A HYBRID ANN-CNN MODEL FOR PREDICTING NON-LINEAR RELATIONSHIP OF COVID-19 CASES BASED ON WEATHER FACTORS

By

YAHAYA MOHAMMED SANI * ANDREW GAHWERA **

*-** *Department of Computer Science, School of Computing and Information Technology, Makerere University, Kampala, Uganda.*

Date Received: 10/10/2023 Date Revised: 09/11/2023 Date Accepted: 29/11/2023

ABSTRACT

With the global increase in the emergence of viral diseases, the most recent being the Coronavirus Disease 2019 (COVID-19) in 2020-2021, it has decimated the world with little understanding of its history and the factors that influence its transmission dynamics. Weather significantly influences the spread of respiratory infectious diseases like influenza, yet the impact of weather on COVID-19 transmission in Nigeria remains unexamined and necessitates further clarification. This study presents and compares the results of six machine learning models, the developed Hybrid ANN-CNN, ANN, CNN, LSTM, LASSO, and Multiple Linear Regression models, aiming to predict the impact of weather factors on COVID-19 cases. The dataset used in this study includes daily datasets of Nigerian COVID-19 cases and seven weather variables collected from May 1, 2020, to April 30, 2021. The results indicate that the developed Hybrid ANN-CNN outperforms the remaining five models based on Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for all cases. Specifically, for confirmed cases, the Hybrid ANN-CNN had an MAE of 0.0274, for recovery cases 0.0257, and for death cases 0.0425. Similarly, for RMSE, the developed Hybrid ANN-CNN had values of 0.0469 for confirmed cases, 0.0813 for recovery cases, and 0.0840 for deaths. This was followed by LASSO with an MAE of 0.01384 and CNN and LSTM with 0.1384 and 0.1385, respectively.

Keywords: Hybrid ANN-CNN, LASSO, ANN, LSTM, CNN, MLR and COVID-19.

INTRODUCTION

With the global surge in the emergence of viral diseases, the most recent being the enigmatic coronavirus disease 2019 (COVID-19) in 2020-2021, the COVID-19 pandemic has infected over 50 million people in more than 100 countries, significantly impacting the world (Abdulkareem et al., 2021). It has devastated the planet, yet our understanding of its origin and the factors influencing its transmission dynamics remains limited (Ayanshina et al., 2020). Weather factors have a significant impact on patterns of respiratory infectious diseases like Middle East Respiratory Syndrome (MERS), Severe Acute Respiratory Syndrome (SARS), and influenza. However, the impact of weather factors on COVID-19 transmission remains controversial and requires further investigation and clarification (Ai et al., 2022; Abdulkareem et al., 2021). Weather variability is likely to have the greatest impact on the transmission and control of infectious diseases such as COVID-19 in developing countries. Nigeria, a developing country, has been severely affected by the global COVID-19 pandemic caused by COVID-19 (Moroh et al., 2023). Weather patterns in Nigeria vary between rainy and dry seasons. Rainfall, humidity, and temperature are all associated with increased host susceptibility to infectious diseases (Ayanshina et al., 2020). Bukhari and Jameel argued against the likelihood of a weather-related slowdown in the spread of COVID-19 in the United States and Europe. They cited a significant number of cases recorded under various weather conditions, such as temperature and humidity, in these countries. These observations are partially correct in Nigeria, where COVID-19 cases have been observed in both weather seasons. Weather changes are factors that can affect many infectious diseases and their geographical distribution (Ayanshina et al., 2020). The discovery of positive correlations between the spread of COVID-19 and weather conditions such as temperature and humidity is surprising and highlights the need for more research into the peculiarities of the COVID-19 epidemic in various areas worldwide, including Nigeria (Ayodele et al., 2021; Khursheed et al., 2021). Seasonal variations may influence disease vectors, pathogens, or human host susceptibility (Ayanshina et al., 2020).

Machine learning technologies have demonstrated their effectiveness in combating infectious diseases. However, there is limited research on applying machine learning to understand the relationship between weather conditions and the spread of COVID-19 (Abdulkareem et al., 2021). This study aims to predict the nonlinear impact of various weather variables-such as wind speed, precipitation, relative humidity, minimum temperature, mean temperature, maximum temperature, and solar radiation-on the daily confirmed cases, recoveries, and deaths caused by COVID-19. We conducted Spearman heat map correlation analysis in six states across Nigeria, with each state representing one of the six geopolitical zones. To this end, this study developed a hybrid ANN-CNN along with five other state-of-the-art machine learning models, including but not limited to LASSO, LSTM, ANNN, CNN, and multiple linear regression models. These models were used to predict the impact of weather risk factors on COVID-19 cases and account for the non-linear relationships between the weather risk factors and the daily confirmed, recovered, and fatal COVID-19 cases in six Nigerian states. Additionally, this study's contributions include determining the weather risk factors associated with COVID-19 infection, recovery, and death cases in Nigeria using the developed hybrid ANN-CNN model, which conducts feature extraction and performs better with limited medical records data, especially in the case of COVID-19 data scarcity. A formal representation of the model was provided. After completing the experiment, the results and model performance were compared based on MAE and RMSE. The developed Hybrid ANN-CNN outperforms the remaining five models in predicting the impact of meteorological variables on COVID-19 confirmed, recovery, and death cases, exhibiting improved performance. The developed Hybrid ANN-CNN can be used in the detection of vulnerable areas for possible COVID-19 community transmission, resulting in more targeted public health measures (Karmokar et al., 2022).

- A hybrid ANN-CNN model was developed to predict the impact of weather factors on the spread of COVID-19 cases with feature extraction and performs better under small amount of data since medical records are scarce, particularly COVID-19 data.
- The formal representation of the hybrid ANN-CNN model using mathematical expressions was developed.
- An architecture of the developed hybrid ANN-CNN model for predicting the impact of weather factors on the spread of COVID-19 cases was designed.
- Spearman correlation heat map was used to visualise the correlations between weather variables and COVID-19 cases.
- The developed hybrid ANN-CNN model outperformed other five state- of-the-art models evaluated based on MAE and RMSE used for models comparison performance evaluation metrics.

1. Related Work

Doni et al. (2022), developed a deep learning model to estimate the number of expected cases in India, taking into account some weather variables and COVID-19 data, that is, recovered, infected and death cases. The model was implemented using Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Bidirectional RNN (BRNN), Long Short-Term Memory (LSTM) and Bidirectional LSTM. Data was collected from the

Indian subcontinent. When compared to the other models, the results showed that the BRNN algorithm produced a better prediction model. However, this model was limited to predicting the association between recovered, infected and deceased individuals and their relationship to two environmental factors, namely humidity and temperature, with- out taking into account other variables such as solar radiation, precipitation and wind speed.

Prata et al. (2020), investigated the association between temperature and COVID-19 infection in Brazilian state capitals. Cumulative data with the daily number of confirmed cases were collected from 27 February to 1 April 2020. The linear and non-linear association between annual mean temperature compensation and confirmed cases was investigated using a Generalized Additive Model (GAM). In the temperature range of 16.8°C to 27.4°C, the GAM response curve showed a negative linear association between temperatures and daily cumulative con- firmed cases of COVID-19. The polynomial linear regression model predicted an R-squared of 0.81053. Temperatures were shown to have a negative linear association with the number of confirmed cases. This study only looked at the association between temperature and COVID-19 confirmed cases, neglecting re- covered cases and deaths, which could also be influenced by weather or other causes.

Abdulkareem et al. (2021), implemented three machine learning models (Convolutional Neural Network (CNN), ADtree Classifier and BayesNet) based on confirmed cases and weather factors such as temperature, humidity, wind and precipitation. The results of the experiments indicated that the CNN had the best accuracy of 99.56%. The ADtree classification approach achieved the second highest accuracy of 97.89% and the BayesNet classifier was the least accurate at 97.01%. As a result of this finding, it was clear that temperature, humidity, wind and rainfall were important factors in influencing the COVID19 confirmed cases. However, this model did not include the relationship between these weather variables and cases of recovery or death, and solar radiation, which is an important predictor variable.

Yuan et al. (2021), evaluated the influence of meteorological conditions on daily new COVID- 19 cases as of 31 August 2020. The log-linear Generalized Additive Model (GAM) was used to investigate the impact of meteorological variables on COVID-19 daily new cases. The authors' data showed that temperature, relative humidity and wind speed are non-linearly related to daily new cases. Wind speed (when greater than 7 m/s) and relative humidity (greater than 70%) were not statistically associated with COVID-19 transmission. The results of this study will be an important addition for public health policy makers to understand the weather dependence of COVID-19. Using only one model to draw conclusions is not appropriate in machine learning research, and also other important weather variables such as rainfall and solar radiation were not taken into account.

The use of machine learning models that specialize in feature extraction and selection based on feature importance and handle the non-linear relationship between weather factors and COVID-19 cases has not been captured in the previous studies reviewed to determine the impact of weather factors on COVID-19 transmission. Furthermore, none of the previous studies used a hybrid ANN- CNN to predict the impact of weather factors on COVID-19 cases with improved model performance, while overcoming the problem of small amount of dataset associated with medical record data, especially COVID-19. This provides a solid foundation for this research to fill the identified gaps, which is what this study aimed to achieve.

