

일급수량 예측을 위한 인공지능모형 구축

Implementation of Daily Water Supply Prediction System by Artificial Intelligence Models

연인성¹ · 전계원^{2,*} · 윤석환¹

Yeon, In-sung¹ · Jun, Kye-won^{2,*} · Yun, Seok-whan¹

1 충북대학교 토목공학과

2 삼척대학교 방재기술전문대학원

(2005년 1월 19일 논문접수; 2005년 8월 11일 최종 수정논문 채택)

Abstract

It is very important to forecast water supply for reasonable operation and management of water utilities. In this paper, water supply forecasting models using artificial intelligence are developed. Artificial intelligence models shows better results by using Temperature(t), water supply discharge (t-1) and water supply discharge (t-2), which are expressed by neural network(LMNNWS; Levenberg-Marquardt Neural Network for Water Supply, MDNNWS; Modular Neural Network for Water Supply) and neuro fuzzy(ANFISWS; Adaptive Neuro-Fuzzy Inference Systems for Water Supply). ANFISWS model which is applied for water supply forecasting shows stable application to the variable water supply data. As results, MDNNWS model shows the highest overall accuracy among proposed water supply forecasting models and the lowest estimation error with the order of ANFISWS, LMNNWS model.

Key words: water supply, forecast, artificial intelligence, neural network, neuro-fuzzy

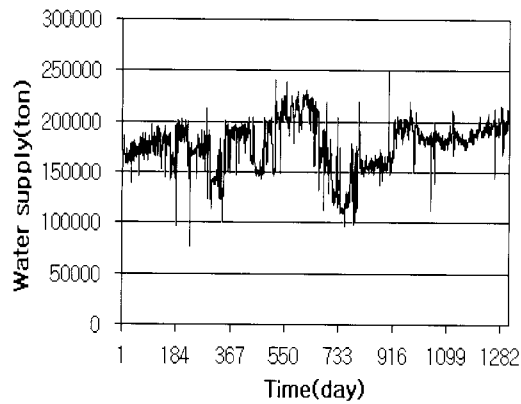
주제어: 급수량, 예측, 인공지능, 신경망, 뉴로-퍼지

1. 서 론

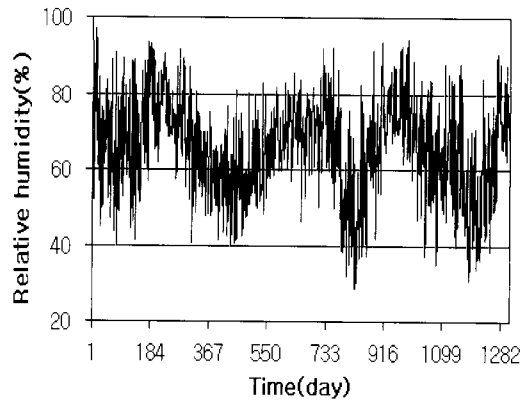
산업구조의 변화와 경제성장은 인구의 도시집중과 도시의 비대화를 초래하고 있다. 이로 인해 수원의 수질악화와 상수도의 부족 현상 등이 야기되었으며, 상수도 급·배수시설의 효율적인 관리와 운용에 많은

관심이 대두되고 있다. 상수도의 합리적인 운용과 관리를 위해서는 급수량 예측이 매우 중요하다. 급수량의 예측방법에는 예측시점에 따라 장기예측과 단기예측으로 나누어진다. 장기예측은 월, 연별 예측으로 장래 상수도시설의 신설 또는 증설계획 수립시 필요하고, 단기예측은 일 또는 시간단위의 수요예측을 의미하며 수도시설의 합리적인 운용 또는 급·배수량의

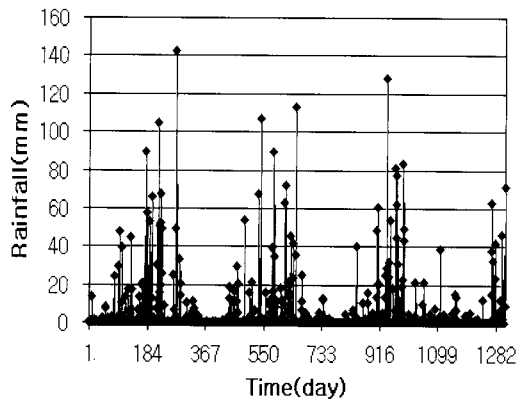
*Corresponding author Tel: +82-33-570-6816, FAX: 82-33-570-6819, E-mail: kwjun@samcheok.ac.kr (JUN, K.W.)



