Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT

KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences

School for Information Technology

Master of Statistics and Data Science

Master's thesis

Short term prediction of COVID-19 outbreak in Ireland: estimation and model selection

Haymanot Zeleke Tadesse

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Biostatistics

SUPERVISOR :

Prof. dr. Ziv SHKEDY

2020
2021

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences
School for Information Technology
Master of Statistics and Data Science

Master's thesis

Short term prediction of COVID-19 outbreak in Ireland: estimation and model selection

Haymanot Zeleke Tadesse
Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Biostatistics

SUPERVISOR :
Prof. dr. Ziv SHKEDY

Acknowledgements

First and foremost, I am thankful to Almighty God and the Blessed Virgin Mary for the blessings and strengths of every day of my life, not just during my studies.

Secondly, I am extremely grateful to my supervisors, Prof. dr. Ziv Shkedy for his invaluable advice, continuous support, and patience during my study. I benefited greatly from your valuable comments, feedbacks, discussions we had and your immediate reply to my emails highly stimulated me throughout this study.

Thirdly, I would like to thank all professors who taught and sharing knowledge at Hasselt University. I would also like to extend my gratitude to VLIR-UOS for providing me financial support to undertake two-year master program in Belgium.

Finally but most importantly, I would like to thank my parents Zeleke Tadesse, Alemitu Wubie, my brothers and sisters; for their constant love, support and encouragement. Without them none of this would have been possible. I cannot begin to thank them enough. You are always there for me. I would also like to extend my gratitude to best friends and all my classmates.

Haymanot Zeleke Tadesse
June 16, 2021
Hasselt University
Belgium

Abstract

Introduction: The COVID-19 pandemic is now a major global health threat. COVID-19 is an infectious disease caused by a newly discovered coronavirus. This newly identified coronavirus, SARS-CoV-2, has caused a worldwide pandemic of respiratory illness. The World Health Organization China country office reported a cluster of pneumonia cases in Wuhan, Hubei Province of China on 31 December 2019. The first confirmed case of the coronavirus in Ireland was reported by the National Public Health Emergency Team on 29 February 2020 and up to now Ireland have seen a three wave pattern in reported cases of COVID-19.

Objectives: The main objectives of this study were to make short term prediction of COVID-19 cases by using the data on the period of 4 March 2020 to 30 June 2020 and to investigation of model selection procedures.

Methodology: The cumulative number of reported COVID-19 cases on the period 04 March to 30 June 2020 were analyzed. Nonlinear growth curves models to the daily cumulative COVID-19 cases were implemented in two basic distribution assumption, namely Poisson and normal. And to achieve the objective of this study AIC and chi-squared were used to select the best model on the estimation period and prediction period respectively.

Results: The non linear models were fit to make a short term prediction. The five parameter logistic model was found to be the best fit model compared to others proposed models in the estimation period during the first wave under the Poisson assumption and Richards was found to be the best fit model in all eleven different estimation period under normal assumption . Nevertheless, Richards and Gompertz models have the best performance to make short term prediction relative to others candidates of non linear growth models for Poisson and normal respectively. In conclusion, the results suggested that to to select the best model according to the performance of prediction to use chi-squared value instead of AIC. Moreover, the non linear models was very important to describe and to make a short term prediction of the daily cumulative cases.

Key Words: *COVID-19, Non linear model, Short term prediction, AIC, Chi-squared, Richards, Gompertz*

Contents

1 Introduction

1.1 Background of the study

The Coronavirus disease (COVID-19) pandemic is now a serious global health threat. COVID-19 is an infectious disease caused by a newly discovered coronavirus. Most of the people infected with the COVID-19 virus will experience mild to moderate respiratory disease and recover without requiring special treatment. Older people, and people with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory illness, and cancer are more likely to develop serious illness.The best way to prevent and slow down transmission is to be informed about the COVID-19 virus, the disease it causes and the way it spreads. Protect yourself and others from infection by washing your hands or using an alcohol based rub frequently and not touching your face. The COVID-19 virus spreads primarily through droplets of saliva or discharge from the nose when an infected person coughs or sneezes, so it's important that you also practice respiratory etiquette (for example, by coughing into a flexed elbow)[16].

A newly identified coronavirus, SARS-CoV-2, has caused a worldwide pandemic of respiratory disease, called COVID-19. The World Health Organization (WHO) China country office reported a cluster of pneumonia cases in Wuhan, Hubei Province of China on 31 December 2019. Case in China are now declining but, there is a rapidly increase in another country. The disease has spread all over the world is alarming and the therefore case number continue to rise. It is almost hard to believe that some countries still haven't any reported cases of the COVID-19. As of MAY 30^{th} 2021, only 7 countries haven't confirmed any cases of COVID-19. Many of these countries without COVID-19 are Pacific Island countries in Oceania [20].

The first cases of 2019-nCoV are reported with in the European Region on 24 January 2020, France has officially notified to the WHO Regional Office for Europe of three confirmed cases of 2019-nCoV. Two patients were detected in Paris and one in Bordeaux. All 3 had travelled from Wuhan, China and are now hospitalized in France[17].The first confirmed case of the coronavirus in Ireland was reported by the National Public Health Emergency Team on 29 February 2020 and the first death by Coronavirus on 11^{th} March. The case is related to travel from an affected area in northern Italy, instead of contact with another confirmed case in Ireland [3]. Within three weeks, cases had been confirmed in all counties. The government shut all schools, colleges, childcare facilities and cultural institutions on 12 March 2020 [10]. On march 27 Ireland will be place a restriction like: non essential shops closed, stay at home except essential worker, banned all non essential travel and contact with other people [5], cancelling all large gathering and all business were shut on march 24 2020.

On the middle of April the spread of COVID-19 has reached as low as it needs to be in Ireland. As the National Public Health Emergency Team reported that the growth rate is close to zero, which is the spread of the virus has been stopped effectively by the restriction[4]. New daily cases and deaths were decreased by June and the restriction were relaxed slowly lifted, whereas for summer break the schools remained closed.

Ireland have seen a three wave pattern in reported cases of COVID-19 up to now. Timeline of the first, second and third waves of of COVID-19 infection in Ireland is on 01 March to 01 August 2020, 02 August to 21 November2020 and 22 November 2020 onwards respectively . In Ireland, from 3 January 2020 to 26 May 2021, there have been 258,968 confirmed cases of COVID-19 with 4,941 deaths, reported to WHO. As of 23 May 2021, a total of 2,349,207 vaccine doses have been administered [18]. For this study the first wave of COVID-19 in Ireland was used from late February up to end of July.

The non linear epidemiological models studies aim to to generate and assess short-term forecasts of the cumulative number of confirmed reported cases. These non linear epidemiological models have previously been used to model other disease outbreaks for instance Zika virus [15], Dengue [9], and human mortality [6]. Recently, the COVID-19 pandemic fitted the generalized logistic model, Richards model and a sub-epidemic model to the cumulative COVID-19 cases in the Hubei province of China and the rest of China and made a short-term forecast of 5, 10 and 15 days ahead for five consecutive days [14]. In recently Reddy and colleagues fitting several nonlinear growth curves (Richards, 3 and 4 parameter logistic, Weibull and Gompertz), to made short term forecasts of 5, 10, 15, 20, 25 and 30-day ahead for COVID-19 cases and deaths at the national level and also as the provincial level in South Africa [12].

In this study non linear epidemiological models are used to produce short term(ten consecutive days) predictions of the total number of reported cases in Ireland. In addition to this short term prediction we studied one, two ,three, four, five and ten days. The short term prediction produced from such models can be useful to guide the allocation of resources that are critical to bring the epidemic under control and to to make a decision for policy maker.

The paper is structured as follows: in section 2 outlines the study data. In section 3, we described the statistical methodology of non linear growth model. In this section detailed about several nonlinear growth curves, model selection and short term predication. Then, section 4 presents the main findings. Finally, Section 5, discusses the main findings and draw conclusion.

1.2 Objectives

The main objectives of this study were to make short term prediction of COVID-19 cases by using the data on the period of 4 March 2020 to 30 June 2020 and to investigation of model selection procedures.

2 Data

2.1 Publicly Available COVID-19 data

The source of data for this study is publicly available daily reported cumulative COVID-19 case [8]. The daily updated data are currently stored at `https://covid19datahub.io/`. The data was read by using R software and drive the daily new COVID-19 case. As mentioned in the previous part for this study the first wave in Ireland were used, from 04 March 2020 to 30 June 2020, there have been 25,473 confirmed cases of COVID-19 with 1,736 deaths, reported to WHO. As of 31 May 2021, the Department of Health had confirmed 261,686 cases and 4,949 deaths. More than 90% of those who have died were aged over 65[11].

2.2 Incidence data and cumulative cases

The publicly available database contains the total daily COVID-19 cases which is updated each days and other variables. A total daily record of COVID-19 cases were extracted and the new daily cases (incidence) calculated from the total daily cases by: Y(t)-Y(t-1) where Y(t) is Cumulative number of cases and t is the time point.

The data which is extracted from the database and derived from the given data set during the period 04 March 2020 to 30 June 2020 is presented in Table 1.

Table 1: The sample of data set which is used in this study.

date	confirmed(y(t))	administrative_area_level_1	confirmed_daily	time(t)
04/03/2020	6	Ireland	6	1
05/03/2020	13	Ireland	7	2
06/03/2020	18	Ireland	5	3
07/03/2020	19	Ireland	1	4
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
29/06/2020	25462	Ireland	23	118
30/06/2020	25473	Ireland	11	119

The daily new cases of COVID-19 in Ireland for the first wave of the pandemic and the cumulative case of COVID-19 are displayed in Figure 1. The spread of COVID-19 was high over the period second week of March to the middle of April. The highest number(peak point) of daily new COVID-19 cases was reported on 2020-04-17 and then spread of COVID-19 starts to decrease.

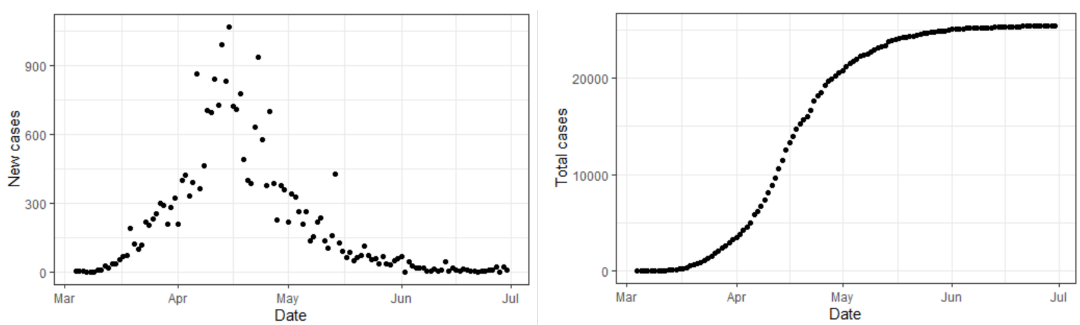

Figure 1: The scatter plot for the daily confirmed COVID-19 cases (left) and cumulative COVID-19 cases(right) for the period 04 March 2020 to 30 June 2020 (first wave)

3 Methodology

3.1 Non-linear growth model

To achieve the objectives of this study, a set of several nonlinear growth curves fitted for the daily cumulative number of reported COVID-19 cases and produce short term prediction of COVID-19 cases which has an important role in shaping public policy and decision-making. The assumption of a Poisson and normal distribution were used for the daily cumulative number of reported COVID-19cases.

The fitted daily new number of COVID-19 cases was obtained from the fitted daily total number by taking the first derivative of the mean ($\frac{d\mu(t)}{dt}$). By consider both Poisson distribution (Y(t)$\sim$ $Poisson(\mu(t))$) and normal distribution ($Y(t) \sim N(\mu(t), \sigma^2)$) where Y(t) is Cumulative number of COVID-19 cases, $\mu(t) = E(Y(t))$ and σ^2 is the variance of Y(t), the Richards, three parameter logistic, five parameter logistic and Gompertz model were fitted to the reported daily total number of cases. For all models α is the final size of the epidemic, γ is the per capita intrinsic growth rate of the infected population, k is the exponent of the deviation from the standard logistic curve η is the turning point and α_0 the initial size. All the analysis was performed by using SAS software for fitting nonlinear growth curves and to make a short term prediction and R software for reading the data and to make some plots and Chi-squared calculation software.

