EpidemiOptim:

A Toolbox for the Optimization of Control Policies in Epidemiological Models

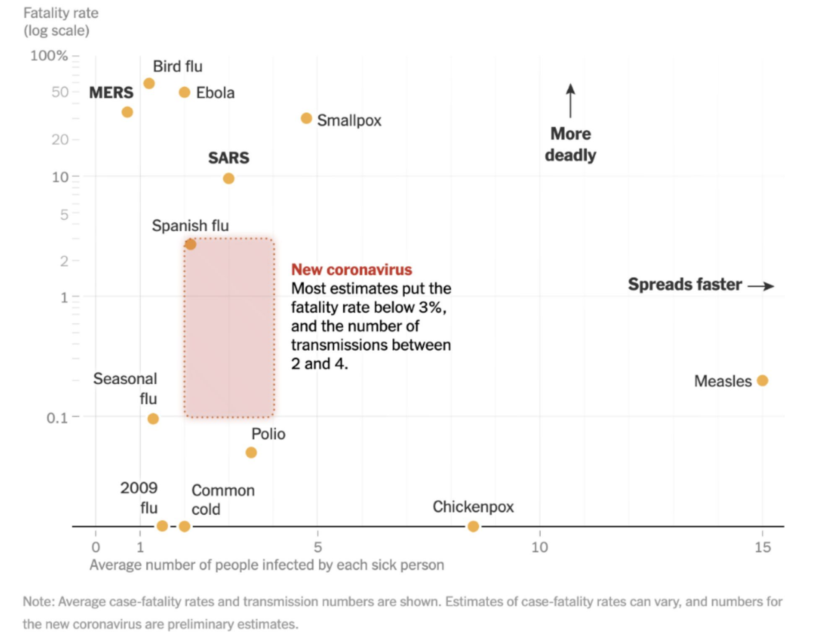
Cédric Colas, Boris Hejblum, Sébastien Rouillon, Rodolphe Thiébaut, Pierre-Yves Oudeyer, Clément Moulin-Frier, **Mélanie Prague**.
Univ. Bordeaux; Inria/Inserm SISTM team; Vaccine research institute; Inria FLOWERS team

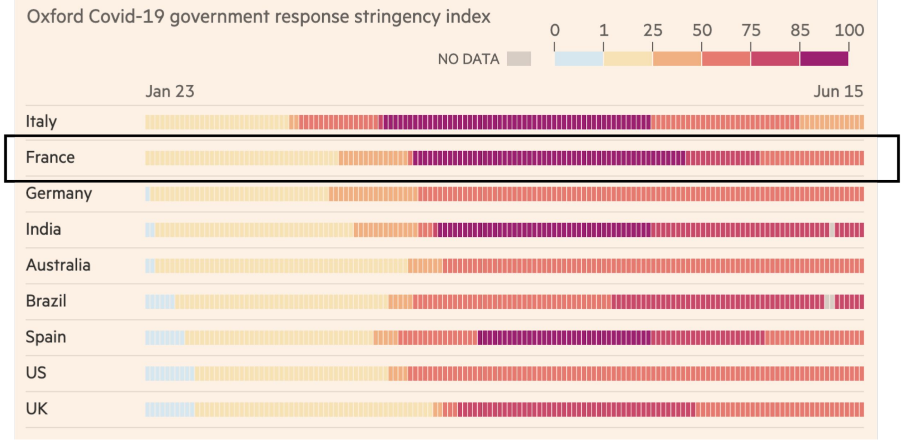
informatiques mathématiques
Inserm

Motivations

SARS-Cov-2 appeared in Wuhan (China) in December 2019
No Vaccine until December 11th 2020

Worldwide implementation of **Non-pharmaceutical Intervention**: from less stringent (masks, hand washing…) to most stringent complete lock-down.





Second, third waves ?
Need for on/off strategies

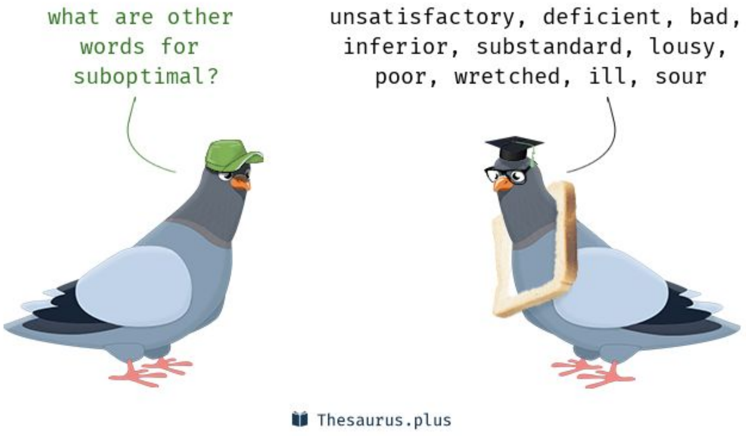
Designing Intervention Strategies is Difficult

Pre-defined strategies are bound to be **suboptimal:**

1) the space of potential strategies can be large, heterogeneous and multi-scale (Halloran et al., 2008);

2) their impact on the epidemic is often difficult to predict (Ferguson et al. 2020; Salje et al. 2020; Prague et al. 2020);

3) the problem is multi-objective by essence: it often involves public health objectives like the minimization of the death toll or the saturation of intensive care units, but also societal and economic sustainability.



Related Approaches - Existing works

High-level contributions
Identification of methods available to solve the problem

(Alamo et al. 2020) **Road-map** from access to data to decision step.

(Shearer et al. 2020) **How to** define social political, ethical, epidemiological (...) costs.

(Yanez et al. 2020) Description of **general framework** for disease spread control based on reinforcement learning in general.

Computational contributions
Implementation of optimization processes

Different epidemiological models(Yaesoubi et al. 2020, Chandak et al. 2020, Kompella et al. 2020).

