

지능형 모델링(ANFIS)을 이용한 상수도 1일 급수량 예측

이경훈¹⁾ · 김강석²⁾ · ○강일환³⁾

1. 서론

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. 퍼지 뉴런이란 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.

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. ANFIS의 구조와 학습알고리즘

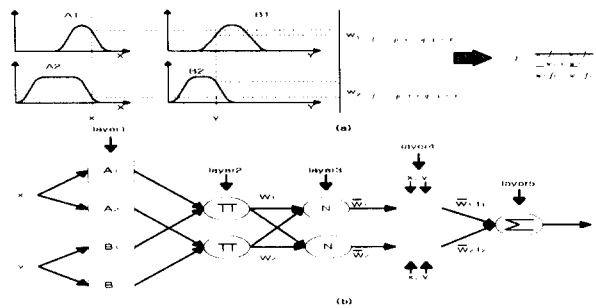


그림 1. (a) Type-3 퍼지 추론; (b) 동가 ANFIS (type-3 ANFIS)

그림 1은 ANFIS모델에 적용된 Type-3의 추론과정과 구조를 보이고 있다. 또한, layer 1에선 입력변수의 퍼지화 과정으로 각각의 변수에 대해 소속함수를 결정한다.

소속함수는 11가지의 종류가 있고, 간단한 형태의 소속함수로 삼각형과 사다리꼴형 함수가 사용되며 가장 많이 쓰이는 소속함수로는 종형과 가우시안(Gaussian) 함수이다. 본 연구에서는 종형

- 1) 전남대학교 공과대학 토목공학과 부교수
- 2) 성화대학 토목과 전임강사
- 3) 전남대학교 공과대학 토목공학과 박사과정

함수가 쓰였고 각 변수에 대한 소속함수는 2개이다.

layer 1~5는 ANFIS 구조의 추론과정을 자세히 보여주고 있다.(Jang, 1993).

3. 모델의 구성

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.

이상에서 살펴보았듯이 광주광역시 1일 급수량 (Q_{t+1}) 예측을 위한 입력변수는 급수인구 (P), 평균기온 (T)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.

표 1. 입력변수와 데이터 범위에 따른 모델의 구성

Model	Range	Input variable	Training data	Checking data
M2		$Q_t, Q_{t-1}, Q_{t-2}, P_{t+1}, T_t, T_{t+1}$	95~96	97
M1		$Q_t, Q_{t-1}, Q_{t-2}, P_{t+1}, T_t, T_{t+1}$	96	97
W2		$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$	95~96	97
W1		$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$	96	97

표 1은 입력변수와 데이터 범위에 따른 ANFIS모델의 구성을 보여준다. 입력변수는 동정된 급수량, 평균기온, 급수인구와 급수량만으로 되어 있고, 데이터의 범위는 훈련자료로 95~96년과 96년을 사용하였고 검증자료로 97년을 사용하였다. 따라서 원래의 모델에 시계열장을 달리하여 모델을 구성하였다.

4. 분석결과 및 고찰

4.1 모델의 비교·분석

훈련자료의 입력변수를 소속함수로 설계하고 훈련하여 검증자료로 모델의 예측치를 구할 수가 있다. 또한 t분포를 통해 모델의 유의한 정도를 확인 할 수가 있다.

그림 2~3은 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 실제 데이터에 대한 퍼지 집합을 결정하는데 이용되며, 그 집합에서의 소속 정도를 나타낸다. 그림 2~3은 퍼지 변수 설계 과정을 자동적으로 할 수 있는 신경망의 학습능력을 이용하여 얻을 수 있다.

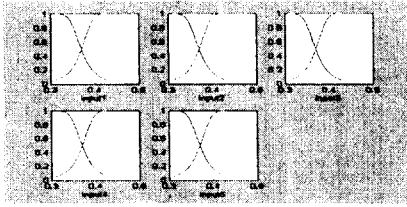


그림 2. W2의 초기 소속함수

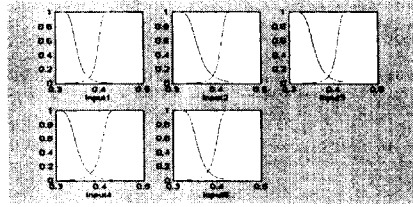


그림 3. W2의 최종 소속함수

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.

