ASKEM Final Evaluation (February 2025):
Epidemiology Independent Questions (IQ) Part 2

Nirupama Bhattacharya, Robyn Kozierok, Zachary Terner, Jack Mitcham, Molly R. Creagar

The views, opinions, and/or findings contained in this report are those of The MITRE Corporation and should not be construed as an official Government position, policy, or decision, unless designated by other documentation. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004.

Each Set will have 2-3 questions of varying difficulty/complexity. All problems are intended to be independent of each other, except for the Stratification questions (in Part 1) which build upon one another, from simple to complex.

Table of Contents

Section 3: Technical Area 3 Functionality 3

Set 3.1: Parameter Calibration 3
- ***General format of inputs, tasks, and outputs*** 3
- **(Q3.1.1)** ***Set 3.1, Question 1: Simple Calibration for 1 Parameter*** 3
- **(Q3.1.2)** ***Set 3.1, Question 2: Medium Complexity Calibration for 2 Parameters*** 5
- **(Q3.1.3)** ***Set 3.1, Question 3: High Complexity Calibration for 4 Parameters*** 6
- ***Set 3.1 Summary Table*** 8

Set 3.2: Forecasting 8
- ***General format of inputs, tasks, and outputs*** 9
- **(Q3.2.1)** ***Set 3.2, Question 1: Simple simulation and forecast*** 9
- **(Q3.2.2)** ***Set 3.2, Question 2: Medium Complexity Forecast*** 10
- **(Q3.2.3)** ***Set 3.2, Question 3: High Complexity Forecast*** 11
- **Set 3.2** ***Summary Table*** 13

Set 3.3: Interventions 14
- ***General format of inputs, tasks, and outputs*** 14
- **(Q3.3.1)** ***Set 3.3, Question 1: Simple Intervention*** 14
- **(Q3.3.2)** ***Set 3.3, Question 2: Medium Complexity Intervention*** 15
- **(Q3.3.3)** ***Set 3.3, Question 3: High Complexity Intervention*** 15
- **Set 3.3** ***Summary Table*** 16

Set 3.4: Optimizing Interventions 17
- ***General format of inputs, tasks, and outputs*** 17
- **(Q3.4.1)** ***Set 3.4, Question 1: Simple Optimization of Intervention*** 17
- **(Q3.4.2)** ***Set 3.4, Question 2: Medium Complexity Optimization of Intervention*** 18
- **(Q3.4.3)** ***Set 3.4, Question 3: Challenging Optimization*** 19
- **Set 3.4** ***Summary Table*** 20

Set 3.5: Ensembles 21
- ***General format of inputs, tasks, and outputs*** 21
- **(Q3.5.1)** ***Set 3.5, Question 1: Simpler Ensemble*** 21
- **(Q3.5.2)** ***Set 3.5, Question 2: Challenging Ensemble*** 22
- **Set 3.5** ***Summary Table*** 22

Section 4: Technical Area 4 Functionality ***23***

Set 4.3: Template Scenarios **23**

General format of inputs, tasks, and outputs 23

(Q*4.3.1) Set 4.3, Question 1: Situational Awareness and Intervention* 23

(Q4.3.2) *Set 4.3, Question 2: Sensitivity Analysis* 24

(Q*4.3.3)* Set 4.3, Question 3: Horizon Scanning 25

Set 4.*3 Summary Table* 26

Section 3: Technical Area 3 Functionality

Set 3.1: Parameter Calibration

Calibrate one or more parameters for a provided model based on data.

Take a given model and dataset and calibrate for requested parameter(s) and compute the error between data and calibrated model output. Problems will vary in complexity from calibrating a single parameter, to simultaneously calibrating multiple parameters.

General format of inputs, tasks, and outputs

Inputs	Tasks	Outputs
• Input model to be calibrated • Dataset and indication of which features to use for calibration. The data sets in this set of questions contain data on Influenza A cases in one influenza season in the United States. • Model configuration, including initial conditions and all parameter values other than those to be calibrated. Known parameters that won't be calibrated, may still have a range of values. • List of parameters to be calibrated with prior distributions for each	• Calibrate model parameter(s) using provided data set • Produce required error measures	• For each calibrated parameter, provide mean parameter value(s) pre- and post-calibration, prior and posterior distributions, and variance pre- and post-calibration • Some measure of goodness-of-fit for the calibrated parameters, such as mean absolute error (MAE) between projected value for one or more state variables, and observational data from input dataset

(Q3.1.1) ***Set 3.1, Question 1**: Simple Calibration for 1 Parameter*

Inputs:

- (Baseline) Model in RMarkdown file titled "Set 3", section Q3.1.1
- (Workbench) "seirh" model in the workflow titled "Q3.1.1"
- "dataset_Q3.1.1" (title of dataset in workbench and csv file)

Begin with the relevant input SEIRH model listed below. The model diagram and set of equations is described below here:

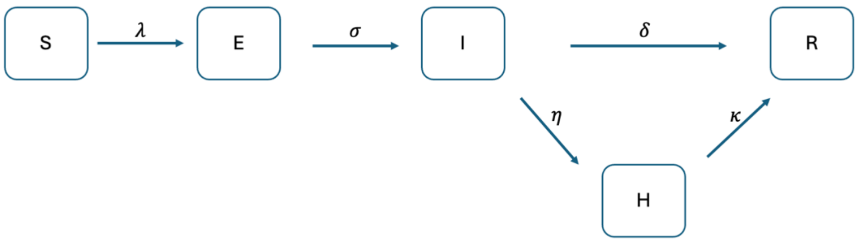

Figure 1 SEIRH model diagram

The equations for the model are given by:

$$\frac{dS}{dt} = -\lambda S$$
$$\frac{dE}{dt} = \lambda S - \sigma E$$
$$\frac{dI}{dt} = \sigma E - \eta I - \delta I$$
$$\frac{dR}{dt} = \delta I + \kappa H$$
$$\frac{dH}{dt} = \eta I - \kappa H$$

Initial state variables:

- S = 50000; susceptible population
- E = 200; latent population
- I = 10; infectious population
- R = 0; recovered population
- H = 0; hospitalized population

All parameter values except for η are known; parameters are defined as the following:

- $\lambda =$ 0.08; force of infection
- $\sigma =$ 0.13; infection rate
- $\delta =$ 0.263; recovery rate, 1/infectious period in days
- $\eta =$ [*unknown, use initial range 0 to 0.04*]; rate of hospitalization
- $\kappa =$ 0.15; rate of recovery after hospitalization, 1/average days in hospitalization

Tasks:

- Use the input dataset to calibrate the parameter η in the SEIRH model using the "new cases" and "new hospitalizations" columns.

