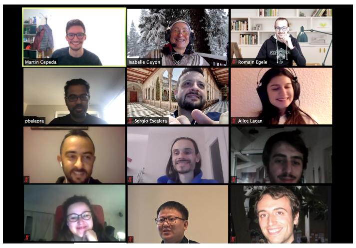


TRANSATLANTIC TEAM FINAL QUALITATIVE SUBMISSION

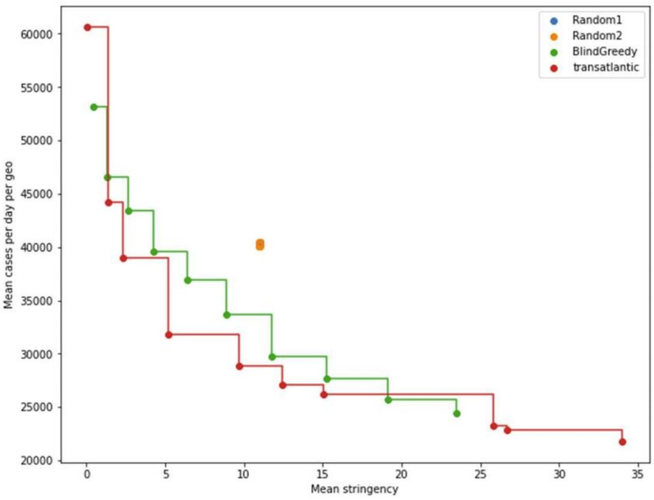
Github repo: https://github.com/transatlantic-team/Prescriptor-Phase

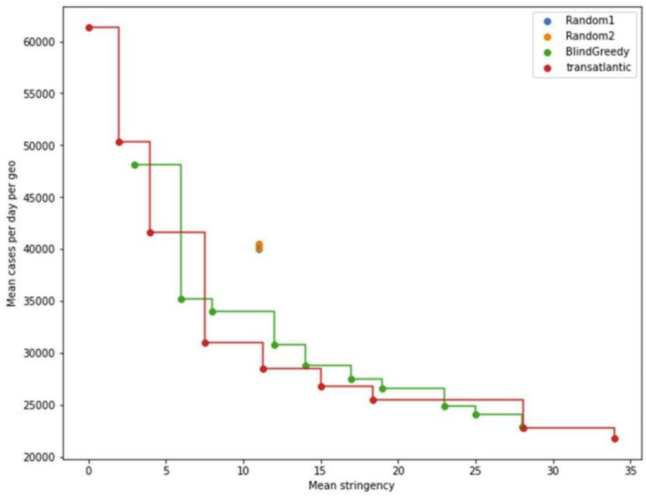
I.Introduction

We are a team of 12 people from 3 continents: Europe, America, and Asia, with our leader (Martin Cepeda) currently based in Paris, France. Our team members include 5 undergraduate students, 1 post-doc, 3 senior researchers, and 3 university professors. Our prescriptor is the product of a collaborative effort and extensive tests, which are detailed in the following sections. Our challenge submission is a fast and general prescriptor, returning up to ten optimal strategies for each country, with the purpose of finding the best NPIs according to given predicted Covid19 cases and Intervention Plan (IP) economical costs. We also investigated (but did not submit) a more advanced method based on Markov Decision Processes (MDP) (see Appendix E). To that end, we developed a reinforcement learning environment following the OpenAI Gym interface, which we are making publicly available.

II.Actionability and usability

We fully address the **multi-objective goal of the challenge, in a flexible way**: The submitted prescriptor is designed to produce a range of predictions, for various time-varying priorities, between the **expected number of new cases** and **the cost of the prescripted IP**. Such priorities can be arbitrarily defined over a fixed timespan, generating up to 10 different strategies per country, modulated by IP costs. The prescriptor, tailored to the underlying predictor provided by XPRIZE, focuses on the optimal NPIs (most reduction of new cases per stringency level), that were inferred from a preliminary analysis of the predictor (see appendix D). The final solution is computationally efficient (under 1 hour for all regions over 90 days) and out-performs the blind-greedy approach. The plots below show the Pareto curve over 90 days (01/08/2020 to 31/10/2020) with uniform random weights (left) and fixed equal weights (right) for Random, BlindGreedy and transatlantic prescriptors:



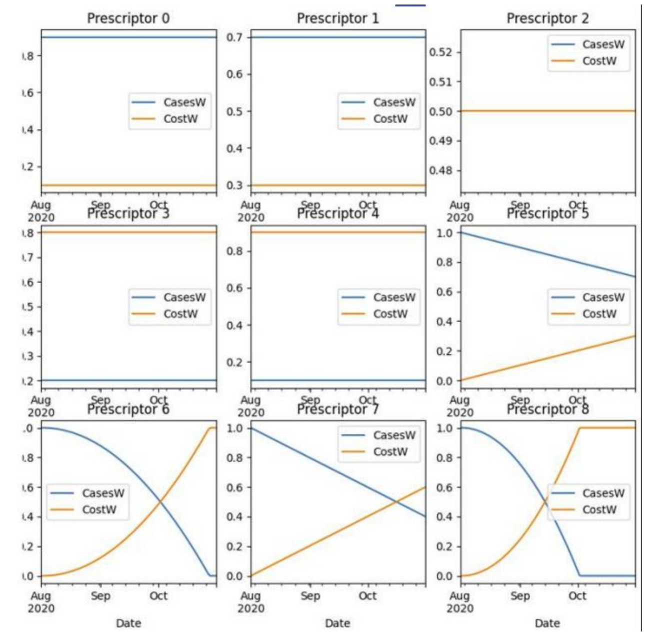


III.Explanation

We designed a **multi-expert prescriptor** (Appendix C) to find a tradeoff between the number of cases and the economical cost. Each expert is designed to find optimal solutions in different economical cost ranges. We provide at least one expert prescriptor for each 5 stringency values: Prescriptor indexed 0 to 3 aim to obtain good results at a low cost (s in [0,5]), those indexed 7 to 9 aim to minimize the number of new cases (s in [24,34]) and those indexed 4 to 6 aim to find balanced prescriptions with a trade off between number of cases and economic cost (s in [5, 20]).

In order to do so, the different IPs are computed as follows (experiment numbers refer to appendix D):

1. Select the top 10 prescriptions from experiments 2 and 4 (precomputed) and compute a global normalized cost in [0,1] for new cases per prescription. Best and worse correspond to the all-max-values and all-zeros prescriptions.
2. Weight prescriptions using the provided costs (computed once). Compute a per region normalized cost in stringency per prescription in [0,1]. 1 to the all-max-values prescriptions, minimum will be always 0.
3. Rank all prescriptions according to a linear combination of previous normalized values for cost in new cases and stringency (computed per each date). Changing the weight we give to each of the two normalized values we define the IP.



