Covid-19 Belgium: Extended SEIR-QD model with nursing homes and long-term scenarios-based forecasts

Nicolas Franco

Namur Institute for Complex Systems (naXys) and Department of Mathematics, University of Namur, Namur, Belgium

Interuniversity Institute of Biostatistics and statistical Bioinformatics (I-BioStat) and Data Science Institute, University of Hasselt, Hasselt, Belgium

November 4, 2020

Abstract

We model the evolution of the covid-19 epidemic in Belgium with an age-structured extended SEIR-QD epidemic model with separated consideration for nursing homes. All parameters of the model are estimated using a MCMC method, except integrated data on social contacts. The model is calibrated on hospitals' data, number of deaths, nursing homes' tests and serological tests. We present the current situation in November 2020 as well as long-term scenarios-based forecasts concerning the second wave and subsequent lifting of measures.

Keywords: SARS-CoV-2, age-structured compartmental SEIR model, hospitalisation and mortality data, social contact patterns, Markov Chain Monte Carlo (MCMC)

1. Introduction

While there are many circulating different models concerning the covid-19 pandemic, it is important to have dedicated models to the specific situation of each country since the evolution of the situation as well as chosen political measures are different. SEIR-type epidemic models [1] are the most suitable for long term forecasting and especially SEIR-QD variants concerning the covid-19 pandemic [2, 3]. We present one of the very few existing extended SEIR-QD model adapted and calibrated on Belgium situation and data. Two similar approaches have been developed by the SIMID COVID-19 team (UHasselt-UAntwerp) [4] and the BIOMATH team (UGent) [5]. All of those models have their own characteristics and are complementary since it is difficult at this time to exactly know how to model the covid-19 in the best way. Another alternative approach has also been developed at the VUB [6] as well as a meta-population model from the SIMID COVID-19 team [7].

Here we summarise the main characteristics of our model. The Belgian population is divided into 8 compartments in order to take account of the different possible stages of the disease as well as the separation between asymptomatic and symptomatic people with a different infectiousness. Each compartment is divided into 5 age classes with different characteristics concerning the behaviour and evolution of the disease. The transmission of the coronavirus between all classes is computed using social contact data at different places (home, work, school, leisure) [8]. Except social contact data, all of the 65 parameters of the model are estimated using a Monte Carlo method, hence there is no assumption coming from studies in other countries. Nursing homes are modelled as isolated entities in order to take account of the different spread timing of the coronavirus compared to the general population. Specific parameters for the situation in nursing homes

Email address: nicolas.franco@unamur.be (Nicolas Franco)

take account of a variable hospitalisation policy based on hospitals load as well as a probability that deaths coming directly from nursing homes are related to the covid-19. There is a specific estimation of potential reimportations coming from travellers during the holiday period. The model is mainly calibrated using hospitalisations and deaths using both incidence and prevalence data (depending on which one is the more appropriate for the considered data) coming from Sciensano's public raw data [9]. Additional constraints on seroprevalence are coming from Sciensano's serological studies on blood donors as reported in Sciensano epidemiological reports [9]. The only positive PCR tests which are taken into consideration are those coming from nursing homes from an overall test campaign in April-May.

In Section 2, we present a description of the model. Technical details are presented in Subsection 2.2 for the equations, Subsection 2.3 for the data and Subsection 2.4 for explanation of the calibration method. Precise details on the timeline used and the full set of estimated parameters of the model are given in Appendix A. The results and discussion Section 3 starts with a presentation of the current estimation from the model in Subsection 3.1. Then we present a test on the validity of the model in Subsection 3.2, a middle-term forecast based on estimation of new policy mesures in Subsection 3.3 and finally some scenarios-based long-term forecasts in Subsection 3.4.

2. Materials and Methods

2.1. General description of the model

We consider the following age classes among the population: 0-24, 25-44, 45-64, 65-74 and 75+. Those classes correspond to public available data, excepted that 75-84 and 85+ classes are merged since it is difficult to know which parts of them are present in nursing homes or homes for the elderly. Hence we assume that the classes 0-24, 25-44, 45-64, 65-74 are only present among the general population, while the remaining is divided between a general 75+ and a specific class of nursing homes residents.

The compartmental model is divided into the following compartments: Susceptible S (people who have never been infected and are a priori susceptible to be infected), Exposed E (people who have just been infected but are without any symptom and still not infectious, hence in latent period), Asymptomatic Infectious AI (this is the part of the exposed people who fall into a continuously asymptomatic disease, which are infectious but with a reduced infectious probability due to their asymptomatic status and directly fall into the recovered status after that period), Presymptomatic Infectious PI (this is the other part of the exposed people who fall into a symptomatic disease, but symptoms do not appear directly, hence there is an intermediate stage where people become infectious but still without any symptom and with an infectious probability still reduced), Symptomatic Infectious SI (real disease period where the infectious probability is higher – people in this compartment will eventually fall either in a recovered status or will be hospitalised, and concerning nursing home, a significant part of them will directly die), Hospitalised Q (hospitalised people are considered as in quarantine for the model, since their contacts are almost inexistant), Deceased D (deaths from the general population are assumed only coming from hospitalised people – there is a small 1% of exceptions which is not taken into consideration here – however, deaths from nursing homes are taken into consideration and separated from deaths coming from hospitals), Recovered R (people who recovered from the disease, from asymptomatic ones, symptomatic ones or from the hospital, and are assumed here immune for the future). All those compartments exist for every age class. We do not consider in this model any subdivision inside the hospital compartment.

In addition, 2000 isolated nursing homes of similar average size are considered with all those compartments. The transmission of infection from the general population to those nursing homes is modelled by a random infection probability inside each nursing home, which is proportional to the number of infectious people and assumed less important since the lockdown period.

A schematic view of the compartments with their relations is presented in Figure 1.

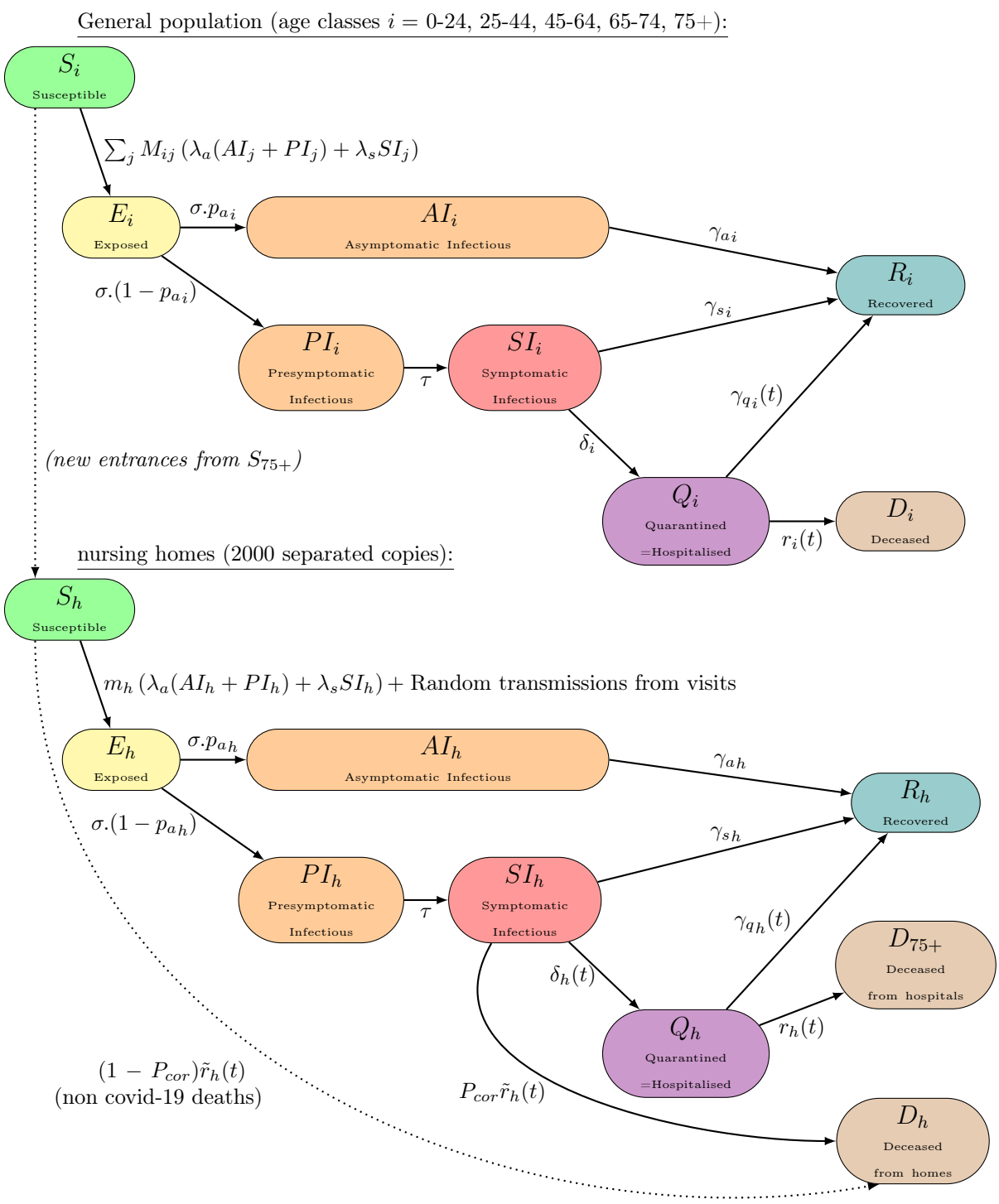


Figure 1: Schematic view of the compartmental model

Several additional or adaptive parameters are considered in order to better catch the reality. A specific probability parameter is catching the fact that only a part of deaths directly coming from nursing homes are due to the covid-19. A specific corrective coefficient is used to correct the new hospitalisations data since patients initially hospitalised for another reason are not officially considered in the data. Recovery and death rated from hospitals are considered variable in time in order to take the continuous improvement of care methods into account. A variable hospitalisation policy is considered for nursing homes during the first wave (period March-June) since residents are less likely to be hospitalised when the hospital load is important. This variable hospitalisation policy is not applied during the second wave in order to catch the reality on the ground.

Policy changes, according to Belgian epidemic' schedule, are monitored using different coefficients for the social contact matrices. Social contacts are divided into 4 categories: home (inside home and nearby family), work (with transport), school and leisure (with other places). All contacts are considered at 100% during the period up to March 14, 2020. Then reduced percentages are estimated by the model for the different periods of lockdown and phases of lift of measures. These reduced percentages are the effect at the same time of mobility restrictions, social distancing, prevention mesures, testing and contact tracing, while it is mathematically impossible to determine the exact part of those effects. Hence new parameters for some or all social contact types are estimated each time there is an important policy change. The timeline and the considered coefficients are described in Appendix A with the full list of estimated parameters. Long-term scenarios-based forecasts are constructed assuming a constant policy and compliance to measures during the future with different realistic possibilities of percentage of social contacts for still unknown policy effects.

