Urban Water Demand Forecasting Using the Stochastic Nature of Short Term Historical Water Demand and supply Pattern

Thewodros G. Mamo [1], Ilan Juran[2], Isam Shahrour [3]

[1,2] Department of Civil & Urban Engineering Polytechnic Institute of New York University, 6 Metro Tech Center, Brooklyn, NY
[*3]Laboratoire de mécanique de Lille, University of Sciences and Technologies of Lille, 59650 Villeneuve d'Arcy, France
[1]tgm232@nyu.edu, P.E., Ph.D. Candidate [2]ijurann@poly.edu , Ph.D. [*3]isam.shahrourr@univ-lille1.fr ,Ph.D.

Abstract

Today's big city water utility companies are experiencing high level of water loss due to various problems in covering a large scale of water supply pipeline networks, therefore any significant improvement of water loss prevention from supply network to treatment plant would require an apprehends stochastic nature of historical water demand and supply pattern. For this reason urban water demand forecasting is one of a key important parameter used when water utility companies are trying to find more efficient and robust ways of supplying water for a large number of urban consumers. Water demand forecasting also plays a significant role in managing and planning of water supply operations and water conservation and optimization strategies.

However traditional forecasting approaches based on only on a set of deterministic design capacity factor or using demand forecasting algorithms without evaluating the relationship between system reliability in response to the stochastic nature of historical water consumption data and supply pattern often become misleading due to the inability to sufficiently forecasting forthcoming events and lack of relevant historical pattern and data. This paper presents an analysis and water demand forecasting demonstration using the stochastic nature of short term historical water demand and supply Pattern for Lille University Z6 pipeline networks research area in France.

Keywords: Urban Water Demand, Time series (AR1), Forecasting, stochastic simulation of demand, Historic demand

1. INTRODUCTION

1.1 The Role of Short Term Water Demand Forecasting

In the context of operation and maintenance management, short term forecasting water demand or water consumption not only playing a vital role for the water utility companies which are trying to find more efficient ways to supply water, for a large number of households, commercial and institutional businesses, but also contribute to control water losses due to leakage from the distribution pipe networks.

Short term urban water demand forecasting system support water distribution operations, as well as budgeting and financing management, and program tracking and evaluation. Billings R. B. and Jones c. V. (2008) forecasting urban water demand pp.7. The purpose of demand forecasting modeling also varies from the simplest historical extrapolation to sophisticated analytical models therefore the choice of an appropriate modeling ought to address the purpose of the forecasting needed by water utility companies, and the quality and quantity of data.

1.2 Overview of forecasting approach Using the Stochastic Nature of Short-Term Historical Water Demand and Supply Pattern

One of the practical applications of time series analysis and stochastic modeling is short term forecasting. In recent years different forecasting method have been developed, which could be used to control water distribution operation system nearly real time based on time series data that are collected sequentially overtime period.

In general the time series of water demand or consumption historical data extracted from SCADA or using different data gathering mechanisms can be divided into deterministic and stochastic components. The deterministic component which is one that can be determined for predictive purpose, the stochastic components, however is constituted by irregular oscillations and random effects which cannot strictly be accounted for physically and which require probabilistic concepts for description. Kottegoda, N.T. (1980). Stochastic water resource technology Pp. 22-23, this character can be identified by time series analysis and modeled by a suitable autoregressive Integrated Moving Average processes which can be described in terms of (p,d,q) notation, where p indicate the autoregressive order, q the moving average order, and d the degree of differencing necessary to achieve stationarity. Wadsworth, Jr. H.M. (2000) handbook of statistical methods for engineers and scientist pp. 18.5

2. Modeling Water Demand Deterministic Components

The modeling of the deterministic components of the time series water demand model consists of periodic components such as seasonal, weekly and hourly historical patterns and persistence components called short-term memory. The periodic components of the time series water demand or consumption data can be modeled by fitting a harmonic series to the annual series of average metered or observed demand of each day of 365 days based on a short-term, pattern based forecasting approach developed by Alvisi, S., Franchini, M. and Marinelli, A. (2007, September). A short-term, pattern-based model for water demand forecasting, Franklin, S. L & Maidment (1986), (1984) an evaluation of weekly and monthly time series forecasts of municipal water use. Wat. Res. Bull. 22 (4), 611–621., Zhou, S.l (2000) Forecasting Daily Urban Water Demand: A Case Study of Melbourne, Journal of Hydrology, and using equation of Kottegoda, N.T. (1980), *Stochastic water resource technology* pp.37.

2.1 The Daily Periodic Component

Therefore, the long-term average daily water demand $\overline{Q}_d^{D,S}$ pattern is modeled using a Fourier series:

$$\overline{Q}_d^{D,S} = \overline{Q} + \sum_{k=1}^{k} [a_k \cos\frac{2\pi k}{365} d + b_k \sin\frac{2\pi k}{365} d] \quad (1)$$

Where:- d = 1, 2, 3….365, $\overline{Q}$ = the mean value of seasonal cycle, k = is the number of harmonic considered, a_k and b_k = are Fourier Coefficients and calculated as:-

$$a_k = \frac{2}{365}\sum_{d=1}^{365}\overline{Q}_d^{D,metr}\cos\frac{2\pi d}{365}k \tag{2}$$

$$b_k = \frac{2}{365}\sum_{d=1}^{365}\overline{Q}_d^{D,metr}\sin\frac{2\pi d}{365}k \tag{3}$$

Where:

$\overline{Q}_d^{D,metr}$, are the observed average daily water demand.