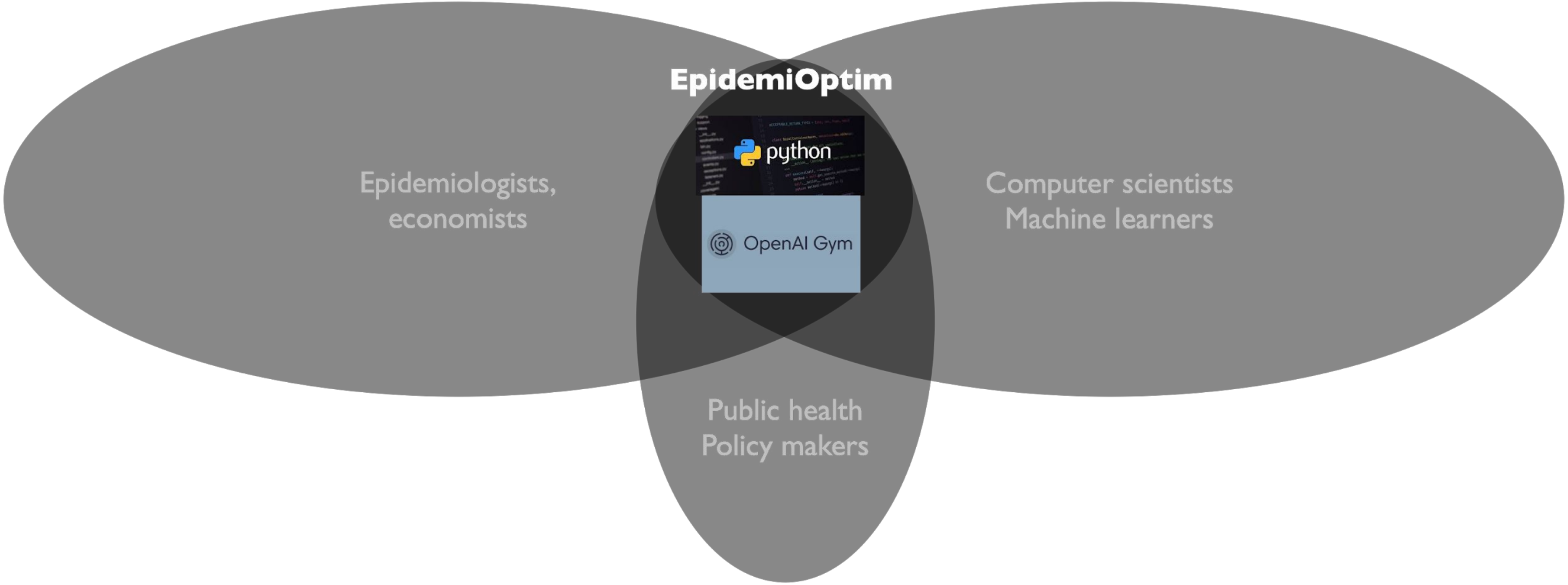
Different optimization methods
(Tarracata et al 2020, Chandak et al. 2020, Arango et al. 2020, Charpentier et al. 2020, Miikakulainen et al. 2020, Elie et al. 2020).

Different cost functions (Libin et al. 2020, Probert et al. 2019, Yaesoubi et al. 2016).

Tool
that can be easily
used, configured and interpreted
by decision-makers

Contribution - The EpidemiOptim ibrary

EpidemiOptim is a **toolbox** that provides a framework to facilitate collaborations between researchers in epidemiology, economics and machine learning.



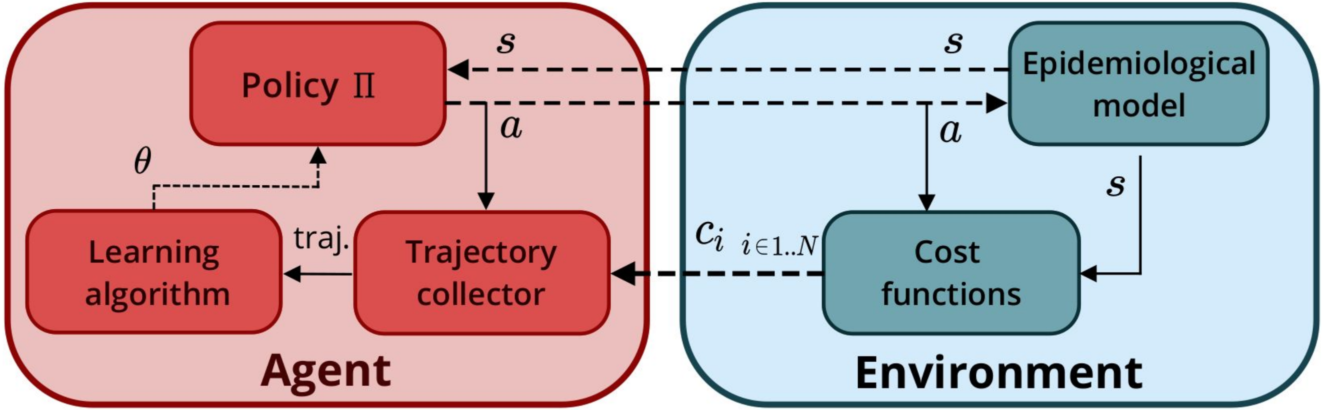
The Epidemic Control Problem

Epidemic Control Problem and RL framework

The Epidemic Control Problem: Find intervention strategies to mitigate the impact of an epidemic.

Requires:

→ one or several epidemiological model(s) (which epidemic?)
→ cost functions to define the objective (what do we want to mitigate?)
→ action modalities (what is the space of intervention strategies?)
→ an optimization algorithm (how do we search the space of intervention strategies?)



Peculiarities of the Epidemic Control Problem

A Multi Objective Problem:

The epidemic control problem is multi-objective by essence (e.g. health and economic costs).
→ using an aggregated cost as convex combinations of costs,
→ using multi-objective algorithms.

Handling Time-Limits:

Optimization algorithm often need finite learning episodes, but epidemics might not be finite. We want to avoid an "after me, the flood" effect.

→ RL agents must be unaware of their position within the episode,
→ RL agents must continue to bootstrap at the last timestep (no "done" signal).

See *Time Limits in Reinforcement Learning* - Pardo et al., 2018 for a discussion.

