

## 서울의 온도 패턴 변화

장학진<sup>1</sup> · 주용성<sup>2</sup>

<sup>1</sup>한양대학교 수학과 · <sup>2</sup>동국대학교 통계학과

접수 2009년 1월 2일, 수정 2009년 1월 14일, 게재확정 2009년 1월 23일

### 요약

이 논문에서 우리는 1961년부터 2008년 사이의 서울지역 온도변화를 스펙트럴 이분산성 모델을 이용하여 연구하였다. 제안한 모델에서 평균 함수는 계절효과를 주기함수를 이용하여 설명하였고, 온도의 전체적인 상승을 이차 회귀 스플라인 곡선을 이용하여 설명하였다. 분산함수 또한 분산의 계절성을 설명하기 위하여 주기함수를 사용하였다. 우리는 연평균온도가 과거 48년 동안 약 1.5도 가량 증가했음을 알 수 있었다. 연평균온도의 상승은 겨울 온도가 상승하는 것에 기인하는 것이었고, 이는 연중 온도변화의 진폭이 줄어들게 만들었다.

주요용어: 분산함수, 스펙트럴 분석, 스플라인.

### 1. 서론

최근 들어 기후변화에 대한 인식과 관심이 크게 높아지며 많은 연구가 국내외에서 진행되고 있다. 기온 상승으로 인한 뚜렷한 평균값의 변화가 있었다는 것 또한 알려지고 있다. IPCC (Intergovernmental Panel on Climate Change, 2001) 보고서에 의하면 20세기 지구의 평균기온이 0.6°C 상승했음이 밝혀졌으며, 이러한 기온 상승은 기존의 여러 연구결과에 의해 우리나라에서도 발생했음이 밝혀져 왔다. 이종범 (1978)은 서울지방의 기온상승경향과 경년변화의 요인을 분석하여 1954년부터 1977년까지 24년간 서울의 평균기온이 0.64°C 증가하는 것을 보였으며, 강용균과 노재식 (1985)은 1954년부터 1983년까지 30년 동안의 기온분석을 통해 우리나라의 14개 지점에서 기온이 상승하고 있다고 하였다. 그 외에도 기상에 관한 많은 통계적 연구들 (Kim, 2002; Joo, 2008)이 국내에서 행해졌다.

본 연구에서는 기상청 종합기상정보시스템 (<http://www.kma.go.kr/>)에 구축된 기후 DB 중 1961년 이후 서울 지역의 월평균, 연평균 기온 자료를 이용하여 서울지역 기온의 변화를 추정하고 이들에 적합한 회귀모형을 제안하고자 한다. 제안된 모형을 이용하여 기온의 추세 변동성을 분석하고 패턴 변화를 파악하고자 한다. 그림 1.1은 서울지역의 과거 48년동안 (1961년~2008년) 연평균 기온 변화를 나타낸다. 특히 1970년대 이후 상당히 강한 기온 상승 추세가 있었다.

전통적인 회귀모형에 있어서 오차항들인  $\epsilon_i$ 는 평균이 0이고, 분산은 어느 고정된 값  $\sigma^2$ 으로 동일하며 독립적인 분포를 갖는 확률변수라고 가정된다. 흔히 나타나는 이분산 현상 (동분산 현상이 파괴되는 현상) 중 하나는 독립변수값이 변화할 때 이에 대응하여 종속변수값들의 분산이 바뀌는 것이다. 서울지역에서 과거 48년동안의 월평균 기온 분포를 나타낸 그림 1.2를 보면, 온도변화 곡선이 이분산성을 가짐을 간접적으로 알 수 있다. 분산이 늦가을 (10월)부터 초봄 (4월)사이에 가장 크게 나타나고, 여름 (6-7월),

<sup>1</sup> (133-791) 서울시 성동구 행당동 17, 한양대학교 수학과, 박사과정.

<sup>2</sup> 교신저자: (100-715) 서울시 중구 필동 3-26, 동국대학교 통계학과, 조교수.