표 2. ANFIS와 다중회귀모델의 오차 및 오차율 비교(1997년 1일 급수량)

Model	Error	MAE(m ³)	RMSE(m ³)	Max Error(%)	Mean Error(%)	Remark
M2		12435	14962	18.81	2.95	95~96
M1		12630	15182	17.99	3.25	96
W2		9226	12933	18.46	2.36	95~96
W1		10860	14672	17.93	2.77	96
Multiple Regression		13093	16413	17.40	3.35	

표 3. ANFIS모델 및 다중회귀모델에 대한 실측치와 예측결과의 상관계수 및 결정계수

Model	Coefficient	Correlation	Determination	Remark
M2		0.849	0.716	95~96
M1		0.835	0.705	96
W2		0.916	0.813	95~96
W1		0.883	0.794	96
Multiple Regression		0.830	0.673	

표 2와 표 3은 제안된 ANFIS모형과 다중회귀모형간의 오차 및 오차율 그리고 상관계수 및 결정계수를 나타내고 있다.

위의 분석을 통해 제안된 ANFIS모형 중 상수도 1일 급수량 예측에는 W2와 W1이 가장 적합한 모델임을 알 수 있다.

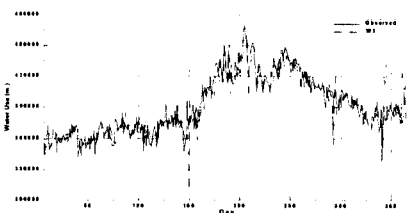


그림 4. W2에 의한 1997년 예측결과

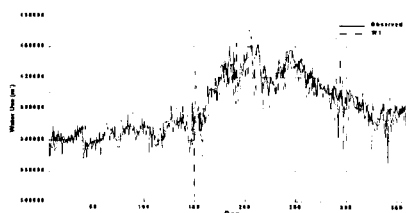


그림 5. W1에 의한 1997년 예측결과

그림 4와 그림 5는 W2와 W1의 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 W2와 W1도 급수량에 급격한 변화를 일으키는 급수관의 파열, 상수도시설의 개수, 청소 등으로 인한 단수 조치들이 가해지면 정확한 예측에 다소 무리가 있는 것으로 판단되었다.

4.2 상세 설계에 의한 모델의 재구성

1일 급수량의 예측에 M1, M2모델보다 적절하다고 판단된 W1과 W2모델의 상세한 설계를 하므로써 보다 나은 결과를 얻고자 모델을 재구성하였다.

표 4. W1모델을 상세 설계한 모델(훈련자료:1996)

모델명	소속함수	학습인자	비고
W1GBDF	일반화된 종형	default	W1GB
W1GBS1	일반화된 종형	Step Size(0.1)	W1GBS
W1GBS2	일반화된 종형	Step Size(0.01)	W1GBSS
W1GSDF	가우스	default	W1G
W1GSS1	가우스	Step Size(0.1)	W1GS
W1GSS2	가우스	Step Size(0.01)	W1GSS
W1TRDF	삼각형	default	W1T
W1TRS1	삼각형	Step Size(0.1)	W1TS
W1TRS2	삼각형	Step Size(0.01)	W1TSS

표 5. W2모델을 상세 설계한 모델(훈련자료:1995~1996)

모델명	소속함수	학습인자	비고
W2GBDF	일반화된 종형	default	W2GB
W2GBS1	일반화된 종형	Step Size(0.1)	W2GBS
W2GBS2	일반화된 종형	Step Size(0.01)	W2GBSS
W2GSDF	가우스	default	W2G
W2GSS1	가우스	Step Size(0.1)	W2GS
W2GSS2	가우스	Step Size(0.01)	W2GSS
W2TRDF	삼각형	default	W2T
W2TRS1	삼각형	Step Size(0.1)	W2TS
W2TRS2	삼각형	Step Size(0.01)	W2TSS

표 4와 표 5는 급수량만으로 구성된 W1과 W2모델을 상세 설계한 모델로, 주어진 소속함수와 Step Size의 값을 재조정된 모델이다. 소속함수로는 주어진 함수인 종형함수 및 가우스, 삼각형 함수를 사용하였으며, Step Size의 증가율과 감소율은 주어진 값인 1.1, 0.9 및 0.1, 0.1 그리고 0.01, 0.01을 사용하였다.

4.3 재구성된 각 모델의 비교·분석

이번 절에서는 W1과 W2모델을 상세 설계한 모델들을 비교·분석해 보기로 한다. W1과 W2 모델을 상세 설계한 모델의 예측결과와 실측치간의 오차 및 오차율, 상관계수 및 결정계수, t검증을 통해 모델을 분석하고 유의정도를 알아보았다.

먼저 훈련자료가 1996년인 W1모델을 상세 설계한 모델들의 예측오차 및 오차율을 살펴보면 표 6과 같다.

