Modeling the Omicron Dynamics and Development in China: with a Deep Learning Enhanced Compartmental Model

Qi Deng [123*]

1. College of Artificial Intelligence
Hubei University of Automotive Technology

2. Division of Business Management,
Beijing Normal University-Hong Kong Baptist University United International College

3. Cofintelligence Fintech Co. Ltd., Hong Kong and Shanghai, China

*. Corresponding author. Email: qideng@uic.edu.cn; qi.deng@cofintelligence.com

Abstract

Background: Compartmental models dominate epidemic modeling. Estimations of transmission parameters between compartments are typically done through stochastic parameterization processes that depend upon detailed statistics on transmission characteristics, which are economically and resource-wide expensive to collect.

Objectives: We apply deep learning techniques as a lower data dependency alternative to estimate transmission parameters of a customized compartmental model, for the purpose of simulating the dynamics of the Omicron phase of the COVID-19 epidemics and projecting its further development in China and subregions within the country.

Methods: We construct a compartmental model, and develop a multivariate, multistep deep learning methodology to estimate the model's transmission parameters. We then feed the estimated transmission parameters to the compartmental model to predict the development of the COVID-19 epidemics in China and subregions within the country for 28 days.

Results: In China (excluding Hong Kong and Taiwan), the daily Omicron infection increase is between 60 and 260 in the 28-day forecast period between June 4 and July 1, 2022. On July 1, 2022, there would be 768,622 cumulative confirmed cases and 591 cumulative deceased cases. The CFR would stabilize at 0.077%±0.00025%. Assuming a 25% infection rate, the total deaths with Omicron would be up to 280,000 without non-pharmaceutical intervention (NPI).

Conclusions: Current compartmental models require stochastic parameterization to estimate the transmission parameters. These models' effectiveness depends upon detailed statistics on transmission characteristics. As an alternative, deep learning techniques are effective in estimating these stochastic parameters with greatly reduced dependency on data particularity.

Introduction

The Omicron variant of COVID-19 started invading China since as early as November 2021, and became a full fledge epidemic since late February 2022. Assuming that all confirmed infections since November 31, 2022 are Omicron cases, as of June 3, 2022, there have been 764,369 confirmed Omicron cases, and 590 patients have died with Omicron infection in China (excluding Hong Kong and Taiwan). The numbers of confirmed and deceased case are 1,214,192 and 9,382, respectively, in Hong Kong. For Taiwan, the numbers are 2,274,666 and 2,663. The hardest hit city in China is Shanghai, which saw 63,035 infections and 595 deaths. The Omicron epidemic is still evolving in China at the time of this study.

Most of the research on Omicron focuses on effectiveness of immunization, vaccination and treatment, with relatively few epidemiological studies on the variant, especially in a Chinese context. A computational simulation-model study utilizing a customized Monte Carlo model to estimate the effect of facemask use before and after different COVID-19 vaccination coverage levels is conducted [1]. A set of posterior statistical models to estimate cumulative infections and cumulative proportion of the populations in worldwide locations is produced [2]. While epidemiological study on Omicron is lacking in general, there have been attempts to model the original variant dynamics in China. A classic SEIR model is used to infer the basic reproduction ratio, and to simulate the Wuhan epidemic [3]. More sophisticated models have been developed to correlate risk levels of foreign countries with their travel exposure to China [4, 5], including a stochastic dual-SEIR approach on both Wuhan population and international travelers to estimate how transmission have varied over time from Wuhan to international destinations [5]. Simulations on international spread after the start of travel ban from Wuhan on Jan 23, 2020 have also been conducted [6], which apply the Global Epidemic and Mobility Model (GLEAM) to a multitude of Chinese and international cities, and a SEIR variety (SLIR) to project the impact of human-to-human transmissions.

Since March 2020, with the outbreak of the original variant winding down in China, researchers have dedicated more efforts in analyzing the effectiveness of containment measures. Mobility and travel history data from Wuhan is used to ascertain the impact of the drastic control measures implemented in China [7]. The spread and control of COVID-19 among Chinese cities with data on human movements and public health interventions is investigated [8]. A transmission model to study the impact of social distancing and school closure is built, which utilizes the contact data for Wuhan and Shanghai and contact tracing information from Hunan Province [9].

In late February 2022, China, especially Shanghai, was hit hard by the Omicron variant. With a much higher initial reproduction rate than that of the original and Delta variants at between 1.72 and 8.2 [10, 11, 12], the Omicron variant presents a potentially game-changing challenge to the country's zero-COVID strategy that was effective against the original and Delta variants. An age-structured stochastic compartmental model (SLIRS) calibrated on the initial growth phase for the 2022 Omicron outbreak in Shanghai is developed [13]. The key contribution of the model is the inclusion of age-specific vaccine coverage data, vaccine efficacy against different clinical endpoints, waning of immunity, different antiviral therapies, and non-pharmaceutical interventions.

The aim of this paper is to establish a class of innovative compartmental models, of which the transmission parameters are estimated by a family of multivariate, multistep deep learning methodologies. The models are then used to predict and simulate the dynamics of Omicron in China (both including and excluding Hong Kong and Taiwan), as well as particularly regions including Shanghai, Hong Kong and Taiwan.