2.3 The Weekly Periodic Component

The weekly pattern of the deterministic components of the water demand or consumption is based on a set of Weekly Correction Factor representing the weekly periodic component of $\overline{P}_{i,j}^{D,W}$ for each week and for each months (Jan….Dec.) and is calculated as:

$$\overline{P}_{i,j}^{D,W} = \overline{Q}_{i,j}^{D} - \overline{Q}_j^{w} \tag{4}$$

Where:

$\overline{Q}_{i,j}^{D}$ = is the mean value of the average daily water demand observed on day (Monday…Sunday) i of the week (i = 1, 2, and 3… 7), the season (winter, spring, summer, fall j of the year j = 1… 4)

$\overline{Q}_j^{w}$ = is the mean value of the average weekly demand in season jb which is winter, spring, summer, fall j of the year j = 1… 4)

2.4 The Hourly Periodic Component

The hourly module, of water demand is a time series forecasting model of hourly water consumption for 24h, the model consists of two modules-daily and hourly. Like the daily module the hourly prediction or forecasting is formulated as a set of equations representing the effects a periodic and persistence component. Therefore the hourly water demand of $Q_{t+y}^{h,for}$ forecasted for the time t hours for y hours ahead is determined using

$$Q_{t+1}^{h,for} = Q_d^{D,for} + \overline{P}_{n,i,j}^{h} + \varepsilon_{t+1} \tag{5}$$

Where: - $Q_d^{D,for}$ = the average daily water demand forecasted using equation (1)

$\overline{P}_{n,i,j}^{h}$ = the hourly deviation representing the daily pattern which is

$$\bar{P}^{h}_{n,i,j} = \overline{Q}^{h}_{n,i,j} - \overline{Q}^{D}_{i,j} \tag{6}$$

Where

$\overline{Q}^{h}_{n,\,i,\,j}$ = represents the mean value of the average hourly water demand measured in hour n (n =1, 2, 3… 24, the hour of the day i in season j)

$\overline{Q}^{D}_{i,\,j}$ = is the mean value of the average daily water demand measured on day i , in season j.

2.5 Modeling Water Demand Stochastic Components

The stochastic components of the water demand data are featured by a series of correlating characteristics which represent the short term memory components. These characteristics can be analyzed using time series and modeled by appropriate autoregressive-moving-average model (ARMA). Therefore the deviation between the average daily water demand and the mean value estimated solely on the basis of the periodic components $\overline{Q}^{D,\,S}_{d}$ and, $\bar{P}^{h}_{n,i,j}$ and is modeled using the autoregressive process AR (1). Alvisi, S., Franchini, M. and Marinelli, A. (2007, September). A short-term, pattern-based model for water demand forecasting, and Box, G. E. P., (1994). Time series analysis forecasting and control, 3rd edition

$$\sigma^{D}_{d} = \sigma^{D}_{d} = \Phi_{1} \cdot \sigma^{D}_{d-1} \tag{7}$$

Where:

σ^{D}_{d} = the deviation between the average daily water demand $\overline{Q}^{D}_{d}$ and the mean value estimated based on the basis of the periodic components.

$\overline{Q}^{D,\,S}_{d}$ and, $\overline{P}^{D,W}_{i,j}$ ➔ (i, j shows both Monday – Friday day 1…7, and winter – summer i.e.1,..4)

Φ = Auto-regressive parameter can be calibrated based on the basis of the observed deviation, in autoregressive (ARMA) processes of order p.

$$\sigma^{D,metr}_{d} = \overline{Q}^{D,metr}_{d} - [\overline{Q}^{D,S}_{d} + \overline{P}^{D,W}_{i,j}] \tag{8}$$

The daily and hourly residual or persistence component are modeled by using AR (1) on errors ε_{t+y-1} and ε_{t+y-24}

$$\varepsilon_{t+y} = \Psi_{1} \cdot \varepsilon_{t+y-1} + \Psi_{24} \cdot \varepsilon_{t+y-24} \tag{9}$$

The coefficient Ψ_{1} and Ψ_{24} depends on the hour of the day $t+y \equiv n = 1,2,\cdots,24$ $t+y$ Is starting from the beginning of the year and calibrated on the bases of the observed (measured) error ε^{metr}_{t} where:

$$\varepsilon_t^{metr} = Q_t^{h,metr} - (Q_d^{D,metr} + \overline{P}_{n,i,j}^{h}) \tag{10}$$

2.6 Daily and Hourly Demand–Forecasting

Based on the above analysis the average estimated mean daily water demand and the hourly water demand can be calculated using the following equations

$$Q_d^{D,for} = \overline{Q}_d^{D,S} + \overline{P}_{i,j}^{D,W} + \sigma_d^{D} \tag{11}$$

Where:- $\overline{Q}_d^{D,S}$ = is the long-term average daily water demand representing the seasonal (s) periodic component.

$\overline{P}_{i,j}^{D,W}$ = is the correction representing the weekly periodic component.

σ_d^{D} = representing the medium term persistence component.

The hourly module, of water demand is a time series forecasting model of hourly water consumption for 24h, the model consists of two modules-daily and hourly. Like the daily module the hourly prediction or forecasting is formulated as a set of equations representing the effects a periodic and persistence component. Therefore the hourly water demand of $Q_{t+y}^{h,for}$ forecasted for the time t hours for y hours ahead is determined using

$$Q_{t+1}^{h,for} = Q_d^{D,for} + \overline{P}_{n,i,j}^{h} + \varepsilon_{t+1} \tag{12}$$

Where: -

$Q_d^{D,for}$ = the average daily water demand forecasted using equation (1)

$\overline{P}_{n,i,j}^{h}$ = the hourly deviation representing the daily pattern which is

$$\overline{P}_{n,i,j}^{h} = \overline{Q}_{n,i,j}^{h} - \overline{Q}_{i,j}^{D} \tag{13}$$

Where: - $\overline{Q}_{n,i,j}^{h}$ = represents the mean value of the average hourly water demand measured in hour

(n=1, 2, 3… 24, the hour of the day i in season j)

$\overline{Q}_{i,j}^{D}$ = is the mean value of the average daily water demand measured on day i , in season j.

The rest of this paper discusses the result and analysis of the Zone 6 research project area.

