Short-term forecasting of monthly water consumption in hyper-arid climate using recurrent neural networks

Abdullah A. Alsumaiei

Civil Engineering Department, College of Engineering and Petroleum, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait
alsumaiei.a@ku.edu.kw

ABSTRACT

Freshwater supply is a major challenge in regions with limited water resources and extremely arid climatic conditions. The objective of this study is to model the monthly water demand in Kuwait using the nonlinear autoregressive with exogenous input (NARX) neural network approach. The country lacks conventional surface water resources and is characterized by extremely arid climate. In addition, it has one of the fastest growing populations. In this study, linear detrending is performed on the water consumption time series for the period from January 1993 to December 2018 to eliminate the influence of population growth before application to the NARX model. Monthly temperature data are selected as exogenous input to the NARX model, because they are strongly associated with the water consumption data. Correlation analyses are performed to determine the input and feedback delays of the NARX model. The results demonstrate that the recurrent NARX model is efficient and robust for forecasting the short-term water demand, with a Nash-Sutcliffe (NS) coefficient of 0.837 in the validation period. Seasonal model assessment shows that the model performs best during the critical summer season. The NARX-based recurrent model is established as a powerful and promising tool for predicting urban water demand. Thus, it can efficiently aid the development of resilient water supply plans.

Keywords: Water supply modeling; NARX model; Linear detrending; Water footprint; Developing countries.

List of Abbreviations:

Nonlinear Autoregressive with Exogenous Input (NARX)
Autoregressive Moving Average (ARMA)
Autoregressive Integrated Moving Average (ARIMA)
Genetic Algorithm (GA)
Support Vector Machine (SVM)
Artificial Neural Network (ANN)
General Regression Neural Network (GRNN)
Autoregressive (AR)
Million Imperial Gallons (MIG)
Feedforward Neural Network (FNN)
Recurrent Neural Network (RNN)
Autoregressive Models with Exogenous Input (ARX)
Levenberg–Marquardt (LM)
Nash–Sutcliffe (NS)

1. INTRODUCTION

Securing sustainable water supply is vital for human life. Global water adequacy is vulnerable to threats induced by rapid population growth and climate change (Dawadi and Ahmad, 2013; Vörösmarty et al., 2000; Wang et al., 2016). Water scarcity has constrained economic development in many countries throughout the world (Blignaut and Van Heerden, 2009; Zhou et al., 2013). The anthropogenic stress on water resources is a major concern, particularly in regions with limited water resources and high population concentration. Hence, addressing the water security problem in such regions is essential and of high priority. The development of computational tools for forecasting water consumption can aid the development of resilient plans to cope with the anticipated stresses that can affect water availability in these regions.

Water consumption forecasting has been modeled using various techniques including linear and nonlinear methods. Among them, the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) methods are the most popular linear time series tools for water consumption modeling (Adamowski, 2008; Khashei and Bijari, 2011; Zhang, 2003). These linear methods are easy to use and interpret, and often do not require detailed data inputs. However, nonlinear methods have gained greater popularity due to higher efficiency in detecting complex trends and pattern recognizing capability. These methods include evolutionary computing algorithms such as the genetic algorithm (GA), support vector machine (SVM), and artificial neural network (ANN) that are among the most extensively used nonlinear methods in this field. The algorithms operate similar to the human brain and are considered robust soft computing modeling tools. Among the various families of ANNs, the nonlinear autoregressive with exogenous input (NARX) network has been applied for modeling water consumption (Banihabib and Mousavi-Mirkalaei, 2019; Bata et al., 2020; Mousavi-Mirkalaei and Banihabib, 2019).

NARX-based models have demonstrated remarkable performances in predicting short-term water demand, with significant superiority over their counterpart linear models (Bata et al., 2020; Cadenas et al., 2016). The efficiencies of several NARX models in predicting water demands were examined for one-day to four-months advance forecasts at the municipal water supply system in Ontario, Canada. The models were forced with different combinations of exogenous meteorological inputs to assess their impact in improving models' predictions. The results indicated that the air temperature had significant influence in enhancing the NARX-model performance. The four-months NARX model exhibited the highest efficiency in water consumption prediction in terms of error reduction, compared to the shorter temporal scale ones, with the same input data. The study highlighted that consumer behavior can challenge NARX applicability and recommended the recalibration of NARX models in compliance with the targeted temporal scale to obtain reliable forecasts.