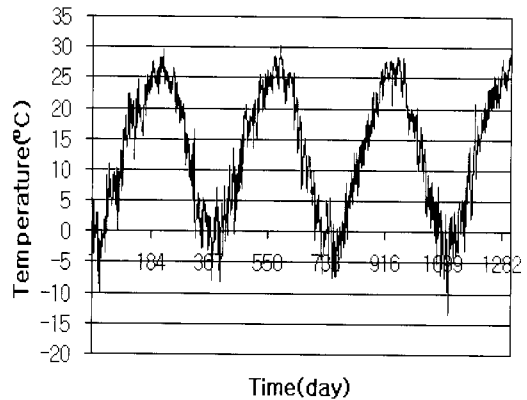
a) 급수량 자료



b) 상대습도



c) 강수량



d) 평균기온

Fig. 1. 대상지역의 자료.

조절에 이용되고 있다(구자용 외, 1997; 현인환 외, 1995)

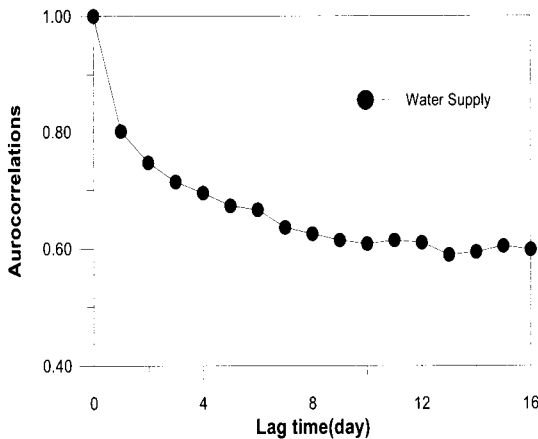
과거 급수량의 수요예측방법에는 크게 다변량 해석방법과 시계열 분석 방법이 있다. 다변량 해석방법의 대표적인 방법으로는 영향요인을 입력변수로 취급하여 모델을 구성하는 다중회귀 모델이 많이 이용되고 있으며, 시계열 분석방법으로는 ARIMA 모델이 대표적인 모델로 사용되고 있다(이경훈 외, 1998).

근래에 인간의 뇌를 모방하고 있는 인공신경망 이론과 사람의 자연적 언어 체계를 표현하는데 효과적인 퍼지이론을 이용하여 기존의 방법으로 해결하기 어려웠던 여러 가지 공학 분야들(패턴인식, 음성인식, 시스템 제어)에 성공적으로 이용되고 있다(공성곤, 1995).

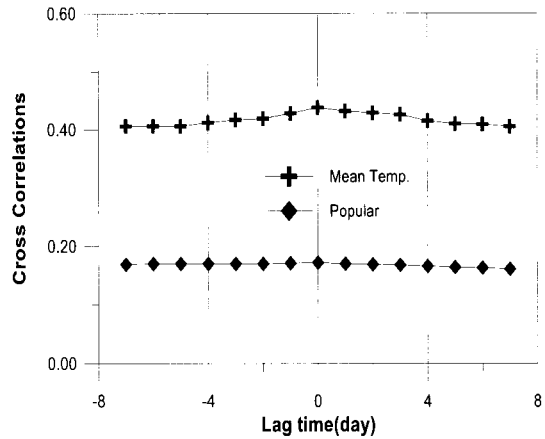
이번 연구에서는 상수급수량 자료와 인구, 경년변화 그리고 기상자료들을 이용하여 상관분석을 실시한 후 입력자료를 결정하고 상수도 급수량의 효율적인 운영과 보다 효과적인 급수량 예측을 위한 모델 개발의 관점에서 비선형 모형인 신경망 모형과 뉴로-퍼지(Neuro-Fuzzy) 모형으로 일급수량 예측 모델을 구성한 후 단기수요 예측방법 모델을 제안하고, 그 적용성을 검토하고자 한다.

2. 대상지역의 자료 및 통계지표

대상지역은 충청북도의 중심도시인 청주시로 선정하였다. 청주시의 인구는 최근 10년 동안 127,610명의 인구가 증가하였으며, 급수인구는 60만 1천 명으



a) 자기상관분석



b) 교차상관분석

Fig. 2. 급수량에 대한 상관분석.

로 전체시민의 96%가 수도물을 사용하고 있고, 시민 1인당 하루 물사용량은 321l로 상수 사용량이 매년 증가하는 추세를 보이고 있다.

1일 급수량 예측을 위해 청주시의 1998년 1월1일부터 2001년 7월 31일까지의 급수량 자료, 인구 및 기후 자료를 활용하였으며, Fig. 1에 급수량, 상대습도, 강수량, 평균기온을 나타내었다.

