

측우기 자료를 포함한 서울 강수량 시계열에 대한 추세분석 및 파엽분석

정현숙¹⁾ 박정수²⁾ 임규호³⁾ 오재호⁴⁾

요약

본 논문에서는 측우기 관측자료와 현대 관측자료로 이루어진 서울 강수량 시계열 자료에 나타난 시간에 따른(년도별 및 계절별) 변화 및 변동 특성을 파악하고자 한다. 이를 위하여 먼저 200여년 간의 강수량 시계열에 어떤 특정한 증감 추세가 있는지를 알아보는 추세분석을 실시하였다. 그리고 추세뿐만 아니라 시간에 따른 강수량의 주기성 및 변동성을 더 자세히 알기 위하여 파엽 변환(wavelet transform)을 실시하여 여러 진동 모드들의 시간에 따른 변화 양상을 분석하였다.

주요용어: 기후변화, 시간변동성, 스피어만의 상관계수, Brillinger의 검정법, Scalogram.

1. 서론

우리 나라의 측우기는 1441년(세종 23년)에 발명된 우량계로서 현대적인 우량계가 도입되기 전인 1907년까지 정량적인 강수량 관측에 사용되었다. 측우기의 관측단위는 寸(치), 分(푼) 이었는데, 정현숙(1999)은 승정원 일기의 측우기 기록을 바탕으로, 이를 미터법으로 단위 환산하여 1777년부터 1907년까지의 자료를 얻었다. 세계 다른 지역에서 강수량의 장기 변동성 연구에 기후 대용지수(proxy climate indices)를 사용한다는 점을 생각하면 서울 지역의 강수량 자료는 한 장소에서 우량계로 관측되어 보존되어 있는 세계 최장기간의 관측 자료로서 과거 한반도의 강수 기후와 한반도에서의 장기적인 강수 변동성을 밝혀 줄 수 있는 귀중한 자료이다. 측우기 자료에 관한 자세한 서술은 조신섭 등(1996), 정현숙, 임규호(1994) 및 정현숙(1999) 등에서 찾아 볼 수 있다.

조신섭 등(1996)은 측우기 자료에 대한 Wada(1917)의 월별계수에 의한 경정법에는 일관성이 없음을 지적하고, 개입(intervention) 모형에 기초한 새로운 경정방법을 제안하였다. 한편 정현숙, 임규호(1994) 및 정현숙(1999)은 승정원 일기로부터 새롭게 측우기 자료를 복원하였으며, 경정되지 않은 원래의 측우기 자료의 신뢰성을 강조하였다.

본 연구에서 사용하는 측우기 관측자료는 1777년-1907년 자료로서 1771년-1776년의 강우기록은 아직 복원되지 못했다. 그런데 1777년-1782년 자료는 강수량의 계속적인 증가추세

1) 기상청 기상연구소 예보연구실, 연구원. E-mail: hsjung@metri.re.kr

2) (500-757) 광주광역시 북구 용봉동 300, 전남대학교 정보통신연구소 및 자연과학대학 통계학과, 부교수.
E-mail: jspark@chonnam.ac.kr.

박정수의 연구는 1999년도 한국대학교육협의회 대학교수 국내교류 연구비 지원에 의한 것이다.

3) 서울대학교 지구환경과학부 대기과학전공, 부교수.

4) 기상청 기상연구소 예보연구실장.

를 보이고 있고 약간의 신뢰도 문제 등으로 추세분석에서는 사용하지 않았다. 또 현대 관측 자료는 1908년-1999년의 기상청 자료이며, 그 중 1950년-1952년의 결측치는 이명인(1996)의 추정값을 이용하였다.

본 연구에서는 측우기 관측자료와 현대 관측자료로 이루어진 서울 강수량 시계열 자료에 나타난 시간에 따른(년도별 및 계절별) 변화 및 변동 특성을 파악하고자 한다. 이를 위하여 먼저 216년간의 강수량 시계열에 어떤 특정한 증감 추세가 있는지를 알아보는 추세분석을 실시하였다. 그리고 추세뿐만 아니라 시간-진동수 영역에서의 강수량 변동성을 더 자세히 알기 위하여 파엽 변환(wavelet transform)을 실시하여 여러 진동 모드들의 시간에 따른 변화 양상을 분석하였다.

2. 시계열에 대한 추세분석

2.1. 추세검정을 위한 통계적 방법

기상 자료의 시계열에 대한 추세분석은 시계열 자료에 “통계적으로 유의한” 어떤 경향이 있는가를 알아보는 문제로 접근 가능하다. 예를 들어 지구 연평균 기온의 과거 100년간의 시계열 자료로부터 통계적으로 유의한 온난화의 여부를 탐지하는 문제이다. 즉 현재까지 관측되는 기온 상승이 (또는 기후 변화가) 단지 기후의 자연적 변동성에 의한 것인지 아니면 어떤 외부적 요인에 의한 통계적 의미에서의 특별한 것인지를 판별해내고자 하는 것이다. 이를 위한 통계적 방법으로는 상관계수에 대한 가설 검정과 “결정적 추세” 검정을 위한 회귀분석 방법과 및 비모수적 방법으로서 Brillinger의 검정법, 세 가지를 이용하였다.

먼저 다음과 같은 선형 모형을 생각해 보자.

$$Y_t = a + bt + E_t. \quad (2.1)$$

여기서 Y_t 는 시간 t 에서의 관측치이고, E_t 는 회귀직선으로 설명되지 않는 부분을 나타내는데 보통 평균 0인 정상과정(stationary process)이라고 가정한다. 여기서 귀무가설 $H_0: b = 0$ 에 대한 가설 검정을 하여 어떤 유의수준에서 귀무가설이 기각되면 우리는 “결정적” 선형추세가 있다고 말할 수 있다. 이때 시간은 연도를 나타내며 시작하는 년도부터 $t = 1, 2, 3, \dots, n$ 으로 표기한다.

