Institute and Faculty of Actuaries

# IFoA COVID-19 Action Taskforce (ICAT) Scenario Modelling Workstream

## Forecasting Disease Spread: A Review of Some of the Available Methods Including References

Ashley Campbell

March 2021

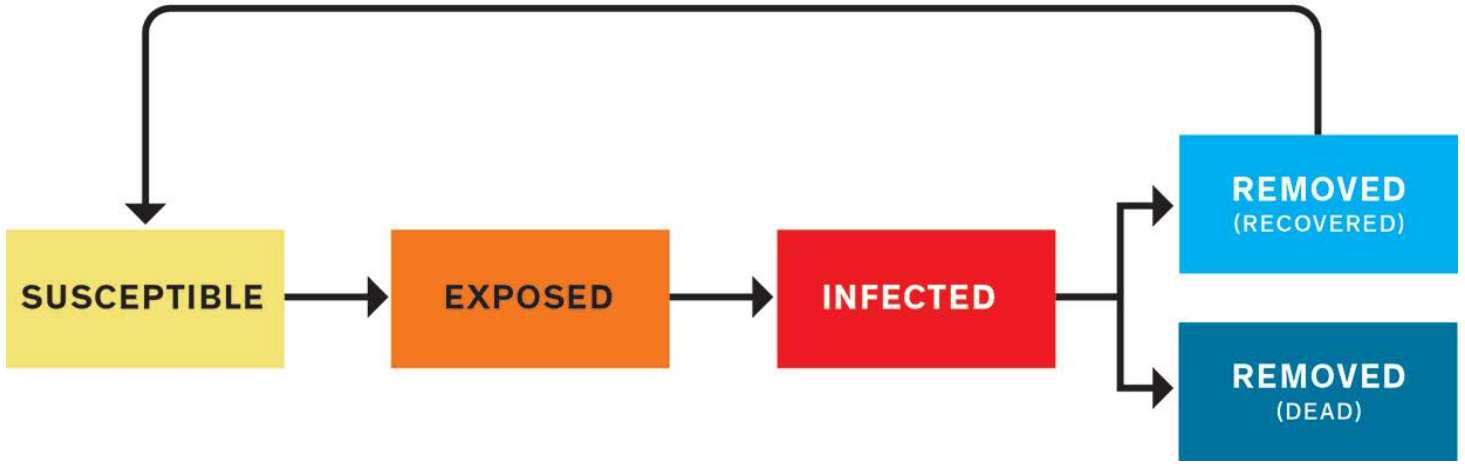
# Content

**A Review of three modelling methodologies for disease spread modelling/forecasting (including References):**

1. Causal / Mechanistic Models - (SEIR) Models
2. Agent-Based / Individual-Based Model
3. Data-Driven Models

# Causal / Mechanistic Models – (SEIR) Models

A SEIRS model puts people into categories: susceptible (S), exposed (E), infected (I), removed from the susceptible population (R), and potentially back to susceptible (S) again, depending on whether a recovered person has immunity from the disease. The modeler's job is to define the equations that determine how people move from one category to the next. Those equations depend on a wide variety of parameters drawn from biology, behavior, politics, the economy, the weather, and more. The example below is a mass action type model, where the population(s) are assumed to be broken up into these homogeneous.




(https://link.springer.com/article/10.1007/s41745-020-00200-6)

# Causal / Mechanistic Models – (SEIR) Models

These models can be further extended by breaking up the population into more homogeneous groups (e.g. by age, sex, occupation, location) to allow for inherent heterogeneity of the underlying populations. This could be imagined as a complex array of vectors or matrices for each compartment above. For these, a significant number of parameters are required to determine mixing patterns in the population. The most complex of these are spatial metapopulation multi-country models, in which the subpopulations are connected through airline and commuter flow networks

(https://link.springer.com/article/10.1007/s41745-020-00200-6)