E-mail : yongsungjoo@dongguk.edu

늦봄 (5월), 초가을 (8-9월) 순으로 작게 나타나는 추세를 보이고 있다. 그림 1.3은 변이가 가장 큰 1월과 가장 작은 9월의 연간 기온 변화를 나타낸다.

월평균 기온 추세선의 분석을 위해서 우리는 다음과 같은 이분산성 스펙트럴 분석 방법을 사용하였다.

$$E(Y_i) = f(t_{1i}, \beta), \quad \text{Var}(Y_i) = g^2(t_{1i}, \gamma). \quad (1.1)$$

여기서  $t_{1i} (= 1, 2, \dots, 576)$ 는 1961년 1월을 기준으로 하여  $i$ 번째 달,  $Y_i$ 는  $i$ 번째 달의 월평균기온을 나타낸다. 예를 들어  $t_{1i} = 1$ 은 1961년 1월을 나타내고  $t_{1i} = 576$ 은 2008년 12월을 나타낸다. 실질적으로  $t_{1i}$ 와  $i$ 가 같지만, 수식을 읽기 편하게 하기 위해서  $t_{1i}$ 가 월을 나타내는 시간변수로 정의되었다.  $f(\cdot)$ 와  $g(\cdot)$ 는 사인과 코사인을 가지는 푸리에 함수이며,  $\beta$ 와  $\gamma$ 는 각각 평균함수  $E(Y_i)$ 와 분산함수  $\text{Var}(Y_i)$ 의 회귀모수이다.

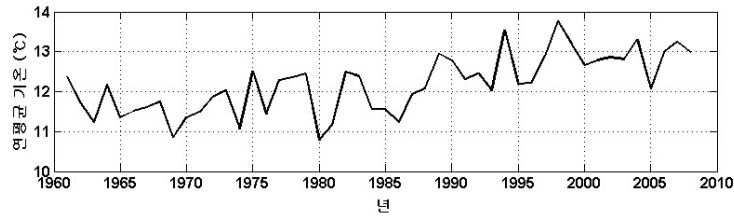


그림 1.1 서울지역 연평균 기온 변화

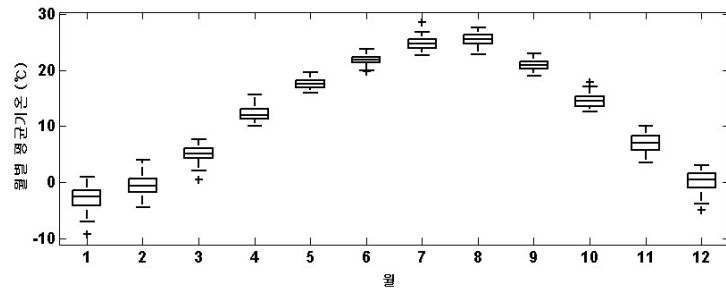


그림 1.2 서울지역 월별 기온 분포

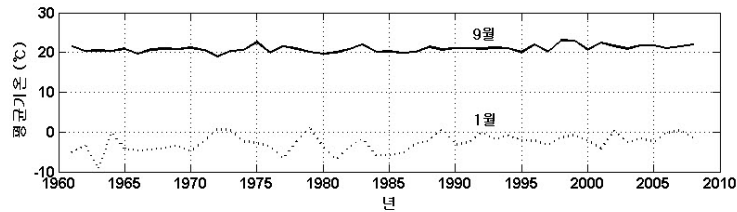


그림 1.3 서울지역 1월, 9월 기온 변화

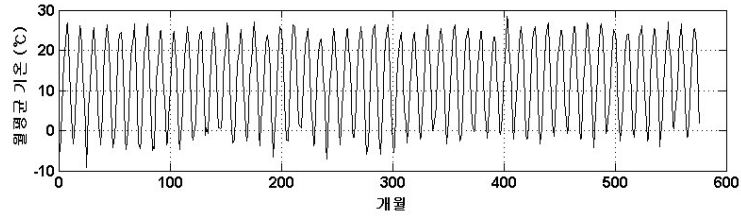


그림 1.4 서울지역 월평균 기온 추세선

본 논문의 구성은 다음과 같다. 2장에서는 스펙트럴 이분산성 모델을 기술하고 모델의 모수 추정 알고리즘을 설명한다. 3장에서는 BIC를 이용한 최적 모델을 결정하고, 기후 자료를 이용한 분석을 수행한다. 마지막으로 4장에서는 이 논문의 결론을 요약한다.

