RESEARCH ARTICLE   **Open Access**

Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models

Daniel Adedayo Adeyinka[1,2*] and Nazeem Muhajarine[1,3]

Abstract

Background: Accurate forecasting model for under-five mortality rate (U5MR) is essential for policy actions and planning. While studies have used traditional time series modeling techniques (e.g., autoregressive integrated moving average (ARIMA) and Holt-Winters smoothing exponential methods), their appropriateness to predict noisy and non-linear data (such as childhood mortality) has been debated. The objective of this study was to model long-term U5MR with group method of data handling (GMDH)-type artificial neural network (ANN), and compare the forecasts with the commonly used conventional statistical methods—ARIMA regression and Holt-Winters exponential smoothing models.

Methods: The historical dataset of annual U5MR in Nigeria from 1964 to 2017 was obtained from the official website of World Bank. The optimal models for each forecasting methods were used for forecasting mortality rates to 2030 (ending of Sustainable Development Goal era). The predictive performances of the three methods were evaluated, based on root mean squared errors (RMSE), root mean absolute error (RMAE) and modified Nash-Sutcliffe efficiency (NSE) coefficient. Statistically significant differences in loss function between forecasts of GMDH-type ANN model compared to each of the ARIMA and Holt-Winters models were assessed with Diebold-Mariano (DM) test and Deming regression.

(Continued on next page)

* Correspondence: daa929@usask.ca
[1]Department of Community Health and Epidemiology, College of Medicine, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada
[2]Department of Public Health, Federal Ministry of Health, Abuja, Nigeria
Full list of author information is available at the end of the article

(Continued from previous page)

Results: The modified NSE coefficient was slightly lower for Holt-Winters methods (96.7%), compared to GMDH-type ANN (99.8%) and ARIMA (99.6%). The RMSE of GMDH-type ANN (0.09) was lower than ARIMA (0.23) and Holt-Winters (2.87). Similarly, RMAE was lowest for GMDH-type ANN (0.25), compared with ARIMA (0.41) and Holt-Winters (1.20). From the DM test, the mean absolute error (MAE) was significantly lower for GMDH-type ANN, compared with ARIMA (difference = 0.11, p-value = 0.0003), and Holt-Winters model (difference = 0.62, p-value< 0.001). Based on the intercepts from Deming regression, the predictions from GMDH-type ANN were more accurate (β_0 = 0.004 ± standard error: 0.06; 95% confidence interval: − 0.113 to 0.122).

Conclusions: GMDH-type neural network performed better in predicting and forecasting of under-five mortality rates for Nigeria, compared to the ARIMA and Holt-Winters models. Therefore, GMDH-type ANN might be more suitable for data with non-linear or unknown distribution, such as childhood mortality. GMDH-type ANN increases forecasting accuracy of childhood mortalities in order to inform policy actions in Nigeria.

Keywords: Sustainable Development Goals, Time series, Under-five mortality rate, Forecasting, Artificial intelligence, Deep learning, GMDH neural network, Autoregressive integrated moving average, Holt-Winters exponential smoothing, Nigeria

Background

Childhood mortality has traditionally been used as an important health indicator for assessing population well-being and consistently gained visibility in the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) [1]. It is a major public health threat in Nigeria and other low-middle-income countries (LMICs). Despite government efforts, the high under-five mortality rate (U5MR)—100 deaths per 1000 live births in 2017 [2], continues to burden the economic and health system of Nigeria.

In the absence of reliable vital registration system of under-five mortalities in most of LMICs, it is difficult for stakeholders to track progress towards achieving the child survival targets of SDG-3, which is aimed at reducing U5MR to 25 deaths per 1000 live births by 2030. To adequately plan for child survival programmes in Nigeria, large investment is required. In the face of the current economic situation of Nigeria, accurate forecasts of childhood mortalities will guide effective use of the limited health resources. On this note, sound modeling approach to improve childhood mortality estimates is needed in Nigeria. Considering the applicability of the traditional time series models for forecasting U5MR, there is little evidence to guide future planning of child health programmes in Nigeria. The argument is that, it is challenging for researchers to choose appropriate time series modeling techniques that can detect non-linear patterns of mortality rates [3–5]. However, some authors have proposed artificial intelligence such as deep learning techniques (e.g., artificial neural networks (ANN), convolution neural networks (CNN), recurrent neural networks (RNN)) [5–8] and machine learning techniques (e.g., support vector machine, random regression forest) [9–11] to improve accuracy of predictive models, while other studies have failed to demonstrate their suitability [12–14]. Unlike the conventional statistical/ mathematical techniques such as Box-Jenkins approach of autoregressive integrated moving average (ARIMA) and Holt-Winters exponential smoothing method, ANN combines both linear and non-linear modeling properties [4, 5]. ANN closely follows the structure and functionality of the human brain and its neurons to solve complex problems faster with minimal human interventions, hence reducing error rates [6]. As ANN is evolving with newer algorithms, few studies [9–11, 15–18] have considered their applicability in population health studies. As far as we know, most of the studies in the fields of population health and medicine have used different deep learning techniques to optimize classification of health outcomes and medical data [19, 20], and disease screening/diagnosis [21, 22]. However, application of deep learning algorithm to forecast long-term childhood mortality is yet to be demonstrated in many LMICs (including Nigeria). Since childhood mortality data from resource-limited countries are often non-linear, noisy, and associated with large degree of uncertainties [2], forecasting with conventional statistical methods is somewhat difficult.

In the fields of engineering, agriculture, finance, and urban planning, group method of data handling (GMDH)—a type of artificial neural net—was observed to improve forecasting, compared with other neural networks. In a recent study, Ghazanfari et al [23] evaluated the performance of multilayer perceptron network (MLP)—a popular ANN algorithm, and GMDH-type ANN in predicting comprehensiveness and workability of concrete. Their study showed more accurate results with GMDH-type ANN. Also, other studies (outside of medicine and population health) have demonstrated the superiority of GMDH-type ANN compared with adaptive neuro fuzzy inference system (ANFIS) and long short-term memory (LSTM) [24, 25]. On this basis, our study focuses on generating accurate estimates and

observing the patterns of U5MR for Nigeria during the SDG-implementation era. This study is in line with the 2014 United Nations' call for data revolution of newer technologies that would improve data for sustainable development [26]. As new approaches are needed for child health programming in resource-limited countries like Nigeria, identifying and demonstrating the use of an appropriate model will ease application of long, time series data for monitoring the attainment of global framework indicators such as SDGs.

