Predicting the Impact of Covid-19 with Modified Epidemiological Model Using Deep Learning

Yixian Chen
Walmart Labs
Sunnyvale, CA, USA
yixian.chen@walmartlabs.com

Jialu You
Walmart Labs
Sunnyvale, CA, USA
jialu.you0@walmartlabs.com

Ivan(Lei) Ji
Tik Tok
Mountain View, CA, USA
ivan_chi@ymail.com

Prakhar Mehrotra
Walmart Labs
Sunnyvale, CA, USA
prakhar.mehrotra@walmart.com

ABSTRACT

In this work, transmission rate and the related impact of COVID-19 have been systematically studied and predicted. We propose an improved modification of the standard Susceptible Exposed Infection and Recovered (SEIR) model factoring with policy-risk-related parameter. By modifying the SEIR epidemic model from deterministic differential equations to stochastic differential equations (SDEs), we improved the model reliability and usability using the extended Kalman Filter. With the time series of reproduction number R calculated from our modified SEIR with extended Kalman Filter, we can then predict the future transmission rate using novel deep learning approaches. Furthermore, to solidly reveal the impact of COVID-19, our work provides detailed and systematical study on two different levels of granularity: national and state-level. Hence, the framework proposed in this work can be effectively and accurately implemented on different region scales.

KEYWORDS

Extended Kalman Filter, Epidemiological Model, LSTM, GRU, Variational Autoencoder

ACM Reference Format:
Yixian Chen, Jialu You, Ivan(Lei) Ji, and Prakhar Mehrotra. 2021. Predicting the Impact of Covid-19 with Modified Epidemiological Model Using Deep Learning. In *epiDAMIK 2021: 4th epiDAMIK ACM SIGKDD International Workshop on Epidemiology meets Data Mining and Knowledge Discovery*. ACM, New York, NY, USA, 6 pages. https://doi.org/xx.xxxx/xxxxxxxxx.xxxxxxx

1 INTRODUCTION

A reliable modeling and forecasting methodology of COVID-19 has been crucial since the outbreak of this novel global disease.

Understanding the transmission behavior of COVID-19 is of utmost importance to be able to respond to the outbreaks and take action against the spread of the disease. Significant research effort has been made world-wide by trying to build the transmission model of COVID-19 in the field of data science. However, the task is complicated especially when governmental regulations can introduce strong interventions to the model's performance and reliability.

The study of epidemic diseases using filtering methods is a recent advancement even if Kalman filters has been widely applied to engineering problems. Filtering techniques are studied to model various dynamic systems, and with SEIR model framework, we are able to derive the transmission rates dynamically.

Our approach combines the modified SEIR using Extended Kalman Filter to estimate the effective reproduction number, R, over time. For the reproduction number forecasting, we build a two-branch neural network structure to select the most accurate network for different levels of seasonal variation in data. Our research modified the SEIR model by adding an additional policy-related and risk-related parameter accounting for factors such as age, population density, local quarantined condition and social-distancing policy. Our research then utilizes Extended Kalman Filter method to the stochastic infection propagation modified SEIR model, in order to filter out the noise and capture the true state change. The EKF method outputs a time series of estimated reproduction number R. According to the time series data, we could then predict the future R using deep learning methodologies. Eventually, our research compares the predictive accuracy of each model versus different data characteristics, such as seasonal variation, and concludes a model selection module to address different model's applicability. And we conduct the study mainly on national-level and state-level.

