Dataset on the COVID-19 Pandemic Situation in Tunisia with application to SIR Model

Authors

Abir Lassoued, Afef Ben Saad, Hela Lassoued, Raouf Ketata[1], Olfa Boubaker [1]

Affiliations

1. University of Carthage, National Institute of Applied Sciences and Technology, Tunis, Tunisia.

Corresponding author(s)

Olfa Boubaker (olfa.boubaker@insat.rnu.tn)

Abstract

April 9, 2020 marks 100 days since first cases of the Coronavirus Disease 2019 in China. In this crucial day with 1 436 198 confirmed infected cases in the world (New confirmed infected cases 82 837) and 85 521 deaths (6286 New daily deaths), the Global Level of the Covid-19 pandemic is evaluated at very high according to the World Health Organization (WHO) situation report. For most people, COVID-19 infection will cause mild illness (fever and at least one sign/symptom of respiratory disease), however for more than 3.4% of people, it can be fatal. Older people and those with pre-existing medical conditions (such as cardiovascular disease, chronic respiratory disease or diabetes) are at risk for severe disease and mortality. The incubation period of the virus is estimated to be between 2 and 14 days but longer incubation are reported. Furthermore, Data published by world authorities shows that statistics are different for different geographical regions and depends on many social and environmental factors. The sad reality of the COVID-19 is that there are currently no medications or vaccines proven to be effective for the treatment or prevention of the disease. The pandemic spread is consequenctly followed by a worldwide panic. Facing this dramatic uncertain situation, implementing a country wide strategy for social distancing and a general logistic policy for critical and life-saving supplies is an urgency for governomental and sanitaire authorities. Several mathematical models have been proposed for developed countries to predict the epidemic spread. However these models can't be adapted to developing countries due to the major differences in geographic, societal, economic and political adopted strategies. Here, we propose the application of the well known SIR model to the case study of Tunisia where Data are collected from three official databases showing differences in reporting methods. The proposed data could be useful to predict the spread of COVID-2019 and design a more reliable model that can estimate the finale size of the spread and help in monitoring infection control.

Keywords

COVID-2019 epidemic; Coronavirus Disease 2019; Mathematical modeling; Epidemic Dynamics; Infection Control; Medical Care in Tunisia, Tunisia.

Specifications Table

Subject	Mathematical modeling of Infectious Diseases
Specific subject area	Mathematical modelling of COVID-2019 to predict spread dynamics, estimate the final epidemic size and control evolution of the infection disease.
Type of data	Chart Graph Figure
How data were acquired	Initial Data are collected from three sources: - The official website of the National Observatory of New and Emerging Diseases of the Ministry of Health of Republic of Tunisia [1]. - The Coronavirus Worldometer [2]. - The World Health Organization(WHO) novel-coronavirus-2019 situation reports [3]. Excel software is used to analyze collected Data. MatLAB software supported by the function fitVirusCV19 [4] is used to generate SIR model [5,6].
Data format	Data are in raw format. An Excel Workbook with row data, figures and statistical data has been collected and analyzed via Excel software. Data have been then uploaded in Matlab to generate SIR Model (figures with extension .tiff).
Parameters for data collection	Parameters of SIR model were : • Transmission rate β • Average recovery rate γ • Population size N • Initial infected cases $I(0)$ These papameters are optimized via the function fitVirusCV19 [4] on the collected data.

