



OPEN

Deep learning forecasting using time-varying parameters of the SIRD model for Covid-19

Arthur Bousquet¹, William H. Conrad², Said Omer Sadat¹, Nelli Vardanyan¹ & Youngjoon Hong³✉

Accurate epidemiological models are necessary for governments, organizations, and individuals to respond appropriately to the ongoing novel coronavirus pandemic. One informative metric epidemiological models provide is the basic reproduction number (R_0), which can describe if the infected population is growing ($R_0 > 1$) or shrinking ($R_0 < 1$). We introduce a novel algorithm that incorporates the susceptible-infected-recovered-dead model (SIRD model) with the long short-term memory (LSTM) neural network that allows for real-time forecasting and time-dependent parameter estimates, including the contact rate, β , and deceased rate, μ . With an accurate prediction of β and μ , we can directly derive R_0 , and find a numerical solution of compartmental models, such as the SIR-type models. Incorporating the epidemiological model dynamics of the SIRD model into the LSTM network, the new algorithm improves forecasting accuracy. Furthermore, we utilize mobility data from cellphones and positive test rate in our prediction model, and we also present a vaccination model. Leveraging mobility and vaccination schedule is important for capturing behavioral changes by individuals in response to the pandemic as well as policymakers.

According to the World Health Organization (WHO), over 5 million people worldwide have died from Covid-19¹. Public health interventions have limited incidence and mortality of this disease from an early stage². Governments, public health institutions, and the public at large benefit from statistical models that help to determine what approaches are effective at controlling the virus, and to predict when it is necessary to take strong measures to slow its transmission. For instance, recent studies have shown the benefits of both voluntary and government-induced social distancing measures³. A key metric to predict epidemic progression is the basic reproduction number (R_0)⁴. Both compartmental models and networked models have been used to predict R_0 ^{5–14}. Compartmental models such as the SIRD (susceptible, infectious, recovered, or dead) model and its variants are used to predict R_0 for infectious diseases^{15,16}, because the number of susceptible, infectious, recovered, and dead people in a population can be readily estimated from publicly available data. The epidemiological parameters β (contact rate), γ (recovery rate), and μ (deceased rate) in the SIRD model can be estimated from the number of the susceptible, infectious, recovered, and dead. These parameters are then used to determine R_0 .

The SIRD model (see Eq. 2) can be solved in different ways. One of the popular methods is to solve the ordinary differential equations (ODEs) using numerical methods. In this case, one needs to know the parameters (β , γ , and μ) of the system of differential equations. However, if these parameters are set to be a time-independent constant, the assumption may not be realistic. For example, the contact rate, β , varies depending on many time-dependent factors such as mobility and lockdown policy. Hence, time-independent parameters may give outdated information, which introduces a prediction error. Another method to solve the equations is to use neural networks by considering the system as a time series¹⁷ with a recurrent neural network¹⁸. This approach does not ensure that the model follows the dynamics of compartmental models, and the neural network is required to predict twice as many variables. More importantly, this approach does not provide the reproduction rate R_0 directly.

Recently, related studies about Covid-19 incorporated mobility datasets to aid in pandemic modeling^{19–21}. For example, James and Menzies¹⁹ used Apple mobility data to examine the relationship between daily Covid-19 cases and national equity index price on a country-by-country basis. Yilmazkuday²⁰ studied the relationship between country-specific changes in mobility, from the Google mobility dataset, and the number of Covid-19 cases. Also, a metapopulation SEIR model was investigated in²¹ that integrated fine-grained dynamic mobility from Safegraph data to simulate the spread of Covid-19. Each of these studies demonstrated that by integrating

¹Department of Mathematics and Data Science, Lake Forest College, Lake Forest, CA, USA. ²Department of Chemistry, Lake Forest College, Lake Forest, CA, USA. ³Department Mathematics, Sungkyunkwan University, Suwon, South Korea. ✉email: hongyj@skku.edu

these mobility data, the SEIR model can accurately fit the real case trajectory, despite substantial changes in the behavior of the population over time.

In this work, we combine a compartmental model with a recurrent neural network that incorporates mobility data as well as the positive test rate. We (1) predict the time-dependent parameters β and μ using a neural network; (2) forecast the infection rates when mobility decreases or increases; and (3) forecast the change in infection rate based on different vaccination schedules. The goal of this paper is to provide a method to predict time-varying parameters β and μ (and hence R_0) as well as to solve the SIRD equations.

The method under consideration in our paper combines the two aforementioned approaches. We first introduce a version of recurrent neural networks to predict the time-varying parameters β and μ . Since γ is assumed to be constant, one can easily find $R_0 = \beta/\gamma$ from the neural network. We then obtain the compartments, S, I, R, and D, by numerically solving the SIRD equation over a certain time period (e.g. 7 days). To test the performance of our approach, we used publicly available data for different countries, France, United Kingdom, Germany, and South Korea, provided by Johns Hopkins University. For more detail, we provide an illustration of the algorithm in Fig. 9. We also include two additional datasets: mobility data from cellphones and the positive test rate. Studies reveal that both mobility and positive test rate have been shown to influence the spread of Covid-19 considerably^{22–25}.

In this paper, we present an accurate computational scheme to predict the reproduction number which enables Covid-19 forecasting. We use this scheme to forecast different scenarios by increasing or decreasing the mobility parameter. In doing so, our model can help study the effect of government-imposed lockdowns on R_0 . Furthermore, we make use of a SIRD model with vaccination to see how vaccination affects the spread of the virus. Among many other vaccination models^{26,27}, our study focuses on the model introduced in^{28,29} as it is sufficient to capture important dynamics in the experiments. By leveraging parameters relative to the vaccination rates, our simulations show how the vaccination rate affects the number of infectious cases. Such experiments can show how different public health interventions may affect the outcome of the epidemic.

Results

In this section, we describe a sequence of numerical experiments of our algorithm further detailed in the Method section below. First, we present the estimated values of our time-dependent parameters β and μ using the Levenberg–Marquardt algorithm. Then, the accuracy of the algorithm is demonstrated using in-sample data, and out-of-sample predictions for the next 10 weeks. Lastly, forecasting depending on mobility and vaccination rate is examined. In summary, our main contributions consist of three key findings; (i) our SIRD–LSTM combined network outperforms classical prediction models; (ii) we incorporate the mobility and vaccination as inputs of our neural network to increase the accuracy of our parameters predictions; (iii) we forecast Covid-19 trends when mobility decreases or increases.

Parameter Estimates. A significant finding of our paper is that treating the parameters β and μ as time-dependent increases model accuracy. Figure 1 shows (β, μ) for four countries (France, United Kingdom, Germany, and South Korea) generated by the Levenberg–Marquardt algorithm. From this, we can find the basic reproduction number, $R_0 = \frac{\beta}{\gamma}$, with $\gamma = 1/14$, which is useful to study the dynamics of the infectious class³⁰. We compare real infection data from France, the United Kingdom, Germany, and South Korea with a SIRD model using constant β or time-dependent β . Figure 2 shows the difference between a β and μ constant, that we estimate using the Levenberg–Marquardt algorithm over one year, with β and μ estimated over just 1 week. The time-dependent model more accurately forecasts the infection rate over seven days across each country regardless of the time period. Therefore, it is necessary to consider β and μ as time-dependent variables.