epiDAMIK 2021, Aug 15, 2021, Virtual

ACM ISBN 978-1-xxxx-XXXX-X…$15.00
https://doi.org/xx.xxxx/xxxxxxxxx.xxxxxxx

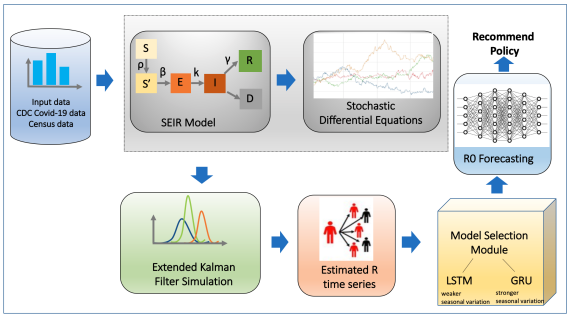

Figure 1: Illustration of Epidemiological Framework. In this framework, a region specific disease parameter space is constructed using surveillance COVID-19 data and census data, followed by the U.S. COVID-19 data consisting of weekly state-level data. A two-branch deep neural network model is trained on the these data. The trained model takes the surveillance data as the input during forecasting.

2 RELATED WORK

2.1 Current SEIR Models and limitations

Various mathematical models and computer simulations have been widely used to conduct analysis or prediction on infectious diseases. The most common approach to model the COVID-19 transmission is the SEIR model which provides a basic method to model the transmission of different kinds of epidemic.

A stochastic discrete-time SEIR model for infectious diseases is developed with the aim of estimating parameters from daily incidence and mortality time series for an outbreak of Ebola in the Democratic Republic of Congo in 1995 [11]. Ndanguza *et al.* [12] performed numerical simulations of the SEIR model in a stochastic point of view using extended Kalman filter methods, and confirmed that the parameters are also identifiable with the stochastic differential equations of the SEIR model.

However, there are several limitations in the current SEIR models. One limitation is that the current COVID-19 SEIR-based models assumes the transmission rate to be constants, which simplifies both mathematical analysis and model fitting. However, in reality, the transmission rates will change with the epidemiological and socioeconomic status and will also be affected by the governmental policies. It's known that the governmental policies include social distancing and quarantine control can effectively reduce the spread of coronavirus. For example, since the early stage of COVID-19 outbreak in the United States, more state and local governments have been imposing curfew orders which effectively slowed down the COVID-19 transmission.

In addition, the SEIR model assumes individuals of the whole population are equally-susceptible. But there are many risk factors such as age, which implies the more senior adults are at higher risk of getting infected by (and recovered from) COVID-19. Therefore, we modified the SEIR model with additional parameters such as age, community location to reflect the difference in terms of risk factors.

2.2 Reproduction Number (R)

The government policy maker would set a COVID-19 alert level, primarily determined by the number of coronavirus cases, and by R, the reproduction number. As Germany's chancellor, Angela Merkel, explained in a widely viewed video in April 2020, an R above one means an outbreak is growing, and below one means that it is shrinking. It doesn't capture the current status of an epidemic and can spike up and down when case numbers are low. It is also an average for a population and therefore can hide the local variations. [7]

2.3 Machine Learning Approaches

Machine Learning approaches have been widely used to model and forecast the infected cases and death cases[8]. Deep learning based forecasting framework in combination with SEIR has been proposed by [18] and validated using synthetic data. In [19], Zeroual et al. applied a simple Recurrent Neural Network (RNN), Long short-term memory (LSTM), Bidirectional LSTM (BiLSTM), Gated recurrent units (GRUs) and Variational AutoEncoder (VAE) algorithms for global forecasting of COVID-19 cases based on a small volume of data. It showed the superior performance of the VAE compared to the other algorithms using COVID-19 data from six countries. Recent researches about non-pharmaceutical interventions focused on mobility data to combine epidemiological models and learning algorithms. [8] provides a Variational-LSTM Autoencoder predictor applied to the virus spread data but also includes factors related to urban characteristics represented in demographic data. Vollmer et al.[16] integrate mobility in a stochastic model. They focused on Italy and suggest that COVID-19 transmission rate and mobility metrics are closely related as well as mobility should be closely monitored in the next weeks and months. Our approach relies on a similar intuition and builds the predictor on reproduction number. Our predictor builds a two-branch neural network structure to take both with large and small seasonal variation situations and finally shows better performance compared to the other approaches in [8, 19]. In this work, we also demonstrate that using the output of our ML approach as a recommendation function for future policy making to effectively suppress the transmission of COVID 19.

