Predicting the COVID-19 Spread Using Compartmental Model and Extreme Value Theory with Application to Egypt and Iraq

Mahmoud A. Ibrahim, Amenah Al-Najafi, and Attila Dénes

1 Introduction

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that is a respiratory pathogen. The disease spreads mainly through respiratory droplets that are produced when an infected person coughs, sneezes, sings, speaks, or breathes. The most common symptoms of COVID-19 are fever, dry cough, fatigue, shortness of breath, sore throat, muscle pain, loss of smell, loss of appetite, headache, and conjunctivitis [1, 2]. Most infected persons (about 80%) develop mild to moderate illness and recover without hospitalization. About 20% become seriously ill and require oxygen, and 5% become critically ill and require intensive care. The background of the disease in Iraq and Egypt can be found in [3].

A variety of mathematical models have been developed to understand the epidemiological features of COVID-19 and the transmission dynamics for many countries, including France [4], Germany [5], Hungary [9], the UK [6], and the USA [7, 8]. Ibrahim and Al-Najafi [3] studied the spread of COVID-19 epidemic in Iraq and Egypt by using compartmental, logistic regression, and Gaussian models, providing a forecast of the spread of COVID-19 in Iraq. Furthermore, we predicted the possible start of the second wave of the COVID-19 epidemic in Egypt using generalized SEIR with time-periodic transmission rate. Here, we

M. A. Ibrahim (✉)
Bolyai Institute, University of Szeged, Szeged, Hungary

Department of Mathematics, Faculty of Science, Mansoura University, Mansoura, Egypt
e-mail: mibrahim@math.u-szeged.hu

A. Al-Najafi · A. Dénes
Bolyai Institute, University of Szeged, Szeged, Hungary
e-mail: amenah@math.u-szeged.hu; densea@math.u-szeged.hu

R. P. Mondaini (ed.), *Trends in Biomathematics: Chaos and Control in Epidemics, Ecosystems, and Cells*, https://doi.org/10.1007/978-3-030-73241-7_4

establish a compartmental mathematical model for the spread of COVID-19, taking into account presymptomatic, mildly, and symptomatically infected individuals. We estimate the parameters that provide the best fit to the incidence data from both countries.

Extreme value theory (EVT) is widely applied in many disciplines, including public health. We refer to some of these studies, Lim et al. [10], in which EVT was used to model the extremes in dengue case counts using provincial-level data in Thailand from 1993 to 2018. Lim et al. [11] analyzed the dengue incidence data in Singapore by using time-varying extreme mixture (tvEM) methods to account for the time dependence of dengue case numbers over extreme and non-extreme time periods. In [12], the annual maxima of pneumonia and influenza deaths were plotted against the return level over the period 1979–2011. Chen et al. [13] used EVT to forecast the probability of outbreaks of highly pathogenic influenza. In more recent research, the EVT has been used to project the future of COVID-19 confirmed cases in Italy, Australia, Iran, South Africa, the USA, and Chile [14]. Here, we estimate the return level and the return period of the COVID-19 epidemic to predict the future of the disease in Egypt and Iraq. We provide several scenarios for the possible peak and its timing using Gaussian2 fit model.

This chapter is organized as follows. Section 2 describes the various methods applied in our work, while the results provided by these methods are given in Sect. 3. This chapter is concluded by a discussion in Sect. 4.

AQ1

2 Methods

2.1 Compartmental Model for COVID-19 Transmission

The population is divided into seven compartments: susceptible (denoted by $S(t)$), exposed ($E(t)$), presymptomatic infected ($P(t)$), symptomatically infected ($I_s(t)$), mildly infected ($I_m(t)$), treated ($I_T(t)$), and recovered individuals ($R(t)$). The total size of the population at any time t is given by $N(t) = S(t) + E(t) + P(t) + I_m(t) + I_s(t) + I_T(t) + R(t)$.

