STOCHASTIC MODELLING OF WATER DEMAND USING A SHORT-TERM PATTERN-BASED FORECASTING APPROACH

Ir. LAM Shing Tim
Development(2) Division, Development Branch, Water Supplies Department.

Abstract: Water demand is an important parameter in the planning of a water supply system. It usually exhibits both deterministic patterns and stochastic components. Traditional planning approach is based on a set of deterministic design capacity factors rather than evaluating system reliability in response to the stochastic nature of water demand and supply patterns. As knowledge of system reliability is increasingly recognised as essential for component sizing, operation optimisation and contingency planning, a new planning approach based on stochastic modelling of water demand using a short-term pattern-based forecasting approach for demand simulation has been developed and tried out in planning for uprating a salt water supply system in Wan Chai. The water supply system planned with this approach results in HK$20 million saving in life-cycle cost without comprising the reliability of the system.

1. INTRODUCTION

1.1 Traditional Approach of Water Demand Modelling

Planning of water supply system aims to ensure that the system has adequate capacity for supplying water to meet various demand scenarios. Water demand is therefore an important parameter in the planning process. Traditional planning approach involves building a water demand model and a hydraulic model of the water supply system for simulation of the performance of system components (including reservoir, pumping station and water mains) in response to various demand scenarios. Unit demands for different consumer categories are established based on historical average consumption record. The design water demand for a water supply system is traditionally modelled by multiplying such unit demands by the corresponding consumer size (e.g. population or area etc) and then superimposed by a peak factor to cater for possible demand fluctuation above the average demand.

1.2 Limitations of Traditional Approach

While this traditional water demand modelling generally bodes well for planning water supply systems where there are little constraints (e.g. site availability limiting the capacity of a service reservoir) or the constraints carry insignificant penalty, on occasions, there may be constraints with severe penalty so much so that a suitable departure from the standard established based on the traditional approach would present a very desirable option. Because of the averaging effect of the traditional approach, the fluctuation in the water demand in real-life situation is masked completely; a more rigorous approach than the traditional approach is called for to deal with such situations.

1.3 New Approach with Stochastic Modelling for Synthetic Demand Generation

To overcome the limitations of the traditional approach, a stochastic model is developed for synthetic demand generation of which the results can generate probable extreme demand and can be input to the hydraulic model of water supply system for continuous extended period simulation (EPS) so as to simulate system performance for reliability analysis. In general, water demand modelling can be based on either a physical approach or a numerical approach. The former involves modelling physical factors that affect water demand directly or indirectly, such as temperature, climate, etc. while the later involves numerical modelling using historical water demand data. Stochastic modelling usually adopts a numerical approach since it is usually formulated by using statistical and probabilistic models that are built on historical data. With historical water demand data of good quality captured by System Control and Data Acquisition (SCADA) system, it is feasible to analyse the deterministic patterns and stochastic components and formulate a suitable stochastic model using the patterns and probabilistic parameters based on the analyses results.

1.4 Application of Forecasting Approach

In the development of pattern-based model for water demand forecasting, Alvis *et al* (2007) illustrated that the daily and hourly water demands demonstrate short-term memory components in a similar manner as the long-term memory effects inherent in traditional streamflow forecasting and flood routing models. It is revealed that deviation between daily demand generated by periodic components tomorrow and the actual average demand of the same day could be correlated with the deviation for daily demand generated today. Similarly deviation between hourly demand generated by periodic components in the next hour and the actual average hourly demand of the same hour could be correlated with the deviations for hourly demand generated 1 hour and even 24 hours ago. In this connection, the short-term pattern-based forecasting model is composed of the deterministic seasonal, weekly and hourly patterns as well as the short-term memory components inherent in the stochastic component. To enable realistic generation of synthetic daily and hourly demands, the stochastic demand models should incorporate both deterministic demand patterns and short-term memory components that are able to take into account short-term memory effect of the demand generated for the previous day or hours, as the case may be.

