مجلة البحوث المالية والتجارية

المجلد (22) – العدد الرابع – أكتوبر 2021

Statistical Analysis of Covid-19 data of Egypt 14 March 2020 - 3 November 2021 and forecasting the fourth wave

By

Samar Ahmed Helmy
Lecturer, Department of Statistics, Mathematics & Insurance
Faculty of Commerce - Port Said University

رابط المجلة: https://jsst.journals.ekb.eg/

Abstract

In this paper, we study the statistical characteristics of the spread of the corona virus epidemic (COVID-19) in Egypt by using spectral analysis and Box-Jenkins analysis of time series , taking into account mildly and symptomatically infected individuals reported daily by the Ministry of Health of Egypt. Many ARIMA models were considered and estimated their parameters, the best model used through the study. The period from 14 March 2020 to 3 November 2021 was studied. Three waves of COVID-19 epidemic are distinguished. A fourth wave is progressing , its main characteristics are presented. The number of new infections through the fourth wave are forecasted. Another approach using Normal distribution model, derived from the actual daily data, is used to study the epidemic wave by wave. Its results cope with those of ARIMA model.

Keywords: **Covid-19, Corona waves, ARIMA model, parameters estimation, periodogram, Normal model**

1. Introduction

The corona virus disease 2019 (COVID-19) is a public health epidemic affecting the whole world, including Egypt. It is caused by the dangerous respiratory syndrome Corona virus 2 (SARS-CoV-2) [(Wu et al 2020) and (Gorbalenya et al 2020)]. The disease spreads among people, most often when physically nearby or over long distances, specially indoors [3,4]. The most important symptoms are fatigue, shortness of breath, loss of smell, loss of appetite, fever, muscle pain, dry cough, and coughing up sputum [5,6].

Egypt's Ministry of Health announced the first case involving a Chinese person on 14 February 2020 at Cairo International Airport. In early March, multiple foreign severe respiratory syndrome corona virus 2 (COVID-19) two cases associated with travel to Egypt were reported – including two cases from the United States, two cases from Tunisia (Plus several potential cases as the two initial cases were part of 1,000 quarantined Tunisian football supporters who visited Egypt from 27 February to 1 March) two cases from France, one case from Canada, and one case from Taiwan.

On 8 March 2020 : A 60-year-old German citizen died in Hurghada,

On 9 March: The World Health Organization (WHO) announced that there were 56 confirmed cases in Egypt. On the same day, the Egyptian Ministry of Tourism discovered 3 cases within the people who work on a floating hotel.

On 13 March: A second Tunisian coming back from Egypt tested positive for the virus and Tunisia officially added Egypt to a list of outbreak areas and closed its borders to it and imposed a quarantine to any person coming from Egypt.

On 24 March 2020, the Egyptian authorities introduced various types of lockdowns (Helmy 2019) to mitigate the effects of the disease.

The ministry of health issues, regularly, a daily corona report as from 14 March 2020 including the number of new cases, deaths and recovered. The peak of COVID-19 occurred in Egypt on 19 June 2020, when the new cases reached 1774, followed by a significant decrease of

the number of cases. This first wave reached its bottom on 22 August 2020 when the new cases dropped to 89 case.

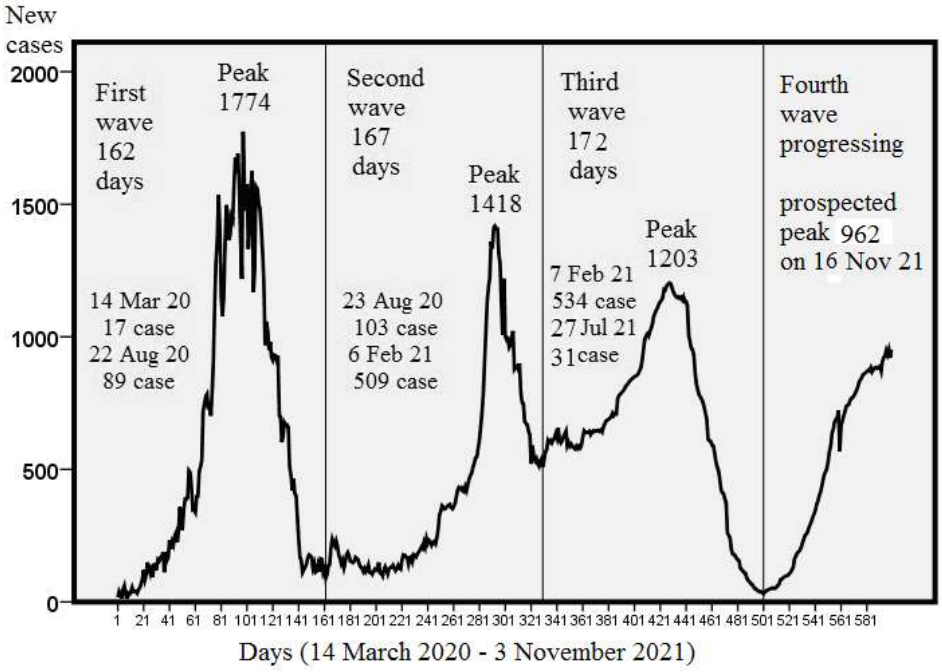
Afterwards, the second wave started. New cases increased gradually to reach its peak at on 31 December 2020 realizing 1418 new cases. New cases went down to reach 509 case on 6 February 2021.

Lastly, the third wave started. New cases increased at a lower rate than the previous waves. It reached its peak at 1203 case on 15 May 2021 This wave ended at 31 case on 27 July 2021.

As from 28 July 2021 new cases increased gradually to reach 399 case on 8 September 2021. It is the start of a new fourth wave .

The following figure (1) shows the new cases of Covid-19 through the three waves and the prospected fourth wave.

Fig (1)



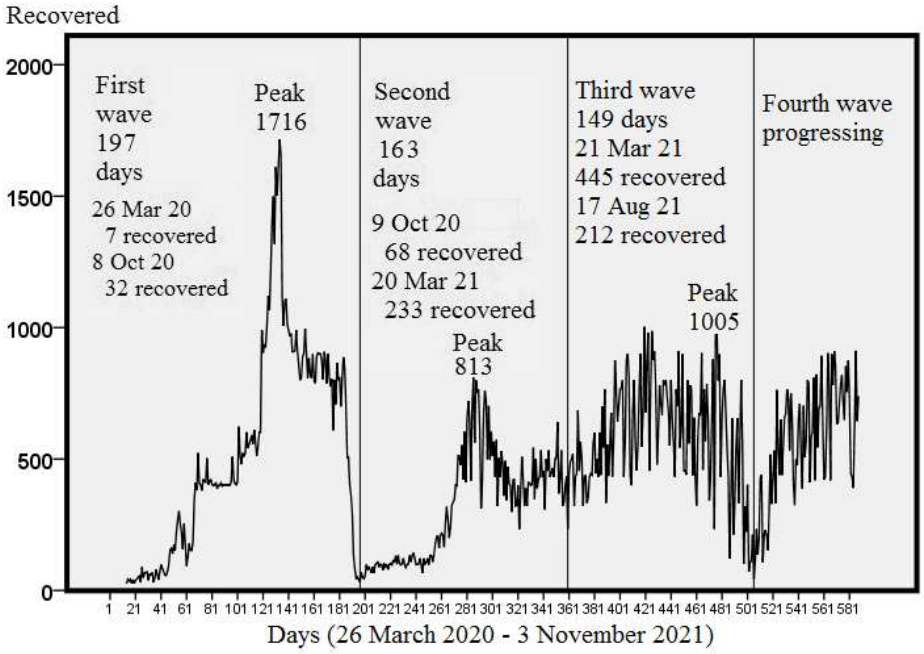
New cases of Covid-19 through the four waves

