Punjab University Journal of Mathematics (2025), 57(12), 1341-1366
https:://doi.org/10.52280/pujm.2025.57(12)04

On the Effectiveness of Data-Driven Models in Forecasting Dynamics of the Infectious Diseases in Comparison to Epidemiological Models

Received
07 October, 2025

Revised
11 May, 2025

Accepted
18 May, 2026

Published Online
11 June, 2026

Shomaila Mazhar
Department of Mathematics,
Islamia College Peshawar, Pakistan

* Zahid Ullah
School of Civil Engineering,
Fuzhou University, Fujian, China

Abstract. For decades, epidemiological models have been fundamental in exploring the dynamics of infectious diseases. They have been primarily used to investigate the infectious nature of diseases and to identify the controlling parameters. With the advancement of machine learning, data-driven models have increasingly gained the interest of researchers. In this regard, several studies have focused on application of data-driven models in transmission of infectious diseases. However, the effectiveness of these models relative to the epidemiological models remains largely unexplored. This study comprehensively evaluates the effectiveness of data-driven models in comparison to epidemiological models. For this purpose, a probability based epidemiological model SI^PRD, and two data-driven models Cov-CNN and Cov-LSTM are developed for forecasting Covid-19 infections. For evaluation and comparison, four existing epidemiological models; SIR, SIRD, SEIR and SEIRD are also considered. The effectiveness of epidemiological and data-driven models is evaluated through short- and long-term training and forecasting of Covid-19 infections in Afghanistan, Iran, and Iraq. In the initial stages of infections, epidemiological models are of prime importance in highlighting the important key controlling parameters. However, in long-term forecasting these models lack the ability to predict the infections accurately. On the other hand, the data-driven models, more specifically Cov-LSTM demonstrate superior performance.

AMS (MOS) Subject Classification Codes: 35S29; 40S70; 25U09
Key Words: Epidemiological Models; Covid-19 forecasting; CNN; LSTM.

**Corresponding Author*: ktk.zahid@gmail.com

1. INTRODUCTION

Throughout history, Infectious diseases have significantly impacted the health of humans across the globe. Historically, human populations have seen a number of epidemics that have swept through populations, causing significant mortalities and eventually disappearing. Some of these epidemics have resurfaced in different years, exhibiting varying degrees of severity and effect on population. Over time, a considerable number of deaths have been caused by some diseases that have become endemic. From ancient pandemics like the Black Death to contemporary pandemic like HIV/AIDS and Covid-19, these illnesses have impacted human history and remain a constant threat to our healthcare systems.

To better understand and effectively minimise the impact and adverse effects of infectious diseases, epidemiologists have been using mathematical models to simulate transmission of infections within populations. Ronald A. Ross and Anderson Grey McKendrick were the first ones to introduce the concept of mathematical modelling to epidemiology. In their models, the total population are classified into multiple compartments with different rates of variations. In the early epidemiological models, the total population was divided into compartments of susceptible and infections, signifying their importance in the transmission of a disease. However, later, extended models were developed by introducing additional compartments and taking into consideration the influence of different factors. The foremost purpose of epidemiological modelling is to describe the infectious nature of diseases and to facilitate policymakers in the development of effective methods to contain the infections. Besides the inherent infectious nature of an infection, its transmission also depends on number of external factors. Understanding of all these factors and their interaction and relation is necessary for development of a mathematical model for a precise representation of an infection. Since, the consideration of all such factors would result in highly expensive computational models and thus not be possible. Therefore, for the simplification of the epidemiological models, underlying assumptions are made, and only the most important parameters taken into consideration. In epidemiology, the simplest and fundamental models include Susceptible-Infected (SI), Susceptible-Infected-Susceptible (SIS), which classify the total population into two and three compartments only. Followed by these models are Susceptible-Infected-Recovered (SIR), Susceptible-Infected-Recovered-Susceptible (SIRS) [29], which signify the importance of the recovery compartment. At present, a number of epidemiological models with emphasis on different compartments have been developed. Examples includes Susceptible-Exposed-Infected-Recovered (SEIR), Susceptible-Exposed-Infectious-Removed-Vaccinated (SEIRV) and Susceptible-Exposed-Infected-Quarantined-Recovered (SEIQR) [35].

In November 2019, first case of coronavirus disease (Covid-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) was reported in Wuhan, China [38, 23]. In a shorter duration of few months, the infections spread globally, resulting in millions of mortalities and thus setting off global pandemic [16]. The date of the first verified (index) case was found to be November 17, 2019 [4]. The Covid-19 outbreak severely damaged the global economy by forcing many nations to impose partial or complete lockdowns and border closures [6, 22, 34]. Agriculture, a primary source of income of the people residing in rural regions of developing countries such as Pakistan, was significantly impacted by the sudden Covid-19 lockdown [13, 51]. The lockdown intervention policy has adversely

affected the low-income populations and migratory labour [33, 30, 32]. To mitigate the effect of the lockdown on low-income populations, the Pakistan government imposed a smart lockdown targeting only the infections hotspots instead of a complete lockdown.

Epidemiological models have played a pivotal role in analyzing the transmission dynamics of infectious diseases. SIR, being the most fundamental epidemiological model, has been the base of several extended infectious diseases models developed and proposed by different researchers in the last several decades [24, 49, 58, 7]. These models also include different models proposed for transmission and forecasting of the most recent highly infectious disease, Covid-19. Researchers have proposed different epidemiological models focusing on various aspects of Covid-19 infections. In [10], a modified SEIR model is introduced to analyze evaluate the influence of the various precautionary measures introduced by the policymakers on the dynamics of the Covid-19 epidemic. In [44] the authors consider a SEIQR model for the propagation of Covid-19. S. Mazhar et al [36] developed a probabilistic model with consideration of non-severe and severe infections for the prediction of Covid 19. In [18], the authors present a novel mathematical model designed to tackle dual diagnostic strategies and the isolation of confirmed cases. An Susceptible-Exposed-Infected-Hospitalized-Recovered-Pathogen (SEIHRP) model is introduced in [53], which considers the shedding effect is utilized to explore the transmission dynamics of the Covid-19 pandemic. In [57], the authors utilize a computational model to investigate the factors contributing to Covid-19 Associated Pulmonary Aspergillosis (CAPA) in immunocompetent patients. In [20], the authors present temporal and spatio-temporal Covid-19 epidemic models considering virus mutations and vaccination influences. Besides the models, researchers have focused on different aspects of the Covid-19 outbreak and developed number of epidemiological models taking into consideration different evolving parameter [52, 42, 61, 1, 41, 63]. In the study [21], the author develops a fractional TB mathematical model using a fractal–fractional differential operator (Mittag–Leffler type) to better capture the complex dynamics of tuberculosis spread in a population, with special emphasis on hospital treatment and public health education.

In recent decades, AI-based technologies have significantly enhanced management, particularly in the healthcare sector. The AI-based algorithms have been of great assistance in detecting infection through image processing, and thus in developing effective epidemic management policies. Recently, the importance and application of AI-based systems in the efficient management of global health emergencies such as Covid-19 have been extensively explored. New approaches, such as machine learning, deep learning, and fuzzy reasoning, have been implemented to accurately predict Covid-19 outbreaks. In [28, 12, 50], authors summarize various AI-based prediction algorithms, focusing on deep learning or machine learning methods. Xavier et al. [60] used least-squares criteria and a gradient descent algorithm for nonlinear regression analysis to predict Covid-19 cases. Alassafi et al., [2] simulated virus transmission across different countries to forecast mortality and confirmed cases. Sanjita et al. [27] employed artificial neural networks to predict overseas visitors, while [17] used a hybrid forecasting approach combining neural networks and ARIMA models to predict daily Covid-19 cases.

Among the different types of machine learning algorithms, the most effective deep learning techniques, that has been used in many research studies to predict Covid-19 cases is, Long Short-Term Memory (LSTM) [11, 47, 8]. Sweeti Sah, et al. [46] and Kumar Shui, et

al. [31] used the LSTM technique to forecast the number of Covid-19 cases in India. By applying the LSTM method in Bandung, Iksan et al. [26] forecasted the Covid 19 cases between March 2, 2020, and November 14, 2020. [54] the bi-directional LSTM method is applied to predict the Covid-19 infection in Indonesia between May 1, 2020, and December 15, 2021. The time series patterns of Covid-19 between the USA and India are compared using the LSTM model in[48]. In study [55], the author develops a mathematical model along with an Artificial Neural Network (ANN) to describe how alcohol use and addiction spread and evolve within a population. Considering the computational power and widespread application of deep learning algorithms in image processing, researchers have extensively explored their use in diagnosing Covid-19. In [56] the authors employed CNN-based image processing techniques to diagnose Covid-19 in its early stage from thermal image data. The authors in [9] proposed an image processing technique to distinguish Covid-19 based pneumonia from other variants of pneumonia using CT image data. Similarly, S. Wang et al.,[59] developed an AI-based model for detection of Covid-19 infection utilizing chest CT scans. In [19], the authors developed a CNN model to identify Covid-19 infections and forecast infections and mortalities with in a population. Similarly, Arora et al., [5] developed a stacked forecasting model named "Convolutional LSTM" for forecasting Covid-19 infections in India. In the study [37], the author develops a Physics-Informed Neural Network (PINN) alongside a novel fractional-order model to effectively capture long-term memory and hereditary effects in neurodegeneration. In [40] Nabi et al. focused on the comparison of multiple neutral networks i.e., LSTM, GRU, CNN and MCNN, in forecasting Covid-19 infections.

