



Article

강 유량의 계절적 변화를 나타내는 통계 모델 개발

박설아, 김근영, 안수현, 문우석*

국립부경대학교 지구환경시스템과학부

Development of a stochastic model representing seasonal variability in river discharge

Seola Park, Geun Yeong Kim, Soohyun Ahn, Woosok Moon*

Division of Earth and Environmental System Sciences, Pukyong National University, Busan 48513, Republic of Korea

Received: July 30, 2024 / Revised: August 22, 2024 / Accepted: August 22, 2024

*Corresponding author: +82-51-629-6643 / E-mail: woosok.moon@gmail.com

요약: 강 유량은 홍수와 가뭄과 같은 극단적 기상현상을 정량화하는 중요한 측정량이므로, 이를 통계적으로 표현하는 것이 중요하다. 통계적인 방법론으로는 ARMA, ARIMA와 같이 현재를 기준으로 과거의 상태와 불확실성의 정도를 결합하여 회귀모형을 만드는 방법을 사용하였다. 계절적 주기성을 제거하여 남은 변동성이 계절성을 갖지 않는다는 가정아래 통계 모형을 적용하지만 남은 변동성에도 계절성은 강하게 존재하므로 회귀 모형의 적용은 한계를 나타낸다. 본 연구에서는 이러한 한계점을 보완한 확률적 모형을 강 유량 데이터에 적용해 강의 안정도를 나타내는 $a(t)$ 와 변동성을 나타내는 $N(t)$ 를 구하였다. 안정도 $a(t)$ 는 양의 값일 때 불안정성이 누적되고 이러한 축적으로 시스템이 최대 분산에 도달하여 불안정성이 해소되는 현상을 기억효과라고 한다. 모델로 구축된 변수를 통해 계절적 변동성에서 장기 강제력을 제거하여 강 시스템이 계절적 변동성에 의해서만 변하는 시스템을 확인하였으며, 강 유량과 온도의 장기 강제력 간의 상관관계를 분석하여 강의 대기 온도에 의한 지역적 영향을 알아냈다. 추후 연구에서 모든 가능한 강에 적용함으로써 강의 시스템을 분류하는 것이 가능함을 시사한다.

주요어: 강 유량, 확률적 모델, 계절적 변동성, 안정도

ABSTRACT: River discharge is a crucial factor in extreme weather events such as floods and droughts, thus constructing stochastic models for river discharge is important for analysis and prediction. Stochastic models such as ARMA and ARIMA are commonly used for this analysis, but even after removing seasonality, the seasonal component itself does not disappear. The stochastic model is applied under the assumption that the remaining variability does not have seasonality by removing the seasonal variability, but since the remaining variability also has strong seasonality, the application of the stochastic model shows limitations. In this study, a stochastic model that complements these limitations is applied to river discharge data to obtain $a(t)$, which represents the stability of the river, and $N(t)$, which represents the variability. When stability $a(t)$ has a positive value, instability accumulates and the phenomenon in which the system reaches maximum variance the instability is resolved through this accumulation is called the memory effect. By removing long-term forcing from the original data, we obtain systems where river flow changes only due to the seasonal stability and the short-term noise. We also analyzed the correlation between the long-term forcing of river discharge and that of surface air temperature to determine the influence of sea surface temperature in specific regions on river discharge. This study suggests that by applying this method to all available rivers, it is possible to classify the variability of river discharge by distinct time-scales.

Key words: river discharge, stochastic model, seasonal variability, stability

1. 서언

강 유량(river discharge)은 홍수와 가뭄을 예측하는 데 있어 중요한 변수로 작용한다(Hirabayashi *et al.*, 2008; Hasanah *et al.*, 2013). Hirabayashi *et al.* (2008)은 기후변화로 인한 강수량 증가와 증발산량 감소로 인해 전세계적으로 강 유량이 증가할 것임을 시사했으며, 이는 강 유량이 기후 변화를 나타내는 중요한 변수임을 나타내고 있다. 하천 흐름 역학은 파악하기 힘든 수많은 프로세스에 영향을 받기 때문에 확률적 특성으로 이해되어야 한다(Zounemat-kermani, 2016). 과거부터 데이터가 가지는 확률적 특성을 재생할 수 있는 확률적 모델링을 통해 유량을 분석하려는 노력이 지속되어 왔다(Modarres, 2007; Bonakdari *et al.*, 2020).

확률적 모델링의 종류에는 대표적으로 ARMA (Autoregressive and Moving Average) 모델과 ARIMA (Autoregressive Integrated Moving Average) 모델이 있다. ARMA 모델은 자기회귀 모델(Auto-Regressive)과 이동평균 모델(Moving-Average)을 결합하여 계절성을 제거한 후, 정상성(Stationary)을 띠는 상태에서 사용 가능했다. ARMA 모델을 활용한 강 유량 연구가 현재까지 활발히 진행되었다(Otache *et al.*, 2011; Liu *et al.*, 2015).

Otache *et al.* (2011)는 ARMA 모델과 PAR (Periodic Autoregressive) 모델로 강 유량 예측 성능을 비교하였으며, FCM (Fuzzy Clustering Method) 방법으로 데이터를 계절에 따라 분할하여 각 AR 모델로 PAR 모델을 구성하였다. 비교 결과, 계절 편차와 계절 상관 구조의 변동성을 설명 가능한 PAR 모델이 ARMA 모델보다 좋은 성능을 보였지만, ARMA 모델은 하천 흐름 역학을 이해하는 기초를 위한 예비 모델로 사용할 수 있는 장점이 존재하였다. 이후, Liu *et al.* (2015)은 ARMA 모델을 확장하여 외부 요인을 함께 고려할 수 있는 외생변수를 포함한 ARMAX 모델을 구축하였다.

반면, 비정상성(Non-stationary)을 띠는 데이터는 차분(Differencing)을 통해 추세를 제거한 후 정상성을 띠도록 변형하여 사용하며 이를 ARIMA (Autoregressive Integrated Moving average) 모델이라 한다. ARIMA는 선형적 과정을 모델링하므로, 강수가 유량으로 변환되는 과정에 내재된 비선형 과정을 나타낼 수 없었다(Nigam *et al.*, 2014).

또한, ARMA와 ARIMA 모델은 데이터에서 계절성을 완전히 제거하였다고 가정된 모델이지만, 데이터에서 계절성을 제거하더라도 계절성 자체가 사라지는 것은 아니다. 계절에 따라서 나타나는 섭동(Fluctuation)의 정도가 달라지는 것을 볼 수 있으며 즉, 월별 편차도 계절성을 가지고 있다. 평균에서도 계절성이 나타나지만 편차에서도 계절성이 강하게 존재하므로 정상성을 가진다고 볼 수 없다. 섭동

(Fluctuation)의 계절성은 정상성(Stationary)이라고 보지 않고 주기 정상성(Cyclostationary)을 가진다고 본다. 그림 1은 시베리아 지역의 Yenisey 강의 원 데이터와 계절성을 제거한 데이터를 나타낸 그림으로, 평균 계절성을 제거한 데이터에도 여전히 계절성의 흔적이 남아있음을 확인할 수 있다. 월별 기후 데이터는 지배적인 계절 싸이클과 계절성을 가지는 섭동의 합으로 구성된다(Moon and Wettlaufer, 2017). 따라서 계절성이 완전히 제거되었다고 가정한 ARMA와 ARIMA 모델은 계절 강 유량 표현과 예측에 적합하지 않다.

