Survey on Forecasting the vulnerability of Covid 19 in Tamil Nadu

P.Ananthi1, S.Jabeen Begum2, V.Latha Jothi3, S.Kayalvili4 and S.Gokulraj5

1M.E student, Computer Science and Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India, ananthi2698p@gmail.com

2Professor, Computer Science and Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India, shabeenbegum@gmail.com

3Professor, Computer Science and Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India, lathajothi.s@gmail.com

4Associate Professor, Computer Science and Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India, kayalvilis@gmail.com

5Associate Professor, Computer Science and Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India, gokulrajs1@gmail.com

Abstract. Predictive and analytic models for forecasting the vulnerability and recovery rate of patients who are affected by COVID 19 are made in this project for good analysis and better decision-making. In this project, linear regression (LR) a Machine Learning model is used to forecast the number of patients will get the infection in near future. By simulating SIRD model, the infection spread and recovery rate of the disease in a geographic region can be predicted. The vulnerability of the disease is checked by observing the transmission of disease over a period. In addition to this many infographic models and graphs are created for easy understanding of data to get more insights about the disease. However, these prediction models enable us to make quick response of pandemic and to bring a conclusion to the disease.

INDEX TERMS: Covid 19, data science, machine learning, prediction, analysis, pandemic, recovery and infection.

1. Introduction
At the end of 2019, from the family of Severe Acute Respiratory Syndrome (SARS), a new strain of corona virus called “Corona virus Disease 2019” was identified in Wuhan a province in China caused a pandemic worldwide. World Health Organization (WHO) on 11th of March 2020 declares this extremely contagious disease as a pandemic disease. Whereas, considering the extent of its spread worldwide, Governments of different countries have imposed border restrictions, aviation restrictions, increasing awareness of hygiene, and social distancing. However, terribly at a very rapid rate the virus remains spreading. While, whereas, most of the individuals infected with the COVID-19 seasoned delicate to moderate respiratory disease, some developed a deadly respiratory illness called pneumonia. There are observations that aged individuals with underlying medical issues like diabetes, cardiovascular disease, chronic respiratory disease, renal, or hepatic diseases and cancer are more seemingly to develop serious unhealthiness. Until now, no specific immunization agent, vaccine or treatment for COVID-19 has been invented. Every patient is treated symptomatically. However, there are many progressing clinical trials evaluating potential treatments. At the time of writing, nearly 33.4 million infected cases were found in more than 215 countries until September 2020, among which
around 1 million deaths, 23.2 million recoveries, 6.8 million mild cases and 60 thousand critical cases were reported.
As a contribution to the current human crisis, this project is an attempt to develop forecasting system for transmission rate, recovery rate and demises rate for Covid 19. The forecasting is done by using mainly four variables: 1) Number of infected cases 2) Number of death cases 3) Number of recovered cases 4) Number of active cases. The forecasting is made as a regression model, for this purpose LR model is used. The SRIF model is used to predict the vulnerability and recovery rate for the infected individuals. This model is intertwined with the dynamic relationships described by nonlinear ordinary differential equations.

2. Literature Survey

2.1. Covid-19 Future Forecasting using Supervised Machine Learning Models
Machine learning (ML) helps to preoperative outcomes to improve the decision making on the future course of actions based forecasting mechanisms have proved their significance anticipation. The forecasting of the number of upcoming patients affected by COVID-19 is done by ML models. Standard forecasting models like least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), linear regression (LR), and exponential smoothing (ES) have been used to forecast the threatening factors of COVID-19. The ES performs best in the current forecasting domain from the results of the study prove that given the nature and size of the dataset. LR and LASSO also perform well for forecasting to some extent to predict death rate and confirm cases. [1]

2.2. Analyzing Recovery from Pandemics by Learning Theory: The Case Of Covid-19
Predicting the recovery time from infectious diseases outbreaks is made in this study. The approach of theory of learning from errors is used here, by imposing countermeasures such as medical treatment, isolation, social distancing etc. specifically to adapt the control of the virus spread by reducing infection rates when these are effective, the infection rate, after reaching a peak, declines following what is the Universal Recovery Curve. The risk of an infectious disease is less when the infection rate slows down, called ‘flattening the curve’. Once the rate peaks the rate, then, should decrease due to successful countermeasures. [2]

