

大氣汚染濃度에 관한 動的確率모델

A Dynamic-Stochastic Model for Air Pollutant Concentration

김 해 경

연세대학교 이과대학 수학과
(원고접수: 1991. 9. 30)

Hae Kyung Kim

Department of Mathematics, College of Sciences, Yonsei University
(Received 30, September 1991)

Abstract

The purpose of this paper is to develop a stochastic model for daily sulphur dioxide (SO_2) concentrations prediction in urban area (Seoul). For this, the influence of the meteorological parameters on the SO_2 concentrations is investigated by a statistical analysis of the 24-hr averaged SO_2 levels of Seoul area during 1989~1990. The annual fluctuations of the regression trend, periodicity and dependence of the daily concentration are also analyzed. Based on these, a nonlinear regression transfer function model for the prediction of daily SO_2 concentrations is derived. A statistical procedure for using the model to predict the concentration level is also proposed.

1. 서 론

대기오염농도의 정확한 예측은 대기질의 개선이나 오염사건의 발생을 예방하는 차원에서 중요하다. 대기오염농도는 다양한 요인으로 방출된 오염물질이 시간의 변화에 따라 여러가지 기상지배인자들의 영향을 받고 확산이동되는 상태에서 결정된다. 대기오염농도나 오염상태의 완전한 설명은 관련된 자연현상 또는 인위적활동 그리고 여기에 영향을 주는 모든 지배인자 사이에서 일어나는 역학현상 등의 종합적인 記述로 이루어질 수 있다. 따라서 이 복잡한 화학 및 물리적 현상이 다양한 기상현상의 영향을 받아 일어나는 이 대기오염현상은 몇가지 간단한 결정모델로는 만족스럽게 설명되지 않는다.

대기오염수준은 늘 변하면서 어떤 규칙성을 지니고 있다. 이러한 현상은 오염수준이 시간의 변화에서 자기 자신에게 從屬되는 自己回歸性을 소유하거나 또는 일정한 週期的패턴을 갖는 어떤 先導因子

의 존재에서 기인될 수도 있다. 더욱이, 이러한 형태의 종속현상은 대개 불확실성이 내존하는 확률현상으로 어떤 決定函數보다는 오히려 確率函數로 기술하는 것이 더 합리적이다.

김해경(1991a)에서 언급한 바와 같이, 오염농도의 확률론적 예측에는 기본적으로 두가지 형태의 모델이 이용될 수 있다. 기상인자와 같이 오염농도에 관련된 주요 지배인자와 여기에 관련된 決定的 또는 確率的 형태를 파악하고 그 관계를 모델개발에 이용하는 이른바 傳達函數確率모델과, 이에 반하여 시간에 따라 내존하는 오염농도의 확률적종속관계를 유도하고 그 관계를 예측에 이용하여 농도의 예측에는 과거와 현재의 오염농도만을 기초로 하는 單變數確率모델이다. 또, 위 논문에서는 도시지역 대기오염현상의 確率 및 統計의 특성을 파악하고 대기오염농도 수준의 예측을 위한 單變數確率모델을 개발하고 통계적절차를 제안하였다. 그러나, 어느 정도의 성과에도 불구하고, 이 모델은 오염농도의 예측모델로서 다음 두 가지의 근본적인 缺點을 지니고 있다. 첫

제, 이 모델에서는 오염농도에 영향을 주는 기온, 풍속 등 오염농도의 주요 지배인자들이 직접 고려될 수 없다. 둘째, 모델의 從屬成分이 自己回歸過程으로 설명되기 때문에 長期豫測에는 예측오차의 조절에 한계를 갖는다.

본 논문의 목적은 대기오염농도의 예측에 있어서 單變數確率모델의 결점을 보완할 수 있는 확률모델을 개발하는데 있다. 이를 위해, 먼저 도시지역 아황산가스의 농도와 이것의 지배인자들 사이에 내존하는 從屬性을 분석하고 이 관계를 이용하여 대기오염농도의 統計 및 確率의 특성을 파악하였다. 다음으로 이 결과들을 기초로 하여 아황산가스 농도수준의 예측을 위한 確率모델을 완성하고, 그 適合度를 확인하였다. 마지막으로 유도된 確率모델을 이용한 오염농도예측의 統計的절차를 제안하고, 결과가 單變數確率모델보다 우수함을 實例를 통하여 확인하였다.

2. 資料와 分析方法

본 분석에서는 1989, 1990년 환경처가 관측한 서울지역(잠실, 광화문, 문래 등 20개동) 시간별 아황산가스(SO₂) 농도 및 풍속, 풍향 등 기상지배인자의 측정값과, 같은 기간동안 기상청에서 관측한 서울지역 일별(최저, 최고, 평균) 기온과 강우량, 습도의 관측치를 이용하였다. 대기오염농도는 다양한 요인으로 방출된 오염물질이 자기자신뿐 아니라 여러가지 기상인자들의 영향을 받고 확산이동되는 상태에서 결정되므로 이 분석에는 기본적으로 지역의 특성과 계절은 물론 기온, 풍속 등 모든 기상지배인자들의 자료를 필요로 한다. 그림 1은 자료중 SO₂의 농도시계열 일부를 나타낸 것이다.

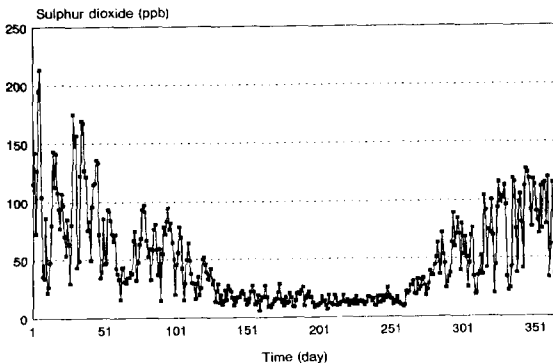


Fig. 1. Original daily SO₂ (ppb) series (Jamsil): 1989. 1. 1~1989. 12. 31.

그림 1에서와 같이, SO₂ 농도는 전반부부터 점차 하강하여 중반부에서는 최저상태를 보이고, 그 후 다시 점차 상승하는 어떤 趨勢를 내포하지만 전후반에서는 중반부에 비해 매우 큰 分散을 가지며 변동하고 있다. 이와같은 특성들은 오염농도가 기온, 풍속 등 이중 어떤 인자들의 영향하에 있음을 암시하며, 농도의 예측에는 오염수준의 확률구조뿐 아니라 이러한 지배인자들의 과거나 현재의 정보도 이용하는 것이 바람직함을 의미한다. 이런 의미에서 농도의 예측모델로는 오염농도의 지배인자들이 그 모델에 포함되는 傳達函數確率모델이 더 이상적일 수 있다.

그러나, 대부분의 지배인자들도 각각 回歸趨勢, 週期性 그리고 從屬性 등 고유의 확률구조를 가지며 또 어떤 시점에서 오염농도의 예측에는 적어도 같은 시점에서 이 변수들의 오차가 동반되는 예측값을 필요로 하게 되므로, 독립변수의 수를 최소화하는 모델의 간소화는 예측의 효율면에서 오히려 바람직하다.

이런 의미에서, 분석의 절차는 먼저 대기오염농도에 영향을 주는 대표적인 독립인자를 결정하여 그 確率構造를 규명하고, 다음으로 SO₂ 濃度가 이 독립변수들로 설명되는 確率모델을 결정하는 것이다.