3. Zone-Six (Z6) Research Area -Lille University

3.1 Introduction

The University Lille 1 was established 1854 in Lille, although its academic roots extend back to 1562. It later moved to Villeneuve d'Ascq in 1967, with 25,000 full-time students plus 15,000 students in continuing education (2011). 1,310 permanent faculty members plus 1,200 staff and around 140 CNRS researchers work there in the different University Lille 1 institutes and 43 research labs. University Lille 1 is a member of the European Doctoral College Lille-Nord-Pas de Calais, which produces 400 doctorate dissertations every year. The university is ranked in the world top 200 universities in mathematics by the Shanghai ranking. (News on Lille 1's webpage, http://www.univ-lille1.fr/Accueil/Actualites?id=26313)

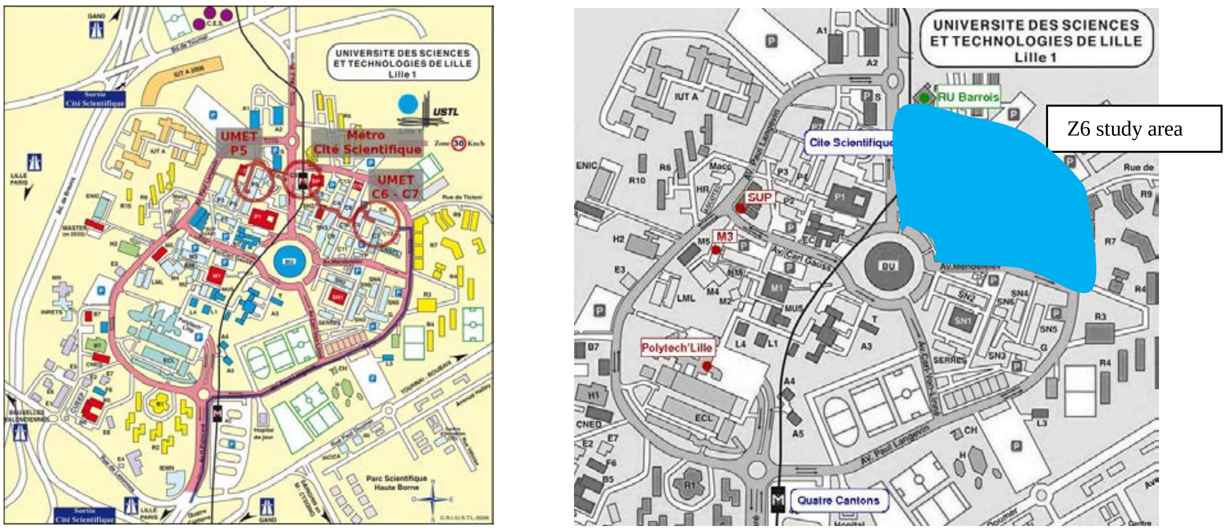


Fig -1- General and project Research Area map for Zone-6 map

3.2 Lille University's Water Network System

The Lille University, drinking water network is divided in to different supply zones Z1, Z2, etc. . Since the university is planning to implement water loss and leakage prevention by using District Metered Area Zone -6 water demand data has been used for this research purposes. The present water consumption at Zone -6 is about 570 m^3 per day, with operating pressure close to 4, 5 bars. The pipe network has about 50 years of age and it is made of cast Iron pipes, 150 mm in diameter. The seasonal water demand pattern of the time series (spring, summer, fall and winter) of Lille University is shown in Figure -3-, while Figure-2- shows the yearly water demand data pattern respectively.

