

자료 지향형 수위예측 모형의 비교 분석

최 승 용* / 한 건 연** / 최 현 구***†

Comparison and analysis of data-derived stage prediction models

Choi, Seung Yong* / Han, Kun Yeun** / Choi, Hyun-Gu***†

요약 : 수위예측을 위해 개념적, 물리적 모형들을 포함한 다양한 유형의 기법들이 사용되고 있다. 그럼에도 불구하고 이러한 기법들 중 수위예측을 위해 단일의 우수한 모형을 선정하는 것은 매우 어려운 일이다. 최근에는 수문학적 과정의 복잡성으로 인해 기존 물리적 기반의 강우-유출 모형이 가지고 있는 단점들을 극복하고자 자료 지향형 수위예측 모형이 널리 도입되고 있다. 본 연구의 목적은 이러한 자료 지향형 모형 중 뉴로-퍼지와 회귀분석 모형의 수위예측에 대한 성능을 비교하는 것이다. 제안된 두 모형을 한강수계의 왕숙천에 대해 적용하였다. 제안된 두 모형의 성능을 평가하기 위해 평균제곱근오차, Nash-Sutcliffe 효율계수, 평균절대오차, 수정 결정계수와 같이 4개의 통계지표들을 사용하였다. 모의결과 뉴로-퍼지 수위예측 모형이 다중선형회귀 수위예측 모형보다 좀 더 나은 예측 결과를 나타내는 것을 확인할 수 있었다. 본 연구결과는 향후 중소하천에서 충분한 선행시간을 확보한 정확도 높은 홍수정보시스템의 구축에 활용할 수 있을 것으로 판단된다.

핵심용어 : 자료 지향형 모형, 뉴로-퍼지, 회귀분석, 수위예측

Abstract : Different types of schemes have been used in stage prediction involving conceptual and physical models. Nevertheless, none of these schemes can be considered as a single superior model. To overcome disadvantages of existing physics based rainfall-runoff models for stage predicting because of the complexity of the hydrological process, recently the data-derived models has been widely adopted for predicting flood stage. The objective of this study is to evaluate model performance for stage prediction of the Neuro-Fuzzy and regression analysis stage prediction models in these data-derived methods. The proposed models are applied to the Wangsukcheon in Han river watershed. To evaluate the performance of the proposed models, fours statistical indices were used, namely; Root mean square error(RMSE), Nash Sutcliffe efficiency coefficient(NSEC), mean absolute error(MAE), adjusted coefficient of determination(R^{*2}). The results show that the Neuro-Fuzzy stage prediction model can carry out the river flood stage prediction more accurately than the regression analysis stage prediction model. This study can greatly contribute to the construction of a high accuracy flood information system that secure lead time in medium and small streams.

Keywords : Data-derived model, Neuro-Fuzzy, regression analysis, stage prediction

† Corresponding author : honorguard99@msn.com

* 정회원 · 국립방재연구원 책임연구원

Principal Researcher, National Disaster Management Institute. Seoul. 121-719, Korea

(e-mail : ecofriend@korea.kr)

** 정회원 · 경북대학교 공과대학 건축·토목공학부 교수

Professor, School of Archi. & Civil Engineering, Kyungpook National Univ., Daegu, 702-701, Korea

(e-mail : kshanj@knu.ac.kr)

*** 비회원 · 교신저자 · 경북대학교 공과대학 건축토목공학부 박사과정

Corresponding author, Candidate of Ph.D., School of Archi. & Civil Engineering, Kyungpook National Univ., Daegu, 702-701, Korea(e-mail : honorguard99@msn.com)

1. 서 론

실질적인 홍수피해를 줄이기 위해서는 하도 정비 및 개수, 홍수 조절지 및 우수지 계획 등의 구조적 대책뿐만 아니라 미래에 발생할 홍수를 사전에 예측할 수 있는 홍수예보 시스템을 개발하는 비구조적 대책의 수립도 절실히 필요하다. 이러한 홍수예보 시스템의 구축을 위해서는 정확한 수위 예측이 필수적이다.

최근 수위 예측에 관한 방법으로 자료 지향형 기반의 모형이 널리 도입되고 있다. 이러한 이유는 기존 수문학적, 물리적 기반의 강우-유출 모형의 경우 많은 양의 자료를 필요로 하고 또한 모형의 검보정을 위한 시간과 노력(VanderKwaak과 Loague, 2001)과 더불어 모형화에 대한 물리적 과정에 대한 이해 등을 요구하기 때문이다(Kokkonen과 Jakeman, 2001). 최근 널리 사용되고 있는 자료 지향형 모형의 기법으로는 회귀모형, 인공신경망, 퍼지추론 시스템, 뉴로-퍼지 등이 있다. 이러한 모형들은 기존의 수문학적, 물리적 기반의 모형과 비교해서 비선형(nonlinear)이거나 동적인 모형의 잡음자료(noisy data)를 처리하는데 용이하며, 모형을 구축하는데 있어서 복잡한 수학적 지식이 필요하지 않기 때문에 상대적으로 모형개발을 위한 절차가 간단하고, 결과 값을 얻기까지의 계산시간 역시 많이 소요되지 않는 장점을 가지고 있다. 또한 자료 지향형 기반의 모형들은 앞선 언급한 모형에 비해 적은 자료를 필요로 함에도 불구하고 수위예측에 있어서 정확도 높은 예측과 적은 계산소요 시간으로 인해 실시간 예보를 위한 적용성이 우수하다(Govindaraju, 2000).

이러한 자료 지향형 모형 중 뉴로-퍼지를 이용한 국외 연구로는 Gautam과 Holz(2001)가 적응형 뉴로-퍼지 기법을 강우-유출해석에 적용하였고, Chang과 Chen(2001)이 Counter Propagation 신경망과 퍼지기법을 융합하여 Counter Propagation Neural Network(CFNN)을 구성하고 하천의 전 시간 유량과 강우관측소의 우량자료를 이용하여

수위표지점에서의 홍수위를 예측하였다. Nayak 등(2005)은 3년간의 호우자료를 이용하여 뉴로-퍼지, 인공신경망, 퍼지 추론시스템을 이용하여 유량을 예측하였다. Aquil 등(2007)은 유량예측을 위해 Levenberg-Marquardt Feed Forward Neural Network(FFNN), Bayesian Regularization-FFNN, 뉴로-퍼지 모형을 비교한 결과 뉴로-퍼지 모형 사용 결과가 가장 우수하다고 제안하였다. Shu와 Ouarda(2008)는 적응형 뉴로-퍼지 추론 시스템을 이용한 미계측 유역에서의 빈도해석을 실시하였고, Kurtulus와 Razack(2009)은 유출 응답에 대한 신경망과 뉴로-퍼지 모형간의 비교 연구를 수행하였으며 Firat 등(2009)는 퍼지 추론 기법간의 비교 연구를 수행하였다.

뉴로-퍼지에 대한 국내 연구로는 진영훈(2000)이 영산강 수계에 대하여 일단위 및 시단위의 유출량 예측을 실시하였으며, 고영준(2001)은 하천의 시유출량 예측을 위해 퍼지-신경망을 적용하여 적용성을 검증하였다. 신성일(2002)은 1997~2002년 발생한 홍수사상에 대해서 금호강 중류부에 위치한 금호수위표지점에서 관측된 자료를 사용하여 홍수량 예측을 실시하였으며, 정대명(2003)은 적응형 뉴로-퍼지 기법을 이용하여 소양강의 유입량을 예측하였다.

회귀 모형을 이용한 국외 연구로는 수문학적 추정을 위한 적용성(Chaloulakou 등(1999), McKerchar와 Delleur(1974)와 Yurekli 등(2005)), 수위예측(Tangborn과 Rasmussen(1976), Phien 등(1990)과 Schilling과 Wolter(2005)) 등과 같은 분야에 적용되었다. Huynh과 Nguyen(2003)은 다중선형 회귀모형 구성에서 주요지점의 수위를 종속변수로 하여 주요지점과 상류지점의 이전시간 수위와 이전시간 강우량 자료를 독립변수로 하여 베트남의 홍강(Red river)에 적용결과 정확도 높은 예측 값을 얻어 적용성을 입증한 바 있다.

회귀 모형을 이용한 국내 연구로는 윤용남과 원석연(1991)이 미계측 중소유역의 월유출량 산정을 위해 다중회귀모형을 사용하였고, 윤강훈과 신현민(1994)은 한강인도교 수위와 영향인자간의

다중회귀분석을 통해 홍수위 예측을 수행한 바 있다. 윤강훈과 김태균(2003)이 레이더 강우를 이용한 다중회귀모형을 이용하여 평창강, 보청천, 위천 유역에 대해 모형의 적용성을 검토하였고, 성지연과 허준행(2009)은 다중선형 회귀 모형의 다중공선성 문제를 해결하기 위해 독립변수를 조정하고 회귀계수를 재산정한 통계적 모형을 이용하여 수위를 예측하였다. 정동국과 이범희(2009)는 대전광역시외의 도심하천 구간에서 현재까지의 강우, 수위 및 상류지역의 수위를 기반으로 수위를 예측할 수 있는 회귀모형을 구성하여 2시간 후의 수위를 예측하였으며, 이범희(2010)는 회귀분석에 의한 도시홍수 예보시스템을 개발하여 수위를 예측하였다.

본 연구에서는 입·출력 자료의 관계만을 이용하여 홍수를 예측하는 자료 기반의 수위예측 방법 중 비선형 시스템의 모의가 가능한 뉴로-퍼지 모형과 모형의 구축과정이 단순하고 기본 이론이 간단하여 수위예측에 자주 사용되고 있는 회귀분석 모형을 이용하여 왕숙천 유역에 적용하고 두 모형간의 수위예측 성능을 비교·분석하였다.

