Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

MODEL PREDICTIVE CONTROL IN OPTIMAL INTERVENTION OF COVID-19 WITH MIXED EPISTEMIC-ALEATORIC UNCERTAINTY

Jinming Wan

Department of Systems Science and Industrial Engineering
Binghamton University
4400 Vestal Parkway East
Binghamton, NY 13902, USA

Saeideh Mirghorbani

Department of Business Analytics and Operations

Binghamton University
4400 Vestal Parkway East
Binghamton, NY 13902, USA

N. Eva Wu

Department of Electrical and Computer Engineering
Binghamton University
4400 Vestal Parkway East
Binghamton, NY 13902, USA

Changqing Cheng

Department of Systems Science and Industrial Engineering
Binghamton University
4400 Vestal Parkway East
Binghamton, NY 13902, USA

ABSTRACT

Non-pharmaceutical interventions (NPI) have been proven vital in the fight against the COVID-19 pandemic before the massive rollout of vaccinations. Considering the inherent epistemic-aleatoric uncertainty of parameters, accurate simulation and modeling of the interplay between the NPI and contagion dynamics are critical to the optimal design of intervention policies. We propose a modified SIRD-MPC model that combines a modified stochastic Susceptible-Infected-Recovered-Deceased (SIRD) compartment model with mixed epistemic-aleatoric parameters and Model Predictive Control (MPC), to develop robust NPI control policies to contain the infection of the COVID-19 pandemic with minimum economic impact. The simulation result indicates that our proposed model can significantly decrease the infection rate compared to the practical results under the same initial conditions.

1 INTRODUCTION

The COVID-19 pandemic has taken a substantial economic and societal toll worldwide (Wan et al. 2022a). Before the mass rollout of vaccination and at the early stages of the outbreak, the infection went rampant globally, owing to human mobility and uncoordinated and ineffective non-pharmaceutical interventions (NPI). Since the pandemic has almost waned and the world has returned to normalcy, it is essential to perform a retrospective analysis to better understand the interplay between intervention policies and the infection dynamics, to enhance the preparedness for future pandemic outbreaks and the resilience of the whole society. Notably, with a lack of knowledge about the novel coronavirus and insufficient and inaccurate data, epistemic and aleatory uncertainty have hindered effective prediction of contagion evolution. Aleatory uncertainty refers to a process's innate randomness and variability, and epistemic uncertainty refers to a lack of knowledge or understanding about the process.

Compartment models have been predominantly used in modeling the infection dynamics of epidemic outbreaks (Chen et al. 2019; Chowell et al. 2003; Jin et al. 2011; Xia et al. 2015). Specifically, the

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

population is divided into non-overlapping groups or compartments representing, for example, those who are susceptible (S), infected (I), recovered (R), and deceased (D). The SIRD model and its variants have been widely used to predict case counts for COVID-19 and the optimal design of NPIs (Armaou et al. 2022; Calafiore et al. 2020). The large number of infections in a short period could crimp the hospital system and exhaust the healthcare resources for non-COVID patients. In an attempt to contain infections, NPIs inevitably strain the economy and social functions (Miller et al. 2020; Polcz et al. 2022). A good NPI strategy seeks a delicate trade-off between economic loss and infections (Lemaitre et al. 2022; Scarabaggio et al. 2022; Sharomi and Malik 2017; She et al. 2022). Model predictive control (MPC) has proven effective in controlling various real-world processes in which future dynamics are highly uncertain. (Armaou et al. 2022; Péni and Szederkényi 2021). Consequently, combining MPC with the compartment model is a promising approach to designing policies to contain the spread of diseases.

As in most complex systems, parameters in a compartmental model could be rather challenging to infer from the available data owing to a lack of knowledge, incomplete and inaccurate data, as well as computational issues (Gallo et al. 2022). Conventional compartmental models, regarded as deterministic, could suffer severely from poor predictive power. Uncertainty quantification in parametric compartment models could also help formalize a more robust design of control policies (Wu and Mortveit 2015). Stochastic compartmental models, moreover, have been developed and applied to the investigation of the COVID-19 pandemic (Mamis and Farazmand 2023) and quantify uncertainties in such models.

In this study, we consider a mixed epistemic-aleatoric uncertainty for parameters in compartmental models in the context of COVID-19 for two reasons. First, the knowledge gap is narrowing as more studies are conducted on COVID-19 and its variants; and second, the uncertainty of infectious system parameters cannot be completely eliminated by acquiring new knowledge or data, due to inherent stochasticity and individual differences. Therefore, it is more reasonable to characterize the extensively studied COVID-19 pandemic in compartmental models using a mixed epistemic-aleatoric parameter. This paper uses probability bounds analysis (PBA) to quantify the uncertainty of the SIRD compartment model. PBA is a collection of mathematical tools that extends the ideas of interval analysis and probability theory, and may be used to quantify both random and deterministic forms of uncertainty in a wide range of scientific endeavors. Since it does not make any optimistic assumptions about parameter values, distribution shapes, or correlations between variables, PBA excels when there is a paucity of information regarding such variables (Gray et al. 2022).

Moreover, conventional compartmental models have limitations in capturing the complexities of the dynamics of the realistic interventional COVID-19 pandemic due not only to the deterministic and constant nature of the model parameters but also to the neglectfulness of human society interventions during the pandemic. Indeed, the effects of human society interventions, as described by various NPI policies implemented by policymakers, are not insignificant. The impact of human society interventions of COVID-19 essentially enters the infection process via feedback from some observables (March et al. 2022), and can significantly alter the population dynamics more than the parameters in the conventional models. Therefore, optimal control to develop practical policies to reduce the realistic pandemic should involve the integration of collected data and the implemented human society interventions to avoid implementing inappropriate policies. One example of this is the incorporation of pandemic data collected on the use of masks into pandemic models that consider the presence of human society interventions. Failure to consider these interventions can result in control policies that do not align with policy design expectations, due to discrepancies between the control effort of the developed policy (treating wearing masks as the minimum intervention) and the control effort of policy implementation (treating non-intervention as the minimum intervention). Therefore, it is inevitable that attempts to capture the complexity of the dynamics of a realistic interventional COVID-19 pandemic will involve some degree of uncertainty in the model parameters and the dynamics of human society interventions. In this study, we introduce additional population feedback parameters to model the impact of human society interventions on a SIRD compartment model, to control a set of models further, and finally assess the simulated achievable robustness in combatting COVID-19 infection.

