MODELLING AND PREDICTION OF COVID-19 EPIDEMIC IN TURKEY AND COMPARING WITH USA AND CHINA

Engin Ergul[a] , **Halil Ibrahim Kurt[b*]** , **Murat Oduncuoglu[c]** , **Necip Fazil Yilmaz[d]**

[a]*Izmir Vocational School, Dokuz Eylul University, Izmir, 35380, Turkey*
engin.ergul@deu.edu.tr

[b]*Department of Metallurgical and Materials Engineering, Engineering Faculty, Gaziantep University, 27310, Gaziantep, Turkey*
*hiakurt@gmail.com (*corresponding author)*

[c]*Technical Sciences, Gaziantep University, 27310, Turkey*
oduncuoglu@gmail.com

[d]*Department of Mechanical Engineering, Engineering Faculty, Gaziantep University, 27310, Gaziantep, Turkey*
nfyilmaz@gantep.edu.tr

Abstract

The aim of the study is to research and compare the influences of the confirmed cases, test number and time range on the death and recovery rates in the United State of America, China, and Turkey, and to find out the effect of the epidemic in the near future of Turkey. The modelling and prediction of effects of the day, case and test numbers of COVID-19 infection in the USA, China and Turkey are carried out using the artificial neural network approach (ANN). The system are trained and tested with the different numbers of neurons, hidden layers and activation functions to increase the reliability and accuracy of model. The proposed models have a high R2 value for China and Turkey. We can say according to the results that the measures taken by the USA are inadequate. The formulation is applied to predict the effect of Covid-19 infection in Turkey. The test number that is an important factor in detecting the cases should be increased. The results show a good fit between the observed data and those obtained by the ANN model. If the precautions are strictly followed, the case number will be decreased significantly after 160 days for Turkey according to result of the proposed model but due to the uncontrolled variables, this time may result in between 200 and 250 days.

Keywords: Covid-19, Modelling, Prediction, USA, China, Turkey

1. Introduction

The new Coronavirus (COVID-19) is a virus identified on January 13, 2020. The virus was detected in patients with fever, cough, and shortness of breath in Wuhan Province, China. The epidemic first spread from person to person, then to other cities, and eventually from china to other countries of the world [1-3]. It is well known that Coronaviruses can lead to respiratory infections, Severe Acute Respiratory Syndrome (SARS) and East Respiratory Syndrome (MERS). In addition to sneezing, touching the mouth, nose and eyes with infected hands increases the risk of coronavirus being transmitted. For this reason, cleaning is one of the most important parameters in prevention of coronavirus. There is no effective treatment for the disease. Depending on the general condition of the patient, the necessary supportive treatment is applied. The effectiveness of some drugs on the virus is being investigated. However, there is currently no virus effective drug.

A COVID-19 study displayed that patients with hypertension, diabetes and coronary heart disease are the most important underlying diseases [4, 5]. It is noted that the patients (>65 years) with comorbidities and ARDS have a higher risk of death. Thus, It can be said that the comorbidities had correlated with the intensity of COVID-19, and could cause poor or death [6].

The first case of Covid-19 epidemic which located across the world by 2019-20 announced on March 10, 2020 by the Republic of Turkey Ministry of Health in Turkey. The first virus-related death in the country occurred on March 15, 2020. The epidemic has caused many to take radical decisions that cause significant impacts and results in social, economic, political, economic, administrative, judicial, military and religious fields. In Turkey, Coronavirus Scientific Board is created, passengers from abroad are subjected to thermal scans, all flights are stopped, border gates are closed, education is interrupted, Entries and exits to 30 metropolitan cities are prohibited and economic measure packages are announced. Environmental, cultural and regional factors can be affected by the epidemiological dynamics of infectious diseases.

Mathematical modeling and prediction approaches have been widely used in a lot of disciplines such as engineering, biology, and medicine [7-15]. In this work, we used the ANN approach to reveal the influences of the confirmed cases, test number, time range, death and recovery rates of COVID-19 in the United State of America, China, and Turkey and compared the results. In addition, we aim to find out the effect of the epidemic in the near future of Turkey.

2. Methods

In this study, we used data given in Table 1. The input variables are the numbers of the days, confirmed cases and tests. The output variables are the numbers of death and recovery.

Table 1. The used variables

Country	Date	Days	Case	Test	Death	Recovery
The USA	14.03.2020-15.06.2020	94	2182833	25259077	118280	881102
China	21.01.2020-15.06.2020	147	83181	830413	4634	78370
Turkey	17.03.2020-15.06.2020	91	179831	2674203	4828	152468

The inputs and outputs are divided into two groups; training and testing sets and they are selected in random systematic way where 70% of data is the training data, and 30% of data is the testing data. All parameters are normalized in the range of 0.0 to 1.0 by using the following equations:

$$X_n = (V_a - V_{min}) / (V_{max} - V_{min}) \tag{1}$$

Where X_n normalized value, V_a is the variables used in the model, V_{min} and V_{max} are the minimum and maximum values of the variables used in the model, respectively. The different neuron numbers (1-5) in one and two hidden layers are used to state the optimal model architecture. Also, Sigmoid and Tangent functions are used as activation functions. In order to increase the reliability and accuracy of the model, different system variables are used. The correlation coefficient (R) is chosen to estimate the performance of the proposed model and, mean square error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are utilized as error-evaluation criteria. The explanations and detailed formulations about statistical values are given in our works [16, 17]. All analyses of the present work are performed using the Matlab program.

3. Results

Tables 2-4 show the statistical results for the USA, China and Turkey. In these tables, samples are coded according to the used system variables. For example, Sig, H2, N1 and N1 represent activation function, the numbers of hidden layers and neurons in hidden layers, respectively. Activation functions is Sigmoid function, the number of hidden layers is two, the number of neuron in the first hidden layer is three and the number of neuron in the second hidden layer is three for "SigH2N3N3" code. The optimum results are obtained with "SigH2N3N3" code sample for USA, "SigH2N4N4" code sample for China and "TanH2N4N4" code sample for Turkey. The correlation coefficient (R) displays the consistency of the results. It is clear that Turkey has the maximum R-value but the USA has the minimum R-value.

