

ANFIS를 이용한 상수도 1일 급수량 예측에 관한 연구

A Study of Prediction of Daily Water Supply Using ANFIS

이 경 훈* / 문 병 식** / 강 일 환***

Rhee, Kyoung Hoon / Moon, Byoung Seok / Kang, Il Hwan

Abstract

This study investigates the prediction of daily water supply, which is a necessary for the efficient management of water distribution system. Fuzzy neuron, namely artificial intelligence, is a neural network into which fuzzy information is inputted and then processed. In this study, daily water supply was predicted through an adaptive learning method by which a membership function and fuzzy rules were adapted for daily water supply prediction. This study was investigated methods for predicting water supply based on data about the amount of water supplied to the city of Kwangju. For variables choice, four analyses of input data were conducted: correlation analysis, autocorrelation analysis, partial autocorrelation analysis, and cross-correlation analysis. Input variables were (a) the amount of water supplied (b) the mean temperature, and (c) the population of the area supplied with water. Variables were combined in an integrated model. Data of the amount of daily water supply only was modelled and its validity was verified in the case that the meteorological office of weather forecast is not always reliable. Proposed models include accidental cases such as a suspension of water supply. The maximum error rate between the estimation of the model and the actual measurement was 18.35 % and the average error was lower than 2.36 %. The model is expected to be a real-time estimation of the operational control of water works and water/drain pipes.

Keywords: Daily Water Use, Fuzzy Nueron, Fuzzy Information, Membership Function, Fuzzy Rule, Adaption Learning

요 지

본 논문에서는 상수도시설을 효율적으로 운영하는 데 필요한 1일 급수량 수요를 예측하는 방식에 대하여 인공지능(Artificial Intelligence)이라 불리는 퍼지 뉴런(fuzzy neuron)을 이용하여 연구하였다. 퍼지뉴런이란 퍼지정보(fuzzy information)를 입력으로 받아들이고 처리하는 퍼지 신경망을 일컫는 말이다. 본 연구에서는 소속함수와 퍼지규칙을 신경망으로 학습하는 기능인 적응식 학습방법을 통하여 1일 급수량을 예측하였으며 연구대상 지역으로는 광주광역시를 선정하였다. 또한 1일 급수량 예측에 있어서 필요한 변수 선택을 위해 입력자료를 상위분석, 자기상관, 부분자기상관, 교차상관 분석 등을 하였으며 동정된 입력변수는 급수량, 평균기온, 급수인구이다. 먼저 급수량, 평균기온, 급수인구로 모델을 구성하였고, 한편으로 기상청의 기후예보자료를 신뢰할 수 없는 경우에도 급수량을 예측할 수 있도록 급수량 차로만으로 모델을 구성하여 그 유효성을 검증하였다. 제안된 모형식은 사고 등의 인위적인 조작(단수 등)이 가해지는 시기를 포함하고도 실측치와 모형의 예측치와의 오차율이 최대 18.46 %, 평균 2.36 % 이내로 나타나, 모형의 결과는 상수도 시설의 운용 및 급·배수관망의 실시간 제어에 많은 도움을 주리라 생각된다.

핵심용어 : 일 급수량, 퍼지 뉴런, 퍼지정보, 소속함수, 퍼지규칙, 적응식 학습

* 전남대학교 공과대학 토목공학과 부교수

Associate Prof., Dept. of Civil Eng., Chonnam National Univ., Kwangju 500-757, Korea

** 서남대학교 공과대학 토목공학과 전임강사

Full-Time Lecturer, Dept. of Civil Eng., Seonam Univ., Namwon, Chonbuk 590-170, Korea

*** 전남대학교 공과대학 토목공학과 석사과정

Graduate Student, Dept. of Civil Eng., Chonnam National Univ., Kwangju 500-757, Korea

1. 서론

산업구조가 기반산업 위주에서 고도의 공업위주로 전환되고, 급격한 경제성장을 이루면서 국민의 생활수준이 크게 향상되고, 인구가 도시에 편중되어 도시의 비대화를 초래하고 있다. 이러한 현대사회는 복합적인 문제점을 유발하고 있으며 물 수요량도 생활용수, 영업용수, 공업용수, 농수산업용수 등 다방면에 걸쳐서 수요량이 증가하여, 기존의 용수로는 수량면에서 수요량을 충족시키지 못하고 있어, 수돗물의 안정된 공급은 아주 중요한 의미를 지니게 되었다.

따라서 상수도의 과학적이고 합리적인 운영을 통해서 수량 및 수질면에서 수요자의 요구에 맞는 질 높은 수도 서비스를 제공하기 위해서는 수도시설을 적절하게 갖추고 경제적으로 운용하지 않으면 안되는데, 이것은 적확(的確)한 수돗물 수요예측이 이루어져야만 가능하게 되는 것이다(이경훈 등, 1994b).

급수량 예측방법은 시점에 따라 단기예측과 장기예측으로 분류할 수 있다. 단기예측은 시간 및 일일예측으로 급·배수량의 실시간 제어에 이용되고, 장기예측은 순, 월, 계절 연별 예측으로 상수도시설을 계획하거나 설계하는데 필요한 자료로 활용되고 있다(Young, 1973).

수돗물 사용량의 단기예측은 중요한 정보를 제공하고 있다. 현재 수요예측은 그 목적에 따라 여러 가지 수법이 제안되었지만, 급·배수제어를 목적으로 한 단기예측에 관한 연구는 많지 않다. 상수도의 합리적인 운용에 필수적인 급·배수제어를 위해서는 일 또는 시간 데이터에 의한 상수사용량의 확률적 예측이 필요하다(이경훈 등, 1994a). 전날 또는 수 시간 전에 장래에 대한 상수 운용계획을 세우면 합리적인 송·배수 펌프의 운용, 밸브의 조정 및 배수지 수위의 설정 등이 가능하게 될 것이다(이경훈 등, 1994c).