Outputs:

- Calibrated parameter η
- Error measure between projected value and data

(Q3.1.2) *Set 3.1, Question 2: Medium Complexity Calibration for 2 Parameters*

Inputs:

- (Workbench) "SEIRHVD" model in workflow "Q3.1.2"
- (Baseline) RMarkdown file titled "Set 3", section "Q3.1.2"
- "dataset_Q3.1.2" (title of dataset in workbench and csv file)

Below is a diagram of the model:

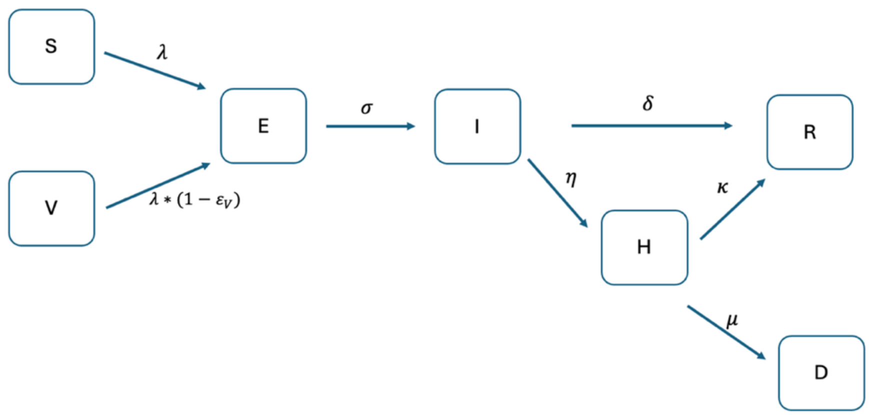

Figure 2. SEIRHVD model diagram

The equations for the model are given by:

$$\frac{dS}{dt} = -\lambda S$$

$$\frac{dE}{dt} = \lambda S + (1-\varepsilon_V)\lambda V - \sigma E$$

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

$$\frac{dR}{dt} = \delta I + \kappa H$$

$$\frac{dH}{dt} = \eta I - \kappa H - \mu H$$

$$\frac{dV}{dt} = -(1-\varepsilon_V)\lambda V$$

$$\frac{dD}{dt} = \mu H$$

Initial state variables:

- S = 50000; susceptible population
- E = 150; latent population
- I = 10; infectious population
- R = 0; recovered population
- H = 0; hospitalized population

- V = 1000; vaccinated population
- D = 0; deceased population

All parameter values except for η and μ are known; parameters are defined as:

- $\lambda =$ 0.09; force of infection
- $\sigma =$ range [0.1-0.16]; infection rate
- $\delta =$ 0.26; recovery rate, 1/ infectious period in days
- $\eta =$ *[unknown, to be calibrated, use initial range 0 to 0.05]*; rate of hospitalization
- $\kappa =$ range [0.09-0.16]; rate of recovery from hospitalization, 1/average days in hospitalization
- ε_V = 0.8765; vaccination efficacy
- $\mu =$ *[unknown, to be calibrated, use initial range 0 to 0.03]*; death rate

Tasks:

- Use the input dataset to calibrate the parameters η, μ in the SEIRHVD model using the "new cases", "new hospitalizations", and "new deaths" columns.

Outputs:

- Calibrated parameters η, μ
- Error measure between projected values and data

(Q3.1.3) *Set 3.1, Question 3: High Complexity Calibration for 4 Parameters*

Record how long it takes you to complete the following task:

Inputs:

- (Workbench) SEIRVTHD model and workflow titled "Q3.1.3"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.1.3
- "dataset_Q3.1.3" (title of dataset in workbench and csv file)

Below is a diagram of the model:

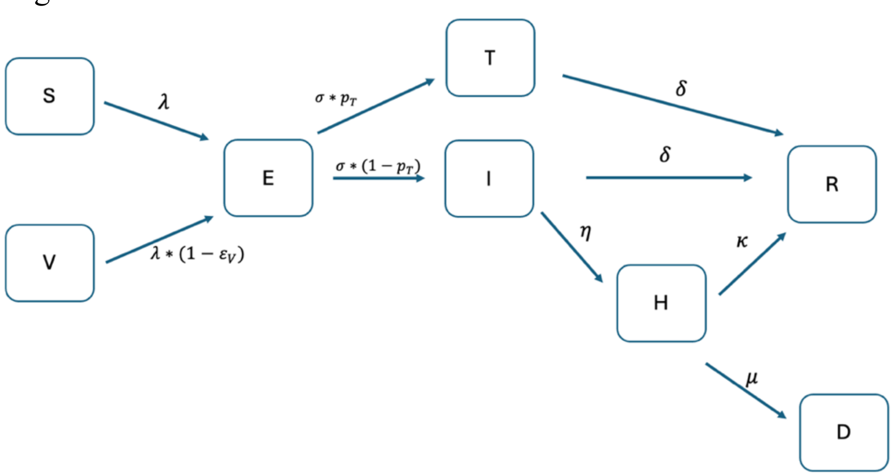

Figure 3. SEIRVTHD model diagram

The equations for the model are given by:

$$\frac{dS}{dt} = -\lambda S$$

$$\frac{dE}{dt} = \lambda S + (1-\varepsilon_V)V - \sigma E$$

$$\frac{dV}{dt} = -(1-\varepsilon_V)\lambda V$$

$$\frac{dI}{dt} = (1-p_T)\sigma E - \delta I - \eta I$$

$$\frac{dT}{dt} = p_T \sigma E - \delta T$$

$$\frac{dH}{dt} = \eta I - \kappa H - \mu H$$

$$\frac{dR}{dt} = \delta T + \delta I + \kappa H$$

$$\frac{dD}{dt} = \mu H$$

Initial state variables:

- S = 40000; susceptible population
- E = 50; latent population
- I = 8; infectious population
- R = 0; recovered population
- H = 0; hospitalized population
- V = 1000; vaccinated population
- D = 0; deceased population
- T = 0; population of latent individuals treated with antivirals to reduce infectiousness

The known and unknown parameters are defined as:

- $\lambda =$ 0.14; force of infection
- $\sigma =$ 0.125; infection rate
- $\delta =$ 0.25; recovery rate, 1/ infectious period in days
- $\eta =$ [*unknown, to be calibrated, use initial range 0 to 0.04*]; rate of hospitalization
- $\kappa =$ 0.13; rate of recovery from hospitalization, 1/average days in hospitalization
- $\varepsilon_V =$[*unknown, to be calibrated, use initial range 0.82 to 0.95*]; vaccination efficacy
- $p_T =$[*unknown, to be calibrated, use initial range 0.25 to 0.45*]*;* proportion of latent individuals treated with antivirals
- $\mu =$ [*unknown, to be calibrated, use initial range 0 to 0.04*]; death rate

Tasks:

- Use the input dataset to calibrate the parameters η, μ, ε_V, p_T in the SEIRHVTD model using the “cases”, “hospitalizations”, “treated”, “deaths”, and “vaccinated” columns.

Outputs:

- Calibrated parameters η, μ, ε_V, p_T
- Error measure between projected values and data

Set 3.1 Summary Table

Question	Inputs (Baseline)	Inputs (Workbench)	Task	Output
Q1	RMarkdown file titled "Set 3", using section Q3.1.1 dataset_set3_1_1.csv	"Q3_1_1" workflow dataset_set3_1_1.csv	Calibrate the SEIRH model using the "new cases" and "new hospitalizations" data columns.	Calibrated parameter η • Mean value pre- and post-calibration • Prior and posterior distributions • Variance pre- and post-calibration • Some measure of goodness-of-fit
Q2	RMarkdown file titled "Set 3", using section Q3.1.2 dataset_set3_1_2.csv	"Q3_1_2" workflow dataset_set3_1_2.csv	Calibrate the SEIRHVD model using the "new cases", "new hospitalizations", and "new death" columns.	Calibrated parameters η, μ • Mean value pre- and post-calibration • Prior and posterior distributions • Variance pre- and post-calibration • Some measure of goodness-of-fit
Q3	RMarkdown file titled "Set 3", using section Q3.1.3 dataset_set3_1_3.csv	"Q3_1_3" model dataset_set3_1_3.csv	Calibrate the SEIRHVTD model using the "new cases", "new hospitalizations", "new treated", and "new death", and "new positive after vacc" columns.	Calibrated parameters η, μ, ε_V, p_T • Mean value pre- and post-calibration • Prior and posterior distributions • Variance pre- and post-calibration • Some measure of goodness-of-fit

Set 3.2: Forecasting

Create probabilistic or deterministic forecasts/simulations with a provided model.

Given a model and configuration to apply, simulate and create plots of all state variables for a given number of timesteps. Three models of increasing complexity will be tested.

The data sets in this set of questions contain data on Influenza A cases in one influenza season in the United States. Influenza season generally happens in the late fall and winter months in the US. The inputs will be a model, initial values for each state variable, and a configuration of parameters. The outputs expected will be plots of all state variables. You will be given a reference point for one of the state variables to check that your simulations are producing correct results.

General format of inputs, tasks, and outputs

Inputs	Tasks	Outputs
• Input model to simulate • Model configuration for forecast/simulation, including initial conditions, parameter values, and simulation parameters (e.g. dt, length of simulation, etc.). For deterministic forecasts/simulations, there will be no uncertainty in initial or parameter values. For probabilistic forecasts/simulation, one or more parameters will be defined as distributions.	• Configure model according to specification • Perform forecast/simulation	• Plots for all state variables, each in a separate plot, and indication that expected output is met (e.g. identification of peaks, etc.). For probabilistic forecasts/simulation, this includes posterior samples of all states.

(Q3.2.1) *Set 3.2, Question 1: Simple simulation and forecast*

Inputs:

- (Workbench) SEIRH model in workflow titled “Q3.2.1”
- (Baseline) SEIRH model in RMarkdown file titled “Set 3”, section Q3.2.1

The equations for the model are given by:

$$\frac{d}{dt}I(t) = -\delta I(t) - \eta I(t) + \sigma E(t)$$

$$\frac{d}{dt}R(t) = \delta I(t) + \kappa H(t)$$

$$\frac{d}{dt}H(t) = \eta I(t) - \kappa H(t)$$

$$\frac{d}{dt}E(t) = -\sigma E(t) + \frac{\lambda I(t) S(t)}{N}$$

$$\frac{d}{dt}S(t) = -\frac{\lambda I(t) S(t)}{N}$$

State variables and initial conditions:

- S = 50000; susceptible population
- E = 200; latent population
- I = 5; infectious population
- R = 0; recovered population
- H = 0; hospitalized population

Parameter values:

- $\lambda = 0.7$; force of infection
- $\sigma = 0.2$; rate of infection, 1/(days in latent state)
- $\delta = 0.263$; recovery rate, 1/(infectious period in days)
- $\eta = 0.02$; rate of hospitalization
- $\kappa = 0.15$; rate of recovery after hospitalization, 1/(average days in hospital)
- $N = S(0) + E(0) + I(0) + R(0) + H(0)$

Simulate the model for 100 days with dt = 1 day

Task:

- Configure and simulate (deterministically) the model using the initial values and parameters listed in the inputs.

Outputs:

- Plots for all state variables over the time range, each graphed in a separate plot. You should expect the peak of state variable I to occur at day 44.

