Analysis and Prediction of COVID-19 using SIR, SEIR, and Machine Learning Models: Australia, Italy, and UK Cases

Iman Rahimi[1], Amir H Gandomi[2,*], Fang Chen[2]

Abstract- The novel Coronavirus disease, known as COVID-19, is an outbreak that started in Wuhan, one of the Central Chinese cities. In this report, a short analysis focusing on Australia, Italy, and the United Kingdom has been conducted. The analysis includes confirmed and recovered cases and deaths, the growth rate in Australia as compared with Italy and the United Kingdom, and the outbreak in different Australian cities. Mathematical approaches based on the susceptible, infected, and recovered case (SIR) and susceptible, exposed, infected, and recovered (SEIR) models were proposed to predict the epidemiology in the countries. Since the performance of the classic form of SIR and SEIR depends on parameter settings, some optimization algorithms, namely, the Broyden–Fletcher–Goldfarb–Shanno (BFGS), conjugate gradients (CG), L-BFGS-B, and Nelder-Mead are proposed to optimize the parameters of SIR and SEIR models and improve its predictive capabilities. The results of optimized SIR and SEIR models are compared with the Prophet algorithm and logistic function as two known ML algorithms. The results show that different algorithms display different behaviours in different countries. However, the improved version of the SIR and SEIR models have a better performance compared with other mentioned algorithms described in this study. Moreover, the Prophet algorithm works better for Italy and the United Kingdom cases than for Australian cases and Logistic function compared with Prophet algorithm has a better performance in these cases. It seems that Prophet algorithm is suitable for data with increasing trend in pandemic situations. Optimization of the SIR and SEIR models parameters has yielded a significant improvement in the prediction accuracy of the models. Although there are several algorithms for prediction of this Pandemic, there is no certain algorithm that would be the best one for all cases.

[1] *Universiti Putra Malaysia, Malaysia*
[2] *Data Science Institute, University of Technology Sydney, Australia,*
corresponding author, email: gandomi@uts.edu.au

Keywords- COVID-19, Analysis, Machine learning, SIR and SEIR models, Optimization.

Introduction: In December 2019, the Chinese government informed the rest of the world that a virus was spreading throughout China. A few months later, it spread very rapidly to some other countries. This virus is the Severe Acute Respiratory Syndrome- Related Coronavirus 2 which causes the disease novel coronavirus known as COVID-19. The United States Centers for Disease Control and Prevention (CDC) identified a seafood market in Wuhan that was suspected to be at the centre of the outbreak. The World Health Organization (WHO) reported a case in Thailand on Jan 13, which was the first time it was identified outside China. On Jan 16, Japan confirmed its first case of this novel coronavirus. On Jan 20, South Korea identified its first confirmed case of the new coronavirus. Nowadays, most countries in the world are affected by this virus.

Putra and Khozin Mu'tamar (2019) used Particle Swarm Optimization (PSO) algorithm to estimate parameters (Susceptible, Infected, Recovered) in the SIR model. The results indicate that the suggested method is precise enough with low error compared to analytical methods. Mbuvha and Marwala (2020) calibrated the SIR model to South Africa after considering different scenarios for R0 (reproduction number) for reporting infections and healthcare resource estimation for the next few days. Qi, Xiao et al. (2020) proposed that both daily temperature and relative humid-ity influenced the occurrence of COVID-19 in Hubei province and insome other provinces.

Salgotra, Gandomi et al. (2020) developed two COVID-19 prediction models based on genetic programming and applied this model in India. Findings from a study by (Salgotra, Gandomi et al. 2020) show genetic evolutionary programming models are highly reliable for COVID-19 cases in India.

In January 2020, the first case of Covid-19 was reported in Australia. In this report, a short analysis focusing on Australia was addressed and reported and continued as a simulation for the next few days.

The manuscript is organized in several sections. Section I presents the research methodology. Section II and III introduce the SIR and SEIR models. Section IV shows the prediction algorithms (logistic function and Prophet algorithm). Sections V shows the results. The conclusion and discussion are provided in the last section.

I. Research methodology

The study was carried out in several phases. For the first step, data were collected from World Health Organization (WHO) and John Hopkins University since they collect data from different organizations. After that, data were analyzed and preprocessed in order to avoid any duplicated and missing values. Numerical tests were performed using Python and R and executed on a computer Intel ® Core i7-4510U 2.0 GHz 8 GB DDR3 Memory (Supplementary file). The flowchart of the research methodology is provided in Figure 1.

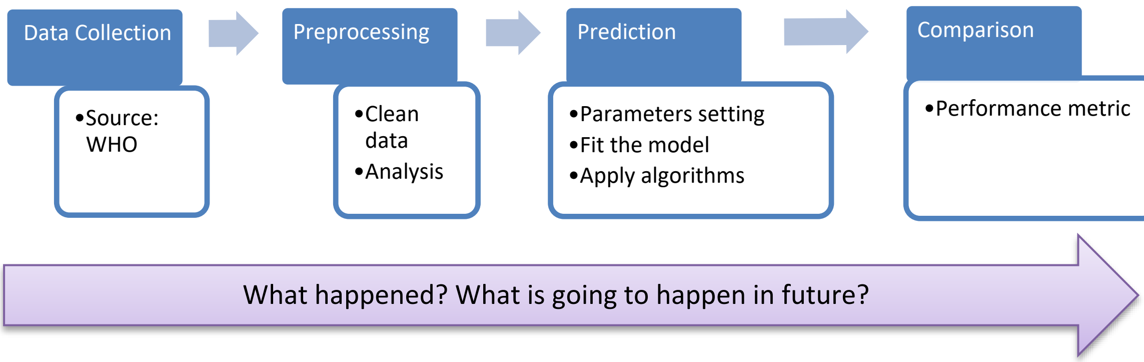


Figure 1 Flowchart of the current research process

II. The SIR model

This section introduces the classic form of the SIR model (Kermack and McKendrick 1932, Capasso and Serio 1978) that is used to describe the transmission of COVID-19 virus in Australia, Italy, and the United Kingdom. The flowchart of SIR model is shown in Figure 2:

Figure 2 SIR model

The SIR model shows how a disease spreads through a population. The equations of SIR model are as shown below (Weiss 2013):

$$\frac{ds}{dt} = -\beta IS \tag{1}$$

$$\frac{dI}{dt} = \beta IS - \gamma I \tag{2}$$

$$\frac{dR}{dt} = \gamma I \tag{3}$$

in which

- S is the number of individual susceptible at time t.
- I is the number of infected individuals at time t.
- R is the number of recovered individuals at time t.
- β and γ are the transmission rate and rate of recovery (removal), respectively.