현재 급수량의 수요예측방법에는 크게 다변량 해석 방법과 시계열 분석방법이 있다. 다변량 해석방법의 대표적인 방법으로는 영향요인을 입력변수로 취급하여 모델을 구성하는 다중회귀모델이 가장 많이 이용되고 있으며, 시계열 분석방법으로는 자기회귀 이동평균모델인 ARIMA모델 (Auto Regressive Integrated Moving Average Model)을 대표적인 것으로 거론할 수 있다.

최근에 생물학적 신경 계통을 모방한 인공 신경망과 사람의 자연적인 언어 체계를 표현하는 데에 효과적인 퍼지 이론은 기존의 방법으로 해결하기 어려웠던 여러

가지 공학분야들(패턴 인식, 음성 인식, 시스템 동경화와 제어)에 성공적으로 이용되고 있다(공성근과 김민수, 1995).

이처럼 다양한 분야에서 신경망과 퍼지 이론이 사용되고 있고 실생활과 아주 밀접한 관계가 있는 급수량 예측을 위한 신경망 적용 연구를 보면 문병석(1998)은 급수량, 급수인구, 기후자료를 이용하여 신경망 구성을 하여 모넨팅 및 적응식 학습률 방법을 이용하여 1일 급수량을 예측하였다.

보다 효과적인 급수량 예측을 위한 모델 개발의 관점에서, 본 논문에서는 인공지능인 퍼지 논리와 신경망의 융합 기술중 하나인 ANFIS를 이용해 1일 급수량을 예측하였다. 하지만, 진정한 의미에서의 인공지능을 실현하기 위해서는 신경망, 퍼지 논리, 전문가 시스템 및 유전 알고리즘을 보다 효과적으로 이용하고 통합해야 가능할 것으로 기대된다.

본 연구의 목적은 상수도의 합리적이고 경제적인 급수량의 운영을 도모하기 위해, 배수제어 실시를 위한 가장 중요하면서도 기본적 자료인 급수량의 단기수요 예측방법의 모델을 제안하고 그 적용성을 검증하고자 하는 것이다.

2. 대상지역 및 연구자료

2.1 대상지역

본 연구의 대상지역으로는 광주광역시를 선정하고, 1995년과 1996년 및 1997년의 급수량, 기후, 급수인구 자료를 활용하였다. 이 자료 중 1995년과 1996년의 급수량, 기후 및 급수인구 자료는 모델의 학습자료로 사용하였고, 1997년의 자료를 검증자료로 활용하였다.

기후자료는 광주지방기상청 일별 기상자료인 평균기온, 최고기온, 상대습도, 평균우량, 일사량, 일조시간을 이용하였고, 급수량 자료는 광주광역시 상수도사업본부가 가지고 있는 급수일지를 활용하였으며, 급수인구는 광주광역시의 통계연보에 있는 자료를 활용하였다. 그러나 통계연보는 연(年)단위의 인구이므로 1일 급수인구로 활용하기 위해 등차급수법으로 인구를 예측하였다.

광주광역시의 1997년 12월말 현재 인구는 1,326,000여명에 이르고 있으며, 급수량을 살펴보면 1997년 말 현재 최대 47만톤/일 규모의 급수량에 달하고 있다. 면적은 약 500 km²로서, 상수사용량이 생활용수를 중심으로 매년 증가하고 있는 지역이다.

2.2 연구자료

본 연구에서는 광주광역시 1일 급수량 예측을 위한 입력변수를 동정하기 위해 통계적 분석에 원시계열 자료를 사용하였다. 또한 변수에 대한 상관분석을 하여 입력변수를 동정하였고, 자기상관 및 교차상관분석을 통해 동정된 변수에 대한 지체시간(lag time)을 결정하여 모델을 구성하였다. 따라서 본 절에서는 상관분석, 자기상관 및 교차상관의 분석 방법을 통해 얻어진 분석결과를 제시하기로 한다.

1995~1997년 자료의 상관계수는 Table 1과 같다. Table 1을 보면 알 수 있듯이 급수량과 급수인구의 상관은 0.533정도이며, 평균기온과 최고기온과의 상관은 각각 0.657, 0.631로 나타났고, 다른 변수와의 상관은 낮음을 알 수 있다.

Fig. 1은 1997년 1년간의 급수량(Water), 급수인구

(Pop) 및 평균기온(MeanT)에 대한 계열상관도(Correlogram)이며, 시계열 자료의 정상성을 검토할 때 유효하다. Fig. 1을 보면 급수량의 자기상관함수는 완만하게 감소하고 있으며 지체시간 30일에서도 0.6이상 유지하고 있다. 이것은 급수량은 자기상관이 매우 크고, 반복성이 없다고 말할 수 있다. 즉, 오늘의 급수량은 전일 그 전전일 등의 급수량으로부터 많은 영향을 받고, 그 주기가 상당히 크다는 것을 알 수 있다. 또한 급수인구 및 평균기온은 급수량과 마찬가지로 자기상관이 크므로 반복주기가 크다는 것을 알 수 있다.

Fig. 2는 급수량과 상관계수가 높은 급수인구(Pop), 평균기온(MeanT)간의 교차상관을 나타낸 것이다. 그림을 보면 알 수 있듯이 일 급수량과 평균기온은 지체시간이 -1과 0일에 교차상관이 높고, 일 급수량과 급수인구는 지체시간 0일에 가장 교차상관이 높은 것을 알 수 있다.