2. Methodology

Nigeria is the 32nd largest country in the world, covering a total area of 923,768 km^2 along the Gulf of Guinea in West Africa. It shares borders with Benin (773 km), Niger (1497 km), Chad (87 km), and Cameroon (1690 km), spanning a length of 4047 km, and boasts 853 kilometers of coastline. Positioned between latitudes 4° and 14° N and longitudes 2° and 15° E, Nigeria features Chappal Waddi, towering at 7936 feet, as its highest point. The primary rivers, Niger and Benue, converge to form the Niger Delta, one of the world's largest river deltas, housing a substantial area of Central African mangroves. This region encompasses the extensive topography of Nigeria, shaped by the valleys of the Niger and Benue rivers, merging into a Y-shaped formation. Southwest of the Niger, the terrain comprises 'rugged' highlands, while to the southeast of the Benue lie hills and mountains constituting the Mambilla Plateau, the nation's highest plateau. This plateau extends to the Cameroonian border, integrating into the Cameroonian Bamenda Highlands. Nigeria is geographically divided into six geopolitical zones, North-East, North-Central, North-West, South-East, South-South, and South-West, along with thirty-six states (Moroh et al., 2023). This study investigates the correlation between weather variables and seasonal fluctuations in the geographical distribution of COVID-19 cases, analyzing one state from each of the country's geopolitical zones: Adamawa State representing the North-East, the Federal Capital Territory (FCT) representing the North-Central, Kano representing the North-West, Enugu representing the South-East, Rivers representing the South-South, and Lagos representing the South-West. Figure 1 shows the map of Nigeria showing study areas.

3. Dataset Collection and Data Description

The data utilized for this study consisted of COVID-19 daily confirmed cases, daily recovery cases, daily deaths, and meteorological variables in Nigeria, sourced from various websites. The COVID-19 datasets encompassed Nigeria's daily confirmed cases totaling 90,786, daily recoveries totaling 70,476, and daily deaths reaching 990. Additionally, seven meteorological variables, wind speed, rainfall, relative humidity, minimum temperature, mean temperature, maximum temperature, and solar radiation were included. The COVID-19 cases were sourced from the Nigeria Centre for Disease Control (NCDC) through the NCDC Coronavirus COVID-19 microsite, while weather data was obtained from the data viewer used for predicting worldwide energy resources (NASA, 2020). The datasets were collected from six Nigerian states: Adamawa, Enugu, FCT, Kano, Rivers, and Lagos.

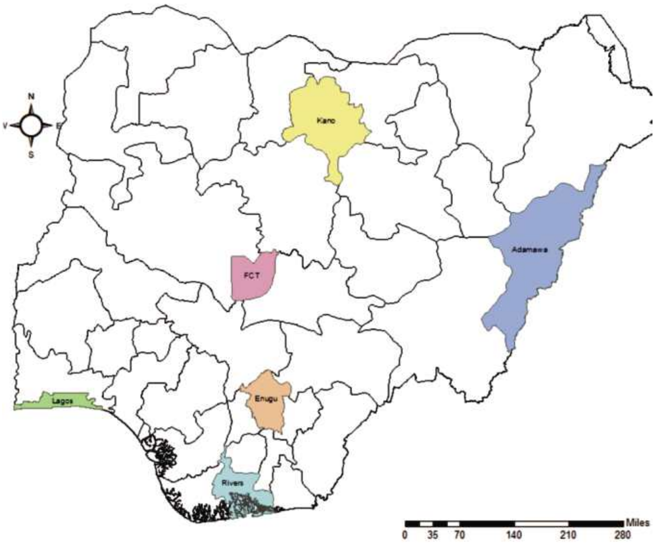


Figure 1. Map of Nigeria Showing Study Areas

3.1 Normalization and Feature Scaling Techniques

To aid the training process, the data values were rescaled using a two-step feature scaling method in this study, which is a hybrid of the max absolute values and minmax feature scalers, and an inverse transformation of the test data. In the most basic terms, the maxabs scaler takes the absolute maximum value of each column and divides each value by the maximum value. This procedure scales the data between (-1, 1), whereas minmax scales the values between (-1, 1), where -1 is the minimum value and 1 is the maximum value. The dataset used in this study was splitted into training and test datasets at a ratio of 80% and 20% for training and testing respectively.

3.2 Modelling Approach

The primary goal of any regression task is to create a model that fits well to both new and previously unseen data. A machine learning model must learn the training data well, without overfitting, in order to generalize well to the test data (Emmert-Streib & Dehmer, 2019).

3.2.1 Multiple Linear Regression Model (MLR)

The multiple linear regression model is a supervised learning method used in machine learning and general data science. Although multiple linear regression models have been known for a long time, several new developments in recent years have greatly extended this model (Emmert-Streib & Dehmer, 2019). One major disadvantage of MLR is that if there are outliers in the data, the model can perform very poorly. Additionally, it cannot distinguish the relationship between the predictor and target variables, which is another aspect that can lead to poor performance. The multiple linear regression equation takes the form:

$$y = \beta 0 + \beta 1x1 + \beta 2x2, .., + \beta pxp + \epsilon \quad (1)$$

Where β0, β1,..., is p + 1, the unknown parameter β0 can be called the regression constant, β1,..., are the regression coefficients. Y is the dependent variable, x1, x2,...,xp is p. Common variables that can be accurately measured and controlled at the same time are called independent variables (Xiao & Jin, 2021).

3.2.2 Least Absolute Shrinkage and Selection Operator (LASSO)

LASSO is a regression analysis method that performs both variable selection and regularization to improve the prediction accuracy and interpretability of the statistical regression model. With a large number of independent variables, we often want to identify a smaller subset of these variables that have the strongest effects. The sparsity of LASSO is mainly seen as an advantage because it leads to a simpler interpretation. A major disadvantage of LASSO is that it has no grouping property, that is, it tends to select only one variable from a group of highly correlated variables.

$$\sum_{i=1}^{n}\left(yi - \sum_{j} xi_j \beta_j\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \quad (2)$$

λ controls the strength of the L1 penalty. λ is basically the amount of shrinkage: if λ = 0, no parameters are dropped. The estimate is equal to the one obtained by linear regression. As λ increases, more and more coefficients are set to zero and dropped (theoretically, if λ = , all coefficients are dropped). As λ increases, the bias increases. As λ decreases, the variance increases. The LASSO in this study used λ = 0.0001 (Xiao & Jin, 2021).

3.2.3 Artificial Neural Network (ANN)

ANN computational approaches have many advantages over traditional machine learning models. An ANN is composed of non-linear parts and can learn an input-output mapping from a learner, adjusting its synaptic weights to adapt to the environment. It can also deal with partial information and provide results under uncertainty. It is worth noting that the analogy with the brain motivates or inspires ANNs, while the urge to create an artificial brain lags far behind. One difficulty in using an ANN model is determining the appropriate characteristics of the training data, the architecture of the network (number of layers and nodes) and the method to avoid overtfitting.

The following equations explain an ANN neuron in mathematical terms,

$$\sum_{j=0}^{n} wk_{jxj} \quad (3)$$

Where w0 is considered as bias and x0=1 (Bikku, 2020; Mas & Flores, 2008). The ANN used in this study adopted a density of three layers of 64, 16 and 1, the activation function was ReLU, loss = MAE, batch size=32 and the experiment was terminated after 100 epochs.

3.2.4 Convolutional Neural Network (CNN)

CNN is a class of deep learning methods that has become dominant in various computer machine learning tasks and is attracting interest in a variety of domains, including nonlinear regression. CNN is a mathematical construct that typically consists of three types of layers (or building blocks), convolution, pooling, and fully connected layers. The first two layers, convolution and pooling, perform feature extraction, while the third, a fully connected layer, maps the extracted features to the final output for both linear and non-linear regression tasks. A convolution layer plays a key role in the CNN, which consists of a stack of mathematical operations such as convolution, a special type of linear operation. Compared to general artificial neural networks, CNN has many advantages: local connections, where each neuron is no longer connected to all the neurons in the previous layer but only to a small number of neurons, effectively reducing parameters and speeding up convergence; weight sharing, where a group of connections can share the same weights, further reducing parameters and dimensionality.

$$yj = f\left(\sum_{i=1}^{n}\left(xj * wij\right) + bj\right) \quad (4)$$

xi represents the input feature; n features are simultaneously input to neuron j; wij represents the weight value of the connection between the input feature xi and neuron j; bj represents the internal state of neuron j, which is the bias value; and yj is the output of neuron j. f (·) is the activation function, which can be a sigmoid function, tanh(x) Rectified Linear Unit Function. Although CNN has many advantages and is widely used, there are many problems that make convolution difficult to handle, such as low generalization, lack of equivariance, and poor results in crowded scenes (Li et al., 2021; Yamashita et al., 2018).

The CNN used in this study adopted Conv1D, a density of three layers of 64, 16 and 1 units, flattening, the activation function was ReLU, loss = MAE, batch size=32 and the experiment was terminated after 100 epochs.

3.2.5 Long Short Term Memory (LSTM)

Long Short-Term Memory (LSTM) models, was based on a specialized deep neural network architecture, have emerged as a key model for time series fore- casting. LSTMs are specifically designed to store very long term temporal dependencies using memory cells containing multiple types of gates. In addition, LSTM can learn non-linearity. A major drawback of LSTMs is that they require more training data to learn more effectively, as the learning parameter cells of LSTMs are computationally expensive (Bolboacă & Haller, 2023; Malakar et al., 2021).

