EURO Journal on Computational Optimization 10 (2022) 100025

Contents lists available at ScienceDirect

EURO Journal on Computational Optimization

journal homepage: www.elsevier.com/locate/ejco

Robot Dance: A mathematical optimization platform for intervention against COVID-19 in a complex network

Luis Gustavo Nonato [a], Pedro Peixoto [b], Tiago Pereira [a], Claudia Sagastizábal [c,*], Paulo J.S. Silva [c]

[a] ICMC, USP, São Carlos, SP, Brazil
[b] IME, USP, São Paulo, SP, Brazil
[c] IMECC/UNICAMP, Campinas, São Paulo, Brazil

ARTICLE INFO

MSC:
90C15
49-04
92C60

Keywords:
Mathematics of COVID-19
Stochastic optimization
Mobility matrix
Complex networks

ABSTRACT

Robot Dance is a computational optimization platform developed in response to the COVID-19 outbreak, to support the decision-making on public policies at a regional level. The tool is suitable for understanding and suggesting levels of intervention needed to contain the spread of infectious diseases when the mobility of inhabitants through a regional network is a concern. Such is the case for the SARS-CoV-2 virus that is highly contagious and, therefore, makes it crucial to incorporate the circulation of people in the epidemiological compartmental models. Robot Dance anticipates the spread of an epidemic in a complex regional network, helping to identify fragile links where applying differentiated measures of containment, testing, and vaccination is important. Based on stochastic optimization, the model determines efficient strategies on the basis of commuting of individuals and the situation of hospitals in each district. Uncertainty in the capacity of intensive care beds is handled by a chance-constraint approach. Some functionalities of Robot Dance are illustrated in the state of São Paulo in Brazil, using real data for a region with more than forty million inhabitants.

1. Context and motivation

The impact that COVID-19 has had in our lives is no news nowadays. Driven by the needs of their population and governments, different countries have reacted against the pandemic in different manners. The diversity observed worldwide in public policies is in stark contrast with the uniform and massive response that the international scientific community has given without distinction of disciplines or frontiers. Throughout the globe, researchers swiftly gathered forces to offer clues and responses to the challenges presented by the pandemic on various fronts. The state of affairs is no different in Brazil, a country with continental dimensions and very heterogeneous society. This last feature poses additional challenges to an already difficult situation, as it puts in question the effectiveness of uniform policies of quarantine adopted in many regions to contain the spread of coronavirus.

Robot Dance is an integrated computational optimization platform developed to assess and estimate the consequences of interventions taken on a regional level when there is a disease outbreak and the spread of the disease is affected by the circulation of people living in the region. A question that all governments face worldwide is how to mitigate (present and future) pandemic waves while limiting collateral economic damage. Robot Dance assists the decision-making, providing a response that is customized to the specific local needs. By "local" we mean any territorial division, of political or administrative nature, into which the area of interest is partitioned (regional health districts, cities, neighborhoods). Robot Dance considers the region as a network with nodes through which the disease propagates, following patterns whose shape and intensity change from night to daytime, particularly during working hours.

The public health of the region depends on the dynamic evolution of several elements that are intertwined in non-trivial patterns. For Robot Dance, these elements are the mobility of inhabitants between the nodes of the network, the epidemiological situation in each district, and the hospitals' capacity in the considered region. Having been designed in a mathematical optimization framework, Robot Dance searches for an efficient assignment of resources, taking into account those three aspects on a regional level, over a horizon of several months. To guide the search, the user can choose among different objective functions, that can aim at maximizing the free circulation, or minimizing the length of consecutive lockdowns, or other options described in Section 5.1 below.

The state of São Paulo, the most populous and the epicenter of the COVID-19 outbreak in Brazil Candido et al. (2020b) is a good example of a complex network where the virus circulates. With a surface comparable to the United Kingdom, and a population similar to the one in Spain,

* Corresponding author.
E-mail addresses: gnonato@icmc.usp.br (L.G. Nonato), ppeixoto@usp.br (P. Peixoto), tiago@icmc.usp.br (T. Pereira), sagastiz@unicamp.br (C. Sagastizábal), pjssilva@unicamp.br (P.J.S. Silva).

https://doi.org/10.1016/j.ejco.2022.100025
Received 5 July 2021; Received in revised form 13 December 2021; Accepted 26 January 2022

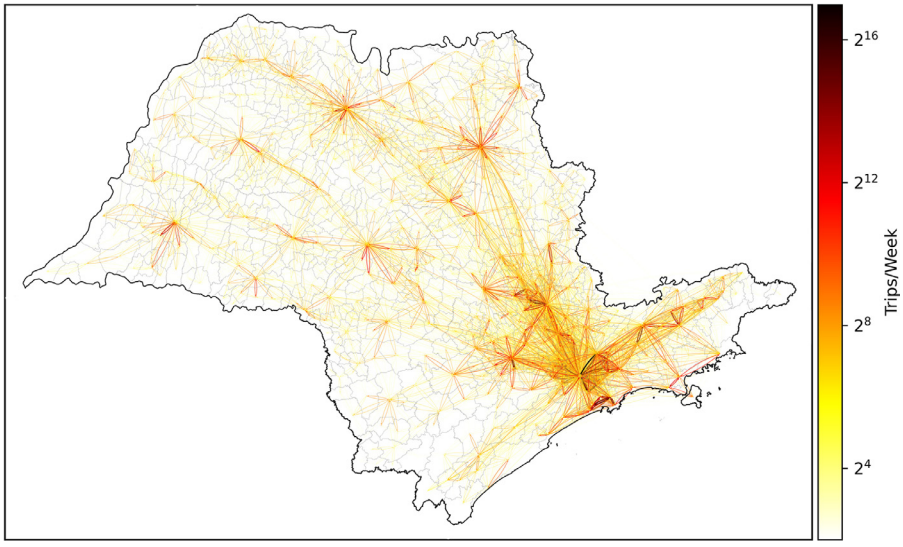

Fig. 1. Number of daily trips in a typical pre-pandemic week in São Paulo state, Brazil (February 2020). The star-shaped lines, in darker colors when the circulation between nearby cities is higher, reveals each star center as an important commuting hub for the network; see also Peixoto et al. (2020).

São Paulo is responsible for about 33% of Brazil's GDP. As illustrated by Fig. 1, many people travel to work, commuting daily from smaller towns to large urban centers. The figure also reveals several such hubs spread across the state.

The situation described by Fig. 1 is typical in Brazil and, more generally, common in regions having a few central points that concentrate most of the economic infrastructure. To go to work, large portions of the population are forced to travel long distances every day between cities (or between nodes in the network). Because of this feature, keeping a global view both in space and time is essential to suitably capture the disease transmission. Besides, those nodes are highly heterogeneous (in terms of wealth, of hospital capacity, of business activities), making the impact of mobility also noticeable beyond the daily scale. The phenomenon was indeed observed for COVID-19 in São Paulo, with the well-off population initially carrying the virus from the hubs with international airports to the rest of the state, progressively more and more to the inland, following a long-term circulation pattern Candido et al. (2020a); Carmo et al. (2020).

To obtain a rigorous mathematical model that represents well the COVID-19 situation, it is crucial to suitably merge epidemiological knowledge with tools from several areas, including optimization, data science, visualization, and computational mathematics. Furthermore, in the considered setting, there is an acute lack of reliable data on the progress of the pandemic, due to subnotification, asymptomatic cases, and changes in how cases are reported over time in the official databases. This handicap is a recurrent difficulty for all the mathematical developments related to COVID-19. To mitigate the impact of this issue in the decision process, in addition to intensive backtesting and cross-validation of the parameters defining the problem, Robot Dance incorporates uncertainty in the use of intensive care unit (ICU) beds. The limit in ICU capacity, a constraint in the optimization problem, is then handled in the form of probabilistic constraints Charnes and Cooper (1959); Dentcheva (2006); Prékopa (1995). In this manner, even if reliable data is scarce, the tool still provides useful insight to compare on a qualitative level the merits of different strategies of containment.

The platform is very versatile and, given appropriate mobility and epidemiological input data, can determine different mitigation protocols that make efficient use of the ICU beds. Those features become particularly useful for a network like the one represented in Fig. 1, having a complex heterogeneous topology. In the case studies presented in Section 5 and thanks to the incorporation of the mobility network in the epidemiological model, Robot Dance is able to pinpoint critical locations where targeted policies, of local nature, have the most impact. For São Paulo state this means that, instead of putting in lockdown the whole state at once, different degrees of social distancing can be imposed in different districts, depending on the local situation of the hospitals, the severity of infections, and the commuting patterns. Another feature of Robot Dance that is presented in our numerical assessment refers to transferring patients to neighboring districts when the local system is close to saturation. Clearly, this is a region-dependent consideration. When the local ICU capacity is attained, it is natural to use resources in the proximity, transferring patients to nearby regions. Robot Dance models this feature by a pool of beds that is shared by several districts. With this mechanism, nodes with spare capacity can reserve a small portion of beds to receive patients from areas with poor hospital infrastructure. The pooling is illustrated for the state capital, São Paulo city, and the five suburban districts that compose its metropolitan area. Since commuting occurs in the whole network, as illustrated in Fig. 1, those functionalities are obtained considering the twenty-two health districts that compose São Paulo state.

Before presenting details about Robot Dance's mathematical formulation, we provide a succinct overview of other mathematical models developed in response to the COVID-19 pandemic.

Some related works, contribution, and terminology

The literature about computational mathematical models that simulate and analyze the impact of COVID-19 outbreak has grown quickly in the past months. We focus the discussion on some works that describe the pandemic dynamics by compartmental models, like Susceptible-Infectious-Removed (SIR) and Susceptible-Exposed-Infectious-Removed (SEIR) models Brauer (2008). Rather than a thorough review (a formidable enterprise), we handpick a few works to contextualize Robot Dance's features. From this perspective, there are two main categories of models in the literature, one considering the pandemic dynamics only, and another group that complements the temporal distribution of the disease with geographical considerations, like Robot Dance.

Temporal methods typically rely on the calibration of parameters of compartmental models to analyze and forecast the behavior of the virus outbreak, drawing conclusions based on such estimates. For example, Aguas et al. (2020) first uses an SEIR compartmental model to estimate the reinfection, transmission, and recovery rates; these parameters are afterwards plugged into simulations, to determine when herd immunity can be reached. In Chen et al. (2020) the number of asymptomatic cases is approximated by fitting the time-dependent transmission and recovery rates for an SIR model. Similarly, the approach Acuña-Zegarra et al. (2020) fits the time-dependent parameters of an SEIR model to determine the impact of distancing measures and the time horizon required

for such measures to take effect. Optimal control formulations of SEIR models are considered in Maurer and do Rosario de Pinho (2016) and Bonnans and Gianatti (2020), for defining strategies of vaccination and confinement, respectively. The latter work defines age-dependent controls to show that performing differentiated confinement can help in minimizing the economic impact of containment measures. Among temporal models, Duque et al. (2020), the closest to Robot Dance, forecasts the severity of confinement protocols that keep the health system from collapsing in the short term. Rather than counting lockdown days, Robot Dance offers the possibility of alternating severe confinement with relaxed social distancing. Such a "dance" can be performed differently for different nodes, for instance keeping open together two nearby districts that have complementary economic activities (or one open and one closed, if better for the business of the region in terms of work patterns).

