Article

Hybrid Model-Based Simulation Analysis on the Effects of Social Distancing Policy of the COVID-19 Epidemic

Bong Gu Kang [1], Hee-Mun Park [2], Mi Jang [2] and Kyung-Min Seo [3,*]

[1] Research Institute of Industrial Technology Convergence, Korea Institute of Industrial Technology (KITECH), Ansan 15588, Korea; bgkang@kitech.re.kr
[2] Department of Computer Engineering, Korea University of Technology and Education (KOREATECH), Cheonan 31253, Korea; phm1321@koreatech.ac.kr (H.-M.P.); dntdmami@koreatech.ac.kr (M.J.)
[3] Department of Future Technology, Korea University of Technology and Education (KOREATECH), Cheonan 31253, Korea
* Correspondence: kmseo@koreatech.ac.kr; Tel.: +82-10-3376-9109

Abstract: This study utilizes modeling and simulation to analyze coronavirus (COVID-19) infection trends depending on government policies. Two modeling requirements are considered for infection simulation: (1) the implementation of social distancing policies and (2) the representation of population movements. To this end, we propose an extended infection model to combine analytical models with discrete event-based simulation models in a hybrid form. Simulation parameters for social distancing policies are identified and embedded in the analytical models. Administrative districts are modeled as a fundamental simulation agent, which facilitates representing the population movements between the cities. The proposed infection model utilizes real-world data regarding suspected, infected, recovered, and deceased people in South Korea. As an application, we simulate the COVID-19 epidemic in South Korea. We use real-world data for 160 days, containing meaningful days that begin the distancing policy and adjust the distancing policy to the next stage. We expect that the proposed work plays a principal role in analyzing how social distancing effectively affects virus prevention and provides a simulation environment for the biochemical field.

Keywords: simulation; SIRD model; discrete-event model; data-based learning; COVID-19 epidemic

1. Introduction

During the initial outbreak of coronavirus disease (COVID-19), the major propagation factors were population movement and cluster infections in groups [1–3]. In South Korea, for example, several situations exploded due to carriers in groups such as churches and hospitals [4]. In Singapore, a serological survey of citizens who traveled abroad confirmed an early spreader [5]. As one of the solutions for COVID-19 prevention, governments worldwide implemented social distancing to prevent the spread of COVID-19. The United States asserted the spreading speed in states conducting social distancing was six times slower than those that did not [6].

To analyze how social distancing is effective in COVID-19 prevention, we utilize modeling and simulation (M&S) methods in this study. Although simulations containing population movement are intrinsically complicated, well-categorized models are understandable, analyzable, and certifiable [7–10]. Therefore, they give helpful insights to analyze and predict COVID-19 situations [11–13]. Based on real-world data regarding confirmed cases in South Korea, the modeling requirements for the COVID-19 simulation are twofold: (1) the movement of the population and (2) the implementation of social distancing policies.

Various M&S methods for the COVID-19 epidemic have been conducted during the past two years. They overcome the limitations of analytical and simulation models. Some researchers have proposed analytical methods with several infection types [14–16], and others have developed simulation models over the partial region [17,18]. The collective

contribution of these studies is to suggest meaningful modeling methods and simulation results for the virus infection. Despite the contribution, they still warrant some improvements. For example, some analytical models are limited in reflecting events, such as policies and population movements; others have focused on the microscopic or macroscopic region, which results in difficulty in conducting the simulation-based analysis at the national level.

This study presents a practical M&S approach to analyze the trends of the COVID-19 epidemic. The proposed model consists of two steps to overcome the limitations of analytical and simulation models: (1) hybrid infection model design and (2) model identification. The proposed infection model utilized real-world data regarding suspected, infected, recovered, and deceased people in South Korea. Based on the data, we design an analytical model known as the SIRD (Susceptible, Infected, Recovered, and Dead) model [16]. Then, we complete an overall simulation model to combine a discrete event model with the SIRD model in a hybrid form. Model coefficients within the overall model contain domestic policies, such as social distancing and regional population movements. We optimized the developed model by updating the coefficients by stages to make the model reflect real-world data.

From a simulation modeling view, we considered an administrative district in South Korea as a fundamental agent. In other words, simulation agents were set in units of administrative districts to express virus transmission within and between the regions. Two main reasons for this simulation modeling are as follows. First, if we considered one person as an agent, as the number of people to be analyzed expanded, simulation agents would also increase, and it may take longer to acquire simulation results. Next, district-level agent modeling helped reflect population movement between the districts.

As a case study, we simulated the COVID-19 epidemic in South Korea to analyze the real-world infection trend and the changes in government policy. A social distancing policy in South Korea is classified into four levels according to the number of confirmed cases per 100,000 people. For infection prevention, the policy compels two primary factors: movement and gathering. For example, in the third level of social distancing, up to four people are permitted for a private gathering. Wearing a mask is not included in the policy; however, South Korea strongly encourages people to wear a mask. In this simulation, we assume that people correctly put on a mask. We used real-world data for 160 days for simulation experiments, containing two meaningful days: (1) a day to begin the distancing policy and (2) a day to adjust the distancing policy to the next stage. Seventeen administrative districts were modeled, and real-world data were used based on the population movement by region announced by the National Statistical Office of South Korea. The simulation results provide insights into pulling and delaying the enforcement of the distance policy.

This study is organized as follows. Section 2 explains related works for virus infection simulation. Section 3 proposes the overall architecture with a hybrid modeling approach, and Section 4 explains and discusses simulation experiments. Finally, Section 5 presents our conclusions.

2. Related Works

For the last couple of years, various studies have been conducted to provide simulation analyses for virus infections, such as the COVID-19 pandemic. In this section, we categorize them into modeling aspects, which are summarized in Table 1.

During virus outbreaks, the population can be classified mainly into five types: (1) susceptible people before infection (S), (2) people exposed to the virus (E), (3) infected people (I), (4) people who have recovered (R), and (5) people who have died (D). These types are used to design analytical models for viral infection analyses

Among the analytical modeling approaches, two primary models have been studied: the SIR model, with S, I, and R factors [14,19–21], and the SEIR model, which adds E factor from the SIR model [15,22–24]. For example, Igor utilized the SIR model to analyze the COVID-19 infection and recovery rates [14]. Because this is an early study that occurred

Int. J. Environ. Res. Public Health **2021**, *18*, 11264

3 of 17

in South Korea, the analysis period is relatively short. Mulder [15] used the SEIR model to analyze the infection exposure, infection, and recovery rates in 36 countries. Mulder's study also cannot analyze the mortality rate. Moreover, the number of exposed people is an inappropriate model for our study based on real-world data because precise real-world data cannot be obtained. The mortality rate also could not be considered.

To improve the SIR model, the SIRD models have been studied, which contain the D factor (i.e., the number of dead people) [16,25–27]. For example, Calafiore et al. [16] predicted the number of S, I, R, and D factors. There is also a modified SIRD model that is more advanced than SIRD. In the case of the previous Ebola virus epidemic, there have been many cases of reinfection after recovery. Therefore, the analysis with the modified SIRD model was suitable for the Ebola virus study [28]. Similar to Calafiore et al. [16], Osemwinyen et al. [28], Basti et al. [29], Shringi et al. [30], and Devosmita et al. [31] were unable to reflect on policy and population movement. Furthermore, in the case of COVID-19, reinfection cases are sporadic in Korea; thus, a modified SIRD model was considered inappropriate.

Table 1. Related works for modeling and simulation studies for virus infection.

Related Work	Modeling Approach	Modeling Method	Description
[14]	Analytical model	Equation-based model (SIR)	- COVID-19 application - Derive the infection rate and recovery rate in Korea - Short analysis period (13 days) - Analysis of mortality is impossible.
[15]	Analytical model	Equation-based mode (SEIR)	- COVID-19 application - Analysis of 36 countries in which COVID-19 has been widely transmitted, including Korea - Derive the probability of exposure to infection, infection rate, and recovery rate - No real-world data can be obtained for the number of people exposed to infection.
[16]	Analytical model	Equation-based model (SIRD)	- COVID-19 application - Derive the infection rate, recovery rate, and mortality rate in Italy - Trend prediction is possible, but special events, such as policy and outbreak of group infection, cannot be reflected.
[28]	Analytical model	Equation-based model (Modified SIR)	- Ebola application - Mathematical analysis of the spread of Ebola in developing countries - Using a modified SIRD model by including reinfection in a typical SIRD model - Model unsuitable for COVID-19 with few cases of reinfection
[17]	Simulation model	Agent-based model	- COVID-19 application - Predict the effectiveness of social distancing policies in the United States
[18]	Simulation model	Agent-based model	- COVID-19 application - Predict infection among people living together in closed spaces (e.g., offices and buildings)

Equation-based models are limited in reflecting events, such as policies and population movements. Simulation models have been used to overcome the limitations of equation-based models. For example, Alagoz et al. [17] predicted the cumulative number of confirmed cases according to the date when the social distancing policy was implemented in three regions of the United States. They predicted how the cumulative number

of infected people would change when the onset of social distancing was accelerated or delayed by a week. They reflected the policy in the simulation, but not the population movement between regions.