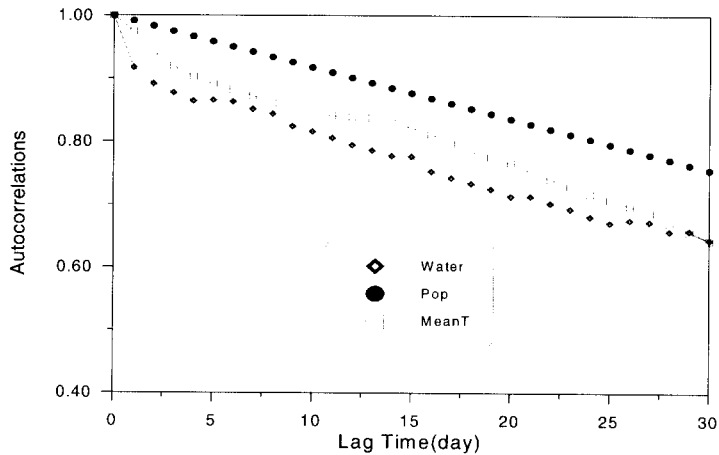


Fig. 1. Auto-correlations of Water Demand, Popular and Mean Temperature for 1997

Table 1. Correlation Coefficient of Variables(1995~1997)

Variable	Water demand	Popular	Mean Temp.	Max Temp.	Humidity	Cloudiness	Precipitation	Shining hours
Popular	0.533							
Mean Temp.	0.657	0.124						
Max Temp.	0.631	0.128	0.978					
Humidity	0.260	0.057	0.324	0.223				
Cloudiness	0.039	-0.060	0.177	0.047	0.607			
Precipitation	-0.131	0.074	0.153	0.095	0.392	0.338		
Shining hours	0.033	0.010	0.027	0.154	-0.648	-0.826	-0.349	
Shining amount	0.168	-0.062	0.385	0.486	-0.490	0.584	-0.201	0.848

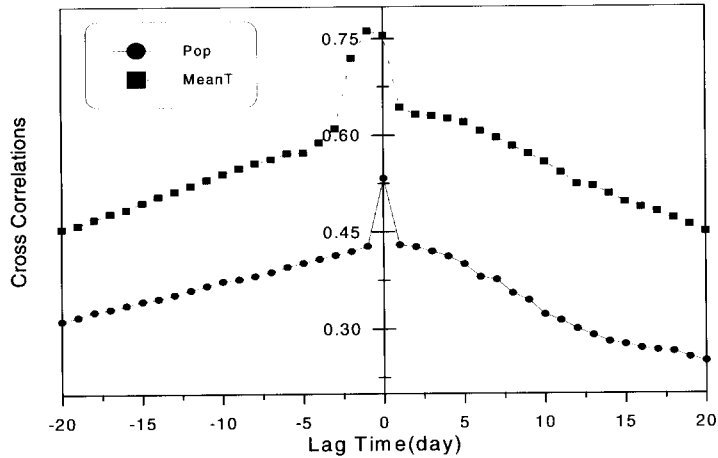


Fig. 2. Cross Correlations of Water Demand, Popular and Mean Temperature for 1997

이상에서 살펴보았듯이 광주광역시 1일 급수량 (Q_{t+1}) 예측을 위한 입력변수는 급수인구 (P), 평균기온 (T), 급수량(Q)으로 동정하여야 하며, 급수인구는 예측일의 급수인구 (P_{t+1}), 평균기온은 예측전날과 예측일 (T_t, T_{t+1})의 평균기온, 급수량은 예측일의 삼일 전부터 예측전날 (Q_{t-2}, Q_{t-1}, Q_t)까지로 모델을 구성해야 할 것으로 분석되었다.

또한 기상청의 기후예보자료를 신뢰할 수 없는 경우에도 급수량을 예측할 수 있도록 급수량 자료만으로 모델을 구성해야 하는데, 모델을 구성하는 경우는 예측 5일전부터 예측전날 ($Q_{t-4}, Q_{t-3}, Q_{t-2}, Q_{t-1}, Q_t$)까지의 급수량 자료로 모델을 구성하여야 할 것으로 판단된다.

3. 급수량 수요예측모델

본 연구에서는 기존의 연구방법이 아닌 인공지능(Artificial Intelligence)중에서 퍼지(Fuzzy)와 신경망(Neural Networks)의 장점을 모두 포함한 ANFIS

(Adaptive-Neural-Based Fuzzy Inference System) 모델을 이용하여 1일 급수량 예측을 하기로 한다.

Table 2는 입력변수와 데이터 범위에 따른 ANFIS 모델의 구성을 보여준다. 입력변수는 앞 절에서 동정된 급수량, 평균기온, 급수인구와 급수량만으로 되어 있고, 데이터의 범위는 훈련자료로 '95~'96년과 '96년을 사용하였고 검증자료로 '97년을 사용하였다. 따라서 원래의 모델에 시계열장을 달리하여 모델을 구성하였다.

3.1 ANFIS(Adaptive-Neural-based Fuzzy Inference System)모형

Fig. 3은 일반적으로 사용되는 퍼지 규칙과 퍼지 추론 메카니즘을 나타내고 있다. 여기서, 퍼지 규칙은 전건부와 후건부로 구성되고, 후건부는 3가지 형태를 취하고 있다. 본 논문의 ANFIS모형에서는 Takagi · Sugeno-type이라 불리는 3번째 형태의 것이 사용되며 Table 3처럼 5단계의 추론과정을 거친다(Jang, 1993). 여기서, A, B, C는 각각의 변수를 소속함수로

Table 2. Construction of Model according to Input Variable and Data Range

Model \ Range	Input variable	Training data	Checking data
Model 1a	$Q_t, Q_{t-1}, Q_{t-2}, P_{t+1}, T_t, T_{t+1}$	'95~'96	'97
Model 1	$Q_t, Q_{t-1}, Q_{t-2}, P_{t+1}, T_t, T_{t+1}$	'96	'97
Model 2a	$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$	'95~'96	'97
Model 2	$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$	'96	'97