For several decades, epidemiological models have been pivotal in describing the transmission dynamics of different infectious diseases. The key objective of these models is to explore the transmission of the infections and to identify effective control parameters. With recent advancement in machine learning, researchers have been increasingly focusing on the application of deep learning models in forecasting infection trajectories. Despite the widespread use of neural networks in different research studies, the effectiveness of these models in comparison to the foundational epidemiological models remains largely unexplored. This study aims to explore the effectiveness of data-driven models in comparison to conventional epidemiological models for understanding the transmission dynamics of infectious diseases. To achieve this, we focus on the recent Covid-19 pandemic and develop a probability-based Covid19 epidemiological model named SI^PRD and two purely data-driven models. The effectiveness of these models is evaluated through comparison with several existing epidemiological models.

The remainder of this article is organized as follows: Section 2 presents a review of fundamental epidemiological models. Section 3 describes the methodology adopted in this study. Section 4 presents the results and discussion, while Section 5 provides the conclusion of the paper.

2. Notations and Preliminaries

The following abbreviations and notations are used throughout this study:

- **SIR**: Susceptible–Infected–Recovered model
- **SIRD**: Susceptible–Infected–Recovered–Deceased model

- **SEIR**: Susceptible–Exposed–Infected–Recovered model
- **SEIRD**: Susceptible–Exposed–Infected–Recovered–Deceased model
- **SI^{p}RD**: Susceptible–Infected–Recovered–Deceased model, where I^p represents the infected population
- **CNN**: Convolutional Neural Network
- **LSTM**: Long Short-Term Memory network
- **ANN**: Artificial Neural Network
- **MAE**: Mean Absolute Error
- **MSE**: Mean Squared Error
- **MAPE**: Mean Absolute Percentage Error
- **RMSE**: Root Mean Squared Error
- **REL**: Relative Error Loss

3. Review of Fundamental Epidemiological Models

During the Covid-19 pandemic, epidemiological models have been vital in predicting disease spread and assessing intervention effectiveness. Key to these models is the reproduction number (R_0), indicating potential spread and guiding public health responses, like social distancing and vaccination strategies. In the following section, we have briefly described the most used epidemiological models i.e., SIR, SIRD, SEIR, and SEIRD, with the aim of their comparison with proposed probability-based and data-driven models. These are the most used and fundamental and relatively simple models. Most of the advanced models proposed by different researchers are based on the extension of these models through incorporation of additional parameters.

3.1. **SIR Model.** The SIR model is the foundational epidemiological model, introduced in 1927. It is a classical mathematical model of infectious diseases, based on the assumption of dividing the total population N into three compartments: susceptible (S), Infected (I), and Recovered (R). and thus, the total population is given by $N = S + I + R$. Figure 1 illustrates the block diagram of the SIR model, showing the flow of individuals among the three compartments. The susceptible compartment changes over time due to interactions between susceptible and infected individuals. Disease transmission occurs at the transmission rate β. Therefore, the dynamics of the susceptible population are given by:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{3. 1}$$

The infected compartment changes over time as susceptible individuals move into the infected class after exposure to the disease. Furthermore, infected individuals leave this compartment upon recovery at the recovery rate γ. Hence, the dynamics of the infected compartment are represented as:

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \tag{3. 2}$$

Finally, the recovered compartment increases as infected individuals recover from the disease and move into the recovered class. The dynamics of the recovered compartment are given by:

$$\frac{dR}{dt} = \gamma I \tag{3. 3}$$

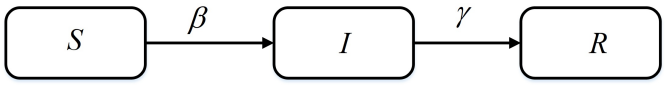

FIGURE 1. Block diagram of SIR Model

The basic reproduction number is a vital parameter of any epidemiological model which represents the expected number of infections that a single infectious host will cause upon introduction to a population of N susceptible hosts. It is a measure of the infectious nature of a disease. The basic reproduction number of SIR model is $R_0 = \frac{\beta}{\gamma}$. Thus, for an infection of transmission rate higher than the recovery rate, reproduction number will be greater than one, indicating endemic.

3.2. **SIRD Model.** The SIRD model is developed by extending the classical SIR model through introducing a Death (D) compartment, and thus dividing the total population into four categories, as illustrated in block diagram in Figure 2. The susceptible compartment behaves similarly to that in the SIR model. The corresponding equation is given by:

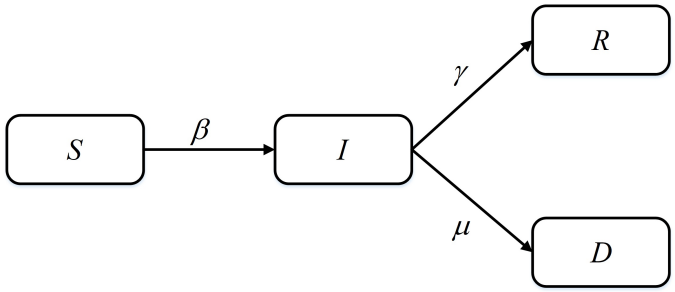

FIGURE 2. Block diagram of SIRD Model

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{3. 4}$$

The infected compartment also changes in a manner similar to the SIR model, where susceptible individuals enter the infected class after disease transmission. However, infected individuals may leave this compartment either through recovery at the rate γ or through disease-related death at the rate μ. Therefore, the dynamics of the infected compartment are represented as:

$$\frac{dI}{dt} = \frac{\beta SI}{N} - (\gamma + \mu)I \tag{3. 5}$$

The recovered compartment is the same as in the SIR model. The corresponding equation is:

$$\frac{dR}{dt} = \gamma I \tag{3. 6}$$

In addition, the SIRD model includes a death compartment, which represents disease-related deaths. Individuals enter this compartment from the infected class at the death

rate μ. The dynamics of the death compartment are given by:

$$\frac{dD}{dt} = \mu I \tag{3. 7}$$

The basic reproduction number of SIRD model is $R_0 = \frac{\beta}{\gamma+\mu}$. The basic reproduction number greater than one shows a higher transmission rate, resulting in an epidemic.

3.3. **SEIR Model.** The SIR and SIRD models describe disease transmission as a whole and epidemic development in general. However, at compartmental level, it does not take into consideration the incubation period. The SEIR model is extension of SIR model through incorporation of latent population [1]. It is among the most used mathematical models for studying the dynamics of infectious diseases in fixed population of size N. In SEIR model, the total population N is divided into four compartments: Susceptible (S), Exposed (E), Infected (I), and Recovered (R), as illustrated in Figure 3.

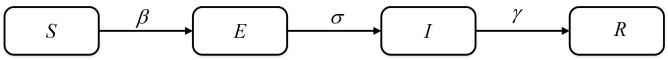

FIGURE 3. Block diagram of SEIR Model

The susceptible compartment behaves similarly to the SIR and SIRD models. The dynamics of the susceptible compartment are given by:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{3. 8}$$

The exposed compartment represents individuals who have been exposed to the disease after interacting with infected individuals but are not yet infectious. These individuals remain in the exposed compartment during the latent period and leave the compartment after completing this period, becoming infected at the rate of σ. The dynamics of the exposed compartment are represented as:

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{3. 9}$$

The infected compartment increases as exposed individuals complete the latent period and move into the infected class. Infected individuals leave this compartment through recovery at the recovery rate γ. Therefore, the dynamics of the infected compartment are given by:

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{3. 10}$$

The recovered compartment behaves similarly to the SIR and SIRD models. The corresponding equation is:

$$\frac{dR}{dt} = \gamma I \tag{3. 11}$$

The basic reproduction number of SIRD model is $R_0 = \frac{\beta}{\gamma+\sigma}$

3.4. **SEIRD Model.** The SEIRD model extends the SEIR model by incorporating infection-induced mortalities through the addition of a Deceased compartment. Thus, SEIRD model offers a comprehensive framework for evaluating the impact of infectious diseases on population by considering mortality (deceased) and recovery compartments. The model is particularly appropriate for diseases with higher mortality rates, allowing policymakers to assess the efficacy of interventions in reducing both morbidity and mortality. The model classifies a total population of size N, into 5 compartments, i.e., Susceptible (S), Exposed (E), Infected (I), Recovered (R) and Deceased (D), as shown in Figure 4.

