

비선형 자기회귀모형을 이용한 남방진동지수 시계열 분석

Nonlinear Autoregressive Modeling of Southern Oscillation Index

권 현 한* / 문 영 일**

Kwon, Hyun-Han / Moon, Young-Il

Abstract

We have presented a nonparametric stochastic approach for the SOI(Southern Oscillation Index) series that used nonlinear methodology called Nonlinear AutoRegressive(NAR) based on conditional kernel density function and CAFPE(Corrected Asymptotic Final Prediction Error) lag selection. The fitted linear AR model represents heteroscedasticity, and besides, a BDS(Brock - Dechert - Sheinkman) statistics is rejected. Hence, we applied NAR model to the SOI series. We can identify the lags 1, 2 and 4 are appropriate one, and estimated conditional mean function. There is no autocorrelation of residuals in the Portmanteau Test. However, the null hypothesis of normality and no heteroscedasticity is rejected in the Jarque-Bera Test and ARCH-LM Test, respectively. Moreover, the lag selection for conditional standard deviation function with CAFPE provides lags 3, 8 and 9. As the results of conditional standard deviation analysis, all I.I.D assumptions of the residuals are accepted. Particularly, the BDS statistics is accepted at the 95% and 99% significance level. Finally, we split the SOI set into a sample for estimating the model and a sample for out-of-sample prediction, that is, we conduct the one-step ahead forecasts for the last 97 values (15%). The NAR model shows a MSEF of 0.5464 that is 7% lower than those of the linear model. Hence, the relevance of the NAR model may be proved in these results, and the nonparametric NAR model is encouraging rather than a linear one to reflect the nonlinearity of SOI series.

keywords : Nonlinear Autoregressive, SOI series, Residual Test, Lag Time

요 지

본 연구에서는 조건부 핵밀도함수와 CAFPE(Corrected Asymptotic Final Prediction Error) 차수결정 방법에 근거한 비매개변수적 비선형 자기회귀 (Nonlinear AutoRegressive, NAR) 모형을 소개하고 이를 SOI(Southern Oscillation Index)에 적용하였다. SOI 자료에 대해서 선형 AR 모형을 적용하였으나 잔차에 대한 검정결과 이분산성(heteroscedasticity)을 나타내었다. 또한 BDS(Brock - Dechert - Sheinkman) 검정에서 비선형성이 존재함을 확인하였다. 따라서 NAR 모형에 SOI 자료를 적용시켰다. CAFPE를 이용하여 가장 적합한 모형으로 지체 1, 2와 4가 선택되었으며 조건부 평균함수를 추정하여 SOI 자료를 모의한 결과 잔차에 대해서 정규성과 이분산성 가정이 Jarque-Bera 검정과 ARCH-LM 검정에서 각각 기각되었으며 또한 조건부 표준편차함수의 최적 차수로 3, 8과 9가 CAFPE를 통해 선택되었다. 조건부 평균함수와 표준편차함수를 모두 고려한 모형에 대한 잔차 검정 결과 잔차의 I.I.D 가정을 만족하였으며 특히, BDS 검정에서 신뢰구간 95%와 99%에서 모두 만족한 결과를 나타내었다. 마지막으로 전체의 15%에 해당하는 SOI 자료에 대해서 One-Step 예측을 수행하였으며 선형 모형에 비해 평균제곱예측오차

* 정회원 · Research Associate Columbia University, NY, USA (e-mail: hk2273@columbia.edu)

** 정회원 · 책임저자, 서울시립대학교 공과대학 토목공학과 부교수 (e-mail: ymoon@uos.ac.kr)

가 7% 적게 나타났다. 따라서, NAR 모형은 여타의 매개변수적 방법과 달리 모형 선택에 있어 자유로우며 비선형성을 고려할 수 있는 모형으로서 SOI 자료와 같은 비선형 자료를 위한 모의방법으로 선형 모형에 비해 많은 장점을 가지고 있다.

핵심용어 : 비선형 자기회귀모형, 남방진동지수, 잔차 검정, 지체시간

1. 서론

시계열 분석은 관측 자료의 스펙트럼분석 및 가설검정 등을 포함하는 다양한 분석을 통하여 첫 번째로 관측된 연속자료에 의해서 표현되는 현상의 특징을 규명하고 두 번째로 규명된 구조를 바탕으로 미래의 사상을 예측하는 목적으로 구분할 수 있다. 두 가지 목적 모두 시계열의 구조를 분석하는 것이 선결되어야 하며 이를 통하여 시계열을 해석하고 모의가 가능하다. 이러한 관점에서 자기회귀(AutoRegressive, AR) 모형은 수문시계열의 많은 실제적인 문제에서 해석에 대한 편리성과 용이함으로 기본적으로 일반적인 방법으로 인식되고 있는 전통적인 방법이라 할 수 있다.

자기회귀과정은 동일 시계열상에서 하나 또는 이전 시간의 관측치와 상관성이 존재한다는 가정을 기본으로 하고 있다. 그러나 만약 우리가 시계열의 비선형 거동에 관심을 갖는다면 지금까지 이용되어온 선형 AR 모형은 시계열의 동적 거동을 반영하기에 무리가 따른다. 수문시계열 분석에서 비선형 이론에 근거한 연구는 최근에 크게 증가하고 있다. 이러한 연구의 일환으로 크게 수문변량에서의 비선형성 검출에 초점을 맞춘 연구와(Hense, 1987; Rodriguez-Iturbe et al., 1989; Jayawardena and Lai, 1994) 수문시계열에 예측 가능성을 평가한 연구가 주를 이루고 있다(Lall et al., 1996; Jayawardena and Xu, 2002; Elshorbagy et al., 2002).

이전 연구에서 Chu and Katz(1985)는 1935년부터 1983년까지의 SOI를 모의하기 위해서 ARMA (Auto-Regressive Moving Average) 모형을 이용하였으며 Trenberth and Hoar(1996)는 ARMA(3,1)를 이용하여 1882년부터 1981년까지의 자료를 모의한 바 있다. 최근에 Ahn and Kim (2005)은 Brock - Dechert - Sheinkman 잔차 검정(BDS, Brock et al., 1991; Brock et al., 1997)을 이용하여 SOI의 비선형성을 검증하고 ARCH (AutoRegressive Conditional heteroscedasticity) 비선형 매개변수 모형을 이용하여 모의를 실시하였다.

그러나, ARCH(Engel, 1982, 1983; Ahn and Kim, 2005) 모형과 같은 매개변수 모형은 적절한 차수의 결정 및 비선형 함수의 추정 등에 있어서 어려움이 존재

한다(Tschernig and Yang, 2000a,b). 이러한 문제점은 차수 결정에 있어서 여러 기준들을 고려해야 함으로서 나타나는 다중의 최적 차수 문제, 또한 ARCH 모형은 백색잡음에 따른 분산의 변동성을 고려하는 비선형 모형이기는 하나매개변수 모형으로서 기존 매개변수 모형이 갖는 분포형 선정의 문제, 매개변수 추정에 관한 문제 등의 단점을 가지고 있다.

이러한 문제점은 시계열에 내재된 특성으로부터 비선형 함수와 차수를 추정할 수 있는 비매개변수적 비선형 모형을 이용하여 보완될 수 있다. 이러한 관점에서 본 논문에서는 단변량 비선형 시계열의 비매개변수적 접근 방법인 비선형 AR(Nonlinear AutoRegressive, NAR) 모형을 제시하고자 한다. 비매개변수적 방법으로서 본 연구에서는 핵밀도함수를 이용하였으며 국내에서 주로 빈도해석, 댐 위험도 분석 및 비선형 시계열 예측 등에서 응용된 사례가 있으나(문영일과 이정규, 1995; 문영일과 권현한, 2004a,b; 권현한과 문영일, 2005; Kwon and Moon, 2006) 이를 비선형시계열 분석에 이용한 사례는 수문기상학 분야에서 극히 드문 실정이며 따라서 이에 대한 적용성 및 타당성을 검증하는데 본 연구의 목적이 있다.