The transmission dynamics is shown in the flow diagram in Fig. 1, and our model takes the form

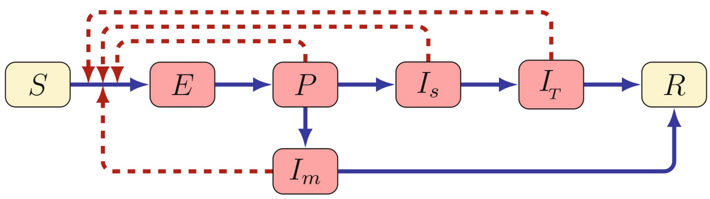


Fig. 1 Follow diagram of the COVID-19 transmission

Table 1 Description of the model (1) parameters

Parameters	Descriptions
β	Transmission rate from infectious classes to susceptible
$\kappa_p, \kappa_m, \kappa_T$	The relative transmissibility of P, I_m and I_T, respectively
θ	Proportion of asymptomatic infections
γ_s	Progression rate from I_s to I_T
γ_m, γ_T	Recovery rates
δ_T	Disease-induced death rate
ν_e, ν_p	Incubation rates

$$
\begin{aligned}
S'(t) &= -\beta \frac{\kappa_p P(t) + \kappa_m I_m(t) + I_s(t) + \kappa_T I_T(t)}{N(t)} S(t),\\
E'(t) &= \beta \frac{\kappa_p P(t) + \kappa_m I_m(t) + I_s(t) + \kappa_T I_T(t)}{N(t)} S(t) - \nu_e E(t),\\
P'(t) &= \nu_e E(t) - \nu_p P(t),\\
I'm(t) &= \theta \nu_p P(t) - \gamma_m I_m(t),\\
I's(t) &= (1-\theta)\nu_p P(t) - \gamma_s I_s(t),\\
I_T'(t) &= \gamma_s I_s(t) - \gamma_T I_T(t) - \delta_T I_T(t),\\
R'(t) &= \gamma_m I_m(t) + \gamma_T I_T(t).
\end{aligned}
\tag{1}
$$

AQ2

The description of the model parameters is listed in Table 1. Susceptibles are those who can be infected through COVID-19. Once a person has been infected with the disease, who moves up to the exposed class, these individuals do not yet have symptoms and can not transfer the virus to susceptible individuals. Exposed individuals progress to presymptomatic class, and these individuals do not yet have symptoms but can transfer the virus. Following the incubation period, presymptomatic individuals move to one of the symptomatically infected class and the mildly infected class, based on whether or not that individual shows symptoms or not. Mildly infected individuals progress to the symptomatically compartment or the recovered class. Symptomatically infected individuals move to the treated compartment, which includes those who reported hospitalized. After the infectious period, the treated persons move to the recovered class. To keep our model simpler, we do not add separate compartments for the quarantined individuals. In particular, β represents the transmission rate from symptomatically infected to susceptible, while $\beta\kappa_p$, $\beta\kappa_m$, and $\beta\kappa_T$ are the transmission rates from presymptomatic, mildly infected, and treated to susceptible, respectively. The length of the latent period for humans is $1/\nu$, while $1/\gamma_m$, $1/\gamma_T$ denote the lengths of the infected period for mildly and symptomatically infected people, respectively. The parameter θ is the fraction of mildly infected among all the infected people.

2.1.1 Derivation of the Basic Reproduction Number

By using the next generation method introduced in [19], we derive a formula for the basic reproduction number of (1). Then by considering the infectious states E,P, I_m, I_s, and I_T in (1) and substituting the values in the disease-free equilibrium $(N, 0, 0, 0, 0, 0, 0)$, we calculate the matrices F and V for the new infection terms and the remaining transfer terms. These two matrices are, respectively, given by

$$F = \begin{bmatrix} 0 & \beta\kappa_p & \beta\kappa_m & \beta & \beta\kappa_T \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } V = \begin{bmatrix} \nu_e & 0 & 0 & 0 & 0 \\ -\nu_e & \nu_p & 0 & 0 & 0 \\ 0 & -\theta\nu_p & \gamma_m & 0 & 0 \\ 0 & -(1-\theta)\nu_p & 0 & \gamma_s & 0 \\ 0 & 0 & 0 & -\gamma_s & \gamma_T + \delta_T \end{bmatrix}.$$

According to [19], the basic reproduction number is the largest absolute eigenvalue of FV^{-1}, and thus, it is given by

$$\mathcal{R}_0 = \rho(FV^{-1}) = \frac{\beta\kappa_p}{\nu_p} + \frac{\theta\beta\kappa_m}{\gamma_m} + \frac{(1-\theta)\beta\nu_e}{\nu_p\gamma_s} + \frac{(1-\theta)\beta\nu_e\kappa_T}{\nu_p(\gamma_T + \delta_T)}. \tag{2}$$

Besides calculating the basic reproduction number $\mathcal{R}_0$ of the model (1), effective reproduction rate $\mathcal{R}_{eff} = \mathcal{R}_0\frac{S(t)}{N}$ can also be estimated by this formula, measuring the average number of secondary cases per infectious case in a population. In addition, the time-dependent reproduction number can be calculated from incidence data (see e.g., [20] for details).

