

## 지질 주제도 작성을 위한 지표 조사 자료와 부가 자료의 지구통계학적 통합

박노옥\*<sup>†</sup> · 장동호\*\* · 지광훈\*

\*한국지질자원연구원 지질자원정보센터, \*\*공주대학교 지리학과

### Geostatistical Integration of Ground Survey Data and Secondary Data for Geological Thematic Mapping

No-Wook Park\*<sup>†</sup>, Dong-Ho Jang\*\*, and Kwang-Hoon Chi\*

\*Geoscience Information Center, Korea Institute of Geoscience and Mineral Resources

\*\*Department of Geography, Kongju National University

**Abstract :** Various geological thematic maps have been generated by interpolating sparsely sampled ground survey data and geostatistical kriging that can consider spatial correlation between neighboring data has widely been used. This paper applies multi-variate geostatistical algorithms to integrate secondary information with sparsely sampled ground survey data for geological thematic mapping. Simple kriging with local means and kriging with an external drift are applied among several multi-variate geostatistical algorithms. Two case studies for spatial mapping of groundwater level and grain size have been carried out to illustrate the effectiveness of multi-variate geostatistical algorithms. A digital elevation model and IKONOS remote sensing imagery were used as secondary information in two case studies. Two multi-variate geostatistical algorithms, which can account for both spatial correlation of neighboring data and secondary data, showed smaller prediction errors and more local variations than those of ordinary kriging and linear regression. The benefit of applying the multi-variate geostatistical algorithms, however, depends on sampling density, magnitudes of correlation between primary and secondary data, and spatial correlation of primary data. As a result, the experiment for spatial mapping of grain size in which the effects of those factors were dominant showed that the effect of using the secondary data was relatively small than the experiment for spatial mapping of groundwater level.

**Key Words :** Kriging, Multi-variate geostatistics, Correlation.

**요약 :** 다양한 지질 주제도는 현지 조사에 의해 부분적으로 샘플링된 공간 자료의 내삽을 통해 작성되어 왔으며, 공간적 예측을 위해 공간적 상관성을 고려하는 지구통계학적 크리깅이 많이 적용되어 왔다. 이 논문에서는 지질 주제도 작성을 목적으로 부분적인 샘플링 자료와 이와 상관된 부가자료를 통합하기 위해 다변량 지구통계 기법을 적용하였다. 다변량 지구통계 기법으로 simple kriging with local means와 kriging with an external drift를 적용하였다. 지하수위 분포도 작성과 퇴적물 입도 분포도 작성의 2가지 사례연구를 수행하였는데, 지하수위 분포도 작성에는 지하수위 분포 샘플링 자료와 수치고도모델을, 퇴적물 입도

분포도 작성에는 입도 샘플링 자료와 IKONOS 원격탐사 자료를 이용하였다. 사례연구 수행결과, 다변량 지구통계 기법이 그동안 많이 이용되어온 단변량 지구통계 기법인 정규 크리깅에 비해 작은 추정 오차를 나타내면서 국소적인 특성을 반영할 수 있었다. 그러나 추정 오차의 정도는 샘플링 밀도, 부가자료와의 상관성과 공간자료 자체의 상관성 정도에 영향을 받는 것으로 나타났는데, 특히 퇴적물 입도 분포도 작성 사례연구에서 이러한 요소들이 상호 영향을 미쳐 부가자료의 이용 효과가 상대적으로 적게 나타났다.

## 1. 서론

그동안 지하수위 분포도, 지화학 원소 분포도, 입도 분포도 등과 같은 지질학적 주제도 작성에는 위치 정보와 고유의 속성 정보를 가지는 현지 조사 자료가 이용되어왔다. 그러나 현지 조사는 비용, 시간, 인력 및 접근성 등의 문제로 조사지역의 모든 지역을 전면적으로 조사하지 못하고 부분적인 샘플링을 수행하게 된다. 예를 들어 지화학 자료는 수계, 암석 혹은 토양 등으로부터의 샘플링을 통해 얻어지게 된다. 그러나 대부분의 2차 분석은 특정 영역의 모든 위치에 대한 정보를 필요로 하기 때문에 이렇게 공간적으로 산재하는 샘플링 자료로부터 전면적인 자료를 얻기 위해 각종 내삽 기법 등을 적용하게 된다.

대부분의 공간자료는 이웃한 자료사이에는 값이 유사한 공간적 상관성을 가지기 때문에 이를 모델링 할 수 있는 지구통계학적 크리깅이 내삽 기법으로 많이 이용되어왔다. 지구통계학적 크리깅은 미지의 값은 공간적으로 상관된 확률 변수(random variable)의 조합으로 모델링된다는 확률 함수의 개념에 기반을 두고 있기 때문에 주변 자료와의 공간적 상관성을 고려하여 미지의 값을 계산하게 된다(Issaks and Srivastava, 1989; Goovaerts, 1997). 추정 혹은 내삽하고자 하는 변수와 상관성을 보이면서 보다 많은 지점에서 샘플링된 부가자료를 이용할 경우 추정 오차를 줄일 수 있다. 기존 연구에 이용되어온 Inverse Distance, TIN, Spline 등은 이러한 부가 자료를 통합할 수 없지만, 크리깅은 부가 자료를 통합할 수 있는 방법론을 제공할 수 있는 장점이 있다. 각종 수치화된 GIS 자료와 다양한 원격탐사 자료의 이용이 가능하기 때문에 이러한 지구통계학적 통합은 적용 가능성이 크다. 예를 들어 강우 자료의 경우에는 수치고도자료나 기상 레이더 자료를 부가자료로 이용할 수 있으며(Raspa *et al.*, 1997; Goovaerts, 2000), 지하수위 자료의 경우에는 수치고도자료가 유용하게 사용될 수 있다

(정상용과 이강근, 1995; 함세영 외, 2001).

이 논문에서는 지질 주제도 작성을 목적으로 부분적인 샘플링 자료와 이와 상관된 부가자료를 통합하기 위해 다변량 지구통계 기법을 적용하고자 한다. 다변량 지구통계 기법으로 Simple Kriging with Local Means (이하 SKLM)와 Kriging with an External Drift (이하 KED)를 적용하였다. 지하수위 분포도 작성과 퇴적물 입도 분포도 작성의 2가지 사례연구를 수행하였는데, 지하수위 분포도 작성에는 지하수위 분포 샘플링 자료와 수치고도모델을, 퇴적물 입도 분포도 작성에는 입도 샘플링 자료와 IKONOS 원격탐사 자료를 이용하였다. 적용 기법의 예측 정확도의 정량적 비교를 위해 교차 검증을 수행하였다.

## 2. 방법론

### 1) 지구통계학적 크리깅

1960년대 이후에 기존 알려진 광상 자료를 이용하여 새로운 광상의 위치를 추정하기 위해 개발된 지구통계학은 기존에 내삽 방법으로 알려진 크리깅으로 대표되어왔다. 지구통계학적 분석은 확률 함수의 개념에 기반을 두고, 고려하고 있는 개개의 노드 혹은 그리드의 미지수를 주변자료의 분포 양상으로 추정한 확률 변수를 이용하여 추정하게 된다. 이때 공간적인 상관성은 거리와 방향을 고려한 베리오그램(variogram)을 기본적으로 이용한다. 즉 크리깅은 이러한 베리오그램을 통해 얻어지는 관측값 사이의 공간적 상관성을 이용하여 공간적 예측을 수행하는 일반화된 최소자승 회귀분석방법론에 해당된다(최종근, 2002; Issaks and Srivastava, 1989; Goovaerts, 1997).

