Predictive Models for Mitigating COVID-19 Outbreak

Antoine Bagula
Department of Computer Science
University of the Western Cape
Cape Town, South Africa
0000-0003-0774-5251

Hloniphani Maluleke
Department of Computer Science
University of the Western Cape
Cape Town, South Africa
0000-0001-8972-0167

Olasupo Ajayi
Department of Computer Science
University of the Western Cape
Cape Town, South Africa
0000-0001-6583-3749

Amani Bagula
Healthcare Division
Healthcare 4.0 for Africa (H4.0.4A)
Cape Town, South Africa
amani.bagula@gmail.com

Nancy Bagula
Healthcare Division
Healthcare 4.0 for Africa (H4.0.4A)
Cape Town, South Africa
nancy.bagula@gmail.com

Moise Bagula
Healthcare Division
Healthcare 4.0 for Africa (H4.0.4A)
Cape Town, South Africa
bagmoise@yahoo.fr

***Abstract*—Reportedly starting in Wuhan, China in 2019, the Corona virus 2019 (Covid-19) is a pandemic that has hit many countries of the world. Globally the number of infections and death cases have been on the rise and Africa is no exception. As of May 2020, South Africa had been the most affected country in Africa with Cape Town being the pandemic's epicentre. Pre-empting the pandemic rather than attempting to cure infected patients is very crucial to Africa, considering its poorer healthcare system compared to developed countries, which have struggled to curtain the ravaging pandemic despite the advanced state of their healthcare systems. This paper proposes two models that i) validate the proposed protective measures using epidemic modelling; ii) pre-empt the evolution of the pandemic through data analytics. The first model builds around the classic SIR model to simulate the main protective measures suggested by the World Health Organisation; while the second built on regression models to predict future confirmed cases. Real Covid-19 data of the city of Cape Town were used for the simulations and results reveal the accuracy of the models and the relevance of combining simulation modelling and data analytics as relevant tools in the fight against the pandemic.**

***Index Terms*—Covid-19, Epidemic Modelling, Ordinary Differential Equations, Healthcare 4.0, Predictive Modelling.**

I. Introduction

The Corona Virus 2019 (Covid-19) which started in Wuhan, China, has recently been declared by the World Health Organisation (WHO) as a global pandemic, after hitting many countries around the world. The number of infections and death cases have continued to rise globally. As of May 2020, South Africa had the highest caseload in Africa; with Cape Town in the Western Cape province of South Africa, being the epicentre of the pandemic. Figure 1 shows the distribution of confirmed cases in the Western Cape, South Africa. It is widely recognised that pre-empting the pandemic is very crucial. This is particularly true for Africa, considering its poorer healthcare system compared to the more developed countries of the World, where the pandemic has caused more casualties despite their advanced healthcare systems. Mathematical epidemiology based on compartmental modelling can be used to define the scale of an epidemic and the rate at which an infection can spread in a community. It can therefore provide useful insights into how the Covid-19 epidemic can be mitigated. It can also be used to simulate the predictive measures suggested by the WHO and their impact on flattening the infectious curve of the virus; in order to avoid surges that can overwhelm healthcare systems.

Studies on mathematical epidemiology started in the early 20's with the work of Kermack and McKendrick pioneering a contribution to the Mathematical Theory of Epidemics [1]. Further research works were conducted to present different aspects of mathematical epidemiology by looking at different theoretical and application aspects of the field. These include [2] that focused on the analysis of equilibrium in malaria, [3] wherein mathematical models for the spread of infectious diseases were presented and [4] which focused on the definition and computation of the basic reproduction number. Further work on mathematical epidemiology emerged later with focus on both theoretical and simulation aspects of epidemics. In [5], a basic introduction to disease modelling is presented with primarily focus on common issues encountered when structuring and analysing SIR models. The work in [6] addressed the interaction of ecological processes, with the objective of modelling these interactions in order to understand the dynamics of infectious agents in communities consisting of interacting host and non-host species. The basic reproduction number is one of the conceptual cornerstones of mathematical epidemiology. It is classically defined as the number of secondary cases generated by a typical infected individual in a fully susceptible population. The basic reproduction was introduced in [7] with an historic reminder describing the steps leading to the statement of its mathematical definition. An explanation of how it can be calculated using the next-