Population growth is a major influencer on water consumption. Recently, Banihabib and Mousavi-Mirkalaei investigated NARX applicability along with the ARIMA model at an arid study site in Tehran, Iran. This study was motivated by the fast-growing population and hot climate, which considerably stressed the urban water demand at the study site. Base prediction models for water demand were constructed based on the meteorological variables. These models were then improved by including additional inputs such as sunshine hours and population growth rates. The study concluded that the impact of population growth was better modeled by the NARX model than the ARIMA (Banihabib and Mousavi-Mirkalaei, 2019) due to the nonlinear pattern of population growth. Hence, the NARX approach was recommended for modeling water consumption at similar sites. Moreover, the study recommended investigating the applicability of NARX at other locations experiencing rapid population growth under arid climatic conditions.

The Arabian Peninsula is one of the largest regions suffering from serious water shortage because of hyper-arid climatic conditions as well as the absence of perennial surface water bodies. The region has witnessed rapid population growth induced by economic development in the past few decades due to the increase in oil production since the middle of the last century. However, this economic development is adversely restricted by limited water resources (Tirth, 2020). The implications of this development on water consumption are serious and need careful attention (Odhiambo, 2017). Most countries in this region have extremely hot and dry climates, with the exception of narrow

coastal areas on the mountain belt. Water shortage has been always a major concern in this region and has substantially limited the livelihood of inhabitants. Currently, seawater desalination plants supply these countries with freshwater; however, this is environmentally problematic and expensive (Odhiambo, 2017). Hence, countries in this region need to enforce new water management policies to secure sustainable water resources.

Unfortunately, only few studies have focused on modeling the water consumption in this region. Previous efforts were mainly focused on employing probabilistic time series techniques or constant rate models (Almedeij and Alotaibi, 2017; Almutaz et al., 2012; Alotaibi and Almedeij, 2007; Mohamed and Al-Mualla, 2010a). Considering the harsh climatic conditions in this region and the rapid population growth, ANN-based methods can be applied as robust modeling tools for modeling the water consumption demand under such conditions. Al-Zahrani and Abo-Monasar assessed the suitability of the ANN in modeling the urban water demand at Al-Khobar city in Saudi Arabia (Al-Zahrani and Abo-Monasar, 2015). They applied the general regression neural network (GRNN) and autoregressive (AR) linear models for forecasting the daily water demand. Initially, the GRNN performance was assessed separately, with meteorological variables including the air temperature, rainfall, wind speed, and relative humidity. The study concluded that air temperature was the main water-demand predictor. The GRNN was then combined with the AR model, and significant improvement was reported; the AR model added autoregression, considered as another strong predictor of water consumption, to the forecasts.

Based on previous modeling studies for water demand forecasting in arid climate, it is argued that air temperature and the previous time step water consumption are the main predictors for water demand. However, the suitability of NARX for water consumption prediction in hyper-arid climates has not been comprehensively investigated. In view of the above, this study aims to investigate the potential of NARX-based models for predicting water demand in such harsh climatic conditions. NARX is a recurrent form of ANNs, which is mainly driven by the previous time value of the desired target variable (e.g., water demand). Furthermore, forcing meteorological variables to NARX, through exogenous input, results in strong water-demand prediction capability.

This paper presents a case study on Kuwait for testing NARX applicability in hyper-arid climatic regimes. The monthly water demand, as well as the meteorological variables, were obtained from the Ministry of Electricity and Water (MEW) for the period from January 1993 to December 2018, for building and testing NARX models. By examining NARX applicability in a unique climatic regime that has not been previously investigated, the results of this study can contribute to advancing water demand forecasting techniques. The country relies mainly on desalinated seawater, which hardly secures the increasing demand for freshwater. Thus, building reliable quantitative predictive models is vital for developing water resource plans. This can be applied to countries with similar climatic and economic conditions as well.

2. STUDY AREA

Kuwait is a small country occupying a total land area of 17,818 km^2, in western Asia. Located in the northwest of the Arabian Gulf, it is bounded by Iraq to the north, Iran to the northeast, and Saudi Arabia to the south. The climate is extremely arid with a long-term average precipitation of 115.2 mm and annual evaporation depths exceeding 3500 mm. As it is a desert country with subtropical climate, the summer months are extremely hot and dry with the maximum day temperatures exceeding 50 °C. The long summer season is generally punctuated by dust storms. On the other hand, the winter season is short and mild, and the recorded temperature rarely drops below 0 °C. The geography of the country is flat, with the highest point in Mutla ridges 306 m above the mean sea level. Generally, the topography gently slopes from west toward the coast. Due to these topographic features, the spatial variation of the climatic conditions is minimal. The meteorological department of the state collects daily meteorological data through 16 weather stations; however, the data from the Kuwait International Airport (KIA) station is generally used for hydrometeorological analysis because of the substantial continuity of data. In this study, daily temperature data were collected for the KIA station from Meteorological Department at the Directorate General of Civil Aviation for a period ranging from 1993–2018.

