

변형함수분석을 이용한 댐수위와 안개지속시간의 인과관계

제갈돈¹

요약 1981년 4월부터 1992년 12월까지 안동댐의 수위와 안동지역의 안개지속시간 사이의 관계를 변형함수분석(transfer function analysis)를 이용하여 조사한 결과, 댐수위의 증가는 4개월의 시차를 두고 단위당(m) 0.89시간씩 안개지속시간을 증가시키고 있음을 확인하였다. 그러므로 댐건설로 인한 인공호수의 생성이 안개발생의 직접적인 원인 중의 하나라고 말할 수 있다. 앞으로는 이같은 객관적인 증거에 기초한 수자원개발정책과 댐관리정책이 수립되어야 할 것이다.

1. 서론

오늘날 인구의 증가와 산업의 급속한 발전에 따라 물의 사용은 전례없이 늘어났고, 수자원을 안정적으로 관리하고 공급하기 위해 대하천 상류에 다목적댐을 건설하여 하천 하류지역에 생활·공업·농업용수를 공급하고 전력생산 및 홍수방지에 집중적인 노력을 투자하고 있다. 그러나 막상 댐을 보유하게 되는 지역은 댐건설로 인하여 상당한 물질적·정신적 피해를 감수하고 있는 것이 현실이다. 이러한 댐주변 지역 주민들이 겪고 있는 피해에 대한 소수의 연구를 종합하여 보면, 가장 심각한 피해요인은 댐주변지역의 기상변화이며, 이에 따른 각종의 피해가 주장되고 있다(남치호 외, 1994; 이종범 외, 1990; 최순호 외, 1989).

댐주변지역의 기상변화에 대한 객관적인 증거에 의하면, 댐건설로 인하여 댐주변지역의 안개발생을 유의하게 증가되었다고 한다(제갈돈, 1995). 안개는 대기 중 수증기의 복사냉각현상에 의하여 발생하기 때문에, 댐의 건설로 인하여 생겨난 거대한 인공호수는 광대한 면적의 수면을 생성하여 댐주변지역에 다량의 수증기를 발생시킴으로써 안개 발생에 직접적인 원인을 제공하였을 것이라고 추측할 수 있다.

그러므로 본 연구는 안동댐의 사례를 분석하여 댐수면의 증가와 안개발생 사이의 인과관계에 대한 가설을 경험적으로 검증하려고 한다. 댐수면에 대한 측정치로는 댐의 수위를 사용하고, 안개의 측정치로는 안개지속시간을 사용할 것이다. 인과관

¹ 안동대학교 행정학과

계의 검증을 위해서는 시계열 사이의 인과관계 검증에 가장 적절한 변형함수분석을 적용한다. 변형함수분석은 댐수위와 안개지속시간 사이의 인과관계를 밝혀줄 뿐만 아니라, 댐수위의 단위당(m) 증가가 안개지속시간에 얼마 만큼의 시간적인 간격을 두고 어느 정도 영향을 미치는가라는 관계의 시차와 크기도 확인할 수 있게 한다. 이러한 사실의 확인을 통해서만이 정확한 수자원개발정책과 댐수위관리정책의 수립을 기대할 수 있을 것이다.

2. 변형함수모형과 분석방법

1. 변형함수모형

선형회귀모형을 이용하여 두 변수 사이의 인과관계를 규명하는 것처럼 투입시계열과 산출시계열 사이의 관계는 변형함수를 이용하여 하나 혹은 그 이상의 시계열을 투입한 후의 산출시계열의 효과를 분석할 수 있다. 일반적으로 변형함수는 지표가 되는 다른 시계열의 과거나 현재의 값을 가지고 어떤 시계열의 현재값을 예측하는데 사용되거나, 어떤 시계열이 공변량(covariate)으로서 다른 시계열에서의 잡음(noise)을 제거하여 개입(intervention)의 효과를 보다 세밀하게 검증하는데 이용되기도 한다.

일반적인 변형함수모형은 다음과 같다.

$$Y_t = f(X_t) + N_t \quad (1)$$

모형(1)에서 Y_t 는 산출시계열이며 X_t 는 투입시계열이다. $f(X_t)$ 는 우리가 구축하고 검증하려는 변형함수이고, N_t 는 ARIMA 모형을 포함하는 잡음항이다. 여기서 변형함수 $f(X_t)$ 의 일반적인 형태는 다음과 같다.

$$f(X_t) = \frac{\omega(B)}{\delta(B)} \cdot X_{t-b} \quad (2)$$

모형(2)에서 $\omega(B)$ 는 개입의 크기에 대한 다항식을 나타내고, $\delta(B)$ 는 개입의 지수적 감소비율을 나타내는 다항식을 의미한다. b 는 투입시계열이 산출시계열에 미치는 효과의 지연(delay)을 표시한다. 일반적인 변형함수모형(2)에서 가장 간단한 모형은 투입시계열이 b 라는 시간간격을 두고 일시적으로 ω_0 만큼 산출시계열에 영향을 주는 경우이다. 이 경우를 식으로 표기하면 다음과 같이 된다.