2.3. Data Mining and Analysis of Scientific Research Data Records on Covid-19 Mortality, Immunity, and Vaccine Development - In the First Wave of the Covid-19 Pandemic
The Covid-19 mortality data, immunity details, and vaccine are obtained through data mining of scientific literature records from the Web of Science Core Collection, using. The analysis is compared with records on all Covid-19 research topics combined and the individual records are analyzed in isolation. Including the Web of Science and R Studio’s Bibliometrix package, data mining tool the data records are analyzed with commutable statistical methods. From historical analysis of scientific data records on viruses, pandemics and mortality, it is identified that Chinese universities have not been leading on these topics historically. We identified a few clusters, containing references to exercise, inflammation, smoking, obesity and many additional factors. [3]

2.4. Epidemic Simulation of H1N1 Influenza Virus Using GIS in South Korea
The global spread of diseases continues to accelerate together with the development of transportation technology. In this paper based on the GIS the examination of the correlation between transportation and infection, the statistical analysis on the highway transport and the domestic flight. By performing the SEIR simulation using logic, the regional infection can be predicted when outbreaks occur to make quick responses of threats of diseases. Through simulation, it can be ascertained that regions which have the large population also have the largest infection. In conclusion, regions which have large inflows rather than outflows have lots of infected people and the infectivity in some regions with poor transportation is relatively low. [4]
2.5. Evaluating Strategies for Pandemic Response in Delhi using Realistic Social Networks
Influenza pandemic in the National Capital Territory of India (NCT-I, including New Delhi and its surrounding areas), analyze on the targeted layered containment strategies are made. A realistic individual-based social contact network for NCT-I is a key contribution of our work is to synthesize by using a wide variety of open source and commercial data. How an influenza-like illness (ILI) is computed by using modeling environment would spread over the NCT-I network. The pharmaceutical and non-pharmaceutical containment strategies are analyzed to control a pandemic outbreak. In using such a network and the powerful epidemic simulation platform EpiFast, a study of various intervention strategies to contain the spread of disease in the city is made. [5]

2.6. Predictive Modeling of Covid-19 Data in US (Adaptive Phase-Space Approach)
SIR has fractions of “Susceptible”, “Infectious” and “Recovered/ Removed” framework that divides a population and defines with using first-order differential equations their dynamic interrelationships. The gradual reduction of infectivity rate is obtained by parameters estimated, although the latest wave is expected to be the largest. Analysis of data from individual states or countries may quantify the distinct effects of different mitigation measures. [6]

2.7. Forecasting the Novel Coronavirus Covid-19
Forecasting requires ample historical data. No prediction is certain as the future. The reliability of the data, variables are being predicted will influence the forecast. Medical predictions are rarely inaccurate, whereas their uncertainty is a serious issue. Predicting and forecasting the pandemics and epidemics future is tedious because the number of cases is huge for study. Here the forecast of confirmed cases of COVID-19 is taken for simple time series. The forecasts are produced using models from the exponential smoothing family. This gives good forecast accuracy among many forecast models especially suitable for short series. Exponential smoothing models have multiplicative or additive patterns. [7]

2.8. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis
To classify data, classification and data mining methods are an effective way. In medical field, they are used in diagnosis and analysis to make decisions. A performance comparison between different machine learning algorithms: Naive Bayes (NB), k Nearest Neighbors (k-NN) Support Vector Machine (SVM) and Decision Tree (C4.5) on the Wisconsin Breast Cancer (original) datasets is conducted. Classifying data with respect to efficiency and effectiveness of each algorithm in terms of accuracy is the main objective, also the precision, sensitivity and specificity. SVM gives the highest accuracy (97.13%) by observing the experimental results show that with lowest error rate. [8]

2.9. Statistical and Machine Learning Forecasting Method (Concerns and Ways Forward)
Machine Learning (ML) methods are used to evaluate performance of various time series in large datasets. The assumptions are made by splitting the data for multiple time series for a huge data set. Machine learning methods with that of eight statistical models are used to compare the post-sample accuracy; it is found that the former models are dominated in all aspects. Moreover, when compared to the statistical model computational requirements are greater in this model. The accuracy of statistical models is greater than of ML models. The forecasting methods produce empirical results to test the performance of that allows definite conclusions and meaningful comparisons. [9]

