Forecast COVID-19 Epidemics by Strengthening Deep Learning Models with Time Series Analysis

Warapree Tangseefa[1], Tepanata Pumpaibool[2], Paisit Khanarsa[3], Krung Sinapiromsaran[4]*
College of Public Health Sciences, Chulalongkorn University, Bangkok, Thailand, 10330[1,2]
Institute of Field Robotics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand, 10140[3]
Department of Mathematics and Computer Science, Chulalongkorn University, Bangkok, Thailand, 10330[4]

***Abstract*—The COVID-19 pandemic has profoundly impacted economic and social structures, directly affecting individuals' lives. Deep learning models offer the potential to forecast future long-term trends and capture the temporal dependencies present in time series data. In this study, we propose leveraging the autocorrelation function (ACF) and the partial autocorrelation function (PACF) series as additional components to enhance the forecasting accuracy of our models. Our proposed method is applied to forecast COVID-19 time series data in twelve countries using the deep learning techniques of Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). When comparing the rankings average of mean absolute error and R-squared, the proposed models demonstrated superior performance in time series forecasting compared to the standard LSTM and GRU model. Specifically, the ACF-PACF-GRU model achieved the best median values for mean absolute percentage error (1.67 per cent for confirmed cases and 2.17 per cent for death cases) and root mean square error (1.92 for confirmed cases and 2.17 for death cases). Therefore, the proposed ACF-PACF-GRU model showed the highest performance in forecasting both confirmed and death cases. This research introduces a novel method for constructing effective time series models aimed at forecasting disease burdens, thereby aiding in epidemic control and the implementation of preventive measures.**

***Keywords*—Forecasting models; COVID-19; long short-term memory; gated recurrent unit; autocorrelation function; partial autocorrelation function**

I. Introduction

In late 2019, a novel β-coronavirus—SARS-CoV-2—emerged in Wuhan, China, posing significant challenges to global health systems [1], [2], [3]. Declared a public health emergency by the WHO in March 2020 [4], COVID-19 has since resulted in over 776.8 million confirmed cases and more than 7 million deaths across 234 countries as of November 10, 2024. Although mortality peaked between 2020 and 2022, global death rates have declined with rising immunity levels [5]. The pandemic's impact spans healthcare, economics, travel, and international policy, highlighting the critical need for accurate forecasting to guide public health interventions and resource allocation [6].

COVID-19 demonstrates dynamic, non-linear transmission patterns influenced by seasonal variation and evolving viral behavior, necessitating adaptive modeling approaches [7]. Time series forecasting has traditionally relied on statistical models like ARIMA and SARIMA. ARIMA has been widely applied to model diseases such as tuberculosis [8], brucellosis [9], and dengue [10]. However, it struggles with capturing seasonality and nonlinear patterns, which SARIMA partially addresses [11], [12], [13], [14], [15]. For COVID-19, ARIMA and SARIMA have been used in diverse contexts, such as Morocco [16], Europe [17], and Canada [18], with varying degrees of accuracy. Nevertheless, their linear assumptions and need for differencing limit performance in more complex, evolving datasets [19].

To overcome these limitations, deep learning architectures have emerged as powerful tools for time series forecasting. Advances in data availability and computational power have facilitated their widespread adoption across domains, with deep neural networks proving particularly effective for regression tasks and temporal pattern modeling [20], [21].

The proliferation of big data, driven by advancements in data acquisition technologies, has created a demand for advanced forecasting models capable of processing complex, multivariate, and high-dimensional time series data [22]. Traditional shallow learning approaches often fail to capture the intricate temporal dependencies and noise inherent in such datasets. In response, deep learning has emerged as a dominant paradigm for time series forecasting due to its ability to model non-linear relationships and long-term dependencies effectively [23].

Among deep learning architectures, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are particularly suited for sequential data. Deep learning models have gained significant attention in COVID-19 forecasting due to their capacity to capture non-linear patterns and temporal dependencies in time series data. Among these, LSTM and GRU networks have emerged as leading architectures for epidemic modeling. Both are types of recurrent neural networks (RNNs) specifically designed to mitigate the vanishing gradient problem, enabling the modeling of long-range temporal dependencies crucial for accurate disease forecasting.

Several studies have demonstrated the effectiveness of LSTM models in predicting COVID-19 cases. Dharani and Bojja [24] proposed an enhanced LSTM architecture with internal gate optimizations to improve the prediction accuracy of COVID-19 time series. Similarly, Shahid et al. [25] compared multiple models—including ARIMA, SVR, LSTM, and Bi-LSTM—for forecasting confirmed, death, and recovered cases across ten countries, finding that Bi-LSTM significantly outperformed other models. Luo et al. [26] further validated the predictive strength of LSTM through a comparative study with XGBoost, highlighting LSTM's superior performance in

*Corresponding authors.

capturing temporal dynamics.

GRU models, which simplify LSTM by using fewer parameters while maintaining comparable performance, have also been applied successfully. ArunKumar et al. [27] utilized GRU, LSTM, and other RNN variants to forecast cumulative confirmed cases in the ten most affected countries, reporting promising results. The comparative advantage of GRU lies in its computational efficiency and similar accuracy, making it suitable for real-time applications in epidemic forecasting.

Moreover, hybrid and ensemble variants such as stacked LSTM, bidirectional GRU, and LSTM-GRU combinations have shown improved performance by leveraging the strengths of both architectures [28], [29], [30]. These models are particularly effective in multi-country forecasting settings where data complexity and temporal variability are high.

Overall, the literature suggests that LSTM and GRU models are well-suited for modeling the complex, nonlinear, and dynamic nature of COVID-19 transmission. Their adaptability, accuracy, and scalability make them powerful tools for public health planning and early warning systems. Given the complex, non-linear nature of time series data, statistical tools like ACF and PACF are often used to analyze temporal dependencies. Their integration into deep learning models, as proposed in this study, is a novel approach that aims to enhance prediction accuracy without altering model architectures [31].

This study integrates ACF and PACF as supplementary inputs to LSTM and GRU models to enhance time series forecasting accuracy. By leveraging these statistical tools to capture temporal dependencies, the proposed approach forecasts COVID-19 confirmed and death cases across twelve countries, and compares results against standard LSTM and GRU models reported in the previous study of Luyu Zhou et al [32]. Unlike prior work, this method augments input features without altering the core architectures. Model performance is evaluated using MAE, MAPE, RMSE, and R-squared.

The study is structured as follows: Section II details LSTM and GRU architectures; Section III outlines the data sources, methodology, preprocessing, and evaluation metrics; Section IV presents results and comparisons; Section V discusses the study results with limitations. Finally, Section VI summarizes the conclusion of the whole study.

II. Methodologies

This study illustrates the effective integration of deep learning architectures with statistical techniques to enhance the accuracy of COVID-19 epidemiological data forecasting. In particular, two widely recognized deep learning approaches, LSTM and GRU, are employed in conjunction with ACF and PACF to capture and model the temporal dependencies inherent within the time series data.

A. Long Short-Term Memory (LSTM)

Deep learning techniques like RNN are well-suited for sequence processing because they capture the temporal dynamics of time series data [33]. However, when dealing with very long time series, passing information from earlier time steps to later ones becomes challenging, leading to difficulties in training due to vanishing or exploding gradients. Hochreiter and Schmidhuber introduced LSTM networks to address this issue by facilitating the flow of previous information across state sequences [34]. The LSTM cell architecture includes a memory cell and three gates: the forget gate (F_t), the input gate (I_t), and the output gate (O_t) shown as Fig. 1.

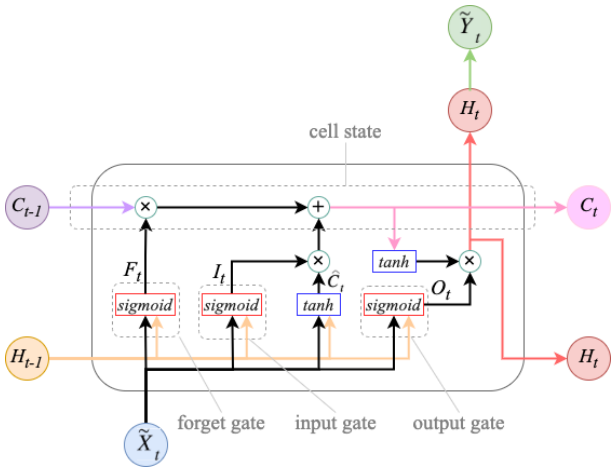

Fig. 1. Architecture of a single LSTM cell.

These gates regulate the flow of information within the cell state: the forget gate controls what information to discard, the input gate determines what information to input to the cell state, and the output gate manages outgoing information [35]. The intermediate cell state ($\hat{C}_t$) serves as the core memory component, propagating information across time steps. Its update at each step is governed by gating mechanisms that selectively regulate the preservation, forgetting, or modification of information. This design enables LSTMs to retain information over long periods, a capability not fully realized in conventional single RNNs. LSTMs excel in capturing long-term dependencies and effectively handling time series data. LSTM network is compute mapping between input sequence and output sequence, i.e. $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_n)$ and $\tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2, ..., \tilde{Y}_n)$. Given the normalized scale of the time series as input $\tilde{X}_t$, as output $\tilde{Y}_t$, and the number of hidden units as h. The gates have the following equations [Eq. (1) to Eq. (7)]:

Input gate:

$$I_t = \text{sigmoid}(W_{xi}\tilde{X}_t + W_{hi}H_{t-1} + b_i) \tag{1}$$

Forget gate:

$$F_t = \text{sigmoid}(W_{xf}\tilde{X}_t + W_{hf}H_{t-1} + b_f) \tag{2}$$

Output gate:

$$O_t = \text{sigmoid}(W_{xo}\tilde{X}_t + W_{ho}H_{t-1} + b_o) \tag{3}$$

Intermediate cell state:

$$\hat{C}_t = \tanh(W_{xc}\tilde{X}_t + W_{hc}H_{t-1} + b_c) \tag{4}$$

Cell state (next memory input):

$$C_t = F_t \odot C_{t-1} + I_t \odot \hat{C}_t \tag{5}$$

New state:

$$H_t = O_t \odot \tanh(C_t) \tag{6}$$

Normalized predicted output:

$$\tilde{Y}_t = H_t \tag{7}$$

where, W_{xi}, W_{xf}, W_{xo}, W_{xc}, and W_{hi}, W_{hf}, W_{ho}, W_{hc} refer respectively to the weight parameters and b_i, b_f, b_o, b_c denote bias parameters for three gates and a memory cell. The symbol $\odot$ indicates the element-wise multiplication. The estimation of C_t relies on the output information from memory cells (C_{t-1}) and the current time step $\hat{C}_t$. The predicted output $\tilde{Y}_t$ is obtained from the hidden state (H_t) at each time t in the current LSTM cell.

B. Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) serves as an alternative version of LSTM designed to enhance the LSTM performance, reduce the number of parameters, and simplify its design [36]. GRU achieves this by amalgamating the input and forget gates of LSTM into a single update gate (Z_t), and utilizing the output gate as a reset gate (R_t). The introduction of reset and update gates represents novel features introduced by GRU models, offering a fresh evaluation method for calculating hidden states in RNN models. In the general architecture of GRU, only the update and reset gates are employed (see Fig. 2).