2. 기본 이론

2.1 뉴로-퍼지 모형

퍼지 이론과 신경망을 결합한 뉴로-퍼지 모형의 기본 이론은 Jang(1992)에 의해 최초로 제안되었으며 그 이후 Brown과 Harris(1994)에 의해 오늘날까지 발전되어 오고 있다. 퍼지 이론은 논리적인 기반 하에서 애매한 정보를 처리, 추론하는데 적합하고 자연 언어적인 표현을 통해 고차원적인 처리에 유용하다. 반면 신경망 이론은 학습능력을 가지고 있어서 시스템의 구성에 유연성이 크며, 데이터 기반의 처리에 우수하다.

현재 퍼지 논리 제어 시스템을 설계하는데 체계적인 방법은 없으며 지금까지 많이 사용되고 있는 방법 중에 하나는 전문가의 know-how를 연구하여 주관적인 방법으로 소속 함수와 규칙을 결

정하고 만족할만한 결과를 얻는지 여부를 시험해보는 것이다. 만약 설계한 제어 시스템이 실제 적용에서 만족할만한 결과를 주지 못할 경우 소속 함수와 규칙들은 재조정 되며, 원하는 제어 시스템을 선정하기 위해서 여러 번의 시행착오와 재조정을 거쳐 소속 함수와 규칙이 결정된다.

이러한 문제를 해결하기 위해서 제시된 모형이 뉴로-퍼지 모형으로 신경망의 구조와 학습능력을 이용하여 제어환경으로부터 얻은 입출력 정보로부터 소속 함수와 규칙을 제어 대상에 맞게 자동으로 조정함으로써 수행된다. 신경망을 이용하여 퍼지 추론을 위한 소속 함수를 결정할 때에는 다음과 같은 장점이 있다.

- ① 첫째, 사용자의 주관적인 판단으로 소속 함수를 결정하지 않고 신경망의 학습으로 소속 함수가 결정되기 때문에 시행착오와 재조정과 같은 과정이 생략되어 시스템의 구축 시간이 단축될 수 있다.
- ② 둘째, 신경망이 비선형성을 가지기 때문에, 비선형성을 갖는 입출력 자료의 관계를 나타내는 비선형성의 소속 함수를 결정함으로써 결과의 정확도를 높일 수 있다.
- ③ 셋째, 신경망의 학습기능을 이용하여 규칙을 자동으로 획득할 수 있다.

2.1.1 뉴로-퍼지 모형의 구조

본 연구에서 사용된 뉴로-퍼지 모형의 구조를 설명하기 위해 퍼지 추론 시스템이 2개의 입력 값과 한 개의 출력 값을 갖는다고 가정하면 1차 Takagi-Sugeno 퍼지 모형은 다음과 같은 2개의 규칙을 갖는다. Takagi-Sugeno 퍼지 모형은 퍼지추론 기법 중 하나로 복잡한 비선형 시스템을 표현하기 위하여 비선형 모델을 여러 개의 구간별 선형 서브 모델로 분해하고 퍼지 규칙을 통해 추론을 하는 기법이다.

Rule 1 : If x is A1 and y is B1, then

$$f_1 = p_1x + q_1y + r_1$$

Rule 2 : If x is A2 and y is B2, then

$$f_2 = p_2x + q_2y + r_2$$

여기서 x 와 y 는 입력 값을 나타내고, A1, B1,

A2, B2는 소속함수, $p_1, q_1, r_1, p_2, q_2, r_2$ 는 선형서브모델의 매개변수를 나타내고 있다. 그림 1은 1차 Takagi-Sugeno 추론 시스템을 나타낸 것이고 그림 2는 신경망을 도입한 뉴로-퍼지 모형의 구조이다.

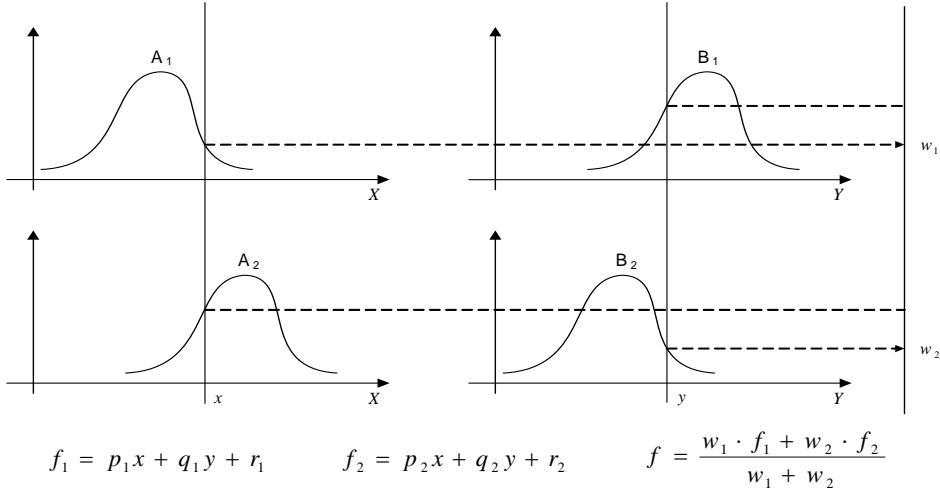


그림 1. 1차 Takagi-Sugeno 추론 기법

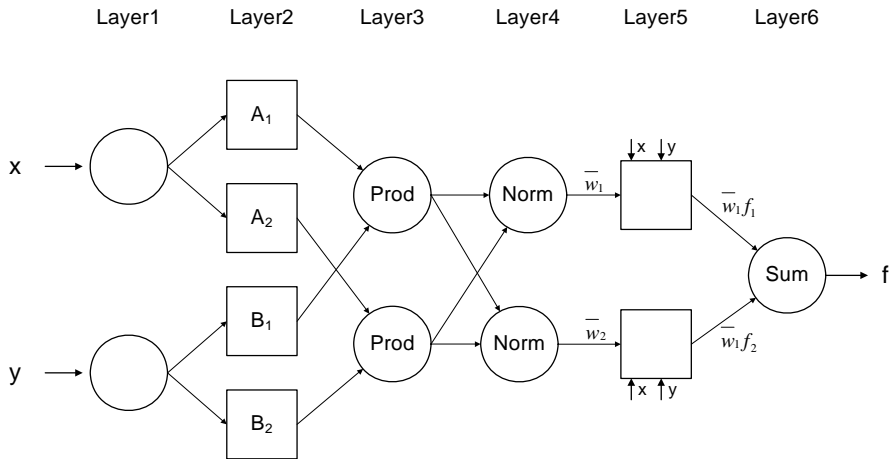


그림 2. 뉴로-퍼지 모형의 구조

뉴로-퍼지의 구조는 $O_{i,i}$ 로 표현되는 각 층(layer) 내에서 노드 출력을 갖는 5개의 층으로 구성되어 있다. 여기서 $O_{i,i}$ 는 층 l 의 i 번째 노드의 출력 값을 나타낸다. 그림 2의 뉴로-퍼지 구조의 각 층에 대해 자세히 살펴보면 다음과 같다.

Layer 1 : 이 층의 모든 노드는 입력 노드로 단지 다음 층으로 외부적 신호만 보낸다.

Layer 2 : 이층의 모든 노드 i 는 적용형 노드로 소속 함수로 작용한다.

$$O_{2,i} = \mu_{A_i}(x), i = 1, 2 \quad (1)$$

$$Q_{2,i} = \mu_{B_{i-2}}(x), i = 3, 4$$

여기서 x 또는 y 는 입력 노드 i 이고, A_i 또는 B_{i-2} 는 이 노드에 관련된 퍼지 집합(언어적 표시)이다. $O_{2,i}$ 는 퍼지 집합(A_1, A_2, B_1, B_2)의 소속 정도를 나타내고 이의 전형적인 소속 함수는 식 (2)와 같다.

$$\mu_{A(x)} = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad (2)$$

여기서 $\{a_i, b_i, c_i\}$ 를 전건부 매개변수라 한다. 이 함수에서 매개변수 값이 변화함에 따라 퍼지 집합 A 의 소속 함수는 다양하게 표현될 수 있다. 즉, 위의 소속 함수에서 a_i, b_i, c_i 의 값을 변화시킴으로써 소속 함수를 사용자가 원하는 다양한 형태로 나타낼 수 있다는 것이다.

Layer 3 : 이층의 모든 노드는 Prod로 표시된 고정노드이며 출력은 모든 입력 신호의 곱으로 식 (3)과 같다.

$$O_{3,i} = w_i = \mu_{A_i(x)} \cdot \mu_{B_i(y)}, i = 1, 2 \quad (3)$$

각 노드의 출력 값은 규칙의 연결강도를 나타낸다.

Layer 4 : 이층의 모든 노드는 Norm이라 표시된 고정 노드이며 i 번째 노드와 관련된 규칙의 정규화된 연결강도를 나타낸다. 즉, i 번째 노드는 전체 연결강도에 대한 i 번째 노드의 연결강도의 비로 식 (4)와 같다.

$$O_{4,i} = \bar{w}_i = \frac{w_i}{\sum w_i} \quad i = 1, 2 \quad (4)$$

Layer 5 : 이층의 모든 노드는 노드 함수를 갖는 적응형 노드로 식 (5)와 같이 가중된 후건부

값을 계산한다.

$$O_{5,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (5)$$

여기서 \bar{w}_i 는 층 4에서의 출력 값이고, $\{p_i, q_i, r_i\}$ 는 이 노드의 후건부 매개변수이다.

Layer 6 : 이 층의 노드는 오직 하나로 함으로 표시되며, 전체 시스템에 대한 추론 결과를 계산하기 위해 식 (6)과 같이 모든 입력 신호의 함으로 표현된다.

$$O_{6,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (6)$$

2.1.2 Subtractive Clustering 알고리즘

Subtractive Clustering 알고리즘은 객체 자료간의 밀집한 정도에 따라 포텐셜이 높게 구해지는 성질을 이용하여 이를 관계 자료로 사용한다. 각 자료 점에서의 밀집도를 계산한 후에 첫 번째 군집을 중심으로 밀집도가 가장 큰 지점을 선택한다. 이 알고리즘의 결과로 군집의 중심에 가장 가까운 객체 자료가 대표되는 군집의 중심으로 구해진다. Subtractive Clustering 알고리즘은 다음과 같은 과정을 수행한다.