Spatio-Temporal methods consider the geographical dissemination of the disease due to human mobility. The typical approach is to distribute the population in groups, setting a compartmental model in each group, adapting the model to accommodate mobility between groups. This is the proposal for the SEIR variant in Wu et al. (2020), one of the first spatio-temporal models in the context of COVID-19. More sophisticated approaches were proposed later on, such as Arenas et al. (2020) which segregates the population in patches and in age strata to assess the performance of containment protocols. In Peixoto et al. (2020), the basis for the illustration in Fig. 1, geolocation mobile phone data is used to estimate the risk of infection in each city.

Robot Dance is based on a spatio-temporal model that splits the population into groups (nodes), representing the mobility as links in the network. The tool can anticipate the geographical evolution of the disease and evaluate the potential impact of containment and prevention strategies. This is fundamental as a response to studies like Kissler et al. (2020), foreseeing resurgences in contagion for at least three more years. Having cast the epidemiological transmission in a complex network, Robot Dance reveals certain epidemiological roads, links that will dominate the spread of the disease, as well as nodes where the need for ICU beds will be more acute. In this way, it is possible to forecast the effect of different containment protocols in a manner that the surges can be handled by the health system in each district, without collapsing. A distinctive feature of Robot Dance is the optimization of the protocols with the possibility of *alternating* the confinement measures in neighboring nodes. Allowing for contiguous districts to face different restriction levels makes it possible not only to keep some nodes relaxed while others undergo more severe measures of social distancing, but also to change the configuration of which districts are open and which ones are closed along the time horizon. It is this automatic choreography, alternating a "hammer" of strict confinement with the "dance" of relaxation, that gave the name to the tool, inspired from the blog post Pueyo (2020).

The significance of alternating hammer and dance is undeniable, as the policy-maker then has a strategy that avoids shutting down the economic activities in a region, all at once. The joint examination of city commute, health infrastructure, and specific containment measures district-wise, gives a global perception of the impact of different actions. For instance, it is possible to analyze the effect on the whole region of "surgical" local measures, such as increasing the number of ICU beds in some key nodes, critical for the network. In the case study in Section 5.3 and Section 5.4, once Robot Dance identifies nearby nodes that are on some critical epidemiological link, better containment policies are computed, putting in place a mechanism of cooperation, to share hospital infrastructure.

Having put the contributions of Robot Dance in context, we are now in a position to give the mathematical formulation of its key ingredients. The mathematical optimization problem under consideration minimizes an objective function over a feasible set with constraints describing the virus transmission, how people commute between cities for work, and the hospital capacity. The corresponding models are detailed below.

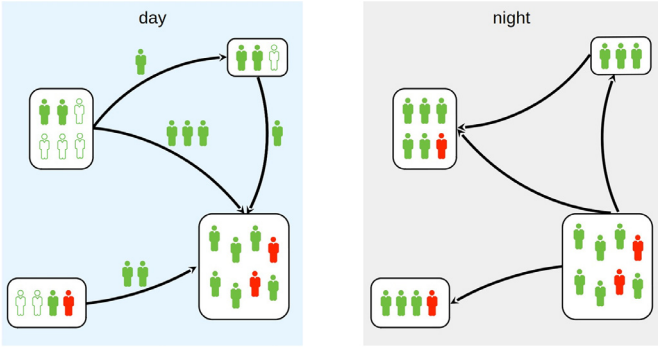

Fig. 2. Commuting between nodes changes the proportions of susceptible and exposed individuals.

2. Basic epidemiological constraints and continuous-time problem

Recall that nodes in the network correspond to cities, or administrative health districts, or some subdivision of the area of interest. The analysis is done for a network with nodes in a set I, with cardinality $|I|$.

Along the time horizon defined by given initial and final times $\mathbb{T}_0$ and $\mathbb{T}_1$, the epidemiological state of the *whole region* is characterized by the population compartments of Susceptible, Exposed, Infected, and Recovered individuals, considered as percentages of the total population in each district. The *Susceptible* individuals represent the people that do not have immunity to the disease and can get sick. The *Exposed* group represents people that were exposed to the pathogen and got infected, but are in an incubating period and do not transmit the disease yet. The *Infected* group represents the population that can actively infect others. Finally, the *Recovered* compartment represents those individuals that are not infecting other people anymore. The compartment definitions emphasize the state of the population groups with respect to the transmission of the disease and not with respect to their overall health state. In particular, the Recovered compartment includes, in addition to individuals that are fully recovered from the disease, sick ones that do not transmit anymore the virus, and also those who died from COVID-19. The capital letters $\mathcal{S}, \mathcal{E}, \mathcal{I}, \mathcal{R}$ denote the corresponding percentages, noting that each compartment is a vector in $[0,1]^{|I|}$. In a continuous-time formulation, the regional state at time t is given by the vector function,

$$x(t) := (\mathcal{S}, \mathcal{E}, \mathcal{I}, \mathcal{R})(t) \in [0,1]^{4\times|I|}, \text{ for } t \in [\mathbb{T}_0, \mathbb{T}_1],$$

noting that $\mathcal{S} + \mathcal{E} + \mathcal{I} + \mathcal{R} = 1$ for each node and time. When convenient, an individual component of the state vector is referred to by a sub-index, putting into brackets its ordinal position. This convention is used in (1) below, where $\mathcal{S}$ and $\mathcal{I}$, the first and third components of the vector x, are denoted by $x_{[1]}$ and $x_{[3]}$, respectively.

We shall also make use of a binary coefficient function $\alpha(t) \in \{0,1\}$ distinguishing time passed in the original node ($\alpha(t) = 1$) from time spent commuting and working in another node ($\alpha(t) = 0$). In Robot Dance, no mobility corresponds to the pandemic evolution only during the night hours, when individuals remain in their own node (typically during one third of the day, so $\alpha(t) = 1$ during $1/3$ of the time in our setting). The distinction is made necessary for the epidemiological model to consider the commuting to work during the day, with individuals traveling between nodes, potentially carrying the disease. The mathematical formulation of how circulation changes the susceptible and exposed compartments among nodes is given in Section 3, and a graphical illustration can be found in Fig. 2.

The effective infection rate $r(t) \in [0, \bar{r}] \subset \mathbb{R}^{|I|}$ defines the way the disease evolves, while T_{inc} and T_{inf} are the incubation and infection periods (the former taken as the lapse between exposure and the time when starting to actively spread the disease). For each node $i \in I$, given an initial condition at time $\mathbb{T}_0$, that is $x_0 = (\mathcal{S}_0, \mathcal{E}_0, \mathcal{I}_0, \mathcal{R}_0)$ with

$\mathcal{S}_0 + \mathcal{E}_0 + \mathcal{I}_0 + \mathcal{R}_0 = 1$, the pandemic evolution during the night is described by a system of ordinary differential equations:

$$\dot{\mathcal{S}}^i(t) = -\frac{\alpha(t)}{T_{\text{inf}}} r^i(t)\mathcal{S}^i(t)\mathcal{I}^i(t) \qquad \dot{\mathcal{E}}^i(t) = \frac{\alpha(t)}{T_{\text{inf}}} r^i(t)\mathcal{S}^i(t)\mathcal{I}^i(t) - \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t)$$
$$\dot{\mathcal{I}}^i(t) = \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t) - \frac{1}{T_{\text{inf}}}\mathcal{I}^i(t) \qquad \dot{\mathcal{R}}^i(t) = \frac{1}{T_{\text{inf}}}\mathcal{I}^i(t).$$

In this dynamical system, nonlinearity stems from the product $\mathcal{S}^i(t)\mathcal{I}^i(t)$, quantifying the amount of susceptible population of node i that gets exposed to infection. Defining the intermediate variables, for each $i \in I$,

$$y^{ii}(t) := \mathcal{S}^i(t)\mathcal{I}^i(t) = x^i_{[1]}(t)x^i_{[3]}(t), \tag{1}$$

the relations for susceptible and exposed compartments become,

$$\dot{\mathcal{S}}^i(t) = -\frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) \quad \text{and} \quad \dot{\mathcal{E}}^i(t) = \frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) - \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t) \tag{2}$$

(the somewhat redundant notation with a double super-index for the new variable will be useful in Section 3).

The additional variable exposes the portion in the dynamics structure that is linear on the state. Specifically, introducing the matrix and vector

$$\mathbb{M} := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{T_{\text{inc}}} & 0 & 0 \\ 0 & \frac{1}{T_{\text{inc}}} & -\frac{1}{T_{\text{inf}}} & 0 \\ 0 & 0 & \frac{1}{T_{\text{inf}}} & 0 \end{pmatrix} \quad \text{and} \quad \mathbb{m} := \frac{1}{T_{\text{inf}}}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3}$$

the system of ordinary differential equations for the night cycle can be written as follows:

$$\dot{x}^i(t) = \mathbb{M}x^i(t) + \mathbb{m}\,\alpha(t)\, r^i(t) y^{ii}(t) \quad \text{for } i \in I. \tag{4}$$

Incidentally, note that, because the columns in the matrix add up to zero and $\mathcal{S}_0 + \mathcal{E}_0 + \mathcal{I}_0 + \mathcal{R}_0 = 1$, the components of both $x^i(t)$ and $y^{ii}(t)$ lie in the cube $[0,1]^{|I|}$. Also, thanks to the introduction of the additional variable (1), the relations in (4) involve, at most, the product of two variables (r and y). Keeping the model in a quadratic format, similar to the one in (4), has a tremendous impact on the numerical solution. We shall come back to this point in Proposition 3.1, after having incorporated the mobility between nodes in Section 3.

In our optimization model, the natural choice for control variable is the infection rate in the network, $(r^i(t), i \in I)$, for which an upper bound $\bar{r} \in \mathbb{R}^{|I|}$ is known. Controlling the infection rate can be seen as a proxy for various kinds of public intervention, including restricting mobility. The objective of containment measures such as social distancing and mask wearing is to eventually reduce the infection rate, see Brauner et al. (2017) and references therein. Similarly to Ferreira et al. (2021), we assume known some rule that links r's levels to non-pharmaceutical interventions. Such rules, which are region-dependent, convert the output of Robot Dance into intervention suggestions that can be implemented by the policy-maker according to the idiosyncrasies of the region under consideration.

In order to lay the groundwork for Robot Dance's optimization model, suppose for the moment there are no more constraints (this is not the case for Robot Dance). If φ and ψ denote two functions defining performance indicators for the state and control, respectively, the problem to be solved in continuous time is

$$\begin{cases} \min\limits_{r(t)\in[0,\bar{r}]} & \displaystyle\int_{\mathbb{T}_0}^{\mathbb{T}_1} \Big(\varphi(x(t)) + \psi(r(t))\Big)dt & \\ \text{s.t.} & (x(t), y(t), r(t)) & \text{satisfy (3)-(4) for a.e. } t \in [\mathbb{T}_0, \mathbb{T}_1] \\ & y(t) & \text{satisfies (1) for a.e. } t \in [\mathbb{T}_0, \mathbb{T}_1] \\ & x(\mathbb{T}_0) = x_0, r(\mathbb{T}_0) = r_0 & \text{are given initial conditions.} \end{cases} \tag{5}$$

This is an optimal control problem with state constraints for which a solution exists, by compactness of the state and control sets. Furthermore, since (4) depends linearly on the control, when the function ψ is linear, the necessary optimality conditions given by Pontryagin maximum principle ensure that any optimal control is a concatenation of bang-bang and singular arcs Clarke (2013).

The full formulation (13) solved by Robot Dance corresponds to the discretization of a control problem having a structure similar to (5), with additional constraints involving only the states or only the control. The setting still corresponds to a so-called pure control-state constrained problem, in which the only explicit relation between the state and the control appears in the dynamical system (combining the night cycle in (4) with the day dynamics described in Section 3 below).