D'Orazio et al. [18] also predicted the trend of virus transmission between humans in a dense space. Although it is possible to analyze the transmission of viruses between humans, there is a limit to targeting a confined space without the population movement. However, when a person is set as an agent, as in D'Orazio et al., it is difficult to expand the analysis scope to the country. When analyzing a confined space without external inflow, it is impossible to express the propagation characteristics of COVID-19 spreading between regions.

To overcome the limitations of analytical and simulation models, we propose a hybrid modeling approach (i.e., simulation models containing the analytical models). The proposed model reflects the social distancing policy, as well as the population movements between regions. We add simulation parameters in the analytical SIRD model to reflect the social distancing policy. We divide administrative districts into a simulation agent, which facilitates representing the population movements between the cities.

3. Proposed Work

3.1. Overall Process Description

For the simulation-based what-if analysis (the third step in Figure 1)—the ultimate goal of this study—it is necessary to construct the simulation model by reflecting the real-world environment, which consists of two steps: (1) hybrid infection model design (first step) and (2) model identification (the second step).
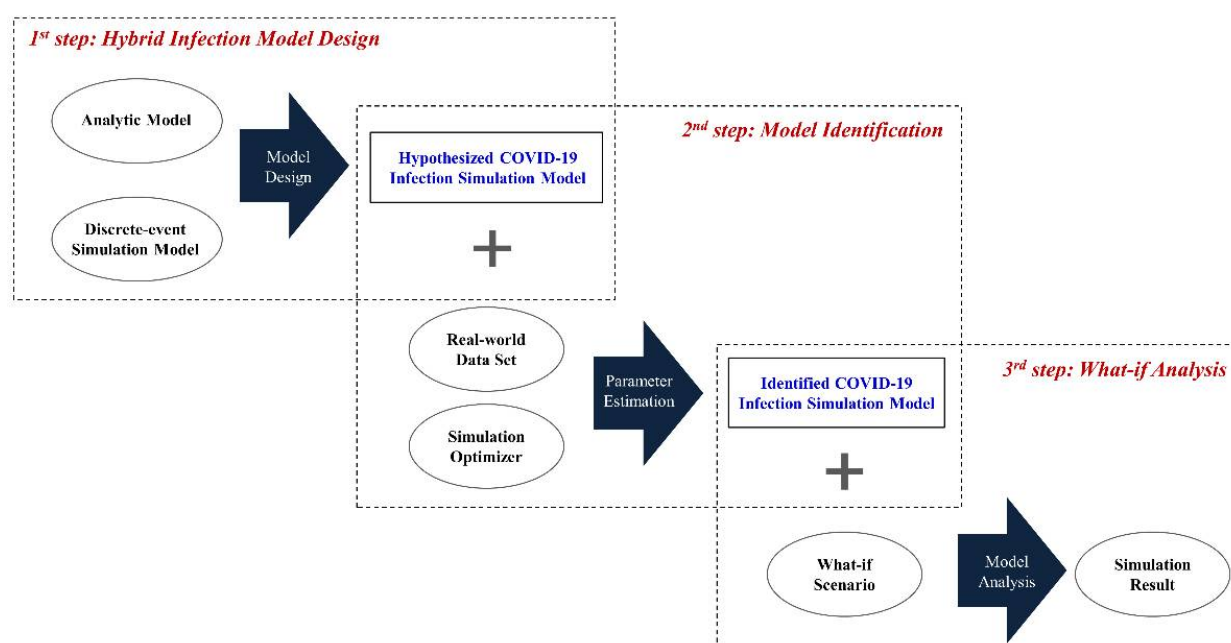

Figure 1. Overall process of proposed simulation model.

The existing simulation models for describing the infection spread process can be categorized into two types: (1) analytical model and (2) discrete-event simulation model [32]. The former focuses on the transmission of infection within a large-scale community using a differential equation-based model (EBM), which can analyze the tendency of infection at a macro level but has the disadvantage of not considering the relationship with other communities. On the other hand, the latter can describe the interaction with other communities using an agent-based model (ABM), but it is limited to a large-scale community.

In the model design step, we constructed a hybrid model to simulate infection, including interactions on a large scale, by combining two model types: the SIRD model (i.e.,

the analytical model) and the infection discrete-event simulation model expressed in the DEVS (discrete event systems specification) formalism and implemented in the DEVSim++ environment [33,34]. The former describes the infection behavior in the inner community, and the latter depicts the infection interaction among communities.

Although we constructed the simulation model at the previous stage, it is impossible to identify all the parameter values in the model. For example, in Table 2, we can know the duration time of the relaxed social distancing policy (t_{rsd}) in the policy parameter set based on the real-world data, but it is impossible to know the exact rate of infection (β) in the infection parameter set in the SIRD model. In the model identification step, to solve this problem, we identified the constructed infection simulation model using a data set acquired from the real-world and simulation optimizer [35], which calibrates the parameters in the simulation model using the optimization algorithms [36–38].

Table 2. Input parameter description of simulation model.

Input Parameter	Notation	Description
Infection parameter set	β	The rate of infection
	γ	The rate of recovery
	μ	The rate of death
	λ	The attenuation coefficient of increase in the rate of infection
	κ	The attenuation coefficient of decrease in the rate of infection
Policy parameter set	t_{rsd}	The duration time of the relaxed social distancing policy
	t_{ssd}	The duration time of the severe social distancing policy

In the final step, we can conduct the what-if analysis using the identified infection simulation model. For example, we can analyze trends of the $S(t)$, $I(t)$, $R(t)$, and $D(t)$ against the policy parameter set of Table 3. This study deals with the first step in the following subsection and the other steps in Section 4.

Table 3. Output parameter description of the simulation model.

Output Parameter	Notation	Description
Variables for SIRD model	$S_m(t)/S_c(t)$	The number of susceptible population in the major city/country at time t
	$I_m(t)/I_c(t)$	The number of the infected population in the major city/country at time t
	$R_m(t)/R_c(t)$	The number of recovered population in the major city/country at time t
	$D_m(t)/D_c(t)$	The number of deaths population in the major city/country at time t

3.2. Hybrid Infection Simulation Model Design

Figure 2 shows the overall architecture of the proposed hybrid infection simulation model (*ISM*). The *ISM* is a discrete-event simulation model, which includes the analytical model as an algorithm. The entire model consists of two sub-models: (1) major city model (*MCM)* and (2) transfer model (*TM*). The *MCM* and *TM* describe the metropolis city and their interaction exchanges, respectively. In the real world, because the infection spread between agents has a discrete-event property and the cities have a modular property, we expressed the entire model using the DEVS formalism. The DEVS expression of *ISM*, *MCM*, and *TM* are represented by Equations (1)–(25). The TM has a role in routing the

infection spread from one MCM to the other MCM, and Equation (7) means their sending and receiving connection. In addition, to prioritize this role, we gave priority to the model as Equation (8).

$$ISM = <X, Y, M, EIC, EOC, IC, Sel> \tag{1}$$

$$X = \varnothing; \tag{2}$$

$$Y = \varnothing; \tag{3}$$

$$M = \{TM\} \cup \cup_{i=1}^{n} \{MCM_i\}; \tag{4}$$

$$EIC = \varnothing; \tag{5}$$

$$EOC = \varnothing; \tag{6}$$

$$IC = \{(MCM_i.to\_TM, TM.from\_MCM), (TM.to\_MCM, MCM_i.from\_TM)\}; \tag{7}$$

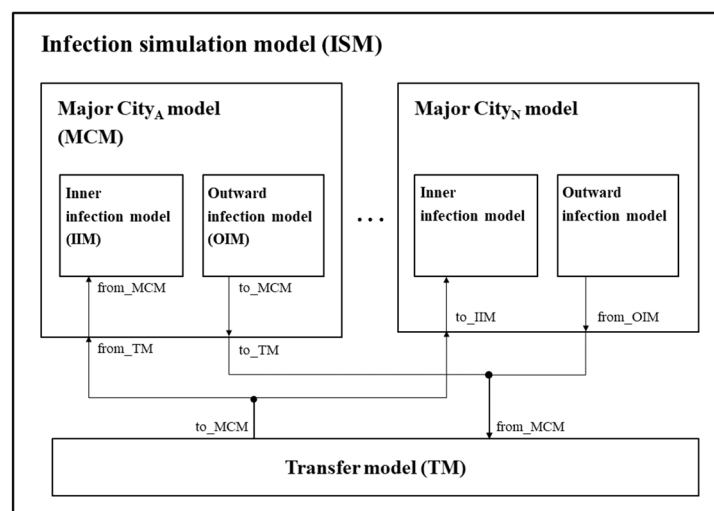$$Sel(\{TM, \cup_{i=1}^{n} \{MCM_i\}\}) = TM. \tag{8}$$

Figure 2. DEVS-coupled model diagram: Infection simulation model (*ISM*).