Table 2. MSE, R and MAE results for the USA

Code	Training			Testing		
	MSE	R	MAE	MSE	R	MAE
SigH2N1N1	0.01832	0.45204	0.09110	0.01031	0.69120	0.08593
SigH2N2N2	0.00937	0.76959	0.05603	0.00306	0.91491	0.04495
SigH2N3N3	0.00992	0.75610	0.05438	0.00306	0.91157	0.04182
SigH2N4N4	0.01095	0.72422	0.06587	0.00333	0.89832	0.04688
SigH2N5N5	0.00781	0.81850	0.06557	0.00733	0.87076	0.05979
TanH2N1N1	0.02643	0.26922	0.11239	0.01754	0.05532	0.11466
TanH2N2N2	0.01161	0.71643	0.07137	0.00566	0.86354	0.05355
TanH2N3N3	0.00977	0.75837	0.05815	0.00391	0.88395	0.04808
TanH2N4N4	0.00923	0.77526	0.05845	0.00390	0.89287	0.05052
TanH2N5N5	0.00207	0.95390	0.02971	0.00755	0.78053	0.05501
SigH1N1	0.01930	0.41295	0.09621	0.01180	0.66326	0.09560
SigH1N2	0.01183	0.75826	0.05940	0.00398	0.91295	0.04725
SigH1N3	0.01229	0.68501	0.07218	0.00405	0.89054	0.04869
SigH1N4	0.01039	0.74167	0.06021	0.00308	0.90787	0.04632
SigH1N5	0.00856	0.79410	0.05857	0.00346	0.90744	0.04679

Code	MSE	R	MAE	MSE	R	MAE
TanH1N1	0.02013	0.35315	0.09834	0.01878	0.17126	0.11435
TanH1N2	0.01194	0.70286	0.06293	0.00452	0.87991	0.05429
TanH1N3	0.01436	0.71329	0.06961	0.00521	0.87806	0.05362
TanH1N4	0.01098	0.72878	0.06934	0.00448	0.86836	0.06083
TanH1N5	0.00773	0.82135	0.05461	0.00386	0.90354	0.04833

Table 3. MSE, R and MAE results for China

Code	Training			Testing		
	MSE	R	MAE	MSE	R	MAE
SigH2N1N1	0.00989	0.18745	0.02821	0.00041	0.73567	0.01778
SigH2N2N2	0.00966	0.23950	0.02213	0.00017	0.97315	0.01267
SigH2N3N3	0.00961	0.25130	0.02092	0.00013	0.93222	0.00972
SigH2N4N4	0.00005	0.99744	0.00403	0.00004	0.97193	0.00383
SigH2N5N5	0.00006	0.99743	0.00436	0.00004	0.97578	0.00408
TanH2N1N1	0.01001	0.15354	0.03102	0.00049	0.71249	0.02019
TanH2N2N2	0.01001	0.15200	0.03117	0.00047	0.72021	0.01939
TanH2N3N3	0.00976	0.21817	0.03020	0.00060	0.54724	0.01926
TanH2N4N4	0.00008	0.99629	0.00514	0.00008	0.96861	0.00539
TanH2N5N5	0.00005	0.99760	0.00400	0.00006	0.95375	0.00440
SigH1N1	0.01012	0.11350	0.03375	0.00064	0.50231	0.02273
SigH1N2	0.01008	0.12660	0.03331	0.00060	0.49510	0.02196
SigH1N3	0.00956	0.25826	0.02239	0.00016	0.95347	0.01153
SigH1N4	0.00393	0.92077	0.01636	0.00026	0.82164	0.01108
SigH1N5	0.00089	0.95966	0.02149	0.00064	0.78063	0.01793
TanH1N1	0.00993	0.17977	0.03169	0.00052	0.70444	0.02098
TanH1N2	0.00981	0.20624	0.02564	0.00027	0.90333	0.01521
TanH1N3	0.00987	0.19284	0.02678	0.00029	0.89508	0.01593
TanH1N4	0.00926	0.31406	0.02492	0.00031	0.87189	0.01536
TanH1N5	0.00013	0.99451	0.00649	0.00021	0.89766	0.00860

Table 4. MSE, R and MAE results for Turkey

Code	Training			Testing		
	MSE	R	MAE	MSE	R	MAE
SigH2N1N1	0.07497	0.38792	0.22838	0.06415	0.45663	0.19812
SigH2N2N2	0.00421	0.97627	0.05397	0.00347	0.98190	0.04593
SigH2N3N3	0.00244	0.98828	0.03873	0.00257	0.98583	0.04302
SigH2N4N4	0.00120	0.99318	0.02828	0.00241	0.98429	0.04037
SigH2N5N5	0.00207	0.98913	0.03575	0.00147	0.99110	0.02959
TanH2N1N1	0.06881	0.46978	0.21559	0.06879	0.39209	0.23117
TanH2N2N2	0.00378	0.97837	0.04838	0.00292	0.98522	0.04290
TanH2N3N3	0.00207	0.98839	0.03900	0.00280	0.98606	0.04218
TanH2N4N4	0.00101	0.99425	0.02509	0.00153	0.98993	0.03292

TanH2N5N5	0.00091	0.99483	0.02388	0.00205	0.98911	0.03490
SigH1N1	0.02834	0.88642	0.13175	0.02445	0.87966	0.11771
SigH1N2	0.00660	0.96194	0.06348	0.00720	0.96361	0.06128
SigH1N3	0.00199	0.98865	0.03724	0.00190	0.99202	0.03512
SigH1N4	0.00467	0.97745	0.05496	0.00278	0.98372	0.04373
SigH1N5	0.00182	0.98968	0.03526	0.00190	0.99108	0.03490
TanH1N1	0.01283	0.92794	0.08771	0.01253	0.91285	0.08670
TanH1N2	0.00710	0.95909	0.06650	0.00841	0.95541	0.06500
TanH1N3	0.00573	0.96708	0.06036	0.00570	0.97035	0.05891
TanH1N4	0.00321	0.98175	0.04347	0.00369	0.98373	0.04721
TanH1N5	0.00310	0.98520	0.04438	0.00240	0.98552	0.03905

