Multi-variant COVID-19 model with heterogeneous transmission rates using deep neural networks

K.D. Olumoyin[a,*], A.Q.M. Khaliq[a], K.M. Furati[b]

[a]*Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN 37132, USA*
[b]*Department of Mathematics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*

Abstract

Mutating variants of COVID-19 have been reported across many US states since 2021. In the fight against COVID-19, it has become imperative to study the heterogeneity in the time-varying transmission rates for each variant in the presence of pharmaceutical and non-pharmaceutical mitigation measures. We develop a Susceptible-Exposed-Infected-Recovered mathematical model to highlight the differences in the transmission of the B.1.617.2 delta variant and the original SARS-CoV-2. Theoretical results for the well-posedness of the model are discussed. A Deep neural network is utilized and a deep learning algorithm is developed to learn the time-varying heterogeneous transmission rates for each variant. The accuracy of the algorithm for the model is shown using error metrics in the data-driven simulation for COVID-19 variants in the US states of Florida, Alabama, Tennessee, and Missouri. Short-term forecasting of daily cases is demonstrated using long short term memory neural network and an adaptive neuro-fuzzy inference system.

Keywords: deep neural network, data-driven simulation, heterogeneous transmission rates, COVID-19, multi-variants

1. Introduction

COVID-19 was first reported in China in 2019 [1], it has since become a global pandemic. In recent months, there have been reports of mutating variants of the virus [2]. In 2021, the dominant mutant variant of COVID-19 was the B.1.617.2 delta variant [3]. Effort to combat the spread of COVID-19 have included combinations of pharmaceutical (vaccination and hospitalization) and non-pharmaceutical (social distancing, contact tracing, and facial mask) measures.

Prior to the onset of COVID-19 mutating variants in the US, the progress seen in the data from several states prompted the ease of the various non-pharmaceutical measures. Amid the news that several states had vaccinated over 70% of its population and a few states had vaccinated between $60\% - 70\%$ of its population, vaccination effort began to slow down in many US states. As a result, the existence of mutating variants resulted in a resurgence in cases of infections. The Center for Disease Control and Prevention (CDC) reported that the dominant variant in the

*Corresponding author
Email address: kayode.olumoyin@mtsu.edu (K.D. Olumoyin)

US in 2021 was the B.1.617.2 delta variant. According to the World Health Organization (WHO), many variants were first reported in the United Kingdom and South Africa and in recent months, the USA, Europe, China, Brazil, and Japan have all reported mutating variant infected cases.

We present a data-driven deep learning algorithm for a model consisting of time-varying transmission rates for each active variant. Using infected daily cases data, we learn the form of the time-varying transmission rates, to reveal a timeline of the impact of mitigation measures on the transmission of COVID-19 [4, 5]. It can also be demonstrated that this algorithm shows improvement on short-term forecasting when combined with a recurrent neural network and an adaptive neuro-fuzzy inference system.

Neural networks are universal approximators of continuous functions [6, 7]. Feedforward neural networks (FNN) have been used to learn approximate solutions of differential equations. In [8], FNN was used to develop differential equation solvers and parameter estimators by constraining the residual. This FNN is called the Physics Informed Neural Network (PINN). PINN has been used to simulate pandemic spread, see [9], where the model parameters were taken to be constants. In [10], an algorithm that combines PINN with Long Short-term Memory (LSTM) is presented to solve an epidemiological model and identify weekly and daily time-varying parameters.

The paper is organized as follows. In Section 2, we introduce and discuss the multi-variant SEIR model and the time-varying transmission rate of each variant. The well posedness of the model is discussed in Section 2.2. The neural network structure of the Epidemiology neural network EINN is presented in 3. Data-driven simulation of COVID-19 data is shown in Sections 4. A comparison of a recurrent neural network based forecast and an adaptive neuro-fuzzy inference system based forecast is presented in 6. The performance error metrics of EINN is discussed in Section 7. The paper is summarized in Section 8.

2. Multi-variant SEIR model

We assume that the total population $N(t) = N$ at any given time is distributed among the following compartments: susceptible (S), exposed (E), Infectious (I_i), $i = 1, \ldots, M$, and recovered (R), where M is the number of different variants. The interaction between the compartments is shown in Figure 1.

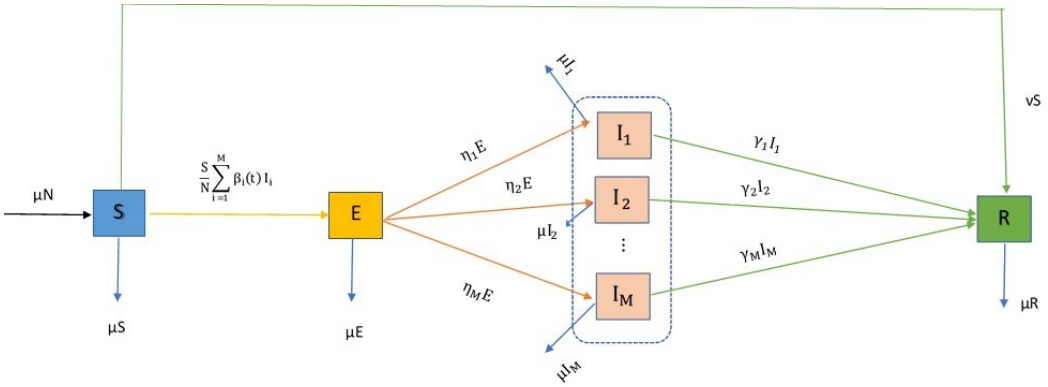


Figure 1: Transfer diagram between the compartments

As shown in Figure (1), the susceptible individuals enter the exposed compartment at the rate $\frac{1}{N}\sum_{i=1}^{M}\beta_i(t)I_i$, where $\beta_i(t)$ is the transmission rate of variant i. The exposed individuals progress to the ith infected compartment at the rate η_i. The ith infected compartment recover at the rate of $\gamma_i I_i$.

We assume a natural death rate μ, given by $\mu = \frac{1}{LFE\times 365}$, where LFE denotes the life expectancy. For simplicity, it is assumed that the birth rate of the population is equal to the death rate. The parameter η is the transmission rate from the exposed to the various infectious sub-compartments, γ_i^{-1} is the mean symptomatic infectious period for the ith variant. The parameter $\nu(t)$ represent the time-dependent removal rate of the vaccinated individuals from the susceptible compartment. We assume that any variant does not super-infect another variant, so there are no interactions between the infectious sub-compartments.

Based on the transfer diagram depicted in Figure 1, the mathematical model for a multi-variant COVID-19 pandemic with heterogeneous transmission rates is given by:

$$\begin{aligned}
\frac{dS}{dt} &= \mu N - \frac{S}{N}\sum_{i=1}^{M}\beta_i(t)I_i - \Big(\nu(t)+\mu\Big)S \\
\frac{dE}{dt} &= \frac{S}{N}\sum_{i=1}^{M}\beta_i(t)I_i - (\eta+\mu)E \\
\frac{dI_i}{dt} &= \eta_i E - (\gamma_i+\mu)I_i, \qquad i = 1,\ldots,M \\
\frac{dR}{dt} &= \sum_{i=1}^{M}\gamma_i I_i + \nu(t)S - \mu R
\end{aligned} \tag{1}$$

subject to non-negative initial conditions

$$S(0) = S_0, \quad E(0) = E_0, \quad I_i(0) = I_{i,0}, i = i,\ldots,M, \quad R(0) = R_0.$$

The parameter η is defined as: $\eta = \sum_{i=1}^{M}\eta_i$, and the total population is

$$N(t) = S(t) + E(t) + \sum_{i=1}^{M} I_i(t) + R(t).$$

The differential equation satisfied by the total population size is obtained by adding all the equations in (1), that is, $\frac{dN}{dt} = 0$ and thus N is constant. The model parameters are summarized in Table 1.

Time-varying transmission rates have been shown to efficiently model the spread of COVID-19 [4, 11]. Next, will discuss the form of the time-varying transmission rates for each variant.

2.1. Variant-based time-varying transmission rates

Time-varying transmission rates in (1) incorporates the impact of governmental actions, and the public response [12]. We consider the transmission rates of the form

$$\beta_i(t) = \beta_i^0 \exp(-\kappa_i t), \qquad 1 \le i \le M. \tag{2}$$

where κ_i in (2) is the infectiousness factor for each i th variant. However, we define $(1+\tau_i)$ to be the factor by which a particular variant is more infectious than the original variant SARS-CoV-2. And so, the following relationship exist between each mutating variant and the SARS-CoV-2 variant (3).

$$\beta_i(t) = \beta_1(t)(1+\tau_i), \qquad 2 \leq i \leq M. \tag{3}$$

In (3), $\beta_1(t)$ is the transmission rate for the original variant SARS-CoV-2. The transmission rates of the subsequent mutating variants are given by $\beta_i(t)$, $i = 2,\ldots,M$, where M represents the number of mutating COVID-19 variants. Although the publicly available data reports the daily infected cases, there were reports that suggest hat the dominant variant, B.1.617.2 delta variant, to be twice as infectious as the original variant SARS-CoV-2. According to CDC reports [3, 13], the delta variant accounted for 1.3% of total infected cases in May, 2021, 9.5% in June, and in August it accounted for 93% of the total infected cases.