Moon and Wettlaufer (2017)에서 계절성은 수많은 물리적 과정의 조합으로 구성되어 있기 때문에 분해하여 파악하기 어렵다는 문제점을 지적하였다. 해당 문제점을 바탕으로 새로운 확률적 모델을 구축하였으며, 기후 데이터에서 안정/불안정 과정, 변동성, 장기 강제력을 한눈에 파악할 수 있도록 하였다. 이후 Moon and Hannachi(2021)는 확률적 모델 식(1)을 Nile강에 적용하여 강에 존재하는 안정성/불안정성을 나타내는 계수 $a(t)$ 와 강의 변동성을 나타내는 변수 $N(t)$ 의 두 가지 주기적 함수의 관계성을 설명하며, Nile 강의 안정성과 변동성을 분석하였다.

$$\frac{d\eta(t)}{dt} = a(t)\eta(t) + N(t)\xi(t) \quad (1)$$

본 연구에서는 식(1)을 시베리아 지역의 Yenisey 강과 미국 Oregon Coast 지역의 Umpqua 강에 적용하여 각 강의 유량과 $a(t)$, $N(t)$ 사이의 관계성을 파악하고 장기 강제력을 제거한 상태의 계절성을 파악해보고자 한다.

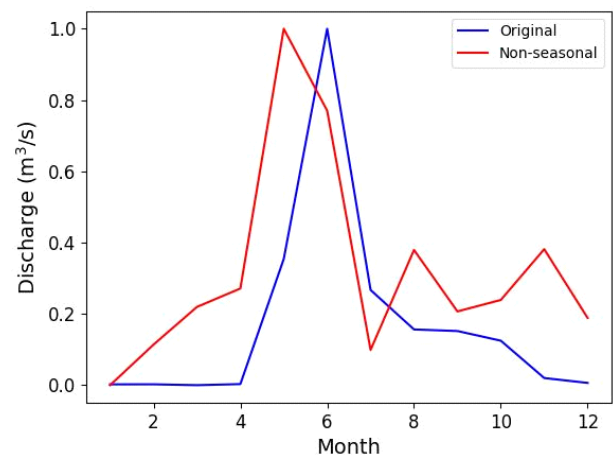


Fig. 1. Comparison of seasonal removed data and original data from the period 1940-2010 of the Yenisey River located in the Siberian region. Non-seasonal data also show that climate cycles exist.

2. 데이터 및 연구방법

2.1. 데이터

강 유량 데이터는 독일 Koblenz에 있는 독일연방수문연구소(BFG, German Federal Institute of Hydrology)의 지구유출량데이터센터(GRDC, Global Runoff Data Centre)에서 운영하는 데이터베이스에서 가져온 것으로, Yenisey 강의 경우, 경도 86.48°, 위도 69.43°지역의 1940년부터 2010년까지의 데이터를 사용하였다. Umpqua 강의 경우, 경도 -123.56°, 위도 43.59°지역의 1940년부터 2020년까지의 데이터를 사용하였다. 온도 데이터는 유럽 중기 예보 센터(ECMWF)에서 제공하는 ERA5재분석 자료를 이용하였으며(Hersbach *et al.*, 2020), 0.25° × 0.25°의 경위도 격자 데이터를 사용하였고, 1940년부터 2020년까지의 데이터를 포함한다. ERA5재분석 자료는 본래 1979년부터 2020년까지 사용할 수 있었던 데이터가 1940년부터 1979년까지의 기간을 back extension하여 최종 수정되었으며, 이는 지상 관측 데이터와 높은 일관성을 보이기에 관측 데이터를 대체하여 사용 가능성을 시사한다(Yilmaz, 2023).

2.2. 연구 방법

2.2.1. 이론적 배경

기후 데이터는 지배적인 계절 진동과 그 외의 변동성의 합으로 구성된다고 가정하여, 원 데이터에서 평균 계절 싸이클을 제거함으로써 기후 변동성을 분석할 수 있다. 계절성을 강하게 가지는 기후 데이터는 일반적으로 다음과 같은 수식으로 표현할 수 있다:

$$\frac{dx}{dt} = F(x,t) + N(t)\xi(t) + D(\tau) \quad (2)$$

여기서 x 는 기후변수로 $F(x,t)$ 는 지배적인 계절성인 주기 강제력을 의미한다. $F(x,t)$ 는 $F(x,t) = F(x,t+T)$ 를 만족하는, $T = 1$ 년인 주기 함수이며, $N(t)\xi(t)$ 와 $D(\tau)$ 는 섭동을 만들어내는 물리과정들을 의미한다. 계절에 비교하여 짧은 시간동안 만들어지는 날씨 관련 프로세스의 영향은 통계적인 잡음 $N(t)\xi(t)$ 으로 표현되었다. 비록, 날씨를 주관하는 대기의 대류 현상은 카오스적인 성질을 지니지만 대류현상의 시간 스케일은 계절에 비해서 상당히 작다. 즉, 두 물리적 시간 스케일 사이에는 분명한 분리를 설정할 수 있다. $\xi(t)$ 는 날씨와 계절 사이에 명확한 분리가 존재한다는 것을 암시하는 가우시안 백색 잡음(White noise)이다. $N(t)$ 는 $\xi(t)$ 와 곱해지면서, 날씨의 강도를 나타내며, 이는 계절에 따라 달라지고, $N(t) = N(t+T)$, $T = 1$ 년인 주기성을 가진다. 날씨의 강도는 중위도 저기압의 활동성과 비례하며 중위도 저기압의 활동성은 적도와 극의 온도차이에 비례한다. 일

반적으로 겨울철 중위도 날씨의 영향력이 여름에 비해서 강하게 나타난다. $D(\tau)$ 는 날씨와 반대로 계절보다 긴 시간 스케일에서 일어나는 물리적 과정에서 비롯된 강제력을 나타낸다. 해양의 순환 혹은 인간이 만들어내는 지구 온난화의 영향력이 그 예가 될 수 있다. 계절적 순환에 비해 굉장히 느리게 일어나기에, 지배적인 시간인 t 와 구별되는 긴 시간인 τ 를 사용한다. 같은 시간 변수임에도 다른 변수를 사용하는 것은 $D(\tau)$ 가 지배적인 시간의 적분과 미분에서 상수처럼 작용할 수 있다는 것을 의미한다.