© IEEE 2020. This article is free to access and download, along with rights for full text and data mining, re-use and analysis

generation matrix method in the case of epidemic models described by Ordinary Differential Equations (ODEs) was also given. The work presented in [8] studied two multiplex contact networks representative of a subset of the Italian and Dutch populations from highly detailed socio-demographic data. The authors used these networks to simulate the infection transmission process of the influenza virus, calibrated on empirical epidemiological data whilst taking its natural history into account. The work revealed that the classical concept of the basic reproduction number is untenable in realistic populations, and does not provide any conceptual understanding of the epidemic evolution. In [9], the authors proposed a Bats-Hosts-Reservoir-People (BHRP) transmission network model to study the Covid-19 from infection sources (probably bats) to humans. The study also considered a simplified Reservoir-People (RP) transmission network model and used the next generation matrix approach to calculate the basic reproduction number $\mathcal{R}_0$ from the RP model to assess the transmissibility of the SARS-CoV-2.

This paper proposes predictive models that build around: i) the SIR model to mimic the main Covid-19 protective measures and ii) machine learning models to predict the evolution of the pandemic with time. Building upon traffic data scrapped from the TomTom system, the proposed SIR model is implemented and simulated as a predictive analytic tool to support the efforts of government in fighting the pandemic. Simulation results reveal the accuracy of the predictive model and the relevance of using such model to provide awareness on network effects and measures that can be taken to mitigate the impact of the pandemic. Furthermore, regression based machine learning models are applied to real Covid data collected from Cape Town; to predict the evolution of the pandemic based on confirmed cases, deaths and recoveries. The models reveal a good performance in the prediction over the period of time considered.

The rest of this paper is organised as follows: following this introductory section is a concise overview of the Corona Virus 2019 in terms of its evolution and public health response in section II. Section III presents the epidemic model while its performance evaluation through simulation modelling is presented in section IV, while the prediction models are discussed in section V. Conclusions and future directions are given in section VI.

II. Corona Virus 2019 Context

Covid-19 or SARS-CoV-2 is a strain of Coronovirus and part of a large family of viruses that cause respiratory illness in humans. Symptoms of Covid-19 are as depicted in Figure 2 and reports have shown the elderly are more susceptible to the virus. As at the time of writing, no known cure or vaccine for the Covid-19 exists. In light of this, different strategies have been adopted by the WHO as public health response to the pandemic. These include:

1) **Interrupt human-to-human transmission** including reducing secondary infections among close contacts and healthcare workers, preventing transmission amplification events, and preventing further international spread;
2) **Identify, isolate and care** for patients early, including providing optimised care for infected patients;
3) **Identify and reduce transmission** from animal sources;
4) **Address crucial unknowns** regarding clinical severity, extent of transmission and infection, treatment options, and accelerate the development of diagnostics, therapeutics and vaccines;
5) **Communicate critical risk and event information** to all communities and counter misinformation;
6) **Minimise social and economic impact** through multisectoral partnerships.

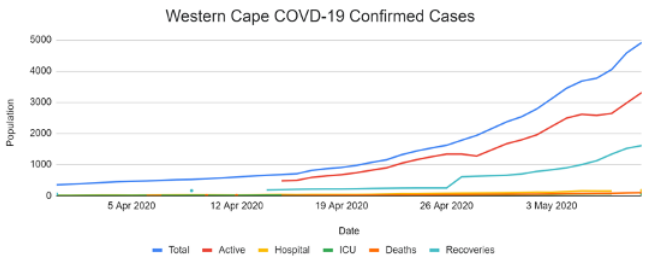


Fig. 1. Covid-19 Confirmed Cases as of 9th May 2020.

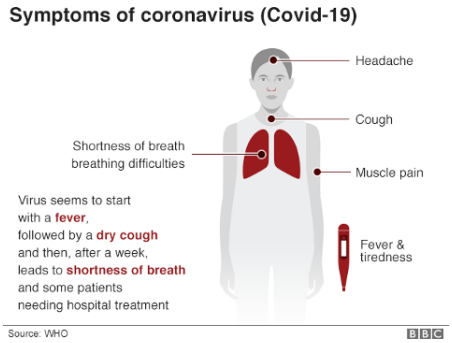


Fig. 2. Covid-19 Symptoms.

