

# 앙상블 유량예측 시스템의 사전 및 사후처리에 관한 연구

## Pre- and Post-Processors of Ensemble Streamflow Prediction System

강태호\*, 김영오\*\*, 홍일표\*\*\*

Tae-Ho Kang, Young-Oh Kim, Il-Pyo Hong

### 요 지

미래 발생 가능한 수문 및 기상현상의 예측과정은 지식의 부족과 자연현상의 다양성으로 인해 불확실성을 포함하게 된다. 하지만 많은 예측들은 아직까지 확정적으로 제공되고 있으며, 결과적으로 예측결과의 불확실성 정도를 제공하지 못하고 있다. 앙상블 유량예측(ESP, Ensemble Streamflow Prediction)은 이러한 불확실성을 고려하여 수자원시스템의 의사결정에 있어 중요한 요소 중 하나인 유량예측을 수행할 수 있는 방법이다. 하지만 ESP의 결과는 기상자료, 유역 초기조건, 수문모형의 매개변수, 단순화된 수문모형에 의해 비교적 큰 불확실성을 포함하게 되며, 따라서 실제적인 현업에서의 사용을 위해서는 불확실성 정도를 줄이기 위한 사전 및 사후처리 과정이 요구된다. 본 연구에서는 국내에서 활용 가능한 기후 예보자료를 사용하여 앙상블 유량예측에 적용할 수 있는 사전처리 방안들을 검토하고, 국내에서 사후처리를 위해 적용되었던 최적선형 보정기법에 더해 다양한 기법들을 강우유출모형인 TANK모형의 모의결과 보정에 적용하였다. 사전 및 사후처리를 적용한 결과 기상자료와 유량예측과정에 존재하는 불확실성을 저감시키는 것이 가능하였다. 특히 사전 및 사후처리가 동시에 적용되었을 경우 그 향상 정도가 단순히 각각의 방법에 의한 향상 정도를 합한 것보다 높게 나타날 수 있음이 확인되었다. 사전 및 사후처리를 동시에 적용한 경우 이수기에는 RPS(Ranked Probability Score) 평가방법 내에서 54 %를, 홍수기에는 8 %를 향상시키는 것이 가능하였다.

**핵심용어** : 불확실성, 앙상블 유량예측, 사전처리, 사후처리, TANK, 오차보정, 기후예보

## 1. 서론

미국 기상청인 National Weather Service을 비롯한 선진국의 기상 및 유량예측 기관에서는 앙상블 예측기법에 기초한 확률론적 예측 시스템을 핵심기법으로 선정하여 이용하고 있다(김영오 등, 2004). 앙상블 유량예측은 현재의 유역 초기조건과 과거 강우 관측기록을 통해 유출앙상블을 생산하는 방법으로, Kim et al.(2001)이 물공급 전망의 개선을 위해 공주지점에 적용한 이후 예측기간에 대해 제시되는 기후예보와의 접목(사전처리, pre-processor)과 수문모형의 오차보정(사후처리, post-processor)에 관한 연구들이 진행되어 왔다. 하지만 사전처리에 대해 국내에서 진행된 연구는 실제 적용에 있어 초보적인 단계에 있으며, 사후처리 또한 정대일 등(2005)이 최적선형 보정기법을 적용한 사례가 있으나 다양한 방법들이 검토되지는 못하였다. 따라서 본 연구에서는 국내에서 활용 가능한 기후 예보자료를 사용하여 앙상블 유량예측에 적용할 수 있는 사전처리 방안들과 최적선형 보정기법에 더해 다양한 사후처리 방안들을 검토한 후 최적의 사전 및 사후처리 방안을 제시하려 한다.

## 2. 사전처리: 앙상블 유량예측을 위한 기후예보의 활용

\* 서울대학교 건설환경공학부 석사과정·E-mail : kangth@snu.ac.kr  
\*\* 정회원·서울대학교 건설환경공학부 부교수·E-mail : yokim05@snu.ac.kr  
\*\*\* 정회원·한국건설기술연구원 수자원연구부 책임연구원·E-mail : iphong@kict.re.kr