The following equations describe the operations of the LSTM network:

$$f(t) = sigm\,(W_f X(t) + U_f h(t-1) + b_f) \tag{5}$$

$$I(t) = sigm\,(W_I X(t) + U_I h(t-1) + b_I) \tag{6}$$

$$\dot{C}(t) = \tanh(W_C X(t-1) + U_C h(t-1) + b_C) \tag{7}$$

$$C(t) = f(t) \,.\, C(t-1) + I(t) \,.\, \dot{C}(t) \tag{8}$$

$$o(t) = sigm\,(W_o X(t) + U_o h(t-1) + b_o) \tag{9}$$

$$h(t) = o(t) \,.\tanh(C(t)) \tag{10}$$

W, U and b denote the weight matrices for the inputs, outputs, hidden layer and bias vector in Equations (6)-(11). The element-wise multiplication operation is denoted by ( ) in the same equations. The two activation functions are the sigmoid (sigm) and the hyperbolic angent (tanh) (Bolboacă & Haller, 2023).

The LSTM used in this work had densities of 64, 16 and 1, an activation function of ReLU, a loss of MAE, a batch size of 32 and the experiment ended after 100 epochs.

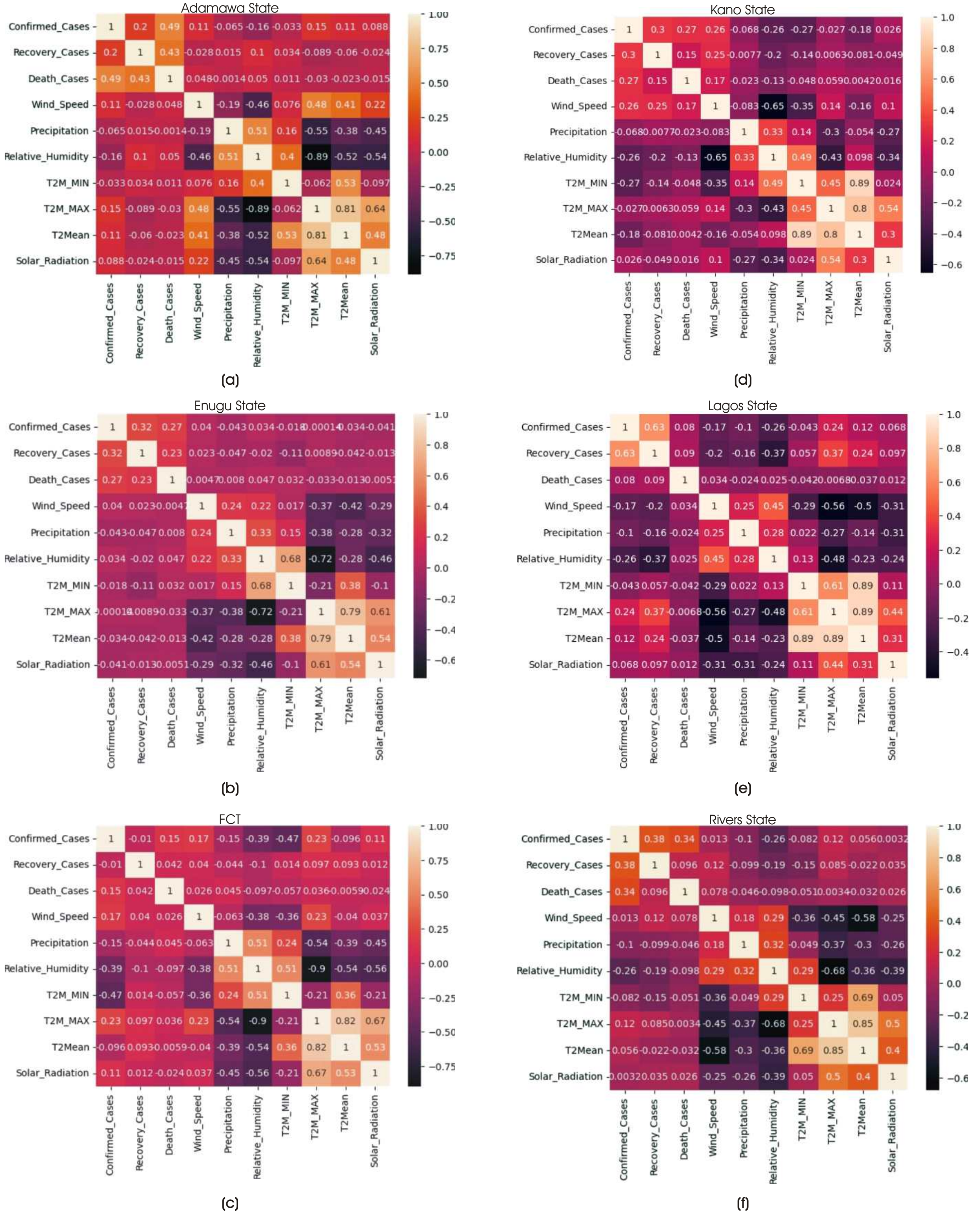
3.2.6 Spearman Correlation Heat Map

A Spearman correlation heat map visualizes the Spearman rank correlation coefficients between variables in a dataset. The Spearman correlation coefficient assesses the strength and direction of monotonic relationships between paired data variables, making it suitable for both linear and nonlinear associations. Figure 2 shows the Spearman Heapmap correlations.

In this study, Spearman heat map correlation analysis was used to determine the correlations between seven independent weather variables including but not limited to wind speed, rainfall, relative humidity, minimum temperature (T 2Min), maximum temperature (T Max), mean temperature (T2Mean) and solar radiation over COVID-19 confirmed, recovered and dead cases which are the dependent variables for the six states of Nigeria studied. The relationship between weather factors and COVID-19 cases showed a non-linear relationship in the Nigerian states.

First, in Adamawa State, wind speed had a weak positive correlation with a p-value of (0.11), negatively correlated with rainfall (-0.065) and relative humidity (-0.16), minimum temperature (-0.033), weak positive correlation with maximum temperature (0.15), mean temperature (0.11) and solar radiation (0.088) for COVID-19 confirmed cases. Similarly, for recovery cases, wind speed had a weak negative correlation with a p-value of (-0.028), a weak positive correlation with rainfall (0.015), relative humidity (0.1), minimum temperature (0.034) and a weak negative correlation with maximum temperature (-0.089), mean temperature (-0.06) and solar radiation (-0.024). Also for deaths, wind speed had a weak positive correlation with a p-value of (0.048), rainfall (0.0014) and relative humidity (0.05), minimum temperature (0.011), and a weak negative correlation with maximum temperature (-0.03), mean temperature (-0.023) and solar radiation (-0.015).

Secondly, in Enugu State, wind speed had a weak positive correlation with a p-value of (0.04), weak negative



(a) (d)

(b) (e)

(c) (f)

Figure 2. Spearman Heapmap Correlations

correlation with rainfall (-0.043), weak positive correlation with relative humidity (0.034), weak negative correlation with minimum temperature (-0.018), maximum temperature (-0.0001), mean temperature (-0.034) and solar radiation (-0.041) for COVID-19 confirmed cases. Similarly, for recovery cases, wind speed had a weak positive correlation with a p-value of (0.023), a negative negative correlation with precipitation (-0.0047) and relative humidity (-0.02), minimum temperature (-0.11), maximum temperature (- 0.089), mean temperature (-0.042) and solar radiation (-0.013). Also for deaths, wind speed had a weak negative correlation with a p-value of (-0.005), a weak positive correlation with rainfall (0.008), relative humidity (0.047), minimum temperature (0.032), a weak negative correlation with maximum temperature (-0.033), mean temperature (-0.013) and solar radiation (-0.005).

Thirdly, for FCT, wind speed had a weak positive correlation with a p- value of (0.17), negatively correlated with precipitation (-0.15), relative humidity (-0.39), minimum temperature (-0.47), positively correlated with maximum temperature (0.23), negatively correlated with mean temperature (-0.096) and positively correlated with solar radiation (0.11) for COVID-19 confirmed cases. Similarly, for recovery cases, wind speed had a weak positive correlation with a p-value of (0.04), negatively correlated with precipitation (-0.044) and relative humidity (-0.01), positively correlated with minimum temperature (0.014), maximum temperature (0.097), mean temperature (0.093) and solar radiation (0.012). Also for deaths, wind speed had a weak positive correlation with p- value of (0.025), precipitation (0.045) and was negatively correlated with relative humidity (-0.097), minimum temperature (-0.057), a weak positive correlation with maximum temperature (0.04), mean temperature (0.006) and was negatively correlated with solar radiation (-0.024).

Fourthly, in Kano State, wind speed had a weak positive correlation with a p-value of (0.26), negatively correlated with rainfall (-0.07), relative humidity (-0.026), minimum temperature (-0.27), positive correlation with maximum temperature (0.27), negatively correlated with mean temperature (-0.18) and positively correlated with solar radiation (0.026) for COVID-19 confirmed cases. Similarly, for recovery cases, wind speed had a positive correlation with a p-value of (0.25), precipitation (0.008), was negatively correlated with relative humidity (-0.02), minimum temperature (-0.014), weakly positively correlated with maximum temperature (0.006), weakly negatively correlated with mean temperature (-0.08) and solar radiation (-0.049). Also for deaths, wind speed had a weak positive correlation with a p-value of (0.17), a weak negative correlation with precipitation (-0.023), relative humidity (-0.13), minimum temperature (-0.048), a weak positive correlation with maximum temperature (0.059), mean temperature (0.0042) and a negative correlation with solar radiation (0.016).

