Techno-computing Journal 2(1) 29-40(2026) Research Paper
https://doi.org/10.71170/tecoj.2025.2.1.29-40


# Influence of Environmental Variables on COVID-19 Pandemic in the Kano State of Nigeria: An Artificial Intelligence-Based Approach

Aliyu Nuhu Salihu[1*], Hatice Bebiş[2], Mariya Nasir Danbatta[3]

*1 Department of Public health Nursing, Faculty of Nursing, Near East University, Nicosia, TRNC.*
*2 Department of Public health Nursing, Faculty of Nursing, Eastern Mediterranean University, Famagusta, TRNC.*
*3 Department of Nursing, Federal University of Health Sciences Azare*
**Corresponding Author Email: aliyunuhuans@gmail.com*

## Abstract

This study investigates the influence environmental variables on COVID-19 pandemic in Kano State, Nigeria through the artificial intelligence-based modelling methods. A descriptive correlational quantitative design was a design based on secondary data collected in a period of four months, between 1st December 2020 and 31st March 2021, with a total of 121 daily observations. The Nigerian Centre for Disease Control (NCDC) and Kano State Ministry of Health provided the COVID-19 epidemiological data (new cases, recovery cases, and death cases) whereas the Nigerian Meteorological Agency (NiMet) and AccuWeather sources were used to retrieve the environmental data (temperature, humidity, and wind speed). The methods of artificial intelligence were used in the MATLAB with three individual models Multiple Linear Regression (MLR), Artificial Neural Network (ANN), and Adaptive Neuro-Fuzzy Inference System (ANFIS). The coefficient of determination ($R^2$), correlation coefficient (R) as well as root mean square error (RMSE) were used to assess model performance. In addition, hybrid models (MLR-ANN and MLR-ANFIS) were also employed to increase predictive performance. The findings showed that the hybrid models outperformed the individual models as they showed better R2 and reduced RMSE during training and testing. The results present the possibility of hybrid artificial intelligence methods in enhancing predictive modeling of COVID-19 dynamics on environmental factors.

**Keywords:** COVID-19, Environmental Factor, Artificial Intelligence, Pandemic, Kano State

## 1. Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic has placed a significant strain on the healthcare system, especially the public health sector. It has also harmed health and the economy around the world, despite the efforts to stop and mitigate the virus' spread [1]. COVID-19 has been confirmed on all continents [2] [3]. The sudden surge of COVID-19 occurred during the winter due to cold weather, high humidity, strong winds, and heavy human traffic. These factors facilitated the rapid dissemination of the virus worldwide. The first case of COVID-19 was found in Wuhan (Hubei, Central China). The case appeared in the winter of 2019. However, in comparison to temperate and polar areas, polar areas were more affected by the virus [4]. Many studies have been published in the literature over the years that show how environmental parameters like humidity, temperature, and wind affect the transmission and survival of this infectious agent. The transmission and survival of airborne and droplet infections are dependent on virus spread in the index host and virus transfer to a secondary host. [5] [6].

(Received 02 Jan 2026; revised 03 March 2026; accepted 14 March 2026; first published online 10 April 2026);

A relationship between the COVID-19 pandemic and environmental factors has been reported in many studies. As a result, knowing how long the coronavirus can survive in the environment at various temperatures and humidity levels is crucial in understanding the viral transmission to the susceptible host [7]. The study discovered that increasing temperatures may have led to a reduction in COVID-19 transmission in 122 cities in China, concluding that this is proof to back up the idea that COVID-19 cases will indeed reduce as temperatures increase [8]. Another study conducted in Jakarta, Indonesia to measure the correlation between COVID-19 and environmental parameters (high, low, and average temperature, humidity, and rainfall) discovered that the local weather is a significant factor in predicting and determining the COVID-19 incidence rate in Jakarta. The mean temperature was found to be significantly correlated with COVID-19 [9]. Another research was carried out by Rendana, who intends to investigate the influence of windy weather on COVID-19 in Indonesia to estimate the direction of the virus's spread. The results show that COVID-19 cases are highly correlated with wind speed, implying that higher COVID-19 cases are associated with lower wind speeds [10]. On the other hand, some studies contradict the previous studies. Kassem conducted the results of this study. The result revealed no relation between COVID-19 transmission and temperature [11]. Another study suggests that the temperature effect has less influence on the COVID-19 pandemic than the cofounding factors and other environmental parameters. The influence of environmental factors on new cases and the death rate is very small; rather, the impact of travel restrictions, lockdowns, and other non-pharmacological interventions (NPIs) measures taken by countries is the most important factor affecting the infection rate [12].

A study conducted by Shuai et al., in Wuhan city discloses that the relationship between COVID-19 recoveries and temperature doesn’t exist, indicating that temperature does not affect COVID-19 recovery. The overall picture of the study indicates that COVID-19 recoveries and temperature e have a weak positive and negative correlation, which means the temperature changes may not have a significant impact on COVID-19 [13][14]. Also, Shuai et al. found that humidity harms COVID-19 recoveries, indicating that there is a significant association [14]. According to their findings, higher humidity does have a substantial influence on improving COVID-19 recoveries, which is in line with the previous survey that found an inverse correlation between higher humidity and daily cases of COVID-19 infection [14] [4]. In Italy, Coccia investigated the relationship between pollution of air and particulate matter in the air, windspeed, and COVID-19 transmission. The study uncovered two significant findings: Cities with high atmospheric windspeeds had a lower number of diseased and dead people, while cities with low atmospheric windspeeds had fewer infected people. Second, cities in hinterland zones (typically on the edges of large urban population centres) with low windspeed and higher atmospheric pollution and particle compounds have a high morbidity and mortality rate[15]. This raised concerns that the local climate, weather changes, and other factors could play a role in the spread of this virus. There is a need to investigate this matter and find the exact association between this virus and weather and other environmental parameters in an attempt to mitigate, contain, and stop the transmission of this virus [14] [16].