$n$ 개의 샘플 자료  $\{z(u_\alpha), \alpha = 1, \dots, n\}$ 를 이용하여 샘플링 되지 않은 위치  $u$ 에서의 크리깅 예측치  $z^*(u)$ 는 주변 관측값의 가중 합으로 표현된다(식 (1)).

$$z^*(u) - m(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}(u)[z(u_{\alpha}) - m(u_{\alpha})] \quad (1)$$

여기서  $m(u)$ 를 정의하는 방식에 따라 단순 크리깅(Simple Kriging), 정규 크리깅(Ordinary Kriging), 경향 크리깅(Kriging a Trend)으로 구별된다(Goovaerts, 1997). 단순 크리깅은 평균값 혹은 경향성이 연구지역에 걸쳐 일정하게 알려져 있다는 가정을 적용한다. 반면 정규 크리깅과 경향 크리깅은 평균값이 알려져 있지 않지만, 특정 이웃한 영역 안에서 일정하다고 가정하여 평균값을 계산하고, 경향 크리깅은 경향성을 좌표계의 함수로 모델링한다. 이러한 크리깅은 특정 그리드에서 추정되는 일종의 점 크리깅(point kriging)에 해당되며, 이외에도 일정 영역을 고려하는 구역 크리깅(block kriging)과 부가자료를 고려할 수 있는 공동 크리깅(cokriging) 등이 있다(최종근, 2002).

## 2) 다변량 지구통계기법

예측하고자 하는 변수  $z$ 의 샘플링 밀도가 너무 작을 경우에, 보다 많은 지점에서 샘플링되었고  $z$ 와 상관성이 있는 부가 변수  $y$ 를 다변량 지구통계기법으로 통합함으로써 변수  $z$ 의 예측 정확도를 향상시킬 수 있다. 이 연구에서는 SKLM과 KED를 다변량 통계기법으로 적용하였는데, 이 두 기법의 자세한 설명은 Goovaerts(1997)을 참조할 수 있으며 이 논문에서는 실질적인 응용 측면만에 대해서만 기술하고자 한다.

SKLM은 식 (1)의 단순 크리깅 예측치의 일정한 평균값  $m$ 을 부가 변수  $y$ 로부터 계산되는 위치마다 변화하는 가변적 지역 평균값으로 대체한다.

$$z_{SKLM}^*(u) - m_{SK}^*(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{SK}(u)[z(u_{\alpha}) - m_{SK}^*(u_{\alpha})] \quad (2)$$

가변적 지역 평균값  $m_{SK}^*(u)$ 는  $z$ 와 부가 변수  $y$ 의 연관성 분석을 통해 얻게 되는데,  $y$ 가 연속형 자료일 경우에는 보통 선형 혹은 비선형 회귀분석을 통해 얻게 된다.

$$z_{SKLM}^*(u) = m_{SK}^*(u) + \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{SK}(u)r(u_{\alpha}) \quad (3)$$

여기서  $r(u_{\alpha})$ 은  $z(u_{\alpha}) - m_{SK}^*(u_{\alpha})$ 으로 정의되는 잔차(residual)를 나타낸다.

식 (3)에서 가중치  $\lambda_{\alpha}^{SK}(u)$ 는 다음 단순 크리깅 방정식의 해로 구하게 된다.

$$\sum_{\beta=1}^{n(u)} \lambda_{\beta}^{SK}(u)C_R(u_{\alpha} - u_{\beta}) = C_R(u_{\alpha} - u), \alpha=1, 2, \dots, n(u_{\alpha}) \quad (4)$$

여기서  $C_R(h)$ 는 잔차  $R(u) = Z(u) - m(u)$ 의 공분산 함수를 나타낸다. 결론적으로 SKLM은 잔차의 크리깅 예측치를 가변적 지역 평균값에 더함으로써 최종적으로 얻게 된다.

KED는 SKLM과 마찬가지로  $z$ 의 평균이나 경향성을 추출하기 위해 부가 변수  $y$ 를 사용한다. 다만 1차 함수와 2차 함수의 경향성이 선형 관계식의 가정 하에 다음 크리깅 방정식을 풀게 된다.

$$z_{KED}^*(u) - m_{KED}^*(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{SK}(u)[z(u_{\alpha}) - m_{KED}^*(u_{\alpha})] \quad (5)$$

여기서  $m_{KED}^*(u) = a_0^*(u) + a_1^*(u)y(u)$ 를 나타낸다.

KED와 SKLM과의 차이점은 평균 혹은 경향성을 이끌어내는 방식에 있다. SKLM에서는 경향성의 계수인  $a_0^*$ 와  $a_1^*$ 를 크리깅 방정식과는 독립적으로 회귀분석을 통해 결정되는 반면, KED에서는 크리깅 방정식을 통해 각 위치별로 다른  $a_0^*(u)$ 와  $a_1^*(u)$ 이 결정되게 된다.

일반적인 KED의 추정치는 식 (6)으로 표현되며,  $(n(u) + 2)$ 개의 선형 방정식의 해로 가중치  $\lambda_{\alpha}^{KED}(u)$ 를 구하게 된다.

$$\begin{aligned} \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{KED}(u)C_R(u_{\alpha} - u_{\beta}) + \mu_0^{KED}(u) + \mu_1^{KED}(u)y(u) \\ = C_R(u_{\alpha} - u), \alpha = 1, 2, \dots, n(u_{\alpha}) \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{KED}(u) = 1 \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{KED}(u)y(u_{\beta}) = y(u) \end{aligned} \quad (6)$$

여기서  $\mu_0^{KED}(u)$ 와  $\mu_1^{KED}(u)$ 은 크리깅 제한조건을 만족시키기 위한 라그랑주 계수를 나타낸다.

지구통계 기법간 예측 정확도의 비교를 위해 one-leave-out 기반의 교차 검증을 수행하였다.  $n$ 개의 샘플링 자료를 고려할 때, 한 개를 제외한  $(n-1)$ 개의 지점의 자료를 이용하여 제외한 지점에서의 크리깅 예측값을 생성하였다. 이 과정을  $n$ 개의 샘플링 지점에 대해 반복하여 샘플링 지점  $z(u_{\alpha})$ 에 대해 이 점을 제외하였을 때의 예측값  $z^*(u_{\alpha})$ 을 계산하였다. 예측값의 정량적 비교 척도를 위해 평균 제곱 오차(Mean Square Error: MSE)와 다변량 지구통계기법인 정규 크리깅의 평균 제곱 오차에 대한 상대적 향상 지수(Relative Index)를 이용하였다.

$$MSE = \frac{1}{n} \sum_{\alpha=1}^n [z(u_{\alpha}) - z^*(u_{\alpha})]^2 \quad (7)$$

$$RI_i = \frac{(MSE_{OK} - MSE_i)}{MSE_{OK}} \times 100$$

여기서  $RI_i$ 는  $i$ 번째 내삽기법의 정규 크리깅의 평균 제곱 오차에 대한 상대적 향상 지수를 나타내며,  $MSE_{OK}$ 와  $MSE_i$ 는 정규 크리깅과  $i$ 번째 내삽기법의 평균 제곱 오차를 각각 나타낸다.

### 3. 사례 연구

이 연구에서는 다변량 지구통계학 기법의 적용성을 평가하기 위해 지하수위 분포도 작성과 퇴적물 입도 분포도 작성의 2가지 사례 연구를 수행하였다. 지하수위 분포도 작성 사례연구에서는 부가자료로 수치고도모형을, 퇴적물 입도 분포도 작성에는 IKONOS 원격탐사 자료를 부가자료로 이용하였다. 다변량 지구통계학 기법의 적용에는 GSLIB (Deutsch and Journel, 1998) 을 기본적으로 이용하였으며, 부가적인 자료처리는 포트란을 이용하여 프로그래밍하였다.

#### 1) 지하수위 분포도 작성

지하수위 분포도 작성 사례 연구는 경기도 이천시를 대상으로 농업진흥청의 광역수맥 조사 보고서 1991년부터 2000년까지의 자료에서 측정된 113개소의 자연수위

를 이용하였다(Fig. 1 (a)). 일반적으로 지하수면의 고도는 지형의 고도에 비례한다는 선행 연구 결과(Williams and Williamson, 1989; 함세영 외, 2001)에 기초하여 부가자료로 1:25,000 수치 지형도로부터 추출한 수치고도모형을 이용하였다(Fig. 1 (b)).