2.2 *Return Level Estimation*

The application of EVT offers different techniques to study the behavior of a sample with very high or very low levels. One of the important techniques of extreme value theory is the idea of the return level. The return level is strongly related to the return period: it is the quantile that will be reached or exceeded once in every year. In this chapter, we will use it to investigate the upper-tail distribution properties of the infection of the COVID-19 epidemic. In this subsection, we follow the methods and definitions given in [15].

AQ3

Let X be a random variable with cumulative distribution function F, and the distribution function of this random variable is called excess distribution function over the threshold u denoted by F_u, defined as

$$F_u(x) = P(X - u \leq x \mid X > u) = \frac{F(u+x) - F(u)}{1 - F(u)}, \qquad x \geq 0,$$

where $1 - F(u)$ is the exceedance probability, and the mean excess function of X $e(u) = E(X - u \mid X > u)$ denotes the mean residual life function. The method is based on exceedances over a specified threshold. Assuming that the appropriate distribution is chosen and then the parameters are estimated, it is useful to calculate the return level. For a given threshold u, assume that the generalized Pareto (GP) distribution with scale σ and shape ξ parameters is a suitable model for exceedances. For sufficiently large u, the distribution function of $(X - u)$, conditional on $X > u$, is therefore approximately

$$H(x) = 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}}, \qquad \xi > 0, \tag{3}$$

AQ4

where $\tilde{\sigma} = \sigma + \xi(u - \mu)$. Let $\zeta_u = P\{X > u\}$, and let x_m be the value that is exceeded once in every m periods on average, and the level x_m will be obtained from

$$x_m = \begin{cases} u + \dfrac{\sigma}{\xi}[(m\zeta_u)^{\xi} - 1] & \xi \neq 0 \\ u + \sigma \log(m\zeta_u) & \xi = 0 \end{cases} \tag{4}$$

provided m is sufficiently large to ensure that $x_m > u$.

To predict the second wave of the COVID-19 epidemic, we apply a Gaussian2 fit model. Let $I(x)$ denote the Gaussian2 function, and it is given by

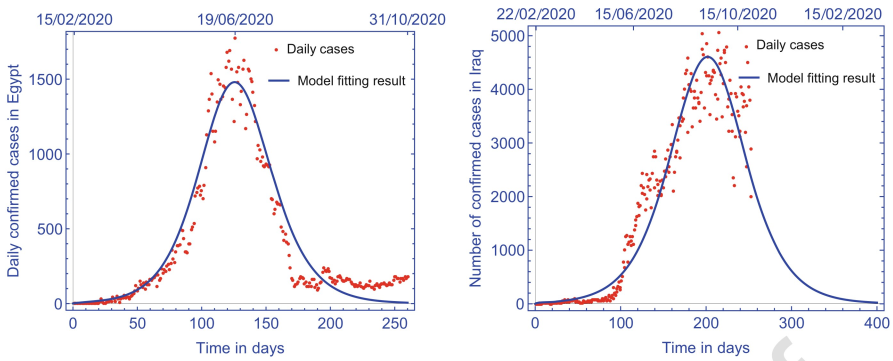
$$I(x) = \sum_{j=1}^{2} I_j \exp\left(-\left(\frac{x - \mu_j}{\sigma_j}\right)^2\right), \tag{5}$$

where I_j is the amplitude, μ_j is the time of the peak, and σ_j is related to the peak width.

3 Results

3.1 Parameter Estimation for Iraq and Egypt

The data were collected from the Worldometer website [16, 17]. We focus on the data from 22 February to 31 October, 2020 in Iraq and from 15 February to 31 October, 2020 in Egypt.