During the last decade, research on artificial intelligence-based technology in healthcare has increased, with technologies showing significant promise in helping and enhancing patient care holistically [17]. AI has been widely used in the public health field. AI is being deployed to track and anticipate the spread of infectious diseases. Researchers hope that by using these technologies, they will be able to prevent more widespread outbreaks in the future [18]. By combining mathematics and computers, AI is used in the public health field to learn from a wide range of data and make forecasts or predictions regarding the health of the population in the community. Healthcare organizations use AI to get the most out of their data and use it to improve disease identification, mitigation, monitoring, and elimination of diseases. Research and AI-based model applications could play a significant role in helping research, health surveillance, early detection of diseases, and, eventually, making decisions, bringing a new era of highly precise public health. Public health nurses (PHNs) can use AI-models to better understand the complicated relationships between genetics, the environment, and disease [18]. However, it is difficult to determine the complex relationships between huge amounts of data about environmental factors and COVID-19 cases, recovery, and deaths. In a single AI-mode study, Mahmoud et al. compared 3 AI-based models (1. adaptive neuro-fuzzy inference system (ANFIS), 2. artificial neural networks (ANN), and 3. multiple linear regression (MLR)) to evaluate the correlation between both the COVID-19 and environmental factors (humidity, temperature, and windspeed). The results revealed that ANN and ANFIS were more favorable than the standard MLR models, which had an average R-value of 0.90 in both the measurement and validation stages. According to the results, ANFIS surpasses the other two models, ANN and MLR [5]. However, the use of a hybrid model, which combines the two models to enhance the performance of a single AI model, was not found in the literature. This is the first time in the literature that a hybrid model of artificial intelligence has been reported for the prediction of the COVID-19 pandemic.

The need to apply AI techniques in predicting the impact of these environmental parameters for the understanding of the COVID-19 pandemic is of paramount importance. The environmental variables such as temperature, humidity and the windspeed are common environmental factors that are reported to influence the transmission dynamics and environmental persistence of respiratory viruses. Temperature has an impact on viral stability and host susceptibility, whereas humidity influences the stability of the aerosol and respiratory droplet durability. The speed of wind might be one of the contributing factors to the atmospheric dispersion and dilution of viruses. Multiple studies have found that there is a strong connection between such environmental factors and the trends of COVID-19 transmission in various parts of the world [6] [8][10]. Thus, these factors were selected as the major predictors in the current study to examine their possible influence on the dynamics of COVID-19 in Kano State, Nigeria. The study aims to see whether there is an association between COVID-19 disease and environmental variables in Kano, as well as to predict three different dependent variables: the number of new cases of COVID-19 NC, death cases DC, and recovery cases RC, using three different environmental variables as the corresponding independent variables; humidity, temperature and windspeed.

## 2. Materials and Methods

This study used a descriptive correlational quantitative design to determine the relationship between environmental factors and COVID-19 dynamics in Kano State, Nigeria. The study employed secondary data that was gathered from December 1, 2020, to March 31, 2021, on COVID-19 in Nigeria. The data consisted of 121 observations per day containing COVID-19 epidemiological variables (new cases, recovery cases, and death cases) and environmental variables (temperature, humidity, and wind speed). Data on COVID-19 were collected through the Nigeria Centre of Disease Control (NCDC) and the Kano State Ministry of Health, whereas the data regarding the environment were obtained through the Nigerian Meteorological Agency (NiMet) and through the AccuWeather database. The datasets were combined and processed to study the influence of environmental variables on the outcome of COVID-19 through artificial intelligence-based modelling techniques.

### *2.1 Instrumentation for COVID-19 PCR test for new cases and recovery cases*

PCR tests are the most accurate and reliable tests among the three (PCR, antigen, and antibody). In Kano, PCR tests were used in this study for the confirmation of COVID-19 positive or negative cases. It’s used to test any suspect person, even if that person didn’t show any symptoms. Likewise, the PCR was used to confirm the recovery of the patients. The positive result reveals that SARS CoV II RNA is present in the host body; In this study, data on new cases of COVID-19 and those who recovered during the study period was publicly released as open data. The data was obtained from PCR results and used as a secondary data source.

### *2.2 Instrumentations for Environmental data*

**Temperature**

The daily temperature has been taken as open data from the Nigerian Meteorological Agency (NIMET). The temperature was measured with a thermometer. The temperature variable was categorized into 3 levels: minimum, average, and maximum.

**Humidity**

The daily humidity rate has been taken as open data from the Nigerian Meteorological Agency (NIMET). The hygrometer was used for measuring humidity. The humidity of water vapour in the air, confined spaces, and soil is measured with a hygrometer. The "wetness and moisture" of the air around us is considered humidity.

**Wind speed**

An anemometer instrument is used for the measurement of wind pressure and wind speed. An anemometer is an important tool for meteorologists for measuring the wind situation in Kano state. The measurement of wind pressure and wind speed data for the study period were used, which were published as open data from the Nigerian Meteorological Agency (NIMET) for daily windspeed.

**Data collection method**

There are two types of data: COVID-19 data and environmental data. The time set for this research was for four months, from December 1st 2020 to March 31st 2021, in the entire Kano state, Nigeria, for COVID-19's second wave. The Ministry of Health in Kano State, Nigeria, and the Nigeria Centre for Disease Control (NCDC) provided the COVID-19 data for NC, RC, and DC. The second data is environmental data (temperature, humidity, and wind speed), generated from the AccuWeather website and the Nigerian Meteorological Agency (NIMET). All the data used was secondary.

### *2.3 Data Analysis Technique*

The data was analyzed using descriptive statistics such as mean, standard deviation SD, mode, median, minimum MIN, and maximum MAX. To assess the significance of the relationship between the factors, correlation analysis was used. These analyses were conducted using Microsoft Excel, the Statistical Package for the Social Sciences (SPSS), and MATrix LABoratory MATLAB software. MATLAB was used to run the AI models MLR, ANN, and ANFIS.

### *2.3.1 Multiple Linear Regression (MLR)*

The purpose of analysing the data in MLR, which is frequently used for prediction, is to create mathematical models that can be used to predict changes in the dependent variables based on independent variables or predictive inputs. The MLR model was used to assess the significance of the relationship between the dependent and independent variables as well as their correlation. Twenty-five percent of the data will be used for testing, while seventy-five percent of the data will be used for training. MatLab and SPSS were used to accurately run the model. The following is the Multiple Linear Regression model (1):

$$y = b0 + b1x1 + b2x2 + \cdots bixi \tag{1}$$

B0 is the Y-intercept, and X1, X2, X3, and so on are the independent variables. Xn and the dependent variables are represented by B1, B2,... The dependent variable is Y, and the independent variable coefficients are Bn, regarding the error term.