Fig. 2는 지하수위 측정 위치에서의 지하수위와 수치 고도의 관계를 나타내고 있다. 두 자료 사이의 상관계수가 0.993으로 아주 높게 나타나서 수치고도모형이 지하수위 분포도 작성에 유용하게 사용될 수 있음을 확인할 수 있다. SKLM에 사용되는 가변적 지역 평균은 선형 회귀분석을 통해 다음 식을 이용하였다.

$$z(u) = a_0^* + a_1^*y(u) = -1.89 + 0.99 y(u) \quad (8)$$

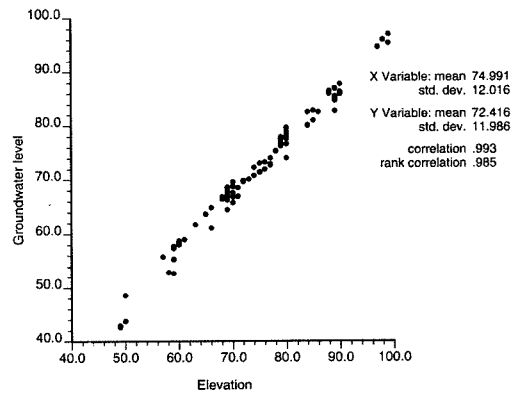


Fig. 2. Scatterplot of groundwater level and elevation at sampling locations.

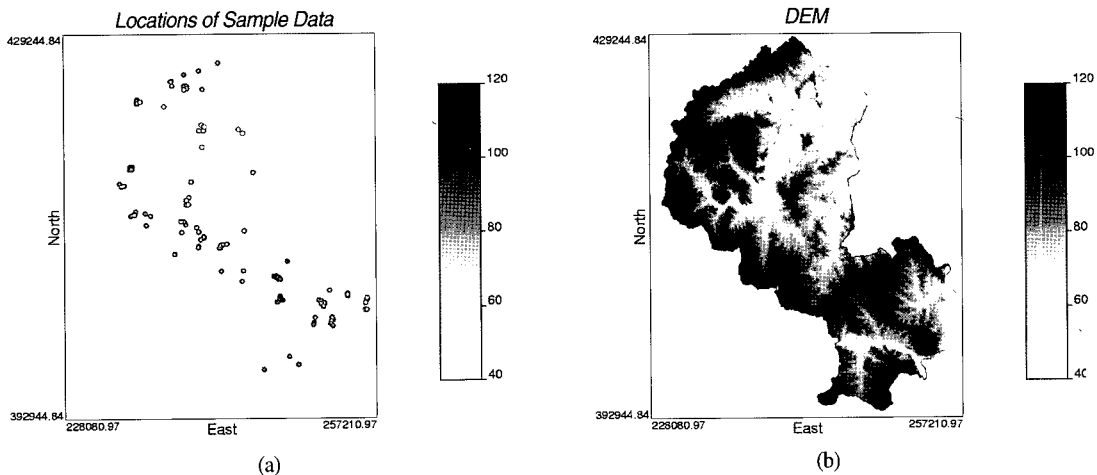


Fig. 1. (a) Location map of sample data, (b) Digital elevation model used for groundwater level mapping. Areas with gray color in (a) denote the study areas considered for groundwater level mapping.

여기서  $z(u)$ 와  $y(u)$ 는 지하수위와 지형 고도값을 각각 나타낸다.

지구통계학적 분석은 기본적으로 베리오그램의 계산과 모델링을 통해 공간적 상관성과 이방성을 우선적으로 고려한다. 여러 방향에 따른 베리오그램을 계산한 결과 특정 방향에 따른 이방성 효과가 나타나지 않아서 전방향(omnidirectional) 베리오그램을 고려하였다. Fig. 3 (a)와 (b)는 지하수위와 수치고도모델의 실험적 베리오그램과 구형(spherical) 모델로 모델링한 베리오그램 모델을 나타내고 있다. 두 경우 모두 대략 8000m 정도까지의 상관거리(range)를 나타내고 있으며, 작은 너겟 효과(nugget effect)를 나타내고 있다. 두 변수의 베리오그램 모델이 거의 유사하게 나오고 있는 양상은 Fig. 2에 나타난 두 변수 사이의 상관관계가 높은 것과 관련이 있으며, 이는 수치고도모델이 지하수위 분포도 작성에 효율적으로 이용될 수 있음을 간접적으로 지시한다. Fig. 3 (c)는 잔차의 실험적 베리오그램과 모델을 나타내고 있는데,

약 5300미터 정도의 상관거리를 보이고 있다.

다변량 지구통계 기법과 기존 단변량 지구통계 기법인 정규 크리깅을 실험적으로 비교하였는데, 크리깅 수행시 최대 24점까지의 샘플링 값을 고려하여 크리깅을 수행하였다. Fig. 4는 크리깅 수행 결과를 나타내고 있다. 112개의 샘플링 지점만을 이용하였을 경우에는 전반적인 값의 범위가 최대 100미터 정도를 나타내고 있다. 연구지역의 가운데 상단 부분이 60미터 미만의 낮은 지하수위 분포를 나타내고 있으며 전반적으로 평활화된 양상을 나타내고 있다. 반면 SKLM과 KED의 결과는 전반적으로 Fig. 1(b)의 수치고도모델의 양상을 보이면서 연구 지역의 좌상단 등의 부근에서 높은 지하수위 분포를 나타내고 있다. 이러한 양상의 차이는 샘플링 값의 분포에 기인한 것으로 판단된다. 즉 Fig. 1(a)에 제시된 것처럼 연구지역의 전체에서 샘플링되지 않고 주변부근에서 주로 샘플링이 되어 있으며, 최대값이 100미터 정도를 나타내고 있기 때문에 정규 크리깅을 수행

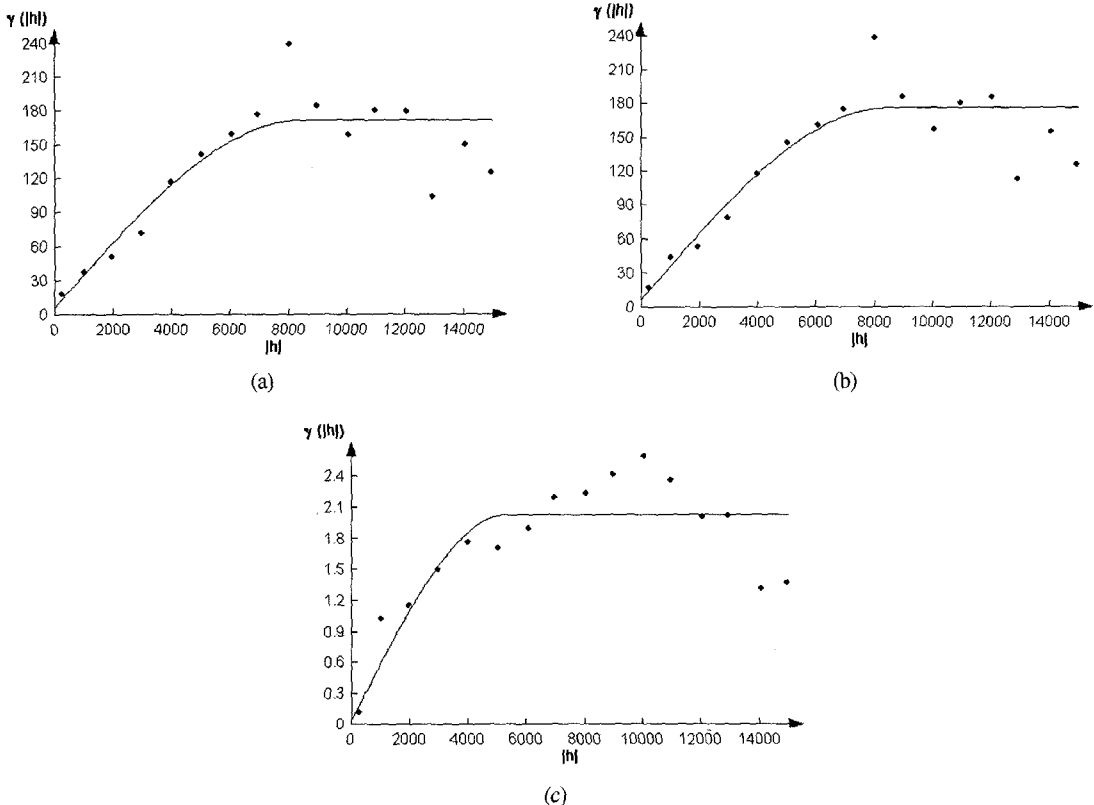


Fig. 3. Experimental variograms of (a) groundwater level, (b) elevation, (c) residual, with model fitted.