(a) Egypt (b) Iraq

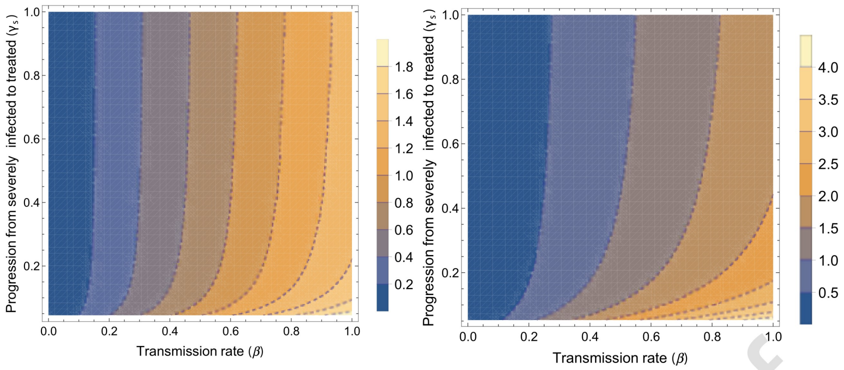
Fig. 2 The model (1) fitted to the daily confirmed cases in (**a**) Egypt and (**b**) Iraq with parameters given in Table 2

Table 2 Parameters and fitted values of model (1) in the case of Iraq and Egypt

Parameters	Value for Iraq $\mathcal{R}_0 = 1.122$	Value for Egypt $\mathcal{R}_0 = 1.129$	Source
β	0.572	0.817	Fitted
κ_p	0.284	0.277	Fitted
κ_m	0.275	0.235	Fitted
κ_T	0.211	0.368	Fitted
θ	0.728	0.805	Fitted
γ_s	0.5	0.255	Fitted
γ_m	0.23	0.203	Fitted
γ_T	0.098	0.336	Fitted
δ_r	0.164	0.191	Fitted
ν_e	0.259	0.155	Fitted
ν_p	0.483	0.93	Fitted

To estimate the model (1) parameters giving the best fit, we applied Latin hypercube sampling, a method used in statistics to measure simultaneous variation of multiple parameters (see e.g., [18] for details). AQ5

Figure 2 shows the model (1) fitted to the daily number of confirmed cases in (a) from Egypt, 15 February 2020 to 31 October 2020, and in (b) from Iraq, 22 February 2020 to 31 October 2020. Our model gives a reasonable good fit for both countries, showing the peak in Egypt and predicting the peak in Iraq. The fitting parameter results are listed in Table 2.



(a) Egypt (b) Iraq

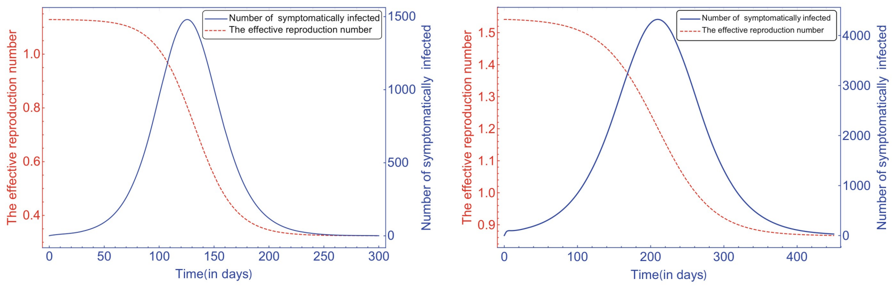
Fig. 3 The contour plot of the basic reproduction number for Iraq and Egypt as a function of transmission rate (β) and progression rate (γ_s) from I_s to I_T

3.2 *Reproduction Numbers*

In order to quantify the effort needed to eradicate infectious diseases, the basic reproduction rate $\mathcal{R}_0$ is an important threshold parameter and is defined as the expected number of secondary infections generated by one infected person in a population where all individuals are susceptible to infection. The basic reproduction number is estimated from the incidence data using exponential growth (EG) method (see e.g., [20] for details), and we found that $\mathcal{R}_0 = 1.047$ for Egypt and $\mathcal{R}_0 = 1.078$ for Iraq. The reproduction number in both countries is greater than one and the disease persists.

Formula (2) gives us the basic reproduction number in any time point by substituting the parameter values into it. To assess the dependence of the basic reproduction number on the parameters that can be subject to control the spread of the virus, the contour plot of the basic reproduction number in terms of the transmission rate (β) and progression rate from symptomatically infected to hospitalized individuals (γ_s) for the two countries is shown in Fig. 3.

