

강수에 측을 위한 퍼지시계열과 뉴로-퍼지 시스템의 적용

나 창 진* / 김 형 수** / 김 중 훈*** / 윤용남****

1. 서 론

물의 순환에 대한 연구, 즉 강수 발생에 관한 연구는 많은 현상 중 특히 관심의 대상이 되어 왔으며 이에 관한 분석 및 연구는 수공학에 있어 기본이 되는 연구활동이라 할 수 있다.

많은 연구자들이 이러한 강수량의 예측의 중요성을 알고 있음에도 불구하고 기존에 그러한 연구가 타 과제에 비하여 적었던 것은 강수의 발생이 일정한 특성 없이 발생함에 그 원인이 있다. 또한 강수에 영향을 주는 인자 역시 불규칙하게 끊임없이 발생한다. 이러한 이유로 기존의 연구의 경우 대략 24시간전의 혹은 12시간, 6시간의 단기간의 강수를 예측하는 것이 대부분이거나 전세계나 큰 대륙의 강수의 거동변화에 대한 연구가 있어왔다. 그러나 수자원 관리 및 계획을 위해서는 월 강수량이나 계절강수량을 정확히 예측하는 것이 필요하다.

효과적인 수자원 계획을 수립하기 위해서는 계절 강수량이나 월강수량의 예측 및 변동에 관한 연구가 이루어져야 한다. 그러나 강수의 예측은 완전할 수는 없는 신의 영역이며 단지 어느 정도의 예측을 통하여 홍수나 가뭄과 같은 자연 재해로부터 피해를 줄이는데 그 목적이 있다고 할 수 있다.

비선형 기법을 이용하여 수문현상의 적용은 최근 다양한 접근이 이루어지고 있으며, 기존의 경우 신경망 기법을 이용한 강우-유출에 관한 연구 및 강수량의 예측등이 주를 이루고 있다(Furundzic, 1998; Lange, 1998; Sajikumar, 1999; Bodri, 2000; Luk, 2001). 또한 비선형 기법인 퍼지를 수문 및 기상에 적용한 연구 역시 많은 부분 진행되고 있다(Ertunga, 1996; Franks, 1998; Ouenes, 2000).

그러나, 뉴로-퍼지 시스템의 수문에의 적용은 아직 많은 연구가 진행되지 않은 상태이다. Chang(2001) 및 Gautam(2001)은 뉴로-퍼지 시스템을 이용하여 강우-유출 모형을 적용한 바 있다.

퍼지 시계열은 Song과 Chissom(1993, 1994)에 의하여 처음으로 정의되어 사회 및 경제 분야의 예측에 활용되었다. Sullivan(1994)는 Song과 Chissom의 모형과 Markov 모형을 비교하였다. Chen(1996, 1998)은 Song과 Chissom의 퍼지 시계열의 개념을 이용하여 새로운 방안을 제시하여 퍼지 시계열의 개념 및 적용을 재정립하였다. Huang(2001)은 Chen의 퍼지 시계열에 Heuristic 개념을 이용하여 예측의 능력을 높인 퍼지 시계열을 제안한바 있으나 수문 분야에는 아직 적용된 바가 없다.

본 연구에서는 이러한 기후 현상의 상호작용을 근간으로 강수에 영향을 미치는 인자를 이용하여 효과적인 강수량 예측 모형 작성하였다. 이를 위하여 뉴로-퍼지 시스템 및 퍼지 시계열을 사용하였다.

2. Heuristic 퍼지 시계열

2.1 퍼지 시계열

퍼지 시계열(fuzzy time series)은 Song과 Chissom(1993)에 의하여 처음으로 정의되어 예측에 활용되었다. Chen(1996)은 Song과 Chissom의 퍼지 시계열의 개념을 이용하여 새로운 방안을 제시하고, 그 개념 및 적용을 재정립하였다.

퍼지 시계열의 개념은 먼저 전통적인 시계열과 비교 할 수 있는데, 전통적인 시계열은 실험 또는 관측치들을 시간에 따라 구성하는데 반해 퍼지 시계열은 퍼지 집합과 fuzzy boundary에 의해 시계열이 구성된다.

