

## Article

# Implementation of Water Demand Forecasting Model to Aid Sustainable Water Supply Chain Management in UAE

Vian Ahmed<sup>1</sup>, Ahmad Saad<sup>2</sup>, Hasan Saleh<sup>3</sup>, Sara Saboor<sup>4\*</sup>, Nikita Kasianov<sup>5</sup> and Tahani Alnaqbi<sup>6</sup>

<sup>1</sup> Affiliation 1; [vahmed@aus.edu](mailto:vahmed@aus.edu)

<sup>2</sup> Affiliation 2; [b00053507@alumni.aus.edu](mailto:b00053507@alumni.aus.edu)

<sup>3</sup> Affiliation 3; [b00048302@alumni.aus.edu](mailto:b00048302@alumni.aus.edu)

<sup>4</sup> Affiliation 4; [g00080300@aus.edu](mailto:g00080300@aus.edu)

<sup>5</sup> Affiliation 5; [b00062665@alumni.aus.edu](mailto:b00062665@alumni.aus.edu)

<sup>6</sup> Affiliation 6; [g00082559@aus.edu](mailto:g00082559@aus.edu)

\* Correspondence: [g00080300@aus.edu](mailto:g00080300@aus.edu); Tel.: +971-50-227-1577 (U.A.E)

**Abstract:** Climate change has become the greatest threat to the survival of world and its ecosystem. With the irreversible impact on the ecosystem, problems like rise in sea level, food-insecurity, natural resources scarcity, seasonal disorders have increased over the past few years. Among these problems, the issue of water scarcity due to the lack of water resources and global warming has plagued several nations. Owing to the rising concerns over water scarcity United Nations (UN) has acknowledged water as a primary resource to the development of societies under the 'Water Goal' of the sustainable development goals. As the changing climate and intermittent availability of water resources pose major challenges to forecast demand, especially in countries like the United Arab Emirates (UAE) which has one of the highest per capita residential water consumption rates in the world. Therefore, the aim of this study is to propose an accurate water demand forecasting technique that incorporates all significant factors to predict the future water demands of the UAE. The forecasting model used is the Long Short Term Memory (LSTM), with the factors considered are mean temperature, mean rainfall, relative humidity, Gross Domestic Product (GDP), Consumer Price Index (CPI) and population growth. The LSTM model predicts the water demand forecasting in the UAE showing that the future demand will decrease from 1821 million m<sup>3</sup> in 2018 to 1809.9 million m<sup>3</sup> in 2027.

**Keywords:** Neural Networks, Long-Short Term Model, Water demand, Forecasting, Sustainable development goals, Water Goal.

## 1. Introduction

Over the past few years, climate change has become the biggest threat to the world which causes an irreversible change in the ecosystem such as rises in sea level, food-insecurity, natural resources scarcity, seasonal disorders and. increased in natural disasters such as wildfires, flooding and droughts [1]. However, among all the problems caused by climate change, water scarcity is believed to be the major troubling factor.

According to the global scale assessment of climate change on water scarcity conducted by [2] using the water crowding index. The study shows that by 2050, 0.5 to 3.1 billion people will be exposed to an increase in water scarcity. In addition, according to a report published by the United Nations in 2018, water scarcity claims to be a potential threat for every continent [3]. The scarcity can be accrued to physical shortages of water or inadequate access to water resources due to a lack of infrastructure. Infrastructure accounts for the storage facilities and distribution networks, which form a crucial part of the water supply chain management system. Moreover, pressure on existing water

resources due to increasing urban developments and escalating population rates, also contribute to an increase in water scarcity.

Thus, due to the growing concerns of water scarcity, United Nations (UN) recognizes water as the significant resource in underpinning equitable, stable and productive societies and ecosystems, which affects the entire development agenda. 'Water connects us all' under the 'Water Goal' was adopted as one of (Goal 6) of the seventeen sustainable development goals to ensure water security [4]. The goal not only incorporates the physical availability of freshwater resources relative to the water demand but also ensures the social and economic factor i.e. safety, quality, accessibility and management of water resources [5]. Thus, this implies that sustainable water management is a significant goal that must be incorporated in the political agenda and by policy makers to ensure the sustainable development of societies and nations.

Water scarcity is a global issue; as [6] conducted a study in metropolitan city of Istanbul that has become overcrowded and experiences a steep rise in the demand for clean water. The study found a significant amount of investment is required for the water supply networks as a considerable water is loss during the water distribution system. Similarly, [7] aims to develop a forecasting model for the precipitation forecasting in the basin of East Azarbaijan, Iran; which has become a significant issue in the water resource management due to climate change and drought. However, the study by [8] reports that 73% of the global affected population inhabit the Asian continent, wherein the thirteen most water-stressed countries are found in the Middle East alone. Amongst the Middle Eastern countries, UAE ranks tenth, with one of the highest per capita residential water consumption rates in the world.

Besides the residential sector, sectors such as the agricultural and industrial sectors also seek water to process their activities. During the last decade due to the development of oil sector, the Gulf region and the UAE particularly has gained enormous attention from the investors, which has shaped once the barren desert into a modern city. With the UAE government investing billions of dollars every year to improve the infrastructure of the country. The construction industry in the UAE utilizes almost half of the resources and raw materials available [9]. In addition, the industry alone accounts for 9 percent of the water consumption, which being poorly acknowledged as the sustainable resource in the construction industry.

Therefore, to meet the water demand of these sectors, UAE harbors several conventional and nonconventional sources of water, deployed within its water supply management system. The conventional natural sources include surface water and groundwater whereas the non-conventional sources represent desalinated and treated wastewater that use fossil fuels as the energy inputs. However, strikingly low availability of natural water resources has encouraged UAE to meet its requirements through desalination plants that account for 22% of the water produced in the UAE [10]. Although demands are satisfied with the current water resources, UAE is set to face challenges in the future owing to the depletion of natural water sources, exhaustion of fossil fuels, population growth, increasing urbanization and the climatic consequences of global warming [11]. This makes forecasting extremely indispensable in determining the demands of UAE in the face of these unprecedented factors, to secure water demand in the future. As forecasting technique helps in developing a pre-notion about the water demand that will help in achieving a breakeven between the water supply and water demand.

Thus, the intent of this study is to propose an accurate water demand forecasting solution that can integrate all important factors relevant to the water management strategies in the UAE.

## **2. Water Demand Forecasting**

Demand forecasting in the water supply networks, sets the groundwork for the functioning of the overall system. It is, therefore, critical to predict the consumer water requirements for the supply network in order to optimize the operation of the water management system. Moreover, [10] asserts that forecasting unveils the consumption patterns, which helps in extrapolating spatial or temporal patterns of water consumption, thereby allowing the appropriate utilization of water resources.

A similar argument was presented by [12], claims that forecasting is deemed indispensable especially in situations of water stress, wherein it becomes mandatory to assess the ability of current water resources to meet water demand in the future. The authors support their argument by imploring the development of a framework to counteract the water crisis through optimization of current management practices or investments in collateral that aim at satisfying water demands in the future. This implies that the accuracy of demand forecasting draws the line between water surplus or water shortage in the future and is considered to be a pivotal factor in developing forecasting techniques. The forecasting techniques can be designed to determine either short-term, immediate-term or long-term water requirements,

However, the focus of this study lies on the existing long-term water forecasting techniques implemented across several countries. According to [13], long term water demand forecasting is used to predict water demands over a period of ten years, which allows the countries to plan improved water supply infrastructures. In addition, the long term demand forecasting emphasized that the forecasting timespan allows room for several factors such as, population, water consumption rates and urbanization to come into play to study their impact on the water demand.

Thus, the following section aims to present several techniques, developed and implemented by water utilities across the world in an attempt to perform robust long-term water demand forecasting.

### *2.1. Forecasting Models*

The section aims to focus on the long term forecasting techniques being used for predicting water demand in several countries. It will aid in identifying the factors and parameters considered in developing the models and limitations associated with each model.

