

Spatial analysis for a real transaction price of land

Jihye Choi^a · Hyang Gon Jin^b · Yongku Kim^{b,1}

^aKorea Appraisal Board; ^bDepartment of Statistics, Kyungpook National University

(Received January 16, 2018; Revised March 13, 2018; Accepted March 15, 2018)

Abstract

Since the real estate reporting system was first introduced, about 2 million real estate transaction per year have been reported over the last 10 years with an increasing demand for real estate price estimates. This study looks at the applicability and superiority of the regression-kriging method to derive effective real transaction prices estimation on the location where information about real transaction is unavailable. Several issues on predicting the real estate price are discussed and illustrated using the real transaction reports of Jinju, Gyeongsangnam-do. Results have been compared with a simple regression model in terms of the mean absolute error and root square error. It turns out that the regression-kriging model provides a more effective estimation of land price compared to the simple regression model. The regression-kriging method adequately reflects the spatial structure of the term that is not explained by other characteristic variables.

Keywords: disclosure price, real transaction price, spatial regression analysis, spatial prediction

1. 서론

2006년 1월 1일 부동산 투기근절, 공평과세 목적으로 부동산 실거래 신고제도가 전격 도입된 지 십여 년이 흐른 지금, 정부에서 운영 중인 부동산실거래관리시스템(Real Estate Transaction Management System; RTMS)에는 연간 약 200만 건의 부동산 실거래 신고자료가 축적되고 있다. 인터넷이 발달하고 정보에 대한 접근성이 높아진 요즘, 부동산 투자에 대한 관심 증가로 부동산 가격정보에 대한 요구도 나날이 증가하고 있으며 현재 국토교통부 부동산실거래가 공개시스템 및 공공데이터포털(data.go.kr) 등을 통해 일부 실거래정보를 공유 및 개방하고 있다. 하지만 이는 단순히 거래사례에 대한 정보만을 제공할 뿐이라 공동주택 실거래의 경우 동, 호수, 토지건물 실거래의 경우 지번을 개인정보보호 등의 이유로 공개하고 있지 않아 실거래의 위치별 정확한 데이터를 구득하기 어려운 실정이라서 정보의 비대칭성이 여전히 존재하고 이러한 부동산 정보의 특수성이 부동산시장에서의 투기가 근절되지 않는 이유 중 하나이다. 또한 부동산 가격정보를 얻더라도 이는 일부 거래된 사례에 대한 가격일 뿐 실거래가 발생하지 않은 지점에 대한 가격 정보를 주지 못하는 한계를 지닌다.

국내의 많은 연구들이 부동산 가격형성요인 및 가격산정 모형에 대해서 분석하였으나 대부분 연구가 분석대상을 공시지가를 대상으로 하였으며 분석지역 역시 대도시에 집중되어 있다. 실거래가격을 활용한

This work is based on part of Jihye Choi's Master thesis.

¹Corresponding author: Department of Statistics, Kyungpook National University, 80 Daehak-Ro, Daegu 41566, Korea. E-mail: kim.1252@knu.ac.kr

연구 역시 주로 가격자료가 풍부하고 분석이 용이한 대도시의 공동주택을 위주로 이루어져 왔다. 하지만 관련법에 의거 공시지가는 적정가격을 기준으로 공시하도록 하고 있으나 지역별 적정가격 포착이 어렵고 가격평가 담당자의 자의성 개입 가능성 등 조사자 편이가 존재하며 과세기준으로 소유자 등으로부터의 민원 등의 사유로 실거래가 수준에 현저히 미치지 못하고 있어 적정 시세가격의 추정에는 적합하지 않다. 또한 사례가 비교적 풍부한 대도시 공동주택의 경우 축적된 데이터를 활용한 실거래가격 추정이 비교적 용이하나 지방도시 내 실거래가 드문드문 발생하는 토지 등의 실거래가격 추정에 대한 연구는 부족한 실정이다.

