

A MULTILAYER NETWORK MODEL IMPLEMENTATION FOR COVID-19

JUAN G. CALVO*, FABIO SANCHEZ*, LUIS A. BARBOZA*, YURY E. GARCÍA†, AND PAOLA VÁSQUEZ†

Abstract. We present a numerical implementation for a multilayer network used to model the transmission of Covid-19 or other diseases with a similar transmission mechanism. The model incorporates different contact types between individuals (*household*, *social contacts*, and *strangers*), which allows flexibility compared to standard SIR type models. The algorithm described in this paper is a simplification of the model used to give public health authorities an additional tool for the decision-making process in Costa Rica, by simulating extensive possible scenarios and projections.

Key words. multilayer network, Covid-19, epidemic models, pandemic

AMS subject classifications. 93A16, 92C60

1. Introduction. Throughout the Covid-19 pandemic, mathematical and statistical models have been used as a way to forecast or analyze the behavior of the disease in various parts of the world [1, 2, 3, 4, 5]. Different mathematical tools have been implemented to understand the disease transmission and to plan for strategic sanitary measures to minimize the burden on the public health system [1, 5, 6]. Deterministic models have been widely used to model infectious diseases in the past [7, 8, 9, 10, 11, 12]. However, due to human mobility, unpredictable human behavior and other factors, adaptable models are better suited to evolve with the ever-changing behavior of the population during the pandemic as sanitary measures are implemented in a particular region of interest.

In this paper, we consider an adaptive multilayer network, capable of managing large populations efficiently; see [13, 14]. A probability of infection is computed on a daily basis for each node independently, permitting a parallel implementation for real-life applications. We count daily interactions per individual and determine if the virus is transmitted depending on the attributes of each interaction.

One of the main advantages of a network model is its flexibility when taking into account individual and social behavior, that allows us to mimic sanitary measures and evaluate possible impacts on the population. As the pandemic progresses, it is possible to incorporate early or ongoing behavioral changes such as reductions in community interactions, mask use, mobility restrictions, network contacts, and lockdowns. We include a Matlab app with the algorithm implementation that can be found in [15], where several attributes and parameters can be modified¹.

The model we present in this paper is a simplification of the model that has been developed and taken into consideration by the health authorities in Costa Rica. It includes the entire population (approximately five million individuals) and the 81² municipalities of the country. The model also considers different scenarios with

*Universidad de Costa Rica, Centro de Investigación en Matemática Pura y Aplicada - Escuela de Matemática, San José, Costa Rica (juan.calvo@ucr.ac.cr, fabio.sanchez@ucr.ac.cr, luisalberto.barboza@ucr.ac.cr).

†Universidad de Costa Rica, Centro de Investigación en Matemática Pura y Aplicada, San José, Costa Rica (paola.vasquez@ucr.ac.cr, yury.garciapuerta@ucr.ac.cr).

¹A full open code will be available as well.

²There are 82 municipalities in the country. However, the 82nd was established in 2018 and there is no historical census data in order to model mobility.

gradual re-openings, mobility restrictions, different percentages for mask use, as well as short and long-term projections of the disease. Results were delivered to the public health authorities as part of the decision-making process [16, 17, 18, 19]. The model takes into account local data, public health policies pertinent to the country and more specific details than the model described in this work.

The rest of the paper is organized as follows. In Section 2, we describe the algorithm. A pseudo-code is presented in Section 3 with a discussion of the implementation. In Section 4, we present two toy examples, in order to show the behaviour of the model. Finally, we include some conclusions and future work in Section 5.

2. Preliminaries. Consider a constant population that lives in a fixed number of counties. Mathematically, we consider a graph (a set of nodes that are connected by edges) where each node represents one individual. Interactions are classified into three *layers*, which represent a different type of contact with particular characteristics that account for the virus transmission. These layers are: (1) a *household network* (people that live in the same house), (2) a *social network* (known contacts such as friends and colleagues), and (3) a *sporadic network* (strangers that you may encounter in short periods of time when you visit random locations).