2. MODEL FORMULATION

2.1 Model Overview

Demand fluctuation usually exhibits deterministic seasonal, weekly and daily patterns, and stochastic daily and hourly components. The stochastic model comprising deterministic and stochastic components is therefore developed. The deterministic component is based on the short-term pattern-based forecasting approach developed by Alvisi *et al* (2007). The stochastic model simulates daily demand and hourly demand of the day by combining simulations of deterministic patterns and stochastic components using time-series forecasting and probabilistic models. Deterministic patterns and parameters of the forecasting and probabilistic models are determined by statistically analysing the historical operation data extracted from database of SCADA system. Residuals can be analysed by an auto-regressive time-series model. Residual values of the daily and hourly nodal demands include random components. Various components of the stochastic demand model are illustrated in Figure 1.

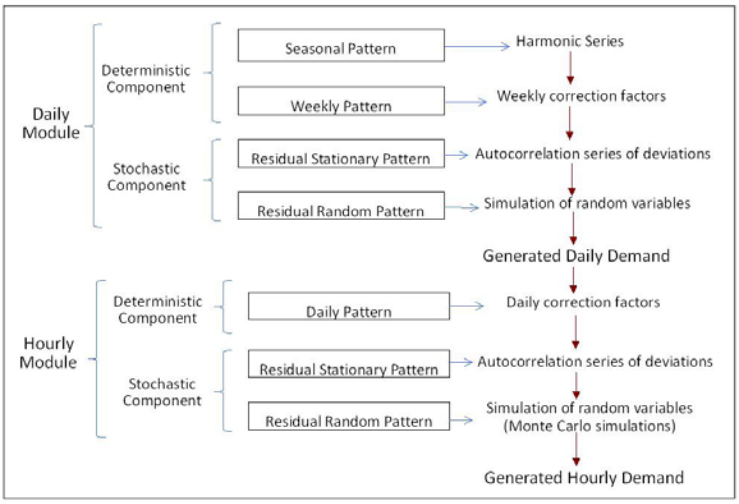

Figure 1 - Structure of Stochastic Demand Model

2.2 Historical Demand Data

The stochastic model of water demand is built on historical water demand data based on the operation database captured by SCADA system for an existing water supply zone. The resolution of zonal demand data is adequate for formulating a stochastic demand model for simulating performance of major system components such as service reservoir, pumping station and trunk mains. For minor system components such as distribution mains or finer resolution at individual villages or service points, it is desirable to break down the constructed zonal demand data into a number of nodal demand data or to make use of historical demand data captured by relevant District Metered Areas or individual bulk meters covering a small area, village or housing estate.

Data cleansing of the SCADA data is necessary in order to discard the SCADA data captured during emergency conditions in which demand patterns deviate from the normal deterministic patterns and affect the short-term memory components. Such abnormal conditions include occurrence of data communication problem, occurrence of pipe bursts, shut down of pumping station, change in supply zone, etc.

2.3 Modelling Deterministic Components

The deterministic component of the demand model comprises seasonal, weekly and hourly patterns. The seasonal pattern is modelled by fitting a harmonic series to the annual series of average observed demand of each day of 365 days. Applying equations 2.10 to 2.13 in Kottegoda (1980), the estimated mean daily demand for days 1 to 365 by using all 182 pairs of harmonic coefficients is modelled using harmonic series –

$$ {}_{av}Q_{\tau}^{d,har} = \mu^{yr,\,obs} + \sum_{i=1\ to\ H}\{{}_{est}\alpha_i \sin(2\pi i\tau/365)\} + \sum_{i=1\ to\ H}\{{}_{est}\beta_i \cos(2\pi i\tau/365)\} $$

where τ = 1 to 365, H = number of harmonics used and the estimated harmonic coefficients are

$$ {}_{est}\alpha_i = (2/365) \sum_{\tau=1\ to\ 365}\{{}_{av}Q_{\tau}^{d,obs} \sin(2\pi i\tau/365)\} \quad and $$
$$ {}_{est}\beta_i = (2/365) \sum_{\tau=1\ to\ 365}\{{}_{av}Q_{\tau}^{d,obs} \cos(2\pi i\tau/365)\} \quad \text{for } i = 1, 2,\ldots,365/2-0.5=182. $$

Maximum 182 harmonics can be used to fully represent an annual daily series of 365 days. However it is unnecessary to fit all of the 182 harmonics. It is sufficient to use a minimum number of harmonics required for fitting the annual daily data. The number of harmonics required is determined by using method of analysis of variances. Fitting with more than the minimum number of harmonics required would bear the risk of incorporating undesirable stochastic component in the deterministic model.

