Celal Bayar University Journal of Science

Covid-19 Death and Case Numbers Forecasting with ARIMA and LSTM Models

Büşra Çetin[1]*, Nida Gökçe Narin[2]

[1]Graduate School of Natural and Applied Sciences, Muğla Sıtkı Koçman University, Muğla, Türkiye
[2]Department of Statistics, Muğla Sıtkı Koçman University, Muğla, Türkiye
*ccetinbusraa@gmail.com
* Orcid No:0000-0002-3628-205X

Received: 10 February 2022
Accepted: 15 February 2023
DOI: 10.18466/cbayarfbe.1070691

Abstract

Covid-19, which quickly became a pandemic, has not yet been fully controlled despite the developed vaccines. The pandemic has caused a global economic crisis over the two-year struggle period. Many countries have lifted the restrictions they implemented in the fight against the pandemic to get out of this crisis. Despite the vaccines, the pandemic still poses a significant threat and maintains uncertainty about when life will return to pre-pandemic and when the economic crisis will be under control. Therefore, analyzing the current situation based on the policies followed so far remains a key issue in accurately predicting the future course of the pandemic. In this study, Covid-19 estimation is made with Auto Regressive Integrated Moving Average (ARIMA) and Long-Short-Term Memory (LSTM) models using daily case and death numbers for Germany, France, Italy, Ireland, Poland, Russia, and Turkey. The root mean squared error, mean absolute percentage error, mean absolute error, adjusted R2, Akaike Information Criterion, and Schwarz Information Criterion metrics were used in model selection. The results showed that the ARIMA and the LSTM models could be used to estimate the number of Covid-19 cases and deaths. It has also been observed that the prediction accuracy of the LSTM models for the countries studied is higher than that of the ARIMA models.

Keywords: ARIMA, Covid-19, Forecasting, LSTM

1. Introduction

The sars-Cov-2 virus, which first appeared in Wuhan, China, in December 2019, quickly spread and caused a pandemic by affecting the whole world in a short time. Although countries apply different policies regarding the fight against the virus, almost all countries have switched to online education, implemented curfews, and suspended social activities for a long time. With these measures taken, the virus spread from time to time decreased. Still, with the loosening of the restrictions for economic reasons, the epidemic intensified again, and this process continued as a cycle.

Studies have been started in many countries for the vaccine, which is seen as an effective tool in the fight against the pandemic. In April 2021, there were 14 vaccines on the market that were approved and put into use by at least one country. Sinovac, Johnson&Johnson, Moderna, Oxford/AstraZeneca, and Pfizer/BioNTech et al. research has proven that the vaccines developed are effective on the virus, but the epidemic remains a threat as the virus mutates.

There are many studies in the literature to predict the progression of the epidemic. Khan F. and Gupta R. estimated the number of cases in India over 50 days using time series and daily case numbers using ARIMA and Nonlinear Auto Regressive (NAR) model-based forecasting [1]. Ozen et al. (2020) used daily validated Covid-19 case data for the United States to predict with machine learning-based Prophet, Polynomial Regression, ARIMA, Linear Regression, and Random Forest models. In the study, mean absolute percent error (MAPE), root means square error (RMSE), and mean absolute error (MAE) metrics were used to compare the performances of the models. The best predictive value according to the MAPE criterion was obtained by