The type of infection spread can be categorized into two cases: within one community and from other communities [39]. Furthermore, because the time interval in which the infection is internally updated in the community can differ from events externally occurring from other communities, we separate the *MCM* model into two atomic models: the inner infection model (*IIM*) and the outward infection model (*OIM*) [40]. The *IIM* describes the inner infection using the extended SIRD model, and the *OIM* describes the moving behaviors between models based on real-world mobility data.

$$MCM = \langle X, Y, M, EIC, EOC, IC, Sel \rangle \tag{9}$$

$$X = \{from\_TM\} \tag{10}$$

$$Y = \{to\_TM\} \tag{11}$$

$$M = \{IIM, OIM\} \tag{12}$$

$$EIC = \{(MCM.from_{TM}, IIM.from\_MCM)\} \tag{13}$$

$$EOC = \{(OIM.to_{MCM}, MCM.to\_TM)\} \tag{14}$$

$$IC = \varnothing \tag{15}$$

$$Sel(\{IIM, OIM\}) = OIM \tag{16}$$

$$TM = <X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta> \tag{17}$$

$$X = \{from\_MCM\}; \tag{18}$$

$$Y = \{to\_MCM\}; \tag{19}$$

$$S = \{WAIT, SEND\}; \tag{20}$$

$$\delta_{ext} : (WAIT) \times (from\_MCM) \rightarrow (SEND); \tag{21}$$

$$\delta_{int} : (SEND) \rightarrow (WAIT); \tag{22}$$

$$\lambda : (SEND) \rightarrow (to\_MCM).; \tag{23}$$

$$ta : (SEND) \rightarrow 0, \tag{24}$$

$$(WAIT) \rightarrow \infty. \tag{25}$$

Many countries have recently implemented policies to prevent the rapid spread of COVID-19 (e.g., social distancing). Despite the apparent effects [41,42], countries have intermittently used social distancing due to the social and economic impacts [43–46]. To reflect this dynamic political change and its delayed time until the effect occurs, we designed the *IIM* by separating four phases: relaxed social distancing transient and steady-state (*RSD_T*, *RSD_S*) and severe social distancing transient and steady-state (*SSD_T* and *SSD_S*), as depicted in Figure 3. The DEVS expression of the designed model is described in Equations (26)–(33). The model internally refreshes the $S(t)$, $I(t)$, $R(t)$, and $D(t)$ through the *executeExtendSIRD()* at the regular interval (i.e., t_{update}) in each phase, updating infection from the other model by dealing with the *from_TM* event. Then, the model transits to another phase after t_{rsd} or t_{ssd} about two weeks.

$$IIM = < X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta > \tag{26}$$

$$X = \{from\_TM\}; \tag{27}$$

$$Y = \varnothing; \tag{28}$$

$$S = \{RSD\_T, \ RSD\_S, \ SSD\_T, \ SSD\_S\} \tag{29}$$

$$\delta_{ext} : \begin{cases} (RSD\_T) \times (from\_TM) \rightarrow (IF) \text{ and execute } updateInteraction() \\ (RSD\_S) \times (from\_TM) \rightarrow (IF) \text{ and execute } updateInteraction() \\ (SSD\_T) \times (from\_TM) \rightarrow (IF) \text{ and execute } updateInteraction() \\ (SSD\_S) \times (from\_TM) \rightarrow (IF) \text{ and execute } updateInteraction() \end{cases} . \tag{30}$$

$$\delta_{int} : \begin{cases} (RSD_T) \rightarrow (SSD_T) \text{ and execute } executeExtendSIRD(), \\ \qquad if \ t_e == (t_{rsd} - 2weeks) \\ (SSD_T) \rightarrow (SSD_S) \text{ and execute } executeExtendSIRD(), \\ \qquad if \ t_e == t_{rsd} \\ (SSD_S) \rightarrow (RSD_T) \text{ and execute } executeExtendSIRD(), \\ \qquad if \ t_e == (t_{ssd} - 2weeks) \\ (RSD_T) \rightarrow (RSD_S) \text{ and execute } executeExtendSIRD(), \\ \qquad if \ t_e == t_{ssd} \end{cases} . \tag{31}$$

$$\lambda : \varnothing; \tag{32}$$

$$ta : \begin{cases} (RSD\_T) \rightarrow t_{update} \\ (RSD\_S) \rightarrow t_{update} \\ (SSD\_T) \rightarrow t_{update} \\ (SSD\_S) \rightarrow t_{update} \end{cases} \tag{33}$$

To be specific, the moment a social distancing policy starts, the number of infected individuals does not immediately decrease but begins to decrease after a certain period, approximately two weeks [47]. In this study, we defined these transient and steady-state regions, respectively. To reflect this dynamic factor over time, this study extends the existing SIRD model corresponding to Equations (35)–(37). Equations (38) and (39) reflect the proposed infection rates (i.e., β_i and β_d) over the remaining time (t_e) in the *RSD* or *SSD* phase using the attenuation coefficients (i.e., λ and κ), which are mapped in Equations (34) and (35).

$$\frac{dS(t)}{dt} = \begin{cases} -\beta_{last}\frac{SI}{N}, & if \ state == RSD\_S \\ -\beta_l\frac{SI}{N}, & if \ state == SSD\_T \\ -\beta_d\frac{SI}{N}, & otherwise \end{cases} \tag{34}$$

$$\frac{dI(t)}{dt} = \beta\frac{SI}{N} - \gamma I - \mu \tag{35}$$

$$\frac{dR(t)}{dt} = \gamma I \tag{36}$$

$$\frac{dD(t)}{dt} = \mu I \tag{37}$$

$$\beta_i(t) = \beta_{last} \times exp^{\lambda(t_e)} \tag{38}$$

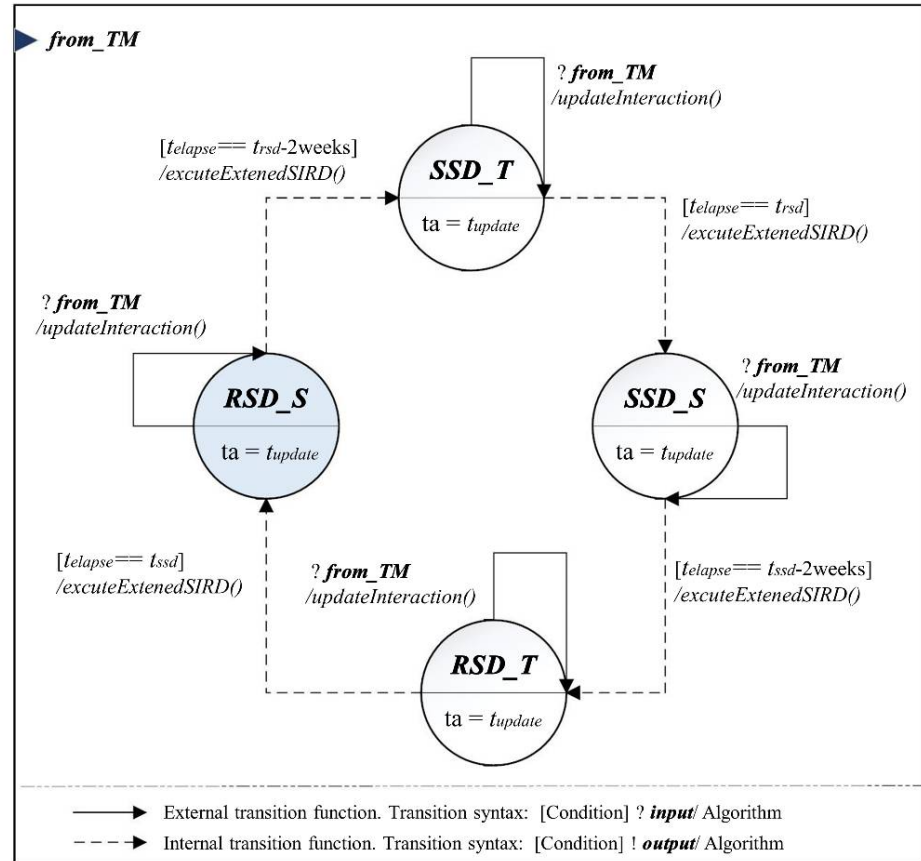$$\beta_d(t) = \beta_{last} \times exp^{-\kappa(t_e)} \tag{39}$$

Figure 3. DEVS-atomic model diagram: Inner infection model (*IIM*).