We propose a modified stochastic SIRD-MPC model to develop robust control policies to combat the spread of disease with minimal social and economic impacts. We use mixed epistemic-aleatoric parameters in a stochastic compartmental model with feedback to emulate the dynamics of human interventions. Subsequently, we apply the MPC technique to determine the optimal NPI control policy while considering uncertainty propagation using probability bounds analysis (PBA) for uncertainty quantification. It is noteworthy that our proposed model possesses general applicability beyond the scope of COVID-19. Additionally, the COVID-19 case provides a valuable means of validating our model.

The following is the structure of this paper. Section 2 introduces the modified stochastic SIRD-MPC model. The data preprocessing and retrospective analysis are introduced in section 3. The simulation result is shown in section 4. The last section is the conclusion and discussion.

2 METHODOLOGY

We use a modified SIRD compartment model and the MPC approach to design optimal containment policies for the COVID-19 pandemic, considering both the epistemic and aleatoric parameter uncertainty.

2.1 Modified SIRD Compartment Model with Parameter Uncertainty

In the conventional SIRD model, the population is divided into four compartments at each time t: susceptible $S(t)$, infected $I(t)$, recovered $R(t)$, and deceased $D(t)$. The intrinsic dynamics of the infection process are fully elucidated by the evolution of $S(t)$, $I(t)$, $R(t)$, and $D(t)$, according to a system of nonlinear differential equations:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t)I(t) \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - \gamma I(t) - \alpha I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) \\ \frac{dD(t)}{dt} &= \alpha I(t) \end{aligned} \tag{1}$$

where β, γ, and α are the transmission rate, recovery rate, and death rate, respectively. Here, $S(t)$, $I(t)$, $R(t)$, and $D(t)$ represent the proportional population of compartment susceptible, infected, recovered, and deceased. Those parameters are assumed fixed and readily available in deterministic SIRD models. However, due to a lack of knowledge and insufficient data on infectious diseases, particularly at the early stage of the pandemic, the aleatory and epistemic uncertainty hinders most models from effectively depicting the COVID-19 evolution. Therefore, there is a dire need for the effective quantification of the parameter uncertainty and its propagation in the model. Though a huge amount of data have been collected so far for the COVID-19 virus and its variations, the inherent stochastic spatiotemporal dynamics of the infection and the population heterogeneity, among other factors, have presented a knowledge gap in the simulation and modeling of COVID-19. Here, parameters in the SIRD compartment model are characterized in a mixed epistemic-aleatoric scheme, where the uncertainty about the probability distribution of a model parameter is expressed in terms of a particular form of the distribution function to an interval bounded by lower and upper bounds on the distribution function parameters. For example, it is generally accepted in the literature that the transmission rate β follows a normal distribution, which is the internal randomness of disease transmission and cannot be eliminated as the aleatory uncertainty. Meanwhile, the transmission rate β exhibits different infectivity in different environments (temperature, altitude, geographical location) (Starke et al. 2021), which is the epistemic uncertainty with the knowledge gaps in understanding the contagion process of disease and the ambient environments. Therefore, we can use the interval to quantify the epistemic uncertainty to characterize the average infectivity under various ambient environments and use the normal distribution to represent the aleatory uncertainty. Specifically, $\beta \sim Normal([0.03,\ 0.90],\ 0.05)$ follows the normal distribution to represent the aleatory uncertainty (internal randomness of disease transmission) with mean $\mu = [0.03{,}0.90]$ and standard deviation $\sigma = 0.05$, where $[a, b]$ represents an interval to quantify the epistemic uncertainty (knowledge gaps in understanding the contagion process of disease) of with lower bound a and upper bound b.

According to the record data on the prevalence rate of COVID-19, policymakers have responded with interventions of varying magnitudes to contain the infection, from shelter-at-home orders to a large-scale lockdown. While effective in bending the infection curve, the lockdown has incurred a huge economic cost. The policymakers design the NPI policy C_t to diminish the interaction in the population and hence the effective transmission rate β according to the prevalence. Here, $C_t \in [0.0, 1.0]$ represents the nominal level of control at time stamp t, which indicates a nominal discount on β. As explained in our recent work (Wan et al. 2022a), the effective control policy U_t is subject to the influence of the population's perception of infection risk and compliance with the intervention policy. For simplicity, we have $U_t = C_t\left(1 - e^{(-\varphi\times I)}\right)$. In this study, the value of each compartment is normalized to represent the fraction of the total population. Therefore, the value of infected compartment I in this SIRD model represents the infection rate (the proportion of infected compartment); $\varphi > 0$ is the scaling factor to steer the strength of the feedback for infection rate I, which can capture the attitude of society and/or government encountering I. Large φ means that the public is prone to comply with the control policy and vice versa. Therefore, the $1 - e^{(-\varphi\times I)}$ represents the effective level of policy implementation, which can characterize overall compliance of control policy for the society and/or government in dealing with the pandemic. In such a manner, as the infection rate I rises, so too will the effective control policy U_t, which will approach the nominal level of control C_t for large scaling factor φ. The effective transmission rate will decrease to become $\beta(1 - U_t)$ if a control strategy is implemented. Therefore, we propose a modified SIRD compartment model by incorporating the feedback control policy U_t and the conventional SIRD compartment model to capture the interventional evolution of the COVID-19 pandemic. Subsequently, discretization of the modified SIRD model is adopted to characterize the compartment evolution, with time step $\Delta t = 1$ day:

$$
\begin{gathered}
S(t+1) = S(t) - (1 - U_t)\beta S(t)I(t) \\
I(t+1) = I(t) + (1 - U_t)\beta S(t)I(t) - \gamma I(t) - \alpha I(t) \\
R(t+1) = R(t) + \gamma I(t) \\
D(t+1) = D(t) + \alpha I(t) \\
U_t = C_t\left(1 - e^{(-\varphi\times I(t))}\right)
\end{gathered}
\tag{2}
$$