본 연구를 통해 얻고자 하는 목적과 연구 진행과정을 요약하여 정리하면 다음과 같다. 1) 비매개변수적 NAR 모형의 기본 이론을 소개한다. 2) 모형의 지체시간과 광역폭(bandwidth)의 추정방법을 제시한다. 3) SOI 자료를 대상으로 제시된 방법론을 근거로 최적의 지체시간과 광역폭을 추정하여 핵밀도함수(kernel function)를 이용한 조건부 함수를 추정한다. 4) 최종적으로 다양한 잔차 검정 방법을 이용하여 NAR 모형의 비선형 자료에 대한 적합성을 평가한다.

2. 이론적 배경

본 연구에서 제시한 방법론의 적합성을 평가하기 위해서 합성자료(synthetic data)를 중심으로 논문을 전개하고자 한다. 따라서, 비선형 특성을 갖는 Exponential AR(Tong, 1990)를 사용하여 자료크기 $n=150$ 을 아래 Eq. (1)을 이용하여 모의하였다.

$$x_n = \sum_{i=1}^p [a_i + b_i \exp(-ax_{n-i}^2)] x_{n-i} + \varepsilon_n \quad (1)$$

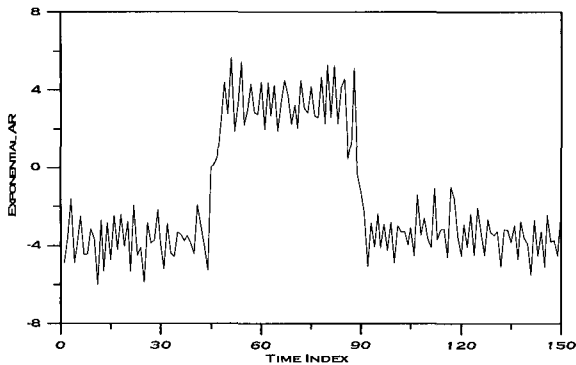


Fig. 1. Time series generation using Exponential AR(1)

여기서, p 는 모형의 차수를 나타내며 ε_n 는 백색잡음을 a, b, α 는 매개변수를 의미한다. 본 연구에 사용된 Exponential AR(1) 모형은 $x_n = (0.3 + 2.2 \exp(-0.1x_{n-1}^2))x_{n-1} + \varepsilon_n$ 이며, Fig. 1은 모의 발생된 $n=150$ 의 자료계열을 나타낸다.

자기회귀 모형은 최근까지도 수문학적 연구에 많이 이용되고 있으며 동일 시계열상에 하나 또는 그 이상의 이전 관측치와 상관성을 갖는 기본적인 가정에 근거를 두고 있다. 다시 말해서, 수문시계열의 특정 변량의 가장 적합한 예측항(predictor)으로서 동일 시계열상의 이전 관측치를 이용하는 것이다. 일반적인 단변량의 AR 모형은 다음과 같이 나타낼 수 있다.

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2)$$

여기서, Y_t 는 시계열을, ε_t 는 백색잡음을 의미하며 δ 는 다음과 같이 나타낼 수 있다(μ 는 평균을 나타낸다).

$$\delta = \left(1 - \sum_{i=1}^p \phi_i \right) \mu \quad (3)$$

즉 AR 과정은 시계열의 이전 값들에 대한 현재 관측치의 선형 회귀분석을 나타내며, 여기서 p 를 AR 모형의 지체시간 또는 차수라고 한다.

본 연구에서는 단변량 시계열의 비선형 특성을 반영할 수 있는 비매개변수적 방법을 소개하고자 한다. 비선형 시계열분석에서, 비매개변수적 해석 접근 방법은 여타의 매개변수적 방법과 달리 함수의 어떠한 가정을

필요로 하지 않으며 모형 선택에 있어 매개변수적 방법에 비해 용이하다 하겠다. 일반적으로 단변량의 추계학적인 과정 y_t 를 위한 비선형 AR 모형은 다음 Eq. (4)와 같이 나타낼 수 있다.

$$y_t = f_\mu(x_t) + f_\sigma(x_t)\xi_t \quad (4)$$

여기서 $x_t = (y_{t-1}, y_{t-2}, \dots, y_{t-11})'$ 는 선택된 지체 1의 (1×1) Vector를 나타내고, ξ_t 는 백색잡음을 나타낸다. $f_\mu(\cdot)$ 와 $f_\sigma(\cdot)$ 는 지체 i_1, \dots, i_1 를 갖는 조건부 평균 함수와 조건부 표준편차 함수를 의미한다. 비선형 AR 과정에서 조건부 함수인 $f_\mu(\cdot)$ 와 $f_\sigma(\cdot)$ 의 추정이 필요하며 또한 적절한 지체시간을 선택해야 하며 이를 추정하기 위한 방법론은 다음과 같다.

2.1 국부선형추정자

시계열 분석에서 이전시점의 자료를 독립변수로 현재의 자료를 종속변수로 하는 회귀방정식을 고려할 수 있는 있으며 이는 앞서 언급했듯이 자기회귀 모형의 기본 가정 사항이다. 여기서 이들의 상관관계를 선형 또는 비선형의 회귀방정식으로 표현할 수 있다. 이러한 관점에서 자료간의 조건부 평균함수를 추정하기 위해서 국부다항식(local polynomial) 방법, 국부선형(local linear) 방법 등의 비선형 방법을 이용할 수 있으며 본 논문에서는 국부선형 방법을 사용하였으며 이를 중심으로 해석 과정을 전개하였다.

조건부 평균함수 $f_\mu(x)$ 의 국부적인 추정값은 각 (1×1) Vector에 대해서 $f_\mu(\cdot)$ 를 개별적으로 추정함으로써 가능하다. 국부적인 추정값을 유도하기 위해서는 $f_\mu(x)$ 의 x 를 중심으로 한 일차 Taylor 전개로 고려될 수 있다.

$$f_\mu(x_t) = f_\mu(x) + \frac{\partial f_\mu(x)}{\partial x} (x_t - x) + O(x_t, x) \quad (5)$$

여기서 $O(x_t, x)$ 는 나머지 항을 나타내며 Eqs. (4) and (5)를 이용하여 다음 Eq. (6)과 같이 표현할 수 있다.

$$y_t = f_\mu(x) + \frac{\partial f_\mu(x)}{\partial x} (x_t - x) + O(x_t, x) + \varepsilon_t \quad (6)$$

일반적으로 조건부 평균함수와 미분계수를 추정하기

위한 매개변수들은 OLS(Ordinary Least Square) 회귀 분석 방법이 적용되어 왔다. 만약, 조건부 평균함수가 선형이 아니라면 나머지 항 $O(x_t, x)$ 은 영(zero)이 아닌 값을 나타내게 된다. 일반적인 OLS 회귀분석은 모든 나머지 항 $O(x_t, x)$ 에 따른 편의 때문에 편이 된 추정치가 불가피하다. 이러한 문제점은 x 근처에 관측치, 즉 이웃 항을 이용함으로써 보완될 수 있다. 비선형 AR 과정에서 핵함수 $K(u)$ 는 이웃 항에 대한 가중함수로 적용될 수 있으며 이웃 항의 크기를 조절하기 위해서 광역폭(bandwidth) h 를 이용하게 된다. 따라서 x 근처에 관측치를 이용하기 때문에 일반적으로 이를 국부 추정자(local estimator)라고 언급하기도 한다. 핵함수는 대칭함수, 양의함수 그리고 $\int k(u)du = 1$ 의 특성을 갖고 있으며 다음과 같이 나타낼 수 있다.

$$K_h(x_t - x) = \prod_{i=1}^I \frac{1}{h^i} K\left(\frac{x_{ti} - x_i}{h}\right) \quad (7)$$

여기서 x_{ti} 는 x_t 의 i 번째 관측치를 의미하며 Eq. (7)은 광역폭 h , 분산 $\sigma_k^2 = \int u^2 k(u)du$ 와 상수 $\|K\|_2^2 = \int k(u)^2 du$ 을 갖는 적(product)으로 나타낼 수 있다. 조건부 평균함수 $f_u(\cdot)$ 의 x 에 대한 추정을 위해서 다음과 같은 행렬을 고려할 수 있다.

$$e = (1, 0_{1 \times 1})', Z(x) = \begin{pmatrix} 1 & \dots & 1 \\ x_{i+1} - x & \dots & x_T - x \end{pmatrix} \quad (8)$$

$$W(x, h) = \text{diag}\{K_h(x_t - x)/T\}_{t=i+1}^T, \quad (9)$$

$$y = (y_{i+1} \dots y_T)'$$

국부선형 추정값을 계산하기 위해서, Plug-In 광역폭 h 를 이용하며 비선형 AR 모형의 조건부 평균함수 $f_u(x)$ 을 위한 국부선형 함수는 다음과 같은 최소자승 추정값으로 나타낼 수 있다.

$$\hat{f}_\mu(x, h) = \quad (10)$$

$$e' \{Z'(x) W(x, h) Z(x)^{-1}\} Z'(x) W(x, h) y$$

국부선형함수를 이용한 회귀분석과 선형회귀분석의 결과를 비교하기 위해서 임의의 자료를 구성하여 회귀 분석 결과를 도시하여 나타내었다. Fig. 2에서 선형회귀 분석과 비선형회귀분석 모두 전반부의 선형부분에서는

결과의 차이를 확인할 수 없지만 후반부의 비선형 구간에서는 국부선형 방법이 자료의 거동을 보다 효과적으로 표현하고 있다. 여기서 국부선형 함수를 추정하기 위해서 광역폭 0.5와 Gaussian 핵함수를 사용하였다.

