

## 단기 물 수요예측 시뮬레이터 개발과 예측 알고리즘 성능평가

### Development of Water Demand Forecasting Simulator and Performance Evaluation

신강욱 · 김주환\* · 양재린 · 홍성택

Gang Wook Shin · Ju-Hwan Kim\* · Jae-Rheen Yang · Sung-Taek Hong

한국수자원공사 K-water연구원

(2011년 3월 21일 접수 ; 2011년 7월 9일 1차수정 ; 2011년 8월 8일 2차수정 ; 2011년 8월 12일 채택)

#### Abstract

Generally, treated water or raw water is transported into storage reservoirs which are receiving facilities of local governments from multi-regional water supply systems. A water supply control and operation center is operated not only to manage the water facilities more economically and efficiently but also to mitigate the shortage of water resources due to the increase in water consumption.

To achieve the goal, important information such as the flow-rate in the systems, water levels of storage reservoirs or tanks, and pump-operation schedule should be considered based on the reasonable water demand forecasting. However, it is difficult to acquire the pattern of water demand used in local government, since the operating information is not shared between multi-regional and local water systems. The pattern of water demand is irregular and unpredictable. Also, additional changes such as an abrupt accident and frequent changes of electric power rates could occur. Consequently, it is not easy to forecast accurate water demands. Therefore, it is necessary to introduce a short-term water demands forecasting and to develop an application of the forecasting models.

In this study, the forecasting simulator for water demand is developed based on mathematical and neural network methods as linear and non-linear models to implement the optimal water demands forecasting. It is shown that MLP(Multi-Layered Perceptron) and ANFIS(Adaptive Neuro-Fuzzy Inference System) can be applied to obtain better forecasting results in multi-regional water supply systems with a large scale and local water supply systems with small or medium scale than conventional methods, respectively.

**Key words** : water operation, demand pattern, forecasting model, water demand

**주제어** : 수운영, 수요패턴, 예측모형, 물수요

\* Corresponding author Tel:+82-42-870-7501, E-mail: juhwan@kwater.or.kr(Kim, J.H.)

## 1. 서 론

우리나라 용수사용량의 증가로 인한 수자원 부족 현상을 해소하고, 수도시설의 효율적 운영관리를 위하여 수운영시스템의 구축이 활발하게 이루어지고 있다. 수도시설의 수량관리를 통한 에너지의 효율적 사용을 위하여 구축된 수운영시스템은 각 급수지역별로 단기 수요량을 예측하여 이를 근거로 유량, 수위 및 펌프운전 계획을 수립함으로써 취송수 제어 및 배급수가 이루어지도록 구성된다. 또한 수요예측을 통해 정수장에서는 적정 생산량을 사전에 예측함으로써 정수장의 운영자동화와 효율성 향상을 도모할 수 있다. 그러나 지자체 수용가 물사용량 패턴의 불규칙성, 지자체 사업장 운영 데이터의 미공유, 사고시 연계운영에 따른 사업장별 운영패턴의 변화 및 잦은 전력요금체계 변경 등의 환경변화로 정확한 수요예측이 어려워 사업장별 최적의 경제적 운영에 한계성이 나타나고 있다. 또한 최적의 물수요 예측량으로 적정 정수생산량을 정확하게 결정함으로써 정수장 최적운영이 가능하게 되나 위와 같은 문제점으로 여전히 단기 물수요 예측 기술이 활용되지 못하므로 국내 특성에 맞는 단기 물수요 예측 기술과 활용방안의 개발이 시급하다.

수운영시스템의 목적은 첫째, 급수구역에 충분한 수량과 적절한 수압을 확보하는 것, 둘째, 사고나 재해의 긴급대응 즉 긴급시에 물 배분의 확보를 위한 운전관리, 셋째, 자동화를 통한 원격제어의 집중화에 의한 전력요금 절감 등 비용저감을 도모하는 것이다. 이 중에서도 자동화 기술이 발전하면서 경영측면에서 볼 때 용수생산단가 절감에 의한 비용절감이 중요사항이다. (Choi, 2009)

본 연구에서는 대규모 물 공급시설인 광역상수도과 중소규모 공급시설인 지자체 물 수요예측에 필요한 수요예측모델을 검토하여 그 결과를 제시하고자 하였다. 이는 광역상수도과 지방상수도의 물사용 규모와 시설운영 특성이 각각 상이한 바 적용할 수 있는 수요예측모델이 다른 성능을 가지므로 모델이 갖는 특성을 평가함으로써 선택적으로 활용할 수 있는 물 수요예측 시뮬레이터를 개발하였다. 이를 통하여 공급 대상지역의 사용패턴 및 운영시설 특성에 적합한 1일 물 수요량 예측과 시간별 수요량 예측을 위한 최적 모델을 도출할 수 있도록 하였다.

## 2. 수요예측 이론 고찰

물 수요예측은 과거 시계열자료가 현존하는 경우 자료의 분석을 통해 시간적 흐름에 따른 사용량의 경향성에 근거하여 예측하는 방법을 토대로 정량화 시킬 수 있다. 정량적 예측기법 중에는 인과모형을 근간으로 한 예측기법을 이용

한 방법들은 물수요량을 직접 측정하지 않고 용수사용량에 영향을 미치는 다른 인자로서 사회·경제적 요인별, 용도별 원단위, 또는 물 사용형태별 예측방법이나 목적에 따라 설명력이 높은 인자를 설명변수 또는 독립변수로 선별적으로 채택하여 간접적으로 수요량을 산정한다. 특히, 수도시설의 운영을 위해서는 단기 수요예측 방법이 필요한데, 대표적으로 사용되는 선형모델로서 회귀분석 및 시계열분석에 근간을 둔 MR(Multi Regression), ARMA(Auto Regressive Moving Average) 모델과 제어이론으로부터 도입된 Kalman 필터와 인공지능이론을 도입한 MLP(Multi Layer Perceptrons), ELM(Extreme Learning Machine), ANFIS(Adaptive Neuro-Fuzzy Inference Systems) 등이 적용되고 있다.

