Data-driven Optimization Model for Global Covid-19 Intervention Plans

Georgian

Chang Liu[1], Akshay Budhkar[1]

[1]Georgian

Problem

We would like to understand the impact of different intervention plans (IP) to combat Covid-19 and prescribe the most beneficial policies. We present an integer program, with parameters computed from historical data, to find the optimal policies.

This work results from participating in the XPRIZE Pandemic Response Challenge[2].

Data

We use the Oxford dataset[3], which contains case numbers and historical IPs took across 280 regions worldwide since the rise of the pandemic.

Standard Predictor

We make use of the predictor[4] made available by XPRIZE to help us estimate parameters and make future prescriptions. The authors use a 21-day lookback window to capture the number of cases and the IPs separately. The model enforces monotonicity on the IPs as increasing their stringency should lead to fewer infection cases. The authors then train a model comprising two separate recurrent LSTM layers as two distinct pathways that eventually lead to a single prediction.

[2] https://www.xprize.org/challenge/pandemicresponse
[3] Thomas Hale, Anna Petherick, Toby Phillips, and Samuel Webster. **Variation in government responses to covid-19**. *Blavatnik school of government working paper*, 31:2020–11, 2020.
[4] Risto Miikkulainen, Olivier Francon, Elliot Meyerson, Xin Qiu, Elisa Canzani, and Babak Hodjat. **From prediction to prescription: Evolutionary optimization of non-pharmaceutical interventions in the covid-19 pandemic**. *arxiv:2005.13766*, May 2020.

Stringency Costs

Let $S_{p,l}$ represent the stringency cost associated with IP p at level l. In this work, we experimented with three sets of stringency costs: fixed, random, and realistic.

The fixed costs assume that the stringency cost for each IP is 1. We generate the random costs at random. The realistic costs are chosen based on the relative costs of implementing these IPs in Canada

Intervention Plan	Restriction Levels	Realistic Stringency Cost
School closing	(0,1,2,3)	9
Workplace closing	(0,1,2,3)	6
Cancel public events	(0,1,2,)	2
Restrictions on gatherings	(0,1,2,3,4,)	5
Close public transport	(0,,2)	8
Stay at home requirements	(0,1,2,3)	7
Restrictions on internal movement	(0,1,2)	7
International travel controls	(0,1,2,3,4)	8
Public information campaigns	(0,1,2)	2
Testing policy	(0,1,2,3)	3
Contact Tracing	(0,1,2)	7
Facial coverings	(0,1,2,3,4)	2

Impact of IPs on the Number of Cases

Let $C_{p,l}$ represent the impact IP p at level l has on the number of infection cases. For each region, first, we use the standard predictor to compute the baseline infection cases (B) by setting all IPs to level zero. Next, for each IP and each restriction level, we use the standard predictor to compute the estimated number of infection cases ($E_{p,l}$) by only activating that specific IP at that restriction level.

$$C_{p,l} = \frac{(E_{p,l} - B)}{B} * 100$$

Integer Program

Variables Let $x_{p,l} = \begin{cases} 1 & \text{if IP } p \text{ at level } l \text{ is prescribed} \\ 0 & \text{otherwise} \end{cases}$

Objective function Normalization We first normalize the stringency costs at level one for the 12 IPs to sum to 1. We then normalize the number of infections to the current case numbers. For each region, let β be the number of current new cases, α be the number of new cases in the previous day, which we compute using the standard predictor. As we prescribe for multiple days, β remains the same α while will vary.

$$\begin{aligned} \text{minimize} \quad & \frac{1}{\beta}\left(\alpha + \sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}_p} C_{p,l} * x_{p,l} * \alpha\right) + \sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}_p} S_{p,l} * x_{p,l} \\ \text{s.t.} \quad & \sum_{l\in\mathcal{L}_p} x_{p,l} = 1 && \forall p \in \mathcal{P} \\ & x_{p,l} \in \{0,1\} && \forall p \in \mathcal{P}, l \in \mathcal{L}_p \end{aligned}$$

The objective function minimizes the total number of new cases and stringency costs. The constraints make sure that for each IP, we select only one restriction level at a time.