On the left figure we observe different weights assigned to a strategy over time (corresponding to step 3). These weights a) can be interpreted as the priority on reducing the number of new cases or the cost of the IP and b) are all based on a common goal: when starting the prescriptions, the priority on reducing the number of new cases is always higher than the priority on reducing the cost of interventions. This is due to the fact that stopping the growth of cases in a pandemic at the beginning allows to delay both the peak of infections and the fastest growth stage of the number of cases (Ferguson, 2020). Please refer to Appendix C for the full specification.

IV. Addressing the challenge

We performed specific analyses and design steps to address all aspects of the challenge: 1) Predictor **sensitivity analysis** 2) Predictor **run time consistency analysis** 3) Predictor **atemporality analysis** 4) **Cost-conservative strategy analysis** and 5) **Development of the prescription strategy**. Details are found in Appendix D. The first 4 steps showed that:

- The number of new cases have a lower bound on the “all values-max” prescription and an upper bound with “all-zero-values” prescription.
- All prescriptions have the same behaviour, after day 2 the rank between prescriptions does not change. Then, it is not necessary to predict the whole month to assess the relative performance of an IP.
- We have good results using only 6 from 12 NPIs (max stringency reduction from 34 to 19).
- Good prescriptions have a large value for C1 and C2.
- For a cost-averse IP (total stringency <=5), we can halve the number of cases w.r.t total inaction (stringency 0 for all NPIs).

These findings allowed us to have a ranked pool of IPs to build the final prescriptor (step 5, see III).

V. Inclusivity and fairness

The prescriptor implements an approach based on the priorities (reduce number of new cases/cost of the intervention plan), without considering any population segmentation, and is in the sense neutral. In Appendix E, we considered a more elaborate MDP framework for policy discovery, but due to space/time complexity of training such a method, we decided to keep the simpler approach.

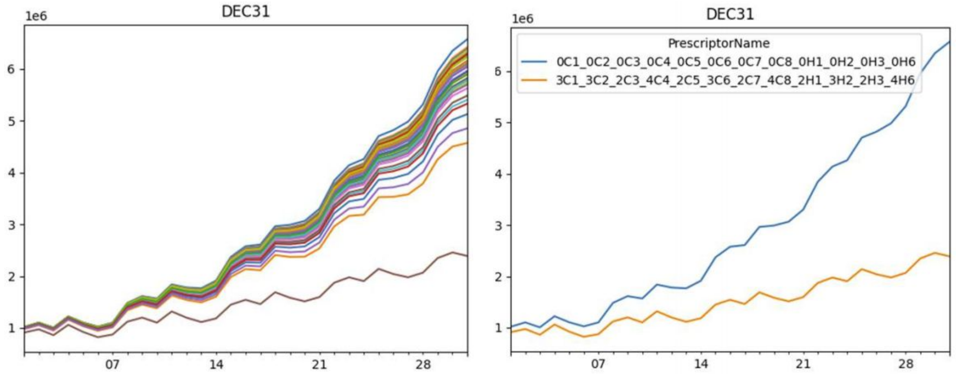
VI. Generality

The presented methodology (sections III, IV) can be replicated with **any other predictor respecting the challenge API**. The transatlantic prescriptor provides IPs based on the country cost weights, using pre-defined priorities over time. In this sense, the prescriptor **does not have any speciality**

regions and can be used in any country where the common predictor has been trained. The choice of making a general prescriptor emanates from a) the generality of the underlying predictor and b) the impossibility to have daily-updated data besides the OxCGRT dataset, which means that including extra features not considered by the common prescriptor can be counter-productive, both constraints present in the challenge according to the organizers.

VII.Consistency

As it is indicated in the Predictor-Prescriptor framework that inspires this challenge (Miikkulainen et al, 2020), the IP prescriptor relies heavily on the quality of the underlying predictor to generate counterfactuals. This is a key aspect both on the significance of the simulated scenarios and the search space of NPIs. In particular, our prescriptor a) ignores not significant NPIs when computing an IP and b) assesses the partial effectiveness of an IP based on only one call to the predictor, over three days. This is due to the following predictor behaviour:



Over a test run of the common predictor over December 2020, we observe that a) all sampled scenarios (IP) give an evolution on the number of cases between the 2 extreme strategies: all NPIs to maximum value and all NPIs to 0, b) each of the runs has the same shape and the ranking of PIs is consistent after the 3rd day of simulation (e.g. if an IP is the i-th most effective one on day 3, it will be the i-th most effective one for any subsequent day) and c) even with maximum stringency over 1 month or 3 months, the number of new cases always increases.

The previous analysis implies that the transatlantic prescriptor relies mainly on the strategy priorities introduced in section III and the country cost weights, the prescriptions being consistent given the common predictor and depending exclusively on input priorities and costs file.

VIII.Transparency and trust

As the different strategies in the prescriptor are associated with a choice of priorities over time (new cases and stringency, see section III), **it is straightforward for a layperson to define custom priority functions over time** and obtain an output IP that will depend on the cost given to the NPIs. This means that the prescriptor is transparent with respect to the given priorities: the higher the priority the user wants to give to new cases/PI cost, the lower the new cases/PI costs will be. As for data usage, our prescriptor only relies on the OxCGRT dataset and the data used by the common predictor, data that is open-sourced and continuously updated for trustworthiness.

IX.Collaborative contributions

Our code and country-specific time independent data are open-sourced. We also provide a Markov Decision Process implementation of prescribing IPs based on (Cepeda, 2020) (see Appendix E).

X.Innovation

We believe our multi-expert transatlantic prescriptor addressing the multi-objective goal of the challenge is novel, flexible, straightforward to use, and offers explainability. Its main features of interest include: (1) using as parameter the priority trade-off over a prescription period; and (2) being combinable with any other predictor. This allows users to a) characterize and analyze the predictor, and b) based on the simulated world behaviour, find the optimal policies constrained by the decision maker's priorities.

APPENDICES

A. Acknowledgements

This work builds upon internships on Covid-19 performed over the summer 2020, sponsored by ChaLearn and INRIA. We are grateful to Alian-Jacques Valleron, Sam Evans, Kristin Bennett, Paola Tubaro and John Erickson for help and advice.