Lille University Seasonal Time Series Water Consumption Pattern

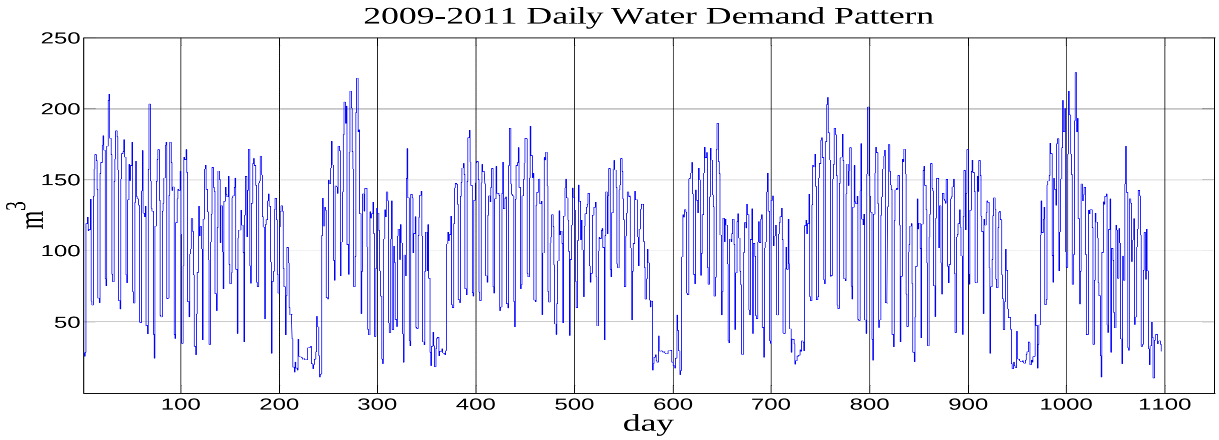
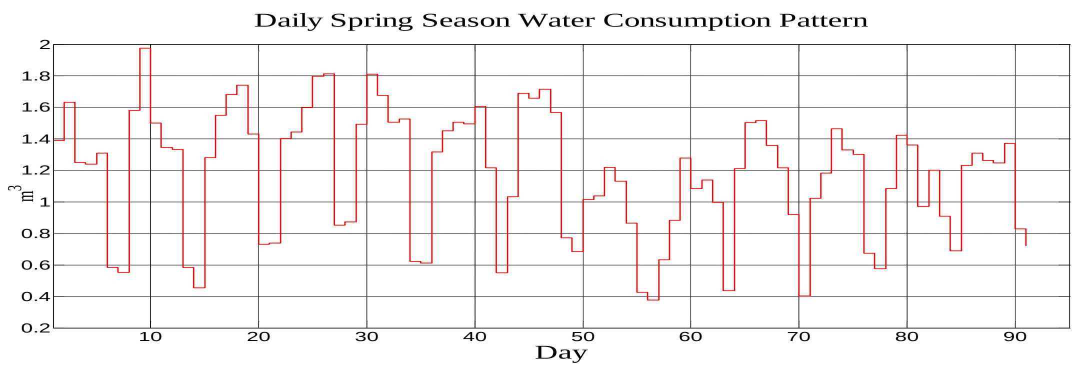
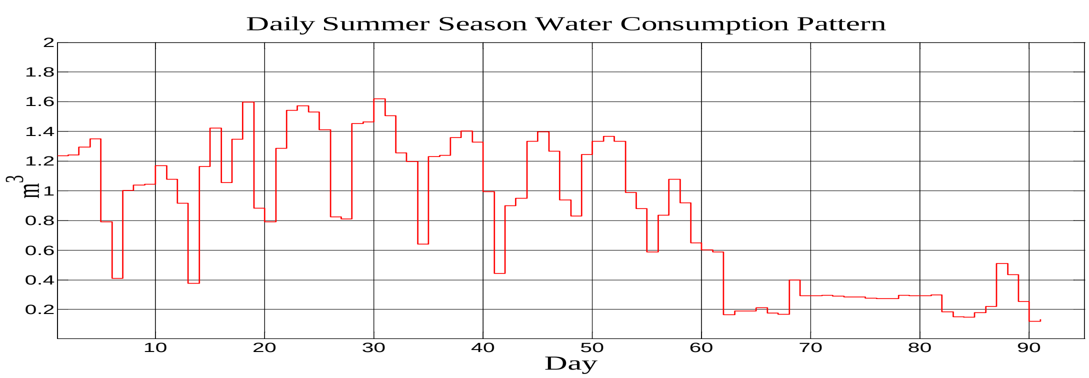
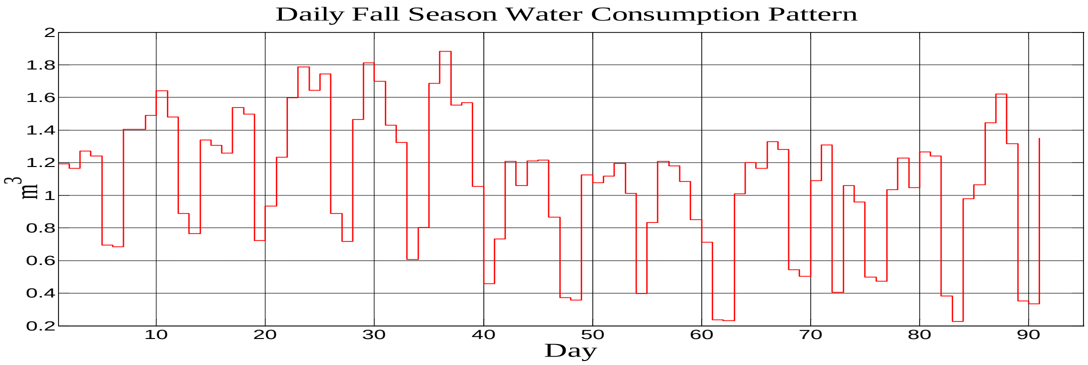


Figure -2- three *years water demand pattern of the time series*







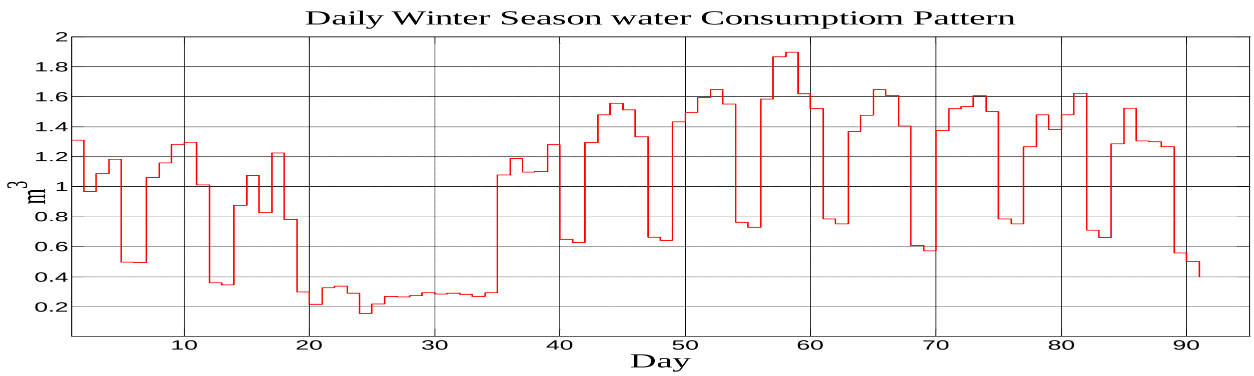


Figure -3- *seasonal water demand pattern of the time series (spring, summer, fall and winter)*

3.3 Model Data

A stochastic water demand model for synthetic generation of water demands has been formulated and applied in modeling and forecasting water demand for Zone 6 water supply Zone project. The model is based on statistical analysis of historical water demand data using a short-term, pattern-based model for water demand forecasting approach (Alvisi et al 2007). The model comprises a deterministic component and a stochastic component. The deterministic components are modeled by a series of seasonal, weekly and daily patterns. The stochastic components are based on time-series forecasting models that take into the short-term memory effect and the random residual components.

The forecasting and Modeling of water demand require reliable data. Reliable data play a key role in the analysis, monitor, and forecast of water system behaviors as bad quality data may result in an erroneous decision scheme. Siao S., Jean l. Bertrand k. (2011), Literature review for data validation methods.

