Epidemiology modeling - systems of ODE (ordinary differential equations) - SEIRQDH model

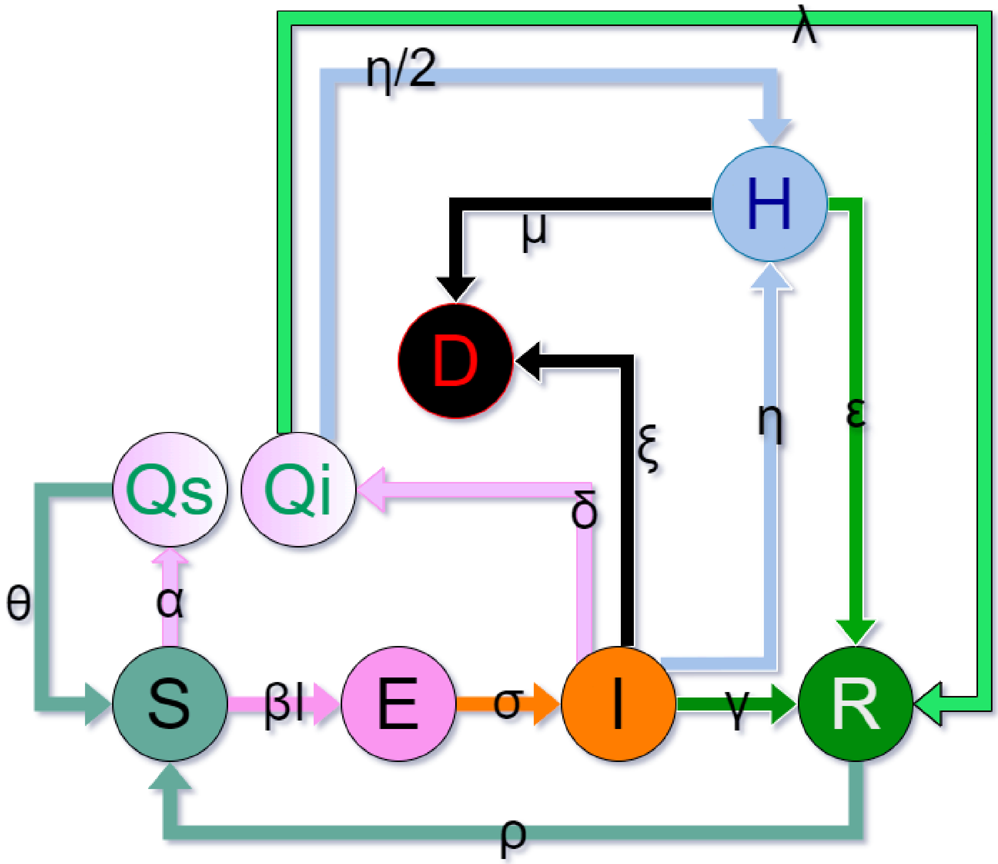
Bartosz Skowronek

Introduction

The aspect I elected to focus my modeling on is social distancing and severity, via introducing compartments: Qi and Qs, (quarantined infectious/susceptible), H(hospitalized) and D(deceased) to a standard SEIR model. The D (deceased) serves mostly an illustrative purpose, because of course a person categorized by that status can no longer change to any other, and the H(hospitalized) compartment allows for more accurate illustration of hospital capacity scenarios. Manipulating the parameters, presented in the next section, enables us to attempt simulating how social distancing can prevent infection spread, or how the severity/mortality of a disease demands a longer period of isolation from a larger proportion of the population to survive. A SEIR+Q+D+H model seems to be a viable representation of the early stages of a pandemic, when the community affected by it attempts various measures to protect themselves, but a vaccine is not yet available.

Model overview

The Qi/Qs compartments take in individuals from compartments S and I that are perceived to be (potentially) infectious and reduce their ability to spread infections. The severity of the modeled disease, described by parameter η(eta), influences how often infected patients require to be hospitalized (in other words, how intense is the flow from compartment I (infectious) to H (hospitalized) . In this model, the term “quarantine” describes a state of social distancing that reduces social contact, due to voluntary isolation of a susceptible person (compartment Qs) or medication at home, in quarantine, of an infected person (Qi). The infected in Qi retain a possibility of hospitalization due to severity, but due to the assumption that staying at home and treating the infection in isolation reduces strain on the immune system of an individual, I reduced severity in this case to ½ of the parameter, even if perhaps this may be either an incorrect assumption or proportion. Below is a graph that illustrates the flow between compartments in the presented model, and next are descriptions of compartments, parameters and the equations that define them. Various studies on the COVID-19 pandemic can provide reference data for this model, for example seasonal mortality rate distributions.