The weekly pattern is based on a set of weekly correction factors which are the difference between the average weekday demand and the average weekly demand and are computed for each weekday for each month. The weekly correction factor for daily demand is defined by

$$ {}_{av}\Delta_{i,j}^{d,w} = {}_{av}Q_{i,j}^{d,w} - {}_{av}Q_{,j}^{w} $$

where ${}_{av}Q_{i,j}^{d,w}$ is the average trend-free daily demand (in L/sec) observed in weekday i (i = 1 to 7) in month j (j = 1 to 12) and ${}_{av}Q_{,j}^{w}$ is the average weekly demand in month j. The weekly correction factor for hourly demand is defined by

$$ {}_{av}\Delta_{n,i,,j}^{h,d} = {}_{av}Q_{\gamma,i,j}^{h} - {}_{av}Q_{i,j}^{d,w} $$

where ${}_{av\text{-}}Q_{\gamma,i,j}^{h}$ is the average trend-free hourly demand (in L/sec) observed in (for hours γ = 1 to 24) observed in day i of the week (i = 1 to 7) and month j (j = 1 to 12) and ${}_{av}Q_{i,j}^{d,w}$ is the average daily demand (in L/sec) observed in day i of the week (i = 1 to 7) and month j (j = 1 to 12).

2.4 Modelling Stochastic Components

By subtracting the deterministic model demand pattern from the observed average demand pattern, the observed deviations $\delta_i^{d,obs}$ are calculated by

$$\delta_i^{d,obs} = {}_{av\text{-}tf}Q_i^{d,obs} - ({}_{av}Q_\tau^{d,har} + {}_{av}\Delta_{i,j}^{d,w}),\ \text{for i = 1 to 365,}$$

similarly, the observed hourly deviations $\xi_t^{h,obs}$ are calculated by

$$\xi_t^{h,obs} = {}_{av\text{-}tf}Q_{t,\,yr}^{h,obs} - ({}_{av\text{-}tf}Q_{i,yr}^{d,obs} + {}_{av}\Delta_{n,i,,j}^{h,d}) ,$$

for i = 1 to 365 days and h = 1 to 8760 hours and yr = the year observed for hourly demand at hour h. The stochastic components are featured by serial correlating characteristics which represent the short-term memory component. These characteristics can be identified by time series analysis and modelled by a suitable type of ARMA (auto-regressive and moving average) model. In general, the appropriate time series models identified for the stochastic components of these daily demand pattern and hourly demand pattern are AR(1) model and AR(1) cum AR(24) model respectively. The daily residuals δ_i^d and hourly residuals ξ_t^h are modelled by

$$\delta_i^d = \phi * \delta_{i-1}^d \quad \text{and} \quad \xi_t^h = \psi_1 * \xi_{t-1}^h + \psi_2 * \xi_{t-24}^h + c_o$$

where ϕ, ψ_1, ψ_2 are parameters of AR models and c_o is a constant.

2.5 Modelling Random Components

The observed residuals of the model demand comprising the deterministic and stochastic components are random in nature. The observed daily residuals $\varepsilon_{i,\,yr}^{d,obs}$ and hourly residuals $\varepsilon_{t,\,yr}^{h,obs}$ represent the random components of the demand patterns and are modelled by

$$\varepsilon_{i,\,yr}^{d,obs} = Q_{i,\,yr}^{d,obs} - Q_i^{d,for} \quad \text{and} \quad \varepsilon_{t,\,yr}^{h,obs} = Q_{t,yr}^{h,obs} - Q_t^{h,est} ,$$

where $Q_i^{d,for} = {}_{av}Q_i^{d,har} + {}_{av}\Delta_{i,j}^{d,w} + \delta_i^{d,est} (=\phi * \delta_{i-1}^{d,obs})$ and $Q_t^{h,est} = Q_i^{d,for} + {}_{av}\Delta_{n,i,,j}^{h,d} + \xi_t^{h,est}$,