An edge between two nodes represents a feasible interaction between two individuals at a particular time, with specific attributes for its interaction depending on the layer. These layers are randomly generated. Layers 1 and 2 are fixed for each simulation and layer 3 can change for each time step, since usually a node has no control on sporadic encounters.

For the first layer, it is assumed that we know the total number of households per county, and the average number of individuals per household per county; the information for Costa Rica is available at [20]. We then group all the individuals into families. The number of members per household is chosen by using a Poisson distribution with mean equal to the average number of individuals per household per county. We assume total connectivity for each family and no edges between different families. It is possible to randomly assign an age group to each family member if data is available, in order to include differences in interactions accordingly; see further details in Section 3.1.

For the second and third layers, choosing the contacts of a node requires some assumptions. It is natural to expect a different number of interactions depending on age, location, density, socioeconomic factors and more. In our implementation, for every node we need to: (1) define the degree of the node (number of contacts) per layer, given by a uniform distribution on an interval with endpoints equal to the minimum and maximum number of allowed contacts per county per layer, and (2) choose randomly its contacts from different counties based on a given connectivity distribution between counties; see Section 3.3.

Each node on the graph includes several attributes. We classify them in two classes: (1) fixed parameters, such as county, household members, age group, degree and graph connectivity for layer 2, and (2) variable attributes such as connectivity for layer 3, epidemiological state, number of days at current epidemiological state, number of interactions with social contacts per day, mask use and self-care behaviour. The former are defined at the beginning of each simulation, and the latter can take random values every time step accordingly.

For simplicity, the propagation in our model is based on an SIR type model with seven compartment states defined by the following epidemiological variables: susceptible (S), exposed (E), diagnosed or observed (O), undiagnosed or not observed (U),

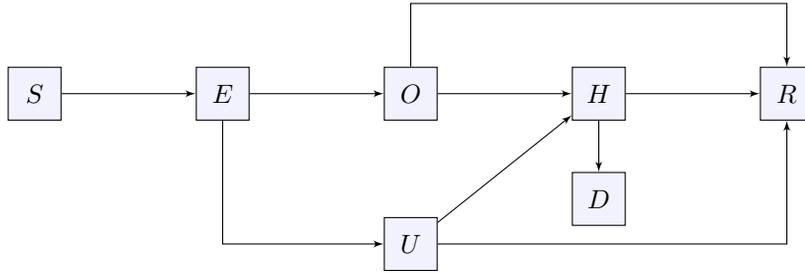


Fig. 2.1: Model compartment states and transitions.

hospitalized (H), recovered (R), and dead (D) states; see Figure 2.1. Modifications with additional states and transitions are straightforward to include.

Once the multilayer network is defined, we model the transmission of the virus by computing probabilities of infection p_i for interactions between infectious and susceptible nodes; see Section 3.4. For a given initial set of exposed nodes, we consider their contacts via the different layers; see Figure 2.2a. Depending on p_i , adjacent nodes of infectious nodes can become exposed; see Figure 2.2b. We then increase the network by including new exposed nodes and their contacts as shown in Figure 2.2c, continuing the process on a daily basis; see Figure 2.2d.

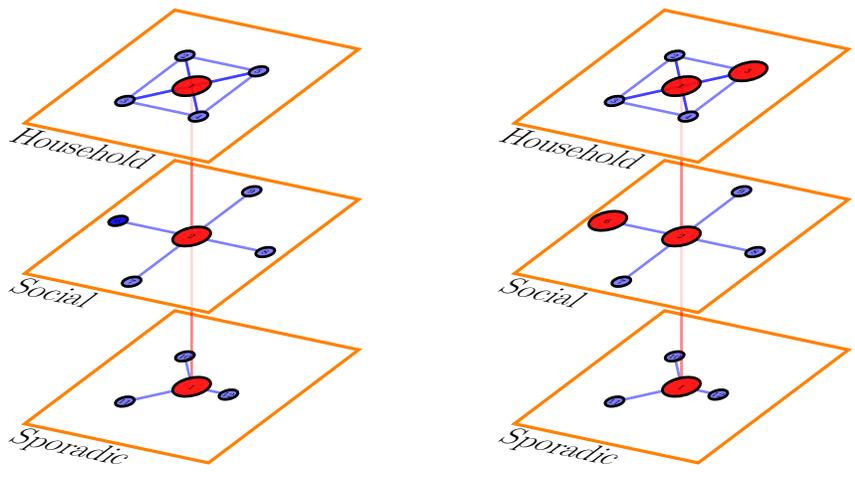
REMARK 2.1. The degree of each node in layer 2 is fixed for each simulation. Nevertheless, on each time step a subset of edges is selected because only a small number of daily contacts is expected to occur. As mentioned before, layer 3 changes daily.

