

입력자료의 불확실성에 따른 다목적 댐 유입량 해석기법

○ 김 성 원*

1. 서론

공학적인 문제의 불확실성을 해석하는데 확률론적 방법, 퍼지이론, 그레이이론(Gray Method) 및 구간산술법(Interval Arithmetic Method)등을 포함한 많은 방법론과 이론이 제시되어 왔다(Moore, 1979; Ayyub and Gupta, 1997; Guyonnet, 1999). 그러나 엄격한 수학적 기초를 둔 확률론적 이론이 불확실성 예측 및 모형화과정에 가장 유용하게 사용되고 있는 실정이다. 확정론적 혹은 확률론적 예측 및 모형화과정과 마찬가지로 신경망을 기본으로 하는 예측 및 모형화과정은 불확실성과 연계가 되어 있다. 이러한 모형화의 불확실성은 모형화 과정뿐만 아니라 신경망의 훈련에 사용된 제한된 자료에 기인한다고 할 수 있다. 일반적으로 수공학에서 모형화의 불확실성은 2가지 요인에 달려있다. 첫째는 유출을 발생시키려는 모형의 종류이고, 두 번째는 모형의 배열(Arrangement)이다.

본 연구에서는 낙동강 상류유역의 병렬 다목적댐군인 안동 및 입하다목적댐의 유입량을 산정하기 위하여 추계학적 모형과 신경망모형을 연계한 공간추계 신경망모형을 다소 수정한 변형된 공간추계 신경망모형(MSS-NNM)을 구성하였다. 그리고 변형된 공간추계 신경망모형의 훈련에 사용될 자료는 추계학적 모형의 모의발생년을 단기간, 중기간 및 장기간으로 모의발생하여 훈련자료의 불확실성에 따른 다목적 댐 유입량의 불확실성을 해석하였다. 즉 추계학적 모형으로 신경망 모형의 훈련자료를 모의발생년에 따라 모의발생시키고, 모의발생된 각 자료를 이용하여 신경망모형을 훈련시켜 모의발생년에 따른 최적 매개변수를 산정한다. 그리고 산정된 매개변수를 이용하여 신경망모형을 검증하여 입력자료의 불확실성에 따른 병렬 다목적댐군의 유입량을 검토하고자 한다.

2. 변형된 공간추계 신경망모형(Modified Spatial-Stochastic Neural Networks Model, MSS-NNM)

2.1 일반화된 회귀신경망모형(GRNNM)의 개요

본 연구에 사용된 일반화된 회귀신경망모형(GRNNM)은 반경기초함수 신경망모형(RBF-NNM)의 변형된 형태의 신경망 모형이다(Caudill, 1993; Specht, 1991; Tsoukalas and Uhrig, 1997). GRNNM은 입력층, 은닉층, 합산층 및 출력층의 4개의 층으로 구성되었으며, 비선형 회귀이론에 기초를 둔 신경망모형이다. 입력층, 은닉층 및 합산층노드는 완전히 연결된 구조로 형성되어 있으나, 출력층의 노드는 합산층의 노드 중 몇 개만 연결이 되어있다. 합산층(Summation Layer)은 여러 개의 합산노드(Summation Node)와 하나의 분지노드(Diversion Node)의 2종류의 노드로 구성되어 있다. 합산노드수는 출력층 노드수와 동일하며, 합산층 노드의 전이함수는 RBF-NNM의 출력층의 전이함수와 본질적으로 동일하다. 분지노드는 어떤 전이함수를 사용하지 않고 은닉층노드의 가중전이값(Weighted Transfer Value)을 합한 것과 같다. GRNNM의 각 출력층노드는 합산층의 합산노드와 분지노드로 연결되어 있으며, 합산층과 출력층 사이는 연결강도로 구성되어 있지 않다. 각 출력층노드의 계산은 합산층의 합산노드로부터 계산된 값을 분지노드로부터 계산된 값으로 나누어 산정한다. 그리고 합산층과 출력층은 근본적으로 출력층 결과치의 정규화과정을 실시하여 GRNNM이 RBF-NNM보다 은닉층노드수의 적절한 선택에 대한 민감도를 감소시키는 역할을 한다. 따라서 각 은닉층노드 반경기초함수의 껍침 현상은 GRNNM에서는 큰 문제가 되지 않는다. GRNNM의 훈련과정은 RBF-NNM과는 상당히 다른 양상을 가지고 있다. 즉 은닉층노드 반경기초함수의 중심과 은닉층노드와 합산층노드의 연결강도는 동시에 산정된다. 따라서 은닉층의 훈련은 RBF-NNM의 경우와 같이 자율

