Foundations of Data Science
©American Institute of Mathematical Sciences
Volume **3**, Number **3**, September **2021**
doi:10.3934/fods.2021022
pp. **479–541**

A STUDY OF DISPROPORTIONATELY AFFECTED POPULATIONS BY RACE/ETHNICITY DURING THE SARS-COV-2 PANDEMIC USING MULTI-POPULATION SEIR MODELING AND ENSEMBLE DATA ASSIMILATION

Emmanuel Fleurantin

Florida Atlantic University
777 Glades Rd., Boca Raton, FL 33431, USA

Christian Sampson*

University of North Carolina
130 Mason Farm Road Chapel Hill, NC 27599, USA

Daniel Paul Maes

University of Michigan
530 Church St, Ann Arbor, MI 48109, USA

Justin Bennett

Arizona State University
1151 S Forest Ave, Tempe, AZ 85281, USA

Tayler Fernandes-Nunez and Sophia Marx

Smith College
Northampton, MA 01063, USA

Geir Evensen

NORCE Norwegian Research Centre AS
Nygårdsporten 112, 5008 Bergen, Norway

Abstract. The disparity in the impact of COVID-19 on minority populations in the United States has been well established in the available data on deaths, case counts, and adverse outcomes. However, critical metrics used by public health officials and epidemiologists, such as a time dependent viral reproductive number (R_t), can be hard to calculate from this data especially for individual populations. Furthermore, disparities in the availability of testing, record keeping infrastructure, or government funding in disadvantaged populations

2020 *Mathematics Subject Classification.* 91F99, 65K10, 65P99, 65K99.

Key words and phrases. SARS-CoV-2, COVID-19, effective reproduction number, ensemble smoothers, data assimilation, multi-population SEIR model, age stratification, ESMDA, parameter estimation, model calibration.

The second author (CSS) is supported by the US Office of Naval Research under grant N00014-18-1-2204. The third author (DPM) is supported by the NSF GRFP DGE 1256260. The seventh author (GE) was supported by internal funding from NORCE.

* Corresponding author.

can produce incomplete data sets. In this work, we apply ensemble data assimilation techniques which optimally combine model and data to produce a more complete data set providing better estimates of the critical metrics used by public health officials and epidemiologists. We employ a multi-population SEIR (Susceptible, Exposed, Infected and Recovered) model with a time dependent reproductive number and age stratified contact rate matrix for each population. We assimilate the daily death data for populations separated by ethnic/racial groupings using a technique called Ensemble Smoothing with Multiple Data Assimilation (ESMDA) to estimate model parameters and produce an $R_t(n)$ for the n^{th} population. We do this with three distinct approaches, (1) using the same contact matrices and prior $R_t(n)$ for each population, (2) assigning contact matrices with increased contact rates for working age and older adults to populations experiencing disparity and (3) as in (2) but with a time-continuous update to $R_t(n)$. We make a study of 9 U.S. states and the District of Columbia providing a complete time series of the pandemic in each and, in some cases, identifying disparities not otherwise evident in the aggregate statistics.

1. **Introduction.** The COVID-19 pandemic has amplified social and economic inequities that impact the wellness of racial and ethnic minority populations. Inequities in socio-economic conditions (e.g. undocumented workers being ineligible for unemployment benefits, types of jobs, square footage per household, ability to use delivery services, access to private transport, access to healthcare) among these populations have led the epidemic to evolve unevenly among them, frequently with catastrophic outcomes. Often times, members of disproportionately affected racial or ethnic minority populations fit the description of "front line workers" who have kept and continue to keep vital services running during the pandemic. Diverse responses and counter-measures implemented by certain states have also affected the way we understand viral spread within different populations and have made it difficult to track viral behavior in regions with differing levels of government intervention and population compliance. Matters are further complicated by the scarcity of COVID-19 data for different ethnic and racial groups, especially minority populations within a particular state. In cases where data alone cannot provide a clear understanding of events, the combination of data and a realistic model can provide a more complete picture of events.

Combining model and data estimates is referred to as data assimilation (DA). DA has been widely and successfully employed in geophysics, numerical weather prediction, and climate studies. Data Assimilation has lead to massive improvements in weather prediction over the years [19]. In addition, reconstructions of global weather patterns are routinely done by marrying available historical data to large scale weather models through DA. Having a time record of global observations is key for advancement, however they are irregular in space, time and quality. By combining data and model, a complete picture of a global weather pattern time series can be formed. An example of such a reanalysis can be found through the European Center for Medium Range Weather Forecasting (ECMWF) [24]. These data sets provide researchers with the tools to understand geophysical processes, changing climate conditions, and develop new tools to improve predictions. There are also instances where neither model nor data alone can provide an accurate picture of a given phenomena. Through the use of both model and data, the prediction of an outcropping event that resulted in a mass fish kill in the Sea of Galilee was predicted 9 hours before the event [31]. This would not have been possible through solitary data analysis or model runs alone. DA has also been used to improve the predictability of cancer models to improve patient outcomes [32].

With DA being such a powerful tool for understanding complex systems, we aim to apply the methodology to new problems in epidemiology. Some notable works have used data assimilation in the context of epidemiology, employing both variational [11] and Kalman filter-like methods [13]. Most of the studies use sequential filtering approaches, e.g., the iterative filter and the Ensemble Kalman filter (EnKF) or the Ensemble Adjustment Kalman Filter (EAKF). Recent studies on SARS-CoV-2 still mostly use a filtering approach. A filter will update bias in the solution, however, it does not adjust the parameters that can lead to that bias. Instead of filtering, we make use of an Ensemble Smoother with Multiple Data Assimilation (ESMDA) updating the parameters that lead to bias in the solution. This method allows for the estimation of time dependent parameters (such as a time dependent viral reproductive number (R_t)) since fitting their values at a given time can be adjusted to predict data at later times. In this way, we are updating the parameters in the past to remove the bias in the solution before it happens. Other examples of iterative ensemble smoothers can be found in [12, 14], in the geosciences in [15] or petroleum reservoir modeling in [29].

For our model, we develop a multi-population SEIR (Susceptible, Exposed, Infected, and Recovered) model that can be used to study how the SARS-CoV-2 virus spreads in different populations, detect inequities, and provide estimates of crucial metrics that can be used to understand any disparities, even those not caught by poorly-collected, inconsistent, or improperly reported data. Our model is an extension of *Evensen et al.*'s SEIR model [18] to a multi-population one. In this model, each population has an age stratified contact rate matrix with different risk factors for severe illness and death for each age group within it. In addition, each population has their own time dependent viral reproductive number $R_t(n)$ driving infections for the population. We assimilate COVID-19 deaths separated by race/ethnicity from the states of Alaska, California, Connecticut, Delaware, Hawaii, Maryland, Michigan, Utah, and Washington, as well as the District of Columbia. These locations are chosen on account of their more complete reporting of deaths by race/ethnicity. The data is assimilated using ESMDA resulting in a complete time series estimate of viral reproductive number, infections, hospitalizations, deaths, and other relevant static parameters for each racial/ethnic population within each location. The result of the data assimilation is referred to as the analysis. We approach the DA in three distinct ways. In the first approach, we assign the same age stratified contact rate matrices to each population as well as the same prior for the time dependent reproductive number, $R_t(n)$. Then we assimilate the daily death data for each population which adjusts its $R_t(n)$ so that the solution of the SEIR model better predicts the observed deaths. This method quantifies potential disparities between populations with no prior assumption on such. Populations whose analysis $R_t(n)$ curves show higher rates of spread over time indicate disparity, especially when they make up a smaller proportion of the total population. In the second approach, we make more assumptions on the causes of disparity between different populations. Specifically, that disparity exists when the bulk of a given population is employed in jobs where they cannot work from home, may live in smaller housing situations with inter-generational households, and may not have access to financial reserves or resources that facilitate the ability to self isolate. We build this assumption into the age stratified contact matrix by increasing the contact rates among working age adults as well as rates among older individuals with the other age groups for those populations. We refer to this matrix as the

Disproportionately Affected Population (DAP) matrix. The decision on whether or not a given population should be assigned this matrix is based only on the implication of disparity in the aggregate statistics as discussed in section 1.1. Each population is given the same prior for $R_t(n)$ which is allowed to be adjusted by the update step, up or down, by the same amount over the time interval we consider. With this piece-wise update, we allow for the uncertainty in $R_t(n)$ to be high, finding an analysis curve for $R_t(n)$ that establishes the bulk difference in the rate of spread for each population. With the third approach, we also use different contact matrices for any DAP, but we use the analysis $R_t(n)$ curves for each population computed in the piecewise update from approach two as a prior. Further, we use a lower uncertainty and apply a *time-continuous update* to $R_t(n)$. It is this time-continuous update that allows us to also update the driving rate of spread in the past to remove bias in the solution before it happens. It is this final approach for which we obtain parameters that produce solutions of the SEIR model most consistent with the data.

The outline of the paper is as follows: Section 1.1 provides our working definition of a DAP, Section 2 describes the SEIR model used, and in Section 3 we give a brief introduction to the ESMDA method itself. In Section 4 we describe how we choose the representative age-stratified contact matrices for a DAP and a Non-Disproportionately Affected Population (NDAP). We also show how the choice of those matrices affects model predictions with free (no DA) model runs. In Section 4.2 we provide the general setup for our ESMDA runs, and present the individual results for the different states we consider. In Section 5 we highlight the importance of the choice of prior estimates of model parameters specifically discussing the results for the State of Connecticut. Finally, in Section 6 we conclude by making an overall assessment of the results obtained across the modeled populations within the 9 states previously mentioned and the District of Columbia, as well as presenting some next steps.

1.1. **Classification of a disproportionately affected population (DAP).** In this work, we consider "Disproportionately Affected Populations" (DAPs) to be sub-populations of the overall population disproportionately affected by the pandemic than other sub-populations. Here our sub-populations will be comprised of the different racial/ethnic populations within a given locality.

We adopt the same criteria used in [7] to identify a DAP. A population is flagged for likely disparity when case counts or deaths meet the following criteria: (1) is at least 33% higher than the census percentage of total population, (2) remains elevated whether including or excluding cases/deaths with unknown race/ethnicity, (3) is based on at least 30 actual cases or deaths. If these criteria are met for a given population we will consider them a disproportionately affected population and refer to them as a DAP. Populations which meet none of the criteria above will be referred to as Non-Disproportionately Affected Populations (NDAPs). It is this criteria that we use to assign a DAP contact matrix with increased contact rates when applying the methodology of approach two or three outlined in the introduction.

It is important to note that these criteria can still fail to detect disparities. For example, one population may be flagged as a DAP in terms of confirmed cases while another may not and a disparity goes unnoticed due to a lack of testing or case reporting for that population. Using ESMDA techniques, we can estimate the actual infections in a population through the assimilation of the death data and detect some of these kinds of possible disparities. Indeed, as is discussed in Section 4.2.6, for the state of Connecticut we see that while the LatinX population is flagged

for a disparity in terms of confirmed cases, we find that the Black population may in fact make up more of the total infections despite representing 6% less of the total population. The Black population, however, was not flagged as having a high percentage of confirmed cases compared to the percentage of the total population. In this way a population that might not be considered a DAP from the aggregate statistics alone may be revealed to be one through the DA analysis.

We would also like to note that in total there are 9 possible racial/ethnic classifications reported in the data, White, Black, LatinX, American Indian and Alaskan Native (AIAN), Asian, Native Hawaiian and Pacific Islander (NHPI), Multiple, Other and Unknown. Each locality that we study may report some data from these populations differently or not at all, and in some cases they may even be double counted. This is particularly true with individuals in the multi-racial grouping. Some states my also count them as members of each grouping to which they belong. It may also be that a multi-racial person chooses one grouping to report themselves as. The Unknown grouping is more a measure of deaths not assigned to any group and is not considered other than to gauge uncertainty as it is not a racial/ethnic classification.

2. **Model.**

2.1. **Evensen *et al.* original (one population) model.**

The original Evensen *et al.* model [18] has n_a compartments/equations for susceptible (S_i), exposed (E_i), and infectious (I_i) individuals in each age group, with $i = 1, \ldots, n_a$. The remaining compartments and equations are for the following infected individuals: quarantined individuals with mild (Q_m), severe (Q_s), or fatal (Q_f) symptoms; hospitalized individuals with severe (H_s) or fatal (H_f) symptoms; individuals with fatal symptoms who are put into a care home (C_f); recovered individuals who had mild (R_m) or severe (R_s) symptoms; and dead individuals (D).

An infectious individual from the age group j (I_j) can infect a susceptible individual in any age group i (S_i). They will do so at a rate $R_{ij}(t)$, which highlights the different contact/infection rates between each age group. These values compose the matrix $\hat{\mathbf{R}}$. We also have a scalar function $R(t)$ (note that this is equivalent to the function $R_t(n)$ mentioned previously with $n = 1$ and with simplified notation) which is the effective reproductive number at time t. Multiplying these two quantities together gives $\mathbf{R}(t) = R(t)\hat{\mathbf{R}}$. This matrix is then re-scaled so that $R(t)$ solely determines the effective reproductive number at any time. The system runs through three different intervention periods: (1) "pre-lockdown" (before any mitigation measures), (2) "lockdown" (when people were asked or required to stay home and avoid gathering), and (3) "post-lockdown" (when a state began to reopen businesses and lessen restrictions on gatherings). For each of these time periods, we can have different $\hat{\mathbf{R}}$ matrices, $\hat{\mathbf{R}}_i$ to reflect different contact rates between different age groups in each time period.

We have the following parameters for this model: the proportion of the population that has mild symptoms and is from age group i (p_m^i), the proportion of the population that has severe symptoms, and is from age group i (p_s^i), the proportion of the population that has fatal symptoms and is from age group i (p_f^i), the proportion of infected individuals with fatal symptoms who are hospitalized (p_h), the average length of the incubation period (τ_{inc}), the average length of the infectious period (τ_{inf}), the average length of hospitalization (τ_{hosp}), the average recovery time for individuals with mild symptoms (τ_{recm}), the average recovery time for individuals

with severe symptoms (τ_{recs}), and the average time to death for individuals with fatal symptoms (τ_{death}).

2.2. **Multi-population model extension.**

2.2.1. *New model assumptions & clarifications.* Now we employ a modified version of the original Evensen *et al.* model, extending it from a single population of interest to n_{c} populations and n_{a} age classes within each population. Here, different populations could represent different countries, states, localities, or a specific population within a locality. For this work we consider each population to be the different racial/ethnic populations in a state. In Equation 1, we use the following notation: n, m are indices running over all n_{c} populations and i, j are indices running over all n_{a} age groups within each population. The total population of any two populations, n and m, are denoted by N_n and N_m respectively.

We model the interaction between groups using the elements of $\mathbf{R}^{\text{C}} \in \mathfrak{R}^{n_{\text{c}} \times n_{\text{c}}}$ (the diagonal must always be 1) and $\mathbf{R}^{\text{A}} \in \mathfrak{R}^{n_{\text{a}} \times n_{\text{a}}}$. The effective reproductive number per population is a scalar function of time, $R_t(n)$, and is a parameter that we estimate. We model the relative differences in infectiousness between age groups using the coefficients in $R^A_{ij}(n)$, which can differ between populations. The model default is "null", such that all elements are set to $R^A_{ij}(n) = 1$, which assumes equal transmission rates among all age groups. In Equation 1, we use the $R^A_{ij}(n)$ for age group i as it interacts with age group j. The only sound alternative to this choice would be to set $R^A_{ij}(n) = 1$ when $m \neq n$, given the number of coefficients we would otherwise need to specify. An example of such a matrix $R^A_{ij}(n)$ can be seen in Figure 1.

The fractions of mildly, fatally, and severely ill (p_{m}, p_{f}, and p_{s}, respectively) can differ between populations. We have used the same hospitalization fraction of fatally ill, p_{h}, for all populations.

In the case with only one population $n = m = 1$, we have $N_m/N_m = 1$, as well as $\mathbf{R}^{\text{C}}(n, m) = 1$. Thus, Equation 1 reduces to the standard SEIR model as in Section 2.1. The use of a multi-compartment model only changes the nonlinear interaction term present in the $\mathbf{S}_i(n)$ and $\mathbf{E}_i(n)$ subequations. Outside of this interaction term, each population evolves independently of every other.

$$\frac{\partial \mathbf{S}_i(n)}{\partial t} = -\sum_{m=1}^{n_c} \frac{N_{\text{m}}}{N_n} R^C_{nm} R_t(n) \left(\sum_{j=1}^{n_a} \frac{R^A_{ij}(n) \mathbf{I}_j(m)}{\tau_{\text{inf}}} \right) \mathbf{S}_i(n) \tag{1a}$$

$$\frac{\partial \mathbf{E}_i(n)}{\partial t} = \sum_{m=1}^{n_c} \frac{N_{\text{m}}}{N_n} R^C_{nm} R_t(n) \left(\sum_{j=1}^{n_a} \frac{R^A_{ij}(n) \mathbf{I}_j(m)}{\tau_{\text{inf}}} \right) \mathbf{S}_i(n) - \frac{1}{\tau_{\text{inc}}} \mathbf{E}_i(n) \tag{1b}$$

$$\frac{\partial \mathbf{I}_i(n)}{\partial t} = \frac{1}{\tau_{\text{inc}}} \mathbf{E}_i(n) - \frac{1}{\tau_{\text{inf}}} \mathbf{I}_i(n) \tag{1c}$$

$$\frac{\partial \mathbf{Q}_{\text{m}}(n)}{\partial t} = \sum_{i=1}^{n_a} \frac{p^i_{\text{m}}(n)}{\tau_{\text{inf}}} \mathbf{I}_i(n) - \frac{1}{\tau_{\text{recm}}} \mathbf{Q}_{\text{m}}(n) \tag{1d}$$

$$\frac{\partial \mathbf{Q}_{\text{s}}(n)}{\partial t} = \sum_{i=1}^{n_a} \frac{p^i_{\text{s}}(n)}{\tau_{\text{inf}}} \mathbf{I}_i(n) - \frac{1}{\tau_{\text{hosp}}} \mathbf{Q}_{\text{s}}(n) \tag{1e}$$

$$\frac{\partial \mathbf{Q}_{\mathrm{f}}(n)}{\partial t} = \sum_{i=1}^{n_a} \frac{p_{\mathrm{f}}^{i}(n)}{\tau_{\mathrm{inf}}} \mathbf{I}_i(n) - \frac{1}{\tau_{\mathrm{hosp}}} \mathbf{Q}_{\mathrm{f}}(n) \tag{1f}$$

$$\frac{\partial \mathbf{H}_{\mathrm{s}}(n)}{\partial t} = \frac{1}{\tau_{\mathrm{hosp}}} \mathbf{Q}_{\mathrm{s}}(n) - \frac{1}{\tau_{\mathrm{recs}}} \mathbf{H}_{\mathrm{s}}(n) \tag{1g}$$

$$\frac{\partial \mathbf{H}_{\mathrm{f}}(n)}{\partial t} = \frac{p_{\mathrm{h}}}{\tau_{\mathrm{hosp}}} \mathbf{Q}_{\mathrm{f}}(n) - \frac{1}{\tau_{\mathrm{death}}} \mathbf{H}_{\mathrm{f}}(n) \tag{1h}$$

$$\frac{\partial \mathbf{C}_{\mathrm{f}}(n)}{\partial t} = \frac{(1-p_{\mathrm{h}})}{\tau_{\mathrm{hosp}}} \mathbf{Q}_{\mathrm{f}}(n) - \frac{1}{\tau_{\mathrm{death}}} \mathbf{C}_{\mathrm{f}}(n) \tag{1i}$$

$$\frac{\partial \mathbf{R}_{\mathrm{m}}(n)}{\partial t} = \frac{1}{\tau_{\mathrm{recm}}} \mathbf{Q}_{\mathrm{m}}(n) \tag{1j}$$

$$\frac{\partial \mathbf{R}_{\mathrm{s}}(n)}{\partial t} = \frac{1}{\tau_{\mathrm{recs}}} \mathbf{H}_{\mathrm{s}}(n) \tag{1k}$$

$$\frac{\partial \mathbf{D}(n)}{\partial t} = \frac{1}{\tau_{\mathrm{death}}} \mathbf{H}_{\mathrm{f}}(n) + \frac{1}{\tau_{\mathrm{death}}} \mathbf{C}_{\mathrm{f}}(n) \tag{1l}$$

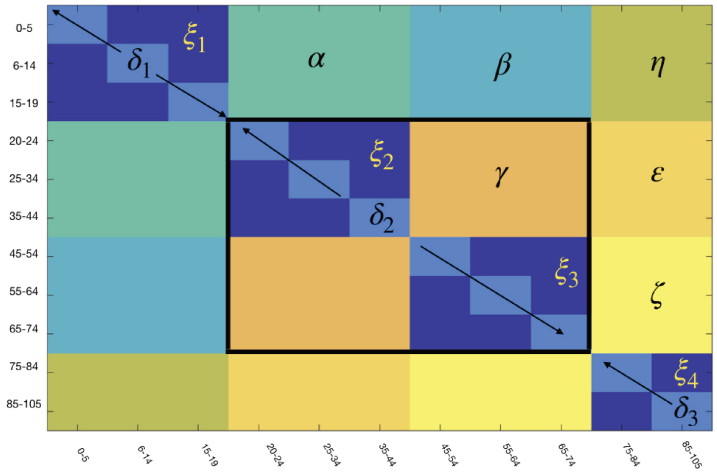

FIGURE 1. The general form of the $R_{ij}^{A}(n)$ contact matrix elements for contact rates between age groups i and j in a given sub-population. The contact matrix is then subdivided into different blocks where parameters $\alpha, \beta, \eta, \gamma, \varepsilon, \zeta, \delta_1, \delta_2, \delta_3, \xi_1, \xi_2, \xi_3$ and ξ_4 control the contact rates between different age groups which generate similar patterns for spreading the disease. In particular, we define γ, δ_2, ξ_2 and ξ_3 to be the parameters for the contact rates of the working class age groups.

3. **Methodology.** A common issue with solving high-dimensional models is the difficulty in accurately estimating parameters. However, optimal estimates can be found using data assimilation methods. In particular, we highlight the use of an ensemble Kalman Filter (EnKF method) for sequential data assimilation. In recent history, ensemble data assimilation has been widely implemented in weather predicting [22], tumor growth and spread [25], and petroleum reservoir history matching e.g. [29]. The existing theory and application in these fields have seen the implementation of ensemble data assimilation in other fields, such as epidemiology. In

particular, we take inspiration from Evensen *et al.*'s use of an ensemble smoother with multiple data assimilation (ESMDA) in their extended SEIR model of COVID-19 upon which we build [18].

3.1. **The inverse problem.** Ensemble smoother techniques can be derived by assuming a perfect forward model.

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) \tag{2}$$

In general, $\mathbf{x}$ is the realization of model parameters, and $\mathbf{y}$ consists of the uniquely predicted measurements. For the case of COVID-19, $\mathbf{x}$ consists of the initial conditions, parameters, and time-reliant effective reproductive numbers. We relate the predictions $\mathbf{y}$, to the parameters $\mathbf{x}$ through the model operator $\mathbf{g}(\mathbf{x})$, where $\mathbf{g}$ is the model in Equation 1.

$$\mathbf{d} \leftarrow \mathbf{y} + \mathbf{e} \tag{3}$$

In our model, $\mathbf{y}$ then consists of the predicted measurements of deaths, cases and hospitalizations given some model error, $\mathbf{e}$. Here, $\mathbf{d}$ is the observed data. To solve the inverse problem, it is efficient to frame it into an equation using Bayes' theorem:

$$f(\mathbf{x} \mid \mathbf{d}) \propto f(\mathbf{d} \mid \mathbf{g}(\mathbf{x}))f(\mathbf{x}) \tag{4}$$

Equation 4 represents the so-called smoothing problem, which can be approximated using ensemble methods. We refer the reader to Evensen's analysis of solving inverse problems for a full derivation and well-constructed analysis in [17]. We also note that Equation 4 is also the basis of sequential filtering problems.

3.2. **ESMDA.** The assimilation method used is an iterative ensemble smoothing method called an ensemble smoother with multiple data assimilation. The method, which is similar to an ensemble smoother, solves the parameter-estimation problem and is formally derived from the Bayesian formulation using a tempering procedure [30]. What separates this method from other similar methods is that it approximates the posterior recursively, gradually introducing information to alleviate the impact of nonlinear approximation. After updating through to the last time step, it begins the assimilation process over again - resampling the perturbation in the vector of perturbed observations, which reduces sampling error [16]. The simplicity and effectiveness of ESMDA are what make it an optimal assimilation method for this particular application.

For simplicity, we lay out the ensemble methods without the mathematical details, as laid out in [18].

- First, sample a large ensemble of realizations of the prior uncertain parameters (age groups, the functions $R_t(n)$, and the initial infected and exposed), given their prescribed first-guess values and standard deviations.
- Integrate the ensemble of model realizations forward in time to produce a prior ensemble prediction, which also characterizes the uncertainty.
- Compute the posterior ensemble of parameters by using the misfit between prediction and observations, and the correlations between the input parameters and the predicted measurements.
- Finally, compute the posterior ensemble prediction by a forward ensemble integration. The posterior ensemble is then the "optimal" model prediction with the ensemble spread representing the uncertainty.

4. **Case studies.** In this section, we highlight how the model behaves when using different age-stratified matrices for two populations, one DAP and one NDAP. During the different intervention periods, we make the assumption that DAPs will have increased contact rates to varying degrees across all age groups. The matrices are described below. We then describe our assumptions and various approaches to the ESMDA problem and interpret the results obtained for each state in our study. First we make some remarks on the model parameters and data chosen for assimilation.

For the model parameters explained in Section 2, we use the same initial guesses as in [18] and show them in Tables 8 and 9 in Appendix A. The initial values for theses parameters are based off values obtained from available data and some initial model tuning experiments. However, the DA will fine tune these parameters if necessary. We use the same initial parameters for all populations including the p-numbers and case fatality rates. While there is some evidence that different racial/ethnic populations may have more members with underlying conditions, this is also true of members of the same population who live in different regions. We do not find enough available data to make confident guesses in any differences in these parameters amongst the different populations or localities. We also believe the differences to be relatively small and that any discrepancies can be accounted for by a small increase or decrease in $R_t(n)$ as estimated by the ESMDA scheme. In general, we find that for our best obtained results, these initial parameters are good estimates and that $R_t(n)$ is the primary differentiator between populations. It should be noted though that some populations such as the AIAN population in Alaska may not have as much access to needed medical care in Arctic rural areas. This could mean a higher CFR and lower $R_t(n)$ is warranted, however the DA will still detect discrepancies through $R_t(n)$.

For this work, we choose only to assimilate death data for each of the racial/ethnic populations. While the model and ESMDA scheme can assimilate case counts and hospitalizations, we find this data to currently be of too poor quality for assimilation. Case counts can be very misleading when the actual percentage of cases captured is unknown and changing through time while hospitals may not report racial/ethnic data at all or quickly enough. However as deaths are recorded by each local government and each reports a deceased individual's race, there is more consistency in the reporting. At the time of writing this manuscript, the race/ethnic data set is somewhat incomplete for most states and is changing as the COVID-19 pandemic unfolds. We choose states for which there is at least 93% reporting up to January 3, 2021, for our analysis. We also increase the uncertainty in the death data to account for unknown cases while keeping it low enough for the ESMDA to actually make informative updates. We believe more analysis should be done when the data is more complete; for example, when more data, such as hospitalizations, can be made readily available. However, we believe there is still much that can be learned in this analysis. The data that we use is compiled by The Covid Tracking Project who themselves compile the data from local government authorities [7].