### *2.3.2 Artificial Neural Network (ANN)*

McCulloch and Pitts were fascinated by neural network systems and the brains of living organisms. They proposed the concept of an "artificial neural network" (ANN). The ANN model is a complex interconnection based on statistics and mathematics that characterizes biological neurons required for human brain functions. The multilayer perceptron (MLP) is frequently used in ANN models. The ANN is made up of three layers. 1. the input layer, where data is dispersed across the network; 2. the output layer, where data is distributed across the network; and 3. the output layer, where data is distributed across the network; 2. the data is processed in the unnoticed layer; and 3. the results for specific inputs are retrieved in the output layer. MATLAB software is used to create the ANN model. It has multiple hidden layers as well as a key factor. The ANN model has been used and widely employed in different fields of study. As a result, a training algorithm is employed to compute the error, which is defined as the variance between the measured and simulated value. The generalized equation is as follows: (2):

$$ep = yd - ya \tag{2}$$

### *2.3.3 Adaptive Neuro-Fuzzy Inference System (ANFIS)*

ANFIS has been proven to be an effective software that includes the fuzzy surgeon model's approach, demonstrating the advantages of both ANN and fuzzy logic in one system. It uses IF-THEN rules to create a mapping between input and output (also known as the Takagi–Sugeno inference model). ANFIS has recently been utilized to predict and analyze complicated datasets. ANFIS is a real-world reliable indicator due to its ability to evaluate real-world functions. Fuzzy logic converts input data into fuzzy values by using membership functions. The digit numbers range from 0 to 1. Moreover, model nodes in ANFIS act as membership functions (MFs), allowing for the simulation of input-output relationships.

The following are the rules for a first-order Sugeno fuzzy, assuming the FIS has two inputs, "X" and "Y," and one output, "F."

Rule 1: if μ(x) is A1 and μ(y) B1then f1= p1x+q1y+r1 (3)

Rule 2: if μ(x) is A2 and μ(y) is B2 then f2= p2x+q2y+r2 (4)

### *2.3.4 Model validation*

Each scientific prediction model must be validated to determine its performance over the dependent variable. Validation of the models has been carried out in different ways. coefficient of determination ($R^2$), Pearson's correlation coefficient (R), root mean square error (RMSE) and mean square error (MSE). The root mean square error (RMSE) and mean square error (MSE) represent the models' average error. Both techniques are negative-oriented values or scores with a range of 0 to 00. As a result, lower RMSE and MSE values indicate better model prediction results. R2, on the other hand, is a statistical parameter that runs from 0 to 1 and has been used to validate the model. The R2 value around 1 indicates a significant relationship between variables, whereas the R2 value near 0 indicates the least significant relationship between variables. In every regression model, to assess the significance of a factor, the standardized coefficient and the changing method of R2 are used. The following are the equations of MAE, RMSE, and R2

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left[(Y)_{obs,j}-(Y)_{com,j}\right]^2}{\sum_{j=1}^{N}\left[(Y)_{obs,j}-\overline{(Y)}_{obs,j}\right]^2} \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Y_{obsi}-Y_{comi})^2}{N}} \tag{6}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(Y_{obsi} - Y_{comi})^2 \tag{7}$$

$$R = \frac{\sum_{i=1}^{N}(Y_{obs}-\bar{Y}_{obs})(Y_{com}-\bar{Y}_{com})}{\sqrt{\sum_{i=1}^{N}(Y_{obs}-\bar{Y}_{obs})^2\sum_{i=1}^{N}(Y_{com}-\bar{Y}_{com})^2}} \tag{8}$$

## 3.0 Results and Discussion

### *3.1 Demographic Variables*

Table 1 illustrates the descriptive analysis of the entire data set used in this study for 4 consecutive months. The study comprises a total of 121 days. The descriptive statistics of the entire data set, which consists of four consecutive months. Environmental factors have been very crucial in viral survival. Many studies have shown the influence of environmental variables, ranging from humidity, temperature, precipitation, wind speed, air pollution, etc. that affect viral survival and transmission [19].

***Table 1:*** *COVID-19 and Environmental Variables Distributions of the Whole Data.*

| | New Cases | Recovery Cases | Death Cases | Min T°C | Average T°C | Max T°C | Humidity% | Windspeed Km/h |
|---|---|---|---|---|---|---|---|---|
| **Mean** | 20.3 | 17.1 | 0.4 | 19.4 | 28.7 | 38.0 | 29.4 | 14.7 |
| **SD** | 25.5 | 18.9 | 0.7 | 3.4 | 2.6 | 2.6 | 8.5 | 4.5 |
| **Median** | 14.0 | 10.0 | 0.0 | 19.0 | 28.5 | 38.0 | 27.0 | 15.9 |
| **Mode** | 1.0 | 0.0 | 0.0 | 18.0 | 28.5 | 36.0 | 42.7 | 16.6 |
| **Minimum** | 0.0 | 0.0 | 0.0 | 13.0 | 23.0 | 32.0 | 17.9 | 6.0 |
| **Maximum** | 179.0 | 80.0 | 4.0 | 29.0 | 34.5 | 43.0 | 50.4 | 23.2 |
| **Range** | 179.0 | 48.0 | 2.0 | 5.0 | 5.0 | 8.0 | 15.4 | 9.5 |
| **Sum** | 2453 | 2067 | 53 | 2348.0 | 3474.5 | 4601.0 | 3562.9 | 1777.8 |

## 3.2 Correlational Analysis

The correlational analysis of this study provides an immense explanation of the data. The results reveal the correlation analysis between all the parameters.