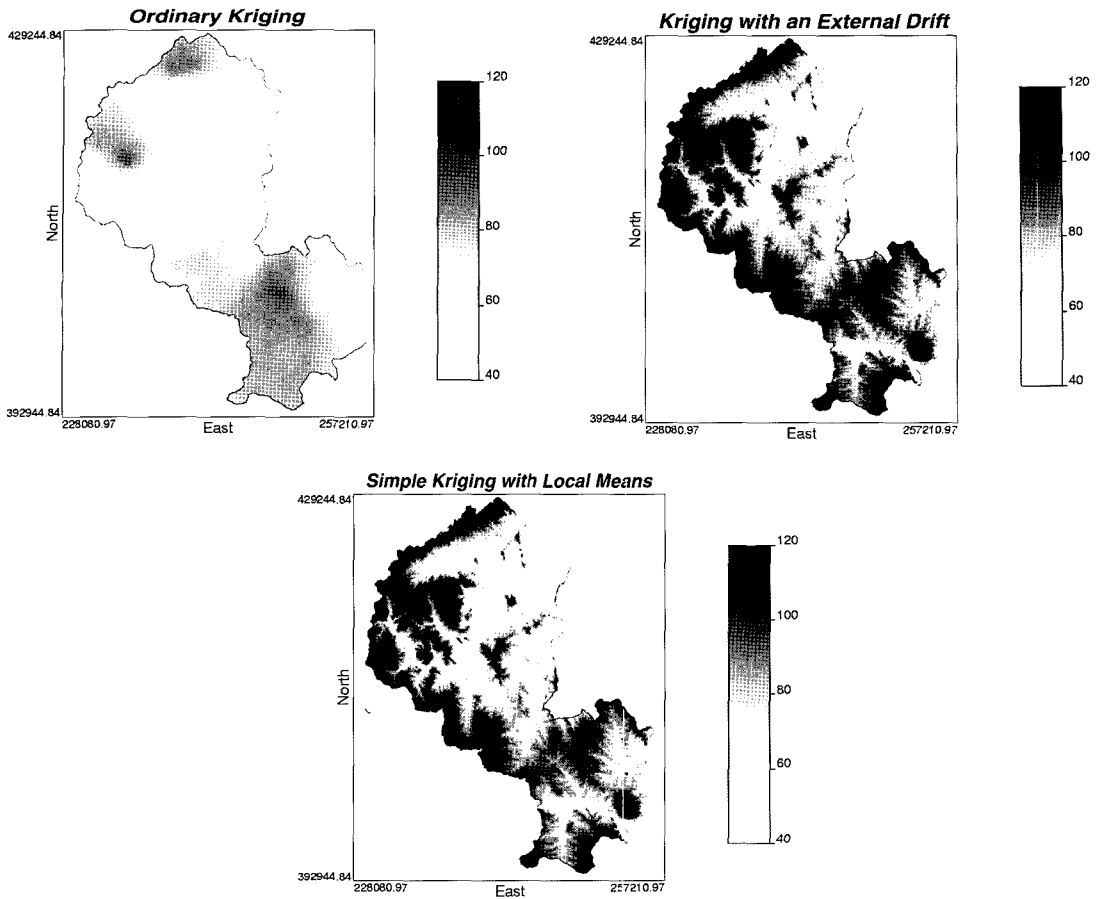


Fig. 4. Groundwater level distribution maps generated by geostatistical algorithms.

하였을 때 전반적으로 평활화되면서 샘플링 값의 최대값과 최소값 분포와 유사한 양상을 보이게 된다. 반면 SKLM과 KED의 경우에는 수치고도모델을 직접적으로 지하수위 분포에 이용하기 때문에 수치고도모델의 양상이 지하수위 분포 양상에 직접적으로 영향을 미쳐 최대값에서 차이를 나타내고 있다. 이러한 결과는 Fig. 5에 제시된 선형 회귀분석 예측치와 크리깅 결과의 차이의 히스토그램 분포를 통해 확인할 수 있다. 정규 크리깅의 경우에는 고도값에 비해 주로 낮은 값을 나타내고 있으며, SKLM과 KED는 거의 유사한 대칭 분포를 나타내고 있다. 다만 선형 회귀분석 예측치와의 차이값의 최대값과 최소값이 SKLM에서 KED에 비해 작게 나오고 있는데, 선형 회귀값의 계수를 결정하는 방식의 차이에 기인한 것으로 판단된다.

Table 1은 단변량 지구통계 기법인 정규 크리깅, 선

형 회귀분석, 다변량 지구통계 기법의 교차 검증 결과를 나타내고 있다. 선형 회귀분석의 경우에는 113개의 모든 자료를 이용한 선형 회귀식에 대한 각 샘플 위치별 편차에 대해 평균 절대 오차와 평균 제곱 오차를 계산하였는데, 그 결과 선형 회귀분석의 오차값은 다소 저추정 (underestimation) 결과를 나타낼 수 있다. 부가자료인 고도자료를 고려하지 않은 정규 크리깅에 비해 다변량 지구통계 기법이 월등히 우수한 예측 오차를 나타내고 있다. 가변적 지역 평균의 방식에 있어서 차이가 있음에도 불구하고 SKLM과 KED는 모두 유사한 예측 능력을 나타내고 있다. 이러한 정량적 결과는 부가자료가 예측하고자 하는 변수와 강한 상관관계를 나타낼 경우 보다 높은 예측 능력을 가지고 샘플링되지 않은 지점에서의 값을 예측할 수 있음을 나타낸다. 샘플링 지점이 접근성 등의 이유로 특정 구역에 집중될 경우 샘플링 되지 않은

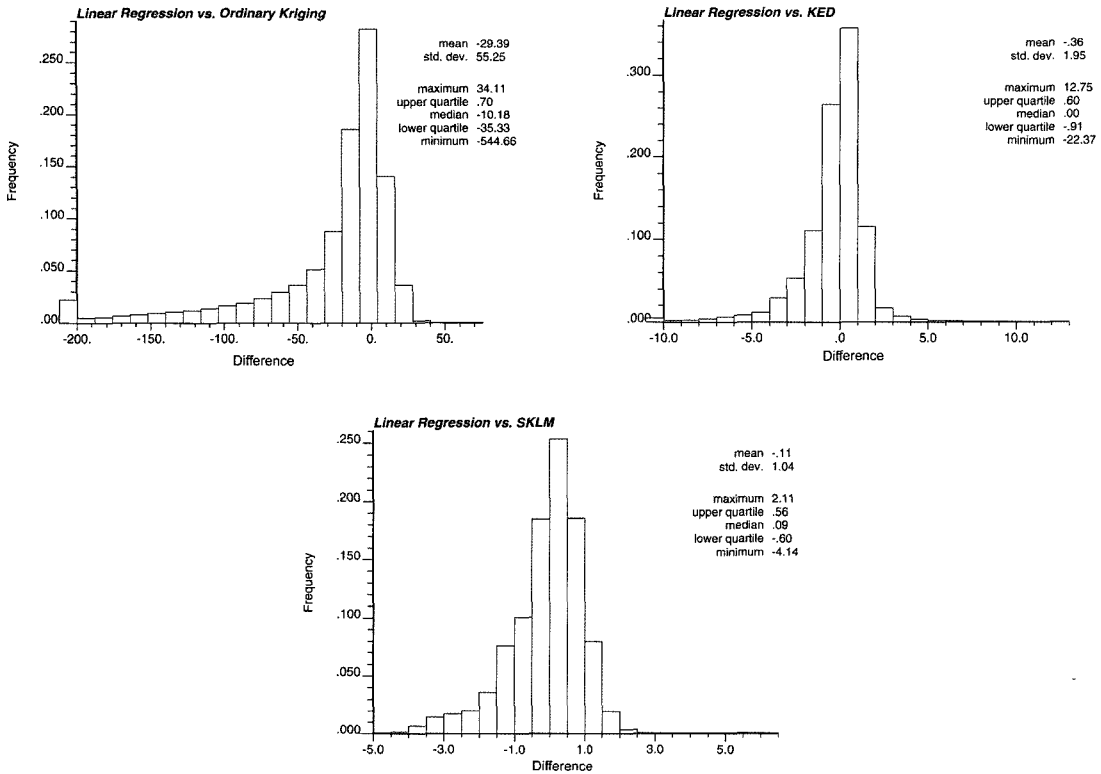


Fig. 5. Histograms of differences between linear regression estimates and kriging results.