그림 1에서도 보듯이 기후 데이터의 계절 주기성은 날씨 혹은 해양의 영향력에 비해 큰 값을 가진다. 본 절의 $\langle \cdot \rangle$ 는 각 월에 해당하는 모든 년도의 평균으로 $F(\langle x \rangle, t) \gg N(t)$ 를 만족한다. 평균장은 다음과 같이 수학적 모델에서 정의할 수 있다:

$$\frac{d\langle x \rangle}{dt} = F(\langle x \rangle, t) + \langle D(\tau) \rangle \quad (3)$$

기후 데이터인 유량은 평균장과 섭동의 합 $x = \langle x \rangle + \eta$ 로 표현할 수 있으므로 식(2)에서 평균장을 제거할 수 있다:

$$\begin{aligned} \frac{d}{dt}(\langle x \rangle + \eta) &= F(\langle x \rangle + \eta, t) + N(t)\xi(t) + \langle D \rangle + d(\tau) \\ &\approx F(\langle x \rangle, t) + \left. \frac{\partial F}{\partial x} \right|_{x=\langle x \rangle} \eta + N(t)\xi(t) + \langle D \rangle + d(\tau) \end{aligned} \quad (4)$$

$$\frac{d\eta}{dt} = a(t)\eta(t) + N(t)\xi(t) + d(\tau) \quad (5)$$

여기서 계절적 안정도인 $a(t) = \left. \frac{\partial F}{\partial x} \right|_{\langle x \rangle}$ 를 정의할 수 있으며, $a(t) = a(t+T)$, $T = 1$ 년인 주기성을 가진다. $a(t)$ 는 그 부호에 따라서 섭동의 시간적 변화의 양상을 결정한다. 만약, $a(t)$ 가 양의 부호를 가지면 섭동은 시간에 따라서 지수적으로 증가하는 것을 의미하며, 반면에 음의 부호를 가지면 지수적으로 감소하여 빠른 속도로 제로로 수렴하는 것을 나타낸다. 즉, 양의 부호를 가지는 것은 그 순간 시스템이 불안정하다라는 것을 나타내며, 음의 부호는 안정됨을 나타내고 있다. $a(t)$ 의 부호와 절대값의 크기에 따라서 계절적 싸이클이 가지는 안정도를 나타낸다. $d(\tau)$ 는 계절적 변동에 비해 매우 느리게 일어나기 때문에 원 데이터에서 계절적 주기성을 제거한 값을 1년 시간 평균하여 $\bar{d} = \overline{d(\tau)\Delta t}$ 로 구할 수 있다:

$$\bar{d} = \overline{\eta(t+\Delta t) - \eta(t) - a(t)\eta(t)\Delta t} \quad (6)$$

본 연구에서는 강 유량 데이터를 활용하여 식(5)를 구성

하는 주기 함수 $a(t)$ 와 $N(t)$ 의 계절 값을 건설하는 것이 우리의 목표이다.

2.2.2. 주기 함수의 건설

강 유량 데이터는 월별 데이터로 존재하기에, $\Delta t = 1/12$ 로 정의할 것이다. 여기서 시간 1은 1년을 의미하며, 무작위 효과(random effect)를 데이터에 적용하기 위해 Δt 동안 움직이는 브라운 입자(Δw , brownian particle)를 $\xi(t) = \frac{\Delta w}{\Delta t}$ 로 정의한다. 따라서 주기 함수 $a(t)$ 와 $N(t)$ 의 값을 건설하기 위해 연속성(continuous)을 지닌 식(5)를 식(7)과 같이 이산화(discretize)시킬 수 있다:

$$\Delta \eta = a(t)\eta(t)\Delta t + N(t)\Delta w + d(\tau)\Delta t \quad (7)$$

기본 식 $\Delta \eta = \eta(t + \Delta t) - \eta(t)$ 에서 분산의 형태를 만들어 주기 위해 식(7)에 $\eta(t)$ 를 곱해주어 정리할 수 있다.

$$\begin{aligned} \eta(t)(\eta(t + \Delta t) - \eta(t)) &= a(t)\eta^2(t)\Delta t + \\ N(t)\eta(t)\Delta w + d(\tau)\eta(t)\Delta t \end{aligned} \quad (8)$$

$$\langle \eta(t)\eta(t + \Delta t) \rangle - \langle \eta^2(t) \rangle = a(t)\langle \eta^2(t) \rangle \Delta t \quad (9)$$

식(9)의 평균을 구하면, $N(t)$ 가 포함된 항과 $d(\tau)$ 가 포함된 항은 백색 가우시안 잡음의 성격으로 말미암아 0이 된다. 또한, $\langle N(t)\Delta w \rangle \Delta t$ 가 백색 잡음(White noise)과정을 나타내기 때문에 소거된다. $\langle d(\tau)\eta(t)\Delta t \rangle$ 경우, $d(\tau)$ 는 식(6)에서 설명한 바와 같이 년도별로 각기 다른 값을 가지므로 월 별로 변하는 계절성에 비해 매우 미미한 변화를 나타내기 때문에 사라진다.

식(9)에 의해 $a(t)$ 는 다음과 같이 건설할 수 있다:

$$a(t) = \frac{1}{\Delta t} \frac{\langle \eta(t)\eta(t + \Delta t) \rangle - \langle \eta^2(t) \rangle}{\langle \eta^2(t) \rangle} \quad (10)$$

식(7)에서 $\eta(t + \Delta t) - \eta(t) - a(t)\eta(t)\Delta t$ 를 $y(t)$ 로 치환하면 다음과 같은 식을 도출할 수 있다:

$$y(t) = N(t)\Delta W_t + d(\tau)\Delta t \quad (11)$$

$$y(t + \Delta t) = N(t + \Delta t)\Delta W_{t+\Delta t} + d(\tau)\Delta t \quad (12)$$

여기서 식(12)는 장기 강제력의 항을 제거하기 위해 도입한 것으로, $d(\tau)$ 는 월별로 변하는 $N(t)$ 와 달리 연별로 천천히 변하기 때문에 두 항의 상관성은 없는 것으로 볼 수 있다. 따라서 $\langle \cdot \rangle$ 의 평균을 사용하여 다음과 같은 식으로 근

사할 수 있다:

$$\langle y^2(t) \rangle \cong N^2(t)\Delta t + d^2(\tau)\Delta t^2 \quad (13)$$

$$\langle y(t)y(t + \Delta t) \rangle \cong d^2(\tau)\Delta t^2 \quad (14)$$

$$N(t) = \sqrt{\frac{\langle y^2(t) \rangle - \langle y(t)y(t + \Delta t) \rangle}{\Delta t}} \quad (15)$$

2.3. 통계적 모델의 적용

$x_{i,k}$ 는 강의 유량 데이터라고 가정하면, i 는 $1 \leq i \leq N$ 값의 년도를, k 는 $1 \leq k \leq 12$ 의 값의 월을 의미한다. 계절적 변동성을 분석하기 위해 원 데이터에서 기후 평균값(계절성)을 제거해야 한다. 기후 평균값(계절성)은 다음과 같이 나타낸다:

$$\langle x \rangle_k = \frac{1}{N} \sum_{i=1}^N x_{i,k} \quad (16)$$

기후 변동성은 $\eta_{i,k} = x_{i,k} - \langle x \rangle_k$ 로 나타내며, 이를 이용하여 데이터의 평균장을 제거하여 섭동장만 확인할 수 있다. $a(t)$ 와 $N(t)$ 는 다음과 같은 분산(Variance)과 자기 상관(Auto-covariance)을 이용하여 구할 수 있다:

$$V_k = \frac{1}{N-1} \sum_{i=1}^N \eta_{i,k}^2$$

$$A_k = \frac{1}{N-1} \sum_{i=1}^N \eta_{i,k}\eta_{i,k+1}$$

이를 통해 a_k 를 다음과 같이 정의한다:

$$a_k = \frac{A_k - V_k}{V_k \Delta t} \quad (17)$$

이제 $y_{i,k}$ 를 $\eta_{i,k+1} - \eta_{i,k} - a_k\eta_{i,k}\Delta t$ 로 정의하여 N_k 를 도출할 수 있다:

$$N_k = \sqrt{\frac{1}{\Delta t} \frac{1}{N-1} \sum_{i=1}^N y_{i,k}^2} \quad (18)$$

$d(\tau)$ 는 정의한 $y_{i,k}$ 를 1년 시간 평균으로 구할 수 있다:

$$d_{i,k} = \sum_{k=1}^{12} y_{i,k} \quad (19)$$

$d_{i,k}$ 는 연평균을 했기 때문에, 같은 해의 모든 월은 동일

한 값을 가진다. 하지만 지역마다 a_k 의 값이 다르기에 지역 적 민감도에 따라 값이 달라진다.