$$f(X_t) = \omega_0 X_{t-b} \quad (3)$$

모형(3)에서 $(b-1)$ 시점에서 유의한 효과가 하나 더 존재하는 경우에는 모형(3)의 다항식이 다음과 같이 변형된다.

$$f(X_t) = \omega_0 X_{t-b} + \omega_1 X_{t-b-1} = (\omega_0 + \omega_1 B) X_{t-b} \quad (4)$$

만약 모형(3)에서 투입시계열과 산출시계열 사이의 인과관계가 지수적인 감소의 효

과가 존재하는 경우 변형함수모형은 다음과 같이 형성된다.

$$f(X_t) = \frac{\omega_0}{(1-\delta B)} X_{t-b}, \quad 0 < \delta < 1 \quad (5)$$

모형(5)에서는 산출시계열에 대한 투입시계열의 효과는 δ 의 크기에 따라 기하급수적인 성격을 지닌 분산적 시차구조를 나타내고 있다. 궁극적으로 X_t 에서의 변화가 초래하는 Y_t 에서의 점근적(asymptotic) 변화는 $\omega_0/(1-\delta)$ 가 될 것이다.

2. 변형함수분석

변형함수분석은 세 가지의 단계로 진행된다. 첫번째 단계는 변형함수를 분석하는 필요한 통계치의 계산을 용이하게 하기 위하여 사전에 투입시계열과 산출시계열을 검토하는 것이다. 두번째 단계에서는 변형함수를 식별하고 잔차 N_t 를 계산한다. 마지막으로 N_t 의 ARIMA 모형을 구축하여 모형 전체의 적합성을 검증하는 단계이다. 이 세단계 중에서 가장 핵심적인 단계는 두번째의 변형함수를 식별하는 단계이다. 이 단계에서 투입시계열과 산출시계열의 관계를 식별하기 위해서는 일반적으로 추정된 교차상관함수(cross-correlation, CCF)를 이론적 교차상관함수와 비교하여 관계의 패턴을 밝혀낸다 (McCleary and Hay, 1980).

CCF(k)는 시차 t 에서의 투입시계열 X_t 와 시차 $t+k$ 에서의 산출시계열 Y_t 사이의 관계를 나타낸다. 그러므로 $k>0$ 일 때 CCF(k)는 X_t 와 미래의 Y_t 값 사이의 상관관계를 나타내고, $k<0$ 일때의 CCF(k)는 X_t 와 과거의 Y_t 값 사이의 상관관계를 나타낸다. ARIMA 방법에서의 자기상관함수(ACF)와 편자기상관함수(PACF)는 CCF와 유사하지만, CCF는 0을 중심으로 대칭을 이루지는 않는다. 즉 CCF(k)는 CCF(-k)와 일치하지 않는다.

CCF는 X_t 와 Y_t 사이의 관계에 대한 추정치이지만, 이것은 X_t 와 Y_t 의 자기상관관계(autocorrelation)로부터 영향을 받는다. 만약 X_t 나 Y_t 가 정상적(stationary)이 아니거나 자기상관관계를 포함하고 있는 경우, CCF는 X_t 와 Y_t 사이의 참(true)관계를 반영하지 못한다. 그러므로 CCF를 계산하기 전에 반드시 자기상관관계를 제거해야만 한다. 가장 광범위하게 사용되는 자기상관관계의 제거방법은 CCF의 계산 전에 시계열을 사전백색화(prewhitening)하는 것이다 (Box and Jenkins, 1976).

사전백색화는 다섯 단계로 이루어지는데, 첫번째 단계는 각 시계열을 정상적으로 만들기 위하여 차분(differencing)을 하는 것이다. 차분된 후의 투입시계열과 산출시계열을 각각 W_t 와 Z_t 라고 하면, 각각의 모형은 다음과 같이 표기된다.

$$W_t = (1 - B)^d X_t \quad (6)$$

$$Z_t = (1 - B)^d Y_t \quad (7)$$

두번째 단계는 차분된 투입시계열 W_t 의 ARIMA 모형을 구축하는 것이다.

$$W_t = \frac{\Theta(B)}{\Phi(B)} a_t \quad (8)$$

세번째는 W_t 를 백색화 하기 위하여 a_t 로 전환하는 것이다.

$$a_t = \frac{\Phi(B)}{\Theta(B)} W_t \quad (9)$$

네번째는 전환된 함수를 통하여 Z_t 를 여과(filtering)하여 b_t 를 생성한다.

$$b_t = \frac{\Phi(B)}{\Theta(B)} Z_t \quad (10)$$

마지막 단계에서는 사전백색화된 a_t 와 사전여과된 b_t 사이의 교차상관함수를 구한다. 여기에서 생성된 CCF는 변형함수 $f(X_t)$ 에서의 $\omega(B)$ 와 $\delta(B)$ 의 다항식을 식별할 수 있는 단서를 찾을 수 있으며, 이 단서에 근거하여 변형함수모형을 완성할 수 있다.