Compartments:

S: Susceptible (normal engagement in social activities)

E: Exposed (susceptible individuals that had contact with virus and are developing an infection)

I: Infectious (no special social distancing measures implemented)

R: Recovered (infection healed, immunity developed)

H: Hospitalized (due to severity of infection; quarantined at hospital)

Qi: Quarantined (infection treatment at home; social distancing)

Qs: Quarantined (social distancing of a susceptible individual)

D: Deceased

Parameters (all parameters are numbers within range $[0, 1]$):

β(beta): Transmission rate $\frac{(average\ number\ of\ contacts\ per\ individual) \bullet (probability\ of\ infection\ on\ contact)}{total\ population}$

α(alpha): % of susceptible population that goes into quarantine (social distancing)

θ(theta): 1/(average time of social distancing of a quarantined susceptible individual)

σ(sigma): Incubation rate (1/ time for exposed individual to become infectious)

γ(gamma): Recovery rate (1/time of recovery from infection) without any distancing measures

δ(delta): % of infected population that goes into quarantine (medication at home)

η(eta): % of infected individuals that require hospitalization (disease severity)

μ(mu): Mortality rate for hospitalized patients

ξ (xi): Mortality rate for infected patients (death before hospitalization)

ρ(rho): Rate at which recovered individuals lose immunity (1/days to become susceptible)

ε(epsilon): Recovery rate for hospitalized patients (1/time of recovery)

λ(lambda): Recovery rate for infected individual in quarantine

Model Equations:

$$\frac{dS}{dt} = -\beta SI + \rho R - \alpha S + \theta Q_s$$

$$\frac{dE}{dt} = \beta SI - \sigma E$$

$$\frac{dI}{dt} = \sigma E - \gamma I - \eta I - \xi I - \delta I$$

$$\frac{dR}{dt} = \gamma I + \epsilon H + \lambda Q_i - \rho R$$

$$\frac{dQ_i}{dt} = \delta I - \lambda Q_i - \frac{\eta Q_i}{2}$$

$$\frac{dQ_s}{dt} = \alpha S - \theta Q_s$$

$$\frac{dH}{dt} = \eta I + \frac{\eta Q_i}{2} - \mu H - \epsilon H$$

$$\frac{dD}{dt} = \xi I + \mu H$$

Key Model Features:

Social distancing is represented by variations in the β parameter. Lower β values indicate stronger social distancing measures.

Disease severity is captured by the η parameter, representing the rate at which infectious individuals require hospitalization.

The quarantine compartment (Q) reduces the number of infectious individuals actively spreading the disease.

The hospitalized compartment (H) allows for tracking of severe cases and healthcare system load.

The deceased compartment (D) provides a way to track mortality over time.

Initial conditions and exploration

To keep things rather simple and easy to imagine or understand, I have chosen to try emulating the beginning of the COVID-19 pandemic in Poland, setting the initial conditions as:

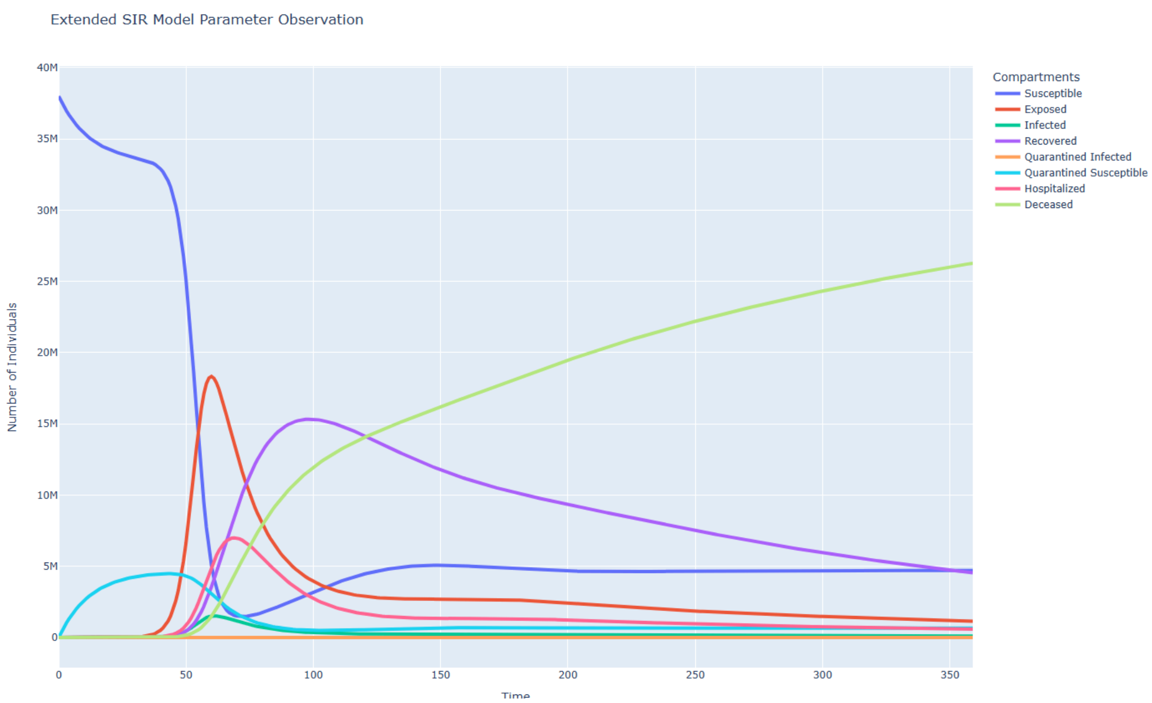
N = 38000000

initial_conditions = [N - 1, 0, 1, 0, 0, 0, 0, 0] # [S0, E0, I0, R0, Qi0, Qs0, H0, D0]

