An investigation of three mathematical models on the COVID-19 pandemic: the case of Turkey

Engin Can[1*]

[1]Sakarya University of Applied Sciences, Faculty of Technology, Department of Basic Sciences of Engineering, 54050 Serdivan/Sakarya, Turkey

Abstract. Mathematical modelling plays a major role in assessing, controlling, and forecasting potential outbreaks. This article consists of implementing mathematical forecasting models by Liang [14], Ma [16], and the curve fitting method with the least squares as a standard approach in regression analysis and compares those calculate results obtained from the Covid-19 outbreak data, which is announced by the Turkish Ministry of Health, between March 10, 2020, and May 15, 2020, in Turkey. For this purpose, we presented an overview of three mathematical models and finally, demonstrated their applications using the Ministry of Health of Turkey's publicly reported data.

1 Introduction

Mathematical modelling in epidemiology started with Bernoulli [1], and his model extended by Lambert [2] and Laplace [3]. However, the first study of the modern mathematical epidemiology was presented by Ross [4]. McDonald extended Ross's model by utilizing modern computer power in which the mathematical model for the dynamics and the control of mosquito-transmitted pathogens provided robust results in real-world applications as Ross–MacDonald models [5]. The first book on mathematical modelling of epidemiological systems is published by Bailey [6]. Moving to the 21st century, a mathematical model on H7N9 influenza, domestic poultry, and humans in China is introduced by Xing et al. [7] and for Zika virus, using logistic growth in human populations is developed by Lee and Pietz [8]. Moreover, a complex dynamical behaviour in epidemiological models is identified by Yu and Lin [9]. A susceptible-exposed-infectious-recovered (SEIR) model to describe the transmission dynamics is introduced by Wu et al. [10]. For the basic reproductive number based on data fitting of a SEIR model is reported by Read et al. [11].

Recently some mathematical models tried to describe the dynamics of Covid-19 was investigated. Some phenomenological models were presented by Roosa et al. [13], which were validated by outbreaks of other diseases different from Covid-19. Liang [14] presented a simple mathematical model with low percent error rates in which a 2-term ordinary differential equation that real cases and simulation result for Covid-19 in four different cities of China were implemented. Ivorra et al. [15] offer a novel approach that considers the fraction θ of detected cases over the real total infected cases, which allows studying the

---

*Corresponding author: ecan@subu.edu.tr

importance of this ratio on the impact of Covid-19. They also introduced a method to estimate some of the unknown parameters of their model, in which they proposed a study of the impact of the percentage of detection of cases and obtained that the magnitude of the pandemic can be drastically reduced when increasing this percentage. Ma [16] investigated a 3-term linear logarithmic equation, which offers the simplest mathematical method. Yang and Wang's model [17] consist of five ordinary differential equations and it has also multi-parameter like θ-SEHIRD model and provides numerical solutions with three linear equations. Chatterjee et al. [18] presented a stochastic mathematical model and investigated in India. They developed SEIR model, using a set of differential equations, which consist of six ordinary differential equations.

The rest of this paper is organized as follows: In the next section, we present an overview of two new mathematical models [14, 16] and the curve fitting method. We conduct a brief mathematical analysis and compare the results. In Section 3, we apply numerical simulations by incorporating the infection of all models for Turkey's publicly reported data [19]. Finally, we conclude the paper with some general results and discussion.

2 Preliminaries

In this section, we summarized the mathematical models taken directly from the published papers [14, 16]. The method of the least squares is a standard approach in regression analysis to approximate the solution of overdetermined systems by minimizing the sum of the squares of the residuals made in the results of every single equation, which is the most important application of the curve fitting [20]. We investigated several polynomial and logarithmic functions to interpret the corresponding coronavirus data.