The country has a developing economy that is mainly based on the petroleum industry. Kuwait maintains one of the largest oil reserves in the world and produces 2.85 million barrels per day. The economy has witnessed huge expansion in the past decades in various sectors. Besides the oil industry, the contribution of the other local industries to the national GDP is minimal and does not exceed 10%. Due to economic development, the population has grown from 1.53 million in 1993 to 4.62 million by the end of 2018, which is one of the highest population growth rates in the world.

Securing of freshwater supplies in accordance with the increasing demand is a serious challenge to the state. Figure 1 demonstrates the historical growth in freshwater demand and the corresponding population growth. The evident fluctuation in the water consumption data is attributed to the seasonality in water demand. During the hot summer season, the demand for freshwater increases mainly due to the increase in the household consumption of freshwater. With no surface water resources, the country relies on seawater desalination to secure this demand. Based on data obtained from the Ministry of Electricity and Water, the country operates eight desalination plants producing a total of 614.8 million imperial gallons (MIG) per day. Table 1 lists the desalination plants in the country and their production capacity (MEW, 2018). These plants supply the majority of the water demand. Being one of the largest oil exporters in the world, the state can supply these plants with the necessary energy. Further, the government subsidizes the cost of desalination and bears more than 90% of the production cost. Freshwater supply subsidy has made the per capita water consumption in Kuwait among the highest in the world (Hamoda, 2001).
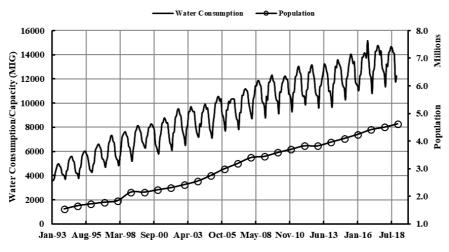

Figure 1. Historical water demand and population growth in Kuwait.

Table 1. Desalination plants in Kuwait (MEW, 2018).

Desalination Plant	Production Capacity in MIG/day (2017)
Al-Shuwaikh	49
Al-Shuaibah (North)	45
Al-Shuaibah (South)	30
Al-Doha (East)	42
Al-Doha (West)	110.4
Al-Zour (South)	140.4
Al-Zour (North)	107
Al-Sabiya	100
Total	**614.8**

3. METHODS

3.1 Recurrent NARX model

3.1.1 Model architecture

NARX neural network is a recurrent model considered as a special class within the ANN family. Inspired by biological neurons, ANNs are composed of nodes connected through artificial neurons in a particular manner. As part of the training, neurons are assigned weights to search for the optimal fit corresponding to the network's input/output data. ANNs are categorized in accordance with the information flow direction (Hydrology, 2000). In feedforward neural networks (FNNs), the nodes have a layered arrangement; the input data enters the input layer and processed through the hidden layers towards the last output layer. Thus, information flow in FNNs is unidirectional; however, in recurrent neural networks (RNNs), the information flow is bidirectional. The RNN output is recycled as the input for the following time step.

Recurrent NARX models are considered as a special form of the RNN. The basic postulate of the NARX model involves relating the current output of a time series to the past output of the same series and to the past/current outputs of an external (exogenous) time series. This renders NARX networks powerful for modeling nonlinear systems (Menezes Jr and Barreto, 2008; Siegelmann et al., 1997). NARX networks are a combination of the ANN and ARX, which is a common statistical tool for analyzing and modeling time series. Together, these NARX elements enable the detection of nonlinear behavior within the modeled data. A schematic illustration of the NARX structure is depicted in Figure 2.
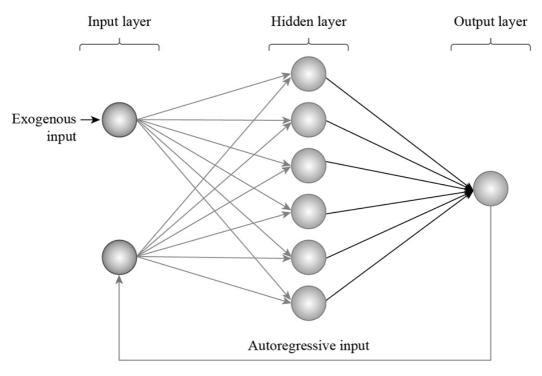

Figure 2. Schematic illustration of the NARX structure.

3.1.2 Training algorithms

For training the NARX network, the Levenberg–Marquardt (LM) algorithm was selected in this study. This algorithm has been frequently used for NARX training in various applications (Goyal et al., 2014; Ngia and Sjoberg, 2000; Rahimi et al., 2018). It represents an interpolation of the most advantageous elements of the steepest descent and Gaussian-Newton methods (Samarasinghe, 2016). The algorithm iteratively runs a search for the function minima and produces the optimal solution. In this study, MATLAB software was used to build and train the NARX models.