***Table 2:*** *COVID-19 and environmental variables relationship Analysis*

| | New Cases | Recovery | Death | Max T°C | Min T°C | Average T°C | Humidity% | Windspeed Km/h |
|---|---|---|---|---|---|---|---|---|
| **New Cases** | 1 | | | | | | | |
| **Recovery** | 0.171 | 1 | | | | | | |
| **Death** | 0.123 | 0.175 | 1 | | | | | |
| **Max T°C** | -0.310 | -0.067 | -0.230 | 1 | | | | |
| **Min T°C** | -0.355 | -0.225 | -0.178 | 0.543 | 1 | | | |
| **Average T°C** | -0.381 | -0.179 | -0.227 | 0.838 | 0.913 | 1 | | |
| **Humidity%** | 0.238 | -0.273 | -0.064 | -0.263 | -0.352 | -0.356 | 1 | |
| **Wind speed Km/h** | -0.180 | 0.158 | -0.083 | 0.053 | 0.194 | 0.152 | -0.661 | 1 |

First, the findings revealed a strong inverse relationship between mean temperature and the COVID-19 confirmed NC, implying that when temperatures are low, the cases of COVID-19 increase. Consequently, the hypothesis proposed by various scientists, such as According to Wang et al., temperature can influence COVID-19 transmission to some extent, and there could be an ideal temperature for the virus to transmit. This could explain why it initially appeared in Wuhan [20]. Abdollahi and Rahbaralam discovered that the temperature and the daily number of infections have an inverse relationship [9]. Many other health professionals stated that when the temperature is high, the COVID-19 cases will eventually decrease. This argument could be accepted, although the correlation is weak, with an R-value of-0.4. As a PHN, there is a need to educate people on how to prevent themselves from being trapped in an indoor area, especially in the winter and colder seasons when the temperature is low. Humidity and cases of COVID-19 have a highly positive relationship, which implies that the higher the humidity, the more COVID-19 cases, and the higher the cases of coronavirus. Chan et al. stated that relative humidity influences the cases of coronavirus [7]. Also, Mecenas et al. conducted a review on the effects of relative temperature and humidity on COVID-19. They stated that humidified environments seem to lower the transmission of COVID-19 [6].

Lastly, the weakest inverse correlation happens to be between windspeed and confirmed NC of COVID-19. This implies that when the windspeed is high, the number of NC will decrease. Rendena conducted a study that supported this argument by stating that windspeed is strongly associated with cases of COVID-19, implying that when the windspeed is low, the number of COVID-19 cases will increase [10]. All of these correlational analyses were discussed previously, and many researchers supported the arguments. This study also proved those previous studies, even though the correlation was very weak, and this happened because of some confounding factors that existed in the study area. In recovery, there is a negative relationship between lower temperature and COVID-19 RC, which means when the temperature goes down, the number of patients recovering from surgery increases. This contradicts the study conducted by Shuai et al. that found that temperature does not affect COVID-19 recovery [14]. However, there's an inverse correlation between the humidity and RC. This implies that the higher the humidity, the higher the recovery time for patients. Shuai et al., who stated that increased humidity has a major impact on COVID-19 recoveries, supported this finding. [14]. In death cases, temperature and DC have an inverse relationship. This means that when the temperature increases, the DC decreases. This finding is supported by the study conducted by Kifer et al., which stated that COVID-19 severity decreased when spring approached. It also points out that when winter is approaching, the disease severity and death rate may both rise dramatically [16]. Moreover, the weakest correlation appears between humidity and death cases with an R-value of -0.06, and windspeed with DC with an R-value of -0.08. Table 2.

### 3.3 Single AI-Based Models Results (MLR, ANN and ANFIS)

In this study, two non-linear regression models, ANFIS and ANN, and the multilinear regression model MLR, were used to assess and predict the COVID-19 cases by using environmental parameters. The results provide a detailed and comprehensive comparison of the best individual model results (MLR, ANN, and ANFIS) of the NC, RC, and DC for predicting the COVID-19 cases by using environmental parameters. The models' performance indicators were determined using four different performance criteria: R2, R, MSE, and RMSE. MATLAB 9.3 (R2019a) was used to simulate the models.

***Table 3:*** *Single Models Results of the MLR, ANN and ANFIS*

| | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | $R^2$ | R | RMSE | MSE | $R^2$ | R | RMSE | MSE |
| MLR-NC | 0.0722 | 0.2687 | 0.1638 | 0.0263 | 0.1311 | 0.3621 | 0.0070 | 0.0000 |
| ANN-NC | 0.2010 | 0.4483 | 0.1410 | 0.0195 | 0.4998 | 0.7070 | 0.0053 | 0.0000 |
| ANFIS-NC | 0.3932 | 0.6270 | 0.1229 | 0.0148 | 0.7673 | 0.8760 | 0.0036 | 0.0000 |
| MLR-RC | 0.1806 | 0.4250 | 0.2692 | 0.0712 | 0.2315 | 0.4812 | 0.0066 | 0.0000 |
| ANN-RC | 0.0736 | 0.2712 | 0.2568 | 0.0648 | 0.2167 | 0.4655 | 0.0067 | 0.0000 |
| ANFIS-RC | 0.7477 | 0.8647 | 0.1340 | 0.0176 | 0.7977 | 0.8932 | 0.0034 | 0.0000 |
| MLR-DC | 0.0523 | 0.2287 | 0.2151 | 0.0455 | 0.0636 | 0.2523 | 0.2151 | 0.0455 |
| ANN-DC | 0.0774 | 0.2781 | 0.2123 | 0.0443 | 0.0698 | 0.2643 | 0.2123 | 0.0443 |
| ANFIS-DC | 0.6225 | 0.7890 | 0.1358 | 0.0181 | 0.6243 | 0.7901 | 0.1358 | 0.0181 |

Table 3. provides a detailed and comprehensive comparison of the best individual model results (MLR, ANN, and ANFIS) of the NC, RC, and DC for forecasting the COVID-19 cases by using environmental parameters. When modelling the ANN, the Levenberg-Marquardt algorithm was used with 1,000 iterations, an RMSE of NC of 0.141023, RC of 0.256791, and a DC of 0.212285, with a learning rate of 0.01, and a momentum coefficient of 0.9 in Table 6. The total number of hidden layers was increased through a trial-and-error method, wherein the greatest model design was chosen. Moreover, suitable and ideal parameter selection for the ANFIS model is vital for selecting the best framework for the model based on the results of the training, validation, and testing phases. Our findings are supported by those of Mahmoud et al., who reported that a single AI-based model was used to predict the influence of the environmental parameters on the COVID-19 pandemic. According to their findings (Mahmoud et al., 2021), ANFIS outperformed MLR and ANN. Additionally, this study reveals the same results: In comparison to the other two models (MLR and ANN), the ANFIS model offers a better fit. In addition, MLR shows the weakest performance among the models.