2.10. SEIR Modeling of the Covid-19 and its Dynamics
Covid 19 is analyzed by SEIR epidemic model for general data and other control strategies such as quarantine, external input and hospitals. Hubei province data is used for particle swarm optimization (PSO) algorithm is applied for estimating the parameters. In this model new parameters such as seasonality and stochastic infection are used. The parameters affect the dynamics of the system, and it is different for different parameters. [10]
3. Proposed Method
In this project, Covid 19 data is analyzed by using statistics and machine learning techniques for the purpose of drawing conclusion about information. The raw data is collected from many resources and then transformed in a way which is appropriate to make models.

It is mainly used for the purpose of better decision making and testing models or theories. The raw requires process of inspecting, cleaning, transforming, modeling, analyzing and interpreting raw data. The main aspects of data preparation are 1) Domain knowledge 2) Statistics and Machine Learning 3) Software. Big data is the pillar behind the idea that one can make useful inference with the large body of data that was not possible with smaller dataset. Extremely large dataset may be analyzed computationally to reveal patterns, trends and associations that are not transparent or easy to identify.

3.1 Dataset
The dataset used for this study has been obtained from official press release of Tamil Nadu State Government through www.stopcoronatn.com website. As the aim of the study is to predict and forecast the cases. All the data of reports has been daily updated for accuracy.

| Date    | Total Deaths | Total Recoveries | Total Confirmed Cases | Active Cases | New Confirmed Cases | New Deaths |
|---------|--------------|------------------|-----------------------|--------------|---------------------|------------|
| 3/7/20  | 0            | 0                | 1                     | 1            | 1                   | 0          |
| 3/8/20  | 0            | 0                | 1                     | 1            | 0                   | 0          |
| 3/9/20  | 0            | 0                | 1                     | 1            | 0                   | 0          |
| 9/20/20 | 8811         | 486479           | 541993                | 46703        | 5516                | 60         |
| 9/21/20 | 8871         | 491971           | 547337                | 46495        | 5344                | 60         |

Table 2 data set has the number of new cases reported daily in all districts of Tamil Nadu. Tamil Nadu has 38 districts and in our case Nagapattinam district includes the Mayiladudurai district information. So the dataset has the information for 37 districts for dates from 7 March to 21 September.

| Date     | Ariyalur | Chengalpattu | Chennai | …………. | Vellore | Villupuram | Virudhunagar |
|----------|----------|--------------|---------|----------|---------|------------|--------------|
| 3/7/20   | 0        | 0            | 0       | …………. | 0       | 0          | 0            |
| 3/8/20   | 0        | 0            | 0       | …………. | 0       | 0          | 0            |
| 9/20/20  | 3487     | 32580        | 155639  | …………. | 13516   | 10394      | 14031        |
| 9/21/20  | 3517     | 32799        | 156625  | …………. | 13651   | 10525      | 14066        |

Table 3 represents the case details of August month with the attributes of total active, discharge, death and confirmed cases reported along with gender and lab count. Table 4 is the dataset of Erode district with the area, subarea and count details.
3.2 Regression Model
3.2.1 Linear Regression: In this model forecasting is made by finding the linear relationship between independent and dependent variables. For a set of input values \( x \), the solution is the prediction of set of values \( y \). Thus regression is how \( y \) is related to \( x \).

\[
y = \beta_0 + \beta_1 * x + \varepsilon
\]

Where \( \varepsilon \) is the error in linear regression which can be equivalent to,

\[
E(y) = \beta_0 + \beta_1 * x
\]

3.2.2 SIRD Model: The SIRD model with standard incidence which contains 4 subpopulations:

- Susceptible(S)
- Infective(I)
- Recovered(R)
- Demised(D)

The susceptible can become infective and the infective can become recovered or Demised, but no other transitions are considered. \( N \) is the total population and it is given by,

\[
N = S + I + R + D
\]

where \( N \) always remains constant

\[
\frac{dS}{dt} = -\beta_0 I
\]

\[
\frac{dI}{dt} = \beta_0 I \frac{S}{N} - (\gamma + \delta) I
\]

\[
\frac{dR}{dt} = \gamma I
\]

\[
\frac{dD}{dt} = \delta I
\]

Then the total new infective infected by all individuals in the infected compartment per unit of time, at time \( t \) is \((\beta_0 * T(N) \frac{S}{N}) * I\), which is called incidence of disease. The movement between the classes by the system of differential equations is described by this model. Where \( \beta_0 \) is the transmission rate, \( \gamma \) is the recovery rate, \( \delta \) is demised rate and \( R_0 = \beta_0 / (N+S) \).

