Journal of Wetlands Research
Vol. 21, No. s-1, November 2019, pp. 61-68

ISSN 1229-6031 (Print) / ISSN 2384-0056 (Online)
DOI https://doi.org/10.17663/JWR.2019.21.s-1.61

장기 기후 변동성을 고려한 인공신경망 앙상블 모형 적용: 한강 유역 댐 유입량 예측을 중심으로

김태림·주경원·조완희*·허준행†

연세대학교 건설환경공학과
*K-Water 통합물관리처

Application of Artificial Neural Network Ensemble Model Considering Long-term Climate Variability: Case Study of Dam Inflow Forecasting in Han-River Basin

Kim, Taereem·Joo, Kyungwon·Cho, Wanhee*·Heo, Jun-Haeng†

Department of Civil and Environmental Engineering, Yonsei university, South Korea
**Integrated River Basin Mnagement Division, K-water, South Korea*
(Received : 07 October 2019, Revised: 21 November 2019, Accepted: 21 November 2019)

요 약

최근 장기적인 기후 변동성을 고려하기 위하여 대기-해양 순환 패턴을 수치화한 기상인자가 수문 변수 예측에 널리 사용되고 있다. 또한 정확하고 안정적인 예측을 위해 인공신경망 기반의 예측 모형이 꾸준히 발전하고 있다. 기상인자를 활용하여 기후 변동성을 고려한 수문량 예측은 수자원 및 환경 보존의 장기적인 관리에 효율적으로 활용될 수 있으므로 수문 변수에 유의한 인자의 파악과 이를 활용한 예측 모형의 적용은 꾸준한 도전이 될 것이다. 본 연구에서는 우리나라 한강 유역 댐 유입량에 통계적으로 유의한 상관성이 있는 대표 기상인자를 선정하고, 이를 인공신경망 앙상블 모형에 적용하여 댐 유입량 예측을 수행하였다. 이를 위해 앙상블 경험적 모드분해법을 활용하여 댐 유입량과 기상인자간의 통계적 상관성을 확인하였으며, 기존 단일 인공신경망 모형의 한계를 보완한 인공신경망 앙상블 모형을 구축하였다. 예측 수행 결과, 5개 댐 상관계수 평균이 훈련 기간에서 0.88, 검증 기간에서 0.68의 예측력을 보이는 것을 확인하였으며, 본 연구에서의 절차를 토대로 우리나라의 다양한 수문 변수와 기후 변동성간의 관계를 활용한 다양한 적용 사례가 나오길 기대한다.

핵심용어 : 기상인자, 기후 변동성, 댐 유입량 예측, 앙상블 경험적 모드분해법, 인공신경망 앙상블

Abstract

Recently, climate indices represented by quantifying atmospheric-ocean circulation patterns have been widely used to predict hydrologic variables for considering long-term climate variability. Hydrologic forecasting models based on artificial neural networks have been developed to provide accurate and stable forecasting performance. Forecasts of hydrologic variables considering climate variability can be effectively used for long-term management of water resources and environmental preservation. Therefore, identifying significant indicators for hydrologic variables and applying forecasting models still remains as a challenge. In this study, we selected representative climate indices that have significant relationships with dam inflow time series in the Han-River basin, South Korea for applying the dam inflow forecasting model. For this purpose, the ensemble empirical mode decomposition(EEMD) method was used to identify a significance between dam inflow and climate indices and an artificial neural network(ANN) ensemble model was applied to overcome the limitation of a single ANN model. As a result, the forecasting performances showed that the mean correlation coefficient of the five dams in the training period is 0.88, and the test period is 0.68. It can be expected to come out various applications using the relationship between hydrologic variables and climate variability in South Korea.

Key words : Artificial Neural Network Ensemble, Climate Index, Climate Variability, Dam Inflow Forecasting, Ensemble Empirical Mode Decomposition

† To whom correspondence should be addressed.
Department of Civil and Environmental Engineering, Yonsei University, South Korea
E-mail: jhheo@yonsei.ac.kr

- **Kim, Taereem** Department of Civil and Environmental Engineering, Yonsei University, South Korea / Postdoctoral researcher (taereem@yonsei.ac.kr)
- **Joo, Kyungwon** Department of Civil and Environmental Engineering, Yonsei University, South Korea / Ph.D. Candidate (kwjy1@yonsei.ac.kr)
- **Cho, Wanhee** Integrated River Basin Management Division, K-Water, South Korea / Research fellow (wanhee@kwater.or.kr)
- **Heo, Jun-Haeng** Department of Civil and Environmental Engineering, Yonsei University, South Korea / Professor (jhheo@yonsei.ac.kr)

1. 서 론

기후 변화의 영향과 함께 보고되고 있는 기후 변동성은 수문변수에 장기적으로 영향을 미치고 있다. 인간 활동으로 인한 기후 변화의 영향뿐만 아니라 대규모 대기-해양 시스템에 내재되어 장기적 기후 변동성을 고려한 수문량 예측은 수자원 및 환경 보존의 효율적인 관리에 매우 중요해질 것으로 보인다. 이에 따라 장기 기후 변동성을 고려하여 예측에 적용하는 연구가 확대되고 있으며, 특히 세계 기후 시스템의 광범위한 기후 모드를 나타내는 미국 해양 대기관리/지구 시스템 연구소(NOAA/ESRL)에서 제공하는 기상인자를 활용한 연구가 활발하게 진행되고 있다(Schepen et al., 2012; Mekanik et al., 2013; Abbot and Marohasy, 2014; Li et al., 2015;. Kim et al., 2018; Jung et al., 2018).