MLR in general fails to model very complex nonlinear processes, which might be related to the fact that it uses the least-squares approach, which simulates the connection between the independent and dependent variables in a linear manner. Another explanation is that the MLR simulation results created a large number of negative values, which may have hindered the model's performance accuracy. Not surprisingly, the ANFIS model demonstrates a powerful and promising ability to solve complex data as a new non-linear system for simulation. In terms of model performance accuracy, the rank order is as follows: ANFIS > ANN > MLR. Quantitative results based on goodness of fit were used to measure the linear regression models. According to R2, the ANFIS model provides a better fit compared with the other two models (ANN and MLR), improving the efficiency of performance during testing and training by up to 3% and 4%, respectively. This means that ANFIS could be a suitable model for PHN in predicting COVID-19 NC, RC, and DC. However, this single AI-model reveals a weak performance for simulating the environmental factors and the COVID-19 pandemic. The hybrid model was proposed to boost the performance of single AI-models.

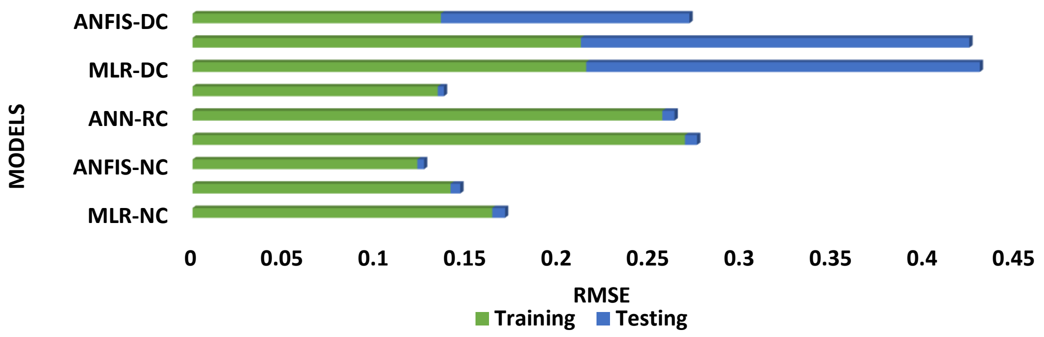


***Figure 1:*** *RMSE Results of a Single Model (MLR, ANN, and ANFIS)*

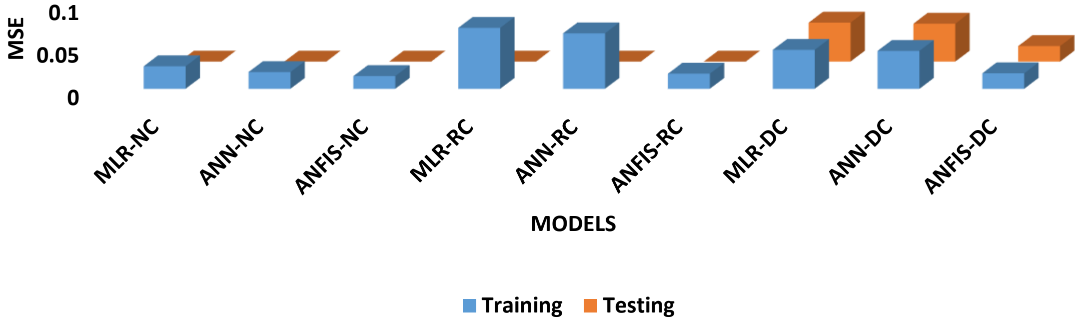


***Figure 2:*** *MSE Results of a Single Model (MLR, ANN, and ANFIS)*

Quantitative results are based on the goodness of fit to measure the linear regression. According to the results $R^2$, the ANFIS model provides a better fit compared with the other two models (ANN and MLR), improving ANN and MLR performance efficiency during testing and training by up to 3% and 4%, respectively. On the one hand, the forecast RMSE and MSE results are also displayed on the bar chart, showing the relationship between the measurement data and the simulation data (see Figures 1 and 2). The bar chart above shows the RMSE and MSE results. Based on those results, ANFIS shows the best match between the calculated and simulated values. The higher the R2 and R, the better the results. The lower the RMSE and MSE, the better the results.

## *3.4 Hybrid Models Results (MLR-ANN and MLR-ANFIS)*

When a single model does not provide the appropriate results, the hybrid model is introduced. A hybrid model is a combination of two or more clustering techniques. This is the first study in the literature in the field of public health predicting COVID-19 and environmental parameters using the hybrid model. Moreover, hybrid-learning algorithms are beginning to attract attention in a variety of fields.

***Table 4****: Hybrid Models Results of the MLR-ANN and MLR-ANFIS*

| | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | $R^2$ | R | RMSE | MSE | $R^2$ | R | RMSE | MSE |
| **MLR-ANN-NC** | 0.6945 | 0.8334 | 0.0872 | 0.0076 | 0.6783 | 0.8236 | 0.0875 | 0.2958 |
| **MLR-ANFIS-NC** | 0.8540 | 0.9241 | 0.1070 | 0.0115 | 0.8645 | 0.9298 | 0.1237 | 0.3517 |
| **MLR-ANN-RC** | 0.9583 | 0.9789 | 0.0604 | 0.0036 | 0.9467 | 0.9730 | 0.0589 | 0.2428 |
| **MLR-ANFIS-RC** | 0.9140 | 0.9560 | 0.2303 | 0.0530 | 0.9235 | 0.9610 | 0.3000 | 0.5478 |
| **MLR-ANN-DC** | 0.7402 | 0.8604 | 0.1770 | 0.0313 | 0.7352 | 0.8574 | 0.1778 | 0.4217 |
| **MLR-ANFIS-DC** | 0.9498 | 0.9746 | 0.1622 | 0.0263 | 0.9325 | 0.9656 | 0.1634 | 0.4042 |

Table 4 shows the performance indexes of the hybrid models MLR-ANN and MLR-ANFIS used in this survey. The comparative analysis shows clearly a better performance of the hybrid models as compared to the single models. Indicatively, when forecasting new cases of the COVID-19 pandemic, MLR-ANFIS hybrid model had an $R^2$ of 0.8645, which is significantly better than the top single model (ANFIS) with R2 = 0.7673. Similarly, on recovery cases, it was found that the MLR-ANN hybrid model had R2 = 0.9467, which outperformed the single ANFIS model ($R^2$ = 0.7977). In the case of death, the MLR-ANFIS hybrid model indicates higher performance with $R^2$ = 0.9325 as compared to the optimal performance of the single model (ANFIS $R^2$ = 0.6243). These findings confirm that hybrid models of artificial intelligence have greater effect on predictive accuracy than single models. This study is the first in the area of public health PH to predict the NC, RC, and DC using a hybrid model based on the COVID-19 pandemic and environmental variables. Learning algorithms are also beginning to attract attention in a variety of fields. The study results show the performance indexes of the hybrid models MLR-ANN and MLR-ANFIS. The results show that hybrid data intelligence algorithms are more effective than individual models. Based on the performance indicators in Table 4, MLR-ANFIS was found to be more efficient than MLR-ANN in NC and DC, while MLR-ANN performed more efficiently in RC than MLR-ANFIS.