3.3 Data Collection and Interpretation
In this current crisis, data collection is difficult at best of its time to bring insights from the collected data. The collected big data is transformed to a system that has the ability to store and work with large amount of data. In Covid 19 situation, the system must be robust, local, and timely are thus critically needed. To design a model it is need to gain knowledge from domains for analysis of high-resolution data and real-time data acquisition. In corona virus, the majority of symptoms are mild that are treated. Reporting focuses on mortality and morbidity, the count of health facilities for testing, or care is easy to calculate. The data sampling result gives idea about the increasing number of cases, while at the same understanding the model gives idea about the spread and distribution of cases. Collecting
accurate data and understanding the limitations of the data that have already been collected is an essential part of understanding the situation. Without good data, policymakers can’t make good decisions.

3.4 Data Modeling and Prediction
Covid 19 data of Tamil Nadu has been collected. Based on the data, descriptive analytics is made for describing the data looks when having a large data. It also describes where the center of the data and how it is distributed. Predictive analytic is used to make predictions about what is going to happen next for a new scenario. It usually describes the relationship between datasets. In our situation, it is needed to not only develop the models, but also to determine in which ways they’re wrong and which ways they’re useful. The decisions that are made to combat the spread of this pandemic with the help of results from the models, along with data. While handling large set of data statistics help to obtain the relationship between the data and its probability. Basically it brings insight in the form of number from data. It is carried out in both population, and sample statistics as well. Bigger is the sample higher would be the approximation.

3.5 Data Visualization and Communication
Info graphics and data visualizations are a useful and helpful way of putting risks and raw numbers into perspective. In our system information into a context that is visually appealing, easy to understand in accurate and useful manner. Visualizations make the information easy to share and have the potential to incite fear and alarm just as much as they have the potential to inform.

4. Results and Discussion
The study attempts to forecast the number of cases will affect in future. The dataset used in this study consists of number of newly infected case, recovered and demised in Tamil Nadu. The datasets are continuously updated every day from 7 March 2020 to 21 September 2020.
From the plot 2, the total death, recovered, confirmed and active cases were reported in timeline. The vertical lines depict the dates of lock downs. The first two solid lines represent the entire shutdown to the state with border restriction and strict rules. The next four dashed lines represent the extended lock with some relaxations and it was not apply to containment zones. The last three dotted lines represent the extension of lockdown without district restrictions and inter-state movement. It is clearly observed from the plot that the confirmed cases continuous to grow day by day. In addition to this the number of recoveries from the disease is also increasing exponentially.

The number active is quite large in the month of July but in turn it drops in few days. The number of death cases remains flat in comparison with other parameters.
From the figure 3, all the parameters are in less count in March month. In August month adversely all parameters got spike following from July month spike, but in September parameters are declining gradually. The percentage gives an idea about how much the month contributes to its total sum.
Moving average is used to analyze the data points by making averages for multiple subsets for the entire dataset in a series. Here the moving average is calculated for a week i.e. for 7 days.

\[ MA = \frac{x_n + x_{n-1} + \cdots + x_{n-(n-1)}}{N}, \quad N = 7 \]  

Here daily confirmed cases and its moving average are made in figure 4 which is quite similar. This means the count of new confirmed cases in a week is similar for all the days with small number of change of count each days. The figure 5 shows the plotted lines are similar. This means that number of death remains similar for a week for all 7 months with some drops in August month. Case completion rate is rate at which the confirmed cases are recovered from the disease. From figure 8 initially the rate was 100% but then decreased below 10% in April month. But in later days the completion rate are becoming good and it is nearly 90%. In figure 9, the trend of fatality rate was from 25 March 2020, the day first death for Covid 19 was recorded. Initially in the month of April the death rate is high but in later days the curve for death is flattened.