3. 땀수위와 안개지속시간의 인과관계

안개지속시간

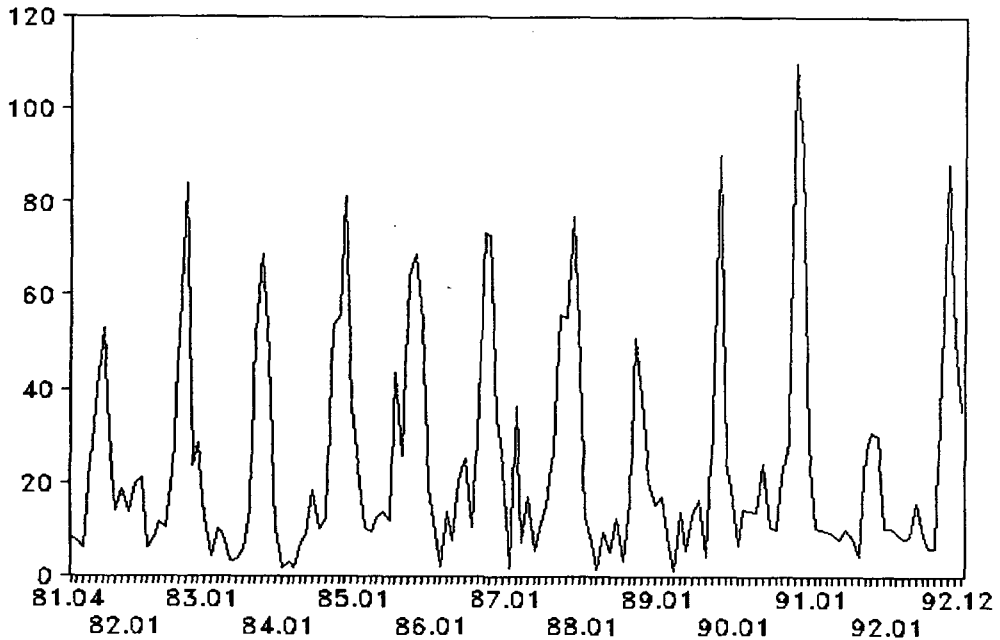


그림 1. 안동지역의 안개지속시간 시계열 그래프

그림 1은 투입시계열인 1981년 4월부터 1992년 12월까지의 안동댐의 수위를 보여주고 있고, 그림 2는 산출시계열인 안동지역의 안개지속시간을 보여주고 있다. 그림에서 보면 두 시계열은 측정기간 동안 계절적인 주기를 가지고 변동하고 있음을 알 수 있다. 이같은 동일한 변동패턴을 보고 연구자들은 쉽게 두개의 시계열이 인과관계를 가지고 있다고 결론을 내릴 수 있으나, 이 결론은 매우 위험하다. 계절적인 변동패턴이 시계열 사이의 인과관계에 대해서는 아무런 단서도 제공해 줄 수 없다. 두 시계열 사이의 인과관계를 검증하기 위해서는 보다 구체적인 변형함수모형을 구축할 필요가 있다.