Figure 1 and 2 illustrate the death number in the training test and the testing set for the countries. The R^2 values in the training test for the USA, China and Turkey are 0.57, 0.99 and 0.98, respectively. USA showed the lowest consistency, while the highest consistency for China in training set. The results in the training test for Turkey are quite high. In testing set, the R^2-values of the USA, China and Turkey are 0.82, 0.94 and 0.98, respectively. Turkey has the maximum consistency in testing set compared to the USA and China.
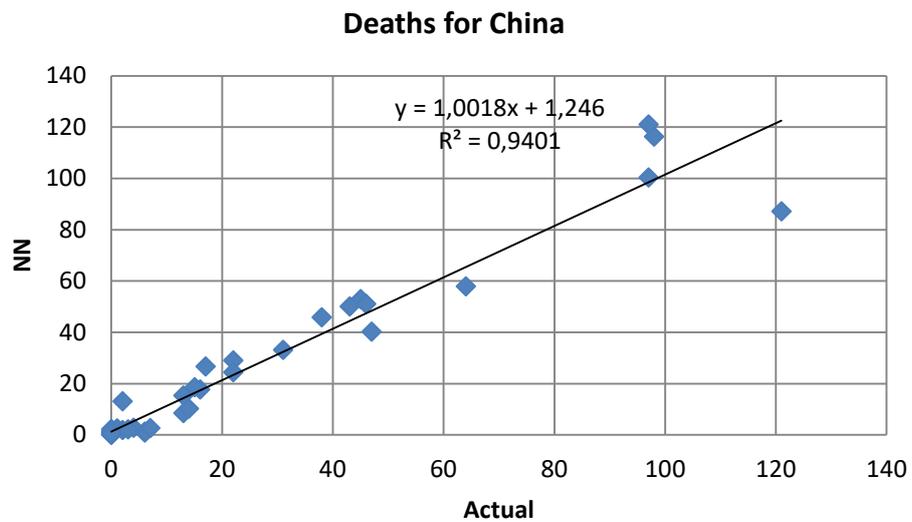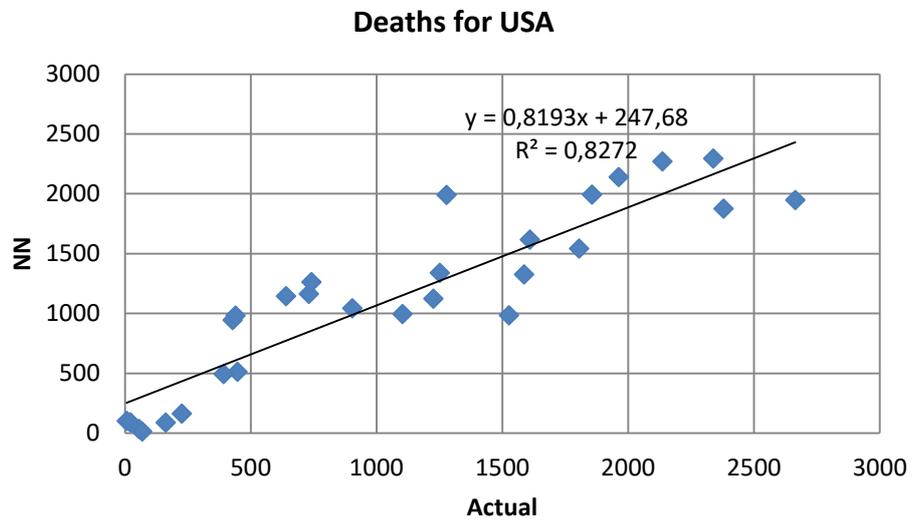

Deaths for USA

$y = 0,5279x + 656,82$
$R^2 = 0,5724$

Deaths for China

$y = 0,9911x + 1,1988$
$R^2 = 0,9951$

Deaths for Turkey

y = 0,9856x + 0,8384
R² = 0,9885

(NN vs Actual scatter plot)
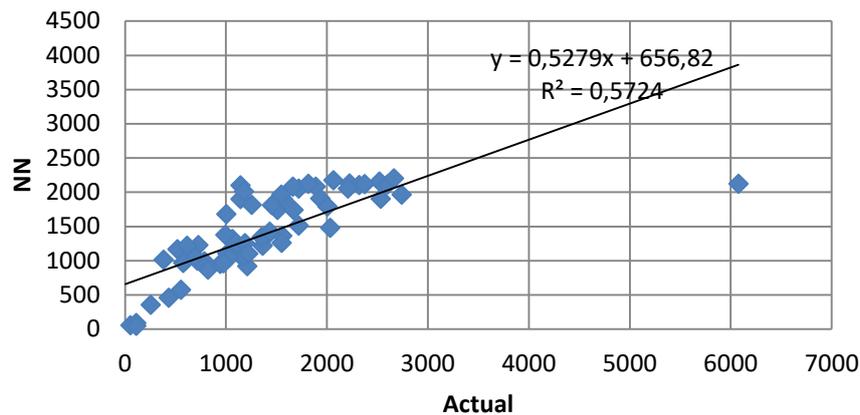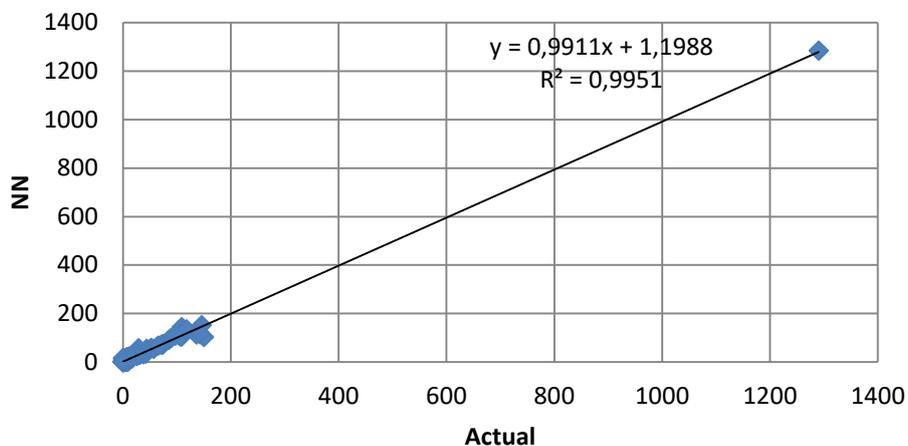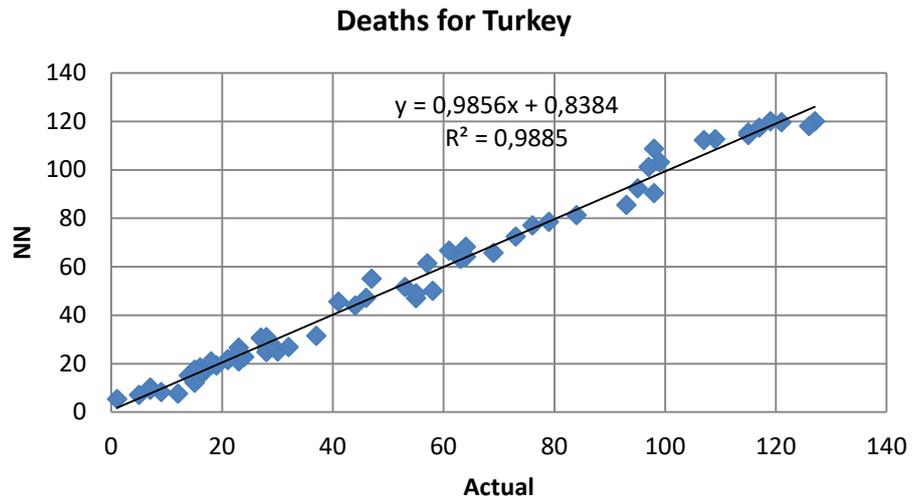
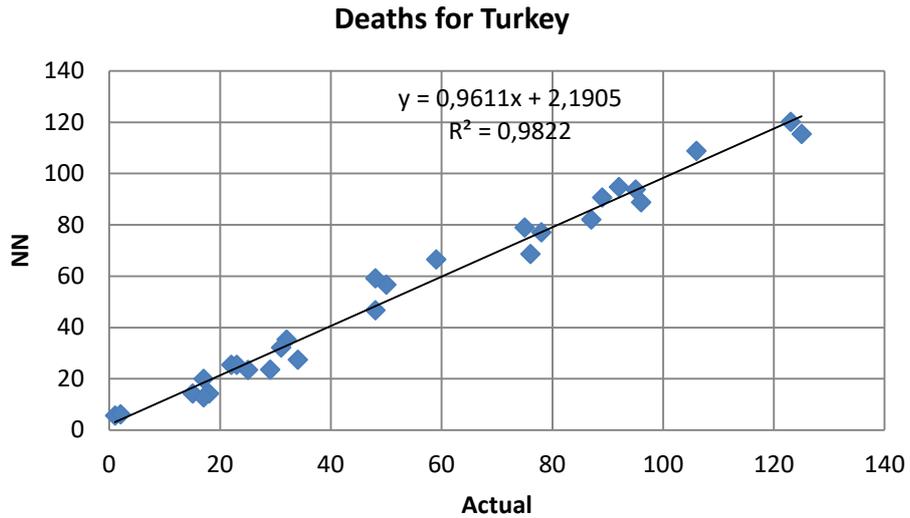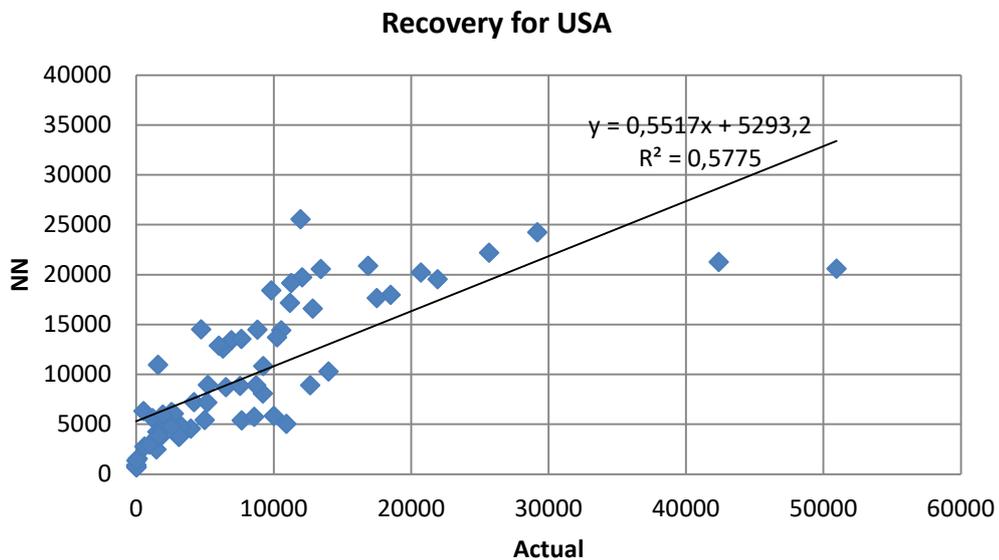**Figure 1.** Death rates in training test for the USA, China and Turkey

