doi.org/10.36721/PJPS.2022.35.1.REG.141-149.1

Forecasting the daily deaths caused by COVID-19 using ARIMA model

Amber Asghar[1], Hina Khan[2], Nida Razzak[1], Zara Waseem[1], Midhat Salman[3] and Tanveer Hussain[4]*

[1]Department of Statistics, Virtual University of Pakistan, Lahore, Pakistan
[2]Department of Statistics, Government College University, Lahore, Pakistan
[3]University of Health Sciences, Lahore, Pakistan
[4]Department of Molecular Biology, Virtual University of Pakistan, Lahore, Pakistan

Abstract: Coronavirus disease (COVID-19) pandemic has recently become a global health crisis. On the basis of this study the data reported from ten different countries on confirmed daily deaths caused by COVID-19. By fitting the linear regression models based on the data from ten countries to find the relationship between the new cases and deaths reported daily. We also used the autoregressive integrated moving average model (ARIMA) to predict the potential number of daily deaths caused by COVID-19 in these countries in the next 3 Months. The R^2value obtained for Iran (0.24) implies that 24% of daily deaths correspond to the daily cases. The $R^2$of Pakistan 0.662 which indicates that 66.2% of daily deaths are explained by our predictor variable. In Turkey 70.2% of daily deaths are explained by daily cases and India recorded the highest number of deaths while UAE had the lowest number of deaths. Our results suggest that the pandemic is under control in China, UAE and Australia. Pakistan, Iran, Germany and Italy however, showed an upward trend in the spread of the disease, which may correlate with a high increase in death rate as the data indicated.

Keywords: Coronavirus, pandemic, ARIMA, prediction, trends.

INTRODUCTION

In 1937, scientists found 15 different types of corona viruses that infected the mammals and birds (Beaudette, 1937). Later in 1960s, the first human corona virus was detected in the nasal sample of a patient with common cold (Hamre and Procknow, 1966). A novel corona virus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a new strain of Coronavirus (a large family of viruses) that was reported in Wuhan City, China in December 2019 possibly from a zoonotic source (https://www.who.int/docs/default-source/ coronaviruse/ situation-reports/20200423-sitrep-94-covid-19.pdf?sfvrsn =b8304bf0_4). The virus is small in size ranging in diameter between 66-81nm (Park *et al.,* 2019). The World Health Organization has (WHO) designated the coronavirus disease caused by the virus as COVID-19. A large number of infections have now been reported outside China, with the infections spread over more than 200 countries worldwide. There is a difference among countries in the epidemiological investigations and detection capacity for the affected cases. The disease is extremely contagious with variable symptoms ranging from appearance of no symptoms to severe respiratory problems, so the prevalence of the virus in a population is difficult to estimate correctly. Since there is currently no established treatment available for the disease, the management of the spread of the disease by taking appropriate protective measures and the preparation of healthcare system to deal with the health crisis is crucial. According to the WHO reports from 8th June 2020, about 7 million COVID-19 cases and 0.4 million deaths have been reported from around the world. The number of COVID-19 cases is still rising vigorously in some parts of the world. The current situation has put immense burden on the health sectors and has also badly deteriorated the global economy due to the strict measures taken by the governments. This unprecedented situation shows the need of the availability of a statistical modeling tool that can help in forecasting the possible number of daily deaths caused by COVID-19. Such a tool will be helpful in managing the pandemic by preparing the health sector for providing suitable services to the patients in order to minimize daily deaths.

Linear regression analysis is the most commonly used technique of all statistical techniques. It is a modeling technique where the regression estimates explain the association between one predicted variable and one or more predictor variables (Kumari K and Yadav S 2018). In clinical or biomedical research, the analysis is used to understand two or more independent variables to predict an outcome variable. ARIMA models were first suggested by (Box and Jenkins, 1970) and have numerous applications in various fields such as their use for forecasting in marketing, industry production, social problems, medical research and economics etc. In medical applications, time series forecasting models have been effectively applied to forecast the prevalence of diseases, estimate the death rate and evaluate the time dependent risk. ARIMA is also a good choice for predicting the hospital outpatient visits using a combinatorial model

**Corresponding author:* e-mail: tanveer.hussain@vu.edu.pk

(Lou *et al*., 2017). In a recent study from Iran, the health expenditures were forecasted using the ARIMA model (Ramezanian *et al*., 2019). ARIMA models are also being used for forecasting the COVID-19 confirmed cases in different countries from January 22th, 2020 to March 1, 2020 and predicted results for the next 17 days (Dehesh, 2020). ARIMA model has advantages over the short term of the series model as it provides the exact forecasting. However, it has the limitation of requiring at least 50 or more observations to forecast.