식 (2.1)을 이용하여 추세검정을 하기 위하여, b 의 추정치 \hat{b} 와 \hat{b} 의 표준오차 추정치 $SE(\hat{b})$ 를 구해서, $T = \hat{b}/SE(\hat{b})$ 를 검정 통계량으로 이용한다. 이때 \hat{b} 와 $SE(\hat{b})$ 를 어떻게 구하느냐의 문제가 생기는데 이것을 구하는 방법에 따라 검정통계량이 달라진다. 만약 E_t 가 각 t 에 대해서 독립이면 T 는 보통의 회귀분석에서의 회귀계수의 검정을 위한 통계량이 된다. 만약 E_t 가 독립이 아니고 어떤 상관관계를 가지면, 즉 예를들어 AR(p)를 따른다면, 거기에 적합한 $SE(\hat{b})$ 를 구하여 사용해야 할 것이다. 그러나 본 연구에서 구성한 서울 강수량 시계열에 대한 독립성 검정을 실시한 결과 모두 독립적이라고 판정이 되어, 단순 회귀분석 방법을 적용하였다.

또한 이러한 독립성에 힘입어 시간 t 와 시계열 Y_t 간의 스피어만(Spearman) 상관계수 ρ 에 대한 가설검정 $H_0: \rho = 0$ 를 실시하여 추세여부를 판정하였다. 또한 Y_t 가 독립적이 아

년 시계열이라면 단위근 검정을 통하여 “확률적 추세” 여부를 판정할 수 있겠으나, 본 연구에서 구성한 시계열들은 모두 독립적이라고 판정되어 단위근 검정을 하지 않았다.

한편 Brillinger(1989)는 μ_t 의 단조성에 대한 비모수적 검정법을 제안하고 있다. 단조성이란 모든 t 에 대해서 $\mu_t \leq \mu_{t+1}$ 또는 $\mu_t \geq \mu_{t+1}$ 가 만족되며 특히 몇 개의 t 에 대해서는 부등호(strict inequality)가 성립되는 경우이다. Brillinger는 단조성과 관련된 추정치로서 $\sum_{t=1}^n c_t Y_t$ 를 이용하는데, 여기서 계수 c_t 는 다음과 같은 Abelson and Tukey(1963)의 공식을 이용한다:

$$c_t = \left[(t-1) \left(1 - \frac{t-1}{n}\right) \right]^{1/2} - \left[t \left(1 - \frac{t}{n}\right) \right]^{1/2}. \quad (2.2)$$

이 값 $\sum_{t=1}^n c_t Y_t$ 는 시계열의 시작과 끝을 대비시키는 값으로서 시계열이 단조 증가이면 이 값은 양수이고 단조 감소이면 이 값은 음수가 된다. 가설 $H_0 : \mu_t = \mu$ 에 대한 Brillinger의 검정통계량은

$$T_{Br} = \frac{\sum_{t=1}^n c_t Y_t}{\left[\sum_{t=1}^n c_t^2 \text{Var}(Y_t) \right]^{1/2}} \quad (2.3)$$

으로서 근사적으로 표준 정규분포를 따른다. 여기서 Y_t 가 독립적이며 같은 분포를 가지면 $\text{Var}(Y_t)$ 대신에 $\hat{\sigma}^2$ 를 이용할 수 있겠다. 한편 Brillinger 방법을 사용했을 때, 강수량에 있어서 증가추세가 탐지되었다면 그 증가량이 얼마인가를 추정하기 위해서 다음과 같은 이동평균을 사용한다. 즉 $t = v+1, \dots, n-v$ 에 대하여

$$\hat{\mu}_t = \frac{1}{2v+1} \sum_{s=-v}^v Y_{t+s} \quad (2.4)$$

를 이용하는데, 증가량은 $\hat{\mu}_{n-v} - \hat{\mu}_{v+1}$ 가 된다. 이때 v 의 선택이 문제인데, 우리는 교차검증량(leave-one-out cross-validation)

$$CV(v) = \frac{1}{n} \sum_{t=v+1}^{n-v} (\hat{\mu}_{-t} - \mu_t)^2 \quad (2.5)$$

를 최소로 하는 v 를 선정하였다. 여기서 $\hat{\mu}_{-t}$ 는 t 번째 관측치를 빼고 나머지만 가지고 t 번째 관측치에 대해 이동평균을 구한 값이다.

2.2. 서울 강수량 시계열의 추세

우리는 1783년 1월부터 1999년 2월까지의 월 강수량 자료에 대해 다음과 같이 2개월씩 이동평균을 구하여 시계열을 만들었다. 즉 (1,2)월 평균 강수량으로 이루어진 시계열, (2,3)월 평균 강수량으로 이루어진 시계열, ..., (12,1)월 평균 강수량으로 이루어진 시계열로서 모두 12개의 시계열을 만들었다(따라서 각 시계열에서 n 은 216이다). 이렇게 함으로써 계절적 특성을 파악할 수 있을 뿐만 아니라, 측우기 자료가 가지는 다음과 같은 오차를 극복할 수 있다. 측우기 자료는 강설량 및 2mm 이하의 강수량은 관측되지 못했으므로 우

리는 겨울의 강수량과 관련된 시계열, 즉 (1,2)월, (2,3)월, (11,12)월, (12,1)월 시계열을 본 추세분석 연구에서 일단 제외시켰다. 따라서 이제 (3,4)월, ..., (10,11)월의 8개 시계열에 대해서만 2.1절에서 기술한 세가지 방법을 적용하여 추세분석을 한다.

먼저 이들에 대해 SAS/ETS 의 AUTOREG 프로시저를 이용하여 자기상관 3차까지의 일반화된 Durbin-Watson 검정(조신섭, 이정형[1997])을 실시하였다. 그 결과 자기상관이 존재하지 않은 것으로 판정되어, 모두 독립적인 시계열로 간주되었다. 이는 예를 들어 올해 봄의 강수량이 내년 봄이나 내후년 봄의 강수량의 다과에 어떤 영향을 미치지 않을 것으로 판단되는 기상학적인 의미와 일치한다.

표 2.1은 스피어만의 상관계수, 회귀분석 및 Brillinger 방법을 적용한 결과이다. Brillinger 방법에서 $\hat{\sigma}^2$ 으로 S^2 을 이용하였고, c_t 값의 특성상 $\sum_{t=1}^n c_t Y_t$ 를 계산하는데 있어서 시계열의 처음과 끝의 극단적 영향력을 줄이기 위하여 양쪽 끝에 각각 3개씩의 영 값을 주고 계산하였다. 회귀분석에서 예를 들어 (3,4)월 및 (4,5)월의 증가량의 계산은 회귀계수 곱하기 200으로 구해졌다. 여기서 200을 곱한 이유는 Brillinger 방법에서 (3,4)월 및 (4,5)월에는 15년 이동평균($v=7$)을 사용했기 때문에 216년에서 15를 빼면 201년인데, 실제로 변화된 년수는 200이기 때문이다.