EpidemiOptim Toolbox Organization

Interface

```python
from epidemioptim.environments.models import get_model
from epidemioptim.environments.cost_functions import get_cost_function
from epidemioptim.environments.gym_envs import get_env
from epidemioptim.optimization import get_algorithm
from epidemioptim.configs.get_params import get_params

config =  'dqn'

# Get the configuration
params = get_params(config_id=config)

# Get the epidemiological model
model = get_model(model_id=params['model_id'], params=params['model_params'])

# Get cost function
cost_function = get_cost_function(cost_function_id=params['cost_id'], params=params['cost_params'])

# Create the optimization problem as a Gym-like environment
env = get_env(env_id=params['env_id'], cost_function=cost_function, model=model, sim_horizon=params['sim_horizon'], seed=params['seed'])

# Get DQN algorithm parameterized by beta
algorithm = get_algorithm(algo_id=params['algo_id'], env=env, params=params)

# Run the training loop
algorithm.learn(num_train_steps=params['num_train_steps'])
```

Utils

Comparison tools:

We want to easily compare approaches, both visually and statistically.

Visualization tools:

We want to explore the results visually, interact with policies and models.

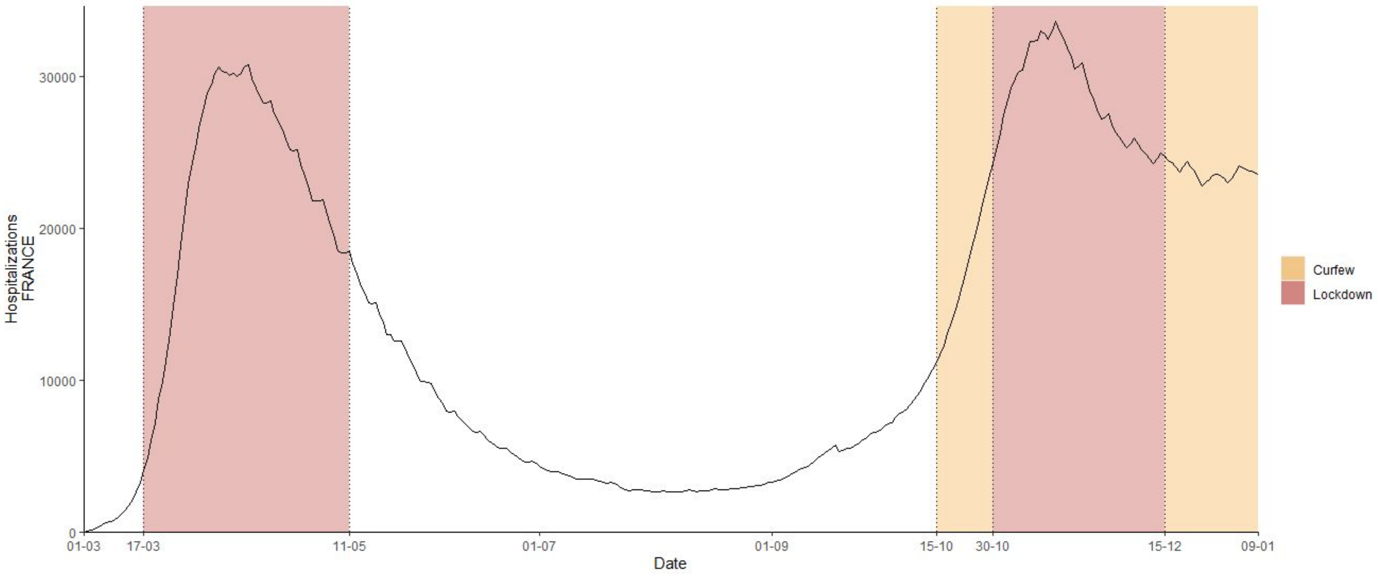
We want accessibility for non-expert users (e.g. general public, decision makers, etc.)