3 METHODOLOGY: MODIFIED DYNAMIC SEIR MODEL USING EXTENDED KALMAN FILTER

3.1 Modified SEIR Model with consideration of policy factors and risk factors

The SEIR epidemic model divides the population into key compartments: susceptible (S), latent or exposed (E), infectious (I), and recovered (R). The transition process among these four compartments follows the form S $\rightarrow$ E $\rightarrow$ I $\rightarrow$ R: susceptible people can become exposed to the virus, then infected, then recovered or dead. However, when non-pharmaceutical interventions such as quarantine or lock-down policies are executed, these compartments will also change accordingly. Therefore, we modify the original SEIR model by adding an additional parameter ρ to account for relative policy factors and risk factors. Due to the epidemiological characteristics of COVID-19 and the local social-distancing policies carried

out by the government, we extend the original SEIR model with a short-term model. Caused by the high complexity of the actual situations and uncertainties, some simplifications are necessary for analyzing the theoretical math model. Thus, our model satisfies the following assumptions:

1) the population is assumed relatively fixed. We don't consider new compartments and dynamics such as new births or non-disease related deaths due to the short-term time window;

2) It's been found that COVID-19 infection induces robust, neutralizing antibody responses that are stable for at least three months [17]. Thus, second infection is not considered in this short-term model;

3) All coefficients involved in this model are positive constants.

4) Homogeneous mixing: any individual in the susceptible compartment is equally likely to contact any other non-quarantined individual in the population on a given day;

5) On the population level, the disease processes are identical for all individuals.

Based on the above assumptions and the actual policy-related and risk-related factors, such as age and population density, the spread of COVID-19 in the populations can be described in Figure 1.

And the corresponding ordinary differential equations of this model depicted in the graphical scheme (See Figure 1 SEIR model module) is formulated as below:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\beta\frac{S'(t)I(t)}{N} = -\beta\frac{\rho S(t)I(t)}{N} \\
\frac{dE(t)}{dt} &= \beta\frac{\rho S(t)I(t)}{N} - kE(t) \\
\frac{dI(t)}{dt} &= kE(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t) \Leftrightarrow R = N - S - I - E
\end{aligned} \tag{1}$$

The population is divided into six groups, where $S(t)$, $S'(t)$, $E(t)$, $I(t)$, $R(t)$, and $D(t)$ denote the original susceptible, new susceptible with considerations of social-distancing policies and other risk factors, exposed, infectious, recovered, and dead respectively. In equation (3.1), we use bilinear incidence rates to describe the infection of COVID-19 and the parameter β marks the contact rate between the susceptible and the exposed. We adopt parameter ρ to present the ultimate parameter accounting for risk factors such as age, population density, local quarantined rate and social-distancing level. In other words, when the original susceptible people in the population at time t is $S(t)$, only the ρ portion of this group, i.e. $\rho S(t) = S'(t)$, is considered to be really susceptible under the particular local social-distancing policies and risk factors. The transition rate of the exposed to the infectious group is denoted as k and the recover rate is denoted as γ.

3.1.1 Reproduction Number R Estimation. We acquired COVID-19 the national-level and state-level confirmed cases and deaths data from publicly reported statistics compiled by the New York Times [15]. In order to estimate the value of basic reproduction number (R_0), we used the daily published data as well as the estimated data to perform modified SEIR model fitting. The smallest time unit is one day. We assumed some parameters according to the epidemic characteristics of COVID-19: $k = \frac{1}{5}$, $\gamma = \frac{1}{14}$. The transition rate between E and I is denoted by k, and the incubation period should be $\frac{1}{k}$ days. Since the median incubation period of COVID-19 is around 5 days from exposure to symptoms onset [1], the parameter k is thus assumed to be $\frac{1}{5}$ in our model. γ is the average rate of recovery or death in infected populations. Because the average recovery time for COVID-19 is 2 weeks, we take the recovery rate γ as $\frac{1}{14}$. Between I and R, the transition rate is assumed to be proportional to the number of infectious individuals which is γI. Thus, the basic reproduction number R at time t is:

$$R_t = \frac{\rho_t \beta_t}{\gamma_t} \tag{2}$$

3.1.2 Ordinary differential equation modeling to stochastic infection propagation modeling. The mathematical representation of SEIR diseases is widely used deterministically: "dynamical biological processes are better modeled by means of systems of deterministic ordinary differential equations (ODE), partial differential equation (PDE), or delay differential equation (DDE) [10]." The deterministic modeling is simple to interpret and analyze in terms of how a system changes or evolves, while the stochastic modeling considers uncertainties in the transmission of disease between individuals. Our model implements the stochastic infection propagation modeling onto the modified SEIR model. The modified SEIR deterministic model has been explained in earlier section **3.1** in **Equation (3.1)**. This section illustrates the mathematical conversion from SEIR ODE to Stochastic Differential Equation (SDE). This conversion is conducted using method found in [5] [3] [4] [13] and is proved in the supplementary materials. The stochastic SEIR has the transition probability as below:

$$\begin{aligned}
&P(s+k, i+j, l+m) \\
=&f(x)\begin{cases}
\frac{\beta SI}{N}\Delta t & (k,j,m) = (-1,1,0) \\
kE\Delta t & (k,j,m) = (0,-1,1) \\
\gamma I\Delta t & (k,j,m) = (0,0,-1) \\
1 - (\frac{\beta SI}{N} + kE + \gamma I)\Delta t & \text{otherwise}
\end{cases}
\end{aligned} \tag{3}$$

And the stochastic SEIR model has the covariance matrix as follows:

$$\begin{aligned}
&E[(\Delta x)(\Delta x)] \\
&= \sum_{i=1}^{3} p_i(\Delta x_i)(\Delta x_i)^T \\
&= \begin{pmatrix}
\beta\frac{x_1x_2}{N} & -\beta\frac{x_1x_2}{N} & 0 \\
-\beta\frac{x_1x_2}{N} & \beta\frac{x_1x_2}{N} + kx_2 & -kx_2 \\
0 & -kx_2 & kx_2 + \gamma x_3
\end{pmatrix}\Delta t
\end{aligned} \tag{4}$$

3.2 Extended Kalman Filter Application

In estimation theory, the extended Kalman filter (EKF) is the nonlinear version of the Kalman filter which linearizes about an estimate of the current mean and co-variance. In EKF, the state transition and observation models don't need to be linear functions of the state but may instead be differentiable functions. For a nonlinear system

as described below, EFK assumes that the random variable w_k captures the process uncertainties in the model and v_k, uncorrelated to w_k, captures the measurement noise.

$$\begin{aligned} x_k &= f_(x_{k-1}, u_k) + w_k \\ z_k &= h(x_k) + v_k \end{aligned} \tag{5}$$

where x_k is the n × 1 State vector ; w_k is the n × 1 Process noise vector ; z_k is the m × 1 Observation vector v_k is the m × 1 Measurement noise vector ; $f(.)$ is the n × 1 Process nonlinear function ; $h(.)$ is the m × 1 Observation nonlinear vector function ; Q_k is the n × n Process noise covariance matrix ; R_k is the m × m Measurement noise covariance matrix.

Here S_t and v_t are the process and observation noises which are both assumed to be zero mean multivariate Gaussian noises with co-variance Q_t and R_t and u_t is the control vector. The function f can compute the predicted state from the previous estimate and the function h can compute the predicted measurement from the predicted, but f and h can't be used for the co-variance directly and instead a matrix of partial derivatives (the Jacobian function) is computed. At each time step, the Jacobian is computed with the current predicted states. These matrics can be used in the Kalman filter equations. This process essentially linearizes the non-linear function around the current estimate.

