

Spread of Covid-19: a Study Case of Honduras, Forecasting with Logistic Model and SIR Model*

Danny Guerrero¹

Abstract

In this compilation, the logistic model and the SIR model are tested to forecast the spread of Covid-19 virus in Honduras by using the software Matlab. The mathematics topics involved are presented in a friendly way to facilitate their reading and to identify some possible didactic practices derived from their application. Certainly, some graphs, data arrangements and comments about the results obtained are given. The forecasts in this document depended on the handling of the data provided by the Hondurean's authorities.

Keywords: Logistic model, SIR model, Covid-19, forecasting, math didactics.

1. Introduction

According to the World Health Organization (WHO) [9], coronaviruses are a large family of viruses known to cause diseases ranging from the common cold to more serious diseases such as Middle Eastern Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). They are positive-stranded RNA viruses, with a crown-like appearance under an electron microscope.

The first official date of the new coronavirus disease 2019 (COVID-19) is December 31, in which the Chinese health authorities had given news of these unusual cases. At the beginning of January 2020, the city Wuham had found dozens of cases and hundreds of people were under observation. In fact, the first

*This work was not economic supported by any institution.

Email address: danny.guerrero@upnfm.edu.hn (Danny Guerrero)

¹Departamento de Ciencias Matemáticas, Universidad Pedagógica Nacional Francisco Morazán, Tegucigalpa, Honduras.

investigations showed that the infected were frequent visitors to the Huanan Seafood Wholesale Market in the city, which has been closed since January 1, 2020, hence the hypothesis that the contagion may have been caused by some animal products sold in the market.

On January 21, the WHO announced that the new coronavirus, probably passed from the animal to the human being (a leap in species), is also transmitted by human contact. Since that date, the virus started spreading uncontrollably to other countries.

The first confirmed case of Covid-19 in Honduras occurred on March 11, 2020, see [4]. Until the date, the National Risk Management System (SINAGER), through a national radio and television network, communicate the daily news of confirmed, dead and recovered cases at the national level.

In this compilation, the aim is to forecast the spread of the Covid-19 in Honduras, by testing the logistic and susceptible-infected-recovered (SIR) mathematical models of evolution of a pandemic. All the simulation will be run in Matlab, Version 2019a. Codes are taken from [3, 6].

Two details must be considered essential when fixing attention on a mathematical model:

1. The forecasts are as good as the data available, so robust forecasts depend on the data.
2. Predicting the evolution of Covid-19, including the number of undetected cases, is difficult. However, in a sense, the confirmed data are important to appreciate the challenges to be face and to identify the progress of the spread of the virus. The risk is to underestimate the actual scale of the problem.

After this short introduction to the Covid-19 pandemic and its forecast of spread, in Sections 2 and 3, respectively, a mathematical description of the logistic model and the SIR model will be presented in detail from a didactic point of view, with the intention of being able to speak about them as math topics in university level classes that adapt to the contents involved. In Section

4, the results obtained from the simulations are presented.

Some strategies could be suggested from this work, for the moment it just be said that the spread of the virus can be controlled to the extend that population adapt to its presence, keep washing their hands and try to stay at home. Deeper conclusions and comments are given in the last Section 5.

2. Logistic Model

Most of the basic mathematical models of growth of a population phenomenon, the growth rate is proportional to the size of the population, generating a model that behaves exponentially. A model of this form is unreal to simulate the spread of Covid-19, because the environment imposes conditions and limits on its spread.

A more accurate model for predicting the behaviour of Covid-19 infections is the logistic model. Let C be the number of confirmed infected cases, the logistic model says that the relative growth rate C'/C decreases when C is approaching the final size of the epidemic. The corresponding equation is the so called logistic differential equation:

$$\frac{dC}{dt} = kC \left(1 - \frac{C}{L}\right), \quad (1)$$

donde:

- C is the accumulated number of cases, the logistic curve.
- $k > 0$ is the infection rate or steepness of the curve C .
- $L > 0$ is the final epidemic size, the curve's maximum value.