Statistical analysis of the observed daily demand residuals and hourly demand residuals revealed that the density distributions of the residuals are normal random distributions with zero mean and the simulated daily residuals $\varepsilon_i^{d,gen}$ and hourly residuals $\varepsilon_t^{h,gen}$ are modelled by

$$\varepsilon_i^{d,gen} = m + \sigma_\varepsilon^d * z_i \quad \text{and} \quad \varepsilon_t^{h,gen} = m + \sigma_\varepsilon^h\ z_t$$

where $m = 0$, σ_ε^d = standard deviation of daily demand, σ_ε^h = standard deviation of daily demand and z_i & z_t = independent standard normal random variables.

2.6 Algorithm of Demand Generation

At the start of a day of a year, the daily seasonal and weekly pattern demand components are calculated to form the daily deterministic demand of the day. With the deterministic daily demand, the hourly deterministic demand of each hour of the day is calculated by applying the hourly demand pattern. For the stochastic daily and hourly components, the time-series component is calculated by the auto-regression equation and random numbers are generated from the corresponding normal distributions to form the randomized component.

A uniform random number (value between 0 and 1) is drawn for the day and another uniform random number is drawn for every hour of the year, and then such random numbers are transformed to normal random numbers for the simulation of the residual components. Summation of the deterministic daily/hourly demand component, the stochastic daily/hourly demand component and the corresponding residual components gives the generated daily/hourly demand. Flow charts showing the generation algorithms for daily and hourly demands are shown in Figures 2 and 3 respectively.

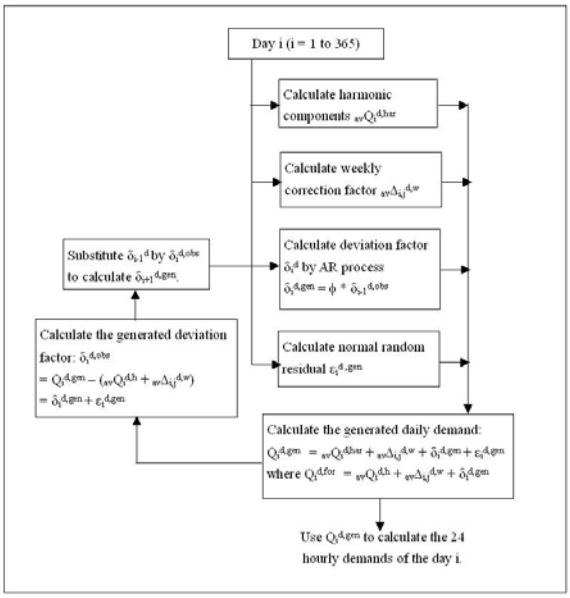

Figure 2 – Algorithm of Daily Demand Generation

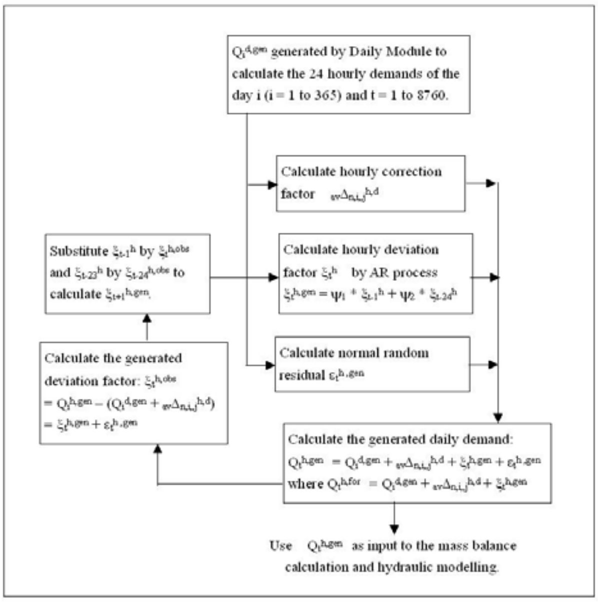

Figure 3 – Algorithm of Hourly Demand Generation

2.7 Demand Generation for Monte Carlo Simulation

Monte Carlo simulation of system performance involves repeating the stochastic demand generation process successively for either 365 days or 8760 hours of a year and continuing the generation process for subsequent years. Due to the stochastic nature of the demand generation process, every annual hourly pattern generated is unique. The entire annual daily and hourly patterns are input at an hourly time-step to the hydraulic model for EPS so as to simulate system responses of the year. With a sufficiently large number of annual daily and hourly demands, the system responses can be statistically analysed for evaluation of system performance under all possible scenarios of demand patterns covering all probability of occurrence.