3.1.1 The Richards model

The Richards model has often been used to model reported cumulative cases in disease outbreaks and widely used growth model. The basic premise of the Richards model is that the daily incidence curve consists of a single peak of high incidence, resulting in an S-shaped epidemic curve and a single turning point of the outbreak. These turning points, defined as times at which the rate of accumulation changes from increasing to decreasing or vice versa, can be easily located by finding the inflection point of the epidemic curve, the moment at which the trajectory begins to decline [19]. It is useful to capture the temporal variations of an outbreak, in particular the turning points (or peaks and valleys of the incidence curve) [15, 14]. The mean structure of the model and incidence respectively are given by:

$$\mu(t) = \alpha(1 + ke^{-r(t-\eta)})^{-1/k},$$
$$\frac{d\mu(t)}{dt} = \frac{\gamma}{k}[(\frac{\mu(t)}{\alpha})^k - 1].$$

The plot in Figure 2 shows an example of the incidence and mean structure of the cumulative daily cases of the Richards model for Nigeria outbreak during March 1-October 31, 2020

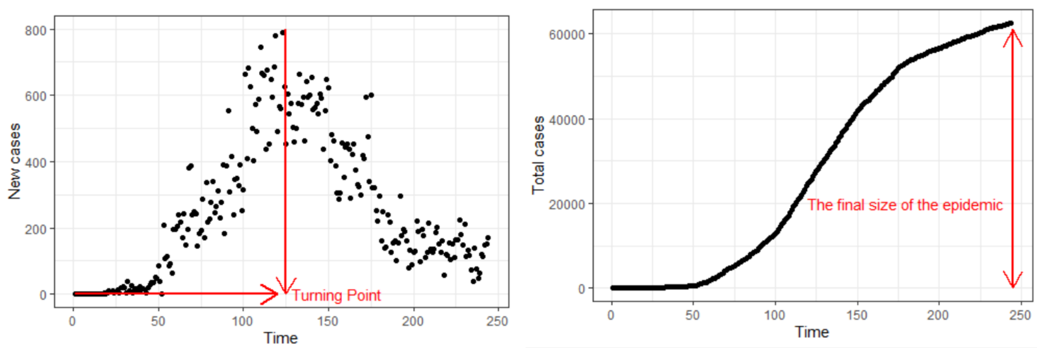

Figure 2: The daily COVID-19 cases incidence curve(left) and the theoretical epidemic curve using Richards model (right) for Nigeria outbreak during March 1-October 31, 2020.

3.1.2 The Three Parameter Logistic model

This model is the special case of the Richards model, obtained when the exponent parameter k = 1. The growth curve is symmetric around turning point and has equal periods of slow and fast growth. The mean structure of the model and incidence respectively are given by:

$$\mu(t) = \frac{\alpha}{1 + e^{-r(t-\eta)})},$$
$$\frac{d\mu(t)}{dt} = \gamma\mu(t)\left[\frac{\mu(t)}{\alpha} - 1\right].$$

3.1.3 The Five parameter Logistic model

The five-parameter logistic model, which includes a fifth parameter, permits asymmetry to be effectively modeled.This model has asymmetric factor k that makes it to be asymmetric.The 5PL model is able to eliminate most of the lack of-fit error present in fitted 4PL models [7]. The mean structure of the model and incidence respectively are given by:

$$\mu(t) = \alpha_0 + \frac{\alpha - \alpha_0}{[1 + (2^{(1/k)} - 1)(\frac{t}{\eta})^{\gamma}]^k},$$
$$\frac{d\mu(t)}{dt} = \frac{k\gamma}{t}[(u(t) - \alpha_0]\left[[\frac{\mu(t) - \alpha_0}{\alpha - \alpha_0}]^{1/k} - 1\right].$$

3.1.4 The Gompertz model

The Gompertz model is one of the most frequently used sigmoid models fitted to growth data and other data [6]. This function is especially useful in describing the rapid growth of a certain population of organisms while also being able to account for the eventual horizontal asymptote. This model is a very flexible model that the mean structure of the model and incidence is given by:

$$\mu(t) = \alpha_0 + (\alpha - \alpha_0)e^{-e-\gamma(t-\eta)},$$
$$\frac{d\mu(t)}{dt} = \gamma[\mu(t) - \alpha_0]ln\left[\frac{\mu(t)-\alpha_0}{\alpha-\alpha_0}\right].$$

3.2 Short term prediction

The main interest of this paper is to use the available data and make a short term prediction to the total number of COVID-19 cases for the ten days ahead. By using four different models and two different assumption of the daily cumulative cases that have been previously used to derive a short term prediction. In this study we used different estimation period(t) which is $t = 70, 71, 72, \ldots, 80$ and 119 and make a ten consecutive days a head prediction of the total number of cases for each estimation period. A ten days a head (t up to t + 10) period is the prediction period.

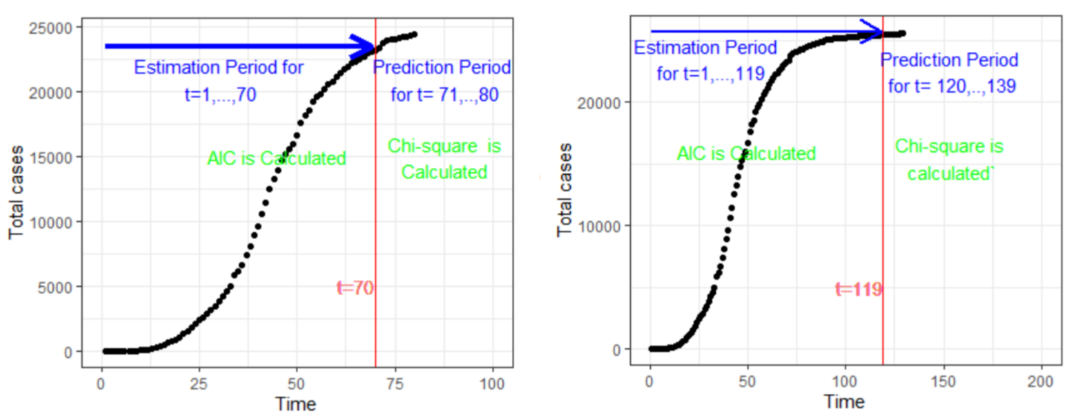

Figure 3: The estimation and prediction period of cumulative COVID-19 cases during March 4- May 12 2020(left) and during March 4-June 30 2020 (right).

Figure 3 shows that the estimation period and predication period of 70 days and for the first wave of COVID-19 in Ireland. The red line indicates the number of days which is used to make an estimation of the model. On both plot, the number of days before red line used for estimation of model and , the number of days after red line , indicates the prediction period. On the estimation period, the AIC value were calculated where as on the prediction period the Pearson chi-squared were calculated which is used for model selection. Furthermore, the estimation of different models with respect to different estimation period and a ten days consecutive ahead prediction will be presented in section 4.

Another way of make a predication in this study is to use different estimation period ($t = 70, 71, 72, \ldots, 80$), which has different non linear growth model for each distribution and then make one days (t+1), two days (t+2), three days (t+3), four days (t+4), five days (t+5) and ten days (t+10) ahead prediction. We predicted each of the eleven different models and collocated a one, two, three, four, five and ten days ahead from each model, respectively.

Table 2 illustrates the prediction of cumulative COVID-19 cases ahead of the selected number of days. First, estimated each of the eleven different models then, made a ten consecutive days ahead prediction from each models finally, collocated a one, two, three, four, five and ten days ahead prediction from each model, respectively.

Table 2: Illustration of one days, two days,three days, four days, five days and ten days ahead prediction.

	Prediction Day					
Estimation period (1-t)	One Days ahead (t+1)	Two Days ahead (t+2)	Three Days ahead (t+3)	Four Days ahead (t+4)	Five Days ahead (t+5)	Ten Days ahead (t+10)
1-70	71	72	73	74	75	80
1_71	72	73	74	75	76	81
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
1-80	81	82	83	84	85	90

3.3 Model selection

Model selection plays a fundamental role in choosing a best model from a series of candidate models for data driven modelling and system identification problems. The goodness of fit of a statistical model describes how well it fits a set of observations. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question. To select the model which has the best goodness fit to the data Akaike Information criteria(AIC) for the estimation period and the Pearson chi-squared value for predilection period were used.

3.3.1 Akaike Information criteria

Among various model selection methods, Akaike information criterion (AIC) is the most popular measures. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, the model with a minimum value is then treated as an optimal choice for the estimation period. AIC can be calculated as [2]:

$$AIC(k) = -2nln(L) + 2k$$

where k is the number of fitted parameters in the model, L is the maximum likelihood estimate for the model and N is the sample size.

3.3.2 Pearson Chi-squared

Pearson chi-squared uses a measure of goodness of fit which is the sum of differences between observed and expected outcome. In this study the Pearson Chi-squared were calculated for the ten days prediction period for each model corresponding to the two assumed distribution. And also Chi-squared were calculated for one days, two days,three days, four days, five days and ten days ahead prediction to select the best model. The smallest the chi-square is provided the best model according to the performance of short term prediction period. Here the Pearson chi-squared value calculated for ten predicted days with the corresponding observed cumulative daily cases. Pearson Chi-squared be calculated as [1]

$$\chi^2 = \sum_{t=t+1}^{t+1o} \frac{(Y(t)-\hat{Y}(t))^2}{\hat{Y}(t)} \ .$$

where t is the estimation period, $Y(t)$ is the observed data, and $\hat{Y}(t)$ is the corresponding predicted value from the given model.

4 Results

4.1 Estimation Periods

The incidences and cumulative cases of COVID-19 are displayed in Figure 4. The read lines in both plot of COVID-19 cases at the days of 70,72,74, . . . , 80 and 119 indicates all the the estimation period which is used in this study to make estimation, model selection and short term prediction.

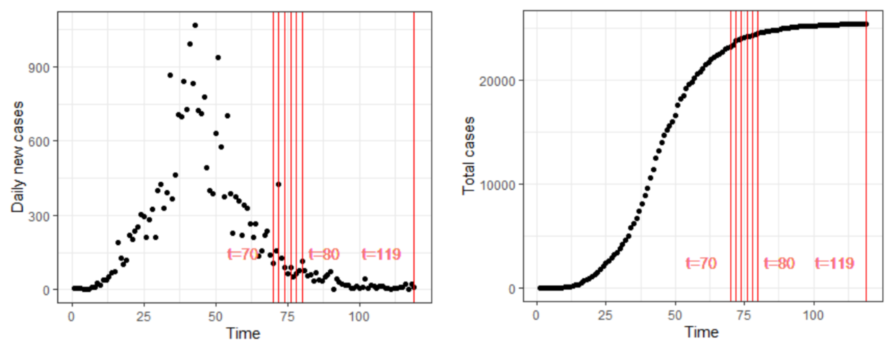

Figure 4: The daily new cases (left) and cumulative confirmed (right) COVID-19 outbreak cases for Ireland during March 4-June 30, 2020.

4.2 Analysis of Non-linear growth model for the assumption of Poisson and Normal distribution

In Table 3, we can see that the AIC value of all the four candidate models were calculated across the assumption of Poisson and normal distribution for the first wave in Ireland. The AIC value suggests a choice of the five parameter logistic, which has the smallest AIC value for both Poisson assumption and normal assumption. On the estimation period the five parameter logistic model is the best fit model compared to others proposed models.