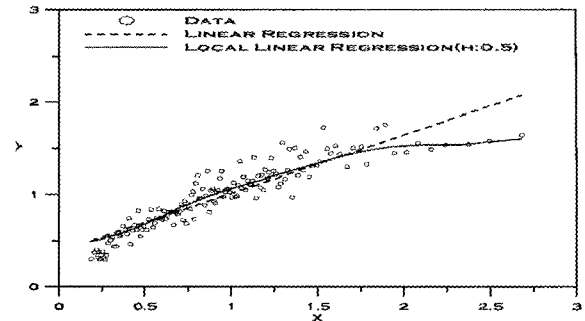


Fig. 2. Comparison between linear regression and local linear regression for the synthetic time series

2.2 지체시간과 광역폭 선택

Tschernig and Yang(2000b)은 지체시간과 광역폭을 비매개변수적으로 추정하기 위한 방법론을 제시하였다. 최대 지체 L 을 추정하기 앞서 가능한 지체시간의 조합의 수를 결정하게 되면 다음 Eq. (11)과 같이 나타낼 수 있다.

$$x_{t,L} = (y_{t-1}, y_{t-2}, \dots, y_{t-L})' \quad (11)$$

Eq. (11)에서 최대 지체 L 을 모두 고려하는 것은 해석상 큰 부담을 야기하므로 최적의 지체시간 결정의 주된 목적은 여분의 지체 Vector를 감소시키는 것이라 할 수 있다. 이러한 최적의 지체시간을 추정하기 위한 방법으로 비선형 AR 과정에서는 평균통합오차에 근간을 둔 Asymptotic Final Prediction Error(AFPE) 또는 이를 수정한 CAFPE(Corrected AFPE)가 이용된다.

지체시간 결정을 위해서 가중함수 $w(\cdot)$ 의 추정이 필요하며, 여기서 무작위로 선택된 지체시간 Vector와 선택된 최적의 지체시간과 구별하기 위해서 기호 "*"를 이용하여 구별하였다. 주어진 광역폭 h 와 지체 Vector i_t 에 대해서 FPE는 다음과 같이 나타낼 수 있다.

$$\begin{aligned} FPE(h, i_1^*, \dots, i_1^*) &= E\left((\tilde{y}_t - \hat{f}_\mu(\tilde{x}_t^*, h))^2 w(\tilde{x}_t, L)\right) \\ &= \int \left[\int \left((\tilde{y}_t - \hat{f}_\mu(\tilde{x}_t^*, h))^2 w(x_L) f(\tilde{y}, \tilde{x}_L) d\tilde{y} d\tilde{x}_L \right) \right. \\ &\quad \left. \times f(y_1, \dots, y_T) dy_1 \dots dy_T \right] \quad (12) \end{aligned}$$

Tschernig and Yang(2000b)는 다음 Eq. (13)과 같은 AFPE를 제안하였다.

$$AFPE(h, i_1, \dots, i_l) = A + b(h)B + c(h)C \quad (13)$$

여기서 A 는 통합분산(integrated variance), $b(h)B$ 는 추정치의 통합분산(integrated variance of estimation) 그리고 $c(h)C$ 는 추정치의 통합제곱편의(integrated squared bias)를 나타낸다.

$$A = \int f_\sigma^2(x)w(x_L)f(x_L)dx_L \quad (14)$$

$$= E[f_\sigma^2(x_t)w(x_t, L)]$$

$$B = \int f_\sigma^2(x)w(x_L)f(x_L)/f(x)dx_L \quad (15)$$

$$= E\left[f_\sigma^2(x_t)\frac{w(x_t, L)}{f(x_t)}\right]$$

$$C = \int \left(\text{tr} \left\{ \frac{\partial^2 f_\mu(x)}{\partial x \partial x'} \right\} \right)^2 w(x_L)f(x_L)dx_L \quad (16)$$

$$= E\left[\left(\text{tr} \left\{ \frac{\partial^2 f_\mu(x_t)}{\partial x \partial x'} \right\} \right)^2 w(x_t, L) \right]$$

$$b(h) = \|K\|_2^{2l}(T-i_1)^{-1}h^{-1}, c(h) = \sigma_k^4 h^4 / 4 \quad (17)$$

여기서 가장 작은 AFPE(.)을 나타내는 지체시간이 최적 차수로 결정된다.

2.2.1. 광역폭 추정

광역폭(bandwidth) h 에 대해서 AFPE를 최소화함으로써 점근적인 최적의 광역폭을 추정할 수 있으며 이는 $b(h)B$ 와 $c(h)C$ 사이의 분산-편의를 추정함으로써 가능하다.

$$h_{opt} = \left\{ \frac{1 \|K\|_2^{2l} B}{(T-i_1)C\sigma_k^4} \right\}^{1/(1+4)} \quad (18)$$

B 의 기대값은 국부추정치로부터 가중제곱오차(weighted squared errors)를 평균함으로써 가능하다.

$$\hat{B}(\hat{h}_B) = \frac{1}{T-i_{l,t=i_1+1}} \sum \{y_t - \hat{f}_\mu(x_t, \hat{h}_B)\}^2 \frac{w(x_t, L)}{\hat{f}(x_t, \hat{h}_B)} \quad (19)$$

여기서 $\hat{f}(\cdot, \cdot)$ 는 핵밀도함수 $f(x)$ 의 Gaussian Kernel Estimator를 나타내고 \hat{h}_B 는 Rule-of-Thumb 광역폭(Silverman, 1986)을 의미한다.

$$\hat{h}_B = \hat{\sigma} \left(\frac{4}{1+2} \right)^{1/(1+4)} T^{-1/(1+4)} \quad (20)$$

여기서 $\hat{\sigma} = \left(\prod_{j=1}^l \sqrt{\text{var}(x_{t,j})} \right)^{1/l}$ 은 Regressor의 표준편차의 기하평균을 나타낸다. 상수 C 는 다음 Eq. (21)로 표현될 수 있다.

$$\hat{C}(h_c) = \frac{1}{T-i_{l,t=i_1+1}} \sum_{j=1}^l \left[\sum_{j=1}^l \hat{\mu}^{jj}(x_t, h_c) \right]^2 w(x_t, L) \quad (21)$$

여기서 $\mu^{jj}(\cdot, \cdot)$ 는 함수 f_μ 의 2차 미분계수를 나타내며 국부 2차 추정자를 이용하여 계산할 수 있다.

$$\{\hat{c}_0, \hat{c}_{11}, \dots, \hat{c}_{1l}, \hat{c}_{21}, \dots, \hat{c}_{2l}\}$$

$$= \text{argmin}_{\{\hat{c}_0, \hat{c}_{11}, \dots, \hat{c}_{1l}, \hat{c}_{21}, \dots, \hat{c}_{2l}\}}$$

$$\sum_{t=i_1+1}^T \left\{ \begin{array}{l} y_t - c_0 \\ -c_{11}(x_{t1} - x_1) - \dots \\ -c_{1l}(x_{t1} - x_1) \\ -c_{21}(x_{t1}, x_1)^2 - \dots \\ -c_{2l}(x_{t1}, x_1)^2 \end{array} \right\}^2 K_h(x_t - x) \quad (22)$$