Case Study

-

Lock-down policies for COVID-19 Epidemic

COVID-19 Case Study

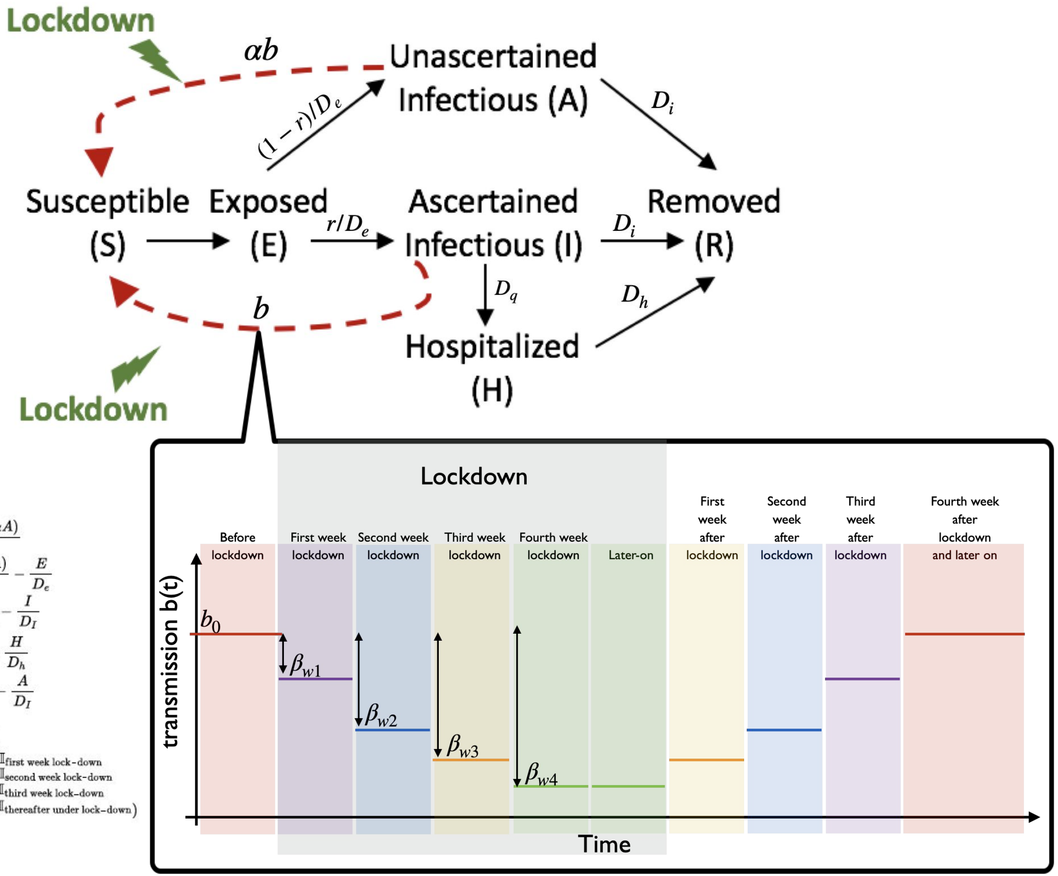


Multiple lockdown on-off lockdown is a worldwide adopted strategy:

- **Reduction in the number of Hospitalization** (and therefore incidence and additional deaths)
- **High economical costs**

→ In France: 2 months lock-down in march led to 35% of activity reduction (INSEE) and 32% of GDP reduction - 120€billion gap (OFCE).

Epidemiological Model



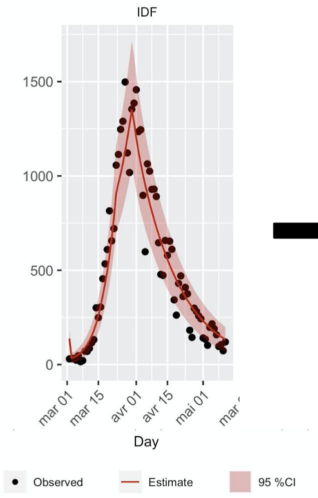
$$
\begin{aligned}
\frac{dS}{dt} &= -\frac{bS(I+\alpha A)}{N}\\
\frac{dE}{dt} &= \frac{bS(I+\alpha A)}{N} - \frac{E}{D_e}\\
\frac{dI}{dt} &= \frac{rE}{D_e} - \frac{I}{D_q} - \frac{I}{D_I}\\
\frac{dR}{dt} &= \frac{(I+A)}{D_I} + \frac{H}{D_h}\\
\frac{dA}{dt} &= \frac{(1-r)E}{D_e} - \frac{A}{D_I}\\
\frac{dH}{dt} &= \frac{I}{D_q} - \frac{H}{D_h}\\
b &= b_0 \exp(\beta_{w1}\,\mathbb{I}_{\text{first week lock-down}}\\
&\quad +\beta_{w2}\,\mathbb{I}_{\text{second week lock-down}}\\
&\quad +\beta_{w3}\,\mathbb{I}_{\text{third week lock-down}}\\
&\quad +\beta_{w4}\,\mathbb{I}_{\text{thereafter under lock-down}})
\end{aligned}
$$

Parameters are **estimated from data (point estimate and uncertainty)** with a population approach using *Monolix* software

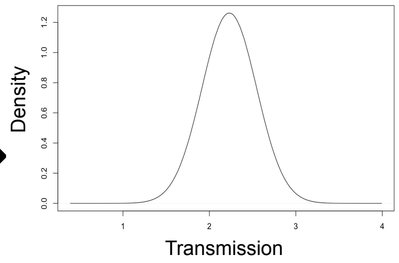
[Prague et al. 2020, MedRXiv]

Distribution of Models

Uncertainty in
data fitting

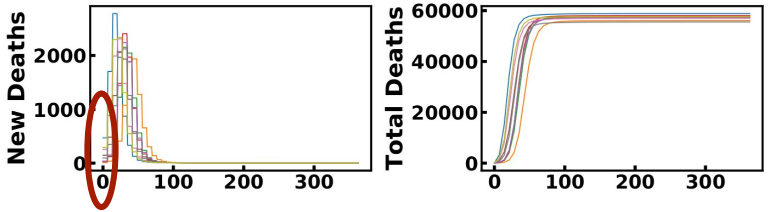


Uncertainty in
parameters values



Sampling from
parameters joint
distribution

Uncertainty in epidemics
Simulations for RL

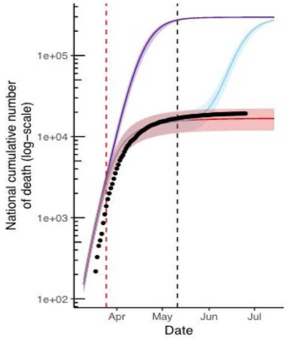


Epidemics is not directly detected.
We assume a delay uniformly distributed
[0;14] days before control policies start.

Cost Functions

Health cost function :

- Minimizing the number of deaths



Number of
recovered from COVID-19

$$C_{\text{health}}(t) = 0.005R(t)$$

Economic cost function (Havik et al.,2014):

- Loss in Gross Domestic Product (GDP)
- Depend on the population unable to work : $G(t) = I(t) + H(t) + 0.005R(t)$

Exogeneous
Technical progress

Employment rate

Level of partial
Employment

Capital
Elasticity

$$C_{\text{economic}}(t) = Y_0 - AK_0^{\gamma_k}(1 - u(t))\lambda(N - G(t))^{1-\gamma_k}$$

Initial
GDP

Initial
Capital stock

Population
size

Deep Q-Networks (DQN) (Mnih et al., 2013)

Alternates between:

→ data collection: interact with epidemics and collect transitions (s, a, s', costs)

→ data exploitation: train a Q network: Q(s, a).

$$C_{\text{aggregated}} = (1-\beta) \times C_{\text{health}} + \beta \times C_{\text{economic}}$$

$$\beta = 0, \quad C_{\text{aggregated}} = C_{\text{health}}$$
$$\beta = 1, \quad C_{\text{aggregated}} = C_{\text{economic}}$$

Playing Atari with Deep Reinforcement Learning - Mnih et al., 2013

DQN variants

Goal DQN (Schaul et al., 2015; Badia et al., 2020)

β is part of the inputs to the Q network. The agent is now "goal-conditioned", where the goal is a particular aggregated objective parameterized by β.