Consecutive Constraint In the real world, it is natural that when an IP is put in place, it will last for at least several days. By analyzing historical IPs across all the regions, we created a minimum number of consecutive days that each intervention plan at a specific level must hold, which we set a ceiling of 7 days. At the time of prescribing for a day, based on the previous seven days of prescribed IPs, we create a new parameter $D_{p,l}$ = 1 if IP p at level l has not been prescribed to enough consecutive days, thus forcing the model to set $x_{p,l}$ = 1, and 0 otherwise.

$$D_{p,l} \leq x_{p,l} \qquad \forall p \in \mathcal{P}, l \in \mathcal{L}_p$$

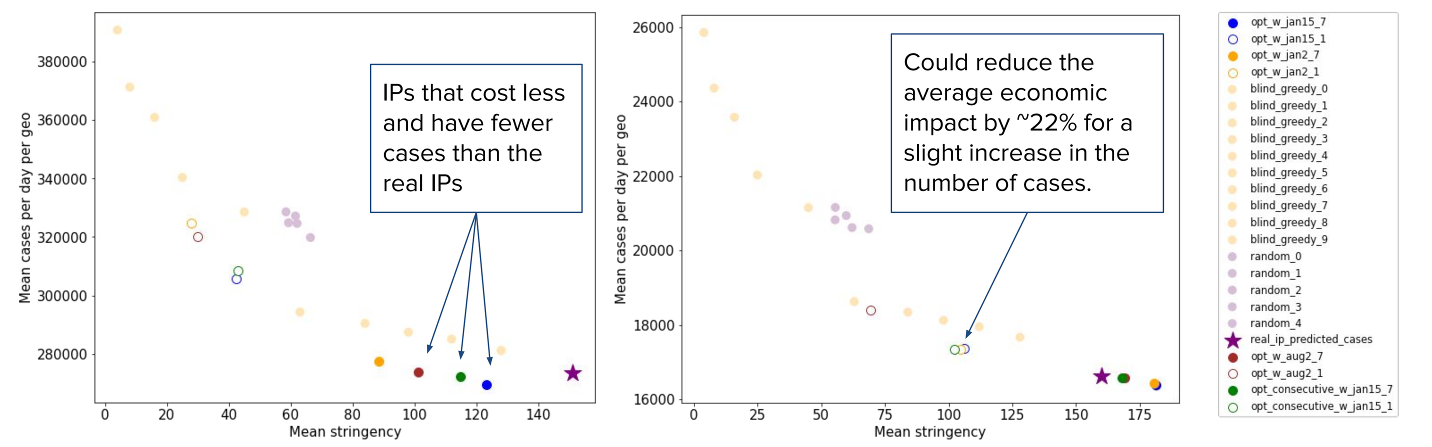
Baseline Models

Blind-greedy heuristic This method does not take into consideration past IPs. It starts by setting all IPs to level zero, then iteratively, increases a restriction level of the next least costly IP.

Random Heuristic For each region and each prescription day, this heuristic assigns a valid level for each IP at random.

Experimental Results

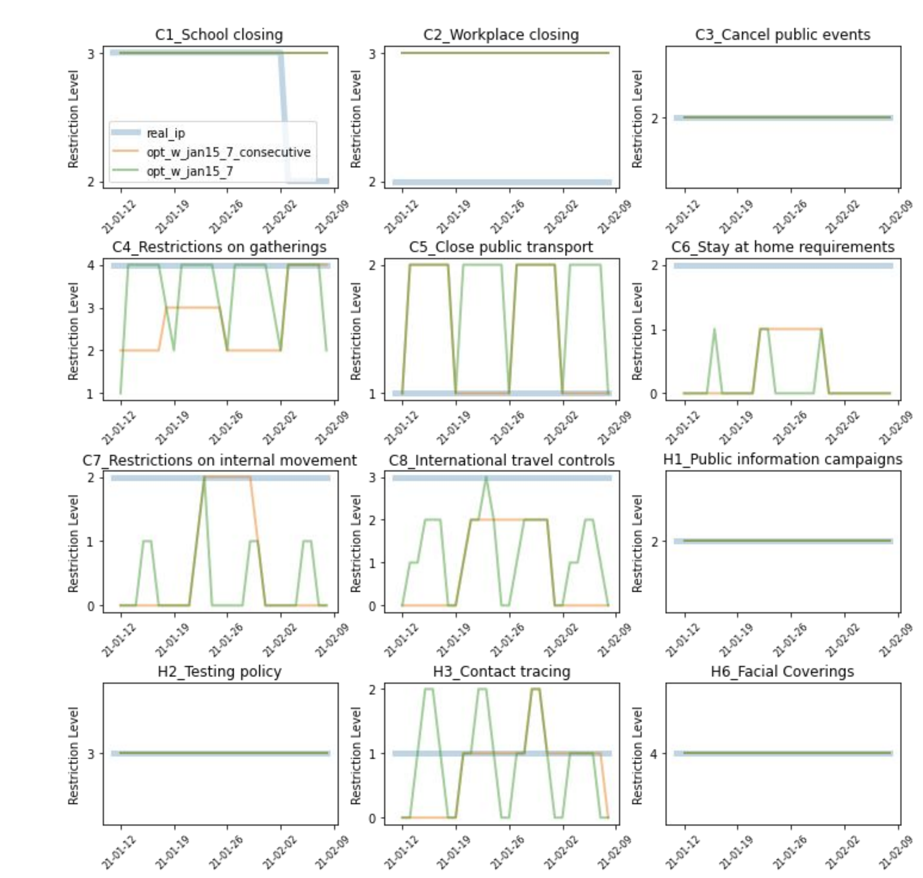
We make prescriptions for 28 days from Jan 12, 2021. We then measure the mean stringency and the mean number of new cases predicted by the standard predictor.