- ① 자료간의 거리를 구한 후 각 객체 자료 X_i 마다 포텐셜 P_i 를 구한다.

$$P_i = \sum_{j=1}^N \exp(-\alpha \|X_i - X_j\|^2), i = 1, 2, \dots, N \quad (7)$$

여기서 $\alpha = 4/r_a^2$ 이고, r_a 는 양의 상수로 각 포텐셜 값의 의미있는 경계를 정해주는 파라미터이다. 즉 r_a 밖의 자료는 포텐셜 값에 영향을 주지 못한다.

- ② 첫 번째 군집의 중심을 구한다. N 개의 포텐

설 중 가장 높은 값을 P_1^* 라 놓고 이때의 자료가 첫 번째 군집의 중심 X_1^* 가 되게 한다.

- ③ 첫 번째 군집의 중심의 영향을 다른 포텐셜 P_i 에서 제거한다. 즉 첫 번째 군집의 중심 근처에서 다음 군집의 중심이 발생하지 않도록 첫 번째 군집의 포텐셜 값을 제거하지 않으면 안 된다. 첫 번째 군집의 중심을 제거한 포텐셜 값 P_i' 는 다음과 같다.

$$P_i' = P_i - P_i^* \exp(-\beta \| X_i - X_1^* \|^2), i = 1, 2, \dots, N \quad (8)$$

여기서 $\beta = 4/r_b^2$ 이고, r_b 는 양의 상수로 $r_b > r_a$ 이며 군집의 중심 근처에서 다음 군집의 중심이 나타나지 않도록 한다.

- ④ 두 번째 클러스터의 중심을 구한다. 즉 P_i' 중 가장 높은 값을 P_2^* 라고 하고 이때의 자료가 두 번째 군집의 중심 X_2^* 가 된다.
- ⑤ 일반적으로 k번째 군집의 중심은 다음과 같이 구한다.

$$P_i' = P_i - P_{k-1}^* \exp(-\beta \| X_i - X_{k-1}^* \|^2), i = 1, 2, \dots, N \quad (9)$$

여기서 가장 높은 값을 갖는 자료 X_k^* 가 먼저 군집의 중심 후보가 되며, 이를 사용하여 다음 군집의 중심의 테스트를 한다.

2.2 회귀분석 모형

회귀분석이란 객관적으로 나타난 자료를 바탕으로 독립변수와 종속변수간의 상호관계(선형, 비선형 관계 등)를 분석하는 기법이다. 이러한 회귀분석은 변수들간의 관계를 정형적으로 나타낼 수 있으므로 독립변수와 종속변수간의 미래 추이를 예측할 경우 유용하다. 회귀분석에 사용되는 변수

는 크게 독립변수(independent variable)와 종속변수(dependent variable)로 구분되는데 하나의 독립변수와 하나의 종속변수를 포함하는 가장 단순한 회귀분석 형태를 단순회귀분석이라고 하고 여러 개의 독립변수가 있는 회귀분석을 다중회귀분석이라고 한다.

종속변수 Y 를 설명하는데 k개의 독립변수인 X_1, X_2, \dots, X_k 를 도입할 때 다중회귀모형은 다음과 같이 정의된다.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \epsilon_i, i = 1, 2, \dots, n \quad (10)$$

여기서, $\beta_0, \beta_1, \dots, \beta_k$ 는 회귀계수 또는 매개변수를 의미하고, ϵ_i 는 서로 독립이고 동일한 분포 $N(0, \sigma^2)$ 을 따르는 오차항이다. $\beta_j (j=1, \dots, k)$ 는 j번째 독립변수 X_j 의 회귀계수(기울기)를 의미하는데, X_j 를 제외한 다른 모든 독립변수의 값이 고정된 상태에서 X_j 의 값이 한단위 증가할 때의 Y 값의 평균변화량을 나타낸다. 그리고 X_{ji} 는 독립변수 X_j 의 i번째 관측치를 의미한다.

다중선형회귀모형은 행렬로 표현함으로써 간단한 형태로 나타낼 수 있다. 식 (10)을 행렬형태로 표현하기 위해 β_0 에 형식적으로 '1'을 넣고 관측번호 i 를 1에서 n 까지 풀어쓰면 다음과 같다.

$$\begin{aligned} Y_1 &= \beta_0 \cdot 1 + \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + \epsilon_1 \\ Y_2 &= \beta_0 \cdot 1 + \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + \epsilon_2 \\ &\vdots \\ Y_n &= \beta_0 \cdot 1 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + \epsilon_n \end{aligned} \quad (11)$$

Y 행, X 행, 모수, 오차를 각각 묶어 벡터와 행렬을 이용하여 다음과 같이 나타낼 수 있다.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (12)$$

여기서 벡터는 모두 열벡터(column vector)인데 벡터 Y 와 ϵ 는 크기 n 의 열벡터이며, β 는 크기 $(k+1)$ 의 열벡터이다. 그리고 X 는 $n \times (k+1)$ 의 행렬이 된다. 이때 다중선형회귀모형은 다음과 같은 간단한 형태로 나타낼 수 있다.

$$Y = X\beta + \epsilon \quad (13)$$

여기서 오차항 벡터 ϵ 는 정규분포를 따르는 n 개의 확률변수 ϵ_i 들로 이루어졌기 때문에 다변량 정규분포를 따른다.

2.2.1 회귀계수 추정

다중선형회귀모형의 회귀계수들의 추정 방법 중 가장 기본적인 방법인 최소자승법이다. 최소자승법은 실제값과 추정치간 차이 즉, 회귀오차의 자승을 합한 값이 최소가 되는 표본 회귀식을 찾는 방법이다. 오차를 최소화하기 위한 회귀계수 $\beta_0, \beta_1, \dots, \beta_k$ 를 구하기 위해서는 $\beta_0, \beta_1, \dots, \beta_k$ 에 대해 각각 편미분한 값이 0이 되면 된다. 오차항의 제곱합 Q 는 다음과 같이 표현할 수 있다.

$$Q = [Y_i - (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})]^2 \quad (14)$$

최소자승법에 의한 회귀계수의 추정은 제곱합 Q 를 각 β_j 에 대하여 편미분하고 이를 0으로 하는 다음과 같은 연립방정식을 풀어 $\hat{\beta}_j$ 들을 구하는 것이다.

$$\begin{aligned} \frac{\partial Q}{\partial \beta_0} &= -2 \sum [Y_i - (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})] = 0 \\ \frac{\partial Q}{\partial \beta_1} &= -2 \sum X_{1i} [Y_i - (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})] = 0 \\ &\vdots \\ \frac{\partial Q}{\partial \beta_k} &= -2 \sum X_{ki} [Y_i - (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})] = 0 \end{aligned} \quad (15)$$

위의 연립방정식을 만족하는 β_j 를 $\hat{\beta}_j$ 라 하면 다

음 관계가 성립하는데 이를 정규방정식이라 한다.

$$\begin{aligned} \hat{\beta}_0 n + \hat{\beta}_1 \sum X_{1i} + \dots + \hat{\beta}_k \sum X_{ki} &= \sum Y_i \\ \hat{\beta}_0 \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \dots + \hat{\beta}_k \sum X_{1i} X_{ki} &= \sum X_{1i} Y_i \\ &\vdots \\ \hat{\beta}_0 \sum X_{ki} + \hat{\beta}_1 \sum X_{ki} X_{1i} + \dots + \hat{\beta}_k \sum X_{ki}^2 &= \sum X_{ki} Y_i \end{aligned} \quad (16)$$

식 (16)을 벡터-행렬식으로 표현하기 위해서 행렬 $X^T X$ 를 다음과 같이 나타낼 수 있다.

$$X = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} & \dots & \sum X_{ki} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i} X_{2i} & \dots & \sum X_{1i} X_{ki} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_{ki} & \sum X_{ki} X_{1i} & \sum X_{ki} X_{2i} & \dots & \sum X_{ki}^2 \end{bmatrix} \quad (17)$$

식 (17)을 이용하면 식 (18)의 정규방정식이 다음과 같이 표현된다.

$$(X^T X) \hat{\beta} = X^T Y \quad (18)$$

따라서 $\hat{\beta}$ 의 최소자승추정량은 다음과 같다.

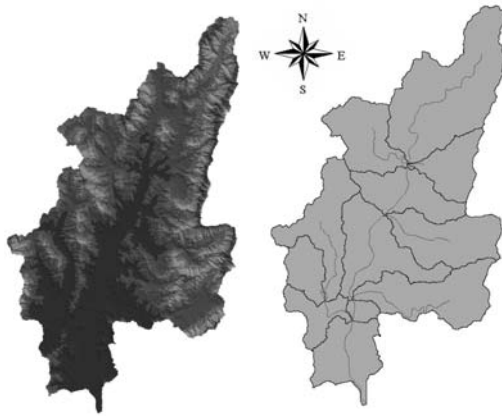
$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (19)$$

3. 모형의 구축

3.1 대상유역 및 자료

왕숙천 유역은 행정구역으로는 경기도의 구리시와 남양주시 및 포천군 일부에 걸쳐 있다. 유역 내 수위관측소는 부평, 연평, 내곡, 퇴계원, 갈매 수위관측소가 있으며, 이 중 왕숙천 유역의 홍수 예경보 지점은 퇴계원 수위관측소 지점으로 1984년부터 관측을 개시하였다.

유역 전체에 대한 평균 표고는 약 189.2 m이며, 평균 경사는 10.11% 정도이다. 왕숙천 유역은 경기도 포천군 내촌면의 수원산(705.0 EL.m)에서 발원하여 남쪽으로 유하하다가 용암천과 우안에서 합류한 후 한강의 본류로 유입된다. 유역 면적은 270.79km², 본류의 유로연장은 37.24km이다(건설교통부, 2005). 그림 3은 대상 유역의 TIN과 유역도를 나타내고 있다.