4.1. **Non-DA runs.** We begin our simulations by working with two "toy" populations to see how the model performs with no DA. For clarity, this run is not assimilating populations by race/ethnicity. This is simply a run to test that our multi-population model is working as expected given certain initial conditions. To this effect, we consider two populations with the following initial conditions: population 1 has an initial exposed of 500, initial infected of 350, and a case fatality ratio (CFR) of 0.009; population 2 has an initial exposed of 1000, initial infected

of 700, and a CFR of 0.01. We use the following parameters for Equation 1 in our simulation, $\tau_{\text{inf}} = 3.8$, $\tau_{\text{inc}} = 5.5$, $\tau_{\text{recm}} = 14$, $\tau_{\text{recs}} = 5$, $\tau_{\text{hosp}} = 6$, and $\tau_{\text{dead}} = 16$. We use a continuous function for $R_t(n)$ with data from [9]. For this run, we use the time varying estimates of the effective reproductive number for the state of Utah obtained from [9] (see bottom left plot of Figures 3, 4). The interactions between age groups are taken into account in the $R^A_{ij}(n)$ matrix entries (see Table 1 and Figure 2). Our decisions on the chosen values are based on assumptions that population 2 will be a DAP while population 1 will be a NDAP. We chose two scenarios with different proportions of the DAP and NDAP present in the total population.

For the non DA runs, each matrix in (i), (ii) of Figure 2, is used to indicate the different "intervention periods" in the model: top ="pre-lockdown", middle ="lockdown", and bottom ="post-lockdown"; (i) represents contact matrices for a NDAP and (ii) represents contact matrices for a DAP. For both (i) and (ii), the top matrices are the same. We used a contact matrix from a statistical survey of Europe [18] to describe the transmission between different age groups as a basis for such matrices. In Table 1, we note the parameter values that were used to scale the top matrix for each intervention period for both the DAP and the NDAP. During the lockdown intervention period, contact rates are decreased across all age groups, but more so for school-age children and older adults, in line with school closures and assumed caution amongst the most vulnerable populations. After a "re-opening" (i.e., entering the "post-lockdown" period), rates are increased across all age groups but more so for working-age adults and children, while assuming caution remains amongst the older population. For a DAP, contact rates between working adults and older adults are higher than for a NDAP. This builds in the assumption that DAP jobs tend to be in places where exposure is more likely (grocery stores, construction, etc) and that there are more inter-generational households connecting working-age adults and children with older adults in such populations. Identifying the scaling parameters is a difficult task and more data would be needed to directly estimate them. We take an ad-hoc approach to estimation but note that the normalization of R^A_{ij} means that the primary control of the rate of spread is in the scaling factor $R_t(n)$ for the DA runs in Section 4.2. We choose scaling factors which provide differentiation in the results of the non DA runs between the populations without being too extreme. In (iii) we see the inter-population contact matrix between our two populations where the diagonal elements, R^C_{nn}, are 1 and the off-diagonal elements are very small ($<< 1$). The results show that the model takes into effect the population size as well as the contact matrices.

In Figure 3, both populations have about the same number of people. Since we are not assimilating data for these runs, we have only included the averages in the plots (which in this case is the same as if we were simply integrating a basic SEIR model). In all plots, the blue curves represent the NDAP while the red is the DAP. We can clearly note that the total amount of projected deaths and cases for the DAP tend to be higher since the intra-population contact rates for each intervention period are also higher for the DAP. In the bottom-left figure, we use data on the state of Utah as the effective reproductive number $R_t(n)$ for this simulation (black curve) retrieved from [9]; the shaded gray area around the curve represents the cone of uncertainty provided by the site [9]. The blue and green curves represents $R_t(n)$ (they are the same in this case since we have no DA). In Figure 4, colors have the same meaning as in Figure 3. The NDAP represents a significant amount of the population (about 70%) in this scenario. The total amount of projected deaths and

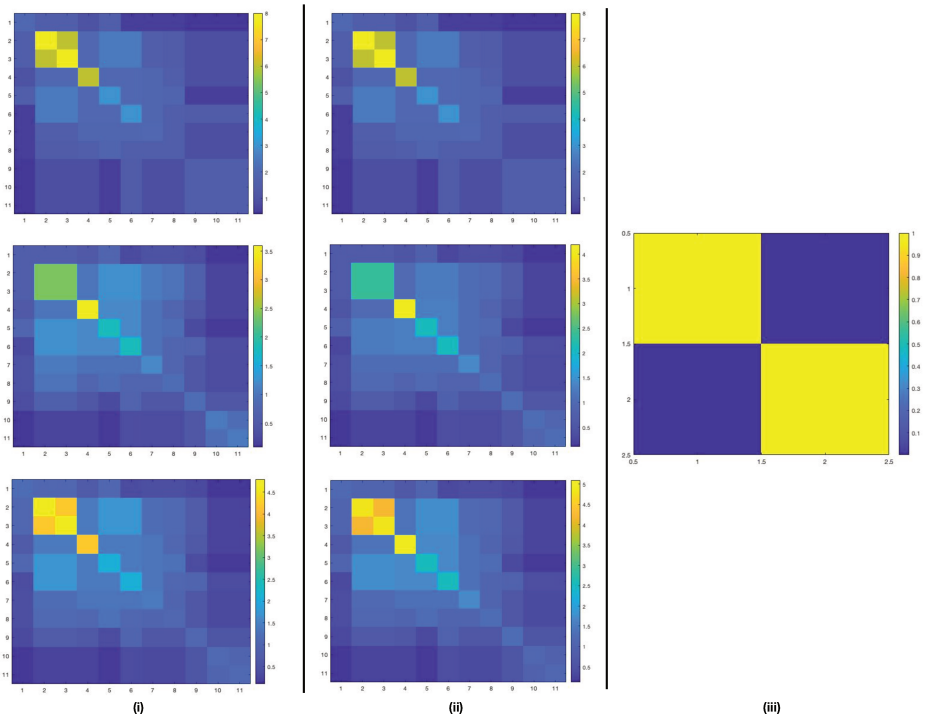

FIGURE 2. $\mathbf{R}^{\mathrm{A}}$ for the NDAP (i) and DAP (ii) in our runs without DA for each intervention period along with the $\mathbf{R}^{\mathrm{C}}$ matrix (iii) used in these simulations. We also used the $\mathbf{R}^{\mathrm{A}}$'s in (i) and (ii) for the DA runs.

cases for the DAP is higher until June when projected deaths and cases for the NDAP surpasses it, but not by much.

Scalings	Matrix Parameters												
	α	β	γ	η	ϵ	ζ	δ_1	δ_2	δ_3	ξ_1	ξ_2	ξ_3	ξ_4
DAP lock	0.5	0.7	0.55	0.25	0.25	0.35	0.3	0.7	0.7	0.4	0.65	0.55	0.6
DAP post-lock	0.7	0.8	0.7	0.3	0.3	0.4	0.6	0.85	0.7	0.85	0.7	0.65	0.65
NDAP lock	0.5	0.7	0.5	0.2	0.2	0.3	0.3	0.6	0.7	0.4	0.6	0.5	0.6
NDAP post-lock	0.7	0.8	0.7	0.25	0.25	0.35	0.6	0.7	0.75	0.7	0.7	0.65	0.65

TABLE 1. Matrix scaling parameters for DAP and NDAP workers in lockdown and post-lockdown time periods.

4.2. **DA runs.** In this section we discuss the sources of the data used in our analysis, explain our choices for inter-population contact matrices, define our experimental setup, and present the results of the ESMDA analysis performed on several states for which racial data reporting was at least 93% or above [7].

4.2.1. *Data.* We choose states which show very apparent disparities between populations (e.g. according to an analysis of disproportionate effects of COVID-19 on racial/ethnic populations in [7]) as well as some for which disparities are less apparent in the aggregate data. Initial conditions on the number of exposed and infected for

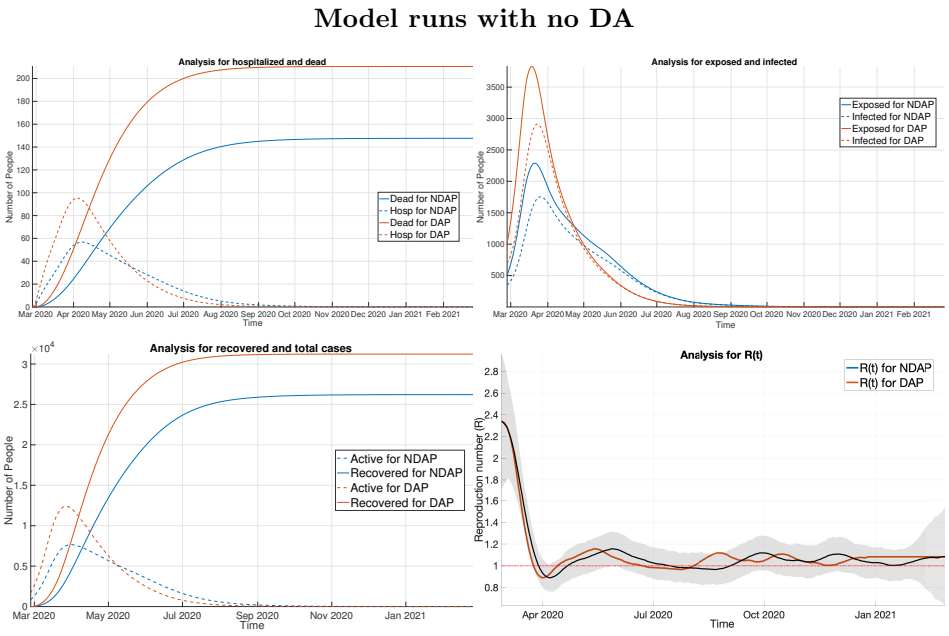

FIGURE 3. Results for non-DA runs with DAP and NDAP at about the same population.

each racial/ethnic population and state are estimated using data from the Institute for Health Metrics and Evaluation (IHME) [5] which provides an estimate of actual cases for each day of the pandemic. We take this estimate from the start date of our simulations and scale it by the proportion of the total population each population represents. This provides an initial guess for the number of infected and then we double this for the number of exposed. We note that these initial conditions are also fine tuned by the ESMDA algorithm. Our population by age group data for each locality comes from the 2018 U.S. Census Bureau estimates. We use the age groupings described in [18], however, this type of data is not available in each racial/ethnic category. As an estimate we take the age grouping data from the specific locality and scale it by the proportion of the total population each population represents. This is not completely ideal, but serves as a reasonable estimate. We also note that some population percentages for a given state may add up to more than 100%. This could be due to double counting in some cases. In the end, population estimates will have inaccuracies in general and we believe these estimates are sufficient for the current work. We anticipate more accurate information on age groups coming from the 2020 U.S. census when it is made available.

4.2.2. *Inter population contact matrices.* In our analysis, we employ three different inter-population matrices for each of the three intervention periods as in Table 2. In the first period before any mitigation measures, $R^C_{nm} = 0.001$ for $n \neq m$; during the lockdown periods, we reduce the off-diagonal terms by an order of magnitude to $R^C_{nm} = 0.0001$; and finally after the lockdown, we take the off-diagonal terms to be $R^C_{nm} = 0.0005$, assuming caution in the population. We choose the off-diagonal

FIGURE 4. Results for non-DA runs with NDAP at 70% of the total population.

mixing terms to be fairly small by the following reasoning. Most of the spread of this virus occurs during mask-less, indoor, and sustained close contact. This implies that, especially during and after the lockdown periods, most of the spread will happen in the home or group settings where precautions are not taken or are infeasible to maintain, an assumption supported by a recent CDC study [21]. This implies that for the majority of the pandemic, members of these specific racial populations are not often having the type of contact conducive to spreading SARS-CoV2 with people outside of their family nor their group of close friends, both of which are likely to consist of members of the same racial/ethnic population. This also implies that these populations may in fact evolve somewhat independently of each other with transmissions between racial populations being less common than within populations. To illustrate, a member of a particular population may become exposed at a place with other populations present, i.e. work or large gathering. This would represent one new infection for this individual and their population. However, they may then spread infection to several immediate or extended family members who likely share race/ethnicity. Thus, one infection across populations can become many infections within the population. This effect would be amplified during times when people are taking precautions when outside of the home.

In practice, when repeating our experiments over a range of values for $10^{-5} \leq R^c_{nm} \leq 1$, we find that we can best explain the data in the range of 10^{-4} to 10^{-2} as measured by the χ-squared statistic. In addition, the $R_t(n)$ curves for each population retain the general trend of the initial prior, which is desirable. For values of large mixing ($\geq 10^{-1}$) the populations with the highest number of infections per capita tend to drive the dynamics with the other populations having values for $R_t(n)$ well below the exponential threshold of $R = 1$. This is unrealistic as it would imply

one population is primarily responsible for the majority of transmission which is incompatible with the fact that most of the transmission happens in the home or close friend or work settings.

This effect can be understood through equations (1a) and (1b). Large values for R^C_{nm} would require a large reduction in $R_t(n)$ to scale down the transmissions from population m (with a large number of infections) to population n (with far fewer infections) so that they are consistent with the number of deaths in the data. Likewise, if R^C_{nm} is too small, $R_t(n)$ may be driven up. However, if a particular population is a very small proportion of the total population, then an $R_t(n)$ below $R = 1$ is realistic as interactions with their own population are less likely and interactions with other populations would be the primary driver of spread. As a result, $R_t(n)$ would not be driven down just because of a large mixing value. We illustrate some of these differences in Figure 5. We also note that parameters that are well understood, such as the CFR, must change somewhat significantly between populations as well in order to fit the data for large mixing values. For all of the reasons above we keep the off-diagonal mixing values around the order of 10^{-4} to 10^{-3}.

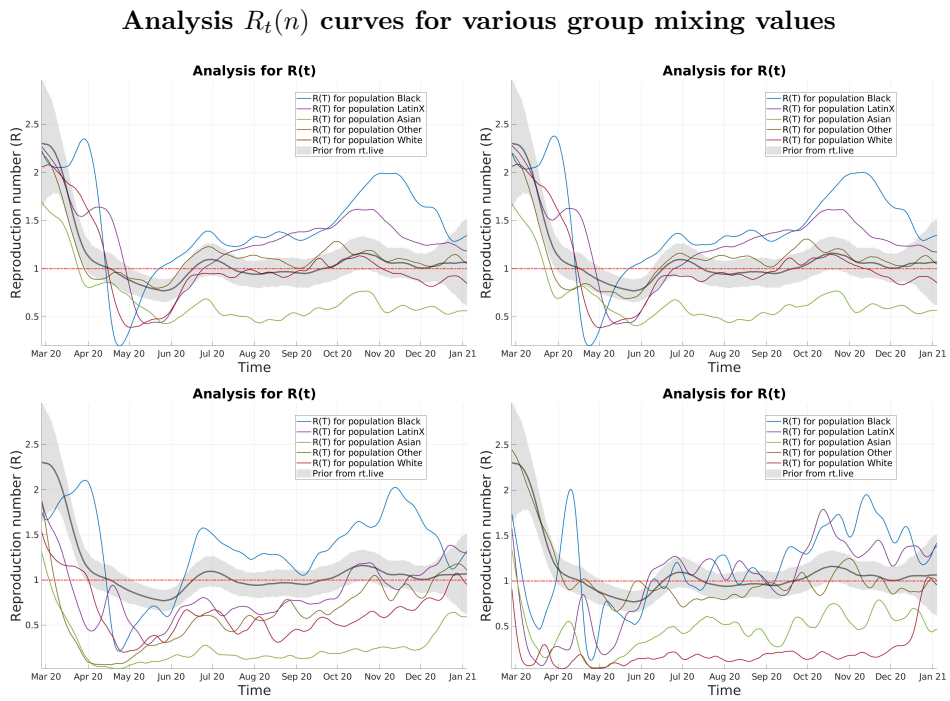

FIGURE 5. Examples of analysis runs for various values of R^C_{nm} from the District of Columbia (DC). Top: left $R^C_{nm} = 0$, right $R^C_{nm} = 10^{-3}$. Bottom: left $R^C_{nm} = 10^{-1}$, right $R^C_{nm} = 6 \times 10^{-1}$.

4.2.3. *Experimental setup.* We take three approaches to our DA analysis. In approach one, we use age contact matrices of all ones ($R^{\mathrm{A}}_{ij}(n) = 1$ for all age groups i, j and populations $1, \ldots, n$) which would remove any possible effect resulting from

differences between the age group contact matrices of the DAPs and the NDAPs. We do this in an effort to detect any possible disparities among populations without making prior assumptions of how their contact rates might differ. We further assume the same piece-wise prior for each population on the effective reproductive scaling factor $R_t(n)$ with large uncertainty ($\sigma = 3$), allowing the ESMDA smoothing to adjust this based on the data. The prior itself is taken from rt.live [9] for each state that we study. By piece-wise we mean that the update to the $R_t(n)$ prior is done on the whole, shifts up or down for the curve given rather than time dependent adjustments. As a result, we would expect the recovered $R_t(n)$ for each population to be somewhat stratified with larger values on average belonging to the DAPs. This is shown in Figure 6 for the District of Columbia where we see a stratification indicating the Black and LatinX populations having the largest values for $R_t(n)$. We do note the somewhat poor data fit when doing piece-wise updates, this is remedied when allowing for a time-continuous update to $R_t(n)$ discussed below.

Case: Piece-wise updates to $R_t(n)$ (DC)

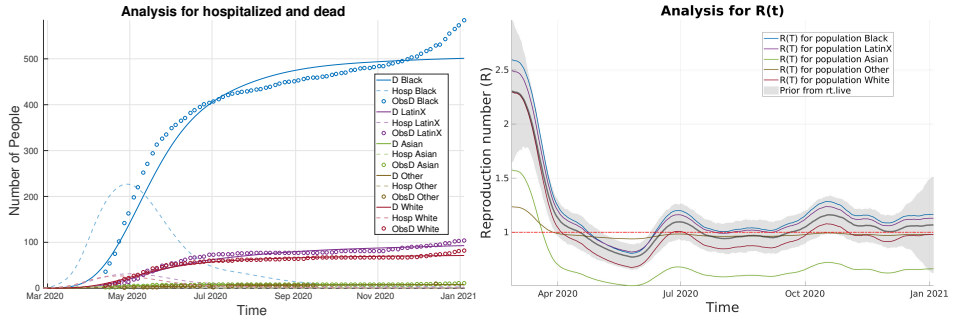

FIGURE 6. Analysis results for the District of Columbia (DC) with $R_{ij}^A = 1$ and piece-wise updates to $R_t(n)$.

In approach two, we repeat the same experiment as above with the exception of using different age contact matrices. Here we use the DAP and NDAP matrices described in Section 4.1. We assign the DAP matrices to populations that meet the criteria for disparity as described in covidtracking.com [7], also outlined in Section 1.1. With approach two, using DAP and NDAP contact matricies, we typically expect the stratification to be somewhat less dramatic as the increased contact rates among working age and elderly groups in the DAPs can explain increased transmissions and deaths without higher values for $R_t(n)$, which here is also updated piece-wise. An example of these runs is shown in Figure 7 where we see a stratification among populations with the analysis $R_t(n)$ values typically a bit lower than for the case without age stratification. This is because of the increased contact rates among adults in the age-stratified matrices, who are more likely to suffer death, as compared to that of children, who are less likely to suffer death from SARS-Cov2. This means that less spread is needed to account for the number of deaths in the data.

This effect can also be understood through equations (1a) and (1b), specifically the product $R_{ij}^A(n)\mathbf{I}_{ij}(m)$ when $n = m$. If contact rates for working-age and older individuals are higher, the number of infections attributed to those age groups is increased. As a result, the scaling factor $R_t(n)$ may need to be reduced if those infections are overestimated and inconsistent with the number of deaths in the data.

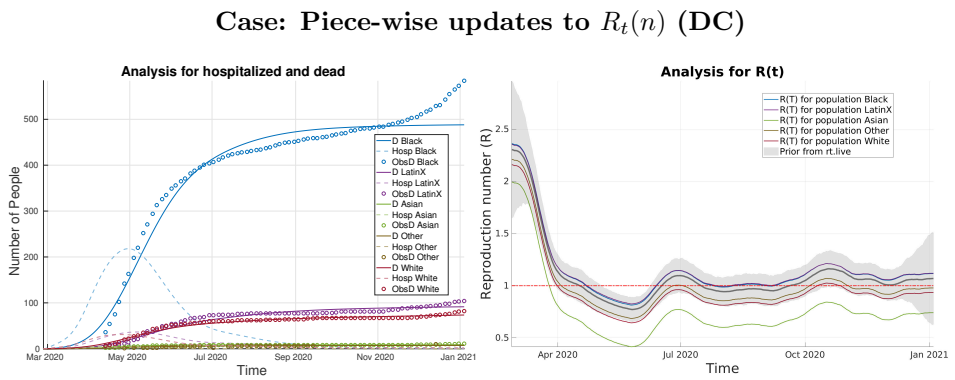

FIGURE 7. Analysis results for the District of Columbia (DC) with age stratified R^A and piece-wise updates to $R_t(n)$.

The specific contact rates are difficult to determine and a study to compute them is encouraged.

Finally, we employ approach three where we allow for a time-continuous estimation of $R_t(n)$ using a decorrelation length of 10 days, the CDC recommended time that an individual should quarantine after exposure [6]. Here the priors for $R_t(n)$ used for each population are taken directly from the analysis $R_t(n)$ that comes from approach two. We decrease the uncertainty to $\sigma = 0.5$ and employ the DAP and NDAP matrices in the same way as previously described in Section 4.1.

For all three approaches, we have three intervention periods where we begin with large contact rates at the beginning of the pandemic, severely reduced contact rates during a given state's mitigation (lockdown) period, and slightly reduced contact rates after a given state re-opening date. How we scale these rates for each age group is discussed in Section 4.1 and given in Table 1. The dates, in addition to the starting date of each simulation, for each of the states we study can be found in Table 2. The start date is chosen to be the first day given by our prior which comes from rt.live. The date of the first intervention is chosen when a state began to close schools and the beginning of the introduction of the White House Coronavirus Task Force's initial "15 days to slow the spread" campaign which prompted many Americans to take precautions ahead of their individual states mandates.

To briefly summarize, we take three approaches to our DA experiments. In approach one, with results shown in appendix B, we are agnostic to the differences in contact rates between age groups among the different racial/ethnic populations as well as government interventions periods during safety restrictions. For this run, the contact matrices for all populations are comprised of all ones for the entire integration window. In approach two, with the results also shown in appendix B, we use contact matrices with different rates of contact between age groups which decrease during times of safety restrictions. In addition, populations which meet the criteria to be considered DAPs are assigned contact matrices with slightly increased contact rates among working age individuals as well as working age and senior individuals. This is to build in the *assumption* that members of DAPs would likely have employment in jobs where work form home is not possible and have a larger proportion of inter-generational households. In approach three we use the analysis

results for $R_t(n)$ from approach two as priors and allow for time-continuous updates to $R_t(n)$. This allows us to correct bias in the solution before it happens and provides results which best fit the available data.

In the following sections we discuss the results obtained for each state and also present the results from approach three. We make observations and at times speculate on the causes of what we see in the analysis results with supporting information when available.

Interventions	Information on Intervention Periods by State and the District of Columbia									
	AK	CA	CT	DC	DE	HI	MD	MI	UT	WA
Start date	3/8/20	2/25/20	2/27/20	2/27/20	3/1/20	2/28/20	2/25/20	2/21/20	2/29/20	1/9/20
1st Phase	3/19/20	3/19/20	3/19/20	3/17/20	3/17/20	3/17/20	3/17/20	3/17/20	3/17/20	3/17/20
2nd Phase	5/27/20	5/27/20	5/27/20	5/29/20	5/31/20	5/15/20	5/15/20	5/19/20	5/5/20	5/28/20

TABLE 2. Date breakdown by intervention periods for all states and the District of Columbia.

Race/Ethnicity	Percentage of Population per State and the District of Columbia									
	AK	CA	CT	DC	DE	HI	MD	MI	UT	WA
AIAN	**0.15**	0.076	X	X	X	X	X	0.005	0.023	**0.01**
Asian	0.06	0.14	0.04	0.0435	0.045	**0.38**	0.06	0.035	0.038	0.08
Black	0.03	0.06	0.1	**0.4453**	0.22	0.02	0.29	**0.14**	0.021	0.04
LatinX	X	**0.39**	**0.16**	**0.113**	**0.09**	X	**0.1**	X	0.142	**0.13**
Multi	X	X	**0.02**	X	0.02	X	X	X	X	0.05
NHPI	**0.01**	0.039	X	X	X	**0.1**	X	X	0.016	**0.008**
Other	0.08	X	0.01	0.01	X	0.24	X	0.03	0.01	0.005
White	0.65	0.37	0.67	0.4196	0.62	0.25	0.51	0.78	0.78	0.69

TABLE 3. Demographic breakdown by race/ethnicity (where data is available) for all states and the District of Columbia. Groups that meet the criteria to be a DAP are in bold. AIAN = American Indian and Alaska Native, NHPI = Native Hawaiian and Pacific Islander.

4.2.4. *The state of Alaska (AK).* Table 3 displays a population breakdown by race/ethnicity for all states that we study. We will frequently refer to this table to make comparisons between population statistics and the model outputs such as the number of exposed, recovered, infected, hospitalized individuals, and deaths. Furthermore, Figure 36 displays various analyses among these populations in the state of Alaska. Immediately striking is a major disparity revealed in the number of deaths between the White and American Indian and Alaskan Native (AIAN) populations. The White population makes up about 65% of the total population of the state while the AIAN population is only about 15%, yet they have comparable numbers of deaths. This can be seen in the analysis for hospitalized and dead plots in Figure 36. This disparity may be due to the large proportion of the AIAN population which lives far north of the Arctic Circle in regions where it is desirable to be indoors (where transmission is more likely) most of the year, particularly in the winter months. There is also far less medical infrastructure in these regions which can contribute to higher death rates. There is also a noticeable disparity amongst the Native Hawaiian and Pacific Islander (NHPI) population which makes up only about

1% of the population. In the analysis of exposed and infected plots of Figure 36, the number of infections of this population is above both that of the Asian and Black populations which make up 6% and 3% of the total population respectively. In the same three figures, we also see that the number of total infections–as estimated by the analysis up to the last data point on January 3, 2021–for the NHPI population is higher than that of the Black population. We also see evidence of these disparities in the analysis of the $R_t(n)$ functions for all three types of runs. In the ensemble of $R_t(n)$ plots for the first two approaches with piece-wise updates and large uncertainty, we see that in Figure 18 (Appendix B) stratification of the curves occurs with the most affected populations above the least effected populations for the majority of the integration interval. In the first scenario, where all entries of the age contact matrices are equal to one, we observe slightly higher values for $R_t(n)$ than in the second case where we employ the age-based contact matrices described in Section 4.1. This is due to the higher contact rates in the contact matrices for more susceptible age groups (adults) in the second case. Otherwise, $R_t(n)$ must be larger to account for the rate of spread. It is also notable that in the first two cases, the rapid increase in death during the winter surge is somewhat missed by the analysis.

