Microscopic Markov Chain Approach for Measuring Mobility Driven SARS-CoV-2 Transmission

Trevor G. Kent[1(✉)], Nolan E. Phillips[1], Ian McCulloh[1,2], Viveca Pavon-Harr[1], and Heather G. Patsolic[1]

[1] Applied Intelligence Accenture Federal Services, 1201 New York Ave NW, Washington, DC 20005, USA
trevor.g.kent@accenturefederal.com
[2] Whiting School of Engineering, Johns Hopkins University, Laurel, MD, USA

Abstract. After more than a year of non-pharmaceutical interventions, such as, lock-downs and masks, questions remain on how effective these interventions were and could have been. The vast differences in the enforcement of and adherence to policies adds complexity to a problem already surrounded with significant uncertainty. This necessitates a model of disease transmission that can account for these spatial differences in interventions and compliance. In order to measure and predict the spread of disease under various intervention scenarios, we propose a Microscopic Markov Chain Approach (MMCA) in which spatial units each follow their own Markov process for the state of disease but are also connected through an underlying mobility matrix. Cuebiq, an offline intelligence and measurement company, provides aggregated, anonymized cell-phone mobility data which reveal how population behaviors have evolved over the course of the pandemic. These data are leveraged to infer mobility patterns across regions and contact patterns within those regions. The data enables the estimation of a baseline for how the pandemic spread under the true ground conditions, so that we can analyze how different shifts in mobility affect the spread of the disease. We demonstrate the efficacy of the model through a case study of spring break and it's impact on how the infection spread in Florida during the spring of 2020, at the onset of the pandemic.

Keywords: Microscopic Markov Chain Approach · Approximate Bayesian Computation · Epidemiology · SARS-CoV-2

1 Introduction

During the bubonic plague outbreak of 1665–1666, the town of Eyam stands out as an example of how limiting mobility can alter the spread of a deadly disease. After the bubonic plague came to Eyam through trade with London, the town and its neighboring communities developed a plan that would enforce

R. M. Benito et al. (Eds.): COMPLEX NETWORKS 2021, SCI 1016, pp. 301–314, 2022.
https://doi.org/10.1007/978-3-030-93413-2_26

the first example of 'curbside pick-up' by shutting down their borders to flow in and out of the community while neighboring communities dropped off much needed supplies outside those borders [4,27]. As author and prominent Youtube content creator John Green points out in a podcast episode of the Anthropocene Reviewed [14], "Eyam's self-quarantine held to a remarkable extent for fourteen months, preventing the spread of the mortality to any nearby towns." While fascinating, the convolutions of human connection did not exist to the extent that it does today. We live in an increasingly interconnected society where technological advancements make travel across vast distances more accessible. The United Nations World Tourism Organization estimates there were nearly 1.4 billion international arrivals in the year 2018; a figure that has more than doubled since the year 2000 [24]. With this increased travel comes more wide-spread transmissibility of communicable diseases which are no longer confined to the locations from which they originate. The SARS-CoV-2 pandemic has particularly illuminated just how connected our world has become, and the effect of human connectivity on disease transmission.

The advent of widespread use of mobile phones and GPS devices has created an unprecedented amount of data which captures the locations of individuals at varying time-points. With more granular details about human mobility patterns, researchers have been able to leverage these data to help explain and predict the spread of disease while developing epidemiological models and algorithms. Researchers in [2] were able to use mobility and social distancing parameters to model spread of COVID-19 during the early stages of the pandemic in Spain. However, due to their limited access to mobility data, they were forced to adopt a rather complex model that introduced dynamic contact parameters by applying deviations from their "baseline" conditions. Our approach differs in that we adopt time varying mobility and contact parameters derived from Cuebiq that mitigates uncertainty arising from variation of human behavior before, during, and after lockdown interventions. We show that access to these data allows us to achieve a similar fit to observed daily cases while maintaining a parsimonious functional form for the probability of infection. Researchers in [5] also adopt a probability of transmission with dependencies on human mobility and contact parameters derived from mobile phones. They were able to show how model fits improved given more realistic mobility networks for 10 major commuting regions in the United States. While this method shows promise, limiting the scope of analysis to a commuting region imposes artificial boundaries on the system. We apply a novel approach that includes an extra node in the mobility matrix which encodes the rest of the U.S. and thereby permits infections initialized elsewhere to spread into the region of interest.