(Q3.2.2) ***Set 3.2, Question 2:*** ***Medium Complexity Forecast***

Inputs:

- (Workbench) SEIRHVD model and workflow titled "Q3.2.2"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.2.2

The SEIRHVD model uses the following equations:

$$\frac{d}{dt}I(t) = -\delta I(t) - \eta I(t) + \sigma E(t)$$

$$\frac{d}{dt}H(t) = \eta I(t) - \kappa H(t) - \mu H(t)$$

$$\frac{d}{dt}R(t) = \delta I(t) + \kappa H(t)$$

$$\frac{d}{dt}D(t) = \mu H(t)$$

$$\frac{d}{dt}S(t) = -\frac{\lambda I(t) S(t)}{N}$$

$$\frac{d}{dt}E(t) = -\sigma E(t) + \frac{\lambda(1 - \epsilon_V) I(t) V(t)}{N} + \frac{\lambda I(t) S(t)}{N}$$

$$\frac{d}{dt}V(t) = -\frac{\lambda(1 - \epsilon_V) I(t) V(t)}{N}$$

State variables and initial conditions:

- S = 50000; susceptible population
- E = 200; latent population
- I = 20; infectious population
- R = 0; recovered population
- H = 0; hospitalized population
- V = 10000; vaccinated population
- D = 0; deceased population

Parameter values:

- λ = 0.65; force of infection
- σ = 0.15; rate of infection, 1/days in latent state
- δ = *unknown, range of 0.2 - 0.3*; recovery rate, 1/ infectious period in days
- η = 0.009; rate of hospitalization
- κ = 0.11; rate of recovery after hospitalization, 1/average days in hospitalization
- ε_V = 0.81; vaccine efficacy
- μ = 0.02; rate of death due to severe illness
- N = S(0) + E(0) + I(0) + R(0) + H(0) +V(0) + D(0)

Tasks:

- Configure the model using the initial values and parameters listed in the inputs.
- Simulate the data (probabilistically) for 110 days with dt = 1 day

Outputs:

- Plots for all state variables over the time range, each graphed in a separate plot. To convey the uncertainty in the parameter value(s), include sample forecasts of all states using sample parameter values. (If using a Bayesian workflow, include posterior samples; if not, include forecasts spanning the range of the uncertain parameter values.)
    - You should expect the peak of state variable E using the average value of δ to occur at day 57.

(Q3.2.3) *Set 3.2, Question 3:* *High Complexity Forecast*

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.2.3"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.2.3

The SEIRHVTD model uses the following set of equations:

$$\frac{d}{dt}S(t) = -\frac{\lambda I(t)S(t)}{N}$$

$$\frac{d}{dt}E(t) = -p_T\sigma E(t) - \sigma(1-p_T)E(t) + \frac{\lambda(1-\varepsilon_V)I(t)V(t)}{N} + \frac{\lambda I(t)S(t)}{N}$$

$$\frac{d}{dt}V(t) = -\frac{\lambda(1-\varepsilon_V)I(t)V(t)}{N}$$

$$\frac{d}{dt}I(t) = -\delta I(t) - \eta I(t) + \sigma(1-p_T)E(t)$$

$$\frac{d}{dt}T(t) = p_T\sigma E(t) - \delta T(t)$$

$$\frac{d}{dt}R(t) = \delta I(t) + \delta T(t) + \kappa H(t)$$

$$\frac{d}{dt}H(t) = \eta I(t) - \kappa H(t) - \mu H(t)$$

$$\frac{d}{dt}D(t) = \mu H(t)$$

Initial state variables:

- S = 18000; susceptible population
- E = 200; latent population
- I = 30; infectious population
- R = 0; recovered population
- H = 0; hospitalized population
- V = 10000; vaccinated population
- T = 0; treated population
- D = 0; deceased population

Initial values of parameters:

- λ = 0.85; force of infection
- σ = *unknown, range of 0.07 - 0.11*; rate of infection, 1/days in latent state
- δ = 0.25; recovery rate, 1/ infectious period in days
- η = *unknown, range of 0 – 0.02*; rate of hospitalization
- κ = 0.1; rate of recovery after hospitalization, 1/average days in hospitalization
- ε_V = 0.65; vaccine efficacy
- μ = *unknown, range of 0 - 0.02*; death rate due to severe illness
- p_T = 0.3; proportion of latent population treated with antivirals
- N = S(0) + E(0) + I(0) + R(0) + H(0) +V(0) +T(0) + D(0)

Task:

- Configure the model using the initial values and parameters listed in the inputs.
- Simulate the data (probabilistically) for 200 days with dt = 1 day

Outputs:

- Plots for all state variables over the time range, each graphed in a separate plot. To convey the uncertainty in the parameter value(s), include sample forecasts of all states using sample parameter values. (If using a Bayesian workflow, include posterior samples; if not, include forecasts spanning the range of the uncertain parameter values.)
  - You should expect the peak of state variable I using the average values of the unknown parameters, to occur around day 87

Set 3.2 *Summary Table*

Question	Inputs (Baseline)	Inputs (Workbench)	Task	Output
Q1	RMarkdown file titled "Set 3", using section Q3.2.1 Initial values and parameters given in question text	"Q3.2.1" workflow Initial values and parameters given in question text	Configure the model as instructed and simulate the data for 100 timesteps.	Separate forecast plots for all state variables graphed over the time range. You should expect the peak of state variable I to occur at timestep 44.
Q2	RMarkdown file titled "Set 3", using section Q3.2.2 Initial values and parameters given in question text	"Q3.2.2" workflow Initial values and parameters given in question text	Configure the model as instructed and simulate the data for 110 timesteps.	Separate forecast plots for all state variables graphed over the time range, conveying the uncertainty introduced by unknown parameter values. You should expect the peak of state variable E to occur around timestep 57, using the average value of δ.
Q3	RMarkdown file titled "Set 3", using section Q3.2.3 Initial values and parameters given in question text	"Q3.2.3" workflow Initial values and parameters given in question text	Configure the model as instructed and simulate the data for 200 timesteps.	Separate forecast plots for all state variables graphed over the time range, conveying the uncertainty introduced by unknown parameter values. You should expect the peak of state variable I (using the average values of the unknown parameters) to occur around day 87.

Set 3.3: Interventions

Implement interventions in a provided model, simulate, and assess impact.

Given a fully configured model, add increasingly complex sets of interventions, simulate, and then create requested plots describing intervention impacts.

The following three questions explore how interventions can alter a disease trajectory. You will compare a disease forecast before and after an intervention is implemented. The inputs will be a model with configured initial values and parameters, and a given intervention value and day of intervention. The outputs expected will be plots of a few state variables and the difference between output values of the forecast before and after intervention. The SEIRHVTD model used in these questions has the same equations as in Q3.2.3 for your reference.