Deaths for USA

y = 0,8193x + 247,68
R² = 0,8272

(NN vs Actual scatter plot)

Deaths for China

y = 1,0018x + 1,246
R² = 0,9401

(NN vs Actual scatter plot)

Deaths for Turkey

$$y = 0{,}9611x + 2{,}1905$$
$$R^2 = 0{,}9822$$

Figure 2. Death rates in testing set for USA, China and Turkey

Figure 3 and 4 explain the recovery number in the training test and testing set for the countries. In the training test, the R^2-values of the USA, China and Turkey are 0.57, 0.99 and 0.97, respectively. The USA demonstrated the lowest consistency, while the highest consistency for China in the training set. The results in the training test for Turkey are quite high. The R^2-values of the USA, China and Turkey in the testing set are 0.22, 0.97 and 0.93, respectively. China has the maximum consistency in the testing set compared to the USA and Turkey.

Recovery for USA

$$y = 0{,}5517x + 5293{,}2$$
$$R^2 = 0{,}5775$$

Recovery for China

Recovery for Turkey
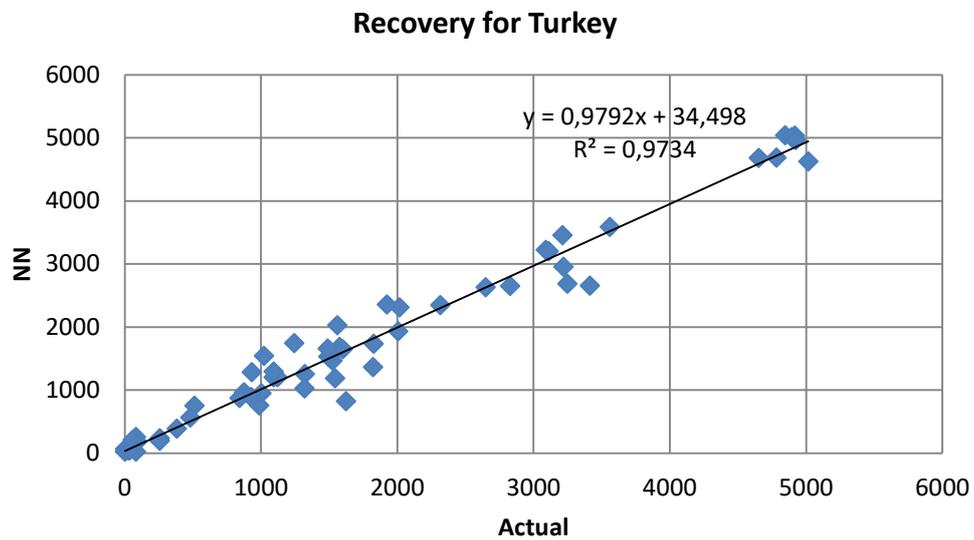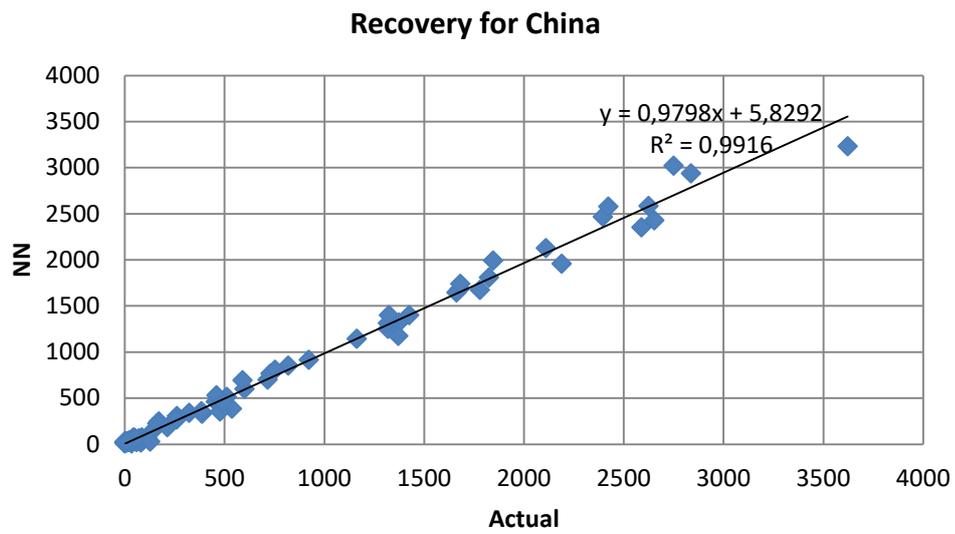
Figure 3. Recovery rates in training test for the USA, China and Turkey