* 정희원, 동양대학교 지구환경시스템공학과, 전임강사(E-mail : swkim68@phenix.dyu.ac.kr)

훈련(Unsupervised Training)으로 구성되고, 특별한 군집알고리즘이 필요로 하며 군집알고리즘의 변경은 훈련이 시작하기 전에 결정되어 져야 한다. 각 은닉층노드의 반경기초함수의 폭은 은닉층노드 반경기초함수의 겹침 현상에 대해 GRNNM의 낮은 민감도현상 때문에 은닉층의 나타난 구조에 따라 설정될 필요는 없다. 위의 과정에 의한 은닉층이 차례차례로 구성되는 것과 동시에 은닉층과 합산층의 연결강도의 값은 지도훈련(Supervised Training)을 통하여 설정된다. 각 은닉층노드와 합산층의 개별적인 합산노드사이의 연결강도는 특별한 은닉층노드 반경기초함수의 중심을 형성한 출력값에 따라서 동일한 값으로 산정되어 있다. 훈련자료의 부가적인 입력값들이 동일한 은닉층노드에 할당된다면, 대응하는 출력값들은 그들 연결강도의 이전 값에 더해진다. 동시에 초기 값으로 영(Zero)로 설정된 각 은닉층노드와 분지노드의 연결강도는 은닉층노드에 할당된 훈련자료로부터 각 입력값에 의해서 하나씩 증가된다. 다음 Fig. 1은 본 연구에 적용될 GRNNM의 구조를 나타낸 것이다.

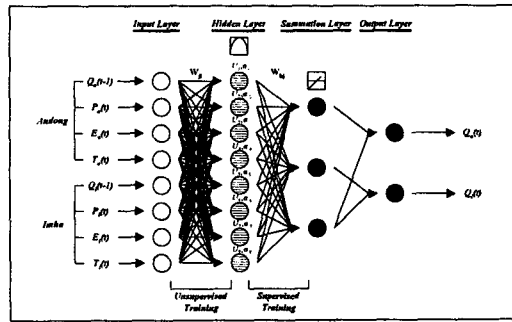


Fig. 1 Proposed Generalized Regression Neural Networks Model

2.2 GRNNM의 구성과정

GRNNM의 입력층과 은닉층사이의 연결함수는 입력벡터요소 x_i 에서 중심(Centers)인 u_{ji} 를 제한 값의 제공에 대한 평방근의 합으로 나타내고 있으며 다음 식[1]과 같이 나타낼 수 있다.

$$R_j = \sqrt{\left(\sum_{i=1}^m (x_i - u_{ji})^2 \right)} \quad [1]$$

여기서 아랫첨자 i, j 는 입력층과 은닉층을 나타내고 있으며, R_j 는 입력벡터들 사이의 반경거리(Radial distance)를 나타낸다. 따라서 입력벡터를 $X = [x_1, x_2, \dots, x_m]^T$ 라 나타내고, 입력층과 은닉층의 연결강도벡터를 보통 반경중심(Radial centers) 혹은 중심(Centers)으로 표시하며 $U^{(j)} = [u_{j1}, u_{j2}, \dots, u_{jm}]^T$ 라고 나타내면 식[1]은 식[2]와 같이 나타낼 수 있다.

$$R_j = \|X - U^{(j)}\| \quad [2]$$

여기서 $\|\cdot\|$ 는 유클리드거리(Euclidean distance)이다. 또한 R_j 를 은닉층의 전이함수인 $\Phi_1(\cdot)$ 에 적용하면 다음 식[3]과 같이 나타낼 수 있다.

$$S_j = \Phi_1(R_j) = \Phi_1(\|X - U^{(j)}\|) \quad [3]$$

보통 은닉층의 전이함수 $\Phi_1(\cdot)$ 는 RBF를 사용하며, 일반적으로 RBF의 종류로는 Thin Plate Spline Function, Gaussian Kernel Function, Multiquadric Function, 및 Inverse-Multiquadric Function등이 있다(Mason et al., 1996). 본 연구에서는 Gaussian Kernel Function(GKF)을 사용하였으며 입력벡터의 값이 0일 때 최대값은 1이고, 입력벡터와 중심사이의 거리가 감소함에 따라 출력값은 증가한다. GKF는 다음 식[4]와 같이 나타낼 수 있다.

$$\Phi_1(\cdot) = \exp(-B_1 R^2) = \exp\left[-\frac{\sum_{i=1}^m (x_i - u_{ji})^2}{2\sigma^2}\right] \quad [4]$$

여기서 B_1 는 $\frac{1}{2\sigma^2}$ 으로 나타낼 수 있으며 보통 일정한 값을 가지며, σ 는 RBF의 폭을 나타낸다.