Figure 4 shows the effective reproduction number along with the number of symptomatically infected in Egypt and Iraq, 2020–2021, showing that the number of infected individuals begins to decline when the effective reproduction number goes below 1. The highest value of the effective reproduction number is calculated to be about $\mathcal{R}_{eff} \approx 1.129$ in Egypt and $\mathcal{R}_{eff} \approx 1.122$ in Iraq.



(a) Egypt (b) Iraq

Fig. 4 The effective reproduction number and the number of symptomatically infected in (**a**) Egypt and (**b**) Iraq, 2020–2021

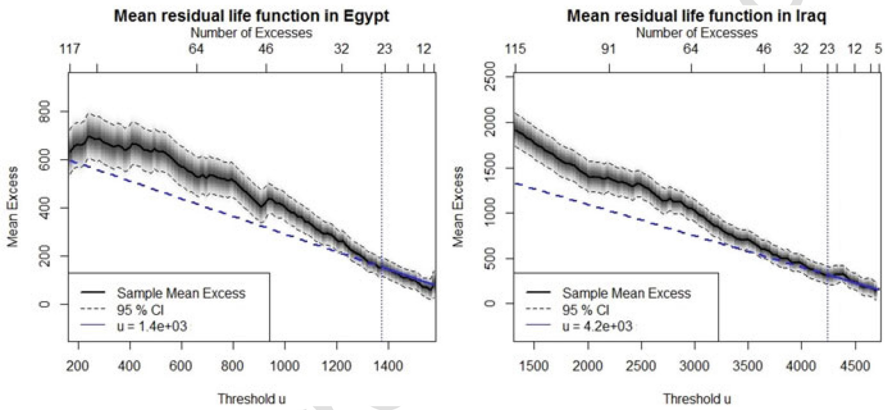


Fig. 5 Mean excess plot with threshold in Iraq and Egypt, 2020

3.3 *Prediction of the Second Wave of the COVID-19 Epidemic*

The application of the return level required choosing an optimal threshold assuming that data exceeding a specified threshold follows a Pareto distribution to determine an accurate return level estimate. It is very important to choose a plausible threshold value because choosing a threshold value that is too small leads to an imprecise estimate and choosing a threshold value that is too high leads to a biased estimate. The results of the empirical mean excess function show the appropriate threshold value for our data and also the peak value for infections, with the peak value in Iraq being 4200 and in Egypt 1400.

Figure 5 shows the peak values selected for infections, which are 4200 and 1400 in Iraq and Egypt, respectively.

The return level for the peaks corresponding to the selected threshold for 2020 and 2021 is shown in Fig. 6. Over the 2021 period, it indicates that 4434, 4468,

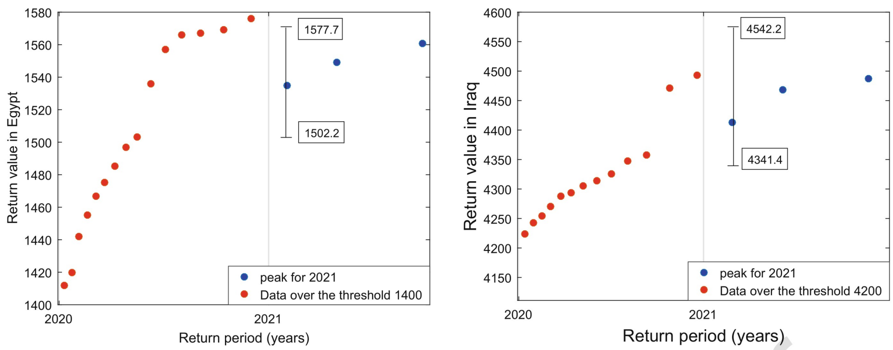


Fig. 6 Return periods and return levels for Egypt and Iraq, 2020–2021

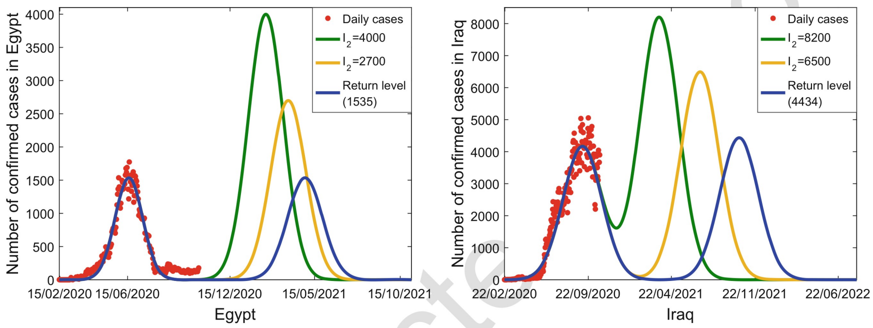


Fig. 7 Two different scenarios with return level to the daily confirmed cases in Egypt and Iraq, 2020–2021

