크리깅 기법을 이용한 단양군의 산림 탄소저장량 추정* - 지상부 바이오매스를 대상으로 -

박현주^{1*} · 신휴석² · 노영희³ · 김경민⁴ · 박기호³

Estimating Forest Carbon Stocks in Danyang Using Kriging Methods for Aboveground Biomass*

Hyun-Ju PARK^{1*} · Hyu-Seok SHIN² · Young-Hee ROH³
Kyoung-Min KIM⁴ · Key-Ho PARK³

요 약

본 연구에서는 크리깅 기법 중 가장 일반적으로 사용되고 있는 정규 크리깅(ordinary kriging)과 여러 보조 자료를 활용한 회귀 모델에 크리깅 기법을 결합한 형태인 회귀 크리깅(regression kriging)을 이용하여 지상부 바이오매스의 탄소저장량 추정을 시도하였다. 분석 결과 단양군의 산림 탄소저장량은 정규 크리깅의 경우 3,459,902 tonC, 회귀 크리깅의 경우 3,384,581 tonC로 추정되었으며 회귀 크리깅에 이용된 회귀 모델의 결정계수는 0.1033으로 나타났다. 표본점을 임상별로(활엽수림, 침엽수림&혼효림) 층화하여 회귀 크리깅을 수행한 경우에는 탄소저장량은 3,336,206 tonC로 가장 낮게 추정되었으며 회귀 모델의 결정 계수는 각각 0.35, 0.18로 표본점 전체를 이용했을 때보다 높은 값을 보였다. 각 기법의 교차 검증(cross validation) 결과 표본점 전체를 이용한 회귀 크리깅의 RMSE(22.32 ton/ha)가 가장 낮았으나 기법간의 차이(0.23 ton/ha)는 크지 않은 것으로 나타났다.

주요어 : 산림, 지상부 바이오매스, 탄소저장량, 회귀 크리깅, 보조 자료, 단양

2011년 11월 17일 접수 Received on November 17, 2011 / 2011년 12월 16일 수정 Revised on December 16, 2011 / 2012년 3월 13일 심사완료 Accepted on March 13, 2012

* 이 연구는 2010년도 국립산림과학원에서 시행한 기후변화 대응 산림정책 및 기술개발 사업에 의하여 수행되었음. 이 논문은 한국지리정보학회 2010년 11월 추계학술대회(신휴석 외, 2010)에서 발표된 내용을 수정 보완한 것임.

1 한국환경정책·평가연구원 환경전략연구본부 Environmental Strategy Research Group, Korea Environment Institute

2 서울대학교 국토문제연구소 Institute for Korean Regional Studies, Seoul National University

3 서울대학교 지리학과 Department of Geography, Seoul National University

4 국립산림과학원 산림자원정보과 Division of Forest Resources Information, Korea Forest Research Institute

* 연락처 E-mail : hjpark@kei.re.kr

ABSTRACT

The aim of this study is to estimate aboveground biomass carbon stocks using ordinary kriging(OK) which is the most commonly used type of kriging and regression kriging(RK) that combines a regression of the auxiliary variables with simple kriging. The analysis results shows that the forest carbon stock in Danyang is estimated at 3,459,902 tonC with OK and 3,384,581 tonC with RK in which the R-square value of the regression model is 0.1033. The result of RK conducted with sample plots stratified by forest type(deciduous, conifer and mixed) shows the lowest estimated value of 3,336,206 tonC and R-square value(0.35 and 0.18 respectively) is higher than that of when all sample plots used. The result of leave-one-out cross validation of each method indicates that RK with all sample plots reached the smallest root mean square error(RMSE) value(22.32 ton/ha) but the difference between the methods(0.23 ton/ha) is not significant.

KEYWORDS : Forest, Aboveground Biomass, Carbon Stock, Regression Kriging, Auxiliary Variables, Danyang

서론

기후변화협약(United Nations Framework Convention on Climate Change, UNFCCC) 및 교토의정서(Kyoto Protocol)의 발효에 따라 탄소배출권에 대한 관심이 증대되면서 탄소의 주요 흡수원인 산림 바이오매스가 주목을 받고 있다. 기후변화에 관한 정부간 패널(Intergovernmental Panel On Climate Change, IPCC)에서는 온실가스 통계 작성에 있어 일정한 지침을 제시하고 있고, 통계 작성 수준을 세 등급(Tier 1~Tier 3)으로 구분하고 있는데 현재 우리나라의 산림 바이오매스 추정 방식은 방법론 측면에서 국가산림조사(National Forest Inventory, NFI) 자료와 국내 산림통계자료를 이용한 Tier 2 수준의 추정이라 평가되고 있다.

하지만 현재의 시·군·구 단위로 집계되는 자료를 통해서는 시·군·구 단위 내에서의 국지적인 공간 분포를 파악하기는 어렵다. 또한 국제적으로 인정받는 우수 지침을 따르기 위해 세분화된 전환계수의 적용, 체계적으로

수집된 공간자료의 활용, 다양한 방법론을 활용한 프로세스 기반 모델의 정의 등 Tier 3 수준에서 요구하는 추정 방법으로 이전되어야 할 필요도 점차 증가하고 있다.

해외에서는 다양한 공간 자료를 통합하여 산림 바이오매스와 관련된 변수 예측의 정확도를 높이고 불확실성 분포를 파악할 수 있는 geostatistics를 활용한 연구가 활발히 진행되고 있다. 미국의 Meng *et al.*(2009)은 비공간적 모델링(nonspatial modeling)과 공간적 모델링(spatial modeling)을 결합한 산림 변수 예측을 위해 현지조사 자료뿐만 아니라 GPS, Landsat ETM+, GIS 자료를 통합하는 체계적인 접근 프로세스를 정의하였다. 또한 교차 검증(cross validation)을 통해 정규 크리깅(ordinary kriging), 공동 크리깅(co-kriging), 보편 크리깅(universal kriging), 회귀 크리깅(regression kriging) 등 여러 크리깅 기법을 비교하였는데, 그 중 회귀 크리깅이 소나무의 흉고단면적 예측에 가장 적합하다는 결론을 내었다. Sales *et al.*(2007)은 최근 아마존의 산림 파괴가 점진적으로 진행되고 있는 가운데 산림 바이오매스의 공간 분포와 변화를 더욱 정확하게 추정하는 것이 중

요하다는 관점에서 geostatistics를 활용한 추정 방법의 효과적인 접근방법을 제안한 바 있다. 또한 잔차 지도를 통해 추정값의 신뢰도를 평가할 수도 있고 향후 자료 수집의 우선순위를 정하거나 자료 수집을 위해 다른 수단을 적용할 것인지를 결정할 수 있다고 보았다.