In this article we fitted linear regression model to evaluate if the independent variable (daily new cases) perform efficiently in predicting a dependent variable (daily deaths). We then used ARIMA model for forecasting the daily deaths caused by COVID-19 in ten countries. This study involved the number of cases and deaths reported from ten different countries that included Spain, Italy, Turkey, China, Australia, Germany, UAE, Iran, India and Pakistan. We have included most of the countries in which the population is highly affected by COVID-19. The best ARIMA model has been identified for each country. The daily cases of deaths have been forecasted for 3 months at 95% Confidence limit.

MATERIALS AND METHODS

Data Source

The data used in this study included the number of COVID-19 cases and deaths reported from July 1, 2020 to September 30, 2020 obtained from the website of worldmetero (https://www.worldometers.info/corona virus/). We randomly selected ten countries most of which had their population highly affected by COVID-19 (table 1).

Mathematical Model

In this study, we have used linear regression analysis to identify the relationship between the new cases and deaths of COVID-19 reported daily in different countries. We have also used autoregressive integrated moving average (ARIMA) model on the data to predict the potential number of daily deaths caused by COVID-19 in different countries. We differenced the data to make it stationary by removing trends or seasonal structures for using it in ARIMA model. Minitab 18 software was used to develop the ARIMA models.

ARIMA Models

These models are the most general class of models for predicting a time series. A non-seasonal ARIMA model is classified as an ARIMA (p, d, q) models. Where, p is number of auto regressive term, d number of non-seasonal differences required for stationarity and q is number of lagged forecast errors in the prediction equation. In ARIMA model, we differenced the data to make it stationary. A model that shows stationarity is one that indicates that there is constancy to the data over time. Most of the time data show trends, so by differencing trends or seasonal structures are removed.

Table 1: Data of COVID-19 cases and death in 10 countries reported on September 30, 2020 (https://www.worldometers.info/coronavirus/).

Countries	Cases	Deaths
Spain	819549	31791
Italy	314861	35894
Germany	292911	9571
Turkey	318663	8195
China	85403	4634
Iran	457219	26169
India	6310267	98708
UAE	94190	419
Australia	27078	886
Pakistan	312263	6479

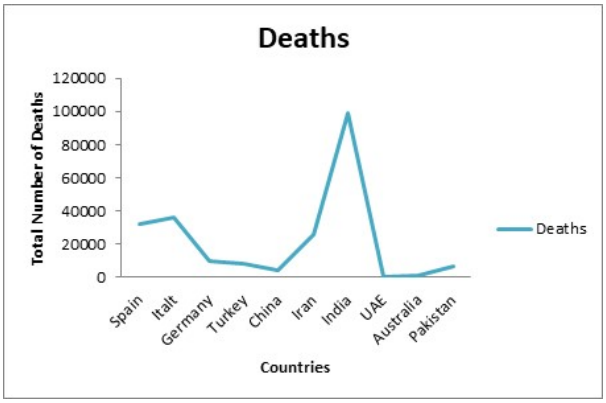


Fig. 1: Total numbers of deaths caused by COVID-19 in 10 different countries

A time series Y_t is said to follow the ARMA (p,q) model if:

$$Y_t = O_o + O_1 Y_{t-1} + O_2 Y_{t-2} + \ldots + O_p Y_{t-p} + \varepsilon_t - w_1\varepsilon_{t-1} - w_2\varepsilon_{t-2} - \ldots - w_q\varepsilon_{t-q}$$

Where p and q are greater than zero, p refers to autoregressive AR, q refers to moving average MA and ε_t is a white noise. An extension of ordinary autoregressive moving average model (ARMA) is autoregressive integrated moving average model ARIMA (p, d, q) follows as: $O(B)(1-B)^{d}Y_t = W(B)\varepsilon_t$

Where, Y_t denote a non-stationary time series at time t, d is the order of differencing and ε_t is a white noise (mean zero and constant variance) and B represent the backward shift operator.

STATISTICAL ANALYSIS

In the present study, we used the linear regression and ARIMA models. Linear regression fits a straight line that minimizes the discrepancies between predicted and actual output values and ARIMA models are used for forecasting the future values based on past results. The details of the models are already discussed in the above section.