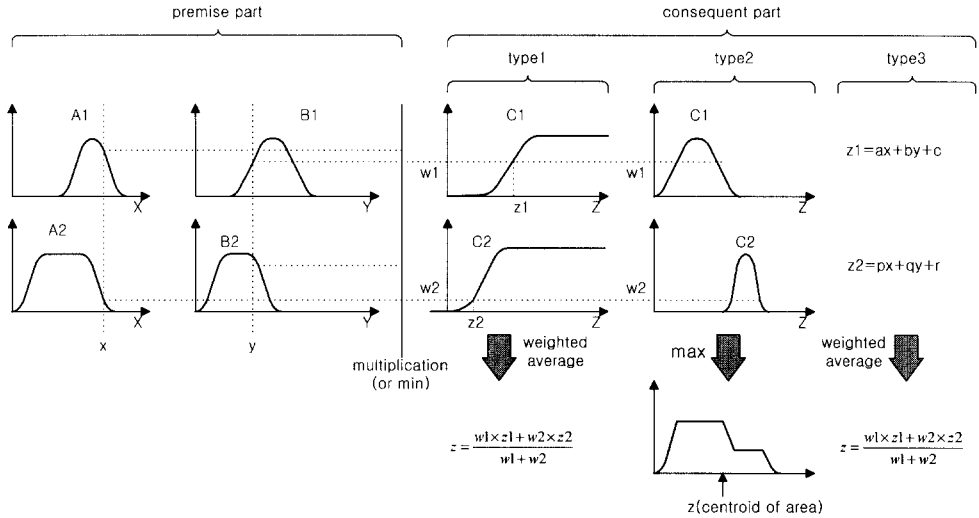


Fig. 3. Commonly used fuzzy if-then rules and fuzzy reasoning mechanisms

Table 3. Fuzzy Inference Systems

Contents	Step	Procedure
	Step 1	Fuzzify Inputs
	Step 2	Apply Fuzzy Operator
	Step 3	Apply Implication Method
	Step 4	Aggregate All Outputs
	Step 5	Defuzzify

표현한 것이다.

탈퍼지화 기법에는 여러 가지 방법들이 제안되고 있으나, 일반적으로 높이법, 면적법이 중심법보다 양호한 결과를 얻는다. 본 연구에서는 탈퍼지화 기법으로 Eqn. (1)의 높이법이 사용되고 있다. 이 방법은 Sugeno 퍼지 추론법의 특별한 경우이다. 즉, $f_i(x_0, y_0) = z_i$ (정수)인 경우이다(Kaufmann과 Gupta, 1985).

$$Z_0 = \frac{Z_1 \cdot W_1 + Z_2 \cdot W_2 + \dots + Z_n \cdot W_n}{W_1 + W_2 + \dots + W_n} \quad (1)$$

여기서 Z_i 는 대표점, W_i 는 퍼지집합의 높이를 의미한다.

3.2 ANFIS의 구조와 학습알고리즘

신경망과 퍼지 논리의 장점을 최대한 살려 이용하고, 각각의 기술이 갖고 있는 단점을 보완하고자 하는 시도로서 여러 가지 형태의 뉴로-퍼지 회로망이 소개되

고 있으며, 또한 여러 가지 응용 분야에서 활발히 이용되고 있다(권순학, 1995).

퍼지이론과 신경망의 외형적인 유사성은 다음과 같다. 첫째 퍼지 추론규칙 조건부(IF 부분)의 퍼지변수와 입력간의 Min 연산은, 뉴런의 입력과 연결선에 부여된 가중치의 곱셈 연산에 대응된다. 둘째 여러 추론규칙의 결론부(THEN 부분)에 대한 Max 연산은, 뉴런의 입력과 가중치를 곱한 것을 합하는 연산과 대응된다. 또한 뉴런의 sigmoid 함수와 같은 전달함수의 모양은 가능성 분포(possibility distribution) 함수의 형태에 관련이 있을 수 있다(강근택, 1995).

ANFIS는 신경망과 퍼지이론을 동시에 가장 쉽게 이용하는 방법으로 신경망을 전처리기로 사용하고 나중에 퍼지처리(fuzzy processing)를 하는 것이다. 이러한 입장에서 패턴처리 등과 같은 일은 신경망을 이용하여 처리하고, 추론 등의 작업은 퍼지논리를 이용한다. 퍼지논리에서는 퍼지변수 또는 언어항(linguistic term)을 포함하는 퍼지추론 규칙을 사용하여 시스템을 설계하거나 기술한다. 그런데 이렇게 기술된 퍼지규칙들이 시스템을 제대로 반영하도록 하려면 규칙에 사용된 퍼지변수의 소속함수 모양 등을 조정해야 한다. 이러한 조정 작업은 시간이 오래 걸리고 지루한 작업이다. 그래서 이러한 소속함수 결정에 신경망의 학습기능을 이용한다(Jang, 1993).

적응식 신경망의 기본적인 학습규칙은 경사하강추적법(gradient descent method)과 연쇄 규칙(chain

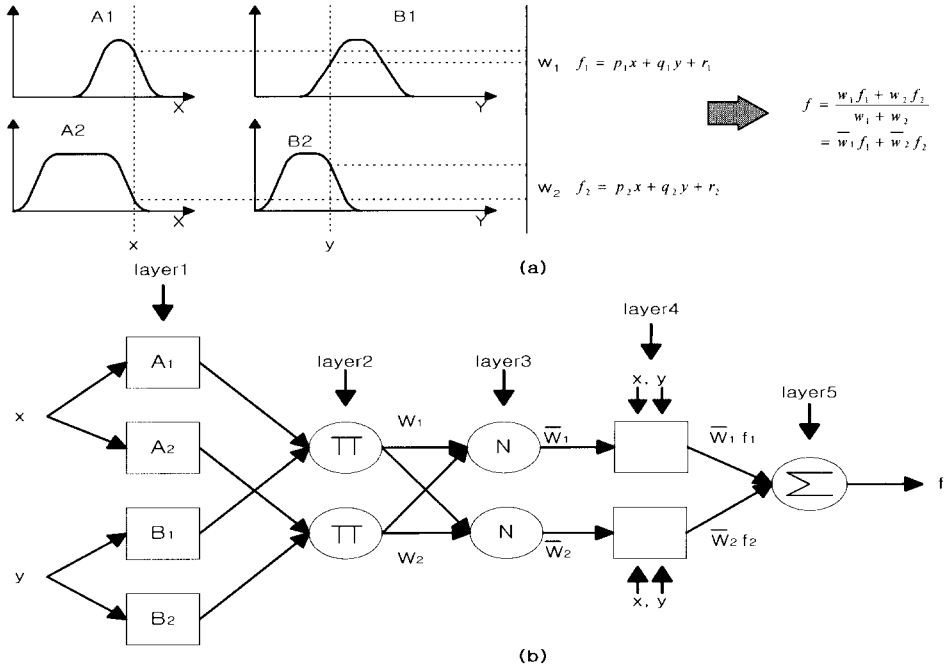


Fig. 4. (a) Type-3 fuzzy reasoning; (b) equivalent ANFIS (type-3 ANFIS)

