Water Supply

Water Supply Vol 22 No 9, 7172 doi: 10.2166/ws.2022.290

Water provision planning on the basis of human population growth forecasts: A case study of the City of uMhlathuze, KwaZulu-Natal Province, South Africa

Thato Julius Mokoma [a,*] and Surafel Luleseged Tilahun [b]

[a] Department of Economics, Faculty of Commerce, Administration, and Law, University of Zululand, Richards Bay, KwaZulu-Natal Province, South Africa
[b] Department of Mathematics, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia
*Corresponding author. E-mail: mokomatj@gmail.com

TJM, 0000-0002-6148-4323

ABSTRACT

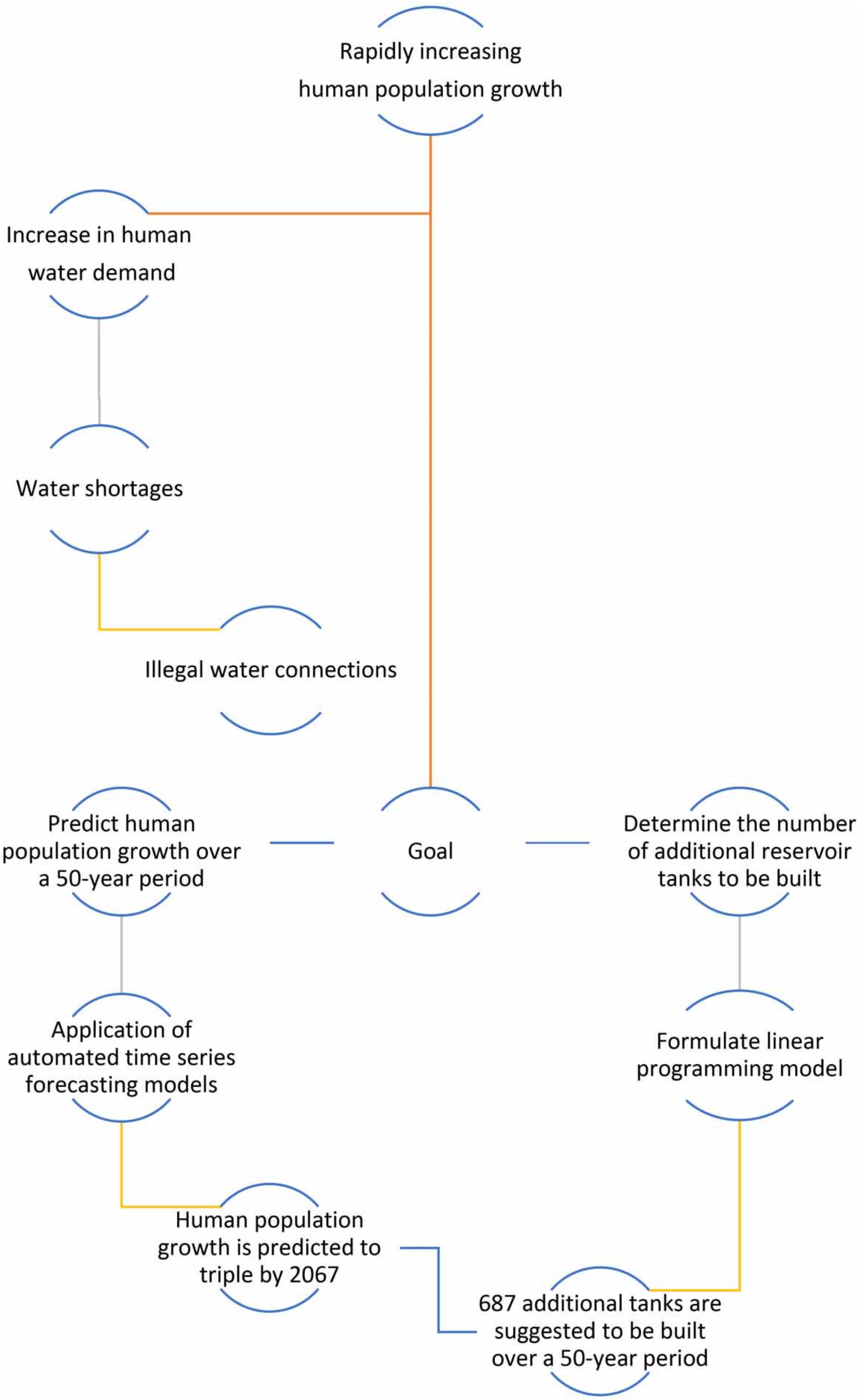
Access to adequate water is a battle the City of uMhlathuze faces daily, with the disadvantaged and impoverished areas being most affected as a result of a rapidly increasing growth in human population. The future water condition will be impossible to manage unless the municipality is able to address the increased demand for water which often leads to illegal water connections. The goal of this study is twofold: first to predict human population growth over a 50-year period and secondly, to formulate a linear programming model to determine the number of additional reservoir tanks which need to be built, in order to satisfy the human water demand for the next half century. This study compares the application of automated time series forecasting models like ARIMA, ETS, and BATS model to predict the growth in human population. The best model to forecast human population growth was then selected using the forecasting assessment criterion of RMSE, MAE, MASE and MAPE. Based on the forecast results obtained in this study, ARIMA(1,1,0) with drift model has shown a better prediction accuracy in terms of the RMSE = 3.725, MAE = 2.265, MASE = 0.211, and MAPE = 0.757, than the ETS(A,Ad,N), and BATS(1, {0,0}, 1,-) model to explain the observed values of a time series. The forecast results derived from ARIMA(1,1,0) with drift model have indicated that the human population growth will triple by 2067, albeit at a higher pace than at any period since 1996. On the basis of human population growth forecasts, it is suggested that 687 additional reservoir tanks with a capacity of 47,500 kilolitres are built within the next 50 years in order to supplement the increased demand for water that comes with an increase in human population. An implication of these findings is the possibility that, should anticipated human population growth materialise, the current reservoir tanks are likely to come under stress due to the increased demand for water.

Key words: automated time series forecasting models, human population growth, linear programming model, reservoir tanks, water supply and distribution

HIGHLIGHTS

- Water provision planning.
- Application of automated time series forecasting models to predict the growth in human population.
- Formulation of a new linear programming model along with its solution to determine the number of additional reservoir tanks to be built for half a century in advance.

GRAPHICAL ABSTRACT



INTRODUCTION

The most critical municipal service that rural and urban communities require is a reliable and adequate supply of water (Dighade *et al.* 2014; Javadinejad *et al.* 2019). In South Africa (SA), water boards, district and local municipalities have the sole mandate, as water services authorities, to supply and distribute the minimum required amount of potable water in their areas of jurisdiction. According to the Constitution of the Republic of SA, every citizen is entitled to a certain amount of water regardless of their ability to pay for it (South African Government 1996).

SA is currently facing a serious water distribution challenge as municipalities across the country are finding it hard to meet increased demands for water, largely driven by an increase in human population growth in both rural and urban areas (Department of Water and Sanitation 2018). As a result, the average water available per day per person becomes limited because, as the human population increases, consumption patterns change, economies develop and pressures on natural water resources increase which in turn affects water distribution in rural and urban areas (Sonaje & Joshi 2015; Oki & Quiocho 2020; Loubser *et al.* 2021).