III. The SEIR model

The SEIR model is an extended version of SIR model (Peng, Yang et al. 2020). It models the interaction of people between different conditions: the susceptible (S), exposed (E), infective (I), and recovered (R). The parameters S, I, and R are same as parameters in SIR model and E presents the fraction of individuals that have been infected but does not show any signs. The SEIR-model diagram is as follows (Fig. 3):

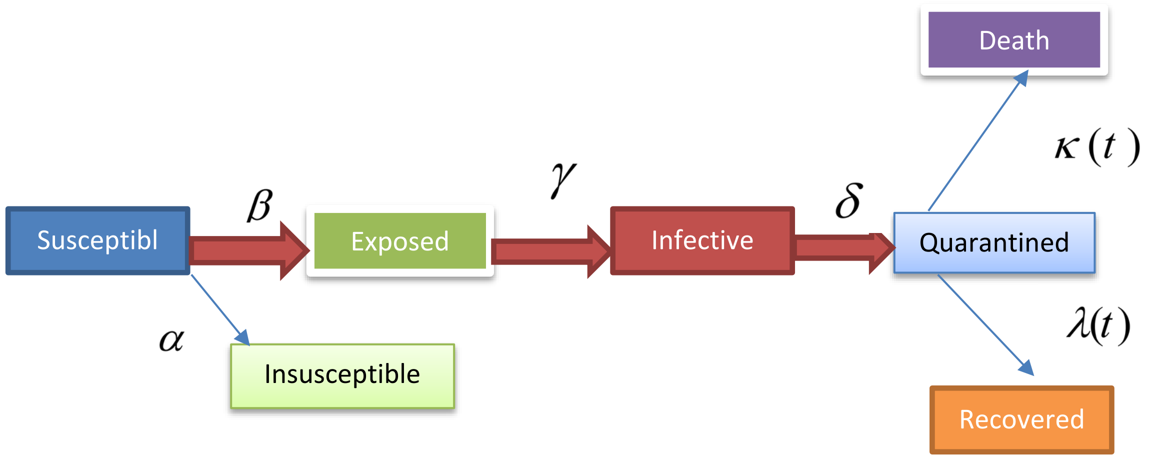


Figure 3 The SEIR diagram (Peng, Yang et al. 2020)

The equations of SEIR model are defined as follows (Eqs. 4-10):

$$\frac{dS(t)}{dt} = -\beta\frac{S(t)I(t)}{N} - \alpha S(t) \tag{4}$$

$$\frac{dE(t)}{dt} = \beta\frac{S(t)I(t)}{N} - \gamma E(t) \tag{5}$$

$$\frac{dI(t)}{dt} = \gamma E(t) - \delta I(t) \tag{6}$$

$$\frac{dQ(t)}{dt} = \delta I(t) - \lambda(t)Q(t) - \kappa(t)Q(t) \tag{7}$$

$$\frac{dR(t)}{dt} = \lambda(t)Q(t) \tag{8}$$

$$\frac{dD(t)}{dt} = \kappa(t)Q(t) \tag{9}$$

$$\frac{dP(t)}{dt} = \alpha S(t) \tag{10}$$

Where

α presents the protection rate, β shows the infection rate, illustrates the inverse of the average latent time, δ displays the inverse of the average quarantine time, λ_0 *and* λ_1 are coefficients used in the time- dependent cure rate, κ_0 *and* κ_1 are coefficients used in the time-dependent mortality rate (Peng, Yang et al. 2020).

IV. Prediction

In the present section, some machine learning techniques were used for COVID-19 case predictions in Australia, Italy, and the United Kingdom. Machine learning is a branch of computer science in which data could teach algorithms. The learning process could be done as supervised-, unsupervised, and/or semi-supervised learning forms (Mitchell 1997, Arkes 2001, Armstrong 2001, Nikolopoulos, Litsa et al. 2015, Maleki, Mahmoudi et al. 2020). In this section, some approaches that are used for prediction of cases (confirmed and deaths) of COVID-19 Pandemic are provided.

a) Logistic function

A logistic function could be defined as follows:

$$f(x) = \frac{L}{1+e^{-k(x-x_0)}} \tag{11}$$

e = Euler's number

$x_0 = Sigmoid's\ midpoint$,

L is the curve’s maximum value,

and K is the logistic growth of the curve

b) Times Series forecasting with the Prophet algorithm

The Prophet algorithm is an open-source tool developed by Facebook’ s Data Science team, and its main goal is business forecasting (Taylor and Letham 2017, Taylor and Letham 2018). The Prophet algorithm works well with time-series data that have seasonal effects and are robust in dealing with missing data (Ndiaye, Tendeng et al.

2020). In the Prophet algorithm, the forecast could be written as shown in Equation 5 (Ndiaye, Tendeng et al. 2020):

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + y_2 + \dots y_T)/T \tag{12}$$

in which $y_1, y_2, \dots, y_T$ are denoted as historical data, and $\hat{y}_{T+h|T}$ is a short-hand to forecast $y_{T+h|T}$ based on available data.

V. Results

a. Analysis

i. New cases

In this sub-section, the confirmed growth rates focusing on Australia, Italy, and the United Kingdom for every day from 2020-04-24 to 2020-05-23 were calculated. Figure 4 depicts the growth rate of confirmed cases in the countries. As can be seen in Figure 4, the growth rate for Australia was always below 0.5 during times of outbreak and just above 0.0 at the of May, while the rate for Italy and the United Kingdom is generally high. The growth rate for the United Kingdom was almost above 2.0 in April and then dramatically declined in May. The rate for Italy fluctuates between 0.5 and 1.5 in April and May.

Figure 5 also presents the growth rate of death cases for the above-mentioned countries daily from 2020-04-24 to 2020-05-23. The growth rate for death cases in Australia fluctuated between 0 and 7 in April and May and was 7 at the end of April (higher than Italy and United Kingdom during the same time), while for Italy, the rate was almost below 2.0 during the same time period and for the United Kingdom, the rate was just below 4.0 at the end of April and just above 0.0 at the end of May.

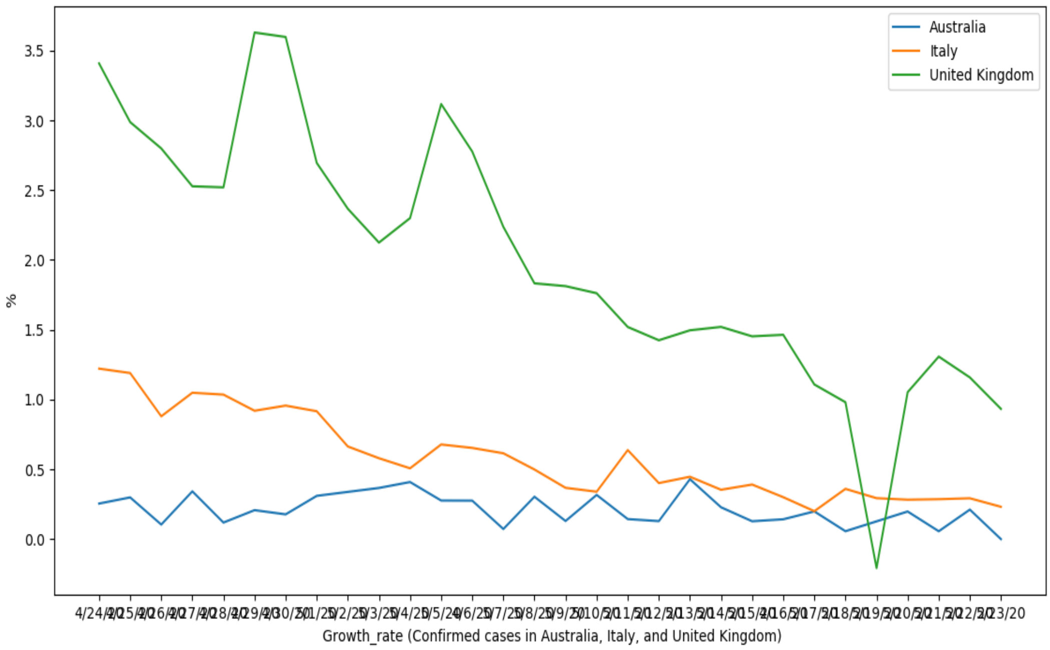


Figure 4 Growth rate (Confirmed cases in Australia, Italy, and the United Kingdom)