This is because its RMSE and MSE scores are low, while R and R2 are also higher. Quantitative results show that the MLR-ANFIS and MLR-ANN hybrid models surpassed all the single models and improved the prediction accuracy of ANN, ANFIS, and MLR in the test phase by up to 7%, 10%, and 12%, respectively, based on goodness of fit (R2). The performance accuracy of the hybrid data intelligence algorithms model has been demonstrated by using a scatter plot. The prediction findings were presented using a graphical representation of the scatter graph to demonstrate the correlation between both the simulated and measured values of the COVID-19 NC using the hybrid models MLR-ANN and MLR-ANFIS. MLR-ANFIS outperformed MLR-ANN in the prediction of COVID-19 NC and RC. In the prediction of DC from COVID-19, MLR-ANN outperformed MLR-ANFIS. Finally, the hybrid models are further analyzed and compared by using a radar chart (MLR-ANN and MLR-ANFIS) of the NC: COVID-19 RC and DC in Kano State, Nigeria. The hybrid correlation coefficients of the two-two models are compared during the training, validation, and testing phases, and the figure shows that MLR-ANFIS outperforms the MLR-ANN in NC and DC, whereby MLR-ANN outperformed MLR-ANFIS in the RC of COVID-19.

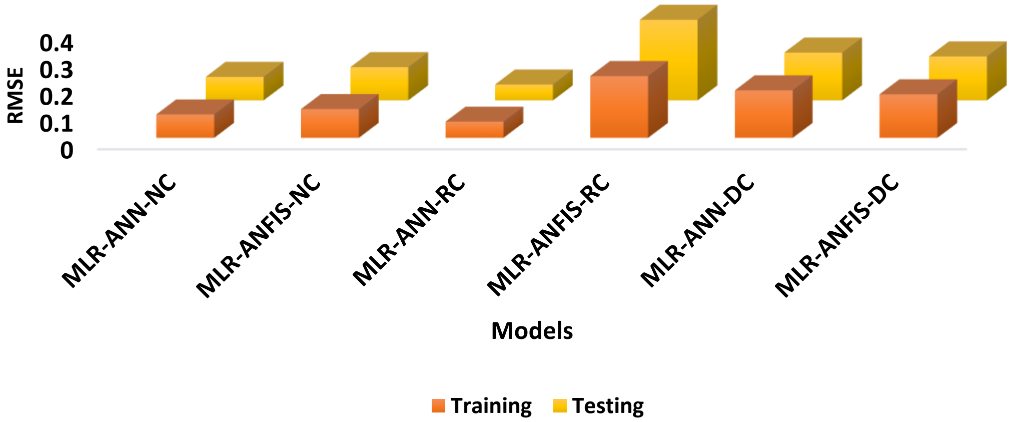


***Figure 3:*** *RMSE Results of the Hybrid Models (MLR-ANN and MLR-ANFIS)*

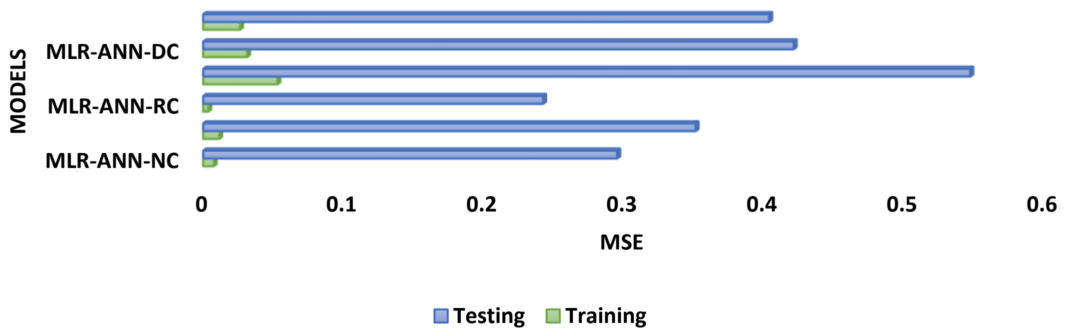


***Figure 4:*** *MSE Results of the Hybrid Models (MLR-ANN and MLR-ANFIS)*

The performance accuracy of the hybrid data intelligence algorithms model has also been demonstrated by using a scatter plot.

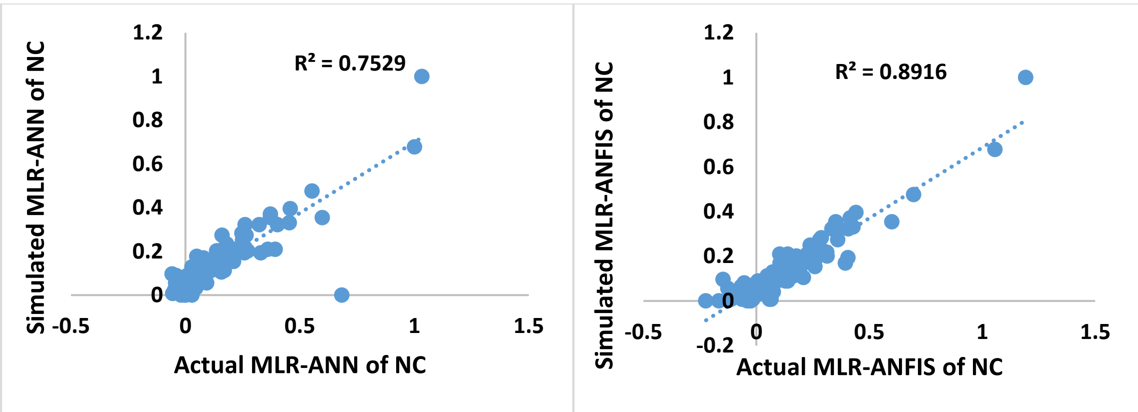


***Figure 5:*** *Scatter Plot of the Hybrid Model on New Cases of COVID-19 A. MLR-ANN B. MLR-ANFIS*

The prediction results have been demonstrated by using a diagrammatic representation of the scatter plot in Figure 5 to show the association that stands between the simulated and measured values of the new cases of COVID-19using the hybrid models MLR-ANN and MLR-ANFIS as shown in Figure 5. The MLR-ANN R2 = 0.7529 and the MLR-ANFIS R2 = 0.8916 in the prediction of NC of COVID-19.