The prediction performances can be affected by the uncertainty brought by the noise, so to select the best model according to prediction performance the χ^2 were calculated for the ten days prediction and select the one that has smallest value. The χ^2 values for the prediction period of all the four candidate models were calculated across the assumption of Poisson and normal distribution for the first wave in Ireland were displayed in Table 3. In the first wave of COVID-19 cases in Ireland Richards model has the smallest Chi-square value relative to others in the Poisson distribution whereas, Gompertz has the smallest Chi-square value relative to others in the normal distribution. so, Richards and Gompertz models have the best performance to make a short term prediction relative to others candidates of non linear growth models.

Table 3: AIC value of four different models and Chi-squared value for the Poisson and Normal distribution in the first wave of Ireland during estimation period & prediction period respectively.

	Estimation Period		Prediction Period	
Model	AIC(Poisson)	AIC(Normal)	chisq(Poisson)	chisq(Normal)
Richards	2440.3	1736.9	**6.482**	19.684
Three parameter logistic	5462.1	2403.7	231.714	82.5808
Five Parameter logistic	**2250.3**	**1721.9**	37.79	2.89513
Gompertz	4683.9	2068.1	88.7678	**1.133**

The estimated parameter value of the selected model which has smallest AIC five parameter logistic model on both Poisson and normal assumption are shown in Table 4 and 5 respectively. The estimated turning point and final size of COVID-19 were 43.9968 and 25224 respectively. In this study the long term parameter estimates are less of interest because the main objectives were to make a short term prediction and investigate on the model selection criteria.

Table 4: Parameter estimates of five parameter logistic model for Poisson assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr $>\|t\|$	95% Confidence Limits	
alpha	25224	30.763	119	819.95	<.0001	25163	25285
alpha0	15.7219	1.7295	119	9.09	<.0001	12.2972	19.1465
k	3.51	0.196	119	17.9	<.0001	3.1218	3.8982
gamma	3.7369	0.01635	119	228.61	<.0001	3.7045	3.7692
eta	43.9968	0.04127	119	1066.18	<.0001	43.9151	44.0785

The observed and predicted incidence, and the total daily reported and predicted number of COVID-19 cases from observed and 5PL during March 4-June 30 2020 under Poisson are displayed in Figure 5. From this graph we can observe that, the estimated tuning point and the estimated final size of COVID-19.

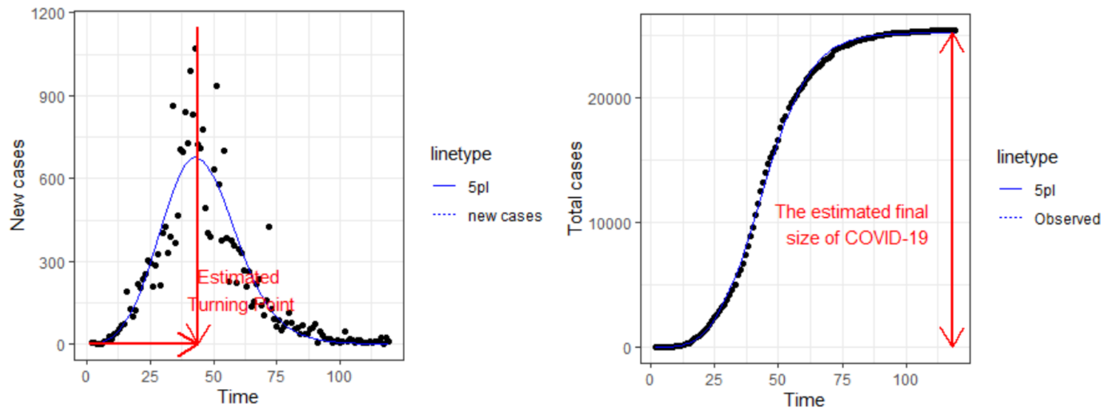

Figure 5: The incidence and predicted incidence(left) and the predicted and daily total reported (right) of COVID-19 cases from observed and 5PL during March 4-June 30 2020.

Table 5: Parameter estimates of five parameter logistic model for Normal assumption during March 4-June 30 2020.

Parameter	**Estimate**	**Std.error**	**DF**	**t Value**	**Pr** $>\|t\|$	**95% Confidence Limits**	
alpha	25522	32.9275	119	775.08	<.0001	25456	25587
k	1.3343	0.05581	119	23.91	<.0001	1.2238	1.4448
alpha0	221.47	28.9167	119	7.66	<.0001	164.22	278.73
gamma	4.5267	0.05526	119	81.91	<.0001	4.4172	4.6361
eta	43.9415	0.04696	119	935.69	<.0001	43.8485	44.0345
se	14030	816.34	119	17.19	<.0001	12414	15646

The incidence and predicted incidence and the predicted and reported daily total number of COVID-19 cases from observed, Richards, 3PL, 5PL and Gompertz during March 4-June 30 2020 for Poisson (the left two) and for normal (the right two) are displayed in Figure 6. From this graph we can observe that, even though the three parameter logistic model does not fit the data well than the others, they tried to handle the peak of the incidence for both Poisson and normal distribution assumption.

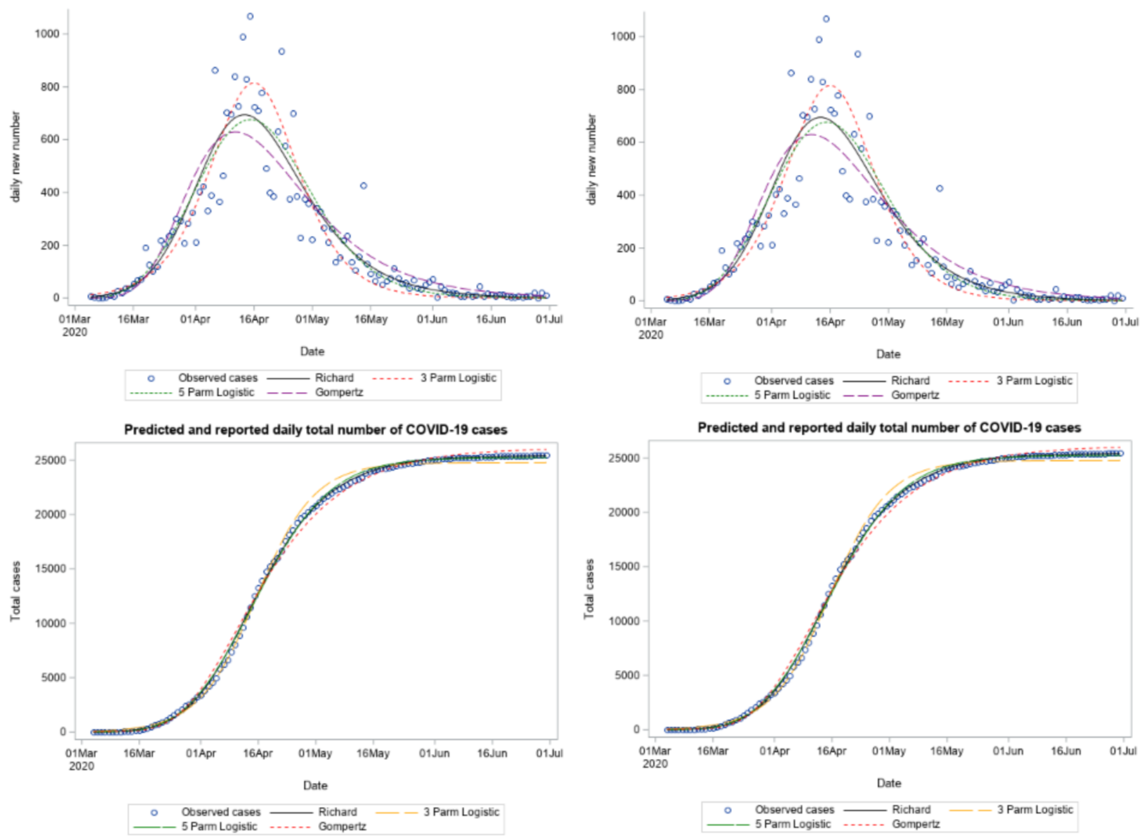

Figure 6: The incidence and predicted incidence(top two) and the predicted and daily total reported (bottom two) of COVID-19 cases from observed, Richards, 3PL, 5PL and Gompertz during March 4-June 30 2020.

In Figure 7 and 9 the fitted and reported daily total number of cases are shown. We can see that, the lines for the estimated model and reported cases follow each other closely, which is all the models fit the data pretty well on both distribution assumption. The fitted daily new number of cases was obtained from the fitted daily total number by taking the first derivative of the mean, shown in Figure 8 and 10 with the reported daily new number of cases. In this graph we can see that a good fit between the reported and the fitted and the three parameter logistic model tried to handle the peak of the incidence relative to others model on both distribution assumption.

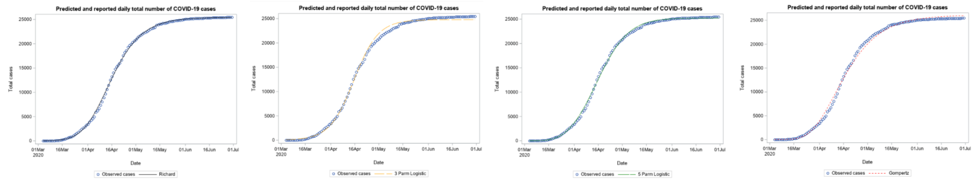

Figure 7: Predicted and reported daily total number of COVID-19 cases for the first wave under Poisson assumption during March 4-June 30 2020.

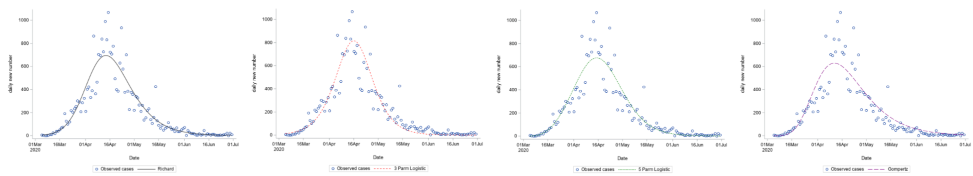

Figure 8: The fitted and reported daily new number of COVID-19 cases in the first wave under Poisson assumption during March 4-June 30 2020.

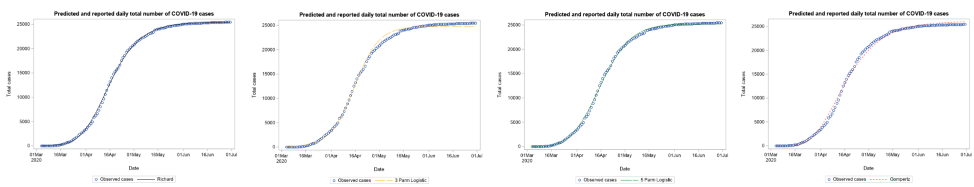

Figure 9: Predicted and reported daily total number of COVID-19 cases for the first wave under normal assumption during March 4-June 30 2020.

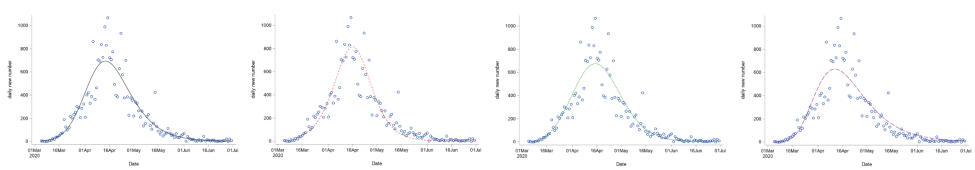

Figure 10: The fitted and reported daily new number of COVID-19 cases in the first wave under normal assumption during March 4-June 30 2020.