rule)이다. 신경망은 여러개의 층으로 이루어진 비선형 네트워크이며 각 층은 여러개의 처리소자로 구성된다. 각 층의 처리소자는 서로 연결되어 있고 영향력의 크기를 나타내는 연결강도를 최적화하여 동작한다. 따라서 각 층에 있는 처리소자간의 최적 매개변수를 예측하여 이것을 신경망의 연결강도로 사용한다(Patterson, 1996).

경사하강추적법은 매개변수를 최적화하는 탐색기법으로 1차 도함수만을 사용한다. 이 알고리즘은 원하는 목표값(d)과 실제 출력값(o) 사이의 오차제곱합으로 정의된 비용함수(cost function) E의 값을 경사하강추적법에 의해 최소화하는 방향으로 학습한다(도용태 등, 1997).

$$E = \sum_p E_p, \quad (E_p = \frac{1}{2} \sum_j (d_{pj} - o_{pj})^2) \quad (2)$$

여기서 p는 p번째 학습패턴을, E_p 는 p번째 패턴에 대한 오차이고, d_{pj} 는 p패턴에 대한 목표 출력의 j번째 요소를, o_{pj} 는 실제 출력의 j번째 요소를 의미한다.

신경망을 이용하여 소속함수를 결정할 때에는 다음과 같은 장점이 있다. 첫째, 사용자가 직접 개입되지

않고 소속함수에 대한 학습이 알고리즘적으로 이루어지기 때문에 시스템의 개발 시간이 단축될 수 있다. 둘째, 신경망이 비선형을 가지기 때문에, 비선형인 소속함수를 결정할 수 있다. 셋째, 신경망의 학습기능을 이용하여 전문가로부터 규칙을 자동으로 획득할 수 있다. 넷째, 신경망의 학습기능에 의해 추론 환경에 능동적으로 적응할 수 있다(이광형과 오길록, 1992).

Fig. 4는 ANFIS모델에 적용된 Type-3의 추론과정과 구조를 보이고 있다. 추론과정은 전술한 3.1.1절과 동일하다. 또한, layer 1에선 입력변수의 퍼지화 과정으로 각각의 변수에 대해 소속함수를 결정한다. Eqn. (3)과 (4)는 각 노드에서의 소속함수에 관한 것이다.

$$O_i^1 = \mu_{A_i}(x) \quad (3)$$

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^b} \quad (4)$$

Eqn. (4)에서의 소속함수는 일반화된 종형 함수이다. 소속함수는 11가지의 종류가 있고, 간단한 형태의 소속함수로 삼각형과 사다리꼴형 함수가 사용되며 가장 많

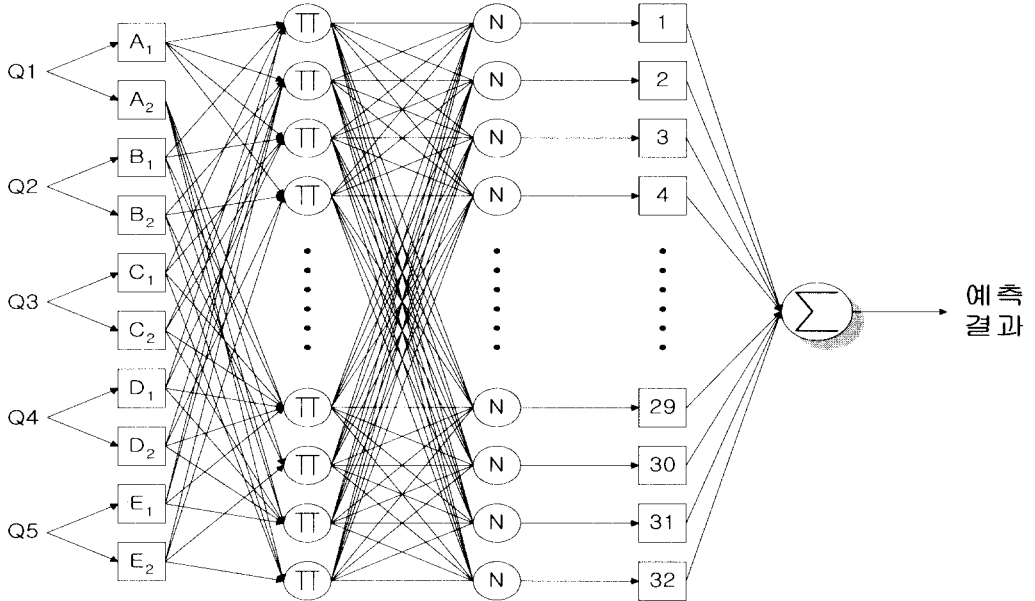


Fig. 5. The ANFIS architecture for Model 2a & Model 2
(The connections from inputs to layer 4 are not shown.)

이 쓰이는 소속함수로는 종형과 가우시안(Gaussian) 함수이다. 본 연구에서는 종형 함수가 쓰였고 각 변수에 대한 소속함수는 2개이다.

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i = 1, 2 \quad (5)$$

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad (6)$$

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (7)$$

$$O_i^5 = \sum_j \bar{w}_j f_j = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (8)$$

layer 1~5는 ANFIS 구조의 추론과정을 자세히 보여주고 있으며, 각 layer에서 사용되는 공식은 Eqn. (3)~(8)에 해당된다(Jang, 1993).

