

현장 조사 자료를 이용한 GIS 기반 주제도 작성을 위한 단변량 크리깅 기법의 비교

박 노 욱[†]

인하대학교 지리정보공학과

Comparison of Univariate Kriging Algorithms for GIS-based Thematic Mapping with Ground Survey Data

No-Wook Park[†]

Dept. of Geoinformatic Engineering, Inha University

Abstract : The objective of this paper is to compare spatial prediction capabilities of univariate kriging algorithms for generating GIS-based thematic maps from ground survey data with asymmetric distributions. Four univariate kriging algorithms including traditional ordinary kriging, three non-linear transform-based kriging algorithms such as log-normal kriging, multi-Gaussian kriging and indicator kriging are applied for spatial interpolation of geochemical As and Pb elements. Cross validation based on a leave-one-out approach is applied and then prediction errors are computed. The impact of the sampling density of the ground survey data on the prediction errors are also investigated. Through the case study, indicator kriging showed the smallest prediction errors and superior prediction capabilities of very low and very high values. Other non-linear transform based kriging algorithms yielded better prediction capabilities than traditional ordinary kriging. Log-normal kriging which has been widely applied, however, produced biased estimation results (overall, overestimation). It is expected that such quantitative comparison results would be effectively used for the selection of an optimal kriging algorithm for spatial interpolation of ground survey data with asymmetric distributions.

Key Words : Interpolation, Log-normal kriging, Indicator kriging, Multi-Gaussian kriging, Ground survey.

요약 : 이 연구의 목적은 비대칭 분포를 가지는 현장 조사 자료로부터 GIS 기반 주제도를 생성하기 위한 공간 내삽 방법으로 단변량 크리깅 기법을 비교하는데 있다. 기존 정규 크리깅과 비선형 자료 변환에 기반을 둔 로그 정규 크리깅, 다중 가우시안 크리깅과 지시자 크리깅을 지화학 원소 비소와 납에 대해 사례 연구를 통해 비교하였다. 예측 능력의 비교 분석을 위해 leave-one-out 기반 교차 검증을 통한 오차 분석을 수행하였으며, 샘플링 밀도의 차이에 따른 오차의 변화 양상도 분석하였다. 비교 분석 결과, 지시자 크리깅이 전반적으로 가장 높은 예측 능력을 나타내었으며, 작은 값과 높은 값의 예측 능력도 우수한 것으로 나타났다. 정규 크리깅에 비해 비선형 자료 변환 기반 크리깅 기법들이 우수한 예측 능력을 나타내었지만, 기존에 많이

적용된 로그 정규 크리깅은 샘플링 밀도와 상관없이 편향 정도가 가장 크게 나타내었다. 이 연구를 통해 얻어지는 정량적 검증 결과는 비대칭 분포를 가지는 현장 조사 자료의 내삽을 위한 크리깅 기법의 선정에 유용하게 이용될 수 있을 것으로 기대된다.

1. 서론

각종 지구과학 주제도는 수치화를 통해 GIS 기반의 데이터베이스로 구축되어 각종 분석에 사용되고 있다. 이러한 GIS 기반 주제도는 대부분 현장 조사 자료나 원격탐사 자료와 같은 부가 자료를 이용하여 제작되고 있다. 특히 연구지역에서 전역적인 정보 획득이 가능한 원격탐사 자료로부터 지형, 토지피복, 식생 정보 등의 다양한 주제정보를 포함하는 주제도 제작이 활발하게 이루어지고 있다. 그러나 원격탐사 자료로부터 지구과학 여러 분야에서 요구하는 모든 주제정보의 획득이 가능한 것은 아니기 때문에, 전통적인 방법인 현장 조사 자료를 이용한 GIS 기반 주제도 작성 역시 중요한 부분을 차지한다.

직접적으로 현장 조사 자료를 주제도 작성에 이용하는 경우, 접근성, 시간, 비용 등의 한계로 일부 지역에서만 현장 자료의 획득이 가능하기 때문에 주제도 작성을 위해 각종 내삽 과정을 통해 주제도를 생성한 후에 2차 분석을 수행하게 된다. 그러므로, 2차 분석의 신뢰성을 보장하기 위해서는 현장 자료가 획득하지 못한 지점에서의 정확한 속성값의 추정이 필수적이다. 일반적으로 부분적으로 획득된 자료로부터 전역적인 래스터 주제도를 생성하기 위해 다양한 내삽 기법들이 적용되어 왔다. 역거리 가중치법(inverse distance weighting: IDW), cubic spine 등 다양한 내삽 방법들이 적용되어 왔는데, 이러한 내삽 방법들은 대부분 GIS S/W에서 기본 기능으로 제공되고 있다. 이러한 일반적인 내삽 방법론 이외에, 샘플 자료의 공간적 자기상관성에 기반을 둔 고급 내삽 방법론으로 간주할 수 있는 크리깅 기법도 많이 적용되고 있다. 일반적인 내삽 방법들과 다르게, 크리깅 기법은 자료에 존재하는 이방성(anisotropy)을 베리오그램 모델링을 통해 내삽 과정에 반영할 수 있으며, 원격탐사 자료와 같은 부가 자료의 통합이 가능한 장점이 있다(Goovaerts, 1997). 최근에는 현장 자료가 획득되지 못한 지점에서의 속성값의 추정에 따르는 불확실성을 줄이기 위해 다변량 크리깅을 이용한 원격탐사 자료,

수치고도자료 등과 같은 부가 자료의 통합 연구가 국내 외에서 활발하게 적용되고 있다(정상용, 이강근, 1995; 박노옥, 장동호, 2008; Goovaerts, 2000; Glenn and Carr, 2003; Sales et al., 2007; Park et al., 2009). 특히 최근에는 ArcGIS 지구통계학 분석 모듈, GSLIB (Deutsch and Journel, 1998), GSTAT (Pebesma and Wesseling, 1998), SGeMS (Remy et al., 2009) 등과 같은 다양한 상용/비상용 지구통계학 분석 S/W의 이용이 가능해짐에 따라 크리깅 분석은 다양한 분야에서의 주제도 작성을 위한 유용한 방법론으로 적용될 수 있다.

하나의 변수만을 고려하는 단변량 크리깅 기법에는 전통적인 단순 크리깅(simple kriging, SK), 정규 크리깅(ordinary kriging, OK), 일반 크리깅(universal kriging, UK) 등이 있다(Deutsch and Journel, 1998). 또한 원 자료의 비선형 변환에 기반을 둔 로그 정규 크리깅(log-normal kriging, LK), 다중 가우시안 크리깅(multi-Gaussian kriging, MG), 지시자 크리깅(indicator kriging, IK) 등이 적용될 수 있다. 이러한 다양한 크리깅 기법의 존재는 사용자 입장에서는 선택의 폭이 넓다는 장점이 될 수도 있지만, 반대로 적절한 크리깅 방법론의 선택이 어려울 수도 있다. 크리깅 방법론의 선택 기준은 크리깅 기법의 이론적 배경의 이해도 필요로 하지만, 처리하고자 하는 자료의 특성도 고려해야 한다. 현장 조사 자료의 분포 특성은 속성값의 특성에 따라 좌우되지만, 정규 분포와 같은 대칭분포를 따르지 않고 양이나 음으로 치우친 비대칭 분포를 나타내는 경우가 많이 존재한다. 또한 관측 센서의 탐지 분해능의 한계로 인해 탐지 한계를 벗어난 관측값은 모두 특정 한계값으로 관측되는 경우도 존재하며, 관측 오차가 아닌 이상값 또한 자료값에 포함될 수 있다.

비대칭 분포나 이상값을 가지는 자료의 경우에 대해 현재까지 기존 일반적인 크리깅 기법과 비선형 변환 기반 크리깅 기법의 비교는 외국에서 일부 수행된바 있지만, 국내에서는 비교 연구가 아직 수행되지 못하였다. Paprits and Dubois(1999)는 중금속 오염도 작성을 위해 정규 크리깅과 비선형 변환 기반 크리깅 기법을 짚나

이프(jackknife) 방법을 통해 비교하였는데, 각 기법별 예측 정확도의 차이가 크지 않으며 현장 자료의 샘플링 밀도에 예측 정확도가 의존적이라는 결과를 보고하였다. 반면, Saito and Goovaerts(2000)는 다이옥신 오염 지역의 600지점에서 획득된 자료에 비선형 변환 기반 크리깅 기법을 적용하고 역시 잭나이프 방법을 통해 예측 능력을 비교하였는데, 로그 정규 크리깅과 지시자 크리깅이 가장 높은 예측 정확도를 나타내었고, 샘플링 밀도의 변화에 정규 크리깅이 가장 크게 영향을 받는다고 보고하였다. 그러나 이러한 결과는 단일 자료에 적용한 분석 결과로 일반화를 위해서는 추가 연구가 필요하다. 국내에서는 조홍래와 정종철(2006, 2009)이 강우 자료와 미세먼지(PM10)의 내삽을 위해 역거리 가중치법, radial basis function과 정규 및 일반 크리깅을 적용하였는데, 크리깅 기법이 다른 내삽방법에 비해 높은 예측 정확도를 나타낸다고 보고하였다. 그러나 이 연구에서는 현장 자료의 분포 특성을 고려한 비선형 변환 기반 크리깅 기법의 비교 분석을 수행하지 않았다.

이 연구에서는 현장 자료를 이용한 GIS 기반 주제도 작성을 위한 크리깅 기법의 적용에서 극단적인 비대칭 분포나 이상값을 가지는 자료의 내삽을 위해 비선형 변환 기반 크리깅 기법의 예측 능력을 비교하고자 하였다. 이를 위해 강원도 태백산 광화대내의 253개 지점에서 획득된 지하화 조사 자료중에서 비소(As)와 납(Pb) 2개의 자료에 대해 비선형 변환 기반 크리깅 기법을 적용한 사례 연구를 수행하였다. 예측 능력의 비교를 위해 leave-one-out 기반의 교차 검증을 수행하였으며, 특히 자료값의 분포 특성을 고려하기 위해 자료 구간별 예측 오차를 추가적으로 비교분석하였다. 또한 현장 자료의 샘플링 밀도에 따른 예측 오차의 변화 양상도 함께 분석하였다. 이러한 분석 과정을 통해 비대칭 분포나 이상값을 가지는 현장 자료를 이용한 공간 내삽에 적합한 크리깅 기법을 제시하고자 하였다.