따라서 본 연구에서는 물 수요예측을 위해 해당 공급지역의 물 사용규모와 시설운영 특성을 고려하여 위에서 열거된 6가지 예측이론을 도입한 시뮬레이터를 개발하여 모의결과에 따른 성능을 평가할 수 있도록 함으로써 물 사용 특성에 적합한 모델을 선택적으로 활용할 수 있도록 하고자 하였다.

일반 선형모형(General linear model)은 추정하고자 하는 모수에 관하여 1차식으로 표시된 모형을 의미한다.  $p$ 개의 독립변수  $x_1, \dots, x_p$ 에 대응하는 종속변수  $y$ 의 선형모형은 다음과 같이 표현된다. (Box, 1970)

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \epsilon \quad (1)$$

시계열모형은 미래의 예측 값이 예측대상이 되는 변수의 과거 자료와 과거 오차에만 의존된다는 가정을 설정하고 과거 관측치의 패턴을 규명하여 이것을 미래로 연장함으로써 미래 값을 예측하는 것으로 영향인자들 대신 입출력 관계를 시간의 함수로 표시한 것이다. 대표적으로 사용되는 방법으로 Box-Jenkins에 의해 제안된 것이 ARMA모델로서 일정변수를 갖는 AR(1)모형의 일반식은  $p$ 차 AR에 대하여 다음 식으로 표시할 수 있다.

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_p X_{t-p} + \epsilon_t \quad (2)$$

여기서,  $X_t, X_{t-1}, \dots, X_{t-p}$ 는 자기회귀과정을 표시하며  $\beta_1, \beta_2, \dots, \beta_{t-p}$ 는 자기회귀계수이고,  $\epsilon_t$ 는 평균이 0이고 분산이  $\sigma^2$ 인 독립시계열로 무작위성분(Random component)을 나타낸다. 차수가 1인 AR(1)모형의 일반식은 다음과 같다.

$$X_t = \beta_1 X_{t-1} + \epsilon_t \quad (3)$$

여기서,  $\epsilon_t$ 는  $\epsilon_t = \sigma z_t$ 이며  $z_t$ 는 평균치가 0이고 분산이 1인 정규분포를 갖는 난수(Normally distributed random number)이다.

MLP는 신경망을 나타내는 모형으로 가장 폭 넓게 사용되어온 것이 Rumelhart등(1986)에 의해 소개된 것이며

이는 특별한 구조나 매개변수의 산정 및 자료의 변화 등이 필요치 않고 자료의 축적에 따라 모형의 능력을 향상시킬 수 있는 장점을 가지고 있다. (Kosko, 1992) 물 수요예측을 위해서는 일련의 시간변화에 따른 수요패턴을 기준이 되는 시점으로부터 향후 일정시간 이후 시간의 수요량을 추정하는 것으로 기준시점 이전의 수요량 시계열이 입력되어 예측하고자 하는 시간에서의 사용량을 출력하는 구조를 갖도록 하였으며 이는 계산 수요량  $o$ 와 실측 수요량  $y$ 의 제곱오차합이 설정된 오차범위까지 최소화 하도록 반복적으로 모형 매개변수를 조정함으로써 학습이 이루어진다.

$$E_p = \frac{1}{2} \sum_k (y_{pk} - o_{pk})^2 \quad (4)$$

신경망모형의 하나인 ELM에서는 가중치들은 반복적인 학습구조 대신 일반화된 역행렬에 의해 계산되는 구조를 갖는 것으로 전형적인 반복 학습 방법들에 비하여 매우 빠른 학습이 가능하다. 또한 퍼지이론을 토대로 Jang에 의해 제안된 ANFIS는 퍼지규칙의 수를 최소로 하면서 복잡한 비선형 시스템의 표현에 뛰어난 능력을 보이는 방법으로 Bezdek에 의해 제안된 FCM(Fuzzy Cluster Method) 알고리즘에 의해 목적함수인 관측치와 모의치간의 오차를 최소화하도록 반복함으로써 주어진 자료로부터 유사한 클러스터를 분류하여 생성된 클러스터에 의해 주어진 자료를 재현한다. (Jang 1993)

FCM에 의해 결정된 각각의 클러스터는 시스템의 특성을 기술하는 퍼지 규칙으로 사용되며 퍼지모델의 추론값  $f^*$ 는 다음 식과 같이 가중평균법(weight average method)에 의해 다음과 같이 계산된다.

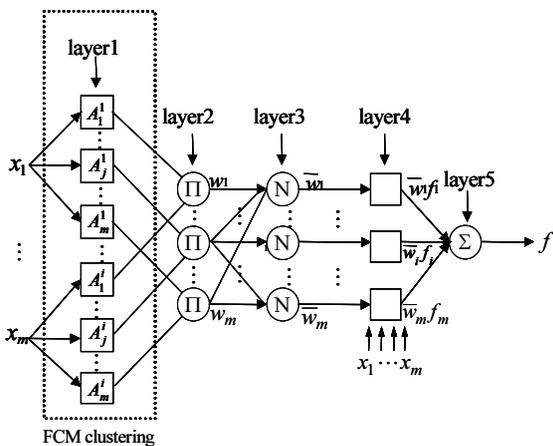


그림 1 ANFIS 구조

$$f^* = \frac{\sum_{i=1}^m w^i f^i}{\sum_{i=1}^m w^i} \quad (5)$$

Kalman 필터에 의한 예측이론은 다음과 같이 상태방정식에 의해 표현되며, 여기서,  $x(k)$ 는  $n \times 1$  시스템 상태벡터,  $A(k)$ 는  $n \times n$  시변상태전이행렬,  $z(k)$ 는  $m \times 1$  측정 벡터,  $C(k)$ 는  $m \times n$  시변 출력 행렬,  $w(k)$ 는  $n \times 1$  시스템 오차, 그리고  $v(k)$ 는  $m \times 1$  측정 오차를 나타낸다. (Kalman, 1960)

