pdf(taha = r), \\ &= Pr.pdf(taha = r | antontan'isa = I) Pr.pdf(antontan'isa = I). \end{aligned}$$

$Pr.pdf(antontan'isa = I | taha = r)$ eto dia herijika fikitrohana (probability density) ahazoana antontan'isa = I raha toa ka manana taha = r .

$Pr.pdf(taha = r | antontan'isa = I)$ dia herijika fikitrohana ahazoana taha = r raha toa ka manana antontan'isa = I . Io herijika fikitrohana io no ikajiana ny tombana hoan'ilay taha miainga amin'ny antontan'isa ananana. Amin'ny modely RP [5, 4, 2], $Pr.pdf(taha = R(t))$ no raisina ho herijika fikitrohana mialoha (prior probability density) hoan'ny taha $R(t)$ ary $Pr.pdf(taha = R(t) | antontan'isa = I)$ no herijika fikitrohana afara (posterior).

Rehefa mitetika (discrete) ilay refy dia misolo ho herijika (Pr) ilay $Pr.pdf$.

Rehefa $\Pr.pdf(\text{antontan'isa} = \text{Ivao} | \text{taha} = r \times w) = \text{Poisson}(\text{antontan'isa} = \text{Ivao}, \text{taha} = r \times w)$ ary $\Pr.pdf(\text{taha} = r) = \text{Gamma}.pdf(r, a, b)$ dia [5, 4]:

$$\Pr.pdf(\text{taha} = r | \text{antontan'isa} = \text{Ivao}) = \text{Gamma}.pdf(r, a + I, 1/(1/b + w)).$$

Mahasedra, 27/09/2020.

References

- [1] Centre de Commandement Opérationnel-Covid-19 Madagascar. *Data ilaina*. May 2020.
- [2] Imperial College COVID-19 response team. *Short-term forecasts of Covid-19 deaths in multiple countries*. COVID-19 planning tools. May 2020. <https://mrc-ide.github.io/covid19-short-term-forecasts/index.html>.
- [3] Imperial College COVID-19 response team. *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf>.
- [4] Kris V. Parag sy Chirstl A. Donnelly. *Optimising Renewal Models for Real-Time Epidemic Prediction and Estimation*. November 2019. Tranokala bioRxiv. <https://www.biorxiv.org/content/10.1101/835181v1.full.pdf>.
- [5] Anne Cori, Neil M. Ferguson, Christophe Fraser, ary Simon Cauchemez. *A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics: Supplementary Data*. American Journal of Epidemiology, Septambra 2013. <https://academic.oup.com/aje/article/178/9/1505/89262>.
- [6] S. Cauchemez et al. *Real-time Estimates in Early Detection of SARS*. Emerging Infectious Diseases Journal, Janoary 2006. https://wwwnc.cdc.gov/eid/article/12/1/05-0593_article.
- [7] Christophe Fraser. *Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic*. PLoS One Journal, August 2007. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950082/>.
- [8] J Wallinga sy M. Lipsitch. *How Generation Intervals Shape the Relationship Between Growth Rates and Reproductive Numbers*. Proceedings of the Royal Society B: Biological Sciences, Novambra 2006. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1766383/>.
- [9] Kyle Siegrist et al. *Renewal process*. Tranokala Random: Probability, Mathematical statistics, Stochastic processes. <https://www.randomservices.org/random/renewal/index.html>. Jolay 2007. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950082/>.
- [10] H. Mahasedra Ratsimbarison. *Tombana momba ny taha-pifindran'ny Covid-19 ao Madagasikara*. Personal notes. Aprily 2020.