

Sensitivity, Estimability and Parameter Identification of a Mathematical Covid-19 Epidemic Model

Ilias Bouchkira^{1,2,3}

¹ Laboratoire Réactions et Génie des Procédés, University of Lorraine, Nancy Cedex, France.

² Modeling, Simulation and Data Analysis Lab, Mohammed VI Polytechnic University , Benguerir, Morocco.

³ Laboratoire Chimie, Physique, Catalyse et Environnement, Université Hassan II de Casablanca, Morocco.

Ilias.Bouchkira@univ-lorraine.fr

Abstract

In this work, a mathematical Covid-19 epidemic transmission network model is investigated for the study of the virus spread dynamics. The purpose is to present a sensitivity-based estimability analysis as well as an accurate parameter identification approaches for reliable mathematical modeling. The most sensitive parameters of the model are identified using a local sensitivity approach, these sensitivities are then used within an orthogonalization algorithm to assess the estimability of the unknown parameters from available data. A database of newly reported infected and recovered people in China is used. The most estimable model parameters are identified, their accuracy is assessed by computing confidence intervals and their numerical values are also used to compare the model predictions to real data. The Pearson Product-Moment coefficient is computed. Its high values show the accuracy of the new model parameterization and the importance of the used sensitivity-based estimability approach.

Keywords: Coronavirus, Modeling, Sensitivity analysis, Estimability analysis, Parameter identification.

1 Introduction

An outbreak of viral-looking pneumonia of unknown etiology emerged in the city of Wuhan, Hubei Province, China, in December 2019. On January 9, 2020, the discovery of a new Coronavirus (first called 2019-nCoV and then officially SARS-Cov2, different from SARS-CoV, responsible for the SARS epidemic in 2003 and MERS-CoV, responsible for an epidemic evolving since 2012 in the Middle East) was officially announced by the Chinese health authorities and the World Health Organization. This new virus is the causative agent of a new respiratory infectious disease called COVID-19. So far, no vaccine have been found which is perfectly normal according to many medical sources. The World Health Organization has stated that it would take at least a year to obtain a widely usable product.

In the meantime, it is very important to develop mathematical models to estimate certain characteristics of this epidemic, in particular, to predict important information such as the evolution of the number of new infected and recovered people, the transmission rate of the virus, the rate of recovery and death from the Covid-19, but also to predict how to minimize the spread of this new virus all over the world. To this end, some earlier mathematical modeling researches have been conducted [1,2]. Most of them are based on transmission network models developed previously for epidemics like Ebola and HIV [3,5].

Recently, many mathematical models were investigated for the study of the transmission dynamics of the Coronavirus in the most infected countries. Singhal et al. [6], Badr et al. [7], Marmarelis [8], Li et al. [9] investigated predictive models for the case of the USA. Sarkar et al. [10], Salgotra et al. [11], Patrikar et al [12], Ganiny and Nisar [13], Bagal et al. [14] discussed the mathematical modeling of the virus spread to predict its dynamics in India. Bastos and Cajueiro [15], Crokidakis et al. [16], Peixoto et al. [17], Tang et al. [18], Konarasinghe [19] presented similar work using Brazilian data. Soukhovolsky et al. [20], Amiranashvili et al. [21] presented some analysis related to Russia. Other works are available, they mainly concern other countries such as France, Italy, United Kingdom and others. Most of these models involve unknown parameters to be identified using accurate data. The accuracy and type of used data are very important for the development of a reliable mathematical model.

One of the proposed models was adopted for the Coronavirus epidemic case by Tian-Mu Chen et al. [1]. Nevertheless, the development of this model as well as the identification of its unknown parameters were carried out based on limited databases due to earlier study of the new disease and to unavailability of information. This work addresses a new parameterization of the proposed model using more data. Indeed, a local sensitivity analysis approach based on a centred finite difference approach is carried out to assess the sensitivity of the model outputs with respect to each parameter. These local sensitivities are then used withing an estimability analysis algorithm developed by Yao et al. [22] to study the estimability of the model parameters from available measurements. The most estimable parameters are identified and used to compare the predictions of the model to available data.

2 Methods

2.1 Database

The reported infected and recovered cases of SARS-CoV-2, which have been named as COVID-19, were collected for the modelling study from the China's National Infectious Disease Information System (IDIS) which requires each COVID-19 case to be reported electronically by the responsible doctor as soon as a case has been diagnosed. Some of these results were also published by researchers from the Johns Hopkins Center for Systems Science and Engineering (CSSE). The investigated database consists of the number of newly infected and newly recovered people in China from 21 January 2020 recorded during 41 days. For the modeling, a time step was taken equal to 2 days.