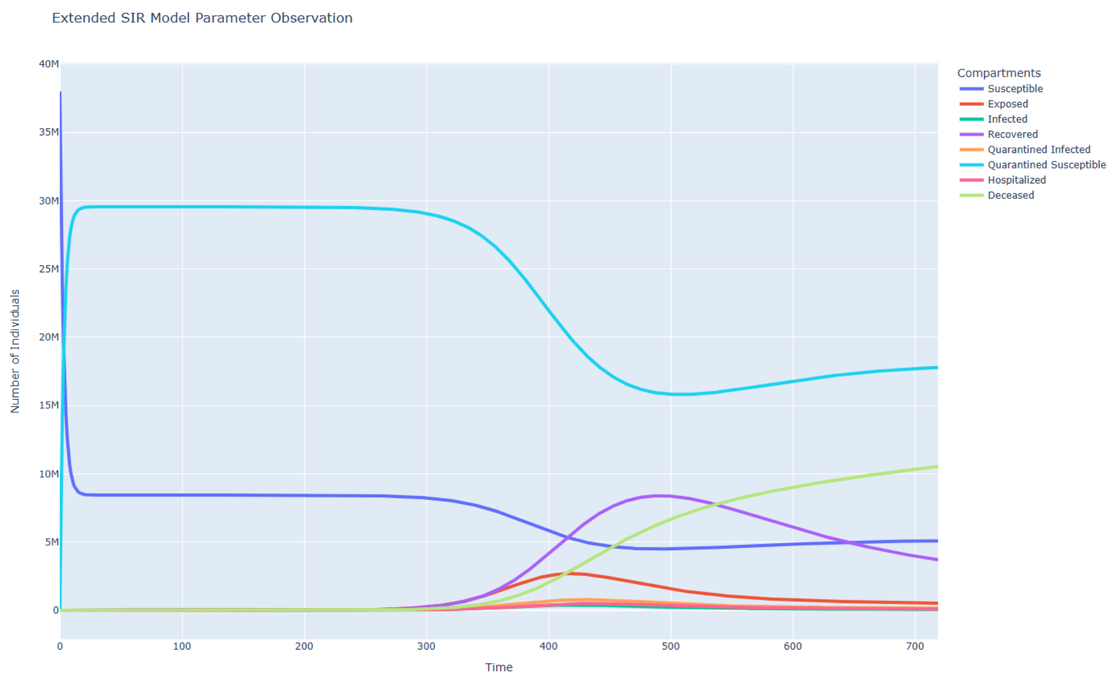
This is not perfect, because we assume that the entire pandemic starts from a single outbreak, and COVID-19 had some characteristics not captured in my model, for example the large amount of asymptomatic infections, which would likely deserve their own compartment for proper modeling. Additionally, my model does not account for comorbidities, constant hospital patient load and other factors. However, it still allows for visualization of some aspects of the pandemic and relations between events, such as quarantining greatly reducing fatality rates. By setting parameters in accordance to some real-world data and my own vague approximations (sources are at the end of this document), we get something that looks like:

```
prob_of_infecting = 1/7
avg_no_contacts_per_individual = 25
beta_true = prob_of_infecting * avg_no_contacts_per_individual / N
sigma_true = 1/7  # Incubation period of ~7 days
gamma_true = 1/14  # Recovery rate (14 days average)
rho_true = 1/90  # Rate of loss of immunity (90 days)
eta_true = 0.1  # Hospitalization rate for infected patients (severity)
mu_true = 0.1  # Mortality rate for hospitalized patients
xi_true = 0.001  # Direct mortality rate from infection (compartment I)
alpha_true = 0.01  # Rate at which susceptible individuals enter quarantine
delta_true = 0.01  # Rate at which infected individuals enter quarantine
epsilon_true = 1/20 # Rate at which individuals leave hospital
theta_true = 1/14  # Rate at which susceptible individuals leave quarantine
lambada_true = 1/14  # Rate at which infected individuals leave quarantine
```

While all of those are difficult in regard to giving a reasonable approximation, the last five quarantine parameters are especially hard to give a general measure for, because in a real case we lived through, they greatly varied due to governmental restrictions or various other things, like the shock factor of the early stage of the pandemic. Anyway, with the above parameters (quarantine rates are minimal) and a time of 360 days, my model presents this result:

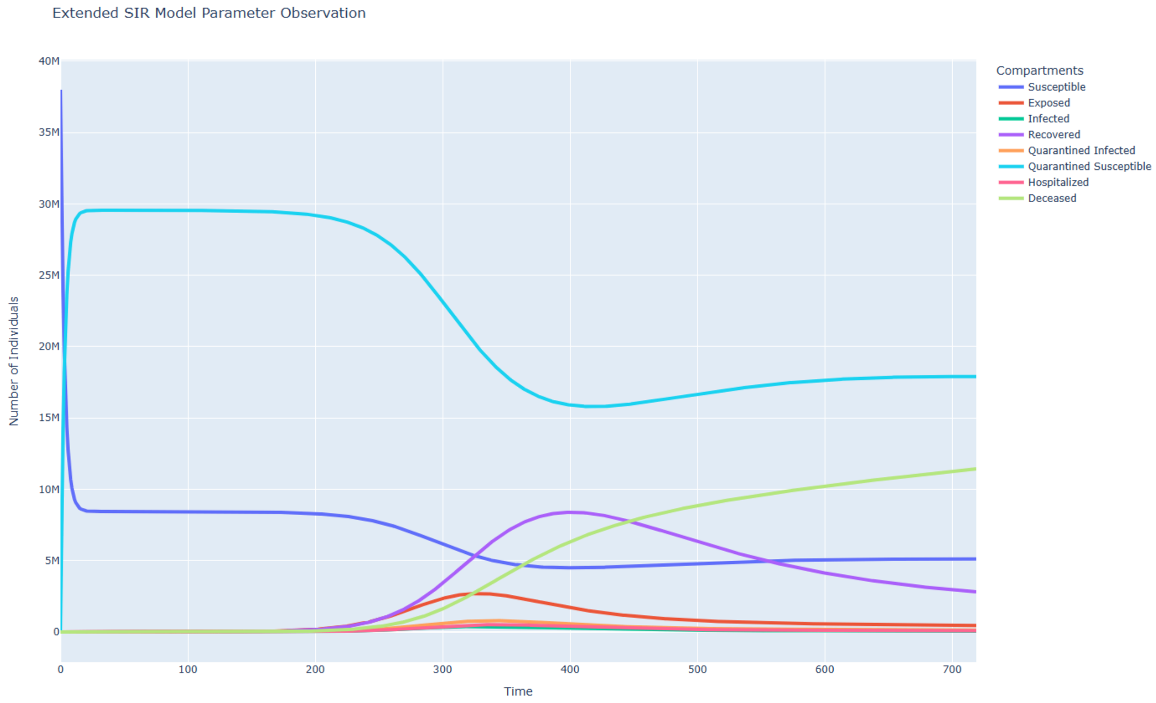


It seems quite radical that the model so rapidly turns hospitalized patients deceased. It is also worth noting that for a country with a theoretical number of about 160 thousand hospital beds, we would fill them pretty quickly (around the 40 day mark hospital demand begins rising astronomically, eventually reaching levels around 7 million beds and that is obviously an impossible threshold). Due to the nature of my model, infection numbers never become very high due to quickly turning into hospitalisations (and later deaths) or recoveries. We see that without any preventive measures, a pandemic like the COVID-19 one can get out of hand very quickly. Obviously this plot may also just be proof of the weakness of my model, but it's not enough for a full judgment. Let's change the parameters a little:



This plot above shows data for the same parameters, but with alpha and delta set to 0.25 instead of 0.1. With a single outbreak infection and just 25% of susceptible and infected individuals choosing to quarantine at each time step (1 day in that case) we can go for almost a year without drastic consequences, about 300 days instead of about 30 for the previous parameter set.