This model takes into consideration potential reimportations of covid-19 from abroad during the holidays period. No reimportation is assumed in June since borders where barely opened. Reimportations are estimated during the period July to September using the following method: According to 2019 travel trends [10], we consider a proportionality of travellers of 36% in July, 26% in August and 21% in September. There is no data available concerning the inhomogeneous repartition inside each month, but we assume a homogeneous one for July and August while a 2 to 1 ratio between the first half of September and the second half. Only the top five countries of destination are considered with the following proportion: France 23%, Spain 11%, Italy 9% and The Netherlands 7% (Belgium is discarded). Then for each of those countries we consider the daily ECDC statistics on cumulative numbers for 14 days of covid-19 cases per 100000 [11]. The reimportation are added using an estimated global coefficient and injected proportionally in the exposed compartiment of the classes 0-24, 25-44, 45-64 and 65-74 and removed from the corresponding susceptible compartments. The estimated reimportations per day are presented in Appendix A.

This model does not take into consideration not officially observed effects like seasonality or cross-immunity. The population is only age-structured and not spatially structured. A spatial refinement of such a model would be really important, but currently the public data officially provided are not of sufficient detail in order to correctly fit a complex spatial-structured model.

2.2. Equations of the model

Equations of the model for the general population are the following ones, with i = 0-24, 25-44, 45-64, 65-74, 75+ depending on the age class:

$$\begin{aligned}
\frac{\mathrm{d}S_i}{\mathrm{d}t} &= -S_i \sum_j M_{ij} \frac{\lambda_a(AI_j + PI_j) + \lambda_s SI_j}{N_j} \\
\frac{\mathrm{d}E_i}{\mathrm{d}t} &= S_i \sum_j M_{ij} \frac{\lambda_a(AI_j + PI_j) + \lambda_s SI_j}{N_j} - \sigma E_i \\
\frac{\mathrm{d}AI_i}{\mathrm{d}t} &= \sigma p_{a_i} E_i - \gamma_{a_i} AI_i \\
\frac{\mathrm{d}PI_i}{\mathrm{d}t} &= \sigma \left(1 - p_{a_i}\right) E_i - \tau PI_i \\
\frac{\mathrm{d}SI_i}{\mathrm{d}t} &= \tau PI_i - \delta_i SI_i - \gamma_{s_i} SI_i \\
\frac{\mathrm{d}Q_i}{\mathrm{d}t} &= \delta_i SI_i - r_i(t) Q_i - \gamma_{q_i}(t) Q_i \\
\frac{\mathrm{d}D_i}{\mathrm{d}t} &= r_i(t) Q_i \\
\frac{\mathrm{d}R_i}{\mathrm{d}t} &= \gamma_{a_i} AI_i + \gamma_{s_i} SI_i + \gamma_{q_i}(t) Q_i
\end{aligned}$$

Parameters are explained in Table 1. Parameters without age class index i are assumed similar for all classes while those with age class index i are class-dependent. Some specific parameters are time-dependent.

Parameter	class-dep.	time-dep.	description
λ_a	no	no	transmission rate from asymptomatic infectious persons
λ_s	no	no	transmission rate from symptomatic infectious persons
σ	no	no	rate at which an exposed person becomes contagious
τ	no	no	rate at which an presymptomatic person becomes symptomatic
M_{ij}	yes	no	social contact matrices (WAIFW)
p_{a_i}	yes	no	probability of a completely asymptomatic disease
δ_i	yes	no	rate at which a symptomatic person develops heavy symptoms and is hospitalised
γ_{a_i}	yes	no	rate at which a person recovers from asymptomatic disease
γ_{s_i}	yes	no	rate at which a person recovers from light symptomatic disease
$\gamma_{q_i}(t)$	yes	yes	rate at which a person recovers from hospital
$r_i(t)$	yes	yes	rate at which a person dies from hospital

Table 1: Main parameters of the model

The transmission is governed by the so-called social contact hypothesis [12]. Social contact matrices M_{ij} (WAIFW) are collected using the SOCRATES online tool [8] and the 2010 dataset [13]. Work and transport categories are merged as well as leisure and otherplace. 4 different parameters which are adapted depending on lockdown/policy mesures are used as coefficients, hence the complete contact matrices are:

$$M_{ij} = C_{\text{home}} M_{ij\,\text{home}} + C_{\text{work}} M_{ij\,\text{work}} + C_{\text{school}} M_{ij\,\text{school}} + C_{\text{leisure}} M_{ij\,\text{leisure}}$$

During the holiday period (July-September), several elements are removed each day from the S_i classes and added to the E_i class according to estimated travellers and estimated average infection in the visited countries with a global coefficient C_{reimp} (as explained in Section 2.1).

The time-dependence of some parameters is computed using a logistic function with estimated parameters P_{recovery}, μ_{recovery} and s_{recovery}:

$$\gamma_{q_i}(t) = \gamma_{q_i} \left(1 + \frac{P_{\text{recovery}}}{1 + e^{-\frac{t-\mu_{\text{recovery}}}{s_{\text{recovery}}}}}\right) \qquad r_i(t) = r_i \left(1 - \frac{P_{\text{recovery}}}{1 + e^{-\frac{t-\mu_{\text{recovery}}}{s_{\text{recovery}}}}}\right)$$

Constants of the model are $N_{0-24} = 3250000$, $N_{25-44} = 3000000$, $N_{45-64} = 3080000$, $N_{65-74} = 1150000$ and $N_{75+} = 870000$ outside nursing homes (with an additional $N_h = 150000$ inside nursing homes) for a total population of $N = 11500000$. Those numbers are round numbers coming from the structure of the Belgian population as provided by the Belgian Federal Government on April 2020 [14]. An initial condition p_0 is proportionally distributed between the E_i on day 1 among the general population (corresponding to March 1 reported situation = February 29 real situation). Nursing homes are assumed not initially infected.

Equations of the model for the specific population in nursing homes follow a variation:

$$\begin{aligned}
\frac{\mathrm{d}S_h}{\mathrm{d}t} &= -S_h m_h \frac{\lambda_a(AI_h + PI_h) + \lambda_s SI_h}{75} - \tilde{r}_h(t)(1 - P_{cor})I_h \\
&\quad - \text{Random transmissions from visits} + \text{New entrances} \\
\frac{\mathrm{d}E_h}{\mathrm{d}t} &= -\frac{\mathrm{d}S_h}{\mathrm{d}t} - \sigma E_h \\
\frac{\mathrm{d}AI_h}{\mathrm{d}t} &= \sigma p_{ah} E_h - \gamma_{ah} AI_h \\
\frac{\mathrm{d}PI_h}{\mathrm{d}t} &= \sigma\left(1 - p_{ah}\right) E_h - \tau PI_h \\
\frac{\mathrm{d}SI_h}{\mathrm{d}t} &= \tau PI_h - \delta_h SI_h - \gamma_{sh} SI_h - \tilde{r}_h(t) P_{cor} SI_h \\
\frac{\mathrm{d}Q_h}{\mathrm{d}t} &= \delta_h SI_h - r_h(t) Q_h - \gamma_{qh}(t) Q_h \\
\frac{\mathrm{d}D_{75+}}{\mathrm{d}t} &+= r_h(t) Q_h \\
\frac{\mathrm{d}D_h}{\mathrm{d}t} &= \tilde{r}_h(t) Q_h \\
\frac{\mathrm{d}R_h}{\mathrm{d}t} &= \gamma_{ah} AI_h + \gamma_{sh} SI_h + \gamma_{qh}(t) Q_h
\end{aligned}$$

Most of the parameters are similar to the general population (but assumed with different values) with some additional considerations. There are 2000 nursing homes considered as separated entities, with a constant population of 75 inside each one, for a total of $N_h = 150000$ residents. New entrances are considered in order to fit the empty places up to 75 residents per nursing home and are removed from the 75+ susceptible class. Transmissions inside a specific nursing home follow usual SEIR-type transmission with a specific coefficient m_h. Transmissions from the general population is computed in a particular way using a daily probability of infection, i.e. each day one additional (integer) infected resident is added with probability $P_{th}\bar{S}_h \sum_j \frac{\lambda_a(AI_j+PI_j)+\lambda_s SI_j}{N}$, where the coefficient is distinguished between the initial phase P_{th} and lockdown and subsequent phases P'_{th}. $\bar{S}_h$ means the number of susceptibles within the specific nursing home. Starting from lockdown, transmissions are only considered from the 25-65 population (i.e. with $j = 25 - 44$ and $45 - 64$) since transmissions are mainly from nursing homes' workers. Potential reverse transmissions are however not monitored here. Deaths from care centres through hospitalisations are counted within the 75+ class. Additional deaths from care homes are monitored using a death rate $\tilde{r}_h$ with a coefficient P_{cor} which captures the probability that the death is covid-19 related. Remaining non-covid-19 related deaths are assumed occurring in the susceptible class (or in the recovered class if the first one is empty). For the first wave only (March 1 to June 30) a variable hospitalisation policy is computed in order to correspond to the reality and using variable parameters of constant sum $\delta_h(t) + P_{cor}\tilde{r}_h(t) = \delta_h$, the proportion being monitored over time by a logistic function depending on hospitals load with an additional delay, with estimated parameters P_{delay}, μ_{hosp} and s_{hosp}:

$$\delta_h(t) = \delta_h - \frac{\tilde{r}_h P_{\text{cor}}}{1 + e^{-\frac{Q(t-\text{delay})-\mu_{\text{hosp}}}{s_{\text{hosp}}}}} \qquad \tilde{r}_h(t) = \frac{\tilde{r}_h}{1 + e^{-\frac{Q(t-\text{delay})-\mu_{\text{hosp}}}{s_{\text{hosp}}}}}$$

This variable hospitalisation policy is nonexistent for the second wave since most of nursing home residents are hospitalised during this period. Hence from July 1, those parameters are considered with the value $Q = 0$.

The basic reproduction number is estimated by the leading eigenvalue of the next-generation matrix [15, 16] (the eigenvalue is real since this matrix is positive definite):

$$R_0 = \text{maxeigenv}\left[\lambda_a\left(\frac{p_{aj}}{\gamma_{aj}} + \frac{1-p_{aj}}{\tau}\right)M_{ij} + \lambda_s\left(\frac{1-p_{aj}}{\gamma_{sj}+\delta_j}\right)M_{ij}\right]_{ij}$$

The effective reproduction number $R_t = R_e$ is estimated by $R_t = R_0 \frac{\sum_i S_i(t)}{N - \sum_i D_i(t)}$.

2.3. Considered data

We consider the following data for the calibration of the model coming from Sciensano's public raw data (October 31, 2020 release):

- New hospitalisations (incidence) with an additional corrective estimated parameter SUPP_{hosp} which estimates the percentage of missing covid-19 patients at the time of admission (hence catching supplementary patients not initially hospitalised for covid-19 and transferred from other units)
- Hospital load (prevalence)
- Released from hospital (cumulative)
- Total deaths from hospital (cumulative)
- Deaths (incidence) from age classes 45-64, 65-74, 75+ (incidence data are more suitable for those classes since there are a percentage of deaths for which the age class is unknown and prevalence data could contain an accumulation of errors)
- Deaths (cumulative) from all age classes (cumulative data are more suitable for the 0-24, 25-44 classes since incidence data are almost zero)
- Deaths from nursing homes (incidence and cumulative)
- Total deaths (incidence and cumulative)

Deaths reported with a specific date are considered on that specific date while situations reported by hospitals are considered to occur up to 24h before the hospital report hence 2 days before the official data communication. Note that graphics are plotted using the dates of Sciensano's communications (1 day delay).