The human population in King Cetshwayo District Municipality (KCDM), particularly in the City of uMhlathuze, has been continuously growing ever since the first Census data collection in the year 1996, followed by Censuses collected in 2001, 2011 and the Community Survey (CS) conducted in 2016 (Statistics South Africa 2016). In just over five years between 2011 and 2016, the human population growth increased by over 22%. With the City of uMhlathuze being one of the fastest developing cities in SA and the third economic and tourism hub in the KwaZulu-Natal (KZN) province, a rapid increase in human population is probable in the coming years, largely to be influenced by its growing industrial development zone (IDZ), adjacent to the Richards Bay harbour (Aurecon South Africa 2015a, 2015b; City of uMhlathuze 2020).

The main economic sectors in the City of uMhlathuze include manufacturing (45.9%), mining and quarrying (11.6%), financial, real estate and business (10.7%), community, social and personal services (10.4%), transport and communication (9.1%), trade (6.3%), as well as agriculture, forestry, and fishing (3.2%) (KZN Provincial Government Communication 2022). Large industrial companies in the municipality include Mondi Richards Bay, South 32/BHP Billiton, Tronox, Fairbreeze mine, Foskor, Hillside and Bayside Aluminium and the Richards Bay Coal Terminal (RBCT).

Other significant industrial companies within the municipality include the Tongaat-Hulett sugar mill, uMfolozi sugar mill and Mpact (previously known as Mondi Felixton), both in Felixton. In addition to the industrial development in the City of uMhlathuze, there are large-scale agricultural activities, which consist principally of citrus and sugarcane, both irrigated, and commercial forestry or plantations owned mostly by Sappi and Mondi.

Water is observed to be available in the reservoir tanks at the City of uMhlathuze; however, it is not sufficient for the demand, as the number of people requiring water is greater than what the reservoir tanks can supply, especially in rural areas (Zulu 2017). The findings of a research conducted by Mthethwa (2018) has also indicated that the City of uMhlathuze is grappling with poor and inadequate water infrastructure. In a case where the demand for water is human population-driven, high population growth levels put pressure on the amount of water available, leading to per capita water shortages (Mohammed & Sahabo 2015; Parks *et al.* 2019; Mnisi 2020; Mulwa *et al.* 2021). Consequently, there is a great demand on water sources in the City of uMhlathuze, especially on the uMhlathuze River, due to water demands from Empangeni and Richards Bay (City of uMhlathuze 2021).

The findings of the studies conducted by Kassa (2017), Parks *et al.* (2019), Oki & Quiocho (2020), Faye (2021), Heidari *et al.* (2021) and Mulwa *et al.* (2021) have indicated that an increase in human population growth also puts pressure on existing water infrastructure that was built decades ago. As a result, water does not reach users due to burst pipes and leaks in the water distribution network (Gunawan *et al.* 2017). SA's water distribution infrastructure is estimated to be 39 years old because it was mostly built during the apartheid regime (Department of Water and Sanitation 2017; South African Institution of Civil Engineering 2017). Not only has the infrastructure aged, but there has been further deterioration as a result of poor maintenance and operation which leads to increased water demand.

It is important to note that water demand will be impossible to manage in the future unless municipalities are able to address the present water distribution challenges; including water security, water demand management, water conservation, equity, water efficiency and sustainable consumption (Department of Water and Sanitation 2018; Development Bank of Southern Africa 2022). Increased water demand leads to water shortages for a few days and the net results are often communities seeking alternatives through illegal water connections (Karuaihe *et al.* 2016). Consequently, illegal water connections affect the pressure or volume supply of water in pipelines, either to or from the public communal standpipe or reservoir tank (Water Integrity Network 2020).

To this end, it is assumed that the demand of water in the City of uMhlathuze will continue to increase as the human population increases. Therefore, greater efforts are needed to ensure that water is adequate for every household in both rural and urban areas. For these reasons, the goal of this study is twofold: first to predict human population growth over a 50-year period and secondly, on the basis of human population growth forecasts, the number of additional reservoir tanks to be used to meet a possible increase in water demand will then be determined.

Prediction of human population growth is important to plan service delivery accordingly, that is to estimate basic human needs such as demand for food, water supply and distribution infrastructure, power supply, and transportation amongst other things (Gulseven 2016). Moreover, it plays an important role in decision making for the socio-economic and demographic development of the region.

Most research on human population prediction involves the application of traditional demographic and statistical models like the Exponential growth model, Verhulst Logistic growth model, Grey prediction model of population growth based on a logistic approach, Malthusian growth model, linear regression and non-linear regression models (e.g., Gompertz growth models), and Box-Jenkins time series models to name a few, which often need to be adjusted (Ofori *et al.* 2013; Kulkarni *et al.* 2014; Omony 2014; Wei *et al.* 2015; Patel & Prajapati 2016; Zabadi *et al.* 2017; Mondol *et al.* 2018; Stephano & Jung 2019; Tong *et al.* 2020; Wang *et al.* 2021).

However, this study proposes the application of machine learning-inspired automated time series models like the Auto-Regressive Integrated Moving Average (ARIMA) model, Error Trend Seasonal (ETS) model, and Box-Cox Transformations, ARMA errors, Trend and Seasonal components (BATS) model, to predict the growth in human population over a 50-year period. The proposed time series forecasting models are automated because they do not need any human judgment and therefore adjust automatically (Anvari *et al.* 2016; Hyndman & Athanasopoulos 2018).

The originality of this study therefore incorporates an important dimension, such as formulating our own linear programming model to determine the provision of reservoir tanks to satisfy human water demand for the next 50 years. The formulation of a new linear programming model will be useful in terms of planning for building new water infrastructure in the future, based on the prediction of human population growth and its direction of expansion so that a basic water consumption requirement of about 25 litres per person per day within a 200 metres (m) radius from the public standpipe is achieved. It is considered a minimum; therefore, it is not considered adequate for a full, healthy, and productive life (Department of Water and Sanitation 2018).

MATERIALS AND METHODS

Description of the study area

The City of uMhlathuze is an administrative area in KCDM, located on the north-east coast of the KZN province in SA, which is about 180 kilometres north-east of Durban (City of uMhlathuze 2016). The municipality was named after the uMhlathuze river, which meanders through the municipal area and symbolically unites the towns, suburbs, and traditional areas (City of uMhlathuze 2019). uMhlathuze local municipality assumed city status on the 5th of December 2000 after the demarcation process and the local government elections of that date. The City of uMhlathuze's boundary was further enlarged after the South African local government elections in 2016 when Ntambanana local municipality was merged into it (City of uMhlathuze 2021). The re-determination of the municipal boundary affected its geographic setting and now the municipality is divided into 34 municipal wards as shown in Figure 1.

The municipality's total land area covers 795,971 km² and incorporates Richards Bay, Empangeni, eSikhaleni, Ngwelezane, eNseleni, and Vulindlela (City of uMhlathuze 2020). Climatic conditions in the municipality are characterised by a warm to hot and humid subtropical climate, with warm, moist winters (City of uMhlathuze 2020). Average daily maximum temperatures range from 29 °C in January to 23 °C in July, and extreme temperatures can reach more than 40 °C in summer. The average annual rainfall is 1228 millimetres (mm), and 80% of the rainfall occurs between spring and summer, that is from September to February (City of uMhlathuze 2021). The remaining 20% occurs between autumn and winter, that is from March to August.