국내에서도 최근 들어 신뢰성 있고 공간적으로 명확한(spatially explicit) 산림 바이오매스 및 탄소저장량 추정을 위한 다양한 노력이 이루어지고 있으며, 추정 방법으로는 회귀모형을 이용하거나(이승호 등, 2004; 장안진과 김형태, 2008; 정상영 등, 2009) 최근된 기법(k-Nearest neighbor, k-NN)을 이용한 추정(임종수 등, 2007; 2009)이 주로 이루어지고 있다. 하지만 국내에서는 해외에서 다양하게 시도되고 있는 geostatistics 기법을 이용한 바이오매스 추정은 거의 이루어지지 않고 있다. 일반적으로 geostatistics 기법은 확률 함수 개념에 기반을 두고 있고 불확실성(uncertainty) 분석의 틀 또한 제공하여(Goovaerts, 1997) 공간 예측과 평가에 있어 다양한 가능성을 내포하고 있기 때문에, 이를 이용한 탄소저장량 추정을 시도하고 검증해 볼 필요가 있다.

따라서 본 연구에서는 geostatistics 기법의 일부인 정규 크리깅과 회귀 크리깅을 이용하여 여러 보조 자료를 결합한 산림 탄소저장량 추정을 시도하였다. 공간적인 이질성이 큰 우리나라의 산림 조건을 고려했을 때 4km×4km 격자 상에 위치한 국가산림조사 표본점의 바이오매스 측정치를 이용하여 전역적인 분포를 추정하는 데에는 한계가 있기 때문에, 다양한 보조 자료를 활용한 경향성 추정과 크리깅이 결합되어 미조사지점 추정의 정확도를 높일 수 있는 회귀 크리깅에 초점을 맞추어 결과를 고찰해보고자 한다.

연구지역 및 사용자료

1. 연구지역

연구지역은 현재 1:5,000 정밀입상도가 완성된 지역인 단양군으로 선정하였다. 단양군

은 충청북도의 동북단에 위치하며, 동남쪽으로는 경상북도 영주시, 북쪽으로는 강원도 영월군, 남쪽으로는 경상북도 문경시, 서쪽으로는 제천시와 각각 경계하고 있는 3도 접경의 도 경계를 형성하고 있다. 범위는 동경 128° 39'에서 동경 128° 13'(연장 38.3 km), 북위 36° 47'에서 북위 37° 09'(연장 39.9 km)에 해당한다(그림 1).

단양군의 지형적 특성은 83.7%가 산악지대이고 경지면적은 11.2%에 불과하여 대부분 산악으로 이루어져 험준한 산세를 형성하고 있다. 주 수계는 남한강 상류로서 본 지역을 관통하여 충주호를 형성한다.

단양군의 기후적 특성은 내륙산간지대에 위치하여 기온의 교차가 심한 내륙성 기후를 나타낸다. 연평균 기온은 12° C, 평균최고 기온은 18.4° C, 평균최저 기온은 6.7° C, 연간 총 강우량은 882.5 mm이다(단양군, 2009).

2. 사용자료

1) 국가산림조사 자료

국가산림조사에서 얻어지는 산림통계 중에서 임목측적은 바이오매스 및 탄소저장량을 산출하는데 사용되는 기본 자료라는 점에서 매우 중요하다. 본 연구에서는 임목조사 자료에서 계산된 임목측적을 기반으로 표본점별 바이오매스를 구하고, 크리깅을 통해 미조사지점의 바이오매스 값을 추정한 후, 탄소전환계수를 적용하여 최종 탄소저장량을 산출하는 방법을 사용하였다. 표본점별 바이오매스를 계산할 때에는 탄소 저장고(carbon pool) 중에서도 지상부 바이오매스만을 대상으로 하였다.

시간적 범위는 그림 1과 같이 제5차 국가산림조사가 이루어지는 2006~2010년 중 2006~2009년의 자료를 사용하였다. 계통추출법에 의해 단양군에 배치된 표본점은 총 195개이지만 2010년에 조사 예정된 36개 표본점과 조사년도 미배치 28개 표본점, 그리고 기타 자료가 불충분한 표본점 11개 등을 제외하고, 총 120개의 표본점을 사용하였다. 표본

FIGURE 1. 단양군의 위치와 Landsat TM 위성영상 범위(좌측) 및 NFI 표본점 배치 현황(우측)

점의 형태는 전국을 $4\text{km} \times 4\text{km}$ 격자로 나눈 후 그 교차점에 4개의 표본점을 집락으로 배치하는 계통적 집락추출법을 채택하였으며(그림 2), 표본 추출 강도는 0.01%(전국 산림 640만 ha 중 640 ha를 조사), 각 표본점의 크기는 0.16 ha이다(국립산림과학원, 2009).

FIGURE 2. 집락표본점 구조 (국립산림과학원, 2009)

본 연구에서는 표본점의 지상부 바이오매스 계산을 위해 임목조사 항목의 수종, 수고, 흉고직경, 수령 등의 자료를 이용하였으며, 바이오매스의 더욱 정확한 추정을 위해 2010년 국립산림과학원에서 개발된 12개의 수종별 목재기본밀도(wood density, WD)와 바이오매스 확장계수(biomass expansion factor, BEF)를 적용하였다. 이렇게 수종별로 세분화된 계수를 적용하는 방식은 IPCC 2006 GL에서도 권장하고 있는 방식이다.

2) 임상도

임상도는 산림의 공간분포를 연속적으로 나타낸 주제도이며 본 연구에서는 단양군 산림 정밀지도 제작 산출물인 1:5,000 정밀임상도를 이용하였다. 이 임상도는 2007년 10월에 촬영된 디지털항공사진과 임상조사 및 표준지 조사를 기초로 제작되었다(단양군, 2010).