REMARK 2.2. If the population has a large number of individuals, it is not recommended to create the whole graph initially due to execution times. Instead, at every time step we add new nodes and their edges on the graph in order to achieve faster running times; see Figure 2.2 and Section 3.3 for further details.

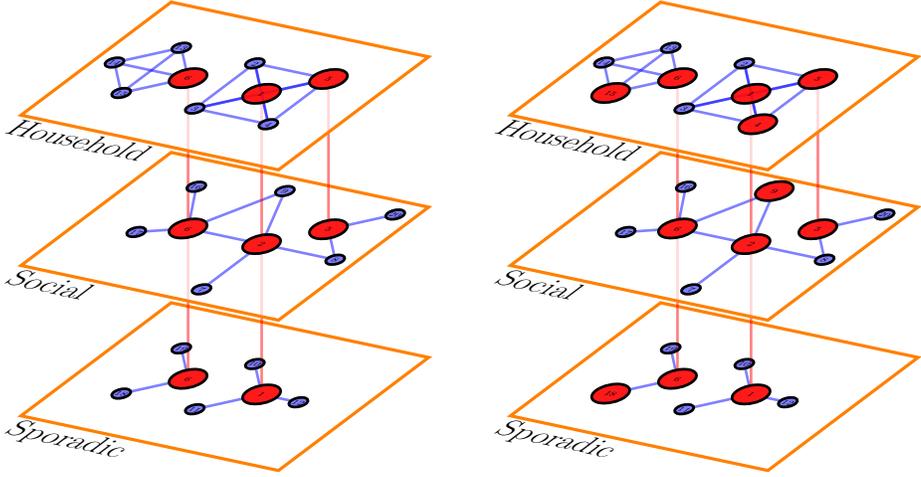
3. Implementation. In this section we present a pseudo-code for the modeling of Covid-19 as described in Section 2; see Algorithm 1. We describe each step in the following sections in a general and simplified setting.

3.1. Creation of household network. In this implementation we assume that the following data is available: total number of households per county, average number of individuals per household per county and age groups; see a toy example in Table 4.1. We then create all the households and assign their inhabitants by using a Poisson distribution with mean equal to the average number of individuals per household per county; see Figure 3.1.

Each individual is labeled with an ID from 1 to N , where N is the size of the population. We create the arrays `IDtoCounty`, `IDtoFamily` and `IDtoAge` that return the county, family and age group of a given ID, and the cells `familyToIDs` and `countyToIDs` that return the IDs for a given family or county, respectively. These arrays allows us to directly access required information to construct the graph.



(a) An infectious node and its network. (b) Two new infected nodes.



(c) Augmented network. (d) Four new infected nodes.

Fig. 2.2: Modeling the transmission of the virus. The network adapts on a daily basis depending on new exposed nodes, imitating the evolution of new cases.

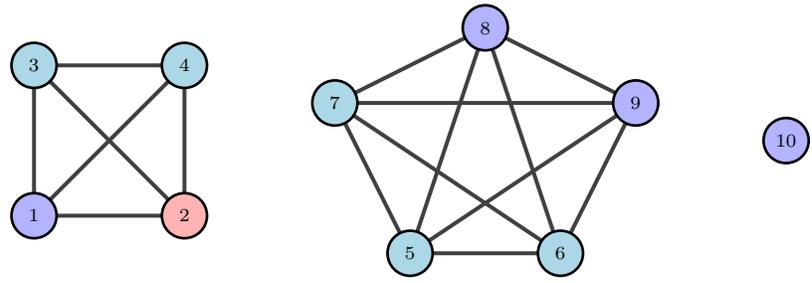


Fig. 3.1: Three families with 4, 5 and 1 members. A fully connected graph for each component is assumed. Colors represent different age groups.