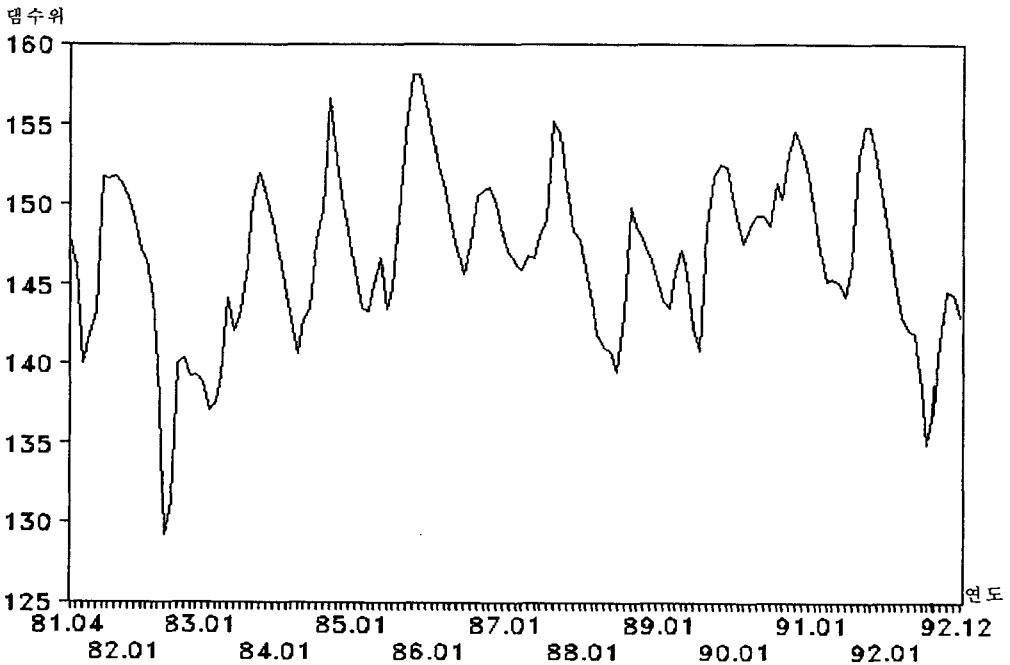


그림 2. 안동댐의 수위 시계열 그래프

변형함수모형을 구축하기 위한 첫번째 단계는 두개의 시계열에 대한 일변량 ARIMA 모형을 구축해야 한다. 사전에 두 시계열에 대한 ARIMA 모형을 구축하는 이유는 두 시계열이 정상적인가를 검사하고, 각각의 시계열이 어떠한 자기회귀(AR)와 이동평균(MA)의 요소를 가지고 있는가를 사전에 확인하기 위해서이다. 만약 시계열이 정상적이 아니거나 다른 특정한 패턴을 가지는 경우에는 두 시계열 사이의 교차상관관계가 지나치게 팽창되어 허위의 관계를 나타낼 가능성이 많기 때문이다. 각각의 시계열에 대한 자기상관함수(ACF)와 편자기상관함수(PACF)를 분석한 후 잠정적으로 식별된 투입시계열(X_t)과 산출시계열(Y_t)의 ARIMA 모형은 다음과 같다. 식별된 모형

(1)과 모형(2)에 대한 모수 추정치를 표 1 에서 보면, 모든 추정치가 통계적으로 유의함을 알 수 있다. 또한 모형진단을 위한 Q 통계량이 통계적으로 유의하지 않음으로 보아 식별된 투입과 산출시계열이 통계적으로 적합하다고 할 수 있다.

$$X_t = \frac{(1-\theta_1 B)(1-\Theta_{12} B^{12})}{(1-B)(1-B^{12})} a_t \tag{11}$$

$$Y_t = \frac{(1-\theta_1 B)(1-\Theta_{12} B^{12})}{(1-\phi_1 B)(1-B^{12})} a_t \tag{12}$$

표 1 모형(1)과 모형(2)의 모수 추정치

모형	모수추정치	t 값 ^{a)}	Q 통계량 ^{b)}
(1)	$\theta_1 = -0.16$	-1.84	$Q_{21} = 16$
	$\Theta_{12} = 0.9$	28.90	
(2)	$\theta_2 = 0.27$	3.32	$Q_{20} = 18$
	$\Theta_{11} = 0.34$	4.43	
	$\Theta_{12} = 0$		

a) 모든 t 값은 0.05 수준에서 유의함 b) 모든 Q 값은 0.1 수준에서 유의하지 않음

이 두 시계열의 인과관계에 관한 가설은 투입시계열(X_t)인 안동댐의 수위가 산출시계열(Y_t)인 안동지역의 안개지속시간에 영향을 미친다는 것이다. 그렇지만 이 같은 인과관계의 순서가 정확한 가설인지, 만약 정확하다면 과연 어느 정도의 시간적인 지연을 두고 댐의 수위가 안개지속시간에 영향을 미치는지는 두 시계열 사이의 교차상관함수(CCF)를 통하여 밝혀질 것이다.

그러나 교차상관함수에 사용되는 시계열이 정상성을 위한 차분 후에도 시계열이 백색잡음이 되지 않은 경우 두 시계열 사이의 교차상관관계는 오류를 범할 가능성이 많다. 따라서 CCF는 반드시 투입시계열이 백색잡음 과정일 때만이 설명력을 가진다. 설사 CCF가 상당히 강한 인과관계의 단서를 제시하더라도 투입시계열이 백색잡음이 안되면 아무런 단서로서의 역할을 할 수 없다. 왜냐하면 투입시계열이 백색잡음이 아니면 CCF는 시계열 사이의 종속성이나 시계열 내의 종속성을 반영하게 된다.

