Modeling of Coronavirus Spread

Sergey Voronin, IAI, 05.2020

1 Overview of the SIR model

The spread of the novel coronavirus which causes the COVID-19 disease can be modeled with a simple differential equation model called SIR [1], where in a whole population of a given region, S stands for the number of susceptible people to infection, I for the number of infections, and R for the number recovered during an epidemic. Note that the spread of this virus occurs directly from person to person without an intermediary such as a mosquito, which makes simple models such as SIR applicable. The default SIR model makes a number of simplifications, in part because it does not assume population change due to birth and death, and age and time/weather dependent factors, although some of these changes can be worked into the model. Susceptible individuals are individuals that have never been infected and are able to catch the disease, which for this virus presents almost the entire population, with the exception of those survived and recovered. This is in contrast to a new strain of the flu, with which most individuals have some experience due to being sick before with a likely different, but similar structured strain. Recovered individuals are assumed to remain immune from developing the disease and are not considered to become possible virus spreaders if exposed. Without vitality stats, the population is fixed so that $S + I + R = N$ and the disease spreads through the interaction of susceptible and infected individuals. There are two parameters $0 < \gamma, \beta < 1$ which control the time evolution of S, I, R. It is assumed that an infected individual recovers at rate γ, so the period of infection is $1/\gamma$ days (for coronavirus this period seems long, on the order of 10 or more days). The second parameter β is the approximate normalized transmission rate and is proportional to the fraction of how often a susceptible-infected contact results in a new infection [2] (for coronavirus this is relatively high, as it is found to be easily transmissive with close contact). Based on these assumptions, the spread can be described by the following differential equations:

$$\frac{dS}{dt} = -\beta \frac{SI}{N}; \quad \frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I; \quad \frac{dR}{dt} = \gamma I \tag{1.1}$$

The term $\beta \frac{SI}{N}$ represents the number of newly infected individuals per unit time, corresponding to homogeneous mixing of the infected and susceptible cases, the negative of which is the rate at which S decreases. The sum of the three equations is zero which represents that the total population doesn't change (a plausible estimate for a large enough region). The rate of change of infections is $\beta \frac{SI}{N}$ minus the rate of recovery, which is γI and equal to the time derivative of R. We can also easily introduce vitality statistics (birth and death) into the differential system. The new births act to increase the susceptible population, while the deaths decrease all the three S, I, R quantities. This logic leads to the system:

$$\frac{dS}{dt} = \mu N - \nu S - \beta \frac{SI}{N}; \quad \frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I - \nu I; \quad \frac{dR}{dt} = \gamma I - \nu R, \tag{1.2}$$

where μ, ν are respectively the normalized birth and death rate parameters. Also possible is to differentiate between mortality rates for the three different groups and account for loss of immunity over time [1].

2 Basic modeling

The system of first order differential equations above can be solved from given initial conditions for S, I, R and the parameter values γ and β. The solution can be obtained at future time by e.g. the Runge-Kutta method. Notice first, that the infection will always die out, no matter the initial conditions. If not, $\frac{dR}{dt} > 0$ at $t \to \infty$, implying $R \to \infty$, a contradiction. The following R code shows the basic steps in the differential system setup, parameter initialization and integration:

```
library(deSolve)
N = 1e7;
integration_range = seq(from=0,to=100,by=0.1)

sir_equations <- function(time, variables, parameters) {
  with(as.list(c(variables, parameters)), {
    dS <- mu*N - beta * I * S/N - nu*S;
    dI <-  beta * I * S/N - gamma * I - nu*I;
    dR <-  gamma * I - nu*R;
    return(list(c(dS, dI, dR)))
  })
}

param_vals <- c(
  beta  = 0.3,
  gamma = 0.1, nu = 0.01, mu = 0.01
)

initial_values <- c(
  S = N,
  I = 30,
  R = 20
)

sir_sys_sol <- ode(
  y = initial_values,
  times = integration_range,
  func = sir_equations,
  parms = param_vals
)

sir_sys_sol = as.data.frame(sir_sys_sol);
plot(integration_range, sir_sys_sol$I, col="red", ylim=c(1,N), main='Urban spread model', xlab
    ='time (days)', ylab='I(t),R(t)')
points(integration_range, sir_sys_sol$R, col="blue", ylim=c(1,N))
lines(integration_range, sir_sys_sol$R, col="blue", ylim=c(1,N))
legend("left", c("infected", "recovered"),
       col = c("red", "blue"), lty = 1, bty = "n")
```

Below, we show some examples from the online resource [2] and the above listing. For example, let us suppose that $S = 10000$, $I = 50$, and $R = 25$. This would mean that we are starting with about

0.5% infection rate, and that some of the infected have already recovered, in given proportion. If we simulate the results for fixed $\gamma = 0.066$ (roughly 1/15) with different values of β, we obtain:

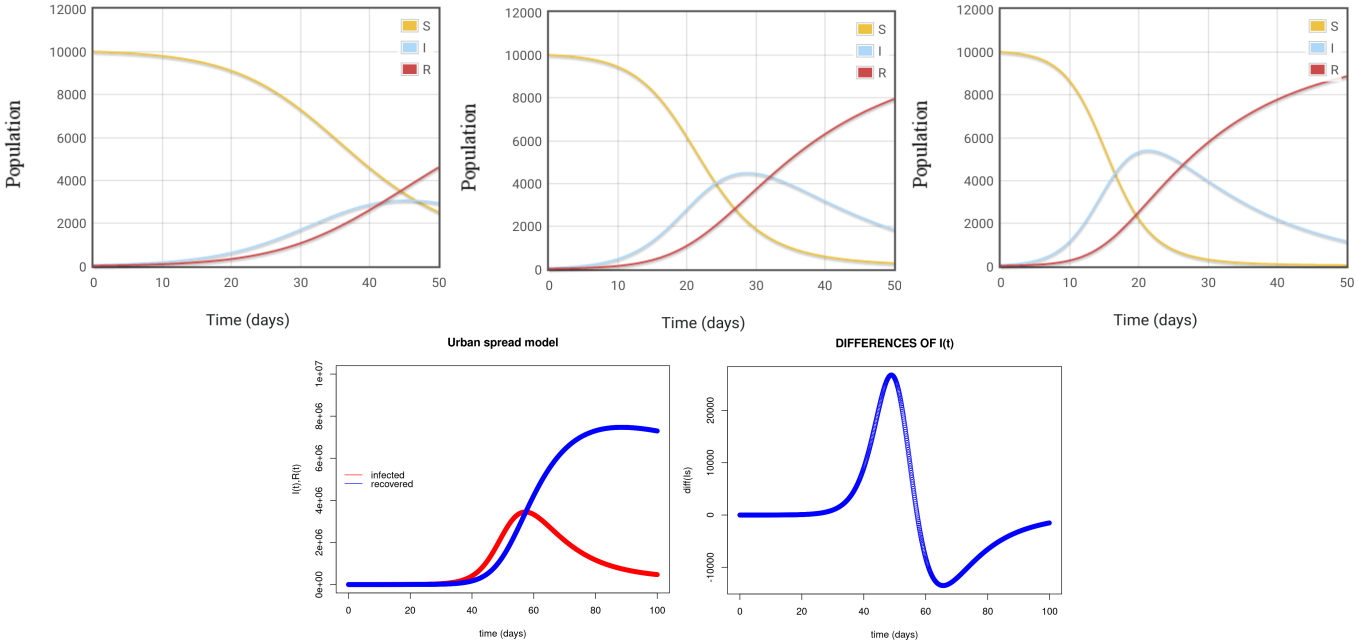

Figure 1: Simulations for evolution of S, I, R over first 50 days with $\beta = 0.2, 0.3, 0.4$ (top) and example simulation with vitality stats for $I, R(t)$ and corresponding diff$(I(t))$.

Notice that in a short time period the value of γ would not change (unless e.g. a fast curing drug is discovered or the virus significantly mutates). For this reason, the only thing we can affect is the value of β, via measures such as social distancing and also environmental factors such as air temperature and humidity. With a lower value of β it takes longer time to reach the peak and the peak of infected individuals (light blue curve) ascends to a lower maximum than for higher values of β. Slightly higher values significantly elevate the peak and shorten the time to achieve it. With $\beta = 0.4$, starting from 50 infections, we get roughly half the population (of 10075) infected in 3 weeks. Notice that overall $I(t)$ numbers reported very official and news media agencies for a given region are a sample of the overall numbers (as only a subset of those infected are typically tested). On the other hand, the rate of change of infections dI/dt can be analyzed and compared to the reported numbers. Parameters (e.g. α, β, μ, ν) can be tuned based on past data (so the calculated dI/dt reflects the observed trend or possibly some processed, statistically imputed, or interpolated data) and the *optimized* model can be used to make predictions into the near future. This can be done by pre-supposing some initial values for the parameters and in a loop, choosing randomly a parameter to tune and varying the value of the chosen parameter, while comparing the resulting dI/dt behavior (as obtained from integration with the current parameter settings) to the observed trend, setting the parameter value corresponding to the closest match in some least squared sense.

Also important is the so called herd immunity containment scenario, in which $I(t)$ does not increase,

reach a maximum and then decrease as in the above plots, but rather decreases monotonically to zero. This occurs, when the so called effective reproductive number is less than 1. If a fraction of the population has been vaccinated, or gains immunity, then $\rho S(0)$ individuals are removed from the susceptible population at the beginning. It can be shown that containment occurs if $\rho \geq \rho_c = 1 - \frac{1}{R_e}$ [1]. The reproductive number R_e for the novel coronavirus is estimated at between $2 - 4$, though with large possible variation [4]. This means that up to about 70% of the population must gain immunity before a containment scenario. While so, it is likely that infection rates will drop in the Summer in the Northern hemisphere with higher air temperatures and more concentrated UV lighting [3], effectively reducing air-to-air transmission outdoors. However, overall, outbreaks may continue until the critical immune threshold is reached via vaccination or natural means.

3 References

[1] Weiss, Howard Howie. 'The SIR model and the foundations of public health.' Materials mathematics (2013): 0001-17.
[2] Hans Nesse. 'Global Health - SIR Model.' http://www.public.asu.edu/~hnesse/classes/sir.html; shinySIR R module https://cran.r-project.org/web/packages/shinySIR/.
[3] Walker, Christopher M., and GwangPyo Ko. 'Effect of ultraviolet germicidal irradiation on viral aerosols.' Environmental science technology 41, no. 15 (2007): 5460-5465.
[4] Liu, Ying, Albert A. Gayle, Annelies Wilder-Smith, and Joacim Rocklöv. 'The reproductive number of COVID-19 is higher compared to SARS coronavirus.' Journal of travel medicine (2020).