## 2.1 기후예보의 평가

본 연구에서는 사전처리를 위해 두 가지의 기후 예보자료를 사용하였다. 그 한 가지는 순(旬)강수전망으로 평년값을 기준으로 3가지 범주(AN; Above Normal, N; Normal, BN; Below Normal)에 대해 예보결과가 제시된다. 2000년부터 2003년 4월까지의 강수전망 실적치는 월간산업기상정보(<http://industry.kma.go.kr>)에서 수집할 수 있으며, 2003년 5월 이후의 월강수전망은 기상청 홈페이지에서 확인할 수 있다(강부식 등, 2007). 순강수전망의 특징을 살펴보면 전국에 대해 평균적으로 제시되고 있어 특정 적용 지역을 순강수전망이 어느 정도 반영할 수 있는지에 대한 우선 평가가 필요하다. 표 1은 순강수전망과 소양강댐 유역 강수량과의 적중률을 평가한 것으로 예보가 AN과 BN인 경우에서 모두 적중률은 42 %로 초보예측 33.3 %보다 높은 것을 확인할 수 있다. 반면 예보가 N인 경우의 적중률은 27 %로 비교적 낮은 값을 보였다.

표 1. 소양강댐 유역에 대한 기상청 순강수전망의 적중률(2004.1~2007.12)

예측 \ 관측	AN	N	BN	범주별 비율
AN	0.42(14)	0.24(8)	0.33(11)	0.23(33)
N	0.34(31)	0.27(25)	0.39(36)	0.64(92)
BN	0.32(6)	0.26(5)	0.42(8)	0.13(19)
범주별 비율	0.35(51)	0.26(38)	0.38(55)	1.00(144)

다른 한 가지의 기후예보자료는 NCAR(National Center for Atmospheric Research)가 주관이 되어 개발한 모형인 CCSM(Community Climate System Model)의 결과를 경계조건으로 다시 상세 지형을 고려할 수 있는 SNURCM(Seoul National University Regional Climate Model)을 통해 예측한 결과이다. SNURCM은 환경부가 주관하는 ‘차세대 핵심환경기술개발사업’ 내에서 개발된 모형으로, 1980년부터 1999년까지의 강수에 대한 예측결과를 10일 단위로 검증한 결과  $R^2$ 는 0.07 정도를 보였으며 관측과 예측의 평균적인 차이는 2.5 mm, RMSE(Root Mean Squared Error)는 17.14 mm임을 보였다.

## 2.2 확률 기후예보

기후예보를 ESP의 예측결과에 반영하기 위해서는 우선적으로 단일 값이나 범주별 예보를 통해 예측시점에 대한 확률적 예보를 계산하는 과정이 요구된다. 국내에서 활용 가능한 순강수전망과 SNURCM의 결과를 사용하여 확률적인 예보를 얻기 위해 Schaake et al.(2007)과 강부식 등(2007)이 제안한 방법을 사용하였다. 본 논문에서는 Schaake et al.(2007)이 제안한 방법을 Conditional Probability Method(CPM)로 강부식 등(2007)이 제안한 방법은 Categorical Method(CM)로 나타내었다.

CPM은 과거에 단일 값으로 제시된 예측과 동 기간에 대한 관측자료를 사용하여 예측시점에 대해 확률적인 예측결과를 제시하는 방법이다.  $X$ 를 과거 제시된 예측 값으로,  $Y$ 를 동 기간에 대한 관측 값이라 하고,  $X$ 와  $Y$ 에 대한 결합분포(joint distribution)를 2변량정규분포(bivariate normal distribution)라 가정한다면 예측에 대한 분포는 식 1과 같이 조건분포(conditional distribution)로 나타낼 수 있다(Schaake et al., 2007). CPM은 이러한 조건분포를 사용하여 예보를 확률적으로 나타내는 방법이다.

$$\begin{aligned} \mu_{Y|X} &= \mu_Y + \rho_{XY}(\sigma_Y / \sigma_X)(x - \mu_X) \\ \sigma_{Y|X} &= \sigma_Y^2(1 - \rho_{XY}^2) \end{aligned} \quad (1)$$

기상청에서 제공되는 순강수전망을 사용하여 확률적으로 예측을 제시하는 CM은 과거 범주별(AN, N, BN)로 제시된 순강수전망과 동 기간에 대한 관측 자료를 정리하고, 각 범주, 예를 들어 AN의 범주에 대해 실제 관측된 자료들을 모아서 이 자료에 대한 분포를 찾아 AN을 대표하는 분포로 사용하는 방법이다.

## 2.3 확률 기후예보의 반영

분포로 제시된 확률 기후예보를 ESP에 반영하는 방법으로 3가지를 사용하였다. 첫 번째로 Schaafe Shuffle(SS)은 예측시점에 제시된 기후예보에 대한 분포를 ESP의 예측에 사용되는 기상양상블과 동일한  $N$ 개의 구간으로 나눈 후, 각 구간의 기댓값을 계산하여 크기 순으로 나열된 기상양상블을 동일한 순번을 가지는 구간의 기댓값으로 대신하는 방법이다(Clark et al., 2004). 이때 각 구간은 동일한 확률  $1/N$ 을 가진다.