Table 1. Cross validation results in a case study for groundwater level mapping.

Algorithm	Mean Square Error (MSE)	Relative Index (%)
Ordinary Kriging	23.48	-
Linear Regression	2.16	90.80
KED	0.70	97.02
SKLM	0.68	97.10

지점에서의 값은 베리오그램을 통해 얻어지는 상관거리에 의존하여 예측할 수밖에 없기 때문에, 이 경우 미지의 점에서의 예측값의 오차가 클 것을 예상할 수 있다. 따라서 이러한 내삽 결과에 기반한 지하수위 분포도를 2차적인 분석에 사용할 경우 예측에 수반되는 불확실성이 커서 2차 분석 결과에도 크게 영향을 미칠 것으로 예상된다. 따라서 샘플링 지점의 개수나 공간적 분포가 균일하지 않을 경우 부가자료를 이용하여 내삽을 수행하는 것이 나올 것으로 판단된다. 또한 수치고도모델의 선형식에 의한 단순 변환에 해당되는 선형 회귀분석에 비

해 다변량 지구통계기법의 예측 오차가 상대적으로 작게 나타나고 있다. 선형 회귀분석에서는 식 (8)을 통해 고도값의 단순 선형 변환을 취하게 되는데, 이때 잔차가 공간적으로 상관되어 있지 않다는 가정이 내포되어 있다. 그러나 Fig. 3(c)에 나타난 것처럼 잔차가 공간적 상관성을 나타내고 있기 때문에 이러한 공간적 상관성을 고려하는 다변량 지구통계기법을 적용함으로써 예측의 정확도를 높힐 수 있던 것으로 판단된다.

## 2) 입도분포도 작성

두 번째 사례연구에서는 퇴적물 입도분포도 작성을 위해 현장에서 채취한 퇴적물의 입도 분석 자료와 고려하고 있는 연구지역에서 모든 값을 얻을 수 있는 고해상도 원격탐사 자료인 IKONOS 자료를 이용하였다. 이 연구에서 과거에 많이 사용되어온 Landsat 자료를 사용하지 않고 IKONOS 영상 자료를 사용한 이유는 퇴적물 현장 샘플지점이 30cm 이내로 Landsat의 공간해상도 30m를 사용하면 현실세계와 다른 결과치를 반영할

수 있어서 고해상도 위성영상을 사용하였다.

대상지역은 안면도 남쪽 해안의 바람아래 해수욕장으로 행정구역상 태안군 고남면 장곡리에 해당한다. 바람아래 해수욕장 주변의 해안지형은 전형적인 리아스식 해안으로 굴곡이 매우 심하여 간석지가 많이 발달되어 있다. 본 간석지는 지형상 작은 만입부에 해당되며, 전면에 할미섬, 섬웃섬 등의 시스택, 파석대 및 해안사구 등이 파랑을 막아주어 대규모 간석지가 발달할 수 있었다. 남쪽 및 서쪽 연안은 연안사취와 사빈 및 해안사구가 발달하였다. 연구지역의 수심은 전체적으로 10m 미만으로 얕으며, 장곡리 내만지역의 수심은 평균 1.5m 정도로 1~5도 정도의 완만한 경사지를 유지하고 있다. 바람아래 해수욕장 전면의 수심은 5m 미만으로 5~10도 정도로 비교적 완만하게 외해로 열려있다(장동호 외, 2003).

바람아래 해수욕장 주변의 입도분포도 작성을 목적으로 현장 실측, 시료채취 및 고해상도 위성자료 분석을 수행하였다. 이 지역의 고해상도 위성자료로는 2001년 2월 26일에 획득된 IKONOS 자료를 분석에 사용하였다(Fig. 6 (a)). 본 연구지역이 포함되는 검조소가 없기 때문에 바람아래 해수욕장과 가장 가까운 보령검조소의 조석 자료를 이용하였는데, 영상 획득 시기와 계절과 조석자료가 유사한 2002년 2월 26일 오전에 현장 실측을 수행하였다. 현장 시료채취는 총 53개 지점에서 획득하였으며(Fig. 6 (b)), 채취된 시료는 실내에서 자연 건조시킨 후, 63 $\mu$ 의 체를 사용하여 습식 분석하여 빨과 모래 및 자갈로 분리시켰으며, 모래와 자갈은 10간격으로 건식 분석하여 입도별 중량백분율을 구하였다. IKONOS 영상 자료에서 할미섬과 섬웃섬은 마스킹 처리하였으

며, 수계영역과 육상 영역은 무감독 분류를 통해 이후 분석에서 제외하였다.

우선 IKONOS 다중 분광 밴드 자료와 입도 분포와의 상관성을 살펴보기 위해 밴드별 산포도를 작성하였다(Fig. 7). 모든 밴드에서 입도분포와 DN값은 음의 상관관계를 보이고 있으며, 밴드 2가 가장 상관관계가 높은 것으로 나타났다. 이러한 상관관계에 기초하여 IKONOS 자료를 부가자료로 이용하여 연구지역의 입도분포도를 작성하였다. KED는 부가자료를 1가지만 이용할 수 있기 때문에, 가장 높은 상관성을 나타낸 IKONOS 밴드 2를 이용하였으며, SKLM의 적용에는 밴드 2만을 이용한 경우와 모든 밴드를 이용한 경우 2가지로 구분하여 각각 적용하였다. Table 2는 SKLM에 선형 회귀분석을 통해 얻어진 결정 계수를 나타내고 있다. 밴드 2만을 이용하였을 때에 비해 모든 밴드의 선형 회귀분석 결과가 상대적으로 적은 잔차 오차를 나타내고 있다.

입도 분포 자료의 공간적 상관성을 살펴보기 베리오그램 분석을 수행하였다. 이방성 분석결과, 135도 방향으로 강한 이방성이 발견되어 135도와 이와 수직된 방향인 45도에 대해서 구역 이방성을 고려한 이방성 모델링을 수행하였다(Fig. 8 (a), (b)). 또한 SKLM을 위해 Table 2에 제시된 밴드 2와 모든 밴드를 가변적 지역 평균을 이용한 경우에 대해 각각 잔차를 계산하여 실험적 베리오그램을 계산하고 모델링을 수행하였다. 135도 방향의 강한 경향성이 사라졌기때문에 잔차의 경우 등방성을 나타내어 전방향 베리오그램을 고려하였다(Fig. 8 (c), (d)).

입도 분포도 작성을 위해 단변량 지구통계 기법으로는 정규 크리깅과 경향 크리깅을 적용하였다. 입도 분포

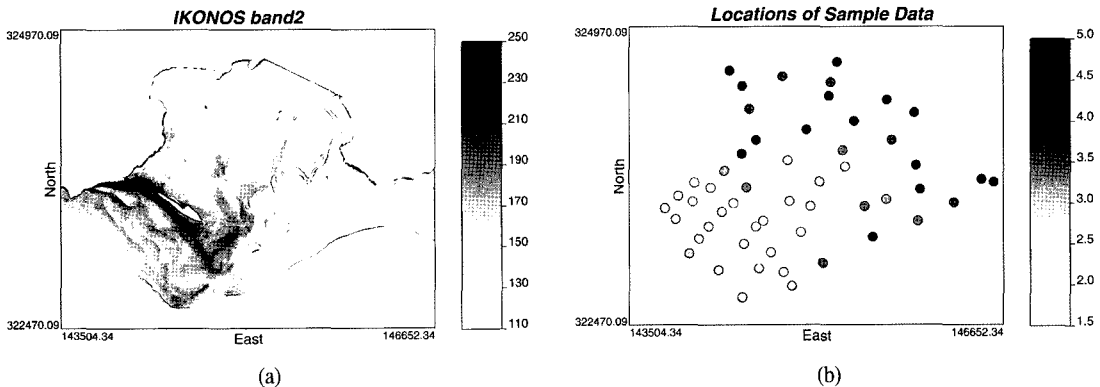


Fig. 6. (a) IKONOS band2 imagery, (b) Location map of sample data. Symbols in (b) are enlarged for visibility.