The process of estimating the parameters is what would lead to the cone of uncertainty in the output predictions. This class of problems are very well suited to a Kalman Filter treatment, since we can explicitly propagate errors (assuming Gaussian distribution of model errors, and a linear state model, which this is). We thus applied this Extended Kalman Filter method to our SEIR stochastic model, in order to filter out the noise and capture the true state change. The Susceptible (S), Exposed (E), and Infected (I) comprise the state vector:

$$x = [S, E, I]^T \tag{6}$$

The state in time k can be predicted by the previous state in time $k-1$ as:

$$x_k = \begin{bmatrix} S_k \\ E_k \\ I_k \end{bmatrix} = \begin{bmatrix} S_{k-1} - \frac{\beta S_{k-1} I_{k-1}}{N} \Delta t \\ E_{k-1} + \frac{\beta S_{k-1} I_{k-1}}{N} \Delta t - k E_{k-1} \Delta t \\ I_{k-1} + k E_{k-1} \Delta t - \gamma I_{k-1} \Delta t \end{bmatrix} \tag{7}$$

The process noise comes from the stochastic nature of the stochastic SEIR model. We use the covariance matrix of the process noise as:

$$Q_k = \begin{bmatrix} \beta \frac{x_1 x_2}{N} & -\beta \frac{x_1 x_2}{N} & 0 \\ -\beta \frac{x_1 x_2}{N} & \beta \frac{x_1 x_2}{N} + k x_2 & -k x_2 \\ 0 & -k x_2 & k x_2 + \gamma x_3 \end{bmatrix} \Delta t \tag{8}$$

And we assume the measurement noise matrix is:

$$R_k = \begin{bmatrix} 100 \\ 10 \\ 1 \end{bmatrix} I_{3\times 3} \tag{9}$$

We use the Extended Kalman Filter to simulate the true SEIR state transition process with arbitrary initial parameter β_0, the initial transmission rate parameter, Q_0, the initial process noise matrix. And we use Cross-Validation to tune the optimal initial transmission rate parameter β_0. The output of EKF is a time series of the simulated transmission rate, with which we can calculate our estimated R0 based on equation (3.6)

3.3 Reproduction Number R Time Series Prediction

Based on the simulated basic reproduction number (R) time series from EKF, we then predict future R using multiple techniques including Long short-term memory (LSTM), Gated recurrent unit (GRU), Variational-LSTM Autoencoder (VAE LSTM), Variational Autoencoder-GRU-based (VAE GRU) and Bayesian Structural Time Series (BSTS). We chose these models as baseline comparisons according to their superior performance in related work. With the applications of and comparisons between the deep learning strategies and Bayesian method, we are able to figure out the advantages and disadvantages of these methods under different situations. Time series models can be broadly categorized as generative and discriminative, depending on how the target is modeled [14]. Generative models, such as ARIMA, ETS and BSTS, assumes that the time series are generated from an unknown stochastic process and can be described by estimating the parameters of this stochastic process with the maximized likelihood. In contrast to these models, neural generative models expresses the parameters as a function of a neural network whose weights are shared among all time series and learned from the whole training data [2]. Discriminative models, on the other hand, model the conditional distribution directly via neural networks. Since discriminative models assume less structural characteristics, they are thus more flexible and applicable to a broader class of application domains compared to generative models [2].

The neural network architectures are described in Figure 2. As shown, we implement the LSTM model with three LSTM layers and one Dense layer. There are 25 neurons in the input LSTM layer with an activation function as ReLU; there are 15 neuron in the first hidden LSTM layer with an activation function as ReLU; there are 10 neurons in the second hidden LSTM layer with an activation function as SeLU. The model is compiled with Mean Absolute Error (MAE) loss function and the efficient Adam version of stochastic gradient descent. This model is fitted with 50 training epochs with a batch size of 20. The GRU model uses a similar architecture except that it is implemented with three GRU layers and one Dense layer. All the hyper-parameters above are tuned by the Grid Search with Cross Validation. The VAE LSTM The model comprises two branches trained in parallel in an end-to-end fashion. "VAE LSTM" is a mixture of VAE and LSTM, which comprises two ches trained in parallel, while "VAE GRU" is a mixture of VAE and GRU. Basically, the VAE is considered as an autoencoder whose training is regularised to bypass the overfitting problem and establish a latent space with suitable properties enabling the generative process.Similar to a traditional autoencoder, a VAE contains both an encoder and a decoder.