2. 단변량 크리깅 기법 및 비교 방법론

지구통계학적 분석의 주요 엔진으로 간주할 수 있는 크리깅은 일반화된 최소자승 회귀분석 방법론에 기반을 둔 내삽기법으로 알려져 있지만, 이론적으로는 미지의

관측지점에서의 속성값은 그 위치에서 정의된 확률 변수의 결과들로 모델링된다는 확률 함수의 개념에 기반을 두고 있다(Isaaks and Srivastava, 1989; Goovaerts, 1997). 이 때, 확률 함수는 공간적 상관성을 가지는 확률 변수의 조합으로 구성되고, 공간적 상관성은 정상성(stationarity)의 특성하에 일반적으로 베리오그램을 통해 크리깅 시스템에 반영된다.

n 개의 샘플 자료 $\{z(\mathbf{u}_\alpha), \alpha = 1, \dots, n\}$ 가 주어졌다고 할 때, 샘플링 되지 않은 미지의 위치 \mathbf{u} 에서의 크리깅 예측치 $z^*(\mathbf{u})$ 는 주변 관측값의 선형 가중 합으로 표현되는데, 적용하는 크리깅 기법에 따라 가중치 결정 등에서 차이가 발생한다. 이 논문에서는 여러 단변량 크리깅 기법중에서 일반적으로 많이 적용되던 정규 크리깅과 비선형 자료 변환 기반 크리깅 기법으로 로그 정규 크리깅, 다중 가우시안 크리깅, 지시자 크리깅을 비교분석하였다. 여기서 비선형 변환은 원 자료에 비선형 변환을 수행했다는 의미이며, 실제 크리깅 적용시에는 주변 관측값의 선형 가중합으로 미지의 값을 예측하게 된다. 이 논문에서는 각 크리깅 기법별 간단한 적용 식에 대해서만 기술하고자 하며, 자세한 이론은 Goovaerts(1997)와 Chil's and Delfiner(1999)를 참조할 수 있다.

1) 정규 크리깅(OK)

미지의 위치 \mathbf{u} 에서의 정규 크리깅 추정치 $z^*_{OK}(\mathbf{u})$ 는 주변 관측값 $z(\mathbf{u}_\alpha)$ 들의 선형 결합으로 식 (1)과 같이 표현된다.

$$z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{OK}(\mathbf{u})z(\mathbf{u}_\alpha) \quad (1)$$

식 (1)에서 정규 크리깅 가중치 $\lambda_{\alpha}^{OK}(\mathbf{u})$ 는 추정 분산 $\text{VAR}\{z^*_{OK}(\mathbf{u}) - z(\mathbf{u})\}$ 을 최소화하면서, 추정의 불편향 조건인 $E\{z^*_{OK}(\mathbf{u}) - z(\mathbf{u})\} = 0$ 을 만족하도록 결정되어야 하는데, 다음 식 (2)의 정규 크리깅 시스템을 통해 결정된다.

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{OK}(\mathbf{u})\gamma(\mathbf{u}_\alpha - \mathbf{u}_\beta) - \mu_{OK}(\mathbf{u}) = \gamma(\mathbf{u}_\alpha - \mathbf{u}) \quad \alpha = 1, \dots, n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{OK}(\mathbf{u}) = 1 \end{cases} \quad (2)$$

여기서 $\mu_{OK}(\mathbf{u})$ 와 $\gamma(\cdot)$ 는 라그랑지안 파라미터와 두 지점 간 거리에 따른 베리오그램 값을 각각 나타낸다.

정규 크리깅을 수행할 경우, 크리깅 예측치 $z^*_{OK}(\mathbf{u})$ 와 더불어 식 (3)에 주어진 오차 분산도 함께 얻어진다.

$$\sigma^2_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{OK}(\mathbf{u})y(\mathbf{u}_{\alpha} - \mathbf{u}) - \mu_{OK}(\mathbf{u}) \quad (3)$$

2) 로그 정규 크리깅(LK)

로그 변환은 극단적으로 큰 값이 존재하고, 0값이 자료에 포함되지 않을 경우, 원자료의 분포를 대칭형태로 변환시키는 일반적인 변환 방법으로 많이 활용되어왔다. 로그 정규 크리깅은 이러한 로그 변환된 자료값에 대해서 정규 크리깅을 적용한 경우에 해당한다.

로그 변환된 공간에서의 관측치 $y^*_{LK}(\mathbf{u})$ 는 역시 로그 변환된 주변 관측값 $y(\mathbf{u}_{\alpha})$ 들의 선형 결합으로 식 (4)로 표현된다. 식 (4)의 크리깅 가중치는 로그 변환된 공간에서 정규 크리깅 식 (2)에 따라 계산되며, 차이점은 베리오그램이 로그 변환된 자료로부터 계산된다는 것이다.

$$y^*_{LK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{LK}(\mathbf{u})z(\mathbf{u}_{\alpha}) \text{ with } y(\mathbf{u}_{\alpha}) = \ln[z(\mathbf{u}_{\alpha})] \quad (4)$$

정규 크리깅을 적용하는데 있어서 원 자료의 분포에 대한 가정(정규 분포 등)을 필요로 하지는 않지만, 일반적으로 원 자료의 분포가 강한 치우침을 보이지 않을 경우 우수한 예측 능력을 나타내는 것으로 알려져 있다(Saito and Goovaerts, 2000). 따라서 극단값의 영향을 로그 변환에 의해 줄임으로써 로그 변환된 공간에서의 정규 크리깅을 적용하게 된다. 그러나 원 자료 공간으로의 역변환을 위해 지수 함수 변환을 적용할 경우 편향된 결과를 나타내는 것으로 알려져 있으며, 역변환은 다음 식을 통해 수행하게 된다(Journal and Huijbregts, 1978).

$$z^*_{LK}(\mathbf{u}) = \exp[y^*_{LK}(\mathbf{u}) + \sigma^2_{OK}(\mathbf{u}) + \mu_{LK}(\mathbf{u})] \quad (5)$$

즉 원 자료 공간으로의 역변환에는 로그 공간에서의 정규 크리깅 결과이외에 분산($\sigma^2_{LK}(\mathbf{u})$)과 라그랑지안 파라미터($\mu_{LK}(\mathbf{u})$)를 필요로 한다. 변환 대상 값이 증가할수록 변환값이 상당히 급격하게 증가하는 지수 함수의 특성상, 역변환된 로그 정규 크리깅 결과는 변환식에 사용되는 3가지 항의 오차에 영향을 많이 받을 가능성이 크다.

3) 다중 가우시안 크리깅(MG)

로그 변환 원자료의 분포를 대칭형태로 변환시키는 일반적인 변환 방법이지만, 로그 변환을 통해 완전한 대칭 분포인 정규 분포를 얻을 수는 없다. 반면 정규 점수

변환(normal score transform)은 원 자료의 크기순 재배열을 통해 얻어지는 누적 분포 함수를 평균 0, 분산 1인 표준 정규 분포에 일대일 대응시킴으로써 원 자료의 분포 모양에 상관없이 표준 정규 분포로의 변환을 수행한다(Deutsch and Journel, 1998).

다중 가우시안 크리깅은 이렇게 정규 점수 변환된 자료에 대해서 단순 크리깅 혹은 정규 크리깅을 수행하는 기법을 의미한다. 이 연구에서는 정규 크리깅과의 비교를 위해 가우시안 공간에서 정규 크리깅을 적용하였다. 정규 점수 변환된 자료값을 $y(\mathbf{u}_{\alpha})$ 라고 할 때, 가우시안 공간에서의 추정치 $y^*_{MG}(\mathbf{u})$ 는 식 (6)으로 주어지며, 베리오그램 모델링을 정규 점수 변환된 자료값에 대해 수행하는 것을 제외하고는 정규 크리깅의 과정과 동일하다.

$$y^*_{MG}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{MG}(\mathbf{u})y(\mathbf{u}_{\alpha}) \quad (6)$$

식 (6)에 주어진 가우시안 공간에서의 추정치는 역 정규점수 변환(normal score back transform)을 통해 원 자료 공간의 추정값 $z^*_{MG}(\mathbf{u})$ 로 변환된다.

4) 지시자 크리깅(IK)

지시자 크리깅은 연속형 변수의 경우 특정 임계치의 초과 혹은 미만 여부에 따라 원 자료값을 0과 1로 이진 변환한 후에 크리깅을 수행하여 특정 임계치보다 작거나 같게 될 조건부 확률을 계산함으로써, 특정 위치에서의 조건부 누적 확률 분포 함수(conditional cumulative distribution function: ccdf)를 모델링하는 방법론에 해당한다(Fig. 1).

우선 지시자 크리깅을 적용하기 위해서는 식 (7)에 의해 원 자료값을 K 개의 특정 임계치 z_k 에 의해 0과 1의 값을 가지도록 하는 지시자 변환을 수행한다.

$$i(\mathbf{u}_{\alpha}; z_k) = \begin{cases} 1 & \text{if } z(\mathbf{u}_{\alpha}) \leq z_k \quad k = 1, \dots, K \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

원 자료의 범위를 이산화시킨 일련의 임계치 z_k 에 대한 조건부 확률을 지시자 크리깅을 통해 얻게 되며, 이로부터 원자료의 ccdf를 모델링하여 미지의 위치 \mathbf{u} 에서의 불확실성을 확률론적으로 표현하게 된다(식 8).

$$\begin{aligned} [i(\mathbf{u}; z_k)]^* &= [F(\mathbf{u}; z_k | (n))]^* = [\text{Prob}\{Z(\mathbf{u}) \leq z_k | (n)\}]^* \\ &= \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}(\mathbf{u}; z_k) i(\mathbf{u}_{\alpha}; z_k) \end{aligned} \quad (8)$$

여기서 (n) 은 특정 위치 \mathbf{u} 주변의 자료값을 의미하며,

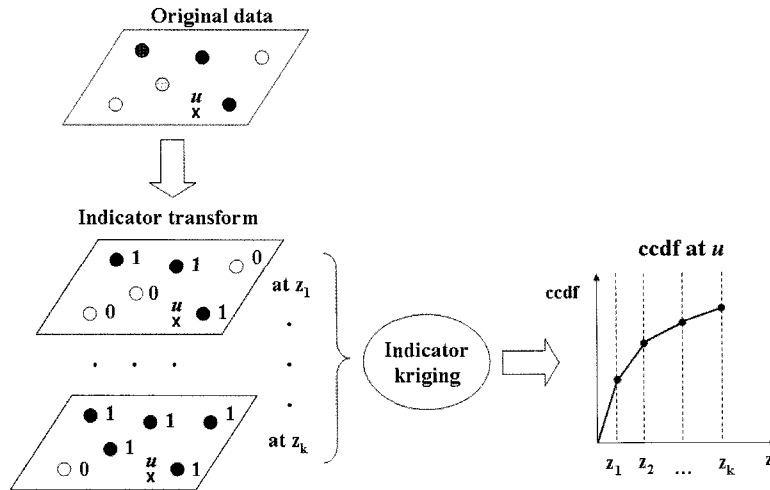


Fig. 1. Procedure for indicator kriging.