In the most realistic case where we allow for a continuous update of $R_t(n)$ with a 10-day decorrelation length, this winter surge is captured and we see a far richer difference in the analysis $R_t(n)$'s between populations. For the White population, we notice a significantly higher $R_t(n)$ from the start of the pandemic through mid-April, which suggests that this was the first population affected, likely because the virus would have arrived in larger cities, such as Anchorage, where a large proportion of the state's White population resides. After that time, we see $R_t(n)$ typically below $R = 1$ while spread becomes exponential for most of the other racial/ethnic populations once the virus reaches beyond cities, into the regions with larger AIAN populations. With the exception of the beginning of the pandemic, we also note that $R_t(n)$ dips further below $R = 1$ and more often for the White population. Notably, $R_t(n)$ remains well below $R = 1$ for the "Other" population, while Other represents 8% of the total population. The reason for this may be that many deaths actually in this population are yet unreported or were reported as members of one of the other populations the state of Alaska considers.

4.2.5. *The state of California (CA).* The populations we consider for California are shown in Table 3. The LatinX and White populations make up 39% and 37% of the population of California respectively. As can be seen in the analysis for hospitalized and dead plots, both populations in Figures 19 (Appendix B) and 37 evolve very similarly early on but begin to diverge from each other after the first set of restrictions are lifted at the end of May 2020. The LatinX population begins to overtake the White population in deaths and infections for all three approaches with a continually widening gap. The LatinX population is the majority in the state of California, however, only by about 3%. The disparity between infections and deaths between the LatinX and White populations after the lifting of restrictions may be related to which type of employment is more common in each population. According to the U.S. Equal Employment Opportunity Commission, White/Asian individuals comprise about 87% and 90% of the high tech jobs in the San Francisco-Oakland-Fremont and Santa Clara County regions, respectively, while Black/Hispanic individuals comprise only about 10% and 8% of the high tech jobs in each of those regions [3]. In March 2020, out of the roughly 19.17 million individuals that make up the labor force in California [4], about 1.87 million of those jobs were in the technology field [1]. That

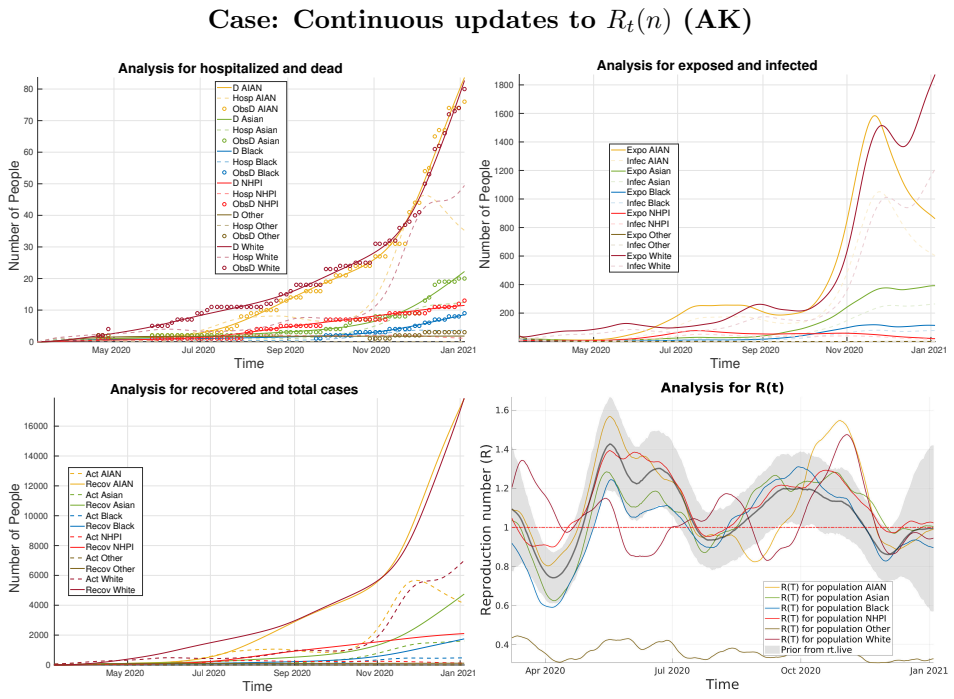

FIGURE 8. Analysis results for the continuous update case for the state of AK.

is, almost 1 in every 10 working-aged individuals in California would have a job considered in the technology sector of the economy. Since such jobs are more easily amicable to being conducted remotely, and because of the large racial disparity we see for the demographics of such jobs, this could be one influencing factor in the disproportionate spread of COVID-19 in the LatinX population.

In the case of California, the LatinX population met the criteria to be classified as a DAP as described in Section 1.1. When we examine the analysis $R_t(n)$ curves for the case of a contact matrix of all ones (Figure 19, top, Appendix B), we do see some stratification occurring among populations with the LatinX population a bit above the others. However, when employing the age-based contact matrices (Figure 19, bottom, Appendix B) this stratification disappears between the two majority populations. In the case of a time-continuous update to the $R_t(n)$ curve (Figure 37), the curve for the LatinX population is, on average, above or similar to the White population.

4.2.6. *The state of Connecticut (CT).* The populations we consider for Connecticut are shown in Table 3. In the state of Connecticut, the LatinX and Multi-Racial populations meet the criteria for the disparity in the disproportionate number of confirmed cases as outlined in Section 4.2. The LatinX population makes up about 16% of the state's population, while the Black and White populations make up 10% and 67%, respectively. All other populations considered in this state's analysis make up less than 5% of the population, with the Multi-Racial population making up about 2%. According to the case count data, 26% of confirmed cases come from

Case: Continuous updates to $R_t(n)$ (CA)

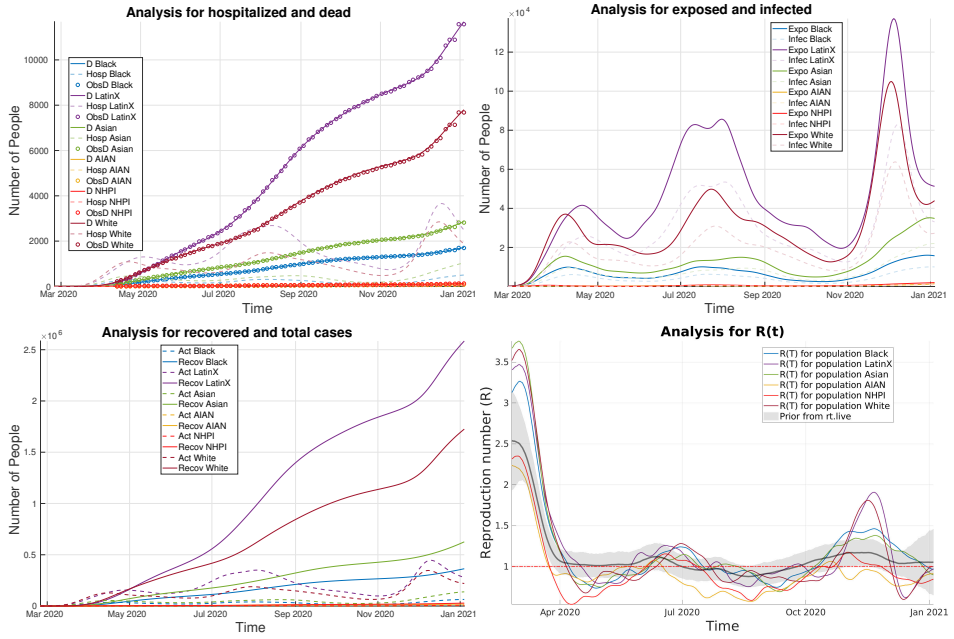

FIGURE 9. Analysis results for the continuous update case for the state of CA.

the LatinX population while only 11% come from the Black population. However the Black population has more deaths, and for all three approaches, they exhibit more overall cases. It is important to again note that confirmed cases may not be representative of all the actual cases, which is the reason why we do not assimilate that data. It may be possible that one population has more access to testing than another or maybe more likely to get tested in general. The fact that the Black population makes up 6% less of the population of Connecticut, yet accounts for 3% more of the deaths (at the time of writing this manuscript) may be an indication of disparities. The low number of confirmed cases compared to the LatinX population suggests either a higher case fatality rate (CFR) or poor access to testing for the Black population. The CFR for each population is estimated in our ESMDA process, however, we do not detect any appreciable difference between these populations in this state for any of the approaches.

When we examine the analysis $R_t(n)$ functions using piece-wise priors (Figure 20, Appendix B), we see that during the first intervention period, $R_t(n)$ is much lower for the White population than all of the others, possibly relating to the types of jobs held by the populations, their access to C.A.R.E.S. act aid, or reserve funds to fall back on. This apparent difference is lessened somewhat in the time-continuous case. In Figure 38 we do see the effect of the first intervention period on $R_t(n)$ between April and May with a sharper dip for all populations with a minimum value amongst the White population. After the reopening, we see an increase in $R_t(n)$ for all populations with the White and Black populations sustaining the highest, and

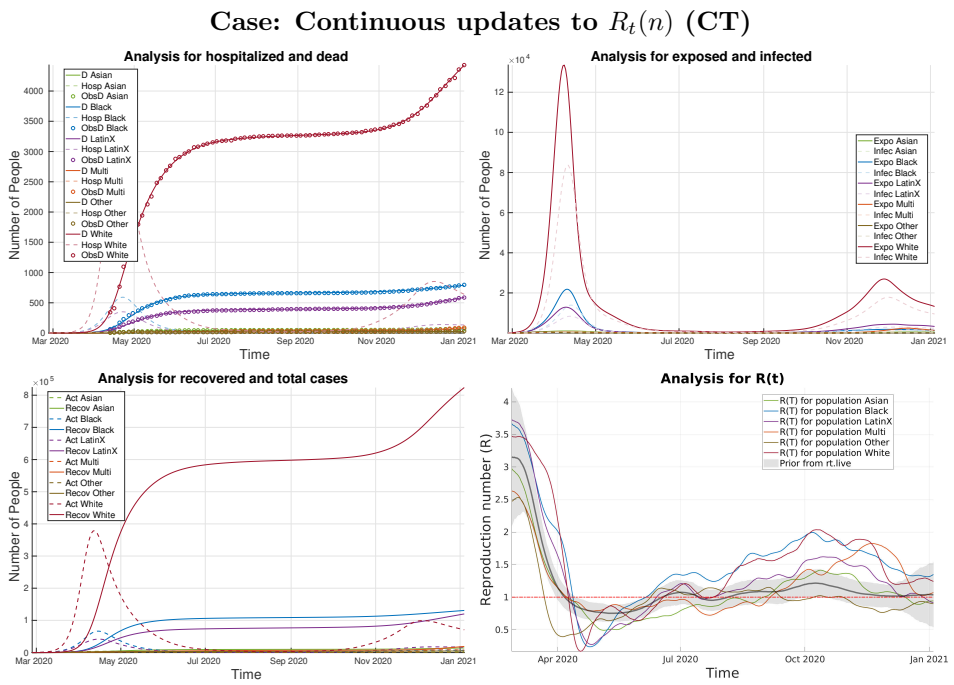

FIGURE 10. Analysis results for the continuous update case for the state of CT.

comparable, values for $R_t(n)$. The LatinX population maintains $R_t(n)$ values less than that of the White and Black populations even though they represent a larger proportion of the total population than that of the Black population. This again suggests the detection of some disparity for the Black population.

4.2.7. *The state of Delaware (DE).* The populations we consider for Delaware are shown in Table 3. The LatinX population fits the criteria described in Section 4.2 for a DAP, making up about 9% of the population but accounting for 18% of confirmed cases. When examining the analysis $R_t(n)$ values for the case with a contact matrix of all ones, in Figure 21 we see a stratification of $R_t(n)$ amongst the populations. During the intervention period, we find that the White population has the lowest values of $R_t(n)$ of the three most crowded populations in the state while the LatinX population has the highest. The very low value for the Asian population is notable given that they are 4% of the total population; however, the state of Delaware includes the NHPI group in this population and some deaths may be reported in the Other or Multi populations. When using the age-based contact matrices and piece-wise updates to $R_t(n)$, with the LatinX population as a DAP, this stratification is lessened during the intervention period (Figure 21, Appendix B). This deeper dip in $R_t(n)$ for the White population during the intervention period is also present in the continuous update case shown in Figure 40. Interestingly, the Black population maintains a much higher value for $R_t(n)$ during this same time period suggesting that they may also be better classified as a DAP in this state even though they are not flagged as being such. This is another example of how reanalysis through DA

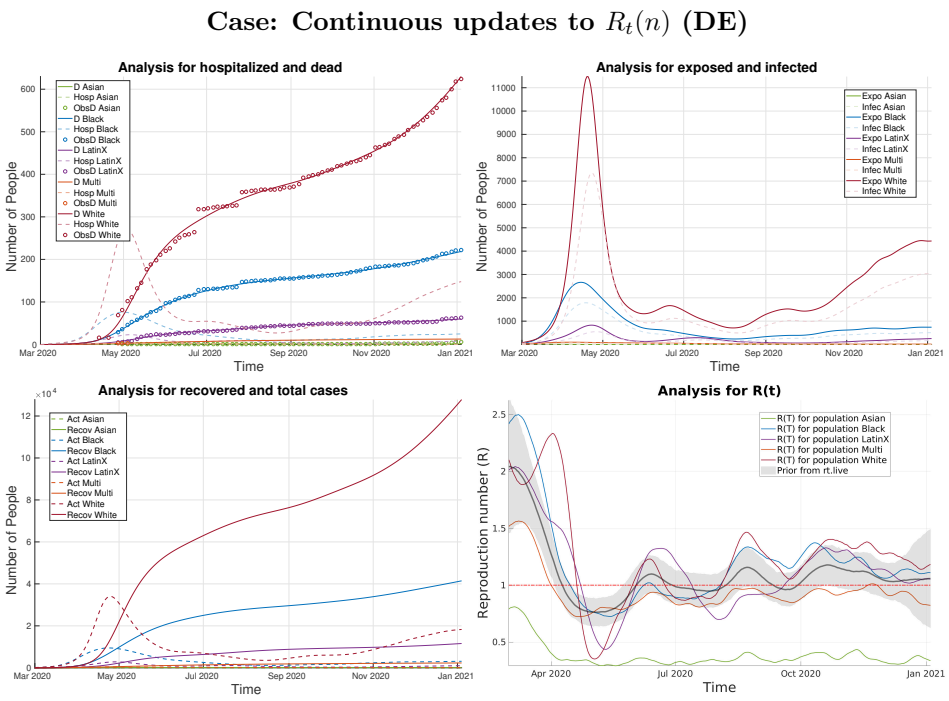

FIGURE 11. Analysis results for the continuous update case for the state of DE.

can be used to understand historical dynamics in further depth than on can through summary statistics.

4.2.8. *The district of Columbia (DC).* The populations we consider for the District of Columbia are shown in Table 3. In the District of Columbia, the Black and LatinX populations meet the criteria of a DAP. The Black population comprises 46% of the population and represents 75% of the deaths, while the LatinX population comprises about 11% of the population and represents 25% of confirmed cases and 13% of deaths. In stark contrast, the White population comprises 41% of the population and makes up only 10% of the deaths. In the cases with piece-wise updates to $R_t(n)$, we see in Figure 22 (Appendix B) that the stratification of the $R_t(n)$ curves with our DAPs are consistently above those for the NDAPs. We note that the Black and LatinX populations have similar analysis $R_t(n)$ curves for these runs, however, the Black population has a much more substantial population total. In terms of the number of active cases and total cases, we see in the relevant analysis plots that the LatinX population overtakes the White population despite being a much smaller proportion of the total population. A possible disparity unique to the LatinX population may be evident in the analysis $R_t(n)$ curves for the continuous-time update runs shown in Figure 39 where we see the Black and White populations' $R_t(n)$ values well below that of the LatinX population's during the intervention period. These disparities may be related to a population's ability to remain in "lockdown" by receiving stimulus money or boosted unemployment coming from the

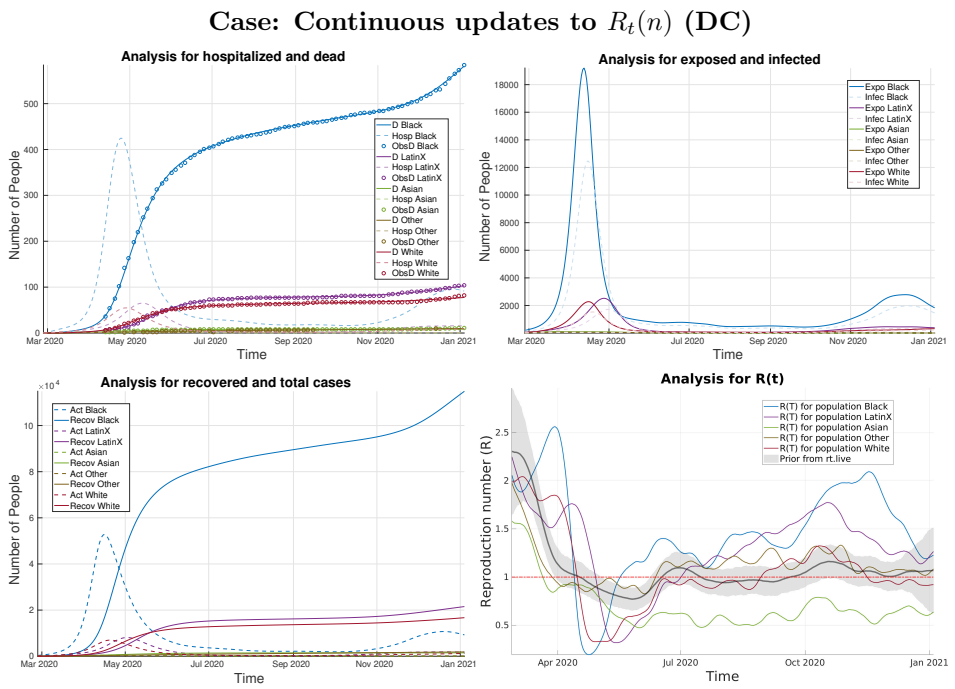

FIGURE 12. Analysis results for the continuous update case for the District of Columbia.

C.A.R.E.S. act. In many cases, undocumented workers were unable to receive this federal assistance [27], hindering one's ability to remain home possibly contributing to increased rates of spread. After the reopening, we see both the LatinX and Black populations' $R_t(n)$ curves go above the exponential spread threshold of $R = 1$, while the White population remains below this threshold until September, where it then moves slightly above that threshold. In the absence of federal financial assistance, populations that have more individuals able to work from home will likely have less spread.

4.2.9. *The state of Hawaii (HI).* The populations we consider for Hawaii are shown in Table 3. The Asian and NHPI populations meet the DAP criteria outlined in Section 4.2. The Asian population makes up 38% of the population and 57% of the deaths, while the NHPI population makes up 10% of the population with 43% of confirmed cases and 31% of the deaths. In the run with contact matrices of all ones, we see in Figure 23 (Appendix B) that there is a stratification in the analysis $R_t(n)$ with the Asian and NHPI populations maintaining higher values of R over all the other populations. The stratification persists when using the age-based contact matrices with the piece-wise updates to $R_t(n)$, as shown in the bottom row of Figure 23. When examining the analysis $R_t(n)$ for the time-continuous update case in Figure 41, the stratification during the initial intervention period is reduced but still present with the NHPI population having the largest values of $R_t(n)$. After the reopening, the stratification persists and increases into the summer months with the Asian population having the largest values of R followed by the NHPI population. After September, $R_t(n)$ is below $R = 1$ for all populations through

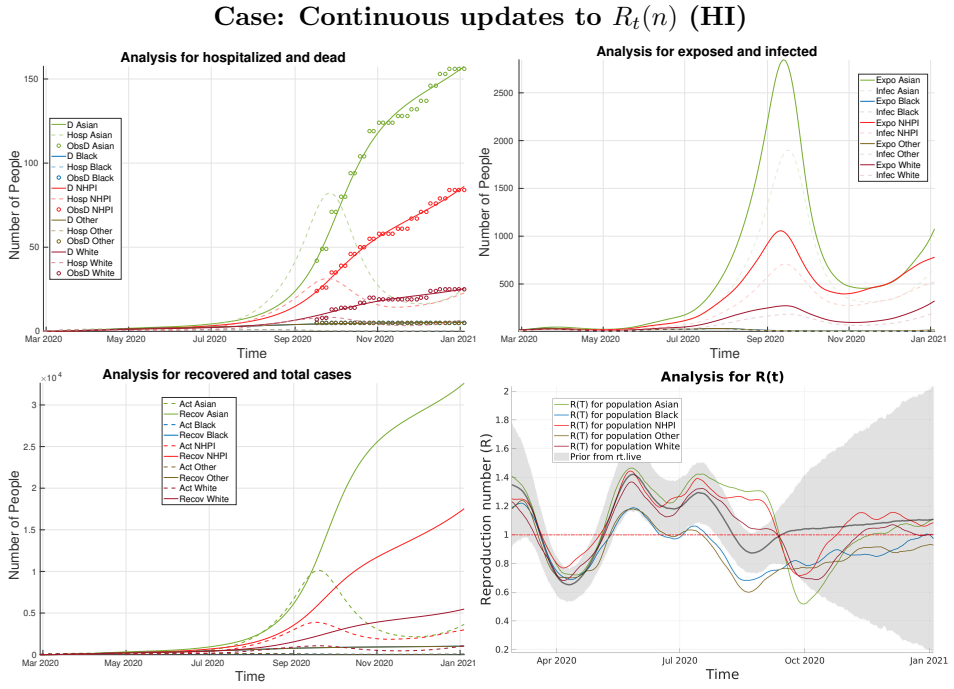

FIGURE 13. Analysis results for the continuous update case for the state of HI.

the end of October before surging again. In this period, the Asian population has the lowest values of $R_t(n)$ possibly due to the period before where they had the largest, causing enough infections to lower the number of susceptible individuals in the population.

4.2.10. *The state of Maryland (MD).* The populations we consider for Maryland are shown in Table 3. For the state of Maryland, the LatinX population meets the threshold for a DAP as outlined in Section 4.2, as they comprise 10% of the population yet account for 20% of the confirmed cases. Indeed, for all three approaches, we see that the LatinX population has the highest values of the analysis $R_t(n)$ during the intervention period, compared to the other populations observable in Figures 24 (Appendix B), and 42. After the intervention period, we do see an almost steady spread around $R = 1$ for this population when piece-wise updates to $R_t(n)$ are used (Figure 24). In the case of continuous-time updates to $R_t(n)$, we see evidence of a fall surge beginning in October for all populations shown in Figure 42. While the Black population did not meet the threshold for a DAP, we see evidence that they could still be one according to our runs and analysis $R_t(n)$. Comprising 29% of the population of the state, they suffer 37% of the deaths compared to that of the White population which makes up 51% of the total population and 49% of the deaths. Indeed, in the analysis for the exposed and infected plots in Figure 42, we see that before April and between May and September, the Black and White populations have roughly the same number of infections and exposures despite the differences in percent make up of the total population. This suggests a disparity that may be tied to the types of work that the majority of people in each population comprise,

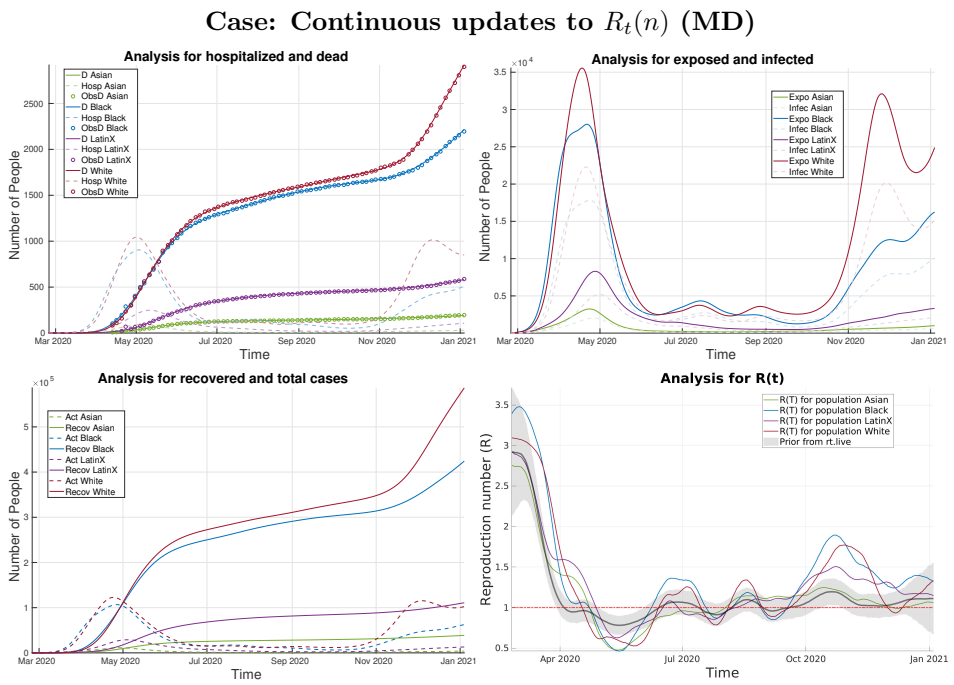

FIGURE 14. Analysis results for the continuous update case for the state of MD.

such as stay-at-home work versus working at a place of high exposure. It may also be that the case fatality rate is higher for one population than the other and there were fewer infections in the Black population than predicted by the assimilation.

4.2.11. *The state of Michigan (MI).* The populations we consider for Michigan are shown in Table 3. In the state of Michigan, the Black population meets the criteria outlined in Section 4.2 to be a DAP. The Black population makes up 14% of the state's population and comprises 24% of the deaths, while the White population makes up 78% of the state's population and comprises 69% of the deaths. In all three approaches, we see that during the intervention period, the White population has the lowest values of the analysis $R_t(n)$ curves, again possibly signaling an ability to stay home due to the types of jobs held or as a result of reserve wealth. This is shown in Figures 25 (Appendix B), and 43. We also see that early on in the pandemic, the Black and White populations have comparable numbers of infections despite the difference in population proportions. This is also shown in the aforementioned figures in the analysis of the exposed and infected as well as the recovered and total cases plots. We can also observe that for all three approaches, in the same figures, that the Black population maintains a higher value of $R_t(n)$ for the vast majority of the data window.