Supplementary table 1

Model Summary				
Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
Pakistan	0.814	0.662	0.659	14.560
Turkey	0.838	0.702	0.698	11.202
Spain	0.669	0.448	0.442	40.777
India	0.939	0.882	0.881	80.330
Italy	0.077	0.006	-0.005	8.496
Australia	0.005	0.000	-0.011	11.145
China	0.139	0.019	0.008	0.473
Iran	0.491	0.241	0.233	31.457
Germany	0.446	0.199	0.190	3.594
UAE	0.235	0.055	0.045	0.963

Predictors: Daily New Cases, For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

Countries	ANOVA[a,b]					
	Model	Sum of Squares	Df	Mean Square	F	Sig.
Pakistan	Regression Residual Total	37424.856 19078.220 56503.076	1 90 91	37424.856 211.980	176.549	.000[b]
Turkey	Regression Residual Total	26551.122 11293.084 37844.207	1 90 91	26551.122 125.479	211.599	.000[b]
Spain	Regression Residual Total	121316.037 149650.833 270966.870	1 90 91	121316.037 1662.787	72.959	.000[b]
India	Regression Residual Total	4353144.979 580759.238 4933904.217	1 90 91	4353144.979 6452.880	674.605	.000[b]
Italy	Regression Residual Total	38.360 6496.803 6535.163	1 90 91	38.360 72.187	0.531	.468[b]
Australia	Regression Residual Total	0.227 11178.512 11178.739	1 90 91	0.227 124.206	0.002	0.966[b]
China	Regression Residual Total	0.397 20.157 20.554	1 90 91	0.397 0.224	1.774	.186[b]
Iran	Regression Residual Total	28270.845 89061.590 117332.435	1 90 91	28270.845 989.573	28.569	.000[b]
Germany	Regression Residual Total	288.812 1162.351 1451.163	1 90 91	288.812 12.915	22.362	.000[b]
UAE	Regression Residual Total	4.888 83.547 88.435	1 90 91	4.888 0.928	5.265	.024[b]

Linear Regression through the Origin, Predictors: Daily New Cases, this total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

Coefficients[a,b]					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std.Error	Beta		
Pakistan	.015	.001	.814	13.287	.000
Turkey	.062	.004	.838	14.546	.000
Spain	.010	.001	.669	8.542	.000
India	.026	.000	.983	58.971	.000
Italy	.001	.001	.077	0.729	.468
Australia	.000	.007	.005	0.043	.966
China	.003	.002	.139	1.332	.186
Iran	.039	.007	.491	5.345	.000
Germany	.003	.001	.446	4.729	.000
UAE	.001	.000	.235	2.295	.024

Dependent Variable, Linear Regression through the Origin

Table 1: Autocorrelation (ACF) and Partial Autocorrelation (PACF) plots of different countries