First recovered cases were recorded on 26 March 2020, i.e 12 days after the first regular registry of new cases. Recovered cases (nearly) follow the same pattern of the new cases. Length of the waves is declining, the first wave continued for 197 days, the second lasted for

163 days, the third went down to 155 days, a fourth wave started as from 17 August 2021 reaching 739 cases on 3 November 2021.

The following figure (2) shows the recovered cases of Covid-19 through the three waves and the prospected fourth wave

Fig (2)



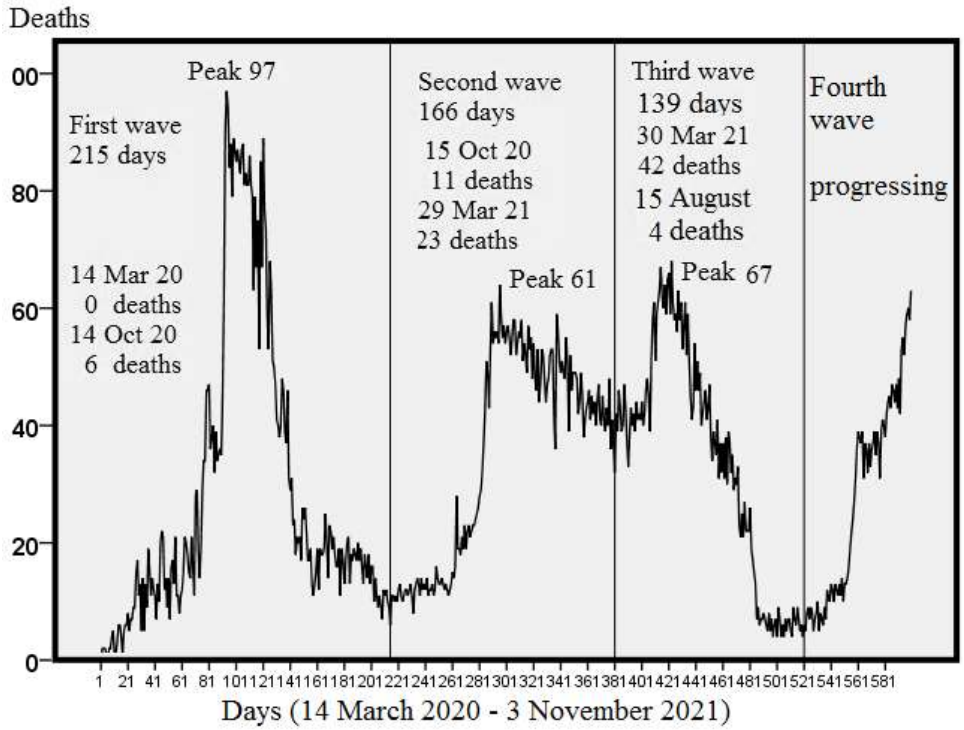
Recovered cases of Covid-19 through the four waves

Deaths also followed the same pattern. Three waves are observed. Similarly, length of the waves is declining from 215 to 166 to 155 day, of course the third wave is continuing

It is expected that recovered and deaths cases follow the same pattern as the new cases, simply because new cases are the source of them. As the recovery and/or death occurs after some time from infection, varying from case to case, the length of recovered and deaths waves are longer than that of the new cases. A new fourth wave started as from 16 August 2021 to reach 63 deaths on 3 November 2021.

The following figure (3) shows the deaths of Covid-19 through the four waves

Fig (3)



Deaths of Covid-19 through the four waves

COVID-19 pandemic in Egypt may be considered a mild case compared to the world and some countries presented in table (1). Concerning the new cases, Egypt comes in the 77th order out of 223 country and location.

Table (1)
COVID-19 pandemic in Egypt, the world and the highest infected countries on 3 November 2021

Order	Country,	Total Cases	Total Deaths	Total Recovered
	World	223,995,028	4,619,829	200,583,033
1	USA	41,561,079	674,547	31,744,880
2	India	33,163,004	442,046	32,321,283
3	Brazil	20,958,899	585,205	20,002,562
4	UK	7,132,072	133,841	5,722,272
5	Russia	7,084,284	190,376	6,340,151
7	Turkey	6,590,414	59,170	6,055,819
8	Iran	5,237,799	112,935	4,481,814
9	Argentina	5,218,993	113,099	5,062,115
11	Spain	4,903,021	85,218	4,529,411
12	Italy	4,590,941	129,766	4,331,257
14	Egypt	4,058,940	93,040	3,793,000
77	**Egypt***	**291,585**	**16,836**	**244,793**
78	Slovenia	273,529	4,462	260,699
79	Bahrain	273,454	1,388	271,127
184	Sint Maarten	3,921	57	3,623
185	New Zealand	3,847	27	3,195
186	Brunei	3,831	17	2,313
223	Micronesia	1	0	1

*** Data include cases before 14 March 2020 and after 3 November 2021**
Source: https://www.worldometers.info

Relative COVID-19 pandemic in Egypt also shows that the case is not as severe as in the world or the countries presented in table (2). New cases in Egypt were 2787 per million representing 9.7% of the world average, so it comes in the 175th place

Table (2)
COVID-19 new cases and deaths per million in a descending order on 8 September 2021