3.2 Linear detrending

The first stage of computing in the proposed recurrent model involves the detrending of the water consumption data. With the increase in population, the water consumption data exhibits an increasing trend. It is difficult to estimate the climatic impact on the water consumption pattern in the presence of the long-term trend. Detrending is applied for eliminating the impact of external influences on the data. Thereby, the detrending procedure effectively improves the NARX model efficiency. An additional advantage of the detrending procedure for raw water consumption data is that it guarantees the specification of significant autocorrelation lags because a highly trended timeseries has high likelihood of significant autocorrelation (Aldrich, 1995). For water consumption modeling, the linear detrending procedure is generally accepted (Altunkaynak and Nigussie, 2018). The least squares criterion is employed to fit a linear regression model with the water consumption data. The detrended water consumption signal can be then applied to the NARX model. The fitted trend is added back to the simulated water consumption outputs in order to reflect the influence of population growth on the water consumption pattern.

3.3 Correlation analysis

A fundamental feature of autoregressive models is the utilization of autocorrelation within the time series to make future predictions. The correlogram test is employed to examine the existence of autocorrelation within the water consumption time series. The aim here is to perform this analysis for the detrended water consumption time series to assess the level to which the adjacent values are correlated. The period of temporal separation is referred to as the lag and represented by the Greek symbol τ.

Initially, standardization of the detrended water consumption time series is computed as follows:

$$Z_j = \frac{C_j - \mu}{\sigma} \tag{1}$$

where C is the detrended water consumption data for month j, μ is the water consumption time series arithmetic average, and σ is the water consumption time series standard deviation. Further, the correlogram r is calculated for standardized data Z_j at lag τ as follows:

$$r_\tau = \frac{n}{n - \tau} \frac{\sum_{j=1}^{n-\tau}(Z_j - Z)(Z_{j+\tau} - \bar{Z})}{\sum_{j=1}^{N}(Z_j - \bar{Z})^2} \tag{2}$$

where $\bar{Z}$ is the arithmetic mean of the n value sample for standardized series Z_j. The correlogram is bounded between -1 and +1. The upper and lower lines are drawn on the correlogram to represent the confidence bounds (C.B.). The confidence lines define the significance level corresponding to $\alpha = 0.05$.

$$C.B. = \pm \frac{z_{1-\alpha/2}}{\sqrt{n}} \tag{3}$$

where z is the quartile function for the standard normal distribution. Based on equation 3, if the computed correlogram exceeds either the upper or lower confidence bounds, significant autocorrelation exists in the data at the designated lag time. Furthermore, cross correlation analysis was performed to examine the appropriateness of the temperature time series selection for exogenous input to the NARX model. The linear dependency between the detrended water consumption and temperature time series was measured using the Pearson correlation. The Pearson correlation statistical metric objectively indicates the robustness of the association of the detrended water consumption data with the selected external input for the recurrent model. In this study, Minitab software was used to perform the correlation analysis.

3.4 Model evaluation and validation

The recurrent model was validated using the customary approach, which involves the chronological separation of the water consumption data into validation and training periods. To ensure reliable performance, the NARX model was constructed with 80% of the data for model training and 20% data for model validation. Both periods were statistically evaluated using three efficiency measures to assess the model reliability. The coefficient of determination (R^2), which measures the strength of association between the simulated and observed targets, was the first metric to be applied. The value of R^2 ranges between 0–1, with 0 indicating the absence of statistical association and 1 indicating perfect correlation between the observed and modeled water consumption data. As R^2 represents the square of the commonly used Pearson correlation coefficient, it is not a direct measure of the prediction accuracy (as given by the Pearson correlation) but an assessment of the accuracy of any predictor built with the model. R^2 can be calculated for both observed and simulated water consumption data values as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(Co_i - \overline{Co})(Cs_i - \overline{Cs})}{\sqrt{\sum_{i=1}^{n}(Co_i - \overline{Co})^2}\sqrt{\sum_{i=1}^{n}(Cs_i - \overline{Cs})^2}}\right)^2 \tag{4}$$

where Co_i is the measured water consumption for month i, Cs_i is the simulated water consumption for month i, $\overline{Co}$ is the mean of the observed water consumption data, $\overline{Cs}$ is the mean of the simulated water consumption data, and n is the number of months within the validation period.