As was found in [20], it is crucial that we examine more granular transmission patterns of a disease to better understand its spread. While the authors of [20] focus on vaccination, which is not relevant for the time frame of this study, looking at how people are mobilizing and coming into contact with one another at a more granular scale than country or state level helps us better understand how things like adherence to stay-at-home protocols and across-region mobility

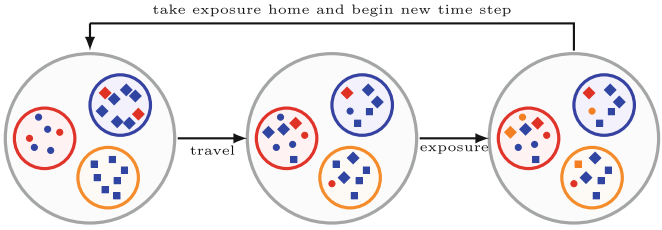


Fig. 1. Considering a population which has three sub-populations, we provide an illustrative example of how susceptible (blue) and infected (red) individuals mobilize and interact during a single time step. Individuals who transition to being exposed during this illustrative time-step are shown in (orange).

impact the spread of a disease. Studies, such as the one presented in [11], have argued that stay-at-home orders have lead to reduced across-region mobility, mitigating the spread of disease. Cuebiq's data allow us to include mobility and contact trends at the county level, for a more refined model.

In this manuscript, we introduce a Microscopic Markov Chain Approach (MMCA) SEIRD compartmental model that incorporates mobility and contact patterns, building upon the work in [2,5]. Using this model, we demonstrate how these data influence the spread of a contagious disease. Our model aims to better account for the effects of mobility on the likelihood of becoming infected. We assume each individual starts in a particular spatial unit; hereafter, referred to as patch. At each time step, they have the ability to move to a different patch, interact with individuals in that patch, and then return to their initial patch. Thus, the probability of a susceptible person becoming infected at a given time point depends on the patch where interactions occur and the demographics of the population in that patch at that time. Figure 1 illustrates an example of how individuals might move around at a given time point in a population with 3 patches. We see three states of infection (red, blue, orange) where blue dots correspond to susceptible individuals and red dots correspond to infected individuals. Reading the figure from left to right, we see that during a time-step individuals start in their patch, travel to other patches where they mingle, and then transition to being exposed (orange) before returning to their original patch for the next time-step. For ease of illustration, we do not include nodes for individuals in the R/D compartments.

The article proceeds as follows. First we will delineate the technical details of the model in Sect. 2. In Sect. 3 we provide a thorough analysis of the spread of COVID-19 in the state of Florida at the county level where we explain the data estimates Sect. 3.1, fitting procedure Sect. 3.2, and scenario analyses Sect. 3.3 explored in this study. Finally, we conclude with a discussion Sect. 4 on findings and explore other potential avenues of research.

2 Technical Details

We have a population of n individuals who are partitioned into patches $\{1, \ldots, N\}$. We will use the notation $\rho_i^m(t)$ to denote the proportion of individuals in patch $i \in \{1, \ldots, N\}$ and state $m \in \{S, E, I, R, D\}$ at time t, where time is measured in days. We adopt a Markov model in which tomorrow's state of disease is only dependent on the current state.

$$\rho_i^S(t+1) = (1 - \Pi_i(t))\rho_i^S(t) \tag{1}$$

$$\rho_i^E(t+1) = \Pi_i(t)\rho_i^S(t) + (1-\sigma)\rho_i^E(t) \tag{2}$$

$$\rho_i^I(t+1) = \sigma\rho_i^E(t) + (1 - \gamma(1-\omega) - \omega\delta)\rho_i^I(t) \tag{3}$$

$$\rho_i^R(t+1) = (1-\omega)\gamma\rho_i^I(t) \tag{4}$$

$$\rho_i^D(t+1) = \omega\delta\rho_i^I(t) \tag{5}$$

We assume static parameters for transitions from $E \to I$, $I \to R$, and $I \to D$. However, due to mobility and social distancing, the primary subject of this study $\Pi_i(t)$, which represents the transition probability from state S at time t to state E at time $t+1$, varies with both time and space. The probability of a randomly selected individual in patch i making such a transition can be expressed as the sum of the probabilities of getting infected in each of the patches multiplied by the probability the individual went to a different patch (Bayes' Rule). Let the matrix $R(t)$ denote the mobility matrix where $R_{ij}(t)$ denotes the proportion of individuals in patch i who mobilize to patch j at time t, and let $P_j(t)$ denote the probability of being infected while in patch j at time t. Note that $\sum_{j=1}^{N} R_{ij}(t) = 1$ for all $i \in \{1, \ldots, N\}$, and all time-steps t. We can express $\Pi_i(t)$ as the convex combination of the proportion of individuals in patch i getting infected by the contagion in patch j, over all j, at time t.