Methods

Theoretical Foundation: Compartmental Models

Compartmental models require detailed statistics on transmission characteristics to estimate the stochastic transmission parameters between compartments. Essentially, these models correlate factors such as geographic distances and contact intensities among heterogeneous subpopulations with gradient probability decay. Technically, transmission parameterization utilizes Bayesian inference methods, such as Marcov Chain Monte Carlo (MCMC) or Gillespie algorithm [14] simulations to form probability density functions (PDFs) on cross-section, in order to estimate transmission parameters for each timestep of a compartmental time series. These detailed statistics on transmission characteristics are economically and resource-wise expensive to collect. As an alternative, some researches [e.g., 13] simply assume the values of these transmission parameters to achieve cost-effectiveness.

We are particularly interested in compartmental models that cover multiple inter-connected and heterogeneous subpopulations [9, 15, 16]. We develop a multistep, multivariate deep learning methodology to estimate the transmission parameters. We then feed these estimated transmission parameters to a class of compartmental models to predict and simulate the development of the Omicron epidemic in China and subregions within the country.

We establish a SIR-derived (SIRD) discrete time series on a daily interval as the theoretical foundation for a deep learning-enhanced compartment model. The reason we do not use more sophisticated models such as the SEIR-variety is the absence of time series data other than the confirmed and deceased cases, which actually strengthens our argument that deep learning is effective when detailed statistics on transmission characteristics are not readily available. A precursor to this study is developed to predict and simulate the dynamics and development of the original COVID-19 variant in the US [17].

The SIRD (Susceptible-Infectious-Removed-Deceased) construct groups a population is into four compartments:

1. Susceptible (S): The susceptible population that progresses into the Infectious compartment.
2. Infectious (I): The infectious individuals that are symptomatic, come from the Susceptible compartment, and progress into the Removed compartment.
3. Removed (R): The recovered individuals, come from Infectious compartment and acquire lasting immunity (there has yet any contradiction against this assumption for Omicron).
4. Deceased (D): The deceased cases come from Infectious compartment.

The SIRD model has a discrete daily ($\Delta t = 1$) multivariate time series construct given by the follow matrix form:

$$\begin{bmatrix} S_{t+1} \\ I_{t+1} \\ R_{t+1} \\ D_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & -\beta & 0 & 0 \\ 0 & 1+\beta-\gamma^R-\gamma^D & 0 & 0 \\ 0 & \gamma^R & 1 & 0 \\ 0 & \gamma^D & 0 & 1 \end{bmatrix} \begin{bmatrix} S_t \\ I_t \\ R_t \\ D_t \end{bmatrix} \tag{1}$$

or:

$$\overrightarrow{V_{t+1}} = \overleftrightarrow{A_t}\,\overrightarrow{V_t} \tag{2}$$

$$where\ \overrightarrow{V_{t+1}} = \begin{bmatrix} S_{t+1} \\ I_{t+1} \\ R_{t+1} \\ D_{t+1} \end{bmatrix}, \quad \overleftrightarrow{A_t} = \begin{bmatrix} 1 & -\beta & 0 & 0 \\ 0 & 1+\beta-\gamma^R-\gamma^D & 0 & 0 \\ 0 & \gamma^R & 1 & 0 \\ 0 & \gamma^D & 0 & 1 \end{bmatrix} and\ \overrightarrow{V_t} = \begin{bmatrix} S_t \\ I_t \\ R_t \\ D_t \end{bmatrix}$$

The Greek letters $\beta, \gamma^R, \gamma^D$ in Equations 1 and 2 are the "susceptible-to-infectious," "infectious-to-removed," and "infectious-to-deceased" transmission parameters, respectively.

Since we need to estimate the transmission parameters, we rewrite and rearrange Equations 1 and 2 to the following matrix representation:

$$\begin{bmatrix} \Delta S_{t+1} \\ \Delta I_{t+1} \\ \Delta R_{t+1} \\ \Delta D_{t+1} \end{bmatrix} = \begin{bmatrix} -I_t & 0 & 0 \\ I_t & -I_t & -I_t \\ 0 & I_t & 0 \\ 0 & 0 & I_t \end{bmatrix} \begin{bmatrix} \beta \\ \gamma^R \\ \gamma^D \end{bmatrix} \tag{3}$$

or:

$$\overrightarrow{\Delta V_{t+1}} = \overleftrightarrow{B_t}\,\vec{\mathrm{I}_t} \tag{4}$$

$$where\ \overrightarrow{\Delta V_{t+1}} = \begin{bmatrix} \Delta S_{t+1} \\ \Delta I_{t+1} \\ \Delta R_{t+1} \\ \Delta D_{t+1} \end{bmatrix}, \overleftrightarrow{B_t} = \begin{bmatrix} -I_t & 0 & 0 \\ I_t & -I_t & -I_t \\ 0 & I_t & 0 \\ 0 & 0 & I_t \end{bmatrix} and\ \vec{\mathrm{I}_t} = \begin{bmatrix} \beta \\ \gamma^R \\ \gamma^D \end{bmatrix}$$

<u>Theoretical Foundation: Parameterization Process with Deep Learning and SIRD Simulation</u>

The transmission parameters $(\beta, \gamma^R, \gamma^D)$ in Equations 1 to 4 are both non-stochastic values in the temporal dimension (*t*), and stochastic variables along three "spatial parameter dimensions," namely population distribution (*S*), population mobility (*L*), and population heterogeneity (*C*). A parameterization process to estimate the transmission parameters at each time step (cross-section in the multivariate SIR time series construct) is therefore required, and has the following expression:

$$F_{tp} = F_{tp}(t, S, L, C) \in (\beta, \gamma^R, \gamma^D) \tag{5}$$

Equation 5 shows that each transmission parameter (F_{tp}) can be modeled in a 4-dimentional tempo-spatial framework. The parameterization process is thus to estimate the in-sample values of cross-sectional F_{tp} at each time step *t* in the SIRD time series construct, and predict its out-of-sample values.

