

지형특성을 고려한 효과적인 DEM 생성

황철수*, 구자용**, 오충원**

*서울대학교 국토문제연구소 연수연구원, **서울대학교 대학원 박사과정

1. 서론

1) 연구목적

수치고도모형(DEM, Digital Elevation Model)은 일반적으로 규칙적인 격자 행렬(grided matrix)을 통해 공간상에 나타나는 지형기복의 변화를 연속적으로 표현하는 모형이라고 정의된다(Burrough, 1986). 2차원 평면으로 표현되는 단위 격자는 위치(x, y 좌표)와 고도값(z 값)을 수치적으로 표시하고 있기 때문에 지형면을 수치적으로 해석하기에 적합한 구조를 가지고 있다. 이와 같은 특성으로 인해 DEM은 컴퓨터로 처리하기에 알맞은 자료구조이고 특히 GIS 환경에서 편리하게 처리할 수 있게 됨에 따라 3차원의 지형면(surface)을 나타내는 대표적 모형으로 평가되고 있다. DEM은 지리학 연구에서도 한 지역의 기본적 지형특성(예를 들어 경사도, 사면방향, 지형단면 등)과 하계망 특성(유역 구분, 하계망의 위상구조 등), 혹은 지형과 관련된 통계적 분석 등에 자주 활용되고 있으며 군사적 목적과 도시계획 혹은 단지설계 등과 같은 분야에서 광범위하게 이용되고 있다. 최근 GIS 또는 컴퓨터 지도학 분야의 연구가 활성화되고 국가적으로 지리정보를 구축하면서 우리 나라에서도 이에 대한 관심이 증대되고 있다. 특히 1995년 이후 국가지리정보체계(NGIS)구축사업이 본격적으로 진행되면서 국가기본지리정보 가운데 하나로 DEM을 구축하여야 한다는 논의가 활발히 진행되고 있다.

한편 DEM을 구축하는 데는 측량조사자료, 수치지도 혹은 종이지도의 등고선 자료, 항공사진측량 자료, 인공위성이미지 등이 자료원(source)으로 이용되고 있다. 그런데 국가적 규모로 DEM을 개발하고자 할 경우에는 대부분 국가에서 수치지도나 종이지도에 표시된 등고선을 활용하고 있다. 이렇게 DEM을 개발하는데 등고선 자료를 원자료로 활용하는 데는 상대적으로 다른 자료들에 비해 비용을 줄이면서 높은 정확도를 유지할 수 있기 때문이다. 또한 등고선에서 DEM을 개발하기 위한 과정에서 내삽 과정(interpolation processes)이 일반적으로 적용된다. 여기에서 내삽 알고리즘은 등고선으로부터 DEM을 개발하는 과정에서 DEM의 정확도와 효율성을 결정짓는 요인으로 작용한다. 따라서

본 연구는 우리나라에서 현재 구축이 진행되고 있는 수치지형도의 등고선 자료를 활용하여 DEM을 개발할 때 지형적 특성을 고려한 효과적인 내삽 방법을 탐구하는데 목적을 두고 있다. 구체적으로 첫째 내삽 알고리즘 간의 정확도를 비교·평가하고, 둘째 지형적 특성에 따라 DEM의 정확도에 영향을 미치는 내삽 알고리즘이 존재하는가를 밝히고, 셋째 내삽 과정의 효율성을 향상시키기 위해서 공간적 표본추출방법에 따른 내삽 알고리즘을 적용한 DEM의 정확도를 평가한다.