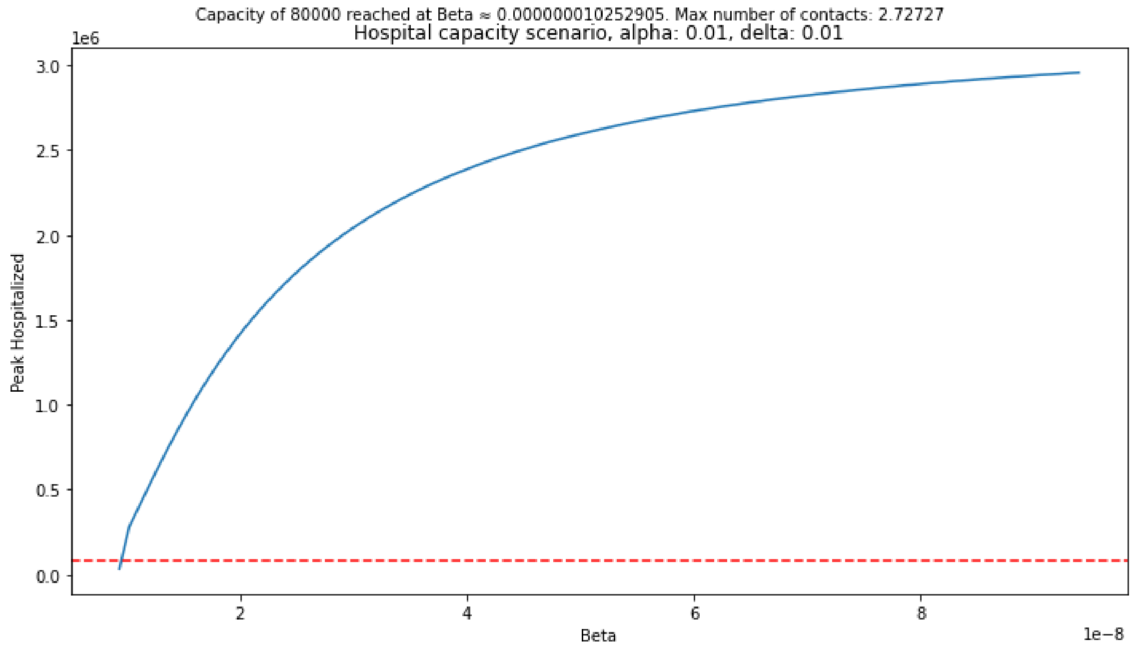
Lastly in this part, another plot with the same parameters, but with the number of initial infections changed to 25. This time, drastic consequences (again apparently greatly overblown by the model) happen about 100 days faster than with a single infection. It is impossible to tell how many simultaneous outbreaks happened in the beginning of the COVID-19 pandemic (in Poland) because of initially weak detection rates and lack of information about the virus, but it is likely that in our country it started in multiple places at once. A conclusion we may come to if we believe in this model's accuracy here is that the scale of pandemic growth isn't really affected by the number of initial infections, but that number has great impact on how fast initial problems appear.



