

Incorporating Future Climatic and Socioeconomic Variables in Water Demand Forecasting: A Case Study in Bangkok

Mukand S. Babel · Nisuchcha Maporn · Victor R. Shinde

Received: 25 July 2013 / Accepted: 16 March 2014 /

Published online: 20 April 2014

© Springer Science+Business Media Dordrecht 2014

Abstract With concerns relating to climate change, and its impacts on water supply, there is an increasing emphasis on water utilities to prepare for the anticipated changes so as to ensure sustainability in supply. Forecasting the water demand, which is done through a variety of techniques using diverse explanatory variables, is the primary requirement for any planning and management measure. However, hitherto, the use of future climatic variables in forecasting the water demand has largely been unexplored. To plug this knowledge gap, this study endeavored to forecast the water demand for the Metropolitan Waterworks Authority (MWA) in Thailand using future climatic and socioeconomic data. Accordingly, downscaled climate data from HadCM3 and extrapolated data of socioeconomic variables was used in the model development, using Artificial Neural Networks (ANN). The water demand was forecasted at two scales: annual and monthly, up to the year 2030, with good prediction accuracy (AAREs: 4.76 and 4.82 % respectively). Sensitivity analysis of the explanatory variables revealed that climatic variables have very little effect on the annual water demand. However, the monthly demand is significantly affected by climatic variables, and subsequently climate change, confirming the notion that climate change is a major constraint in ensuring water security for the future. Because the monthly water demand is used in designing storage components of the supply system, and planning inter-basin transfers if required, the results of this study provide the MWA with a useful reference for designing the water supply plan for the years ahead.

Keywords ANN · Climate change · Climate downscaling · Sensitivity analysis · Thailand · Water demand forecasting

1 Introduction

The reliable supply of safe drinking water is the primary objective of any water utility and over the years rapid advances have been made in this regard worldwide. However, with time, the

M. S. Babel (✉) · N. Maporn · V. R. Shinde

Water Engineering and Management, Asian Institute of Technology, Klong Luang, Pathumthani 12120, Thailand

e-mail: msbabel@ait.asia

challenges faced by the water sector have changed. In recent years, making provisions for maintaining a steady and regular supply under climate change regimes has become a major concern for water utilities. There is an increasing evidence which indicates that the trend of water availability under climate change will be significantly affected (Majumdar 2013; McFarlane et al. 2012; Beck and Bernauer 2011; Bates et al. 2008; IPCC 2007 etc.), which will have direct repercussions on the water utilities' ability to meet consumer demands.

An integral precursor of ensuring a reliable supply of drinking water is forecasting the water demand, which then forms the basis for planning supply/demand side management measures. Among the various techniques used to forecast the water demand, recent studies highlight the superiority of Artificial Neural Network (ANN) (Bennett et al. 2013; Campisi-Pinto et al. 2012; Babel and Shinde 2011; etc.). However, irrespective of the technique, success depends upon the type and number of explanatory variables used to make the forecast. In literature, variables affecting water demand are generally grouped into two classes: climatic, and socio-economic and demographic variables. While climatic variables—temperature, rainfall, relative humidity, wind speed, sunshine hours etc.—are usually associated with small-scale (daily, weekly) water demand (e.g. Jain et al. 2001), socio-economic and demographic variables like population, household connections, household income, education level etc. are found to affect the medium to large-scale (monthly, annual) water demand more (e.g. Gato et al. 2011). Interestingly, despite climate change concerns, the use of Global Climate Model (GCM) data, which provides forecasts for future climate conditions, in water demand forecasting has hardly been explored in literature. In perhaps the only study yet in this regard, Khatri and Vairavamoorthy (2009) used only precipitation and temperature data from the HadRM3 model to develop models to forecast the water demand for Birmingham, UK. The focus of their study, however, was to address uncertainties associated with not only climate change but also population and economic growth. Further, because of lack of adequate data they were unable to quantify the effect of future climate variables on the water demand.

This study seeks to plug in knowledge gaps in water demand forecasting by exploring the use of five future climate variables—Precipitation, maximum temperature, minimum temperature, evaporation, relative humidity (RH)—in forecasting the water demand. The objectives of this study were to (a) forecast, until 2030, the water demand for the Metropolitan Waterworks Authority (MWA) in Bangkok, using future climatic and socioeconomic variables, at two different scales: Annual and Monthly, and (b) develop a sensitivity index to quantify the effect of climate change on each scale of demand. The study employed ANN and Sensitivity Analysis to achieve these objectives.

2 Data Collection

The MWA is the primary water utility in Bangkok Metropolis - Thailand, supplying water to the domestic, industrial and irrigation sectors, covering 3,195 km². It receives raw water from two main intake sources: Chao Phraya River (up to 60 m³/s, depending on need) and inter basin transfer from Mae Klong River (up to 45 m³/s, depending on need). Two categories of variables have been used to forecast the water demand of MWA: Climatic, and socioeconomic. Climatic variables include rainfall, evaporation, relative humidity, minimum temperature, and maximum temperature. From a previous study (Babel and Shinde 2011), for the same study area, it was found that the average temperature did not correlate well with the water demand, hence only maximum and minimum temperatures were used in the current study. Socio-economic variables include: per capita Gross Provincial Product (GPP), population, number of houses, number of household connections and water tariff. The water tariff in MWA is

collected for individual connections, and has an increasing block structure. For example, for consumption up to 30 m³, the price is *Thai Baht* 8.50/m³, which increases to *Thai Baht* 10.03/m³ for consumption between 31 and 40 m³ (MWA 2009). Climatic data was collected from the Thai Meteorological Department for Bangkok Metropolis station, a meteorological station which lies within the MWA jurisdiction, while the socioeconomic data was collected from official publications of various government Ministries, and the MWA. In order to forecast the water demand until the year 2030, it was essential to project/forecast the values of each of these explanatory variables until the year 2030 so that they could be used as inputs in the model development. The next section describes the methods used to project these variables, and the subsequent outcomes.