Description of data collection	The Data of COVID-19 in Tunisia from March 02, 2020 to April 19, 2020 were collected from three sources [1–3] and an Excel file was used to build a time-series database [7]. The database includes for each souce the total infected cases, the infected cases/day, the total Deaths and the New daily deaths. The daily Tests and the total recovered are only collected from the country official source [1]. Time series are analyzed via Excel and then uploaded in fitVirusCV19 [4] in order to generate the Data related to the SIR Model [5,6].
Data source location	Institution: University of Carthage, National Institute of Applied Sciences and Technology City: Tunis Country: Tunisia
Data accessibility	Raw data can be retrieved from a public repository [7]: Repository name: Mendeley Data Data identification number: 10.17632/v8fhn6wsh2.3 Direct URL to data: https://data.mendeley.com/datasets/v8fhn6wsh2/3

Value of the Data

- These data are useful because they represent a simple approach to evaluate the transmission dynamics of COVID-2019 in developing countries for which very few results are reported.
- Governmental authorities as well public health institutes can benefit from these data since they can predict the epidemic spread and implement a country wide strategy for social distancing and logistic policy of critical and life-saving supplies. Such tools can be also used for medical education and biomedical engineering training.
- The additional value of these data lies in the fact to estimate the country unknown parameters related to the SIR model since these parameters are special for different geographical regions and depend on many social and environmental factors.
- The tool can be very useful to design key strategies to tackle the spread. It can help to verify whether the extensive measures imposed by authorities like social distancing, quarantine, lockdown, and curfew and so on are proficient. The tool also enables the projecting of essential

supply needs, including the estimation of biomedical equipment, drugs for supportive care, and consumable medical supplies.

- The data can be used to design a more reliable model including every societal fact and the entire measures taken by authorities in order to estimate the finale size of the spread and propose an objective tool for monitoring infection control.

Data Description

The Data of COVID-19 in Tunisia from March 02, 2020 to April 19, 2020 were collected from the website of three official sources: the National Observatory of New and Emerging Diseases of the Tunisian Ministry of Health [1], the Coronavirus Worldometer [2] and The World Health Organization (WHO) [3]. An Excel file was used to build a time-series database [7] including for each source the time series related to the total infected cases, the infected cases/day, the total deaths and the new daily deaths/day. The daily tests and the total recovered are only collected from the country official source. In this paper, the times series of the total infected cases are shown by Fig.1 for three sources and exhibiting differences in reported data. Due to differences in reporting methods, retrospective data consolidation, and reporting delays, the number of new cases presents differences between the official source report and the two other sources using different inclusion criteria and different data cutoff. Data are analyzed via statistical functions and shows different standard deviation for the three sources [7]. Fig.2. shows time series of the total infected cases, the total Deaths and the total recovered collected from the country official source whereas Fig.3 shows the time series of the daily Tests and the infected Cases/day. To understand the magnitude of the risk posed by this novel Coronavirus for the country at this date, the Transmission Rate (R_0) computed at the date of April 19, 2020 and representing the number of newly infected people from a single case and the Fatality Rate (CFR) computed for the same date are found [7]:

R_0 (%) = (Total cases/Total tests) *100= 5.11 %

CFR (%) = (Total death /Total cases) *100= 4.29%

SIR model [5,6] is applied to the dataset [7] consisting of 45 number determinations. Fig.4 shows Actual and predicted total number of cases by the SIR model in Tunisia. Fig.5 shows the actual and predicted daily number of cases. Though the estimated decline begins on April 14, the curve shows a tail till at May 10, 2020 for zero new infection. The estimated basic reproduction number $\bar{R}_0$ is equal to 2.025 where the epidemic peak have occurred on or around first April 2020.