4.3 Short term prediction of the total number of reported COVID-19 cases

The ten consecutive days prediction of the total number of reported COVID-19 cases, under Poisson and normal assumption for Richards, 3PL, 5PL and Gompertz model respectively are presented in Table 6 and 7. The predicted total COVID-19 cases on 01 July were 25394 (95% C.I 25341,25446), 24769 (95% C.I 24726, 24812), 25217 (95% C.I 25341,25446) and 25976 (95% C.I 25341, 25446) for Richards, 3PL, 5PL and Gompertz model respectively under the Poisson assumption. The predicted total COVID-19 cases on 01 July were 25300 (95% C.I 25258, 25341), 25071 (95% C.I 25035, 25107), 25426 (95% C.I 25376, 25476) and 25467 (95% C.I 25427, 25507) for Richards, 3PL, 5PL and Gompertz model respectively under the normal assumption. Under the normal assumption the prediction based on Gompertz model were more close to the observed total COVID-19 cases than others model. However, on the assumption of Poisson Richards model were more close to the observed total COVID-19 cases than others model.

Table 6: A short term predictions of the total number of reported COVID-19 cases under Richards, 3PL, 5PL and Gompertz on the Poisson distribution during the estimation period March 4-June 30 2020.

		Richards model			3 Parameter Logistic			5 Parameter Logistic			Gompertz		
Date	Observed	Prediction	Prediction Interval		Prediction	Prediction Interval		Prediction	Prediction Interval		Prediction	Prediction Interval	
01/07/2020	25477	25394	25341	25446	24769	24726	24812	25217	25341	25446	25976	25341	25446
02/07/2020	25489	25396	25343	25449	24769	24726	24812	25218	25343	25449	25984	25343	25449
03/07/2020	25498	25398	25345	25451	24769	24726	24812	25218	25345	25451	25992	25345	25451
04/07/2020	25509	25399	25346	25453	24769	24726	24813	25219	25346	25453	25999	25346	25453
05/07/2020	25527	25401	25348	25454	24770	24726	24813	25219	25348	25454	26005	25348	25454
06/07/2020	25531	25403	25349	25456	24770	24726	24813	25220	25349	25456	26011	25349	25456
07/07/2020	25538	25404	25351	25457	24770	24727	24813	25220	25351	25457	26017	25351	25457
08/07/2020	25542	25405	25352	25459	24770	24727	24813	25221	25352	25459	26023	25352	25459
09/07/2020	25565	25406	25353	25460	24770	24727	24813	25221	25353	25460	26028	25353	25460
10/07/2020	25589	25408	25354	25461	24770	24727	24813	25221	25354	25461	26032	25354	25461

The predictive accuracy of the models to predict the total COVID-19 cases beyond the estimation period are displayed graphically in Figure 11 over a Poisson and normal assumption, for 10 days ahead period prediction form all models. In both graphs after the red vertical line it is prediction period and before it is the estimation period. The 3P and 5P models underestimate and Gompertz overestimate the cumulative cases for Poisson assumption. We observe that the Richards model yields substantially best prediction for ten days ahead than the others in case of Poisson assumption. The 3P, 5P and Richards models underestimate while, Gompertz yields best prediction for ten days ahead of the total COVID-19 cases under normal assumption. For the estimation period 70,71,. . . , 80, predict the total COVID-19 cases beyond the estimation period are displayed in Figure 14 & 15 in the Appendix B under Poisson and normal assumption respectively.

Table 7: A short term predictions of the total number of reported COVID-19 cases under Richards, 3PL, 5PL and Gompertz on the normal distribution during the estimation period March 4-June 30 2020.

		Richards model			3 Parameter Logistic			5 Parameter Logistic			Gompertz		
Date	Observed	Prediction	Prediction Interval		Prediction	Prediction Interval		Prediction	Prediction Interval		Prediction	Prediction Interval	
01/07/2020	25477	25300	25258	25341	25071	25035	25107	25426	25376	25476	25467	25427	25507
02/07/2020	25489	25301	25259	25343	25072	25036	25107	25430	25380	25481	25471	25431	25511
03/07/2020	25498	25303	25261	25345	25072	25036	25108	25435	25384	25486	25474	25433	25514
04/07/2020	25509	25304	25262	25346	25072	25036	25108	25439	25387	25490	25476	25436	25517
05/07/2020	25527	25305	25263	25347	25073	25037	25109	25443	25391	25495	25479	25439	25519
06/07/2020	25531	25306	25264	25349	25073	25037	25109	25446	25394	25499	25481	25441	25522
07/07/2020	25538	25307	25265	25350	25073	25037	25109	25450	25397	25503	25484	25443	25524
08/07/2020	25542	25308	25266	25351	25073	25037	25109	25453	25400	25507	25486	25445	25526
09/07/2020	25565	25309	25267	25352	25073	25038	25109	25456	25403	25510	25488	25447	25528
10/07/2020	25589	25310	25267	25352	25074	25038	25110	25459	25405	25513	25489	25449	25530

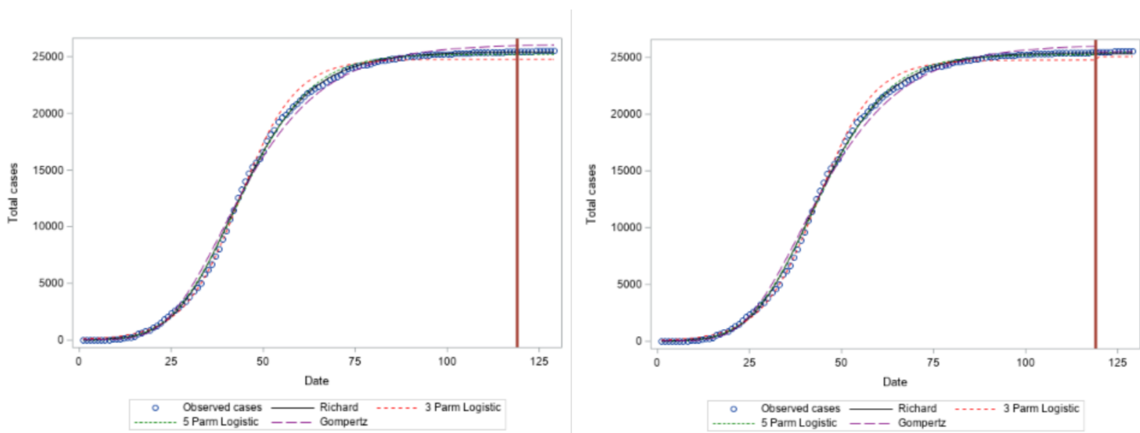

Figure 11: Predicted daily total number of COVID-19 cases from Richards, 3PL, 5PL and Gompertz and observed under the Poisson (left) and normal (right).

4.4 Goodness of fit for different estimation period and prediction Period

4.4.1 Goodness of fit on all eleven estimation period

The AIC and chi-squared for all four models on eleven different estimation period (70, 71, ..., 80 days) were displayed under Poisson and normal assumption are presented in Table 8. According to the AIC, the 5P logistic model had the smallest value under all eleven estimation period relative to others candidate model under Poisson distribution assumption. However, under normal distribution assumption Richards model had the smallest AIC value on all eleven estimation period relative to others candidate model. The smallest AIC indicating that the 5P logistic model and Richards model to be the best fit model in the estimation period for Poisson and normal assumption respectively.

4.4.2 Goodness of fit on all eleven estimation period for the prediction period

To select the best model for prediction, we used Pearson chi-squared for prediction period. According to the chi-squared value in Table 8 for Poisson distribution assumption, Richards had the smallest value under all eleven estimation period, which suggests that the best model for prediction of ten successive days relative to others models. Conversely, for normal distribution assumption, Gompertz had the smallest value in Table 8 under all eleven estimation period, which indicates that the best model for prediction of ten consecutive days relative to others. This results indicates that the best fit to the cumulative COVID-19 cases can be predict poorly due to the uncertainty brought by the noise.

Table 8: AIC and chi-squared for four models, two distribution on 11 different estimation period (70, 71, . . . , 80 days).

		Estimation Period		Prediction Period	
No. of days in the estimation period	Model	AIC(Poisson)	AIC(Normal)	chisq(Poisson)	chisq(Normal)
70	Richards	1833.3	**987.2**	**48.82559**	185.202
	3p logistic	3230.4	1012.7	902.9766	327.1458
	5p logistic	**1372.3**	1063.3	217.7025	294.9895
	Gompertz	2591.3	1449.9	1267.752	**44.15149**
71	Richards	1853	**1010.7**	**27.42302**	187.6401
	3p logistic	3272.7	1044.8	911.6323	338.9701
	5p logistic	**1386.5**	1094.3	236.5153	251.4118
	Gompertz	2680.5	1522.3	1395.035	**29.39806**
72	Richards	1865.4	**1030.4**	**26.56455**	150.2046
	3p logistic	3340.3	1075.7	861.9739	309.1327
	5p logistic	**1408.2**	1110.8	232.0423	179.4014
	Gompertz	1120.395	1486.6	1123.3	**31.28429**
73	Richards	1877.9	**1057**	**25.80068**	119.5072
	3p logistic	3407.4	1117.5	815.0648	285.7456
	5p logistic	**1429.7**	1137.5	257.1313	117.4597
	Gompertz	2797.6	1500.7	1087.155	**32.90451**
74	Richards	1890.5	**1080.3**	**24.37115**	97.23934
	3p logistic	3471.6	1156.2	773.4841	263.2989
	5p logistic	**1454.3**	1160.1	247.5568	84.72195
	Gompertz	2858	1513.9	1050.647	**33.79318**
75	Richards	1903.5	**1100.2**	**21.08806**	84.13319
	3p logistic	3531.3	1190.9	743.0804	247.8134
	5p logistic	**1471.3**	1178	257.5778	62.60566
	Gompertz	2921.7	1527.8	1002.933	**32.02975**
76	Richards	1916.5	**1131.8**	**18.07222**	73.57644
	3p logistic	3589.6	1245.8	715.2119	236.0881
	5p logistic	**1493.1**	1214.4	260.3409	51.15754
	Gompertz	2985.6	1586.1	955.6551	**30.57808**
77	Richards	1930	**1148.6**	**14.37939**	67.44428
	3p logistic	3644.2	1277.5	693.6695	213.485
	5p logistic	**1514.3**	1229.5	264.8965	43.42842
	Gompertz	3052	1601.2	903.7591	**27.00348**
78	Richards	1943.5	**1164.4**	**10.51942**	64.03319
	3p logistic	3696.8	1305.3	678.6021	213.8908
	5p logistic	**1535.6**	1243.6	271.5711	38.88799
	Gompertz	3118.8	1616.7	849.0743	**23.26585**
79	Richards	1957.3	**1179.9**	**7.189958**	61.94266
	3p logistic	3748.9	1338	655.8355	220.2228
	5p logistic	**1557.5**	1257.3	268.0032	35.61862
	Gompertz	3184.2	1632.1	794.0964	**19.08338**
80	Richards	1969.4	**1182.7**	**5.293357**	59.20057
	3p logistic	3804.2	1344.8	649.8869	215.8939
	5p logistic	**1581**	1252.6	249.4869	31.48529
	Gompertz	3244.2	1600.9	745.5794	**16.23784**

4.5 Goodness of fit for one, two,three, four, five and ten days ahead Prediction

In this study also make a prediction for one, two, three, four, five and ten days ahead by using eleven different estimation period(t=70,71 . . . , 80). This type of prediction were done by first, fit all this eleven models second, make t+1 prediction for each estimation period third, collect this predicted value from each estimation period consecutively from t=70,71 . . . , 80 finally, all this predicted value from each models gives one days ahead prediction. Applying the same procedure for the remaining prediction except the second step which is changed to t+2, t+3, t+4, t+5 and t+10 for two, three, four, five and ten days ahead prediction respectively. To select the best model according to prediction performance chi-squared were calculated for each days ahead prediction. The calculated chi-squared is of each days ahead prediction for all four proposed different models under the assumption of Poisson and normal distribution assumption presented in Table 9. Likewise the previous section the model with the smallest chi-squared is best model relative to others according prediction performance. The Richards model had the smallest chi-squared value under the Poisson distribution assumption while, Gompertz model had the smallest chi-squared value under the normal distribution assumption. So, the Richards and Gompertz model are the best model in terms of prediction performance under Poisson and normal distribution assumption respectively.