또한 합산층의 합산노드와 분지노드는 S_j 와 은닉층과 합산층의 연결강도벡터의 연산 결과치를 입력벡터로 받으며, 다음 식[5]와 같이 나타낼 수 있다.

$$T_k = \sum_{j=1}^n W_{kj} \cdot S_j = \sum_{j=1}^n W_{kj} \cdot \Phi_1(\|X - U^{(j)}\|) \quad [5]$$

여기서 W_{kj} 는 은닉층과 합산층사이의 연결강도벡터이고, 각 k에 대하여 식[5]는 반경기초함수를 나타내고 있으며, 합산층의 전이함수 $\Phi_2(\cdot)$ 에 나타내어 합산층의 결과치를 표시하면 다음 식[6]-[8]과 같이 나타낼 수 있다.

$$SN_1(t) = \Phi_2\left(\sum_{j=1}^n W_{1j} \cdot \Phi_1(\|X - U^{(j)}\|)\right) \quad [6]$$

$$SN_2(t) = \Phi_2\left(\sum_{j=1}^n W_{2j} \cdot \Phi_1(\|X - U^{(j)}\|)\right) \quad [7]$$

$$DN(t) = \Phi_2\left(\sum_{j=1}^n W_{3j} \cdot \Phi_1(\|X - U^{(j)}\|)\right) \quad [8]$$

여기서 SN_1 은 합산층 첫 번째 합산노드의 결과치, SN_2 은 합산층 두 번째 합산노드의 결과치이고 DN 은 합산층 분지노드의 결과치이다. 합산층의 전이함수 $\Phi_2(\cdot)$ 는 전이함수 중에서 선형전이함수(Pure Linear Transfer Function, PLTF)를 사용하였다. 그리고 출력층의 결과치를 표시하면 다음 식[9]-[10]과 같이 나타낼 수 있다.

$$Q_a(t) = \frac{SN_1(t)}{DN(t)} \quad [9]$$

$$Q_i(t) = \frac{SN_2(t)}{DN(t)} \quad [10]$$

여기서 $Q_a(t)$ 는 안동 다목적댐의 유입량을 나타낸 것이고 $Q_i(t)$ 는 임하 다목적댐의 월평균유입량을 나타낸 것이다. 다음 그림 2는 GRNNM 은닉층의 입·출력과정 및 GKF를 나타낸 것이다.

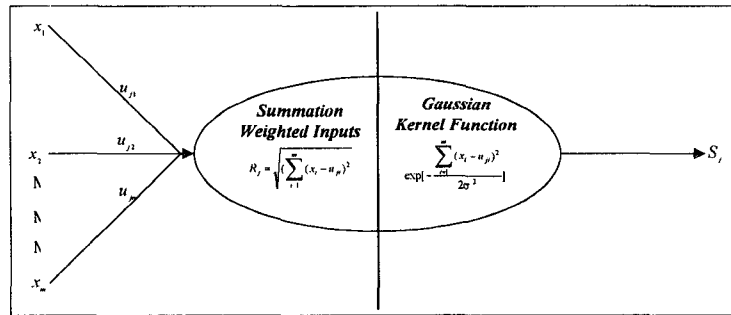


Fig. 2 Hidden Layer in GRNNM(김성원 외, 2001)

3. 추계학적 모형(Stochastic Model)

본 연구의 공간추계 신경망모형의 훈련자료는 부족한 자료 시계열을 모의발생시킨 자료를 이용하였다. 따라서 부족한 자료시계열은 추계학적 모형의 일종인 PARMA(1,1)모형에 의하여 월평균유입량, 월면적강우량, 월별 증발접시 증발량과 월평균 기온자료를 모의발생하였으며, 제시된 PARMA(1,1)모형은 다음 식[11]과 같이 나타낼 수 있다(김성원, 2003).

$$y_{\nu, \tau} = \mu_{\tau} + \phi_1(y_{\nu, \tau-1} - \mu_{\tau-1}) + \varepsilon_{\nu, \tau} - \theta_1 \varepsilon_{\nu, \tau-1} \quad [11]$$

여기서 ν 는 년(Year), τ 는 계절(Season)이고 $\tau = 1, 2, \dots, \omega$ 이며 계절의 수를 나타내고 있다. 일반적으로 PARMA(1,1)모형은 월유출량을 모의하는데 사용된다(Salas et al., 1980). 본 연구의 PARMA(1,1)모형은 SAMS(Salas J. D., 1998)