Parameter	**Notation**	**Range**	**Remark**	**Reference**
Baseline transmission rate for each ith variant	β_i^0	[0,1)	fitted using daily cases data	[4]
Emigration rate	μ	$\frac{1}{LFE\times 365}$	constant	[14]
Mean latent period	η^{-1}	$2-14(days)$	constant	[14]
recovery rate for each ith variant	γ_i	$[0,1)$	constant	
infectiousness factor for each ith variant	κ_i	$[0,1)$	constant	

Table 1: Summary table of parameters in model (1)

2.2. Well-posedness of the model

Definition 1 ([15], Locally Lipschitz continuity)**.** *Let $d_1, d_2 \in \mathbf{N}$ and $\mathbf{S}$ be a subset of $\mathbf{R}^{d_1}$. A function $\mathbf{F} : \mathbf{S} \to \mathbf{R}^{d_2}$ is Lipschitz continuous on $\mathbf{S}$ if there exists a nonnegative constant $L \geq 0$ such that*

$$|\mathbf{F}(x) - \mathbf{F}(y)| \leq L|x-y|, \quad x, y \in \mathbf{S}. \tag{4}$$

Let $\mathscr{U}$ be an open subset of $\mathbf{R}^{d_1}$, and let $\mathbf{F} : \mathscr{U} \to \mathbf{R}^{d_2}$. We shall call $\mathbf{F}$ locally Lipschitz continuous if for every point $x_0 \in \mathscr{U}$ there exists a neighborhood V of x_0 such that the restriction of $\mathbf{F}$ to V is Lipschitz continuous on V.

We consider a more general framework of model (1)

$$z'(t) = G(z(t)), \qquad z(0) = z_0, \tag{5}$$

where $z(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T$ and $G(z(t)) = (g_1(z(t)), g_2(z(t)), \ldots, g_n(z(t)))^T$, the initial condition $z_0 \in \mathbf{R}^n$. We state the following theorem.

Theorem 2.1 ([15]). *If $\mathbf{G}:\mathbf{R}^n \to \mathbf{R}^n$ is locally Lipschitz continuous and if there exist nonnegative constants B, K such that*

$$|\mathbf{G}(z(t))| \leq K\ |z(t)| + B, \quad z(t) \in \mathbf{R}^n, \tag{6}$$

then the solution of the initial value problem (5) *exists for $t > 0$, and*

$$|(z(t))| \leq |z_0| \cdot \exp(K\ |t|) + \frac{B}{K} \cdot (\exp(K \cdot |t|) - 1), \quad t > 0, \tag{7}$$

Lemma 2.2. *For each ith variant, $i \in \{1,\ldots,M\}$, the time varying transmission rates $\nu, \beta_i : [0,\infty) \to [0,\infty)$ are Lipschitz continuous and contnuously differentiable. There exists $\beta_{min}, \beta_{max} > 0$ and $\nu_{min}\nu_{max} > 0$ such that $\beta_{min} \leq \beta_i(t) \leq \beta_{max}$, $\nu_{min} \leq \nu(t) \leq \nu_{max}$ for all t.*

Theorem 2.3. *The nonlinear first order system of differential equations* (1) *has at least one solution which exists for $t \in [0,\infty)$.*

Proof. Let $z(t) = (S(t), E(t), I_1(t), \ldots, I_M(t), R(t))^T$, we can set

$$G:\mathbf{R}^{M+3} \to \mathbf{R}^{M+3}, \qquad z(t) \to \begin{pmatrix} \mu N - \frac{S}{N}\sum_{i=1}^{M} \beta_i(t) I_i - \Big(\nu(t)+\mu\Big)S \\ \frac{S}{N}\sum_{i=1}^{M} \beta_i(t) I_i - (\eta+\mu)E \\ \eta_i E - (\gamma_i+\mu)I_i, \quad i = 1,\ldots,M \\ \sum_{i=1}^{M} \gamma_i I_i + \nu(t)S - \mu R \end{pmatrix} \tag{8}$$

G is locally Lipschitz continuous, using supremum norm $||f(t)|| := \sup_{t\in[a,b]} |f(t)|$, we have

$$\begin{aligned}
||G(z(t))|| &= \sup_{t\in[0,\infty)} \left\{ \left| \mu N - \frac{S(t)}{N}\sum_{i=1}^{M} \beta_i(t) I_i(t) - \Big(\nu(t)+\mu\Big)S(t) \right|, \left| \frac{S(t)}{N}\sum_{i=1}^{M} \beta_i(t) I_i - (\eta+\mu)E(t) \right|, \right. \\
&\qquad \left. \Big|\eta_i E(t) - (\gamma_i+\mu)I_i(t), \quad i = 1,\ldots,M\Big|, \left| \sum_{i=1}^{M} \gamma_i I_i(t) + \nu(t)S(t) - \mu R(t) \right| \right\} \\
&\leq \sup_{t\in[0,\infty)} \left\{ \mu N + \beta_{max} \left| \frac{S(t)}{N}\sum_{i=1}^{M} I_i(t) \right| + (\nu_{max}+\mu)|S(t)|, \beta_{max} \left| \frac{S(t)}{N}\sum_{i=1}^{M} I_i(t) \right| + (\eta+\mu)|E(t)|, \right. \\
&\qquad \left. \eta_i|E(t)| + (\gamma_i+\mu)|I_i(t)|, \gamma_i \sum_{i=1}^{M} |I_i(t)| + \nu_{max}|S(t)| + \mu|R(t)| \right\} \\
&\leq \sup_{t\in[0,\infty)} \Big\{ \mu N + \beta_{max}|S(t)| + (\nu_{max}+\mu)|S(t)|, \beta_{max}|S(t)| + (\eta+\mu)|E(t)|, \\
&\qquad \eta_i|E(t)| + (\gamma_i+\mu)|I_i(t)|, \gamma_i \sum_{i=1}^{M} |I_i(t)| + \nu_{max}|S(t)| + \mu|R(t)| \Big\} \\
&\leq K||z(t)||.
\end{aligned}$$

So by Theorem 2.1, and the boundedness of the time-varying nonlinear functions from Lemma 2.2, the nonlinear initial value problem 1 has a solution for all time. □

2.3. Basic reproduction number and equilibria stability

The basic reproduction number $\mathscr{R}_0$ is the expected number of secondary infections that a single infectious individual will generate on average within a susceptible population.

Definition 2. *The disease-free equilibrium of* (1) *is given by*

$$(S^*, E^*, I_1^*, \ldots, I_M^*, R^*) = (S_0, 0, 0, \ldots, 0, 0).$$

The basic reproduction number $\mathscr{R}_0$ is calculated for the case when $\beta_i(t) = \beta_i^0$, $i \in \{1, \ldots, M\}$. Applying the next-generation operator approach [16], the reproduction number $\mathscr{R}_0$ is obtained as the spectral radius of the next generation matrix FV^{-1}, where

$$F = \begin{pmatrix} 0 & \beta_1^0 & \ldots & \beta_M^0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}, \quad V = \begin{pmatrix} \eta+\mu & 0 & 0 & \ldots & 0 \\ -\eta_1 & \gamma_1+\mu & 0 & \ldots & 0 \\ -\eta_2 & 0 & \gamma_2+\mu & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\eta_M & 0 & 0 & \ldots & \gamma_M+\mu \end{pmatrix}.$$

The basic reproduction number $\mathscr{R}_0$ is computed as follows in (9)

$$\mathscr{R}_0 = \sum_{i=1}^{M} \frac{\beta_i^0 \eta_i}{(\gamma_i+\mu)(\eta+\mu)}. \tag{9}$$

Next, we analyze the local asymptotic stability of the disease-free equilibrium in Definition 2.

Theorem 2.4. *The disease-free equilibrium* $(S^*, E^*, I_1^*, \ldots, I_M^*, R^*)$ *of* (1) *is locally asymptotically stable if* $\mathscr{R}_0 < 1$.

