

지가 추정을 위한 공간내삽법의 정확성 평가

전병운^{1*}

Evaluating the Accuracy of Spatial Interpolators for Estimating Land Price

Byong-Woon JUN^{1*}

요 약

지금까지 지가나 주택가격을 추정하는데 회귀모형 기반 공간내삽법과 크리깅(Kriging) 기반 공간내삽법이 많이 사용되었지만, 이들 공간내삽법의 성능을 서로 비교한 연구는 거의 없는 실정이다. 따라서 본 연구는 대구시 달서구를 사례로 지가를 추정하는데 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법을 적용해 보고 그 정확성을 평가하였다. 회귀모형 기반 공간내삽을 위해 최소자승모형(OLS), 공간지체모형(SLM), 공간오차모형(SEM), 지리가중회귀모형(GWR)을 사용하였고, 크리깅 기반 공간내삽을 위해 단순 크리깅(SK), 정규 크리깅(OK), 일반 크리깅(UK), 공동 크리깅(CK)을 이용하였다. 먼저, 전역적 정확성 지수인 평균 제곱근 오차(RMSE), 수정된 평균 제곱근 오차(adjusted RMSE), 분산지수(COD)를 이용하여 그 정확성을 통계적으로 평가하였다. 다음으로, 3차원 잔차도와 산점도를 이용하여 그 정확성을 시각적으로 서로 비교하였다. 통계적 및 시각적 분석결과에 의하면, 공간적 의존성을 반영할 수 있는 공간회귀모형(SAR)과 크리깅 기법들 보다 공간적 이질성을 고려할 수 있는 GWR이 사례지역에서 지가를 추정하는데 상대적으로 정확한 공간내삽방법인 것으로 나타났다. 본 연구의 결과는 지가를 통해 도시의 공간구조를 분석하는 이차적 연구에 기여할 것이다.

주요어 : 지가, 공간적 의존성, 공간적 이질성, 공간회귀모형, 지리가중회귀모형, 크리깅

ABSTRACT

Until recently, regression based spatial interpolation methods and Kriging based spatial interpolation methods have been largely used to estimate land price or housing price, but less attention has been paid on comparing the performance of these spatial interpolation methods. In this regard, this research applied regression based spatial interpolators and Kriging based spatial interpolators for estimating the land prices in Dalseo-gu, Daegu

2017년 9월 19일 접수 Received on September 19, 2017 / 2017년 9월 23일 수정 Revised on September 23, 2017 /
2017년 9월 23일 심사완료 Accepted on September 23, 2017

1 경북대학교 지리학과 Department of Geography, Kyungpook National University

* Corresponding Author E-mail : bwjun@knu.ac.kr

metropolitan city and evaluated the accuracy of eight spatial interpolators. OLS, SLM, SEM, and GWR were used as regression based spatial interpolators while SK, OK, UK, and CK were employed as Kriging based spatial interpolators. The global accuracy was statistically evaluated by RMSE, adjusted RMSE, and COD. The relative accuracy was visually compared by three-dimensional residual error map and scatterplot. Results from statistical and visual analyses indicate that GWR reflecting the spatial non-stationarity was a relatively more accurate spatial predictor to estimate land prices in the study area than SAR and Kriging based spatial interpolators considering the spatial dependence. The findings from this research will contribute to the secondary research into analyzing the urban spatial structure with land prices.

KEYWORDS : Land Price, Spatial Dependence, Spatial Non-Stationarity, SAR, GWR, Kriging

서론

지가는 토지 과세의 기준이면서 도시의 토지 이용을 결정하는 중요한 요인으로서 도시의 공간구조와 그 변화를 분석하기 위한 지표로 많이 활용이 된다(Kim *et al.*, 2005; Kim, 2006; Park and Koo, 2011; Kim, 2012; Kim and Lee, 2013). 따라서 도시 내에서 지가를 정확하게 추정하는 것은 정확하고 신속한 과세를 위해 필요할 뿐만 아니라 도시의 공간구조를 정확하게 포착하기 위해 매우 중요하다.

1989년 이래 우리나라는 표준지 공시지가와 개별공시지가를 제공하고 있다. 표준지 공시지가는 감정평가사의 정밀평가에 의해 산정되지만 개별공시지가는 표준지를 이용한 비교방식에 의해 산정된다. 표준지를 이용한 비교방식에 의한 개별공시지가 산정방법은 비교 표준지와 대상 필지의 토지특성을 비교하여 토지가격비준표로부터 추출한 가격배율을 비교 표준지 공시지가에 곱하여 산출한다. 개별공시지가의 대량산정 방법은 효율적 처리를 위해 다수의 표준지 공시지가를 기준으로 다수의 토지가격비준표를 활용한다(Kim and Kim, 2004). 이러한 토지가격비준표를 활용한 개별공시지가의 대량산정방법에 대한 대안으로 최소자승(OLS: Ordinary Least Squares) 기반 헤도닉 회귀모형(hedonic regression)이 널리 활용되고 있다(Seo and Lee, 2

001; Choi and Lee, 2006; Song and Park, 2012). OLS 기반 헤도닉 회귀모형은 표준지 공시지가를 종속변수로 표준지의 구조적 및 주변적 특성(형상, 지세, 용도지역, 도로조건, 토지 이용상황 등)을 설명변수로 하는 OLS 기반 회귀모형을 적용하여 지가를 추정하는 방법이다. 그러나 이 방법은 지가가 가지는 공간적 의존성(spatial dependence)과 공간적 이질성(spatial non-stationarity)이라는 공간적 효과(spatial effect)를 고려할 수 없기 때문에 비효율적이고 편기된 모형을 설정할 수 있다는 단점이 있어 지가 추정의 정확성을 떨어뜨릴 수 있다(Seo and Lee, 2001; Choi and Lee, 2006; Kim and Jun, 2012; Song and Park, 2012; Lee and Park, 2013). 지가의 공간적 의존성은 한 지점의 지가는 멀리 떨어진 지점의 지가 보다 인근 지점의 지가에 의해 영향을 더 많이 받는다는 것이다. 반면에, 지가의 공간적 이질성은 지가와 이에 영향을 주는 토지의 구조적 및 주변적 특성의 관계가 공간상에서 일정하지 않고 상이하다는 것이다. 사실상 도시 내에서 지가 추정을 위해 공간적 효과를 배제한다는 것은 비현실적이다.

이러한 OLS 기반 헤도닉 회귀모형의 단점을 보완하기 위하여 공간회귀모형(SAR: Spatial Auto-Regression)과 지리가중회귀모형(GWR: Geographically Weighted Regression) 그리고 크리깅(Kriging) 기반 공간내삽법이 보편적으로 활용되었다. 일부 연구는 지가 추정에 있어서 공간

적 의존성을 반영하기 위해 SAR(Seo and Lee, 2001; Choi and Lee, 2006; Song and Park, 2012; Lee and Park, 2013)와 크리깅 기반 공간내삽법(Kim, 2012; Lee *et al.*, 2014; Choi and Kim, 2015; Sohn and Lee, 2016)을 활용하였다. 다른 연구는 지가 추정에 있어서 공간적 이질성을 고려하기 위해 GWR을 적용하였다(Seo, 2003; Kim and Jun, 2012). 이들 방법은 크게 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법으로 구분할 수 있다. 공간내삽법은 일정한 지역에서 관측지점의 값을 이용하여 미관측 지점의 값을 추정하는 공간분석기법이다(Yang and Hodler, 2000; Kim *et al.*, 2013). 이 기법은 소수의 표준지 지가를 이용하여 미관측 지점에서 다수의 개별필지 지가를 신속하게 추정할 수 있게 한다. SAR와 GWR은 흔히 모수 추정이나 예측을 위해 사용되는 회귀모형이고 지가라는 공간자료를 예측하는데 적용되었기 때문에 지가 추정을 위한 회귀모형 기반 공간내삽법으로 간주할 수 있다. 반면에 크리깅 기반 공간내삽법은 지가 추정을 위해 지구통계학적 모델링(geostatistical modeling)을 할 수 있는 크리깅 기법을 적용한 것이다. 지가뿐만 아니라 주택가격을 추정하는데도 회귀모형 기반 공간내삽법(Park *et al.*, 2003; Militino *et al.*, 2004; Chrostek and Kopczewska, 2013; Jin and Lee, 2014; Chun, 2016)과 크리깅 기반 공간내삽법(Chica-Olmo, 2007; Montero and Larraz, 2011; Kim *et al.*, 2013; Kuntz and Helbich, 2014)이 많이 활용되었다.