As the population movement increase is a significant factor in the spread of the infection, it is essential to reflect this phenomenon in the simulation model [48,49]. Although the β in *IIM* indirectly contains the movement within the community, it does not include the external infection. To reflect this property, we designed the outward infection model (*OIM*), which calculated the movement of people using *calculateMovingInfection*() and sends them to the other models using the *to_TM* event at the irregular interval (i.e., t_{move}) in the *MOVE* phase of Figure 4. Equations (40)–(47) are the DEVS expressions of the *OIM*.

$$OIM = < X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta > \tag{40}$$

$$X = \varnothing; \tag{41}$$

$$Y = \{to\_TM\} \tag{42}$$

$$S = \{MOVE,\ STOP\}; \tag{43}$$

$$\delta_{ext} : \varnothing; \tag{44}$$

$$\delta_{int} : (MOVE) \rightarrow (MOVE)\ if\ t_e < t_{finish},$$
$$(SSD) \rightarrow (RSD)\ and\ execute\ simulationStop()\ if\ t_e == t_{finish}; \tag{45}$$

$$\lambda : (MOVE) \rightarrow (to\_TM)\ and\ execute\ calculateMovingInfection() \tag{46}$$

$$ta : (MOVE) \rightarrow t_{move},\ (STOP) \rightarrow \infty. \tag{47}$$
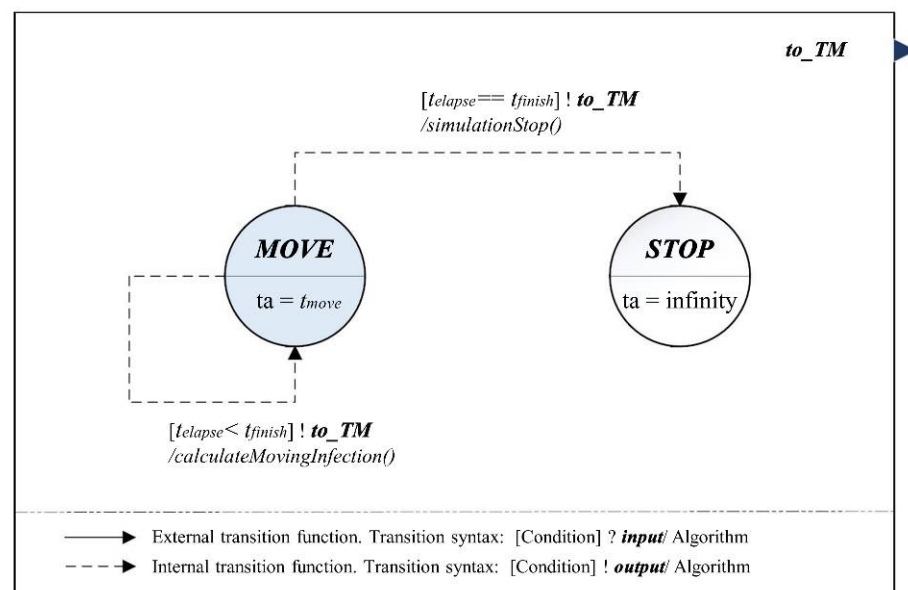
Figure 4. DEVS-atomic model diagram: Outward infection model (*OIM*).

Based on the movement location data of subscribers provided by the mobile carrier [50,51], this study defined Equations (51)–(53) to infer the number of susceptible (S_{ij}), infected (I_{ij}), and recovered (R_{ij}) population moving from major city i to j. In these equations, M means the average number of movement of people to another city per day; r_{mc} means the ratio of total city movement to another major city. The data of mobile carriers indicates that M and r_{mc} are 0.028 and 0.076, respectively. Considering the situation in which people move from major city i to j, the number of the population entering major city j can be defined under the assumption that the proportion of the incoming population is the same across the entire area. Furthermore, assuming that the amount of movement is proportional to the population of the central city, the amount of population sent from major city i to j is equal to $M \times r_{mc} \times N_j \times N_i / \sum(N - N_j)$. Considering the sending population type is proportional to the city's current state, as shown in Equations (48)–(50), each movement amount can be expressed as Equations (51)–(53) to note sending population type. The *IIM* calculates the S_{ij}, I_{ij}, and R_{ij} via the *calculateMovingInfection*() and sends them to the *TM* with the output port *to_TM*.

$$r_s = \frac{S_m(t)}{S_m(t) + I_m(t) + R_m(t)} \tag{48}$$

$$r_I = \frac{I_m(t)}{S_m(t) + I_m(t) + R_m(t)} \tag{49}$$

$$r_R = \frac{R_m(t)}{S_m(t) + I_m(t) + R_m(t)} \tag{50}$$

$$S_{ij} = M \times r_{mc} \times N_j \times \frac{N_i}{\sum N - N_j} \times r_s \tag{51}$$

$$I_{ij} = M \times r_{mc} \times N_j \times \frac{N_i}{\sum N - N_j} \times r_I \tag{52}$$

$$R_{ij} = M \times r_{mc} \times N_j \times \frac{N_i}{\sum N - N_j} \times r_R \tag{53}$$

3.3. Model Identification and What-If Analysis

Although we designed the simulation model, the model included logical behaviors but did not reflect the real-world parameters. Figure 5 shows the process to update the parameters of the model based on real-world data, which we call model identification.
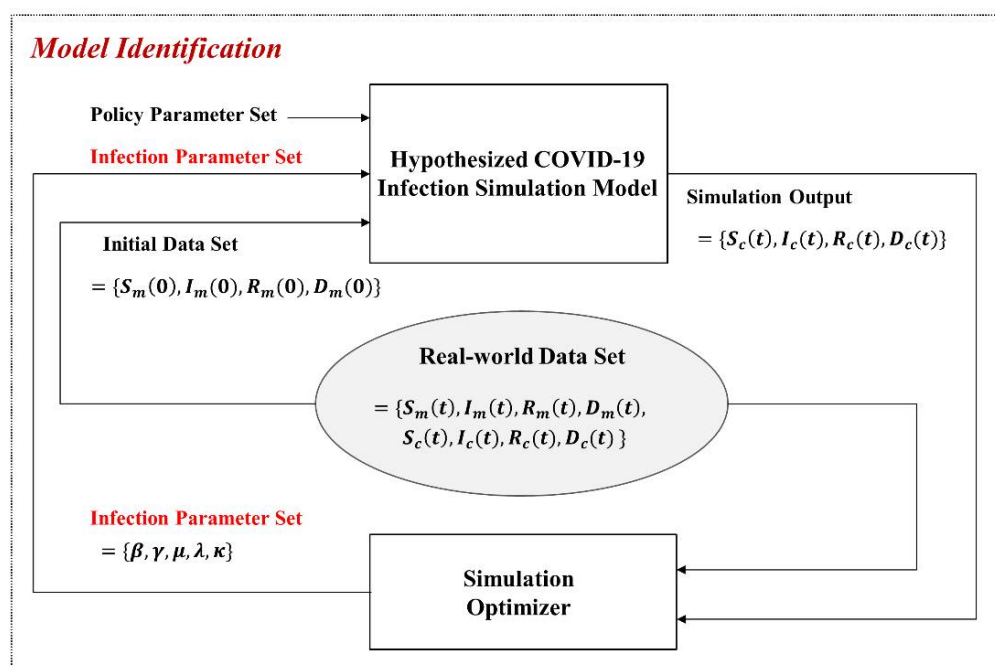
Int. J. Environ. Res. Public Health **2021**, *18*, 11264

10 of 17

Figure 5. Model identification with simulation-based optimization.