Description of data

Human population data (secondary) of the City of uMhlathuze for the first Census data collection in the year 1996, followed by Censuses collected in 2001, 2011 and the Community Survey (CS) conducted in 2016 was obtained from Statistics SA

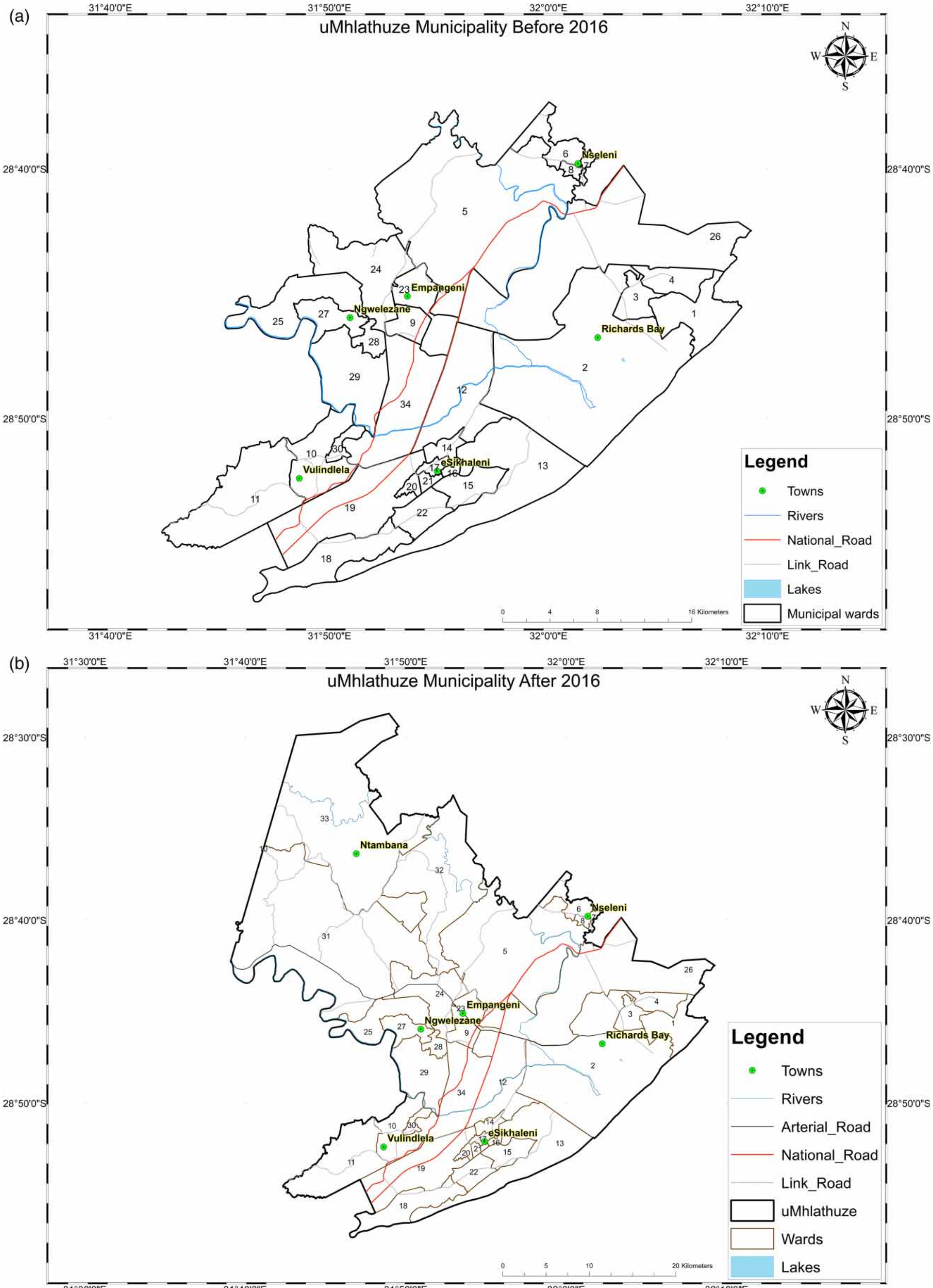


Figure 1 | Map showing the geographical location of the City of uMhlathuze: (a) Municipal wards before 2016 local government elections; (b) enlarged municipal area following the inclusion of three wards (31, 32, and 33) from the former Ntambanana local municipality post the 2016 local government elections.

(StatsSA) as indicated in Table 1. The polygon features (in the form of shapefiles) used to draw the map of the City of uMhlathuze (municipal boundary, municipal wards, river lines or lakes, roads, and towns) were obtained from the Human Sciences Research Council (HSRC). Analysis of data was carried out using RStudio software package version 3.4.4 (R Core Team 2020), to interpolate data as well as to predict human population growth, while Microsoft Excel Solver was used to solve the linear programming model which was used to determine the number of additional reservoir tanks to be built in future. Finally, ArcGIS version 10.5 was used to draw the graphical location of the City of uMhlathuze.

Identifying gaps in the dataset

Time series data is often recorded and analysed to predict future values, understand phenomena and to understand the behaviour of variables, among other things (Lepot *et al.* 2017). However, for several reasons, the data contains missing values, gaps, irregular time steps of recordings, or removed data points that often need to be filled for data analysis. Census human population data of the City of uMhlathuze contains gaps because the first Census data collection in SA was in 1996, followed by Censuses conducted in 2001, and 2011, and the CS conducted in 2016. Censuses are not annually implemented; therefore, estimates are needed for the intercensal period (Fukuda 2010).

Interpolation method

In order to fill missing values or gaps in time series data for the City of uMhlathuze, and to predict future values, a linear interpolation method was used to fill gaps in the time series data. The method is commonly used for producing intercensal population estimates across the social sciences and in general because of its simplicity and accurate estimates (Chen *et al.* 2012; Crowder *et al.* 2012). It searches for a straight line that passes through the end points y_A and y_B. The method is mathematically expressed in Equations (1a) and (1b):

$$Y_i = \frac{y_A - y_B}{a - b}(i - b) + y_B \tag{1a}$$

$$Y_i = (1 - \alpha)y_B + \alpha y_A \tag{1b}$$

where α is the interpolation factor and varies from 0 to 1. Interpolated datasets are bound between y_A and y_B true values are, on average, underestimated: this is strongly dependent on the distribution of data (on which side the distribution is tailored – left or right) and should be verified for each dataset (Musial *et al.* 2011).

Automated time series forecasting models

The following automated time series forecasting models were applied in this study with the aim of predicting the human population of the City of uMhlathuze. The reason for using multiple automated time forecasting models was mainly to compare their forecast results and select the best forecasting model, and to maximise accuracy and minimise bias.