#### **A. Forecasting Model: Unit Water Demand Analysis**

This model referred to as unit water demand takes into consideration the number of users in the future, the time and the space variabilities, to forecast the water demand. This technique assumes that future users will consume water at the same rate as the present users, thereby highlighting the backdrop of this technique [14]. Built on this assumption, this technique presents an elementary approach to achieving water forecast and this is further supported by [13], as the most commonly implemented technique. The technique is applied in the absence of significant data and allows water management systems to develop a rough estimate of the future water demand. Due to its feasibility, unit water demand analysis can be flagged as a qualitative approach to determining water demand, which does not rely on 'empirical' data as demonstrated in [13,14]. This in turn might raise concerns regarding the validity of the results, therefore, bringing into question the accuracy of the forecasting.

Hence, the biggest drawback of this technique is attributed to its underlying assumption, i.e. consumer behavior remains unchanged in the future, which renders it superfluous in the face of changing consumer habits due to economic developments and urbanization. Although, unit-water demand analysis provides a good estimate for future planning of water resources, while doing so, it compromises accuracy and provides limited predictive capability, which renders its application.

#### **B. Forecasting Model: Univariate Time Series Analysis**

The univariate time series technique uses time series models to predict the future water demand. Chen and Boccelli, 2014 defines time series analysis as a "...classic statistical method for analyzing discrete, infinite series of random variables". It appears that their claim can be applied to water demand forecasting where the variables are considered to be the factors that affect water demand. As opposed to unit water demand analysis, the term "time series" reflects a dependency on past observations that can be used to develop a temporal relationship between the data and the time [15].

However, though this technique provides an upper hand in terms of accuracy in comparison to unit water demand analysis, it hinders its flexibility by not considering changes in the socio-economic context such as tariffs, employment, population and urban patterns in developing the model. Therefore, acting as the biggest weakness of this approach, while undermining the overall efficiency of the technique.

### C. Forecasting Model: Time Series Regression Analysis

Regression analysis is a statistical technique for investigating and modeling the relationship between variables [16]. In water demand forecasting, demand is assigned as the dependent variable and the factors affecting it are referred to as the independent variables. These factors can address population size, current water prices, and average income and precipitation rates. Other than the characteristic of regression, this model is similar to the univariate time series analysis as argued by [17]; that both models use time series data and deal with irregularities and fluctuations that occur due to government policies and emerging events that happened within the range of study. Therefore, both models make use of smoothing techniques employed to reduce irregularities in the historical data in order to improve the accuracy of results. The study also proposed a multiple regression model that used high correlation factors such as annual population, annual gross domestic product (GDP) and previous annual water consumption, to forecast annual water demand. Following the model development, they implemented the Hodrick-Prescott (HP) filter technique to smoothen trends in the time series, allowing them to minimize forecast errors and thereby yield high forecasting accuracy.

Despite its advantages, this technique's main weakness is that it was applied to few water systems and it is not ensured that these assessments are representative of other water utilities in general.

### D. Forecasting Model: Scenario-Based Analysis

Scenario-based approaches are scenario specific regression models that take into account a certain combination of independent variables in determining their impact on the water demand [13]. This approach calls for developing scenarios that can be investigated under certain constraints, in order to find the best strategy for water management systems in the future. For example, this approach was implemented in northwestern China to study the impact of land use on surface runoff and water yield by [18]. It focused on land use owing to the intrinsic relationship between water resources and land use patterns, given that different land structures have varying water demands. The investigation depicted that higher water utilization positively impacts urbanization while negatively impacting water availability due to prolonged use. This creates a pinch in the existing water resources that turns into a vicious cycle over time by depleting the available water resources.

Although the technique provided solutions for water resource allocation in China, it lacks the depth needed to understand the impact of socio-economic activities on water resources as thereby posing a potential research gap in the area of study.

### E. Forecasting Model: Decision Support System Analysis (DSS analysis)

The scenario-based analysis can be further extended for the development of decision support systems. Where decision support systems are custom-made and vary based on a combination of certain factors such as population, infrastructure and land use pattern. However, scenario-based analysis is based on regression models whereas decision support systems can be developed for any type of mathematical model, be it a linear, non-linear, single objective, multi-objective or regression model.

One such example is evident from the forecasting carried out in the emirates of Al Ain and Umm-Al-Quwain in the UAE by [19]. The implemented techniques used a constant water use rate model that primarily accounted for the population in the future indicating that consumption rates will remain the same in the future, with only the number of users varying. On the other hand, the alternate model known as the "specify forecasting model" relied on various number of factors such as population, temperature and rainfall.

Although the results by [19,20] are highly indicative of the fact that decision support systems may prove to be reliable in gauging the demand for water in the future, this strategy faces the same limitations as the scenario-based analysis. Moreover, not much research was found regarding this technique.

### F. Forecasting Model: Long Short-Term Memory (LSTM)

The long short-term memory (LSTM) is an architecture or model that uses an artificial recurrent neural network which is used to facilitate deep structural learning [21]. As defined by Zhang, the LSTM is an explicit recurrent neural network used to forecast or model impermanent sequences together with their long-term dependencies more precisely than conventional recurrent neural networks. Characteristically, RNNs have 'temporary' memory, and they mostly rely on previous information to perform current tasks. Thus, the LSTM model is designed to prevent the problem of long-term dependency [22]. Such problems must be avoided in water demand forecasting because conditions are always changing. Therefore, the model can remember information for a long time such as the RNNs in the LSTM model create a chain of reiterating modules as shown in Figure 1, where  $x_t$  acts as an input vector,  $h_t$  acts as hidden layer vector. Each module stands for a single time step, while combining all the modules creates the overall LSTM model.

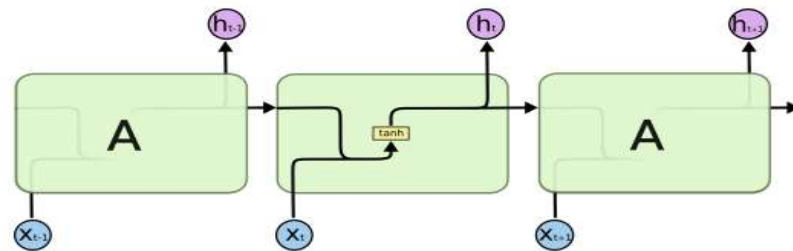


Figure 1. RNN Repeating Model.

A typical LSTM unit is made up of an activation function (output gate), input gate, and a cell (neuron). The role of a cell is to keep the memory of values over subjective time breaks [21]. The input gate and the output gate, on the other hand, regulate how data flows in and out of the cell. All the neurons, input and output gates make up the overall number of hidden layers and the overall network structure as can be seen in Figure 2.

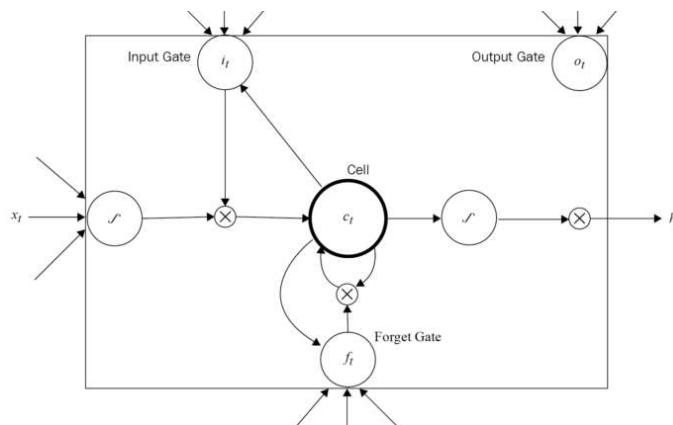


Figure 2. LSTM Network/Model. Source [21]

The core of an LSTM model is its cell state, which acts as the conveyor belt to flow the information along the entire chain [23]. The model operates in four steps such as: First, the activation function (noted as forget gate) of the LSTM structure decides the type of information to let out of the cell state (neuron) as shown in equation 1(a).

Secondly, the input gate layer (activation function) determines the type of information to let into the cell state and updates the value as in equation 1 (b). The tanh layer then generates a vector of the novel information to add to the cell state.

Thirdly, the LSTM system combines the two added information. It updates the content of the cell accordingly to create a new cell state. An operation is then generated to scale the state value and offer potential solutions in equation 1 (c).