## 2. 분석 방법

### 2.1. 스펙트럴 이분산성 모델

수식 (1.1)에 나와 있는 평균함수와 분산함수는 그림 1.1, 그림 1.2, 그림 1.3 그리고 그림 1.4를 관찰한 후 설정되었다. 월평균 기온 추세선은 12개월 주기로 반복되고 있지만 전체적으로 온도가 약간 상승한 것을 그림 1.1에서 볼 수 있다. 특히 상승하는 패턴이 1970년경부터 변화함을 알 수 있다. 이러한 이유 때문에, 우리는 완전 모델 (full model)의 평균함수에 12개월 주기함수를 포함하였고, 비모수적인 온도상승을 설명하기 위해  $t_{2i}, t_{2i}^2, \max\{t_{2i} - k_0, 0\}^2$ 항을 사용하였다 (Ruppert 등, 2003). 그림 1.4를 관찰해보면 진폭이 줄어드는 듯한 모습을 관찰할 수 있다. 이러한 진폭의 변화는 12개월 주기 코사인, 사인 함수와 시간 ( $t_{2i}, t_{2i}^2$ )과의 교차항을 사용함으로써 설명할 수 있다. 분산은 6개월 주기와 12개월 주기를 가진 주기함수를 사용하였다. 하지만 분산의 변화가 지구온난화와 함께 일어 날수 있기 때문에  $t_{2i}, t_{2i}^2$ 와 교차항들을 완전모델에서 고려하였다.

이러한 함수들은 다음과 같이 수식을 통하여 구체화 되었다.

$$E(Y_i) = f(x_i, \beta) = x_i\beta, \quad Var(Y_i) = g^2(w_i, \gamma) = w_i\gamma.$$

여기서,

$$x_i = \left( 1, \sin\left(\frac{\pi t_{1i}}{6}\right), \cos\left(\frac{\pi t_{1i}}{6}\right), t_{2i}, t_{2i}^2, \max\{t_{2i} - k_0, 0\}^2, \right. \\ \left. t_{2i} \sin\left(\frac{\pi t_{1i}}{6}\right), t_{2i} \cos\left(\frac{\pi t_{1i}}{6}\right), t_{2i}^2 \sin\left(\frac{\pi t_{1i}}{6}\right), t_{2i}^2 \cos\left(\frac{\pi t_{1i}}{6}\right) \right), \\ w_i = \left( 1, \sin\left(\frac{\pi t_{1i}}{6}\right), \cos\left(\frac{\pi t_{1i}}{6}\right), \sin\left(\frac{\pi t_{1i}}{3}\right), \cos\left(\frac{\pi t_{1i}}{3}\right), t_{2i}, t_{2i}^2, \right. \\ \left. t_{2i} \sin\left(\frac{\pi t_{1i}}{3}\right), t_{2i} \cos\left(\frac{\pi t_{1i}}{3}\right), t_{2i}^2 \sin\left(\frac{\pi t_{1i}}{3}\right), t_{2i}^2 \cos\left(\frac{\pi t_{1i}}{3}\right) \right),$$

$\beta = (\beta_0, \beta_1, \dots, \beta_9)$ ,  $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_{10})$ ,  $t_{2i}(= 1, 2, \dots, 48)$ 는  $i$ 번째 달이 속해있는 몇 번째 년도임을 나타낸다. ,

## 2.2. 모델의 점추정

위의 모델은 다음과 같은 알고리즘 (Davidian과 Carroll, 1987)을 이용하여 추정되었다.