임상도의 속성 중 바이오매스 추정에 영향

을 미칠 수 있는 변수로 경급, 영급, 수관밀도를 이용하였다.

단양군의 총 탄소저장량을 계산하기 위해서는 단양군의 산림 지역에 해당하는 면적을 알아야 한다. 2009년 산림기본통계에 의하면 단양군의 산림 면적은 62,745ha로 80.38%의 산림율을 보였으며 임목축적은 6,883,958m³, 평균임목축적은 109.71m³/ha였다(산림청, 2010). 하지만 본 연구에서는 산림기본통계상의 면적을 이용하지 않고 단양군 정밀임상도를 셀 크기 30m×30m의 래스터로 변환한 후 산림면적을 별도로 계산하였다. 산림의 정의는 ‘산림자원의 조성 및 관리에 관한 법률’의 규정에 따라 임목지와 무림목지를 포함하며, 이에 따른 단양군의 산림 면적은 총 64,932 ha였다. 따라서 이 면적을 기준으로 단양군의 바이오매스와 탄소저장량을 산출하였다.

3) 위성영상

본 연구에서는 대면적 산림자원조사에서 널리 이용되는 중해상도 위성인 Landsat-5 TM 영상을 이용하였다(그림 1). 이 영상이 취득된 시기는 2006년 3월 14일이며 공간해상도가 30m로 대면적 산림자원조사의 야외표본점 크기와 유사하다. 영상의 밴드 중에서 공간해상도가 120m인 6번 밴드를 제외하고, 공간해상도가 30m인 나머지 6개 밴드의 분광값을 분석에 이용하였다. 위성영상은 1:25,000 수치지형도를 이용하여 기하보정 되었다.

위성영상의 활용은 임종수 등(2009)의 연구에서 산림 바이오매스 추정 회귀모델에 사용된 7개의 밴드울 지수를 참고하였다. 7개의 밴드울 지수는 각각 비율식생지수(Ratio Vegetation Index, RVI), 수분스트레스지수(Moisture Stress Index, MSI), 식생상태지수(Vegetation Condition Index, VCI), 정규식생지수(Normalized Difference Vegetation Index, NDVI), 습윤지수(Normalized Difference Moisture Index, NDMI), MIRV1과 MIRV2(Hosseini *et al.*, 2004)이며, 각각의 계산식은 표 1에서 설명하고 있다. 래스터

연산은 ArcGIS® Spatial Analyst를 이용하여 이루어졌다.

4) DEM 자료

DEM 자료는 30m×30m의 셀 크기를 가지며 원자료의 고도값뿐만 아니라 DEM으로부터 추출한 지형 변수를 추가적으로 포함하였다. 지형변동성지수(Vector Ruggedness Measure, VRM)는 복잡한 경사 변화를 보이는 지형특성을 알 수 있게 해 주는 지수이며(Sappington *et al.*, 2007), 지형습윤지수(Terrain Wetness Index, TWI)는 한 지점에 물이 흘러 들어가는 경향을 추정하여 지형의 습윤 정도를 측정해 주는 지수이다(Sørensen *et al.*, 2006).

분석 방법

본 연구에서는 Geostatistics를 이용한 탄소저장량 추정을 위해 정규 크리깅과 회귀 크리깅의 두 가지 방법을 이용하여 그 결과를 비교하였다. 회귀 크리깅시 이용되는 회귀분석과 크리깅 분석 작업은 오픈소스 통계 패키지인 R x64 2.11.1 버전(<http://cran.r-project.org>)을 사용하였으며, 구체적인 절차는 다음과 같다.

1. 회귀모델링

바이오매스의 추정에 보조 자료가 미칠 수 있는 영향을 평가하기 위해 회귀모델링을 시도하였다. 적합한 회귀모델을 이용하여 경향성(trend)을 제거한 후 그 잔차를 이용한 베리오그램 모델링과 크리깅이 수행되었다.

1) 회귀모델링에 사용된 변수

일반적으로 모형 구축의 과정은 간단한 모형으로부터 시작하여 필요에 의해 점차 복잡한 모형으로 발전하게 된다. 잔차 구조에 이상이 없는 한 간명성(parsimony)의 원칙에 의해 가장 간단한 모형을 선택하는 것이 일반

TABLE 1. 회귀 모델링에 사용된 변수

변수 형태	자료 형태	사용 변수	비고
종속변수	NFI 표본점	바이오매스 (Biomass)	단위: ton/ha
		영급 (AGE)	1-10
		경급 (DBH)	0-3
설명변수	임상도	수관밀도 (DEN)	1-3
		고도 (DEM)	단위: m
		지형변동성 (VRM)	
	DEM	지형습윤지수 (TWI)	
		비율식생지수 (RVI)	TM4/TM3
		수분스트레스지수 (MSI)	TM5/TM4
	Landsat-5 TM 영상	식생상태지수 (VCI)	TM7/TM4
		정규식생지수 (NDVI)	(TM4 - TM3)/(TM4 + TM3)
		습윤지수 (NDMI)	(TM4 - TM5)/(TM4 + TM5)
		MIRV1	(TM7 - TM3)/(TM7 + TM3)
		MIRV2	(TM5 - TM3)/(TM5 + TM3)

적이다.

회귀분석에 사용 가능한 모든 변수들을 정리하면 표 1과 같다. 이들 변수 중에서 바이오매스 추정에 유의미한 변수를 선택한 후 회귀모델을 적합하게 된다.

2. 회귀모형 적합 및 최종 모델 선정

바이오매스 추정에 유의한 변수 선택 방법은 단계적 선택법 (stepwise selection method)을 택하였다. 단계적 선택법은 전진 선택법 (forward selection method)을 보완한 방법으로 전진 선택법의 각 단계마다 선택된 변수가 모형에 추가됨으로써 제거시킬 변수가 존재하는지를 확인하는 과정을 포함하고 있다. 모형 선택 기준으로는 AIC(Akaike Information Criteion) 통계량을 이용하였다.