2) 연구지역과 방법

본 연구에 사용한 수치지형도의 축척은 1:25,000이며, 연구의 대상지역은 1:25,000 지형도에서 도영명 공주에 해당하는 지역으로 지형이 산지와 계곡, 그리고 평지가 함께 분포하는 특성을 갖고 있다. 우리 나라 수치지형도는 지도 사상(事象)들을 주제별로 분류하고 코드를 부여하여 레이어(layer)라는 단위로 관리하고 있다. 국가 수치지형도 표준안에 의하면 등고선에 해당하는 코드는

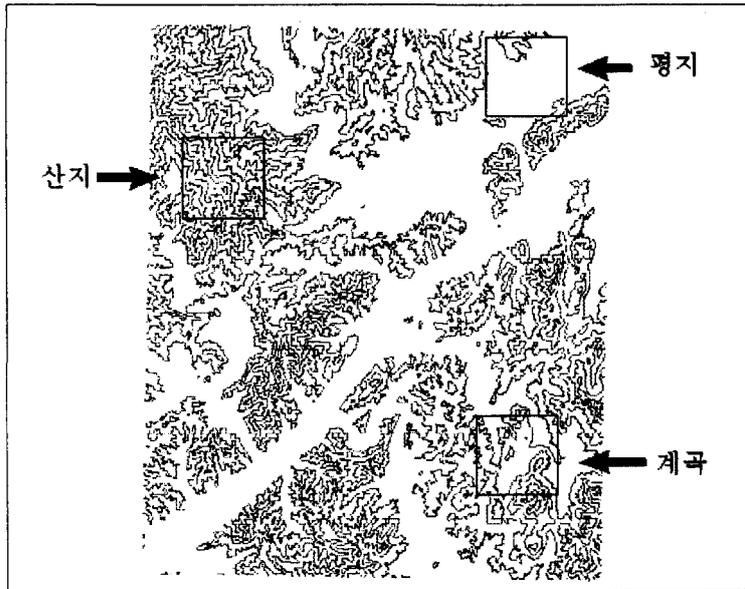


그림 1. 연구 표본 지역

711로 시작하는 레이어이다. 본 연구지역의 수치지형도에 이 범주에 속하는 레이어는 7111(주곡선)과 7114(계곡선)이며, 두 가지 등고선 레이어를 추출하여 수치고도모형의 개발을 위한 기본자료로 이용하였다.

표본지역은 지형특성을 고려하여 2km×2km 크기로 선정하였다. 전체 대상지역에서 지형특성에 따라 평지, 계곡, 산지의 세가지 지역을 선정하였다. 즉 지형의 기록이 거의 없는 지역을 평지 표본으로, 경사면과 계곡을 가지고 있는 지역을 계곡 표본으로, 산봉우리와 능선으로 이루어진 지역을 산지 표

본으로 선정하였다. 공주 도엽 중에서 선정된 표본지역의 위치는 그림 1과 같다.

한편, 본 연구에서는 내삽을 통해 개발되는 DEM의 정확도를 평가하기 위한 척도로 RMSE(root mean square error)를 사용하였다. RMSE는 각각의 오차값들이 갖는 오차의 평균량을 평가하는 일

$$RMSE_z = \sqrt{\frac{\sum (Z_i - Z_n)^2}{N}} \quad Z_i : \text{측정 자료의 } i\text{번째 위치의 고도값}$$

Z_n : 정확도가 높은 참조자료의 i 번째 위치의 고도값
 N : 자료의 개수

반적 방법으로, 지도와 같은 공간적 자료에 대해 적용할 경우 동일한 좌표계를 가지는 두 개의 자료 집단을 직접 비교하여 계산함으로써 해당 자료 집단의 정확도를 단일 척도로 계산한다(Maling, 1989). 계산된 RMSE 값이 크면, 참조자료(Reference Dataset)와 비교자료(Test Dataset)에서의 속성값의 차, 즉 오차가 더 크다는 것을 의미한다. RMSE는 관측점의 고도값과 내삽결과의 해당지점 고도값의 직접비교가 가능하고 계산과정이 간단하여 각국의 지도제작기관에서 보편적인 정확도 평가지표로 사용하고 있다(João, 1998). RMSE의 계산식은 다음과 같다(FGDC, 1998).

본 연구에서 RMSE를 구하는데 사용한 참조값은 우리나라 국가기본도를 제작할 때 적용하는 국가기준점(수치지형도 코드 731)과 항측기준점(수치지형도 코드 732)을 이용하였다.