Accuracy of our model. To test the forecasting capability of the SIRD–LSTM combined network, we compare the number of predicted confirmed Covid-19 cases under various measures for within sample scenarios. The in-sample fit of the model is an essential indicator for the validity of the model's prediction of the parameters, whereas the out-of-sample forecasts can provide an important guideline for decision/policymakers. Figure 3 depicts the prediction of the time varying parameters (β, μ) compared with (β, μ) from the dataset. We randomly choose N_T test data amongst 365 days, and make use of them as a test set. To measure the accuracy, we use the relative- L^2 errors of $\beta, \mu, S, I, R,$ and D such that

$$\text{Errors} = \frac{1}{N_T} \sum_{i=1}^{N_T} \frac{\|Y_{true}^i - Y_{pred}^i\|_{L^2}}{\|Y_{true}^i\|_{L^2}}, \quad N_T = \text{number of test days (test set)}, \quad (1)$$

where Y_{true}^i is the i th true dataset of $\beta, \mu, S, I, R,$ or D , and Y_{pred}^i is i th predicted values from our algorithm. We observe that the predicted and true parameters are close to each other. Table 1 demonstrates quantitative results on accuracy of our computation.

Table 1 shows the relative L^2 error of β is between 3.13×10^{-3} and 6.29×10^{-2} , the relative L^2 error of μ is between 9.26×10^{-2} and 1.73×10^{-1} . The relative L^2 error with $N_T = 14$ (2 weeks), of the compartments, $S, I, R,$ and D , is also displayed in Table 1. Figure 4 depicts mobility, positive test rate, cumulative infectious individuals, and contact ratio β against the time. The positive test rate and cumulative infectious individuals follow similar trends as opposed to mobility and the positive test rate. The countries under consideration enforce lockdowns as cumulative infectious individuals increased. Hence, the trend plots reveal that greater mobility leads to an increase in infectious individuals.

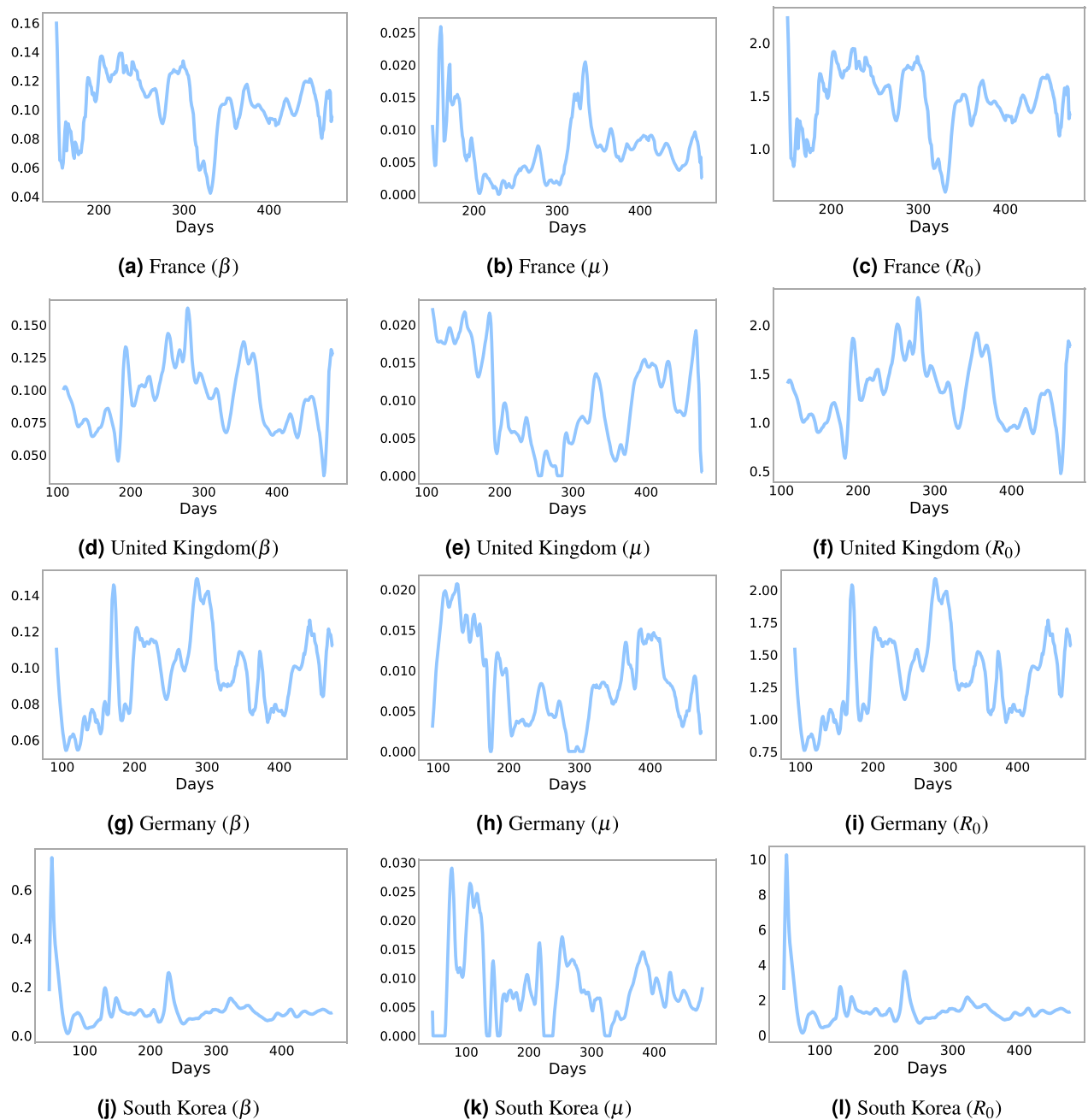


Figure 1. Predicted time varying parameters β (contact rate), μ (deceased rate), and R_0 reproduction number for each country derived from the Levenberg–Marquardt algorithm.

Out-of-sample forecast. We next conduct an out-of-sample forecast analysis of our SIRD–LSTM combined model. Figure 5 demonstrates a prediction of R_0 of each country using β generated by the LSTM networks. By forecasting β , in Fig. 6, we show a short-term prediction of the SIRD model up to 10 weeks. In the simulation, we assume that the positive test rate and mobility are the same as the final observation from the dataset. Both the SIRD and vaccinated SIRD models are computed and demonstrated in Fig. 6. In France, Germany, and South Korea, the depicted curves of the infections for the next 10 weeks are increasing, while the infection curve for the next 10 weeks tends to slightly decrease in the United Kingdom. In fact, it has been reported from various sources in May 2021 that the vaccination strategy and lockdowns in the United Kingdom were successful³¹.