Note that the interruption of human-to-human transmission can be achieved through a combination of public health measures such as: rapid identification, diagnosis and management of the cases, identification and follow up of contacts, infection prevention and control in health care settings, implementation of health measures for travellers, awareness-raising in the population and risk communication.

III. Covid-19 Epidemic Modelling

In modelling the Covid-19 epidemic, we introduce three state categories for individuals in any given population namely: Susceptible (S), Infected (I) and Recovered (R). Figure 3 presents a finite state machine of our epidemic model for a system partitioned into m population groups or locations. It reveals the states of individuals and for each state its associated transitions, as well as the actions that trigger transitions from one state to another.

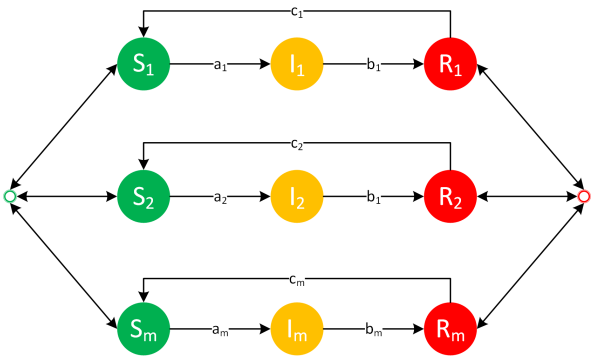


Fig. 3. State Diagram of the Epidemic Model.

We assume that susceptible individuals S_i in a population group or location may be infected at the rate a_i, while the infected individuals I_i move to the Recovered state R_i at a rate b_i. While recovered individuals might move back to the Susceptible state at a rate of c_i, this case has not been considered in this work. Susceptible individuals might also move from a location i to another location j with rate a_{ij}. Movement across population groups (locations) is synonymous to people moving across provinces. The differential equation (equation 1) is a result of these assumptions.

$$\begin{cases} S_i(t+1) = S_i(t) - a_i(t)S_i(t) - a_{j,i}S_i(t) \\ I_i(t+1) = I_i(t) + a_iS_i(t) - b_i(t)I_i(t) + a_{j,i}S_j(t) \\ R_i(t+1) = R_i(t) + b_i(t)I_i(t) \\ N_i(t) = S_i(t) + I_i(t) + R_i(t) \end{cases} \tag{1}$$

Where $S_i(t)$, $I_i(t)$ and $R_i(t)$ are functions of time t for each location i and N_i denotes the size of the population at location i. The negative rates in the model represents a decrease, while the positive ones represent an increase in population. The parameter a_i stands for the transmission rate between the susceptible and infected population while the parameter b_i represents the recovery rate from infection. The parameter a_i depends on the number of susceptible individuals S_i and infected individuals I_i as well as the actions taken to contain the pandemic in terms of social distancing. It therefore makes sense to relate a_i with three other measures:

1) The random transmission rate of each individual at a location i at time t, denoted by $\beta_i(t)$.
2) The fraction of infected population in location i at time t, denoted by $x_i(t) = \frac{I_i}{N_i(t)}$.
3) A Human-to-Human transmission parameter ρ expressed by α for lockdown or δ for social distancing. $\rho = 1$ depicts a base case scenario, $\rho = 0$ depicts total lockdown (or quarantine), while $0 < \rho < 1$ depicts in-between both ends.

Therefore, a_i can be computed using the following formula:

$$a_i(t) = \alpha\beta_i(t)x_i(t) \tag{2}$$

It is important to note the following:

1) The third term $a_{j,i}S_i(t)$ of the first expression in (1) is the same as the fourth term in the second expression. It is used to express the shared infection between locations represented by the fraction of infected people that have arrived from other location j into the location i, weighted by their respective transmission rates $\beta_j(t)$. It is expressed by

$$a_{j,i}S_j(t) = \alpha S_j(t)\frac{\sum\limits_{j\neq i} w_{i,j}(t)x_j(t)\beta_j(t)}{N_i + \sum\limits_{j\neq i} w_{i,j}(t)}$$

.

where α **is a lockdown effect parameter** representing the intensity of traffic in a location, which can be used to reveal the impact of the lockdown on the model. $w_{i,j}(t)$ is the mobility from location j to location i, while $x_j(t) = \frac{I_j(t)}{N_j}$ is fraction of infected population at location j at time t.