Auto-regressive integrated moving average (ARIMA) model

The Auto-Regressive Integrated Moving Average (ARIMA) model is the most frequently used and recognized time series forecasting model, which consists of three components: namely the auto-regressive (AR), integral (I), and moving average (MA)

Table 1 | Description of data and sources

Period	Description	Document type	Format	Source
1996–2016	Human population	Microsoft Excel	Electronic	StatsSA
2016	uMhlathuze municipal boundary	Shapefile	Vector	HSRC
2016	uMhlathuze municipal wards	Shapefile	Vector	HSRC
2016	uMhlathuze roads	Shapefile	Vector	HSRC
2016	River lines/lakes	Shapefile	Vector	HSRC
2016	Location of towns	Shapefile	Vector	HSRC
2016	Wet logged areas	Shapefile	Vector	HSRC
2016	Protected areas	Shapefile	Vector	HSRC

(Jain & Mallick 2017). The AR component in Equation (2) models the association between the value of a variable at a specified time with its value in previous time(s), for example if $P = 1$, then each observation is a function of only one previous observation (Karabiber & Xydis 2019):

$$y_t = c + \phi_1 Y_{t-1} + e_t \tag{2}$$

where y_t represents the observed value at time t, Y_{t-1} represents the previous observed value at time $t-1$, e_t represents the random error at time t, c and ϕ_1 are both constants. Other observed values of the series can be included in the right-hand side of the equation if $p > 1$. The integrated (I) component in Equation (3) comes into consideration when the time series becomes stationary after the first (or second) difference:

$$y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots \phi_p Y_{p-1} + e_t \tag{3}$$

The MA component in Equation (4) models the association between the values of error term of a variable at a specified time with its error term value in previous time(s), for example if $q = 1$, then each observation is a function of only one previous error:

$$y_t = c + \theta_1 e_{t-1} + e_t \tag{4}$$

where e_t represents the random error at time t, e_{t-1} represents the previous random error at time $t-1$. Other errors can be included in the right-hand side of the equation if $q > 1$. Combining the components discussed in Equations (2)–(4) leads to the ARIMA model in Equation (5) (Hyndman & Athanasopoulos 2018):

$$y'_t = c + \phi_1 y'_{t-1} + \phi_2 y'_{t-2} + \ldots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q} + e_t \tag{5}$$

where y'_t is the differenced time series. The 'predictions' on the right-hand side include both lagged values of y_t and lagged errors. This is called ARIMA (p, d, q) model, where p is the order of the autoregressive part, d is the degree of first differencing involved and q is the order of the moving average part.

The automated ARIMA model uses an automatic forecasting model known as the 'auto.arima function', which is provided through the forecast package in RStudio (Hyndman *et al.* 2008). By using the 'auto.arima function', the model overcomes a limitation of selecting the appropriate parameters and being considered subjective by automatically selecting the best number of time lags of the autoregressive model, the degree of differencing needed to reach stationarity as well as the order of the moving average model (Hyndman & Athanasopoulos 2018).

Error trend seasonality (ETS) model

The Error Trend Seasonal (ETS) model was proposed in the 1950s and is often used to forecast univariate time series (Pegels 1969). The simple ETS model has the following form:

$$\hat{y}_{t+1|t} = \hat{y}_{t|t-1} + \alpha(y_t - \hat{y}_{t|t-1}), \quad \text{or} \quad \hat{y}_{t+1|t} = \alpha y_t + (1-\alpha)\hat{y}_{t|t-1} \tag{6}$$

where y_t is a time series, $\hat{y}_{t|t-1}$ is the forecast value for$y_t$by taking account of all previous values, $y_1, y_2, \ldots, y_{t-1}$, and α is a smoothing parameter between 0 and 1. For longer range forecasting by the simple ETS model, the forecast formula could be written as:

$$\hat{y}_{t+h|t} = \hat{y}_{t+1|t} \quad \text{where} \quad h = 1, 2, 3 \ldots, \tag{7}$$

where h means the number of periods ahead (Lepot *et al.* 2017). A method that is suitable for a specific time varies with trend and seasonality. The trend component includes five possibilities: None (N), Addictive (A), Additive damped (Ad), Multiplicative (M), and Multiplicative damped (Md), while the seasonal component includes three possibilities: None (N), Additive (A), and Multiplicative (M).

The Error Trend Seasonality (ETS) model uses an automatic forecasting model known as the 'ets function' which is provided through the forecast package in R programming code (Hyndman *et al.* 2008). By using the 'ets function', the model overcomes a limitation of failing to provide a method for easy calculation of prediction intervals (Hyndman & Athanasopoulos 2018).

From the forecast package, the model considers the error, trend, and seasonal components in choosing the best exponential smoothing model from over 30 possible options by optimising initial values and parameters using the Maximum Likelihood Estimation (MLE), and selecting the best model based on the Akaike Information Criterion (AIC) (Hassani *et al.* 2017).

Box-Cox transformations, ARMA errors, trend and seasonal components (BATS) model

The Box-Cox Transformation, Auto-Regressive Moving Average (ARMA) errors, Trend and Seasonal components (BATS) model, is an exponential smoothing state space model, which includes Box-Cox Transformation (model for non-linear data), ARMA model for residuals, Trend and Seasonal components and Trigonometric Seasonality (Hassani *et al.* 2018). The model relies on trigonometric functions; therefore, it can handle complex seasonal time series variations. The BATS model has the following form:

$$y_t^{(\omega)} = \begin{cases} \dfrac{(y_t^{\omega} - 1)/\omega}{\log Y_t} \end{cases} \quad \text{for when } \omega \neq 0 \text{ and when } \omega = 0 \tag{8}$$

$$y_t^{(\omega)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^{n} s_{t-mi}^{i} + d_t \tag{9}$$

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t \tag{10}$$

$$b_t = (1-\phi)b + \phi b_{t-1} + \beta d_t \tag{11}$$

$$s_t^{(i)} = s_{t-mi}^{(i)} + \gamma_i d_t \tag{12}$$

where $\omega \in R$ is the Box-Cox Transformation parameter, $m_1, \ldots\ldots, m_n$ denote the constant periods of the n seasonal components, b is the long run trend, d_t is an ARMA (p, q) process with Gaussian white noise innovations having zero mean and constant variance, $t = 1, \ldots, T$, l_t is the local stochastic level, b_t is the short term trend and s_t^i *is* the stochastic level of the $i - ith$ seasonal component.

The BATS model uses an automatic forecasting model known as the 'BATS function' which is provided through the forecast package in R programming code (Hyndman *et al.* 2008). By using the 'BATS function', the model overcomes the limitation of being slow to estimate, especially with long time series (Hyndman & Athanasopoulos 2018). However, on the positive side, the model has the option of implementing a multi-seasonality analysis without too many parameters, it can work under non-integer seasonality, and it can work under high frequency data (Hassani *et al.* 2018). Furthermore, the model allows for any autocorrelation in the residuals to be considered; and lastly, it involves a much simpler, yet more efficient estimation procedure.