Additional constraints are considered coming from Sciensano's epidemiological reports [9] (those constraints determine the set of admissibles parameters). Serological studies on blood donors are considered in the following way: the ratio between immune people (for the classes 25-44, 45-64 and 65-74) coming directly from the asymptomatic compartment ($\sum_i AI \to R$) and the total asymptomatic population who has not developed a symptomatic covid-19 disease ($\sum_i S + E + AI + [IA \to R] + PI$) should be respectively between 0.5% and 2.8% 7 days before March 30 and between 3.5% and 6.2% 7 days before April 14, April 27 and May 11 (the 7 day delay is here to take the needed time to build a detectable immunity into account). There are also constraints on parameters as e.g. $p_{a0-24} > p_{a25-44} > \dots$ in order to reproduce the more severity of the covid-19 on older persons as well as trivial constraints to avoid negative or out-of-bound parameters.

Additional constraints are imposed on nursing homes coming from the result of massive PCR test on April-May: the average percentage of infected people should be $8\% \pm 3\%$ during the period April 15-30 and less than $2\% \pm 2\%$ during the period May 15-31. Those percentages are estimated from Sciensano's epidemiological reports using a calculated incidence between each week. Additionally, the average percentage of asymptomatic residents (including presymptomatic ones) among infected should be $75\% \pm 10\%$.

2.4. Calibration method

All parameters are estimated using a MCMC Metropolis–Hastings algorithm [17]. Two different modes are used for the calibration and the statistical analysis:

- Best-fit burning mode: an optimised First-choice hill climbing algorithm using weighted least-squares performed on one parameter at a time (i.e. one neighbour = variation of one parameter), with downhill moves allowed up to 0.2% in order to avoid local optima and with a quick best fit search performed on accepted neighbours
- MCMC mode: a Metropolis–Hastings algorithm performed on all parameters (i.e. one neighbour = variation of all parameters) where the likelihood fonction is constructed using the property that the empirical variance from n data follows up to a coefficient a χ^2 distribution (we use $\frac{nS^2}{\sigma^2} \sim \chi^2(n) \cong \mathcal{N}(n, 2n)$ with σ estimated from the burning period)

Weights for least-squares are defined in the following way: For each set of data (and each age class), the weight is chosen such that a best fit search considering only those data (hence a best fit discarding all other data) gives similar empirical variance. Hence each sort of data gives a similar contribution to the likelihood. For data with both incidence and prevalence numbers considered, the contribution of incidence numbers is favoured at ∼80%. From this method, we obtained the following weights: hospitalisations prevalence 1, released 1, deaths from hospital 10, deaths 0-24 5000, deaths 25-44 1000, deaths 45-64 2000 incidence 2 prevalence, deaths 65-74 500 incidence 1 prevalence, deaths 75+ 200 incidence 0.5 prevalence, deaths from nursing homes 100 incidence 0.2 prevalence, total deaths 100 incidence 0.2 prevalence. New hospitalisations data (indicence) have a specific period-dependent weight since the noise is very different between wave periods and low hospitalisations periods: 2 first wave period (March 1-June 7), 100 intermediate period (June 8 - August 31), 10 second wave period (September 1-. . .).

The program is written is C language. The full ODEs are solved by numerical integration using the GNU gsl odeiv2 librairy and a Runge-Kutta-Fehlberg45 integrator. The computation is performed on the HPC cluster Hercules2 [1].

The method of approaching global minima is particular due to the presence of a very high number of estimated parameters, a large number of numerical integrations to be performed (due to the presence of 2000 separated nursing homes) and the fact that the algorithm should not take several days to complete in order to quickly produce previsions of the epidemic. Hence the method is separated into different steps. In a first time, a set of best-fit search is performed from a very large distribution (100 times usual standard deviation) using 5000000 iterations with a 10 times average step and a special trick to increase the rapidity of the algorithm: instead of 2000 different nursing homes, only 100 nursing homes are considered with each time 20 copies of each. This approximation is suitable as long as the algorithm is still far from the best-fit. In a second time, the best-fit search is pursued for 20000 iterations using the complete 2000 different nursing homes in order to affine parameters. A set of at least 250 different priors is collected from the best obtained results. In a third time, the MCMC algorithm is performed from the 250 priors, with 50000 iterations retaining every 5000 iterations. This produces a set of 2500 samples coming from potentially different local minima zones which avoids a too high autocorrelation of the results. Since the first step is very time-consuming, further runs of the model are performed without a complete recalibration, by reusing the previous 2500 samples as a set from which 250 random priors are taken and running a 20000 burning best-fit period before the MCMC step. This is a kind of data assimilation process. Recalibration should however be performed each time there is a major change in the model or in the policy.

[1] "Plateforme Technologique de calcul Intensif" (PTCI) located at the University of Namur, Belgium.

3. Results and Discussion

3.1. Current estimations

We present in this Section the result of the calibration of the model as on November 1, 2020, with considered data up to October 31, 2020. Results are presented in the figures with medians, 5% and 95% percentiles, hence with a 90% confidence interval.

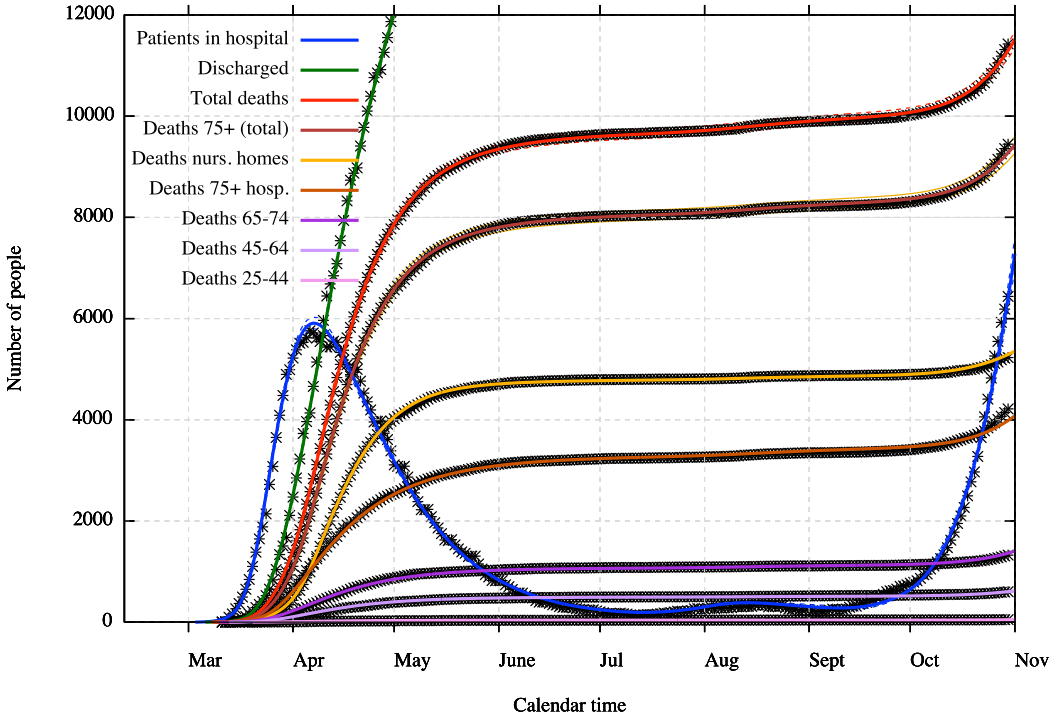


Figure 2: General view on prevalence data and estimations

In Figure 2, we have a general representation of the evolution of the epidemic in Belgium with hospitalisations, people discharged from hospitals and deaths coming from hospitals and from nursing homes, all in prevalence or cumulative numbers. The interest in modelling the epidemic within nursing homes separately from the general population can clearly be seen on this figure. Indeed, the form of the death curve for nursing homes is really different from the ones for the general population since the epidemic started a few time later in nursing homes but took a bigger proportion. Note that the very small percentage of deaths occurring at another places is not taken into consideration in this model.

The model calibration fits the real prevalence data with a very good exactitude (excluding of course data noises) despite that fact that the calibration is mostly done on incidence data. The comparison between the model and some incidence data are presented in Figure 3 for the general incidence data in hospitalisations and deaths and in Figure 4 for incidence data in deaths with age class repartition. The consideration of a continuous improvement of care was needed in order to correctly fit the death curves among the different periods. However, the figure 2 presents a slightly larger increase than expected for the deaths within the class 75+. This could be the result of a small decrease in the quality of care during the second wave due to the huge load of the hospitals (but still better than during the first wave).

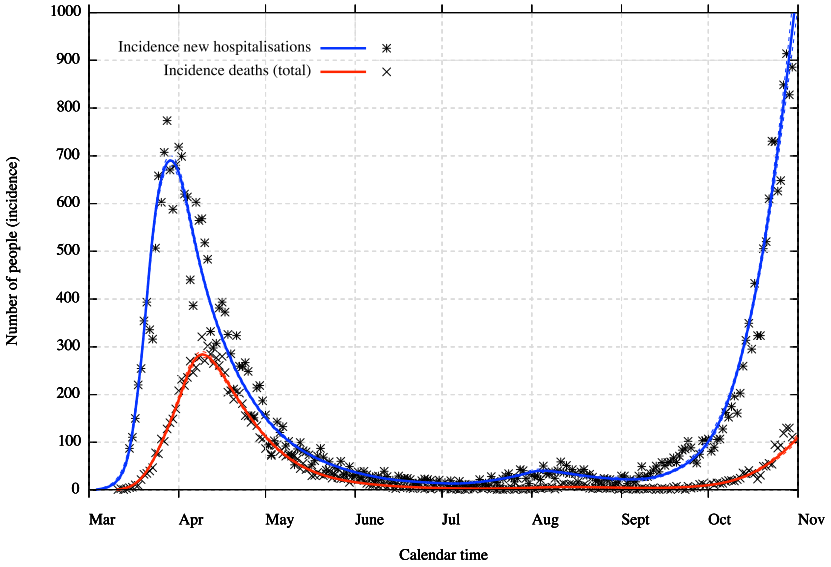


Figure 3: Incidence in new hospitalisations (with underreporting correction) and deaths

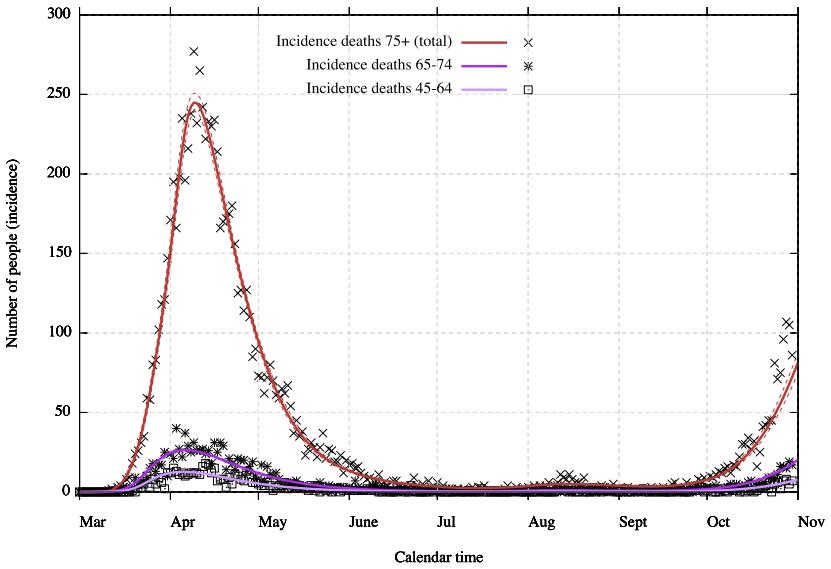


Figure 4: Incidence deaths within each age class

We must remark that in order to obtain a correct fit we needed to introduce specific elements within the model. Incidence data on new hospitalisations are underroported. Indeed, somme patients are initially admitted for another reason than covid-19 and only transferred in a covid-19 section afterwards. This results in more people going out from hospitals than officially entering. The model estimates those supplementary patients at 23.1% [16.4% ; 28.7%]. Real data in Figure 3 are plotted with the estimated correction.