급수량 자료는 청주시 물관리사업소에서 관리하는 상수도 운용일지를 이용하였으며, 기후자료는 청주시 상관측소의 통계자료를 활용하였다. 인구자료는 청주시 통계연보와 통계청 자료를 이용하였으며, 1일 급수인구를 추정하기 위해 등차급수법을 사용하였다.

입력변수를 결정하기 위해 급수량과 기후 및 인구 자료들의 상관분석을 실시하였다. 상관계수의 크기는 평균기온, 인구, 강수량, 상대습도 순으로 분석되었고, 강수량 및 상대습도는 다른 변수들에 비해 상대적으로 상관도가 낮았다. 따라서 급수량, 평균기온, 인구 자료를 입력변수를 결정하였다. 급수량 자료에 대한 자기상관 분석과 평균기온, 인구자료간의 교차상관분석을 수행하여 Fig. 2에 나타내었다.

Fig. 2에서 알 수 있듯이 급수량은 지체시간이 증가할수록 상관도가 감소하였으며, 지체시간 10일 이후 까지 계열상관도가 0.6 이상의 값을 나타내었고, Fig. 2의 교차상관 분석에서는 평균기온의 경우 지체시간이 0일에 교차상관이 가장 높았으며, 급수인구도 0일에 교차상관이 가장 높게 나타났다. 따라서 모형의

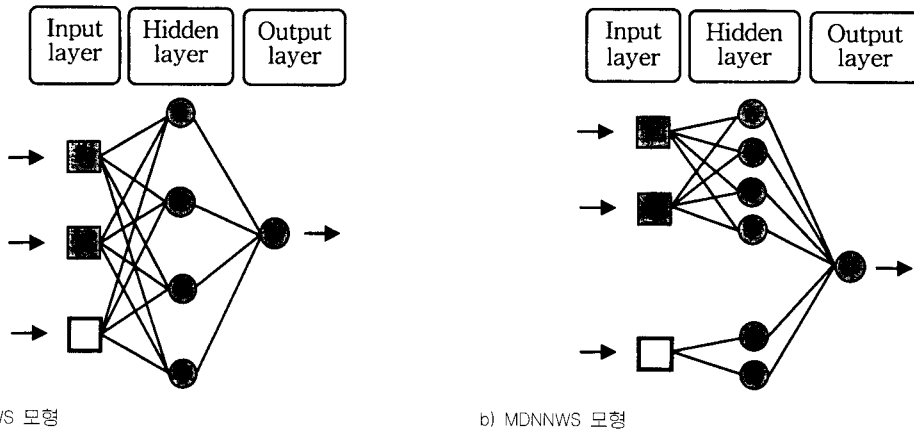
Table 1. 통계지표

Item	Relation Equation
Root Mean Square Error (RMSE)	$\left\{ \frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2 \right\}^{\frac{1}{2}}$
Root Relative Square Error (RRSE)	$\left\{ \frac{1}{N} \sum_{i=1}^N \left(\frac{X_i - Y_i}{X_i} \right)^2 \right\}^{\frac{1}{2}}$
Mean Average Relative Error (MARE)	$\frac{1}{N} \sum_{i=1}^N \left \frac{X_i - Y_i}{X_i} \right $
Percent Error in Volume (PEV)	$\left(\frac{Y_v - X_v}{X_v} \right) \times 100$
Percent Error in Peak (PEP)	$\left(\frac{Y_p - X_p}{X_p} \right) \times 100$

간소화와 기상 예보의 난이함을 고려할 때 급수량의 경우는 예측일의 1~2일 전의 자료를 사용하고, 평균기온과 급수인구는 예측 해당일이나 전일의 자료를 사용해야 할 것으로 판단된다.

연구결과에 대한 분석을 위해 Table 1의 통계지표를 사용하였다.

여기서, N은 자료의 수, X_i는 실측 급수량, Y_i는 계산 급수량, X_v는 실측 급수량의 용적, Y_v는 계산 급수량의 용적, X_p는 실측 급수량의 최대치, Y_p는 계산 급



a) LMNWS 모형

b) MDNWS 모형

Fig. 3. 적응 신경망 모형의 구조.

수량의 최대치이다.

3. 적응 모형

3.1. 신경망 모형

본 연구에서 사용한 다층 신경망 모형은 크게 네트워크의 구조 형태에 따라 2가지로 구분하였다. 은닉층이 분리되지 않아 입력자료의 위치만을 고려하여 학습하는 일반적인 형태의 신경망 모형과 입력자료의 다른 성질을 구분하여 학습할 수 있도록 은닉층을 분류하여 구성한 모형으로 구분하였으며, 은닉층이 분류된 Modular 신경망 모형의 일반적인 구조를 Fig. 3에 나타내었다.