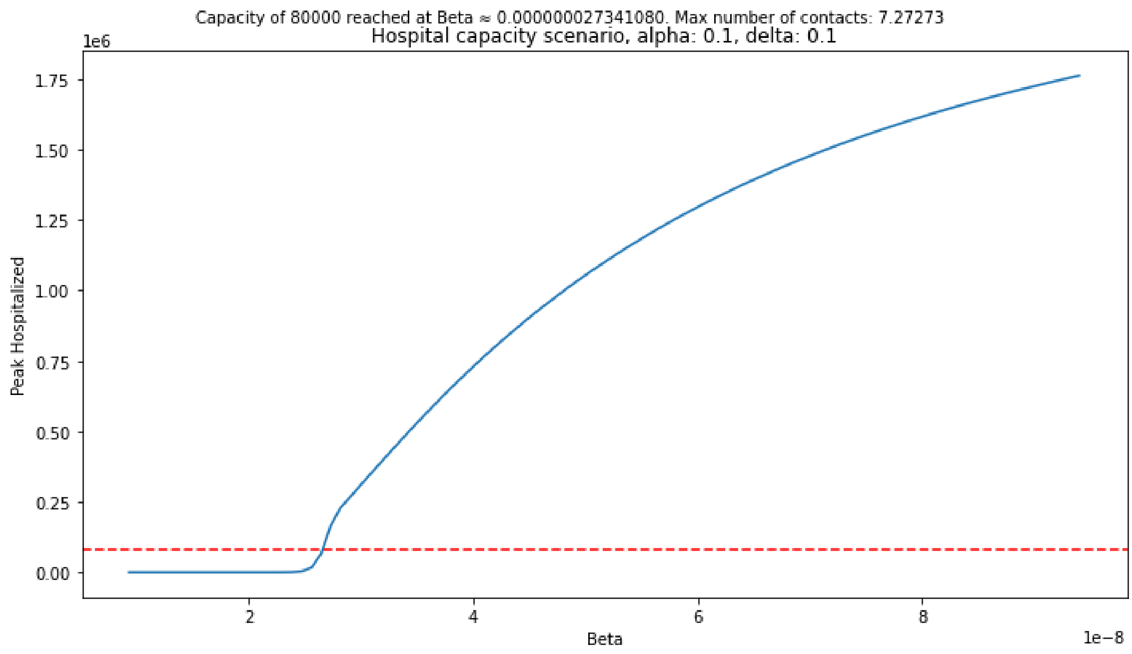
Exploratory scenarios

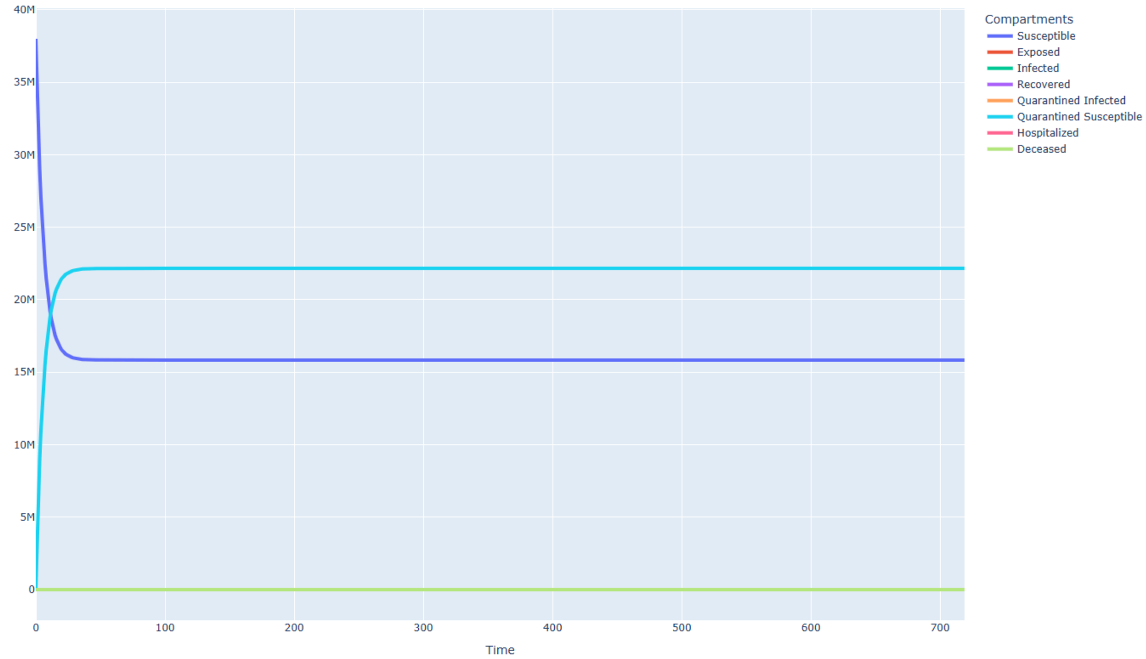
1: Hospital capacity

Due to the fact that my model doesn't have a simple "infected" compartment to measure, a threshold of hospital capacity other than (0.1*population) must be found. A figure I found on the internet (again, sources with descriptions are at the end of this document) says that Poland has about 160 thousand hospital beds. It is, however, somewhat difficult to find how many of those are occupied on a regular basis, so I will use a vague approximation based on the fact that in the anomalous year of 2020, about 55% beds in Polish hospital beds were occupied, and that somehow that was a smaller percentage than the previous year. I will resort to a simple estimate of 50% beds being taken. That would leave us with a hospital capacity threshold of 80 thousand free beds.

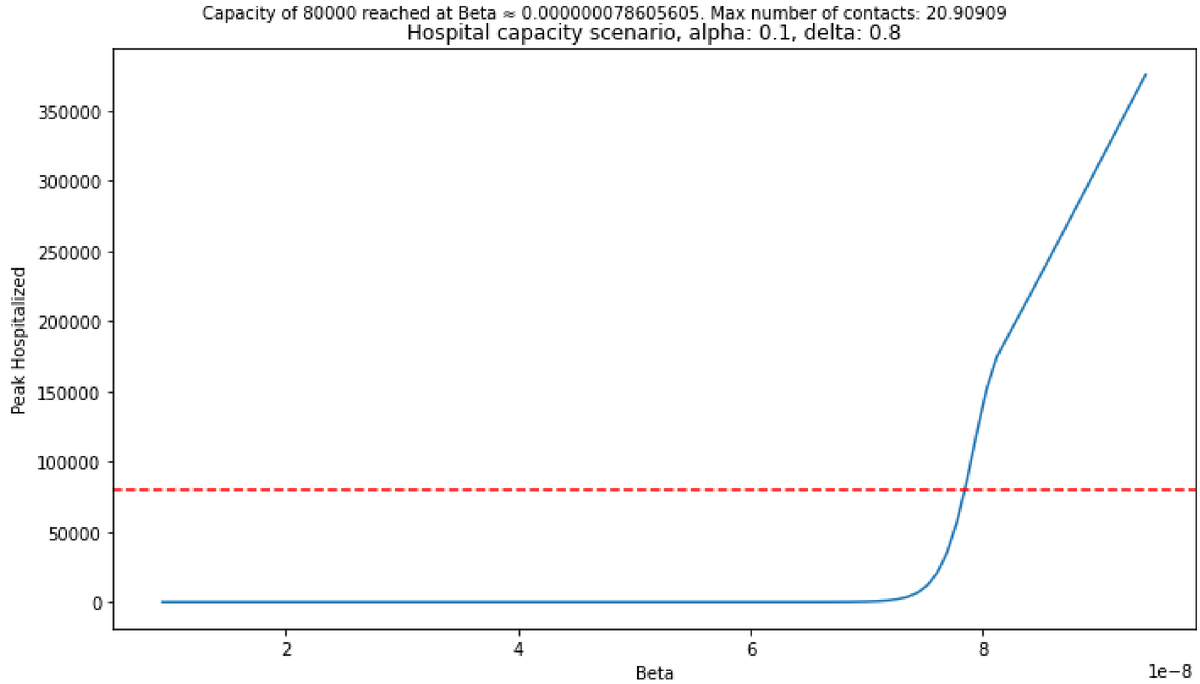


For minimal alpha and delta, the number of human contacts to stay safe (below hospital capacity threshold) is very small. However, a small percentage of quarantining efforts can be enough to roughly triple the safe number of contacts:



