Optimization of Stochastic Epidemiological Models for Disease Control and Prediction

Dr. Khalil Al Handawi

Joint work with: Prof. Michael Kokkolaras

Journées de l'optimisation 2022

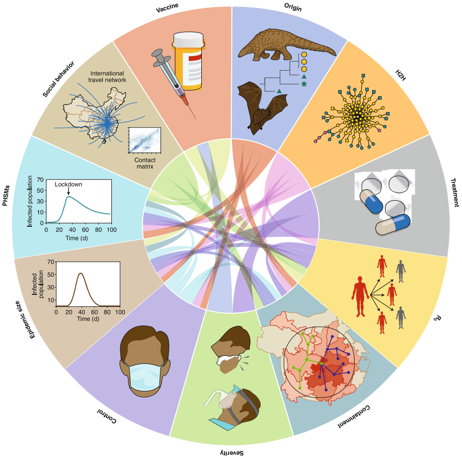
July 24, 2023

Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

In the early stages of epidemics, we have access to:

Simulation models can be developed based on this information and used in public health and social measures (PHSMs) policy-making.



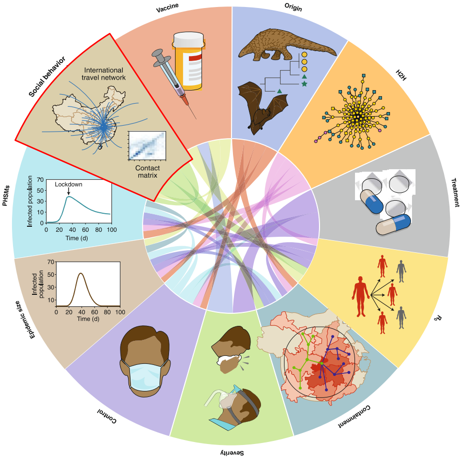
Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

In the early stages of epidemics, we have access to:

- **Social behaviour and demographics**

Simulation models can be developed based on this information and used in public health and social measures (PHSMs) policy-making.



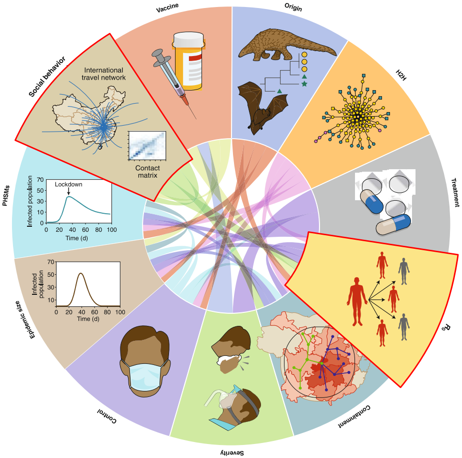
Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

In the early stages of epidemics, we have access to:

- Social behaviour and demographics
- **Incidence rate data**

Simulation models can be developed based on this information and used in public health and social measures (PHSMs) policy-making.



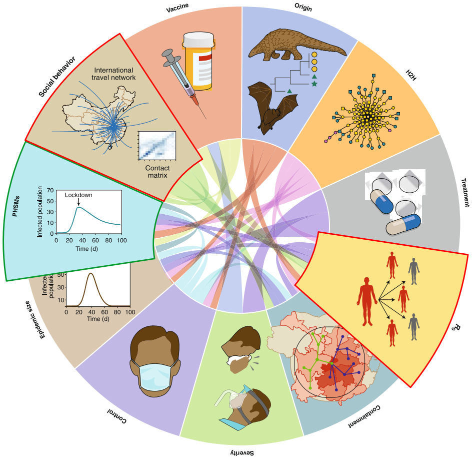
Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

In the early stages of epidemics, we have access to:

- Social behaviour and demographics
- Incidence rate data

Simulation models can be developed based on this information and used in **public health and social measures (PHSMs)** policy-making.

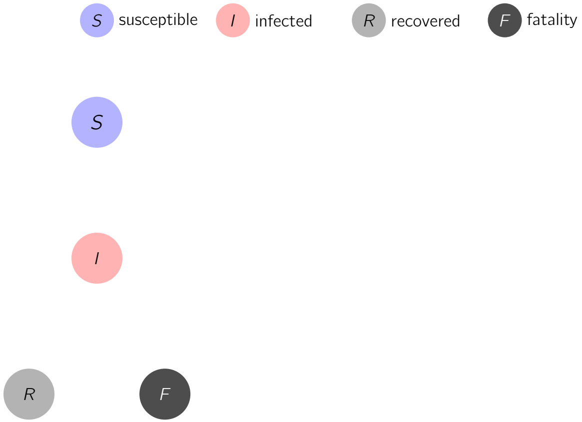


Simulation and modeling of pandemics in early stages

Joint work with: Prof. Michael Kokkolaras

Background: epidemiological models

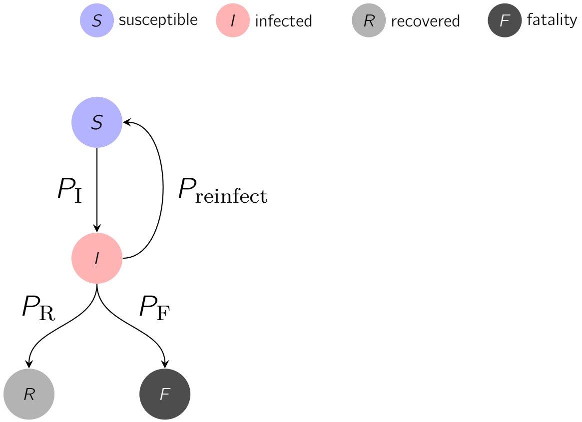
What are compartmental epidemiological models?



Background: epidemiological models

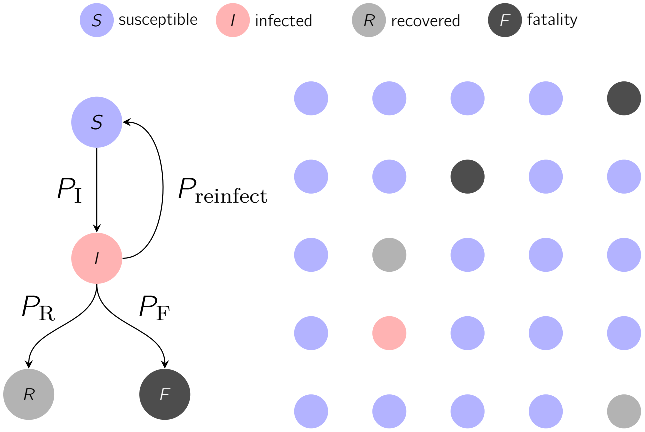
What are compartmental epidemiological models?

- Described by a *stochastic* process



What are compartmental epidemiological models?

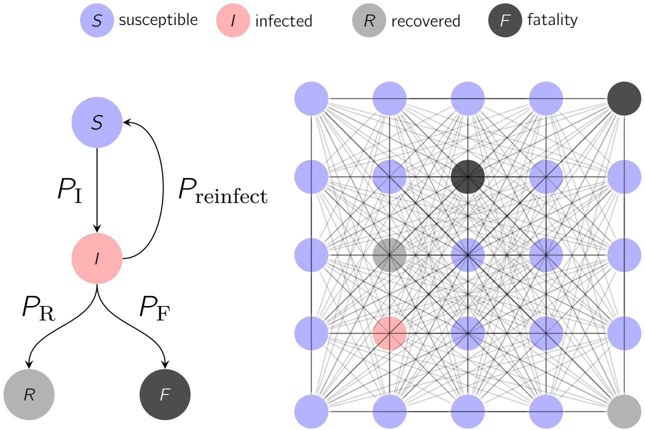
- Described by a *stochastic* process



Background: epidemiological models

What are compartmental epidemiological models?

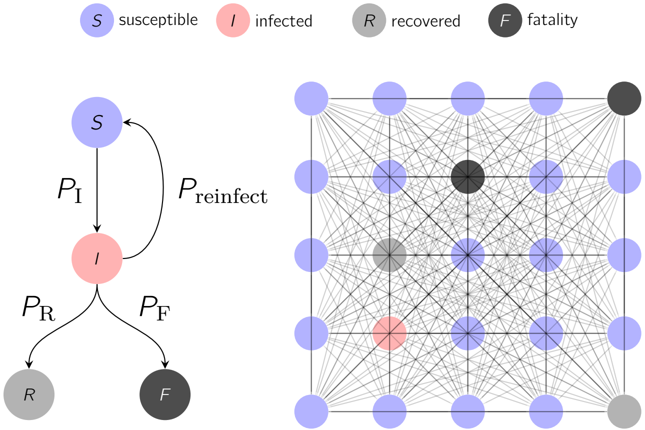
- Described by a *stochastic* process
- Assumes *homogenous* interaction



Background: epidemiological models

What are compartmental epidemiological models?

- Described by a *stochastic* process
- Assumes *homogenous* interaction
- Deterministic response for large N



Background: epidemiological models

What are compartmental epidemiological models?

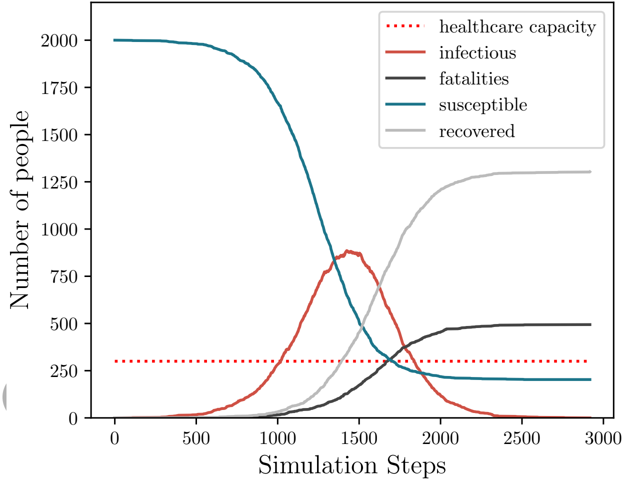
- Described by a *stochastic* process
- Assumes *homogenous* interaction
- Deterministic response for large N

$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$

$$\frac{dR}{dt} = \gamma I,$$

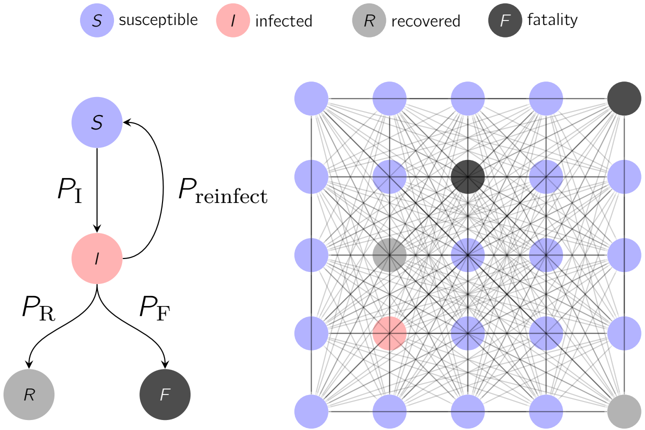
where $N = S + I + R$, β controls infection spread, and γ controls recovery rate



Background: epidemiological models

What are compartmental epidemiological models?

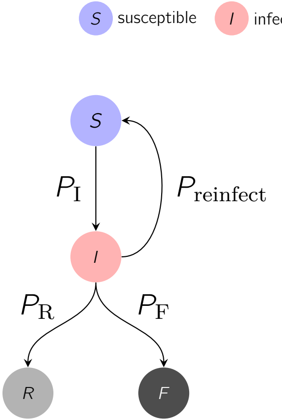
✓ Analytical solutions are available

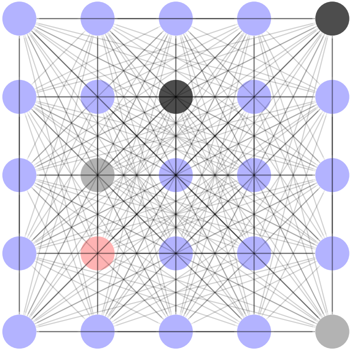


Background: epidemiological models

What are compartmental epidemiological models?

- ✓ Analytical solutions are available
- ✓ Captures large-scale population dynamics



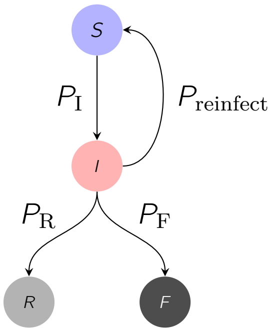


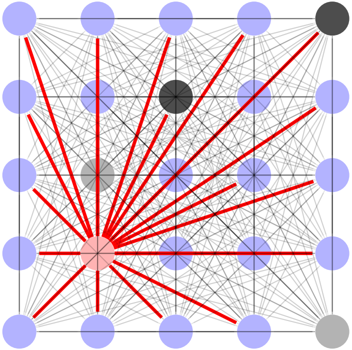
Background: epidemiological models

What are compartmental epidemiological models?

- ✓ Analytical solutions are available
- ✓ Captures large-scale population dynamics
- ✘ Does not account for *geography*

S susceptible I infected R recovered F fatality

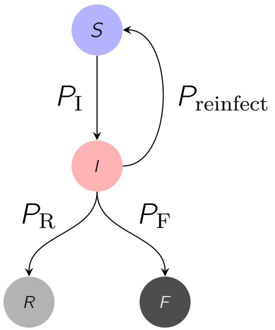


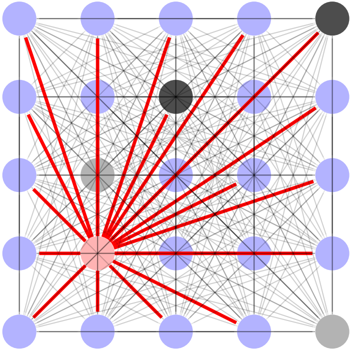


Background: epidemiological models

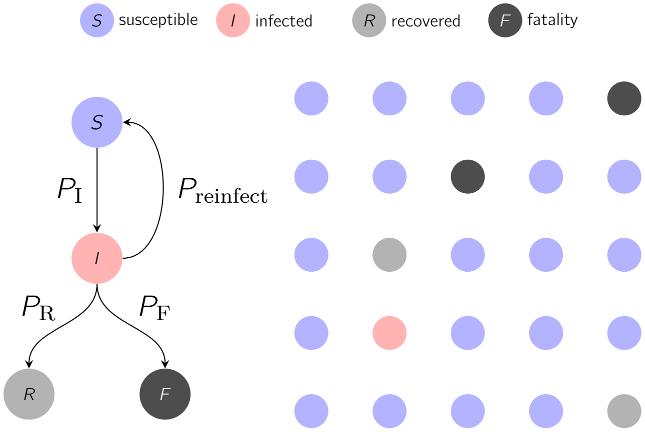
What are compartmental epidemiological models?

- ✓ Analytical solutions are available
- ✓ Captures large-scale population dynamics
- ✘ Does not account for *geography*
- ✘ Cannot model effect of intervention policies





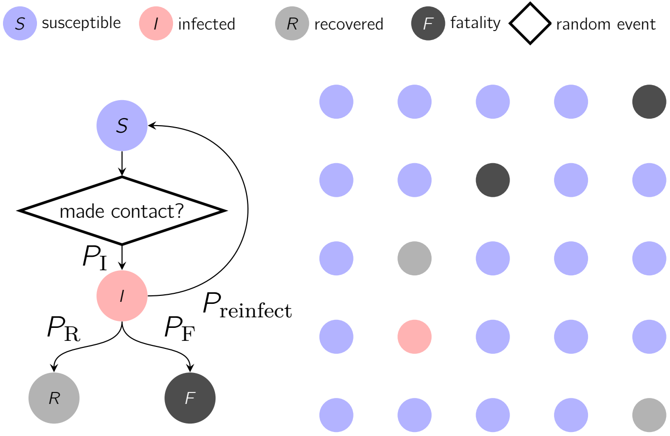
What are agent-based epidemiological models?



Background: epidemiological models

What are agent-based epidemiological models?

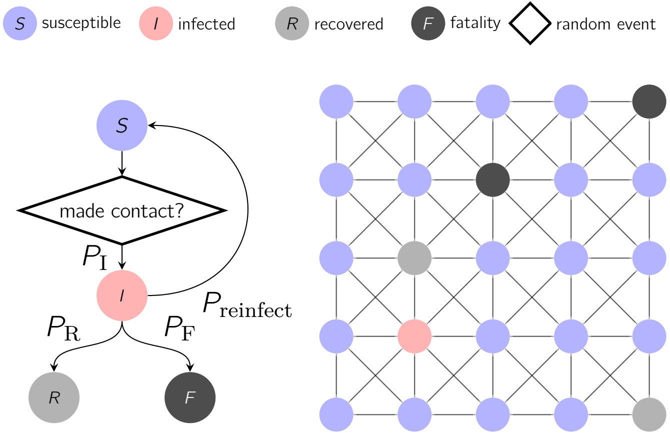
- *Stochastic* process



Background: epidemiological models

What are agent-based epidemiological models?