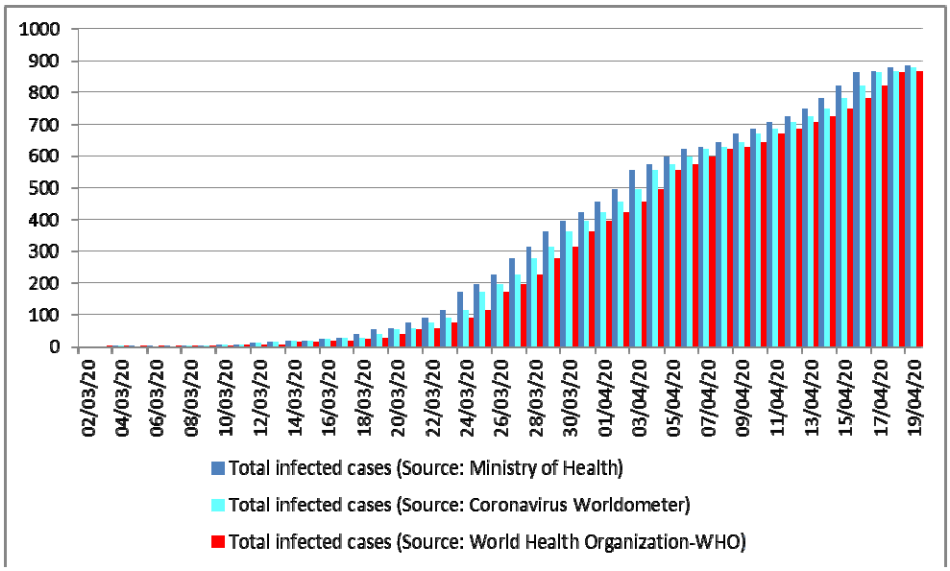

Figure 1. Times series of the total infected cases reported from March 02, 2020 to April 19, 2020.

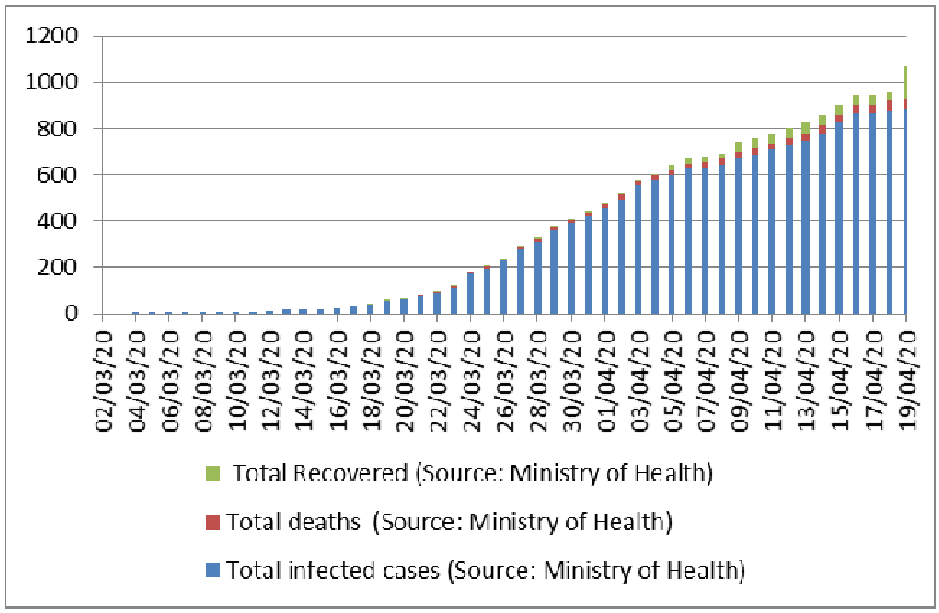

Figure 2. Times series of the total infected cases, total deaths and total recovered reported from March 02, 2020 to April 19, 2020.