수문 자료에 내재된 장기 변동성을 분석하기 위하여 신호처리 분석 기법으로 알려진 앙상블 경험적 모드 분해법(Ensemble empiricap mode decomposition; EEMD)을 적용할 수 있다(Wu et al., 2007; Wu and Huang, 2009). EEMD는 시계열 안에 내재된 주기적인 패턴과 장기 경향성을 추출할 수 있으며 자료가 비선형적, 비정상성인 경우에도 효과적으로 분해할 수 있기 때문에 장기적 기후 변동성 및 변화의 영향을 받는 수문 자료에 효율적으로 적용할 수 있다는 장점이 있어 다양한 연구에 활용되고 있다(Kim and Cho, 2016; Kim et al., 2018; Lee and Ouarda, 2019).

또한 장기 예측에 활용될 수 있는 수문 모형은 전통적으로 물리적 모형과 통계적 모형이 있으며, 최근에는 인공신경망(Artificial Neural Networks; ANN)을 활용한 모형이 널리 활용되고 있다. ANN 모형은 다양한 분야에서 높은 적용성과 예측력을 보이지만, 모형의 구조적인 문제로 인하여 발생되는 한계를 내포하고 있다(Kolen and Pollack 1991; Santos-García et al. 2004). 가장 많이 언급되는 부분으로는 초기 가중치 매개변수의 무작위성으로 인한 불확실성이 있다. 이는 인공신경망 모형 구축 시 초기 가중치가 무작위로 생성됨에 따라 동일한 입력 자료임에도 불구하고 예측 결과의 성능이 크게 달라질 수 있어 매우 민감하다는 것이다. 또다른 문제점은 은닉층의 최적 노드 수의 선정에 관한 불확실성이다. ANN 모형의 은닉층 최적 노드 수를 선정하는 방법은 일반적으로 가능한 모든 노드 수를 고려하여 예측을 수행한 후 가장 좋은 결과를 제공하는 경우를 최적 노드 수로 선정하는 trial-and-error 방법을 사용하기 때문에 많은 계산 시간이 소요된다. 이러한 한계를 극복하기 위하여 최근 앙상블 개념을 활용한 ANN 모형의 개발이 확대되고 있으며 다양한 적용이 시도되고 있다(Cannon and Whitfield, 2002; Anctil and Nicolas , 2004; Boucher et al., 2009; Zaier et al., 2010; Tiwari et al., 2013; Kim and Seo, 2015).

본 연구에서는 국내 수문 자료에 영향을 미치는 장기 기후 변동성을 고려한 예측을 수행하고자 한다. 이를 위해 EEMD를 활용하여 댐 유입량 자료의 내재된 주기를 추출한 후 다양한 기상인자와의 상관성 분석을 통해 한강 유역의 5개 댐 유입량 시계열과 유의한 상관성을 보이는 대표 기상인자를 선정하였다. 최종적으로 선정된 기상인자를 입력 변수로 하는 ANN 앙상블 모델을 구축하고 예측 성능을 살펴보았다.

2. 연구방법

2.1 앙상블 경험적 모드 분해법

EEMD는 시계열 자료에 백색잡음을 추가하고 체거름 알고리즘(sifting algorithm) 을 통해 자료를 분해한 뒤 최종적으로 앙상블의 평균을 사용하여 각 분해된 요소를 추출하는 방법이다. 여기서 체거름 알고리즘을 간략하게 나타내면 아래과 같다. 먼저 주어진 시계열($y(t)$, $t=1,2,3,\cdots,n$) 의 극값(국소 최댓값 및 최솟값)을 구하고, 삼차 스플라인(cubic spline)을 활용하여 상위선(upper envelope, $y_u(t)$)과 하위선(lower envelope, $y_l(t)$)을 구한다. Eq.(1)과 같이 상위선과 하위선의 평균선(mean envelope)을 구한다.

$$y_m(t) = (y_u(t) + y_l(t))/2 \tag{1}$$

원 시계열($y(t)$)에서 평균선을 공제한 후($y(t) - y_m(t)$)의 내재모드함수의 조건을 만족할 경우, 이를 내재모드함수(intrinsic mode function; IMF)라 정의한다. 이후 추출된 IMF를 제거한 후의 시계열을 원 시계열로 재설정 후($y(t)$), 남은 자료계열이 단조함수 또는 하나의 극값만 존재하여 더 이상 새로운 계열이 추출되지 않을 때까지 반복적인 알고리즘을 통해 N개의 IMF를 추출한다. 따라서 주어진 시계열은 Eq.(2)와 같이 내재된 주기를 나타내는 유한개의 IMF와 장기적인 경향성을 나타내는 하나의 잔여값(residue)으로 분해된다.

$$y(t) = \sum_{i=1}^{N} IMF_i + residue \tag{2}$$

EEMD에 관한 상세한 설명은 다음 문헌을 참고할 수 있다(Huang et al., 1998; Wu and Huang, 2004; Kim et al., 2017; Kim et al., 2018; Jung et al., 2018).