두 번째는 분위사상법(Quantile Mapping, QM)으로 ESP의 입력 자료인 기상양상블과 기후예보의 확률분포를 사용하여 기상양상블 멤버들이 기후예보의 확률분포와 동일한 분포를 가지도록 하는 방법으로, Hashino et al.(2002)이 수문모형의 확률 편이(Bias)를 보정하기 위해 적용한 사례가 있다.  $F_O$ 와  $F_F$ 를 각각 관측된 기상양상블과 기후예보의 분포라 하고  $Y_i$ 를 기상양상블의 멤버라 한다면 보정된 기상양상블 멤버  $Y_i^*$ 는 다음의 식을 통해 계산된다.

$$Y_i^* = F_F^{-1}(F_O(Y_i)) \quad (2)$$

세 번째 방법은 Stedinger et al.(2002)이 제안한 Pdf-Ratio(PR)라는 방법으로 앞의 두 방법과는 다르게 기상양상블과 기후예보의 확률밀도함수 비를 사용하여 각 멤버들의 확률을 변화시키는 방법이다. 기후예보가 반영된 앙상블 멤버의 확률  $P_i$ 는 다음의 식에 의해 계산될 수 있다.

$$P_i = (1/N)[F_F(Y_i)/F_O(Y_i)] \quad (3)$$

## 3. 사후처리: 수문모형의 계통적 오차 보정

사후처리는 유역 초기조건, 수문모형의 매개변수, 단순화된 수문모형에 의한 불확실성을 저감시키는 과정으로 3가지 방법이 본 연구에서 적용되었다. 첫 번째 방법은 Smith et al.(1992)에 의해 제안된 Event Bias Correction Method(EBC)로 같은 기상 입력 자료에 의해 발생하는 편이의 크기는 일정하다는 가정에서 시작한다. 과거 관측된 유출량을  $Q_{O_i}$ 라 하고 동 기간의 기상자료를 사용하여 모의된 유출량을  $Q_{S_i}$ , 예측시점의 유역 초기조건 하에서 모의된 유출량을  $Q_{F_i}$ 라 한다면 EBC에 의해 보정된 유출량  $Q_{R_i}$ 는 식 4에 의해 계산된다.

$$Q_{R_i} = (Q_{O_i}/Q_{S_i}) \times Q_{F_i} \quad (4)$$

두 번째 방법은 Theli(1971)에 의해 처음 제안된 최적선형 보정기법(Optimal Linear Correction, OLC)으로, 국내에서는 정대일 등(2005)이 앙상블 유량예측을 위한 SSARR(Streamflow Synthesis and Reservoir Regulation)모형의 모의유량 보정에 적용한 사례가 있는 방법이다. 예측결과인  $Q_{F_i}$ 는 다음의 선형회기식을 통해 간단히 보정될 수 있으며, 회귀계수  $\alpha$ 와  $\beta$ 는  $Q_{O_i}$ 와  $Q_{S_i}$ 의 자료를 사용한 최소제곱법에 의해 추정될 수 있다.

$$Q_{R_i} = \alpha + \beta Q_{F_i} \quad (5)$$

마지막은 기후예보를 ESP의 결과에 반영하기 위해 사용했던 분위사상법으로, 모의된 유출량  $Q_{S_i}$ 는 관측된 유출량  $Q_{O_i}$ 와 동일한 분포를 가져야 한다는 가정에서 예측된 유출량  $Q_{F_i}$ 를 보정하는 방법이다.  $Q_{S_i}$ 에 대한 분포를  $F_S$ ,  $Q_{O_i}$ 에 대한 분포를  $F_O$ 라 한다면  $Q_{R_i}$ 는 식 6을 통해 보정 가능하다.

$$Q_{R_i} = F_O^{-1}(F_S(Q_{F_i})) \quad (6)$$

## 4. 사전 및 사후처리의 적용

### 4.1 평가방법

앙상블 유량예측의 결과를 평가하기 위해 확률예측의 정확도를 평가할 수 있는 AHS(Average Hit Score)와 RPS(Ranked Probability Score)를 사용하였다. AHS는 실측유량이 발생한 구간의 예측확률을 검증기간 동안 평균한 값이며, 예측구간이 3개인 초보예측의 경우 33.3 %를 갖는다. 거리민감도(sensitivity to distance)를 측정할 수 있는 방법 중 하나인 RPS는 관측과의 거리에 따라 벌점을 부여하는 방법으로 완벽예보에서는 RPS =

0이며, 정확성이 낮아질 경우 점차적으로 큰 양의 값을 가진다(Wilks, 1995). RPS는 식 7과 같이 계산된다.

$$RPS = \sum_{m=1}^J (Z_m - O_m) \quad (7)$$

식 7의  $J$ 는 확률구간 수이며  $Z_m$ 은 예측에 대한,  $O_m$ 은 관측에 대한 누적확률이다.