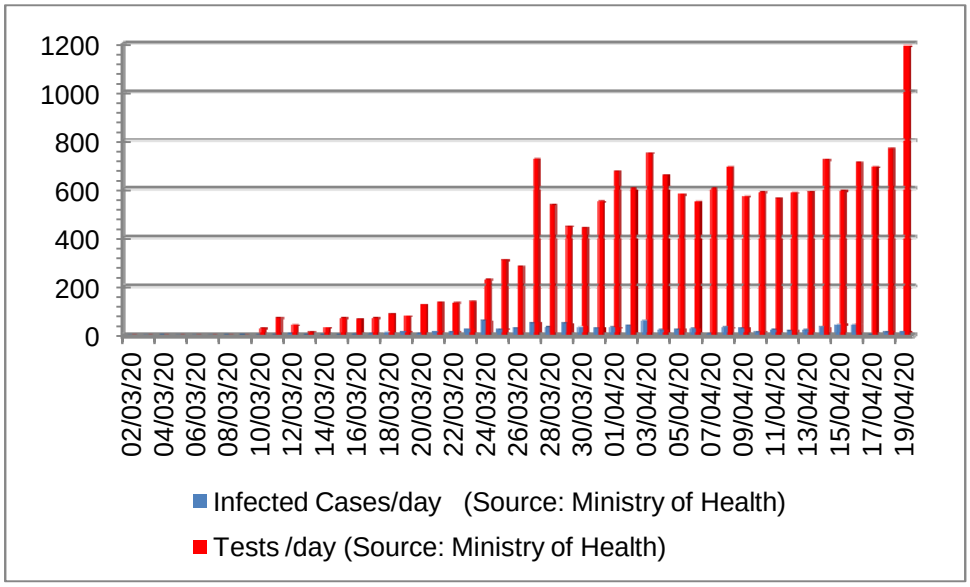

Figure 3. Time series of the daily Tests and the infected Cases/day reported from March 02, 2020 to April 19, 2020.

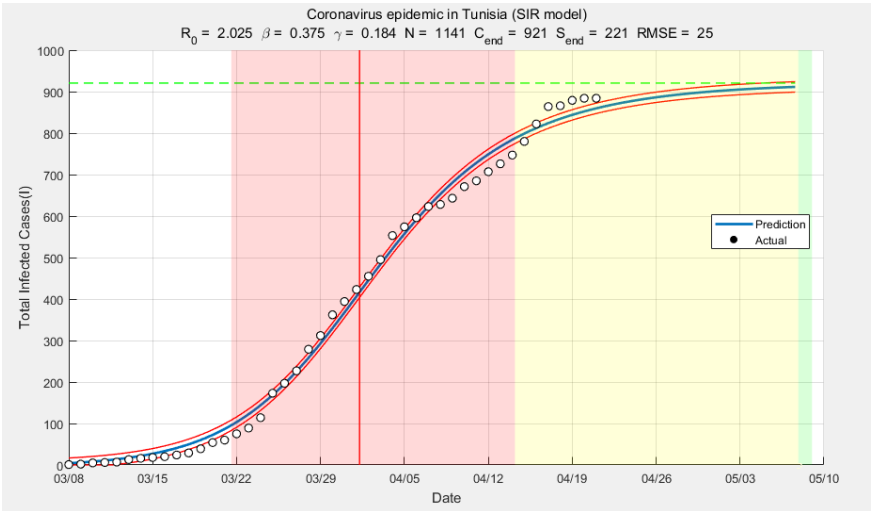

Figure 4. Actual and predicted total number of cases by the SIR model in Tunisia.

C_{end} represents the epidemic size and S_{end} represents the number of susceptible individuals left. RMSE is the Root Mean Square Error.

