NCSP'21
RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing 2021
March 1-2, 2021

Simulation of Infectious Number of COVID-19 in JAPAN by Using SIR Model with Neural Networks

Aika Ohno, Yoko Uwate, and Yoshifumi Nishio

Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima 770-8506, Japan
Email: {a-ohno, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract

In this study, we propose a model to simulate infectious number of COVID-19 in Japan by using a mathematical model that is called SIR model. We use time variables for SIR model, and Echo State Networks (ESN) is applied to simulate the parameters of SIR model. Then, we use the simulated parameters, and infectious number in Japan is estimated.

1. Introduction

COVID-19 is confirmed for the first time in the city of Wuhan in December 2019. As everyone knows, COVID-19 spread all over the world, and still now on confirmed cases are increasing day by day. According to Jonhs Hopskins University's report, more than 96,000,000 infected cases and 2,000,000 deaths had been confirmed globally until January 20th 2021 [1]. Then, many cities and countries are locked down to prevent spreading COVID-19, and some of cities are enforced lock down before pandemic or outbreak situation. As a result, they could prevent pandemic. Therefore, it is important to predict infectious numbers, and to know how the infectious number would go on.

Mathematical models for infectious diseases such as SIR model and SEIR model have been used for the simulation of infectious diseases. SIR model and SEIR model are used for Severe Acute Respiratory Syndrome and seasonal influenza as well [2],[3]. Neural networks are also applied for the simulation of COVID-19 [4],[5].

Echo State Networks (ESN) is one kind of neural networks, and ESN is used for time-series data computing.

In this study, we simulate infectious number of COVID-19 in Japan by using SIR model. Then, ESN is used to train and test the parameters of SIR model, and we aim at simulating the infectious graph that has more realistic waves.

2. SIR model

SIR model is a popular mathematical model for infectious diseases. SIR is the initial letters of susceptible, infectious and removed. Susceptible indicates the statement of the people who do not have antibodies of that infection. In this study, as to COVID-19, we regard all of the people at the first point as susceptible. Infectious means the statement of the people who have infectious capacity. Removed means the statement of the people who recover from that infection and also people who pass away. In SIR model, susceptible people move to the statement of infectious with the infectious rate α, and the infectious people move to the statement of removed with the removed rate γ. Figure 1 shows the statement structure of SIR model.

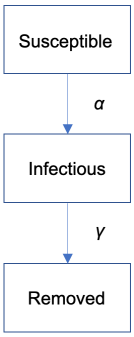

Figure 1: The statement structure of SIR model.

This SIR model can be expressed by the following differential equations. $S(t)$, $I(t)$ and $R(t)$ mean the population of susceptible, the population of infectious and the population of removed at time t, respectively. Equation (1) shows the differential equations of SIR model.

$$\begin{cases} \dfrac{dS}{dt} = -\alpha S(t)I(t) \\ \dfrac{dI}{dt} = \alpha S(t)I(t) - \gamma I(t) \\ \dfrac{dR}{dt} = \gamma I(t) \end{cases} \tag{1}$$

In SIR model, all of the population are divided in these three statements which are susceptible, infectious and removed. Then, whole population N is the addition of the three statements as shown in Eq. (2), and we regard N as constant.

$$N = S(t) + I(t) + R(t) \tag{2}$$

Then, Eq. (1) is solved by the Runge-Kutta method. Figure 2 shows a typical graph of SIR model.

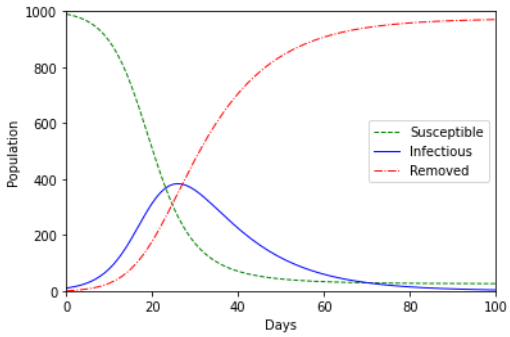

Figure 2: A graph of SIR model.