$$x(k+1) = A(k)x(k) + w(k) \quad (6)$$

$$z(k) = C(k)x(k) + v(k) \quad (7)$$

이는 관측치와 모의치간의 공분산을 최소화시킴으로써 모형의 매개변수를 추정하는 것으로 Kalman 이득방정식과 상태예측 방정식 및 오차 공분산 방정식은 각각 다음식과 같다.

$$K(k) = [A(k)P(k)C^T(k)] [C(k)P(k)C^T(k) + Q_1]^{-1} \quad (8)$$

$$\hat{x}(k+1) = A(k)\hat{x}(k) + K(k)[z(k) - C(k)\hat{x}(k)] \quad (9)$$

$$P(k+1) = [A(k) - K(k)C(k)]P(k) [A(k) - K(k)C(k)]^T + K(k)Q_2K^T(k) \quad (10)$$

여기서,  $P(k)$ 는 오차 공분산 행렬,  $\hat{x}(k)$ 는  $k$  시점의 상태 예측값,  $Q_1$ 은 측정오차의 공분산 행렬, 그리고  $Q_2$ 는 오차의 공분산 행렬을 각각 나타낸다.

### 3. 물수요 예측 시뮬레이터 개발

#### 3.1 시뮬레이터의 구성

물 수요량 예측을 위한 시뮬레이터의 개발은 타 정보시스템으로부터 실시간 자료의 취득을 위한 연계 프로그램, 자

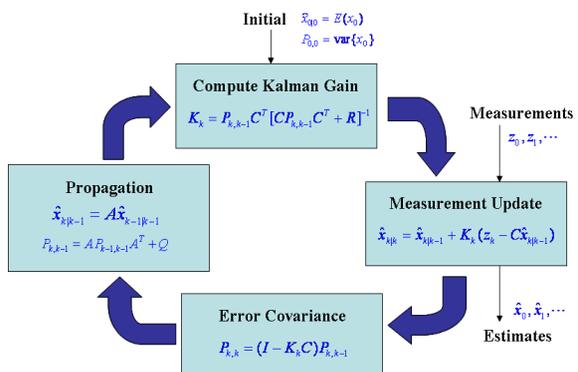


그림 2 Kalman 필터의 계산순서

료형식 변환 프로그램, 수요예측을 위한 알고리즘 프로그램, 그리고 모델 매개변수 및 예측결과의 평가 및 분석프로그램으로 이루어지도록 하였다. 수도시설의 운영을 위해 구축된 타 정보시스템과의 연계는 대상 급수지역의 수요량 자료의 취득을 목적으로 데이터베이스와의 연계는 물론, 기상자료 확보를 위한 실시간 물관리시스템과 연계되도록 구성하였다. 자료의 취득주기는 매 1시간 단위로 설정하여 특정기간 동안 타 정보시스템의 자료와 이력자료의 확인이 가능하도록 하였으며 사용량 및 기상자료의 시간별 자료와 일자료가 하나의 파일로 구성되어 매일 1시부터 24시까지의 시간자료와 일예측을 위해 해당 일의 일적산 유량과 일평균 기상자료를 순차적으로 저장하여 사용한다. 또한 시뮬레이터에서 6가지 모델에 대한 모의가 가능하도록 사용자가 입력변수를 선정할 수 있도록 하였고 입력변수는 물 사용량에 대한 과거이력과 기상자료를 선택하여 다양한 조건에서 모의가 가능하도록 하였다.

각 예측모델에 대한 모의결과는 그래프 형식과 텍스트 형식으로 표현되도록 구성되었다. 시뮬레이터 예측실행 후 결과는 분석하고자 하는 선택구간 선정과 추세 분석을 위하여 그래프 형식으로 표현되고, 그래프 형식에서 선택된 구간에 대하여 각 모델의 일별 예측 오차값과 시간별 예측 오차값을 제시하도록 구성하였다. 그래프 형식에서는 모의된 예측 모델 중 특정 예측모델만 표시하거나 몇 개의 예측모델을 중첩하여 분석할 수 있도록 구성하였다. 이를 통하여 임의의 기간에 대한 운영 데이터의 최적 알고리즘을 선정할 수 있으며, 적용코자 하는 각 사업장에 대한 예측 알고리즘에 대한 검증이 가능하다. 모의결과는 검증하기 위하여 다음의 절대평균오차(MAPE; Mean Absolute Percent Error)를 사용하였으며 이는 다음과 같이 표현된다.

$$MAPE(\%) = \frac{1}{N} \sum_{k=1}^N \frac{|y_k - \hat{y}_k|}{y_k} \times 100 \quad (11)$$

여기서,  $y_k$ 는 실제 물 공급량,  $\hat{y}_k$ 는 예측 물 공급량이다.

그림 3에서는 모의기간 등 초기조건을 설정하고 각 예측이론별 알고리즘을 선정실시간 데이터베이스와 연계한 실시간 모의, 모의결과에 대하여 MAPE에 의한 성능평가 등 시뮬레이터의 실시간 물 수요예측 절차를 나타낸 것이다.