입력층과 은닉층의 Unit을 모두 연결하는 일반적인 신경망 네트워크와는 달리 Modular 네트워크는 전체 네트워크를 부분적인 네트워크(sub-network)로 분류하여 성질이 다른 입력변수를 구분하고 해당은닉층에만 영향을 미치도록 구성되어 있어 입력 자료별로 나누어서 학습을 진행시키므로 학습 시간을 단축시킬 수 있을 뿐 아니라 자료의 특성을 살려 동시에 학습시킬 수 있다. 급수량 예측을 위한 2가지 신경망 모형은 모두 출력값을 생성하는 전달함수는 시그모이드 함수를 사용하였으며, 학습 방법으로 역전파 알고리즘을 학습향상을 위해 Levenberg-Marquardt 방법을 사용하였다. 학습과정에서 연결강도를 조정할 때 이용되는 Levenberg-Marquardt update rule은 그 자체가 일련의 선형방정식들의 해인 탐색방향을 사용하며,

식 (1)과 같이 나타낼 수 있다.

$$\Delta W = (J^T J + \eta I)^{-1} J^T e \quad (1)$$

여기서, J 는 각 연결강도에 대한 오차 미분의 Jacobian matrix, η 는 적응학습률 그리고, I 는 학습률 변화를 반영한 단위행렬, e 는 오차벡터이며 적응 학습률 η 의 변화를 통해 학습시 빠르게 수렴한다.

3.2. ANFISWS 모형

ANFISWS(Adaptive Neuro-Fuzzy Inference Systems for Water Supply)는 신경회로망과 퍼지이론이 결합된 형태로서 신경회로망의 구조와 학습능력을 이용하여 제어환경으로부터 얻은 입출력 정보로부터 소속함수와 제어규칙을 제어 대상에 맞게 자동 조절하는 방식의 모형이다.

ANFISWS의 기본적인 학습방법은 각 노드의 출력에 대한 제공된 오차의 미분 즉, 오차 신호를 계산하는 역전파 알고리즘에 의해 후방향 경로에서 전제부 매개변수를 조정한다. 이 학습방법은 기존의 다층 퍼셉트론에서 사용된 역전파 학습방법과 동일한 방법이다. 전제부 매개변수를 학습하기 위해 MSE(Mean Square Error)를 최소화하기 위한 목적함수는 목표치 d_i 와 출력된 계산치 y_i^* 에 의해 식 (2)와 같이 나타낼 수 있다.

$$E = \frac{1}{2} \sum_{i=1}^n (d_i - y_i^*)^2 \quad (2)$$

소속함수의 모양이 전제부 매개변수에 의해 정의되어지기 때문에 목적함수 E 는 결론부 매개변수가 고정된다면 소속함수의 폭과 중심을 계산하는 a_i, c_i 의 매개변수로 결정되어진다. 여기서 a_i 는 소속함수의 폭, c_i 는 소속함수의 중심을 나타낸다. 목적함수 E 의 값을 감소시키기 위해서는 전제부 매개변수에 대해서 다음과 같이 최대 경사법(steepest descent method)에 의해 반복적으로 이루어진다.

$$a_i(t+1) = a_i(t) - \eta \frac{\partial E}{\partial a_i} \quad (3)$$

$$c_i(t+1) = c_i(t) - \eta \frac{\partial E}{\partial c_i} \quad (4)$$

$$\eta = \frac{k}{\sqrt{\sum_{\alpha} \left(\frac{\partial E}{\partial \alpha} \right)^2}} \quad (5)$$

여기서, k 는 매개변수 공간에서 경사벡터의 이동한 거리로 수렴 속도 변화에 영향을 미친다. 전제부 매개변수가 주어질 때 전체적인 출력 f 는 결론부 매개변수인 $\{p_j, q_i, r_j\}$ 의 선형적인 조합에 의해 식 (6)과 같이 표현되어질 수 있다. 여기서 \bar{w} 는 정규화된 발화 강도를 의미한다. Fig. 4에 ANFISWS 모형의 구성도를 나타내었다. Fig. 4에서 알 수 있듯이 ANFISWS는 데이터가 입력되면 소속함수의 제어와 신경망의 학습률을 적용하여 매개변수를 학습하여 결과를 출력한다.

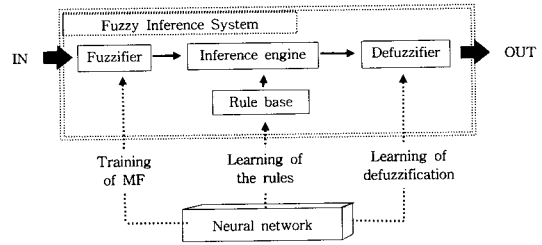


Fig. 4. ANFISWS 모형의 구조.