Fifthly, in Lagos State, Wind Speed had a weak negative correlation with p-value of (-0.17), Precipitation (-0.01), Relative Humidity (-0.26), Minimum Temperature (-0.043), positively correlated with Maximum Temperature (0.24), Mean Temperature (0.12) and Solar Radiation (0.068) for COVID-19 confirmed cases. Similarly, for recovery cases, wind speed had a weak negative correlation with a p-value of (-0.2), precipitation (-0.016), positive correlation with relative humidity (0.37), minimum temperature (0.057), maximum temperature (0.37), mean temperature (0.24) and solar radiation (0.097). Also for deaths, wind speed had a weak positive correlation with a p-value of (0.034), precipitation (0.024), relative humidity (0.025), a negative correlation with minimum temperature (-0.042), a weak positive correlation with maximum temperature (-0.0068), mean temperature (-0.037) and a weak positive correlation with solar radiation (0.012).

Finally, in Rivers State, for COVID-19 confirmed cases, wind speed had a weak positive correlation with p-value of (0.013), weak negative correlation with precipitation (-0.1), relative humidity (-0.26), minimum temperature (-0.082), weak positive correlation with maximum temperature (0.12), mean temperature (0.0056) and solar radiation (0.003). Similarly, for recovery cases, wind speed had a weak positive correlation with p-value of (0.12), negative correlation with precipitation (-0.0099), relative humidity (-0.19), minimum temperature (-0.15), weak positive correlation with maximum temperature (0.085), weak negative correlation with mean temperature (-0.022) and solar radiation (-0.035). Similarly, for deaths, wind speed had a weak positive correlation with a p-value of (0.078), and was negatively correlated with precipitation (-0.046), relative humidity (-0.098), minimum temperature (-0.051), weakly positively correlated with maximum temperature (0.0034), weakly negatively correlated with mean temperature (-0.032) and weakly positively correlated with solar radiation (0.026).

3.2.7 The Developed Hybrid ANN-CNN Model

The developed hybrid ANN-CNN is an integration of ANN and CNN models to form a consistent whole model. ANN prides itself for performing robust non-linear machine learning operation and even with small amount of data, since medical records data are scare, especially COVID-19 data. CNN on the other hand perform feature extraction, and the fully connected layer, maps the extracted features to the final output, for both linear and non-linear regression tasks. Each neuron in the CNN is no longer connected to all the neurons in the previous layer, but only to a small number of neurons, which is effective in reducing parameters and speeding up convergence and weight distribution, and both ANN and CNN work well with time series data. These characteristics of the two models overcome the drawbacks of each individual model discussed above, with improved performance in the prediction of weather factors against COVID-19 daily confirmed and recovery cases based on MAE and RMSE models comparison performance evaluation metrics. The mathematical equation for the developed hybrid ANN-CNN is as follows:

$$\text{Suppose}X(t) \rightarrow Y(t) \tag{11}$$

$$X(t) = (xw1(t), xw2(t).......xwn(t)) \tag{12}$$

$$Y(t) = (y1(t), y2(t)..........yn(t)) \tag{13}$$

Where X(t) and Y(t) are the Hybrid ANN-CNN inputs and outputs variables pairs at time t, respectively.

Suppose:

$$X(t) = D(tw),\ 1 \leq t \leq T \tag{14}$$

D(tw), is the daily weather variables including rainfall, relative humidity, wind speed, minimum temperature, mean temperature, maximum temperature and solar radiation on day t as independent variables against the corresponding dependent variables which are number of daily confirmed, recovery and death cases and T is the number of days for the training datasets for both independent and dependent variables.

The developed hybrid ANN-CNN network was trained using time series data to predict the impact of weather factors consisting of seven weather variables rain- fall, relative humidity, wind speed, minimum temperature, mean temperature, maximum temperature and solar radiation as independent variables against COVID-19 confirmed, recovery and death cases. The network architecture was implemented using ANN neurons and a ConV1D layer (each with output vectors of size 365 elements). The network consists of three linear layers of 64, 16 and 1 respectively. The developed Hybrid ANN-CNN model adopted ADAM optimiser, using a time window of 365 days (Dtw = 365) and the activation function was ReLU for non-linear operation, kernel regularizer = L1 to overcome the problem of overfitting, loss = MAE, batch size=32 and the experiment was terminated after 100 epochs. For each dataset used in this study, both weather and COVID-19 datasets from 1st May, 2020 up to 292 days, which is 80% of the dataset, were used for training and the remaining dataset for 73 days, which is 20% of the dataset, was used for testing. The testing phase was repeated to estimate a set of future days prediction as follows:

$$\hat{Y}i = \text{Etimate}(x1wi(t), x2wi(t)......xnwi(t)),\ 1 \leq i \leq N \tag{15}$$

Where N denote the number of days in the test dataset.

Mean Absolute Values (MAE) were used for the loss functions and the math. Figure 3 shows the Architecture of hybrid ANN-CNN Model for the predicting weather factors against COVID-19 confirmed Recovery and death cases and mathematical equation was as follows:

$$\frac{1}{N}\sum_{i=1}^{D}\left|x_i - x\right| \tag{16}$$

where n = the number of errors, Σ = the summation

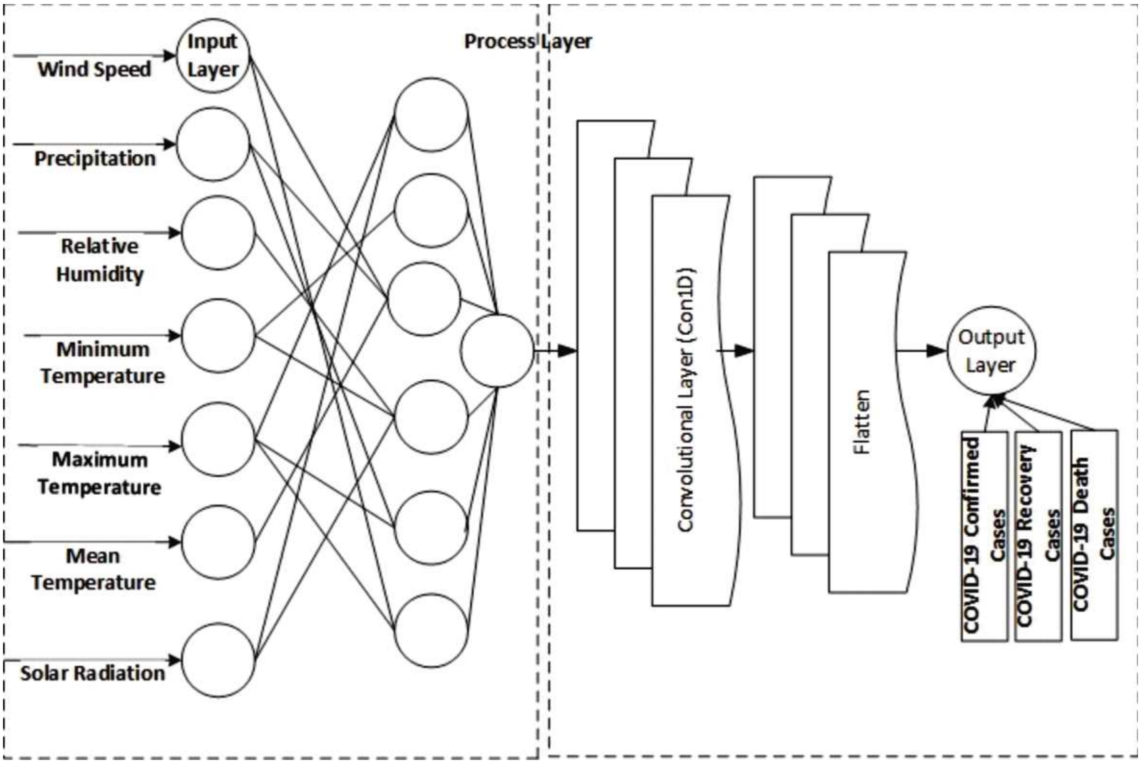


Figure 3. Architecture of Hybrid ANN-CNN Model

symbol and —xi - x— = the absolute errors.

And Root Mean Square Error (RMSE) which is the square root of Mean Square Error (MSE) was also used to further analyze the depth of error and justify the better performance of the proposed hybrid ANN-CNN model. The mathematical equation for RMSE is given below,

$$\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{N}} \qquad (17)$$

4. Results

4.1 One Year Trend for COVID-19 Confirmed, Recovery and Deaths Cases

Figure 4 shows the one year trend for COVID-19 Confirmed, Recovery and Deaths Cases, Adamawa, Enugu, FCT, Kano, Lagos and Rivers states in Nigeria.