Forecasting depending on mobility. Policymakers have sought to decrease the rate of infection in their populations by decreasing population mobility through lockdowns and, more recently, increasing vaccinations. Here, we model the effect of decreasing mobility and increasing vaccination rate on infection rate. If the mobility is increased by 30% of the normal mobility (baseline mobility), the model shows that the peak of infectious individuals increases drastically, see Fig. 7. The data show how visits to places are changing compared to the baseline. A baseline day represents a normal value for that day of the week. The baseline day is the median value

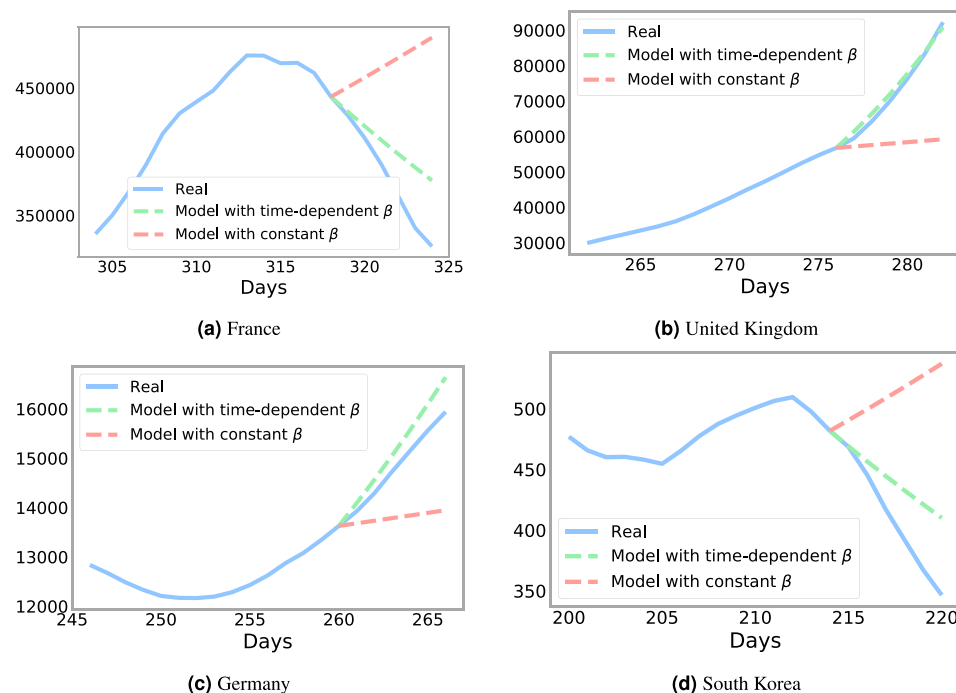


Figure 2. Comparison between the number of infected, I , from our data with a predicted I using the SIRD model using a constant β or a time-dependent β .

from the 5 weeks Jan 3–Feb 6, 2020; for more information, see e.g.³² Figure 7 shows that in France, South Korea, and Germany, increased mobility results in a drastic change in the number of new Covid-19 cases. On the other hand, if mobility restrictions are decreased to 30% normal mobility, the model predicts that the peak of infectious individuals decreases compared to the baseline mobility.

Forecasting depending on the vaccination rate. In addition, with vaccination, the Covid-19 cases are noticeably decreasing for all of the countries under study in our work. The countries whose reproduction number ($R_0 = \beta/\gamma$) is close to 1 such as the United Kingdom and South Korea, have a better vaccination effect than the other countries. Figure 8 displays forecasting of infectious cases under various vaccination schedules within 70 days. In the experiment, we assume that the vaccine is evenly distributed with respect to time. The plots reveal that high vaccination rates are important in reducing the number of infectious cases. Figure 7 shows the models' forecast for infections with different mobility levels in each country. Given mobility information, the combined SIRD–LSTM model can predict the time-varying parameters (β, μ). With those predicted parameters, the number of infectious individuals are implemented with or without vaccination. Based on the projected forecasts, we observe that a continuation of quarantine level mobility will result in low case counts.

Discussion

We introduced a novel algorithm that incorporates deep learning and compartmental models allowing for forecasts and evaluation of the current Covid-19 outbreak worldwide. We combined the SIRD model with the LSTM network and observed advantages of real-time forecasting and parameter estimation. The new algorithm integrates the forecasting accuracy of LSTM networks with the epidemiological model dynamics of the SIR-type model. Compared to the classical SIRD model in the literature, we forecast time-varying parameters predicted by the LSTM neural network. To forecast the parameters, mobility and positive test rate data are used in the architecture. We find that these inputs are important in improving the model's ability to fit the data. In addition, incorporating these data is essential for capturing behavioral changes by individuals in response to the pandemic as well as to observe the effect of policy decisions to increase vaccination and decrease mobility. As in other approaches, we conduct our research on publicly available datasets. We demonstrate how a new algorithm can be developed to better exploit quantitative measures in the fight against Covid-19. By utilizing reliable metrics and infection dynamics, we provide an approach that is deeply data-driven and computer-based. The proposed simulations can provide a tool for forecasting the effects of different mobility scenarios. Furthermore, as the proposed algorithm is compatible and generalizable, this allows for additional compartments in the SIR model or additional input datasets in the network which makes the method accessible to policymakers.

Our developments point towards several extensions of great importance. In particular, we evaluated the impact of the imposition and relaxation of lockdown measures by inputting these changes into the LSTM neural network. We found that employing lockdown rules for each country can help to capture interesting regional dynamics of the Covid-19, and may give specific information to the policymakers. Another direction is to study