3. RESULTS

3.1 Water Demand Generation

Based on historical hourly demand data captured by SCADA system from 2005 to 2008, stochastic water demand models are established for a fresh water supply zone in Tai Po and a flushing water supply zone in Wan Chai for generation of water demands in year 2009. The model parameters are summarized in Table 1. As illustrated by the small R^2 values in Table 1, the short-term memory effect ϕ in stochastic component of the daily demand is not significant as compared with the relatively stronger effects ψ_1 and ψ_1 in the hourly demand. This implies that daily demand is dominated by deterministic seasonal and weekly patterns rather than the stochastic components that would subject to random fluctuations. This is reflected by the fluctuation patterns shown in Figures 4 and 5 in which the generated hourly patterns fluctuate more than the generated daily patterns.

Model Parameters	Fresh Water Supply Zone	Flushing Water Supply Zone
Number of Harmonics	21	22
ϕ	0.037	0.246
$R^2 =$	0.001	0.061
σ_ε^d	9.37	27.63
ψ_1	0.520	0.476
ψ_2	0.390	0.121
c_o	-1.157	0.417
$R^2 =$	0.677	0.270
σ_ε^h	27.95	31.82

Table 1 - Model Parameters

Results of 2009 Demand	Fresh Water Supply Zone			Flushing Water Supply Zone		
	Generated	Actual	% Difference	Generated	Actual	% Difference
Peak Hourly Demand (L/s)	208.9	191.2	9.3%	517.0	496.3	4.2%
Peak Daily Demand (L/s)	134.2	121.9	10.1%	374.5	353.2	6.0%
Mean Daily Demand (L/s)	88.8	87.7	1.3%	282.4	282.2	0.1%
Peak Daily Factor	1.51	1.39	8.7%	1.33	1.25	6.0%
Peak Hourly Factor	2.35	2.18	7.9%	1.83	1.76	4.1%