The Errors by Epochs of LSTM and GRU on USA R0 time series forecasting is shown in Figure below:

We first fit the three models (LSTM, GRU, BSTS) on estimated USA basic reproduction number R (daily R time series from April 1st, 2020 to **October 8th, 2020**) and compared their performance (See Table 1).

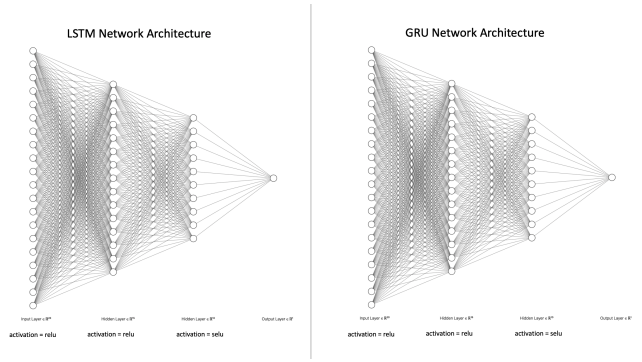

Figure 2: Neural Network Architecture

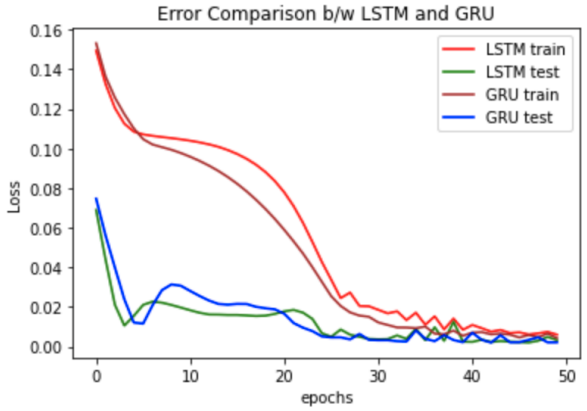

Figure 3: Errors of LSTM and GRU Comparison

Errors	LSTM	GRU	BSTS	vae lstm	vae gru
MAPE	0.0061	0.0049	0.0823	0.0081	0.0081
RSME	0.0097	0.0081	0.1144	0.0111	0.0109

Table 1: Errors of different models on USA R forecasting

As shown in the error table, the neural network models are much better than the baseline BSTS model and better than the VAE encoded neural network models. And the GRU-based model performs slightly better than the LSTM-based model.

To further analyze the applicability and predictive accuracy of each model based on data characteristics such as seasonal variation, we run the models on state-level data. As a result, for data with stronger seasonal variation, our GRU neural network performs better than our LSTM neural network; for data with smaller seasonal variation, our LSTM neural network performs better. We illustrate this using the 52 US state R time series data. The seasonal variation indicator $SVI$is calculated using the formula below:

$$\textbf{SVI} = \frac{1}{n-6}\sum_{i=7}^{n}(x_i - MA_{i-6})^2$$

The MA_j in the formula represents the moving average by 7-day window $MA_j = \frac{1}{7}(x_j + x_{j+1} + ... + x_{j+6})$, j = 1, 2, ..., n-6.

The states with data characteristics of stronger seasonal variation (28 states out of 52 states) has lower MAPE error using GRU neural network instead of LSTM neural network has lower MAPE error using LSTM neural network instead of GRU neural network. The difference between the mean seasonal variations indicator of the two groups are statistically significant (p = 0.0138 < 0.05) (See Table 2). Therefore, we can conclude the applicability of our models that, we should use LSTM neural network for data with weaker seasonal variation (Seasonal Variation Indicator < 0.1); and we should use GRU neural network for data with strong seasonal variation (Seasonal Variation IndicatorSVI >= 0.1). Thus, we establish our model selection model with this logic (See Figure 5) to use the most accurate model based on data characteristics.