2.2 교차 상관성 분석

교차 상관성 분석은 두 시계열($X_1(t)$, $X_2(t)$) 간의 선형 상관성의 세기를 표현하는 통계적 기법으로 두 시계열의 교차상관계수(r_k)가 -1 또는 1일 경우 두 시계열 간의 강한 음의 상관성 또는 양의 상관성을 나타내며, 교차상관계수는 Eq. (3)과 같이 계산된다.

$$r_{X_1(t),X_2(t)}(k) = \frac{\sum_{t=1}^{T-k}(X_1(t)-\overline{X_1}(t))(X_2(t+k)-\overline{X_2}(t))}{\sqrt{\sum_{t=1}^{T}(X_1(t)-\overline{X_1}(t))^2}\sqrt{\sum_{t=1}^{T}(X_2(t)-\overline{X_2}(t))^2}} \tag{3}$$

여기서, $\overline{X_1}(t)$와 $\overline{X_2}(t)$는 각 시계열의 평균, k는 시간 지연(시차)을 나타낸다.

2.3 인공신경망 앙상블 모형

ANN 모형은 인공지능(Artificial intelligence) 기법의 한 분야로 뇌의 신경세포의 작동 원리를 모방하여 수학적인 방식으로 나타낸 모델이다. 일반적으로 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성되어 있으며, 각 층에는 뉴런을 나타내는 노드가 존재한다. 이러한 노드들은 뉴런을 연결하는 시냅스의 역할과 같이 연결 강도를 매개변수로 하여 이어지고 비선형 형태의 함수를 통해 출력되는 학습 과정을 통해 예측을 수행한다(Moradi and Lee, 2018).

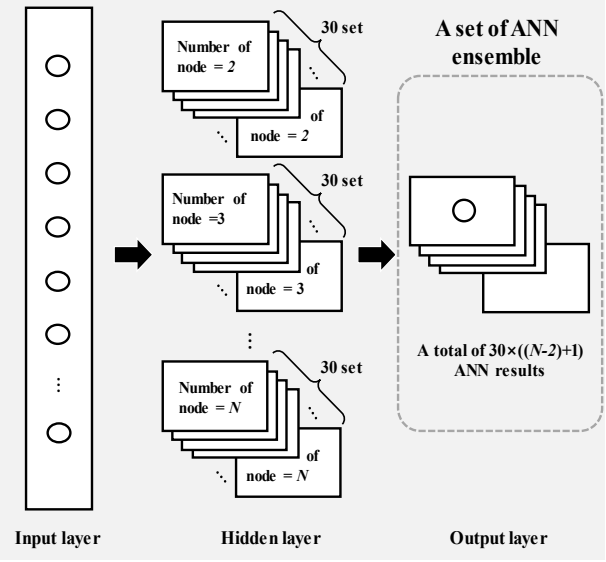

Fig. 1. Structure of ANN ensemble model used in this study.

본 연구에서는 단일 ANN 모형의 두 가지 한계점을 보완하기 위하여 ANN 앙상블 모형을 구축하였다(Fig. 1). n개의 입력 자료가 있다고 가정했을 때 초기 가중치의 불확실성을 고려하기 위하여 단일 ANN 모형을 30번 구축하여 앙상블을 생성하고, 은닉층의 최적 노드 수 선정의 불확실성을 고려하기 위하여 최소 노드 수(2) 부터 최대 노드 수(N) 까지 각각 단일 ANN 모형을 구축하고 앙상블을 생성하였다. 여기서 최대 노드 수(N)는 입력 자료의 개수로 설정하였다. 최종적으로 구축된 앙상블의 중앙값으로 예측 결과를 산정하였다.

3. 적용 및 결과

3.1 대상 지점 및 자료

본 연구에서는 우리나라 한강 유역에 영향을 미치는 기상인자를 고려한 댐 유입량 예측을 수행하기 위해 미국 해양 대기관리/지구 시스템 연구소(NOAA/ESRL)에서 제공하는 월 단위 기상인자와 한국수자원공사 관할 5개 댐의 월 평균 유입량 자료를 이용하였다. Fig. 2는 대상 댐의 위치를 지도상에 나타낸 것이며, 댐의 기본 정보 및 기상인자 목록은 Table 1, 2에 나타내었다.

Table 1. Basic information of five dams used in this study.

Name of dam (Abbreviation)	Type	Basin area(㎢)	Data period
Soyangang dam(SY)	Multi-objective	2,703	1974.01-2017.12
Chungju dam(CJ)	Multi-objective	6,648	1986.01-2017.12
Hoengseong dam(HS)	Multi-objective	209.0	1998.01-2017.12
Gwangdong dam(GD)	Water supply	125.0	1989.01-2017.12
Dalbang dam(DB)	Water supply	29.4	1990.01-2017.12

Table 2. List of climate indices used in this study.