2차 미분계수의 추정치는 $\hat{\mu}^{jj}(x, h) = 2\hat{c}_{2j}, j=1, \dots, l$ 에 의해서 계산되고 Rule-of-Thumb 광역폭 \hat{h}_c 는 다음과 같이 나타낼 수 있다(Yang and Tschernig, 1999).

$$\hat{h}_c = 2\hat{\sigma} \left(\frac{4}{1+4} \right)^{1/(1+6)} T^{-1/(1+6)} \quad (23)$$

추정된 B 와 C 에 근거한 Plug-In 광역폭은 Eq. (24)로 주어진다.

$$\hat{h}_{opt} = \left\{ \frac{1 \|K\|_2^{2l} \hat{B}(\hat{h}_B)}{(T-i_1)\hat{C}(\hat{h}_c)\sigma_k^4} \right\}^{1/(1+4)} \quad (24)$$

2.2.2. 지체시간선정

B 와 C 가 이미 광역폭 선택에서 추정이 되었기 때문

에 지체시간을 결정하기 위해서 통합분산만 추정하면 되며 Eq. (25)로 나타낼 수 있다.

$$\hat{A}(h) = \frac{1}{T - i_1} \sum_{t=i_1+1}^T \{y_t - \check{\mu}(x_t, h)\}^2 w(x_{t,L}) \quad (25)$$

$\hat{A}(h)$ 의 1차 편의를 제거하기 위해서 점근적인 편의를 감하게 되면 다음과 같은 점근적으로 불편의 된 통합오차를 추정할 수 있다.

$$\hat{A}(h_{opt}) = \hat{A}(\hat{h}_n) - \left\{ b(\hat{h}_n) - \frac{2K(0)^l}{(T - i_1)\hat{h}_n^l} \right\} \hat{B}(\hat{h}_B) - c(\hat{h}_n)\hat{C}(\hat{h}_C) \quad (26)$$

따라서, 지금까지 추정된 모든 추정치를 동시에 고려하면 Eq. (12)에 의해서 주어진 AFPE를 추정할 수 있으며 Eq. (27)로 주어진다.

$$AFPE = \hat{A}(\hat{h}_{opt}) + \frac{2K(0)^l}{(T - i_1)\hat{h}_{opt}^l} \hat{B}(\hat{h}_B) \quad (27)$$

그러나 AFPE는 지체시간 결정시에 과다추정의 문제점이 발생할 수 있으며 이러한 문제점을 개선하기 위해서 수정된 CAFPE가 제안되었다(Tschernig and Yang, 2000b).

$$CAFPE = AFPE \left\{ 1 + l(T - i_1)^{-4/(1+l)} \right\} \quad (28)$$

2.2.3. 모의자료에 대한 지체시간 결정 및 조건부 평균함수 추정

본 연구에서 제시한 비선형 시계열의 지체시간 결정 방법인 CAFPE와 기존의 선형 모형에 주로 사용되어온 지체시간 결정방법을 모의 발생시킨 Exponential AR(1) 자료를 대상으로 비교하여 분석하였다. 여기서 선형 모형을 위한 지체시간 선정은 여러 분야에서 일반적으로 사용되는 4가지 방법으로 검토하였다. 즉, AIC (Akaike, 1973), HQ (Hannan and Quinn, 1979), SC (Schwarz, 1978; Rissanen, 1978), 마지막으로 FPE (Akaike, 1969)를 이용하였다. Table 1은 지체시간 결정 결과를 나타낸다.

지체시간을 결정하기 위해서 최대 지체 10과 가능한

조합의 수 10개를 고려하여 각 모형을 검토하였으며 조합의 수의 따른 최적의 지체 Vector와 추정값을 표 1에 나타내었다. 모의 발생된 자료계열은 차수 1을 갖는 비선형 AR 모형이기 때문에 최적의 지체시간으로 1이 선택되어야 한다.

그러나 모의 발생된 자료가 1차의 지체시간을 갖는 시계열임에도 불구하고 각 방법들은 다른 차수의 모형을 최적의 모형으로 추정하고 있다. 즉 AIC, FPE와 HQ는 AR(p₁, p₂, p₃, p₇)을 최적의 모형으로 제시하고 있으며, 반면 SC는 AR(p₁, p₂, p₃, p₇)를 최적의 모형으로 추정하고 있다. 따라서 비선형 특성을 갖는 자료의 구조를 효과적으로 반영하지 못한 결과로 사료되며 본 연구에서 제시한 CAFPE를 이용하여 같은 조건에서 시계열 자료에 지체 Vector를 다시 추정해 보았다. Table 2는 CAFPE를 통해 추정된 모형의 최적 지체시간과 광역폭을 나타낸다.

Table 2에서 최적의 지체시간으로 모의된 자료와 동일한 지체 1이 최적의 모형으로 추정되었다. 따라서, 비선형 특성을 갖는 시계열의 차수결정에 대한 타당성을 확인할 수 있으며 이를 토대로 모의된 자료의 조건부 평균함수를 추정하여 Fig. 3에 나타내었다.

그림에서 두꺼운 실선은 모의된 자료에 실제 함수를 나타내며 두꺼운 점선은 비선형 AR 모형을 통해 추정된 조건부 평균함수를 의미한다. 즉 1차의 지체시간을 갖는 자료와 현재시점 자료의 비선형 관계를 연결해 주는 함수로서 비선형 AR 모형을 통해 추정된 조건부 함수가 실제 함수의 거동을 거의 동일하게 추정하고 있음을 확인할 수 있다.

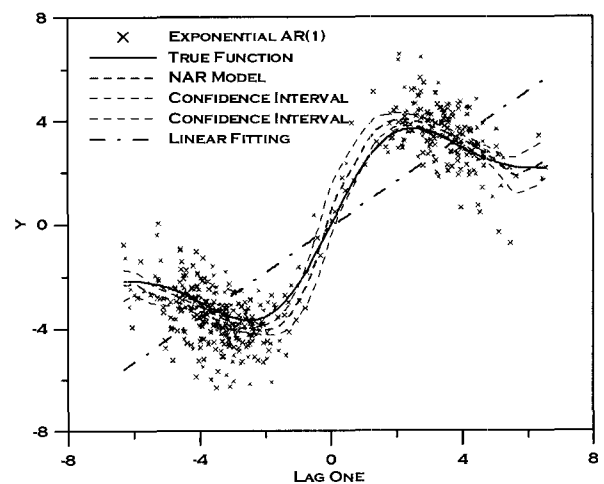


Fig. 3. Conditional mean function of the synthetic exponential AR(1) time series with the associated uncertainty bound

$$\varepsilon_t = y_t - f_\mu(x_t), \quad t = i_1 + 1, \dots, T \quad (29)$$

Eq. (2)에 근거해서 Eq. (4)의 비선형 AR방정식은 $\varepsilon_t = f_\sigma(x_t)\xi_t$ 로 대체할 수 있으며 다음 Eq. (30)으로 변형할 수 있다.

$$\varepsilon_t^2 = f_t^2(x_{\sigma,t}) - f_t^2(x_{\sigma,t})(\xi_t^2 - 1), \quad (30)$$

$$t = i_{\sigma,l} + 1, \dots, T$$

여기서 $i_{\sigma,l} = \max(i_l, i_{l'})$ 를 의미하고 i_{σ} 는 x_σ 의 가장 큰 지체를 나타낸다. 또한 Eq. (10)을 조정하여 Eq. (31)로 나타낼 수 있다.

$$\hat{\varepsilon}_t(h) = y_t - \hat{f}_\mu(x_t, h) \quad (31)$$

만약, $\hat{f}_\mu(x_t, h)$ 가 Eq. (10)처럼 국부선형방법을 고려하면 조건부 분산 $f_\sigma^2(x_t)$ 의 국부선형 추정자는 Eq. (10)의 y_t 를 $\hat{\varepsilon}^2$ 로 대체함으로써 표현될 수 있다.

$$\hat{f}_\sigma^2(x_\sigma, h_\sigma) =$$

$$e' \left\{ Z'(x_\sigma) W(x_\sigma, h_\sigma) Z(x_\sigma) \right\}^{-1} Z'(x_\sigma) W(x_\sigma, h_\sigma) \hat{\varepsilon}^2 \quad (32)$$

여기서, $\hat{\varepsilon}^2 = (\hat{\varepsilon}_{i_1+1}^2, \dots, \hat{\varepsilon}_T^2)'$ 을 의미하며 Nadaraya-Watson 추정자는 다음 Eq. (33)으로 변형될 수 있다.

$$\hat{\sigma}_{NW}^2(x_\sigma, h_\sigma) = \frac{\sum_{t=i_{\sigma,l}+1}^T K_{h_\sigma}(x_{\sigma,t} - x_\sigma) \hat{\varepsilon}_t^2}{\sum_{t=i_{\sigma,l}+1}^T K_{h_\sigma}(x_{\sigma,t} - x_\sigma)} \quad (33)$$

조건부 표준편차함수의 광역폭과 지체시간은 변수와 Eq. (30)을 수정함으로써 가능하다.