시계열 내의 상관관계 즉 자기상관관계를 제거하기 위한 가장 실체적인 방법은 시계열을 사전백색화하는 것이다. 분석 이전에 시계열을 사전백색화 함으로써 CCF에서 시계열의 자기상관관계를 제거할 수 있다. 사전백색화를 위해서는 투입시계열 X_t 에 대한 일변량 ARIMA모형을 변형하여 두 시계열에 적용한다. 모형(11)에 나타나 있는 투입시계열을 사전백색화 하면 다음과 같이 된다.

$$a_t = \left[\frac{(1+0.16B)(1-0.90B^{12})}{(1-B)(1-B^{12})} \right]^{-1} X_t \tag{13}$$

투입시계열의 ARIMA 모형을 산출시계열에 적용하여 사전여과시키면 사전여과된 산출시계열은 다음과 같다.

$$b_t = \left[\frac{(1+0.16B)(1-0.90B^{12})}{(1-B^{12})} \right]^{-1} Y_t \quad (14)$$

모형(13)과 모형(14)에 나타나 있는 시계열 사이의 CCF 는 두 시계열 사이의 분명한 인과관계에 대한 단서를 제공할 것이다. 그림 3의 교차상관함수를 보면 +4의 시차에서 통계적으로 유의한 하나의 스파이크가 존재하고 있음을 볼 수 있다. 이 사실

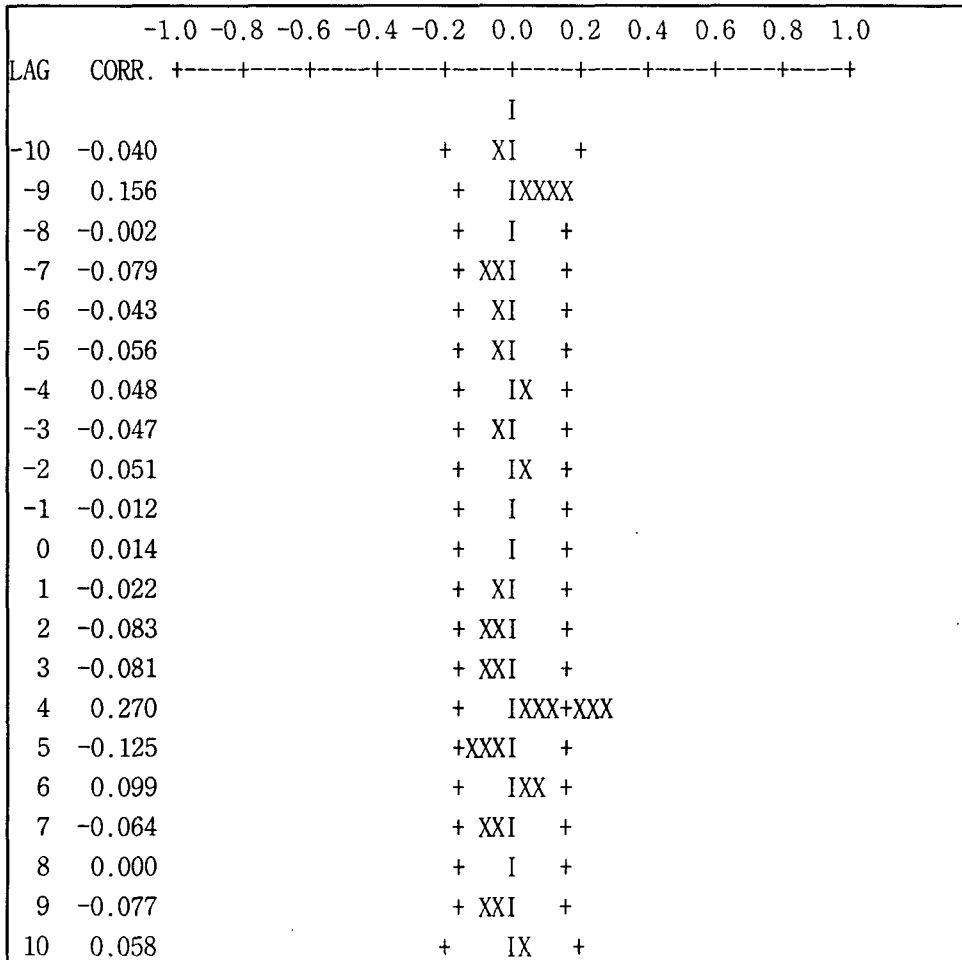


그림 3. 사전백색화된 댐수위와 안개지속시간시계열의 교차상관함수

은 댐의 수위가 정확하게 4 개월만큼 안개지속시간을 선행하여 유도하고 있음을 의미한다. 따라서 우리가 제시했던 가설처럼 댐의 수위가 투입시계열이고 안개지속시간이 산출시계열이라고 할 수 있다. 만약 통계적으로 유의한 스파이크가 음의 CCF 에서 나타났다면 우리가 세웠던 인과관계의 순위는 반대로 바뀌어서 안개지속시간이 투입시계열이 되고 댐의 수위가 산출시계열이 되는 문제가 발생했을 것이다.