To solve infinite-dimensional problems like (5) we adopt the approach called "discretize-and-optimize". That is, we consider a discrete approximation and use nonlinear programming methods to solve the discretized optimal control problem. The configurations of Robot Dance that promote alternating between periods of open economy with severe containment have a nonconvex quadratic objective function, for which the nonlinear programming methods provide only local minima.

Convergence of the solution of the discrete optimization problem to a reference solution of the continuous formulation depends on certain conditions related to smoothness, independence, controllability, and coercivity, stated in Dontchev et al. (2000). These conditions hold for Robot Dance.

Discretization of the differential equations amounts to choosing an approximation scheme for the left-hand side in (4). In our model, central finite differences appeared as a good compromise,

$$\dot{x}^i(t) \approx \frac{x^i_{t+1} - x^i_{t-1}}{2}. \tag{6}$$

On the right-hand side, the time dependence is a sub-index, because we are dealing with vectors, no longer with functions. Regarding the auxiliary variable defined in (1), this means that now one of its sub-indices refers to the vector component and another one to the time step. Accordingly, given x^i, y^{ii}, r^i with vectorial time components x^i_t, y^{ii}_t, r^i_t for $t \in \mathcal{T} := \{\mathbb{T}_0, \mathbb{T}_0 + 1, \ldots, \mathbb{T}_1\}$, the discretized system has the form

$$\begin{cases} x^i_{t+1} = x^i_{t-1} + 2\mathbb{M}x^i_t + 2\mathbb{m}\,\alpha_t\, r^i_t y^{ii}_t & t \in \mathcal{T}, \quad i \in I \\ y^{ii}_t = x^i_{[t,1]}x^i_{[t,3]} & t \in \mathcal{T}, \quad i \in I, \end{cases} \tag{7}$$

for suitable initial conditions x^i_{-1}, x^i_0 with components adding up to 1, and an initial reproduction rate r^i_0.

The version of Robot Dance presented here does not consider age groups. Further compartments, in particular discriminating age groups, can be treated similarly, expanding the state vector x and its companion additional variable y into the corresponding new components (keeping in mind that such extensions will increase the dimension of a problem which is already very large-scale). We refer to Silva et al. (2021b) and Silva et al. (2021a) for variants of Robot Dance with age groups that additionally introduce vaccination and testing as control variables, respectively.

3. Modeling daily commute

As shown by Fig. 2, during the day, transit between nearby nodes changes the dynamics of the susceptible and exposed compartments. An $|I| \times |I|$ mobility matrix with entries $p^{ij} \in [0,1]$ represents the percentage of inhabitants of node i traveling from node i to node j. Among several sources of mobility data that became available for COVID-19 research purposes,[1] we make use of high resolution smartphone geolocation provided by the Brazilian company InLoco.[2] The company gathers movement patterns of more than one-fifth of the inhabitants in São Paulo state, providing a social distance index with the so-called Hexagonal Hierarchical Spatial Index level 8 spatial resolution, accounting for

[1] such as https://www.google.com/covid19/mobility/ and https://covid19.apple.com/mobility.

[2] https://mapabrasileirodacovid.inloco.com.br.

hexagons of approximately 460 m edge length.[3] The index is the percentage of users that stayed home at a given day divided by the number of users in the company database for a given area. Also, from Peixoto et al. (2020) and similarly to Fig. 1, we have InLoco mobility data for São Paulo state with high resolution (approximately 1 km^2), in the form of daily origin-destination movement counts.

Diagonal elements in the mobility matrix represent the proportion of people that did not leave the district. The values p^{ii} are given by the social distance index, added to the proportion of people that moved at least 1 km, but stayed in the district. The complement of those values, considering people who left the district i, is distributed among the off-diagonal terms p^{ij}, proportionally to the trip count recorded between nodes i and j.

For the model to take into account mobility, the susceptible and exposed dynamics in (2) must incorporate the impact of commuting. First, those nodes whose value $r^i(t)$ is smaller than the natural reproduction number of the disease without intervention r_0^i are under current non-pharmacological intervention, like social distancing or even lockdown. It is, then, natural to assume that inbound travel should be discouraged, say by a factor $\zeta^i r^i(t) \in [0,1]$. For example, in the implementation we used $\zeta^i := 1/r_0^i$, so that inbound travel is decreased proportionally to the decrease in the natural reproduction number required by $r^i(t)$. Letting N^i stand for the total population of node i, this is reflected by the entries of the *effective mobility* matrix and the *effective population*, defined by

$$\mathbb{p}^{ki}(t) := \zeta^i r^i(t) p^{ki} \quad \text{and} \quad \mathbb{N}^i(t) := \sum_{k\in I} \mathbb{p}^{ki}(t) N^k \quad \text{for } i \in I, \tag{8}$$

respectively. The second modification is related to the intermediate variable defined in (1) to represent the product

$$\mathcal{S}^i(t)\mathcal{J}^i(t),$$

the amount of susceptible population of node i that gets exposed to infection. Because commuting modifies the circulation flow, with mobility the dynamics must now consider the products

$$\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\mathbb{I}^j(t),$$

the amount of susceptible population of node i that gets exposed to infection at other nodes due to commuting there. The products involve the following *effective infection* ratio:

$$\mathbb{I}^j(t) := \frac{1}{\mathbb{N}^j(t)} \sum_{k\in I} \mathbb{p}^{kj}(t)\mathcal{J}^k(t)N^k \quad \text{for } j \in I. \tag{9}$$

Putting all these ingredients together gives the following ordinary differential equations for the susceptible and exposed compartments:

$$\begin{aligned}
\dot{\mathcal{S}}^i(t) &= -\frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) - \frac{(1-\alpha(t))}{T_{\text{inf}}} \sum_{j\in I} r^j(t)\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\mathbb{I}^j(t) \\
\dot{\mathcal{E}}^i(t) &= \frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) + \frac{(1-\alpha(t))}{T_{\text{inf}}} \sum_{j\in I} r^j(t)\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\mathbb{I}^j(t) - \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t).
\end{aligned}$$

where the factor $\alpha(t)$ is used to weigh the night cycle (2) with the mobility during day time (in fractions equal to 1/3 and 2/3, for example). The other equations, defining $\dot{\mathcal{J}}^i(t)$ and $\dot{\mathcal{R}}^i(t)$, are as before.

The following diagram summarizes all the state variables and their transition dependencies.

transition	depending on
$\mathcal{S}^i \to \mathcal{S}^i$	$\mathcal{S}^i, \alpha, r, \zeta, p^{ij}, N, T_{\text{inf}}$
$\mathcal{S}^i \to \mathcal{E}^i$	same as above
$\mathcal{E}^i \to \mathcal{E}^i$	$\mathcal{E}^i, T_{\text{inc}}$
$\mathcal{E}^i \to \mathcal{J}^i$	$\mathcal{J}^i, T_{\text{inf}}$
$\mathcal{J}^i \to \mathcal{J}^i$	$\mathcal{J}^i, T_{\text{inf}}$
$\mathcal{J}^i \to \mathcal{R}^i$	T_{inf}

[3] https://h3geo.org/.

Since keeping a quadratic format as in (4) is crucial for implementation, we show below how to extend the device in (1) to suitably represent the 4th-degree monomials in the summations above.

Proposition 3.1 (Quadratic reformulation). *For all $i, j \in I$, the mobility elements $p^{ij}, \zeta^i \in [0,1]$ are given parameters, as well as the night and day cycle weights $\alpha(t) \in (0,1)$ for $t \in \mathcal{T}$.*

The SEIR model with mobility can be reformulated as follows:

$$\begin{aligned}
\dot{\mathcal{S}}^i(t) &= -\frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) - \frac{(1-\alpha(t))}{T_{\text{inf}}} \sum_{j\in I} \zeta^j p^{ij} u^j(t) z^{ij}(t) \\
\dot{\mathcal{E}}^i(t) &= \frac{\alpha(t)}{T_{\text{inf}}} r^i(t) y^{ii}(t) + \frac{(1-\alpha(t))}{T_{\text{inf}}} \sum_{j\in I} \zeta^j p^{ij} u^j(t) z^{ij}(t) - \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t) \\
\dot{\mathcal{J}}^i(t) &= \frac{1}{T_{\text{inc}}}\mathcal{E}^i(t) - \frac{1}{T_{\text{inf}}}\mathcal{J}^i(t) \\
\dot{\mathcal{R}}^i(t) &= \frac{1}{T_{\text{inf}}}\mathcal{J}^i(t).
\end{aligned} \tag{10}$$

In the dynamics (10) we introduced the additional state variables

$$y^{ik}(t) := \mathcal{S}^i(t)\mathcal{J}^k(t) = x^i_{[1]}(t)x^k_{[3]}(t) \quad \text{and} \quad z^{ij}(t) := \frac{\sum_{k\in I} p^{kj} y^{ik}(t) N^k}{\sum_{k\in I} p^{kj} N^k}, \tag{11}$$

defined for $i, j, k \in I$, as well as the following additional control variable, defined for $j \in I$:

$$u^j(t) := \big(r^j(t)\big)^2. \tag{12}$$

Proof. To prove the claim we shall rewrite the rightmost terms in the susceptible and exposed compartments as follows

$$r^j(t)\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\mathbb{I}^j(t) = \zeta^j p^{ij} u^j(t) z^{ij}(t).$$

Plugging the expression (8) for the effective mobility in both the numerator and denominator in (9) gives that

$$\begin{aligned}
\mathbb{N}^j(t) &= \zeta^j r^j(t) \sum_{k\in I} p^{kj} N^k \quad \text{and} \quad \sum_{k\in I} \mathbb{p}^{kj}(t)\mathcal{J}^k(t)N^k \\
&= \zeta^j r^j(t) \sum_{k\in I} p^{kj}\mathcal{J}^k(t) N^k.
\end{aligned}$$

In turn, the corresponding equivalent expression for the probability of local infections

$$\mathbb{I}^i(t) = \frac{1}{\sum_{k\in I} p^{ki} N^k} \sum_{k\in I} p^{ki}\mathcal{J}^k(t) N^k.$$

Shortening the denominator to $D := \sum_{k\in I} p^{ki}\mathcal{J}^k(t)N^k$ for the sake of clarity, the following algebraic manipulations give the stated result:

$$\begin{aligned}
r^j(t)\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\mathbb{I}^j(t) &= r^j(t)\mathbb{p}^{ij}(t)\mathcal{S}^i(t)\frac{1}{D}\sum_{k\in I} p^{kj}\mathcal{J}^k(t)N^k \\
&= r^j(t)\mathbb{p}^{ij}(t)\frac{1}{D}\sum_{k\in I} p^{kj}\mathcal{S}^i(t)\mathcal{J}^k(t)N^k \\
&= r^j(t)\mathbb{p}^{ij}(t)\frac{1}{D}\sum_{k\in I} p^{kj} y^{ik}(t)N^k \\
&= \zeta^j p^{ij} u^j(t) z^{ij}(t).
\end{aligned}$$

□

Notice that in (10) the equality $\dot{\mathcal{S}}^i(t) + \dot{\mathcal{E}}^i(t) + \dot{\mathcal{J}}^i(t) + \dot{\mathcal{R}}^i(t) = 0$ holds for all i and t. This ensures that the sum of these state variables is invariant in time (we omit writing the superfluous constraint $\mathcal{S}(t) + \mathcal{E}(t) + \mathcal{J}(t) + \mathcal{R}(t) = 1$).