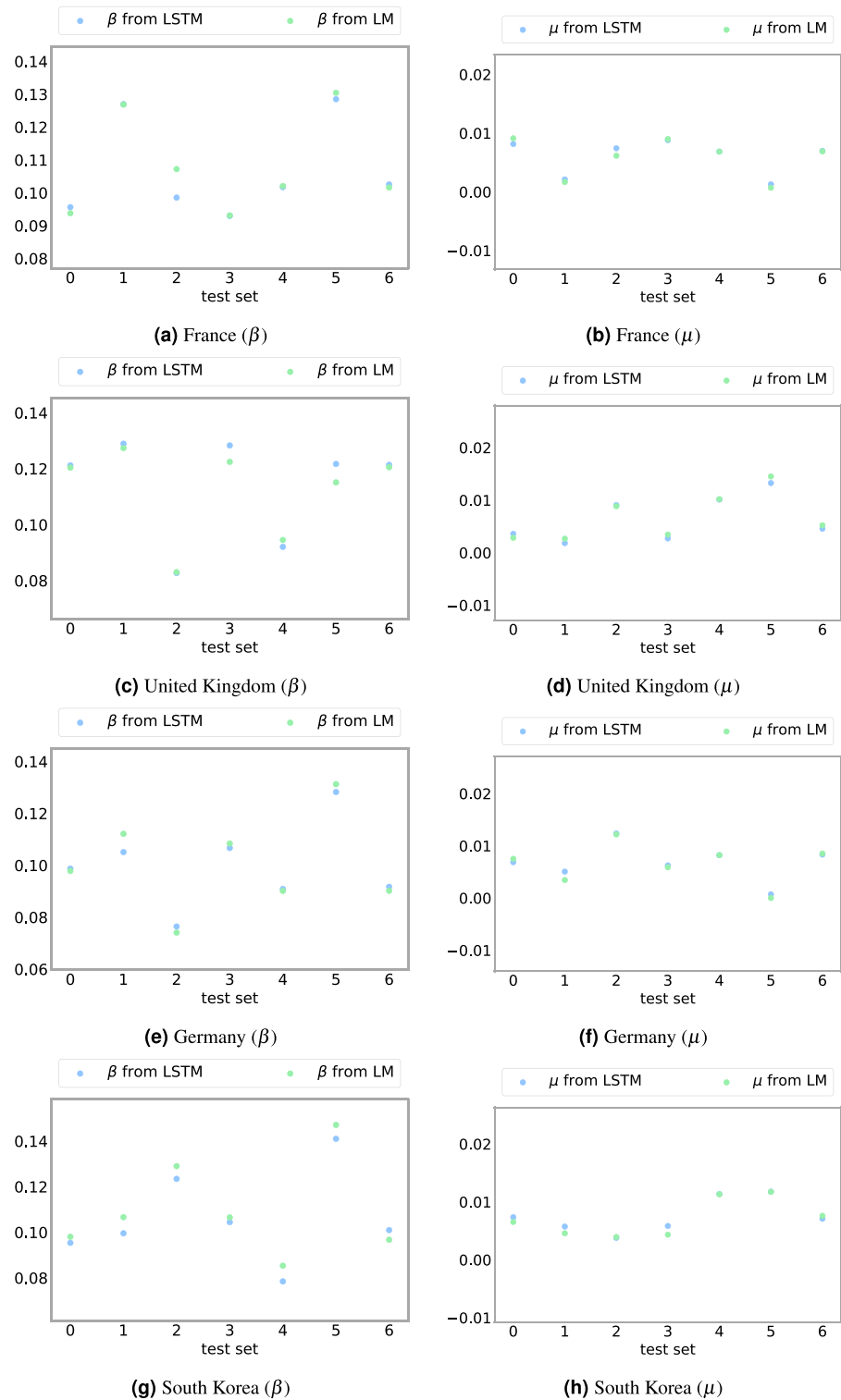


Figure 3. Comparison between (β, μ) computed by the Levenberg–Marquardt (LM) algorithm which is considered as our true data, and (β, μ) predicted from the LSTM networks for France, United Kingdom, Germany, and South Korea.

an advanced deep learning architecture such as attention mechanism or transformer³³. These modern architectures can provide better investigation on not only the increase in forecasting performance but also on how the

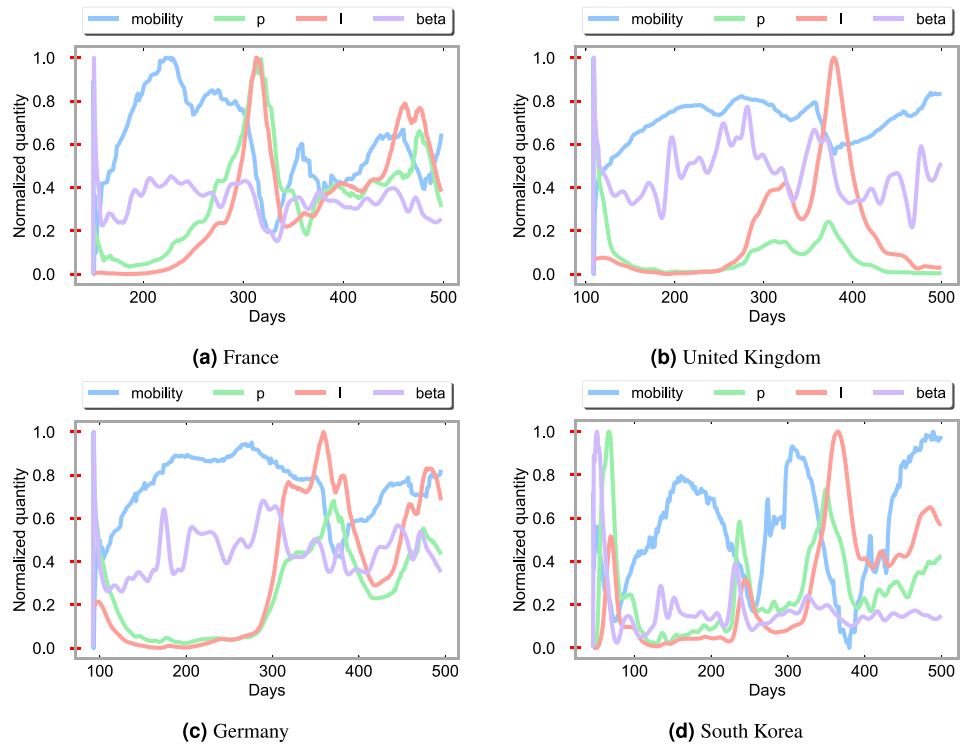


Figure 4. Plots of the normalized mobility values, positive test rate (p), infectious individual (I), and contact ratio (β) against the time (Days) for France, United Kingdom, Germany, and South Korea.

Country	β	μ	S	I	R	D
France	3.17×10^{-3}	1.02×10^{-1}	9.4×10^{-5}	4.2×10^{-2}	6.3×10^{-3}	6.3×10^{-2}
United Kingdom	3.13×10^{-3}	9.26×10^{-2}	1.7×10^{-4}	3.8×10^{-2}	1.8×10^{-3}	2.2×10^{-2}
Germany	3.11×10^{-3}	9.71×10^{-2}	5.4×10^{-5}	2.1×10^{-2}	1.1×10^{-2}	1.3×10^{-1}
Korea	6.29×10^{-2}	1.73×10^{-1}	8.0×10^{-7}	9.4×10^{-3}	2.0×10^{-3}	3.4×10^{-2}

Table 1. Relative L^2 errors with $N_T = 14$ (2 weeks) for β and μ , and for the SIRD implementation with the LSTM networks. SIRD stands for susceptible (S), infectious (I), removed (R), deceased (D) individuals. The definition of relative L^2 error is stated in (1).

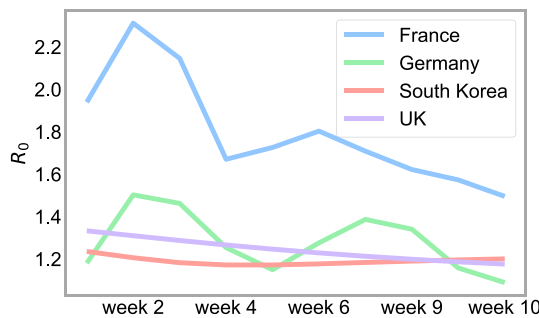


Figure 5. 10 weeks prediction of the reproduction number, R_0 , using β generated by the LSTM network for each country.