Polynomial Regression [2]. Sevli and Başer (2021) predicted that the prognosis for the future period might be high in the case of estimation with the prophet model [3]. Awan and Aslam (2020) used the automatic ARIMA model in their study and estimated 10-day cases in European countries (Italy, France, Spain, and Germany) [4]. Kırbaş et al. (2020) compared ARIMA, Nonlinear Auto Regressive Neural Networks (NARNN), and LSTM models with estimated confirmed cases of Covid-19 in Denmark, Belgium, Germany, France, the United Kingdom, Finland, Switzerland, and Turkey. They showed that LSTM is the most appropriate approach. They also predicted a decrease in the rate of increase in cases [5]. Hernandez-Matamoros et al. (2020) modeled the spread of Covid-19 to six continents worldwide with ARIMA. They also investigated whether there is a relationship between countries in the same geographical region regarding the virus's behavior [6].
I
Roy et al. (2021) examined the example of India, which has a high number of daily cases, using the ARIMA model based on the data obtained between January 30 and May 11, 2020 [7]. Moftakhar et al. (2020) made new cases estimation with the Artificial Neural Network (ANN) and ARIMA models they created using the Iranian new case numbers observed between February 19 and March 30, 2020. They showed that ARIMA's estimation performance was higher [8]. In their study, Alzahrani et al. (2020) estimated cases with Auto-Regressive (AR), Moving Average (MA), Auto Regressive Moving Average (ARMA), and ARIMA models for the next four weeks using the number of new cases seen in Saudi Arabia between March 2 and April 20, 2020 [9].

Elsheikh et al. (2021) proposed an LSTM model to estimate the total number of confirmed cases, totally recovered cases, and total deaths in Saudi Arabia. The accuracy of the proposed model was compared with ARIMA and nonlinear autoregressive neural network (NARANN) and used to estimate the total number of confirmed cases, total recovered cases, and total deaths for Brazil, India, Saudi Arabia, South Africa, Spain, and the USA [10]. Ala'raj et al. (2021) predicted the development of the Covid-19 outbreak in the United States using SEIRD (Susceptible, Exposed, Infectious, Recovered, and Deceased) and ARIMA models [11]. Eroğlu (2020) used ANN and LSTM models to predict the number of cases 7 days in advance by using the official data of Turkey and concluded that ANN models performed better in predictions [12]. There are many more studies in the literature on the Covid-19 pandemic. However, most of the 2020 studies used relatively fewer data.

Despite the vaccines developed, the fight against the pandemic remains important due to its different variants. As the number of observations increases in data learning systems, the prediction performance of the models to be created also increases. For this reason, forward-looking predictions are made in this study by examining the course of the epidemic exceeding one year. The daily death and case numbers between March 2020 and October 2021 were modeled using ARIMA and LSTM methods for Germany, France, Ireland, Italy, Poland, Russia, and Turkey. In addition, 10-day cases and deaths are estimated using ARIMA models used in short-term time series estimation and LSTM models based on deep neural networks.

2. Material and Methods

Data collected at certain time intervals are called time series (TS) data. Different models have been developed to make predictions for the future with the help of TS data, which contains the observation values of past periods. The most widely used among these models are the ARIMA models, which can model the linear relationship between the data forming the series without needing any prior knowledge about the structure and general trend of the series. ARIMA models can be successfully applied to stationary series or series that can be made stationary by taking the difference. LSTM is a deep learning model used in solving time series problems. It was developed to solve the gradient descent problem that occurs in the backpropagation process in Recurrent Neural Network. LSTM networks consist of input gate, cell state, forget gate, and output gates, which use sigmoid or tahn activation function to solve short-term memory problems. Thanks to these gates that control and regulate the flow of information, it can receive the information from the previous time and transmit it to the next time. With training, the model can decide which information to remember and which to forget. ARIMA and LSTM models were used in this study due to their widespread use in time series modeling and their successful forecasting performance.
The data used in the study were obtained from the Covid-19 dataset published by Our World in Data [13]. For Germany, France, Ireland, Italy, Poland, Russia, and Turkey, the probable number of cases and deaths for the next ten days is predicted with the prediction models created using the daily confirmed deaths and new cases numbers until 30 October 2021.