Proof. The Jacobian of the right hand side of (1) at the equilibrium point is given by

$$J = \begin{pmatrix} -(\nu+\mu) & 0 & -\beta_1^0 & -\beta_2^0 & \ldots & -\beta_M^0 & 0 \\ 0 & -(\eta+\mu) & \beta_1^0 & \beta_2^0 & \ldots & \beta_M^0 & 0 \\ 0 & \eta_1 & -(\gamma_1+\mu) & 0 & \ldots & 0 & 0 \\ 0 & \eta_2 & 0 & -(\gamma_2+\mu) & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \eta_M & 0 & 0 & \ldots & -(\gamma_M+\mu) & 0 \\ \nu & 0 & \gamma_1 & \gamma_2 & \ldots & \gamma_M & -\mu \end{pmatrix}$$

If $M = 1$, the eigenvalues of the Jacobian matrix are given as follows:

$$\lambda_1 = -\mu, \lambda_2 = -(\nu+\mu), \lambda_3 = \frac{1}{2}\Big(-A-B\Big), \lambda_4 = \frac{1}{2}\Big(A-B\Big).$$

where

$$A = \sqrt{4\beta_1^0\eta_1 + (\eta_1 - \gamma_1)^2},$$
$$B = \eta_1 + 2\mu + \gamma_1.$$

Clearly, $A, B > 0$ and $A < B$, so that $\lambda_1, \lambda_2, \lambda_3, \lambda_4 < 0$. Similarly, we can show negative eigenvalues for $M \geq 2$. So the disease-free equilibrium is locally asymptotically stable. □

3. Epidemiology Informed Neural Network (EINN)

A Feedforward Neural Network (FNN) composed of L layers, t inputs and an output $\mathscr{N}$ can be represented as the following function

$$\mathscr{N}(t;\Sigma) = \sigma(W_L\sigma(\ldots\sigma(W_2\sigma(W_1t + b_1) + b_2)\ldots) + b_L), \tag{10}$$

where $\Sigma := (W_1, \ldots, W_L, b_1, \ldots, b_L)$. The neural network weight matrices are W_l, $l = 1, \ldots, L$, while the bias vectors are b_l, $l = 1, \ldots, L$. Here, σ is the activation function. Given a collection of sample pairs (t_k, u_k), $k = 1, \ldots K$, where u is some target function, the goal is to find Σ^* by solving the following optimization problem

$$\Sigma^* = \arg\min \frac{1}{K}\sum_{k=1}^{K} ||\mathscr{N}(t_k;\Sigma) - u_k||_2^2. \tag{11}$$

The function $\frac{1}{K}\sum_{k=1}^{K} ||\mathscr{N}(t_k;\Sigma) - u_k||_2^2$ on the right hand side of (11) is called the mean squared error (MSE) loss function. A major task in training a network is to determine the suitable number of layers and the number of neurons per layer needed, the choice of activation function, and an appropriate optimizer for the loss function [17].

EINN is a form of Feedforward Neural Network that includes the known epidemiology dynamics in its loss function. EINN is adapted for the SEIR model (1), where the Mean Square Error (MSE) of this neural network's loss function includes the known epidemiology dynamics such as a lockdown, while other mitigation measures such as social distancing, and contact tracing are detected by the time-varying transmission rate. The output of EINN are the learned solutions to the SEIR model (1) denoted by $S(t_j;\psi;\rho)$, $E(t_j;\psi;\rho)$, $I(t_j;\psi;\rho)$, $R(t_j;\psi;\rho)$, $j = 1, \ldots, T$. Here, ψ represent the neural network weights and biases while ρ represent the epidemiology parameters and T is the number of days in our dataset. Next, we set-up time-varying transmission rate networks whose outputs are $\beta_i(t_j;\pi_i;\kappa_i)$, $j = 1, \ldots, T$, for $i = 1, \ldots, M$. Each π_i represent the weights and biases of each ith network and κ_i is the infectiousness factor for each ith variant. The training data is generated using cubicspline and denoted by $\tilde{I}(t_j)$ and $\tilde{V}(t_j)$, $j = T_v, \ldots, T$. Here T_v is an integer that correspond to the vaccination start date in the dataset. The B.1.617.2 delta variant was first reported in the USA in May, T_δ is an integer that correspond to May $4th$, 2021. We observe that training data is not available for all the compartments in the SEIR model, however, EINN is able to capture the epidemiology interactions between the compartments because the residual of equation (1) is included in the MSE loss function.

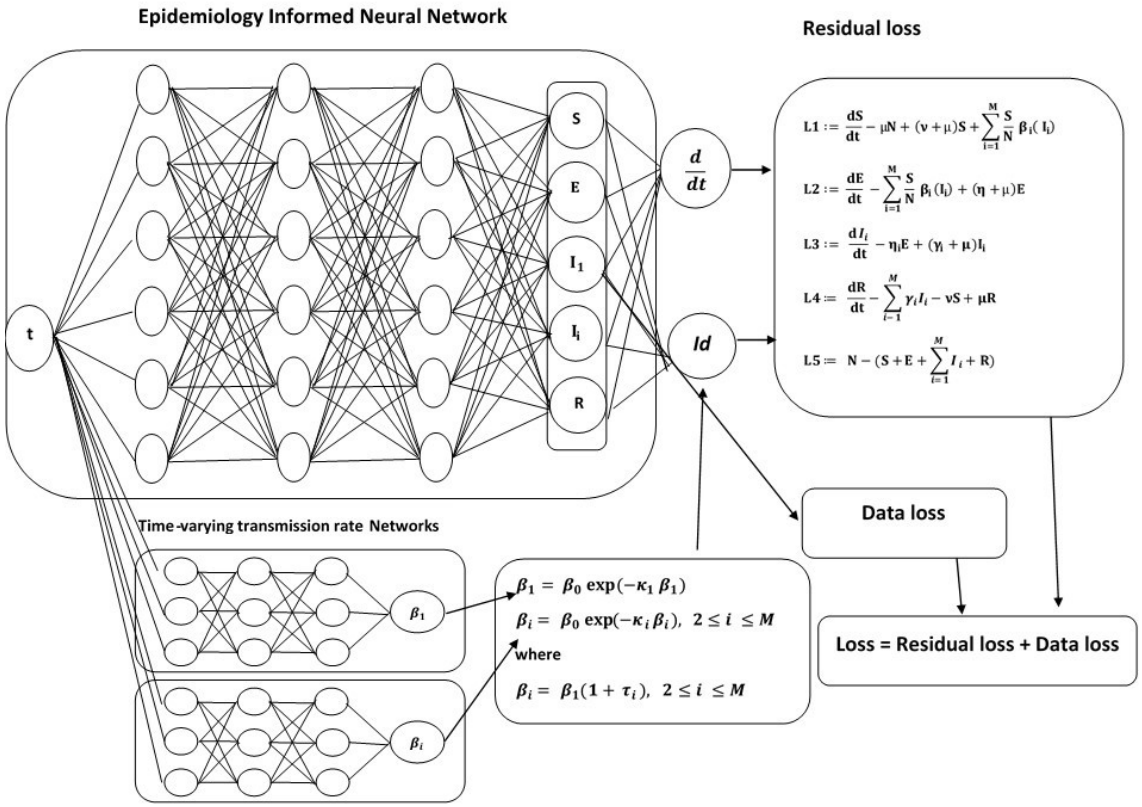


Figure 2: Schematic diagram of the Epidemiology Informed Neural Network with nonlinear time-varying transmission rate.

The MSE loss function for EINN is given by,

$$
\begin{aligned}
MSE = & \frac{1}{T_\delta}\sum_{j=1}^{T_\delta}\Big|I_1(t_j;\psi;\rho)-\tilde{I}(t_j)\Big|^2+\frac{1}{|T-T_\delta|}\sum_{j=T_\delta}^{T}\Big|\sum_{i=1}^{M}I_i(t_j;\psi;\rho)-\tilde{I}(t_j)\Big|^2\\
&+\frac{1}{T_\nu}\sum_{j=1}^{T_\nu}\Big|\nu(t_j;\psi;\rho)-0\Big|^2+\frac{1}{|T-T_\nu|}\sum_{j=T_\nu}^{T}\Big|\nu(t_j;\psi;\rho)S(t_j;\psi;\rho)-\tilde{V}(t_j)\Big|^2\\
&+\sum_{i=2}^{M}\Big|I_i(t_{\delta_1};\psi;\rho)-p_{\delta_1}\tilde{I}(t_{\delta_1})\Big|^2+\sum_{i=2}^{M}\Big|I_i(t_{\delta_2};\psi;\rho)-p_{\delta_2}\tilde{I}(t_{\delta_2})\Big|^2\\
&+\frac{1}{T_\delta}\sum_{i=2}^{M}\sum_{j=1}^{T_\delta}\Big|\beta_i(t_j;\pi_i;\kappa_i)-0\Big|^2\\
&+\frac{1}{|T-T_\delta|}\sum_{i=2}^{M}\sum_{j=T_\delta}^{T}\Big|\beta_1(t_j;\pi_1;\kappa_1)(1+\tau_i)-\beta_i(t_j;\pi_i;\kappa_i)\Big|\\
&+\sum_{l=1}^{5}L_l,
\end{aligned}
\tag{12}
$$