$F(\cdot)$ 는 특정 임계치 z_k 에서의 ccdf를 나타낸다. 식 (8)의 우변에서의 크리깅 가중치는 정규 크리깅을 적용하는 과정에 동일하며, 지시자 변환값의 베리오그램을 대입하는 차이가 있다.

K 개의 임계치에 대한 크리깅을 각각 수행하게 되면, 특정 위치에서의 ccdf를 얻게 되는데, 이로부터 추정값은 일반적으로 ccdf의 평균값을 통해 얻게 된다. 이 연구에서는 다중 가우시안 크리깅에서 적용한 방식과 마찬가지로 식 (9)에 의해 전체 자료 구간을 100개로 이산화시킨 후 얻어지는 분위수 $z_p(u)$ 의 합을 통해 추정값을 계산하였다.

$$z^*_{IK}(u) = \frac{1}{100} \sum_{p=1}^{100} z_p(u) \text{ with } p = l/100 - 0.5/100 \quad (9)$$

지시자 크리깅을 수행하기 위해서는 K 개의 베리오그램 모델링과 K 개의 크리깅 시스템을 풀어야 한다. 베리오그램 모델링을 한 번만 수행하는 다른 크리깅 방법과의 실제 수행면에서 번거로움이 다르게 되는데, 이 연구에서는 지시자 크리깅과 더불어 임계치를 중앙값만을 사용함으로써 베리오그램 모델링과 크리깅 시스템을 한 번만 수행하는 중앙값 지시자 크리깅(median indicator kriging: MIK)을 함께 적용하였다.

5) 정량적 예측 능력 비교 방법론

크리깅 기법은 이미 자료값이 알려진 원 자료의 위치에서 그 값이 그대로 재생되기 때문에, 별도의 검증 자료가 있지 않는한 기법별 예측 능력을 비교할 수 없다.

이 연구에서는 각 크리깅 기법별 예측 능력의 비교를 위해 leave-one-out 기반의 교차 검증을 수행하였다. 이 방법에서는 n 개의 샘플 자료 $\{z(u_\alpha), \alpha = 1, \dots, n\}$ 에서 임의로 1개를 제거한 후에, 나머지 $(n-1)$ 개의 자료를 가지고 제거한 지점에서의 값을 유추하는 과정을 n 번 반복한다. 이러한 과정을 통해 n 개의 지점에서의 예측값 $\{z^*(u_\alpha)\}$ 을 얻게 되고, 이 값과 원 참값을 비교하여 추정 오차를 계산한다.

이러한 추정 오차를 정량적으로 계산하기 위해서 이 연구에서는 평균 오차(mean error: ME)와 평균 절대 오차(mean absolute error: MAE)를 식 (10)에 의해 계산하였다. 평균 오차를 계산함으로써, 과추정 혹은 저추정 양상 즉 편향 정도를 파악할 수 있으며, 평균 절대 오차 계산을 통해 오차의 크기를 판단할 수 있다.

$$ME = \frac{1}{n} \sum_{\alpha=1}^n z^*(u_\alpha) - z(u_\alpha) \quad (10)$$

$$MAE = \frac{1}{n} \sum_{\alpha=1}^n |z^*(u_\alpha) - z(u_\alpha)|$$

3. 연구 지역 및 사용 자료

이 연구에서는 다양한 크리깅 기법의 비교를 위해 이진수 외(1998)에 의해 강원도 태백산 광화대 내의 예미, 호명, 임계, 정선 1:50,000 지질도폭에 포함되는 지역의 253개 하천수에서 획득된 지하화 조사 자료 분석 결과 중에서 비소(As)와 납(Pb) 2개의 원소에 대한 사례연

구를 수행하였다 일반적으로 지화학도 작성에 사용되는 매체 중에서 하천수는 하나의 시료가 시료채취지점 상류지역 지질 및 오염 정보 등을 가장 잘 반영하는 대표성이 다른 종류의 시료에 뛰어나 최근 광역 지화학 조사에 가장 널리 사용되는 매체로 알려져 있다(이진수 외, 1998). 이진수 외(1998)에 따르면, 시료채취 위치 설정을 위해 실내 위치 설정 후 야외 현장에서 시료채취 지점 주변의 지세, 지질, 광화작용 여부, 하천 및 퇴적물의 제반 현황과 오염상태 및 오염원을 파악한 후에, 되도록 인위적인 오염원을 피하여 시료를 채취하였다고 보고되었다. 두 원소의 측정 단위는 ppb로, 두 원소 모두 연구 지역의 북쪽과 남쪽 지역에서 낮은 값들이 존재하고, 중앙부에서 상대적으로 높은 값들이 분포하고 있다. 또한 연구 지역의 우측 하단에 이상값들이 존재하고 있다 (Fig. 2). 여기서 주의할 점은 이 사례연구의 목적은 연구 대상지역의 지화학 분포도 작성을 통한 대상지역의 오염원, 토양 및 암석 등의 영향 파악이 아니라, 비대칭

분포를 나타내는 현장 조사 자료를 이용한 주제도 작성에 적용되는 단변량 크리깅 기법의 비교 검토에 있다.

크리깅 기법의 적용 이전에, 원 자료의 통계치 분석을 통해 자료의 특성을 파악하고자 하였다. Fig. 3은 비소와 납 원소의 히스토그램과 기본 통계치를 나타내고 있다. 두 원소 모두 센서 분석 한계치인 1 ppb이하 자료가 나타나고 있으며, 대부분 하위 25%의 값이 중앙값보다 작은 18 ppb와 13 ppb로 각각 나타나고 있다. 또한 두 원소 모두 평균과 표준편차가 거의 유사하게 나타나고 있는데, 이는 대부분의 원소값보다 상당히 큰 이상값의 영향때문에 나타난 결과로 이러한 이상값 때문에 두 원소 모두 양으로 치우친 분포로 양의 왜도(skewness)값을 나타내고 있다. 두 원소의 기본 통계값의 분포 양상은 유사하지만, 납 원소와 다르게 비소 원소의 히스토그램 분포는 이봉 구조를 나타내고 있다. 요약하면, 두 원소 모두 자료값들이 분석 한계치 주변의 값과 이상값의 영향으로 대칭 분포가 아닌 양으로 치우친 비대칭 분포를

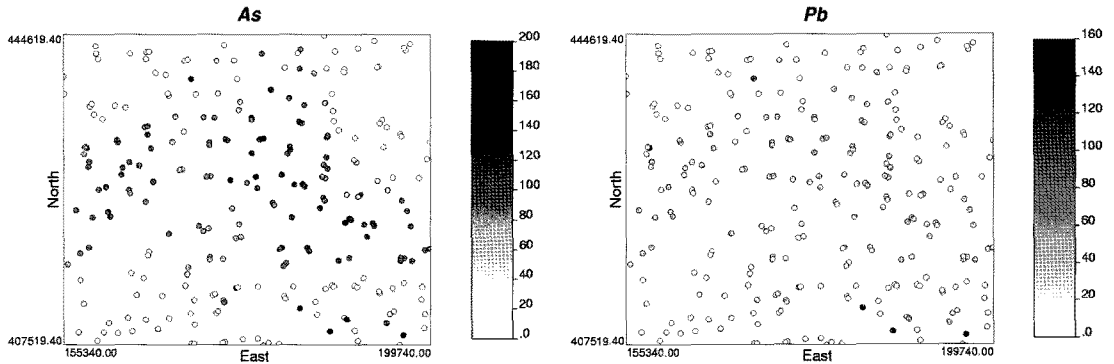


Fig. 2. Location maps of As and Pb elements in the study area.

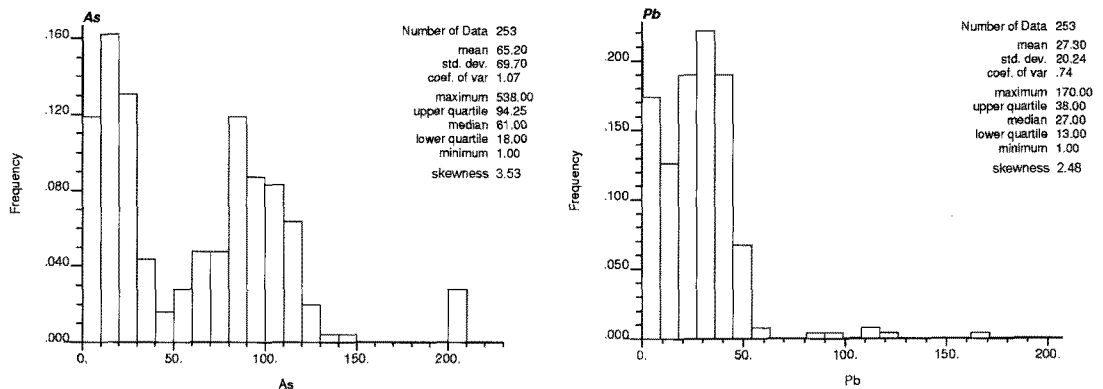


Fig. 3. Histogram and descriptive statistics of As and Pb elements.

나타내고 있다. 이러한 비대칭 분포는 원자료를 그대로 사용하는 정규 크리깅과 비선형 자료 변환에 기반한 크리깅 기법의 예측 능력에 영향을 미칠 것으로 예상된다.

4. 비교 분석 결과

크리깅 기법의 적용 이전에, 각 원소별로 원 자료, 로그 변환된 자료, 정규점수 변환된 자료에 대해 실험점 베리오그램을 작성한 후에, 베리오그램 모델링을 수행하였다. 지시자 변환을 위한 임계치는 100분위수로 설정하여 9개의 임계치별로 베리오그램 모델링을 수행하였다. 중앙값 지시자 크리깅에서는 중앙값에 대한 베리오그램 모델링만 수행하였다. 실험적 베리오그램을 작성한 결과, 두 원소 모두에서 이방성은 나타나지 않아서 등방성 베리오그램 모델링을 수행하였다. 크리깅 수행

시 최대 사용 자료의 개수는 자료의 공간적 분포를 고려하여 12개로 설정하였으며, 최대 탐색 반경은 각 베리오그램의 상관거리를 고려하여 최대 20,000미터로 하였다. 다중 가우시안 크리깅과 지시자 크리깅에서 원 자료 값 사이에서의 값 추정과 최소값과 최대값을 넘어서는 경우에 필요한 추정값과 cdf의 내삽에는 선형 모델을 이용하였다. 이러한 크리깅 분석에는 포트란 프로그래밍과 GSLIB(Deutsch and Journel, 1998)을 이용하였다.