As shown in Fig. 2, the infectious graph of SIR model has only one peak. SIR model has the advantage that is quick simulation and the simple structure. However, SIR model has the disadvantage that SIR model cannot express multiple waves with constant parameters. Therefore, we aim at expressing multiple peaks in infectious simulation graph by applying ESN which is one kind of neural networks and which is introduced in the following section.

3. Echo State Networks

ESN is one kind of reservoir computing, and it can handle time-series data [6]. ESN has the feature of quick learning. Therefore, the advantage of SIR model would not be lost by using this network. ESN has reservoir layer between input and output. Figure 3 shows the structure of ESN.

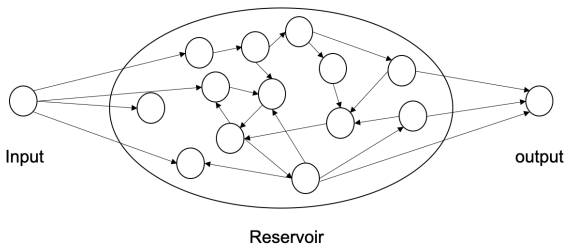

Figure 3: The structure of Echo State Networks.

In the reservoir layer there is a lot of reservoir nodes, and the reservoir nodes are connected randomly. The input weights and reservoir weights are not updated, only the output weights are updated.

$x(n)$ is a vector of reservoir neuron activation states at time step n, and n is the discrete time. Tanh function is used as the activation function as shown in Eq. (3). $u(n)$ is the input data, a is the leaking rate and W_{in} is the input weight.

The update equations are shown in Eq. (3).

$$\begin{cases} \tilde{x}(n) = tanh(W_{in}[1; u(n)] + Wx(n-1) \\ x(n) = (1-a)x(n-1) + a\tilde{x}(n) \end{cases} \tag{3}$$

Root Mean Squared Percentage Error (MSE) is used for validation. N, $Y_{simulated}$ and Y^t means the number of the data, the simulation value and the target value, respectively.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Y^t - Y_{simulated})^2 \tag{4}$$

W_{out} is the output weights, reg is the regularization coefficient, and I is the identity matrix in Eq. (5). X indicates $[1;u(n);x(n)]$.

$$W_{out} = Y^t X^T (XX^T + regI)^{-1} \tag{5}$$

In ESN, W_{out} is learned and updated to minimize MSE. In this study, ESN is used for simulating of the parameters $\alpha(t)$ and $\gamma(t)$.

4. Data

We use open data from Ministry of Health, Labour and Welfare of Japanese government [7]. From the open data source, positive cases, recovered cases and death toll in Japan are obtained.

Then, $S(t)$, $I(t)$ and $R(t)$ are calculated by the following equations. $P(t)$, $R_e(t)$ and $D(t)$ means the reported positive cases, reported recovered cases and reported death toll at time t.

$$\begin{cases} S(t) = N - I(t) - R(t) \\ I(t) = \sum_{i=1}^{t} P(t) - \sum_{i=1}^{t} R(t) \\ R(t) = \sum_{i=1}^{t} R_e(t) - \sum_{i=1}^{t} D(t) \end{cases} \tag{6}$$

In this study, N is 126,000,000 that is the whole population of Japan.

5. Proposed Method

First, the parameters $\alpha(t)$ and $\gamma(t)$ are calculated from SIR model differential equations by using difference equations method.

$$\begin{cases} \alpha(t) = \dfrac{S(t) - S(t+1)}{S(t)I(t)} \\ \gamma(t) = \dfrac{R(t+1) - R(t)}{I(t)} \end{cases} \tag{7}$$

Second, the parameters $\alpha(t)$ and $\gamma(t)$ are learned and tested by ESN. It is iterated 30 times with random number for initial weight, and average output is used.

Third, $I(t)$ is calculated by using simulated parameters $\alpha(t)$ and $\gamma(t)$.

$$\begin{cases} S(t+1) = S(t) - \alpha(t)S(t)I(t) \\ R(t+1) = \gamma(t)I(t) + R(t) \\ I(t+1) = N - S(t+1) - R(t+1) \end{cases} \tag{8}$$

As to real data, we use first 100 data for training and the next 100 data for testing.