Evaluating prediction performance

To evaluate the performance of the automated time forecasting models, the root mean squared error (RMSE), mean absolute error (MAE), mean absolute scaled error (MASE) and mean absolute percentage error (MAPE) were used. The methods are mathematically described as follows:

$$ME = \frac{1}{N}\sum_{t=1}^{N}(f_i^{obs} - \hat{f}_i^{pred}), \tag{13}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N}(f_i^{obs} - \hat{f}_i^{pred})^2}, \tag{14}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{i=N}|f_i^{obs} - \hat{f}_i^{pred}|, \tag{15}$$

$$MPE = \frac{100\%}{N}\sum_{t=1}^{N}\frac{(f_i^{obs} - \hat{f}_i^{pred})}{f_i^{obs}}, \tag{16}$$

$$MAPE = \frac{100\%}{N}\sum_{t=1}^{N}\left|\frac{f_i^{obs} - \hat{f}_i^{pred}}{f_i^{obs}}\right| \tag{17}$$

where N is the number of datum points in the test period, f_i^{obs} is the ith observed f value, and $\hat{f}_i^{pred}$ is the ith forecasted f value. The 'accuracy()' function from the RStudio forecast package was utilized when calculating the model performance indicators. This function automatically calculates the forecast accuracy measures (Karabiber & Xydis 2019).

Determining the number of additional reservoir tanks to be built in future

One of the goals of this study is to predict the human population growth of the City of uMhlathuze for the next 50 years (from 2017 until the year 2067), therefore water provision for future human population must be made. As such, a linear programming model from Equations (18)–(20) was formulated with the aim of determining the number of additional reservoir tanks to be built in the future, based on the prediction of human population growth. The model considered the following input notations:

- $c_i =$ the unit cost of building a reservoir tank $(c_1, c_2, c_3, \ldots, c_m)$,
- $r_i =$ reservoir tank type to be built in future $(r_1, r_2, r_3, \ldots, r_m)$,
- $w_i =$ capacity of reservoir tanks to be built in future $(w_1, w_2, w_3, \ldots, w_m)$,
- $m =$ the total number of additional reservoir tank type to be built in future,
- $z_i =$ human population growth (demand) in the next 50 years,
- $k =$ minimum amount of potable water available per person per day,
- $y =$ capacity of current reservoir tanks,
- $i =$ the index of potential facility sites.

Decision variable

- $r_i = (r_1, r_2, r_3, \ldots, r_m)$, where $i = 1, 2, 3, \ldots, m$

Using the above notations, a linear programming model can be formulated as follows:

$$\text{Minimize} \quad f(x) = \sum_{i=1}^{m} c_i r_i \tag{18}$$

$$\text{Subject to:} \quad f(x) = \sum_{i=1}^{m} z_i k \le \sum_{i=1}^{m} w_i r_i + y \tag{19}$$

$$r_i \ge 0 \quad \text{integer} \quad \forall i \tag{20}$$

The objective function in Equation (18) seeks to minimize the total unit cost of building an additional reservoir tank in the future while satisfying all the supply and demand restrictions. The constraint in Equation (19) requires that demand from human population (demand) should be less or equal to water supplied per person per day. Lastly, the constraint in Equation (20) is an integer decision variable.

RESULTS AND DISCUSSION

Automated time series forecasting results

The human population growth of the City of uMhlathuze reached its first peak in 2001, when it was increasing by 6.728% per year on average, placing the municipality as one of the fastest growing municipalities, as illustrated in Figure 2(a). Such growing human populations escalate the demand for water infrastructure. From 2001 the human population growth curve slightly changed as the human population was growing at a far slower growth rate of 1.465% per year on average until it reached its second peak in 2011.

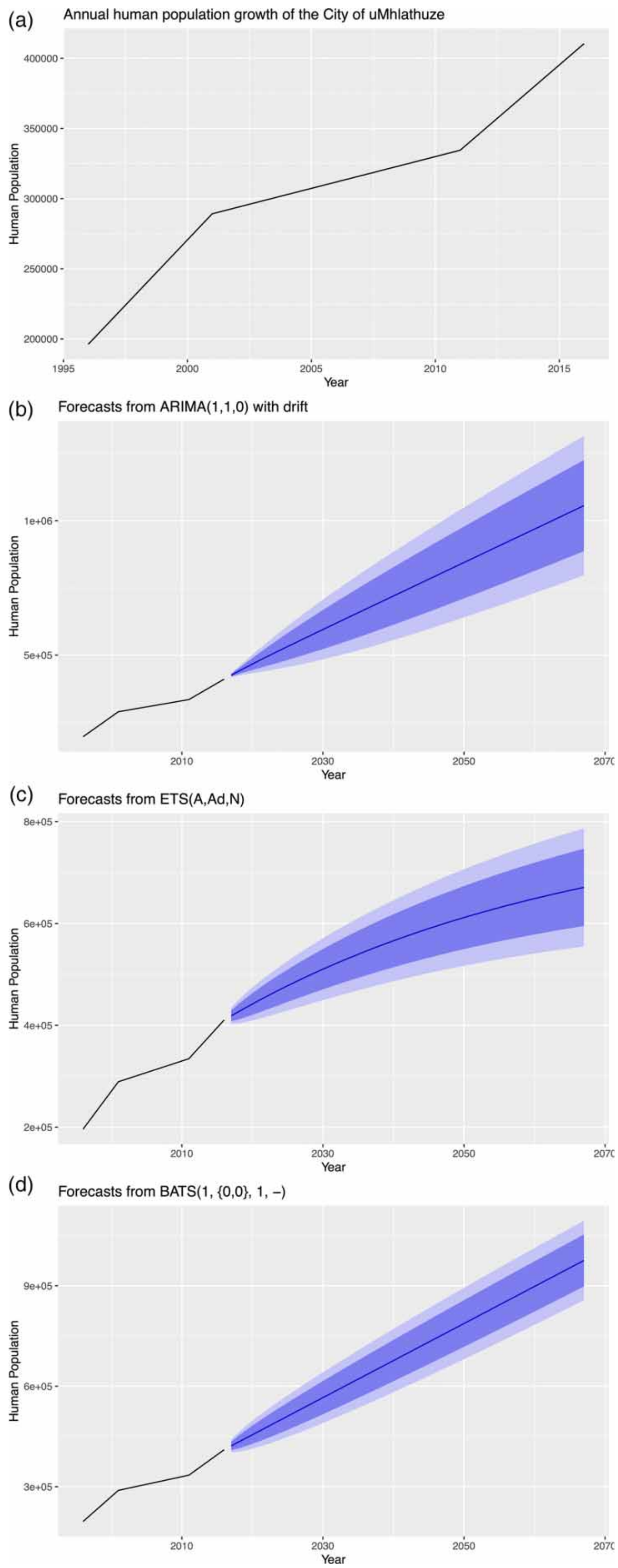


Figure 2 | (a) Human population growth of the City of uMhlathuze for 1996, 2001, and 2011 Census periods and 2016 CS; Forecast graphs for human population growth; (b) ARIMA(1,1,0) with drift model; (c) ETS(A,Ad,N) model; and (d) BATS(1, {0, 0}, 1, -) model.

The factors that could have influenced this annual slow growth rate include the cholera outbreak as well as the human immunodeficiency virus (HIV) and acquired immune deficiency syndrome (AIDS) cumulative deaths experienced by the KZN province and its municipalities between 2000 and 2001 (Department of Health KwaZulu Natal: Epidemiology Unit 2001; Dorrington *et al.* 2002), according to information presented in Tables 2 and 3 respectively. The HIV, AIDS and cholera outbreak resulted in many premature deaths which affected the human population in rural and urban areas of the KZN province and the City of uMhlathuze at large (Eshowe, Nkandla, uMfolozi, and Ngwelezane).