From the real world, we can acquire the $S(t)$, $I(t)$, $R(t)$, and $D(t)$ and the policy parameter set, but we cannot directly know the infection parameter set related to the *SIRD executeExtendSIRD*() and can identify it through simulation optimization, including optimization algorithms. At first, the hypothesized model is executed using the initial data set, known policy parameter set, initialized infection parameter set, and the $S(t)$, $I(t)$, $R(t)$, and $D(t)$. Next, the simulation optimizer compares the real-world data set and simulation output and updates the infection parameter set, iteratively simulating until the error converges to an acceptable range. We call the hypothesized model the updated last parameter set identified model.

After the model identification, we can conduct a new analysis of the simulation output against the policy parameter set, which is fixed at the model identification step. The process of model identification finds the unknown part of the entire parameter set; the process of what-if analysis conducts a simulation by changing a part of the entire parameter set.

4. Simulation Experiment

This section confirms that the implemented simulation-based model reflects the real-world infection trend and changes in government policy or infection coefficient.

We conducted the experiments using real-world data for 160 days from 6 May 2020 to 12 October 2020. The start day was when South Korea's distance policy began, and the final day was when the distance policy was upgraded to the second stage.

Simulation results were optimized using the Matlab R2018 version. The algorithm used for optimization was a genetic algorithm. The desktop CPU used in the experiment was a 3.7 GHz Intel i7-8700 K processor, and the memory was 32 GB RAM. The simulator used DEVSim++ and was developed in C++.

4.1. Simulation-Based Model Identification

We define the infection coefficients at the model design step and predict the infected population through model simulation. Because the simulation results regarding the infected population are different from the real-world data, infection coefficients should be optimized to reflect real-world trends and applied to the simulation model. In this study, we use a genetic algorithm to find SIRD coefficients and analyze infection trends, especially social

distancing effects. The optimization phase consists of three steps: (1) error calculation, (2) offspring generation, and (3) mutation generation.

First, we define a value error by a difference between the real-world data of the COVID-19 infected population and our simulation results. We use root mean square error (RMSE) to measure the error quantitatively. The error scale is adjusted to the original scale; thus, it is easy to interpret the error results with simple numbers. Next, for the progeny generation step, the SIRD value is calculated by generating progeny using a part of the parent chromosome. Finally, mutations are used to prevent convergence to local optimization points. It is created with a 0.05% chance of mutation. Equation (54) expresses this process as an equation.

$$p = arg\min_{p} \frac{1}{n} \sqrt{\sum_{t=a}^{n} \{f(t) - y_{real}(t)\}^2} \tag{54}$$

Figure 6 shows the coefficients generated in the process of optimization. The x-axis represents the number of iterations performed, and the y-axis represents the value of each item. Figure 6a–e shows the SIRD coefficient and social distancing policy coefficient used in this study, and Figure 6f is the process of optimizing the above parameters and indicates the error that has occurred. Table 4 shows the optimized infection coefficients.
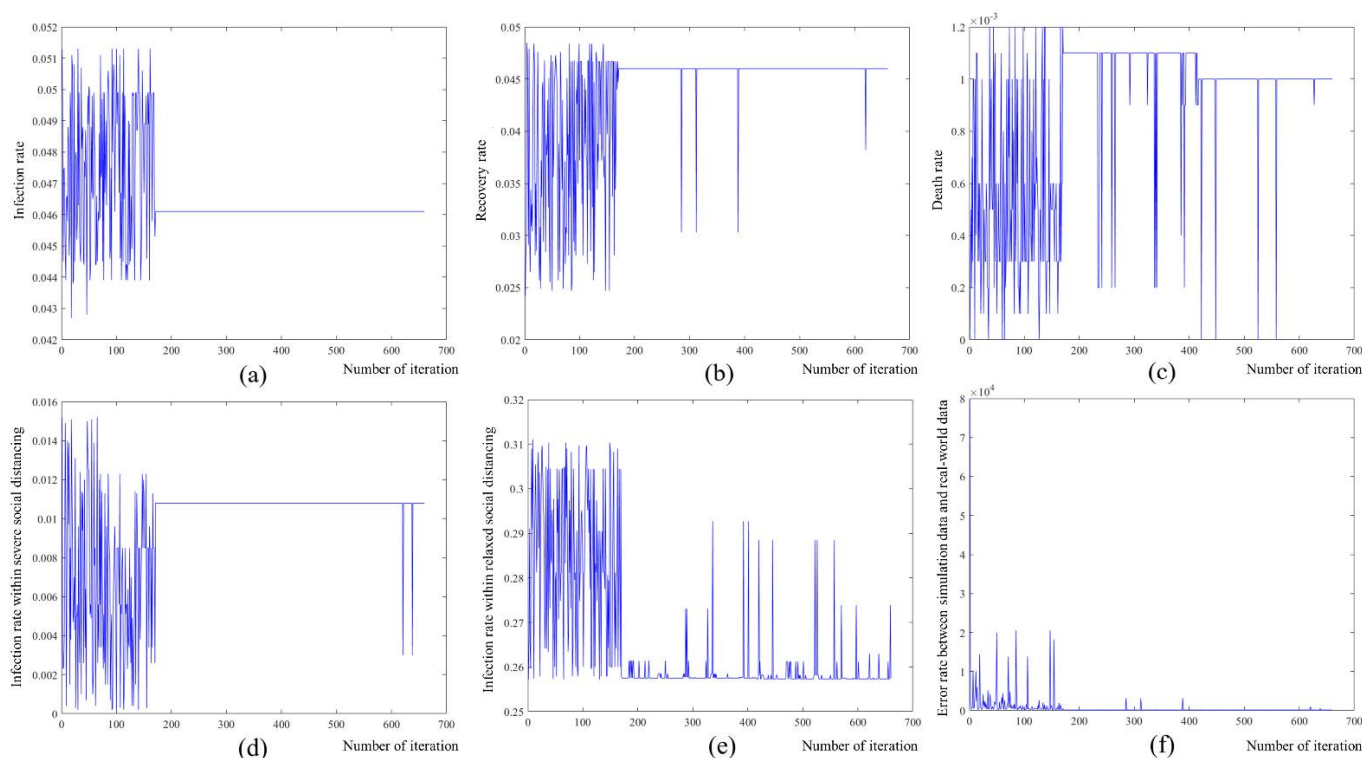

Figure 6. Changes of model parameters via simulation-based optimization with real-world data: (**a**) infection rate; (**b**) recovery rate; (**c**) death rate; (**d**) infection rate within severe social distancing period; (**e**) infection rate within relaxed social distancing period; (**f**) error rate between simulation data and real-world data.

Table 4. Optimized infection coefficients.

Parameter	Optimized Coefficient
β	0.0461
γ	0.0460
μ	0.0010
λ	0.0108
κ	0.2573

Int. J. Environ. Res. Public Health **2021**, 18, 11264

12 of 17

Genetic algorithms pass their genes to the next generation to find solutions that fit well. This process finds an optimal solution by repeating the process of generating chromosomes and calculating the fitness of the offspring. Because of this process, the value of each coefficient in Figure 6 fluctuates and then converges to an optimum value. The optimization process was conducted 660 times and then converged.