2.2 Incorporation of Stochastic Compartment Model and MPC

In order to construct control strategies to combat the spread of COVID-19, we incorporate optimal control method into the modified stochastic SIRD compartment model, considering the mixed epistemic-aleatoric uncertainty. Therefore, to execute an optimal control approach to reduce the infection of COVID-19, it is required to quantify the uncertainty of the crucial variables of the compartment model, which are $S(t), I(t), R(t)$, and $D(t)$ at time stamp t. PBA method (implementation details of uncertainty quantification are provided in the appendix section) for uncertainty propagation is used to assess the possible outputs/compartment states and the effect of uncertainty on decision-making for our proposed model. Meanwhile, cumulative uncertainty makes effective optimal control impractical due to the time propagation of uncertainty with model dynamics. To address this issue, MPC is applied because it involves continuously updating the best strategy to make up for performance losses predicted over lengthy time horizons (Lemaitre et al. 2022). We use a two-stage approach to implementing MPC in this investigation: (a) solving the optimization problem for a fixed predictive horizon N_p using the system's states $h(t_0) = (S(t_0), I(t_0), R(t_0), D(t_0))$, achieved from the collected epidemiological data (true data or estimated stochastic data), as the initial conditions at the start of the optimization problem and (b) putting into the first step of optimal design C_{t_0} for the compartment model (equation (2)) to achieve the next time stamp states $h(t_1)$ and starting the next new prediction horizon ($[t_1, t_{N_p+1}]$) until the final terminal time. Therefore, the

MPC procedures aim at optimizing the objective within the predictive horizon N_p. The general MPC problem with cost function $f(h(t), U_t)$ starting from time stamp $t = k$ can be represented as:

$$\boldsymbol{J}\big(h(k)\big) = \min_{\boldsymbol{C}(\boldsymbol{k})} \sum_{t=k+1}^{k+N_p} f(h(t), U_t) \tag{3}$$

s.t **constraints**

where $J(k)$ denotes the total estimated cost in MPC from time stamp k to $k + N_p$. $f(h(t), U_t)$ is the general form of the cost function, reliant on $h(t)$ and U_t at the time stamp t. In other words, the control cost depends on the contagion dynamics and the control magnitude, and different functional forms have been attempted in the literature (Wan et al. 2022b). We elaborate on the cost function in this study in the next section. $h(t) = h(t|k)$ is the predicted state at time stamp t given state $h(k)$ at time stamp k. $\boldsymbol{C}(\boldsymbol{k}) = \{C_k, \dots C_{k+N_p-1}\}$ is a vector of manipulated variables in a prediction horizon N_p days start from time stamp k, and C_{min} and C_{max} are the minimum and maximum values of the nominal level of control C_t. In this paper, the **constraints** are defined as follows:

$$\hat{S}(t+1) = \hat{S}(t) - (1 - U_t)\beta\hat{S}(t)\hat{I}(t)$$
$$\hat{I}(t+1) = \hat{I}(t) + (1 - U_t)\beta\hat{S}(t)\hat{I}(t) - \gamma\hat{I}(t) - \alpha\hat{I}(t)$$
$$\hat{R}(t+1) = \hat{R}(t) + \gamma\hat{I}(t)$$
$$\hat{D}(t+1) = \hat{D}(t) + \alpha\hat{I}(t)$$
$$U_t = C_t\big(1 - e^{(-\varphi\times\hat{I}(t))}\big)$$
$$0 \le C_t \le 1.0$$

where $\hat{S}, \hat{I}, \hat{R}$, and $\hat{D}$ represent the estimated results for S, I, R, and D, respectively.

3 MODEL EVALUATION

3.1 Data Preprocessing

In this study, we use the COVID-19 data from the open-access database "The COVID Tracking Project" (available at https://covidtracking.com/data/download) that compiles and makes available information on COVID-19 in the United States. The data is collected from January 2020 to March 2021. We get the cumulative information on "positive", "recovered", and "death". Based on the data definitions provided by "The COVID Tracking Project", which is available at https://covidtracking.com/about-data/data-definitions. The "positive" refers to the cumulative confirmed cases plus probable cases reported of COVID-19, the "death" to the cumulative deaths associated with a confirmed or probable case diagnosis of COVID-19, and the "recovered" to the cumulative number of people who have been confirmed and then recovered from COVID-19. Therefore, the daily infected cases should be the cumulative confirmed cases minus the cumulative recovered and cumulative deceased cases. The susceptible cases are the total population minus the cumulative confirmed cases. In this study, we consider all the data normalized for each compartment susceptible (S), infected (I), recovered (R), and deceased (D). To begin with, we adopt a 7-day moving average (MA) filter to smooth the time series record for the 4 compartments. In this study, the preprocessed data obtained through the use of the MA technique is treated as true data and is utilized in the analysis.

3.2 Retrospective Analysis

Analyzing past events, data, or results to better understand their causes and effects is called retrospective analysis. In this work, the purpose of the retrospective analysis is to understand the interplay between such intervention policies and infection dynamics. Here, we compare the nominal level of control C_t and the observed infection rate I to analyze whether the proposed feedback control intervention in our model can capture the human society interventions during the pandemic, while maintaining the minimum gap between the estimated result with the true data. In such a manner, we construct the cost function with the absolute value as a penalty term for the difference between the estimated states ($\hat{I}(t), \hat{R}(t)$, and $\widehat{D}(t)$) and the true data in order to evaluate whether our proposed model can capture the realistic intervention dynamics of COVID-19. We also consider the estimated results from the final day to match the trend of the trajectories. Due to the parameter uncertainty, we apply the mean of the estimated data to the objective function. Hence, the optimization problem can be formulated by:

