On Forecasting Water Consumption in Davao City Using Autoregressive Integrated Moving Average (ARIMA) Models and the Multilayer Perceptron Neural Network (MLPNN) Processes

Jamil L. Lipae, Dr. Eveyth P.Deligero

***Abstract*—**Water plays a significant role in human existence. Hence, understanding about its importance and usage is also necessary. This study sought to examine the characteristics, trends, and forecasts of the monthly water consumption in cubic meters in Davao City as a whole and as four different sectors such as residential, commercial, governmental and bulk or seller. The time series statistical forecasting processes such as Autoregressive Integrated Moving Average (ARIMA) models and Multilayer Perceptron Neural Network (MLPNN) models were used to analyze the data from January 1994 to October 2010 and predict the future values from November 2010 to December 2013 of the monthly water consumption in Davao City. As a result, in the ARIMA processes, it was shown that the most appropriate models are ARIMA(0,1,5)X(0,1,1)$_{12}$, ARIMA(0,1,5)X(0,1,1)$_{12}$, ARIMA(0,1,2)X(0,1,1)$_{12}$, ARIMA(0,1,1)X(0,1,1)$_{12}$ and ARIMA(0,1,1)X(0,1,1)$_{12}$ for water consumption as a whole, residential, commercial, governmental, and bulk or seller, respectively. In the MLPNN processes, on the other hand, the most appropriate models are ANN(36,6,1) with sigmoid function for water consumption as a whole, ANN(48,7,1) with hyperbolic tangent for residential, ANN(36,6,1) with bipolar sigmoid for commercial, ANN(48,7,1) with sigmoid function for governmental, and ANN(24,5,1) with sigmoid function for bulk or seller. Both processes agreed that water consumption in Davao City will continue to increase in the future.

***Keywords*—**Autoregressive Integrated Moving Average (ARIMA) Models, Forecasting, Monthly Water Consumption, Multilayer Perceptron Neural Network (MLPNN) Processes, Time Series Analysis.

I. Introduction

THROUGHOUT the years, mathematics has been the center of human development. It became essential in human civilization in the ancient era and still essential in human existence today. The role of mathematics is very much obvious and useful from the very simple human activities to the most complicated ones. Further, the development of mathematics from simple arithmetic to the higher concepts and applications is so much needed to explain and solve real world problems and happenings. One of the problems that a human being may encounter is the lack of resources in the future in terms of food supplies, shelters, amount of land for settlement, clothing, and the like. Because of this unavoidable and inevitable circumstance, every government in the world needs to think of possible and concrete solutions to avoid worst problems to occur in the future, especially in the distribution of resources.

Resources include water, whether it is used for intake or any other usage. Each individual needs water in their daily living, and it is much obvious throughout human history. It is even considered as one basic need of every human and even of animals and other creatures. The United Nation recommends that people need a minimum of 50 liters of water a day for drinking, washing, cooking and sanitation [12]. However, nowadays, water had been a problem in many countries due to the issues of cleanliness and sanitations, improper distribution, and lack of supplies that causes water shortage. It is said that around the world, water access issues are reaching crisis point. It was reported further that in 1990, over a billion people did not have the amount of water needed for their daily consumption due to water shortage [11].

In the Philippines, water crisis is also a big issue. It is said that Metro Manila is now experiencing a water shortage crisis with millions enduring water supply rationing due to series of unexpected calamities [9]. One reported that Metro Cebu in the Visayas and Davao City in Mindanao could experience a water crisis by 2015, as well as a sanitation crisis. Current water supplies in Cebu and Davao fall well short of their populations' needs with the situation set to worsen in the coming years as the numbers rise [13].

Davao City, in particular, is tackling the issue whether the water source will be converted to Tamugan River. However, there's a question if Tamugan river enough to sustain the needs of water by the people in Davao City.

The problems on water shortage and water distribution can only be addressed if each government in each country would know the availability and demand of water in the future so as the increase of population. Hence, statistical forecasting is so much appropriate knowledge in mathematics that could address these particular problems. Statistical forecasting is a method that uses the past to predict the future by identifying trends and patterns within the data to develop a forecast. This forecast is referred to as a statistical forecast because it uses mathematical formulas to identify the patterns and trends while testing the results for mathematical reasonableness and

Jamil L. Lipae is with the Department of Mathematics, Ateneo de Davao University, Davao City 8000, Philippines (phone: +63 932 3466191; e-mail: jamlipae@ymail.com / jllipae@addu.edu.ph).

Dr. Eveyth P. Deligero is with the College of Arts and Sciences, University of Southeastern Philippines, Davao City 8000, Philippines (e-mail: evedeligero@yahoo.com).

confidence [20]. With this continuing problem on water crisis, the researchers decided to conduct a study with regards to forecasting monthly water consumption in Davao City, Philippines from Nov 2010 to Dec 2013.

II. MATERIALS AND METHODS

The data that was used in this study is the recorded total monthly billed water consumption in Davao City, Philippines from Jan 1994 to Oct 2010 registered at Davao City Water District (DCWD). The data for water consumption were stratified into five categories namely water consumption as a whole, residential, commercial, governmental, and bulk or seller.