Table 4 shows the results of the optimization: the optimized infection coefficients. Figure 7 shows the results of comparing the simulation results with the SIRD values of South Korea by using the optimized infection coefficients. These simulation results show a similar trend to the real SIRD figures in South Korea, meaning that the implemented simulation-based model reflects the real-world environment well.
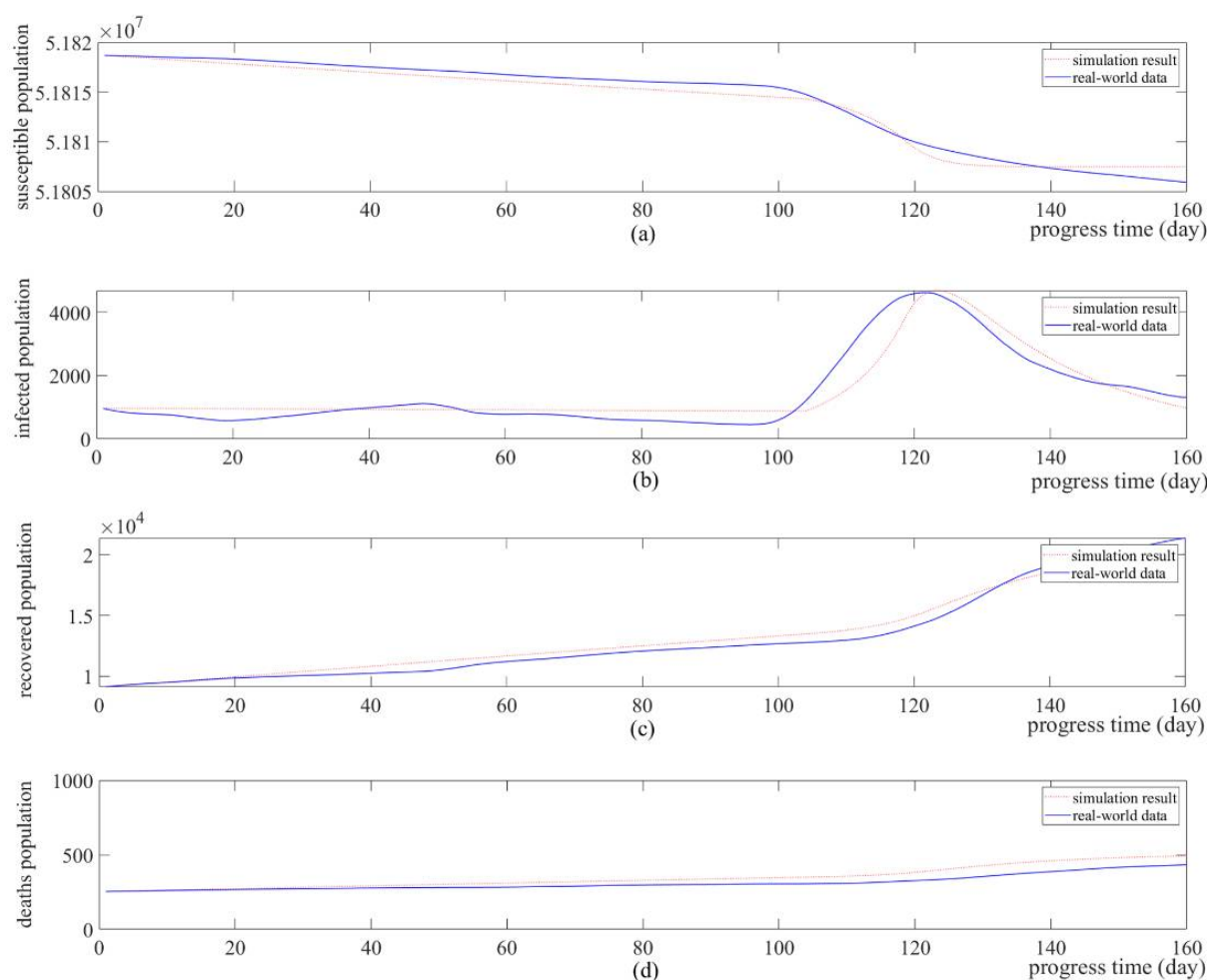

Figure 7. Comparison of simulation results and real-world data. (**a**) number of susceptible population; (**b**) number of infected population; (**c**) number of recovered population; (**d**) number of deaths population.

$\beta\gamma\mu\lambda\kappa$Figure 7a,c,d shows no significant difference compared to the real-world data. However, Figure 7b has a more significant error than the other graphs because the infection rate beta is approximated by an exponential function according to the social distancing policy, as shown in Equations (38) and (39). Thus, real-world data with multiple slopes can be accurately calculated. Therefore, it cannot be approximated.

4.2. What-If Analysis: Simulation of Schedule Change in Social Distancing

The what-if analysis is utilized to optimize infection coefficients in Section 4.1. It shows a comparison of each scenario in which the social distancing policy start date changes. Figure 8 shows the simulation result using the optimized infection coefficient. In this graph, the x-axis represents the simulation time, and the y-axis shows the number of infected

Int. J. Environ. Res. Public Health **2021**, *18*, 11264

13 of 17

people. Alt.1 and Alt.2 represent simulating a scenario in which the starting point of social distancing is advanced by seven days and three days, respectively, and Alt.3 and Alt.4 are when the starting point of social distancing is delayed. The maximum number of infected people and the period of maintaining social distancing are different for each scenario, depending on when social distancing is started.
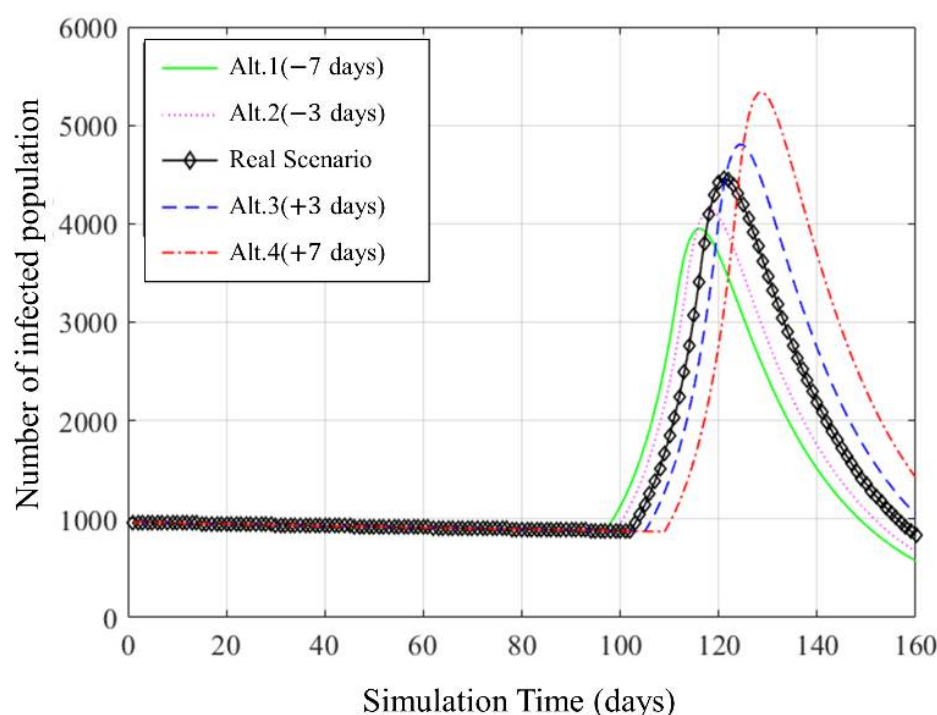

Figure 8. Infected population by varying start time for social distancing policy.

Figure 9a shows the change in the maximum number of infected people that occurs as the starting point of the social distancing policy is changed based on the results confirmed in Figure 8. In the Alt.1 scenario, where social distancing was conducted seven days earlier, up to 3800 people were infected. In Alt.4, which implemented a social distancing policy seven days later, a maximum of 5481 people were infected. Thus, if the social distancing policy is postponed for one day, an additional 250 people could be infected every day. As the simulation data are approximated by an exponential function, if the social distancing policy is delayed further, such as 15 days or 30 days, more infections can occur. On the other hand, if social distancing policies are implemented faster, more infections can be prevented. This shows that the implementation of the distancing policy is closely related to the number of infected people.

Figure 9b represents the period from which the social distancing policy is initiated until the recovery. In Figure 9b, If the social distancing policy is implemented seven days earlier, reaching the sedation period is 56 days. If the social distancing policy is implemented seven days later, it is 64 days to reach the sedation period. The sedation phase has a close relationship with the maximum number of infected populations, as the infected population must reach the threshold of the sedation phase. By delaying the implementation period of social distancing, the maximum infected population increases, and the period of social distancing policy increases. On the other hand, if social distancing is performed in advance, the maximum infected population and the social distancing period can be shortened.
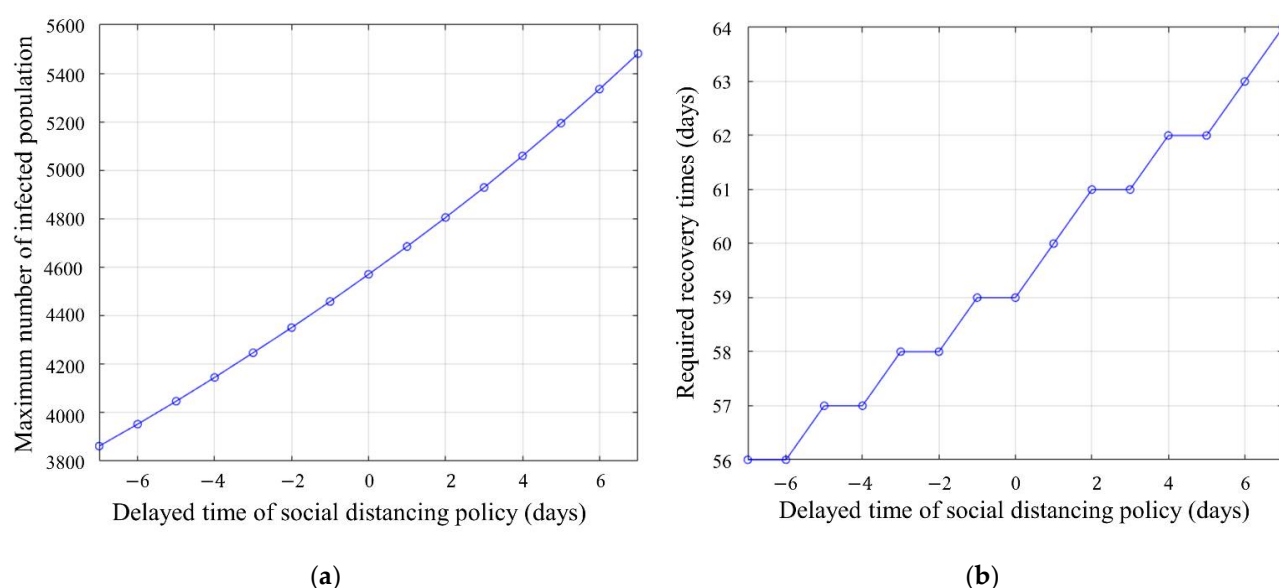
Int. J. Environ. Res. Public Health **2021**, *18*, 11264