$$\boldsymbol{J}\big(h(k)\big) = \sum_{t=k+1}^{k+N} \left(\min_{C(t-1)} \sum_{i=t}^{t+N_p} Z(i) + Z\big(t+N_p\big) \right) \quad (4)$$

s.t **constraints**

where N is the total estimation time to estimate each compartment for a given initial condition $h(k)$ and we solve an optimization to predict coming N_p days at each time stamp t. $Z(i) = g\big(I(i)\big) + g\big(R(i)\big) + g\big(D(i)\big)$ and $g(\cdot)$ is the absolute difference between the mean of the estimated value and the true value of compartments (I, R, and D). For example, $g(I(t)) = \left\| \mu\left(\hat{I}(t)\right) - I(t) \right\|_1$, where $\mu(\cdot)$ is the mean of the variable, calculated based on the PBA method, and $\hat{I}(t)$ and $I(t)$ are the estimated and true results for the infected compartment at the time stamp t. This optimization problem includes N small optimization problems and each small optimization problem can achieve a vector of optimized variables $\boldsymbol{C(t-1)} = \{C_{t-1}, \ldots, C_{t-1+N_p}\}$ by minimizing $\sum_{i=t}^{t+N_p} Z(i) + Z\big(t+N_p\big)$. Only the first control variable C_{t-1} and the first day of the state $h(t) = (S(\text{t}), I(\text{t}), R(\text{t}), D(\text{t}))$ are keeping for time stamp t. Based on equation (4), we can achieve the implemented nominal level of control set $\boldsymbol{C_{imp}} = \{C_k, .., C_{k+N}\}$, where C_i, $i = k, \ldots, N$, is the first element of vector $\boldsymbol{C(i)} = \{C_i, \ldots, C_{i+N_p}\}$. The achieved nominal level of control set $\boldsymbol{C_{imp}} = \{C_k, .., C_{k+N}\}$ can represent the policy design of the government from time stamp k to $k+N$.

Here, we assume that the policymaker designs the control actions consistently for each prediction horizon, which means that $\boldsymbol{C(i)} = \left\{C_i, \ldots, C_{i+N_p}\right\}$ with $C_a = C_i$, $a = i+1, \ldots, i+N_p$. Moreover, we set the estimation time $N = 7$ days for reinitializing because we assume that the policymakers readjust/ re-estimate the intervention policy weekly based on the latest collected information. We initialize the value of 4 compartments to the values of September 13, 2020, in Texas, and the total population is a constant value of 29,360,000. Thus, the normalized initial condition $S(0) = 0.97157$, $I(0) = 0.002641$, $R(0) = 0.019720$, and $D(0) = 0.000482$. Consequently, the initial condition of 4 compartments in the next period is set to the value of September 20, 2020. Additionally, we set the prediction horizon $N_p = 7$ and the scaling factor $\varphi = 1000$. The parameter uncertainty for transmission rate β, recovery rate γ, and death rate α are set as follows based on studies (Pacheco and de Lacerda 2021; Pei and Zhang 2021; Sebbagh and Kechida 2022):

$$\beta \sim Normal([0.03,\ 0.90],\ 0.05)$$
$$\gamma \sim Normal([0.001,\ 0.45], 0.01)$$
$$\alpha \sim Normal([0.0001,\ 0.07], 0.01)$$

We compare the estimated result of our model and the true data in 7 weeks ($7N$), from September 14, 2020, to November 1, 2020 (applied in all the figures in this work), as shown in Figure 1. The shaded area represents the possible evolution result of each compartment under the given parameter uncertainty. Because uncertainty propagation accumulates uncertainty from prior days, the shaded area gradually expands. As time progresses, the uncertainty will grow to be enormous. As a result, the estimation time cannot be made too long lest the prediction becomes meaningless. In this case, we suppose that policymakers reassess the intervention strategy every week in light of the latest collected information, causing uncertainty to build up again every $N = 7$ days. As a result, in the course of 7 days, there is an incremental increase in uncertainty, followed by a significant drop as a result of the implementation of the latest certain information. The result in Figure 1 shows that the predictive interval covers the observed numbers, which indicates that our proposed model can capture the realistic interventional evolution of the

COVID-19 pandemic. Furthermore, we compare the implemented nominal level of control C and the effective control U, as obtained from equation (4), with the practical infection rate I to assess the correlation between the two. This evaluation enables us to evaluate whether the proposed feedback control intervention in our model can characterize the human society interventions during the pandemic. Our underlying assumption is that an increase in the level of control leads to a reduction in infections, while a decrease in control results in a rise in infections under a relatively stable scenario. The intuitive result, shown in Figure 2, indicates that the C has a negative influence on the infection rate I, which conforms to the expectation of our proposed model. The correlation of C and I is -0.92.

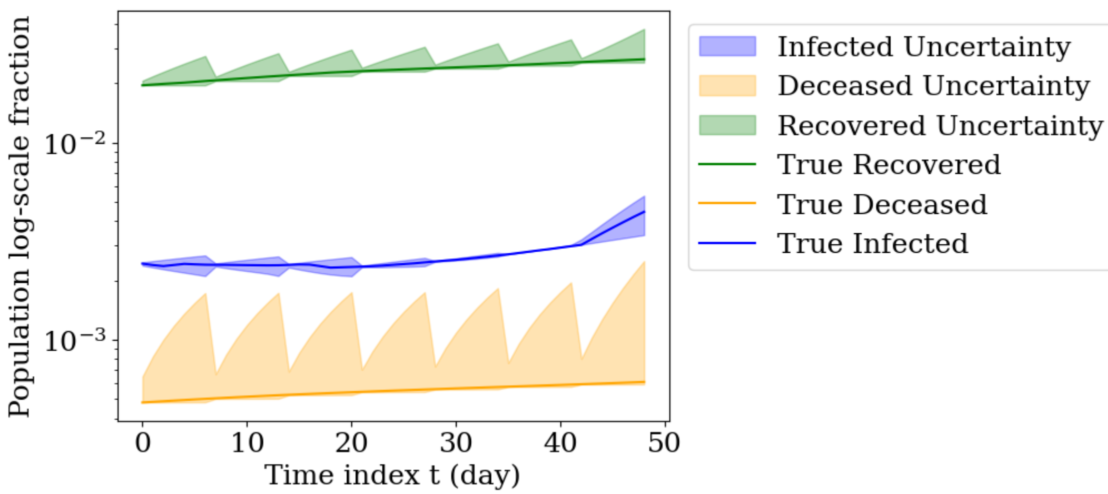

Figure 1: Comparing the estimated time series resulted from proposed model and true data starting from date September 14, 2020. The solid lines are the true data for infected, recovered, and deceased compartments presented with blue, green, and orange colors, respectively. The shaded areas are the estimated interval for infected, recovered, and deceased compartments presented with blue, green, and orange colors, respectively.