본 연구의 목적은 축적된 실거래 신고가격 데이터를 활용하여 실거래 미발생 지점에 대한 시세가격 추정 모형을 도출하는 것으로, 부동산 가격이 지리적 위치에 따라 결정되는 특수성을 가지는 것을 고려하여 공간구조가 반영될 수 있도록 공간회귀 모형을 통한 추정 토지 시세가격의 정확도를 살펴보았다. 거래사례가 많지 않은 지역의 토지 시세가격 추정모형 및 효과 분석을 위하여 지방중소도시인 경상남도 진주시를 대상지역으로 선정하여 분석하되, 토지가격수준이 현저히 상이한 용도지역을 도시지역(주거/상업/공업/녹지지역)과 비도시지역(관리/농림/자연환경보전지역)으로 구분, 분석하여 실거래 미발생 지점의 시세가격을 추정하였다. 특히 변수 간 선형성을 기반으로 한 회귀모형과 공간구조를 반영한 공간회귀 모형을 비교분석함으로써 토지의 가격이 동일 특성임에도 공간구조(위치정보)에 따라 달라진다는 것을 보여주었다. 또한 토지 시세가격 추정 모형을 실거래기반 가격공시제도 도입 시 활용하거나 실거래 신고가격 진단기준으로서의 활용방안을 제시하였다.

2. 연구의 범위와 방법

본 연구의 공간적 범위는 경상남도 진주시를 대상으로 선정하였다. 진주시는 전형적인 도농복합도시로 속성정보(용도지역, 이용상황 등)에 따라 부동산 가격이 형성되나 지역성숙도 등 위치정보에 따라 가격이 달리 형성되는 지역으로, 공간구조가 잘 드러나는 지역 특성을 지녀 본 연구의 분석지역으로 선정하였다. 분석에 사용한 자료는 2014년부터 2016년 10월까지 실거래 신고된 토지 거래가격을 활용하였다.

연구의 수행을 위하여 크게 다섯 단계로 나누어 진행하였다. 첫째, 시세가격 수준 포착에 대한 사회적 요구 증가와 그에 따른 실거래 신고제도의 도입배경을 살펴본다. 더불어 전통적인 토지가격 산정방법에 대해 알아보고, 부동산 가격형성요인과 부동산 가격에 대한 관계를 연구한 문헌을 검토하여 토지 시세가격 산정에 관한 실질적인 연구의 범위를 결정하였다. 둘째, 선행연구를 바탕으로 토지 실거래 가격과 가격을 결정짓는 가격형성요인을 분석하기 위한 주요 변수를 설정하였다. 실거래 가격과 여러 요인 간의 관계를 살펴본 후 실거래가격(추정 시세가격)을 종속변수로 설정하고, 토지 실거래가격에 영향을 미치는 다양한 요인을 독립변수로 설정하였다. 셋째, 모형 도출에 앞서 토지 가격 수준과 가격형성요인이 현저히 상이하며 지역적으로 구분된 도시지역과 비도시지역을 나눈 후, 각각의 지역에 대하여 전통적인 회귀분석을 실시하여 토지 시세가격 추정에 대한 적합 모형을 찾았다. 넷째, 회귀모형의 잔차 크리깅 분석을 통해 공간회귀 모형을 찾고, 회귀모형과 공간회귀 모형 결과를 비교분석하여 공간회귀 모형의 예측능력을 검증하였다. 마지막으로 공시지와 추정 시세가격을 실거래가격과 비교분석한 결과를 토대로 향후 공시가격의 균형성 제고방안을 모색하고, 실거래 신고가격의 적정성 진단기준으로서의 활용가능성을 검토해 보았다.

