

공간자료에 대한 지리적 가중회귀 모형과 크리깅의 비교*

김선우¹⁾ 정애란²⁾ 이성덕³⁾

요약

공간상관이 존재하는 지리통계 자료(geostatistical data)에 대하여 일반적으로 널리 사용되는 Kriging 모형과 통계학적 공간자료 분석모형인 지리적 가중회귀 모형을 고려하고, 미지의 위치에 대한 예측력을 비교해 본다. 두 모형의 예측력을 검토하기 위하여 환경부 자료를 실증사례로 활용한다. 전국의 116개 대기오염 측정망에서 얻은 1999년의 월별 일산화탄소(Co/ppm) 자료의 평균을 구하여 Kriging모형과 지리적 가중회귀 모형에 적합하고 미지의 위치를 예측하여 예측오차제곱합(PRESS)으로 각각의 방법에 대한 예측성능을 비교한다.

주요용어: 공간통계학, 크리깅, 세미베리오그램, 지리적가중회귀모형, 예측오차제곱합

1. 서론

공간통계학에서 사용되는 공간자료는 특정 위치 또는 특정 지역에서의 관측값이나 측정값으로 구성되며 지리통계 자료(geostatistical data), 격자 자료(lattice data), 공간 점 패턴 자료(spatial point patterns data)의 세가지 유형으로 분류된다. 지리통계 자료는 연속적인 공간상의 고정된 위치에서 얻은 측정값들의 집합으로 그 위치는 일반적으로 공간상에서 연속적이고, 격자 자료는 공간 지역과 관계되는 자료로 공간 지역은 규칙적이거나 불규칙적이다. 공간 점 패턴 자료는 위치 자체가 관심의 대상이 되는 변수로서, 공간 지역 내에서 관측된 유한한 위치의 집합으로 구성된다. 본 논문에서는 지리통계 자료를 중심으로 연구하고자 한다. Matheron(1963)은 Kriging을 이용한 예측방법에서 베리오그램을 추정할 때 실제 표본 관측값을 거리의 상관성으로 나타내주는 표본 베리오그램을 제시하였고, Cressie와 Hawkins(1980)는 이상점의 효과를 최소로 할 수 있는 로버스트 베리오그램을 제안하였다. 한편 공간자료에 대한 통계적 모형에 의한 분석의 연구로 McMillen과 McDonald(1998)는 몬테 카를로에 의한 증명과 실제 자료분석을 통하여 지리적 가중 회귀 모형의 가중치로 사용되는 국소적 가중 최대 우도 추정량을 제안하였고, Brunson(1996)은 거리에 기초하는 가중치를 이용한 지리적 가중회귀 모형을 제안하였으며, Lesage(2001)는 지리적 가중회귀

* 본 연구는 산학협동재단의 2003년도 학술 연구비 지원에 의해 연구되었음.

1) (361-763) 충북 청주시 흥덕구 개신동 12번지, 충북대학교 정보통계학과 박사과정

E-mail: statg023@hanmail.net

2) (361-763) 충북 청주시 흥덕구 개신동 12번지, 충북대학교 정보통계학과 박사과정

E-mail: arjeong@netian.com

3) (361-763) 충북 청주시 흥덕구 개신동 12번지, 충북대학교 정보통계학과, 기초과학연구소 교수

E-mail: sdlee@chungbuk.ac.kr

모형의 문제점을 해결하기 위하여 베이지안 방법을 고려하였고 지리적 가중 회귀 모형에 의하여 추정된 최소 제곱법의 결과를 향상시켰다. 최승배와 강창완(2001)은 변수가 2개인 경우 교차베리오그램에 대한 감도분석의 결과와 제안된 주성분점수를 이용한 감도분석의 결과를 모의실험과 실제자료분석을 통하여 비교하였다. 장지희 등(2003)의 논문은 보통크리깅이 거리역수 크리깅보다 우수한 공간예측방법임을 서울시 27개지역 2000년 오존도를 가지고 보였다. 공간 자료 분석의 주요 문제는 n 개의 관측값 $z(S_\alpha)$ 의 정보에 기초해서 관측되지 않은 값 $z(S_0)$ 를 예측하는데 있기 때문에 본 논문은 환경부의 일산화탄소 자료를 이용하여 크리깅 방법과 지리적 가중회귀 분석방법과의 예측정도를 비교하고자 한다. 본 논문의 2장에서는 Kriging에 의한 공간자료 분석을 위하여 공간 자료의 모형을 먼저 제시하고, Kriging의 세 단계를 통하여 공간예측의 문제를 해결한다. 3장에서는 지리적 가중회귀 모형에 의한 공간 자료분석을 다루기 위해 지리적 가중회귀 모형을 소개한다. 4장에서는 실제 환경자료를 사용하여 Kriging에 의한 공간 자료 분석 방법과 지리적 가중회귀에 의한 공간자료 분석방법으로 공간예측을 하고 PRESS 기준을 통하여 두 방법간의 예측력을 비교한 후 마지막으로 5장에서 결론 및 향후 연구과제를 제시한다.