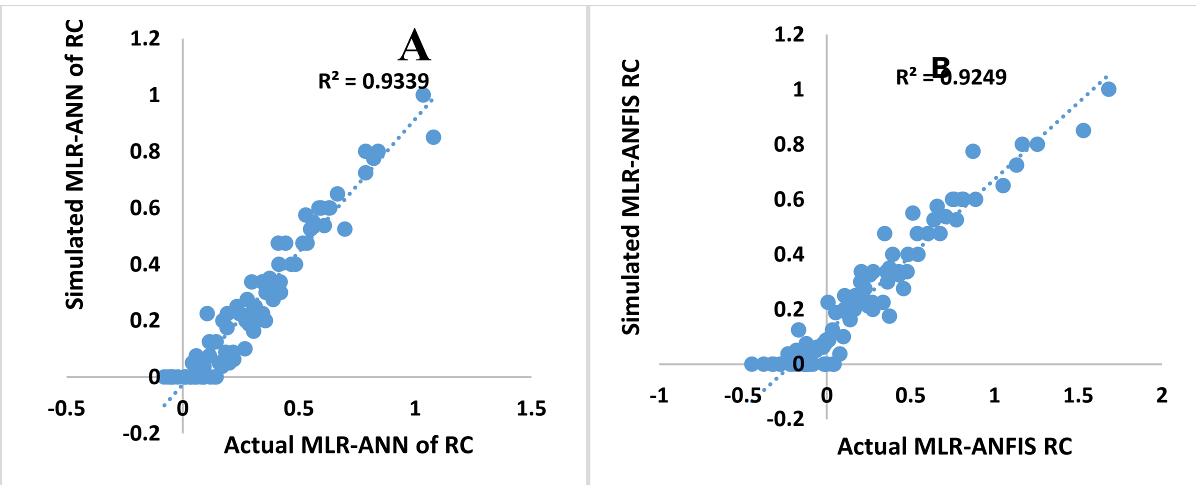


***Figure 6:*** *Scatter Plot of the Hybrid Model on Recovery Cases of COVID-19 A. MLR-ANN B. MLR-ANFIS*

The prediction results have been demonstrated by using the graphical illustration of the scatter plot in Figure 6 to show the association that stands between the simulated and measured values of the RC of COVID-19 using hybrid models MLR-ANN and MLR-ANFIS. As shown in the figure, the MLR-ANN R2 = 0.9339 and the MLR-ANFIS R2 = 0.9249 in the prediction of RC of COVID-19.

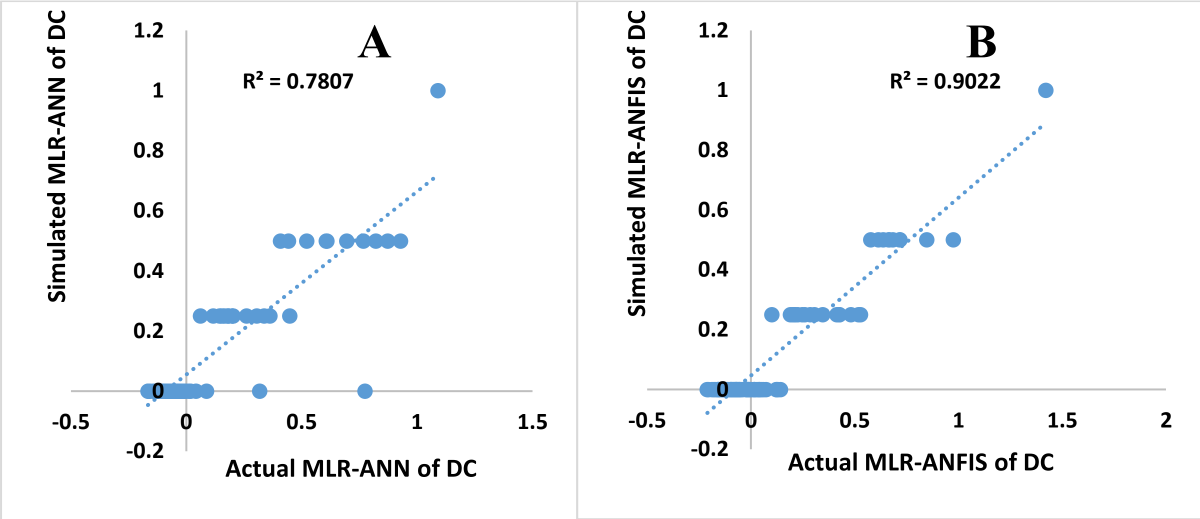


***Figure 7:*** *Scatter Plot of the Hybrid Model on Death Cases of COVID-19 A. MLR-ANN B. MLR-ANFIS*

The predicted results have been demonstrated by using the graphical illustration of the scatter plot in Figure 7 to show the association that exists between the simulated and measured values of the DC of COVID-19 using the hybrid models MLR-ANN and MLR-ANFIS as shown in the figure. The MLR-ANN R2 = 0.7807 and the MLR-ANFIS R2 = 0.9022 in the prediction of the DC of COVID-19.