1 단계 :

$X = (x_1^T, x_2^T, \dots, x_n^T)^T$ ,  $Y = (Y_1, Y_2, \dots, Y_n)^T$ 라 하자.  $k = 1$ 이라 설정한 후  $\hat{\beta}^{(0)}$ 와  $e^{(0)}$ 를 다음과 같이 계산한다.

$$\begin{aligned}\hat{\beta}^{(0)} &= (X^T X)^{-1} X^T Y, \\ e^{(0)} &= Y - X \hat{\beta}^{(0)}.\end{aligned}$$

2 단계 :

$W = (w_1^T, w_2^T, \dots, w_n^T)^T$ 라 하자. 이 때  $\hat{\gamma}_0^{(k)}$ ,  $\hat{\psi}^{(k)}$ 과  $\hat{\gamma}^{(k)}$ 는 다음과 같이 계산된다.

$$\begin{aligned}\hat{\gamma}_0^{(k)} &= (W^T W)^{-1} W^T (e^{(k-1)})^2, \\ \hat{\psi}^{(k)} &= [\hat{\psi}_{ij}^{(k)}], \quad \hat{\psi}_{ij}^{(k)} = \begin{cases} g^4(w_i, \hat{\gamma}_0^{(k)}) & i=j \\ 0 & i \neq j \end{cases}, \\ \hat{\gamma}^{(k)} &= (W^T (\hat{\psi}^{(k)})^{-1} W)^{-1} W^T (\hat{\psi}^{(k)})^{-1} (e^{(k)})^2.\end{aligned}$$

3 단계 :

$\hat{\Sigma}^{(k)}$ ,  $\hat{\beta}^{(k)}$ 과  $e^{(k)}$ 를 다음과 같이 계산한다.

$$\begin{aligned}\hat{\Sigma}^{(k)} &= [\hat{\Sigma}_{ij}^{(k)}], \quad \hat{\Sigma}_{ij}^{(k)} = \begin{cases} g^2(w_i, \hat{\gamma}^{(k)}) & i=j \\ 0 & i \neq j \end{cases}, \\ \hat{\beta}^{(k)} &= (X^T (\hat{\Sigma}^{(k)})^{-1} X)^{-1} X^T (\hat{\Sigma}^{(k)})^{-1} Y, \\ e^{(k)} &= Y - X \hat{\beta}^{(k)}.\end{aligned}$$

$k = k + 1$ 으로 설정하고 2 단계와 3단계를 추정값들이 수렴할 때까지 반복한다.

위의 알고리즘은 잔차제곱의 평균  $E(e_i^2) = g^2(w_i, \gamma)$ 이고, 잔차제곱의 분산  $Var(e_i^2) = g^4(w_i, \gamma)$ 이라는 사실을 바탕으로 만들어 졌다. 2단계에서는 잔차를 결과변수로 취급하면서 가중 최소제곱법을 이용하여 분산함수를 추정한다. 3단계에서는 2단계에서 구해진 분산함수를 이용하여 평균함수를 가중 최소제곱법을 이용하여 구한다. 이렇게 2-3단계를 반복적으로 실행하면서 추정된 평균함수와 분산함수의 모수 추정치가 일치적 (consistent)이라는 것이 Davidian과 Carroll (1987)에 설명되어있다.

## 3. 결과분석

2.1에서 제안한 스펙트럴 이분산성 모델을 큰 모델 (full model)로 고려하고 여기에서 가장 적절한 설명변수 (모수)를 갖는 모델을 선택하기 위하여 BIC를 사용하였다. BIC는 다음과 같이 정의된다.

$$BIC = -2 \sum_{i=1}^n \ln L(\beta, \gamma) + (p_\beta + p_\gamma) \ln(n).$$

여기서,  $\ln L(\beta, \gamma)$ 는 로그우도함수이고,  $p_\beta$ 는 평균함수에 대한 모수  $\beta$ 의 개수이며,  $p_\gamma$ 는 분산함수에 대한 모수  $\gamma$ 의 개수를 의미한다. 첫 번째 항은 모델적합 정도의 척도로서 자료를 가장 잘 설명할 수 있는 상세한 자료의 모델을 찾도록 유도하는 성분이다. 두 번째 항은 모수의 개수에 대한 패널티 항으로 볼 수 있다. BIC는 모수의 수가 증가할수록 우도의 값이 점진적으로 증가하지만 모수의 수가 증가함에 따

른 패널티 항에 의해 상쇄되므로 서로 조화되는 타협점에서 최선의 모델 상태가 결정되는 것을 제시한다. 따라서 모델 적합이 잘 될수록 BIC값이 작아지게 된다.