where the residual L_l, $l=1,\dots 5$, is as follows

$$
\begin{aligned}
L_1 &= \frac{1}{T}\sum_{j=1}^{T}\Big|\frac{dS(t_j;\psi;\rho)}{dt_j}+\frac{S(t_j;\psi;\rho)}{N}\Big(\sum_{i=1}^{M}\beta_i(t_j;\pi_i;\kappa_i)I_i(t_j;\psi;\rho)\Big)+\Big(\nu(t_j;\psi;\rho)+\mu\Big)S(t_j;\psi;\rho)-\mu N\Big|^2\\
L_2 &= \frac{1}{T}\sum_{j=1}^{T}\Big|\frac{dE(t_j;\psi;\rho)}{dt_j}-\frac{S(t_j;\psi;\rho)}{N}\Big(\sum_{i=1}^{M}\beta_i(t_j;\pi_i;\kappa_i)I_i(t_j;\psi;\rho)\Big)+(\eta+\mu)E(t_j;\psi;\rho)\Big|^2\\
L_3 &= \frac{1}{T}\sum_{i=1}^{M}\sum_{j=1}^{T}\Big|\frac{dI_i(t_j;\psi;\rho)}{dt_j}-\eta_i E(t_j;\psi;\rho)+(\gamma_i+\mu)I_i(t_j;\psi;\rho)\Big|^2\\
L_4 &= \frac{1}{T}\sum_{j=1}^{T}\Big|\frac{dR(t_j;\psi;\rho)}{dt_j}-\sum_{i=1}^{M}\gamma_i I_i(t_j;\psi;\rho)-\nu(t_j;\psi;\rho)S(t_j;\psi;\rho)+\mu R(t_j;\psi;\rho)\Big|^2\\
L_5 &= \frac{1}{T}\sum_{j=1}^{T}\Big|N-(S(t_j;\psi;\rho)+E(t_j;\psi;\rho)+\sum_{i=1}^{M}I_i(t_j;\psi;\rho)+R(t_j;\psi;\rho))\Big|^2.
\end{aligned}
\tag{13}
$$

where $\eta=\sum_{i=1}^{M}\eta_i$, $i=1,\dots M$.

The daily infected cases, the vaccinated cases, the known COVID-19 variants facts and the transmission rates are enforced in the mean square error (MSE) (12), see Figure (2). For instance, p_{δ_1} and p_{δ_2} correspond to the proportion of daily cases that was due to the mutating variants as reported by the CDC [13].

Algorithm 1 EINN algorithm for SEIR model with time-varying transmission rate

1: Construct EINN

specify the input: t_j, $j = 1,\dots,T$

Initialize EINN parameter: ψ

Initialize the mathematical model parameters: $\rho = [\gamma_i]$, $i = 1,\dots,M$.

Output layer: $S(t_j;\psi;\rho)$, $E(t_j;\psi;\rho)$, $I(t_j;\psi;\rho)$, $R(t_j;\psi;\rho)$, $j = 1,\dots,T$

2: construct neural networks: β_i, $j = 1,\dots,M$

specify the input: t_j, $j = 1,\dots,T$

Initialize the neural network parameter: ϕ

Specify β_i^0 obtained by fitting daily cases

Output layers : $\beta_i(t_j;\pi_i;\kappa_i)$

$$\beta_i(t_j;\pi_i;\kappa_i) = (1+\tau_i)\beta_1(t_j;\pi_1;\kappa_1), \qquad i \geq 2 \tag{14}$$

3: Specify EINN training set

Training data: using cubicspline, generate $\tilde{I}(t_j)$ and $\tilde{R}(t_j)$, $j = 1,\dots,T$.

4: Train the neural networks

Specify an MSE loss function:

$$\begin{aligned} MSE = {} & \frac{1}{T_\delta}\sum_{j=1}^{T_\delta}\left|I_1(t_j;\psi;\rho)-\tilde{I}(t_j)\right|^2 + \frac{1}{|T-T_\delta|}\sum_{j=T_\delta}^{T}\left|\sum_{i=1}^{M}I_i(t_j;\psi;\rho)-\tilde{I}(t_j)\right|^2 \\ & + \frac{1}{T_v}\sum_{j=1}^{T_v}\left|v(t_j;\psi;\rho)-0\right|^2 + \frac{1}{|T-T_v|}\sum_{j=T_v}^{T}\left|v(t_j;\psi;\rho)S(t_j;\psi;\rho)-\tilde{V}(t_j)\right|^2 \\ & + \sum_{i=2}^{M}\left|I_i(t_{\delta_1};\psi;\rho)-p_{\delta_1}\tilde{I}(t_{\delta_1})\right|^2 + \sum_{i=2}^{M}\left|I_i(t_{\delta_2};\psi;\rho)-p_{\delta_2}\tilde{I}(t_{\delta_2})\right|^2 \\ & + \frac{1}{T_\delta}\sum_{i=2}^{M}\sum_{j=1}^{T_\delta}\left|\beta_i(t_j;\pi_i;\kappa_i)-0\right|^2 \\ & + \frac{1}{|T-T_\delta|}\sum_{i=2}^{M}\sum_{j=T_\delta}^{T}\left|\beta_1(t_j;\pi_1;\kappa_1)(1+\tau_i)-\beta_i(t_j;\pi_i;\kappa_i)\right| \\ & + \sum_{l=1}^{5}L_l, \end{aligned} \tag{15}$$

Minimize the MSE loss function: compute $\arg\min_{\{\psi;\pi_i\}}(MSE)$ using an optimizer such as the $L-BFGS-B$.

5: return EINN solution

$S(t_j;\psi;\rho)$, $E(t_j;\psi;\rho)$, $I_i(t_j;\psi;\rho)$, $R(t_j;\psi;\rho)$, $j = 1,\dots,T$, $i = 1,\dots,M$.

epidemiology parameters: γ_i, $i = 1,\dots,M$.

vaccination parameter: v

6: return time-varying epidemiology parameter:

$\beta_i(t_j;\pi_i;\kappa_i)$, $j = 1,\dots,T$, $i = 1,\dots,M$.

Infectiousness factor: κ_i, $i = 1,\dots,M$.

parameter: τ_i, $i = 2,\dots,M$.

4. Data-driven simulation of COVID-19 variants

We present results of the implementation of the EINN algorithm in Figure (2) for COVID-19 data from Alabama, Missouri, Tennessee, and Florida. We consider data from March 2020 to September 2021, during which there were two dominant variants;the original variant SARS-CoV-2 and the delta variant (B.1.617.2). CDC report indicate that 1.3% of the total infected cases were due to the delta variant in May 4th 2021 [13]. The EINN algorithm learns the infected cases, and the time-varying transmission rates due to each variant. In Table (2)–(5), pre-γ_1, post-γ_1, post-γ_2 denote the recovery rate of people infected due to the original variant SARS-CoV-2 before the onset of the delta variant, recovery rate of people infected due to the original variant SARS-CoV-2 after the onset of the delta variant, and recovery rate of people infected due to the delta variant after the onset of the delta variant respectively.

The CDC reports that by July 31st, 2021, the proportions of infected cases that are due to the B.1.617.2 delta variant in Alabama was 82.6%, Tennessee was 67.4%, Missouri 53.9%, and in Florida, it was 86.4% [3]. The CDC also reported that in the USA, the delta variant accounted for about 1.3% of the infected cases.

We seek to learn τ_i for an ith mutating variant. For the simulations in this section, we observed that the delta variant is a dominant mutating variant therefore we included only two variants, the SARS-CoV-2 and the delta variant.

Parameters	Mean	Std
pre- γ_1	0.02423	0.01266
post- γ_1	0.00395	0.00717
post- γ_2	0.00463	0.00768
η_1	0.12437	0.04933
η_2	0.20893	0.04933
κ_1	1.07385	0.06271
κ_2	1.13052	0.02809
$(1+\tau)$	1.22391	0.10176

Table 2: Using Alabama daily cases from March 2020 to September 2021, the EINN Algorithm (1) learns the model parameters