GMDH algorithm is a self-organizing inductive modeling and forecasting technique that extracts important information from the data to build a multilayered model through supervised learning [27]. A well-known problem with all time series methods, is that inadequately preprocessed input data can result in poor forecasting. Unlike the traditional statistical methods, no a priori knowledge of series stationarity and randomness is required for GMDH algorithm [28]. GMDH neural network can automatically learn from the data and uncover hidden processes not detectable by the conventional methods [29]. On the other hand, implementation of GMDH ANN turns out to be tricky because there is currently no theoretical guidelines for designing GMDH architectural layers in order to improve prediction accuracy [7]. Since, it is important to generate more accurate U5MR for Nigeria during the SDG-era, this study aimed to model long-term U5MR with GMDH-type ANN, and compare the forecasts with the most commonly used conventional statistical methods—ARIMA regression and Holt-Winters exponential smoothing models.

Methods

This study was exempted from ethical review by the University of Saskatchewan Behavioural Ethics Committee (ID# 904) as it relied on a publicly available aggregated de-identified dataset [30]. The dataset used is the historical aggregated yearly U5MR of Nigeria for 1964–2017 (Supplementary file 1). The dataset was obtained from the official website of the World Bank [31]. The dataset was based on the reconciled country-level estimates from different data sources by the United Nations Inter-agency Group for Child Mortality Estimation team (UN IGME) [31].

We applied ARIMA regression, Holt-Winters exponential smoothing, and GMDH neural nets to predict annual U5MR. The historical mortality data span from 1964 to 2017, giving a total of 54 observations, which was adequate to fit ARIMA regression (i.e., at least 50 non-missing data points) [32]. Furthermore, we generated long-term forecasts to determine U5MR for Nigeria by 2030 (to coincide with attainment of SDGs). GMDH-type ANN was purposefully selected from the class of deep learning algorithms because of its robustness

against incorrect, noisy, and small dataset [33]. Also, recent studies in other disciplines have demonstrated its superiority over RNN and LSTM [24, 25, 34, 35]. *P*-value < 0.05 and 95% confidence intervals (CI) were used to assess statistical significance.

Fitting ARIMA model

We utilized Stata™ version 15.1 software [36] to fit ARIMA regression model. The model construction is in four iterative steps: model identification, parameter estimation, diagnostic checking, and prediction. As the first step, data preprocessing geared towards understanding the underlying patterns in the data and data transformation was ensured. The stationarity of the aggregated U5MR was assessed by plotting line graph (Fig. 1a). It was observed that the assumption of stationarity for time series analysis was violated as evident by the non-seasonal downward trend of the overall under-five mortality rates. After different calibrations, third-order differencing was appropriate in removing the observed trend (Fig. 1b). Next, Dickey-Fuller (DF) test with drift was used to assess the stationarity of the differenced data (DF = − 9.02, lag order = 2, *p*-value< 0.001), and the absolute value of t-statistic was greater than the critical value at 5% level (9.02 vs. 1.68, p-value< 0.001).

The autocorrelation function (ACF) and partial autocorrelation (PACF) plots were also checked to determine the structure of the correlation between time lags of the differenced data (Fig. 1c and d). The ACF plot had a significant spike at the second lag, indicating second order moving averages (MA (2)) or ARIMA (0,3,2). Furthermore, ARIMA (0,3,2) model had the lowest Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC), hence it was considered suitable for fitting the actual under-five mortality rates. The model with smallest possible number of parameters (principle of parsimony) was selected to represent the distribution of the data. Also, the model comparison with AIC and BIC is valid, because the candidate models fitted the same data (U5MR) [37]. Following this principle, the specified ARIMA (0,3,2) model is expressed as:

$$\Delta_3 y_{t=} b_1 * \varepsilon_{t-1} + b_2 * \varepsilon_{t-2} \tag{1}$$

$$\Delta_3 y_t = y_t - y_{t-3} \tag{2}$$

y_t and ε_t were the actual value and random error at time period t, respectively. b_1 and b_2 where the model parameters of moving averages at lag 1 and lag 2 (− 0.35 and 0.62 respectively) with standard deviation $\delta = 0.22$.

The adequacy of the fitted model was determined by the randomness of the model residuals (Fig. 2). Also, all the eigenvalues for stability of estimates were less than one and the inverse roots of MA polynomial visually
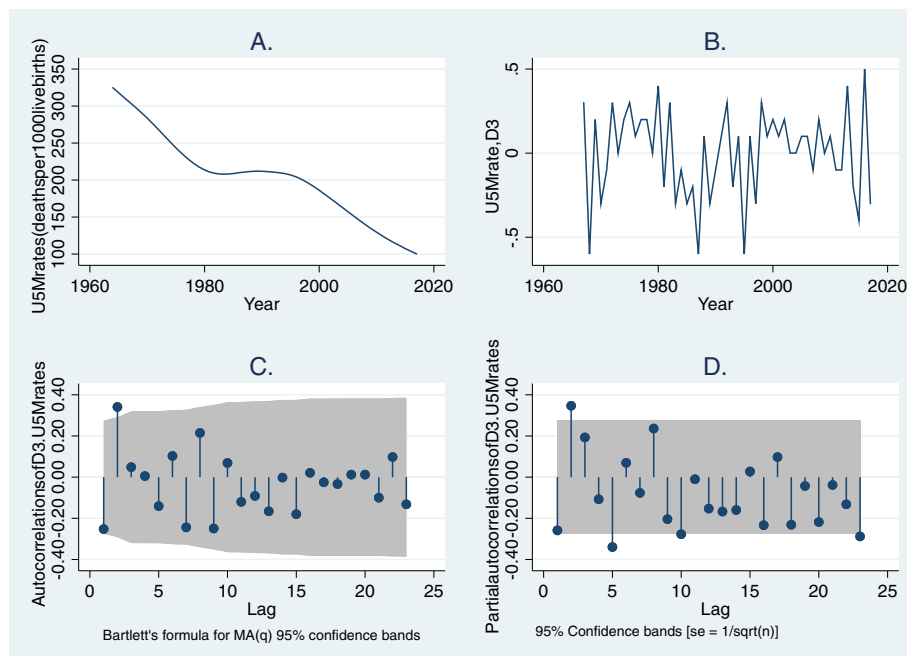
Fig. 1 (**a**) Time series plot of under-five mortality rates in Nigeria for ARIMA modeling, 1964–2017 (B) Third order difference of yearly under-five mortality rates (**c**) Autocorrelation function plot of third order differenced under-five mortality rates (**d**) Partial autocorrelation function plot of third order differenced under-five mortality rates. D3: third order differencing, U5M: under-five mortality, grey color in plots C and D: 95% Confidence Interval
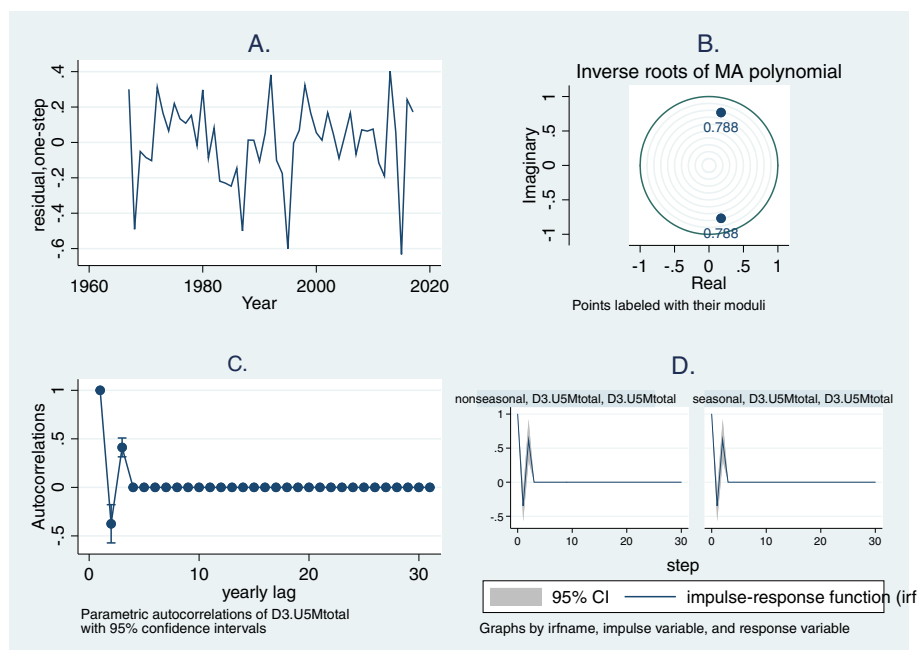