지금까지 지가나 주택가격을 추정하는데 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법이 많이 사용되었지만, 이들 공간내삽법의 성능을 체계적으로 서로 비교한 연구는 거의 없는 실정이다. 일부 연구는 지가나 주택가격을 추정하는데 있어서 SAR 및 GWR과 같은 회귀모형 기반 공간내삽법의 정확성을 서로 비교하였다(Park *et al.*, 2003; Song and Park, 2012; Chrostek and Kopczewska, 2013; Chun, 2016). 반면에, 다른 연구는 지가나 주택가격 추정을 위해 크리깅 기반 공간내삽법의 정확성을

상대적으로 평가하였다(Chica-Olmo, 2007; Kim *et al.*, 2013; Kuntz and Helbich, 2014; Choi and Kim, 2015; Sohn and Lee, 2016). 하지만, 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법의 정확성을 상대적으로 서로 비교한 연구는 아직 없는 실정이다. 이들 중에서 어떤 공간내삽법이 지가를 추정하는데 가장 정확한 방법인지에 대한 실증적 연구가 필요하다. 따라서 본 연구는 대구시 달서구를 사례로 지가를 추정하는데 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법을 적용해 보고 그 정확성을 평가하고자 한다.

지가 및 주택가격 추정을 위한 공간내삽법

지가나 주택가격 추정에 관한 기존의 연구들은 크게 회귀모형 기반 공간내삽법을 활용한 연구와 크리깅 기반 공간내삽법을 적용한 연구로 나눌 수 있다. 먼저, 지가 추정을 위해 회귀모형 기반 공간내삽법을 활용한 연구는 다시 지가의 공간적 의존성을 고려하기 위해 SAR를 활용한 연구(Seo and Lee, 2001; Choi and Lee, 2006; Song and Park, 2012; Lee and Park, 2013)와 지가의 공간적 이질성을 감안하기 위해 GWR을 적용한 연구(Seo, 2003; Kim and Jun, 2012)로 구분할 수 있다.

Seo and Lee(2001)는 서부산권의 7개 행정구를 사례로 지가를 추정하기 위해 공간적 의존성과 주택하위시장을 고려할 수 있는 공간지체모형(SLM: Spatial Lag Model) 기반 헤도닉 회귀모형을 적용하였다. 이들은 SLM이 OLS 보다 지가 추정에 있어서 모형의 적합도가 우수하다는 것을 밝혔다. SLM은 SAR의 일종으로서 종속변수인 지가의 공간적 의존성을 다룰 수 있는 회귀모형이다. Choi and Lee(2006)는 창원시를 사례로 주거지역의 지가를 추정하기 위하여 공간적 의존성과 주변 용도지역의 접근성을 고려할 수 있는 SLM 및 공간오차모형(SEM: Spatial Error Model) 기반 헤도닉 회귀모형을 적용하였다. 그 결과, SEM이 OLS나 SLM 보

다 주거지 내에서 지가 추정을 위한 모형의 적합도에서 높게 나타났다. SEM은 SAR의 한 유형으로서 지가 추정에 있어서 잔차의 공간적 의존성을 다룰 수 있는 회귀모형이다. Song and Park(2012)는 경기도 광주시를 사례로 농지가격을 추정하기 위하여 SLM, SEM, 공간회귀혼합모형(SAC: Spatial Autoregressive Confused), GWR 기반 헤도닉 회귀모형을 적용하고 그 정확성을 서로 비교하였다. 그 분석 결과, GWR이 가장 정확한 것으로 나타났다. SAC은 SAR의 한 유형으로서 SLM과 SEM을 혼합한 회귀모형이다. Lee and Park(2013)은 인천시의 7개 행정구를 사례로 지가를 추정하기 위해 SLM과 SEM 기반 헤도닉 회귀모형을 적용하였다. 이들은 공간가중행렬을 구성할 때에 인근지역의 범위를 좁게 설정할수록 SLM과 SEM의 모형 적합도가 향상되었다는 것을 확인하였다.

Seo(2003)는 서부산권의 7개 행정구를 사례로 지가 추정에 있어서 공간적 이질성을 설명하기 위하여 GWR 기반 헤도닉 회귀모형을 적용하였다. 그는 GWR이 OLS 보다 나은 모형 적합도를 보였다고 논증하였다. Kim and Jun(2012)는 서울시 강남구를 사례로 공간구문론을 이용하여 접근성 변수를 개발하고 지가 추정에 있어서 공간적 이질성을 고려하기 위하여 GWR 기반 헤도닉 회귀모형을 적용하였다. 이들은 공간구문론을 활용하여 개발한 접근성 변수가 유의하고 GWR이 OLS 보다 향상된 모형 적합도를 나타낸 것을 확인하였다.

다음으로 지가 추정에 있어서 공간적 의존성을 반영하기 위해 크리깅 기반 공간내삽법을 적용한 연구는 Kim(2012), Lee *et al.*(2014), Choi and Kim(2015), Sohn and Lee(2016) 등이 있다. Kim(2012)은 대구시를 사례로 1995년부터 2011년까지 지가의 시공간적 변화를 분석하기 위하여 정규 크리깅(OK: Ordinary Kriging)을 이용하여 지가면(land value surface)을 추정하였다. Lee *et al.*(2014)은 서울시를 사례로 상업용 토지의 실거래가를 추정하기 위하여 전문가 지식과 공간적 의존성을 반영할 수 있는 베이저안 크리깅(BK: Bayesian Kriging)을 적용하

였다. 이들은 전문가 지식과 공간적 의존성을 모두 반영한 BK가 다른 크리깅 모형 보다 가격 예측력에 있어서 향상된 것을 확인하였다. Choi and Kim(2015)는 경북 영천시를 사례로 토지 실거래가를 추정하기 위하여 OK와 공동 크리깅(CK: Co-Kriging)을 활용하였다. 그 결과, CK가 OK 보다 토지 실거래가를 효과적으로 추정한 것으로 나타났다. Sohn and Lee(2016)는 전북 김제시를 사례로 농지가격을 추정하기 위하여 회귀 크리깅(RK: Regression Kriging)을 적용하였다. 이들은 RK 모형이 OLS 기반 헤도닉 회귀모형 보다 상대적으로 정확하고 농지가격의 공간적 의존성을 반영한 것을 확인하였다.

한편, 주택가격을 추정하기 위하여 회귀모형 기반 공간내삽법을 활용한 연구는 Park *et al.*(2003), Militino *et al.*(2004), Chrostek and Kopczewska(2013), Jin and Lee(2014), Chun(2016) 등이 있다. Park *et al.*(2003)은 서울시 광진구의 6개 아파트 단지를 사례로 아파트 가격을 추정하기 위하여 OLS, SLM, SEM 기반 헤도닉 회귀모형을 적용하였다. 그 결과, SLM이 상대적으로 정확한 것으로 나타났다. Militino *et al.*(2004)는 스페인 Pamplona시를 사례로 아파트 가격을 추정하는데 있어서 공간적 의존성을 반영하기 위한 대안적 방법으로 조건부 자기회귀(CAR: Conditional Autoregressive) 모형과 선형혼합효과(LME: Linear Mixed Effects) 모형을 제안하였다. Chrostek and Kopczewska(2013)는 폴란드 Wroclaw시를 사례로 주택가격을 추정하기 위하여 OLS, 공간확장모형(spatial expansion model), SLM, SEM, GWR 등 다양한 헤도닉 회귀모형을 적용하고 그 정확성을 평가하였다. 이들은 그 중에서 GWR이 가장 정확한 방법이라고 제시하였다. Jin and Lee(2014)는 서울시를 사례로 아파트 단지의 평균가격을 추정하기 위하여 지리가중라소(GWL: Geographically Weighted LASSO) 모형을 제안하였다. 이들은 GWL은 GWR이 가지는 독립변수의 다중공선성 문제를 보완할 수 있다고 주장하였다. Chun(2016)은 서울시를 사례로 아파트 실거래가를 추정하기 위하여 OLS, SLM, SEM, GWR

기반 헤도닉 회귀모형을 적용하고 모형의 적합도를 상대적으로 비교하였다. 그 분석 결과, GWR이 모형 적합도에서 가장 높은 것으로 나타났다.