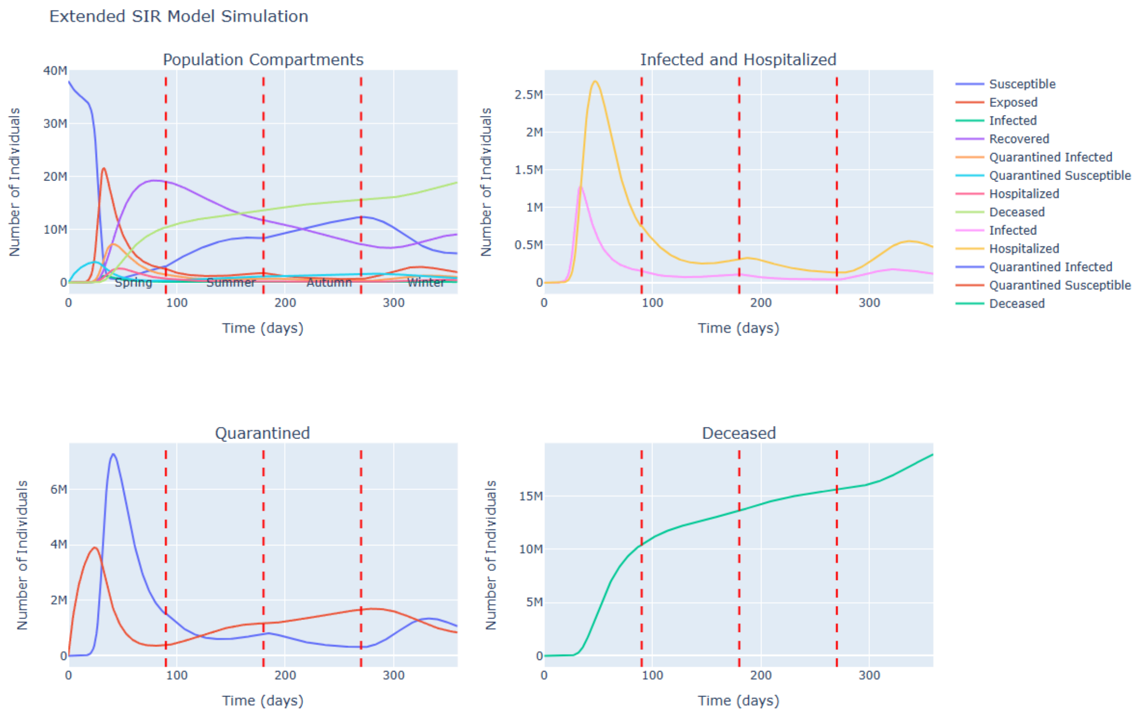


This is the plot of compartments for a beta just below the threshold. It appears like if there's no cure (like a vaccine), long periods of quarantining work for survival. The number of people hospitalized is not visible because it is too low in scale.



If we keep quarantining the infected (delta = 0.8), the number of contacts gets pretty close to a normal number ("normal number" as referenced [here]).