표 6의 모델에서 알 수 있듯이 W1모델을 상세 설계한 모델에서는 입력측 변수가 종형함수이고, 학습인자가 조정된 W1GBS1과 W1GBS2모델이 절대평균오차 10774m³, 평균제곱근오차 14627m³, 최대오차율 16.31%, 평균오차율 2.74%로 가장 양호한 결과를 보이고 있다. 반면에 입력측 변수가 가우스 함수나 삼각 함수인 모델에서는 최대오차율이 낮게 나타남을 알 수 있다.

최종적으로 W1모델과 W1모델을 상세 설계한 모델을 비교하면 W1GBS1 및 W1GBS2모델쪽이 절대평균오차 10774m³, 평균제곱근오차 14627m³, 최대오차율 16.31%, 평균오차율 2.74%로 W1모델의 절대평균오차 10860m³, 평균제곱근오차 14672m³, 최대오차율 17.93%, 평균오차율 2.77%보다 모든 면에서 양호한 결과를 갖고 있음을 알 수 있다.

이상에서 살펴본 오차 및 오차율, 상관계수 및 결정계수, t검증결과 등을 고려하여 종합적으로 판단해 보면, W1모델을 상세 설계한 모델중에서는 W1GBS1 및 W1GBS2모델이 가장 양호한 예측결과를

표 6. W1모델을 상세 설계한 모델의 예측오차 및 오차율

모델 \ 오차	절대평균 오차(m ³)	평균제곱근 오차(m ³)	최대 오차율(%)	평균 오차율(%)
W1GBDF	10863	14773	19.55	2.76
W1GBS1	10774	14627	16.31	2.74
W1GBS2	10774	14627	16.31	2.74
W1GSDF	10970	14841	15.72	2.79
W1GSS1	10970	14841	15.72	2.79
W1GSS2	10970	14841	15.72	2.79
W1TRDF	10970	14841	15.72	2.79
W1TRS1	10970	14841	15.72	2.79
W1TRS2	10970	14841	15.72	2.79

표 7. W2모델을 상세 설계한 모델의 예측오차 및 오차율

모델 \ 오차	절대평균 오차(m ³)	평균제곱근 오차(m ³)	최대 오차율(%)	평균 오차율(%)
W2GBDF	9226	12934	18.47	2.36
W2GBS1	9226	12934	18.47	2.36
W2GBS2	9226	12934	18.47	2.36
W2GSDF	9531	13189	21.64	2.43
W2GSS1	9658	13138	18.50	2.45
W2GSS2	9675	13146	18.18	2.46
W2TRDF	9531	13189	21.64	2.43
W2TRS1	9658	13138	18.50	2.45
W2TRS2	9675	13146	18.18	2.46

보였고 또한 1일 급수량 예측을 위한 모델에서도 W1모델보다 뛰어난 예측력을 보이고 있음을 알 수 있다.

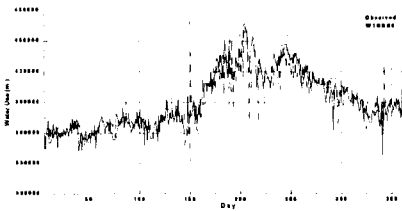


그림 6. W1GBS1모델에 의한 1997년 예측결과

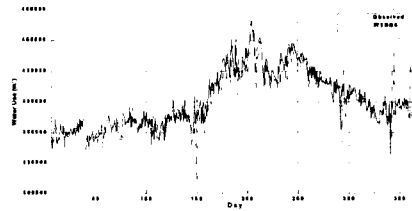


그림 7. W2GBS1모델에 의한 1997년 예측결과

그림 6은 W1모델을 상세 설계한 모델 중 가장 양호한 예측결과를 보이는 W1GBS1모델로 입력변수의 소속함수는 종형함수이고 Step Size는 조정된 값인 0.1, 0.1이며, 반복횟수 5인 모델이다. 위 예측결과에서 알 수 있듯이 6월 초순이나 12월 중순경의 일 급수량이 갑자기 변하는 곳에서는 많은 오차가 발생하고 있다. 이와 같이 인위적인 조치로 인하여 1일 급수량의 변화가 심한 곳에서는 예측치가 실측치를 따라가지 못하고 있음을 알 수 있다.

또한, 훈련자료가 1995~1996년인 W2모델을 상세 설계한 모델들의 예측오차 및 오차율을 살펴 보면 표 7과 같다.