### 3.2 시뮬레이터의 적용

본 시뮬레이터에서 적용된 예측모델은 앞서 언급한 AR, MR, KF, MLP, ELM, ANFIS 모델로서 일별 단기 물 수요예측 알고리즘에 대하여 6가지 모델을 구분하여 먼저 일별 수요예측량이 도출되면 이를 이용하여 예측일의 24시간에 대하여 시간별 예측을 수행토록 하였다. 시간별 예측은 고정 분배방식과 시간별 오차 가중 분배방식을 적용할 수 있

도록 하였고 각 알고리즘의 매개변수 선정 및 이를 적용한 각 알고리즘별 모의를 통해 자동적으로 예측결과에 대하여 정해진 매 시간마다 정확도를 확인할 수 있도록 함으로써 모델의 적정성을 평가할 수 있도록 하였다.

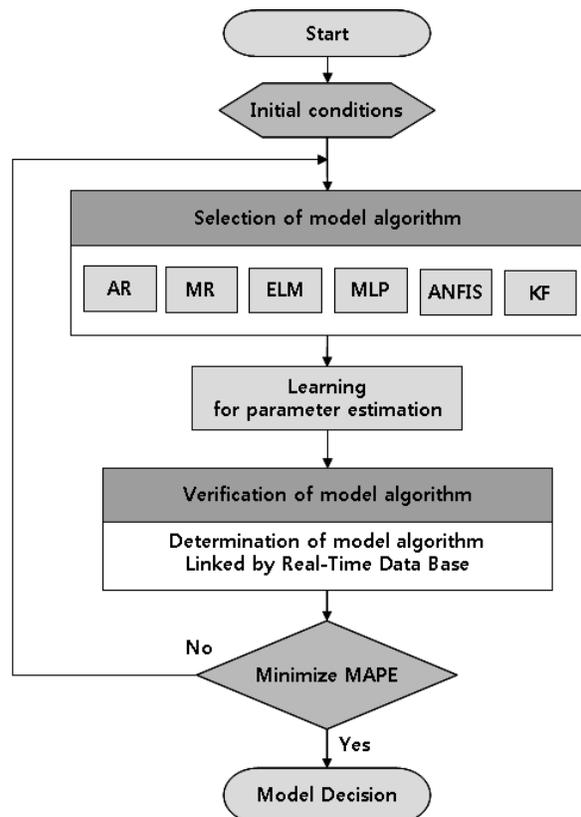


그림 3 수요예측 시뮬레이터 절차

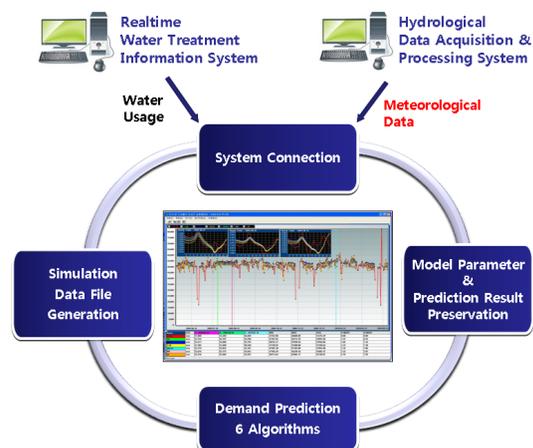
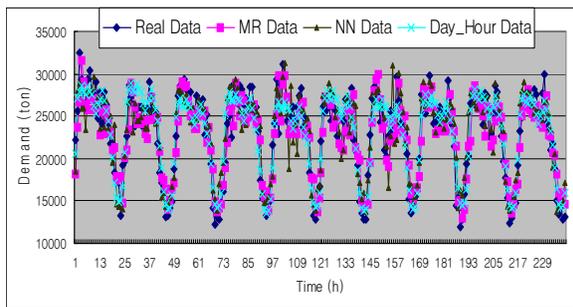


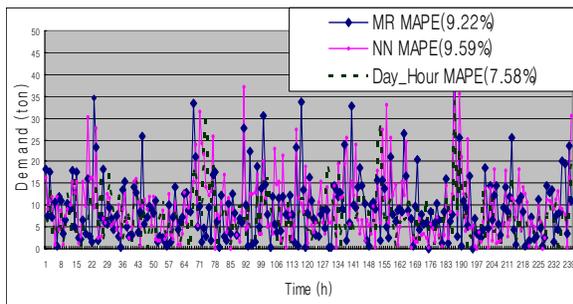
그림 4 수요예측 시뮬레이터의 자료흐름 및 구성도

특히, 시간별 물 사용량 예측을 위해서는 MR 모델과 MLP 모델을 적용하였다. 또한, 일별 예측량을 구한 후 시간 예측량을 도출하는 방법을 제시하여 상호 비교분석을 통하여 최적의 예측 알고리즘을 제안하였다. 먼저 다중선형회귀모델과 신경망모델을 적용한 결과 **그림 4**에서와 같이 각 모델의 예측 MAPE값이 9.22%, 9.59%로 나타나 다중선형회귀모델이 신경망모델보다 예측특성이 양호함을 알 수 있었다. 반면, 일별 예측 후 시간별 예측한 경우 MAPE값이 7.58%로 나타나 시간기준 예측모델인 다중선형회귀모델과 신경망모델의 결과보다 양호한 결과를 얻었다.