The data employed for Zone -6 study area were consist of hourly water demand in meter cube per day (m^3/d). The water demand data were available for a period of 2009 to 2011; however, during the evaluation of the water demand data for the completeness, correctness process to avoid questionable data and erroneous we found that some of the data were scattered with a missing records, as a result, those missing data has been adjusted and some of them has been completely omitted and the remaining corrected water demand data were considered for model development and testing. All the data were divided into two sets: modeling set consisting the 2009-2010 years of data, and a testing set consisting of the remaining 2011 year of data. The overall task for this research involves the illustration and analysis of Water Demand Forecasting Using the Stochastic Nature of Short Term Historical Water Demand and supply Pattern methodologies for Lille University Z6 using the 2009-2011 water demand / consumption data. The includes an evaluation of considered forecasting methodologies including graphical plots of forecasted vs. observed, time series plots, and relevant statistical measures (e.g., correlation coefficient, residual analyses,..etc.).

3.4 Forecasting Results Analysis and Discussions

This study demonstrates how Water Demand Forecasting Using the Stochastic Nature of Short Term Historical Water Demand and supply Pattern models are useful to study and forecast short term water demand for water utility companies. This paper demonstrates also the data analysis results obtained and shown graphically in Figures 4 through 5 which confirmed that the model is considerably doing well and gave reliable forecasts for daily and hourly basis of the water demand. The modeling results also indicate that the synthetic patterns, fluctuations and statistics of the generated demands are consistent with the actual ones.

To measure the forecasting accuracy Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Tracking signals (TR) were used and the result on Table-1- shows the summary of analysis. Based on the calculations results of accuracy and sensitivity measures the MAE is 3.28% for daily and 1.63% for hourly respectively the ideal MAE is zero which would mean there is no forecasting error, the larger the MAE, the less the accurate the resulting model. The Tracking signal (TR) also calculated to measure how often our estimations have been above or below the actual vale and to decide where to re-evaluate. Based on the analysis made the TR value has been found that 0.54 for daily and 0.85 for hourly water demand respectively. Most of the Positive Tracking signal values show the actual values are above forecasted values, in other hand the Negative Tracking signal values shows the actual values are below the forecasted values, and most of the time it ranges from 4 and -4. ($4 \leq \mathrm{TR} \leq -4$). The squared correlation coefficient or the coefficient of determination has been calculated to evaluate the strength of a relationship between forecasted and observed water demand data, based on the analysis it has been found that 0.98 for daily and 0.97 for hourly respectively and this value shows there is strong relation between the forecasted and observed water demand value.