Recovery for USA

Recovery for China

$$y = 1{,}0274x + 30$$
$$R^2 = 0{,}9786$$

Recovery for Turkey
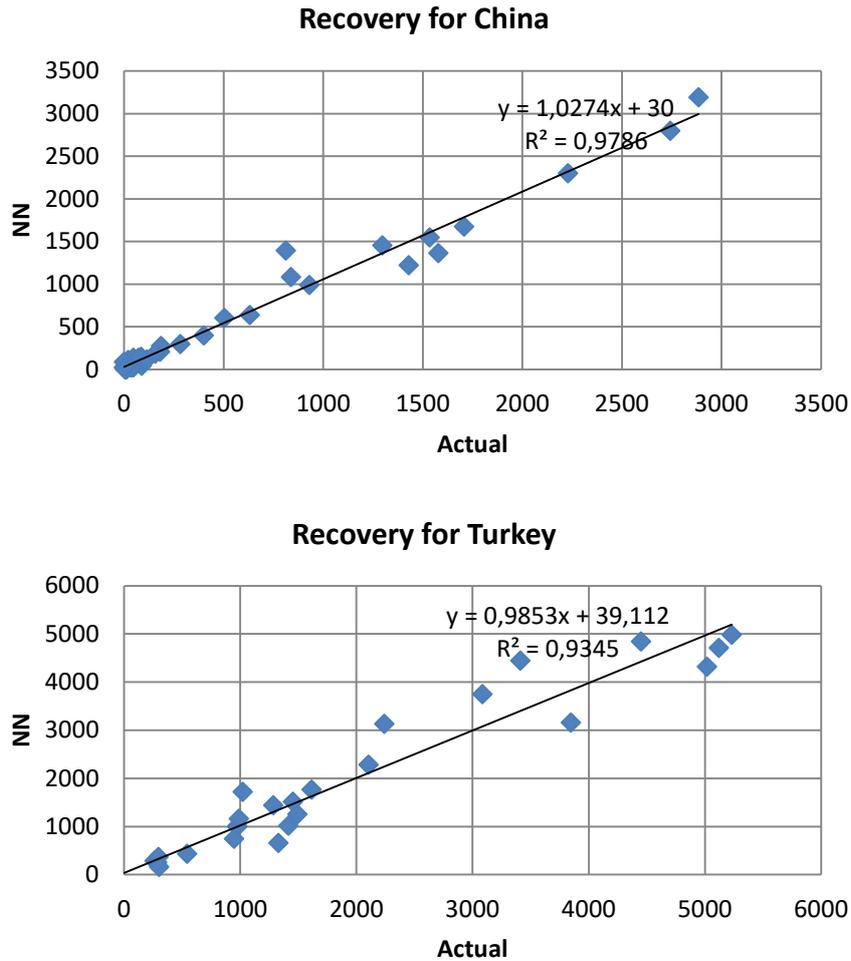

$$y = 0{,}9853x + 39{,}112$$
$$R^2 = 0{,}9345$$
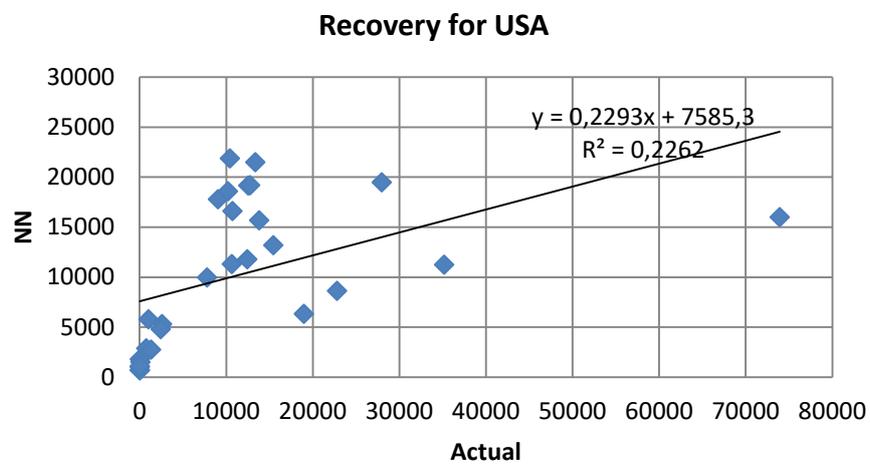
Figure 4. Recovery rates in testing set for the USA, China and Turkey

The rate of case/tested is charted in Figure 5. The number of tests and confirmed cases does not refer to the same in the three countries. The average case number of the USA, China and Turkey is 23169, 567 and 1979, and the average test number for these countries is 253700, 5634 and 29246, respectively. The highest case/tested rate is observed in the USA while the lowest case/tested rate is in China.
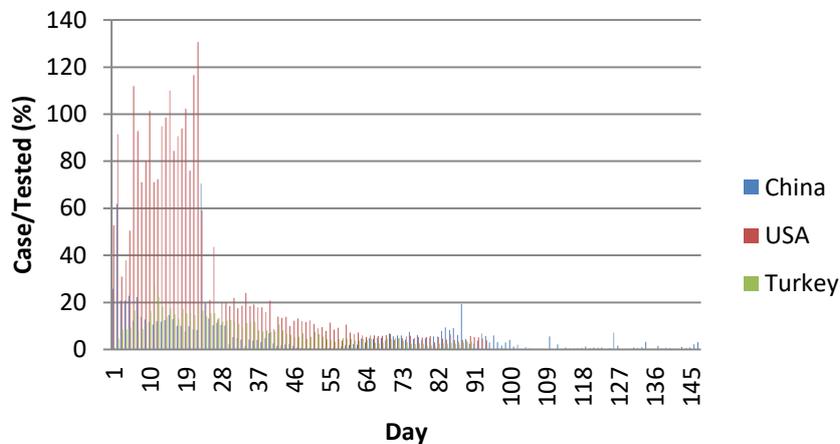
Figure 5. Case/ Tested number of the USA, China and Turkey