2.2 뉴로-퍼지 시스템

뉴로-퍼지 시스템은 기존의 퍼지 논리 시스템(fuzzy logic system)에 신경망의 학습능력을 도입한 것이다. 즉 전문가지식기반의 퍼지논리 시스템에 학습이라는 유연한 기능을 부가하여 기존의 개념으로 해결하기 어려운 점들을 풀어가는 방식이다(Jang, 1996). 본 연구에서 사용하고자 하는 ANFIS 모형은 Jang(1993)에 의해 제안된 뉴로-퍼지 모델로서 1차 Sugeno 시스템에 관한 퍼지 시스템 파라미터들을 최적화시키기 위한 혼합된 학습 규칙들을 갖는 1차 Sugeno 모형을 말한다.

* 고려대학교 토목환경공학과 석사과정
** 선문대학교 건설토목환경공학과 부교수
*** 고려대학교 토목환경공학과 부교수
**** 고려대학교 토목환경공학과 교수

3. 적용 방법

3.1 Chen의 퍼지 시계열을 이용한 강수량의 예측(case 1)

Chen의 예측 모형은 앞서 언급되었으며, Chen의 모형을 강수량 예측에 적용하기 위한 절차는 다음과 같다. 그림 1.은 다음의 절차를 도식화 한 것이다.

- 1) 강수량 자료계열의 상한과 하한을 결정한다. 이는 자료계열의 minimum값과 maximum값을 산정후 조정한다.
- 2) 결정된 상한과 하한을 일정 간격으로 분류하고 각 구간에 대한 언어적인 값을 부여한다. 즉 구간의 강수에 많다 혹은 적다 등의 의미를 부여한다. 즉 시계열을 fuzzify 한다.
- 3) fuzzify 한 시계열의 경우 시점 t-1에서 t로 가는 logic을 분류하여 그룹화 한다.
- 4) 예측 기간의 시계열을 fuzzify 한다.
- 5) 과정 3)에서 만든 그룹을 예측 구간에 적용하여 t 시점의 강수량을 예측한다.
- 6) Chen의 모형의 경우 예측시 분류된 구간의 중간값의 평균을 활용한다.

3.2 Heuristic 퍼지 시계열을 이용한 강수량의 예측(case 2)

Huang의 예측 모형은 Chen의 모형에 일부를 수정하였을 뿐 모형은 Chen의 모형을 그대로 사용한다. Huang은 자료계열의 그룹화시 heuristic 기법을 사용하여 적절한 그룹화를 위한 기준을 선정하였다. 이는 모형의 예측 능력을 증대시킬 수 있다. 그림 2.는 Huang에 의하여 수정된 Chen의 퍼지 시계열 모형을 간략히 설명하고 있다.

먼저 자료계열의 그룹화를 완성 후 그룹을 시점 t-1과 t를 비교하여 오차의 정도를 가지고 다시 그룹화한다. 그러나 Huang의 예측 모형의 경우 시점 t와 t-1의 오차정도가 모형에 들어가므로 진정한 의미의 예측이라 할 수 는 없다. 그러므로 본 연구에서는 예측 과정에 있어 case 2의 확률개념을 활용한다.

시점 t-1의 상태가 A_7 일 경우 시점 t에서 $\{A_8, A_5, A_7, A_9\}$ 의 상태가 될 수 있다고 가정할 경우, t와 t-1의 오차 정도에 따라 A_7 그룹은 $A_7 \rightarrow \{A_8, A_7, A_9\}$ $A_7 \rightarrow \{A_8, A_5, A_7, A_9\}$ $A_7 \rightarrow \{A_5, A_7\}$ 로 다시 그룹화된다.

Huang의 모형의 경우 세부 그룹을 알고 예측을 하나 이는 진정한 의미의 예측이라 할 수 없으므로 세부 그룹에 대한 확률 즉, A_7 그룹의 경우 3/9, 4/9, 2/9의 확률을 갖는다. 이러한 확률개념을 이용하여 출력된 3개의 output을 확률평균한다. 각 모형에 output은 Chen의 모형과 마찬가지로 구간의 중간값을 평균하는 방법을 이용한다.

3.3 뉴로-퍼지 시계열을 이용한 강수량의 예측(case 3)

Chen이나 Huang의 모형의 경우 퍼지 시계열을 이용하여 예측의 가능성을 보이고 있으나 정작 실제적인 예측의 방법론에 있어 평균법을 이용하는 등 빈약함을 보이고 있다. 그래서 본 연구에서는 이러한 문제를 보완하기 위하여 예측에 있어 뉴로-퍼지 모델을 사용한다. 이는 Chen의 방법에 뉴로-퍼지 모델을 이용하여 예측하는 방법과 Huang의 과정에 뉴로-퍼지 모델을 이용하는 방법으로 분류된다. Chen의 모형의 경우 그룹화 시킨 계열을 그림 3.과 같이 전 단계의 입력을 통하여 출력을 발생시키는 방식으로 예측을 한다(Chen case 3). Huang의 과정의 경우 case 2에서 설명한 바와 같이 확률개념을 이용하여 case 2와 같은 방식으로 강수량을 예측한다(Huang case 3). 그림 3. Huang의 과정에 뉴로-퍼지 모델을 활용한 모형(Huang case 3)을 보이고 있다.