1) 예측 능력의 비교 분석

Fig. 4는 비소와 납 원소의 교차 검증을 통해 얻어진 일부 크리깅 기법별 예측값과 참값의 산점도를 나타내고 있다. 두 원소에서 모두 오차 분산을 최소화 하는 과정에서 작은 값들은 크게, 큰 값들은 작게 예측하면서 전반적인 분산이 참값에 비해 작게 나타나는 크리깅의 일반적인 특성이 나타나고 있다. 결과적으로 크리깅 기

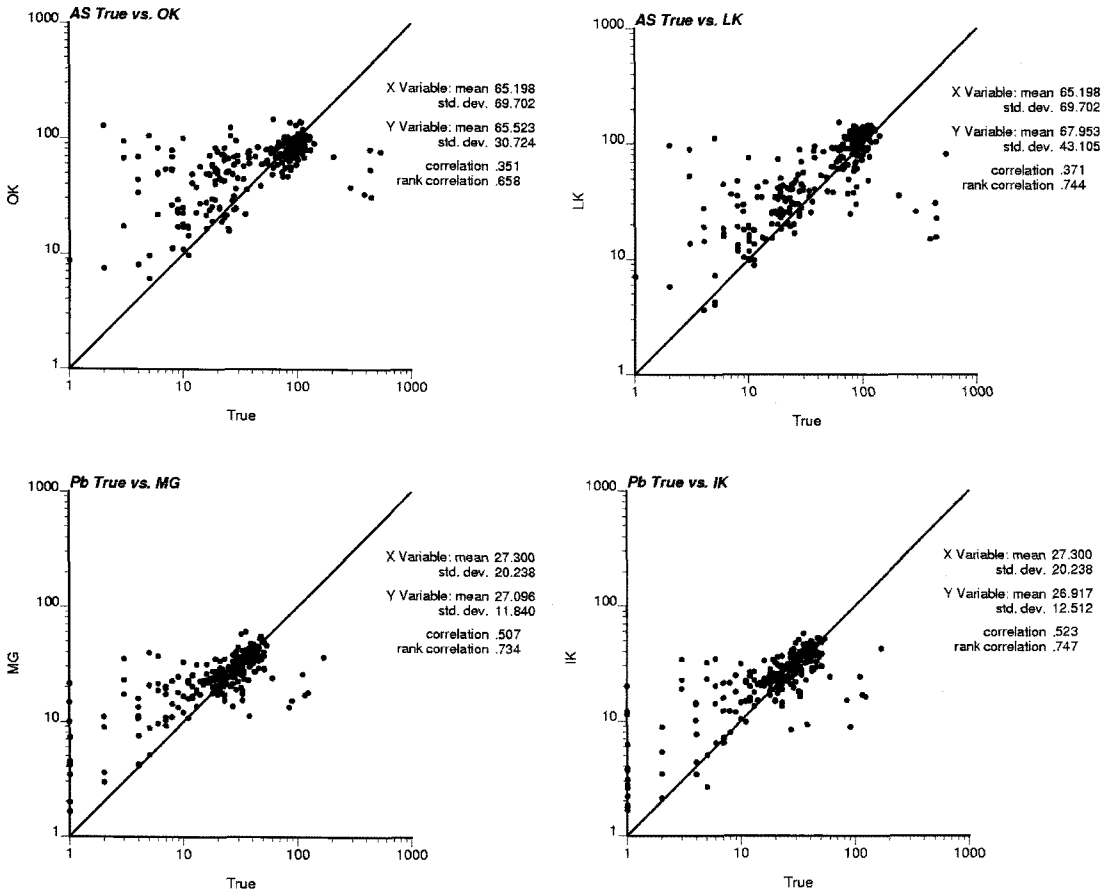


Fig. 4. Log-scale scatterplots of estimated versus observed values for some selected kriging algorithms.

Table 1. Quantitative measures related to prediction capability of each kriging algorithm based on leave-one-out cross validation

	As			Pb		
	ME	MAE	Correlation	ME	MAE	Correlation
OK	0.326	30.772	0.351	0.298	9.391	0.484
LK	2.755	28.985	0.371	6.720	13.013	0.497
MG	-2.019	27.196	0.371	-0.204	8.682	0.507
IK	-0.039	27.101	0.403	-0.383	8.454	0.523
MIK	0.842	28.712	0.374	-0.450	8.876	0.474

법의 예측 능력은 이러한 오차 분산을 최소화하는 과정에서 참값을 어느 정도 가깝게 예측하느냐에 좌우되게 된다.

Table 1은 교차 검증을 통해 계산된 비소와 납 원소의 ME, MAE 및 참값과의 선형상관계수를 각각 나타내

고 있다. 편향 정도를 파악할 수 있는 ME의 경우, 비소에서는 로그 정규 크리깅이 평균적으로 참값을 과추정하는 것으로 나타났다. 반면 다중 가우시안 크리깅은 로그 정규 크리깅과 반대로 저추정하는 양상을 나타내고 있으며, 지시가 크리깅의 ME값은 0에 가까워 가장 편향 정도가 약한 것으로 나타났다. 납 원소의 경우에는, 비소와 마찬가지로 로그 정규 크리깅이 참값의 과추정 정도가 가장 큰 것으로 나타났으며, 나머지 크리깅 기법들의 편향 정도는 상대적으로 작게 나타났다. 참값 예측 오차의 크기를 나타내는 MAE의 경우, 비소에서는 다중 가우시안 크리깅과 지시자 크리깅이 가장 작은 값을 나타냈으며, 정규 크리깅이 가장 크게 나타났다. 예측값과 참값의 선형상관계수는 두 원소 모두에서 지시자 크리깅이 상대적으로 제일 높게 나타나서, 참값과의 경향성

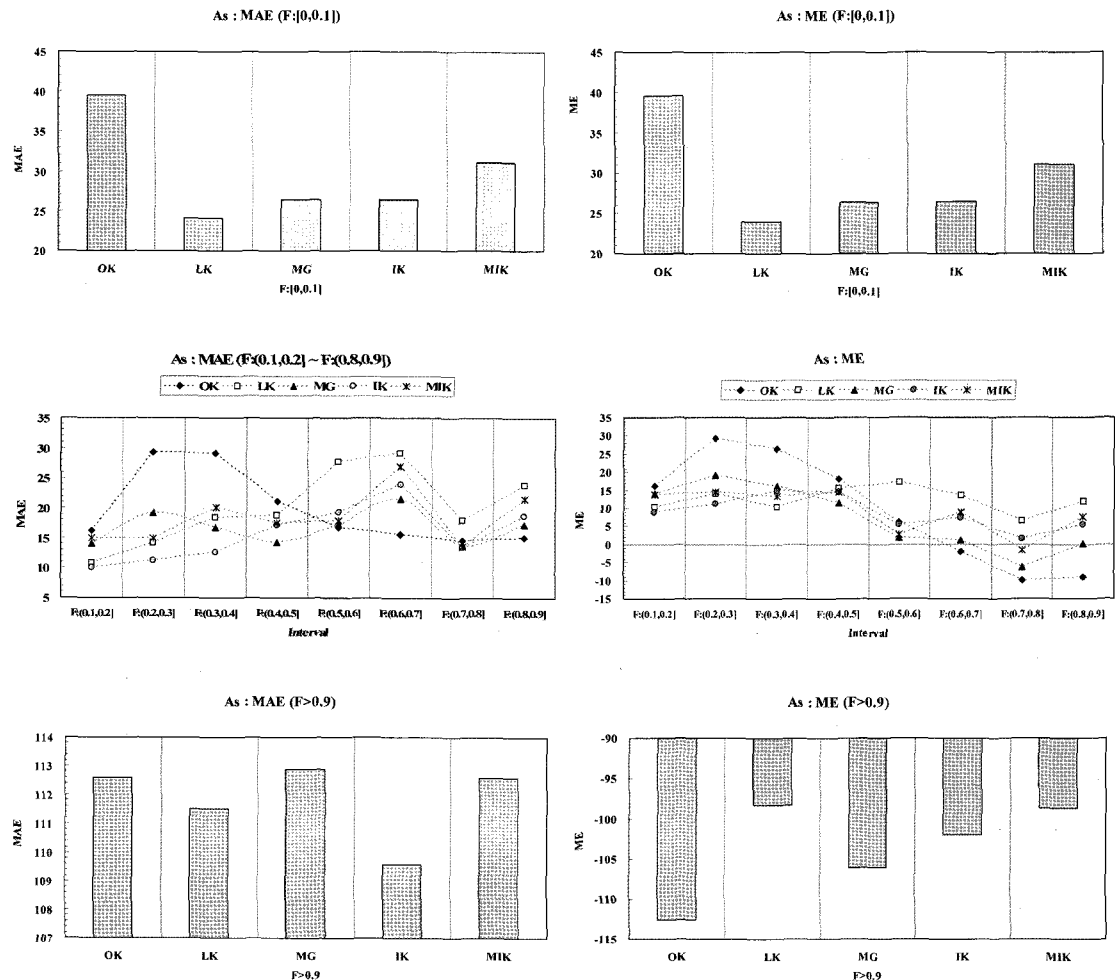


Fig. 5. MAE and ME of each kriging algorithm for percentile classes of As. The left column for MAE and the right column for ME.

이 가장 높은 것으로 나타났다. 이러한 분석 결과를 종합하면, 지시자 크리깅이 편향 정도, 예측 오차의 크기 및 참값과의 상관성 측면에서 가장 우수하며, 그 다음으로 다중 가우시안 크리깅과 중앙값 지시자 크리깅이 우수하였다.

Table 1에 제시된 값들은 전반적인 경향을 나타내기 때문에, 자료값의 분포 특성을 고려하기 위해 백분위수 구간별 예측 능력을 추가적으로 비교분석하였다. Fig. 5의 As 원소의 구간별 MAE와 ME를 살펴보면, 우선 10% 백분위수 미만의 자료값에 대해서는 모든 크리깅 기법들이 과추정하는 양상을 보였으며, 정규 크리깅의 오차가 가장 크게 나타났다. 로그 정규 크리깅의 MAE가 가장 작게 나타났으며, 그 다음으로 지시자 크리깅과 다중 가우시안 크리깅은 유사한 오차를 나타내었다.