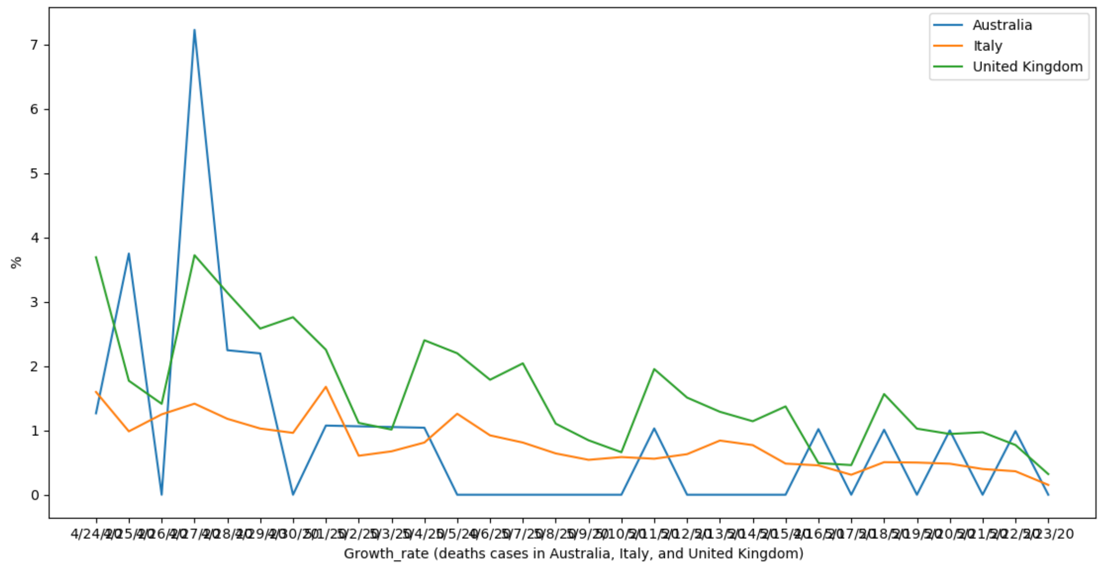


Figure 5 Growth rate (death cases in Australia, Italy, and the United Kingdom)

ii. Overall growth rate

This section shows numbers of active cases in these three countries. The active cases were calculated using the following equation:

Active_cases=confirmed_cases - deaths_cases -recovered_cases (13)

From equation (13), the overall growth rate could be calculated according to Equation 14:

Overall growth rate[i]=((active cases[i]-active case[i-1])/active case[i-1]) *100 (14)

In equation (14), the index i presents day. Figure 6 illustrates the overall growth rate for confirmed cases in the countries. Negative numbers show that people recovering are faster than those getting sick and that would be good news. The rate for Australia in the time period was almost below zero and changed from −15 at the end of April to just below −5 at the end of May and for Italy fluctuated between just above −7.5 and just above 0.0, while the rate for the United Kingdom was almost always positive number in the time horizon (00.0 and 3.0). Figure 7 illustrates the number of death cases in Australia compared with the two other countries, and it is clear that the number in Australia is significantly lower than other two.

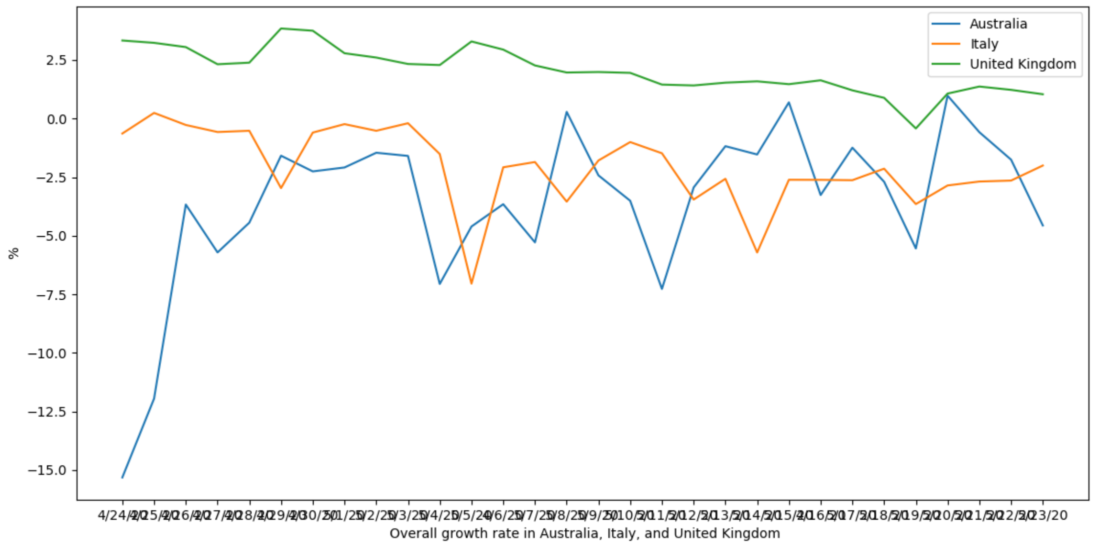


Figure 6 Overall growth rate for confirmed cases in Australia, Italy, and the United Kingdom

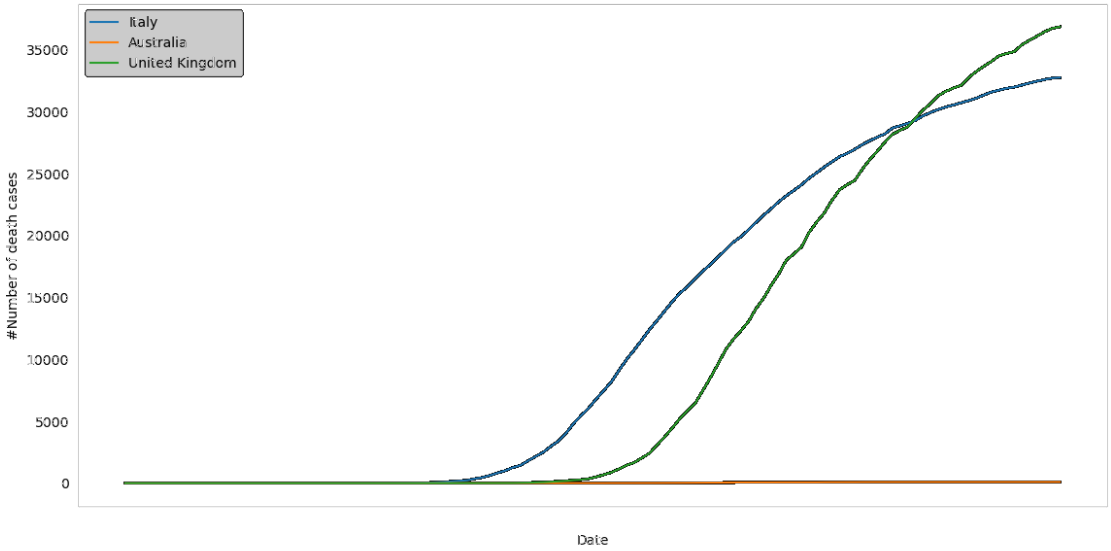


Figure 7 Number of death cases in Australia compared with Italy and the United Kingdom