Package를 이용하였고, 계절의 수는 12계절로 하였다. 그리고 모의발생한 표본(Sample)은 250년의 단기간(Short-term), 550년의 중기간(Mid-term), 1050년의 장기간(Long-term)으로 각각 2개의 표본이며, 월평균유입량, 월면적강우량, 월별 증발접시 증발량과 월평균 기온자료를 모의발생하였다. 모의된 자료중 첫 번째 표본자료에서 초기에 발생하는 편차를 제거하기 위하여 초기의 50년 자료는 제거하고 각각 200년, 500년, 1000년의 자료만을 이용하여 변형된 공간추계 신경망모형의 훈련 자료로 사용하였다.

4. 다목적 댐 유입량자료의 불확실성 분석

본 연구의 변형된 공간추계 신경망모형의 각 훈련자료를 이용하여 훈련결과 반경기초함수의 중심과 폭 및 연결강도벡터를 산정하였다. 또한 산정된 매개변수를 이용하여 변형된 공간추계 신경망모형의 검증결과 발생된 자료에 따른 다목적 댐 유입량자료의 불확실성을 검토하였으며, 발생된 자료의 수가 많을수록 변형된 공간추계 신경망모형의 훈련 및 검증결과가 양호한 것으로 나타났다.

5. 결론

본 연구에서는 변형된 공간추계 신경망모형을 이용하여 훈련자료의 모의발생년이 단기간, 중기간 및 장기간에 대하여 최적매개변수인 반경기초함수의 중심과 폭 및 연결강도벡터를 산정하였으며, 산정된 매개변수를 이용하여 변형된 공간추계 신경망모형을 검증하였다. 검증결과 변형된 신경망모형의 훈련자료의 발생년에 따른 다목적 댐 유입량의 불확실성을 검토하였으며, 상관성을 유도하였다. 따라서 변형된 공간추계 신경망모형은 병렬다목적 댐 유입량의 불확실성 해석을 위한 잠재성이 내포된 모형이며, 지금까지의 연구결과로서 신경망모형은 어떤 복잡한 함수적인 응답(Functional Response)의 접근성에 장점이 있는 것으로 검증되었기 때문에 본 연구에서 제시된 변형된 공간추계 신경망모형은 복잡한 병렬 다목적 댐 유입량의 산정에 있어서 적용성이 높은 것으로 분석되었다. 따라서 본 연구를 기초로 하여 변형된 공간추계 신경망모형을 이용한 병렬다목적댐 운영을 개선을 위한 지속적인 연구가 필요한 것으로 사료된다.

참고문헌

1. 김성원(2003). "추계확적모형과 신경망모형을 연계한 병렬저수지군의 유입량산정.", **한국수자원학회 논문집 게재예정.**
2. 김성원, 이순탁, 조정석(2001). "중소하천유역에서 Hybrid Neural Networks에 의한 수문학적 예측.", **한국수자원학회 논문집**, 제 34권, 제 4호, pp. 303-316.
3. Ayyub, B.M., and Gupta, M.M.(1997). *Uncertainty analysis in engineering and sciences : Fuzzy logic, statistics and neural network approach*, Kluwer Academic, Boston.
4. Caudill, M.(1993). "GRNN and Bear it." *AI Expert*, Vol. 8, No. 5, pp. 28-33.
5. Guyonnet, D., Come, B., Perrochet, P., and Parriaux, A.(1999). "Comparing two methods for addressing uncertainty in risk assesment." *J. of Environ. Eng.*, Vol. 125, No. 6, pp. 660-666.
6. Mason, J.C., Price, R.K., and Tem'ne, A.(1996). "A neural network model of rainfall-runoff using radial basis functions." *Hydro. Sci. J.*, Vol. 41. No. 3, pp. 399-417.
7. Moore, R.E.(1979). *Method and applications of interval analysis*, SIAM, Philadelphia.
8. Salas, J.D.(1998). *SAMS ; Stochastic analysis, modeling, and simulation user manual*, Colorado State University, Fort Collins, CO.
9. Salas, J.D., Delleur, J.R., Yevjevich, V., and Lane, W.L.(1980). *Applied modeling of hydrologic time series*, Water Resources Publications, Littleton, CO.
10. Specht, D.F.(1991). "A general regression neural network." *IEEE Transactions on Neural Networks*, Vol. 2, No. 6, pp. 568-576.
11. Tsoukalas, L.H., and Uhrig, R.H.(1997). *Fuzzy and neural approaches in engineering*, John Wiley & Sons Inc.