2. DEM 생성 내삽 알고리즘 평가

자료점이 분포하는 지역내에서 조사되지 않은 지점의 변수값을 추정하는 내삽법(interpolation)은 크게 관측점의 자료가 갖는 영향권을 대표할 수 있는 이항 함수를 설정하고 그 함수의 변수를 이용하여 자료면을 추정하는 '적합함수접근법(fitted function interpolation)'과 어떤 범위에 속하는 관측점들의 영향치를 직접 계산하여 특정 지점의 자료값을 얻는 '가중평균법(weighted average

interpolation)'으로 분류할 수 있다(Watson, 1994). 그리고 가중평균법은 거리에 기초한 가중평균법(distance-based)과 면적에 기초한 가중평균법(area-based)으로 다시 구분할 수 있다. 본 연구에서는 이러한 범주로 분류되는 내삽법 가운데 가장 일반적으로 적용되는 역거리가중법(inverse distance weighting method), 크리깅법(Kriging method), TIN 방법, 운형법(spline method)을 적용하여 DEM을 개발하였다. 평지, 계곡, 산지로 구분하여 선정된 표본지역에 대해 4 가지 내삽법을 적용하여 개발된 DEM은 실제 지상 해상력을 기준으로 20 미터의 격자 크기를 갖는다.

1) 내삽 알고리즘

① 역거리가중 내삽법(IDW, inverse distance weighting)

서로 가까이 위치한 점들이 멀리 떨어져 있는 지점보다 더 유사한 경향이 있다는 톨블러의 법칙(Tobler's law of geography)을 이용하여 관측점과 내삽지점 사이에 가중치를 주는 방법이다. 가중치로는 일반적으로 역거리제곱을 사용하며, 평균을 위해 윈도우를 사용하게 된다. 내삽지점 주위의 약 10여개의 서로 이웃하고 있는 관측점을 기초로 내삽과정을 거쳐 연속면을 작성한다. 작성된 연속면에서 일정간격의 격자의 고도값을 추정하여 수치고도모형을 생성한다.

② 크리깅 내삽법(Kriging)

적합함수에 의한 내삽접근법 가운데 대표적 방법인 크리깅은 지역화 변수이론에 기초하여 어떠한 변수의 공간적 변동을 전체경향의 구조적 성분, 무작위 공간적 자기상관, 노이즈 및 잔차항이라는 세 개의 주요한 성분 합으로 표현한 후, 분산도(semivariogram or variogram)를 이용하여 내삽에 필요한 최적 가중치를 결정한다. 결정된 최적 가중치를 이용하여 내삽과정을 거쳐 연속면을 작성하고, 작성된 연속면에서 일정간격의 고도값을 추정하여 수치고도모형을 생성한다. 크리깅은 임의의 규칙에 근거하여 가중치를 선택하는 것이 아니라 확률함수가 공간상에서 어떻게 변화하는가에 따라 선택하기 때문에 다른 방법들에 비해 유연하게 적용할 수 있다는 장점을 갖는다.

③ TIN 내삽법(triangle irregular network based)

TIN 내삽법은 면적에 기초한 가중내삽법의 일종으로 임의의 꼭지점(vertex)에서 시작하여 주변에 있는 두 개의 꼭지점을 이용하여 델로니 삼각형(Delaunay triangle)을 구성하고 삼각형 내부의 값은 평면상의 상대적인 위치로 결정하는 방법이다. 구성된 삼각형 망에 일정간격의 격자(grid)를 씌우고 삼각형과 중첩되는 격자의 고도값을 삼각형 평면상의 고도로 추정하여 수치고도모형을 생성한다.

④ 운형 내삽법(spline)

직사각형의 패치에 기초한 가중내삽법으로 분류되는 스피라인 내삽은 자료점 사이에 가중치를 설정하여 구해진 연속적인 구분 다항함수(piecewise polynomial function)를 통해 자료면을 구한다. 일반적으로 운형 내삽법에 의해 생성되는 자료는 평활화(smoothing)되는 효과가 있으며, 최대값과 최소값이 반드시 자료점과 일치하지 않는 결과가 생길 수도 있다.