Accordingly, the optimal control problem with mobility has the following format

$$\begin{cases} \min\limits_{r(t)\in[0,\bar{r}]} & \int_{\mathbb{T}_0}^{\mathbb{T}_1} \Big(\varphi(x(t)) + \psi(r(t))\Big)dt & \\ \text{s.t.} & (x(t), y(t), z(t), r(t), u(t)) & \text{satisfy (10) for a.e. } t \in [\mathbb{T}_0, \mathbb{T}_1] \\ & (y(t), z(t)) & \text{satisfy (11) for a.e. } t \in [\mathbb{T}_0, \mathbb{T}_1] \\ & u(t) & \text{satisfies (12) for a.e. } t \in [\mathbb{T}_0, \mathbb{T}_1] \\ & x(\mathbb{T}_0) = x_0, r(\mathbb{T}_0) = r_0 & \text{are given initial conditions;} \end{cases} \tag{13}$$

to be compared with (5).

Given an initial state, the discretized epidemiological constraints (10), derived using central finite differences as in (6)-(7), involves at most the product of two variables. The same holds for the additional state variables in (11) and (12). As a result, the feasible set in the optimization problem solved by Robot Dance is nonconvex and quadratic. This structure is often exploited by modeling languages and optimization solvers like JuMP Dunning et al. (2017) and IPOPT Wächter and Biegler (2005), the packages used by Robot Dance for modeling and optimization respectively. Keeping the quadratic structure for the optimization problem was crucial for efficiency.

4. Representing hospital capacity

In addition to the epidemiological constraints with mobility, Robot Dance considers the capacity that hospitals in the network have to attend to individuals sick with the virus. As mentioned, a critical question is the quality of the available data, mostly with respect to the number of infections. In order to somehow mitigate this issue, Robot Dance sets a probabilistic constraint for the use of ICU beds in each district. Thanks to the structure of the considered uncertainty, the chance constraint can be cast into a deterministic equivalent reformulation, see Charnes and Cooper (1962, 1963). The procedure is explained below. Before, we note that uncertainty is present in many aspects of the problem, but for scalability reasons, only the most relevant one, critical for the overall survival rate, was included in the model.

To define the capacity constraint, let v_t^i be given values for the ICU bed capacity for the ith district, for each $i \in I$ and $t \in \mathcal{T}$. Let $\texttt{icu}_t$ denote the percentage of the infected population that needs intensive care (the precise definition will be given in (21)). We consider a new variable V_t^i, representing an accumulation of sick individuals, over τ days prior to t:

$$V_t^i := \sum_{k=t-\tau}^{t} \frac{\gamma_k^i}{T_{\text{inf}}} =: \mathbb{V}_t^{i\top} x. \tag{14}$$

The value of τ, usually ranging in [7,10], see de Souza Noronha et al. (2020), corresponds to the average number of days infected individuals typically spend in the ICU. On the right-hand side in (14) the new variable is expressed as the inner product of the state variable x with a vector $\mathbb{V}_t^i \in \mathbb{R}^{|I||\mathcal{T}|}$, suitably defined. This writing emphasizes the fact that V_t^i depends linearly on the state variable of the optimization problem.

With these definitions, a deterministic constraint limiting the ICU attention would have the expression

$$\texttt{icu}_t V_t^i \le v_t^i. \tag{15}$$

In this simple version, $\texttt{icu}_t$ is considered constant, say $\mathbb{E}[\texttt{icu}_t]$, the mean value over the nodes in the network. To make a more realistic modeling of this constraint, crucial for the problem under consideration, we represent the ratio by a stochastic process that we shall approximate by a time series, Durbin and Koopman (2012).

In order to determine the parameters of the time series, we make an estimation using historical records of intensive care unit beds and new infected individuals in the region, see Section 5.2. The data was scaled to the range [0,1], using the minimum and maximum records, respectively denoted by $\underline{\texttt{icu}}$ and $\overline{\texttt{icu}}$, so that a time-series model is adjusted for the values

$$\mathcal{R}_t := \frac{\texttt{icu}_t - \underline{\texttt{icu}}}{\Delta} \quad \text{for} \quad \Delta := \overline{\texttt{icu}} - \underline{\texttt{icu}}. \tag{16}$$

By examining several indicators of model adequacy, the best fit was an autoregressive model of order p:

$$\mathcal{R}_t = c_0 + c_1 t + \sum_{j=1}^{p} \phi_j \mathcal{R}_{t-j} + \omega_t, \text{ where } \omega_t \sim \text{iid}\, \mathcal{N}(0, \sigma_\omega^2). \tag{17}$$

In the rightmost expression, the white noise ω_t is a random variable that is independent and identically distributed according to a normal distribution with zero mean and variance given by σ_ω^2.

The explicit formulation of the probabilistic expression for the constraint (15) involves some algebraic manipulations. The result below writes the recursion as a function of the callibration parameters. The expression depends on the p starting values, $\left(\texttt{icu}_{\mathbb{T}_0-j}, j = 1, \dots, p\right)$, a given data, called the tendency of the series.

Proposition 4.1 *[Explicit expression for the time series]. Given the scaled ratios (16) for $t \ge \mathbb{T}_0 + 1$ and the autoregressive model of order p in (17), consider the $p \times p$ matrix defined by*

$$A := \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ 1 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{pmatrix}$$

and let a_{ij}^k denote the entries of A^k, the kth power of the matrix A, for $k = 0, \dots, t$.

The following holds for the time series of the original data:

(i) *the expected value is*

$$\mathbb{E}[\texttt{icu}_t] := \left(1 - \sum_{j=1}^{p} a_{1j}^t\right)\underline{\texttt{icu}} + \Delta \sum_{k=0}^{t-1} a_{11}^k c_0 + \Delta \sum_{k=0}^{t-1} a_{11}^k (t-k) c_1 + \sum_{j=1}^{p} a_{1j}^t \texttt{icu}_{\mathbb{T}_0-j}.$$

(ii) $\texttt{icu}_t(\omega) = \mathbb{E}[\texttt{icu}_t] + \Delta \sum_{k=0}^{t-1} a_{11}^k \omega_{t-k}$, for $t \ge \mathbb{T}_0 + 1$.

Proof. We start by writing in vectorial form,

$$\text{the tendency is } S_0 := \begin{pmatrix} \mathcal{R}_0 \\ \mathcal{R}_{-1} \\ \dots \\ \mathcal{R}_{1-p} \end{pmatrix} \text{ and the time series is}$$

$$S_t := \begin{pmatrix} \mathcal{R}_t \\ \mathcal{R}_{t-1} \\ \dots \\ \mathcal{R}_{t-p} \end{pmatrix} \text{ for } t \ge 1.$$

Extending the intercept, drift and noise with zeros, to vectors in $\mathbb{R}^p$:

$$C_0 := \begin{pmatrix} c_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad C_1 := \begin{pmatrix} c_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \varepsilon_t := \begin{pmatrix} \omega_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and recalling the definition of the matrix A, we can write the whole series as a vector: $S_t = C_0 + C_1 t + A S_{t-1} + \varepsilon_t$. A recursive application for $t \ge 1$ results in the identity

$$S_t = \sum_{k=0}^{t-1} A^k C_0 + \sum_{k=0}^{t-1} (t-k) A^k C_1 + A^t S_0 + \sum_{k=0}^{t-1} A^k \varepsilon_{t-k}.$$

In particular, since the extended noise has mean 0, we obtain the expression for the expected value of the first component of the scaled series, which is the component of interest in our case:

$$\mathbb{E}[\mathcal{R}_t] := \sum_{k=0}^{t-1} a_{11}^k c_0 + \sum_{k=0}^{t-1} a_{11}^k (t-k) c_1 + \sum_{j=1}^{p} a_{1j}^t \mathcal{R}_{-j}, \text{ for } \mathcal{R}_t \text{ defined in (16).}$$

Scaling back using (16), yields the expression stated in item (i) for the expected value of the time series. By a similar procedure, the recursion and the identity in (16) yield the expression for the time series given in item (ii). □

Notice that the explicit expression derived for the time series in Proposition 4.1 needs computing the entries a_{ij}^k of the matrices A^k, the power of A. These calculations can be done efficiently once, before the optimization process starts. Plugging those explicit relations in the probabilistic formulation of (15), namely in

$$\mathbb{P}\left[\mathtt{icu}_t(\omega) V_t^i \le v_t^i\right] \ge 1 - \mathrm{p}, \text{ where } \mathrm{p} \in \left(0, \tfrac{1}{2}\right), \tag{18}$$

yields a linear inequality constraint; see for instance Ackooij et al. (2010); Prékopa (1995). More precisely, the explicit deterministic equivalent formulation of the chance constraint (18), written as a function of the state variable x, is

$$\mathbb{E}[\mathtt{icu}_t]\mathbb{V}_t^{i\top} x \le v_t^i - F^-(1-\mathrm{p})\sigma_\omega \Delta \mathbb{V}_t^{i\top} x \sqrt{\sum_{k=0}^{t-1} {a_{11}^k}^2} \quad \text{for } i \in I, t \in \mathcal{T}, \tag{19}$$

where F^- is the inverse cumulative distribution function of the standard Gaussian distribution.

When compared to the initial deterministic inequality (15), the probabilistic constraint (18) remains linear, with a smaller right-hand side, confirming the "robustifying" effect of the chance constraint. It should be noted that the model adopted for uncertainty sets a probability constraint for each time step separately. As such, the chance constraints disregard potential correlations between time steps for the ratio icu. A joint constraint, in which the probability holds for the whole time horizon, might be preferable, but its deterministic counterpart would make the feasible set less tractable (with more couplings and, possibly, a conic structure).

Probabilistic constraints are appealing in real-life problems, because of their straightforward and natural interpretation. The computational solution of probabilistically constrained optimization problems, exploiting structural properties in different settings, has been addressed in various forms in the stochastic programming literature, Adam and Branda (2016); Klein Haneveld and van der Vlerk (2006). The model (19) is related to the approach in Ackooij and Sagastizábal (2014). Works based on sampling average approximations dealt with by integer programming techniques are Luedtke and Ahmed (2008) and Pagnoncelli et al. (2009). Augmented Lagrangians that effectively compute p-efficient points were considered in Dentcheva and Martinez (2012); see also van Ackooij et al. (2017). For other methodologies and solvers, we refer to Pflug and Pichler (2011) and Branda (2012).

5. Robot Dance in action

Robot Dance is written in Julia Bezanson et al. (2017) using the JuMP modeling language Dunning et al. (2017). The source code, data, and scripts used in our experiments are available as open-source code in https://github.com/pjssilva/Robot-dance. We present two case studies that show the versatility of the platform for different configurations. The first experiment illustrates that it is possible to control the disease, enforcing alternation between loose and strong containment measures (playing the hammer and the dance from Pueyo (2020)). Such a strategy is meant to alleviate the detrimental effect of long lockdowns (on both the economy and the population's psychological mood). The second case, dealing with mobility and ICU capacity, highlights the interest of allowing neighboring regions to share their ICU beds.

Before reporting these results, we provide some details on the computational implementation of the model.