From 2011 the human population growth increased at a pace of 4.181% per year on average until it reached another peak in 2016. This indicates that quite a large number of people migrated (in-migration) to the KZN province as well as the City of uMhlathuze between the years 2011 and 2016 (Statistics South Africa 2016), according to the in-migration statistics presented in Table 4.

The human population growth of the City of uMhlathuze was then predicted from 2017 until 2067 as illustrated in Figure 2(b)–2(d). The human growth curves for ETS(A,Ad,N) model, ARIMA(1,1,0) with drift model and BATS(1, {0,0}, 1, -) model have shown an upward trend from the year 2017 until 2067. Based on the performance results of these automated time series forecasting models, the human population growth is expected to triple by 2067, at a higher pace than at any period since 1996 according to the ARIMA(1,1,0) with drift model. The findings are in agreement with those obtained by Aurecon South Africa (2015a, 2015b) and the City of uMhlathuze (2019, 2020, 2021), that with the City of uMhlathuze being the fastest developing municipality in SA, an increase in human population is probable in the coming years; largely to be influenced by an enlarged municipal area following the inclusion of three wards from the former Ntambanana, post the 2016 local government elections.

Table 2 | Cholera outbreak in KZN between August 2000 and November 2001

Number	Region	Number of reported cases and deaths
1	Ugu Region – South Coast	8650
2	Pietermaritzburg – Ndlovu	4425
3	Ladysmith	454
4	Ulundi	24,310
5	Eshowe – Nkandla	31,462
6	uMfolozi – Ngwelezane	23,951
7	Durban – Unicity	2456
8	Newcastle – Nqutu	2088
9	Jozini	406
10	Stanger – uNtunjambili	7672
	Total	**157,874**

Table 3 | HIV and AIDS deaths in KZN province

Year	Cumulative HIV and AIDS deaths
2000	105,340
2001	159,216

Table 4 | KZN province in-migration

Year	In-migration
2011	232,248
2016	243,439

Economic opportunities from the growing industrial centre as well as agricultural activities might be factors that contribute to a continuously growing human population in the City of uMhlathuze in the near future, as more people will be coming (in migration) from neighbouring rural and urban communities seeking employment (City of uMhlathuze 2021). Increased demand for water may also emanate from water users located outside the municipality's catchment but supplied from it; and this may lead to the number of people requiring water being more than what the reservoir tank can handle (Aurecon South Africa 2015a).

The findings also confirm a theory developed by Malthus (1798), that the human population will grow exponentially (for instance, doubling or tripling with each cycle) while basic means of subsistence like food, water supply and distribution, and electricity amongst other things, will grow at an arithmetic rate (for instance, by the repeated addition of a uniform increment in each uniform interval of time). Growth in human population does not only indicate the changes in human population size, but also indicates the changes in the distribution of basic services such as water (Kılıç 2020). It also means mounting demand and competition for water for domestic, industrial, agricultural, and municipal uses. As a result, the challenges to meet user demands for basic services may also increase.

An implication of these findings is the possibility that, should anticipated human population growth materialise, the current reservoir tanks are likely to come under stress due to increased water demand (Mulwa *et al.* 2021). This is supported by a recent finding from studies conducted by Parks *et al.* (2019), Heidari *et al.* (2021), Mulwa *et al.* (2021), that a continuous increase in urbanisation and human population growth affects water infrastructure resources, and this may in turn cause water shortages in rural and urban communities.

Evaluating prediction performance

Table 5 presents the forecast performance of ARIMA(1,1,0) with drift model, ETS(A,Ad,N) model, and BATS(1, {0,0}, 1, -) model by means of the RMSE, MAE, MAPE, and MASE. For each of these measures, a smaller value indicates higher prediction accuracy of the model (Karabiber & Xydis 2019). As shown in Table 5, ARIMA(1,1,0) with drift model outperformed the ETS(A,Ad,N) model, and BATS(1, {0, 0}, 1, -) model, because it has the smallest errors across four forecast evaluation measures. This indicates that the ARIMA(1,1,0) with drift model is the best model to forecast human population growth of the City of uMhlathuze.

Reservoir tanks to be built in the City of uMhlathuze in the next 50 years

A linear programming model was formulated with the goal of determining the provision of reservoir tanks in the City of uMhlathuze to satisfy the possible human water demand for the next 50 years. The analysis takes into consideration the estimated unit cost of building a reservoir tank, human population growth forecasts (demand) in the next 50 years, existing reservoir tank capacity as well as the minimum amount of potable water (25 litres) available per person per day, amongst other things.

The results of a linear programming model presented in Table 6 have indicated that a total number of 687 additional water reservoir tanks with a capacity of 47,500 kilolitres should be built annually in the City of uMhlathuze within the period of 50 years in order to supplement the anticipated increase in demand for water as a result of a rapid increase in human population growth and urbanization. Below is the description of the notations used in Table 6:

RN* - Reservoir number,
EBC* - Estimated building costs in SA Rands (R),
RCK* - Reservoir capacity in kilolitres,
DV* - Decision variables,
TC* - Total costs (x),
TCK* - Total capacity in kilolitres

Table 5 | Results from evaluating automated time series forecasting methods

Time Series Forecasting Method	RMSE	MAE	MAPE	MASE
ARIMA(1,1,0) with drift	3725	2265	0.757	0.211
ETS(A,Ad,N)	7221	6708	2.354	0.626
BATS(1, {0, 0}, 1, -)	9731	7176	2.732	0.670

Table 6 | Additional reservoir tanks to be built over a 50-year period

RN*	EBC*	RCK*	DC*	TC*	TCK*
1	R3,507,551.65	3400	0	0	0
2	R5,139,979.05	5000	0	0	0
3	R3,707,551.65	3500	0	0	0
4	R7,139,979.05	6500	0	0	0
5	R2,276,270.85	2252	0	0	0
6	R1,248,878.15	900	0	0	0
7	R653,311.60	248	0	0	0
8	R3,018,033.25	2700	0	0	0
9	R3,11,335.30	100	0	0	0
10	R18,579,916.20	20,000	0	0	0
11	R341,335.30	114	0	0	0
12	R341,335.30	114	0	0	0
13	R2,220,033.25	1600	0	0	0
14	R625,000.05	320	0	0	0
15	R9,037,958.10	10,230	0	0	0
16	R605,000.05	300	0	0	0
17	R1,420,033.25	1275	0	0	0
18	R4,896,979.05	4500	0	0	0
19	R18,579,916.20	20,000	0	0	0
20	R8,867,958.10	10,000	0	0	0
21	R865,724.60	500	0	0	0
22	R9,279,958.10	9000	0	0	0
23	R9,279,958.10	9000	0	0	0
24	R8,867,958.10	10,000	0	0	0
25	R1,437,142.35	1000	0	0	0
26	R3,889,979.05	4000	0	0	0
27	R311,335.30	100	0	0	0
28	R1,918,033.25	1500	0	0	0
29	R453,311.60	200	0	0	0
30	R2,018,033.25	1880	0	0	0
31	R8,867,958.10	10,000	0	0	0
32	R5,139,979.05	5000	0	0	0
33	R18,579,916.20	20,000	0	0	0
34	R18,579,916.20	20,000	0	0	0
35	R311,335.30	100	0	0	0
36	R453,311.60	200	0	0	0
37	R1,278,878.15	924	0	0	0
38	R41,832,811.45	47,500	687	28,739,141,466	32,632,500
39	R705,000.05	350	0	0	0
40	R331,335,30	115	0	0	0
41	R374,335.30	140	0	0	0
42	R503,311.60	250	0	0	0