Figure 6 shows the number of death in the USA, China and Turkey, and Figure 7 displays the rate of death/population. The number of death and the rate of death/population of the USA is 118160 and 0.0358% which is the highest among the three countries.
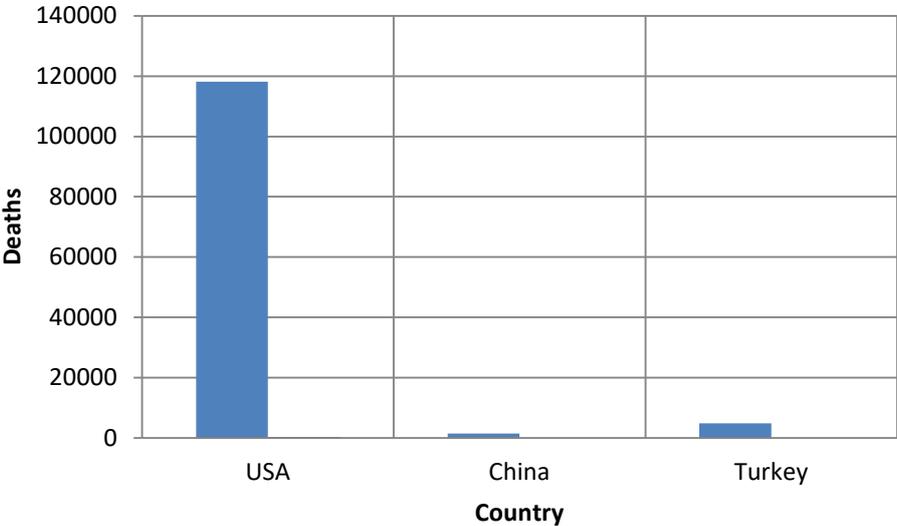

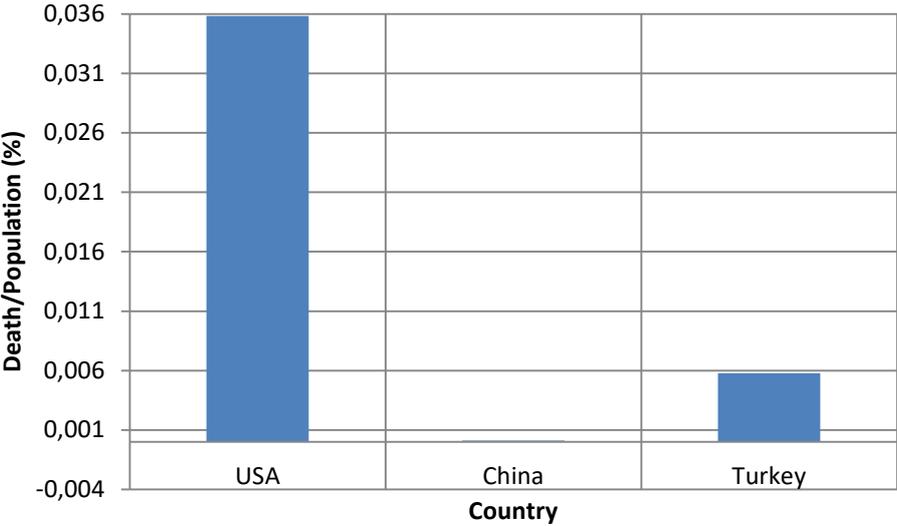**Figure 6.** Death number of the USA, China and Turkey

Figure 7. Death/Population rates of the USA, China and Turkey

Figure 8 and 9 indicate the number of recoveries and the rate of recovery/population in the USA, China and Turkey. The highest recovery number and recovery/population rate are observed in the USA as 881818 and 0.0267%, respectively.
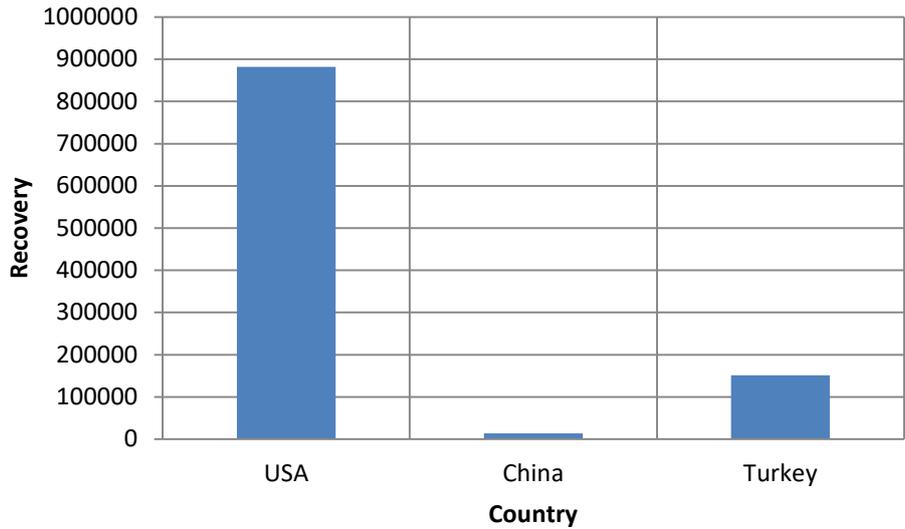
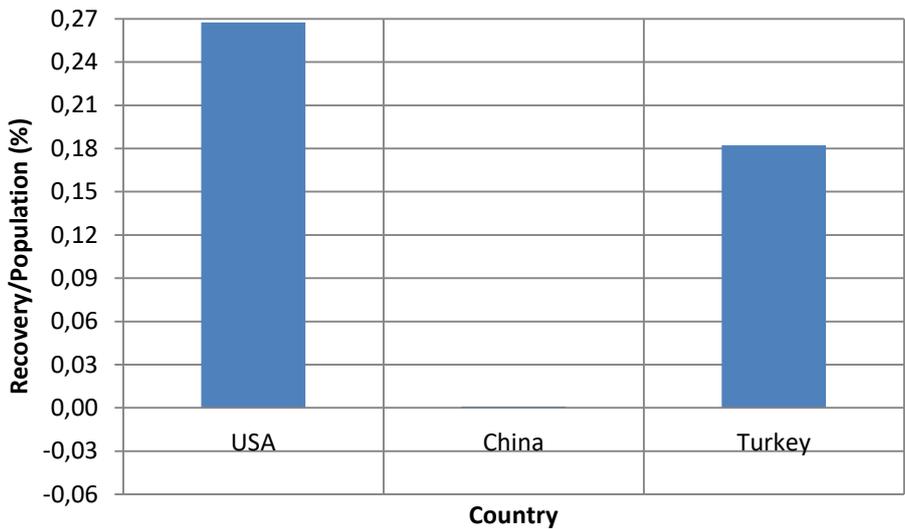**Figure 8.** Recovery number of the USA, China and Turkey