This can be expressed as follows:

$$\Pi_i(t) = \sum_{j=1}^{N} R_{ij}(t)P_j(t), \tag{6}$$

where $P_j(t)$ captures the probability that a person becomes infected by the contagion while in patch j. This, again, depends on mobility, since the probability a person becomes infected while in patch j depends on how people are mobilizing into and out of that patch, how many other individuals a susceptible individual can expect to come into contact with while in patch j, and what proportion of those individuals will be carriers of the disease.

Let $X(t; j \to i)$ denote the number of susceptible individuals who are exposed by an infected individual who came to patch i from patch j at time t. Further assume exposure is independent across individuals. We can model $X(t; j \to i)$ as a Binomial random variable with parameters $Q(t; j \to i)$ and β where $Q(t; j \to i)$ represents the number of contacts between individuals in patch i with individuals coming into patch i from patch j at time t and β represents the rate of transmission, which is assumed to be fixed over time and patches.

$$P_j(t) = 1 - \prod_{k=1}^{N} P\left(X(t; k \to j) = 0\right)$$
$$= 1 - \prod_{k=1}^{N} (1-\beta)^{Q(t;k\to j)}.$$

To compute $Q(t; k \to j)$, which is the proportion of infected individuals coming into patch j from patch k at time t scaled by the expected number of contacts, we need to multiply the average number of contacts for someone in patch j at time t, $c_j(t)$, by the proportion of *infected* individuals in patch j who mobilized from patch k at time t. Thus, we can express $Q(t; k \to j)$ as follows:

$$Q(t; k \to j) = c_j(t) \frac{n^I_{k\to j}(t)}{n^{eff}_j(t)}, \tag{7}$$

where

$$n^I_{k\to j}(t) = \rho^I_k(t) n_k R_{kj}(t) \eta_k(t)$$

represents the proportion of infected individuals from patch k mobilizing to patch j at time t. Here $\eta_k(t)$ denotes the proportion of infectious individuals in patch k who choose to mix with the general population,

$$\eta_k(t) = (1 - \alpha_k(t))[\lambda_0 - \lambda_c \mathcal{H}(t - t_c)] \tag{8}$$

where,

$$\mathcal{H}(x) = \begin{cases} 1, & x \geq 0 \\ 0 & x < 0 \end{cases}$$

and finally,

$$n^{eff}_j(t) = \sum_{k=1}^{N} n_k R_{kj}(t)$$

represents the effective population of patch j at time t, that is, the total number of individuals in patch j at time t taking into consideration individuals staying in the patch and coming in from other patches.

Taken together, the probability of becoming infected by the contagion while in patch j is

$$P_j(t) = 1 - \prod_{k=1}^{N} (1-\beta)^{Q(t;k\to j)} \tag{9}$$
$$= 1 - (1-\beta)^{\frac{c_j(t)\sum_{k=1}^{N} n^I_{k\to j}(t)}{n^{eff}_j(t)}}. \tag{10}$$

Thus, $\Pi_i(t)$ can be computed as follows.

$$\Pi_i(t) = \sum_{j=1}^{N} R_{ij}(t) \left(1 - (1-\beta)^{\frac{c_j(t)\sum_{k=1}^{N} n^I_{k\to j}(t)}{n^{eff}_j(t)}}\right). \tag{11}$$

The above formulation does not take into consideration those individuals who are staying home, and assumes that people who are infected mobilize outside the resident patch at a rate proportional to the population. We incorporate the proportion of individuals staying home with the individuals remaining in their respective patches at time t. Allow $Z(t)$ to be a matrix such that row i represents the distribution of mobility of individuals across patches out of those who are mobilizing and are from patch i, while row i of $R(t)$ represents the distribution of individuals across patches out of those individuals who begin time t being in patch i. This means that $R_{ij}(t) = \kappa_i(t)(I_N)_{ij} + (1 - \kappa_i(t))Z_{ij}(t)$ where $\kappa_i(t)$ denotes the proportion of individuals staying home at time t and $(I_N)_{ij}$ represents an element of the identity matrix with $N \times N$ dimensions. Then $Z_{ii}(t)(1 - \kappa_i(t))$ denotes the proportion of individuals who are not staying home at time t but are mobilizing *within* patch i (Fig. 2).