교차상관함수(CCF) 상에서 통계적으로 유의한 스파이크가 시차 +4 에서 오직 한개만이 존재하고 있다는 사실을 확인한 후의 변형함수모형은 모형(3)에서 제시된 가장 간단한 모형으로 다음과 같다.

$$(1-B^{12})Y_t = (1-B)(1-B^{12})\omega_0 X_{t-4} + N_t \quad (15)$$

모형(15)에서 보면 댐의 수위의 변화는 4 개월 후 안개지속시간에 ω_0 의 비율만큼 영향을 미친다는 사실을 알 수 있다. 시차 +4 에서 오직 하나의 통계적으로 유의한 스파이크가 존재하기 때문에 이 두 시계열 사이의 인과관계의 비율에는 다른 역동적(dynamic) 변형의 효과가 존재하지 않는다. 즉 일시적인 영향 이외의 지수적인 감소 혹은 증가 효과는 없다.

두번째 단계는 변형함수모형(15)에서 잡음항인 N_t 를 식별하는 것이다. N_t 를 식별하는 데는 여러 가지 방법이 있으나 가장 적합한 방법은 N_t 를 변형함수의 잔차에서 식별하는 것이다. 이를 위해서는 먼저 N_t 를 백색잡음으로 가정하는 것이다. N_t 를 백색잡음으로 가정한 후의 모형은 다음과 같다.

$$(1-B^{12})Y_t = (1-B)(1-B^{12})\omega_0 X_{t-4} + a_t \quad (16)$$

모형(16)의 잔차에 대한 ACF는 그림 4에 나타나 있다. 그림 4의 ACF에서 N_t 에 대한 ARIMA 모형은 다음과 같다.

$$N_t = \frac{(1-\Theta_{11}B^{11}-\Theta_{12}B^{12})}{(1-\phi_1B)} a_t \quad (17)$$

모형(17)의 N_t 를 모형(15)에 대입시킨 후 잠정적으로 확정된 완전한 변형함수모형은 다음과 같이 된다.

$$Y_t = (1-B)\omega_0 X_{t-4} + \frac{(1-\Theta_{11}B^{11}-\Theta_{12}B^{12})}{(1-\phi_1B)(1-B^{12})} a_t \quad (18)$$

모형(18)에 대한 모수 추정치를 표 2에서 보면, 모든 추정치가 0.05 수준에서 통계적으로 유의하게 나타나고 있다. 이 사실로 미루어 추정단계에서는 변형함수모형이 일단 적합함을 알 수 있다. 보다 정확한 모형의 적합성을 검증을 위하여 모형(18)의 잔차에 대한 ACF를 그림 5에서 보면, 그래프 상에서 통계적으로 유의한 스파이크가 하나도 존재하고 있지 않음으로 미루어 모든 자기상관함수가 통계적으로 유의하지 않다.