Table 2. 4가지 경우의 입출력 구성

Case	Input	Output
Case I	Q_t, Q_{t-1}, T_{t+1}	Q_{t+1}
Case II	Q_t, Q_{t-1}, T_t	Q_{t+1}
Case III	$Q_t, Q_{t-1}, T_{t+1}, P_{t+1}$	Q_{t+1}
Case IV	Q_t, Q_{t-1}, T_t, P_t	Q_{t+1}

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2$$

$$= (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + (\bar{w}_1) r_1 + (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 + (\bar{w}_2) r_2 \quad (6)$$

4. 적용 및 결과

4.1. 모형의 구성

청주시 1일 급수량 예측을 위해 은닉층을 분리하지 않은 일반적인 신경망 모형(LMNNWS)과 은닉층을 분리한 Modular 신경망 모형(MDNNWS), 그리고 신경망과 퍼지를 결합시킨 ANFISWS 모형을 적용하였으며, 자료 분석을 통해 결정한 입력변수에 따라 Table 2와 같이 구성되었다. $t+1$ 일의 급수량(Q_{t+1}) 예측시에 사용한 자료들은 $t-1$ 일의 급수량(Q_{t-1}),

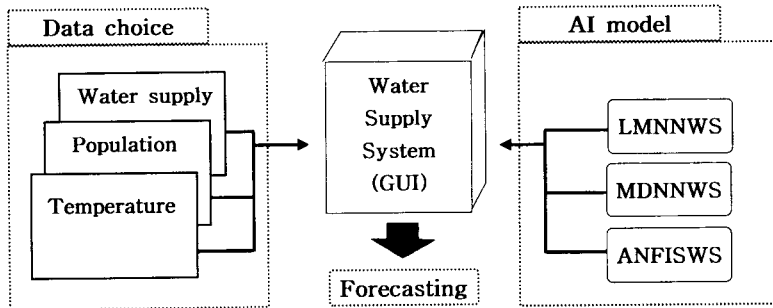


Fig. 5. 상수도 급수량 예측 시스템 구성도.

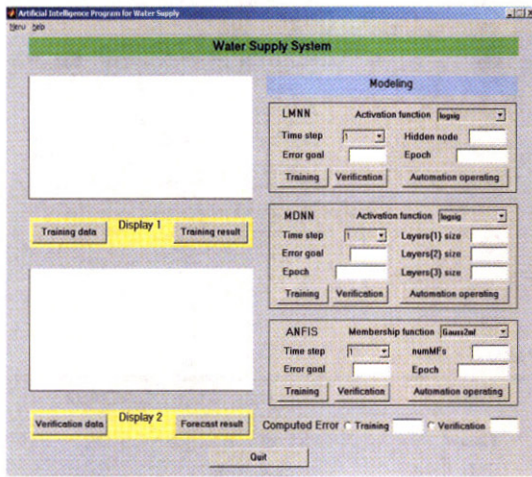


Fig. 6. GUI 구현.

t 일의 급수량(Q_t), t 일의 수온(T_t), $t+1$ 일의 수온(T_{t+1}), t 일의 인구(P_t), $t+1$ 일의 인구(P_{t+1}) 자료를 조합하여 4가지 경우에 대해 모의할 수 있도록 하였다.

구성된 신경망 모형들은 Matlab을 통하여 GUI로 구현된 시스템을 통해 학습에서 예측까지 운영이 가능하도록 연계되었으며, 상수도 급수량 예측을 위한 시스템 구성도는 Fig. 5에, 구현된 GUI는 Fig. 6에 나타내었다. GUI 내에서는 인공지능 모형들의 구조를 변경할 수 있도록 은닉층 노드수 및 소속함수의 수 (Number of Membership Function), 학습 목표오차, 학습회수에 대한 기능과 운영상의 편의를 위해 자동 수행할 수 있는 기능을 추가하였으며, 사용자가 학습 및 예측 결과를 확인할 수 있도록 하였다.

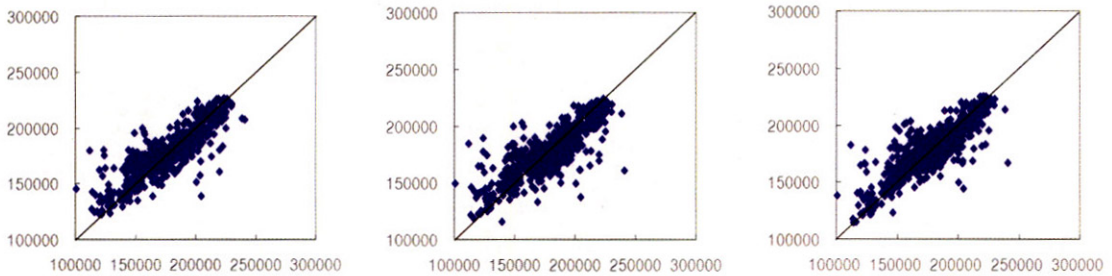


Fig. 7. 각 모형에 대한 학습결과.