세가지 방법 모두에서 (3,4)월 및 (4,5)월에 5% 유의수준에서 증가추세를 판정하였다. (7,8)월 및 (8,9)월의 증가량이 상당한데도 추세가 있다고 판정하지 않은 이유는 연도별 표준편차가 상대적으로 더 크기 때문이며, 강수량의 자연변동성 내에서의 증가량으로 해석할 수 있다.

그런데 이 결과는 2mm 이하를 관측하지 않은 측우기 자료에 대한 것이므로, 200여년 간의 강수량 증가추세가 2mm 이하의 강수현상 때문인지 아니면 정말 기후변화에 의한 것인지 구분을 할 수 없게 된다. 이 문제를 해결하기 위해, 1908년부터 1998년까지 88년간(1950-1952년 제외) 해당 달에 2mm 이하 강수현상의 분포를 살펴본 결과 표 2.2와 같았다. 특히 (3,4)월, (4,5)월의 평균을 측우기 자료에 일괄적으로 더해서 새로운 시계열을 만든 뒤, 이들에 대해서 위에서와 같은 세가지 방법으로 증가추세에 대한 가설 검정을 실시하였다. (여기서 1908년 이후의 일 강수량에서 2mm 이하인 경우를 빼고 분석하는 것도 생각할 수 있다. 그러나 이 경우 1908년 이전과 이후 강수량 모두가 실제 강수량이 아니라는 문제가 생긴다. 따라서 우리는 적어도 한쪽이라도 실제강수량을 이용한다는 의미에서 위와같이 1908년 이전의 강수량에 얼마간의 평균을 일괄적으로 더했다.) 그 결과 세가지 방법 모두에서 (3,4)월 및 (4,5)월에 5정도였고 (4,5)월의 증가량은 14mm에서 16mm 정도였다. 즉 봄철에 약간의 비가 더 내리는 것으로 해석할 수 있다. 그러나 월 평균 강수량에서 약 8mm에서 16mm 정도의 (200년 동안의)증가량은 비록 통계적으로 유의한 증가라고 하더라도 기상학적으로 큰 의미를 갖지는 않은 것으로 보인다. 한편 3개월 단위의 이동평균에 대해서도 같은 분석을 해 본 결과, (3,4,5)월과 (4,5,6)월의 증가량이 5% 유의수준에서 유의하였다. 그 증가량은 월 평균 각각 8.4mm와 4.8mm였다 (구체적 내용 생략).

표 2.1: 측우기 자료를 포함한 서울 강수량 자료에 대한 2개월 이동평균에 대한 세가지 방법(스피어만 상관계수, 회귀분석, Brillinger 검정)에 의한 추세분석 결과 (* : 유의확률 < 0.05).

2개월 시작달	통계량 중위수	표준편차	스피어만 상관계수	회귀계수 및 월 평균 증가량	Brillinger 방법 월 평균 증가량
3	48.98	31.21	0.181 *	0.0819 (16.36) *	10.11 *
4	72.18	40.45	0.124 *	0.0836 (16.73) *	18.03 *
5	97.68	59.49	0.068	0.0822 (16.44)	23.38
6	232.70	115.20	0.005	0.0127 (2.54)	-3.82
7	277.00	153.21	0.045	0.0211 (4.22)	36.13
8	183.52	107.99	0.035	0.0165 (3.3)	48.42
9	78.07	59.22	0.072	0.0515 (10.3)	11.36
10	39.94	25.49	0.410	0.0278 (5.56)	10.34

표 2.2: 1908년부터 1998년 사이의 일 강수량이 2mm 이하인 강수현상의 월 평균 강수량 및 강수일수.

통계량 월	3	4	5	6	7	8	9	10	11
강수량 평균	2.35	2.16	2.21	2.49	3.26	2.66	2.18	1.89	2.98
강수일수 평균	3.54	3.00	3.40	3.63	4.38	3.85	2.99	2.67	4.05

표 2.3: 측우기 자료에 강수량 2mm 이하의 강수현상을 고려하기 위해 평균을 더한 뒤의 증가추세에 대한 가설검정 결과 (* : 유의확률 < 0.05).

통계량 시작달	스피어만 상관계수	회귀계수 및 증가량	Brillinger 증가량(월 평균)
3	0.148 *	0.0666 (13.32) *	7.86 *
4	0.094 *	0.0688 (13.76) *	15.85 *

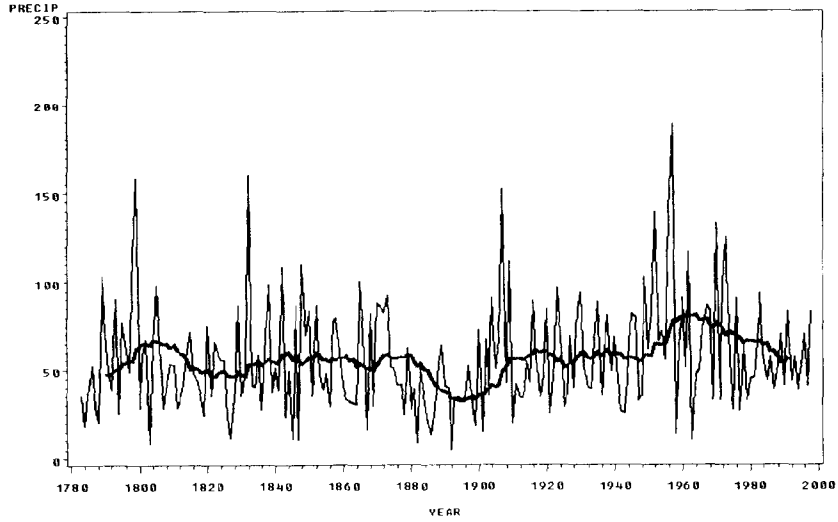


그림 2.1: (3,4)월 평균 강수량의 시계열과 이동평균($v=7$ 을 이용한 식 (2.6)). 측우기 자료에는 2mm 이하의 강수현상을 고려하여 평균을 더한 것임.

