A SEIRD Model for Control of COVID-19: Case of Azerbaijan

Tarana Aliyeva [1], *Ulviyya* Rzayeva [1*], and *Reyhan* Azizova [1]

[1] Azerbaijan State University of Economics, Department of Digital Economy and ICT, Istiglaliyyat 6, AZ1001 Baku, Azerbaijan

Abstract.

Research background: The study uses the key parameters of the spread of the epidemic, dividing the population into several groups S - susceptible, E - exposed, I - infectious, R - recovered, D - dead. It is found that the model behaves differently depending on the R_0 indicator - the average number of people that one infected manages to infect. Measures to suppress the epidemic undertaken by Azerbaijan and their effectiveness have been considered.

Purpose of the article: The aim of the article is to model the current dynamics of the disease for future forecasting. The model takes into account all the main parameters of the epidemic: the proportion of severe patients and the mortality rate depending on the age of the patients, the duration of the incubation period and the infectious phase of the disease; incomplete registration of infected people due to the high prevalence of asymptomatic disease and insufficient testing; possible measures to contain and suppress the epidemic and their impact on R_0.

Methods: The article uses the linear regression method, which consists in finding estimates of unknown parameters and the formation of a functional relationship between the sickness rate and the factors determining it.

Findings & Value added: The constructed model analyzes the growth of patients in the country after removing the restrictive measures taken in early May on the basis of real statistics.

Keywords: *COVID-19; SEIRD; regression model; differential equations; exponential growth*

JEL Classification: *C150; C510; C670*

[*] Corresponding author: ulviyya.rzayeva@unec.edu.az

1 Introduction

The world is swept by a pandemic of the cruel and dangerous coronavirus, capable infecting more than half of humanity and killing tens of millions of people. Scientists around the world investigate a new Chinese virus. Moreover, this virus is studied not only by virologists, microbiologists and physicians, but also by representatives of other areas of world science. The focus is on the impact that COVID-19 can have on biota, flora and fauna, as well as on various spheres of human life.

This article discusses the SEIRD model, which reasonably describes the spread of disease in humans [1]. The study covers the period from May 4 – the day of the lifting of severe restrictive measures in Azerbaijan – to June 30.

2 Research methodology

There are various approaches to modelling the processes occurring in complex systems. One of the most common methods is simulation based on systems of differential equations describing both individual subsystems and the object as a whole [2]. For numerical modelling, the article uses an extrapolation method based on the implicit Euler method for the numerical solution of differential equations [3, 4].

Since the most convenient in this case is the tabular way of entering and visualizing calculations [5, 6, 7], the authors used Excel spreadsheets with a high level of functionality, as the suitable tool for calculations.

This study of the forecasting the exponential growth of people infected with coronavirus in retrospective time by SEIRD method is based on a multivariable regression model.

The information framework of the research is largely based on data published by specialists from the WHO, the Chinese, American and European Centers for Disease Control, as well as materials on the treatment and prevention of this infection [8].

3 Building a SEIRD Epidemic Model for Azerbaijan

The SEIRD epidemic model belongs to the class of compartmental models, the essence of which is to divide the population into several groups, in our case: S – susceptible, E – exposed, I – infected, R – recovered, D – dead.

The modified version of the SEIRD model is presented as follows:

$$
\begin{aligned}
dS/dt &= -\beta S(I+\theta E)/N \,, \\
dE/dt &= \beta S(I+\theta E)/N - kE \,, \\
dI/dt &= kE - \gamma I - \mu D \,, \\
dR/dt &= \gamma I \,, \\
dD/dt &= \mu I \,,
\end{aligned}
\tag{1}
$$

where β is infection rate, γ is recovery rate, k is rate of transition of the disease from the incubation stage to the open, θ is contagion rate of latent carriers of infection. Assuming the illness lasts an average of 14 days (at least as long as the mild form lasts, which accounts for up to 80% of cases), find the value $\gamma = 1/14 = 0.0714$. Let's accept $\beta = 3/14 = 0.2143$. Taking into account the average duration of the incubation period of 3 days, we take $k = 1/3 = 0.33$. The value of $\theta = 0.6$ we borrow from [1]. Parameter μ – death rate – determined by

modelling (μ=0.012), since this parameter depends on the strengthening measures taken to combat the epidemic and various factors related to human health [9, 10].