Climate Index (Abbreviation)	Classification	Climate Index (Abbreviation)	Classification
NINO 1+2 (NINO12)	ENSO/SST:Pacific	Tropical Northern Atlantic Index (TNA)	SST:Atlantic
NINO 3 (NINO3)	ENSO/SST:Pacific	Tropical Southern Atlantic Index (TSA)	SST:Atlantic
NINO 4 (NINO4)	ENSO/SST:Pacific	Carribbean SST Index (CAR)	SST:Atlantic
NINO 3.4 (NINO34)	ENSO/SST:Pacific	Pacific Decadal Oscillation (PDO)	Teleconnections
Bivariate ENSO Timeseries (BEST)	ENSO	Northern Oscillation Index (NOI)	Teleconnections
Multivariate ENSO Index (MEI)	ENSO	Pacific North American Index (PNA)	Teleconnections
Trans-Nino Index (TNI)	SST:Pacific	Western Pacific Index (WP)	Teleconnections
Western Hemisphere Warm Pool (WHWP)	SST:Pacific/SST:Atlantic	Eastern Atlantic/Western Russia (EAWR)	Teleconnections
Oceanic Nino Index (ONI)	SST:Pacific	North Atlantic Oscillation (NAO)	Teleconnections
Atlantic Multidecadal Oscillation (AMO)	SST:Atlantic	Southern Oscillation Index (SOI)	Atmosphere
Atlantic Meridional Mode (AMM)	SST:Atlantic	Quasi-Biennial Oscillation (QBO)	Atmosphere
North Tropical Atlantic SST Index (NTA)	SST:Atlantic	Artic Oscillation (AO)	Atmosphere

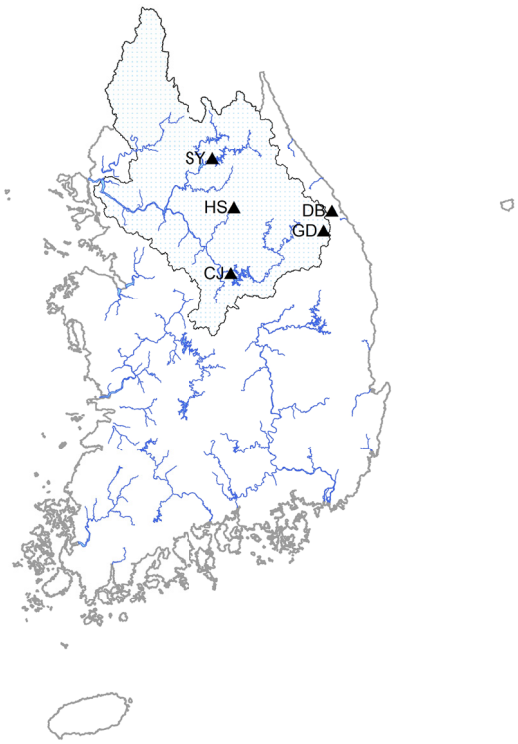

Fig. 2. Geographic location of five dams in Han-River basin.

3.2 대표 기상인자 선정

한강 유역 댐 유입량과 기상인자 간의 통계적 상관성을 분석하기 위한 절차로 댐 유입량 시계열을 EEMD를 통해 분해하였다. 각 댐별 IMF와 기상인자와 교차 상관성 분석을 통해 IMF별로 가장 높은 상관 계수를 갖는 기상인자를 선정하였고, 이 중에서 가장 많이 선정된 상위 3개 기상인자를 한강 유역을 대표하는 최적 기상인자로 선정하였다.

Fig. 3은 EEMD를 통해 분해된 각 댐별 IMF를 보여주고 있다. 각 댐의 자료 길이에 따라 7~9개의 IMF가 분해되었으며 마지막 IMF는 residue를 나타낸다. 분해가 진행됨에 따라 다양한 주기의 IMF가 추출되고, 최종적으로 장기 경향을 나타내는 residue가 남게된다. 따라서 EEMD를 통해 추출된 IMF는 댐 유입량 시계열에 내재된 다양한 주기 및 장기 경향의 정보가 있으며 이를 통해 원자료 시계열보다 정확한 통계적 분석이 가능하다.

Table 3에 각 IMF와 가장 높은 상관 계수를 가지는 기상인자를 나타내었다. 그 중에서도 댐 유입량의 계절적 변동을 반영하는 세 번째 IMF(IMF3)와 NINO12인자는 높은 상관 계수를 보인다. NINO12인자는 태평양 해수면 온도와 관련한 기상인자로 적도 동태평양 지역(0S-10S, 90W-80W)에서 측정된 평균 해수면 온도를 지수화 한 값이다. 따라서 본 결과는

Table 3. Climate indices and the highest correlation coefficients with the IMFs in five dams.