Table 2 - Demand Comparison --- Generated versus Actual

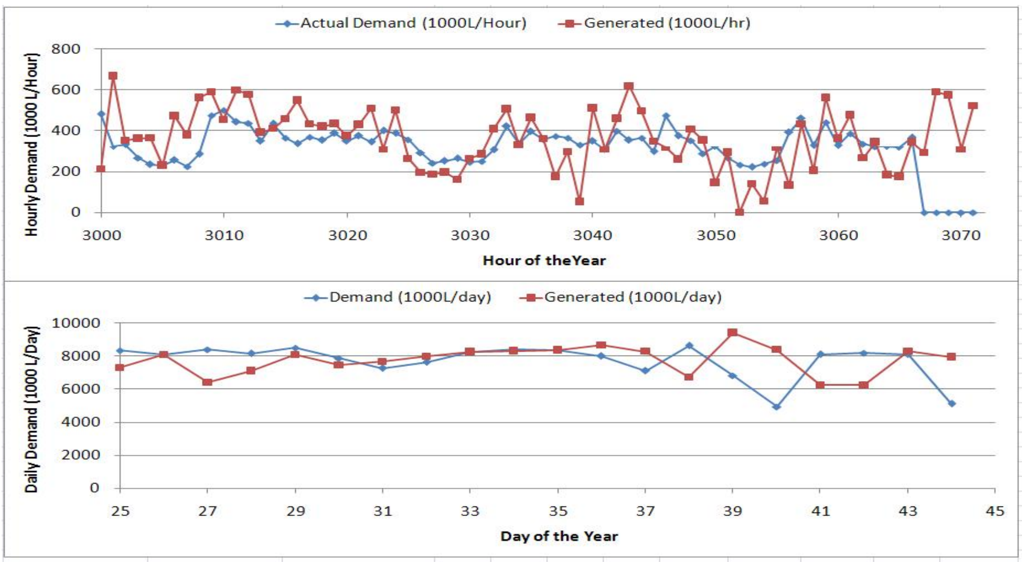

Figure 4 - Fresh Water Supply Zone Demand Patterns --- Actual versus Generated

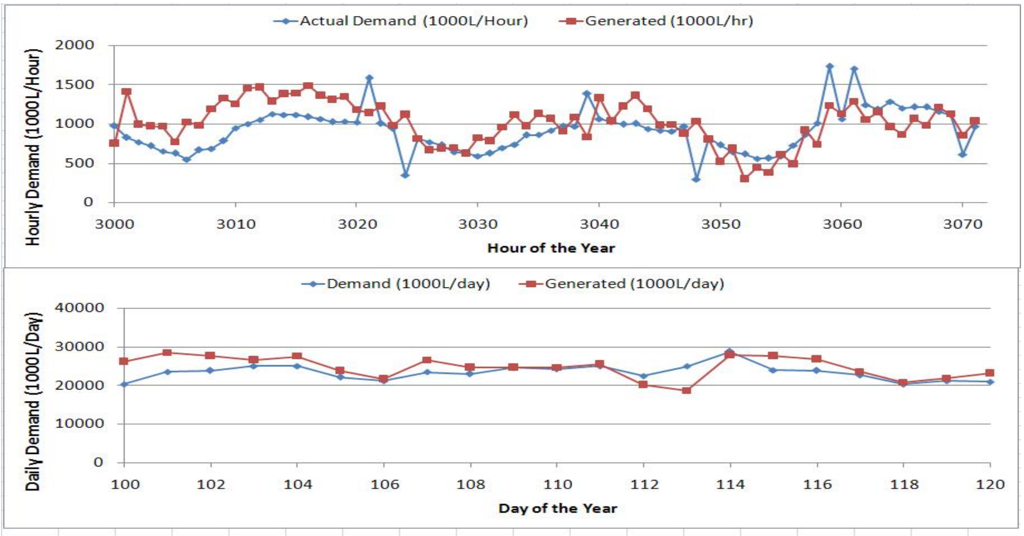

Figure 5 - Flushing Water Supply Zone Demand Patterns --- Actual versus Generated

3.2 Model Validation

The validation of the stochastic model has been carried out by means of generating daily and hourly demand continuously for a year. Several statistics and factors of the generated demands are computed and compared with the ones computed for actual demands of the year. The validation results of the fresh water model and flushing water model mentioned above are summarised in Table 2. As seen in Table 2, the generated mean demand and peak demand are slightly higher than the actual ones. This is likely attributed by possible over-estimation of the randomness by the model. To check the randomness of the deviation between the generated demand and the actual one, the frequency distribution of the deviations is plotted in Figure 6. As shown on the normal quartile plots, the deviations follow normal distributions. It can be concluded that the deviations from the actual demands as well as any over-estimation of the peak demand are mainly contributed by the random components of the stochastic model. Such relatively higher randomness means that the model would output conservative results as shown in Table 2.

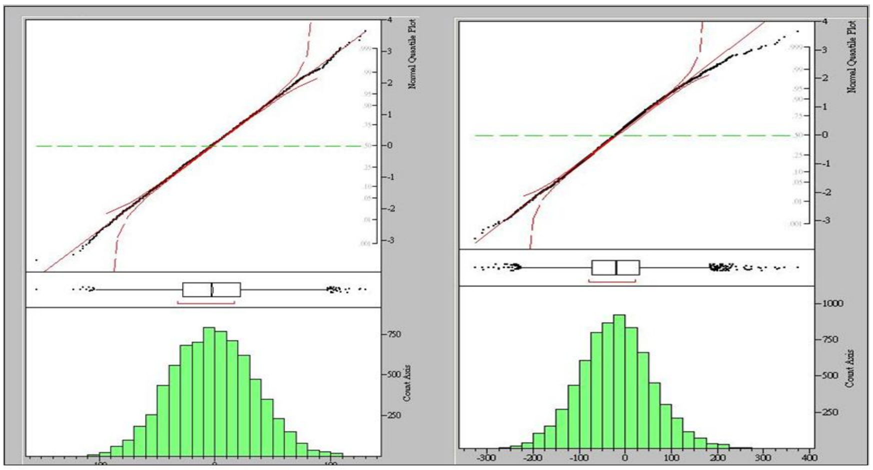

Figure 6 - Normal Quantile Plot and Frequency Distribution of the Generation Errors