2) The shared infection expression above can be expressed by $a_{j,i}S_j(t) = \alpha\alpha_{j,i}S_j(t)$ where

$$\alpha_{j,i}(t) = \frac{\sum\limits_{j\neq i} w_{i,j}(t)x_j(t)\beta_j(t)}{N_i + \sum\limits_{j\neq i} w_{i,j}(t)}$$

.

3) Given that the total population is constant as expressed by the fourth equation in (1), the portion of shared infection (infected group from other locations) is subtracted from the population of susceptible individuals, that are to be added to the infected group in the first equation. It is also added to the population of infected individuals that move to the recovered state in the second equation.
4) The parameter $b_i(t)$, which in our model represents the recovery rate, can be considered as time-independent as expressed by $b_i(t) = \gamma$ or time-dependent when expressed as a function of time by $b_i(t) = \gamma(t)$.

The pandemic prediction problem consists of finding for each location i, the evolution of the function set over time $L_i(t) = \{S_i(t), I_i(t), R_i(t)\}$; while taking into account the transformation above and the difference equations in (1). It is defined as follows:

$$\begin{array}{ll} Find & L_i(t) = [S_i(t), I_i(t), R_i(t)] \\ subject\ to & \end{array} \tag{3}$$

$$\begin{cases} S_i(t+1) = S_i(t) - \delta\beta_i(t)x_i(t)S_i(t) - \alpha\alpha_{j,i}(t)S_j(t) & (3.1) \\ I_i(t+1) = I_i(t) + \delta\beta_i(t)x_i(t)S_i(t) - \gamma I_i(t) + \alpha\alpha_{j,i}(t)S_j(t) & (3.2) \\ R_i(t+1) = R_i(t) + \gamma I_i(t) & (3.3) \\ N_i(t) = S_i(t) + I_i(t) + R_i(t) & (3.4) \\ \mathcal{R}_0 = \frac{\beta}{\gamma} & (3.5) \end{cases}$$

Note that equations (3.1), (3.2) and (3.3) form a system of Ordinary Differential Equations (ODEs) expressing the dynamics of the compartmental epidemic model by showing how the disease is disseminated from one compartment to another. On the other hand, equation (3.4) is a constraint that reveal that the population remains constant throughout, which is an indication of the absence of birth or immigration, death or emigration. In our model the equation (3.5) has been used to reveal how the basic reproduction number $\mathcal{R}_0$, a parameter that defines how quickly the disease spreads, is related to the

other parameters of the model. $\mathcal{R}_0$ is a parameter that plays a crucial role in determining the fate of an outbreak. If its value is greater than the ratio of total population to the susceptible cases at time zero ($\mathcal{R}_0 > \frac{N(0)}{S(0)}$), then it is assumed that the outbreak will turn into a full-fledged epidemic. If otherwise, that is ($\mathcal{R}_0 <= \frac{N(0)}{S(0)}$), then the outbreak will not result in an epidemic.

IV. Performance Evaluation

We conducted a set of experiments to simulate the outcome of some of the public health protective measures proposed by the WHO and implemented by the South African Government as responses to the pandemic mitigation. Two measures were considered: i) locking down the city in order to reduce traffic intensity and ii) social distancing with the expectation of reducing the probability of infection through contacts. As suggested earlier, both measures are designed to flatten the infectious curve and prevent local healthcare systems from being overwhelmed.

A. Initial Model Parameters

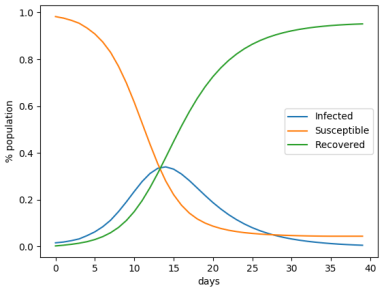
Given that in the case of Covid-19, the value of the basic reproduction number $\mathcal{R}_0$ varies across countries, ranging between 1.4 and 4.0; we selected as initial value $\mathcal{R}_0 = 4$ to consider the worst case scenario. The value of γ was considered based on the average infectious rate in Cape Town for the first 40 days of the outbreak. It was set to 0.23, while the value of β was derived from equation (3.5). Our initial population was set to $S(0) = 3,776,000$, considering the entire population of Cape Town as being susceptible to the Covid-19 outbreak. The initial time T was set to be day 1, the day the first cases were confirmed in Cape Town. The simulation was used to predict the number of susceptible, infected and recovered cases over a period of 40 days starting from day 1. We use the first 40 days because this duration cuts across the first few weeks of the outbreak and the first few weeks of the government imposed lockdown and social distancing regulations.