## 4.2 사전 및 사후처리 평가

사전 및 사후처리를 TANK모형을 강우유출모형으로 사용하는 ESP 시스템에 적용하였다. 적용유역은 1974년부터 관측 자료가 존재하는 한강의 소양강댐으로 선정하였으며, 2001년부터 2007년까지의 기간을 홍수기와 이수기로 나누어 사전 및 사후처리가 적용된 ESP 시스템의 순단위 결과를 평가하였다. 순강수전망을 ESP의 예측결과에 반영하기 위해서 Schaake et al.(2007)이 제안한 CPM을 사전처리로써 사용하였으며, SNURCM을 결과를 반영하여 위해서는 강부식 등(2007)이 제안한 CM을 사용하였다(표 2). 표 3은 RPS와 AHS의 지표를 사용한 평가결과이며, 짙은 색으로 표시된 부분은 각각의 기간 및 평가방법 내에서 다른 방법에 비해 우수한 결과를 보이는 것이다.

표 2. 예보 형식에 따른 사전처리 방법

	예보형식	자료기간	사전처리
순강수전망	범주별	2000~현재	CM
SNURCM	확정적	1980~2019	CPM

우선 사전처리의 결과만을 보면 홍수기에는 CPM과 QM을 적용한 경우 RPS와 AHS가 각각 0.364와 0.484로 다른 방법에 비해 우수한 결과를 보였다. 이수기에는 CM과 PR을 적용하였을 때 RPS와 AHS가 각각 0.767과 0.344로 다른 방법보다 양호한 결과를 보였다. 이러한 결과를 토대로 현재 제시된 기후예보의 정확성만으로도 ESP의 결과를 향상시키는 것이 가능함을 확인할 수 있었다. 사후처리만을 적용하였을 경우 홍수기에는 EBC가 가장 적은 RPS 결과 0.348을 보였으며, OLC는 AHS 내에서 0.478로 우수하였다. 이수기에는 EBC의 RPS가 0.378로 다른 방법에 비해 좋은 결과를 보였으며, AHS의 평가방법에서는 QM이 0.463으로 가장 높은 결과를 보였다. 적용이 간단한 것에 비해 홍수기와 이수기에서 모두 높은 향상가능성을 보였으며, 특히 이수기에는 사후처리가 적용된 경우 RPS를 0.769에서 0.378까지 줄이는 것이 가능하였다.

사전 및 사후처리를 동시에 적용한 경우 때로는 단순히 각각의 방법을 통해 향상된 정도를 합한 것보다 높은 향상 정도를 보여 두 방법을 동시에 적용해야 하는 필요성을 보이고 있었다. 예를 들어 이수기에 사전처리과정인 CM과 PR에 의한 RPS의 향상 정도는 0.002이며, 사후처리 EBC에 의한 향상 정도는 0.391인데 반해 CM, PR, EBC를 동시에 적용하였을 경우 0.402 정도가 향상되었다. 홍수기에는 사전처리로써 CPM과 QM을 사후처리로 OLC를 적용한 경우 RPS와 AHS가 각각 0.343과 0.489로 다른 방법에 비해 우수하였다. 이수기에는 CM, PR, EBC를 적용한 경우 RPS의 평가방법 내에서 0.367로 가장 높은 향상 정도를 보였으나, AHS의 평가방법에서는 CM, PR, QM을 사용하였을 때 0.484로 높은 값을 보였다.

## 5. 결론

본 연구는 사전처리 내에서 국내에서 활용 가능한 기후예보인 순강수전망과 SNURCM의 예측결과를 사용하여 ESP의 입력자료인 기상자료의 불확실성을, 사후처리를 통해 유역 초기조건, 수문모형의 매개변수, 단순화된 수문모형에 의한 불확실성을 저감할 수 있었다. 특히 사전 및 사후처리가 동시에 적용되었을 경우 그 향상 정도가 단순히 각각의 방법에 의한 향상 정도를 합한 것보다 높게 나타날 수 있음을 확인하였다.