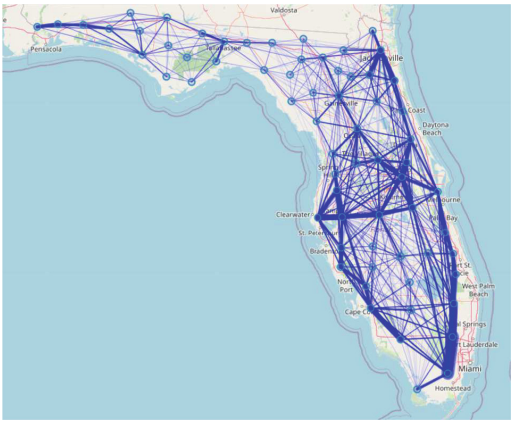


Fig. 2. The probability of infection updates each day according to a Markov Process and patches are linked through an underlying mobility matrix $R(t)$. Displayed here using scikit-mobility [23] is the mobility network for a snapshot in time where the thickness of the connections are motivated by the proportion of people from the source county that travel to a target county. Note that the extra node in the mobility matrix to denote the rest of the U.S. is not shown here.

In practice it is important to note that the region under consideration is not a closed system. The boundary issue is a long-established problem within network and spatial analyses [18,28]. Rather than imposing an artificial boundary that would explicitly assume infections cannot be imported from areas outside the focal area, we include a "catch all" node to incorporate mobility patterns from all other areal units. The mobility and contact parameters are averaged over all exogenous patches. Though this is still an abstraction from adopting a model over the region at large, it is an improvement over approaches that impose boundaries around the area of study. In our case study, introduced in Sect. 3, we show how

the MMCA SEIRD model can be applied to the spread of SARS-CoV-2 in Florida and add an extra node to encapsulate mobility between Florida and the rest of the U.S.

3 Florida and Mobility During the Pandemic

We demonstrate the utility of our model through a case study of SARS-CoV-2 spread from January through May of 2020 at the county level in Florida. We focus on the months of March and April, surrounding spring break where we observe larger volumes of people in Florida than normal.

3.1 Data and Parameter Estimates

The majority of parameters in the model are either estimated directly from Cuebiq [1] or pulled from the literature. To estimate the unknown parameters, we employ an Approximate Bayesian Computation (ABC) algorithm as defined in [8]. This family of algorithms is advantageous when the likelihood function is expensive or intractable to calculate, which is the case here since we would otherwise have to sum over all latent paths from which infections could have arisen. The unknown parameters of our model consist of the transmission rate β, initial number of latent infectious seeds e_0, initial number of infectious seeds i_0, proportion of infectious individuals who choose to mix with the general population despite being infectious λ_0 and then following the onset of lockdown restrictions this figure is reduced by λ_c. The method yields means for the posterior distributions of each variable.

Parameter	Value	Description	Source
$c_j(t)$	–	Average contacts	[1]
$R_{ij}(t)$	–	Mobility matrices	[1]
$\kappa_i(t)$	–	Proportion staying home	[1]
$\alpha_k(t)$	–	Social distancing index	[29]
$p_i(t)$	–	Prevalence ratio	[15]
ζ	14	Reporting lag	[15]
σ	1/5.1	Incubation rate	[17]
γ	1/21	Recovery rate	[26]
δ	1/17.8	Death rate	[26]
ω	0.0066	Infection fatality ratio	[26]
β	0.0649	Transmission rate	Calibrated
e_0	24.988	Initial pre-infectious seeds	Calibrated
i_0	9.849	Initial infectious seeds	Calibrated
λ_0	0.8991	Infectious non-compliance pre-lockdown	Calibrated
λ_c	0.2499	Non-compliance adjustment post-lockdown	Calibrated

The U.S. department of health and human services reports data on important intervention dates like containment or reopening events at the state and county level [7]. They report that many counties in Florida didn't officially shutdown until late March or early April. Thus we chose to impose our lockdown restriction t_c at March 30, 2020. The first official case in the United States was reported on January 21, 2020 [10], yet there is still uncertainty surrounding when the first infections appeared and how many. We chose to start the pandemic on December 30, 2019 (the day before the first confirmed case reported in China [10]). We estimate the initial infectious seeds i_0 as 10 (10 times the number known in China), and initial exposed cases e_0 as 25 in the region outside Florida. The ratio of initial exposed to initial infectious seeds falls within the credible range of what R_0 for SARS-CoV-2 is estimated to be [21,30] (Fig. 3).