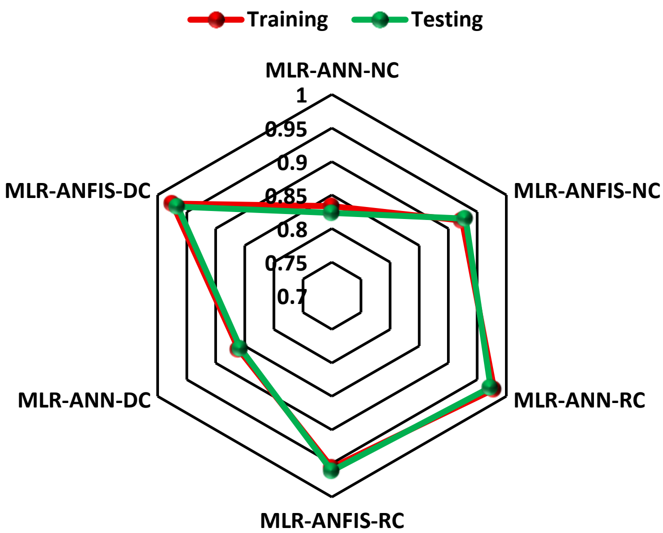


***Figure 8:*** *Radar Chart of the Hybrid Model of New Cases, Recovery and Death*

The hybrid models are further analyzed and compared by using a radar chart (MLR-ANN and MLR-ANFIS) of the NC, RC and DC that have been identified in Kano State, Nigeria. The hybrid correlation coefficients of the two-two models are compared during the training, validation, and testing phases.

## 4.0 Conclusion

According to this study, both artificial intelligence-based models inform of ANN and ANFIS and a classical linear regression model, MLR, were applied in the simulation of the environmental factors to predict three different variables, namely, NC, RC, and DC. The AI-models showed superiority over the classical MLR model, though with weak performances. Therefore, the hybrid-based models inform of MLR-ANN and MLR-ANFIS were used to boost the performance efficiency of the single AI-models in both testing and training stages. The correlation analysis results showed an inverse correlation between the temperature and cases of COVID-19. This means that as the temperature decreases, the COVID-19 cases increase. It is recommended that health education, awareness, and enlightenment be given to people by PHN on the environmental effects of COVID-19. Other environmental conditions with variations in different seasons, such as in summer, spring, and winter, should also be equally considered during the simulation of these environmental factors. Other techniques, such as ensemble machine learning, can equally be used to boost the performance efficiency of single models.

**Competing Interests:** The authors declare that they have no competing interests.

**Data Availability Statement:** The supported data associated with this researcher is available upon request from the corresponding author.

## References

[1] B. Oliveiros, L. Caramelo, N. C. Ferreira, and F. Caramelo, "Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases," *MedRxiv*, pp. 2020–03, 2020.
[2] WHO, "Compendium of WHO and other UN guidance on health and environment," 2021.
[3] O. A. Adegboye *et al.*, "Change in outbreak epicentre and its impact on the importation risks of COVID-19 progression: A modelling study," *Travel Med. Infect. Dis.*, vol. 40, p. 101988, 2021.
[4] Q. Li *et al.*, "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia," *N. Engl. J. Med.*, vol. 382, no. 13, pp. 1199–1207, 2020.
[5] K. Mahmoud *et al.*, "Prediction of the effects of environmental factors towards COVID-19 outbreak using AI-based models," *IAES Int. J. Artif. Intell.*, vol. 10, no. 1, p. 35, 2021.
[6] P. Mecenas, R. T. da R. M. Bastos, A. C. R. Vallinoto, and D. Normando, "Effects of temperature and humidity on the spread of COVID-19: A systematic review," *PLoS One*, vol. 15, no. 9, p. e0238339, 2020.
[7] K.-H. Chan, J. M. Peiris, S. Lam, L. Poon, K. Yuen, and W. H. Seto, "The effects of temperature and relative humidity on the viability of the SARS coronavirus," *Adv. Virol.*, vol. 2011, no. 1, p. 734690, 2011.
[8] J. Xie and Y. Zhu, "Association between ambient temperature and COVID-19 infection in 122 cities from China," *Sci. Total Environ.*, vol. 724, p. 138201, 2020.
[9] A. Abdollahi and M. Rahbaralam, "Effect of temperature on the transmission of COVID-19: A machine learning case study in Spain," *MedRxiv*, pp. 2020–05, 2020.
[10] M. Rendana, "Impact of the wind conditions on COVID-19 pandemic: A new insight for direction of the spread of the virus," *Urban Clim.*, vol. 34, p. 100680, 2020.
[11] A. Z. E. Kassem, "Does temperature affect COVID-19 transmission?," *Front. Public Health*, vol. 8, p. 554964, 2020.
[12] K. Chiyomaru and K. Takemoto, "Global COVID-19 transmission rate is influenced by precipitation seasonality and the speed of climate temperature warming," *MedRxiv*, pp. 2020–04, 2020.
[13] M. M. Iqbal, I. Abid, S. Hussain, N. Shahzad, M. S. Waqas, and M. J. Iqbal, "The effects of regional climatic condition on the spread of COVID-19 at global scale," *Sci. Total Environ.*, vol. 739, p. 140101, Oct. 2020, doi: 10.1016/j.scitotenv.2020.140101.
[14] Z. Shuai *et al.*, "Climate indicators and COVID-19 recovery: A case of Wuhan during the lockdown," *Environ. Dev. Sustain.*, vol. 24, no. 6, pp. 8464–8484, 2022.
[15] M. Coccia, "How high wind speed can reduce negative effects of confirmed cases and total deaths of COVID-19 infection in society," 2020.
[16] D. Kifer *et al.*, "Effects of environmental factors on severity and mortality of COVID-19," *Front. Med.*, vol. 7, p. 607786, 2021.
[17] H. von Gerich *et al.*, "Artificial Intelligence-based technologies in nursing: A scoping literature review of the evidence," *Int. J. Nurs. Stud.*, vol. 127, p. 104153, 2022.
[18] R. K. Gupta and R. Kumari, "Artificial intelligence in public health: opportunities and challenges," *JK Sci.*, vol. 19, no. 4, pp. 191–192, 2017.
[19] N. C. Peeri *et al.*, "The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned?," *Int. J. Epidemiol.*, vol. 49, no. 3, pp. 717–726, Jun. 2020, doi: 10.1093/ije/dyaa033.
[20] M. Wang *et al.*, "Temperature Significantly Change COVID-19 Transmission in 429 cities," Feb. 25, 2020, *Infectious Diseases (except HIV/AIDS)*. doi: 10.1101/2020.02.22.20025791.