We aim to build a multi-step, multivariate deep learning method to estimate the transmission parameters, utilizing both the standard deep neural network (DNN) and the advanced recurrent neural network – long short-term memory neural network (RNN-LSTM) methodologies. We propose the following steps to achieve this goal:

1. Constructing the in-sample SIRD time series using observed Omicron data.
2. Calculating in-sample daily transmission parameters from the in-sample SIRD time series constructed in Step 1.

3. Decomposing Equation (5) as: $F_{tp} = F_{tp}(t, S, L, C) = F_{tp}(t)\Psi(S, L, C)$. That is, at given time step t, along the temporal dimension, the non-stochastic value of the transmission parameter is $F_{tp}(t)$; along the spatial dimensions (S, L, C), the cross-sectional probability distribution of the transmission parameter is $\Psi(S, L, C)$.
4. Applying deep learning algorithms (DNN and RNN-LSTM) to fit the in-sample decomposed transmission parameters in step 3. Deep learning is performed on both $F_{tp}(t)$ and $\Psi(S, L, C)$ to calibrate the in-sample values of $F_{tp}(t, S, L, C)$ along both temporal and spatial dimensions, respectively.
5. With the in-sample transmission parameters obtained in step 4, applying DNN and RNN-LSTM algorithms again, in both progressive and recursive manners, to predict the out-of-sample transmission parameters for multiple scenarios.
6. Simulating out-of-sample Omicron dynamics recursively through the SIRD time series, using the out-of-sample transmission parameters predicted in step 5.

The methodological innovation of our research is mainly reflected in deep learning of the cross-sectional probability distribution $\Psi(S, L, C)$. Along the three spatial parameter dimensions of S, L, and C, the change of $\Psi(S, L, C)$ can be expressed "traditionally" as follows:

Continuous: $d\Psi(S, L, C)dSdLdC$ (6.1)

Discrete: $\Delta\Psi(S, L, C) = \frac{\partial\Psi(S,L,C)}{\partial S\partial L\partial C}\Delta S\Delta L\Delta C$ (6.2)

Equations (6.1, 6.2) have the following three characteristics:

1. $\Psi(S, L, C)$is differentiable in the (S, L, C) domain.
2. $d\Psi(S, L, C)$, $\Delta\Psi(S, L, C)$are first-order expression (linear), rather than higher-order expansions (polynomial).
3. Out-of-sample $d\Psi(S, L, C)$, $\Delta\Psi(S, L, C)$ are deterministic, and their "probability distribution" only comes from the in-sample randomness of $\Psi(S, L, C)$.

The above 2nd and 3rd characteristics make it meaningless to use Equations 6.1 and 6.2 to predict changes in $\Psi(S, L, C)$: the 2nd characteristic makes it impossible for the model to extract high-dimensional (non-linear) features that the observed data can provide; and the 3rd characteristic makes the out-of-sample cross-sectional prediction non-stochastic, thus the model can only be "mechanical." As such, using deep learning algorithms to fit and predict changes in $\Psi(S, L, C)$ becomes a natural choice. First of all, deep learning is essentially nonlinear (higher-order polynomial) fitting, so that the high-dimensional information contained in the observed data can be fully extracted; second, calibration of hyperparameters in deep learning preserves the stochastic probability distribution of out-of-sample predictions.

The DNN decomposition of the parameter f_{tp} is expressed as:

$$F_{tp} = F_{tp}(t, S, L, C) = F_{tp}(t)\Psi(S, L, C) \in (\beta, \gamma^R, \gamma^D) \tag{7.1}$$

$$f_{tp} = \ln(F_{tp}) = \ln\left(F_{tp}(t)\Psi(S, L, C)\right) = \ln\left(F_{tp}(t)\right) + \ln(\Psi(S, L, C)) \tag{7.2}$$

$$\Longrightarrow f_{tp} = f_{tp}(t) + \psi(S, L, C) \tag{7.3}$$

$$f_{tp} = a + \Theta\left(\sum_{j\in[1,J]} w_j x_{jt}\right) + \Phi\left(\sum_{k\in(S,L,C)} w_k x_{kp}\right) + \epsilon_t \tag{7.4}$$

$$\Longrightarrow f_{tp}(t) = a + \Theta\left(\sum_{j\in(1,J)} w_j x_{jt}\right) \tag{7.5}$$

$$\Longrightarrow \psi(S, L, C) = \Phi\left(\sum_{k\in(S,L,C)} w_k x_{kp}\right) + \epsilon_t \tag{7.6}$$

f_{tp} is the observed (actual) parameter value (logarithm) at time t, $f_{tp}(t)$ is its time dimension decomposition, and $\psi(S, L, C)$ is its parameter dimension decomposition;
a is the regression intercept in the time range T ($t\in[0,T]$), and ϵ_t is the regression residual at time t;
x_{jt} is the value of the jth time factor ($j\in[1,J]$) at time t, and w_j is its regression coefficient, i.e., the weight;
x_{kp} is the value of the kth factor ($k\in(S,L,C)$) in the parameter dimension, w_k is the regression coefficient, i.e., the weight;
$\Theta(\cdot)$ is the "overall" activation function in the time dimension and $\Phi(\cdot)$ is the "overall" activation function in the parameter dimension.