Algorithm 1 Multilayer network model

```
1: procedure NETWORKMODEL
2:   Create household network. ▷ See Section 3.1
3:   Define initial conditions. ▷ See Section 3.2
4:   for day = 1, 2, 3... do
5:     for all new exposed nodes  $E$  do ▷ See Section 3.3
6:       if family has not been added then
7:         Read connections in the family layer.
8:         Store edge connectivity.
9:       end if
10:      if node requires more contacts in layer 2 then
11:        Create list of eligible contacts.
12:        Choose random nodes for layer 2.
13:        Store edge connectivity for layer 2.
14:      end if
15:      end for
16:      for all infectious nodes do
17:        Create list of eligible contacts.
18:        Choose random nodes for layer 3.
19:      end for
20:      for all susceptible nodes  $S$  do ▷ See Section 3.4
21:        Compute probability of infection.
22:        Determine if current node has been infected.
23:      end for
24:      Determine transitions between states. ▷ See Section 3.5
25:      Store results for current day.
26:    end for
27: end procedure
```

3.2. Initial conditions. For simplicity, we assume an initial graph given by all the exposed individuals at time zero and their contact networks; see Figure 2.2a for the case of one initial exposed node. If data is available, we can include chains of transmission for a fixed time as an initial condition, by creating edges among identified clusters. Before community transmission in Costa Rica, the public health authorities were able to keep track of chains of transmission for approximately four months. This data allowed us to create accurate contact networks for the early stages of the pandemic.

3.3. Updating the network. As mentioned in Remark 2.1, creating a particular graph and its structure for a large population is both time and memory consuming. Instead, our graph evolves over time. In order to keep track of the transmission of the disease, we only need to include the edges between new exposed individuals E and their contacts.

We choose layer 2 contacts for new added nodes as follows. We first create a list of eligible nodes based on a connectivity distribution between counties. For instance, density, demographic information or mobility data could be used for this purpose. For simplicity, we consider a binary symmetric matrix where its entry (m, n) is equal to 1 if there are connections between counties m and n , and 0 otherwise; see a toy example in Table 4.1. This means that the eligible list for a node includes all IDs

from counties that have connectivity with its county. We then randomly choose the remaining number of contacts for each node from the eligible list. Further restrictions can be added when defining the eligible list, such as age group, density and distance between counties. Edges on layer 3 are daily updated in a similar way, excluding preassigned contacts.

3.4. Probability of infection. If the probability that node j infects a susceptible node i at a given day is β_{ij} , then the probability of infection p_i for node i is given by

$$p_i = 1 - \prod_{j \neq i} (1 - \beta_{ij})$$

where j includes all the indices of nodes that can infect node i that particular day. The value β_{ij} can depend on the layer, epidemiological state, social distancing, mask use and self-care behaviour, among other factors. Transmission of the virus can also include imported cases where infection occurs in an external location. Non-periodic events can also be considered (as massive events or the creation of clusters due to super spreader events) to study their impact and possible scenarios, where different probabilities p_i can be defined for each particular case.

In our implementation, we consider two binary random variables `useMask` and `selfCare` that indicate whether or not an encounter between two nodes includes mask use and self-care, respectively. We assume that a fixed fraction of diagnosed cases do not isolate and that hospitalized individuals cannot infect susceptible nodes. We then have different values for β_{ij} depending on the interaction between the two nodes and their attributes. We only need to check susceptible nodes that have infectious contacts. Once p_i is determined, a uniformly distributed pseudo-random number $r_i \in (0, 1)$ is generated. If $r_i < p_i$, then node n_i is marked as an exposed individual at the current time step.

3.5. Transitions between states. At each time step, every node has one of the following states: susceptible, exposed, diagnosed, undiagnosed, hospitalized, recovered or dead; see Figure 2.1. The event where a susceptible becomes exposed is given by the probability of infection as discussed in Section 3.4. Any other event depends on a *probability of transition*. Every node has a different probability to move to a feasible next state depending on the period of time at its current epidemiological state. If such event should happen, we then store the new state and day; if not, the state remains unchanged.