B. References

(Yu et al, 2020) Yang Yu, Yu-Ren Liu, Fan-Ming Luo, Wei-Wei Tu, De-Chuan Zhan, Guo Yu, Zhi-Hua Zhou, COVID-19 Asymptomatic Infection Estimation. medRxiv, 2020 https://www.medrxiv.org/content/10.1101/2020.04.19.20068072v1

(Miikkulainen et al, 2020) Risto Miikkulainen, Olivier Francon, Elliot Meyerson, Xin Qiu, Elisa Canzani, Babak Hodjat, From Prediction to Prescription: Evolutionary Optimization of Non-Pharmaceutical Interventions in the COVID-19 Pandemic. arXiv e-print, 2020. https://arxiv.org/abs/2005.13766

(Cepeda, 2020) Covid-19 risk mitigation. Master thesis. https://github.com/cepedus/COVID19-Risk-Mitigation. Submitted to JDSE2021.

(Pearl, 2009) Judea Pearl. Causality. Cambridge University Press, 2009.

(Ferguson, 2020) Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani. Report 9 - Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. MRC Centre for Global Infectious Disease Analysis, Imperial College, 2020. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/

(Sutton 2018) Sutton, R.S. and Barto, A.G. Reinforcement Learning. An Introduction. 2nd Edition, A Bradford Book, Cambridge, 2018. incompleteideas.net/book/RLbook2020.pdf

(Petersen, 2020) Eskild Petersen, Marion Koopmans, Unyeong Go, Davidson H Hamer, Nicola Petrosillo, Francesco Castelli, Merete Storgaard, Sulien Al Khalili, Lone Simonsen, Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics, The Lancet Infectious Diseases, 2020.

(Goodman-Bacon, 2020) Andrew Goodman-Bacon and Jan Marcus. Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies. Techreport, 2020. https://papers.ssrn.com/abstract=3603970.

(Alkire, 2018) Blake C. Alkire, Alexander W. Peters, Mark G. Shrime, and John G. Meara. The Economic Consequences Of Mortality Amenable To High-Quality Health Care In Low- And Middle-Income Countries. Health Affairs, 2018.

(Javaid, 2015) Kiran Javaid et al. Morbidity and Economic Growth. Tech report Evans School Policy Analysis and Research Group, 2015. https://evans.uw.edu/policy-impact/epar/research/morbidity-and-economic-growth.

(Brodeur, 2020) Abel Brodeur and Andrew E. Clark and Sarah Fleche and Nattavudh Powdthavee. Assessing the impact of the coronavirus lockdown on unhappiness, loneliness, and boredom using Google Trends. ArXiv preprint, 2020. https://arxiv.org/abs/2004.12129

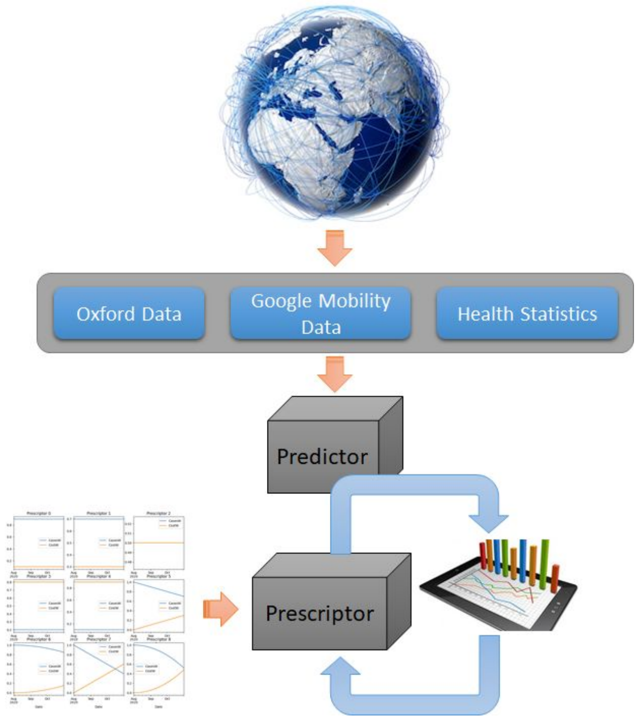
(Aron, 2020) Janine Aron and John Muellbauer. A pandemic primer on excess mortality statistics and their comparability across countries. Our World in Data, 2020. https://ourworldindata.org/covid-excess-mortality

(Woolf, 2020) Steven H. Woolf, Derek A. Chapman, Roy T. Sabo, Daniel M. Weinberger, and Latoya Hill. Excess Deaths From COVID-19 and Other Causes, JAMA, 2020. https://doi.org/10.1001/jama.2020.11787.

(Bai, 2018) Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. ArXiv preprint, 2018. https://arxiv.org/abs/1803.01271

(Roser, 2020) Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina and Joe Hasell. Coronavirus Pandemic (COVID-19). https://ourworldindata.org/coronavirus. Our World in Data, 2020.

C. Transatlantic prescriptor schema



Our approach can be defined using the left image:

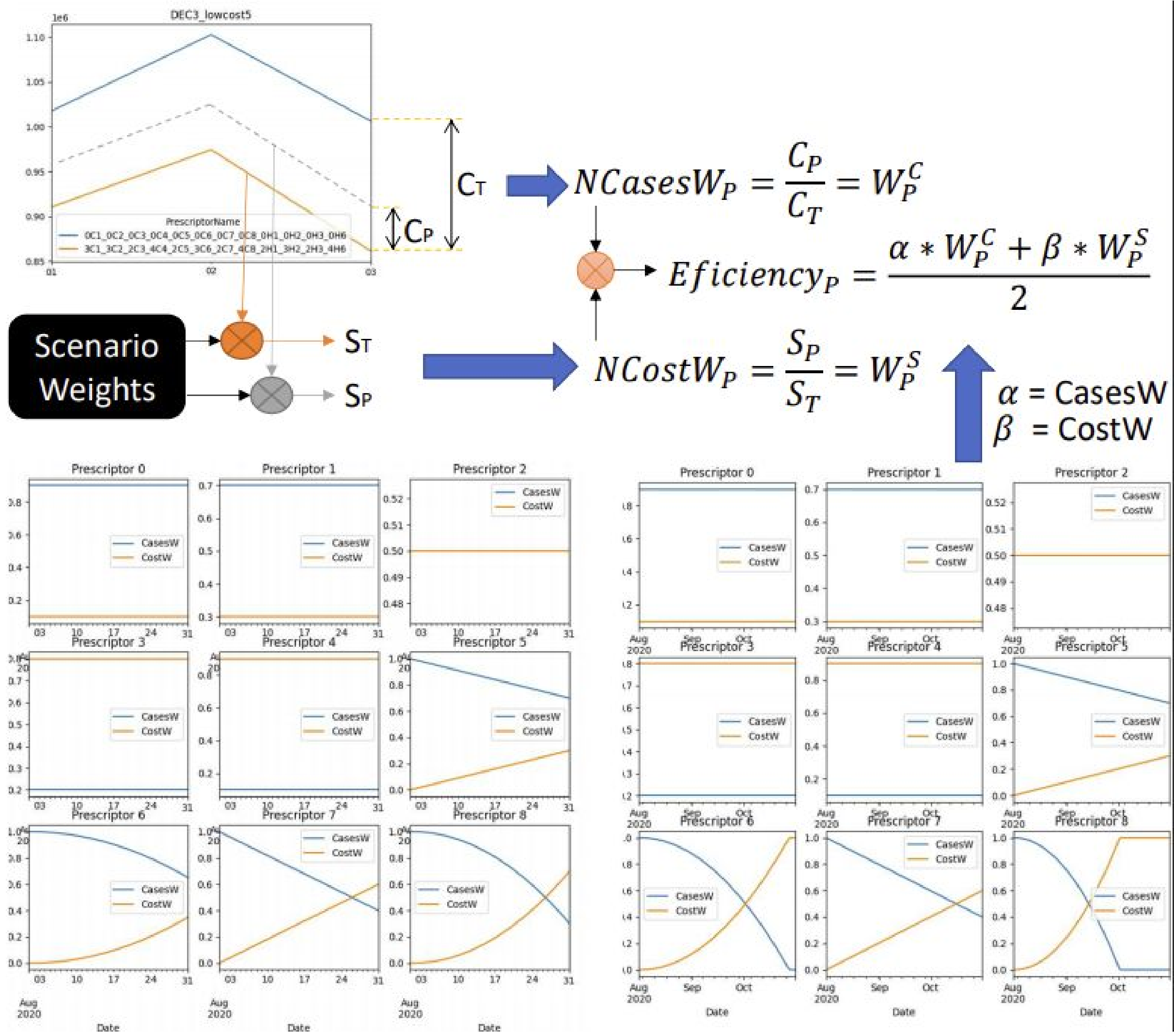
- We have a view of what is happening in the real world using the different data sources
- We model the evolution of the pandemic and the impact of the taken decisions with the predictor. The predictor is our model of the real-world.
- The prescriptor takes different policies and uses the predictor to evaluate the impact of different strategies following those policies, finding the best combination of prescriptions. Using policies, we tell the prescriptor our priorities.

Full specification:

1. We selected the top 10 prescriptions from experiments 2 and 4 (appendix D). [PRECOMPUTED]
- We compute a global normalized cost in new cases per prescription as: (Prescriptor_NewCases- Best_prescriptor_newCases) / (Worse_prescriptor_newCases - Best_prescriptor_newCases). Best and worse correspond to the all-max-values and all-zeros prescriptions.
2. We weigh prescriptions using the provided costs. [COMPUTED ONCE].
- We compute a per region normalized cost in stringency per prescription as: Prescriptor_Stringency / Max_prescriptor_stringency. Max corresponds to the all-max-values prescriptions. Minimum will be always 0.
3. We rank all prescriptions according to a linear combination of previous normalized values for cost in new cases and stringency. [COMPUTED FOR EACH DATE]. Changing the weight we give to each of the two normalized values we define the strategy.

Prescription Strategy:
Each IP is one of the following:
- Constant weights: Weights do not evolve during the period
- Linear weights evolution: Weights change linearly during the period (decrease the relevance on new cases and increase the relevance on the stringency)
- Exponential weights evolution: Weights change exponentially during the period (decrease the relevance on new cases and increase the relevance on the stringency)

The prescription strategy is depicted in the following figure:



D. Common predictor analysis

Prior to the development of the submitted solution, we performed several experiments on the common predictor provided by XPRIZE:

- **Experiment 1:**

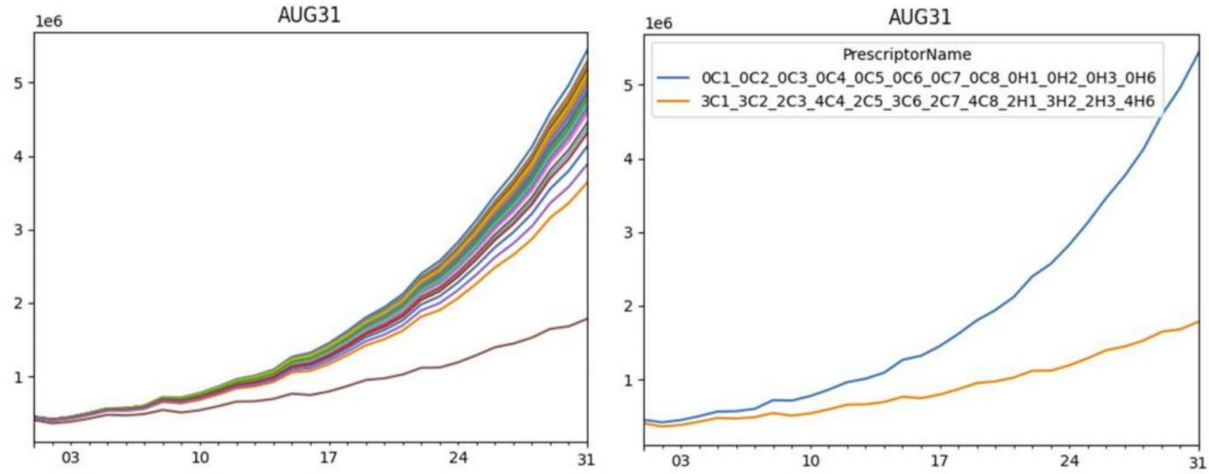
1) We generated a first set of prescriptions using a single IP (Intervention Plan) with all valid values (0 to MAX_VAL)