附錄에서는 대기오염농도와 이에 영향을 미치는 주요인자들 특히 기상인자들 사이에 존재하는 통계적, 확률적 종속관계를 분석하였다. 이 분석에 의하면 대기오염농도에 영향을 주는 대표적인 독립인자는 氣溫과 風速이다. 그러나, 겨울철의 경우와는 달리 일년 전체에서는 풍속이 기온에 비해 오염농도에 대한(線型 또는 非線型的) 回歸關係가 모두 희박하므로 사계절 SO₂ 濃度의 예측은 (평균)기온만으로 설명되는 확률모델이 바람직하다.

따라서, 오염농도에 영향을 주는 기온을 독립변수로 갖는 오염농도의 확률모델의 개발은 예측오차를 감소시키고 효율을 증대시킬 수 있다. 그러나, 일반적으로 두 시계열 사이에 존재하는 종속관계를 함수식으로 쉽게 설명할 수 없다. 그것은 각 시계열에 내존하는 回歸趨勢, 週期性 또는 自己從屬性 등이 두 시계열에 존재하는 종속관계의 강도를 나타내는 相關係數나 交叉相關係數를 실제보다 확대하거나 축소할 수 있기 때문이다(Bartlett, 1935). 특히, 각 시계열의 分散들이 두 시계열의 연합분포의 分散보다 클 때는 交叉共分散 등 종속관계가 각 시계열의 單獨分散에 의하여 매몰되는 경우가 흔히 있다. 따라서, 이러한 분석에는 관련된 시계열의 事前白場化(Box & Jenkins, 1976)를 필요로 한다. 시계열의 事前白場化는 시간의 변화에서 나타나는 그 변량의

전체적 변동상태의 관찰과 분석을 통하여 그 시계열에 내존하는 回歸趨勢와 週期性 등을 제거하여 白場 시계열로 만드는 과정으로 그 시계열의 單變數確率 모델화 과정으로 볼 수 있다. 왜냐하면, 기온이 독립인자로서 대기오염농도의 결정에 강한 영향을 주는 先導因子라면 대기오염농도의 확률모형은 기온의 확률모형과 유사하게 되고, 따라서 이 모델화 과정에서 얻어지는 양측의 마지막 殘差는 모두 백장으로 볼 수 있기 때문이다. 물론 종속인자인 대기오염농도는 기온뿐 아니라 자신의 과거수준에도 종속되어 있기 때문에 대기오염농도에 대응하는 잔차는 近似的 白場으로 보는 것이 타당하다. 이런 의미에서 事前白場化를 위해서 먼저 최저기온 시계열을 單變數確率모델화한다. 그리고 이 모델을 SO₂ 오염농도 시계열에 적용하고 여기에서 얻어진 두 殘差의 종속성을 분석한다.

3. 平均氣溫의 單變數確率모델

日日最低, 最高 그리고 平均氣溫 사이에는 (時系列) 相關係數가 모두 0.93 이상임을 감안하여 본 분석에서는 평균기온을 일일기온의 대표기온으로 사용한다.

앞에서 말한 바와 같이, 平均氣溫의 전체적인 경향은 전반부부터 점차 상승하여 7~8월경에는 최고상태를 보이고, 그후 다시 점차 하강하는 어떤 趨勢를 내포하며, 동시에 趨勢주위를 천천히 상하진동하며 어떤 기간을 두고 반복되는 週期性이 나타난다. 또한 氣溫은 연속변화로서 순간적 도약을 갖지 않기 때문에 농도수준 상호간의 從屬性의 존재가 가정된다.

따라서, 일일 平均氣溫 T_t(t는 시간)는 다음 기본假定式으로 표현된다.

$$T_t = D_t^{(T)} + C_t^{(T)} + S_t^{(T)} + E_t^{(T)}$$

이다. 여기에서 D_t^(T)는 回歸趨勢성분, C_t^(T)는 週期성분, S_t^(T)는 確率성분 그리고 E_t^(T)는 측정오차를 포함하는 確率誤差성분이다. 또, 확률성분 S_t^(T)와 측정오차 E_t^(T)의 합, Z_t^(T)=S_t^(T)+E_t^(T)은 보통 定常 또는 非定常 ARMA(p, q) 모델

$$\sum_{j=0}^p a_j Z_{t-j} = \sum_{k=0}^q b_k e_t \quad (p, q \text{는 양의 정수}) \quad (3.1)$$

로 설명된다. 이때 a_j와 b_k는 적당한 상수(단 a₀=b₀=1)이고 e_t는 (正規)白色誤差이다. 따라서, 이 기본가정식은 T_t와 식(3.1)을 만족하는 Z_t를 사용하여

$$T_t = D_t^{(T)} + C_t^{(T)} + Z_t^{(T)} \quad (3.2)$$

로 표시된다.

평균기온의 단변수 확률모델화는 가정식 (3.2)의 각 성분 즉 回歸趨勢, 週期 그리고 從屬 성분의 最適函數 형태를 결정하고 또 그 確率構造를 파악하는데 있다. 김해경(1991b)에서와 같은 과정에 의하면 서울지방의 경우 D_t^(T)의 적절한 함수식은 4차 多項式이다. 실례로 1989년 잠실의 경우 2차와 3차 다항식에서 발생하는 잔차제곱합(SSE)이 각각 5416.09, 3995.07인데 반하여 4차 다항식의 경우에는 2852.65로 감소됨을 볼 수 있다. 보편화된 최소제곱방법으로 이 함수식의 母數를 推定하고 그 有意性を 檢定할 수 있다. 또, 측정값과 이 回歸成分에서 얻어진 殘差 R_{T,t}⁽¹⁾, R_{T,t}⁽²⁾=T_t-D_t^(T)의 피리오도그램이나 스펙트럼을 기초로 하여 週期成分 C_t^(T)의 함수형태

$$C_t^{(T)} = \sum_{i=1}^p \alpha_{i0} \sin(\alpha_{i1} + \alpha_{i2}t)$$

를 결정하고, 또 殘差 R_{T,t}⁽²⁾, R_{T,t}⁽²⁾=T_t-(D_t^(T)+C_t^(T))를 이용하여 종속성분 Z_t^(T)의 모델

$$Z_t^{(T)} = \sum_{i=1}^q a_i Z_{t-i}^{(T)} + e_t^{(T)}$$

의 階數와 未知母數를 결정한다. 여기에서 週期성분과 從屬성분의 階數 p, q는 해당년도의 固有 整數이다. 따라서, 평균시계열 T_t는 t≥q+1에 대하여

$$T_t = a_0 + a_1t + a_2t^2 + a_3t^3 + a_4t^4 + \sum_{i=1}^p \alpha_{i0} \sin(\alpha_{i1} + \alpha_{i2}t) + \sum_{i=1}^q a_i Z_{t-i}^{(T)} + e_t^{(T)}$$

이다. 여기에서 e_t^(T)는 正規(0, σ²) 確率變數이고, Z_t^(T)는 殘差 R_{T,t}⁽²⁾로 近似化될 수 있는 確率變數이다. 1989년도의 자료의 경우, 殘差시계열 {R_{T,t}⁽¹⁾}의 主, 副週期가 121.7일, 12.6일로 각각 나타나고, 이중 主週期 121.7일만을 이용한 주기성분함수로 얻어진 殘差시계열 {R_{T,t}⁽²⁾}는 階數 2의 自己回歸過程(AR(2))를 따른다. 결과적으로, 평균기온의 확률모델은

$$T_t = a_0 + a_1t + a_2t^2 + a_3t^3 + a_4t^4 + a_{10} \sin(\alpha_{11} + \alpha_{12}t) + a_1 Z_{t-1}^{(T)} + a_2 Z_{t-2}^{(T)} + e_t^{(T)}$$