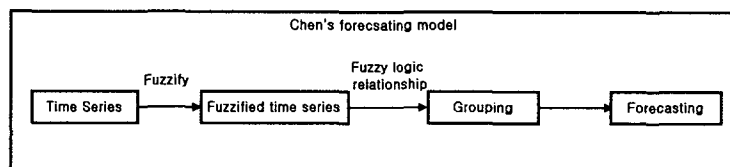


그림 1. Chen의 퍼지 시계열 예측 모형

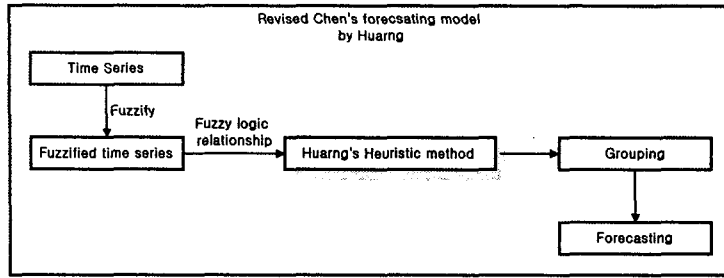


그림 2. Huarng의 퍼지 시계열 예측 모형

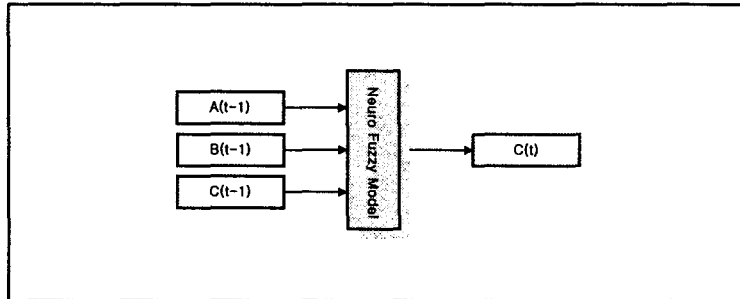


그림 3. 뉴로-퍼지를 이용한 강수량 예측 모형

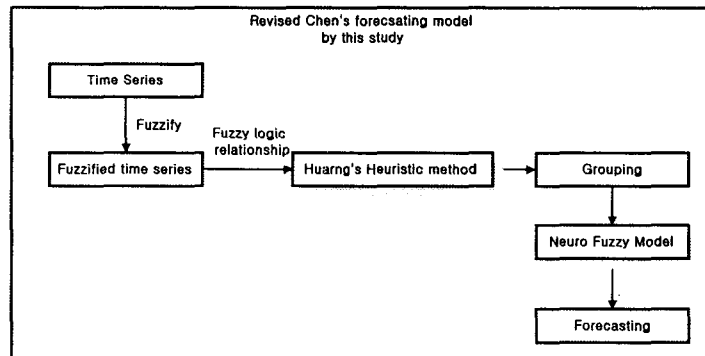


그림 4. 뉴로-퍼지 시계열 예측 모형

4. 적용 결과

4.1 미국 일리노이주에 적용

4.1.1 Chen의 퍼지 시계열을 이용한 강수량의 예측(case 1, 3)

퍼지 시계열을 강수량 예측에 적용하기 위하여 Chen(1993)에 의하여 제안된 Chen의 퍼지 시계열 모형을 적용하였다.

Chen의 모델에서 평균법을 이용한 예측은 전 단계 강수량 자료에서 전이 가능한 상태의 총 평균으로 나타난다. Chen의 모델에 의한 강수량 예측 결과는 그림 5와 같다.