Deaths coming directly from nursing homes are not all due to covid-19 since many PCR tests are lacking. The model estimates that only 83.6% [66.7% ; 89.4%] of those deaths are really due to covid-19. The ratio between deaths coming directly from nursing homes and deceased patients in hospitals coming initially from nursing homes seems to be not constant, and it was necessary to introduce a variable hospitalisation policy. The best answer found was to monitor hospitalisations from nursing homes through a logistic function depending on general hospital load but with a specific delay. Hence, when hospital load starts to become too important, less people from nursing homes are hospitalised and the reverse effect occurs when hospital load gets lower, but each time with a delay estimated at 10.5 days [9.0 ; 12.4]. However, this variable hospitalisation policy turned out to be unsuitable for the second wave period since there a is better attention to residents in nursing homes.

Initially the model overestimated the number of deaths from the end of the first wave. It was not possible to calibrate constant death rates throughout all phases of the epidemic. This is the consequence of both a care improvement in hospitals and a lower aggressiveness of the virus. Hence death and recovered rates within each age class are also monitored by a logistic function depending on time. The current improvement (in comparison to the very beginning of the epidemic) is estimated as 62.6% [54.6% ; 67.1%], hence 62.6% of patients which should have died in March are now recovering from hospitals. We must remark that it is impossible to know which part is due to care improvement and which part is due to lower aggressiveness of the virus and that the death rate seems to restart becoming a bit higher in October.

The basic reproduction number R_0, representing the average number of cases directly generated by one infectious case in a population which is assumed totally susceptible, is estimated in average for each period (we consider this number dependent on lockdown measures) and computed as the leading eigenvalue of the next-generation matrix (cf. Section 2.2 for details). The effective reproduction number R_t (or R_e) represents the average number of cases directly generated by one infectious case taking account of the already immune population, hence varying over a period. The estimations are presented in Table 2.

	R_0	R_t (at the end of the period)
Pre-lockdown: March 1 → March 13	4.13 [3.89 ; 4.39]	4.08 [3.85 ; 4.34]
School and leisure closed: March 14 → March 18	2.24 [2.13 ; 2.35]	2.17 [2.07 ; 2.27]
Full lockdown: March 19 → May 3	0.65 [0.61 ; 0.72]	0.61 [0.57 ; 0.66]
Phase 1-2: May 4 → June 7	0.79 [0.75 ; 0.83]	0.73 [0.70 ; 0.76]
Phase 3: June 8 → June 30	0.99 [0.91 ; 1.07]	0.91 [0.84 ; 0.99]
Phase 4: July 1 → June 28	1.40 [1.29 ; 1.53]	1.29 [1.19 ; 1.40]
Phase 4bis: July 29 → August 31	0.75 [0.63 ; 0.88]	0.69 [0.58 ; 0.81]
Second wave: September 1 → October 31	1.73 [1.62 ; 1.85]	1.32 [1.25 ; 1.40]
Nursing home (average): March 1 → October 31	0.73 [0.06 ; 1.67]	0.41 [0.04 ; 0.99]

Table 2: Estimations of R_0 and R_t values

The reproduction number of the pre-lockdown period is a bit overestimated. This is probably due to the fact that the model does not take explicitly account of infections coming from abroad travellers at this particular time and this results in an estimated R_0 slightly above 4. For the period phase 1A-1B-2, since there were policy changes almost every weeks, we only provide here the estimated R_0 at the end of this period. The second wave R_0 does not take account of the new measure applied in October 19 whose effets should only be visible on November. We can see that the estimated R_0 inside nursing homes is very low. This seems logical since nursing homes with only one infected resident do not necessary end with a majority of infections. This implies however that the important spreading of the covid-19 inside some nursing homes is due in a great part to several importations, probably due to staff members. However, we must remark that the uncertainty concerning this reproduction number is very huge, probably the result of disparities among nursing homes.

The infection fatality rate (IFR) can be estimated using the total set of recovered people according to the model (hence including untested and asymptomatic people). Due to the consideration of variable death rates, the IFR is different from the early period of the epidemic and the last months. Estimations are presented in Table 3. The mean and last period are limited to September since October data need some consolidation regarding the number of deaths.

	General IFR	March-April period	July-September period
Overall population	0.96% [0.84% ; 1.09%]	1.08% [0.93% ; 1.24%]	0.29% [0.28% ; 0.31%]
0-24	0.00% [0.00% ; 0.00%]	0.00% [0.00% ; 0.00%]	0.00% [0.00% ; 0.00%]
25-44	0.01% [0.01% ; 0.01%]	0.02% [0.01% ; 0.02%]	0.01% [0.00% ; 0.01%]
45-64	0.19% [0.17% ; 0.21%]	0.21% [0.19% ; 0.23%]	0.08% [0.07% ; 0.09%]
65-74	1.72% [1.54% ; 1.95%]	1.85% [1.67% ; 2.09%]	0.86% [0.75% ; 0.99%]
75+ (nurs. homes included)	7.84% [6.85% ; 9.06%]	9.25% [8.01% ; 10.74%]	1.89% [1.69% ; 2.17%]

Table 3: Infection fatality rate estimations

Table 4 presents some estimations concerning some characteristics of the disease coming from the model. The set of asymtomatic people probably includes mild symptomatic people. The model cannot really detect the exact time when symptoms appear, hence the end of the incubation period merely corresponds to the estimated time when the infectiousness becomes more important. The total disease duration for symptomatic peoples is only for those who are not hospitalised. The hospitalisation duration is the average until discharged or deceased (no distinction is provided) at the beginning of the epidemic, hence before care improvement. The duration for asymptomatic nursing homes' residents cannot really be estimated by the model. Indeed, once a single nursing home (or an isolated part of it) is completely infected, asymptomatic infected residents can remain forever ill without any new possible contamination, hence there is no boundary on such duration coming from the available data, so this excessive duration must be considered as an outlier.

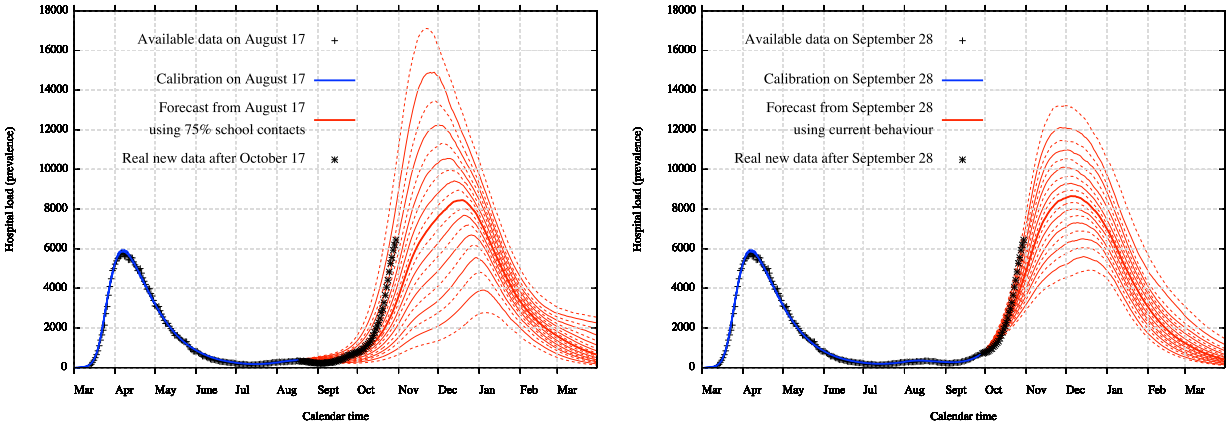
	0-24	25-44	45-64	65-74	75+	nursing homes
Latent (pre-infectious) period	1.3 days [1.1 ; 1.8]					
Presymptomatic period	6.8 days [4.9 ; 8.0]					
Total disease duration asymptomatic people	4.7 days [4.1 ; 5.5]	5.3 days [4.4 ; 5.9]	5.7 days [4.7 ; 6.4]	6.3 days [5.4 ; 7.5]	7.8 days [6.1 ; 9.9]	26.7 days [16.0 ; 58.7]
Total disease duration symptomatic people	11.1 days [9.6 ; 12.4]	11.6 days [10.2 ; 13.1]	12.2 days [10.8 ; 13.6]	12.7 days [11.4 ; 14.1]	13.2 day [11.7 ; 14.4]	13.8 days [12.1 ; 15.2]
Hospitalisation duration (before care improvement)	16.1 days [13.2 ; 18.9]	18.3 days [16.8 ; 19.9]	16.5 days [15.4 ; 17.5]	12.1 days [11.0 ; 13.3]	11.5 days [10.7 ; 12.4]	10.6 days [8.9 ; 11.5]
Overall percentage of asymptomatic people	93.9% [81.3 ; 99.0]	87.3% [68.9 ; 96.2]	74.6% [60.7 ; 82.2]	56.1% [42.6 ; 65.3]	34.2% [23.1 ; 49.9]	24.2% [12.1 ; 40.0]

Table 4: Some characteristics of the covid-19 as estimated by the model. All durations are average durations, and the given uncertainties are uncertainties on those averages, not on the individual values.