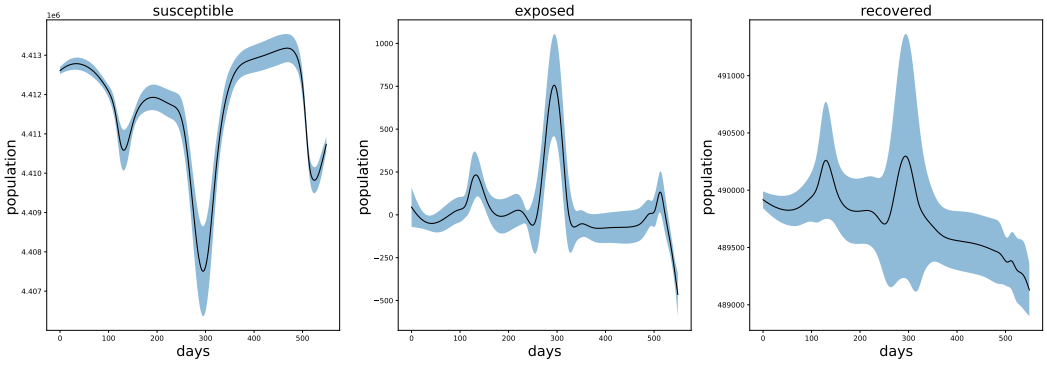


Figure 3: learned Alabama Susceptible, Exposed, and Recovered daily population

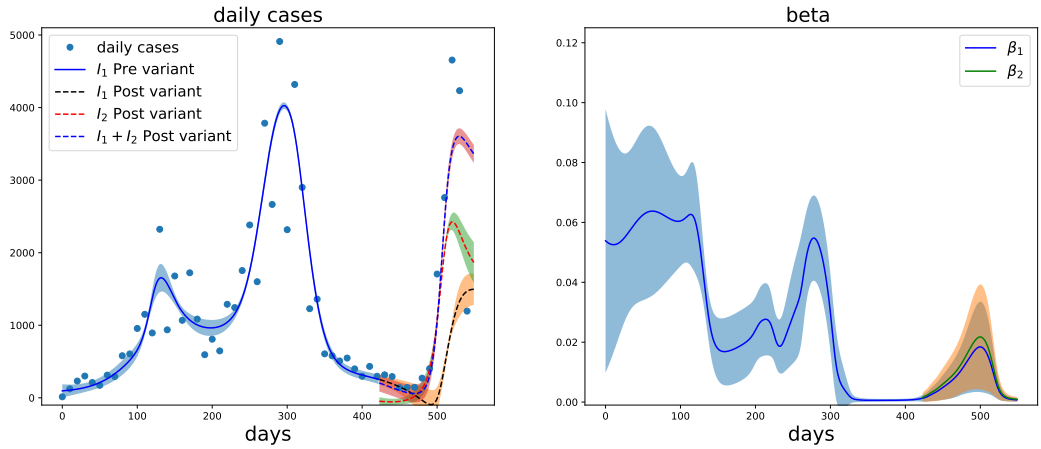


Figure 4: Alabama daily cases and time-varying transmission rates

Parameters	Mean	Std
pre- γ_1	0.02344	0.00613
post- γ_1	0.00911	0.00426
post- γ_2	0.02095	0.01794
η_1	0.15912	0.03381
η_2	0.17420	0.03383
κ_1	1.01114	0.03561
κ_2	1.10474	0.01358
$(1+\tau)$	1.15537	0.08817

Table 3: Using Missouri daily cases from March 2020 to September 2021, the EINN Algorithm (1) learns the model parameters

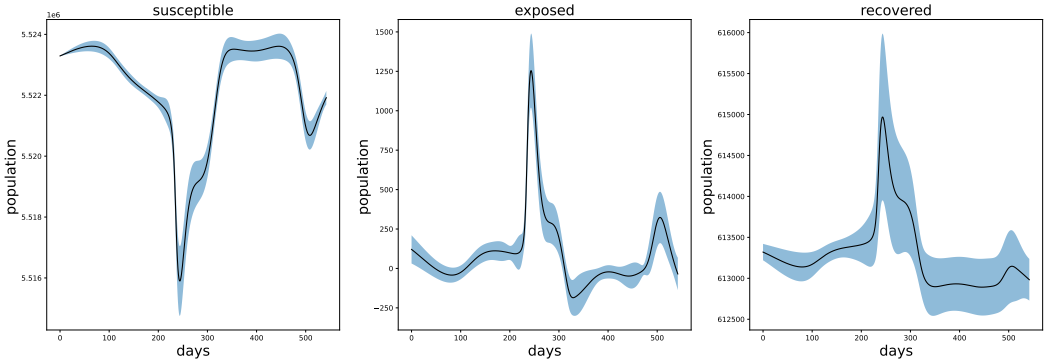


Figure 5: learned Missouri Susceptible, Exposed, and Recovered daily population

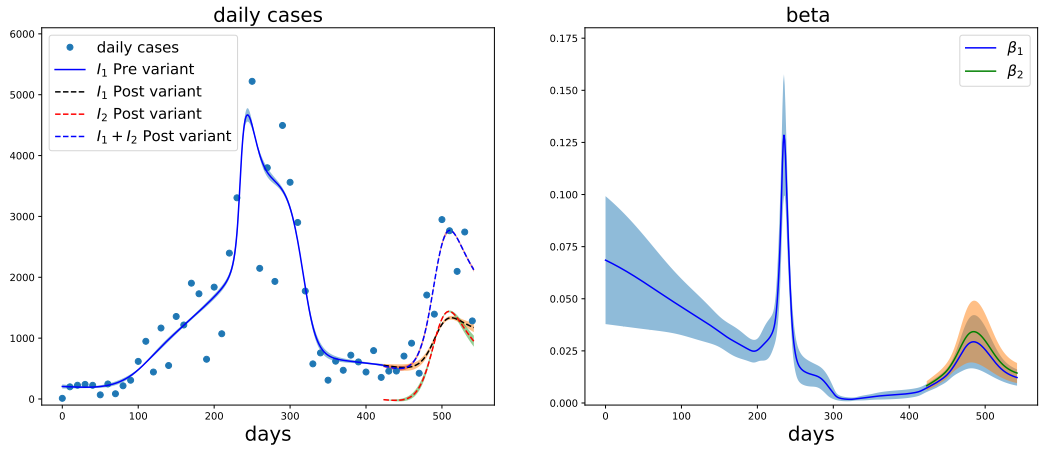


Figure 6: Missouri daily cases and time-varying transmission rates

Parameters	Mean	Std
pre- γ_1	0.01259	0.01111
post- γ_1	0.00587	0.00821
post- γ_2	0.00849	0.01081
η_1	0.13721	0.03429
η_2	0.19611	0.03427
κ_1	1.04761	0.03035
κ_2	1.13552	0.02867
$(1+\tau)$	1.09879	0.09738

Table 4: Using Tennessee daily cases from March 2020 to September 2021, EINN Algorithm (1) learns the model parameters

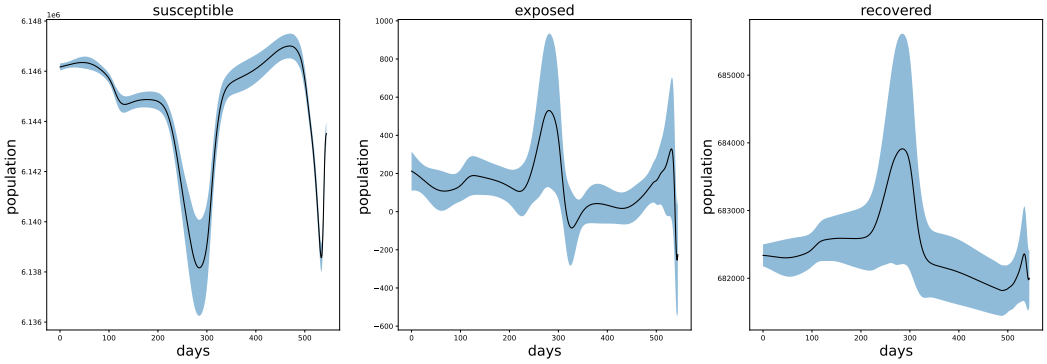


Figure 7: learned Tennessee Susceptible, Exposed, and Recovered daily population

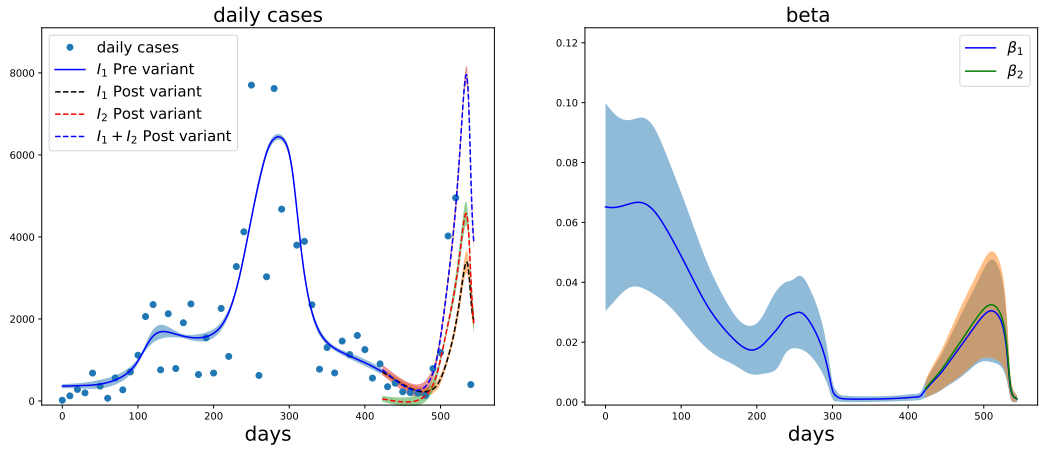


Figure 8: Tennessee daily cases and time-varying transmission rates

Parameters	Mean	Std
pre- γ_1	0.02968	0.01594
post- γ_1	0.00943	0.00985
post- γ_2	0.00576	0.00516
η_1	0.09304	0.06144
η_2	0.24027	0.06143
κ_1	1.03508	0.02477
κ_2	1.13773	0.00892
$(1+\tau)$	1.12553	0.11431

Table 5: Using Florida daily cases from March 2020 to September 2021, EINN Algorithm (1) learns the model parameters