Finally, the model determines what to give as an output or solution based on the filtered version of what it has in the cell state. The output gate layer (with its activation function) does complete this task in an organized fashion such that it only produces what it intended to output as shown in equation 1 (d)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1a)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1b)$$

$$C_t = f_t * C_{t-1} + i_t * C'_t \quad (1c)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) , h_t = o_t * \tanh(C_t) \quad (1d)$$

Source: [23]

Whereas as, in eq 1 (a-d)  $x_t$  acts as an input vector,  $h_t$ : hidden layer vector,  $o_t$ : output vector,  $W, b, C$ : parameter matrices and vector and  $\sigma$ : Activation functions.

The overall structure of an LSTM is almost identical to a neural network architecture with a set of hidden layers containing neurons with activation functions as discussed previously. The only difference lies in the fact that a different time instance carries its own short-term memory states (parameters and weights), and by combining the overall states makes up the whole LSTM model.

Therefore, an assumption could be made that memory is one of the shortcomings of the time series model as they face keeping memory while handling circulating forecasts with time series or sequential data. However, this is where the specialty of the LSTM model comes in [24]. The model can keep memories in the long terms; thus, beating long-term dependency problems. LSTM uses the information generated from past's tasks to complete the current tasks; thereby, creating a perfect correlation and offering accurate predictions for water demand issues. Similarly, studies by [25, 26] supports this claim by illustrating models that used ANNs to predict water demand, using factors like temperature, rainfall occurrence, previous water demand rates and global climate change.

Although a number of models exist in the literature, to adopt a forecasting model it is important to identify the factors relevant to the region. Therefore, the subsequent section will focus on identifying factors affecting the water demand forecasting in the UAE

section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation as well as the experimental conclusions that can be drawn.

## 2.2. Factors Affecting Water Demand Forecasting In UAE

The section focusses on identifying the factors that are likely to impact on the water demand forecasting in the United Arab Emirates, by reviewing the previous existing case studies in the literature.

Mohamed and Al-Mualla conducted a study in 2010, in the city of Umm-Al Quwain to forecast the water demand for the next twenty-five years. The research used two databases for water consumption provided by Umm Al-Quwain Water Department from 1980 to 2007 and 2000 to 2007 respectively. The data includes several factors such as the population, temperature and rainfall. The results revealed that population was the only significant factor in both databases. In addition, the research was conducted in 4 scenarios that examined the effect of additional independent variables such as population, average annual income, unmetered water demand and the effect of mega projects on the water demand. For the first scenario, the research expected an increase of water demand of 50% by the year of 2015 and forecasted the water demand to double in 2025. For the average income, it expected that the demand will increase by 40% in the next ten years. For the remaining scenarios,



the results predicted that the demand will increase in the future as these variables are predicted to increase in the future [10].

Another study (Mohamed and Al-Mualla, 2010) in the emirate of Umm Al-Quwain uses the same IWR-MAIN software. It uses two databases containing water consumption for different sectors, population, rainfall and temperature data while the first database contains the average daily water consumption in Umm Al-Quwain from 1980 to 2000 and the second contains meter units' readings from 2000 to 2007. The results in database one predicted that the water demand will reach 20,000 m<sup>3</sup> and 30,000 m<sup>3</sup> per day in 2014 and 2030 respectively. While for database two the demands are forecasted to be 12,000 m<sup>3</sup> and 19,000 m<sup>3</sup> per day for 2014 and 2030 respectively [12].

Moreover, [20] adopts the same approach to forecast the long term demands for seven sectors : residential, agricultural, commercial, government, industrial, public services and non-metered services in Al Ain. The data adopted consists of the water consumption for Al-Ain in the seven sectors from 1998 to 2012 as well as other factors affecting the demand and they are population, temperature and rainfall data for the same period. The results revealed that the consumption of water in 2030 will be double of that in 2015 and the water consumption is highest in the month of august. It also revealed that commercial, government and non-metered services sectors have the largest consumption in august while agricultural, industrial, public services and residential sectors have highest consumption in October, June, April and March respectively, with a 45% increase was expected in 2030.

Therefore, it can be concluded from the previous water demand forecasting studies conducted in the UAE that the factors that impacts the water consumption are temperature, rainfall, humidity, Consumer Price Index (CPI), GDP and population.

Table 1 summarizes the factors affecting the water demand forecasting in the UAE, and classify them into three categories environmental, economic and social.

**Table 1. Factors Affecting Water Demand Forecasting, Source: Authors**

Sector	Factors
Environmental	Mean Temperature
	Mean Rainfall
	Relative Humidity
Economical	CPI
	GDP
Social	Population growth

Thus, the factors deduced from the literature aids the study to identify the significant predictors of water demand forecasting in the UAE by adopting regression analysis to explore the relationship between the factors and water demand.

In addition, literature reports on several studies that have adopted various model to forecast the water demand. However, each model displays a set of limitations in predicting the long term demand. Whereas, as indicated by [27], Long Short Term Memory (LSTM) model, a form of Recurrent Neural Networks (RNNs) is an advanced model. It is able to learn and memorize the data as it arrives as well as being able to predict texts and draw relationships between variables with substantial accuracy. Therefore, the intent of this study is to propose a Long Short Term Memory (LSTM) model on factors identified.

### 3. Methodological Steps

To address the challenges in the future owing to the depletion of natural water sources, exhaustion of fossil fuels, population growth, increasing urbanization and the climatic consequences of global warming in UAE. The study adopts a quantitative approach to propose an accurate water demand forecasting technique that incorporates all significant factors to predict the future water demands of the UAE. The methodological approach was based on three main steps;

**Step I** - The study obtains its primary data for water consumption from the Federal Competitiveness and Statistics Authority over the period of 10 ten years ranging from 2007 till 2017.

**Step II** – Regression analysis (Hypothesis Testing) is adopted to identify the significant factors that impact the water demand forecasting in UAE. A set of six hypothesis is drawn to determine the significant impact of temperature, rainfall, humidity, CPI, GDP and population growth to on water demand forecasting in the UAE.

**Step III** – A Long-Short Term Model (LSTM) is proposed to develop a forecasting model to that predicts the future water demand in the UAE.

#### 4. Data Collection and Analysis

This section describes the steps followed for data collection and analysis in order to predict the water forecasting demand in the UAE.

##### 4.1. Water Consumption Data in the UAE

The data of water consumption in the UAE was obtained from the Federal Competitiveness and Statistics Authority over the period of 10 ten years ranging from 2007 till 2017 [28]. The data considered several variables or factors, such as the installed capacity of desalinated water plants by entity, quantity of produced water by producer, percentage distribution of produced water quantity by source of water, quantity of used water, desalinated water quantity provided to authorities by provider entity, and water tariff by daily consumption slab rate and sector. It also reflects the association of water production and consumption across different sectors. The plants under investigation includes the department of energy Abu Dhabi, Dubai Electricity and Water Authority (DEWA), Sharjah Electricity and Water Authority (SEWA), and the Federal Electricity and Water Authority (FEWA). The data shows a significant increase of water consumption over the years, as consumption was quite low in 2007 as compared to 2017 as shown in Table 2.

**Table 2. Quantity of Used Water in United Arab Emirates, 2007-2017 (MCM) Source: [28]**

Year	Water Used
2007	1,365.0
2008	1,469.1
2009	1,554.3
2010	1,556.0
2011	1,581.4
2012	1,680.8
2013	1,714.1
2014	1,786.4
2015	1,847.2
2016	1,828.8
2017	1,829.5

Thus, it can be assumed that the increase may be an indication of a probable increase in the growth of GDP and population of the United Arab Emirates. \_

##### 4.2. Regression Analysis (Hypothesis Testing)



The study adopts regression analysis to deduce a causal relationship between the independent variables (rainfall, humidity, temperature, CPI, GDP and population growth) and dependent variable (water demand).

Firstly, it is important to prepare the data into a single scale using MATLAB software. As the meteorological attributes such as temperature, rainfall and humidity are given in months, while the other factors such as CPI, GDP, population growth and water consumption are expressed in years. Thus, it is necessary to use a single scale of measurement. Secondly, regression analysis will be used to present a causal relationship between a dependent variable and a collection of independent variables by adopting the hypothesis testing.