3.2. Confrontation of previous calibrations

One way to test the robustness of a model is to confront previous predictions with current reality. This model can provide forecasts in two different ways. When new policy interventions are expected or a specific behaviour change is planned due to the calendar, it is possible to extrapolate the future transmission of the covid-19 (monitored here by the number of contacts) using relative percentage of transmission in comparaison to the pre-lockdown phase. This percentage can only be a vague estimate of what could be the real transmission and it is sometimes suitable to look at several different scenarios. On the other hand, when no new policy intervention is expected for a certain time, it is possible to have a more precise forecast based on current estimated contacts (what we call the current behaviour), forecast which is only valid up to the next policy intervention.

We present two previous forecasts from the model. The first one, presented in Figure 5a is a 2 1/2 months old forecast based on the specific scenario that the transmission at school from September 1 would be at a level of 75% due to sanitary measures like masks wearing. The second one presented in Figure 5b is a 1 month old forecast based on the current behaviour and the estimation from the model of the number of contacts, which was estimated at that time to be 69.7% [44.2% ; 88.6%]. Those previous confrontation highlight the fact that the uncertainty must be taken into consideration for any forecast.



(a) Forecast from August 17

(b) Forecast from September 28

Figure 5: Previous forecast from August 17 based on the scenario of a 75% transmission at school from September 1 and from September 28 based on the continuation of current behaviour. The strong line represents the median, continuous lines represent deciles (10% percentiles) while dashed lines represent ventiles (5% percentiles).

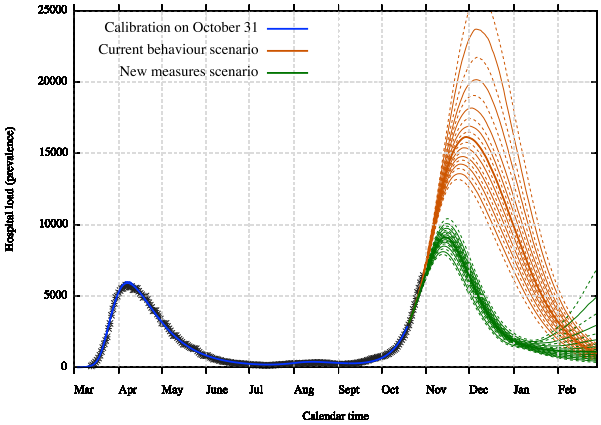
3.3. Middle-term scenario-based forecast

Every forecast is hypothetical. New measures that have not been tested cannot really be estimated on the level of their impact and it is impossible to predict evolution in compliance to them from the population as well as future policy changes. This is why any realistic forecast must rely on the assumption of a perfect continuity of measures and compliance for elements which are a priori not suspected to change soon and on different hypothetical scenarios for untested modifications of measures.

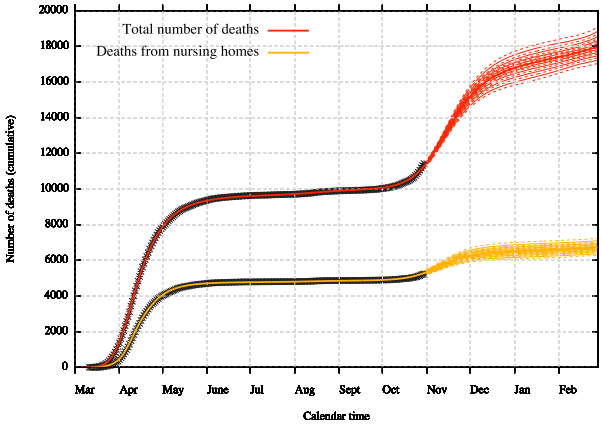
In response to the large second wave in Belgium, authorities have decided to enforce new measures on October 19 as closing bars and restaurants, reducing the allowed social contact (known as bubble) to one person, promoting teleworking and establishing a curfew during the night. On November 2, a soft lockdown is put into place, with closure of non-essential shops, teleworking mandatory, leisure mostly reduced and social contacts even more reduced. Schools are closed during 2 weeks and then reopen with a 5/6 attendance (except for universities).

While it is impossible to know with precision the impact from those measures, we can estimate that the effect from the soft lockdown could be comparable to the effect of the first lockdown, since the small remaining liberties could be balanced by generalised sanitary measures like mask wearing. The effects from October 19 measures is more incertain, but should be situated between September behaviour and lockdown behaviour. Hence the most realistic middle-term scenario is to consider a half reduction on contacts (work, leisure and family) from September situation in comparaison to the first lockdown on October 19, with a full reduction applied from November 2 until the December 13 planned deadline. Schools are considered at 0% transmission from November 2 to November 15 and at 5/6 thereafter. Every contacts are assumed to be restored at September level after December 13 (except for usual school closures).

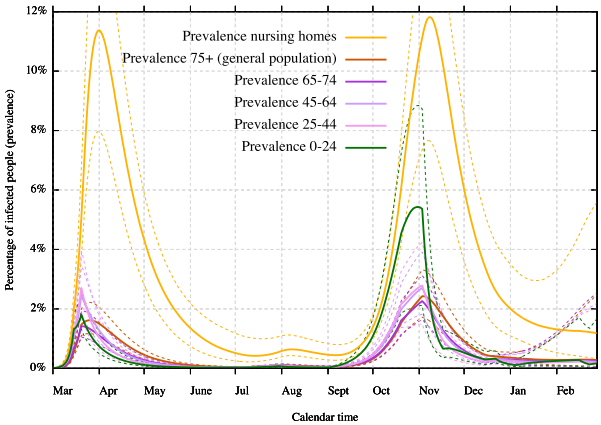
In Figure 6a, we present the estimated effect on hospital load from those measures. We must note that, according to those measures and to the model, the theoretical maximum capacity of 10000 hospital beds in Belgium should be almost reached but not exceeded, at least in an average national level. Figure 6b presents the expected mortality in case of the new measures scenario. We must remark that this expected mortality relies on a constant quality of care that may not be maintained.



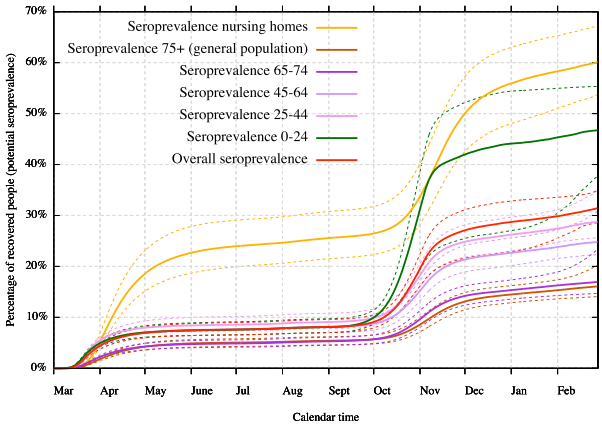
(a) Hospital load for two scenarios (with ventiles)



(b) Deaths forecast for new measures scenario (with ventiles)



(c) Estimated prevalence (confidence interval 90%)



(d) Potential seroprevalence (confidence interval 90%)

Figure 6: Middle-term scenario with potential effects from new measures applied on October 19 and November 2. The first figure presents a comparison of the hospital load with or without the effects from the new measures. The others figures present the previsions on mortality, prevalence and seroprevalence.

From the model, we can also extrapolate the evolution of the virus through the whole population over time. In Figure 6c, we present the estimated percentage of infected people over time for each age class. We can clearly see the effect of mid-March lockdown measures on children and working people. The effect of lockdown measures on older people (especially 75+) is less important since the curve is broken in a less effective manner. Concerning the second wave, we can see that the virus is really present among the very young population due to two complete months of school opening. This prevalence is completely shut down by the two weeks closure and should be brought at a lower level than other age classes.

In Figure 6d, we present the estimated percentage of recovered people, hence the estimated percentage of immunity acquired within each age class if we make the assumption that a constant immunity is granted to recovered people. Such a constant immunity is not guaranteed for the moment, but recent studies show that antibodies should be present after several months for a large majority of the population [18]. The seroprevalence is calibrated using blood donors tests results (around 1.3% on March 30 and 4.7% on April 14) [9]. Since those tests where only performed on an (almost) asymptomatic population which have not developed covid-19 symptoms from the past 4 weeks, the model extrapolates immunity coming also from the symptomatic population and from nursing homes. Note that we allow a 7 days delay in our model after recovering to be sure of the detectability of the anticorps. Table 5 presents the detail of some seroprevalence estimation.

	global immunity	among asymptomatic	inside nursing homes
March 30	2.68% [2.26% ; 3.01%]	2.39% [2.00% ; 2.66%]	1.49% [1.22% ; 1.93%]
April 14	5.45% [4.68% ; 6.18%]	4.58% [3.91% ; 4.93%]	9.51% [8.29% ; 11.21%]
October 31	18.49% [15.42% ; 21.16%]	10.37% [9.09% ; 10.90%]	32.94% [32.44% ; 34.17%]
January 1	28.76% [22.92% ; 32.89%]	18.95 % [16.91% ; 20.15%]	56.17% [54.41% ; 56.99%]

Table 5: Seroprevalence estimations

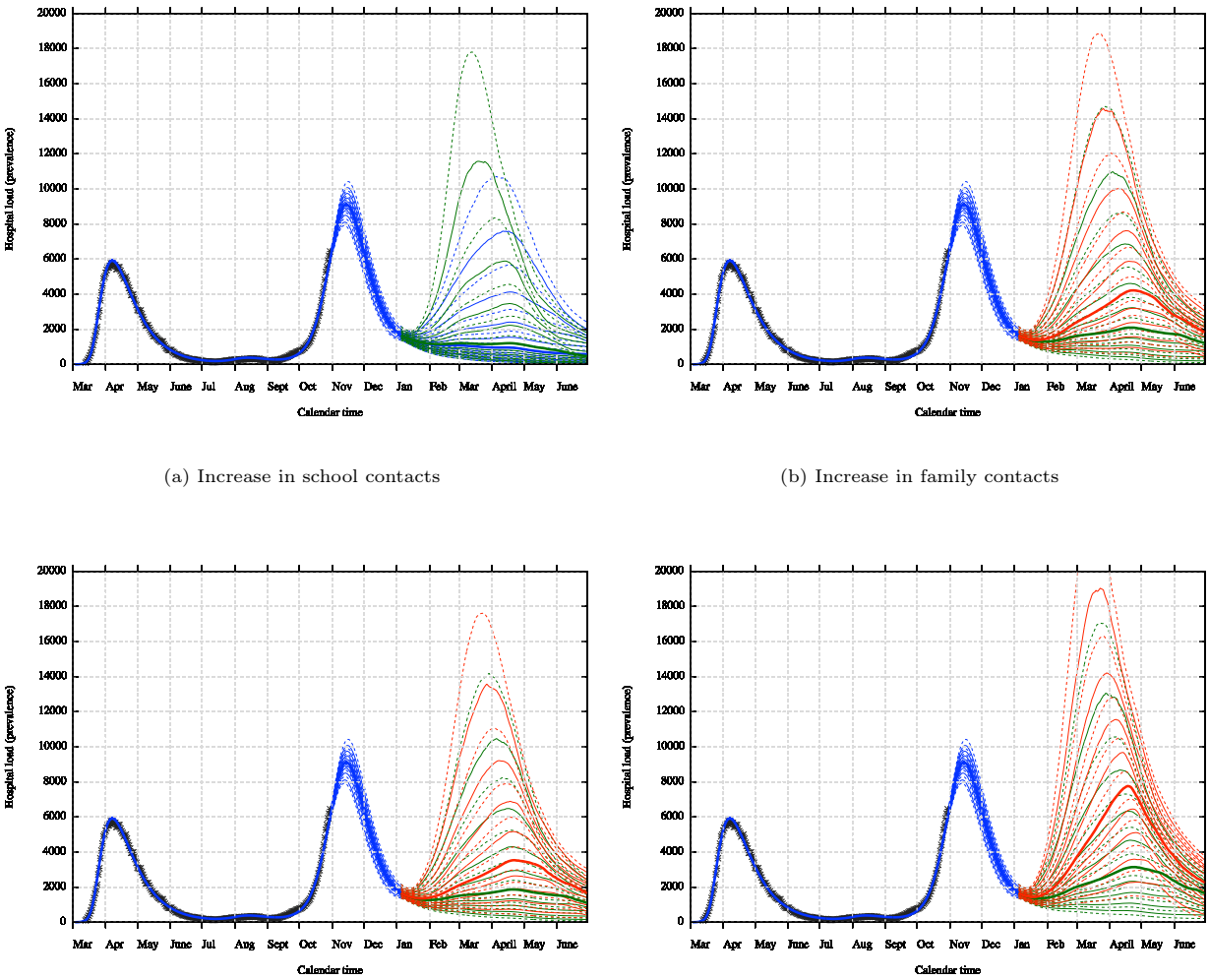
3.4. Long-term scenarios-based forecasts