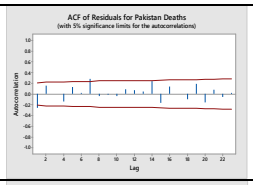
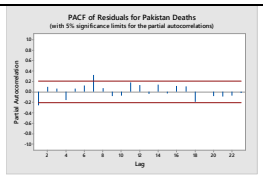
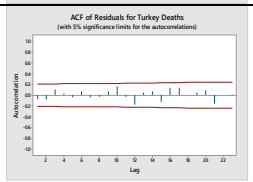
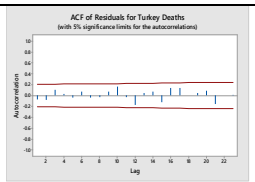
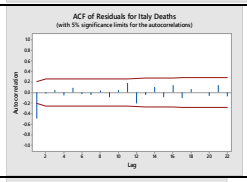
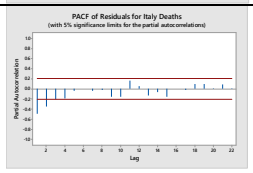
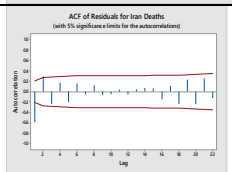
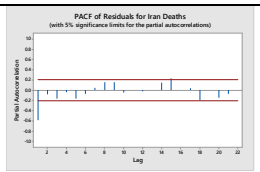
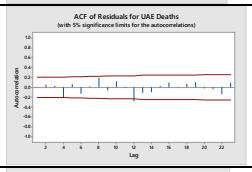
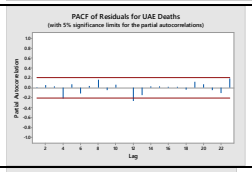
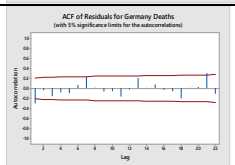
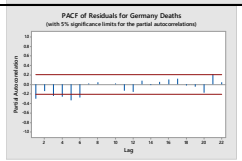
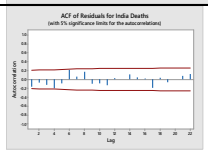
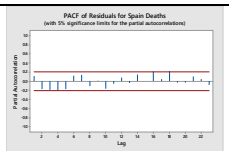
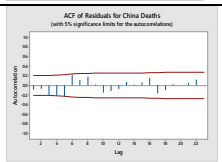
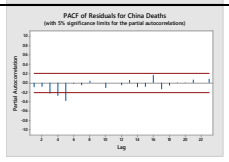
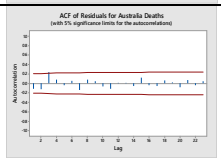
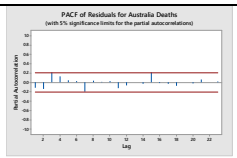
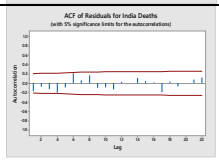
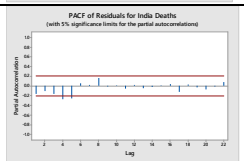
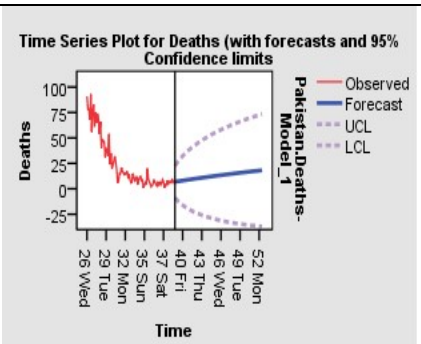
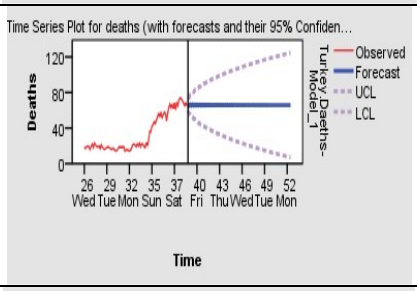
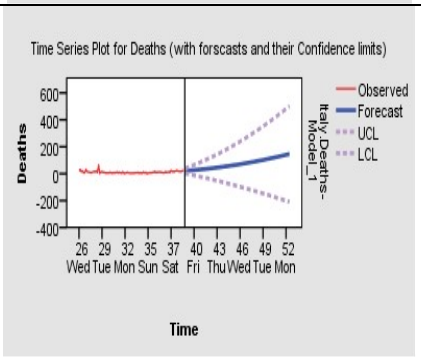
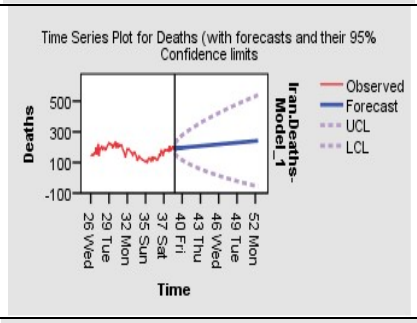
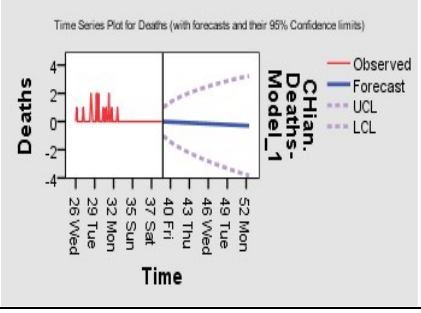
Countries	Models	ACF	PACF
Pakistan	ARIMA(1,0,1)		
Turkey	ARIMA(1,1,1)		
Italy	ARIMA(0,2,3)		
Iran	ARIMA(2,2,2)		
UAE	ARIMA(1,0,1)		
Germany	ARIMA(0,2,1)		
Spain	ARIMA(2,0,1)		
China	ARIMA(4,1,0)		
Australia	ARIMA(2,0,1)		
India	ARIMA(4,2,3)		

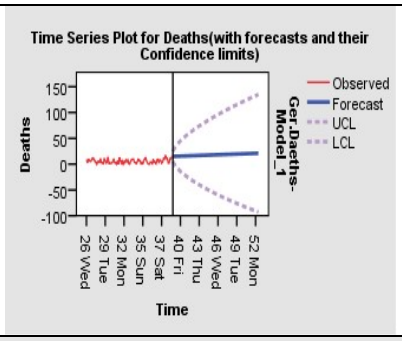
Table 2: Time series plot of deaths of different countries (with forecast and their 95% Confidence limit)

Countries	Estimates of Parameters	Plots
Pakistan	Type Coef SE Coef T-Value P-Value AR 1 0.9987 0.0107 93.69 0.000 MR 1 0.5808 0.0985 5.90 0.000	
Turkey	Type Coef SE Coef T-Value P-Value AR 1 -0.9637 0.0687 -14.02 0.000 MA 1 -0.8478 0.1170 -7.25 0.000	
Italy	Type Coef SE Coef T-Value P-Value MA 1 1.1003 0.0054 205.38 0.000 MA 2 -0.2307 0.1148 -2.01 0.048 MA 3 0.1454 0.1180 1.23 0.221	
Iran	Type Coef SE Coef T-Value P-Value AR 1 -0.2564 0.1399 -1.83 0.000 AR 2 0.0917 0.1343 0.68 0.070 MA 1 0.6847 0.0188 36.37 0.000 MA 2 0.3620 0.0490 7.38 0.000	
China	Type Coef SE Coef T-Value P-Value AR 1 -0.8299 0.1039 -7.98 0.000 AR 2 -0.5792 0.1333 -4.34 0.000 AR 3 -0.2795 0.1333 -2.10 0.039 AR 4 -0.2395 0.1038 -2.31 0.023	