General format of inputs, tasks, and outputs

Inputs	Tasks	Outputs
• Model with pre-intervention configuration applied (including initial conditions, parameter values, and simulation parameters) • One or more interventions to implement, specified as changing the value of one or more parameters, and creating and applying a new model configuration. Interventions may also be specified as changing the values of a dataset attached to the model configuration process (e.g. contact matrix).	• Simulate model with pre-intervention configuration to get baseline results with respect to the outputs of interest • Create new model configuration (intervention configuration) with the described intervention(s) • Simulate model with intervention applied • Calculate impact of intervention with respect to the outputs of interest	• Simulation output plots for outputs of interest, before and after intervention was implemented • Impact of intervention with respect to outputs of interest, in a metric or quantitative output (e.g. approximate difference in number of hospitalizations between baseline and after intervention policy applied).

(Q3.3.1) *Set 3.3, Question 1: Simple Intervention*

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.3.1"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.3.1

Tasks:

- Simulate the baseline model with no intervention applied.
- Create an intervention policy to change the proportion of individuals p_T treated with antivirals to 0.4 starting at day 10.
- Simulate the intervention for 100 days.

Outputs:

- Separate forecast plots for state variables I, D, T, pre- and post-intervention.
- Approximate difference in cumulative number of deaths at day 100 between baseline and after intervention policy applied.

(Q3.3.2) ***Set 3.3, Question 2*: Medium Complexity Intervention**

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.3.2"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.3.2

Tasks:

- Simulate the baseline model with no intervention applied.
- Suppose a town wants to hold a large vaccination clinic and public health officials have enough people signed up to decrease the total remaining susceptible population to 17890 (which would increase the vaccinated population to 15890). They are planning this clinic for day 5 – you may assume the vaccine is immediately effective, so the intervention policy can start on day 5. Create an intervention policy that reflects this and also includes an intervention that changes the proportion of individuals treated with antivirals to p_T=0.4 if the number of infected individuals I reaches 500.
- Simulate the intervention policy for 100 days.

Outputs:

- Separate forecast plots for state variables I, D, T, pre- and post-intervention.
- Approximate difference in number of deaths at day 100 between baseline and after intervention policy applied.

(Q3.3.3) ***Set 3.3, Question 3*: High Complexity Intervention**

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.3.3"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.3.3

Tasks:

- Simulate the baseline model with no intervention applied.
- The local public health office wants you to simulate what would happen if they were to vaccinate with a more effective vaccine throughout the entirety of the timespan. Suppose a more effective vaccine with efficacy $\varepsilon_V = 0.9$ is created and could be introduced on Day 1. Create an intervention policy that reflects this. Also include an intervention that changes the proportion of individuals treated with antivirals to $p_T = 0.9$ if the number of infected individuals I reaches 1000.

- Finally, you should implement a quarantine policy that reduces λ to $\lambda = 0.6$ if the number of hospitalized individuals reaches 500.
- Simulate the model with these intervention policies for 100 days.

Outputs:
- Separate forecast plots for state variables I, D, T, pre- and post-intervention.
- Approximate difference in number of deaths at day 100 between baseline and after intervention policy applied.

Set 3.3 *Summary Table*

Question	Inputs (Baseline)	Inputs (Workbench)	Task	Output
Q1	RMarkdown file titled "Set 3", using section Q3.3.1	"Q3.3.1" workflow	Intervene to change the proportion of individuals treated with antivirals to 0.4 starting at day 10. Simulate for 100 days.	• Separate forecast plots for state variables I, D, T, pre- and post-intervention • Change in end state of the variables I, D, T • Approximate difference in number of deaths at day 100 between baseline and after intervention policy applied
Q2	RMarkdown file titled "Set 3", using section Q3.3.2	"Q3.3.2" workflow	Intervene to decrease susceptible population to 20650 people (which would increase the vaccinated population to 25035) beginning at day 5, and to change the proportion of individuals treated with antivirals to p_T=0.4 if the number of infected individuals reaches 500. Simulate for 100 days.	• Separate forecast plots for state variables I, D, T, pre- and post-intervention • Change in end state of the variables I, D, T • Approximate difference in number of deaths at day 100 between baseline and after intervention policy applied.
Q3	RMarkdown file titled "Set 3", using section Q3.3.3	"Q3.3.3" workflow	Intervene to change vaccine efficacy to ε_V =0.9 beginning at day 1. Also include a quarantine intervention that sets the parameter $\lambda = 0.01$ if the number of hospitalized individuals reaches 500. Simulate for 100 days.	• Separate forecast plots for state variables I, D, T, pre- and post-intervention • Change in end state of the variables I, D, T • Approximate difference in number of deaths at day 100 between baseline and after intervention policy applied.

Set 3.4: Optimizing Interventions

Find optimal interventions based on provided goals.

Given a configured model and a goal, optimize interventions and create requested plots to prove that the optimized intervention satisfies the goal. Optimization can be value-based (intervention parameter values), time-based (find the earliest/latest time for an intervention, or minimize the total amount of time an intervention is in effect), and can involve a single or multiple objectives.

The following three questions explore how interventions can alter a disease trajectory. You will optimize the intervention policy to achieve the given goal. The inputs will be a model with configured initial values and parameters, an initial intervention value and day of intervention, and a goal for the intervention outcome. The outputs expected will be plots of a few state variables and an indicated value from the optimal intervention policy. The SEIRHVTD model used in these questions has the same equations as in set 3.2.3 for your reference.

General format of inputs, tasks, and outputs

Inputs	**Tasks**	**Outputs**
• Model with initial unoptimized configuration applied (initial conditions, parameter values, and simulation parameters) • Objective function with one or more objectives in terms of variables or observables to optimize. E.g. find minimal mask efficacy required to achieve outcome X (defined as one or more constraints); find the latest time a social distancing policy can be implemented and still achieve outcome Y; minimize cumulative infections AND minimize cost according to some cost function. • Constraints on the optimization problem, in terms of parameters, state variables, or observable functions. Constraints can apply for the entire simulation, or just for certain time periods.	• Simulate model with unoptimized configuration to get baseline results with respect to the objective • Perform value- or time-based optimization (or some combination of the two), with single or multiple objectives • Simulate model with optimal intervention	• A solution to the objective function (e.g., to achieve X, the minimum required mask efficacy is ρ, t_{int} is the latest time that intervention policy Z can be implemented, etc.) • A verification that the provided solution solves the optimization question (e.g. plot showing results with respect to objective and constraints, pre- and post-optimization, etc.)