Statistical Predictive Software like R was also used in this paper. It was used mainly in analyzing the characteristics of the data gathered and in coming up with the time series plots namely the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots, and in identifying and validating the Autoregressive Integrated Moving Average (ARIMA) models. Other statistical software, the Zaitun Time Series, was utilized in MLPNN processes.

For ARIMA models, the procedure was divided into three. These are (1) data analysis, (2) model identification and estimation, and (3) model validation and forecasting.

To analyze the data gathered, the time series plot for the data was first determined. After that, the tests for stationarity were done to check whether the series is stationary or not. The tests were the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test. The data is only stationary if the null hypothesis in KPSS test is accepted, and both null hypotheses in ADF and PP tests are rejected. The null hypothesis of KPSS test indicates that the time series is stationary, and the ADF and PP tests indicate that there is an existence of the unit root on the series. If the series has a unit root, the series is not stationary [16]. If the data is not stationary, then first regular differencing must be applied to meet the condition for stationarity, and if by doing this the data is still not in a stationary condition, the second regular differencing will be applied [2].

Moreover, in identifying the possible and appropriate Autoregressive Integrated Moving Average (ARIMA) model, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) were plotted using the R statistical software. Then, the Box-Jenkins assumptions were used as the basis in identifying the appropriate models. Secondly, using the same statistical software, the coefficients of the parameters and residuals were determined to come up with a final ARIMA model that would fit the original data. The final ARIMA models were further examined to check if all the parameters are significant using the assumptions of the criterions and p-value. The models with the smallest Akaike Information Criterion (AIC) and or Root Mean Squared Error (RMSE) were chosen as the final models to be used in forecasting water consumption [19].

Lastly, for model validation and forecasting, the Autocorrelation Function (ACF) of the residual series was plotted to distinguish if the residuals are uncorrelated or independent. If some of the lags are significantly different from zero, then the model is appropriate. Thus, the model could give good estimates and true values in the future. However, if this condition is not met, then the process will repeat from data analysis until the appropriate model is identified [2].

Moreover, the residual series was analyzed using the Box-Pierce Q-statistics and Box-Ljung test to make clear that the values are normally distributed. If the residual series meets the condition of normal distribution, then the model could be used in predicting future values [5].

For MLPNN processes, the procedure was divided into four. These are (1) variable selection, (2) formation of training, testing, and validation sets, (3) architecture, and (4) model verification and forecasting.

The variable selection was software based. This means that the software was the one that chose the variables for training, testing, and validation sets.

The training set was the largest set and was used by the networks to learn patterns presented in the data. The testing set was used to evaluate the generalization ability of a supposedly trained work. A final check on the performance of the trained network was made using validation set [15].

The Neural Network architecture defined its structure including number of hidden layers, number of hidden nodes, and number of output nodes. The architecture is classified into four steps. These are (1) identifying the number of hidden layers, (2) identifying the number of hidden nodes, (3) identifying the number of output nodes, and (4) determining the right activation function. The hidden layer/s provide/s the network with its ability to generalize. In most studies, a Neural Network with one or occasionally two hidden layers is used since they have the tendency to perform very well [1],[17]. There is no specific formula to be used in finding the optimum number of neurons in the hidden layer/s. Yet, for a three-layered network, Masters proposed a Geometric Pyramid Rule which is $h = \sqrt{nm}$ where h be the number of hidden neurons, n the number of input neurons and m the number of output neurons [7],[8]. As for output nodes, Neural Networks with multiple outputs, especially if these outputs are widely spaced, will produce inferior results compared to the network with a single output. Lastly, activation function is very important in building a model. Activation functions are mathematical formulae that determine the output of a processing node. The purpose of the activation function is to prevent output from reaching very large value which can paralyze Neural Networks and therefore inhibit training [6]. In this paper, four kinds of activation functions were examined. These are linear, sigmoid, bipolar sigmoid, and hyperbolic tangent activation function.

The Neural Network that was used in this study is the Multilayer Perceptron Neural Network (MLPNN). The Network constructed a model based on examples of data with known outputs. It has to build the model up solely from the examples presented, which are together assumed to implicitly contain the information necessary to establish the relation. To verify the indicated model, the residual series was plotted and tested to check whether the series are uncorrelated or not. When the residuals were shown to be uncorrelated, the model selected was used to forecast the water consumption in Davao City.