and 4498 infection cases per day are expected to be exceeded in next year in Iraq with confidence intervals (4341.4, 4564.2), (4321, 4704.2), and (4302, 4858.5), respectively, while 1534 (1502.2, 1577.7), 1549 (1511.2, 1598), and 1560 (1523.6, 1661.7) infection cases per day are expected to be exceeded once in the next year in Egypt. The upper and lower confidence intervals for peaks 4468 and 4498 in Iraq and 1549 and 1560 in Egypt indicate low precision and high uncertainty, while the confidence intervals to the peaks 4434 and 1534 for Iraq and Egypt, respectively, revealed narrower and less uncertainty. To predict the spread of COVID-19 in Iraq and Egypt, we apply the Gaussian2 model (5) to estimate the value and time of the expected peak for two different scenarios and estimate the time of the peak that we obtained from return level. Figure 7 shows the daily cases with three expected maximum peak values at its timing in Iraq and Egypt. Table 3 shows the parameters that were used to obtain each scenario and return level estimation. The return level peak timing is estimated to occur on 12 October 2021 with $R^2 = 0.9574$ for Iraq, while on 18 April 2021 in Egypt with $R^2 = 0.9578$. The second wave peak timing is estimated to occur between 21 March and 4 July, 2021 in Iraq, while in Egypt it is estimated to occur between 17 February and 29 March, 2021.

Table 3 Estimated parameter results for two scenarios and return level of the Gaussian model to Iraq and Egypt

Parameters	Gaussian2 model		
	Scenario one	Scenario two	Return level
Iraq			
Estimated peak day cases	8200	6500	4434
Estimated peak date	21/3/2021	04/07/2021	12/10/2021
Goodness of fit (R^2)	0.9675	0.9544	0.9574
Root-mean-square error (RMSE)	364.8	364.9	365
Egypt			
Estimated peak day cases	4000	2700	1535
Estimated peak date	17/02/2021	29/03/2021	28/04/2021
Goodness of fit (R^2)	0.9498	0.9779	0.9578
Root-mean-square error (RMSE)	111.3	111.4	111.7

4 Discussion

We have studied the spread of COVID-19 epidemic in Egypt and Iraq by using compartmental (generalized SEIR) model considering presymptomatic, mildly, and severely infected individuals. We estimated the parameters that best fit the incidence data. Our model provides a reasonable good fit to the incidence data in both countries.

The reproduction number was estimated based on the cumulative confirmed cases by using the exponential growth (EG) method and was found to be 1.078 and 1.047 for Iraq and Egypt, respectively. Using our compartmental model, we obtained a formula for the basic reproduction number that allowed us to calculate the value of $\mathcal{R}_0$. Using the estimated parameter set resulting from fitting our model to the incidence data in both countries, we found that $\mathcal{R}_0 = 1.122$ and $\mathcal{R}_0 = 1.129$ for Iraq and Egypt, respectively. The basic reproduction number is greater than one, indicating that the virus still persists in both countries. The highest value of the effective reproduction number is estimated to be about 1.129 in Egypt and 1.122 for Iraq (see Fig. 4). The contour plots of the basic reproduction number (see Fig. 3) suggest that to control the spread of the COVID-19 outbreak, both countries should work to decrease the transmission rate enough by making more restrictions and precaution measures in the cities that have large numbers of infected people.

The return level for the peaks indicates that infection cases per day are expected to be exceeded once in next year and corresponds to a number of 4434 and 1535 infection cases with narrower and less uncertain confidence intervals in Iraq and Egypt, respectively. The Gaussian2 fit model was used to obtain statistical predictions for the spread of COVID-19 pandemic in Iraq and Egypt, and we fitted the Gaussian2 model to the daily confirmed cases to estimate the value and timing of the expected peak for two different scenarios and to determine the timing of the peak that we obtained from the return level for both countries. The results of the return level in Iraq illustrate that the predicted daily cases are estimated to be 4434, while the peak values of scenario one and scenario two are expected to be 8200 and 6500 on March 21, 2021 and July 4, 2021, respectively. In Egypt, the predicted daily cases are estimated to be 1535, while the peaks of scenario one and scenario two are expected to be 4000 and 2700 on 17 February and 29 March, 2021, respectively.