New cases per Million population			Deaths per Million population		
Order	World	28,736	Order	World	592.7
1	Seychelles	207,861	1	Peru	5,926
8	Bahrain	154,362	2	Hungary	3,123
9	Maldives	149,455	3	Bosnia and Herz.	3,063
10	Georgia	143,955	4	North Macedonia	2,968
15	USA	124,692	5	Gibraltar	2,880
16	Israel	122,875	6	Montenegro	2,842
20	Argentina	114,230	7	Czechia	2,834
25	France	105,093	9	Brazil	2,730
28	UK	104,407	11	Argentina	2,475
32	Brazil	97,776	20	USA	2,024
36	Kuwait	94,495	22	Tunisia	2,009
49	Qatar	83,423	23	Poland	1,995
59	Italy	76,064	24	UK	1,959
62	UAE	72,446	28	Spain	1,822
82	Oman	57,585	30	France	1,763
84	Tunisia	56,836	50	Russia	1,304
93	Russia	48,520	56	Lebanon	1,201
94	Egypt	48,262	65	Jordan	1,019
97	Iraq	46,968	78	Israel	784
123	India	23,754	79	Bahrain	784
134	Saudi Arabia	15,391	80	Oman	776
138	Japan	12,722	97	Kuwait	559
142	Djibouti	11,748	104	Indonesia	499
148	Brunei	8,658	117	India	317
151	Mauritania	7,197	122	Saudi Arabia	243
161	S. Korea	5,212	123	Qatar	215
165	Algeria	4,449	124	UAE	205
175	**Egypt***	**2,787**	**136**	**Egypt***	**161**
192	Benin	1,529	137	Djibouti	156
199	Nigeria	932	142	Japan	131
200	Sudan	837	145	Algeria	123
210	Yemen	270	203	China	3
223	*MS Zaandam*		223	Micronesia	

* **Data include cases before 14 March 2020 and after 8 September 2021**
Source: https://www.worldometers.info

Deaths were 161 per million making 27.2% of the world average. Table (2) also shows the relative COVID-19 deaths in Egypt, the world and

some countries. Egypt comes in the 136th place among all countries and places of the world.

2. Review

Quite a number of research papers were published studying the epidemiological features of COVID-19 and how to mitigate its effects on public health. (Tosepu et al. 2020) studied the correlation between COVID-19 pandemic and the atmospheric temperature in Jakarta, Indonesia. (Batista 2020) used the logistic growth model to predict the final size and the peak of the COVID-19 epidemic in China, South Korea, and the rest of the world. In their paper (Boldog, et al 2020), a computational tool was introduced to assess the risks of new coronavirus outbreaks outside China. (Almeshal, et al. 2020) predicted the size of the corona epidemic in Kuwait based on the confirmed corona data and used stochastic and deterministic modeling approaches. (Kuniya 2020), investigated the corona epidemic in Japan and forecasted the peak of infection using the traditional SEIR model . (Röst et al. 2020), proposed an epidemiological and statistical study of the early phase of the COVID-19 outbreak in Hungary and developed an age-structured compartmental model to study alternate post-lockdown scenarios. In their work (Bantan, et al 2020) the authors developed an exponential family of continuous distributions to provide new statistical models. Several researchers have studied COVID-19 in Iraq (see, e.g., [Sarhan, et al 2020, A.R.; Flaih, et al 2020 and Jebril, N, 2020]).

Using several regression models, (Amar et al. 2020) searched the actual COVID-19 database of Egypt from 15 Feb. till 15 Jun. 2020 to forecast the number of COVID-19-infected patients and measured the

final epidemic scale. Using the epidemic SIR model (Hasab et al. 2020) studied the epidemiological distribution and modeling of the COVID-19 epidemic in Egypt . The SEIR compartmental model was used to study COVID-19 epidemic in Egypt and to predict the time of the peak of this epidemic based on the daily reported cases and deaths of the ministry of Health (Anwar, W.A.; AbdelHafez, A.M. 2020). Using SIR and SEIR compartment models. El Desouky 2020 studied the corona epidemic in Egypt to predict the time of the possible peak and to model the changes induced by the social behavior of the Egyptians during the Holy Month of Ramadan

3. Data and Methods

3.1. COVID-19 Data

The data for new cases, recovered and deaths were collected from the daily corona report published on https://twitter.com/mohpegypt website. of the Ministry of Health. This work focuses on the data from 14 March 2020 till 8 September 2021, i.e 544 daily data.

The data used in the study is restricted to those registered in the official quarantine hospitals. So, the data do not include infected people who were homely quarantined, recovered or dead.

The data for daily maximum and minimum temperature collected from the web site https://accuweather.com.

Compared data were collected from https://Worldometer.info website

3.2. Methods

Spectral analysis :

Any time series can be expressed as a combination of cosine and sine waves with differing periods (how long it takes to complete a full cycle) and amplitudes (maximum/minimum value during the cycle). This fact can be utilized to examine the periodic (cyclical) behavior in a time series.

A periodogram is used to identify the dominant periods (or frequencies) of a time series. This can be a helpful tool for identifying the dominant cyclical behavior in a series, particularly when the cycles are not related to the commonly encountered monthly or quarterly seasonality.

Properties of a Cosine Function:

For discrete time (meaning time t = integer values), these definitions are useful for a cosine (or sine) wave:

Period

(T) is the number of time periods required to complete a single cycle of the cosine function.

Frequency

Is ω = 1/T. It is the fraction of the complete cycle that's completed in a single time period.

Imagine fitting a single cosine wave to a time series observed in discrete time. Suppose that we write this cosine wave as

$$x_t = A \cos(2\pi\omega t + \phi)$$

- **A is the amplitude. It determines the maximum absolute height of the curve.**
- **ω is the frequency. It controls how rapidly the curve oscillates.**
- **ϕ is the phase. It determines the starting point, in angle degrees, for the cosine wave.**

Also the Box-Jenkins procedure will be used to analyze the data. The Box-Jenkins methodology is a strategy for identifying, estimating and forecasting autoregressive integrated moving average models.

Other statistical methods will be used when needed.

SPSS-20 is used for necessary calculations

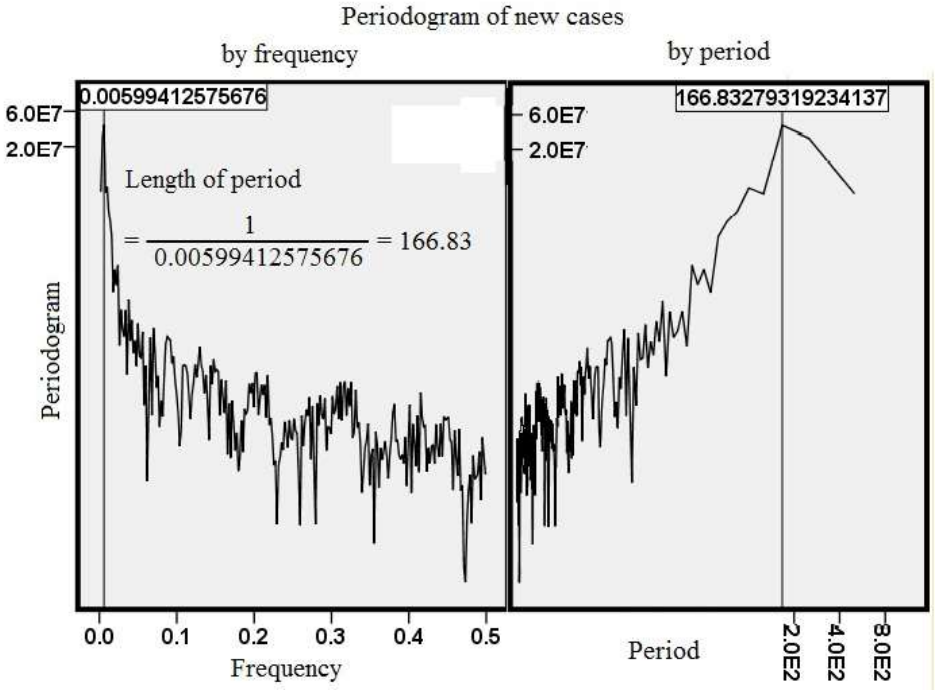
4 - New cases

Spectral analysis of the new cases shows that the length of the period is 166.83 days The next Fig (4), as concluded earlier in Fig (1), also Fig (5) confirms this fact.