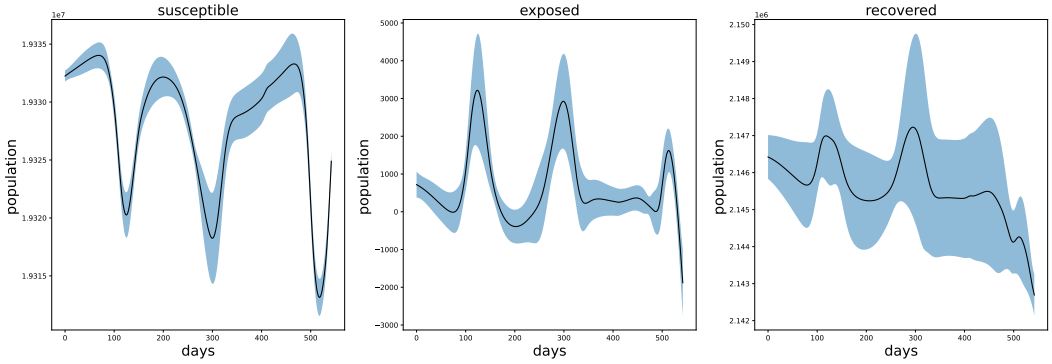


Figure 9: learned Florida Susceptible, Exposed, and Recovered daily population

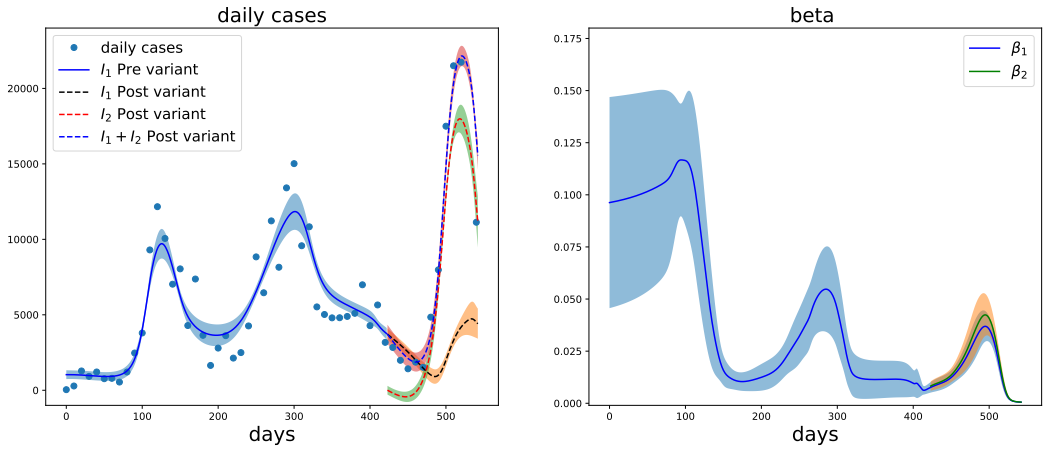


Figure 10: Florida daily cases and time-varying transmission rates

5. Forecasting daily cases

Forecasting the spread of infectious diseases in many studies are based on multiple linear regression (MLR), ordinary least squares regression (OLSR), principal component regression (PCR) and partial least squares regression (PLSR) and statistical methods such as the Auto Regressive Moving Average (ARIMA) and its many variants [18, 19, 20]. These statistical methods are not optimal for nonlinear predictive task. This has motivated a shift towards techniques that rely on neural networks and neuro-fuzzy models [21]. In this Section, we present an hybrid neural network that combines the simplicity and nonlinear learning capabilities of the Epidemiology-informed neural network (EINN) as well as the fuzzy inference system (ANFIS).

Adaptive neuro-fuzzy inference system (ANFIS), an hybrid neural network itself, is a combination of fuzzy logic and a feedforward neural network. It incorporates the advantages of both methods including learning capabilities, interpretability, quick convergence, adaptability and high accuracy. ANFIS displays excellent performance in approximation and prediction of nonlinear relationships in various fields [22].

The Adaptive Neuro-Fuzzy Inference System (ANFIS) was introduced in [23]. It combines a neural network with a fuzzy inference system (FIS) based on "IF-THEN" rules. One major advantage of FIS is that it does not require knowledge of the main physical process as a pre-condition. ANFIS combines FIS with a backpropagation algorithm. These techniques provide a method for the fuzzy modeling procedure to learn from the available dataset, in order to compute the membership function parameters that best allow the fuzzy inference system to track the given input/output data.

To forecast the transmission of a multi-variant COVID-19, we present an efficient deep learning forecast model which combines two neural networks, we solve the ODE system using an Epidemiology Informed Neural Network (EINN) and we forecast using an adaptive neuro-fuzzy system (ANFIS), which we called the EINN-ANFIS model.

6. Comparison of Forecasting techniques

We present results of the implementation of ANFIS, EINN-ANFIS, LSTM, EINN-LSTM for COVID-19 data from Alabama, Missouri, Tennessee, and Florida from March 2020 to September 2021. In the ANFIS approach, We used 4 regressors, 12 membership rules, and learning rate of 0.002. Training was done using 300 epochs, where we used the adams optimizer and for the loss function, we used the mean square error. The EINN-ANFIS is a hybrid neural network, where EINN is first used to train the daily cases dataset and a second round of training is done using ANFIS. In the LSTM approach, we used 4 input layers which corresponds to the daily cases at times t, $t+1$, $t+2$, and $t+3$. The adams optimizer is also used in training the LSTM with 20 epochs and the loss function also uses the mean square error. In the EINN-LSTM approach, a first batch of training is done using the EINN algorithm and then a second batch of training is done using LSTM. In Tables (6)–(9), we present the validation loss of each method.

Method	Mean	Std
ANFIS	0.00048	0.00098
EINN-ANFIS	0.00032	0.00050
LSTM	0.00141	0.00004
LSTM-EINN	0.00110	0.00006

Table 6: Validation loss in the ANFIS, EINN-ANFIS, LSTM, and LSTM-EINN forecasting technique for Alabama daily cases from March 2020 to September 2021.

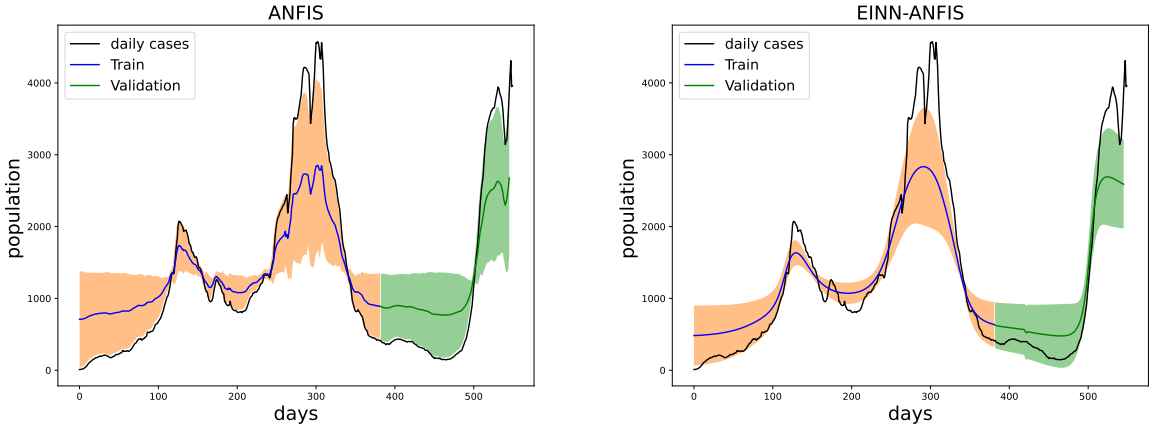
Method	Mean	Std
ANFIS	0.00011	0.00019
EINN-ANFIS	0.00004	0.00006
LSTM	0.00333	0.00003
LSTM-EINN	0.00118	0.00012

Table 7: Validation loss in the ANFIS, EINN-ANFIS, LSTM, and LSTM-EINN forecasting technique for Missouri daily cases from March 2020 to September 2021.

Method	Mean	Std
ANFIS	0.00061	0.00125
EINN-ANFIS	0.00033	0.00056
LSTM	0.00267	0.00011
LSTM-EINN	0.00183	0.00009

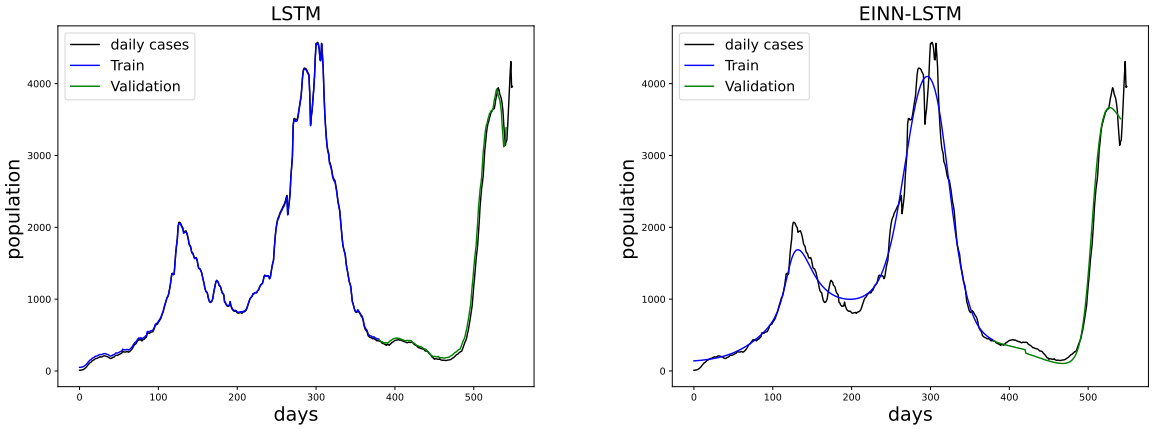
Table 8: Validation loss in the ANFIS, EINN-ANFIS, LSTM, and LSTM-EINN forecasting technique for Tennessee daily cases from March 2020 to September 2021.