(Q3.4.1) *Set 3.4, Question 1: Simple Optimization of Intervention*

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.4.1"
- (Baseline) RMarkdown file titled "Set 3", section Q3.4.1

- Constraint: The goal of this intervention is to keep the current number of hospitalizations, H, below 80 at each day of the simulation, in at least 95% of all simulations.
- Objective: Find minimal value of p_T (proportion of individuals who are treated with antivirals), that will meet the constraint with 95% confidence

Tasks:

- Simulate the unoptimized, configured model for 90 days to obtain a baseline plot of H before intervention optimization.
- Set an initial intervention policy to start at day 10 by setting the proportion of individuals who are treated with antivirals $p_T = 0.35$. The intervention end time is at day 90.
- Optimize the intervention to find the optimal parameter value of p_T to keep the number of hospitalized individuals under 80 at all times (since the local hospital system cannot support more than that), with 95% confidence, while keeping p_T as low as possible. Note that health officials don't think it is reasonable to assume they will have treatment available for more than 50% of the eligible population at any given time in the simulation.

Outputs:

- Minimum value for p_T if the intervention starts on day 10
- Plot of H before and after optimal intervention policy applied, demonstrating the solution to the optimization problem meets constraint

(Q3.4.2) *Set 3.4, Question 2: Medium Complexity Optimization of Intervention*

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.4.2"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.4.2
- Constraint: The goal of this intervention is to keep the number of infections I below 1100 for each day and the number of hospitalizations H under 75 for each day in the simulation, with 95% confidence.
- Objective: Find latest start time for intervention on p_T (proportion of individuals who are treated with antivirals), that will meet the constraint with 95% confidence

Tasks:

- Simulate the unoptimized, configured model for 100 days to obtain a baseline plot of I before intervention and optimization.
- Set an initial intervention policy to start at day 8 by setting the proportion of individuals who are treated with antivirals p_T to 0.4.
- Optimize the intervention policy to find the optimal day to start the intervention of $p_T = 0.4$, with end time of day 100. The goals is to keep the number of infections to under 1100 and the number of hospitalizations H under 75 for each day, with 95% confidence. Note that officials want to start the treatment as late as possible while still meeting the goal.

Outputs:

- Latest possible day to start the intervention if $p_T = 0.4$
- Plot of I and H before and after optimal intervention policy applied (four plots total), demonstrating the solution to the optimization problem meets constraint

(Q3.4.3) *Set 3.4, Question 3: Challenging Optimization*

Inputs:

- (Workbench) SEIRHVTD model and workflow titled "Q3.4.3"
- (Baseline) RMarkdown file titled "Set 3", using section Q3.4.3
- Constraint: The goal of this intervention is to keep the number of infections I below 125 people for each day and the number of hospitalizations H under 40 for each day in the simulation, with 95% confidence.
- Objective: Find the latest possible day to start intervention on p_T, the lowest possible value of p_T, and latest possible day to introduce new quarantine policy, in order to meet constraint with 95% confidence

Record how long it takes you to complete the following task.

Tasks:

- Simulate the unoptimized, configured model for 90 days to obtain a baseline plots of H and I before intervention and optimization.
- Set an initial intervention policy to start at day 10 by setting the proportion of individuals who are treated with antivirals p_T to 0.3. Also, suppose starting at day 13, public health officials introduce a new quarantine policy that reduces the effective λ to 0.5; incorporate this intervention assumption.
- Optimize the intervention policy with end time of day 90, to find the latest possible day to start the intervention of p_T (while still meeting the intervention goal), the lowest possible value of p_T to meet the goal (while still meeting the intervention goal), and the latest possible day to introduce the new quarantine policy (while still meeting the intervention goal). These interventions should be optimized with the goal to keep the number of infections I under 125 and the number of hospitalizations under 40 for each day in the simulation, with 95% confidence.

Outputs:

- Latest day to start the intervention of p_T and the accompanying lowest possible value of p_T
- Latest possible time to start the quarantine intervention of reducing λ to 0.5
- If there is more than one optimal solution set, please make a set of plots for each optimal solution. If there are more than a few optimal solution sets, please document the ranges of values in each of the optimal solution set and make a set of plots corresponding to the lowest possible value of p_T to meet the goal, and choose values for the start day for p_T and the start day for the quarantine policy, to plot with the lowest possible value of p_T.

- Plots of I before after optimal intervention policy (two separate plots), demonstrating the solution to the optimization problem meets constraint
- Plots of H before and after optimal intervention policy (two separate plots), demonstrating the solution to the optimization problem meets constraint