3. 이론적 배경 및 선행연구검토

3.1. 이론적 배경

부동산 실거래가격이란 특정부동산에 대한 당사자 간의 합의에 따라 실제 지불된 금액을 의미하는 것으로

로, 매도자와 매수자간에 목적 부동산에 대해 상호간 사실상 합의된 금액 그 자체를 의미한다. 따라서 부동산 실거래가격은 시세보다 현저히 낮거나 높아 객관적 입장에서 타당성이 결여될 수 있으며 이런 경우라도 계약자유의 원칙에 의해 실거래가격으로 인정되나 (Jang, 2008), 이러한 가격을 인근의 보편적인 시세가격으로 인정할 수는 없을 것이다. 앞서 살펴본 실거래가격의 특수성 때문에 2005년까지 정확한 실거래가격이 아닌 공시가격으로 신고 가능하였고, 이중계약서 작성 등이 거래관행처럼 여겨져 실제 부동산의 거래가격과는 괴리된 채 양도소득세 등이 과세되어 왔었다. 하지만 부동산 가격이 급등하고 부동산 세금이 실제 가치보다 낮게 부과되어 부동산 투기를 부추기고 탈세의 온상으로 지목되는 등 사회적 문제가 심화되자, 부동산 거래 투명성 제고, 실거래가격에 의한 정상과세 목적 달성, 세부절차 전자화를 통한 민원 간소화, 행정 효율화 향상을 주요 골자로 한 부동산 실거래가격 신고의무제도가 2006년 1월 1일 전격 도입되었다.

관련법에 의거 신고된 부동산 실거래가격은 허위 신고 여부 등에 대해 한국감정원의 가격 적정성 검증을 거치게 되고, 거래내역 및 검증결과는 국세청 및 관할 시군구청에 통보되어 과세자료로 활용된다. 이러한 신고가격의 검증은 실거래가격의 적정성을 판단하는 기준 설정이 무엇보다 중요할 것이나, 적정 실거래가격 수준을 판단하는 기준을 설정하기 어렵고 정확한 진단을 위해서는 비용이 과다하게 드는 등의 사유로 인하여 현재 공시가격을 기준으로 일정비용의 범위를 적정 판단기준으로 활용하는 진단모형을 적용하여 평가하고 있다. 하지만 공시가격은 과세를 목적으로 평가 또는 산정된 가격으로서 실거래가격과 괴리되어 있으며 공시가격의 불형평성이 존재하는 바 (Yoo 등, 2014), 실제 실거래가격의 적정 수준을 진단하기 어려워 실거래가격 진단기준으로 한계를 가지므로, 적정 시세가격 추정 방법 및 진단기준으로서의 활용방안에 대한 별도의 논의가 필요하다.

보통 토지의 시세가격은 감정평가를 통하여 찾을 수 있는데, 감정평가는 토지 등의 경제적 가치를 판정하여 그 결과를 가액(價額)으로 표시하는 것을 말한다. 일반적인 감정평가는 토지 등이 통상적인 시장에서 충분한 기간 동안 거래를 위하여 공개된 후 그 대상물건의 내용에 정통한 당사자 사이에 신중하고 자발적인 거래가 있을 경우 성립될 가능성이 가장 높다고 인정되는 대상물건의 가액인 시장가치를 찾는 과정이다.

감정평가에서 시장가치는 가치의 3원성인 비용성, 수익성 및 시장성에 의하여 도출되며, 투입비용에 따른 가치산정방법인 비용접근법, 해당 부동산으로 얻을 수 있는 수익성에 따른 가치산정방법인 수익접근법, 유사한 부동산이 시장에서 거래되는 가격에 근거하여 가치를 산정하는 시장접근법, 총 3가지 방법에 의해 찾을 수 있다. 부동산의 가치를 산출하는데 있어 부동산 운영으로 인한 수익이 발생하는 경우 투자수익률에 근거한 부동산 가격이 형성될 것이므로 수익접근법이 우수할 것이나, 임야 및 휴경지 등 별도의 수익이 발생하지 않는 토지의 경우는 이를 적용하기 곤란하다. 또한 비용접근법은 투입 원가(건물 등)에 근거하여 부동산의 가격을 산출하는 방법으로, 토지는 위치불변인 자연의 산물로서 더 이상 증가될 수 없고 이전하여 생산할 수 없는 고유한 특성인 부증성 및 고정성이 있으므로 토지의 가격이란 비용접근법에 의한 추정하는 것이 제한적이다. 반면, 시장접근법은 대상물건의 시장성에 따라 평가하는 방법으로, 부동산의 가격은 인근지역 내에 대상과 유사한 부동산의 과거 거래된 가격과 유사하게 형성된다는 보편원칙에 근거하고 있다. 즉, 유사한 최신의 거래사례와 대상물건과의 차이(규모, 위치, 특성 등)를 품등비교하여 적정시세를 산출하는 방법으로, 거래사례라는 시장증거에 의해 가격을 산출하므로 설득력이 높은 방식이나, 거래가 드물거나 직접적인 품등비교가 불가능한 경우에는 적용하기 어려운 방법이다.