2.1. Auto Regressive Integrated Moving Average

Auto Regressive Integrated Moving Average (ARIMA) is a univariate time series modeling approach that combines both auto regressive (AR) and moving average (MA) models. The model based on using the autocorrelation structure of the time series was proposed by George Box and Gwilym Jenkins [13]. ARIMA models are defined as ARIMA(p,d,q). AR(p) represents p-order autoregressive process, MA(q) represents a q^{th} degree of moving average process. If the processed time

Celal Bayar University Journal of Science
Volume 19, Issue 1, 2023, p 39-46
Doi:10.18466/cbayarfbe.1070691

series is not stationary, d is the number of differencing required to make the time series stationary.
An AR model is assumed to be generated from a linear function of the past values of the time-series. The AR models are named according to the number of past observations they contain, and the model containing *p* past values is shown as AR(p). In the AR(p) model, the current output x_t is expressed by previous values and parameters, as formulated in Eq. (2.1), where *t* is the time, *p* is the order of the parameters, ϕ_i are the autocorrelation coefficients, *c* is a constant value, and ε_t is the Gaussian white noise series with mean zero and variance σ_ε^2 [14].

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t \quad \textbf{(2.1)}$$

The MA models are models in which the observation value in a certain period of the time series is expressed as a linear combination of the error terms of the same period and the error terms of the q number of past periods. In the MA(q) model, the current output x_t is expressed by previous values and parameters, as formulated in Eq. (2.2), where *t* is the time, *q* is the order of the parameters, θ_i are the partial autocorrelation coefficients, ε_t is the Gaussian white noise series with mean zero and variance σ_ε^2 [15].

$$x_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \quad \textbf{(2.2)}$$

The ARMA models, a combination of AR and MA models, are also frequently used to model stationary time series [14]. The ARMA(p,q) model is defined as in Eq (2.3).

$$x_t = \mathrm{c} + \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \quad \textbf{(2.3)}$$

Non-stationary time series are made stationary by taking the difference. The models applied to the series converted to stationary by the difference process are called integrated models. For example, if the time series has a linear trend, the first difference series is stationary. On the other hand, if the time series has a curvilinear trend, the second difference series is stationary by retaking the differences. The model used in this case is expressed as ARIMA(p,d,q). Here, "d" refers to the stationarization parameter of the series [15,17]. The ARIMA(p,d,q) model is defined as in Eq (2.4).

$$x_t = (1 - x_{t-d}) + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \, \mathbf{(2.4)}$$

Model selection is based on AIC, SIC, and adjusted R^2 values. The model with the lower absolute values of the AIC and SIC information criteria and the higher adjusted R^2 is the most appropriate model.

2.2. Long Short Term Memory

Long-short-term memory (LSTM) is a deep learning approach with recurrent neural networks (RNN) architecture proposed by Hochreiter and Schmidhuber [18]. RNN is also used to handle long-term dependencies, but it can cause losses by taking into account previous periods' learning. This situation is known as the vanishing gradient problem. The vanishing gradient problem is due to backpropagation optimization used in neural network training.

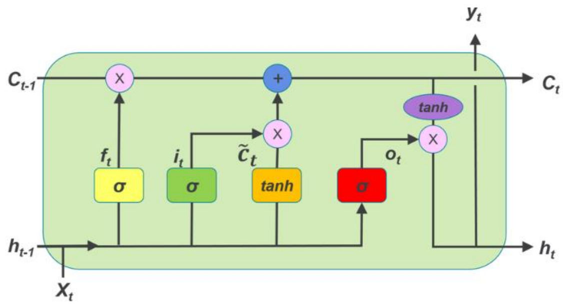

Figure 1. The LSTM Architecture [20]

Backpropagation optimization ensures that the weight and bias values in different layers of a neural network are updated by stepping in the gradient of the weight/bias values for the loss function. However, in the case of the vanishing gradient, it is not possible to update the weights/bias values or perform the learning. LSTM has been developed to solve the gradient vanishing problem encountered in RNNs and is widely used in time series estimation and natural language processing [19].