Continue…

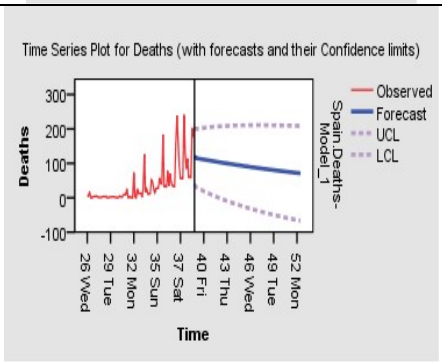
Germany

Type	Coef	SE Coef	T-Value	P-Value
MA 1	1.0025	0.0091	109.93	0.000



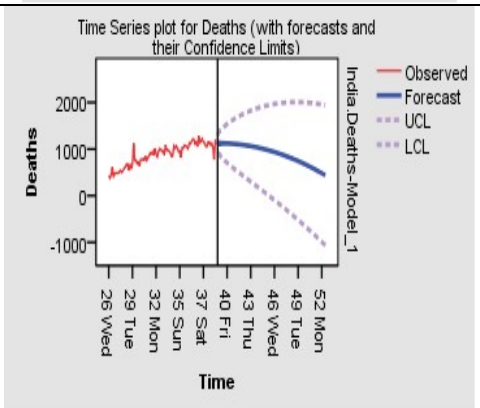
Spain

Type	Coef	SE Coef	T-Value	P-Value
AR 1	1.1408	0.1450	7.87	0.000
AR 2	-0.1266	0.1363	-0.93	0.355
MR 1	0.8901	0.1179	7.55	0.000



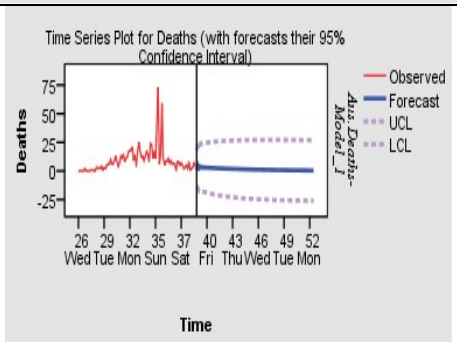
India

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-2.1244	0.1271	-16.72	0.000
A R 2	-1.7863	0.2843	-6.28	0.000
AR 3	-0.7413	0.2835	-2.61	0.011
AR 4	-0.2441	0.1218	-2.00	0.048
MA 1	-0.8034	0.0934	-8.60	0.000
MA 2	-0.7607	0.0981	7.76	0.000
MA 3	-0.9568	0.0655	14.60	0.000



Australia

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.4585	0.1297	3.54	0.001
AR 2	0.5181	0.1228	4.22	0.000
MA 1	-0.1234	0.1014	-1.22	0.227
MA 2	0.7528	0.0799	9.42	0.000



UAE

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.9999	0.0030	335.75	0.000
MA 1	0.9692	0.0476	20.38	0.000

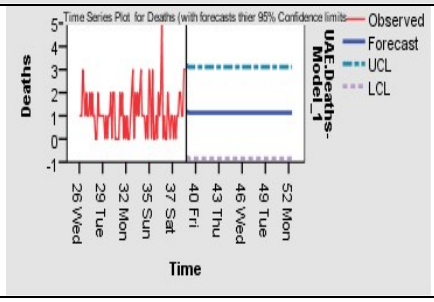


Table 3: Prediction of Covid-19 Daily death cases for 3 Months according to ARIMA model with 95% Confidence limit