Let us describe the real situation generated by the COVID-19 pandemic in Azerbaijan in the considered period from May 4 to June 30 (Fig. 1):
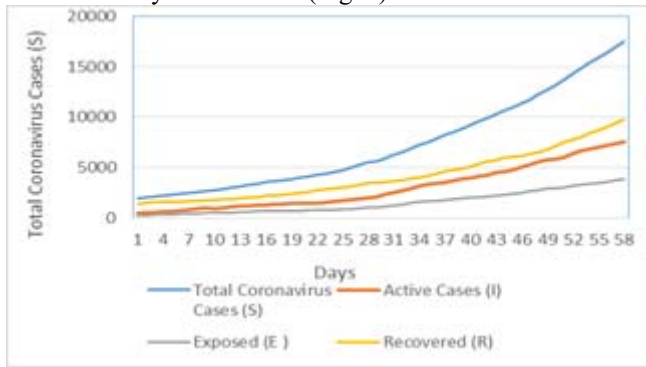

Fig. 1. The number of active cases I, exposed E and recovered R from May 4th to June 30th.

Starting from June 20, the share of active cases in the total population during the week increased from 3.04% to 4.27%. This warns of COVID-19's further spread in the republic (Fig.2):
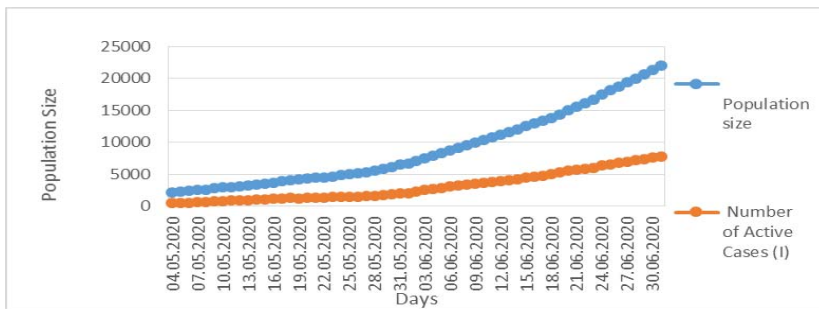

Fig. 2. Behaviour of Number of Active Cases I in relation to the considered population N.

Simulating the behaviour of the virus spread by differential equations' systems (1) is carried out by the numerical Euler method with given initial values and a step of 1 day in the considered period [11]. As the initial conditions, we use the data on Azerbaijan as of May 4 [12]:

$S_0 = 1984$,
$I_0 = 478$,
$E_0 = 0.5*I_0$

To further model the rate of infection spread, we introduce the variables N_d – the number of cases on a particular day.

The probability of infection from each contact is calculated using the formula:

$$p = (1 - {N_d}/{pop.size}) \cdot \qquad (2)$$

Then ΔN_d – the number of new cases during the day:

$$\Delta N_d = E \cdot p \cdot \Delta N_d$$

Therefore,

$$\Delta N_{d+1}=N_d+E{\cdot}p{\cdot}N_d=(1+E{\cdot}p)\ N_d \rightarrow N_d=(1+E{\cdot}p)^d{\cdot}N_0 ,$$

and the growth factor of the disease will look like this:

$$Growth\ factor = {}^{\Delta N_d}\!/_{\Delta N_{d-1}} \tag{3}$$

By the end of June, due to the increase in the number of new cases, the probability (2) that each contact would become an infection was equal to one (Fig.3). This means that the development of such probability of infection can lead to the peak of the first wave and a smooth transition to the second.
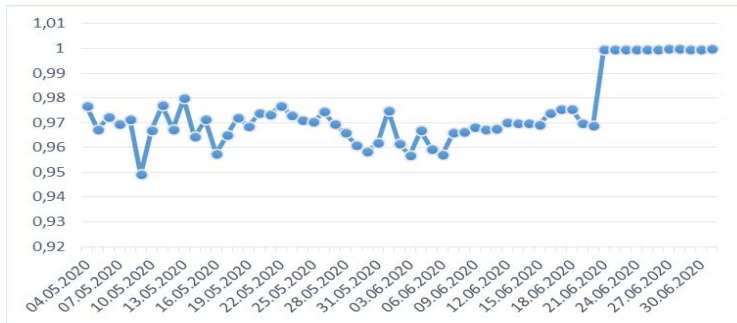

Fig. 3. The probability of each contact with an infected person becoming an infection.