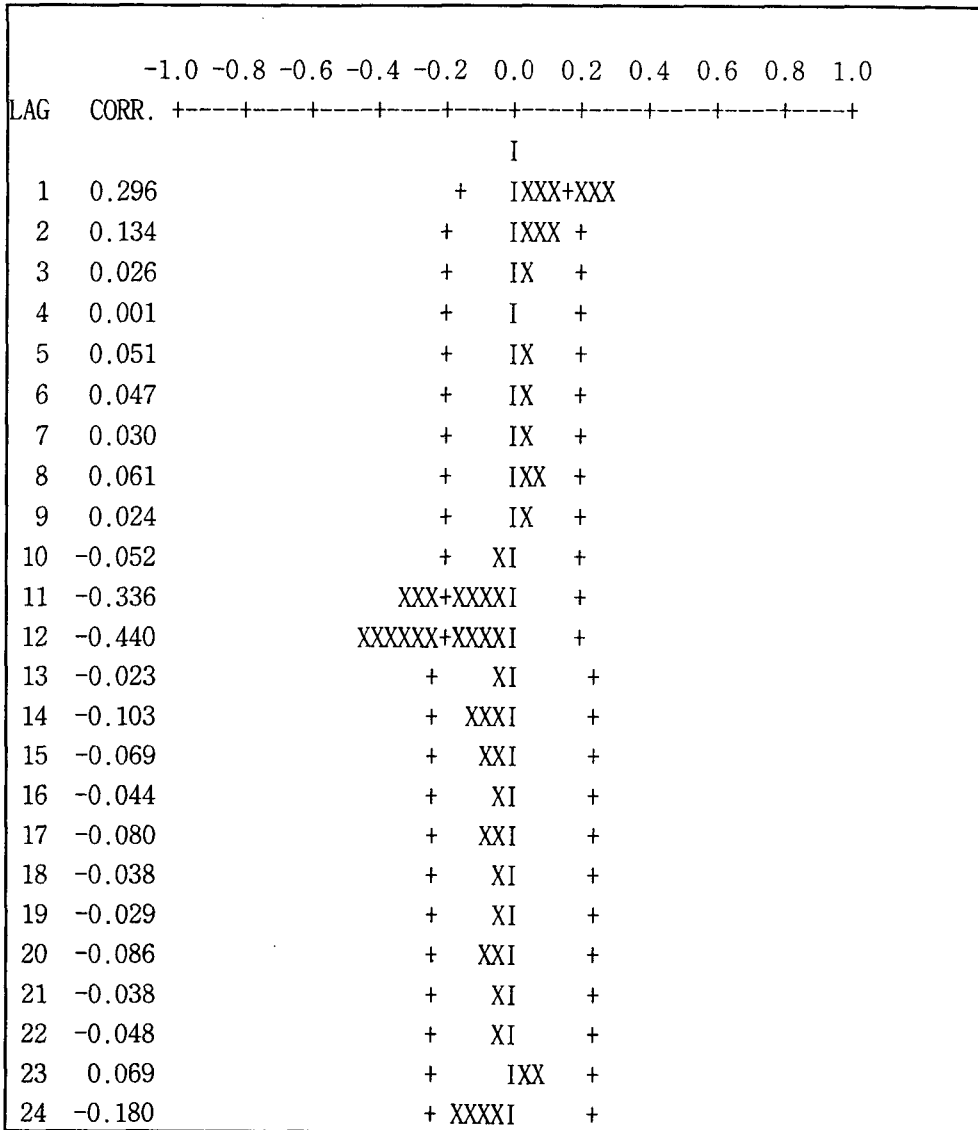


그림 4. 교차상관 후의 잔차 ACF

않다고 할 수 있다. 또한 모형진단을 위한 Q 통계량도 0.1 수준에서 통계적으로 유의하지 않게 나타나고 있어, 모형(18)은 통계적으로 적합한 모형이라고 확정할 수 있다. 이 모형의 적합성이 검증되었기 때문에, 댐 수위가 안개지속시간에 영향을 미친다는 결론을 내릴 수 있다. 그리고 이들 인과관계의 시차는 4개월이며 크기는 0.89 시간임을 알 수 있다.

표 2. 변형함수모형(18)의 모수 추정치

모형	모수 추정치	t 값 ^{a)}	Q 통계량 ^{b)}
(6)	$\phi_1 = 0.28$	3.43	Q ₁₉ = 15
	$\theta_{11} = 0.33$	4.08	
	$\theta_{12} = 0.90$	36.15	
	$\omega_0 = 0.89$	1.85	

a) 모든 t 값은 0.05 수준에서 유의함. b) 모든 Q 값은 0.1 수준에서 유의하지 않음.