The model allows to construct long-term scenarios which are very suitable to study the potential impact from a specific measure. The possibilities are numerous but we present in this section a simple study of the potential impact of an increase in contacts at a specific place (school, family, work and leisure). The increase is perform from January 4, 2021 up to June 30, 2021, when the risk of an emerging third wave is present. We work here with the assumption that there is no modification on the set of susceptible people except from natural infection, hence with the assumption that a lasting immunity is granted to recovered people. This hypothesis could be modified negatively in the future if the probability of a reinfection is important or positively is the immunity is artificially increased by the arrival of a vaccine.

The baseline scenario is the restart of all activities on January 4 with similar transmissions/contacts as in September. Those estimated contacts percentage are 91.6% [39.5 %; 99.3%] for school contacts, 51.5% [46.8 %; 54.6%] for family contacts, 8.9% [5.8 %; 13.6%] for work contacts and 30.7% [18.1 %; 56.7%] for leisure contacts. We remind here that those percentages do not correspond to the exact number of contact as determined by the attendance, but to the reduced transmission in comparison to the pre-lockdown period as the result of decrease of contacts but also of sanitary measures. This could explain while the transmission is estimated at a very low level at work since sanitary measures and social distancing are more respected than during leisures or among family. The high transmission percentage at school does not necessarily mean that schools are the engine of the virus transmission since most of the student are asymptomatic with a reduced infectiousness, and the uncertainty is still very important concerning this percentage.

The baseline scenario is presented in Figure 7a together with the potential impact of a full transmission at school, hence a transmission without any sanitary measure as well as without any quarantine imposed by the testing and tracing process. We can see that the baseline scenario itself provides a non-zero probability of a third wave but still particularly low. The full contacts at school scenario increases a bit this probability to a reasonable extent.

Increases in family contacts, work contacts and leisure contacts are presented in Figures 7b, 7c and 7d with each time a hypothetical increase of 10% or 20%. Those increase must be understood as a non-proportional increase (e.g. a work increase of 10% corresponds to $8.9\% + 10\% = 18.9\%$). We can clearly see that an increase in leisure contact has the most important effect on the evolution of the epidemic and could lead to a potentially problematic third wave. Full transmission scenarios for family, work or leisure cannot be taken as realistic since they would provide a complete explosion in the absence of vaccine.



(a) Increase in school contacts

(b) Increase in family contacts

(c) Increase in work contacts

(d) Increase in leisure contacts

Figure 7: Long-term scenarios with potential isolated effect from an increase in contacts at a specific place (with ventiles).

3.5. Conclusion

We have presented an age-structured SEIR-QD type model with a large number of ameliorations as a specific consideration for nursing homes, variable parameters and reimportation from travellers. Those ameliorations were important in order to catch the specificity of the epidemic in Belgium.

The model allows to have a good study of the current behaviour of the epidemic, with an estimation of hidden elements like the real prevalence of the virus and the potential evolution of the immunity. More important, the model allows to construct scenarios-based forecasts in order to estimate the potential impact from new policy measures and can explicitly serve, in complement of others models, to policy makers.

However, the model suffers from several limitations which would be important to try to solve in order to better catch the evolution of the epidemic. In particular, we can say that the lack of spatial consideration is only a huge approximation of the reality, even if the Belgian country is small and very connected. The compartmental distinction is limited to asymptomatic and symptomatic while there are several variations of the severity and hospitals are considered as a unique homogeneous element. Furthermore, the lack of

refinement inside age classes is a brake on the study of interesting scenarios, as e.g. studying the separated impact from transmission at primary school, secondary school or university. We must remark however that such a distinction is impossible without sufficiently refined data, and those are not publicly released in Belgium, which is very problematic for quality scientific research.

Acknowledgement

The author wants to acknowledge the different members of the Walloon consortium on mathematical model of the covid-19 epidemic for the numerous discussions, especially Sebastien Clesse, Annick Sartenaer, Alexandre Mauroy, Timoteo Carletti as well as Germain van Bever for statistical discussions. The author wants also to acknowledge the members of the Flemish consortium for the very useful exchanges, models' comparisons and helps on improvement, especially the members of the SIMID-COVID-19 team (UHasselt-UAntwerp) and the BIOMATH team (UGent).

This work was supported by the Namur Institute for Complex Systems (naXys) and the Department of Mathematics of the University of Namur, Belgium. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region.

References

[1] K. Rock, S. Brand, J. Moir, M. J. Keeling, Dynamics of infectious diseases., Rep Prog Phys 77 (2) (2014) 026602. doi:10.1088/0034-4885/77/2/026602.

[2] L. Peng, W. Yang, D. Zhang, C. Zhuge, L. Hong, Epidemic analysis of covid-19 in china by dynamical modeling, medRxiv (2020). doi:10.1101/2020.02.16.20023465.
URL https://www.medrxiv.org/content/early/2020/02/18/2020.02.16.20023465

[3] W. Yang, D. Zhang, L. Peng, C. Zhuge, L. Hong, Rational evaluation of various epidemic models based on the covid-19 data of china, medRxiv (2020). doi:10.1101/2020.03.12.20034595.
URL https://www.medrxiv.org/content/early/2020/03/16/2020.03.12.20034595

[4] S. Abrams, J. Wambua, E. Santermans, L. Willem, E. Kuylen, P. Coletti, P. Libin, C. Faes, O. Petrof, S. A. Herzog, P. Beutels, N. Hens, Modeling the early phase of the belgian covid-19 epidemic using a stochastic compartmental model and studying its implied future trajectories, medRxiv (2020). doi:10.1101/2020.06.29.20142851.
URL https://www.medrxiv.org/content/early/2020/07/01/2020.06.29.20142851

[5] T. W. Alleman, J. Vergeynst, E. Torfs, D. Illana Gonzalez, I. Nopens, J. M. Baetens, A deterministic, age-stratified, extended seird model for assessing the effect of non-pharmaceutical interventions on sars-cov-2 spread in belgium, medRxiv (2020). doi:10.1101/2020.07.17.20156034.
URL https://www.medrxiv.org/content/early/2020/07/20/2020.07.17.20156034

[6] C. W. Barbé K., Blotwijk S., Data-driven epidemiological model to monitor the sustainability of hospital care, Tech. Rep. Covid19 Technical Note No. ICDS043020., VUB (2020).

[7] P. Coletti, P. Libin, O. Petrof, L. Willem, A. Steven, S. A. Herzog, C. Faes, J. Wambua, E. J. Kuylen, P. Beutels, N. Hens, A data-driven metapopulation model for the belgian covid-19 epidemic: assessing the impact of lockdown and exit strategies, medRxiv (2020). doi:10.1101/2020.07.20.20157933.
URL https://www.medrxiv.org/content/early/2020/07/25/2020.07.20.20157933

[8] L. Willem, T. Van Hoang, S. Funk, P. Coletti, P. Beutels, N. Hens, Socrates: an online tool leveraging a social contact data sharing initiative to assess mitigation strategies for covid-19, BMC Research Notes 13 (1) (2020) 293. doi:10.1186/s13104-020-05136-9.
URL https://doi.org/10.1186/s13104-020-05136-9

[9] Sciensano: Datasets and epidemiological reports [online] (2020).

[10] Abto: Association of belgian travel organisers, september 2019 travel trends [online] (2019).

[11] Ecdc: European centre for disease prevention and control, daily number of new reported cases of covid-19 by country worldwide [online] (2020).

[12] J. Wallinga, P. Teunis, M. Kretzschmar, Using Data on Social Contacts to Estimate Age-specific Transmission Parameters for Respiratory-spread Infectious Agents, American Journal of Epidemiology 164 (10) (2006) 936–944. doi:10.1093/aje/kwj317.
URL https://doi.org/10.1093/aje/kwj317

[13] L. Willem, K. Van Kerckhove, D. L. Chao, N. Hens, P. Beutels, A nice day for an infection? weather conditions and social contact patterns relevant to influenza transmission, PLOS ONE 7 (11) (2012) 1–7. doi:10.1371/journal.pone.0048695.
URL https://doi.org/10.1371/journal.pone.0048695

[14] Statbel: Structure of the population [online] (2020).

[15] O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio r0 in models for infectious diseases in heterogeneous populations, Journal of Mathematical Biology 28 (4) (1990) 365–382. doi:10.1007/BF00178324.
URL https://doi.org/10.1007/BF00178324

[16] O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, Journal of The Royal Society Interface 7 (47) (2010) 873–885. doi:10.1098/rsif.2009.0386.
URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2009.0386

[17] W. K. Hastings, Monte carlo sampling methods using markov chains and their applications, Biometrika 57 (1) (1970) 97–109. doi:10.2307/2334940.
URL http://www.jstor.org/stable/2334940

[18] A. Wajnberg, F. Amanat, A. Firpo, D. R. Altman, M. J. Bailey, M. Mansour, M. McMahon, P. Meade, D. R. Mendu, K. Muellers, D. Stadlbauer, K. Stone, S. Strohmeier, V. Simon, J. Aberg, D. L. Reich, F. Krammer, C. Cordon-Cardo, Robust neutralizing antibodies to sars-cov-2 infection persist for months, Science (2020). doi:10.1126/science.abd7728.
URL https://science.sciencemag.org/content/early/2020/10/27/science.abd7728

Appendix A. Timeline and estimated parameters

In this Appendix, we provide some technical details concerning the calibration of the model. Table A.6 shows the estimated number of reimportations of covid-19 per day during the holidays period. The complete Belgian policy timeline and corresponding social contact matrices coefficients are presented in Table A.8. The complete list of estimated parameters from the calibration on October 31, 2020 data is given in Table A.8.