2.3.1. 장기 강제력의 제거

모델을 통해 얻은 식에는 장기 강제력 $d(\tau)$ 가 포함되어 있어 계절성만을 명확히 보기 어렵다. 식(5)를 통해 해를 구할 수 있다:

$$\begin{aligned} \eta(t) = & \exp\left(\int_0^t a(s)ds\right) \int_0^t N(t') \\ & \exp\left(-\int_0^t a(s)ds\right) d\tau + \\ & \exp\left(\int_0^t a(s)ds\right) \int_0^t d(\tau) \exp\left(-\int_0^t a(s)ds\right) dt' \end{aligned} \quad (20)$$

식(20)에서 $d(\tau)$ 는 계절적 순환 t 에 비해 굉장히 느리게 일어나기에, 앞절에서 언급한 바와 같이 지배적인 시간의 적분에서 상수처럼 작용한다. 따라서 다음과 같은 식을 통해 장기 강제력을 제거하여 계절성을 확인할 수 있다:

$$\begin{aligned} \eta(t) - d(\tau) \exp\left(\int_0^t a(s)ds\right) \int_0^t \exp\left(-\int_0^t a(s)ds\right) dt' \\ = \exp\left(\int_0^t a(s)ds\right) \int_0^t N(t') \exp\left(-\int_0^t a(s)ds\right) dt' \end{aligned} \quad (21)$$

위 식에서 $\exp\left(\int_0^t a(s)ds\right) \int_0^t \exp\left(-\int_0^t a(s)ds\right) dt'$ 는 강의 불안정성에 따라 장기 강제력이 반응하는 정도를 나타내는 항이다. $a(t)$ 는 계절적 주기성을 가지며, $0 \leq \tilde{t} < 1$ 이기 때문에 $t = nT + \tilde{t}$ 를 사용하여 해당 식을 계산할 수 있다(Moon and Wettlaufer, 2013). T 는 1년을 의미하기 때문에, 양의 상수 γ 를 다음과 같이 정의한다:

$$\gamma = -\int_0^T a(s)ds \quad (22)$$

따라서,

$$\begin{aligned} \exp\left(\int_0^t a(s)ds\right) = \exp\left(n \int_0^T a(s)ds + \int_0^{\tilde{t}} a(s)ds\right) \\ = e^{-n\gamma} \exp\left(\int_0^{\tilde{t}} a(s)ds\right) \end{aligned} \quad (23)$$

$$\begin{aligned} \int_0^t \exp\left(-\int_0^t a(s)ds\right) dt' = e^{\gamma n} \int_0^{\tilde{t}} \exp\left(-\int_0^t a(s)ds\right) dt' \\ + \frac{1 - e^{\gamma n}}{\gamma} \int_0^T \exp\left(-\int_0^t a(s)ds\right) dt' \end{aligned} \quad (24)$$

식(23)과 식(24)를 곱하여 최종 식(25)로 나타낸다:

$$\begin{aligned} \exp\left(\int_0^t a(s)ds\right) \int_0^t \exp\left(-\int_0^t a(s)ds\right) dt' \\ = \exp\left(\int_0^{\tilde{t}} a(s)ds\right) \left\{ \int_0^{\tilde{t}} \exp\left(-\int_0^t a(s)ds\right) dt' \right. \\ \left. + \frac{e^{\gamma n} - 1}{\gamma} \int_0^T \exp\left(-\int_0^t a(s)ds\right) dt' \right\} \end{aligned} \quad (25)$$

$e^{\gamma n}$ 은 η 가 커짐에 따라 0으로 수렴하기 때문에, 위 식으로 장기 강제력이 반응하는 정도를 계산할 수 있다.

2.3.2. 모델의 신뢰도 검증

모델은 계절적 주기와 날씨의 효과만을 포함한 주기성을 확인하기 위해 장기 강제력을 제거한 식(21)을 기반으로 식(26)을 구성하였다.

$$\frac{d\eta}{dt} = a(t)\eta(t) + N(t)\xi(t) \quad (26)$$

2.2.2 절과 마찬가지로, 식(26)을 이산화(discretize)시키고 Δt 를 양 변에 곱하여 다음과 같은 모델의 기본 식을 도출하였다:

$$\Delta \eta = a(t)\eta \Delta t + N(t) \Delta W \quad (27)$$

$$\eta_{n+1} = (1 + a_n \Delta t)\eta_n + N_n \Delta W_n \quad (28)$$

날씨의 무작위성(Randomness)을 모델에 반영하기 위해 평균이 0이고 분산이 $\Delta t = 1/12$ 를 따르는 정규분포를 브라운 입자(Brownian particle) ΔW_n 에 적용하였다. 원 데이터의 첫번째 값을 바탕으로 식(28)에서 도출된 값을 이용하여 모델의 표준편차를 계산하였다. 이러한 무작위성을 고려하면 그림 2a와 2c로 모델링되며, (a)는 Yenisey 강의 장기 강제력을 제거하지 않았을 때의 모델링이고, (b)는 제거하였을 때의 모델링이다. 이러한 두 모델의 신뢰도는 모델의 표준편차와 실제 데이터의 표준편차를 비교함으로써 확인할 수 있다. 표준편차는 데이터의 분포를 나타내며, 두 값이 비슷하면 모델이 유의미함을 의미한다. 모델의 신뢰도는 모델의 표준편차와 원 데이터의 표준편차를 비교하여 평가할 수 있다.

그림 2b와 2d를 통해 모델의 표준편차를 분석할 수 있으며, (b)는 강의 장기 강제력을 제거하지 않은 모델의 표준편차와 원 데이터의 표준편차를 비교한 것이고, (d)는 장기 강제력을 제거한 모델의 표준편차와 원 데이터의 표준편차를 비교한 것이다. 장기 강제력을 제거하여 계절적 주기와

날씨의 효과만을 포함한 모델이 장기 강제력을 제거하지 않은 모델보다 데이터의 분포를 더 정확하게 반영하고 있음을 알 수 있다.

3. 결과

식(1)에서 첫번째 항 $F(x,t)$ 는 계절에 따라 변하는 강물의 유입과 유출을 나타내는 지배적인 항으로, 강물의 증발(유출), 강수량(유입) 등을 의미한다. 특히, 주기성을 가지며, $F(x,t) = F(x,t + T)$, $T = 1$ 년을 만족한다. 두번째 항 $N(t)\xi(t)$ 는 날씨의 효과를 나타내는 항으로, 중위도 저기압의 영향으로 내리는 강수의 영향을 나타낸다. 세번째 항인 $D(\tau)$ 는 계절적 순환에 비해 느리게 발생하는 항으로, 엘니뇨나 라니냐에 의한 강 유량의 변화 또는 지구 온난화로 인한 강 유량의 변화를 의미한다. 예를 들어, 엘니뇨 시기에 강 유량이 크게 증가하거나 라니냐 시기에 강 유량이 크게 감소하는 상황을 모델에 반영한다.