Fig. 2 Diagnostic plots for ARIMA (0,3,2) of under-five mortality rates, Nigeria (1964–2017) (**a**) Residual plot (**b**) Inverse roots of MA polynomial (**c**) Autocorrelations of differenced rates (**d**) Impulse-response function plot

indicates that the eigenvalues were within the unit circle (modulus of 0.79). This suggests that the MA parameters satisfied invertibility condition (Fig. 2b). This was further confirmed with Portmanteau (Q) test for white noise (Q = 32.3, p-value = 0.095).

Fitting Holt-Winters exponential smoothing model
Holt-Winters non-seasonal smoothing (often referred to as triple exponential smoothing) was used to predict the overall under-five mortality rates. According to Chatfield [38], it is the most advanced method in the category of smoothing methods. The smoothing parameters were automatically generated with Stata™ version 15.1 software [36] prior modeling with Holt-Winters method. The α (level) and β (slope) of trend should lie between 0 and 1, with values closer to 0 implying that the estimates at the current/future time points are based on recent observations [38]. The optimal smoothing weights were computed as α=0.91 and β=0.51. The residual plots after fitting under-five mortality rates for Nigeria, using Holt-Winters exponential model are shown in Fig. 3.

Fitting GMDH-type artificial neural network model
The time series artificial neural network was implemented by the GMDH-type neural network core algorithm in GMDH Shell DS version 3.8.9 software [39]. We used the built-in time series pre-processing features of GMDH-type algorithm [40] to automatically remove the under-five mortality trend. The target variable (U5MR) was automatically transformed into cube root, with a minimum of zero lag and maximum of 6 lags. The input variables included time and lags of the transformed mortality rates. The polynomial neuron function of GMDH-type model is as follows:

$$f\left(x_i, x_j\right) = a_o + a_1.x_i + a_2.x_j + a_3.x_i.x_j \qquad (3)$$

where $x = (x_i, x_2...),$ *the input variables vector, and A* $= (a_0, a_1, a_2, ..)$ *the vector of weights.*

To avoid under/over-fitting of U5MR arising from improper training of dataset, the network was implemented with the dataset randomly partitioned using Pareto principle [41, 42]—80% was used for training and 20% was the test dataset for evaluating model accuracy. We designed an optimal neural-type time series model based on best performing hyper parametrization with polynomial neural networks of GMDH-type [43] (Fig. 4). Following the rule of thumb that the number of hidden neurons should be less than twice the input layer size, we developed the neural architecture [44, 45]. After different calibrations of neural architecture, the parameters for the network was configured with maximum number of network layers of 60 and initial neurons of 25. Similar to a method used by Banica et al. [46], we stopped creating new neural layers when: (1) the new layer failed to improve the model accuracy, compared with preceding layer, (2) changes in testing error was less than 1%, and (3) configuration limit for the number of layers has been reached. Adequacy of the model was further confirmed with criterion value and residual plots (Fig. 4). With low criterion value of 1.01e-5, the final neural architecture adequately fits the data. The parameters and coefficients of equation for GMDH-type model are given as:

$$Y_{[t]} = 5.206e^{-05} - N_{25}*0.077 + N_2*1.07734 \qquad (4)$$

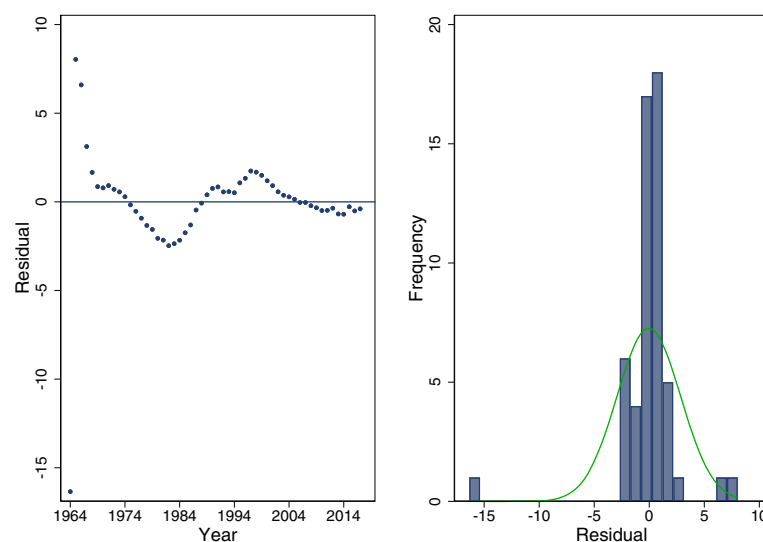*Where Y corresponds to year of forecast, and N indicates neurons 2 and 25.*