The rate of infection spread depends on many factors, including the variable *Growth factor*. For the situation with a pandemic in Azerbaijan, the calculated growth factor using formula (3) is 1.05 on average, i.e. above one, which signals exponential growth.

Next, in our model, we need to calculate the spread of coronavirus R_t – an indicator that determines the average number of people that one patient infects before isolation. It is calculated based on data on the increase in new cases for the last 8 days.

The spread of coronavirus is calculated using the formula:

$$R_t = (N\_d_8+N\_d_7+N\_d_6+N\_d_5)\ /\ (N\_d_4+N\_d_3+N\_d_2+N\_d_1),$$

where N_{di}, $i = 1,\ldots,8$ – number of registered patients for the corresponding day.

The relationship between *Growth factor* and R_t is shown in the following graph (Fig. 4):
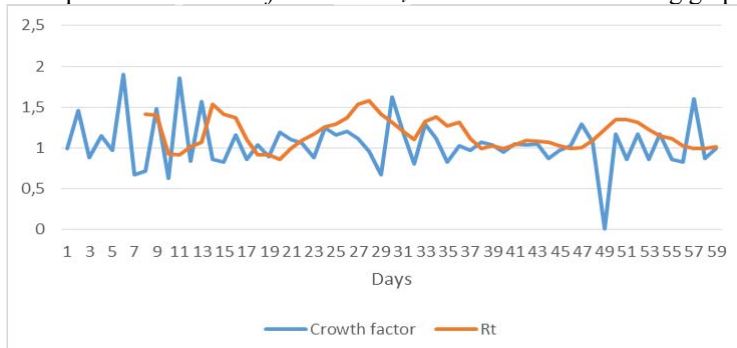

Fig. 4. Description of the relationship between *Growth factor* and the spread of coronavirus.

R_t shows how many people on average manage to infect one person before he is isolated. If it is equal or less than 1, the authorities can begin to ease quarantine measures. If the coefficient drops to 0.8 or less, the region can proceed to the second stage, with a

value of 0.5 or less - to the third. As can be seen from the graph (Fig.4), for the considered period in Azerbaijan, this coefficient is on average 1.17, that is, more than 1.

The next step in predicting the values of a variable, we use a regression analysis toolbox [13, 14]. Its goal is to develop a statistical model to predict the values of the dependent variable or response from the values of at least one independent or explanatory variable [15].

Consider a linear regression model:

$$y=a+b*t+u, \tag{4}$$

where a – regression shift, b – regression slope, u – random error.

The performed analysis shows that the calculated value of the function y describes its initial value relatively well. To determine the linear relationship between the number of daily observations of the infection spread and the corresponding day using the linear regression method (Fig.7), the values in (4) were found: $a = -4.3185$ and $b = 9.285038$. Let's explain the meaning of these values. The calculated slope $b = 9.285038$ means that as the variable t increases by one, the mean value of the variable y increases by 9.285 units.

The interpretation of the parameter a equal to -4.3139 should be treated with care. This regression value is the mean of the y variable at $t = 0$. This means that if the number of periods is zero, the number of infections per day will decrease by 4% more. Since the number of periods cannot be zero, the shift can be considered as the proportion of all infections, depending on other factors (intensity or communications' development, the nature of restrictive measures, weather, and geographic location).

Using a trend line

$$y=9.285t-4.3139 \tag{5}$$

it is possible to predict the situation by the number of daily new infections in the considered observation period - from May 4 to June 30, where the mixed correlation coefficient $R^2 = 0.9277$ estimates the proportion of variation in the variable y, which is explained by the independent variable t in the regression model. R^2 – the value of the reliability of the approximation is close to unity (Fig.8) and indicates a strong positive linear relationship between the two variables. Only 7.2% (the result of the mixed correlation coefficient) in the sample of daily infection rates is due to other factors not accounted in the regression model.