For in-sample data, f_{tp} is the actual (observed) transmission parameter value (logarithmic) of the SIRD time series, $f_{tp}(t)$ is the non-stochastic parameter value (logarithmic) in the temporal dimension, and $\psi(S, L, C)$ is the cross-sectional probability distribution in the spatial dimensions (logarithmic). The importance of decomposing f_{tp} into $f_{tp}(t)$ and $\psi(S, L, C)$ is to independently calibrate the temporal non-stochastic value and the spatial probability distribution of f_{tp}, in order to further estimate its out-of-sample values.

The RNN-LSTM decomposition of f_{tp} is expressed as:

$$F_{tp} = F_{tp}(t, S, L, C) = F_{tp}(t)\Psi(S, L, C) \in (\beta, \gamma^R, \gamma^D) \tag{8.1}$$

$$f_{tp} = \ln(F_{tp}) = \ln\left(F_{tp}(t)\Psi(S, L, C)\right) = \ln\left(F_{tp}(t)\right) + \ln(\Psi(S, L, C)) \tag{8.2}$$

$$\Longrightarrow f_{tp} = f_{tp}(t) + \psi(S, L, C) \tag{8.3}$$

$$f_{tp} = a + \Gamma\left(\sum_{\tau\in[1,t-1]} w_{t-\tau} f_{tp}(t-\tau)\right) + \Theta\left(\sum_{j\in[1,J]} w_j x_{jt}\right) + \Phi\left(\sum_{k\in(S,L,C)} w_k x_{kp}\right) + \epsilon_t \tag{8.4}$$

$$\Longrightarrow f_{tp}(t) = a + \Gamma\left(\sum_{\tau\in[1,t-1]} w_{t-\tau} f_{tp}(t-\tau)\right) + \Theta\left(\sum_{j\in(1,J)} w_j x_{jt}\right) \tag{8.5}$$

$$\Longrightarrow \psi(S, L, C) = \Phi\left(\sum_{k\in(S,L,C)} w_k x_{kp}\right) + \epsilon_t \tag{8.6}$$

f_{tp} is the observed (actual) parameter value (logarithm) at time t, $f_{tp}(t)$ is its time dimension decomposition, and $\psi(S, L, C)$ is its parameter dimension decomposition;
a is the regression intercept in the time range T ($t\in[0,T]$), and ϵ_t is the regression residual at time t;
x_{jt} is the value of the jth time factor ($j\in[1,J]$) at time t, and w_j is its regression coefficient, i.e., the weight;
x_{kp} is the value of the kth factor ($k\in(S,L,C)$) in the parameter dimension, w_k is the regression coefficient, i.e., the weight;
$\Theta(\cdot)$ is the "overall" activation function in the time dimension and $\Phi(\cdot)$ is the "overall" activation function in the parameter dimension;
$f_{tp}(t-\tau)$ is the observed (actual) parameter value (logarithm) ($\tau\in[1,t-1]$) value at time t-τ, and $w_{t-\tau}$ is its autoregressive coefficient, or the weight;
$\Gamma(\cdot)$ is the temporal autocorrelation "overall" activation function of $f_{tp}(t)$, representing a nonlinear higher-order polynomial fit of the autoregression.

The difference between Equations 7 and 8 is that RNN-LSTM includes a temporal autocorrelation of $f_{tp}(t)$ itself, so that it is not only a sequence in the parameter dimensions, but also a time series.

Data

We collect the following COVID-19 datasets from two sources:

1. Dataset 1: A JHU CSSE dataset (https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/), which tracks confirmed cases and deceased cases. We use the confirmed/deceased dataset to form training data for deep learning.
2. Dataset 2: A Tencent dataset (https://view.inews.qq.com/g2/getOnsInfo?name=disease_other), which updates daily records (confirmed, active, deceased, recovered, etc.). We use these detailed case data to construct the compartmental model.

In general, both datasets have some reporting discrepancies, with certain extreme outliers in both directions, thus we run a 7-day moving average on the datasets to smooth out these data irregularities.

Modeling Methodology

We then conduct the following step-by-step operations to model the Omicron epidemic in China (including and excluding Hong Kong and Taiwan), Shanghai, Hong Kong and Taiwan. Fig. 1 is the flowchart to illustrate the modeling methodology.