90% 백분위수 이상의 자료값의 경우에는 모든 크리깅 기법들이 저추정하는 양상을 보였으며, 지시자 크리깅의 오차가 가장 작게 나타났다. 10% 백분위수와 90% 백분위수 사이 값들의 예측에서는 중앙값인 50% 백분위수까지는 지시자 크리깅의 오차가 가장 작게 나타났으며, 그 이상에서는 정규 크리깅의 오차가 가장 작게 나타났다. ME값을 통한 편향 정도를 비교하면, 정규 크리깅은 60% 백분위수까지는 과추정하다가 그 이후부터는 저추정하는 양상으로 바뀌고 있다. 반면, 로그 정규 크리깅은 모두 과추정하는데, 50% 백분위수 이상부터는 과추정 양상이 가장 두드러지게 나타나고 있다. Fig. 6의 Pb 원소의 구간별 MAE와 ME를 살펴보면, 10% 백분위수 미만의 자료값에 대해서는 As 원소와 마찬가지로 과추정 양상이 나타났으며, 지시자 크리깅의 MAE가

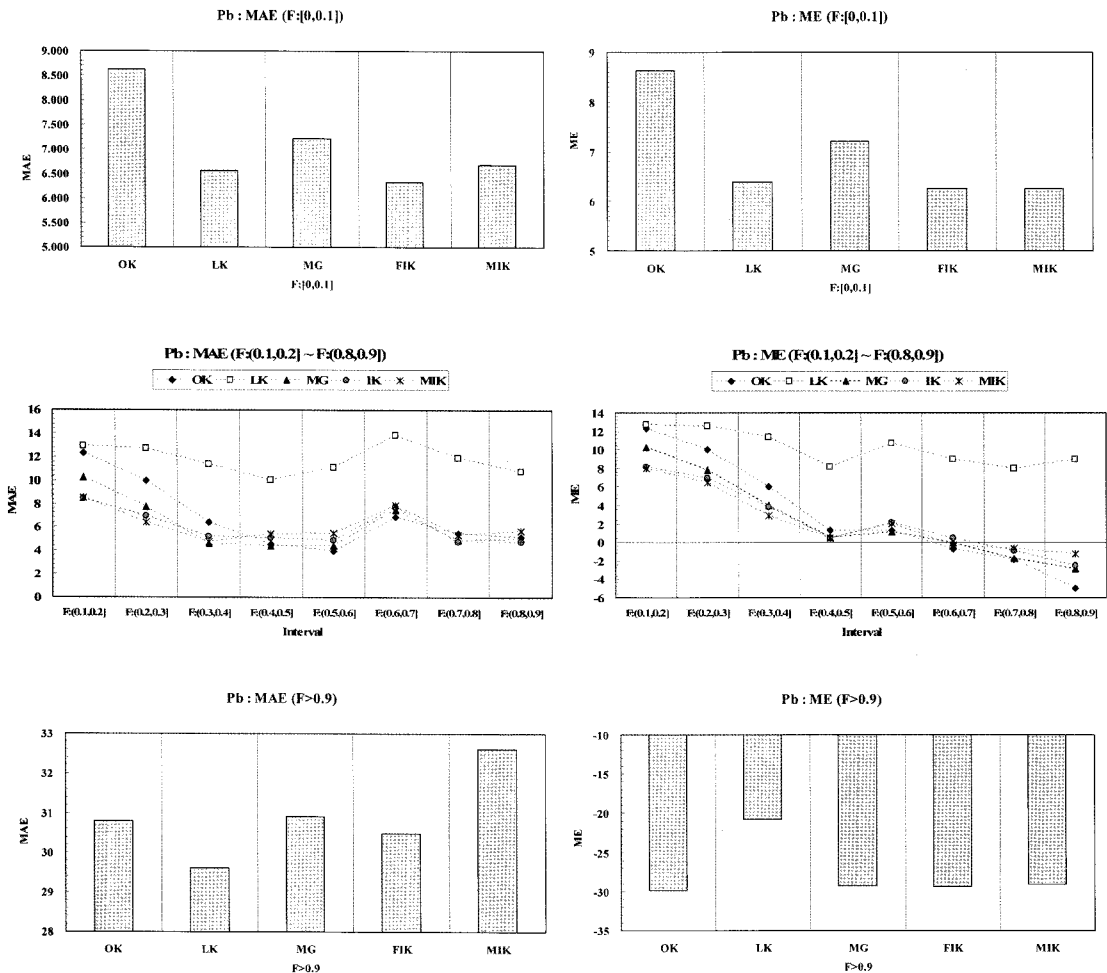


Fig. 6. MAE and ME of each kriging algorithm for percentile classes of Pb. The left column for MAE and the right column for ME.

가장 적게 나타났다. 90% 백분위수 이상의 자료값의 경우에는, 모든 크리깅 기법들이 저추정하였으며 로그 정규 크리깅의 오차가 가장 작았고, 그 다음으로 지시자 크리깅으로 나타났다. 10% 백분위수와 90% 백분위수 사이 구간의 값들의 예측 양상을 살펴보면, 로그 정규 크리깅은 모두 과추정하는 양상을 나타냈으며, 다른 크리깅 기법들은 60% 백분위수부터 저추정 양상으로 바뀌었다. 그리고 지시자 크리깅과 중앙값 지시자 크리깅의 오차가 가장 작게 나타났으며, 그 다음으로 다중 가우시안 크리깅의 오차가 작았다.

이러한 두 원소의 구간별 예측 능력을 종합하면, 작은 값과 큰 값의 예측 양상에는 크리깅의 일반적인 특성인 각각 과추정과 저추정 양상이 나타나지만, 종합적으로는 지시자 크리깅의 예측 능력이 제일 우수한 것으로 나타났다. 그리고 이 사이 구간에서의 예측의 경우, 구간별 예측 오차의 차이는 존재하지만 전반적으로 역시 지시자 크리깅이 두 원소 모두에서 공통적으로 가장 우수한 것으로 나타났으며, 이 구간내에서 높은 값들은 로그 정규 크리깅의 과추정 양상이 가장 크게 나타났다. 로그 정규 크리깅의 경우, 식 (5)의 역변환 과정에서 분산값과 라그랑지안 계수의 값이 지수함수의 지수에 더해지는 형태이기 때문에, 지수 함수의 특성상 높은 값들이 과대하게 부각되는 것으로 판단된다. 또한, 로그 정규 크리깅을 제외한 나머지 크리깅 기법의 Table 1에 제시된 작은 ME값은 대부분 50% 혹은 60% 미만까지의 값에 대한 과추정과 그 이상 범위의 값에 대한 저추정이 균형을 이루어서 0 주변의 작은 값을 나타낸 것이다.

2) 샘플링 밀도에 따른 예측 능력의 비교 분석

샘플링 밀도 즉 현장 관측 수에 따른 예측 능력의 변화 양상을 살펴보기 위해, 원 자료의 75%, 50%, 25%를 각각 샘플링한 자료에 대해 교차 검증을 수행하였다. 전체 자료에 대해 임의 샘플링을 수행할 경우, 특정 구간의 값이 적거나 많을 수 있기 때문에 100분위수별로 75%, 50%, 25%를 각각 임의 샘플링하여 분석하였다.

Fig. 7은 두 원소의 샘플링 밀도의 차이에 따른 ME, MAE 및 원 자료값과의 상관계수를 나타내고 있다. Pb의 경우, 샘플링 밀도가 작아질수록 오차가 커지고, 상관계수 또한 전반적으로 감소하는 일반적인 결과를 나타내고 있다. 그러나 As의 경우, 샘플링 밀도의 차이에

따른 오차 및 상관계수의 크기 차이가 두드러지게 나타나지는 않았다. 이러한 Pb와는 다른 As의 양상은 Fig. 3에 제시된 히스토그램에서 확인할 수 있듯이 100미만의 값에 값들이 많이 몰려있어서 특정 값들이 샘플링을 통해 자료에서 제외되더라도 다른 값들이 이 값의 영향을 대체하였기 때문인 것으로 판단된다. 샘플링 밀도의 차이에 따른 예측 능력을 나타내는 MAE의 양상을 살펴보면, 일관되게 지시자 크리깅의 예측 오차가 가장 작게 나타났으며, 그 다음으로는 다중 가우시안 크리깅 혹은 중앙값 지시자 크리깅으로 나타났다. 정규 크리깅의 오차가 Pb 원소의 일부 샘플링 밀도를 제외하고는 모두 가장 크게 나타났다. 예측의 편향 정도를 나타내는 ME의 양상을 살펴보면, As의 25% 샘플링을 제외하고는 지시자 크리깅의 편향 정도가 가장 작게 나타났으며, 일관되게 로그 정규 크리깅의 과추정이 두드러졌다. 다중 가우시안 크리깅의 경우, As의 원소의 예측에서 25% 샘플링을 제외하고는 모두 저추정하는 양상을 나타내었다. 이러한 다중 가우시안 크리깅의 결과는, 정규 점수 변환에 기인한 것으로 판단된다. 정규 점수 변환 과정에서 원 자료를 크기순으로 오름차순으로 재배열한 후에, 이에 대응하는 표준 정규 분포의 값으로 치환을 하게 되는데, 원 자료의 절대적인 값의 차이보다는 자료의 순위 즉 재배열 순서만이 중요하게 작용하기 때문에 역변환을 하더라도 이러한 차이를 최대한 고려할 수 없는 단점이 존재하게 되며, 이러한 영향이 결과적으로 저추정 양상을 가져온 것으로 판단된다. 또한 일관되게 원 자료값과의 상관계수에서 나타나는 지시자 크리깅의 상대적으로 높은 상관계수는 지시자 크리깅이 다른 크리깅 기법에 비해 낮게 나타난 MAE 및 ME의 결과를 뒷받침해 주고 있다.

Fig. 8과 Fig. 9는 As 원소의 구간별 샘플링 밀도에 따른 자료 구간별 MAE와 ME의 변화 양상을 나타내고 있다. 샘플링 밀도가 변하더라도 50% 백분위수인 중앙값 미만의 값의 예측에서는 정규 크리깅의 오차가 가장 큰 것으로 나타났으며, 중앙값 이상부터는 로그 정규 크리깅의 오차가 가장 컸다. 90% 백분위수 이상의 값에 대한 예측 능력은 샘플링 밀도에 따른 영향이 가장 작게 나타났지만, 지시자 크리깅의 예측 능력이 제일 우수하였다. Fig. 9의 ME의 양상에서는 90% 백분위수 이상의 값에 대해서는 샘플링 밀도와 크리깅 기법의 차이에 관