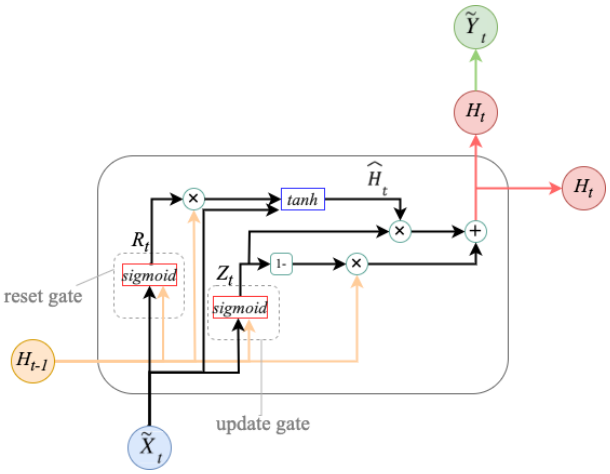

Fig. 2. Architecture of a single GRU cell.

The update gate determines the extent of previously retained memory, while the reset gate governs the manner in which current inputs combine with past memory. The GRU network compute mapping between input sequence and the output sequence in normalized form, that is, $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_n)$ and $\tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2, ..., \tilde{Y}_n)$. The mathematical relationships between the various GRU components are presented in Eq. (8) to Eq. (12):

Update gate:

$$Z_t = \text{sigmoid}(W_{xz}\tilde{X}_t + W_{hz}H_{t-1} + b_z) \tag{8}$$

Reset gate:

$$R_t = \text{sigmoid}(W_{xr}\tilde{X}_t + W_{hr}H_{t-1} + b_r) \tag{9}$$

Cell state:

$$\hat{H}_t = \tanh(W_{xh}\tilde{X}_t + W_{hh}(R_t H_{t-1}) + b_h) \tag{10}$$

New state:

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \hat{H}_t \tag{11}$$

Normalized predicted output:

$$\tilde{Y}_t = H_t \tag{12}$$

where, W_{xz}, W_{xr}, W_{xh} and W_{hz}, W_{hr}, W_{hh} refer respectively to weight parameters and b_z, b_r, b_h denote bias parameters. The symbol $\odot$ indicates the element-wise multiplication. For a given time step t, the current update gate Z_t is used to combine the previous hidden state H_{t-1} and current candidate hidden state $\hat{H}_t$. The normalized predicted output $\tilde{Y}_t$ is derived from the hidden state (H_t) at each time step t within the current GRU cell.

C. Autocorrelation Function and Partial Autocorrelation Function

Time series data often harbors hidden characteristics, encompassing both linear and non-linear components that profoundly impact analysis, modeling, and forecasting. In this context, the autocorrelation function (ACF) and partial autocorrelation function (PACF) emerge as indispensable tools in time series analysis, offering valuable insights into its structure and dynamics. The ACF and PACF serve to uncover the presence and characteristics of autocorrelation within the data. Autocorrelation and partial autocorrelation represent statistical measures that assess the relationship between a time series and its preceding values, known as lag. These measures illuminate the temporal dependencies and patterns inherent in the time series, thereby facilitating a deeper understanding of its behavior and informing predictive modeling endeavors.

The ACF measures the correlation between observations in a time series and their lagged values across different time intervals. This helps to identify patterns and temporal dependencies within the time series data. The ACF calculates the correlation between the series (X_t) and the same series lagged by k time periods (X_{t-k}). Mathematically, the ACF at lag-k is expressed in Eq. (13) to Eq. (14) [37]:

$$\text{Corr}_{(X_t, X_{t-k})} = \frac{\text{Cov}(X_t, X_{t-k})}{\sqrt{\text{Var}(X_t) \cdot \text{Var}(X_{t-k})}} \tag{13}$$

$$\text{Corr}_{(X_t, X_{t-k})} = \frac{\sum_{t=k+1}^{n}(X_t - \bar{X}) \cdot (X_{t-k} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2} \quad (14)$$

where, X_t represents the value of a time series at time t, $\bar{X}$ means the mean of the time series, and n denotes the number of observations in the time series. The variable k represents the time gap being examined and is termed the lag, where k ranges from 1 to n. In essence, a lag-k autocorrelation refers to the correlation between values that are k time periods apart.

Unlike autocorrelation, partial autocorrelation isolates the direct correlation between an observation in a time series and its past values, eliminating the influence of other values at intermediate time lags [31]. For instance, the partial autocorrelation for lag-3 only considers the correlation that lag-1 and lag-2 do not account for. In a time series, the partial autocorrelation between X_t and X_{t-k} is defined as the conditional correlation between X_t and X_{t-k}, conditioned on X_{t-1}, ... , X_{t-k+1}, representing the observations between time points t and $t-k$. The PACF at lag-1 is defined to equal the ACF at lag-1. Subsequently, the partial autocorrelation at lag-k is expressed mathematically as follows [37]:

$$\text{Corr}_{(X_t, X_{t-k} | X_{t-1}, \ldots, X_{t-k+1})} = \frac{\text{Cov}(X_t, X_{t-k} | X_{t-1}, \ldots, X_{t-k+1})}{\sqrt{\text{Var}(X_t | X_{t-1}, \ldots, X_{t-k+1}) \cdot \text{Var}(X_{t-k} | X_{t-1}, \ldots, X_{t-k+1})}} \quad (15)$$

In Eq. (15) X_t represents the value of a time series at time t, and n signifies the total number of observations in the time series. The variable k denotes the time gap under consideration and is termed the lag, with k ranging from 1 to n. Essentially, a lag-k autocorrelation indicates the correlation between values that are k time periods apart.

D. Proposed Methods with ACF, PACF Series

Deep learning techniques, such as LSTM and GRU, have been successfully applied in the context of time series forecasting. The performance of these methods relies heavily on having a sufficient amount of data to appropriately fit their parameters [17]. However, the number of samples extracted from a short time series may be insufficient to achieve an optimal model [38]. To prevent overfitting, these methods should be properly regularized. One challenge in time series forecasting is that, even with a long series and ample data, observations from the distant past often provide fewer cues for prediction. Recent observations from an individual series tend to be more useful in forecasting. This could be attributed to the evolving nature of patterns within a series.

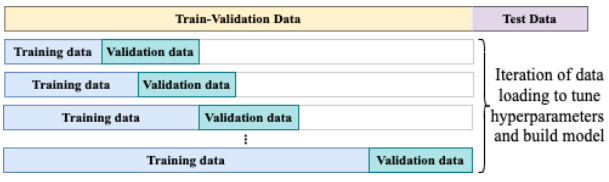

Fig. 3. Train-validation-test split of time series dataset.

In Fig. 3, a given time series is divided into in-sample and out-sample segments using a certain ratio, such as 70/30. The out-sample portion serves as testing data to evaluate the model. The in-sample segment is further divided into training and validation data. However, using separate validation data means excluding recent observations from the training data, potentially missing recent patterns. One simple solution is to include the validation data in model training. However, this may lead to overfitting, reducing accuracy on test data. Specifically, we are determined to apply the statistical properties of the ACF, PACF to avoid model overfitting and enhance the learning potential of deep learning methods for time series forecasting.

The primary concept of the proposed models in this study is to integrate the ACF and PACF series from lag-1 to lag-7, as another new series into deep learning models based on LSTM and GRU for sequential predictions of COVID-19 cases, as depicted in Fig. 4. A part of the training data, and validation data are utilized to build the ACF and PACF series from lag-1 to lag-7, separately in each time period. Both ACF, PACF series uncover the temporal dependencies that exist in the original series. Therefore, the original part of training data and its ACF/PACF series are passed to the standard LSTM or GRU model for generate one proposed model in the model training process, such as ACF-LSTM, PACF-LSTM, ACF-PACF-LSTM, ACF-GRU, PACF-GRU, and ACF-PACF-GRU, as shown in Fig. 5. In the validation process, the original part of training data and its ACF/PACF series are employed with an optimization algorithm to fine-tune the hyperparameters of each model that fit the trained model. Moreover, the testing data from the original time series is used to generate the ACF and PACF series, which are then employed to predict values during the testing phase, as illustrated in Fig. 6.

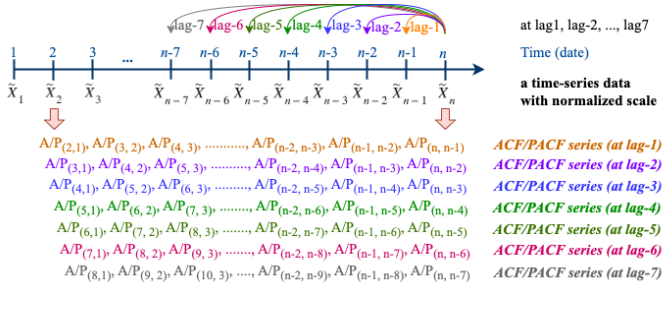

Fig. 4. ACF and PACF series at lag1 to lag7.

III. Materials

This section describes the source of the dataset, design experiments for COVID-19 forecasting, data preprocessing, and the evaluation criteria for model predictive performance.

A. Data Source of COVID-19

This study aims to compare deep learning prediction models using COVID-19 datasets from twelve countries with the highest number of COVID-19 cases. The COVID-19 dataset was sourced from the WHO International website, providing daily cumulative cases of COVID-19 [39], in .csv format. The dataset contains the daily cumulative number of confirmed,

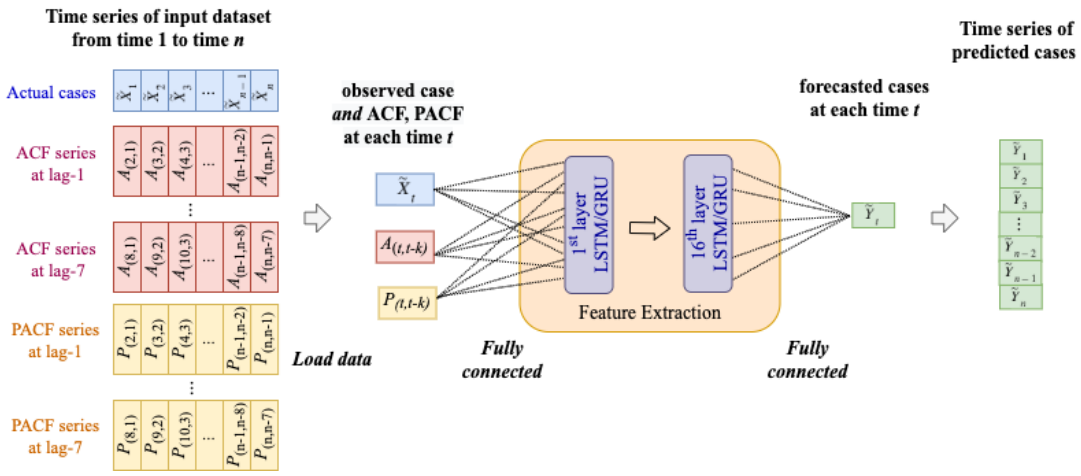

Fig. 5. COVID-19 data, ACF and PACF series into proposed model.