III. RESULTS AND DISCUSSIONS

a. Characteristics of Water Consumption

The data showed that the average yearly consumption of water consumption as a whole from 1994 to 2009 in Davao City is 44,655,821 cu.m with the median of 43,662,618 cu.m and a standard deviation of 8,990,787 cu.m, 37,325,552 cu.m for water consumption as residential with the median of 36,761,484 cu.m and a standard deviation of 7,992,421 cu.m, 6,120,955 cu.m for water consumption as commercial with the median of 6,334,116 cu.m and a standard deviation of 996,049 cu.m, 1,157,548 cu.m for water consumption as governmental with the median of 1,051,125 cu.m and a standard deviation of 269,725 cu.m, and 52,936 cu.m for water consumption as bulk or seller with the median of 43,401 cu.m and a standard deviation of 26,871 cu.m. Moreover, the rate of increase or average percentage of the water consumption as a whole, residential, commercial, governmental, and bulk are 5.11%, 5.25%, 5.45%, 2.48%, and 6.84%, respectively.

b. Forecasting Using ARIMA Models

Figure 1 shows the time series plot of water consumption as a whole. The figure shows an increasing trend which suggests that the series is not in stationary condition and there's a need to apply the differencing technique to make the series stationary. Moreover, to support this claim, the test for stationarity was applied. On the first test, KPSS = 0.01 < 0.05, ADF = 0.5735 > 0.05, and PP = 0.01 < 0.05. Both KPSS and ADF tests showed that the series is not stationary. Thus, there was a need to apply the differencing technique. After applying the technique, the series was in stationary condition with KPSS = 0.1 > 0.05, ADF = 0.01 < 0.05, and PP = 0.01 < 0.05. Hence, the series was ready for model identification. Both the ACF and PACF plots as shown in Figure 2 and Figure 3, respectively, display a sharp cut-off and the lag 1 of the ACF is negative. This implies an order of moving average (MA) term. The feasible number of MA term is the lag that is zero or approximately equivalent to zero [18]. Table I presents the possible ARIMA models with corresponding criterions, the Root Mean Squared Error (RMSE) and the Akaike Information Criterion (AIC). In this case, the model with the least criterion for water consumption as a whole is ARIMA (0,1,5) X $(0,1,1)_{12}$ with the AIC value of 5011.86 and RMSE of 1.274e+5. Hence, this is the model that was used in forecasting water consumption in Davao City as a whole from Nov 2010 to Dec 2013. When the residuals were examined, they manifested no autocorrelation. This claim was supported by the ACF and the PACF plots of the residuals. This claim was given more strength by the Box-Pierce and the Box-Ljung Tests such that p-values are greater than 0.05 which implied that null hypotheses of the tests had to be accepted, that is the lags are independent from each other. This further concludes that the indicated model was indeed appropriate to forecast future values of water consumption. The chosen model for water consumption as a whole, ARIMA(0,1,5) X $(0,1,1)_{12}$ has coefficients per term which are MA1 = -1.1329, MA2 = 0.2643, MA3 = 0.3212, MA4 = -0.2947, MA5 = 0.1234, and SMA1 = -0.8651. Thus, the explicit equation is

$$Y(t) = \frac{(1+1.1329B-0.2643B^2-0.3212B^3+0.2947B^4-0.1234B^5)(1+0.8651B^{12})}{(1-B)(1-B^{12})}a_t \qquad (1)$$

where B as the backshift operator and a_t as the white noise.

Further, Figure 4 presents the actual and predicted plot of water consumption as a whole. It is clear that the predicted values follow the high-and-lows of the actual values. Figure 5 presents the time series plot of water consumption as a whole with the 38 month forecasts. It is shown that the forecasts also follow the trend of the actual values.

With the same analysis applied to the water consumption in four sectors, namely residential, commercial, governmental, and bulk, the models and the predictive equations are as follows:

Residential: ARIMA (0,1,5) X $(0,1,1)_{12}$

$$Y(t) = \frac{(1+1.1243B-0.2567B^2-0.2949B^3+0.3058B^4-0.1433B^5)(1+0.8412B^{12})}{(1-B)(1-B^{12})}a_t \qquad (2)$$

Commercial: ARIMA (0,1,2) X $(0,1,1)_{12}$

$$Y(t) = \frac{(1+0.6354B-0.1798B^2)(1+0.9401B^{12})}{(1-B)(1-B^{12})}a_t \qquad (3)$$

Governmental: ARIMA (0,1,1) X $(0,1,1)_{12}$

$$Y(t) = \frac{(1+0.6808B)(1+0.8941B^{12})}{(1-B)(1-B^{12})}a_t \qquad (4)$$

Bulk or Seller: ARIMA (0,1,1) X $(0,1,1)_{12}$

$$Y(t) = \frac{(1+0.4691B)(1+0.8409B^{12})}{(1-B)(1-B^{12})}a_t \qquad (5)$$

Figure 6, Figure 7, Figure 8, and Figure 9 show the time series plots of water consumption (black) with forecasted values (red) for residential, commercial, governmental, and bulk, respectively.

Furthermore, Table II shows the 38 month forecasted values of water consumption in cubic meter from Nov 2010 to Dec 2013 in Davao City, Philippines. The expected total water consumption as a whole for 38 months is 215,349,797 cu.m with an average monthly consumption of 5,667,010 cu.m and an average yearly percentage increase of 3.44%, 182,832,272 cu.m with an average monthly consumption of 4,811,376 cu.m and an average yearly percentage increase of 3.60% for residential, 26,148,958 cu.m with an average monthly consumption of 688,130 cu.m and an average yearly percentage increase of 3.13% for commercial, 6,169,952.3 cu.m with an average monthly consumption of 162,367.17 cu.m and an average yearly percentage increase of 3.23%, and 150,364 cu.m with an average monthly consumption of 3,957 cu.m and an average yearly percentage decrease of 5.38% for bulk or seller.

c. Forecasting Using MLPNN Processes

Several models on the water consumption on all sectors were tested heuristically with the use of the Zaitun Time Series software.