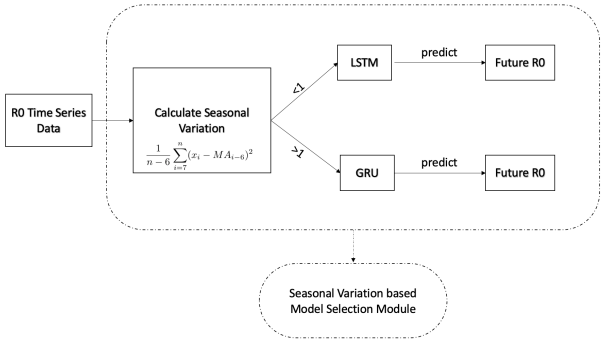

Figure 4: Model Selection Module

	Value
LSTM Mean SVI (24 states)	0.0838
GRU Mean SVI (28 states)	0.1354
T test score	-2.5529
T test p value	0.0138

Table 2: Model Performance vs Seasonal Variation

After we put the data seasonal variation characteristics into the consideration of model selection, the predictive accuracy was significantly increased. We calculate the average MAPE errors of 52 USA states R time series prediction with different models (See Table 3). The baseline BSTS model has the largest MAPE error at 6.09%; the LSTM neural network and GRU neural network have similar average MAPE errors – 2.26%, 2.18% respectively. But our seasonal-variation-based model selection module decreased the average MAPE error to 1.41%, which is a significant increase in predictive accuracy. In this case, even though the VAE-autoencoded GRU neural network has the lowest average MAPE Error (1.71%) as a single model, our seasonal-variation-based model selection module still beats it.

Model	BSTS	LSTM	GRU
Avg Mape	0.0609	0.0226	0.0218
Model	VAE LSTM	VAE GRU	LSTM/GRU
Avg Mape	0.0354	0.0171	0.0141

Table 3: Error Comparison across models

3.4 Real-life Applications and Policy Recommendation

The basic reproduction number, R_0, one of the most well-known thresholds in deterministic epidemic theory, predicts a disease outbreak when $R_0 > 1$ and predicts a disease extinction when $R_0 < 1$ [9]. While we have implemented the stochastic view of SEIR model using Extended Kalman Filter to simulate the true R trend (instead of R_0), we need to find the R threshold in terms of stochastic epidemic theory that predict a major outbreak. Allen and Lahodny in their research have found that "If $R_0 > 1$" and i infectious individuals are introduced into a susceptible population, then the probability of a major outbreak is approximately $1 - (1/R_0)^i$ [6]. Based on this formula, as well as the particular data features of USA Covid-19 data, we set our R threshold as follows:

$$\begin{cases} R < 0.8 & \text{Extinction of Disease} \\ 0.8 \leq R \leq 1.2 & \text{Boarderline} \\ R > 1.2 & \text{Disease Outbreak} \end{cases} \tag{10}$$

Accordingly, we suggest policy recommendations as below:

$$\begin{cases} R < 0.8 & \text{Minimum level lockdown and social distancing policy} \\ 0.8 \leq R \leq 1.2 & \text{Medium level lockdown and social distancing policy} \\ R > 1.2 & \text{Maximum level lockdown and social distancing policy} \end{cases} \tag{11}$$

In addition, when $0.8 \leq R \leq 1.2$, we can further consider two scenarios, including an upward trend and a downward trend in the recent 15 days. When it is trending upward in the past 15 days, we suggest stricter controlling measures against COVID-19 as it's more likely to go beyond 1.2 soon. However, when there is a downward trend, we suggest looser controlling measures as it could go below 0.8. Here we only talk about the high-level ideas as it really depends on the real situation to act on.

4 DISCUSSION

One challenging part of our work is the data. As COVID-19 is an emerging, rapidly evolving situation, the limited number of data hampers modeling and forecasting R0. To better forecast COVID-19, one further research is to generate more data points and model using the generated dataset. We have done some research and experiments on modeling generated dataset, and we will explain it in greater detail in another paper.