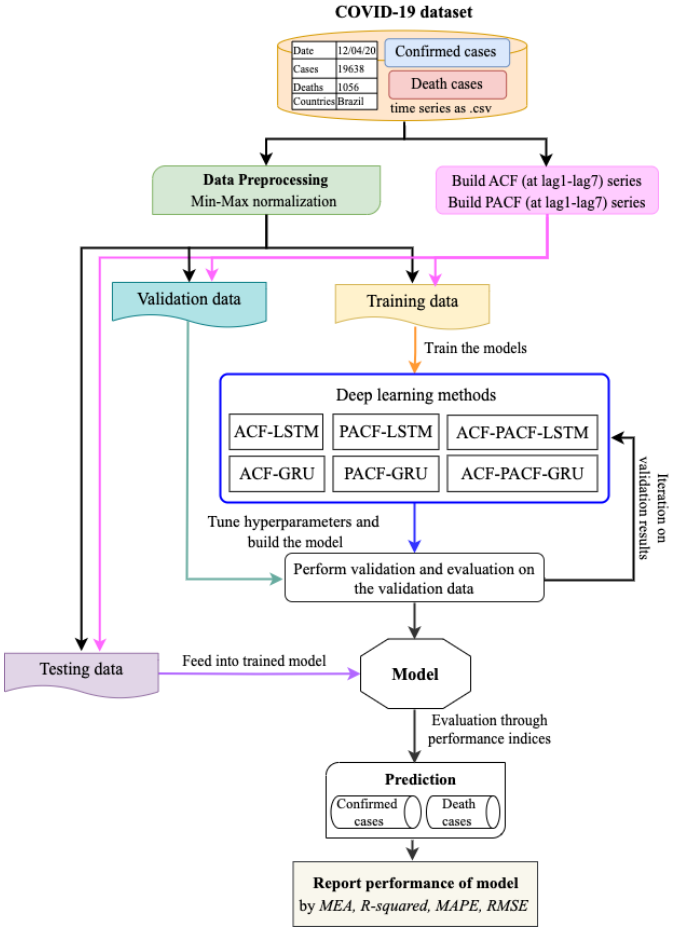

Fig. 6. Framework of the proposed scheme.

death, and recovered cases in the time series format with temporal patterns. In this study, the experiments were conducted using the time series data of confirmed and death cases for twelve countries with the highest number of confirmed cases from January 6, 2020, to June 6, 2022. These countries are the USA, India, Brazil, France, Germany, the UK, Russia, Italy, Turkey, Spain, Vietnam, and Argentina. These specific countries were chosen based on a selection process proposed in a prior publication [32]. This COVID-19 data set includes the counts of confirmed cases, and deaths cases shown in Table I. For training and validation, 70% of each time series is allocated, with the remaining 30% reserved for case prediction in Fig. 3. Before developing a proposed model, the data series undergoes preprocessing.

TABLE I. OVERALL COVID-19 CASES

Countries	Confirmed cases	Death cases
USA	83,894,969	1,000,512
India	43,181,335	524,701
Brazil	31,153,069	666,997
France	28,647,902	149,202
Germany	26,558,190	143,131
UK	22,342,463	199,612
Russia	18,355,200	379,584
Italy	17,505,973	166,949
Turkey	15,077,320	99,004
Spain	12,520,475	108,488
Vietnam	10,726,045	43,081
Argentina	9,276,618	128,973

B. Data Preprocessing

The dataset consists of two features: confirmed and death cases. Unscaled data can impede the convergence process. To enhance model performance and ensure consistency in evaluating deep learning models, preprocessing raw data through normalization is essential. We applied normalization to the data using the formulation $\tilde{X}_t = (X_t – min)/(max – min)$, where $\tilde{X}_t$ represents the normalized actual value at time point t, ensuring it falls within a fixed range typically from zero to one, and X_t denotes the actual value observed at time point t. After constructing the forecasting model at each time point t and obtaining the forecasted values on the normalized scale $\tilde{Y}_t$, we transformed them back to the original data scale using the following formula: $Y_t = \tilde{Y}_t * (max - min) + min$. In this

context, Y_t represents the predicted data at the time point t, rescaled to the raw data format. In Python, the observed raw data can be normalized using the *MinMaxScaler* object from scikit-learn [40]. MinMaxScaler subtracts the smallest feature value and then divides by the feature range, where the range is the difference between the original maximum and original minimum values. This scaler preserves the shape of the original data distribution, retaining the information embedded in the original dataset without significantly altering it or diminishing the importance of outliers.

C. Design Experiments for COVID-19 Forecasting

This study proposes a deep-learning framework for COVID-19 time-series forecasting, integrating the ACF, and PACF series from lag-1 to lag-7 into LSTM and GRU models. Six deep learning models are employed for forecasting daily confirmed and death cases: ACF-LSTM, PACF-LSTM, ACF-PACF-LSTM, ACF-GRU, PACF-GRU, and ACF-PACF-GRU in Fig. 5. The proposed models offer a comprehensive exploration of deep learning techniques. A visual representation of the proposed scheme is depicted in Fig. 6.

Given the progressive spread of COVID-19, the cumulative number of confirmed and death cases is forecasted at specific points in time using preceding values. The COVID-19 time series datasets in this study are segmented into training, validation, and testing datasets, as illustrated in Fig. 6. Monitoring the model's generalization capabilities and selecting appropriate hyperparameters is crucial. To address this, the data is partitioned based on time for training and validation purposes.

COVID-19 forecasting entails two main stages: training and testing. During the initial stage, raw data is preprocessed, standardized, and utilized to construct the deep learning model. Parameters for the deep learning models are adjusted to minimize the loss function during training, employing the Adam optimizer for this purpose. Subsequently, in the testing stage, the previously constructed models with the selected parameters are deployed to forecast the number of COVID-19 cases. The hyperparameter settings for the deep learning models—ACF-LSTM, PACF-LSTM, ACF-PACF-LSTM, ACF-GRU, PACF-GRU, and ACF-PACF-GRU—are outlined in Table II.

TABLE II. PROPOSED SCHEME WITH HYPERPARAMETERS AND THEIR VALUES

Methods	Hyperparameters	Values
ACF-LSTM/	Layers	1
ACF-GRU/	Timestep	7
PACF-LSTM/	Hidden units	16
PACF-GRU	Feature(s)	8
	Loss	MAE
	Last activation	tanh
	Batch size	64
	Epochs	10000
	Optimizer	Adam
	ACF/PACF	lag-1, lag-2, lag-3,..., lag-6, lag-7
ACF-PACF-LSTM/	Layers	1
ACF-PACF-GRU	Timestep	7
	Hidden units	16
	Feature(s)	15
	Loss	MAE
	Last activation	tanh
	Batch size	64
	Epochs	10000
	Optimizer	Adam
	ACF	lag-1, lag-2, lag-3,..., lag-6, lag-7
	PACF	lag-1, lag-2, lag-3,..., lag-6, lag-7

Sixteen cells of LSTM or GRU algorithms are incorporated into a single hidden layer within the proposed models illustrated in Fig. 7. Model accuracy is evaluated by comparing measured data with real data using standard statistical performance measures including MAE, MAPE, RMSE, and R-squared. Simulations and experiments of the proposed models of deep learning are implemented with Keras [41], the Python deep learning library. The primary objective of this study is to assess the predictive capabilities of the six proposed models for forecasting the number of COVID-19 cases, which will be compared with the standard LSTM and GRU models in terms of MAE, and R-squared presented in the previous study. The specific steps involved are as follows:

Step1. The data of COVID-19 are normalized, and the time series is divided into training sets and testing sets;

Step2. Introduce the ACF and, or PACF series (at lag-1, ..., and lag-7) to LSTM, or GRU algorithm;

Step3. Define the hyperparameters in an proposed model in training data;

Step4. In the training step, using the training set data to determine the optimal solution of the model to predict the forecasted values in the training data;

Step5. In the validation step, using validation data to evaluate the fit model from training step;

Step6. In the testing step, using testing data to test the trained model, and evaluate the accuracy of the model by using performance indices. Given the progressive spread of COVID-19, the cumulative number of confirmed and death cases is forecasted at specific points in time using preceding values. The COVID-19 time series datasets in this study are segmented into training, validation, and testing datasets, as illustrated in Fig. 3. Monitoring the model's generalization capabilities and selecting appropriate hyperparameters is crucial. To address this, the data is partitioned based on time for training and validation purposes.

D. Performance Metrics

In this study, the mean absolute error (MAE), and the coefficient of determination (R-squared or R^2) are utilized as predictive performance metrics to evaluate forecast accuracy and compare between proposed models and the standard deep learning models (LSTM, and GRU). Whereas comparing the accuracy forecasting among the six proposed models is considered by the mean absolute percentage error (MAPE), and the root mean square error (RMSE) [42], [43]. These performance metrics are estimated from a normalized form of the actual and predicted values at time point t presented as $\tilde{X}_t$ and $\tilde{Y}_t$,

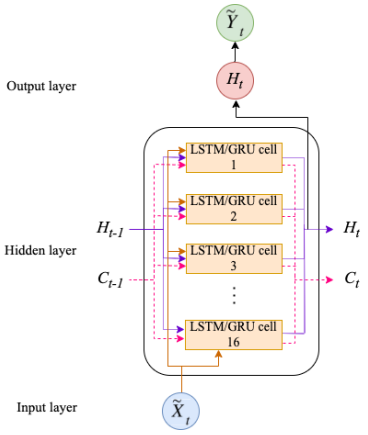

Fig. 7. Architecture of 16 LSTM/GRU cells.

respectively. $\bar{\tilde{X}}$ represents the mean of the normalized actual values, and n denotes the total number of data time points. A model demonstrates superior performance when MAE, MAPE, and RMSE are closer to 0. Additionally, a model is considered more efficient when R-squared approaches 1 [44], [45]. Mean absolute error (MAE) enables the measurement of the average magnitude of errors for a set of predictions, irrespective of their direction. It is expressed mathematically in Eq. (16):

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|\tilde{Y}_t - \tilde{X}_t\right| \tag{16}$$

MAPE quantifies the magnitude of errors as a percentage and can be expressed mathematically in Eq. (17):

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\tilde{X}_t - \tilde{Y}_t}{\tilde{X}_t}\right| \times 100 \tag{17}$$

Root mean square error (RMSE) reveals the discrepancies in amplitude between actual and forecasted data. RMSE is often flavored for evaluating short-term forecasting [46], [47], and can be calculated as Eq. (18):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\tilde{Y}_t - \tilde{X}_t)^2} \tag{18}$$

To demonstrate the proportion of the variation between actual value and forecasted value, R-squared is presented in Eq. (19):

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(\tilde{X}_t - \tilde{Y}_t)^2}{\sum_{t=1}^{n}(\tilde{X}_t - \bar{\tilde{X}})^2} \tag{19}$$

IV. Results

This section describes the evaluation criteria for model predictive performance, results, and discussion about those results. To evaluate the performance of deep learning methods for COVID-19 time-series forecasting, this study compares six proposed models against the standard LSTM and GRU models used as baselines presented as the results from the previous study of Zhou et al. (2023). Our study expands upon the methodologies used by Zhou et al. (2023) by incorporating ACF and PACF as additional input features to the LSTM and GRU models, which has led to substantial improvements in forecasting accuracy. Moreover, the results of this study cover the forecasted trends of cumulative confirmed, and cumulative death of the COVID-19 cases are in using all six proposed models, including ACF-LSTM model, PACF-LSTM model, ACF-PACF-LSTM model, ACF-GRU model, PACF-GRU model, and ACF-PACF-GRU model. The top-12 countries based on cumulative confirmed cases, as of June 6, 2022, are the United States (USA), India, Brazil, France, Germany, United Kingdom (UK), Russia, Italy, Turkey, Spain, Vietnam, and Argentina. The dataset for each country comprises two time series data sets: cumulative confirmed cases, and cumulative death cases. The forecast was done for each period of the two time series data of the individual country and its ACF, and/or PACF with lag-1, lag-2, ..., and lag-7 by independently feeding those time series data into the optimized LSTM and GRU based RNN networks. A total of 144 simulations were performed.

A. Cumulative Confirmed Cases

The study evaluates forecasting accuracy by comparing the proposed models against the standard LSTM and GRU models using the same time series data of COVID-19 confirmed cases. Performance measures for eight models illustrate the actual and predicted cases across twelve countries using deep learning models based on LSTM and GRU. The rankings of MAE and R-squared values for each country, comparing the six proposed models with the standard models, are presented in Table III and Table IV, respectively.