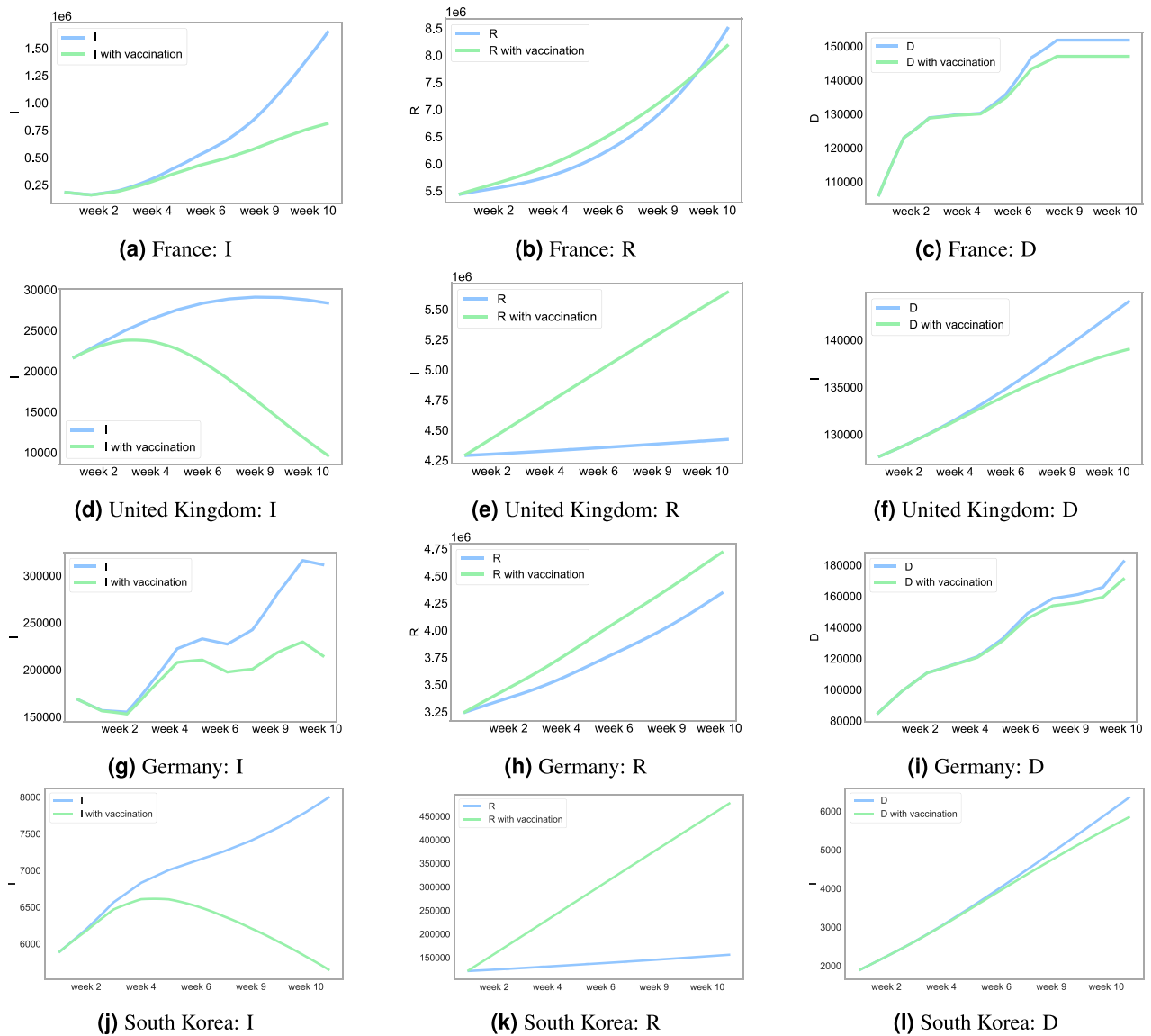


Figure 6. The SIRD prediction with the LSTM network for France, United Kingdom, Germany, and South Korea. The plots display I, R, and D prediction against the time (Days) for 10 weeks.

highly nonlinear capabilities of the neural network can be used to conduct inference on latent parameters of the SIR model.

Methodology

In this section, we explore our numerical method and prediction algorithm considered in this research. To begin, we describe the compartmental models, the SIRD equations, and the Runge–Kutta method. Then, we present the Levenberg–Marquardt algorithm. Lastly, we illustrate the combined SIRD–LSTM architecture which is the heart of our approach. We confirm that all methods were performed in accordance with the relevant guidelines and regulations.

Compartmental model: SIRD model. In this study, we represent the spread of Covid-19 using the susceptible–infected–recovered–dead (SIRD) model. Compartmental models have been used to simplify the mathematical modeling of infectious diseases^{34,35}. One of the well-known (and simplest) models is the SIR model, and many models including SIRD are derivatives of this basic form^{36–38}. The SIRD model predicts how a disease spreads, the total number infected, or the duration of an epidemic, and estimate important epidemiological parameters such as the reproductive number. Regarding the compartmental model, the population is assigned to compartments with labels:

- $S(t)$: the number of individuals susceptible of contracting the infection at time t ,
- $I(t)$: the number of individuals that are alive and infected at time t ;

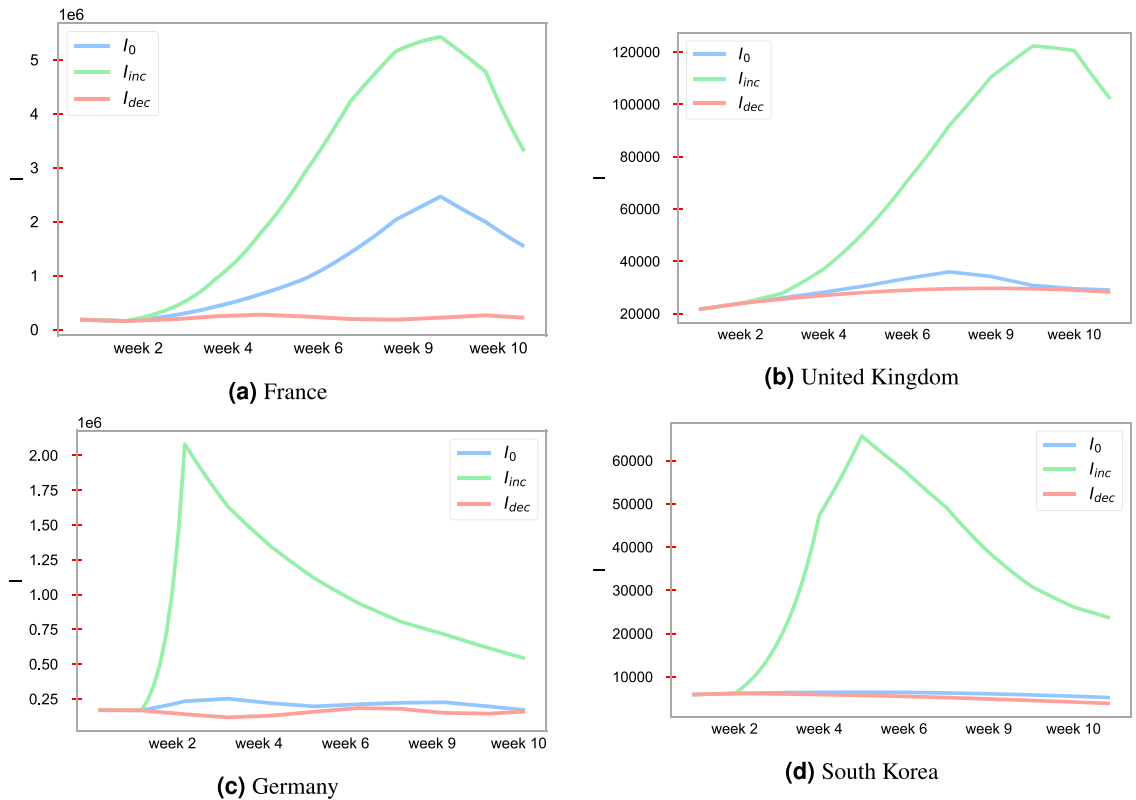


Figure 7. Forecasting the number of Covid-19 infections for France, the United Kingdom, Germany, and South Korea under 30% increased and decreased mobility to normal mobility (baseline mobility). Mobility data is real-time cell phone/mobile device location for each country collected from³². Here, I_0 , I_{inc} , and I_{dec} stand for infections with normal, 30% increased, 30% decreased mobility, respectively. The vaccination model is used for the simulations.

- $R(t)$: the cumulative number of individuals that recovered from the disease up to time t ;
- $D(t)$: the cumulative number of individuals that deceased due to the disease, up to time t .