Table III presents the possible MLPNN models for water consumption as a whole. It can be seen that the model with the lowest criterion, RMSE = 1.29e+5, was the ANN (36,6,1) with sigmoid function as the activation function. This is a three-layered model with 36 nodes in the input layer, 6 nodes in the hidden layer, and 1 node in the output layer. This is the model that was used for forecasting water consumption. The model was validated to be appropriate since the residuals were uncorrelated and random. Figure 10 presents the actual and predicted plot of water consumption as a whole. It is very obvious that the predicted graph (red) follow the pattern of the actual plot (blue).

The same analysis for MLPNN processes were applied to the water consumption in four sectors.

As a result, for residential, the model with the lowest criterion, RMSE = 2.28e+4, was the ANN (48,7,1) with hyperbolic tangent function as the activation function. For commercial, the model with the lowest criterion, RMSE = 4.25e+2, was the ANN (48,7,1) with bipolar sigmoid function as the activation function. However, this model failed on the normality test so as the models with RMSE = 4.34e+2 and RMSE = 1.20e+3. That's why the model with the fourth least value of criterion, RMSE = 5.04e+3, is the one that was utilized, and this was ANN (36,6,1) with bipolar sigmoid function as the activation function. For governmental, the appropriate model was the ANN (48,7,1) with sigmoid function as the activation function and RMSE = 7.84e+3. For the bulk or seller, the most appropriate model was the ANN (24,5,1) with sigmoid function as the activation function and RMSE = 6.49e+2. Figure 12 to Figure 15 show the actual and forecasted plots of the four sectors.

Table IV shows the 38 month forecasted values of water consumption in cu.m from Nov 2010 to Dec 2013 in Davao City using the MLPNN processes. The expected total water consumption as a whole for 38 months is 210,327,970 cu.m with an average monthly consumption of 5,534,946.6 cu.m and an average yearly percentage increase of 1.70%, 177,145,838 cu.m with an average monthly consumption of 4,661,732.6 cu.m and an average yearly percentage increase of 6.41% for residential, 25,902,436 cu.m with an average monthly consumption of 681,643.1 cu.m and an average yearly percentage increase of 1.66% for commercial, 6,167,412 cu.m with an average monthly consumption of 162,300 cu.m and an average yearly percentage increase of 4.78% for governmental, and 227,567 cu.m with an average monthly consumption of 5,988.6 cu.m and an average yearly percentage increase of 2.90% for bulk or seller.

In addition, the comparison between the forecasted values of water consumption as a whole and the total of the other four sectors showed no significant difference, both for ARIMA and MLPNN processes. Yet, in comparing the forecasted values generated by ARIMA and the ones generated by MLPNN processes, they showed no significant difference either. However, by looking at their RMSE, MLPNN showed smaller values. Thus, MLPNN manifested to be a better method in forecasting water consumption in Davao City, Philippines.

TABLE I: POSSIBLE ARIMA MODELS FOR WATER CONSUMPTION AS A WHOLE

MODEL	MSE	RMSE	AIC
ARIMA(0,1,1)X(0,1,1)	2.031e+10	1.425e+5	5047.39
ARIMA(0,1,2)X(0,1,1)	1.707e+10	1.307e+5	5016.32
ARIMA(0,1,3)X(0,1,1)	1.711e+10	1.308e+5	5017.70
ARIMA(0,1,4)X(0,1,1)	1.639e+10	1.280e+5	5012.50
ARIMA(0,1,5)X(0,1,1)	**1.622e+10**	**1.274e+5**	**5011.86**
ARIMA(0,1,1)X(1,1,0)	2.814e+10	1.677e+5	5093.55
ARIMA(0,1,2)X(1,1,0)	2.369e+10	1.539e+5	5064.52
ARIMA(0,1,3)X(1,1,0)	2.313e+10	1.521e+5	5062.37
ARIMA(0,1,4)X(1,1,0)	2.251e+10	1.500e+5	5059.10
ARIMA(0,1,5)X(1,1,0)	2.183e+10	1.477e+5	5055.99

TABLE II: FORECASTED VALUES OF THE WATER CONSUMPTION FROM NOV 2010 TO DEC 2013 (ARIMA)