B. Simulating the city lockdown

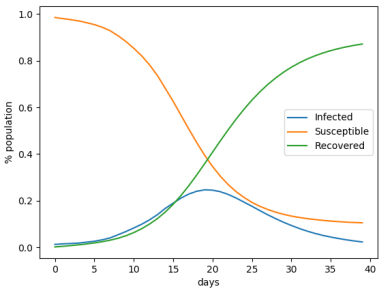
To simulate the city lockdown, the traffic intensity parameter α was varied in the lockdown experiments to evaluate its impact of the epidemic curves. We conducted the first set of experiments using different traffic intensity within the city and compared the pandemic curves obtained to assess if and when the lockdown enabled the flattening of the infectious curve. Different traffic intensities were considered in our experiments with values set to α= $0.9, 0.8$ for minimal lockdown, and α =$\{0.6, 0.4\}$ for stringent lockdown. We did not go beyond 40% lockdown as essential services remained functional during the lockdown. The results are presented in Figures 4 and 5.

The results under minimal lockdown depicted in Figure 4 revealed that:

For $\alpha = 0.9$, the infection curve was nearly normally distributed and peaking at about 40% by the 15th day. This implies that in two weeks, about 40% of the population would be infected. The recovery curve reveals an exponential growth with starts to plateau at the 30th day, revealing that almost a full recovery could be seen after about a month.



(a) $\alpha = 0.9$.



(b) $\alpha = 0.8$.

Fig. 4. Minimal City Lockdown: $\alpha = 0.9$ and $\alpha = 0.8$.

For $\alpha = 0.8$, the slight reduction in traffic intensity resulted in a slightly flatter infection curves. This reveals that a reduction of the traffic intensity by 10% had major impact on the shape of the epidemic curve, reducing its peak to about 25%. It is important to note that, the susceptible curve never gets to zero. This is as expected, as individuals who are locked down are still susceptible and a person can only be in one of the three states at any given time.

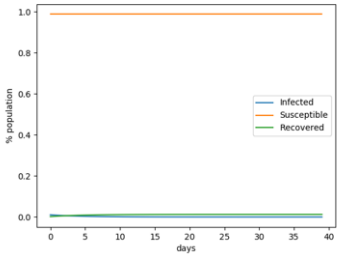
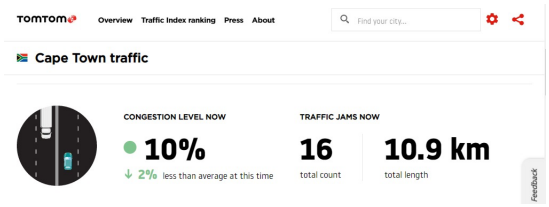


Fig. 5. Stringent City Lockdown: $\alpha \leq 0.5$.

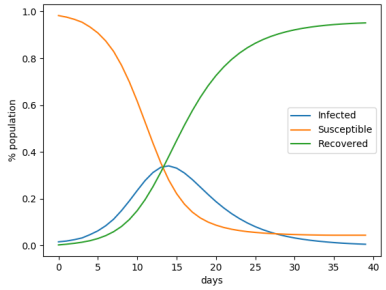
The results under stringent lockdown in Figure 5 revealed that: for $\alpha = 0.5$, the infectious curve was completely flat. This is as expected, because if stringent lockdown had been in place right from the first day, then no one would have been infected. Similarly, with no one infected, there would be no recoveries, hence the flat recovery curve. The susceptibility curve remains

at the peak, because as earlier stated we assumed everyone in the city is susceptible at the beginning. Besides an individual would still be considered susceptible even if in isolation. Similar results were obtained for values of $\alpha \leq 0.5$, hence these graphs are not shown.

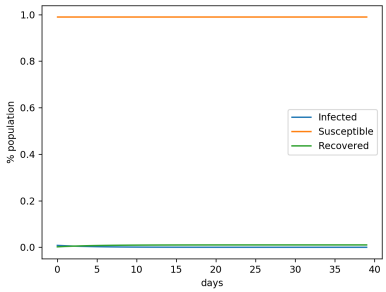
From these results, it can be inferred that, if the city had been put on any form of lockdown from the first few days, the outbreak would have been quickly contained. Even a 10% restriction of mobility would have translated to less than 25% of the population being infected with quicker recovery.