The second statistical measure applied was the MAE, which provides an objective quantification of the variation of the simulated targets from the observations according to the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Co_i - Cs_i|. \tag{5}$$

The Nash–Sutcliffe (NS) coefficient, which is a widely accepted hydrological model performance metric, was the third employed criterion (Nash and Sutcliffe, 1970). This coefficient ranges from $-\infty$–1; when the value is lower than zero, the observed timeseries mean is a superior predictor to the simulated values (Krause et al., 2005). Employment of the NS for model assessment guarantees that the modeled system variability has adequate representation at any point, whenever the NS values exceed zero. The equation for NS calculation is as follows:

$$NASH = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(Co_i - Cs_i)^2}{\frac{1}{m}\sum_{i=1}^{m}(Co_i - \overline{Co})^2} \tag{6}$$

where m is the number of observed months for the water consumption time series.

4. RESULTS AND DISCUSSION

4.1 Recurrent model application

Preliminary analysis on the water consumption data commenced with detrending the water consumption time series. The results of the linear detrending procedure are shown in Figure 3. The raw water consumption time series exhibits an obvious increasing trend induced by population growth. Detrending the water consumption data ensures that the computed correlogram reflects significant data autocorrelation by excluding the influence of the long-term trend. Moreover, this reduces the inputs to the NARX model, eventually minimizing the computational cost and avoiding overfitting. The detrended water consumption signal was then tested for autocorrelation, and cross correlation analysis was performed with the temperature data. Figure 4 displays the correlation analysis results. Figure 4A shows the computed correlogram for the water consumption remainder signal. The depicted upper and lower boundaries define the significance limits for autocorrelation. It is evident that the water consumption reminder signal has significant autocorrelation up to the 2nd lag. Thus, the feedback delay term in the NARX model was selected as two.
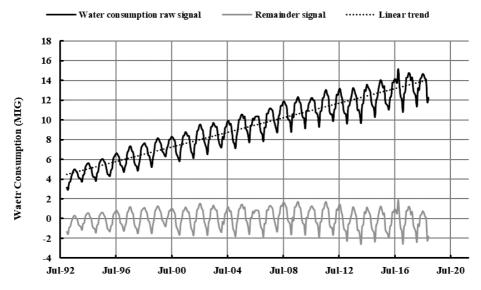

Figure 3. Results of linear detrending procedure application to raw water consumption data.

This implies that water consumption in the previous two months will be used for predicting the water demand for the current month. The lag term of two months for the detrended water consumption data proves that autoregressive models can used for short-term forecast of the water demand. Figure 4B indicates that the water consumption reminder signal is strongly correlated with the temperature data series, with a correlation coefficient of +0.89. The direction of association is positive, indicating that peak water consumptions occurs in the summer months with the increase in temperature. Hence, an input delay of zero was selected for exogenous input to the NARX model. The hidden-layer number was then manually calibrated by optimizing the statistical metrics values to achieve the best model performance, and the NARX performance was assessed at each run. In this study, the number of hidden layers that produce the best efficiency in terms of the statistical measures was found to be 30. Table 2 summarizes the NARX model parameters used in this study.
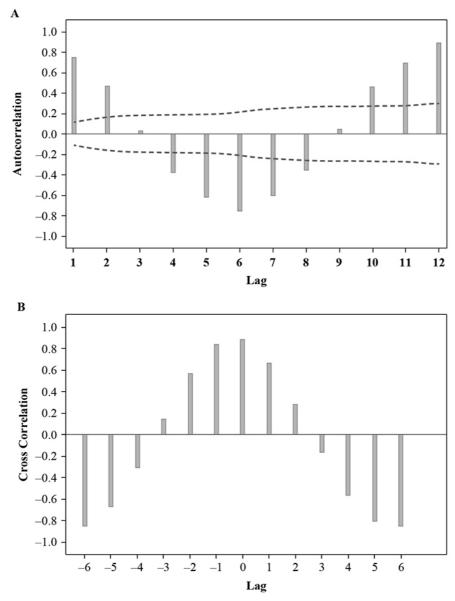
Figure 4. Correlation analysis results. **A)** The computed correlogram for the water consumption remainder signal.
B) Water consumption-temperature data cross correlation analysis.

The statistical evaluation measures for the calibrated NARX model are listed in Table 3. To avoid subjective evaluation, it important to consider all the performance efficiency indicators. The NS coefficients are 0.996 and 0.927 for the training and validation periods, respectively. These high NS values emphasize the soundness of the NARX model in representing water-demand variability. The other statistical measures support the model efficiency findings for forecasting. The correlation coefficient and coefficient of determination values are greater than 0.80 for the training as well as validation periods, indicating that the NARX model can be used for water-demand forecasting with high confidence.

Table 2. NARX model parameters for modeling water consumption in Kuwait.

NARX parameter	Value	Correlation
Feedback delay	2	1st lag= +0.75; 2nd lag = +0.47
Input delay (temperature data)	0	0 lag = +0.89
Hidden neurons	30	

Table 3. Statistical evaluation measures for NARX model assessment.