2. Kriging에 의한 공간자료 분석

2.1. 공간자료

공간자료는 확률과정(stochastic process) $Z(s)$ 의 실현치 이고, s 가 D 안에서의 위치이며, D 는 R^d 의 부분집합으로 위치 s 의 전체집합이다. R^d 를 d 차원의 유클리드 공간이라 할 때, 확률벡터 $z(s)$ 는 다음과 같이 표현된다.

$$\{z(s) : s \in D \subset R^d\} \quad (2.1)$$

여기서 d 는 보통 1,2,3의 값을 가진다. 공간자료는 일반적으로 $(z_i, s_i), i = 1, \dots, n$ 으로 표현되고 여기서 z_i 는 위치 s_i 에서 흥미 있는 현상의 i 번째의 관측 값이다.

공간자료의 분석시, 추론을 위하여 다음 식(2.2)의 정상성 가정이 필요하다.

$$\begin{aligned} E(Z(s+d) - Z(s)) &= 0 \\ \text{Var}(Z(s+d) - Z(s)) &= 2\gamma(d), \quad s, s+d \in D \end{aligned} \quad (2.2)$$

여기서 $2\gamma(d)$ 는 베리오그램을 나타낸다.

2.2. 베리오그램 추정

표본 베리오그램은 각 표본간의 거리 및 이들 관측값의 상관성을 나타낸 것으로, 각 표본들을 나열한 후 추정이 필요한 지점에서 거리가 가까운 표본위치 지점부터 가장 먼 표본위치 지점까지 거리에 포함된 표본들간의 상관관계를 나타내어준다. 이 표본 베리오그램은 등방성이라는 가정하에서 실제 표본 관측값을 거리의 상관성으로 나타내어준 것으로

Matheron(1963)이 제안하였으며, 다음과 같다.

$$\hat{\gamma}(d) = \frac{1}{2N_d} \sum_{N(d)} (z(s_i) - z(s_j))^2 \tag{2.3}$$

여기서, $N(d)$ 는 유클리드 거리 $|s_i - s_j| = d$ 를 가지는 모든 쌍들의 집합으로서 $N(d) = \{(i, j) : |s_i - s_j| = d\}$ 이고, N_d 는 $N(d)$ 에 포함되는 서로 다른 원소 쌍들의 수이고, $z(s_i)$ 와 $z(s_j)$ 는 각각의 공간 위치 s_i, s_j 에서의 관측값이다. 베리오그램이 불등방일 때, 방향성 표본 베리오그램은 식 (2.3)에서 d 대신에 벡터 \underline{d} 를 사용하여 계산된다. 만약 자료에 이상점이 존재한다면 이러한 이상점을 제거하지 않으면서 그것의 효과를 최소화할 수 있도록 Cressie와 Hawkins(1980)는 다음의 식 (2.4)와 같은 로버스트 베리오그램 추정량을 제안하였다.

$$\bar{\gamma}(d) = \frac{\left[\frac{1}{2N_d} \sum_{N(d)} |z(s_i) - z(s_j)|^{1/2} \right]^4}{0.457 + 0.494/N_d} \tag{2.4}$$

2.3. 이론적인 베리오그램 모형의 적합

표본 베리오그램의 정보를 이용하여 공간 데이터를 추정하기 위해 이론적인 베리오그램 모형을 적합시킨다. 대부분의 이론적인 베리오그램 모형에는 동일한 위치의 표본간에 존재하는 분산인 Nugget 효과, 세미베리오그램이 더 이상 증가되지 않는 때의 값으로 분산을 나타내는 Sill과 더 이상 상관이 없어지는 거리인 Range와 같은 몇몇의 모수들이 사용된다. 이론적인 베리오그램 모형은 다음과 같다.