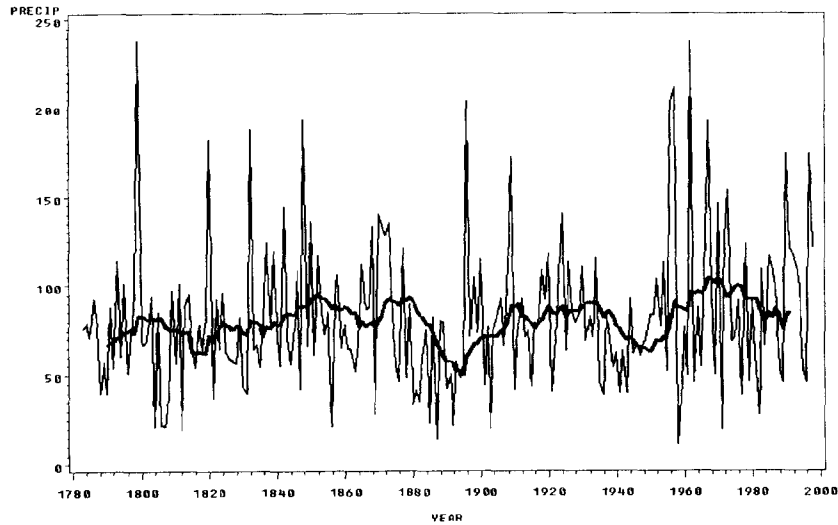


그림 2.2: (4,5)월 평균 강수량의 시계열과 이동평균($v=7$ 을 이용한 식 (2.6)). 측우기 자료에는 2mm 이하의 강수 현상을 고려하여 평균을 더한 것임.

3. 파엽 분석

3.1. 파엽 변환

파엽 변환(wavelet transform)은 비교적 최근에 개발되어 영상처리에 주로 사용되어 왔으며, 통계학에서도 최근 이에 대한 상당한 연구가 진행되고 있다(Antoniadis and Oppenheim [1995], Ogden [1997]). 환경과학 자료에 적용된 최근의 연구는 예를 들어 Percival and Mofjeld(1997)를 들 수 있다. 파엽 변환은 시계열 자료의 분석에도 자주 적용되는데(Priestley, 1996), 그 이유는 고주파 영역에서는 창이 좁아지고 저주파 영역에서는 창이 넓어지게 되어 진동수에 따라 탄력적인 창(variable window width)을 가짐으로써 푸리에 변환(FT)을 이용한 스펙트럼 분석의 결점을 보완할 수 있어서, 특히 제한된 공간과 시간 영역에서 다중 규모적이고 동적인 특성을 가지는 현상들을 연구하기에 적합하기 때문이다(Lau and Weng, 1995).

파엽 변환은 파엽(wavelet)이라는 일반화된 일시적 기본함수와 “늘림 모수(a, dilation parameter)” 및 “병진 모수(b, translation parameter)”를 연속적으로 사용한다(식 3.1).

$$\Psi_{b,a}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right). \tag{3.1}$$

여기서 Ψ 는 모 파엽(mother wavelet)이고 $\Psi_{b,a}$ 는 파엽이다. 시계열 $f(t)$ 의 파엽 변환은 $f(t)$ 와 파엽 사이의 곱말림(convolution) 적분으로 정의되며, 따라서 파엽 계수(wavelet coefficient)는 다음과 같이 표현된다.

$$W(b, a) = \int_{-\infty}^{\infty} f(t)\Psi_{b,a}(t) dt = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) dt. \tag{3.2}$$

여기서 $\Psi(t)$ 가 복소함수이면 $\Psi_{b,a}(t)$ 대신에 복소공액 $\Psi_{b,a}^*(t)$ 를 사용한다. 또한 a 와 b 는 실수이고 a 는 양의 값을 가진다. 관측 자료수가 많은 경우 식 (3.2)를 사용하여 $0 < a \leq I$, $0 < b < J$ 에 대해 수치 적분을 하면 되지만 많은 계산 시간을 필요로 한다($O(IJ^2)$). 한편 식 (3.3)과 같이 $f(t)$, $\Psi(t)$ 의 FT 형태와 FFT(Fast Fourier Transform)를 사용하여 진동수 영역에서 계산하면 계산 시간을 $O(IJ \log_2 J)$ 로 줄일 수 있다는 장점을 갖는다.

$$W(b, a) = a^{1/2} \int_{-\infty}^{\infty} \hat{f}(\omega)\hat{\Psi}(a\omega) e^{-ib\omega} d\omega. \tag{3.3}$$

여기서 \hat{f} 및 $\hat{\Psi}$ 은 각각 f 와 Ψ 의 푸리에 변환함수이다.

$\Psi_{b,a}(t)$ 이 몇 가지 기본 조건을 만족하면 우리는 원래의 시계열을 역으로 다시 얻을 수 있게 된다(Chan, 1995). 즉

$$f(t) = \frac{1}{C_{\Psi}} \int \frac{1}{a^2} \int \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) W(b, a) db da \tag{3.4}$$

이다. 여기서

$$C_{\Psi} = \int_0^{\infty} \frac{|\hat{\Psi}(\omega)|^2}{\omega} d\omega < +\infty \tag{3.5}$$

표 3.1: 자료에 적절한 모 파엽을 선택하기 위해 고려하는 모 파엽들의 종류: 이름과 파엽 식 및 특성.

이름	$\Psi(t)$	특성
Morlet ($\omega_0 =$ 진동수)	$\pi^{-1/4} \exp(-i\omega_0 t - e^{-\omega_0^2 t^2/2}) e^{-t^2/2}$	복소수
Paul ($m =$ 차수)	$\frac{2^m i^m m!}{\sqrt{\pi(2m)!}} (1 - it)^{-(m+1)}$	복소수
DOG ($m =$ 미분 차수)	$\frac{(-1)^{m+1}}{\sqrt{\Gamma(m+\frac{1}{2})}} \frac{d^m}{dt^m} (e^{-t^2/2})$	실수

이다. 파엽 변환은 그 과정에서 FT처럼 에너지가 보존되므로(식 3.6) 지역 파엽 스펙트럼 에너지 $E(a, b)$ 를 식 (3.7)과 같이 표현할 수 있다.

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{C_{\Psi}} \int_0^{+\infty} \frac{1}{a^2} \int_{-\infty}^{\infty} |W(b, a)|^2 db da, \quad (3.6)$$

$$E(a, b) = \frac{1}{C_{\Psi}} \frac{|W(b, a)|^2}{a^2}. \quad (3.7)$$

통계학 문헌에서는 이 에너지를 시간 축과 주기 축 상에 등고선도(contour)로 그린 그림을 scalogram 이라고 부르는데(Ogden[1997]), 이 그림을 통하여 우리는 단순한 시계열 그림이나 FT를 통한 스펙트럼 그림에서 알기 어려운 주기적 변동성을 파악할 수 있다(그림 3.2부터 그림 3.5까지 참조).