2.1 The Liang's model

Liang [14] was given a mathematical modelling that establishes propagation growth model in which examines the number of people infected with daily values with following equation:

$$\frac{dN}{dt} = r_0 \left(1 - \frac{N}{N_{max}}\right) N \tag{1}$$

where
- $N(t)$ is a time t = day expressed the daily number of infectious cases ($N(t_0) = N_0$)
- N_{max} is the maximum number of $N(t)$,
- r_0 is the growth rate before any measure was taken,
- r_0N, of differential equation (1) expresses the natural pandemic trend of infectious diseases in the absence of any prevention and control measures,
- $-r_0N^2/N_{max}$, shows the effect of prevention and control measures for communicable diseases.

Liang presented a simple mathematical model with low percent error rates in which a 2-term ordinary differential equation (1) and the simulation of the model consists of the solution of this equation. The success of the model is achieved when the number N reaches $N_{max}/2$. In other words, this model is useful if the N_{max} is known exactly.

2.2 The Ma's model

Ma [16] used the PLEC (power law with exponential cutoff) model for the growth of Covid-19 infections in which the number of the cumulative value of infected people examined with the following log-linear equation:

$$ln\, I = ln\, c + w ln\, T + dT \tag{2}$$

where

- T is days (starting from a specified date),
- I is the number of cumulative infections,
- $c,\, d,$ and w are three parameters to be estimated from the observed data. c is the initial infection number, d is of particular interests [see 16], w is the growth rate of infections over time.

   Ma's model has a 3-term linear logarithmic equation (2) and reported$< 5\%$ error rate, which offers the simplest mathematical method. It is the simplest model among other models and the error rate is low.

2.3 The curve fitting model

According to Rolle's Theorem [21], we consider the pandemic process as a single peak with a 2nd degree polynomial. The several peak values can be expressed using 3rd and higher degree polynomials. A suitable 2nd degree polynomial function and the logarithmic function in R^2 calculated by the method of least squares as follows:

2.3.1 *2nd degree polynomial function*

Let

$$f(x) = y = c_1 + c_2 x + c_3 x^2$$

be the 2nd degree polynomial where $c_1,\, c_2,\, c_3 \in R$. Then

$$E(c_1, c_2, c_3) = \sum_{k=1}^{n} (c_1 + c_2 x + c_3 x^2 - y_k)^2$$

is minimum where y_k are the forecast values for all $k = 1, ..., n$.
   Thus, by using the necessary conditions

$$\frac{\partial E}{\partial c_1} = 0, \qquad \frac{\partial E}{\partial c_2} = 0, \qquad \frac{\partial E}{\partial c_3} = 0$$

linear equation system is obtained. Solving these linear equations system gives $c_1,\, c_2,\, c_3.$
   Further information on 3rd and higher degree polynomial functions with curve fitting can be found in [20].

2.3.2 *Logarithmic function*

We can give this curve fitting model, which is used when the relationships between dependent and independent variables are not linear according to the parameters, as follows:

Let

$$f(x) = y = a_1 e^{a_2 x}$$

be the logarithmic function with the least squares.

$$E(a_1, a_2) = \sum_{k=1}^{n} (a_1 e^{a_2 x_k} - y_k)^2$$

is minimum where y_k are the forecast value for all $k = 1,...,n$.

Thus, by using the necessary conditions

$$\frac{\partial E}{\partial a_1} = 0, \qquad \frac{\partial E}{\partial a_2} = 0$$

we obtain following nonlinear equation system:

$$\begin{cases} a_1 \sum_{k=1}^{n} x_k e^{2(a_2 x_k)} - \sum_{k=1}^{n} x_k y_k e^{a_2 x_k} = 0 \\ a_1 \sum_{k=1}^{n} e^{2(a_2 x_k)} - \sum_{k=1}^{n} y_k e^{a_2 x_k} = 0 \end{cases}$$

For solving these nonlinear equations system by taking logarithm we obtain

$$ln y = ln a_1 + a_2 x$$

with substitution *Y=ln y, A=ln a₁*

$$Y = A + a_2 x.$$

3 Applying all models to the same data

In this section, we apply all models to the reported data by the Turkish Ministry of Health [19].

