ASKEM 18-Month Milestone: Hackathon Scenarios

USE CASE: EPIDEMIOLOGICAL MODELING, REGNETS, STOCK AND FLOW

Nirupama Bhattacharya, Zachary Terner, Sam Malloy, Robyn Kozierok

To prepare for the 18-month evaluation scenarios, we have developed a series of hackathon scenarios that are representative of and exercise similar functionality as our target expectations for the evaluation. These questions are meant to help guide and prioritize critical development for success in the evaluation. The goal is to address as much as possible within the Terarium workbench (including the interactive notebook environment). Please be sure to use the logging features of Terarium to ensure that we get accurate timing information.

Table of Contents

Scenario 1: Calibrating models and estimating causal effects

In this scenario, we are interested in determining the effects of masking and social distancing on Covid-19 infections using simulated data for four different countries: Afghanistan, Colombia, France, and the United Kingdom. The simulations use contact matrices and populations subdivided into age groups for each of the countries. In these simulations, each country implements interventions at the same time, and we are interested in understanding how the effects of the interventions differ between countries. The data are generated from an SEIRD model.

In these questions, we provide the contact matrices and population data for each country, as well as the outputs of the simulated SEIRD model. We ask you to calibrate a model to estimate R_0, and to estimate the causal effects of interventions.

In the data provided, 'UK_population.csv' and 'UK_contact_matrix.csv' contain data on the population counts and the contact matrix for each age group in the UK. 'UK_compartments.csv' has compartments S, E, ICase, IMild, R, and D. All patients in IMild transition to R, whereas all patients in ICase transition to the hospital (IHospital, not currently included in this output). From there, some patients recover and others transition to D. 'UK_infections.csv' has total infections for each day. There are analogous files for the other countries.

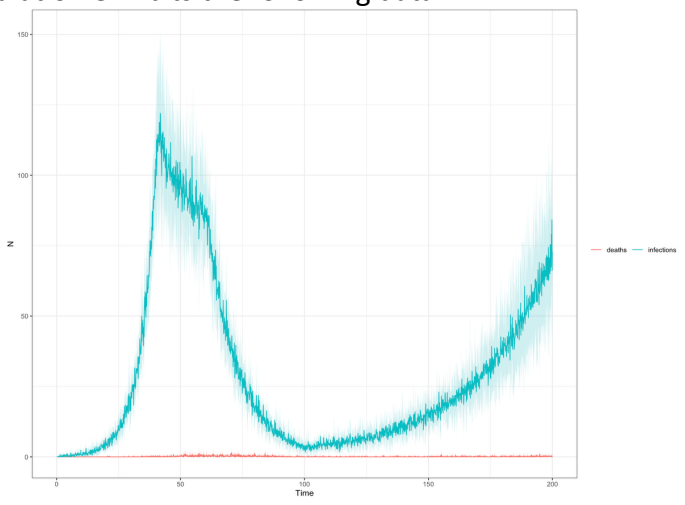
In each country:

1. From t = 0 to t = 40 days, there are no interventions in place, and Covid spreads unabated.
2. At t = 40 days, people begin wearing masks, decreasing the spread of covid.
3. At t = 100 days, mask fatigue sets in. People start to mask less, but more than they did from t = 0 to t = 40. Cases increase again.

During this same time frame, each country begins implementing a social distancing policy at t = 60 days. Due to the policy, contact decreases to 30% of its original levels for the interactions of people over the age of 65 with others over the age of 65 (i.e., the 3x3 matrix on the bottom right of the contact matrix) and it decreases to 60% of its original levels otherwise.

Because this model is stochastic, (in particular, the number of people transitioning between compartments is randomly drawn from a binomial distribution), we run each simulation ten times for each country. The dt parameter is set to 0.2 so that there are 5 time steps per day and 1000 time steps in each 200-day simulation.

For example, one simulation exhibits the following data:



Using the data provided, which includes contact matrices and population counts for all four countries, along with the output from running 10 simulations for each country:

1. Calibrate models to estimate R_0 for each country in each of the four time intervals: [0,40), [40,60), [60,100), [100, 200].
2. Estimate the causal effects of masking alone, the combined effect of masking and social distancing, and the effect of social distancing alone on infections. Include uncertainty in the estimated effects.

(Optional)

3. In each country, what is the maximum value R_0 can be over the last time interval ([100, 200]), to ensure that there are no more than 50 infections at t = 200 days with 90% confidence? Are these different across countries?
4. Can R_0 be changed in one of the three preceding intervals ([0,40), [40,60), [60,100)), without changing it in the fourth, to ensure that infections at t = 200 days stay below ¾ of their simulated value at t = 200 with 95% confidence?

Scenario 2: Mediation Analysis

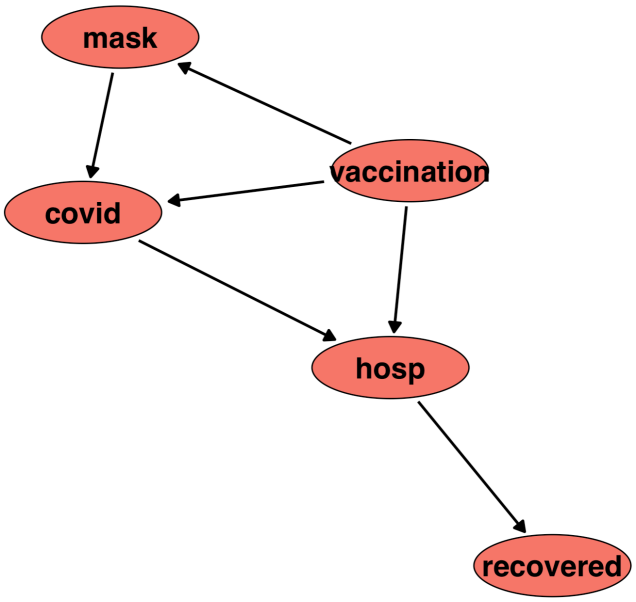


Figure 1. Causal DAG for Scenario 2.