3 Forecasting Explanatory Variables

3.1 Forecasting Climatic Variables

Climatic variables are forecasted with the help of GCMs, which are advanced tools to simulate the response of the global climate to increasing greenhouse gas concentrations, thereby providing estimates of future climatic conditions. A number of GCMs are in use but the use of the HadCM3 (Hadley Center (UK) Coupled Model, Version 3) is quite popular because of its good resolution (2.5°×3.75° latitude by longitude) and ability to make forecasts for different Special Report on Emissions Scenarios (SRES). This GCM has been used successfully for several studies in Thailand (e.g. Artlert et al. 2013; Thompson et al. 2013; Trisurat et al. 2011). The SRES scenarios, established by the IPCC (2001) are categorized into four storylines (A1, A2, B1, B2) which are built on sets of assumptions about possible future conditions. The A2 and B2 scenarios were considered for this study because it was endeavored to forecast the water demand under two diverse trends of future development. The A2 scenario portrays a very heterogeneous world where economic development is regionally oriented and per capita economic growth and change are more fragmented. In the B2 scenario, the emphasis is on local solutions to economic, social and environmental sustainability with intermediate levels of economic development (IPCC 2001).

Projection of the future climate conditions was done by statistical downscaling of GCM data. The purpose of downscaling is to generate regionally relevant data by developing quantitative relationships between predictors (large-scale atmospheric variables) and predictands (local surface variables). The Statistical Downscaling Model (SDSM, Wilby et al. 2002) was used in this study for downscaling GCM data. Downscaling essentially involves three stages — Screening the predictors (to identify the more pertinent predictors), Calibration (to develop the relationship between the predictors and predictands), and Validation (to test the relationship developed during calibration). This relationship is then applied to the GCM output to simulate future local climate.

The data for the five climatic variables was procured for the period 1961–2010. Predictor screening was done by examining the partial correlations and scatter plots between each predictor and the desired predictand for a set time duration. The choice of the number of predictors is usually subjective (Mahmood and Babel 2012) but the aim should be to choose predictors which display good strength of association with the predictand by both visual examination and statistical analysis. The chosen variables for each predictand were then used for calibration, and subsequent validation. Data corresponding to the period between 1961 and 1990 was used for calibration, while that for the period 1991–2010 was used for validation. The goodness of fit in calibration and validation was measured by the Coefficient of

determination (R^2), and Root Mean Square Error (RMSE). Table 1 presents the list of predictors that were chosen for each predictand, along with the calibration and validation results.

Calibration results were found satisfactory for all predictands, with R^2 ranging between 0.735 (maximum temperature) and 0.983 (monthly precipitation), coupled with low RMSE values. The validation results confirm that there is a very good agreement with the modeled and observed data, with $R^2 \geq 0.85$ and low RMSE for all predictands, suggesting that the relationship between each predictand and the corresponding predictors can be used for future projections. Hence, projections of the five predictands, using this relationship and GCM data, were made up to 2030, for both A2 and B2 scenarios, as shown in Fig. 1a, b and c.

3.2 Forecasting Socioeconomic Variables

Data for the five socioeconomic variables was procured for the period 1987–2010. The per capita GPP and population data was obtained from The National Economic and Social Development Board website, while the rest of the data was collected from the MWA head office. While all this data was used for the model development, only the data for number of household connections was used for the forecasting the water demand until 2030. This is because among all the socioeconomic variables the best-fit model for each scale of demand (as will be seen later in the paper) required only the number of household connections in order to produce the maximum prediction accuracy, thus rendering the remaining variables redundant. The projections for the number of household connections until 2030 were made by extrapolating the existing best-fit trend. Water sales records were used as observed water demand, which were available on a monthly basis from 1987 until 2010. This data was fed into the ANN models as expected output (i.e. observed demand), and the model performance was evaluated on how close the computed output (predicted demand) was to this expected output.

4 ANN Model Development

4.1 Methodology

ANN was used to model the water demand and then forecast it up to 2030, using the projected climatic and socioeconomic variables as explanatory variables. ANN attempts to simulate the

Table 1 NCEP predictors, calibration and validation results for the various predictands

Predictand	Predictor	Calibration		Validation	
		R^2	RMSE	R^2	RMSE
Maximum temperature (°C)	p5_u, r850, temp	0.735	0.577	0.961	0.286
Minimum temperature (°C)	shum, temp	0.935	0.409	0.985	0.495
Monthly precipitation (mm)	p_u, p_zh, p5_z, r500	0.983	18.06	0.964	25.74
Monthly evaporation (mm)	p_u, p_zh	0.903	7.857	0.906	0.467
Relative humidity (%)	p_u, p_zh, p8zh, shum, temp	0.852	1.249	0.850	3.724

p5_c 500 hPa zonal velocity, *r850* relative humidity at 850 hPa, *temp* mean temperature at 2 m, *shum* surface specific humidity, *p_u* surface zonal velocity, *p_zh* surface divergence, *p5_z* 500 hPa vorticity, *r500* relative humidity at 500 hPa, *p8zh* 850 hPa divergence, *p_f* surface airflow strength