마지막으로, 주택가격을 추정하기 위하여 크리깅 기반 공간내삽법을 활용한 연구는 Chica-Olmo(2007), Montero and Larraz(2011), Kim *et al.*(2013), Kuntz and Helbich(2014) 등이 있다. Chica-Olmo(2007)는 스페인 Granada시에서 아파트 가격을 추정하기 위하여 CK를 적용하였다. 그 결과, CK가 OK 보다 상대적으로 정확한 것으로 나타났다. Montero and Larraz(2011)는 스페인 Toledo시를 사례로 주택과 상가 가격을 추정하기 위하여 IDW, OK, CK를 적용하고 그 정확성을 평가하였다. 이들은 CK가 OK 보다 정확하고 OK가 역거리가중(IDW: Inverse Distance Weighting) 내삽법 보다 정확하다는 것을 증명하였다. Kim *et al.*(2013)은 서울시를 사례로 아파트 실거래가를 추정하기 위하여 OK를 적용하고 최근된 내삽법 및 IDW와 그 정확성을 서로 비교하였다. 그 결과, 이론과 달리 IDW가 OK 보다 정확하다는 것을 확인하였다. Kuntz and Helbich(2014)는 오스트리아 Vienna시를 사례로 단독주택가격을 추정하기 위하여 탈경향 크리깅(DK: Detrended Kriging), 일반 크리깅(UK: Universal Kriging), 탈경향 공동 크리깅(DCK: Detrended Co-Kriging), 일반 공동 크리깅(UCK: Universal Co-Kriging)을 적용하고 그 정확성을 서로 비교하였다. 이들은 DK와 UK는 그 정확성이 대체로 유사하고 UCK는 DCK 보다 상대적으로 정확하다는 것을 주장하였다.

현재까지 지가나 주택가격을 추정하기 위하여 다양한 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법이 널리 활용이 되었다. 하지만, 이들 공간내삽법의 정확성은 사례지역에 따라 상이하게 나타났다. 또한, 동일한 사례지역에서 두 그룹의 공간내삽법의 정확성을 상대적으로 비교한 실증적 연구는 아직 학계에 보고되지 않은 상황이다. 지가 추정을 위한 가장 정확한 공간내삽법을 선정하기 위해 이들 방법을 비교·평가할 필요가 있다.

연구자료 및 방법

본 연구의 사례지역은 대구광역시 달서구이다(그림 1). 달서구는 2012년부터 2016년까지 최근 5년간 대구광역시 구·군 중에서 토지거래량이 평균적으로 가장 많은 지역이었기 때문에 사례지역으로 선정하였다. 또한, 최근 5년 동안 달서구의 토지거래가 가장 활발하게 일어났던 2013년을 본 연구의 시간적 범위로 설정하였다.

지가 추정을 위한 공간내삽법의 정확성을 비교·평가하기 위해서 본 연구에서는 국토교통부 부동산 공시가격 알리미 웹사이트에서 제공하는 달서구의 1,380개의 표준지 공시지가 자료를 이용하였다. 표준지 공시지가는 지번 주소로 소재지가 표시되어 있다. 이를 경위도 좌표로 변환하기 위해 지오코딩 프로그램인 GeoCoder-Xr 2.0을 사용하였다. 그러나 GeoCoder-Xr 2.0은 도로명 주소를 경위도 좌표로 변환하기 때문에 먼저 1,380개의 표준지 공시지가의 소재지를 도로명 주소로 변환하였다. 1,380개의

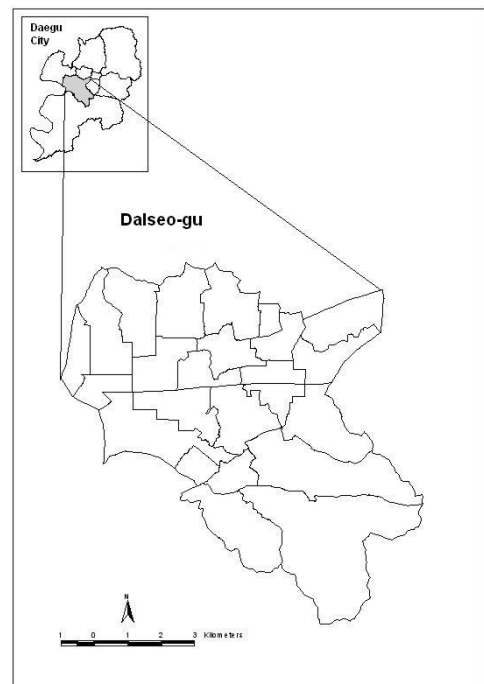


FIGURE 1. Location of study area

소재지 중에서 126개의 소재지는 도로명 주소를 아직 부여받지 못했거나 토지주가 도로명 주소를 신청하지 않는 등의 이유로 도로명 주소가 아직 사용되고 있지 않았다. 그래서 해당 지점들은 V-World에서 제공하는 X-Ray Map을 이용하여 지번 주소를 입력한 다음 해당 지번의 경위도 좌표를 직접 취득하였다. 이렇게 취득한 1,380개의 표준지 공시지가의 경위도 좌표를 ArcGIS 10.1을 이용하여 세계측지계의 TM (Transverse Mercator) 좌표로 변환하였다.

그림 2는 본 연구의 전반적인 분석과정을 보여준다. 먼저, 총 1,380개의 표준지는 1,095개의 자료점과 285개의 검사점으로 나누었다(그림 3). 총 1,380개의 표준지 중에서 약 20%에 해당되는 285개의 표준지를 교차검증을 위해 검사점으로 추출하였다. 285개의 검사점을 달서구 전체에 균일하게 분포시키기 위하여 24개 행정동별로 층화비례배분하여 임의로 추출하였다. 그 나머지인 1,095개의 표준지를 자료점으로

추출하고 그 자료점에서 공시지가를 바탕으로 회귀모형 공간내삽법과 크리깅 공간내삽법을 적용하여 지가를 추정하였다. 회귀모형 기반 공간내삽법으로 OLS, SLM, SEM, GWR 기반 헤도닉 회귀모형을 사용하였다. 반면에, 크리깅 기반 공간내삽법으로 단순 크리깅(SK: Simple Kriging), OK, UK, CK를 이용하였다. 이들 회귀모형 기반 공간내삽법과 크리깅 공간내삽법은 기존의 선행연구에서 많이 사용되었기 때문에 본 연구에서 선정하였다.

본 연구에서 사용된 지가 추정을 위한 OLS 기반 헤도닉 회귀계수는 식 1과 같이 표현된다. OLS는 전역적으로 회귀계수를 추정하여 지가의 전반적인 경향을 파악할 수 있게 해 준다.