4. 잔차 검정

일반적으로, 추정된 모형이 시계열의 내재된 특성을 잘 반영하는지에 대한 검토가 필요하며 단변량 시계열에 대해서 특히 잔차(residual) 검정이 요구된다. 즉 잔차가 IID(independent and identically distributed) 가정을 만족하는지 검정이 필요하며 이를 위해서 오차의 자기회귀과정, 정규성과 ARCH 효과 등의 검토를 하게

되며 이를 만족하지 못할 경우 모형을 다시 구성해야 한다. 이들 잔차 검정은 크게 시각적인 검토와 통계적인 가설검정으로 구분할 수 있다. 첫 번째로 모형의 오류를 검토하는 가장 쉽고 전통적인 방법으로 자기상관함수(AutoCorrelation Function, ACF)와 편자기상관함수(Partial AutoCorrelation Function, PACF)를 이용할 수 있으며 두 번째로 Portmanteau Test와 같은 가설검정 방법이 있다. 본 연구에서는 두 가지 방법을 모두 고려하였으며 가설검정방법으로는 오차의 자기상관성, 정규성, 이분산성과 잔차의 비선형성을 검토하였다. 이들 방법론에 대해서 간단히 정리하여 나타내었다.

4.1 Portmanteau 검정

잔차의 자기상관성을 검토하는 가장 전통적이고 일반적인 방법으로 Portmanteau Test는 다음과 같은 가설검정을 통해 이루어진다.

$$H_0 : \rho_{r,1} = \dots = \rho_{r,l} = 0$$

$$\text{Versus } H_1 : \rho_{r,l} \neq 0 \quad l = 1, \dots, n_r$$

여기서 $\rho_{r,l}$ 은 잔차 $(u_{r,t}, u_{r,t-1})$ 의 자기상관계수를 나타낸다. 잔차에 대한 Portmanteau 검정과 수정된 Portmanteau 검정(Ljung and Box, 1978)는 다음과 같다.

$$Q_{n_r} = T \sum_{l=1}^{n_r} \hat{\rho}_{r,l}^2 \quad (34)$$

$$Q_{n_r} = T^2 \sum_{l=1}^{n_r} \frac{1}{T-1} \hat{\rho}_{r,l}^2 \quad (35)$$

여기서, T는 자료의 수, $\rho_{r,l}$ 은 지체 $l = n_r$ 에서의 자기상관계수를 의미한다. 귀무가설이 유지되려면 통계치가 $\chi_{1-\alpha, n_r}^2$ 분포를 따라야 한다.

4.2 Jarque and Bera Test

잔차의 비정규성 검정은 3차 4차 모멘트 즉, 왜곡도와 첨도를 이용한다 (Lomnicki, 1961; Jarque and Bera, 1987). 즉 표준정규분포는 0의 왜곡도와 3의 첨도를 따르므로, 귀무가설과 대립가설을 다음과 나타낼 수 있다.

$$H_0 = E(u_{r,t}^s)^3 = 0 \quad E(u_{r,t}^s)^4 = 3$$

$$\text{Versus } H_0 : E(u_{r,t}^s)^3 \neq 0 \quad E(u_{r,t}^s)^4 \neq 3$$

$$JB = \frac{T}{6} \left[\frac{1}{T} \sum_{t=1}^T (\hat{u}_{r,t}^s)^3 \right]^2 + \frac{T}{24} \left[\frac{1}{T} \sum_{t=1}^T (\hat{u}_{r,t}^s)^4 - 3 \right]^2 \quad (36)$$

T는 자료의 수를 의미하며 $\hat{u}_{r,t}^s$, $\hat{u}_{r,t}^s$, 은 각각 자료의 왜곡도와 첨도를 나타낸다. 귀무가설이 유지되려면 통계치가 $\chi^2(2)$ 분포를 따라야 한다.

4.3 ARCH-LM(Lagrange multiplier) Test

이분산성(heteroscedasticity) 검정은 잔차를 ARCH(p) 모형에 적용함으로써 가능하다. 시간에 따라 분산(σ_t^2) 이 변하는 ARCH 모형은 계수 w 와 α 에서 결정되는 선형함수로 나타낼 수 있다.

$$\sigma_t^2 = w + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \quad (37)$$

여기서 이분산성 검정을 위한 가설은 다음과 같다.

$$H_0 : \alpha_1 = \dots \alpha_p = 0$$

$$\text{Versus } H_1 : \alpha_1 \neq 0 \text{ 또는 } \dots, \text{ 또는 } \alpha_p \neq 0$$

잔차의 정규성 가정하에서 회귀방정식(Eq. (37))의 결정계수 R^2 로 추정된다.

$$ARCH_{LM}(p) = T(R^2) \quad (38)$$

비이분산성의 귀무가설이 유지되려면(Engle, 1982), 통계치는 $\chi^2(p)$ 분포를 따라야 한다.

4.4 BDS(Brock - Dechert - Sheinkman) Test

BDS Test는 Brock et al.(1991), Brock et al.(1997) and Kim et al.(2003) 등에 의해서 적용된바 있으며 일반적으로 비선형 구조의 특성을 평가하기 위한 대표적인 검정방법이다. BDS 통계치는 다음 Eq. (39)와 같은 상관적분으로부터 추정된다.

$$BDS(m, \varepsilon) = T^{1/2} \{ C_m(\varepsilon) - [C_1(\varepsilon)]^m \} / b_m \quad (39)$$

여기서 $C_m(\varepsilon)$ 은 자료계열에 대한 상관적분을 나타내고

b_m 은 Embedding Dimension m에 따른 BDS 통계치의 표준편차를 나타낸다. IID의 귀무가설은 주어진 ε 과 $m > 1$ 에 대해서, $C_m(\varepsilon) - [C_1(\varepsilon)]^m$ 이 0일 때 기각할 수 없다.

5. 남방진동지수

본 연구에서는 1950년부터 2004년까지의 월 단위 남방진동지수(southern oscillation index, SOI)를 NAR 모형에 적용하였다. 일반적으로, SOI는 선형추계학적 모형 보다는 비선형추계학적 모형이 시계열의 특성을 표현하는데 유리한 것으로 알려지고 있다(Ahn and Kim, 2005). 따라서 선형 AR 모형과 본 연구에서 제시한 비매개변수적 NAR 모형을 비교하였다. 분석에 앞서 SOI 자료의 주기특성을 검토하기 위해서 Wavelet Transform을 실시하여 Fig. 4에 나타내었다. 여기서 강한 Wavelet Power가 2-8년 주위에서 나타나고 있음을 확인할 수 있으며 최근에 경년변화와 10년 이상의 저빈도의 특성이 강하게 나타나고 있다. 왼쪽 그림은 Wavelet Power 스펙트럼을 나타내며 오른쪽은 주기에 따른 평균 Wavelet Power를 나타낸다. 두꺼운 검은 선은 95% 신뢰구간을 나타낸다. 여기서 Morlet Wavelet and Zero Padding을 이용하였다(Torrence and Compo, 1998).

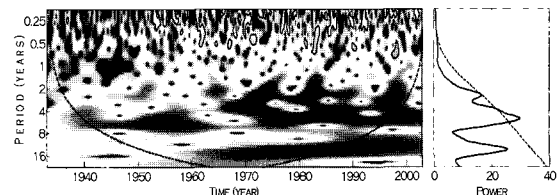


Fig. 4. Continuous Wavelet Transform analysis of Southern oscillation Index

6. 선형 AR 모형

이전 연구에서 Chu와 Katz(1985)는 SOI를 ARMA 모형에 적용하였으며 Akaike Information Criteria (Akaike, 1973)를 이용하여 ARMA(1,7:1) 모형을 선택하였다. 여기서 (1,7)의 의미는 지체 1과 지체 7만을 이용한 것을 의미하며 즉, 지체 2-6의 자료는 모형에서 제거된다. 최근에 비선형 ARCH 모형과 비교하기 위해서 Ahn과 Kim(2005)은 참조모형으로 ARMA와 비교하였다.