Table 3. 인공지능 모형의 학습결과 분석

Error Model		Structure	RMSE	MARE	PEV	PEP
LMNNWS	Case I	3-6-1	14529.77	0.0568	0.00331	7.955
	Case II	3-6-1	14534.12	0.0577	0.00351	9.284
	Case III	4-8-1	14539.45	0.0579	0.00374	9.497
	Case IV	4-8-1	14536.26	0.0581	0.00362	9.372
MDNNWS	Case I	3-[3]-[3]-1	14449.33	0.0432	0.00013	6.501
	Case II	3-[3]-[3]-1	14449.92	0.0492	0.00025	6.777
	Case III	4-[2]-[2]-1	14450.38	0.0591	0.00112	6.945
	Case IV	4-[2]-[2]-1	14450.12	0.0506	0.00093	6.805
ANFISWS	Case I	3-2-1	14526.74	0.0511	0.00045	6.484
	Case II	3-2-1	14529.68	0.0524	0.00062	6.756
	Case III	4-2-1	14535.32	0.0599	0.00253	7.128
	Case IV	4-2-1	14533.38	0.0563	0.00138	6.767

주) LMNNWS:Inputnode-Hiddennode-Outputnode
 MDNNWS:Inputnode-[Hiddenlayer1node]-[Hiddenlayer2node]-Outputnode
 ANFISWS:Inputnode-Membership Function-Outputnode

4.2. 결과

구성된 LMNNWS, MDNNWS, ANFISWS의 인공지능 모형을 이용하여 1998년 1월~1999년의 12월 자료를 학습하였으며, 각 모형에 대한 구조 및 학습결과를 Fig. 7과 Table 3에 나타내었다. 학습결과에서 나타난 바와 같이 각각의 인공지능 모형들은 충분한 학습을 통해 실측치의 경향을 잘 인식하였다. 특히 입력자료로서 전일급수량(Q_t), 전전일 급수량(Q_{t-1}), 예측일의 평균기온(T_{t+1})을 사용한 Case I의 학습결과가 RMSE, MARE, REV, PEP 등 통계분석에서 가장 우수하게 분석되었으며, Case II~IV의 경우도 우수한 결과를 나타내었다.

인공지능 모형들의 단기 급수량 예측모형의 예측력을 검증하기 위해 2000년 1월~2001년 7월까지의 자료를 검증자료로 사용하여 1일 급수량의 단기예측을 실시하였다. Fig. 8은 상수도 급수량 예측을 위해 GUI로 구현된 시스템의 운영 결과이며, Table 4와 Fig. 9-11에 인공지능 모형들에 의한 급수량 예측 결과를 나타내었다. Fig. 9-11에서 알 수 있듯이 검증자료인 2001년 1월~2001년 7월까지의 실측 급수량 자료를 인공지능 모형들로 예측한 결과와 비교하면 인위적 조작(단수)등이 가해지는 6월 초순, 10월 중순 및 12월 중순경의 시기를 제외하면 예측력이 비교적 우수한 것으로 나타났다.

입력자료의 형태별로 분석한 결과, Case I의 경우가 가장 좋은 결과를 보였으며, Case II, Case IV, Case III의 순으로 나타나 급수량의 변화 성향과 예측

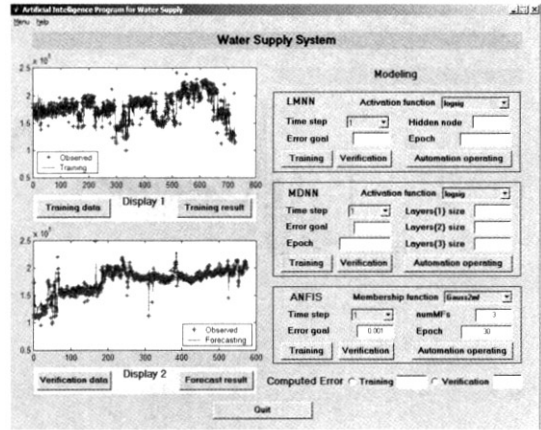


Fig. 8. 급수량 예측을 위한 인공지능시스템의 GUI 구현 결과.