2) 내삽 알고리즘 정확도 평가

내삽 알고리즘의 정확도를 비교하기 위하여 3 개 표본지역을 대상으로 역거리가중법(inverse distance weighting method), 크리깅법(Kriging method), TIN 방법, 운형법(spline method) 등 4 가지 내삽기법을 적용하여 총 12 개의 DEM을 작성하였다. DEM은 모두 20 미터의 지상 해상력을

갖고 있도록 개발하였고 그 레이어(커버리지)는 100 x 100 행렬의 격자로 구성되었다(그림 2, 3, 4).

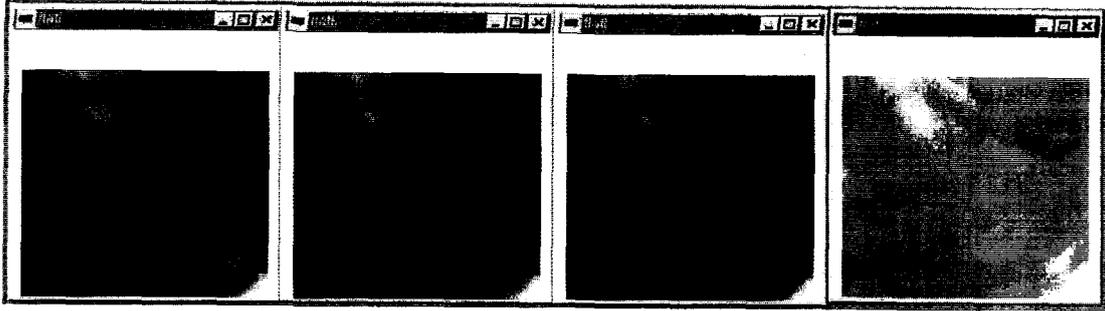


그림 2. 평지 표본지역에 대해 4 가지 내삽법을 적용한 DEM (IDW-Kriging-TIN-Spline)

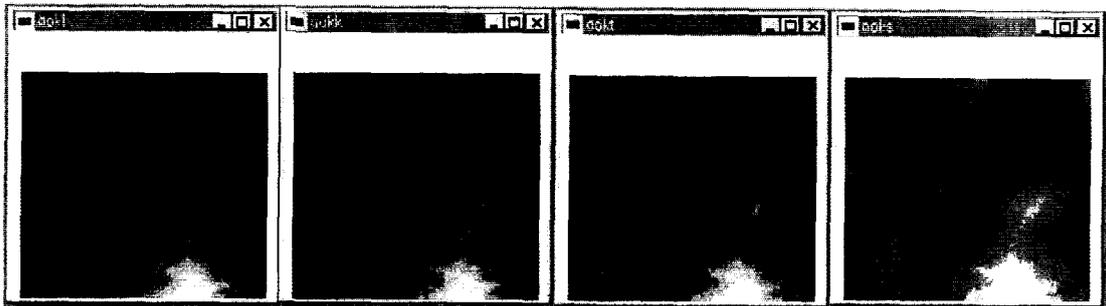


그림 3. 계곡 표본지역에 4 가지 내삽법을 적용한 DEM (IDW-Kriging-TIN-Spline)