Fig. 5는 ANFIS모형 중 입력변수가 급수량(5차)으로 구성된 Model 2a 및 Model 2에 적용된 ANFIS구조로 각 입력변수를 2개의 소속함수로 표현하였고, 입력 값으로부터 layer 4로의 연결은 그리지 않았다. 또한 Fig. 5는 type-3의 ANFIS로 $2^5 = 32$

개의 규칙이 생성된다.

이러한 ANFIS모형은 수학적 모델을 만들지 않아도 되므로 급수량을 예측하는데 신속하게 대처할 수 있으며 소속 함수의 상세 설계를 통한 연구로 정도가 높은 예측을 할 수 있으리라 판단된다.

4. 분석결과 및 고찰

본 연구에서는 상수도 급수량의 ANFIS모형을 개발하기 위해 모형의 구성에 필요한 입력변수로 급수량, 평균기온, 급수인구를 통정하였다. 따라서 급수량 (Q_t, Q_{t-1}, Q_{t-2}), 급수인구 (P_{t+1}) 및 평균기온 (T_t, T_{t+1})으로 Model 1a와 Model 1을 구성하였으며, 기후자료를 신뢰할 수 없는 경우에도 급수량 자료 ($Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$)만을 가지고 Model 2a와 Model 2를 구성하였다. 따라서, 훈련자료의 입력 변수를 소속함수로 설계하고 훈련하여 검증자료로 모델의 예측치를 구할 수가 있다. 또한 t분포를 통해 모델의 유의한 정도를 확인 할 수가 있다.

Fig. 6~7은 초기 소속함수와 학습을 통해 얻어진 최종 소속함수를 나타낸다. 퍼지 논리를 이용할 때, 먼저 언어적인 변수에 의해 표현되는 입력측 변수를 소속함수들에 의해 표현해야 한다. 소속함수들은 주어진

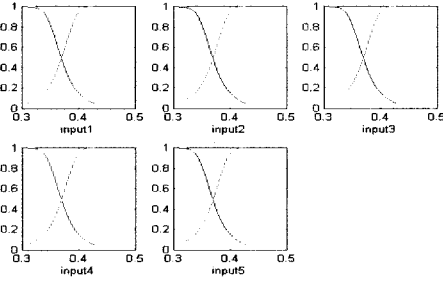


Fig. 6. Initial Membership Functions of Model 2a

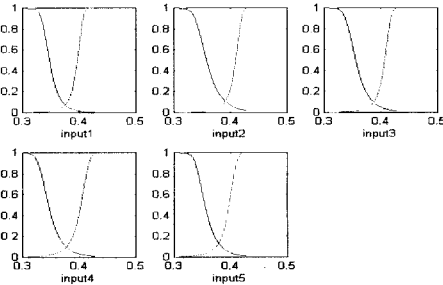


Fig. 7. Final Membership Functions of Model 2a

실제 데이터에 대한 퍼지 집합을 결정하는데 이용되며, 그 집합에서의 소속정도를 나타낸다. Fig. 6과 Fig. 7은 퍼지 변수 설계 과정을 자동적으로 할 수 있는 신경망의 학습능력을 이용하여 얻을 수 있다. 또한 각 모델에 대한 규칙 수는 2ⁿ 개가 되는데 Model 1a와 Model 1은 입력변수가 6개이므로 64개의 규칙이, Model 2a와 Model 2는 입력변수가 5개이므로 32개의 규칙이 만들어진다.

Fig. 8과 9는 각각 Model 2a와 Model 2의 RMSE(Root Mean Square Error)에 대한 그래프이며, 반복횟수에 따른 훈련자료와 검증자료의 RMSE의

변화를 도시한 것으로 적절한 반복횟수를 정할 수 있다. 검증자료의 RMSE변화가 어느 일정한 값을 가질 때 적절한 반복횟수로 정한다. 여기서 Model 2a와 Model 2의 적절한 반복횟수는 5, 27회이다. 이 반복횟수로 검증자료의 예측치를 구하였다.

급수량 예측 ANFIS모형의 적용타당성을 평가하기 위하여 각 모형에서 계산된 결과를 오차 및 통계적인 검증방법을 이용하여 검토하였다. 본 연구에서는 시계열 자료에 대한 적용타당성 문제를 다룰 때 가장 일반적으로 사용되고 있는 RMSE, MAE, 절대 오차율, 상관계수 및 결정계수를 사용하였다.

또한 1일 급수량 예측자료 집단과 실측자료 집단간의 동질성(homogeneity)은 비모수 검정인 2표본 가설 검정시험으로 조사할 수 있다. 두 모분산이 미지인데 표본규모가 작아 정규분포에 의한 근사치 계산이 곤란한 경우에는 t분포를 이용하여 검증할 수 있다. 본 논문에서는 95 % 유의수준에서 예측 급수량과 실측 급수량과의 통계적 동질성을 검증하였다.

Table 4는 제안된 ANFIS모형 및 다중회귀모형의 예측오차 및 오차율을 보이고 있다. ANFIS모형에서 가장 예측이 잘된 모델은 급수량 자료만을 가지고 모형을 구성한 Model 2a로 절대평균오차(MAE)가 9226 m³, 평균제곱근오차(RMSE)가 12933 m³, 최대오차율 18.46 %, 평균오차율이 2.36 %로 나타났다. 또한 Model 2도 Model 2a에는 못 미치지만 절대평균오차가 10860 m³, 평균제곱근오차가 14672 m³, 최대오차율 17.93 %, 평균오차율이 2.77 %로 회귀모형보다는 정확한 예측치를 보이고 있다. 또한 Model 2a 및 Model 2에서 최대오차율은 큰 값이지만 인위적인 조작(단수 등)이 가해졌기 때문에 큰 오차가 발생하고 있다고 생각된다. 그러나, 제안된 ANFIS모형 중 급수량, 평균기온, 급수인구로 구성된 Model 1a 및 Model 1은 회귀모형보다도 좋은 예측결과를 보이고 있으나, 회

Table 4. Errors of ANFIS and Regression Model Estimating the Daily Urban Water Demand for the Year of 1997