2.2 Mathematical modeling

A mathematical model was recently proposed by Tian-Mu Chen et al. [1], The latter considers the appearance of the Coronavirus in a bat population. The virus is supposed to be subsequently transmitted to hosts that are hunted and subsequently transported to the seafood market of Wuhan city. For the sake of simplification, the transmission of the virus from the bat population to hosts of the seafood market was assumed to be very quick, and the model assumes only transmission of the virus from the new main source (seafood market) to humans.

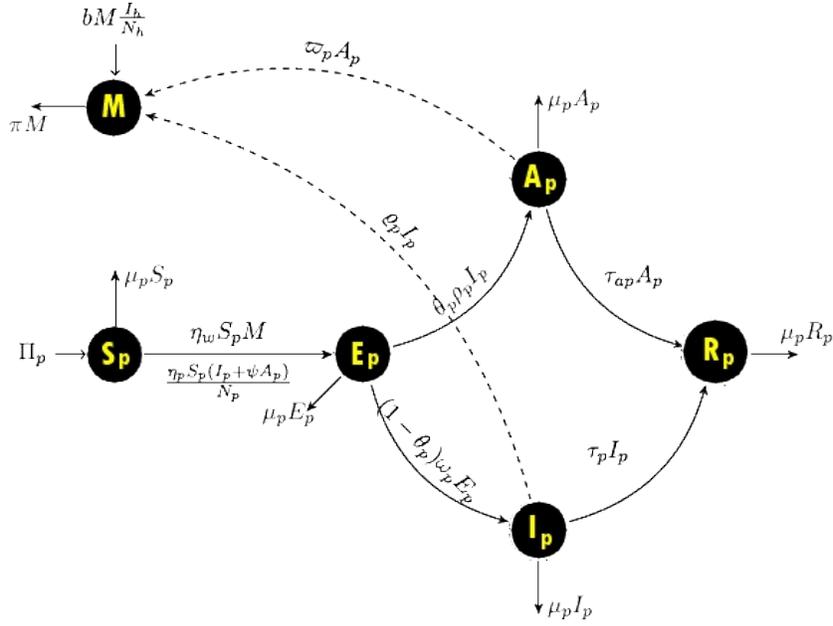


Figure 1: Interaction among the people and the main source M.

Fig. 1 shows the interactions among the people and the main source. Based on the considered interactions, the Coronavirus transmission model can be described as a system of differential equations as presented in (Eqs. 1-8). It is assumed that the total population of people denoted by N_p is classified further into five subgroups S_p , E_p , I_p , A_p and R_p which represent respectively, the susceptible, exposed, infected (symptomatic), asymptotically infected, and the recovered or the removed people. As very well described in [1,2] several parameters are used. Among them, natural birth and death of the human from the population are denoted Π_p and μ_p respectively. Each person from the susceptible class S_p risks being infected through considered contact with another person from the infected class I_p through the term given by $n_p S_p I_p$, where n_p presents the transmission rate of the COVID-19. Asymptotically infected people are also able to transfer the disease, through the term given by $\psi n_p S_p A_p$, where ψ is equal to 0 if the person from the A_p class is not asymptomatic, ψ is equal to 1 otherwise. θ_p is the proportion of asymptomatic infections. Once the incubation period ρ_p is complete, people from the A_p class join the I_p class through a transmission rate denoted w_p . People from the I_p and A_p classes can be recovered and join the R_p class with the rates τ_p and

τ_{ap} . The main source of the virus is denoted M . Susceptible people from the S_p class risk getting infected through considerable contact with the main source M through the term given by $n_w S_p M$, where n_w is the transmission rate of the disease from the main source M to the susceptible people. Finally, people from the I_p and A_p classes can transfer the virus to the main source through the rates q_p and \hat{w}_p respectively and the virus is removed from the main source with a rate denoted π . Assuming that the virus can be imported in short time to the main source the term bMI_h/Np which expresses the rate of the visiting of the main source by the host is likely to vanish.