The LSTM architecture consists of an entry gate, an exit gate, and a forget gate. These gates regulate the flow of information entering and leaving the cell. They also allow cells to enable data to be passed or discarded optionally. The mathematical representation of the LSTM architecture given in Figure 1 is as follows.

$$f_t = \sigma(W_{f,x} * X_t + W_{f,h} * h6_{t-1} + b_f) \quad \textbf{(2.5)}$$
$$i_t = \sigma(W_{i,x} * X_t + W_{i,h} * h_{t-1} + b_{i)}) \quad \textbf{(2.6)}$$
$$\hat{c} = tan_h\, \sigma(W_{c,x} * X_t + W_{c,h} * h_{t-1} + b_c) \quad \textbf{(2.7)}$$
$$C_t = (C_{t-1} * f_t + i_t * \hat{C}_t) \quad \textbf{(2.8)}$$
$$o_t = \sigma(W_{o,x} * X_t + W_{o,h} * h_{t-1} + b_o) \quad \textbf{(2.9)}$$
$$h_t = o_t * tan_h\,(C_t) \quad \textbf{(2.10)}$$

Where, W represents the weight matrix, b represents the bias vector, i_t represents the input gate, f_t represents the forget gate, and o_t represents the output gate. Eq (2.5)

Celal Bayar University Journal of Science
Volume 19, Issue 1, 2023, p 39-46
Doi:10.18466/cbayarfbe.1070691

decides what information to enter through the gate using X_t and h_{t-1} as inputs. With the help of the sigmoid activation function (σ) used in the forget gate, it determines whether the information to be entered will pass to the next stage (if its value is 1) or not (if its value is 0).

Eq (2.6) decides which information to update. With Eq (2.7), a vector of new candidate information is created in the tanh layer. Then, with Eq (2.8), new information is created. Finally, using Eq (2.9) and Eq (2.10), the output is obtained. [21, 22].

2.3 Model Selection

The dates of the first cases observed in the countries used in this study differ. In addition, many factors affect the total number of patients and the number of daily cases and deaths, such as countries' health policies, the quarantine measures they take, the restrictive standards they apply, and the hygiene practices. However, all these factors were ignored in the study. The problem is treated as a univariate time series modeling problem using only observed death and case numbers.

The performance of a forecasting model is determined by comparing the observed values with the predicted values. In this study, RMSE, MAPE, MAE metrics, which are frequently preferred, are used to assess the performance of time series forecasting models. The RMSE measures the distance between the actual value and the predicted value, that is, the spread of the prediction error. The MAE is a linear score that measures the mean size of all individual errors in the predictions, regardless of their direction. MAPE expresses the reliability of the estimation methods as a percentage. If there are zeros among the actual values, the MAPE value cannot be calculated due to a division by zero error. The percentage error cannot exceed 100% for very low forecast values, but there is no upper limit to the percent error for very high forecast values. MAPE is biased in that it systematically chooses a method with very low estimates. RMSE, MAPE, and MAE are calculated using the formulas given in Eq (2.11), Eq (2.12), and Eq (2.13) respectively [23].

$$\boldsymbol{RMSE} = \sqrt{\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i})^{\boldsymbol{2}}} \qquad \textbf{(2.11)}$$

$$MAPE = \frac{\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i}}{n} * 100 \qquad \textbf{(2.12)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \textbf{(2.13)}$$

3. Case Study

As a case study, the daily number of Covid-19 cases and daily deaths in seven countries (Germany, France, Ireland, Italy, Poland, Russia, and Turkey) were examined. Data from March 2020 to October 2021 were used in the modeling performed with ARIMA and LSTM. Then, the 10-day number of cases and deaths between October 31, 2021, and November 9, 2021, for each country was estimated.

3.1 Forecasting with ARIMA

Eviews 10 program was used in ARIMA modeling. First of all, for the number of cases and deaths by country, stability control was carried out with the ADF (Augmented Dickey Fuller) test. ARMA models were applied to the stationary data sets. Non-stationary data sets were made stationary by taking the first and second differences. AR and MA levels in the models were determined by examining the ACF (autocorrelation) and PACF (partial autocorrelation) graphs, respectively. Models selected according to AIC, SIC, and Adjusted R^2 criteria are given in Table 1. The models with the lowest AIC and SIC values and the highest Adjusted R^2 values were selected as the best model among the possible model. Four data sets (France-Death, Ireland-Case, Turkey-Death and Case) are modeled with ARMA as they are stationary at the level. All other data sets are modeled with ARIMA since they become stationary at first difference.