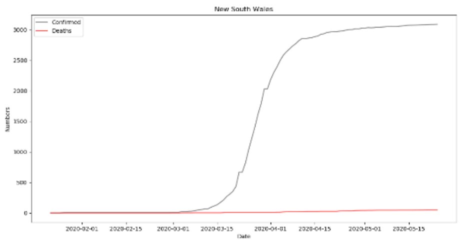


Figure 8 (a) New South Wales

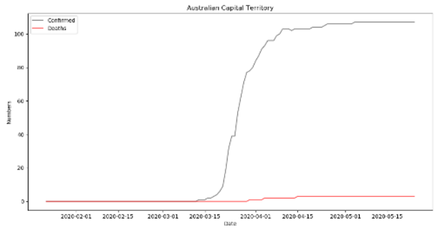


Figure 8 (b) Australian Capital Territory

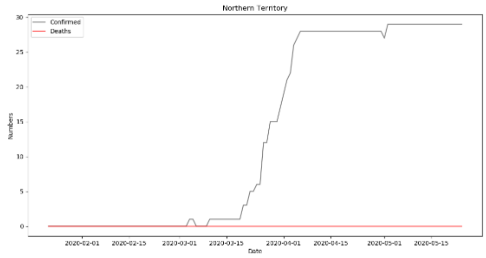


Figure 8(c) Northern Territory

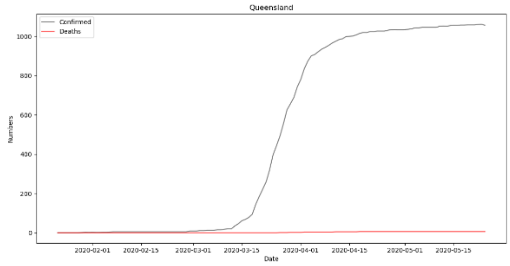


Figure 8(d) Queensland

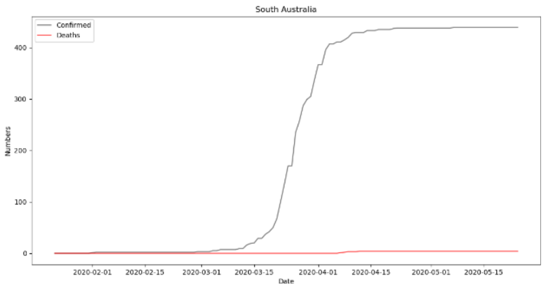


Figure 8(e) South Australia

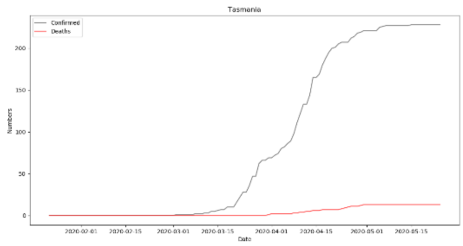


Figure 8(f) Tasmania

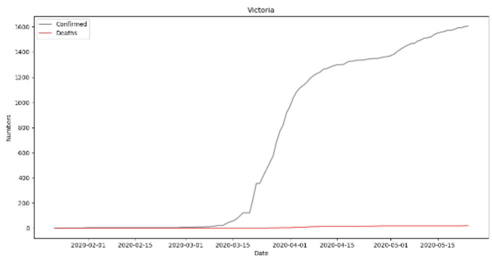


Figure 8 (g) Victoria

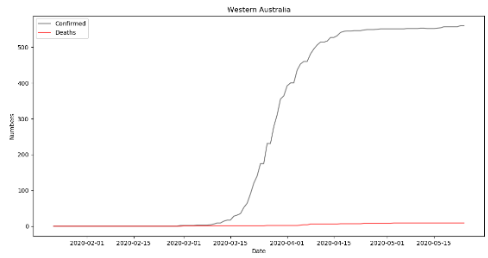


Figure 8(h) Western Australia

Figure 8 Confirmed versus death cases in different Australian states

Figure 8 (a–h) shows confirmed versus deaths cases in each individual Australian state. By now (2020-05-23), New South Wales and Northern Territory possess the most and least number of confirmed and death cases in Australia, respectively. From Figure 8(a–h), the number of confirmed and death cases in

New South Wales significantly differed from other states in Australia and increased dramatically, while the Northern Territory experienced some fluctuation during the study time period.

With the aim of forecasting, the logistic function is defined in Equation (11) and was applied to collected data (Time horizon: start of outbreak in the countries) and results have been illustrated in Figures 9-14. As it is shown in Figures 9-14, the logistic function is fitted until the trend of cases is increases and to evaluate the performance of metric R2 scores used for confirmed and death cases. Results are presented in Table 2. Another metric that has been used in experiments is the root mean square error (RMSE), and the results of RMSE I depicted in Table 2. The best RMSE value belongs to the Australian cases (confirmed and deaths).

Table 1 R2 score fore different countries, different cases

countries	Confirmed cases	Deaths cases
Australia	0.87	0.67
United Kingdom	0.92	0.97
Italy	0.93	0.95

Table 2 Root mean square error (RMSE) values for different countries and different cases

countries	Confirmed cases	Deaths cases
Australia	8.22	0.88
United Kingdom	21.94	6.97
Italy	23.24	8.00

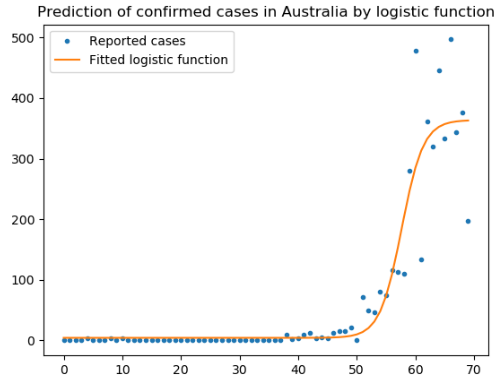


Figure 9 Prediction of confirmed cases by logstic function (Australia)

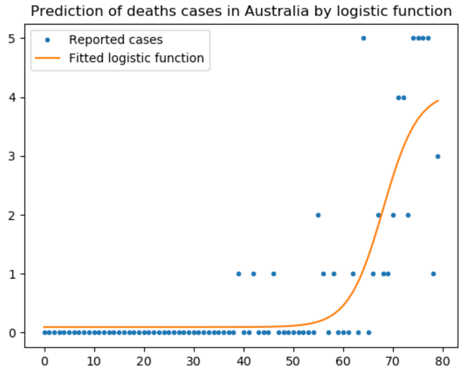


Figure 10 Prediction of death cases by logistic function (Australia)

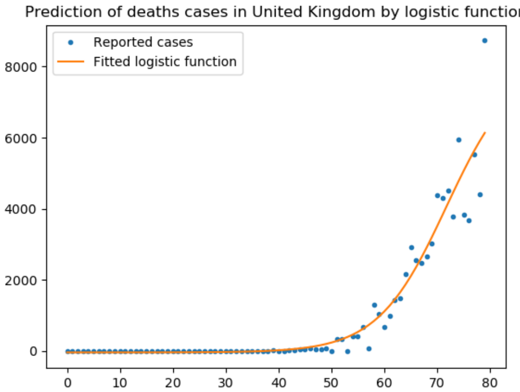


Figure 11 Prediction of confirmed cases by logistic function (United Kingdom)

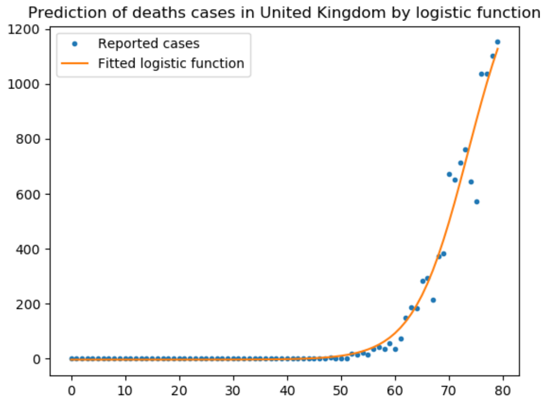


Figure 12 Prediction of deaths cases by logstic function (United Kingdom)