Set 3.4 *Summary Table*

Question	Inputs (Baseline)	Inputs (Workbench)	Task	Output
Q1	• RMarkdown file titled "Set 3", using section Q3.4.1 • Constraint: keep H < 80 • Objective: Find minimal value of p_T that will meet constraint with 95% confidence	• "Q3.4.1" workflow • Constraint: keep H < 80 • Objective: Find minimal value of p_T that will meet constraint with 95% confidence	Set up the initial intervention policy as described. Optimize the intervention policy with end time of 90 days to find the optimal parameter value of p_T to keep the number of hospitalizations H under 80 each day.	• Minimum value for p_T • Plot of H before and after optimal intervention policy
Q2	• RMarkdown file titled "Set 3", using section Q3.4.2 • Constraint: keep I < 1100 each day, H <75 each day • Objective: Find latest start time for intervention on p_T that will meet constraint with 95% confidence	• "Q3.4.2" workflow • Constraint: keep I < 1100 each day, H <75 each day • Objective: Find latest start time for intervention on p_T that will meet constraint with 95% confidence	Set up the initial intervention policy as described. Optimize the intervention policy with end time of 100 days to find the optimal day to start the intervention of p_T at 0.4 to keep the number of infections I under 1100 each day and the number of hospitalizations H under 75 each day	• Latest start time to start the antiviral treatment intervention of $p_T = 0.4$ • Plot of I and H before and after optimal intervention policy
Q3	• RMarkdown file titled "Set 3", using section Q3.4.3 • Constraint: keep I < 125 and H <40 each day • Objective: Find *the latest possible day* to start intervention on p_T, *the lowest possible value* of p_T, and *latest possible day to introduce new quarantine policy,* in order to meet constraint with 95% confidence	• "Q3.4.3" workflow • Constraint: keep I < 125 and H <40 each day • Objective: Find *the latest possible day* to start intervention on p_T, *the lowest possible value* of p_T, and *latest possible day to introduce new quarantine policy,* in order to meet constraint with 95% confidence	Set up the initial intervention policy as described. Optimize the intervention policy with end time of day 90 to find the optimal day to start the intervention of p_T, the accompanying optimal value of p_T, and the optimal day to introduce the quarantine policy, with the treatment intervention policy and quarantine policy both starting as late as possible, and the treatment intervention value set as low as possible.	• Latest possible time to start intervention of p_T • Accompanying minimum value of p_T • Latest possible time to start quarantine policy, seen through a change in λ • Plots of I and H before and after optimal intervention policy

Set 3.5: Ensembles

Create an ensemble of provided models.

Given models and calibration data, create an ensemble of the models that best fits the calibration data and then perform a forecast.

General format of inputs, tasks, and outputs

Inputs	Tasks	Outputs
• Three or more models that we would like to create ensemble with • Dataset and indication of which features to use for calibration • Forecast configuration (e.g. initial conditions, number of timesteps)	• Create an ensemble, which can be weighted or non-weighted • Perform forecast with ensemble according to configuration	• Calibration results (weights of component models and measures of error / calibration fit, calibrated parameters for component models) • Forecast output for all state variables, including uncertainty intervals • Mean absolute error (MAE) between projected value for one or more state variables from the ensemble model, and observational data from Dataset.csv

(Q3.5.1) *Set 3.5, Question 1: Simpler Ensemble*

Inputs:

- (Workbench) Collection of models in workflow "Q3.5.1"
- (Baseline) RMarkdown file titled "Set 3", section "Q3.5.1"
- Dataset: "dataset_Q3.5.1_ensemble" (Workbench) and "dataset_Q3.5.1_ensemble.csv" (Baseline)

Tasks:

- The goal of this task is to combine three models into an ensemble of the models. The models are provided for you.
- Calibrate each of the models individually using the dataset provided. For parameters with unknown values, use the data to calibrate the model's parameters to be within the ranges provided. Calibrate the individual models to minimize the difference in cases, hospitalizations, or both relative to the data provided. You may use any error criteria you like. Simulate each of the models individually.
- Once all three models are calibrated (and have scalar values for all parameters), create an ensemble model. To do this, find weights which sum to 1 to minimize the same error that was minimized in the step above. (In other words, if you calibrated to minimize hospitalizations above, do that here; if you calibrated to minimize cases above, do that here). Simulate the ensemble model.
- Compare the error that the ensemble model computes to the error generated by the individual models.

Outputs:
- Calibrated parameters for each of the three individual models that were calibrated
- Calibrated weights for the ensemble of the three calibrated models
- Errors (as defined as being relative to cases, hospitalizations, or both) of the three calibrated models and the ensemble model
- Ensemble simulation outputs with uncertainty

(Q3.5.2) *Set 3.5, Question 2: Challenging Ensemble*

Inputs:
- (Workbench) Collection of models titled “3.5.2 ensemble”
- (Baseline) RMarkdown file titled “Set 3”, using section Q3.5.2
- (Both) d3_ensemble (in Terarium) and dataset_ensemble.csv (Baseline)

Tasks:
- The tasks for 3.5.2 are the same as for 3.5.1, but you are now using five models in the ensemble instead of three. In this question, all parameters for the individual models are calibrated, so you do not need to calibrate those but do need to calibrate the ensemble as a whole.

Outputs:
- The output instructions for 3.5.2 are the same as for 3.5.1, but you are now using five models in the ensemble instead of three.

Set 3.5 *Summary Table*

Question	Inputs (Baseline)	Inputs (Workbench)	Task	Output
Q1	RMarkdown file titled “Set 3”, using section Q3.5.1 Data: dataset_Q3.5.1_ensemble.csv	(Workbench) Collection of models titled “3.5.1 ensemble” Data: d3_ensemble	Combine (3 or 5) models into an ensemble. If any parameters are not set as constants, first calibrate those parameters using the data provided. Then, build an ensemble by finding the model weights which minimize the error in the ensemble. You may use any error criteria you like. Simulate ensemble.	• Calibrated parameters for each of the (3 or 5) individual models that were calibrated • Calibrated weights for the ensemble of the (3 or 5) calibrated models • Errors (relative to hospitalizations, cases, or both) of each calibrated model and the ensemble model • Ensemble simulation outputs with uncertainty
Q2	RMarkdown file titled “Set 3”, using section Q3.5.2 Data: dataset_ensemble.csv	(Workbench) Collection of models titled “3.5.2 ensemble” Data: d3_ensemble		

Section 4: Technical Area 4 Functionality

Set 4.3: Template Scenarios

For Set 4.3, there are no prepopulated workflows for the workbench teams to use. Instead, for each of the tasks below, you will need to make a new workflow using Terarium. Each workflow should use the appropriate template for that question – for example, Q4.3.1 should use the "Situational Awareness" template. Please label the new workflows that you make using the format of the other workflows (e.g., "Q4.3.1").