본 연구는 통계적 모형에 의해 토지 시세가격을 추정하는 방법에 대한 것으로, 실거래 가격에 기반하고 있다는 점에서 시장접근법에 의한 방식이라고 할 수 있으며 모형에 의하여 시세가격이 산정되기 때문에 거래사례와의 비교 시 자의성이 배제되고, 거래가 드문 지역이라고 하더라도 모형의 설명변수로 실거래

미관측 지점에 대한 가격을 도출할 수 있다는 점에서 기존 시장접근법의 단점을 획기적으로 개선할 수 있을 것으로 기대된다. 다만 사례가 현저히 부족하거나 모형으로 설명하지 못하는 부분은 추정값에 반영되지 못하는 한계를 가지는 바, 통계적 모형이 시세가격 추정방법으로서 정책 등 실무에서 활용되기 위하여 향후 개선해야 할 과제다.

3.2. 선행연구 검토

토지가격 산정방식에 대한 연구의 상당수가 감정평가를 통한 가격 추정에 대한 것이었고, 실거래신고제도가 도입된 2006년 이후 실거래 데이터 활용방안으로 통계적 모형에 의한 가격추정이 본격 논의되기 시작하였으나 부동산 특성의 헤도닉모형에 의한 토지가격 추정이 대부분이었다. 현행법 상 개별공시지가는 표준지공시지가 평가금액에 기반한 토지가격비준표(헤도닉 모형)를 사용하여 대량 산정하도록 기준하고 있다. 하지만 헤도닉 모형에 의한 부동산 가격 산정은 간편하나, 토지특성 항목 설명력 부족 및 다중공선성 문제, 회귀모형 및 더미변수의 한계, 부동산 가격의 대표적 특성인 공간구조가 반영되지 못한 채 가격이 산정되는 등 실제 가격과 괴리되는 한계를 지닌다 (Seo와 Kwak, 2014).

Sohn 등 (2003)은 VAR모형을 활용하여 부동산 시장여건변화를 반영한 부동산가격 예측모형을 구축하여 IMF 전후기간에 대한 충격반응 및 분산분해 분석을 통해 시장기본가치의 중요성이 얼마나 높아졌는지를 검증하여 ARIMA모형에 대한 우수성을 보여주었고, Ku와 Shin (2015)은 헤도닉 가격 모형을 분석하되, 분산팽창지수 및 능형회귀분석을 이용해 다중공선성을 파악하고 유의한 설명변수를 선별하는 데에 활용하여 더 정확하고 적절한 헤도닉 가격 모형을 구축하는 결과를 보여주었다. Lee 등 (2014)은 상업용 토지의 가격을 정확히 추정하고자 크리깅 방법을 적용함으로써 지가의 공간적 상관성을 명시적으로 고려하였고, 자료부족의 한계를 극복하기 위하여 전문가 지식을 사전 확률분포의 형태로 모형에 반영할 수 있는 베이저안 크리깅 방법을 활용하여 거래 자료가 희소한 상황에서의 부동산 가격추정 가능성을 확인하였다. 특히 크리깅을 통한 가격산정과 관련된 선행연구로, Sohn과 Lee (2016)는 공간의존성을 반영한 공간회귀의 우수성을 회귀모형과 비교분석함으로써 공간회귀의 우수성을 검증하였으나, 분석 데이터를 농지의 실거래가격이 아닌 전문가에 의해 정제된 과세목적의 표준지공시지가를 활용함으로써 시세가격을 추정하지는 못하였다. Choi와 Kim (2015)은 실거래 미발생 지점에 대한 효과적인 가격추정을 위해 실거래 가격 및 고도, 경사도를 통합한 공동크리깅의 적용을 통한 토지가격 산정 가능성을 연구하였고, 하이브리드 기법으로서 공간회귀 모형의 가능성을 제안하였다. 그 외의 관련 연구로는 Lee와 Park (2013) 그리고 Lee와 Kim (2013) 등이 있다.