In addition, N is the total number of people in the area at time t with $N = S(t) + I(t) + R(t)$. The SIRD model is given by the following expressions¹⁵:

$$\begin{aligned}
 \frac{dS}{dt} &= -\frac{\beta IS}{N}, \\
 \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I - \mu I, \\
 \frac{dR}{dt} &= \gamma I, \\
 \frac{dD}{dt} &= \mu I,
 \end{aligned}
 \tag{2}$$

where the parameter β , called the contact ratio, represents the effective contact rate, i.e. expected number of people infected by an infectious person, and γ is defined as recovery rate, i.e. expected number of people removed from the infected state. The ratio of β and γ is called as reproduction number, i.e. $R_0 = \beta/\gamma$. The reproduction number (R_0) shows the average number of secondary infections coming from an infected person. The parameter μ is defined as a deceased rate. We assume that the recovered subjects are no longer susceptible to infection; the number of deaths due to other reasons is neglected. Further, the region under consideration is assumed to be isolated from other regions. This is a reasonable assumption as containment measures such as travel restriction have been enforced in most countries.

By introducing the vaccination rate,

$$\nu = \frac{\text{target population}}{\text{completion time}}
 \tag{3}$$

the $S(t)$ and $R(t)$ terms can be modified for the vaccination model. We add the vaccination rate, ν , and the vaccine efficacy factor, ε , into our SIRD model to study an extended SIRD model with vaccination. For instance, $\varepsilon = 0.95$ for the Moderna and Pfizer vaccine³⁹. More precisely, we introduce a multiplier factor $\delta = (1 - \varepsilon)$.

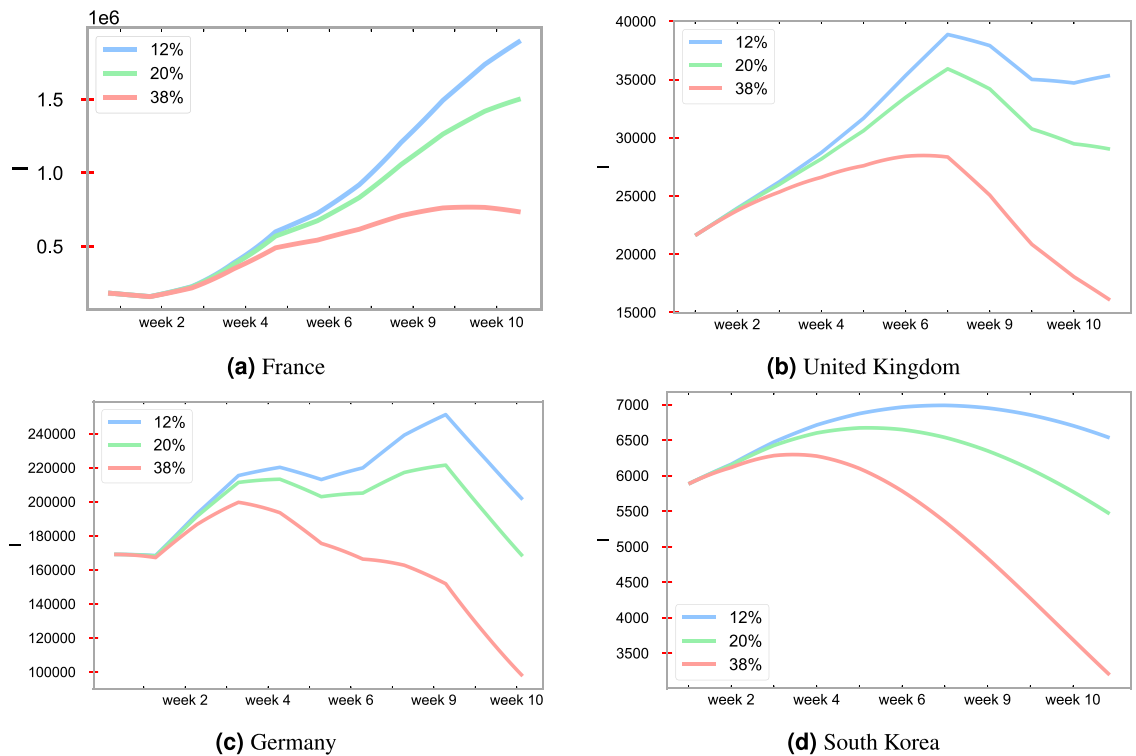


Figure 8. Forecasting of the number of Covid-19 infections for France, the United Kingdom, Germany, and South Korea under various vaccination schedules. Here, “12%”, “20%”, and “38%” mean 12%, 20%, 38% of the population is vaccinated, respectively.

We now write the following SIRD model which incorporates vaccination^{28,29}

$$\begin{aligned} \frac{dS}{dt} &= -v\delta - \frac{\beta IS}{N}, \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I - \mu I, \\ \frac{dR}{dt} &= v(1 - \delta) + \gamma I, \\ \frac{dD}{dt} &= \mu I. \end{aligned}$$

With the SIRD model, we generate a deep neural network to predict β and μ . Subsequently, the SIRD with vaccination model provides the dynamics of the vaccination with predicted parameters β and μ .

The contact rate, $\beta = \beta(t)$, and death rate, $\mu = \mu(t)$, of many acute infectious diseases varies significantly in time and frequently exhibits significant seasonal dependence^{40,41}. Epidemiological models can be used to predict contact and death rate, which are important for measuring the spread of disease. A substantial body of research predicts the contact and death rate, β and μ , of infectious diseases via the discrete compartmental model^{42–44}. The rest of this section introduces an algorithm to compute the time-dependent parameters directly from our data and the discrete SIRD model.

Levenberg–Marquardt algorithm. To estimate the contact rate, β , and the death rate, μ , we use the Levenberg–Marquardt algorithm. To apply the algorithm, we solve the SIRD equations using a numerical approximation. In the present study, we use the fourth-order Runge–Kutta methods (RK4) which give the following discrete version of the SIRD model. For simplicity, we set

$$y = \begin{pmatrix} S \\ I \\ R \\ D \end{pmatrix}, \quad F = \begin{pmatrix} -\frac{\beta IS}{N} \\ \frac{\beta IS}{N} - \gamma I - \mu I \\ \gamma I \\ \mu I \end{pmatrix},$$

then (2) can be recast

$$\frac{dy}{dt} = \mathbf{F}(t, \mathbf{y}). \quad (4)$$

The RK4 of (2) can be written as

$$\begin{aligned} \mathbf{y}_{n+1} &= \mathbf{y}_n + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \\ \mathbf{k}_1 &= \mathbf{F}(t_n, \mathbf{y}_n), \quad \mathbf{k}_2 = \mathbf{F}(t_n + h/2, \mathbf{y}_n + h\mathbf{k}_1/2), \\ \mathbf{k}_3 &= \mathbf{F}(t_n + h/2, \mathbf{y}_n + h\mathbf{k}_2/2), \quad \mathbf{k}_4 = \mathbf{F}(t_n + h, \mathbf{y}_n + h\mathbf{k}_3). \end{aligned}$$

Given a dataset $(\mathbf{y}(t))$, using the Levenberg–Marquardt algorithm, we aim to find the parameters $(\beta_n, \mu_n) := (\beta(t_n), \mu(t_n))$ of the model curve with the least-squares curve-fitting⁴⁵,

$$(\hat{\beta}, \hat{\mu}) = \arg \min_{\beta, \mu} \sum_{i=1}^m \left[\mathbf{y}_{n+1,i} - \left(\mathbf{y}_{n,i} + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) \right) \right]^2. \quad (5)$$

We note that the Covid-19 dataset for each country is obtained from the Google mobility report³².

Neural network architecture. Long short term memory networks—so-called LSTM—are variants of recurrent neural network (RNN), capable of learning long-term dependencies. They were introduced by Hochreiter and Schmidhuber⁴⁶, and are widely used in many fields such as time series prediction⁴⁷, speech recognition⁴⁸, and robot control⁴⁹ among many other applications.