YEAR	MONTHS	WHOLE	RES'L	COM'L	GOV'L	BULK
2010	November	5,508,578	4,654,981	668,650	162,796	4,468
	December	5,161,768	4,359,497	631,102	150,752	3,861
2011	January	5,642,918	4,802,124	684,574	158,576	4,687
	February	5,206,218	4,394,636	643,349	152,944	3,675
	March	5,281,167	4,472,944	644,238	152,392	4,068
	April	5,576,584	4,747,331	669,028	151,401	4,296
	May	5,498,620	4,677,304	664,355	150,182	4,356
	June	5,446,567	4,632,644	655,168	153,517	4,303
	July	5,552,308	4,712,809	674,167	165,699	4,504
	August	5,584,546	4,729,661	685,085	161,470	4,180
	September	5,566,171	4,717,914	683,136	161,204	4,359
	October	5,583,823	4,734,779	684,634	159,046	3,634
	November	5,633,110	4,763,194	688,227	167,950	4,250
	December	5,305,209	4,490,415	652,333	155,906	3,643
2012	January	5,889,667	5,018,388	705,805	163,730	4,470
	February	5,372,017	4,534,761	664,581	158,098	3,458
	March	5,475,624	4,644,124	665,469	157,548	3,851
	April	5,771,041	4,918,510	690,259	156,555	4,079
	May	5,693,077	4,848,484	685,586	155,336	4,138
	June	5,641,024	4,803,824	676,399	158,671	4,086
	July	5,746,766	4,883,989	695,398	170,853	4,286
	August	5,779,004	4,900,841	706,317	166,624	3,962
	September	5,760,628	4,889,094	704,367	166,358	4,141
	October	5,778,281	4,905,958	705,866	164,200	3,416
	November	5,827,567	4,934,374	709,458	173,104	4,033
	December	5,499,666	4,661,595	673,564	161,060	3,426
2013	January	6,084,125	5,189,568	727,036	168,884	4,252
	February	5,566,474	4,705,940	685,812	163,252	3,240
	March	5,670,081	4,815,304	686,700	162,702	3,633
	April	5,965,498	5,089,690	711,491	161,709	3,861
	May	5,887,534	5,019,664	706,817	160,490	3,921
	June	5,835,481	4,975,004	697,630	163,825	3,868
	July	5,941,223	5,055,169	716,630	176,007	4,069
	August	5,973,461	5,072,020	727,548	171,778	3,745
	September	5,955,086	5,060,273	725,598	171,512	3,924
	October	5,972,738	5,077,138	727,097	169,354	3,199
	November	6,022,024	5,105,553	730,690	178,258	3,815
	December	5,694,123	4,832,774	694,796	166,213	3,208

TABLE III: POSSIBLE MLPNN MODELS FOR WATER CONSUMPTION AS A WHOLE

MODELS	RMSE			
	ACTIVATION FUNCTION			
	Semi-linear	**Sigmoid**	**Bipolar Sigmoid**	**Hyperbolic Tangent**
ANN (12,4,1)	1.34e+6	1.46e+5	1.37e+5	1.50e+5
ANN (24,5,1)	1.43e+6	1.40e+5	1.35e+5	2.35e+5
ANN (36,6,1)	1.35e+6	1.29e+5	1.92e+5	1.72e+5
ANN (48,7,1)	1.20e+6	1.33e+5	1.37e+5	1.72e+5

TABLE IV: FORECASTED VALUES OF THE WATER CONSUMPTION FROM NOV 2010 TO DEC 2013 (MLPNN)

YEAR	MONTHS	WHOLE	RES'L	COM'L	GOV'L	BULK
2010	November	5,314,145	4,159,097	654,140	168,746	6,540
	December	5,332,810	4,702,080	672,737	169,418	5,420
2011	January	5,497,598	4,766,288	702,755	136,136	6,589
	February	5,225,563	4,220,771	656,895	138,945	6,018
	March	5,404,747	4,401,379	652,816	167,386	6,169
	April	5,391,365	4,813,711	694,345	147,148	5,268
	May	5,373,535	4,004,773	647,397	129,841	5,392
	June	5,529,328	4,619,984	667,023	149,018	5,063
	July	5,550,532	4,803,544	691,296	180,133	6,355
	August	5,517,755	4,524,075	656,355	197,794	5,601
	September	5,590,170	4,428,048	638,697	177,381	7,233
	October	5,450,547	4,762,220	699,215	138,953	6,495
	November	5,547,995	4,646,822	664,468	177,224	6,670
	December	5,408,504	4,475,768	669,518	196,425	4,866
2012	January	5,665,584	4,857,286	709,184	180,899	5,595
	February	5,355,914	4,553,840	671,516	118,303	4,450
	March	5,477,074	4,620,409	698,995	125,397	5,995
	April	5,460,236	4,754,311	706,634	151,153	5,829
	May	5,685,609	4,831,616	692,200	175,552	7,485
	June	5,451,012	4,503,029	663,586	125,872	6,217
	July	5,642,535	4,856,855	711,576	147,387	6,999
	August	5,622,169	4,755,652	653,726	163,784	5,162
	September	5,614,082	4,727,278	674,020	173,230	6,020
	October	5,587,256	4,757,019	687,493	158,843	4,366
	November	5,672,190	4,729,587	682,645	182,882	6,492
	December	5,380,904	4,510,192	677,782	162,385	5,776
2013	January	5,715,272	4,915,092	706,611	159,641	7,116
	February	5,474,604	4,688,156	676,403	144,607	6,153
	March	5,556,330	4,714,002	690,852	160,630	6,814
	April	5,676,748	4,865,731	711,346	170,572	5,504
	May	5,664,298	4,803,903	693,338	155,238	5,601
	June	5,588,234	4,610,370	677,921	142,878	4,475
	July	5,682,798	4,928,909	710,704	185,241	6,440
	August	5,655,559	4,861,607	687,979	193,677	5,559
	September	5,698,206	4,596,318	679,363	186,096	7,352
	October	5,664,893	4,927,944	711,336	172,648	6,160
	November	5,650,834	4,854,695	680,061	172,421	7,022
	December	5,551,032	4,593,477	679,510	183,528	5,303