3.1 Application of the Liang's model

The first Covid-19 infection was detected in Turkey on March 10, 2020. After the first case detection, the government started regular Covid-19 tests and the first results were released on March 17, 2020. As it is seen in Fig. 1, the number of daily-infected people reached its peak on April 11, 2020.

For application of this model, r_0 must be calculated in advance. Hence, we need a daily growth rate before March 17, 2020. This is 7th day with 47 daily infected population number in [19], thus we obtain

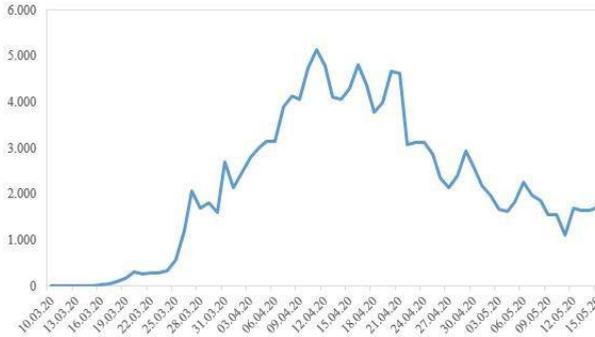$$r_0 = \left(\frac{47}{1}\right)^{\frac{1}{7}} - 1 \cong 0.73.$$

Fig. 1. Daily infected population number in Turkey.

This coefficient allows us with r_0N, how much more infected people would be per day if no action was taken. For example, on the 35th day, the cumulative number of infected people would be 105797 instead of 61049.

As reported in [19], t_{32} so on the 33rd day is the daily infected number reached the maximum value on April 11, 2020, with $N_{max} = 5138$. By substituting in eq. (1),

$$\frac{dN}{dT} = r_0N - \frac{r_0N^2}{N_{max}} = 0.$$

3.2 Application of the Ma's model

This model estimates the peak value where the T_{max} and I_{max} values of infections are most monotonically increased.
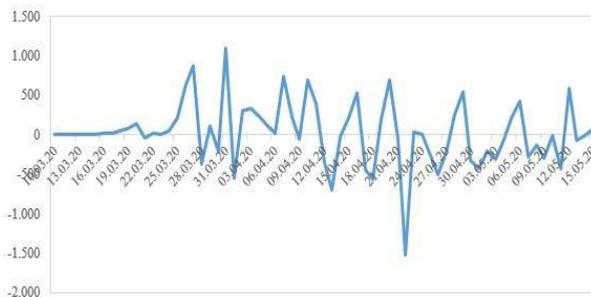

Fig. 2. Daily infected population difference in Turkey.

According to [19] and Fig. 2, the most monotonically increased values are on 22nd day, i.e. March 31, 2020. Hence, $T_{max} = 22$ and $I_{max} = 2704$. With substituting (see [16] eqn. (6))

$$\ln 13531 = \ln 2704 + w \times \ln 22 \Rightarrow w \cong 0.52$$

and, from (2) we obtain

$$\ln 13531 = \ln 2704 + 0.52 \times ln\ 22 + d \times 22 \Rightarrow d \cong 0.0001.$$

As in [19], the infected number reaches its peak on the 33rd day, with substituting

$$\ln I = \ln c + w \times ln\ T + d \times T$$

$$\ln I = \ln 2704 + 0.52 \times ln\ 33 + 0.0001 \times 33 \cong 9.72$$

$$I = e^{9.72} \cong 16647.$$

Hence, error rate by Ma

$$E(\%) = (I_{max} - I)/I_{max} = (16647 - 13531)/16647 \cong 0.18$$

and finally, Ma's model provided approximately

$$P(\%) = 1 - E(\%) = 1 - 0.18 = 0.82 = 82\%$$

accuracy rates on the Turkish reported data.