Chen의 평균법 예측 모형의 예측 정도를 높이기 위하여 그룹별 뉴로-퍼지 모델을 구성하여 Chen case 3 모형을 작성한다. 이때 입력으로는 강수량, 기온, 토양수분이 들어간다. 또한 검정 기간의 분류 상태인 표 1에 따라 구축된 그룹별 모형을 적용하여 강수량을 예측할 수 있다. 그림 6.에서는 그 예측 결과를 보이고 있다

표 1. 1996년 강수량 자료계열 분류(case 4, 5, 6)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
그룹	A ₅	A ₂	A ₂	A ₄	A ₃	A ₁₄	A ₁₆	A ₁₀	A ₃	A ₁₀	A ₄	A ₄	A ₆	A ₃	A ₇

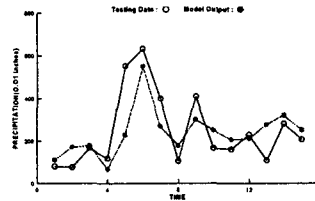
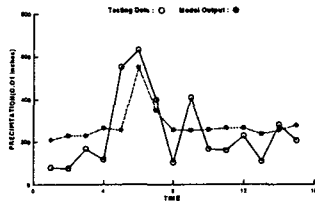


그림 5. case 1 모형에 의한 일리노이주 강수량 예측 그림 6. Chen case 3 모형에 의한 일리노이주 강수량 예측

4.1.2 Heuristic 퍼지 시계열을 이용한 강수량의 예측(case 2, 3)

Huang(2001)이 제안한 Heuristic 퍼지 시계열은 Chen의 퍼지 시계열 그룹화를 자세히 하여 Chen의 예측 모형의 정도를 높였다. Huang의 모형의 경우 그룹화에 있어 Heuristic 기법을 사용하여 그룹 분류의 기준을 선정하는 것이다. Huang의 기법에 의하여 더욱 많은 그룹이 발생하게 된다. Huang의 모형은 출력의 특성치가 모형에 들어감으로서 실제적인 예측이라 할 수는 없다. 그러므로 학습자료를 통하여 그룹별 전이확률개념을 이용하여 Huang의 모형의 문제점을 개선하였다. 이는 오차정도를 통한 재그룹화에 있어 오차정도의 기준은 자료의 양의 오차, 음의 오차를 기준으로 하여 자료를 분류하였다. 퍼지 시계열 그룹화는 Chen의 모형과 같으나 이를 오차정도를 기준으로 재그룹화 하였다. 재그룹화된 퍼지 시계열과 각 그룹의 발생확률을 산정하였다.

그림 7.은 자료계열의 분류와 표 1.의 그룹을 활용하여 전이 가능한 분류를 평균법을 이용하여 1997년의 강수량을 예측한 결과이다. 모형에 있어서 예측 결과는 양의 오차를 갖는 경우와 음의 오차를 갖는 경우 두가지의 경우에 대한 결과가 발생하며, 두가지 경우의 발생확률은 출력된 결과에 확률가중평균되어 최종 예측 결과를 보인다.

그림 8.은 표 1.에 나타난 그룹에 따른 뉴로-퍼지 모델을 구성하여 Huang case 3모형을 작성한 후 표 2.에 나타난 1997년 자료계열의 분류에 따라 모형을 적용하여 강수량을 예측한 결과이다. 그림 8.은 그림 6.의 결과와 마찬가지로 양의 오차와 음의 오차를 갖는 경우의 발생확률에 대한 확률가중평균을 하여 최종 예측 결과를 발생시켰다.

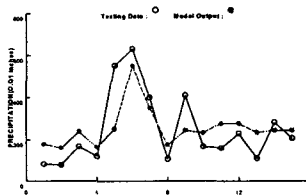
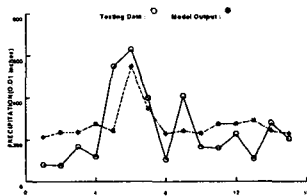


그림 7. case 2 모형에 의한 일리노이주 강수량 예측 그림 8. Huang case 3 모형에 의한 일리노이주 강수량 예측

4.2 국내에의 적용

4.2.1 Chen의 퍼지 시계열을 이용한 강수량의 예측(case 1, 3)

4.1.1절의 적용과 동일하게 국내의 월강수량 자료를 예측하였다.

국내 강수량 자료계열을 퍼지 시계열로 작성하기 위하여 상한, 하한을 각각 0, 450(mm)으로 결정한 후 일정간격 30(mm)를 기준으로 퍼지 집합 $A_1 \sim A_{15}$ 으로 분류하였다. 분류된 자료계열을 전이 그룹을 나누었으며, 표 2.는 1998, 1999년 입력자료계열의 분류를 나타내고 있다.