Model	Error	MAE (m ³)	RMSE (m ³)	Max Error(%)	Mean Error(%)	Remark
Model 1a		12435	14962	18.81	2.95	95~96
Model 1		12630	15182	17.99	3.25	96
Model 2a		9226	12933	18.46	2.36	95~96
Model 2		10860	14672	17.93	2.77	96
Multiple Regression		13093	16413	17.40	3.35	

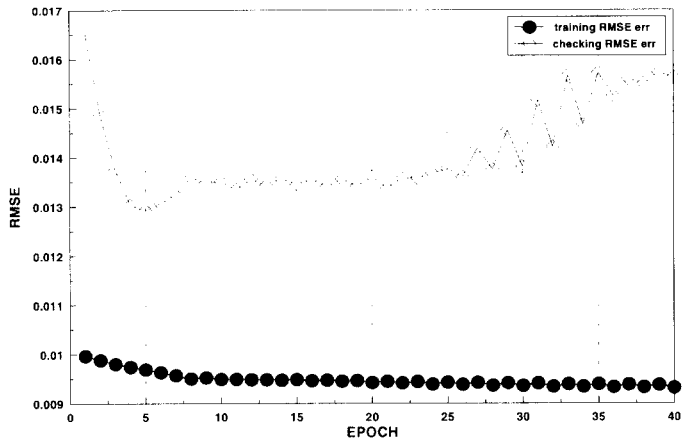


Fig. 8. RMSE of Model 2a

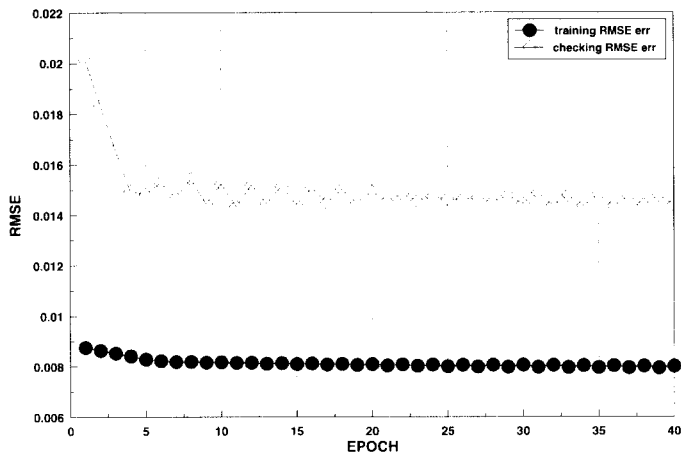


Fig. 9. RMSE of Model 2

귀모형과 차이가 거의 없어 1일 급수량의 예측에 적합치 않는 모델임을 알 수 있다.

예측결과와 실측자료의 상관계수 및 결정계수를 나타낸 것이다. 따라서 ANFIS모형 중 Model 2a가 상관계수 0.916, 결정계수 0.813으로 가장 좋은 결과를 보이고

Table 5는 제안된 ANFIS모형과 다중회귀모형의

Table 5. Coefficient of Correlation and Multiple Determination between the Estimation and the Actual Value

Coefficient \ Model	Correlation	Determination	Remark
Model 1a	0.849	0.716	'95~'96
Model 1	0.835	0.705	'96
Model 2a	0.916	0.813	'95~'96
Model 2	0.883	0.794	'96
Multiple Regression	0.830	0.673	

있고, 그 다음이 Model 2, Model 1a, Model 1 및 다중회귀모형 순으로 나타나 있다.

Table 6은 제안된 ANFIS모형 및 다중회귀모형에

대한 t검증 결과이다. 이 결과를 보면 모든 모형이 95 % 신뢰구간 내에서 귀무가설이 채택되어 실측자료와 예측결과가 통계적 유의성을 갖는다고 판단 할 수 있

Table 6. Analysis of T-test for ANFIS and Regression Model(1997)

Model	T-test		P	DF	Remark
	$ t $	$t_{\alpha/2}$			
Model 1a	1.48	1.96	0.31	712	Selection
Model 1	1.76	1.96	0.26	706	Selection
Model 2a	0.83	1.96	0.41	716	Selection
Model 2	0.57	1.96	0.57	715	Selection
Multiple Regression	0.94	1.96	0.35	714	Selection

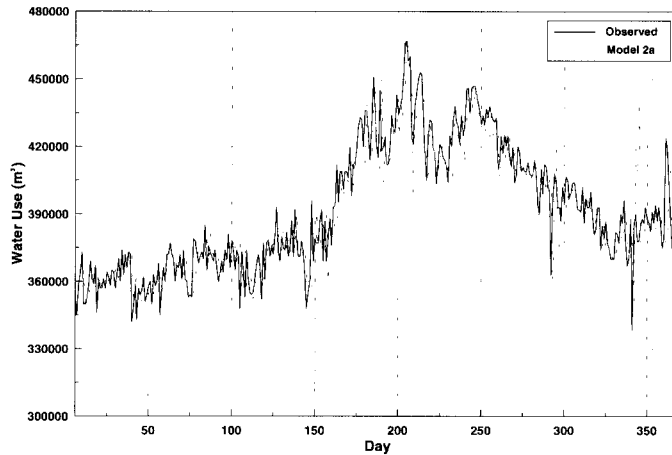


Fig. 10. Estimation of Daily Urban Water Demand Using ANFIS Model 2a, 1997

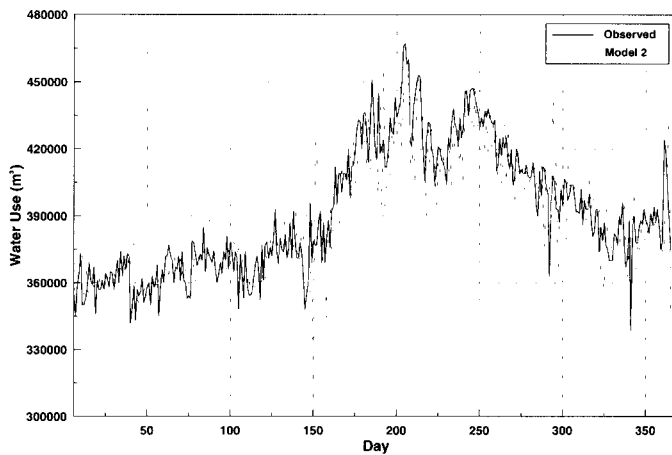


Fig. 11. Estimation of Daily Urban Water Demand Using ANFIS Model 2, 1997

다.