사전 및 사후처리를 동시에 적용한 경우 이수기에는 RPS 평가방법 내에서 54 %를, 홍수기에는 8 %를 향상시키는 것이 가능하였다. 현재 전 세계적으로 다양한 수치예보모델이 개발되고 있으며, 결과적으로 향상된 기후예보결과를 사전처리를 통해 앙상블 유량예측 결과에 반영하는 것이 가능하므로 이러한 결과는 앞으로 보다 진보될 것으로 예상된다. 기상청의 순강수전망 또한 점차 많은 자료가 누적되고 있어 향후 앙상블 유량예측에 있어 보다 높은 적용 가능성을 보일 것이다.

표 3. 사전 및 사후처리를 적용한 앙상블 유량예측 평가결과

Verification method		RPS				AHS				
Post-processor		No Post	EBC	QM	OLC	No Post	EBC	QM	OLC	
Pre-processor										
Wet Season	No Pre	0.373	0.348	0.360	0.353	0.471	0.444	0.450	0.478	
	CPM	QM	0.364	0.346	0.359	0.343	0.484	0.450	0.456	0.489
		SS	0.369	0.355	0.362	0.348	0.483	0.444	0.452	0.482
		PR	0.383	0.356	0.372	0.366	0.466	0.441	0.444	0.473
	CM	QM	0.388	0.355	0.378	0.370	0.453	0.429	0.421	0.452
		SS	0.407	0.381	0.389	0.387	0.460	0.421	0.421	0.453
PR		0.385	0.369	0.380	0.367	0.470	0.436	0.446	0.477	
Dry Season	No Pre	0.769	0.378	0.472	0.540	0.341	0.437	0.463	0.439	
	CPM	QM	0.774	0.389	0.484	0.554	0.341	0.433	0.452	0.429
		SS	0.780	0.387	0.485	0.549	0.339	0.436	0.452	0.433
		PR	0.767	0.377	0.468	0.538	0.341	0.437	0.466	0.442
	CM	QM	0.767	0.389	0.465	0.530	0.337	0.433	0.463	0.447
		SS	0.781	0.415	0.491	0.553	0.330	0.420	0.440	0.427
PR		0.767	0.367	0.460	0.528	0.344	0.446	0.484	0.457	

### 감 사 의 글

본 연구는 국토해양부의 ‘앙상블 모형을 이용한 확률적 유량예측’ 연구사업과 안전하고 지속가능한 사회 기반건설사업단 SIR BK21의 기획연구과제 지원에 의해 수행되었습니다.

### 참 고 문 헌

1. 강부식, 유승엽, 고익환(2007). 강수전망 가중앙상블을 이용한 장기 확률유량예측, 대한토목학회 논문집, 제 27권, 제2B호, pp. 183-191.
2. 김영오, 정대일(2004). 국내 앙상블 유량예측 연구 5년, 한국수자원학회 학술발표회, 한국수자원학회, CD.
3. 정대일, 이재경, 김영오(2005). 최적선정보정을 이용한 앙상블 유량예측 시스템의 개선, 한국수자원학회 논문집, 한국수자원학회, 제38권, 제6호, pp. 473-485.
4. Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R.(2004). The Schaake Shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields, *J. Hydrometeorol.*, 5(1), pp. 243-262.
5. Hashino, T., Bradley, A. A., and Schwartz, S. S.(2002). Verification of probabilistic streamflow forecasts, Tech. Rep. IIHR Report No.427, IIHR-Hydrosci. Eng.
6. Kim, Y.-O., Jeong, D. I., and Kim, H. S.(2001). Improving water supply outlooks in Korea with ensemble streamflow prediction, *Water International*, 26(4), pp. 563-568.
7. Schaake, J., Demargne, J., Hartman, R., Mullusky, M., Welles, E., Wu, L., Herr, H., Fan, X., and Seo, D. J.(2007). Precipitation and temperature ensemble forecasts from single-value forecasts, *Hydrol. Earth Syst. Sci. Discuss.*, 3, pp. 655-717.
8. Smith, J. A., Day, G. N. and Kane, M. D.(1992). Nonparametric framework for long-range streamflow forecasting, *Journal of Water Resources Planning and Management*, 118(1), pp. 82-91
9. Stedinger, J. and Kim, Y.-O.(2002). Updating ensemble probabilities based on climate forecasts, 2002 *Conference on Water Resources Planning and Management*, ASCE. Roanoke, USA, CD.
10. Theil, H.(1971). Applied economic forecasting, North Holland Publishing Company, Amsterdam, Netherlands.
11. Wilks, D. S.(1995). Statistical Methods in the Atmospheric Sciences, volume 59 of International Geophysics Series, Academic Press, San Diego.