Goal DQN with constraints (Goal DQN-C) (modified from Badia et al., 2020)

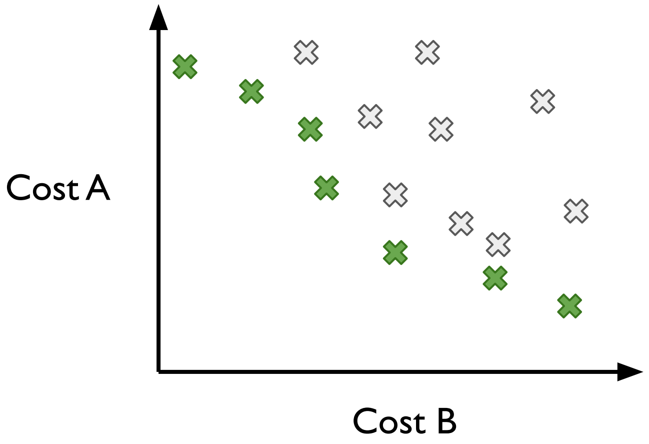
Goals are now parameterized by constraints expressed as maximal number of deaths M_{health} and maximal economic cost M_{economic}.

$$M_{\text{health}} : \text{if } C^{1:t}_{\text{health}} > M_{\text{health}}, \ \text{then } C^{t}_{\text{aggregated}} = 1000$$
$$M_{\text{economic}} : \text{if } C^{1:t}_{\text{economic}} > M_{\text{economic}}, \ \text{then } C^{t}_{\text{aggregated}} = 1000$$

Universal value function approximators - Schaul et al., 2015
Never Give Up - Badia et al., 2020

NSGA-II (Deb et al., 2002)

NSGA-II is a state-of-the-art multi-objective algorithm based on a genetic algorithms.
It produces a Pareto front of policies instead of a single policy.

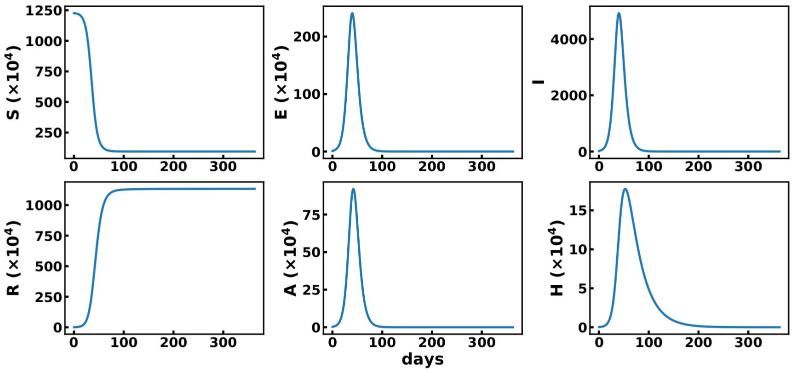


Pareto front: set of undominated solutions.

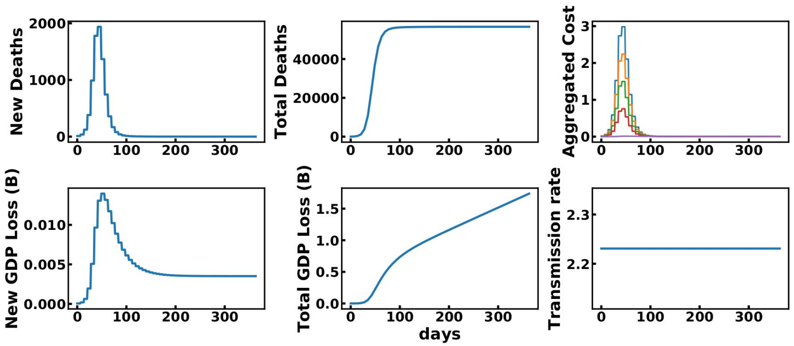
A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II
Deb et al., 2002

Results - DQN

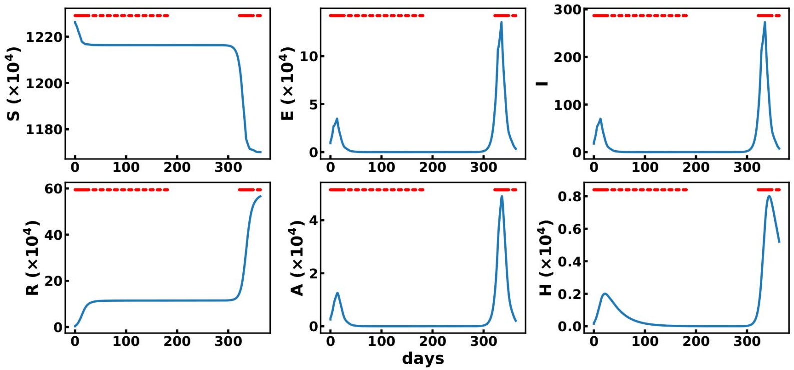
Beta = 0.8



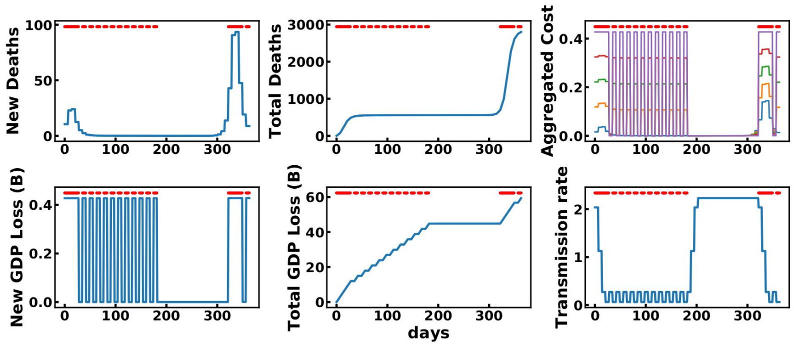
Eco cost: 1.74 B, Death Cost: 56621, Aggregated Cost: 31.62