$a(t)$ 는 월 별 안정도를 나타내는 변수로, 계절적 민감도의 정보를 제공해준다.

$$\frac{d\eta}{dt} = a(t)\eta \quad (29)$$

식(29)의 해는 $\eta(t) = \eta_0 e^{\int_0^t a(t)dt}$ 로, $a(t)$ 의 값이 양수라면

유량의 섭동이 기하급수적으로 늘어나 원래 상태에서 벗어나 불안정해지는 것을 나타낸다. 반대로, $a(t)$ 의 값이 음수라면 유량의 섭동이 0으로 수렴하는데 이는 원래 상태를 유지하려는 시스템의 성질을 반영한 것으로 안정된 상태를 반영한다.

$N(t)$ 는 날씨의 강도를 나타내기 때문에, 큰 값을 가지면 $N(t)\xi(t)$ 의 값이 커져 짧은 시간 동안의 예측할 수 없는 변동이 매우 커진다. 이는 기상 이변의 변동성과 잠재성이 커져 극단적인 날씨인 홍수나 가뭄의 확률이 증가함을 의미한다. 예를 들어, $N(t)$ 가 큰 우기의 경우 날씨의 변동성이 커지고 홍수 발생 확률이 증가한다. 반면, $N(t)$ 가 작은 건기의 경우 날씨의 변동성이 작아지고 예측 가능하고 안정적인 상태가 유지된다.

전체적인 변동성은 안정성을 나타내는 $a(t)$ 와 날씨의 변

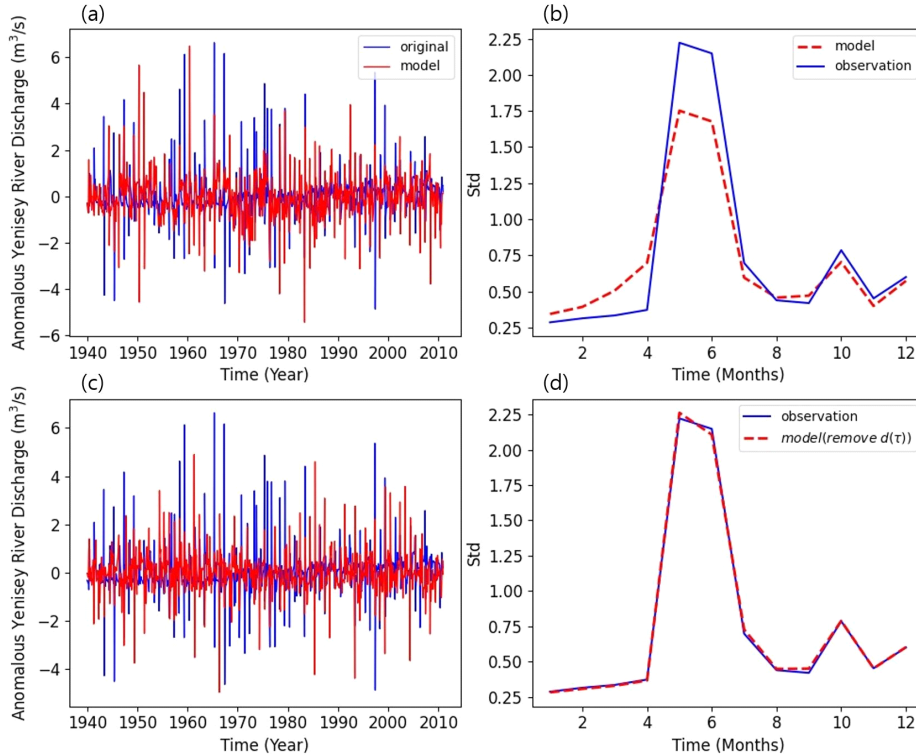


Fig. 2. We used river flow data from the Yenisei River in Siberia with climatological mean seasonal cycle removed from 1940 to 2020. The top two figures (a) and (b) show the results when the long-term forcing $D(\tau)$ is not removed from the original data, and the bottom two figures (c) and (d) show the results when the long-term forcing $D(\tau)$ is removed from the original data. (a) Time series of the river discharge data (blue), and a model simulation without long-term forcing removed (red). (b) Monthly standard deviations for observed data (blue) and model without long-term forcing removed (red). (c) Time series of the river discharge data (blue), and a model simulation with long-term forcing removed (red). (d) Monthly standard deviations for observed data (blue) and model with long-term forcing removed (red).

동성을 조율하는 $N(t)\xi(t)$ 의 선형적인 결합을 통해서 결정된다. 날씨의 변동성이 커지고 계절 안정도가 불안정하면 전체적인 변동성이 커진다. 그리고 계절 안정도가 안정하면 날씨의 변동성이 커지더라도 전체적인 변동성은 작아질 수 있다. 안정도와 날씨의 변동성의 결합을 통해서 전체적인 변동성이 계절에 따라서 어떻게 변하는 지를 이해해야 한다.

강 유량 데이터는 단위 시간에 강의 단면적을 통과하는 부피로, 강 유량 값이 클수록 단위 시간에 강의 단면적을 통과하는 부피가 증가한다. 예를 들어, 예년보다 많은 강수량이 발생하면, 강에 존재하는 물의 부피가 증가하기 때문에, 강의 유량이 늘어난다. 반대로, 예년보다 적은 강수량이 발생하면, 강에 존재하는 물의 부피가 감소하기 때문에, 강의 유량이 줄어든다. 낮은 유량에서는 물이 깊은 채널 내에 갇혀 있기 때문에, 작은 유속 변화로도 유량이 빠르게 변한다. 높은 유량 조건에서는 유량이 증가하면서 물이 깊은 채널을 넘어 강의 평탄한 부분(예를 들어, 강둑)으로 퍼지게 된다. 이 경우, 유량이 조금 증가하더라도 유효폭이 크게 증가하여 물이 넓은 범위로 퍼지면서, 강의 너비가 급격하게 증가한다(Ashmore and Sauks, 2006). 유효폭은 강의 특정

구간에서 물이 차지하고 있는 전체 표면적을 구간의 총 길이로 나눈 값으로, 강의 폭과 유량 간의 관계를 분석하는데 유용하다.

이를 바탕으로 강에 대한 피드백 작용을 알 수 있으며, 우리는 극단적인 환경에서 강의 피드백 작용을 알아볼 것이다. 강이 높은 유량 조건에서 강수가 발생하면 유량이 증가하면서 유효폭이 크게 증가하고, 강의 너비가 급격하게 증가한다. 강의 채널이 너비가 증가함에 따라 변하면서 유량이 더욱 급격하게 증가할 수 있는 환경이 조성된다. 이러한 환경에서 추가적인 강수가 발생하면 강의 채널이 더 변하게 되고 유량이 더욱 급격하게 증가할 수 있는 더 좋은 환경이 조성된다. 즉, 높은 유량 조건에서 강수가 발생하면 강의 유량이 증가하는 양의 피드백이 발생한다. 반면, 강이 낮은 유량 조건에서 강의 퇴적물이 쉽게 가라앉고(Bobrovitskaya et al., 1996), 퇴적물의 침전이 증가해 강 바닥과의 마찰이 증가하고 유량이 감소한다. 낮은 유량 조건에서 강수가 발생하지 않으면, 강의 유량이 더 감소하게 된다는 음의 피드백이 발생한다.