SOI 자료의 경우 원자료에서 계절성을 제거한 Anomaly 자료를 사용하였다. 본 연구에서는 NAR 모형과 비교의 목적으로 SOI 자료를 선형 AR ($p, \dots, p_{\text{maximum}}$) 모형에 적합 시켰으며 선형 모형을 위

Table 4. Results of the goodness fit test of the linear AR model

Portmanteau Test		Null of Hypothesis	Jarque-Bera Test		Null of Hypothesis	ARCH-LM Test		Null of Hypothesis
Port	8.3730	Accepted	Test Statistic	4.145	Accepted	Test Statistic (Lag:4)	5.622	Accepted
P-Value	0.9370					P-Value	0.061	
L&B	25.410	Accepted	P-Value	0.126		Test Statistic (Lag:8)	16.2907	Rejected
P-Value	0.9321					P-Value	0.0384	

Table 5. Results of the BDS test of the linear AR model

Embedding Dimension	SOI	Null of Hypothesis		Residual	Null of Hypothesis	
$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$		$\alpha = 0.01$	$\alpha = 0.05$
$m = 2$	25.873	Rejected	Rejected	2.519	Accepted	Rejected
$m = 3$	28.459	Rejected	Rejected	3.799	Rejected	Rejected
$m = 4$	31.518	Rejected	Rejected	4.276	Rejected	Rejected
$m = 5$	34.096	Rejected	Rejected	3.981	Rejected	Rejected

마지막으로 Brock et al.(1991)이 제시한 BDS 검정을 $r = 0.5\sigma$ 와 Embedding Dimension 2에서 5를 대상으로 추정하여 Table 5에 나타내었다. 통계치가 매우 큰 값을 나타내고 있으며 따라서, 자료의 비선형성이 존재함을 확인할 수 있다.

Table 5에서 IID의 귀무가설은 95%와 99%의 신뢰 구간에서 기각되고 있으며 따라서 $AR(p_1, p_2, p_3, p_5, p_{10})$ 이 최적의 모형으로 선택되었지만 SOI의 비선형성을 효과적으로 반영할 수 없으며 비선형성을 고려할 수 있는 모형이 필요하다 하겠다.

7. Nonlinear Autoregressive Model

선형 AR 모형의 잔차검정 결과, SOI는 비선형특성을 고려할 수 있는 모형이 필요하다. 따라서 NAR 모형에 SOI를 적용하였으며 해석과정을 간략하게 정리하면 다음과 같다. 1) 모형의 지체시간 추정, 2) 조건부 평균 함수 추정, 3) 잔차의 검정, 4) 만약 잔차가 이분산성을 나타내게 되면 조건부 표준편차함수 추정의 순서로 연구를 진행하게 된다. Fig. 7은 SOI 원자료의 PACF를 나타낸다. 그림에서 보듯이 SOI 자료는 작은 지체를 갖는 모형이 적합할 것으로 예상된다.

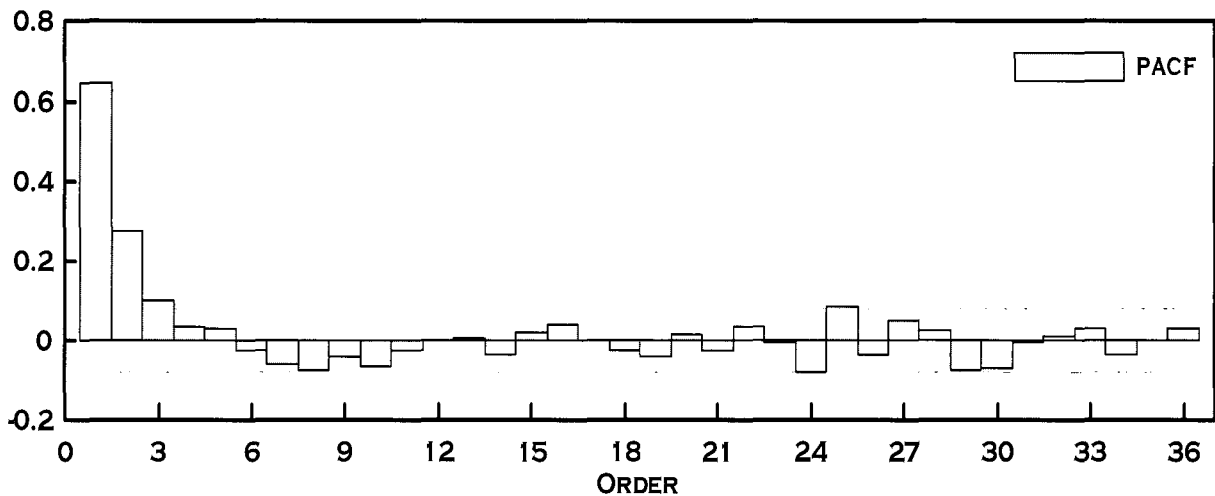


Fig. 7. Partial Autocorrelation Function (PACF) of Southern Oscillation Index

Table 6. Determination of lag time of the conditional mean function in the Nonlinear AR model

Conditional Mean Function													
No	Lag Vector										CAFPE	Bandwidth	
1	1											0.5818	0.7343
2	1	2										0.5270	0.6785
3	1	2	4									0.5135	0.7050
4	1	2	3	4								0.5311	0.7989
5	1	2	3	6	8							0.5681	0.8353
6	1	2	3	4	8	10						0.6200	0.8889
7	1	2	3	4	5	8	10					0.7069	0.8793
8	1	2	3	4	5	6	7	8				0.8272	0.9135
9	1	2	3	4	5	6	7	8	10			1.0804	0.9396
10	1	2	3	4	5	6	7	8	9	10		1.9418	0.9337

CAFPE를 이용하여 가능한 지체시간의 조합을 검토함으로써 최적의 지체 Vector를 추정할 수 있다. 본 연구에서는 최대 지체 10을 갖는 10개의 가능한 지체 Vector 조합을 검토하였으며 Table 6에 정리하여 나타내었다. 추정결과 지체 1, 2와 4를 갖는 조합에서 가장 작은 값을 나타내고 있으며 이를 토대로 조건부지체 1과 2의 평균함수를 추정하여 Figs. 8 and 9에 나타내었다. 그림에서 비선형성이 존재함을 확인할 수 있으며 특히, 지체 2와 4에서 강한 비선형성을 나타내고 있다.

적용된 조건부 평균함수의 적합성을 평가하기 위해서 잔차에 대한 검정을 실시하였다. 잔차에 대한 독립성은 Portmanteau Test에서 기각되었으나 정규성과 비이분산성에 대한 귀무가설은 각각 Jarque-Bera Test와 ARCH-LM Test에서 기각되었다. 무엇보다도 지체 8을

갖는 Jarque-Bera Test와 ARCH-LM Test에서 p값이 각각 0.022와 0.047를 나타내고 있다. 더욱이 CAFPE를 이용하여 조건부 표준편차함수의 지체 Vector를 추정한 결과 3, 8과 9를 갖는 조합이 최적의 모형으로 선택되어 이분산성이 존재하고 있음을 다시 한 번 확인할 수 있었으며 Table 7에 추정결과를 나타내었다. 따라서 추정된 지체시간을 토대로 조건부 표준편차함수를 추정하여 Figs. 10 and 11에 나타내었으며 강한 비선형성을 확인할 수 있었다.