(a) 대상유역 TIN (b) 유역도

그림 3. 대상유역

실시간 수위예측을 위한 지점으로 왕숙천 하류 지점인 퇴계원 지점을 선정하였고 모형을 구축하기 위한 입력 자료로는 16개의 호우사상을 선정하여 각 호우사상의 대상 유역에 대한 강우관측소

의 강우자료와 퇴계원 수위관측소의 수위를 사용하였다. 강우자료는 대상 유역의 강우관측소로부터 얻어진 각각의 강우량으로부터 유역의 평균 강우량을 산정하여 사용하였다. 이것은 왕숙천 유역은 총 유역면적 270.79km²로 소규모 유역에 속하므로 각 관측소의 강우량을 각각 서로 다른 입력 변수로 고려할 만큼 지체시간이 길지 않다고 판단되었고, 따라서 강우자료는 유역 전체를 대표하는 하나의 강우량 값을 산정하여 사용하였다. 유역의 평균 강우량은 각 강우관측소에서 측정된 값을 티센법을 이용하여 면적가중치를 주어 유역전체의 평균값으로 사용하였다.

표 1은 모의를 위해 선정된 16개의 각 호우사상의 강우시작 및 종료 시각, 총 강우량, 호우원인 등의 특성을 나타내고 있으며 향후 모의에 대한 분석 등의 설명에 대한 이해를 돕고 표기를 간단하게 하기 위해 표에서 보는 바와 같이 각 호우사상에 대한 이름을 부여하였다.

표 1. 선정된 호우사상의 특성

호우사상이름	강우시작	강우종료	총강우량(mm)	강우지속 시간(hr)	호우원인	비고
W-1	03/05/07 05:00	03/05/07 24:00	95	19	기압골	Training
W-2	03/07/21 00:00	03/07/22 04:00	222	28	장마전선	Training
W-3	03/08/06 12:00	03/08/07 09:00	106	21	태풍(아타우)	Training
W-4	03/08/19 18:00	03/08/20 14:00	123	20	기압골	Training
W-5	03/08/23 05:00	03/08/26 08:00	246	51	기압골	Training
W-6	03/08/27 13:00	03/08/28 08:00	84	19	기압골	Training
W-7	03/09/18 05:00	03/09/19 01:00	170	20	기압골	Training
W-8	04/06/19 07:00	04/06/20 20:00	101	37	기압골	Training
W-9	04/07/11 21:00	04/07/13 08:00	174	35	장마전선	Testing
W-10	04/08/17 18:00	04/08/20 13:00	96	43	태풍(메기)	Testing
W-11	05/06/26 23:00	05/06/27 19:00	140	20	기압골	Checking
W-12	05/07/01 01:00	05/07/01 11:00	109	10	장마전선	Checking
W-13	05/07/28 02:00	05/07/28 20:00	105	18	기압골	Checking
W-14	05/08/02 22:00	05/08/03 06:00	86	8	기압골	Checking
W-15	05/08/10 15:00	05/08/11 16:00	159	25	기압골	Testing
W-16	05/09/13 06:00	05/09/14 04:00	95	19	기압골	Testing

3.2 뉴로-퍼지 수위예측 모형 구축

3.2.1 뉴로-퍼지 수위예측 모형

뉴로-퍼지 모형 수위예측 모형을 구축하기 위해서는 training, checking, testing 세 단계를 거치게 되는데, 이 중 training과 checking 과정을 통해 소속 함수의 매개변수를 결정하게 된다. 소속함수의 매개변수는 각 소속함수의 형태를 결정짓는 변수로 예를 들어 본 연구에서 사용한 가우스 함수의 경우 2개의 매개변수를 가지며 이 매개변수 값에 의해 소속함수의 폭과 중앙값이 결정된다. training과 checking 과정에 사용하는 호우

사상에 따라 소속 함수의 매개변수 값이 달라지게 되며 뉴로-퍼지 수위예측 모형의 결과 또한 차이가 나게 된다. 따라서 최적의 수위예측 모형을 구축하기 위해서는 선정된 16개의 호우사상을 이용해 training과 checking을 위한 다양한 조합을 구성하여 각 조합에 따른 결과를 비교하여 training과 checking을 위한 호우사상을 선정하여야 한다. 그러나 본 연구의 목적은 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형과의 수위예측에 대한 성능을 비교하는 것이므로 두 모형간 training과 checking을 위한 사상을 표 1과 같이 선정하여 모형을 구축하였다.

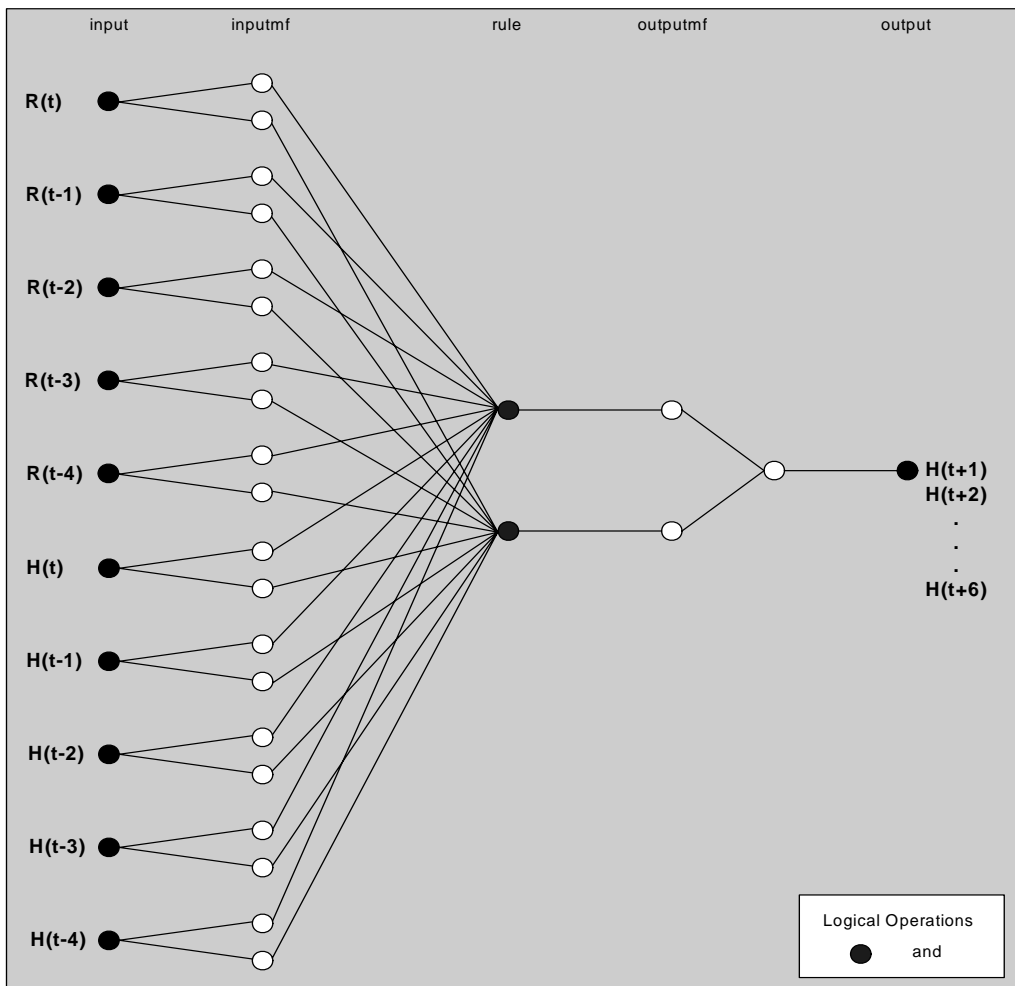


그림 4. 구축된 뉴로-퍼지 모형

뉴로-퍼지 모형을 이용한 수위예측에 있어서 중요한 사항은 수위에 영향을 미치는 강우와 수위의 시간적 분포를 입력자료 조합에 어떻게 구성할 것인지 또 얼마만큼의 자료를 사용할 것인지이다. 예를 들면 현재 시간을 t시간이라 하면, 선행시간 t+1, t+2, t+3 시간의 수위를 예측하는데 있어서 수위 자료는 t, t-1 시간의 자료, 강우자료는 t, t-1, t-2 시간의 자료를 사용할 것인지 혹은 수위 자료는 t, t-1, t-2 시간의 자료, 강우자료는 t, t-1 시간의 자료를 사용하거나 혹은 더 많은 자료들을 사용할 것인지를 결정하는 일은 좀 더 정확한 수위를 예측하는데 있어서 중요하다.

그러나 이러한 최적 입력자료 조합의 선정을 위해서는 상당한 시간이 소요된다. 또한 본 연구의 목적은 앞선 설명한 바와 같이 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형의 수위예측 성능을 비교하는 것으로 두 모형간의 정확한 비교 검토를 위해서는 두 모형의 입력자료의 시간적 분포를 동일하게 해야 한다. 따라서 본 연구에서는 수위와 강우의 시간분포를 t-4까지로 결정하여 입력자료를 구성하였다.

모의 시간의 단축 및 최적의 입력 공간 분할을 위해서 subtractive clustering 기법을 추가적으로 적용하여 그림 4와 같이 각 입력 변수당 2개의 소속 함수와 2개의 규칙을 이용한 뉴로-퍼지 모형을 구축하였다. 또한 입력 자료(수위와 강우의 조합)와 출력 자료(수위)의 비선형성을 고려하여 소속 함수의 형태로는 가우스 함수를 사용하였으며 수위예측의 선행시간은 t+1(30분), t+2(60분), t+3(90분), t+4(120분), t+5(150분), t+6(180분)의 6가지 경우로 하여 각 선행시간별 소속 함수의 매개변수를 결정하였다.