Date	Spain	Italy	Germany	Iran	Pakistan
10/1/2020	122(41,203)	21(4,39)	15(6,2)	198(162,234)	7(-9,22)
10/2/2020	116(32,200)	22(1,42)	15(2,29)	189(148,229)	7(-10,24)
10/3/2020	115(30,200)	22(-1,45)	15(-1,32)	197(146,247)	7(-11,25)
10/4/2020	114(28,201)	23(-3.49)	15(-4,34)	191(137,246)	7(-12,27)
10/5/2020	114(26,201)	24(-5,52)	15(-6,37)	197(135,258)	7(-13,28)
10/6/2020	113(24,202)	24(-7,56)	15(-8,39)	193(127,260)	8(-14,29)
10/7/2020	112(23,202)	25(-9,59)	15(-10,41)	197(125,269)	8(-15,30)
10/8/2020	112(21,203)	26(-10,62)	16(-12,43)	195(119,271)	8(-15,31)
10/9/2020	111(19,203)	27(-12,66)	16(-14,45)	198(117,278)	8(-16,32)
10/10/2020	111(18,204)	28(-14,69)	16(-16,47)	197(112,281)	8(-17,33)
10/11/2020	110(16,204)	28(-16,73)	16(-17,49)	198(109,288)	8(-18,34)
10/12/2020	109(15,204)	29(-18,76)	16(-19,50)	198(105,291)	8(-18,35)
10/13/2020	109(13,205)	30(-20,80)	16(-20,52)	199(102,297)	9(-19,36)
10/14/2020	108(11,205)	31(-22,84)	16(-22,54)	199(98,300)	9(-19,37)
10/15/2020	108(10,205)	32(-24,87)	16(-23,55)	200(95,305)	9(-20,38)
10/16/2020	107(8,206)	33(-26,91)	16(-24,57)	200(92,309)	9(-20,39)
10/17/2020	107(7,206)	34(-28,95)	16(-26,58)	201(89,317)	9(-21,39)
10/18/2020	106(6,206)	34(-30,99)	16(-27,59)	202(83,321)	9(-21,40)
10/19/2020	105(4,207)	35(-32,102)	16(-28,61)	203(80,325)	9(-22,41)
10/20/2020	105(3,207)	36(-34,106)	16(-30,62)	203(78,329)	10(-22,42)
10/21/2020	104(1,207)	37(-36,110)	16(-31,64)	204(75,333)	10(-23,42)
10/22/2020	104(0,208)	38(-38,114)	16(-32,65)	204(72,337)	10(-23,43)
10/23/2020	103(-1,208)	39(-40,118)	17(-33,66)	205(69,340)	10(-24,44)
10/24/2020	103(-3,208)	40(-42,122)	17(-34,68)	205(67,344)	10(-24,44)
10/25/2020	102(-4,208)	41(-44,127)	17(-36,69)	206(64,348)	10(-24,45)
10/26/2020	102(-5,208)	42(-46,131)	17(-37,70)	207(62,351)	10(-25,46)
10/27/2020	101(-7,209)	43(-48,135)	17(-38,71)	207(59,355)	11(-25,46)
10/28/2020	100(-8,209)	45(-50,139)	17(-39,73)	208(57,358)	11(-26,47)
10/29/2020	100(-9,209)	46(-52,143)	17(-40,74)	208(57,358)	11(-26,46)
10/30/2020	99(-10,209)	47(-54,143)	17(-41,75)	209(55,362)	11(-26,48)
10/31/2020	99(-12,209)	48(-56,152)	17(-44,76)	209(52,365)	11(-27,48)
11/1/2020	98(-13,209)	49(-59,157)	17(-45,80)	210(489,372)	11(-27,49)
11/2/2020	98(-14,210)	50(-61,161)	17(-46,81)	210(45,375)	11(-27,50)
11/3/2020	97(-15,210)	51(-63,166)	17(-47,82)	211(43,379)	11(-27,50)
11/4/2020	97(-16,210)	53(-65,170)	17(-48,83)	211(41,382)	12(-28,51)
11/5/2020	96(-18,210)	54(-67,175)	18(-49,84)	212(39,385)	12(-28,52)
11/6/2020	96(-19,210)	55(-70,180)	18(-50,85)	212(36,389)	12(-28,52)
11/7/2020	95(-20,210)	56(-72,184)	18(-51,87)	213(34,392)	12(-29,53)
11/8/2020	95(-21,210)	57(-74,189)	18(-52,88)	214(32,395)	12(-29,53)
11/9/2020	94(-22,210)	59(-76,194)	18(-53,89)	214(30,398)	12(-29,54)
11/10/2020	94(-23,210)	60(-79,199)	18(-54,90)	215(28,401)	12(-29,55)
11/11/2020	93(-24,210)	61(-81,204)	18(-55,91)	215(26,405)	13(-30,55)
11/12/2020	93(-25,211)	63(-83,209)	18(-56,92)	216(24,,408)	13(-30,55)
11/13/2020	92(-26,211)	64(-86,214)	18(-57,93)	216(22,411)	13(-30,56)
11/14/2020	92(-27,211)	65(-88,219)	18(-58,94)	217(20,414)	13(-30,56)
11/15/2020	91(-28,211)	67(-90,224)	18(-59,95)	217(18,417)	13(-31,57)
11/16/2020	91(-30,211)	68(-93,229)	18(-60,96)	218(16,420)	13(-31,57)
11/17/2020	90(-31,211)	69(-95,229)	18(-60,97)	218(14,423)	13(-31,58)
11/18/2020	90(-32,211)	71(-97,239)	18(-61,98)	219(12,426)	13(-31,58)