해당일의 평균기온이 중요한 인자로 나타났다. 적용 모형별로 보면 MDNNWS 모형은 RRSE가 0.0891, MARE는 0.0488의 오차를 나타내며 가장 좋은 결과를 보였으며, 다음으로 ANFISWS(RRSE: 0.0908, MASE: 0.0527), LMNNWS(RRSE: 0.1004, MASE: 0.0588), 모형순으로 작은 오차를 보였다. 신경망과 퍼지이론을 조합한 ANFISWS 모형은 침투치 퍼센트 오차(PEP)가 7.245%로 신경망 모형들(MDNNWS: 7.973%, LMNNWS: 9.623%)에 비해 작게 나타나 자료의 변화가 심한 초반부에 안정된 예측 결과를 보였으나 후반부에는 과소 예측된 경향을 보였다. 신경망 모형 중에서는 은닉층을 구분하여 학습한 MDNNWS 모형이 좋은 결과를 보였으며, 출력값과

Table 4. 각 모형에 의한 급수량 예측결과

Error		RRSE	MARE	PEV	PEP
Model					
LMNNWS	Case I	0.1004	0.0588	2.102	9.623
	Case II	0.1030	0.0592	2.466	11.951
	Case III	0.1650	0.1003	4.661	14.337
	Case IV	0.1589	0.1035	4.335	13.901
MDNNWS	Case I	0.0891	0.0488	0.223	7.973
	Case II	0.0899	0.0502	0.434	8.829
	Case III	0.1017	0.0613	0.692	9.781
	Case IV	0.0926	0.0514	0.572	9.125
ANFISWS	Case I	0.0908	0.0527	0.313	7.245
	Case II	0.0929	0.0538	0.526	7.266
	Case III	0.1009	0.0605	1.163	8.691
	Case IV	0.0965	0.0574	0.996	7.664

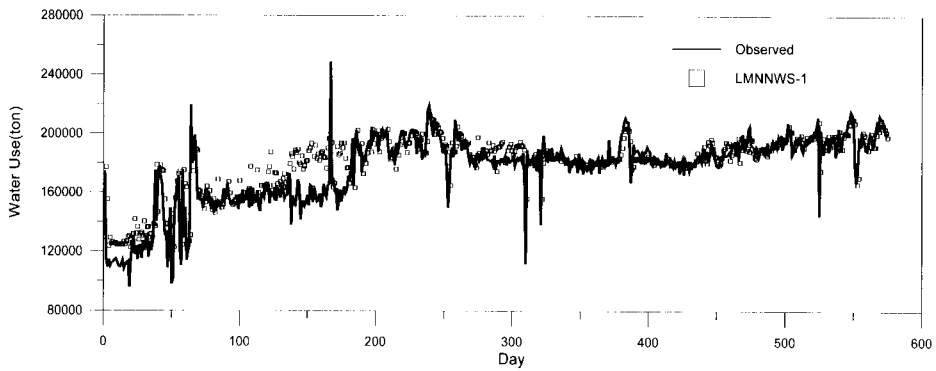


Fig. 9. LMNWS-1 모형의 1일 급수량 예측결과.

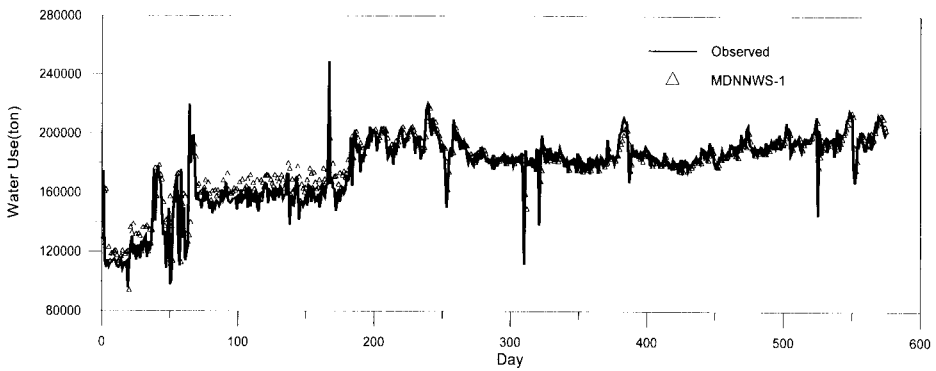


Fig. 10. MDNWS-1 모형의 1일 급수량 예측결과.