As can be observed from these Tables (6)–(9) EINN-ANFIS is an imporovement over ANFIS and similarly, EINN-LSTM is an improvement over LSTM.



(a) ANFIS

(b) EINN-ANFIS



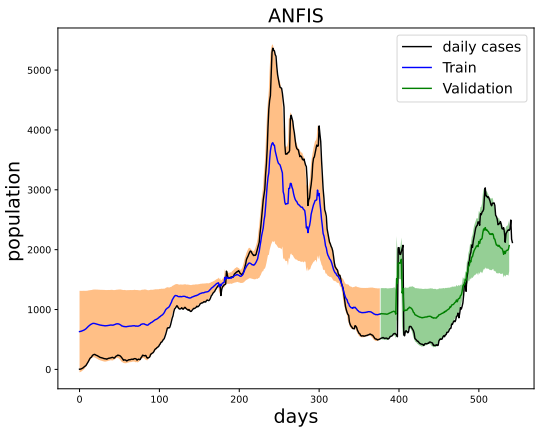
(c) LSTM

(d) EINN-LSTM

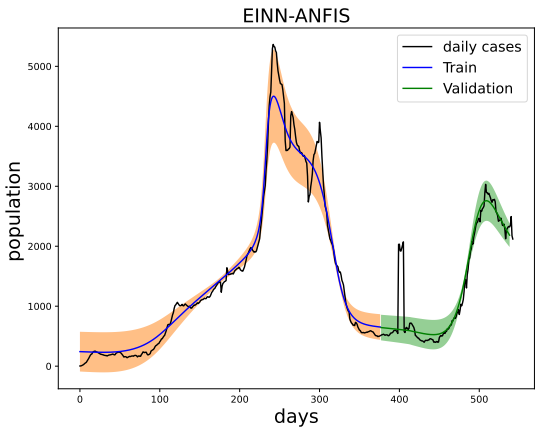
Figure 11: Alabama daily cases forecasting using ANFIS, EINN-ANFIS, LSTM, LSTM-EINN

Method	Mean	Std
ANFIS	0.00199	0.00284
EINN-ANFIS	0.00249	0.00347
LSTM	0.00169	0.00009
LSTM-EINN	0.00149	0.00014

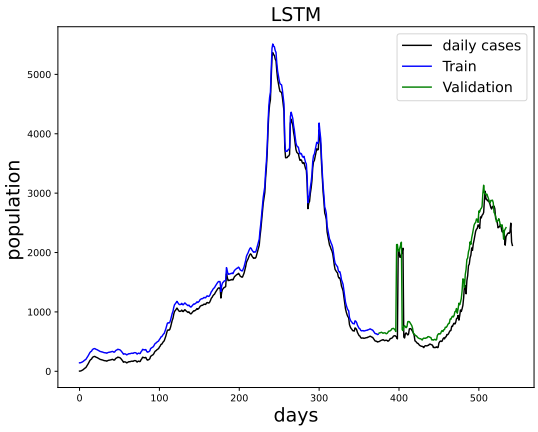
Table 9: Validation loss in the ANFIS, EINN-ANFIS, LSTM, and LSTM-EINN forecasting technique for Florida daily cases from March 2020 to September 2021.



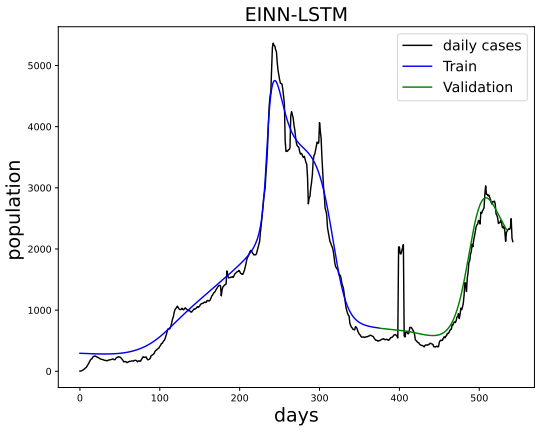
(a) ANFIS



(b) EINN-ANFIS

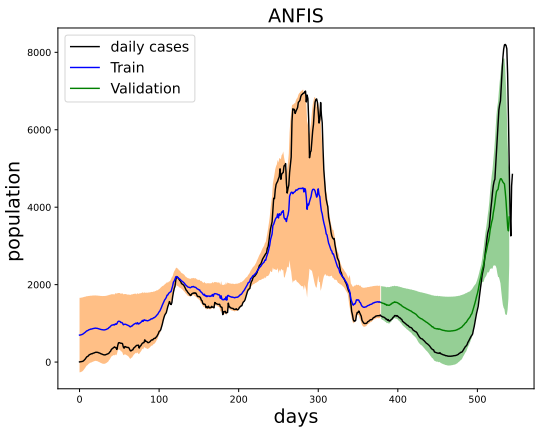


(c) LSTM

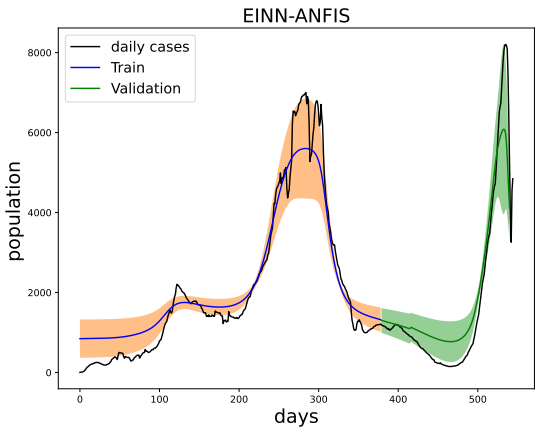


(d) EINN-LSTM

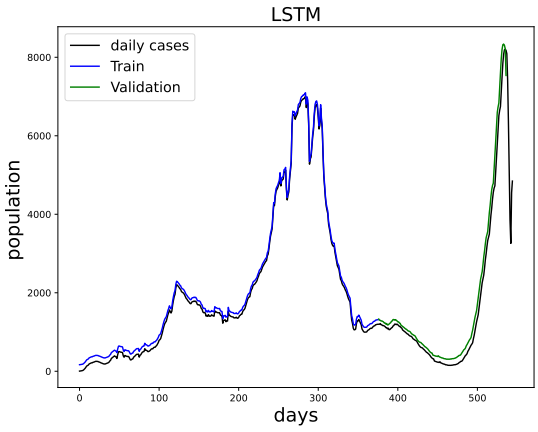
Figure 12: Missouri daily cases forecasting using ANFIS, EINN-ANFIS, LSTM, LSTM-EINN