The significance of this step is to remove the redundant factors as such features could lead to an increase in errors in the model that is going to be trained/fitted. Therefore, based on the findings of the literature, the study draws 6 hypothesis to determine the significance of factors such as temperature, rainfall and humidity, CPI, GDP and population growth on the water demand in the UAE as shown in equation 2

$$Y(\text{Water Consumption}) = \beta_0 + a\beta_1 + b\beta_2 + c\beta_3 + d\beta_4 + e\beta_5 + d\beta_6 \quad (2)$$

Equation 2. Regression Analysis. Source: [16]

**Hypothesis 1:** *The change in temperature has a high impact on water demand in UAE.*

To determine the impact of temperature on water demand in UAE, regression analysis was conducted and plotted as seen in Figure 3. Thus, from the graph it can be seen that the changes in temperature does not presents a causal relationship with the water demand. It was found that R square value is equal to 0.001, which is lower than accepted limit of 0.5 and the p- value is greater than 0.05.

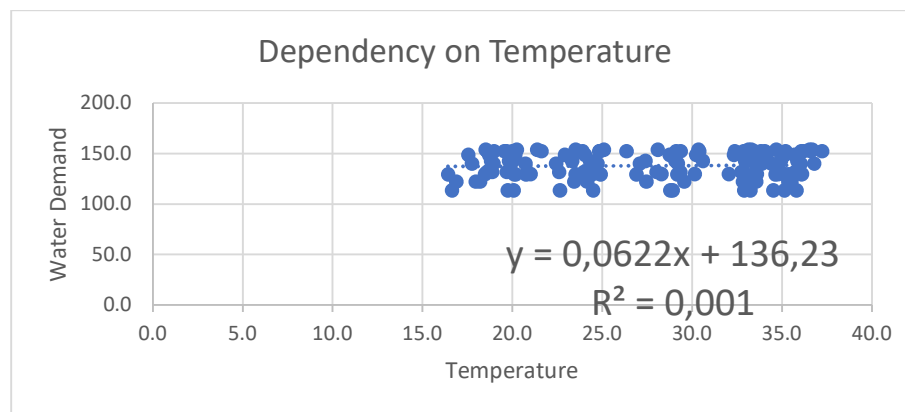


Figure 3. The relationship between Temperature and water demand. Source: Authors

Therefore, it can be concluded that changes in temperature do not significantly affect the changes in water demand, even though it still might be the cause of high-water demand in the UAE.

**Hypothesis 2:** *The change in rainfall amount has a high impact on water demand in UAE.*

The relationship between water demand and rainfall, can be seen in the figure 4, that represents the changes in the rainfall do not show the causal relationship with the water demand. The R square value is equal to 0.0012, which is lower than accepted limit of 0.5 and the p- value is greater than 0.05.

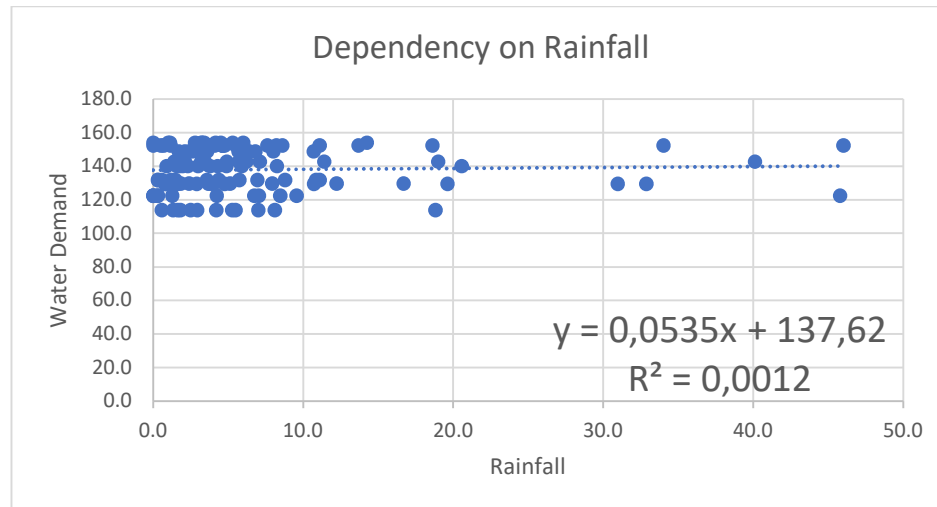


Figure 4. The relationship between Rainfall and water demand. Source: Authors

Therefore, it can be concluded that rainfall does not affect the changes in water demand significantly even though it might still be the cause of high-water demand in the UAE.

**Hypothesis 3:** *The change in humidity has a high impact on water demand in UAE.*

Figure 5 represents the relationship between humidity and the water demand. As it can be seen on the graph the changes in humidity percentage do not have strong causal relationship with the water demand. The R square value is equal to 0.0004, which is lower than accepted limit of 0.5 and the p-value is greater than 0.05.

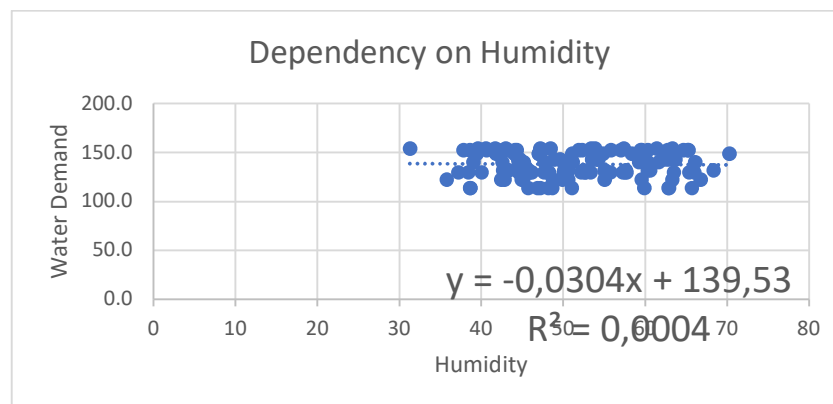


Figure 5. The relationship between Humidity and water demand. Source: Authors

Therefore, it can be concluded that humidity level does not affect the fluctuation of water demand significantly, even though it still might be the cause of high-water demand in the UAE.

**Hypothesis 4:** *The change in Gross Domestic Product has a high impact on water demand in UAE.*

Records of Water demand are plotted against GDP on Figure 6. As shown, the GDP changes alongside with water demand. This relationship is represented by the line with big positive slope. The R square value is equal to 0.853, which is higher than the accepted limit of 0.5 and the p-value is greater than 0.05.

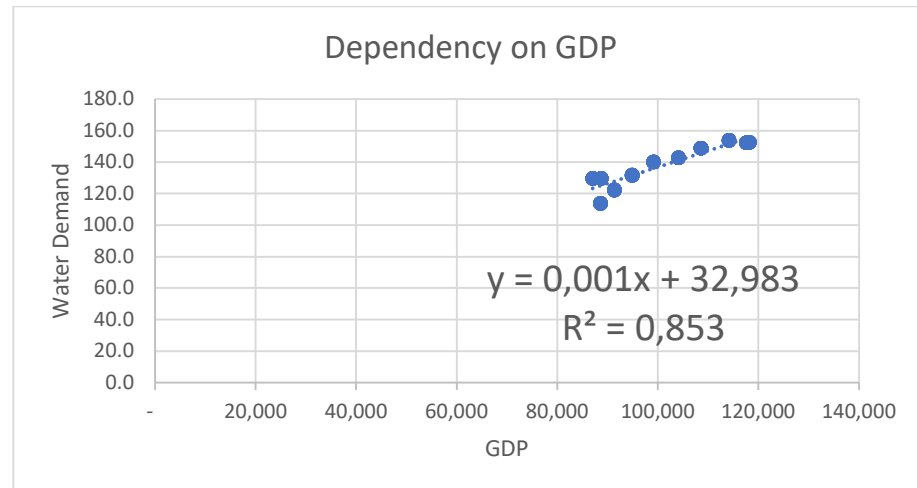


Figure 6. The relationship between GDP and water demand. Source: Authors

Therefore, it can be concluded that the changes in water demand happen along with the changed of GDP, even though GDP may not be the only cause of the high-water demand in the UAE.