표 7의 모델에서 알 수 있듯이 W2모델을 상세 설계한 모델에서는 입력측 변수로 종형함수를 갖는 W2GBDF, W2GBS1, W2GBS2모델이 동일한 결과치인 절대평균오차 9226m³, 평균제곱근오차 12934 m³, 최대오차율 18.47%, 평균오차율 2.36%로 가장 양호한 결과를 보이고 있다.

최종적으로 W2모델과 W2모델을 상세 설계한 모델을 비교하면 W2GBDF, W2GBS1 및 W2GBS2모델은 W2모델의 절대평균오차 9226m³, 평균제곱근오차 12933m³, 최대오차율 18.46%, 평균오차율 2.36%와 거의 같은 값을 가짐을 알 수 있다.

이상에서 살펴본 오차 및 오차율, 상관계수 및 결정계수, t검증결과 등을 고려하여 종합적으로 판단해 보면, W2모델을 상세 설계한 모델 중에서는 W2GBDF, W2GBS1 및 W2GBS2모델이 가장 양호한 예측결과를 보이고 있음을 알 수 있다. 하지만 W2모델과 거의 유사한 예측결과를 가지고 있으므로 1일 급수량 예측을 위한 모델은 주어진 값을 사용하는 W2모델이 바람직하다고 판단된다.

그림 7은 W2모델을 상세 설계한 모델 중 가장 양호한 예측결과를 보이는 W2GBS1모델로 입력변수의 소속함수는 종형함수이고 Step Size는 조정된 값인 0.1, 0.1이며, 반복횟수 5인 모델이다.

위 예측결과에서 알 수 있듯이 일 급수량이 갑자기 변하는 곳에서는 많은 오차가 발생하고 있다. 즉, 6월 초순 및 12월 중순경에 인위적인 조치로 인하여 1일 급수량의 변화가 심한 곳에서는 예측치가 실측치를 따라가지 못하고 있음을 알 수 있다.

5. 결 론

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

모든 결과들을 종합해 보면 급수량, 평균기온, 급수인구로 구성된 M1, M2모델보다 급수인구만으로 구성된 W1, W2모델이 1일 급수량의 예측에 적합한 모델임을 알 수 있었다. 그리고 훈련자료가 1996년인 W1모델보다는 훈련자료가 1995~1996년인 W2모델이 훨씬 양호한 모델임을 알 수 있다.

또한 W1, W2모델의 상세 설계를 통해 입력변수의 소속함수를 gbell(Generalized Bell)함수로 하는 것이 gauss나 tri함수로 하는 것 보다 양호한 결과를 가져옴을 알 수 있었다. 그리고 W1모델과 W1모델을 상세 설계한 W1GBS1 및 W1GBS2모델을 비교한 결과 W1모델을 상세 설계한쪽이 더욱 양호한 예측결과를 갖는 모델임을 알았고, W2모델은 W2모델을 상세 설계한 모델과 거의 비슷한 예측결과를 갖고 있어 주어진 값을 모델로 한 W2모델이 1일 급수량 예측에 훨씬 적합함을 알 수 있다.

전체적인 관점에서 보면 1일 급수량의 예측에는, 훈련자료가 1995~1996년이고 입력측 변수의 소속함수는 일반화된 종형인 W2모델이 적합함을 알 수 있다.

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.

1일 급수량에 급격한 변화를 일으키는 급수관의 파열, 상수도시설의 개수, 청소 등으로 인한 단수 조치들이 가해지면 본 연구에서 연구한 ANFIS모델로도 정확한 예측에 상당한 무리가 있는 것으로 판단된다. 이러한 문제점을 해결할 수 있는 방안으로, 보다 효과적인 모델 개발의 관점에서 신경망, 퍼지 논리, 전문가 시스템 및 유전 알고리즘을 보다 효과적으로 이용하고 통합해야 가능할 것으로 기대된다.

참 고 문 헌

1. 강근택 (1995). "퍼지제어 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 74-82.
2. 권순학 (1995). "퍼지집합, 퍼지척도 및 퍼지적분 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 39-51.
3. 공성곤, 김민수 (1995). "적응 퍼지제어 : 뉴로 퍼지 제어 기술특집." 제어·자동화·시스템공학회지, 제어·자동화·시스템공학회, 제1권, 제3호, pp. 101-108.
4. Dan W. Patterson (1996). "Artificial neural Networks Theory and Application." Prentice Hall, pp. 20-36.
5. Jyh-shing Roger Jang (1993). "ANFIS: Adaptive-Neural-Network-Based Fuzzy Inference system." IEEE Trans. on System, Man and Cybernetics, vol. 23, no 3, pp. 665-685.
6. L. A. Zadeh, "Fuzzy Sets", Information and Control, Vol. 8, 1965. pp. 338-353.