(a) Tomtom Data.



(b) $\alpha = 0.9$.



(c) $\alpha = 0.1$.

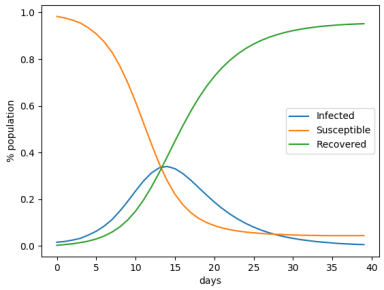
Fig. 6. Traffic intensity comparison: $\alpha = 0.9$ and $\alpha = 0.1$.

To verify this claim, another experiment was conducted to reveal the impact of the model when using real traffic intensity data. Traffic data for the city of Cape Town was web-scrapped from TomTom and applied to the model to evaluate the impact on the curves. As revealed by Figure 6, the real traffic intensity in the city at that point in time was 0.1. When we applied this traffic intensity to the model, the following was observed:

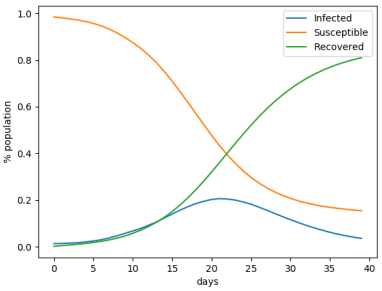
As in the case with $\alpha \leq 0.5$, the infectious and recovery curves were completely flat. Similarly, the susceptibility curve remained at 100%, implying that everyone in the city was still vulnerable to the virus.

C. Simulating the Impact of Social Distancing in the City

The social distancing impact was simulated in a manner similar to the city lockdown by considering different values of the δ parameter for what were considered as both minimal and stringent social distancing. These values were used in the social distancing experiments to reveal their impact on the epidemic curves. Four different distance values were considered in our experiments, viz.: $\delta = \{1.0, 0.8\}$ for minimal social distancing and $\delta = \{0.5, 0.4\}$ for stringent social distancing. The results are presented in Figures 7 and 8.



(a) $\delta = 1.00$.



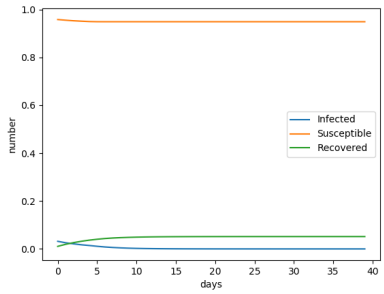
(b) $\delta = 0.80$.

Fig. 7. Minimal Social distancing: $\delta = 1.0$ and $\delta = 0.80$.

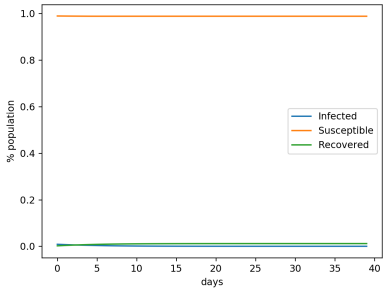
Figure 7 reveals that for $\delta = 1.0$, with no social distancing, a peak of about 37% was reached in less than 2 weeks of the pandemic. Interestingly, there was an interception between the number of recovered individuals, number of susceptible and number of infected on the 12th day. The recovery curve shows that recovery would rise gradually and start to plateau after the 30th day. However, it never gets to 100%, implying that full recovery would not be achieved in this time frame. This reveals that aggressive social distancing would be required to tackle the pandemic.

For $\delta = 0.80$, when minimal social distancing starts to take effect, the peak of infection drops from 40% to 20% on the 21st day of the pandemic. It then starts decreasing but never gets to zero. The recovery curve rises gradually and only started to plateau after the 37th day; however, it never reaches the 100% mark. This implies that with only minimal social distancing, about 20% of the population would still remain infected. It is important to note that, the peak of the infectious curve was almost halved, revealing that with even minimal

social distancing, the percentage of the infected population can be cut by almost half.



(a) $\delta = 0.50$.



(b) $\delta \leq 0.40$.