**Hypothesis 5:** *The change in Consumer Price Index has a high impact on water demand in UAE.*

The relationship between CPI and the water demand can be seen in Figure 7. The graph represents a strong causal relationship between the changes in CPI values and the water demand, that is illustrated by the line with big slope. The R square value is equal to 0.9504, which is almost twice higher than the accepted limit of 0.5 and the p- value is less than 0.05.

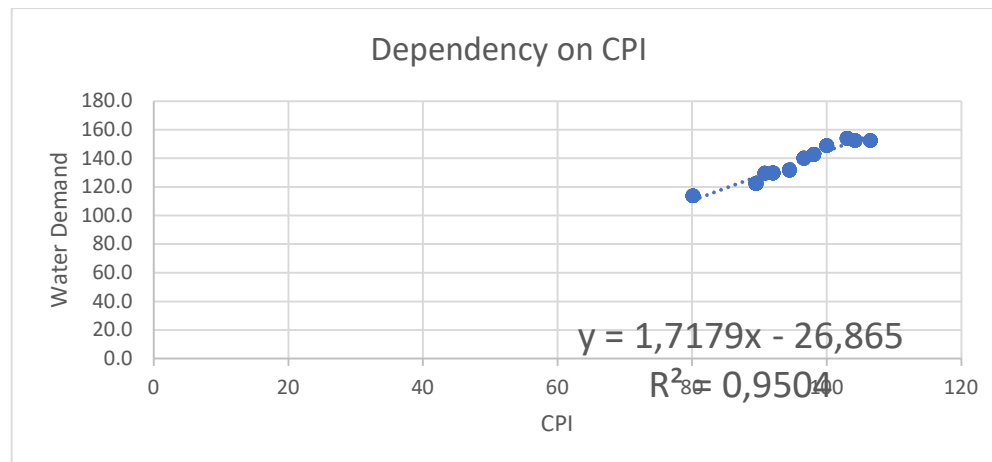


Figure 7. The relationship between CPI and water demand. Source: Authors

Therefore, it can be concluded that the changes in water demand happen along the changes of CPI factor, even though CPI may not be the only cause of the high-water demand in the UAE.

**Hypothesis 6:** *The change in Population Growth has a high impact on water demand in UAE.*

Figure 8 represents the relationship between the Population growth and the water demand. As it can be seen on the graph the changes in population growth rate do have strong causal relationship with the water demand that is represented by the line with big negative slope. The R square value is equal to 0.8637, which is almost twice higher than the accepted limit of 0.5 and the p value is less than 0.05.

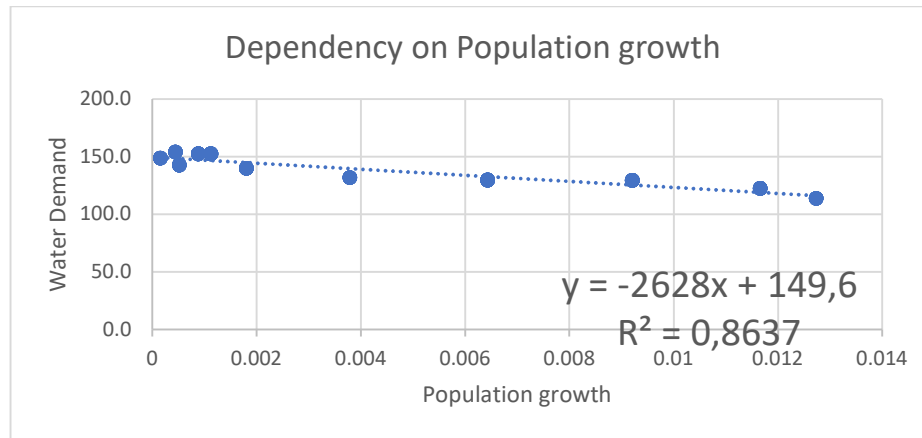


Figure 8. The relationship between Population growth and water demand. Source: Authors

Therefore, it can be concluded that the changes in water demand happen along the changes of population growth rate, even though population growth rate may not be the only cause of the high-water demand in the UAE.

In summary, CPI, GDP and population growth are the significant factors that impact water demand forecasting in UAE. Whereas factors such as temperature, humidity and rainfall were removed from the data as they have little or no effect on the water demand in the UAE.

#### 4.3. LSTM Model

The step aims to propose a forecasting model based on the Long-short Term Model (LSTM) taking into consideration the factors that have an impact on the water demand forecasting in the UAE.

##### 4.3.1. Data Restructuring

The pre-filtering Regression analysis (Hypothesis testing) stage aids in deducing average CPI, GDP and population growth as the only significant factors, thus a restructuring of the data is required. First, the prevailed attributes CPI, GDP and Population growth along with monthly total water consumption data were restructured as yearly values as they were originally obtained. To justify, these attributes were structured in monthly basis to comply with mean rainfall, relative humidity and mean temperature attributes that were structured in months and since the last were eliminated, annual values were calculated back. Accordingly, the month attribute was deleted. Therefore, the complete data set after restructuring includes Instance Number, year, average yearly CPI values, GDP values in years, yearly population growth and total yearly water consumption.

##### 4.3.2. Training

The stage aids to train the LSTM model based on the parameters considered as a significant predictors of water demand in the UAE. The training stage includes the following steps importing, normalizing, dividing the data to training and validation data, defining the model parameter and launching the fitting command. The first step is importing the restructured data where the columns like the instance number and the year, which is irrelevant for fitting the model will be eliminated. Secondly, the best practices for training neural networks is to normalize the data before fitting and training. As normalizing the data speeds up learning and aids faster convergence of error. Moreover, since the tanh activation function was used, it is more preferable to use normalized data because this will speed up the training [29]. After normalization, the data was divided into training and validation/testing data. Training data was used exclusively to train the model, while the testing data was used to compute the RMSE and thus evaluate the reliability of the model. A 9/11 percentage split was used. Since the database consists of 11 years (2007 to 2017), the years from 2007 to 2015 will be

used for training/fitting and 2016 - 2017 for testing. Next, the parameters discussed earlier were defined in MATLAB software such as *random number was set to 1, number of features as 4, number of hidden layers between 1 to 200, max training epochs to 250 and activation function as tanh*. The number of hidden layers varied between 1 to 200 as trial and error to find the best network architecture. Finally, after getting the data ready and defining the model parameters, the 'trainNetwork' function of Matlab was used to start the internal fitting/training algorithm to obtain file, which defines the finally fitted LSTM model as shown in figure 9. Network architectures between 1 and 200 hidden layers were fitted and the best architecture was chosen based on the RMSE value.

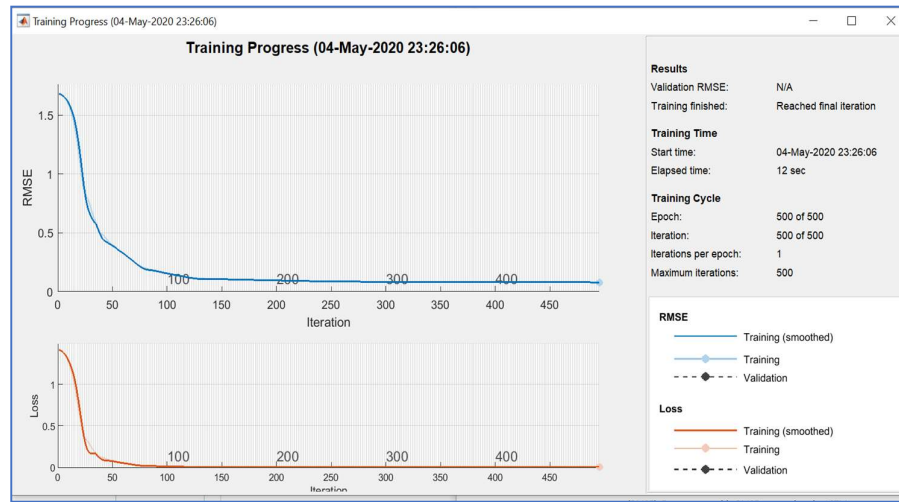


Figure 9. Training Curves for RMSE and Loss. Source: Authors

#### 4.3.3. LSTM Implementation

The LSTM is a model that uses an artificial recurrent neural network used to facilitate deep structural learning. However, different network architectures (number of hidden layers) lead to different models with varying levels of reliability. This stage, therefore, discusses the process that was used to test different network architectures before finally reaching a conclusion about which architecture is the most superior/ reliable. A 9/11 percentage split was used, since the database consists of 11 years (2007 to 2017), the years from 2007 to 2015 were used for training/fitting and 2016 and 2017 for testing. To evaluate the reliability of each model, the RMSE is calculated for both the model as a whole as well as the water demand. The RMSE of the model as whole is calculated by evaluating the error between predicted and actual for all 4 features (since all features are output due to the nature of the LSTM operating as a recurrent) as shown in equation 3 [29].