4.2.12. *The state of Utah (UT).* The populations we consider for Utah are shown in Table 3. In the state of Utah, none of the populations reported on meet the criteria to be a DAP as described in 4.2. In Figures 26 (Appendix B), and 44, we see a stratification of the analysis $R_t(n)$ curves amongst the populations, with the White and Asian populations usually having the lowest values compared to the

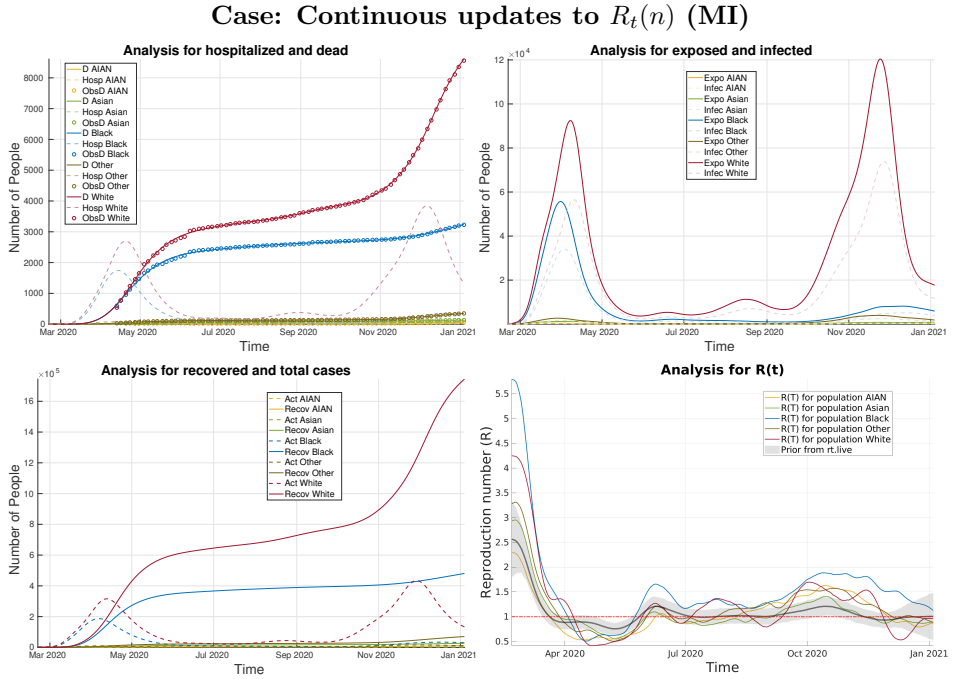

FIGURE 15. Analysis results for the continuous update case for the state of MI.

others. In the analysis plot for $R_t(n)$ in the time-continuous case, we see a striking rise in the reproductive rate beginning at the end of the intervention period for the AIAN population (see Figure 44). Utah has a sizable amount of Native American reservation land, a population that was shown to be disproportionately affected by the pandemic. This spike is less apparent in the cases with piece-wise updates to $R_t(n)$, and we note that the time-continuous update runs are better able to track time-dependent differences in the spread between populations. We also see a large spike in the continuous case at the end of the summer for the Other population. Also interesting here is that the disparity between different populations during the intervention period is somewhat less apparent than in other states. One possible reason for this is the large population of members of the Latter Day Saints (LDS) faith. The LDS church often offers financial and food assistance to its members. This could act as a cushion for LDS members of the LatinX population who were ineligible to receive assistance from the C.A.R.E.S. act, allowing them to remain home longer. While we were unable to find the specific proportions of the LatinX population in Utah that are also Mormon, we do note that according to the LDS church, 68% [10] of the state is LDS. We cannot be sure how much of an effect this would have however.

4.2.13. *The state of Washington (WA).* The populations we consider for the state of Washington are shown in Table 3. In Washington, the LatinX, NHPI, and AIAN populations meet the criteria for a DAP, as described in Section 4.2. The LatinX population comprises 13% of the total population with 34% of the confirmed cases; the NHPI population is less than 1% of the total population yet makes up 2%

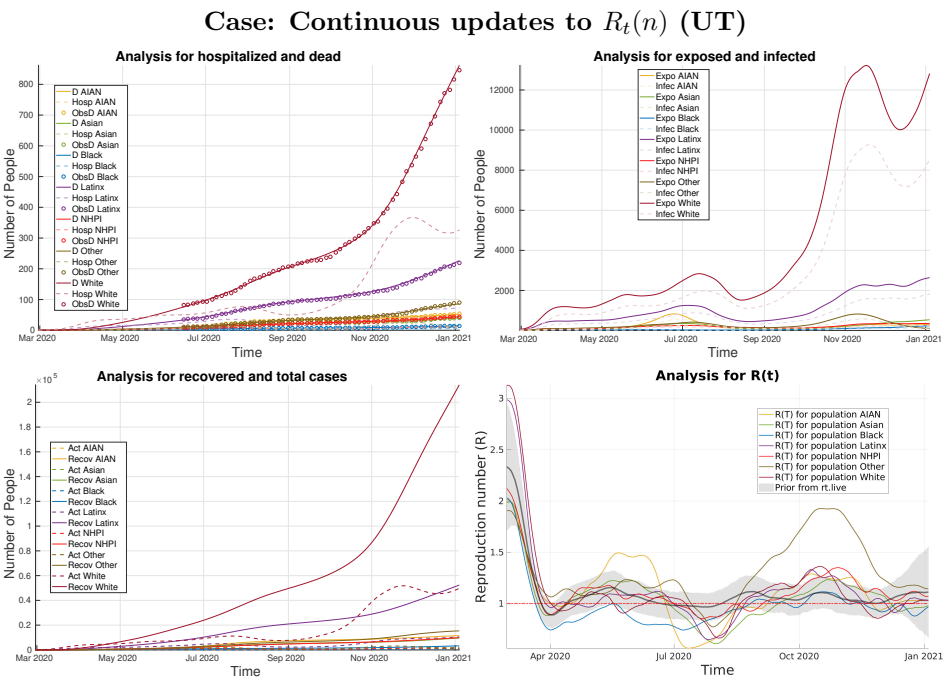

FIGURE 16. Analysis results for the continuous update case for the state of UT.

of the deaths; and the AIAN population comprises 1% of the population while accounting for 2% of the deaths. The disparity for the AIAN and NHPI populations is evidenced in all three approaches with their analysis $R_t(n)$ values predominately above $R = 1$ in all three cases shown (Figures 27, Appendix B, and 45). We also observe in the analysis for $R_t(n)$ plots in the continuous update case (Figure 45) that the LatinX population has higher values for the reproductive rate than all other populations–except for the NHPI and AIAN populations–during the intervention period.

5. **A remark on priors.** We would like to make a remark on the importance of priors used in DA and how they can be used to explore a dynamical system and test assumptions. For simplicity, we focus on a specific example from this work stemming from the state of Connecticut. In section 4.2.6 we note the possible detection of a disparity otherwise not apparent in the aggregate statistics available at the time. There, the Black population did not meet the criteria for disparity as defined in section 1.1 (in terms of the number of confirmed cases) while the LatinX population did. Yet the Black population had a higher number of deaths despite making up a smaller part of the population than the LatinX population. This presents two distinct scenarios: either the Black population has a higher case fatality rate than the other populations or the number of actual cases was under-sampled. The latter could be due to a lack of access to testing or an avoidance of testing due to fear of cost or ability to take time off of work. In all of the previous runs we take our prior on the CFR to be 0.009, just under one percent. While we refer to this as the CFR, because of the way our model is constructed, this is really the *infection fatality rate*

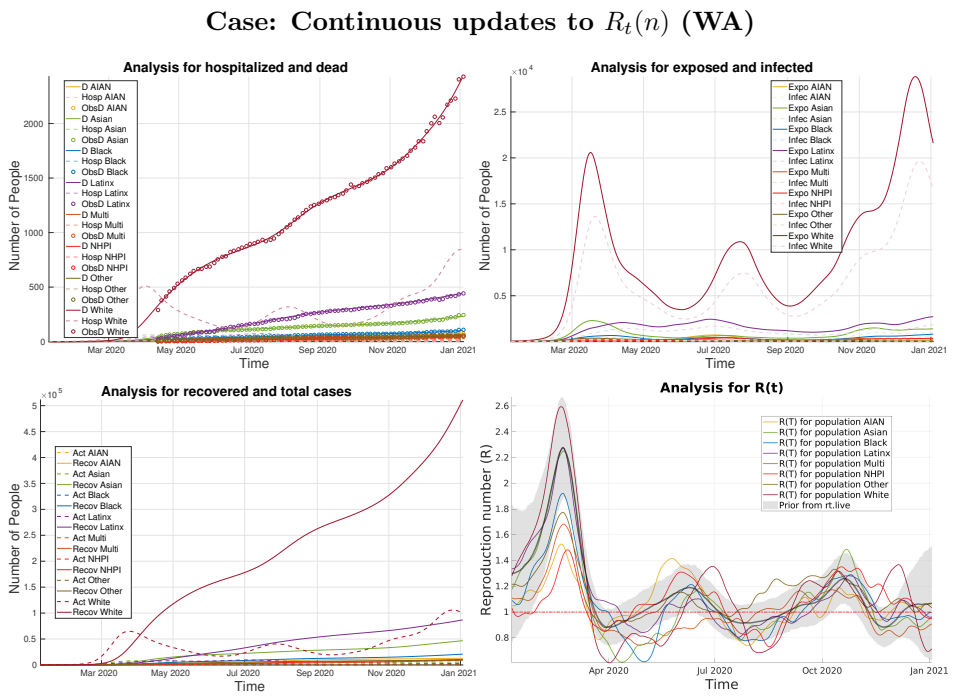

FIGURE 17. Analysis results for the continuous update case for the state of WA.

(IFR). The IFR is the percent of deaths that occur out of all infections, detected or not. The value of 0.009 is the average over all age populations and is inline with W.H.O. estimates [23] and seroprevalence studies [20]. While the CFR for each population could be estimated directly from the data, this is currently difficult. If there is disparity in access to testing then the CFR estimated directly from the data will clearly be skewed. A study in New York observed that the number of total tests significantly increased with the increasing proportion of White residents [26]. In one study [28], wildly varying CFRs between racial/ethnic populations were observed ranging from 7.35% for the non-Hispanic White population to 1.39% for the Hispanic population in Ohio. Further, data on the race of persons tested is much more difficult to curate than that of deaths. This has to do primarily with the volume of testing but also to do with the fact that death reporting is far more standardized by state laws. In situations where some data is more reliable than others, DA is a powerful tool for investigation. With the use of a well designed model, inconsistencies in data can be tempered to produce reasonable and feasible posterior estimates consistent with the reality reflected in the data.

To investigate the two scenarios discussed above, we complete several different runs for Connecticut using various priors and compare the results. In the first set we take the prior for $R_t(n)$ to be that given by rt.live for all populations and assign a standard deviation of 1.5. Using this we then complete assimilation runs with CFR priors of 0.02, 0.009, and 0.001 with large uncertainties. The priors, posteriors and observation χ^2 statistic for this first experiment are shown in Table 5 and Figures 28,

TABLE 4. CFR estimates coming from the run presented in Section 4.2.6

Population	Prior,$\sigma = 0.009$	Post,$\chi^2 = 31$
Asian	0.0900	0.0089
Black	0.0900	0.0100
LatinX	0.0900	0.0090
Multi	0.0900	0.0093
Other	0.0900	0.0094
White	0.0900	0.0094

TABLE 5. Results for the CFR using the same $R_t(n)$ prior for all populations with $\sigma_{R(t)} = 1.5$.

Population	Prior,$\sigma = 0.05$	Post,$\chi^2 = 112$	Prior,$\sigma = 0.05$	Post,$\chi^2 = 92$	Prior,$\sigma = 0.03$	Post,$\chi^2 = 87$
Asian	0.009	0.0074	0.020	0.0093	0.001	0.0070
Black	0.009	0.0126	0.020	0.0157	0.001	0.0084
LatinX	0.009	0.0213	0.020	0.0246	0.001	0.0147
Multi	0.009	0.0103	0.020	0.0104	0.001	0.0086
Other	0.009	0.0065	0.020	0.0042	0.001	0.0065
White	0.009	0.0229	0.020	0.0271	0.001	0.0134

TABLE 6. Results for the CFR using $R_t(n)$ priors taken from piece-wise assimilation first assuming a CFR of 2% for all populations.

Population	Prior,$\sigma = 0.05$	Post,$\chi^2 = 89$	Prior,$\sigma = 0.002$	Post,$\chi^2 = 41$
Asian	0.020	0.0282	0.020	0.0199
Black	0.020	0.0337	0.020	0.0199
LatinX	0.020	0.0317	0.020	0.0200
Multi	0.020	0.0316	0.020	0.0202
Other	0.020	0.0299	0.020	0.0204
White	0.020	0.0259	0.020	0.0195

TABLE 7. Results for the CFR assuming a CFR of 2% for only the Black Group.

Population	Prior,$\sigma = 0.05$	Post,$\chi^2 = 302$	Prior,$\sigma = .1x$	Post,$\chi^2 = 38$	Prior, $\sigma = 0.05$	$\chi^2 = 83$
Asian	0.0090	0.0101	0.0090	0.0091	0.0090	0.0249
Black	0.0200	0.0226	0.0200	0.0201	0.0200	0.0371
LatinX	0.0090	0.0211	0.0090	0.0092	0.0090	0.0299
Multi	0.0090	0.0097	0.0090	0.0092	0.0090	0.0285
Other	0.0090	0.0053	0.0090	0.0093	0.0090	0.0234
White	0.0090	0.0248	0.0090	0.0090	0.0090	0.0256

29 and 30. Here all of the figures relating to this section can be found in appendix C. In the second set of runs, we repeat the original process outlined in section 4.2 substituting in a prior of 0.02 over 0.009 for the CFR with an uncertainty of 0.002 to obtain a prior for the continuous update. The $R_t(n)$ resulting from this is then used with a prior CFR of 0.02 for all populations with uncertainties of 0.05 and 0.002 with the results displayed in Table 6 and Figures 31 and 32. Finally, we assume a CFR of 0.009 for all populations except the Black population to which we assume a

prior of 0.02 (results in Table 7 and Figures 33, 34 and 35). It is in this last case alone that we find a solution for which case counts for the Black population can be lower while deaths can be higher. This single solution requires low uncertainty in the CFR with the CFR of the Black population being at least twice that of the others (columns 4 and 5 of Table 7 with the $R_t(n)$ curves shown in Figure 34). In all other scenarios a disproportionate under-counting of cases is supported by the solution. In this last experiment, the second and third columns in table 7 come from a run where all populations have the same $R_t(n)$ prior and large uncertainty while the last four instead use different priors for each population coming from an initial piece-wise assimilation assuming a CFR of 0.02.

We now make some general observations on the results of the experiments described above. In the cases where the uncertainty in $R_t(n)$ and the CFR were large, we obtain our poorest fits with the posterior CFR tending to be larger for the three populations making up the largest proportion of the total population, White, LatinX and Black. When all populations are given the same prior for $R_t(n)$, the posterior CFR for the Black population is lower than that of the White and LatinX populations in all cases (see Table 5). The resulting posteriors for an individual populations are similar whether a prior of 0.009, 0.02 or 0.001 are chosen. For the three largest populations, the posterior CFR obtained is larger than that of the W.H.O. estimate of 0.009. This however is a result of the fact that the $R_t(n)$ prior assigned is averaged over all of the populations and for some will be too low. In this case, this may drive the CFR up to account for the number of deaths observed with less infections. In general, larger priors for the CFR drive $R_t(n)$ estimates down and for lower CFR priors the opposite is true. This illustrates why a good estimate of the CFR is important when reconstructing the time varying $R_t(n)$ curves.

In our primary analysis of each state, first we take a prior for $R_t(n)$ from rt.live and assign this to each population with large uncertainty assuming a CFR of 0.009 with a low uncertainty. We then perform a piece-wise assimilation of the data to obtain new priors for each population to be used in the time continuous case. This is done to roughly separate the $R_t(n)$ curves based on which populations would have had more spread given the assumption of our CFR. The new priors used in the continuous assimilation are assigned lower uncertainty than before and the same CFR prior is then used. In these cases, we always obtained our best fits to the data. In a parallel experiment, we repeat this process substituting a prior CFR of 0.02 but add a continuous assimilation with larger uncertainty in the CFR. In the case where only the CFR is changed, we see a reasonably good fit to the data and note that the posterior $R_t(n)$ generally takes on lower values due to the increase in CFR. When assigning a larger uncertainty to the CFR of 0.05, we see the posterior CFRs increase to 3% for all populations. This has to do with a sampling bias that is introduced through a requirement that the CFR be greater than zero. With a standard deviation of 5% and an assumed mean of 2% sampling which throws away negative or zero values otherwise obtained from a Gaussian distribution, we will end up with a bias toward larger values not discarded. The results of these experiments are shown in Table 6 and Figures 31 and 32.

Finally we looked at a scenario where the death rate was only higher for the Black population, 0.02 compared to 0.009. In the first instance, we assign the same $R_t(n)$ prior to all populations with large uncertainty for this and the CFR priors. A poor fit is obtained with posterior CFRs near 0.025 for the most crowded populations, the increase again likely due to sampling bias as discussed above. In the next two

instances, piece-wise assimilations are again performed first to provide new priors and then continuous assimilation done with low and high uncertainty in the CFR. The best fit is obtained using a lower uncertainty for the CFR and is the only case where a solution which allows the number of infections in the Black population to be less than that of the LatinX population but with more deaths. Incidentally, the posterior $R_t(n)$ curve for the Black population takes on lower values than that obtained under the assumption of a CFR of 0.009.

While our best results for Connecticut, in terms of data fit, comes from the assumption of a CFR of 0.009 for all populations (Table 4 and Figure 38), we do obtain a comparable quality of fit in one case when assuming a CFR of twice that of all other populations using low uncertainty (columns 4 and 5 of Table 7 and Figure 34) which is also the only case that does not suggest an under-counting of infections for the Black population. While there are estimates of higher CFR rates for minorities, they are difficult to pin down and vary wildly by location. Part of the reason for this is that the CFR estimated directly from data depends on having a good estimate of the actual number of cases. If a population is under-served in testing, a CFR may appear higher as less over all cases were captured compared to deaths. In addition, the CDC currently estimates the risk of death for the LatinX population to be 2.3 times that of White's while the Black population is 1.9 times more at risk [8]. This would not support the prior which assumes a CFR for the Black population twice that of the LatinX population and instead suggest that the Black population was under-served in terms of testing.

The ability to test a wide array of assumptions and view the implications of those assumptions through the lens of both model and data is a major feature of the methods of Data Assimilation. Under the prior of a CFR around 1% consistent with W.H.O. estimates, the detection of a disparity not apparent in the aggregate statistics for the Black population was possible. Through the marriage of model and data, information otherwise not determinable can be brought to light. For these reasons the authors believe DA can be used to complete a rigorous reanalysis of the course of the pandemic to provide other scientists invaluable data to improve responses and design mitigation measures for future events. As more and more data become available, tools like the ones outlined in this manuscript can be applied with more accuracy.

6. **Conclusion and discussion.**

We have developed a multi-population SEIR model with age classes and employed an ESMDA scheme to perform an initial re-analysis of the spread of SARS-CoV2 among different racial/ethnic populations. We find disparities in the rate of spread for different populations and estimate that rate as a function of time. We believe the primary driver of these disparities is the ability of these populations to self-quarantine and avoid exposure. Factors that impact one's ability to self-quarantine include an individual's type of employment (work from home or not), access to reserve funds such as generational wealth, and access to healthcare or the general infrastructure in their specific locality. Groups for which many members cannot work from home or do not have reserve funds naturally will be exposed more often to the disease, eventually bringing the virus home to their family and close friends where spread happens robustly and rapidly [21].

In our analyses, we find that typically the Black, LatinX, American Indian and Alaskan Native (AIAN), and Native Hawaiian and Pacific Islander (NHPI) populations exhibit $R_t(n)$ curves suggestive that they fit the criteria to be considered

a DAP. This is typically most notable during and following the intervention period (lockdown). For the LatinX population, many were ineligible to receive financial assistance from the C.A.R.E.S. act due to immigration status making isolation at home during these periods difficult. Stimulus funds may also be less deliverable to DAPs, such as the Black population, which is often underbanked complicating fund distribution. After these intervention periods, DAPs risked increased exposure by returning to work in jobs that cannot be done from home. We consistently find that the White population, while a larger proportion of the overall population, had lower $R_t(n)$ values until the beginning of the Fall surge, around October. This suggests that this population was better able to self isolate for a far longer period of time than the other populations we consider. This is consistent with what is known about wealth disparities between the White population and minority populations in the United States [2].

In some cases, we find evidence that a population is a DAP even though they do not fit the specific criteria outlined in Section 1.1. This is the case in the state of Connecticut where we see evidence that the LatinX population meets the criteria to be a DAP in terms of a disproportionate number of confirmed cases while the Black population does not. However, the analysis suggests more infections among the Black population than the LatinX population, despite being a smaller percentage of the total population. This is based on the assimilation of the number of deaths and suggests that the Black population in the state may have poor access to testing or are less likely to seek a test. With the time-continuous analysis, we can also detect when spread began for a specific population. We note that in the state of Alaska, spread begins in the White population and later into the AIAN population. This suggests major cities like Anchorage were hit first then spread to the rural regions later. The disparity in the number of deaths between the majority White population and the AIAN population also suggests how important healthcare infrastructure is in reducing the impact of the virus. Many of the far north rural populations in Alaska have little healthcare infrastructure and deliveries of supplies can be extremely difficult. We argue that lessons can be learned from these types of analyses to improve planning for a future pandemic.

Our main goals in this work were to study disparity in the spread of SARS-Cov2 and to demonstrate the utility of models and data assimilation to aid in understanding how the pandemic evolved among different racial/ethnic populations in various regions of the United States. This kind of analysis can provide important insight into successes and failures of policy as well as highlight causes for disparities. Using techniques such as ESMDA can allow us to "fill in the data gaps" presented when only considering things like confirmed cases or looking at general statistics. Armed with more complete information and better policy, more effective planning can be implemented to help avoid such disparities and reduce the loss of life in a future pandemic. We note that at this stage the data available is a bit sparse and somewhat incomplete. While we make some conjectures as to the causes of disparities we find in our analyses, we stress that direct causation is not assured. However, as the data is better curated in the coming months and years, analyses such as these can be a powerful tool for social scientists, epidemiologists, and other experts to understand how and why events unfolded as they did and to find better ways to prepare in the future.

Acknowledgments. The authors would like to thank the AIM & MCRN summer program (2020) for their support in supervising, guiding, and helping fund this project. In particular, our mentors Nancy Rodriguez at CU Boulder and Chris Jones at UNC-Chapel Hill for their guidance during the summer program. We would like to also acknowledge the front line workers who keep society functioning at the risk of their own lives. C. Sampson was supported by the US Office of Naval Research under grant N00014-18-1-2204. Daniel P. Maes was supported by the NSF GRFP DGE 1256260 grant. G. Evensen was supported by internal funding from NORCE. Finally, we are thankful for the substantial and constructive reviews that helped us significantly improve our paper.

Appendix A. **Model parameter priors.** In Table 8 we show the first guess parameters used for all populations and in Table 9 we show the p-factors and CFR initially chosen and used for all populations.

Parameter	First guess	Description
τ_{inc}	5.5	Incubation period
τ_{inf}	3.8	Infection time
τ_{recm}	14.0	Recovery time mild cases
τ_{recs}	5.0	Recovery time severe cases
τ_{hosp}	6.0	Time until hospitalization
τ_{death}	16.0	Time until death
p_{f}	0.009	Case fatality rate
p_{s}	0.039	Hospitalization rate (severe cases)
p_{h}	0.4	Fraction of fatally ill going to hospital

TABLE 8. The table gives a set of first-guess model parameters. As we could not find scientific estimates of these parameters, we set their values based on available information from the internet and initial model-tuning experiments. We leave it to the data assimilation system to fine-tune the parameter values.

Age group	1	2	3	4	5	6	7	8	9	10	11
Age range	0–5	6–12	13–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–105
Population	351159	451246	446344	711752	730547	723663	703830	582495	435834	185480	45230
p–mild	1.0000	1.0000	0.9998	0.9913	0.9759	0.9686	0.9369	0.9008	0.8465	0.8183	0.8183
p–severe	0.0000	0.0000	0.0002	0.0078	0.0232	0.0295	0.0570	0.0823	0.1160	0.1160	0.1160
p–fatal	0.0000	0.0000	0.0000	0.0009	0.0009	0.0019	0.0061	0.0169	0.0375	0.0656	0.0656

TABLE 9. The p-numbers indicate the fraction of sick people in an age group ending up with mild symptoms, severe symptoms (hospitalized), and fatal infection.

Appendix B. **Piece-wise assimilation.** Here we present the analysis for hospitalized and dead as well as $R_t(n)$ when applying piece-wise updates to $R_t(n)$ in the cases that $R^A_{ij}(n) = 1$ and when using our age stratified matrices for the DAPs and NDAPs discussed in Section 4.1.