In Table III, the MAE ranking of the standard LSTM model across twelve countries typically falls between 7$^{\text{th}}$ and 8$^{\text{th}}$, with exceptions in Argentina (2$^{\text{nd}}$), Italy (3$^{\text{rd}}$), and France (4$^{\text{th}}$). Zhou et al. (2023) found that LSTM models consistently performed at a moderate level for COVID-19 case predictions. In our study, the ACF-PACF-GRU model consistently outperformed other models, achieving the lowest average MAE ranking (2$^{\text{nd}}$), with superior accuracy in countries like Brazil, France, Turkey, Spain, Vietnam, and Argentina. The standard GRU model ranks between 1$^{\text{st}}$ and 7$^{\text{th}}$. Among the proposed LSTM-based models, the ACF-LSTM model generally ranks between 5$^{\text{th}}$ and 8$^{\text{th}}$, except for the USA, UK, and Russia, which share the 2$^{\text{nd}}$ position. The PACF-LSTM model ranks between 2$^{\text{nd}}$ and 5$^{\text{th}}$, while the ACF-PACF-LSTM model primarily falls between 6$^{\text{th}}$ and 8$^{\text{th}}$, except for Brazil and India, which both rank 5$^{\text{th}}$.

At the country level, the proposed GRU-based models show varying performance. The ACF-GRU model generally ranks between 1$^{\text{st}}$ and 5$^{\text{th}}$, except for Brazil (7$^{\text{th}}$) and Argentina (8$^{\text{th}}$). The PACF-GRU model ranks between 3$^{\text{rd}}$ and 7$^{\text{th}}$, with India as an exception at 1$^{\text{st}}$ place. The ACF-PACF-GRU model, which demonstrates superior performance, primarily ranks between 1$^{\text{st}}$ and 3$^{\text{rd}}$ in MAE, except for Russia and Italy, which both rank 4$^{\text{th}}$. This model achieves the highest forecasting accuracy in six countries: Brazil, France, Turkey, Spain, Vietnam, and Argentina.

TABLE III. ANALYSIS RESULTS OF MAE PERFORMANCE METRICS FOR COVID-19 CONFIRMED CASES ACROSS TWELVE COUNTRIES

Country	LSTM*	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	GRU*	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[8] 0.0211	[2] 0.002325	[4] 0.004666	[7] 0.012734	[5] 0.0073	[1] **0.001303**	[6] 0.007466	[3] 0.004307
India	[8] 0.0057	[6] 0.003872	[3] 0.002278	[5] 0.002675	[7] 0.0046	[4] 0.002605	[1] **0.001567**	[2] 0.001741
Brazil	[8] 0.0089	[6] 0.007829	[2] 0.001885	[5] 0.005264	[4] 0.0039	[7] 0.008784	[3] 0.003358	[1] **0.001028**
France	[4] 0.0168	[6] 0.027764	[3] 0.015254	[8] 0.028122	[2] 0.0131	[5] 0.020826	[7] 0.027992	[1] **0.008854**
Germany	[8] 0.0478	[7] 0.027316	[3] 0.016762	[6] 0.027117	[1] **0.0067**	[4] 0.018798	[5] 0.022447	[2] 0.011984
UK	[8] 0.0085	[2] 0.002272	[4] 0.003317	[8] 0.012526	[5] 0.0067	[1] **0.001142**	[6] 0.007669	[3] 0.002476
Russia	[7] 0.0062	[2] 0.002109	[5] 0.004959	[8] 0.013515	[3] 0.0035	[1] **0.001597**	[6] 0.005153	[4] 0.004294
Italy	[3] 0.0110	[8] 0.029610	[2] 0.009979	[7] 0.022999	[1] **0.0080**	[5] 0.019571	[6] 0.020897	[4] 0.013893
Turkey	[7] 0.0096	[6] 0.008251	[2] 0.004501	[8] 0.013838	[5] 0.0070	[3] 0.005817	[4] 0.006010	[1] **0.002264**
Spain	[8] 0.0153	[7] 0.012144	[2] 0.005996	[6] 0.010588	[4] 0.0081	[5] 0.009843	[3] 0.007193	[1] **0.005131**
Vietnam	[8] 0.0263	[5] 0.016750	[4] 0.010404	[6] 0.018506	[7] 0.0186	[2] 0.007740	[3] 0.010370	[1] **0.005800**
Argentina	[2] 0.0041	[7] 0.012601	[3] 0.004301	[6] 0.010753	[5] 0.0052	[8] 0.014584	[4] 0.004530	[1] **0.003561**
Avg. Ranking	[6.50]	[5.33]	[3.08]	[6.67]	[4.08]	[3.83]	[4.50]	[2.00]
Median	0.0103	0.0102	0.0048	0.0131	0.0069	0.0083	0.0073	0.0043

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models and standard deep learning models (LSTM*, GRU*) cited from the previous study of Luyu Zhou et al.

TABLE IV. ANALYSIS RESULTS OF R-SQUARED PERFORMANCE METRICS FOR COVID-19 CONFIRMED CASES ACROSS TWELVE COUNTRIES

Country	LSTM*	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	GRU*	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[8] 0.79	[2] 0.999514	[5] 0.998138	[6] 0.996355	[7] 0.82	[1] **0.999901**	[4] 0.998232	[3] 0.998652
India	[8] 0.69	[5] 0.998422	[3] 0.998596	[6] 0.998372	[7] 0.81	[2] 0.999254	[1] **0.999595**	[4] 0.998565
Brazil	[8] 0.78	[5] 0.995175	[2] 0.999668	[4] 0.997976	[7] 0.83	[6] 0.995075	[3] 0.999116	[1] **0.999820**
France	[7] 0.83	[6] 0.993040	[2] 0.997088	[5] 0.993198	[8] 0.72	[3] 0.995699	[4] 0.993328	[1] **0.997891**
Germany	[8] 0.72	[5] 0.991069	[2] 0.997044	[6] 0.990451	[7] 0.83	[3] 0.995867	[4] 0.994361	[1] **0.997991**
UK	[7] 0.83	[2] 0.999760	[4] 0.999546	[6] 0.998228	[8] 0.72	[1] **0.999930**	[5] 0.999315	[3] 0.999711
Russia	[7] 0.87	[2] 0.999738	[5] 0.998436	[6] 0.996569	[8] 0.79	[1] **0.999840**	[3] 0.998838	[4] 0.998813
Italy	[7] 0.82	[6] 0.991217	[1] **0.998369**	[4] 0.995043	[8] 0.79	[3] 0.996498	[5] 0.994902	[2] 0.998267
Turkey	[8] 0.72	[5] 0.998472	[4] 0.998901	[6] 0.996464	[7] 0.83	[2] 0.999157	[3] 0.998985	[1] **0.999708**
Spain	[7] 0.83	[6] 0.997367	[4] 0.997940	[3] 0.998465	[8] 0.78	[5] 0.997814	[2] 0.999352	[1] **0.999553**
Vietnam	[7] 0.81	[5] 0.996667	[3] 0.997697	[6] 0.996010	[8] 0.69	[2] 0.998190	[4] 0.997547	[1] **0.998866**
Argentina	[8] 0.79	[5] 0.993972	[3] 0.997978	[4] 0.996731	[7] 0.87	[6] 0.993932	[1] **0.999478**	[2] 0.998541
Avg. Ranking	[7.50]	[4.50]	[3.17]	[5.17]	[7.50]	[2.92]	[3.25]	[2.00]
Median	0.8000	0.9970	0.9983	0.9965	0.8000	0.9980	0.9989	0.9987

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models and standard deep learning models (LSTM*, GRU*) cited from the previous study of Luyu Zhou et al.

TABLE V. ANALYSIS RESULTS OF MAPE PERFORMANCE METRICS FOR COVID-19 CONFIRMED CASES ACROSS TWELVE COUNTRIES

Country	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[2] 0.303071%	[4] 0.617389%	[6] 1.496590%	[1] **0.173846%**	[5] 0.904907%	[3] 0.599810%
India	[6] 0.413458%	[3] 0.252859%	[5] 0.295722%	[4] 0.283855%	[1] **0.171580%**	[2] 0.196526%
Brazil	[5] 0.992938%	[2] 0.232107%	[4] 0.653093%	[6] 1.017317%	[3] 0.362137%	[1] **0.124322%**
France	[5] 4.023079%	[2] 2.542633%	[6] 4.031528%	[3] 3.163861%	[4] 4.014180%	[1] **1.729816%**
Germany	[6] 4.178394%	[2] 2.757850%	[5] 4.066854%	[3] 3.108251%	[4] 3.535206%	[1] **1.733134%**
UK	[2] 0.345353%	[4] 0.520922%	[6] 1.525532%	[1] **0.175605%**	[5] 0.972250%	[3] 0.369609%
Russia	[2] 0.301561%	[5] 0.662813%	[6] 1.546049%	[1] **0.227717%**	[4] 0.658677%	[3] 0.584758%
Italy	[6] 4.174280%	[1] **1.759937%**	[5] 3.201211%	[4] 2.891370%	[3] 2.628040%	[2] 2.015938%
Turkey	[5] 0.880689%	[2] 0.578842%	[6] 1.539030%	[3] 0.642271%	[4] 0.714556%	[1] **0.297870%**
Spain	[6] 1.618311%	[3] 0.980700%	[4] 1.246353%	[5] 1.397544%	[2] 0.851835%	[1] **0.655080%**
Vietnam	[6] 4.506387%	[5] 3.810372%	[3] 3.537982%	[2] 2.466952%	[4] 3.541360%	[1] **1.455438%**
Argentina	[5] 1.621819%	[3] 0.564103%	[4] 1.306167%	[6] 1.752075%	[2] 0.482833%	[1] **0.450001%**
Avg. Ranking	[4.67]	[3.00]	[5.00]	[3.25]	[3.42]	[1.67]
Median	1.306%	0.640%	1.532%	1.207%	0.878%	0.592%

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models.

TABLE VI. ANALYSIS RESULTS OF RMSE PERFORMANCE METRICS FOR COVID-19 CONFIRMED CASES ACROSS TWELVE COUNTRIES

Country	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[2] 0.004424	[4] 0.008354	[6] 0.015735	[1] **0.001864**	[5] 0.010148	[3] 0.007133
India	[6] 0.004335	[3] 0.003575	[5] 0.003894	[2] 0.002783	[1] **0.002112**	[4] 0.003780
Brazil	[5] 0.008499	[2] 0.002280	[4] 0.005525	[6] 0.010168	[3] 0.004581	[1] **0.001665**
France	[4] 0.035989	[2] 0.021375	[6] 0.036731	[3] 0.027494	[5] 0.036429	[1] **0.015350**
Germany	[5] 0.039105	[2] 0.023398	[6] 0.039550	[3] 0.026441	[4] 0.031702	[1] **0.017866**
UK	[2] 0.003656	[4] 0.005250	[6] 0.015484	[1] **0.001978**	[5] 0.009525	[3] 0.004472
Russia	[2] 0.003764	[5] 0.009854	[6] 0.018624	[1] **0.002953**	[4] 0.008916	[3] 0.008283
Italy	[6] 0.037787	[1] **0.014468**	[5] 0.029706	[3] 0.024618	[4] 0.028783	[2] 0.017685
Turkey	[5] 0.011119	[2] 0.007022	[6] 0.017426	[3] 0.007553	[4] 0.008216	[1] **0.003545**
Spain	[6] 0.016057	[3] 0.011521	[5] 0.013918	[4] 0.013410	[2] 0.009497	[1] **0.006811**
Vietnam	[5] 0.023661	[3] 0.018179	[6] 0.028106	[2] 0.016426	[4] 0.018929	[1] **0.013277**
Argentina	[5] 0.015487	[3] 0.008682	[4] 0.012727	[6] 0.017869	[1] **0.006048**	[2] 0.007632
Avg. Ranking	[4.42]	[2.83]	[5.42]	[2.92]	[3.50]	[1.92]
Median	0.013	0.009	0.017	0.012	0.010	0.007

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models.