4. APPLICATIONS

4.1 Reliability-based Planning for Upgrading Wan Chai Salt Water Supply System

Reliability has been regarded as a useful tool for water supply planning (Lansey *et al* 2004). A practical reliability-based planning approach is now made possible and developed with application of the stochastic water demand modelling. Driven by the need to overcome site constraints and minimise environmental impacts, the reliability-based planning approach has been tried out in the project for upgrading the Wan Chai salt water supply system in Hong Kong (Leung *et al* 2009).

Based on the traditional planning approach, expansion of the existing service reservoir, which is very difficult due to site constraints, is required under the project to meet increased demand. To enable evaluation of alternative system configurations, stochastic water demand model is formulated based on historical water demand data of the existing Wan Chai salt water supply system and incorporated into the system hydraulic model with probabilistic pipe failure model for Monte Carlo simulation and evaluation of system performance. It is found that a simulation of 1000 years could give reliable converging results of system performance and operation parameters.

Results of Monte Carlo simulations enable quantification and evaluation of the reliability level of the upgraded system with standard configurations based on the traditional planning approach first. Under the original system configuration, the probability of system failure is 18.4 hours of system outage per annum at 98% confidence level. Such level of system is regarded as the design reliability benchmark for assessing the acceptability of the alternative system configuration where no additional storage is required based on the stochastic approach. The alternative configuration results in reduction of the probability of system failure to 15.2 hours of system outage per annum at 98% confidence level, thereby achieved no worse than or even better level of system reliability than the original system configuration.

A life-cycle cost assessment reveals that the present value of the life-cycle cost of the alternative configuration is less than that of original configuration by 20 million Hong Kong dollars. Since the alternative configuration is found to be more economical while the reliability level is not compromised, this configuration has been adopted in the upgrading project.

5. CONCLUSION

A stochastic water demand model for synthetic generation of water demands has been formulated and applied in modelling water demand generated in both fresh water and flushing water supply zones and planning for the Wan Chai salt water supply system upgrading project. The model is based on statistical analysis of historical water demand data using a short-term pattern-based forecasting approach (Alvisi *et al* 2007). The model comprises a deterministic component and a stochastic component. The deterministic components are modelled 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 modelling results indicate that the synthetic patterns, fluctuations and statistics of the generated demands are consistent with the actual ones.

Application of the model for Monte Carlo simulation of system performance has enabled planning for the Wan Chai salt water supply system upgrading project with a new reliability-based planning approach that has minimised impacts on the adjacent environment and optimised the system configuration. It is a useful tool for assessment of optimal system reliability in quantified terms and optimisation of system configuration and operation (cost and energy saving) without compromising operation requirement and system reliability.

References

Alvisi, S., Franchini, M. and Marineli, A. (2007). A short-term, pattern-based model for water demand forecasting. *Journal of Hydroinformatics* 09.1 pp. 39-50.

Kottegoda, N. T. (1980). *Stochastic Water Resources Technology*. pp. 34-46, 98-104, 111-172. Macmillan, London.

Lansey K., Mays L. W. and Tung Y. K. (2004). Reliability and Availabuility Analysis of Water Distribution Systems. *Urban Water Supply Management Tools*. pp. 6.1-6.31. McGraw-Hill, New York.

Leung, W. L. and Lam, S. T. (2009). Planning of a flushing water supply system in Hong Kong with a stochastic analysis method. *Proceedings of the 8th Shenzhen-Hong Kong-Zhuhai-Macau Water Supply Conference cum the 5th Across-Strait Conference on Safe Water Quality Control Technology and Management 28-29 October 2009.*

Water Supplies Department (1985). Departmental Instruction No. 1309 - Design Criteria (Revised July 1985). *Departmental Instructions*, Water Supplies Department, HKSAR Government, China.