We show a toy example of probabilities of transition for the diagnosed class O in Table 3.1. In this case, if a node has one day as O , it has a 40% chance that it will be hospitalized. Otherwise, it will remain as O on day 2. On day three, there are three possibilities: remain at O (50%), move to H (10%) or recover R (40%). Note that each row adds up to 1 (one state should always be chosen according to these probabilities), and eventually all nodes that have been exposed should end up as R or D . This approach allows us to include variable stay periods on each class if data is available; average stay periods can be considered otherwise.

4. Numerical results. In this section we present results for two toy examples obtained with a parallel code implemented in Matlab and a Lenovo SR650 Server with two 2.20GHz Intel Xeon Plata 4214 processors. We remark that a generalization of this implementation has allowed us to study disease projections and different scenarios

time in O	S	E	O	U	H	R	D
1			0.6		0.4		
2			1.0				
3			0.5		0.1	0.4	
4			0.1		0.3	0.6	
5						1.0	

Table 3.1: Toy example for the probability of transitions between states from the observed class O ; omitted entries are zero. From compartment O , it is possible to move to states H or R as shown in Figure 2.1.

in Costa Rica, where the total population is approximately five million people with feasible running times.

4.1. Fixed parameters. We consider four counties with a population of 1300 nodes; its demographic distribution and the connectivity matrix are shown in Table 4.1. We consider initially E_0 exposed nodes randomly chosen. Parameters used in these simulations appear in Table 4.2. We assume that 70% of the population wears a face mask, 35% respects self-care and 60% of exposed nodes become diagnosed. Values for the probability of infection are shown in Table 4.2 for layer 2; probabilities are halved for layer 3. Finally, the probabilities of transitions between states are shown in Appendix 1.

County	HH	ANIH	AG1	AG2	AG3	County	1	2	3	4
1	140	3.0	20	50	30	1	1	1	1	1
2	100	3.5	28	44	28	2	1	1	0	1
3	80	4.0	20	40	40	3	1	0	1	0
4	60	3.5	40	40	20	4	1	1	0	1

Table 4.1: (left) Toy demographic distribution considered in Section 4.1. We present the number of households (HH), the average number of individuals per household (ANIH) and age groups (AG) percentage distribution. (right) Connectivity matrix. Non-zero entries indicate possible connections among counties.

Parameter	Interval	Face mask	Self care	β
Contacts in social network	[5, 15]	No	No	0.21
Contacts in sporadic network	[0, 15]	No	Yes	0.15
Contacts per day, county 1	[5, 15]	Yes	No	0.08
Contacts per day, county 2	[3, 13]	Yes	Yes	0.05
Contacts per day, county 3	[1, 10]			
Contacts per day, county 4	[1, 10]			

Table 4.2: (left) Network parameters in toy example and (right) probability of infection depending on the type of encounter between two nodes for layer 2. Values on layer 3 are halved; see Section 4.1.

We present a simulation for $E_0 = 1$ (only one initial exposed node) in Figure

4.1a. In Figure 4.1b, we present 1000 curves for the cumulative number of cases. We remark that in this scenario with only one initial exposed node, there is no outbreak in 7% of the simulations. Meanwhile, in the remaining 93% there are 700 cumulative number of cases on average.

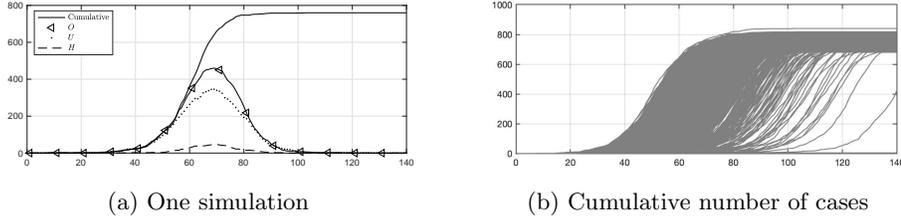


Fig. 4.1: Experiments with $E_0 = 1$. (left) Cumulative number of cases, O , U and H for one simulation. (right) Cumulative number of cases for 1000 simulations; see Section 4.1. The variability is due to the fact that there is only one initial exposed node.