The internal validation of the performance of prediction for the Richards, 3PL, 5PL and Gompertz models under the Poisson and normal distribution assumption are displayed in Figure 12 and 13 respectively. The Pearson correlation coefficient was high in all proposed models for one days ahead prediction. The Richards model for one days ahead prediction under the Poisson distribution had a higher correlation and we can also see it from Figure 12 it is much closer to the observed cases. However, the Gompertz model for one days ahead prediction under the normal distribution had a higher correlation and as we can see from Figure 13 it is much closer to the observed COVID-19 cases.

Table 9: Chi-squared of a predication period for four different models and six different day ahead.

Day ahead	Model	chisq(Poisson)	chisq(Normal)
One days ahead	Richards	**19.50548**	58.92673
	3p logistic	532.8995	149.585
	5p logistic	109.9854	52.94483
	Gompertz	637.93	**18.97963**
Two days ahead	Richards	**14.44419**	77.84179
	3p logistic	628.5141	187.2363
	5p logistic	148.0867	75.21441
	Gompertz	713.256	**16.54887**
Three days ahead	Richards	**17.2091**	86.80644
	3p logistic	531.8038	219.2704
	5p logistic	179.4335	86.05263
	Gompertz	824.5353	**20.79948**
Four days ahead	Richards	**20.37062**	94.83963
	3p logistic	745.632	242.4324
	5p logistic	211.3298	95.75746
	Gompertz	943.3673	**25.80146**
Five days ahead	Richards	**23.29913**	104.2195
	3p logistic	805.0675	268.1092
	5p logistic	246.9054	106.5046
	Gompertz	1060.973	**30.66412**
Ten days ahead	Richards	**26.4612**	180.809
	3p logistic	1141.44	433.015
	5p logistic	469.087	196.597
	Gompertz	1605.72	**45.2913**

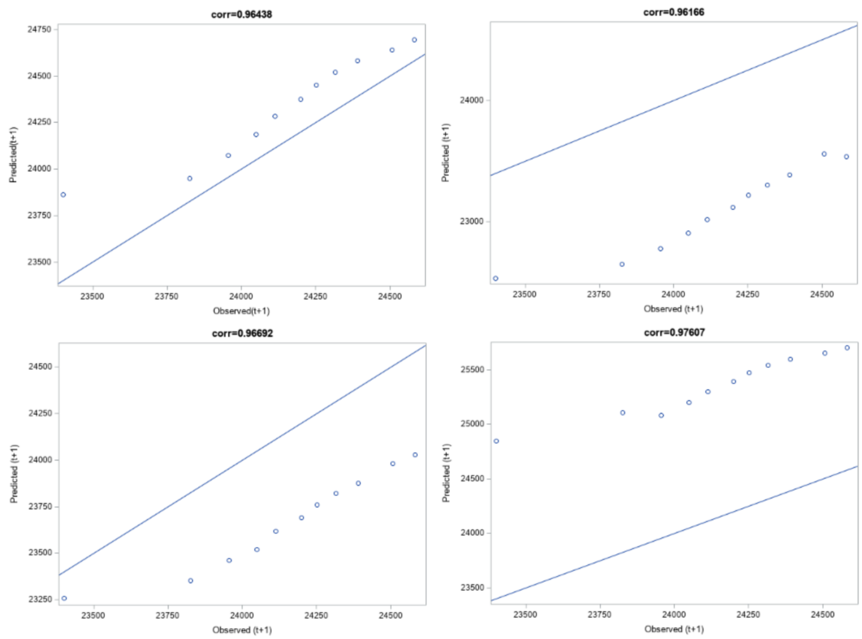

Figure 12: Internal validation for prediction of one days ahead from Richards, 3PL, 5PL, Gompertz and observed under the Poisson distribution.

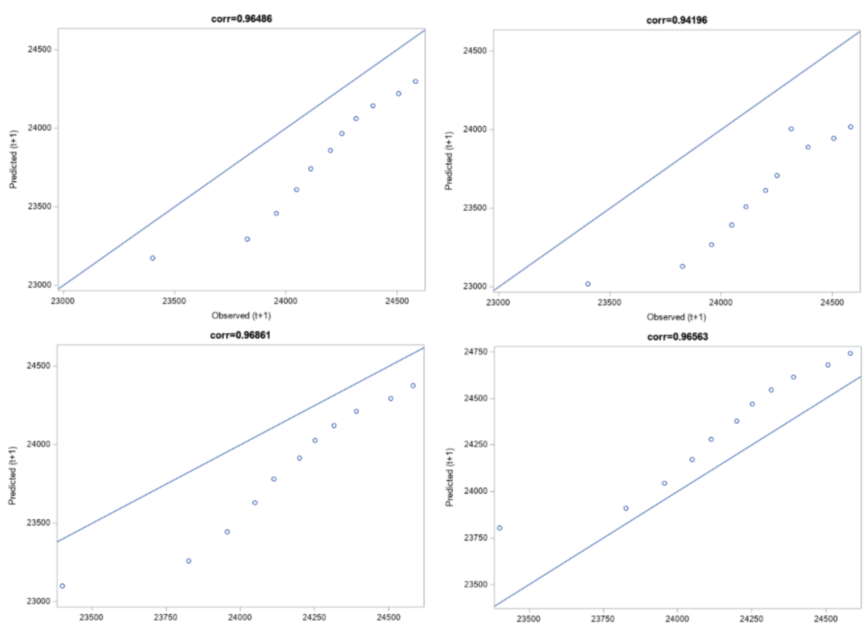

Figure 13: Internal validation for prediction of one days ahead from Richards, 3PL, 5PL and Gompertz and observed under the normal distribution.

5 Discussion and Conclusion

The main aim of this study was to use several previously validated approach of phenomenological models to make short term prediction of COVID-19 cases, and to develop a criteria for model selection for this purpose. To address these main objectives, the publicly available COVID-19 during March 4-June 30 2020 data were analysed using several non linear growth models under the Poisson and normal distribution assumption. This method requires that an initial value for each parameter be estimated, Starting value specification is one of the most difficult problems encountered in estimating parameters of non linear models. This impact also on the convergence process and obtaining a Positive definite Hessian matrix. Therefore, we fitted four models in each distribution for each estimation period. In both distribution and all estimation period, we used Richards, three parameter logistic, five parameter logistic and Gompertz model.

The first main objective of this study to make a short term prediction such as, ten consecutive days ahead, and one, two, three, four, five and ten days ahead prediction which is generated from eleven different estimation period. On a ten days ahead prediction, first fit the models for each different eleven estimation period and also for the first wave and then make a ten consecutive days prediction for each models under both assumption. Whereas, on one, two, three, four, five and ten days ahead prediction the predicted value were collected from each estimation period. And the other main aim of this study was model selection out of all proposed model on the estimation and prediction period under both distribution. So to made a model selection, we used AIC and the Pearson chi-squared value for estimation and prediction period respectively.

Based on the results presented in this study the five parameter logistic model was found to be the best fit model compared to others proposed models in the estimation period during the first wave under the Poisson assumption and Richards was found to be the best fit model in all eleven different estimation period under the normal assumption. On the other hand, a several non linear growth models were fitted to COVID-19 cases during the period 5 March 2020 to 22 June 2020 in South Africa [12], and they saw that the Richards models favored within the estimation period according to the AIC.

Nevertheless, Richards and Gompertz models have the best performance to make a short term prediction relative to others candidates of non linear growth models for Poisson and normal respectively during the first wave. This result agreed with that was fitted by Roosa to the COVID-19 pandemic China from February 5th to February 24th, 2020 [14] ,and the GLM and Richards models was favored to 5-day ahead forecasts generated in Hubei. A variety of phenomenological growth models were fitted to the COVID-19 pandemic during 22 January to 13 February 2020 for provinces of Guangdong and Zhejiang, China by Roosa and and colleagues [13], and they saw that the GLM and Richards model provide best prediction comparable mean estimates and prediction intervals, while the sub-epidemic model forecasts exhibit significantly greater uncertainty.

In this study we fitted a several models for eleven different estimation period and made a short term prediction. So, form this eleven different estimation period the 5P logistic model and Richards model were the best fit model in each estimation period based on their AIC under Poisson and normal distribution respectively. Nevertheless, Richards and Gompertz had the smallest chi-squared value on the prediction period which is the best model according to the performance of prediction for Poisson and normal respectively. The models that have the best goodness to fit within the estimation period can predict poorly beyond the estimation period were also observed by Reddy and colleagues [12].

In this particular study, the results suggested that the five parameters logistic model was found to be the best fit model on the estimation period during the first wave in both assumption.The five parameter logistic model and Richards was found to be the best fit model on the in all different eleven estimation period.Nevertheless, Richards and Gompertz models have the best performance to make a short term prediction relative to others candidates of non linear growth models under Poisson and normal assumption respectively during the first wave and in all different eleven estimation period. In conclusion, we have seen that Richards and Gompertz model have the best performance to make a short term prediction, but five parameter logistic model was found to be the best fit on estimation period under the Poisson and normal distribution assumption respectively. This suggested in general that to use AIC and chi-squared model selection criteria to select the best model relative to others proposed models on the estimation and prediction period respectively. Finally, we recommended to extend this study for up to date Covid-19 cases as well as deaths data including others non linear models.

References

[1] Agresti, Alan. Categorical Data Analysis. John Wiley Sons. 2013

[2] Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974;19(6):716–23.

[3] BBC News. Available: https://www.bbc.com/news/world-europe-51693160.

[4] Brennan,Cianan."Spread of Covid-19 infection has reached a 'plateau' with growth rate 'close to zero'". 16 April 2020, Available: https://www. irishexaminer.com/news/arid-30994503.html

[5] Citizens information. New measures to combat COVID-19 from midnight on 27 March 2020. Available: https://whatsnew.citizensinformation. ie/2020/03/27/new-measures-to-combat-covid-19-from-midnight-on-27-march-2020/

[6] Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 1825; 182:513-583.

[7] Gottschalk PG, Dunn JR. The five-parameter logistic: a characterization and comparison with the four-parameter logistic. Anal Biochem. 2005 Aug 1;343(1):54-65. doi: 10.1016/j.ab.2005.04.035. PMID: 15953581.

[8] Guidotti, E., Ardia, D. "COVID-19 Data Hub", Journal of Open Source Software 5(51):2376,doi:10.21105/joss.02376. 2020, Available:https://covid19datahub.io/

[9] Hsieh Y-H, Chen CWS. Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks. Trop Med Int Heal. 2009;14(6):628–38 [cited 2020 Jul 26].

[10] Leahy, Pat; Cullen, Paul; Lynch, Suzanne; Kelly, Fiach (12 March 2020). "Coronavirus: Schools, colleges and childcare facilities in Ireland to shut". The Irish Times. Retrieved 12 March 2020.

[11] Power, Jack. Over-65s account for 90% of coronavirus deaths recorded in Ireland. 13 April 2020.

[12] Reddy, T., Shkedy, Z., Janse van Rensburg, C. et al. Short-term real-time prediction of total number of reported COVID-19 cases and deaths in South Africa: a data driven approach. BMC Med Res Methodol 21, 15 (2021). https://doi.org/10.1186/s12874-020-01165-x

[13] Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, et al. Short-term Forecasts of the COVID-19 Epidemic in Guangdong and Zhejiang, China:February 13–23, 2020. J Clin Med. 2020;9(2):596 [cited 2020 Jun 24].

[14] Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infect Dis Model. 2020;5:256–63.

[15] Sebrango-Rodríguez CR, Martínez-Bello DA, Sánchez-Valdés L, Thilakarathne PJ, Del Fava E, Van Der Stuyft P, et al. Real-time parameter estimation of Zika outbreaks using model averaging. Epidemiol Infect. 2017;145(11):2313–23.

[16] World Health Organization. Available: https://www.who.int/health-topics/ coronavirus / (Accessed on 25 May 2021).

[17] World Health Organization. Available: https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/01/2019-ncov-outbreak-first-cases-confirmed-in-europe.

[18] World Health Organization. Available:https://covid19.who.int/region/euro/ country/ie. (Accessed on 27 May 2021).