The reduced system is presented below.

$$\frac{dS_p}{dt} = \Pi_p - \mu_b S_p - \frac{n_p S_p (I_p + \psi A_p)}{Np} - n_w S_p M \quad (1)$$

$$\frac{dE_p}{dt} = \frac{n_p S_p (I_p + \psi A_p)}{Np} + n_w S_p M - (1 - \theta_p) w_p E_p - \theta_p \rho_p E_p - \mu_p E_p \quad (2)$$

$$\frac{dI_p}{dt} = (1 - \theta_p) w_p E_p - (\tau_p + \mu_p + \zeta_p) I_p \quad (3)$$

$$\frac{dA_p}{dt} = \theta_p \rho_p E_p - (\tau_{ap} A_p + \mu_p) A_p \quad (4)$$

$$\frac{dR_p}{dt} = \tau_p I_p + \tau_{ap} A_p - \mu_p R_p \quad (5)$$

$$\frac{dM}{dt} = q_p I_p + \hat{w}_p A_p - \pi M \quad (6)$$

The feasible region and the initial conditions for the resulting model are given by:

$$\Omega = \left\{ (S_p(t), E_p(t), I_p(t), A_p(t), R_p(t)) \in R^{5+} : N_p(t) \leq \frac{\Pi_p}{\mu_p} \right\} \quad (7)$$

$$S_p(0) \geq 0, E_p(0) \geq 0, I_p(0) \geq 0, A_p(0) \geq 0, R_p(0) \geq 0 \quad (8)$$

3 Sensitivity analysis

Sensitivity analysis was carried to study the influence of the model parameters on the model outputs. In particular, on the number of newly reported cases of infections I_p . Indeed, a local sensitivity method based on the one-at-time (oat) principle was implemented and then used to compute the local sensitivities of the twelve model parameters (listed in Table 1).

The sensitivity matrix is given below [22-24], where subscripts i, j refer respectively to time in days and the number of unknown parameters. S is the model output. This matrix is computed by influencing the value of each parameter P_j by 10% forward and backward, then, the centered finite difference method is used to approximate its elements.

$$\tau_{i,j} = \begin{pmatrix} \frac{\partial S_1}{\partial P_1} & \cdots & \frac{\partial r_1}{\partial P_j} \\ \vdots & \ddots & \vdots \\ \frac{\partial S_i}{\partial P_1} & \cdots & \frac{\partial S_i}{\partial P_j} \end{pmatrix} \quad (9)$$

The results of sensitivity analysis are presented in Fig. 2. Analysis of these results show that the parameters which contribute to the spread of the new Coronavirus and consequently to increase the number of newly reported infections are parameters 3, 4, 9, and 10 (with the highest sensitivity values), they correspond respectively to the contact rate, the proportion of asymptomatic cases, the rate of transmission of the virus between the I_p class and the source M , the rate of transmission of the virus between the A_p class and the source M source. In fact, high contact rate leads to a quick spread of the virus and thus causes more newly infected cases who will subsequently infect other people even when they are asymptomatic. The second parameter that influences the number of new infected cases is the rate or number of asymptomatic persons, we believe that this is the most important and dangerous parameter, since an asymptomatic person infects other people before the symptoms of the disease appear and consequently the true number of infected people increases invisibly. The last two parameters have almost the same effect, in fact, the main source M is supposed to be the middle where many people gather (the seafood market) in the case of Wuhan city, therefore, the more the source of the virus spread grows, the more the number of infected people increases too. Moreover, the sensitivity of certain parameters is Gaussian in form, these parameters are supposed to be the main controllable factors having contributed to the increase of the number of newly infected people. Their control after $t=25$ days (where we can see that their sensitivity is decreasing) reduced the infection rates. Other parameters whose sensitivities are not Gaussian are not controllable as they are still sensitive even in the stable regime of the outbreak i.e, (After $t=25$ days). They correspond to the incubation period of the virus, the cure rate and the death rate from the virus which are currently not controllable.