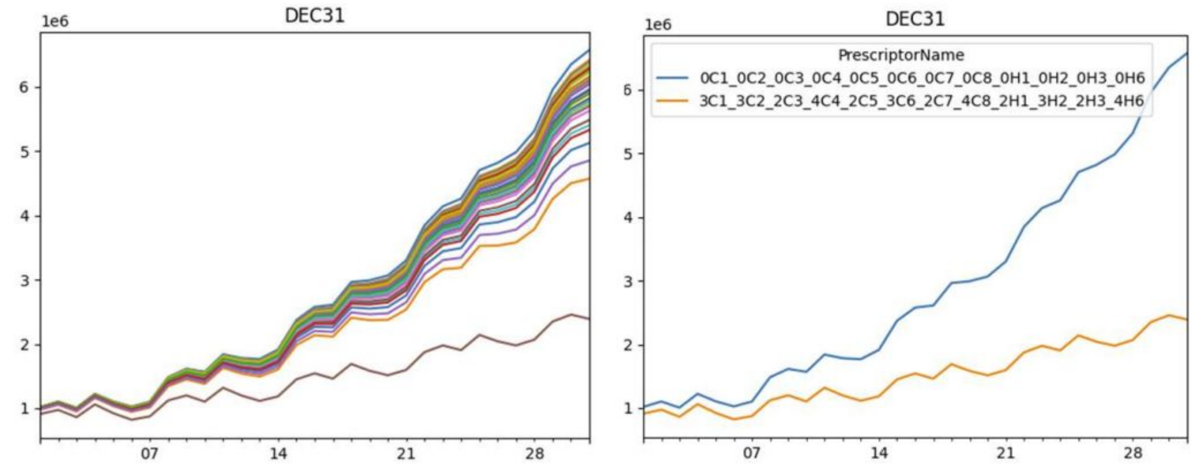
2) We selected 2 months of 2020 as case examples (August and December) and predicted the per day new cases for all months with a fixed prescription.

3) We note AUG31 for the predicted new cases during August 2020 and DEC31 for the predicted new cases during December 2020. As limits we use a prescription with all values to zero and one with all values to their maximum value.

4) We use the following prescription codification: [Vc1]C1_[Vc2]C2_ … _ [VH6]H6 where [VIP]is the value assigned to the prescripton IP.

Results:



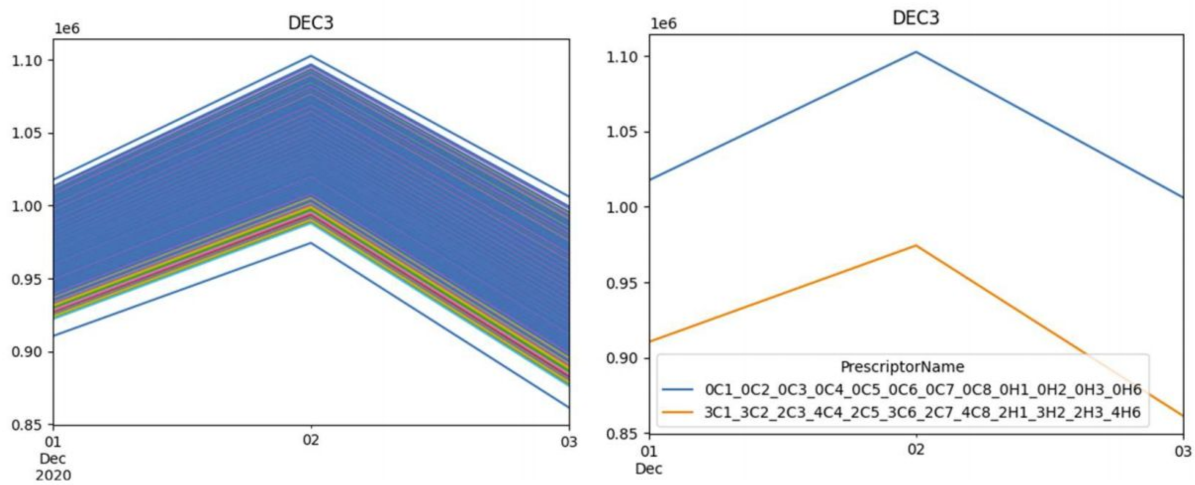


At the left we have the new cases evolution for all single IP prescriptions plus the all IPs at maximum value. At right we have only the extreme prescriptions (all zero and all max value). We can see that a) new cases tend always to increase and we only affect the speed, b) Extreme prescriptions seem to provide us with the best and worst possible values at each moment and c) the ranking of the most effective IP (in reduction of the number of cases) is stable from the 3rd day of predictions until the last.

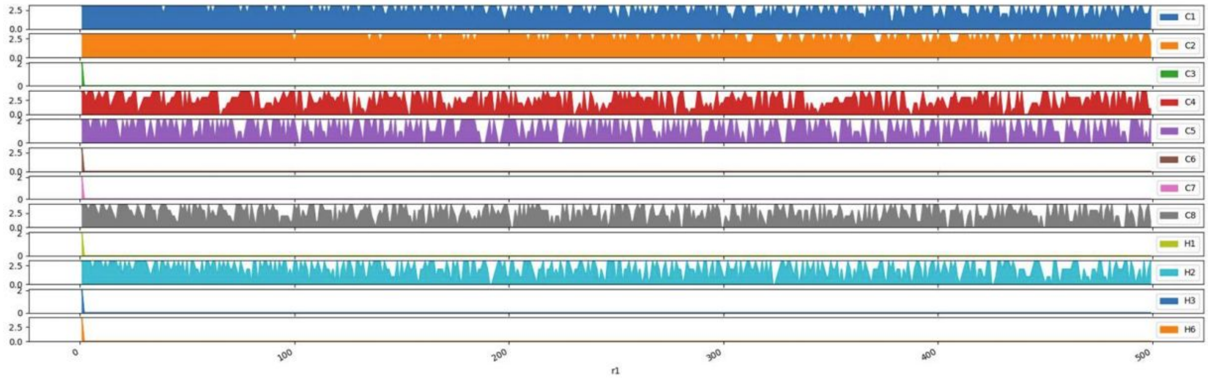
- **Experiment 2:**

1) We select the 6 most significative IPs from experiment 1 and generate all possible prescriptions with only those 6: C1, C2, C4, C5, C8, H2 => 4.800 prescriptions (40h computation)
2) We evaluate all those prescriptions in a small scenario of 3 days: DEC3: Predicted new cases for only the first three days of December 2020.

Results:



At the left we have the new cases evolution for all 4800 IPs. At right we have only the extreme prescriptions (all zero and all max value). Below this paragraph: Distribution of the 12 different NPIs for the top-500 PIs. We observe the preponderance of C1, C2, C4, C5, C8 and H2, which can be considered as the most effective NPIs according to the common predictor.