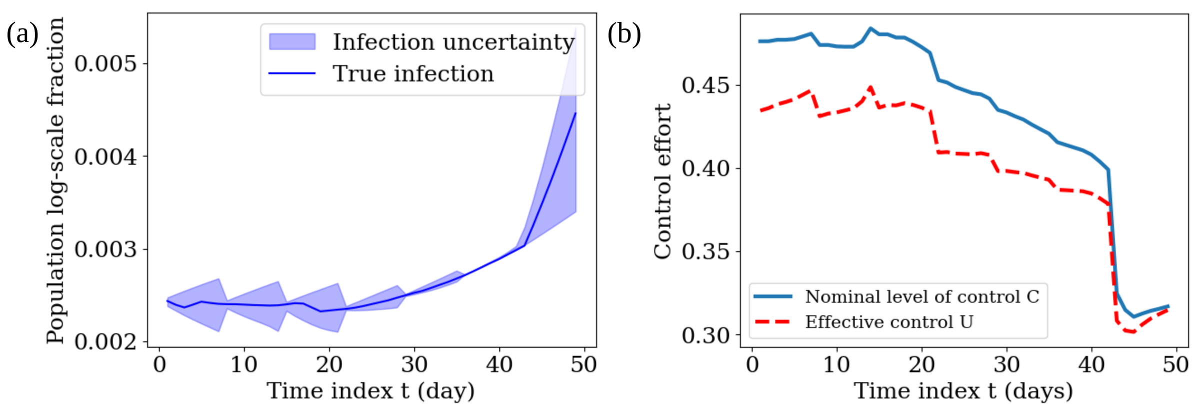

Figure 2: Comparing the achieved time series nominal level of control C_t and the practical infection rate $I(t)$ to check the trend of $I(t)$ and C_t starting from date September 14, 2020. The shaded area in (a) is the estimated interval of the infection rate and the solid line is the observed infection rate. The solid lines in (b) show the achieved nominal level of control C and effective control U.

4 SIMULATION RESULT

This work targets constructing a modified stochastic SIRD-MPC model to develop robust optimal control policies to provide guidance to combat the spread of the pandemic. In this section, we implement a control scheme into our proposed model to determine the robust optimal control policy so as to decrease the maximum infection rate cost-effectively. The control scheme can be characterized by the detail optimization problem described below:

$$\boldsymbol{J}\big(h(k)\big) = \min_{\boldsymbol{C(k)}} \sum_{t=k+1}^{k+N_p} w_1 \mathbb{U}\left(\hat{I}(t)\right) + w_2 \mathbb{L}(U_t) \tag{5}$$

s.t **constraints**

where $\mathbb{U}(\cdot)$ and $\mathbb{L}(\cdot)$ are the maximum and minimum values of the variable; w_1 and w_2 are weights for each term. As introduced in Section 2, the control action U_t is an effective control policy imposed on society based on the nominal level of control C_t designed by policymakers and the effective level of policy implementation characterized by $1 - e^{(-\varphi \times I)}$. Therefore, U_t is also an uncertainty due to the existence of $\hat{I}$ in U_t. Here, the cost function considers the worst scenario of infected compartment $\mathbb{U}\left(\hat{I}(t)\right)$ and worst effective control policy $\mathbb{L}(U_t)$ to design the nominal level of control policy C_t so as to avoid significant infection and economic costs. Compared to the cost function in equation (4) minimizing the gap between the estimated result with true data, we desire to decrease the infection rate to achieve a better control result than true data.

In this case, we set the w_1 =1000 and $w_2 = 1$ to minimize the maximum estimated infection rate $\hat{I}$ even with a large economic cost. Meanwhile, to curb the uncertainty cumulation, we reset the MPC every 7 days, similar in section 3.2, with new initial condition $h(k) = \mu\left(\hat{h}(k)\right)$, where $\hat{h}(k)$ is the estimated system state based on our proposed model. The simulation result is shown in Figure 3 where (a) shows the true and estimated infection rate and (b) presents the achieved nominal level of control and the effective control based on our proposed model and control scheme. The result indicates that our proposed model can effectively suppress the infection rate even in the worst-case scenario where uncertainty in model parameters leads to the infection rate at each time stamp t. It is worth noting that due to the low infection rate, the effective control is significantly less than the nominal level of control. Therefore, policymakers should consider designing stricter control policies to tame the pandemic.

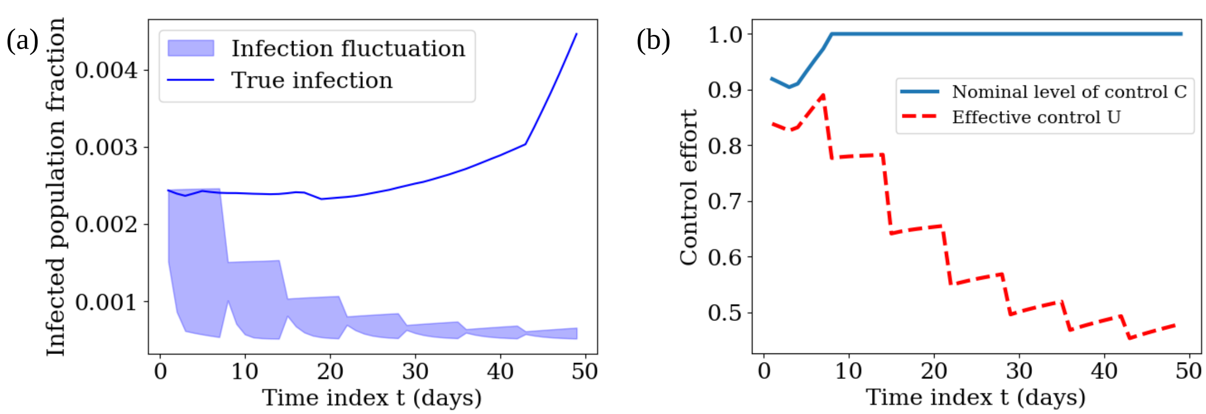

Figure 3: Comparing the estimated infection rate and true infection rate starting from date September 14, 2020: (a) comparison of the true and estimated infection rate; and (b) the achieved ideal level of control C and the effective control U based on the proposed model.