Case: Piece-wise updates to $R_t(n)$ (AK)

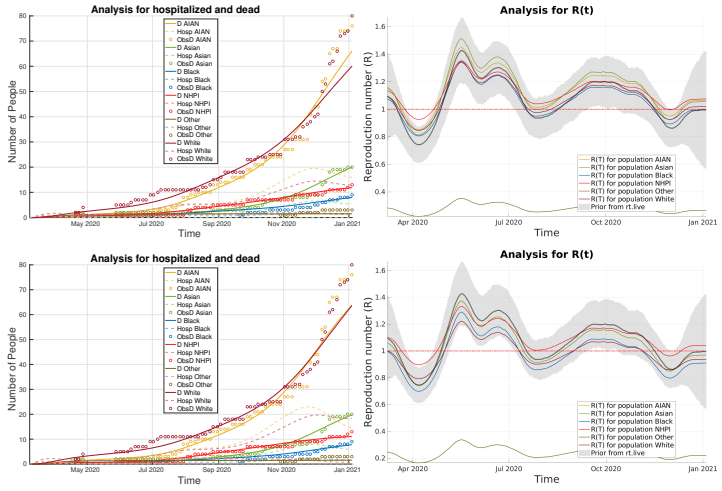

FIGURE 18. Analysis results when using piece-wise updates to $R(t)$ for AK. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

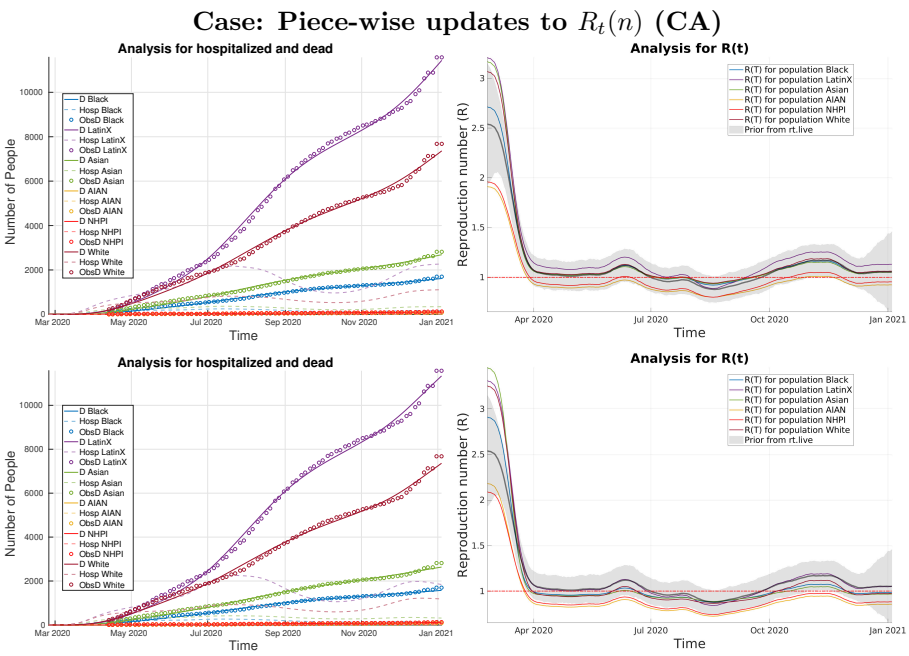

FIGURE 19. Analysis results when using piecewise updates to $R(t)$ for CA. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

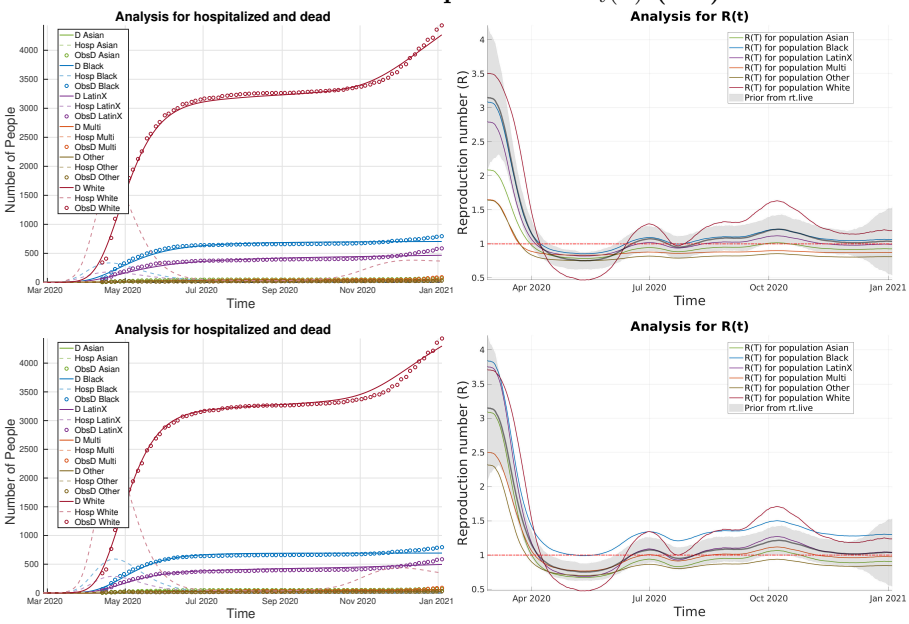

FIGURE 20. Analysis results when using piecewise updates to $R(t)$ for CT. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

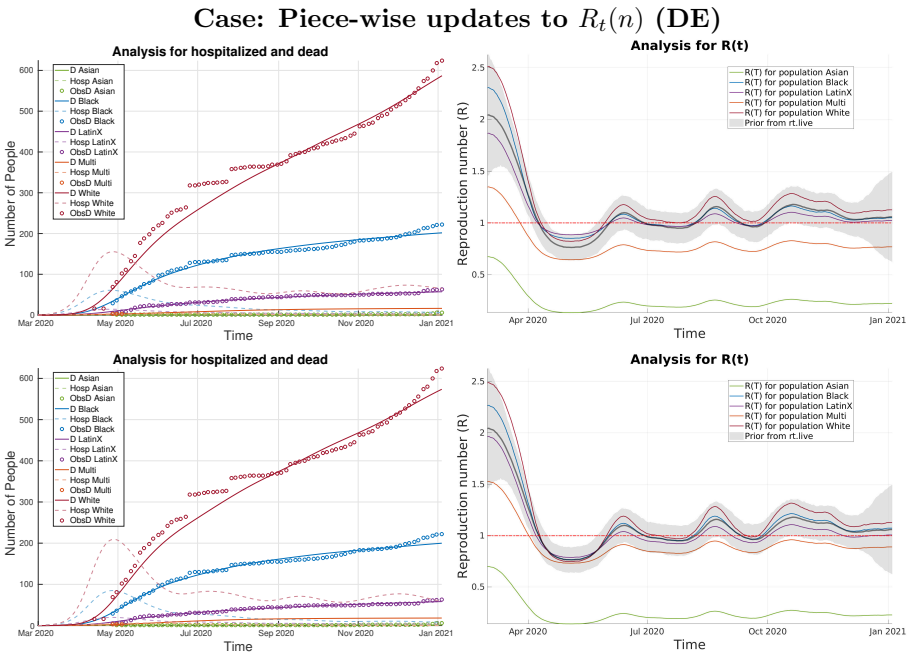

FIGURE 21. Analysis results when using piecewise updates to $R(t)$ for DE. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

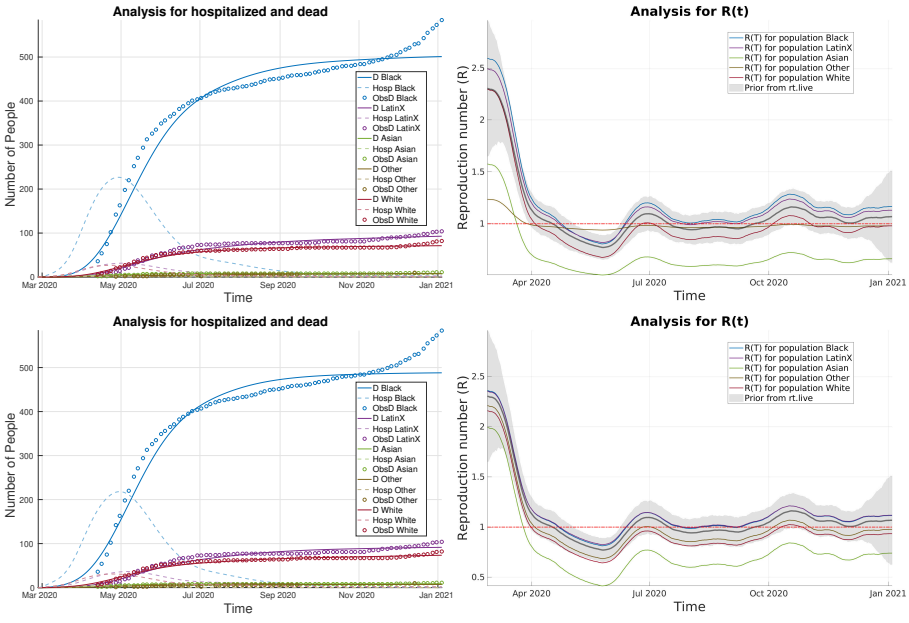

FIGURE 22. Analysis results when using piecewise updates to $R(t)$ for DC. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

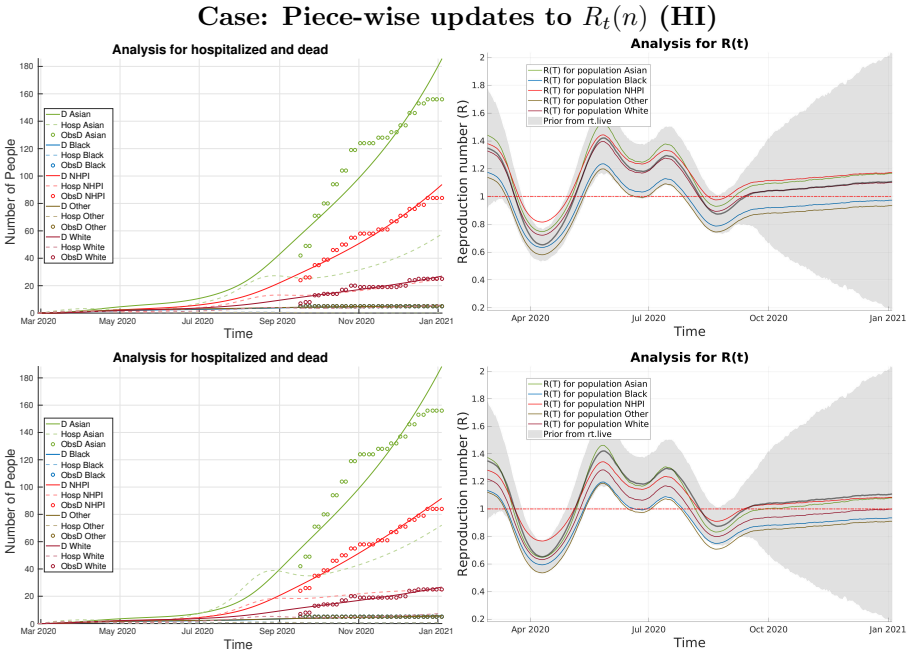

FIGURE 23. Analysis results when using piecewise updates to $R(t)$ for HI. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

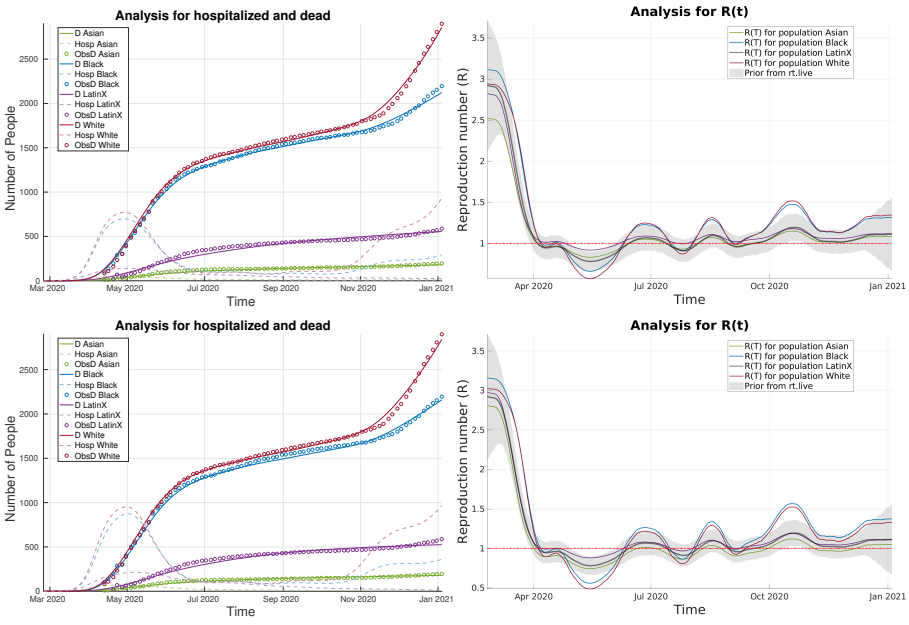

FIGURE 24. Analysis results when using piecewise updates to $R(t)$ for MD. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

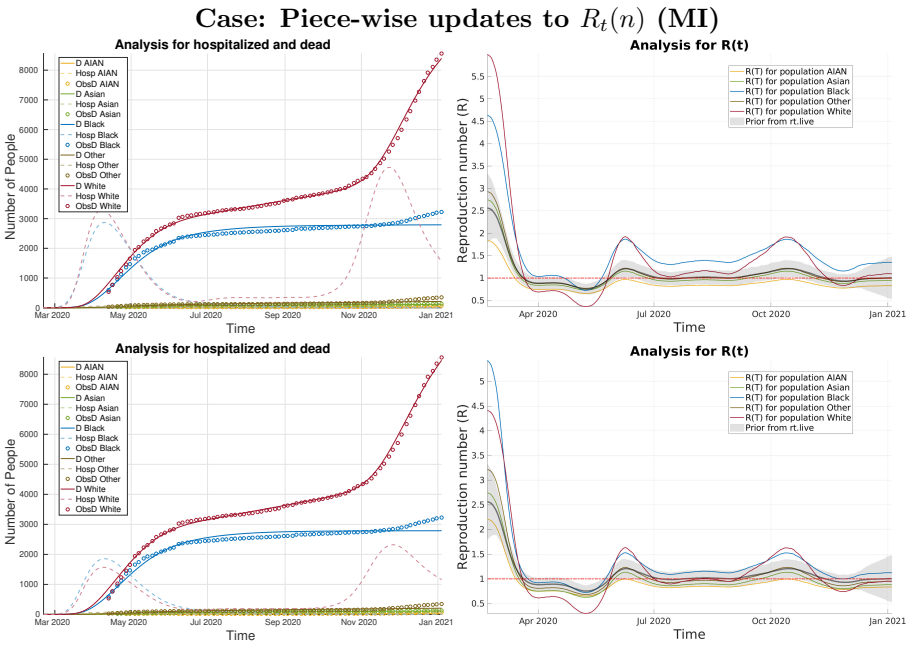

FIGURE 25. Analysis results when using piecewise updates to $R(t)$ for MI. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

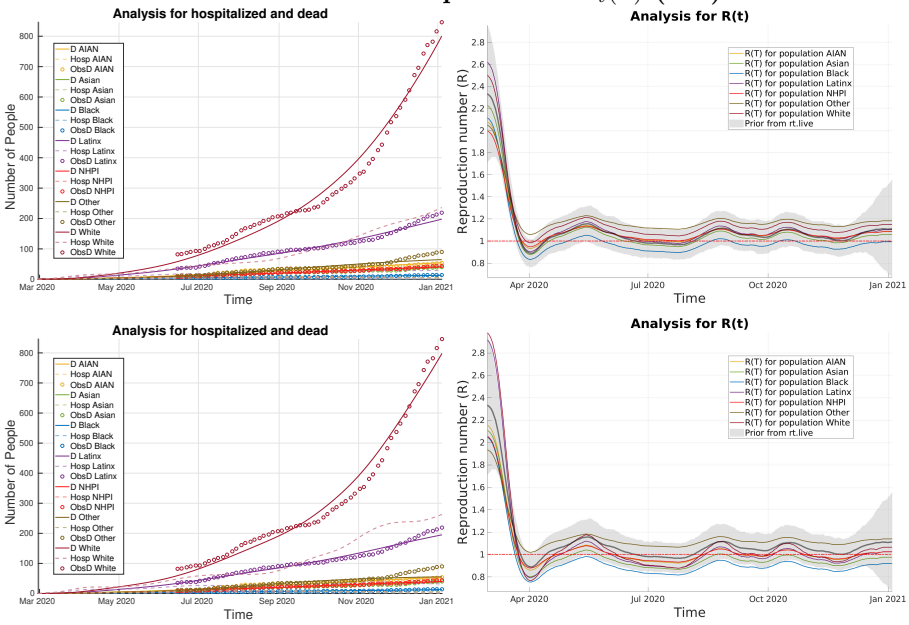

FIGURE 26. Analysis results when using piecewise updates to $R(t)$ for UT. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

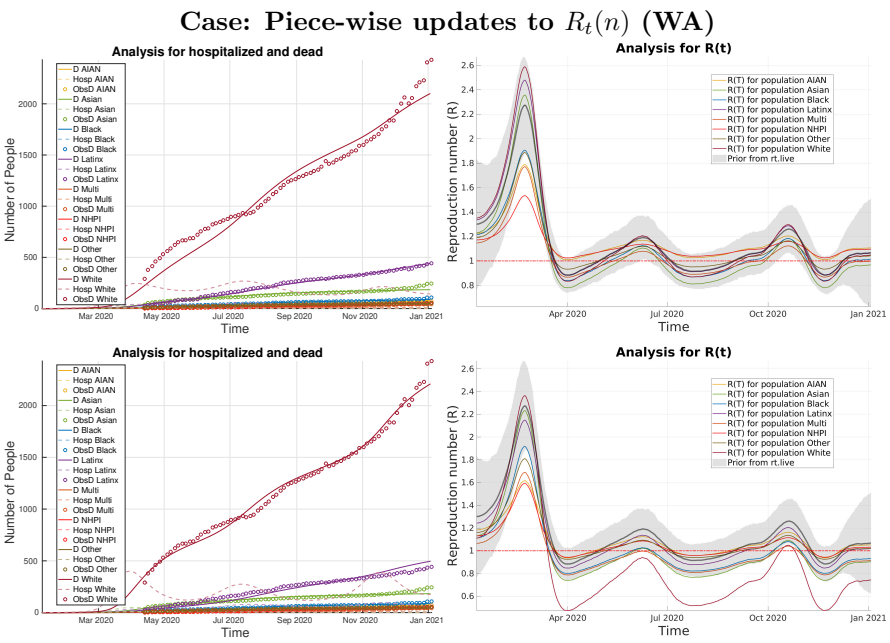

FIGURE 27. Analysis results when using piecewise updates to $R(t)$ for WA. Top row: R^A with entries of all ones. Bottom Row: R^A for DAPs and NDAPs.

Appendix C. Prior experiments for Connecticut.

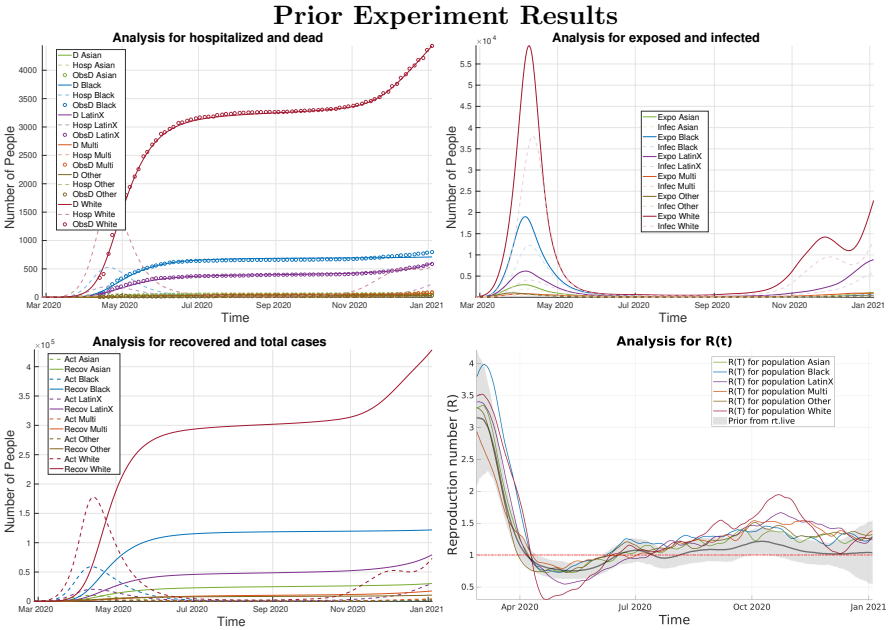

FIGURE 28. Analysis results with a CFR prior of 0.009 ($\sigma_{CFR} = 0.05$) and the same $R_t(n)$ ($\sigma_{R(t)} = 1.5$) prior for all populations. This figure corresponds to columns 2 and 3 in Table 5

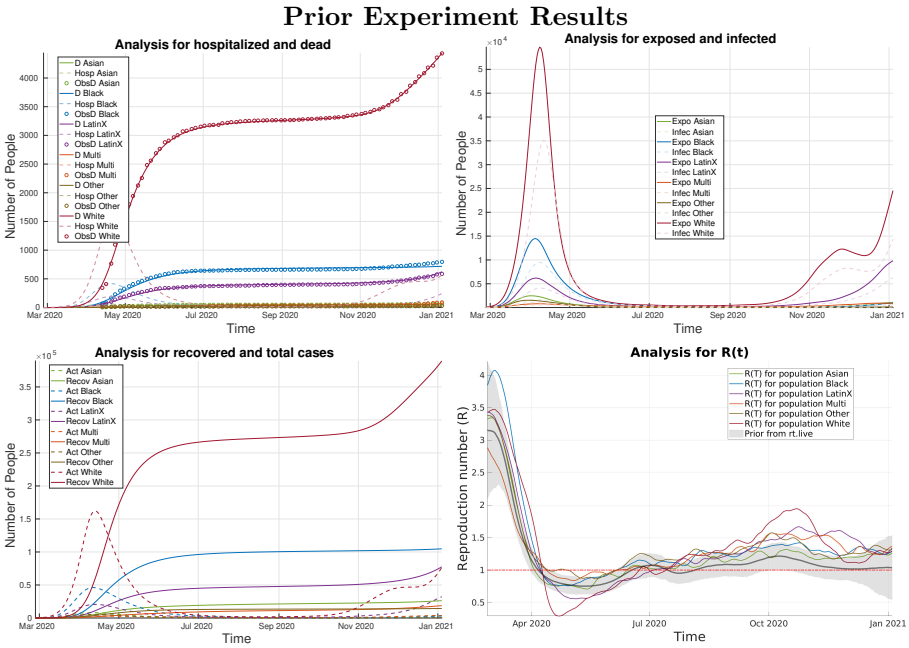

FIGURE 29. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.05$) and the same $R_t(n)$ ($\sigma_{R(t)} = 1.5$) prior for all populations. This figure corresponds to columns 4 and 5 in Table 5.

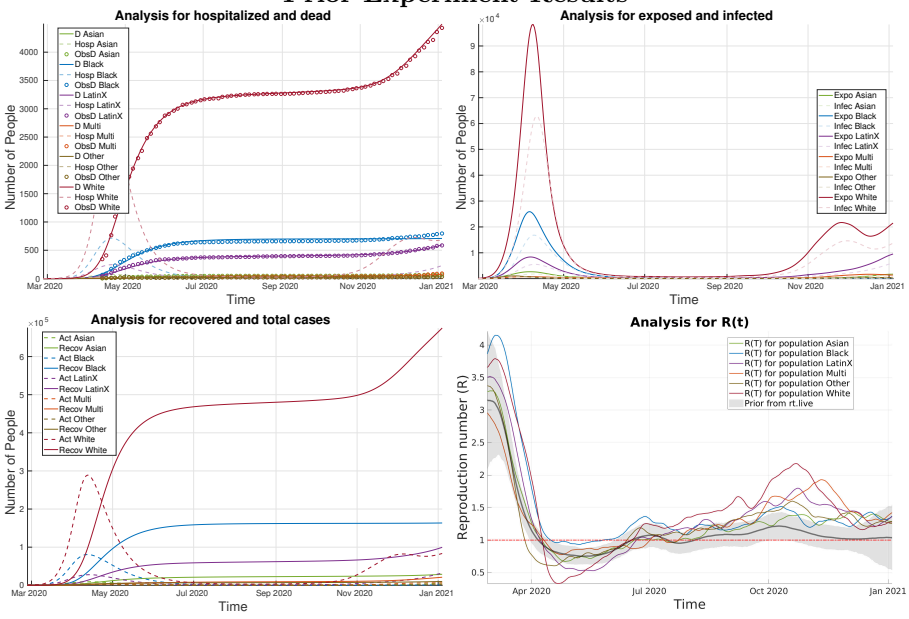

FIGURE 30. Analysis results with a CFR prior of 0.001 ($\sigma_{CFR} = 0.03$) and the same $R_t(n)$ ($\sigma_{R(t)} = 1.5$) prior for all populations. This figure corresponds to columns 6 and 7 in Table 5.

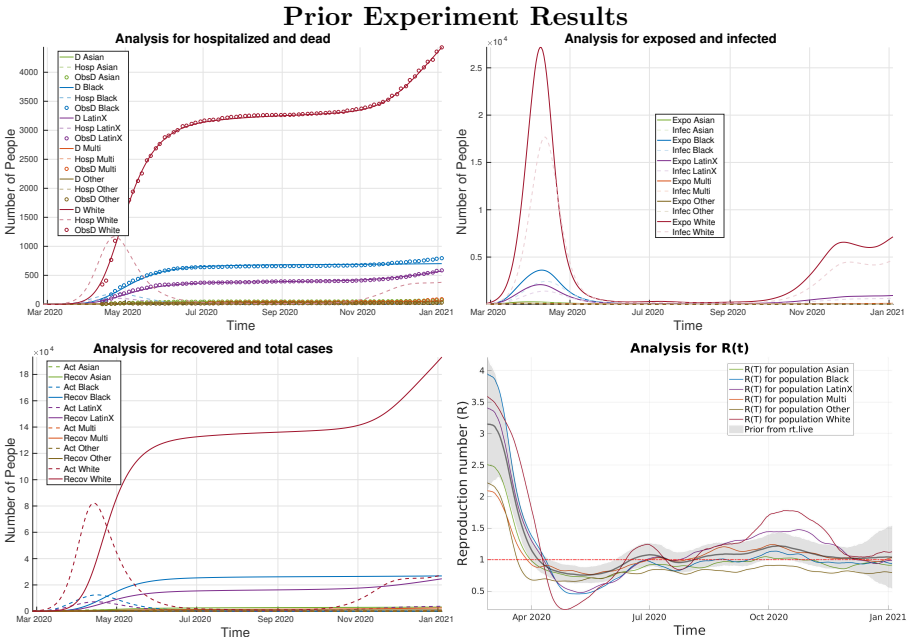

FIGURE 31. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.05$) and prior $R_t(n)$ ($\sigma_{R(t)} = 0.5$) curves coming from initial piece-wise assimilation also assuming a 0.020 CFR prior. This figure corresponds to columns 2 and 3 in Table 6.

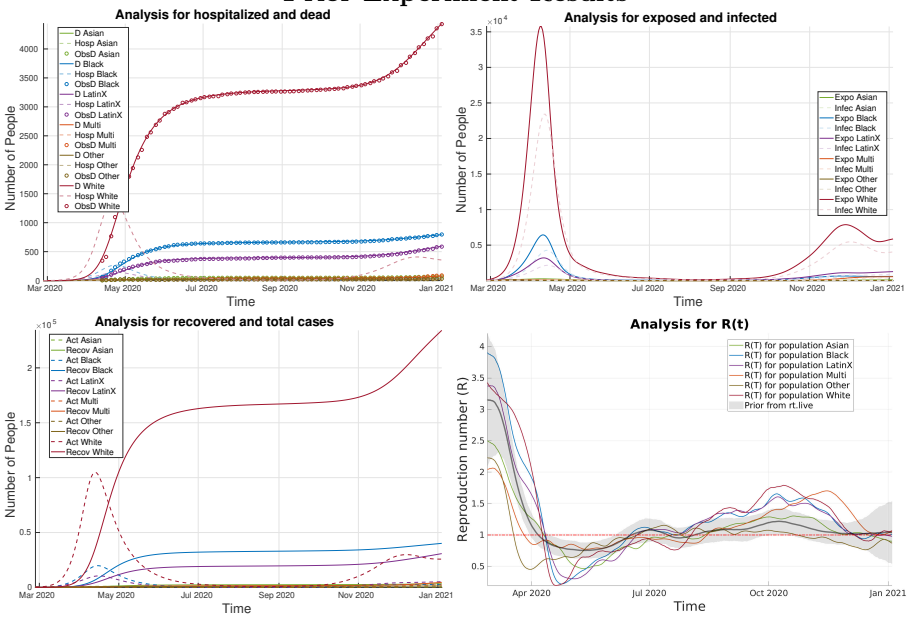

FIGURE 32. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.002$) and prior $R_t(n)$ ($\sigma_{R(t)} = 0.5$) curves coming from initial piece-wise assimilation also assuming a 0.020 CFR prior. This figure corresponds to columns 4 and 5 in Table 6.

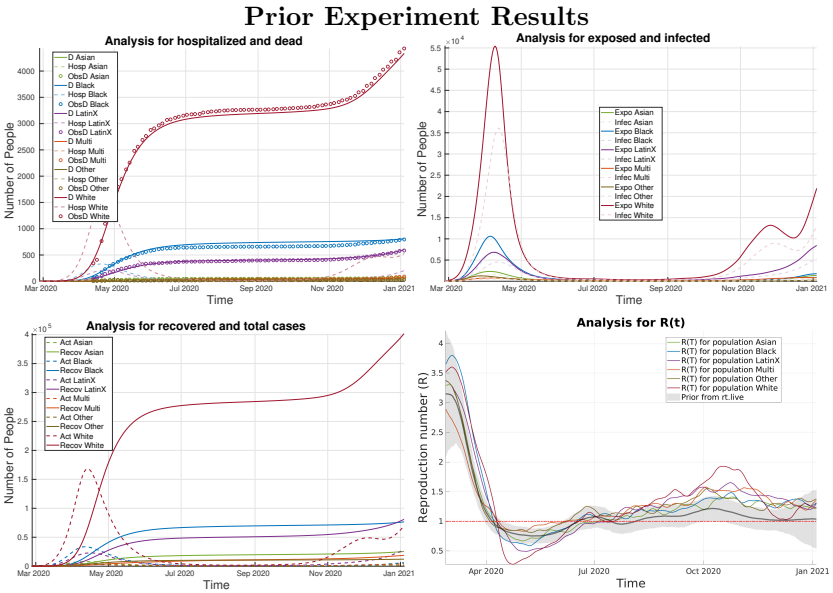

FIGURE 33. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.05$) for the Black population only and a CFR Prior of 0.009 ($\sigma_{CFR} = 0.05$) for all other populations. The prior $R_t(n)$ ($\sigma_{R(t)} = 1.5$) curves are the same for all populations. This figure corresponds to columns 2 and 3 in Table 7.