3.3 회귀분석 수위예측 모형 구축

회귀모형에 있어서 유의성검정, 즉 회귀성 검정은 독립변수의 기울기에 해당하는 모든 회귀계수 (β_0 은 제외)가 0의 값을 갖는가 또는 그렇지 않은

가를 검정하는 것이다. 독립변수의 기울기에 해당하는 모든 회귀계수가 0일 때 회귀성이 없다고 말하며, 적어도 하나의 기울기가 0이 아니라고 판단되면 회귀성이 있다고 말하게 된다. 종속변수 Y의 총변동 SST는 추정된 회귀식에 의해 설명되는 변동 SSR과 잔차에 의한 변동 SSE로 분해된다. 이러한 변동으로부터 다음의 다중회귀모형의 분산분석표를 얻을 수 있다.

표 2. 다중회귀모형의 분산분석표

요인	자유도	제곱합	제곱 평균	F-값
다중선형회귀	k	SSR	MSR=SSR/k	$F_0=MSR/MSE$
잔차	n-k-1	SSE	MSE=SSE/(n-k-1)	
계	n-1	SST		

분산분석표에 주어진 F-값은 설정된 회귀식의 유의성 검정에 사용되는데, 통상적으로 F-값이 1보다 큰 값이면 회귀식이 유의하다고 할 수 있다. 표 3은 회귀분석 수위예측 모형의 분산분석 결과를 나타낸 것으로 표에서 보는 바와 같이 본 회귀식이 유의함을 알 수 있다.

회귀분석 결과와 최소자승법을 이용한 회귀계수 산정 방법을 통해 회귀분석 수위예측 모형을 구축하였다. 표 4는 각 선행 예측시간별 수위예측 모형의 회귀계수 산정결과와 각 독립변수의 개별회귀계수 t-검정 값을 나타내고 있다. 각 독립변수의 종속변수 Y에 대한 영향력의 대소를 판단하기 위한 개별회귀계수의 유의성 검토에는 표에서 보는 바와 같이 t-검정 값을 이용한다. t-검정 값이 높은 변수일수록 종속변수 Y를 예측하는데 공헌도가 높다고 판단하며 식 (20)과 같이 계산한다.

$$t\text{-검정 값} = \frac{\text{개별회귀계수 값}}{\text{계수의 표준오차}} \quad (20)$$

표 3. 회귀분석 수위예측 모형의 분산분석 결과

모형	요인	자유도	제곱합	제곱 평균	F-값
H(t+1) 예측모형	다중선형회귀	10	99.67021	9.967021	33,613.93
	잔차	1,067	0.316381	0.000297	
	계	1,077	99.98659		
H(t+2) 예측모형	다중선형회귀	10	98.25883	9.825883	8,705.147
	잔차	1,067	1.20437	0.001129	
	계	1,077	99.4632		
H(t+3) 예측모형	다중선형회귀	10	95.91665	9.591665	3,396.062
	잔차	1,067	3.013581	0.002824	
	계	1,077	98.93023		
H(t+4) 예측모형	다중선형회귀	10	92.59696	9.259696	1,703.073
	잔차	1,067	5.801333	0.005437	
	계	1,077	98.3983		
H(t+5) 예측모형	다중선형회귀	10	88.37533	8.837533	993.2676
	잔차	1,067	9.493562	0.008897	
	계	1,077	97.86889		
H(t+6) 예측모형	다중선형회귀	10	83.56202	8.356202	647.4978
	잔차	1,067	13.77003	0.012905	
	계	1,077	97.33206		

표 4. 회귀분석 수위예측 모형 구축 결과

	H(t+1) 예측모형		H(t+2) 예측모형		H(t+3) 예측모형	
	계수	t-검정	계수	t-검정	계수	t-검정
절편	-0.0031	-4.8076	-0.0076	-6.0299	-0.0112	-5.6579
R(t)	0.0026	7.6171	0.0111	16.7335	0.0200	19.0635
R(t-1)	0.0053	13.3685	0.0080	10.3641	0.0100	8.2193
R(t-2)	-0.0006	-1.5016	-0.0006	-0.7475	-0.0001	-0.0562
R(t-3)	0.0002	0.5668	0.0003	0.3646	0.0008	0.6410
R(t-4)	-0.0008	-2.1433	-0.0019	-2.5711	-0.0033	-2.8326
H(t)	1.5658	47.0441	1.7650	27.1780	1.6180	15.7505
H(t-1)	-0.6656	-10.7340	-1.0847	-8.9659	-1.1042	-5.7701
H(t-2)	-0.0350	-0.5334	0.1054	0.8240	0.2177	1.0754
H(t-3)	0.0684	1.1899	0.0259	0.2309	-0.0705	-0.3973
H(t-4)	0.0420	1.5152	0.1250	2.3125	0.2285	2.6712
	H(t+4) 예측모형		H(t+5) 예측모형		H(t+6) 예측모형	
	계수	t-검정	계수	t-검정	계수	t-검정
절편	-0.0131	-4.7457	-0.0129	-3.6674	-0.0114	-2.6880
R(t)	0.0275	18.9152	0.0325	17.4399	0.0356	15.8881
R(t-1)	0.0107	6.3227	0.0113	5.2395	0.0109	4.1746
R(t-2)	0.0013	0.6961	0.0018	0.7898	0.0031	1.1052
R(t-3)	0.0007	0.3615	0.0014	0.5823	0.0017	0.6058
R(t-4)	-0.0037	-2.3121	-0.0041	-1.9709	-0.0036	-1.4557
H(t)	1.3350	9.3664	1.0441	5.7263	0.7514	3.4219
H(t-1)	-0.9179	-3.4572	-0.7254	-2.1358	-0.4588	-1.1215
H(t-2)	0.2405	0.8564	0.3176	0.8838	0.3119	0.7208
H(t-3)	-0.1236	-0.5024	-0.2274	-0.7224	-0.3286	-0.8667
H(t-4)	0.3072	2.5887	0.3868	2.5482	0.4768	2.6082

4. 모의결과 및 비교검토

구축된 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형을 이용하여 수위예측을 수행하였다. 모형의 검증을 위해 사용한 4개의 호우 사상에 대한 결과 중 W-9와 W-10 호우사상의 각 수위예측 모형에 대한 선행시간별 예측 결과를 그림 5와 그림 6에 나타내었다. 그림 5에서 보는 바와 같이 30분 선행 기간의 경우에는 본 연구에서 구축한 뉴로-퍼지 수위예측 모형과 회귀분석 수위

예측 모형 모두 예측 수위와 관측 수위가 상당히 잘 일치하고 있음을 알 수 있고 특히 침투예측에 있어서도 두 모형 모두 상당히 만족할 만한 결과를 나타내고 있다. 그러나 예측 선행시간이 길어질수록 회귀분석 수위예측 모형을 이용한 모형의 경우 예측 정확도가 뉴로-퍼지 수위예측 모형에 비해 떨어지는 것을 확인할 수 있고, 특히 침투예측에 있어서 다소 과소 예측되는 경향을 보이고 있음을 확인할 수 있다.

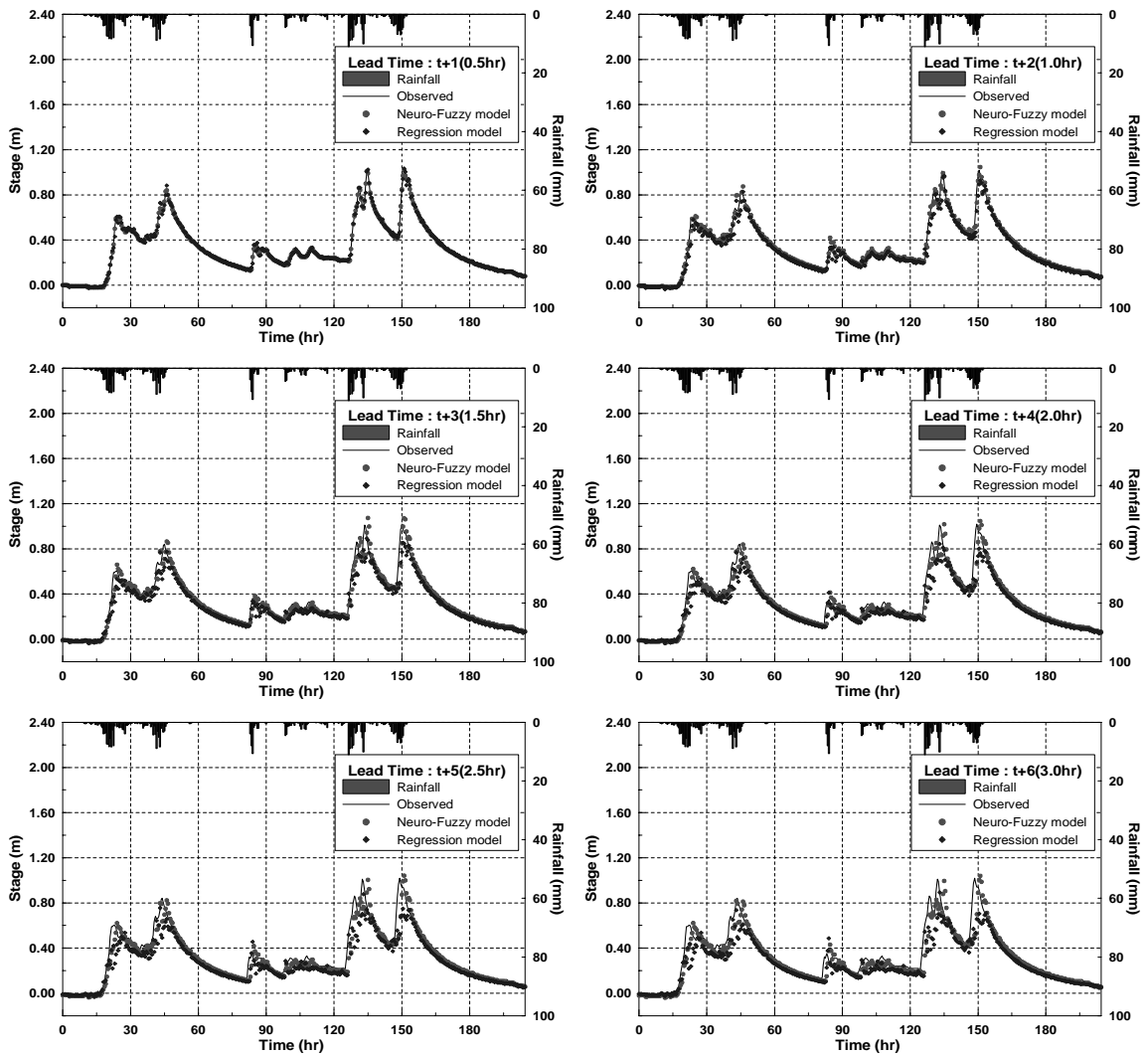


그림 5. 뉴로-퍼지와 회귀분석에 의한 선행시간별 수위예측 결과(W-9)