Now we are facing a new fourth wave , will continue for 167 days ? When it will reach its peak ?

If we consider the previous three waves, the peaks were 1774 , 1418 and 1203 respectively. The peak of the second wave was 84.84% of the peak of the first wave.

Fig. (4)



Spectral analysis of new cases

The peak of the third wave was 79.93% of the peak of the second wave. If we apply theses two ratios, then the peak of the fourth wave is expected to lie between 962 and 1020 new case. This wave is expected to last for 167 days as analyzed above.

Correlation between new cases and the maximum daily temperature is negative very week non significant, but it is a little higher with the minimum daily temperature and also negative and non significant.

Table (3)
Correlation new cases and max and min daily temperature

		Max Temp**	Min Temp
New cases	Pearson Correlation	-.059-	-.205-**
	Sig. (2-tailed)	.179	.000
	N	527	527

****. Correlation is significant at the 0.01 level (2-tailed).**

The Box-Jenkins method is an iterative application of the next three steps:

1. Identification.

Using graphical presentation of the data, autocorrelations, correlogram, partial autocorrelations, and other information, a class of simple ARIMA models is selected. This helps to estimate appropriate values for p, d, & q

2. Estimation.

The parameters of the selected model are estimated using maximum likelihood techniques, backcasting, etc., as outlined in Box-Jenkins (1976).

3. Diagnostic Checking.

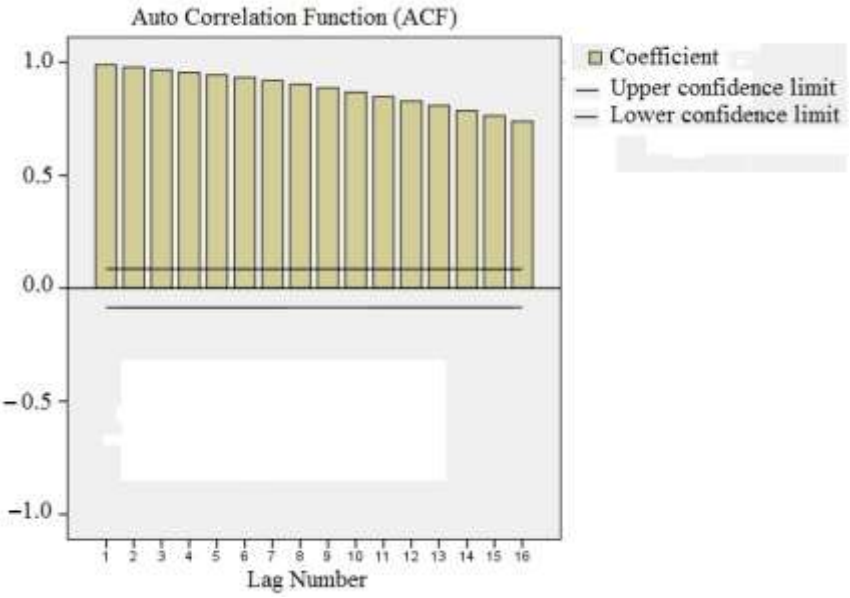
The autocorrelations of the residual series (the series of residual, or error, values) are used to check for inadequacies of the model.

These steps are applied until step three does not produce any improvement in the model. Now we go over these steps in detail

1 - Model identification

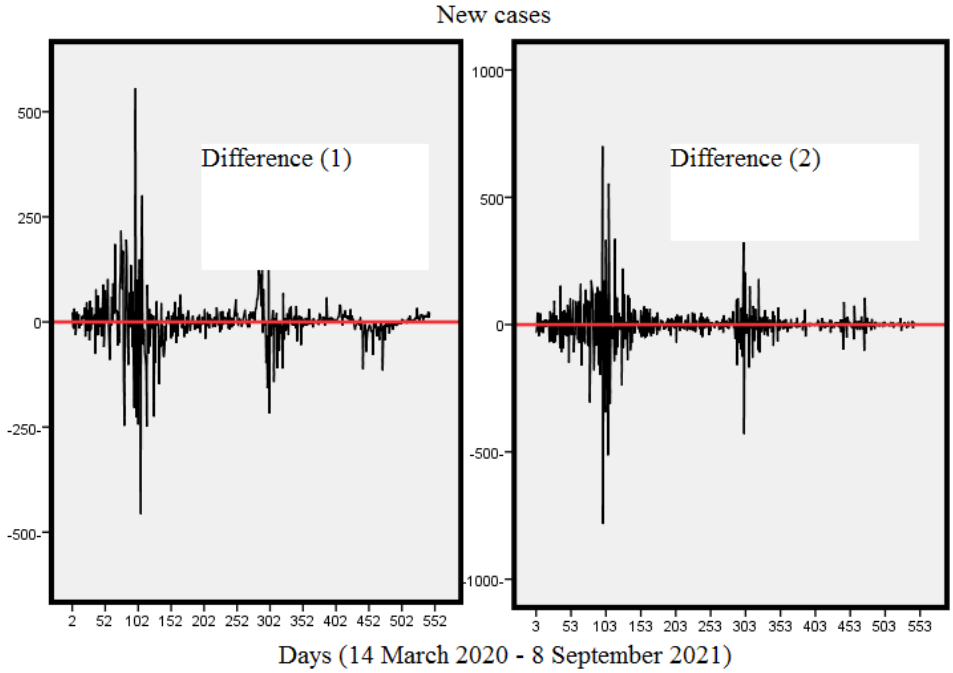
Autocorrelations calculated using SPSS are:

Fig, (5)