5.1. Objective functions and implementational details

Given initial conditions x_0 and r_0, an upper bound $\bar{r} \in \mathbb{R}^{|I|}$, and a time horizon $\mathcal{T} = \{\mathbb{T}_0, \dots, \mathbb{T}_1\}$ typically discretized using daily time steps, the mathematical optimization problem solved by Robot Dance has the form

$$\begin{cases} \min\limits_{r \in [0,\bar{r}]} & \sum\limits_{t \in \mathcal{T}} \Big(\varphi(x_t) + \psi(r_t)\Big) & \\ \text{s.t.} & (x, y, z, r, u) & \text{satisfy versions of (10)-(12) in discrete time} \\ & x & \text{satisfies (19)} \\ & r \in \mathcal{P} & \text{for some polyhedral set } \mathcal{P} \subset \mathbb{R}^{|I|\times(\mathbb{T}_1 - \mathbb{T}_0)}. \end{cases} \tag{20}$$

The set $\mathcal{P}$ is useful to bound some variables as well as to impose temporal patterns on the control variable. For example, if instead of a "hammer and dance" alternation of lockdowns and free circulation, the decision-maker prefers a policy that gradually relaxes the containment measures, constraints that prevent too abrupt changes in consecutive weeks can be incorporated in $\mathcal{P}$. Additionally, in our implementation the controls r_t^i only change every two weeks, in accordance to São Paulo's public policy on COVID-19. This is achieved by using in the code the same variable for fourteen days in a row.

In (20) the objective functions φ and ψ can be chosen by the user among several possibilities. In our runs, the state is not considered in the objective function ($\varphi \equiv 0$). As for the terms ψ, typical measures for assessing the control performance are often a combination of the following functions:

- If ensuring a maximal circulation is a priority then, given weights $w_i := \dfrac{N^i}{\sum_{j\in I} N^j}$,

  $$\psi^{\textsc{MaxCirc}}(r_t) := \sum_{i \in I} w_i(\bar{r} - r_t^i)$$

  keeps the rate as close as possible to the upper bound $\bar{r}$, understood as the one with free circulation under a "new normal", see Section 5.2 below.
- Given a minimum achievable reproduction number $\underline{r}$, the function

  $$\psi^{\textsc{Pattern}}(r_t) := -\sum_{i \in I} \delta_t^i w_i (\bar{r} - r_t^i)(\underline{r} - r_t^i)$$

  follows the sequence of confinement and relaxation induced by the scalars δ_t^i. Such weights are chosen by the decision-maker, to promote a geographical pattern to open (or close) certain areas, according to certain demographic and economic criteria.
- To prevent lockdowns (severe confinement) from being too long in a single node, Robot Dance sets

  $$\psi^{\textsc{Short}}(r_t) := -\sum_{i \in I} \delta_t^i w_i (r_t^i - r_{t-1}^i)^2.$$

- To encourage alternation of the confinement levels between specific nodes,

  $$\psi^{\textsc{Altern}}(r_t) := -\sum_{i \in I} \sum_{i \ne j \in I} \min(\delta_t^i, \delta_t^j) \min(w_i, w_j)(r_t^i - r_t^j)^2.$$

  This objective function is necessary to enforce controls of bang-bang type because, with mobility, the SEIR dynamics (10) is no longer linear on the control (the new control variable u in (12) is quadratic on r).

In our implementation, the parameters δ_t^i above were defined as the ith region population divided by the mean population in all regions, the same value for all t.

The details listed below were important for improving the performance of the numerical method.

1. The discretized ODE system uses a central difference scheme with daily time steps. In order to properly take into account the $\alpha(t)$ parameter function along a full day, we make a weighted sum of the two terms that define the derivative $\dot{S}^i$ in (10), as follows. The weight of each term corresponds to the portion of the day that is spent in the considered state. The first term, representing the evolution of the

disease during the night, is multiplied by $1/3$. The weight for the second term, representing day activities, is $2/3$. The same mechanism was applied in $\hat{\mathcal{E}}^i$.

2. As already stated, the original nonconvex, non-quadratic model is recast as a nonconvex quadratic model thanks to a thoughtful addition of auxiliary dummy variables. The choice is not unique, as it depends on which sub-expressions are selected to be represented by intermediate variables. We tried several alternatives, and taking y, z and u given in Proposition 3.1 proved to have the best numerical performance. This reformulation was the most important modification made to the original model, in terms of solution speed. Thanks to the quadratic nature of (10), we were able to instruct JuMP to compute Hessians only once. In turn, JuMP exposed this information to IPOPT, and the solver could then use the (cheap) second-order knowledge very efficiently.
3. All the variables in the SEIR model are between 0 and 1 and add up to one. By this token, the system (10) is well scaled. In the implementation, a similar scaling was employed for all the variables and constraints. In particular, the number of ICU beds in Section 4 was rescaled to represent a percentual value of the overall population.
4. To avoid extra couplings between the variables, the value of V_t^i, the number of sick individuals over τ in (14), was approximated as follows,

$$V_t^i := \sum_{k=t-\tau}^{t} \frac{\mathcal{I}_k^i}{T_{\text{inf}}} \approx \tau \frac{\mathcal{I}_t^i}{T_{\text{inf}}}$$

This approximation is reasonable because the number of sick individuals should not vary too fast in short time windows. Moreover, the resulting approximate versions of the deterministic capacity constraint (15) and its respective chance-constrained counterpart tend to be more stringent than the original ones (using (14) can be interpreted as requiring a constraint to hold on a rolling mean, while the implemented version limits each term defining the mean individually).

5. To ensure nonemptiness of the feasible set, an initial "hammer" was enforced in (20). This action corresponds to an initial short period of severe confinement that brings the percentage of the infected population down to levels that are acceptable for the hospital capacity, in all the regions.

In all of our runs, IPOPT proved to be very robust and did not need any special initial point to start the solving process. This phenomenon might be explained that the fact that once the control variables are fixed, the state variables solving the SEIR model define a unique trajectory.

With this setting, the largest nonconvex quadratic optimization instance of (20) dealt with by Robot Dance had 275,033 variables, 274,452 equality, and 7980 inequality constraints. Such problem was solved in 9 min and 34 s of wall time using IPOPT with the MA97 parallel linear solver Group (2013) on a desktop computer with an AMD Ryzen 1700X processor, that has 8 cores running at 3.4 GHz, and 64 GB of RAM.

5.2. Benchmark information, parameters, and data fitting

The state of São Paulo (SP) is partitioned into 22 health districts, represented in Robot Dance by a complex network with 22 nodes intertwined by the mobility links shown in Fig. 1. For all the experiments the optimization is performed for the whole SP state, over the period July 1, 2020 – July 28, 2021. In (19), setting $\mathrm{p} = 0.1$ ensures satisfaction of the ICU capacity constraints on any given day in 90% of the cases.

Following Wu et al. (2020), the incubation and infected times were set to $T_{\text{inc}} = 5.2$ and $T_{\text{inf}} = 2.9$, respectively. In (20) the upper bound on the control, set to $\bar{r} = 1.8$, represents the virus reproduction number when there are no stringent restrictions on circulation, but considering a "new normal" setting. An estimate equal to 2.5, from Liu et al. (2020), was suitable for the beginning of the pandemic. The smaller value employed by Robot Dance is meant to capture the decrease in the reproduction rate that was observed as a result of behavioral changes in the society (stricter hygiene habits, protective measures not associated with quarantines, like masks, etc). On the other end, we used 0.8 as the lowest achievable reproduction number under a severe lockdown. This value is in line with infection rates observed under the most stringent restrictions in countries like Germany, Italy, and Spain Hotz et al. (2020). Such values need to be revised, as new information becomes available; in particular, the basal reproduction number equal to 2.5 might be updated to represent the prevalence of more contagious strains of the virus.

The parameters of the time series described in Section 4 were calibrated from historical records of ICU beds occupancy and new cases from https://www.seade.gov.br/coronavirus/, discriminating the metropolitan area and the interior of the state, Hallal et al. (2020). The ratio was defined taking moving averages as follows:

$$\rho_t := \frac{\text{Moving Average}\left[(\text{ICU beds})_t\right]}{\text{Moving Average}\left[\sum_{\ell=t-7}^{t} (\text{Corrected new cases})_\ell\right]}. \tag{21}$$

In the denominator, $\tau = 7$ as suggested by Aziz et al. (2020), using a proxy of (14) that corrects the available data with a subnotification factor Γ and taking a mean over three days to smooth the intermittency observed in the records, after a weekend:

$$(\text{Corrected new cases})_\ell := \frac{\Gamma}{3} \sum_{k=\ell-3}^{\ell} (\text{new cases})_k. \tag{22}$$

It is well known that the official data suffers from severe subnotification in Brazil Silveira et al. (2020). In order to find a suitable value for the compensation Γ, we proceeded as follows. Using public records for the city of São Paulo we defined a trajectory for the $\mathcal{R}$-compartment until July 29th, 2020. This value was compared with the value estimated in a serological inquiry made in the whole city.[4] We observed that to arrive at the same value, it was necessary to multiply the official data by $\Gamma := 11.6$. Such an astonishing factor was confirmed experimentally in a field research Hallal et al. (2020). We used such correcting factor in all the calculations, including the estimation of the initial conditions for the runs. The value of $x_0 = (\mathcal{S}_0, \mathcal{E}_0, \mathcal{I}_0, \mathcal{R}_0)$, was determined running a variation of Robot Dance to fit the data publicly available for São Paulo state.[5]

For the calibration, the records prior to $\mathbb{T}_0$ = July 1, 2020 were considered, excluding the first days, whose standard deviation exhibits an unusual cusp. Using a training set with 75% of the remaining data, various statistical tests indicated that adjustment with a Box-Jenkins autoregressive process was suitable. We estimated models with lag, differencing and moving average parameters $(p, d, q) \in \{1, 4\} \times \{0, 1\} \times \{0, 1, 2\}$. The best fit was obtained with a pure autoregressive model of lag 2 ($p = 2$ and $d = q = 0$).

The three graphs on the left in Fig. 3 show the data available for defining the ratio in (21), approximately 100 days since May 19, 2020. The red lines correspond to the denominator (top), numerator (middle) and quotient (bottom). Each line is shadowed by its standard deviation. The blue line in the top represents the true new cases (without the correction (22) and without accumulation over seven days, scaled using a factor of 7Γ, to represent them together with the denominator in (21); and likewise for the standard deviation). The fact that new cases are incorporated in the official database by "clumps" is very clear in the graph, as well as the significant variation this data exhibits from day to day. Incidentally, the big jump by the end of July was produced by a change of the testing policy in São Paulo state. Thanks to an increased investment of the government in testing, all of a sudden numerous new cases were uncovered in the state. The new numbers were all incorporated as individuals infected the same day in the database. This change

[4] https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/17_9_2020_PPT_COLETIVAADULTO_FASE%205.pdf.

[5] https://raw.githubusercontent.com/seade-R/dados-covid-sp.