Fig. 8. Stringent Social distancing

The results under stringent social distancing depicted in Figure 8 revealed that for $\delta = 0.5$, the infection was completely contained after about 10 days. With a peak of about 5% at day 1. This implies that by practicing social distancing immediately the virus broke out, the outbreak could have been defeated within the first two weeks. In response, the recovery curve peaked around the 10th day and plateaued afterwards. This was also expected, as the outbreak would have been contained by the second week. Similarly, the susceptibility curve dropped slightly from 100% to 95%. This 5% drop translates to the individuals who moved from susceptible to infected, when the virus first broke out.

For $\delta = 0.40$, the infection and recovery curves were completely flat. This means that if social distance was aggressively implemented earlier, the pandemic breakout would have been prevented. Furthermore, all individuals would have remained in the susceptible state.

V. Predicting the Pandemic's Evolution

Experiments were also conducted to predict the evolution of the pandemic in Cape Town. For these experiments, two regression-based machine learning models were used, namely: Multi-Linear Regression (MLR) and Support Vector Regression (SVR). The models were trained using publicly available data about the epidemic in Cape Town, for the first 40 days of the outbreak. Due to the small size of the dataset, all were used in training the models. New sets of data were then generated to test the models using the population growth formula in 4.

$$P(t) = P_0 E^{rt} \tag{4}$$

where P is the total population after time t, P_0 is the starting population, r is the percentage rate of growth, t is the time in hours and E is the Euler number = 2.7182. This new data set spanned the first 60 days and allowed us test the predictive performance of our models, for the 20 days after the training data.

A. Prediction Models & Metrics

MLR is a variant of linear regression analysis, wherein multiple independent variables are used to predict a dependent variable. Support vector regression (SVR) on the other hand is a machine learning regression model based on Support Vector Machine [10] which seeks to minimize the predictor coefficient to a value $\leq$ threshold (ϵ). For our SVR implementation, the following parameters were used $C = 10^7, \gamma = 1e^{-10}$, rbf kernal. These were obtained after performing grid search on SVR. For both models, Recovery P_{rec}, deaths P_{dea} and infections P_{inf} rates were used as multiplicative scalars for the independent variables and expressed in equation 5.

$$Y_p = Intercept + P_{inf}X_0 + P_{rec}X_1 + P_{dea}X_2 \tag{5}$$

where $Intercept = 35.088$, $P_{inf} = 0.779$, $P_{rec} = 0.720$, $P_{dea} = 5.995$

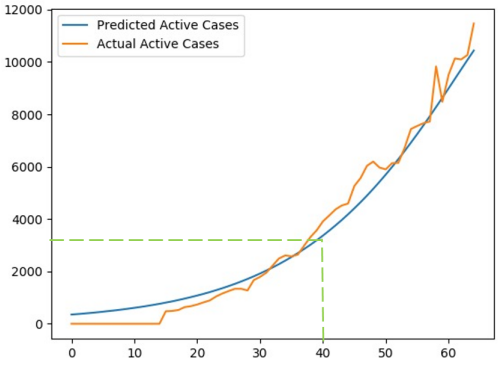
In evaluating the resulting prediction, we used the Root Means Square Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2). RMSE is the square root of the averaged squared difference between the actual and predicted values. MAE measures the absolute difference between the actual and predicted values, while R^2 is used to benchmark a selected model against a baseline (mean of the data).

B. Prediction Results

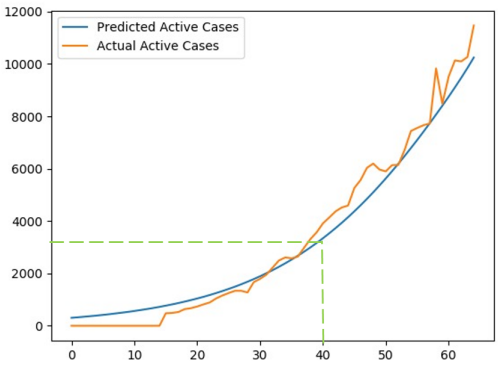
A comparison between reported (actual) numbers of Covid-19 cases and the predicted numbers are shown in Figure 9. Both figures are plots of days (x-axis) against number of cases (y-axis); with the orange and blue lines depicting the actual and predicted data respectively.