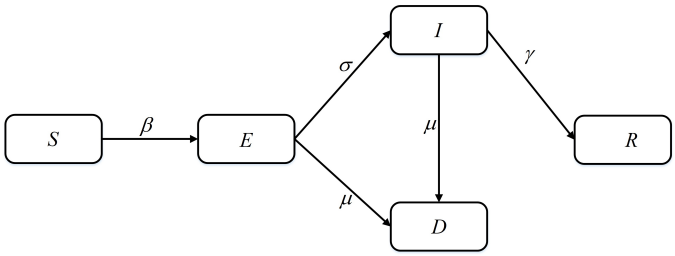

FIGURE 4. Block diagram of SEIRD Model

The susceptible compartment remains the same as in the SIR, SIRD and SEIR model. The susceptible population decreases due to interactions with infected individuals at the transmission rate β. The dynamics are given by:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{3. 12}$$

The Exposed compartment remains the same as in the SEIR model. The only difference is that the individual leaves the compartment either by becoming infectious at the rate σ or through disease-induced death at the rate μ. Therefore, the exposed compartment dynamics are:

$$\frac{dE}{dt} = \frac{\beta SI}{N} - (\sigma + \mu)E \tag{3. 13}$$

The infected compartment receives individuals from the exposed class once they complete the latent period. In addition to recovery at the rate of γ, infected individuals may also die due to the disease at rate μ. Hence, the infected compartment is described by:

$$\frac{dI}{dt} = \sigma E - (\gamma + \mu)I \tag{3. 14}$$

The recovered compartment remains the same as in previous models, where infected individuals recover at rate γ and move into the recovered class:

$$\frac{dR}{dt} = \gamma I \tag{3. 15}$$

Finally, the death compartment accounts for disease-induced mortality. Individuals enter this compartment from both the exposed and infected classes due to disease-related death at rate μ. The dynamics are given by:

$$\frac{dD}{dt} = \mu E + \mu I \tag{3. 16}$$

The Basic reproduction number for the SEIRD model is $R_0 = \frac{\beta\sigma}{(\sigma+\mu)(\gamma+\mu)}$.

4. METHODOLOGY

In this study, a broad comparison is performed between classical epidemiological models and artificial neural network (ANN) models. On one side, the basic epidemiological models, namely SIR model, SIRD model, SEIR model, SEIRD model, and a novel proposed epidemiological model, the SI^pRD model, are considered. On the other side, the widely used ANN models, belonging to the class of deep learning architectures, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), are employed. The performance of the epidemiological models is then compared with that of the ANN-based models. The rationale behind this comparison is that epidemiological models describe disease transmission dynamics and provide insight into the progression of an outbreak. These models can estimate important epidemiological quantities, such as the reproduction number, and help determine whether a disease is likely to become epidemic or endemic. In contrast, ANN models are primarily data-driven approaches that learn patterns directly from historical data and are capable of capturing complex nonlinear relationships without relying on epidemiological assumptions. However, they generally require large amounts of data for effective training. Therefore, this comparison helps in understanding whether ANN-based models can provide improved short-term forecasting performance compared to deterministic epidemiological models. For the analysis, each epidemiological model was calibrated using reported epidemic data to perform short-term fitting and prediction. Similarly, the CNN and LSTM models were trained on the same datasets for forecasting purposes. The predictive performances of all models were then compared using appropriate error measures and statistical indicators.

4.1. **Dataset Description.**

The dataset used in this study consists of the reported infected cases obtained from the World Health Organization for three countries: Afghanistan, Iran, and Iraq. The data include daily confirmed COVID-19 cases recorded from 1 March 2020 to 29 January 2022. Prior to model implementation, the datasets were pre-processed to remove inconsistencies and prepare the data for model fitting and validation. The processed data were then used for calibration, training, and performance evaluation of both the epidemiological and deep learning models. All epidemiological as well as ANN-based models were trained using 80 percent of the data, while the remaining 20 percent of the data were reserved for validation and prediction performance assessment.

4.2. **Proposed Probability-Based SI^pRD Epidemiological Model.**

Most of the epidemiological models effectively describe the transmission dynamics of infections and are useful in identification of controlling parameters in the infectious diseases. However, these models lack the capability to accurately forecast infection trends over time. This limitation hinders their effectiveness in predicting future outbreaks and understanding the potential trajectory of disease spread, which is critical for public health planning and intervention strategies. For this purpose, a novel epidemiological model named SI^pRD that uses probability-based factors to describe the transmission of infections among different compartments, is proposed. The SI^{p}RD model consists of four compartments i.e., Susceptible (S)-Infected (I)-Recovered (R)-Death(D), as shown in the block diagram in Figure 5. The mortalities in susceptible and recovered compartments are denoted by a parameter b. The mortality due

to infections is assumed to occur within the infected compartment only, denoted by μ. For simplification purposes, the natural death and fertility rate are assumed equal, and both are denoted by b. Thus, the total population remains constant and is denoted by N, and at any instant of time t, it is equivalent to the sum of populations in all compartments. The flow of the populations among different compartments is depicted in Figure 5.

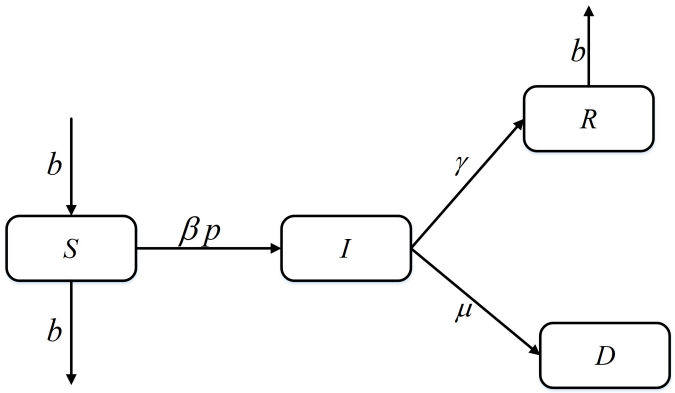

FIGURE 5. Block diagram of SI^{p}RD Model

The transmission of Covid-19 infections depends on exposure of a susceptible individual to an infected individual. Let p be the probability of a susceptible individuals getting infectious upon their contact with an infected person. Similarly, q denotes the probability of the susceptible individual remaining susceptible despite its exposure to infection. Thus, the probability that to become an infectious individual is $\frac{SpI}{pI+qS}$ and similarly the probability of the individual to remains susceptible is $\frac{qS}{pI+qS}$. If $\boldsymbol{\beta}$ is the transmission rate, the rate at which susceptible individuals become infected and transition into the infected compartments is determined by the incidence function $\frac{\beta SpI}{pI+qS}$. Based on the assumptions, the following model is constituted: The susceptible compartment changes over time due to the incidence function describing disease transmission. In addition, individuals enter the population through recruitment at rate b, while natural death also occurs in the population at rate b. Therefore, the dynamics of the susceptible compartment are given by:

$$\frac{dS}{dt} = bN - \frac{\beta SpI}{pI + qS} - bS \tag{4. 17}$$

The infected compartment receives individuals from the susceptible compartment through the incidence term. Infected individuals leave the compartment either by recovery at rate γ or by ,disease-induced death at rate μ. Hence, the infected dynamics are:

$$\frac{dI}{dt} = \frac{\beta SpI}{pI + qS} - (\gamma + \mu)I \tag{4. 18}$$

The recovered compartment consists of individuals who recover from the infected compartment at rate γ. In addition, natural death also affects this compartment at rate b. Therefore,

the recovered compartment is given by:

$$\frac{dR}{dt} = \gamma I - bR \tag{4. 19}$$

The death compartment changes over time by accounting for disease-related mortality occurring in the infected compartment. Specifically, individuals who die due to the infection are transferred from the infected compartment to the death compartment at the rate μ. Therefore, the dynamics of the death compartment are given by:

$$\frac{dD}{dt} = \mu I \tag{4. 20}$$

Equation (4. 17)-(4. 20) consists of system of first order differential equations which represents the time rate of change of population in susceptible, infected, Recovered and Death compartments. To compute the basic reproduction number, the Next Generation Matrix approach is used. For this purpose, the infected compartment is set to zero, which leads to the disease-free equilibrium point of the model. At this equilibrium, the susceptible population is given by $S^* = N$. The infected compartment of the model is given by the following equation:

$$\frac{dI}{dt} = \frac{\beta SpI}{pI + qS} - (\gamma + \mu)I \tag{4. 21}$$

At the disease-free equilibrium point, where $S^* = N$, the equation reduces to:

$$\frac{dI}{dt} = \frac{\beta pI}{q} - (\gamma + \mu)I \tag{4. 22}$$

The new infection term be $F = \frac{\beta p}{q} I$ and the transition term be $V = (\gamma + \mu)I$. So, the basic reproduction number R_0 is calculated as the dominant eigenvalue of the matrix FV^{-1}. Hence, the basic reproduction number is given by the following:$R_0 = \frac{\beta p}{q(\gamma+\mu)}$. The system of first order differential equations needs to be numerically solved to determine the population in each compartment of the model at each instant of time. In this study, fourth order Runge-Kutta method is used to solve the differential equations. The initial population (value) in each compartment is obtained from the historically Covid-19 data. Besides the initial values, the model also requires the initial values of the unknown parameter i.e., $\beta,\ q,\ \gamma,\ \mu,\ p$. For this purpose, initial guess of the unknown parameters is also provided and are iteratively updated to determine the optimum values through minimization of the objective function. The optimum values of the parameters are used for Covid-19 forecasts. An in-house MATLAB based on the Runge-Kutta method is developed to perform the fitting and forecasting of the Covid-19 infections.