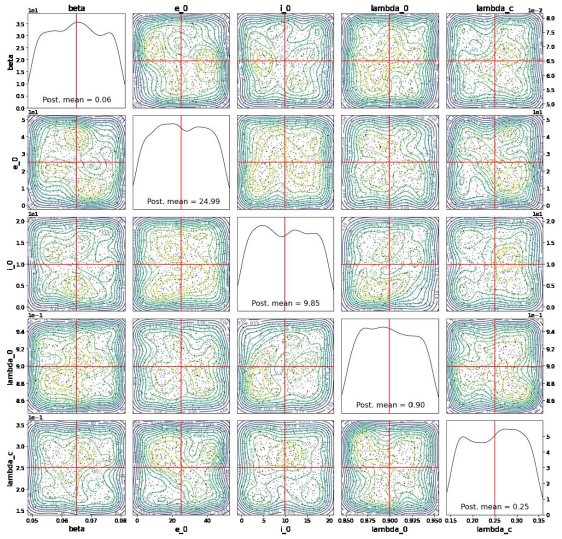


Fig. 3. The posterior distributions of the unknown parameters in the model. We use the Python package abcpy [9] with 1000 samples.

The University of Maryland reports a data series $\alpha_k(t)$ that measures the extent to which social distancing restrictions are being followed at the county level [29]. We use this parameter in Eq. 8 along with λ_0 and λ_c to measure the proportion of infectious individuals who are not adhering to social distancing restrictions before and after lockdown at time t_c. Before the pandemic, the United States fostered a culture that encouraged individuals to go to work whenever possible, including when sick. A prior CDC study found that roughly 82.7% of individuals went to work even though they were exhibiting flu-like symptoms [12]. We surmise that during the early months of the pandemic prior to lockdown restrictions, people who were infectious likely went about their business

as usual, only staying home when they felt they could afford to take time off. Once citizens became more aware that the disease was a more dangerous threat, lock-downs started being mandated and policies enacted to ensure people would get tested and stay home while sick.

3.2 Residuals: True Estimates and the Markov Model

Due to the asymptomatic nature of the disease [19] and limited testing availability early on in the pandemic [25], a significant quantity of cases went undetected [22,25]. These missing observations present challenges for estimating unknown parameters in an epidemiological model. Several researchers have found success in fitting to COVID-19 attributed death counts instead [2,22]. The authors in [22] implement a Bayesian augmentation algorithm and were able to validate the model with sero-prevalence data from New York state. However, this approach cannot be applied to the earliest stages of the pandemic at more granular spatial areas since not enough deaths had been reported. An independent researcher Youyang Gu provides an alternative approach reliant on testing data, which became one of the most widely-cited models in academia, media, government, and industry [15]. Along with his infections estimates, he provides a prevalence ratio, which we will denote as $p_i(t)$, and represents the ratio of true new daily infections to daily reported cases. We choose to fit to reported new daily cases after correcting for an assumed $\zeta = 14$ day lag in reporting and applying the prevalence ratio. Thus to arrive at our estimate of new daily cases we first calculate infectious incidence $N_i^{E\to I}(t)$ like so:

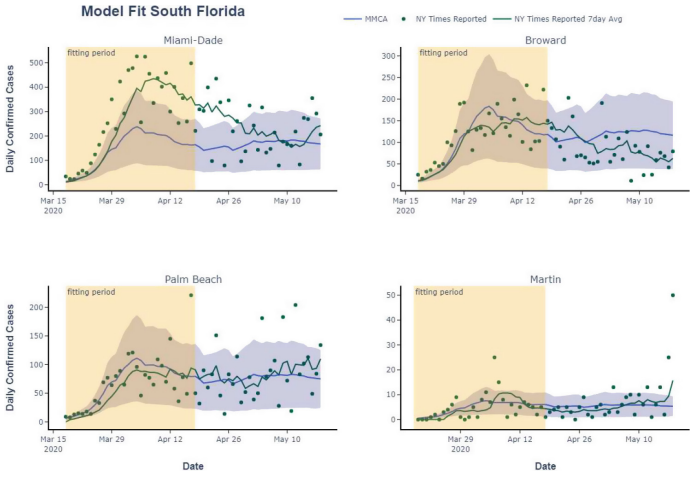


Fig. 4. For South Florida, we show reported new daily cases from *New York Times* (green) vs new daily cases estimated from our MMCA SEIRD model (blue). We choose to fit the model from 03-18-2020 to 04-18-2020. The shaded region represents the 95% confidence interval for our estimates.