This equation can be solved by separation of variables as

$$\int \frac{dC}{C(1 - \frac{C}{L})} = \int k dt.$$

To solve the left hand, partial sums are used to obtain

$$C \left(1 - \frac{C}{L}\right) = \frac{L}{C(L - C)} = \frac{1}{C} + \frac{1}{L - C}.$$

Solving the equation:

$$\begin{aligned}\int \frac{dC}{C} + \int \frac{dC}{L-C} &= \int k dt, \\ \ln|C| - \ln|L-C| &= kt + B, \\ \ln\left|\frac{L-C}{C}\right| &= -kt - B, \\ \left|\frac{L-C}{C}\right| &= e^{-kt-B}, \\ \frac{L-C}{C} &= Ae^{-kt},\end{aligned}$$

where B is an integration constant and $A = \pm e^{-C}$. From the last equality in the previous equation, using the initial condition $C(0) = C_0 > 0$, it is obtained that $A = \frac{L-C_0}{C_0}$. Finally, solving for C :

$$C = \frac{L}{1 + Ae^{-kt}}.$$

Some details to consider are:

- At the beginning of the contagious phase the curve C behaves as the natural growth (exponential) $C \approx C_0 e^{kt}$.
- When $t \rightarrow \infty$, the number of cases tends to the Weibull function:

$$C \approx L \left(1 - e^{-k(t-t_0)}\right).$$

- The growth rate $\frac{dC}{dt}$ reaches its maximum when its derivative equals 0, that is, $\frac{d^2C}{dt^2} = 0$.
- From the last condition, the maximum of the growth rate (peak) occurs when $t_p = \frac{\ln A}{k}$, giving a number of cases equals to $C_p = \frac{L}{2}$ and a growth rate $\left(\frac{dC}{dt}\right)_p = \frac{kL}{4}$.
- A common question is to find the doubling time, this can be done by the reader and it should be found:

$$\text{Doubling time} = \frac{\ln 2}{k} - \frac{1}{k} \ln \left(\frac{1}{A} - e^{-kt}\right) - t.$$

For the implementation of the logistic model to predict the spread of the Covid-19 in Honduras, it will be used the technique given by Milan Batista in [1]. For the practical calculation of the parameters L , k and A , it is used least square fit with the MATLAB functions *lsqcurvefit* and *fitnlm*. The implemented Matlab code was taken from [3].

3. Susceptible-Infected-Recovered (SIR) Model

The SIR model is one of the simplest compartmental models [8], and many models are derivatives of this basic form. The model consists of three compartments: S for the number of susceptible, I for the number of infectious, and R for the number of recovered or deceased (or immune) individuals. This model is reasonably predictive for infectious diseases which are transmitted from human to human, and where recovery confers lasting resistance, such as Covid-19. The SIR system can be expressed by the following set of ordinary differential equations:

$$\frac{dS}{dt} = -\frac{\beta IS}{N}, \quad (2)$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I, \quad (3)$$

$$\frac{dR}{dt} = \gamma I, \quad (4)$$

where $N = S + I + R$, β and γ correspond to the transmission and recovery rates respectively.

This system is non-linear, however it is possible to derive its analytic solution in implicit form, see [5].

Two observations can be done at this moment:

- Firstly, note that from $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$, (by adding equations (2)-(4)) it follows that: $S(t) + I(t) + R(t) = N$, expressing in mathematical terms the constancy of population N . Note that the previous relationship implies that one need only study the equation for two of the three variables.

- Secondly, note that the dynamics of the infectious class depends on the following ratio $\mathfrak{R}_0 = \frac{\beta}{\gamma}$, the so-called basic reproduction number (also called basic reproduction ratio). This ratio is derived as the expected number of new infections (these new infections are sometimes called secondary infections) from a single infection in a population where all subjects are susceptible. This ratio \mathfrak{R}_0 biologically means that if this number is less than or equal to 1 the disease goes extinct, whereas if this number is greater than 1 the disease will remain permanently endemic in the population.

The implementation of the SIR model will be done following the description in [2] and using the Matlab code in [6]. The description of the implementation of the model it is summarized as follows.