Evaluation measure	Training	Validation
Pearson correlation	0.996	0.927
Coefficient of determination (R^2)	0.992	0.859
NS coefficient	0.992	0.837
MAE (MIG)	170.61	370.63

It is also crucial to assess the model efficiency by examining the MAE values. In many cases, good performances with respect to the NS, correlation, and R^2 coefficients do not necessarily indicate good model performance. However, in this study, the computed MAE values for the training and validation periods demonstrate that the margin of error is insignificant. This renders the model predictions useful in water-supply management for developing appropriate short-term water-supply strategies. It is worth noting that all the error metrics in the validation period are slightly lesser than those in the training period. This reasonable drop in metrics eliminates the possibility of model overfitting.

The modeled water consumption time series is shown in Figure 5. Apparently, the model can represent the long-term trend of the water consumption data with seasonal variations. The fitted linear trend of the raw water consumption data was added back to the NARX output, enhancing the model capability for capturing the long-term data trend. Moreover, the seasonal variation of the exogenous input (temperature data) efficiently reflects the seasonal variations of the water consumption data. Efficiency in forecasting the seasonal water consumption pattern is crucial. As the water-consumption demand generally peaks during summer, it is essential for any forecasting model to show satisfactory performance during this critical period.
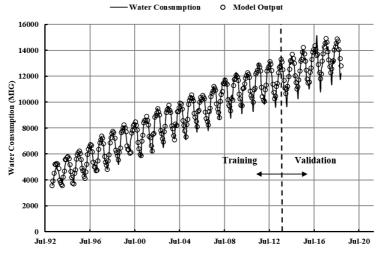

Figure 5. Modeled and observed water consumption data from 1993–2018.

Figure 6 displays the seasonal variations for the absolute forecasting error in the training and validation periods. The performance of the NARX model is the best during the summer season in both periods, demonstrating its robustness in this critical season. Water consumption forecasting in the summer season can aid water supply management by planning for short-term water supply alternatives. Higher absolute error values are noticed in the winter season forecasts; improving the model performance in this season requires additional exogenous inputs correlated to the water consumption during this period. However, additional exogenous inputs may expose the model to the overfitting problem. Thus, it would be meaningless to improve model performance in the winter season, exposing the model performance to potential efficiency drop in the summer season.
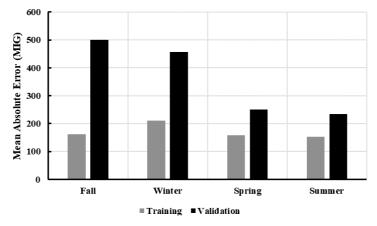

Figure 6. Seasonal variation of the absolute error in the recurrent model forecast.

The reliability of a forecasting model is strongly affected by predication bias. Hence, it is essential to examine the model bias by observing the forecasting error pattern. Figure 7 depicts the distribution of the computed errors with respect to the zero-error line. The error distribution is normally distributed with the maximum error frequencies scattered near the zero-error divide. The positive (model underestimating the targets) and negative (model overestimating the targets) errors are comparably distributed on the right and left of the zero-error divide. This systemic error distribution implies that the model does not consistently overestimate or underestimate the target water-demand value. Thereby, the bias in model prediction is eventually eliminated, ensuring more reliable model forecasts.
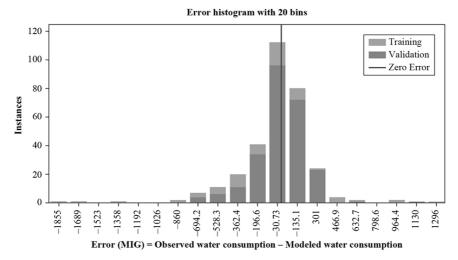

Figure 7. Distribution of the computed errors with respect to the zero-error line.

4.2 Comparison with previous studies

Numerous studies have investigated the various modeling techniques for modeling water consumption under various conditions. It is important to establish suitable modeling frameworks appropriate for the water supply system characteristics of each system. In this study, two characteristics distinguish the water supply system in the study area from the other counterpart systems: (i) The extremely harsh and arid environment, which stresses the water demand during the summer season and (ii) the rapid population growth induced by the developing economy in the state. These two factors shape a unique system that is comprehensively investigated in this study. Al-Zahrani and Abo-Mansour (Al-Zahrani and Abo-Monasar, 2015) utilized ANNs for modeling the urban water demand in Al-Khobar city, Saudi Arabia. This city shares the same factors that govern the study area. In the model developed by Al-Zahrani and Abo-Monasar (2015), $R^2 = 0.676$ for the best performing model in the validation period. This is significantly less than the same performance metric (0.859) in this study. Exclusion of the influence of population growth through the trending procedure for the NARX model input resulted in this major improvement in the proposed model. Modeling the detrended water consumption time series assisted the NARX model in recognizing seasonal water demand patterns without being affected by the long-term increasing trend induced by rapid population growth.