그림 3a는 Yenisey 강의 안정성 지표인 $a(t)$ 와 변동성 지

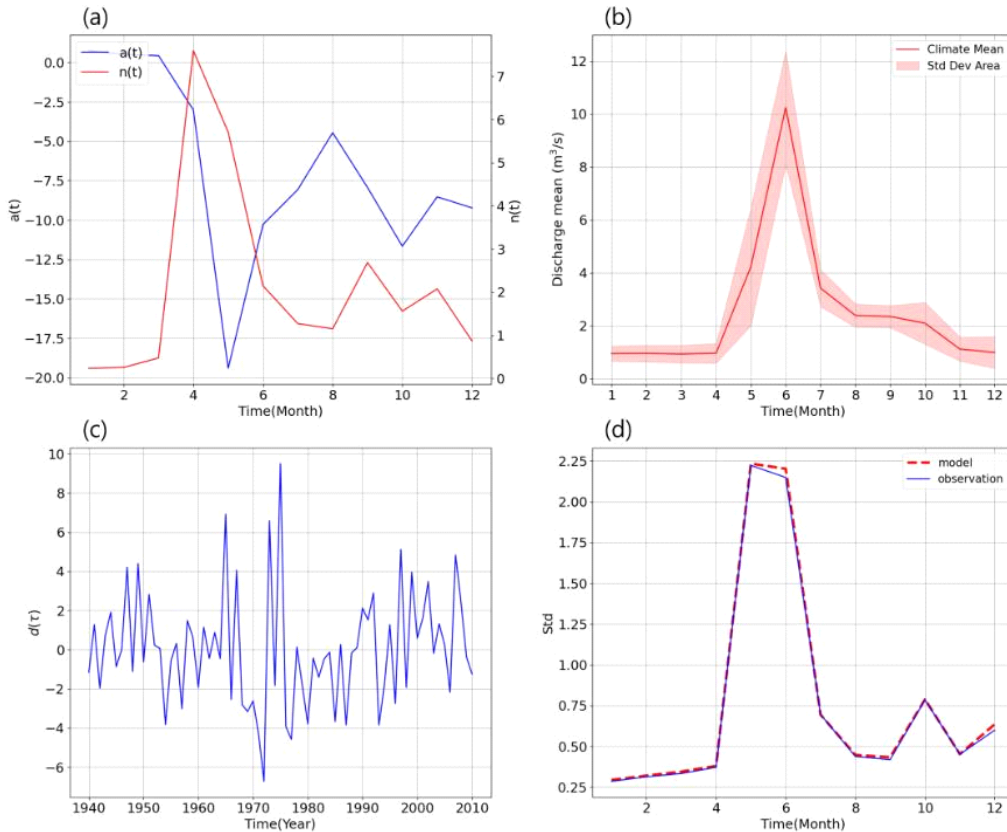


Fig. 3. Stochastic model construction based on the river discharge of the Yenisey for the period 1940-2010. Equations (9), (10) are used to construct $a(t)$, $N(t)$. (a) Two periodic functions: the stability function $a(t)$, noise amplitude $N(t)$. (b) Climatological mean with error bars. (c) The long-term forcing $D(\tau)$. (d) Monthly standard deviation of the observation data(blue) and the model (red).

표인 $N(t)$ 의 값을 월별로 나타낸 그래프이다. $a(t)$ 의 값이 1월, 2월, 3월까지 양수를 기록하고, 4월에 급격히 감소하는 양상을 보인다. $a(t)$ 가 양수일 때, 유량의 섭동($\eta(t)$)은 기하 급수적으로 증가하여 시스템이 불안정해진다. 반면, $a(t)$ 가 음수일 때, 유량의 섭동($\eta(t)$)은 0으로 수렴하여 시스템이 안정화된다. 따라서 $a(t)$ 가 양수일 경우 강은 불안정한 상태를 유지하며, $a(t)$ 가 음수일 경우 강은 안정한 상태를 유지한다.

그림 3b는 Yenisey 강 유량의 월별 기후평균을 나타내며, 5월에 유량이 급증, 6월에 최대 유량에 도달한 후 7월에 급감하는 경향을 보인다. 이는 $a(t)$ 와 $N(t)$ 로 설명된다. 그림 3a에서 $a(t)$ 가 양수 값을 가지는 1월, 2월, 3월은 강이 불안정한 상태를 가지고 있기 때문에 해당 기간 동안 강의 불안정성이 누적되어 최대 분산을 가지는 “기억 효과”가 발생한다(Moon and Wettlaufer, 2017). 이는 $N(t)$ 가 최대값에 도달했을 때 해소된다. Yenisey 강의 경우, $a(t)$ 의 부호가 양수에서 음수로 전환되는 시기가 3월과 4월 사이에 존재하며, 4월에 $N(t)$ 의 값이 최대에 도달한다. 최대값을 가진 $N(t)$ 는 유량의 변동성이 가장 크다는 것을 의미하며, 이는 5월 유량의 강 유량의 편차가 최대가 됨을 의미한다. 이

는 5월 유량의 급증을 초래하며, 이는 6월에 최대에 도달한다. 최대의 유량은 홍수와 같은 극단적인 기상 현상을 통해 강의 불안정성이 해소된다.

이러한 현상은 Yenisey 강뿐만 아니라 Umpqua 강에서도 나타난다. 그림 4a에서 $a(t)$ 의 값이 9월, 10월에 양수이고, 11월에 음수로 변하는 것을 확인할 수 있다. 2개월 동안 Umpqua 강의 불안정성이 누적되고, $N(t)$ 가 최대값에 도달했을 때 해소된다. Umpqua 강의 경우, $a(t)$ 의 부호가 양수에서 음수로 전환되는 시기가 10월과 11월 사이에 존재하며, 11월에 $N(t)$ 의 값이 최대가 된다. 이는 12월 유량의 급증을 야기하며, 1월에 최대 유량에 도달한다.

그림 3b, 그림 4b에서 두 강 모두 많은 양의 비가 내리는 우기와 비가 거의 내리지 않는 건기가 존재하고 있음을 알 수 있으며, 건기에는 비가 거의 내리지 않고 우기에 집중되어 비가 많이 내리고 있음을 그래프에서 확인할 수 있다. 즉, 강의 전체 과정은 안정하지만, 강이 불안정한 기간은 전체 기간에 비해 짧고, 안정한 기간들이 강력하게 안정하기 때문에, 이러한 불안정은 강력한 안정한 기간에 상쇄되어 강의 전체 과정은 안정하다는 것을 알 수 있다.