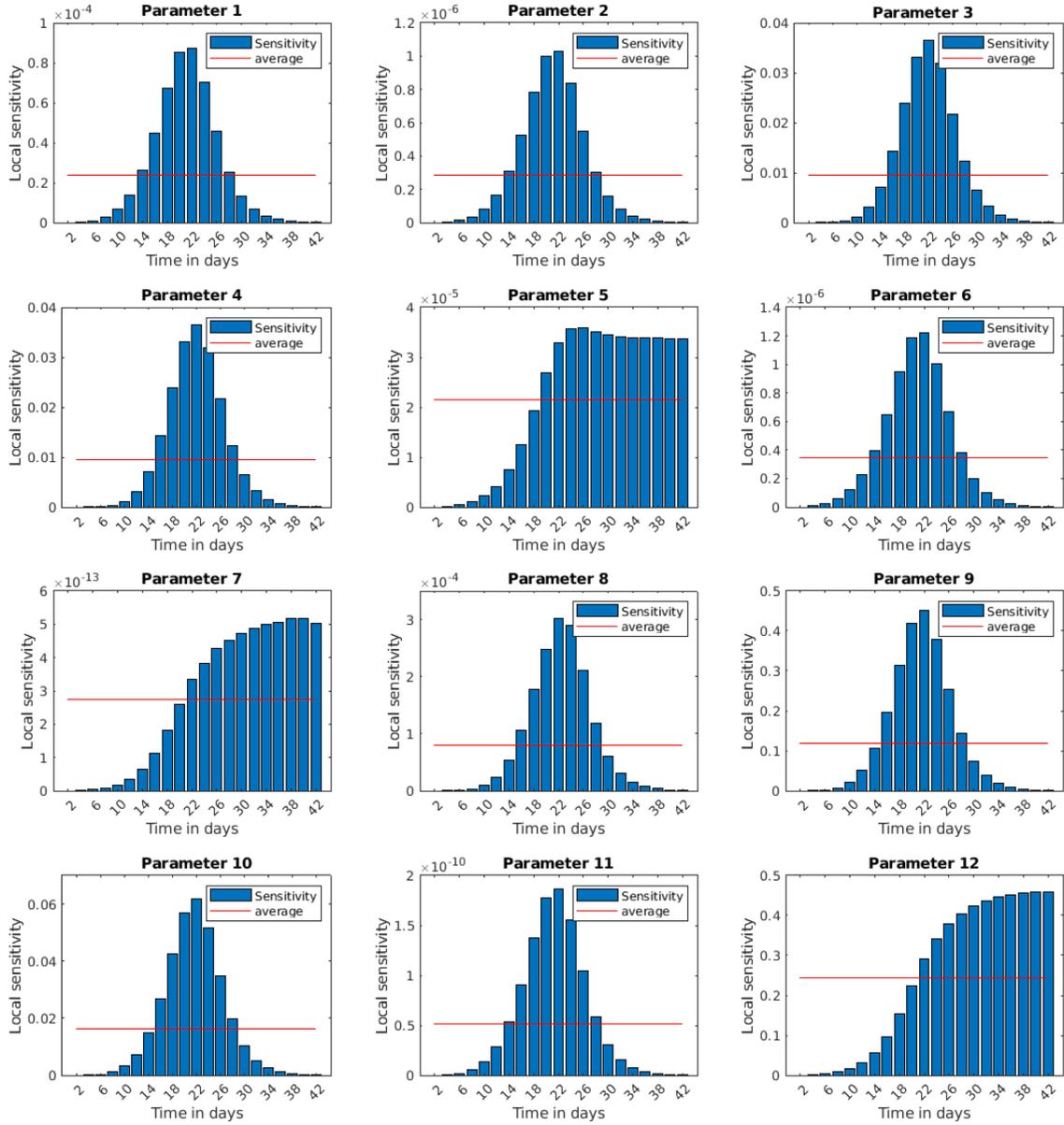


Figure 2: Local sensitivity analysis of the twelve model parameters with respect to the time

4 Estimability analysis

The objective of the estimability analysis is to identify how many of the model parameters can be estimated accurately from the available data, while the ones with low estimability potential can be set to certain nominal values without degrading the prediction capability of the model [25]. To this end, some estimability analysis methods are proposed in the literature. Among them, the orthogonalization-based methods have proven to be particularly relevant to rank the parameters from the most to the less estimable [22-24]. The method used in this work is based on a sensitivity coefficient matrix Z that is computed using the previously calculated sensitivities.

$$Z = \begin{pmatrix} S_{1,1} |_{d_1} & \cdots & S_{1,P} |_{d_1} \\ \vdots & \ddots & \vdots \\ S_{N,1} |_{d_N} & \cdots & S_{N,P} |_{d_N} \end{pmatrix} \quad (10)$$

In the matrix Z , the rows correspond to the number of data and the columns refer to the number of the model unknown parameters. The matrix Z is then implemented within the estimability algorithm developed by Yao et al. [22] in order to rank the parameters from the most to the less estimable ones. In this case, the estimability analysis involves a vector P of 12 unknown parameters and a vector \mathbf{y} of 2 measured outputs i.e. the number of new infected and recovered people I_p and R_p . Moreover, the elements of the sensitivity matrix Z are normalized as:

$$S_{i,j}^* = \frac{\overline{P_j}}{\overline{y_i}} \cdot S_{i,j} \quad (11)$$

where $\overline{P_j}$ and $\overline{y_i}$ scaling factors. The elements $S_{i,j}^*$ are therefore of the same order of magnitude. The Magnitude values of twelve parameters are also reported in table 1. Their high values show that the database contains enough information to estimate all of them. Indeed, Yao et al. [22] proposed an estimability cutoff threshold equal to 0.04 to say that unknown parameters with the lowest magnitude are not identifiable, consequently, their values must be fixed either from the literature or from previous works. It is to precise that the value of 0.04 means that a parameter is estimable if a variation of its value of 10% causes at least 2% of the model output.