$$(RMSE|RMSE) = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}} \quad (3)$$

RMSE Equation, Source: [29]

The RMSE of the water demand is specific for the focus feature (water demand). This is calculated by measuring the RMSE between predicted and actual for water demand solely. To consider both RMSE, a weighted score for both RMSE was calculated. Each error component was assigned a weight of 0.5 (equal importance), and the sum product was obtained as a weighted score. For all the models, the weighted score RMSE was calculated and sorted starting with the model with the least error as shown in Figure 10.

	A	B	C	D	E
1	N hidden layers	Water demand RMSE	Model RMSE	Weighted score	
2	18	0.015094428	0.23135442	0.123224424	
3	14	0.006204872	0.24492233	0.125563601	
4	5	0.001441093	0.2698265	0.135633796	
5	17	0.050370008	0.22497737	0.137673689	
6	7	0.010679631	0.26781377	0.139246701	
7	19	0.035547644	0.24480891	0.140178277	
8	26	0.078197271	0.20332852	0.140762896	
9	12	0.038432747	0.24893227	0.143682509	
10	8	0.01526299	0.28415877	0.14971088	
11	25	0.085007459	0.21706831	0.151037885	
12	23	0.085495144	0.21740362	0.151449382	
13	6	0.070830256	0.23392841	0.152379333	
14	15	0.074789792	0.23061863	0.152704211	
15	16	0.055522949	0.2523925	0.153957725	
16	20	0.10301235	0.21417356	0.158592955	
17	21	0.10682037	0.21324751	0.16003394	
18	13	0.085105926	0.23531592	0.160210923	
19	28	0.11940554	0.20650202	0.16295378	
20	9	0.09126392	0.23547643	0.163370175	

Figure 10. Evaluating best architecture. Source: Authors

As seen in Figure 10, the architecture with 18 hidden layers has the least weighted RMSE score of approximately 0.12, thus considered the most superior/reliable. A graph representing the water demand, model and weighted RMSE was generated as seen below in Figure 11.

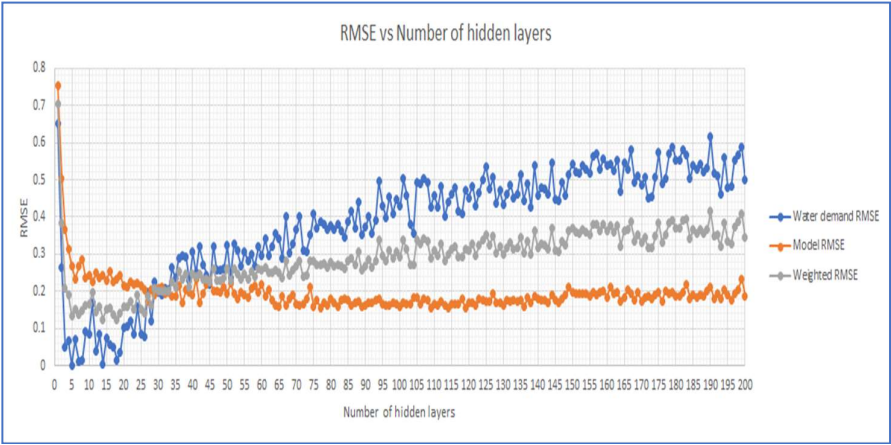


Figure 11 (a). RMSE vs number of hidden layers (full range). Source: Authors

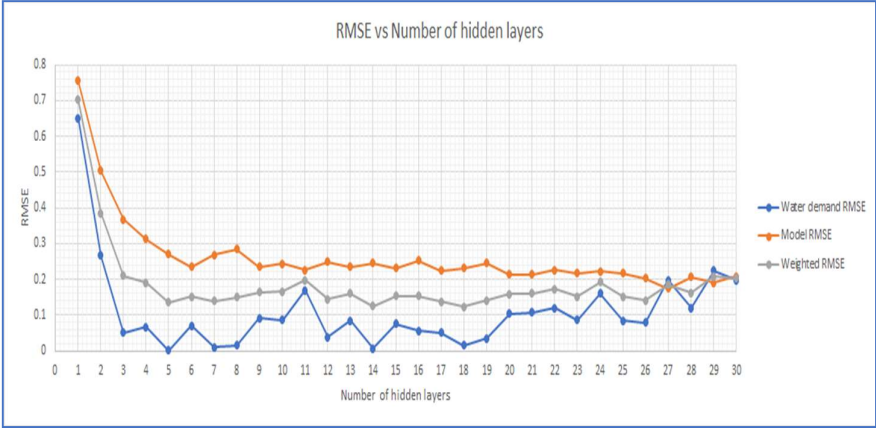


Figure 11 (b) RMSE vs number of hidden layers (1-30). Source: Authors



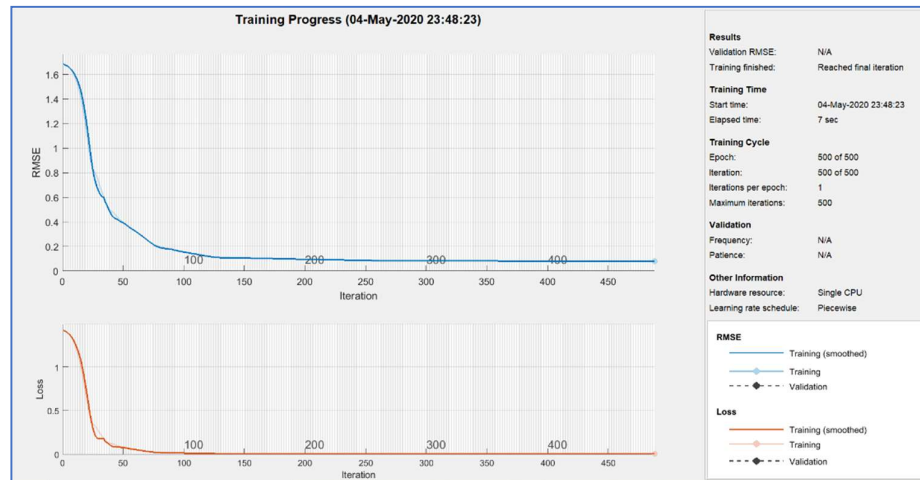


Figure 11 (c) Training curves for most superior architecture. Source: Authors

The blue line in Fig. 11 (a) represents the water demand RMSE, the orange line represents model RMSE and the grey line represents the weighted RMSE. As seen above, models between 4 and 30 hidden layers have the minimum errors in terms of water demand as well as the weighted error. For better visualization, emphasis is put on architectures between 1-30 hidden layers as shown below in Figure 11 (b). As a side note, RMSE values below 0.2 are considered highly reliable, while values below 0.5 are considered acceptable [30]. All architectures between 4 and 30 hidden layers are considered excellent because their weighted RMSE score is below 0.2. Nevertheless, the network architecture with 18 hidden layers proves to be the superior/reliable architecture. Furthermore, figure 11 (c) portrays the training process for the most superior architecture (18 hidden layers). The loss function diverged to almost zero and the RMSE was almost below 0.1. This shows that the training error was very low and that the training reliability is high.

Taking into consideration that training RMSE was below 0.1, model RMSE was 0.23, weighted RMSE 0.12 and most importantly the water demand RMSE 0.015, it can be concluded that the 18 hidden layers network architecture is highly superior and in fact the most superior from all other architectures. Therefore, the chosen LSTM model can be confidently considered to be reliable and accurate enough to be used later for forecasting the water consumption for the next 10 years.