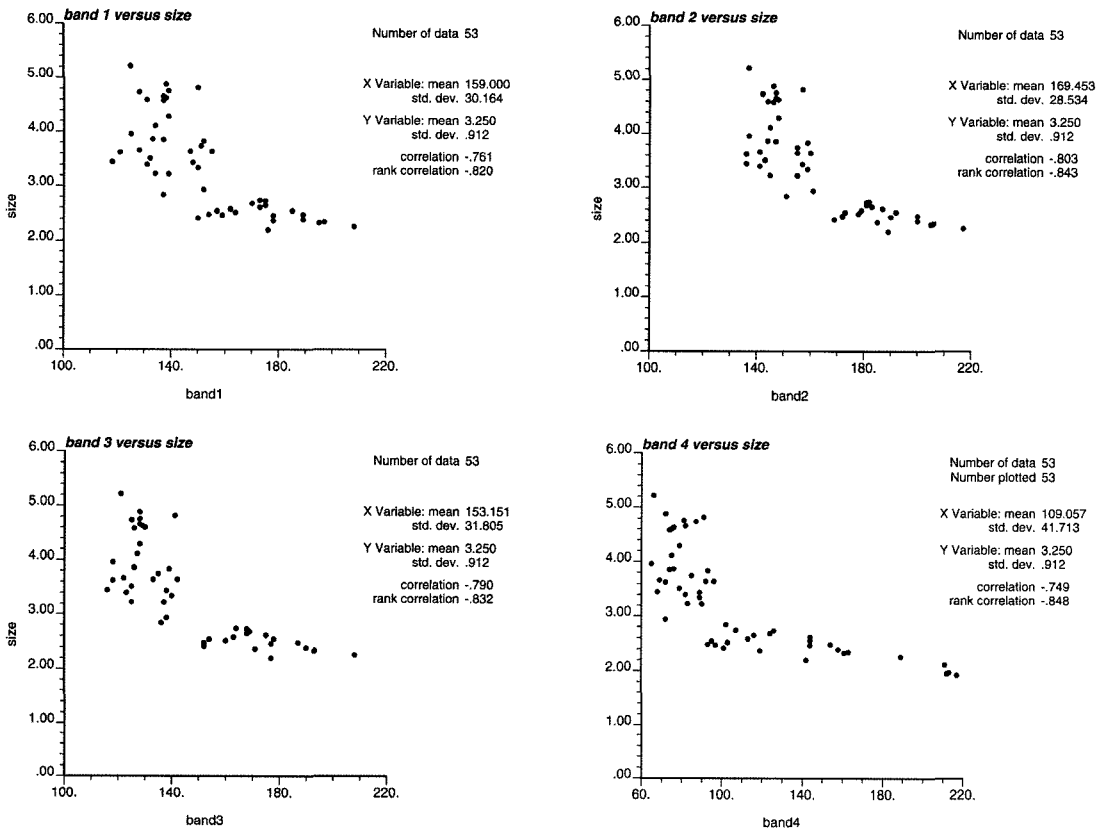


Fig. 7. Scatterplots of grain size and DN values of each IKONOS imagery at sampling locations.

Table 2. Linear regression coefficients for the SKLM approach.

Case	Equation	Residual standard error
Band2	$z(u) = 7.5971 - 0.0257 y_{band2}(u)$	0.554
All bands	$z(u) = 11.8356 - 0.0898 y_{band1}(u) - 0.2063 y_{band2}(u) + 0.0806 y_{band3}(u) - 0.0024 y_{band4}(u)$	0.448

가 특정 방향으로 이방성을 나타냈기 때문에 이러한 경향성을 자체적으로 모델링을 할 수 있는 경향 크리깅을 새로이 적용하였다. 다변량 지구통계 기법으로는 IKONOS 밴드 2를 부가자료로 이용하는 KED, IKONOS 밴드 2와 모든 밴드를 각각 부가자료로 이용한 SKLM를 고려하였다(Fig. 9). 정규 크리깅과 경향 크리깅 결과를 살펴보면 전반적으로 평활화된 유사한 양상을 보이고 있으며, 경향 크리깅 결과에서 연구지역 상단 가운데 부근에서 높은 값을 나타내고 영역이 연결되어 나타나고 있다. 크리깅은 일종의 최소오차 회귀분석 방법이므로 자료가

존재하지 않는 임의의 지점에 대해 분산을 최소화 하는 값을 추정하기 때문에 평활화된 공간 분포를 보이게 된다. 반면 IKONOS 자료를 부가자료로 이용한 KED와 SKLM의 경우에는 국소적인 입도 분포의 특성을 반영하고 있다. 바람아래 해수욕장 주변의 입도의 공간적인 분포를 정성적으로 해석하면 다음과 같다. 정규 크리깅과 경향 크리깅은 유사한 공간적 입도분포 패턴을 나타낸다. 특히 연구지역의 만입지를 중심으로 점토의 평균 함유량이 60%이상인 sandy mud가 광범위하게 나타난다. 퇴적물의 유형은 사질과 실트가 혼재하며, 평균입도는 4.0~4.5이다. 이들 sandy mud 전면에는 평균입도 3.0~3.5의 실트질 모래로 구성된 muddy sand가 외해와 연결되면서 가장 넓은 면적에 걸쳐 분포한다. 이밖에 할미섬을 중심으로는 모래의 평균함유량이 90%이상인 sand 퇴적물이 해변 및 사주의 형태로 나타난다. 평균입도는 1.5~2.00 세립사로 구성되어 있으며, 이들 지역의 퇴적물의 평균입도의 전체적인 변화 양상은 외해로

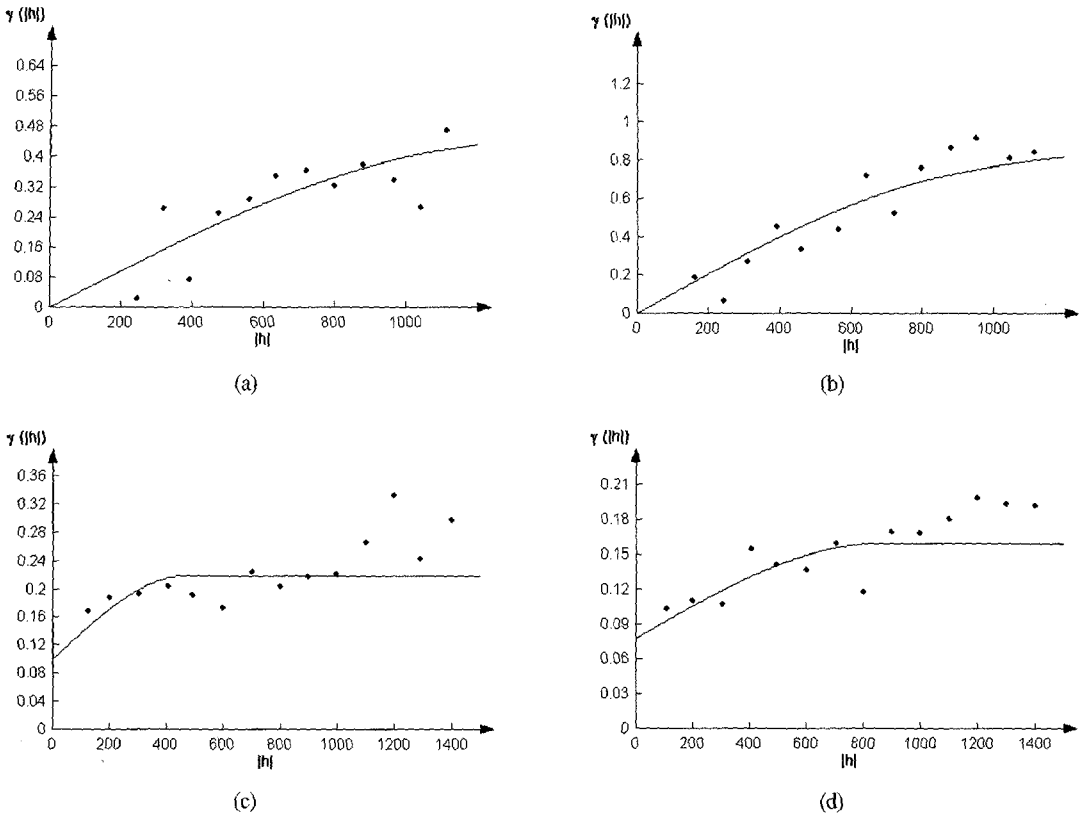


Fig. 8. Experimental variograms of (a) grain size sample data in 135 degree direction, (b) grain size sample data in 45 degree direction, (c) residual of the use of IKONOS band 2, (d) residual of the use of IKONOS all bands, with model fitted.