Classic RNNs can keep track of arbitrary long-term dependencies in the input sequences. However, there is a computational drawback to the standard RNNs. In standard RNNs, this repeating module will have a very simple structure, such as a single layer. When training a classical RNN with back-propagation, the gradients which are back-propagated may tend to zero (vanish gradient problem) because the RNN remembers data for just a small duration of time. In other words, if we need the information after a small-time it may be reproducible, but once a lot of information is fed in, this information may get lost somewhere. This issue can be resolved by applying a variant of RNNs such as the LSTM network. The LSTMs are explicitly designed to avoid the long-term dependency problem as remembering information for long periods is practically their default behavior.

The compact forms of the LSTM with a forget gate can be described by the following system of equations:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \quad (6)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i), \quad (7)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o), \quad (8)$$

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c), \quad (9)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t, \quad (10)$$

$$h_t = o_t \circ \sigma_h(c_t), \quad (11)$$

where x_t is input vector, c_t is a memory cell, and $\{i_t, f_t, o_t\}$ denote the input, forget, and output gates, respectively; for more details, see for instance^{46,50,51}. Here, the operator \circ denotes the Hadamard product (element-wise product), and subscript t indexes the time step.

In the proposed neural network, we couple the SIRD model (2) and the LSTM network. By the Levenberg–Marquardt algorithm, predictions on β and μ are made by curve-fitting methods. With this, input data consists of $x_t = \{S_t, I_t, R_t, D_t, p_t, m_t, \beta_t, \mu_t\}$ where p_t is a positive rate (the percentage of all coronavirus tests performed that are actually positive) and m_t is a mobility trend at time t obtained from Google's mobility report. The reports chart movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential. The parameters β_t and μ_t are predicted by the Levenberg–Marquardt algorithm. The output of the LSTM network is (β_{t+1}, μ_{t+1}) . When implementing cost functions, we apply a mean-squared forecasting error metric as well as mean-absolute percentage errors.

The network structure and activation of each hidden unit in the hidden layers are determined by the neurons in the previous layers. The activity of each layer is given by the nonlinear activation function σ such as a sigmoid function or ReLU function. The final output of the coupled model is obtained by combining the network output of confirmed cases with the SIR model forecast. More precisely, the collective dataset generated from the SIRD model is used as inputs for the LSTM whose outputs provide the parameters β and μ for the next time period. By predicting the parameters, we are able to solve the SIRD model, which gives $\{S, I, R, D\}$ for the next time period. The coupled models given in Fig. 9 illustrate the Neural LSTM–SIRD architecture. The network architecture we use is an LSTM with ReLU activation functions, and is trained by using Adam optimizer with a mean-squared error loss function. The model is not constrained to a particular setup and we could search over various hyper-parameters to manipulate the number of neurons, with similar results.

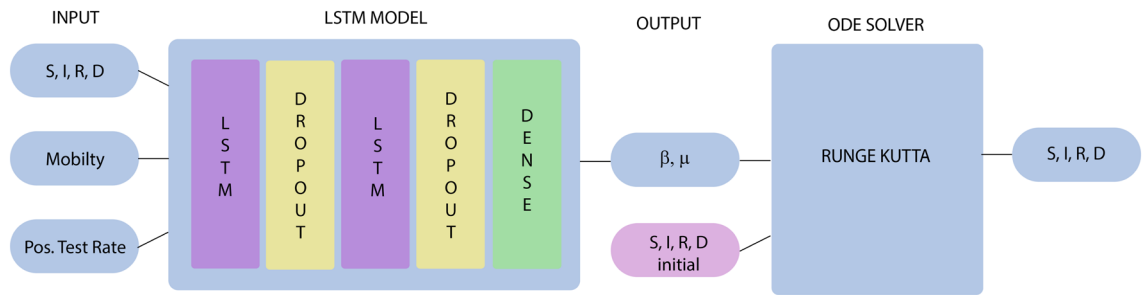


Figure 9. A description of the combined SIRD–LSTM model structure with Covid-19 community mobility (mobility) and positive test rate (Pos. Test Rate) to generate forecasts of time varying parameters (β , μ). The ODE solver based on the Runge–Kutta fourth order method makes use of the predicted parameters in the numerical discretization.

Data. We collected data from the following sources:

- Covid-19 data repository by the center for systems science and engineering (csse) at Johns Hopkins University, <https://github.com/CSSEGISandData/COVID-19> (see⁵²).
- Our World In Data, <https://github.com/owid/covid-19-data/tree/master/public/data> (see⁵³).
- Google Mobility Report, <https://www.google.com/covid19/mobility/>.

Received: 6 June 2021; Accepted: 10 February 2022

Published online: 22 February 2022

References

1. Who coronavirus (Covid-19) dashboard.
2. Tian, H. *et al.* An investigation of transmission control measures during the first 50 days of the Covid-19 epidemic in China. *Science* **368**, 638–642. <https://doi.org/10.1126/science.abb6105> (2020).
3. Yan, Y. *et al.* Measuring voluntary and policy-induced social distancing behavior during the Covid-19 pandemic. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.2008814118> (2021).
4. Rothman, K., Greenland, S. & Lash, T. *Modern Epidemiology* (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2015).
5. Sen, D. & Sen, D. Use of a modified SIRD model to analyze Covid-19 data. *Ind. Eng. Chem. Res.* **60**, 4251–4260. <https://doi.org/10.1021/acs.iecr.0c04754> (2021).
6. Tolles, J. & Luong, T. Modeling epidemics with compartmental models. *JAMA* **323**, 2515–2516. <https://doi.org/10.1001/jama.2020.8420> (2020).
7. Anderson, R. M., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T. D. How will country-based mitigation measures influence the course of the Covid-19 epidemic?. *The Lancet* **395**, 931–934. [https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5) (2020).
8. Keeling, M. J. & Eames, K. T. Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307. <https://doi.org/10.1098/rsif.2005.0051> (2005).
9. Nadini, M., Rizzo, A. & Porfiri, M. Epidemic spreading in temporal and adaptive networks with static backbone. *IEEE Trans. Netw. Sci. Eng.* **7**, 549–561. <https://doi.org/10.1109/TNSE.2018.2885483> (2020).
10. Pastore, Y., Pionti, A., Gomes, M. F. D. C., Samay, N., Perra, N. & Vespignani, A. The infection tree of global epidemics. *Netw. Sci.* **2**, 132–137. <https://doi.org/10.1017/nws.2014.5> (2014).
11. Yang, Z. *et al.* Modified SEIR and AI prediction of the epidemics trend of Covid-19 in China under public health interventions. *J. Thorac. Dis.* **12**, 165 (2020).
12. Giordano, G. *et al.* Modelling the Covid-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med.* **26**, 855–860. <https://doi.org/10.1038/s41591-020-0883-7> (2020).
13. Sharma, S., Volpert, V. & Banerjee, M. Extended SEIQR type model for Covid-19 epidemic and data analysis. *Math. Biosci. Eng. MBE*. <https://doi.org/10.3934/mbe.2020386> (2020).
14. Oliveira, T. d. P. & Moral, R. d. A. Global short-term forecasting of Covid-19 cases. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-87230-x> (2021).
15. Bailey, N. *The Mathematical Theory of Infectious Diseases and Its Applications*. Mathematics in Medicine Series (Griffin, 1975).
16. Anderson, R. M. Discussion: The Kermack–McKendrick epidemic threshold theorem. *Bull. Math. Biol.* <https://doi.org/10.1007/BF02464422> (1991).
17. Bjornstad, O. N., Finkenstadt, B. F. & Grenfell, B. T. Dynamics of measles epidemics: estimating scaling of transmission rates using a time series sir model. *Ecol. Monogr.* **72**, 169–184 (2002).
18. Gers, F. A., Eck, D. & Schmidhuber, J. Applying LSTM to time series predictable through time-window approaches. In Tagliaferri, R. & Marinaro, M. (eds.) *Neural Nets WIRN Vietri-01*, 193–200 (Springer, London, 2002).
19. James, N., Menzies, M. Efficiency of communities and financial markets during the 2020 pandemic. *Chaos: An Interdiscipl. J. Nonlinear Sci.* **31**, 083116. <https://doi.org/10.1063/5.0054493> (2021).
20. Yilmazkuday, H. Stay-at-home works to fight against Covid-19: International evidence from google mobility data. *J. Hum. Behav. Soc. Environ.* **31**, 210–220. <https://doi.org/10.1080/10911359.2020.1845903> (2021).
21. Chang, S. *et al.* Mobility network models of Covid-19 explain inequities and inform reopening. *Nature* **589**, 82–87. <https://doi.org/10.1038/s41586-020-2923-3> (2021).
22. Schlosser, F. *et al.* Covid-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proc. Natl. Acad. Sci.* **117**, 32883–32890. <https://doi.org/10.1073/pnas.2012326117> (2020).
23. Soares, N. *et al.* Sirnet: Understanding social distancing measures with hybrid neural network model for Covid-19 infectious spread (2020).