그림 3c와 그림 4c는 1년 마다 변하는 장기 강제력을 나

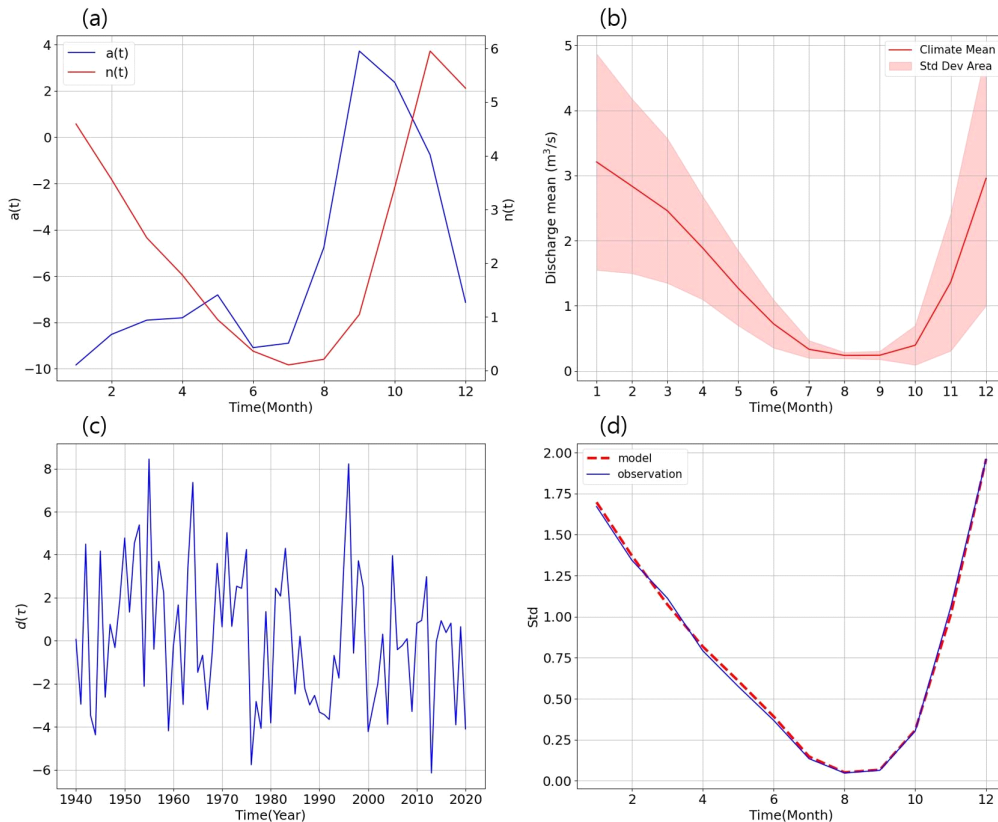


Fig. 4. Stochastic model construction based on the river discharge of the Umpqua for the period 1940-2020. Equations (9), (10) are used to construct $a(t)$, $N(t)$. (a) Two periodic functions: the stability function, noise amplitude $N(t)$. (b) Climatological mean with error bars. (c) Monthly standard deviation of the observation data (blue) and the model (red).

타낸 그림으로, 긴 주기성을 가지고 있음을 나타낸다. 그림 3d와 그림 4d는 모델의 표준편차와 실제 데이터의 표준편차를 비교한 그림으로, 모델과 데이터의 표준편차가 거의 일치하여 전체적인 데이터의 분포가 비슷함을 알 수 있다. 강의 장기 강제력을 제거하여 모델링한 그래프이기 때문에, 지배적인 주기성과 날씨의 효과만을 포함한다. 장기 강제력 없이 계절적으로 강이 어떻게 변하는지를 알 수 있다. 강의 장기 강제력 $d(\tau)$ 는 식(12)와 같이 필터링 가능하므로, 2 m 대기 온도의 $f(\tau)$ (강의 장기 강제력과 구분하기 위하여 2 m 대기 온도의 장기 강제력은 $f(\tau)$ 를 사용하였다)와 상관관계를 구하여 어느 지역의 온도가 해당 강의 장기 강제력에 얼마나 영향을 미치는지 알 수 있다. 그림 5a는 그림 3c의 $d(\tau)$ 와 2 m 대기 온도 $f(\tau)$ 상관관계를 시각화한 그림으로, 시베리아와 북극 전체 지역에서 약한 상관관계가 있음을 보여준다. 유량이 봄철 눈 녹음 기간동안 북극해로 유출(Ahmed *et al.*, 2020)되기 때문에 온도가 올라가면 유량이 증가(Overeem and Syvitski, 2010)하는 약한 양의 상관관계가 나타나는 것으로 보인다. 그림 5a와 5c를 비교해 보면, Yenisey 강에 PDO(Pacific Decadal Oscillation) 패턴이 약한 음의 상관성을 보임을 알 수 있다.

그림 5b는 그림 4c의 $d(\tau)$ 와 2 m 대기 온도 $f(\tau)$ 의 상관관계를 시각화한 그림으로, 강이 캘리포니아 지역에 위치하고 있으므로 주변 지역에서 약한 음의 상관관계가 나타나고 있고, Umpqua 강도는 PDO와 약한 음의 상관관계를 보이고 있음을 확인할 수 있다. 두 강의 유량 모두 PDO의 영향을 받는 것으로 보았을 때, PDO가 강 유량에 중요한 부분을 차지하고 있다고 기대해볼 수 있다(Neal *et al.*, 2002). 한편 PDO는 태평양에서 중위도와 열대 해수면온도의 가장 지배적인 모드로 첫번째 경험적 직교함수에 해당한다.

강 유량의 $d(\tau)$ 가 어느 기후 요소와 연관되어 있는지 파악하는 것은 전 세계 각 지역의 극단적 날씨를 이해하는데 중요한 지표가 된다. 이는 강의 장기적 변화를 예측하는데 유용한 정보를 제공하며, 강의 $d(\tau)$ 와 2 m 대기 온도의 $f(\tau)$ 의 상관성을 계산하여, 어느 기후 요소의 영향이 가장 강한

지를 파악한다면, 이는 그 기후 요소로 강에 존재하는 시스템을 이해하는 것이 가능해진다.

4. 결론

강 유량은 홍수, 가뭄 예측에 있어 중요한 변수로, ARMA와 ARIMA 모델을 활용하여 유량 예측 시도가 이루어졌다. ARMA와 ARIMA 모델은 계절성을 제거한 데이터로 모델링을 구현하지만 기후데이터에서는 계절성이 가장 주요한 변동으로서, 데이터에서 평균 계절 주기를 제거하더라도 계절성은 여전히 남아있다. 계절성 문제를 보완하기 위해 Moon and Hannachi (2021)에선 강 유량 데이터가 지배적인 계절적 주기성과 날씨, 장기 강제력과 같은 아주 작은 섭동의 합으로 구성되며, 이는 계수가 계절 주기성을 가지는 확률적 미분 방정식 식(1)과 같이 나타낼 수 있는 것을 확인하였다.

이를 바탕으로 본 연구는 주기 함수 $a(t)$ 와 $N(t)$ 를 도출하였고 각각 강 유량의 안정성에 대한 함수와 강 유량의 변동성 효과를 의미한다. 양의 $a(t)$ 일 때는 불안정한 상태이고 음의 $a(t)$ 일 때는 안정한 상태를 나타냄을 알 수 있다. 또한, $N(t)$ 가 매우 큰 값을 가질 때는 날씨의 변동성이 크다는 것을 의미하며, 이는 홍수와 같은 극단적인 기상현상의 발현 확률이 증가함을 시사한다. $N(t)$ 는 강의 불안정성이 계속 해서 누적되고, 이를 해소할 때 가장 큰 값을 가진다. 기후 데이터에 존재하는 장기적으로 영향을 미치는 항은 식(16)에 의해 제거할 수 있으며, 강 유량의 장기적 기후 영향성을 제거한 계절성을 파악할 수 있어 매우 의미 있는 자료이다. 장기적 기후 영향성을 제거한 데이터를 식(23)을 이용하여 모델링의 검증이 가능하며, 우리가 구축한 모델은 계절적으로 변하는 표준편차를 데이터와 거의 비슷하게 재생하며 유의미한 모델임을 보여주었다.