Continue…

11/19/2020	89(-33,211)	72(-100,245)	18(-62,99)	220(10,429)	14(-31,58)
11/20/2020	89(-34,211)	74(-102,250)	18(-63,100)	220(8,432)	14(-32,59)
11/21/2020	88(-34,211)	75(-105,255)	19(-64,101)	221(6,435)	14(-32,60)
11/22/2020	88(-35,211)	77(-107,261)	19(-65,102)	221(5,438)	14(-32,60)
11/23/2020	87(-36,211)	78(-110,266)	19(-66,103)	222(3,441)	14(-32,61)
11/24/2020	87(-37,211)	80(-112,271)	19(-66,104)	222(1,444)	14(-32,61)
11/25/2020	86(-38,211)	81(-115,277)	19(-67,105)	223(-1,447)	14(-33,61)
11/26/2020	86(-39,211)	83(-117,283)	19(-68,106)	223(-3,449)	14(-33,62)
11/27/2020	85(-40,211)	84(-119,288)	19(-69,107)	224(-4,452)	14(-33,62)
11/28/2020	85(-41,211)	86(-122,294)	19(-70,108)	224(-6,452)	15(-33,62)
11/29/2020	84(-42,211)	87(-124,299)	19(-71,109)	225(-8,458)	15(-33,63)
11/30/2020	84(-43,211)	89(-127,305)	19(-71,110)	225(-10,461)	15(-33,63)
12/1/2020	83(-45,211)	91(-129,311)	19(-72,111)	226(-11,464)	15(-34,63)
12/2/2020	83(-45,211)	92(-132,317)	19(-73,112)	226(-13,466)	15(-34,64)
12/3/2020	82(-46,211)	94(-135,323)	19(-74,112)	227(-15,469)	15(-34,64)
12/4/2020	82(-47,211)	96(-137,328)	19(-75,114)	227(-17,472)	15(-34,65)
12/5/2020	81(-48,211)	97(-140,334)	19(-76,115)	228(-18,475)	15(-34,65)
12/6/2020	81(-49,211)	99(-142,340)	20(-77,116)	228(-20,477)	16(-34,65)
12/7/2020	80(-50,211)	101(-145,346)	20(-78,117)	229(-21,480)	16(-34,66)
12/8/2020	80(-50,211)	102(-147,352)	20(-78,118)	230(-23,483)	16(-35,65)
12/9/2020	80(-51,211)	104(-150,358)	20(-79,119)	230(-25,486)	16(-35,66)
12/10/2020	79(-52,210)	106(-153,365)	20(-80,120)	231(-26,488)	16(-35,67)
12/11/2020	79(-53,210)	108(-155,371)	20(-81,121)	232(-28,491)	16(-35,67)
12/12/2020	78(-54,210)	110(-158,377)	20(-81,121)	232(-30,494)	16(-35,67)
12/13/2020	78(-54,210)	111(-160,383)	20(-82,122)	233(-31,496)	16(-35,68)
12/14/2020	78(-55,210)	113(-163,389)	20(-83,123)	233(-33,499)	16(-35,68)
12/15/2020	77(-56,210)	115(-166,396)	20(-84,124)	234(-34,502)	17(-35,68)
12/16/2020	77(-57,210)	117(-168,402)	20(-84,125)	234(-36,504)	17(-36,69)
12/17/2020	76(-57,210)	119(-171,409)	20(-85,126)	235(-37,512)	17(-36,69)
12/18/2020	76(-58,210)	121(-174,415)	20(-86,127)	235(-39,509)	17(-36,69)
12/19/2020	75(-59,210)	123(-176,422)	20(-87,127)	236(-40,512)	17(-36,70)
12/20/2020	75(-60,210)	124(-179,428)	21(-87,128)	236(-42,515)	17(-36,70)
12/21/2020	75(-60,210)	126(-182,435)	21(-88,129)	237(-43,517)	17(-36,70)
12/22/2020	74(-61,210)	128(-184,441)	21(-89,130)	237(-45,520)	17(-36,71)
12/23/2020	74(-62,210)	130(-187,448)	21(-90,132)	238(-46,522)	17(-36,71)
12/24/2020	73(-63,209)	132 (-190,455)	21(-91,133)	238(-48,525)	17(-36,71)
12/25/2020	73(-63,209)	134(-193,461)	21(-91,133)	239(-49,527)	18(-37,72)
12/26/2020	73(-64,209)	136(-195,468)	21(-92,134)	239(-51,530)	18(-37,72)
12/27/2020	72(-65,209)	138(-198,475)	21(-93,135)	240(-52,533)	18(-37,72)
12/28/2020	72(-64,209)	140(-201,482)	21(-93,135)	240(-53,535)	18(-37,73)
12/29/2020	72(-66,209)	142 (-204,488)	21(-94,136)	241(-55,538)	18(-37,73)
12/30/2020	71(-66,209)	145(-206,495)	21(-94,136)	241(-55,539)	18(-37,73)
12/31/2020	71(-67,209)	147(-209,502)	21(-95,137)	242(56-,540)	18(-37,73)

RESULTS

Fig. 1 represents the total deaths caused by COVID-19 in ten different countries. According to the plot, India recorded the highest number of deaths while UAE had the lowest number of deaths.