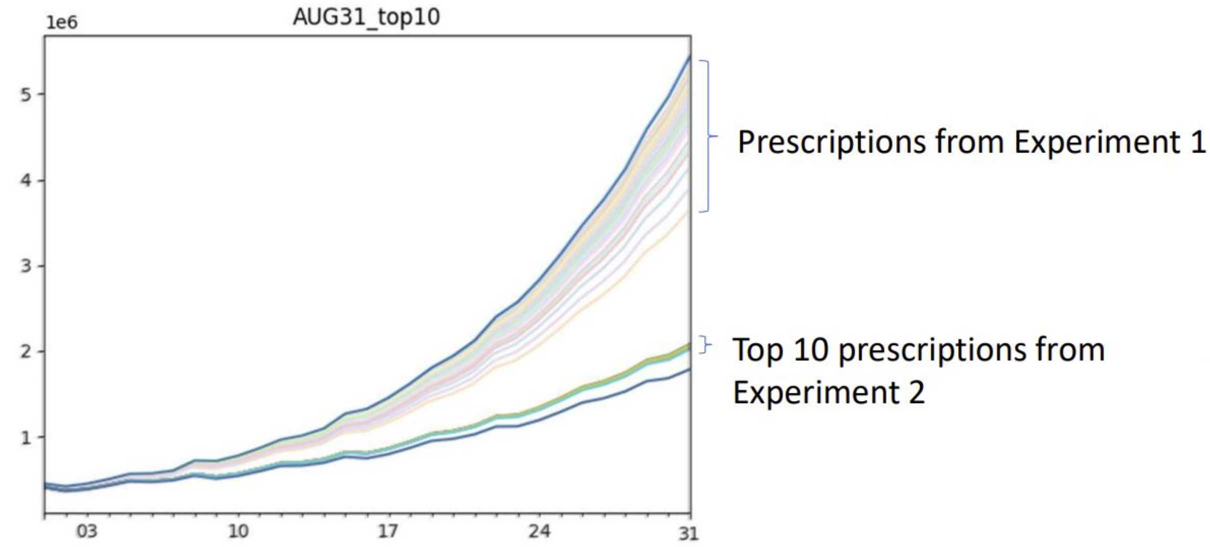
	PrescriptorName	PredictedDailyNewCases	r0	r1
	3C1_3C2_2C3_4C4_2C5_3C6_2C7_4C8_2H1_3H2_2H3_4H6	915460.756800	1.0	1.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	928784.222831	2.0	2.0
	3C1_3C2_0C3_3C4_2C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	929971.169797	3.0	3.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	930076.385843	4.0	4.0
	3C1_3C2_0C3_4C4_1C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	930881.802999	6.0	5.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_4C8_0H1_2H2_0H3_0H6	930897.886746	5.0	6.0
	3C1_3C2_0C3_2C4_2C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	931240.972957	7.0	7.0
	3C1_3C2_0C3_3C4_2C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	931331.828409	8.0	8.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_2C8_0H1_3H2_0H3_0H6	931493.924676	9.0	9.0
	3C1_3C2_0C3_3C4_1C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	932161.084098	11.0	10.0
	3C1_3C2_0C3_3C4_2C5_0C6_0C7_4C8_0H1_2H2_0H3_0H6	932190.411643	10.0	11.0
	3C1_3C2_0C3_4C4_1C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	932305.663740	13.0	12.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_3C8_0H1_2H2_0H3_0H6	932302.928104	12.0	13.0
	3C1_3C2_0C3_1C4_2C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	932601.585889	14.0	14.0
	3C1_3C2_0C3_2C4_2C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	932677.644549	15.0	15.0
	3C1_3C2_0C3_3C4_2C5_0C6_0C7_2C8_0H1_3H2_0H3_0H6	932828.878639	16.0	16.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_1C8_0H1_3H2_0H3_0H6	933057.577079	17.0	17.0
	3C1_3C2_0C3_4C4_1C5_0C6_0C7_4C8_0H1_2H2_0H3_0H6	933160.102512	18.0	18.0
	3C1_3C2_0C3_4C4_0C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	933184.065599	20.0	19.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_4C8_0H1_1H2_0H3_0H6	933256.541459	19.0	20.0
	3C1_3C2_0C3_2C4_1C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	933531.591312	22.0	21.0
	3C1_3C2_0C3_2C4_2C5_0C6_0C7_4C8_0H1_2H2_0H3_0H6	933574.222098	21.0	22.0
	3C1_3C2_0C3_3C4_1C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	933662.854933	24.0	23.0
	3C1_3C2_0C3_3C4_2C5_0C6_0C7_3C8_0H1_2H2_0H3_0H6	933670.256726	23.0	24.0
	3C1_3C2_0C3_4C4_2C5_0C6_0C7_2C8_0H1_2H2_0H3_0H6	933846.960946	25.0	25.0
	3C1_3C2_0C3_4C4_1C5_0C6_0C7_2C8_0H1_3H2_0H3_0H6	933874.207733	26.0	26.0
	3C1_3C2_0C3_0C4_2C5_0C6_0C7_4C8_0H1_3H2_0H3_0H6	934064.905052	27.0	27.0
	3C1_3C2_0C3_1C4_2C5_0C6_0C7_3C8_0H1_3H2_0H3_0H6	934124.576478	28.0	28.0
	3C1_3C2_0C3_2C4_2C5_0C6_0C7_2C8_0H1_3H2_0H3_0H6	934263.497830	29.0	29.0

At the left: ranking of the top-29 among the 4800 IPs. We observe that they all have C1 and C2 to the maximum value.

- **Experiment 3:**

1. We generated a set of prescriptions using top ranked prescriptions from experiment 2 (10 best).
2. We apply best prescriptions from experiment 2 (from December 2020) to August 2020. We note AUG31_2: Predicted new cases during August 2020.

Results:



We verify that the rankings are stable across 2 different months: prescriptions in December will perform similarly in August.

Key findings from Experiments 1-3:

a) The number of new cases have a lower bound on the "all-values-max" prescription and a higher bound with "all-zero-values" prescription.

b) All prescriptions have the same behaviour, after day 2 the rank between prescriptions does not change. It is not necessary to predict all the month.