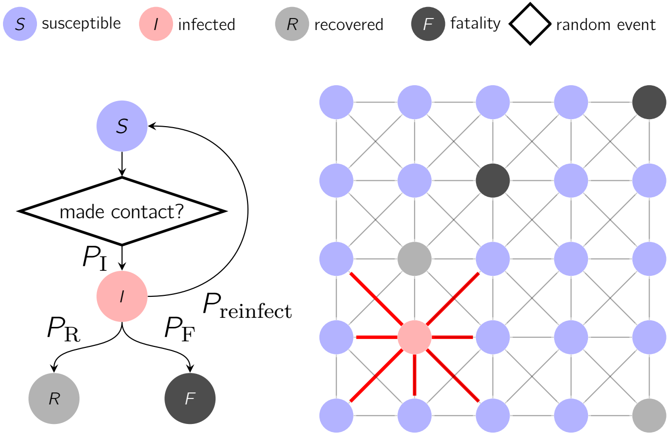
- *Stochastic* process
- Assume *heterogenous* interaction



Background: epidemiological models

What are agent-based epidemiological models?

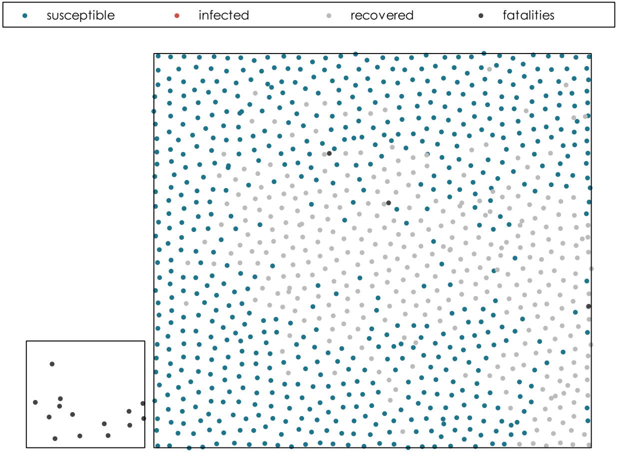
- *Stochastic* process
- Assume *heterogenous* interaction



What are agent-based epidemiological models?

- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response

Realization 1

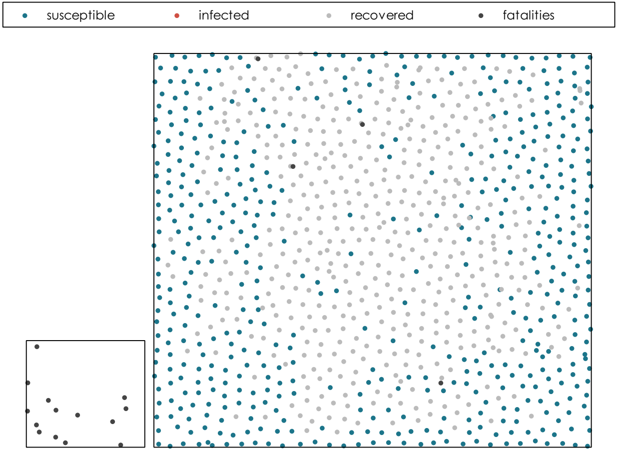


Background: epidemiological models

What are agent-based epidemiological models?

- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response

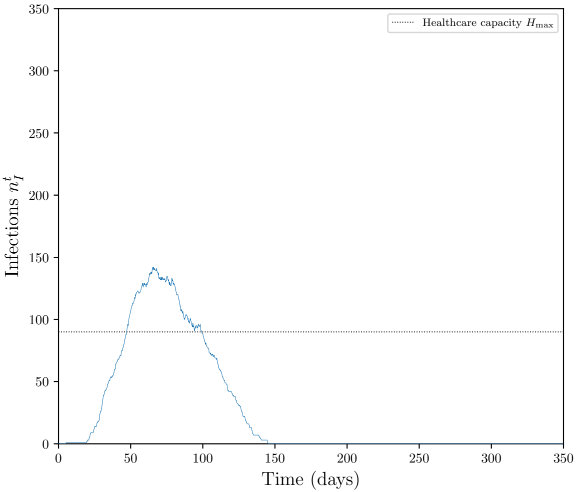
Realization 2



What are agent-based epidemiological models?

- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response

Realization 1

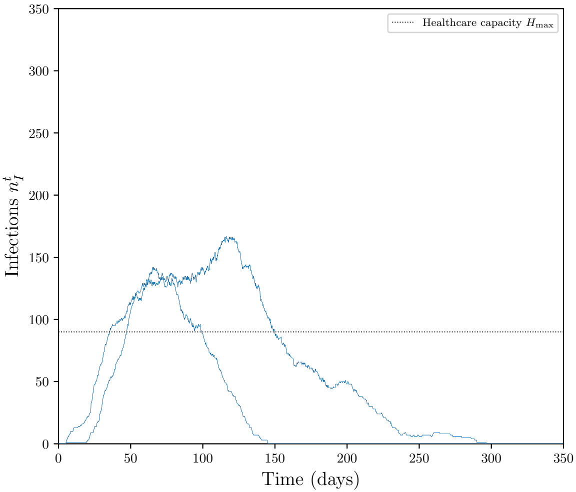


Background: epidemiological models

What are agent-based epidemiological models?

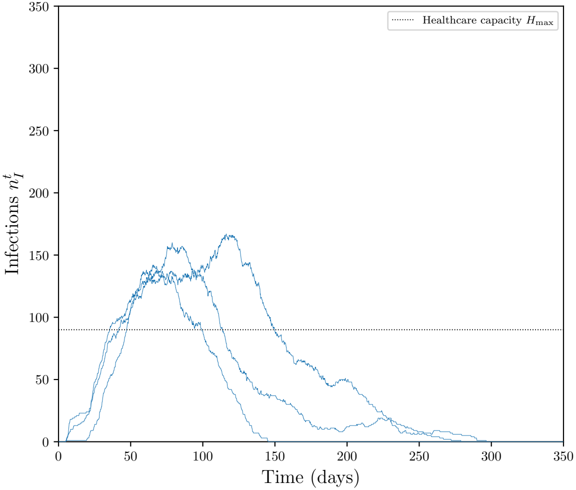
- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response

Realization 2



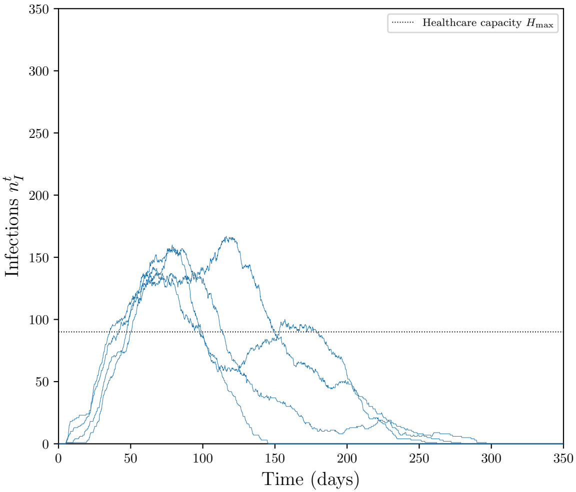
What are agent-based epidemiological models?

- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response



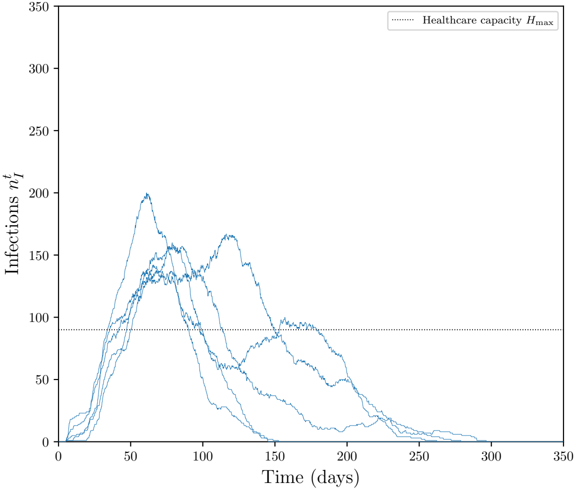
What are agent-based epidemiological models?

- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response



What are agent-based epidemiological models?

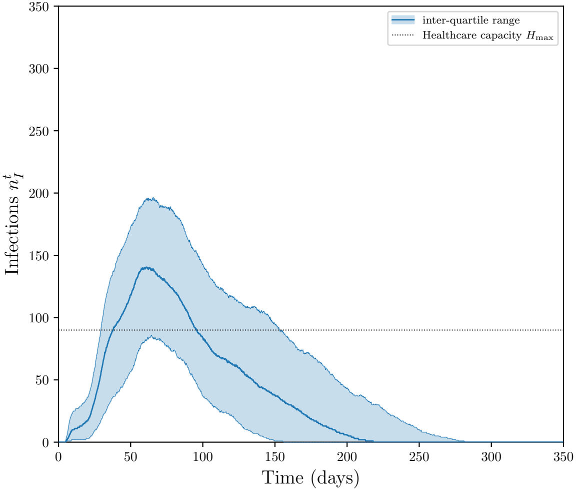
- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response



Background: epidemiological models

What are agent-based epidemiological models?

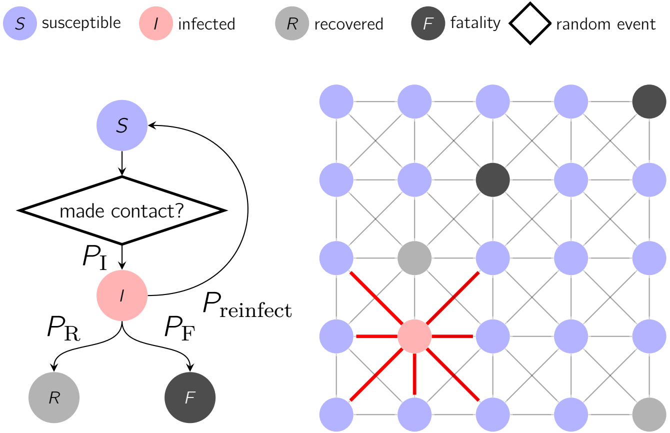
- *Stochastic* process
- Assume *heterogenous* interaction
- Stochastic response



Background: epidemiological models

What are agent-based epidemiological models?

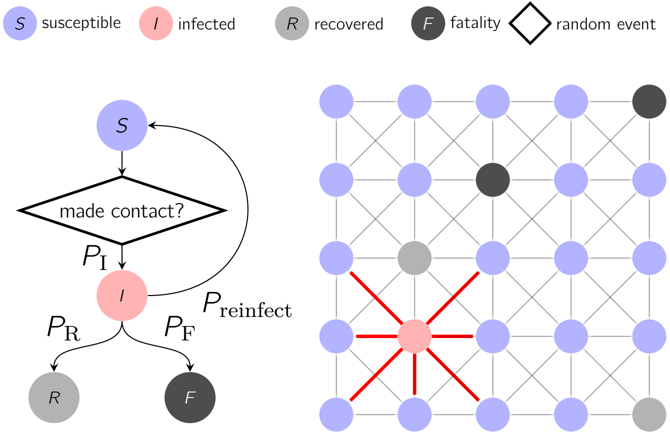
✓ Account for *geography* and *demographics*



Background: epidemiological models

What are agent-based epidemiological models?

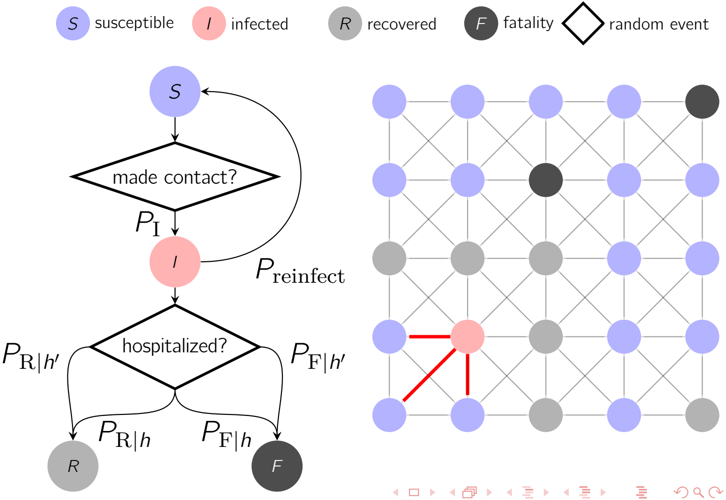
- ✓ Account for *geography* and *demographics*
- ✓ Describe local phenomena



Background: epidemiological models

What are agent-based epidemiological models?

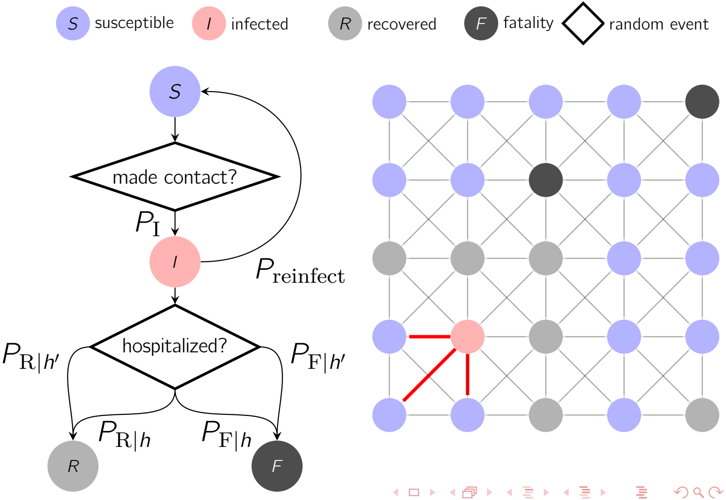
- ✓ Account for *geography* and *demographics*
- ✓ Describe local phenomena
- ✓ Can be used to model intervention policies



Background: epidemiological models

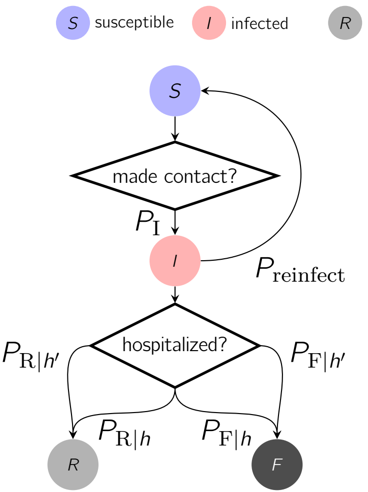
What are agent-based epidemiological models?

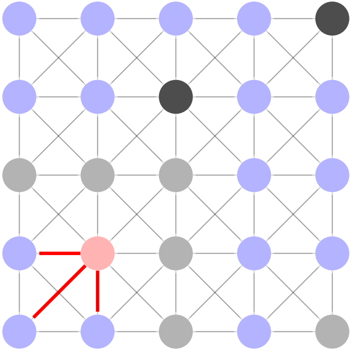
- ✓ Account for *geography* and *demographics*
- ✓ Describe local phenomena
- ✓ Can be used to model intervention policies
- ✘ Stochastic response makes decision making challenging



What are agent-based epidemiological models?

- ✓ Account for *geography* and *demographics*
- ✓ Describe local phenomena
- ✓ Can be used to model intervention policies
- ✗ Stochastic response makes decision making challenging
- ✗ Computationally expensive

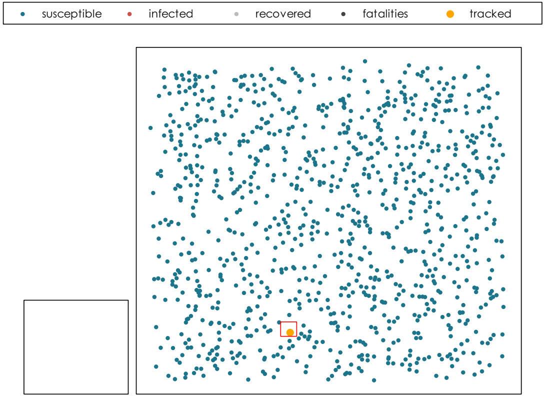




Public health policy-making problem formulation

What is the **cost** of public health interventions?