5 CONCLUSION AND DISCUSSION

In this present study, we propose a modified stochastic SIRD compartment model with mixed epistemic-aleatoric parameter uncertainty to characterize the COVID-19 dynamics and incorporate it with the optimal control technique MPC to develop the optimal NPI control policy to contain virus transmission with minimum economic impact. The retrospective analysis indicates that our proposed modified SIRD-MPC model can capture the realistic interventional COVID-19 dynamics with the proposed feedback NPI control policy implementation. Moreover, the simulation result in section 4 suggests that our proposed model can achieve a better containment of the infection rate compared to the practical control result based on the control scheme. In a realistic control design, the effective control may be significantly smaller than the nominal control, so the control scheme should be more stringent.

Possible future research will investigate how to design a prediction horizon N_p, such as inconstant N_p, in our proposed model. Moreover, although our proposed model achieves good performance by model validation, we will consider more factors for the feedback mechanism to improve our model. Likewise, designing dynamic scaling factor φ and exploring various control schemes are also worth investigating in order to better understand and apply our proposed model to combat the pandemic infection in a realistic situation.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (Award Number 2119334) of the United States and Interdisciplinary Collaborations Grant of Binghamton University.

APPENDIX

This paper employs the PBA method to quantify the mixed epistemic-aleatoric uncertainty by constructing probability boxes (p-boxes) that capture the propagated uncertainty. The p-box serves as a versatile framework that combines intervals and probability distributions into a unified structure, allowing for the characterization of interval bounds on probability distributions. This framework facilitates the rigorous treatment of both epistemic and aleatory uncertainty during computational processes.

Generally, a p-box, denoted as $P_b(x)$ can be represented by a left bound and right bound function as

$$P_b(x) = [\underline{P_b}(x), \overline{P_b}(x)], \overline{P_b}(x) \geq \underline{P_b}(x)$$

where $\underline{P_b}(x)$ and $\overline{P_b}(x)$ represent the left and right bound distribution functions, respectively. In the PBA method, these functions are approximated based on the inverse of the cumulative distribution function of the left and right bounds. The process involves selecting N_s points between 0 and 1 to construct a cumulative probability dataset $\boldsymbol{y} = (y_1, \dots y_{N_s})$ to record each cumulative probability y_i, where $i = 1, \dots, N_s$. Then, the minimum and maximum values $(\underline{x_i}, \overline{x_i})$ of variable x for each cumulative probability y_i are calculated using the inverse of the cumulative distribution functions. Finally, the left and right bound distribution functions $\underline{P_b}(x)$ and $\overline{P_b}(x)$ are formulated based on the cumulative probability dataset $\boldsymbol{y}$ and its coordinated left and right bound datasets $\underline{\boldsymbol{x}} = (\underline{x_1}, \dots \underline{x_{N_s}})$ and $\overline{\boldsymbol{x}} = (\overline{x_1}, \dots, \overline{x_{N_s}})$, achieved from the inverse of the cumulative distribution function. Therefore, the whole process allows for the determination of interval bounds on probability distributions within the p-box framework.

According to the description above, the most important step is to calculate the minimum and maximum values $(\underline{x_i}, \overline{x_i})$ for each cumulative probability y_i. To illustrate this process, we consider a general case where $x \sim Normal([\mathrm{a}, \mathrm{b}], [\mathrm{c}, \mathrm{d}])$, $a \leq b$, and $c \leq d$. We can reformulate the variable x into four distributions: $x1 \sim Normal(\mathrm{a}, \mathrm{c})$, $x2 \sim Normal(\mathrm{a}, \mathrm{d})$, $x3 \sim Normal(\mathrm{b}, \mathrm{c})$, and $x4 \sim Normal(\mathrm{b}, \mathrm{d})$. Denoting the cumulative distribution functions of $x1$, $x2$, $x3$, and $x4$ are $y1 = P_r1(x)$, $y2 = P_r2(x)$, $y3 = P_r3(x)$, and $y4 = P_r4(x)$ for any given value x, respectively. The inverse of the cumulative distribution functions of $x1$, $x2$, $x3$, and $x4$ are denoted as $x1 = P_r1^{-1}(y)$, $x2 = P_r2^{-1}(y)$, $x3 = P_r3^{-1}(y)$, and $x4 = P_r4^{-1}(y)$ for any given cumulative probability y. Hence, we can obtain the left bound and right bound points for a given cumulative probability y_i as follows:

$$\underline{x_i} = \min(P_r1^{-1}(y_i), P_r2^{-1}(y_i), P_r3^{-1}(y_i), P_r4^{-1}(y_i))$$
$$\overline{x_i} = \max(P_r1^{-1}(y_i), P_r2^{-1}(y_i), P_r3^{-1}(y_i), P_r4^{-1}(y_i))$$

To quantify the uncertainty of variables, it is necessary to select a confidence interval $I_\alpha = [I_{\alpha1}, I_{\alpha2}]$, where $0 \leq I_{\alpha1} \leq I_{\alpha2} \leq 1$. The precision of the left bound and right bound functions approximation relies on the total number of selected points, denoted as N_s. In this study, we choose N_s =200 evenly space values between $I_{\alpha1} = 0.0001$ and $I_{\alpha2} = 0.9999$. Consequently, the cumulative probability dataset $\boldsymbol{y} = (0.0001, \dots, 0.9999)$.