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i \quad (1)$$

y_i : 표준지 공시지가

x_{1i} : 용도지역

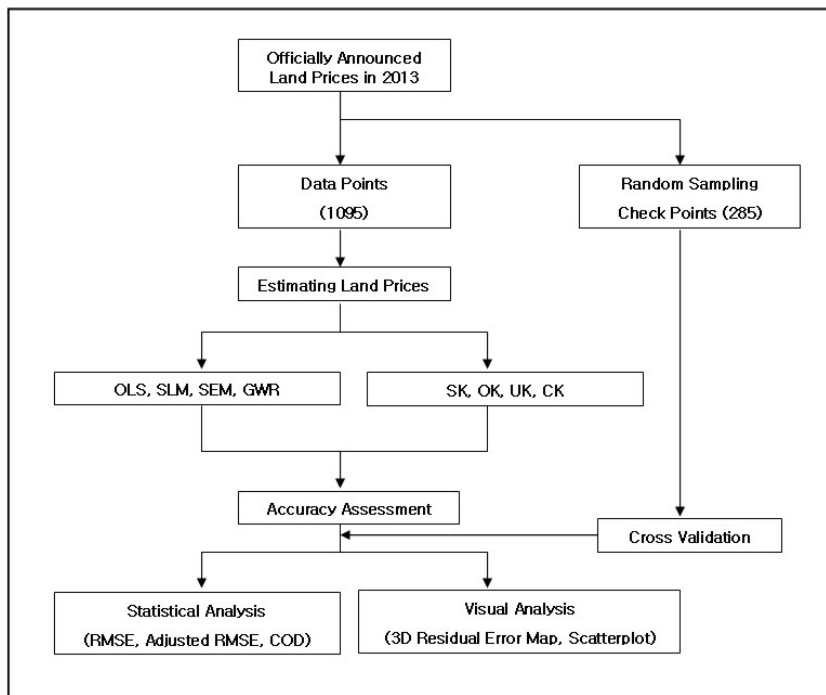


FIGURE 2. Research flow

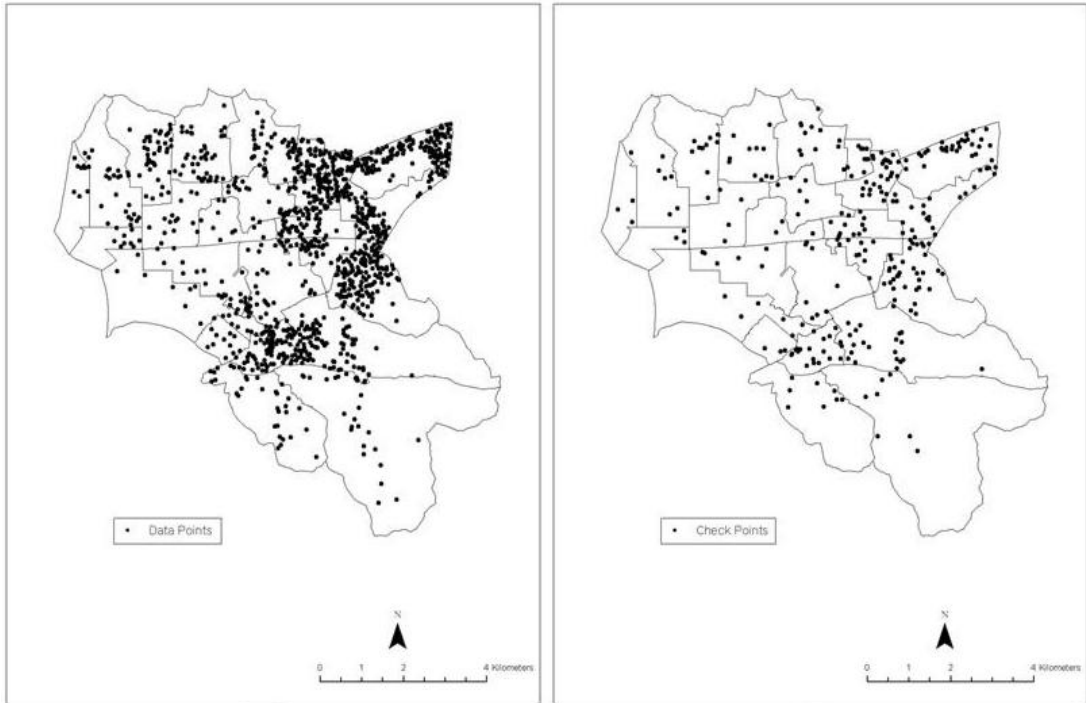


FIGURE 3. The spatial distribution of data points and check points

x_{2i} : 도로조건
 α : 상수
 β_1, β_2 : 회귀계수

여기서, 독립변수인 용도지역과 도로조건은 원래 범주형 자료로 제공이 되었지만, 본 연구에서는 분석의 편의상 가중치를 부여하여 연속형 자료로 변환하여 사용하였다. 기존의 연구에서 범주형 독립변수는 보통 더미변수로 처리하였지만, 이것은 또 다른 문제점을 초래한다. 예를 들면, 많은 더미독립변수를 헤도닉 회귀모형에 투입하면 독립변수간의 공선성 문제가 발생할 수 있고 결정계수를 증가시켜 모수를 과대 추정할 수 있는 단점이 있다.

중속변수인 지가의 공간적 의존성을 고려하기 위한 SLM 기반 헤도닉 회귀모형은 식 2와 같다. 본 연구에서 공간가중행렬(spatial weighted matrix)은 지수형((exponential) 반베리오그램

(semi-variogram) 모형을 이용하여 탐색한 상관거리(range)인 930m를 인접지역의 탐색거리로 설정하여 구축하였다.

$$y_i = \alpha + \rho Wy + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i \quad (2)$$

ρ : 공간자기회귀계수
 Wy : 공간지체변수

지가 추정에 있어서 그 잔차의 공간적 의존성을 반영하기 위한 SEM 기반 헤도닉 회귀모형은 식 3과 같다. 여기서도 위에서 언급한 동일한 공간가중행렬이 적용되었다.

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \lambda W\epsilon + \nu \quad (3)$$

λ : 공간자기회귀계수
 $W\epsilon$: 공간오차변수

ν : $N(0, \sigma^2)$ 인 서로 독립적인 오차

지가 추정에 있어서 공간적 이질성을 감안하기 위한 GWR 기반 헤도닉 회귀모형은 식 4와 같다. GWR은 OLS와 유사하지만, 표준지의 소재지에 따라 국지적으로 회귀계수를 산출할 수 있다는 점이 다르다.

$$y_i = \alpha(\mu_i, \nu_i) + \beta_1(\mu_i, \nu_i)x_{1i} + \beta_2(\mu_i, \nu_i)x_{2i} + \epsilon_i \quad (4)$$

(μ_i, ν_i) : 표준지의 좌표

지가 추정에 있어서 공간적 의존성을 고려하기 위한 크리깅 모형의 일반식은 식 5와 같다. 크리깅은 제한된 수의 표준지 공시지가 자료와 그 공간적 상호관계를 이용하여 표준지에 인접한 지역의 개별공시지가를 효율적으로 추정할 수 있게 하는 지구통계학적 방법이다. 이 기법은 주로 기온, 강수, 오염농도 등과 같은 자연환경자료를 추정하는데 사용되었지만, 지가와 같은 사회경제적 자료를 추정하는데 최근에 와서 많이 사용이 되고 있다.

$$Z(l) = m(l) + u(l) + \epsilon \quad (5)$$

$Z(l)$: 추정 지가

$m(l)$: 전반적 경향

$u(l)$: 임의적이지만, 공간적으로 자기상관적인

$m(l)$ 로부터의 편차

ϵ : 잔차

여기서, $m(l)$ 를 추정하기 위한 다양한 방법이 사용되고 있다. 이들 방법은 기본적 전제조건과 가중치를 산출하는 방법에 따라 SK, OK, UK, CK로 나누어진다. SK는 단지 예측오차를 최소로 하는 가중치를 산출하여 예측값을 추정한다. OK는 자료 분포의 등질성을 전제로 편기되지 않은 추정식을 산출하고 오차분산을 최소화하면서 가중치를 산출하여 예측값을 추정한다.