#### 4.3.4. Generated Output

The model will be fully trained before being used for forecasting. As discussed previously, the available database contains 11 years of training data, thus, all 11 years (2007-2017) of training data was used in this step to fully train the LSTM model. To be specific, the chosen architecture (18 hidden layers) was trained again using all the previously discussed steps, but rather considered all data as training data with no data kept for validation. Finally, the chosen LSTM model was run for the next ten years from 2018 to 2027. Table 3 below shows the forecasted water consumption for years between 2018 and 2027.

Table 3. Forecasted water demand for next 10 years (2018 to 2027). Source: Authors

Year	Forecasted water consumption Million M <sup>3</sup>
2018	1821.66125488281
2019	1819.83398437500
2020	1818.40258789063
2021	1817.30456542969
2022	1816.19335937500
2023	1815.05126953125

2024	1813.84448242188
2025	1812.57812500000
2026	1811.25830078125
2027	1809.89636230469

After conducting the regression analysis and rejecting three null hypothesis out of six due to p value results, the LSTM model with 18 hidden layers and three factors as inputs was proven to have the least error among other models with different architectures. According to the predictions done using the chosen LSTM deep learning model, the water consumption in the UAE for the next ten years is expected to decrease. As shown in Table 3 and Figure 12, water consumption is decreasing from 1821.6 Million Cubic meters in 2018 to 1809.9 million cubic meter in 2027.

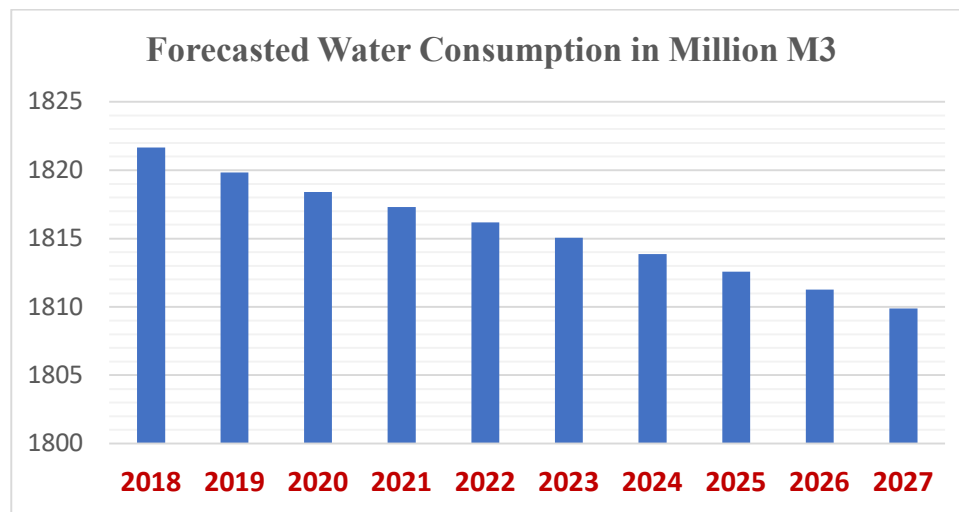


Figure 12 Forecasted Water consumption for the next 10 years. Source: Authors

By examining the restructured data set that was used to fit the model, it is inferred that although the Average CPI and GDP values were increasing and the population growth was positive, the water consumption values were decreasing for the last two years. To justify, the LSTM model gives a higher priority to recent data values, and by looking into the last two values for the water consumption that are for year 2016 and 2017, it is seen that they were decreasing. The decrease in the forecasted water consumption especially after 2016 can be justified by the great efforts put by the UAE government to raise awareness regarding sustainability, establishments of water saving programs as well as imposing regulations for illegal water use. According to [31], the UAE government had imposed a new regulation regarding illegal groundwater use and required all farms to install meters. Another initiative in Dubai called “Badia Farms” was launched in March 2018, which is a vertical farm that recycles the water used in farms and uses hydroponic technology and 95% less water compared to normal farms [32]. Moreover, in 2017 the ministry of Energy and Industry launched the UAE water security strategy that targets reducing the yearly water consumption per capita in the UAE to the half by 2036 [33]. Therefore, the decreasing trend shown in Figure 12 is attributed to factors that are more influential to water demand compared to CPI, GDP and population growth. Be that as it may, the model predicted the decrease in the water demand which proves the accuracy of the model.

## 5. Conclusion

The United Arab Emirates has been suffering from water stress due to the lack of surface water, lack of rainfall and the excessive pumping of groundwater. Thus, due to this it has shifted to desalination projects to help satisfy future demands which is costly in terms of maintenance and operation. Therefore, the need to have an accurate water demand forecasting technique was deemed essential to optimize the water management system and thereby tackling the research gap. Several

water demand forecasting techniques such as unit water demand analysis, univariate time series analysis, time series regression analysis, scenario-based analysis and decision support system analysis were identified and analyzed. However, it was found that the use of these techniques was hindered due to wrong assumption and lack of depth in understanding the socio-economic contexts. Therefore, a new model was introduced which is the LSTM model to forecast water demand in UAE. In addition, previous case studies of water demand forecasting in UAE were reviewed to identify the relevant factors such as mean temperature, mean rainfall, relative humidity, CPI, GDP and population growth. However, to identify the significant predictors of water demand forecasting, six hypotheses were formed, that aid in identify CPI, GDP and population growth as the significant factors affecting the water demand in UAE. The data from an open source provided from the Federal Competitiveness and Statistics Authority over the period of 2007 to 2017 was used to train the LSTM model using MATLAB software. The LSTM model was trained with 9/11 percentage split of the data (2007 -2015) and tested with 2/11 percentage split of the data (2016 – 20017). Several architectures were trained, tested and evaluated based on Root Mean Square Error and the architecture with 18 hidden layers was found to be the most accurate. The selected model with the best architecture was again trained for the whole data period from 2007 to 2017 and used to predict the future water demand in the UAE for the next ten years. The results of the LSTM model showed a decreasing trend from 1821 million m<sup>3</sup> in 2018 to 1809.9 million m<sup>3</sup> in 2027. The water demand was increasing in the data set from 2007 till 2015 but started to decrease for the last two years although the CPI, GDP and population growth were increasing for the whole period. It can be justified due to the model tendency to prioritize more recent values. The fact that the future water consumption showed a decreasing trend can be attributed to other factors not considered in the study and affected by reducing the water demand such as the 2016 government regulation on the agriculture sector, that prohibits the illegal connections and required all farms to install meter. Similarly, initiative to reuse the water in farms and the target of UAE to reduce the yearly water consumption to the half by 2036. Therefore, the results of these actions are manifested on the predicted water demand in the UAE for the next ten years and thereby the model results are considered accurate and achieve the aim of the study. The decrease of water demand is considered beneficial to the UAE as it is now possible for the government to reduce the efforts put on costly desalination plants and utilize the water resource according to the forecasted results.