부동산 가격 산정모형에 대한 기존의 연구들은 대부분 회귀모형에 의한 산정방식으로 적정 설명변수를 찾는 과정과 회귀식에 대한 설명력에 초점을 맞추고 있다. 본 연구는 회귀모형에서 반영되지 못한 공간구조를 반영한 크리깅을 접목한다는 점과 실제 시장에서 거래된 실거래가격을 활용하여 시세가격 도출한다는 점에서 기존의 다른 연구와 차이가 있으며, 실제 회귀모형과 공간회귀 모형 비교를 통하여 공간회귀의 우수성을 가시적으로 보여준다는 점에서 의미가 있다. 그간 실거래자료를 활용한 선행연구는 대부분 실거래 사례가 풍부하고 자료 구득이 용이한 수도권 지역 및 가격형성요인이 비교적 간단하고, 사례수집이 용이한 공동주택에 대한 분석은 많았던데 비해 지방도시와 토지에 대한 분석은 많지 않았다. 본 연구는 부동산 가격을 추정함에 있어 지방 중소도시의 토지가격에 대한 연구라는 점에서 의의가 있을 것이다. 또한 공간회귀 모형에 의해 도출된 토지 시세가격 예측값의 제도적 활용방안을 제안하고자 한다. 즉, 실제 공시가격과 토지 시세가격 예측값 비교 분석을 통해 공시가격의 실거래가격 대비 현실화를 제고방안을 제시하고, 실거래가격에 기반한 시세가격 추정을 통해 실거래 신고가격 적정성 진단기준으로 활용할 수 있음을 살펴보았다.

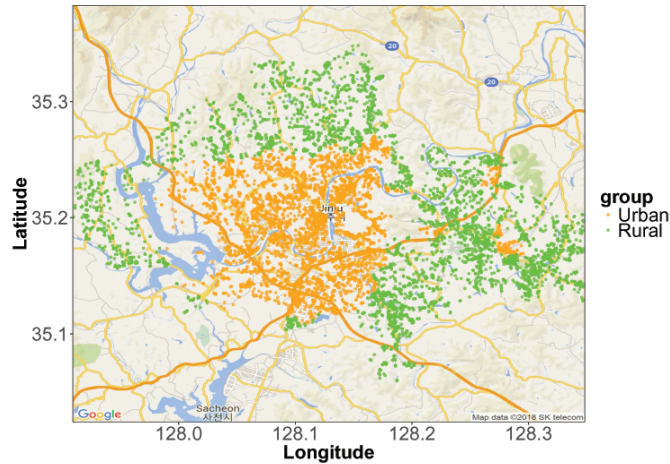


Figure 4.1. Observed locations of study area (urban area vs. rural area).