가면서 세립화되는 경향을 보인다. KED 결과에서는 정규 크리깅 및 경향 크리깅과 유사한 입도 공간분포 특징을 나타내지만, 패턴 및 미세한 부분의 표현에서 현실세계를 보다 많이 반영하고 있는 특징을 보인다. 특히, 일부 갯골지역을 중심으로는 기존 도면에서 나타나지 않았던 gravel의 함량이 80% 이상인 역질퇴적물이 협소하게 분포하며, 평균입도는 2.5~3.0 $\phi$ 이다. 또한 할미섬을 중심으로 발달한 해변 및 사주는 분포패턴이 기존 도면과 다르게 명확히 구분되어 나타나며, 그 배후에는 실트질 모래로 구성된 muddy sand와 sandy mud가 차례로 나타난다. 또한 간석지 내에서 조류의 방향에 의해 생겨나는 갯골들이 여러 개 구분되는 특징을 보인다. IKONOS 자료를 부가 정보로 이용한 SKLM에서는 가장 현장의 입도분포도와 유사한 공간분포 패턴이 나타난다. 특히 만입지의 sandy mud가 협소하게 분포하며, 평균입도 3.0~3.5 $\phi$ 의 muddy sand가 가장 넓은 면적에 걸쳐 분포한다. 해변 및 사주 역시 현장자료와 가장 유

사한 분포패턴을 나타내지만, 일부 지역에서는 다른 입도분포 패턴으로 나타난다. IKONOS 밴드 2를 이용한 SKLM에서는 해변 및 사주가 명확한 입도분포의 경계를 가지면서 나타나지만, 모든 밴드를 다 이용한 경우에는 불분명하게 분포한다. 또한 갯골 역시 IKONOS 밴드 2를 이용한 SKLM에서는 여러 개의 갯골 패턴이 서로 다른 입도분포 특징으로 보이지만, 모든 밴드를 다 이용한 경우에는 잘 나타나지 않는다. 그러나 세파대 (swash zone) 부근에서 파랑의 에너지가 집중되어 입도분포가 조립화되는 특징을 관찰할 수 있다.

다양한 지구통계 기법의 정량적 예측 능력의 비교를 위해 교차 검증을 수행하였고, 결과는 Table 3에 제시되어 있다. 지구통계기법과의 비교를 위해 IKONOS 밴드 2와 모든 밴드를 각각 선형 회귀분석에 사용한 결과를 함께 제시하였다. 부가자료를 이용하는 다변량 지구통계기법을 적용한 경우가 상대적으로 작은 추정 오차를 보이고 있다. 그러나 지하수위 분포도 작성 사례연구

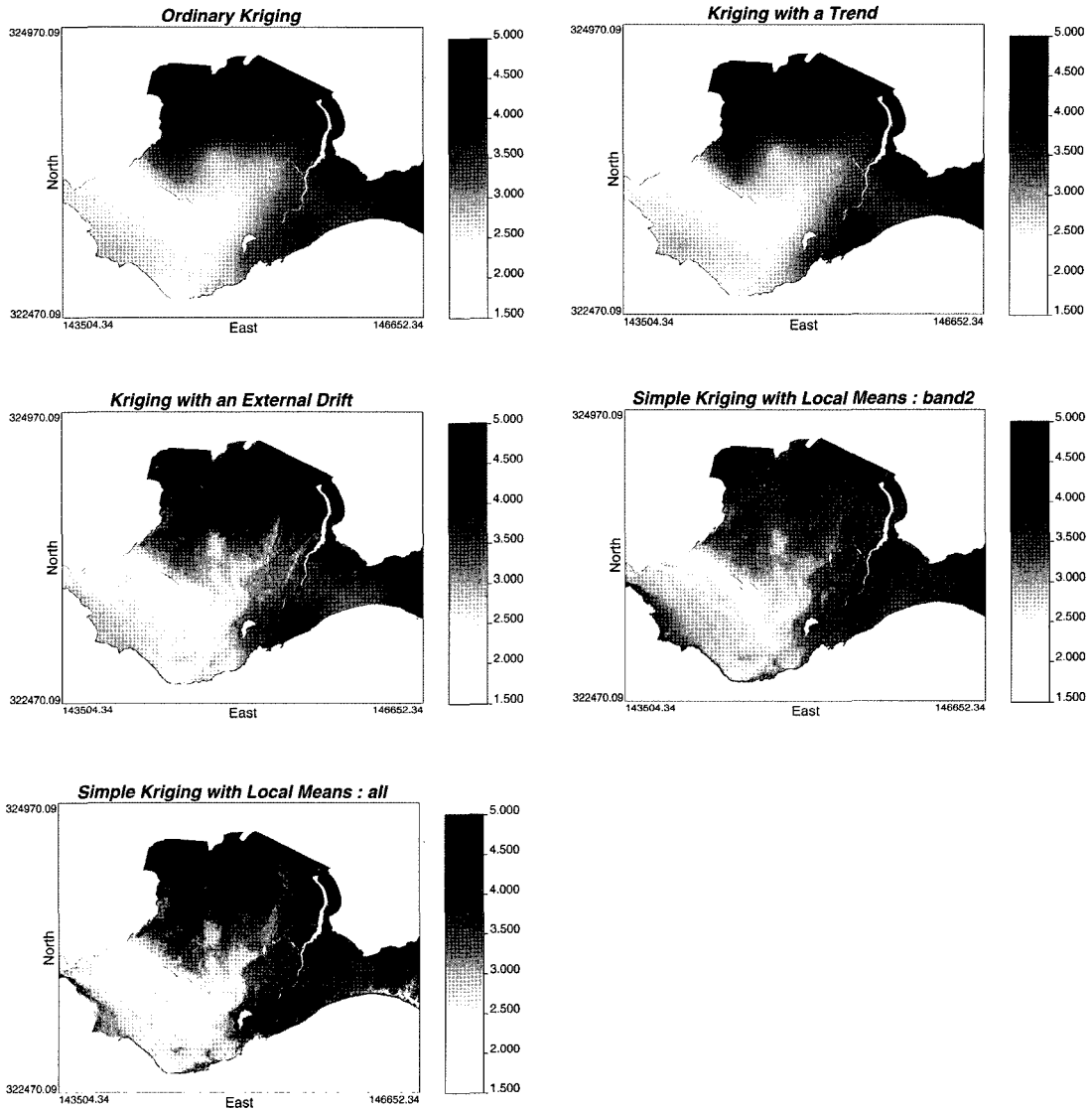


Fig. 9. Grain size distribution maps generated by geostatistical algorithms.

Table 3. Cross validation results in a case study for grain size mapping.