1. The simulated data provided was generated by the causal Directed Acyclic Graph (DAG) in Figure 1 above, which represents the relationships between vaccination status, wearing masks, contracting Covid-19, hospitalization, and recovery. The data were generated via a discrete event simulation. Note that in the causal model above:
   - Whether you wear a mask affects your probability of contracting Covid.
   - Whether you are vaccinated affects your probability of wearing a mask, your probability of contracting Covid, and your probability of getting hospitalized **once you have Covid.**
   - You cannot become hospitalized without contracting Covid first.
   - Each node represents a discrete random variable with two possible values (True, False).
   - Everyone who gets hospitalized fully recovers.

   In the graph above, one should interpret an edge between nodes A and B (A → B) as "the value of A affects the probability of B occurring" and NOT "A must occur for B to occur". In that sense, "vaccination → mask" should be understood as someone's vaccination status affecting the probability that they decide to wear a mask. Similarly, "vaccination → hosp" should be interpreted as someone's vaccination status affects the probability that they become hospitalized (if they contract Covid-19.)

   In this simulation, there are 10000 individuals in a population and a fixed proportion are eligible to be vaccinated every day. The simulation runs for 1000 days. Whether an individual masks or not is influenced by their vaccination status, but it can also change day-to-day.

There is one vaccine in circulation. The vaccine takes effect immediately. The vaccine provides full immunity immediately upon vaccination, but the immunity declines linearly (to no immunity) over the course of 90 days. One cannot be vaccinated a second time.

There are two files associated with this question:

- 'full-*fixed*-vax.csv' contains the full simulation output (in long format), where the proportion of people who are eligible to be vaccinated every day is *fixed*.
- 'daily-summary-*fixed*-vax.csv' takes the output from 'full-*fixed*-vax.csv', groups it by time, and computes sums of those with covid, those hospitalized, those vaccinated, and those recovered every day (i.e. "daily summaries" of those four states).

Using the information and data provided, estimate:

a. The (direct) effect of vaccination on the likelihood of becoming infected with Covid
b. The (indirect) effect of vaccination on the likelihood of becoming infected with Covid
c. The effect of masking as a mediator on the likelihood of becoming infected with Covid
d. The causal effects of vaccination on the likelihood of becoming hospitalized, if infected with Covid

2. In Question 1, a fixed proportion of people are eligible to be vaccinated every day. In this question, the proportion of people who can be vaccinated can change every day. How does this impact the estimated effects above?

There are two files associated with this question, which are analogous to the two described in Question 1 above:

- 'full-*random*-vax.csv' contains the full simulation output (in long format), where the proportion of people who are eligible to be vaccinated every day is *random*.
- 'daily-summary-*random*-vax.csv' takes the output from 'full-*random*-vax.csv', groups it by time, and computes sums of those with covid, those hospitalized, those vaccinated, and those recovered every day (i.e. "daily summaries" of those four states).

Scenario 3 [*Preparation for Decision-making Confidence Metric]:* Supporting decisionmakers at various phases of the Covid-19 pandemic

1. Pretend that it is ***November 1st, 2020***, in the first year of the Covid-19 pandemic, when the main preventive measure was masking. This timepoint is about a month before vaccines first became available in the United States. You are supporting a federal decisionmaker interested in forecasting what might happen over the next few weeks, and what kinds of interventions would need to be put in place to limit negative population health outcomes (number of Covid-19 cases, hospitalizations, and deaths).

    a. **Search and Select Model**: Search for and select an appropriate model for this time period and location (United States country level). The model should be able to support decisionmaker questions about masking and social distancing policies, and their impacts on cases, hospitalizations, and deaths. It should not have irrelevant concepts and variables not relevant to the time period (in other words, there should not be anything related to vaccination and multiple variants of Covid). Model Requirements:
        - can support masking and social distancing interventions
        - outputs cases, hospitalizations, deaths
    b. Please provide information about the literature corpus or git repositories you searched over to find this model.
    c. What are the assumptions, limitations, and strengths of the chosen model?
    d. **Model Comparison**: What are the key differences between the chosen model and one other candidate model from the literature or other sources? In your answer, include:
        - differences in model assumptions
        - comparison of limitations and strengths
        - structural differences
        - explanation of why you selected the chosen model over the alternative
    e. **Model Comparison**: Consider MechBayes, a well-performing model (according to WIS score for forecasted deaths, for November 2020) submitted to the CDC ForecastHub for this time period. See MechBayes code repository and model specification. What are the key differences between the model chosen in 1a, and MechBayes? In your answer, include:
        - differences in model assumptions
        - comparison of limitations and strengths
        - structural differences
    f. Given the differences between your chosen model and the ForecastHub model, how well do you expect your model will perform in comparison, for a near-term forecasting task?
    g. **Find Parameters**: Find relevant parameter values for the chosen model (relevant to this time period and for the United States at a national level), and fill in the following information about sources and quality. If relevant, you may include multiple rows for the same parameter (e.g. perhaps you find different values from different reputable

sources), with a 'summary' row indicating the final value or range of values you decide to use.

Parameter	Parameter Definition	Parameter Units	Parameter Value or Range	Uncertainty Characterization	Sources	Modeler assessment on source quality

h. **Model Extraction**: Extract the chosen model from the source material. Time the entire process to extract the model and curate the results until you are confident the model represented in the workbench is correct.

i. **Single Model Forecast**: Now use the extracted model in the workbench to do a 4-week forecast of cases, hospitalizations, and deaths, from the starting date of November 1st, 2020.
    - For each outcome, plot the 4-week forecast trajectory from your selected model, the 4-week forecast trajectory from the ForecastHub model considered in 1e (MechBayes), and the actual observational data for this period. How do the trajectories compare? For actual observational data to evaluate your forecasts, use the sources described here: https://github.com/reichlab/covid19-forecast-hub/blob/master/data-truth/README.md. You can find forecast data from all models in the ForecastHub here: https://github.com/reichlab/covid19-forecast-hub/tree/master/data-processed
    - In comparing with observational data, calculate Absolute Error and optionally any other error metrics used by the CDC Forecasting Hub, as described here: https://delphi.cmu.edu/forecast-eval/.
    - If your model forecast did not perform as expected, why do you think that is?
    - (Optional) If your model forecast was not optimal, consider going back to 1g, and calibrating some of the more uncertain parameters based on historical data from right before the forecast period.