Date	people infected	Date	people infected	Date	people infected
07/01/20	40.9 [22.3 ; 56.9]	08/01/20	83.7 [45.7 ; 116.4]	09/01/20	180.4 [98.4 ; 250.8]
07/02/20	41.6 [22.7 ; 57.9]	08/02/20	84.6 [46.1 ; 117.6]	09/02/20	185.5 [101.2 ; 257.9]
07/03/20	42.2 [23.0 ; 58.7]	08/03/20	91.1 [49.7 ; 126.6]	09/03/20	191.9 [104.7 ; 266.8]
07/04/20	41.1 [22.4 ; 57.2]	08/04/20	101.1 [55.2 ; 140.6]	09/04/20	196.8 [107.3 ; 273.6]
07/05/20	38.3 [20.9 ; 53.3]	08/05/20	105.6 [57.6 ; 146.9]	09/05/20	202.8 [110.6 ; 282.0]
07/06/20	39.5 [21.5 ; 54.9]	08/06/20	110.2 [60.1 ; 153.3]	09/06/20	209.3 [114.2 ; 291.0]
07/07/20	42.7 [23.3 ; 59.3]	08/07/20	116.1 [63.3 ; 161.5]	09/07/20	218.5 [119.2 ; 303.8]
07/08/20	42.6 [23.2 ; 59.2]	08/08/20	119.9 [65.4 ; 166.7]	09/08/20	223.1 [121.7 ; 310.3]
07/09/20	44.0 [24.0 ; 61.2]	08/09/20	120.9 [65.9 ; 168.1]	09/09/20	229.4 [125.1 ; 318.9]
07/10/20	46.6 [25.4 ; 64.7]	08/10/20	125.5 [68.4 ; 174.5]	09/10/20	234.9 [128.1 ; 326.6]
07/11/20	42.3 [23.1 ; 58.9]	08/11/20	134.4 [73.3 ; 186.9]	09/11/20	241.9 [132.0 ; 336.4]
07/12/20	41.8 [22.8 ; 58.1]	08/12/20	139.3 [76.0 ; 193.8]	09/12/20	245.5 [133.9 ; 341.4]
07/13/20	45.8 [25.0 ; 63.7]	08/13/20	150.4 [82.0 ; 209.1]	09/13/20	252.7 [137.9 ; 351.4]
07/14/20	47.6 [26.0 ; 66.2]	08/14/20	157.7 [86.0 ; 219.3]	09/14/20	258.9 [141.2 ; 360.0]
07/15/20	46.9 [25.6 ; 65.2]	08/15/20	162.0 [88.4 ; 225.3]	09/15/20	264.8 [144.4 ; 368.2]
07/16/20	49.0 [26.7 ; 68.1]	08/16/20	170.3 [92.9 ; 236.8]	09/16/20	135.9 [74.1 ; 189.0]
07/17/20	50.8 [27.7 ; 70.6]	08/17/20	189.6 [103.4 ; 263.6]	09/17/20	139.0 [75.8 ; 193.3]
07/18/20	51.7 [28.2 ; 71.8]	08/18/20	182.6 [99.6 ; 253.9]	09/18/20	143.7 [78.4 ; 199.8]
07/19/20	51.9 [28.3 ; 72.2]	08/19/20	191.0 [104.2 ; 265.6]	09/19/20	147.1 [80.3 ; 204.6]
07/20/20	59.3 [32.3 ; 82.4]	08/20/20	200.8 [109.5 ; 279.2]	09/20/20	151.0 [82.4 ; 209.9]
07/21/20	64.1 [35.0 ; 89.2]	08/21/20	213.7 [116.5 ; 297.1]	09/21/20	156.4 [85.3 ; 217.5]
07/22/20	67.0 [36.5 ; 93.1]	08/22/20	219.2 [119.6 ; 304.8]	09/22/20	158.3 [86.3 ; 220.1]
07/23/20	72.9 [39.8 ; 101.4]	08/23/20	228.0 [124.4 ; 317.1]	09/23/20	162.3 [88.5 ; 225.6]
07/24/20	77.8 [42.4 ; 108.1]	08/24/20	255.8 [139.5 ; 355.7]	09/24/20	165.9 [90.5 ; 230.6]
07/25/20	79.8 [43.5 ; 110.9]	08/25/20	255.3 [139.3 ; 355.0]	09/25/20	170.9 [93.2 ; 237.7]
07/26/20	80.1 [43.7 ; 111.3]	08/26/20	265.4 [144.8 ; 369.0]	09/26/20	175.9 [96.0 ; 244.6]
07/27/20	89.6 [48.9 ; 124.6]	08/27/20	275.7 [150.4 ; 383.4]	09/27/20	179.4 [97.9 ; 249.4]
07/28/20	95.4 [52.1 ; 132.7]	08/28/20	290.7 [158.5 ; 404.1]	09/28/20	185.0 [100.9 ; 257.2]
07/29/20	100.9 [55.0 ; 140.3]	08/29/20	301.3 [164.4 ; 418.9]	09/29/20	185.2 [101.0 ; 257.5]
07/30/20	106.0 [57.8 ; 147.4]	08/30/20	306.5 [167.2 ; 426.1]	09/30/20	186.5 [101.7 ; 259.3]
07/31/20	113.3 [61.8 ; 157.5]	08/31/20	323.9 [176.7 ; 450.3]		

Table A.6: Estimation of reimportation per day of covid-19 during the holidays period

Timeline	Summary	Home	Work (+ transport)	School	Leisure (+ others)
Pre-lockdown: March 1 → March 13	everything is open	1	1	1	1
Half-lockdown: March 14 → March 18	schools and all leisures closed	1	1	0	$C_{\text{leisurelock}}$
Full lockdown: March 19 → May 3	teleworking + travel restrictions	C_{homelock}	C_{worklock}	0	$C_{\text{leisurelock}}$
Phase 1A: May 4 → May 10	shops partially reopen + few contacts allowed	$\frac{C_{\text{homelock}}+C_{\text{homeunlock}}}{2}$	$\frac{C_{\text{worklock}}+C_{\text{workunlock}}}{2}$	0	$C_{\text{leisurelock}}$
Phase 1B: May 11 → May 17	all shops and compagnies reopen (except bars/restaurants)	$\frac{C_{\text{homelock}}+C_{\text{homeunlock}}}{2}$	$C_{\text{workunlock}}$	0	$C_{\text{leisurelock}}$
Phase 2: May 18 → May 24	progressive partial opening of schools	$\frac{C_{\text{homelock}}+C_{\text{homeunlock}}}{2}$	$C_{\text{workunlock}}$	$0.2C_{\text{schoolunlock}}$	$C_{\text{leisurelock}}$
Phase 2: May 25 → June 1	progressive partial opening of schools	$\frac{C_{\text{homelock}}+C_{\text{homeunlock}}}{2}$	$C_{\text{workunlock}}$	$0.4C_{\text{schoolunlock}}$	$C_{\text{leisurelock}}$
Phase 2: June 2 → June 7	progressive partial opening of schools	$\frac{C_{\text{homelock}}+C_{\text{homeunlock}}}{2}$	$C_{\text{workunlock}}$	$0.6C_{\text{schoolunlock}}$	$C_{\text{leisurelock}}$
Phase 3: June 8 → June 30	Schools partially opened + leisures + bars/restaurants	$C_{\text{homeunlock}}$	$C_{\text{workunlock}}$	$C_{\text{schoolunlock}}$	$C_{\text{leisurejune}}$
Phase 4: July 1 → June 28	Cultural events + Social contacts (15 persons bubble)	$2C_{\text{homeunlock}} - C_{\text{homelock}}$	$C_{\text{workunlock}}$	0	$C_{\text{leisurejuly}}$
Phase 4bis: July 29 → August 31	Social contacts restricted (5 persons bubble)	$C_{\text{homeunlock}}$	$C_{\text{workunlock}}$	0	$C_{\text{leisureaug}}$
September 1 → September 25	bubble optional	$2C_{\text{homeunlock}} - C_{\text{homelock}}$	$C_{\text{workunlock}}$	$C_{\text{schoolsept}}$	$C_{\text{leisuresept}}$
September 25 → October 19	bubble mandatory	$C_{\text{homeunlock}}$	$C_{\text{workunlock}}$	$C_{\text{schoolsept}}$	$C_{\text{leisuresept}}$

Table A.7: Belgian policy timeline and corresponding social contact matrices coefficients