With the help of (5), it is possible to predict the behaviour of the infection spread for another 50 days in advance. Our model predicts the number of infections for that day:

$$y (108) = 998.$$

Given that linear regression represents a more accurate forecast for 2 weeks ahead, we have 72 days after the abolition of strict quarantine:

$$y(72) = 664.$$

4 Discussion and conclusion

The spread of infectious diseases is a complex phenomenon with many interacting factors. The key role of mathematical epidemiology is to create models for the spread of pathogens. These models serve as a mathematical basis for understanding the complex dynamics of disease spread.

The paper considers the SEIRD model of the spread of coronavirus infectious disease.

Based on real data on coronavirus disease in Azerbaijan, the model was numerically implemented in MS Excel.

The article discusses the application of a linear regression model to predict the process of exponential growth of infected. However, the following precautions must be taken into account:
1. The linear model is only an approximate function of exponential growth and has a certain margin of error.
2. We have studied the exponential growth of the epidemic; the exponential growth corresponds to the epidemic only in its initial stages. At some point, recovered people will no longer spread the infection. When (almost) everyone is infected or cured, growth will stop.

References

1. Shi, P., Cao, S., Feng, P. (2020). SEIR Transmission dynamics model of 2019 nCoV coronavirus with considering the weak infectious ability and changes in latency duration. *Medrxiv*.
2. Eisenhammer, T., Hübler, A., Packard, N. (1991). Modelling experimental time series with ordinary differential equations. *Biological Cybernetics*, 65, 107-112.
3. Worldometer (2020). Azerbaijan. Retrieved from : https://www.worldometers.info/coronavirus/country/azerbaijan/
4. Allen, C., Tsou, M.H., Aslam, A., Nagel, A., Gawron, J.M. (2016). Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza. *PLOS ONE*, 11(7), 157734.
5. Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature,* 457, 1012-1014.
6. Gu, Y., Chen, F., Liu, T., Lv, X., Shao, Z., Lin, H., Liang, C., Zeng, W., Xiao, J., Zhang, Y. (2015). Early detection of an epidemic erythromelalgia outbreak using Baidu search data. *Scientific Reports*, 5(1), 12649.
7. Lampos, V., Miller, A. C., Crossan, S., Stefansen, C. (2015). Advances in nowcasting influenza-like illness rates using search query logs. *Scientific Reports*, 5(1), 12760.
8. Akpa, O. M., Oyejola, B. A. (2010). Modeling the transmission dynamics of HIV/AIDS epidemics: an introduction and a review. *The Journal of Infection in Developing*, *Countries*, 4, 597-608.
9. Alaniz, A. J., Bacigalupo, A., Cattan, P. E. (2017). Spatial quantification of the world population potentially exposed to Zika virus. *International Journal of Epidemiology*, 46, 966-975.
10. Bisanzio, D., Bertolotti, L., Tomassone, L., Amore, G., Ragagli, C., Mannelli, A., Giacobini M., Provero P. (2010). Modeling the spread of vector-borne diseases on bipartite networks. *PLOS ONE*, 5(11), 13796.
11. Alonso, W. J., Viboud, C., Simonsen, L., Hirano, E. W., Daufenbach, L. Z., Miller, M. A. (2007). Seasonality of influenza in Brazil: a traveling wave from the Amazon to the subtropics. *American Journal of Epidemiology*, 165(12), 1434-1442.
12. Chowell, G., Hengartner, N. W., Castillo-Chavez, C., Fenimore, P. W., Hyman, J. M. (2004). The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *Journal of Theoretical Biology*, 229(1), 119-126.
13. Gog, J. R., Ballesteros, S., Viboud, C., Simonsen, L., Bjornstad, O. N., Shaman, J., Chao, D. L., Khan, F., Grenfell, B. T. (2014). Spatial transmission of 2009 pandemic influenza in the US. *PLOS Computational Biology*, 10(6), 1003635.

14. Lofgren, E., Fefferman, N. H., Naumov, Y. N., Gorski, J., Naumova E. N. (2007. Influenza seasonality: underlying causes and modeling theories. *Journal of Virology*, 81(11), 5429-5436.

15. Simonsen, L. (1999). The global impact of influenza on morbidity and mortality. *Vaccine*, 17(1), 3-10.