The initial conditions are $S(0) = S_0$, $I(0) = I_0$ and $R(0) = R_0$ (note that this is different from the basic reproduction ratio \mathfrak{R}_0). Eliminating I from Equations (2) and (4) yields

$$S = S_0 \exp \left[-\frac{\beta}{\gamma N} (R - R_0) \right].$$

From this equation, note that when $t \rightarrow \infty$, the number of susceptible individuals S_∞ is

$$S_\infty = S_0 \exp \left[-\frac{\beta}{\gamma N} (R_\infty - R_0) \right], \quad (5)$$

where R_∞ is the final number of recovered individuals. Since the final number of infected individuals is supposed to be 0, recalling that $N = S + I + R$, hence $N = S_\infty + R_\infty$. From this and (5), the equation for R_∞ is

$$R_\infty = N - S_0 \exp \left[-\frac{\beta}{\gamma N} (R_\infty - R_0) \right].$$

In order to use the model, the parameters β , γ and the initial values S_0 and I_0 must be estimated from available data (for example, $R_0 = 0$ and $C_0 = C_1$).

Now the available data is a time series of the total number of cases C , i.e., $C = I + R$. We can estimate the parameters and initial values by minimizing the difference between the actual and predicted number of cases, i.e., by minimizing

$$\|C_t - \hat{C}_t(\beta, \gamma, S_0)\|^2,$$

where $C_t = (C_1, C_2, \dots, C_n)$ are given number of cases in times C_1, C_2, \dots, C_n and $\hat{C}_t = (\hat{C}_1, \hat{C}_2, \dots, \hat{C}_n)$ are corresponding estimates calculated by the model. To estimate the parameters is used the MATLAB's function *fminsearch*. To solve the model equation is used the MATLAB's function *ode45*.

There exists a lot of variations of the SIR model, for instance, the Susceptible-Expose-Infected-Recovered model (SEIR) which assumes that the virus-infected person is not infectious during the incubation period and that the infected person did not take isolation measures during the illness, some works can already be found using this model, for example see [7].

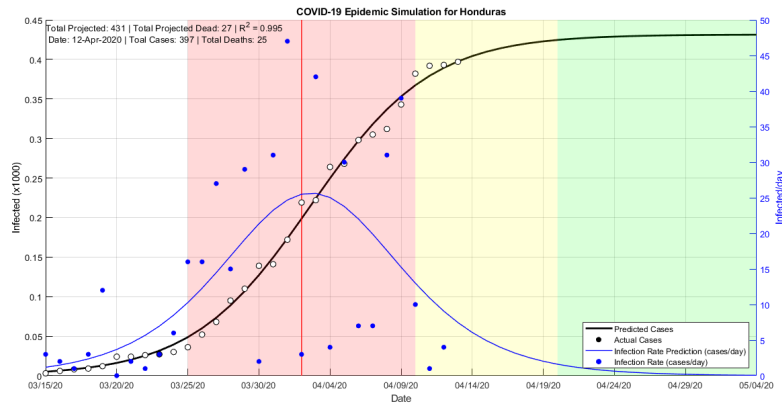
4. Forecasting Covid-19 in Honduras

The results will be presented in this section with data confirmed until April 12, 2020. Keep in mind the following recommendations to interpret the figures:

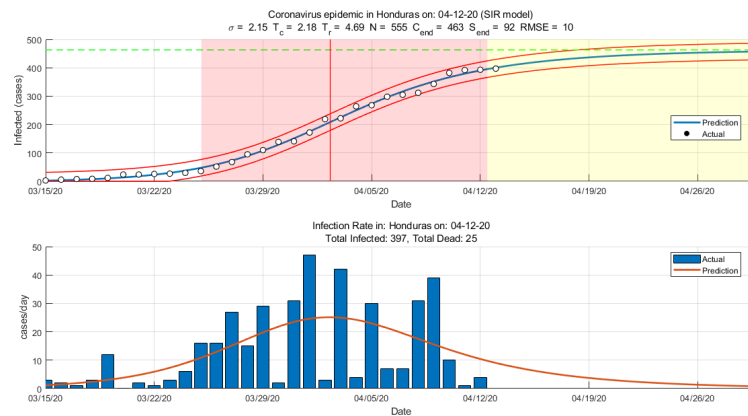
- In figures with caption “**a**”, it can be observed in the blue curve the prediction of cases per day with the **logistic model** and with blue points the real data. In the same figure, the black curve represents the prediction of accumulated cases, with logistic model too, and with white points the current accumulated declared cases. It is important to note that the total projected (predicted) number of cases is shown in the top left side of the figure.
- In figures with caption “**b**”, it can be observed results from the **SIR model**. In the first sub-plot, the blue curve represents the prediction of accumulated cases and the white points the current accumulated declared cases. On the other hand, in the second sub-plot, the red curve stays for the forecast cases per day and with blue bars the real data. Notice that the final predicted number of cases is denoted by C_{end} and the basic reproduction ratio $\mathfrak{R}_0 = \sigma$.