3.3 Application of the curve fitting method

From the reported data of the National Health Commission (NHC) of Turkey between March 17, 2020 and May 15, 2020 as seen in Fig. 3, we obtain 2nd degree polynomial with the least squares applying daily infected population number as
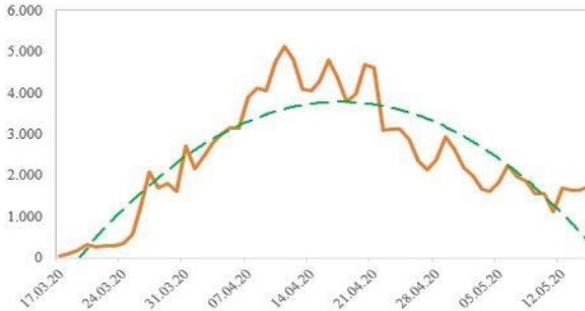
$$y = -4.4x^2 + 286.2x - 840.2$$

Fig. 3. 2nd degree polynomial curve fitting of estimating peak value.

that the date of the peak value calculated as follows

$$\frac{-c_2}{2c_3} \cong 33th\ days.$$

with substitution this value in 2nd degree polynomial we obtain

$$\text{Infected population number in peak value date} = N\ (33) \cong 3813.$$

The second zero point of the parabola above is May 17, 2020, which is the approximate end date of the pandemic.

Now we study what happens if the strict measures will not be taken:

It is obvious that the number of infected people increases as well as 3rd or higher degree polynomial with several peak values as displayed in Fig. 4. So we obtain 3rd degree polynomial with the least squares as
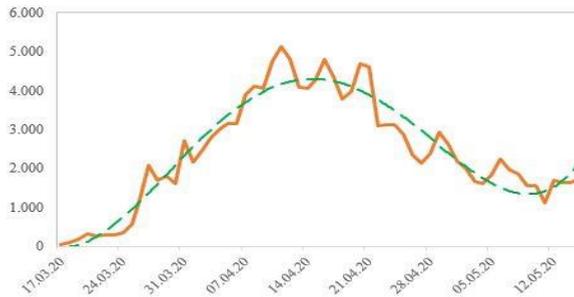
$$y = -0.8x^3 + 25.5x^2 - 65.7x + 263$$

Fig. 4. 3rd degree polynomial curve fitting of estimating if no measures are taken in Turkey.

On the other hands, it is possible to work the whole through

$$\mu = \frac{\text{total infected}}{\text{total recovered}}.$$

If the value of μ approximately equal to 1 or not which also suggest whether the measures were taken to reduce the spread of the outbreak.

$\mu = 1$ means all infected population is recovered at the moment also here are no intubated, dead, quarantined, exposed more. In Fig. 5, the situation of Turkey was demonstrated. (From the reported data of NHC of Turkey between March 27, 2020 i.e. 20 pandemic day and May 15, 2020).

The nonlinear function with the least squares is

$$y = 78.858e^{-0.092x}$$

$$ln\, y = -0.092x + ln\, 78.858$$

$$Y = -0.092x + 4.367$$

and solving the equation by substitution of $Y = \mu = 1$

$$1 = -0.092x + 4.367$$

we obtain

$$x \cong 37$$

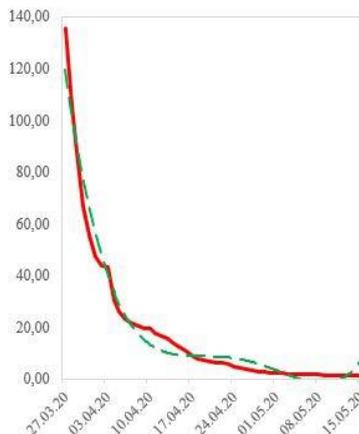that means 57 days later, which is on May 3, 2020, this outbreak would have ended.

Fig. 5. The ratio of total infected/total recovered with curve fitting.