For $E_0 = 10$ and the given parameters, there is a one-wave pattern in the numerical results. We present the mean of 1000 simulations with their 5th and 95th percentiles in Figure 4.2. Results can also be desegregated by county and/or age group, but we omit these figures for the sake of brevity. They can be obtained by using the app given in [15].

4.2. Modeling changes in behaviour. We now consider a population of one million individuals distributed in 15 counties. We impose restrictions on social behaviour by increasing the percentage of individuals that wear a mask and decreasing the number of daily interactions on an period of two months, starting on day 20, emulating the effect of mild regulations. We present results for 1000 simulations in Figure 4.3, where we compare the effect of such restrictions.

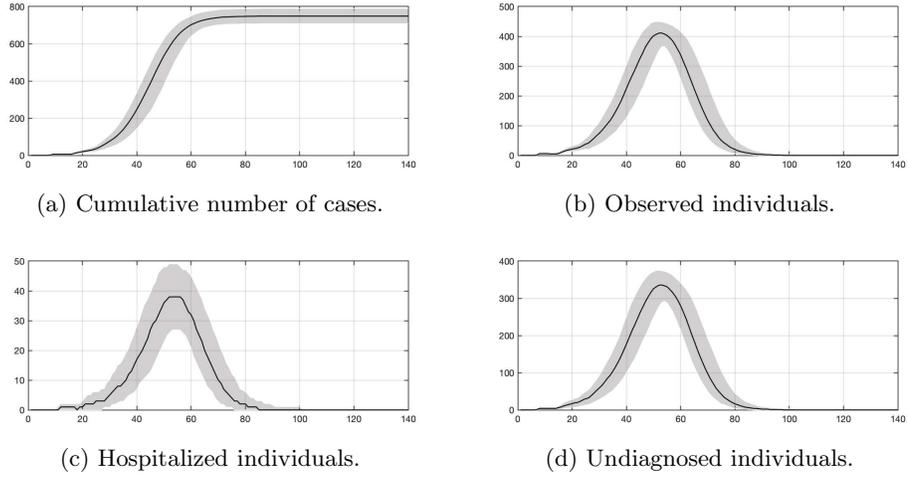


Fig. 4.2: Mean (black line) for 1000 simulations with $E_0 = 10$. The shaded region corresponds to values between the 5th and 95th percentiles of 1000 simulations; see Section 4.1.

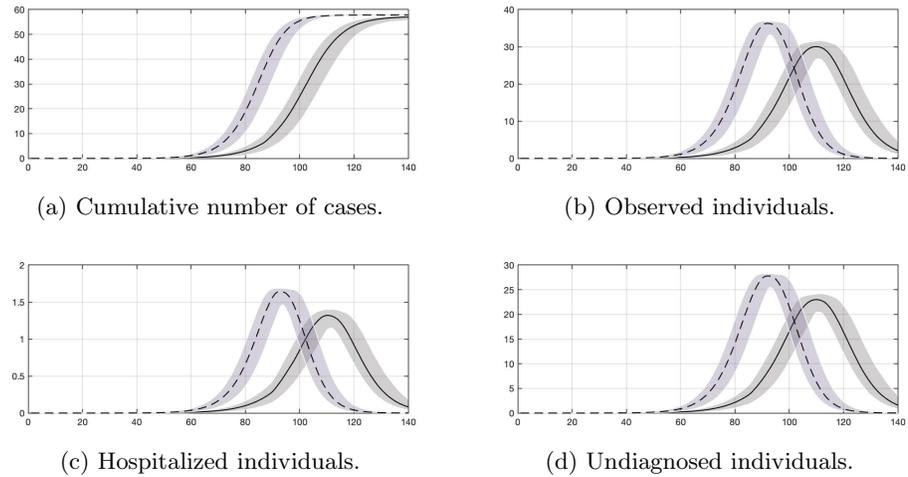


Fig. 4.3: Percentage of cases with no measures (dashed line) and mild restrictions (solid line) for two months starting on day 20. We consider a reduction on the number of daily contacts and an increment in use of masks with $E_0 = 10$. The shaded region corresponds to values between the 5th and 95th percentiles of 1000 simulations; see Section 4.2.