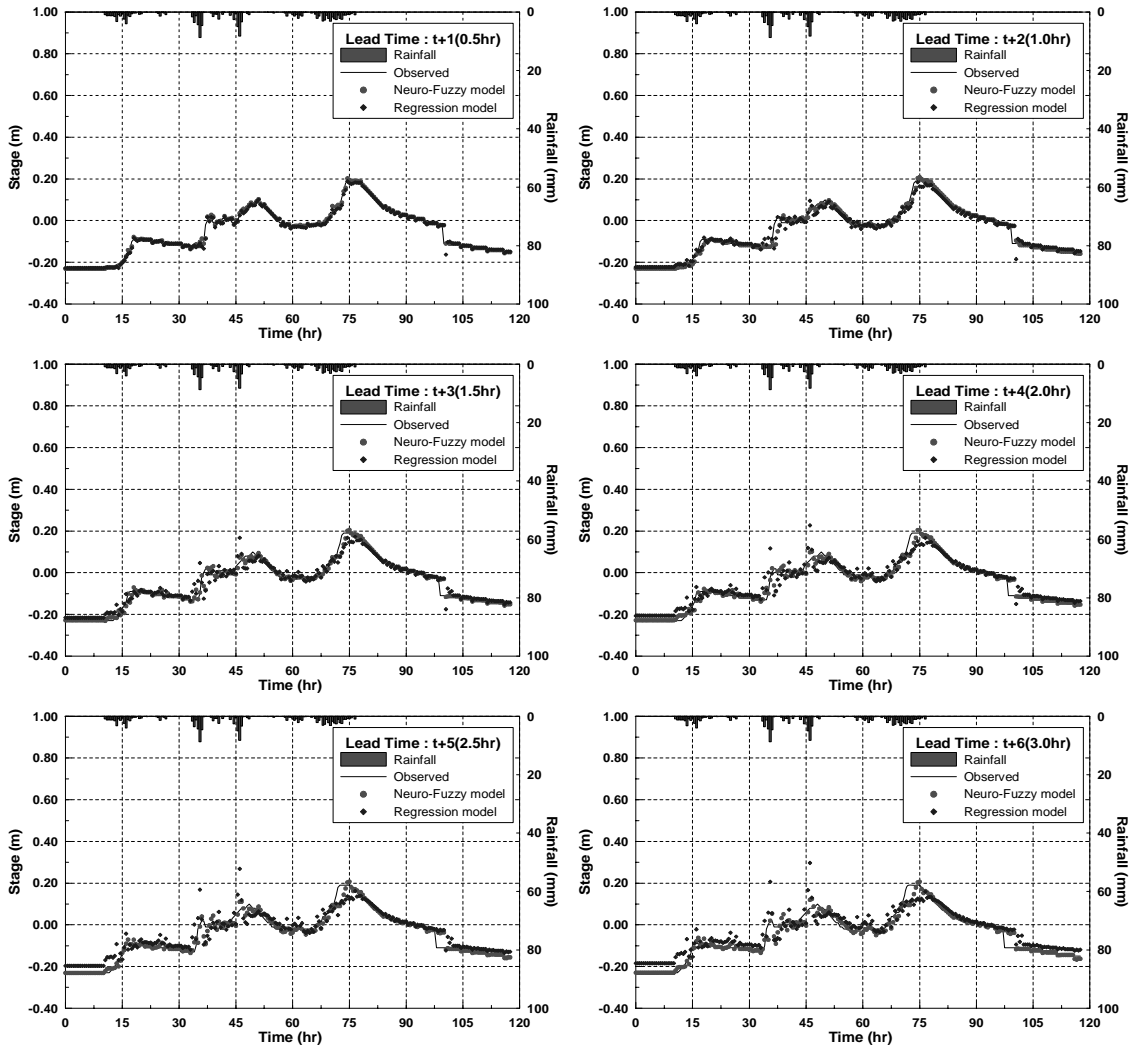


그림 6. 뉴로-퍼지와 회귀분석에 의한 선행시간별 수위예측 결과(W-10)

W-10의 수위예측 결과인 그림 6을 살펴보면 30분 선행시간의 경우 두 모형 모두 전체적인 수위형상에 있어 예측수위와 관측수위가 상당히 잘 일치하고 있으나 회귀분석 수위예측 모형의 경우 예측수위의 진동(oscillation)현상이 나타나는 것을 확인할 수 있고 또한 수위가 급격하게 하강하는 부분에서 관측수위와 예측수위가 상당히 차이가 남을 알 수 있다. 예측 선행시간이 길어질수록 두 모형간의 수위예측 결과는 확연히 차이가 나는 것을 확인할 수 있는데, 수위가 상승하는 구간에서는 회귀분석 수위예측 모형의 예측 결과가 관측수위에 비해

다소 과대 예측되고 있고, 침두수위에 있어서는 회귀분석 수위예측 모형이 뉴로-퍼지 수위예측 모형과 비교해서 과소 예측되는 것을 알 수 있다.

본 연구에서 구축된 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형간의 보다 정확한 비교를 위해 W-10 호우사상에 대한 모의결과에 대해 그림 7과 같이 t+1(0.5hr), t+3(1.5hr), t+6(3.0hr)의 선행시간에 대한 예측치와 관측치를 도시하고 수정결정계수 값을 산정하여 함께 나타내었고 그림 8에 각 선행시간별 예측치와 관측치간의 오차를 나타내었다. 그림 7에서 보는바와 같이 0.5시

간에 대한 예측 결과를 살펴보면 뉴로-퍼지 수위 예측 모형과 회귀분석 수위예측 모형의 수정 결정계수가 각각 0.990과 0.989로 그 차이가 0.001로 아주 작게 나타남을 알 수 있다. 그러나 선행 예측시간이 증가함에 따라 두 모형간 예측 정확도가 확연히 차이나 나는 것을 알 수 있는데, 1.5시간에 대한 예측 결과를 보면 뉴로-퍼지 수위예측 모형의 경우 수정 결정계수가 0.971인데 반해 회귀분석 수위예측 모형의 수정 결정계수는 0.929

로 두 모형간 수정 결정계수의 차이가 0.042가 됨을 알 수 있다. 선행 예측시간이 3.0시간인 경우에는 뉴로-퍼지 수위예측 모형의 수정 결정계수가 0.950이고 회귀분석 수위예측 모형의 수정 결정계수는 0.851로 차이가 0.099로 더 증가하고 있음을 확인할 수 있다. 이러한 결과를 통해 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 수위예측 능력이 우수한 것을 확인할 수 있다.

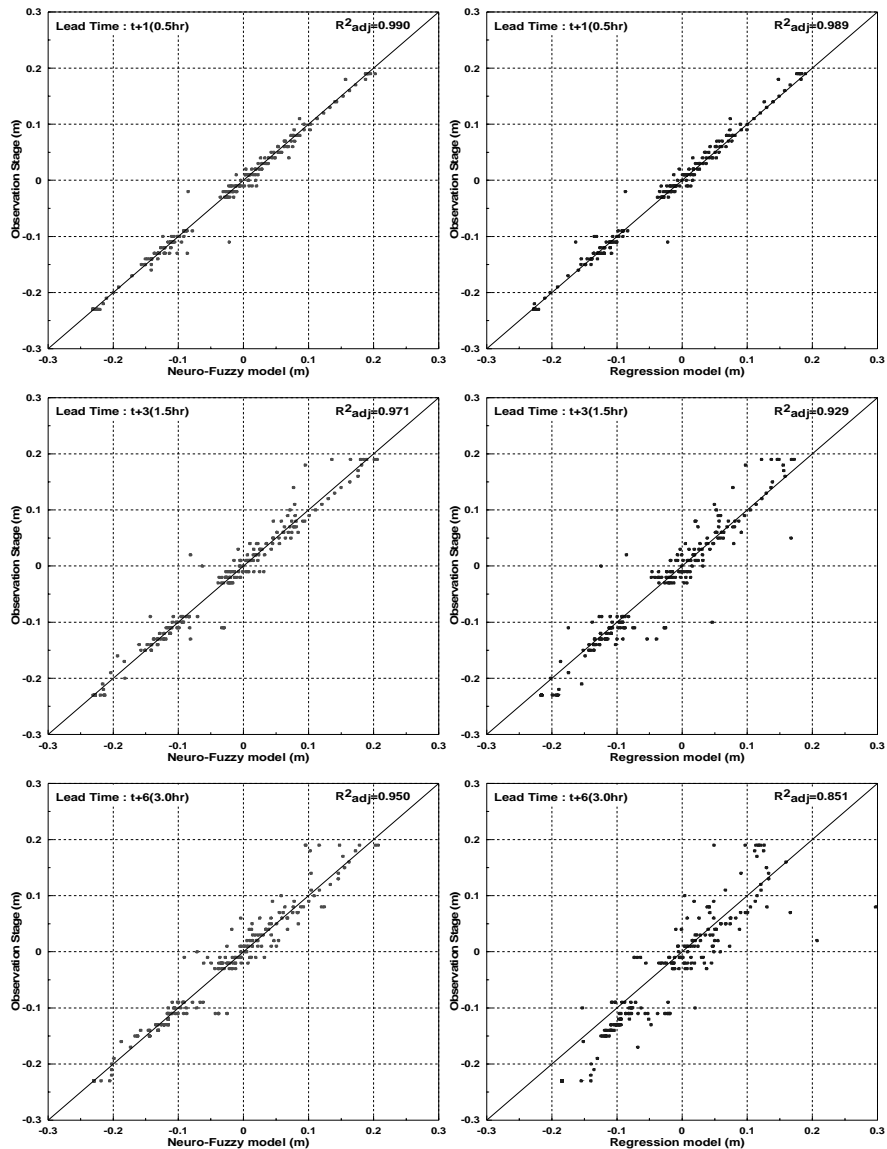


그림 7. 뉴로-퍼지와 회귀분석 수위예측 모형간 관측치와 예측치 결과 비교(W-10)