Name of dam	IMF	Climate index	Correlation coefficient	Lag(month)
Soyangang dam	IMF1	NINO12	0.1295	10
	IMF2	NINO12	0.4741	5
	IMF3	NINO12	0.7368	4
	IMF4	QBO	0.3409	1
	IMF5	SOI	0.1960	1
	IMF6	AMM	0.2901	1
	IMF7	AMM	0.2252	4
	IMF8	AMM	0.2282	12
	Residue	AMO	0.3229	1
Chungju dam	IMF1	NINO12	0.1431	10
	IMF2	NINO12	0.4774	5
	IMF3	NINO12	0.7534	5
	IMF4	QBO	0.3994	1
	IMF5	AMO	0.1977	1
	IMF6	AMM	0.2006	1
	IMF7	TNA	0.2845	12
	Residue	EAWR	0.1602	12
Hoengseong dam	IMF1	NINO12	0.1741	10
	IMF2	NINO12	0.4808	5
	IMF3	NINO12	0.7687	4
	IMF4	QBO	0.2621	1
	IMF5	TNA	0.4583	5
	IMF6	TNA	0.4411	7
	Residue	TNI	0.3522	1

Name of dam	IMF	Climate index	Correlation coefficient	Lag(month)
Gwangdong dam	IMF1	PNA	0.1140	9
	IMF2	NINO12	0.4177	5
	IMF3	NINO12	0.7146	5
	IMF4	QBO	0.2987	3
	IMF5	MEI	0.2947	12
	IMF6	NINO4	0.3272	1
	IMF7	AMO	0.3795	12
	Residue	AMM	0.1683	2
Dalbang dam	IMF1	NINO12	0.1145	11
	IMF2	NINO12	0.3791	5
	IMF3	NINO12	0.6488	5
	IMF4	ONI	0.1691	12
	IMF5	ONI	0.3254	12
	IMF6	AMO	0.3745	12
	IMF7	AMO	0.2273	12
	Residue	EAWR	0.1896	12

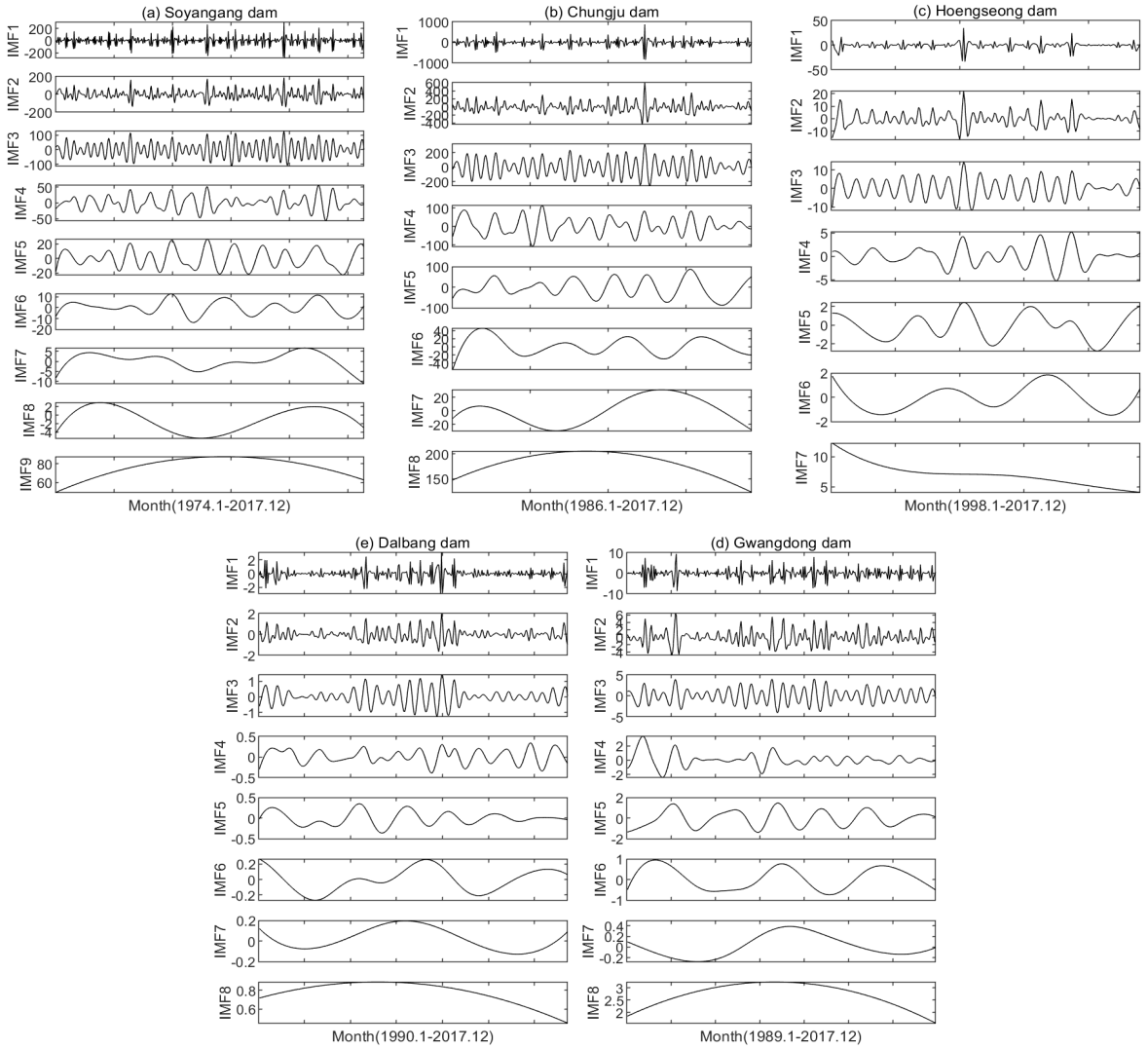

Fig. 3. Results of EEMD in five dams. (a) Soyangang dam, (b) Chungju dam, (c) Hoengseong dam, (d) Gwangdong dam, (e), Dalbang dam.