Figure 9. Recovery/Population rate of the USA, China and Turkey

Figure 10 and 11 demonstrate the effect of the number of days on the number of death and recovery in Turkey. The number of cases is determined as 1000 and the number of tests is as 29386 that is the average test number for Turkey in Figure10. In Figure 11, the case is fixed as 1000 and the test is 15000 for Turkey. It is clear from these Figures that as the number of day increases, the number of deaths and the recovery decreases in Turkey. It is well known that the test number is one of the most important factors in detecting Covid-19 pandemic. So, the influence of the test number on the number of death is researched in Figure 12 for Turkey. The number of days and cases is designated as 90 and 100, respectively. It is observed that the number of death reduces with increasing in the average test number.
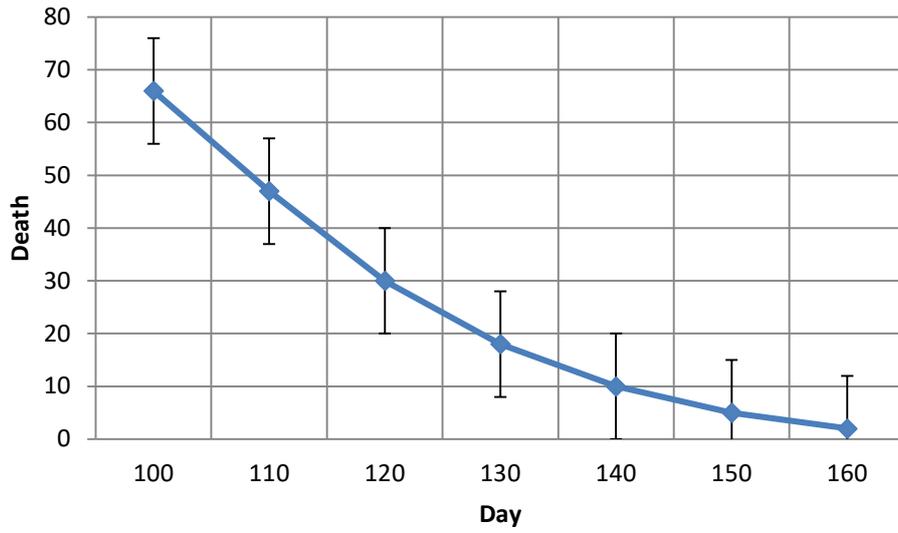
Figure 10. The relationship between the number of days and the number of death in Turkey
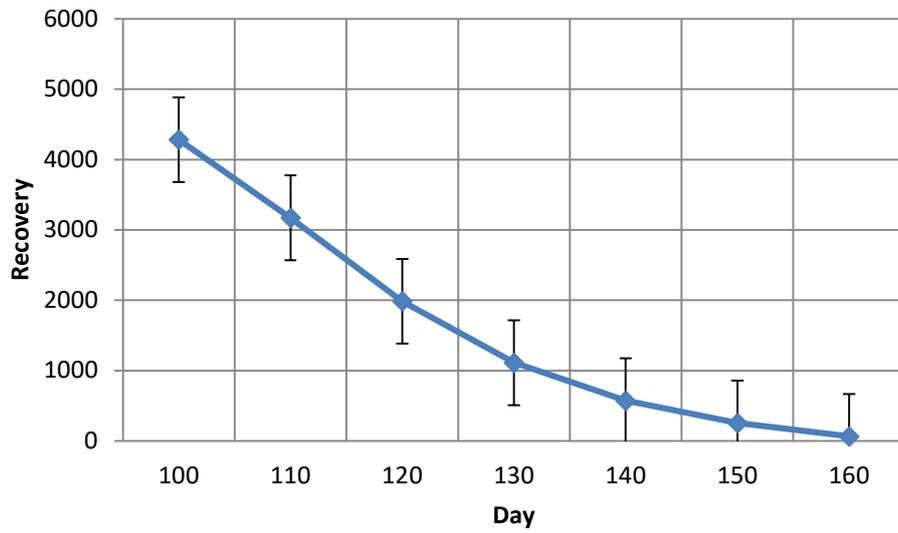

Figure 11. The relationship between the number of days and the number of recovery in Turkey
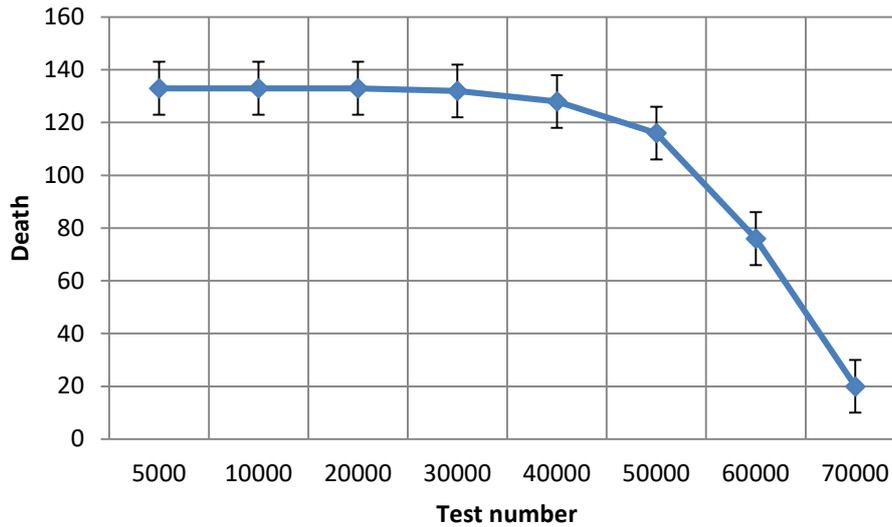
Figure 12. The relationship between the number of tests and the number of death in Turkey

4. Discussion

In the artificial neural network model, MSE, MAE, R and R^2 parameters are used to find out the performance of the studied model. The effect of input variables on the death and recovery numbers is revealed. In order to get the best performance, the various system parameters are used and the different NN architectures are obtained according to the results of the error criteria and correlation coefficient for each country. This shows the effect of the system variables on the results. The system is trained with the used data in the training test and the learning ability is studied with the data in the test set. The correlation coefficients of the USA, China and Turkey in test set are bigger than 0.90. This explains that the proposed model has the confidence. It is concluded that the obtained optimal architecture for each country exhibits the minimum error values and the maximum correlation. The R^2 value is a statistical indicator usually applied to multiple regression analysis and it checks the accuracy of the model in which the prediction is just the mean of all of the results. If the R^2 value is 1, there is a perfect fit between actual and NN results. If this value approaches zero, this means that the fit is bad between the results. It is concluded that the proposed model could be predicted death rates with high accuracy for Turkey.