다. UK는 공간적으로 변화하는 평균을 제거하지 않고 가중치를 산출하여 예측값을 추정한다. 이러한 공간적 이질성을 반영하기 위하여 본 연구에서는 1차 경향 모형을 적용하여 내삽하였다. CK는 두 개 이상의 변수를 이용하면서 통상 OK나 UK를 기반으로 가중치를 산출하여 예측값을 추정한다. 본 연구에서는 지가의 공간적 이질성을 고려할 수 있는 UK를 기반으로 가중치를 산출하여 CK를 수행하였다. 용도지역과 도로조건이 보조설명변수로 사용되었다. 일반적으로, 크리깅 모형은 지가 추정에 있어서 공간적 의존성을 다루기 위하여 베리오그램(variogram)이나 공분산을 활용한다. 선행연구에서 지가함수로 음지수모형을 주로 사용하였기 때문에(Kim and Lee, 2013), 본 연구에서는 지수형 반베리오그램 모형을 적용하여 크리깅 내삽을 실시하였다. 이외에, 크리깅 기반 공간내삽법에서 모든 파라미터는 표준화를 위해 기본값을 선정하여 내삽을 실시하였다.

회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법의 정확성은 교차검증을 통해 통계적으로 그리고 시각적으로 평가되었다. 먼저, 전역적 정확성 지수인 평균 제곱근 오차(RMSE: Root Mean Square Error), 수정된 평균 제곱근 오차(adjusted RMSE), 분산지수(COD: Coefficient of Distribution)를 이용하여 그 정확성을 통계적으로 평가하였다. RMSE는 식 6을 이용하여 산출하였다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - a_i)^2} \quad (6)$$

n : 검사점 표준지의 수

z_i : 추정 지가

a_i : 실제 지가

수정 RMSE는 식 7에 의하여 정량화하였다. RMSE와 수정 RMSE의 차이는 실제 지가에 의한 표준화의 여부이다. 즉, 수정 RMSE는 추정 지가와 실제 지가의 편차를 실제 지가로 표준화하여

RMSE가 큰 표준지의 영향을 최소화한 것이다.

$$\text{수정 RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{z_i - a_i}{a_i}\right)^2} \quad (7)$$

COD는 식 8을 이용하여 계산하였다. COD는 추정 지가에 대한 실제 지가의 예측비율의 편차를 계산한 것으로 COD가 작을수록 그 정확성이 상대적으로 뛰어나다고 판단할 수 있다. 흔히, COD가 20~25 이하이면, 과세표준을 산정할 때 수용할 수 있다(Lee *et al.*, 2014; Sohn and Lee, 2016).

$$\text{COD} = \frac{\sum_{i=1}^n |PR_i - \text{median}(PR_i)|}{\text{median}(PR_i)} \quad (8)$$

PR_i : 예측비율(추정 지가/실제 지가)

$\text{median}(PR_i)$: 예측비율의 중위수

다음으로, 3차원 잔차도(three-dimensional residual error map)와 산점도(scatterplot)를 이용하여 그 정확성을 시각적으로 서로 비교하였다. RMSE와 수정 RMSE는 절대적 오차(absolute error)를 단일지수로 보여 주는 전역적 정확성 지수이다. 전역적 정확성 지수는 오차가 공간상에서 어떻게 변화하는지를 시각화해 줄 수 없는 단점이 있다. 이와는 달리, 상대적 오차(RE: Relative Error)는 오차가 공간상에서 집중되는지 또는 분산되는지를 시각화해 줄 수 있다(Yang and Hodler, 2000). RE는 검사점으로 사용된 각 표준지의 잔차를 절대값으로

변환하여 산출하였다(식 9). 따라서, 각 표준지의 RE는 국지적 및 상대적 정확성 지수로 간주할 수 있다.

$$RE = |z_i - a_i| \quad (9)$$

효율적인 시각화를 위해 크리깅 기법을 이용하여 RE에 대한 공간내삽면을 생성하고 이를 3차원 잔차도로 나타내었다. 마지막으로, 8개의 공간내삽법의 정확성을 산점도를 통해 시각화하여 상대적으로 평가하였다.

결과 및 고찰

1. 통계적 정확성 평가

표 1은 8개의 공간내삽법의 정확성을 통계적으로 평가한 결과를 제시하고 있다. 먼저, 회귀모형 기반 공간내삽법에 대한 R^2 와 AIC (Akaike Information Criterion)를 통해 모형의 적합도를 평가하였다. R^2 와 AIC에 의하면, 지가의 공간적 이질성을 고려할 수 있는 GWR 기반 헤도닉 회귀모형이 지가 추정을 위한 가장 적합한 방법인 것으로 나타났다. 이러한 결과는 기존의 일부 연구(Song and Park, 2012; Chrostek and Kopczewska, 2013; Chun, 2016)와도 일치한다. 이와 달리, 크리깅 기반 공간내삽법은 모형의 적합도에 관한 통계적 결과를 제시하지 못하는 단점이 있었다.

그림 4는 285개의 검사점 표준지에 대한 국지적 R^2 을 나타낸다. 국지적 R^2 가 사례지역에 걸쳐 상이하게 분포하는 것을 알 수 있다. 특히, 지하철 1호선 송현역 및 월촌역 부근, 중부내륙고속도로 남대구 IC 및 성서 IC 부근, 지하철 2호선

TABLE 1. Statistical results for accuracy assessment

	OLS	SLM	SEM	GWR	SK	OK	UK	CK
R^2	0.63	0.64	0.67	0.79	-	-	-	-
AIC	30993	30976	30902	30529	-	-	-	-
RMSE	309979	307575	298255	262651	454482	445106	439687	461553
Adj. RMSE	36.77	43.25	18.24	10.5	56.81	20.66	19.71	29.19
COD	3.84	5.04	3.52	1.55	5.54	2.30	2.27	3.39

내당역, 두류역, 감삼역, 죽전역, 용산역 부근에서 국지적 R^2 가 높게 나타났다. 이는 대구광역시 달서구에서 지가의 공간적 이질성이 강하게 존재하고 있다는 것을 직접적으로 증명해 준다.

전역적 정확성 지수인 RMSE 및 수정 RMSE에 따르면, 8개의 공간내삽법 중에서 지가의 공간적 이질성을 반영하는 GWR 기반 헤도닉 회귀모형이 가장 정확한 방법인 것으로 나타났다. 또한, 회귀모형 기반 공간내삽법이 크리깅 기반 공간내삽법 보다 상대적으로 정확한 것으로 나타났다. 기존의 일부 연구는 이론과 달리 크리깅 기법이 다른 공간내삽법 보다 실제로 정확하지 않다고 주장한다(Yang and Hodler, 2000; Kim *et al.*, 2013). 크리깅 기반 공간 내삽법 중에서는 UK가 상대적으로 정확한 방법이었다. 이것은 CK가 상대적으로 정확하다는 기존의 연구

결과(Chica-Olmo, 2007; Montero and Larraz, 2011; Kuntz and Helbich, 2014; Choi and Kim, 2015)와 상충되는 것이다. 이러한 결과는 사례지역에서 GWR와 같이 지가의 공간적 이질성을 고려할 수 있는 모형이 상대적으로 정확한 방법으로 확인된 것과 일맥상통한다. 왜냐하면, 크리깅 기법 중에서 상대적으로 자료의 공간적 이질성을 반영하여 추정할 수 있는 기법이 UK이기 때문이다. 달서구에서 지가의 공간적 의존성을 설명할 수 있는 SLM 및 SEM 그리고 크리깅 기반 공간내삽법 보다 지가의 공간적 이질성을 고려할 수 있는 GWR이 상대적으로 정확한 방법으로 확인되었다. 이는 사례지역에서 지가의 공간적 의존성 보다 공간적 이질성이 강하게 나타나기 때문이다. 실제로, 종속변수인 지가에 대한 Moran의 I가 0.1032로 나타났고, 그