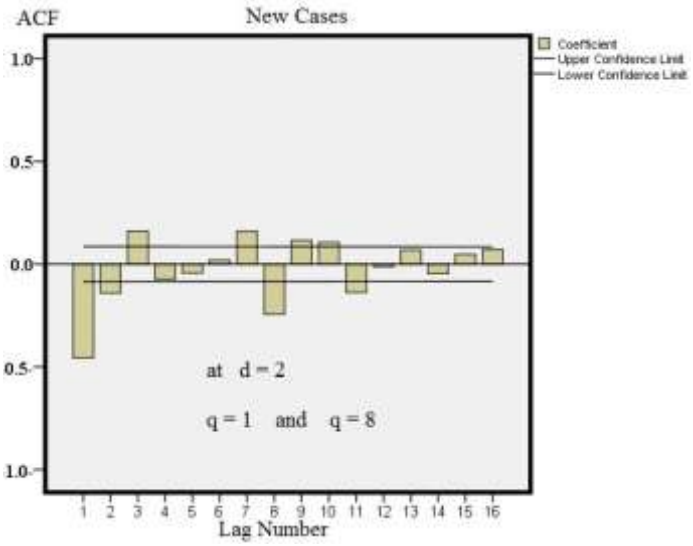
It is clear that the series is non stationary, so, we try the first difference to achieve stationarity, it did not work. Taking the second difference, the series reached stationarity as shown in the following figure,

Fig (6)



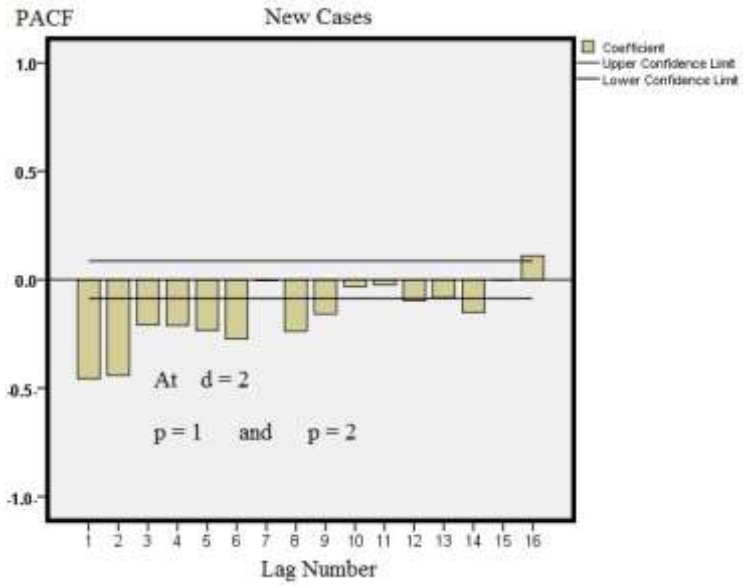
The autocorrelation function (ACF) shows that q = 1 and 8 when d = 2

Fig (7)



the partial auto correlations shows that p = 1 and 2

Fig (8)



Thus we have four models to estimate to pick the best of them,

Model (1) ARIMA(1, 2, 1)
Model (2) ARIMA(2, 2, 1)
Model (3) ARIMA(1, 2, 8)
Model (4) ARIMA(2, 2, 8)

2 - Model estimation

Table (4)
Comparison of the estimated parameters of the four models

Model	Parameters	Estimate	SE	t	Sig.	Remarks
ARIMA (1,2,1)	Constant	-.092-	.249	-.371-	.711	Constant & AR parameters are significant
	AR (Lag 1)	-.041-	.045	-.913-	.362	
	MA (Lag 1)	.955	.014	69.629	.000	
ARIMA (2,2,1)	Constant	-.084-	.317	-.264-	.792	Constant & AR(1) parameters are significant
	AR (Lag 1)	-.072-	.045	-1.594-	.111	
	AR (Lag 2)	-.156-	.045	-3.498-	.001	
	MA (Lag 1)	.929	.018	51.969	.000	
ARIMA (1,2,8)	Constant	-.083-	.320	-.258-	.797	Constant & AR & 5 MA parameters are significant
	AR (Lag 1)	-.377-	.258	-1.463-	.144	
	MA (Lag 1)	.597	.256	2.336	.020	
	MA (Lag 2)	.484	.250	1.936	.053	
	MA (Lag 3)	-.107-	.061	-1.768-	.078	
	MA (Lag 4)	.044	.067	.656	.512	
	MA (Lag 5)	-.048-	.061	-.793-	.428	
	MA (Lag 6)	-.096-	.060	-1.589-	.113	
	MA (Lag 7)	-.104-	.050	-2.077-	.038	
	MA (Lag 8)	.148	.043	3.449	.001	
ARIMA (2,2,8)	Constant	-.051-	.071	-.726-	.468	Constant & 7 MA parameters are significant AR parameters are not
	AR (Lag 1)	-.737-	.098	-7.543-	.000	
	AR (Lag 2)	-.689-	.097	-7.140-	.000	
	MA (Lag 1)	.242	1.216	.199	.842	
	MA (Lag 2)	.129	.919	.140	.888	
	MA (Lag 3)	.604	.762	.793	.428	
	MA (Lag 4)	.076	.059	1.283	.200	
	MA (Lag 5)	-.092-	.080	-1.154-	.249	
	MA (Lag 6)	-.009-	.069	-.123-	.902	
	MA (Lag 7)	-.161-	.078	-2.070-	.039	
	MA (Lag 8)	.210	.262	.803	.422	

The above table shows a comparison of the estimated parameters of the four models:

The four models reveal some significant parameters and the others are not. We study the forecasts of the four models compared to actual values to decide which model will be used in the rest of the study.

Table (5)
Comparison of the forecasts of the four models to actual data

Model	15/9	16/9	17/9	18/9	19/9	20/9	21/9	22/9	23/9	24/9
ARIMA(1,2,1)	516	528	541	554	567	580	593	606	619	633
Diff from actual	-2.8	-7.2	-8.0	-13.0	-13.2	-14.6	-13.8	-12.4	-14.3	11.4
ARIMA(2,2,1)	519	535	550	565	581	596	612	628	644	660
Diff from actual	-2.3	-6.0	-6.5	-11.3	-11.0	-12.2	-11.0	-9.2	-10.8	16.2
ARIMA(1,2,8)	519	533	551	568	585	599	616	631	648	664
Diff from actual	-2.3	-6.3	-6.3	-10.8	-10.4	-11.8	-10.5	-8.8	-10.2	16.9
ARIMA(2,2,8)	509	511	515	520	523	524	529	530	531	534
Diff from actual	-4.1	-10.2	-12.4	-18.4	-19.9	-22.8	-23.1	-23.4	-26.5	-6.0
Actual	531	569	588	637	653	679	688	692	722	568

It is clear the best forecasts are those of the model ARIMA(1,2,8). In addition, 7 out of 10 parameters are significant, the model is:

$$Y_t = -0.083 - 0.377\, Y_{t-1} + \varepsilon_t - 0.597\, \varepsilon_{t-1} - 0.484\, \varepsilon_{t-2} + 0.107\, \varepsilon_{t-3} - 0.044\, \varepsilon_{t-4} + 0.048\, \varepsilon_{t-5} + 0.096\, \varepsilon_{t-6} + 0.104\, \varepsilon_{t-7} - 0.148\, \varepsilon_{t-8}$$

Where:

Y daily new corona cases
t time in days
ε the white noise ~ $N(0,\sigma^2)$

3 - The fourth wave

The above model will be used to characterize the fourth wave which started on the 17th of July 2021. from the above spectral analysis, the length of the period is expected to last for 167 days. it is expected to reach its peak at 962 new cases on 16th November 2021, i.e after 122 days from its start which is almost half of the expected length of the fourth wave.