4 Results and conclusion

In this study, we presented three mathematical models, which were used to investigate the Covid-19 outbreak that started in Wuhan, China, and spread to the world in a short time period. The growth rate of infectious diseases determines their prevalence in the early stage, and the infection inhibition constant depends on the prevention and control measures adopted by different regions. On the other hand, the success of presented mathematical methods is evident in revealing approximately, when the outbreak will end, if the strict measures are taken and followed by the countries. Therefore, the analysis of the models reasonably explains the characteristics of their transmission and the measures taken in different places during the outbreak.

For this purpose, we provided brief detailed results of all models. The most important difference between the Liang's and Ma's models is that Liang worked with the daily number of infected people that estimates the end time of the outbreak if the peak is reached, and Ma worked with the cumulative number of infected people that estimates the peak using the values where the number of infections increased monotonically.

On the other hand, outbreak data can be compared and interpreted using the curve fitting model by adding also susceptible/infected, susceptible/recovered ratios.

Finally, all three models were applied to the Turkish Ministry of Health's publicly reported data. The obtained results are given in Table 1.

Table 1. Expected daily and cumulative peak dates.

Expected daily and cumulative peak dates, if all measures were taken	
Daily peak value by Liang's Model	11 April 2020
Cumulative peak value by Ma's Model	15 May 2020
Daily peak value with 2nd polynomial curve fitting	17 May 2020
Daily peak value with logarithmical curve fitting	3 May 2020

References

1. D. Bernoulli, Math. Phys. Acad. Roy. Sci. **1** (1766)
2. J.H. Lambert, Buch. der Realsch. **3** (1772)
3. P.S. Laplace, Theorie analytique des probabilites (Paris, 1812)
4. R. Ross, The prevention of malaria (John Murray, London, 1911)
5. S. Blower and D. Bernoulli, Rev. Med. Virol **14** (2004)
6. N.J.T. Bailey, The Math. Theo. of Inf. Dis. and its Appl. (Griffin, London, 1957)
7. Y. Xing, L. Song, G.Q. Sun, Z. Jin, J. Zhang, Appl. Math. Comp. **309** (2017)
8. E.K. Lee, Y. Liu, F.H. Pietz, *A compartmental model for zika virus with dynamic human and vector populations*, in Proceedings of American Medical Informatics Association, AMIA Annual Symposium (2016)
9. P. Yua, W. Linb, Jour. Biol. Dyn. **10**, 1 (2016)
10. K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J.M. Hyman, P. Yan, G. Chovell, Inf. Dis. Model. **5** (2020)
11. J.T. Wu, K. Leung, G.M Leung, Lancet **395** (2020)
12. J.M. Read, J.R.E. Bridgen, D.A.T. Cummings, A. Ho, C.P. Jewell, medRxiv (to be published)
13. N. Imai, A. Cori, I. Dorigatti, M. Baguelin, C. A. Donnelly, S. Riley, N. M. Ferguson, (Imperial College London Source 2020)
14. K. Liang, Inf. Gen. and Evol. **82** (2020)
15. M.R. Ivorra, M. R. Ferrández, M. Vela-Pérez and A.M. Ramos, Commun. Nonlin. Sci and Numer. Simul. **8** (2020)
16. Z. Ma, MedRxiv (to be published)
17. C. Yang and J. Wang, Math. Biosci. and Eng. **17,** 3 (2020)
18. K. Chatterjee, K. Chatterjee, A. Kumar, S. Shankar, Med J. Arm. For. India **76**, 2 (2020)
19. Turkish Ministry of Health: https://covid19bilgi.saglik.gov.tr/tr/
20. E.S. Türker and E. Can, Bilgisayar uygulamalı sayısal analiz yöntemleri (Değişim Yayınları, Sakarya, Turkey, 2007)
21. G.B. Thomas, Thomas Calculus (12th edition, Addison-Wesley, USA, 2006)