BIC를 이용하여 적합한 모델 선정 결과, 다음과 같은 설명변수  $x_i$ 와  $w_i$ 에 대응하는 각각의 모수  $\beta$ 와  $\gamma$ 를 가질 때 가장 적합하였다.

$$x_i = \left( 1, \sin\left(\frac{\pi t_{1i}}{6}\right), \cos\left(\frac{\pi t_{1i}}{6}\right), \max\{t_{2i} - k_0, 0\}^2, t_{2i} \sin\left(\frac{\pi t_{1i}}{6}\right), t_{2i} \cos\left(\frac{\pi t_{1i}}{6}\right) \right),$$

$$w_i = \left( 1, \sin\left(\frac{\pi t_{1i}}{6}\right), \cos\left(\frac{\pi t_{1i}}{6}\right), \sin\left(\frac{\pi t_{1i}}{3}\right), \cos\left(\frac{\pi t_{1i}}{3}\right) \right).$$

또한  $\max\{t_{2i} - k_0, 0\}^2$ 의  $k_0$  ( $k_0 = 1, 2, \dots, 47$ ) 결정을 위하여 BIC를 계산하였다. 결과는 표 3.1에 나타내었다. 표 3.1을 보면 BIC값이  $k_0 = 3$ 일 때 가장 낮게 나타나므로 우리는 최종 모델에서  $k_0 = 3$ 으로 고정하였다. 또한 분석결과  $\max\{t_{2i} - k_0, 0\}^2$ 의 계수가 양의 값으로 나타나므로  $k_0$  번째 해부터 변화하는 온도상승 패턴을 설명한다는 것을 알 수 있다. 평균함수에서 온도상승이나 하강 패턴을 더욱 자세히 설명하기 위하여 여러 개의 매듭항들 ( $\max\{t_{2i} - k_1, 0\}^2, \max\{t_{2i} - k_2, 0\}^2, \dots$ )를 추가적으로 두어 보았으나 베이지안 정보기준 (Bayesian Information Criterion; BIC) (Schwartz, 1978)를 이용하여 적합한 모델 선정 결과 매듭항이 하나일 경우 가장 적합한 것으로 나타났다.

표 3.1  $k_0$ 에 따른 BIC

$k_0$	1	2	3	4	5
BIC	2197.0491	2197.0039	2196.9910	2197.0119	2197.0728
$k_0$	10	15	20	30	40
BIC	2198.1990	2200.9192	2205.2814	2221.4425	2240.4292

BIC를 이용하여 선정된 주요 설명변수들에 상응하는 모수들 ( $k_0 = 3$ )의 추정치를 표 3.2에 기술하였다. 표 3.2를 보면, 예상대로 평균함수에서의 12개월 주기성과 분산함수에서의 6개월 및 12개월 주기 특성이 뚜렷하게 나타난다. 평균함수에서  $t_{2i}$ 에 대응하는 모수가 선택되지 않은 것은 1961년부터  $k_0 = 3$ 인 1964년까지는 온도의 상승이 없었음을 나타내고,  $\max\{t_{2i} - 3, 0\}^2$ 에 대응하는  $\beta_3$ 가 양의 값을 가진다는 것은 1964년 이후 온도가 상승했음을 나타낸다. 하지만, 이 결과가 1961년 이전에 온도 상승이 없었음을 나타내는 것은 아니다. 연평균 경향선 ( $\hat{\beta}_0 + \hat{\beta}_3 \cdot \max\{t_{2i} - 3, 0\}^2$ )은 그림 3.1에 나타나 있다.  $k_0 = 3$ 에 상응하는 1964년 이후 온도상승이 보이며 육안으로는 1970년대 이후 급격한 온도상승을 볼 수 있다. 1961년에 비해 2008년도의 연평균 기온이 대략 1.5°C 정도 상승하였음을 알 수 있다.