Table 1. Selected Models for Countries

Countries		ARMA/ARIMA	Adj. R^2	AIC	SIC
Germany	Case	(4,1,4)	0.57	19.49	19.56
	Death	(6,1,2)	0,82	12,11	12,17
France	Case	(5,1,7)	0.59	21.25	21.35
	Death	(4,4)	0.72	12.57	12.65
Ireland	Case	(2,2)	0.84	14.87	14.91
	Death	(3,1,2)	0.45	8.09	8.14
Italy	Case	(5,1,10)	0.57	17.26	17.38
	Death	(5,1,2)	0.94	10.93	10.99
Poland	Case	(5,1,4)	0.94	17.76	17.84
	Death	(2,1,8)	0.68	11.13	11.22
Russia	Case	(6,1,5)	0.35	15.61	15.71
	Death	(5,1,2)	0.48	-0.57	-0.50
Turkey	Case	(4,0)	0.98	17.90	17.94
	Death	(2,0)	0.98	7.80	7.83

3.2 Forecasting with LSTM

Each country's death and case numbers are modeled through the LSTM architecture using the Keras library with the Python programming language. 70% of the datasets are reserved for training and 30% for testing. The LSTM architecture is built with 100 neurons in each hidden layer and one neuron in the output layer.

Table 2. Optimal Values for LSTM Models

Countries		Epoch	Batch Size	Look Back
Germany	Case	100	40	28
	Death	80	100	28
France	Case	100	15	28
	Death	100	30	28
Ireland	Case	100	70	28
	Death	150	80	28
Italy	Case	100	100	28
	Death	100	80	28
Poland	Case	200	18	28
	Death	150	30	28
Russia	Case	130	20	28
	Death	100	50	28
Turkey	Case	180	150	28
	Death	100	50	28

Optimum values of the epoch, batch size, and look-back parameters were determined separately for each model by manual search method. The model was trained with randomly different parameter groups, and the performances were observed until the best results were obtained by making the prediction error as small as possible. The parameter values determined at the end of trial-and-error execution procedures are given in Table 2.

3.3 Performance Comparison of Models

The number of cases and deaths between 31.10.2021 and 9.11.2021 for all countries was estimated by the models created with ARIMA and LSTM. The prediction performances of the models were compared using RMSE, MAPE, and MAE metrics. The results obtained are given in Table 3. The metrics used can range from 0 to ∞, but values close to 0 represent better performance.

The results obtained in this study generally showed that the prediction performance of LSTM is higher. According to the RMSE metric, the difference between the performances of the two models was much higher, while it was relatively small compared to the MAE metric. In addition, the daily changes in the forecast values produced by the models are given graphically in Figure 2.

Tablo 3. Performance Comparison of Models

		RMSE		MAPE		MAE	
Countries		ARIMA	LSTM	ARIMA	LSTM	ARIMA	LSTM
Germany	Case	8494.88	**7902.24**	289.56	**231.77**	**7347.9**	7445.6
	Death	59.22	**54.44**	**310.46**	435.31	**46.9**	48.2
France	Case	3549,14	**3012.22**	81.36	**69.72**	3284	**2743.3**
	Death	28.10	**21.14**	152.97	**73.85**	24.2	**16.6**
Ireland	Case	1330,49	**796.48**	28.83	**14.88**	1060.7	**546.7**
	Death	16,79	**16.65**	NAN	NAN	12.3	**7.1**
Italy	Case	983.79	**937.64**	**18.90**	19.32	**746.7**	790.3
	Death	**12.31**	16.52	**32.32**	47.88	**11.6**	14.8
Poland	Case	4538.88	**958.17**	34.30	**24.52**	3921.7	**2510.2**
	Death	78.53	**0.95**	357.76	**205.88**	64.1	**51.5**
Russia	Case	3009.30	**991.00**	6.55	**2.08**	2531.5	**810.6**
	Death	**44.97**	51.99	**3.32**	4.35	**38.5**	50.2
Turkey	Case	3456.58	**1874.75**	11.41	**5,28**	3261.1	**1517.7**
	Death	17.74	**17.20**	**5.75**	6.39	**12.8**	13.9

The solid red line in the graphs shows the actual number of deaths and cases. The green dashed line shows the ARIMA prediction values, and the blue dashed line shows the LSTM values. When the graphs are examined, it is seen that the models are successful in estimating the general trend but cannot catch extreme changes.