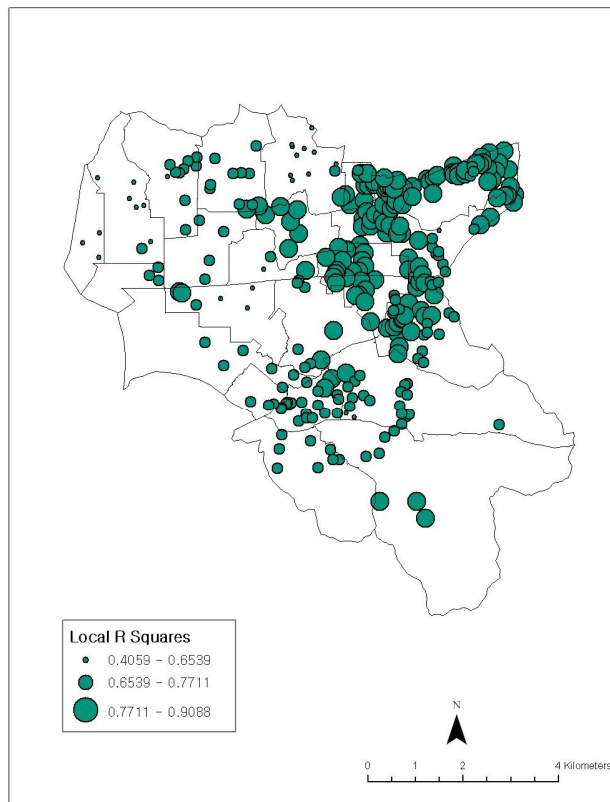


FIGURE 4. Local R square values for check points

잔차에 대한 Moran의 I는 0.0974로 나타나 공간적 자기상관, 즉 공간적 의존성이 미약한 것으로 나타났다.

8개의 공간내삽법의 COD는 모두 20~25이하로 나타나 정확한 것으로 판단할 수 있었으나, 그 중에서 상대적으로 가장 정확한 방법은 역시 GWR이었다. 전역적 정확성 지수인 RMSE, 수정 RMSE, COD와 모형의 적합도 지수인 R^2 와 AIC에 의한 통계적 평가 결과를 종합하면, 8개의 공간내삽법 중에서 GWR이 가장 적합하고 정확한 공간내삽법으로 밝혀졌다.

2. 시각적 정확성 평가

지가 추정을 위한 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법의 정확성을 3차원 잔차도와 산점도를 통해서 시각적으로 평가하였다. 그림 5는 8개의 공간내삽법의 정확성을 3차원 잔차도를 통해 나타낸 것이다. 국지적 및 상대적 정확성 지수인 RE가 작을수록 상대적으로 정확하다는 것을 의미한다. 만약 공간내삽법의 정확성이 동일하다면, RE가 사례지역에서 균일하게 분포하는 것이 상대적으로 정확하다고 할 수 있다. 대체로, RE가 큰 지역(지하철 1호선 상인역 인근, 지하철 2호선 두류역 인근, 대덕산 부근)은 그림 3에서 제시된 것처럼 자료점 표준지가 충분히 제공되지 않은 지역임을 확인할 수 있었다. 이는 지가 추정을 위한 공간내삽법의 정확성을 제고하기 위해서는 많은 수의 표준지 공시지가가 넓은 지역에 걸쳐 제공되어야 한다는 것을 시사한다. GWR을 위한 3차원 잔차도에서 전반적으로 RE가 작고 넓게 분포하고 있어서 상대적으로 가장 신뢰할 만한 방법이 GWR임을 재확인할 수 있었다. 하지만, 크리깅 기반 공간내삽법 간에 시각적으로 뚜렷한 차이는 없었다.

그림 6은 8개의 공간내삽법의 정확성을 산점도를 통해 나타낸 것이다. 점들이 회귀선을 중심으로 근접하면 상대적으로 정확하다고 판단할 수 있다. GWR 기반 헤도닉 회귀모형은 R^2 가 0.743으로 가장 높았고 점들이 회귀선을 중심으로 가장 근접한 것으로 나타났다. 다시 말해

서, GWR이 8개의 공간내삽법 중에서 가장 정확한 것으로 확인되었다. 크리깅 기반 공간내삽법은 R^2 가 회귀모형 기반 공간내삽법 보다 매우 낮았고 점들이 회귀선을 중심으로 흩어져 있어 그 정확성이 낮은 것으로 나타났다. 사례지역에서 크리깅 기반 공간내삽법은 지가의 시각화에는 용이하였지만, 지가 추정에 있어서 그 정확성은 상대적으로 떨어지는 것으로 확인되었다.

결론

본 연구는 대구시 달서구를 사례로 지가 추정을 위한 회귀모형 기반 공간내삽법(OLS, SLM, SEM, GWR)과 크리깅 기반 공간내삽법(SK, OK, UK, CK)의 정확성을 통계적으로 그리고 시각적으로 평가하였다. 그 분석결과, 사례지역에서 공간적 이질성을 고려할 수 있는 GWR 기반 헤도닉 회귀모형이 공간적 의존성을 반영할 수 있는 SLM 및 SEM과 크리깅 기반 공간내삽법들 보다 지가를 추정하는데 있어서 상대적으로 정확한 방법인 것으로 판명되었다. 이론적 견고성과 달리, 크리깅 기반 공간내삽법은 지가를 쉽게 시각화할 수 있게는 했지만, 지가를 추정하는데 있어서 그 정확성은 GWR 보다 훨씬 떨어지는 것으로 확인되었다. 따라서, 지가 추정을 위해 크리깅 기반 공간내삽법을 적용할 경우, 그 기본적 전제조건을 면밀히 검토할 필요가 있다. 또한, 지가 추정을 위한 회귀모형 기반 공간내삽법과 크리깅 기반 공간내삽법의 정확성은 사용된 표본의 수와 그 공간적 분포에 의해 많이 영향을 받는다는 것을 밝혀냈다.

하지만, 본 연구는 몇 가지 방법론적 이슈를 고려할 필요가 있다. 먼저, 본 연구에서 가장 정확한 공간내삽법으로 확인된 GWR은 다른 회귀모형 기반 공간내삽법 처럼 일부 검사점에서 음의 지가 추정치를 산출하였다. 반면에, UK를 제외한 나머지 크리깅 기반 공간내삽법은 모든 검사점에서 양의 지가 추정치를 산출하였다. GWR 기반 헤도닉 회귀모형을 지가산정 업무에 활용하기 위해서는 이러한 음의 지가 추정치를