The results of NARX application in the study site demonstrated comparable efficiency with the other ANN-based model applications at various climatic conditions (Bougadis et al., 2005; Liu et al., 2003). Further, the developed NARX outperformed the constant rate models applied at study sites with similar climatic and demographic settings (Mohamed and Al-Mualla, 2010a, 2010b). The application of the NARX model resulted in better performance in terms of the prediction errors and R^2 values, implying that recurrent NARX-based models can be used as water management tools for water supply management in this region. The simplicity of the data requirement is a key advantage for NARX models. However, the applicability of the proposed approach is limited by the climatic, economic, and demographic features of the study site. Thus, the mobility of model application should be addressed with caution and is limited to similar conditions principally found in the Arabian Peninsula.

4.3 Role of NARX models in developing resilient water supply plans in hyper-arid regions

This study presented a NARX-based methodology for forecasting water consumption in a hyper-arid study site. It is well known that water demand is directly associated with population growth; however, climatic influencers also play a significant role. This study examines the suitability of monthly air temperatures in establishing reliable short-term water consumption forecasts in extremely harsh environmental conditions as those found in Kuwait. The novelty of the current study stems from the unique application of NARX models for water consumption forecasting in the harsh climatic conditions of the study area. The current study represents a first attempt in modeling water consumption system using the NARX method in such hyper arid environment. The implications of the obtained results for the anticipated water consumption demand can definitely aid water planners in developing efficient water supply plans in hyper-arid regions to cope with future freshwater needs on a short-term scale. This includes setting maintenance programs for desalination plants in order to utilize their full production capacity during critical seasons. Further, the results of short-term water demand forecasting can assist official authorities in making desalination plant production capacity upgrade decisions to secure the water supply needs.

5. CONCLUSION

This study investigated the suitability of NARX-based recurrent models for forecasting water consumption at a hyper-arid study area in Kuwait. The proposed recurrent model demonstrated strong modeling capability in forecasting short-term water demands. The detrended water consumption data exhibited strong autocorrelation up to the 2nd lag. In addition, cross correlation analysis showed that the average monthly temperatures have a significant Pearson correlation of +0.89 with the detrended water consumption time series. This strong association with the detrended water consumption data makes the average monthly temperature, which is an exogenous input to the recurrent model, a strong predictor of the water demand.

Optimal NARX performance was achieved by calibrating the number of hidden layers to 30. Errors associated with the model forecasts were normally distributed around the zero-error point. This emphasizes the model reliability and eliminates predictions bias. Seasonal analysis of the model performance showed that the model performed best during the critical summer season. Comparisons with previous modeling studies under similar climatic conditions indicated that recurrent NARX models are superior to conventional general regression neural networks. Future research efforts may be dedicated to validating the presented methodology at other study sites with emphasis on improving the model predictions by considering other exogenous inputs associated with water demand.

6. REFERENCES

Adamowski, J.F. 2008. Peak daily water demand forecast modeling using artificial neural networks. J. Water Resour. Plan. Manag. **134**: 119-128.

Al-Zahrani, M.A. & Abo-Monasar, A. 2015. Urban residential water demand prediction based on artificial neural networks and time series models. Water Resour. Manag. **29**: 3651-3662.

Aldrich, J. 1995. Correlations genuine and spurious in Pearson and Yule. Stat. Sci. **10**: 364-376.

Almedeij, J. & Alotaibi, S. 2017. Periodic and random components of water consumption in Kuwait. Proceedings of the Institution of Civil Engineers-Water Management. Thomas Telford Ltd, pp. 207-216.

Almutaz, I., Ajbar, A., Khalid, Y. & Ali, E. 2012. A probabilistic forecast of water demand for a tourist and desalination dependent city: Case of Mecca, Saudi Arabia. Desalination. **294**: 53-59.

Alotaibi, S. &Almedeij, J. 2007. Fresh water consumption in Kuwait: modeling the variable pattern. Kuwait J. Sci. Eng. **34**: 183.

Altunkaynak, A. & Nigussie, T.A. 2018. Monthly water demand prediction using wavelet transform, first-order differencing and linear detrending techniques based on multilayer perceptron models. Urban Water J. **15**: 177-181.

Banihabib, M.E. & Mousavi-Mirkalaei, P. 2019. Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region. Sustain. Cities Soc. **48**: 101585.