4. Conclusion

Despite the vaccines developed and all the precautions taken, the Covid-19 pandemic still continues. This study aims to predict the future values of the number of Covid-19 cases and deaths daily. Many factors affect the spread of Covid-19, the number of cases and deaths, such as age, chronic diseases, and social relations. It has been effective on the number of cases and deaths in vaccines applied to improve herd immunity, but Covid-19 still poses a significant threat to human and public health due to emerging new variants. In this study, ignoring all other effects, we focused only on the number of cases and deaths and predicted the number of cases and deaths for the next ten days with univariate time series models.

For the purpose of the study, seven countries with the highest number of daily cases and deaths in the first peak period of the pandemic were selected, and their data were modeled using ARIMA and LSTM methods.

Celal Bayar University Journal of Science
Volume 19, Issue 1, 2023, p 39-46
Doi:10.18466/cbayarfbe.1070691

In the literature, there are studies carried out with a similar purpose. However, many of these studies were published in the first year of the pandemic. In general, they examined a single country as a case study. However, the course of the epidemic differs from country to country, and as the amount of data used in modeling increases, the performance of time series models increases. Our study differs from those in the literature regarding both the countries studied and the amount of data used in the modeling. For the seven countries covered in the study, separate estimation models were obtained using the number of cases and deaths from March 2020 to October 2021. These models estimated the 10-day cases and deaths from 31 October 2021.