5. Conclusions and final remarks. We have presented an implementation for a multilayer network to support public health authorities during the Covid-19 pandemic. Capturing the transmission dynamics of the SARS-CoV-2 virus is essential

in order to provide information about disease projections and specific data-driven scenarios.

The model exhibits a flexible structure that allowed us to include particular attributes for each node and their contact networks, describing the complexities of social behavior and the specific disease. In this article, the implementation of the general structure of the model is described which can be used in other social and epidemic contexts.

With a health emergency such as the one caused by the SARS-CoV-2 virus, public health authorities were forced to make short-term decisions. Mathematical models help authorities to systematize the transmission mechanisms of the virus and project the expected behavior in the short and medium-term under some assumptions. In particular, a network model is a valuable tool as an input in the decision-making process for health authorities. It has the capability to include social factors and an epidemiological structure, enabling a greater level of detail in the study of the disease. Despite the flexibility of the model, there are limitations that mostly pertain to data availability and its quality.

Implementing and executing these models can be challenging. The availability of information is closely related to the specificity of the question that needs to be answered. Information on the initial conditions for the model and parameter value estimates can be a limiting factor when data is not available or its quality is not optimal. Efficient and fast algorithms are necessary due to the on-demand health scenarios that need to be simulated frequently, and execution times are limited by computational resources.

In an rapidly evolving epidemic as Covid-19, the model must be calibrated frequently, as social behavior is constantly changing and, in many cases, it is unpredictable. The model we have discussed in this paper allows us to straightforwardly adapt and include frequent changes on variables that mimic social behavior. Additionally, the model can be generalized to include other factors, such as the effect of herd immunity and the role of vaccination, as countries take the necessary steps to get back to the new normal.

Acknowledgments. The authors would like to thank support by the Research Center in Pure and Applied Mathematics and the Department of Mathematics at Universidad de Costa Rica.

REFERENCES

- [1] Center for Disease Control. COVID-19 Forecasts: Deaths — CDC. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html>, February 2021. (Accessed on 02/17/2021).
- [2] Adiga, A., Dubhashi, D., Lewis, B., Marathe, M., Venkatramanan, S., Vullikanti, A. Mathematical models for Covid-19 pandemic: a comparative analysis. *Journal of the Indian Institute of Science*, **100**:793-807, 2020.
- [3] Torrealba-Rodriguez, O., Conde-Gutiérrez, R.A., Hernández-Javier, A.L. Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons & Fractals*, **138**:109946, 2020.
- [4] Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., Sun, F., Jit, M., Munday, J.D. et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The lancet infectious diseases*, **20**(5):553-558, 2020.
- [5] Centre for the Mathematical Modelling of Infectious Diseases LSHTM. <https://www.lshtm.ac.uk/research/centres/centre-mathematical-modelling-infectious-diseases>, February 2021. (Accessed on 02/17/2021).
- [6] Star, L., Moghadas, S.M. The role of mathematical modelling in public health planning and