3.2. 모 파엽의 선택과 실제 계산

여러 종류의 모 파엽들(mother wavelets) 중에서 본 연구의 목적에 가장 적합한 것을 선택하기 위해 (통계학자들에게는 잘 알려져 있지는 않지만) 기상 자료 분석에 유용한 것으로 알려진 세 종류의 모 파엽(표 3.1)을 실제 관측 자료에 적용하여 각각의 장·단점을 알아보고, 본 연구의 목적에 부합되는 하나의 모 파엽을 선택하였다.

DOG(Derivatives Of Gaussian)에서 $m = 2$ 일 때를 Mexican hat 파엽(그림 3.1c)이라고 하는데, 이는 시간 영역에서의 분해능이 뛰어나지만, 진동수 영역상의 분해능은 좋지 못하고, 일강수량 자료를 사용한 경우 10일보다 짧은 진동 모드는 분해하지 못하는 단점이 있다. Morlet 파엽(그림 3.1a)은 시간 영역에서의 분해능은 조금 낮지만, 고주파에 해당하는 10-60일 사이 진동수 영역에서의 분해능은 뛰어나다. 한편, Paul 파엽(그림 3.1b)은 시간 및 진동수 영역에서 모두 Mexican hat 파엽과 Morlet 파엽의 중간 정도의 분해능을 가진다. 이 결과는 다른 연도의 관측 자료뿐만 아니라 임의로 합성한 자료에서도 동일하게 나타남으로써 스펙트럼 연구에는 Morlet 파엽이 좋고, 시계열의 시간 영역에서 경향을 필터하기 위해서는 Mexican hat 파엽이 유용하다는 Baliunas 등(1997)의 의견과 일치하였다.

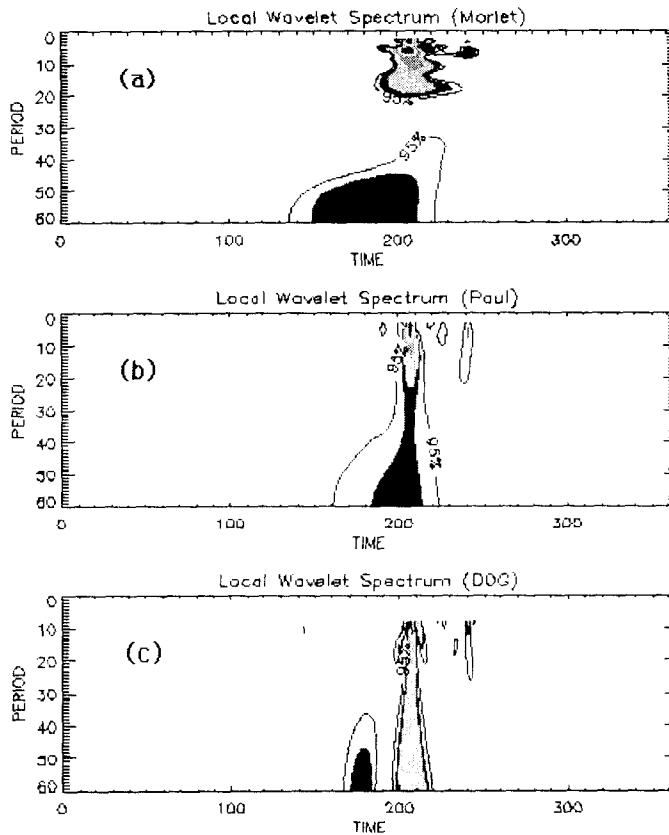


그림 3.1: 1987년 일강수량 자료에 대한 (a) Morlet 파엽, (b) Paul 파엽, (c) Mexican hat 파엽을 사용한 지역 파엽 스펙트럼. 가로 축은 시간을 나타내고 세로 축은 주기를 나타내며, 그림의 두꺼운 등고선은 95% 신뢰영역을 나타낸다.

본 연구의 목적은 강수 예측이 아니라 이미 발생한 강수 현상에서 나타나는 주요 진동 모드들을 조사하고 이들의 시간 변동성 및 관련된 대기 현상을 알아보는 것이므로, 본 논문에서는 모 파엽으로 진동수 영역에서의 분해능이 뛰어난 Morlet 파엽을 사용하였다. 이를 위해서 Gu and Philander(1995)의 알고리즘과 포트란 프로그램을 이용하였으며, 결과 그림은 GrADS(Grid Analysis and Display System, <http://grads.iges.org/grads/> 참조)을 이용하여 그렸다.

표 3.1의 Morlet 파엽은 허용 조건을 만족하기 위해서 $\omega \geq 5$ 이어야 하는데, 이 경우에 다음과 같이 근사된다.

$$\Psi(t) = \pi^{-1/4} \exp(-i\omega_0 t - t^2/2) . \tag{3.8}$$

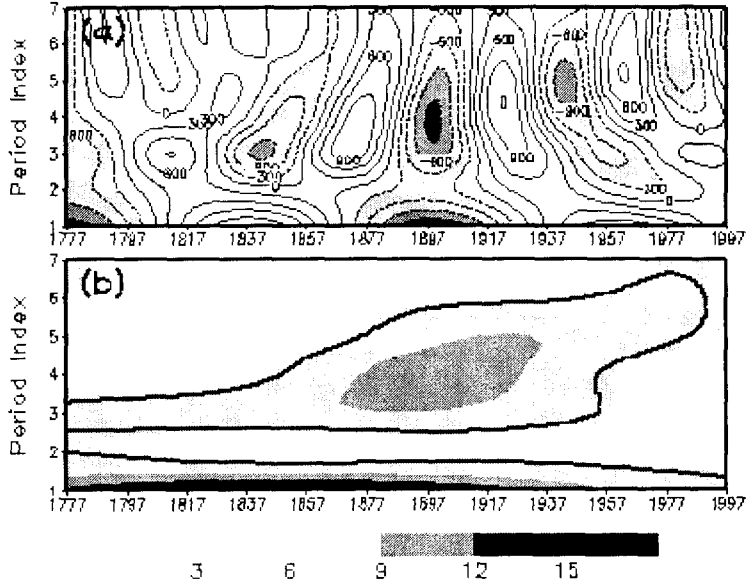


그림 3.2: Morlet 파엽을 이용한 측우기를 포함한 서울 강수량 자료의 파엽 변환(30년-120년 진동 모드). (a) 파엽 계수의 실수부, (b) 표준화된 지역 파엽 스펙트럼. 주기지수 1,2,..., 6,7은 각각 주기 120년, 80년, 60년, 48년, 40년, 34년, 30년을 나타낸다.