로 주어진다. 회귀성분과 주기성분의 합의 함수를 기온자료에 적용하여 얻어진 최소제곱법 추정값(SSE=2552.41)은

$$\begin{aligned} \hat{a}_0 &= 2.1681(0.4212) & \hat{\alpha}_1 &= -0.1363(0.0028) \\ \hat{a}_2 &= 0.0036(0.00001) & \hat{\alpha}_3 &= -0.000016(0.0000013) \\ \hat{a}_4 &= 0.000000018(0.000000000) & \hat{a}_{10} &= -1.3846(0.2006) \\ \hat{\alpha}_{11} &= -0.6599(0.0330) & \hat{\alpha}_{12} &= 0.0509(0.0016) \end{aligned}$$

이다. 또 殘差 $\{R_{\tau,t}^{(2)}\}$ 로부터 계산된 a_1, a_2 의 (條件) 최소제곱추정치는

$$\hat{a}_1 = 0.8907(0.04979), \quad \hat{a}_2 = -0.3171(0.04982)$$

이다. 괄호안은 推定의 標準誤차를 표시한다.

4. 氣溫과 汚染濃度 사이의 動的相關

大氣污染濃도가 자기의 과거뿐 아니라 일일 平均 氣溫에서도 영향을 받을 때, 기온과 이 오염농도 사이에 존재하는 動的종속관계는 函數로 표현되는 근사적 線型濾過函數로 설명하는 것이 편리하다. 본 절에서는 이 두 변량사이의 전달함수를 유도하는데 기초가 되는 交叉相關을 분석한다.

일일 平均 SO_2 濃度 X_t 도 가정식 (3.2)의 형태로 표시할 수 있다. 그러나, 시간에 따라 일정치 않은 오염농도의 분산은 분석결과와 효율을 감소시킬 뿐 아니라 분석방법의 기본가정인 분산의 定常을 위한 변수변환이 필요하다. 적당한 변수변환은 SO_2 농도 X_t 의 對數變換 $\ln X_t$ 이다(그림 5(a) 참조).

따라서, 對數 SO_2 농도의 측정값 $Y_t = \ln X_t$ (t = 시간)는 다음 기본假定式으로 표현된다.

$$Y_t = D_t^{(v)} + C_t^{(v)} + Z_t^{(v)}$$

단, $D_t^{(v)}, C_t^{(v)}$ 그리고 $Z_t^{(v)}$ 는 각각 Y_t 의 回歸趨勢, 週期 그리고 從屬 成分이다.

平均 SO_2 농도의 單變數確率모델화는 김해경 (1991a)의 결과를 참고하면, 서울지방 SO_2 농도수준의 確率모델은

$$\ln X_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \sum_{i=1}^q \beta_{0i} \sin(\beta_{1i} + \beta_{2i} t) + \sum_{i=1}^q b_{0i} Z_{t-i}^{(v)} + e_t^{(v)}$$

이다. 1989년도 잠실자료의 경우, 有意的인 4次回歸 多項式 $D_t^{(v)}$, 主週期 91.3일만을 사용한 $p=1$ 의 週期函數式 $C_t^{(v)}$ 그리고 $q=1$ 인 AR(1)의 從屬成分 $Z_t^{(v)}$ 들에 대한 최소제곱추정값(SSE=65.579)은

$$\begin{aligned} \hat{\beta}_0 &= 4.0402(0.1236) & \hat{\beta}_1 &= 0.0259(0.0045) \\ \hat{\beta}_2 &= -0.0004(0.00005) & \hat{\beta}_3 &= 0.000002(0.0000002) \\ \hat{\beta}_4 &= -0.00000003(0.00000000) & \hat{\beta}_{10} &= -0.1863(0.0332) \\ \hat{\beta}_{11} &= -0.9659(0.3782) & \hat{\beta}_{12} &= -0.0667(0.0018) \\ \hat{b}_1 &= 0.48112(0.0459) \end{aligned}$$

이다.

앞에서 말한 바와 같이, 平均氣溫을 독립인자로 하는 SO_2 汚染濃度の 전달함수모델을 유도하기 위

해서는 먼저 두 변량에서 각각 자기의 회귀추세와 주기성을 제거하여 平均(零)과 分散이 일정한 定常時系列을 만든다. 그것은 각 변량의 分散들이 두 변량의 共分散보다 클 때는 交叉共分散 등 종속관계가 각 변량의 單獨分散에 의하여 매몰되기 때문이다. 따라서 먼저 殘差 $R_{\tau,t}^{(2)}$ 을 이용하여 평균기온의 종속성분 $\{Z_t^{(T)}\}$ 을 ARIMA 모델에 의하여 白場化하고, 다음에 같은 모델을 오염농도의 종속성분 $\{Z_t^{(V)}\}$ 에 적용하여 새로운 殘差를 얻는다. 전달함수모델은 이 과정에서 얻어진 두 殘差 사이에 존재하는 交叉相關關係로부터 설명되기 때문이다.

위에서 밝혀진 바와 같이, 1989년 잠실자료의 경우 殘差 $R_{\tau,t}^{(2)}$ 에 의하여 근사화된 氣溫의 從屬成分 $Z_t^{(T)}$ 는 AR(2) :

$$Z_t^{(T)} = 0.8907 Z_{t-1}^{(T)} - 0.3171 Z_{t-2}^{(T)} + e_t^{(T)}$$

를 따르고, 白場時系列 $\{e_t^{(T)}\}$ 은 다음식

$$R_{\tau,t}^{(3)} = R_{\tau,t}^{(2)} - 0.8907 R_{\tau,t-1}^{(2)} + 0.3171 R_{\tau,t-2}^{(2)} \quad (t=3, 4, \dots, n)$$

로 정의되는 세번째 殘差 $\{R_{\tau,t}^{(3)}\}$ 로 근사화된다. 이 결과를 변수 Y_t 의 종속성분 $Z_t^{(V)}$ 에 적용시키면

$$Z_t^{(V)} = 0.8907 Z_{t-1}^{(V)} - 0.3171 Z_{t-2}^{(V)} + e_t^{(V)}$$

이고, $\{R_{y,t}^{(2)}\}$ 가 $Z_t^{(V)}$ 의 근사시계열로 Y_t 의 두번째 殘差일 때

$$R_{y,t}^{(3)} = R_{y,t}^{(2)} - 0.8907 R_{y,t-1}^{(2)} + 0.3171 R_{y,t-2}^{(2)} \quad (t=3, 4, \dots, n)$$

로 정의하여 Y_t 의 세번째 잔차 $\{R_{y,t}^{(3)}\}$ 을 얻는다. 그리고 $\{R_{\tau,t}^{(3)}\}$ 와 $\{R_{y,t}^{(3)}\}$ 로부터 추정되는 $\{e_t^{(T)}\}$ 와 $\{e_t^{(V)}\}$ 의 交叉共分散과 交叉相關을 구하고, 이것을 이용하여 다시 $\{Z_t^{(T)}\}$ 와 $\{Z_t^{(V)}\}$ 사이의 衝擊反應值 v_k 를 공식 (4.1)를 통해 얻는다.

$$v_k = \frac{C_{T,Y}(k)}{S_T^2} \quad (4.1)$$

단, $C_{T,Y}(k)$ 는 時差 k 에서 $\{e_t^{(T)}\}$ 와 $\{e_t^{(V)}\}$ 의 交叉共分散 그리고 S_T^2 는 $\{e_t^{(T)}\}$ 의 分散이다. 時差에 따라 계산된 v_k 의 값과 형태는 그림 2에 주어진다.