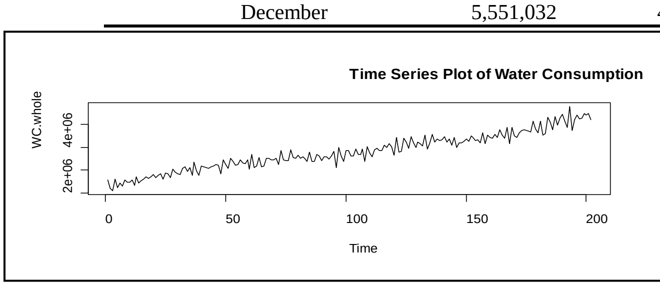

Fig. 1: Time Series Plot of Water Consumption as a Whole

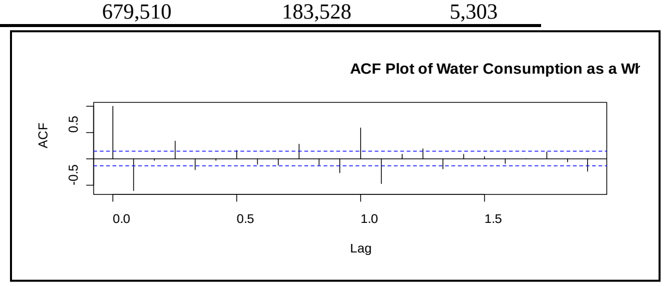

Fig. 2: ACF Plot of Water Consumption as a Whole with First Differencing

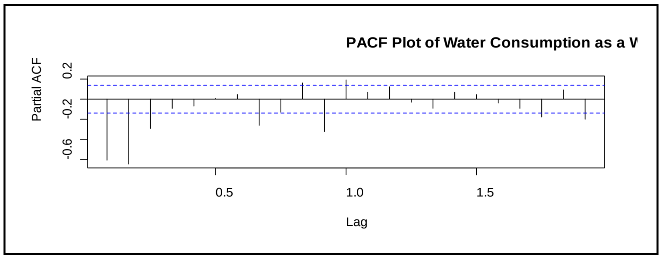

Fig. 3: PACF Plot of Water Consumption as a Whole with First Differencing

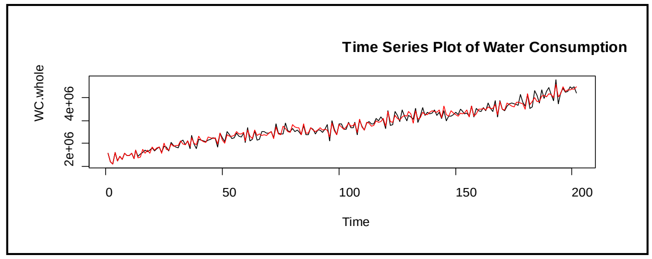

Fig. 4: Time Series Plot of Water Consumption as a Whole with the Predicted Values (ARIMA)

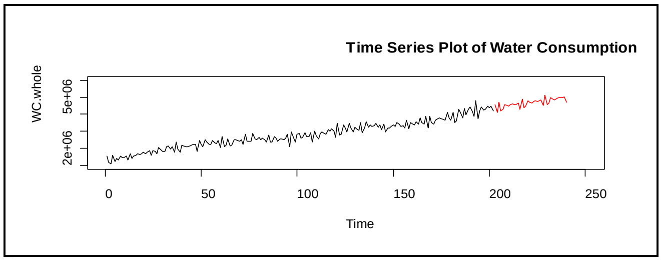

Fig. 5: Time Series Plot of Water Consumption as a Whole with 38 Month Forecasts (ARIMA)

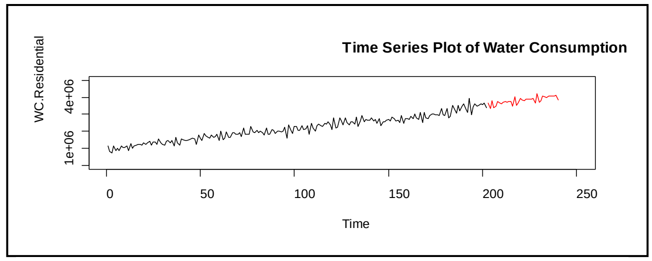

Fig. 6: Time Series Plot of Water Consumption as Residential with 38 Months Forecasts (ARIMA)

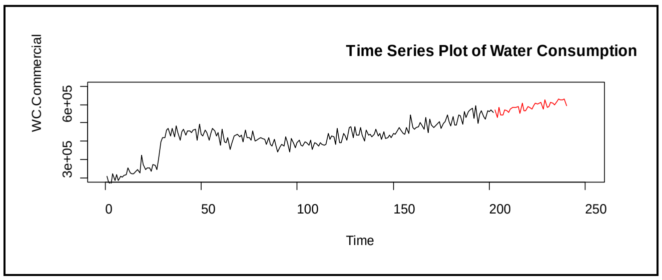

Fig. 7: Time Series Plot of Water Consumption as Commercial with 38 Months Forecasts (ARIMA)