When comparing LSTM-based and GRU-based models, the proposed GRU-based models exhibit a lower average ranking than their LSTM counterparts. The top three models in terms of average MAE ranking are ACF-PACF-GRU (2.00), PACF-LSTM (3.08), and ACF-GRU (3.83). The MAE performance results indicate that incorporating the PACF series into LSTM models significantly enhances forecasting accuracy compared to the standard LSTM. Similarly, integrating the ACF series into GRU models improves predictive performance over standard GRU. According to the R-squared rankings in Table IV, the standard LSTM and GRU models consistently rank between 7^{th} and 8^{th} across twelve countries. Among the proposed LSTM-based models, the ACF-LSTM model generally ranks between 2^{nd} and 5^{th}, except for Italy, which ranks 1^{t}. The PACF-LSTM model falls between 4^{th} and 5^{th}, with exceptions in the USA, UK, and Russia, where it ranks 2^{nd}. The ACF-PACF-LSTM model typically ranks between 4^{th} and 6^{th}, except for Spain, which is ranked 3^{rd}.

For the GRU-based models at the country level, the ACF-GRU model generally ranks between 1st and 3rd, except in Spain (5^{th}), Brazil (6^{th}), and Argentina (6^{th}). The PACF-GRU model typically ranks between 3^{rd} and 5^{th}, with exceptions in India (1^{st}), Argentina (1^{st}), and Spain (2^{nd}). The ACF-PACF-GRU model, demonstrating superior performance, primarily ranks between 1^{st} and 3^{rd}, except in India and Russia, where it ranks 4^{th}. This model achieves the highest forecasting accuracy in six countries: Brazil, France, Germany, Turkey, Spain, and Vietnam.

Among all models, the proposed GRU-based models exhibit a lower average R-squared ranking than their LSTM counterparts. The top three models in terms of average ranking are ACF-PACF-GRU (2.00), ACF-GRU (2.92), and PACF-LSTM (3.17). The R-squared values indicate that incorporating ACF and PACF series into both LSTM and GRU models significantly enhances forecasting accuracy compared to using only the standard LSTM or GRU models. Therefore, this study demonstrated the MAE and R-squared rankings for ACF-PACF-GRU consistently outperformed other models, showing significant improvement over traditional LSTM and GRU models as used by Zhou et al. in forecasting cumulative confirmed cases across twelve countries.

To summarize the overall performance of the eight models across twelve countries, Fig. 8 presents the median values and boxplots for each model based on different performance metrics. The ACF-PACF-GRU model has the lowest median MAE of 0.004, indicating superior accuracy. In terms of R-squared, the PACF-GRU model achieves the highest median value of 0.9989, followed closely by the ACF-PACF-GRU model at 0.9987. Zhou et al. (2023) highlighted the efficacy of deep learning models like LSTM in predicting COVID-19 cases, though their model did not incorporate ACF/PACF features. Our results align with their conclusion that deep learning models are effective, but we extend their work by showing that incorporating ACF and PACF series as input features significantly improves forecasting accuracy. In particular, the ACF-PACF-GRU model, which combines both ACF and PACF, demonstrated the best performance, significantly lowering the MAE compared to the models used by Zhou et al. (2023), which did not utilize these time-series dependencies.

The MAE and R-squared results indicate that incorporating ACF and/or PACF as additional input series in deep learning models enhances the forecasting accuracy of time series data. To determine which proposed model provides the most accurate COVID-19 case predictions, we analyzed the rankings of MAPE and RMSE across six proposed models for each country, as shown in Tables V and VI.

Overall, the top three models based on average rankings for MAPE and RMSE are ACF-PACF-GRU (with 1.67, and 1.92, respectively), PACF-LSTM (with 3.00, and 2.83), and ACF-GRU (with 3.25, and 2.92). As shown in Table V, the ACF-PACF-GRU model consistently outperforms other models in MAPE across seven countries: Brazil, France, Germany, Turkey, Spain, Vietnam, and Argentina. Similarly, Table VI highlights that this model achieves the lowest RMSE in six countries: Brazil, France, Germany, Turkey, Spain, and Viet-

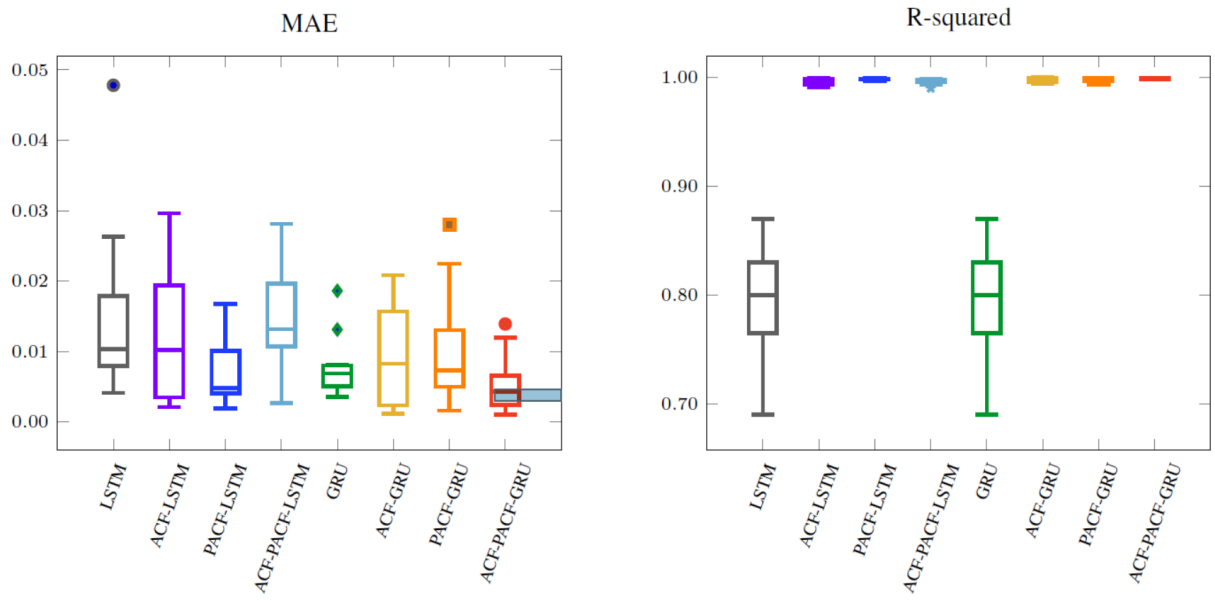

Fig. 8. Analysis results of MAE and R-squared for confirmed cases, compared eight proposed models with the standard deep learning models.

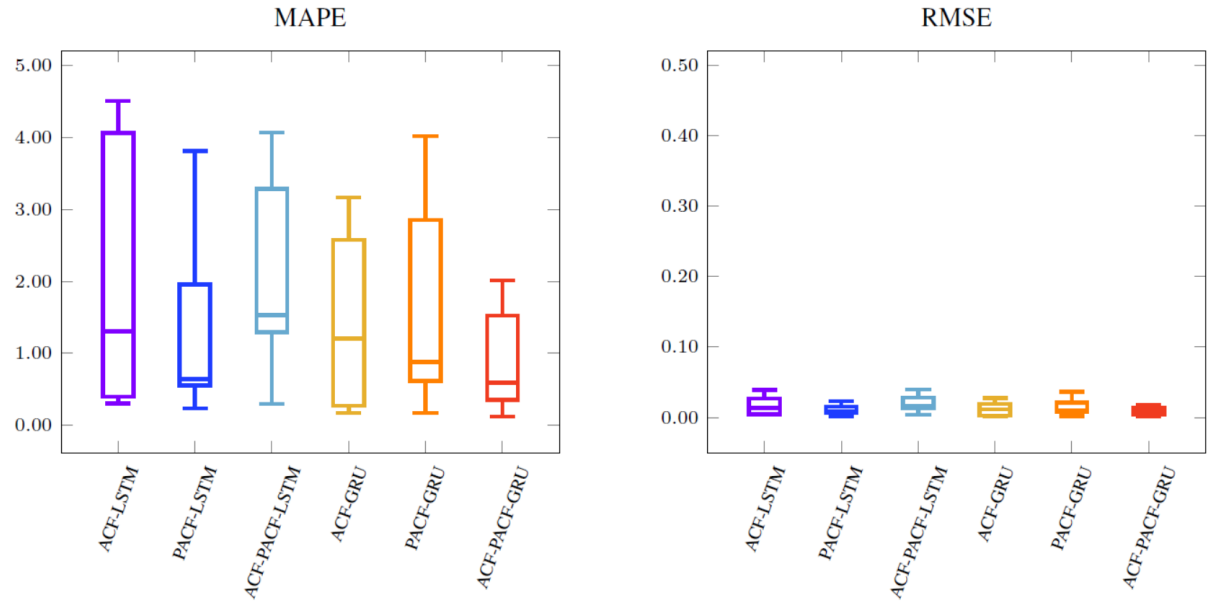

Fig. 9. Analysis results of MAPE and RMSE for confirmed cases compared among six proposed models.

nam. At the country level, the ACF-GRU model ranks within the top three positions more frequently than the PACF-GRU model in both MAPE and RMSE.

These findings suggest that for LSTM-based models, incorporating only the PACF series yields better forecasting performance than adding either the ACF series alone or both series together. Conversely, for GRU-based models, integrating both ACF and PACF series leads to superior forecasting accuracy compared to including only one of the series. Among all models, the ACF-PACF-GRU model demonstrates the highest potential for forecasting COVID-19 confirmed cases.

To provide a comprehensive overview of the results across all twelve countries, Fig. 9 presents the median values and boxplots for the six proposed models, categorized by performance metrics. The median values of MAPE and RMSE in Fig. 9 confirm that the ACF-PACF-GRU model demonstrates the highest predictive accuracy, with the lowest median MAPE at 0.592% and RMSE at 0.007. These findings highlight the effectiveness of the ACF-PACF-GRU model, indicating that the proposed approach significantly improves the forecasting performance for confirmed cases.

B. Cumulative Death Cases

Using the same time-series data for COVID-19 death cases, the forecasting accuracy of the six proposed models is compared to the standard LSTM and GRU models. The

TABLE VII. ANALYSIS RESULTS OF MAE PERFORMANCE METRICS FOR COVID-19 DEATH CASES ACROSS TWELVE COUNTRIES

Country	LSTM*	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	GRU*	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[8] 0.019620	[2] 0.000922	[5] 0.001147	[6] 0.004033	[7] 0.006790	[1] **0.000889**	[4] 0.000991	[3] 0.000978
India	[8] 0.005300	[6] 0.001803	[5] 0.001533	[3] 0.001247	[7] 0.004280	[4] 0.001366	[2] 0.000976	[1] **0.000791**
Brazil	[8] 0.008280	[6] 0.001946	[1] **0.000451**	[4] 0.000625	[7] 0.003630	[5] 0.001575	[2] 0.000471	[3] 0.000509
France	[8] 0.015620	[6] 0.006603	[1] **0.002567**	[4] 0.004213	[7] 0.012180	[5] 0.005486	[3] 0.003727	[2] 0.002902
Germany	[8] 0.044450	[5] 0.005379	[2] 0.002347	[4] 0.004396	[6] 0.006230	[7] 0.007754	[3] 0.004262	[1] **0.002249**
UK	[8] 0.007910	[5] 0.001435	[3] 0.000398	[4] 0.000698	[7] 0.006230	[6] 0.001858	[1] **0.000326**	[2] 0.000330
Russia	[8] 0.005770	[3] 0.000696	[4] 0.000714	[6] 0.002520	[7] 0.003260	[1] **0.000523**	[5] 0.000791	[2] 0.000597
Italy	[8] 0.010230	[6] 0.004871	[1] **0.000281**	[5] 0.004385	[7] 0.007440	[4] 0.003320	[2] 0.001520	[3] 0.001878
Turkey	[8] 0.008930	[5] 0.002190	[3] 0.001275	[6] 0.005586	[7] 0.006510	[4] 0.001922	[2] 0.001088	[1] **0.000755**
Spain	[8] 0.014230	[6] 0.004767	[1] **0.002161**	[4] 0.004239	[7] 0.007530	[5] 0.004738	[3] 0.003460	[2] 0.002798
Vietnam	[7] 0.024460	[3] 0.010312	[1] **0.005568**	[8] 0.027172	[5] 0.017300	[2] 0.008209	[4] 0.014931	[6] 0.021237
Argentina	[7] 0.003810	[6] 0.002313	[4] 0.000842	[3] 0.000650	[8] 0.004840	[5] 0.001997	[2] 0.000505	[1] **0.000457**
Avg. Ranking	[7.83]	[4.92]	[2.58]	[7.45]	[6.83]	[4.08]	[2.75]	[2.25]
Median	0.0096	0.0023	0.0012	0.0041	0.0064	0.0020	0.0010	0.00009

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models and standard deep learning models (LSTM*,GRU*) cited from the previous study of Luyu Zhou et al.