5 Parameter identification

The unknown parameters of the Coronavirus model were identified from the available data using Matlab optimization environment. More specifically, `fmincon` optimizer based on the interior-point algorithm [11] was used to solve the identification problem to possibly reach the global optimality. An objective function was defined as:

$$F(P) = \sum_{i=1}^N \left(\frac{I_{p_i}^{obs} - I_{p_i}^{mod}}{I_{p_i}^{obs}} \right)^2 + \sum_{i=1}^N \left(\frac{R_{p_i}^{obs} - R_{p_i}^{mod}}{R_{p_i}^{obs}} \right)^2 \quad (12)$$

$I_{p_i}^{obs}$ and $I_{p_i}^{mod}$ corresponds to the reported new infected cases and the model new infected predictions. $E_{p_i}^{obs}$ and $E_{p_i}^{mod}$ corresponds to the reported new recovered cases and the model new recovered predictions. Subscript i refers to the time in days. On the other hand, the accuracy of the identified parameters is assessed through the calculation of confidence intervals computed as follows [27] since the objective function can be written as:

$$F(P) = r(P)^T \cdot W \cdot r(P) \quad (13)$$

where $r(P)$ is the residue vector, i.e. the difference between the predicted and observed values of the outputs, W is a weighing matrix and P corresponds to the vector of the unknown parameters. Assuming that the errors of recording the observed data are normally distributed and independent, the co-variance matrix C for the least squares problems is expressed as:

$$C = \frac{F(P^*)}{d - n} (J^T W J)^{-1} J^T W^2 J (J^T W J)^{-1} \quad (14)$$

P^* is the vector of optimal parameters which minimize the objective function $F(P)$ and J is the Jacobian matrix of the vector $r(P)$. We note that (Eq.14) is an approximation for nonlinear problems and it is more precise when non-linearities are not strong. In this work, W is equal to the identity matrix, i.e. all used observations have the same weight in the objective function. The Jacobian matrix is given by:

$$J = \begin{pmatrix} \frac{\partial r_1}{\partial P_1} & \cdots & \frac{\partial r_1}{\partial P_j} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_i}{\partial P_1} & \cdots & \frac{\partial r_i}{\partial P_j} \end{pmatrix} \quad (15)$$

Subscripts i and j correspond to the number of outputs and the number of unknown parameters respectively. The Jacobian matrix was computed using a local one-at-a-time (OAT) method, it consists in disturbing the value of each parameter P_j^* 10% forward and backward, then, the centred finite difference method is used to approximate the elements of J . This required to perform 1008 simulations (12 parameters x 42 observed data x 2). The uncertainty on each estimated parameter j is given by:

$$\epsilon_{P_j} = \pm \frac{\sqrt{c_{jj}} t_{1-\alpha/2, \nu}}{P_j^*} .100\% \quad (16)$$

c_{jj} is the j^{th} diagonal element of the C matrix, $t_{1-\alpha/2, \nu}$ is deduced from the student t-distribution with ν degrees of freedom and corresponds to the probability of $1 - \alpha/2$ that the true values of the parameter lies within the confidence interval given as:

$$P_j \in [P_j^* - \sqrt{c_{jj}} t_{1-\alpha/2, \nu}; P_j^* + \sqrt{c_{jj}} t_{1-\alpha/2, \nu}] \quad (17)$$

In this work the probability $1 - \alpha/2$ is taken equal to 95%. Is to precise that although the confidence intervals are an approximation of the accuracy of the identified parameters, particularly when strong non-linearities are involved in the model, they show whether the parameters are determined with good precision or not. On the other hand, the Pearson product-moment coefficient is also computed and reported as [27]:

$$r = \frac{\sum_{i,j} (S_{ij}^m - \overline{S_{ij}^m}) (S_{ij}^e - \overline{S_{ij}^e})}{\sqrt{\left(\sum_{i,j} (S_{ij}^m - \overline{S_{ij}^m})\right)^2 \left(\sum_{i,j} (S_{ij}^e - \overline{S_{ij}^e})\right)^2}} \quad (18)$$

It allows to compare the model predictions to real reported cases used for the parameter identification. S_{ij}^m is the model prediction, S_{ij}^e is the corresponding observed values and $\overline{S_{ij}}$ are their average values.