j. **Ensemble Forecast**: Can you improve upon your single-model forecast in 1i, by creating an ensemble model forecast with the chosen model from 1a, MechBayes, and one other model from the literature that you compared against in 1d? You can parameterize these models, or the combined ensemble model, using any approach you'd like. For each outcome (cases, hospitalizations, deaths), plot the 4-week forecast trajectory from the ensemble model you've created, and compare performance against the observational data, and one of the ForecastHub ensembles (e.g. 'COVIDhub-4_week_ensemble'). You can find forecast data from the ForecastHub ensembles here: https://github.com/reichlab/covid19-forecast-hub/tree/master/data-processed

2. **Comparing and Optimizing Interventions**: Still considering the same timepoint of November 1st, 2020, the decisionmaker you're supporting is exploring masking and social distancing policies.
    - Assume the mask policy will have an 80% compliance rate, and results in 75% decrease in transmissibility parameters for the compliant population.
    - Assume the social distancing policy will have a 50% compliance rate, and results in a 50% decrease in contact rates between individuals in the compliant population.
    - Assume each policy goes into effect right at the start of the period.

    a. What is the impact of each policy by itself, on the trajectories for Covid-19 cases, hospitalizations, and deaths, over the next 8 weeks?
    b. Now assume that you have the flexibility to choose the start dates for these policies. *Considering each policy by itself*, when is the latest each could be implemented, in order to ensure that hospitalizations never exceed a national threshold of 60k during November and December of 2020?
    c. *Now considering a combination of policies*, when is the latest each could be implemented in order to ensure that hospitalizations never exceed a national threshold of 60k during November and December of 2020?
    d. If there was uncertainty in the results, what is the source, and is the growth of uncertainty over time as expected?

3. Fast forward to July 15th, 2021, during the upswing of the Covid wave caused by the arrival of the Delta variant. Vaccines were available at this time. Do not consider specific demographic groups for this question.
    a. **Model Update**: Now update the selected model from Q1 to include vaccinations and be able to support interventions around vaccinations (e.g. incorporate a vaccination policy or requirement, which increases rate of vaccination). Please be sure to use the logging features of Terarium to ensure that we get accurate timing information.
    b. **Model Comparison**: Please explain how the model was updated and how it compares to the original starting model from Q1-2. Provide model comparison diagrams where appropriate.
    c. **Find Parameters**: Considering the updated model with additional variables, and new time period, what is the updated parameter table that you will be using?

Parameter	Parameter Definition	Parameter Units	Parameter Value or Range	Uncertainty Characterization	Sources	Modeler assessment on source quality

    d. **Model Checks**: Implement common sense checks on the model structure and parameter space to ensure the updated model and parameterization makes physical sense. Explain the checks that were implemented. For example, under the assumption that the total population is constant for the time period considered:
        - Demonstrate that population is conserved across all categories. If there are birth and death processes in the model, ensure these are accounted for in your assessment of population conservation.

- Ensure that the total unvaccinated population over all states in the model, can never increase over time, and the total vaccinated population over all states in the model, can never decrease over time.
- What other common-sense checks did you implement? Are there others you would have liked to implement but were too difficult?

e. **Single Model Forecast**: Now use the updated model to do a 4-week forecast of cases, hospitalizations, and deaths, from the new date, July 15th, 2021. How do the results compare with forecasts from MechBayes for the same 4-week time period?

4. **Stratification Challenge**: Still considering the same timepoint as Q3, the decisionmaker you're supporting is exploring targeted vaccination policies to boost vaccination rates for specific subpopulations. To support these questions, you decide to further extend the model by considering several demographic subgroups, as well as vaccination dosage. Stratify the model by the following dimensions:
    - Vaccination dosage (1 or 2 doses administered)
    - Age group
    - Sex
    - Race/Ethnicity

    To inform initial conditions and rates of vaccination, efficacy of vaccines, etc., consider the subset of vaccination datasets from the starter kit listed in 'Scenario3_VaccinationDatasets.xlsx'. Where initial conditions are not available for a specific subgroup, make a reasonable estimate based on percentages from Census sources (e.g. https://www.census.gov/quickfacts/fact/table/US/PST045223). Where parameters for specific subgroups are unavailable, generalize based on the ones that are available. Choose the number of age and race/ethnicity groups based on the data that is available.

Scenario 4: Microbial Regnets

The human gut microbiome is an increasingly important area of study for multiple infectious diseases and chronic conditions, including COVID-19 and long COVID. However, characterizing the dynamics of microbial ecology, which includes the interactions between microbes and with their environment, is a major challenge. This scenario is a proof-of-concept to demonstrate that ASKEM can utilize regulatory network models to grapple with such questions.

1. Consider a simple system of 4 microbial species with a variety of competitive and mutualistic interactions amongst each other. Given a set of initial microbial populations, a set of inferred growth rates for each species, and an interaction matrix representing the negative (competitive or inhibitory) or positive (mutualistic or beneficial) effects that the presence of one species has on another in the microbial ecosystem… construct a model of the system utilizing the generalized Lotka-Volterra equation below:

$$\frac{dx_i}{dt} = x_i\left(r_i + \sum_{j=1}^{n} a_{ij}x_j\right)$$

This system represents a community of n species, where x_i is the population of species i, r_i is the intrinsic growth rate of species i, and a_{ij} represents the interaction coefficient between species i and j.