2. 베리오그램

바이오매스 또는 회귀식을 통해 경향성을 제거한 바이오매스 잔차에 대해 공간적 다양성 (spatial variability) 또는 공간적 의존성 (spatial dependancy)을 반영하기 위해 베리오그램을 적합하였다. 베리오그램은 이격된 거리에 따른 자료들 간의 관계를 분석하여 연구지역의 공간적 자료 분포 구조를 파악한 후

미조사 지점의 자료값을 추정하는데 활용되며, 본 연구에서는 구형 모형 (spherical model) 과 지수형 모형 (exponential model)을 이용하였다.

3. 크리깅

크리깅은 국지적인 보간 방법 중의 하나로 미조사지점의 예측값을 주변의 관측지점의 값의 가중선형조합으로 산출하는 대표적인 Geostatistics 방법 중의 하나이다. 크리깅에는 다양한 종류가 있는데 이중 정규 크리깅은 환경과학 분야에서 가장 많이 사용되고 있으며 (Webster and Oliver, 2007) 단순 크리깅의 단점인 추정식이 편향이 된다는 점을 극복하기 위한 방법이다 (최종근, 2007).

크리깅을 수행할 때에는 일반적으로 적용되는 가정들이 있다. 첫 번째는 대상 변수가 정규분포를 이루어야 한다는 것이다. 정규분포를 만족하지 않을 경우에는 예측변수와 종속변수의 선형관계를 만족시키기 위해 변수 변환이 이루어지기도 하며, 크리깅을 통해 예측된 값은 해석을 위해 쉽게 역변환 될 수 있다. 하지만 분산의 추정치는 역변환하기 어려운 단점이 있다 (Hengl *et al.*, 2004). 두 번째 가정은 자료가 공간적으로 불변성

(stationarity)을 유지해야 한다는 것이다. 자료에 경향성이 존재하면 크리깅 수행에 필수적인 공간적 의존성에 모형설정오류(misspecification)가 발생하여 추정치에 편(bias)을 가질 수 있다(Karl, 2010).

따라서 크리깅을 수행하기 전에 경향성을 제거하는 여러 방법들이 제안되었다. 먼저, 자료에 공간적인 경향성이 없을 때 사용하는 크리깅 방법으로 단순 크리깅(simple kriging)과 앞에서 살펴본 정규 크리깅이 있다. 단순 크리깅은 자료의 평균이 알려져 있을 때, 정규 크리깅은 자료의 평균이 알려져 있지 않을 때 사용한다(Bailey and Gatrell, 1995).

다음으로 공간 예측과 경향성의 추정을 함께 수행하는 방법으로 보편 크리깅이 있다. 경향성을 포함한다는 이유로 보편 크리깅의 정의를 확대하여 사용하는 경우가 많았으나, 최근에는 여러 학자들이 보편 크리깅의 정의에 대해 합의점을 찾은 것으로 보인다. 즉, 경향성을 모델링할 때 다른 보조 자료의 이용 없이 샘플링 지점의 좌표 값만을 이용하는 경우를 보편 크리깅이라 할 수 있다(Hengl *et al.*).

만약 경향성을 모델링할 때 좌표 값만이 아닌 다른 변수를 활용한다면 이를 외부 추이를 이용한 크리깅(Kriging with external drift, KED)이라 한다. 외부 추이를 이용한 크리깅은 샘플 자료로부터 일반화최소제곱(Generalized Least Square, GLS)을 통해 생성된 회귀식을 이용하여 한꺼번에 공간 예측을 수행한다(Hengl *et al.*, 2004). 하지만 경향성과 잔차의 추정이 개별적으로 이루어진 다음 그 결과가 합쳐져서 공간 예측을 할 수도 있는데 이를 회귀 크리깅이라 한다. 이러한 방법은 Ahmed and de Marsily(1987)가 최초로 제안하였고, 이후 Odeh *et al.*(1994; 1995)이 토양 변수의 예측에서 이 방법을 사용하면서 회귀 크리깅이라고 부르기 시작하였다(Hengl, 2009). 이러한 회귀 크리깅은 외부 추이를 이용한 크리깅보다 안정적이고 다양한 회귀 접근법을 사용할 수 있다는 장점을

가진다(Hengl *et al.*, 2004).

이러한 회귀 접근법은 샘플링 위치에서의 관측값을 이용하여 보조 변수 사이의 관계를 모델링하여 미조사 지점의 값을 추정해 내는 방법이다. 환경적인 보조 변수에는 토지피복도, 원격탐사 영상, 토양도, 토지이용도 등이 있을 수 있다. 가장 일반적인 회귀 접근법 중 하나는 다중 선형 회귀(linear multiple regression) 방법이 있으며(Hengl *et al.*, 2007) 이 방법 또한 가중 평균값을 이용하여 예측값을 결정한다.

$$\hat{z}(s_0) = \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0); \quad q_0(s_0) \equiv 1 \quad (1)$$

식 (1)에서 $q_k(s_0)$ 는 대상 위치에서의 보조 변수 값, $\hat{\beta}_k$ 는 추정된 회귀계수, p 는 예측 변수 또는 보조 변수의 개수를 나타낸다(지리적인 좌표값과 혼동을 피하기 위하여 예측 변수를 x 대신 q 로 표기).

회귀 크리깅은 이러한 회귀 접근법과 크리깅을 결합한 것이라 할 수 있으며, 설명 가능한 변이를 회귀 기법으로 적합한 후 잔차(residual), 즉 설명되지 않은 변이에 대해서는 기대 값이 0인 단순 크리깅을 이용하여 내삽하는 방법이다(Hengl *et al.*, 2004). 회귀 크리깅은 크리깅 가중치를 결정할 때 보조 변수를 예측변수로 직접 사용하며, 보편 크리깅이나 외부 추이를 이용한 크리깅과 내삽 방법 측면에서는 수학적으로 동일하다고 할 수 있다(Hengl *et al.*, 2007).

$$\begin{aligned} \hat{z}(s_0) &= \hat{m}(s_0) + \hat{e}(s_0) \\ &= \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0) + \sum_{i=1}^n \lambda_i \cdot e(s_i) \end{aligned} \quad (2)$$

식 (2)에서 $\hat{m}(s_0)$ 은 적합된 추이(drift)를 나타내며, $\hat{e}(s_0)$ 은 내삽된 잔차이다. $\hat{\beta}_k$ 는 추정된 추이 모형의 계수($\hat{\beta}_k$ 는 추정된 절편), λ_i 는 잔차의 공간적 의존 구조를 반영한 크리깅 가중치, $e(s_i)$ 는 샘플링 위치 s_i 에서의 잔

차를 나타낸다. 회귀계수 $\hat{\beta}_k$ 는 최소제곱 (Ordinary Least Square, OLS) 혹은 일반화 최소제곱 (Generalized Least Square, GLS) 과 같은 방법을 이용하여 각 표본점으로부터 추정될 수 있으며, 개별적인 관측값 사이의 공간적 상관관계가 고려된다(Cressie, 1993).