TABLE VIII. ANALYSIS RESULTS OF R-SQUARED PERFORMANCE METRICS FOR COVID-19 DEATH CASES ACROSS TWELVE COUNTRIES

Country	LSTM*	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	GRU*	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[8] 0.7347	[2] 0.999868	[5] 0.999795	[6] 0.998705	[7] 0.7626	[1] **0.999880**	[4] 0.999833	[3] 0.999846
India	[8] 0.6417	[6] 0.997680	[5] 0.998281	[4] 0.998810	[7] 0.7533	[3] 0.998986	[2] 0.999065	[1] **0.999315**
Brazil	[8] 0.7254	[6] 0.996445	[1] **0.999656**	[4] 0.999443	[7] 0.7719	[5] 0.997545	[2] 0.999608	[3] 0.999560
France	[7] 0.7719	[6] 0.992627	[1] **0.998876**	[4] 0.997605	[8] 0.6696	[5] 0.995249	[3] 0.998167	[2] 0.998824
Germany	[8] 0.6696	[3] 0.998863	[2] 0.999333	[4] 0.998483	[7] 0.7719	[6] 0.996506	[5] 0.998353	[1] **0.999481**
UK	[7] 0.7719	[5] 0.999244	[3] 0.999943	[4] 0.999795	[8] 0.6696	[6] 0.998779	[1] **0.999957**	[2] 0.999944
Russia	[7] 0.8091	[5] 0.999951	[1] **0.999978**	[6] 0.999880	[8] 0.7347	[4] 0.999972	[2] 0.999970	[3] 0.999974
Italy	[7] 0.7626	[6] 0.997360	[1] **0.999979**	[5] 0.997804	[8] 0.7347	[4] 0.998670	[2] 0.999731	[3] 0.999559
Turkey	[8] 0.6696	[3] 0.999770	[5] 0.999711	[6] 0.998569	[7] 0.7719	[4] 0.999717	[2] 0.999793	[1] **0.999895**
Spain	[7] 0.7719	[6] 0.995523	[2] 0.998576	[4] 0.997383	[8] 0.7254	[5] 0.996040	[3] 0.997925	[1] **0.998809**
Vietnam	[7] 0.7533	[5] 0.998189	[2] 0.999137	[6] 0.993226	[8] 0.6417	[1] **0.999473**	[3] 0.999031	[4] 0.998211
Argentina	[8] 0.7347	[6] 0.994614	[4] 0.999266	[3] 0.999416	[7] 0.8091	[5] 0.997782	[1] **0.999693**	[2] 0.999584
Avg. Ranking	[7.50]	[4.92]	[2.67]	[4.67]	[7.50]	[4.08]	[2.50]	[2.17]
Median	0.7440	0.9979	0.9995	0.9986	0.7440	0.9987	0.9997	0.9996

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models and standard deep learning models (LSTM*,GRU*) cited from the previous study of Luyu Zhou et al.

performance measures of eight models, shown in Tables VII and VIII, display the actual and forecasted death cases across twelve countries using deep learning models based on LSTM and GRU. The rankings of MAE and R-squared values for each country are presented for the six proposed models and the standard models (LSTM and GRU) in Tables VII and VIII, respectively.

In Table VII, the MAE ranking for the standard LSTM and GRU models among twelve countries generally falls between 7^{th} and 8^{th} place, similar to the results observed in Zhou et al.'s (2023) study. However, our study shows a clear advantage of the proposed GRU-based models over their LSTM counterparts.

For the standard GRU model, the rankings are mostly between 7^{th} and 8^{th}, except for Vietnam (5^{th}) and Germany (6^{th}). For the proposed LSTM-based models, the ACF-LSTM model typically ranks between 5^{th} and 6^{th}, with exceptions in the USA (2^{nd}), Russia (3^{rd}), and Vietnam (3^{rd}). The PACF-LSTM model ranks between 1^{st} and 3^{rd}, except for Argentina (4^{th}), Russia (4^{th}), India (5^{th}), and the USA (5^{th}). The ACF-PACF-LSTM model mostly ranks between 3^{rd} and 6^{th}, with the exception of Vietnam (8^{th}).

For the proposed GRU-based models at the country level, the ACF-GRU model generally ranks between 4^{th} and 7^{th}, except for Russia (1^{st}), the USA (1^{st}), and Vietnam (2^{nd}). The PACF-GRU model ranks between 2^{nd} and 5^{th}, with an exception in the UK (1^{st}). The ACF-PACF-GRU model ranks 1^{st} to 3^{rd} in MAE for most countries, including India, Germany, Turkey, and Argentina. In contrast, the standard LSTM and GRU models performed at lower rankings for these countries. This improvement is particularly evident when considering the MAE and R-squared values for death cases in countries like

TABLE IX. ANALYSIS RESULTS OF MAPE PERFORMANCE METRICS FOR COVID-19 DEATH CASES ACROSS TWELVE COUNTRIES

Country	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[2] 0.108838%	[5] 0.137860%	[6] 0.448506%	[1] **0.104958%**	[4] 0.118071%	[3] 0.116503%
India	[6] 0.194973%	[5] 0.167504%	[3] 0.134351%	[4] 0.147819%	[2] 0.104303%	[1] **0.084415%**
Brazil	[6] 0.206977%	[1] **0.047485%**	[4] 0.066193%	[5] 0.165993%	[2] 0.049643%	[3] 0.053557%
France	[6] 0.761604%	[1] **0.291291%**	[4] 0.454209%	[5] 0.622381%	[3] 0.408169%	[2] 0.315537%
Germany	[5] 0.628394%	[2] 0.296073%	[3] 0.486435%	[6] 0.888655%	[4] 0.512760%	[1] **0.268329%**
UK	[5] 0.158320%	[3] 0.044199%	[4] 0.073488%	[6] 0.211555%	[1] **0.036055%**	[2] 0.037579%
Russia	[3] 0.084860%	[4] 0.092779%	[6] 0.283885%	[1] **0.063388%**	[5] 0.100405%	[2] 0.074391%
Italy	[6] 0.525774%	[1] **0.032344%**	[5] 0.471992%	[4] 0.366857%	[2] 0.168702%	[3] 0.210783%
Turkey	[5] 0.232556%	[3] 0.141059%	[6] 0.587429%	[4] 0.214168%	[2] 0.118658%	[1] **0.085518%**
Spain	[6] 0.559273%	[1] **0.249926%**	[4] 0.460365%	[5] 0.535909%	[3] 0.387958%	[2] 0.308229%
Vietnam	[3] 1.157108%	[1] **0.691349%**	[6] 2.957375%	[2] 0.946599%	[4] 1.699260%	[5] 2.379124%
Argentina	[6] 0.245255%	[4] 0.089956%	[3] 0.068889%	[5] 0.207647%	[2] 0.053661%	[1] **0.048188%**
Avg. Ranking	[4.92]	[2.58]	[4.50]	[4.00]	[2.83]	[2.17]
Median	0.239%	0.139%	0.451%	0.213%	0.118%	0.101%

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models.

TABLE X. ANALYSIS RESULTS OF RMSE PERFORMANCE METRICS FOR COVID-19 DEATH CASES ACROSS TWELVE COUNTRIES

Country	ACF-LSTM	PACF-LSTM	ACF-PACF-LSTM	ACF-GRU	PACF-GRU	ACF-PACF-GRU
USA	[2] 0.001234	[5] 0.001526	[6] 0.004537	[1] **0.001177**	[4] 0.001384	[3] 0.001328
India	[6] 0.002388	[5] 0.002115	[4] 0.001708	[3] 0.001626	[2] 0.001515	[1] **0.001312**
Brazil	[6] 0.002159	[1] **0.000653**	[4] 0.000831	[5] 0.001750	[3] 0.000743	[2] 0.000734
France	[6] 0.007181	[1] **0.002826**	[4] 0.005207	[5] 0.006047	[3] 0.004367	[2] 0.003495
Germany	[5] 0.006297	[2] 0.002971	[4] 0.005446	[6] 0.009024	[3] 0.005061	[1] **0.002624**
UK	[5] 0.001620	[3] 0.000460	[4] 0.000947	[6] 0.002270	[1] **0.000385**	[2] 0.000441
Russia	[4] 0.000958	[3] 0.000890	[6] 0.002853	[1] **0.000733**	[5] 0.001071	[2] 0.000764
Italy	[6] 0.006214	[1] **0.000435**	[5] 0.005660	[4] 0.003911	[2] 0.001755	[3] 0.002111
Turkey	[5] 0.002708	[3] 0.002088	[6] 0.006989	[4] 0.002211	[2] 0.001767	[1] **0.001222**
Spain	[5] 0.005452	[1] **0.003074**	[4] 0.005129	[6] 0.005542	[3] 0.004095	[2] 0.003252
Vietnam	[3] 0.010946	[1] **0.006718**	[6] 0.029411	[2] 0.008528	[4] 0.015452	[5] 0.022089
Argentina	[6] 0.003026	[4] 0.001157	[3] 0.001002	[5] 0.002087	[1] **0.000726**	[2] 0.000909
Avg. Ranking	[4.92]	[2.50]	[4.67]	[4.00]	[2.75]	[2.17]
Median	0.003	0.002	0.005	0.002	0.002	0.001

Note: Bold numbers indicate the best results. Superscript numbers in brackets indicate ranking among six proposed models.

Brazil and India.

When comparing the LSTM-based and GRU-based models, the proposed GRU models generally have a lower average ranking than the LSTM models. Among the eight models, the top three in terms of average MAE ranking are ACF-PACF-GRU (2.25), PACF-LSTM (2.58), and PACF-GRU (2.75). The MAE performance results show that incorporating the PACF series into LSTM models enhances forecasting performance more effectively than using the standard LSTM alone. However, integrating both ACF and PACF series into GRU models improves forecasting accuracy compared to using the standard GRU.

As shown in Table VIII, the R-squared ranking for the standard LSTM and GRU models across twelve countries generally falls between 7^{th} and 8^{th}. For the proposed LSTM-based models, the ACF-LSTM model typically ranks between 5^{th} and 6^{th}, with exceptions in the USA (2^{nd}), Germany (3^{rd}), and Turkey (3^{rd}). The PACF-LSTM model ranks between 1^{st} and 3^{rd}, except for Argentina (4^{th}), the USA (5^{th}), India (2^{nd}), and Turkey (5^{th}). The ACF-PACF-LSTM model usually ranks between 4^{th} and 6^{th}, with the exception of Argentina (3^{rd}).