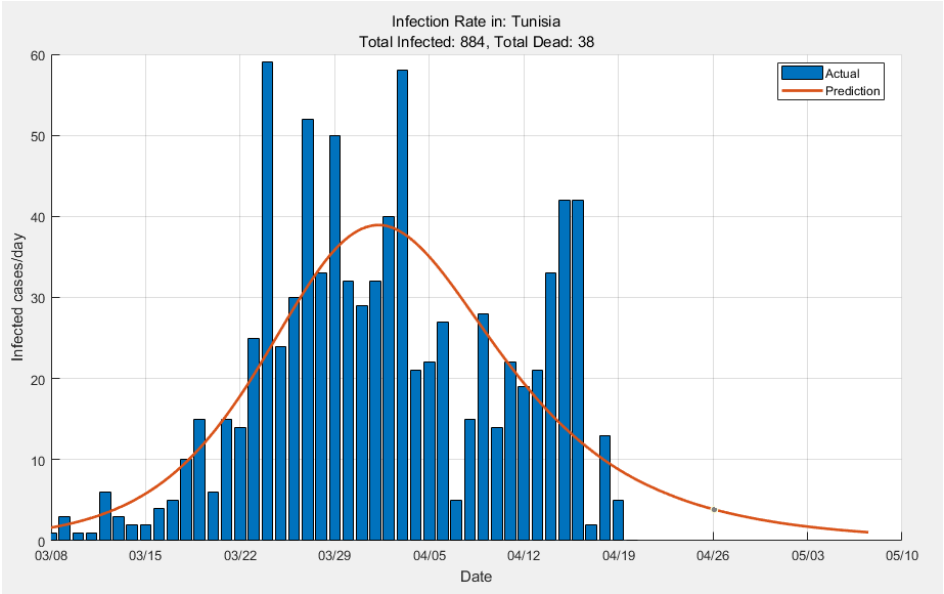

Figure 5. Actual and predicted Daily number of cases by the SIR model in Tunisia.

Experimental Design, Materials, and Methods

Susceptible-Infectious-Recovered (SIR) model is a compartmental model described by three ordinary differential equations that can estimate the number of susceptible (S), the number of infectious (I), and the number of recovered patients (R) [5,6]. The model is characterized by four parameters : the transmission rate β, the average recovery rate γ, the population size N and the Initial number of infected cases $I(0)$. The model is considered under a number of assumptions; N is constant and not represents the population size. It is estimated as $N= S + I + R$; $R(0)$ is assumed zero and the final number of infected people is also assumed to be zero. For the implementation of the SIR model an open-source MATLAB code, the *fitViruscv19v3* [4], is used in order to provide graphical visualizations and short-term forecasts. First of all, collected data from the database [7] are introduced in the MatLAB code to substitute the existing data in the program. Then, estimation of the parameters and initial values of the SIR model using the *fitViruscv19v3* is done by minimizing the differences between the actual and predicted numbers of infected cases. The code uses the predefined MATLAB function *fminsearch* to solve the dynamic optimization problem. The parameters of the model are obtained by minimization of the objective function, which is the sum of squares for residuals of values and sum of squares for residuals of values differences. The weights of summands are selected automatically. Optimization Toolbox via the fminsearch function is used to compute optimal values of unknown model parameters. The SIR differential system is finally solved using MATLAB function *ode45*.

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

References

[1] National Observatory of New and Emerging Deseases of (Ministery of Health of Republic of Tunisia). covid-19.tn. Ministery of Health 2020. https://covid-19.tn/fr/tableau-de-bord/ (accessed April 19, 2020).

[2] Wordometer. worldometers.info. COVID-19 Coronavirus Pandemic 2020:https://www.worldometers.info/coronavirus/. https://www.worldometers.info/coronavirus/ (accessed April 19, 2020).

[3] World Health Organization (WHO). Novel Coronavirus (2019-nCoV) situation reports. WHO Bulletin 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reportsports (accessed April 19, 2020).

[4] Milan Batista. fitVirusCOVID19. MATLAB Central File Exchange 2020. https://www.mathworks.com/matlabcentral/fileexchange/74658-fitviruscovid19 (accessed April 19, 2020).

[5] Hethcote HW. The Mathematics of Infectious Diseases. SIAM Review 2000;42:599–653. https://doi.org/10.1137/S0036144500371907.

[6] Batista M. Estimation of the final size of the COVID-19 epidemic. MedRxiv 2020. https://doi.org/10.1101/2020.02.16.20023606.

[7] [Data] Lassoued A, Ben Saad A, Lassoued H, Ketata R, Boubaker O. Dataset on the COVID-19 Pandemic Situation in Tunisia with application to SIR Model. Mendeley Data 2020. https://doi.org/10.17632/v8fhn6wsh2.3.