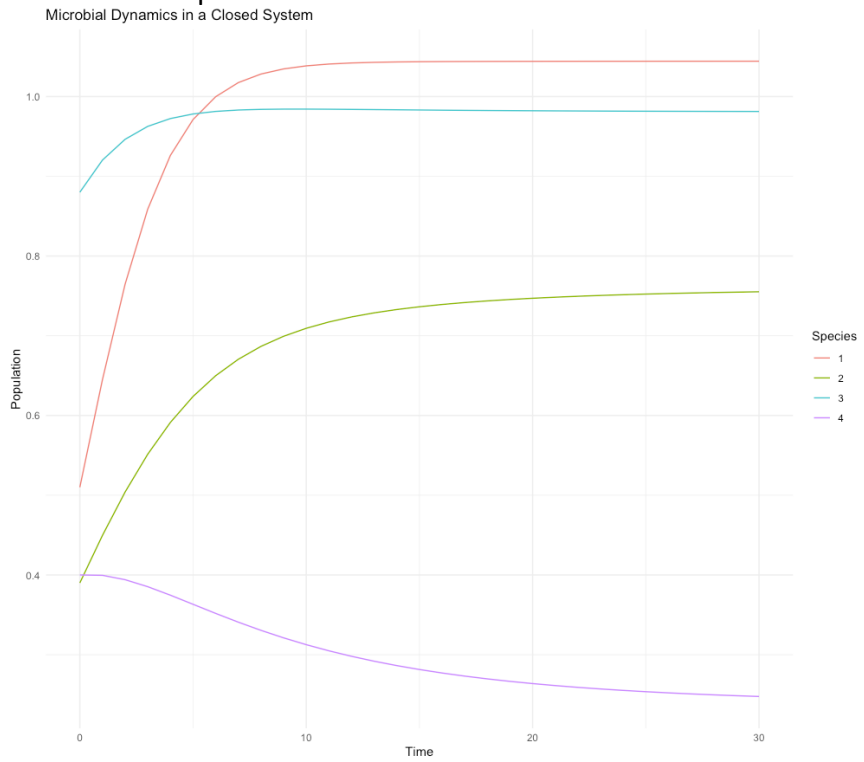
2. Simulate the interactions between each species for a 30-day period, with time steps of 1 day. Plot the population size of each species over the 30-day time period. Use the following parameter and initial conditions data to configure the model:

	Species 1	**Species 2**	**Species 3**	**Species 4**
Growth rate (CFU/day)	0.53	0.42	0.49	0.33
Initial conditions (in CFU/arbitrary unit area)	0.51	0.39	0.88	0.4

Interaction matrix (effect on growth rate, CFU/day)

	Species 1	**Species 2**	**Species 3**	**Species 4**
Species 1	-0.5	-0.01	0.002	-0.009
Species 2	0	-.5	0	-0.169
Species 3	-0.002	-0.003	-0.5	0.02
Species 4	0	-0.226	-0.04	-0.5

The output should match the plot below.