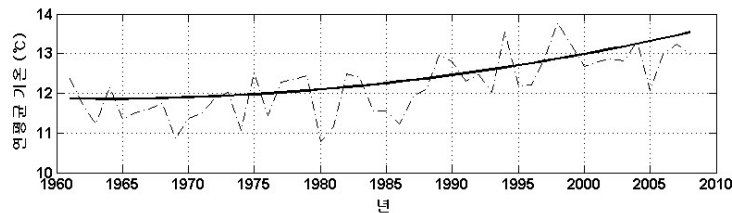


그림 3.1 서울지역 연평균 기온 경향선

표 3.2 회귀모수 추정치

	모수	추정치
평균함수	$\beta_0$	11.8543
	$\beta_1$	-8.6650
	$\beta_2$	-11.2059
	$\beta_3$	0.0008
	$\beta_4$	0.0158
	$\beta_5$	0.0173
분산함수	$\gamma_0$	2.7548
	$\gamma_1$	1.4055
	$\gamma_2$	1.1938
	$\gamma_3$	0.7546
	$\gamma_4$	0.6362

$t_{2i} \sin\left(\frac{\pi t_{1i}}{6}\right)$ 와  $t_{2i} \cos\left(\frac{\pi t_{1i}}{6}\right)$ 에 대응하는  $\beta_4$ 과  $\beta_5$ 가 유의한 것은 월평균 기온의 평균함수의 진폭의 변화가 있었음을 나타낸다. 평균함수에서의 주기성을 나타내는 함수는 다음과 같다.

$$y_p = (\beta_1 + \beta_5 t_{2i}) \sin\left(\frac{\pi t_{1i}}{6}\right) + (\beta_2 + \beta_6 t_{2i}) \cos\left(\frac{\pi t_{1i}}{6}\right). \quad (3.1)$$

$y_p$ 의 진폭의 변화를 확인하기 위하여, 수식 (3.1)을

$$y_p(t_{1i}) = C^* \cos\left(\frac{\pi t_{1i}}{6} - \eta\right)$$

의 형태로 쓴다. 여기서  $C^*$ 는  $y_p$ 의 진폭,  $\eta$ 는 위상각이라 하며, 다음의 관계식을 만족한다.

$$C^* = \sqrt{(\hat{\beta}_1 + \hat{\beta}_5 t_{2i})^2 + (\hat{\beta}_2 + \hat{\beta}_6 t_{2i})^2},$$

$$\tan \eta = \frac{\hat{\beta}_2 + \hat{\beta}_6 t_{2i}}{\hat{\beta}_1 + \hat{\beta}_5 t_{2i}}.$$

그림 3.2에 나와 있듯이 진폭은 대략 1.2°C 정도 줄어들었음을 알 수 있다. 그림 1.3과 그림 1.4를 자세히 보면 1961년 이후 겨울 온도는 약간 상승하였지만 여름 온도는 거의 비슷함을 관찰할 수 있다. 다시 요약하여 설명하자면 1.5°C 가량의 연평균 온도 상승은 주로 겨울철 기온의 상승에 의한 것이고 이로 인하여 진폭이 줄어들었다고 할 수 있다.

분산함수 추정식 ( $Var(\hat{Y}) = W\hat{\gamma}$ )에 의한 추정값은 그림 3.3에 나타나 있다. 그림 3.3을 관찰해 보면 겨울철과 여름철에 분산이 크게 나타나는 계절성이 유의함을 볼 수 있다. 과거 48년 동안 분산은 그림 3.3과 같은 패턴을 가졌음을 확인할 수 있었다.