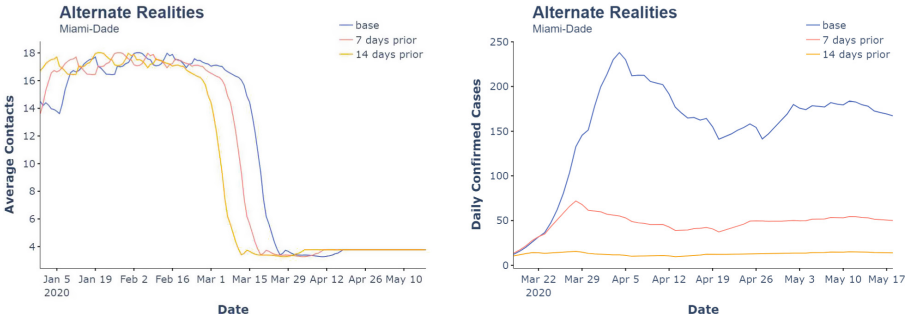
$$N_i^{E\to I}(t) = \sigma \rho_i^E(t) n_i \tag{12}$$

The estimated daily cases is then given by:

$$\overline{N}_i^{E\to I}(t) = \frac{N_i^{E\to I}(t-\zeta)}{p_i(t)} \tag{13}$$

In Fig. 4 we show how our estimated daily cases $\overline{N}_i^{E\to I}(t)$ compares to the reported 7 day average of new daily cases from the *New York Times*. We demonstrate that under baseline population behaviour derived from Cuebiq, the model is able to fit well within the 1 month fitting period from March 18th to April 18th, as well as a 1 month held out period from April 18th to May 18th. We first apply our model to data from March 18th since it is the earliest date that we have prevalence ratio data for all counties in South FL.

3.3 Alternate Realities



(a) Average Contacts (b) Daily Confirmed Cases

Fig. 5. Displayed here are model estimates of average contacts $c_j(t)$ (a) and new daily reported cases $\overline{N}_i^{E\to I}(t)$ (b) given two alternative scenarios. The scenario where lock down occurs 7 days prior to t_c is shown in (salmon) while the scenario where lock-down occurs 14 days prior to t_c is shown in (orange). The scenario given real data inputs is shown in (blue). Note that we replace the true input parameters with the last known input parameters for all time points after the last day of the fitting period.

Once the default parameters are set, a user can investigate the impact of alternate realities, such as, never shutting down or allowing lockdown restrictions to persist over a much longer period. Here, we examine the scenarios where lockdown restrictions occur 7 days and 14 days prior to the official enforcement dates. All data derived input parameters are lagged to simulate the effect of the shift in human behaviour as a result of awareness campaigns or enforcement of containment policies occurring earlier. In Fig. 5 we see an example of how $c_j(t)$

was adjusted for each scenario. We also impose a shifted lockdown date t_c so that the fitted λ_c parameter is applied earlier. This approach inherently assumes that the effect of lockdown restrictions would be the exact same when applied earlier in the pandemic.

Marginal changes in intervention dates can lead to vastly different outcomes in disease transmission. In Fig. 5, we see that our model suggests applying restrictions even a week earlier would have reduced the daily reported cases by a wide margin. On April 4, the first peak in our model, we find under the condition where lockdown occurs 7 days prior that the daily case count would reduce by 182 cases or 76.8% change. Over the entire two month period, the total number of cases would reduce by 6633 (69.2%) under the same alternative scenario. Locking down prior to spring break could have not only prevented the influx of a large volume of people into Florida, but could also likely have had a huge impact on how the pandemic played out in the rest of the U.S. due to a subset of the influx of people returning back to their resident counties with the disease. More careful selection of when lockdown restrictions are applied could have drastically reduced the strain on our public health system.