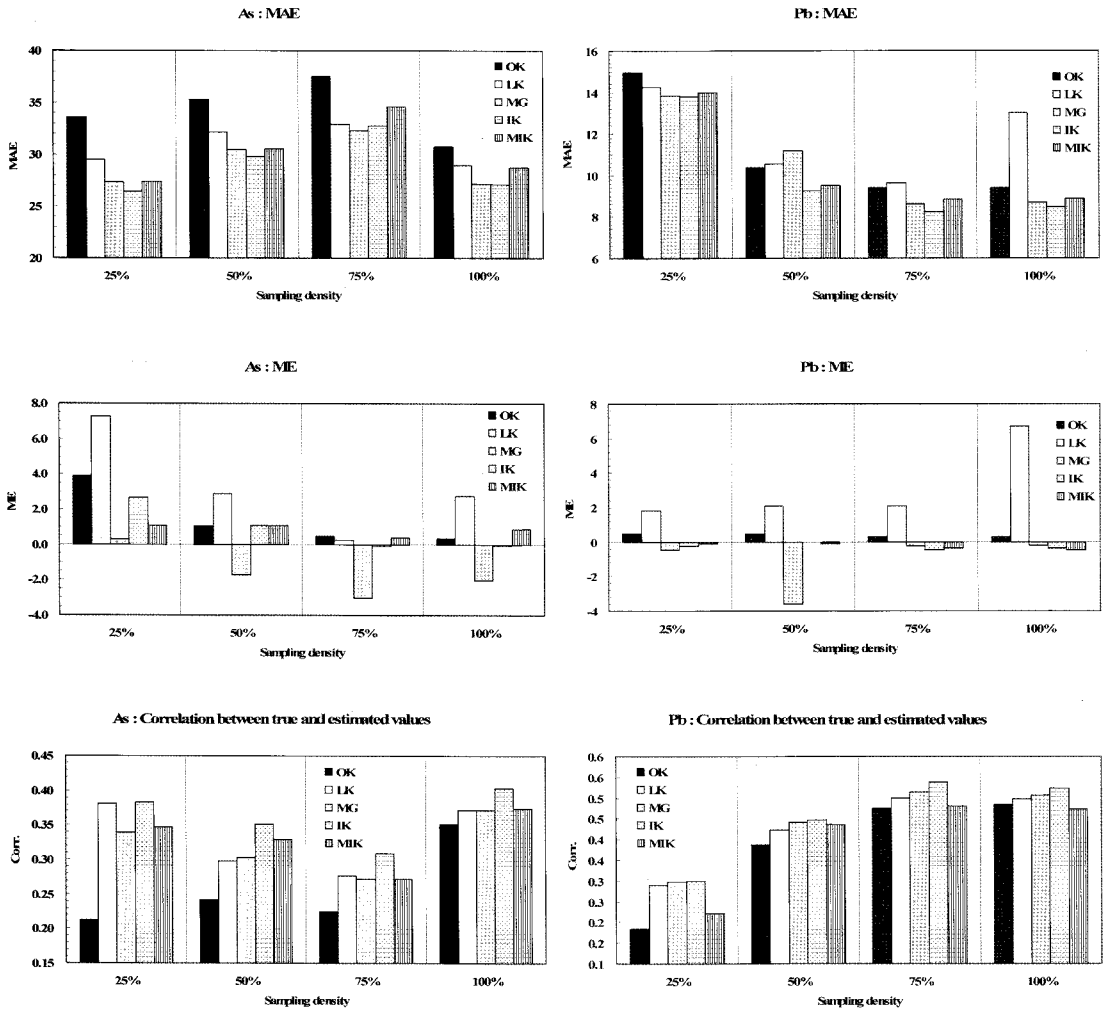


Fig. 7. Impact of sampling density and kriging algorithms on the ME, MAE and correlation coefficient for As and Pb. The left column for As and the right column for Pb.

계없이 모두 저추정하는 것으로 나타났다. 또한 60% 백분위수 이상값의 과추정 혹은 저추정의 양상은 샘플링 밀도에 따라 다소간의 차이가 나타나서 일관된 특성을 확인할 수는 없었다. 그러나 로그 정규 크리깅의 과추정 양상은 모두 두드러지게 나타났으며, 정규 크리깅이 50% 백분위수 미만의 값을 샘플링 밀도와 상관없이 가장 크게 과추정함을 확인할 수 있었으며, 이러한 양상은 이 구간 이하에서의 높은 MAE 값을 나타내게 되었다.

Pb 원소의 구간별 샘플링 밀도에 따른 자료 구간별 MAE의 변화 양상을 살펴보면(Fig. 10), As 원소와의 공통된 양상은 정규 크리깅의 높은 MAE값과 중앙값부터 높은 값의 예측에서 로그 정규 크리깅의 과추정과 높은 MAE이다. 그러나 10% 백분위수 이하의 값의 예측

에서는 로그 정규 크리깅의 MAE가 가장 낮게 나타났다. 또한 As와 달리 다중 가우시안 크리깅의 지지자 크리깅에 비해 상대적으로 높은 오차가 샘플링 밀도가 감소할수록 60% 백분위수 이상의 값의 예측에서 나타나고 있다. 중앙값 지지자 크리깅의 경우, 낮은 값의 예측에서 특히 샘플링 밀도가 작을수록 커지거나 다중 가우시안 크리깅과 유사한 오차 정도를 나타내고 있다. Fig. 11의 ME의 변화 양상을 살펴보면, 로그 정규 크리깅의 중앙값 이상의 값의 예측에서 과추정이 샘플링 밀도의 변화에 상관없이 두드러짐을 확인할 수 있다. 또한 다중 가우시안 크리깅의 높은 값 범위에서의 저추정 양상도 역시 샘플링 밀도의 변화와 상관없이 나타나는 일반적인 특성인 것으로 나타났다.

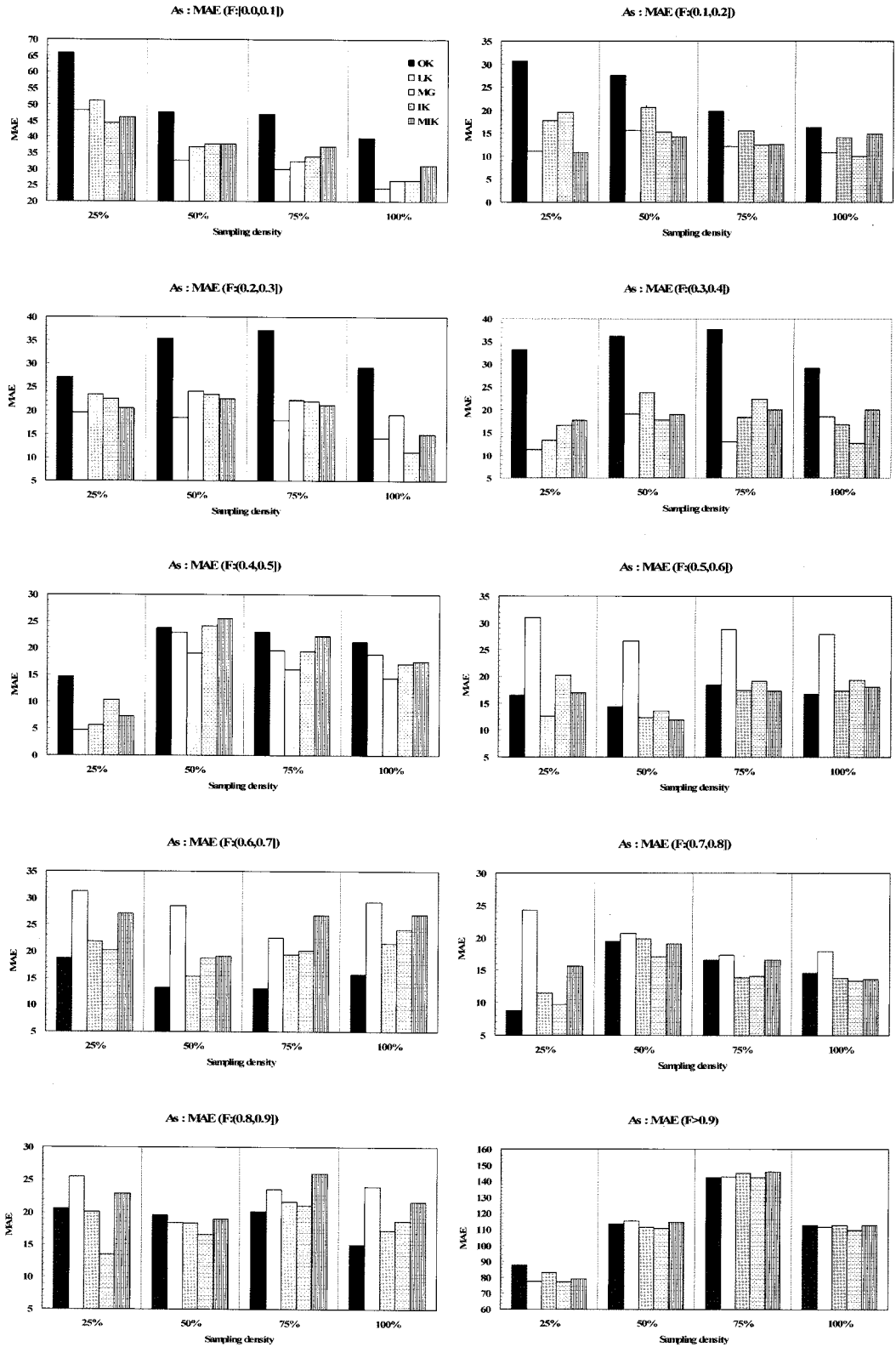


Fig. 8. Impact of sampling density and kriging algorithms on the MAE for percentile classes of As.

Comparison of Univariate Kriging Algorithms for GIS-based Thematic Mapping with Ground Survey Data