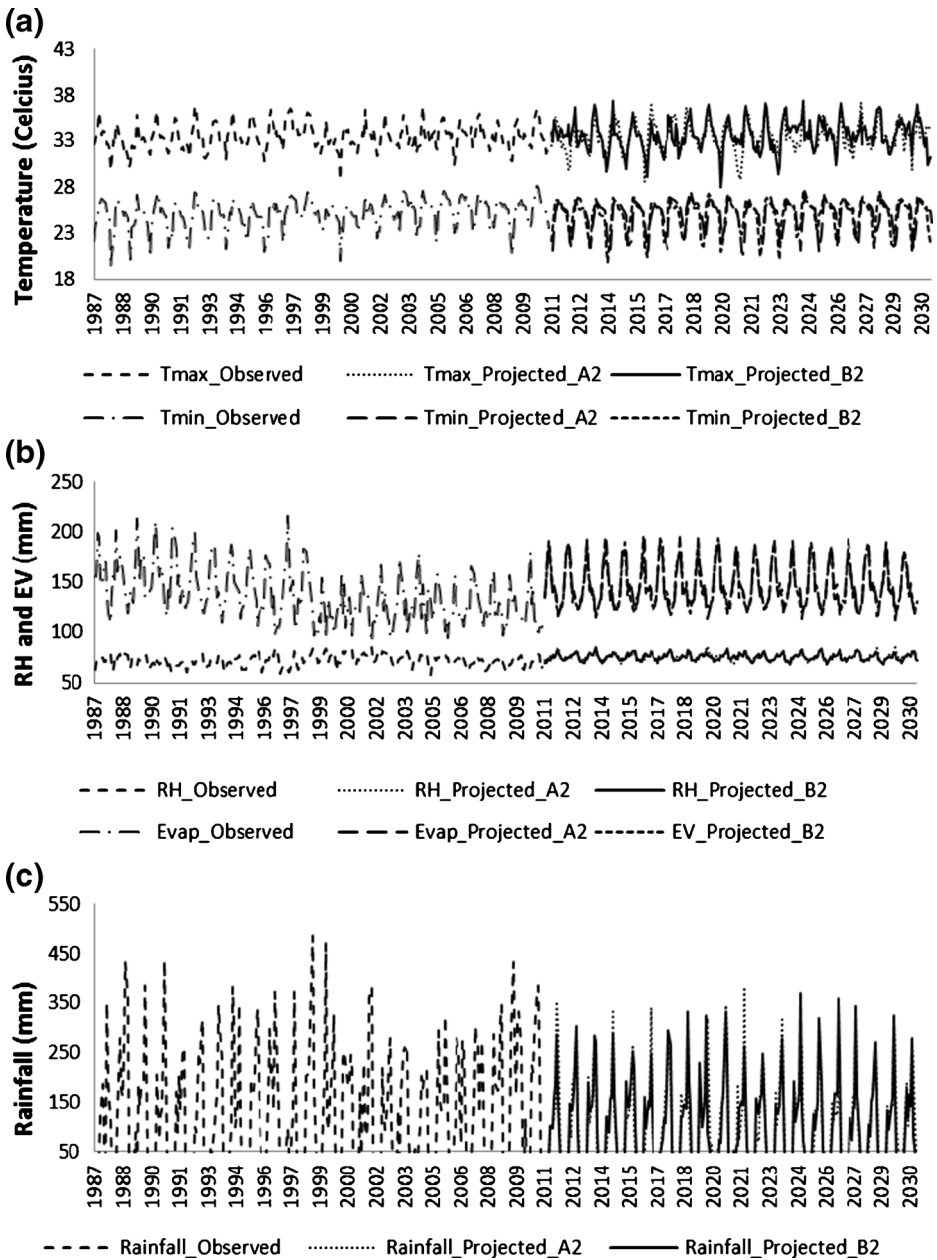


Fig. 1 Observed and forecasted trends of climatic variables in MWA service area

workings of the neurons in the brain by using a network of artificial neurons organized in layers, which receive a stimulus and, via a transfer function, mathematically convert it into an output signal (Babel and Shinde 2011). Developing an ANN model requires designing three major aspects: (a) Choosing an ANN architecture (which describes the flow of information in the model), (b) Determining the number of neurons (which are the basic building blocks of

ANN) and (c) Choosing an activation function (a non-linear function that translates input into output). Details of ANN structure, types of architecture and relevant terminology can be found elsewhere (e.g. Flood and Kartam 1994).

A detailed analysis of potential inputs variables is the first, and crucial, step in ANN modeling to minimize information loss and save computation time. Like Adamowski (2008) and Babel and Shinde (2011), this study used rank correlation analysis to select the relevant input variables. A correlation matrix was developed where the correlation coefficients (r) between each variable and water demand were calculated. Among variables that were highly correlated (magnitude greater than 0.8) to the observed demand, and to each other, only the variable which had the greatest correlation with the demand was chosen. Variables having lower correlation coefficient (less than 0.8) with the water demand were all included in the model.

The water demand was forecasted for two scales — Annual and monthly. All input data for the models corresponded to the period 1987–2030, out of which the data from 1987 to 2010 was used for training and testing. The testing set was approximately 20 % of the total observed data. The input data for the period 2010–2030, which was used to forecast the water demand, was obtained from the GCM models and extrapolation techniques described earlier. This study employed NeuroShell2, a popular ANN software, for model development. The performance of the models was evaluated by three common performance indicators (PIs) — R^2 , RMSE, and Average Absolute Relative Error (AARE). First, the best-fit architecture was identified by evaluating the performance of various ANN architectures against the three PIs. Then, the best-fit model for the best architecture was identified by checking the performance of the model by omitting the input variables in a systematic manner (described in the next section). With this best-fit model for each scale, the water demand up to 2030 was forecasted by using the projected values of the climatic variables for both A2 and B2 scenarios, and the extrapolated values of the socioeconomic variables.