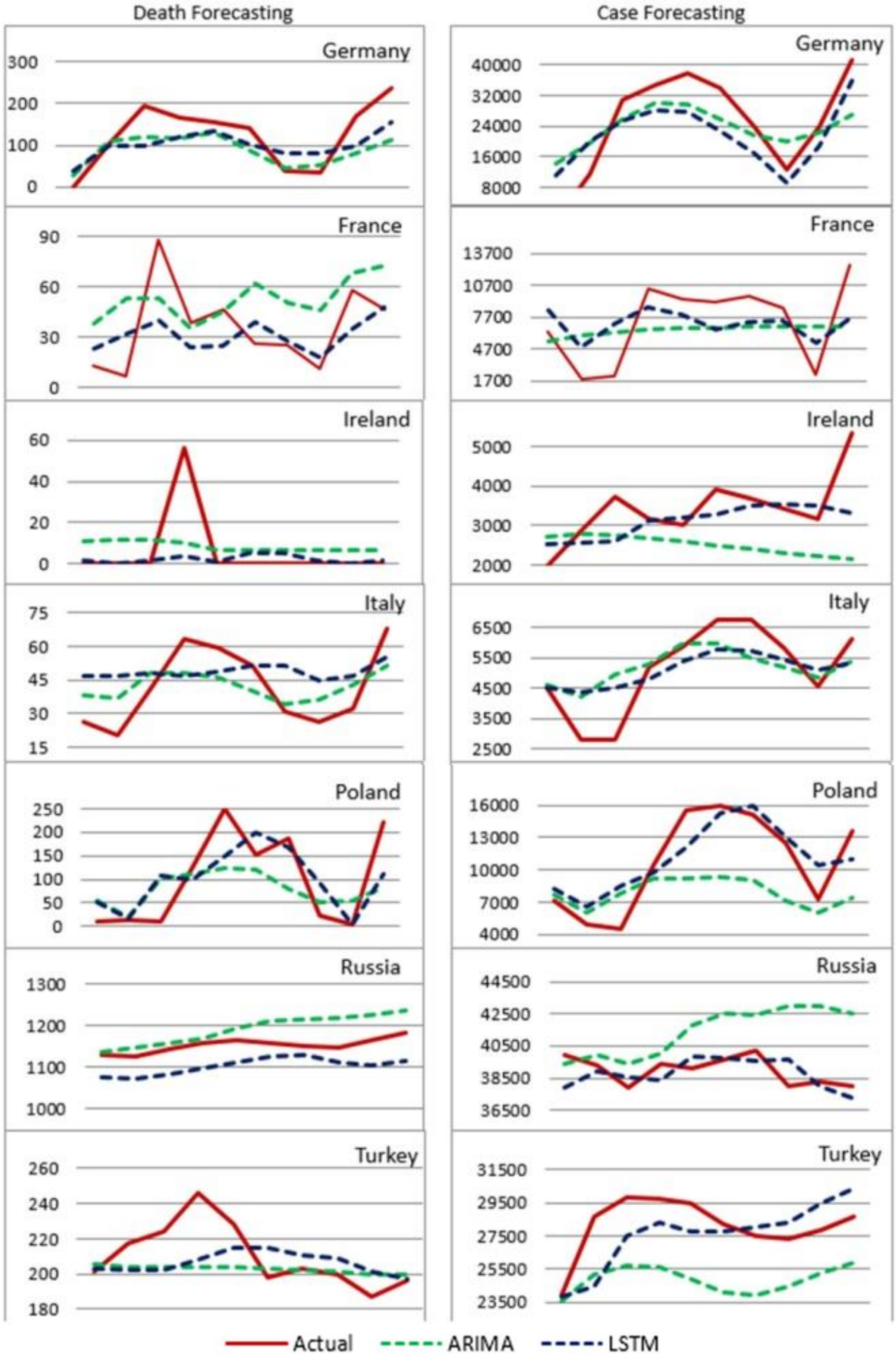

Figure 2. Graph of ARIMA and LSTM forecast versus actual values for 10-days cases and deaths.

In the results in Table 3, while the ARIMA models were more successful in estimating the Italian and Russian death numbers according to the RMSE, the LSTM's prediction success was higher in the remaining twelve models. According to MAPE values, while ARIMA's prediction performance was good for five models, LSTM was more successful in eight models. MAPE values could not be calculated as there were no deaths on some days in Ireland during the forecast period. According to MAE values, LSTM produced successful predictions with six models versus eight models. When the results are evaluated in terms of all metrics, while the prediction performance of LSTM is higher for seven models, ARIMA produced more successful predictions in only two models.

When the actual values and the estimated values of the numbers of deaths and cases given in Figure 1 are examined, it is seen that the LSTM models are more successful than ARIMA in forecasting the structure of the accurate data. However, due to the factors that were not taken into account in the univariate time series analysis, the extreme values occurring in the real data could not be captured in both models. ARIMA models estimated the numbers of deaths and cases for Russia and cases for Turkey and Ireland as far greater than their actual values. The LSTM model, on the other hand, estimated the number of deaths in Russia smaller than their actual values.

The results of the study showed that the ARIMA and LSTM models were successful in predicting the short-term future of the Covid-19 outbreak. In future studies, multivariate time series methods and different deep learning methods can be used to capture shock effects and increase prediction accuracy.

Author's Contributions

Büşra ÇETİN: Drafted and wrote the manuscript, performed the experiment and result analysis.
Nida GÖKÇE NARİN: Assisted in analytical analysis on the structure, supervised the experiment's progress, result interpretation and helped in manuscript preparation.

Ethics

There are no ethical issues after the publication of this manuscript.

References

[1]. Khan F, Gupta R. 2020. ARIMA and NAR based prediction model for time series analysis of COVİD19 cases in India *Journal of Safety Science & Resilience*, 12-18.

[2].Özen N, Saraç S. ve Koyuncu M. 2021 Prediction of COVID-19 Cases in the United States of America with Machine Learning Algorithms, *European Journal of Science and Technology Special Issue 22, pp.* 134-139.

[3]. Sevli, O. & Başer, V. G. 2020 Machine Learning Based Case Estimation Using Prophet Model with Time Series Data for Covid-19 Outbreak, *European Journal of Science and Technology* No. 19, 827-835.