1. We construct a confirmed/deceased time series starting from March 1, 2022 (in-sample) from Dataset 1. March 1, 2022 is generally accepted as the outbreak point of Omicron epidemic in China [13].
2. We apply two deep learning approaches, the standard DNN (Deep Neural Networks) and the advanced RNN-LSTM, in both progressive and recursive fashions, to fit the confirmed/deceased in-sample time series from Step 1, and predict the further development of confirmed/deceased cases for 28 days (out-of-sample).
3. We construct an in-sample SIRD time series starting March 1, 2022 from Dataset 2.
4. We use the in-sample SIRD time series constructed in Step 3 to come up with in-sample sequences for the SIRD daily transmission parameters (β, γ^R and γ^D).
5. We then use the confirmed/deceased time series (in-sample and out-of-sample) from Step 2 as training data, the in-sample β, γ^R and γ^D sequences from Step 4 as training label, and apply the DNN and RNN-LSTM techniques to predict β, γ^R and γ^D for 28 days (out-of-sample).
6. Finally, we use the predicted (out-of-sample) transmission parameters (β, γ^R and γ^D) from Step 5 to simulate the 28-day progression (out-of-sample) of the SIRD model in a recursive manner, starting from the last timestep from the in-sample SIRD time series from Step 4.
7. We then repeat Steps 1-6 for Shanghai (hardest hit area by Omicron within China), Hong Kong (hit by Omicron roughly 1 month before Shanghai), and Taiwan (hit by

Omicron roughly 1 month before Shanghai), in order to test the robustness of the model with data from different phases (and populations) of the epidemic.

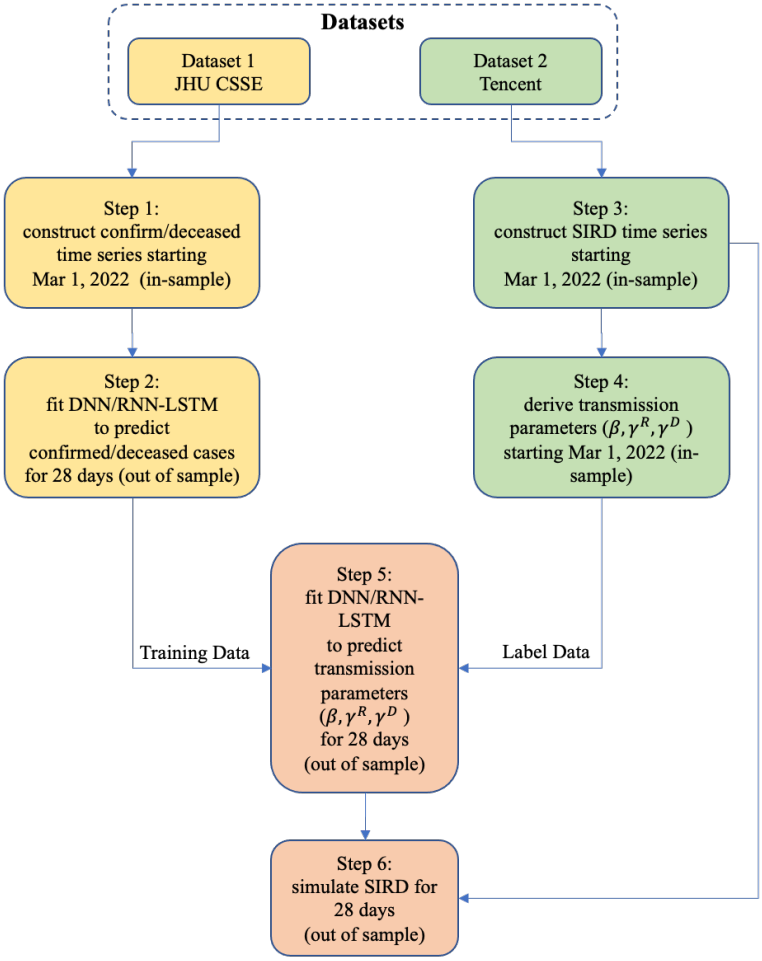

Fig. 1. The methodology flowchart

Results

The average results of eight models (scenarios with different learning hyperparameters) based on data up to June 3, 2022 are illustrated in Figs. 2 – 5 (28-day forecast).

We predict that, in the 28-day forecast period, in China (excluding Hong Kong and Taiwan), the daily Omicron infection increase is between 60 and 260. On July 1, 2022, there would be 768,622 cumulative confirmed cases and 591 cumulative deceased cases, with a case death rate, or case fatality ratio (CFR) at 0.07688%. Furthermore, the predicted CFR is stable over our 28-day forecast time-period (0.077%±0.00025%). The significance of this finding is that, based on the predicted CFR, we can estimate the worst-case scenario of the total number

of infections of Omicron, as well as the total number of deaths with Omicron in China. With a total population of 1.4 billion, and an assumption of a 25% infection rate without nonpharmaceutical intervention (this assumption is based on the observed US infection rate at around 25% by May 31, 2022), the total Omicron-infected population of China (NPI) would eventually reach 350 million. The total deaths with Omicron would be up to 280,000 (assuming a more severe 0.08% CFR than the 0.077% estimate).

Using the same methods, we predict that, on July 1, 2022, in Shanghai, there would be 60,540 cumulative confirmed cases and 589 cumulative deceased cases, with a stabilized CFR at 0.9730%. The CFR may seem high, but the number of asymptotic infections in Shanghai is about 10 times as many as the number of confirmed cases. Therefore, the infection fatality ratio (IFR) is around 0.1%. We further predict that the number of daily infection increase in Shanghai has dropped to its lowest level in early June, 2022, but would maintain a low high climb towards at least the end of June, 2022, up to around 120 new daily infections.