The first scenario that will be shown is the prediction of the spread of COVID-19 with the current data confirmed by SINAGER, results are presented

in Figure 1. It can be confirmed the affirmation of the WHO that Honduras is in the fourth phase of the propagation. The final predicted number of cases with both models are closed, 431 with the logistic model and 463 with the SIR model. Note that it is expected to finish the spread by April, 24, 2020. This supports (in a risky way) the end of the government proposed house confinement until April, 19.



(a) Forecast with logistic model.

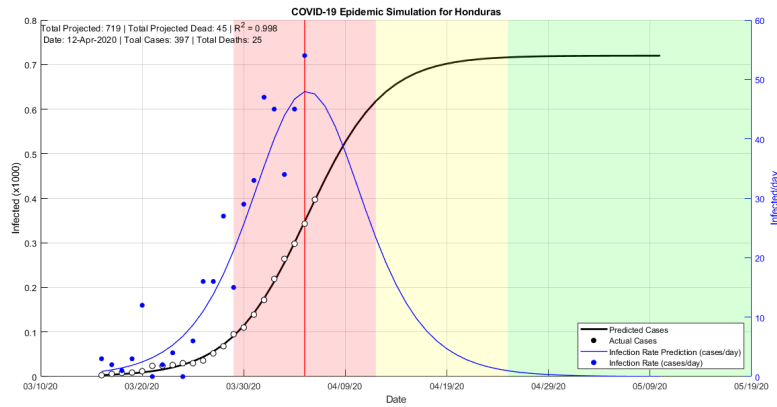


(b) Forecast with SIR model.

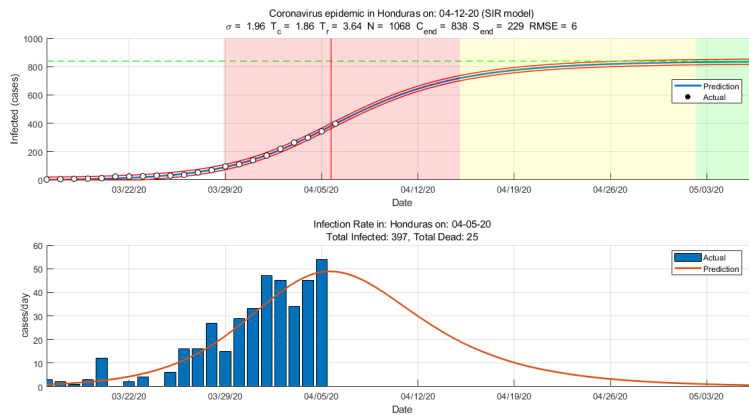
Figure 1: Prediction of the spread of COVID-19 with the current data confirmed by SINAGER, Honduras

The second scenario that will be considered for Honduras is to group the data in two-day packages, from the date there was the first decrease in the report of

new cases. The results obtained under these conditions are presented in Figure 2. Before commenting on the results, it is worth mentioning that Honduras is the only country that has experienced more than two relatively abrupt decreases in the amount of data reported, creating suspicion about its veracity. It can be seen that now the total number of infected by covid-19 exceeds 800 cases with both models. In addition, the end of the spread is expected to be May 3.



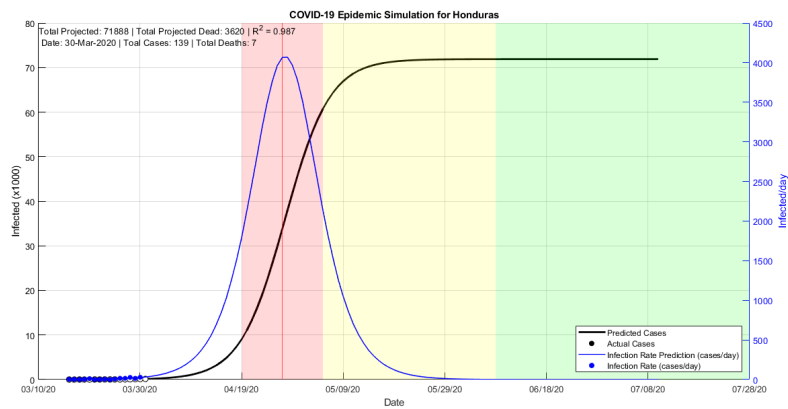
(a) Forecast with logistic model.