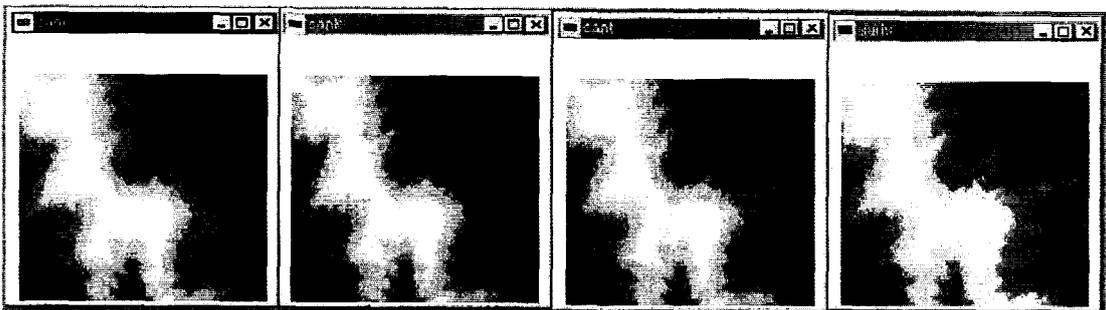


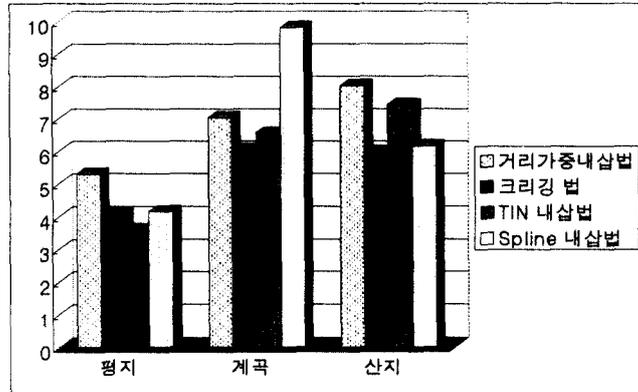
그림 4. 산지 표본지역에 4 가지 내삽법을 적용한 DEM (IDW-Kriging-TIN-Spline)

한편, 12 개의 DEM이 갖는 정확도를 비교하기 위해 각각의 표본지역에 속해 있는 표고점과 삼각점을 기준 참조자료로 하여 RMSE를 계산하였다(표 1). 표본 지역별로 내삽 알고리즘의 정확도를 살펴보면 공통적으로 자료값의 분산도가 낮은 평지에 비해 분산도가 높아짐에 따라 RMSE가 증가하는 경향을 띄고 있다. 평지표본에서는 TIN 방법이 그리고 계곡표본과 산지표본에서는 크리깅 방법이 가장 정확도가 높게 나타나고 있다. 특히 크리깅 방법은 평지표본의 경우에도 TIN 방법에 따른 정확도와 크게 차이를 두고 있지 않고 있다. 그것은 앞서 밝힌 바와 같이 크리깅의 특성상 공간 상에서 확률함수가 갖는 분포특성에 따라 가중치가 결정됨에 따라 그만큼 지형적 특성에 유연하게 대처하고 있다고 평가할 수 있다. 그리고 운형 내삽법의 경우는 계곡표본에서 다른 방법에 비해

정확도가 크게 떨어지는 특징을 보이고 있다.

표 1. 표본지역별 내삽 알고리즘의 정확도 비교

표본지역	기준 참조 표본수	역거리가중	크리깅	TIN	운형
평 지	24	5.293	4.112	3.617	4.148
계 곡	9	7.063	6.077	6.583	9.827
산 지	32	8.041	6.009	7.444	6.186



3. 공간적 표본추출에 기초한 내삽법을 적용한 DEM

DEM에서 필요한 모든 지점의 고도를 원자료에서 모두 측정하는 방법은 비용이나 시간적인 측면에서 비효율적 요소가 있다. 따라서, 원자료로부터 적당한 표본지점들의 고도를 직접 측정하고, 그 외 지점들의 고도는 이들 표본 고도점들로부터 내삽에 의해 구하는 방법이 일반적이다. 본 연구에서 사용한 수치지형도에서 하나의 등고선 자료는 실제 동일한 고도값을 갖는 점(vertex)들의 집합으로 표현되고 있으며, 대개의 경우 수작업 디지털라이징 혹은 해석도화에 의해 취득되며 따라서 무작위 분포 특성을 갖는다. 여기에서는 한 지역의 DEM을 개발하고자 하는 경우 등고선을 구성하는 점들의 모집단 전체에 대해 내삽을 수행하는 방법과 공간적 표본추출에 의해 구성된 표본 점(sampled vertex)에 의한 내삽을 적용한 방법을 비교하여 효과적 DEM 개발의 방안을 제시하고자 한다.