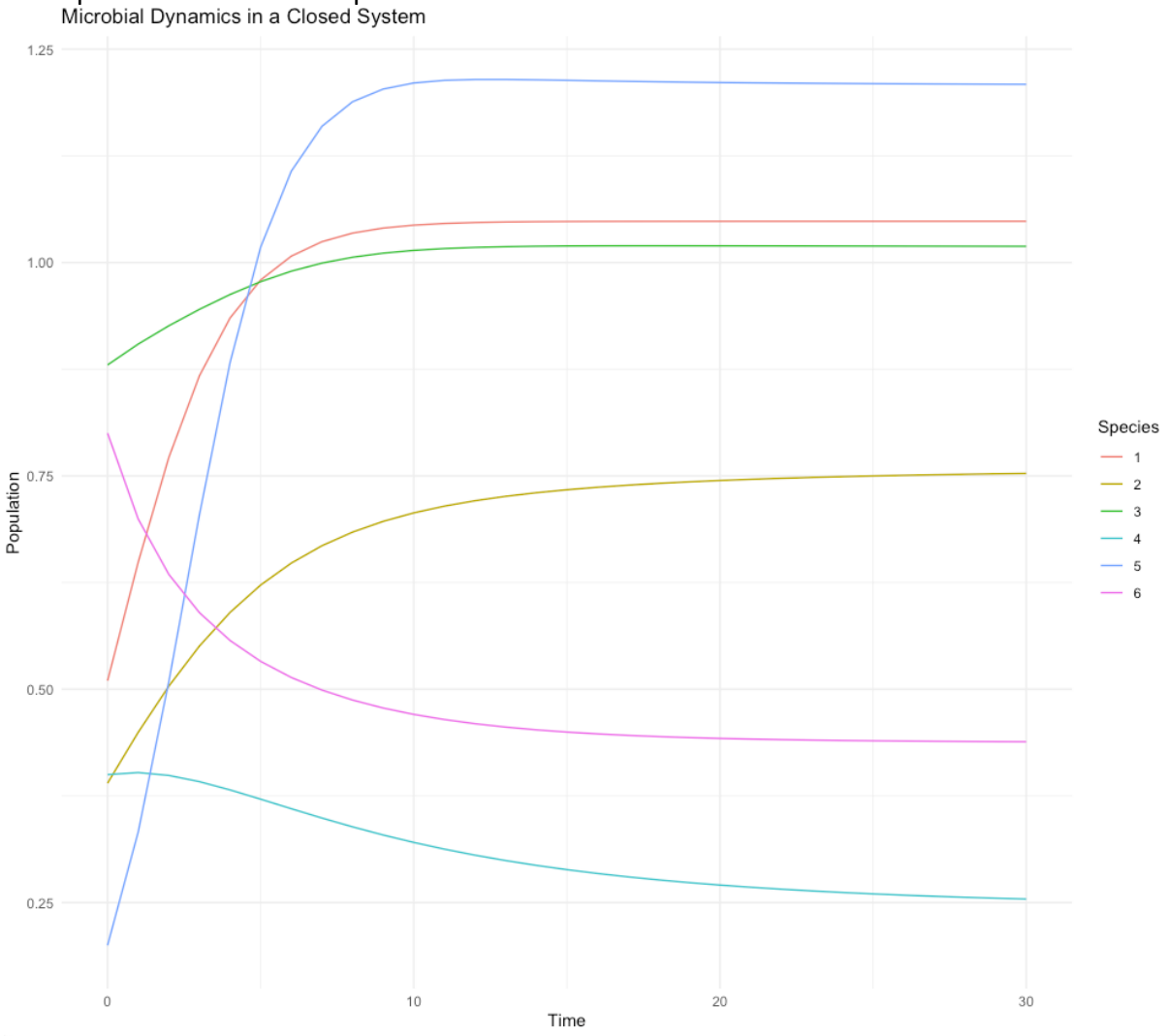


3. Update the model to reflect a six species model, using the parameter and initial conditions data below:

	Species 1	**Species 2**	**Species 3**	**Species 4**	**Species 5**	**Species 6**
Growth rate (CFU/day)	0.53	0.42	0.49	0.33	0.7	0.3
Initial conditions (in CFU/arbitrary unit area)	0.51	0.39	0.88	0.4	0.2	0.8

	Species 1	**Species 2**	**Species 3**	**Species 4**	**Species 5**	**Species 6**
Species 1	-0.5	-0.01	0.002	-0.009	-0.002	0.01
Species 2	0	-0.5	0	-0.169	0	0
Species 3	-0.002	-0.003	-0.5	0.02	0.03	-0.04
Species 4	0	-0.226	-0.04	-0.5	0	0.01
Species 5	0	-0.1	-0.02	0	-0.5	0
Species 6	0	-0.04	-0.05	0	0	-0.5

The output should match the plot below:

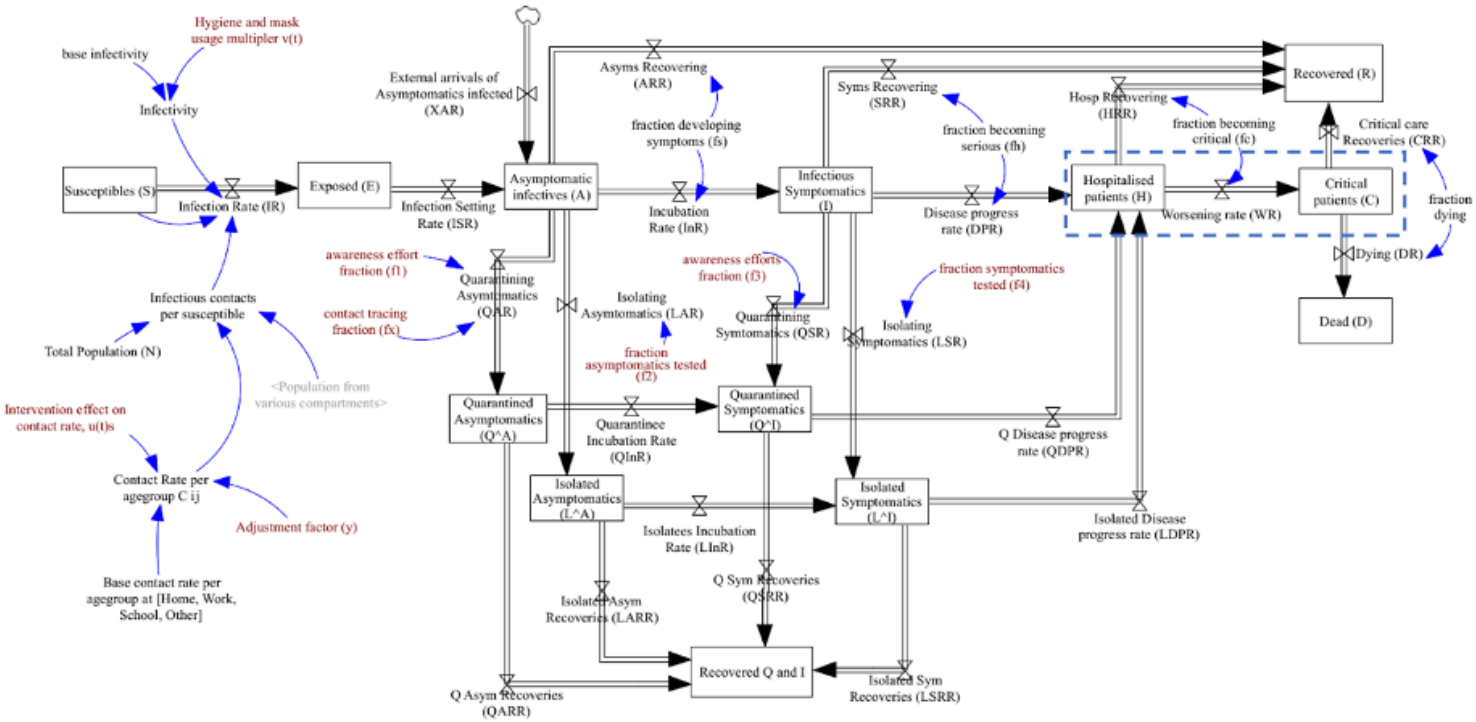


Scenario 5: Stock and flow testing and diagnostics model (relevant for ASPR demo)

One of the major challenges that ASKEM aims to address is the limited reproducibility of models that are created using proprietary or specialized modeling software. Within the epidemiological modeling use case, the use of proprietary modeling software is particularly common within the system dynamics ("stock and flow") component of the modeling ecosystem. System dynamics models are useful for exploring multifaceted scenarios, such as supply chain connections with disease dynamics.

The goal of this scenario is to demonstrate that ASKEM can ingest and simulate a model that is representative of the types of system dynamics approaches that were implemented in the early stage of the COVID-19 pandemic.