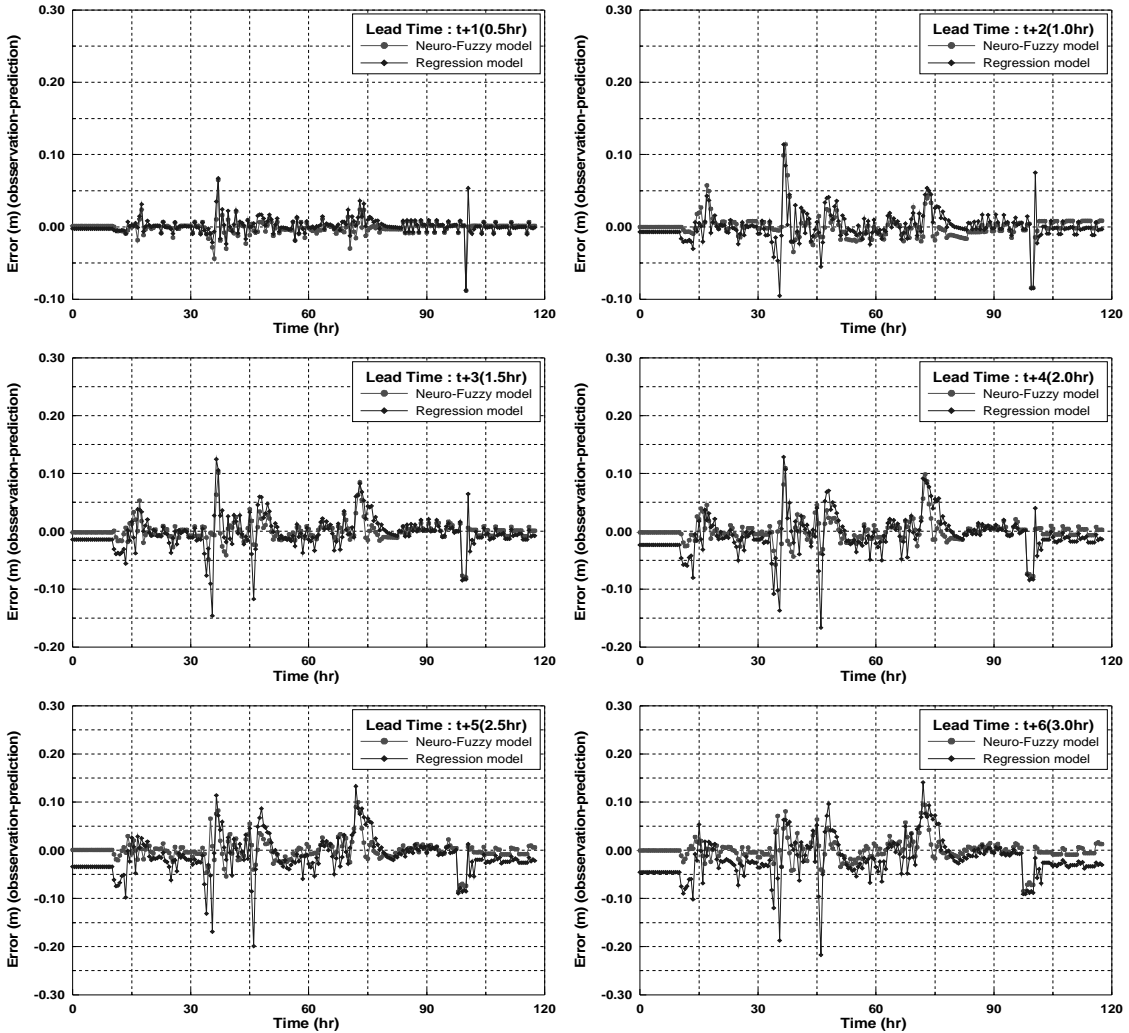


그림 8. 다중선형회귀 수위예측 모형간 오차의 비교(W-10)

이러한 결과는 예측치와 관측치간의 오차를 나타낸 그림 8을 살펴보면 더욱 명확하게 알 수 있다.

그림 8에서 0.5시간의 예측치와 관측치간의 오차를 나타낸 그림을 살펴보면 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 오차가 작게 나타나고 있음을 알 수 있고 이러한 결과는 선행 예측시간이 증가할수록 명확히 알 수 있다. 1.5시간의 선행시간에 대한 결과를 보면 그림에서 보는 바와 같이 회귀분석 수위예측 모형의 경우 예측치와 관측치간의 오차 변동폭이 약 -15cm~13cm의 범위를 보이는데 반해 뉴로-퍼지 수위예측 모

형의 경우 오차 변동 폭이 약 -8cm~10cm로 회귀 분석 수위예측 모형에 비해 약 30% 작게 나타나는 것을 확인할 수 있다. 3.0시간의 선행시간의 경우에도 그림에서 보는 바와 같이 회귀분석 수위예측 모형의 경우 오차의 변동 폭이 약 -22cm~15cm인데 반해 뉴로-퍼지 수위예측 모형의 경우 -10cm~10cm로 회귀분석 수위예측 모형에 비해 오차가 상당히 작게 나타나고 있음을 확인할 수 있다. 이를 통해 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 우수한 수위예측 능력을 보이고 있음을 확인할 수 있었다.

표 5. 정량적 평가를 위한 통계지표

통계지표	식
수정 결정계수 (R ^{*2})	$1 - \frac{N-1}{N-k-1} \frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$
평균제곱근오차 (RMSE)	$\sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}$
Nash-Sutcliffe 효율계수 (NSEC)	$\frac{\sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$
평균절대오차 (MAE)	$\frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i $

여기서 y_i 와 \hat{y}_i 는 각각 관측치와 예측치, \bar{y} 와 $\bar{\hat{y}}$ 는 관측치와 예측치의 평균, e_i 는 관측값과 예측값의 차, N 은 자료의 개수, k 는 모형 구축에 사용된 변수의 개수(본 연구에서는 10)이다.

본 연구에서 구축된 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형간의 4개의 검증 사상에 대한 수위 예측에 대한 정확도를 정량적으로 평가하기 위해 수정 결정계수(R^{*2}, adjusted coefficient of determination)와 평균제곱근오차(Root Mean Square Error), Nash-Sutcliffe 효율계수(Efficiency Coefficient), 평균절대오차(Mean Absolute Error)와 같은 통계지표를 사용하였다. 사용된 통계지표를 간단히 설명하면 표 5와 같다.

검증사상 4개의 수위예측 결과에 대해 평균제곱근오차, 효율계수, 평균절대오차, 수정 결정계수를 각각 계산하여 표 6에 나타내었다. 표에서 보는 바와 같이 평균제곱근오차의 경우 뉴로-퍼지 수위예측 모형은 0.5시간의 예측 선행시간에서 최소 0.011m, 3.0시간의 예측 선행시간에서 최대

표 6. 다중선형회귀 수위예측 모형간 성능 평가를 위한 통계지표 값의 비교

호우사상	모형	RMSE						NSEC					
		t+1	t+2	t+3	t+4	t+5	t+6	t+1	t+2	t+3	t+4	t+5	t+6
W-9	뉴로-퍼지	0.016	0.017	0.030	0.057	0.067	0.076	0.995	0.982	0.938	0.915	0.888	0.855
	회귀분석	0.023	0.046	0.066	0.080	0.094	0.106	0.990	0.959	0.918	0.876	0.829	0.783
W-10	뉴로-퍼지	0.011	0.018	0.019	0.022	0.023	0.024	0.990	0.972	0.971	0.961	0.957	0.951
	회귀분석	0.012	0.021	0.029	0.035	0.040	0.046	0.989	0.964	0.929	0.898	0.862	0.822
W-15	뉴로-퍼지	0.026	0.054	0.090	0.106	0.128	0.146	0.994	0.974	0.925	0.895	0.848	0.799
	회귀분석	0.044	0.076	0.091	0.112	0.133	0.153	0.982	0.947	0.924	0.885	0.836	0.781
W-16	뉴로-퍼지	0.011	0.018	0.032	0.038	0.041	0.051	0.995	0.986	0.957	0.939	0.926	0.888
	회귀분석	0.017	0.028	0.032	0.040	0.048	0.061	0.987	0.968	0.955	0.931	0.900	0.839
호우사상	모형	MAE						R ²					
		t+1	t+2	t+3	t+4	t+5	t+6	t+1	t+2	t+3	t+4	t+5	t+6
W-9	뉴로-퍼지	0.010	0.017	0.029	0.034	0.041	0.047	0.995	0.982	0.941	0.920	0.897	0.871
	회귀분석	0.013	0.026	0.039	0.049	0.058	0.067	0.992	0.968	0.943	0.921	0.897	0.873
W-10	뉴로-퍼지	0.006	0.011	0.011	0.013	0.014	0.015	0.990	0.973	0.971	0.960	0.956	0.950
	회귀분석	0.006	0.013	0.019	0.024	0.029	0.035	0.989	0.964	0.929	0.902	0.875	0.851
W-15	뉴로-퍼지	0.014	0.027	0.045	0.053	0.064	0.075	0.994	0.973	0.927	0.897	0.853	0.809
	회귀분석	0.019	0.034	0.047	0.059	0.072	0.083	0.982	0.947	0.929	0.894	0.849	0.798
W-16	뉴로-퍼지	0.006	0.010	0.014	0.017	0.019	0.023	0.995	0.988	0.958	0.939	0.927	0.891
	회귀분석	0.007	0.013	0.018	0.026	0.035	0.046	0.987	0.967	0.956	0.935	0.909	0.855

0.146m의 오차를 나타내고 있는데 반해 회귀분석 수위예측 모형의 경우에는 0.5시간의 예측 선행시간에서 최소 0.012m, 3.0시간의 예측 선행시간에서 0.153m로 뉴로-퍼지 모형에 비해 큰 평균제곱근오차를 보이고 있음을 알 수 있다.