Algorithm	Mean Square Error (MSE)	Relative Index (%)
Ordinary Kriging	0.241	-
Kriging with a Trend	0.233	3.32
Linear Regression of IKONOS band 2	0.295	-22.41
Linear Regression of IKONOS all bands	0.230	4.56
KED	0.162	32.78
SKLM with IKONOS band 2	0.183	24.07
SKLM with IKONOS all	0.138	42.74

와 비교하였을 때 단변량 지구통계기법을 적용한 경우가 선형 회귀분석 적용 결과에 비해 유사하거나 작은 추정 오차를 나타내고 있으며, 다변량 지구통계기법사이의 차이가 두드러지지 않고 있다. 입도 분포값의 범위가 작은 점을 고려하더라도 이는 부가자료와 입도자료와의 상관성의 차이에 기인한 것으로 판단된다. 지하수위 분포도 작성의 경우, 부가자료로 사용한 수치고도모델과 지하수위 분포의 상관성이 0.99로 상당히 높았으나, 입도 분포도 작성에 사용한 IKONOS 자료는 상대적으로 작은 밴드별로 0.7에서 0.8 사이의 상관성을 보였다. 따라서 부가자료와의 상관성이 아주 높지 않을 경우에 부가자료를 사용하지 않고 샘플링 값들의 공간적 상관성을 고려하는 크리깅 기법과 선형 회귀분석 혹은 다변량 지구통계기법 사이의 차이가 두드러지지 않은 것으로 판단된다. 또한 입도 분포값과 잔차의 공간적 상관성에도 기인한 것으로 판단된다. Fig. 8에 제시된 베리오그램을 살펴보면, 입도 분포값은 너겟값이 거의 0에 가까운 값으로 잡음 효과가 적음을 나타내고 있다. 그러나 밴드 2와 모든 밴드를 이용한 경우의 잔차들은 모두 상대적으로 큰 너겟값을 보이고 있으며 또한 전체 공분산에 대한 상대적인 너겟 효과가 약 48%로 높게 나타나고 있다. 이러한 높은 상대적인 너겟 효과는 잔차 자료사이의 공간적 상관성이 약함을 의미하기 때문에 부가자료를 이용한 효과가 상대적으로 크지 않은 것으로 판단된다. 강우 예측을 위해 수치고도모델을 이용한 Goovaerts (2000)에서는 강우와 고도사이의 상관계수가 0.75 이하일 경우 정규 크리깅이 선형 회귀분석보다 정확한 예측 능력을 보였으며, SKLM의 적용시 잔차 자료사이의 상관성이 떨어질수록 부가자료의 이용 효과가 작아진다고 보고하였다. 따라서 선형 연구 결과와 유사하게 부가자료와의 상관성과 공간적 상관성의 차이로 인해 상대적으로 부가자료 이용의 효과가 크지 않은 것으로 판단된다. 지하수위 분포도 작성 사례연구와 대별되는 또다른 특징은 샘플링 밀도이다. 지하수위 분포도 작성에 사용한 지하수위 샘플링 자료는 접근성 등의 문제로 연구지역에 골고루 분포하는 것이 아니라 특정 지역에 집중되어 나타나고 있다. 따라서 이러한 경우에는 부가자료가 예측에 많은 정보를 제공할 수 있다. 입도 분포 작성 사례연구에서는 연구지역에 골고루 샘플링 자료가 분포되어 있어 부가자료의 효과가 상대적으로 두드러지지 않은 것으로

판단된다.

#### 4. 토의 및 결론

공간적으로 샘플링 밀도가 높지 않은 자료를 이용한 주제도 작성의 정확도 향상을 위해 부가자료를 통합할 수 있는 다변량 지구통계기법을 적용하여 활용성을 검토하였다. 지구통계학은 공간자료의 상관성을 내삽에 이용할 수 있을 뿐만 아니라 부가자료를 통합할 수 있는 방법론을 제시한다. 지하수위 분포도 작성과 입도 분포도 작성의 사례연구를 통해 다변량 지구통계기법은 기존에 많이 사용되어온 단변량 지구통계기법에 비해 공간적인 분포에 있어서 크리깅의 특징인 평활화 효과를 완화하여 국소적인 특성을 반영할 수 있으며, 예측 능력 면에서도 우수한 것으로 나타났다. 그러나 다변량 지구통계 기법의 효과는 고려하고 있는 변수와 부가자료와의 상관성, 개별 자료들의 공간적 상관성과 샘플링 밀도에 영향을 받는 것으로 나타났다. 지하수위 자료와 상관성이 매우 높은 수치고도모델을 이용할 경우, 지하수위 자료 자체와 잔차의 높은 공간적 상관성과 결합되어 샘플링 위치가 연구지역에 골고루 위치하지 못한 자료만을 이용하는 단변량 지구통계 기법에 비해 우수한 예측 능력을 나타내었다. 그러나 부가자료와의 상관성이 상대적으로 떨어지고, 잔차의 낮은 공간적 상관성으로 인해 입도 분포도 작성에서는 이러한 효과가 두드러지지 않았다. 앞으로 다양한 사례연구를 통해 전술한 요소들의 효과에 대한 연구가 필요할 것으로 판단된다.

방법론 측면에서 다변량 지구통계기법으로 적용한 KED와 SKLM은 부가자료를 가변적 지역 평균으로 사용하는 방식의 차이가 있지만, 여러 개의 부가자료를 이용해야 할 경우에는 차이가 있다. 하나의 부가자료를 가변적 지역 평균으로 사용하고 선형관계를 만족해야만 하는 KED에 비해 다양한 부가자료를 변환에 의해 이용할 수 있는 SKLM의 적용성이 더 크다고 간주할 수 있다.

이 논문에서는 적용한 지구통계기법은 샘플링되지 않은 지점에서의 값을 예측하는 것이 목적이지만, 예측에 수반되는 불확실성을 모델링할 수는 없다. 크리깅의 결과로 예측값과 오차 분산값이 함께 계산이 되지만, 오

차 분산값은 자료의 배치 및 구성에 관한 정보를 제공하지만 예측과 관련된 불확실성의 척도가 될 수는 없다. 부가자료를 이용하였을 경우와 이용하지 않을 경우의 효과는 불확실성 분석을 통해 교차 검증에 의한 추정 오차 분석과 더불어 보다 정량적으로 비교될 수 있다. 앞으로 국소적인 불확실성 분석을 수행할 수 있는 비모수적 접근인 지시자 크리깅(indicator kriging)이나 모수적 접근인 다중 가우시안 크리깅(multi-Gaussian kriging)을 적용하여 부가자료에 의한 불확실성 감소 효과를 분석할 예정이다.

## 사 사

이 연구는 한국항공우주연구원의 다목적실용위성 3호 개발사업 위탁과제 “지질환경 변화정보 추출 및 분석 기술 개발”과 한국지질자원연구원 기본사업의 연구비 지원에 의해 수행되었습니다. 지하수위 분포도 작성에 사용된 자료를 제공해준 구자혁 님께 감사드립니다.

## 참고문헌

장동호, 김만규, 김광훈, 2003. 안면도 바람아래 해수욕장 주변의 퇴적환경 변화 연구, 한국지형학회지, 10(2): 143-155.  
 정상용, 이강근, 1995. 난지도 매립지 일대의 지하수위 분포 추정을 위한 복합 크리깅의 응용, 지하수환경, 2(2): 58-63.

최종근, 2002. 공간정보 모델링: 크리깅과 최적화 기법, 구미서관.  
 함세영, 이병대, 성익환, 정재열, 2001. 영천담 도수로 터널 굴착 이전의 자연수위분포, 자원환경 지질학회 춘계 공동학술발표회 논문집, 30-32.  
 Deutsch, C. V. and A. G. Journel, 1998. *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York.  
 Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York.  
 Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology*, 228: 113-129.  
 Issaks, E. H. and R. M. Srivastava, 1989. *An Introduction to Applied Geostatistics*, Oxford University Press, New York.  
 Raspa, G., M. Tucci, and R. Bruno, 1997. Reconstruction of rainfall fields by combining ground raingauges data with radar maps using external drift method, In: Baafi, E. Y. and N. A. Schofield (Eds.), *Geostatistics Wollongong '96*, Kluwer Academic, Dordrecht, 941-950.  
 Williams, T. A. and A. K. Williamson, 1989. Estimating water-table altitudes for regional ground-water flow modeling, US. Gulf coast 1989, *Ground Water*, 27(3): 333-340.