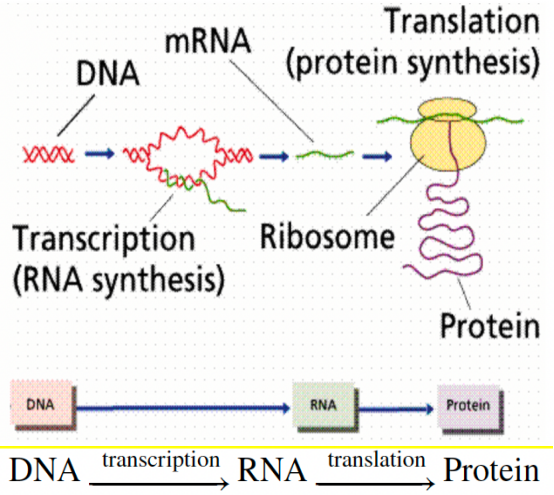
1. Ingest the file "IndiaNonSubscriptedPulsed.mdl", which contains a system dynamics model that was used to assess the effectiveness of testing, isolation, and quarantine on the potential trajectory of COVID-19 in early 2020. The preprint for the model, *Effectiveness of Testing, Tracing, Social Distancing and Hygiene in Tackling COVID-19 in India: a System Dynamics Model* is available at: arXiv:2004.08859
2. Inspect the flow diagram to ensure that it is consistent with the model as implemented in its original environment (the modeling platform Vensim). The key stocks, and flows between them, should match the diagram in Figure 2:



3. Simulate the model using the base configurations in the .mdl file. The simulation output for the base case should indicate a peak around day 250, at 5 million cases/day.
4. Demonstrate that you can adjust rates in the simulation configuration, such as "default delay disease diagnosis" and "testing impact on delay", which affect the flow into hospital demand.
5. In their original documentation, the developers of this model expressed an interest in introducing uncertainty into model components related to testing. The accuracy of tests in the base configuration of the model is assumed to be absolute. Demonstrate that you can introduce uncertainty into (1) the accuracy of test results and (2) the amount of time it takes to receive a test result. Show that you can set multiple model configurations to fully explore the uncertainty range in the relevant parameters, and show how simulating across this range affects the peak day and peak caseload.
6. Demonstrate that you can modify the model structure by adding an additional testing modality, rapid antigen tests. The base configuration assumes only one type of testing with one set of rates, and adding antigen-based testing will allow a modeler to explore the proportion of fast, but less reliable, antigen-based tests, compared with slower but more reliable nucleic acid amplification tests. Include the uncertainty added in step 5 above for both test types. *Note: this step is meant to be exploratory; there is no "correct" answer other than demonstrating that the ingested SD model structure can be meaningfully modified in the workbench.*
7. Consider extending the model to account for other structural features not included in the original model, such as vaccination and the potential for reinfection. Incorporate uncertainty in the new features using evidence drawn from literature review in the workbench. *Note: this step is meant to be exploratory; there is no "correct" answer other than demonstrating that the ingested SD model structure can be meaningfully modified in the workbench.*

Scenario 6: Regnets for gene expression and regulation

A fundamental theory in molecular biology, known as the 'the central dogma', describes gene expression as the flow of information from DNA to RNA (through a step called transcription), and from RNA to proteins (through a step called translation). The processes of transcription, translation, and protein degradation, depend on the amount of RNA (specifically, mRNA), and proteins, and mathematically can be described as the rate of change of the concentrations of these materials. This scenario will explore several simple models of gene expression and regulation.



For this scenario, we consider the following models.

Chen Model: An early linear transcription model of gene expression was proposed by Chen et. al. in 1999, which considered only a feedback loop from proteins, to transcription, but ignores any feedback to translation. For the purpose of this scenario, we make the following modifications to the original model:

- We consider three types of mRNA and proteins
- Type 1 mRNA produce Type 1 proteins, which are involved in the transcription of all types of mRNA.
- Type 2 mRNA produce Type 2 proteins, and Type 3 mRNA produce Type 3 proteins. Type 2 and 3 proteins are not involved in transcription.