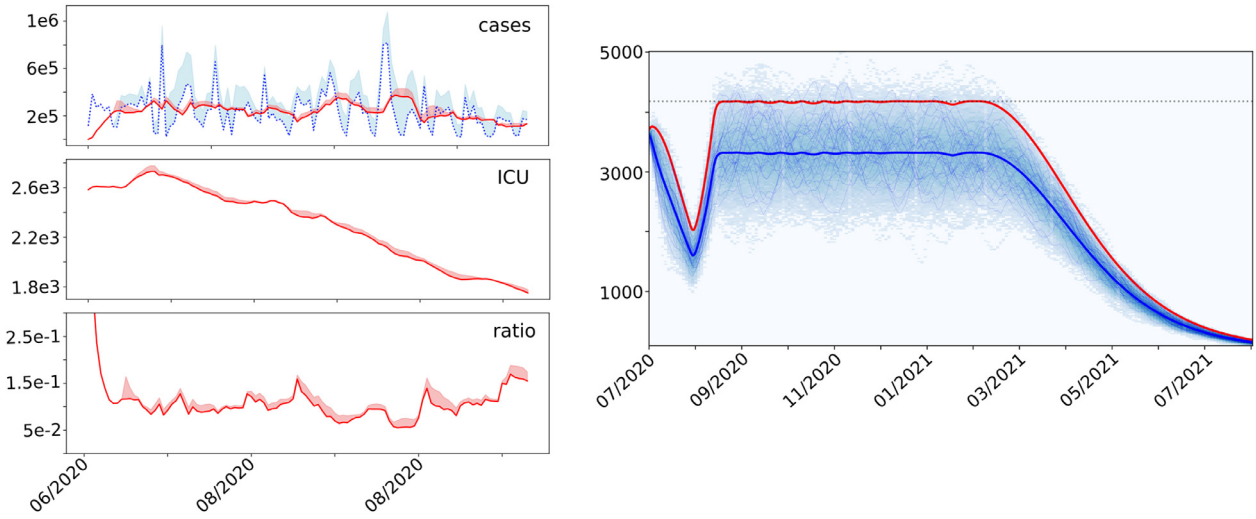

Fig. 3. On the left, moving average and standard deviation of data in (21), defining the denominator (top), numerator (middle), and ratio (bottom). The blue line in the top plot corresponds to the original data (scaled with a factor 7Γ), shadowed by its standard deviation. On the right, trajectories forecasting the use of ICU beds in São Paulo city, using a time series, as described in Section 4, calibrated with the data on the left. The dotted line represents the ICU capacity and a distancing protocol suggested by the Robot Dance code. The blue line is the expected value of the ICU demand. It sits comfortably below capacity as it was designed to sustain the demand up to the 90% quantile, represented by the red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of paradigm produced an abnormal increase in the record of that day. The right plot in Fig. 3 shows different trajectories of use of ICU beds in São Paulo city, forecasted from the time series described in Section 4, with tendency taken from the history of records.

5.3. Alternation of distancing protocols

In the network from Fig. 1 considered by Robot Dance, SP city is the state capital and has a population of 11.9 million inhabitants. The rest of the SP metropolitan area, with more than 10 million inhabitants, gathers a considerable number of small and medium-sized cities. Some of those cities host important factories and financial centers, while others are essentially dormitory towns. Because of these features, many people commute daily in the metropolitan SP area, often traveling a long time every day. In order to best capture the interaction of all the nodes in the state, Robot Dance considers its whole, with 22 health districts. For the sake of conciseness, we only report the results for the six districts that form the metropolitan SP area: SP city and five clusters of cities located to the east, west, north, southeast, and southwest of the capital, denoted by E, W, N, SE, and SW, respectively.

To illustrate different choices of the policy-maker, two different configurations were considered, one ensuring maximal circulation and another enforcing alternation of the confinement levels, every two weeks. The first option can be seen as a policy that favors a gradual relaxation of confinement measures. The second one, by contrast, switches from free circulation to lockdowns. This policy might be applicable in populations willing to endure two weeks of severe confinement, as long as the confinement is planned in advance (for example for locations where, after enjoying two weeks of fully open business, shop owners or factories can organize themselves to replenish stock or to do inventory during the lockdown periods). The respective objective functions in (20) are

$$\psi^{\text{MaxCirc}} \quad \text{and a weighed sum of all terms described in Section 5.1.} \tag{23}$$

In order to ease the interpretation of the results, which vary both in time and space, the output of Robot Dance optimization process was organized in a diagram condensing all the information. In the visualizations shown on Fig. 4, each row corresponds to one of the six health districts in the metropolitan SP area, using a color pattern for the severity of confinement. The consecutive rectangular blocks report the level (color) and duration (width) of the distancing protocols in the considered district. The black curve therein displays the dynamics of infected individuals. On the right of the diagram, a column indicates the highest percentage of the population that was sick (computed excluding the initial "hammer" described in Section 5.1).

On the left visualization in Fig. 4, distancing protocols are progressively relaxed over time in almost all regions. However, four out of six regions (E, W, N, SW), remain between "severe" and "elevated" containment over almost six months. Such measures have a very negative economic impact, mainly on non-essential activities such as bars and restaurants, and also on the educational and health system. The situation improves on the right, where the alternation mechanism was put in place. We note that the city of São Paulo is not largely affected when the intermittent mechanism is turned on. However, now the regions N, W, and E greatly benefit from the "dance" that switches from two-week periods with loose protocols (fully open in the case of the SE) followed by 14 days of more rigid containment. Such a swing clearly mitigates the economic impact of the pandemic in those regions, as activities can be planned to happen during the relaxation periods, including replenishing stocks or promoting sales to take advantage of the open business. Moreover, when comparing the black lines for the SE district in both graphs, we notice that alternation also relieves the stress on the medical staff (the health system is less pressured during the red block periods on the right).

Finally, notice that neither the gradual relaxation nor the alternation strategies succeeded in improving the situation for the SW district. In both configurations, that region endures about eight consecutive months of strict containment. This observation triggered the studies in the next subsection.

5.4. Identifying critical links and bed sharing

As shown in Fig. 4, even though the E, W, N, and SE benefit from alternation, the SW district remains under severe measures for an extended period of time. Such a phenomenon goes somehow against the expectation and was only detectable after running Robot Dance and analyzing carefully its output.

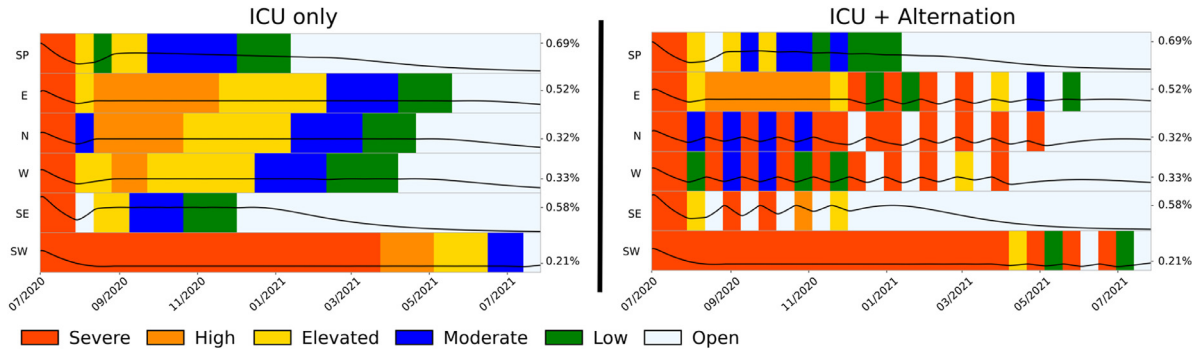

Fig. 4. Comparison of two dynamics of distancing protocols in SP metropolitan area. On the left, when ensuring maximal circulation without exceeding the ICU capacity, the distancing protocols are progressively relaxed, but remain between severe and elevated levels for about six months in four regions (E, W, N, SW). By contrast, on the right, a protocol of alternation that also respects the ICU capacity has a clear positive effect, most notably in the west suburbs (W), during the months 09 to 12 in 2020. The column on the right of each diagram shows the maximal percentage of sick people attended by the health system, after an initial lockdown phase used to reduce the number of infected individuals, if their proportion was too high for the hospital capacity in the district. The percentages on the right vertical axes represent the highest infection ratio estimated in each population after the initial lockdown brings the disease to acceptable levels. Controls r_t^i lie in the interval $[0.8, 1.8]$, that was split into five equally spaced subintervals, each one having one color associated, from red to green. White corresponds to $r_t^i = 1.8$, a fully open society in a "new normal" setting. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In order to understand the interplay in what appears to be a problematic link, we report results only for the SW district and SP city. Robot Dance new runs still consider all the 22 regions in the state with the objective function in (23) and probability level $p = 0.1$ in (19). The different configurations are obtained by modifying some aspects of mobility in the link between the SW district and SP city, or of the distribution of hospital resources in the metropolitan area.

Recall that, to compute the output in Section 5.3, each region is assumed to treat infected cases using only their own hospitals. It appears that this approach is detrimental for the entire area, because the number of ICU beds in the SW is very low (less than 90 for 1 million population). By "detrimental" we mean that all the six metropolitan districts are forced into containment measures that could be lighter if patients could be transferred to SP city.

We now explain the case study. Having identified the SW district as a critical node thanks to the output in Fig. 4, the decision-maker needs to evaluate the effect of different manners to address the problematic issue. To assess if it would be worth to simply forbid all commuting between SW and SP city, or rather transfer some patients to the state capital, the following three alternatives were considered by Robot Dance:

1. the original configuration from Fig. 4, ensuring maximal circulation without exceeding the local ICU capacity in the whole state;
2. the same, but eliminating the link SW-SP city, no commuting is allowed between those two nodes (there is still mobility in the whole state); and
3. as in the first configuration, but implementing a pool of beds in which the five districts in the metropolitan area have access to a small fraction of SP city's hospital capacity (very large, over 3400 beds).

The output of Robot Dance for SW and SP city for these three runs is shown on the right of Fig. 5. Therein, the three alternatives are respectively referred to as "with critical link and no ICU sharing" (top), "without critical link and no ICU sharing" (middle), and "with critical link and ICU sharing" (bottom). In particular, the top diagram reproduces the left one in Fig. 4, extracting the rows for SP city and the SW district.

In the top right visualization in Fig. 5, the rightmost percentages, of the population that could be attended at local ICU units, reveal the very poor hospital capacity in the SW. Even under the most severe measures, this district could not afford to attend more than 0.17% inhabitants while SP city dealt with a much higher value, 0.47% of its own population. This feature triggers more than eight months of severe lockdown for the SW district, until the end of March 2021, as shown in the top diagram (otherwise the SW's health system would collapse). Such a severe protocol is not realistic from an economic point of view, especially considering that the nearby SP city, less than 80 km away, fully opens its activities at the beginning of January 2021. The second alternative, that plainly forbids commuting between the SW and SP city. reported in the middle graph, confirms the crucial role that the SW region has in the network. Without the flow of workers circulating in the SW-SP city link, the latter would be fully open one month earlier, from December 2020. Moreover, in the configuration, the attention to sick individuals in SP city improves slightly (increases to 0.55%, from 0.47% in the top). The SW region can also relax a bit the strength of its containment from November 2020, facing one and a half months of severe measures, followed by almost five months of elevated measures before starting the alternation of hammer (red) and dance (white). The SW capacity to attend surges in its own hospitals still remains low, 0.19% of the local population.

Clearly, it is impractical to completely cut the circulation to the capital for long periods, as advocated by the second alternative. The third policy, by contrast, is much more effective. The percentages on the bottom right visualization in Fig. 5 speak by themselves: SP city maintains its high values to respond to hospital surges (0.52%) while in the SW the percentage jumps from 0.18% to 0.59%.

In order to let the spare capacity in SP city absorb part of the patients from the suburbs, Robot Dance was run with the constraints (19) grouped in a single pool of ICU beds for the entire SP metropolitan area. The positive impact of this measure is clear in the visualization "with critical link and ICU sharing" in the bottom right of Fig. 5. Notice that, while containment measures do not change substantially in SP city (the three less severe containment measures only alternate more), for the SW district the situation changes significantly. Now the duration of containment measures drops to less than half, and SW is fully open from early January 2021. Moreover, prior to that date, severe distancing measures alternate with more relaxed protocols, thereby maintaining alive the economic and social activities in the district.