표 2. 1998, 1999년 강수량 자료계열 분류(case 1, 2, 3)

	1	2	3	4	5	6	7	8	9	10	11	12
그룹	A_2	A_2	A_2	A_2	A_6	A_4	A_{10}	A_{10}	A_{15}	A_9	A_3	A_2
	13	14	15	16	17	18	19	20	21	22	23	24
그룹	A_1	A_1	A_1	A_4	A_3	A_5	A_6	A_{10}	A_{14}	A_{12}	A_4	A_1

Chen의 모델에서 평균법을 이용한 예측은 전 단계 강수량 자료에서 전이 가능한 상태의 총 평균으로 나타난다. Chen의 모델에 의한 강수량 예측 결과는 그림 9.와 같다.

Chen의 평균법 예측 모형의 예측 정도를 높이기 위하여 표 4.에 나타난 그룹별 뉴로-퍼지 모델을 구성하여 Chen case 3 모형을 작성한 후 검정기간의 분류 상태에 따른 그룹별 모형을 적용하여 강수량을 예측하였다. 예측결과는 그림 10.과 같다.

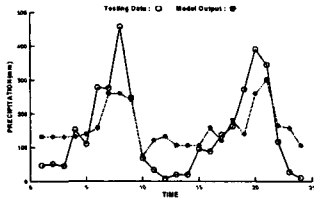


그림 9. case 1 모형에 의한 국내 강수량 예측

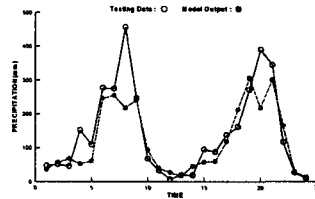


그림 10. Chen case 3 모형에 의한 국내 강수량 예측

4.2.2 Heuristic 퍼지 시계열을 이용한 강수량의 예측(case 2, 3)

4.1.2절에서 적용한 것과 동일한 방법으로 국내 강수량에 적용한다. 오차정도의 기준을 자료의 양의 오차, 음의 오차를 기준으로 하여 자료를 분류하였으며, 퍼지 시계열 그룹화는 Chen의 모형과 같으나 이를 오차정도를 기준으로 재그룹화 하였다. 재그룹화된 퍼지 시계열과 각 그룹의 발생확률을 산정하였다.

그림 11은 자료계열의 분류와 표 2의 그룹을 활용하여 전이 가능한 분류를 평균법을 이용하여 1998년~1999년의 강수량을 예측한 결과를 나타내고 있다. 예측 결과는 양의 오차를 갖는 경우와 음의 오차를 갖는 경우와 두가지 경우의 발생확률을 이용한 확률가중평균되어 최종 예측 결과를 보인다.

그림 12는 표 2에 나타난 그룹에 따른 뉴로-퍼지 모델을 구성하여 작성한 Huarng case 3 모형으로 1998, 1999년의 강수량을 예측한 결과를 나타내고 있다.

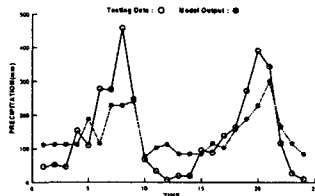


그림 11. case 2 모형에 의한 국내 강수량 예측

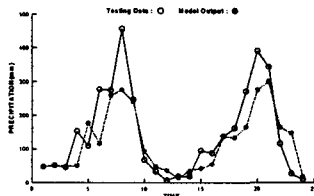


그림 12. Huarng case 3 모형에 의한 국내 강수량 예측

5. 요약 및 결론

한 지역의 기후는 각종 인자들의 복합적인 상호작용에 의하여 그 특성이 결정되는 것이며 간단한 방법으로 이를 예측하기란 대단히 어려운 것이다. 따라서 비교적 장기간 동안의 기상관측 자료를 획득하여 통계학적으로 분석함으로써 기후에 대한 예측을 할 수 있다.

본 연구에서는 각종 인자들의 복합적인 상호작용 및 기후의 비선형적 특성의 문제를 해결하기 위하여 애매함(fuzziness)을 수리적으로 취급이 가능하도록 하는 퍼지이론을 활용하였다. 또한 세부적인 퍼지이론의 적용은 전통적인 시계열을 fuzzify한 퍼지시계열을 사용하는 퍼지시계열 모형과 신경망의 학습기법을 이용한 뉴로-퍼지 모형을 적용하였다.