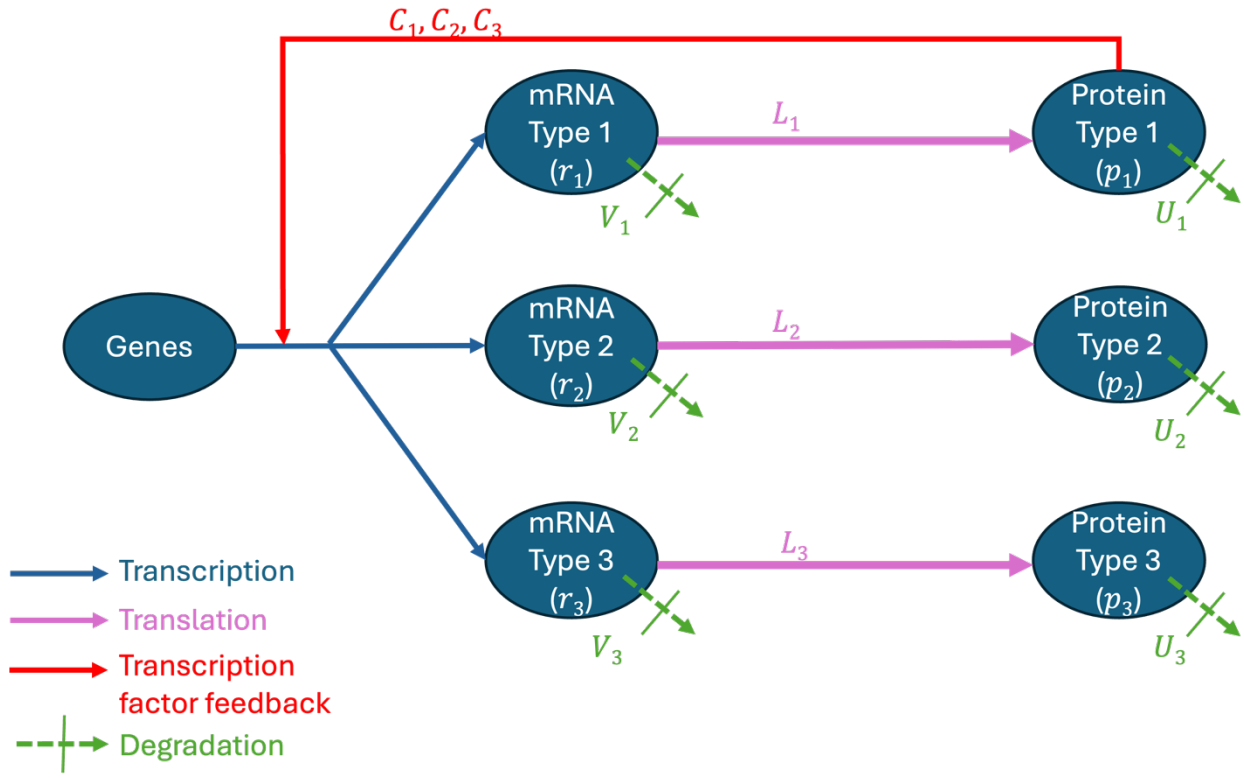


Figure 33. Modified version of the model by Chen et. al., which describes a dynamic system of gene regulation with feedback on transcription, but not translation.

The modified Chen model is described in Figure 33 and can be expressed by the following set of differential equations:

$$\frac{dr_1}{dt} = C_1 p_1 - V_1 r_1$$
$$\frac{dr_2}{dt} = C_2 p_1 - V_2 r_2$$
$$\frac{dr_3}{dt} = C_3 p_1 - V_3 r_3$$
$$\frac{dp_1}{dt} = L r_1 - U_1 p_1$$
$$\frac{dp_2}{dt} = L_2 r_2 - U_2 p_2$$
$$\frac{dp_3}{dt} = L_3 r_3 - U_3 p_3$$

Where:

- r_i represent the concentration of Type i mRNA, in units *nMolar*
- p_i represent the concentration of Type i protein, in units *nMolar*
- C_i represents the transcription rates for Type i mRNA, in units $\frac{nMolar\ of\ mRNA}{nMolar\ of\ protein \cdot minute}$
- L_i represents the translation rates for Type i protein, in units $\frac{nMolar\ of\ protein}{nMolar\ of\ mRNA \cdot minute}$

- V_i is the natural degradation rate of Type i mRNA, in units $\frac{1}{minute}$
- U_i is the natural degradation rate of Type i protein, in units $\frac{1}{minute}$

Hunt Model: In this scenario, we also consider an extended model of gene expression, proposed by Hunt et. al., which incorporates multiple additional feedback loops into the protein synthesis process:

- Type 2 proteins help stabilize all mRNA against degradation (in effect lowers degradation rate)
- Each type of protein regulates the production of its own protein type, during both transcription and translation processes

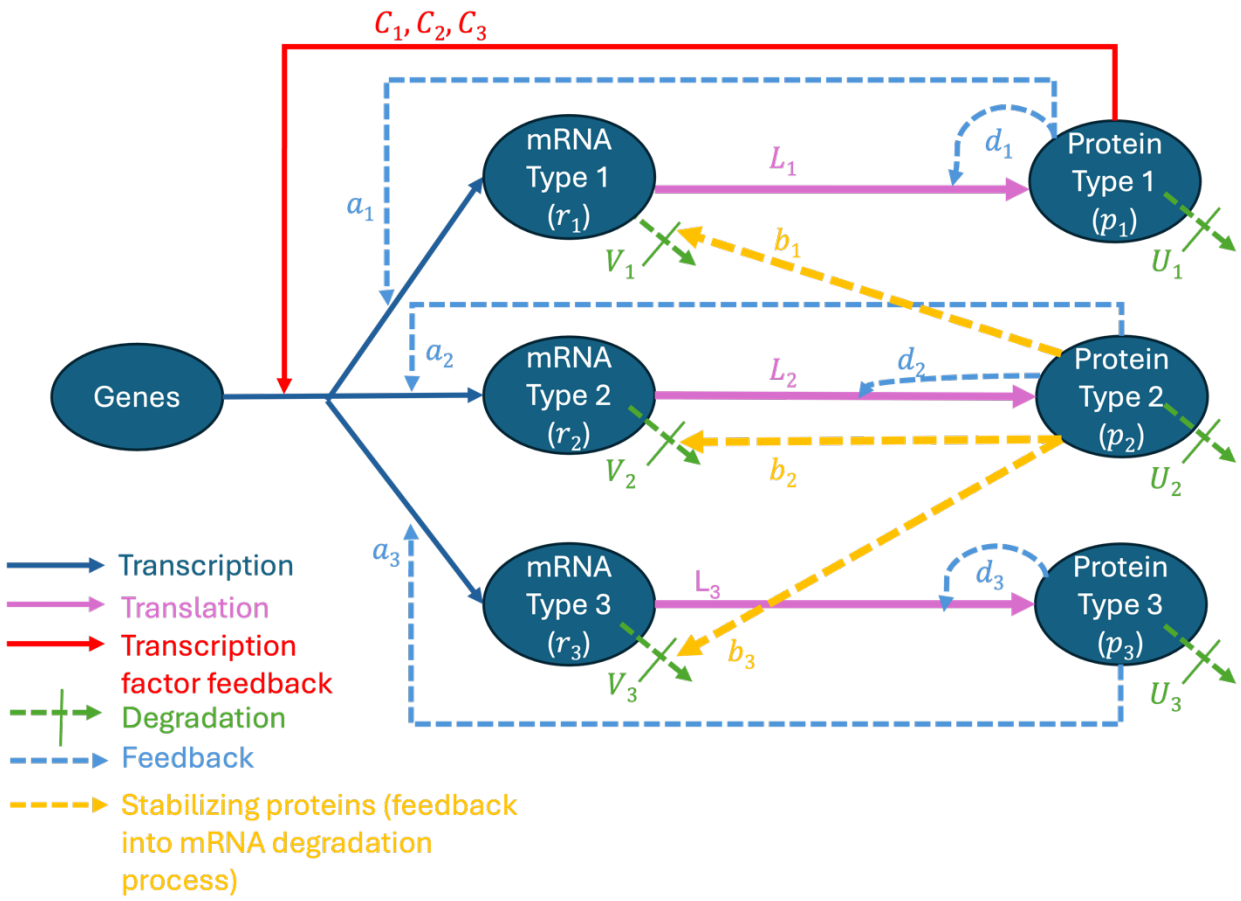


Figure 44. Representation of the Hunt Model.