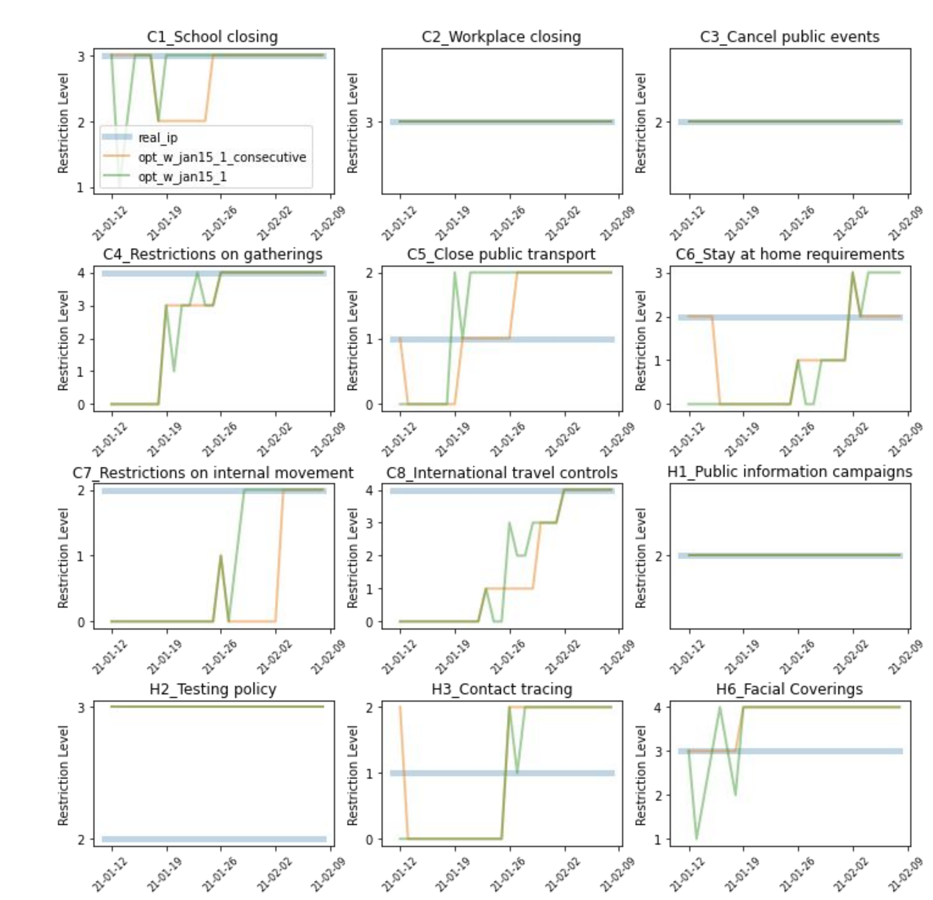
Intervention Policies

Real IPs, our model with and without the consecutive constraint

United States



Canada



Conclusions and Future Work

- We proposed a simple, data-driven integer program approach to find optimal intervention plan prescriptions for any region globally
- We have made two interactive dashboards available to the public to review these different IPs and their impact on regions of choice
- We can improve the normalization of the two costs in the integer program objective function by implementing algorithms from the multi-objective optimization field
- We can also improve on the estimated $C_{p,l}$ by performing a linear regression on the historical IPs
- The performance of the prescription model is heavily dependent on the performance of the infection cases predictor, and we aim to improve it with better modeling techniques and more available data

← Interactive dashboards allowing users to study results from all different models.