General format of inputs, tasks, and outputs

Inputs	**Tasks**	**Outputs**
Model to construct scenario around	Execute scenario as described in each question. Different questions may describe different types of scenarios (e.g. sensitivity analysis, horizon scan, etc.)	Provide scenario outputs as a collection of plots

(Q4.3.1) Set 4.3, Question 1: Situational Awareness and Intervention

Inputs:

- (Workbench) Configured SEIRHVTD model in a new workflow that you make
- (Baseline) RMarkdown file titled "Set 4", using section Q4.3.1
- Dataset
    - Baseline: dataset_4.3.1_sit_awareness_trunc.csv
    - Workbench: d3_1_3_trunc50
- Configuration to apply to the model:
    - Initial conditions:
        - S = 40000; susceptible population
        - E = 150; latent population
        - V = 12000; vaccinated population
        - I = 5; infectious population
        - Tr = 0; treated population
        - H = 0; hospitalized population
        - R = 0; recovered population
        - D = 0; deceased population
    - Parameters:
        - λ= 0.19; force of infection
        - ε_V = 0.85; vaccine efficacy
        - σ= 0.19; rate of infection, 1/days in latent state
        - δ= 0.25; recovery rate, 1/ infectious period in days
        - η= range (.005, .03); rate of hospitalization
        - p_T= 0.1; proportion of latent individuals treated with antivirals
        - κ = 0.13; rate of recovery after hospitalization, 1/average days in hospitalization
        - μ = 0.01
        - N = 52,155 people

Tasks:

- Apply the given configuration to the model
- Calibrate the model to the first 50 days of data, create a simulation for the next 50 days of data, and assess the impact of two potential interventions on hospitalizations and cases.
    - Intervention 1: reduce the transmission parameter λ by half starting on day 60
    - Intervention 2: reduce the parameter sigma by half starting on day 60
- In the workbench, this should be done with a situational awareness template.

Outputs:

- Graphs of hospitalizations and cases for the next 50 days of data 1) assuming no intervention, 2) assuming the lambda intervention was applied, and 3) assuming the sigma parameter was cut in half starting at day 60.

(Q4.3.2) *Set 4.3, Question 2: Sensitivity Analysis*

Inputs:

- (Workbench) Q4.3.2 SEIR-with-N model in a new workflow that you make
- (Baseline) RMarkdown file titled "Set 4", using section Q4.3.2

Tasks: In this scenario, there is a hypothetical disease spreading among a population of N=10,000 people. The goal of this task is to determine which parameter (β, σ, or γ) has the greatest influence on the peak infected population given a specific range of uncertainty.

The initial conditions are as follows:

- S = 9990 people
- E = 10 people
- I = 0 people
- R = 0 people

The parameter values are as follows. These are all uniform distributions.

Table 1 Sensitivity Analysis Parameter Values

Parameter	Initial value	Lower end of range	Upper end of range
β	0.2/day	0.1/day	0.5/day
σ	0.25/day	0.167/day	0.5/day
γ	0.14/day	0.1/day	0.2/day

Perform a sensitivity analysis on each parameter across the given range to determine the impact on the peak number of infected (I_{peak})

- Plot I_{peak} against a range of values for each parameter to visualize the relationship
- Identify the parameter that most strongly influences I_{peak}
- Provide a brief analysis (a sentence or two) of your findings

Outputs:

- Three plots, showing sensitivity of I_{peak} to each parameter
- A brief analysis (a sentence or two) showing which parameter most strongly influences I_{peak}

(Q4.3.3) Set 4.3, Question 3: Horizon Scanning

Inputs:

- (Workbench) SEIRHVTD model in a new workflow that you make
- (Baseline) RMarkdown file titled "Set 4", using section Q4.3.1
- Configuration to apply:
  - Initial Conditions
    - S = 40000; susceptible population
    - E = 150; latent population
    - I = 5; infectious population
    - R = 0; recovered population
    - H = 0; hospitalized population
    - V = 12000; vaccinated population
    - Tr = 0; treated population
    - D = 0; deceased population
  - Parameters
    - λ = 0.19; force of infection
    - ε_V= 0.81; vaccine efficacy
    - σ = 0.19; rate of infection, 1/days in latent state
    - δ = 0.25; recovery rate, 1/ infectious period in days
    - η = 0.02; rate of hospitalization
    - p_T= 0.1; proportion of latent individuals treated with antivirals
    - κ = 0.14; rate of recovery from hospitalization, 1/average days in hospitalization
    - μ = 0.02; death rate
    - N = 52,155 total people

Tasks:

- Apply the provided configuration to the model
- Do a horizon scan and create four versions of this model by varying the λ and ε_V parameters to be the following combinations:

Table 2. Horizon Scanning with λ and ε_V

	Low λ	**High ε_V**
Low λ	(.05, .50)	(.25, .50)
High ε_V	(.05, .95)	(.25, .95)

- Simulate the model using each of the above configurations for 100 days. All other parameters should be the same values as found in Q3.3.1.

Outputs:

- Forecast plots for H, I, V, and D for each of the four configurations above.

Set 4.*3 Summary Table*

Question	Input Baseline	Input Workbench	Task	Output
Q1	RMarkdown file titled "Set 4", using section Q4.3.1 Dataset: dataset_4.3.1_sit_awareness_trunc.csv	Model: SEIRHVTD model in a new workflow that you make Dataset: d3_1_3_trunc50	Calibrate the model to the first 50 days of data, create a simulation for the next 50 days of data, and assess the impact of two potential interventions on hospitalizations and cases.	Graphs of hospitalizations and cases for the next 50 days of data 1) assuming no intervention, 2) assuming the lambda intervention was applied, and 3) assuming the sigma parameter was cut in half starting at day 60.
Q2	RMarkdown file titled "Set 4", using section Q4.3.2	Q4.3.2 SEIR-with-N model in a new workflow that you make	Perform a sensitivity analysis on each parameter across the given range to determine the impact on the peak number of infected (I_{peak}) • Plot I_{peak} against each parameter to visualize the relationship. • Identify the parameter that most strongly influences I_{peak} • Provide a brief analysis (a sentence or two) of your findings.	Three plots, showing sensitivity of I_{peak} to each parameter A brief analysis (a sentence or two) showing which parameter most strongly influences I_{peak}
Q3	RMarkdown file titled "Set 4", using section Q4.3.3 Model configuration	SEIRHVTD model in a new workflow that you make Model configuration	Do a horizon scan and create four versions of the model by setting the λ and ε_V to be combos of $\lambda = \{.05, .25\}$ and ε_V to be $\{.50, .95\}$ Simulate the model using each of the above configurations for 100 days. All other parameters should be the same values as found in Q3.3.1.	Forecast plots for H, I, V, and D for each of the four configurations.