지수모형	$\hat{\gamma}(d; \underline{\theta}) = \begin{cases} 0, & d = 0 \\ c_0 + c_1[1 - \exp(-d/a)], & d \neq 0 \end{cases} \tag{2.5}$ $\underline{\theta} = (c_0, c_1, a), \quad c_0 : \text{Nugget effect}, \quad c_1 : \text{Sill}, \quad a : \text{Range}$
구형모형	$\hat{\gamma}(d; \underline{\theta}) = \begin{cases} c_0 + c_1 \left[\left(\frac{3}{2} \right) \left(\frac{d}{a} \right) - \left(\frac{1}{2} \right) \left(\frac{d}{a} \right)^3 \right], & d \leq a \\ c_0 + c_1, & d > a \end{cases} \tag{2.6}$ $\underline{\theta} = (c_0, c_1, a), \quad c_0 : \text{Nugget effect}, \quad c_1 : \text{Sill}, \quad a : \text{Range}$
가우시안모형	$\hat{\gamma}(d; \underline{\theta}) = \begin{cases} 0, & d = 0 \\ c_0 + c_1[1 - \exp(-d^2/a^2)], & d \neq 0 \end{cases} \tag{2.7}$ $\underline{\theta} = (c_0, c_1, a), \quad c_0 : \text{Nugget effect}, \quad c_1 : \text{Sill}, \quad a : \text{Range}$
선형모형	$\hat{\gamma}(d; \underline{\theta}) = \begin{cases} 0, & d = 0 \\ c_0 + c_1 d, & d \neq 0 \end{cases} \tag{2.8}$ $\underline{\theta} = (c_0, c_1), \quad c_0 \geq 0, \quad c_1 \geq 0$
파워모형	$\hat{\gamma}(d; \underline{\theta}) = \begin{cases} 0, & d = 0 \\ c_0 + b d^\lambda, & d \neq 0 \end{cases} \tag{2.9}$ $\underline{\theta} = (c_0, b, \lambda), \quad c_0 \geq 0, \quad b \geq 0, \quad \lambda < 2, \quad (\text{단}, \lambda \neq 0)$

2.4. 공간예측

공간 분석에서 중요한 문제 중에 하나는 관측된 자료를 근거로 하여 관측되지 않은 위치 값을 예측하는 것이다. Kriging은 확률변수의 공간적 변화성을 설명하는 베리오그램에 따라 선택된 가중치와 자료간의 선형결합형태로 임의의 공간상 위치에서의 값을 추정하는 선형추정방법중의 하나로 가중치는 모형의 평균오차가 이라는 가정하에, 오차의 분산이 최소가 되도록 모형화한다.(Isaaks and Srivastava, 1989) 공간자료를 분석하기 위한 Kriging의 일반선형 모형은 다음과 같다.

$$z(s) = m(s) + \epsilon(s) \quad (2.10)$$

여기서, $m(s)$ 는 방향(direction)이나 추세를 의미하는 모형의 평균으로 $m(s)$ 의 선형구조에 따라 Universal Kriging, Ordinary Kriging, Simple Kriging등으로 분류되고, $\epsilon(s)$ 는 오차항이다. Ordinary kriging은 모형의 평균을 알수 없지만 일정하다는 가정 하에 추정하는 방법으로 널리 이용되고 있다.

$$z(s) = \mu + \epsilon(s) \quad (2.11)$$

미지의 관측장소 s_0 에서 Z 의 추정치 $\hat{Z}(s_0)$ 를 추정하기 위하여 다음의 세 가지 조건을 가정한다.

① $\hat{Z}(s_0)$ 는 n 개의 관측위치 s_1, \dots, s_n 에서의 관측값 $Z(s_1), \dots, Z(s_n)$ 의 선형함수이다.

$$\hat{Z}(s_0) = \sum_{i=1}^n w_i Z(s_i) \quad (2.12)$$

② 예측모형의 계수의 합은 1이다.

$$\sum_{i=1}^n w_i = 1 \quad (2.13)$$

③ 예측값 $\hat{Z}(s_0)$ 은 가장 작은 평균제곱 예측오차(MSPE) $E(Z(s_0) - \hat{Z}(s_0))^2$ 를 갖는다.