이를 위해 본 연구에서 사용한 벡터형(vector type)의 수치지형도의 등고선 자료를 우선 1 미터 격자 크기로 래스터화(rasterization)한 다음 이를 기초로 x, y 양방향에서 1 미터 간격의 벡터형 자료로 재구성하였다. 재구성된 자료는 각각의 공간적 표본추출법에서 모집단 성격을 갖고 이로부터 공간적 해상력을 달리한 표본에 따라 내삽이 이루어진다. 한편, 이와 같은 공간적 표본추출을 이용한 내삽의 효과와 효율성을 평가하는데는 앞서의 내삽 알고리즘 평가에서 정확도가 가장 우수한 크리깅을 이용하였다.

1) 공간적 표본 추출법

본 연구에서 적용한 공간적 표본 추출법은 '체계적 표본추출법(systematic sampling)'과 '군집형 무작위 표본추출법(cluster random sampling)' 그리고 '복합 점진적 표본추출법(hybrid progressive sampling)'으로 구분할 수 있다. 첫번째 '체계적 표본추출'은 가장 간편한 방법으로서 균등하게 표본이 추출되는 특성을 갖는다. 여기에서는 1 미터 간격의 벡터형 자료로 재구성된 모집단 자료로부터

거리의 단위를 각각 5 미터, 10 미터, 20 미터, 30 미터, 40 미터, 50 미터로 설정하여 표본을 추출한 다음 이를 이용해 크리깅 내삽을 실시하였다. 두 번째 '군집형 무작위 추출법'은 일정한 군집(cluster) 지역을 설정하여 그 영역내에서 표본을 무작위로 추출하는 방법으로 본 연구에서는 그 군집 지역의 크기를 앞서 체계적 표본추출에서의와 같은 5 미터, 10 미터, 20 미터, 30 미터, 40 미터, 50 미터로 각각 설정하고 그 표본수 역시 동일하게 적용하였다. 마찬가지로 이러한 표본에 따라 크리깅 내삽을 수행하여 DEM을 개발하였다.

한편, 체계적 추출과 군집형 무작위 추출은 정도의 차이는 있지만 규칙적인 표본격자를 이용한 표본추출이라고 할 수 있다. 이런 방식은 일반적으로 기복량이 적은 지역에서 필요 이상으로 많은 자료점들이 추출될 수 있고 기복의 변화가 큰 지역에서는 표본수가 너무 적을 수도 있기 때문에 지표기복의 변화를 충실히 반영하는 데에는 비효율적이라는 지적을 받는다. 점진적 표본추출법은 지표기복의 특성을 최대한 반영하기 위한 방법으로 Makarovic(1973)의 연구 이후 변형된 방법이 상당수 소개되고 있다. 본 연구에서도 변형된 점진적 추출법을 제시하고 이를 이용해 내삽한 DEM을 개발하였다.

세 번째 '복합 점진적 추출법'은 표본을 일정크기의 격자로 나눈 후 각 격자의 범위(최대값-최소값)나 분산도를 검사하여 표본을 추출하는 방식이다. 분산도가 일정 한도 내에 위치하면, 그 격자의 중심에 가까이 위치한 점을 표본으로 설정하고, 분산도가 기준값보다 크면 그 격자를 다시 하위 격자로 나누고 분산도를 검사하게 된다. 이 과정은 모든 격자들의 분산도가 기준값을 만족할 때까지 반복된다. 지표의 기복이 심한 지역은 많은 양의 표본이 추출되고 기복이 적은 지역은 상대적으로 적은 양의 표본이 추출된다. 이 표본추출법은 고도값의 공간적 분포를 고려하여 표본양을 조절하므로 지형의 특성을 반영하는 효과적인 방식이라 할 수 있다. 본 연구에서는 전체지역을 5*5 크기의 격자로 나누고, 각 격자의 분산도를 검사하여, 분산도가 0 보다 크면 그 격자를 다시 5*5 개의 하위격자로 나누어 점진적으로 표본추출을 행하였다.