By using linear regression analysis, we identified the relationship between the new daily cases and deaths of COVID-19 in 10 different countries. Results of linear regression model are given in Supplementary table 1.

The R^2value obtained for Iran (0.24) implies that 24% of daily deaths correspond to the daily cases. The $R^2$of Pakistan 0.662 which indicates that 66.2% of daily deaths are explained by our predictor variable. In Turkey 70.2% of daily deaths are explained by daily cases. Only 5.5% of daily deaths are explained by UAE. Only 1.9% of China daily deaths are explained by its daily cases. In Germany 19.9% of daily deaths are explained by daily cases. In India 88.2% of daily deaths are explained by daily cases. Only 0.6% of daily deaths are explained by Italy. Only 0% of daily deaths are explained by Australia. The

R^2value obtained for Spain (0.44) implies that 44% of daily deaths correspond to the daily cases. The reasons of unexplained death cases in the reported data for all countries may be the diagnosed cases are not true indicative of the extent of the spread of the disease in the population or testing capacity may not be enough.

The ANOVA table shows the usefulness of the linear regression model so, we expected the p-value to be significant (<0.05) and the values were indeed significant for all the countries, which indicates the usefulness of our linear regression model. Coefficient table provides the quantification of the relationship between daily cases and daily deaths. We fitted best the ARIMA model for COVID-19 daily death data from the ten countries to be implemented for forecasting.

Table 1, we develop the graph of the residual of ACF and PACF to check whether our model meet the assumption that residuals are independent or not. As we see that the spikes of ACF and PACF of Australia, Turkey and Spain are within the bands so there is no significant correlation and conclude that residuals are independent so our model fits the data. In the plot of Pakistan, Italy, Iran, UAE and China one spike is slightly outside the band that shows significant correlation but this may be due to random error.

Table 2 shows the estimates of parameters and the forecast plots of ARIMA model at 95 % Confidence limit for different countries with high number of daily death cases. It is observed that P is less than the significant value which is 0.05 so we conclude that these ARIMA models are best fit model.

Tables 3 represent the forecast for 3 months with 95% confidence interval for ten different countries. This table indicates that India have shown a downward trend after 10th November. China shows almost control the pandemic and zero death. Pakistan, Iran, Germany and Italy indicate the upward trend for next 3 months which show more deaths. Australia and UAE show the stable trend. Turkey shows stable trend after 10th October.

DISCUSSION

Our results suggest that the pandemic is under control in China, UAE and Australia. Pakistan, Iran, Germany and Italy however, showed an upward trend in the spread of the disease, which may correlate with a high increase in death rate as indicated by the data. It is essential to stress that our rough estimate is based upon the past available data and there may be a change based upon the daily number of new cases reported in future. This study has been performed to provide a model that may be used for predicting the death rate caused by COVID-19 to help the government authorities to take necessary actions for managing the current health crises in the country.

REFERENCES

Beaudette FR (1937). Cultivation of the virus of infectious bronchitis. *J. Am. Vet. Med.,* **90**: 51-60.

Box GE and Jenkins GM (1970). Time series analysis: Forecasting and control Holden-Day. San Francisco.

Dehesh T, Mardani-Fard HA and Dehesh P (2020). Forecasting of covid-19 confirmed cases in different countries with Arima models. *Med Rxiv*. doi: https://doi.org/10.1101/2020.03.13.20035345.

Hamre D and Procknow JJ (1966). A new virus isolated from the human respiratory tract. *Proceedings of the Society for Experimental Biology and Medicine,* **121**(1): 190-193.

Kumari K and Yadav S (2018). Linear regression analysis study. *J. Pract. Cardiovasc. Sci.,* **4**(1): 33-36.

Luo L, Luo L, Zhang X and He X (2017). Hospital daily outpatient visits forecasting using a combinatorial model based on ARIMA and SES models. *BMC Health Serv. Res.,* **17**(1): 469.

Park WB, Kwon NJ, Choi SJ, Kang CK, Choe PG and Kim JY (2019). Virus isolation from the first patient with SARS-CoV-2 in Korea. *J. Korean Med. Sci,,* 35(7): e84.

Ramezanian M, Haghdoost AA, Mehrolhassani MH, Abolhallaje M, Dehnavieh R and Najafi B et al (2019). Forecasting health expenditures in Iran using the ARIMA model (2016-2020). *Med. J. Islam. Repub. Iran,* **33**(1): 143-146.