4.2 ANN Models to Forecast the Water Demand

4.2.1 Annual Water Demand Models

Annual demand forecasting is required to plan activities like operation of reservoirs, pricing structuring, water allocation, implementing demand management measures, etc. Annual demand models in this context are the models which forecast the yearly water demand of the MWA. A correlation matrix between the annual water demand and potential explanatory variables was developed as presented in Table 2, after which the initial set of input variables for the ANN models was selected based on the conditions outlined earlier in Section 4.1.

As observed in Table 2, the per capita GPP, population, number of household connections, number of households, and tariff are all highly correlated with the annual water demand (expressed by correlation coefficients greater than 0.8). Further, the inter-correlation among these variables is also high ($r > 0.8$), suggesting that the inclusion of any one variable in the model development should suffice. The number of household connections was selected because it is most strongly correlated to the demand ($r = 0.98$). Hence, the initial input data set used for the first annual demand model comprised of: number of household connections, rainfall, evaporation, relative humidity, minimum temperature and maximum temperature.

With this initial input data set ANN models were developed (training) using five different architectures: Standard backpropagation, Recurrent network with feedback from input layer, Recurrent network with feedback from hidden layer, Recurrent network with feedback from output layer, and backpropagation with two hidden slabs having different activation functions

Table 2 Correlation matrix between potential explanatory variables for predicting MWA’s annual water demand

Variable	Demand	GPP	Pop	Conn.	Houses	Tariff	Rainfall	EV	RH	Tmin	Tmax
Demand	1	0.975	0.907	0.980	0.952	0.822	0.404	-0.725	-0.105	0.644	0.325
GPP		1	0.855	0.946	0.914	0.781	0.366	-0.687	-0.099	0.606	0.329
Pop			1	0.952	0.944	0.935	0.944	-0.808	0.118	0.662	0.139
Conn.				1	0.990	0.901	0.459	-0.792	0.028	0.681	0.283
Houses					1	0.901	0.457	-0.824	0.122	0.680	0.292
Tariff						1	0.346	-0.862	0.324	0.654	0.100
Rainfall							1	-0.500	0.240	0.053	-0.279
EV								1	-0.435	-0.383	0.100
RH									1	-0.036	-0.370
Tmin										1	0.631
Tmax											1

GPP Gross Provincial Product, *Pop* Population, *Conn* Number of household connections, *EV* Evaporation, *RH* Relative humidity, *Tmin* Minimum temperature, *Tmax* Maximum temperature

Table 3 Results of annual water demand prediction models

Model	Architecture	Neurons	Activation function	R^2			AARE (%)	RMSE (MCM)
				Train	Test	Total		
Y1	3-layer BackProp net	6-8-1	Linear [-1,-1], logistic, logistic	0.925	0.731	0.890	7.924	80.34
Y2	Recurrent net with input layer feedback	6-8-1-6	Linear [-1,-1], logistic, logistic	0.974	0.854	0.929	6.336	64.64
Y3	Recurrent net with hidden layer feedback	6-8-1-8	Linear [-1,-1], logistic, logistic	0.974	0.906	0.949	5.613	55.62
Y4	Recurrent net with output layer feedback	6-8-1-1	Linear [-1,-1], logistic, logistic	0.974	0.941	0.960	4.993	48.34
Y5	BackProp with 2 hidden slabs with different activation functions	6-4-4-1	Linear [-1,-1], Gaussian, Gaussian comp, logistic	0.960	-0.655	0.774	10.873	112.2
Y4(1)	Recurrent net with output layer feedback	5-7-1-1	Linear [-1,-1], logistic, logistic	0.974	0.957	0.963	4.705	46.07
Y4(2)		4-7-1-1		0.974	0.987	0.966	4.918	44.05
Y4(3)		3-6-1-1	0.992	0.977	0.966	4.762	44.40	
Y4(4)		2-6-1-1	0.954	0.979	0.950	5.315	48.59	
Y4(5)		2-6-1-2	0.963	0.981	0.952	5.486	48.86	
Y4(6)		1-5-1-1	0.958	0.979	0.951	5.347	47.96	

Highlighted fields represent best fit network, and best fit model of that network

Y4(1): number of connections, evaporation, relative humidity, minimum temperature and maximum temperature

Y4(2): number of connections, RH, maximum temperature and minimum temperature

Y4(3): number of connections, maximum temperature and minimum temperature

Y4(4): number of connections, maximum temperature

Y4(5): number of connections, minimum temperature

Y4(6): number of connections

(Readers are referred to Flood and Kartam 1994 for detailed descriptions of the architectures). The number of neurons in each layer was decided by trial and error. Table 3 presents the results of the analysis, where (in the top half of the Table) it is observed that that recurrent network with output layer feedback provides the best results against the specified PIs, with a high R^2 of 0.974 and 0.941 in training and testing sets respectively. Further, this network produced the lowest AARE (4.993 %) and RMSE (48.34 MCM) in both training and testing sets respectively.