4.3. **Data-driven Covid-19 Models.** For decades, epidemiological models have been of great assistance to policy makers in public health and their importance cannot be denied. However, recent years have witnessed a surge in interest among researchers towards data-driven models, particularly in the realm of infectious diseases. In this section, we have formulated two distinct data-driven models, named Cov-CNN and Cov-LSTM, particularly designed for prediction of Covid-19 infections. The former model is based on Convolutional Neural Network (CNN) architecture to assess its efficacy in capturing the transmission dynamics of the infections. The later one is based on Long Short-Term Memory

(LSTM) neural network architecture to explore its effectiveness in modeling the dynamics of infections.

4.3.1. *Cov-CNN.* Convolutional Neural Network (CNN) is a class of deep learning algorithms, initially introduced in 1980 by LeCun [53], and primarily used for analyzing visual data. Nowadays, the CNN based neural networks are commonly used in the problems of classification through object detection and image processing. It is widely used in different fields of science and technology. Despite being a powerul tool in classification and object detection, CNN based models have also been employed in forecasting of time series data due to their robust feature extraction capabilities, examples include forecasting stock prices, air quality, and energy load [54-56, 62]. In the current study, a CNN based algorithm named Cov-CNN is developed with the aim of its application in forecasting of Covid-19 infections trajectory within a population. The Cov-CNN architecture comprises of Con-volutional, pooling and finally a fully connected layer. The convolutional layer identifies patterns in sequential data through filters and the pooling layer down-samples and combines data to reduce dimensionality while preserving essential the features. The fully connected layer captures complex patterns and relationships within the extracted features, facilitating the model's comprehensive understanding and prediction capabilities. The general archi-tecture of Cov-CNN is illustrated in Figure 6.

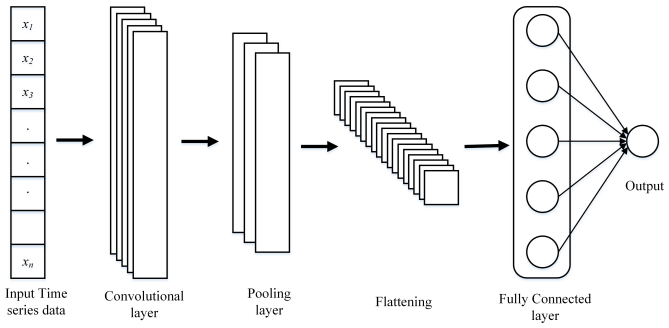

FIGURE 6. Schematic illustration of Cov-CNN Architecture

The Cov-CNN model architecture begins with a 1D convolutional layer with 64 kernels of size 3, designed to identify patterns in daily infection cases. The model uses ReLU activation for non-linearity. The convolutional layer is followed by a pooling layer with the purpose of the reducing the size and thus the computational cost of the model. In Cov-CNN, the down-sampling is performed through a 1D max-pooling layer with a pool size of 2. The input is divided into one-dimensional pooling region, and the maximum of each region is computed. The pooling layer is followed by a fully connected dense layer and thus its acts as bridge between the convolutional and fully connected layer. The convolutional layer consists of 50 neurons, and it uses ReLU activation function. The ReLU function enables the neural network to learn complex patterns and variation in the data through the introduction of non-linearity. The flattened output from the previous is received as input by the fully connected dense layer.

4.3.2. *Cov-LSTM.* Long Short-Term Memory (LSTM) is an extension of Recurrent Neural Network (RNN) primarily designed to overcome the shortcomings of the conventional RNN. It was first presented in 1997 by [25]. The primary limitation of RNN is its inability to store information for longer duration, when it might be required in forecast of an output. Thus, the conventional RNN is incapable of handling long-term dependencies in data. The LSTM neural network was designed to effectively overcome the vanishing gradient limitation of the RNN. The LSTM neural networks have been found to be very effective in time-series forecasting and have been extensively used in different fields of science and technology. In the current study, a Cov-LSTM i.e., an LSTM based architecture is developed to evaluate its effectiveness in time short and long-term forecasting of Covid-19 infections. Figure 7 illustrates the single LSTM cell with three gates to regulate the flow of information i.e., the forget, input and output gates. Detailed explanations of the three gates are provided as follows:

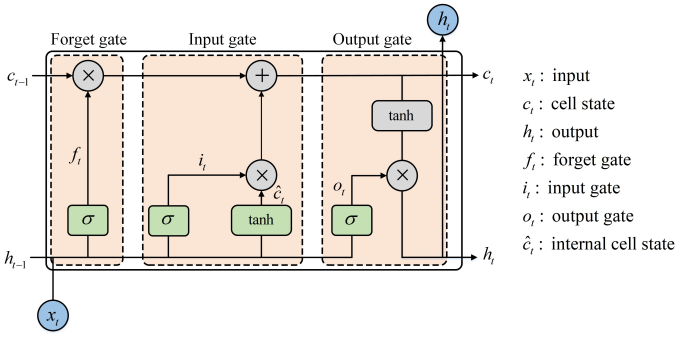

FIGURE 7. LSTM single cell

The forget gate determines which information to discard by using the input x_t and previous hidden state $h_(t-1)$ with a sigmoid function to produce the forget gate

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{4. 23}$$

The input gate decides which information to store, processing x_t and $h_{(t-1)}$ through a sigmoid function and a tanh function to create a vector $\tilde{c}_t$, given as

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{4. 24}$$

$$\tilde{c}_t = \tanh\left(W_c\left[h_{t-1}, x_t\right] + b_c\right) \tag{4. 25}$$

The cell state is then updated:

$$c_t = f_t\, c_{t-1} + i_t\, \tilde{c}_t \tag{4. 26}$$

The output gate determines the output information using x_t and $h_{(t-1)}$ with another sigmoid function. The output gate activation values are determined as

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{4. 27}$$

The output gate regulates and controls the flow of information from the current cell state. Point-by-point multiplication is performed on the two outputs to compute h_t at time t, given

as

$$h_t = o_t \tanh(c_t) \tag{4. 28}$$

The architecture of Cov-LSTM model consists of a number of layers to detect long-term dependencies in sequential data. The initial Cov-LSTM layer has 100 units with a sigmoid activation function, followed by other layers of 75- and 50-units using tanh and ReLU activations, respectively. The final output of the model is computed by a dense layer single unit. The schematic illustration of the Cov-LSTM architecture is shown in Figure 8. The model uses the Adam optimiser based on the stochastic gradient descent method and mean squared error loss function, and it is trained for 300 epochs, ensuring effective parameter updates and accurate evaluation. The detailed features of the Cov-CNN and Cov-LSTM are given in Table 1.

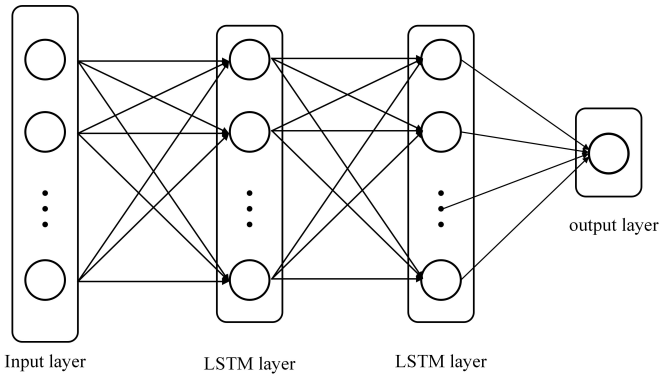

FIGURE 8. Schematic illustration of Cov-LSTM Architecture

TABLE 1. Model configuration of Cov-LSTM and Cov-CNN

Parameter	**Cov-LSTM**	**Cov-CNN**
Optimization method	Adam (stochastic gradient descent)	Adam (stochastic gradient descent)
Nonlinear Activation	Sigmoid, Tanh, ReLU	ReLU
Kernel		3
Hidden Units	100, 75, 50, 50	50
Hidden Layers	4	1
Number of Epochs	300	300
Convolutional Filters		64

4.4. **Data Processing and Models Training.** In the development of data-driven models such as Cov-CNN and Cov-LSTM, the initial step is data collection from a database and its preprocessing. The construction and monitoring of the WHO Covid-19 database consists of a complex data flow process from multiple sources. In this study, Covid-19 is data is collected from Our World in Data [14], a database coordinated with WHO Covid-19

surveillance database [14, 15, 3]. The acquired data is normalized to a standard range and is divided into two sets, i.e., training and testing datasets. The training dataset is used for training the Cov-LSTM and Cov-CNN models, while the testing dataset is used for validation testing purposes. The trained model can then be used for prediction of Covid-19 infections. The stepwise training process is illustrated in Figure 9.