Table (6)
Forecasts of the fourth wave of new cases using ARIMA(1,2,8)

Date	Actual	Forecasts	Date	Actual	Forecasts
15 Sep 21	531	519	10	837	861
16	569	533	11	846	864
17	588	551	12	857	867
18	637	568	13	861	870
19	653	585	14	865	873
20	679	599	15	869	876
21	688	616	16	874	879
22	692	631	17	871	882
23	722	648	18	883	885
24	568	664	19	868	888
25	667	680	20	877	892
26	680	697	21	883	895
27	702	713	22	885	898
28	718	730	23	881	902
29	738	747	24	886	905
30	741	764	25	889	908
1 Oct 21	745	781	26	871	911
2	761	790	27	907	914
3	768	799	28	923	917
4	771	808	29	927	920
5	778	820	30	948	923
6	788	829	31	933	926
7	799	835	1 Nov. 21	951	929
8	811	845	2	921	932
9	831	851	3	951	935

The above table (6) shows the expected new cases up to the peak on 16 November 2021.

We may summarize the aspects of the fourth wave as follows:

Start of the wave : 28 July 2021

End of the wave : 25 February 2022

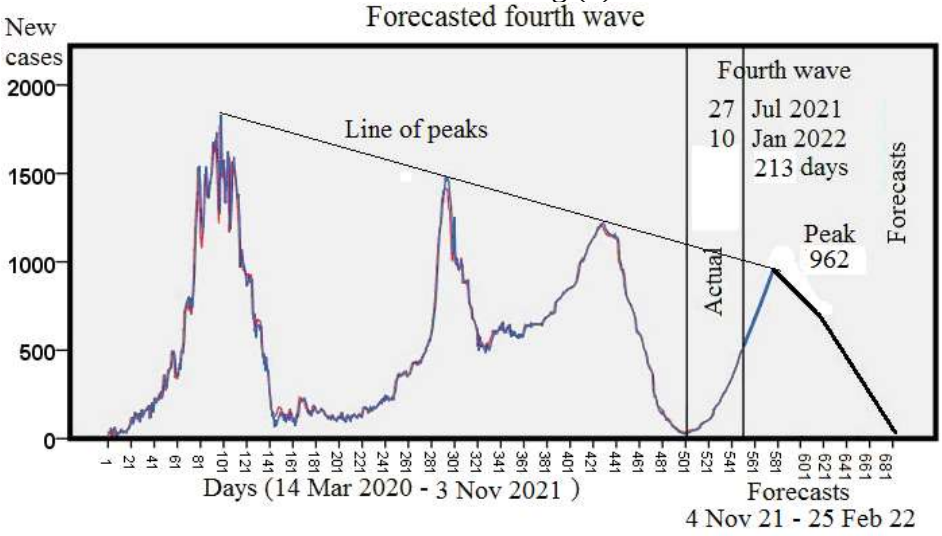
Length of the wave : 213 days

Peak : 962 new case

Expected date of the peak : 16 November 2021

The following Fig (9) shows the prospected fourth wave which is expected to end on 10th January 2022.

Fig (9)
Forecasted fourth wave



It is expected (hoped) that the new cases start declining after the peak of 16th November 2021
The question is: will corona epidemic end after the fourth wave, or there will be another fifth wave of new cases ?

4 - Deaths and recovered cases

Infections may recover, die or stay under treatment. This classification differed from wave to another as shown in the following table (7)

Table (7)
Distribution of total cases in the three waves
14 March 2020 - 3 November 2021

Wave	Total infected	Recovered	Deaths	Under care
First	97178	75367	5244	16567
%	100.00	77.55	5.40	17.05
Second	71869	61083	4360	6426
%	100.00	84.99	6.07	8.94
Third	114979	98092	6897	9990
%	100.00	85.31	6.00	8.69
Fourth*	49747	52479	2328	-5060
%	100.00	105.49	4.68	- 10.17
Total	333773	287021	18829	27923
%	100.00	85.99	5.64	8.37

* Only the period from 27 July 2021 to 3 November 2021

The first recovered cases were reported on 26 March 2020 i.e 12 days after the first reported new cases infected. These 12 days are considered the average period necessary for the treatment of infected

person to recover from covid-19. Hence, the new cases were lagged 12 days to cope with the recovered cases. Deaths occur any time from infection, the first deaths were reported on 15 March 2020, i.e one day after the first reported infected cases, so they are considered contemporaneous (almost) to the new infected cases.

It is clear from the above table that Egypt gained some experience dealing with corona since the percentage recovery increase from 77.55% through the first wave to 85.31% in the third wave. Cases under care (treatment) decreased from 17.05% to 8.69%. Generally speaking, one may consider the three case are the same.

As for the fourth wave started 17 July 2021, the data look confusing since the recovered cases are more than the new cases. However, the negative under care cases mean that people were infected through previous waves are recovered through this wave.

The data on the fourth wave cover the period from 27 July 2021 to 3 November 2021. Recovered cases reported up to 3rd of November 2021 are shifted 12 days above so they are considered up to 22 October 2021

5 - Normal Distribution Model

The spread of the infection in a populations is a random process. Each person has a probability of infection (*p*). It is a discrete distribution (Bernoully) of the form :

$$f(n;p) = p^n q^{1-n} \ , \ n = 0, 1 \qquad , p + q = 1 \qquad (2)$$

where:
p **is the probability of infection**
q **is the probability of not infected**

So, n = 1 at the day of infection and 0 on all other days.

If an area in the country of about 50 persons so we have daily infections k and non infections 50 - k, so we move from Bernoully distribution to a binomial distribution of the form:

$$\Pr(X=k)=\binom{n}{k} p^k\, q^{n-k} \quad , \; k=1,\ldots,n \qquad \textbf{(3)}$$

where n = 50 for this area.

If we consider whole country, where the population is over 70 Millions, i.e we may consider n → ∞.

Now consider *p* , the probability of a person is infected in a certain day, is calculated as the average daily infections (333773/600 = 556.29) divided over the population = 556.29/70,000,000 = 0.000 008 is very small

under the following conditions:

1 - n independent random variables

2 - n → ∞ .