Table 1 shows the results of predicting weather variables over COVID- 19 confirmed, recovery and death cases for Adamawa, Enugu, FCT, Kano, Lagos and Rivers states based on MAE performance metrics. To this end, the developed Hybrid ANN-CNN model outperformed the remaining five state-of-the-art models for all cases. For confirmed cases, the developed Hybrid ANN-CNN had MAE of (0.0274), LASSO (0.0277), LSTM (0.0280), MLR (0.0281), ANN (0.0296) and CNN (0.0296) for Adamawa. Similarly, for recovery cases, Hybrid ANN-CNN had an MAE of (0.0257), LSTM (0.0364), ANN (0.0368), CNN (0.0376), LASSO (0.0409) and MLR (0.0423). And for deaths, Hybrid ANN- CNN had MAE of (0.0425), LSTM (0.0425), MLR (0.0438), LASSO (0.0440), ANN (0.0442) and CNN (0.0457) respectively. In Enugu State, Hybrid had MAE of (0.0833) for confirmed cases, MLR (0.1120), LASSO (0.1127), ANN (0.1138), CNN (0.1156) and LSTM (0.1169). For recovery cases, Hybrid ANN- CNN had MAE of (0.0552), LASSO (0.0554), MLR (0.0563), LSTM (0.0566), ANN (0.0573) and CNN (0.0574). Similarly, for deaths, the Hybrid ANN-CNN had MAE of (0.0207), LSTM (0.0212), ANN (0.0213), CNN (0.0220), LASSO (0.0224) and MLR (0.0226). For FCT, the hybrid ANN-CNN had an MAE of (0.0946) for confirmed cases, followed by LASSO (0.0964), CNN (0.1063), LSTM (0.1072), ANN (0.1072) and MLR (0.5288). For recovery cases, Hybrid ANN-CNN had MAE of (0.0060), LSTM (0.0146), LASSO (0.0183), MLR (O.O186), ANN (0.0197)

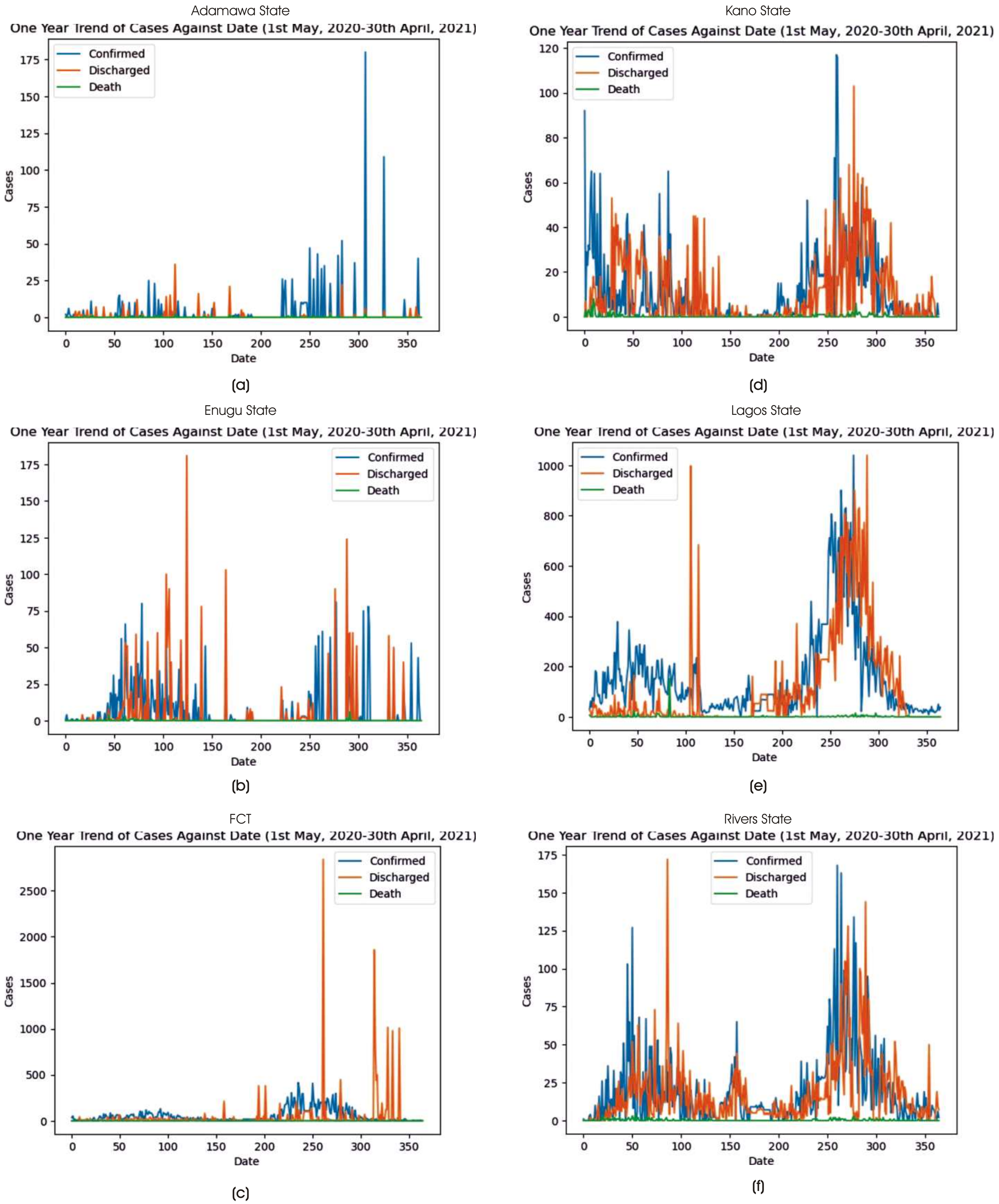


Figure 4. One Year Trend For COVID-19 Confirmed, Recovery and Death Cases

State	Cases	MLR	ANN	CNN	LSTM	LASSO	Hybrid
Adamawa	Confirmed	0.0281	0.0296	0.0296	0.0280	0.0277	0.0274
	Recovery	0.0423	0.0368	0.0376	0.0364	0.0409	0.0257
	Death	0.0438	0.0442	0.0457	0.0425	0.0440	0.0425
Enugu	Confirmed	0.1120	0.1138	0.1156	0.1169	0.1127	0.0833
	Recovery	0.0563	0.0573	0.0574	0.0566	0.0554	0.0552
	Death	0.0226	0.0213	0.0220	0.0212	0.0224	0.0207
FCT	Confirmed	0.5288	0.1072	0.1063	0.1072	0.0964	0.0946
	Recovery	0.0186	0.0197	0.0207	0.0146	0.0183	0.0060
	Death	0.0684	0.0697	0.0687	0.0644	0.0682	0.0519
Kano	Confirmed	0.6286	0.1016	0.1012	0.1007	0.1013	0.0988
	Recovery	0.1113	0.1112	0.1121	0.1115	0.1138	0.1111
	Death	0.0697	0.0722	0.0745	0.0732	0.0731	0.0636
Lagos	Confirmed	0.1106	0.1106	0.1109	0.1110	0.1107	0.1098
	Recovery	0.5906	0.1240	0.1247	0.1253	0.1207	0.1194
	Death	0.1528	0.0134	0.0135	0.0133	0.0144	0.0131
Rivers	Confirmed	0.0988	0.1007	0.1038	0.0992	0.0985	0.0982
	Recovery	0.0771	0.0804	0.0802	0.0750	0.0801	0.0737
	Death	0.6989	0.1063	0.1081	0.1064	0.1062	0.1057

Table 1. Weather Factors Vs. Confirmed, Recovery and Death Cases (MAE)

and CNN (0.0207). For deaths, the Hyrid ANN- CNN model had MAE of (0.0519), LSTM (0.0644), LASSO (0.0682), MLR (0.0684), CNN (0.0687) and ANN (0.0697). The hybrid in Kano State had MAE of (0.0988) for confirmed cases, LSTM (0.1007), CNN (0.1012), LASSO (0.1013), ANN (0.1016) and MLR (0.6286). For recovery cases, Hybrid ANN- CNN had MAE of (0.1111), ANN (0.1112), MLR (0.1113), LSTM (0.1115), CNN (0.1121) and LASSO (0.1138). For deaths, Hybrid ANN-CNN had MAE of (0.0636), MLR (0.0697), ANN (0.0722), LASSO (0.0731), LSTM (0.0732) and CNN (0.0745).

In Lagos State, Hybrid ANN-CNN had MAE of (0.1098) for confirmed cases, MLR (0.1106), ANN (0.1106), LASSO (0.1107), CNN (0.1109) and LSTM (0.1110). For recovery cases, the Hybrid model had MAEs of (0.1194), LASSO (0.1207), ANN (0.1240), CNN (0.1247) and LSTM (0.1253).

Similarly, for deaths, Hybrid ANN-CNN had MAE of (0.0131), LSTM (0.0133), ANN (0.0134), CNN (0.0135) and LASSO (0.0144). In Rivers State, Hybrid ANN-CNN had an MAE of (0.0982) for confirmed cases, followed by LASSO (0.0985), MLR (0.0988), LSTM (0.0992), ANN (0.1007) and CNN (0.1038). For recovery cases, Hybrid ANN-CNN had MAE of (0.0737), LSTM (0.0750), MLR (0.0771), LASSO (0.0801), CNN (0.0802) and ANN (0.0804) respectively.

For the case of death, the hybrid ANN-CNN had MAE of (0.1057), LASSO (0.1062), ANN (0.1063), LSTM (0.1064), CNN (0.1081) and MLR (0.6989) respectively.