[19] Y.-H. Hsieh,Richards Model: A Simple Procedure for Real-time Prediction of Outbreak Severity.31 May 2014.

[20] Zoe S. List of countries without coronavirus on May 30 2021. Available: https://koryogroup.com/blog/are-there-countries-without-coronavirus.

Appendix A: Tables

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	25419	27.4285	119	926.75	<.0001	25365	25474
k	0.3604	0.008566	119	42.08	<.0001	0.3435	0.3774
gamma	0.08735	0.000606	119	144.1	<.0001	0.08615	0.08855
eta	41.0674	0.05733	119	716.3	<.0001	40.9539	41.1809

Table 10: Parameter estimates of Richards model for Poisson assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	24770	21.7997	119	1136.26	<.0001	24727	24813
gamma	0.1318	0.000336	119	391.94	<.0001	0.1311	0.1324
eta	43.5155	0.03729	119	1167.01	<.0001	43.4416	43.5893

Table 11: Parameter estimates of three parameter logistic model for Poisson assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	26102	27.4821	119	949.79	<.0001	26048	26157
alpha0	40.3101	2.8055	119	14.37	<.0001	34.7549	45.8654
gamma	0.06562	0.000183	119	357.84	<.0001	0.06526	0.06599
eta	38.7703	0.03698	119	1048.27	<.0001	38.697	38.8435

Table 12: Parameter estimates of Gompertz model for Poisson assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	25318	21.7747	119	1162.72	<.0001	25275	25361
k	0.4277	0.01956	119	21.86	<.0001	0.389	0.4665
gamma	0.09216	0.00085	119	108.33	<.0001	0.09048	0.09385
eta	41.4269	0.1156	119	358.34	<.0001	41.198	41.6558
se	14023	793.32	119	17.68	<.0001	12452	15594

Table 13: Parameter estimates of Richards model for normal assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	25075	18.196	119	1378.05	<.0001	25039	25111
gamma	0.1158	0.00046	119	254.12	<.0001	0.1149	0.1167
eta	44.1022	0.03936	119	1120.43	<.0001	44.0243	44.1802
se	14043	448.36	119	31.32	<.0001	13155	14931

Table 14: Parameter estimates of three parameter logistic model for normal assumption during March 4-June 30 2020 2020.

Parameter	Estimate	Std.error	DF	t Value	Pr >$\|t\|$	95% Confidence Limits	
alpha	25511	21.0137	119	1214.02	<.0001	25469	25553
alpha0	405.76	26.5084	119	15.31	<.0001	353.27	458.25
gamma	0.07855	0.00037	119	213.61	<.0001	0.07782	0.07928
eta	39.1494	0.04912	119	796.97	<.0001	39.0522	39.2467
se	14041	553.55	119	25.37	<.0001	12945	15137

Table 15: Parameter estimates of Gompertz model for normal assumption during March4-June 30 2020 2020

Appendix B: Graphs

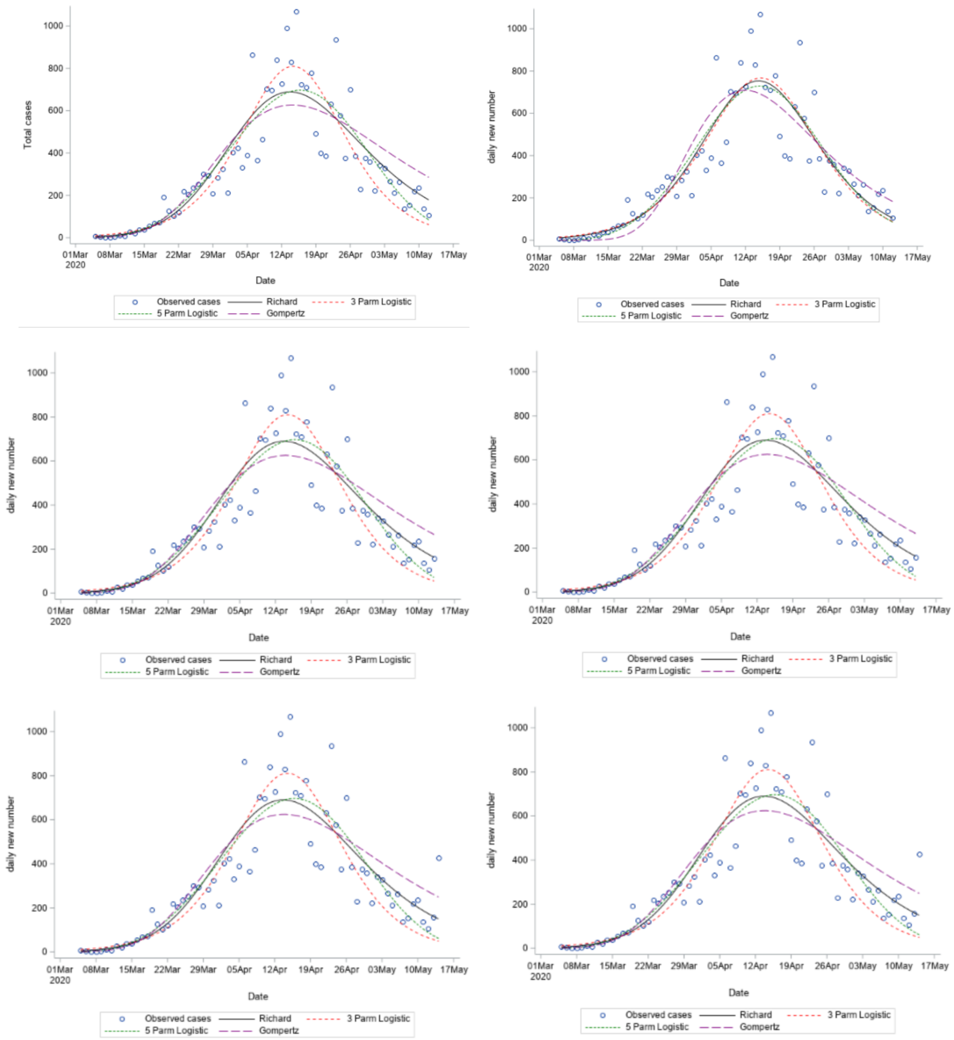

Figure 14: The incidence and predicted incidence of a COVID-19 cases from observed, Richards, 3PL, 5PL and Gompertz for 7o,71 and 72 days estimation period(row wise) under Poisson(left) and normal(right).

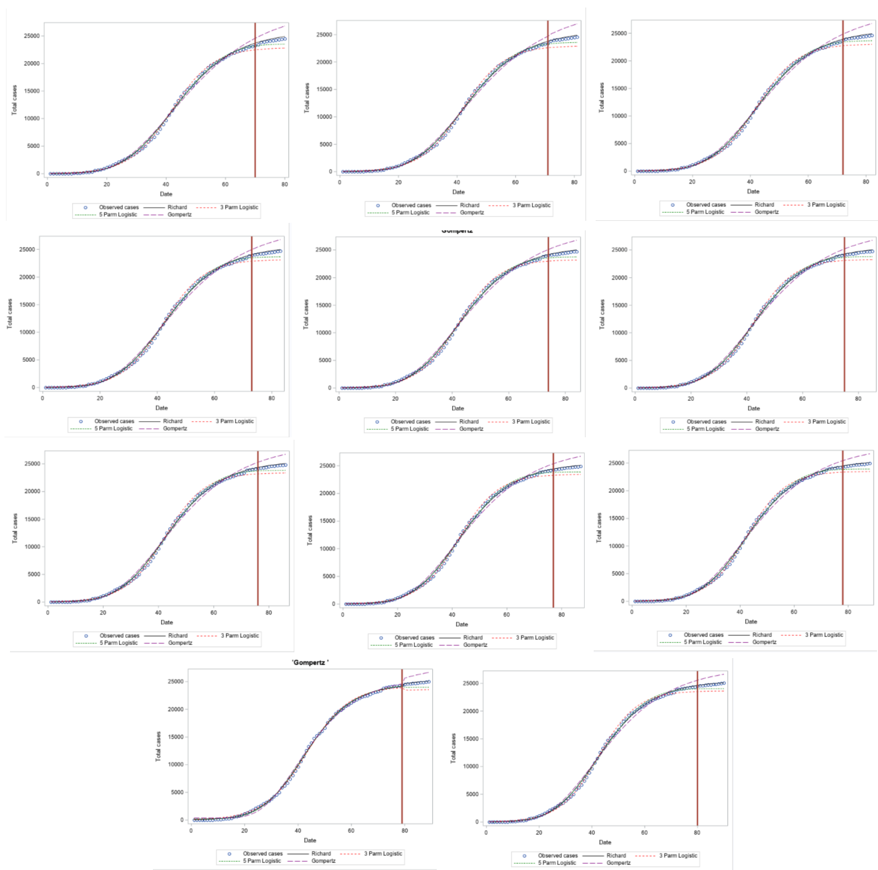

Figure 15: A ten day Predicted cumulative COVID-19 cases in Ireland from the Richards, 3P,5P and Gompertz model under Poisson on the estimation periods of 70, 71, 72, ..., 80 days.

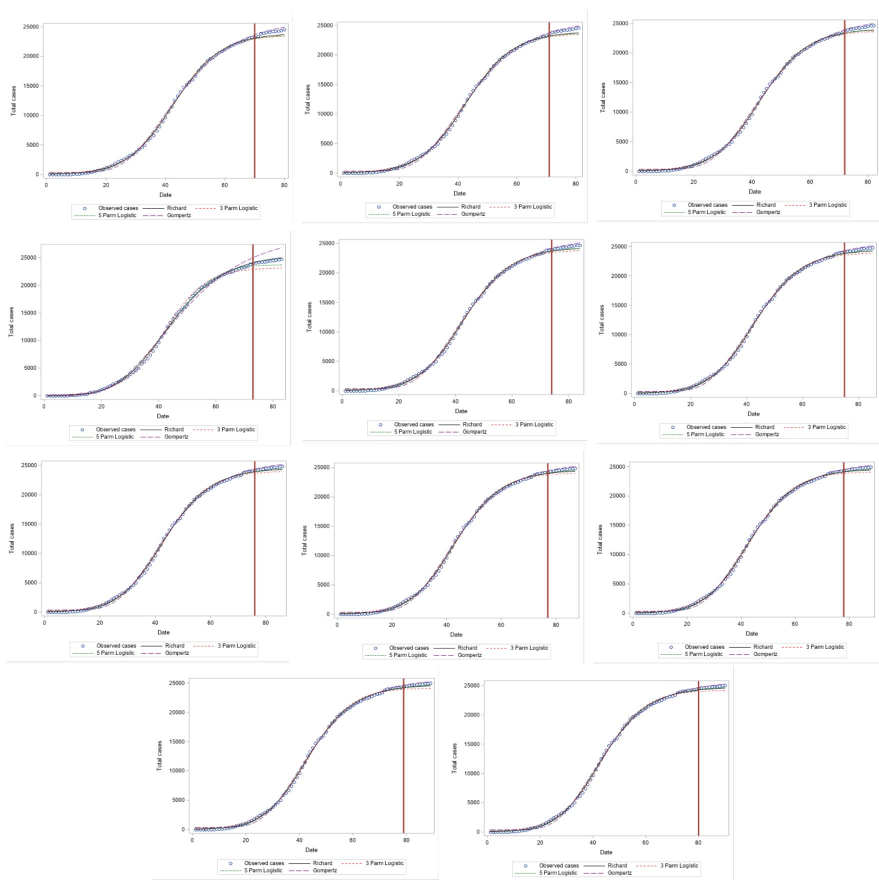

Figure 16: A ten day Predicted cumulative COVID-19 cases in Ireland from the Richards, 3P,5P and Gompertz model under normal on the estimation periods of 70, 71, 72, . . . , 80 days.