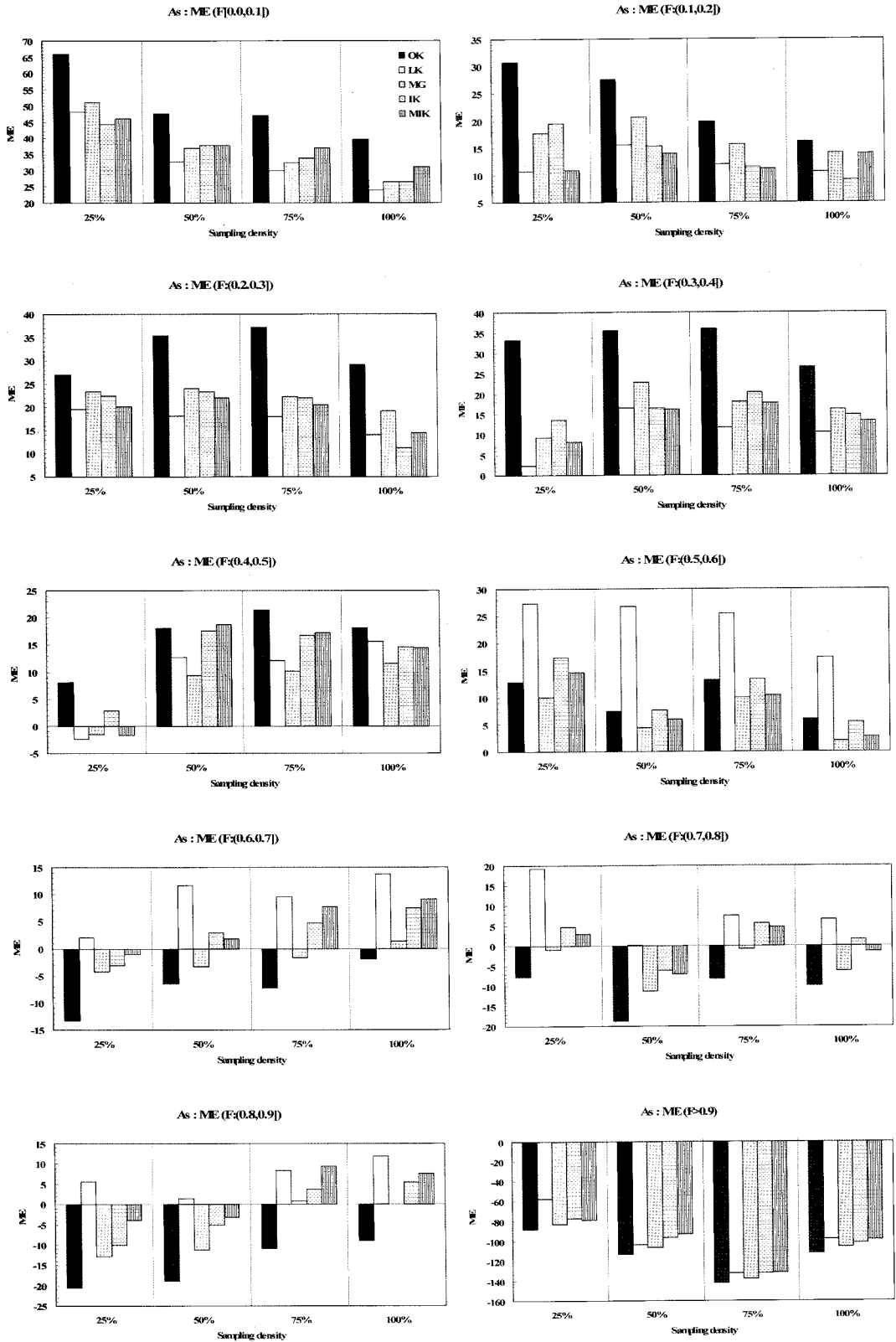


Fig. 9. Impact of sampling density and kriging algorithms on the ME for percentile classes of As.

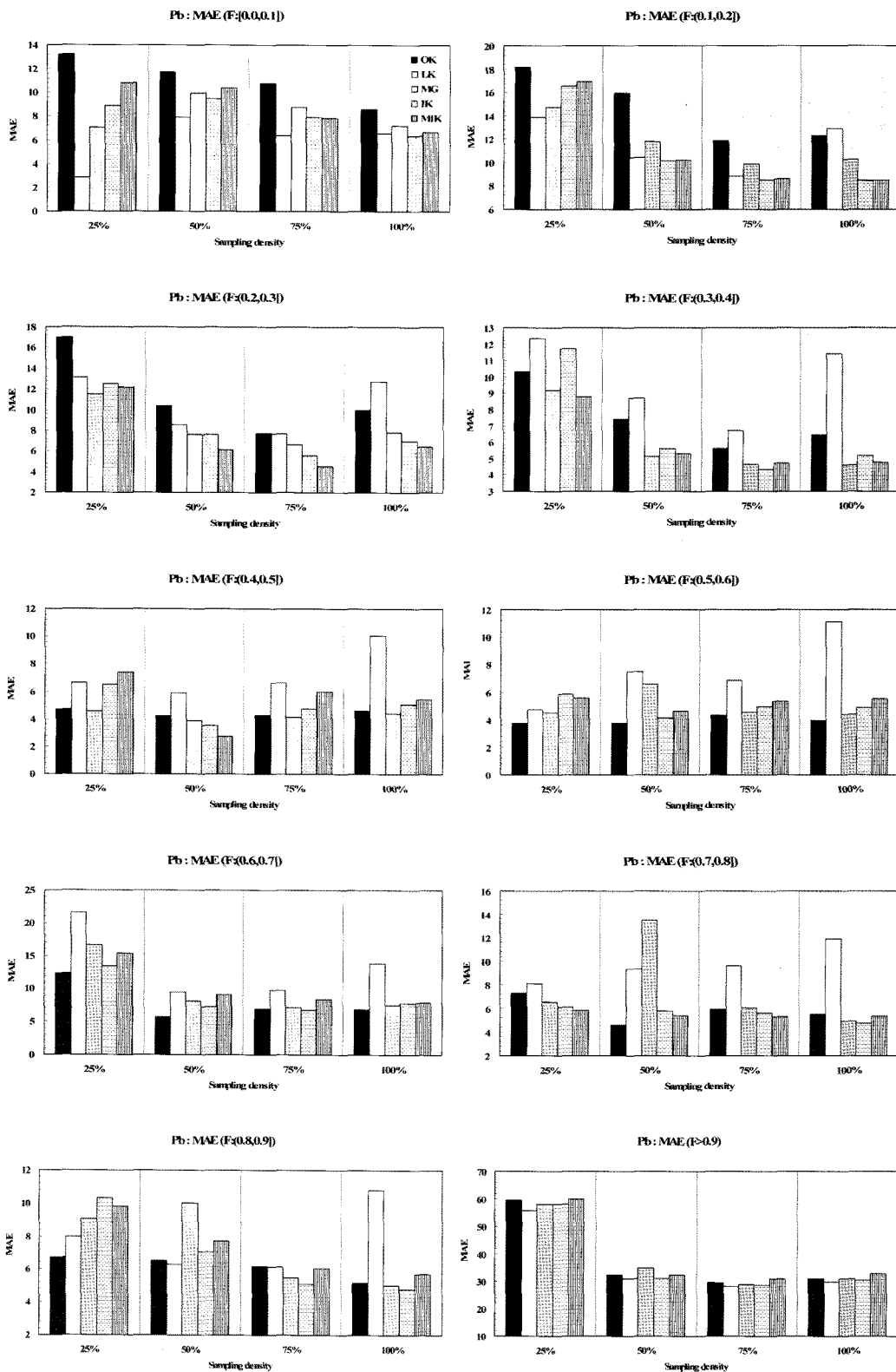


Fig. 10. Impact of sampling density and kriging algorithms on the MAE for percentile classes of Pb.

Comparison of Univariate Kriging Algorithms for GIS-based Thematic Mapping with Ground Survey Data

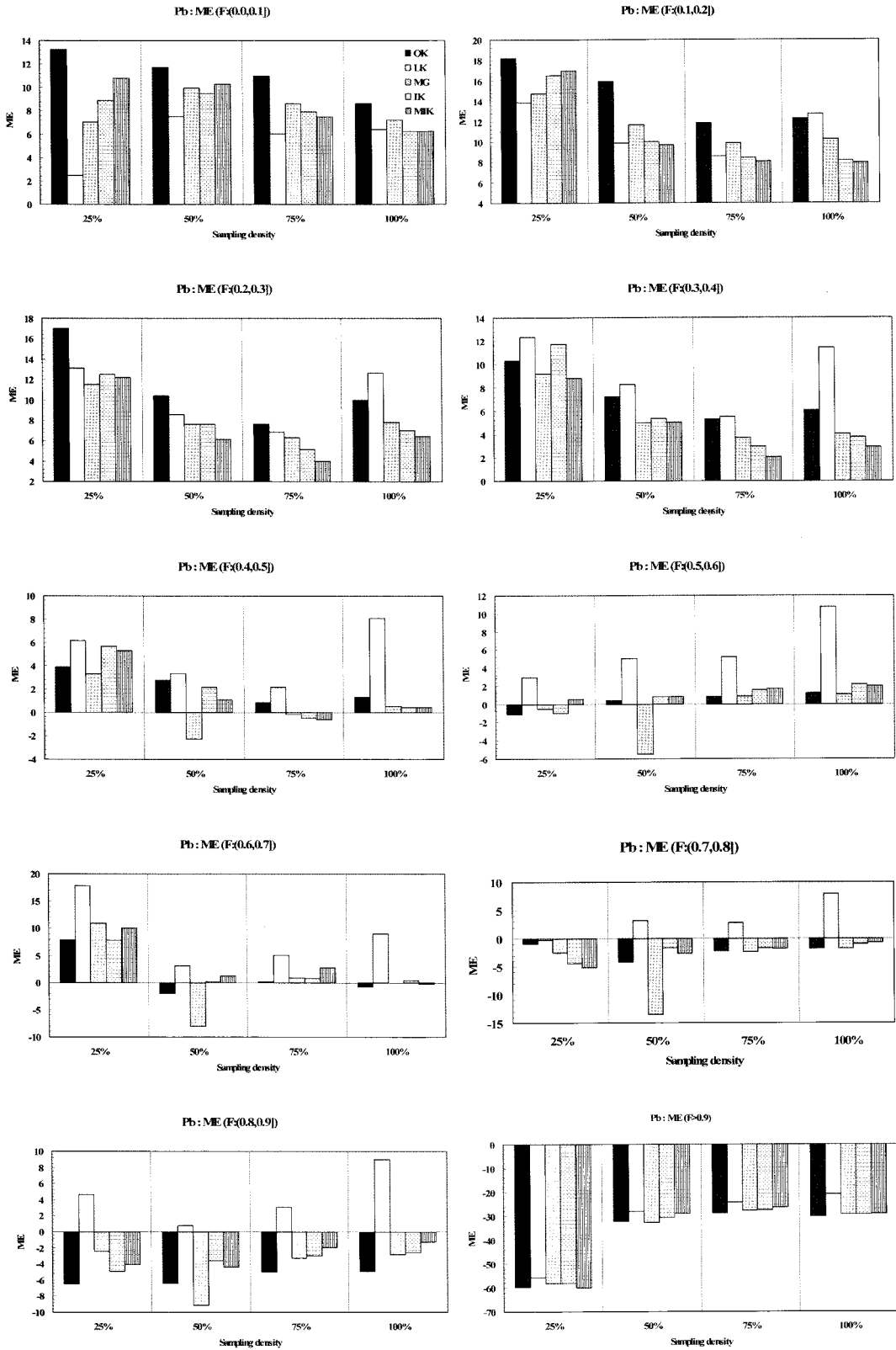


Fig. 11. Impact of sampling density and kriging algorithms on the ME for percentile classes of Pb.