For the proposed GRU-based models at the country level, the ACF-GRU model generally ranks between 3^{rd} and 6^{th}, except for the USA and Vietnam, where it ranks 1^{st}. The PACF-GRU model typically ranks between 1^{st} and 3^{rd}, except for the USA (4^{th}) and Germany (5^{th}). The ACF-PACF-GRU model, which achieves superior results, ranks between 1^{st} and 3^{rd} in most countries, except for Vietnam (4^{th}). This model demonstrates the highest forecasting accuracy in six countries: India, Germany, Turkey, and Spain.

Among the various models, the proposed GRU-based models tend to have a lower average R-squared ranking compared

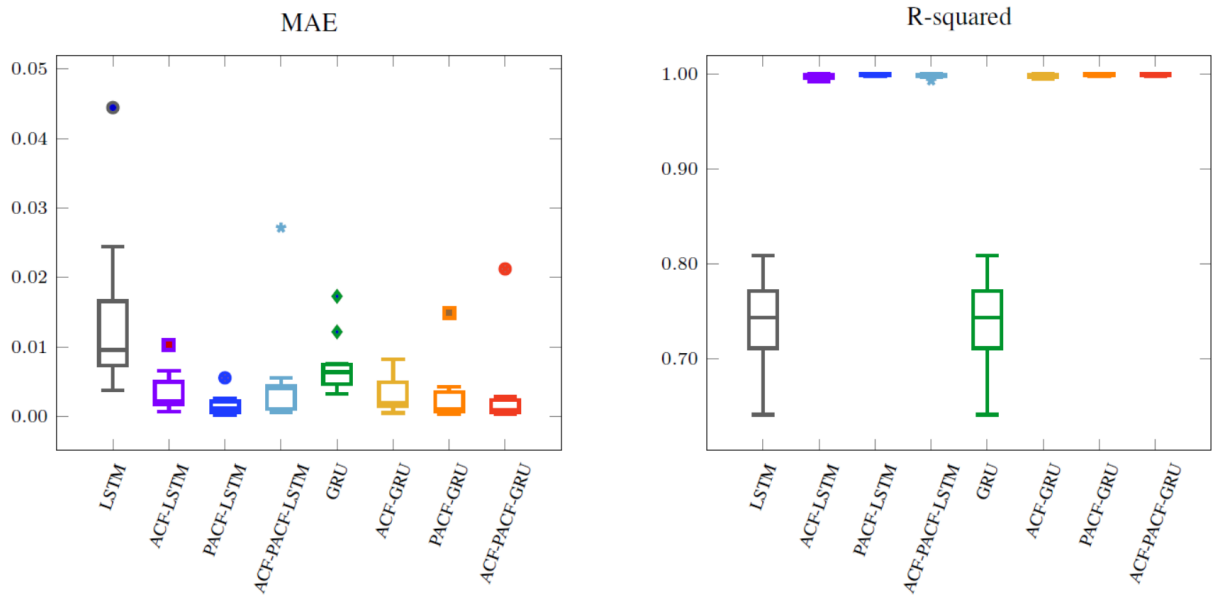

Fig. 10. Analysis results of MAE and R-squared for death cases, compared eight proposed models with the standard deep learning models.

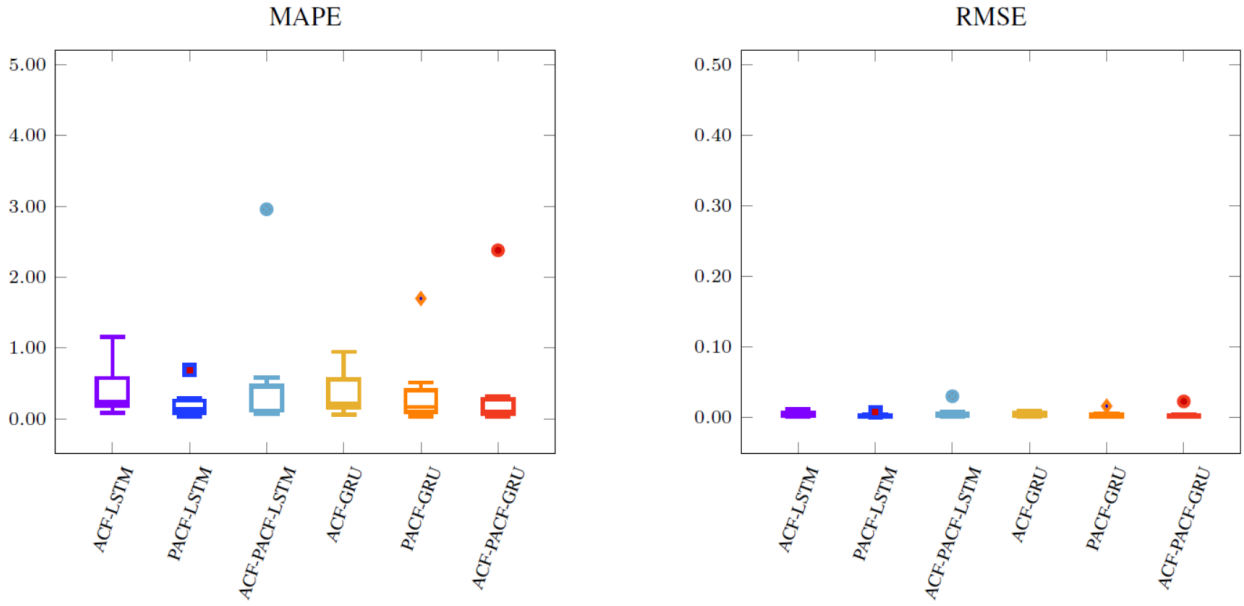

Fig. 11. Analysis results of MAPE and RMSE for death cases compared among six proposed models.

to the LSTM-based models. The top three models based on average R-squared ranking are ACF-PACF-GRU (2.17), PACF-GRU (2.50), and PACF-LSTM (2.67). The R-squared values indicate that incorporating both ACF and PACF series into GRU models can significantly enhance forecasting accuracy compared to models based solely on LSTM or GRU, similar to the improvement observed when adding the PACF series to either LSTM or GRU.

To provide a comprehensive overview of the results for all twelve countries, Fig. 10 presents the median values and boxplots for eight models, organized by performance metrics. The median values in Fig. 10 indicate that the ACF-PACF-GRU model for death cases has the highest predictive ability, with the lowest MAE value of 0.0009 and the highest R-squared value of 0.9996. However, the PACF-GRU model has a slightly higher median R-squared value of 0.9997.

Next, we examine which proposed model demonstrates the greatest forecasting accuracy for COVID-19 death cases. The MAPE and RMSE values for the six proposed models were ranked for each country, as shown in Table IX and Table X. Overall, the top 3 models based on average rankings for MAPE and RMSE are ACF-PACF-GRU (2.17, and 2.17, respectively), PACF-LSTM (2.58, and 2.50), and ACF-GRU (2.83, and 2.75). In Table IX, the ACF-PACF-GRU model, with an average MAPE ranking of 2.17, consistently outperforms the other models in four countries, including India, Germany, Turkey,

and Argentina. Similarly, in Table X, the ACF-PACF-GRU model, with an average RMSE ranking of 2.17, is the best model in six countries, including India, Germany, and Turkey.

To provide a comprehensive performance comparison across all twelve countries, Fig. 11 illustrates the median values and corresponding boxplots for the six proposed models, evaluated using key performance metrics. The ACF-PACF-GRU model consistently outperformed other models, achieving the lowest median MAPE of 0.101% and RMSE of 0.001. These results highlight the superior forecasting capability of the ACF-PACF-GRU model, demonstrating the effectiveness of integrating autocorrelation and partial autocorrelation features in enhancing prediction accuracy for COVID-19 death cases.

Therefore, the forecasting performance of the LSTM model, when only the PACF series is added, is better than when either the ACF series or both series are incorporated into the same algorithm. On the other hand, adding only the ACF series or both series into the GRU model leads to better forecasting accuracy compared to adding only the PACF series. However, the MAPE and RMSE results for death cases confirm that the ACF-PACF-GRU model has the highest potential for accurately forecasting COVID-19 cases.

V. Discussions

This study introduces a novel approach to improve the forecasting accuracy of COVID-19 confirmed and death cases by integrating ACF and PACF series into LSTM and GRU models. The key motivation for choosing this approach is to address the limitations of existing models, which primarily rely on standard LSTM and GRU networks. While LSTM and GRU models are adept at capturing long-term dependencies in time-series data, they fall short in modeling non-linear relationships and temporal dependencies [20], particularly in the context of irregular and complex epidemic data like COVID-19. The incorporation of ACF and PACF enhances the model's ability to capture both short-term and long-term dependencies, providing a more accurate representation of the underlying data dynamics, which is essential for infectious disease forecasting.

This study demonstrates that our proposed models, especially the ACF-PACF-GRU model, outperform standard LSTM and GRU models in forecasting accuracy, as evidenced by the improved performance across key metrics such as MAE, R-squared, MAPE, and RMSE. The incorporation of ACF and PACF features provides additional insights into the temporal dependencies of the data, which were absent in previous studies, including the work by Zhou et al. (2023). This approach not only improves the predictive accuracy for COVID-19 but is also applicable to other epidemic forecasting tasks where similar temporal dependencies exist.

A. Implications of Results

The findings of this study have significant implications for the forecasting of infectious disease trends. The unique structure of LSTM and GRU models, when combined with ACF and PACF series, enhances the ability of the models to effectively capture both linear and non-linear temporal dependencies present in epidemic data. The median MAPE values for the ACF-PACF-GRU model, which are 0.592% for confirmed cases and 0.101% for death cases, underscore the efficacy of the proposed method. These values highlight how the integration of ACF and PACF into deep learning models can significantly improve the forecasting performance over existing models. Furthermore, the improved accuracy achieved by our models supports previous research emphasizing the importance of time-series dependencies in epidemic forecasting. Our proposed methodology, which extends the work of Zhou et al. (2023) by incorporating ACF and PACF features, offers a more robust framework for predicting COVID-19 trends and can be adapted to other infectious diseases in future research.

B. Practical Implications

The proposed deep learning models provide valuable insights for forecasting the cumulative number of confirmed and death cases of COVID-19 by incorporating ACF and PACF series into LSTM and GRU models. By treating the forecasting task as a time-series problem, leveraging historical data to predict future trends, the models can provide more accurate predictions, thereby aiding public health decision-making. The integration of ACF and PACF into GRU models, in particular, improves model performance, making them more suitable for real-time epidemic prediction. The enhanced accuracy of the ACF-PACF-GRU model, when compared to the methods used by Zhou et al. (2023), demonstrates the potential for further improving deep learning-based forecasting models with additional time-series features. These forecasts are crucial for governments and public health officials, enabling them to make informed decisions about epidemic control measures, resource allocation, and intervention strategies.

C. Limitations

While LSTM and GRU models excel at capturing long-term dependencies, they do not inherently account for the temporal dependencies critical to epidemic forecasting. The addition of ACF and PACF significantly enhances the models' ability to capture these temporal patterns, which are often key to understanding disease transmission. However, this study focused solely on time-series data of the disease, excluding other variables such as vaccination rates and country-specific interventions, which could also influence epidemic outcomes. Future research could incorporate these additional features to provide a more comprehensive forecasting model. Additionally, the complex architecture of deep learning models requires careful hyperparameter selection, which may vary depending on the dataset and problem. Thus, the proposed models in this study, while more accurate than existing models, still require optimization for different contexts and datasets.