Table 1: Numerical values of twelve unknown parameters of the Corona virus model along with their estimability magnitude (EM) and their confidence intervals

Parameter	Description	Value	(EM)	Confidence interval
Π_p	Birth rate	$\mu_p \cdot N_p(0)$	-	-
μ_p	Natural mortality rate	$1/(76.79 \times 365)$	-	-
n_p	Contact rate	0.897	4.36	[0.790;1.607]
ψ	Transmissibility multiple	0.156	$0.16 \cdot 10^2$	[0.133;0.176]
n_w	Disease transmission coefficient	$1.39 \cdot 10^{-6}$	3.17	$[1.23 \cdot 10^{-6}; 1.39 \cdot 10^{-6}]$
θ_p	The proportion of asymptomatic infection	$1.31 \cdot 10^{-5}$	$7.15 \cdot 10^4$	$[2.28 \cdot 10^{-6}; 1.31 \cdot 10^{-5}]$
w_p	Incubation period	$2.55 \cdot 10^{-5}$	$0.24 \cdot 10^4$	$[1.78 \cdot 10^{-5}; 2.55 \cdot 10^{-5}]$
ρ_p	Incubation period	0.416	$0.95 \cdot 10^4$	[0.107;0.460]
τ_p	Removal or recovery rate of I_p	$3.66 \cdot 10^{-7}$	$0.42 \cdot 10^4$	$[3.64 \cdot 10^{-7}; 3.6 \cdot 10^{-7}]$
τ_{ap}	Removal or recovery rate of A_p	0.030	$0.17 \cdot 10^4$	[0.013;0.031]
q_p	Contribution of the virus to M by I_p	0.996	1.31	[0.322;1.317]
\hat{w}_p	Contribution of the virus to M by A_p	0.540	5.94	[0.301;0.703]
π	Removing rate of virus from M	$8.10 \cdot 10^{-6}$	$0.11 \cdot 10^2$	$[8.03 \cdot 10^{-6}; 8.10 \cdot 10^{-6}]$
ζ	Death rate of infected people	0.344	$0.56 \cdot 10^4$	[0.218;0.419]

For the parameter identification, the birth rate and the natural mortality rate were taken from the literature [28]. The other unknown parameters are estimated from the number of newly reported infected and recovered people from 21 January to 11 March 2020 in China. The total population of China for the year 2019 was estimated to 1.386 billion [28]. The life expectancy in China for the year 2019 is 76.79, therefore, the parameters μ_p is given by $1/76.79$ per year. The value of Π_p is deduced as $\Pi_p = \mu_p \cdot N(0)$. At $t=0$. $N_p(0)$ is equal to the population of China, i.e., 1.386 billion. We suppose that at $t=0$, there are no recovered cases yet. The other values are chosen to satisfy $N_p(0) = S_p(0) + E_p(0) + I_p(0) + R_p(0) + A_p(0)$. Table 1 shows the identified parameters along with their confidence intervals. It is noteworthy that there are no confidence intervals that exhibit boundaries with opposite sign, meaning that the parameters values are likely to vanish. The new parametrization is used to compare the model outputs to the reported infections and recovered cases (Figs. 3). The resulting Pearson product-moment coefficient r is found to be higher than 0.98 hence showing the good quality of the model predictions.

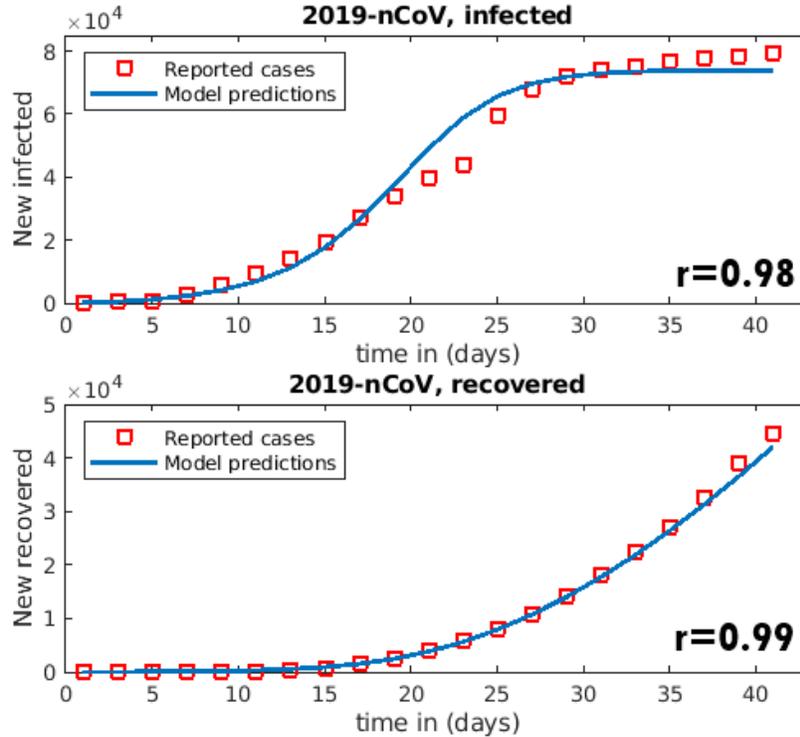


Figure 3: Reported cases of 2019-nCoV new infected and recovered in China vs model predictions.