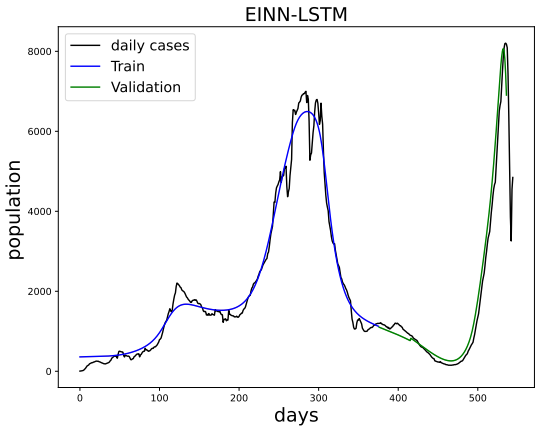
(a) ANFIS



(b) EINN-ANFIS

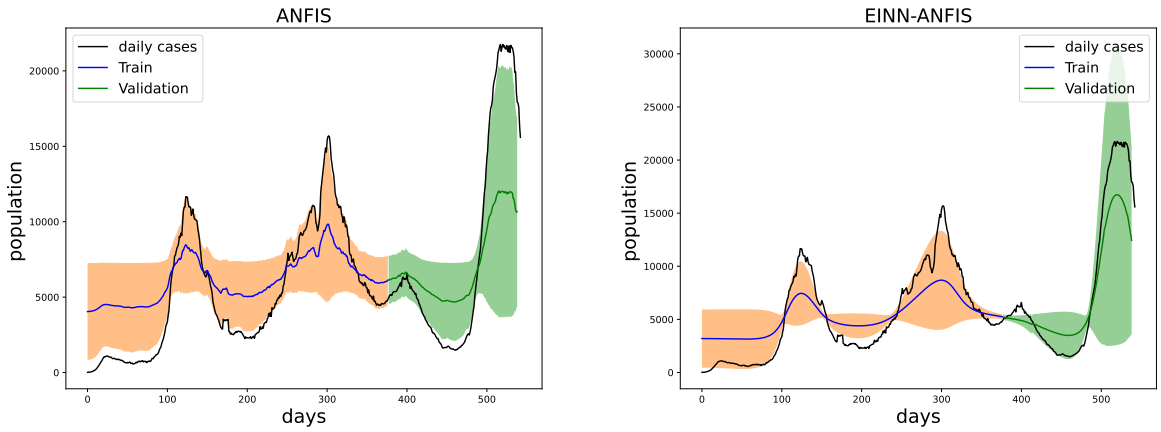


(c) LSTM



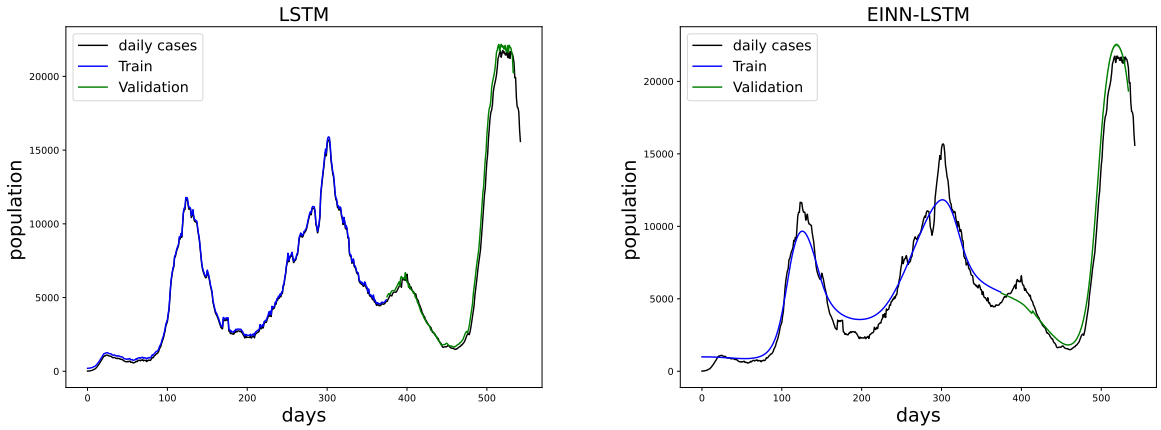
(d) EINN-LSTM

Figure 13: Tennessee daily cases forecasting using ANFIS, EINN-ANFIS, LSTM, LSTM-EINN



(a) ANFIS

(b) EINN-ANFIS



(c) LSTM

(d) EINN-LSTM

Figure 14: Florida daily cases forecasting using ANFIS, EINN-ANFIS, LSTM, LSTM-EINN

7. Performance analysis of error metrics

The following error metrics are used in our data driven simulation:

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(Y_i - \tilde{Y}_i)^2}$$

,

where Y and $\tilde{Y}$ are the predicted and original values, respectively.

- Mean Absolute Error (MAE):

$$MAE = \frac{1}{N_s}\sum_{i=1}^{N_s}|Y_i - \tilde{Y}_i|$$

.

- Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{N_s}\sum_{i=1}^{N_s}|\frac{Y_i - \tilde{Y}_i}{Y_i}|$$

.

- Root Mean Squared Relative Error (RMSRE):

$$RMSRE = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(\frac{Y_i - \tilde{Y}_i}{Y_i})^2}$$

,

N_s represents the sample size of the data.

In Table 10 We provide a comparison of error metrics for EINN using random splits for the training and test data.

State	*RMSE*	*MAE*	*MAPE*	*RMSRE*
Florida	0.00768642	0.0868862	0.62593067	1.4576211
Tennessee	0.01003919	0.07956579	1.41574097	3.11372066
Alabama	0.00772795	0.08293544	0.76998496	0.24813209
Missouri	0.0083841	0.09985308	1.65703082	0.44569263

Table 10: Error metrics for random split

8. Conclusion

We have presented a data-driven deep learning algorithm that learns time-varying transmission rates of multi-variant in an infectious disease such as COVID-19. The algorithm we presented learns the nonlinear time-varying transmission rates without a pre-assumed pattern as well as predict the daily cases and daily recovered populations. We learn these population groups using only daily cases data. This approach is found useful when the dynamics of an epidemiological model such as an SEIR model is impacted by various mitigation measures. The algorithm presented in this paper can be adapted to most epidemiology models. Using US daily cases data, we demonstrate that the algorithm presented in this work can be combined together with recurrent neural networks and ANFIS for an improved short-term forecast. This study is seen useful in the event of a pandemic such as COVID-19, where public health interventions and public response and perceptions interfere in the interaction of the compartments in an epidemiology model.

The computer codes will be available at `https://github.com/okayode/EINN-COVID`.

References

[1] World Health Organization (WHO), Archived: WHO Timeline-COVID-19, `https://www.who.int/news/item/27-04-2020-who-timeline--covid-19` (Accessed: 2021-08-12).

[2] E. Callaway, Making sense of coronavirus mutaions, Nature 585 (2020) 174–177.

[3] Centers for Disease Control and Prevention (CDC), Delta variant: What we know about the science, `https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html` (Accessed: 2021-08-12).

[4] K. Olumoyin, A. Khaliq, K. Furati, Data-driven deep-learning algorithm for asymptomatic covid-19 model with varying mitigation measures and transmission rate, Epidemiologia 2 (2021) 471–489.

[5] K. D. Olumoyin, A. Q. M. Khaliq, K. M. Furati, Data-driven deep learning algorithms for time-varying infection rates of covid-19 and mitigation measures, arXiv`doi:10.48550/ARXIV.2104.02603`.
URL `https://arxiv.org/abs/2104.02603`

[6] G. Cybenko, Approximation by superposition of a sigmoidal function, Mathematics of control, signals and systems.

[7] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks 4 (2) (1991) 251–257.

[8] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics informed deep learning: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics 378 (2019) 686–707.

[9] M. Raissi, N. Ramezani, P. Seshaiyer, On parameter estimation approaches for predicting disease transmission through optimization, deep learning and statistical inference methods, Letters in Biomathematics 6 (2) (2019) 1–26.

[10] J. Long, A. Khaliq, K. Furati, Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach, International Journal of Computer Mathematics 98 (2021) 1617–1632.

[11] M. Jagan, M. S. deJonge, O. Krylova, D. J. Earn, Fast estimation of time-varying infectious disease transmission rates, PLoS Computational Biology 16 (9) (2020) e1008124. doi:10.1371/journal.pcbi.1008124.

[12] D. He, J. Dushoff, T. Day, J. Ma, D. Earn, Inferring the causes of the three waves of the 1918 influenza pandemic in england and wales, Proc. R. Soc. 280 (2013) 20131345.

[13] Centers for Disease Control and Prevention (CDC), Variant proportions, https://covid.cdc.gov/covid-data-tracker/#variant-proportions (Accessed: 2021-08-20).

[14] K. Furati, I. Sarumi, A. Khaliq, Fractional model for the spread of COVID-19 subject to governmental intervention and public perception, Applied Mathematical Modelling 95 (2021) 89–105.

[15] D. Schaeffer, J. Cain, Ordinary Differential Equations: Basics and beyond, Springer, New York, 2016.

[16] P. Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Mathematical Biosciences 180 (2002) 29–48.

[17] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, Cambridge, 2016.

[18] M. Eftekhari, A. Yadollahi, A. Shojaeiyan, M. Ayyari, Development of an artificial neural network as a tool for predicting the targeted phenolic profile of grapevine Vitis vinifera foliar wastes, Frontiers in plant science (2018) 9:837.

[19] C. Du, J. Wei, S. Wang, Z. Jia, Genomic selection using principal component regression, Heredity 121 (1) (2018) 12–23.

[20] V. Chimula, L. Zhang, Time series forecasting of COVID-19 transmission in canada using lstm networks, Chaos Solitons Fractals 135 (2020) 109864.

[21] M. Hossain, S. Mekhilef, F. Afifi, L. Halabi, L. Olatomiwa, M. Seyedmahmoudian, B. Horan, A. Stojcevski, Application of the hybrid ANFIS models for long term wind power density prediction with extrapolation capability, PLoSONE 13 (4) (2018) e0193772.

[22] A. Vacilopoulos, R. Bedi, Adaptive neuro-fuzzy inference system in modelling fatigue life of multidirectional composite laminates, Computational Materials Science 43 (4) (2008) 1086–1093.

[23] J.-S. R. Jang, Anfis: adaptive-network-based fuzzy inference system, IEEE 23 (1993) 665–685.