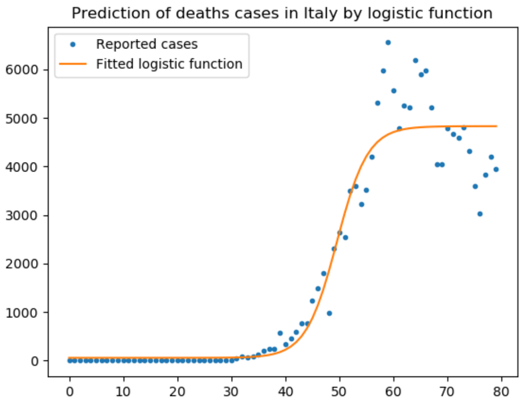


Figure 13 Prediction of confirmed cases by logistic function (Italy)

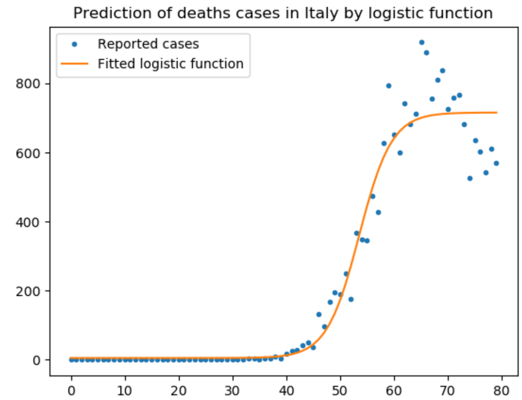


Figure 14 Prediction of deaths cases by logstic function (Italy)

Figures 15–17 present the results of the classic SIR model. As previously mentioned, controlling β parameters indicate the level of disease transmission, and γ is the recovery (removal) period indicating how much peope could recover in a period. First, all parameters were initially added to the SIR model and applied it to real data, but it can be seen from Figures 15–17 and Table 3 (RMSE values) in which the classic form was not suitable for prediction of the COVID-19 pandemic in these three countries. In order to fit the SIR models to Australia, Italy, and the United Kingdom, an optimizer was needed to find the unknown parameters (β and γ) from equation R_0 ($R_0 = \frac{\beta}{\gamma}$) since these parameters could be estimated. Before the start of the outbreak, it is essential to address whether the number of susceptible cases is equal to the number of people in these countries because no antibodies exist, and no vaccines for the disease have been developed. At first, R_0=2.7 was fixed (reported by Australian Government: Department of Health) as the the median number, $\beta = 0.378$, and $\gamma = 0.14$. Figure 19 (a-c) present the confirmed cases provided by the optimized SEIR model with the above-mentioned decriptions in the three countries (See Figure 18).

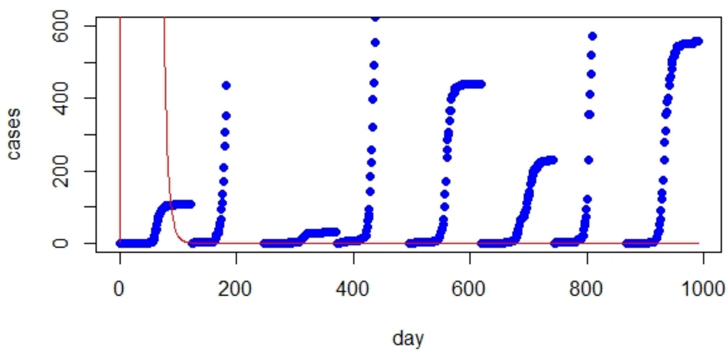


Figure 15 Predicted cases in Australia using the susceptible, infected, recovered (SIR) model (blue: real confirmed cases, red: SIR model)

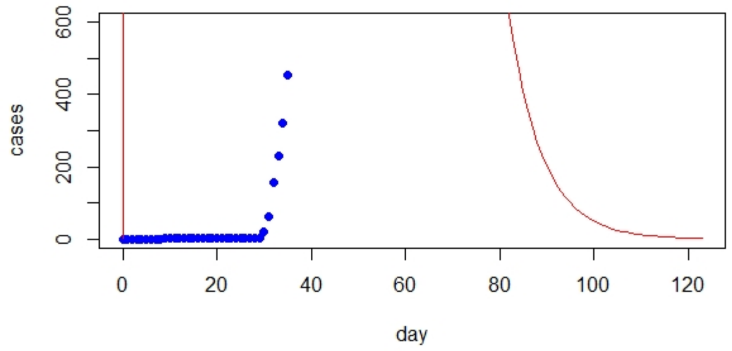


Figure 16 Predicted cases in Italy based on the SIR model (blue: real confirmed cases, red: SIR model)

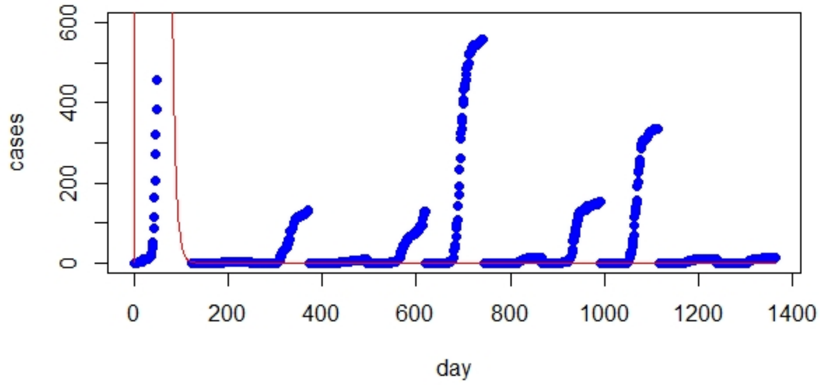


Figure 17 Predicted cases in UK based on the SIR model (blue: real confirmed cases, red: SIR model)

Table 3 RMSE values obtained by SIR model (before optimization of parameters)

Italy	**United Kingdom**	**Australia**
18.75	15.45	831.84

Real data were used to estimate the values of β and γ. An optimizer was used to find the best estimation of β and γ. The optimization algorithms were the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (Fletcher 1987), L-BFGS-B (Byrd, Lu et al. 1995), conjugate gradients (CG), (Fletcher and Reeves 1964), and Nelder-Mead (Nelder and Mead 1965). The parameter settings are provided in Table 3. The flowchart of the improved SIR and SEIR versions and parameter settings for the above-mentioned algorithms are addressed in Figure 18 and Table 4, respectively.