Nash-Sutcliffe 효율계수의 경우에도 모든 선행시간에 대해 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 최소 0.001에서 최대 0.129 정도의 높은 값을 나타내고 있음을 확인할 수 있었다. 평균절대오차와 수정 결정계수에 대한 두 모형간 결과를 비교해보면 예측 선행시간별로 다소 차이는 있으나 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 우수한 수위 예측 성능을 보이고 있음을 확인할 수 있다.

5. 결 론

본 연구에서는 기존의 개념적, 수문학적, 물리적 기반 모형 등의 단점을 극복하고자 최근 수위예측을 위해 널리 사용되고 있는 자료지향형 모형 중 뉴로-퍼지와 회귀분석 모형을 이용하여 수위예측 모형을 구축하고 실제하천에 적용하여 모의결과를 비교·검토하였다. 본 연구의 결과를 요약하면 다음과 같다.

- (1) 뉴로-퍼지와 회귀분석 수위예측 모형 구축을 위해 수위와 강우의 입력자료에 대한 시간분포를 동일하게 한 후 모형을 구축하였다. 뉴로-퍼지 수위예측 모형 구축을 위해서는 모의 시간의 단축 및 최적의 입력 공간 분할을 위해서 subtractive clustering 기법을 추가적으로 적용하였으며, 회귀분석 수위예측 모형 구축을 위해서는 모형의 유의성을 검토하기 위해 분산분석을 통한 회귀성 검정을 수행하고 개별회귀계수에 대한 t-검정을 수행하여 모형을 구축하였다.
- (2) 구축된 뉴로-퍼지 수위예측 모형과 회귀분석 수위예측 모형을 왕숙천 유역의 다양한 사상에 대해 적용하였다. 두 모형간의 수위

예측에 대한 성능을 비교분석하기 위해 각 선행시간별 수위예측 결과와 관측치와의 오차를 그림을 통해 비교하였다. 비교 결과 뉴로-퍼지 수위예측 모형이 회귀분석 수위예측 모형에 비해 우수한 수위 예측능력을 보이고 있음을 확인할 수 있었다.

- (3) 본 연구에서 제시한 두 모형의 수위 예측에 대한 성능을 정량적으로 평가하게 위해 평균제곱근오차, Nash-Suttcliffe 효율계수, 평균절대오차, 수정 결정계수와 같은 4개의 통계지표들을 산정하여 제시하였다. 모의결과 모든 예측 선행시간에 대해 뉴로-퍼지를 이용한 수위 예측 모형이 회귀분석 수위예측 모형보다 우수한 예측 결과를 보이고 있음을 확인할 수 있었다. 이러한 결과는 실제 수문학적 과정에서의 복잡성과 강우-유출의 비선형적 관계를 뉴로-퍼지 수위예측 모형의 경우 비선형의 소속함수를 도입함으로써 모의를 수행하는데 반해 회귀분석 수위예측 모형의 경우 입력력의 선형적인 관계를 도입하여 모의를 수행하는데 따른 것으로 판단된다. 이러한 결과는 향후 도달시간이 짧아 충분한 선행시간을 확보하기 어려운 중소하천, 특히 도심하천의 홍수 예경보에 활용될 수 있을 것으로 판단된다.

감사의 글

본 연구는 국토해양부가 출연하고 한국건설교통기술평가원에서 위탁 시행한 건설기술혁신사업(08기술혁신F01)에 의한 차세대홍수방어기술개발연구단의 연구비 지원에 의해 수행되었습니다.

참 고 문 헌

고영준. 2001. 하천의 시유출량 예측을 위한 퍼지 신경회로망의 적용. 석사학위 논문, 전남대학교. pp. 1-6, pp. 14-17.

- 건설교통부. 2005. 주요지천 홍수예보 프로그램 개선.
- 신지연, 허준행. 2009. 통계적 모형을 이용한 지천 홍수예측. 한국수자원학회 학술발표회논문집, pp. 1524-1527.
- 신성일. 2002. 신경회로망과 뉴로-퍼지를 이용한 홍수량예측에 관한 연구. 석사학위논문, 경일대학교.
- 윤강훈, 김태균. 2003. 레이더 예측 강우를 이용한 다중회귀 유출 모형의 개발. 대한토목학회 학술발표회 논문집, pp. 2187-2198.
- 윤강훈, 신현민. 1994. 한강인도교 수위와 영향인자간의 다중회귀분석에 의한 홍수위 예측모형. 한국수자원학회지, 제27권, 제3호, pp. 55-69.
- 윤용남, 원석연. 1991. 미계측 중소유역의 월유출량 산정을 위한 다중회귀모형 연구. 한국수자원학회지, 제24권, 제3호, pp. 71-82.
- 이범희. 2010. 회귀분석에 의한 도시홍수 예보시스템의 개발. 대한토목학회논문집, 30(4-B): 347-359.
- 정동국, 이범희. 2009. 회귀기법을 이용한 도시홍수위 예측모형의 개발. 한국수자원학회논문집, 43(2): 221-231.
- 진영훈. 2000 하천의 유출량 예측을 위한 인공신경망 이론의 적용. 석사학위 논문. 전남대학교.
- 정대명. 2003 월 담유입량 예측을 위한 뉴로-퍼지시스템의 적용. 석사학위 논문, 세종대학교. pp. 2-5.
- Aquil, M., Kita, I., Yano, A., and Nishiyama, S. 2007. A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behavior of runoff. Journal of Hydrology, 337(1-2), 22-34
- Brown, M. and Harris, C. 1994. Neuro-fuzzy adaptive modeling and control. Prentice Hall, New York.
- Chaoulakou, A., Assimacopoulos, D. and Lekkas, T. 1999. Forecasting daily maximum ozon concentration in the Athens basin. Environmental Monitoring and Assessment, 56: 97-112.
- Chang, F.J., and Chen, Y.C. 2001. A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. Journal of Hydrology, 245: 153-164.
- Firat, M., Turan, M.E., and Yurdusev, M.A. 2009. Comparative analysis of fuzzy inference systems for water consumption time series prediction. Journal of Hydrology, 374: 235-241.
- Gautam, D.K., and Holz, K.P. 2001. Rainfall-runoff modeling using adaptive neuro-fuzzy systems, Journal of Hydroinformatics, pp. 3-10.
- Govindaraju, R.S. 2000. Artificial neural networks in hydrology II: hydrologic applications. Journal of Hydrologic Engineering, 5(2): 124-137.
- Huynh N. P. and Nguyen D. A. K. 2003. Flood forecasting for the upper reach of the Red River Basin, North Vietnam. <http://www.wrc.org.za>
- Jang, J.-S. 1992. Self-learning fuzzy controllers based on temporal backpropagation. IEEE Trans Neural Netw, 3(5): 714-723.
- Kokkonen, T.S. and Jakeman, A.J. 2001. A comparison of metric and conceptual approaches in rainfall-runoff modeling and its implications. Water Resources Research, 37(9): 2345-2353.
- Kurtulus, B. and Razack, M. 2009. Modeling daily discharge responses of a large karstic aquifer using soft computing methods Artificial neural network and neuro fuzzy. Journal of Hydrology, 375: 146-162.

- McKerchar, A.I. and Delleur, J.W. 1974. Applications of seasonal parametric linear stochastic models to monthly flow data. *Water Resources Research*, 10: 246-255.
- Nayak, P.C., Sudheer, K.P., Rangan, D.M., and Ramasastri, K.S. 2005. Short-Term Flood Forecasting with a Neurofuzzy Model. *Water Resources Research*, 41(4), W04004.
- Phien, H.N., Huong, B.K. and Loi, P.D. 1990. Daily forecasting with regression analysis. *Water SA*, 16(3): 179-184.
- Schilling, K.E. and Wolter, C.F. 2005. Estimation of streamflow, baseflow and nitrate-nitrogen loads in Iwoa using multiple regression models. *Journal of American Water Resources*
- Shu, C. and Ouarda, T.B.M.J. 2008. Regional flood frequency analysis at ungauged sites using the adaptive neuro fuzzy inference system. *Journal of Hydrology*, 349: 31-43.
- Tangborn, W.V. and Rasmussen, L.A. 1976. Hydrology of North Cascades region, Washington-part 2: a proposed hydrometeorological streamflow prediction method. *Water Resources Research*, 12: 203-216.
- VanderKwaak, J.E. and Loague, K. 2001. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resources Research*, 37(4): 999.
- Yurekli, K., Kurung, A. and Ozturk, F. 2005. Testing residuals of an ARIMA model on the Cekerek stream watershed in Turkey. *Turkish Journal of Engineering and Environmental Sciences*, 29: 61-74.

- 논문접수일 : 2011년 08월 30일
- 심사의뢰일 : 2011년 08월 31일
- 심사완료일 : 2011년 11월 29일