앞의 세 조건하에서 얻어진 추정량은 최량선형불편예측량, 즉 BLUP(best linear unbiased predictor)임이 Robinson(1991)에 의해 증명되었다.

3. 지리적 가중회귀 모형에 의한 공간자료 분석

지리적 가중회귀 모형(GWR : Geographically weighted regression model)은 커널 회귀에 기초하고 있는데, 공간에서의 모든 점의 국소적 선형 회귀를 구하기 위해서는 자료의 이웃해 있는 관측값들의 거리를 가중치로 사용해야만 한다.

GWR 모형은 다음과 같이 나타낼 수 있다.

$$W_i y = W_i X \beta_i + \epsilon, \quad i = 1, \dots, n \quad (3.1)$$

여기서, y 는 공간상의 n 개 위치에서 뽑힌 종속변수 관측치들의 $n \times 1$ 벡터 이고, X 는 독립변수의 $n \times k$ 행렬이며, ϵ 은 상수 분산을 갖는 정규분포를 따르는 $n \times 1$ 벡터 이다. W_i 는 i 번째 관측값과 다른 모든 관측값들 사이의 거리를 반영하는 i 번째 관측값의 거리에 기초한 가중치 d_i 벡터를 포함하는 $n \times n$ 대각 행렬로 표현되고, β_i 는 i 번째 관측치에 관련있는 $k \times 1$ 모수 벡터이다. 가중회귀 최소제곱 추정량 $\hat{\beta}_i$ 는 다음과 같다.

$$\hat{\beta}_i = (X'W_i^2X)^{-1}(X'W_i^2y) \quad (3.2)$$

i 번째 관측값의 GWR모형에서 사용되는 거리에 기초한 가중치는 표본에서 i 번째 관측치와 다른 모든 관측치 사이의 거리 벡터에 기초해서 결정되는 벡터 $W_{i,(j)}$, $j = 1, 2, 3$ 의 식으로 표현된다. 거리 퇴화 모수(distance decay parameter)를 가지는 이 거리벡터는 공간적인 자료 표본에서 이웃에 있는 관측값들로부터의 상대적인 가중치를 나타내는 가중 함수를 만드는데 사용된다.

여러 가중 함수를 수립하는 방법 중에서 Brunsdon et al (1996)에 의해 제안된 다음의 가중함수를 사용하는 지리적 가중회귀모형을 GWR1로 쓴다.

$$W_{i,(1)} = \sqrt{\exp(-d_i/\theta)} \quad (3.3)$$

여기서 모수 θ 는 퇴화모수(decay parameter) 혹은 bandwidth 모수로서 이러한 bandwidth의 변화는 지수함수의 모양을 변화시킨다.

또 다른 가중치 설계로 McMillen과 McDonald(1998)에 의해 제안된 tri-cube 함수를 사용하는 모형은 GWR2 모형이라 하기로 한다.

$$W_{i,(2)} = (1 - d_i/q_i)^3 I(d_i < q_i) \quad (3.4)$$

여기서 q_i 는 i 번째 관측값의 q 번째 가까이에 있는 관측값과 i 번째 관측값간의 거리이고, $I()$ 는 지시함수이다.

Lesage(2001)는 Gaussian 함수 ϕ 에 의존하는 가중회귀의 가중치를 제안하였고, 이 가중치를 사용하는 모형을 GWR3으로 명명하고자 한다.

$$W_{i,(3)} = \phi(d_i/\sigma\theta) \quad (3.5)$$

여기서 ϕ 는 표준 정규 분포이고, σ 는 거리 벡터 d 의 표준편차이며, d_i 는 i 번째 관측치와 다른 모든 관측값들간의 거리 벡터로 표현된다.

모수 θ 의 값은 국소적 선형 회귀 방법에서 종종 사용되는 교차분석(cross validation)을 사용해서 결정할 수 있다.

$$\sum_{i=1}^n [y_i - \hat{y}_{\neq i}(\theta)]^2 \quad (3.6)$$

여기서 $\hat{y}_{\neq i}(\theta)$ 는 측정 과정으로부터 생략된 i 번째 위치의 관측점을 가지는 y_i 의 적합된 값이다.