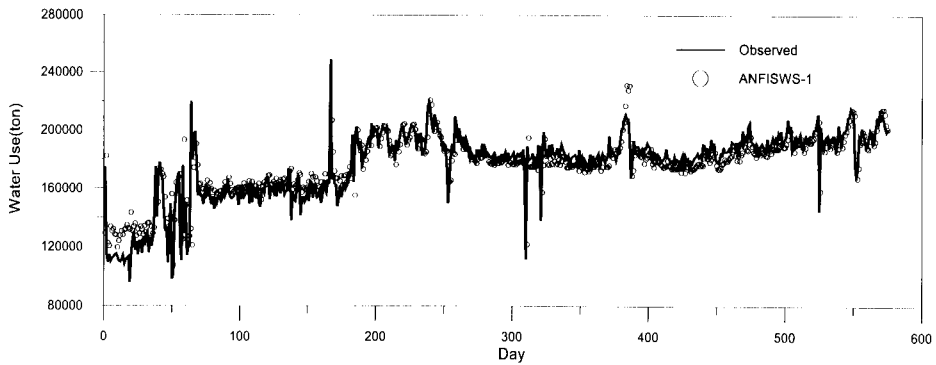


Fig. 11. ANFISWS-1 모형의 1일 급수량 예측결과.

성격이 다른 입력자료를 사용할 경우 은닉층을 분리하는 것이 효과적인 것으로 분석되었다.

5. 결 론

상수도 1일 급수량 예측을 위해 Levenberg-

Marquardt 학습방법을 사용하는 일반적인 형태의 신경망 모형(LMNWS)과 입력자료의 성격을 구분하여 학습한 Modular 신경망 모형(MDNWS) 그리고 신경망과 퍼지이론을 조합한 ANFISWS 모형 등의 인공지능 모형들을 적용하고 다음과 같은 결론을 얻었다.

1) 각 모형 모두 급수량 예측을 위한 입력자료로 예측 해당일의 기온(T_{t+1})과 예측일의 이전 급수량자료(Q_t, Q_{t-1})로 구성된 Case I 모형이 좋은 결과를 나타냈다.

2) 모형별로는 MDNNWS 모형, ANFISWS 모형, LMNNWS 모형 순으로 실측치에 근사하였으며, MDNNWS 모형과 ANFISWS 모형은 근소한 차이를 나타냈다. ANFISWS 모형은 자료의 변동폭이 심한 경우에도 안정된 예측 결과를 보여 불규칙하게 공급되는 급수량 예측에 적합한 모형으로 분석되었고, 급수량 총량을 예측시에는 MDNNWS 모형이 적합한 모형으로 분석되었다.

3) 일급수량 예측을 위한 시스템 구축은 Matlab을 이용하였으며 GUI로 구현된 시스템을 통해 학습에서 예측까지 운영이 가능하도록 연계하였으며 이를 이용하면 수도시설을 운영 관리할 때 앞으로의 급수량을 예측함으로써 각종 시설의 운영자가 예측적인 업무를 수행할 수 있고 전력비 등 효율적인 운전에도움을 줄 것이다.

참고문헌

공성곤, 김민수 (1995) 적응 퍼지제어(뉴로-퍼지제어 기술

특집), *제어·자동화·시스템 공학회지*, 1(3), pp. 101-108.

구자용, 소천명, 도원토요노, 급수량 단기 수요예측에 관한 연구, *상하수도학회지*, 11(1).

김주환, 김종우, 박재홍 (1999) 최적화 기법을 이용한 광역 상수도 관로시스템 설계, *상하수도학회지*, 13(1), pp. 95-112.

연인성(2005) 실시간 유량-수질관리를 위한 인공지능시스템 개발, 충북대학교 대학원박사학위논문, pp. 15-40.

이경훈, 강일환, 문병석(2000) 상수도 1일 급수량 예측을 위한 ANFIS적용, *상하수도학회지*, 14(3), pp. 281-290.

이경훈, 문병석, 오창주 (1998) 역전파 알고리즘을 이용한 상수도 일일 급수량 예측, *상하수도학회지*, 12(4), pp. 43-52.

전계원(2004) 홍수유출과 수질예측을 위한 Web기반 시스템의 개발, 충북대학교 대학원 박사학위논문, pp. 13-39.

현인환, 목동우(1995) 상수 급수량의 단기 예측, *상하수도학회지*, 9(1), pp. 96-106.

J.-S. R. JANG, C.-T. SUN, E. MIZUTANI (1997) *Neuro-Fuzzy and Sofi Computing*, Prentice Hall International, Inc.

Simon Haykin (1999) *Neural Networks* 2nd Ed. Prentice Hall International, Inc.

Tay, J.H. and Zhang, X.(2000) A Fast Prediction Neural Fuzzy Model for High-rage Anaerobic Wastewater Treatment Systems, *Water Resources Research*, 34(11), pp. 2849-2860.