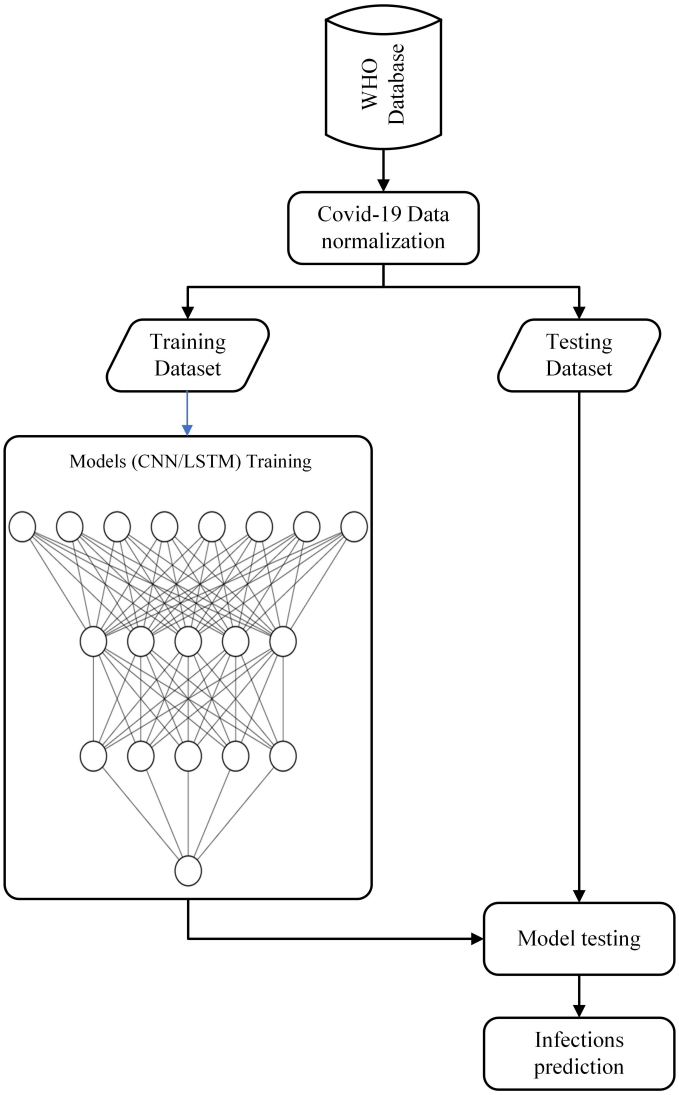

FIGURE 9. Data processing and training flowchart of data-driven model

Following the training, the performance of Cov-CNN and Cov-LSTM in terms of their ability to forecast the infections are assessed through comparison with the actual Covid-19

infections. For this purpose, four different evaluation metrics i.e., Mean Absolute Error (MAE), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Relative Error (REL), are employed to evaluate the accuracy of the model's forecast.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_h^i - y_p^i\right|}{y_h^i} \times 100\% \tag{4. 29}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_h^i - y_p^i\right| \tag{4. 30}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_h^i - y_p^i\right)^2} \tag{4. 31}$$

$$\text{REL} = \sum_{i=1}^{n}\frac{\left(y_h^i - y_p^i\right)^2}{\left(y_h^i\right)^2} \tag{4. 32}$$

where historical Covid-19 data is denoted by y_h^i and forecasted Covid-19 infections are denoted by y_p^i. The total number of data points are denoted by n. The purpose of using multiple evaluation matrices is to provide a more detailed understanding of the strength and weakness of each of the two data-driven models. Each evaluation matrix provides insights into different aspects of the model such as precision, sensitivity, and accuracy. MAPE computes the mean percentage difference between the historical value and the forecasted value of the infection. Due to its error computation in terms of percentage, therefore, it is quick to understand and interpret and is independent of the scale. MAE provides the mean absolute difference between the historical value and the forecasted value of the infection. It provides errors in the unit of the input data. RMSE provides the square root of the mean of square difference between the historical value and the forecasted value of the infection. It focuses on the error aspect of the model and is more appropriate for optimization algorithm with undesirable large errors. The final evaluation matric REL simply provides an average squared difference between the historical value and the forecasted value of the infection.

5. Results and Discussions

For the analysis of epidemiological and data-driven models, Covid-19 data of three Asian countries, i.e., Afghanistan, Iran, and Iraq, is acquired from [59]. To evaluate the effectiveness of the data-driven models, a comparison of data fitting and prediction is conducted in two stages, i.e., short- and long-term. The purpose of the short-term data fitting and prediction is to identify the most effectives models that can be further analysed in the long-term prediction.

5.1. **Short-term models analysis.** For evaluation of epidemiological and data-driven models in short-term data analysis, Covid-19 data of 600 days, i.e., March 01, 2020, till October 22, 2021, is used for fitting and training purposes. The subsequent 100 days Covid-19 data, i.e., October 23, 2021, till January 29, 2022, is used for forecasting purposes, as given in

Table 2. The fitting accuracy of the models significantly varies across different epidemiological models. In this study, we have evaluated the basic models, SIR and SIRD, intermediate models, SEIR and SEIRD, and advanced models SI^pRD and data-driven models, Cov-CNN and Cov-LSTM, as discussed in the following sections.

TABLE 2. short-term training and forecasting daily data

Data set	**Days**	**Date**
Training	600	01/03/2020 – 22/10/2021
Forecasting	100	23/10/2021 – 29/01/2022

The SIR and SIRD models are fundamental in understanding transmission dynamics of infections; however, demonstrates poor fit to the Covid-19 infection data across all three countries. These models are basic and simple as they do not consider the latent period and detailed recovery, death, and re-infections, which likely contributes to their poor performance. Figure 10(a), 11(a) and 12(a) shows short-term fitting of the SIR and SIRD models across three different countries i.e., Afghanistan, Iran and Iraq, respectively. These models oversimplify the disease dynamics and thus do not perform well in fitting and forecasting. The SEIR and SEIRD models, which introduce an exposed compartment into the model, prior to infections of individuals, are relatively better in fitting to the actual infections in comparison to the basic models. The introduction of the exposed compartment allows these models (SEIR and SEIRD) to account for incubation period of the infection, which adds to the better fitting performance of the models. The improved fitting of the SEIR and SERID models across all three countries, as shown in Figure 10(b), 11(b) and 12(b), highlights the importance of the latent period of infections.

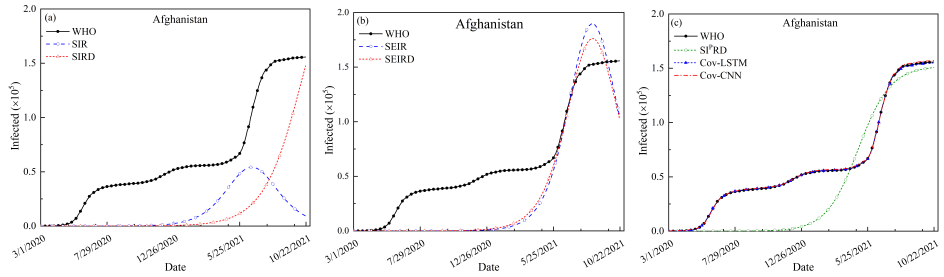

FIGURE 10. Performance comparison of epidemiological and data-driven models fitted to Afghanistan's COVID-19 short-term dataset over 600 days.

Although the intermediate models such as SEIR and SEIRD performs well than the fundamental models, they still lack the dynamics of infections in detail. In this regard, our proposed SI^PRD demonstrates improved performance over the traditional epidemiological models (SIR, SIRD, SEIRD and SEIRD). The proposed model incorporates additional parameters to better capture the dynamics of disease transmission and progression, resulting in a more accurate fit to the actual infection data of three countries, as shown in Figure

10(c), 11(c) and 12(c). The data-driven models, i.e., Cov-CNN and Cov-LSTM outperformed all the considered epidemiological models in terms of fitting to the Covid-19 infection data. These models capture complex dynamic behavior and learn the patterns in the data that are not accurately captured by the epidemiological models. Among the data-driven models the Cov-LSTM model is superior in terms of fitting to the Covid-19 data in comparison to the Cov-CNN, across all three countries, as illustrated in Figure 10(c), 11(c) and 12(c).

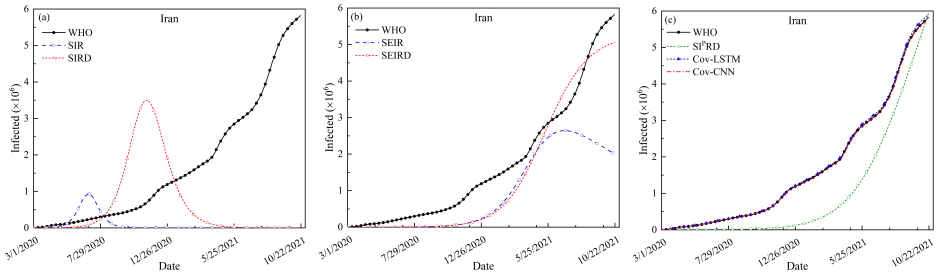

FIGURE 11. Performance comparison of epidemiological and data-driven models fitted to Iran's COVID-19 short-term dataset over 600 days.