- decision making. *Purple Paper, National Collaborative Center for Infectious Diseases*, 2010.
- [7] Perelson, A.S., Ribeiro, R.M. Modeling the within-host dynamics of HIV infection. *BMC biology*, **11**(1):1-10, 2013.
- [8] Dietz, K., Schenzle, D. Mathematical Models for Infectious Disease Statistics. In: Atkinson A.C., Fienberg S.E. (eds), *A Celebration of Statistics*. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-8560-8_8.
- [9] Etbaigha, F., Willms, A.R., Poljak, Z. An SEIR model of influenza A virus infection and reinfection within a farrow-to-finish swine farm. *PLoS One*, **13**(9):e0202493, 2018.
- [10] Hethcote, H.W. The mathematics of infectious diseases. *SIAM Review*, **42**(4):599-653, 2000.
- [11] Choisy, M., Guégan, J.-F., Rohani, P. Mathematical modeling of infectious diseases dynamics. *Encyclopedia of infectious diseases: modern methodologies*, **379**, 2007.
- [12] Chen, S.-C., Hsieh, M.-H. Modeling the transmission dynamics of dengue fever: implications of temperature effects. *Science of the total environment*, **431**:385-391, 2012.
- [13] Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A. Multilayer networks. *Journal of Complex Networks*, **2**(3):203-271, 2014.
- [14] De Domenico, M., Granell, C., Porter, M.A., Arenas, A. The physics of spreading processes in multilayer networks. *Nature Physics*, **12**(10):901-906, 2016.
- [15] EpiMEC. Covid-19 app. www.github.com/epimec, 2021.
- [16] Pan American Health Organization (PAHO/WHO). Costa Rica: Pandemia covid-19. Informe estratégico mensual N°3. <https://www.paho.org/es/documentos/costa-rica-pandemia-covid-19-informe-estrategico-mensual-no-3>, 2020. (Accessed on 03/09/2021).
- [17] Pan American Health Organization (PAHO/WHO). Costa Rica: Pandemia covid-19. Informe estratégico mensual N°4. <https://www.paho.org/es/documentos/costa-rica-pandemia-covid-19-informe-estrategico-mensual-no-4>, 2020. (Accessed on 03/09/2021).
- [18] Pan American Health Organization (PAHO/WHO). Costa Rica: Pandemia covid-19. Informe estratégico mensual N°5. <https://www.paho.org/es/documentos/costa-rica-pandemia-covid-19-informe-estrategico-mensual-no-5>, 2020. (Accessed on 03/09/2021).
- [19] Pan American Health Organization (PAHO/WHO). Costa Rica: Pandemia covid-19. Informe estratégico mensual N°6. <https://www.paho.org/es/documentos/costa-rica-pandemia-covid-19-informe-estrategico-mensual-no-6>, 2020. (Accessed on 03/09/2021).
- [20] INEC. Población — Instituto Nacional de Estadística y Censos. <https://www.inec.cr/poblacion>. (Accessed on 02/17/2021).

Appendix A. Probabilities of transition. We present the probabilities of transition between epidemiological states used in the toy examples of Section 4. We remark that there are different probabilities that depend on the age group; see Tables A.1, A.2, A.3, A.4.

Days	1	2	3	4	5	6	7
E	1.0	1.0	1.0	1.0	1.0	0.8	
O						0.1	0.6
U						0.1	0.4

Table A.1: Probabilities from E to O and U depending on the number of days at state E ; empty entries are zero.

Days	1	...	12	13	14	15	16	17	18
U	1.0	...	1.0	0.9	0.8	0.7	0.5	0.1	
R				0.1	0.2	0.3	0.5	0.9	1.0

Table A.2: Probabilities from U to R depending on the number of days at state E ; empty entries are zero.

Days	Age group 1			Age group 2			Age group 3		
	O	H	R	O	H	R	O	H	R
1	1.000			1.000			1.000		
2	1.000			1.000			1.000		
3	1.000			1.000			1.000		
4	0.978	0.022		0.953	0.047		0.629	0.371	
5	1.000			1.000			1.000		
...		
17	1.000			1.000			1.000		
18			1.000			1.000			1.000

Table A.3: Probabilities from O to H and R depending on the number of days at state E ; empty entries are zero.

Days	Age group 1			Age group 2			Age group 3		
	H	R	D	H	R	D	H	R	D
1	1.00			1.00			1.00		
2	1.00			1.00			1.00		
3	1.00			1.00			1.00		
4	1.00			1.00			1.00		
5	1.00			1.00			1.00		
6	1.00			1.00			0.75		0.25
7	1.00			1.00			0.75		0.25
8	1.00			1.00			0.75		0.25
9	1.00			1.00			1.00		
10		1.00		1.00			1.00		
11		1.00			0.96	0.04	1.00		
12		1.00			0.96	0.04	1.00		
13		1.00			0.96	0.04	1.00		
14		1.00			0.96	0.04		0.75	0.25

Table A.4: Probabilities from H to R and D depending on the number of days at state H ; empty entries are zero.