

기상 확률예보를 이용한 앙상블 예측

○ 정대일¹⁾, 김영오²⁾

1. 서론

최근 ENSO 정보 또는 기상 확률예보를 중장기 유량예측에 활용하여 수문예측의 불확실성을 극복하려는 연구가 활발히 진행되고 있다. Piechota et al.(1998)은 호주 동부의 기상변화와 ENSO의 상관관계를 이용하여 계절유량에 대한 확률예측 모형을 구축하였으며, Hamlet and Lettenmaier(1999)는 ENSO와 더불어 Pacific Decadal Oscillation(PDO) 지표를 기준으로 과거 강우자료를 몇 개의 범주로 나눈 후 앙상블 예측기법을 미국 Columbia River의 유량 확률예보에 적용하였다. 미국 기상청(National Weather Service)의 수문예측 핵심기법인 앙상블 예측(Ensemble Streamflow Prediction)은 Kim et al. (2001)에 의해 우리나라의 건설교통부 하천계획과와 한국건설기술연구원이 매월 초 발간하고 있는 '물공급전망'의 개선을 위해 이미 적용되었다.

본 연구에서는 앙상블 예측을 이용하여 충주댐 월 유입량 시나리오를 생성하고, 여기에 기상 확률예보를 이용하는 방안으로 Croley방법(1996, 1997, 2000, 2001)과 Steginger and Kim(2002)이 제시한 PDF-Ratio방법을 적용하여 그 실용성을 검증하고자 한다.

2. 앙상블 예측과 가중값

앙상블 예측은 강우-유출모형에 미래에 일어날 가능성이 있는 모든 강우 시나리오(즉, 강우 앙상블)를 입력하여 그 결과물로 다수의 유량 시나리오(즉, 유량 앙상블)를 얻는 기법이다. 강우-유출모형의 초기조건은 예측하는 시점에 따라 변하므로, 같은 강우 시나리오를 사용한다 하여도 초기조건에 따라 유량 시나리오가 다르게 생성된다(김영오 등, 2001).

앙상블 예측에서 가장 핵심이 되는 사항은 생성된 유량 시나리오에 가중값을 부과하는 과정이며, 이를 통해 입력자료의 불확실성을 극복할 수 있다. 가중값을 부여하는 가장 간단한 방법으로는 모든 시나리오의 중요도가 동일하다고 가정하는 방법, 즉 예를 들어 30개의 유량예측 시나리오가 있다면, 각각 1/30의 동일한 확률을 부과하는 방법이다. 조금 발전된 방법은 예측대상의 수문사상과 비슷한 과거의 수문사상을 가진 연도에 해당하는 시나리오에 보다 높은 가중값(또는 확률)을 부여하는 방법이다. 수문학적 상황을 판단하는 근거에는 강우나 기온의 예측정보, ENSO 정보 등이 있을 수 있다.

가중값에 관한 연구로, Croley는 그의 책(2000)과 연속된 연구(1996, 1997, 2001)에서 NOAA(National Oceanic and Atmospheric Administration)의 기상 확률예보를 고려하여 예측유량 시나리오에 가중값을 부여하는 구체적인 기법을 제안하였다. Croley방법은 n 개의 시나리오와 m 개의

1) 서울대학교 지구환경시스템공학부 박사과정

2) 서울대학교 지구환경시스템공학부 조교수

기상 확률예보를 이용해, 가중값(w_i)과 1의 차의 제곱을 최소로 하는 목적함수와 기상 확률예보가 제약조건인 식(1)의 최적화문제를 풀어 가중값을 구한다.

$$\begin{aligned} \min \quad & \sum_{i=1}^n (w_i - 1)^2 \\ \text{subject to} \quad & \sum_{i=1}^n a_{k,i} w_i = e_k \quad k = 1, \dots, m \end{aligned} \quad (1)$$