Though both models gave similar output curves, their values for each of the metrics differed. For MLR, the RMSE was 511.87, which was lower than SVR's value of 545.83. Similarly, MLR also resulted in a lower MAE value at 422.75 compared to SVR's 438.54. These results indicate that MLR's predicted values were slightly more accurate (closer to the actual values) than those of SVR. This was further buttressed by the R^2 values, with MLR's 0.976 being closer to 1 than SVR's 0.973.

These differences are however only marginal and both figures show that the predicted data for confirmed cases follows the trajectory of the reported data from the national database. It can therefore be concluded that, there is a strong correlation



(a) Prediction curves - MLR.



(b) Prediction curves - SVR.

Fig. 9. Predictive Evolution of Confirmed Cases.

between the predicted data and the reported trends within the actual data $(r = 0.999, p < 0.001)$.

VI. Conclusions and Future Work

The goals of the study were to: i) produce simulation results that could help in validating the measures put in place to mitigate the impact of the Covid-19 virus and ii) use machine learning based data analytics to produce models that can predict the evolution of the pandemic based on the actual numbers of recovered, infected and confirmed cases. Building upon the classic SIR model, this study's focus modelled the Covid-19 epidemic in Cape Town, South Africa; with the objective of estimating the number of infections, the peak infection day, the rate of increase of infections and the resolution of the end-point of the epidemic. The SIR model included two parameters used to respectively mimic the lockdown and social distancing measures. By varying these parameters in a set of carefully selected simulation experiments, the model revealed the relevance of the preventive measures by showing: i) how the peak of the infectious curve could be flattened; ii) how the infection rate could be reduced to prevent overwhelming the healthcare system; iii) how the scale of the pandemic could be reduced in time, with full immunity of the population at the end of the pandemic. Two predictive data analytic models complemented the simulation results by revealing (with high precision) the evolution of the confirmed cases based on the actual number of cases, recoveries and deaths.

A number of other epidemic models could be used to mimic the Covid-19 pandemic, in particular the Susceptible-Exposed-Infected-Removed (SEIR) model, which accounts for a period of exposition to the virus before infection. The extension of the current work using this model and the underlying simulations for the other Covid-19 affected cities of South Africa as well as tracking the migration of affected individuals across these cities are directions for future research. At the time of writing, Covid-19 related data for other African countries, including the Democratic Republic of Congo, have been collected for the purpose of replicating this study in those countries. This is another avenue for future research work and an extension to the framework proposed in [11] for supporting healthcare in Africa specifically and the developing nations in general.

References

[1] Kermack, W.; McKendrick, A. (1927). A Contribution to the Mathematical Theory of Epidemics, in Proceedings of the Royal Society A. 115 (772), pp. 700–721. doi:10.1098/rspa.1927.0118.

[2] McDonald, G. (1952). The analysis of equilibrium in malaria, Trop. Diseases Bull., 49, pp. 813–829.

[3] Hethcote, H. (1975). Mathematical models for the spread of infectious diseases, in Epidemiology, Utah Conference Proceedings, SIAM, Philadelphia, pp. 122–131.

[4] Diekmann, O., Heesterbeek, J. and Metz, J. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations, J. Math. Biol., 28, pp. 365-382.

[5] Blackwood, J. and Childs, L. (2018) An introduction to compartmental modelling for the budding infectious disease modeller, Letters in Biomathematics, 5(1), pp. 195-221, DOI: 10.1080/23737867.2018.1509026

[6] Roberts, M. and Heesterbeek, J. (2012). Characterizing the next-generation matrix and basic reproduction number in ecological epidemiology, in Journal of Mathematical Biology, DOI 10.1007/s00285-012-0602-1.

[7] Perasso, A. (2018). An introduction to the Basic Reproduction Number in mathematical Epidemiology, in ESAIM: Proceedings and Surveys, 62, pp. 123-138.

[8] Quan-Hui L. et al (2018). Measurability of the epidemic reproduction number in data-driven contact networks, In PNAS Journal, 115(50), pp. 12680 -12685.

[9] Tian-Mu C. et al (2020). A mathematical model for simulating the phase-based transmissibility of a novel coronavirus, in Chen et al. Infectious Diseases of Poverty, pp. 9-24.

[10] Jakkula, V. (2006). Tutorial on support vector machine (svm). School of EECS, Washington State University, 37.

[11] Bagula, A., Mandava, M. and Bagula, H. (2018). A framework for healthcare support in the rural and low income areas of the developing world, in Journal of Network and Computer Applications, 120, pp. 17-29.