2) 실험 결과의 비교와 평가

'체계적 표본추출법'이나 '군집형 무작위 표본추출법'을 적용해 개발한 DEM의 정확도를 평가한 결과 매우 특이한 경향과 현상을 공통적으로 발견할 수 있다(표 3, 4). 그것은 일반적 예상과 달리 표본의 밀도가 가장 높은 경우보다 오히려 구체적으로 명확히 규정하기 힘들지만 어느 한계(threshold)공간적 거리를 어느 정도 둔 표본을 기준으로 개발한 DEM의 정확도가 오히려 높다는 점이다. 이것은 DEM을 개발하기 위하여 전체 등고선을 구성하는 점(vertex)들을 전부 사용할 필요성이 없다는 가능성을 제기하고 있다. 특히 RMSE라는 절대치를 기준으로 비교할 경우, 앞장에서 전체 모집단 규모를 적용하여 내삽한 여러 결과에 비해 표본의 공간적 밀도가 낮은 30, 40, 50 미터를 적용한 두 가지 표본추출법에 의한 내삽 결과와 크게 차이를 두고 있지 않는다.

지형적 특성을 고려해 선정한 표본지역별로 그 결과를 살펴보면, 평지의 경우 두 가지 표본추출법의 정확도가 거의 차이가 없으며 또한 공통적으로 표본의 밀도가 감소(즉 표본의 공간적 단위거리가 증가)하더라도 뚜렷하게 정확도가 손상되지 않는다. 계곡의 경우 체계적 추출법과 군집형 무작위 추출법에서 각각 30 미터와 20 미터를 기준으로 상대적으로 뚜렷한 정확도 변화가 나타났다. 산지의 경우는 전반적으로 체계적 추출법이 군집형 무작위 추출법에 비해 정확도가 높은 특성을 보였다. 따라서 전반적으로 체계적 표본추출과 군집형 무작위 표본추출을 비교한 결과 공간상에 균등한 분포 패턴을 갖는 체계적 표본추출을 적용하여 내삽을 실시하는 방법을 DEM의 정확도란 측면에서 우수하게 평가할 수 있다.

마지막으로 지표기복의 변화를 충실히 반영하기 위해 제안한 '복합 점진적 표본추출법'을 적용한 DEM의 경우는 그 정확도가 다른 두 가지 추출법 뿐 만 아니라 전체 모수에 의한 방법보다도

오히려 정확한 결과를 보이고 있다. 다만, 이 방법은 다른 방법에 비해 표본추출에 필요한 분산도의 측정과 이를 기준으로 반복(iteration) 계산이 필요하기 때문에 상대적으로 복잡하다.

표 2. 체계적 표본추출법을 적용한 DEM의 RMSE 변화

표본 간격	1 미터		5 미터		10 미터		20 미터		30 미터		40 미터		50 미터	
	표본 갯수	RM SE												
평지	35,604	4.709	9,667	4.632	4,977	4.492	2,458	4.083	1,496	3.879	1,061	4.276	784	4.311
계곡	105,582	6.445	28,193	6.181	14,362	5.862	6,349	6.219	3,372	7.697	2,093	7.697	1,405	9.580
산지	155,804	6.549	40,580	6.419	20,568	6.079	8,686	5.865	4,215	6.948	2,481	6.948	1,597	7.261

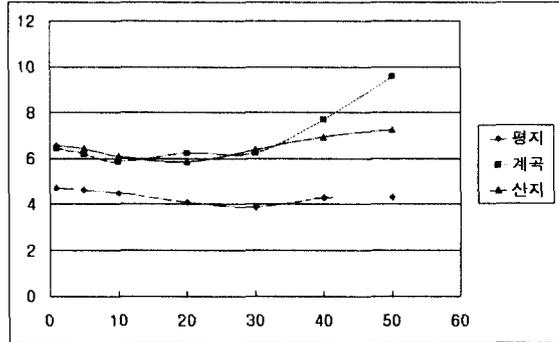


표 3. 군집형 무작위 표본추출법을 적용한 DEM의 RMSE 변화

표본 크기 기준	1 미터		5 미터		10 미터		20 미터		30 미터		40 미터		50 미터	
	표본 갯수	RM SE												
평지	35,604	4.709	9,667	4.609	4,977	4.538	2,458	4.321	1,496	3.912	1,061	4.217	784	4.922
계곡	105,582	6.445	28,193	6.559	14,362	5.885	6,349	5.728	3,372	7.402	2,093	8.436	1,405	9.560
산지	155,804	6.549	40,580	6.414	20,568	6.167	8,686	6.337	4,215	7.162	2,481	8.200	1,597	10.906