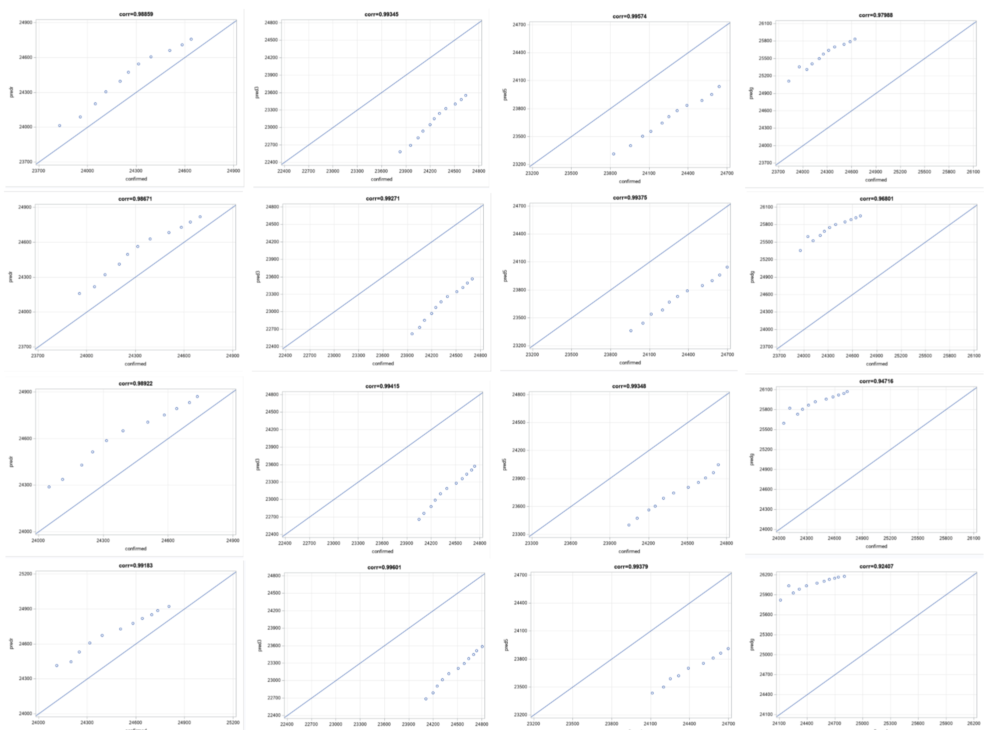

Figure 17: Internal validation for predictions 2, 3, 4, and 5 days ahead prediction for the estimation period 70,71,. . . , 80 days under Poisson assumption.

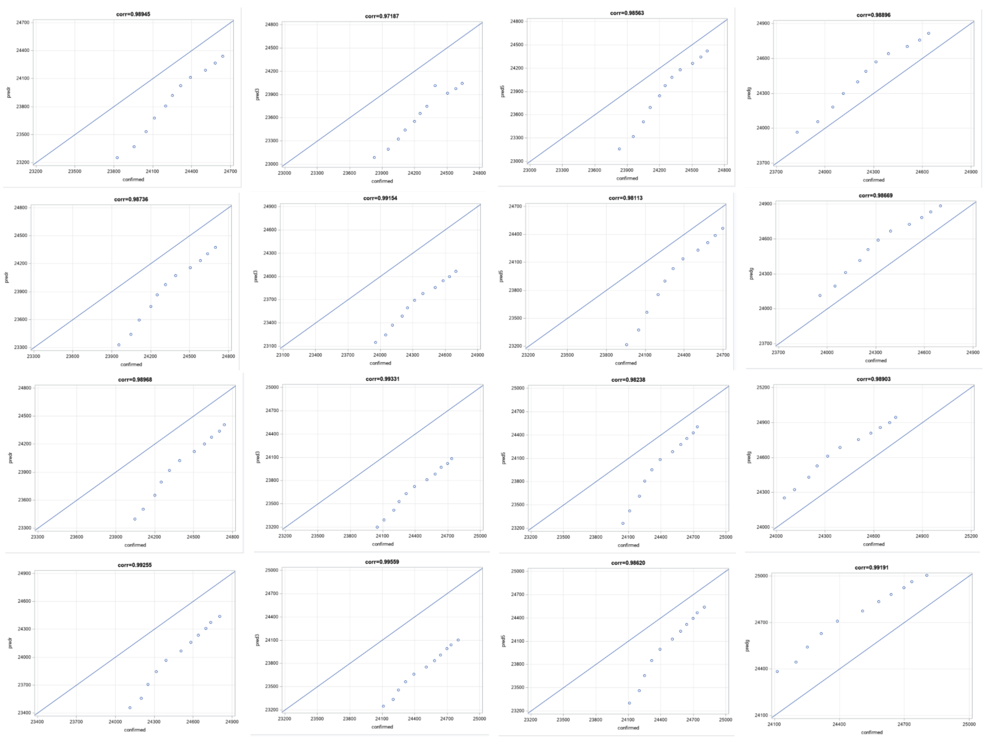

Figure 18: Internal validation for predictions 2, 3, 4, and 5 days ahead prediction for the estimation period 70,71,..., 80 days under normal assumption.

Code for basic models

```
proc import datafile="C:\Users\Haymanot\Desktop\thesis2021\Idata.csv"
out=dataId dbms=csv;
run;
/**** Richards Model for Ireland */
proc nlmixed data=dataId;
title 'Richards model';
parms alpha=26000 k=0.5 gamma=0.05 eta=45;
bounds K>0, alpha>0, eta>0, gamma>0;
mu = alpha*((1+k*exp(-gamma*(time-eta)))**(-1/k));
model confirmed ~ poisson(mu);
predict mu out= rpred;
run;
/**** 3 parameter logistic Model for Ireland****/
proc nlmixed data=dataId;
title '3 param logistic';
parms alpha = 26000, gamma = 0.05, eta = 45;
bounds alpha>0, eta>0, gamma>0;
mu = alpha/(1+exp(-gamma*(time-eta)));
model confirmed ~ poisson(mu);
predict mu out=pred3;
run;
/**** 5 parameter logistic Model for Ireland ****/
proc nlmixed data=dataId;
title '5 param logistic';
parms alpha = 26000, alpha0=1, k = 0.005, gamma = 0.05, eta = 45;
bounds k>0, alpha>0, alpha0>0, eta>0, gamma>0;
mu = alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*((time/eta)**gamma))**k;
model confirmed ~ poisson(mu);
predict mu out=pred5;
run;
/**** Gompertz Model for Ireland ****/
proc nlmixed data=dataId;
title 'Gompertz ';
parms alpha = 26000, alpha0 = 1, gamma = 0.05, eta = 45;
bounds alpha > 0, alpha0> 0, gamma > 0, eta > 0;
mu = alpha0+((alpha-alpha0)*exp(-exp(-gamma*(time-eta))));
model confirmed ~ poisson(mu);
predict mu out=gopred;
```

```
run;
/******plote of predictive and observed ****/
data datar; set rpred;
keep id time date confirmed pred confirmed_daily
administrative_area_level_1;
rename pred=predr;
run;
data data3; set pred3;
keep time pred;
rename pred=pred3;
run;
data data5; set pred5;
keep time pred;
rename pred=pred5;
run;
data datag; set gopred;
keep time  pred;
rename pred=predg;
run;
data Ireland_poisson;
merge datar data3 data5 datag;
by time;
run;
proc print data=Ireland_poisson; run;
proc export
  data=Ireland_poisson
  dbms=CSV
  outfile="C:\Users\Haymanot\Desktop\thesis2021\Ireland_poisson.csv"
  replace;
run;
proc sgplot data= Ireland_poisson;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=date y=pred3 / lineattrs=(color= orange PATTERN= 5)
LEGENDLABEL = "3 Parm Logistic";
series x=date y=pred5 / lineattrs=(color=green PATTERN= 6)
LEGENDLABEL = "5 Parm Logistic";
series x=date y=predg / lineattrs=(color=red PATTERN= 2)
```

```
LEGENDLABEL = "Gompertz";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_poisson;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_poisson;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed /
LEGENDLABEL = "Observed cases";
series x=date y=pred3 / lineattrs=(color= orange PATTERN= 5)
LEGENDLABEL = "3 Parm Logistic";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_poisson;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=pred5 / lineattrs=(color=green PATTERN= 6)
LEGENDLABEL = "5 Parm Logistic";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_poisson;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predg / lineattrs=(color=red PATTERN= 2)
LEGENDLABEL = "Gompertz";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
/****plots of incidence and predicted incidence****/
data Ireland_poisson_model_1;
```

```
    set Ireland_poisson;
    by time;
    prev_rpred = lag(predr);
    prev_pred3 = lag(pred3);
prev_pred5 = lag(pred5);
prev_predg = lag(predg);

    inc_rpred = predr - prev_rpred;
inc_pred3 = pred3 - prev_pred3;
    inc_pred5 = pred5 - prev_pred5;
inc_predg = predg - prev_predg;
if time =1 then delete;
run;
title " ";
proc sgplot data= Ireland_poisson_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_rpred  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=date y=inc_pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
series x=date y=inc_pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
series x=date y=inc_predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = " Gompertz";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_poisson_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_rpred  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_poisson_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
```

```
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_poisson_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_poisson_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = " Gompertz";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
/********** Normal **********/
/**** Richards Model for Ireland****/
proc nlmixed data=dataId;
title 'Richards model';
parms alpha=26000 k=0.5 gamma=0.04 eta=45, se=14000;
bounds K>0, alpha>0, eta>0, gamma>0;
mu = alpha*((1+k*exp(-gamma*(time-eta)))**(-1/k));
model confirmed ~ normal(mu,se);
predict mu out=rpredn;
run;
/**** 3 parameter logistic Model for Ireland****/
proc nlmixed data=dataId;
title '3 param logistic | normal';
parms alpha = 26000, gamma = 0.04, eta = 45, se=14000;
bounds alpha>0, eta>0, gamma>0;
mu = alpha/(1+exp(-gamma*(time-eta)));
model confirmed ~ normal(mu,se);
predict mu out=pred3n;
run;
/**** 5 parameter logistic Model for Ireland****/
proc nlmixed data=dataId ;
```

```
title '5 param logistic';
parms alpha = 26000, k = 0.5, alpha0=1, gamma = 0.04, eta = 45, se=14000;
bounds k>0, alpha>0, eta>0, gamma>0;
mu = alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*((time/eta)**gamma))**k;
model confirmed ~ normal(mu, se);
predict mu out=pred5n;
run;
/**** Gompertz Model for Ireland ****/
proc nlmixed data=dataId;
title 'Gompertz ';
parms alpha = 26000 alpha0 = 1 gamma = 0.04 eta = 45 se=14000;
bounds alpha0 >0, alpha>0, eta>0, gamma>0;
mu = alpha0+((alpha-alpha0)*exp(-exp(-gamma*(time-eta))));
model confirmed ~ normal(mu, se);
predict mu out=gopredn;
run;
/******plote of predictive and observed ****/
data datar; set rpred;
keep id time date confirmed pred confirmed_daily
administrative_area_level_1;
rename pred=predr;
run;
data data3; set pred3;
keep time pred;
rename pred=pred3;
run;
proc print data=pred3;run;
data data5; set pred5;
keep time pred;
rename pred=pred5;
run;
data datag; set gopred;
keep time  pred;
rename pred=predg;
run;
data Ireland_normal;
merge datar data3 data5 datag;
by time;
run;
proc export
```

```
  data=Ireland_normal
  dbms=CSV
  outfile="C:\Users\Haymanot\Desktop\thesis2021\Ireland_normal.csv"
  replace;
run;
proc sgplot data= Ireland_normal;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=date y=pred3 / lineattrs=(color= orange PATTERN= 5)
LEGENDLABEL = "3 Parm Logistic";
series x=date y=pred5 / lineattrs=(color=green PATTERN= 6)
LEGENDLABEL = "5 Parm Logistic";
series x=date y=predg / lineattrs=(color=red PATTERN= 2)
LEGENDLABEL = "Gompertz";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_normal;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_normal;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=pred3 / lineattrs=(color= orange PATTERN= 5)
LEGENDLABEL = "3 Parm Logistic";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_normal;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=pred5 / lineattrs=(color=green PATTERN= 6)
LEGENDLABEL = "5 Parm Logistic";
```

```
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
proc sgplot data= Ireland_normal;
title "Predicted and reported daily total number of COVID-19 cases";
scatter x=date y=confirmed / LEGENDLABEL = "Observed cases";
series x=date y=predg / lineattrs=(color=red PATTERN= 2)
LEGENDLABEL = "Gompertz";
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
/****plots of incidence and predicted incidence****/
data Ireland_normal_model_1;
    set Ireland_normal;
    by time;
    prev_rpred = lag(predr);
    prev_pred3 = lag(pred3);
prev_pred5 = lag(pred5);
prev_predg = lag(predg);

    inc_rpred = predr - prev_rpred;
inc_pred3 = pred3 - prev_pred3;
    inc_pred5 = pred5 - prev_pred5;
inc_predg = predg - prev_predg;
if time =1 then delete;
run;
title " ";
proc sgplot data= Ireland_normal_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_rpred  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=date y=inc_pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
series x=date y=inc_pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
series x=date y=inc_predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = " Gompertz";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
```

```
title " ";
proc sgplot data= Ireland_normal_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_rpred  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_normal_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
title " ";
proc sgplot data= Ireland_normal_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;

title " ";
proc sgplot data= Ireland_normal_model_1 noborder;
scatter x=date y=confirmed_daily / LEGENDLABEL = "Observed cases";
series x=date y=inc_predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = " Gompertz";
YAXIS LABEL = 'daily new number';
XAXIS LABEL = 'Date';
run;
```