Table 2 shows the results of predicting weather variables over COVID- 19 cases for Adamawa, Enugu, FCT, Kano, Lagos and Rivers states based on RMSE performance metrics. To this end, the developed Hybrid ANN-CNN model outperformed the remaining five state-of-the-art models for all cases. For confirmed cases, the developed Hybrid ANN-CNN had RMSE of (0.0469), LASSO (0.0489), LSTM (0.0495), ANN (0.0495), CNN (0.0495) and MLR (0.0497) for Adamawa. Similarly, for recovery cases, Hybrid ANN-CNN had an RMSE of (0.0813), CNN (0.0853), LSTM (0.0854), ANN (0.0855), LASSO (0.0879) and MLR (0.0887). And for deaths, Hybrid ANN- CNN had RMSE of (0.0840), MLR (0.0904), CNN (0.0905), LASSO (0.0905), ANN (0.0906) and LSTM (0.0907) respectively. In Enugu State, Hybrid had RMSE of (0.1626) for confirmed cases, MLR (0.1734), LASSO (0.1737), CNN (0.1741), LSTM (0.1741) and ANN (0.1742). For recovery cases, Hybrid ANN- CNN had RMSE of (0.1272), LASSO (0.1272), CNN (0.1279), LSTM (0.1284), ANN (0.1284) and MLR (0.1325). Similarly, for deaths, the Hybrid ANN-CNN had RMSE of (0.0207), LSTM (0.0212), ANN (0.0213), CNN (0.0220), LASSO (0.0224) and MLR (0.0226). For FCT, the hybrid ANN-CNN had an MAE of (0.1105) for confirmed cases, followed by MLR (0.1182), LASSO (0.1196), CNN (0.1424), LSTM (0.1427) and ANN (0.1429). For recovery cases, Hybrid ANN- CNN had RMSE of (0.0163), CNN (0.0381),

State	Cases	MLR	ANN	CNN	LSTM	LASSO	Hybrid
Adamawa	Confirmed	0.0497	0.0495	0.0495	0.0495	0.0489	0.0469
	Recovery	0.0887	0.0855	0.0853	0.0854	0.0879	0.0813
	Death	0.0904	0.0906	0.0905	0.0907	0.0905	0.0840
Enugu	Confirmed	0.1734	0.1742	0.1741	0.1741	0.1737	0.1626
	Recovery	0.1325	0.1284	0.1279	0.1284	0.1272	0.1272
	Death	0.0385	0.0381	0.0383	0.0379	0.0384	0.0357
FCT	Confirmed	0.1182	0.1429	0.1424	0.1427	0.1196	0.1105
	Recovery	0.0390	0.0388	0.0381	0.0383	0.0387	0.0163
	Death	0.0865	0.0843	0.0843	0.0843	0.0861	0.0843
Kano	Confirmed	0.9987	0.1413	0.1413	0.1413	0.1413	0.1347
	Recovery	0.1390	0.1330	0.1343	0.1342	0.1384	0.1324
	Death	0.1418	0.1416	0.1421	0.1422	0.1416	0.1366
Lagos	Confirmed	0.1656	0.1689	0.1689	0.1690	0.1689	0.1543
	Recovery	0.9151	0.1790	0.1788	0.1784	0.1788	0.1588
	Death	0.4664	0.0234	0.0231	0.0230	0.0241	0.0225
Rivers	Confirmed	0.1338	0.1310	0.1310	0.1303	0.1304	0.1282
	Recovery	0.1184	0.1233	0.1231	0.1210	0.1232	0.0938
	Death	0.1436	0.1409	0.1384	0.1385	0.1384	0.1347

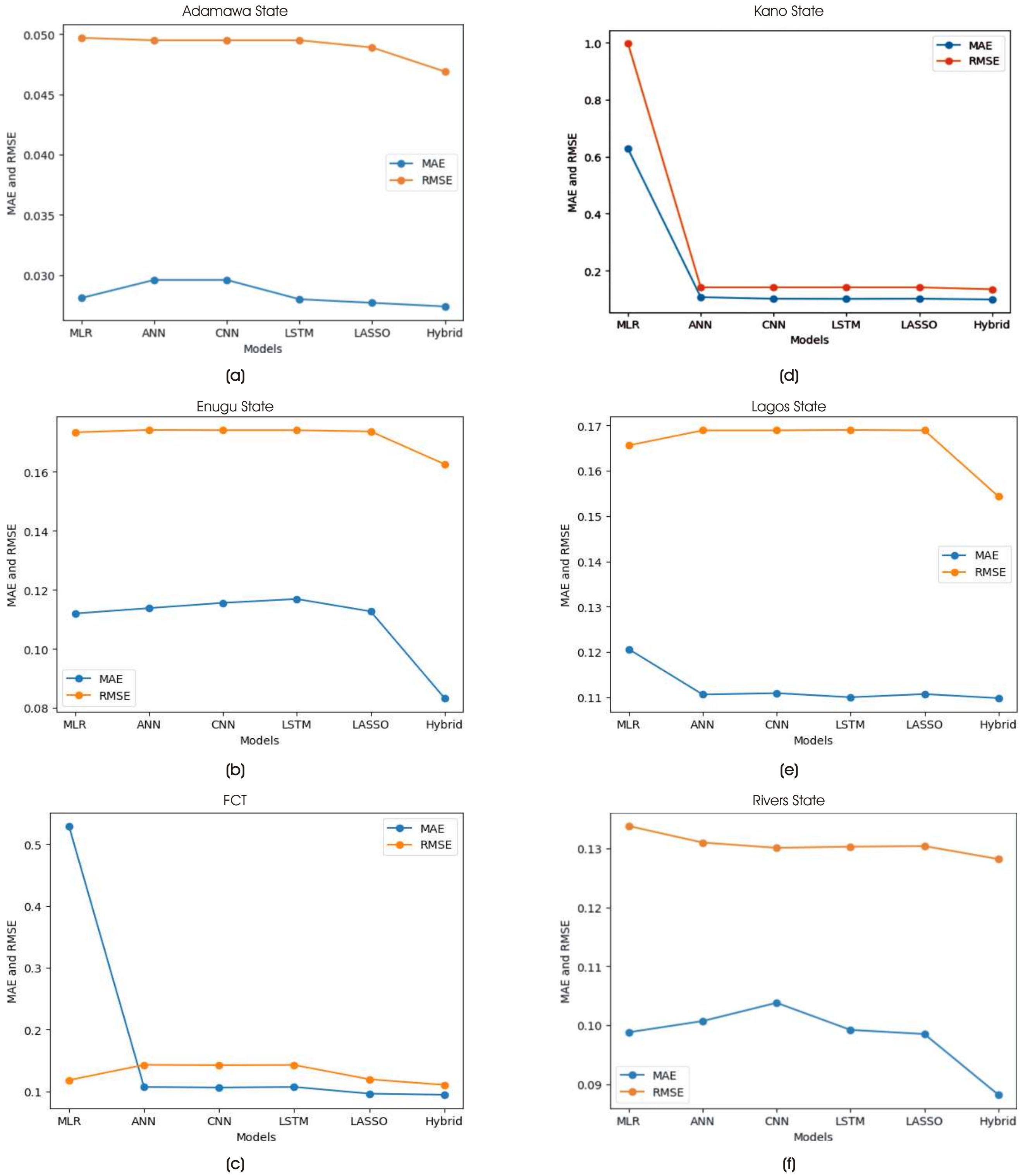
Table 2. Weather Factors Vs. Cases (RMSE)

LSTM (0.0383), LASSO (0.0387), ANN (0.0388) and MLR (O.O390). For deaths, the Hyrid ANN-CNN model had RMSE of (0.0843), LSTM (0.0843), ANN (0.0843), CNN (0.0843), LASSO (0.0861) and MLR (0.0865). The hybrid in Kano State had RMSE of (0.1347) for confirmed cases, LSTM (0.1413), CNN (0.1413), LASSO (0.1413), ANN (0.1413) and MLR (0.9987). For recovery cases, Hybrid ANN-CNN had RMSE of (0.1324), ANN (0.1330), LSTM (0.1342), CNN (0.1343), LASSO (0.1384) and MLR (0.1390). For deaths, Hybrid ANN-CNN had RMSE of (0.1366), ANN (0.1416), LASSO (0.1416), MLR (0.1418), CNN (0.1421) and LSTM (0.1422).

In Lagos State, Hybrid ANN-CNN had RMSE of (0.1543) for confirmed cases, MLR (0.1656), ANN (0.1689), LASSO (0.1689), CNN (0.1689) and LSTM (0.1690). For recovery cases, the Hybrid model had RMSE of (0.1588), LSTM (0.1784), CNN (0.1788), LASSO (0.1788), ANN (0.1790) and MLR (0.9151). Similarly, for deaths, Hybrid ANN-CNN had RMSE of (0.0225), LSTM (0.0230), ANN (0.0231), CNN (0.0234), LASSO (0.0241) and MLR (0.4664). In Rivers State, Hybrid ANN-CNN had an RMSE of (0.1282) for confirmed cases, followed by LSTM (0.1303), LASSO (0.1304), ANN (0.1310) CNN (0.1310) and MLR (0.1338). For recovery cases, Hybrid ANN-CNN had RMSE of (0.0938), LSTM (0.1210), CNN (0.1231), LASSO (0.1232), ANN (0.1233) and MLR (0.1184) respectively. For the case of death, the hybrid ANN-CNN had RMSE of (0.1347), LASSO (0.1384), CNN (0.1384), LSTM (0.1385), ANN (0.1409) and MLR (0.1436) respectively.

5. Discussion

This study developed a hybrid ANN-CNN model and five other state-of-the-art machine learning models to predict the impact of weather variables on COVID- 19 cases spread in six Nigerian states. The developed hybrid ANN-CNN model outperforms other models and helps to identify vulnerable areas for possible COVID-19 transmission. Previous studies have shown a positive correlation between COVID-19 spread and weather conditions in some countries, highlighting the need for further research into the peculiarities of the epidemic in countries like Nigeria (Ayanshina et al., 2020; Ayodele et al., 2021). This study predicts the non-linearity of seven weather variables, wind speed, precipitation, relative humidity, minimum temperature, mean temperature, maximum temperature and solar radiation against daily confirmed, recovered and death cases in six Nigerian states. This study used Spearman heat map correlation analysis to examine the correlations between seven weather variables, including wind speed, rainfall, relative humidity, minimum temperature, maximum temperature, mean temperature and solar radiation, over COVID-19 confirmed, recovered and death cases in six Nigerian states. The results showed a non-linear relationship between weather factors and COVID-19 cases in the six states, which include but are not limited to; Adamawa, Enugu, FCT, Kano, Lagos and Rivers states respectively. Figure 5 shows the line plots showing models performance comparison for confirmed cases for Adamawa, Enugu, FACT, Kano, Lagos and Rivers based on MAE and RMSE.