We predict that, on July 1, 2022, in Hong Kong, there would be 1,220,352 cumulative confirmed cases and 9,282 cumulative deceased cases, with a CFR at 0.7759%. We further predict that the number of daily infection increase in Hong Kong would be flat and at a very low level till at least the end of June, 2022, between 350 and 1,100. Effectively, the Omicron epidemic phase in Hong Kong is over.

On the other hand, we predict that on July 1, 2022, in Taiwan, there would be 3,842,576 cumulative confirmed cases and 4,482 cumulative deceased cases, with a CFR at 0.1174%. While the Omicron epidemic in Taiwan is far from being over, the number of daily infection increase has peaked by the end of May at about 83,000, and would drop to roughly 41,750 by July 1, 2022. It appears that the CFR would eventually stabilize at about 0.12%, and the total number of deaths would be between 4,800 and 8,400 (based on the assumption that the total number of Omicron infections in Taiwan would be around 4,000,000 and 7,000,000).

Since dataset 2 only provides detailed time series case data for the entire country of China (including Hong Kong and Taiwan), we are only able to construct the SIRD times series for the whole country, not for the subregions. We predict that the numbers of cumulative confirmed/deceased cases for the whole country (again with Hong Kong and Taiwan included) would be 5,587,799 and 15,380, respectively, with a CFR stabilizing at 0.2765% on July 1, 2022. We then forecast the transmission parameters, and afterwards, simulate the dynamics and development of the Omicron epidemic with the SIRD time series construct (Figs. 6 and 7). For the 28-day time-period ending on July 1, 2022, we find that the daily infection increase has already peaked at the end of May, and would drop steadily to near zero around June 14, 2022, but then rise at a low level between 150 and 4,000 afterwards. Effectively, by mid-June, 2022, the epidemic phase of Omicron in China, including Hong Kong and Taiwan, would end.

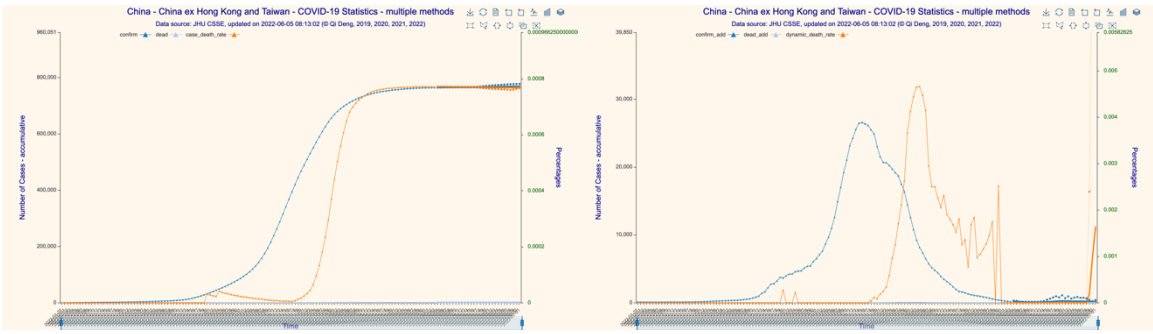

Fig. 2. 28-day forecast for confirmed/deceased cases in China (excluding Hong Kong and Taiwan)