4. 실증분석

본 장에서는 실제 환경자료를 사용하여 앞에서 제안된 Kriging에 의한 공간 자료 분석 방법(Ordinary Kriging, Universal Kriging)과 지리적 가중회귀에 의한 공간자료 분석방법(GWR)으로 공간예측을 하고 PRESS 기준을 통하여 두 방법간의 예측력을 비교하고자 한다.

4.1. 자료설명

본 논문에서는 환경부 홈페이지의 우리나라 116개 관측소에서 관측된 월 평균 일산화탄소(CO)자료를 실증분석에 이용하였다. 자료는 116개 관측소의 위치를 나타내는 TM좌표와 측정된 일산화탄소의 양으로 구성되어 있다. 자료의 유형은 격자화 되지 않은 지리통계 자료(nongridded geostatistical data)로서 관측값 간의 거리에 따라 상관관계가 다른 자료의 유형이다. Kolmogorov-Smirnov검정 결과 유의확률이 0.066으로서 우리나라 일산화탄소 자료가 정규분포를 따르고 있음을 알 수 있고, 그 자세한 분포는 아래의 그림 4.1의 히스토그램에서 확인할 수 있다. 그림 4.2는 대기 오염도를 측정하기 위한 전국 116곳의 측정망의 위치를 나타내고 있다.

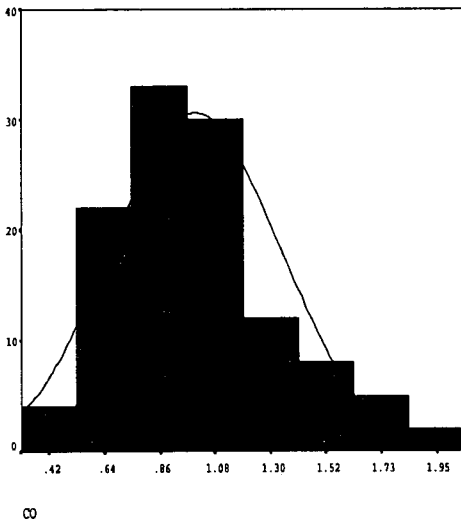


그림 4.1: 일산화탄소 히스토그램

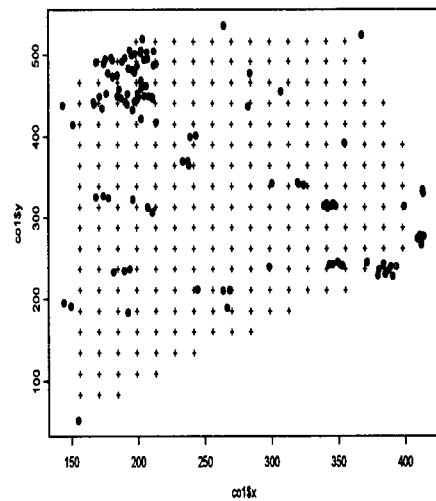


그림 4.2: 대기오염 측정망 위치

4.2. 베리오그램 추정

그림 4.3은 모든 방향을 고려한 베리오그램 추정그림이고, 그림 4.4는 식 (2.4)의 로버스

트 베리오그램 추정량을 이용한 로버스트 베리오그램 결과이다. 그림 4.3과 그림 4.4를 비교해보면 자료의 변동폭이 줄어드는 것을 확인할 수 있다.

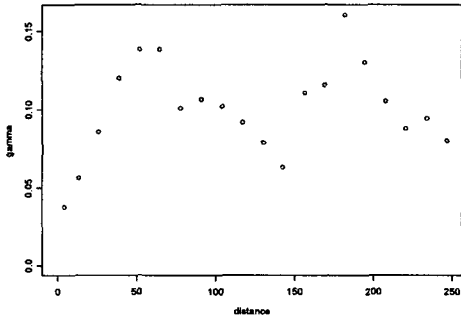


그림 4.3: 일산화탄소 자료의 베리오그램 추정

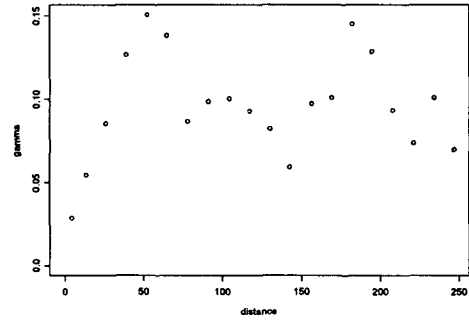


그림 4.4: 일산화탄소 자료의 로버스트 베리오그램

그림 4.5는 실제 자료의 0°, 45°, 90°, 135°방향에 대한 방향성 베리오그램 결과를 보여주고 있다. 방향이 다른 4개의 그림이 모두 비슷한 결과를 보이고 있으므로 일산화탄소 자료는 등방성을 만족함을 알 수 있다.