이렇게 공간적 다양성을 모델링하고 지도화 하기 위해 추정 변수의 관측치에 의존하여 내삽하는 방법과 공간적으로 완전한 보조 정보와 추정 변수와의 회귀에 기반하여 내삽하는 방법, 이 두 가지의 개념적으로 다른 접근법을 결합하는 내삽 기법을 하이브리드(hybrid) 내삽 기법이라 부르기도 한다(Hengl, 2009).

4. 예측능력 비교

표본점만을 이용한 정규 크리깅과 여러 보조 자료를 결합한 회귀 크리깅의 모델의 예측 능력 비교 및 베리오그램 모형의 조합에 따른 비교를 위해 Leave-one-out 교차 검증을 통한 평균제곱근오차(RMSE)를 산출하였다. Leave-one-out 교차 검증 방법은 표본점 전체 n 개 가운데 한 표본점의 바이오매스를 예측하기 위해 나머지 $n-1$ 개의 표본점을 사

용하여 추정하는 방법이다. 이러한 방식으로 모든 표본점에 대한 예측치와 실측치의 차이를 이용하여 RMSE를 계산하였다.

분석 결과

1. 정규 크리깅 결과

NFI 표본점에서 관측된 바이오매스 값을 이용하여 베리오그램을 모델링하였으며 구형 모형과 지수형 모형을 이용하였다. 그림 3의 베리오그램 모형 비교에서 구형 모형을 이용한 베리오그램의 공간적 상관관계의 범위는 약 7285m였고 상대 너깃(relative nugget)은 39.0%였으며, 지수형 모형의 공간적 상관관계의 범위는 약 3,883m였고(효력 범위는 약 8,951m) 상대 너깃은 35.4%였다.

너깃 효과(nugget effect)는 값을 알고 있는 관측지점들 간의 거리보다 더 짧은 거리에서의 상관관계에 대한 정보가 부족하기 때문에 발생하는 효과이며, 짧은 거리 내에서 값의 변화가 급격할 때 이를 적절하게 설명하지 못하는 원인이 될 수 있다(Isaaks and Srivastava, 1989; 최중근, 2007). 단양군의

FIGURE 3. 바이오매스의 베리오그램: 구형 모형(좌측), 지수형 모형(우측)(단위: m)

FIGURE 4. 정규 크리깅을 이용한 바이오매스 예측치와 표준편차(구형 모형)

FIGURE 5. 베리오그램 모형에 따른 정규 크리깅 예측치의 차이, 구형 모형(좌측), 지수형 모형(중간), 모형간 예측치의 차이(구형-지수형)(우측)

산림 바이오매스 베리오그램에서 상대 너깃 효과가 비교적 크게 나타나는 이유는 표본점 간의 거리가 멀어 관측 지점의 값이 관측되지 않은 지점의 값을 대표하기 힘들기 때문으로 볼 수 있다.

그림 4는 앞에서 도출된 베리오그램 중 구형 모형을 이용하여 공간 구조를 파악한 후 정규 크리깅을 적용한 결과이다. 바이오매스의 추정치와 표준편차를 나타내며 표본점이 위치했던 지역에서 표준편차가 확연히 낮게 나타나는 것을 확인할 수 있다. 추정된 바이오매스의

최소값은 19.60ton/ha, 최대값은 204.85ton/ha, 평균은 106.57ton/ha를 나타냈으며, 이에 따른 단양군 산림면적 64,932ha에 해당하는 지상부 바이오매스의 총량은 6,919,803ton, 탄소전환계수 0.5를 적용한 탄소저장량은 3,459,902tonC로 추정되었다. 지수형 모형을 이용한 바이오매스의 추정치 역시 표본점이 위치했던 지역에서 표준편차가 낮게 나타나는 것을 확인할 수 있었으며, 추정된 바이오매스의 최소값은 20.37 ton/ha, 최대값은 206.10ton/ha, 평균은

105.74ton/ha를 나타냈고, 지상부 바이오매스의 총량은 6,865,910ton, 탄소저장량은 3,432,955tonC로 추정되었다.

이렇게 베리오그램 모형에 따른 탄소저장량의 차이는 26,947tonC로 나타났으며, 그림 5와 같이 구형과 지수형 모형의 추정치의 차이를 비교해 보면 표본점이 위치하지 않은 지역일수록 그 차이가 커지는 것을 확인할 수 있었다(붉은색과 푸른색 부분). 따라서 표본점이 위치하지 않은 지역은 추정 방법에 따라 추정치의 변동이 발생할 가능성이 커진다고 볼 수 있다.

2. 회귀 크리깅 결과

회귀 크리깅은 가장 먼저 표본점별 바이오매스와 보조 자료간의 회귀모델링이 수행된 다음, 회귀 잔차를 이용하여 베리오그램 모델링과 크리깅이 수행된다. 먼저 표본점 전체를 이용하여 회귀 크리깅을 수행하고, 다음으로 회귀모델의 설명력을 높이기 위해 표본점을 임상별로 층화하여 회귀 크리깅을 수행하였다.