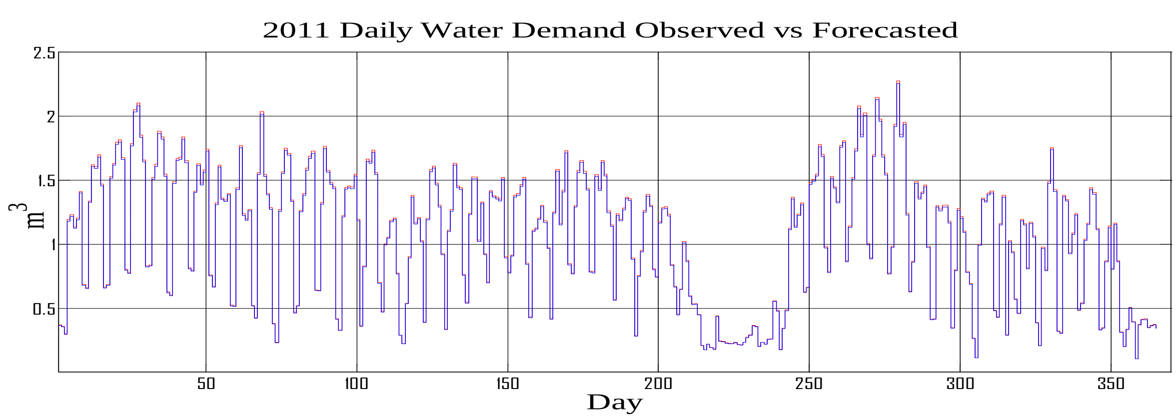


Figure-4- observed (red) daily and forecasted (blue) water demand

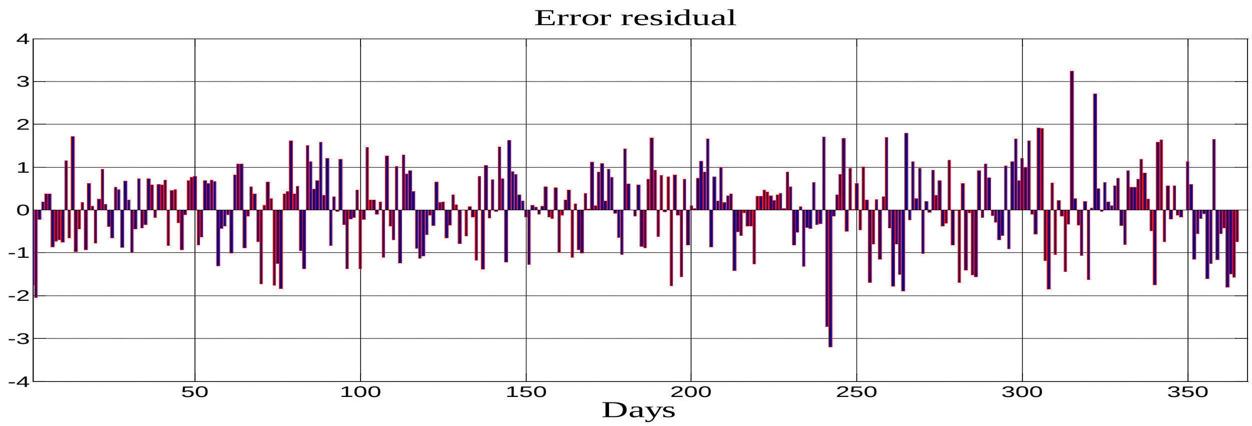


Figure-5- 2011 daily forecasted and observed Normalized Error Residual

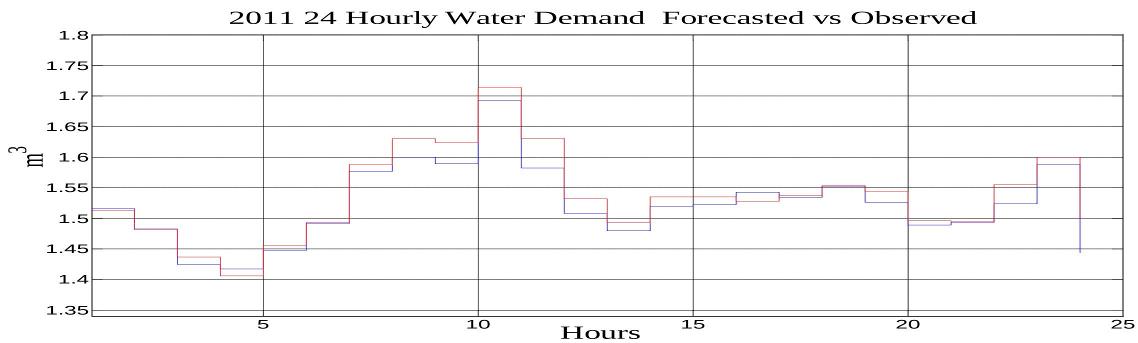


Figure-6- observed (red) hourly and forecasted (blue) water demand

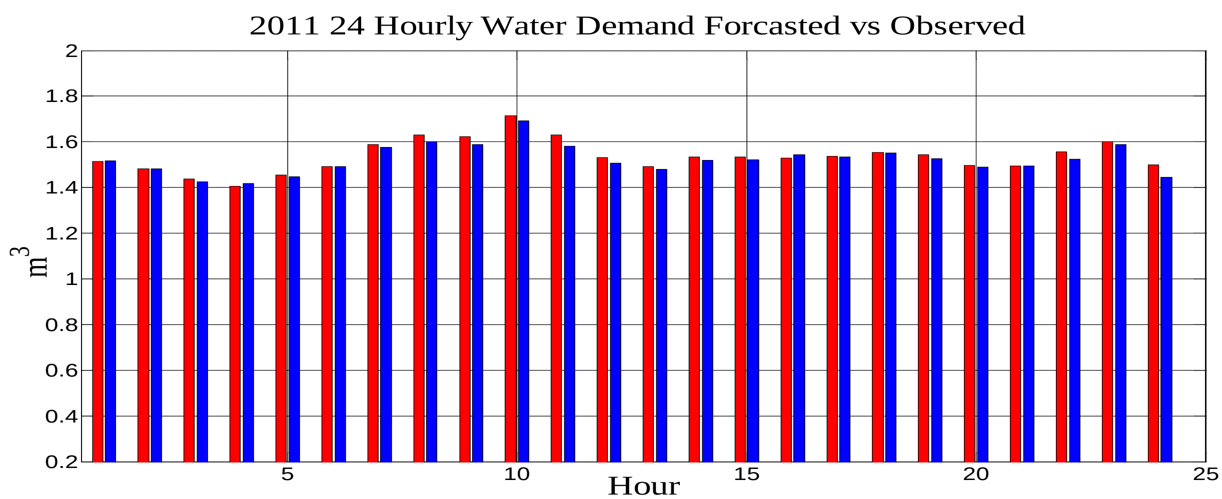


Figure-7- observed (red) hourly and forecasted (blue) water demand comparison

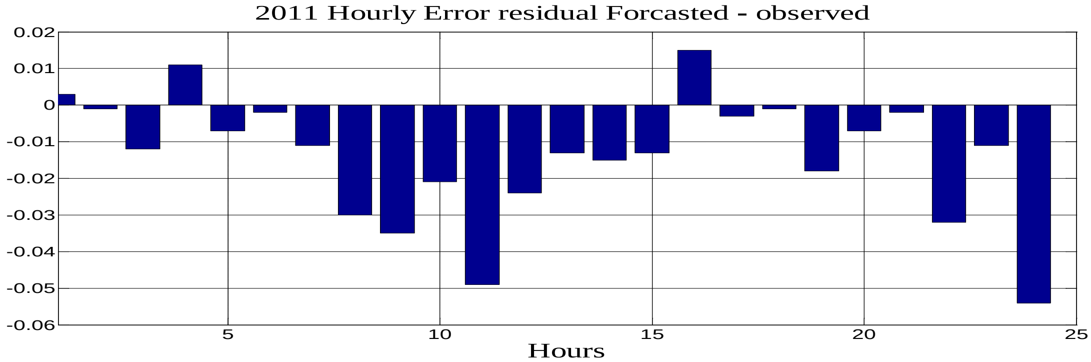


Figure-8- 2011 Hourly forecasted and observed Error Residual

3.5 Forecasting Accuracy, Validation and Sensitivity Analysis

The practical use of any model not only depends on the forecasted output accuracy and reliability, but also varies depends on the accuracy requirements from one water utility companies to another and from application to application. For this Research the validation comparisons were performed by establishing whether or not the model forecasted in 2011 water consumption output satisfactorily match with the actual observed water consumption data in the year 2011. Even if the usual practice of validation typically requires a quantitative metric be satisfied in order for validation to be confirmed, but we can also show a graphic comparison between the model-forecasted measures of short term water demand and the actual measures data. Below figure-9- shows