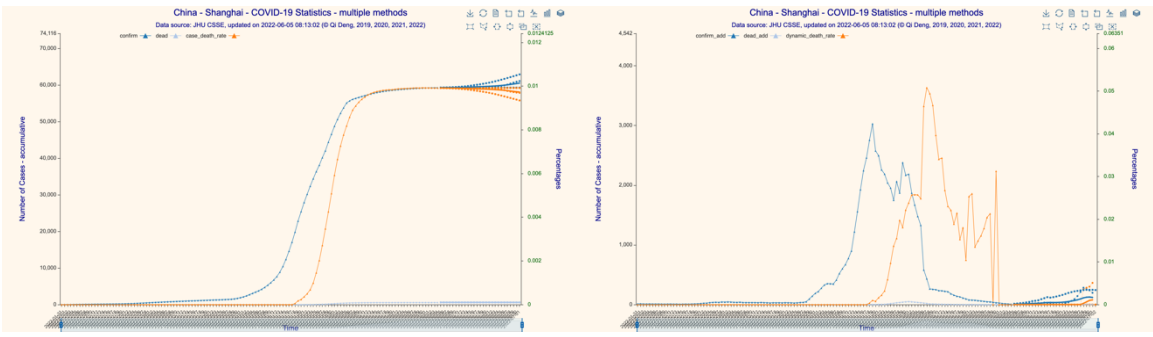

Fig. 3. 28-day forecast for confirmed/deceased cases in Shanghai

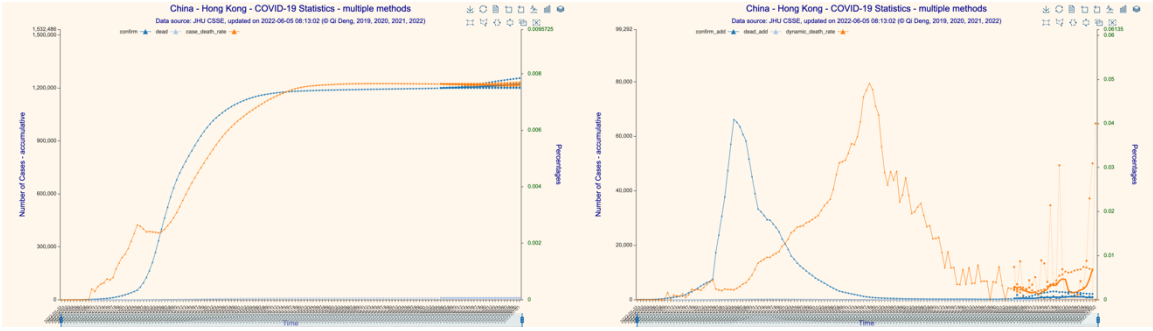

Fig. 4. 28-day forecast for confirmed/deceased cases in Hong Kong

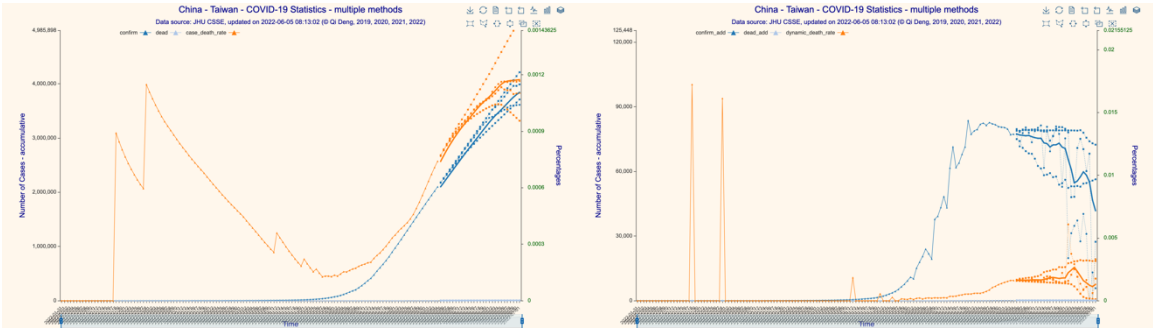

Fig. 5. 28-day forecast for confirmed/deceased cases in Taiwan

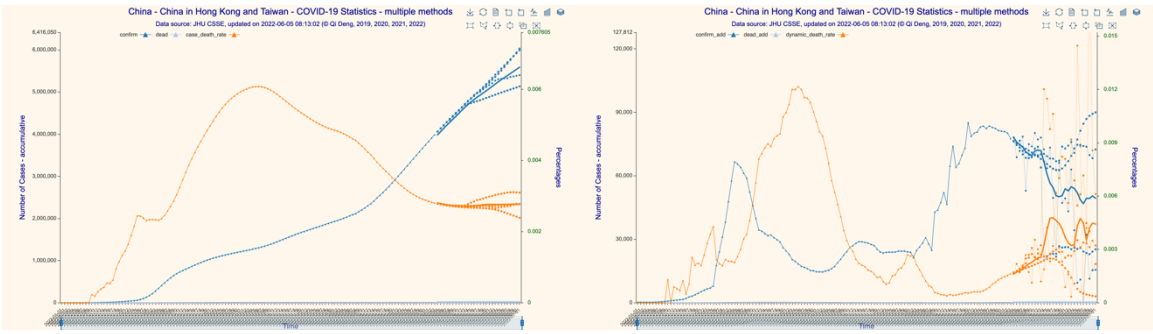

Fig. 6. 28-day forecast for confirmed/deceased cases in China (including Hong Kong and Taiwan)

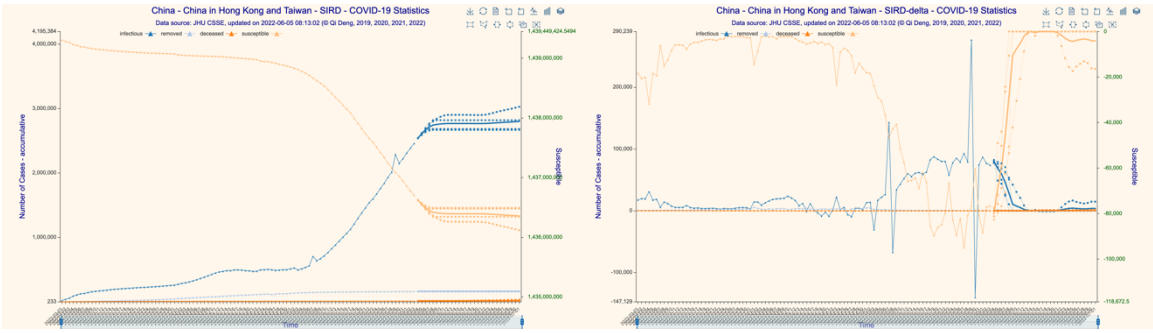

Fig. 7. 28-day SIRD simulation for cases in China (including Hong Kong and Taiwan)

Discussion

We apply DNN and RNN-LSTM techniques to estimate the stochastic transmission parameters for a SIRD model with a discrete time series construct. We then use the model to

forecast the further development of Omicron in China (including and excluding Hong Kong and Taiwan), Shanghai, Hong Kong and Taiwan.