따라서, 위의 결과를 바탕으로 본 연구에서는 2% ~ 3%의 일 예측 결과를 바탕으로 시간별 예측은 시간별 고정 분배 방식과 시간별 오차 가중 분배방식을 제안하였다. 일별 예측결과에 따른 시간별 예측은 과거 데이터를 이용하여 24시간 가중치를 고려하여 예측하고자 한다. 이는 과거 60일 데이터를 이용하여 각 시간대별 평균을 구하고 이를 각 시간대별 가중치로 환산하여 해당 시간에 고정 분배하는 방식을 말한다. 이와 같은 고정 분배방식을 모의 실험한 결과, 전체 예측오차는 MAPE 기준 약 8.1%로 나타났으며, 이를 각 요일별 오차율로 분석하면 다음 **그림 7**과 같이 일요일 예측결과가 약 11%로 가장 큰 오차를 보인다. 반면 화요일의 경우 약 6.9%로 가장 양호한 예측결과를 얻었다.



(a) 물 사용량 예측결과(사용량)



(b) 물 사용량 예측결과(MAPE)

그림 5. 시간기준 물 수요예측

시간별 오차 가중분배 방식은 실제 1일 물 사용량이 식 (12)와 같고, 시간별 고정 분배방식에 의하여 예측된 결과 가 식 (13)과 같다고 가정하자.

$$y_1 = x_1 + x_2 + \dots + x_{23} + x_{24} \quad (12)$$

$$y'_1 = x'_1 + x'_2 + \dots + x'_{23} + x'_{24} \quad (13)$$

여기서, 첫 번째 시간예측인 1시 데이터의 경우 실제 사용량  $x_1$ 에 대하여 예측량  $x'_1$ 을 식 (4)와 같이 오차가 발생한다면, 고정 분배방식에서 구한 2시 이후 유량 예측값에 오차값이 포함된 식 (15)를 구할 수 있다.

$$e_1 = x_1 - x'_1 \quad (14)$$

$$y'_2 = x'_1 + x'_2 + \dots + x'_{23} + x'_{24} + e_1 \quad (15)$$

이와 같이 구한 2시 이후의 유량 예측값은 식 (16)에서와 같이 각 시간대별 가중치를 곱하여 예측값이 가변되도록 한다.

$$x'_i = \frac{x'_i}{\sum_{j=2}^{24} x'_j} r_1, \quad i=2,3,\dots,24. \quad (16)$$

여기서,  $r_1 = x'_2 + \dots + x'_{23} + x'_{24} + e_1$

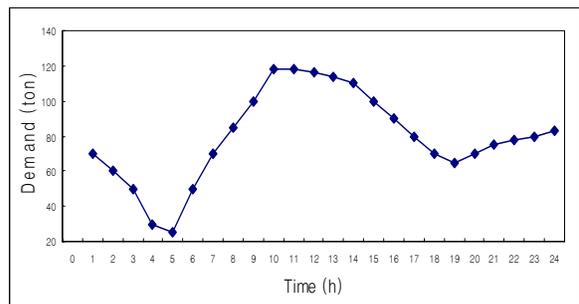


그림 6. 시간대별 물 공급 가중치

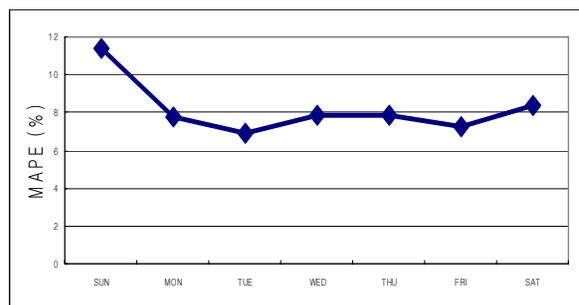


그림 7 고정 분배방식에 의한 요일별 오차율

아래 **그림 8**은 매 24시에 예측된 다음날의 시간대별 예측값이 실제 물 공급량과 오차가 발생되면, 오차가 발생한 시간부터 그날 24시까지의 예측값을 고정 분배방식에서 구한 가중치에 따라 가중치를 곱하여 예측값을 변경하도록 구성된 것이다.

시간별 오차 가중 분배방식을 적용하여 시간별 예측오차를 구한 결과 약 7.9%를 얻었으며, 요일별 오차율에 대한 결과는 **그림 9**와 같다. 일요일의 경우 10.7% 오차율을 나타내어 가장 낮은 예측율을 보였으며, 화요일과 금요일의 경우 약 6.8%로 매우 양호한 예측결과를 얻을 수 있었다.

#### 4. 적용결과 및 고찰

본 연구에서 개발된 단기 물 수요예측 시뮬레이터를 통하여 물 사용량이 많은 광역상수도 계통과 물 사용량이 상대적으로 적은 지방상수도 계통에 대한 시간별 예측을 실험하였다. 두 지역에 대한 물 사용량 특성과 예측결과에 대한 상대비교를 통하여 시뮬레이터의 유용성을 검증하고자 하였다.

#### 4.1 대규모 공급시설 적용 결과

대규모 공급시설인 광역상수도 A정수장의 일일 공급 가능한 급수는 약 100 만톤이며 급수 인구는 약 200만 명인 정수장이다. A정수장에서의 연간 물공급량을 분석하기 위하여 2006년부터 2008년까지 3년 동안의 물 공급데이터를 조사하였다. **그림 10**은 일일 공급량을 도시한 것으로 그림에서와 같이 여름철 물 공급량이 대체로 많고, 년도별 평균 공급량에서는 2006년도 44만톤, 2007년도 47만7천톤, 그리고 2008년도 약 50만톤의 물이 공급되어 년평균 약 3만톤 규모의 공급량이 증가된 것을 알 수 있다. 시간대별로는 새벽 0시부터 5시까지 물 공급량이 감소하는 것을 알 수 있다. 또한 낮 12시경 물 소비량이 가장 많은 것으로 나타나 일일 물 공급량의 최대치를 나타내었다.