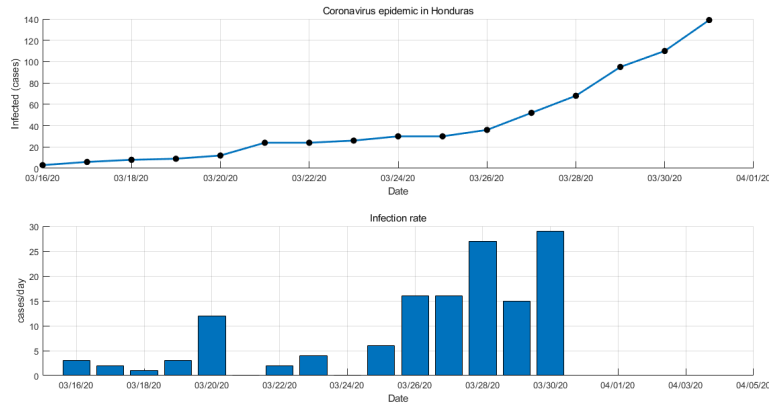
(b) Forecast with SIR model.

Figure 2: Prediction of the spread of COVID-19 with the current data confirmed by SINAGER, Honduras. Grouping the data in two-day packages, from the date there was the first decrease in the report of new cases.

A final scenario analysed for Honduras is to present the predictions taking as a deadline the day before the first decrease in reported cases. In Figure 3, the results show a much more alarming situation than the previous scenarios. It can be seen that the contagion peak is expected on day April 24, with a total number cases of 71883 and end of the contagion phase by May 30. Also note that the SIR model failed to make any prediction, so it only plotted the confirmed cases.



(a) Forecast with logistic model.



(b) Forecast with SIR model.

Figure 3: Prediction of the spread of COVID-19 with the current data confirmed by SINAGER, Honduras. Taking as a deadline the day before the first decrease in reported cases.

It could be suspected that the presented models do not adapt to the propagation of the Covid-19 in Honduras. Anyway, in Figure 4, the reader can interpret the predictions and observe how they fit to the real cases when analysing USA, Italy, Ecuador and Costa Rica data. This is only mentioned to emphasize the effectiveness of these models and perhaps reject the data received in Honduras.