The fact of sharing ICU beds is beneficial for the whole metropolitan area, not only the critical SW node. Without bed sharing, the right visualization in Fig. 4 shows only SP city and SE district fully open from January 2021 on. With the pool of beds, by contrast, similarly to what is reported in Fig. 5 bottom right for the SW, all the six districts in the metropolitan area are fully open from January 2021 (we do not include the diagram for the whole state, to save space).

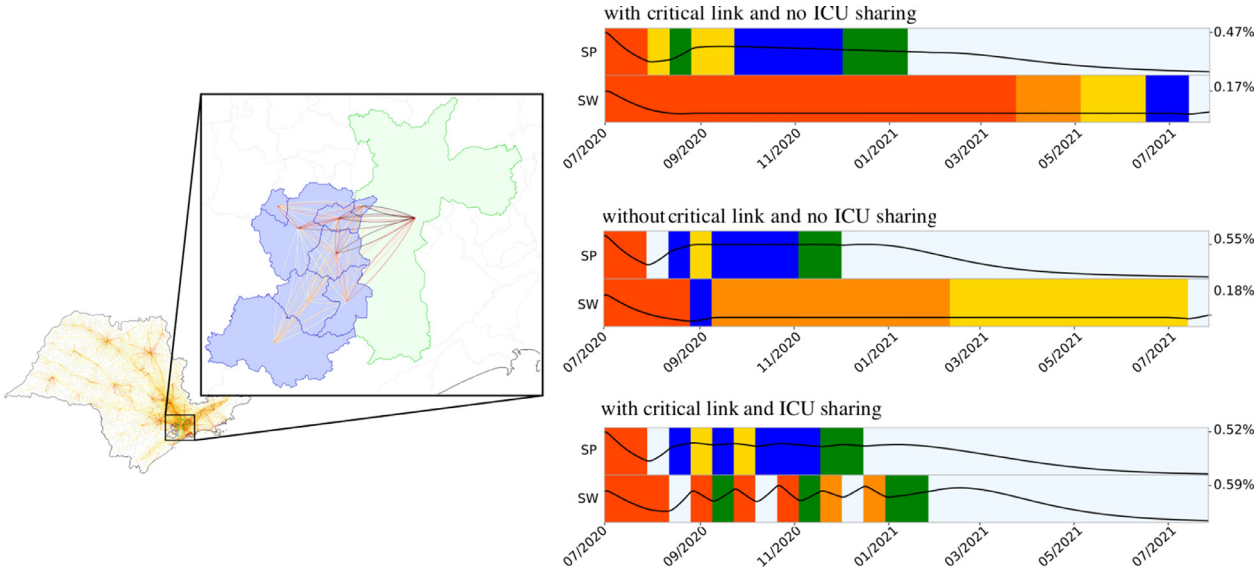

Fig. 5. The left map zooms in those connections between the SP city and the SW district. The link SW-SP city is critical for the network, but forbidding commuting in the link is not the solution; it is preferable to put in place a mechanism to transfer patients to SP city. This can be seen by comparing the two bottom visualizations. The middle graph shows that cutting the link circulation has practically no effect on the situation neither in SW nor in SP city. By contrast, the policy reported in the bottom graph, that keeps the link and implements a pool of ICU beds, improves significantly the overall state for SW without affecting much the situation in SP city.

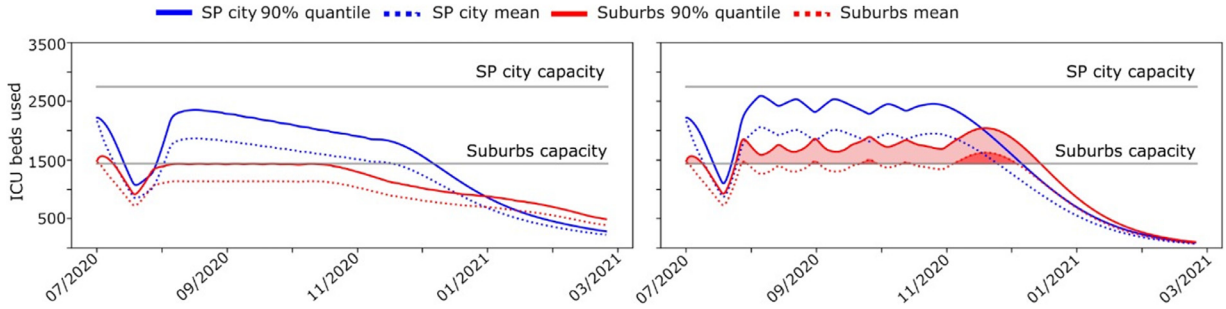

Fig. 6. Having access to a pool of beds is beneficial for the whole metropolitan area of São Paulo. For the top and bottom policies in Fig. 5, here represented in the left and right plots, we report the mean use of ICU beds (dotted line) and its 90% quantile (continuous line). Horizontal lines indicate the maximum capacity in SP city and suburbs (the other five metropolitan districts together). The possibility to transfer patients from the suburbs to SP city creates a reserve of additional capacity for the former. This loss does not affect much the capacity of SP city to attend to its own patients. On the right, the reserved beds for the suburbs, representing about 300 additional beds, are plotted in the light shadowed area. The number of transfers, about 100 and reported in the dark area, mostly occurr between November and December 2021.

Our final Fig. 6 reports on the dynamic use of the pool of beds in the metropolitan region, separately for SP city (that loses capacity) and the other five districts together (that gain capacity). The dotted, continuous, and horizontal lines therein show, respectively, the number of ICU beds used in average, its 90% quantile corresponding to the chance constraint (19) with $p = 0.1$, and the capacity. The difference between the left and right graphs is on the policies, respectively obtained by Robot Dance with and without a pool of beds (the top and bottom graphs in Fig. 5). We see that SP city large availability of ICU beds is never reached. By contrast, the suburbs are close to saturation, the main culprit being the critical SW district. Without bed sharing, the suburbs' left curve remains stuck against the horizontal line (hospitals functioning at maximum capacity) for many months.

In Fig. 6, the area shadowed in light red on the right plot represents the capacity gained by the suburbs (about 300 beds). This additional reserve of beds is actually employed by the suburbs, particularly the SW district, at those points where the dotted line (mean use) gets higher than the horizontal line of local capacity. Those points are represented by the darker shadowed areas in Fig. 6 (about 100 beds). Note in passing that the pool induces an effect similar to the alternation mechanism discussed in Section 5.3 (this phenomenon is perceptible on the right plots in Fig. 5, where the lines go up and down, making peaks).

To finish, recall that Robot Dance applies an initial "hammer" to guarantee nonemptiness of the feasible set in (20). This is noticeable in the graphs in Fig. 4, where the first red column in all visualizations is meant to be disregarded. The hammer also materializes in the initial gully shape of trajectories forecasting the ICU bed use in the right of Fig. 3, and in the curves in Fig. 6.

6. Final comments

Social distancing measures have proven effective to contain the spread of COVID-19, but there is no universal recipe that can be applied throughout the globe. In each country, and for each region in a country, idiosyncrasies that depend on social and economic factors, individual to each place, must be taken into account. This issue, particularly important in Brazil, has increased relevance when the economic infrastructure is concentrated in a megalopolis, forcing a large portion of the population to commute daily to work.

Robot Dance was formulated keeping these important considerations in mind. Its open-source code has the versatility required to analyze and compare several configurations without much ado. Thanks to its SEIR representation with network mobility, Robot Dance is able to consider explicit interventions in specific nodes. Nodes can be neighborhoods in a city, or cities in a state, or states in a country. Section 5.4, comparing the model with and without a mechanism for sharing ICU beds in the SP metropolitan area, while analyzing the state as a whole, illustrates well this feature.

The problem solved by Robot Dance takes into account effects that are not directly observable, or intuitive, at first sight. This is explained by the mathematical optimization setting, that goes beyond the usual simulations typical in the epidemiological area. Section 5.4 is also a good example of this situation: some crossed spatio-temporal effects make critical the link between the state capital and one much smaller suburban district; but simply eliminating the link, removing the circulation between the two nodes, does not resolve the issue. A much better policy is to allow for the transfer of patients. Incidentally, as pointed out by one reviewer, Robot Dance's output with bed sharing gives an indication of the importance of properly sizing hospital capacity in each geographic region rather than in aggregate.

In Silva et al. (2021b) and Silva et al. (2021a), respectively, Robot Dance was used to determine how to best deploy vaccination and testing campaigns. Optimization is particularly useful when resources are scarce, and such is the case in many countries regarding COVID-19, whether it is due to an insufficient number of vaccines or testing material, or to lack of technical staff. Our studies are in agreement with the proposal in Kissler et al. (2020), that calls for an urgent need of projecting how the coronavirus will unfold in the coming years, to prevent recurrent wintertime outbreaks. The computational optimization framework of Robot Dance could help in this respect, as the open-source code provides an output that is reproducible and public.

Regarding the predictive power of the platform, a word of caution is in order. The results obtained with a numerical tool like Robot Dance should not be seen as precise and accurate, they are prone to limitations. First, the quality of the input data impairs the calibration of the initial conditions of the differential equation and affects the estimation of the whole epidemic situation. Second, a mathematical model fails to grasp several complex phenomena in the pandemic evolution. Changes in society behavior are represented, neither in the reaction to the disease evolution, nor to lockdowns or measures restricting the free circulation, that alter the mobility patterns of the population. Introducing those changes involves modelling the basal reproduction number and the mobility matrix varying with time, which in turn means re-calibrating the involved parameters along the optimization process. Those modifications require handling uncertainty in a multistage setting, as new data is revealed. Third, population is not homogeneous, but composed of different groups, professions, levels of income, and other characteristics that are not covered in the current model. Projections computed by Robot Dance are indicative, rather than assertive. They are approximations of the future evolution of the pandemic, keeping in mind that the accuracy decreases as the considered time gets farther in the future.

For all these reasons, Robot Dance needs to be constantly re-run with new data. This is not a difficult task since, the careful computational implementation, that exploits the strength of JuMP as modelling language and IPOPT as optimization solver, provides results for an area with over forty million inhabitants, like the state of São Paulo, in less than ten minutes. In spite of the aforementioned limitations, tools like Robot Dance support the decision-making of public administrators and politicians with insight and a better understanding of the effect of different types of interventions on the pandemic.

Finally, to keep the platform aligned with the constant new information that is learned about COVID-19, Robot Dance's short-term projections should be compared to the actual development of the disease, introducing model enhancements, new parameters, and technologies like an effective medicine or a vaccine as they become available.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the reviewers and editor for beneficial comments. Research supported by CNPq-Brazil (303552/2017-4, 304301/2019-1, 301778/2017-5), the São Paulo Research Foundation (FAPESP) – Brazil (2013/07375-0, 2016/18445-7, 2016/04190-7, 2018/24293-0), Royal Society, UK-grant No. NAF\R1\180236, and Serrapilheira Institute (Grant No. Serra-1709-16124).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejco.2022.100025

References

Ackooij, W.v., Berge, V., de Oliveira, W., Sagastizbal, C., 2017. Probabilistic optimization via approximate p-efficient points and bundle methods. Comput. Oper. Res. 77, 177–193.

Ackooij, W.v., Henrion, R., Müller, A., Zorgati, R., 2010. On probabilistic constraints induced by rectangular sets and multivariate normal distributions. Math. Method. Oper. Res. 71 (3), 535–549.