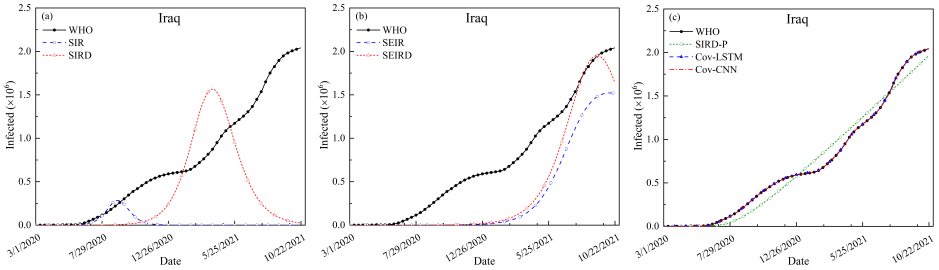

FIGURE 12. Performance comparison of epidemiological and data-driven models fitted to Iraq's COVID-19 short-term dataset over 600 days.

Short-term fitting of the models shows that the Epidemiological models such as SIR, SIRD, SEIR and SERID do not fit well to the actual Covid-19 data, and thus they are not advanced enough to forecast the trend of the infections accurately. Since the proposed SIPRD, Cov-CNN and Cov-LSTM perform well in terms of fitting, thus we have only considered these models to analyze the effectiveness of these models in forecasting the infections. Figure 14 shows forecasted trajectories of Covid-19 infections computed through SIPRD, Cov-CNN, and Cov-LSTM. The results show that the epidemiological model SI^pRD forecasts the general trend of the infection transmission, providing a broad understanding of how the infection spreads and peaks. The data-driven models Cov-CNN

and Cov-LSTM significantly outperform the SI^pRD model in terms of forecasting accuracy. Among the two data-driven models, the Cov-LSTM has superior forecasting accuracy to the counter Cov-CNN model. The difference in the forecasting accuracy of the Cov-CNN and Cov-LSTM is clearer in Figure 13. These results show that data-driven approaches are a promising tool for real-time infections forecasting.

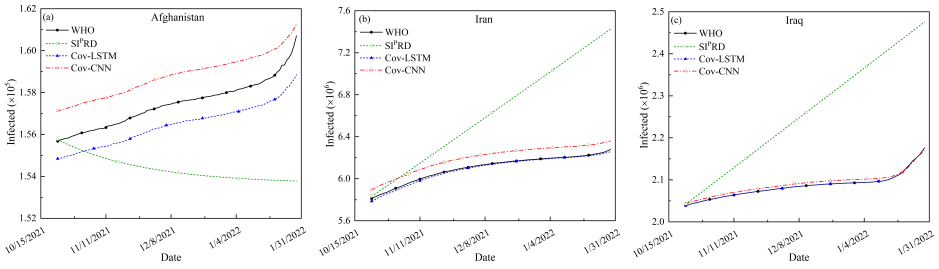

FIGURE 13. Short-term (100-day) forecasting of infected cases.

5.2. **Long-Term models analysis.** Following the short-term fitting and forecasting analysis, we intended to evaluate the effectiveness of the SI^PRDdata-driven models in long-term analysis. For this purpose, we have conducted the data fitting and training of the selected models for 944 days with a subsequent forecasting of infections for 405 days. The details of training data and forecasting data are provided in Table 3.

TABLE 3. Daily training and validation data long-term

Data set	**Days**	**Date**
Training	944	01/03/2020 – 30/09/2022
Forecasting	405	01/10/2022 – 09/11/2023

In the long-term performance evaluation, the proposed SI^pRD model shows limitation in terms of fitting accurately to the historical Covid-19 infection data of three countries. The epidemiological model SI^pRD shows the global trend of the Covid-19 infections; however, it does not represent the local variation in the infections, as shown in Figure 14. The results highlight the limitation of the epidemiological models in representing the dynamics of infectious diseases over an extended period. The presumptions made in epidemiological models limit their ability to take into consideration the evolving factors and their influence on disease transmission for an extended period. In contrast to the epidemiological model, the data-driven models perform very well and fit accurately to the Covid-19 data of all the three countries, as shown in Figure 14. The data-driven models capture the variation in the disease dynamics with significant accuracy. This highlights the strength of the data-driven models over epidemiological models.

To evaluate the effectiveness of the epidemiological model, i.e., SI^PRD and data-driven models, Cov-CNN and Cov-LSTM, in long-term forecasting, a 405-day forecast was conducted and plotted against the actual infection data of Afghanistan, Iran, and Iraq.

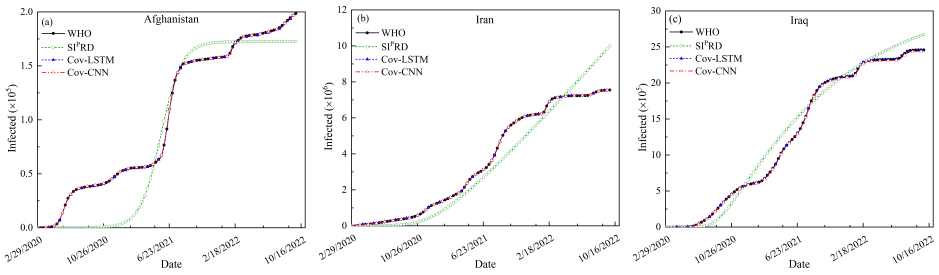

FIGURE 14. Fitting of SIPRD, Cov-LSTM, and Cov-CNN of 944 days

The SI^PRD model, while able to show the general transmission trend, lacks accuracy and reliability in long-term forecasting, as shown in Figure 15. The inherent limitation in the epidemiological models is due to the underlying assumptions and the static optimised parameters. The Cov-CNN and Cov-LSTM model demonstrations significantly improved performance in long-term forecasting of Covid-19 infections across all three countries, as shown in Figure 15. Among the two data-driven models, the Cov-LSTM model exhibits superior forecasting accuracy due to its capability of retaining long-term dependencies, which allows it to adapt to variation in infections. Cov-CNN also forecasts the transmission of infections with acceptable accuracy.

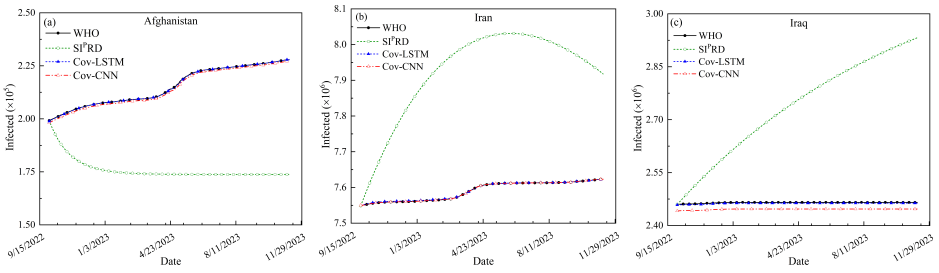

FIGURE 15. Long-term prediction of models

5.3. **Training Behavior and Error Metrics.** For the training and validation of the Cov-CNN and Cov-LSTM, total data is divided into training and validation dataset. Where the training dataset consists of 80 percent of the total infections data and the remaining 20 percent is used for validation purposes. For training of Cov-CNN and Cov-LSTM 300 epochs are used. The training and validation losses of the Cov-CNN are shown in Figure 16, which illustrates the training and validation loss of the Cov-CNN model for Afghanistan, Iran, and Iraq over 300 epochs. In all three countries, the training loss shows a steady decreasing trend, indicating effective learning of the underlying data patterns. Similarly, the validation loss decreases alongside the training loss and gradually stabilises, suggesting consistent convergence of the model. The close alignment between training and validation loss curves indicates good generalisation performance and minimal overfitting. Overall, the

Cov-CNN model demonstrates stable learning behaviour and strong predictive capability across all three datasets, confirming its robustness for modelling Covid-19 dynamics in different regional contexts.

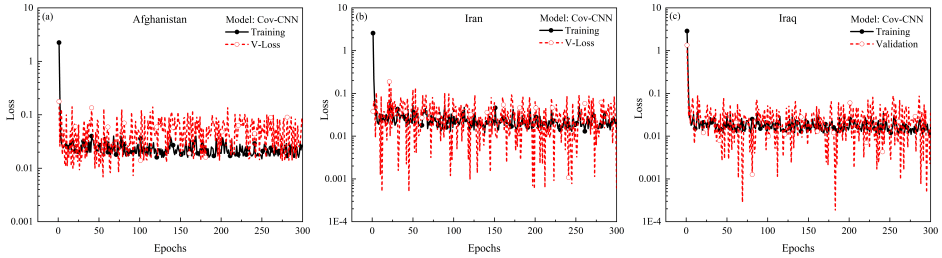

FIGURE 16. Long-term prediction of models

Furthermore, the training and validation loss of the Cov-LSTM model is shown in Figure 17 for Afghanistan, Iran, and Iraq over 300 epochs exhibit a similar parallel behaviour in all three countries. This indicates that the model effectively learns the underlying data patterns while maintaining stable generalisation performance. The close alignment between training and validation loss curves suggests consistent convergence and minimal overfitting. Overall, the Cov-LSTM model demonstrates improved and more stable performance compared to the Cov-CNN model, yielding better predictive results across all datasets.