4 Discussion

We have shown that behavioral changes drive the differences in COVID-19 transmission. We use the lock-down dates as a proxy for when people change their behavior most; however, these changes are likely dynamic and heavily influenced by increased awareness and adherence to local guidelines. We do find evidence that the shift in behavior that occurred around March 30, 2020 was influential in determining how the pandemic spread in Florida. It is unclear whether the change of behavior we observe is a direct result of the lock-down restrictions themselves, awareness of the disease, or some other factor. Research is beginning to emerge on the co-evolution of awareness and epidemic progression; see, for example, [13] in which the authors adopt a multiplex network and MMCA approach. While assessing the effects of lock-downs and other formal policies on human behavior is beyond the scope of this paper, future research should seek to account for public awareness of a disease and its effect on human interactions over the course of a pandemic within this modeling framework.

There is significant evidence that a large portion of the population experienced infectiousness without showing symptoms. In fact, one study reported that 39.7 and 44.8% of positive cases sampled for two different survey periods in the Italian municipality of Vo' were found to be asymptomatic [19]. Parameter estimation can be computationally intensive, which constrains how much realism can be captured within the functional form. Though we did not include an asymptomatic compartment, we did address the fact that asymptomatic individuals are likely to mobilize differently than their symptomatic counterparts. We argue that symptomatic individuals are more aware of their illness and thus more likely to isolate themselves as a result. We regulate the extent to which discordant pairs can be formed by incorporating the stay home parameter into

the calculation of the mobility matrices. Equation 8 also captures the fact that a subset of the population will continue to mix with the population despite being infectious. This parameter plays an important role in capturing how differently asymptomatic individuals may respond to infectiousness. However, this approach does not account for the fact that the infectiousness profile may be different for asymptomatic vs symptomatic patients, and existing research proves that transmission events can occur 2 to 3 days before symptom onset [16]. Future research should seek to add this additional complexity to the model by introducing additional compartments for asymptomatic and pre-symptomatic patients.

Our model is notably less accurate at quantifying the number of cases for Miami-Dade than the other counties in Florida early in the pandemic. There is another layer of human mobility that is not captured in the commuting flows; namely airline traffic. Miami, an international hub, should have increased numbers of people and contacts, thus infections, imported from outside of the commuting region. This is an established problem in epidemiological modeling and prior research has shown the importance of accounting for airline travel as another vector of transmission [3,6]. With airline data, our approach could be extended to account for international travelers and even different rates of positivity of those travelers.

Furthermore, the exposure and transmission rates change with the introduction of vaccines and variants over the course of a pandemic. Our approach could be adjusted to account for this heterogeneity by dividing the pandemic into epochs. The end conditions of one epoch become the seed conditions for the next and the unknown parameters associated with the primary drivers of disease during that period of the pandemic could be estimated. This would require creating additional exposed and infected compartments for the variants. Similarly, an additional compartment could be created for vaccinated individuals since "break through" cases have been identified. Thus, more of the nuances of the current (or future) pandemics could be accounted for using this modeling paradigm. Though it should be noted that increased complexity will make the model take longer to run and could introduce additional errors when estimating unknown parameters.

This work enhances our ability to model future communicable diseases and the potential effects of non-pharmaceutical interventions. The approach accounts for many of the drivers of the SARS-CoV-2 pandemic, yet remains adaptable to emerging conditions. Thus, this framework empowers policy and decision makers with data-driven estimates of how the pandemic will affect specific regions and enable those decision makers to enact more efficacious policies.

Acknowledgements. This model was built in collaboration with Applied Intelligence Accenture Federal Services. The past year we served as members of their COVID-19 response team where we built and deployed models for decision support of various federal clients. We also appreciate Cuebiq's Data for Good program for providing the mobility data needed for this analysis.