**Figure 4.** New confirmed cases vs. MA

**Figure 5.** New death cases vs. MA

**Figure 6.** New recovered cases vs. MA

**Figure 7.** Daily positive sampling rate vs. positive sampling rate
Number of confirmed cases in all 37 districts has been recorded and plots are made based on the number of the total confirmed cases in a district. At the time data collection Mayiladudurai district was along with Nagappatinam district. The Nagapattinam district data contains both districts data. The dataset here used does not contains the airport surveillance cases. For better understanding of the spread of disease across the regions they are divided into 4 groups:

1. Group one has the district which has more than one lakh cases; Chennai is the only district which has more than 1 lakh cases.
2. The second group has the districts has counts between ten thousand and one lakh. Most of the districts are under this category. Third group has the districts which has the cases between 5000 and 10000.
This category has 10 districts. The fourth group holds the district which has less than 5000 cases. This category also has 10 districts.

From the four graphs, cases in Chennai inclined to growth June. Similarly Chenngalpattu and Thiruvannamalai districts are also inclined in June which are geographically near to Chennai. The rest of the districts has more cases after relaxation of district border restrictions.

**Figure 14.** Average number of cases per day in August month

**Figure 15.** Gender distribution of August month dataset

As it is understand from figure 2 that August month has more number of new confirmed, active, death and recovered cases. August month dataset is chosen for getting more insights about the disease. A particular district is taken into account to understand the regional spread of infection. Figure 14 describes the number of cases in aspect of confirmed, recovered, death, active and gender occurs per day. From figure 15 it is clear that male is more affected than other genders. Nearly 60% of infected patients are male and rest is female.

**Figure 16.** Forecast of confirmed cases

**Figure 17.** Forecast of death cases

**Figure 18.** Trend of confirmed cases

**Figure 19.** Trend of death cases

Figure 16 shows the forecast of confirmed cases based on August month data. The linear line depicts that the cases would get possibly increase by days. And the dim area over the linear line represents the
possible fall and rise of case in future. Figure 17 represents that the number of death cases will fall in near future. Figure 18 & 19 shows that the trend of deaths is decreasing by days and the trend of confirmed cases are increasing by days.

Table 5. Error estimation

| Attribute    | Mean Square Error | Root Mean Square Error | $R^2$ |
|--------------|-------------------|-----------------------|-------|
| Confirmed cases | 21035.92          | 145.03                | 0.820 |
| Death cases  | 193.13            | 13.89                 | 0.63  |

Erode district dataset for August month is considered here. Figure 20 show the geographical region or districts limits of patients admitted in Erode district for treatment. The analysis shows that most of the patients are from Erode. Nearly 2% patients are from other districts and state. Figure 21 describes that the residence location of the patients in order to identify the containment zone. While measuring the containment zones it is vital to know the location of the patients, for this purpose location is analyzed here but some patients gave their work location instead of residence which will cause trouble in making zones. For knowing which area has more number of cases the total confirmed cases in August month data is clustered with its area and sub area, figure 22 shows that erode corporation limit has the most number of cases then Bhavani and Sathyamangalam taluks has the more cases. But the conflict occur hear was that sub area name was not clearly mentioned in the corporation limit area. The urban and suburban areas are mostly affected than rural areas. All the area in the top 10 list was either a urban or a suburban area.
5. Conclusion
An ML-based prediction system has been proposed for predicting the risk of COVID-19 outbreak globally. The system analyses dataset containing the day-wise actual past data and makes predictions for upcoming days using machine learning algorithms. From the models and info graphics various insights are obtained. The vulnerability of infection is concluded by comparing the forecast model and actual result when using original data. The accuracy of the model is obtained as well. From the observation we can say that the number of cases and dates are positively correlated. As the days increases the number of cases is also increasing irrespective of the location and economical condition of the districts. District border relaxation and inter-state movements are the major cause of increase in new cases so far. As the forecast and prediction model gives that number of death is less than recovered cases in our geographical region. This study will be enhanced continuously in the future course, next plan is to explore the prediction methodology using the updated dataset and use the most accurate and appropriate ML methods for forecasting. Current data live forecasting will be one of the primary focuses in our future work.