VI. Conclusions

In this study, we introduced a novel deep learning approach that incorporates ACF and PACF to improve time-series forecasting models for predicting COVID-19 confirmed and death cases. By leveraging the ACF and PACF series in conjunction with original time-series data, our models were able to capture both short-term and long-term temporal dependencies, leading to improved forecasting accuracy. The ACF-PACF-GRU model, in particular, demonstrated superior performance across a range of metrics compared to both standard LSTM and GRU models. These findings suggest that incorporating time-series dependencies, such as ACF and PACF, significantly

enhances the performance of deep learning models in epidemic forecasting. Future research should explore the application of this methodology to other infectious diseases, such as influenza and dengue fever, to further validate the robustness and generalizability of the approach.

ACKNOWLEDGMENT

This research was part of the graduation requirement for the philosopher degree at College of Public Health Sciences, Chulalongkorn University, which was supported by multidisciplinary researchers. The authors extend their heartfelt thanks to the anonymous reviewers for their insightful and constructive feedback. The authors acknowledge the assistance of Gemini in refining the grammar and clarity of this manuscript. However, the scientific content and conclusions are entirely the authors' responsibility.

1) Credit authorship contribution statement:

a) Warapree tangseefa: Conceptualization, Data curation, Methodology, Software, Formal analysis, Validation, Writing – original draft, Writing – review & editing.

b) Tepanata pumpaibool: Conceptualization, Review & editing.

c) Paisit khanarsa: Conceptualization, Software, Formal analysis, Visualization, Writing – review & editing.

d) Krung sinapiromsara: Conceptualization, Methodology, Writing – review & editing, Funding acquisition, Investigation, Supervision.

2) Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

REFERENCES

[1] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, "The species severe acute respiratory syndrome-related coronavirus: classifying 2019-ncov and naming it sars-cov-2," *Nature microbiology*, vol. 5, no. 4, pp. 536–544, 2020.

[2] B. Vellingiri, K. Jayaramayya, M. Iyer, A. Narayanasamy, V. Govindasamy, B. Giridharan, S. Ganesan, A. Venugopal, D. Venkatesan, H. Ganesan *et al.*, "Covid-19: A promising cure for the global panic," *Science of the total environment*, vol. 725, p. 138277, 2020.

[3] T. M. Abd El-Aziz and J. D. Stockand, "Recent progress and challenges in drug development against covid-19 coronavirus (sars-cov-2)-an update on the status," *Infection, Genetics and evolution*, vol. 83, p. 104327, 2020.

[4] D. Cucinotta and M. Vanelli, "Who declares covid-19 a pandemic," *Acta bio medica: Atenei parmensis*, vol. 91, no. 1, p. 157, 2020.

[5] World Health Organization, "Covid-19 epidemiological update, edition 174, 24 december 2024," 2024, Accessed: March. 12, 2025. [Online]. Available: https://iris.who.int/handle/10665/379898

[6] H. Abbasimehr and R. Paki, "Prediction of covid-19 confirmed cases combining deep learning methods and bayesian optimization," *Chaos, Solitons & Fractals*, vol. 142, p. 110511, 2021.

[7] Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, and M. Wang, "Presumed asymptomatic carrier transmission of covid-19," *Jama*, vol. 323, no. 14, pp. 1406–1407, 2020.

[8] Z. Zuo, M. Wang, H. Cui, Y. Wang, J. Wu, J. Qi, K. Pan, D. Sui, P. Liu, and A. Xu, "Spatiotemporal characteristics and the epidemiology of tuberculosis in china from 2004 to 2017 by the nationwide surveillance system," *BMC Public Health*, vol. 20, pp. 1–12, 2020.

[9] V. Rahmanian, S. Bokaie, K. Rahmanian, S. Hosseini, and A. T. Firouzeh, "Analysis of temporal trends of human brucellosis between 2013 and 2018 in yazd province, iran to predict future trends in incidence: A time-series study using arima model," *Asian Pacific Journal of Tropical Medicine*, vol. 13, no. 6, pp. 272–277, 2020.

[10] K. Shashvat, R. Basu, and A. P. Bhondekar, "Application of time series methods for dengue cases in north india (chandigarh)," *Journal of Public Health*, vol. 29, pp. 433–441, 2021.

[11] Y. Chen, K. Leng, Y. Lu, L. Wen, Y. Qi, W. Gao, H. Chen, L. Bai, X. An, B. Sun *et al.*, "Epidemiological features and time-series analysis of influenza incidence in urban and rural areas of shenyang, china, 2010–2018," *Epidemiology & Infection*, vol. 148, p. e29, 2020.

[12] W. Hu, S. Tong, K. Mengersen, and D. Connell, "Weather variability and the incidence of cryptosporidiosis: comparison of time series poisson regression and sarima models," *Annals of epidemiology*, vol. 17, no. 9, pp. 679–688, 2007.

[13] L.-P. Yang, S.-Y. Liang, X.-J. Wang, X.-J. Li, Y.-L. Wu, and W. Ma, "Burden of disease measured by disability-adjusted life years and a disease forecasting time series model of scrub typhus in laiwu, china," *PLoS neglected tropical diseases*, vol. 9, no. 1, p. e3420, 2015. [Online]. Available: https://doi.org/10.1371/journal.pntd.0003420

[14] M. S. Park, K. H. Park, and G. J. Bahk, "Combined influence of multiple climatic factors on the incidence of bacterial foodborne diseases," *Science of the total environment*, vol. 610, pp. 10–16, 2018.

[15] X. Zhang, Y. Liu, M. Yang, T. Zhang, A. A. Young, and X. Li, "Comparative study of four time series methods in forecasting typhoid fever incidence in china," *PloS one*, vol. 8, no. 5, p. e63116, 2013. [Online]. Available: https://doi.org/10.1371/journal.pone.0063116

[16] I. CHOUJA, S. SAOUD, and M. SADIK, "New sarima approach model to forecast covid-19 propagation: case of morocco," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 12, 2021.

[17] K. Bandara, C. Bergmeir, and S. Smyl, "Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach," *Expert systems with applications*, vol. 140, p. 112896, 2020. [Online]. Available: https://doi.org/10.1016/j.eswa.2019.112896

[18] J. Sun, "Forecasting covid-19 pandemic in alberta, canada using modified arima models," *Computer methods and programs in biomedicine update*, vol. 1, p. 100029, 2021. [Online]. Available: https://doi.org/10.1016/j.cmpbup.2021.100029

[19] M. Zhai, W. Li, P. Tie, X. Wang, T. Xie, H. Ren, Z. Zhang, W. Song, D. Quan, M. Li *et al.*, "Research on the predictive effect of a combined model of arima and neural networks on human brucellosis in shanxi province, china: a time series predictive analysis," *BMC Infectious Diseases*, vol. 21, pp. 1–12, 2021.

[20] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200209, 2021. [Online]. Available: https://doi.org/10.1098/rsta.2020.0209

[21] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting methods and applications*. USA: John wiley & sons, 2008.

[22] M. Längkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern recognition letters*, vol. 42, pp. 11–24, 2014.

[23] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[24] N. Dharani and P. Bojja, "Analysis and prediction of covid-19 by using recurrent lstm neural network model in machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, 2022.

[25] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for covid-19 with deep learning models of lstm, gru and bi-lstm," *Chaos, Solitons & Fractals*, vol. 140, p. 110212, 2020. [Online]. Available: https://doi.org/10.1016/j.chaos.2020.110212

[26] J. Luo, Z. Zhang, Y. Fu, and F. Rao, "Time series prediction of covid-19 transmission in america using lstm and xgboost algorithms," *Results in Physics*, vol. 27, p. 104462, 2021. [Online]. Available: https://doi.org/10.1016/j.rinp.2021.104462

[27] K. ArunKumar, D. V. Kalaga, C. M. S. Kumar, G. Chilkoor, M. Kawaji, and T. M. Brenza, "Forecasting the dynamics of cumulative covid-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (arima) and seasonal auto-regressive integrated moving average (sarima)," *Applied soft computing*, vol. 103, p. 107161, 2021. [Online]. Available: https://doi.org/10.1016/j.asoc.2021.107161

[28] H. Verma, S. Mandal, and A. Gupta, "Temporal deep learning architecture for prediction of covid-19 cases in india," *Expert Systems with Applications*, vol. 195, p. 116611, 2022. [Online]. Available: https://doi.org/10.1016/j.eswa.2022.116611

[29] N. Ayoobi, D. Sharifrazi, R. Alizadehsani, A. Shoeibi, J. M. Gorriz, H. Moosaei, A. Khosravi, S. Nahavandi, A. G. Chofreh, F. A. Goni *et al.*, "Time series forecasting of new cases and new deaths rate for covid-19 using deep learning methods," *Results in physics*, vol. 27, p. 104495, 2021. [Online]. Available: https://doi.org/10.1016/j.rinp.2021.104495

[30] S. Mehrmolaei, M. Savargiv, and M. R. Keyvanpour, "Hybrid learning-oriented approaches for predicting covid-19 time series data: A comparative analytical study," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106754, 2023. [Online]. Available: https://doi.org/10.1016/j.engappai.2023.106754

[31] F. Kamalov, F. Thabtah, and I. Gurrib, "Autocorrelation for time series with linear trend," in *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*. IEEE, 2021, pp. 181–185.

[32] L. Zhou, C. Zhao, N. Liu, X. Yao, and Z. Cheng, "Improved lstm-based deep learning model for covid-19 prediction using optimized approach," *Engineering applications of artificial intelligence*, vol. 122, p. 106157, 2023. [Online]. Available: https://doi.org/10.1016/j.engappai.2023.106157

[33] A. Ntakaris, G. Mirone, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Feature engineering for mid-price prediction with deep learning," *Ieee Access*, vol. 7, pp. 82 390–82 412, 2019.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[35] A. Graves, "Generating sequences with recurrent neural networks," *arXiv preprint arXiv:1308.0850*, 2013. [Online]. Available: https://arxiv.org/pdf/1308.0850

[36] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014. [Online]. Available: https://arxiv.org/pdf/1406.1078

[37] J. D. Cryer, *Time series analysis*. Boston: Duxbury Press, 1986.

[38] W. Yan, "Toward automatic time-series forecasting using neural networks," *IEEE transactions on neural networks and learning systems*, vol. 23, no. 7, pp. 1028–1039, 2012.

[39] World Health Organization, "Daily frequency reporting of new covid-19 cases and deaths by date reported to who," Accessed: May. 22, 2024. [Online]. Available: https://srhdpeuwpubsa.blob.core.windows.net/whdh/COVID/WHO-COVID-19-global-daily-data.csv

[40] J. Brownlee, "How to use standardscaler and minmaxscaler transforms in python," 2020, Accessed: March. 12, 2025. [Online]. Available: https://machinelearningmastery.com/standardscaler-and-minmaxscaler-transforms-in-python/

[41] F. Chollet, "Keras," 2015, Accessed: May. 22, 2024. [Online]. Available: https://keras.io // https://github.com/fchollet/keras

[42] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," *International journal of forecasting*, vol. 8, no. 1, pp. 69–80, 1992. [Online]. Available: https://doi.org/10.1016/0169-2070(92)90008-W

[43] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[44] H. Zang, L. Cheng, T. Ding, K. W. Cheung, M. Wang, Z. Wei, and G. Sun, "Application of functional deep belief network for estimating daily global solar radiation: A case study in china," *Energy*, vol. 191, p. 116502, 2020.

[45] S. G. Gouda, Z. Hussein, S. Luo, and Q. Yuan, "Model selection for accurate daily global solar radiation prediction in china," *Journal of cleaner production*, vol. 221, pp. 132–144, 2019.

[46] J. Fan, L. Wu, F. Zhang, H. Cai, X. Ma, and H. Bai, "Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of china," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 168–186, 2019.

[47] K. Bakirci, "Correlations for estimation of daily global solar radiation with hours of bright sunshine in turkey," *Energy*, vol. 34, no. 4, pp. 485–501, 2009.