6. Simulation Results

The following parameters are used for ESN. The size of reservoir is 50, the leaking rate is 0.03, the spectral radius coefficient is 0.999 and the regularization coefficient is 0.5. As a result, MSE of $\alpha(t)$ is 0.0023511. Figure 4 shows the test part of the real data and the simulation result of $\alpha(t)$.

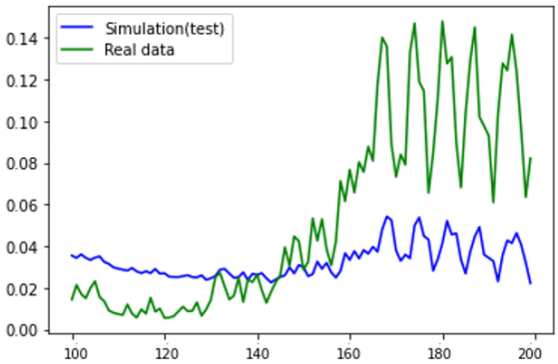

Figure 4: The real data and the test result of $\alpha(t)$.

The test result graph seems not to be following with the real data graph enough. Especially, after 160th days there is difference between simulation values and real values. However, the shape of the graph seems to be following.

As to the simulation of $\gamma(t)$, the following parameters are used for ESN. The size of reservoir is 50, the leaking rate is 0.8, the spectral radius coefficient is 0.999 and the regularization coefficient is 0.005. As a result, MSE of $\gamma(t)$ is 0.0030737. Figure 5 shows the test part of the real data and the simulation result of $\gamma(t)$.

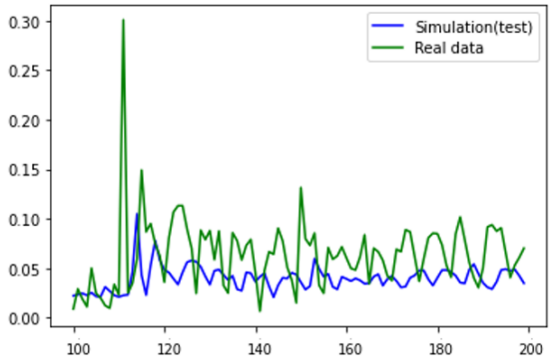

Figure 5: The real data and the test result of $\gamma(t)$.

The test result graph seems to be following with the real data graph. The serrate waveform which is shown in the real graph is expressed in the test result graph.

Figure 6 shows the simulation result of infectious number by using test result of $\alpha(t)$ and $\gamma(t)$.

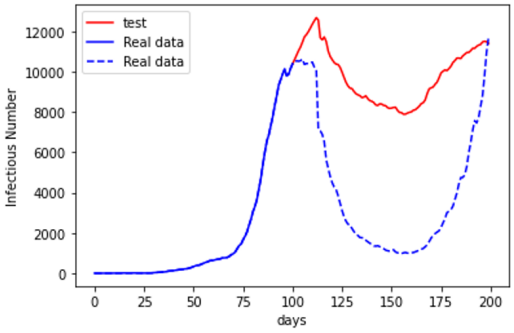

Figure 6: The simulation result of infectious number.

The waveform is not like a smooth curb, it is serrate waveform. These features are not shown in the typical infectious graph of SIR model. Moreover, the valley of the simulation test graph is the same with the valley of real data. However, the number of simulation test result is not following with the real data.

7. Additional Simulation

As an additional simulation, we simulate the infectious numbers by using real data and test result mixed.

Figure 7 shows the simulation result of infectious number by using the real data of $\alpha(t)$ and the test result of $\gamma(t)$.