그림 2로부터 두 시계열 사이에 존재하는 몇가지 확률적 특징이 관찰된다. 먼저, 交叉相關係數의 급속한 감소는 적용된 두 殘差時系列의 定常狀態를 확인해 준다. 다음으로, 陽의 時差에서 주로 有意的인 交叉相關의 값이 나타나고 있음은 氣溫이 大氣汚染의 어느정도의 先導指數가 됨을 의미한다. 그리고, 交叉相關이 時差 1에서 가장 큰 有意的인 값을 가지며, 더욱이 陰으로 나타나고 있다. 이것은 두 변량

Lag	Cross		Plot of Cross correlations and standard errors																			Impulse response	
	Covar	Corr																					
			-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
0	0.17137	0.22860													*****								0.05007
1	-0.23300	-0.31080										*****			.								-0.06807
2	-0.04810	-0.06423										*			.								-0.01406
3	-0.03740	-0.04995										*			.								-0.01094
4	-0.05730	-0.07642										**			.								-0.01673
5	-0.00400	-0.00529										.			.								-0.00115
6	0.00994	0.01325										.			.								0.00290
7	-0.04150	-0.05532										.	*		.								-0.01212
8	0.01419	0.01893										.	.		.								0.00414
9	0.02180	0.02908										.	.	*	.								0.00636
10	-0.01370	-0.01828																	0.00400

"." marks two standard errors

Variance and standard deviation of transformed series:

$$s_y^2 = 0.164193, \quad s_r^2 = 3.422528$$

$$s_y = 0.405207, \quad s_r = 1.850007$$

Fig. 2. Cross covariance, cross correlation and impulse response estimates between $\{e_t^{(T)}\}$ and $\{e_t^{(V)}\}$.

사이에는 時差 1에서 陰의 交叉相關이 내존하는 것으로, 내일의 대기오염농도는 오늘의 기온에 의하여 영향을 받으며 오늘의 기온하락(상승)은 내일의 SO₂ 오염농도의 상승(하락)을 가져옴을 의미한다. 그러나, 미세하지만 時差 0에서 나타나는 有意의 太陽의 相關은 오늘의 기온 상승이 당일인 오늘의 오염농도를 증가시키는 요인도 다소 내포하고 있음을 말한다. 이러한 현상에 대한 역학적인 설명은 氣象學으로 미룬다.

5. SO₂ 汚染濃度の 傳達函數모델

본 절에서는 기온과 오염농도와의 動的從屬을 설명하는 傳達函數모델을 유도하고 미지수의 推定, 모델의 適合性 그리고 예측방법 등을 實例를 통하여 생각한다.

5.1 傳達函數의 決定

殘差 $\{R_{T,t}^{(3)}\}$ 와 $\{R_{Y,t}^{(3)}\}$ 로 계산된 그림 2의 標本衝擊反應函數 패턴을 관찰하면 시계열 $\{Z_t^{(V)}\}$ 와 $\{Z_t^{(T)}\}$ 에 존재하는 傳達函數모델은

$$Z_t^{(V)} = (A_0 - A_1 B) Z_t^{(T)} + U_t$$

(단, $BZ_t^{(T)} = Z_{t-1}^{(T)}$)으로 주어진다. 더욱이 衝擊反應函數와 傳達函數와의 관계(Box & Jenkins, 1976)

로 부터 계산된 계수는

$$\hat{A}_0 = v_0 = 0.0501$$

$$\hat{A}_1 = -v_1 = 0.0681$$

이다. 따라서, 구하는 傳達函數式은

$$Z_t^{(V)} = (0.0501 - 0.0681 B) Z_t^{(T)} + U_t$$

이다.

한편, 誤差項 $\{U_t\}$ 는 3절에서와 같은 방법으로 이 전달함수식에서 얻어지는 殘差의 標本 自己相關係數와 部分自己相關係數로부터 AR(1)을 따르고 있음이 확인된다. 구체적으로

$$U_t = -0.1742 U_{t-1} + w_t \text{ 또는 } U_t = \frac{w_t}{(1 + 0.1742 B)}$$

가 된다. 위 두 결과로부터

$$Z_t^{(V)} = (0.050 - 0.068 B) Z_t^{(T)} + \frac{w_t}{(1 + 0.1742 B)}$$

을 얻는다. 결과적으로, 구하는 傳達函數 모델은

$$Z_t^{(V)} = (A_0 - A_1 B) Z_t^{(T)} + \frac{w_t}{(1 + \phi_1 B)} \quad (5.1)$$

이다. 이때 $\{w_t\}$ 는 평균이 零인 正規白場誤差이다.

5.2 母數의 推定

각 성분을 분리적용하여 얻어진 전달함수의 추정

값

$$\hat{A}_0=0.0501, \hat{A}_1=0.0681, \hat{\phi}_1=0.1742$$

들은 誤差를 비교적 많이 포함한다. 원성모델 전체를 殘差 $\{R_{T,t}^{(2)}\}$ 와 $\{R_{Y,t}^{(2)}\}$ 에 다시 적용하면 추정오차를 최소화할 수 있다. 이 때, 위 값들은 初期값으로 하는 非線型최소제곱방법이 적당하다. 여기서, 비선형최소제곱추정값이란 (5.1)에서 얻어지는

$$w_t = Z_t^{(1)} + \phi_1 Z_{t-1}^{(1)} - A_0 Z_t^{(1)} - (A_0 \phi_1 - A_1) Z_{t-1}^{(1)} + A_1 \phi_1 Z_{t-2}^{(1)}$$

의 條件殘差제곱합

$$Q(A_0, A_1, \phi_1) = \sum_{t=3}^n w_t^2(A_0, A_1, \phi_1 | \{Z_t^{(1)}\}) = \{R_{Y,t}^{(2)}\}, \{Z_t^{(1)}\} = \{R_{T,t}^{(2)}\}$$

을 최소화하는 A_0, A_1, ϕ_1 의 값을 말한다. 컴퓨터 패키지(예 : SAS)를 이용하여 추정값을 계산하면 잔차제곱합을 68.16에서 42.77으로 감소시키는 추정값

$$\begin{aligned} \tilde{A}_0 &= 0.0720 (0.0097) \\ \tilde{A}_1 &= 0.0788 (0.0096) \\ \tilde{\phi}_1 &= -0.4671 (0.0473) \end{aligned}$$

이 구해진다.

앞 절의 결과들을 (5.1)에 대입하여 전달함수모델이 완성된다 :

$$Y_t = \tilde{D}_t^{(1)} + \tilde{C}_t^{(1)} + (0.0720 - 0.0788 B) (T_t - \tilde{D}_t^{(1)}) - \tilde{C}_t^{(1)} + \frac{w_t}{(1 - 0.4671B)} \quad (5.2)$$

$$\begin{aligned} \text{단, } \tilde{D}_t^{(1)} &= 4.0402 + 0.0259t - 0.0004t^2 + 0.000002t^3 \\ &\quad - 0.000000003t^4, \\ \tilde{C}_t^{(1)} &= -0.1863 \sin(-0.9659 - 0.0667t), \\ \tilde{D}_t^{(1)} &= 2.1680 - 0.1363t + 0.0036t^2 - 0.000016t^3 \\ &\quad + 0.000000018t^4, \\ \tilde{C}_t^{(1)} &= -1.3840 \sin(-0.6599 + 0.0509t). \end{aligned}$$

5.3 모델의 適合度

(5.2)가 평균기온을 독립인자로 갖는 SO₂ 오염농도의 예측모델로 적합한가를 확인할 필요가 있다. 일반적으로, 모델의 적합성은 이것을 실제자료에 적용하여 얻어진 결과와 여기에서 얻어지는 殘差들의 분석으로 확인된다. 특히 a) 분석의 效率를 위한 모델의 簡素化 b) 기온을 독립인자로 하는 傳達函數의 安靜性과 확률성분의 定常性, 그리고 c) 測定誤差를 설명하는 確率誤差의 白場性 등에 관심을 갖는다.