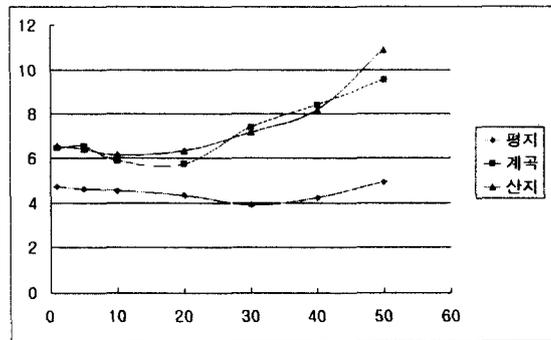


표 4. 복합 점진적 표본추출법을 적용한 DEM의 정확도

표본 지역 구분	RMSE	표본 갯수
평지	3.579	1866
계곡	5.978	10847
산지	6.027	17834

4. 결론

본 연구는 우리 나라 수치지형도의 등고선 자료를 대상으로 하고 있지만 실제 고도값 자료의 유형이 점(vertex) 형태로 분포하는 경우에 일반적으로 적용할 수 있다. 본 연구를 통해 밝혀낸 연구 결과와 시사점을 간략하게 요약하면 다음과 같다.

첫째, 지형의 기복을 달리하여 구분한 표본지역에 대한 4가지 DEM 생성 내삽 알고리즘의 정확도를 비교한 결과 크리깅 내삽을 제외한 나머지 내삽법들은 표본지역들 사이에서 RMSE의 차이가 크게 나타나는 특징을 보였다. 이를 통해 크리깅 내삽을 지형기복의 변화에 상관없이 정확도를 유지하면서 안정적으로 DEM을 생성할 수 있는 방법으로 평가할 수 있다.

둘째, DEM 개발에 내삽 알고리즘을 적용할 때 공간적 표본 밀도를 달리한 실험을 통해서 내삽에 동원되는 원자료의 규모에 관한 흥미로운 사실을 발견하였다. 즉 과도한 밀도를 갖는 표본을 적용한 내삽보다 어느 한계영역 내에서는 오히려 공간적 표본 밀도를 낮게 적용할 때 DEM의 정확도가 높게 나타났다. 또한 본 연구에서 실험한 공간적 표본 거리, 즉 5 미터에서 50 미터까지의 표본 자료에서 한계영역을 벗어난 경우에도 지나치게 정확도가 떨어지는 현상은 전반적으로 발견할 수 없었다. 이러한 실험적 연구는 DEM의 개발을 위한 내삽과정에서 표본의 크기를 원자료 전체보다는 추출된 표본으로 이용하는 방안이 정확도와 계산상의 효율성 측면에서 효과적일 가능성을 시사하고 있다.

셋째, 공간적 표본추출에 의한 내삽의 정확성을 비교한 결과 지표기복의 변화를 충실히 반영하는 '복합 점진적 표본추출법'이 계산의 단계가 복잡하지만 가장 우수한 결과를 나타냈다. 그러나 이 방법은 점진적 단계(progressive steps)를 어느 기준에 따라 설정하여야 하는가에 대한 지표를 찾아 내기가 쉽지 않다.

참고문헌

- Burrough, P.A., 1986, *Principles of Geographical information systems for Land resources assessment*, Clarendon: Oxford.
- Federal Geographic Data Committee, 1998, *Geospatial Positioning Accuracy Standards: Part3: National Standard for Spatial Data Accuracy*, Subcommittee for Base Cartographic Data
- João, E.M., 1998, *Causes and Consequences of Map Generalisation*, Taylor & Francis: Bristol.
- Maling, D.H., 1989, *Measurements from Maps: Principles and Methods of Cartometry*, Pergamon Press: Oxford.
- Wood, J.D., 1998, "Modelling the continuity of surface form using Digital Elevation Models," *Proceedings 8th International Symposium on Spatial Data Handling*, pp.725-36.