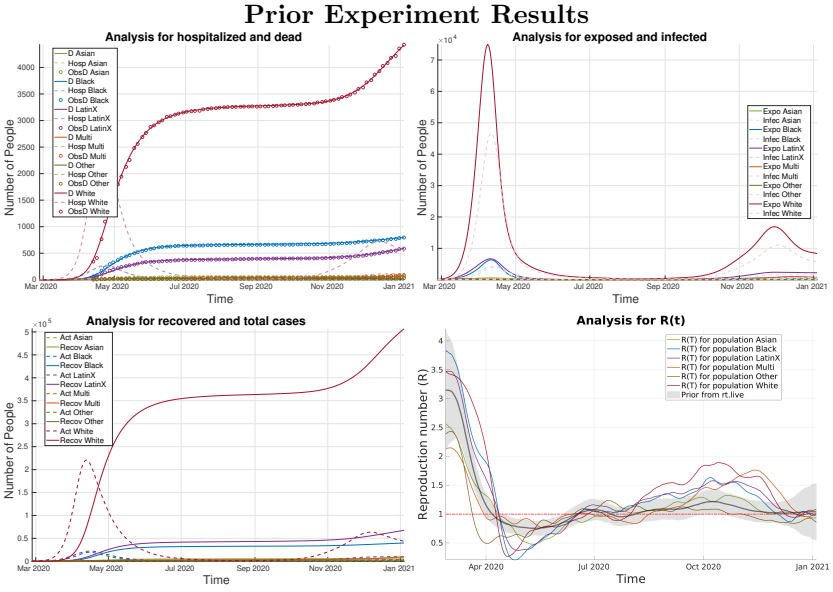

FIGURE 34. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.002$) for the Black population only and a CFR Prior of 0.009 ($\sigma_{CFR} = 0.0009$) for all other populations. The prior $R_t(n)$ ($\sigma_{R(t)} = 0.5$) curves coming from initial piece-wise assimilation also assuming a 0.020 CFR prior. This figure corresponds to columns 4 and 5 in Table 7 and is the only case where Black infections can be lower than that of the LatinX population and still have more deaths.

Prior Experiment Results

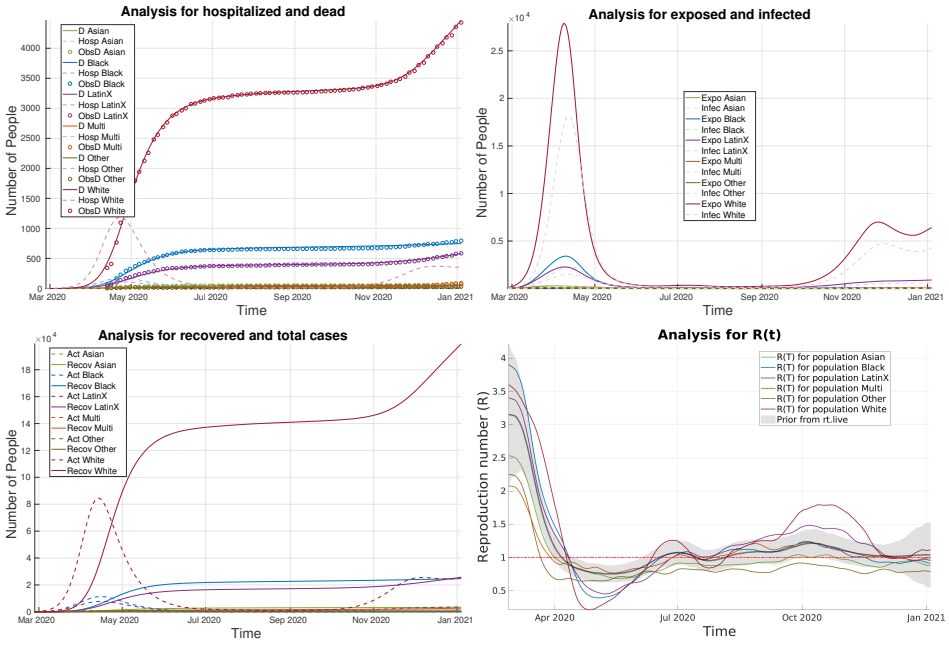

FIGURE 35. Analysis results with a CFR prior of 0.020 ($\sigma_{CFR} = 0.05$) for the Black population only and a CFR Prior of 0.009 ($\sigma_{CFR} = 0.05$) for all other populations. The prior $R_t(n)$ ($\sigma_{R(t)} = 0.5$) curves coming from initial piece-wise assimilation also assuming a 0.020 CFR prior. This figure corresponds to columns 6 and 7 in Table 7

Appendix D. Main results figures enlarged.

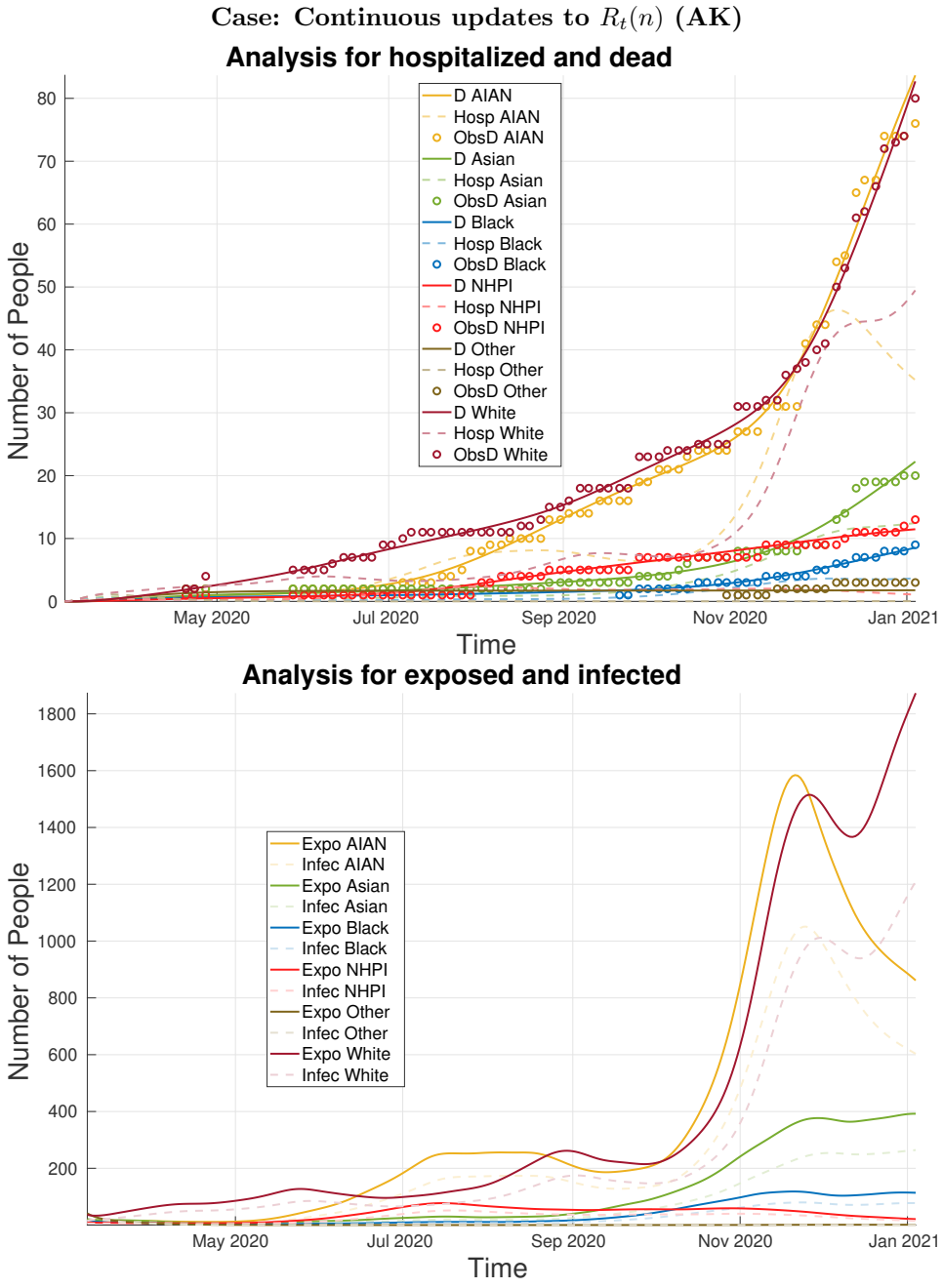

FIGURE 36. Analysis results for the continuous update case for the state of AK.

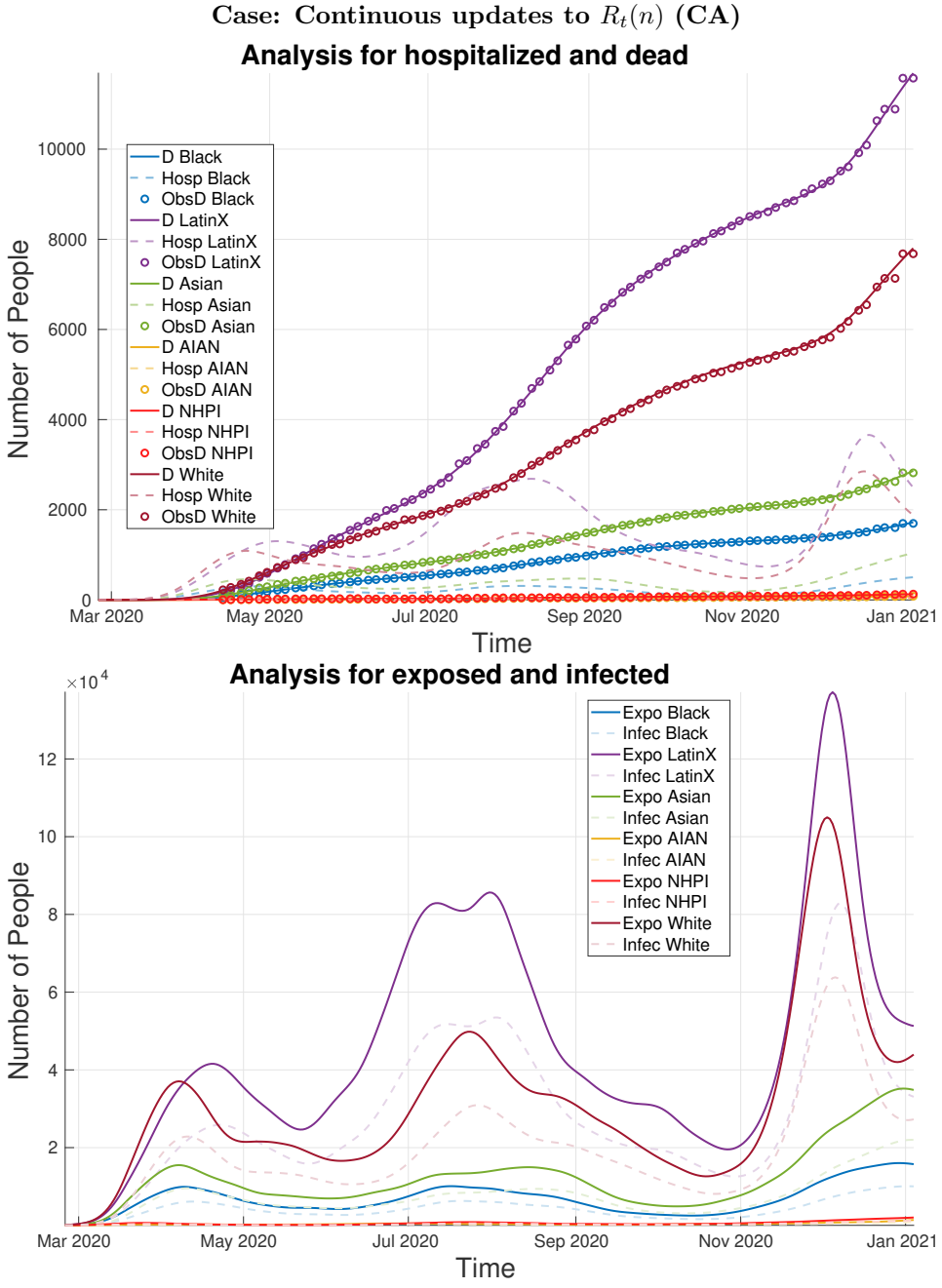

FIGURE 37. Analysis results for the continuous update case for the state of CA.

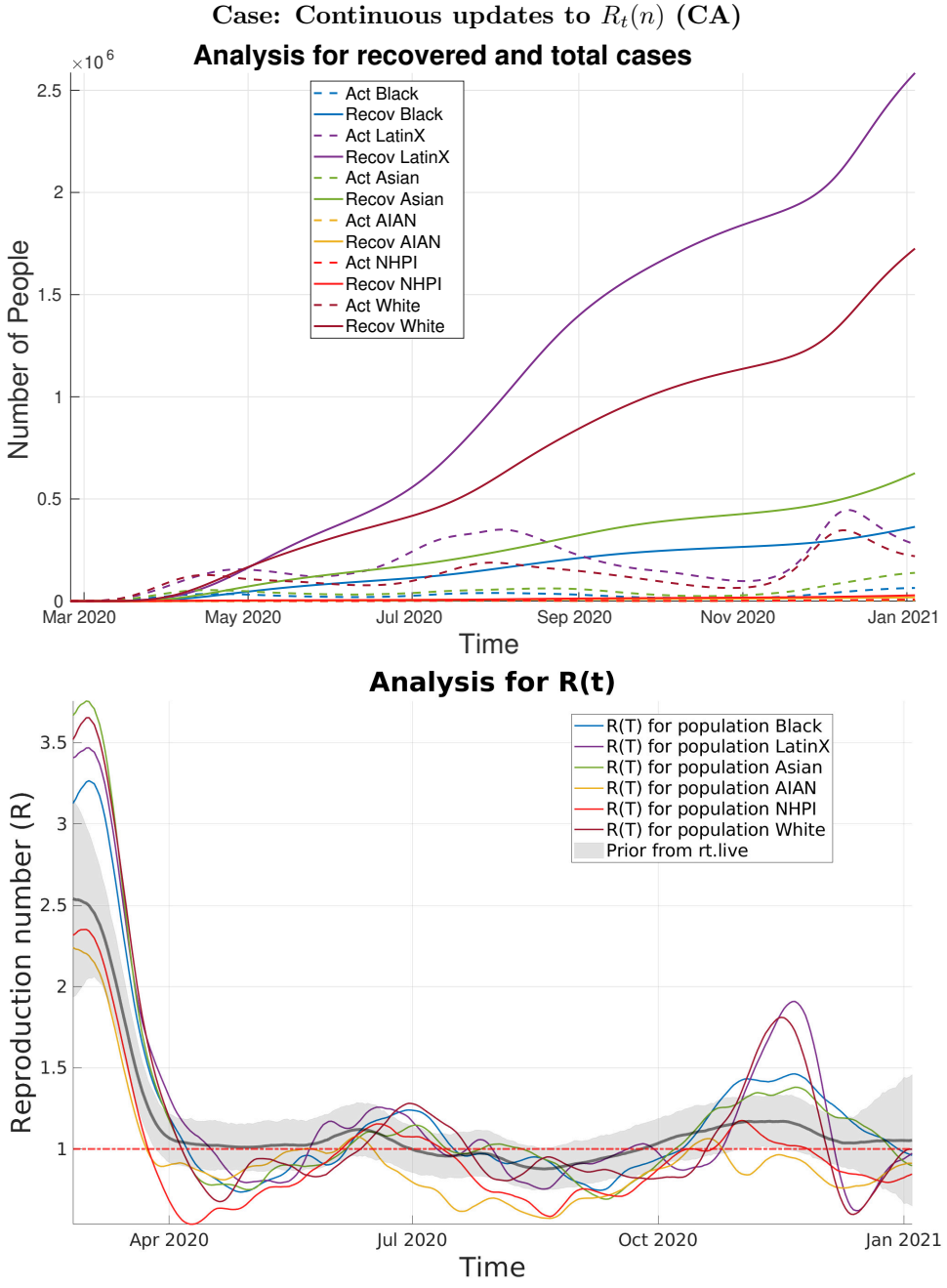

FIGURE 37. Analysis results for the continuous update case for the state of CA.

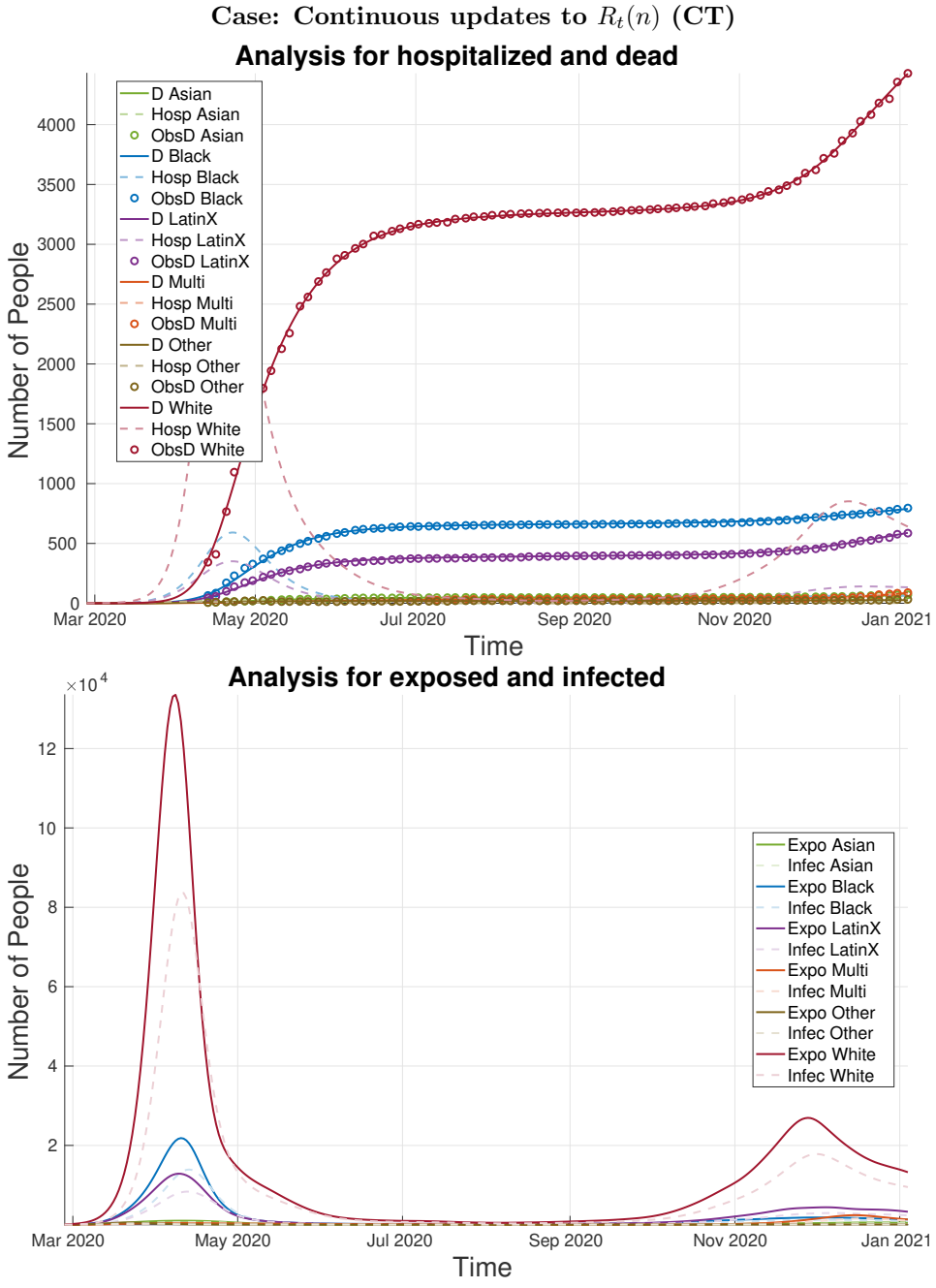

FIGURE 38. Analysis results for the continuous update case for the state of CT.

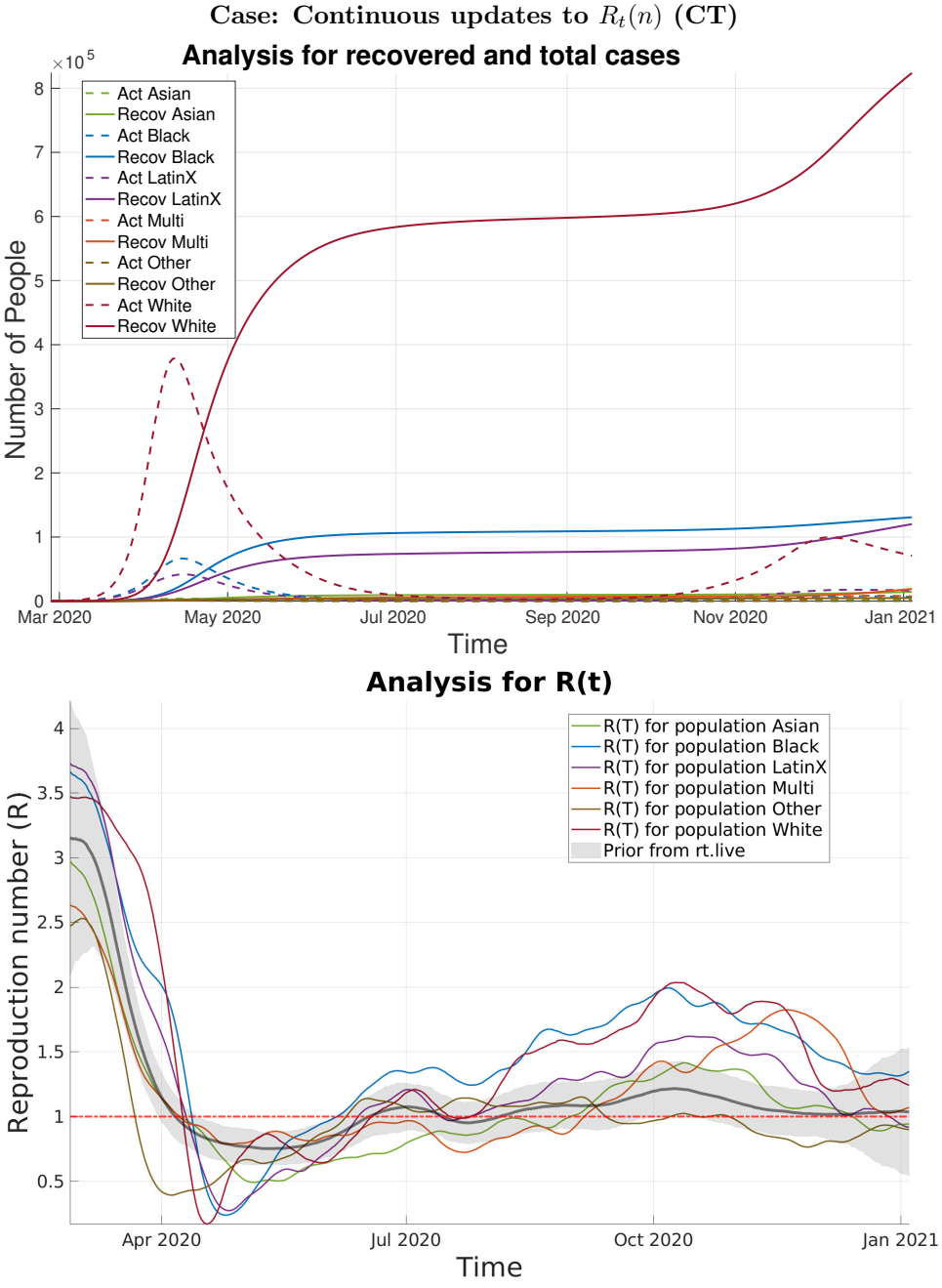

FIGURE 38. Analysis results for the continuous update case for the state of CT.

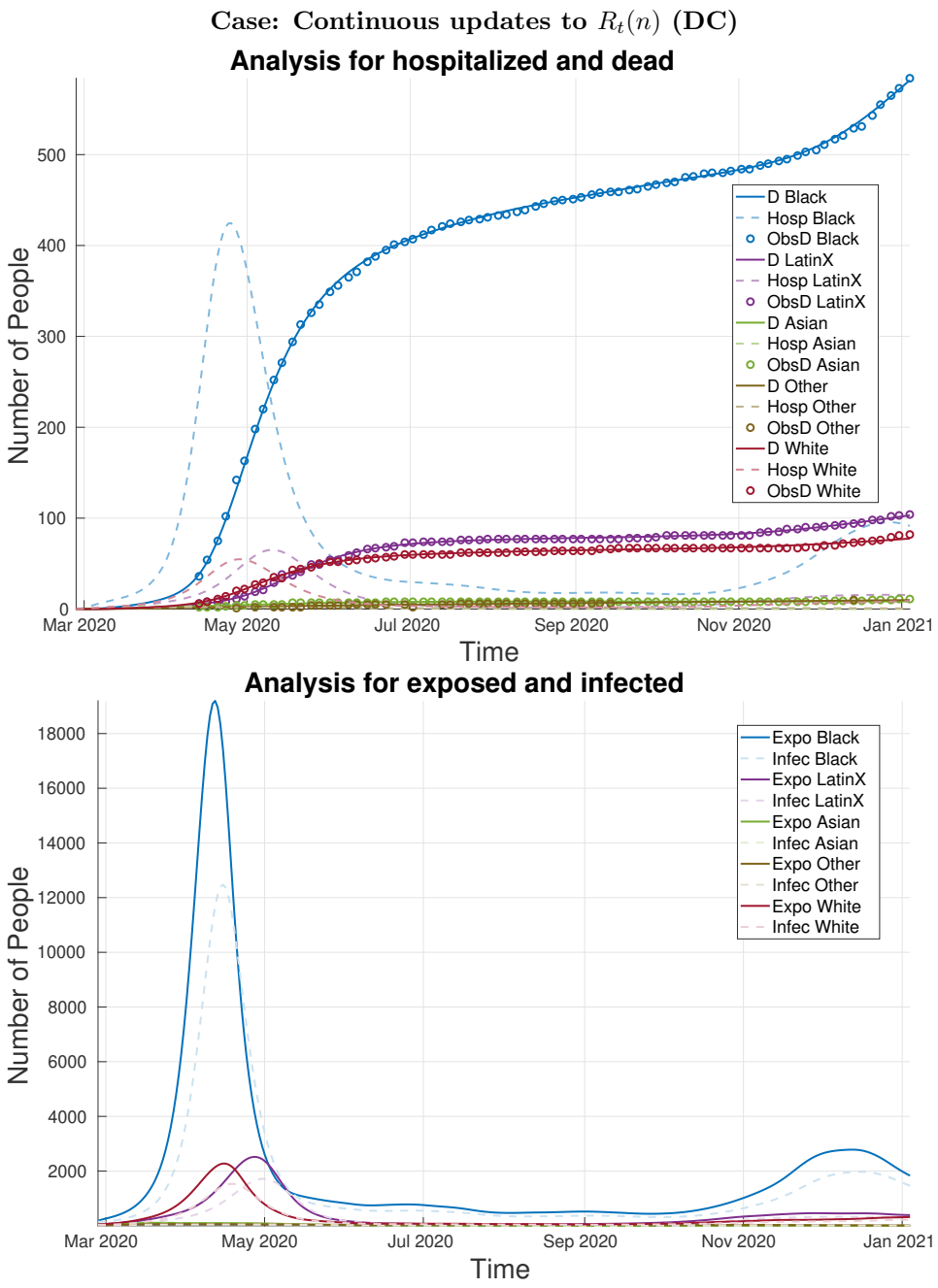

FIGURE 39. Analysis results for the continuous update case for the state of DC.