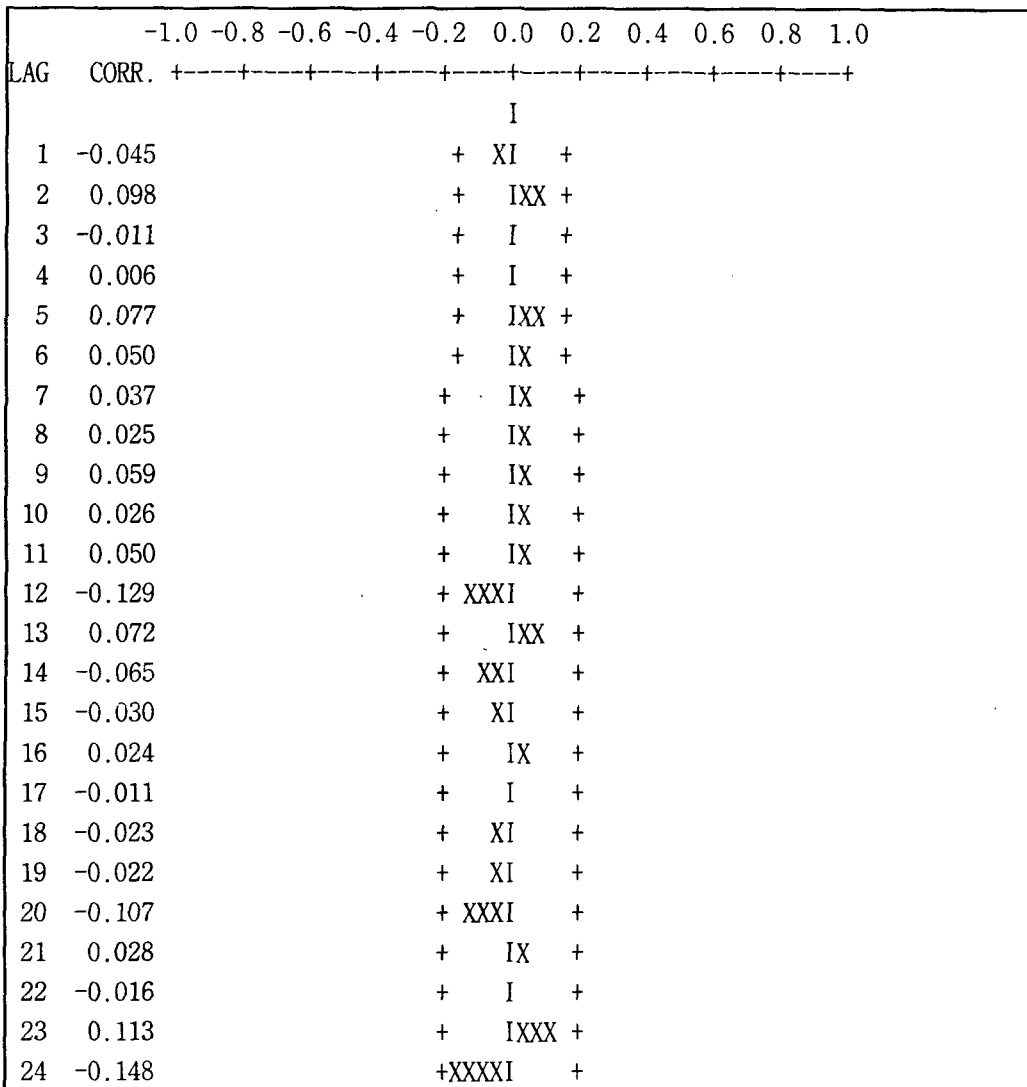


그림 5. 변형함수모형(18)의 잔차 ACF

4. 결 론

변형함수분석을 통하여 안동댐의 수위와 안동지역의 안개지속시간을 조사한 결과 댐수위가 안개지속시간에 영향을 미친다는 사실을 확인하였다. 구체적인 인과관계의 패턴을 보면, 댐수위의 증가가 4 개월의 시차를 두고 안개지속시간의 증가에 영향을 미치고 있고, 그 영향의 크기는 댐수위 1m 증가당 0.89 시간임을 확인하였다. 따라서 이같은 객관적인 증거를 바탕으로 앞으로의 수자원개발정책과정에서 댐의 규모에 따른 안개발생 가능성을 충분히 사전에 고려하여야 할 것이다. 또한 댐을 관리하는 수자원공사도 댐의 수위에 따라 댐주변지역의 안개 발생이 어느 정도일 것인가를 예측하여 댐을 관리할 필요가 있다.

비록 본 연구는 안개지속시간에 미친 댐의 수위만을 인과관계 상의 투입시계열로 파악하고 있다는 한계는 있으나, 댐수위뿐만 아니라 다른 원인요소들도 고려하여 다수의 투입시계열을 삽입한 변형함수분석을 통한 보다 완전한 인과관계의 검증은 앞으로의 과제일 것이다. 또한 이 분석모형을 댐건설로 인한 다른 부정적인 영향 즉 대기오염, 질병 등의 분석에 적용하여 보다 포괄적인 댐건설로 인한 영향평가에 기여할 수 있을 것이다.

참 고 문 헌

- 기상청(1981-1992), [기상연보].
 수자원공사 안동다목적댐 관리사무소(1981-1992), [안동다목적댐 통계자료].
 남치호 외(1994), [다목적댐의 사회경제적 영향분석], 안동대학교 안동지역개발연구소.
 이종범 외(1990), "댐건설에 따른 국지기후의 변화실태," [기상연구논문집], 17 권 1호, 기상연구소.
 제갈돈(1993), [간여시계열 실험과 분석], 아카데미.
 제갈돈(1995), "간여시계열분석을 통한 안동댐의 기상영향평가," [한국행정학보], 29 권 1호, 한국행정학회.
 최순호 외(1989), "안개상습지에서 기상요인이 수도수량에 미치는 영향," [농사시험보고서], 31 권 3호, 경상북도농촌진흥원.
 Box, G. E. P. and Jenkins, G. M. (1976), *Time Series Analysis: Forecasting and Control*, rev. ed. San Francisco: Holden-day.
 Glass, G. V., Wilson, V. L. and Gottman, J. M. (1975), *Design and Analysis of Time Series Experiments*, Boulder, CO: Colorado Associate University Press.
 McCleary, R. and Hay, R. A. Jr. (1980), *Applied Time Series Analysis for the Social Sciences*, Beverly Hills, CA: Sage.

The Causal Relationship between Dam Water Level and Fog Time

- A Transfer Function Analysis -

Don Jaegal²

Abstract As a result of the transfer function analysis on the casual relationship between Andong Dam water level and fog time in Andong City from March of 1981 to December of 1992, it should be referring to monthly increases in the water level having an effect on 0.89 hour per m increases in the fog time with 4 months delay. It is evident, hence, the artificial lake of the dam is one of the direct causes of fog. Based upon this objective evidence, water resources development policy and dam managment policy must be formulated.

² Andong National University