Fig. 3 Residual plots for Holt-Winters exponential smoothing for overall under-five mortality rates, Nigeria (1964–2017)
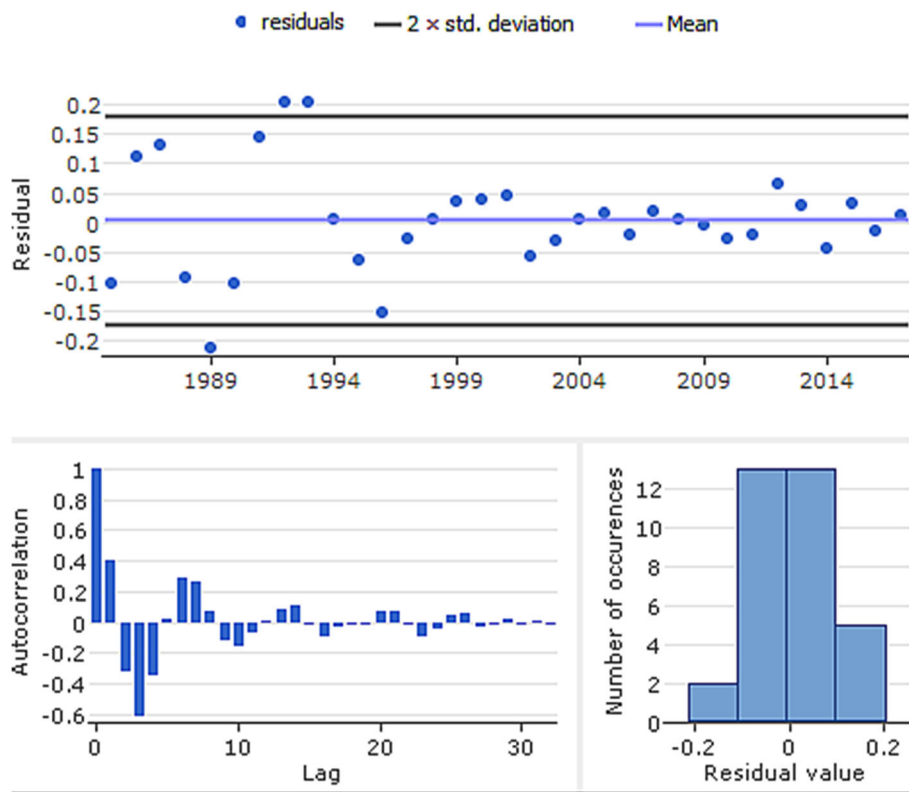
Fig. 4 Residual plots for GMDH-type neural network for overall under-five mortality rates, Nigeria (1964–2017)

Model comparison

The foremost problem with measuring prediction accuracy is the identification of key performance indicators. Although mean absolute percentage error (MAPE), mean squared error (MSE) and RMSE are the commonly used accuracy metrics, they are prone to asymmetrical distribution of errors [47]. MAPE is generally not considered as a good performance indicator because of its disadvantages— (1) only accurate for ratio-scaled data, and (2) it disfavors models when the predicted values are more than the actual (historical) values [48]. On the other hand, a benefit of using RMSE is that it is more appropriate if large errors are anticipated. Even though strength of performance measurements vary, we selected root mean absolute error (RMAE) because of its robustness against outliers [49]. In addition, RMSE was chosen because it minimizes the effects of bias, and measures dispersion of prediction errors (i.e., model stability) [49]. The model with the lowest RMAE and RMSE values provides good fit for U5MR in Nigeria.

Furthermore, modified Nash-Sutcliffe model efficiency coefficient (NSE) was calculated to address the challenges of over-estimating extreme values, arising from squared differences of actual and predicted values in original Nash-Sutcliffe efficiency equation [50]. Efficiency coefficient of ≥0.9 implies highly accurate prediction, and < 0.8 implies inaccurate prediction [51].

Relative to the observed rates, statistically significant differences in loss function between forecasts of each of ARIMA and Holt-Winters models were compared to GMDH-type ANN, based on absolute value error from Diebold-Mariano (DM) test [52]. DM test is a statistical test for comparing two competing forecasts based on loss-differentials, given the historical (observed) values. In estimating the predictive accuracy, squared error was not used because of its tendency to overestimate errors [53]. Also for the long-run variance of the differenced series from its autocovariance function, a maximum lag order of 9 was selected by Schwert criterion and the weights of Bartlett kernel (i.e., zero autoregression).

The measurements are expressed as [54]:

$$RMAE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |y_i - \widehat{y}|} \tag{5}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y})^2} \tag{6}$$

$$modified\ NSE = 1 - \frac{\Sigma|y_i - \hat{y}|^j}{\Sigma|y_i - \overline{y}|^j} \tag{7}$$

Where $\hat{y}$ = predicted value of y, $\overline{y}$
= mean value of y, $j = 1$

To further test the equivalence of the in-sample predictions (from the individual methods) with the observed historical values, Deming regression—an extension of errors-in-variables regression was performed. Deming regression assumes that forecasting errors are caused by the methods used [55, 56], (sample size > 20) [56], and the measurement error variance ratios (λ) are constant [55, 57]. When $\lambda = 1$, Deming regression is like orthogonal regression. An intercept (β_0 or constant) with a confidence interval including 0 (i.e., intercept is not significantly different from zero) suggests no systematic difference/ bias between the measurements [56]. Also, slope coefficient (β_1) with a confidence interval including 1 (i.e, slope is not significantly different from 1) indicates absence of proportional differences [56]. The Deming regression model utilized jack-knife replications to estimate the standard errors (SE) and 95%CI of the coefficients.