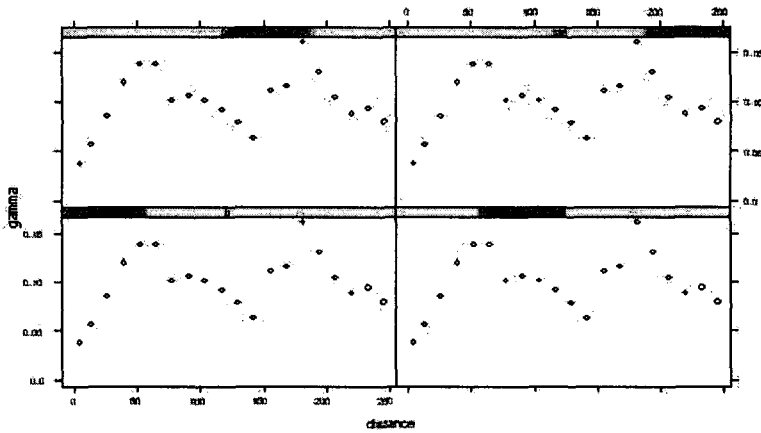


그림 4.5: 일산화탄소 자료의 방향성 베리오그램

4.3. 베리오그램의 이론적 모형 적합

추정된 베리오그램을 가장 잘 설명하는 이론적 모형을 찾기 위하여 모수를 변화시키면

서 구형모형, 지수모형, 가우시안 모형에 적합시킨 결과 range=40, sill=0.09, nugget=0.02인 구형 모형 일 때 최소 오차 제곱합을 갖는다. 따라서 다음의 구형모형을 이론적 베리오그 램 모형으로 사용한다.

$$\hat{\gamma}(d, \theta) = \begin{cases} 0.02 + 0.09 \left[\left(\frac{3}{2} \right) \left(\frac{d}{40} \right) - \left(\frac{1}{2} \right) \left(\frac{d}{40} \right)^3 \right], & d \leq 40 \\ 0.02 + 0.09, & d > 40 \end{cases} \quad (4.1)$$

4.4. 예측 및 예측력 비교

공간통계 분석 방법과 일반통계 분석 방법에 대해 각 모형별로 전체 116개 자료 중 1개를 제외한 115개 자료를 이용하여 모형을 적합하고 그 나머지 1개의 자료를 예측하였다. 이것을 반복 수행하여 전체 116개의 자료를 예측하였다. Kriging에 의한 분석방법(Ordinary Kriging, Universal Kriging)과 지리적 가중회귀에 의한 분석방법의 예측성능을 다음의 식 (4.2)와 같은 예측 오차 제곱합(PRESS : PRediction Error Sum of Square)기준을 이용하여 비교한다.

$$PRESS = \sum_{i=1}^n \{Z(s_i) - \hat{Z}(s_i)\}^2 \quad (4.2)$$

여기서, $Z(s_i)$ 는 i 번째 위치에 있는 관측값이고, $\hat{Z}(s_i)$ 은 i 번째 위치에 있는 관측값을 제외하고 나머지 관측값으로 i 번째 위치에 있는 값을 예측한 값이다. 공간통계 분석 방법과 지리적 가중회귀분석 방법 각각에 대한 PRESS 값은 다음의 표 4.1와 같다. Ordinary Kriging과 Universal Kriging의 Kriging에 의한 공간자료 분석방법이 GWR1이나 GWR의 지리적 가중회귀에 의한 공간자료 분석방법보다 PRESS가 더 작음을 알 수 있으나 $W_{i,(2)}$ 를 가중치로 사용하는 GWR 모형의 PRESS 값은 Kriging 결과보다 더 작다. 이것은 거리에 따라 상관관계가 달라지는 자료의 특성을 가중치 $W_{i,(2)}$ 가 잘 반영하기 때문인 것으로 보인다.