14 of 17

Figure 9. Relationship between infected population and distance policy: (**a**) maximum number of infection population by varying social distancing period; (**b**) required recovery time by varying social distancing period.

4.3. Discussion

We designed a hybrid model using DEVSim++ and optimized the simulation model based on Korean data. The what-if analysis experiment in Section 4.2 was conducted through the designed model. The experiment can check the trend of the COVID-19 infected population through the interaction between simulation models that include *MCM* and *TM*. *MCM* calculates SIRD using infection parameters. It also physically sends the number of moving persons between the cities to the *TM*. *TM* transmits the infected population received from the *MCM* to a pre-determined destination to spread the infection.

Here, we discuss two further applications for the developed simulators. First, The simulator overcomes the limitations of the existing standalone SIRD model by modeling the infected population between cities through the *TM*. The simulator overcomes the limitations of the existing standalone SIRD model by modeling the infected population between cities through the *TM*. In the experimental stage, optimization was performed using statistical data sets of daily and monthly infected populations in major cities in South Korea. Therefore, the infection coefficient used in this study is applied only to the case of South Korea. However, as long as there is a data set of the country you want to apply, it can be commonly applied to all countries.

Next, it shows the correlation between the maximum number of infected people and the length of time to reach recovery. The simulator of this study can infer the recovery period from the maximum infected population according to the time of distance as shown in Figure 9. As a result, it shows that it can be a reference for social distancing policy.

5. Conclusions

The proposed method includes two sections: (1) hybrid modeling and (2) simulation-based optimization. The form describes the interaction among large-scale cities by integrating the existing differential equation-based model and agent-based model. The latter makes the model reflect the real world by updating the inner states of the hybrid model through the use of public data.

The proposed model simulates the number of suspected, infected, recovered, and deceased COVID-19 patients in South Korea using the SIRD infection coefficient and policy coefficient. The model is optimized with the results of the SIRD model output from the previous step to fit the data of the real-world SIRD model. In the last step, as a what-

Int. J. Environ. Res. Public Health **2021**, *18*, 11264

15 of 17

if analysis, we used the policy coefficient to simulate pulling and delaying the Korean government's distance policy enforcement, and we experimented to confirm the result.

We expect that the proposed work plays a principal role in analyzing how social distancing effectively affects virus prevention and provides a simulation environment for the biochemical field.

Abbreviations

ABM	Agent-Based Model
D	Dead
DEVS	Discrete Event System Specifications
EBM	Equation-Based Model
EIC	External Input Coupling Relation
EOC	External Output Coupling Relation
I	Infected
IC	Internal Coupling Relation
IIM	Inner Infection Model
ISM	Infection Simulation Model
MCM	Major City Model
OIM	Outward Infection Model
R	Recovered
RMSE	Root Mean Square Error
RSD	Relaxed Social Distancing
RSD_S	Relaxed Social Distancing Steady State
RSD_T	Relaxed Social Distancing Transient State
S	Susceptible
SEIR	Susceptible, Exposed, Infected, and Recovered
SIR	Susceptible, Infected, and Recovered
SIRD	Susceptible, Infected, Recovered, and Dead
SEIR	Suspected, Exposed, Infected, and Recovered
SSD	Severe Social Distancing
SSD_S	Severe Social Distancing Steady State
SSD_T	Severe Social Distancing Transient State
TM	Transfer Model

References

1. Kim, Y.; Jiang, X. Evolving transmission network dynamics of COVID-19 cluster infections in South Korea: A descriptive study. *MedRxiv* **2020**. [CrossRef]
2. Shukla, P.; Kundu, R.; Arivarasi, A.; Alagiri, G. A Social Distance Monitoring System to ensure Social Distancing in Public Areas. In Proceedings of the 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 17–18 March 2021; pp. 96–101.
3. Roncone, R.; Giusti, L.; Mammarella, S.; Salza, A.; Bianchini, V.; Lombardi, A.; Prosperocco, M.; Ursini, E.; Scaletta, V.; Casacchia, M.; et al. "Hang in There!": Mental Health in a Sample of the Italian Civil Protection Volunteers during the COVID-19 Health Emergency. *Int. J. Environ. Res. Public Health* **2021**, *18*, 8587. [CrossRef] [PubMed]
4. Choi, S.; Ki, M. Analyzing the effects of social distancing on the COVID-19 pandemic in Korea using mathematical modeling. *Epidemiol. Health* **2020**, *42*, e2020064. [CrossRef] [PubMed]
5. Yong, S.E.F.; Anderson, D.E.; Wei, W.E.; Pang, J.; Chia, W.N.; Tan, C.W.; Teoh, Y.L.; Rajendram, P.; Toh, M.P.H.S.; Poh, C.; et al. Connecting clusters of COVID-19: An epidemiological and serological investigation. *Lancet Infect. Dis.* **2020**, *10*, 809–815. [CrossRef]