3 - p is very small

The binomial distribution → normal distribution of the form

$$I(t)=I_o\, e^{-\left(\frac{t-\mu}{\sigma}\right)^2} \qquad \textbf{(1)}$$

***I*(*t*)**	**the number of daily infections**
t	**time**
I_0	**the maximum value at time μ**
σ	**the standard deviation of the normal distribution**

The model is characterized by three independent parameters:

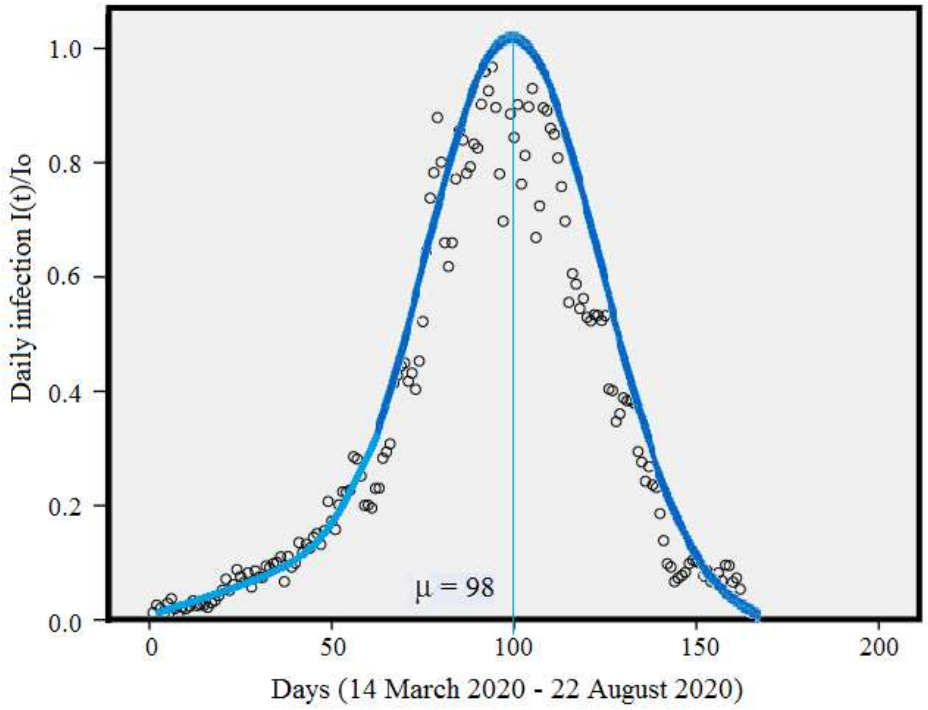
1 - the variance,

2 - the maximum height, and

3 - the time of maximum height (peak date)

Despite its simplicity, the Normal model has significant predictive power. We use it to predict the daily infection and when we reach the peak number of infections. In addition, a numerical study has explained that the Normal model is a special version of the SEIR model. The introduction of social distancing leads to a decrease in the infection rate. As we derived the Normal model, epidemics are initially binomial; later, as all distributions of the population are added, they result in the bell-shaped daily quantities. So it starts from minimum, reaching a maximum somewhere, then drops back to reach minimum consisting a single wave.

Fig (10)



Prediction of the time evolution $I(t)/I_0$ for Egypt for the best fit parameters μ = 98 days and σ = 26.9 days using SPSS20. The time t = 0 corresponds the date 14 March 2020.

By monitoring the new daily infections of the first wave one easily derive the relative change in the daily infections $R(t)$ where we used the distribution (1):

$$R(t)=\frac{dI(t)/dt}{I(t)}=\frac{d\ln I(t)}{dt}=2\frac{\mu-t}{\sigma^2} \qquad (2)$$

5.1 Total number of infections

According to CAPMAS (2021), during the first wave of the virus, population over the age of 15, (about 70 percent of the total population) are infected, if nothing is done to reduce the number of infections. Scaling the total population in units of 10^6N_6, (N_6 = 102 since the population of Egypt is 102×10^6) we estimate that $0.7q10^6N_6$ are infected during the duration of the first virus wave, where the quarantining factor q accounts for the effect of the currently taken different types of lockdowns to mitigate the effect of the virus

Integrating the Normal (1) over all times we then obtain

$$0.7\times10^6\ p\,N_6 = I_o \int_{-\infty}^{\infty} e^{-\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

put $\frac{t-\mu}{\sigma} = x$ **then**

$$= I_o\,\sigma \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\, I_o\,\sigma \qquad \textbf{(3)}$$

Since $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

Equation (3) yields for the maximum value

$$I_o = \frac{7\times10^5\ p\,N_6}{\sigma\sqrt{\pi}} = \frac{7\times10^5\ p\,N_6}{\sigma\times1.772} = \frac{3.95\times10^5\ p\,N_6}{\sigma} \qquad \textbf{(4)}$$

p **is the ratio of total infections to 70% of total population = 0.00094, and σ = 21.2 ,**
Substitute in (4) , I_o **(maximum infections) = 1786.4 compared to the actual maximum of 1774 case.**

6 - Conclusion

Three different statistical methods are used to analysis the behaviour of Covid-19. The main target is to shed some light on the near future of this epidemic, i.e its fourth wave.

Applying the above analysis, and gathering the results from the spectral analysis, ARIMA model and Normal model most important prospected aspects of the fourth wave may be summarized as follows:

1 - Length of the wave 167 days (spectral analysis), so it is expected to reach its minimum (or vanishes) on 10 January 2022

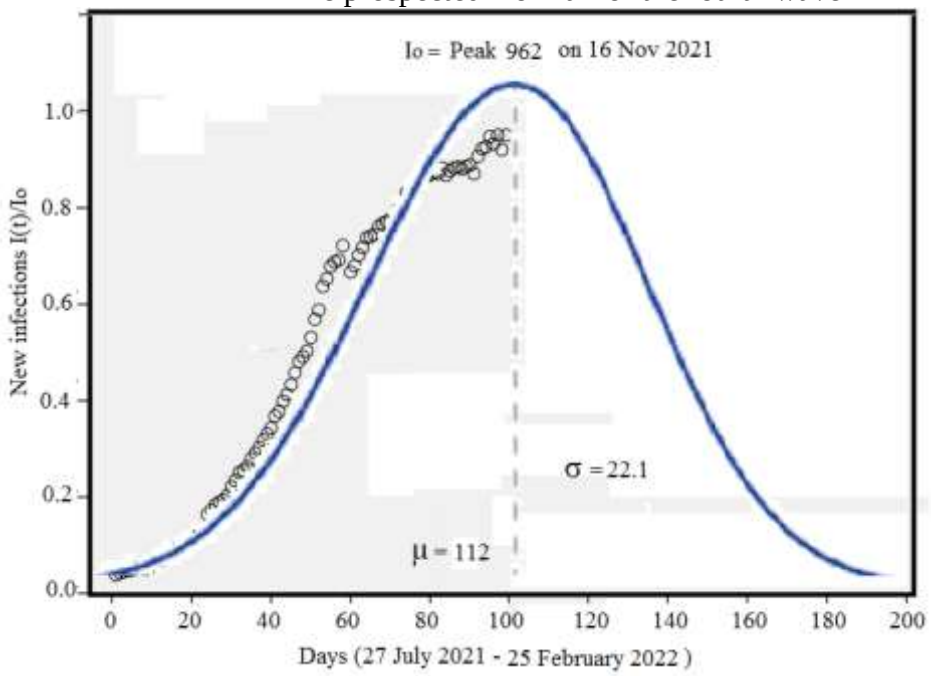
2 - It reaches 1007 new infections as its maximum on 16 November 2021 after 112 days from its beginning (table 6)

3 - Decline of infections after the maximum will be little faster compared to the increasing trend before the maximum. It took 112 days

to rise from 31 case to the prospected peak of 962 case, while it is expected to continue for 55 days to reach its minimum.