적도 동태평양 지역의 평균 해수면 온도가 4개월 또는 5개월 차이로 우리나라 유입량의 계절적 주기와 유의한 통계적 상관성이 있다는 것을 의미한다. 그 다음으로 5개 댐 중 3개 댐에서 QBO 인자와 AMO 인자가 각각 중기 및 장기 경향을 나타내는 IMF와 통계적 상관성이 있음을 보였으며, 시차는 대체적으로 1개월과 12개월이 선정되었다. 이러한 결과를 통해 적도 지역의 30mb 고도에서의 위도 간 바람장으로 정의되는 QBO인자와 우리나라 유입량에 내재된 중기 경향과의 비교적 짧은 지연 시간을 통한 상관성 및 북대서양 해수면 온도 관련 기상인자인 AMO인자와 유입량간의 장기적인 상관성을 확인할 수 있다.

최종적으로 5개월 지연된 NINO12인자, 1개월 지연된 QBO 인자, 12개월 지연된 AMO인자(NINO12(−5), QBO(−1), AMO(−12))가 한강 유역의 댐 유입량과 통계적 상관성을 가지는 대표 기상인자로 선정되었으며, 이는 우리나라 강우 자료와 기상인자와의 상관성을 발견한 Kim et al. (2018)의 선행 연구 결과와 비슷한 결과를 보인다.

3.3 ANN 앙상블 모형 적용 및 예측 결과

각 댐별로 3.2장에서 선정된 대표 기상인자(NINO12(−5), QBO(−1), AMO(−12)) 와 12개월, 24개월, 36개월 지연된 자체 유입량을 입력 변수로 ANN 앙상블 모형을 구축하고 댐 유입량 예측을 수행하였다. 모형의 훈련 및 검증을 위해 관측 자료의 시작부터 2014년 12월까지를 훈련 기간으로 설정하고, 2015년 1월부터 2017년 12월까지(최근 3년)를 검증 기간으로 설정하였다.

Fig. 4는 각 댐별 ANN 앙상블 모형의 훈련 기간(training period)과 검증 기간(test period)의 구축 결과이다. 여기서, 검정 실선은 관측된 댐 유입량 자료, 회색 실선은 ANN 앙상블 멤버 예측 결과, 파란 점선은 앙상블의 중앙값을 나타낸다. ANN 앙상블 멤버는 각 댐별로 총 150(초기 가중치 고려한 반복 30회 × 노드 수 5개)개가 생성이 되었다. 먼저 훈련 기간을 살펴보면, 전체 댐에서 ANN 앙상블 예측 결과가 실제 관측된 유입량의 흐름을 잘 반영하고 있는 것을 볼 수 있다. 따라서 한강 유역의 댐을 기준으로 선정된 대표 기상인자와 자체 유입량은 ANN 앙상블 모형 구축에 적절한 입력 자료로 사용될 수 있음을 확인하였다. 검증 기간에서의 예측력을 살펴보면, 앙상블 밴드가 다소 넓어져 단일 인공신경망에 포함된 불확실성을 확인할 수 있다. 그러나 높은 불확실성에도 불구하고 극심한 가뭄이 기록된 2015년을 제외하고는 앙상블의 중앙값을 통한 예측에서 관측 유입량과 상당히 근사한 예측 결과를 보이는 것을 시각적으로 확인할 수 있다. 이러한 결과를 자세히 살펴보기 위해 r(correlation coefficient) 및 RMSE(root mean square error)를 통해 예측력을 확인하였다. r값의 경우, 식 (3)에서 k=0일 경우와 동일하게 계산되며, RMSE는 Eq.(4)와 같이 구할 수 있다.

$$\mathrm{RMSE}=\sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_2(t)-X_1(t))^2} \tag{4}$$

여기서 T는 자료 길이, X_1과 X_2는 각각 관측 유입량과 예측 유입량을 의미한다. 또한 r(correlation coefficient)는 −1 또는 1에 가까워질수록 예측 성능이 좋음을 의미하고 RMSE(root mean square error)는 값이 작을수록 예측 성능이 좋음을 의미한다.

Table 4에서 나타낸 바와 같이 한강 유역 5개 댐의 ANN 앙상블 예측 성능은 훈련 기간에서 r값의 평균이 0.88, 검증 기간에서는 r값의 평균이 0.68인 예측력을 보이고 있다. 특히 소양강댐의 경우 검증기간의 r값이 0.7이상으로 다섯 댐 중 가장 높은 성능을 보이고 있다. RMSE를 통한 성능 평가에서도 소양강 댐의 경우 훈련기간과 검증기간의 RMSE값이 큰 차이가 나지 않으며 특히 충주댐의 경우 검증기간의 RMSE값이 더 낮으므로 두 댐의 경우 비교적 높은 예측 성능을 보인다. 그러나 횡성, 광동, 달방댐의 경우 상대적으로 높은 RMSE를 보임에 따라 소양, 충주댐보다는 다소 낮은 예측 성능을 보이고 있다. 전반적으로 한강 유역의 댐 유입량 예측 결과는 대표 기상인자의 사용과 함께 비교적 정확한 예측력을 보이고 있으며, 향후 수문량에 영향을 미치는 다른 요인들과 함께 세부적인 개선을 수행한다면 실제 댐 유입량 예측에 활용될 수 있을 것이라 판단된다.

