# Multi-Variable Wuhan Coronavirus Outbreak Model

*Hart Traveller Massie-Keller | Jan 28, 2020*

**Reference**

Code: https://github.com/harttraveller/OBsim
Contact: https://www.harttraveller.com/contact-1

If you would like to suggest more parameters for the model, you find bugs I need to fix, or you think the source code should be modified please use the contact form on my website to reach me


**Abstract**

Current popular exponential models of the Wuhan coronavirus spread are incomplete, and thus offer a limited picture of possible future outbreak developments. This can be rectified with a more complex model, that while not conclusive on future events, can be used to more effectively model the range of possibilities. Furthermore, current apparent exponential growth of the virus could be due to a time lag between infection and reporting - the virus might not be spreading as quickly as it appears. In this report, an alternative model is suggested, and links to the program are provided so the reader may experiment with different parameter settings on their own.

*This report does not aim to provide a single conclusive model for evaluating the future of the outbreak, but rather a model for readers to use and experiment with to better understand the range of future possibilities.*

**Introduction**

The recent coronavirus outbreak in Wuhan, China (and as of Jan 28th in several other Chinese provinces/countries) has fueled internet panic, as well as speculation about its potential to become a global pandemic. One way in which misinformation has spread under the guise of data analysis, is that of overly simplistic exponential models of infection growth. For instance, earlier today a gif was posted to the subreddit /r/dataisbeautiful which extrapolates to future infection spread given current trends [Fig. 1], as of 6:51 PM Jan 28th, 2020, it sits at ~24,000 upvotes.

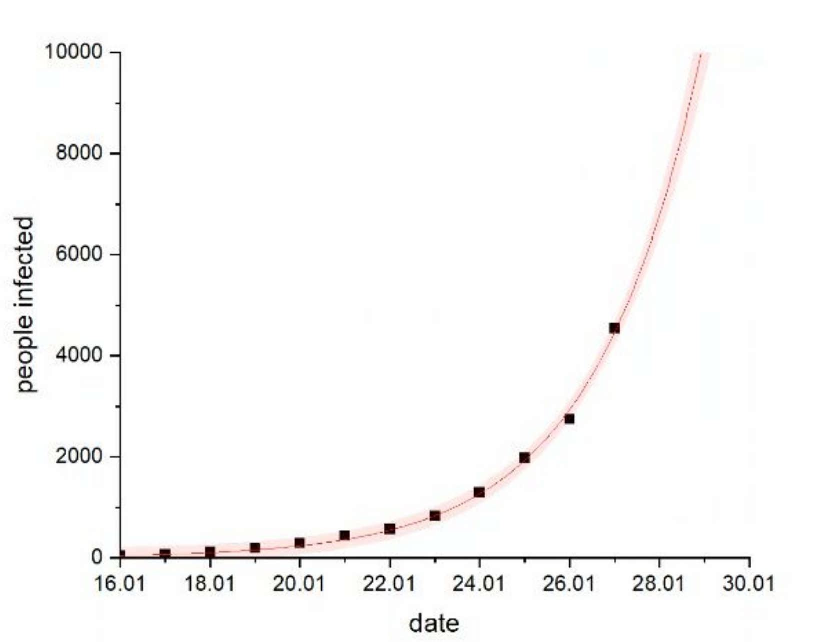



*Figure 1: Simplistic Model*

A number of other models have appeared in which individuals plug infection numbers into excel, and then use the mean daily percent increase to extrapolate into the future indefinitely. Such models estimate the global number of infected at well over a million within weeks. The problem with these models, though they are not *necessarily* incorrect, is that they do not account for other variables, and they overlook reasons why growth could appear to be exponential, and thus treating them as concretely factual (as many are) is disingenuous.

One reason growth might appear to be exponential is that the number of people actually infected is significantly higher than the number of known infected, and because of the heightened awareness the reported cases are quickly catching up to the actual infection figures. Counterintuitively, the longer this infection has been in circulation, the better, as this allows for lower overall infection rates while maintaining current estimates. Hypothetically, it would be better if over the course of 100 days the infection grew from 1 to 50,000, than if over the course of 10 days it grew from 1 to 5000. If current academic models assume the increase in reported figures represents the true underlying growth curve, then they may be flawed. Note that I am not asserting these exponential models are incorrect, but rather that they are incomplete in their considerations, and as such should be informed by a broader analysis.