References

1. Cuebiq. https://help.cuebiq.com/hc/en-us
2. Arenas, A., et al.: Modeling the spatiotemporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions. Phys. Rev. X **10**(4), 041055 (2020)
3. Bajardi, P., Poletto, C., Ramasco, J.J., Tizzoni, M., Colizza, V., Vespignani, A.: Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PloS one **6**(1), e16591 (2011)
4. Beaumont, P.: Eyam recalls lessons from 1665 battlewith plague. The Observer (2020)
5. Chang, S., et al.: Mobility network models of COVID-19 explain inequities and inform reopening. Nature **589**(7840), 82–87 (2021)
6. Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A.: The role of the airline transportation network in the prediction and predictability of global epidemics. Proc. Natl. Acad. Sci. **103**(7), 2015–2020 (2006)
7. COVID-19 State and County Policy Orders. [Dataset]. U.S. Department of Health and Human Services (2021). https://catalog.data.gov/sr_Latn/dataset/covid-19-state-and-county-policy-orders-9408a
8. Del Moral, P., Doucet, A., Jasra, A.: An adaptive sequential Monte Carlo method for approximate Bayesian computation. Stat. Comput. **22**(5), 1009–1020 (2012)
9. Dutta, R., Schoengens, M., Onnela, J.-P., Mira, A.: ABCpy: a user-friendly, extensible, and parallel library for approximate Bayesian computation. In Proceedings of the Platform for Advanced Scientific Computing Conference, pp. 1–9 (2017)
10. Editors, H.: First confirmed case of COVID-19 found in U.S. (2020). https://www.history.com/this-day-in-history/first-confirmed-case-of-coronavirus-found-in-us-washington-state
11. Engle, S., Stromme, J., Zhou, A.: Staying at home: The mobility effects of COVID-19. VOX, CEPR Policy Portal (2020). https://voxeu.org/article/staying-home-mobility-effects-covid-19
12. Cdc, P., et al.: Experiences with influenza-like illness and attitudes regarding influenza prevention-united states for disease control, 2003–04 influenza season. MMWR. Morbidity Mortality Weekly Rep. **53**(49), 1156–1158 (2004). https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5349a3.htm
13. Granell, C., Gómez, S., Arenas, A.: Competing spreading processes on multiplex networks: awareness and epidemics. Phys. Rev. E **90**(1), 012808 (2014)
14. Green, J.: Plague. [podcast] The Anthropocene Reviewed (2020). https://www.wnycstudios.org/podcasts/anthropocene-reviewed/episodes/anthropocene-reviewed-john-green-plague, Accessed 5 Oct 2021
15. Gu, Y.: COVID-19 Projections for Miami-Dade, Fl, USA (2020). https://covid19-projections.com/us-fl-miami-dade
16. He, X., et al.: Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med. **26**(5), 672–675 (2020)
17. Lauer, S.A., et al.: The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Internal Med. **172**(9), 577–582 (2020)
18. Laumann, E.O., Marsden, P.V., Prensky, D.: The boundary specification problem in network analysis. Res. Methods Soc. Netw. Anal. **61**, 87 (1989)
19. Lavezzo, E., et al.: Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. Nature **584**(7821), 425–429 (2020)

20. Masters, N.B., Eisenberg, M.C., Delamater, P.L., Kay, M., Boulton, M.L., Zelner, J.: Fine-scale spatial clustering of measles non vaccination that increases outbreak potential is obscured by aggregated reporting data. Proc. Natl. Acad. Sci. **117**(45), 28506–28514 (2020)
21. Li, Q., et al.: Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New Engl. J. Med. **382**, 1199–1207 (2020)
22. McCulloh, I., Kiernan, K., Kent, T.: Inferring true COVID19 infection rates from deaths. Front. Big Data **3**, 37 (2020)
23. Pappalardo, L., Simini, F., Barlacchi, G., Pellungrini, R.: scikit-mobility: a Python library for the analysis, generation and risk assessment of mobility data (2019)
24. Roser, M.: Tourism. Our World in Data (2017). https://ourworldindata.org/tourism
25. Silverman, J.D., Hupert, N., Washburne, A.D.: Using influenza surveillance networks to estimate state-specific prevalence of SARS-CoV-2 in the United States. Sci. Transl. Med. **12**(554), 1–9 (2020)
26. Verity, R., et al.: Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect. Dis. **20**(6), 669–677 (2020)
27. Whittles, L.K., Didelot, X.: Epidemiological analysis of the Eyam plague outbreak of 1665–1666. Proc. Royal Soc. B Biol. Sci. **283**(1830), 20160618 (2016)
28. Wang, Q., Phillips, N.E., Small, M.L., Sampson, R.J.: Urban mobility and neighborhood isolation in America's 50 largest cities. Proc. Natl. Acad. Sci. **115**(30), 7735–40 (2018)
29. Zhang, L., et al.: An interactive COVID-19 mobility impact and social distancing analysis platform. medRxiv (2020)
30. Zhao, S., et al.: Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. Int. J. Infect. Dis. **92**, 214–217 (2020). PMID: 32007643; PMCID: PMC7110798. https://doi.org/10.1016/j.ijid.2020.01.050