The epidemic curves presented in Figs. 2 present a comparison between the model predictions and the observed data which are generated using China’s National Infectious Disease Information System (IDIS) which requires each COVID-19 case to be reported electronically by the responsible doctor as soon as a case has been diagnosed. It illustrates that the epidemic rapidly grew from 10-25 days, which is very consistent with the sensitivity analysis results. We believe that the model parameters whose sensitivities have decreased after $t=25$ days are the most controllable. Their control allowed the decrease the spread of the virus. In particular, parameters 1, 2, 3, 9 and 10 which reflect the human-to-human transmission of the virus were controlled in China by applying an appropriate level of emergency management protocols such as total containment and returning to basic public health policies for more than 60 million people. In addition, the proportion of asymptomatic people reflected in this work by parameter 4 was controlled by expanding surveillance to detect COVID-19 transmission chains, by testing all patients with atypical pneumonia, conducting screening in some patients with upper respiratory illnesses and/or recent COVID-19 exposure, and adding testing for the COVID-19 virus to existing surveillance systems.

6 Conclusion

Until researchers find a vaccine to combat the spread of the new Coronavirus, it is very important to develop mathematical models to predict the process and key parameters that contribute to its spread. This work aimed to the modeling of the new Coronavirus (n-coV 2019). The model under development involves twelve unknown parameters to be identified from available data i.e., the number of newly reported infected and recovered patients for a period of 41 days. First, the sensitivities of the twelve parameters were used to classify the most influential parameters of the model. Then, these sensitivities were used in an estimability algorithm to determine whether the available data allows the identification of all the unknown parameters of the model or only a subset. The sensitivity and estimability results were used to define the most estimable parameters. Their values along with their interval confidence intervals were calculated and reported. Finally, the results of the model prediction and the observed cases are compared and found to be in good agreement.

Nevertheless, this model was investigated based on data from only from China, its use to predict similar information for other countries may give erroneous results. For this reason, a new study will be carried out in the future to take into account as many data as possible, especially from countries still suffering from the spread of the virus for further development of the model under study.

References

- [1] Chen, T. M., Rui, J., Wang, Q. P., Zhao, Z. Y., Cui, J. A., Yin, L. (2020). A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infectious Diseases of Poverty*, 9(1), 1-8.
<https://doi.org/10.1186/s40249-020-00640-3>
- [2] Khan, M. A., Atangana, A. (2020). Modeling the dynamics of novel coronavirus (2019-nCov) with fractional derivative. *Alexandria Engineering Journal*.
<https://doi.org/10.1016/j.aej.2020.02.033>
- [3] Chowell, G., Nishiura, H. (2014). Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC medicine*, 12(1), 196. <https://doi.org/10.2807/1560-7917.ES2014.19.40.20920>
- [4] Althaus, C. L., Low, N., Musa, E. O., Shuaib, F., Gsteiger, S. (2015). Ebola virus disease outbreak in Nigeria: transmission dynamics and rapid control. *Epidemics*, 11, 80-84.
<https://doi.org/10.1016/j.epidem.2015.03.001>