본 연구에서는 주기적 함수 $a(t)$ 와 $N(t)$ 의 효과를 확인하기 위해, 시베리아 지역에 위치하는 Yenisey 강과 미국 Oregon Coast의 Umpqua 강의 유량 데이터를 이용하여 분석하였다. 두 강 모두 비가 거의 내리지 않는 건기와 비가 내리는

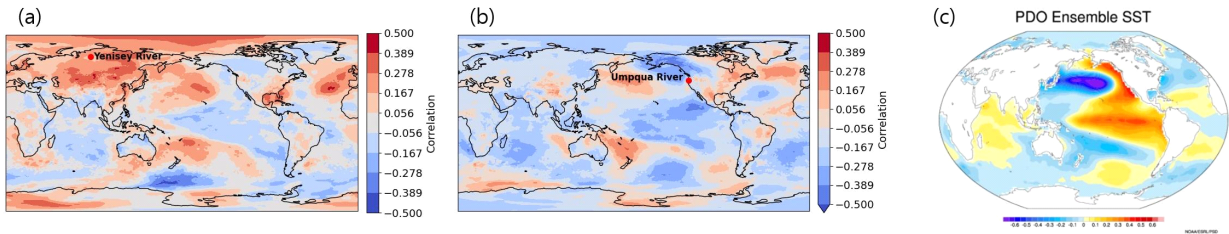


Fig. 5. The correlation between $D(\tau)$ of a river and $f(\tau)$ of 2 m temperature. During the period 1940-2010, the correlation of river and long-term forcing was obtained at all points, where (a) is the Yenisey river and (b) is the Umpqua river. (c) The PDO patterns and series from various SST datasets. This is a picture taken from the Plots & Data part of the Physical Sciences Laboratory.

우기, 두가지 기간이 존재하며, 주기 함수 $a(t)$ 가 양의 상태를 유지하다가 음의 상태로 바뀔 때 최대 분산을 가지는 기억 효과에 의해 $N(t)$ 가 최대가 되어 극단적인 기상현상의 발현이 증가하는 공통점을 발견하였다. 또한, 강에 작용하는 장기적 강제력을 제거하여 오직 강의 지배적 계절 주기성과 날씨의 효과만을 포함한 상황을 그래프로 나타낼 수 있었다. 이를 2 m 대기 온도의 장기적 강제력과 각 지점별로 상관(Point-correlation)한 결과, PDO패턴이 공통적으로 나타남을 알 수 있다. PDO 패턴이 강 유량과 깊은 관계성이 있음을 시사하며, 추후 연구에서 밝혀야 할 중요한 과제이다.

향후 연구에서는 지구에 존재하는 모든 강의 장기적 강제력 $d(\tau)$ 와 2 m 대기 온도에 존재하는 $f(\tau)$ 를 점 상관하여 특정 지역의 강이 어떤 기후요소에 영향을 받는지 분류할 수 있을 것이다. 이러한 기후 요소로 분류된 강들의 과정을 이해함으로써, 해당 강에서 발생하는 홍수와 가뭄 같은 극단적인 기상현상을 기후 인덱스를 통해 예보할 수 있는 가능성을 기대할 수 있을 것이다.

감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2023-00281017). 이 논문은 2024년도 정부(교육부)의 재원으로 한국연구재단의 G-램프(LAMP) 사업 지원을 받아 수행된 연구임(No. RS-2023-00301702).

REFERENCES

- Ahmed, R., Prowse, T., Dibike, Y., Bonsal, B. and O'Neil, H., 2020, Recent trends in freshwater influx to the Arctic Ocean from four major Arctic-draining rivers. *Water*, 12, 1189.
- Ashmore, P. and Sauks, E., 2006, Prediction of discharge from water surface width in a braided river with implications for at-a-station hydraulic geometry. *Water Resources Research*, 42.
- Bobrovitskaya, N.N., Zubkova, C. and Meade, R.H., 1996, Discharges and yields of suspended sediment in the Ob and Yenisey Rivers of Siberia. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, 236, 115-124.
- Bonakdari, H., Binns, A.D. and Gharabaghi, B., 2020, A comparative study of linear stochastic with nonlinear daily river discharge forecast models. *Water Resources Management*, 34, 3689-3708.
- Claps, P., Rossi, F. and Vitale, C., 1993, Conceptual-stochastic modeling of seasonal runoff using autoregressive moving average models and different scales of aggregation. *Water Resources Research*, 29, 2545-2559.
- Hasanah, Y., Herlina, M. and Zaikarina, H., 2013, Flood prediction using transfer function model of rainfall and water discharge approach in Katulampa dam. *Procedia Environmental Sciences*, 17, 317-326.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. and Thépaut, J.N., 2020, The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999-2049.
- Hirabayashi, Y., Kanae, S., Emori, S., Oki, T. and Kimoto, M., 2008, Global projections of changing risks of floods and droughts in a changing climate. *Hydrological Sciences Journal*, 53, 754-772.
- Liu, Y., Wu, J., Liu, Y., Hu, B.X., Hao, Y., Huo, X. and Wang, Z.L., 2015, Analyzing effects of climate change on streamflow in a glacier mountain catchment using an ARMA model. *Quaternary International*, 358, 137-145.
- Meade, R., Bobrovitskaya, N. and Babkin, V., 2000, Suspended-sediment and fresh-water discharges in the Ob and Yenisey rivers, 1960-1988. *International Journal of Earth Sciences*, 89, 461-469.
- Modarres, R., 2007, Streamflow drought time series forecasting. *Stochastic Environmental Research and Risk Assessment*, 21, 223-233.
- Moon, W. and Hannachi, A., 2021, River Nile discharge, the Pacific Ocean and world climate-a seasonal synchronization perspective. *Tellus A: Dynamic Meteorology and Oceanography*, 73, 1-12.
- Moon, W. and Wettlaufer, J.S., 2013, A stochastic perturbation theory for non-autonomous systems. *Journal of Mathematical Physics*, 54, <https://doi.org/10.1063/1.4848776>.
- Moon, W. and Wettlaufer, J.S., 2017, A unified nonlinear stochastic time series analysis for climate science. *Scientific reports*, 7, 44228.
- Neal, E.G., Walter, M.T. and Coffeen, C., 2002, Linking the pacific decadal oscillation to seasonal stream discharge patterns in Southeast Alaska. *Journal of Hydrology*, 263, 188-197.
- Nigam, R., Nigam, S.M.S.K. and Mittal, S.K., 2014, The river runoff forecast based on the modeling of time series. *Russian Meteorology and Hydrology*, 39, 750-761.
- Otache, M.Y., Mohammed, A.S. and Ahaneku, I.E., 2011, ARMA modelling of Benue River flow dynamics: comparative study of PAR model. *Open Journal of Modern Hydrology*, 1, 1-19.
- Overeem, I. and Syvitski, J.P., 2010, Shifting discharge peaks in Arctic rivers, 1977-2007. *Geografiska Annaler. Series A, Physical Geography*, 92, 285-296.
- Yilmaz, M., 2023, Accuracy assessment of temperature trends from ERA5 and ERA5-Land. *Science of the Total Environment*, 856, 159182.
- Zounemat-Kermani, M., 2016, Investigating chaos and nonlinear forecasting in short term and mid-term river discharge. *Water Resources Management*, 30, 1851-1865.