1) 표본점 전체를 이용한 회귀 크리깅

표본점 전체 바이오매스에 대한 회귀모델링 결과는 표 2와 같다. 단계별 선택법으로 유의미한 변수는 DBH, MIRV1, TWI만 선택되었으며 수정된 결정계수는 0.1033이었다. 비교적 낮은 결정계수로 인해 보조 변수가 표본점 전체 바이오매스의 분포를 그리 잘 설명하지 못한다고 볼 수도 있지만, 복잡한 임분 구조

와 환경 조건으로 인해 회귀모델의 설명력이 낮게 나타나는 현상은 기존 연구에서도 밝혀진 바 있으며(Lu, 2006; Rahman *et al.*, 2008; 임종수 등, 2009), 기존의 국내 연구(임종수 등, 2009)에 비한다면 높은 편이라 할 수 있다. 하지만 DBH가 큰 임목일수록 바이오매스의 저장량이 높을 것으로 예상되었으나 실제 분석에서 음(-)의 부호가 나온 점 등은 추후 더욱 정교한 회귀모델링을 통해 보정해야 할 부분으로 보인다. 이는 4km×4km 간격의 표본점 배치에 비해 연구지역의 범위가 협소하고 표본점의 개수 또한 충분하지 않은 것에 기인한 것으로 판단된다.

위의 회귀모델을 이용하여 표본점별 바이오매스와 보조 자료와의 경향성을 반영하고 그 잔차를 이용하여 베리오그램 모형을 적합한 결과는 그림 6과 같다. 구형 모형을 이용한 베리오그램의 공간적 상관관계의 범위는 약 6,463m였고 상대 너깃(relative nugget)은 39.4%였으며, 지수형 모형의 공간적 상관관계의 범위는 약 2,882m였고(효력 범위는 약 7,287m) 상대 너깃은 37.1%였다. 정규 크리깅과 같이 너깃 효과가 비교적 큰 편이긴 하지만 표본점 전체를 이용한 회귀 잔차의 자기상관이 잘 드러난다고 할 수 있다. 이 베리오그램 모형을 통해 공간적 다양성을 반영한 회귀 크리깅 결과는 그림 7과 같다. 정규 크리깅에 비해 바이오매스 예측치나 표준편차의 분포에서 국지적인 변이를 더욱 상세하게 보여주는 것을 알 수 있다.

회귀 크리깅의 결과는 최대/최소값이 원자

TABLE 2. 표본점 전체 바이오매스에 대한 회귀모델링 결과

Predictor Variables	Coefficients	Std. Error	t Value	Pr(> t)
(Intercept)	187.524	30.131	6.224	9.24e-09 ***
DBH	-22.403	14.747	-1.519	0.1316
MIRV1	73.451	31.467	2.334	0.0214 *
TWI	-4.099	1.672	-2.452	0.0158 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 64.4 on 109 degrees of freedom

Multiple R-squared: 0.1273, Adjusted R-squared: 0.1033

F-statistic: 5.3 on 3 and 109 DF, p-value: 0.001895

FIGURE 6. 단양군 지역 표본점 전체 회귀 잔차 베리오그램: 구형 모형(좌측), 지수형 모형(우측)(단위: m)

FIGURE 7. 회귀 크리깅을 이용한 바이오매스 예측치와 표준편차 (표본점 전체, 구형 모형)

료의 최대/최소값과 비교했을 때 차이가 발생할 수 있는데, 과대/과소 추정치를 제거해 주기 위해 1.96 표준편차(95%)에 해당하는 값만을 추출하여 지도화하였다. 이렇게 추정된 바이오매스의 최소값은 29.93 ton/ha, 최대값은 178.56 ton/ha, 평균은 104.25 ton/ha를

나타냈으며, 이에 따른 단양군 산림면적 64,932 ha에 해당하는 지상부 바이오매스의 총량은 6,769,162 ton, 탄소전환계수 0.5를 적용한 탄소저장량은 3,384,581 tonC로 추정되었다.

TABLE 3. 활엽수림 바이오매스에 대한 회귀모델링 결과

Predictor Variables	Coefficients	Std. Error	t Value	Pr(> t)
(Intercept)	1666.41299	591.53861	2.817	0.006465 **
AGE	-29.15203	11.24926	-2.591	0.011863 *
DBH	-56.15458	20.34167	-2.761	0.007549 **
DEN	21.59333	11.58767	1.863	0.067058 .
DEM	0.08403	0.05204	1.615	0.111362
MIRV1	149.62088	38.00704	3.937	0.000209 ***
TWI	-3.97095	1.86533	-2.129	0.037185 *
VRM	-56.13494	23.36579	-2.402	0.019245 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 55.54 on 63 degrees of freedom

Multiple R-squared: 0.4162, Adjusted R-squared: 0.3513

F-statistic: 6.416 on 7 and 63 DF, p-value: 1.037e-05

2) 임상별 층화를 이용한 회귀 크리깅

다음으로 공간적으로 다양한 산림 바이오매스의 설명력을 높이기 위해 표본점을 침엽수림, 활엽수림, 혼효림으로 구분하여 회귀모델 적합을 시도하였다. 하지만 임상을 구분하였을 경우 표본점의 개수가 침엽수림, 활엽수림, 혼효림이 각각 24개, 75개, 18개로 매우 적어지므로 베리오그램 모델링에 적합하지 않고 바이오매스 예측치도 과대/과소 추정되는 현상이 있었다. 따라서 활엽수림 75개, 침엽수림과 혼효림을 합한 42개 표본점의 두 그룹으로 나누어 모델을 적합하였다. 이렇게 도출된 회귀식을 이용하여 임상별 바이오매스를 각각 추정한 다음, 임상별 면적에 대한 바이오매스를 합산하여 전체 바이오매스를 구하는 데 이용하였다.

활엽수림에 대한 회귀모델링 결과는 표 3과 같았다. AGE, DBH, DEN, DEM, MIRV1, TWI, VRM이 유의미한 변수로 선택되었고

수정된 결정계수는 약 0.35로 표본점 전체를 이용했을 때보다 높게 산출되었다. 이는 활엽수림 단일 임상만을 포함하고 있고, 다양한 변수를 설명변수로 활용할 수 있었기 때문으로 볼 수 있다. 하지만 표본점 전체를 이용했을 때와 마찬가지로 DBH와 AGE가 큰 임목일수록 바이오매스의 저장량이 높을 것으로 예상되었으나 실제 분석에서 음(-)의 부호가 나온 점 등은 추후 지역 특성과 수종 등을 고려한 더욱 정교한 회귀모델링을 통해 보정해야 할 부분이라 할 수 있다.