Results

The mean annual U5MR was 203.84 (standard deviation: 58.02) deaths per 1000 live births, ranging between 324.8 deaths per 1000 live births in 1964 and 100.2 deaths per 1000 live births in 2017. From Fig. 5a, the in-sample prediction of U5MR for 1964–2017 from ARIMA, Holt-Winters and GMDH-type ANN were close to the observed historical rates. Also, similar out-of-sample rates were observed from 2018 to 2020 for the three models (Fig. 5b). However, the out-of-sample forecasts from 2021 to 2030 for each model were different (Fig. 5b). The difference is greatest for Holt-Winters model compared to GMDH-type ANN and ARIMA regression models. The forecast obtained from GMDH-type ANN model is higher than others—85.89 (95% prediction interval (PI): 85.72–86.08) deaths per 1000 live births by 2030. Holt-Winters method generated smallest mortality rate for 2030, 51.20 (95%PI: 50.66–51.73) per 1000 live births. For 2030, ARIMA model generated a rate closer to the rates for the GMDH-type ANN model—80.17 (95%PI: 79.64–80.71) deaths per 1000 live births. According to the GMDH-type model, U5MR was observed to rise from 2028 to 2030 (Fig. 5b).

The modified NSE coefficient was slightly lower for Holt-Winters methods (96.7%), compared to GMDH-type ANN (99.8%) and ARIMA (99.6%). Further comparison between GMDH-type ANN and ARIMA with RMSE, RMAE and DM test indicated that GMDH-type ANN's performance was better (Table 1). The RMSE of GMDH-type ANN (0.09) was lower than ARIMA (0.23) and Holt-Winters (2.87). Similarly, RMAE was lowest for GMDH-type ANN (0.25), compared with ARIMA (0.41) and Holt-Winters (1.20). From the DM test, the mean absolute error (MAE) was significantly lower for GMDH-type ANN, compared with ARIMA (difference = 0.11, *p*-value = 0.0003), and Holt-Winters model (difference = 0.62, p-value< 0.001).
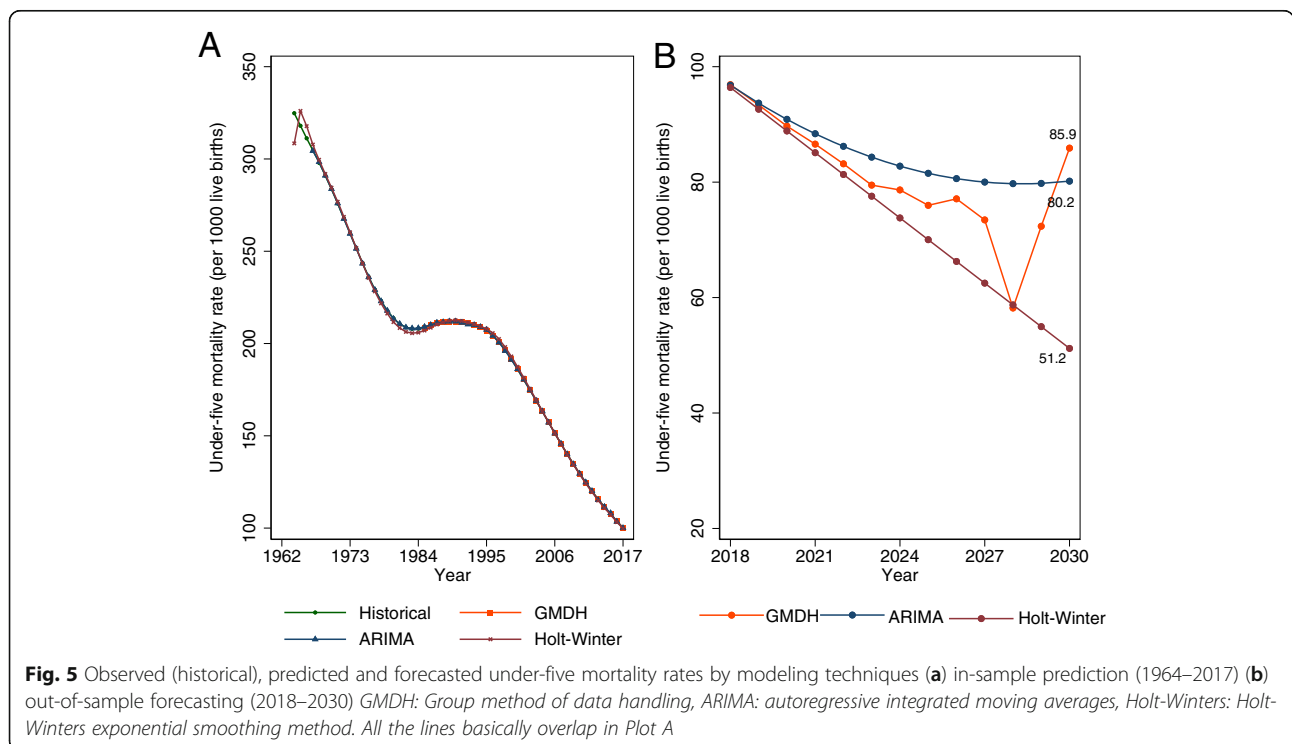

Fig. 5 Observed (historical), predicted and forecasted under-five mortality rates by modeling techniques (**a**) in-sample prediction (1964–2017) (**b**) out-of-sample forecasting (2018–2030) *GMDH: Group method of data handling, ARIMA: autoregressive integrated moving averages, Holt-Winters: Holt-Winters exponential smoothing method. All the lines basically overlap in Plot A*

Table 1 Performance measures of time series techniques for under-five mortality rates in Nigeria

Model	GMDH-type ANN	ARIMA	Holt-Winters exponential smoothing
Best parameters	Training set = 80%, testing set = 20%	$p = 0$, d = 3, q = 2	α=0.91, β=0.51
RMSE	0.09	0.23	2.87
RMAE	0.25	0.41	1.20
Modified NSE	0.998	0.996	0.967
DM test statistic	Reference	−3.608*	−4.474*

*significant at p-value < 0.05; p = number of autoregressive terms, d = number of differencing, q = number of moving average terms; RMSE: Root mean squared error; RMAE: Root mean absolute error; α=coefficient for the level smoothing; β= coefficient for the trend smoothing; modified NSE: modified Nash-Sutcliffe model efficiency coefficient; DM: Diebold-Mariano test

As shown in Table 2, the coefficients of slopes and intercepts suggest there were no proportional and systematic differences between the predicted rates for the three models and the observed (historical) rates. While the slopes (proportional difference) for the three methods were similar, the intercepts (systematic difference) and standard errors were different. The lowest coefficient of intercept and SE were obtained with GMDH-type ANN (β_0 = 0.004 ± SE: 0.06; p-value = 0.940)—implying that GMDH-type predictions were closest to the observed (historical) rates than ARIMA (β_0 = 0.03 ± SE: 0.16; p-value = 0.865), and Holt-Winters (β_0 = 0.89 ± SE: 2.35; p-value = 0.706).