To arrive at the best fit model for this network (Y4), one variable at a time was dropped from the input set, and the remaining variables were used for the model development. To elaborate on this consider the lower half of Table 3, which presents the results of the series of models developed for the Y4 network. First, the Y4(1) model was developed by omitting rainfall from the input data set. This model performed better (total $R^2=0.963$, AARE=4.705 % and RMSE=46.07 MCM) when compared to the Y4 model. It should be pointed out that if the model performance had not improved then rainfall would have been reinstated in the input set and another variable (e.g. RH) would have been omitted. Model results for this new set would have been then examined to check for improvement. Next, the Y4(2) model was developed by omitting evaporation and hence the input data set now had four variables (number of household connections, RH, maximum temperature and minimum temperature). Compared to Y4(1), this model performed better with respect to two out of the three PIs: total $R^2=0.966$ and RMSE=44.05 MCM. Seeking further improvement, the Y4(3) model was developed by omitting RH from the input data set. The AARE reduced to 4.762 %, while the total R^2 remained the same at 0.966, suggesting an overall improvement. Now, the Y4(4) model was developed by omitting minimum temperature from the input set. However, no improvement was observed against any of the PIs ($R^2=0.950$, AARE=5.315 %, and RMSE=48.59 MCM). Because there was no improvement, minimum temperature was reinstated in the Y4(5) model and maximum temperature was removed. Again, the model performance failed to improve ($R^2=0.952$, AARE=5.486 %, and RMSE=48.86 MCM). Finally, both minimum and maximum temperatures were removed and the model was developed with only the number of household connections as the input, which also did not cause any improvement ($R^2=0.951$, AARE=5.347 %, and RMSE=47.96 MCM) when compared to the results of Y4(3), which is the best-fit model for this scale of demand. Hence, only the number of household connections, minimum temperature and maximum temperature were used to forecast the annual demand of MWA.

Figure 2 shows a good fit between the observed and predicted trend of the annual water demand using model Y4(3). Also shown is the forecasted water demand for the period 2011–2030 for both A2 and B2 scenarios using the projected values of minimum and maximum temperature, and the extrapolated values of the number of household connections. There is no significant difference in forecasts for the A2 and B2 scenarios, and the maximum deviation in

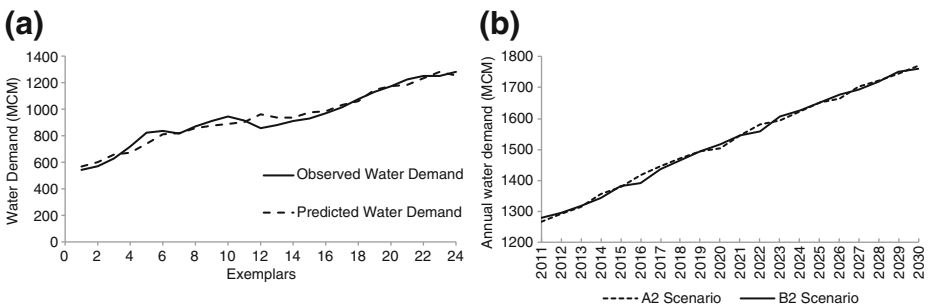


Fig. 2 Observed vs. predicted, and forecasted annual water demand using the Y4(3) model

the forecasts occurs in 2016, corresponding to 23 MCM, which is less than 2 % of the annual water demand. In both scenarios, the water demand is forecasted to increase by around 39 % in 2030 when compared to 2010, which is quite significant from a planning point of view.

4.2.2 Monthly Water Demand Models

Monthly demand forecasts are integral in planning storage facilities, inter-basin transfers, addressing seasonal fluctuations in water availability, etc. Further, because Thailand is a tourist country monthly demand is generally higher during the peak tourism season, thereby making a case for accurate monthly forecasts of the water demand. Using the same procedure, as that for the annual demand model development, the initial input data set for the monthly demand models was fixed based on correlating potential explanatory variables with the observed monthly water demand, as presented in Table 4. Accordingly, the initial input data set for the monthly demand models comprised of number of household connections, rainfall, evaporation, relative humidity, minimum temperature and maximum temperature.

Table 5 presents the monthly prediction results, where (in the top half of the Table) it is observed that that backpropagation with two hidden slabs with different activation functions (M5) outperforms the other architectures, with a R^2 of 0.958 and 0.918 in training and testing sets respectively, and lowest AARE (4.823 %) and RMSE (3.88 MCM).

Like for the annual demand models, the best-fit model of the M5 network was also identified by exploring the effect of omitting different explanatory variables as described in Table 5. It is clear that removing any variable from the initial data set does not improve the model performance, instead the performance deteriorates when even a single variable is removed, as seen from the results of models M5(1) to M5(6). The PI statistics for the M5 model are better than that for any of the sub-models, suggesting that all the six variables are required to make an accurate prediction of the MWA's monthly water demand. Figure 3 shows the ability of M5 model in predicting the monthly water demand. Further, an example of the forecast is also provided for the month of April using both A2 and B2 scenarios.