4. 결 론

장기 기후 변동성에 대한 관심이 증가하면서 이를 고려한 수문 예측 등의 연구가 지속적으로 진행되고 있다. 이에 따라, 통계적 및 인공지능 기법과 모형은 다양한 적용을 통해 발전하고 있다. 본 연구에서는 우리나라 한강 유역 댐 유입량에 영향을 미치는 대표 기상인자를 선정하고 이를 활용한 ANN 앙상블

Table 4. Statistical measurement of model performance during training and test periods in five dams.

Name of Dam	r		RMSE	
	Training	Test	Training	Test
Soyangang dam	0.83	0.73	62.85	71.09
Chungju dam	0.89	0.67	128.41	95.37
Heongseong dam	0.98	0.63	2.67	115.98
Gwangdong dam	0.83	0.69	2.21	117.72
Dalbang dam	0.86	0.66	0.61	119.40
Average	0.88	0.68	39.35	103.91

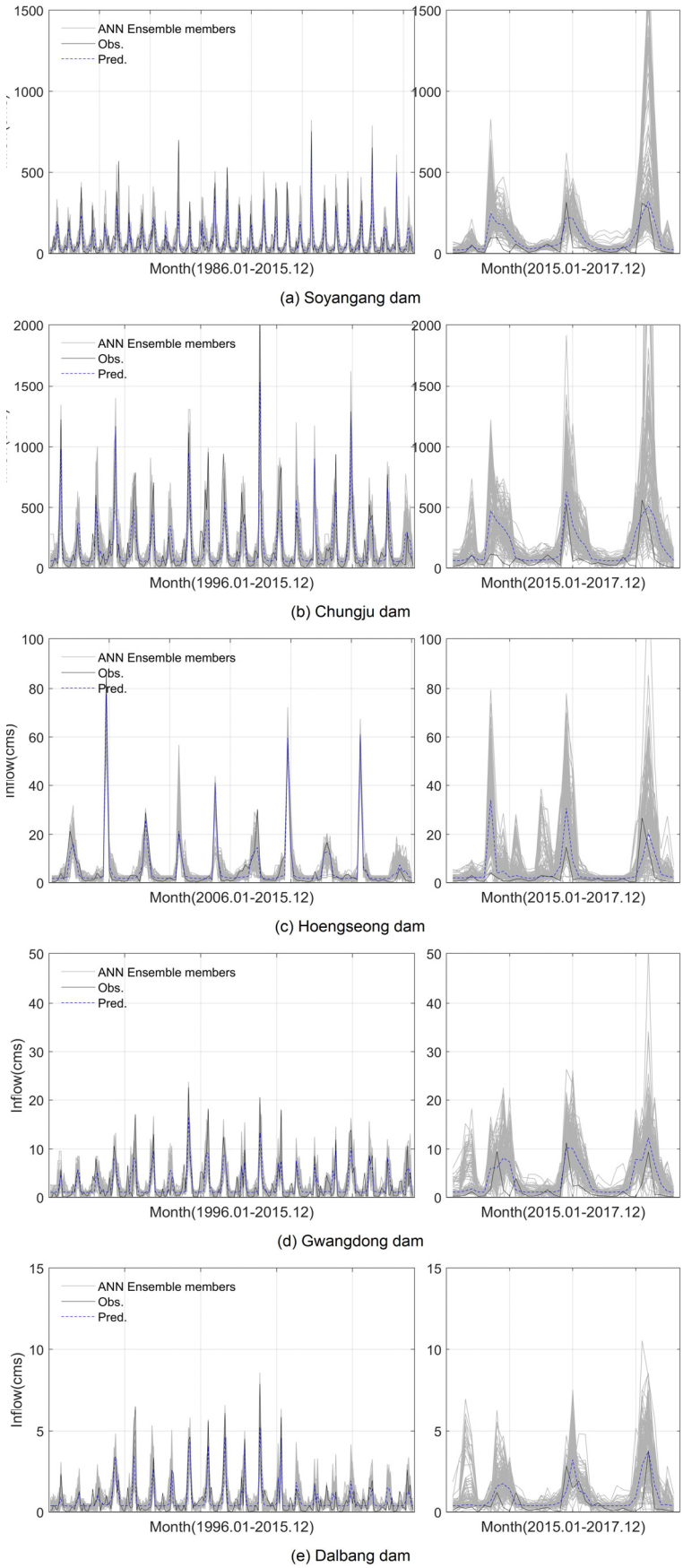

Fig. 4. Results of inflow forecasting using ANN ensemble model during training and test periods. (a) Soyangang dam, (b) Chungju dam, (c) Hoengseong dam, (d) Gwangdong dam, (e), Dalbang dam.