실제로 본 연구에서는 식 (3.8)의 한 형태인 식 (3.9)를 사용하였다.

$$\Psi(t) = e^{-i2\pi t} \exp \left[-\left(\frac{2\pi}{k_\Psi} \right)^2 |t|^2 / 2 \right]. \quad (3.9)$$

여기서, k_Ψ 는 모 파엽의 가우스 외피(envelope)의 넓이를 정의하는 값으로서, 분석 중에는 일정한 상수 값으로 고정된다. $|k_\Psi|$ 가 증가하면, 파엽 계수의 시간 축에서의 분해능이 감소하게 되므로, 파엽으로서의 허용 조건을 만족하는 범위 내에서는 될 수 있으면 작은 값을 취하는 것이 좋다. 본 연구에서는 6.0을 사용하였다.

또한 파엽 스펙트럼의 신뢰구간은 Torrence and Compo(1998)가 몬테카를로 방법을 이용하여 제안한 분포를 이용하였다. 즉, 주어진 s 시간, n 진동수에서 시계열의 분산(σ^2)로 표준화한 지역 파엽 스펙트럼 $|W_n(s)|^2$ 은 식 (3.10)과 같은 카이 제곱 분포를 가진다.

$$\frac{|W_n(s)|^2}{\sigma^2} \Rightarrow \frac{1}{2} P_k \chi_2^2. \quad (3.10)$$

여기서 P_k 는 진동수 k 에서의 푸리에 스펙트럼의 평균을 의미한다. 그러면 지역 파엽 스펙트럼의 $100 \times (1 - \alpha)\%$ 신뢰구간은 다음과 같이 구해진다.

$$\frac{2}{\chi_2^2(\alpha/2)} |W_n(s)|^2 \leq \varpi_n^2(s) \leq \frac{2}{\chi_2^2(1 - \alpha/2)} |W_n(s)|^2. \quad (3.11)$$

본 연구에서와 같이 일정한 시간 간격으로 관측된 시계열 자료를 분석할 때는 식 (3.1)의 a와 b는 불연속적인 값을 가지게 된다. 시계열 자료의 양쪽 끝 부분에서 발생하는 가장자리 효과를 줄여주기 위해서는 실제 관측 자료를 덧붙여(padding) 사용하였다. 예를 들면 1990년을 분석하기 위해서는 1989년-1991년 자료를 입력자료로 사용하였다. 결측인 해는 결측 전후의 자료를 이용하여 시간 영역에서 선형 내삽하였다.

3.3. 파엽 분석의 결과: 기상학적 해석

그림 3.2는 220년 동안의 서울 지역 연강수량 자료로 파엽 변환을 실시한 결과로 30년 이상의 장주기 진동 모드에 대해 보여준다. 그림 3.2a는 식 (3.2)를 이용하여 구한 파엽 계수의 실수 부분으로 파엽 계수는 특정한 시간-진동수에서 파엽으로 설명할 수 있는 분산에 비례한다. 그림 3.2b는 시계열의 분산으로 표준화된 지역 파엽 스펙트럼을 보여준다(식 3.7과 식 3.10). 굵은 실선은 식 (3.11)을 이용하여 계산한 95% 이상의 신뢰도를 갖는 부분이다. 또 세로축의 주기 지수 1, 2, ..., 7은 식 (3.1)의 a 값에 역비례 하는 값으로서 각각 120년, 80년, 60년, 48년, 40년, 34년, 30년 주기를 나타낸다. 이 그림에서 1900년을 전후한 건초기 동안 50-60년 주기에서 아주 큰 분산이 나타나고 있으며, 후기 조선시대에서 현대로 오면서 분산으로 표현되는 에너지가 점점 짧은 주기로 이동하는 것을 볼 수 있다.

다음은 131년간의 측우기 관측과 89년간의 현대 관측 일강수량 자료에 파엽 변환을 적용하여 2-20년의 장주기 및 경년 변동을 살펴보았다(그림 3.3과 그림 3.4). 각각 적용한 이유는 자료 개수가 늘어나면 계산에 필요한 시간이 급격히 늘어나는 문제점이 있기 때문이다. 주기 지수 1, 2, ..., 7, 8은 각각 23.3년, 11.6년, 7.8년, 5.8년, 4.9년, 3.7년, 3.0년, 2.5년의 주기를 나타낸다. 20년 정도의 장주기 성분은 1800년 이전과 1900-1920년과, 1960년대 후반부터 1980년대 후반 사이를 제외하면 아주 약하게 나타난다. 흑점활동과 관련이 있는 11.6년 주기는 분석 전 구간에서 통계적으로 유의하게 나타나지 않으므로 서울 지역 강수량 변동과 흑점 활동을 관련짓기는 어렵다고 판단된다. 한편 엘니뇨 변동 시간 규모(2-7년)에서는 5.8년 주기가 가장 우세하지만 시간에 따른 변동성이 매우 커서 서울 지역 강수량 변동과 엘니뇨의 상관 관계를 설명하는 것은 어려울 것으로 사료된다.

측우기 자료 중 여름철(6-9월) 강수량이 평균보다 많았던 해를 여름 몬순이 강했던 해로 정의하고, 파엽 변환을 이용하여 30-60일 주기의 계절안 진동(MJO)과 이보다는 짧고 저·고기압 등의 종관규모 진동보다는 긴 주기(SISO: Short-term Intraseasonal Oscillation)의 시간 변동성을 조사하였다. 1833년을 예로 보면(그림 3.5의 왼쪽) 그림에서 상단은 일강수량을, 중간은 파엽 변환의 실수부를, 하단은 분산으로 표준화된 지역 파엽 스펙트럼(식 3.10)을 보여 준다. 가로축은 시간(day)을, 세로축은 주기 지수로서 1, 2, 3, ..., 14는 각각 60일, 52일, 46일, 41일, 37일, 34일, 31일, 25일, 21일, 18일, 16일, 15일, 14일, 12일 주기에 해당한다. 여름철 동안 모든 계절안 진동 모드에서 분산이 강하게 나타나고, 특히 일강수량이 300mm에 달하는 집중호우와 관련하여 20-30일 주기와 15일 이하의 짧은 주기에서 분산이 강화되었다.