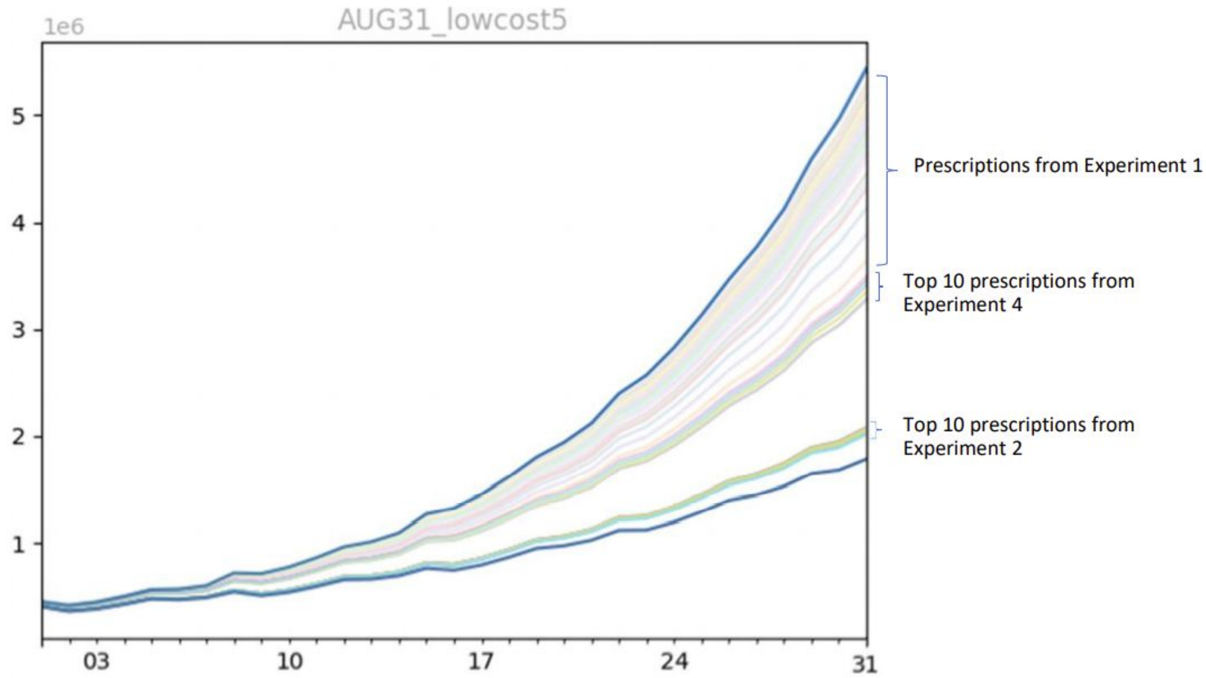
c) We can have good results using only 6 from 12 Ips (max stringency reduction from 34 to 19.

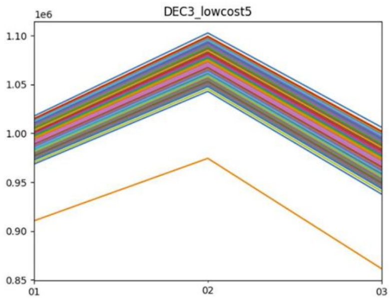
d) Good prescriptions have a large value on C1 and C2.

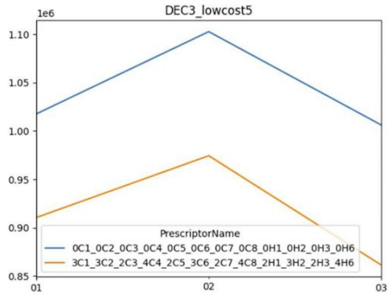
● **Experiment 4:**

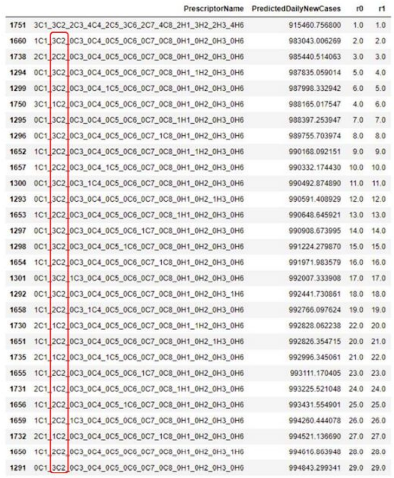
1. We select all possible prescriptions with maximum stringency of 5: 4.800 prescriptions.
2. We evaluate all those prescriptions in a small scenario of 3 days. We note AUG3: Predicted new cases for only the first three days of August 2020.

Results:





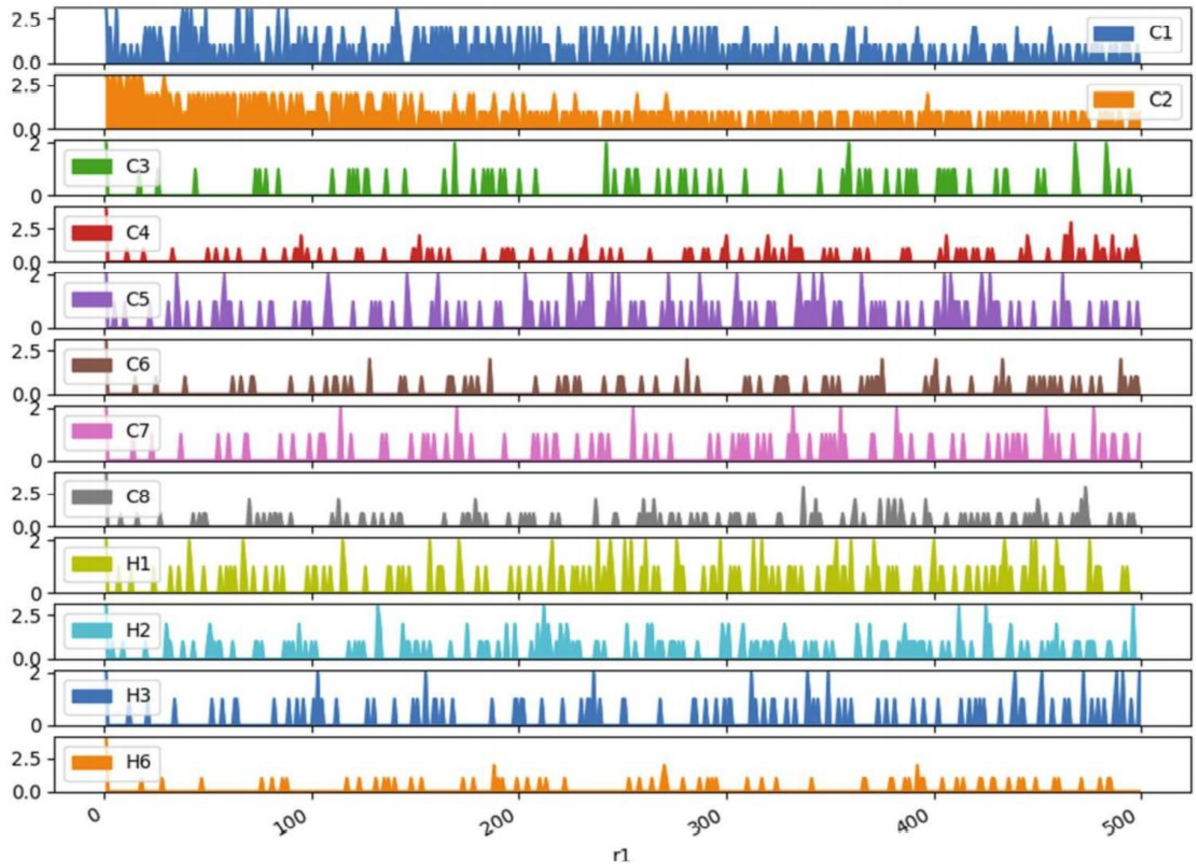




	PrescriptorName	PredictedDailyNewCases	r0	r1
1751	3C1_3C2_2C3_4C4_2C5_3C6_2C7_4C8_2H1_3H2_2H3_4H6	915460.756800	1.0	1.0
1660	1C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	983043.006269	2.0	2.0
1738	2C1_2C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	985440.514063	3.0	3.0
1294	0C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_1H2_0H3_0H6	987835.059014	5.0	4.0
1299	0C1_3C2_0C3_0C4_1C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	987998.332942	6.0	5.0
1750	3C1_1C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	988165.017547	4.0	6.0
1295	0C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_1H1_0H2_0H3_0H6	988397.253947	7.0	7.0
1296	0C1_3C2_0C3_0C4_0C5_0C6_0C7_1C8_0H1_0H2_0H3_0H6	989755.703974	8.0	8.0
1652	1C1_2C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_1H2_0H3_0H6	990168.092151	9.0	9.0
1657	1C1_2C2_0C3_0C4_1C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	990332.174430	10.0	10.0
1300	0C1_3C2_0C3_1C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	990492.874890	11.0	11.0
1293	0C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_1H3_0H6	990591.408929	12.0	12.0
1653	1C1_2C2_0C3_0C4_0C5_0C6_0C7_0C8_1H1_0H2_0H3_0H6	990648.645921	13.0	13.0
1297	0C1_3C2_0C3_0C4_0C5_0C6_1C7_0C8_0H1_0H2_0H3_0H6	990908.673995	14.0	14.0
1298	0C1_3C2_0C3_0C4_0C5_1C6_0C7_0C8_0H1_0H2_0H3_0H6	991224.279870	15.0	15.0
1654	1C1_2C2_0C3_0C4_0C5_0C6_0C7_1C8_0H1_0H2_0H3_0H6	991971.983579	16.0	16.0
1301	0C1_3C2_1C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	992007.333908	17.0	17.0
1292	0C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_1H6	992441.730861	18.0	18.0
1658	1C1_2C2_0C3_1C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	992766.097624	19.0	19.0
1730	2C1_1C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_1H2_0H3_0H6	992828.062238	22.0	20.0
1651	1C1_2C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_1H3_0H6	992826.354715	20.0	21.0
1735	2C1_1C2_0C3_0C4_1C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	992996.345061	21.0	22.0
1655	1C1_2C2_0C3_0C4_0C5_0C6_1C7_0C8_0H1_0H2_0H3_0H6	993111.170405	23.0	23.0
1731	2C1_1C2_0C3_0C4_0C5_0C6_0C7_0C8_1H1_0H2_0H3_0H6	993225.521048	24.0	24.0
1656	1C1_2C2_0C3_0C4_0C5_1C6_0C7_0C8_0H1_0H2_0H3_0H6	993431.554901	25.0	25.0
1659	1C1_2C2_1C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	994260.444078	26.0	26.0
1732	2C1_1C2_0C3_0C4_0C5_0C6_0C7_1C8_0H1_0H2_0H3_0H6	994521.136690	27.0	27.0
1650	1C1_2C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_1H6	994616.863948	28.0	28.0
1291	0C1_3C2_0C3_0C4_0C5_0C6_0C7_0C8_0H1_0H2_0H3_0H6	994843.299341	29.0	29.0

At the left we have the new cases evolution for all 4800 IPs. At center we have only the extreme prescriptions (all zero and all max value). At right: ranking of top-29 PIs, we observe that C2 remains very present.

Below this paragraph: Distribution of the 12 different NPIs for the top-500 PIs. We observe that with the constraint of maximum stringency 5, we observe a more balanced distribution of the different NPIs.



E. OpenAI gym environment for policy discovery

We developed a reinforcement learning environment `CovidEnv` following the OpenAI Gym interface. This environment was not used to submit our final strategy because of the long training time required by the on-policy learning algorithm we tried (i.e., Proximal Policy Optimization). However, we think that releasing this kind of environment could help the AI community contribute even more to this research. The main idea is to encompass the `Predictor` in the RL environment. To initialize the environment there are different arguments:

- `predictor_script_path`: path to the predictor.py Python script following the exact interface of the challenge (so generic).
- `oxford_csv_path`: path to the Oxford data.
- `future_days`: the number of future days predicted by the predictor at each step.

- `lookback_days`: the window size of past information provided in each observation.
- `episode_length`: the length of each RL trajectory played by the agent.
- geoids: a list of geoids from which will be drawn one for each episode.
- `weights_constant`: if None, it will be drawn randomly following the given constraints. Otherwise a constant value can be used (e.g., 0 if the economic impact should not be taken into consideration).

The **action space** is a multi-dimensional discrete space where each dimension corresponds to a given NPI column and is defined by the minimal and maximal value of this same NPI. An action represents a decision for all NPIs on a given day.

The **observation space** is a dictionary composed of 3 keys:

- `npis`: past NPIs information with shape (`lookback_days`,number of NPIs)
- `new_cases`: past information about daily new cases with shape (`lookback_days`,)
- `weights`: economic weights of each NPI with shape (number of NPIs,)

The **reward** (maximised by the agent) is computed from (npis, weights, new_cases) with the following formula:

$$r_t = -[C_{eco}(x_t^{npi}) + \gamma C_{cases}(x_t^{cases})]$$

Where C_{eco} is the economic cost described in the challenge guidelines and C_{cases} is the sum of estimated r_0 (reproduction number) on the future new number of cases provided by the predictor.

This environment is working and easy to plug in RL libraries such as Tensorforce.