Discussion

This study compared the predictive ability of artificial intelligence technique with traditional statistical methods in view of forecasting U5MR for Nigeria from 1964 to 2030. With lowest error rates of RMAE and RMSE from this comparative analysis, we demonstrated that deep learning algorithms such as GMDH-type neural nets might be more suitable for long-term forecasting of U5MR than ARIMA regression and Holt-Winters exponential smoothing methods. Similarly, Deming regression suggests more accurate prediction of U5MR with GMDH-type ANN. Using the high efficiency coefficients (> 90%) and overlapping of the predicted rates as the criteria, all three models performed well with in-sample predictions of U5MR for Nigeria. For the period from 1964 to 2017 (in-sampling prediction) and out-of-sample forecasting from 2018 to 2020, all three models had similar results, however, for the longer out-of-sample forecasting period (2021–2030), the rates were significantly different. Also, Nigeria will not achieve child survival targets of SDG by 2030. Furthermore,

GMDH-type ANN showed that U5MR will start increasing by 2028. Further analysis with age-specific mortality rates suggests that the surge in U5MR from 2028 to 2030 is due to increasing trend in neonatal mortality rates between 2028 and 2029, and child mortality rates from 2029 to 2030 (results are not shown).

In line with other studies [58, 59], our findings showed that ARIMA regression might not be suitable for long-term forecasting of U5MR, in this case for Nigeria. According to Koutsoyiannis [58, 59], ARIMA regression may not be ideal for data that exhibit long-range dependencies because of its slow decay of autocorrelation structure with lag time, making it less sensitive to tipping-points. In addition, ARIMA and Holt-Winters models assume normality of time series data, whereas under-five mortality data for Nigeria showed a non-linear trend. Also, unlike the two traditional approaches, GMDH-type ANN avoids overfitting by dropping nodes with insufficient predictive power (i.e., fully automatic structural and parametric optimization) [33]. As opposed to ARIMA and Holt-Winters methods, GMDH time series also allows for detection of recent changes in data (arising from natural behavior, policy changes and interventions), and weighs recent data more than past data during model training [29]. These detailed patterns might be easily missed by the conventional methods.

Given that more accurate results were obtained with the GMDH-type algorithm, projecting childhood mortality rates based on neural network would provide better evidence to guide prevention strategies to accelerate gains in child survival for Nigeria. A similar pattern of results was obtained from previous studies that predicted health outcomes with other artificial neural networks. Purwanto et al [60] and Zernikow et al [61]

Table 2 Deming regression for comparing GMDH-type ANN, ARIMA and Holt-Winters models

Reference: Observed historical U5MR	Proportional difference (slope)			Systematic difference (intercept)		
	β_1 (SE)	95% LCL, UCL	P-value	β_0(SE)	95% LCL, UCL	P-value
GMDH-type ANN	1.000 (0.0004)	0.999, 1.001	< 0.001	0.004 (0.058)	−0.113, 0.122	0.940
ARIMA	1.000 (0.001)	0.998, 1.002	< 0.001	0.027 (0.160)	−0.293, 0.348	0.865
Holt-Winters	1.000 (0.013)	0.969, 1.023	< 0.001	0.890 (2.349)	−3.822, 5.602	0.706

LCL Lower Confidence Limit, *UCL* Upper Confidence Limit, *SE* Jack-knife standard errors

showed that multilayer perceptron ANN was superior to linear regression for predicting infant and preterm neonatal deaths, respectively.

More generally, this study indicates that, though U5MR in Nigeria continues to decline from 100.2 deaths per 1000 live births in 2017 to 85.9 deaths per 1000 live births in 2030, Nigeria might not achieve the SDG-3 target that aims to reduce the U5MR to 25 deaths per 1000 live births by 2030 [62]. More importantly, the government of Nigeria needs policy innovations to address the observed rise in U5MR by 2028. On evidence such as indicated in this paper, the government of Nigeria should use reliable estimates to improve the design and accelerate the implementation of child health programmes in order to attain the SDG-3 targets for under-five mortality by 2030.

To our knowledge, this is the first published comparative study of ARIMA, Holt-Winters and GMDH-type neural nets on childhood mortality—in Nigeria. Also, the time series used data that covered long period of time—54 years, making the models more stable. Given the high validation accuracy (93.9%) and low RMSE (0.09) of GMDH-type ANN, there is no evidence to suggest that the observed fluctuating patterns of U5MR from 2026 to 2030 is due to overfitting. Although more datapoints are needed to generate more stable models, the forecasts from this GMDH-ANN model seem adequate because of non-seasonality of the dataset [63]. As more datapoints are available in the future, it is necessary to fine-tune the GMDH-type ANN model. As often encountered with ANN modeling, a major gap is paucity of evidence for optimization of neurons for generating ANN architecture [7]. We relied on calibrations that could give maximum predictive power. Also, further assessment of GMDH-type ANN model with leave-one-out and multiple 3-fold cross-validations showed consistent findings with Pareto principle, and no sign of under/overfitting. To determine the robustness of GMDH-type ANN, further sensitivity analyses with RNN and LSTM models were performed in Jupyter notebook for Python 3 with TensorFlow interface [64]. Using Adam optimizer to train the models at different calibrations, error rates from RNN and LSTM were higher (i.e., RMSE ranged from 14 to 20), compared with GMDH-type ANN (RMSE of 0.09). In line with our observations, many studies [24, 25, 34, 35] have also demonstrated the superiority of GMDH-type ANN over both RNN and LSTM. In addition, we observed that RNN and LSTM algorithms might be less suitable because of the few data points available for this study, coupled with the problems of gradient vanishing and gradient explosion (i.e., accumulation of large error gradients leading to unstable models). To ensure generalizability, the GMDH-type ANN model was further tested on neonatal mortality, and sex-specific mortality data for Nigeria, obtained from World Bank [31]. There was no indication to suggest under(over-fitting) of data. The unexpected rise in U5MR from 2028 to 2030 warrants further investigation. It is also important to note that it is somewhat challenging to accurately estimate data preprocessing time for time series models because they are based on trial and error approach. In addition, computational time heavily relies on computer hardware efficiency such as central processing unit (CPU) and random-access memory (RAM). To generate interventions for improving child survival programmes in Nigeria, we prioritized model accuracy over time.