Bata, M.H., Carriveau, R., Ting & D.S.K. 2020. Short-Term Water demand forecasting using nonlinear autoregressive artificial neural networks. J. Water Resour. Plan. Manag. **146**: 4020008.

Blignaut, J. & Van Heerden, J. 2009. The impact of water scarcity on economic development initiatives. Water Sa. 35.

Bougadis, J., Adamowski, K. & Diduch, R. 2005. Short-term municipal water demand forecasting. Hydrol. Process. An Int. J. **19**: 137-148.

Cadenas, E., Rivera, W., Campos-Amezcua, R. & Heard, C. 2016. Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. Energies **9**: 109.

Dawadi, S. & Ahmad, S. 2013. Evaluating the impact of demand-side management on water resources under changing climatic conditions and increasing population. J. Environ. Manage. **114**: 261-275.

Goyal, M.K., Bharti, B., Quilty, J., Adamowski, J. & Pandey, A. 2014. Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert Syst. Appl. **41**: 5267-5276.

Harnoda, M.F. 2001. Desalination and water resource management in Kuwait. Desalination. **138**: 285-393.

Hydrology, A.T.C. on A. of A.N.N. in. 2000. Artificial neural networks in hydrology. I: Preliminary concepts. J. Hydrol. Eng. **5**: 115-123.

Khashei, M. & Bijari, M. 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl. Soft Comput. **11**: 2664-2675.

Krause, P., Boyle, D.P. & Bäse, F. 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. **5**: 89-97.

Liu, J., Savenije, H.H.G. & Xu, J. 2003. Forecast of water demand in Weinan City in China using WDF-ANN model. Phys. Chem. Earth, Parts A/B/C. **28**: 219-224.

Menezes Jr, J.M.P. & Barreto, G.A. 2008. Long-term time series prediction with the NARX network: An empirical evaluation.

Neurocomputing. **71**: 3335-3343.

Minstry of Electricity and Water (MEW). 2018. Yearly Statistics Report, Ministry of Electricity and Water, State of Kuwait.

Mohamed, M.M. & Al-Mualla, A.A. 2010a. Water demand forecasting in Umm Al-Quwain using the constant rate model. Desalination. **259**: 161-168.

Mohamed, M.M. & Al-Mualla, A.A. 2010b. Water demand forecasting in Umm Al-Quwain (UAE) using the IWR-MAIN specify forecasting model. Water Resour. Manag. **24**: 4093-4120.

Mousavi-Mirkalaei, P. & Banihabib, M.E. 2019. An ARIMA-NARX hybrid model for forecasting urban water consumption (case study: Tehran metropolis). Urban Water J. **16**: 365-376.

Nash, J.E. & Sutcliffe, J.V. 1970. River flow forecasting through conceptual models part I- A discussion of principles. J. Hydrol. **10**: 282-290.

Ngia, L.S.H. & Sjoberg, J. 2000. Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm. IEEE Trans. Signal Process. **48**: 1915-1927.

Odhiambo, G.O. 2017. Water scarcity in the Arabian Peninsula and socio-economic implications. Appl. Water Sci. **7**: 2479-2492.

Rahimi, Z., Shafri, H.Z.M. & Norman, M. 2018. A GNSS-based weather forecasting approach using Nonlinear Auto Regressive Approach with Exogenous Input (NARX). J. Atmos. Solar-Terrestrial Phys. **178**: 74-84.

Samarasinghe, S. 2016. Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition. Auerbach publications.

Siegelmann, H.T., Horne, B.G. & Giles, C.L. 1997. Computational capabilities of recurrent NARX neural networks. IEEE Trans. Syst. Man, Cybern. Part B. **27**: 208-215.

Tirth, V. (2020). Nexus of Water Footprint with Energy and GDP of Saudi Arabia and Solution for Sustainable Water Usage. In Advances in Water Resources Engineering and Management (pp. 43-57). Springer, Singapore.

Vörösmarty, C.J., Green, P., Salisbury, J. & Lammers, R.B. 2000. Global water resources: vulnerability from climate change and population growth. Science (80-). **289**: 284-288.

Wang, X., Zhang, J., Shahid, S., Guan, E., Wu, Y., Gao, J. & He, R. 2016. Adaptation to climate change impacts on water demand. Mitig. Adapt. Strateg. Glob. Chang. **21**: 81-99.

Zhang, G.P. 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing. **50**:159-175.

Zhou, F., Xu, Y., Chen, Y., Xu, C.Y., Gao, Y. & Du, J. 2013. Hydrological response to urbanization at different spatio-temporal scales simulated by coupling of CLUE-S and the SWAT model in the Yangtze River Delta region. J. Hydrol. **485**: 113-125.