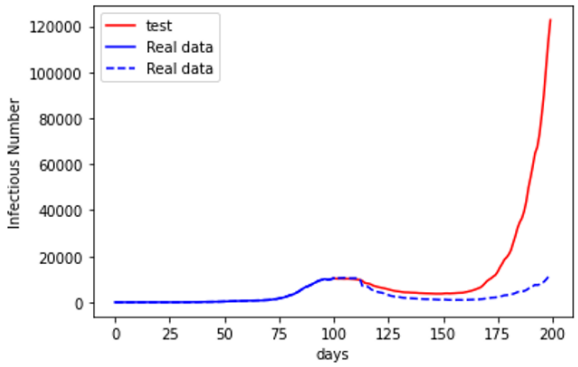

Figure 7: The simulation result using the real data of $\alpha(t)$ and the test result of $\gamma(t)$.

Obviously, the simulation result is far different from the real infectious number. Then, Figure 8 shows the simulation result of infectious number by using the test result of $\alpha(t)$ and the real data of $\gamma(t)$.

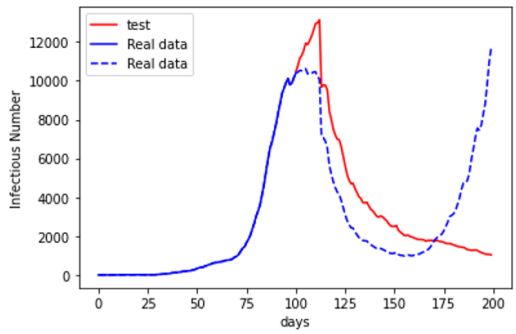

Figure 8: The simulation result using the test result of $\alpha(t)$ and the real data of $\gamma(t)$.

The simulation result graph is following better with the graph of the real infectious number compared with Fig. 6 and 7.

From these results, we can know the simulation of $\gamma(t)$ is more important and influential than the simulation of $\alpha(t)$. We consider that it is because $\gamma(t)$ has the larger digits of 0.1 than the digits of $\alpha(t)$ of 1.0×10^{-9}. Therefore, a simulation of $\gamma(t)$ is more important to have more accurate simulation of infectious number.

8. Conclusion

In this study, the infectious number of COVID-19 was simulated with time variables $\alpha(t)$ and $\gamma(t)$ by using SIR model method and ESN to simulate the parameters $\alpha(t)$ and $\gamma(t)$. Then, we could express serrate waveform in the test part of infectious number simulation. From the additional simulation result, we could show the importance of the simulation of $\gamma(t)$.

For future works, we need to improve the test accuracy of the parameters $\alpha(t)$ and $\gamma(t)$. Especially, we need to improve $\gamma(t)$ simulation. In this study, though we used 100 data for training and other 100 data for testing, these data size might be too small data size for ESN. However, to increase the data size is not good solution. We will keep using small data size to predict for unknown infectious diseases besides COVID-19. To improve $\gamma(t)$ simulation, we will try to use other neural networks such as other recurrent neural networks as well.

Acknowledgement

This work was supported by Center for Design-Oriented AI Education and Research Grant-in-Aid for Students.

References

[1] JOHNS HOPSKINS UNIVERSITY OF MEDICINE, “COVID-19 Map”, https://coronavirus.jhu.edu/map.html, 2021-1-20.

[2] Ng, Tuen Wai, Gabriel Turinici and Antoine Danchin, “A double epidemic model for the SARS propagation”, BMC Infectious Diseases 3.1 (2003): 19.

[3] Osthus, Dave, et al, “Forecasting seasonal influenza with a state-space SIR model”, The annals of applied statistics 11.1 (2017): 202.

[4] Zeroual, Abdelhafid, et al. “Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study.” Chaos, Solitons Fractals 140 (2020): 110121.

[5] Wieczorek, Michał, Jakub Siłka, and Marcin Woźniak. “Neural network powered COVID-19 spread forecasting model.” Chaos, Solitons Fractals 140 (2020): 110203.

[6] Lukoševičius, Mantas, “A practical guide to applying echo state networks”, Neural networks: Tricks of the trade. Springer, Berlin, Heidelberg, 2012.

[7] Ministry of Health, Labour and Welfare, “Open data”, https://www.mhlw.go.jp/stf/covid-19/open-data.html, 2020-10-03.