먼저, 回歸成分 $D_t^{(1)}$ 와 $D_t^{(1)}$ 의 4차 다항식 결정은 매우 타당하다. 3절에서 말한 바와 같이 2차와 3차식에서 발생하는 殘差제곱합이 4차에 비하면 너무 크기 때문이다. 이러한 결과는 母數推定에 대한 標準誤差의 검토 또는 回歸分析의 分散分析表(김해경, 1991b)에서도 확인된다.

다음으로, 정해진 형태로 볼 때 전달함수는 安靜하며 그리고 ϕ_1 의 추정에서 확률성분의 定常성을 위한 부등식 $|\phi_1| = |-0.4671| < 1$ 도 쉽게 성립된다.

마지막으로, 모델의 적용에서 얻어지는 마지막 殘差의 스펙트럼은 물론이고 그림 3에서 주어진 自己相關係數와 部分自己相關係數는 모두 白場時系列의 특성을 나타내고 있다. 더욱이, 그림 3로부터 계산되는 自由度 d.f. = k - 1 = 10 - 1 = 9를 갖는 Box-Pierce 카이-제곱검정량의 값이

$$\begin{aligned} Q(r) &= (N - k - p) \sum_{i=1}^{k=10} r_i^2 = (365 - 10 - 1) (0.0275) \\ &= 9.69 \end{aligned}$$

(단, r_i = 時差 i에서의 自己相關係數, N = 관찰값의 갯수, k = 自己相關係數가 零이 아닌 最大時差 그리고 p = 推定된 母數의 갯수)로 주어지고, 有意水準 0.05에서 같은 自由度 d.f. = 9의 이론적 카이-제곱값이 16.9임을 참작할 때 確率誤差의 白場性을 부정할 근거가 없다.

더욱이, 그림 4는 위에서 이용한 殘差와 시계열 $\{Z_t^{(1)}\}$ 사이의 交叉相關을 나타낸 것으로 時差에 관계없이 有意的이 아님을 보여준다. 또, 自由度 d.f. = h - r - s = 31 - 0 - 1 = 30를 갖는 카이-제곱 검정량의 값

$$\begin{aligned} Q(c) &= (N - h - p) \sum_{i=0}^{h=30} c_i^2 = (365 - 30 - 1) (0.1165) \\ &= 38.93 \end{aligned}$$

(단, c_i 는 時差 i에서의 交叉相關係數, k = 交叉相關係數가 零이 아닌 最大時差 그리고 r, s는 (6.1)에서 정의됨)을 有意水準 0.05의 이론분포의 값 43.77에 비교해 볼 때, 이러한 사실이 더욱 확실해진다. 이와같은 사실들이 잔차 $\{w_t\}$ 와 시계열 $\{Z_t^{(1)}\}$ 사이에 교차상관이 없음을 의미하므로 결국 모델의 적합성이 확인된다.

5.4 豫測方法의 數的事例

개발된 모델은 기온과 오염농도의 과거자료를 기초로 하여 미래의 오염농도수준을 결정하는데 이용된다. 이 예측모델은 回歸成分, 週期成分 그리고 確率成分으로 구성되지만 長期豫測에는 회귀 및 주기

Mean = -0.0018, Standard deviation = 0.3423

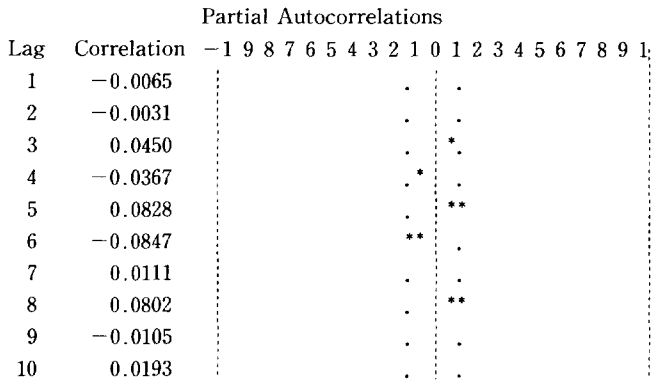
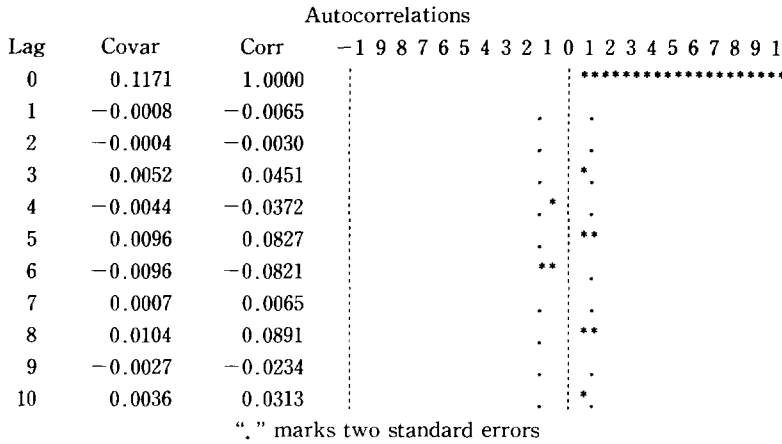


Fig. 3. Autocorrelation, partial autocorrelation of residuals from the fitted last model (5.1).

성분이 그리고 短期豫測에는 전달함수의 구조를 가진 종속성분이 적절이 대응하게 된다. 실제로, 장기 예측에서 종속성분은 큰 역할을 하지 못한다. 그리고 회귀성분과 주기성분만으로는 단기예측에서 만족스러운 결과를 기대하기 어렵다.

(5.2)로부터

$$(1-0.4671B) Y_t = (1-0.4671B) (\tilde{D}_t^{(y)} + \tilde{C}_t^{(y)}) + (1-0.4671B) (0.0720 - 0.0788B) [T_t - (\tilde{D}_t^{(T)} + \tilde{C}_t^{(T)})]$$

또는

$$Y_t = 0.4671 Y_{t-1} + P_t^{(y)} - 0.4671 P_{t-1}^{(y)} + 0.0720 R_t^{(T)} - 0.1124 R_{t-1}^{(T)} + 0.0368 R_{t-2}^{(T)} \tag{5.3}$$

단, $P_t^{(y)} = \tilde{D}_t^{(y)} + \tilde{C}_t^{(y)}$

$$= 4.4042 + 0.0259t - 0.0004t^2 + 0.000002t^3 - 0.000000003t^4 - 0.1803 \sin(-0.9659 - 0.0667t),$$

$$R_t^{(T)} = T_t - [\tilde{D}_t^{(T)} + \tilde{C}_t^{(T)}] = T_t - [2.1680 - 0.1363t + 0.0036t^2 - 0.0000160t^3 + 0.000000018t^4 - 1.3840 \cdot \sin(-0.6599 + 0.0509t)]$$

를 얻는다.