4. 모형 및 분석결과

4.1. 분석자료 및 변수 설명

부동산의 거래가격이란 실제 매도자와 매수자간의 거래가 완료된 이후 사후적으로 관측되는 결과로서, 분석에 활용하는 사례는 경상남도 진주시에서 2014년~2016년 10월까지 거래되어 부동산 실거래로 신고된 약 1만 건을 대상으로 하였다. 토지가격의 경우 여러가지 가격형성요인들에 의해 결정되나, 특히 개발의 밀도 및 이용상황을 결정짓는 용도지역은 가격을 형성하는 가장 큰 특성으로 용도지역에 따라 가격수준이 현저히 다르고 이전에 거래된 인근의 유사한 부동산 실거래가격 수준에 따라 달라지는 지역적 특성을 지니는바 정확한 분석을 위하여 개발밀도수준과 가격형성요인이 상이한 도시지역과 비도시지역으로 분리하여 연구를 실시하였다.

본 연구에서 활용한 자료 중 실거래사례의 소재지 및 거래가격은 RTMS에 기준하되, 실거래 신고자료의 특성상 신고주의에 의해 수집된 토지특성 정보가 정확하지 않을 수 있는 점을 고려하여 용도지역, 이용상황 등의 특성자료는 지방자치단체 담당공무원이 토지대장 및 토지이용 계획확인서 등 관련 공부에 근거하여 조사한 해당 토지의 개별공시지가 데이터를 활용하였고 고저, 형상, 경사도, 도로 및 철도와의 거리 등은 항공지적 자료를 기초로 수치화하여 활용하였다. 또한 이용상황, 도로접면 등 일부 범주형 변수는 분석 결과의 유의성을 높이기 위하여 유사한 항목을 통합하여 분석하였고, 상권과의 거리는 상업부동산 임대동향조사(한국감정원)에서 조사·공표하고 있는 진주시 중심상권으로부터의 거리를 측정하여 활용하였다.

Figure 4.1과 같이 도시 및 비도시지역의 거래사례 분포를 살펴보면, 진주시를 가로지르는 남강 중심부를 위주로 도시지역이 형성되어 있고 외곽을 둘러싸고 비도시지역이 위치하고 있다. 비교적 고밀도 개발이 가능한 도시지역의 경우 분포범위가 넓은 녹지지역의 거래건수가 많았고 이용상황 역시 농지의 비율이 가장 컸으나, 대지의 거래 비율이 33%로서 비도시지역의 14%에 비하여 상대적으로 높았다. 또한 개발가능한 토지가 주로 거래되는 바 도로접면이 넓은 거래사례가 많고 맹지(접한 도로가 없는 토지)의 거래사례는 없었다. 반면, 비도시지역의 경우 도시지역과 마찬가지로 도로접면 폭이 큰 거래가 거의 없고, 농로의 도로접면 폭인 세로불(자동차 통행이 불가능하나 이륜자동차의 통행이 가능한 도로) 이하의

거래가 72%에 달하였다.

토지의 시세가격이란 일반적인 상식에서 고밀도(용적률 및 건폐율)로 개발 가능할수록 높고 이용상황이 상업용일수록 비싸고 농지 또는 임야 등 조방적으로 이용하는 용도일수록 가격이 낮으며, 형상이 정방형에 가까워 이용 시 토지 손실이 적은 토지가 가격이 높게 형성된다. 하지만 이러한 특성들이 동일하더라도 지리적 위치가 어디냐에 따라 가격이 현저하게 다를 수 있으며 이러한 사유로 토지 시세가격 산정 모형이 공간구조를 지닐 수밖에 없으므로 공간구조를 반영할 수 있는 방법에 대한 연구가 필수적이라 할 것이다.