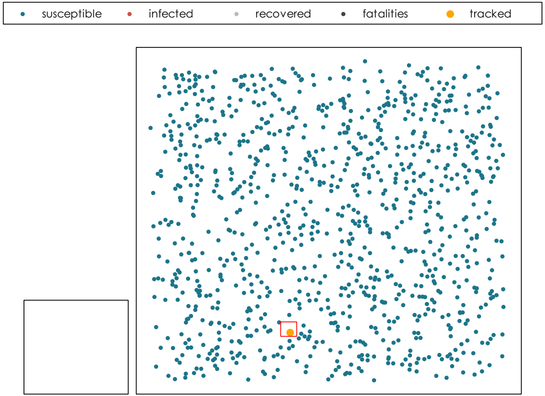
No interventions applied

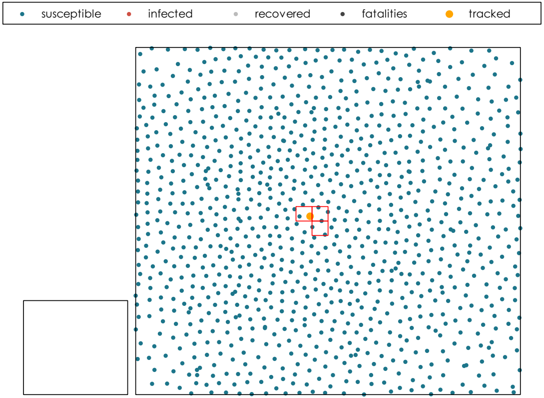


Public health policy-making problem formulation

What is the **cost** of public health interventions?

No interventions applied

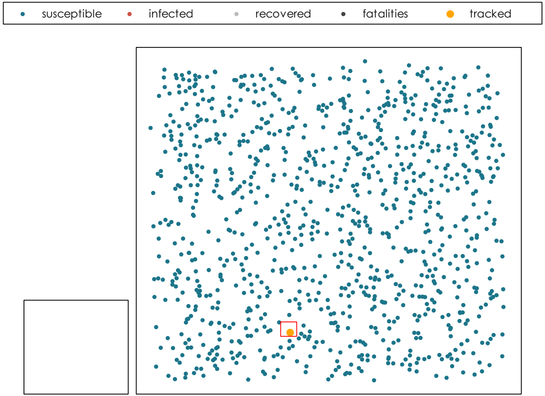




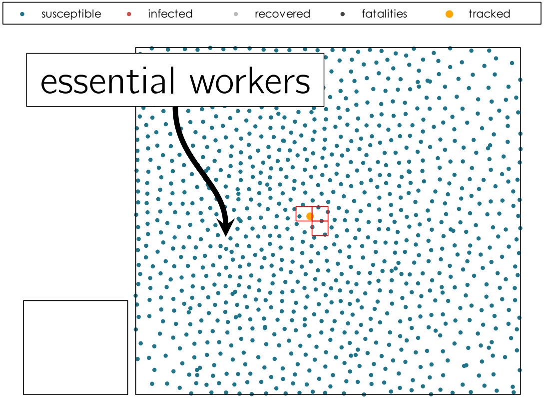
Public health policy-making problem formulation

What is the **cost** of public health interventions?

No interventions applied



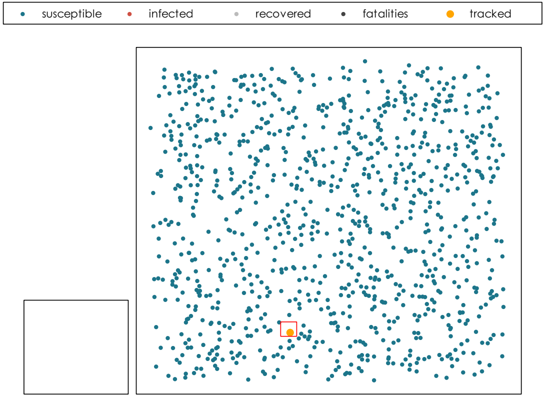
with intervention



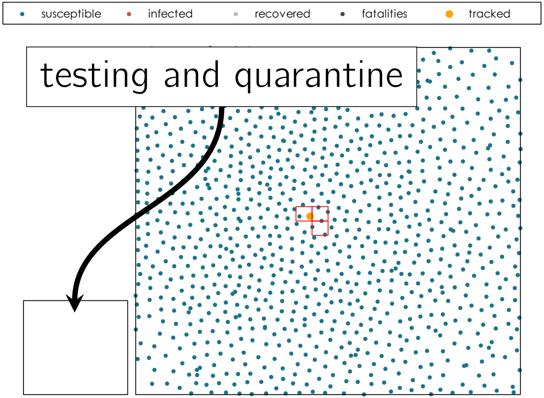
Public health policy-making problem formulation

What is the **cost** of public health interventions?

No interventions applied

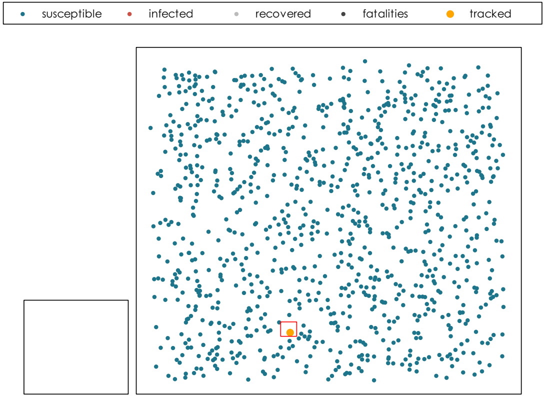


with intervention

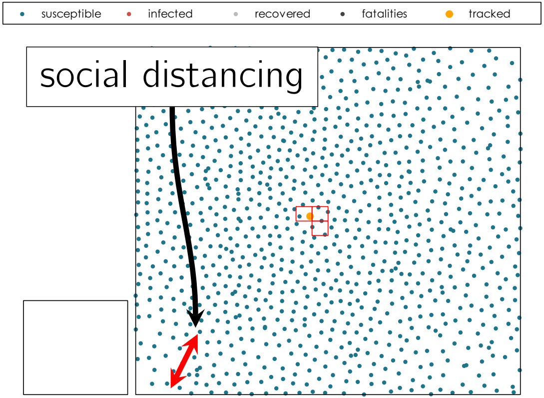


What is the **cost** of public health interventions?

No interventions applied



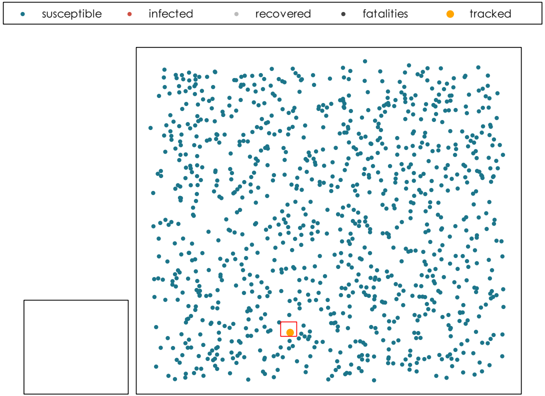
with intervention



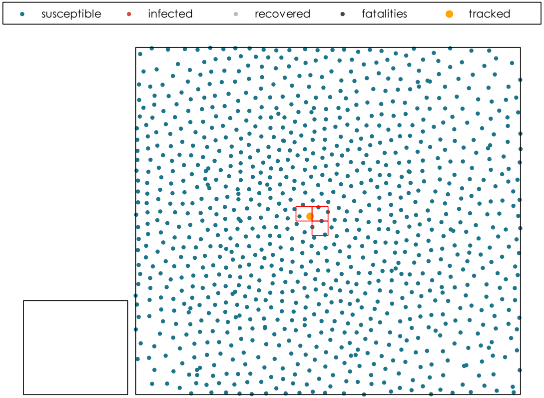
Public health policy-making problem formulation

What is the **cost** of public health interventions?

No interventions applied



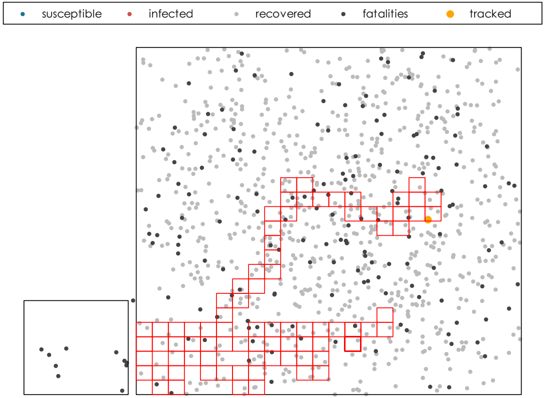
with intervention



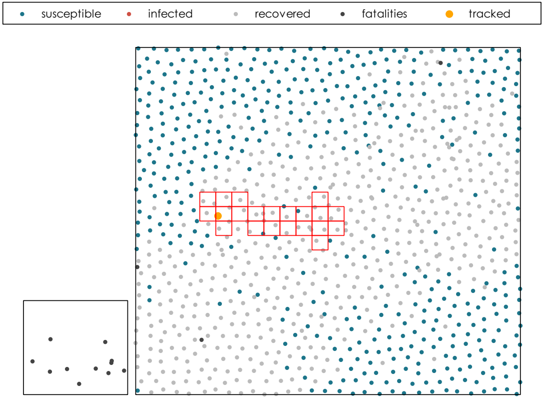
Public health policy-making problem formulation

What is the **cost** of public health interventions?

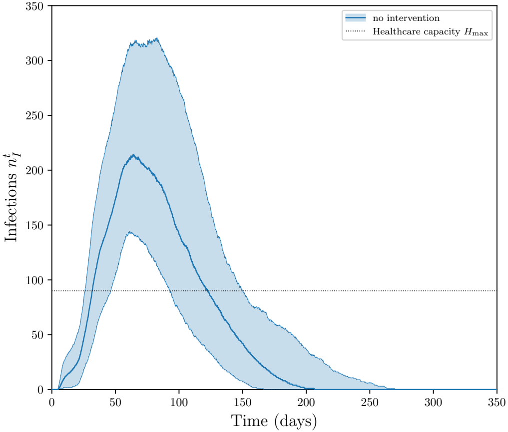
No interventions applied

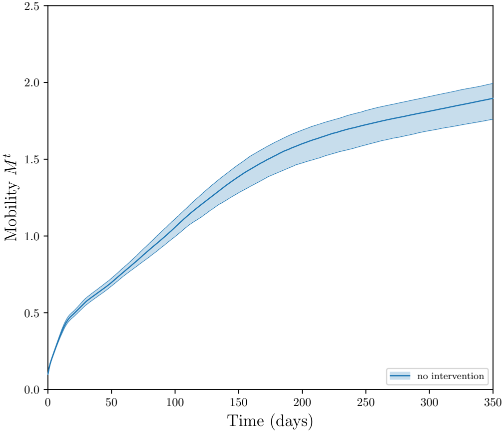


with intervention



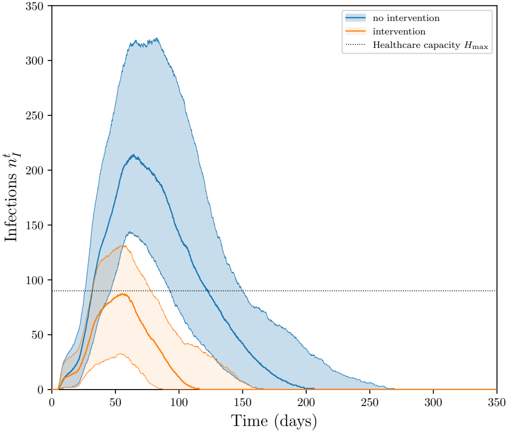
What is the **cost** of public health interventions?



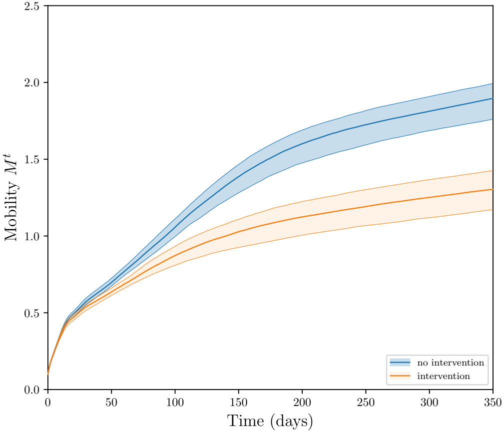


What is the **cost** of public health interventions?

infections ↓



mobility ↓



Optimization problem

Objective and constraints

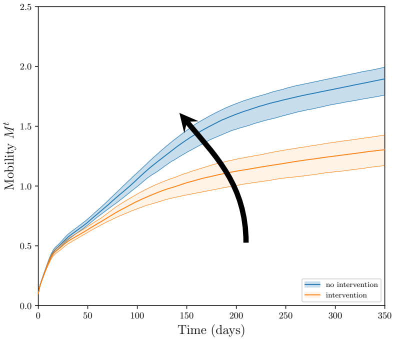
$$\min_{\mathrm{x}} \quad f(\mathrm{x}) = -M^T$$

subject to

where $\quad \mathrm{x} = [n_E, S_D, n_T]^T$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily



Optimization problem

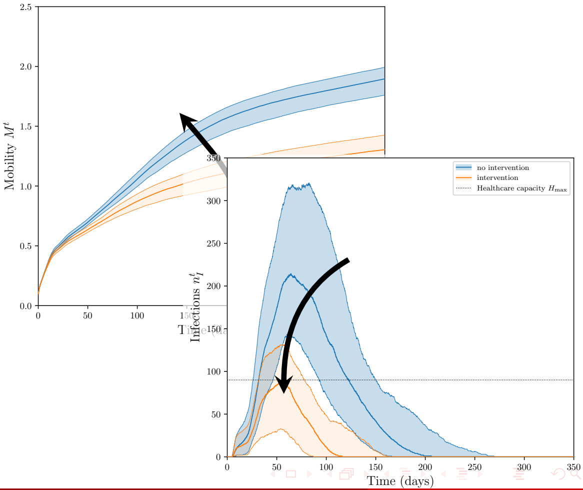
No gradient information available, blackbox is expensive and noisy

Objective and constraints

$$\begin{array}{ll} \min\limits_{x} & f(x) = -M^T \\ \text{subject to} & c(x) \equiv n_{I,\max} - H_{\max} \leq 0 \\ \text{where} & x = [n_E, S_D, n_T]^T \end{array}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily



Optimization problem

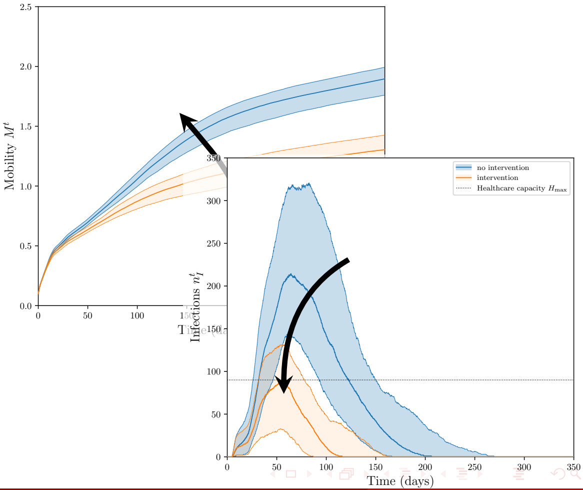
No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

$$\begin{aligned} \min_{\mathrm{x}} \quad & f(\mathrm{x}) = -M^T \\ \text{subject to} \quad & c(\mathrm{x}) \equiv n_{I,\max} - H_{\max} \leq 0 \\ \text{where} \quad & \mathrm{x} = [n_E, S_D, n_T]^T \end{aligned}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily



Optimization problem

No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

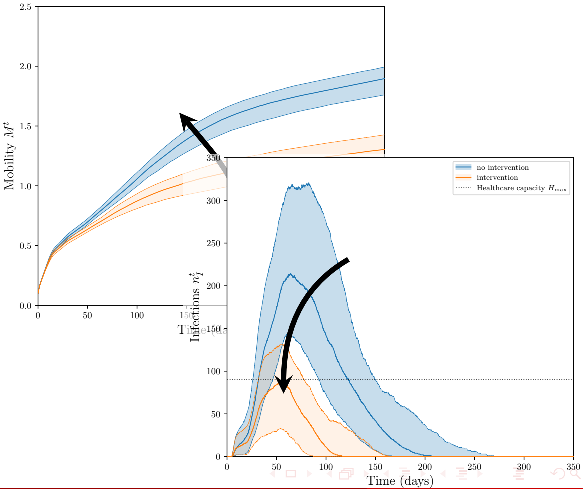
$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = -M^T\right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_\Theta\left[c_\Theta(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0$$

$$\text{where} \quad x = [n_E, S_D, n_T]^T, \ \Theta\text{: realizations}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily



Optimization problem

No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x) = -M^T\right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_{\Theta}\left[c_{\Theta}(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0$$