일 평균기온과 SO₂ 오염농도를 기초하여 유도된 이 방정식은 두 가지 방법 즉 “관찰값 이용방법” 그리고 “예측값 이용방법”으로 농도수준의 전방예측에 이용할 수 있다. 경험적으로, 1일-전방예측에는 前者의 방법이, 그리고 1일을 초과한 k일(k>1)-전방예측에는 後者가 적당하다. 그림 5(a)에서는 1989년도 잠실자료의 對數시계열과 예측예정기간을 표시하였다. 1년중 전반부 334일(1월 1일~11월 31일)의 측정값을 이용하여 335일부터 1개월간(12월 1일~12월 31일)을 위 방법들을 통하여 전방예측하고, 이 기간동안의 두 가지 예측값을 실제값과 동시에 그림 5(b)에 나타내었다. 이때 계산된 전방예측이란

Variance of input=6.991637

Lag	Covar	Corr	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1		
0	0.0318	0.03522												*											
1	0.0652	0.07208											.	*											
2	0.1397	0.15437										.	.	***											
3	0.1485	0.16416										.	.	***											
4	0.0841	0.09301										.	.	**											
5	0.0197	0.02186										.	.	.											
6	0.0058	0.00648										.	.	.											
7	-0.0152	-0.01683										.	.	.											
8	-0.0092	-0.01012										.	.	.											
9	0.0330	0.03650										.	.	*											
10	0.0744	0.08223										.	.	**											
11	0.0879	0.09716										.	.	**											
12	0.1001	0.11059										.	.	**											
13	0.0300	0.03317										.	.	*											
14	-0.0360	-0.03982										.	*	.											
15	-0.0027	-0.00302										.	.	.											
16	0.0087	0.00962										.	.	.											
17	0.0068	0.00761										.	.	.											
18	0.0195	0.02165										.	.	.											
19	0.0133	0.01473										.	.	.											
20	0.0577	0.06378										.	.	*											
21	0.0596	0.06585										.	.	*											
22	0.0440	0.04867										.	.	*											
23	0.0345	0.03818										.	.	*											
24	0.0063	0.00705										.	.	.											
25	0.0112	0.01241										.	.	.											
26	0.0027	0.00303										.	.	.											
27	-0.0265	-0.02927										.	*	.											
28	-0.0475	-0.05250										.	*	.											
29	-0.0064	-0.00710										.	.	.											
30	-0.0178	-0.01970										.	.	.											

"." marks two standard errors

Fig. 4. Cross covariance, cross correlation between $\{R_{T,t}^{(2)}\}$ and residuals from the fitted last model (5.1).

예를 들어 1일-전방의 경우에 먼저 334일까지의 자료를 이용하여 335일의 예측값을 계산하고, 또 다음 336일의 예측값은 334일까지의 자료에 335일의 (예측값이 아닌) 실제 측정값을 예측에 다시 참여시키는 방법으로 계산한다. 또, k일-전방($k=1, \dots, 31$) 예측값은 334일의 측정값과 이것을 기초로 하여 단계적으로 계산된 k-1일까지의 예측값들을 동시에 사용하여 얻는다. 그림 5(c)는 그림 5(b)의 예측결과를 指數變換에 의하여 본래 단위로 환원한 것이다.

傳達函數確率모델과 單變數確率모델의 예측효과는 수치결과로도 비교된다. 같은 자료에 대한 그림

5(b)의 결과를 김해경(1991a, 그림 5.8)의 결과를 비교하면 전달함수의 경우가 우월함이 확인된다. 이러한 전달함수모델의 우월성은 근본적으로 기온이 오염농도의 지배 또는 선도인자로서 무시될 수 없음을 의미한다.

6. 濃度豫測의 統計的 筋次

앞 절에서와 같이, 모델(5.2)는 평균기온을 기초로 한 일평균 SO₂ 수준의 豫測방정식을 구하는데 이용될 수 있다. 그러나, 김해경(1991a, b)에서 밝혀

진 바와 같이 氣溫과 汚染濃度の 回歸趨勢나 週期成分 그리고 從屬成分들은 일반적으로 각각 해당지역과 예측년의 고유한 특성을 갖게 되므로, 豫測방정식은 기본적으로 해당지역 예측년의 자료만으로 결정하는 것이 바람직하다.

豫測年 전반기 평균기온 $\{T_t\}$ 과 SO_2 의 對數농도 $\{Y_t\}$, $t=0, 1, \dots, n-1$ 을 기초로 한 후반기 오염농도의 일반적인 豫測節次는 다음과 같다.

1단계 : 일일 平均氣溫과 SO_2 汚染濃度の 각 回歸趨勢성분

$$D_t^{(v)} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4,$$

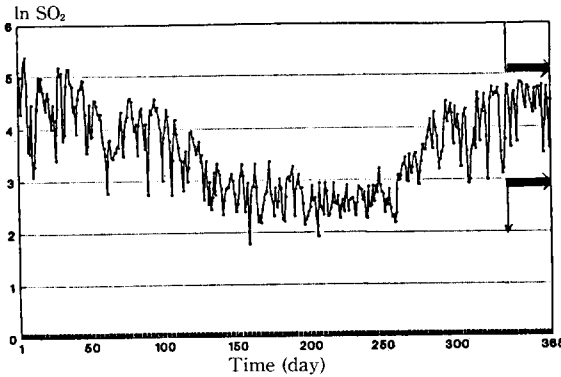
$$D_t^{(T)} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4$$

들의 係數 α_i, β_j ; $i, j=0, 1, 2, 3, 4$ 들에 대한 최소제곱 추정값을 구한다. 이때, 이 趨勢성분들은 非確率성분으로 간주하여 가능한 과거자료 모두를 이용한다.

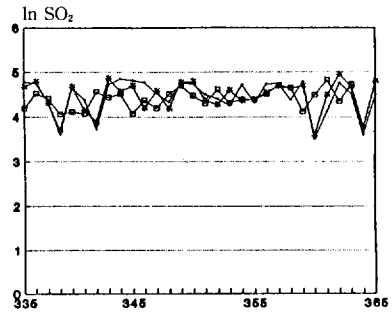
2단계 : 殘差 $\{R_{T,t}^{(v)}\}$ 와 $\{R_{Y,t}^{(v)}\}$, $R_{T,t}^{(v)} = T_t - D_t^{(T)}$, $R_{Y,t}^{(v)} = Y_t - D_t^{(v)}$ 의 피리오도그램이나 스펙트럼을 기초로하여 각 週期成分 $C_t^{(T)}, C_t^{(Y)}$ 의 함수형태

$$C_t^{(Y)} = \sum_{i=1}^{P_1} \alpha_{i0} \sin(\alpha_{i1} + \alpha_{i2} t),$$

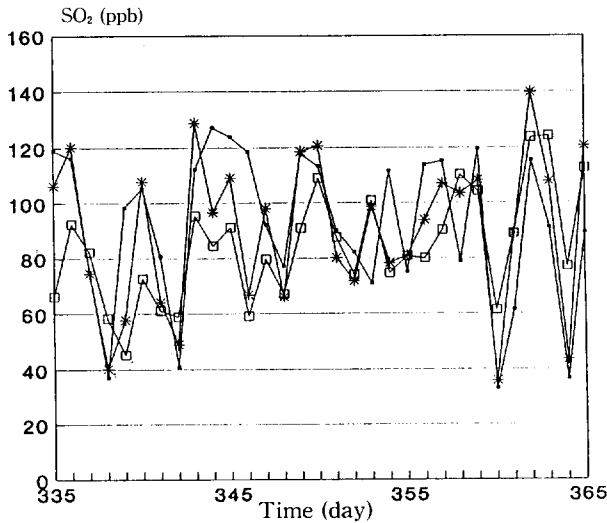
$$C_t^{(T)} = \sum_{j=1}^{P_2} \beta_{j0} \sin(\beta_{j1} + \beta_{j2} t)$$



(a)



(b)



(c)

Fig. 5. (a) Original series ($\ln SO_2$) (—●—) and prediction region: 1989, Jamsil.
 (b) Original series ($\ln SO_2$) (—●—) and predicted series using observations (—●—) and predicted values (—□—).
 (c) Original SO_2 series (—●—) and predicted series using observations (—●—) and predicted values (—□—).