Again, it is apparent that the trend and magnitude of water demand forecasted for the two scenarios is quite similar. The maximum deviation in magnitude is 1.5 MCM (Year 2029),

Table 4 Correlation matrix between potential explanatory variables for predicting MWA's monthly water demand

Variable	Demand	GPP	Pop	Conn.	Houses	Tariff	Rainfall	EV	RH	Tmin	Tmax
Demand	1	0.958	0.910	0.963	0.933	0.823	0.130	-0.329	-0.014	0.295	0.161
GPP		1	0.881	0.942	0.909	0.792	0.068	-0.373	-0.048	0.201	0.107
Pop			1	0.960	0.952	0.938	0.095	-0.439	0.052	0.211	0.029
Conn.				1	0.991	0.906	0.098	-0.448	0.022	0.211	0.071
Houses					1	0.924	0.099	-0.468	0.069	0.215	0.071
Tariff						1	0.071	-0.463	0.150	0.211	0.017
Rainfall							1	-0.201	0.717	0.373	-0.001
EV								1	-0.240	0.370	0.664
RH									1	0.437	0.028
Tmin										1	0.761
Tmax											1

GPP Gross Provincial Product, *Pop* Population, *Conn* Number of household connections, *EV* Evaporation, *RH* Relative humidity, *Tmin* Minimum temperature, *Tmax* Maximum temperature

Table 5 Results of monthly water demand prediction models

Model	Architecture	Neurons	Activation function	r^2			AARE (%)	RMSE (MCM)
				Train	Test	Total		
M1	3-layer BackProp net	6-19-1	Linear [-1,-1], logistic, logistic	0.947	0.928	0.942	5.107	4.84
M2	Recurrent net with input layer feedback	6-19-1-6	Linear [-1,-1], logistic, logistic, logistic	0.947	0.829	0.933	5.317	4.93
M3	Recurrent net with hidden layer feedback	6-19-1-19	Linear [-1,-1], logistic, logistic, logistic	0.944	0.932	0.938	4.847	4.42
M4	Recurrent net with output layer feedback	6-19-1-1	Linear [-1,-1], logistic, logistic, logistic	0.946	0.930	0.940	5.169	4.92
M5	BackProp with 2 hidden slabs with different activation functions	6-10-10-1	Linear [-1,-1], Gaussian, Gaussian comp, logistic	0.958	0.918	0.951	4.823	3.88
M5(1)	BackProp with 2 hidden slabs with different activation functions	5-9-1-1	Linear [-1,-1], Gaussian, Gaussian comp, logistic	0.915	0.609	0.867	8.366	7.21
M5(2)		4-7-1-1		0.957	0.870	0.930	5.786	4.89
M5(3)		3-6-1-1		0.954	0.722	0.924	5.727	4.96
M5(4)		2-6-1-1		0.956	0.879	0.938	5.212	4.50
M5(5)		2-6-1-2		0.953	0.921	0.937	5.367	4.62
M5(6)		1-5-1-1		0.921	0.960	0.915	5.739	5.19
M5(6+)								

Highlighted fields represent best fit network, and best fit model of that network

M5(1): number of connections, evaporation, RH, maximum temperature, minimum temperature

M5 (2): number of connections, rainfall, RH, maximum temperature, minimum temperature

M5 (3): number of connections, evaporation, rainfall, maximum temperature, minimum temperature

M5 (4): number of connections, evaporation, rainfall, RH, maximum temperature

M5 (5): number of connections, evaporation, rainfall, RH, minimum temperature

M5 (6): number of connections

M5 (6+): number of connection, maximum temperature, minimum temperature

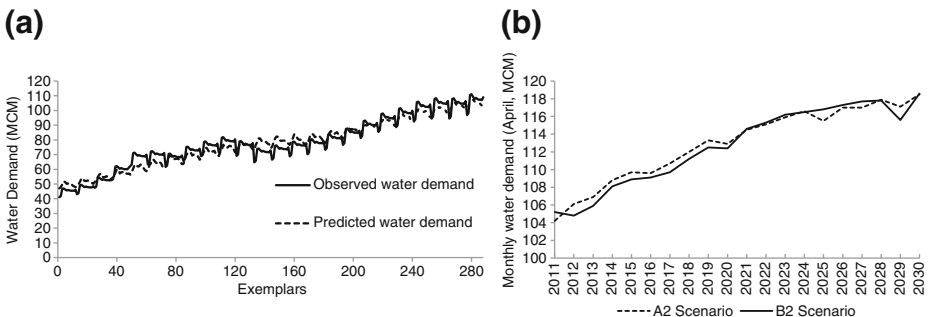


Fig. 3 Observed vs. predicted water demand, and forecasted monthly water demand for April using the M5 model

which corresponds to less than 1.3 % of the monthly demand. The monthly demand is forecasted to increase by around 15 % in 2030, compared to the base period of 2010.

The outcomes of this forecasting exercise clearly indicate that the water demand at both scales will increase with time, which has a heavy bearing on expansion endeavors. However, expansion activities in water supply are capital intensive, in which a major portion of the expenditures are taken up in the installation phase. Hence, initial investment needs are high and decision makers will need to carefully consider all aspects of the expansion plans, including the forecasts, before any financial commitment is made. Because the projected GCM climatic data is integral to making accurate forecasts, and because climate modeling is a developing science, it is crucial to understand, and quantify, the effect of climatic variables on the water demand at each scale, which will then help in developing response measures to deal with the uncertainty associated with the climate projections. The next section describes the technique used in this study to assess the impact of the climatic variables on the water demand for each scale of demand.

5 Sensitivity Analysis of Explanatory Variables

Sensitivity analysis was performed to identify the variables which are most likely to affect future water demand, for each scale of demand. The numeric value of each explanatory variable in the testing data set was iteratively increased and decreased by 10 and 30 % respectively, and the corresponding change in the output (water demand) was observed. Variables causing more change in the magnitude of the output (positive or negative) were deemed to be more sensitive. The sensitivities of the explanatory variables were quantified by developing a sensitivity index, which also would facilitate comparison between variables. In context of this study, the sensitivity index of an explanatory variable is the standard deviation of the percentage change in demand caused by varying the magnitude of the variable by ± 10 and ± 30 %.