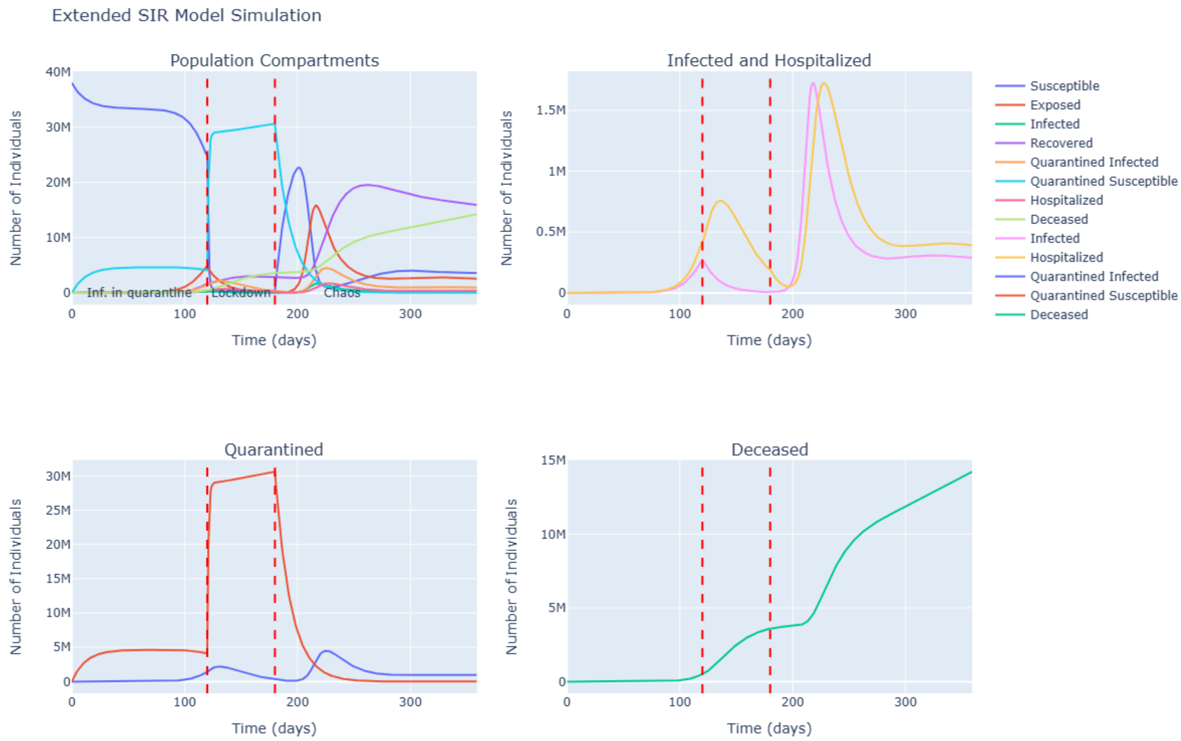
2: Seasonal variations



Reference interactive plot is in file "wykres_seasonal.html"

In this scenario, there are 4 time compartments in a year, to try representing seasons. The difference between those 4 compartments is their beta value, which is modulated by the "prob_of_infecting" variable, that is set to ¼ for spring and autumn, ½ for winter and 1/7 for summer. A better and more accurate seasonal model could be achieved by splitting the year into 365 compartments, each having a moving average of temperature from three days (temperature data acquired from existing sources) and modulating the "prob_of_infecting" variable based on the temperature value, because it has been proven that temperature can be a good proxy for modeling the seasonal variability of COVID-19 (source). In my model, the difference is not very big, but still noticeable. The summer compartment slows down infectious spread and flattens the curve of death count, whereas the winter compartment noticeably speeds up the death toll.

3: Custom scenario: complete lockdown for two months



Reference interactive plot is in file "wykres6.html"
The custom scenario introduces three phases, for the first 120 days there's standard pandemic development, but with a focus on quarantining infected persons, but when the numbers of fatalities start rising, a government-imposed lockdown begins and lasts for 60 days. The lockdown is simulated by reducing the number of average interactions to 0.3 (as close to zero as possible) and increasing quarantine rates to 1 each, meaning a 100% quarantine rate for both infected and susceptible individuals. The quarantine seems to be slowing down the tempo of the deceased count increasing and drastically reduces the number of infected and hospitalized individuals, but it is too short to impactfully prevent further problems. After the lockdown parameters are changed to simulate a more relaxed social interactions profile to attempt simulating mass anti-government chaotic protests, with 45 average interactions up from the standard 25 and quarantine rate for susceptible persons down to 0, while the quarantine rate for the infected is 0.4, reflecting how in real conditions people voluntarily quarantine when infected to protect others, even if it is not expected of them. Shortly after the chaotic phase begins, there is a massive spike in infections and hospitalizations, which then falls due to the excess infections turning into deaths.

Data generation tool

The function “generate_data” allows for data generation when the parameter “generation” is set to True, and additionally artificial gaussian (normal distribution) noise is added to the data if the parameter “add_noise” is set to True. There is a seasonal variation in the function as well, using the parameters seasonal_amplitude=200000, seasonal_period=365 the data simulates an increase in infections every seasonal_period, when set to 365 and a high enough amplitude, it becomes another way of simulating how an infectious disease becomes easier to spread during the winter.

Prediction using optimization

While I have attempted using scipy.integrate.solve_ivp to solve this problem, I have found no success. GPT-generated code and variations of the solutions found in 06_compex_generation.py did not manage to work on this model, usually getting stuck in infinite loops while attempting, sometimes even with iteration limiting there was no success. I may have not exhausted the potential options, but I don’t know how to proceed.

Various sources used for reference in this report

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9877962/ - for probability of infecting on contact
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8478104/ - for number of social contacts per individual per day

  https://www.researchgate.net/figure/Daily-average-number-of-contacts-per-person-in-age-group-j-The-average-number-of_fig2_228649013 - for number of social contacts per individual per day
- https://www.webmd.com/covid/coronavirus-incubation-period for covid19 incubation period
- https://www.verywellhealth.com/how-long-does-covid-last-8547620 for a bold approximation of covid19 recovery time

- https://www.aamc.org/news/had-covid-recently-here-s-what-know-about-how-long-immunity-lasts-long-covid-and-more for an idea of length of immunity waning time
- https://www.nature.com/articles/s41598-023-35413-z for what % of infected with COVID require hospitalization
- https://www.gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2 for how many infections result in death
- https://www.statista.com/statistics/1111959/poland-general-hospitals-and-hospital-beds/ For hospital capacity of Poland
- https://www.medexpress.pl/ochrona-zdrowia/gus-ogromne-tapniecie-liczby-hospitalizacji-w-2020-roku-83673/ for % of occupied hospital beds