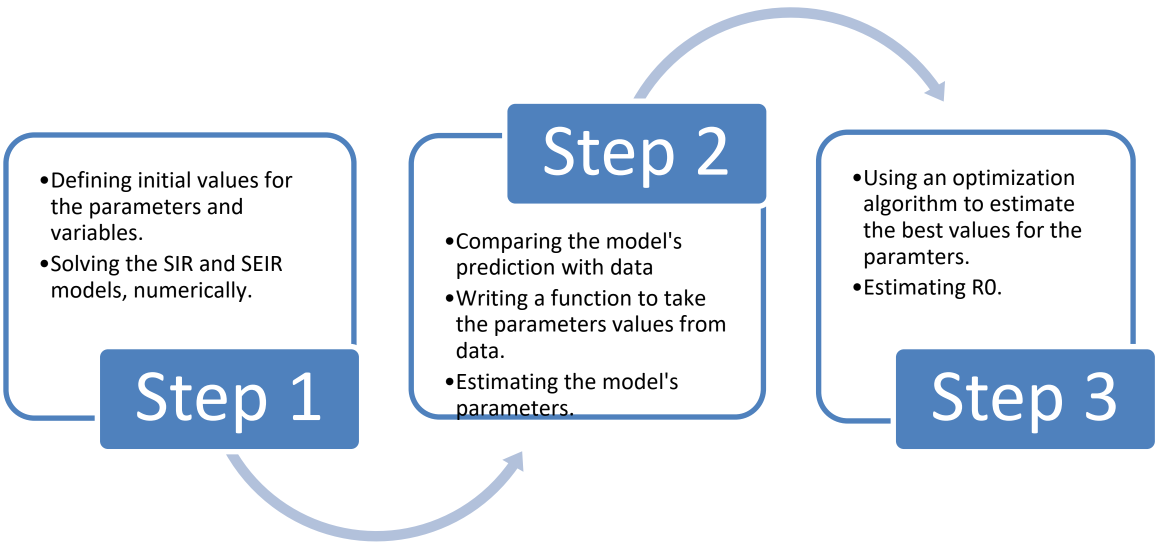


Figure 18 Flowchart of improved version of SIR and SEIR models

Table 4 Parameter settings

Algorithm	**Parameters setting**
BFGS	Maxit=100, reltol*=$1e^{-8}$
Nelder-Mead	Maxit=500, reltol=$1e^{-8}$, alpha=1, beta=0.5, gamma=2.0
L-BFGS-B	Maxit=100, reltol=$1e^{-8}$, lmm**=5, factr***=$1e^{7}$
CG	Maxit=100, reltol=$1e^{-8}$

*Reltol= Relative convergence tolerance, **lmm= number of BFGS updates retained, ***factr=convergence factor

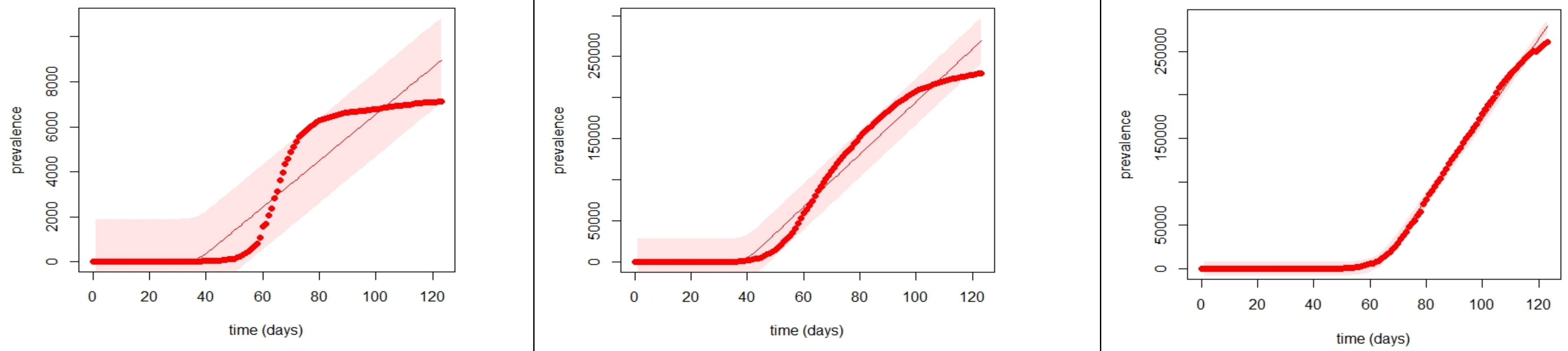
Table 5 shows the optimized values obtained by different algorithms (SIR model). The best values for the parameters were found using the Nelder—Mead algorithm (for SIR model) and L-BFGS-B algorithm (for SEIR model). This method is illustrated in Figure 18. As was mentioned earlier, before the start of the outbreak, the number of susceptible cases was equal to the number of people in these countries because no antibodies exist, and no vaccine for the disease is available. From Wikipedia, the populations of Australia, Italy, and the United Kingdom are 25^{06}, 60^{06}, and 67^{06}, respectively. Table 6 illustrates the RMSE values obtained by the algorithms (for SIR and SEIR models) showing improvements in significantly reducing the values.

Table 5 Median values of SIR parameters determined by the departments of health in each country

COUNTRY	β				γ				R_0			
Algorithm	BFGS	Nelder-Mead	L-BFGS-B	CG	BFGS	Nelder-Mead	L-BFGS-B	CG	BFGS	Nelder-Mead	L-BFGS-B	CG
Australia	0.014	0.014	0.378	0.37	0.22	0.22	0.14	0.14	0.063	0.063	2.64	2.64
United Kingdom	0.37	3.84701^{-3}	0.37	0.37	0.14	1.94^{-1}	0.14	0.14	2.64	0.02	2.64	2.64
Italy	0.37	1.083555^{-3}	0.37	0.37	0.14	3.9088^{-1}	0.14	0.37	2.64	0.01	2.64	2.64

Table 6 RMSE values obtained based on the improved SIR model considering a 0.99 confidence interval

Model	**Italy**	**United Kingdom**	**Australia**
SIR model	1.41	1.01	1.13
SEIR model	1.12	1.23	1.04



(a) Australia (b) Italy (c) UK

Figure 19 Prediction done by optimized SEIR model

Table 7 Predicted cumulative confirmed cases in Australia (cross-validation matrix)

y	**ds**	$\hat{y}$	$\hat{y}_{lower}$	$\hat{y}_{upper}$	**cutoff**
7095	2020-05-21	21309.752	18998.140	23829.955	2020-04-04
7099	2020-05-22	21630.708	19245.072	24269.904	2020-04-04
7114	2020-05-23	21959.985	19424.097	24640.939	2020-04-04
7114	2020-05-24	22326.688	19766.194	25093.353	2020-04-04

Table 8 Predicted cumulative confirmed cases in the United Kingdom (cross-validation matrix)

y	**ds**	$\hat{y}$	$\hat{y}_{lower}$	$\hat{y}_{upper}$	**cutoff**
252246	2020-05-21	143776.53	126702.28	162413.93	2020-04-04
255544	2020-05-22	146462.83	128526.68	165539.80	2020-04-04
258504	2020-05-23	148818.88	130813.85	168216.41	2020-04-04
260916	2020-05-24	150344.39	131476.87	170004.00	2020-04-04

Tables 7–9 present the results of the predicted cumulative confirmed cases obtained using the Prophet algorithm in the three countries. In the presented tables, **y** represents the true values of confirmed cases, **ds** is time, $\hat{y}$ is the forecasted values, $\hat{y}_{lower}$ and $\hat{y}_{upper}$ are the lower and upper bounds for the forecasted values, respectively. It should be noted , the forecasted values were made between the cutoff and cutoff + horizon. Tables 7–9 are also called cross-validation matrices that are used to find the error values between **y** and $\hat{y}$ after which the RMSE values can be obtained (Figure 23 a–c). Figures 20–22 visualize forcasted values obtained using the Prophet algorithm, indicating the mentioned algorithm is fitted for the cases of Italy and the United Kingdom but with errors for Australia.