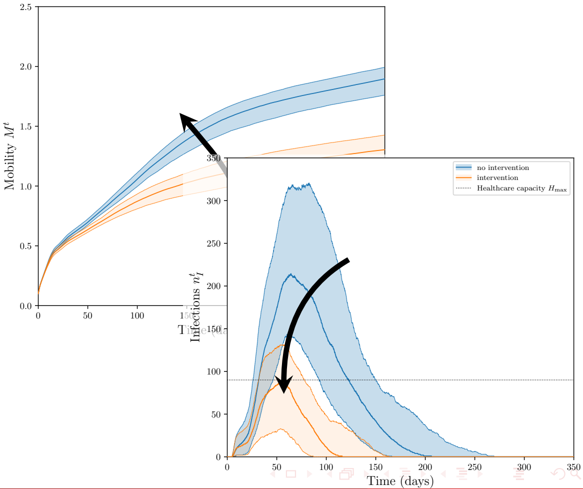
$$\text{where} \quad x = [n_E, S_D, n_T]^T, \ \Theta\text{: realizations}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics



Optimization results

StoMADS, `NOMAD`, and genetic algorithms were used to solve the problem

Objective and constraints

$$\begin{array}{ll} \min\limits_{x} & f(x) = \mathbb{E}_\Theta \left[f_\Theta(x) = -M^T \right] \\ \text{subject to} & c(x) = \mathbb{E}_\Theta \left[c_\Theta(x) \equiv n_{I,\max} - H_{\max} \right] \leq 0 \\ \text{where} & x = [n_E, S_D, n_T]^T, \Theta\text{: realizations} \end{array}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

StoMADS, `NOMAD`, and genetic algorithms were used to solve the problem

Objective and constraints

$$\begin{array}{ll} \min_{x} & f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = -M^T\right] \\ \text{subject to} & c(x) = \mathbb{E}_\Theta\left[c_\Theta(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0 \\ \text{where} & x = [n_E, S_D, n_T]^T, \Theta\text{: realizations} \end{array}$$

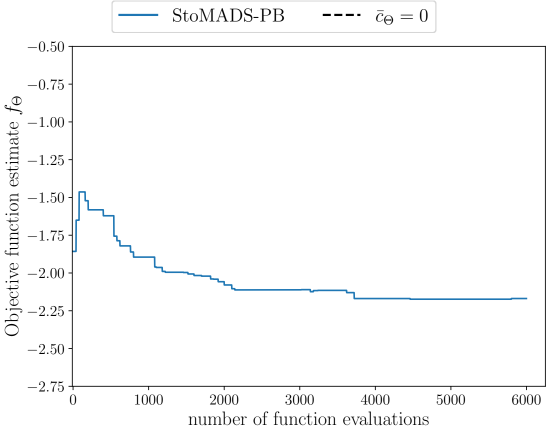
Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

sampling rate $n^k = 4$



Optimization results

StoMADS, NOMAD, and genetic algorithms were used to solve the problem

Objective and constraints

min over x: $f(x) = \mathbb{E}_\Theta \left[f_\Theta(x) = -M^T \right]$

subject to: $c(x) = \mathbb{E}_\Theta \left[c_\Theta(x) \equiv n_{I,\max} - H_{\max} \right] \leq 0$

where: $x = [n_E, S_D, n_T]^T$, Θ: realizations

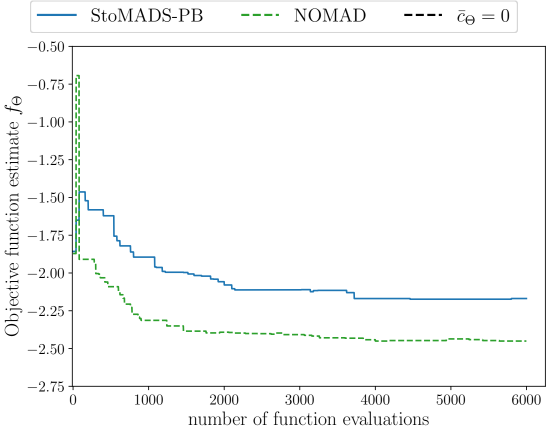
Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

sampling rate $n^k = 4$



StoMADS, NOMAD, and genetic algorithms were used to solve the problem

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta \left[f_\Theta(x) = -M^T \right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_\Theta \left[c_\Theta(x) \equiv n_{I,\max} - H_{\max} \right] \le 0$$

$$\text{where} \quad x = [n_E, S_D, n_T]^T, \Theta\text{: realizations}$$

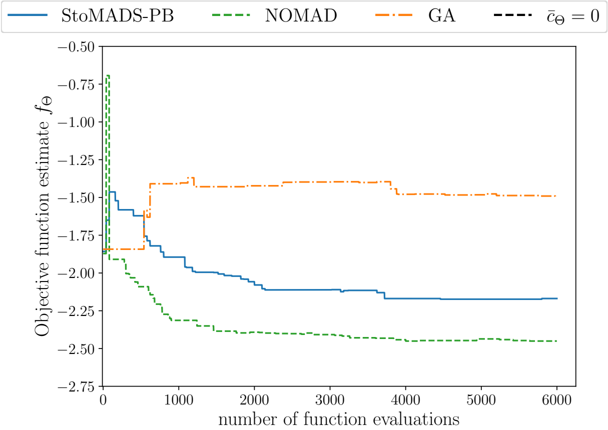
Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

sampling rate $n^k = 4$



Optimization results

StoMADS, `NOMAD`, and genetic algorithms were used to solve the problem

Objective and constraints

$\min\limits_{x}$	$f(x) = \mathbb{E}_\Theta \left[f_\Theta(x) = -M^T \right]$
subject to	$c(x) = \mathbb{E}_\Theta \left[c_\Theta(x) \equiv n_{I,\max} - H_{\max} \right] \leq 0$
where	$x = [n_E, S_D, n_T]^T$, Θ: realizations

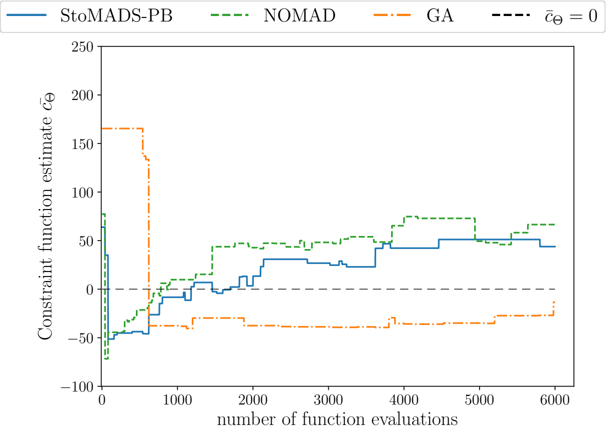
Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

sampling rate $n^k = 4$



StoMADS, `NOMAD`, and genetic algorithms were used to solve the problem
best known solution

Objective and constraints

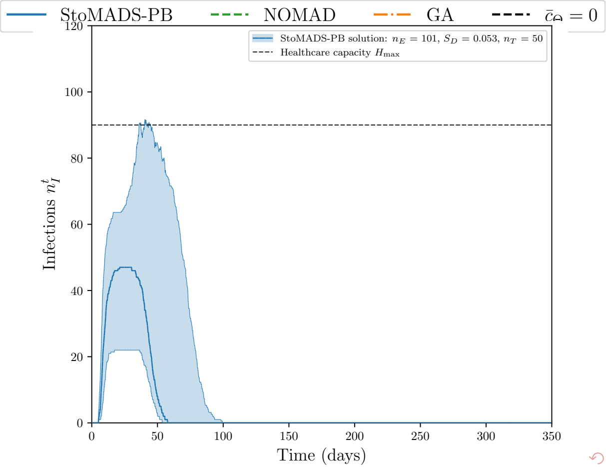
$\min_{x}$	$f(x) = \mathbb{E}_\Theta \left[f_\Theta(x) = -M^T \right]$
subject to	$c(x) = \mathbb{E}_\Theta \left[c_\Theta(x) \equiv n_{I,\max} - H_{\max} \right] \leq 0$
where	$x = [n_E, S_D, n_T]^T$, Θ: realizations

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics



StoMADS, NOMAD, and genetic algorithms were used to solve the problem

best known solution

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = -M^T\right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_\Theta\left[c_\Theta(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0$$

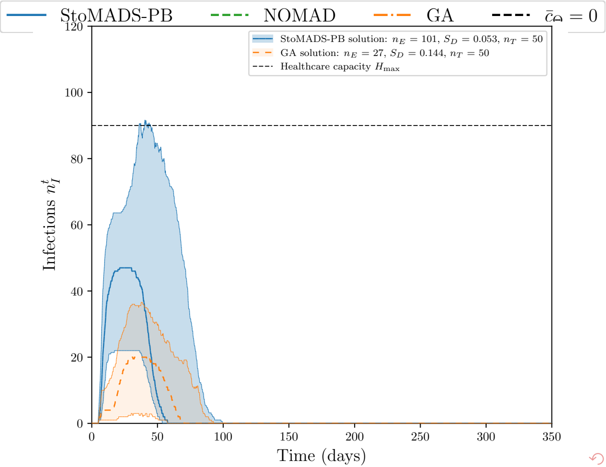
$$\text{where} \quad x = [n_E, S_D, n_T]^T, \Theta\text{: realizations}$$

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics



StoMADS, `NOMAD`, and genetic algorithms were used to solve the problem

best known solution

Objective and constraints

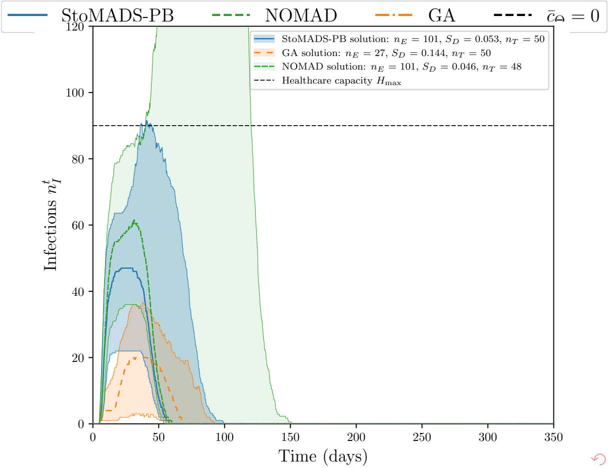
$\min_{x}$	$f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x) = -M^{T}\right]$
subject to	$c(x) = \mathbb{E}_{\Theta}\left[c_{\Theta}(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0$
where	$x = [n_E, S_D, n_T]^T$, Θ: realizations

Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics



Optimization results

StoMADS, NOMAD, and genetic algorithms were used to solve the problem

Objective and constraints

min$_{x}$ $\quad f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x) = -M^{T}\right]$

subject to $\quad c(x) = \mathbb{E}_{\Theta}\left[c_{\Theta}(x) \equiv n_{I,\max} - H_{\max}\right] \leq 0$

where $\quad x = [n_E, S_D, n_T]^T$, Θ: realizations

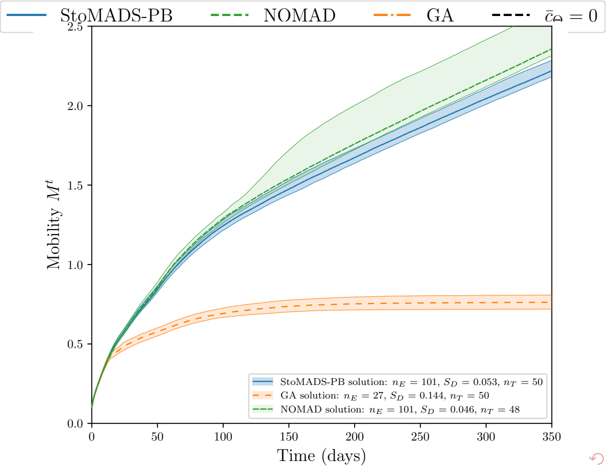
Design variables

- n_E : Number of essential workers
- S_D : Social distancing factor
- n_T : Number of tests daily

Randomly seeded parameters

- Initial conditions
- Interactions, demographics

best known solution

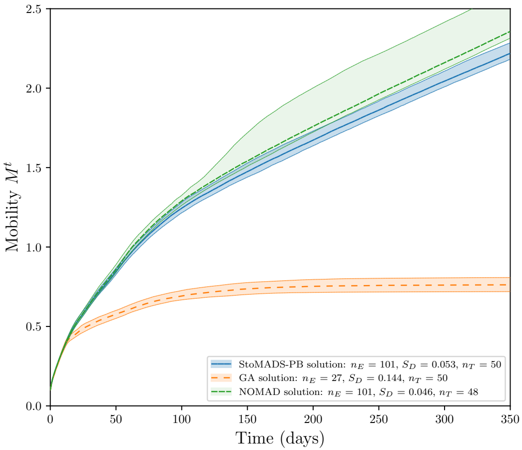


Conclusion

Formulated and solved public health policy-making problems

- Identified a public health policy that favored

best known solution

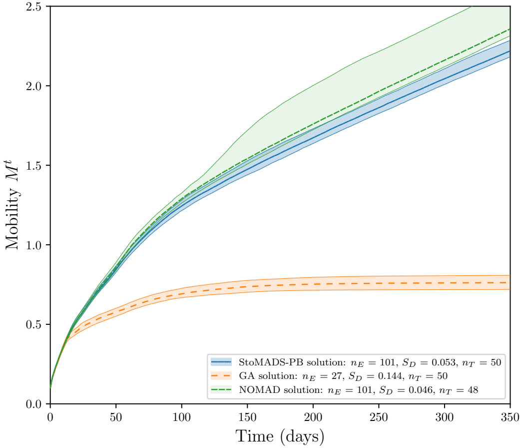


Conclusion

Formulated and solved public health policy-making problems

- Identified a public health policy that favored
  - High testing capacity n_T

best known solution

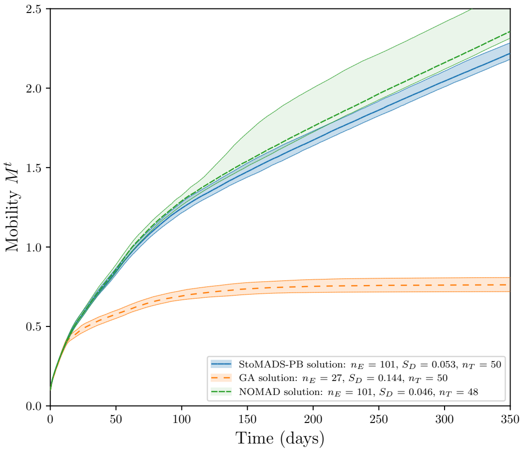


Conclusion

Formulated and solved public health policy-making problems

- Identified a public health policy that favored
  - High testing capacity n_T
  - Large number of essential workers n_E

best known solution

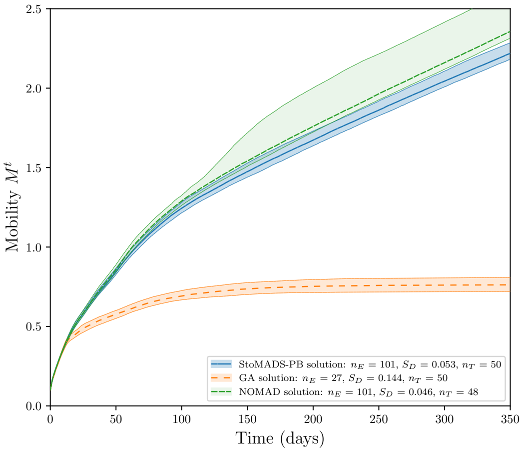


Conclusion

Formulated and solved public health policy-making problems

- Identified a public health policy that favored
  - High testing capacity n_T
  - Large number of essential workers n_E
  - Modest social distancing S_D

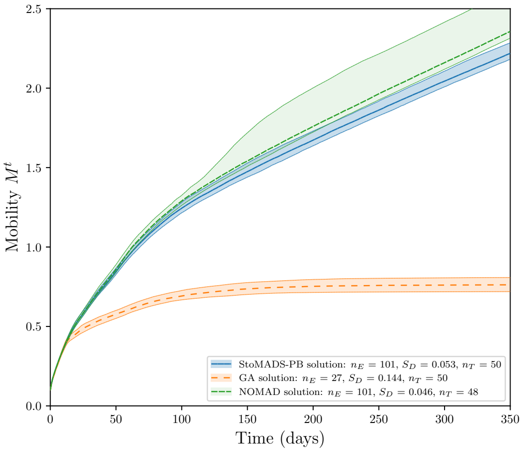
best known solution



Formulated and solved public health policy-making problems

- Identified a public health policy that favored
  - High testing capacity n_T
  - Large number of essential workers n_E
  - Modest social distancing S_D
- StoMADS outperformed GAs and NOMAD on active constraints