강수량 예측 모형은 강수의 특성인 각종 인자들의 복합적인 상호작용을 고려하기 위하여 다양한 대기 순환 인자를 이용하여 모형을 구축하였다. 본 연구에서는 강수량, 기온, 토양수분, 해수면 온도, 해면 기압의 인자들이 사용되었다.

제안된 예측 모형은 미국 일리노이주의 1996년 21일 간격의 1년간 강수량 및 국내 1988, 1999년의 강수량을 예측하였다. 소개된 case 1, case 2, case 3을 적용하였다.

일리노이주의 경우 대기순환인자로서 강수량, 기온, 토양수분을 이용하였으며 국내의 경우 강수량, 기온, 해수면 온도, 해면 기압을 이용하였다. case 1 적용결과, 대기순환 입력인자로 강수량만이 들어가며 최종 출력 생성 능력의 한계를 보였으며, case 2의 경우 실제적인 예측을 위하여 확률개념의 도입이 불가피 하며 case 2와 마찬가지로 최종 출력 생성 능력의 한계를 보였다. 본 연구에서 제안한 case 3의 경우 입력인자로 앞서 언급한 다양한 대기순환인자를 입력으로 사용하였으며 기존의 퍼지시계열 모형인 case 1, case 2의 문제점을 해결하였다. 표 3은 이러한 모형의 예측 결과를 보이고 있다.

표 3. 예측결과 오차

	미국 일리노이주				국내			
	case 1	case 2	Chen case 3	Huarng case 3	case 1	case 2	Chen case 3	Huarng case 3
P.E(%)	73.32	75.93	46.38	52.42	213.41	172.32	41.44	65.17
RMSE	130.22	136.42	115.25	113.39	87.65	86.24	68.82	72.03

이러한 결과를 통하여 다음과 같은 결론을 내릴 수 있었다.

- 1) 기존에 제안된 퍼지 시계열 예측 모형의 경우 다양한 인자에 영향을 받는 강수의 예측에는 미흡한 점이 많다.
- 2) 뉴로-퍼지 모형의 경우 복합적인 작용을 고려하는데 용이하므로, 본 연구에서 제안한 퍼지 시계열과 뉴로-퍼지 모형을 이용한 예측 모형은 강수와 같은 자연현상의 예측에 기존의 퍼지시계열 모형에 비하여 효과적임을 알 수 있었다.

6. 참고문헌

- Ahmed Ouenes (2000). "Practical application of fuzzy logic and neural networks to fractured reservoir characterization", *Computers Geosciences*, Vol. 26, p. 953~962
- D. K. Gautam, K. P. Holz (2001). "Rainfall-runoff modelling using adaptive neuro-fuzzy systems", *Journal of Hydroinformatics*, March, p. 3~10
- Drasko Furundzic (1998). "Application example of neural networks for time series analysis: rainfall-runoff modeling", *Signal Process*, Vol. 64, p. 383~396
- Fi-John Chang, Yen-Chang Chen (2001). "A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction", *Journal of Hydrology*, Vol. 245, p. 153~164
- Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani (1996). *Neuro-fuzzy and soft computing : A Computational Approach to Learning and Machine Intelligence*, Prentice Hall
- Kin C. Luk, J. E. Ball, A Sharma (2001). "An application of artificial neural networks for rainfall forecasting", *Mathematical and Computer Modeling*, Vol. 33, p. 683~693
- L. Bodri, V. Cermak (2000). "Prediction of extreme precipitation using a neural network: application to summer flood occurrence in Moravia", *Advances in Engineering Software*, Vol. 31, p. 312~321
- N. Sajikumar, B.S. Thandaveswara (1999). "A non-linear model using an artificial neural network", *Journal of Hydrology*, Vol. 216, p. 32~55
- N. T. Lange (1998). "New Mathematical Approach Hydrological modeling-An Application of Artificial Neural Network", *Phys. Chem. Earth*, Vol. 24, No. 1-2, p. 31~35
- Stewart W. Franks, Philippe Gineste, Keith J. Beven, Philippe Merot (1998). "On constraining the predictions of a distributed model : The incorporation of fuzzy estimates of saturated areas into the calibration process", *Water Resources Research*, Vol. 34, No. 4, p. 787~797
- Qiang Song, Brad S. Chissom (1993). "Forecasting enrollments with fuzzy time series-Part 1", *Fuzzy Sets and System*, Vol. 53, p. 1~9
- Qiang Song, Brad S. Chissom (1994). "Forecasting enrollments with fuzzy time series-Part 2", *Fuzzy Sets and Systems*, Vol. 62, p. 1~8