JSI vs COVID prescriptor

1 Description of the approach

In the second competition phase we had to create intervention plans for different time periods and for different regions. Such intervention plans should have good trade-offs between the stringency of the interventions and the projected infections that result from them. This task can be formulated as a multi-objective optimization problem with two conflicting objectives. Such problems are commonly tackled with multi-objective evolutionary algorithms (MOEAs) that imitate the biological evolution to search the space of possible intervention plans, evaluate them in terms of their stringency and the number of infections, and find plans with good trade-offs between the objectives. Using an MOEA based on Nondominated Sorting Genetic Algorithm II (NSGA-II), we achieved good results, but this method turned out to be computationally too expensive, as each call to the standard predictor needed a few seconds to evaluate each region. We thus developed two methods derived from this standard multi-objective optimization approach, and then combined them at the end.

Our first method was to compute several plans in advance, and then for a specific region select the plan that is the most appropriate for the current situation in that region. This plan is then adjusted to the provided intervention weights. As an example of how to interpret this method, imagine we have a region with infections quickly rising, we then select a plan that has stringent interventions at the beginning to counteract them, and finally select specific interventions whose weights add up to the planned stringency. Details of this method are found in Section 1.1.

Our second method was to use a surrogate model instead of the standard predictor, i.e., a model that generates very similar predictions but is much faster. We used a Susceptible-Exposed-Infective-Recovered (SEIR) epidemiological model for the task, as it is not only fast and can be tuned to fit the predictions made by the standard predictor well, but is also explainable and semantically suitable for the domain. It allowed us to do the multi-objective optimization for each region directly in the allotted time. The details of this method are found in Section 1.2.

After obtaining ten intervention plans from each method, we evaluated them all with the standard predictor and then chose ten that maximize the hypervolume. The hypervolume is the area bounded by the obtained Pareto front approximation; the Pareto front approximation consists of solutions that are non-dominated (i.e., no other solution was found that has both lower stringency and results in fewer infections). Aside from removing dominated solutions, maximizing the hypervolume tends to make the selected solutions evenly spaced.

1.1 Pre-computed plans

As previously stated, several intervention plan templates were pre-computed using NSGA-II with a custom solution representation. We identified 12 different situations and for each of them created 40 different plans – with different stringency/infections trade-offs. Ten of these plans were 90 days in length, and we added ten for each of the 45, 60 and 75 days durations.

1.1.1 Possible situations

The plan chosen to be used for a specific region is based on the infection data leading to the start date of the intervention planning and the population of that region. To choose it, we first estimated the slope of the change of the number of daily infections as either stagnant, raising, raising fast, falling, falling fast, or close to zero. These six categories were further split depending if the region's population is larger or smaller compared to the median population number. This results in twelve combinations that categorize different situations that one can address with different plans.

1.1.2 Representation of the optimization solutions

The intervention plans were represented as vectors, where the *i*-th variable represents what stringency to use on the *i*-th week. We decided to optimize with the granularity of one week for two reasons: 1) it is unrealistic to expect real-life policies to change with a higher frequency and 2) the quality of the solutions did not substantially improve when using a smaller granularity. It is of note that this granularity parameter is adjustable and can be set to a different duration if a decision maker so desires.

To evaluate such a vector of stringencies during the optimization process, they are first extended so that each variable represents one day, then each stringency is replaced by a combination of twelve intervention values (i.e., one value per intervention type), and finally the resulting matrix is sent to the standard predictor. The intervention combination with which to replace each stringency is selected as the one with the lowest projected infections out of those that have the appropriate stringency. The details of this replacement are discussed in the next section.

It is of note that instead of using the stringency vector representation, we could instead be working with the *m* x *n matrix* representation, where *n* is the number of weeks and *m* is the number of individual interventions. This is the standard multi-objective optimization approach, but – as mentioned at the beginning – it converged more slowly than the stringency vector approach. However, it can still be used if the properties of the standard predictor change in the future.

1.1.3 Selecting specific interventions

Our empirical tests with the standard predictor indicated that its output is partially predictable. As an example, if using only intervention A (e.g., *C6_Stay_at_home_requirements* = 3) reduced the infections by 3% compared to using no interventions, and using only intervention B reduces the infections by 2%, then using them both combined reduces the infections by nearly 5%.

This motivated us to create a linear model that can predict the infection percentage drop, given an arbitrary intervention combination. While this model turned out to be relatively inaccurate in its absolute predictions, it conserved the relative order of intervention combinations well. This allowed us to rank every one of the roughly 7 million combinations and store them ordered. When needing to transform a stringency into an intervention combination, we simply took the highest-ranking combination given the stringency limit. One can interpret this procedure as having expert knowledge on which interventions are generally more effective than others.

The stringency-to-intervention replacement happens in two places: 1) during the optimization to evaluate the intervention plans and 2) when applying a pre-computed intervention plan (in the stringency vector representation) to a new region with different intervention weights.

1.2 Optimization with the SEIR model

The basic optimization was done in almost the same way as described in Section 1.1 – using weekly granularity and replacing the stringency with the appropriate intervention combination. The only two exceptions were that this optimization was not done in advance, but directly for the regions and time intervals of interest, and that a fast surrogate model was used instead of the standard predictor.