Ackooij, W.v., Sagastizábal, C., 2014. Constrained bundle methods for upper inexact oracles with application to joint chance constrained energy problems. SIAM J. Optim. 24 (2), 733–765.

Acuña-Zegarra, M.A., Santana-Cibrian, M., Velasco-Hernandez, J.X., 2020. Modeling behavioral change and COVID-19 containment in a trade-off between lockdown and compliance. Math. Biosci. 108370.

Adam, L., Branda, M., 2016. Nonlinear chance constrained problems: optimality conditions, regularization and solvers. J. Optim. Theory Appl. 170 (2), 419–436.

Aguas, R., Corder, R.M., King, J.G., Goncalves, G., Ferreira, M.U., Gomes, M.G.M., 2020. Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. medRxiv.

Arenas, A., Cota, W., Gomez-Gardenes, J., Gómez, S., Granell, C., Matamalas, J.T., Soriano-Panos, D., Steinegger, B., 2020. A mathematical model for the spatiotemporal epidemic spreading of COVID 19. medRxiv.

Aziz, S., Arabi, Y.M., Alhazzani, W., Evans, L., Citerio, G., Fischkoff, K., Salluh, J., Meyfroidt, G., Alshamsi, F., Oczkowski, S., Azoulay, E., Price, A., Burry, L., Dzierba, A., Benintende, A., Morgan, J., Grasselli, G., Rhodes, A., Møller, M.H., Chu, L., Schwedhelm, S., Lowe, J.J., Bin, D., Christian, M.D., 2020. Managing ICU surge during the COVID-19 crisis: rapid guidelines. Intensive Care Med. 46 (7), 1303–1325.

Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: a fresh approach to numerical computing. SIAM Rev. 59 (1), 65–98.

Bonnans, J.F., Gianatti, J., 2020. Optimal control techniques based on infection age for the study of the COVID-19 epidemic. Math. Model. Nat. Phenom. 15–48 doi:10.1051/mmnp/2020035.

Branda, M., 2012. Stochastic programming problems with generalized integrated chance constraints. Optimization 61 (8), 949–968.

Brauer, F., 2008. Compartmental Models in Epidemiology. In: Mathematical epidemiology. Springer, pp. 19–79.

Brauner, J.M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, J., Gaveniak, T., Stephenson, A.B., Leech, G., Altman, G., Mikulik, V., Norman, A.J., Monrad, J.T., Besiroglu, T., Ge, H., Hartwick, M.A., Teh, Y.W., Chindelevitch, L., Gal, Y., Kulveit, J., Inferring the effectiveness of government interventions against COVID-19. Science 371(6531). 33323424.

Candido, D.D.S., Watts, A., Abade, L., Kraemer, M.U., Pybus, O.G., Croda, J., de Oliveira, W., Khan, K., Sabino, E.C., Faria, N.R., 2020. Routes for COVID-19 importation in Brazil. J. Travel. Med. 27 (3).

Candido, D.S., Claro, I.M., de Jesus, J.G., Souza, W.M., Moreira, F.R.R., Dellicour, S., Mellan, T.A., du Plessis, L., Pereira, R.H.M., Sales, F.C.S., Manuli, E.R., Thézé, J., Almeida, L., Menezes, M.T., Voloch, C.M., Fumagalli, M.J., Coletti, T.M., da Silva, C.A.M., Ramundo, M.S., Amorim, M.R., Hoeltgebaum, H.H., Mishra, S., Gill, M.S., Carvalho, L.M., Buss, L.F., Prete, C.A., Ashworth, J., Nakaya, H.I., Peixoto, P.S., Brady, O.J., Nicholls, S.M., Tanuri, A., Rossi, Á. D., Braga, C.K., Gerber, A.L., de C. Guimarães, A.P., Gaburo, N., Alencar, C.S., Ferreira, A.C., Lima, C.X., Levi, J.E., Granato, C., Ferreira, G.M., Francisco, R.S., Granja, F., Garcia, M.T., Moretti, M.L., Perroud, M.W., Castiñeiras, T.M.P.P., Lazari, C.S., Hill, S.C., de Souza Santos, A.A., Simeoni, C.L., Forato, J., Sposito, A.C., Schreiber, A.Z., Santos, M.N.N., de Sá, C.Z., Souza, R.P., Resende-Moreira, L.C., Teixeira, M.M., Hubner, J., Leme, P.A.F., Moreira, R.G., Nogueira, M.L., Ferguson, N.M., Costa, S.F., Proenca-Modena, J.L., Vasconcelos, A.T.R., Bhatt, S., Lemey, P., Wu, C.-H., Rambaut, A., Loman, N.J., Aguiar, R.S., Pybus, O.G., Sabino, E.C., Faria, N.R., 2020. Evolution and epidemic spread of SARS-CoV-2 in Brazil. Science 369 (6508), 1255–1260. doi:10.1126/science.abd2161.

Carmo, R.F., Nunes, B.E., Machado, M.F., Armstrong, A.C., Souza, C.D., 2020. Expansion of COVID-19 within Brazil: the importance of highways. J. Travel Med. 27 (5).
Charnes, A., Cooper, W., 1959. Chance-constrained programming. Manag. Sci. 6, 73–79.
Charnes, A., Cooper, W., 1962. Chance constraints and normal deviates. J. Amer. Statist. Assoc. 57, 134–148.
Charnes, A., Cooper, W., 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. Oper. Res. 11, 18–39.
Chen, Y.-C., Lu, P.-E., Chang, C.-S., Liu, T.-H., 2020. A Time-Dependent SIR Model for COVID-19 With Undetectable Infected Persons. in IEEE Transactions on Network Science and Engineering 11 (4), 3279–3294. doi:10.1109/TNSE.2020.3024723.
Clarke, F., 2013. Functional Analysis, Calculus of Variations and Optimal Control. Springer London doi:10.1007/978-1-4471-4820-3.
Dentcheva, D., 2006. Optimization Models with Probabilistic Constraints. In: Probabilistic and Randomized Methods for Design under Uncertainty. Springer-Verlag, pp. 49–97. doi:10.1007/1-84628-095-8_2.
Dentcheva, D., Martinez, G., 2012. Regularization methods for optimization problems with probabilistic constraints. Math. Program. 138 (1–2), 223–251. doi:10.1007/s10107-012-0539-6.
Dontchev, A.L., Hager, W.W., Malanowski, K., 2000. Error bounds for Euler approximation of a state and control constrained optimal control problem. Numer. Funct. Anal. Optim. 21 (5–6), 653–682. doi:10.1080/01630560008816979.
Dunning, I., Huchette, J., Lubin, M.,2017. JuMP: a modeling language for mathematical optimization 59 (2), 295–320. doi:https://doi.org/10.1137/15M1020575
Duque, D., Morton, D.P., Singh, B., Du, Z., Pasco, R., Meyers, L.A., 2020. Timing social distancing to avert unmanageable COVID-19 hospital surges. Proc. Natl. Acad. Sciences 117 (33).
Durbin, J., Koopman, S.J., 2012. Time Series Analysis by State Space Methods. Oxford University Press doi:10.1093/acprof:oso/9780199641178.001.0001.
Ferreira, C.P., Marcondes, D., Melo, M.P., Oliva, S.M., Peixoto, C.M., Peixoto, P.S., 2021. A snapshot of a pandemic: the interplay between social isolation and COVID-19 dynamics in Brazil. Available at SSRN 3837638.
Group, C. M., 2013. HSL(2013). A collection of Fortran codes for large scale scientific computation. http://www.hsl.rl.ac.uk.
Hallal, P.C., Horta, B.L., Barros, A.J.D., Dellagostin, O.A., Hartwig, F.P., Pellanda, L.C., Struchiner, C.J., Burattini, M.N., da Silveira, M.F., Menezes, A.M.B., Barros, F.C., Victora, C.G., 2020. Evolução da prevalência de infecção por COVID-19 no Rio Grande do Sul, Brasil: inquéritos sorológicos seriados. Ciência & Saúde Coletiva 25 (suppl 1), 2395–2401. doi:10.1590/1413-81232020256.1.09632020.
Hotz, T., Glock, M., Heyder, S., Semper, S., Böhle, A., Krämer, A., 2020. Monitoring the spread of COVID-19 by estimating reproduction numbers over time. arXiv preprint arXiv:2004.08557.
Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H., Lipsitch, M., 2020. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368 (6493), 860–868. doi:10.1126/science.abb5793.
Klein Haneveld, W., van der Vlerk, M., 2006. Integrated chance constraints: reduced forms and an algorithm. Comput. Manag. Sci. 3, 245.
Liu, Y., Gayle, A.A., Wilder-Smith, A., Rocklöv, J., 2020. The reproductive number of COVID-19 is higher compared to coronavirus. J. Travel Med. 27 (2). doi:10.1093/jtm/taaa021.
Luedtke, J., Ahmed, S., 2008. A sample approximation approach for optimization with probabilistic constraints. SIAM J. Optim. 19, 674.
Maurer, H., do Rosario de Pinho, M., 2016. Optimal control of epidemiological SEIR models with l_1-objectives and control-state constraints. Pacif. J. Optim. 12 (2), 415–436.
Pagnoncelli, B., Ahmed, S., Shapiro, A., 2009. Sample average approximation method for chance constrained programming: theory and applications. J. Optim. Theory Appl. 142, 399.
Peixoto, P.S., Marcondes, D., Peixoto, C., Oliva, S.M., 2020. Modeling future spread of infections via mobile geolocation data and population dynamics. an application to COVID-19 in Brazil. PLoS ONE 15 (7), e0235732.
Pflug, G.C., Pichler, A., 2011. Approximations for Probability Distributions and Stochastic Optimization Problems. In: Stochastic Optimization Methods in Finance and Energy. Springer New York, pp. 343–387.
Prékopa, A., 1995. Stochastic programming. Mathematics and its Applications, 324. Kluwer Academic Publishers Group, Dordrecht.
Pueyo, T., 2020. Coronavirus: The hammer and the dance. https://medium.com@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56.
Silva, P.J.S., Pereira, T., Sagastizábal, C., M, C.M., Nonato, L.G., Struchiner, C.J., 2021. Smart testing and critical care bed sharing for COVID-19 control. PLOS ONE 16 (10).
Silva, P.J.S., Sagastizábal, C., Nonato, L.G., Struchiner, C.J., Pereira, T., 2021. Optimized delay of the second COVID-19 vaccine dose reduces ICU admissions. Proc. Natl. Acad. Sci. 118 (35).
Silveira, M., Barros, A., Horta, B., Pellanda, L., Victora, G., Dellagostin, O., Struchiner, C., Burattini, M., Valim, A., Berlezi, E., et al., 2020. Repeated population-based surveys of antibodies against SARS-CoV-2 in southern Brazil. medRxiv.
de Souza Noronha, K.V.M., Guedes, G.R., Turra, C.M., Andrade, M.V., Botega, L., Nogueira, D., Calazans, J.A., Carvalho, L., Servo, L., Ferreira, M.F., 2020. Pandemia por COVID-19 no Brasil: análise da demanda e da oferta de leitos hospitalares e equipamentos de ventilação assistida segundo diferentes cenários. Cadernos de Saúde Pública 36 (6). doi:10.1590/0102-311x00115320.
Wächter, A., Biegler, L.T., 2005. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Program. 106 (1), 25–57. doi:10.1007/s10107-004-0559-y.
Wu, J.T., Leung, K., Leung, G.M., 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 395 (10225), 689–697.