모형을 구축하여 댐 유입량 예측을 수행하였다. 선정된 NINO12(-5), QBO(-1), AMO(-12)인자는 우리나라 댐 유입량에 통계적으로 유의한 상관성을 가지고 보였다. 해당 기상인자를 입력 변수로 활용하여 ANN 앙상블 모형을 적용한 결과, 5개 댐 상관계수 평균이 훈련 기간에서 0.88, 검증 기간에서 0.68로 나타났으며 ANN 앙상블의 범위를 통해 기존 단일 ANN 모형이 가지고 있는 한계를 보완하고 보다 안정적인 예측 결과를 제공하는 것을 확인하였다. 추후 본 연구에서 제시한 절차를 기반으로 우리나라 장기 수문량 예측 분석 및 개선에 널리 활용하고 다양한 적용 사례가 나오길 기대한다.

사 사

본 연구는 K-Water의 장기 유량 예측기술 개발 및 용수공급 분석체계 구축 용역으로 수행되었으며 이에 감사드립니다. 본 결과물은 환경부의 재원으로 한국환경산업기술원의 물관리 연구사업의 지원을 받아 연구되었습니다. (과제번호 83081)

Reference

Abbot, J, and Marohasy, J (2014). Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks, *Atmospheric Research*, 138, pp. 166-178.

Anctil, F, and Nicolas, L (2004). Generalisation for neural networks through data sampling and training procedures, with applications to streamflow predictions, *Hydrology and Earth System Sciences*, 8(5), pp. 940-958.

Boucher, MA, Perreault, L, and Anctil, F (2009). Tools for the assessment of hydrological ensemble forecasts obtained by neural networks, *J. of Hydroinformatics*, 11.3-4: 297-307.

Cannon, AJ, and Whitfield, PH (2002). Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models, *J. of Hydrology*, 259, pp. 136-151.

Huang, NE, Shen, Z, Long, SR, Wu, MC, Shih, HH, Zheng, Q, Yen, NC, Tung, CC, and Liu, HH (1998). The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 354, pp. 903-995.

Jung, TH, Kim, H, Kim, H, and Heo, JH (2019). Selection of climate indices for nonstationary frequency analysis and estimation of rainfall quantiles, *J. of the Korean Society of Civil Engineers*, 39(1), pp. 165-174. [Korean Literature]

Kim, H, Kim, T, Shin, H, and Heo, JH (2017). A study on a tendency of parameters for nonstationary distribution using ensemble empirical mode decomposition method, *J. of Korea Water Resources Association*, 50(4), pp. 253-261. [Korean Literature]

Kim, SE, and Seo, IW (2015). Artificial neural network ensemble modeling with conjunctive data clustering for water quality prediction in rivers, *J. of Hydro-Environment Research*, 9 (3), pp. 325-339

Kim, T, Shin, JY, Kim, S, and Heo, JH (2018). Identification of relationships between climate indices and long-term precipitation in South Korea using ensemble empirical mode decomposition, *J. of Hydrology*, 557, pp. 726-739.

Kim, Y, and Cho, K (2016). Sea level rise around Korea: analysis of tide gauge station data with the ensemble empirical mode decomposition method, *J. of Hydro-environment Research*, 11, pp. 138-145.

Kolen, JF, and Pollack, JB (1991). Back propagation is sensitive to initial conditions, *Advances in Neural Information Processing Systems*, 3, pp. 860-867.

Lee, T, and Ouarda, TBMJ (2019). Multivariate nonstationary oscillation simulation of climate indices with empirical mode decomposition, *Water Resources Research*, 55. https://doi.org/ 10.1029/2018WR023892.

Li, J, Liu, X, and Chen, F (2015). Evaluation of nonstationarity in annual maximum flood series and the associations with large-scale climate patterns and human activities, *Water Resources Management*, 29(5), pp. 1653-1668.

Mekanik, F, Imteaz, MA, Gato-Trinidad, S, and Elmahdi, A (2013). Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes, *J. of Hydrology*, 503, pp. 11-21.

Santos-García, G, Varela, G, Novoa, N, and Jiménez, MF (2004). Prediction of postoperative morbidity after lung resection using an artificial neural network ensemble, *Artificial Intelligence in Medicine*, 30, pp.61-69.

Schepen, A, Wang, QJ, and Robertson, D (2012). Evidence for using lagged climate indices to forecast Australian seasonal rainfall, *J. of Climate*, 25, pp. 1230-1246.

Tiwari, MK, and Adamowski, J (2013). Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models, *Water Resources Research*, 49, pp. 6486-6507.

Wu, Z, and Huang, NE (2004). A study of the characteristics of white noise using the empirical mode decomposition method, *Proceedings of the Royal Society of London. Series A*, 460, pp. 1597-1611.

Wu, Z, Huang, NE, Long, SR, and Peng, CK (2007). On the trend, detrending, and variability of nonlinear and nonstationary time series, *PNAS*, 104(38), pp. 14889-14894.

Wu, Z, and Huang, NE (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method, *Advances in Adaptive Data Analysis*, 1(1), pp. 1-41.

Zaier, I, Shu, C, Ouarda, TBMJ, Seidou, O, and Chebana, F (2010). Estimation of ice thickness on lakes using artificial neural network ensembles, *J. of Hydrology*, 383, pp. 330-340.