표 4.1: PRESS를 이용한 예측성능 비교결과

예측방법	Kriging 모형		지리적 가중회귀		
	Ordinary Kriging	Universal Kriging	GWR1	GWR2	GWR3
PRESS	8.40605	8.350389	10.450796	7.4006622	10.453901

5. 결론

우리나라 일산화탄소 자료는 nugget effect=0.02, sill=0.09, range=40인 구형 모형으로 가장 잘 적합되었다. 지리적 가중회귀에 의한 분석방법과 Kriging에 의한 분석방법의 예측

성능을 PRESS 기준으로 비교한 결과 Ordinary Kriging과 Universal Kriging의 Kriging에 의한 공간자료 분석방법이 GWR1이나 GWR3의 지리적 가중회귀에 의한 공간자료 분석 방법보다 PRESS가 더 작게 나타났다. 그러나 $W_{i,(2)}$ 를 가중치로 사용하는 GWR2 모형은 Kriging 결과보다 더 작은 PRESS값을 가졌다. 이것은 거리에 따라 상관관계가 달라지는 자료의 특성을 가중치 $W_{i,(2)}$ 가 잘 반영하기 때문인 것으로 보인다. 실증 분석 결과 거리에 따라 상관이 있는 지리 통계 자료의 경우에는 Kriging에 의한 공간자료 분석방법을 사용하는 것이 지리적 가중회귀에 의한 공간자료 분석방법을 사용하는 것보다 더 타당하다라는 결과를 얻었다. 그러나 본 논문에서 활용된 tri-cube 함수를 가중치로 사용하는 GWR2모형은 Kriging에 의한 분석방법보다 더 작은 PRESS 값을 나타냄으로써 공간자료 분석에 적합한 것으로 보인다.

참고문헌

- 장지희, 홍태영, 남궁평 (2003), Spatial data analysis using the kriging method, <한국통계학회 논문집>, 10, 423-432.
- 최승배, 강창완 (2001), 주성분점수를 이용한 이변량 공간자료에 대한 감도분석, <응용통계연구>, 14, 415-427.
- Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. (1996). Geographically weighted regression : A method for exploring spatial non stationarity, *Geographical analysis*, 28, 281-298.
- Cressie, N. and Hawkins, D. M. (1980). Robust estimation of the variogram, *Mathematical Geology*, 12, 115-125.
- Isaaks, E. h. and Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*. Oxford university press., New York.
- Lesage, J. P. (2001). A family of Geographically Weighted Regression Models, Technical Report.
- Matheron, G(1963). Principles of geostatistics, *Economic Geology*, 58, 1246-1266.
- McMillen, Daniel P. and John F. McDonald. (1998). Locally weighted maximum likelihood estimation : Monte Carlo evidence and an application, Paper presented at the Regional Science Association International meetings, Santa Fe, NM.
- Robinson, G.K. (1991). That BLUP is a good thing: The estimation of random effects, *Statistical Science*, 6, 15 -51.

[2004년 8월 접수, 2004년 12월 채택]

Comparison between Kriging and GWR for the Spatial Data*

Sun-Woo Kim ¹⁾ Ae-Ran Jeong ²⁾ Sung-Duck Lee ³⁾

ABSTRACT

Kriging methods as traditional spatial data analysis methods and geographical weighted regression models as statistical analysis methods are compared. In this paper, we apply data from the Ministry of Environment to spatial analysis for practical study. We compare these methods to performance with monthly carbon monoxide observations taken at 116 measuring area of air pollution in 1999.

Keywords: Spatial statistics, Kriging, Semiveriogram, Geographically weighted regression model, PRESS

* This research was supported by Korea Sanhak Foundation(2003).

- 1) Graduate Student, Department of Information and Statistics, Chungbuk National University, 12 Gaesin-dong, Cheongju, Chungbuk, 361-763, Korea
E-mail: statg023@hanmail.net
- 2) Graduate Student, Department of Information and Statistics, Chungbuk National University, 12 Gaesin-dong, Cheongju, Chungbuk, 361-763, Korea
E-mail: arjeong@netian.com
- 3) Professor, Department of Information and Statistics, Basic science research institute, Chungbuk National University, 12 Gaesin-dong, Cheongju, Chungbuk, 361-763, Korea
E-mail: sdlee@chungbuk.ac.kr