5 CONCLUSION

In this study, we propose an modification of the standard SEIR model with policy and risk factors. We re-formulate SEIR model from deterministic differential equations to SDEs, apply extended Kalman Filter to calculate the reproduction number at each time-step and then predict the future reproduction number with various deep learning models. Furthermore, we conduct our research on two different granular levels including US national and state-level and provide policy recommendations in order to help policy makers to better understand the current situation and prepare for controlling the COVID-19.

REFERENCES

[1] 2020. Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19). *HEALTHCARE WORKERS* (2020). https://doi.org/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html

[2] Benidis K. Schneider M. B. Flunkert V. Gasthaus J. Januschowski T. Maddix D. Rangapuram S. Salinas D. Schulz J. Stella L. Turkmen A. C. Wang Y. Alexandrov, A. 2019. Neural time series models with GluonTS Time Series Workshop ICML 2019. (2019).

[3] Allen L.J.S. Arciniega A. Allen, E.J. and P.E. Greenwood. 2008. Construction of Equiva- lent Stochastic Differential Equation Models. *Stochastic Analysis and Applications* (2008).

[4] E.J. Allen. 2007. *Construction of Equivalent Stochastic Differential Equation Models.*

[5] L.J.S. Allen and E.J. Allen. 2003. A Comparison of Three Different Stochastic Population Models with Regard to Persistence Time. *Theoretical Population Biology* (2003).

[6] Lahodny G. E. Allen, L.J.S. 2012. Extinction thresholds in deterministic and stochastic epidemic models. *Journal of Biological Dynamics* 6, 2 (2012), 590–611.

[7] Sebastian Funk and Rosalind M. Eggo. 2020. Early dynamics of transmission and control of 2019-nCoV: a mathematical modelling study. (2020).

[8] 2020. Variational-LSTM Autoencoder to forecast the spread of coronavirus across the globe. (2020).

[9] O. Jong, M. C. M.Diekmann. 1992. A method to calculate—for computer-simulated infections—the threshold value, R0, that predicts whether or not the infection will spread. *Preventive Veterinary Medicine* 12, 3–4 (1992), 269–285.

[10] P.E. Kloeden and E. Platen. 1999. *Numerical Solution to Stochastic Differential Equations.* Springer, Berlin.

[11] Phenyo E. Lekone and Bärbel F. Finkenstädt. 2006. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics* 62, 4 (2006), 1170–1177.

[12] I. S. Mbalawata Ndanguza, D. and J. P. Nsabimana. 2016. Analysis of sdes applied to seir epidemic models by extended kalman filter method. *Applied Mathematics* 7, 17 (2016), 2195.

[13] Mbalawata2 I. S. Nsabimana1 J. P. Ndanguza1, D. 2016. Analysis of SDEs Applied to SEIR Epidemic Models by Extended Kalman Filter Method. *Applied Mathematics* (2016).

[14] A. Y. Ng and M. I. Jordan. [n.d.]. .

[15] New York Times. 2020. Coronavirus (Covid-19) Data in the United States. (2020). https://doi.org/nytimes/covid-19-data

[16] et al Vollmer, Michaela AC. 2020. A sub-national analysis of the rate of transmission of COVID-19 in Italy. (2020).

[17] Amanat F. Firpo A. Altman D. Bailey M. Mansour M. McMahon M. Meade P. Mendu D. R. Muellers K. Stadlbauer D. Stone K. Strohmeier S. Aberg J. Reich D. Krammer F. Cordon-Cardo C. Wajnberg, A. 2020. SARS-CoV-2 infection induces robust, neutralizing antibody responses that are stable for at least three months. *medRxiv* (2020). https://doi.org/10.1101/2020.07.14.20151126

[18] Jiangzhuo Chen Wang, Lijing and Madhav Marathe. 2019. Deep learning based epidemic forecasting with synthetic information. *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (2019).

[19] et al Zeroual, Abdelhafid. 2020. Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons Fractals* (2020).