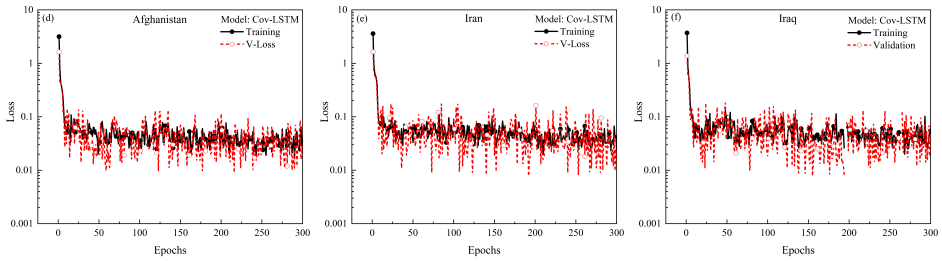

FIGURE 17. Long-term prediction of models

To further evaluate the performance of both ANN-based models, Cov-CNN and Cov-LSTM, a comparative analysis of error metrics has been presented using bar plots in Figure 18. The figure illustrates the performance comparison in terms of MAE, MAPE, RMSE, and REL for Afghanistan, Iran, and Iraq. In the visualisation, the yellow bars represent the Cov-CNN model, while the pink bars correspond to the Cov-LSTM model. The results indicate that both models achieve promising predictive performance across all three countries. However, the Cov-LSTM model consistently outperforms the Cov-CNN model in terms of lower error values, as also confirmed in the corresponding results in Table 4. This demonstrates the superior accuracy and robustness of the Cov-LSTM approach for modelling and forecasting Covid-19 dynamics.

TABLE 4. Evaluation metrics values for Cov-LSTM and Cov-CNN

Country	Models	MAE	MAPE	RMSE	REL
Afghanistan	Cov-LSTM	82.19763	0.0386%	132.88517	0.0001581
	Cov-CNN	817.2139	0.3813%	826.6836248	0.006052709
Iran	Cov-LSTM	808.7599	0.0107%	1020.4604	0.0000073513
	Cov-CNN	51900	0.683%	51800	0.0188
Iraq	Cov-LSTM	2046.176	0.083%	2045.2473	0.00027815
	Cov-CNN	18662.17	0.757%	18639.95175	0.0231

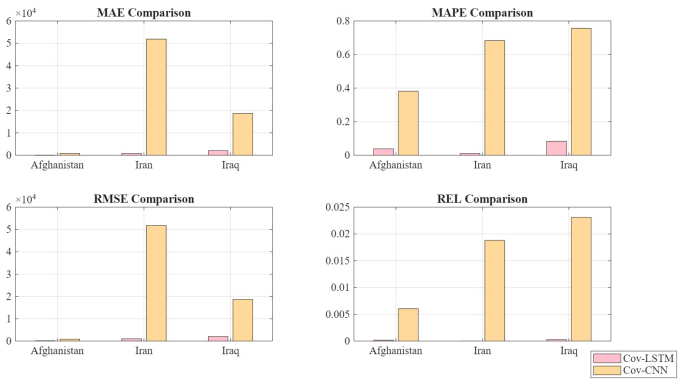

FIGURE 18. Comparative visualization of Cov-LSTM and Cov-CNN model performance across Afghanistan, Iran, and Iraq using MAE, MAPE, RMSE, and REL error metrics

6. CONCLUSIONS AND FUTURE WORK

For decades, epidemiological models have been extensively used for modeling the infectious diseases and their dynamics, thus facilitating the policymakers in the mitigation of infections. In recent years, a significant increase has been seen in the application of data-driven models for forecasting infections. In this study, we have evaluated the effectiveness of data-driven models in forecasting infections in comparison to the conventional epidemiological models. For this purpose, we have used the Covid-19 infections data of Afghanistan, Iran, and Iraq. Based on the comparison of short-term and long-term fitting and forecasting of Covid-19, the following conclusions are drawn:

i. The basic epidemiological models such as SIR and SIRD, although foundational in understanding the disease dynamics, demonstrates poor fitting performance, and thus lacks the ability to forecast infections. The intermediate models SEIR and SEIRD show improved performance in terms of fitting, highlighting the importance of the latent period. However, it fails to forecast the Covid-19 infections.

ii. The proposed SI^PRD model demonstrates improvement over the SIR, SIRD, SEIR and SEIRD models in terms of fitting and forecasting. The SI^PRD model exhibits the ability to forecast the trend of Covid-19 infections for a short time i.e., up to 100 days. However, it lacks accuracy in forecasting for over extended periods.

iii. The data-driven models, i.e., Cov-CNN and Cov-LSTM, exhibit excellent performance in terms of short- and long-term forecasting of infections. Both the data-driven models outperform the epidemiological models in terms of forecasting.

iv. Among the two data-driven models, the Cov-LSTM outperforms the Cov-CNN in accurately forecasting the Covid-19 infections. This is attributed to its ability to retain long-term dependencies and thus capturing the effect of evolving factors in infection dynamics over an extended period.

v. Epidemiological models can play a significant role in initial stages of infections to identify important control parameters and thus help policymakers. The data-driven models depend on a significant amount of infections data and thus are not as useful in the initial stages of the infectious disease due to the unavailability of data.

Overall, the proposed study presents a comprehensive comparison between classical epidemiological models and ANN-based models. Classical epidemiological models are useful for providing early-stage insights into disease dynamics; however, they generally struggle to capture long-term fitting and forecasting behaviour. Although more detailed epidemiological formulations can improve performance by incorporating additional information, they require large and often unavailable datasets. In contrast, ANN-based models demonstrate strong performance in both short-term and long-term prediction tasks, making them more effective for epidemic forecasting. For future work, an agent-based modelling framework will be developed to capture individual-level interactions, including mobility and vaccination effects. Furthermore, sensitivity analysis can be effectively applied to agent-based models to quantify the influence of key parameters on epidemic dynamics and to identify the most critical factors driving disease transmission. In addition, large language models (LLMs) will be integrated to enhance interpretation and support decision-making. A comparative analysis between agent-based models with mobility and vaccination components and LLM-enhanced approaches will be conducted to evaluate their effectiveness in epidemic modelling and forecasting.

CREDIT AUTHORSHIP CONTRIBUTION'S STATEMENT

Shomaila Mazhar: contributed to the conceptualization of the study, mathematical modeling, implementation of data-driven models, data analysis, simulations, interpretation of the results, and manuscript writing.

Zahid Ullah: contributed through research supervision, methodological guidance, critical review of the mathematical framework, revision of the manuscript, and overall supervision of the study.

DECLARATIONS

Conflict of Interest: The authors declared that there is no conflict of interest regarding the publication of this paper.

Funding: The authors received no specific funding for this research.

REFERENCES

[1] I.O. Abdallah, P.M.T. Djomegni and M.S.D.Haggar. Optimal control strategies of cell infections in a Covid-19 model, Alexandria Engineering Journal, **69** (2023): 747-757.

[2] M.O. Alassafi, M. Jarrah and R. Alotaibi. Time series predicting of COVID-19 based on deep learning, Neurocomputing, **468** (2022): 335-344.
[3] M. Allan et al. The World Health Organization COVID-19 surveillance database, International Journal for Equity in Health, **21**, No. Suppl 3 (2022): 167.
[4] K.G. Andersen et al. The proximal origin of SARS-CoV-2, Nature Medicine, **26**, No. 4 (2020): 450-452.
[5] P. Arora, H. Kumar and B.K. Panigrahi, Prediction and analysis of COVID-19 positive cases using deep learning models, Chaos, Solitons and Fractals, **139**, No. 110017 (2020): 1-17.
[6] A. Atkeson, What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios, National Bureau of Economic Research Working Paper No. 26867, (2020). 1-27. DOI 10.3386/w26867
[7] H. Azad, The Jacobi Identity, Journal of Mathematics, **16** (1983): 9–30.
[8] R. Bala and R.P. Singh, Financial and non-stationary time series forecasting using LSTM recurrent neural network for short and long horizon, IEEE ICCCNT, No. 01 (2019): 1-7.
[9] C. Butt, J. Gill, D. Chun and B. A. Babu. Deep learning system to screen coronavirus disease 2019 pneumonia, Applied Intelligence, No. 01 (2020): 1-11.
[10] J.M. Carcione et al., A simulation of a COVID-19 epidemic based on a deterministic SEIR model, Frontiers in Public Health, **8**, No. 230 (2020): 1-13
[11] V.K.R. Chimmula and L. Zhang, Time series forecasting of COVID-19 transmission in Canada using LSTM networks, Chaos, Solitons and Fractals, **135**, No. 10 (2020): 109864.
[12] C. Comito and C. Pizzuti, Artificial intelligence for forecasting and diagnosing COVID-19 pandemic, Artificial Intelligence in Medicine, **128**, No. 102286 (2022): 1-14.
[13] N.M. Cook, The Impact of COVID-19 on Iowa's Corn, Soybean, Ethanol, Pork, and Beef Sectors, Iowa State University, Working Paper No. 20-WP 606 (2020): 1-14.
[14] *COVID-19 Data Explorer* (2024). Our World in Data. https://ourworldindata.org
[15] *COVID-19 cases: WHO COVID-19 dashboard.*(2024). https://data.who.int/dashboards/covid19/cases-.
[16] D. Cucinotta and M. Vanelli, WHO declares COVID-19 a pandemic, Acta Bio Medica Atenei Parmensis, **91**, No. 1 (2020): 157.
[17] C.C. da Silva et al.Covid-19 dynamic monitoring and real-time spatio-temporal forecasting, Frontiers in Public Health, **9**, 641253 (2021): 1–14.
[18] D.K. Das, A. Khatua, T.K.Kar and S. Jana. The effectiveness of contact tracing in mitigating COVID-19 outbreak, Applied Mathematics and Computation, **404**, No. 126207 (2021): 1-12.
[19] M. Diqi, S.H. Mulyani and R. Pradila. DeepCov: effective prediction model of COVID-19 using CNN algorithm, SN Computer Science, **4**, No. 4 (2023): 396.
[20] A.A. Elsadany et al. Spatio-temporal dynamics of COVID-19 epidemic, Electronic Journal of Mathematical Analysis and Applications, **11**, No. 2 (2023): 1-31.
[21] M. Farman, D. Amilo, M.Ghannam, K.S. Nisar and M. Hafeez. Stability and optimizing treatment control of tuberculosis model via numerical approach, Results in Control and Optimization, **18** No. 100650 (2025): 1–16.
[22] N. Fernandes. Economic effects of coronavirus outbreak (COVID-19) on the world economy. IESE Business School, SSRN, No. 3557504 (2020): 1–33
[23] A. Gorbalenya et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nature Microbiology, **5**, No. 4 (2020): 536-544.
[24] D. Guanghong, L. Chang, G. Jianqiu, W. Ling, C. Ke and Z. Di. SARS epidemical forecast research in mathematical model, Chinese Science Bulletin, **49**, No. 21 (2004): 2332-2338.
[25] S. Hochreiter and J. Schmidhuber. Long short-term memory, Neural Computation, **9**, No. 8 (1997): 1735-1780.
[26] M.N. Iksan, Y. Hidayat and A. Bachrudin. Prediksi jumlah kasus aktif COVID-19 di Kota Bandung menggunakan LSTM, E-Prosiding Seminar Nasional Statistika, No. 01, (2020): 1-8.
[27] S. Jaipuria, R. Parida and P. Ray. The impact of COVID-19 on tourism sector in India, Tourism Recreation Research, **46** No. 2 (2021): 245-260.
[28] M.B. Jamshidi et al. A review of AI approaches to forecasting COVID-19, AI, **3** No. 2 (2022): 493-511.
[29] W.O. Kermack and A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London Series A, **115** No. 772 (1927): 700-721.
[30] H.H.P. Kluge, Z. Jakab, J. Bartovic, V. D'Anna and S.Severoni. Refugee and migrant health in the COVID-19 response, The Lancet, **395** No. 10232 (2020): 1237-1239.