The test number is an important factor in identifying people with Covid-19 disease. It is possible to say that the death rate will decrease significantly with the increase in the number of average tests within 90 days. It is well known that Covid-19 disease is conveyed through coughing or sneezing. After touching the contaminated surfaces, touching the eyes, nose or mouth is risky and extreme care should be taken to clean it. So, people have to use wear masks, pay attention to cleanliness and maintain social distance. In Turkey, we have to be very careful when complying with customs and traditions. People with diseases such as heart disease, hypertension, diabetes, chronic respiratory disease and cancer are at great risk. Strict precautions to these patients will reduce the death rate. China carried out strict precautions and so the death rate for this country is quite low. The case numbers may decrease in the summer months since the people going out to open areas like picnic, park and garden will increase the social distance. It is reported that temperatures above 8 to 10 °C are associated with decreased COVID-19 daily confirmed cases rate [18]. The vaccine or specific drug for

the therapy of Covid-19 pandemic still has not been developed. A lot of factors affect the therapy process. It can be said that the environmental and nutritional interventions are a key technique to reach the best clinical outcome [19]. The Vitamin D deficiency, environmental variables like air pollution and travel frequency affect the death and recovery rates [20-23]. As a result of detailed analysis, we have to learn to live Covid-19 until a vaccine is found. To reduce the death rate and to increase the recovery rate, we must strengthen the immune system because there is no proof a specific method to help stop the virus.

5. Conclusion

In order to estimate the effect of days, COVID-19 cases and tested numbers on the death and recovery rates in the USA, China and Turkey, the ANN approach is performed. The results are compared for the three countries. The detailed analysis is carried out for Turkey. The results of the modeling show a good fit between the estimated and the observed data given as a result of an R^2 of 0.82, 0.94 and 0.98 for death rates and 0.22, 097 and 0.93 for recovery rates in USA, China and Turkey in the testing set. The death/population rates of the USA, China and Turkey are 0.0358%, 0.0001% and 0.0057%. China is the most successful country compared to the USA and Turkey, and the USA is the most unsuccessful country according to these results. The predicted results from the models could differ in a significant way from the observed ones because the countries apply different strategies. The re-opening of the community, stay-home and social distancing policies will affect the death rates. The test number is an important factor to detect the pandemic, and countries should increase the test number and perform strict measures to control the spreading of Covid-19 pandemic. If the precautions are strictly implemented, we can say that the two-digit case numbers can be reached after 160 days for Turkey. But, we can say that this may be about within 200 – 250 days. Because it is important to whom the tests are applied as much as the number of tests.

References

[1]   Cooper, I., Mondal, A., Antonopoulos, C. G., "A SIR model assumption for the spread of COVID-19 in different communities", Chaos, Solitons & Fractals, (2020) : 110057.

[2]   Tezer, H., Demirdağ, T. B., "Novel coronavirus disease (COVID-19) in children", Turkish Journal of Medical Sciences 50(SI-1) (2020) : 592-603.

[3]   Leal-Neto, O. B., et al., "Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning", International Journal of Medical Informatics 143 (2020) : 104263.

[4]   Guan, W.-j., et al., "Clinical characteristics of coronavirus disease 2019 in China", New England Journal of Medicine 382 (2020) : 1708-1720.

[5]   Celik, I., Saatci, E., Eyüboğlu, A. F., "Emerging and reemerging respiratory viral infections up to Covid-19", Turkish Journal of Medical Sciences 50(SI-1) (2020) : 557-562.

[6]   Yang, X., et al., "Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: A single-centered, retrospective, observational study", The Lancet Respiratory Medicine, 8(5) (2020) : 475-481.

[7]   Kurt, H., Oduncuoglu, M., "Formulation of the effect of different alloying elements on the tensile strength of the in situ Al-Mg2Si composites", Metals 5(1) (2015) : 371.

[8]   Kurt, H., Oduncuoglu, M., Kurt, M., "A mathematical formulation to estimate the effect of grain refiners on the ultimate tensile strength of Al-Zn-Mg-Cu alloys", Metals 5(2) (2015) : 836.

[9]   Togun, N., et al., "Formulation of effects of atropine, pralidoxime and magnesium sulfate on cardiac tissue levels of nitric oxide, malondialdehyde and glutathione in organophosphate poisoning using artificial neural network", Computers in Biology and Medicine 40(1) (2010) : 29-36.

[10] Ardakani, A. A., et al., "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks", Computers in Biology and Medicine 121 (2020) : 103795.

[11] Ankarali, H., et al., "A statistical modeling of the course of COVID-19 (SARS-CoV-2) outbreak: A comparative analysis", Asia Pacific Journal of Public Health 32(4) (2020) : 155-160.

[12] Ghosh, S., "Predictive model with analysis of the initial spread of COVID-19 in India", International Journal of Medical Informatics 143 (2020) : 104262.

[13] Mollalo, A., et al., "Predicting the hotspots of age-adjusted mortality rates of lower respiratory infection across the continental United States: Integration of GIS, spatial statistics and machine learning algorithms", International Journal of Medical Informatics 142 (2020) : 104248.

[14] Wollenstein-Betech, S., Cassandras, C. G., Paschalidis, I. C., "Personalized Predictive Models for Symptomatic COVID-19 Patients Using Basic Preconditions: Hospitalizations, Mortality, and the Need for an ICU or Ventilator", International Journal of Medical Informatics 142 (2020) : 104258.

[15] Eroglu, E., Bozkurt, E., Esenpinar, A. A., Tek, S., "Mathematical analysis of Covid-19 Phenomenon", Journal of Engineering Technology and Applied Sciences 5(2) (2020) : 59-64.

[16] Kurt, H. I., Oduncuoglu, M., "Effects of temperature, time, magnesium, and copper on the wettability of Al/TiC system", Mathematical Problems in Engineering 710526 (2015) : 6.

[17] Kurt, H. I., Oduncuoglu, M., "Application of a neural network model for prediction of wear properties of ultrahigh molecular weight polyethylene composites", International Journal of Polymer Science 315710 (2015).

[18] Shi, P., et al., "Impact of temperature on the dynamics of the COVID-19 outbreak in China" Science of The Total Environment 728 (2020) : 138890.

[19] Aldaco, R., et al., "Environmental and nutritional impacts of dietary changes in Spain during the COVID-19 lockdown", Science of The Total Environment 748 (2020) : 140524.

[20] Whittemore, P. B., "COVID-19 fatalities, latitude, sunlight, and vitamin D" American Journal of Infection Control 48(9) (2020) : 1042-1044.

[21] Berman, J. D., Ebisu, K., "Changes in U.S. air pollution during the COVID-19 pandemic", Science of The Total Environment 739 (2020) : 139864.

[22] Abid, K., et al., "Progress of COVID-19 epidemic in Pakistan", Asia Pacific Journal of Public Health 32(4) (2020) : 154-156.

[23] Tufan, A., Güler, A. A., Matucci-Cerinic, M., "COVID-19, immune system response, hyperinflammation and repurposing antirheumatic drugs", Turkish Journal of Medical Sciences 50(SI-1) (2020) : 620-632.