3년간의 물공급량 자료에 대하여 요일별 일원분산분석과 특이일에 대한 일원분산분석을 실시한 결과 평일과 주말의 경우 평균값은 거의 변화가 없으나 설 연휴와 추석 연휴 등의 특이일의 경우 약 15 % 물 공급량이 감소되었음을 알 수 있었다.

2006년과 2007년도 기상자료와 공급량 자료의 상관관계

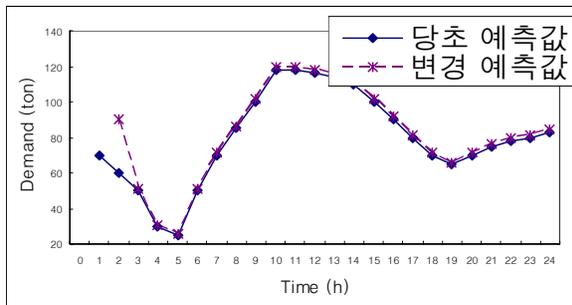


그림 8 시간별 오차 가중 분배방식

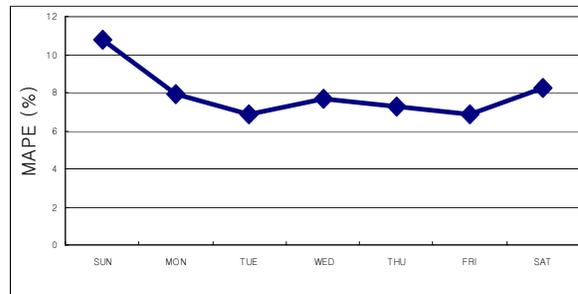
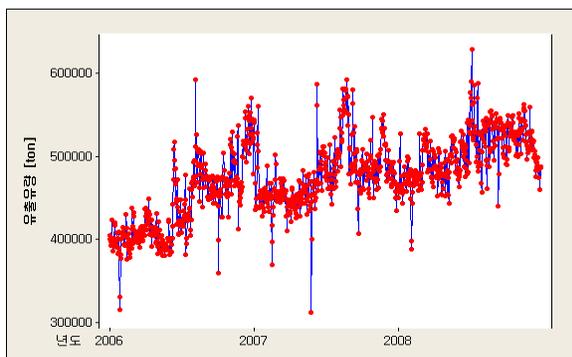
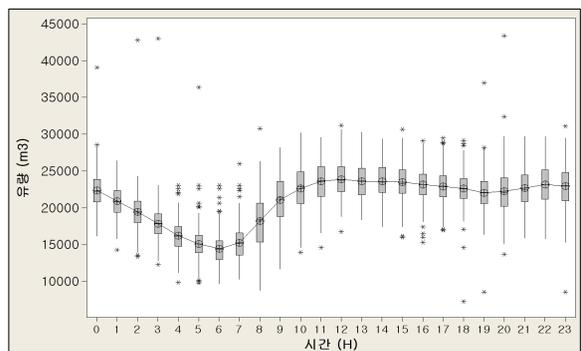


그림 9 오차가중 분배방식에 의한 요일별 오차율 산정 결과



(a)



(b)

그림 10 A정수장 일별(a) 시간별(b) 수요량 변화

표 1 기상자료 상관계수

기상 변수	최고기온	평균기온	최저기온	평균풍속	운량
A정수장	0.304	0.316	0.320	-0.087	0.019
B정수장	0.161	0.183	0.205	0.240	-
기상변수	상태습도	일조시간	강수량	적설량	전일유량
A정수장	0.124	-0.021	0.013	-0.009	0.742
B정수장	0.082	0.015	-	-	0.669

표 2 각 예측모델별 통계적 특성 비교

구분	평균	표준 평균오차	표준편차	최소값	Q1 (1/4분위값)	중앙값	Q3 (3/4분위값)	최대값
Oserved	21842	108.0	4992	8900	18206	22832	25700	32300
ANFIS	21842	95.4	4407	11970	18879	23610	25166	28752
ARMA	21837	95.5	4415	11946	18846	23551	25195	28755
ELM	21793	94.9	4384	12194	18969	23667	25168	28084
KF	21839	96.6	4466	9754	18941	23543	25192	29987
MLP	21791	95.3	4407	12040	18815	23580	25143	28275
MR	21837	95.5	4415	11942	18853	23555	25194	28559

는 최고기온, 평균기온, 최저기온, 평균풍속, 상태습도, 일조시간, 운량, 강수량, 적설량으로 분석한 결과는 표 1에서 볼 수 있으며, 이 중 기온의 상관성이 약 0.3으로 다른 기상 인자에 비해 약간 높게 나타난다.

대규모 공급시설인 광역상수도 A정수장에 대한 수요예측을 위해 개발된 시뮬레이터를 활용하여 일별 및 시간별 예측결과를 비교 분석하였다. 그림 11에서는 A정수장의 시간 단위 예측결과를 도시하였으며 예측률이 가장 양호한 MLP 모델의 경우 일 예측 MAPE가 2.39%, 시간별 예측 MAPE는 9.12%를 얻었으며, 다음으로 MR모델로서 일 예측 MAPE값이 2.59%와 시간 예측 MAPE값이 9.20%를 각각 얻었다. 반면 ELM 모델의 경우에는 일 예측의 경우 MAPE