Figure 4 portrays the sensitivities of each variable for both annual and monthly water demand forecasts. For the annual water demand, it is seen that the number of household connections is far more sensitive than the other variables. The sensitivity index of number of household connections is almost ten times that of both maximum and minimum temperatures, which suggests the redundancy of climatic variables in forecasting MWA's annual water demand. Conversely, the climatic variables have a significant effect on the monthly water demand, among which the maximum temperature and evaporation variables have the highest sensitivities. The maximum temperature variable is particularly significant because it account for the highest sensitivities during the summer season (March–June), usually associated with water shortage. Further, climate change is expected to intensify the summer period in this region, which may lead to extended periods of water shortage if adequate preventive measures are not in place.

The results of the study bring up some interesting points of discussion. First, climate change is unlikely to affect MWA's future annual water demand because climate variables have very little influence on this scale of demand. The socioeconomic variables, especially the number of household connections, appear to have a greater impact on the demand. This observation is in line with that made by Khatri and Vairavamoorthy (2009), who also reported that the effect climatic variables on Birmingham city's forecasted water demand for 2035 is negligible. Second, as opposed to the annual demand, the monthly water demand forecasted for MWA is significantly affected by climatic variables, and subsequently climate change. A possible reason for this phenomenon is that while climate change will affect the weather pattern throughout the year, the change will be more severe in certain months, thereby making the monthly demand more sensitive to changes in climate regimes. To explore this aspect further,

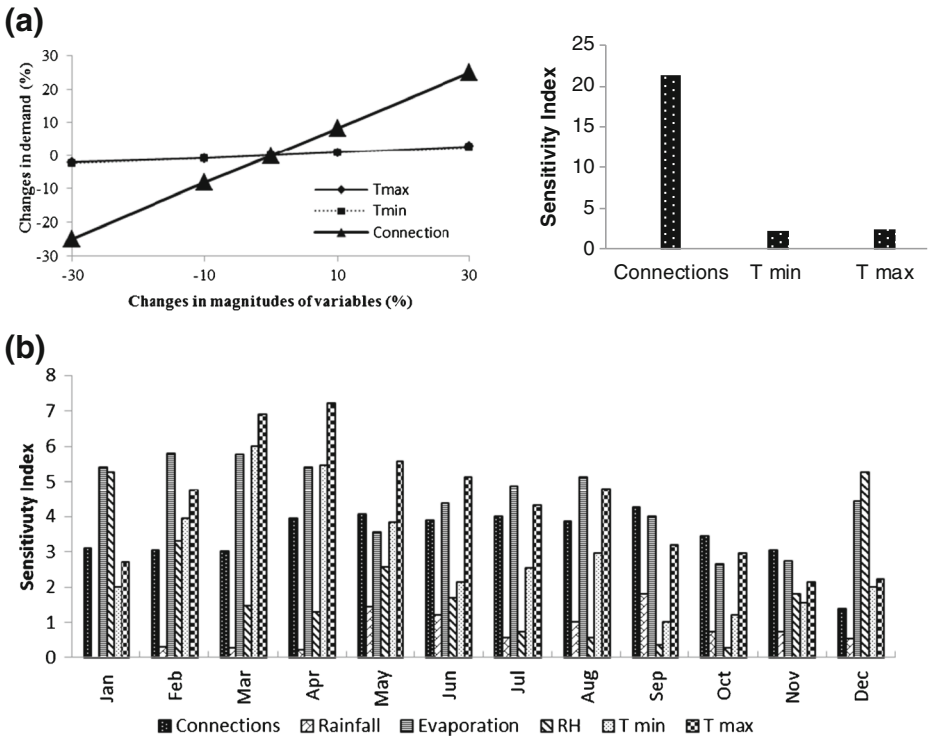


Fig. 4 Sensitivities of explanatory variables used for forecasting the **a** annual and **b** monthly water demands

two additional models were developed with the same ANN architecture but with different explanatory variables. Only the number of household connections was used in the first model, while the second model used the number of connections along with the maximum and minimum temperatures. Monthly forecasts up to 2030 were made for each model, and the two sets of forecasts were compared to investigate the influence of the temperature variable on the demand. As seen in Fig. 5, it is quite evident that the inclusion of the temperature variable

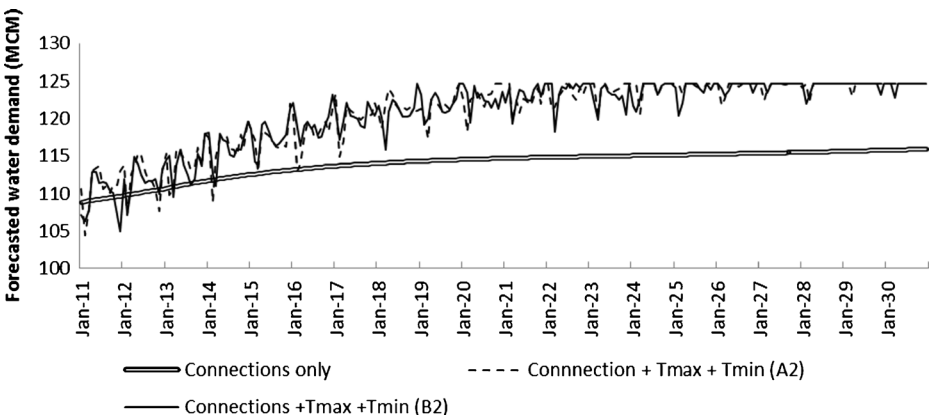


Fig. 5 Effect of the temperature variable on the monthly forecast of MWA's water demand

in the models results in larger values of the forecasted water demand (for both A2 and B2 scenarios), compared to the forecasts made by considering only the number of household connections. This clearly underlines the significance of including climatic variables (temperature, in this case) to make monthly demand forecasts of the MWA. Moreover, the models which used the temperature variables display a zigzag trend of the forecasted demand, which indicates increased demand in certain months of the year. This reinforces the notion suggested earlier that the effect of climate change is more pronounced on a monthly scale rather than the annual scale.