Stedinger and Kim(2002)은 Croley방법과는 달리 확률밀도함수를 이용하여 가중값을 구하는 PDF-Ratio방법을 제안하였다. 강우 등, 과거 기상변수 X 에 대한 확률밀도함수를 $f_0(x)$ 라하고 수문학적 정보(H)에 따른 X 의 예측 확률밀도함수를 $f_1(x) = f(X|H)$ 라고 가정하면 구하려는 가중값, 즉 시나리오 x_i 에 대응하는 가중값 w_i 는 식(2)를 이용하여 구할 수 있다.

$$w_i = (f_1(x_i) / f_0(x_i)) \quad (2)$$

3. 앙상블 예측의 적용

앙상블 예측의 국내 적용 가능성을 검토하기 위해 충주댐 유역을 연구대상으로 선정하였다. 우리나라와 같이 강우자료가 유량자료보다 상대적으로 풍부한 경우 앙상블 예측을 이용하면 강우자료만큼의 유량자료를 생성할 수 있다는 장점을 살릴 수 있다. 충주댐은 1978년 6월에 착공하여 1985년 10월에 완공되었으며 1986년부터 유입량이 실측되었다. 그러나 충주댐 유역의 강우자료는 1965년부터 존재하므로 앙상블 예측을 적용하기에 적당한 대상이라고 할 수 있다.

탱크모형을 이용하여 1996~1999년까지 4년간의 예측을 실시하여 예측유입량 앙상블을 생성하고 이를 실측유입량과 비교, 검토하였다. 1966년부터 1995년까지의 30개의 강우자료를 탱크모형에 입력하여 예측년도(1996~1999)의 매월마다 30개의 유입량 앙상블을 생성하였다. 유입량 앙상블 생성단계에서 예측시점의 토양함수 등의 초기조건이 자동으로 결정되도록 하였다. 예를 들어 1996년 5월의 유입 앙상블을 생성하는 경우는, 1995년 5월부터 1996년 4월까지의 단일한 과거 강우자료를 이용하여 탱크모형을 초기화시킨 후, 1966년부터 1995년까지의 30개의 강우 앙상블을 각각 탱크모형에 입력함으로써 96년 5월의 30개 유입량 앙상블을 생성하였다. 마지막으로 유입량 앙상블에 Croley방법과 PDF-Ratio방법을 이용하여 각 유입량 시나리오에 해당하는 가중값을 각각 구한 후, 유량 확률예측의 정확성 개선 여부를 검토하고 오차분석을 실시하였다.

본 연구에서 사용한 탱크모형은 장기 일 유출량자료 구축을 위해 한국건설기술연구원(강경석 등, 2001)에서 사용한 모형을 이에 필요한 매개변수와 함께 그대로 사용하였다. 이 모형은 4개의 탱크로 이루어진 4단 탱크모형으로서, 각 탱크의 유출공(5개), 침투공(3개), 유출공 높이(4개), 초기저류고(4개) 등 16개의 매개변수로 구성되어 있다.

탱크모형의 검정을 위해 충주댐 유역의 1996~1999년 실측강우를 탱크모형을 이용하여 모의한 후 충주댐의 실측유입량과 비교하였다. 두 유입량의 계절별 Bias, RMSE(Root Mean Square Error), Relative Bias 등을 이용하여 비교한 결과, 봄, 여름 가을에 비해 겨울철의 모의유입량은 정확성이 높지 않음을 알 수 있었다. 따라서 유입량 확률예측의 정확성 평가는 겨울철을 제외한 3월~11월에 한하여 실시하였다.