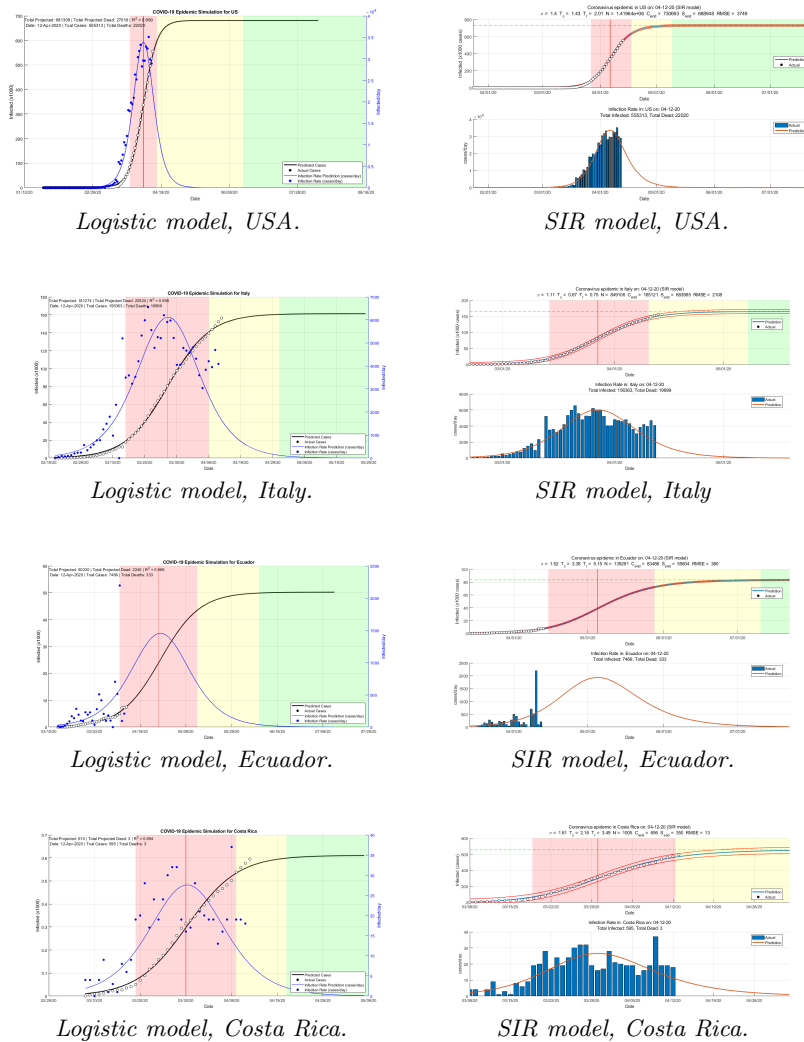


Figure 4: Prediction of the spread of COVID-19 for USA, Italy, Ecuador and Costa Rica data.

5. Conclusions

Any mathematical model for predicting the spread of an epidemic depends on the actual data provided by the corresponding authorities. In this sense, the forecasts in this document focused on the handling of the data provided by the country's authorities. However, some hypotheses were made in order to group the confirmed data in a more homogeneous way. In any of these scenarios, without a doubt, the spread of Covid-19 will continue for at least the next three weeks. The final figures vary according to the data and it has not been considered a strategy for the treatment of possible undetected data.

Many more details could be considered, such as the ratio between confirmed cases and the number of tests performed. In Honduras, so far, barely 2,000 tests have been carried out to detect 393 cases.

It is expected that the work presented, could serve as motivation for students and teachers of subjects such as calculus, linear algebra, differential equations and numerical analysis.

It is remarkable how both models presented fit precisely and robustly to the study of the situation in other countries, such as those presented in the results section.

I apologize for the possible errors found in the drafting of the document and you are invited to make your suggestions about the work to the email on the first page. In a future version of this document all these details will be considered. Furthermore, more exercise-type activities and deeper predictions will be presented.

Acknowledgement

I would like to acknowledge the persons whom supported this work for their helpful comments that substantially contributed to improve it.

References

- [1] Batista, Milan. (2020). Estimation of the final size of coronavirus epidemic by the logistic model. DOI: 10.13140/RG.2.2.36053.37603.
- [2] Batista, Milan. (2020). Estimation of the final size of the coronavirus epidemic by the SIR model. ResearchGate.
- [3] Batista, Milan. (2020). fitVirus (COVID-19 LOGISTIC Model) (<https://www.mathworks.com/matlabcentral/fileexchange/74411-fitvirus>), MATLAB Central File Exchange. Retrieved April 11, 2020.
- [4] Gobierno de la República de Honduras. (2020). <https://www.presidencia.gob.hn/index.php/sala-de-prensa/>
- [5] Harko, Tiberiu; Lobo, Francisco S. N.; Mak, M. K. (2014). Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*. 236: 184194.
- [6] McGee, Joshua. (2020). fitVirusCV19v3 (COVID-19 SIR Model) (<https://www.mathworks.com/matlabcentral/fileexchange/74676-fitviruscv19v3-covid-19-sir-model>), MATLAB Central File Exchange. Retrieved April 11, 2020.
- [7] Tang, Zhou; Li, Xianbin, Li Houqiang. (2020). Prediction of New Coronavirus Infection Based on a Modified SEIR Model. doi: <https://doi.org/10.1101/2020.03.03.20030858>
- [8] Wikipedia contributors. (2020, April 9). Compartmental models in epidemiology. In Wikipedia, The Free Encyclopedia. from <https://en.wikipedia.org>
- [9] World Health Organization (WHO). (2019). All Rights Reserved. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>