Because the monthly water demand is extensively used in planning and designing the supply system, it can be inferred that climate change will have a heavy bearing in this regard in designing future expansion plans. Also given the fact that the influence of the climatic variables on the water demand is the strongest during the summer season, storage facilities will need to be carefully designed to ensure that there is no shortage of water. Thirdly, rainfall does not seem to significantly influence the water demand at any scale. This is quite significant because between the two key variables expected to be affected by climate change—rainfall, and temperature—only effects of temperature need to be taken into account to make accurate forecasts of the water demand.

6 Conclusions

This study was carried out to explore the use of future climatic and socioeconomic variables in forecasting the water demand for the Metropolitan Waterworks Authority (MWA) in Thailand. Accordingly, downscaled climate data from HadCM3 and extrapolated data for socioeconomic variables was used in the model development, using Artificial Neural Networks (ANN). The water demand was forecasted at the annual and monthly scales, up to the year 2030, with good prediction accuracy (AAREs: 4.76 and 4.82 % respectively). While this prediction accuracy is good enough for all practical purposes, further improvement may be possible if certain other techniques like bootstrapping were used to project the socioeconomic variables. Sensitivity analysis of the explanatory variables used in the model development revealed that the number of household connections is the most crucial variable in forecasting the annual water demand, while the climatic variables have virtually no affect. However, climatic variables, especially the maximum temperature evaporation, play a significant role in forecasting the monthly water demands. This is because while climate change will affect the weather pattern throughout the year, the change will be more severe in certain months, thereby making the monthly demand more sensitive to changes in climate regimes. This study confirms the notion that climate change is a major constraint in ensuring water security for the future: Planning for future water supply measures must consider the effects of climate change.

References

- Adamowski JF (2008) Peak daily water demand forecasting modeling using artificial neural networks. *J Water Resour Plan Manag* 134(2):119–128
- Artlert K, Chaleeraktragoon C, Nguyen V-T-V (2013) Modeling and analysis of rainfall processes in the context of climate change for Mekong, Chi, and Mun River Basins (Thailand). *J Hydro Environ Res* 7(1):2–17
- Babel MS, Shinde VR (2011) Identifying prominent explanatory variables for water demand prediction using artificial neural networks: a case study of Bangkok. *Water Resour Manag* 25:1653–1676
- Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (eds) (2008) Climate change and water: technical paper of the intergovernmental panel on climate change. IPCC Secretariat, Geneva, 210 pp

- Beck L, Bernauer T (2011) How will combined changes in water demand and climate affect water availability in the Zambezi river basin? *Glob Environ Chang* 21(3):1061–1072
- Bennett C, Stewart RA, Beal CD (2013) ANN-based residential water end-use demand forecasting model. *Expert Syst Appl* 40(4):1014–1023
- Campisi-Pinto S, Adamowski J, Oron G (2012) Forecasting urban water demand via wavelet-denoising and neural network models. Case study: city of Syracuse, Italy. *Water Resour Manag* 26(12):3539–3558
- Flood I, Kartam N (1994) Neural networks in civil engineering. I: Principles and understanding. *J Comput Civil Eng* 8(2):131–148
- Gato S, Jayasuriya N, Roberts P (2011) Understanding urban residential end uses of water. *Water Sci Technol* 64(1):36–42
- IPCC (2001) *Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, 881pp
- IPCC (2007) *Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change*, Geneva, Switzerland. pp 104
- Jain A, Varshney AK, Joshi UC (2001) Short term water demand forecast modeling at IIT Kanpur using artificial neural networks. *Water Resour Manag* 22(6):299–321
- Khatri KB, Vairavamoorthy K (2009) Water demand forecasting for the city of the future against the uncertainties and the global change pressures: case of Birmingham. In proceeding of EWRI/ASCE: 2009 conference, Kansas, USA May 17–21
- Mahmood R, Babel MS (2012) Evaluation of SDSM developed by annual and monthly sub models for downscaling temperature and precipitation in the Jhelum Basin, Pakistan and India. *Theor Appl Climatol*. doi:10.1007/s00704-012-0765-0
- Majumdar PP (2013) Climate change: a growing change for water management in developing countries. *Water Resour Manag* 27(4):953–954
- McFarlane D, Stone R, Martens S, Thomas J, Silberstein R, Ali R, Hodgson G (2012) Climate change impacts on water yields and demands in south-western Australia. *J Hydrol* 475:488–498
- MWA (2009) Annual report. Metropolitan Waterworks Authority, Bangkok
- Thompson JR, Green AJ, Kingston DG, Gosling SN (2013) Assessment of uncertainty in river flow projections for the Mekong River using multiple GCMs and hydrological models. *J Hydrol* 486:1–30
- Trisurat Y, Shrestha RP, Kjelgren R (2011) Plant species vulnerability to climate change in Peninsular Thailand. *Appl Geogr* 31(3):1106–1114
- Wilby RL, Dawson CW, Barrow EM (2002) SDSM—a decision support tool for the assessment of regional climate change impacts. *Environ Model Softw* 17(2):145–157