다음 단계로 충주댐 유역의 강우 확률예보를 생성하였다. 충주댐 유입량 앙상블에 기상예보를 이용하기 위해서는 예보자료가 확률예보 형태로 존재해야 한다. 우리나라의 경우 기상청에서 발행하는 “월간산업기상정보”에 월별과 순별(상순, 중순, 하순)로 기온과 강우에 대한 예보를 실시하고 있으나, 이는 확률예보가 아닌 과거 30년간의 기상자료의 평균값을 기준으로 ‘평년보다 조금 적겠음’, ‘평년과 비슷하겠음’ 등의 예보형식이며, 조금 적음과 많음의 기준도 명확하지 않아 예측유량 시나리오의 가중값을 구하기 위한 기상 확률예보자료로 정량화하여 이용할 수 없었다. 그러나 본 연구의 목적이 앙상블 예측의 적용성에 대한 검토이므로 충주댐 유역의 실측 강우자료를 이용하여 기상 확률예보를 다음과 같이 인위적으로 생성하였다.

충주댐 유역의 강우자료(1966~1999)를 이용하여 예측기간인 1966~1995년까지의 자료를 가지고, 각각 한달, 상순, 중순, 하순에 대한 누가확률 33.3%, 66.7%에 해당하는 강우를 구하여 이를 기준으로 세 구간을 설정하였다. 다음으로 1996~1999년까지 4년동안 매월 실측강우가 위치하는 구간의 확률을 33.3%가 아닌 50%로 강우가 발생한 부분의 확률을 높여주었다. 그리고 나머지 두 구간은 동일하게 25%로 부과하였다.

생성한 충주댐 유역의 강우 확률예보를 이용하여 Croley방법과 PDF-Ratio방법으로 각 유입량 예측 시나리오의 가중값을 구하였다(그림 1). 가중값의 합은 시나리오 개수와 같은 30이어야 하며 가중값을 30으로 나눈 값이 각 연도에 해당하는 유량 시나리오의 확률이다. 두 방법의 가중값이 일치하지는 않으나 크고 작음의 경향이 많이 유사함을 확인할 수 있다.

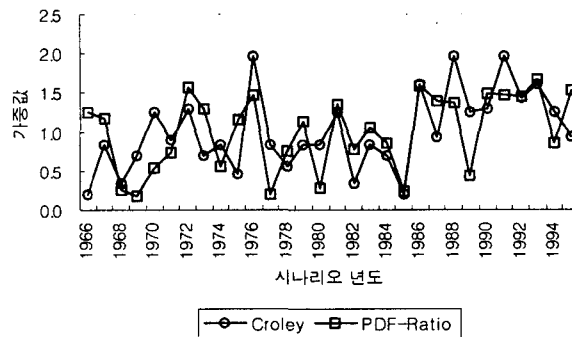


그림 1. Croley방법과 PDF-Ratio방법의 가중값 비교(1996년 4월)

4. 결과분석

유량의 확률예보는 통상 유량을 몇 개의 구간으로 나누고 각 구간의 발생확률을 예보하는 과정이다. 예측지점에 대한 과거유량의 적합분포를 선정하고 누가확률 33.3%와 66.7%에 해당하는 유량(각각 $q_{33.3}$ 과 $q_{66.7}$)을 기준으로 저수(low flow), 평수(normal flow), 풍수(high flow)로 나누었다. 만일 별다른 예측기법이 존재하지 않는다면 미래유량은 이러한 과거자료에 의존하여 각 구간에 대하여 동일하게 33.3%의 확률로 예보하여야 할 것이며, 이를 “초보예측(Naive Forecasts)”이라 한다(김영오 등, 2001). 기상 예측정보에 근거하여 예측유량의 확률분포를 갱신하고 예측구간별 발생확률을 제시하여 충주댐 유입량 확률예측을 실시하였다.

각 예측에서 실측유량이 발생한 구간의 발생확률을 예측점수라고 하였으며, 전체 예측월의 예측점수를 평균한 값을 평균 예측점수라고 하였다. 만약 평균 예측점수가 33.3%보다 크다면 전체적인 앙상블 예측은 초보예측보다 정확하여 예측기법으로 유효하다고 할 수 있다.