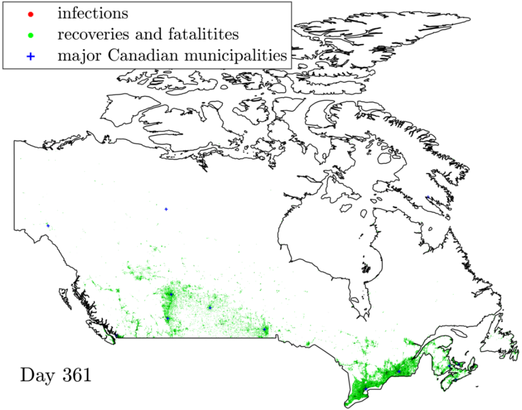
best known solution



Conclusion

Formulated and solved public health policy-making problems

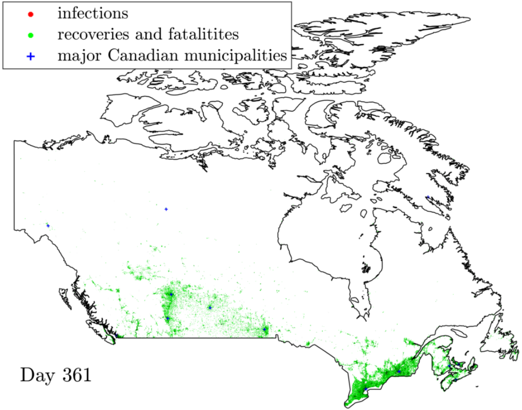
- Identified a public health policy that favored
  - High testing capacity n_T
  - Large number of essential workers n_E
  - Modest social distancing S_D
- StoMADS outperformed GAs and NOMAD on active constraints
- Applicable to other large-scale agent-based models[1]



Conclusion

Formulated and solved public health policy-making problems

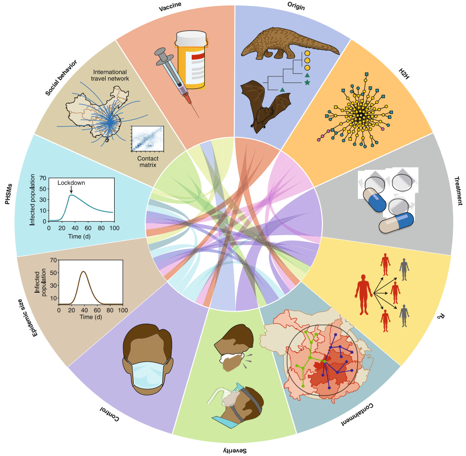
- Identified a public health policy that favored
  - High testing capacity n_T
  - Large number of essential workers n_E
  - Modest social distancing S_D
- StoMADS outperformed GAs and NOMAD on active constraints
- Applicable to other large-scale agent-based models[1]



Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

Inputs for forecasting **epidemic size:**

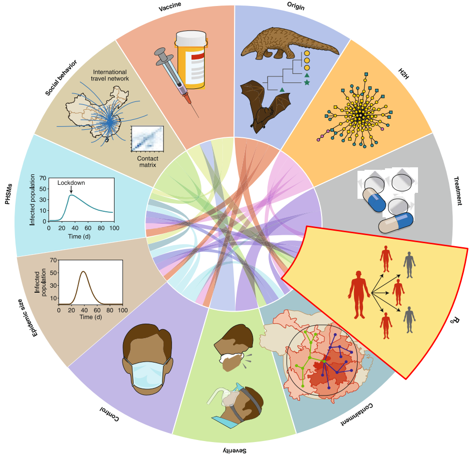


Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

Inputs for forecasting **epidemic size:**

- Growth rate indicators

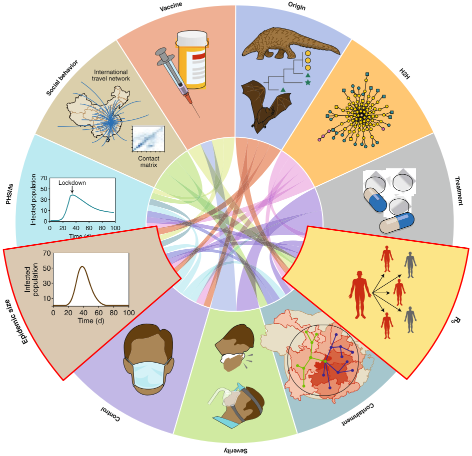


Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

Inputs for forecasting **epidemic size:**

- Growth rate indicators
- Historical incidence rate data

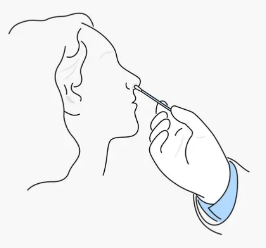


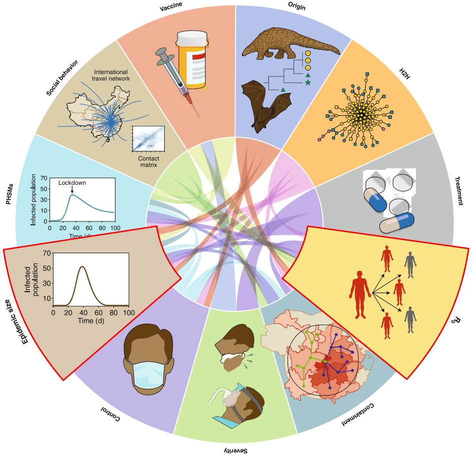
Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

Inputs for forecasting **epidemic size:**

- Growth rate indicators
- Historical incidence rate data
- Serologic assays



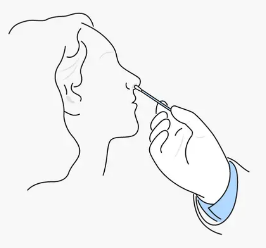


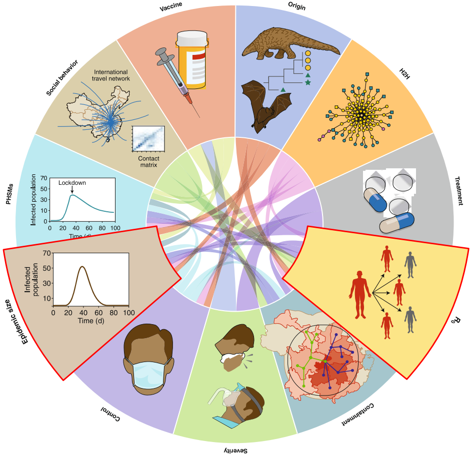
Research displinces in epidemiology

Forecasting novel epidemics is a multidisciplinary field involving multiple *targets*

Inputs for forecasting **epidemic size:**

- Growth rate indicators
- **Historical incidence rate data**
- **Serologic assays**

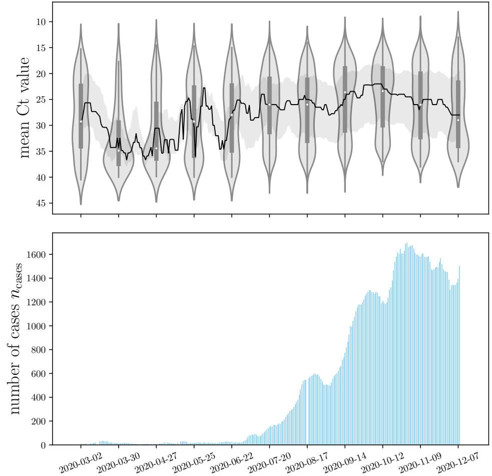




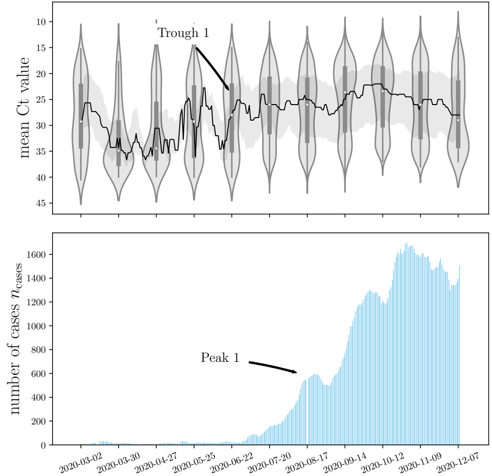
Nowcasting epidemic trajectory using statistical models

Joint work with: Dr. Ibrahim Chamseddine, Dr. Athar Khalil, and Prof. Michael Kokkolaras

How to *design* a machine learning model for forecasting a time-series?

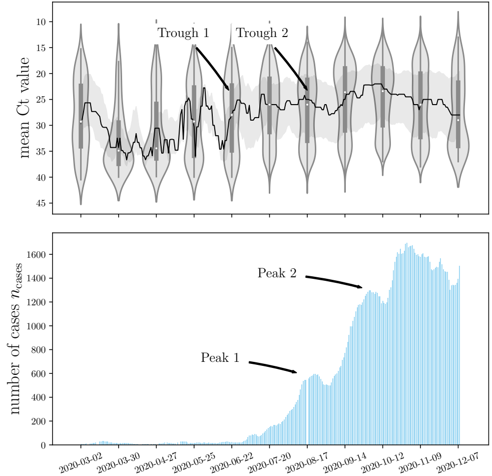


How to *design* a machine learning model for forecasting a time-series?

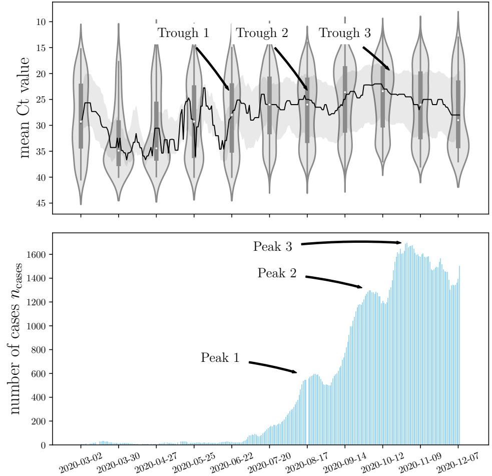


COVID-19 incidence forecasting

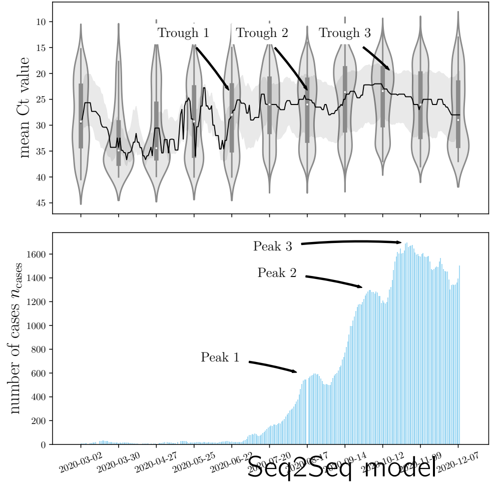
How to *design* a machine learning model for forecasting a time-series?

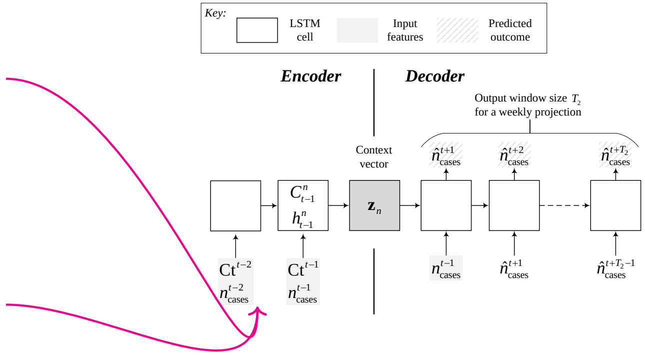


How to *design* a machine learning model for forecasting a time-series?



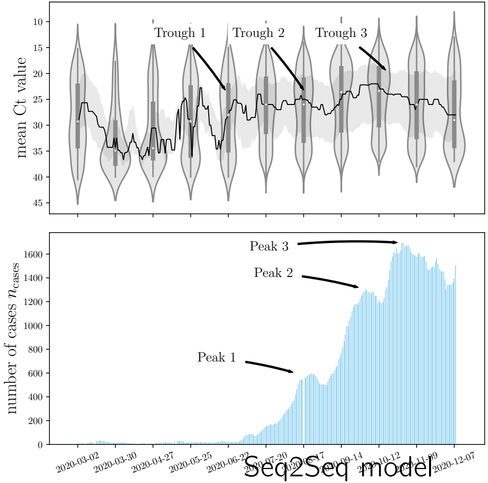
How to *design* a machine learning model for forecasting a time-series?

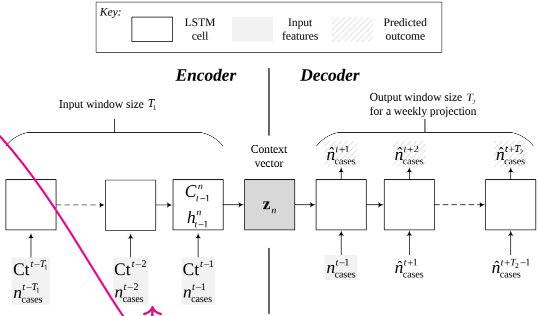




Seq2Seq model

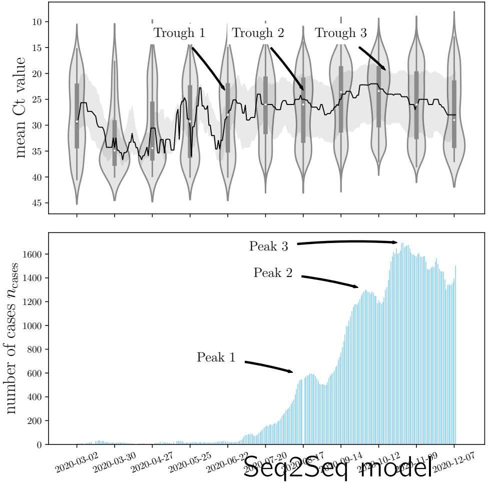
How to *design* a machine learning model for forecasting a time-series?

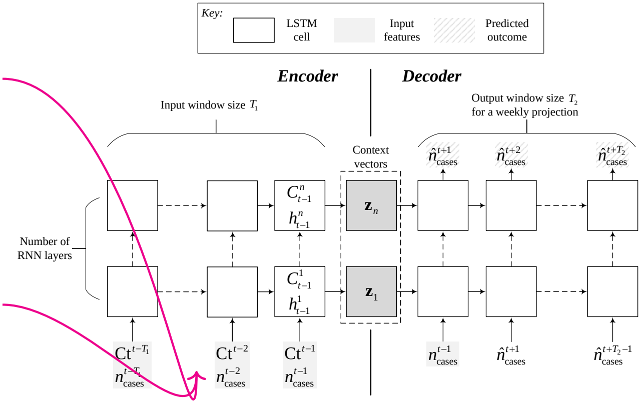




Seq2Seq model

How to *design* a machine learning model for forecasting a time-series?



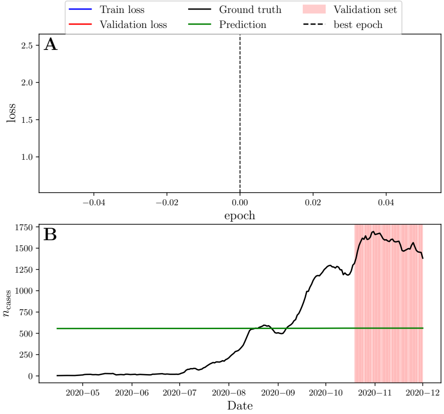


Seq2Seq model

Training the model

There are several challenges associated with hyperparameter optimization

- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting

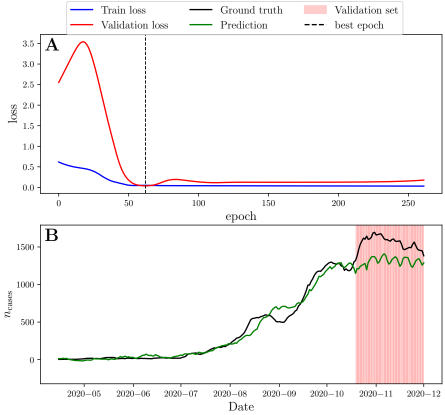


Effect of number of epochs on testing loss

Training the model

There are several challenges associated with hyperparameter optimization

- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting



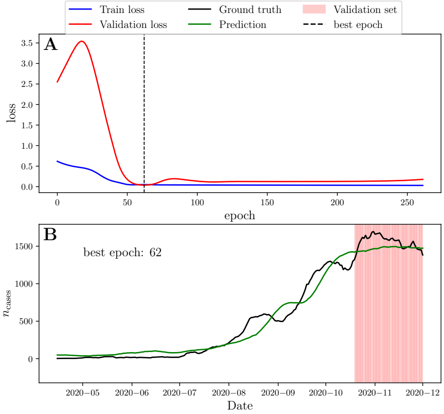
Effect of number of epochs on testing loss

Training the model

There are several challenges associated with hyperparameter optimization

- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting

However,



Effect of number of epochs on testing loss

Training the model

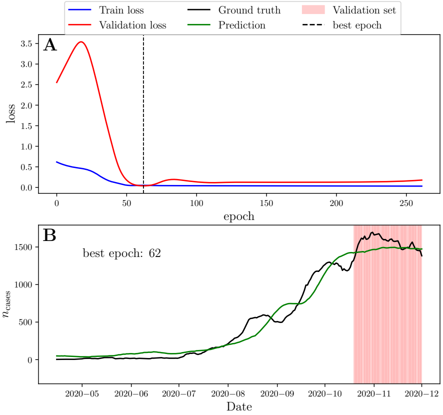
There are several challenges associated with hyperparameter optimization

- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting

However,

1) Other hyperparameters to tune

- Dropout
- Training batch size
- Activation function(s)



Effect of number of epochs on testing loss

Training the model

There are several challenges associated with hyperparameter optimization

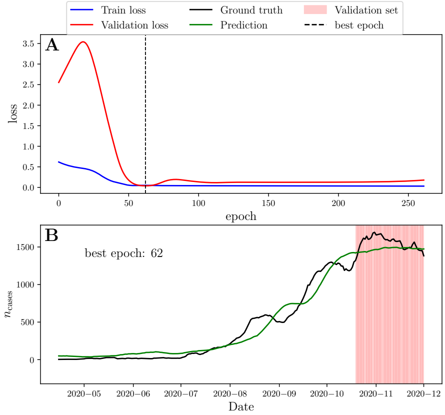
- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting

However,

1) Other hyperparameters to tune

- Dropout
- Training batch size
- Activation function(s)

2) Training can be **expensive**



Effect of number of epochs on testing loss

Training the model

There are several challenges associated with hyperparameter optimization

- The number of epochs can be tuned using *early stopping*
- This is a form of *regularization* to reduce over-fitting

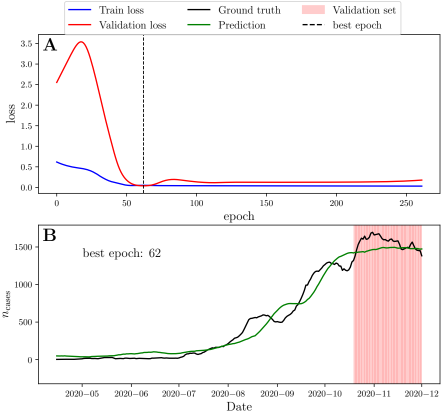
However,

1) Other hyperparameters to tune

- Dropout
- Training batch size
- Activation function(s)

2) Training can be **expensive**

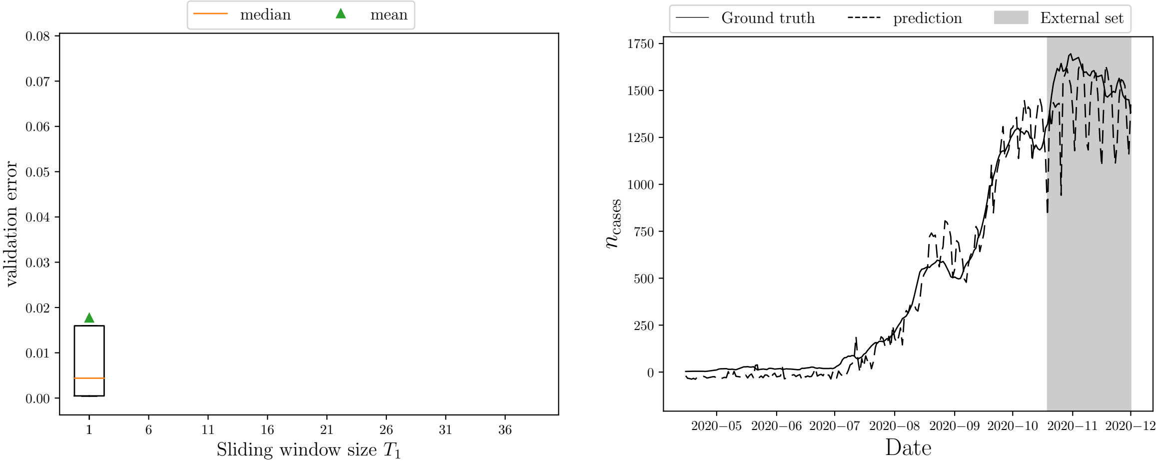
3) Backpropagation is **stochastic**



Effect of number of epochs on testing loss

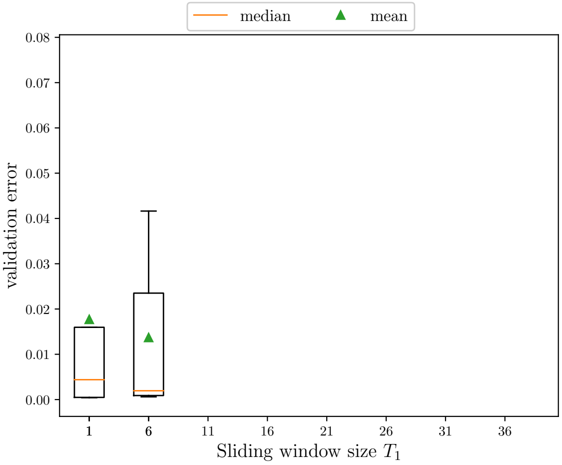
Hyperparameter tuning

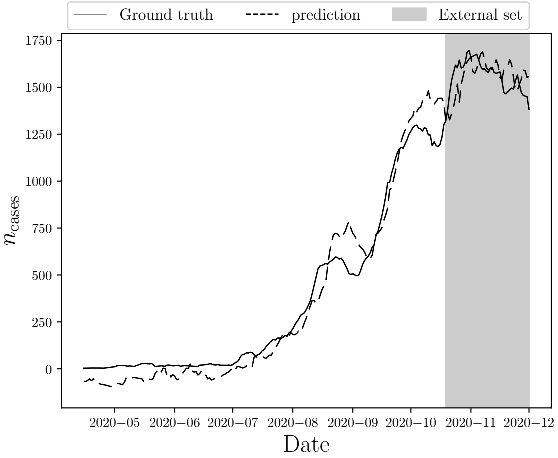
We can use StoMADS to solve such hyperparameter optimization problems



Hyperparameter tuning

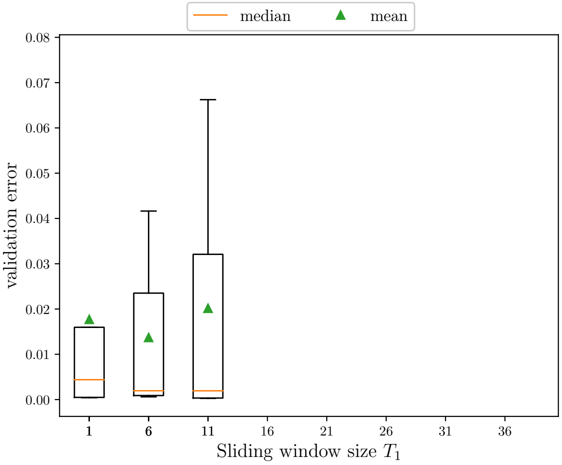
We can use StoMADS to solve such hyperparameter optimization problems

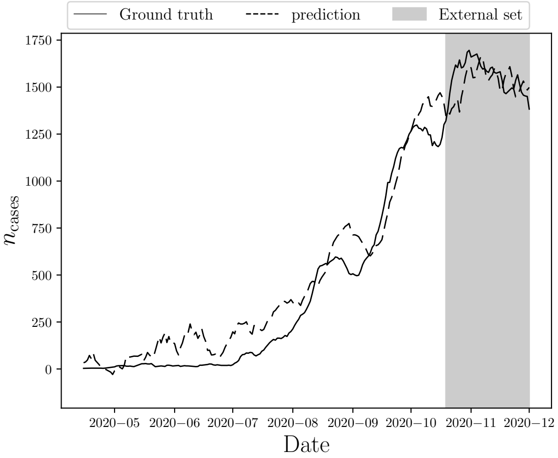




Hyperparameter tuning

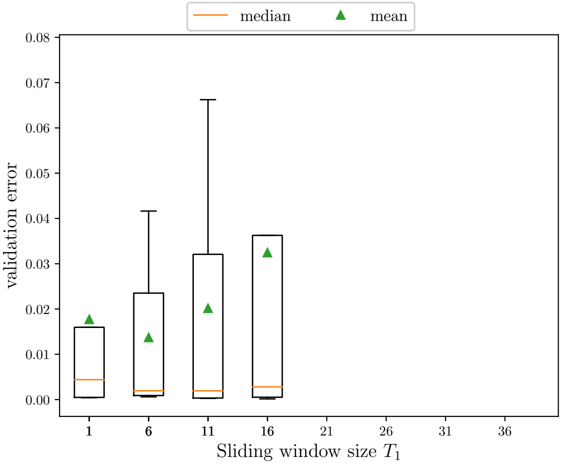
We can use StoMADS to solve such hyperparameter optimization problems

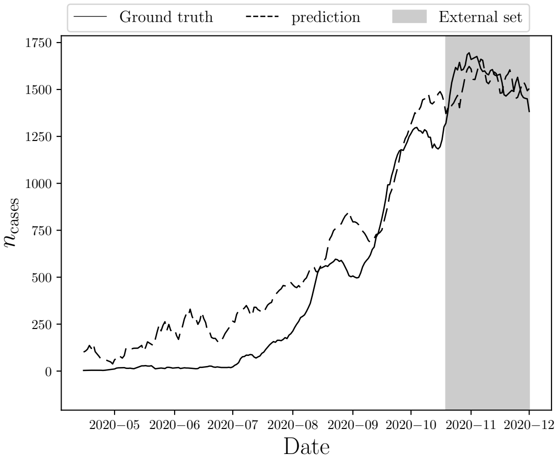




Hyperparameter tuning

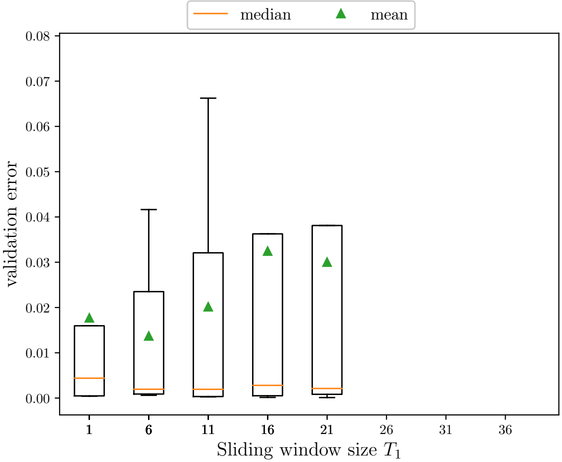
We can use StoMADS to solve such hyperparameter optimization problems

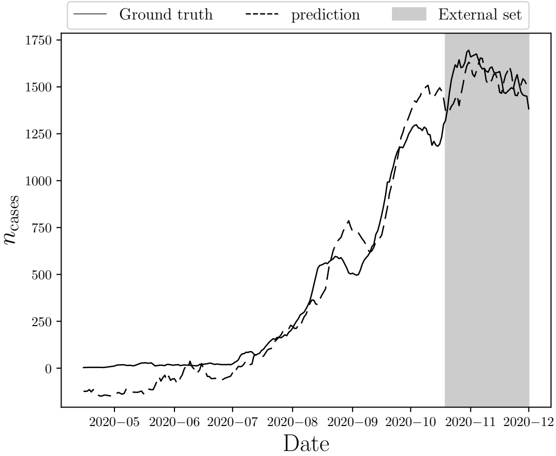




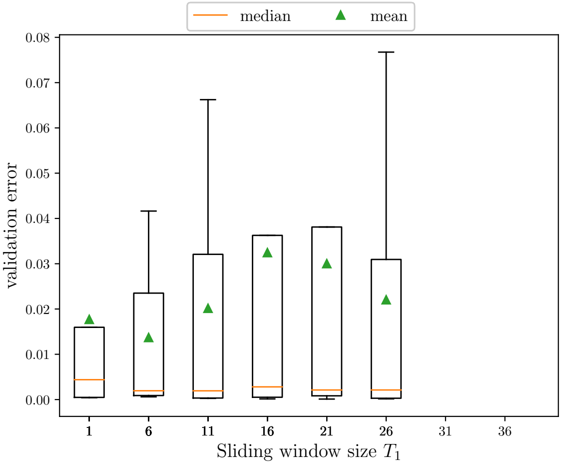
Hyperparameter tuning

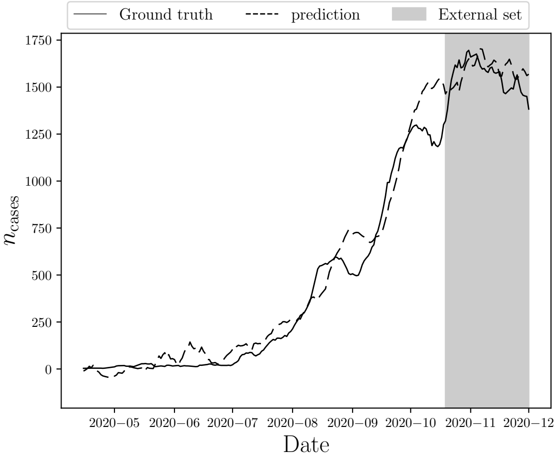
We can use StoMADS to solve such hyperparameter optimization problems





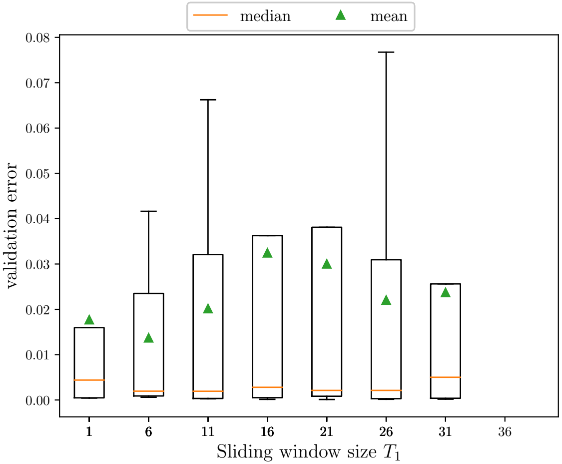
We can use StoMADS to solve such hyperparameter optimization problems

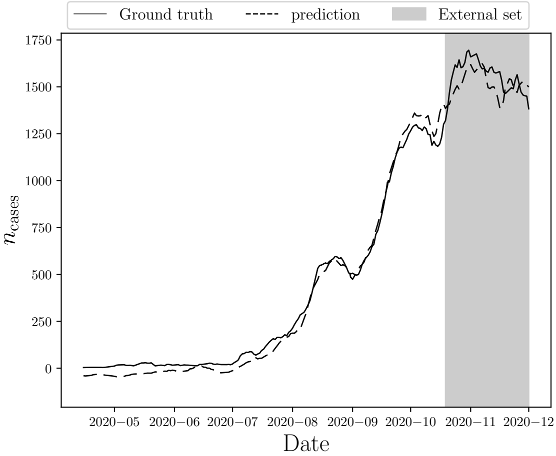




Hyperparameter tuning

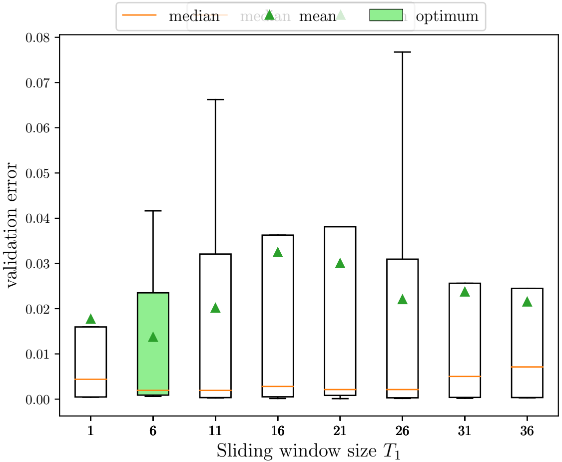
We can use StoMADS to solve such hyperparameter optimization problems

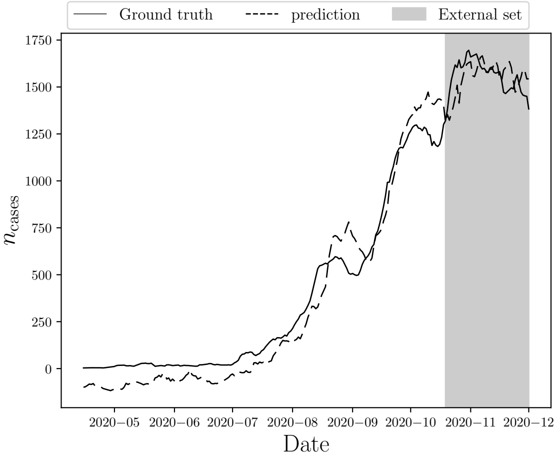




Hyperparameter tuning

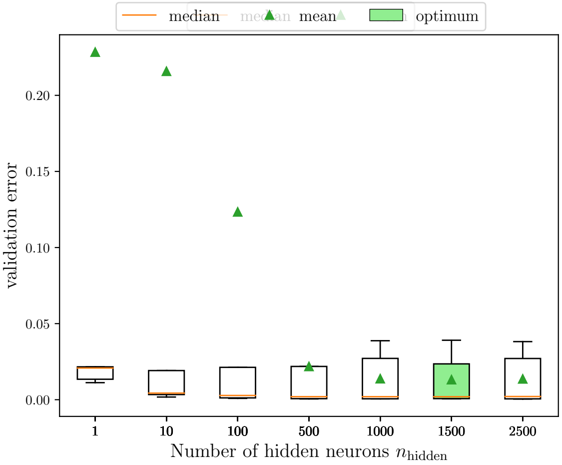
We can use StoMADS to solve such hyperparameter optimization problems