1.2.1 The surrogate model

For the surrogate model we used the same epidemiological SEIR model that we submitted and described in the first phase of the competition, with two differences. First, its parameters were fitted to the standard predictor's outputs instead of ground-truth infections. And second, it was rewritten in Cython (static compiler for Python) to be faster.

This SEIR model needs a vector of beta values in order to correctly predict the infections for a given interval. The beta value dictates how many people each infected person infects (it is the probability of transitioning from the susceptible to the exposed compartment in the model). We assumed that the beta is a function of intervention values. We therefore trained a machine-learning model (Random Forest) that maps from these intervention values, region population and the recent number of infections to the appropriate beta values (again similarly to the first phase of the competition).

Each week can potentially have different intervention values, and thus a different beta. Altogether the surrogate model has three steps: 1) replace the stringency with the actual interventions, 2) determine beta values based on the interventions and 3) predict the infections using the SEIR model.

1.3 Putting everything together

To recap the whole pipeline: for each region we first retrieve the data for the three weeks leading to the start date. This data is either stored in a historical file or is computed with the standard predictor if it is not pre-stored. Based on this data, the region is classified into one of the twelve situations.

If the infection data always equals 0, we prescribe to use no interventions. Otherwise, we run both methods. The pre-computation method chooses a plan based on the situations and the number of days required, while the optimization uses the surrogate model to create a new intervention plan from scratch.

Both methods are comparable in the quality of intervention plans, but using both and combining the results frequently increases the overall quality of the obtained Pareto front approximation. In case that our methodology is taking more time per region than allotted by the competition rules, only the faster pre-computation method is used for the next few regions.

Having the twenty solutions – ten from each method – we finally select ten best one for the final submission. This step was done using the greedy Hypervolume Subset Selection (gHSS) method.

2 Addressing the qualitative evaluation criteria

2.1 Actionability and usability

To be able to interpret the prescribed solutions, we have prepared a visualization web page available here: http://docker-e9.ijs.si:4444/visualization. The page shows the intervention plans in the objective space (daily infections vs stringency), how infections and stringency change with time, and how individual interventions in the plan vary with time. Furthermore, the web page offers a tool for comparing intervention plans. Each intervention plan can also be modified and the result compared with other plans, so that decision-makers can customize plans to account for additional real-life restrictions they encounter. Finally, we are exploring advanced visualizations using more than two dimensions to show the interplay between infections, stringency, time and individual interventions. We may also link external sources of relevant information such as scientific papers and social media in the future.

2.2 Explanation

In addition to the description provided here (which is limited to three pages), we plan to publish a scientific paper describing our approach in detail.

While MOEAs are inherently stochastic, the generated solutions can be easily interpreted using the previously outlined visualization tools. In addition, by defining twelve semantically relevant situations and pre-calculating intervention plan templates for them, the policy maker can observe the different responses best suited for each situation. The actual interventions picked to fill the templates for a specific region are dependent solely on the effectiveness of the interventions and their weight – their choice being deterministic and transparent.

2.3 Addressing the challenge

The intervention plans we propose are almost completely determined by the intervention weights, predictor and time period for changing interventions. The weights are an input and can be as realistic as a decision maker can make them. We use the standard predictor (and a surrogate model that approximates it), but we could easily use another one if it works better in the real world. The interventions change on a weekly basis by default, which is quite realistic, and can also be modified.

2.4 Inclusivity and fairness

Because our prescriptor's operation is almost completely determined by its inputs, it is as inclusive and fair as the inputs. Setting intervention weights fairly goes way beyond our work. If data on infections among vulnerable groups were available, the predictor could output infections weighted by the impact on these groups, and our prescriptor would take this into account automatically.

MOEAs search the space of possible intervention plans stochastically with little bias of any kind other than preferring solutions with low stringency and resulting in few infections. The same algorithm configuration was used for all countries and regions guaranteeing the same quality of solutions everywhere.

2.5 Transparency and trust

The intervention plans are fairly self-explanatory, and we provide a web application for visualizing them, which we believe to be reasonably accessible to a layperson. Our approach for obtaining these intervention plans is somewhat more complex, but all the methods and data will be made available in a public repository, together with a more thorough description of the used methodology. We believe that with a good explanation, the essence of all the methods used could be understood even by a layperson.

2.6 Collaborative contributions

We shared the web application for visualization, which we found quite useful for understanding the solutions our prescriptor produced, with other competitors. This was done on the competition's Slack as soon as the application was in good enough shape to do so.

2.7 Innovation

Two key insights enabled us to develop our prescriptor. The first was that optimizing the overall stringency (rather than individual interventions) leads to effective intervention plans. This enabled us to drastically reduce the search space, making it possible for a state-of-the-art MOEA to find good solutions in a reasonable time. The second insight was that a simple linear model can be used to combine the effect of individual interventions, which was a prerequisite for efficiently replacing stringencies with the best combination of interventions (this problem is analogous to the knapsack problem and thus NP hard if tackled without this insight). While the two insights are critical for our prescriptor, we believe they are also interesting discoveries useful outside this particular task.