Beta = 0.55

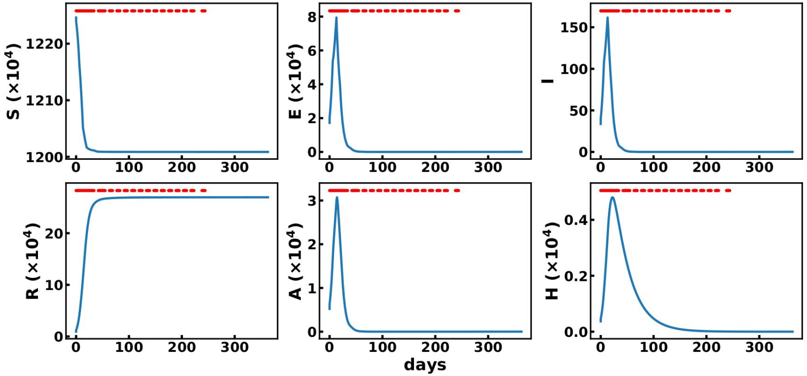


Eco cost: 59.41 B, Death Cost: 2804, Aggregated Cost: 34.86

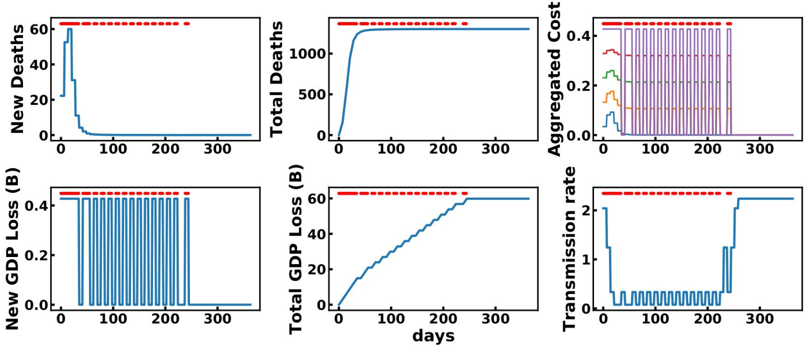


Results - Goal DQN

Beta = 0.65

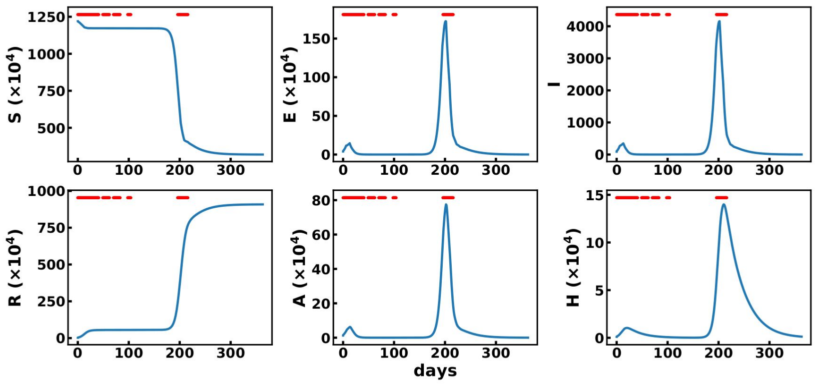


Eco cost: 59.84 B, Death Cost: 1301, Aggregated Cost: 28.03

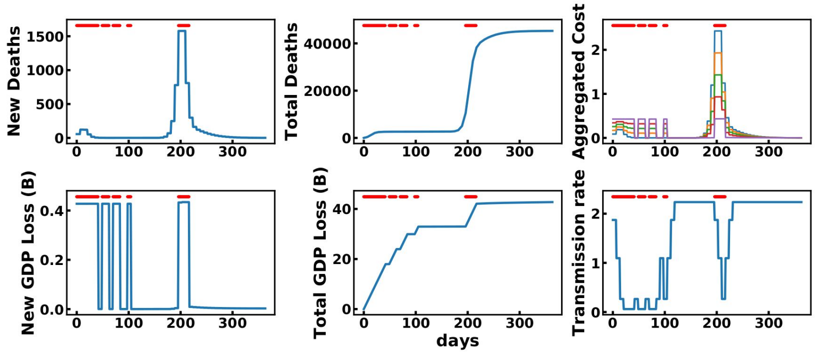


Results - Goal DQN with Constraints

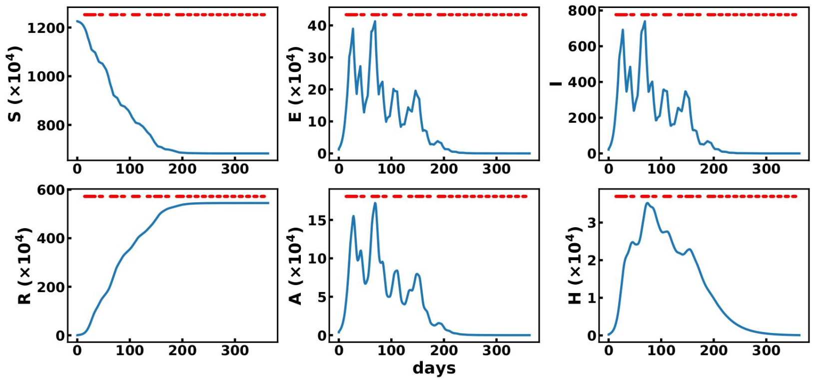
Beta = 0.3 Max economic cost: 55 B



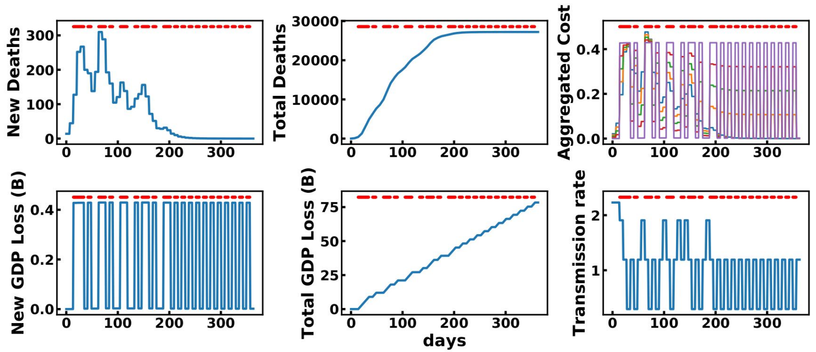
Eco cost: 42.74 B, Death Cost: 45278, Aggregated Cost: 61.58