본 연구에서는 진주시 총 거래사례 중 결측치 및 이상치를 제거하고 남은 9,147건을 도시지역(4,461건), 비도시지역(4,686건)으로 구분하여 분석모형에 활용하였으며, 랜덤샘플링을 통해 각 지역별 4%를 모형 적합에 사용(training data)하였고, 나머지 96%는 모형 성능 검증에 활용(test data)하였다. 전반적으로 도시, 비도시지역 모두에서 공간적인 치우침 없이 데이터가 고르게 분포하도록 하였다. 참고로 도시 및 비도시지역 사례의 기초통계량으로 도시지역의 거래단가 중위수는 116,279원/m²로 비도시지역 26,705원/m²로 보다 약 4.4배 높았다. 최저지가는 임야의 거래가격 수준으로 유사하였으나 최고지가는 약 7.1배가량 차이가 나서 이용상황에 따라 가격차이 폭이 큰 것을 알 수 있다. 2017년도 공시지가 역시 거래단가와 유사한 차이가 발생하였으며, 도로 및 철도, KTX, 상권과의 거리는 전체적으로 도시지역에서 낮게 나타났으며, 고도 및 경사도는 비도시지역에서 높게 나타났다.

4.2. 토지 시세가격 예측을 위한 공간회귀 모형

토지 시세가격 예측을 위해 고려되는 주요 변수는 일반적으로 토지 개발의 밀도를 결정하는 용도지역과 현재 이용 중인 용도를 나타내는 이용상황, 상권과의 거리와 같은 환경적인 요소, 부동산으로서 가지는 고유한 특성인 고저, 형상, 방위 및 도로접면 등이 있다. 이러한 특성이 토지 시세가격에 미치는 영향은 도시지역인지 비도시지역인지에 따라 상이하고, 지역별로는 용도지역에 위치가 구분되기 때문에 도시지역과 비도시지역을 구분하여 변수를 결정하였다. 도시지역은 비교적 고밀도로 개발가능한 지역으로 용도지역(Zoning: 주거, 상업, 공업, 녹지지역), 이용상황(Use), 상권과의 거리(Dist-Market), 주간선도로(도로법 의한 국도/지방도/시도/군도)와의 거리(Dist-Road) 등을 설명변수로 고려하다. 비도시지역의 토지 시세가격 예측을 위해서는 저밀도 개발이 가능한 용도지역(Zoning: 관리, 농림, 자연환경보전지역), 이용상황(Use), 고도(Elev) 및 경사도(Slope), 도로접면의 폭(Width), 형상 등을 설명변수로 사용하였다. 참고로 용도지역, 이용상황, 그리고 도로접면의 폭은 범주형 변수로 사용하였다.

토지의 경우 이러한 속성정보와 더불어 위치가 갖는 공간구조를 고려할 필요가 있는데, 보간기법인 크리깅을 활용하여 해당 위치로 인해 가격이 달라지는 요소인 공간의존성을 반영할 수 있다. 예를 들면, 동일한 특성(주거지역 내 단독주택으로 이용 중인 토지로 유사 도로접면을 가지는 평지)의 토지라고 하더라도 위치(강남구인지 강북구인지)에 따라 그 값이 달라지는 것이 공간의존성이며, 이러한 공간의존성은 기존 회귀모형의 잔차 크리깅을 통해 미처 회귀모형에 반영되지 못한 공간구조를 고려한 가격산정 모형을 만들 수 있는 것이다. 즉, 관측값이 없는 지점에서의 토지 시세가격 예측값을 도출하기 위해서 기존 선형회귀 모형에 공간구조를 가지는 확률효과항을 추가하여 토지 시세가격을 추정하였으며, 토지 시세가격 예측을 위한 통계적 모형은 다음과 같이 표현할 수 있다 (Jin 등, 2017; Kim 등, 2010).

$$y(s) = X(s)\beta + \eta(s), \quad (4.1)$$

여기에서 $y(s)$ 는 반응변수인 s 지점에서의 토지 시세가격, $X(s)$ 는 s 지점에서의 다양한 특성 변수를 나타내는 독립변수 행렬, $\eta(s)$ 는 s 지점에서의 토지 시세가격 추정에 대한 모형 오차로써 공간모형을 가정

Table 5.1. Estimated coefficient (Coef.) and standard error (SE) values of linear regression model for urban and rural areas