- [5] Case, K. K., Ghys, P. D., Gouws, E., Eaton, J. W., Borquez, A., Stover, J., ... Hallett, T. B. (2012). Understanding the modes of transmission model of new HIV infection and its use in prevention planning. *Bulletin of the World Health Organization*, 90, 831-838A.
- [6] Singhal, A., Singh, P., Lall, B., Joshi, S. D. (2020). Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos, Solitons Fractals*, 138, 110023.
<https://doi.org/10.1016/j.chaos.2020.110023>.
- [7] Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases*.
[https://doi.org/10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3).
- [8] Marmarelis, V. Z. (2020). Predictive modeling of Covid-19 data in the US: Adaptive phase-space approach. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 207-213.
<https://doi.org/10.1109/OJEMB.2020.3008313>
- [9] Li, Y., Zhang, R., Zhao, J., Molina, M. J. (2020). Understanding transmission and intervention for the COVID-19 pandemic in the United States. *Science of the Total Environment*, 748, 141560.
<https://doi.org/10.1016/j.scitotenv.2020.141560>
- [10] Sarkar, K., Khajanchi, S., Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in India. *Chaos, Solitons Fractals*, 139, 110049.
<https://doi.org/10.1016/j.chaos.2020.110049>.
- [11] Salgotra, R., Gandomi, M., Gandomi, A. H. (2020). Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming. *Chaos, Solitons Fractals*, 109945.
<https://doi.org/10.1016/j.chaos.2020.109945>.
- [12] Patrikar, S., Poojary, D., Basannar, D. R., Faujdar, D. S., Kunte, R. (2020). Projections for novel coronavirus (COVID-19) and evaluation of epidemic response strategies for India. *Medical Journal Armed Forces India*, 76(3), 268-275.
<https://doi.org/10.1016/j.mjafi.2020.05.001>.
- [13] Ganiny, S., Nisar, O. (2020). Mathematical Modeling and a Month Ahead Forecast of the Coronavirus Disease 2019 (COVID-19) Pandemic: An Indian Scenario. *medRxiv*.
- [14] Bagal, D. K., Rath, A., Barua, A., Patnaik, D. (2020). Estimating the parameters of susceptible-infected-recovered model of COVID-19 cases in India during lockdown periods. *Chaos, Solitons Fractals*, 140, 110154.
<https://doi.org/10.1016/j.chaos.2020.110154>.
- [15] Bastos, S. B., Cajueiro, D. O. (2020). Modeling and forecasting the Covid-19 pandemic in Brazil. *arXiv preprint arXiv:2003.14288*.
- [16] Crokidakis, N. (2020). Data analysis and modeling of the evolution of COVID-19 in Brazil. *arXiv preprint arXiv:2003.12150*.
- [17] Peixoto, P. S., Marcondes, D., Peixoto, C., Oliva, S. M. (2020). Modeling future spread of infections via mobile geolocation data and population dynamics. An application to COVID-19 in Brazil. *PloS one*, 15(7), e0235732.
<https://doi.org/10.1371/journal.pone.0235732>.

- [18] Tang, Y., Serdan, T. D., Masi, L. N., Tang, S., Gorjao, R., Hirabara, S. M. (2020). Epidemiology of COVID-19 in Brazil: using a mathematical model to estimate the outbreak peak and temporal evolution. *Emerging microbes infections*, 9(1), 1453-1456.
<https://doi.org/10.1080/22221751.2020.1785337>
- [19] Konarasinghe, K. M. U. B. (2020). Modeling COVID-19 epidemic of India and Brazil. *J. New Front. Health-care Biol. Sci*, 1(1), 15-25.
- [20] Soukhovolsky, V., Kovalev, A., Pitt, A., Kessel, B. (2020). A new modelling of the COVID 19 pandemic. *Chaos, Solitons Fractals*, 139, 110039.
<https://doi.org/10.1016/j.chaos.2020.110039>.
- [21] Amiranashvili, A. G., Khazaradze, K. R., Japaridze, N. D. (2020). Analysis of twenty-week time-series of confirmed cases of New Coronavirus COVID-19 and their simple short-term prediction for Georgia and Neighboring Countries (Armenia, Azerbaijan, Turkey, Russia) in amid of a global pandemic. medRxiv.
- [22] Yao, K. Z., Shaw, B. M., Kou, B., McAuley, K. B., Bacon, D. W. (2003). Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polymer Reaction Engineering*, 11(3), 563-588.
<https://doi.org/10.1081/PRE-120024426>
- [23] Benyahia, B., Latifi, M. A., Fonteix, C., Pla, F. (2013). Emulsion copolymerization of styrene and butyl acrylate in the presence of a chain transfer agent. Part 2: Parameters estimability and confidence regions. *Chemical Engineering Science*, 90, 110-118.
<https://doi.org/10.1016/j.ces.2009.09.036>
- [24] McLean, K. A., McAuley, K. B. (2012). Mathematical modelling of chemical processes—obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. *The Canadian Journal of Chemical Engineering*, 90(2), 351-366.
<https://doi.org/10.1002/cjce.20660>
- [25] Fysikopoulos, D., Benyahia, B., Borsos, A., Nagy, Z. K., Rielly, C. D. (2019). A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers Chemical Engineering*, 122, 275-292.
<https://doi.org/10.1016/j.compchemeng.2018.09.007>
- [26] Ye, Y. (2011). *Interior point algorithms: theory and analysis* (Vol. 44). John Wiley Sons.
- [27] Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.
- [28] China Population 1950-2020, [https:// www.macrotrends.net/cities/20712/wuhan/population](https://www.macrotrends.net/cities/20712/wuhan/population).