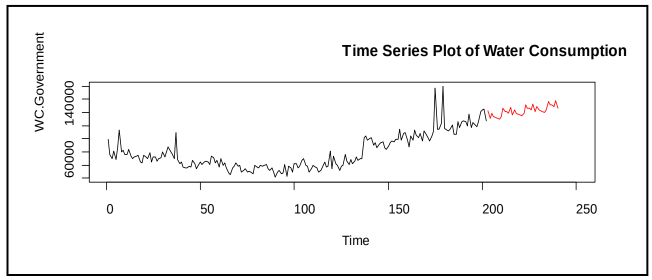

Fig. 8: Time Series Plot of Water Consumption as Governmental with 38 Months Forecasts (ARIMA)

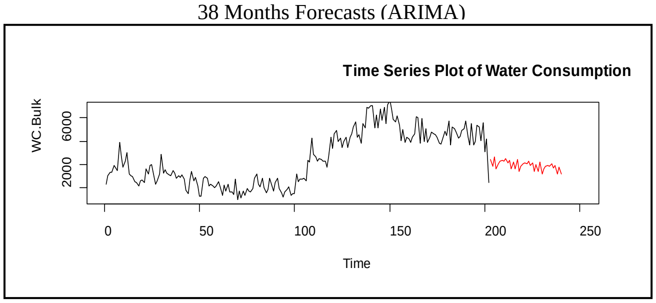

Fig. 9: Time Series Plot of Water Consumption as Bulk with 38 Months Forecasts (ARIMA)

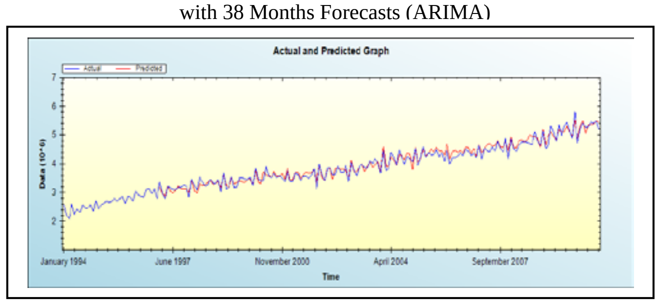

Fig. 10: The Actual and Predicted Plot of Water Consumption as a Whole (MLPNN)

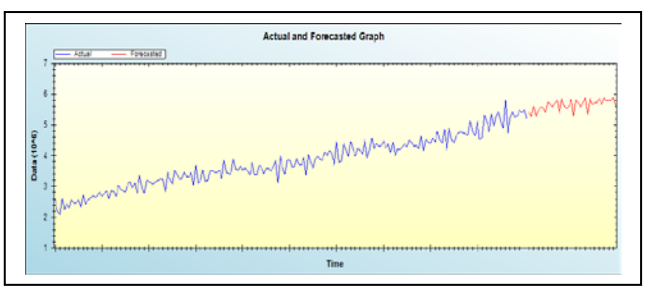

Fig. 11: The Actual and Forecasted Plot of Water Consumption as a Whole (MLPNN)

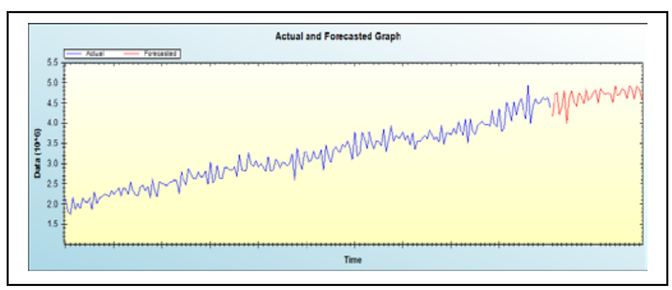

Fig. 12: The Actual and Forecasted Plot of Water Consumption as Residential (MLPNN)

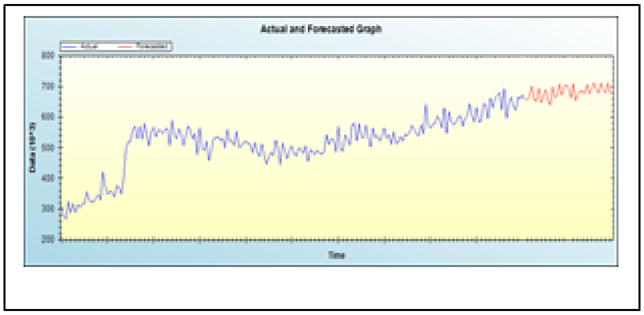

Fig. 13: The Actual and Forecasted Plot of Water Consumption as Commercial (MLPNN)