를 결정한다. 여기에서 p_1, p_2 는 각각 對數 SO_2 농도와 평균기온에 있어서 週期성분의 갯수를 나타내는 해당년도의 固有 整數들이다.

3단계 : 殘差 $\{R_{T,t}^{(2)}\}$, $R_{T,t}^{(2)} = T_t - (D_t^{(T)} \hat{C}_t^{(T)})$ 로 기온의 從屬成分 $Z_t^{(T)}$ 를 AR(q) 모델

$$Z_t^{(T)} = \sum_{u=1}^q a_u Z_{t-u}^{(T)} + e_t^{(T)}$$

로 白場化하고, 같은 모델을 $\{R_{Y,t}^{(2)}\}$, $R_{Y,t}^{(2)} = Y_t - (D_t^{(Y)} \hat{C}_t^{(Y)})$ 에 적용하여 오염농도의 從屬成分 $\{Z_t^{(Y)}\}$ 도

$$Z_t^{(Y)} = \sum_{u=1}^q a_u Z_{t-u}^{(Y)} + e_t^{(Y)}$$

으로 白場化한다.

4단계 : 3단계에서 事前白場化된 $\{e_t^{(T)}\}$ 와 $\{e_t^{(Y)}\}$ 의 근사시계열들로 $\{R_{T,t}^{(3)}\}$ 와 $\{R_{Y,t}^{(3)}\}$ 을 두 식

$$R_{T,t}^{(3)} = R_{T,t}^{(2)} - \sum_{u=1}^q \hat{a}_u R_{T,t-u}^{(2)}$$

$$R_{Y,t}^{(3)} = R_{Y,t}^{(2)} - \sum_{u=1}^q \hat{a}_u R_{Y,t-u}^{(2)}$$

으로 계산하고, 이것들로 부터 추정되는 $\{e_t^{(T)}\}$ 와 $\{e_t^{(Y)}\}$ 의 交叉共分散을 이용하여 $\{Z_t^{(T)}\}$ 과 $\{Z_t^{(Y)}\}$ 사이의 衝擊反應值 $v_k (k=0, 1, \dots)$ 를 공식 (4.1)를 통해 얻는다.

5단계 : 殘差 $\{R_{T,t}^{(3)}\}$ 와 $\{R_{Y,t}^{(3)}\}$ 을 이용하여 얻어진 衝擊反應函數의 형태를 관찰하여 두 시계열 $\{Z_t^{(Y)}\}$ 와 $\{Z_t^{(T)}\}$ 사이에 존재하는 傳達函數의 모델

$$Z_t^{(Y)} = \frac{A(B)}{C(B)} Z_{t-b}^{(T)} + U_t$$

(단, $A(B) = (A_0 - A_1B - A_2B^2 - \dots - A_sB^s)$, $C(B) = (1 - C_1B - C_2B^2 - \dots - C_rB^r)$ 그리고 $BZ_t^{(T)} = Z_{t-1}^{(T)}$)의 階數 (b, r, s)을 결정한다. 그리고 衝擊反應函數와 傳達函數와 관계로부터 계수 母數 $A_i, C_j; i=0, \dots, s; j=1, \dots, r$ 의 값을 추정한다. 또, 여기에서 얻어지는 殘差의 標本 自己相關係數와 部分自己相關係數를 기초로 하여 誤差時系列 $\{U_t\}$ 에 ARMA모델

$$\phi(B)U_t = \theta(B)w_t$$

(단, $\phi(B), \theta(B)$ 는 각각 B의 다항식)을 적용하고 미지수를 추정한다.

6단계 : 傳達函數성분의 결정과정에서 얻어진 값들을 初期값으로 하여, 모델 (6.1)의 非線型최소제곱추정값들을 구하고, 1,2단계에서 결정된 각 성분

들로 豫測方程式을 완성한다.

$$\theta(B)^{-1} \phi(B) C(B) (Y_t - D_t^{(Y)} - C_t^{(Y)}) = \theta(B)^{-1} \phi(B) A(B) (T_t - D_t^{(T)} - C_t^{(T)}) \quad (6.1)$$

$$\text{단, } D_t^{(Y)} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4,$$

$$C_t^{(Y)} = \sum_{i=1}^{p_1} \alpha_{0i} \sin(\alpha_{1i} + \alpha_{2i} t),$$

$$D_t^{(T)} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4,$$

$$C_t^{(T)} = \sum_{i=1}^{p_2} \beta_{0i} \sin(\beta_{1i} + \beta_{2i} t).$$

7. 결 론

서울지방 대기중 SO_2 濃度の 統計 및 確率的 특징은 다음과 같다.

첫째, 일년 사계절중 기온과 SO_2 오염농도 사이에는 강한 지수감소의 非線型回歸關係가 존재한다. 겨울철은 풍속과 오염농도 사이에도 강한 지수감소의 非線型回歸關係가 존재하지만, 다른 계절에는 線型 또는 非線型的 어떤 回歸關係도 희박하다. 일년 전체로 보아, 기온은 大氣汚染濃度の 대표적인 독립인자이며, 겨울철에는 氣溫과 風速이 모두 大氣汚染濃度の 주요 독립인자가 된다.

둘째, 일년중 (對數) SO_2 汚染濃度는 4차다항식의 回歸趨勢를 가지며, 121.7일의 主週期和 해당년에 따라 변하는 副週期를 갖는 週期性도 내포한다. 또한, 이 오염농도는 階數가 1인 自己回己過程으로 설명되는 "1日從屬性"을 갖는다. 이는 SO_2 오염농도의 영향 지속기간이 평균적으로 1일 정도임을 의미한다.

셋째, 氣溫과 大氣汚染濃度 사이에는 時差 1에서 陰의 交叉相關이 내존하며 時差 0에서도 弱한 有意的인 陽의 相關이 존재한다. 이것은 오늘의 오염농도는 어제의 기온에 의하여 직접적인 영향을 받으며 어제의 기온하락(상승)은 오늘의 오염농도의 상승(하락)을 가져온다. 그러나, 오늘의 기온상승이 당일인 오늘의 오염농도를 증가시키는 (기상학적인) 요인도 다소 내포하고 있음을 의미한다.

넷째, SO_2 汚染濃度の 確率모델은 (6.1)로 주어지며, 이 모델을 이용한 농도에측의 統計的 절차는 5節에서와 같다.

<본 논문은 연세대학교 학술연구비(91-104)의 지원으로 이루어졌음>

비하여 여름과 겨울에 특히 심하기 때문이다.

이다.

1. 氣溫과 SO₂ 濃度와의 回歸關係

기온과 SO₂ 농도의 상관계수는 계절에 따라 다소 차이는 있지만 모두 아주 낮은 값을 나타내며 이 사실은 두 변량 사이의 종속관계가 비선형관계를 의미하기도 한다. 그림 A.1은 최저기온과 SO₂ 농도의 對數값을 나타낸 것이다. 여기에서 보는 바와 같이 기온과 對數농도 사이에는 사계절 전체적으로 강한 선형의 회귀관계가 존재함을 볼 수 있다. 다른 지역의 자료에서도 같은 결과를 얻을 수 있다. 이러한 사실은 기온(T)과 오염농도(S) 사이에는 지수감소의 비선형 회귀관계인

$$S = k_1 e^{-k_2 T}$$

가 존재함을 의미한다.