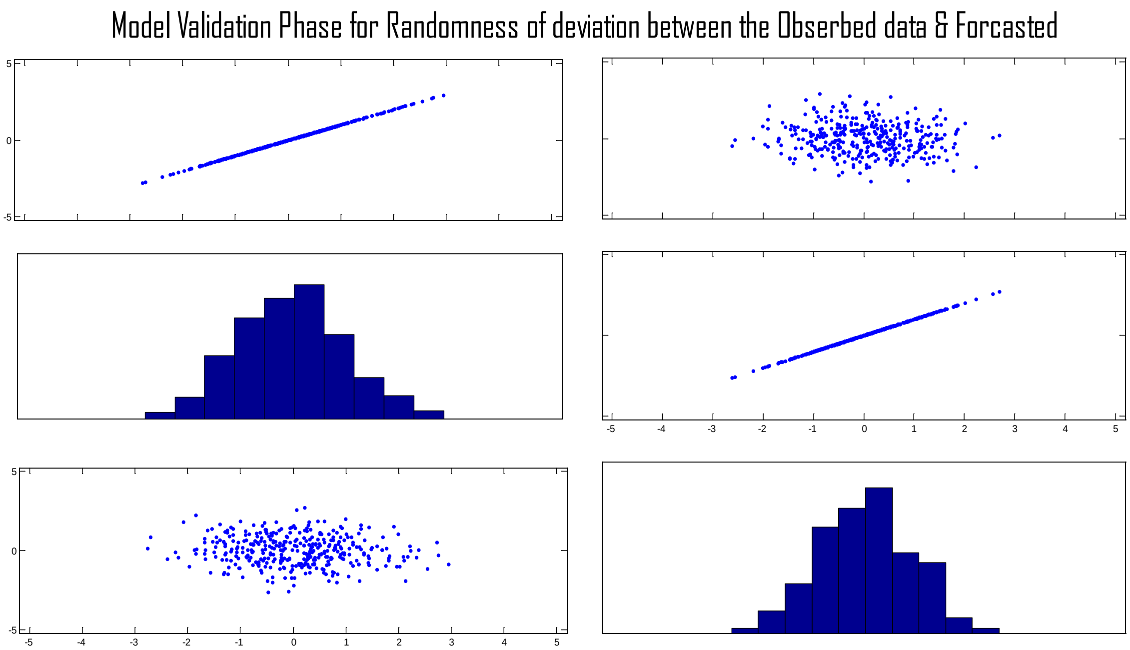


Figure-9- comparison between the model-forecasted measures of short term water demand

4. Sensitivity Analysis

Sensitivity analysis (SA), broadly defined, is the investigation of these potential changes and errors and their impacts on conclusions to be drawn from the model. Pannell D.J (2011). Sensitivity analysis: strategies, methods, concepts, examples. In this research the sensitivity analysis (SA) performed to test the model forecasting validity or accuracy and the forecasting model pattern and persistence components for the daily and hourly 2011 water demand forecasts, taking into account, the periodic and the periodic and persistence component together. The accuracy and sensitivity analysis were calculated and plotted as shown for both daily and hourly 2011 forecasted values. The residual error between forecasted and actual water demand is

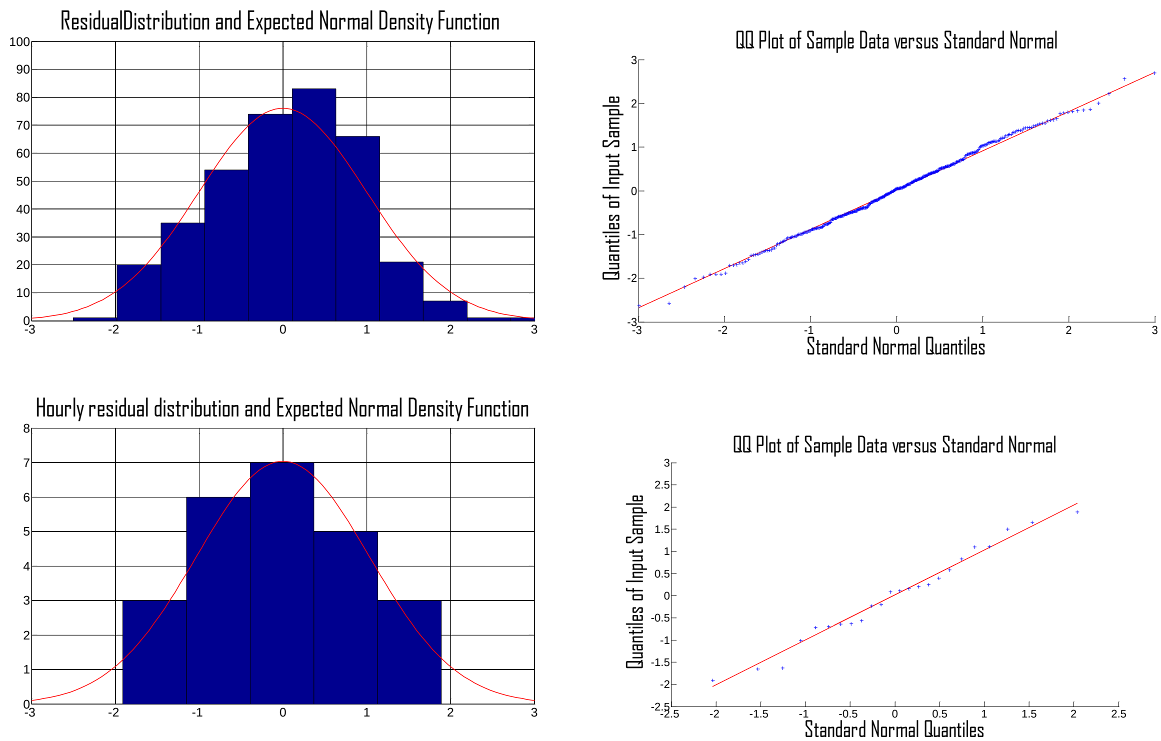


Figure-10- The residual error between forecasted and actual water demand

modeled as a normal distribution with mean of zero and standard deviation of the residual error results, for these reasons the forecasting error were measured using, Average Absolute Relative Error (AARE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Tracking signal (TR). Tracking signal is a measure of how often our estimations have been above or below the actual vale. It is used to decide where to re-evaluate using a model. Most of the Positive Tracking signal values show the actual values are above forecasted values, in other hand the Negative Tracking signal values shows the actual values are below the forecasted values, and most of the time it ranges from 4 and -4. ($4 \leq TR \leq -4$).