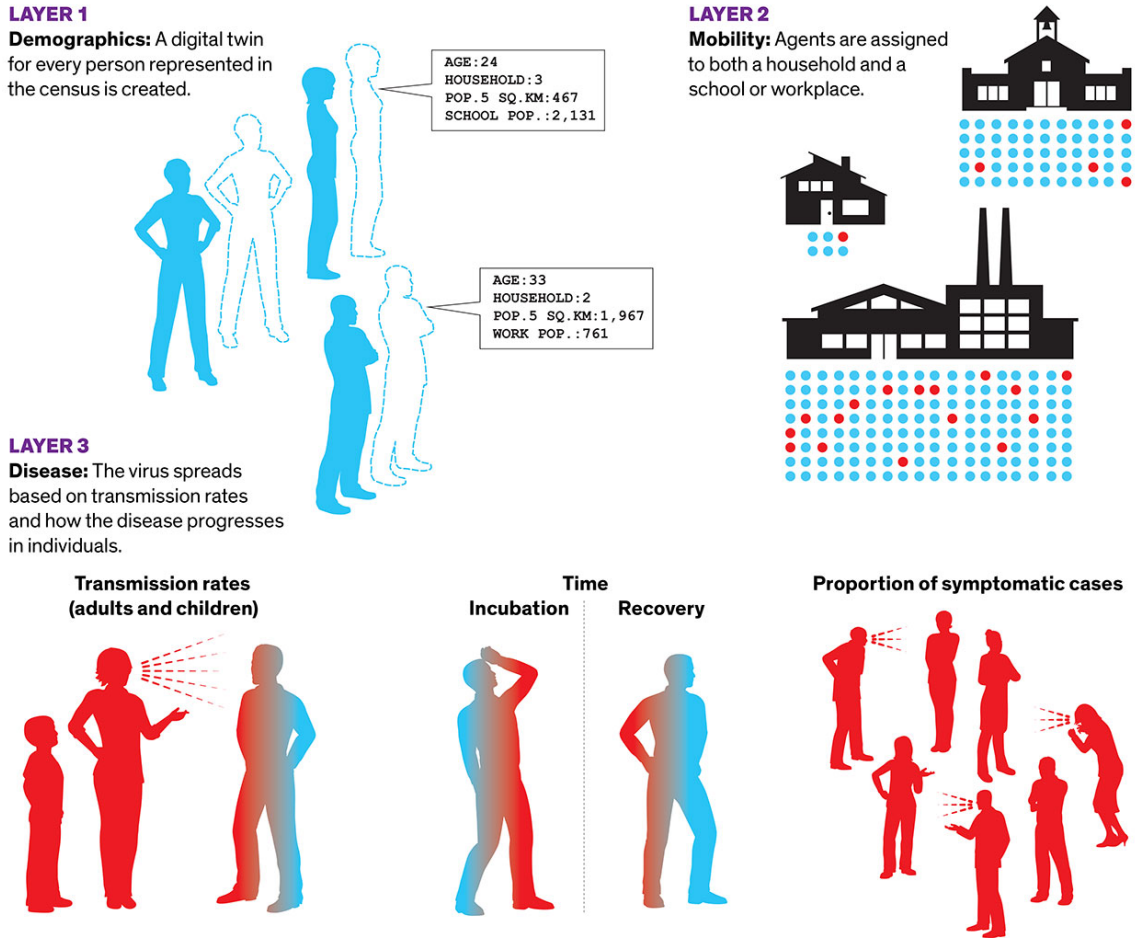
# Causal / Mechanistic Models – (SEIR) Models

## Examples & references:

- (https://link.springer.com/article/10.1007/s41745-020-00200-6)
- Prominent examples of COVID-19 Projections using SEIR models (regularly updated):
- Public Health Authority of Sweden SEIR Model: https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/e/estimates-of-the-number-of-infected-individuals-during-the-covid-19-outbreak/
- GLEAM Metapopulation Multi-country Model: https://covid19.gleamproject.org/
- STEM Metapopulation Model (US) by Iowa State: https://covid19.stat.iastate.edu/longtermproj.html
- Useful tools using this type of modelling approach:
- Complex Compartmental SEIR Models allowing for key determinants of spread such as Asymptomatic, various clinical stanges, hospitalisation and death:
  - Deterministic: https://alhill.shinyapps.io/COVID19seir/
  - Stochastic: http://covid-measures.stanford.edu/
- Deterministic Age Specific Simplified SEIR Model: https://cbdrh.shinyapps.io/covoidance/ (More appropriate for Actuaries)
- Spatio-temporal Metapopulation Models (software):
  - GLEAM: http://www.gleamviz.org/
  - STEM Metapopulation Multi-country Model: https://wiki.eclipse.org/STEM
  -
- Youtube video about simulating COVID-19: https://www.youtube.com/embed/gxAaO2rsdIs
- COVID-19 Resources, Data, peer reviewed parameters and more: https://midasnetwork.us/covid-19/#resources