[4]. Awan T. M., Aslam F. 2020. Prediction of daily COVID-19 cases in European counties using automatic ARIMA model J Public Health Res. 9(3),1765.

[5].Kırbaş İ., Sözen A., Tuncer A.D., Kazancıoğlu F.Ş. 2020 Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches *Chaos Solitons Fractals* 110015

[6]. Hernandez-Matamoros, A., Fujita, H., Hayashi, T., & Perez-Meana, H. 2020. Forecasting of COVID19 per region using ARIMA models and polynomial functions. *Applied Soft Computing*, 96, 106610.

[7]. Roy, S., Bhunia, G. S., & Shit, P. K. 2021. Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Modeling earth systems and environment*, *7*(2), 1385-1391.

[8]. Moftakhar, L., Mozhgan, S. E. I. F., & Safe, M. S. 2020. Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. *Iranian Journal of Public Health*.

[9]. Alzahrani, S. I., Aljamaan, I. A., & Al-Fakih, E. A. 2020. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of infection and public health*, 13(7), 914-919.

[10]. Elsheikh, A. H., Saba, A. I., Abd Elaziz, M., Lu, S., Shanmugan, S., Muthuramalingam, T., ... & Shehabeldeen, T. A. 2021. Deep learning-based forecasting model for COVID-19 outbreak in Saudi Arabia. *Process Safety and Environmental Protection*, 149, 223-233.

[11]. Ala'raj, M., Majdalawieh, M., & Nizamuddin, N. 2021. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infectious Disease Modelling*, 6, 98-111.

[12]. Eroğlu Y. 2020. Forecasting Models For Covid-19 Cases of Turkey Using Artificial Neural Networks and Deep Learning,*Journal of Industrial Engineering* 31(3), 354-372.

[13]. URL https://github.com/owid/covid-19-data/tree/master (accessed at 12.11.2021).

[14]. Box, G.E.P. and Jenkins, G.M. 1970. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

[15]. Özmen A., 1986 Zaman Serisi Analizinde Box-Jenkins Yöntemi ve Banka Mevduat Tahmininde Uygulama Denemesi, Anadolu Üniversitesi Yayınları, 207, Eskişehir.
(Özmen A.Box-Jenkins Method in Time Series Analysis and Application Trial in Bank Deposit Estimation, Anadolu University Press, 207, Eskisehir 1986)

[16]. Akdi, Y. Zaman Serileri Analiz, *Genelleştirilmiş2. Baskı, Gazi Kitabevi, Ankara.2010*
(Akdi, Y. Time Series Analysis, *Generalized 2nd Editionı, Gazi Bookstore, Ankara.2010)*

[17]. Hamzaçebi, C. ve Kutay, F., 2004.Electric Consumption Forecasting of Turkey Using Artificial Neural Networks up to Year

2010", Journal of The Faculty of Engineering and Architecture of Gazi University, (19), No 3, 227-233.

[18]. Hochreiter & Schmidhuber, 1997. Long Short-Term Memory, Neural Computation 9(8):1735-1780.

[19]. Kara A. 2019. Global Solar Irradiance Time Series Prediction Using Long Short-Term Memory NetworkJournal of Science, Gazi UniversityGU J Sci, Part C, 7(4): 882-892.

[20]. T. W. C. B. Aya Abdelsalam Ismail, 2018. Improving Long-Horizon Forecasts with Expectation-Biased,arXiv:1804.06776 [cs.LG]

[21]. M. Yuan, Y. Wu, L. Lin, 2016.Fault Diagnosis and Remaining Useful Life Estimation of Aero Engine Using LSTM Neural Network, IEEE International Conference on Aircraft Utility Systems (AUS), 135–140.

[22]. Olah, C. 2015. Understanding LSTM Networks. August 7,2021, colah.github.io: colah.github.io/posts/2015-08-Understanding-LSTMs

[23]. Emang, D., Shitan, M., Abd Ghani, A. N., & Noor, K. M. 2010. Forecasting with univariate time series models: A case of export demand for peninsular Malaysia's moulding and chipboard. Journal of Sustainable Development, 3(3), 157.