침엽수림과 혼효림에 대한 회귀모델링 결과는 표 4와 같았다. AGE, RVI이 유의미한 변수로 선택되었고 수정된 결정계수는 약 0.18로 활엽수림만을 사용한 것보다 낮게 산출되었다. 이는 침엽수림과 혼효림이 혼재하여 임분의 구조가 복잡하고 표본점의 수가 42개로 작기 때문인 것으로 볼 수 있다.

침엽수림과 혼효림 그룹, 활엽수림 그룹의

TABLE 4. 침엽수림과 혼효림 바이오매스에 대한 회귀모델링 결과

Predictor Variables	Coefficients	Std. Error	t Value	Pr(> t)
(Intercept)	-149.57	86.37	-1.732	0.09123 .
AGE	45.14	13.69	3.296	0.00209 **
RVI	69.48	41.30	1.683	0.10046

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 55.27 on 39 degrees of freedom

Multiple R-squared: 0.2214, Adjusted R-squared: 0.1815

F-statistic: 5.546 on 2 and 39 DF, p-value: 0.007588

FIGURE 8. 단양군 지역 침엽수림과 혼효림의 잔차 베리오그램(좌측)과 활엽수림의 잔차 베리오그램(우측)(단위: m)

회귀 잔차에 대해 적합된 베리오그램 모형은 그림 8과 같다. 두 그룹 모두 공간적 상관관계의 범위(total range)가 약 33,266m와 14,818m로 매우 크게 나타났으며 너짓 효과도 크게 나타났다. 특히 침엽수림과 혼효림 그룹에서는 회귀모형으로 설명된 경향을 제거한 잔차에 대해 공간적 상관관계가 잘 드러나

지 않고 약간의 선형관계만을 보이고 있다. 이들 베리오그램을 이용한 회귀 크리깅 결과는 그림 9, 그림 10과 같다.

그림 11은 회귀 크리깅을 이용하여 임상별로 추정한 결과를 취합하여 침엽수, 활엽수, 혼효림 전체에 대한 ha당 바이오매스를 추정한 결과이다. 정규 크리깅이나 표본점 전체를

FIGURE 9. 회귀 크리깅을 이용한 바이오매스 예측치와 표준편차(침엽수림과 혼효림 그룹)

FIGURE 10. 회귀 크리깅을 이용한 바이오매스 예측치와 표준편차(활엽수림 그룹)

FIGURE 11. 회귀 크리깅을 이용한 바이오매스 예측치와 표준편차(임상별 추정 결과 취함)

이용한 회귀 크리깅에 비해 더욱 상세한 임분 경계를 반영하고 있으며, 표준편차 역시 단양군의 외곽 지역에서 크게 나타나는 것을 확인할 수 있다. 이 역시 과대/과소 추정치를 제거해 주기 위해 1.96 표준편차(95%)에 해당하는 값만을 추출하여 지도화하였다. 이렇게 추정된 바이오매스의 최소값은 10.13ton/ha, 최대값은 198.83ton/ha, 평균은 102.76ton/ha

를 나타냈으며, 이에 따른 단양군 산림면적 64,932ha에 해당하는 지상부 바이오매스의 총량은 6,672,412ton, 탄소전환계수 0.5를 적용한 탄소저장량은 3,336,206 tonC로 추정되었다. 그림 12는 임상별로 층화한 회귀 크리깅을 이용하여 추정된 단양군의 지상부 바이오매스 탄소저장량을 지도화한 것이다.

FIGURE 12. 단양군의 지상부 바이오매스 탄소저장량

3. 탄소저장량 및 예측능력 비교 결과

표 5는 앞에서 분석한 크리깅 방법과 베리오그램 모형에 따른 단양군의 탄소저장량과 RMSE를 비교한 것이다. 그 결과 베리오그램 모형에 따른 추정치의 차이는 정규 크리깅을 사용하였을 경우 26,947 tonC, 회귀 크리깅을 사용하였을 경우 23,701 tonC로 나타났다. 또한 크리깅 방법에 따라서는 탄소저장량 추정치의 차이가 더욱 크게 발생하였는데, 표본점 전체에 대해 구형 모형을 이용하였을 경우 회귀 크리깅 추정치가 정규 크리깅 추정치보다 약 75,321 tonC 적게 추정되었고 지수형 모형을 이용하였을 경우에도 회귀 크리깅 추정치가 72,075 tonC 적게 추정되었다. 또한 크리깅 방법과 베리오그램 모형 조합의 예

측능력 비교를 위해 Leave-one-out 교차검증을 통한 RMSE를 계산하였으며, 회귀 크리깅의 경우가 정규 크리깅에 비해 RMSE가 약 0.23 tonC/ha 작게 나타났다.

단양군의 탄소저장량 추정을 위해 본 연구에서는 총 120개의 국가산림조사 표본점을 사용하였는데 표본점 전체를 사용한 회귀 크리깅 추정에서는 RMSE가 상대적으로 작게 나왔다. 그러나 좀 더 세밀한 분석을 수행하고자 표본점을 임상별로 층화할 경우, 회귀모델의 설명력은 높아졌지만 표본점의 수가 적어 베리오그램 적합에 적절치 않게 되므로 잔차의 공간 구조를 파악하는 데에 어려움이 있었고, RMSE도 상대적으로 높아지는 경향이 있었다.

TABLE 5. 크리깅 방법에 따른 단양군의 탄소저장량과 RMSE 비교

크리깅 방법	베리오그램 모형	탄소저장량(tonC)	RMSE(tonC/ha)
정규 크리깅 (표본점 전체)	구형	3,459,902	22.55
	지수형	3,432,955	22.56
회귀 크리깅 (표본점 전체)	구형	3,384,581	22.32
	지수형	3,360,880	22.33
회귀 크리깅 (임상별 층화)	구형	3,336,206	23.02(침엽수림&혼효림)/ 23.40(활엽수림)
	지수형	3,348,543	22.89(침엽수림&혼효림)/ 23.43(활엽수림)

결론 및 한계

본 연구에서는 크리깅 기법 중에서 정규 크리깅과 회귀 크리깅을 이용하여 지상부 바이오매스의 탄소저장량 추정을 시도하였다. 4 km×4 km 간격의 국가산림조사 표본점만을 이용하는 정규 크리깅에 비해 회귀 크리깅은 여러 보조 자료를 활용할 수 있는 하이브리드 내삽 방식으로, 공간적인 이질성이 큰 우리나라의 산림 조건을 고려했을 때 더욱 상세한 바이오매스의 공간 분포를 파악할 수 있다는 특징이 있었다. 또한 표본점 전체를 이용하였을 경우를 비교해 보면, 회귀 크리깅의 경우 정규 크리깅에 비해 RMSE가 미약하게나마 감소하는 것을 알 수 있었다.