Parameter	Short description	Prior (SD)	Step (SD)	Mean	Median	90% confidence interval
p_0	initial value	$0.0002 \pm 2 \times 10^{-5}$	2×10^{-7}	0.000113368994	0.000104424136	[0.000062181282 ; 0.000198631384]
λ_a	transmission (asympt)	$0.08 \pm 5 \times 10^{-3}$	5×10^{-5}	0.060268239436	0.060182797843	[0.055010180068 ; 0.065547515499]
λ_s	transmission (sympt)	$0.08 \pm 5 \times 10^{-3}$	5×10^{-5}	0.066798120666	0.066101930904	[0.059215177314 ; 0.076805307798]
σ	latent period^{-1}	$0.5 \pm 5 \times 10^{-2}$	5×10^{-4}	0.741615943591	0.743380841466	[0.544207668970 ; 0.913951980829]
τ	presympt period^{-1}	$0.2 \pm 2 \times 10^{-2}$	2×10^{-4}	0.154888026479	0.148102503019	[0.125576130577 ; 0.203169318208]
$p_{a\,(0-24)}$	proba asympt	$0.8 \pm 5 \times 10^{-2}$	5×10^{-4}	0.922441754810	0.939022625212	[0.813783358812 ; 0.989693159324]
$p_{a\,(25-44)}$	proba asympt	$0.7 \pm 5 \times 10^{-2}$	5×10^{-4}	0.857461864066	0.873482939446	[0.689120141848 ; 0.961513803394]
$p_{a\,(45-64)}$	proba asympt	$0.6 \pm 5 \times 10^{-2}$	5×10^{-4}	0.732269126548	0.745976612363	[0.607273667363 ; 0.821782470908]
$p_{a\,(65-74)}$	proba asympt	$0.5 \pm 5 \times 10^{-2}$	5×10^{-4}	0.555341972840	0.560621256502	[0.426239709412 ; 0.652934947580]
$p_{a\,(75+)}$	proba asympt	$0.4 \pm 5 \times 10^{-2}$	5×10^{-4}	0.345643458114	0.342198290076	[0.231839682773 ; 0.498479545273]
$p_{a\,h}$	proba asympt	$0.3 \pm 5 \times 10^{-2}$	5×10^{-4}	0.251661417437	0.241727089429	[0.121947614486 ; 0.394492431025]
$\delta_{(0-24)}$	hospitalisation rate	$0.04 \pm 5 \times 10^{-3}$	5×10^{-5}	0.004276655227	0.002565509161	[0.000621793655 ; 0.014263901584]
$\delta_{(25-44)}$	hospitalisation rate	$0.045 \pm 5 \times 10^{-3}$	5×10^{-5}	0.011833469832	0.010207740274	[0.002468370337 ; 0.025202152996]
$\delta_{(45-64)}$	hospitalisation rate	$0.05 \pm 5 \times 10^{-3}$	5×10^{-5}	0.025187215051	0.024850075987	[0.011695303552 ; 0.041505994872]
$\delta_{(65-74)}$	hospitalisation rate	$0.055 \pm 5 \times 10^{-3}$	5×10^{-5}	0.046411873232	0.046061263574	[0.032030062485 ; 0.060056416554]
$\delta_{(75+)}$	hospitalisation rate	$0.06 \pm 5 \times 10^{-3}$	5×10^{-5}	0.055717824037	0.055703774371	[0.044779038393 ; 0.066511832951]
δ_h	hospitalisation rate	$0.065 \pm 5 \times 10^{-3}$	5×10^{-5}	0.063843874360	0.063559736447	[0.053493601934 ; 0.075264455015]
$\gamma_{a\,(0-24)}$	recover rate (asympt)	$0.29 \pm 2 \times 10^{-2}$	2×10^{-4}	0.302108249861	0.297495929961	[0.251088065843 ; 0.366696165211]
$\gamma_{a\,(25-44)}$	recover rate (asympt)	$0.27 \pm 2 \times 10^{-2}$	2×10^{-4}	0.262581866806	0.256149792676	[0.225665979385 ; 0.323947575923]
$\gamma_{a\,(45-64)}$	recover rate (asympt)	$0.25 \pm 2 \times 10^{-2}$	2×10^{-4}	0.237537428835	0.233069142743	[0.196547244453 ; 0.295714533913]
$\gamma_{a\,(65-74)}$	recover rate (asympt)	$0.23 \pm 2 \times 10^{-2}$	2×10^{-4}	0.205632003429	0.206296147692	[0.163577782890 ; 0.246254297181]
$\gamma_{a\,(75+)}$	recover rate (asympt)	$0.21 \pm 2 \times 10^{-2}$	2×10^{-4}	0.159798596725	0.157967491484	[0.119053929419 ; 0.206141358522]
$\gamma_{a\,h}$	recover rate (asympt)	$0.19 \pm 2 \times 10^{-2}$	2×10^{-4}	0.040587075878	0.039564469664	[0.017441682597 ; 0.068086158490]
$\gamma_{s\,(0-24)}$	recover rate (sympt)	$0.29 \pm 2 \times 10^{-2}$	2×10^{-4}	0.332372010113	0.333335488558	[0.249831098523 ; 0.406846129049]
$\gamma_{s\,(25-44)}$	recover rate (sympt)	$0.27 \pm 2 \times 10^{-2}$	2×10^{-4}	0.279746438042	0.286372909828	[0.207458403881 ; 0.341645828782]
$\gamma_{s\,(45-64)}$	recover rate (sympt)	$0.25 \pm 2 \times 10^{-2}$	2×10^{-4}	0.243338352468	0.243841916160	[0.188861519459 ; 0.300352011445]
$\gamma_{s\,(65-74)}$	recover rate (sympt)	$0.23 \pm 2 \times 10^{-2}$	2×10^{-4}	0.213161146903	0.210642358298	[0.174399129196 ; 0.257620075234]
$\gamma_{s\,(75+)}$	recover rate (sympt)	$0.21 \pm 2 \times 10^{-2}$	2×10^{-4}	0.197711661444	0.196063477516	[0.160976734803 ; 0.239803866724]
$\gamma_{s\,h}$	recover rate (sympt)	$0.19 \pm 2 \times 10^{-2}$	2×10^{-4}	0.176664741023	0.175000325262	[0.144078759282 ; 0.215512674568]
$\gamma_{q\,(0-24)}$	recover rate (hosp)	$0.07 \pm 5 \times 10^{-3}$	5×10^{-5}	0.059861070920	0.058865374932	[0.050580802241 ; 0.072207536513]
$\gamma_{q\,(25-44)}$	recover rate (hosp)	$0.06 \pm 5 \times 10^{-3}$	5×10^{-5}	0.051907199676	0.051234240200	[0.047661615915 ; 0.057810366395]
$\gamma_{q\,(45-64)}$	recover rate (hosp)	$0.05 \pm 5 \times 10^{-3}$	5×10^{-5}	0.049817133672	0.049367990800	[0.046703134479 ; 0.053982406054]
$\gamma_{q\,(65-74)}$	recover rate (hosp)	$0.04 \pm 5 \times 10^{-3}$	5×10^{-5}	0.048657698861	0.048289009179	[0.045722741138 ; 0.052677871088]
$\gamma_{q\,(75+)}$	recover rate (hosp)	$0.03 \pm 5 \times 10^{-3}$	5×10^{-5}	0.047467640429	0.047217602527	[0.044412518378 ; 0.051402013552]
$\gamma_{q\,h}$	recover rate (hosp)	$0.02 \pm 5 \times 10^{-3}$	5×10^{-5}	0.045455769754	0.045445674791	[0.041329432099 ; 0.049668417253]
$r_{(0-24)}$	death rate (hosp)	$0.01 \pm 2 \times 10^{-3}$	2×10^{-5}	0.003093127861	0.002784074969	[0.000998533003 ; 0.006282600141]
$r_{(25-44)}$	death rate (hosp)	$0.015 \pm 2 \times 10^{-3}$	2×10^{-5}	0.004903301298	0.004593462199	[0.001897187961 ; 0.008872988043]
$r_{(45-64)}$	death rate (hosp)	$0.02 \pm 2 \times 10^{-3}$	2×10^{-5}	0.010854880708	0.010627530825	[0.007112430461 ; 0.015264237337]
$r_{(65-74)}$	death rate (hosp)	$0.025 \pm 2 \times 10^{-3}$	2×10^{-5}	0.033886110393	0.034404181173	[0.024066873224 ; 0.042028958878]
$r_{(75+)}$	death rate (hosp)	$0.03 \pm 2 \times 10^{-3}$	2×10^{-5}	0.039689014307	0.039589356086	[0.032584509945 ; 0.046646558158]
r_h	death rate (hosp)	$0.035 \pm 2 \times 10^{-3}$	2×10^{-5}	0.050147311516	0.049003596494	[0.042855048286 ; 0.063504868388]
$\tilde{r}_h$	death rate (homes)	$0.02 \pm 2 \times 10^{-3}$	2×10^{-5}	0.061042831984	0.061050137195	[0.055026813127 ; 0.068058304338]
P_{recovery}	care improvement	$0.7 \pm 5 \times 10^{-2}$	5×10^{-4}	0.618514776960	0.625761510799	[0.546767581812 ; 0.670626235826]
μ_{recovery}	care improvement	200 ± 2	2×10^{-2}	41.259452081425	39.530762846223	[32.786453311874 ; 54.143038942044]
s_{recovery}	care improvement	15 ± 2	2×10^{-2}	22.432425246804	22.273608444500	[18.639582941529 ; 27.056474257277]
SUPP_{hosp}	supplementary entries	$1.15 \pm 5 \times 10^{-2}$	5×10^{-4}	1.229075541493	1.230978392911	[1.164086324477 ; 1.286787328051]
μ_{hosp}	variable hosp. policy	$4000 \pm 5 \times 10^{2}$	5	2374.7047320510	2356.7118217416	[1627.3245987566 ; 3194.5276470489]
s_{hosp}	variable hosp. policy	$2000 \pm 2 \times 10^{2}$	2	1538.8859443431	1527.6682063241	[1138.1099920022 ; 1940.3570980731]
P_{delay}	variable hosp. policy	$15 \pm 2 \times 10^{1}$	2×10^{-1}	10.569847738582	10.522726546392	[8.992775626779 ; 12.365338193794]
P_{cor}	covid-19 related deaths	$0.8 \pm 5 \times 10^{-2}$	5×10^{-4}	0.816063556891	0.835792105479	[0.667127011067 ; 0.893383135406]
P_{th}	transmission to homes	10 ± 1	1×10^{-2}	23.603399118070	23.614857934583	[20.568887295821 ; 26.651097079192]
P'_{th}	transmission to homes	10 ± 1	1×10^{-2}	21.140383014990	21.215327140092	[17.620640758977 ; 24.208883463396]
m_h	transmission in homes	$0.5 \pm 5 \times 10^{-2}$	5×10^{-4}	0.144093350715	0.141920572610	[0.012155471652 ; 0.297114479397]
C_{homelock}	contacts coefficient	$0.5 \pm 1 \times 10^{-2}$	1×10^{-4}	0.443288045160	0.441864888875	[0.407500946378 ; 0.487308764732]
C_{worklock}	contacts coefficient	$0.1 \pm 1 \times 10^{-2}$	1×10^{-4}	0.051192248398	0.051518749011	[0.012914792219 ; 0.093259494623]
$C_{\text{leisurelock}}$	contacts coefficient	$0.1 \pm 1 \times 10^{-2}$	1×10^{-4}	0.083325151860	0.082858778168	[0.053645758400 ; 0.113611637499]
$C_{\text{homeunlock}}$	contacts coefficient	$0.55 \pm 1 \times 10^{-2}$	1×10^{-4}	0.512736736472	0.514801083503	[0.468430352821 ; 0.545259867787]
$C_{\text{workunlock}}$	contacts coefficient	$0.15 \pm 1 \times 10^{-2}$	1×10^{-4}	0.093177145136	0.089110671731	[0.058533515846 ; 0.135677122198]
$C_{\text{schoolunlock}}$	contacts coefficient	$0.15 \pm 1 \times 10^{-2}$	1×10^{-4}	0.234115496497	0.234737320444	[0.197760012192 ; 0.267278948097]
$C_{\text{leisurejune}}$	contacts coefficient	$0.15 \pm 1 \times 10^{-2}$	1×10^{-4}	0.151363734810	0.153027368754	[0.098282435860 ; 0.205158712162]
$C_{\text{leisurejuly}}$	contacts coefficient	$0.3 \pm 1 \times 10^{-2}$	1×10^{-4}	0.452302725093	0.457057350417	[0.376056023732 ; 0.522722938875]
$C_{\text{leisureaug}}$	contacts coefficient	$0.2 \pm 1 \times 10^{-2}$	1×10^{-4}	0.068445627754	0.066289733412	[0.004546102689 ; 0.135752954915]
$C_{\text{schoolsept}}$	contacts coefficient	$0.2 \pm 4 \times 10^{-2}$	5×10^{-4}	0.840552782093	0.915953818158	[0.395750875972 ; 0.992653165749]
$C_{\text{leisuresept}}$	contacts coefficient	$0.25 \pm 4 \times 10^{-2}$	5×10^{-4}	0.330735302961	0.306690031848	[0.181685661643 ; 0.567081229338]
C_{reimp}	reimportation coefficient	50 ± 5	5×10^{-2}	27.333163660351	27.485307829580	[14.993036871298 ; 38.215754758771]

Table A.8: Complete list of estimated parameters from October 31, 2020 data