We make use of two datasets. The first dataset, from JHU CSSE, includes time series tracking confirmed/deceased Omicron cases, which we use to construct training data for deep learning. The second dataset, from Tencent, includes detailed daily Omicron records, from which we construct the SIRD model. We run a 7-day moving average on both datasets to smooth out extreme outliers due to reporting irregularities.

We then apply DNN and RNN-LSTM deep learning techniques to fit the confirmed/deceased time series to predict further development of confirmed/deceased cases, as well as to predict the transmission parameters ($\beta, \gamma^R, \gamma^D$) for 28 days. Finally, we use the predicted transmission parameters to simulate the Omicron dynamics for 28 days.

Our results show that, in China (excluding Hong Kong and Taiwan), the daily Omicron infection increase is between 60 and 260 in the 28-day forecast period. On July 1, 2022, there would be 768,622 cumulative confirmed cases and 591 cumulative deceased cases. The CFR would be stable over our 28-day forecast time-period (0.077%±0.00025%), and the total deaths with Omicron would be up to 280,000. Furthermore, the numbers of cumulative confirmed/deceased cases for the whole country (with Hong Kong and Taiwan included) would be 5,587,799 and 15,380, respectively, with a CFR stabilizing at 0.2765% on July 1, 2022. The daily infection increase has already peaked at the end of May, and would drop steadily to near zero around June 14, 2022, but then rise at a low level between 150 and 4,000 afterwards. Effectively, by mid-June, 2022, the epidemic phase of Omicron in China, including Hong Kong and Taiwan, would end. It is time for the country to scrap its stringent zero-COVID policy and let people get back to their normal personal and economic life.

With the introduction of the deep learning-enhanced compartmental model, we provide an effective and easy-to-implement alternative to prevailing stochastic parameterization, which estimates transmission parameters through either probability likelihood maximization, or Marcov Chain Monte Carlo simulation, or simply assumptions. The effectiveness of the prevalent approach depends upon detailed statistics on transmission characteristics among heterogeneous subpopulations, and such statistics are economically and resource-wide expensive. On the other hand, deep learning techniques uncover hidden interconnections among observed data, reducing prediction's dependency on data particularity. Future research on deep learning's utilities in epidemic modeling can further enhance its forecasting power.

References and Notes

[1] Bartsch SM, O'Shea KJ, Chin KL, et al. Maintaining face mask use before and after achieving different COVID-19 vaccination coverage levels: a modelling study. Lancet Public Health 2022; 7:e356–65; Published Online, March 8, 2022; https://doi.org/10.1016/ S2468-2667(22)00040-8.

[2] COVID-19 Cumulative Infection Collaborators. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. The Lancet 2022; Published Online April 8, 2022; https://doi.org/10.1016/S0140-6736(22)00484-6.

[3] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 2020; https://doi.org/10.1016/S0140-6736(20)30260-9.

[4] Gilbert M, Pullano G, Pinotti F, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. The Lancet 2020; https://doi.org/10.1016/S0140-6736(20)30411-6.

[5] Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The Lancet 2020; https://doi.org/10.1016/S1473-3099(20)30144-4.

[6] Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020; https://doi.org/10.1126/science.aba9757.

[7] Kraemer MUG, Yang CH, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020; 368:493–497.

[8] Tian T, LiuY, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 2020; 368:638–642.

[9] Zhang J, Lou J, Ma Z, Wu J. A compartmental model for the analysis of SARS transmission patterns and outbreak control measures in China. Applied Mathematics and Computation 2005; 162(2):909–924.

[10] Liu Y, Rocklöv J. The effective reproduction number for the omicron SARS-CoV-2 variant of concern is several times higher than Delta. J Travel Med 2022 Mar 9; taac037; https://doi.org/10.1093/jtm/taac037.

[11] Ito K, Piantham C, Nishiura H. Relative instantaneous reproduction number of Omicron SARS-CoV-2 variant with respect to the Delta variant in Denmark. J Med Virol 2022 May; 94(5):2265-2268. https://doi.org/10.1002/jmv.27560.

[12] Kim D, Ali ST , SunKim S, et al. Estimation of Serial Interval and Reproduction Number to Quantify the Transmissibility of SARS-CoV-2 Omicron Variant in South Korea. Viruses 2022 Mar 4;14(3):533. https://doi.org/10.3390/v14030533.

[13] Cai J, et al. Modeling transmission of SARS-CoV-2 Omicron in China. Nature Medicine. https://doi.org/10.1038/s41591-022-01855-7 (2021).

[14] Gillespie DT. Exact Stochastic Simulation of Coupled Chemical Reactions. J Physical Chemistry 1977; 81(25):2340–2361.

[15] Chen TM, Rui J, Wang QP, et al. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. Infectious Diseases of Poverty 2020; 9:24.

[16] Naheed A, Singh M, Lucy D. Numerical study of SARS epidemic model with the inclusion of diffusion in the system. Applied Mathematics and Computation 2014; 229:480–498.

[17] Deng Q. Dynamics and Development of the COVID-19 Epidemics in the US: a Compartmental Model with Deep Learning Enhancement. Journal of Medical Internet Research, 2020; 22(8), e21173, https://doi.org/10.2196/21173, PMID: 32763892.

Supplementary Materials

[1] Dataset_1 time series (constructed)

[2] Dataset_2 time series (constructed)