위의 분석을 통해 제안된 ANFIS모형 중 상수도 1일 급수량 예측에는 Model 2a와 Model 2가 가장 적합한 모델임을 알 수 있다.

Fig. 10과 Fig. 11은 Model 2a와 Model 2의 검증자료인 97년의 실측치와 예측치를 나타낸 것으로 인위적인 조작(단수)등이 가해지는 6월 초순, 10월 중순 및 12월 중순경의 시기를 제외하면 높은 예측력을 보이고 있다.

이상에서 살펴보았듯이 제안된 ANFIS모형 중 양호한 예측력을 보인 Model 2a와 Model 2도 급수량에 급격한 변화를 일으키는 급수관의 파열, 상수도시설의 개수, 청소 등으로 인한 단수 조치등이 가해지면 정확한 예측에 다소 무리가 있는 것으로 판단되었다. 따라서 소속함수의 상세 설계(소속함수의 개수 증가 및 학습인자의 조정)를 통한 ANFIS연구가 필요하며, 보다 나은 인공지능의 실현을 위해 신경망, 퍼지 논리, 전문가 시스템 및 유전 알고리즘에 대한 추가적인 연구가 필요하다고 생각된다.

5. 결 론

본 연구에서는 상수도 1일 급수량 예측을 위하여 ANFIS모델을 구성하고, 대상지역으로 광주광역시를 선정하여 조사한 결과 다음과 같은 결론을 얻었다.

(1) 시계열장이 1995~1996년인 Model 2a의 최대 오차율 및 평균오차율은 각각 18.46 % 및 2.36 %이고, 시계열장이 1996년인 Model 2의 최대오차율 및 평균오차율은 각각 17.93 % 및 2.77 %를 보이고 있어 Model 2a쪽이 1일 급수량 예측에 다소 적합함을 알 수 있다. 또한, Model 2a와 Model 2의 상관계수와 결정계수를 살펴보면 각각 0.916, 0.813 및 0.883, 0.794로 Model 2a가 Model 2 보다 높은 예측력을 보임을 알 수 있다.

(2) 시계열장의 길이에 따른 모델(Model 1a와 Model 1 및 Model 2a와 Model 2)의 예측력을 비교해 보면, 96년으로 구성된 모델보다 95~96년으로 구성된 모델이 다소 좋은 결과를 보인다. 이와 같은 결과는 적정변수의 수보다 훈련자료의 수가 몇 배 정도 커야만 좋은 성과를 얻는데서 기인한다고 판단된다.

(3) 제안된 ANFIS모형과 다중회귀모형을 예측오차 및 오차율 또한 상관계수와 결정계수를 통해 비교·검토한 결과 ANFIS모형 중 Model 2a와 Model 2는 다중회귀모형보다 훨씬 월등한 결과를, Model 1a 및

Model 1은 다중회귀모형보다 다소 나은 결과를 보이므로 1일 급수량 예측을 위한 ANFIS모형에는 급수량만으로 구성된 Model 2a 및 Model 2가 급수량, 급수인구 및 평균기온으로 구성된 Model 1a 및 Model 1보다 적합함을 알 수 있다.

(4) 제안된 ANFIS모델에서도 일 급수량의 변화가 심한 곳에서는 예측치가 실측치를 따라가지 못함을 알 수 있었다. 이러한 문제를 해결하기 위해 소속함수의 상세 설계(멤버십 함수와 학습인자 조정)를 통한 ANFIS연구로 정도가 높은 예측을 할 수 있으리라 판단하며, 이에 대한 연구가 활발해져야 할 것으로 생각된다.

참 고 문 헌

- 강근택 (1995). "퍼지제어 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 74-82.
- 권순환 (1995). "퍼지집합, 퍼지척도 및 퍼지적분 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 39-51.
- 공성곤, 김민수 (1995). "적용 퍼지제어 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 101-108.
- 도용태, 김일곤, 김종완, 박창현 (1997). 인공지능 개념 및 응용. 희중당, pp. 217-246.
- 문병석 (1998). 신경망 이론을 이용한 1일 급수량 예측. 박사학위 논문, 전남대학교, pp. 50-67.
- 이광형, 오길록 (1992). 퍼지 이론 및 응용. 홍릉과학출판사, pp. 61-77.
- 이경훈, 박성진, 문병석 (1994a). "상수도의 급수량 및 배수지 용량변화에 관한 연구." 대한상하수도학회지, 제9권, 제2호, pp. 118-126.
- 이경훈, 이삼노, 문병석 (1994b). "상수도 1일 급수량 변화와 확률분포 모형에 관한 연구." 제36회 수공학연구발표회 논문집, 한국수문학회, pp. 405-410.
- 이경훈, 이삼노, 문병석 (1994c). "상수도 1일 급수량의 시간적 변화의 특성에 관한 연구." 한국수문학회지, 제27권, 제2호, pp. 135-143.
- Jang J.-S.R. (1993). "ANFIS: adaptive network-based fuzzy inference system."

- IEEE Trans. on System, Man and Cybernetics*, Vol. 23, No 3, pp. 665-685.
- Kaufmann, A. and Gupta M.M. (1985). *Introduction to fuzzy arithmetic*. Van Nostrand Reinhold Company, pp. 9-14.
- Patterson, D.W. (1996). *Artificial neural networks theory and Application*. Prentice Hall, pp. 20-36.
- Young, R.A. (1973). "Price elasticity of demand for municipal water: A case study of Tuscon, Arizona." *Water Resources Research*, Vol. 9, No. 4, pp. 1068-1072.
- (논문번호:98-043/접수:1998.07.23/심사완료:1998.11.27)