The Hunt Model is described in Figure 44, and defined by the following set of differential equations:

$$\frac{dr_1}{dt} = \frac{1}{1+\frac{p_1^2}{a_1^2}} C_1 p_1 - \frac{1}{1+\frac{p_2}{b_1}} V_1 r_1$$

$$\frac{dr_2}{dt} = \frac{1}{1+\frac{p_2}{a_2}} C_2 p_1 - \frac{1}{1+\frac{p_2}{b_2}} V_2 r_2$$
$$\frac{dr_3}{dt} = \frac{1}{1+\frac{p_3}{a_3}} C_3 p_1 - \frac{1}{1+\frac{p_2}{b_3}} V_3 r_3$$
$$\frac{dp_1}{dt} = \frac{1}{1+\frac{p_1}{d_1}} L_1 r_1 - U_1 p_1$$
$$\frac{dp_2}{dt} = \frac{1}{1+\frac{p_2}{d_2}} L_2 r_2 - U_2 p_2$$
$$\frac{dp_3}{dt} = \frac{1}{1+\frac{p_3}{d_3}} L_3 r_3 - U_3 p_3$$

Where:

- r_i represent the concentration of Type i mRNA, in units *nMolar*
- p_i represent the concentration of Type i protein, in units *nMolar*
- C_i represents the transcription rates for Type i mRNA, in units $\frac{nMolar\ of\ mRNA}{nMolar\ of\ protein \cdot minute}$
- L_i represents the translation rates for Type i protein, in units $\frac{nMolar\ of\ protein}{nMolar\ of\ mRNA \cdot minute}$
- V_i is the natural degradation rate of Type i mRNA, in units $\frac{1}{minute}$
- U_i is the natural degradation rate of Type i protein, in units $\frac{1}{minute}$
- a_i is an effectiveness factor for feedback on transcription (i.e. suppresses the rate of transcription), for Type i mRNA. The bigger the value of a_i, the smaller the impact of feedback and the less the transcription rate is decreased. The units are *nMolar.*
- d_i is an effectiveness factor for feedback on translation (i.e. suppresses the rate of translation), for Type i proteins. The bigger the value of d_i, the smaller the impact of feedback and the less the translation rate is decreased. The units are *nMolar.*
- b_i is an effectiveness factor for feedback on mRNA degradation (i.e. Type 2 proteins act as stabilizers that lower the rate of degradation, for Type i mRNA). The bigger the value of b_i, the smaller the impact of feedback and the less the degradation rate is decreased. The units are *nMolar.*

For the following questions, use the initial and parameter values indicated in Table 1.

Table 1. Initial and parameter values for Questions 1-2

i	$r_i(0)$ (nM)	$p_i(0)$ (nM)	C_i $\left(\frac{nM\ of\ mRNA}{nM\ of\ protein \cdot min}\right)$	L_i $\left(\frac{nM\ of\ protein}{nM\ of\ mRNA \cdot min}\right)$	$V_i\left(\frac{1}{min}\right)$	U_i $\left(\frac{1}{min}\right)$	a_i (nM)	b_i (nM)	c_i (nM)
1	3	100	0.03	2	0.03	0.15	60	120	120
2	6	500	0.03	2	0.03	0.15	140	140	150
3	5	1	0.024	2	0.03	0.015	170	180	260

1. **Chen Model**: Using the relevant values in Table 1 (not all will be applicable), simulate the Chen Model for a time period of 10 hours, and plot the trajectories of all mRNA and protein types. Comment on the nature of the trajectories and biological realism of the outcomes. Do the trajectories make sense given the model?

2. **Simulations with the Hunt Model**: For all parts, use the relevant values in Table 1, simulate for a time period of 10 hours, and plot the trajectories of all mRNA and protein types.
    a. Simulate the full Hunt Model. How do the outcomes compare with Q1?
    b. Now consider a Special Case 1, where we assume $b_i, \gg p_2 \; and \; d_i \gg p_i$ (i.e., the stabilization of mRNA, and feedback from proteins to translation, are negligible). What does the set of differential equations reduce to? Repeat the simulation for this Case 1 Hunt Model. How do the outcomes compare with Q2a (the Full Hunt Model)? Do the results make sense given the assumptions made?
    c. Consider Special Case 2, where we assume $a_i, d_i \; \gg pi$ (i.e., assume the feedback from proteins to transcription and translation, is negligible). What does the set of differential equations reduce to? Simulate this Case 2 Hunt Model. How do the outcomes compare with Q2a,b? Do the results make sense given the assumptions made?
    d. Consider Special Case 3, where we assume $a_i, \gg p_i \; and \; b_i \gg p_2$ (i.e., assume the feedback from proteins to transcription and stabilization of mRNA, are negligible). What does the set of differential equations reduce to? Simulate this Case 3 Hunt Model. How do the outcomes compare with Q2a-c? Do the results make sense given the assumptions made?
    e. Consider Special Case 4, where we assume $d_i, \gg p_i$ (i.e., assume the feedback from proteins to translation is negligible). What does the set of differential equations reduce to? Simulate this Case 4 Hunt Model, and if needed, extend the simulation length to ensure you can determine whether the system reaches an equilibrium state or not. How do the outcomes compare with Q2a-d? Do the results make sense given the assumptions made?

3. For Questions 1-2, some of the simulation outcomes should include exponential growth trajectories. For those cases, explore changes in Table 1 parameters that will result in changing the outputs from exponential growth trajectories to more bounded behavior. What is the biological interpretation of these parameter changes? (Note that there are multiple sets of parameters that will fulfill this criteria, and there is no single correct answer).

4. (Optional Challenge). For one or more of the systems considered in Q1-2, do a formal stability analysis (i.e., determine equilibrium points, their stability, and their biological interpretation).