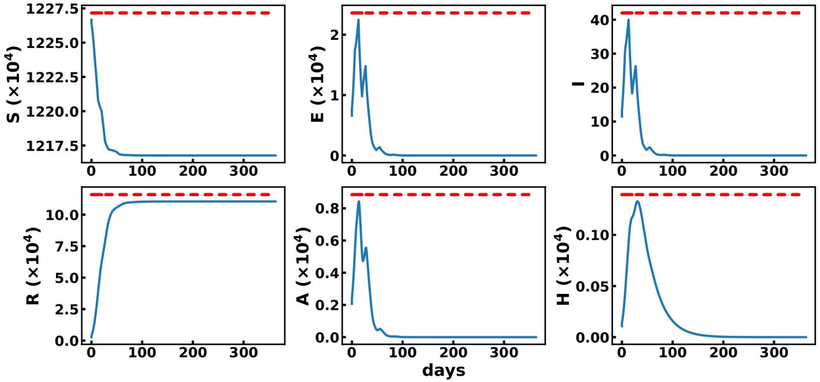
Beta = 0.7 Max # deaths = 30000



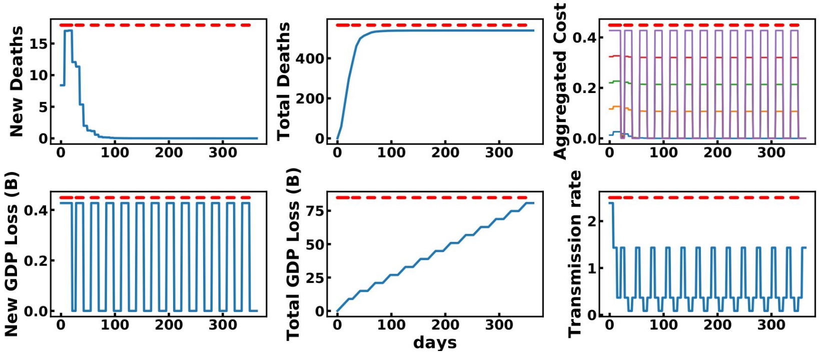
Eco cost: 78.37 B, Death Cost: 27214, Aggregated Cost: 67.42