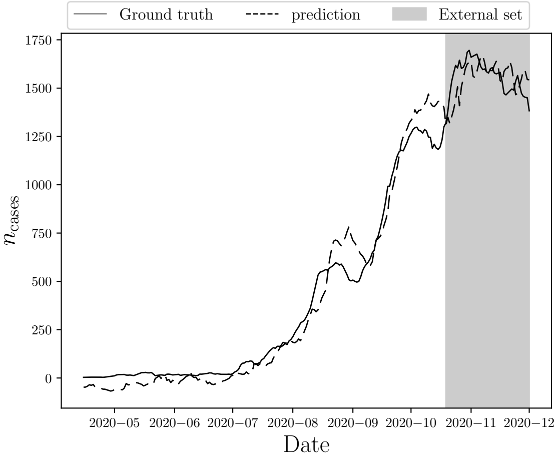




Hyperparameter tuning

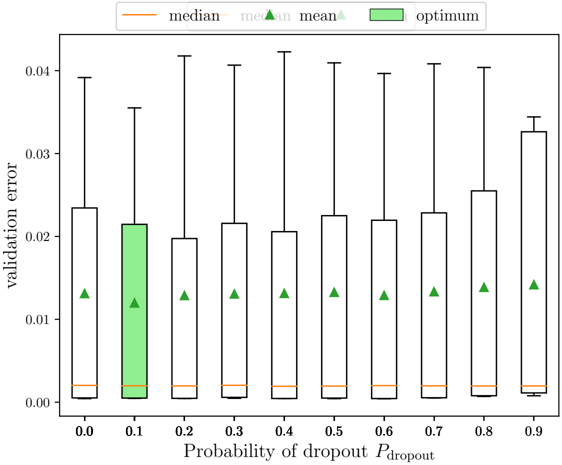
We can use StoMADS to solve such hyperparameter optimization problems

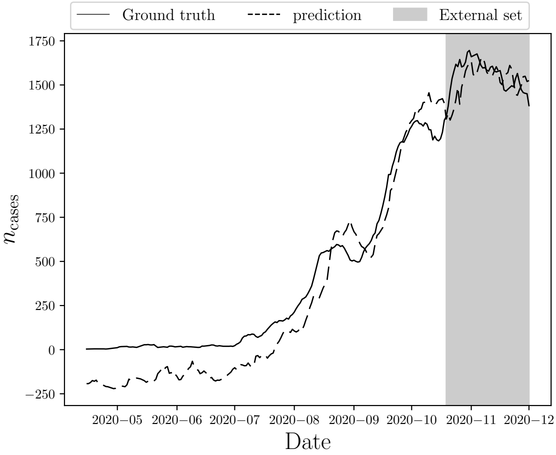




Hyperparameter tuning

We can use StoMADS to solve such hyperparameter optimization problems





Hyperparameter tuning

We can use StoMADS to solve such hyperparameter optimization problems

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = \mathrm{error}_{\mathrm{CV}}\right]$$

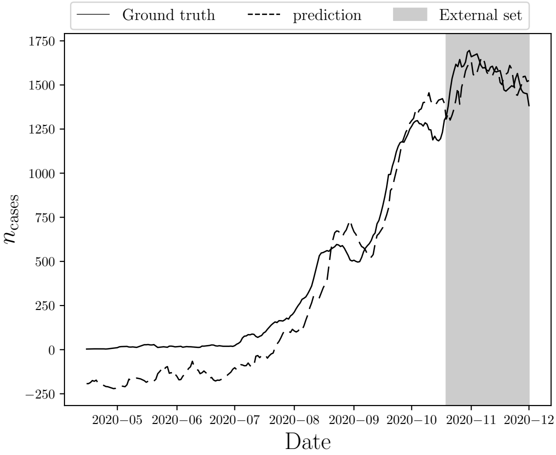
where Θ: realizations

Design variables ($\mathbf{x}$)

- T_1 : Input dimension
- n_{hidden} : Number of hidden neurons
- P_{dropout} : Probability of dropout, etc.

Randomly seeded parameters

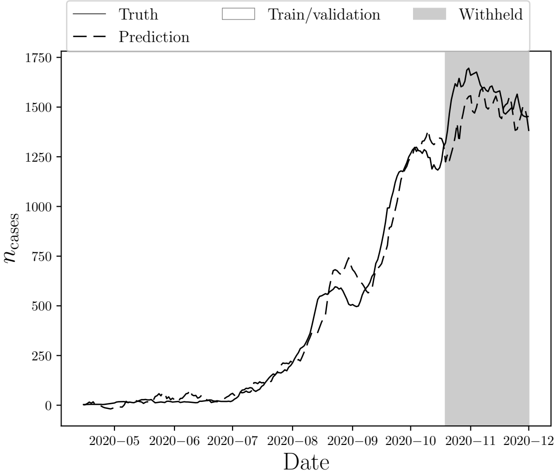
- Initial weights
- Gradient descent steps



Results: Optimal hyperparameters

Optimal hyperparameters for the *Seq2Seq* model:

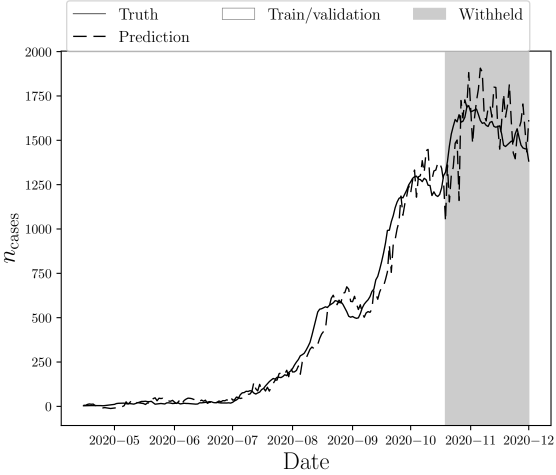
Hyperparameter		**Value**
Sliding window size	T_1	6
Number of hidden neurons	n_{hidden}	1500
Probability of dropout	P_{dropout}	0.8
Number of hidden layers	n_{hidden}	2
Teacher forcing probability	P_{teacher}	0.3
Learning rate	l_{rate}	1×10^{-4}
batch size	b_{size}	32



Results: Optimal hyperparameters

Optimal hyperparameters for the *support vector machine regression (SVR)* model:

Hyperparameter		Value
Sliding window size	T_1	6
Ridge factor	λ	1×10^{-4}
Margin of tolerance	ϵ	1×10^{-2}
Stopping criteria tolerance	ϵ_{tol}	0.1
Learning rate	l_{rate}	1×10^{-5}

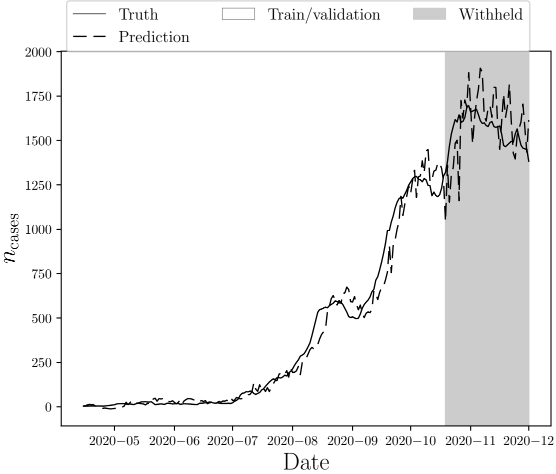


Results: Optimal hyperparameters

Optimal hyperparameters for the *support vector machine regression (SVR)* model:

Hyperparameter		**Value**
Sliding window size	T_1	6
Ridge factor	λ	1×10^{-4}
Margin of tolerance	ϵ	1×10^{-2}
Stopping criteria tolerance	ϵ_{tol}	0.1
Learning rate	l_{rate}	1×10^{-5}

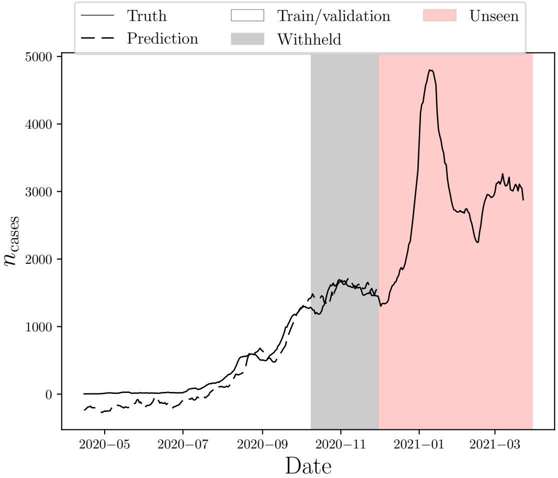
Support vector machine models have **deterministic** performance



Results: Prospective validation

Performance of models on **unseen** data (first 4 months of 2021):

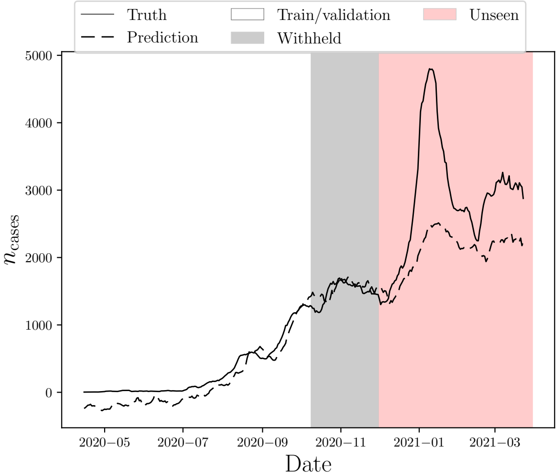
Model	Test error
Seq2Seq	0.571
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	0.168
Gradient boosting	1.444
Linear regression	0.160



Results: Prospective validation

Performance of models on **unseen** data (first 4 months of 2021):

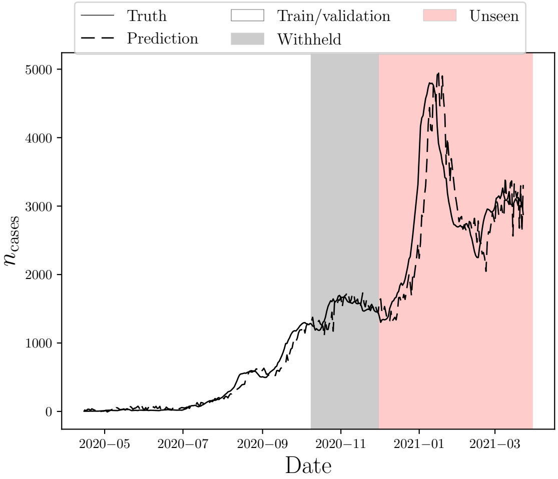
Model	Test error
Seq2Seq	**0.571**
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	0.168
Gradient boosting	1.444
Linear regression	0.160



Results: Prospective validation

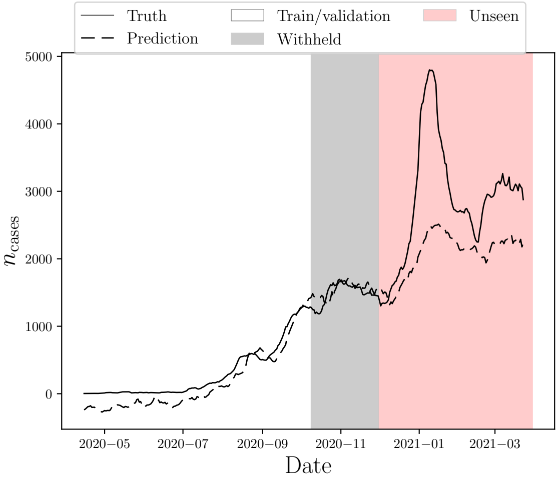
Performance of models on **unseen** data (first 4 months of 2021):

Model	Test error
Seq2Seq	0.571
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	**0.168**
Gradient boosting	1.444
Linear regression	0.160



Effect of increasing number of training days (Adding 1 month of data):

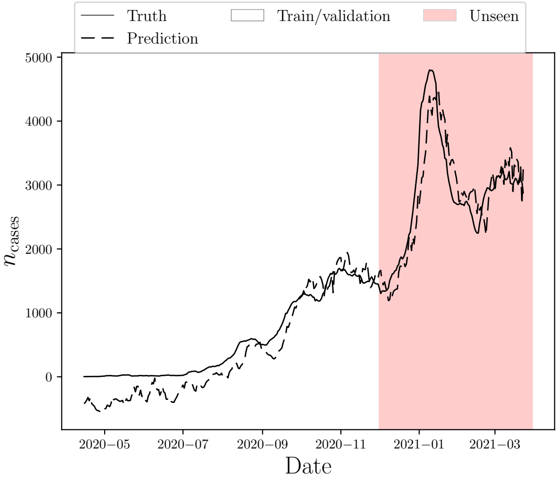
Model	**Test error**
Seq2Seq	**0.571**
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	0.168
Gradient boosting	1.444
Linear regression	0.160



Results: Prospective validation

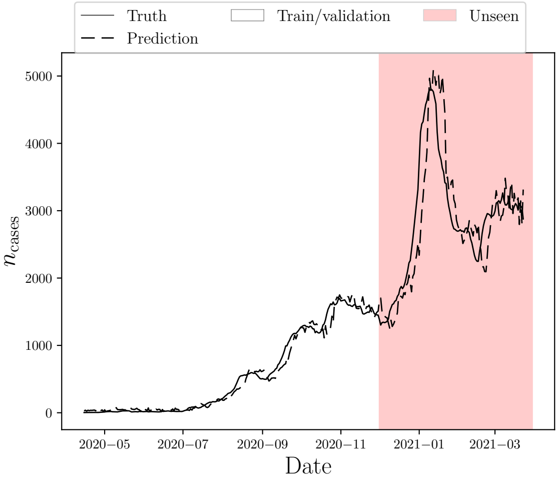
Effect of increasing number of training days (Adding 1 month of data):

Model	Test error
Seq2Seq	**0.106**
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	0.168
Gradient boosting	1.444
Linear regression	0.160



Effect of increasing number of training days (Adding 1 month of data):

Model	Test error
Seq2Seq	**0.106**
Long short term memory (LSTM) cell	0.326
feedforward neural network	0.255
Support vector machine	**0.140**
Gradient boosting	1.444
Linear regression	0.160



Conclusion and future directions

Model discovery and development facilitated by hyperparameter optimization

Conclusion and future directions

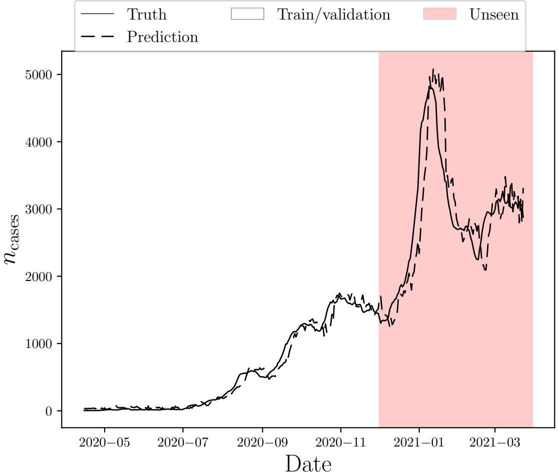
Model discovery and development facilitated by hyperparameter optimization

- Cycle Threshold (Ct) is a useful feature
  for incidence projection

Conclusion and future directions

Model discovery and development facilitated by hyperparameter optimization

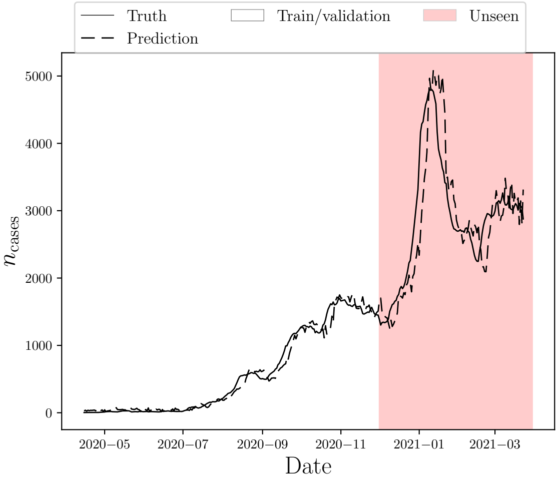
- Cycle Threshold (Ct) is a useful feature for incidence projection
- Model that generalizes well on unseen data