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|F_t - A_t| \qquad (14)$$

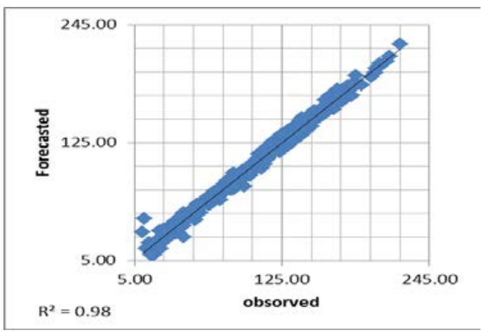
The ideal MAE is zero which would mean there is no forecasting error. The larger the MAE, the less the accurate the resulting model, if errors are normally distributed, then $e_\varepsilon = 1.25$ MAE. A data set with a smaller mean absolute deviation has data values that are closer to the mean than a data set with a greater mean absolute deviation.

$$\mathrm{NPE} = \frac{\bar{e}}{\frac{1}{n}\sum_{t=1}^{n} A_t} \tag{15}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\left[\sum_{t=1}^{n} e_t^2\right]} \tag{16}$$

$$\mathrm{TR} = \frac{\sum_{t=1}^{n}(A_t - F_t)}{\frac{1}{n}\sum_{t=1}^{n}|F_t - A_t|} \tag{17}$$

Where $F_{i.t}$ forecast for series (i) at time(t), $A_{i.t}$ actual for series (i) at time(t), $\bar{e}_i$ average error for series(i), MAE mean absolute error or deviation, NPE net percent error, RMSE root mean square error, TR tracking signal. The results of the calculation are summarized below table



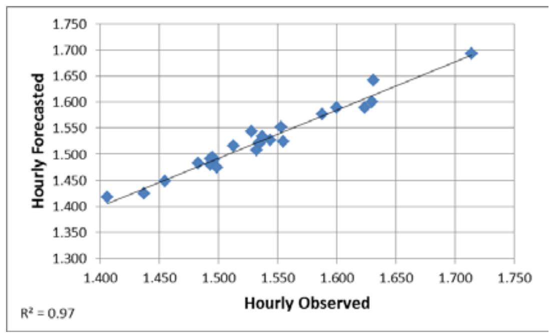


Figure -11- Figure -11- comparison plot of observed versus predicted values

Table-1-

Results of forecasted accuracy measurement	2011 Daily	2011 Hourly
MAE %	3.28	1.63
MAD	3.28	1.63
NEP	1.86	1.1
RMSE	1.86	1.27
TR	0.54	0.85
CORR	0.98	0.97
R^2	0.987	0.97

The Table-1- above shows the result of accuracy measurement and sensitivity analysis R^2

5. CONCLUSIONS

In this paper, we presented Water Demand Forecasting Using the Stochastic Nature of Short Term Historical Water Demand and supply Pattern, and demonstrate the data analysis results obtained and shown graphically, which confirmed that the model is considerably doing well and gave reliable forecasts for daily and hourly basis water demand. The modeling results also indicate that the seasonal patterns and statistics of the generated demands trends are consistent with the actual. This approach could be easily implemented for short term water demand forecasting system to support day to day operations, as well as budgeting and financing management, and program tracking and evaluation.

REFERENCES

[1] Alvisi, S., Franchini, M. and Marineli, A. (2007). A short-term, pattern-based model for water demand forecasting. Journal of Hydro informatics

[2] Billings B.R, and Jones V.C. (2008) Forecasting Urban Water Demand. American Water Works Associations (AWWA) pp. 7

[3] Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. 1994 Time Series Analysis Forecasting and Control, 3rd edn. Prentice Hall Englewood Cliffs, NJ.

[4] Franklin, S. L. & Maidment, D. R. 1986 an evaluation of weekly and monthly time series forecasts of municipal water use. Wat. Res. Bull. 22 (4), 611–621.

[5] Jamieson, D. G., Shamir, U., Martinez, F. & Franchini, M. 2007 Conceptual design of a generic, real-time, near optimal control system for water distribution networks. J. Hydro informatics 9 (1), 3–14.

[6] Kottegoda, N. T. (1980). Stochastic Water Resources Technology pp. 22-23,34-46, 98-104, 111-172. Macmillan, London.

[7] Maidment, D. R. & Parzen, E. 1984a Time patterns of water uses in six Texas cities. J. Wat. Res. Plann. Mngmnt., ASCE 110 (1), 90–106.

[8] David J. Pannell (2011). Sensitivity analysis: strategies, methods, concepts, examples.

[9] Saleba, G.S., Water Demand Forecasting, Proceedings of AWWA Seminar on Demand Forecasting and Financial Risk Assessment, Denver, Colorado, 1985.

[10] Siao S., Jean l. Bertrand k. (2011), Literature review for data validation methods

[11] Smith, J. A. 1988. A Model of Daily Municipal Water Use for Short-Term Forecasting Wat. Res. Res. 24 (2), 201–206.

[12] Shvarster, L., Shamir, U., and Feldman, M., Forecasting Hourly Water Demands by Pattern Recognition Apprach, ASCE Journal of Water Resources Planning and Management, 119(6), 611- 627, 1993.

[13] Wadsworth, J. R., H.M (2000) Hand Book of Statistical Method for Engineers and Scientists. pp.,18.5 –20.5. McGraw-Hill

[14] Zhou, S.L., McMahon, T.A., Walton, A., and Lewis, J., Forecasting Daily Urban Water Demand: A Case Study of Melbourne, Journal of Hydrology, 236, 153-164, 2000.