24. Hisaka, A. *et al.* Global comparison of changes in the number of test-positive cases and deaths by coronavirus infection (Covid-19) in the world. *J. Clin. Med.* <https://doi.org/10.3390/jcm9061904> (2020).
25. Wang, H., Ghosh, A., Ding, J., Sarkar, R. & Gao, J. Heterogeneous interventions reduce the spread of Covid-19 in simulations on real mobility data. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-87034-z> (2021).
26. Dashbali, M. & Mirzaie, M. A compartmental model that predicts the effect of social distancing and vaccination on controlling Covid-19. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-86873-0> (2021).
27. Shirin, A., Lin, Y. T. & Sorrentino, F. Data-driven optimized control of the Covid-19 epidemics. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-85496-9> (2021).
28. Yongzhen, P., Shuping, L., Changguo, L. & Chen, S. The effect of constant and pulse vaccination on an sir epidemic model with infectious period. *Appl. Math. Model.* **35**, 3866–3878. <https://doi.org/10.1016/j.apm.2011.02.012> (2011).
29. Wong, W., Juwono, F. & Chua, T. SIR simulation of Covid-19 pandemic in Malaysia: Will the vaccination program be effective? (2021).
30. Goldberg, T. L. & Wolfe, N. D. Infectious diseases in primates: Behavior, ecology and evolution oxford series in ecology and evolution by Charles L. Nunn and Sonia Altizer. *Q. Rev. Biol.* **82**, 289. <https://doi.org/10.1086/523174> (2007).
31. UK-Government. Covid-19 vaccines: further evidence of success. <https://www.gov.uk/government/news/covid-19-vaccines-further-evidence-of-success> (10 May 2021).
32. Google-Report. Covid-19 community mobility reports. <https://www.google.com/COVID19/mobility/>. Accessed on 05/10/2021.
33. Kim, Y., Denton, C., Hoang, L. & Rush, A. M. Structured attention networks. In: 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings (OpenReview.net, 2017).
34. Ross, R. An application of the theory of probabilities to the study of a priori pathometry Part I. *Proc. Roy. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Charact.* **92**, 204–230. <https://doi.org/10.1098/rspa.1916.0007> (1916).
35. Kermack, W. O., McKendrick, A. G. & Walker, G. T. A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Charact.* **115**, 700–721. <https://doi.org/10.1098/rspa.1927.0118> (1927).
36. Harko, T., Lobo, F. S. & Mak, M. Exact analytical solutions of the susceptible–infected–recovered sir epidemic model and of the sir model with equal death and birth rates. *Appl. Math. Comput.* **236**, 184–194. <https://doi.org/10.1016/j.amc.2014.03.030> (2014).
37. Hethcote, H. W. The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653. <https://doi.org/10.1137/S0036144500371907> (2000).
38. Parshani, R., Carmi, S. & Havlin, S. Epidemic threshold for the susceptible–infectious–susceptible model on random networks. *Phys. Rev. Lett.* **104**, 258701. <https://doi.org/10.1103/PhysRevLett.104.258701> (2010).
39. Mw, T., SM, O. & Self W. H., E. A. Effectiveness of Pfizer–BioNTech and Moderna vaccines against Covid-19 among hospitalized adults aged larger than 65 years—United States. *MMWR Morb. Mortal Wkly. Rep.* **70**, 674–679 <https://doi.org/10.15585/mmwr.mm7018e1> (2021).
40. Becker, N. G. & Britton, T. Statistical studies of infectious disease incidence. *J. Roy. Stat. Soc. Ser. B (Stat. Methodol.)* **61**, 287–307 (1999).
41. Capistrán, M. A., Moreles, M. A. & Lara, B. Parameter estimation of some epidemic models. The case of recurrent epidemics caused by respiratory syncytial virus. *Bull. Math. Biol.* **71**, 1890. <https://doi.org/10.1007/s11538-009-9429-3> (2009).
42. Ponciano, J. M. & Capistran, M. A. First principles modeling of nonlinear incidence rates in seasonal epidemics. *PLOS Comput. Biol.* **7**, 1–14 (2011).
43. He, D., Ionides, E. L. & King, A. A. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J. Roy. Soc. Interface* **7**, 271–283 (2010).
44. Cauchemez, S. & Ferguson, N. M. Likelihood-based estimation of continuous-time epidemic models from time-series data: Application to measles transmission in London. *J. Roy. Soc. Interface* **5**, 885–897 (2008).
45. Nocedal, J. & Wright, S. J. *Numerical optimization*, 2nd edn. Springer Series in Operations Research and Financial Engineering (Springer, New York, 2006).
46. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735> (1997).
47. Vermaak, J. & Botha, E. C. Recurrent neural networks for short-term load forecasting. *IEEE Trans. Power Syst.* **13**, 126–132. <https://doi.org/10.1109/59.651623> (1998).
48. Alex, G. & Jurgen, S. Framework phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **18**, 602–610 (2005).
49. Mayer, H. *et al.* A system for robotic heart surgery that learns to tie knots using recurrent neural networks, in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 543–548. <https://doi.org/10.1109/IROS.2006.282190> (2006).
50. Gers, F. A., Schmidhuber, J. & Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Comput.* **12**, 2451–2471 (2000).
51. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, London, 2016).
52. E, D., H, D. & L., G. An interactive web-based dashboard to track Covid-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
53. Hasell, J. *et al.* A cross-country database of Covid-19 testing. *Sci. Data.* <https://doi.org/10.1038/s41597-020-00688-8> (2020).

Acknowledgements

The work of Y. Hong was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A2C1093579). The work of Nelli Vardanyan was supported by the Richter Scholar Program and the work of Said Omer Sadat was supported by the James Rocco Scholarship.

Author contributions

A.B., W.C., and Y.H. conceived the experiments, and A.B. and Y.H. conducted the experiments. O.S. and N.V. collected data. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022