In the interest of providing this broader analysis, I stayed home from school today and coded a program to model infectious disease outbreak that accounts for actual infection and known infection rates. Furthermore, by updating the parameters input to this model based on real world data, one will get a more complete picture of the expected course of the outbreak, than simply modelling an exponential curve in excel.

The program is created in Python and can be found at the GitHub repository link on the first page under "references". A tutorial jupyter notebook for the program can also be found in the GitHub repository. Line graphs, interactive/animated bar charts, and data export capability are included in the Python Class. Parameters are listed on the next page.

# PARAMETERS

**starting_infected**

descr: number of individuals infected on day 1 of outbreak, suggested input is 1
range: [ 0, inf ]

**infected_growth_rate**

descr: daily percentage growth rate of infection, realistic inputs range from 0.01 to 0.1
range: [ 0, inf ]

**awareness**

descr: starting awareness of virus, ranges from 0 to 1, increases based on multiplier so a 0 value will never increase
range: [ 0, 1 ]

**awareness_growth_rate**

descr: Daily percentage growth rate of awareness value
range: [ 0, inf ]

**mortality_day_rate**

descr: Daily percentage of infected that die
range: [ 0, 1 ]

**recovery_day_rate**

descr: daily percentage of infected that recover
range: [ 0, 1 ]

**recovery_multiplier**

descr: if people are known to be infected, they may recover faster, this represents a multiplier for the speed they recover if they are known to be infected
range: [ 0, inf ]

**containment_effectiveness**

descr: what percentage of known infected patients are prevented from infecting others, if 0, containment is 0 percent effective, if 0.95, containment is 95% effective
range: [ 0, 1 ]

**cycles**

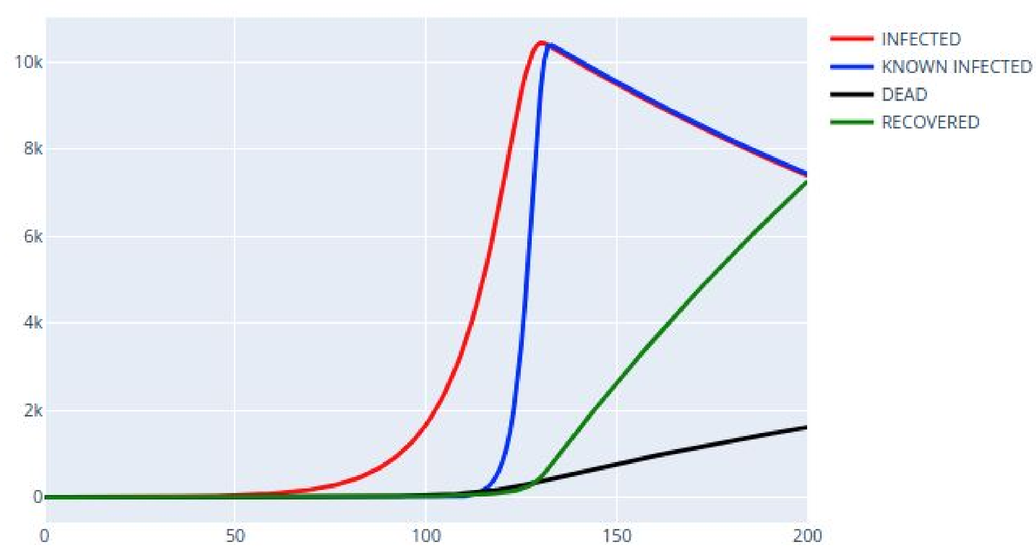
descr: equivalent to number of days the simulation runs for
range: [ 0, inf]

**Scenario 1**

First I wanted to test the hypothesis that apparent growth may appear to be a high degree exponential where real growth is a low degree exponential function. Indeed, there are certain combinations of

The following parameter set yielded the described situation, and seemed to be a potentially accurate conservative model of the situation. In general, simulations following the above described pattern had higher awareness growth rates than infected growth rates with extremely low initial awareness values. Note that the program allows for animated and interactive bar charts of the infection growth which can not be represented here.

starting_infected=1
infected_growth_rate=0.08
awareness=1/1e15
awareness_growth_rate=0.3
mortality_day_rate=0.002
recovery_day_rate=0.001
recovery_multiplier=10
containment_effectiveness=0.9
cycles=200



**Scenario 2**

starting_infected=1
infected_growth_rate=0.1
awareness=0.001
awareness_growth_rate=0.05
mortality_day_rate=0.01
recovery_day_rate=0.01,recovery_multiplier=3
containment_effectiveness=0.95
cycles=200