6. References
[1] Furqan Rustam, Aijaz Ahmad Reshi, Arif Mehmood, Saleem Ullah, Byung-Won On, Waqar Aslam and Gyu Sang Choi may.2020 COVID-19 Future Forecasting Using Supervised Machine Learning Models” IEEE ACCESS pp 101489-101499
[2] Romney B Duffey and Enrico Zio may.2020 Analysing Recovery From Pandemics by Learning Theory: The Case of CoVid-19 vol.8 IEEE Engineering in Medicine and Biology Society Section pp 110789-110795
[3] Petar Radanliev, David De Roure and Rob Walton jun.2020 Data mining and analysis of scientific research data records on Covid-19 mortality, immunity, and vaccine development - In the first wave of the Covid-19 pandemic vol.3 Science direct pp 1121-1132,
[4] Jooyeon Park, Jinhwa Jang and Insung Ahn Oct.2017 Epidemic Simulation of H1N1 Influenza Virus using GIS in South Korea IEEE, 2017 International Conference on Information and Communication Technology Convergence (ICTC) pp 58-60
[5] Huadong Xia, Kalyani Nagaraj, Jiangzhuo Chen and Madhav Marathe sep.2013 Evaluating strategies for pandemic response in Delhi using realistic social networks IEEE International Conference on Healthcare Informatics pp 121-130
[6] Vasilis Z. Marmarelis 2020 Predictive modeling of Covid-19 data in the US: Adaptive phase-space approach IEEE Open Journal of Engineering in Medicine and Biology, pp 207-211
[7] Fotios Petropoulos and Spyros Makridakis Mar.2020 Forecasting the novel coronavirus COVID-19 ( ARGENTINA Lidia Adriana Braunstein, Universidad, Nacional de Mar delPlata ) pp 1-8
[8] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime and Thomas Noel 2016 Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis The 6th International Symposium on Frontiers in Ambient and Mobile Systems ScienceDirect pp 1064 – 1069.
[9] Spyros Makridakis, Evangelos Spiliotis and Vassilios Assimakopoulos Mar.2018 Statistical and Machine Learning forecasting methods: Concerns and ways forward Plos one (Alejandro Raul Hernandez Montoya, Universidad Veracruzana MEXICO) pp 1-26
[10] Shaobo He, Yuexi Peng and Kehui Sun 2020 SEIR modeling of the COVID-19 and its dynamics Springer Nature B.V
[11] Sharma and Milan 10 April 2020 Coronavirus: Second ICMR report on random sampling test results shows possible community transmission India Today
[12] Tamil Nadu State Health & Family Welfare Department daily Bulletin, www.stopcornatn.in, Tamil Nadu Government
[13] Chen Y., Liu Q. and Guo, D.: Emerging coronaviruses: genome structure, replication, and pathogenesis. J. Med. Virol.92 pp 418–423
[14] Zhu N., Zhang D and Wang W., 2019 A Novel Coronavirus from patients with pneumonia in China New Engl.J Med pp 727-733
[15] Fanelli D., and Piazzia F 2020 Analysis and forecast of Covid 19 spreading in China, Italy and France Chaos Solitons Fractuals pp 109761-109761
[16] Ceraolo C. and Giorgi, F.M 2020 Genomic variance of the 2019-ncov coronavirus. J. Med. Virol pp 522–528
[17] G. Bontempi, S.B. Taieb, and Y.-A. Le Borgne 2012 Machine learning strategies for time series forecasting Proc. Eur. Bus Intell. Summer School. Berlin, Germany Springer pp 62-77
[18] Ji, W., Wang, W., and Zhao x. 2020 Cross-species transmission of the newly identified coronavirus 2019-ncov. J. Med. Virol. 92, pp 433–440
[19] J.C. Miller 2017 Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes Infectious Disease Modelling 2 pp 35-55.
[20] J. M. Carcione , J. E. Santos, C. Bagaini and J. Ba. May 2020 A simulation of a COVID-19 epidemic based on a deterministic SEIR model Front. Public Health
[21] Hyndman R and Khandakar Y. 2008 Automatic time series forecasting: the forecast package for R. Journal of Statistical Software pp 1-22