1989. 1. 1~1989. 12. 31 자료에 대한 최소제곱직선은

$$\log S = 4.1968 - 0.0073 T$$

이고, 그리고 그것의 유의성검정결과는 표 A.2에서 주어진다. 따라서 최저기온과 SO₂ 농도 사이의 최소제곱 회귀지수곡선은

$$S = 66.473 e^{-0.0073 T}$$

2. 風速과 SO₂ 농도와의 回歸關係

풍속과 SO₂ 농도의 상관계수도 계절에 따라 다소 차이는 있지만 대부분 낮은 값을 나타낸다. 특히, 對數 SO₂ 농도와 최저기온 사이의 상관계수는 사계절 각각 -0.29(봄), -0.46(여름), -0.19(가을), -0.78(겨울)으로 겨울철만이 비교적 강한 선형관계를 보이고 있다. 이 사실은 겨울철에는 SO₂ 농도와 기온 사이에 어떤 감쇄율을 갖는 지수함수의 비선형회귀관계가 존재함을 의미한다. 실제로 잠실 자료의 경우, 그림 A.2에서 보는 바와 같이 對數 SO₂ 농도와 풍속(W) 사이에는 강한 선형관련이 존재하고, 이때 계산된 최소제곱추정직선은

$$\ln S = 5.1772 - 0.0809 W$$

이며, 이 결과로 얻어지는 풍속과 SO₂ 농도사이의 지수함수의 회귀관계는

$$S = 177.1859 e^{-0.0809 W}$$

이다. 따라서, 겨울철의 경우 풍속과 SO₂ 농도와의 일반적 회귀관계는

$$S = c_0 e^{-c_1 W}$$

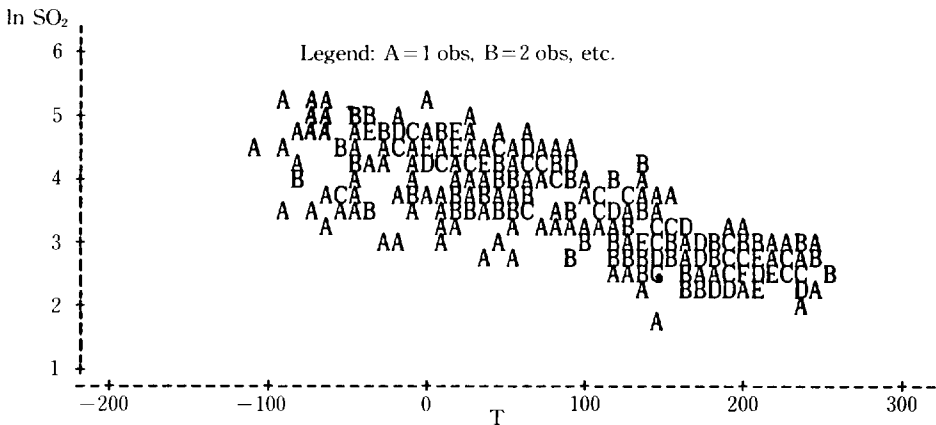


Fig. A.1 Plot of ln SO₂ and temperature T.

Table A.2 SAS output for estimation of regression parameters.

Parameter	Estimate	T for H ₀ : Parameter=0	Pr> T	Std Error of Estimate
INTERCEPT	4.1968	117.59	0.0001	0.03569
t(=TEM)	-0.0073	- 26.36	0.0001	0.00027

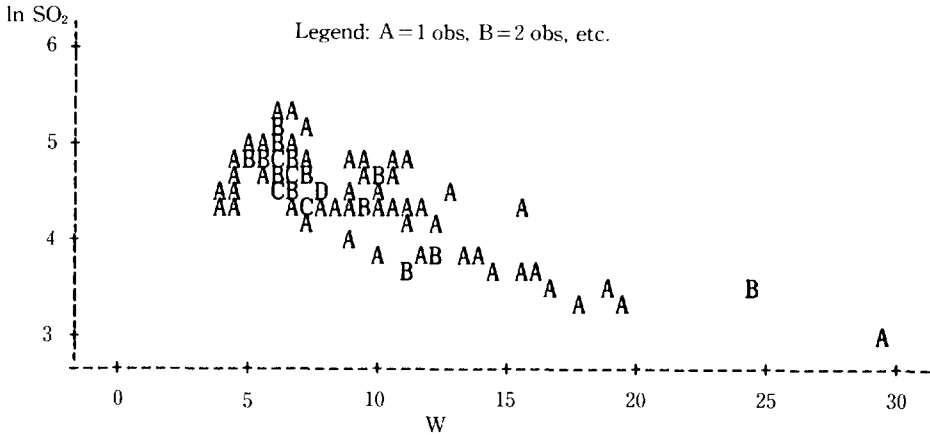


Fig. A.2 Plot of $\ln SO_2$ and wind speed W in winter.

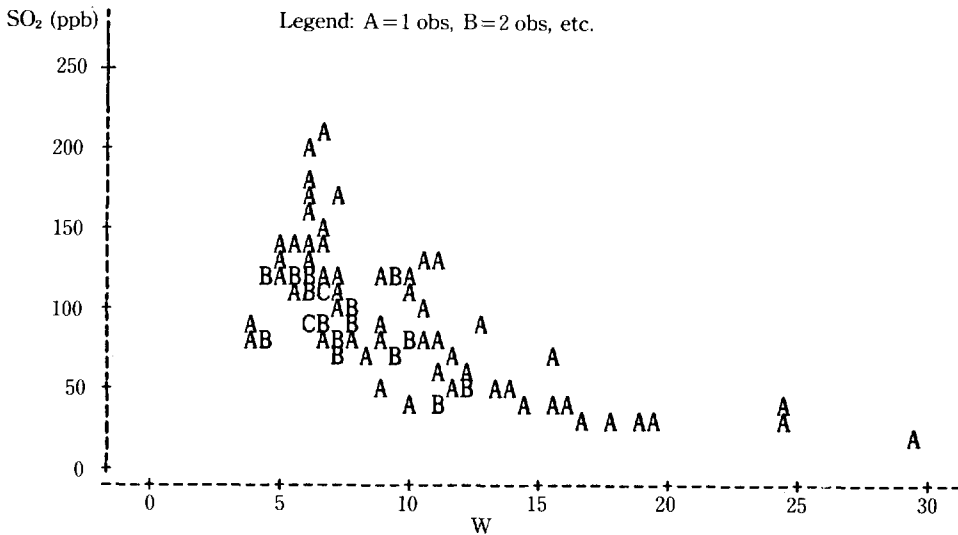


Fig. A.3 Plot of SO_2 and wind speed W in winter.

이다. 그러나 이것이 사계절전체의 종속관계를 설명하지는 않는다.

참 고 문 헌

김해경 (1991a) 大氣汚染濃度에 관한 確率모델. 韓國大氣保全學會誌, 第 7 卷 2號, 127 - 136.

김해경, 김태수 (1991b) 서울지방 겨울기온에 관한 확률 모델. 응용통계연구. (인쇄중)

Bartlett, M.S. (1935) Some aspects of the time-correlation problem in regard to tests of significance. Jour. Royal Stat. Soc. 98, 536.

Box, G.E.P. & Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco: Holdon-Day, 337 - 380.