Table 8에 잔차의 자기상관성, 정규성, 이분산성에 대한 검정결과를 제시하였다. 잔차에 대한 모든 IID 가정이 만족하고 있음을 확인할 수 있다. 마지막으로 잔차에 비선형성을 검토하기 위해서 BDS 검정을 실시하여 Table 9에 나타내었다. 잔차에 대한 IID 가정 모

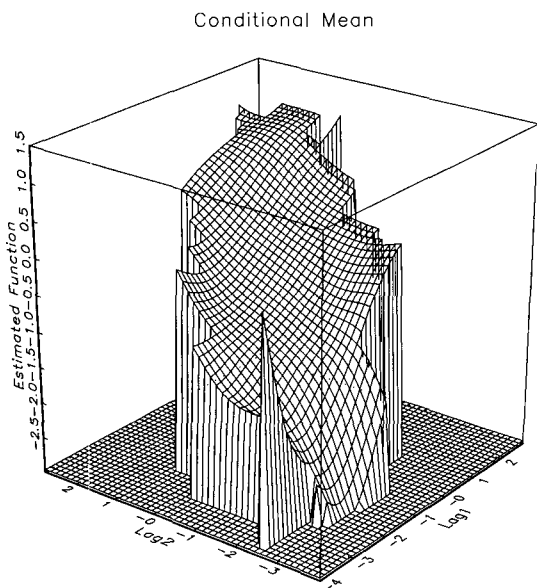


Fig. 8. Conditional mean function between lag-1 and lag-2

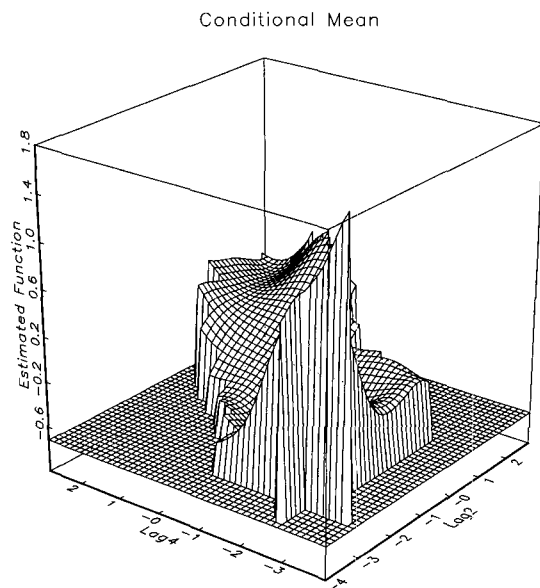


Fig. 9. Conditional mean function between lag-2 and lag-4

Table 7. Determination of lag time of the conditional standard deviation function in the Nonlinear AR model

Conditional Standard Deviation Function													
No	Lag Vector										CAFPE	Bandwidth	
1	4											0.4940	0.6215
2	3	9										0.4886	0.6964
3	3	8	9									0.4856	0.7560
4	3	4	8	9								0.5085	0.7696
5	3	4	6	9	10							0.5483	0.7776
6	1	3	5	6	8	9						0.6254	0.8427
7	1	2	3	4	8	9	10					0.7471	0.9203
8	1	2	3	5	6	8	9	10				0.9857	0.9199
9	1	2	3	4	5	6	8	9	10			1.3990	0.9080
10	1	2	3	4	5	6	7	8	9	10		2.8263	0.9232

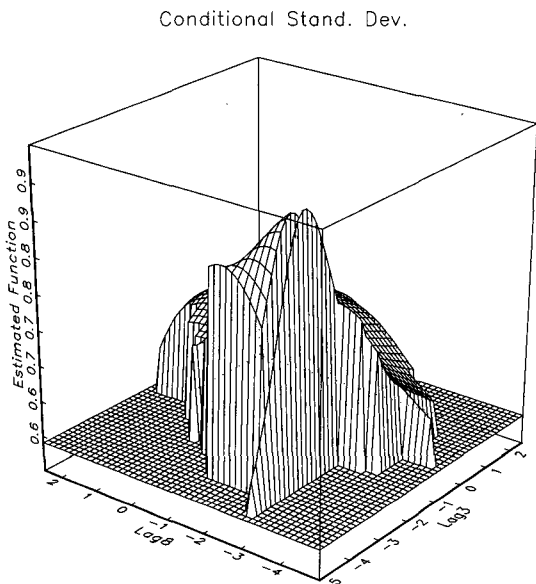


Fig. 10. Conditional standard deviation function between lag 3 and lag 8

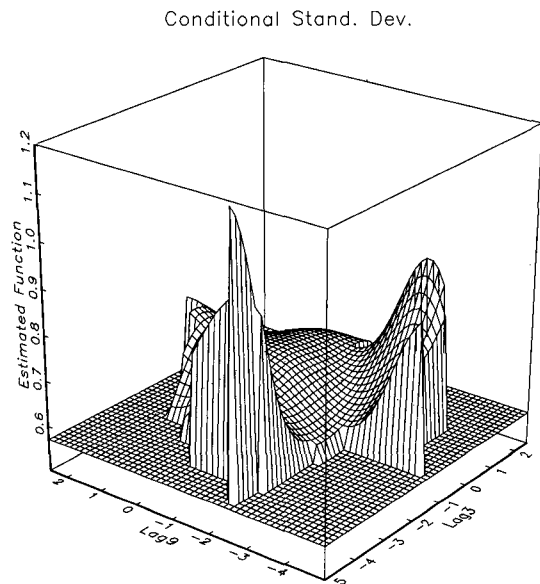


Fig. 11. Conditional standard deviation function between lag 3 and lag 9

Table 8. Results of the goodness fit test of the nonlinear AR model

Portmanteau Test		Null of Hypothesis	Jarque-Bera Test		Null of Hypothesis	ARCH-LM Test		Null of Hypothesis
Port	19.6032	Accepted	Test Statistic	5.1992	Accepted	Test Statistic (Lag:4)	3.7884	Accepted
<i>P</i> -Value	0.2386					<i>P</i> -Value	0.4354	
L&B	19.8560	Accepted	<i>P</i> -Value	0.0743		Test Statistic (Lag:8)	14.3869	Accepted
<i>P</i> -Value	0.2268					<i>P</i> -Value	0.0722	

두 신뢰구간 95%와 99%에서 만족하고 있으며 비매개변수적 NAR 모형이 SOI 자료의 특성을 표현하는데 더 유리함을 증명하고 있다.

본 연구에서는 모형의 적합성을 평가하기 위한 방법으로 One-Step 예측을 실시하여 비교하였다. One-Step 예측은 이용 가능한 j 시점까지 시계열 y_{T_0} 부터

Table 9. Results of the BDS test of the nonlinear AR model

EmbeddingDimension	SOI	Null of Hypothesis		Residual	Null of Hypothesis	
		$\alpha = 0.01$	$\alpha = 0.05$			
$\alpha = 0.01$	$\alpha = 0.05$					
$m = 2$	25.873	Rejected	Rejected	1.3266	Accepted	Accepted
$m = 3$	28.459	Rejected	Rejected	1.4974	Accepted	Accepted
$m = 4$	31.518	Rejected	Rejected	1.7729	Accepted	Accepted
$m = 5$	34.096	Rejected	Rejected	2.0230	Rejected	Accepted

$y_{T_0+1}, \dots, y_{T_0+j}$ 까지 반복적으로 예측을 수행하게 되며 Eq. (40)의 평균제곱예측오차(Mean Squared Prediction Error, MSEPE)를 가지고 예측결과를 평가하였다.

$$F_{error} = \frac{1}{T - T_0} \sum_{j=T_0+1}^{T_0+j} (\hat{y}_j - y_j)^2 \quad (40)$$

예측을 위해서 자료를 분리하여 전체자료의 15%를 예측대상으로 예측을 실시하였으며 NAR 모형에서 0.5464의 MSEPE를 나타내고 있으며 이는 선형 AR 모형에 비해 7% 작은 오차를 나타낸다. 무엇보다도 NAR 모형은 여타의 매개변수적 방법과 달리 모형의 선택에 있어 자유로우며 비선형성을 고려할 수 있는 모형으로서 SOI 자료와 같은 비선형 자료를 위한 모의방법으로 선형 모형에 비해 유리하다 하겠다.

8. 결 론

본 연구에서는 조건부 핵밀도함수와 CAFPE 지치시 간 결정방법을 토대로 비선형 자기회귀모형(Nonlinear AutoRegressive, NAR)을 이용한 비매개변수적 수문시계열 모의방법을 제시하였다. 선택된 선형 AR ($p_1, p_2, p_3, p_5, p_{10}$) 모형은 잔차에 대한 비자기상관성과 정규성의 가정은 받아들여지나 이분산성이 존재하고 있음을 ARCH-LM Test에서 확인할 수 있었고 더욱이, 비선형성이 BDS 검정에서 확인되었다. 이는 최근에 Ahn and Kim (2005)과 같은 결과를 나타낸다.