**Author Contributions:** For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used “Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.”, please turn to the [CRediT taxonomy](#) for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

**Funding:** “This research received no external funding”.

**Conflicts of Interest:** “The authors declare no conflict of interest.”

## References

1. Al-Qawasmi, J., Asif, M., El Fattah, A.A. and Babsail, M.O., 2019. Water Efficiency and Management in Sustainable Building Rating Systems: Examining Variation in Criteria Usage. *Sustainability*, 11(8), p.2416. <https://doi.org/10.3390/su11082416>
2. Gosling, S.N. and Arnell, N.W., 2016. A global assessment of the impact of climate change on water scarcity. *Climatic Change*, 134(3), pp.371-385. <https://doi.org/10.1007/s10584-013-0853-x>
3. Boretti, A. and Rosa, L., 2019. Reassessing the projections of the world water development report. *NPJ Clean Water*, 2(1), pp.1-6. DOI:[10.1038/s41545-019-0039-9](https://doi.org/10.1038/s41545-019-0039-9)
4. Ait-Kadi, M., 2016. Water for development and development for water: realizing the sustainable development goals (SDGs) vision. *Aquat. Procedia*, 6, pp.106-110. doi: 10.1016/j.aqpro.2016.06.013
5. Gain, A.K., Giupponi, C. and Wada, Y., 2016. Measuring global water security towards sustainable development goals. *Environmental Research Letters*, 11(12), p.124015. doi: 10.1088/1748-9326/11/12/124015
6. Yalçıntaş, M., Bulu, M., Küçükvar, M., & Samadi, H. (2015). A framework for sustainable urban water management through demand and supply forecasting: The case of Istanbul. *Sustainability*, 7(8), 11050-11067. <https://doi.org/10.3390/su70811050>
7. Banadkooki, F. B., Ehteram, M., Ahmed, A. N., Fai, C. M., Afan, H. A., Ridwam, W. M., ... & El-Shafie, A. (2019). Precipitation forecasting using multilayer neural network and support vector machine optimization based on flow regime algorithm taking into account uncertainties of soft computing models. *Sustainability*, 11(23), 6681. <https://doi.org/10.3390/su11236681>
8. Alsharhan A.S., Rizk Z.E. (2020) Water Resources and Water Demands in the UAE. In: Water Resources and Integrated Management of the United Arab Emirates. World Water Resources, vol 3. Springer, Cham. [https://doi.org/10.1007/978-3-030-31684-6\\_27](https://doi.org/10.1007/978-3-030-31684-6_27)
9. Govindan, K., Shankar, K.M. and Kannan, D., 2016. Sustainable material selection for construction industry—A hybrid multi criteria decision making approach. *Renewable and Sustainable Energy Reviews*, 55, pp.1274-1288. <https://doi.org/10.1016/j.rser.2015.07.100>
10. Murad, A.A., Al Nuaimi, H. and Al Hammadi, M., 2007. Comprehensive assessment of water resources in the United Arab Emirates (UAE). *Water Resources Management*, 21(9), pp.1449-1463. doi: 10.1007/s11269-006-9093-4.
11. Shahin, S.M. and Salem, M.A., 2015. The challenges of water scarcity and the future of food security in the United Arab Emirates (UAE). *Nat Resour Conserv*, 3(1), pp.1-6. doi: 10.13189/nrc.2015.030101.
12. Mohamed, M.M. and Al-Mualla, A.A., 2010. Water demand forecasting in Umm Al-Quwain (UAE) using the IWR-MAIN specify forecasting model. *Water Resources Management*, 24(14), pp.4093-4120. DOI: <https://doi.org/10.1007/s11269-010-9649-1>
13. Rinaudo, J.D., 2015. Long-term water demand forecasting. In *Understanding and managing urban water in transition* (pp. 239-268). Springer, Dordrecht. doi: 10.1007/978-94-017-9801-3\_11
14. Arunkumar, M. and Mariappan, V.N., 2011. Water demand analysis of municipal water supply using epanet software. *International Journal on Applied Bioengineering*, 5(1), pp.9-19., doi: 10.18000/ijabeg.10072
15. Chen, J. and Boccelli, D.L., 2014. Demand forecasting for water distribution systems. *Procedia Engineering*, 70, pp.339-342. doi: <https://doi.org/10.1016/j.proeng.2014.02.038>.
16. Lawrence, K.D., 2019. *Robust regression: analysis and applications*. Routledge.
17. Wu, Li & Zhou, Huicheng. (2010). Urban water demand forecasting based on HP filter and fuzzy neural network. *Journal of Hydroinformatics - J HYDROINFORM*. 12. Doi: 10.2166/hydro.2009.082.
18. Shang, X., Jiang, X., Jia, R. and Wei, C., 2019. Land use and climate change effects on surface runoff variations in the upper Heihe River basin. *Water*, 11(2), p.344. <https://doi.org/10.3390/w11020344>
19. Mohamed, M.M. and Al-Mualla, A.A., 2010. Water demand forecasting in Umm Al-Quwain using the constant rate model. *Desalination*, 259(1-3), pp.161-168. doi: <https://doi.org/10.1016/j.desal.2010.04.014>
20. Younis, H.I., 2016. Water Demand Forecasting In Al-Ain City, United Arab Emirates. Available at: [https://scholarworks.uaeu.ac.ae/all\\_theses/188](https://scholarworks.uaeu.ac.ae/all_theses/188)
21. Paul, I.J.L., Sasirekha, S., Vishnu, D.R. and Surya, K., 2019, April. Recognition of handwritten text using long short term memory (LSTM) recurrent neural network (RNN). In *AIP Conference Proceedings* (Vol. 2095, No. 1, p. 030011). AIP Publishing LLC. <https://doi.org/10.1063/1.5097522>
22. Zhang, J., Zhu, Y., Zhang, X., Ye, M. and Yang, J., 2018. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *Journal of hydrology*, 561, pp.918-929. <https://doi.org/10.1016/j.jhydrol.2018.04.065>

23. Park, S. and Kim, Y., 2019, June. A Method for Sharing Cell State for LSTM-Based Language Model. In *International Conference on Intelligence Science* (pp. 81-94). Springer, Cham. Doi: [https://doi.org/10.1007/978-3-030-25213-7\\_6](https://doi.org/10.1007/978-3-030-25213-7_6)
24. Rahman, L., Mohammed, N. and Al Azad, A.K., 2016, September. A new LSTM model by introducing biological cell state. In *2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)* (pp. 1-6). IEEE. DOI: [10.1109/ICEEICT.2016.7873164](https://doi.org/10.1109/ICEEICT.2016.7873164)
25. House-Peters, L., Pratt, B. and Chang, H., 2010. Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in hillsboro, oregon1. *JAWRA Journal of the American Water Resources Association*, 46(3), pp.461-472. doi: 10.1029/2010WR009624.
26. Adamowski, J., Fung Chan, H., Prasher, S.O., Ozga-Zielinski, B. and Sliusarieva, A., 2012. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*, 48(1). Doi: <https://doi.org/10.1029/2010WR009945>
27. Van, L.P., De Praeter, J., Van Wallendael, G., De Cock, J. and Van de Walle, R., 2015, January. Machine learning for arbitrary downsizing of pre-encoded video in HEVC. In *2015 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 406-407). IEEE. DOI: [10.1109/ICCE.2015.7066464](https://doi.org/10.1109/ICCE.2015.7066464)
28. Fcsa.gov.ae, 2020. "Home",. [Online]. Available: <https://fcsa.gov.ae/en-us>. [Accessed: 07- Apr- 2020].
29. Gerardnico.com, 2020, "Data Mining - Root mean squared (Error|Deviation) (RMSE|RMSE) [Gerardnico - The Data Blog]". Available: [https://gerardnico.com/data\\_mining/rmse](https://gerardnico.com/data_mining/rmse). [Accessed: 04- May- 2020].
30. ResearchGate, 2020. "What's the acceptable value of Root Mean Square Error (RMSE), Sum of Squares due to error (SSE) and Adjusted R-square?", [Online]. Available: [https://www.researchgate.net/post/Whats\\_the\\_acceptable\\_value\\_of\\_Root\\_Mean\\_Square\\_Error\\_RMSE\\_Sum\\_of\\_Squares\\_due\\_to\\_error\\_SSE\\_and\\_Adjusted\\_R-square?](https://www.researchgate.net/post/Whats_the_acceptable_value_of_Root_Mean_Square_Error_RMSE_Sum_of_Squares_due_to_error_SSE_and_Adjusted_R-square?) [Accessed: 07- Apr- 2020].
31. Bardsley, D., "UAE making major efforts to overcome water conservation's 'many challenges'," <https://www.thenational.ae>, 02-Feb-2018. [Online]. Available: <https://www.thenational.ae/uae/uae-making-major-efforts-to-overcome-water-conservation-s-many-challenges-1.692196>. [Accessed: 05-May-2020].
32. Stanley, I. "How the UAE secures its future water needs," Events – Gulf News, 07-Nov-2018. [Online]. Available: <https://gulfnews.com/going-out/events/how-the-uae-secures-its-future-water-needs-1.2192231>. [Accessed: 05-May-2020].
33. The UAE Water Security Strategy 2036 - The Official Portal of the UAE Government. [Online]. Available: <https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/federal-governments-strategies-and-plans/the-uae-water-security-strategy-2036>. [Accessed: 05-May-2020].