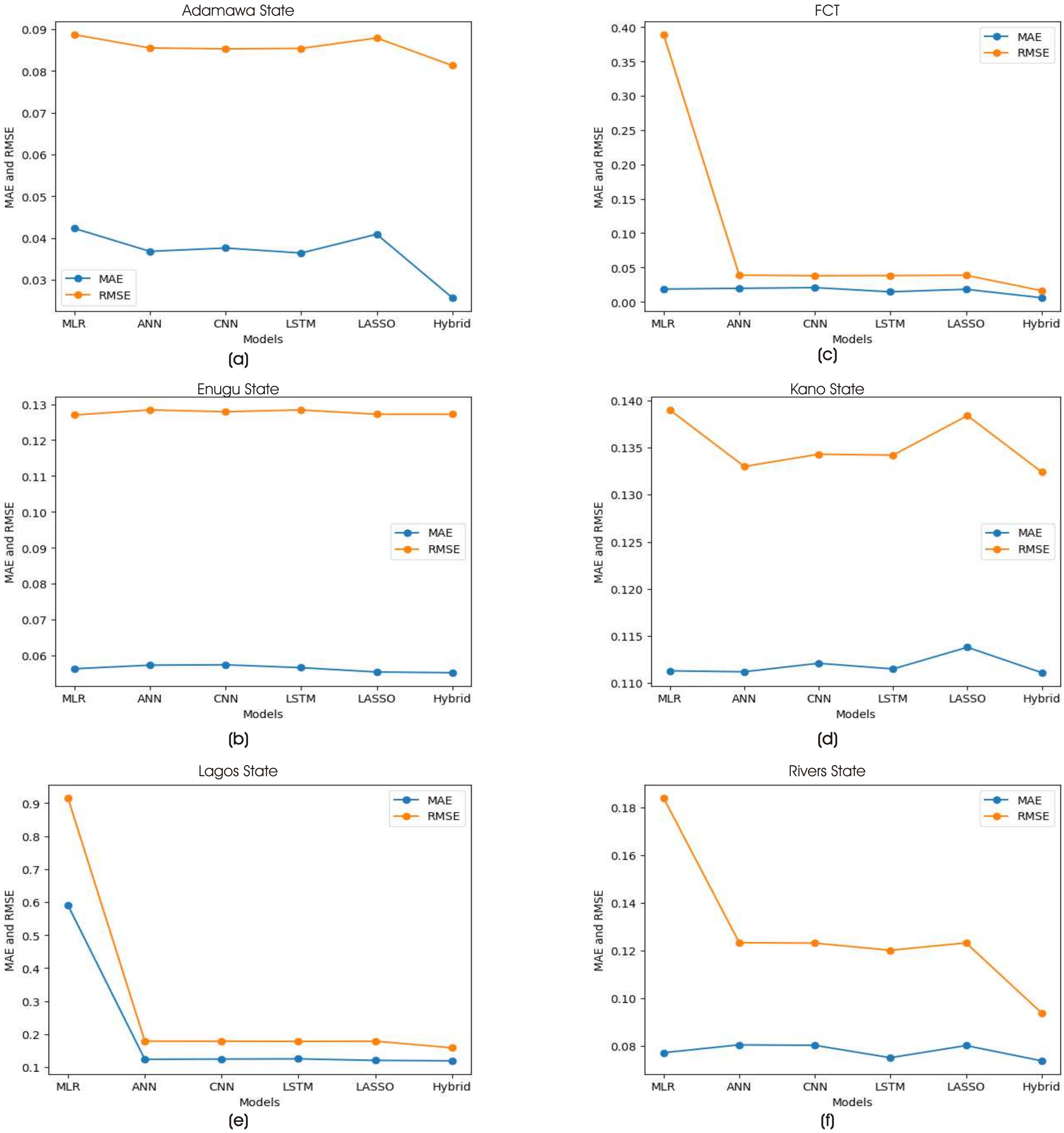
(a) (d)

(b) (e)

(c) (f)

Figure 5. Line Plots Showing Models Performance Comparison for Confirmed Cases for Adamawa, Enugu, FACT, Kano, Lagos and Rivers Based on MAE and RMSE

The results of prediction of weather variables over COVID-19 confirmed, recovery and death cases for Adamawa, Enugu, FCT, Kano, Lagos and Rivers states and the models performance were compared based on MAE and RMSE models performance evaluation metrics. The developed hybrid ANN-CNN model out- performed the remaining five models which include, MLR, ANN, CNN, LSTM and LASSO based on MAE and RMSE for confirmed, recovery and death cases in all the six states respectively. The outcome of this developed model can be used as an alarm system to detect vulnerable regions for possible COVID-19 community transmission and in need of urgent public interventions. Figure 6 shows the line plots showing models performance comparison for recovery cases for



(a) (c) (b) (d) (e) (f)

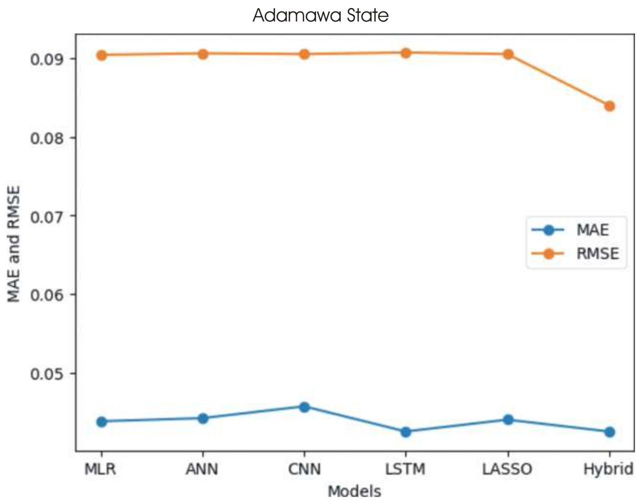
Figure 6. Line Plots Showing Models Performance Comparison for Recovery Cases for Adamawa, Enugu, FACT, Kano, Lagos and Rivers Based on MAE and RMSE

Adamawa, Enugu, FACT, Kano, Lagos and Rivers based on MAE and RMSE.

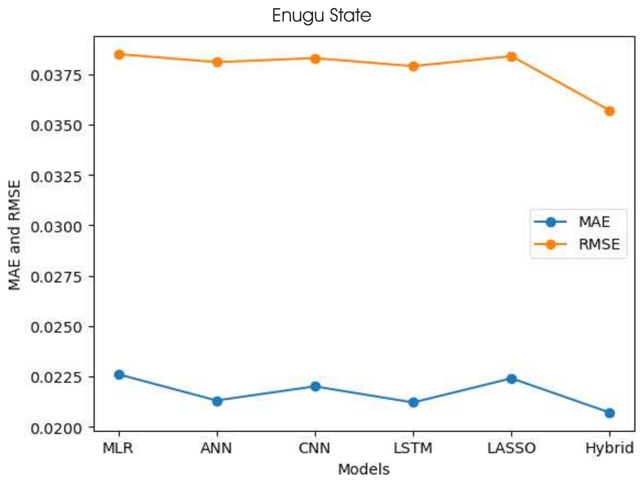
The results showed a non-linear relationship between weather factors over COVID-19 cases in the six states namely, Adamawa, Enugu, FCT, Kano, Lagos and Rivers states. The performance of the models was compared based on MAE and RMSE performance evaluation metrics. The developed hybrid ANN-CNN model performed better than the remaining five models include MLR, ANN, CNN, LSTM and LASSO based on MAE and RMSE for confirmed, recovery and death cases in all the six states respectively. The contributions of this study includes the identification of weather risk factors associated with COVID-19 infection, recovery and death cases in Nigeria, and the developed hybrid ANN-CNN model performs better and also extracts relevant features and fits well with small amounts of data. A formal representation of the developed Hybrid ANN-CNN model has also been provided with Figure 7 showing the line plots for models performance comparison for death cases for Adamawa, Enugu, FACT, Kano, Lagos and Rivers based on MAE and RMSE.