(*Continued.*)

Table 6 | Continued

RN*	EBC*	RCK*	DC*	TC*	TCK*
43	R41,832,811.45	47,500	0	0	0
44	R311,335.30	100	0	0	0
45	R605,000.05	300	0	0	0
46	R1,220,033.25	1200	0	0	0
47	R952,897.75	620	0	0	0
			687		
			Objective function	28,739,141,466	
			Current supply	284,532	
		Supply	>=	**Demand**	
		32,917,032		32,889,125	

CONCLUSIONS AND FUTURE WORKS

This study has succeeded in applying machine learning-inspired automated time series models such as ARIMA(1,1,0) with drift, ETS(A,Ad,N), and BATS(1, {0, 0}, 1, -) model to predict the growth in human population of the City of uMhlathuze over a 50-year period. Significant increases in human population raises many concerns since human population growth, without corresponding expansion of the reservoir tanks, might cause a shortfall in terms of water provision and distribution in future, unless a feasible plan based on the prediction of human population growth forecasts is developed (Mohammed & Sahabo 2015; Parks *et al.* 2019; Mulwa *et al.* 2021).

This is the first study of its kind to demonstrate a linear programming model to be used for water provision planning for half a century in advance. The findings of this study could be useful as knowledge of how human population will grow in future and how many additional reservoir tanks will be needed to satisfy the human water demand is extremely important in terms of guiding the City of uMhlathuze towards policy formulation, development, implementation as well as facilitating planning with the aim of accommodating a continuously growing population and the competing demands from human settlements. The study also highlighted the importance of managing the increased demand for water in order to avoid water shortages in rural and urban communities.

Future work to be derived from this study may include formulating the facility location problem to determine the optimal locations where additional reservoir tanks can be built, based on how settlements are growing and their spatial direction of expansion; also, where demand is likely to be located in future, so that water can be distributed to all applicable rural or urban communities. Moreover, hybrid forecasting models can also be developed and applied for predicting the human population growth of the City of uMhlathuze and therefore the forecast accuracy can be compared in terms of their reliability and robustness. Some of the suggested hybrid combinations include:

- Hybrid neural network auto-regression (NNAR) + ARIMA model,
- Hybrid ARIMA + ETS model,
- Hybrid ETS + NNAR model, and
- Hybrid ARIMA + ETS + NNAR model.

The combination of different time series forecast methods should allow researchers to maximise the chance of capturing both the linear (if any) and nonlinear patterns and may achieve better performance and forecast accuracy (Panigrahi & Behera 2017).

FUNDING

This work is based on research supported wholly by the National Research Foundation of South Africa (Grant Number; 110415).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Anvari, S., Tuna, S., Canci, M. & Turkay, M. 2016 Automated Box–Jenkins forecasting tool with an application for passenger demand in urban rail systems. *Journal of Advanced Transportation* **50** (1), 25–49.

Aurecon South Africa 2015a *Water Reconciliation Strategy for Richards Bay and Surrounding Towns, Water Requirements Report: 109343-9170*. Aurecon South Africa, Durban, South Africa.

Aurecon South Africa 2015b *Water Reconciliation Strategy for Richards Bay and Surrounding Towns, Water Requirements Report: 109343-9168*. Aurecon South Africa, Durban, South Africa

Chen, Z., Fan, P. & Sun, M. 2012 A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioural data. *European Journal of Operational Research* **223** (2), 461–472.

City of uMhlathuze 2016 *uMhlathuze Local Municipality: Final IDP Review 2015/2016. Integrated Development Plan DMS: 1059655*. City of uMhlathuze, Richards Bay, South Africa.

City of uMhlathuze 2019 *uMhlathuze Local Municipality: Draft Integrated Development Plan (IDP) Review 2019–2020. Integrated Development Plan, DMS: 1325466*. City of uMhlathuze, Richards Bay, South Africa.

City of uMhlathuze 2020 *uMhlathuze Local Municipality: Draft Integrated Development Plan (IDP) Review for 2020/21. Integrated Development Plan DMS: 1418729*. City of uMhlathuze, Richards Bay, South Africa.

City of uMhlathuze 2021 *Final IDP Review: May 2021. Integrated Development Plan DMS: 1466045*. City of uMhlathuze, Richards Bay, South Africa.

Crowder, K., Pais, J. & South, S. 2012 Neighborhood diversity, metropolitan constraints, and household migration. *American Sociological Review* **77** (3), 325–353.

Department of Health, KwaZulu Natal: Epidemiology Unit 2001 *Cholera Update in KwaZulu-Natal, November 2001, Department of Health, KwaZulu Natal: Epidemiology Unit*.

Department of Water and Sanitation 2017 *National Norms and Standards for Domestic Water and Sanitation Services. Government Gazette 41100* Department of Water and Sanitation, Pretoria, South Africa.

Department of Water and Sanitation 2018 *National Water and Sanitation Master Plan: Volume 1. Master Plan, WP11240*. Department of Water and Sanitation, Pretoria, South Africa.

Development Bank of Southern Africa 2022 *Water & Sanitation: Every Drop, Every Bit Counts. Technical Report, Development Bank of Southern Africa*.

Dighade, R. R., Kadu, M. & Pande, A. 2014 Challenges in water loss management of water distribution systems in developing countries. *International Journal of Innovative Research in Science, Engineering and Technology* **3** (6), 13838–13846.

Dorrington, R., Bradshaw, D. & Budlender, D. 2002 *HIV/AIDS Profile in the Provinces of South Africa. Indicators for 2002, Centre for Actuarial Research, University of Cape Town*.

Faye, C. 2021 Water resources and their management in an increasing urban demography: The case of Dakar City in Senegal. In: Chandrasekaran, P. T., Javaid, M. S. & Sadiq, A. (eds). *Resources of Water*. InTechOpen, UK, pp. 101–120.

Fukuda, K. 2010 Interpolation and forecasting of population census data. *Journal of Population Research* **27** (1), 1–13.

Gulseven, O. 2016 Forecasting population and demographic composition of Kuwait until 2030. *International Journal of Economics and Financial Issues* **6** (4), 1429–1435.

Gunawan, I., Schultmann, F. & Zarghami, S. 2017 The four Rs performance indicators of water distribution networks: a review of research literature. *International Journal of Quality and Reliability Management* **34** (5), 720–732.

Hassani, H., Silva, E., Antonakakis, N., Filis, G. & Gupta, R. 2017 Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research* **63**, 112–127.

Hassani, H., Silva, E., Gupta, R. & Segnon, M. 2018 Forecasting the price of gold. *Applied Economics* **47** (39), 4141–4152.