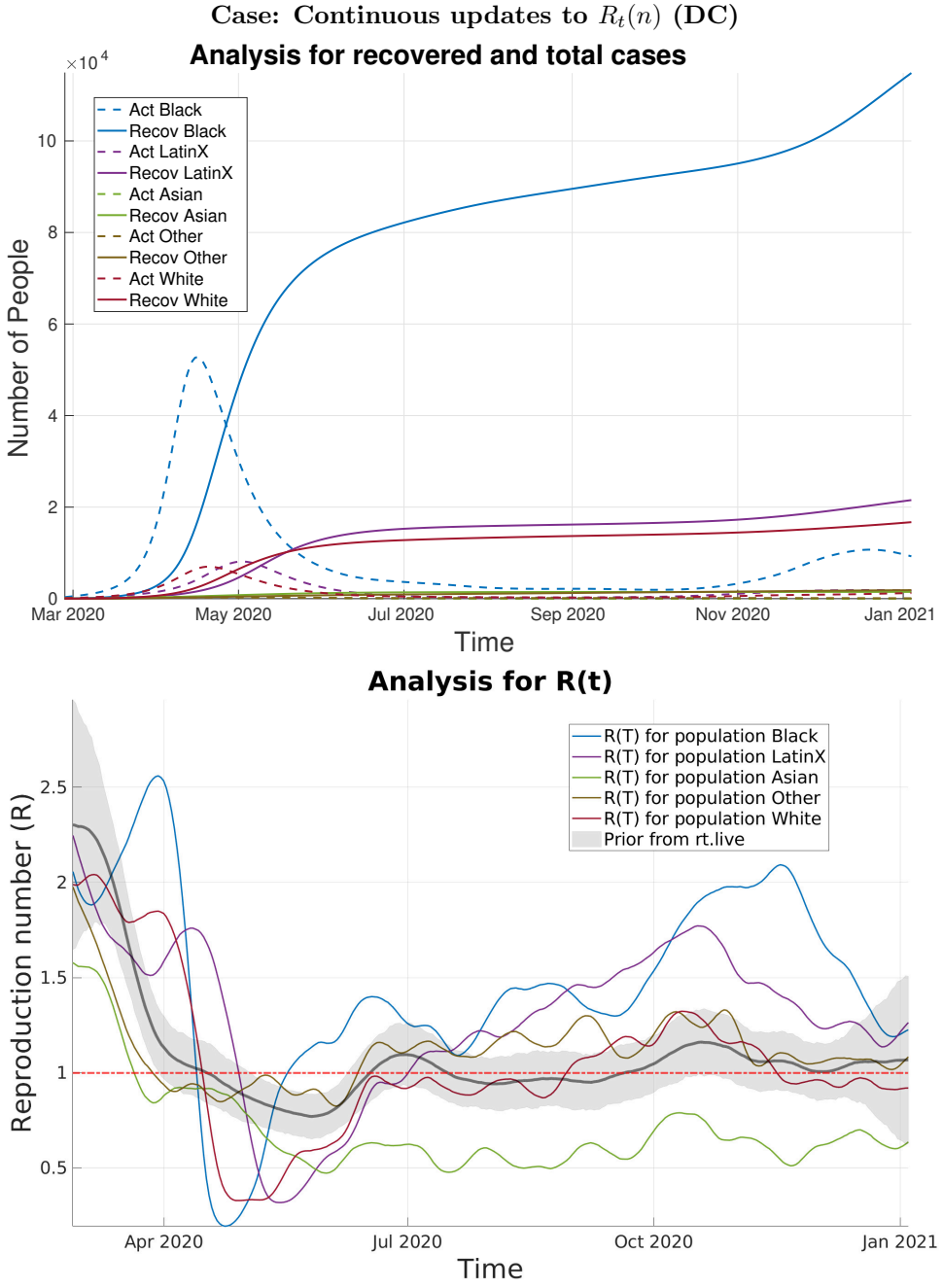

FIGURE 39. Analysis results for the continuous update case for the state of DC.

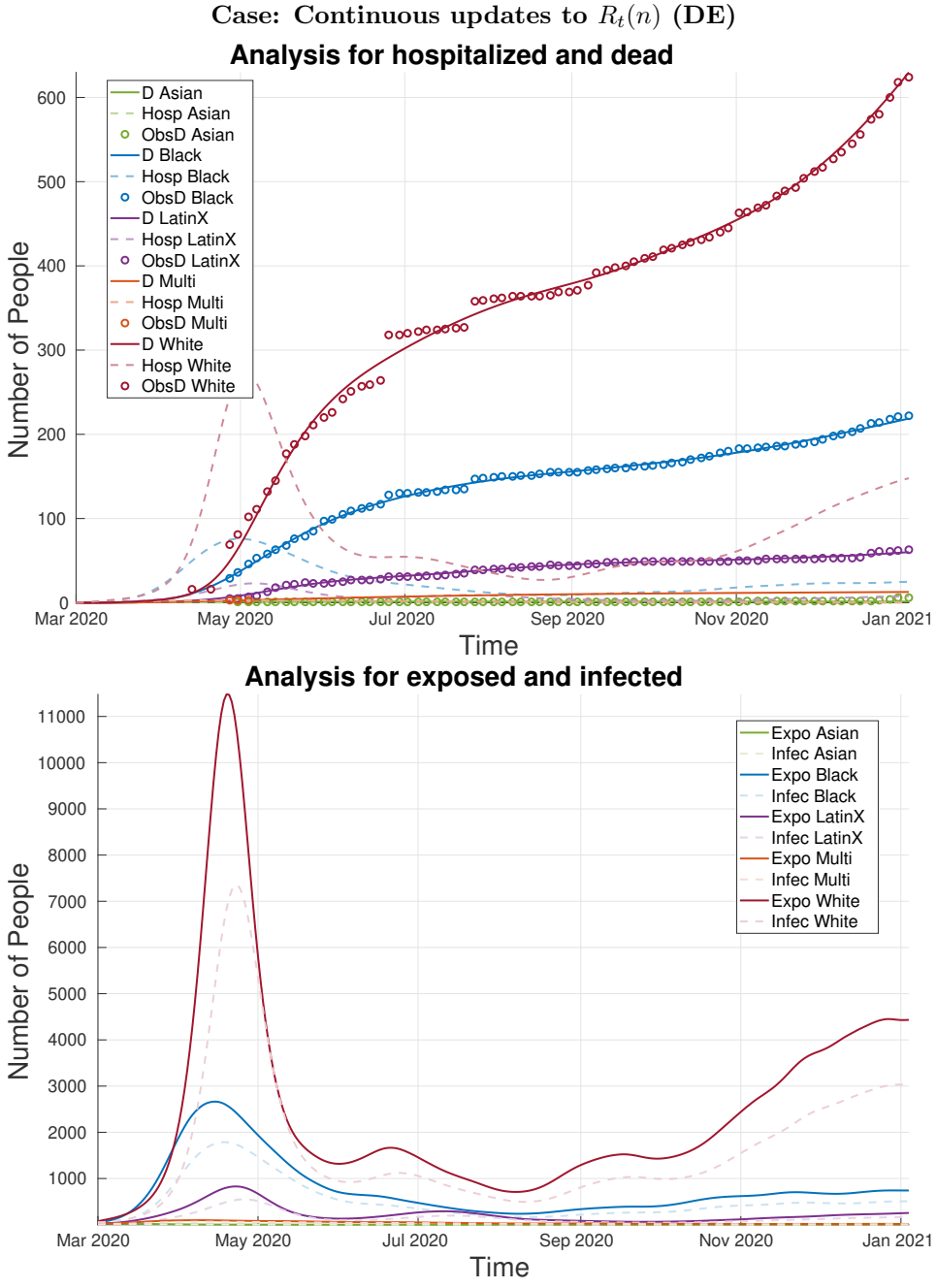

FIGURE 40. Analysis results for the continuous update case for the state of DE.

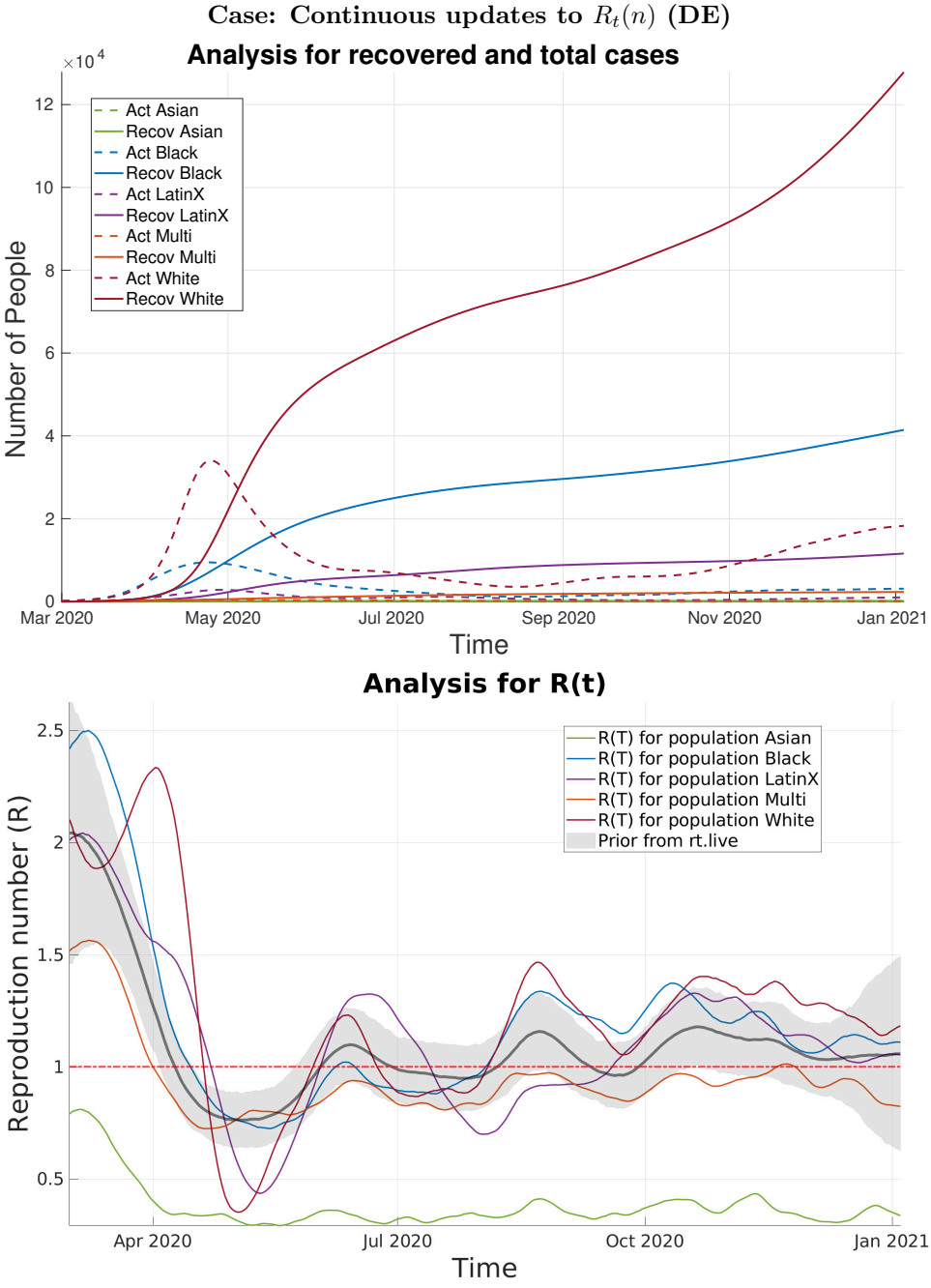

FIGURE 40. Analysis results for the continuous update case for the state of DE.

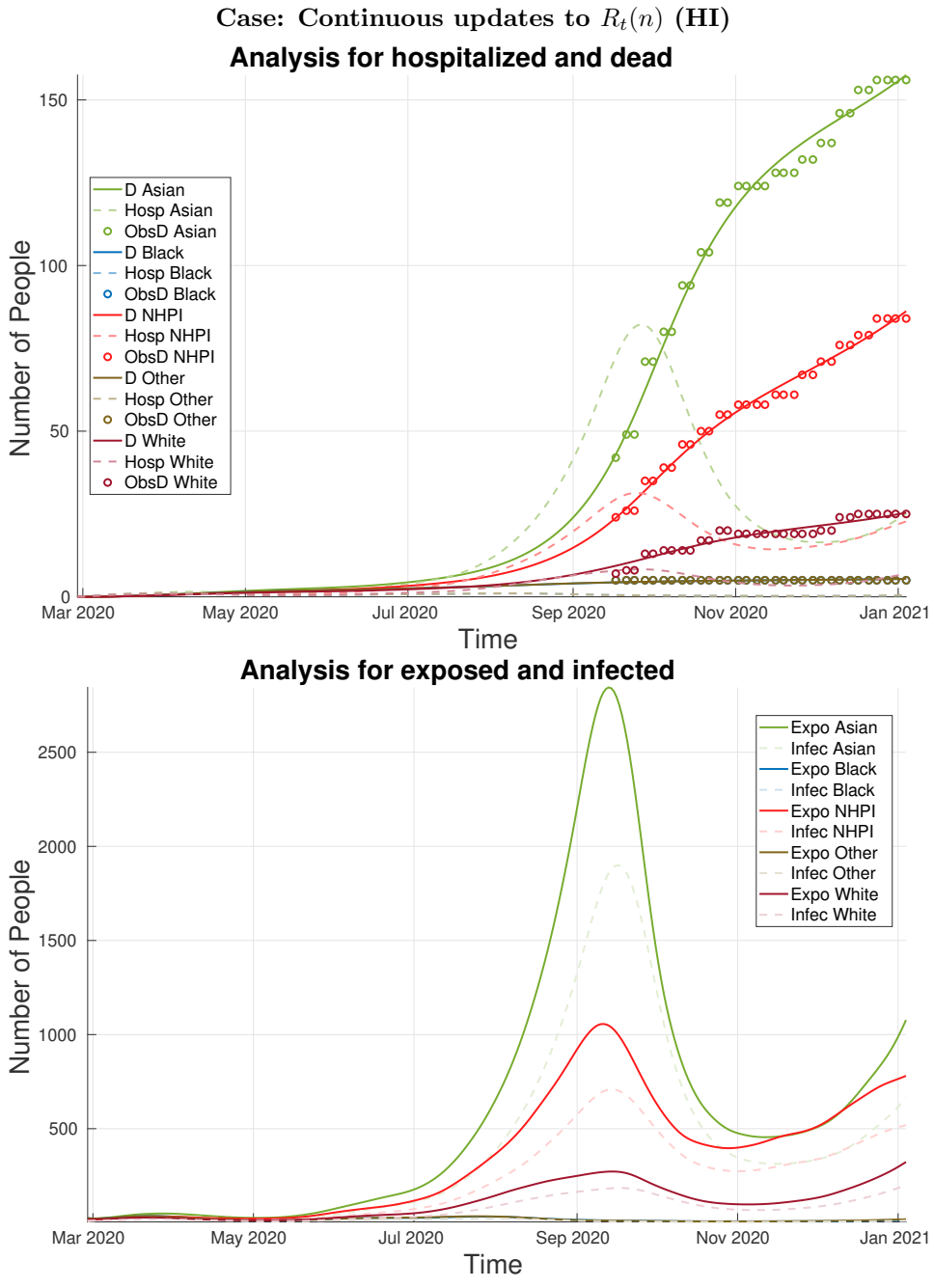

FIGURE 41. Analysis results for the continuous update case for the state of HI.

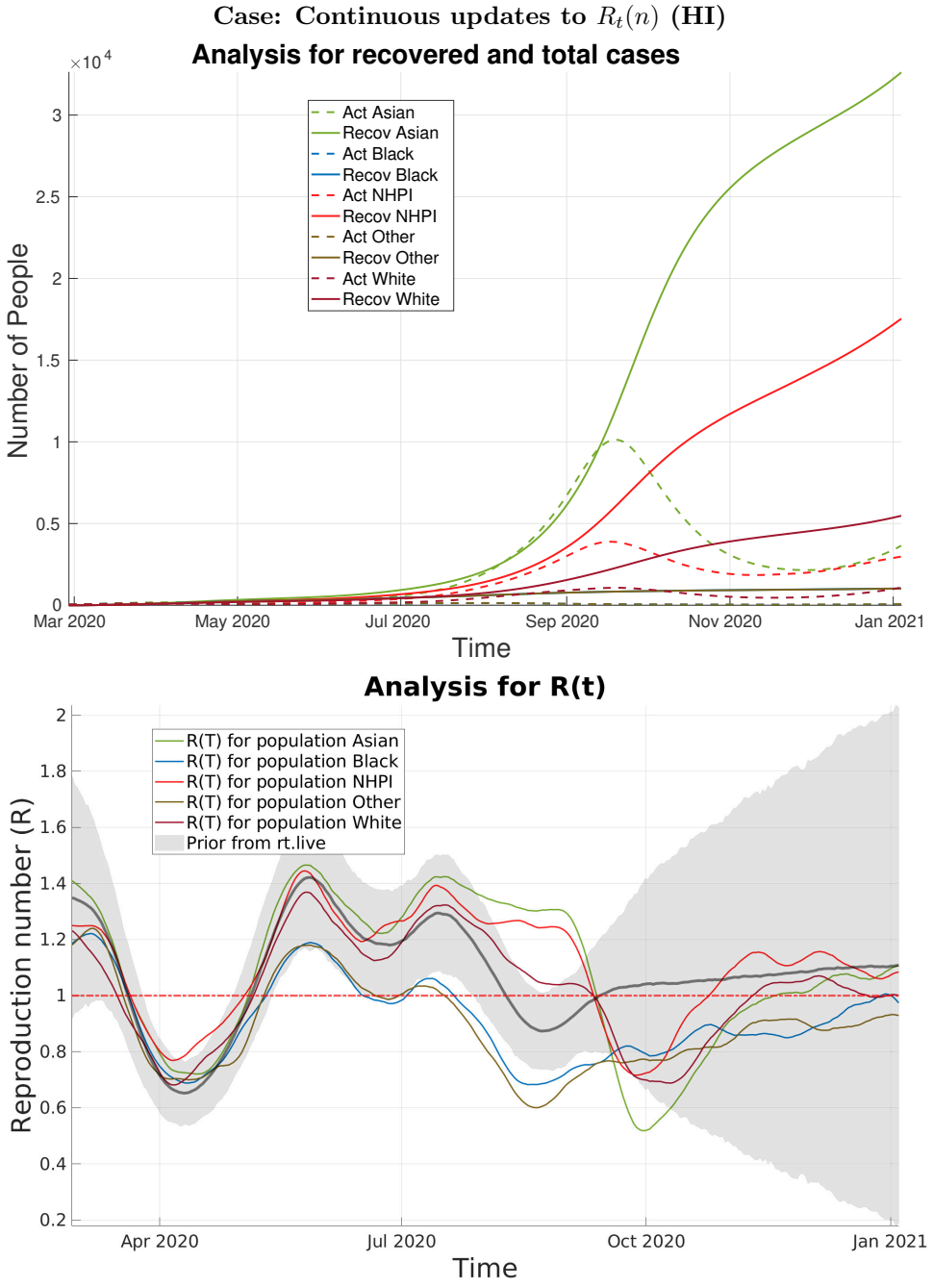

FIGURE 41. Analysis results for the continuous update case for the state of HI.

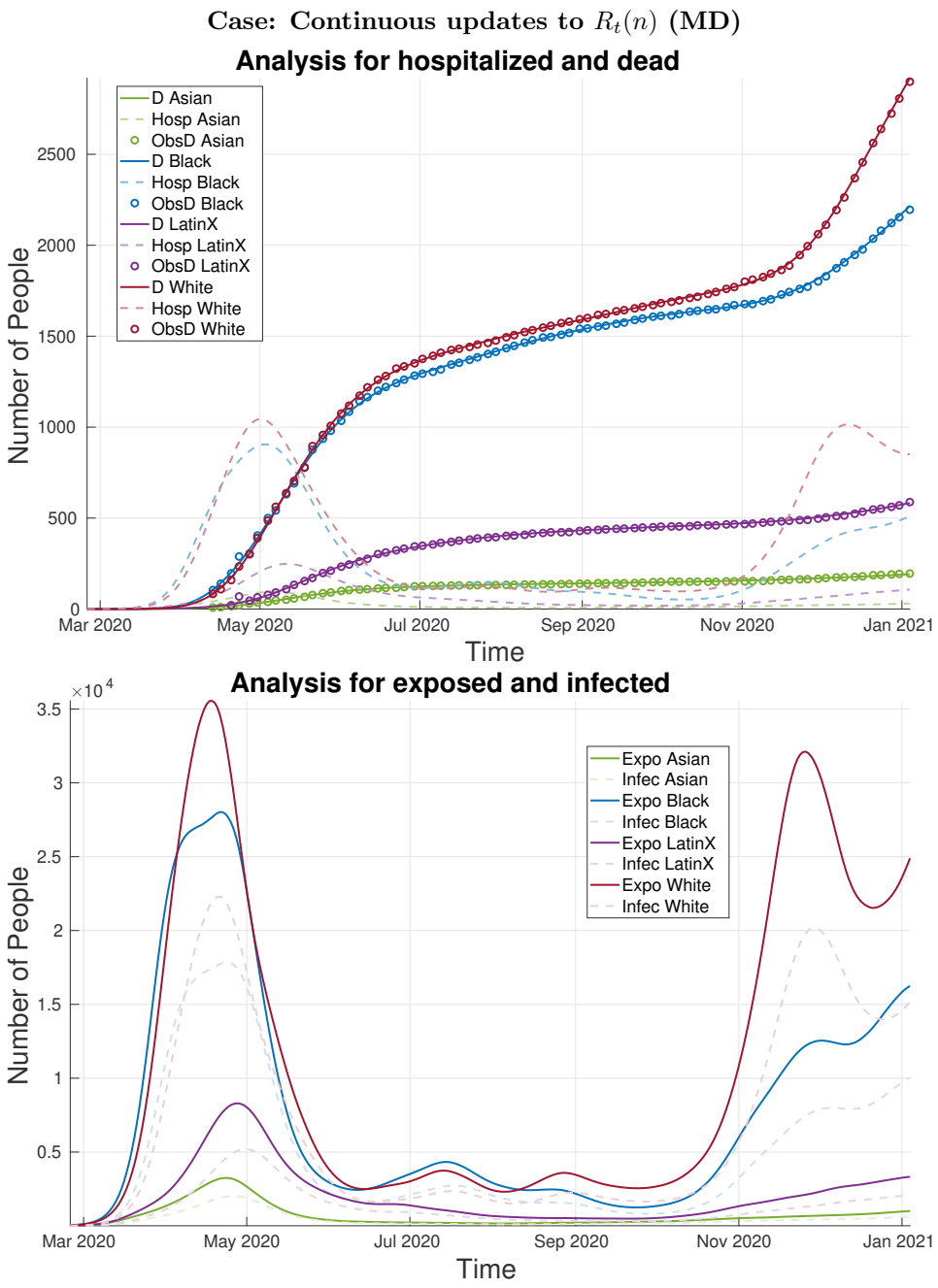

FIGURE 42. Analysis results for the continuous update case for the state of MD.

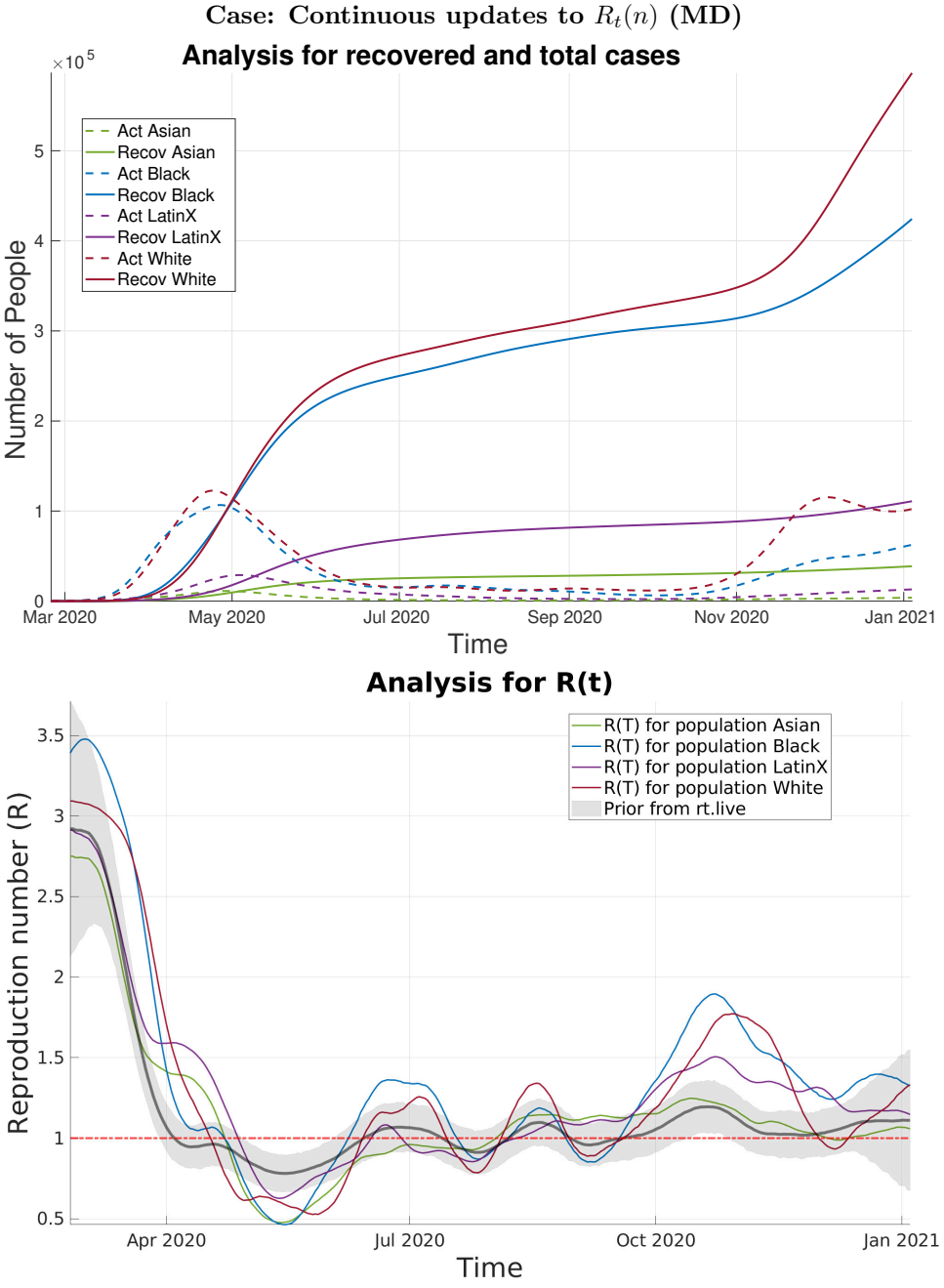

FIGURE 42. Analysis results for the continuous update case for the state of MD.