code for prediction

```
/*********Prediction for Poisson**********/
/**** Richards Model for Ireland */
proc nlmixed data=dataId;
title 'Richards model';
```

```
parms alpha=26000 k=0.5 gamma=0.05 eta=45;
bounds K>0, alpha>0, eta>0, gamma>0;
mu = alpha*((1+k*exp(-gamma*(time-eta)))**(-1/k));
model confirmed ~ poisson(mu);
predict mu out= rpred;
estimate "120" alpha*((1+k*exp(-gamma*(120-eta)))**(-1/k));
estimate "121" alpha*((1+k*exp(-gamma*(121-eta)))**(-1/k));
estimate "122" alpha*((1+k*exp(-gamma*(122-eta)))**(-1/k));
estimate "123" alpha*((1+k*exp(-gamma*(123-eta)))**(-1/k));
estimate "124" alpha*((1+k*exp(-gamma*(124-eta)))**(-1/k));
estimate "125" alpha*((1+k*exp(-gamma*(125-eta)))**(-1/k));
estimate "126" alpha*((1+k*exp(-gamma*(126-eta)))**(-1/k));
estimate "127" alpha*((1+k*exp(-gamma*(127-eta)))**(-1/k));
estimate "128" alpha*((1+k*exp(-gamma*(128-eta)))**(-1/k));
estimate "129" alpha*((1+k*exp(-gamma*(129-eta)))**(-1/k));
run;
/**** 3 parameter logistic Model for Ireland****/
proc nlmixed data=dataId;
title '3 param logistic';
parms alpha = 26000, gamma = 0.05, eta = 45;
bounds alpha>0, eta>0, gamma>0;
mu = alpha/(1+exp(-gamma*(time-eta)));
model confirmed ~ poisson(mu);
predict mu out=pred3;
estimate "120" alpha/(1+exp(-gamma*(120-eta)));
estimate "121" alpha/(1+exp(-gamma*(121-eta)));
estimate "122" alpha/(1+exp(-gamma*(122-eta)));
estimate "123" alpha/(1+exp(-gamma*(123-eta)));
estimate "124" alpha/(1+exp(-gamma*(124-eta)));
estimate "125" alpha/(1+exp(-gamma*(125-eta)));
estimate "126" alpha/(1+exp(-gamma*(126-eta)));
estimate "127" alpha/(1+exp(-gamma*(127-eta)));
estimate "128" alpha/(1+exp(-gamma*(128-eta)));
estimate "129" alpha/(1+exp(-gamma*(129-eta)));
run;
/**** 5 parameter logistic Model for Ireland ****/
proc nlmixed data=dataId;
title '5 param logistic';
parms alpha = 26000, alpha0=1, k = 0.005, gamma = 0.05, eta = 45;
bounds k>0, alpha>0, alpha0>0, eta>0, gamma>0;
```

```
mu = alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*((time/eta)**gamma))**k;
model confirmed ~ poisson(mu);
predict mu out=pred5;
estimate "120" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((120/eta)**gamma))**k;
estimate "121" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((121/eta)**gamma))**k;
estimate "122" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((122/eta)**gamma))**k;
estimate "123" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((123/eta)**gamma))**k;
estimate "124" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((124/eta)**gamma))**k;
estimate "125" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((125/eta)**gamma))**k;
estimate "126" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((126/eta)**gamma))**k;
estimate "127" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((127/eta)**gamma))**k;
estimate "128" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((128/eta)**gamma))**k;
estimate "129" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((129/eta)**gamma))**k;
run;
/**** Gompertz Model for Ireland ****/
proc nlmixed data=dataId;
title 'Gompertz ';
parms alpha = 26000, alpha0 = 1, gamma = 0.05, eta = 45;
bounds alpha > 0, alpha0> 0, gamma > 0, eta > 0;
mu = alpha0+((alpha-alpha0)*exp(-exp(-gamma*(time-eta))));
model confirmed ~ poisson(mu);
predict mu out=gopred;
estimate "120" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(120-eta))));
estimate "121" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(121-eta))));
estimate "122" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(122-eta))));
estimate "123" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(123-eta))));
estimate "124" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(124-eta))));
estimate "125" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(125-eta))));
estimate "126" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(126-eta))));
estimate "127" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(127-eta))));
```

```
estimate "128" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(128-eta))));
estimate "129" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(129-eta))));
run;
proc import datafile="C:\Users\Haymanot\Desktop\thesis2021\ipp.csv"
out=ipp dbms=csv;
run;
title " ";
proc sgplot data=ipp;
scatter x=time y=confirmed / LEGENDLABEL = "Observed cases";
*scatter x=date y= obs/ markerattrs=(color=Gold symbol=Star)
LEGENDLABEL = "New observed";
series x=time y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=time y=pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
series x=time y=pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
series x=time y=predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = "Gompertz";
refline 119
 /axis=x   lineattrs=GraphData2(thickness=3);
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
    /*********Prediction for normal**********/
/**** Richards Model for Ireland */
proc nlmixed data=dataId;
title 'Richards model';
parms alpha=26000 k=0.5 gamma=0.04 eta=45, se=14000;
bounds K>0, alpha>0, eta>0, gamma>0;
mu = alpha*((1+k*exp(-gamma*(time-eta)))**(-1/k));
model confirmed ~ normal(mu,se);
predict mu out=rpredn;
estimate "120" alpha*((1+k*exp(-gamma*(120-eta)))**(-1/k));
estimate "121" alpha*((1+k*exp(-gamma*(121-eta)))**(-1/k));
estimate "122" alpha*((1+k*exp(-gamma*(122-eta)))**(-1/k));
estimate "123" alpha*((1+k*exp(-gamma*(123-eta)))**(-1/k));
estimate "124" alpha*((1+k*exp(-gamma*(124-eta)))**(-1/k));
estimate "125" alpha*((1+k*exp(-gamma*(125-eta)))**(-1/k));
estimate "126" alpha*((1+k*exp(-gamma*(126-eta)))**(-1/k));
```

```
estimate "127" alpha*((1+k*exp(-gamma*(127-eta)))**(-1/k));
estimate "128" alpha*((1+k*exp(-gamma*(128-eta)))**(-1/k));
estimate "129" alpha*((1+k*exp(-gamma*(129-eta)))**(-1/k));
run;
/**** 3 parameter logistic Model for Ireland****/
proc nlmixed data=dataId;
title '3 param logistic | normal';
parms alpha = 26000, gamma = 0.04, eta = 45, se=14000;
bounds alpha>0, eta>0, gamma>0;
mu = alpha/(1+exp(-gamma*(time-eta)));
model confirmed ~ normal(mu,se);
predict mu out=pred3n;
estimate "120" alpha/(1+exp(-gamma*(120-eta)));
estimate "121" alpha/(1+exp(-gamma*(121-eta)));
estimate "122" alpha/(1+exp(-gamma*(122-eta)));
estimate "123" alpha/(1+exp(-gamma*(123-eta)));
estimate "124" alpha/(1+exp(-gamma*(124-eta)));
estimate "125" alpha/(1+exp(-gamma*(125-eta)));
estimate "126" alpha/(1+exp(-gamma*(126-eta)));
estimate "127" alpha/(1+exp(-gamma*(127-eta)));
estimate "128" alpha/(1+exp(-gamma*(128-eta)));
estimate "129" alpha/(1+exp(-gamma*(129-eta)));
run;
/**** 5 parameter logistic Model for Ireland ****/
proc nlmixed data=dataId ;
title '5 param logistic';
parms alpha = 26000, k = 0.5, alpha0=1, gamma = 0.04, eta = 45, se=14000;
bounds k>0, alpha>0, eta>0, gamma>0;
mu = alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*((time/eta)**gamma))**k;
model confirmed ~ normal(mu, se);
predict mu out=pred5n;
estimate "120" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((120/eta)**gamma))**k;
estimate "121" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((121/eta)**gamma))**k;
estimate "122" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((122/eta)**gamma))**k;
estimate "123" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((123/eta)**gamma))**k;
estimate "124" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
```

```
((124/eta)**gamma))**k;
estimate "125" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((125/eta)**gamma))**k;
estimate "126" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((126/eta)**gamma))**k;
estimate "127" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((127/eta)**gamma))**k;
estimate "128" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((128/eta)**gamma))**k;
estimate "129" alpha+(alpha0-alpha)/(1+((2**(1/k))-1)*
((129/eta)**gamma))**k;
run;
/**** Gompertz Model for Ireland ****/
proc nlmixed data=dataId;
title 'Gompertz ';
parms alpha = 26000 alpha0 = 1 gamma = 0.04 eta = 45 se=14000;
bounds alpha0 >0, alpha>0, eta>0, gamma>0;
mu = alpha0+((alpha-alpha0)*exp(-exp(-gamma*(time-eta))));
model confirmed ~ normal(mu, se);
predict mu out=gopredn;
estimate "120" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(120-eta))));
estimate "121" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(121-eta))));
estimate "122" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(122-eta))));
estimate "123" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(123-eta))));
estimate "124" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(124-eta))));
estimate "125" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(125-eta))));
estimate "126" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(126-eta))));
estimate "127" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(127-eta))));
estimate "128" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(128-eta))));
estimate "129" alpha0+((alpha-alpha0)*exp(-exp(-gamma*(129-eta))));
run;
proc import datafile="C:\Users\Haymanot\Desktop\thesis2021\ipn.csv"
out=ipn dbms=csv;
run;
proc print data=ipn;run;
title " ";
proc sgplot data=ipn;
scatter x=time y=confirmed / LEGENDLABEL = "Observed cases";
*scatter x=date y= obs/ markerattrs=(color=Gold symbol=Star)
LEGENDLABEL = "New observed";
```

```
series x=time y=predr  / lineattrs=(color=black PATTERN= 1)
LEGENDLABEL = "Richard";
series x=time y=pred3 / lineattrs=(color= red PATTERN= 2)
LEGENDLABEL = "3 Parm Logistic";
series x=time y=pred5 / lineattrs=(color=green PATTERN= 3)
LEGENDLABEL = "5 Parm Logistic";
series x=time y=predg / lineattrs=(color=purple PATTERN= 4)
LEGENDLABEL = "Gompertz";
refline 119
 /axis=x   lineattrs=GraphData2(thickness=3);
YAXIS LABEL = 'Total cases';
XAXIS LABEL = 'Date';
run;
```