Table 9 Predicted cumulative confirmed cases in Italy (cross-validation matrix)

y	**ds**	$\hat{y}$	$\hat{y}_{lower}$	$\hat{y}_{upper}$	**cutoff**
228006	2020-05-21	373982.5	336940.1	415612.7	2020-04-04
228658	2020-05-22	379300.7	340862.6	422338.4	2020-04-04
229327	2020-05-23	384792.4	344957.8	429120.3	2020-04-04
229858	2020-05-24	390481.8	349482.8	436663.2	2020-04-04

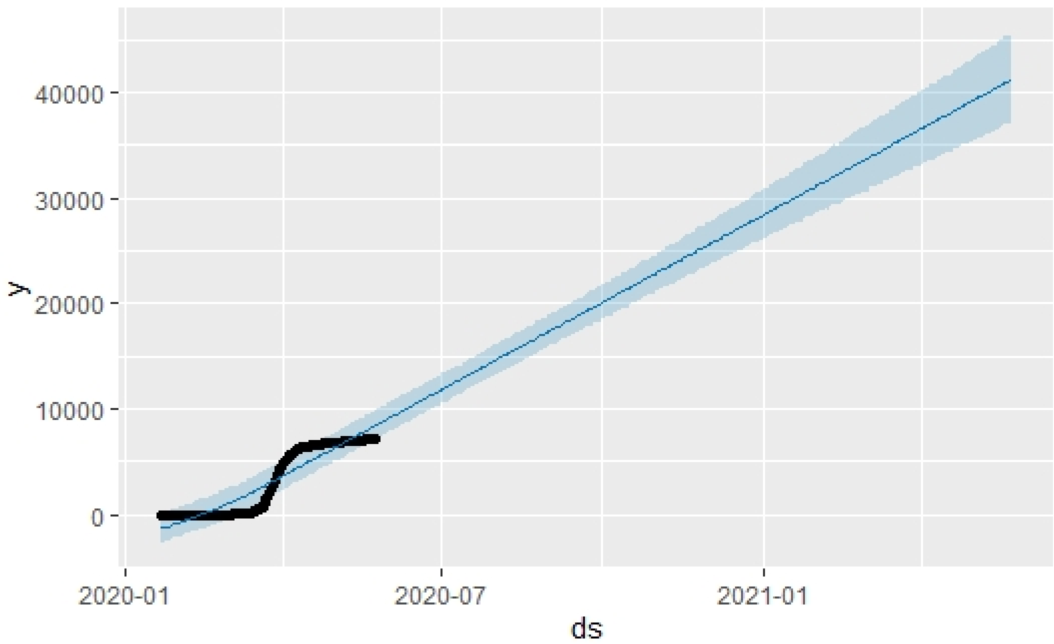


Figure 20 Forcasting by Prophet for the next year (Confirmed cases in Australia)

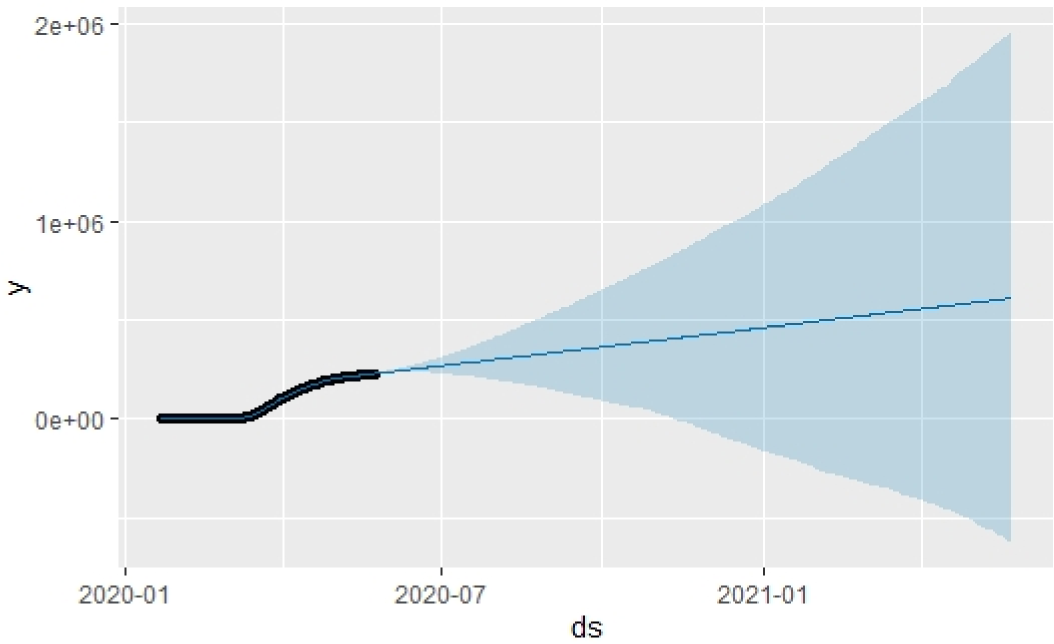


Figure 21 Forcasting by Prophet by the next year (Confirmed cases in Italy)

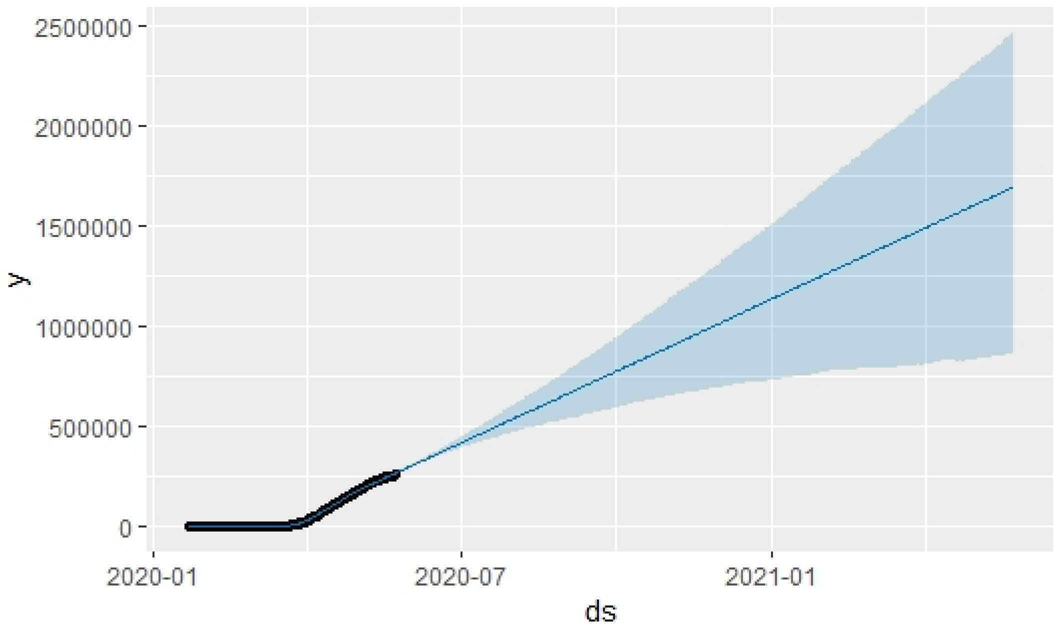
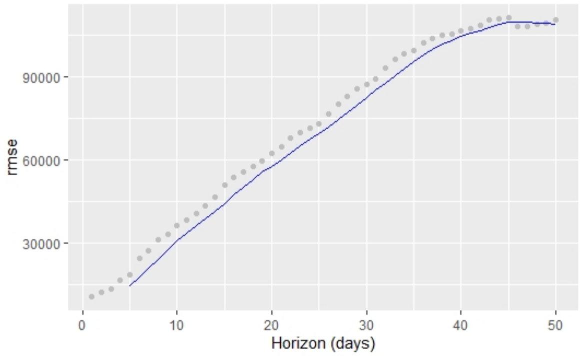
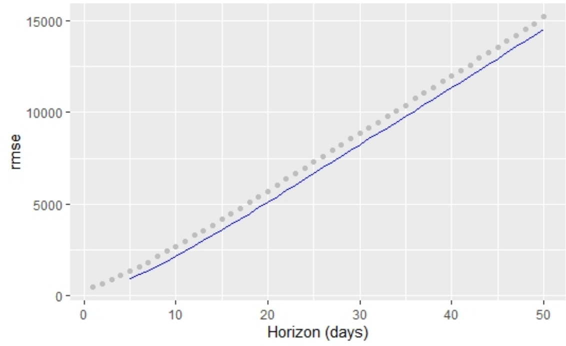