본 연구를 통해 회귀 크리깅을 수행할 때의 유의점도 살펴볼 수 있었다. 첫째, 회귀 크리깅은 컴퓨팅 시간이 많이 소요되므로 고사양의 자동화 환경이 필요하고 시스템 환경에 따른 적정 공간 해상도의 선택이 필요하다. 본 연구에서는 64bit 운영체제, 4GB RAM에서 30 m×30 m의 셀 크기로 작업이 가능했다. 두 번째는 여러 보조 자료의 사전 가공 노력이 필요하다는 것이다. 자료들 간의 좌표체계 일치와 형태간의 변환, 공간 해상도를 일치시키기 위한 리샘플링과 같은 일반적인 처리뿐만 아니라 표본점과 국가산림조사 자료의 속성 결합, 영상 자료의 밴드연산, 지형 변수 생성을 위한 분석 작업 등이 그것이다. 세 번째는 회귀모델링을 별도로 수행해야 하며 회귀 모델링 결과에 따라 예측 결과가 달라질 수

있다는 것이다. 따라서 회귀모델의 변수로 사용된 보조 자료의 정확성이 요구되며 차후에는 더욱 정교한 회귀모델링에 대한 고민이 필요하다고 할 수 있다. [KAGIS](#)

참고문헌

- 국립산림과학원. 2009. 제5차 국가산림자원조사, 현지조사 매뉴얼. 54쪽.
- 김기영, 전명식, 강현철, 이성건 공역. 2009. 예제를 통한 회귀분석, 제4판. 자유아카데미. 393쪽.
- 단양군. 2009. 통계연보(제2장 토지 및 기후).
- 단양군. 2010. 단양군 산림정밀지도 제작 및 활용시스템 통합 완료보고서. 154쪽.
- 손영모, 이경학, 김래현. 2007. 우리나라 산림 바이오매스 추정. 한국임학회지 96(4):477-482.
- 이승호, 김철민, 원현규, 김경민, 조현국. 2004. Landsat TM 위성영상을 이용한 산림자원량 산정. 한국임학회 2004 학술연구 발표논문집 제1권. 250-252쪽.
- 임종수, 공지수, 김성호, 신만용. 2007. kNN 기법을 이용한 강원도 평창군의 산림주제도 작성과 산림통계량 추정. 한국임학회지 96(3):259-268.
- 임종수, 한원성, 황주호, 정상영, 조현국, 신만용. 2009. 위성영상자료 및 국가 산림자원조

- 사 자료를 이용한 산림 바이오매스 추정, 대한원격탐사학회지 25(4):311-320.
- 장안진, 김형태. 2008. 항공사진과 Lidar 데이터를 이용한 산림지역의 바이오매스 추정에 관한 연구. 한국지리정보학회지 11(3):166-173.
- 정상영, 임종수, 조현국, 정진현, 김성호, 신만용. 2009. 산림 바이오매스 변환표와 위성영상을 이용한 무주군의 산림 바이오매스추정. 한국임학회지 98(4):409-417.
- 최종근. 2007. 지구통계학. 시그마프레스. 199쪽.
- Ahmed, S., G. de Marsily. 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research 23(9):1717-1737.
- Bailey, T.C. and A.C. Gatrell. 1996. Interactive Spatial Data Analysis. Prentice Hall, London, UK. 432pp.
- Cressie, N. 1993. Statistics for Spatial Data. John Wiley & Sons, NY, USA. 416pp.
- Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation(Applied Geostatistics). Oxford University Press, NY, USA. 496pp.
- Hengl, T. 2009. A Practical Guide to Geostatistical Mapping(2nd). University of Amsterdam, Netherlands. 291pp.
- Hengl, T., G. Heuvelink and A. Stein. 2004. A generic framework for spatial prediction of soil variables based on regression kriging. Geoderma 122(1-2):75-93.
- Hengl, T., G. Heuvelink and D. Rossiter. 2007. About regression-kriging: From equations to case studies. Computers & Geosciences 33:1301-1315.
- Hosseini, S., S. Khajeddin and H. Azarnivand. 2004. Application of ETM+ data for estimating rangelands cover percentage. 20th ISPRS Congress Youth Forum. Istanbul, Turkey, July 12-23, 2004, pp.198-201.
- Isaaks, E. and R. Srivastava. 1989. An Introduction to Applied Geostatistics. Oxford University Press, NY, USA. 542pp.
- Lu, D. 2006. The potential and challenge of remote sensing based biomass estimation, International Journal of Remote Sensing 27:1297-1328.
- Meng, Q., C. Cieszewski and M. Madden. 2009. Large area forest inventory using Landsat ETM+: a geostatistical approach. ISPRS Journal of Photogrammetry and Remote Sensing 64(1):27-36.
- Odeh, I., A. McBratney and D. Chittleborough. 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63(3-4):197-214.
- Odeh, I., A. McBratney and D. Chittleborough. 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67(3-4):215-226.
- Rahman, M., E. Csaplivos. and B. Koch. 2008. Satellite estimation of forest carbon using regression model. International Journal of Remote Sensing

- 29(23):6917-6936.
- Sales, M., C. Souza Jr, P. Kyriakidis, D. Roberts and E. Vidal. 2007. Improving spatial distribution estimation of forest biomass with geostatistics: a case study for Rondônia, Brazil. *Ecological Modelling* 205(1-2):221-230.
- Sappington, J.M., K.M. Longshore and D.B. Thomson. 2007. Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave desert, *Journal of Wildlife Management* 71(5):1419-1426.
- Sørensen, R., U. Zinko and J. Seibert. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences Discussions* 10:101-112.
- Webster, R. and M.A. Oliver. 2007. *Geostatistics for Environmental Scientists, Statistics in Practice*(2nd). John Wiley & Sons, Chichester, England. 330pp.