따라서 비매개변수적 NAR 모형에 SOI 자료를 적용하였으며 조건부 평균함수에 대해서 CAFPE를 이용하여 1, 2와 4를 최적의 지체 Vector로 추정하였다. 모의 결과잔차에 대한 비정규성과 이분산성이 확인되었다. 또한 CAFPE를 이용한 조건부 표준편차함수의 최적 지체 Vector로 3, 8과 9가 선택되어 SOI 자료의 이분산성이 존재함을 다시 한 번 확인할 수 있었다.

이를 토대로 조건부 평균함수와 표준편차함수를 모두 고려한 모의 결과 잔차에 대해서 모든 IID 가정이

유지되고 있으며 특히 BDS 검정에서 신뢰구간 95%와 99%에서 귀무가설이 받아들여졌다. 마지막으로 모형의 적합성을 평가하기 위해서 One-Step 예측을 수행하여 평균제곱오차를 중심으로 평가한 결과 NAR 모형이 선형 모형에 비해 대략 7% 개선된 결과를 나타내었다.

따라서, NAR 모형은 여타의 매개변수적 방법과 달리 모형에 선택에 있어 자유로우며 비선형성을 고려할 수 있는 모형으로서 SOI 자료와 같은 비선형 자료를 위한 모의방법으로 선형 모형에 비해 많은 장점을 가지고 있다. 그러나 모형의 계산상의 어려움으로 인해 다변량 시계열분석에 제약이 따르며 또한 광역폭 산정에 있어 여전히 보완이 필요하다 하겠다.

감사의 글

본 연구는 건설교통부 한국건설교통기술평가원의 이상기후대비시설기준강화 연구단에 의해 수행되는 2005 건설기술기반구축사업(05-기반구축-D03-01)에 의해 지원되었습니다.

참 고 문 헌

권현한, 문영일 (2006). “상태-공간 모형과 Nearest Neighbor 방법을 통한 수문시계열 예측에 관한 연구.”, **대한토목학회 논문집**, 대한토목학회, Vol.25 No.4B, pp. 275-283

문영일, 이정규 (1995). “변동핵 추정법에 의한 비매개변수적 홍수빈도 분석.”, **대한토목학회 논문집**, 대한토목학회, Vol.15 No.2, pp. 413-424.

문영일, 권현한 (2004). “수리·수문학적 댐 위험도 분석 (I) -비매개변수적 LHS Monte Carlo Simulation 위험도 해석 기법 개발.”, **대한토목학회 논문집**, 대한토목학회, Vol.24 No.5B, pp. 477-487.

문영일, 권현한 (2004). “수리·수문학적 댐 위험도 분석 (II) - 수리·수문학적 댐 위험도 분석(II)-댐 위험도 해석 예시.”, **대한토목학회 논문집**, 대한토목학회, Vol.24 No.5B, pp. 489-500.

- Akaike, H. (1969). "Fitting autoregressive models for prediction.", *Annals of the Institute of Statistical Mathematics*, Vol. 21, pp. 243-247.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, in B. Petrov & F. Csáki (eds.), *2nd International Symposium on Information Theory*, Akadémiai Kiadó, Budapest, pp. 267-281.
- Ahn, J.H. and Kim, H. S. (2005). "Nonlinear Modeling of El Nino/Southern Oscillation Index.", *Journal of Hydrologic Engineering*, Vol. 10, No. 1. pp. 8-15.
- Brock, W. A., Hsieh, D. A., and LeBaron, B. (1991). *Nonlinear dynamics, chaos, and instability: Statistical theory and economic evidence*, MIT Press, Cambridge, USA.
- Brock, W. A., Dechert, W. D., LeBaron, B., and Scheinkman, J.A. (1997). "A Test for Independence Based on the Correlation Dimension.", *Econometric Reviews*, Vol. 15, pp. 197-235.
- Chu, P. S., and Katz, R. W. (1985). "Modeling and forecasting the southern oscillation: A time-domain approach.", *Mon. Weather Rev.*, Vol. 113, pp. 1876-1888.
- Elshorbagy, A., Simonovic, S. P. and Panu, U. S. (2002). "Estimation of missing streamflow data using principles of chaos theory.", *J. Hydrol.*, Vol. 255, pp. 123-133.
- Engel, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.", *J. Econometr.*, Vol. 50, pp. 987 - 1007.
- Engel, R. F. (1983). "Estimates of the variance of U.S. inflation based upon the ARCH model.", *J. Money, Credit, Banking*, Vol. 15, pp. 287 - 301.
- Hannan, E. J. and Quinn, B. G. (1979). "The determination of the order of an autoregression.", *Journal of the Royal Statistical Society*, pp. 190-195.
- Hense, A. (1987). "On the possible existence of a strange attractor for the southern oscillation.", *Beitr. Phys. Atmos.*, Vol. 60, pp. 34-47.
- Jayawardena, A.W. and Lai, F. (1994). "Analysis and prediction of chaos in rainfall and stream flow time series.", *J. Hydrol.*, Vol. 153, pp. 23-52.
- Jarque, C. M. and Bera, A. K. (1987). "A test for normality of observations and regression residuals.", *International Statistical Review*, Vol. 41, pp. 1781-1816.
- Jayawardena, A. W., Li, W. K. and Xu P. (2002). "Neighborhood selection for local modeling and prediction of hydrological time series." *J. Hydrol.*, Vol. 258, pp. 40-57.
- Kim, H. S., Kang, D. S., and Kim, J. H. (2003). "The BDS statistic and residual test." *Stoch Environ. Res. Risk Assess.*, Vol. 17, pp. 104 - 115.
- Kwon, H.-H. and Y.-I. (2006). Moon, "Improvement of Overtopping Risk Evaluations Using Probabilistic Concepts for Existing Dams.", *Stochastic Environmental Research and Risk Assessment*, Vol. 20, pp. 223-237.
- Lall, U., Sangoyomi, T., and Abarbanel, H. D. I. (1996). "Nonlinear dynamics of the Great Salt Lake: Nonparametric short-term forecasting." *Water Resour. Res.*, Vol. 32, pp. 975-985.
- Lall, U. and Mann, M. E. (1995). "The Great Salt Lake: A barometer of low-frequency climatic variability." *Water Resour. Res.*, Vol. 31, pp. 2503-2515.
- Ljung, G. M and Box, G. E. P. (1978). "On a measure of lack of fit in time-series models.", *Biometrika*, Vol. 65. pp. 297-303.
- Lomnicki, Z.A. (1961). "Tests for departure from normality in the case of linear stochastic processes.", *Metrika*, Vol. 4, pp. 37-62.
- Rissanen, J. (1978). "Modeling by shortest data description.", *Automatica*, Vol. 14, pp. 465 - 471.
- Rodriguez-Iturbe I., De Power, F. B., Sharifi, M. B. and Georgakakos, K. P. (1989). "Chaos in rainfall.", *Water Resour. Res.*, Vol. 25, pp. 1667-1675.
- Schwarz, G. (1978). "Estimating the dimension of a model.", *Annals of Statistics*, Vol. 6, pp. 461 - 464.
- Silverman, B. W. (1996). "*Density Estimation for Statistics and Data Analysis*.", Chapman & Hall, London.
- Terrence C. and Compo GP. (1998). "A practical guide to wavelets analysis.", *Bulletin of the American Meteorological Society*, Vol. 79, No. 1, pp. 61-78.
- Trenberth, K. E., and T. J. Hoar. (1996). "The 1990 - 1995 El Niño - Southern Oscillation event: Longest

on record.", *Geophys. Res. Lett.*, 23, 57 - 60.

Tschemig, R. and Yang, L. (2000a). Nonparametric estimation of generalized impulse response functions, SFB 373 Discussion Paper 89, 2000, Humboldt-University, Berlin.

Tschemig, R. and Yang, L. (2000b). "Nonparametric lag selection for time series.", *Journal of Time*

Series Analysis, Vol. 21, pp. 456-487

Yang, L. and Tschemig, R. (2000b). "Multivariate bandwidth selection for local linear regression.", *Journal of the Royal Statistical Society, Series B*, Vol. 61, pp. 793-815.

(논문번호:05-86/접수:2005.06.21/심사완료:2006.10.30)