Conclusions

The GMDH-type ANN predicted U5MR for Nigeria more accurately, compared to ARIMA and Holt-Winters smoothing models. Also, it does not require complicated assumptions needed for traditional time series models. The ARIMA regression and Holt-Winters methods might not be suitable for long-term forecasting of U5MR for Nigeria. Therefore, GMDH-type ANN might be more suitable for data with non-linear or unknown distribution, such as childhood mortality. GMDH-type ANN increases forecasting accuracy of childhood mortalities to inform policy actions in Nigeria and similar settings.

Authors' contributions

DAA conceived the study, analyzed and interpreted the data, and wrote the first draft of the paper. NM assisted in the design, data interpretation, and critically reviewed the manuscript. NM supervised this study. All authors read and approved the final manuscript.

Author details
[1]Department of Community Health and Epidemiology, College of Medicine, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada. [2]Department of Public Health, Federal Ministry of Health, Abuja, Nigeria. [3]Saskatchewan Population Health and Evaluation Research Unit, Saskatoon, Saskatchewan, Canada.

References

1. World Health Organization. Health in 2015: from MDGs, Millennium Development Goals to SDGs, Sustainable Development Goals [Internet]. 2015 [cited 2019 Mar 2]. Available from: https://apps.who.int/iris/bitstream/handle/10665/200009/9789241565110_eng.pdf;jsessionid=EF41ECFD78C867C3DA33E6DC9D133CC6?sequence=1.
2. UNICEF. Levels and Trends in Child Mortality [Internet]. 2018 [cited 2019 Mar 14]. Available from: https://data.unicef.org/wp-content/uploads/2018/09/UN-IGME-Child-Mortality-Report-2018.pdf.
3. Zhang G, Eddy Patuwo BY, Hu M. Forecasting with artificial neural networks: The state of the art. Int J Forecast. 1998;14(1):35–62.
4. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing. 2003;50:159–75.
5. Al-Maqaleh BM, Al-Mansoub AA, Al-Badani FN. Forecasting using artificial neural network and statistics models. Educ Manag Eng. 2016;3:20–32.
6. Kihoro J, Otieno R, Wafula C. Seasonal time series forecasting: a comparative study of Arima and ann models. African J Sci Technol. 2006;5(2):41–9.
7. Aladag CH. A new architecture selection method based on tabu search for artificial neural networks. Expert Syst Appl. 2011;38(4):3287–93.
8. Shi L, Wang XC, Wang YS, Shi L, Wang XC, Wang YS. Artificial neural network models for predicting 1-year mortality in elderly patients with intertrochanteric fractures in China. Brazilian J Med Biol Res. 2013;46(11):993–9.
9. Hsieh MH, Hsieh MJ, Chen CM, Hsieh CC, Chao CM, Lai CC. Comparison of machine learning models for the prediction of mortality of patients with unplanned extubation in intensive care units. Sci Rep. 2018;8(1):1–7.
10. Sakr S, Elshawi R, Ahmed AM, Qureshi WT, Brawner CA, Keteyian SJ, et al. Comparison of machine learning techniques to predict all-cause mortality using fitness data: the Henry Ford exercise testing (FIT) project. BMC Med Inform Decis Mak. 2017;17(1):174.
11. Son YJ, Kim HG, Kim EH, Choi S, Lee SK. Application of support vector machine for prediction of medication adherence in heart failure patients. Healthc Inform Res. 2010;16(4):253–9.
12. Naim I, Mahara T. Comparative analysis of Univariate forecasting techniques for industrial natural gas consumption. Image, Graph Signal Process. 2018;5:33–44.
13. Vochozka M. Practical comparison of results of statistic regression analysis and neural network regression analysis. Littera Scr. 2016;9(2):156–68.
14. Verplancke T, Van Looy S, Benoit D, Vansteelandt S, Depuydt P, De Turck F, et al. Support vector machine versus logistic regression modeling for prediction of hospital mortality in critically ill patients with haematological malignancies. BMC Med Inform Decis Mak. 2008;8(1):56.
15. Parsaeian M, Mohammad K, Mahmoudi M, Zeraati H. Comparison of logistic regression and artificial neural network in low Back pain prediction: second National Health Survey. Iran J Public Health. 2012;41(6):86–92.
16. Khan MT, Kaushik AC, Ji L, Malik SI, Ali S, Wei D-Q. Artificial neural networks for prediction of tuberculosis disease. Front Microbiol. 2019;10:395.
17. Blagojević M, Papić M, Vujičić M, Šućurović M. Artificial Neural Network Model for Predicting Air Pollution. Case Study of the Moravica District, Serbia. Environ Prot Eng. 2018;44(1):129–39.
18. Oustimov A, Vu V. Artificial neural networks in the Cancer genomics frontier. Transl Cancer Res. 2014;3(3):191–201.
19. Raut RD, Dudul S V. Arrhythmias classification with MLP neural network and statistical analysis. In: Proceedings - 1st International Conference on Emerging Trends in Engineering and Technology. 2008. p. 553–558.
20. Abdel-Aal RE. GMDH-based feature ranking and selection for improved classification of medical data. J Biomed Inform. 2005;38(6):456–68.
21. Kondo T, Pandya A, Zurada JM. GMDH-type neural networks and their application to the medical image recognition of the lungs, Proceedings of the SICE Annual Conference International Session Papers (IEEE Cat No99TH8456), Morioka, Japan; 1999. p. 1181–6.
22. Usher D, Dumskyj M, Himaga M, Williamson TH, Nussey S, Boyce J. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabet Med. 2004;21(1):84–90.
23. Ghazanfari N, Gholami S, Emad A, Shekarchi M. Evaluation of GMDH and MLP networks for prediction of compressive strength and workability of concrete. Bull la Société R des Sci Liège. 2017;86:855–68.
24. Do QH, Yen TTH. Predicting primary commodity prices in the international market: an application of group method of data handling neural network. J Manag Inf Decis Sci. 2019;4:471–82.
25. Lopes MNG, Da Rocha BRP, Vieira AC, De Sá JAS, Rolim PAM, Da Silva AG. Artificial neural networks approaches for predicting the potential for hydropower generation: a case study for Amazon region. J Intell Fuzzy Syst. 2019;36(6):5757–72.
26. United Nations Data Revolution Group. A world that counts: Mobilizing the data revolution for sustainable development. 2014. Available from: www.undatarevolution.org. [cited 2020 Feb 25].
27. Farlow SJ. The GMDH algorithm of Ivakhnenko. Am Stat. 1981;35(4):210–5.
28. Onwubolu G. GMDH-Methodology and Implementation in MATLAB GMDH-methodology and implementation in MATLAB; 2016. p. 284.
29. Anastasakis L, Mort N. The development of self-organization techniques in modelling: a review of the group method of data handling (GMDH). 2001. Available from: https://gmdhsoftware.com/GMDH_Anastasakis_and_Mort_2001.pdf . [cited 2020 Mar 6].
30. Government of Canada. Interagency Advisory Panel on Research Ethics. 2018. Available from: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter2-chapitre2/#ch2_en_a2.4. [cited 2018 Nov 25].
31. UN Inter-agency Group for Child Estimation. Data bank [Internet]. The World Bank. 2019 [cited 2019 Jul 20]. Available from: https://data.worldbank.org/indicator/SH.DYN.MORT?end=2017&start=1968&type=shaded&view=map.
32. Box GEP, Tiao GC. Intervention analysis with applications to economic and environmental problems. J Am Stat Assoc. 1975;70(349):70–9.
33. Pradeepkumar D, Ravi V. Forecasting financial time series volatility using particle swarm optimization trained Quantile regression neural network. Appl Soft Comput J. 2017;58:35–52.
34. Kim WJ, Jung G, Choi SY. Forecasting CDS term structure based on Nelson-Siegel model and machine learning. Complex Econ Bus. 2020;2020:1–23.
35. Stefenon SF, Dal Molin Ribeiro MH, Nied A, Mariani VC, Coelho dos LS, Menegat da Rocha DF, et al. Wavelet group method of data handling for fault prediction in electrical power insulators. Int J Electr Power Energy Syst. 2020;123:106269.
36. Stata version 15.1 [Internet]. 2017 [cited 2019 May 30]. Available from: https://www.stata.com/order/.
37. Kuha J. AIC and BIC: comparisons of assumptions and performance. Sociol Methods Res. 2004;33(2):188–229.
38. Chatfield C. The Holt-winters forecasting procedure. Appl Stat. 1978;27(3):264–79.
39. GMDH L. GMDH Shell for Data Science. 2019. Available from: https://gmdhsoftware.com/signup-ds. [cited 2019 Jul 20].
40. GMDH Shell. Preprocess: GMDH Shell Documentation. 2017. Available from: https://gmdhsoftware.com/docs/preprocess. [cited 2020 Sep 13].
41. Macek K. Pareto principle in Datamining: an above-average fencing algorithm. Acta Polytech. 2008;48(6):55–9.