Heidari, H., Arabi, M., Warziniack, T. & Sharvelle, S. 2021 Effects of urban development patterns on municipal water shortage. *Frontiers in Water* **3**, 694817. doi: 10.3389/frwa.2021.694817.

Hyndman, R. & Athanasopoulos, G. 2018 *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, pp. 290–472.

Hyndman, R., Koehler, A., Ord, J. & Snyder, R. 2008 *Forecasting with Exponential Smoothing: the State Space Approach*. Springer Science and Business Media, Berlin, Germany, p. 260.

Jain, G. & Mallick, B. 2017 A study of time series models ARIMA and ETS. *Modern Education and Computer Science* **4**, 57–63.

Javadinejad, S., Ostad-Ali-Askari, K. & Jafary, F. 2019 Using simulation model to determine the regulation and to optimize the quantity of chlorine injection in water distribution networks. *Modeling Earth Systems and Environment* **5** (3), 1015–1023.

Karabiber, O. & Xydis, G. 2019 Electricity price forecasting in the Danish say-ahead market using the TBATS, ANN and ARIMA methods. *Energies* **12** (5), 928.
Karuaihe, S., Molokomme, M., Mathebula, J. & Sakoane, L. 2016 *Rural Water Access and Management in Southern Africa: Does Community Involvement Offer Alternatives? Policy Brief: 2016, Human Sciences Research Council*.
Kassa, M. 2017 Evaluation of water supply and demand: the case of Shambu town, Western Oromia, Ethiopia. *International Journal of Water Resources and Environmental Engineering* **9** (5), 96–101.
Kılıç, Z. 2020 The importance of water and conscious use of water. *International Journal of Hydrology* **4** (5), 239–241.
Kulkarni, S., Kulkarni, R. & Patil, S. 2014 Analysis of population growth of India and estimation for future. *International Journal of Innovative Research in Science, Engineering and Technology* **3** (9), 843–850.
KZN Provincial Government Communication 2022 *City of uMhlathuze Local Municipality: Geography, History and Economy*. KZN Provincial Government Communication, Pietermaritzburg, South Africa.
Lepot, M., Aubin, J. & Clemens, F. 2017 Interpolation in time series: an introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water* **9** (10), 796.
Loubser, C., Chimbanga, B. M. & Jacobs, H. 2021 Intermittent water supply: a South African perspective. *Water SA* **47** (1), 1–9.
Malthus, T. R. 1798 *An Essay on the Principle of Population*. St. Paul's Church-Yard, London.
Mnisi, N. 2020 *Water Scarcity in South Africa: A Result of Physical or Economic Factors?* HSF Briefs, Helen Suzman Foundation, Johannesburg, South Africa.
Mohammed, A. & Sahabo, A. 2015 Water supply and distribution problems in developing countries: a case study of Jimeta-Yola, Nigeria. *International Journal of Applied Sciences and Engineering* **1**, 473–483.
Mondol, H., Mallick, U. K. & Biswas, M. H. A. 2018 Mathematical modeling and predicting the current trends of human population growth in Bangladesh. *Advances in Modelling and Analysis A* **55** (2), 62–69.
Mthethwa, N. I. 2018 *Evaluating Water Conservation and Water Demand Management in an Industrialised City: A Case Study of the City of uMhlathuze in Richards Bay*. *Master's Thesis*, University of Cape Town.
Mulwa, F., Li, Z. & Fangninou, F. F. 2021 Water scarcity in Kenya: current status, challenges and future solutions. *Open Access Library Journal* **8** (1), 1–15.
Musial, J., Verstraete, M. & Gobron, N. 2011 Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series. *Atmospheric Chemistry and Physics* **11** (15), 7905–7923.
Ofori, T., Ephraim, L., Nyarko, F., Keshtavar, A., Moeinaddin, M., Dehnavi, H., Uw, V., Akpan, I., Akpan, A. & Akpan, U. 2013 Mathematical model of Ghana's population growth. *International Journal of Modern Management Sciences* **2** (2), 57–66.
Oki, T. & Quiocho, R. E. 2020 Economically challenged and water scarce: identification of global populations most vulnerable to water crises. *International Journal of Water Resources Development* **36** (2–3), 416–428.
Omony, J. 2014 Modelling and parameter sensitivity analysis of population dynamics in Uganda. *Journal of Advances in Mathematics and Computer Science* **4** (1), 153–161.
Panigrahi, S. & Behera, H. S. 2017 A hybrid ETS–ANN model for time series forecasting. *Engineering Applications of Artificial Intelligence* **66**, 49–59.
Parks, R., McLaren, M., Toumi, R. & Rivett, U. 2019 Experiences and lessons in managing water from Cape Town. Grantham Institute Briefing Paper, 29.
Patel, A. & Prajapati, M. 2016 Estimation for future population growth of China by using logistic model. *International Journal of Scientific Development and Research* **1** (9), 52–56.
Pegels, C. 1969 Exponential forecasting: some new variations. *Management Science* **15** (5), 311–315.
R Core Team 2020 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Sonaje, N. P. & Joshi, M. G. 2015 A review of modelling and application of water distribution networks (WDN) softwares. *International Journal of Technical Research and Applications* **3** (5), 174–178.
South African Government 1996 *Constitution of the Republic of South Africa. Constitution of the Republic of South Africa, South African Government*.
South African Institution of Civil Engineering 2017 *2017 Infrastructure Report Card for South Africa. Technical Report, South African Institution of Civil Engineering*.
Statistics South Africa 2016 *Community Survey 2016: Statistical Release, Technical Report: P0301*. Statistics South Africa, Pretoria, South Africa
Stephano, M. & Jung, I. 2019 Estimation and prediction of population using mathematical models in Tanzania. *IOSR Journal of Mathematics* **15**, 73–82.
Tong, M., Yan, Z. & Chao, L. 2020 Research on a grey prediction model of population growth based on a logistic approach. *Discrete Dynamics in Nature and Society* **2020**, 2416840. https://doi.org/10.1155/2020/2416840.
Wang, P., Huang, X., Mango, J., Zhang, D., Xu, D. & Li, X. 2021 A hybrid population distribution prediction approach integrating LSTM and CA models with micro-spatiotemporal granularity: a case study of Chongming District, Shanghai. *ISPRS International Journal of Geo-Information* **10** (8), 544.
Water Integrity Network 2020 *Money Down the Drain: Corruption in South Africa's Water Sector*. Technical report. Corruption Watch, Johannesburg, South Africa.

Wei, H., Jiang, Y. & Zhang, Y. 2015 A review of two population growth models and an analysis of factors affecting the Chinese population growth. *Asian Journal of Economic Modelling* **3** (1), 8–20.

Zabadi, A., Assaf, R. & Kanan, M. 2017 Mathematical and statistical approach for predicting the population growth. *World Wide Journal of Multidisciplinary Research and Development* **3** (7), 50–59.

Zulu, L. 2017 *Nkandla Municipal Water Infrastructure Asset – GIS Report for King Cetshwayo District Municipality*. King Cetshwayo District Municipality, Richards Bay, South Africa.

First received 5 May 2022; accepted in revised form 31 July 2022. Available online 5 August 2022