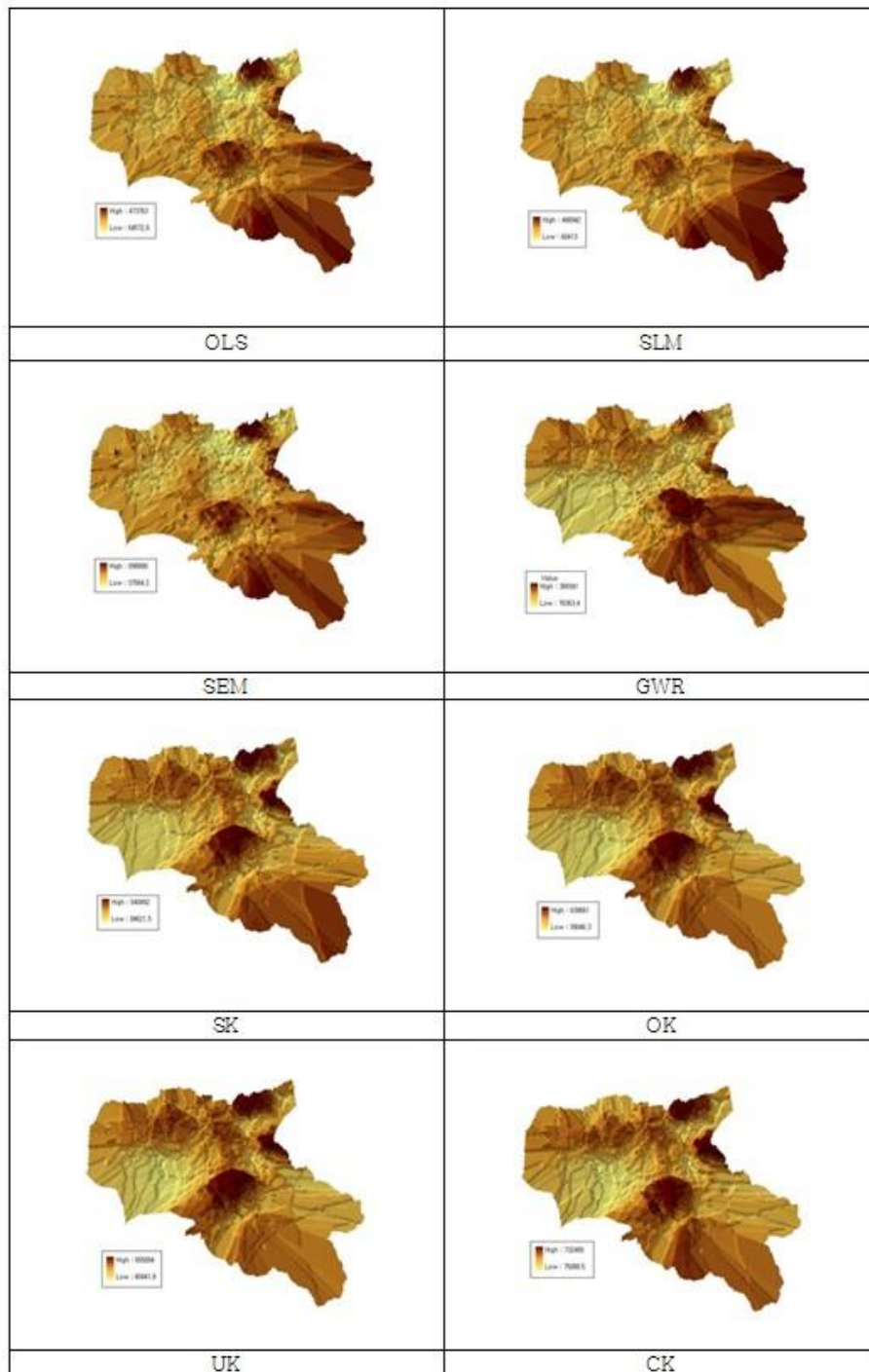


FIGURE 5. Three-dimensional residual error maps

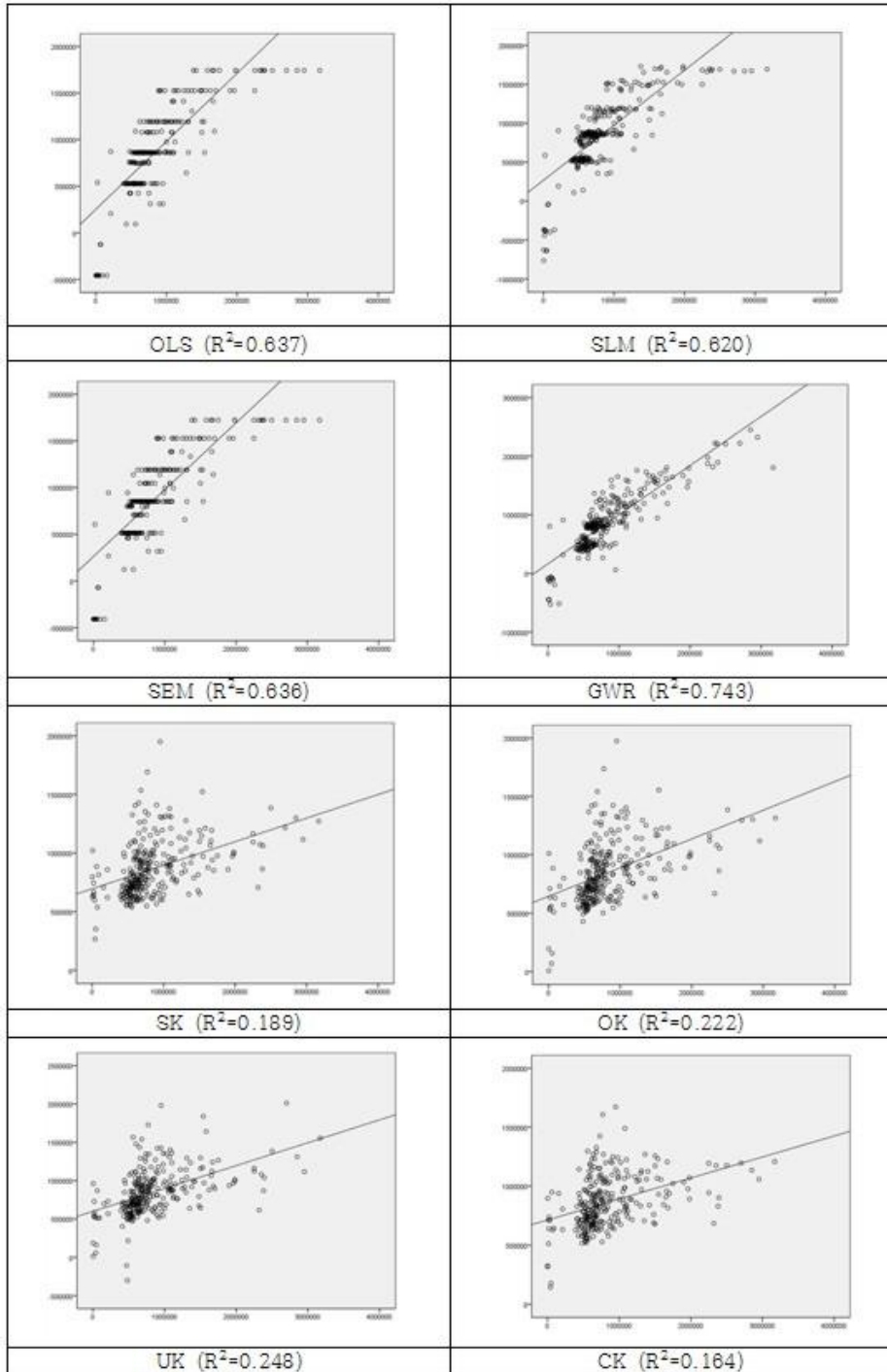


FIGURE 6. Scatterplots
(x-axis: actual land price, y-axis: estimated land price)

보완하기 위한 방법을 모색해야 한다. 다음으로, 표준지의 위치표현방법인 지오코딩(geocoding)이 공간내삽법에 의한 지가 추정의 정확성에 미치는 영향을 분석할 필요가 있다. 현재 우리나라는 지번 주소와 도로명 주소를 사용하고 있다. 표준지에 대한 지오코딩의 차이에 따라 지가에 대한 공간내삽의 결과가 달라질 수 있다. 마지막으로, 향후에 지가 추정을 위한 GWR, BK, RK 그리고 Yoo and Kyriakidis(2009)에 의해 제안된 A2PKED(Area-to-Point Kriging with External Drift)의 정확성을 상대적으로 비교·평가해 볼 필요가 있다.