측우기 관측에 나타난 20-30일 주기와 관련된 기상 현상을 알아보기 위해 현대 자료에서 20-30일 주기가 강하게 나타났을 때의 기상 상태를 알아보았다. 20-25일 모드와 관련된

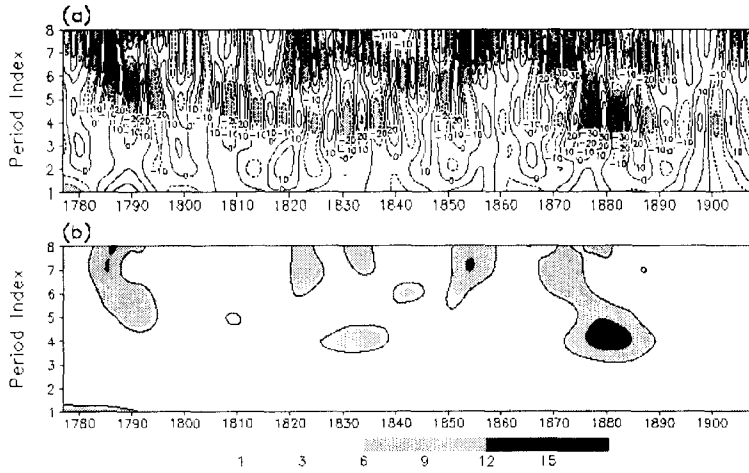


그림 3.3: Morlet 파업을 이용한 측우기 자료의 파업 변환 (2년-23년 진동 모드). (a) 파업 계수의 실수부. (b) 표준화된 지역 파업 스펙트럼. 주기지수 1, 2, ..., 7, 8은 각각 주기 23.3년, 11.6년, 7.8년, 5.8년, 4.9년, 3.7년, 3.0년, 2.5년에 해당된다.

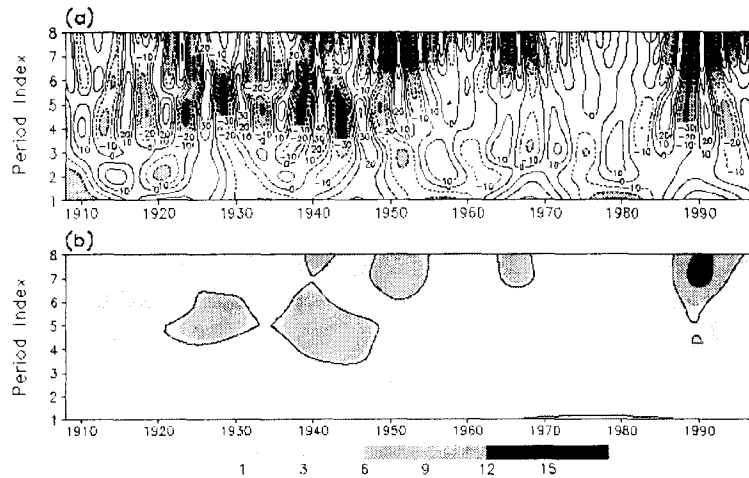


그림 3.4: Morlet 파업을 이용한 현대 우량계 자료의 파업 변환 (2년-23년 진동 모드). 나머지는 그림 3.3과 같다.

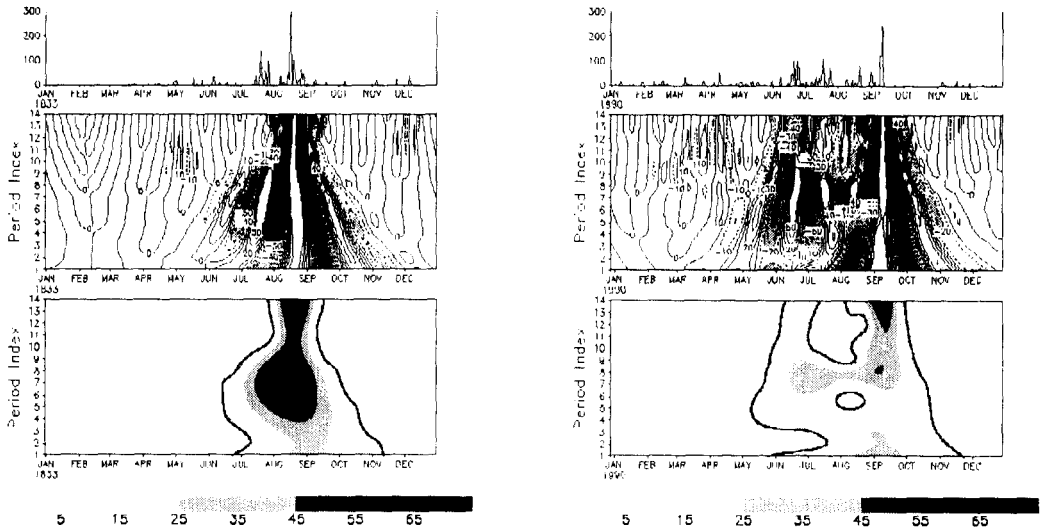


그림 3.5: 1833년(왼쪽)과 1990년(오른쪽) 일강수량 자료의 파엽 변환(12일-60일 진동 모드). 맨 위는 각 해의 일강수량 시계열이고, 중간은 Morlet 파엽 변환에 의한 파엽계수의 실수부이며, 맨 아래는 표준화된 지역 파엽 스펙트럼이다. 주기지수 1,2, ...,13,14는 각각 주기 60일, 52일, 46일,41일, 37일, 34일, 31일, 25일, 21일, 18일, 16일, 15일, 14일, 12일을 나타낸다.

여 태풍의 영향을 받았을 때 SISO가 강화되는 것을 볼 수 있다(1990년, 그림 3.5의 오른쪽). 그러나 태풍의 영향이 적은 서울의 경우 태풍만으로 20-30일 주기에 나타나는 스펙트럼을 충분히 설명할 수는 없으며 실제로 태풍의 영향이 아니더라도 장마전선이나 강한 기압골이 통과할 때 SISO가 강화되었다. SISO가 강한 기간에 현대 자료뿐만 아니라 축우기 자료에서도 이동성 고·저기압과 관련된 종관규모 진동 모드가 강해지는 것을 볼 수 있었다(그림은 생략).