먼저 예측점수를 이용하여 앙상블 예측에 대해 평가를 실시하였다. 충주댐의 과거유입량을 이용하여 각 예측월의 적합분포를 선정하고 $q_{33.3}$ 과 $q_{66.7}$ 을 구하고 예측구간을 설정하였다. 표 1은 가중값을 부여하지 않은 경우, 다시 말해서 각 예측유입량 시나리오에 모두 균일한 가중값(1/30)을

부여한 경우(이하 균일가중값 부여방법)의 결과이다. 표 2는 Croley방법과 PDF-Ratio방법에 의한 가중값을 예측유입량 시나리오에 적용하여 평균 예측점수를 계산한 결과이며 균일가중값 부여방법보다 약 8%정도 증가하였다.

양상불 예측분포의 중앙값과 실측유입량을 확정론적 예측의 평가방법으로 비교하였다. 표 3은 각 예측유입량 양상불의 중앙값과 실측유입량을 비교하여 그 오차를 Bias, RMSE, $\sqrt{\text{Variability}}$ 로 나타낸 결과이다. 균일가중값 부여방법보다 Croley방법과 PDF-Ratio방법의 중앙값이 실측유입량과의 오차에 있어 더 작음을 확인할 수 있다.

표 1. 균일가중값 부여방법의 예측점수(%)

	3월	4월	5월	6월	7월	8월	9월	10월	11월	평균
1996	28.1	40.7	12.1	21.5	39.7	33.0	48.3	73.1	0.0	32.9
1997	14.6	9.3	12.5	44.8	33.7	25.6	69.9	38.1	9.3	28.6
1998	12.7	7.3	50.7	21.2	47.6	37.8	53.9	100.0	0.0	36.8
1999	82.2	12.5	55.1	65.1	29.4	39.1	25.9	90.5	100.0	55.5
평균	34.4	17.4	32.6	38.2	37.6	33.9	49.5	75.4	27.3	38.5*

표 2. 가중값 부여방법별 평균 예측점수(%)

	균일가중값	Croley	PDF-Ratio
평균 예측점수	38.5	46.6	45.9

표 3. 예측분포의 중앙값과 실측유입량 비교

	균일가중값	Croley	PDF-Ratio
Bias	-42.92	-47.52	-47.11
RMSE	173.80	162.50	163.11
$\sqrt{\text{Variability}}$	170.20	157.04	157.81

수문예측에서 발생하는 불확실성은 크게 수문모형이 불완전하기 때문에 발생하는 수문학적 불확실성(Hydrological Uncertainty)과 수문모형의 입력변수로 사용되는 미래의 발생 가능한 강우량과 기온 등의 기상변수에 의해 발생하는 기상학적 불확실성(Meteorological Uncertainty)으로 나눌 수 있다. 강우-유출모형이 완벽하다면 각 예측월의 실측유입량(Q_o)과 실측강우를 탱크모형으로 모의한 유입량(Q_s)은 같아야한다. Q_o 와 Q_s 의 차이를 강우-유출모형에서 발생된 수문학적 오차(E_h)라고 가정하였다. 각 예측유입량 양상불($q_i, i=1,2,\dots, n$)에 E_h 를 더한 후, 실측유입량과 예측유입량 양상불 사이에서 발생한 오차는 실측강우가 아닌 과거의 강우자료들을 이용했기 때문에 생긴 오차(E_m), 즉 기상학적 불확실성에 의한 오차라 할 수 있다. 즉,

$$\begin{aligned}
 E_h &= Q_o - Q_s \\
 Q_{f, med} &= (q_i + E_h)_{med} \quad (i=1, 2, \dots, n) \\
 E_m &= Q_o - Q_{f, med}
 \end{aligned}
 \tag{3}$$

여기서 $Q_{f, med}$ 는 예측유입량 양상불에서 E_h 만큼을 더한 후 구한 중앙값이다.