3) 다중 가우시안 크리깅과 지시자 크리깅의 불확실성 모델링 결과 비교

다른 크리깅 기법과 달리 다중 가우시안 크리깅과 지시자 크리깅의 경우, 예측 위치에서의 ccdf가 모델링될 수 있기 때문에 이를 기반으로 불확실성 모델링이 가능하다. 전술한 바와 같이, 전반적으로 지시자 크리깅이 가장 우수한 예측 능력을 나타내었고, 그 다음으로 다중 가우시안 크리깅이 우수한 결과를 나타내어 이 크리깅 기법의 ccdf로부터 계산이 가능한 불확실성을 비교분석하여 성능을 비교하였다. Ccdf로부터의 불확실성 모델링에는 Deutsch(1997)에 의해 제안된 정확도 그림(accuracy plot)과 예측 우수도(goodness)를 계산하였다. 정확도 그림은 교차 검증을 통해 확률 구간별 실제 포함되는 예측값의 기대 비율을 나타내는 그림으로, 그림의 대각선에 가까울수록 모델의 추정 정확도가 높아진다. 예측 우수도는 식 (11)로 계산되며, 이러한 정확도 그림의 추정 및 기대 비율의 일치정도를 나타내는 통계

치로 높을수록 모델의 예측값이 정확한 것을 나타낸다.

$$G = 1 - \frac{1}{K} \sum_{k=1}^K w_k |p_k^* - p_k| \quad (11)$$

여기서 K 는 연속적인 ccdf의 이산화 갯수를 나타내며, p_k^* 는 기대 비율 p_k 에 대한 추정 비율을 나타낸다. w_k 는 일종의 가중치로 p_k^* 가 p_k 보다 클 경우 1, 아니면 2를 가진다.

Fig. 12는 두 원소의 정확도 그림과 예측 우수도 통계치를 나타내고 있다. 비소의 경우, 예측 우수도 통계치는 지시자 크리깅과 중앙값 지시자 크리깅이 상대적으로 다중 가우시안 크리깅보다 우수하며, 실제 기대 비율도 45도 대각선에 근접하게 나타나고 있다. 납 원소의 경우, 지시자 크리깅이 가장 대각선에 근접하였으며, 중앙값 지시자 크리깅의 경우 기대 비율에 못 미치는 부정확한 예측 능력을 나타내었다. 예측 우수도 통계치에서도 지시자 크리깅이 가장 높은 값을 나타내었으며, 중앙값 지시자 크리깅이 가장 낮게 나타났다. 이러한 결과를

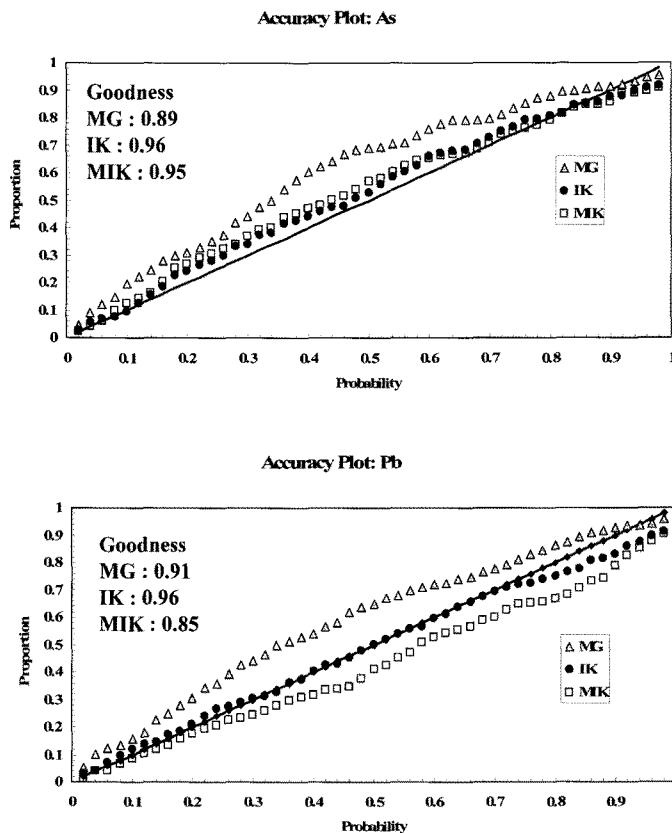


Fig. 12. Accuracy plots and goodness statistics for multi-Gaussian kriging and indicator kriging for As and Pb elements.

종합하면, 지시자 크리깅이 다중 가우시안 크리깅에 비해 예측 우수도가 높으며, 예측에 수반되는 불확실성도 작게 나타났다.

5. 결론

현장 조사 자료로부터 주제도를 작성하기 위해서는 적절한 내삽 기법의 적용이 필수적이며, 특히 비대칭 분포를 가지는 자료의 특성을 고려할 수 있는 내삽 기법의 선택 또한 중요하다. 이 연구에서는 비대칭 분포를 가지는 현장 조사 자료로부터 GIS 기반 주제도를 작성하기 위한 내삽 기법으로 단변량 크리깅 기법 중에서 기존 정규 크리깅과 비선형 자료 변환에 기반한 로그 정규 크리깅, 다중 가우시안 크리깅 및 지시자 크리깅을 비교분석하였다. 비소와 납 원소를 이용한 분포도 작성 사례 연구를 통해 교차 검증 기반 예측 능력의 비교와 더불어 샘플링 밀도에 따른 차이 및 자료의 구간별 예측 능력의 차이를 비교 분석하여 다음과 같은 결론을 얻었다.

기존 정규 크리깅에 비해 비선형 변환 기반 크리깅이 상대적으로 우수한 예측 능력을 나타내었다. 적용 크리깅 기법 모두에서 작은 값은 과추정, 큰 값은 저추정하는 특성을 나타내었는데, 로그 정규 크리깅의 과추정 양상이 가장 두드러졌다. 작은 값들의 분포가 중요한 역할을 하는 주제도를 작성하는 경우, 이러한 로그 정규 크리깅의 과추정 양상은 잘못된 해석 결과를 유도할 가능성이 있을 수 있다. 전반적인 예측 능력면에서 지시자 크리깅이 가장 우수한 예측 능력을 나타내었으며, 다음으로는 다중 가우시안 크리깅과 중앙값 지시자 크리깅 순이었다. 90% 백분위수 이상의 높은 값들은 모두 저추정하였으나, 지시자 크리깅이 상대적으로 가장 작은 예측 오차를 나타내었다. 지시자 크리깅 다음으로 상대적으로 우수한 예측 능력을 보인 다중 가우시안 크리깅의 경우, 불확실성 모델링 결과 지시자 크리깅에 비해 상대적으로 부정확한 것으로 나타났다. 지정된 임계치별로 베리오그램 모델링과 크리깅 시스템을 풀어야 하는 지시자 크리깅의 실제 적용상의 번거로움을 완화한 중앙값 지시자 크리깅의 경우, 다중 가우시안 크리깅과 비교하였을 때 예측 능력이 유사하거나 떨어지는 것으로 나타났다. 이러한 결과는 중앙값 지시자 크리깅이 다른 크

리깅 기법과 마찬가지로 한번의 베리오그램 모델링을 수행하는 실제 적용상의 편리성을 가지지만, 중앙값을 임계치로 설정하는 과정에서 각 임계치별로의 전환 확률을 완벽하게 모델링하는데 있어서 한계가 있기 때문이다.

결론적으로 극단적인 이상값 혹은 탐지 기계 분해능 미만의 자료값이 존재하는 비대칭 현장 조사 자료를 이용한 주제도 작성시, 실제 적용상의 번거로움이 존재하지만 지시자 크리깅의 적용 가능성이 가장 크며, 그 다음으로 다중 가우시안 크리깅이 우수한 결과를 얻을 수 있었다. 이 두 크리깅 기법은 미지의 위치에서의 값 추정과 더불어 예측에 수반되는 불확실성 모델링이 가능하기 때문에, 주제도 작성과 더불어 작성된 주제도의 질 혹은 불확실성을 모델링하는데 유용하게 적용될 수 있을 것으로 판단된다.

사 사

이 연구는 한국지질자원연구원 기관목적사업 “GIS 기반 국토지질정보시스템 실용화 기술 개발”의 위탁과제 연구비 지원에 의해 수행되었습니다.

참고문헌

- 박노욱, 장동호, 2008. 수치표고모델과 다변량 크리깅을 이용한 기온 및 강수 분포도 작성, 대한지리학회지, 43(6): 1002-1015.
- 이진수, 서효준, 황인호, 1998. 지화학조사연구 (1:250,000 강릉도폭 광역 지화학도), 한국자원연구소.
- 정상용, 이강근, 1995. 난지도 매립지 일대의 지하수위 분포 추정을 위한 복합 크리깅의 응용, 지하수환경, 2(2): 58-63.
- 조홍래, 정종철, 2006. 강우자료에 대한 공간보간 기법의 적용, 한국GIS학회지, 14(1): 29-41.
- 조홍래, 정종철, 2009. 공간보간기법에 의한 서울시 미세먼지(PM10)의 분포 분석, 환경영향평가, 18(1): 31-39.
- Chilès, J.-P. and P. Delfiner, 1999. *Geostatistics:*

- Modeling Spatial Uncertainty*, John Wiley & Sons, Inc.
- Deutsch, C. V., 1997. Direct assessment of local accuracy and precision, In: Baafi, E.Y. and N.A. Schofield (Eds.), *Geostatistics Wollongong '96*, Springer, New York, pp. 115-125.
- Deutsch, C. V. and A. G. Journel, 1998. *GSLIB: Geostatistical Software Library and User's Guide*, 2nd Edition, Oxford University Press.
- Glenn, N. F. and J. R. Carr, 2003. The use of geostatistics in relating soil moisture to RADARSAT-1 SAR data obtained over the Great Basin, Nevada, USA, *Computers & Geosciences*, 29(5): 577-586.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*, Oxford University Press.
- Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology*, 228(1-2): 113-129.
- Isaaks, E. H. and R. M. Srivastava, 1989. *An Introduction to Applied Geostatistics*, Oxford University Press.
- Journel, A. G. and C. J. Huijbregts, 1978. *Mining Geostatistics*, Academic Press.
- Papritz, A. and J.-P. Dubois, 1999. Mapping heavy metals in soil by (non-)linear kriging: an empirical validation, In: Gómez-Hernández, J., A. Soares, and R. Froidevaux (Eds.), *geoENV II - Geostatistics for Environmental Applications*, Kluwer, Dordrecht, pp. 429-440.
- Park, N.-W., D.H., Jang, and K.-H. Chi, 2009. Integration of IKONOS imagery for geostatistical mapping of sediment grain size at Baramarae beach, Korea, *International Journal of Remote Sensing*, in press.
- Pebesma, E. J. and C. G. Wesseling, 1998. GSTAT: a program for geostatistical modelling, prediction and simulation, *Computers & Geosciences*, 24(1): 17-31.
- Remy, N., A. Boucher, and J. Wu, 2009. *Applied Geostatistics with SGeMS: a User's Guide*, Cambridge University Press.
- Saito, H. and P. Goovaerts, 2000. Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site, *Environmental Science & Technology*, 34(19): 4228-4235.
- Sales, M. H., C. M. Souza Jr., P. C. Kyriakidis, D. A. Roberts, and E. Vidal, 2007. Improving spatial distribution estimation of forest biomass with geostatistics: a case study for Rondonia, Brazil, *Ecological Modeling*, 205(1-2): 221-230.