References

1. World Health Organization (WHO). Coronavirus. https://www.who.int/health-topics/coronavirus#tab=tab_1
2. Centers for Disease Control and Prevention (CDC). Symptoms of Coronavirus. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html
3. M.A. Ibrahim, A. Al-Najafi, Modeling, Control, and Prediction of the Spread of COVID-19 Using Compartmental, Logistic, and Gauss Models: A Case Study in Iraq and Egypt. *Processes.* **8**(11), 1400 (2020).
4. Di, L. Domenico, G. Pullano, C. E. Sabbatini, P. Y. Bolle, and V. Colizza, Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Med.* **18**, 240 (2020).
5. M.V. Barbarossa, J. Fuhrmann, J. Heidecke, H.V. Varma, N. Castelletti, J.H. Meinke, S. Krieg, T. Lippert, A first study on the impact of current and future control measures on the spread of COVID-19 in Germany. *medRxiv.* (2020).
6. N.M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, et al. Report 9—Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, *Imperial College London: London, UK.* (2020). Available online: https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19 (accessed on 31 October 2020).
7. S.M. Moghadas, A. Shoukat, M.C. Fitzpatrick, C.R. Wells, P. Sah, A. Pandey, J.D. Sachs, Z. Wang, L.A. Meyers, B.H. Singer, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc. Natl. Acad. Sci. USA*. **117**, 9122–9126 (2020).
8. J.S. Weitz, COVID-19 epidemic risk assessment for Georgia. *Github: San Francisco, CA, USA.* (2020). Available online: https://github.com/jsweitz/covid-19-ga-summer-2020 (accessed on 31 October 2020).
9. G. Rst, F.A. Bartha, N. Bogya, P. Boldog, A. Dnes, T. Ferenci, K.J. Horvth, A. Juhsz, C. Nagy, T. Tekeli, Z. Vizi, B. Oroszi, Early phase of the COVID-19 outbreak in Hungary and post-lockdown scenarios. *Viruses.* **12**, 708 (2020).
10. J.T. Lim, B.S.L. Dickens, A.R. Cook, Modelling the epidemic extremities of dengue transmissions in Thailand. *Epidemics.* **33**, 100402 (2020).
11. J.T. Lim, Y. T. Han, B. Sue Lee Dickens, L.C. Ng, A.R. Cook, Time varying methods to infer extremes in dengue transmission dynamics. *PLOS Computational Biology.* **16**(10), e1008279 (2020).
12. M. Thomas, M. Lemaitre, M.L. Wilson, C. Viboud, Y. Yordanov, H. Wackernagel, F. Carrat, Applications of extreme value theory in public health. *PloS one.* **11**(7), e0159312 (2016).
13. J. Chen, X. Lei, L. Zhang, B. Peng, Using extreme value theory approaches to forecast the probability of outbreak of highly pathogenic influenza in Zhejiang, China. *PloS one.* **10**(2), e0118521 (2015).
14. M. Aadhityaa, K.S. Kasiviswanathan, I. Ilampooranan, B. Soundharajan, M. Balamurugan, J. He, A global scale estimate of novel coronavirus (COVID-19) cases using extreme value distributions. *medRxiv* (2020).

15. S. Coles, J. Bawa, L. Trenner and P. Dorazio, An introduction to statistical modeling of extreme values. *London: Springer.* **208**, 208 (2001).
16. Worldometer. Available online: https://www.worldometers.info/coronavirus/country/iraq/ (accessed on 31 October 2020).
17. Worldometer. Available online: https://www.worldometers.info/coronavirus/country/egypt/ (accessed on 31 October 2020).
18. M.D. McKay, R.J. Beckman, W.J. Conover, Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics.* **21**, 239–245 (1979).
19. O. Diekmann, J.A.P. Heesterbeek, M.G. Roberts, The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface.* **7**, 873–885 (2010).
20. T. Obadia, R. Haneef, P. Boëlle, The R_0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Med Inform Decis Mak.* **12**, 147 (2012).