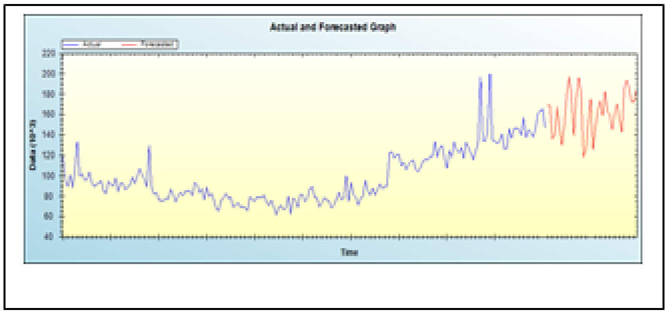

Fig. 14: The Actual and Forecasted Plot of Water Consumption as Governmental (MLPNN)

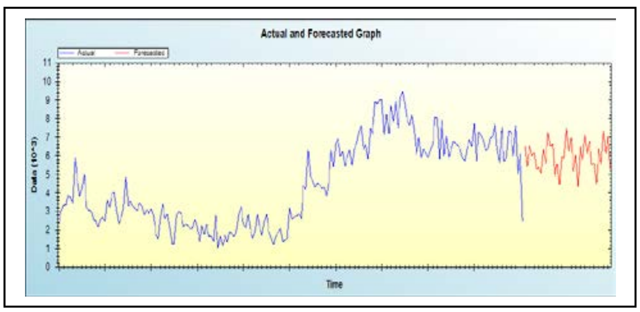

Fig. 15: The Actual and Forecasted Plot of Water Consumption as Bulk or Seller (MLPNN)

IV. Conclusion

It was shown from the results of this study that the water consumption for all four sectors will continuously increase. Further, it was also shown that the ARIMA models and the MLPNN processes are good statistical methods in forecasting time series data, especially the monthly water consumption. However, it came out that both models possess different weaknesses and strengths. Hence, proper investigation on the data and the model should be done.

As for the question which between ARIMA processes and MLPNN processes is the better process in forecasting water consumption in Davao City, it was shown here that only in water consumption as a whole that the ARIMA model is better, and the rests are for MLPNN models. Hence, MLPNN models are the more appropriate models in forecasting water consumption in Davao City, Philippines. The claim that tells that MLPNN processes are better than ARIMA models was also shown by some of the existing studies [3],[4],[10],[14].

Acknowledgment

I would like to thank the Natural Sciences and Mathematics Division of the Ateneo de Davao University through the Faculty Development Program and the Natividad Galang Fajardo Foundation for the Financial Assistance they render and Dr. Eveyth P. Deligero for guiding me through this paper.

References

[1] Azoff, E.M. "Neural Network Time Series Analysis Forecasting of Financial Markets". John Wiley & Sons Ltd, England, 1994.

[2] Box, G.E.P., and Jenkins, G.M. "Time series analysis: Forecasting and control". San Francisco, Holden-Day, 1970.

[3] Chakraborty, K., Mehrotra, K., Mohan, C.K., and Ranka, S., "Forecasting the Behavior of Multivariate Time Series using Neural Networks". Neural Networks, Volume 5, Issue 6, Pages 961-970, November - December 1992.

[4] Ghiassi, M., Zimbra, D.K., Saidane, H. "Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model". Journal of Water Resources Planning and Management, 2008.

[5] Granger, C.W.J., and Newbold, P. "Forecasting Economic Time Series". Economic Theory and Mathematical Economics, Academic Press Incorporated, New York, San Francisco, London, 1977.

[6] Sherrod, P.H. "DTREG: Predictive Modeling Software". DTREG Manual, 2003-2010.

[7] Jha, G.K. "Artificial Neural Networks". Indian Agricultural Research Institute, New Delhi, India, 2007.

[8] Masters, T. "Practical Neural Network Recipes in C++". Academic Press, New York, 1993.

[9] Mongaya, K.M. "Philippines: Manila Water Crisis". Global Voices, July 2010.

[10] Qi, C., and Chang, N.B., 2010. "System Dynamics Modeling for Municipal Water Demand Estimation in an Urban Region under Uncertain Economic Impacts". Journal of Environmental Management, Volume 92, Issue 6, Pages 1628-1641, June 2011.

[11] Shah, A. "Water and Development". Global Issues, June 2010.

[12] United Nations Press Release, "World Population will Increase by 2.5 Billion by 2050". POP/952, March 2007.

[13] Van Klaveren, P. "ADB, France Target Water Upgrades to Avert Potential Crisis in Philippine Cities". Asian Development Bank, November 2010.

[14] Yao, J., and Tan, C.L., 1997. "A Case Study on using Neural Networks to Perform Technical Forecasting of Forex". Neurocomputing, Volume 34, Issues 1-4, Pages 79-98, September 2000.

[15] Retrieved from http://en.wikipedia.org, "Artificial Neural Network".

[16] Retrieved from http://en.wikipedia.org, "Box-Jenkins ARIMA Process".

[17] Retrieved from http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report

[18] Retrieved from http://www.duke.edu, "Summary of Rules for Identifying ARIMA Models". Decision 411 Forecasting.

[19] Retrieved from http://www.realoptionsvaluation.com, "ARIMA: Advanced Forecasting Techniques and Models".

[20] Retrieved from http://www.statisticalforecasting.com, "Statistical Forecasting", 2006.