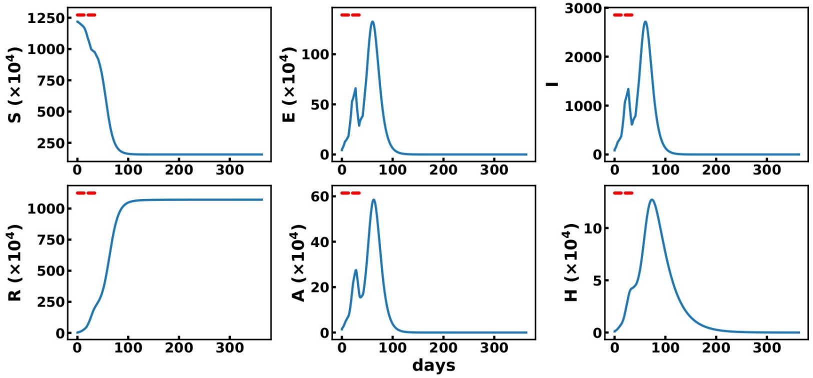


Results - NSGA-II

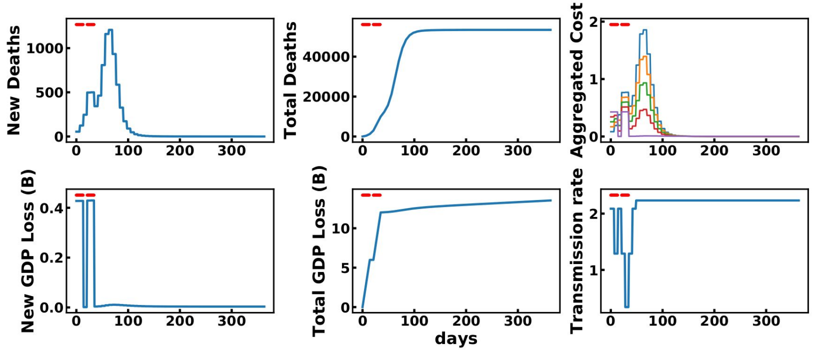


Eco cost: 80.76 B, Death Cost: 538, Aggregated Cost: 40.79

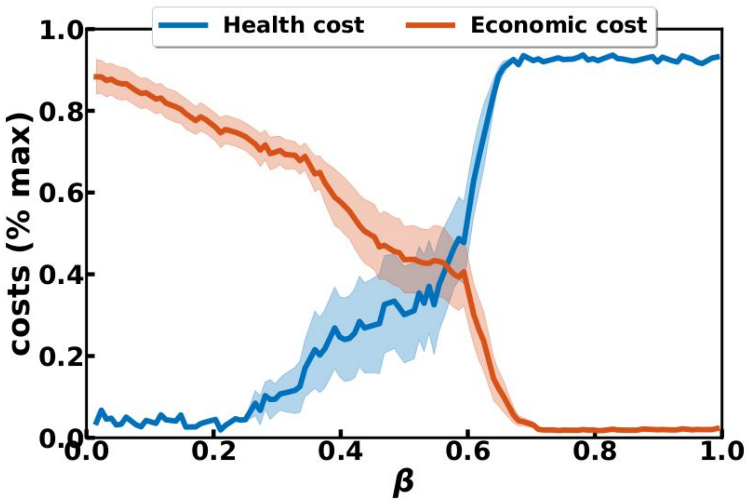




Eco cost: 13.53 B, Death Cost: 53414, Aggregated Cost: 47.85



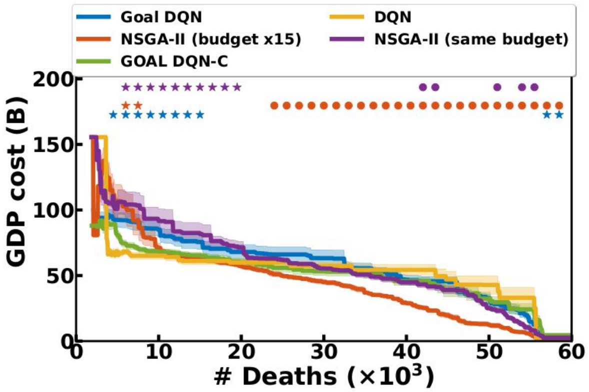
Results - Goal DQN: influence of beta



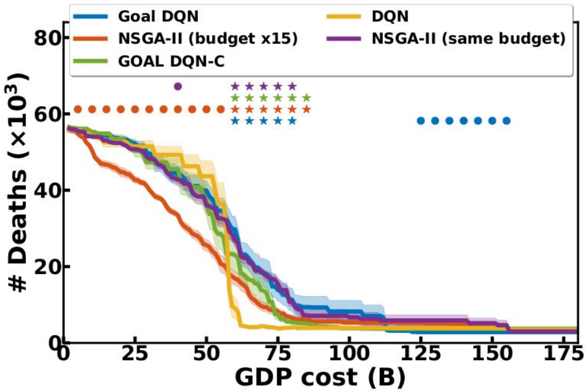
→ Low β
→ Health cost is more important
→ Goal DQN minimizes it (blue)

→ High β
→ Economic cost is more important
→ Goal DQN minimizes it (orange)

Results - Algorithms Comparison



(a)



(b)

1D comparisons

● Significantly better than DQN

★ Significantly worse than DQN

2D comparisons

p-values (uncorrected)

	DQN	GOAL DQN	GOAL DQN-C	NSGA-II (same)	NSGA-II (x15)
DQN	N/A	0.069	0.55	0.018	0.045
GOAL DQN		N/A	0.046	0.84	0.0057
GOAL DQN-C			N/A	0.21	0.35
NSGA-II (same)				N/A	3.2×10^4
NSGA-II (x15)					N/A

Interactive Visualization

https://epidemioptim.bordeaux.inria.fr/

Case Study Conclusion

Algorithm comparison

→ DQN performs better in low death toll regime
→ NSGA-II performs better in low economic cost regime

Mainly two types of strategies

→ Early lockdown or none at all: relies on herd immunity (e.g. Sweden)
→ Short-term lockdowns to control epidemic waves (most European countries)

Lockdowns implemented by countries are longer than ours (probably political and practical reasons).

Many approximations and simplifications

→ This case-study demonstrate the importance of EpidemiOptim.
We do not make any real-world recommendation!

Discussion

Discussion

Automatic optimization for decision making

→ Optimization algorithms should not replace decision makers.
→ Explicit models are better than mental models (exposed assumptions, explainability, easier to discuss).
→ Optimization can help integrate long term effect and explore spaces of intervention strategies.
→ Diversifying models and optimization algorithms reduces model- and algorithm-induced biases.

Collaborative toolbox

→ To be extended to more epidemiological models, cost functions, optimization algorithms, visualization tools, etc.

General approach

→ The same approach can be used to study optimization in any dynamical models (e.g. vaccination processes, economic models, ODE systems etc.)

Resources

Paper: https://arxiv.org/abs/2010.04452

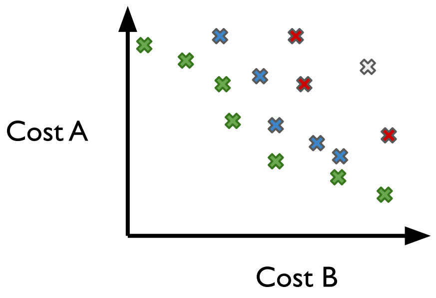
Code: https://github.com/flowersteam/EpidemiOptim

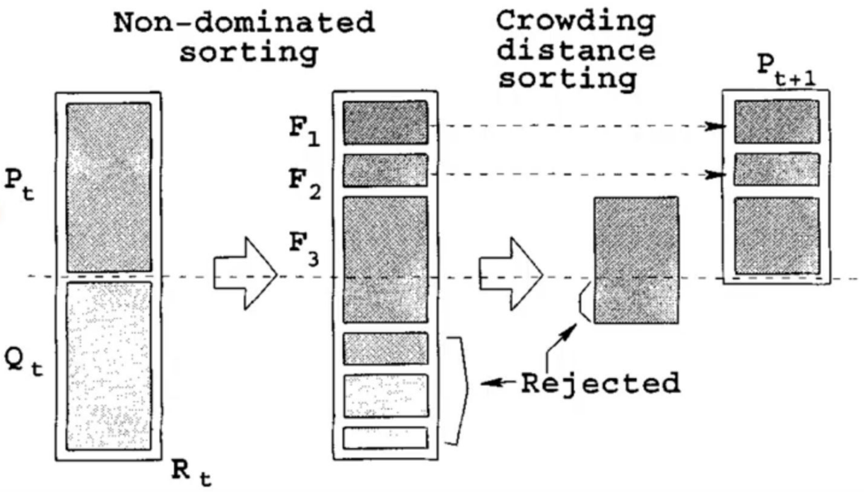
Interactive demo (coming soon): https://epidemioptim.bordeaux.inria.fr/

Contact: cedric.colas@inria.fr - melanie.prague@inria.fr

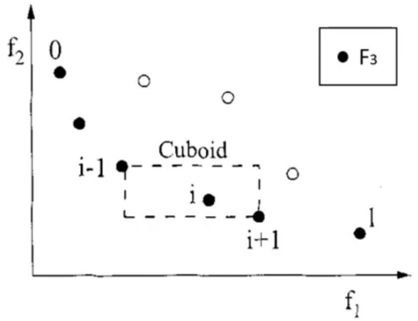
NSGA-II

1. Tournament Selection
2. Cross-over
3. Mutations





1. Create offspring and a combined population R_t
2. Rank and sort offspring due to performance on defined target indicators
3. Take best members to create new population including a good spread in solutions



A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II
Deb et al., 2002