그림 2는 각 예측월의 총 오차 중에서 E_h 와 E_m 을 각 예측월별로 막대그래프를 이용하여 그린 것이다. 겨울철인 12월, 1월, 2월의 경우는 총 오차 중에서 탱크모형으로 인한 오차가 강우에 의한 오차보다 훨씬 더 큼을 보여 주고 있다. 겨울철 이외의 경우는 입력변수인 강우에 의한 오차

가 더 큼을 알 수 있다.

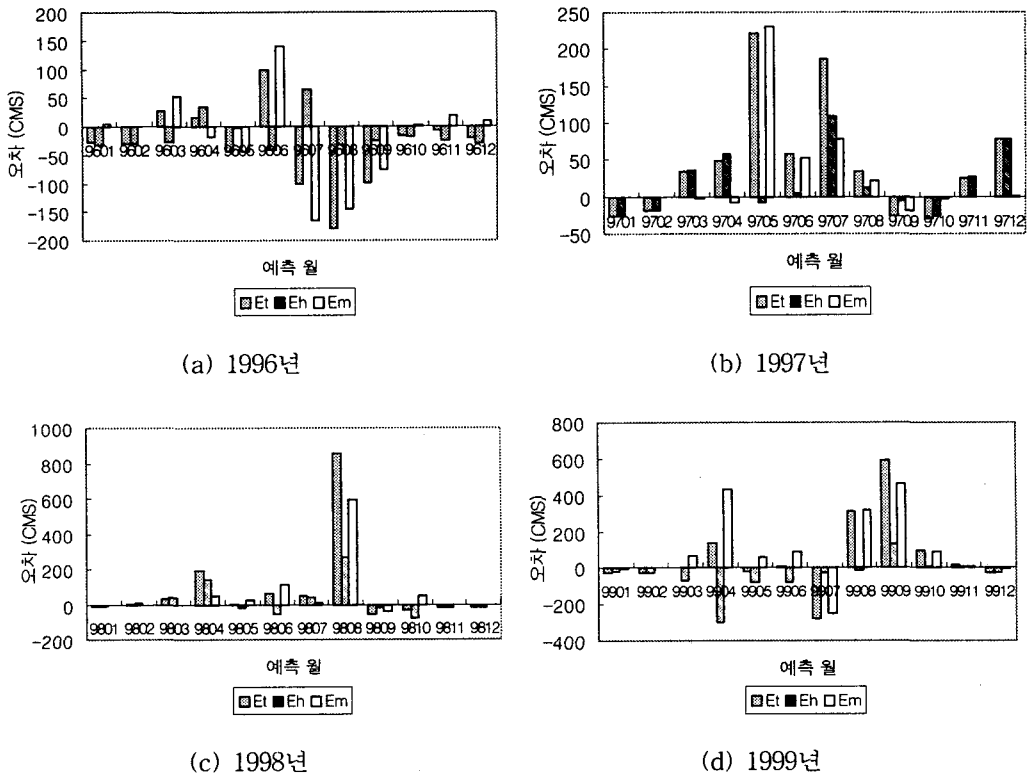


그림 2. 불확실성 분석결과

4. 결론 및 향후연구

양상불 예측을 통해 1996~1999년 동안 실시한 증주담 월 유입량의 확률예보는 가중값을 부여하지 않아도 평균 예측점수가 38.5%로서 초보예측 보다는 우수함이 입증되었다(표 1). 또한 강우 확률예보를 이용하여 예측유입량 시나리오에 가중값을 부여할 경우, 평균 예측점수는 Croley 방법에서 8.1%, PDF-Ratio방법에서 7.4%의 개선효과를 보였다(표 2). 두 방법의 평균 예측점수 차이는 0.7%로 Croley방법의 평균 예측점수가 약간 높았으나 Croley방법이 더 우수하다고 할만큼 큰 차이는 아니었다. 또한 Croley방법은 예측기간이 한달, 상순, 중순, 하순인 4가지의 강우 확률예보를 이용한 반면 PDF-Ratio방법은 한달 강우 확률예보만을 이용하여도 가중값을 구할 수 있다는 장점을 확인하였다.