As illustrated above, the uncertainty quantification based on the PBA method should be conducted within the framework of p-box $[\underline{P_b}(x), \overline{P_b}(x)]$. In fact, probability distributions, intervals, and constants can all be considered as special cases of p-boxes, as explained below:

1. The probability distribution of a variable x^p with known cumulative distribution functions $y = f^p(x^p)$, the left and right bound datasets for its p-box can be written as $\underline{\boldsymbol{x^p}} = \overline{\boldsymbol{x}}^{\boldsymbol{p}} = (f^{p-1}(0.0001), \dots, f^{p-1}(0.9999))$ and its p-box format is $P_b(x^p) = \left[\underline{P_b}(x^p), \overline{P_b}(x^p)\right]$, where $\underline{P_b}(x^p) = \overline{P_b}(x^p)$ because as $\underline{\boldsymbol{x^p}} = \overline{\boldsymbol{x}}^{\boldsymbol{p}}$ in our study.
2. The interval variable x^I with interval $[a, b]$ can achieve the left and right bound datasets for its p-box $\underline{\boldsymbol{x}}^{\boldsymbol{I}} = (a, \dots a)$ and $\overline{\boldsymbol{x}}^{\boldsymbol{I}} = (b, \dots, b)$, respectively.
3. The constant ς can achieve the left and right bound datasets for its p-box $\underline{\boldsymbol{x}}^\varsigma = \overline{\boldsymbol{x}}^\varsigma = (\varsigma, \dots, \varsigma)$.

These representations allow for the flexible application of the p-box framework to various types of uncertainty, encompassing probability distributions, intervals, and constants. The standard arithmetic operations can be performed on p-boxes. Any two p-boxes $P_b1(x1) = [\underline{P_b1}(x1), \overline{P_b1}(x1)]$ and $P_b2(x2) = [\underline{P_b2}(x2), \overline{P_b2}(x2)]$ can conduct standard arithmetic operations as follows: $P_b3(x3) = P_b1(x1) \circ P_b2(x2) = [\underline{P_b3}(x3), \overline{P_b3}(x3)]$, where $\circ \in (+, -, *, /)$.

As we mentioned above, the left and right bound functions $\underline{P_b3}(x3)$ and $\overline{P_b3}(3x)$ should be formulated based on the cumulative probability dataset $\boldsymbol{y}$ and its coordinated achieved left and right bound datasets $\underline{\boldsymbol{x3}} = (\underline{x_13}, \dots \underline{x_{N_s}3})$ and $\overline{\boldsymbol{x3}} = (\overline{x_13}, \dots, \overline{x_{N_s}3})$. In this study, we assume that there is a perfect positive association between the two variables. Therefore, the left and right bound datasets $\underline{\boldsymbol{x3}}$ and $\overline{\boldsymbol{x3}}$ are calculated as follows:

$$\underline{\boldsymbol{x3}} = \underline{\boldsymbol{x1}} \circ \underline{\boldsymbol{x2}} = (\underline{x_11} \circ \underline{x_12}, \dots \underline{x_{N_s}1} \circ \underline{x_{N_s}2})$$
$$\overline{\boldsymbol{x3}} = \overline{\boldsymbol{x1}} \circ \overline{\boldsymbol{x2}} = (\overline{x_11} \circ \overline{x_12}, \dots, \overline{x_{N_s}1} \circ \overline{x_{N_s}2})$$

where, $\underline{\boldsymbol{x1}} = (\underline{x_11}, \dots \underline{x_{N_s}1})$ and $\overline{\boldsymbol{x1}} = (\overline{x_11}, \dots, \overline{x_{N_s}1})$ are the achieved left and right bound datasets for p-box $P_b1(x1)$, based on selected cumulative probability dataset $\boldsymbol{y} = (y_1, \dots y_{N_s})$, respectively. Similarly, $\underline{\boldsymbol{x2}} = (\underline{x_12}, \dots \underline{x_{N_s}2})$ and $\overline{\boldsymbol{x2}} = (\overline{x_12}, \dots, \overline{x_{N_s}2})$ are the achieved left and right bound datasets for p-box $P_b2(x2)$, based on selected cumulative probability dataset $\boldsymbol{y} = (y_1, \dots y_{N_s})$, respectively.

In the PBA framework, the arithmetic operations of p-boxes provide a means to calculate the p-box for a dependent variable by leveraging the established arithmetic relationships of other given uncertainty parameters and/or variables. Through the application of these arithmetic operations, it becomes possible to determine the p-box associated with the dependent variable and consequently quantify its corresponding uncertainty. The handling of cumulative uncertainty involves an iterative process that continuously updates the relevant uncertainty (p-box) of parameters and/or variables, allowing for a comprehensive assessment of the uncertainty pertaining to the dependent variable.

REFERENCES

Armaou, A., B. Katch, L. Russo, and C. Siettos. 2022. "Designing Social Distancing Policies for the COVID-19 Pandemic: A Probabilistic Model Predictive Control Approach". *Mathematical Biosciences and Engineering* 19(9):8804–8832.

Calafiore, G. C., C. Novara, and C. Possieri. 2020. "A Time-Varying SIRD Model for the COVID-19 Contagion in Italy". *Annual Reviews in Control* 50:361–372.

Chen, C. Y., J. P. Ward, and W. B. Xie. 2019. "Modelling the Outbreak of Infectious Disease Following Mutation from a Non-Transmissible Strain". *Theoretical Population Biology* 126:1–18.

Chowell, G., P. W. Fenimore, M. A. Castillo-Garsow, and C. Castillo-Chavez. 2003. "SARS Outbreaks in Ontario, Hong Kong and Singapore: the Role of Diagnosis and Isolation as A Control Mechanism". *Journal of Theoretical Biology* 224(1):1–8.

Gallo, L., M. Frasca, V. Latora, and G. Russo. 2022. "Lack of Practical Identifiability May Hamper Reliable Predictions in COVID-19 Epidemic Models". *Science Advances* 8(3):eabg5234.