[31] S. Kumar, R. Sharma, T. Tsunoda, T. Kumarevel, A. Sharma. Forecasting the spread of COVID-19 using LSTM network, BMC Bioinformatics, **22** No. 6 (2021): 1-9.
[32] The Lancet. India under COVID-19 lockdown, The Lancet, **395**, No. 10233 (2020): 1315.
[33] A. Liem, C. Wang, Y. Wariyanti, C. A. Latkin and B. J. Hall. The neglected health of international migrant workers in the COVID-19 epidemic, The Lancet Psychiatry, **7**, No. 4 (2020).
[34] M. Maliszewska, A. Mattoo and D. Van Der Mensbrugghe, The potential impact of COVID-19 on GDP and trade: A preliminary assessment, World Bank Policy Research Working Paper,No. 9211 (2020): 1–15.
[35] M. Martcheva, An introduction to mathematical epidemiology, Springer, **61** (2015): 1–392.
[36] S. Mazhar, Z. Ullah, S. I. A. Shah and N. Badshah. Development of a probabilistic model for Covid-19 dynamics with consideration of non-severe and severe infections, Alexandria Engineering Journal, **82** No. 01 (2023): 126-138.
[37] A. Mehmood et al. Physics-informed neural network framework for fractional order modeling of Alzheimer's disease, Frontiers in Neuroinformatics, **20** No. 1748481 (2026): 1–14.
[38] V. Monteil et al. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2, Cell, **181** No. 4 (2020): 905-913.
[39] N. Muthukrishnan et al., Brief history of artificial intelligence, Neuroimaging Clinics, **30** No. 04 (2020): 393-399.
[40] K.N. Nabi, M.T. Tahmid, Abdur Rafi, M. E. Kader, M. A. Haider. Forecasting COVID-19 cases: a comparative analysis between recurrent and convolutional neural networks, Results in Physics, **24** No. 104137 (2021): 1–12.
[41] M.A. Noor et al. Stochastic COVID-19 pandemic model: computational analysis, Alexandria Engineering Journal, **61** No. 1 (2022): 619-630.
[42] M.M. Ojo and E.F.D. Goufo, Impact of COVID-19 on a malaria dominated region, Alexandria Engineering Journal, **65** No. 01 (2023): 23-39.
[43] L. Di Persio and O. Honchar, Artificial neural networks architectures for stock price prediction, International Journal of Circuits, Systems and Signal Processing, **10** No. 04 (2016): 403–413.
[44] M. Rafiq, , J.E. Macías-Díaz, A. Raza and N. Ahmed. Design of a nonlinear model for the propagation of COVID-19 and its efficient nonstandard computational implementation, Applied Mathematical Modelling, **89** No. 11 (2021): 1835–1846.
[45] M.H.D.M. Ribeiro et al., Short-term forecasting COVID-19 cumulative confirmed cases, Chaos, Solitons and Fractals, **135** No. 109853 (2020): 1–10.
[46] S. Sah et al. Forecasting COVID-19 pandemic using Prophet, ARIMA, and hybrid stacked LSTM-GRU models in India, Computational and Mathematical Methods in Medicine, No. 01 (2022): 1–14.
[47] F. Shahid, A. Zameer, and M. Muneeb. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM, Chaos, Solitons and Fractals, **140** No. 110212 (2020): 1–12.
[48] S. Shastri, K. Singh, S. Kumar, P. Kour, V. Mansotra. Time series forecasting of Covid-19 using deep learning models, Chaos, Solitons and Fractals, **140** No. 10 (2020): 110227
[49] Y. Shi. Stochastic dynamic model of SARS spreading, Chinese Science Bulletin, **48** No. 12 (2003): 1287-1292.
[50] S. Shujahuddin, M. S. Khan and H. Ali. Mathematical Model for Single and Multiple Object Extraction, Punjab Univ. J. Math., **53** No. 6 (2021): 387.
[51] R. Siche. What is the impact of COVID-19 disease on agriculture?, Scientia Agropecuaria, **11** No. 1 (2020): 3-6.
[52] M. Sinan et al. Analysis of mathematical model of cutaneous leishmaniasis disease, Alexandria Engineering Journal, **72** No. 02 (2023): 117-134.
[53] A. Singh and P. Deolia. COVID-19 outbreak: a predictive mathematical study incorporating shedding effect, Journal of Applied Mathematics and Computing, **69**No. 1 (2023): 1239-1268.
[54] B.A. Sunjaya, S.D. Permai and A.A.S. Gunawan. Forecasting of Covid-19 positive cases in Indonesia using LSTM, Procedia Computer Science, **216** No. 01 (2023): 177-185.
[55] R. Tariq, S. A. Zanib, M. Farman and M. Ghannam. On sensitivity analysis of a nonlinear mathematical approach of alcohol consumption, Mathematical Methods in the Applied Sciences, No. 1 (2026): 1–18.
[56] E.D. Udayanti, E. Kartikadharma, F. Firdausillah and N. Ikhsan. COVID-19 suspects monitoring system based on symptom recognition using deep neural network, International Journal of Engineering and Computer Science Applications, **2** No. 1 (2023): 1-8.

[57] M. Vo, Z. Feng, J. W. Glasser, K. E. N. Clarke and J. N. Jones. Analysis of metapopulation models of SARS-CoV-2, Journal of Mathematical Biology, **87** No. 2 (2023): 24.
[58] J. Wang et al. Spatial dynamics of an epidemic of SARS in an urban area, Bulletin of the World Health Organization, **84** No. 12 (2006): 965-968.
[59] S.H. Wang, X. Zhang and Y.D. Zhang. DSSAE: Deep stacked sparse autoencoder analytical model for COVID-19 diagnosis, ACM Transactions on Management Information Systems, **13** No. 1 (2021): 1-20.
[60] A. Xavier Jr., A C++ code for predicting COVID-19 cases by least-squares fitting of the Logistic model, ResearchGate Preprint, (2020) 1–10.
[61] C. Xu, M. Farman, A. Hasan, A. Akgül, M. Zakarya, W. Albalawi and C. Park. Lyapunov stability and wave analysis of Covid-19 omicron variant, Alexandria Engineering Journal, **61** No. 12 (2022): 11787-11802.
[62] R. Yan et al., Multi-hour and multi-site air quality index forecasting using CNN and LSTM, Expert Systems with Applications, **169** No. 114513 (2021): 1–12.
[63] S. Zeb, M. Bilal, M. Rafiq, S. A. M. Yatim, A. Ahmad, M. Irfan and M. S. Ehsan, Mathematical Modeling for Transmission Dynamics of Hepatitis B Virus, Punjab Univ. J. Math., **57** No. 9 (2025): 964–991.