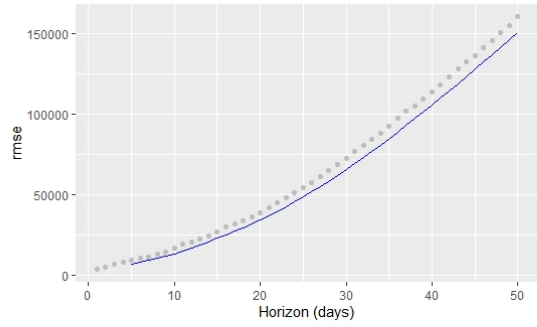
Figure 22 Forcasting by Prophet for the next year (Confirmed cases in United Kingdom)



(a) United Kingdom



(b) Australia



(C) Italy

Figure 23 Visualization of performance metric for Prophet for the countries (considering RMSE)

VI. Conclusion and discussion

COVID-19 is a family of Coronaviruses that has affected the life of billions of people worldwide. The first phase of the paper started with a short analysis of COVID-19, focusing on Australia, Italy, and the United Kingdom. The analysis presents confirmed and death growth rates in Australia, a comparison between Australia, Italy, and the United Kingdom, and also, a short analysis in different states of Australia. The analysis shows that generally Australia is in a good position compared with two other countries. However, the situation in different cities of Australia are completely complicated; for example, New South Wales has the most confirmed and deaths cases, while Northern Territory shows the least confirmed and death cases (it is valuable to mention that New South Wales has more population).

Mathematical approaches based on SIR and SEIR were proposed to predict the epidemiology in Australia, Italy, and the United Kingdom. Since the classic form of SIR and SEIR are deterministic, an improved version based on parameter optimization was suggested to improve the prediction. The results are compared with logistic function and Prophet algorithm and summarized as follows:

- Comparison between the classic form of SIR model with real data showed a significant gap. However, initializing the parameters of the SIR model significantly improved the prediction.
- The classic form of SIR model worked better for the United Kingdom, while the SIR model was not suitable for Australia case (regarding RMSE values).
- The logistic function was a good model for the United Kingdom with an r2_score of 0.97, while this score for Australia was 0.67 and Italy was 0.95.
- The best RMSE value belonged to the Australia cases (confirmed and deaths).
- Optimization of parameters of the SIR and SEIR models significantly improved the prediction accuracy of the models.

- Improved version of SEIR has better performance compared with SIR model (Regarding RMSE values and Figures).
- Optimized SEIR model has better prediction for UK and Italy compared with Australia.
- The best values for the parameters were found using the Nelder—Mead algorithm for SIR model and L-BFGS-B algorithm for SEIR model.
- The Prophet algorithm worked better for Italy and the United Kingdom cases than for Australian cases.
- Logistic function compared with Prophet algorithm had a better performance in these cases.
- The improved version of the SIR and SEIR model had a better performance compared with logistic function, Prophet algorithm, and classic form of SIR model.

In this paper, all forecasting was addressed without considering of scenario of social distancing and quarantine that makes it valuable as a future direction. This paper presents SIR and SEIR as epidemiology models; it would interesting to test other epidemiology models. Moreover, it is worthwhile to combine the mathematical model with other observations such as Policy intervention, human behavior, and constraints.

Compliance with Ethical Standards:

- Sources of Funding: The authors confirm that there is no source of funding for this study.

- Conflict of Interest: The authors declare that they have no conflict of interest.
- Human Participants and/or Animals: None.

References

Arkes, H. R. (2001). Overconfidence in judgmental forecasting. Principles of forecasting, Springer**:** 495-515.

Armstrong, J. S. (2001). Standards and practices for forecasting. Principles of Forecasting, Springer**:** 679-732.

Byrd, R. H., P. Lu, J. Nocedal and C. Zhu (1995). "A limited memory algorithm for bound constrained optimization." SIAM Journal on scientific computing **16**(5): 1190-1208.

Capasso, V. and G. Serio (1978). "A generalization of the Kermack-McKendrick deterministic epidemic model." Mathematical Biosciences **42**(1-2): 43-61.

Fletcher, R. (1987). "Practical methods of optimization. 1987." John and Sons, Chichester.

Fletcher, R. and C. M. Reeves (1964). "Function minimization by conjugate gradients." The computer journal **7**(2): 149-154.

Kermack, W. O. and A. G. McKendrick (1932). "Contributions to the mathematical theory of epidemics. II.—The problem of endemicity." Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character **138**(834): 55-83.

Maleki, M., M. R. Mahmoudi, D. Wraith, K.-H. J. T. M. Pho and I. Disease (2020). "Time series modelling to forecast the confirmed and recovered cases of COVID-19." 101742.

Mbuvha, R. R. and T. Marwala (2020). "On Data-Driven Management of the COVID-19 Outbreak in South Africa." medRxiv.

Mitchell, T. M. (1997). Machine learning, McGraw-hill New York.

Ndiaye, B. M., L. Tendeng and D. Seck (2020). "Analysis of the COVID-19 pandemic by SIR model and machine learning technics for forecasting." arXiv preprint arXiv:2004.01574.

Nelder, J. A. and R. Mead (1965). "A simplex method for function minimization." The computer journal **7**(4): 308-313.

Nikolopoulos, K., A. Litsa, F. Petropoulos, V. Bougioukos and M. Khammash (2015). "Relative performance of methods for forecasting special events." Journal of Business Research **68**(8): 1785-1791.

Peng, L., W. Yang, D. Zhang, C. Zhuge and L. Hong (2020). "Epidemic analysis of COVID-19 in China by dynamical modeling." arXiv preprint arXiv:2002.06563.

Putra, S. and Z. Khozin Mu'tamar (2019). "Estimation of Parameters in the SIR Epidemic Model Using Particle Swarm Optimization." American Journal of Mathematical and Computer Modelling **4**(4): 83-93.

Qi, H., S. Xiao, R. Shi, M. P. Ward, Y. Chen, W. Tu, Q. Su, W. Wang, X. Wang and Z. J. S. o. t. T. E. Zhang (2020). "COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis." 138778.

Salgotra, R., M. Gandomi and A. H. Gandomi (2020). "Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming." Chaos, Solitons & Fractals: 109945.

Taylor, S. and B. Letham (2017). "prophet: Automatic forecasting procedure." R package version 0.2 **1**.

Taylor, S. J. and B. Letham (2018). "Forecasting at scale." The American Statistician **72**(1): 37-45.

Weiss, H. H. (2013). "The SIR model and the foundations of public health." Materials matematics: 0001-0017.