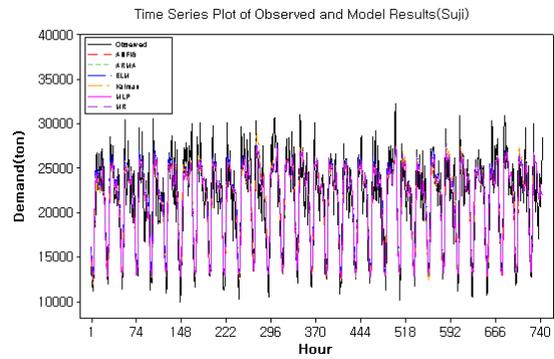
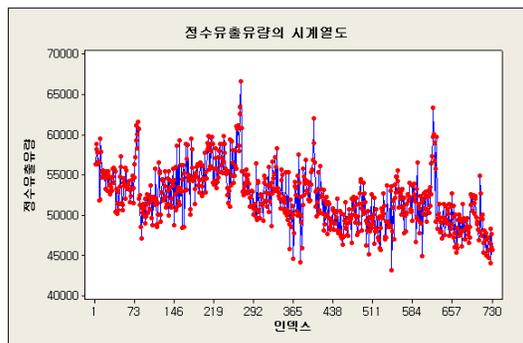
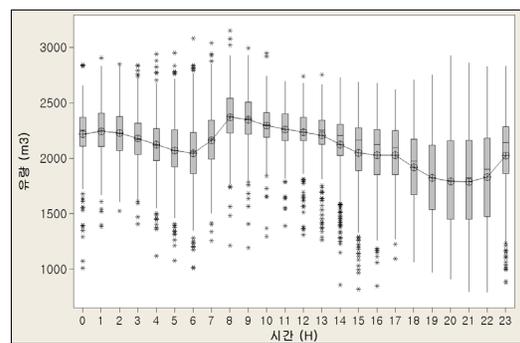


그림 11. A정수장 수요예측모델 비교



(a)



(b)

그림 12 B정수장 일별(a) 시간별(b) 수요량 변화

표 3 각 예측모델별 통계적 특성 비교

구분	평균	표준 평균오차	표준편차	최소값	Q1 (1/4분위값)	중앙값	Q3 (3/4분위값)	최대값
Oserved	2207.7	8.81	236.5	1508.0	2104.0	2256.0	2368.0	2608.0
ANFIS	2203.1	5.41	145.1	1732.4	2098.5	2242.5	2293.0	2671.3
ARMA	2197.8	5.33	142.9	1735.1	2095.4	2233.5	2284.4	2671.2
ELM	2185.8	5.21	139.8	1729.0	2088.6	2216.8	2266.7	2671.0
KF	2215.8	5.75	154.4	1749.2	2103.7	2250.2	2323.3	2674.5
MLP	2158.7	4.98	133.8	1713.1	2072.5	2176.2	2236.1	2667.7
MR	2199.0	5.34	143.3	1735.4	2096.0	2234.9	2286.5	2671.3

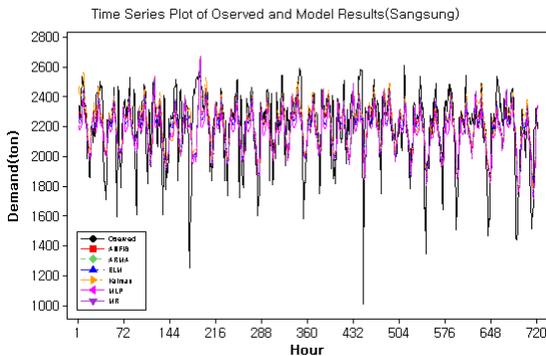


그림 13 B정수장 수요예측모델 비교

값이 4.72%, 시간별 예측한 결과의 MAPE는 9.45%로 나타나 광역상수도과 같은 대규모 공급시설의 일별, 시간별 수요예측을 위해서는 MLP와 MR모델이 가장 적합한 것으로 판명되었다.

#### 4.2 중소규모 시설 적용결과

중소규모 공급시설인 지방상수도 B정수장의 시설용량은 일일 약 9만톤이며, 급수지역은 중소도시로서 급수인구는 약 20만 명이다. 2007년과 2008년 2년간 자료를 분석한 결과, 물 공급량이 줄어들고 있는데 이는 유수율 증대를 위해 수행된 관망정비사업의 개선효과로 볼 수 있으며 시간별 물 수요량 분석결과에서는 새벽시간보다 저녁시간대에 더 적게 사용하는 특성을 보여주고 있다(그림 12참조).

요일별 일원분산분석과 특이일에 대한 일원분산분석 결과에서는 평일과 주말의 경우 평균값은 거의 변화가 없으나 특이일 전과 특이일의 경우 약 15% 물 공급량이 증가된 것으로 나타났다. 표준편차는 평일과 주말보다 특이일의 경우 비교적 많은 차이를 보이고 있음을 알 수 있었다. 또한 표 1에서와 같이 2006년과 2007년도 기상자료와 수요량 간 상관관계 분석결과에서는 기온의 상관성이 약 0.2, 그리

고 풍속의 경우 약 0.24 로서 기상인자에 대한 상관성이 매우 낮은 것으로 나타났다. 중소규모 공급시설인 지방상수도 B정수장의 시간단위별 수요예측 결과를 그림 13에 도시하였으며 각 모델별로 시간단위 수요예측결과를 비교하기 위한 통계적 특성치를 표 3에 제시하였다.