42. Allen DE, Hooper VJ. Generalized correlation measures of causality and forecasts of the VIX using non-linear models. Sustainability. 2018;10(2695): 132–46.
43. GMDH Shell. Solver [Internet]. 2017 [cited 2019 Sep 21]. Available from: https://gmdhsoftware.com/docs/solver#core_algorithm.
44. Berry MJ, Linoff GS. Data Mining Techniques. New York: Wiley; 1997.
45. Xu S, Cheng L. A Novel Approach for Determining the Optimal Number of Hidden Layer Neurons for FNN's and Its Application in Data Mining, 5th International Conference on Information Technology and Applications; 2008. p. 683–6.
46. Banica L, Pirvu D, Hagiu A. Intelligent financial forecasting, the key for a successful management. Int J Acad Res Accounting, Financ Manag Sci. 2012;2(3):192–206.
47. Armstrong JS, Collopy F. Error measures for generalizing about forecasting methods: empirical comparisons. Int J Forecast. 1992;8(1):69–80.
48. Makridakis S. Accuracy measures: theoretical and practical concerns. Int J Forecast. 1993;9(4):527–9.
49. Vandeput N. Data science for supply chain forecast; 2018. p. 223.
50. Phogat V, Ma S, Jw C, Simunek J. Statistical assessment of a numerical model simulating agro hydro-chemical processes in soil under drip Fertigated mandarin tree. Irrig Drain Sys Eng. 2016;5(1):1–9.
51. Tapak L, Rahmani A, Moghimbeigi A. Prediction the groundwater level of Hamadan-Bahar plain, west of Iran using support vector machines. J Res Health Sci. 2014;14(1):81–6.
52. Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat. 1995;13(3):253–63.
53. Pavlicek J, Kristoufek L. Nowcasting unemployment rates with google searches: Evidence from the Visegrad Group countries. PLoS One. 2015; 10(5):e0127084.
54. Pai P-F, Lin C-S. A hybrid ARIMA and support vector machines model in stock price forecasting. Omega. 2005;33:497–505.
55. Martin RF. General Deming regression for estimating systematic Bias and its confidence interval in method-comparison studies. Clin Chem. 2000;46(1):100–4.
56. Francq BG, Govaerts BB. Measurement methods comparison with errors-in-variables regressions. From horizontal to vertical OLS regression, review and new perspectives. Chemom Intell Lab Syst. 2014;134(15):123–39.
57. Sárbu C, Liteanu V, Bâldea M. Evaluation and validation of analytical methods by regression analysis in: reviews in analytical chemistry. Rev Anal Chem. 2000;19(6):467–88.
58. Koutsoyiannis D, Yao H, Georgakakos A. Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods. Hydrol Sci J. 2008;53(1):142–64.
59. Koutsoyiannis D. Uncertainty, entropy, scaling and hydrological stochastics. Hydrol Sci J. 2005;50(3):381–404.
60. Purwanto D, Eswaran C, Logeswaran R. A Comparison of ARIMA, neural network and linear regression models for the prediction of infant mortality rate, 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation; 2010. p. 34–9.
61. Zernikow B, Holtmannspoetter K, Michel E, Pielemeier W, Hornschuh F, Westermann A, et al. Artificial neural network for risk assessment in preterm neonates. Arch Dis Child Fetal Neonatal Ed. 1998;79(2):F129–34.
62. United Nations Economic and Social Council. Progress towards the Sustainable Development Goals [Internet]. 2018 [cited 2019 Mar 3]. Available from: http://unstats.un.org/sdgs.
63. GMDH Shell. Preparing your data: GMDH Shell Documentation [Internet]. 2017 [cited 2020 Sep 13]. Available from: https://gmdhsoftware.com/docs/setting_data.
64. Project Jupyter. Jupyter notebook [Internet]. 2020 [cited 2020 Oct 17]. Available from: https://jupyter.org/.

Publisher's Note