# Agent-Based / Individual-Based Model

In agent-based models, a simulated world accounts for the actions of all the people in a given population. Researchers at the University of Sydney created a COVID-19 model with three layers: demographics, with a digital twin for every person counted by the census; mobility, with the agents moving among households and schools and offices; and the characteristics of the disease




(https://spectrum.ieee.org/artificial-intelligence/medical-ai/why-modeling-the-spread-of-covid19-is-so-damn-hard)
)

# Agent-Based / Individual-Based Model

## Examples & References

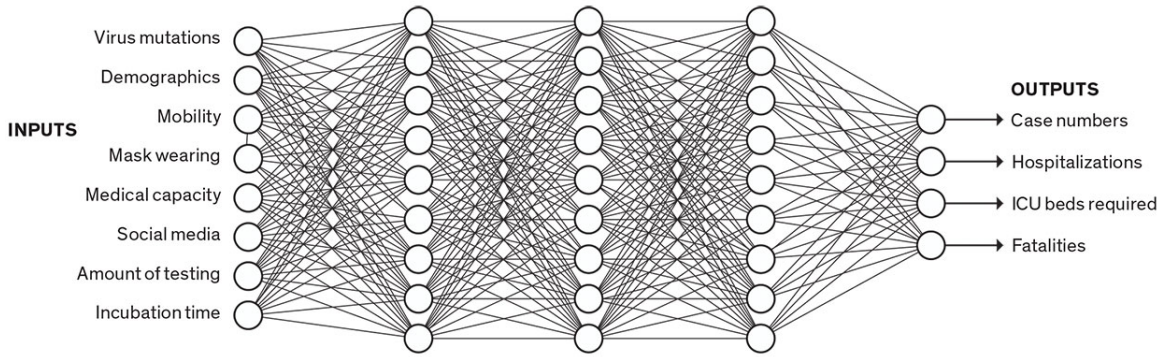
These models are another extension of metapopulation models, They have been used to understand the combined economic and human impact of an outbreak and the government response by designating roles to agents within the economy. These models can help us better understand the effects of compliance to authority within a society.

- Professor Neil Fergusson @ Imperial College's Advanced Individual based SEIR model: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf
- This model was unfairly criticised for unrealistic projections, overestimating mortality and promoting fear tactics. These criticisms were unjustified and to date, Prof Fergusson has been pretty spot on in his estimates, and especially in his approach to the impact of NPIs
- The Sydney Group reported that that social distancing helps very little if only 70 percent of people practice it, but successfully squashes COVID-19 incidence if 80 percent of people can manage it over a span of a few months: https://arxiv.org/abs/2003.10218
- Here is an example open source Agent Based SEIR Model which simulates health and economic effects of social distancing interventions https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7340090/

In general, these models are extremely complex, computationally expensive and data intensive requiring high powered computer clusters. Not only this, but parameterisation of these models require a multi-disciplinary group of experts, adding to the cost.

# Data-Driven Models

Data-driven models don’t sort people into categories; they just crunch numbers. Some data-driven models use neural networks. In this simplified example, a trained neural network infers complex relationships among a broad set of inputs to predict certain outputs. A neural network is a “black box,” as the modeler can’t know or understand the thousands of parameters being used in the prediction.




(https://link.springer.com/article/10.1007/s41745-020-00200-6).

# Data-Driven Models

- It is data-driven models Actuaries may be more familiar with. They do not model the disease dynamics but attempt to find patterns in the available data and combine them appropriately to make short-term forecasts.
- In such data-driven models, it is hard to incorporate interventions directly; hence, the machine is presented with a variety of exogenous data sources such as mobility data, hospital records, etc. with the hope that its effects are captured implicitly.
- Popular purely statistical methods for forecasting influenza-like illnesses (that includes COVID-19) include, e.g., generalized linear models (GLM), autoregressive integrated moving average (ARIMA), and generalized autoregressive moving average (GARMA)
  - (https://link.springer.com/article/10.1007/s41745-020-00200-6).
- Within the community of modelers focused on COVID-19 projections, hybrid approaches seem to be viewed as the state of the art solution. These models aim to combine these data driven techniques with SEIR type models through the use of machine learning.

# Data-Driven Models

## Examples & references:

- The purely statistical: https://covid-19.bsvgateway.org/
- The Hybrid Youyang Gu (YYG) model: https://github.com/youyanggu/covid19_projections
- The Hybrid IHME Model: https://covid19.healthdata.org/united-states-of-america?view=total-deaths&tab=trend

There are a number of other purely statistical models available on the forecasting hub