6. McGrail, D.J.; Dai, J.; McAndrews, K.M.; Kalluri, R. Enacting national social distancing policies corresponds with dramatic reduction in COVID19 infection rates. *PLoS ONE* **2020**, *15*, e0236619. [CrossRef]
7. Kim, B.S.; Nam, S.; Jin, Y.; Seo, K.-M. Simulation Framework for Cyber-Physical Production System: Applying Concept of LVC Interoperation. *Complexity* **2020**, *2020*, 4321873. [CrossRef]
8. Kang, B.G.; Seo, K.-M.; Kim, T.G. Model-based Design of Defense Cyber-Physical Systems to Analyze Mission Effectiveness and Network Performance. *IEEE Access* **2019**, *7*, 42063–42080. [CrossRef]
9. Seo, K.-M.; Hong, W.; Kim, T.G. Enhancing model composability and reusability for entity-level combat simulation: A conceptual modeling approach. *Simulation* **2017**, *93*, 825–840. [CrossRef]
10. Fan, R.; Wang, Y.; Lin, J. Study on Multi-Agent Evolutionary Game of Emergency Management of Public Health Emergencies Based on Dynamic Rewards and Punishments. *Int. J. Environ. Res. Public Health* **2021**, *18*, 8278. [CrossRef]
11. Hussein, T.; Löndahl, J.; Thuresson, S.; Alsved, M.; Al-Hunaiti, A.; Saksela, K.; Aqel, H.; Junninen, H.; Mahura, A.; Kulmala, M. Indoor model simulation for covid-19 transport and exposure. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2927.
12. Qi, B.; Tan, J.; Zhang, Q.; Cao, M.; Wang, X.; Zou, Y. Unfixed Movement Route Model, Non-Overcrowding and Social Distancing Reduce the Spread of COVID-19 in Sporting Facilities. *Int. J. Environ. Res. Public Health* **2021**, *18*, 8212. [CrossRef]
13. Chen, Z.; Shu, Z.; Huang, X.; Peng, K.; Pan, J. Modelling Analysis of COVID-19 Transmission and the State of Emergency in Japan. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6858. [CrossRef] [PubMed]
14. Nesteruk, I. Estimations of the coronavirus epidemic dynamics in South Korea with the use of SIR model. *Preprint* **2020**. [CrossRef]
15. Mulder, L.J. A set of operational parameters for the SEIR model of the coronavirus for the world and some of the more heavily affected countries. *J. Infect. Disease Model.* **2020**. [CrossRef]
16. Calafiore, G.C.; Novara, C.; Possieri, C. A Modified SIR Model for the COVID-19 Contagion in Italy. In Proceedings of the 2020 59th IEEE Conference on Decision and Control (CDC), Jeju, Korea, 14–18 December 2020; Volume 59, pp. 3889–3894. [CrossRef]
17. Alagoz, O.; Sethi, A.K.; Patterson, B.W.; Churpek, M.; Safdar, N. Effect of Timing of and Adherence to Social Distancing Measures on COVID-19 Burden in the United States. *Ann. Intern. Med.* **2021**, *174*, 50–57. [CrossRef]
18. D'Orazio, M.; Bernardini, G.; Quagliarini, E. How to restart? An agent-based simulation model towards the definition of strategies for COVID-19 "second phase" in public buildings. *arXiv* **2020**, arXiv:2004.12927.
19. Cooper, I.; Mondal, A.; Antonopoulos, C.G. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos Solitons Fractals* **2020**, *139*, 110057. [CrossRef] [PubMed]
20. Malavika, B.; Marimuthu, S.; Joy, M.; Nadaraj, A.; Asirvatham, E.S.; Jeyaseelan, L. Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clin. Epidemiol. Glob. Heal.* **2021**, *9*, 26–33. [CrossRef] [PubMed]
21. Chen, X.; Li, J.; Xiao, C.; Yang, P. Numerical solution and parameter estimation for uncertain SIR model with application to COVID-19. *Fuzzy Optim. Decis. Mak.* **2020**, *20*, 189–208. [CrossRef]
22. Piovella, N. Analytical solution of SEIR model describing the free spread of the COVID-19 pandemic. *Chaos Solitons Fractals* **2020**, *140*, 110243. [CrossRef]
23. Zhou, X.; Zhou, X.; Ma, X.; Hong, N.; Su, L.; Ma, Y.; He, J.; Jiang, H.; Liu, C.; Shan, G.; et al. Forecasting the Worldwide Spread of COVID-19 based on Logistic Model and SEIR Model. *MedRxiv* **2020**. [CrossRef]
24. Mahmud, A.; Lim, P.Y. Applying the SEIR model in forecasting the COVID-19 trend in malaysia: A preliminary study. *MedRxiv* **2020**. [CrossRef]
25. Ferrari, L.; Gerardi, G.; Manzi, G.; Micheletti, A.; Nicolussi, F.; Biganzoli, E.; Salini, S. Modeling provincial Covid-19 epidemic data in Italy using an adjusted time-dependent SIRD model. *arXiv* **2020**, arXiv:2005.12170v2.
26. Chatterjee, S.; Sarkar, A.; Chatterjee, S.; Karmakar, M.; Paul, R. Studying the progress of COVID-19 outbreak in India using SIRD model. *Indian J. Phys.* **2020**. [CrossRef]
27. Fanelli, D.; Piazza, F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Solitons Fractals* **2020**, *134*, 109761. [CrossRef]
28. Osemwinyen, A.C. Mathematical Modelling of the Transmission Dynamics of Ebola Virus. *Appl. Comput. Math.* **2015**, *4*, 313–320. [CrossRef]
29. Basti, B.; Hammami, N.; Berrabah, I.; Nouioua, F.; Djemiat, R.; Benhamidouche, N. Stability Analysis and Existence of Solutions for a Modified SIRD Model of COVID-19 with Fractional Derivatives. *Symmetry* **2021**, *13*, 1431. [CrossRef]
30. Shringi, S.; Sharma, H.; Rathie, P.; Bansal, J.; Nagar, A. Modified SIRD Model for COVID-19 Spread Prediction for Northern and Southern States of India. *Chaos Solitons Fractals* **2021**, *148*, 111039. [CrossRef] [PubMed]
31. Devosmita, S.; Debasis, S. Use of Modified SIRD Model to Analyze COVID-19 Data. *Ind. Eng. Chem. Res.* **2021**, *60*, 4251–4260. [CrossRef]
32. Kemeny, A.; Özmen, Ö.; Nutaro, J.J.; Pullum, L.L.; Ramanathan, A. Analyzing the impact of modeling choices and assumptions in compartmental epidemiological models. *Simulation* **2016**, *92*, 459–472.
33. Zeigler, B.P.; Muzy, A.; Kofman, E. *Theory of Modeling and Simulation: Discrete Event & Iterative System Computational Foundations*; Academic Press: Cambridge, CA, USA, 2018.
34. Kim, T.G.; Sung, C.H.; Hong, S.Y.; Hong, J.H.; Choi, C.B.; Kim, J.H.; Seo, K.M.; Bae, J.W. DEVSim++ Toolset for Defense Modeling and Simulation and Interoperation. *J. Def. Model. Simul.* **2011**, *8*, 129–142. [CrossRef]
35. Venkatramanan, S.; Lewis, B.; Chen, J.; Higdon, D.; Vullikanti, A.; Marathe, M. Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics* **2018**, *22*, 43–49. [CrossRef]

36. Hong, J.H.; Seo, K.M.; Kim, T.G. Simulation-based optimization for design parameter exploration in hybrid system: A defense system example. *Simulation* **2013**, *89*, 362–380. [CrossRef]

37. Kang, B.G.; Choi, S.H.; Kwon, S.J.; Lee, J.H.; Kim, T.G. Simulation-based optimization on the system-of-systems model via model transformation and genetic algorithm: A case study of network-centric warfare. *Complexity* **2018**. [CrossRef]

38. Choi, S.H.; Bae, J.W. An Effective Adjustment to the Integration of Optimal Computing Budget Allocation for Particle Swarm Optimization in Stochastic Environments. *IEEE Access* **2020**, *8*, 173654–173665. [CrossRef]

39. Shim, E.; Tariq, A.; Choi, W.; Lee, Y.; Chowell, G. Transmission potential and severity of COVID-19 in South Korea. *Int. J. Infect. Dis.* **2020**, *93*, 339–344. [CrossRef] [PubMed]

40. Wei, Y.; Wang, J.; Song, W.; Xiu, C.; Ma, L.; Pei, T. Spread of COVID-19 in China: Analysis from a city-based epidemic and mobility model. *Cities* **2021**, *110*, 103010. [CrossRef]

41. Courtemanche, C.; Garuccio, J.; Le, A.; Pinkston, J.; Yelowitz, A. Strong social distancing measures in the united states reduced the covid-19 growth rate. *Health Aff.* **2020**, *39*, 1237–1246. [CrossRef]

42. Moosa, I.A. The effectiveness of social distancing in containing Covid-19. *Appl. Econ.* **2020**, *52*, 6292–6305. [CrossRef]

43. McKibbin, W.J.; Fernando, R. The Global Macroeconomic Impacts of COVID-19: Seven Scenarios. *SSRN Electron. J.* **2020**, *19*, 1–30. [CrossRef]

44. Brotherhood, L.; Kircher, P.; Santos, C.; Tertilt, M. An Economic Model of the Covid-19 Epidemic: The Importance of Testing and Age-Specific Policies. *IZA Inst. Lab. Econ. Discuss.* **2020**, *13265*, 1–71.

45. Zhao, S.Z.; Wong, J.Y.H.; Wu, Y.; Choi, E.P.H.; Wang, M.P.; Lam, T.H. Social distancing compliance under covid-19 pandemic and mental health impacts: A population-based study. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6692. [CrossRef]

46. Banks, J. Institute for Fiscal Studies: The mental health effects of the first two months of lockdown and social distancing during the Covid-19 pandemic in the UK. *Inst. Fisc. Stud.* **2020**. [CrossRef]

47. Morato, M.M.; Pataro, I.M.L.; Americano da Costa, M.V.; Normey-Rico, J.E. A parametrized nonlinear predictive control strategy for relaxing COVID-19 social distancing measures in Brazil. *ISA Trans.* **2020**. [CrossRef] [PubMed]

48. Armando, C.; Luigi, D.F.; Maria, M. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Sci. Total Environ.* **2020**, *741*, 140489. [CrossRef]

49. Badr, H.S.; Du, H.; Marshall, M.; Dong, E.; Squire, M.M.; Gardner, L.M. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 1247–1254. [CrossRef]

50. Zhou, Y.; Xu, R.; Hu, D.; Yue, Y.; Li, Q.; Xia, J. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: A modelling study using mobile phone data. *Lancet Digit. Heal.* **2020**, *2*, 417–424. [CrossRef]

51. Chang, S.; Pierson, E.; Koh, P.W.; Gerardin, J.; Redbird, B.; Grusky, D.; Leskovec, J. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **2021**, *589*, 82–87. [CrossRef] [PubMed]