본 연구는 지가 추정의 정확성을 향상시키기 위해 적절한 공간내삽법, 표본의 크기 및 분포, 공간적 의존성과 공간적 이질성과 같은 공간효과를 주의 깊게 탐색할 필요가 있다는 시사점을 제시한다. 본 연구의 결과가 향후에 지가를 통해 도시의 공간구조를 분석하는 이차적 연구에 기여할 것으로 기대된다. **KAGIS**

REFERENCES

- Chica-Olmo, J. 2007. Prediction of housing location price by a multivariate spatial method: cokriging. *Journal of Real Estate Research* 29:91-114.
- Choi, J.H. and B.J. Kim. 2015. A study for applicability of cokriging techniques for estimating the real transaction price of land. *Journal of the Korean Society for Geospatial Information Science* 23(1):55-63 (최진호, 김봉준. 2015. 토지 실거래가격 추정을 위한 공동 크리깅기법의 적용가능성 연구. *한국지형공간정보학회지* 23(1):55-63).
- Choi, Y. and B.H. Lee. 2006. A study on the estimation of land price considering characteristic of the adjacent land use and spatial autocorrelation in residential zone. *Journal of Korea Planning Association* 41(5):45-60 (최열, 이백호. 2006. 공간자기상관과 주변 용도지역에서 접근성을 고려한 주거지 내 지가추정에 관한 연구. *국토계획* 41(5):45-60).
- Chrostek, K. and K. Kopczewska. 2013. Spatial prediction models for real estate market analysis. *Ekonomia* 35:25-43.
- Chun, H.J. 2016. A study on estimation of housing price using GWR model and GIS technology. *Residential Environment* 14(1):1-12 (전해정. 2016. GWR모형과 GIS를 이용한 주택가격 추정에 관한 연구, *주거환경* 14(1):1-12).
- Jin, C.W. and G.H. Lee. 2014. Spatial hedonic modeling using geographically weighted LASSO model. *Journal of the Korean Geographical Society* 49(6):917-934 (진찬우, 이진학. 2014. GWL을 적용한 공간 헤도닉 모델. *대한지리학회지* 49(6):917-934).
- Kim, G.J., B.K. Lee, and B.U. Park. 2013. A comparative analysis on spatial interpolation techniques for price estimation of housing facilities. *Journal of the Korean Association of Professional Geographers* 47(2):119-127 (김길중, 이병길, 박병욱. 2013. 거주시설 가격 추정을 위한 공간보간법의 비교분석. *국토지리학회지* 47(2):119-127).
- Kim, G.Y., C.S. Chung, J.H. Ryu, and J.H. Kim. 2005. A study on the correlativity between land price and residential environmental elements. *Journal of the Korean Association of Geographic Information Studies* 8(4):52-60 (김가야, 정창식, 류재홍, 김정호. 2005. 지가와 주거환경 요소의 상관성에 관한 연구. *한국지리정보학회지* 8(4):52-60).
- Kim, H.K. 2006. Changes of spatial structure in Busan metropolitan using GIS - with special reference to population, employment and land price. *Journal of*

- the Korean Association of Geographic Information Studies 9(4):204-214 (김홍관. 2006. GIS를 활용한 부산시 공간구조변천에 관한 연구 - 인구, 고용, 지가분포를 중심으로. 한국지리정보학회지 9(4):204-214).
- Kim, H.Y. and C.M. Jun. 2012. Land value analysis using space syntax and GWR. *Journal of the Korean Association of Geographic Information Studies* 15(2):35-45 (김혜영, 전철민. 2012. 공간구문론 및 지리적 가중회귀 기법을 이용한 지가분석. 한국지리정보학회지 15(2):35-45).
- Kim, K.Y. 2012. Exploring spatio-temporal variations of land price in Daegu metropolitan city. *Journal of the Korean Association of Regional Geographers* 18(4):414-432 (김감영. 2012. 대구시 지가의 시공간적 변화 탐색. 한국지역지리학회지 18(4):414-432).
- Kim, K.Y. and Y.M. Lee. 2013. Exploring the urban spatial structure of Korean cities using a land value function. *Journal of the Korean Urban Geographical Society* 16(1):99-112 (김감영, 이윤미. 2013. 지가함수를 이용한 우리나라 도시 공간구조 탐색. 한국도시지리학회지 16(1):99-112).
- Kim, Y.S. and B.R. Kim. 2004. The issues of assessment valuation on buildings in Korea. *Real Estate Policy Research* 5(1):21-37 (김양수, 김병량. 2004. 현행 건물과세 평가상의 논점. 부동산정책연구 5(1):21-37).
- Kuntz, M. and M. Helbich. 2014. Geostatistical mapping of real estate prices: an empirical comparison of kriging and cokriging. *International Journal of Geographical Information Science* 28(9):1904-1921.
- Lee, C.R. and K.H. Park. 2013. The effects of neighborhood segmentation on the adequacy of a spatial regression model. *Journal of the Korean Geographical Society* 48(6):978-993 (이창로, 박기호. 2013. 인근지역 범위 설정이 공간회귀모형 적합에 미치는 영향. 대한지리학회지 48(6):978-993).
- Lee, C.R., Y.S. Eum, and K.H. Park. 2014. A bayesian estimation of price for commercial property: using subjective priors and a kriging technique. *Journal of the Korean Geographical Society* 49(5):761-778 (이창로, 엄영섭, 박기호. 2013. 상업용 토지 가격의 베이시안 추정: 주관적 사전지식과 크리깅 기법의 활용을 중심으로. 대한지리학회지 49(5):761-778).
- Militino, A.F., M.D. Ugatte, and L. Garcia-Reinaldos. 2004. Alternative models for describing spatial dependence among dwelling selling prices. *Journal of Real Estate Finance and Economics* 29(2):193-209.
- Montero, J. and B. Larraz. 2011. Interpolation methods for geographical data: housing and commercial establishment markets. *Journal of Real Estate Research* 33:233-244.
- Park, H.S., S.Y. Jyoung, and T.U. Rho. 2003. A spatial econometric analysis of apartment sale prices and spatial effects. *Journal of Korea Planning Association* 38(5):115-125 (박현수, 정수연, 노태욱. 2003. 공간계량경제모형을 이용한 아파트가격과 공간효과분석. 국토계획 38(5):115-125).
- Park, J.U. and D.H. Koo. 2011. Trends and issues of land values studies in Korea: focusing on geography and related disciplines. *Journal of the Korean Association of Professional Geographers* 45(3):409-426 (박지은, 구동희. 2011. 우리나라 지

- 가연구의 동향과 과제: 지리학 및 인접분야를 중심으로. 국토지리학회지 45(3):409-426).
- Seo, K.C. 2003. The measurement of spatial variation by nonparametric hedonic method in the estimation of real estate price: the case of western Busan metropolitan city. Journal of the Korean Regional Development Association 15(2):183-203 (서경천, 2003. 부동산 가격 추정에 있어서 비모수적 헤도닉 방법을 도입한 공간적 변동성 측정: 부산광역시 서부산권을 중심으로. 한국지역개발학회지 15(2):183-203).
- Seo, K.C. and S.H. Lee. 2001. The efficient estimation of land value using spatial autoregressive model and land market segmentations. Journal of Korea Planning Association 36(4):77-94 (서경천, 이성호. 2001. 공간적 자기회귀모델과 토지시장 분할에 의한 효율적 지가추정에 관한 연구. 국토계획 36(4):77-94).
- Sohn, H.G. and C.R. Lee. 2016. Estimating farmland prices using a regression-kriging model. The Korea Spatial Planning Review 89(2):39-53 (손학기, 이창로. 2016. 회귀-크리깅 모형을 활용한 농지가격의 추정. 국토연구 89(2):39-53).
- Song, Y.C. and H.S. Park. 2012. A study on the estimation of farmland price using spatial econometrics approach: focused on urban fringe in Seoul metropolitan area. The Korea Spatial Planning Review 72(1):121-140 (송용철, 박헌수. 2012. 공간계량경제 접근방법을 이용한 농지가격추정에 관한 연구: 수도권 근교농지를 중심으로. 국토연구 72(1):121-140).
- Yang, X. and T. Hodler. 2000. Visual and statistical comparisons of surface modeling techniques for point-based environmental data. Cartography and Geographic Information Science 27(2):165-175.
- Yoo, E.H. and P.C. Kyriakidis. 2009. Area-to-point Kriging in spatial hedonic pricing models. Journal of Geographical Systems 11:381-406. **KAGIS**