Gray, N., S. Ferson, M. De Angelis, A. Gray, and F. B. de Oliveira. 2022. "Probability Bounds Analysis for Python". *Software Impacts* 12:100246.

Jin, Z., J. Zhang, L. Song, G. Sun, J. Kan, and H. Zhu. 2011. "Modelling and Analysis of Influenza A (H1N1) on Networks". *BMC Public Health* 11(1):1–9.

Lemaitre, J. C., D. Pasetto, M. Zanon, E. Bertuzzo, L. Mari, S. Miccoli, R. Casagrandi, M. Gatto, and A. Rinaldo. 2022. "Optimal Control of the Spatial Allocation of COVID-19 Vaccines: Italy as A Case Study". *PLoS Computational Biology* 18(7):e1010237.

Mamis, K., and M. Farazmand. 2023. "Stochastic Compartmental Models of the COVID-19 Pandemic Must Have Temporally Correlated Uncertainties". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 479(2269):20220568.

March, D., J. Bond, and G. Buzi. 2022. "A Feedback SAIR Model for the Spread of Infectious Disease with Application to COVID-19 Pandemic". In *Proceedings of the 2022 American Control Conference (ACC 2022)*, edited by B. Ferri, 1330–1335. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Miller, I. F., A. D. Becker, B. T. Grenfell, and C. J. E. Metcalf. 2020. "Disease and Healthcare Burden of COVID-19 in the United States". *Nature Medicine* 26(8):1212–1217.

Pacheco, C. C., and C. R. de Lacerda. 2021. "Function Estimation and Regularization in the SIRD Model Applied to the COVID-19 Pandemics". *Inverse Problems in Science and Engineering* 29(11):1613–1628.

Pei, L., and M. Zhang. 2021. "Long-Term Predictions of COVID-19 in Some Countries by the SIRD Model". *Complexity* 2021: 1–18.

Péni, T., and G. Szederkényi. 2021. "Convex Output Feedback Model Predictive Control for Mitigation of COVID-19 Pandemic". *Annual Reviews in Control* 52:543–553.

Polcz, P., B. Csutak, and G. Szederkényi. 2022. "Reconstruction of Epidemiological Data in Hungary Using Stochastic Model Predictive Control". *Applied Sciences* 12(3):1113.

Starke, K. R., R. Mauer, E. Karskens, A. Pretzsch, D. Reissig, A. Nienhaus, A. L. Seidler, and A. Seidler. 2021. "The Effect of Ambient Environmental Conditions on COVID-19 Mortality: A Systematic Review". *International Journal of Environmental Research and Public Health* 18(12):6665.

Scarabaggio, P., R. Carli, G. Cavone, N. Epicoco, and M. Dotoli. 2022. "Nonpharmaceutical Stochastic Optimal Control Strategies to Mitigate the COVID-19 Spread". *IEEE Transactions on Automation Science and Engineering* 19(2):560–575.

Sebbagh, A., and S. Kechida. 2022. "EKF-SIRD Model Algorithm for Predicting the Coronavirus (COVID-19) Spreading Dynamics". *Scientific Reports* 12(1):13415.

Sharomi, O., and T. Malik. 2017. "Optimal Control in Epidemiology". *Annals of Operations Research* 251(1):55–71.

She, B., S. Sundaram, and P. E. Paré. 2022. "A Learning-Based Model Predictive Control Framework for Real-Time SIR Epidemic Mitigation". In *Proceedings of the 2022 American Control Conference (ACC 2022)*, edited by B. Ferri, 2565–2570. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wan, J, G. Ichinose, M. Small, H. Sayama, Y. Moreno, and C. Cheng. 2022. "Multilayer Networks with Higher-Order Interaction Reveal the Impact of Collective Behavior on Epidemic Dynamics". *Chaos, Solitons & Fractals* 164:112735.

Wan, J., Y. Che, Z. Wang, and C. Cheng. 2022. "Uncertainty Quantification and Optimal Robust Design for Machining Operations". *Journal of Computing and Information Science in Engineering* 23(1): 011005.

Wu, S., and H. S. Mortveit. 2015. "A General Framework for Experimental Design, Uncertainty Quantification and Sensitivity Analysis of Computer Simulation Models". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, V. W. K. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 1139–1150. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Xia, Z., S. Wang, S. Li, L. Huang, W. Zhang, G. Sun, Z. Gai, and Z. Jin. 2015. "Modeling the Transmission Dynamics of Ebola Virus Disease in Liberia". *Scientific Reports* 5(1):13857.

AUTHOR BIOGRAPHIES

JINMING WAN is a Ph.D. student in the Department of Systems Science and Industrial Engineering at the State University of New York at Binghamton. His research focuses on modeling and analysis, uncertainty quantification, machine learning, and optimal control techniques. Specifically, he is investigating coordinated policy design to enhance preparedness for pandemics, with the goal of providing compelling guidelines to help mitigate the impact of future outbreaks. His email address is jwan8@binghamton.edu.

SAEIDEH MIRGHORBANI is an assistant professor of Business Analytics and Operations and a faculty fellow in atrocity prevention at Binghamton University. Her research primary interest is the applications of management science and operations management in healthcare systems. In addition to the healthcare system, she is also interested in applications of operations management and business analytics in the fight for social justice, and racial and gender disparity in food insecurity and other social issues. Her email address is smirghor@binghamton.edu.

N. EVA WU is a Professor in the Department of Electrical and Computer Engineering at State University of New York at Binghamton. Her research interests include fault tolerant control, and reliability of large scale dynamic systems. Her email address is evawu@binghamton.edu.

CHANGQING CHENG is an associate professor in the Department of Systems Science and Industrial Engineering at Binghamton University. His research interests include nonlinear dynamics, sensing and data-driven modeling, simulation and analytics for process monitoring, quality control, and performance optimization of complex systems. He serves as associate editor for the journal of IISE Transactions on Healthcare Systems Engineering. His email address is ccheng@binghamton.edu.