4 - The expected shape of the fourth wave could be:

Fig (11)
The prospected Normal for the fourth wave



Prediction of the time evolution $I(t)/I_0$ for Egypt for the best fit parameters μ = 112 days and σ = 21.1 days using SPSS20. The time $t = 0$ corresponds with the date 27 July 2021.

5 - Using equation (4) to predict infection in some selected days shown in the following table (8)

Table (8)
Forecasts of the fourth wave of new cases using Normal distribution

Date	Forecasts	Date	Forecasts	Date	Forecasts
7 Nov 21	950.33	29	566.17	1 Feb 22	215.13
16 Peak	962.56	31	548.25	4	188.25
22	920.25	1 Jan 22	539.78	8	152.78
30	891.35	7	467.12	12	116.12
1 Dec 21	882.22	14	395.68	16	80.68
7	782.16	21	323.23	20	44.23
15	710.26	28	251.16	24	8.16
22	638.36	31	224.43	25	0.43

References

Al-Hussein, A.B.A.; Tahir, F.R. Epidemiological characteristics of COVID–19 ongoing epidemic in Iraq. *Bull. World Health Organ. E-Pub* 2020, *6*. [Google Scholar] [CrossRef]

Almeshal, A.M.; Almazrouee, A.I.; Alenizi, M.R.; Alhajeri, S.N. Forecasting the Spread of COVID–19 in Kuwait Using Compartmental and Logistic Regression Models. *Appl. Sci.* 2020, *10*, 3402. [Google Scholar] [CrossRef]

Amar, L.A.; Taha, A.A.; Mohamed, M.Y. Prediction of the final size for COVID–19 epidemic using machine learning: A case study of Egypt. *Infect. Dis. Model.* 2020, *5*, 622–634. [Google Scholar] [CrossRef] [PubMed]

Anderson T. W.. 1971. The Statistical Analysis of Time Series. John Wiley & Son. Canada. p. 386.

Anwar, W.A.; AbdelHafez, A.M. Forecasting the peak of novel coronavirus disease in Egypt using current confirmed cases and deaths. *medRxiv* 2020. [Google Scholar] [CrossRef]

Bantan, R.A.R.; Chesneau, C.; Jamal, F.; Elgarhy, M. On the Analysis of New COVID–19 Cases in Pakistan Using an Exponentiated Version of the M Family of Distributions. *Mathematics* 2020, *8*, 953. [Google Scholar] [CrossRef]

Batista, M. Estimation of the final size of the second phase of coronavirus epidemic by the logistic model. *MedRxiv* 2020. [Google Scholar] [CrossRef]

Boldog, P.; Tekeli, T.; Vizi, Z.; Dénes, A.; Bartha, F.A.; Röst, G. Risk Assessment of Novel Coronavirus COVID–19 Outbreaks Outside China. *J. Clin. Med.* 2020, *9*, 571. [Google Scholar] [CrossRef]

Burkhard B.. 2000. Spectral Analysis and Filter Theory in Applied Geophysics. Springer-Verlag Berlin Heidelberg. p. 101.

Centers for Disease Control and Prevention (CDC). How COVID–19 Spreads. Available online: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html (accessed on 8 October 2020).

Centers for Disease Control and Prevention (CDC). Symptoms of Coronavirus. Available online: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html (accessed on 8 October 2020).

Central Agency of Public Mobilization And Statistics (CAPMAS), Egypt between the first wave and the second wave, Cairo, Feb 2021 (in Arabic).

El Desouky, E.D. Prediction of the Epidemic Peak of Covid19 in Egypt. *medRxiv* 2020. [Google Scholar] [CrossRef]

European Centre for Disease Prevention and Control. Transmission of COVID–19. Available online: https://www.ecdc.europa.eu/en/COVID--19/latest-evidence/transmission (accessed on 8 October 2020).

Gorbalenya, A.E.; Baker, S.C.; Baric, R.S.; de Groot Raoul, J.; Drosten, C.; Gulyaeva, A.A.; Haagmans, B.L.; Lauber, C.; Leontovich, A.M.; Neuman, B.W. The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* 2020, *5*, 536–544. [Google Scholar]

Hasab, A.A.; El-Ghitany, E.M.; Ahmed, N.N. Situational Analysis and Epidemic Modeling of COVID–19 in Egypt. *J. High Inst. Public Health* 2020, *50*, 46–51. [Google Scholar] [CrossRef]

Helmy, Samar 2020 The impact of lockdowns to mitigate the effects of corona on the Egyptian GDP through the period April-June 2020 (in Arabic). Mansoura University, Faculty of Commerce, the Egyptian Journal for Commercial Studies, Sept 2020

Kuniya, T. Prediction of the epidemic peak of coronavirus disease in Japan. *J. Clin. Med.* 2020, *9*, 789. [Google Scholar] [CrossRef]

Röst, G.; Bartha, F.A.; Bogya, N.; Boldog, P.; Dénes, A.; Ferenci, T.; Horváth, K.J.; Juhász, A.; Nagy, C.; Tekeli, T.; et al. Early Phase of the COVID–19 Outbreak in Hungary and Post-lock-down Scenarios. *Viruses* 2020, *12*, 708. [Google Scholar] [CrossRef]

Schlickeiser, R.; Schlickeiser, F. A Normal Model for the Time Development of the Sars-Cov-2 Corona Pandemic Disease. Predictions for Germany Made on 30 March 2020. *Physics* 2020, *2*, 164–170. [Google Scholar] [CrossRef]

Schüttler, J.; Schlickeiser, R.; Schlickeiser, F.; KrSchöttlerger, M. COVID–19 Predictions Using a Gauss Model, Based on Data from April 2. *Physics* 2020, *2*, 197–212. [Google Scholar] [CrossRef]

Tosepu, R.; Gunawan, J.; Effendy, D.S.; Lestari, H.; Bahar, H.; Asfian, P. Correlation between weather and Covid–19 pandemic in Jakarta, Indonesia. *Sci. Total Environ.* 2020, *725*, 138436. [Google Scholar] [CrossRef] [PubMed]

World Health Organization (WHO). Coronavirus. Available online: https://www.who.int/health-topics/coronavirus#tab=tab_1 (accessed on 8 October 2020).

Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID–19) outbreak in China: Summary of a report of 72,314 cases from the Chinese center for disease control and prevention. *J. Am. Med. Assoc.* 2020, *323*, 1239.e42. [Google Scholar] [CrossRef] [PubMed]