#### 4. 결론

본 연구에서는 1961년~2008년의 서울지역 기온변화를 분석하였다. 그 결과 월평균 기온은 12개월 주기로 반복되며 1964년 이후 유의한 온도 상승을 보였다. 전체적으로는 48년 동안 연평균 온도는 대략 1.5°C 정도 상승하였으며 연내 진폭은 14.2°C (1961년)에서 13.0°C (2008년)으로 대략 1.2°C 정도 줄어들

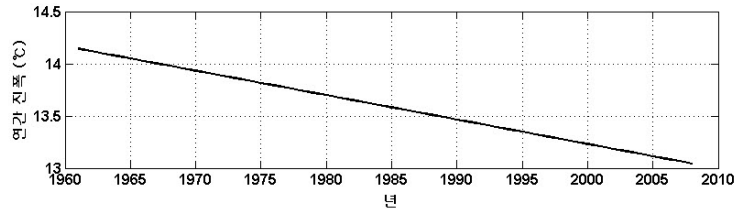


그림 3.2 서울지역 기온 진폭 변화

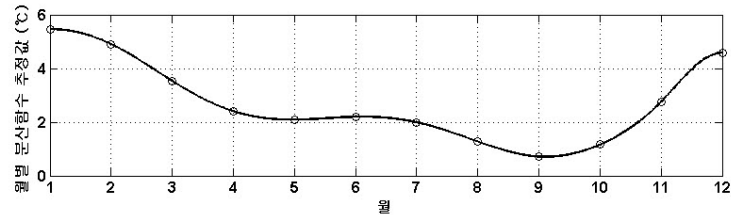


그림 3.3 월별 분산 패턴

었다. 연평균 온도 상승은 주로 겨울철 기온의 상승에 의한 것이며 이로 인하여 진폭이 줄어든 것으로 사료된다.

### 참고문헌

- 강용균, 노재식 (1985). 과거 30년간 (1954-1983) 한국 기온의 연변화 및 연별 변동. <한국기상학회지>, **21**, 1-10.
- 이중범 (1978). 서울지방의 기온 경년변화와 그 요인에 관하여. <한국기상학회지>, **14**, 29-35.
- Davidian, M. and Carroll, R. J. (1987). Variance function estimation. *Journal of the American Statistical Association*, **82**, 1079-1091.
- IPCC (2001). Climate change 2001, the scientific basis. *Houghton and Ding (Eds.)*, Cambridge Press, 881.
- Kim, B., Cho, C., Chung, H., Park, J., Shin, S. and Lee, Y. (2002). Impact of the additional observation data on the weather analysis. *Proceedings of Joint Conference of Korean Data & Information Science Society and Korean Data Analysis Society 2002*, 1-4.
- Joo, Y., Lee, K. and Jung, H. (2008). Estimation of interval censored regression spline model with variance function. *Journal of Korean Data Analysis Society*, **19**, 1247-1253.
- Ruppert, D., Wand, M. and Carroll, R. (2003). *Semiparametric regression*, Cambridge University Press, Cambridge.
- Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **5**, 461-464.

## Change of temperature patterns in Seoul

Hakjin Jang<sup>1</sup> · Yongsung Joo<sup>2</sup>

<sup>1</sup>Mathematics Dept, Hanyang University

<sup>2</sup>Statistics, Dongguk University

Received 2 January 2009, revised 14 January 2009, accepted 23 January 2009

### Abstract

We examined the characteristics of temperature variation in Seoul between 1961 to 2008 using the spectral heteroscedastic model. The mean function in the proposed model explains the season effect using periodic functions and the overall increase using the quadratic regression spline. The variance function also had periodic functions to explain the seasonality of variance. We found that there has been annual mean temperature increase by about 1.5°C for the last 48 years. The increase of annual mean temperature was mainly caused by the increase in winter, which made the amplitude decreased.

*Keywords:* Spectral analysis, spline, variance function.

---

<sup>1</sup> PhD student, Mathematics Dept, Hanyang University, Hangdang 17, Seoul 133-791, Korea.

<sup>2</sup> Corresponding Author: Assistant Professor, Statistics, Dongguk University, Joonggu, Pildong 3-26, Seoul 100-715, Korea. E-mail: yongsungjoo@dongguk.edu