Conclusion

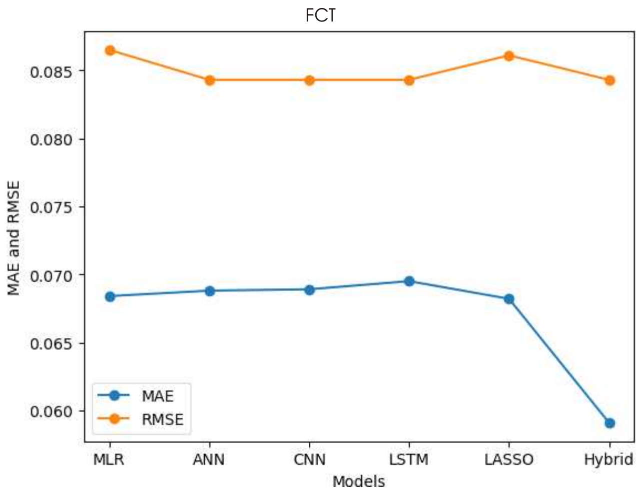
This study developed a hybrid ANN-CNN model and compared five other state-of-the-art machine learning models to predict the effects of weather variables on COVID- 19 prevalence in six states of Nigeria. The study predicted the non-linearities of seven weather variables: wind speed, precipitation, relative humidity, minimum temperature, mean temperature, maximum temperature



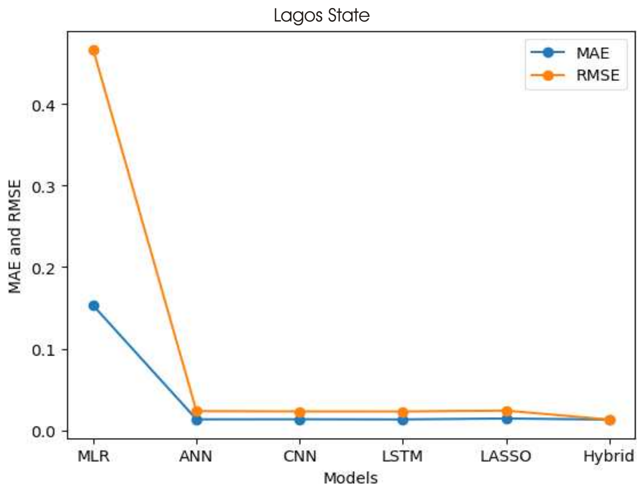
(a)



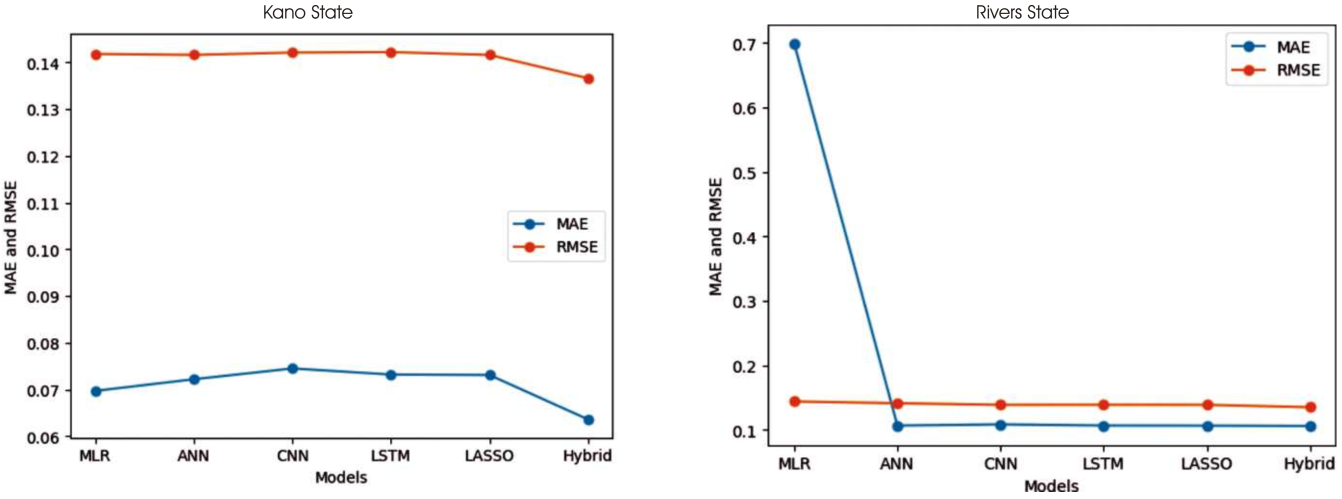
(b)



(c)



(e)



(d)

(f)

Figure 7. Line Plots Showing Models Performance Comparison for Death Cases for Adamawa, Enugu, FACT, Kano, Lagos and Rivers Based on MAE and RMSE

and solar radiation against daily confirmed, recovery and death cases. A Spearman heat map correlation analysis was conducted between seven weather variables including wind speed, precipitation, relative humidity, minimum temperature, maximum temperature, mean temperature and solar radiation against COVID-19 confirmed, recoveries and deaths.

References

[1]. Abdulkareem, A. B., Sani, N. S., Sahran, S., Alyessari, Z. A. A., Adam, A., Abd Rahman, A. H., & Abdulkarem, A. B. (2021). Predicting COVID-19 based on environmental factors with machine learning. *Intelligent Automation & Soft Computing,* 28(2), 305-320. https://doi.org/10.32604/iasc.2021.015413

[2]. Ai, H., Nie, R., & Wang, X. (2022). Evaluation of the effects of meteorological factors on COVID-19 prevalence by the distributed lag nonlinear model. *Journal of Translational Medicine,* 20(1), 1-9. https://doi.org/10.1186/s12967-022-03371-1

[3]. Ayanshina, O. A., Adeshakin, A. O., Afolabi, L. O., Adeshakin, F. O., Alli-Balogun, G. O., Yan, D., & Wan, X. (2020). Seasonal variations in Nigeria: Understanding COVID-19 transmission dynamics and immune responses. *Journal of Global Health Reports,* 4, e2020084.

[4]. Ayodele, K. P., Jimoh, H., Fagbamigbe, A. F., & Onakpoya, O. H. (2021). The dynamics of COVID-19 outbreak in Nigeria: A sub-national analysis. *Scientific African,* 13, e00914. https://doi.org/10.1016/j.sciaf.2021.e00914

[5]. Bikku, T. (2020). Multi-layered deep learning perceptron approach for health risk prediction. *Journal of Big Data,* 7(1), 1-14. https://doi.org/10.1186/s40537-020-00316-7

[6]. Bolboacă, R., & Haller, P. (2023). Performance analysis of long short-term memory predictive neural networks on time series data. *Mathematics,* 11(6), 1432. https://doi.org/10.3390/math11061432

[7]. Doni, A.R., Praba, T.S., & Murugan, S. (2022). Weather and population based forecasting of novel COVID-19 using deep learning approaches. *International Journal of System Assurance Engineering and Management,* 13(1), 100-110. https://doi.org/10.1007/s13198-021-01272-y

[8]. Emmert-Streib, F., & Dehmer, M. (2019). High-dimensional LASSO-based computational regression models: Regularization, shrinkage, and selection. *Machine Learning and Knowledge Extraction,* 1(1), 359-383. https://doi.org/10.3390/make1010021

[9]. Karmokar, J., Islam, M. A., Uddin, M., Hassan, M. R., & Yousuf, M. S. I. (2022). An assessment of meteorological parameters effects on COVID-19 pandemic in Bangladesh using machine learning models.

Environmental Science and Pollution Research, 29(44), 67103-67114. https://doi.org/10.1007/s11356-022-20196-z

[10]. Khursheed, A., Mustafa, F., & Akhtar, A. (2021). Investigating the roles of meteorological factors in COVID-19 transmission in Northern Italy. *Environmental Science and Pollution Research,* 28(35), 48459-48470. https://doi.org/10.1007/s11356-021-14038-7

[11]. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems.* https://doi.org/10.1109/TNNLS.2021.3084827

[12]. Malakar, S., Goswami, S., Ganguli, B., Chakrabarti, A., Roy, S. S., Boopathi, K., & Rangaraj, A. G. (2021). Designing a long short-term network for short-term forecasting of global horizontal irradiance. *SN Applied Sciences,* 3, 1-15. https://doi.org/10.1007/s42452-021-04421-x

[13]. Mas, J. F., & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing,* 29(3), 617-663. https://doi.org/10.1080/01431160701352154

[14]. Moroh, J. E., Innocent, D. C., Chukwuocha, U. M., Vasavada, A., Kumar, R., Siddiq, M. A., & Sah, R. (2023). Seasonal variation and geographical distribution of COVID-19 across Nigeria (March 2020–July 2021). *Vaccines,* 11(2), 298. https://doi.org/10.3390/vaccines11020298

[15]. NASA. (2020). *Prediction of Worldwide Energy Resources (POWER).* Retrieved from https://data.nasa.gov/Earth-Science/Prediction-Of-Worldwide-Energy-Resources-POWER-/wn3p-qsan/data

[16]. Prata, D. N., Rodrigues, W., & Bermejo, P. H. (2020). Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil. *Science of the Total Environment,* 729, 138862. https://doi.org/10.1016/j.scitotenv.2020.138862

[17]. Xiao, Y., & Jin, Z. (2021). The forecast research of linear regression forecast model in national economy. *Open Access Library Journal,* 8(8), 1-17. https://doi.org/10.4236/oalib.1107797

[18]. Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into imaging,* 9, 611-629. https://doi.org/10.1007/s13244-018-0639-9

[19]. Yuan, J., Wu, Y., Jing, W., Liu, J., Du, M., Wang, Y., & Liu, M. (2021). Non-linear correlation between daily new cases of COVID-19 and meteorological factors in 127 countries. *Environmental Research,* 193, 110521. https://doi.org/10.1016/j.envres.2020.110521

ABOUT THE AUTHORS

Yahaya Mohammed Sani works as a Lecturer in the Department of Information Technology at the Federal University of Technology, Minna. He is a Ph.D. Fellow in the Department of Computer Science at the School of Computing and Information Technology, Makerere University, Kampala, Uganda.

Andrew Gahwera is a Part-Time Lecturer in the School of Computing and Information Technology. He is a Ph.D. Fellow in the Department of Information System, Makerere University, Kampala, Uganda.