4. 결론

축우기 자료를 포함한 약 216년간의 서울 강수량 자료에 대해 통계적 방법을 적용한 추세분석 결과, (3,4)월 및 (4,5)월에 5% 유의수준에서 증가추세를 보였고, 나머지 달에서는 증가추세를 보이지 않았다. 약 200년 동안의 강수량의 증가량은 (3,4)월에는 월 평균 약 8mm에서 13mm 정도, (4,5)월에는 약 14mm에서 16mm 정도였다.

또한 시간-진동수 영역에서의 강수량 변동성을 더 자세히 알기 위하여 파엽 변환을 실시한 결과, 여름 몬순의 장주기 변동과 관련된 20년 이상의 진동은 후기 조선시대에서 현대로 오면서 점점 짧은 주기에서 큰 분산이 나타났다. 흑점활동과 관련이 있는 11.6년 주기는 거의 대부분 구간에서 약하게 나타나 서울 지역 강수량과 흑점활동을 관련짓기는 어려

을 것으로 판단된다. 30-60일 주기의 계절안 진동(MJO)은 거의 모든 해에 여름 몬순 기간 동안 나타났으며, 20-30일 진동은 MJO와 구별되는 모드로서 큰 계절변화와 연변동을 가지고 태풍 및 강한 저기압의 통과와 관련되어 집중호우 때 뚜렷하게 나타났다.

측우기 강수량 자료에서 나타난 장주기 변동에서 종관규모 진동까지 모두 현대 자료에서 나타나고 있는 점은 측우기 자료의 기상학적인 가치를 높여 주었다. 서울지역 및 한반도의 강수 기후 변화의 이해는 물론 강수 예측 성능을 향상시키고 나아가서는 수자원을 효율적으로 관리하기 위하여, 측우기 자료라는 세계적으로 보기 드물게 귀중한 자료에 대한 기상통계학적 연구가 지속되어야 한다.

감사의 글

이 연구는 과학기술부 지원으로 수행하는 중점국가연구개발사업 (온실가스저감기술연구)의 일환으로 수행되었습니다.

참고문헌

- [1] 이명인 (1996). <한반도의 도시화에 의한 기온상승>, 석사학위논문, 서울대학교 대기과학과, 서울.
- [2] 정현숙, 임규호 (1994). 서울 지역 강수량과 강수일의 관계, 1770-1907, <한국기상학회지>, 제30권, 487-505.
- [3] 정현숙 (1999). <서울 지역 강수량 시계열에 나타난 시간 변동성 해석>, 박사학위논문, 서울대학교 대기과학과, 서울.
- [4] 조신섭, 이정형 (1997). <SAS/ETS를 이용한 경제시계열분석>, 자유아카데미, 서울.
- [5] 조신섭, 이정형, 김병수 (1996). 시계열 모형을 이용한 측우기 자료의 분석, <응용통계연구>, 제9권, 25-43.
- [6] 和田雄治 (1917). <朝鮮古代觀測記錄調查報告>, 日韓印刷株式會社.
- [7] Abelson, R. and Tukey, J. (1963). Efficient utilization of non-numerical information in quantitative analysis: General theory and the case of simple order, *Annals of Mathematical Statistics*, vol. 34, 1347-1369.
- [8] Antoniadis, A. and Oppenheim, G. (Eds.) (1995). *Wavelets and Statistics*, Springer, Berlin.
- [9] Baliunas, S., Frick, P., Sokoloff, D. and Soon, W. (1997). Time scales and trends in the Central England Temperature data(1659-1990) : A wavelet analysis, *Geophysical Research Letters*, vol. 24, 1351-54.

- [10] Brillinger, D. (1989). Consistent detection of a monotonic trend superposed on a stationary time series, *Biometrika*, vol. 76, 23-30.
- [11] Chan, T.T. (1995). *Wavelet Basics*, Kluwer Academic Publishers, New York.
- [12] Gu, D. and Philander, S.G.H. (1995). Secular changes of annual and interannual variability in the tropics during the past century, *Journal of Climate*, vol. 8, 864-876.
- [13] Lau, K.-M., and Weng, H.Y. (1995). Climate signal detection using wavelet transfer; How to make a time series sing, *Bulletin of American Meteorological Society*, vol. 76, 2391-2402.
- [14] Ogden, R.T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhauser, Boston.
- [15] Percival, D.B. and Mofjeld, H.O. (1997). Analysis of subtidal coastal sea level fluctuations using wavelets, *Jour. Amer. Stat. Assoc.*, vol. 92, 868-880.
- [16] Priestly, M.B. (1996). Wavelets and time-dependent spectral analysis, *Journal of Time Series Analysis*, vol. 17, 85-103.
- [17] Torrence, C. and Compo, G.P. (1998). A Practical Guide to Wavelet Analysis, *Bulletin of American Meteorological Society*, vol. 79, 61-78.

[1999년 12월 접수, 2000년 8월 채택]

Trend analysis and wavelet transform of time series of precipitation including the Chukwookee observations in Seoul *

Hyun-Sook Jung¹⁾ Jeong-Soo Park²⁾ Gyu-ho Lim³⁾ Jai-Ho Oh⁴⁾

ABSTRACT

Characteristics of precipitation in Seoul have been examined by using long-term observational data. Precipitation records from modern rain gauges were used for 1908-1996, together with the traditional Korean rain gauge (called Chukwookee) observations for 1777-1907. A linear trend analysis of seasonal total rainfall shows no significant trends over the last 200 years. A wavelet transform analysis was performed to figure out the transient variations of precipitation.

Keywords: Brillinger's test; Climatic change; Scalogram; Spearman's correlation coefficient; Transient variations.

* This research was performed for the Greenhouse Gas Research Center, one of the Critical Technology-21 Programs, funded by the Ministry of Science and Technology of Korea.

1) Researcher, Forecast Research Lab., Meteorological Research Institute. E-mail: hsjung@metri.re.kr

2) Associate Professor, Information and Telecommunication Research Institute, and Department of Statistics, Chonnam National University. E-mail: jspark@chonnam.ac.kr

3) Associate Professor, Department of Atmospheric Sciences, School of Earth and Environmental Sciences, Seoul National University.

4) Director, Forecast Research Lab., Meteorological Research Institute.