여기에서는 9개월 운영자료를 대상으로 적용하였으며, 그 결과, 일 예측율이 가장 양호한 특성을 나타낸 모델은 ANFIS 모델로서 MAPE값이 3.16%로 가장 양호한 결과를 보였으며 다음으로 MR모델로서 MAPE값이 3.31%로 나타났다. 시간별 예측의 경우 일 예측이 우수한 ANFIS 모델보다 양호한 6.91%의 결과를 얻은 반면 KF 모델의 경우에는 MAPE가 가장 크게 나타났다. 또한 일별 예측결과에서는 KF 모델의 MAPE값이 5.19%, 시간별 예측 MAPE의 경우 7.59%로 산정되었다. 따라서, 중소규모 수도시설로부터 공급되는 수요량의 경우 표 4에서와 같이 일별 예측 및 시간 예측 결과가 가장 양호한 ANFIS모델과 AR모델을 선정하여 예측하는 것이 가장 바람직한 것으로 판단되었다. 그리고, 대규모 수도시설과 중소규모 수도시설의 시간별 수요 패턴이 상이한 것은 물 수요지역의 특성에 따라 시간별 수요패턴을 달리함을 알 수 있다. 즉 대규모 수도시설을 이용하는 수요지역의 경우 공업시설이 널리 분포되어 있어 주거시설이 주로 분포되어 있는 중소규모 수도시설 공급지역과 상이함을 알 수 있었다.

### 5. 결 론

우리나라 용수 사용량의 증가로 인한 수자원 부족 현상을 해소하고, 상수도시설의 운영관리 효율화를 위하여 수운영시스템이 활발하게 구축되고 있는 실정으로 수운영시스템의 기본이 되는 수요예측을 위해 지금까지 적용된 예측이론을 도입하여 적용한 결과 다음과 같은 결론을 얻을 수 있었다.

(1) 본 연구로부터 얻은 수요예측 결과는 수운영시스템에서 관망해석, 시설의 운영률 도출 및 펌프 스케줄링 등에

표 4 예측모델별 MAPE(%)에 의한 성능비교

모델	광역상수도		지방상수도	
	일예측결과	시간예측결과	일예측결과	시간예측결과
ANFIS	2.59	9.20	3.16	7.06
ARMA	2.46	9.15	3.31	6.91
ELM	4.72	9.45	4.33	7.38
KF	4.21	9.34	5.19	7.59
MLP	2.39	9.12	4.84	7.36
MR	2.43	9.15	3.33	7.07

기본적으로 활용될 수 있는 정보로서, 실시간 적용이 가능한 수요예측 시뮬레이터를 개발하였다.

(2) 시뮬레이터는 공급시설 규모별 특성에 적합한 수요 예측모델의 선정이 가능토록 6가지 모형을 탑재하였으며 적용결과를 비교, 분석하여 사용자가 선택적으로 활용할 수 있도록 구축하였다.

(3) 대규모 공급시설과 중소규모 공급시설에 적용한 결과 물수요 패턴이 다른 만큼 각기 다른 특성을 보이고 있는 바, 대규모 공급시설인 광역상수도의 수요예측을 위해서는 MLP 모델이 일 예측값과 시간예측값이 2.39%와 9.12%로서 가장 양호한 결과를 얻었고, 중소규모 공급시설인 지방상수도에 적합한 모델로는 ANFIS 모델이 3.16%과 6.91%를 얻음으로써 각 지역별로 다양한 모델이 제시될 수 있음을 알 수 있었다.

(4) 향후 본 연구 결과를 바탕으로 수운영시스템에 적용될 수 있는 최적의 수요예측 모델을 도출함은 물론, 예측율이 높은 시뮬레이터를 제작하기 위하여 현장자료의 신뢰성을 높이고 국내 실정에 적합한 수요예측이 이루어질 수 있도록 수운영시스템의 완성도를 높여나가야 할 것이다.

### 참고문헌

최기선, 신강욱, 임상희, 전명근 (2009) 테이타 마이닝과 칼만필터링에 기반한 단기 물 수요예측 알고리즘, *제어자동화시스템 공학회*, **15(10)**, pp. 1156-1161.

S. Chenthur Pandian, K. Duraiswamy, C. Chistober Asir Rajan, and N. Kanagaraj, (2006) Fuzzy approach for short term load forecasting, *Electric power systems research*, **Vol. 76**, pp. 541-548.

H. S. Hwang, (2006) Fuzzy GMDH-type neural network model and its application to forecasting of mobile communication, *Computer & Industrial Engineering*, **50**,

pp. 450-457.

J. L. Torres, A. Garcia, M. De Blas, and A. De Francisco, (2005) Forecast of hourly average wind speed with ARMA models in Navarre(Spain), *Solar Energy*, **79**, pp. 65-77.

G.B. Huang, Q.Y. Zhu, C.K. Siew, (2006) Extreme learning machine: Theory and applications, *Neurocomputing*, **70**, pp. 489-501.

R.E. Kalman (1960), A New Approach to Linear Filtering and Prediction Problems, *Transaction of the ASME-Journal of Basic Engineering*, **82**, pp. 35-45.

J.S. Roger Jang, (1993), ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Trans. on Systems, Man, and Cybernetics*, **Vol. 23**, pp. 665-685.

Box,G.E.P and G.M. Jenkins (1970), **Time Series Analysis, Forecasting and Control**, Holdenday, San Francisco.

Andres S. Weigend, et al. (1990), "Back-Propagation, Weight-Elimination and Time Series Prediction," *SMC, Connectionist Models Proceedings of the 1990 summer school*, pp.105-116.

Barto Kosko (1992), *Neural Networks and Fuzzy Systems*, Prentice Hall.

Ramesh Sharda, Rajendra B. Patil (1990), Neural Networks As Forecasting Experts: An Empirical Test, *IJCNN*, June, **Vol. 2**, pp.II-491-II-494.

12. Benito Fernandez, et al. (1990), Nonlinear Dynamics System Identification using Artificial Neural Networks, *IJCNN*, June, **Vol. 2**, pp. II-133-II-142.