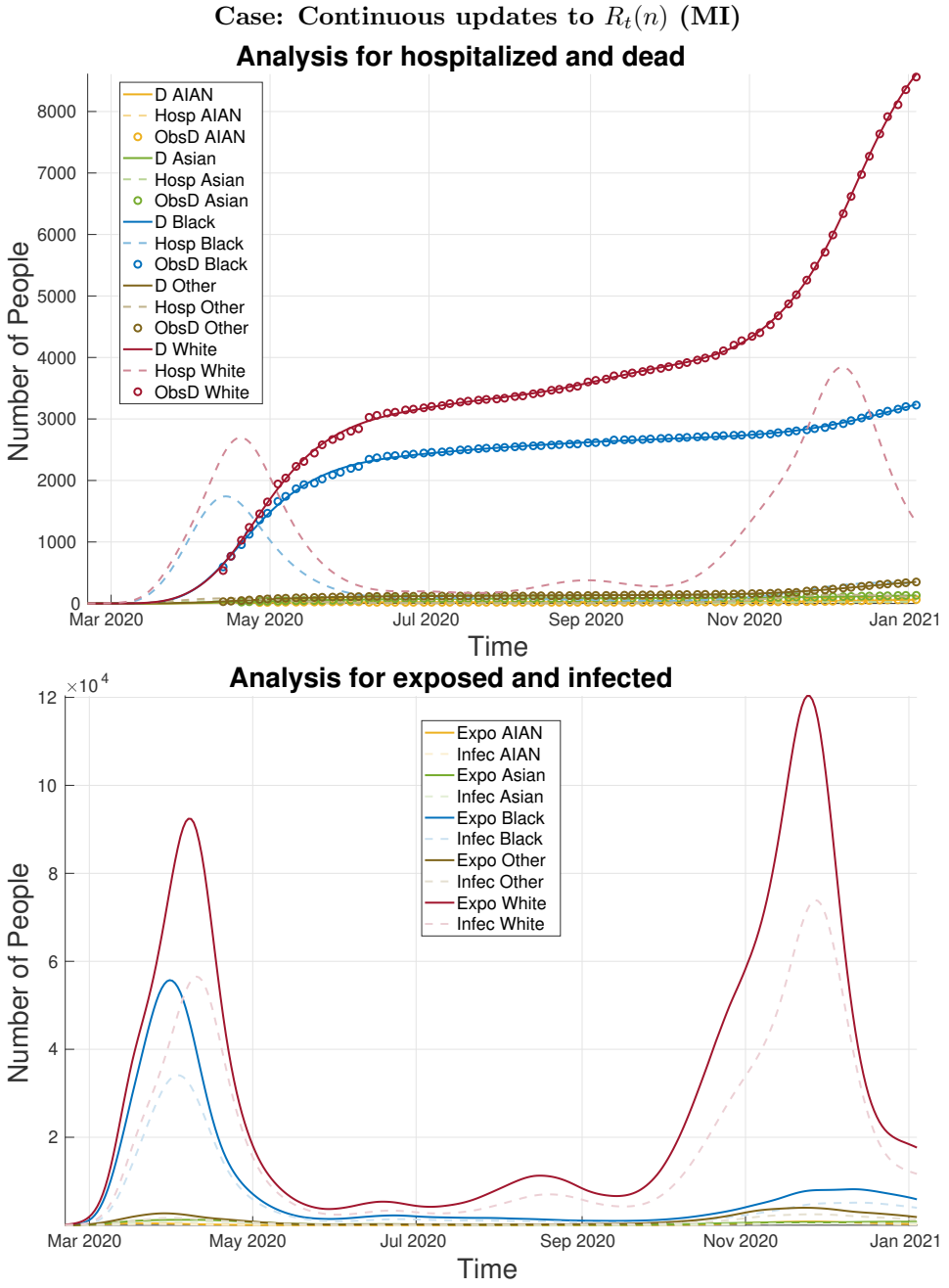

FIGURE 43. Analysis results for the continuous update case for the state of MI.

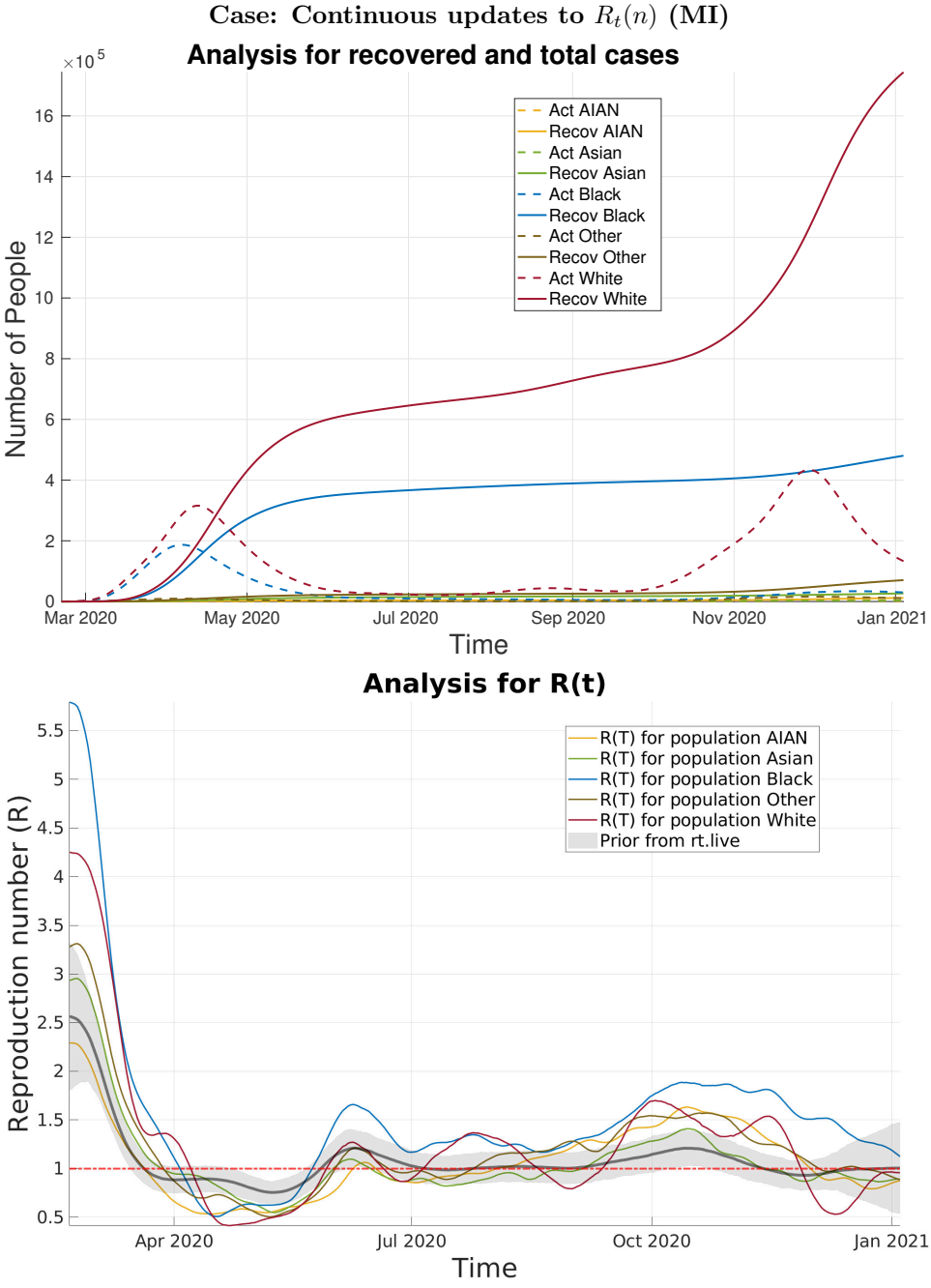

FIGURE 43. Analysis results for the continuous update case for the state of MI.

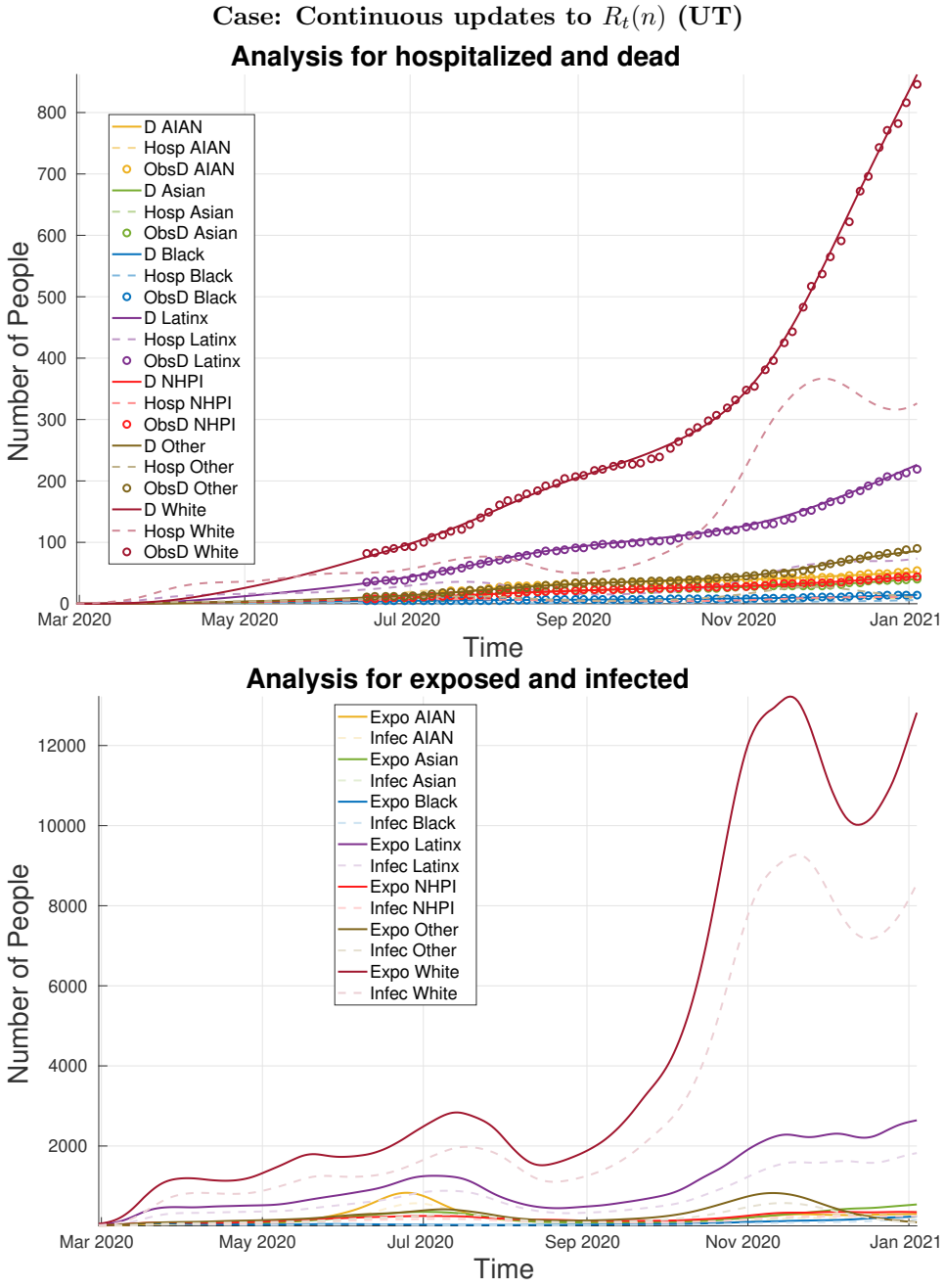

FIGURE 44. Analysis results for the continuous update case for the state of UT.

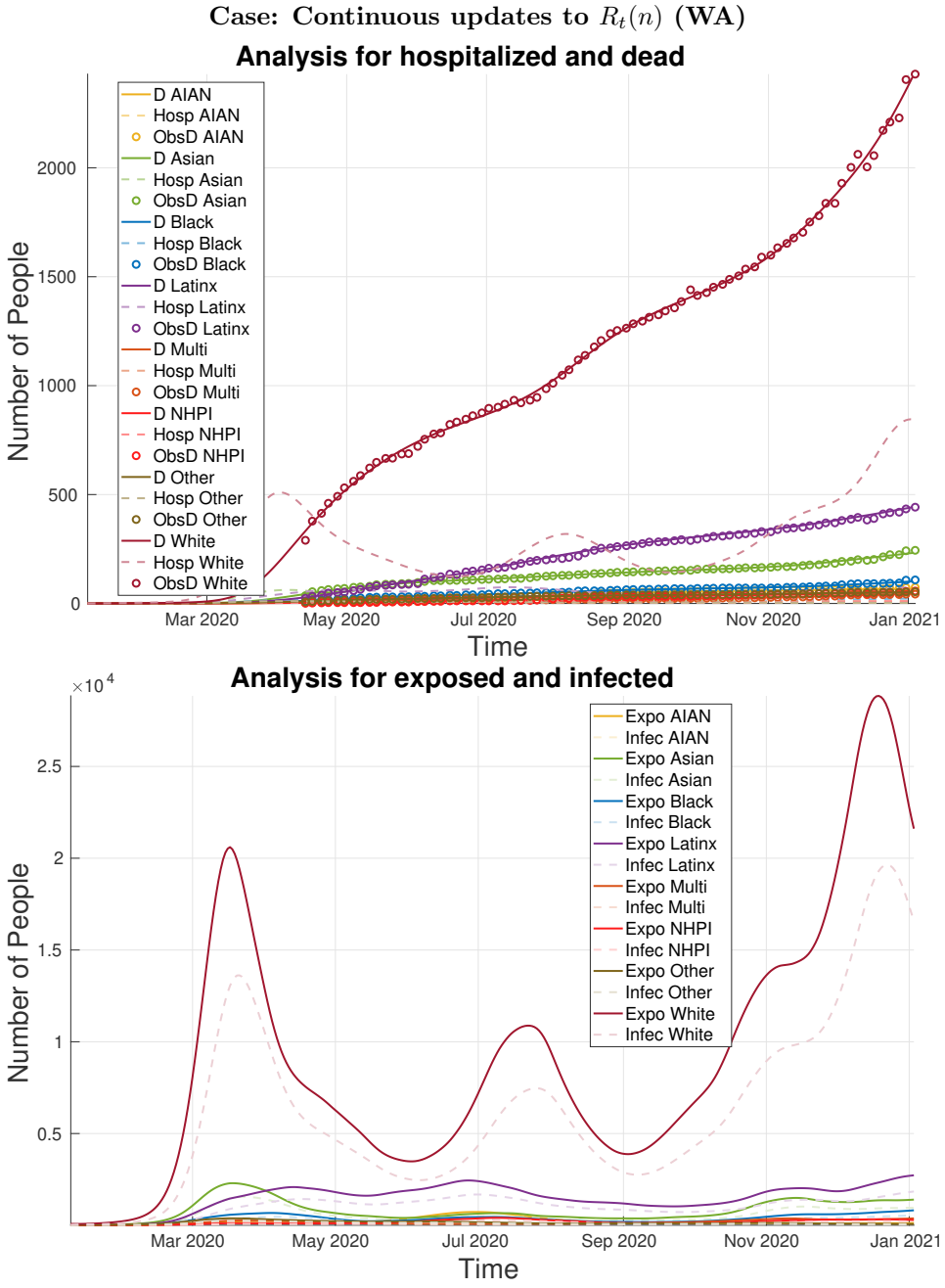

FIGURE 45. Analysis results for the continuous update case for the state of WA.

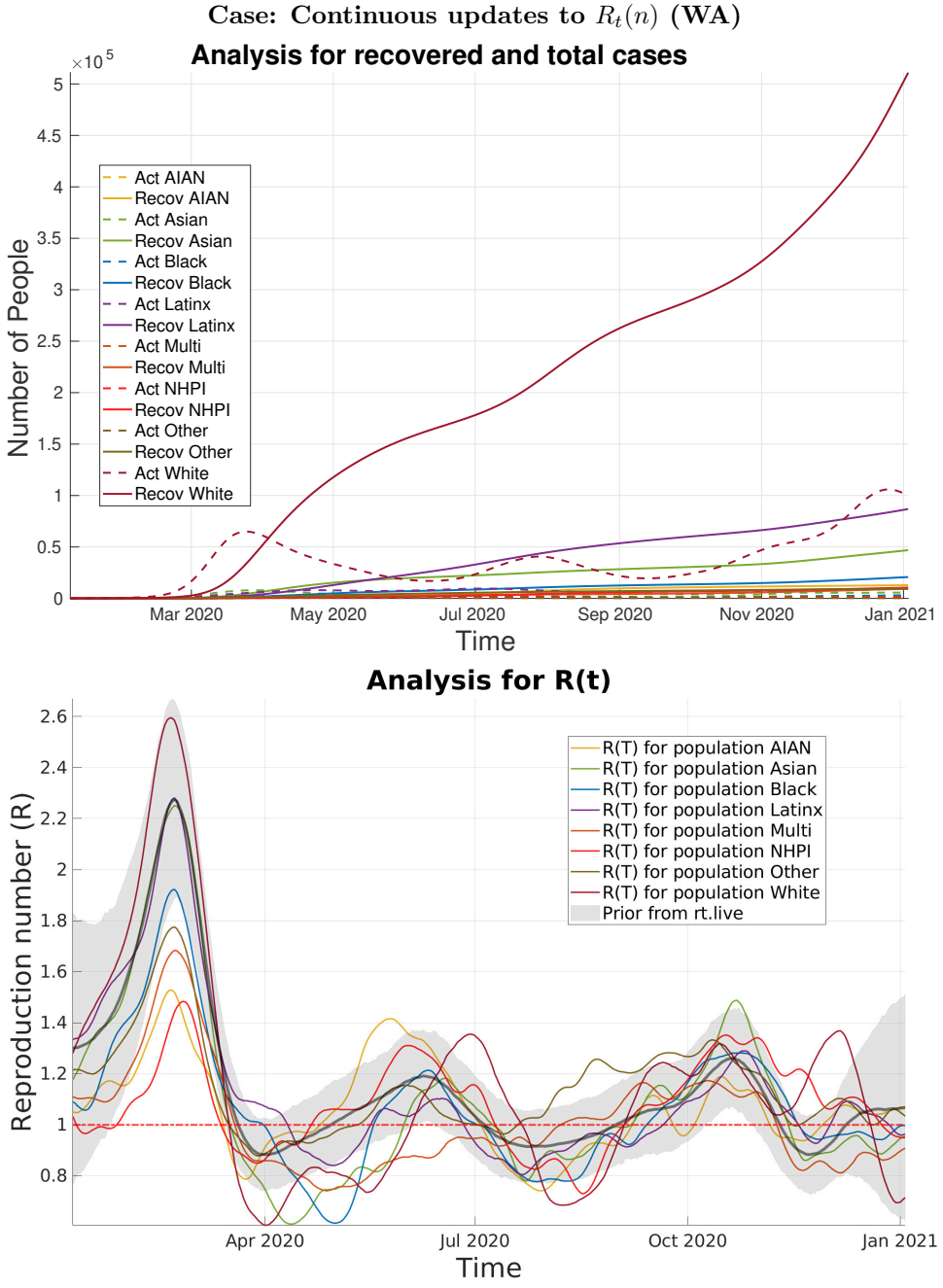

FIGURE 45. Analysis results for the continuous update case for the state of WA.

REFERENCES

[1] Cyberstates 2020: The definitive guide to the U.S. tech industry and tech wrokforce, URL https://www.cyberstates.org, Last accessed 2021-04-13.

[2] Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances, URL https://www.federalreserve.gov/econres/notes/feds-notes/disparities-in-wealth-by-race-and-ethnicity-in-the-2019-survey-of-consumer-finances-20200928.htm, Last accessed 2021-04-13.

[3] Diversity in high tech, URL https://www.eeoc.gov/special-report/diversity-high-tech, Last accessed 2021-04-13.

[4] Economy at a Glance: California, URL https://data.bls.gov/timeseries/LASST060000000000006?, Last accessed 2021-04-13.

[5] IHME COVID-19 estimates, URL http://www.healthdata.org/covid/data-downloads, Last accessed 2021-04-13.

[6] Options to Reduce Quarantine for Contacts of Persons with SARS-CoV-2 Infection Using Symptom Monitoring and Diagnostic Testing, URL https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-options-to-reduce-quarantine.html, Last accessed 2021-04-13.

[7] Racial Data Dashboard, 2021, URL https://covidtracking.com/race/dashboard, Last accessed 2021-04-13.

[8] Risk for covid-19 infection, hospitalization, and death by race/ethnicity, URL https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html, Last accessed 2021-08-03.

[9] Rt COVID-19, URL https://rt.live/, Last accessed 2021-04-13.

[10] Statistics and Church Facts | Total Church Membership, URL http://newsroom.churchofjesuschrist.org/facts-and-statistics/state/utah, Last accessed 2021-07-31.

[11] E. Armstrong, M. Runge and J. Gerardin, Identifying the measurements required to estimate rates of COVID-19 transmission, infection, and detection, using variational data assimilation, *Infectious Disease Modelling*.

[12] M. Asch, M. Bocquet and M. Nodet, *Data Assimilation: Methods, Algorithms, and Applications*, SIAM, Society for Industrial and Applied Mathematics, 2016.

[13] L. M. A. Bettencourt, R. M. Ribeiro, G. Chowell, T. Lant and C. Castillo-Chavez, Towards real time epidemiology: Data assimilation, modeling and anomaly detection of health surveillance data streams, *Lecture Notes in Computer Science Intelligence and Security Informatics: Biosurveillance*, 79–90.

[14] M. Bocquet and P. Sakov, An iterative ensemble Kalman smoother, *Quarterly Journal of the Royal Meteorological Society*, **140** (2013), 1521–1535.

[15] A. Carrassi, M. Bocquet, L. Bertino and G. Evensen, Data assimilation in the geosciences: An overview on methods issues and perspectives, *WCC*, **9** 2018.

[16] A. A. Emerick and A. C. Reynolds, Ensemble smoother with multiple data assimilation, *Computers & Geosciences*, **55** (2013), 3–15.

[17] G. Evensen, Analysis of iterative ensemble smoothers for solving inverse problems, *Computational Geosciences*, **22** (2018), 885–908.

[18] G. Evensen, J. Amezcua, M. Bocquet, A. Carrassi, A. Farchi, A. Fowler, P. L. Houtekamer, C. K. Jones, R. J. de Moraes, M. Pulido, C. Sampson and F. C. Vossepoel, An international initiative of predicting the SARS-CoV-2 pandemic using ensemble data assimilation, *Foundations of Data Science*, (2020).

[19] J. R. Eyre, S. J. English and M. Forsythe, Assimilation of satellite data in numerical weather prediction. part i: The early years, *Quarterly Journal of the Royal Meteorological Society*, **146** (2019), 49–68.

[20] A. L. Garcia-Basteiro, G. Moncunill, M. Tortajada, M. Vidal, C. Guinovart, A. Jiménez, R. Santano, S. Sanz, S. Méndez, A. Llupià, R. Aguilar, S. Alonso, D. Barrios, C. Carolis, P. Cisteró, E. Chóliz, A. Cruz, S. Fochs, C. Jairoce, J. Hecht, M. Lamoglia, M. J. Martínez, R. A. Mitchell, N. Ortega, N. Pey, L. Puyol, M. Ribes, N. Rosell, P. Sotomayor, S. Torres, S. Williams, S. Barroso, A. Vilella, J. Muñoz, A. Trilla, P. Varela, A. Mayor and C. Dobaño, Seroprevalence of antibodies against SARS-CoV-2 among health care workers in a large spanish reference hospital, *Nature Communications*, **11** (2020), Article number: 3500.

[21] C. G. Grijalva, M. A. Rolfes, Y. Zhu, H. Q. McLean, K. E. Hanson, E. A. Belongia, N. B. Halasa, A. Kim, C. Reed, A. M. Fry and H. K. Talbot, Transmission of SARS-COV-2 infections in households — Tennessee and Wisconsin, April–September 2020, *MMWR. Morbidity and Mortality Weekly Report*, **69** (2020), 1631–1634.
[22] P. L. Houtekamer and F. Zhang, Review of the ensemble kalman filter for atmospheric data assimilation, *Monthly Weather Review*, **144** (2016), 4489–4532.
[23] J. P. A. Ioannidis, Infection fatality rate of COVID-19 inferred from seroprevalence data, *Bulletin of the World Health Organization*, **99** (2020), 19–33F.
[24] J. Jeppesen, Fact sheet: Reanalysis, URL `https://www.ecmwf.int/en/about/media-centre/focus/2020/fact-sheet-reanalysis`, 2020, Last accessed 2021-07-31.
[25] E. J. Kostelich, Y. Kuang, J. M. Mcdaniel, N. Z. Moore, N. L. Martirosyan and M. C. Preul, Accurate state estimation from uncertain data and models: An application of data assimilation to mathematical models of human brain tumors, *Biology Direct*, **6** (2011), 64.
[26] W. Lieberman-Cribbin, S. Tuminello, R. M. Flores and E. Taioli, Disparities in COVID-19 testing and positivity in new york city, *American Journal of Preventive Medicine*, **59** (2020), 326–332.
[27] N. Narea, Immigrants have helped keep essential services running. But those without legal status have no financial safety net, URL `https://www.vox.com/2020/5/5/21244630/undocumented-immigrants-coronavirus-relief-cares-act`, 2020, Last accessed 2021-07-31.
[28] I. Pathak, Y. Choi, D. Jiao, D. Yeung and L. Liu, Racial-ethnic disparities in case fatality ratio narrowed after age standardization: A call for race-ethnicity-specific age distributions in state covid-19 data, *MedRxiv*, (2020).
[29] J. Skjervheim, G. Evensen, J. Hove and J. G. Vabø, An ensemble smoother for assisted history matching, *SPE*, (2011), 141929.
[30] A. S. Stordal and A. H. Elsheikh, Iterative ensemble smoothers in the annealed importance sampling framework, *Advances in Water Resources*, **86** (2015), 231–239.
[31] G. Vernieres, A. Anis, R. N. Miller and L. L. Ehret, Generalized inversion of thermistor-chain data and a layer model of lake kinneret, *Ocean Modelling*, **12** (2006), 112–139.
[32] Z. Wu, T. Phan, J. Baez, Y. Kuang and E. J. Kostelich, Predictability and identifiability assessment of models for prostate cancer under androgen suppression therapy, *Mathematical Biosciences and Engineering*, **16** (2019), 3512–3536.

Received January 2021; 1st revision May 2021; 2nd revision August 2021; early access September 2021.

E-mail address: `efleurantin2013@fau.edu`
E-mail address: `christian.sampson@gmail.com`
E-mail address: `dmaes@umich.edu`
E-mail address: `jkbenne6@asu.edu`
E-mail address: `fernandesnunez.t@northeastern.edu`
E-mail address: `sophiaemarx@gmail.com`
E-mail address: `geev@norceresearch.no`