Conclusion and future directions

Model discovery and development facilitated by hyperparameter optimization

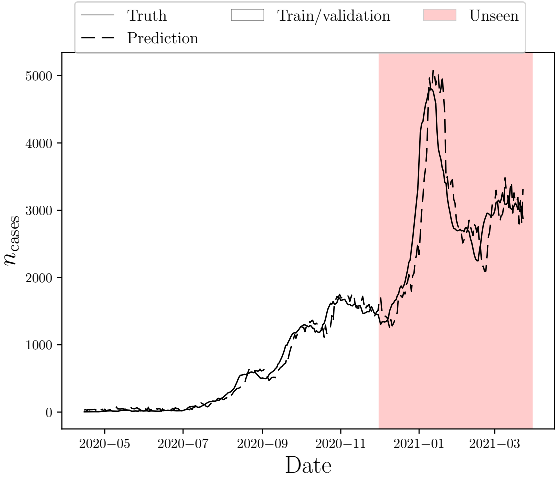
- Cycle Threshold (Ct) is a useful feature for incidence projection
- Model that generalizes well on unseen data
- Simpler models perform well when historical data is limited



Conclusion and future directions

Model discovery and development facilitated by hyperparameter optimization

- Cycle Threshold (Ct) is a useful feature for incidence projection
- Model that generalizes well on unseen data
- Simpler models perform well when historical data is limited
- Works well on other datasets[1]



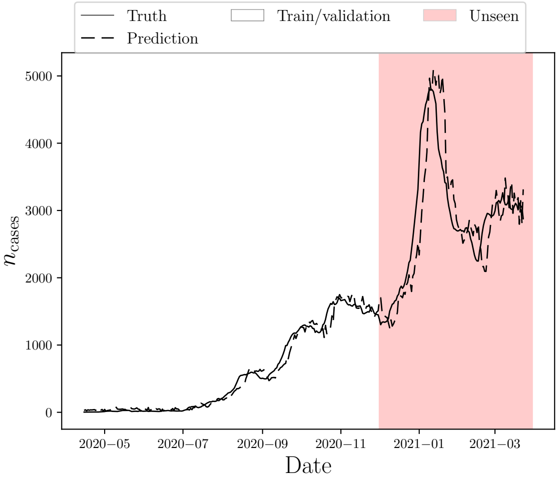
Conclusion and future directions

StoMADS can be improved to solve a wide variety of HPO problems

- Can be used to meet deployment targets

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = \mathrm{error}_{\mathrm{CV}}\right]$$
$$\text{subject to} \quad c(x) = \text{inference time} - \text{threshold} \leq 0$$
$$\text{where} \quad \Theta\text{: realizations}$$



Conclusion and future directions

StoMADS can be improved to solve a wide variety of HPO problems

- Can be used to meet deployment targets

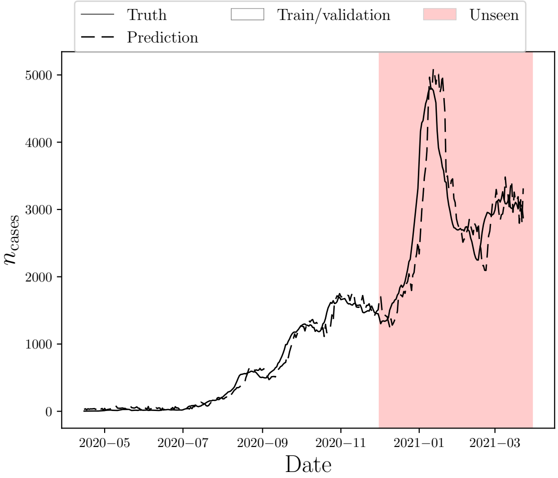
Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x) = \text{error}_{\text{CV}}\right]$$

$$\text{subject to} \quad c(x) = \text{inference time} - \text{threshold} \leq 0$$

$$\text{where} \quad \Theta\text{: realizations}$$

- Should be benchmarked against HyperNOMAD[1], Bayesian optimization[2]



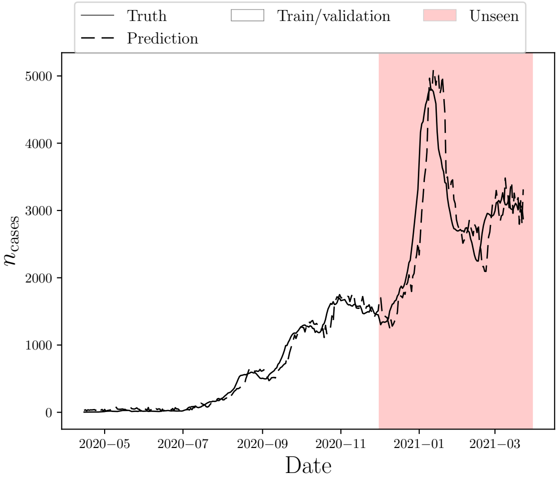
Conclusion and future directions

StoMADS can be improved to solve a wide variety of HPO problems

- Can be used to meet deployment targets

Objective and constraints

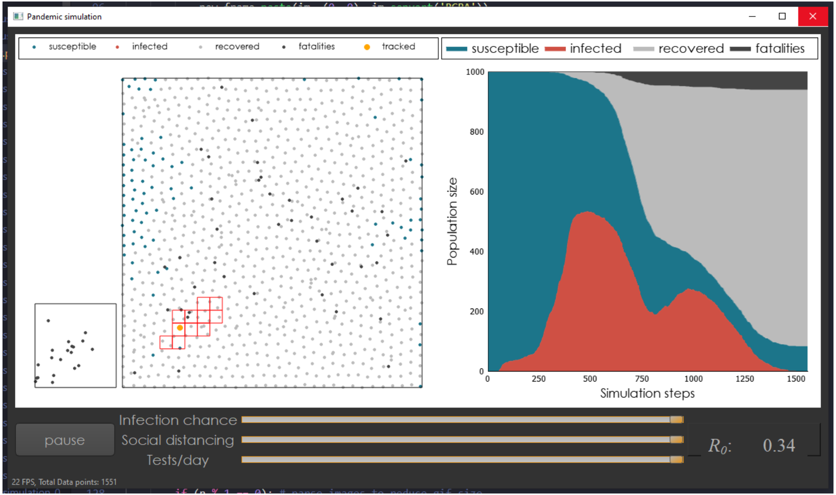
$$\begin{array}{lll} \min\limits_{x} & f(x) = \mathbb{E}_\Theta\left[f_\Theta(x) = \mathrm{error}_{\mathrm{CV}}\right] \\ \text{subject to} & c(x) = \text{inference time} - \text{threshold} \leq 0 \\ \text{where} & \Theta\text{: realizations} \end{array}$$



- Should be benchmarked against HyperNOMAD[1], Bayesian optimization[2]
- Mixed variable version is needed

Thank you for your time

Open source GPU implementation of agent-based modeling



No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x)\right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_\Theta\left[c_\Theta(x)\right] \le 0$$

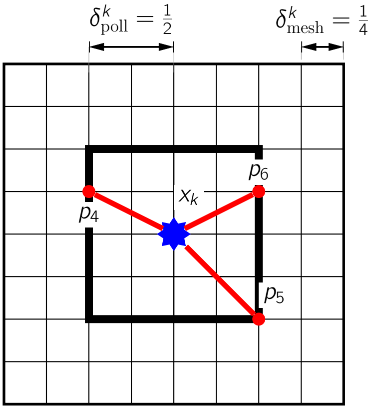
where x : variables Θ: realizations

- Constructs estimates of objective:

$$f^k = \frac{1}{n^k}\sum_{i=1}^{n^k} f_{\Theta_{0,i}}(x_k)$$

- n^k is the sampling rate

Poll **failure**



Overview of stochastic mesh adaptive direct search

No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

$$\min_{x} \quad f(x) = \mathbb{E}_\Theta\left[f_\Theta(x)\right]$$

$$\text{subject to} \quad c(x) = \mathbb{E}_\Theta\left[c_\Theta(x)\right] \le 0$$

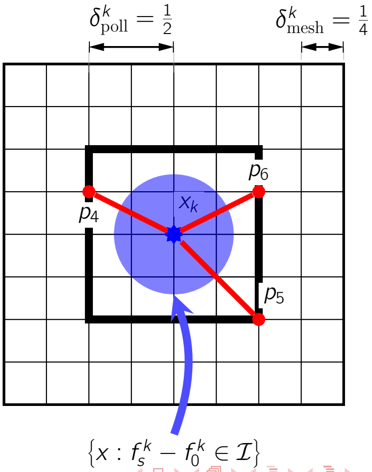
where x : variables Θ: realizations

- Constructs estimates of objective:

$$f^k = \frac{1}{n^k}\sum_{i=1}^{n^k} f_{\Theta_{0,i}}(x_k)$$

- n^k is the sampling rate
- Tracks uncertainty interval $\mathcal{I}(\delta^k_{\text{poll}})$ in the estimate

Poll **certain failure**



Overview of stochastic mesh adaptive direct search

No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

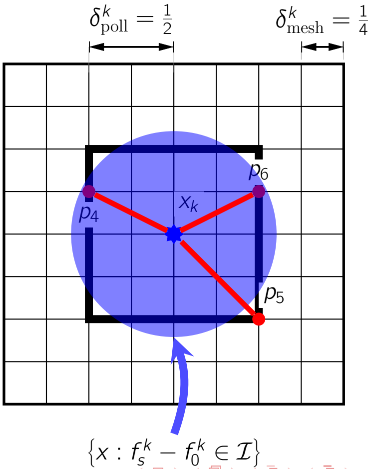
$$\begin{array}{lll} \min_{x} & f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x)\right] \\ \text{subject to} & c(x) = \mathbb{E}_{\Theta}\left[c_{\Theta}(x)\right] \leq 0 \\ \text{where} & x: \text{ variables } \Theta\text{: realizations} \end{array}$$

- Constructs estimates of objective:

$$f^k = \frac{1}{n^k} \sum_{i=1}^{n^k} f_{\Theta_{0,i}}(x_k)$$

- n^k is the sampling rate
- Tracks uncertainty interval $\mathcal{I}(\delta^k_{\text{poll}})$ in the estimate

Poll **uncertain failure**



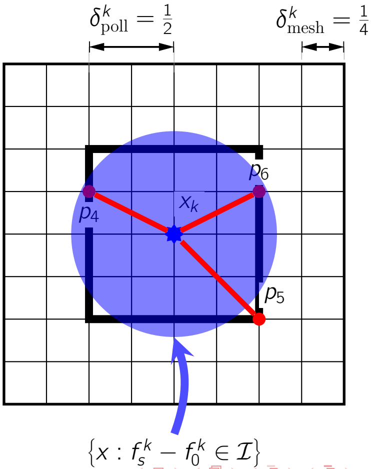
No gradient information available, blackbox is expensive and **noisy**

Objective and constraints

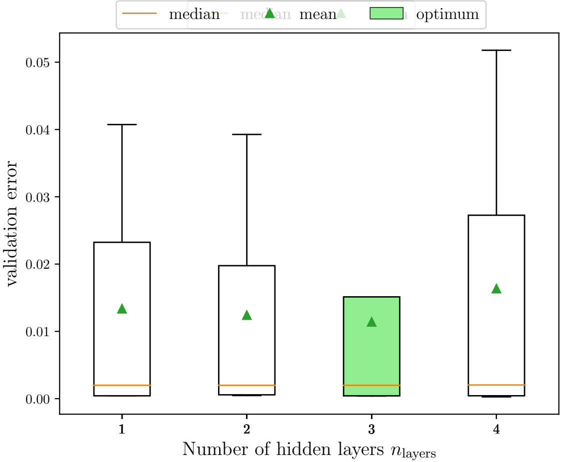
$$\begin{array}{lll} \min\limits_{x} & f(x) = \mathbb{E}_{\Theta}\left[f_{\Theta}(x)\right] \\ \text{subject to} & c(x) = \mathbb{E}_{\Theta}\left[c_{\Theta}(x)\right] \leq 0 \\ \text{where} & x: \text{ variables } \Theta\text{: realizations} \end{array}$$

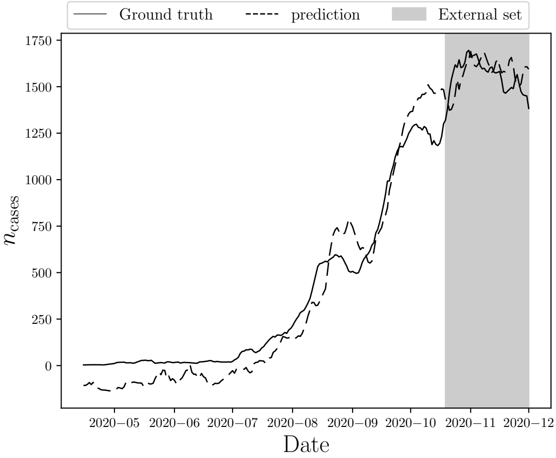
Constraint handling using the *progressive barrier* approach [2]

Poll **uncertain failure**



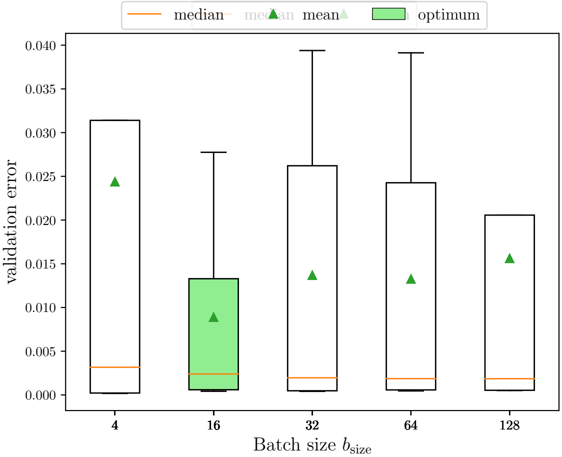
We can use StoMADS to solve such hyperparameter optimization problems

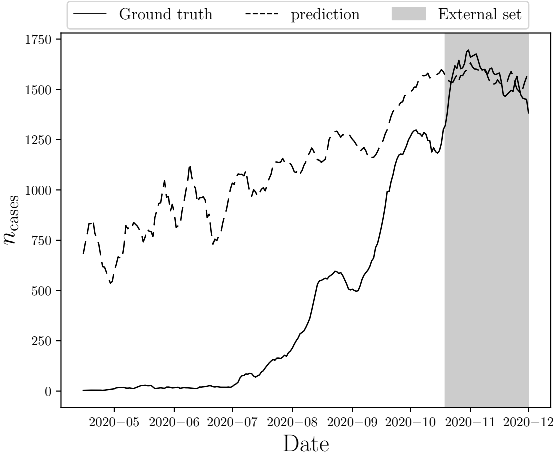




Hyperparameter tuning: other hyperparamters

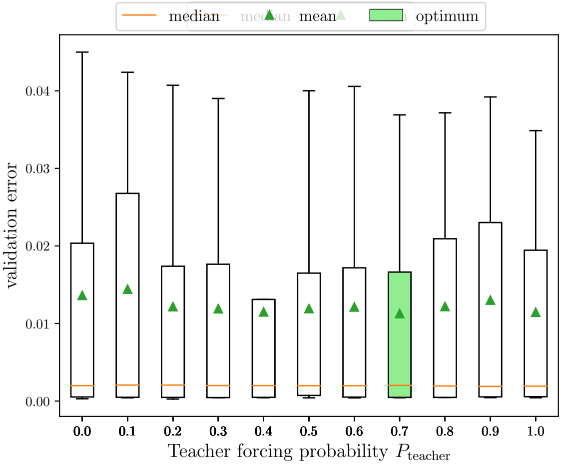
We can use StoMADS to solve such hyperparameter optimization problems

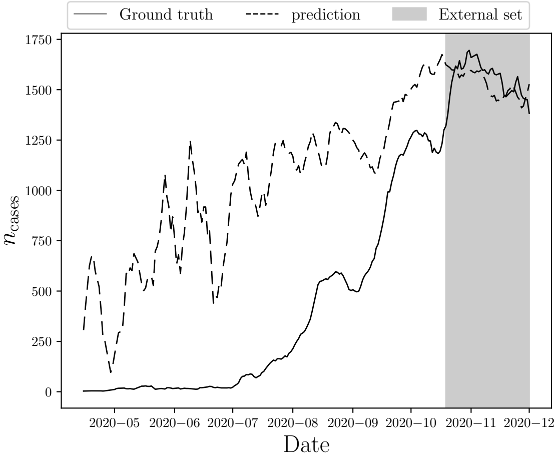




Hyperparameter tuning: other hyperparamters

We can use StoMADS to solve such hyperparameter optimization problems





Hyperparameter tuning: other hyperparamters

We can use StoMADS to solve such hyperparameter optimization problems