예측분포의 중앙값과 실측유입량을 Bias 등의 오차지표로 비교한 결과 균일가중값 부여방법의 오차지표가 가중값 부여방법의 오차지표보다 크게 나타나, 가중값을 부여할 경우 예측분포의 중앙값과 실측유입량의 오차가 작아짐을 확인하였다(표 3).

예측분포의 중앙값, 탱크모형의 모의유입량, 그리고 실측유입량을 이용하여 실시한 불확실성에 관한 분석에서, 강우가 적은 겨울철의 경우 탱크모형의 불확실성에서 발생하는 오차가 큰 반면, 상대적으로 강우가 많은 봄, 여름, 가을철의 경우에는 탱크모형에 의한 오차보다는 입력변수의 불확실성에서 발생하는 오차가 매우 크게 나타났다. 따라서, 겨울철의 예측오차를 감소시키기 위

해서는 강우-유출모형의 개선이 필요하며, 봄, 여름, 가을의 경우는 정확한 기상 확률예보를 이용하여 강우 입력변수의 불확실성을 감소시켜야 한다는 결론에 이르게 되었다.

본 연구에서 사용한 강우 확률예보는 인위적으로 생성한 것이므로 실제 충주댐 유입량 확률예보에 적용하기 위해서는 충주댐 유역의 강우 확률예보가 필수적이다. 그러나 현재 우리나라의 기상청 등을 비롯한 기상 예보기관에서는 강우 확률예보를 실시하고 있지 않는 실정이다. 따라서, 강우 확률예보의 실시와 확보가 가중값을 부여하기 위해 시급히 해결되어야 할 문제이며 현재 우리나라의 기상 예측정보의 실정에 맞는 가중값 부여방안에 관한 연구도 함께 진행되어야 한다.

참고문헌

- 강경석(2001) 탱크모형을 이용한 장기 일 유출량 자료 구축, 2001년 학술발표회 논문집(I), 한국수자원학회, pp. 245-250.
- 김영오, 정대일, 김형섭, 이길성(2001) 앙상블 예측을 통한 물공급전망 개선방안. 2001년 학술발표회 논문집(1), 한국수자원학회, pp. 304-308.
- Croley II, T. E. (1996) Using NOAA's new climate outlooks in operational hydrology. *Journal of Hydrologic Engineering*, ASCE, 1(3), pp. 93-102.
- Croley II, T. E. (1997) Mixing probabilistic meteorology outlooks in operational hydrology. *Journal of Hydrologic Engineering*, ASCE, 2(4), pp. 161-168.
- Croley II, T. E. (2000) *Using Meteorology Probability Forecasts In Operational Hydrology*, ASCE Press.
- 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.
- Croley II, T. E. (2001) Climate-biased storm-frequency estimation. *Journal of Hydrologic Engineering*, ASCE, 6(4), 275-283.
- Hamlet, A. F. and Lettenmaier, D. P. (1999) Columbia river streamflow forecasting based on ENSO and PDO climate signals, *Journal of Water Resources Planning and Management*, ASCE, 125(6), pp. 333-341.
- Piechota, T. J., Dracup, J. A., Chiew, F. H. S., and McMahon, T. A. (1998) Seasonal streamflow forecasting in Eastern Australia and the El Nino-southern oscillation, *Water Resources Research*, 34(11), pp. 3035-3044.
- Stedinger, J. R. and Kim, Y.-O. (2002) Updating ensemble probabilities based on climate forecasts, Conference on Water Resources Planning and Management and Symposium on Managing the Extremes Floods and Droughts, EWRI, ASCE, Roanoke, Virginia, USA. (to be presented).