Covariate category	Urban area		Rural area	
	Coef.	SE	Coef.	SE
Constant	14.68497	0.218185	13.29205	0.725481
Factor(Use)2	-0.59611	0.148307	-0.47402	0.167081
Factor(Use)3	-2.00158	0.270834	-2.24983	0.251250
Factor(Width)3	-0.45423	0.228749	-1.45721	0.779138
Factor(Width)4	-0.60750	0.219311	-1.96579	0.739600
Factor(Width)5	-0.90445	0.221688	-2.22689	0.739238
Factor(Width)6	-0.80702	0.246719	-2.19556	0.750093
Dist-Market	-9.72E-05	1.56E-05	-	-
Factor(Zoning)2	0.41555	0.201980	-	-
Factor(Zoning)3	-0.04692	0.348006	-	-
Factor(Zoning)4	-0.64437	0.173105	-	-
Slope	-0.07592	0.010197	-0.01903	0.010194
Dist-Road	-0.00075	0.000233	-	-
Dist-Rail	-3.99E-05	1.26E-05	-	-
Elev	-	-	-0.00778	0.001764

한다.

$$(\eta(s_1), \dots, \eta(s_N))' \sim \text{MVN}(\mathbf{0}, \Sigma(\sigma^2, \theta)). \quad (4.2)$$

일반적으로 공분산행렬은 공간오차의 분산(σ^2)과 상관행렬($R_{\theta, \nu}(s, s')$)의 곱인 $\Sigma = \sigma^2 R_{\theta, \nu}(s, s')$ 로 표현되며 실제 토지 시세가격 \mathbf{y} 의 공간구조를 설명한다. 여기에서 공간자기상관 함수로 Matérn 함수를 사용하였고, 그 식은 다음과 같다.

$$R_{\theta, \nu}(s, s') = \frac{\left(\frac{d(s, s')}{\theta}\right)^\nu}{2^{\nu-1} \Gamma(\nu)} K_\nu \left(\frac{d(s, s')}{\theta}\right). \quad (4.3)$$

Matérn 함수는 평활모수 ν , 척도모수 θ 와 각 개체사이의 거리인 $d(s, s')$ 에 의해 정의되며, 일반적으로 평활모수 ν 는 값을 고정하여 추정한다. 여기에서 평활모수 $\nu = 0.5$ 인 경우, Matérn 함수는 가장 일반적이고 간단한 지수함수로 축소된다.

5. 분석결과

Table 5.1은 도시지역과 비도시지역의 회귀 적합 결과를 보여주는 것으로, 다양한 변수 중 통계적으로 유의미한 변수만으로 분석을 진행하기 위해 Stepwise 방법을 사용하였다. 용도지역이 주거지역(Zoning1)인 경우에 비하여 상업지역(Zoning2)인 경우 가격수준이 현저히 높으며, 공업지역(Zoning3) 및 녹지지역(Zoning4)인 경우 더 낮았으며 이용상황이 대지일 경우보다 농지, 임지로 갈수록 가격이 낮아져 토지 거래관행과 일치하였다. 다만, 공업지역의 경우 전체 거래건수가 현저히 미미하고 가격수준이 주거지역과 유사하며 통계적 유의성이 없다고 도출되었다. 상권과의 거리가 가깝고 도로와의 거리가 가까울수록 가격이 높아지는 것은 접근성 향상 때문인 것으로 볼 수 있으며, 그 외 도로접면의 폭(Width), 철도와의 거리(Dist-Rail), 경사도(Slope) 등은 모두 일반적인 상식에 부합하는

Table 5.2. Estimated spatial parameters for urban and rural areas

Area	nugget	partial sill	range
Urban	0.219047	0.174440	0.011461
Rural	0.272529	0.239127	0.017538

Table 5.3. Model comparisons in terms of RMSE, MAE, and SR

Area	Model	RMSE	MAE	SR	
				Mean	Median
Urban	Regression	321288	151692	1.39523	1.06617
	Spatial regression	296951	136664	1.29880	1.03610
Rural	Regression	43478	19909	1.25825	0.94428
	Spatial regression	40419	18811	1.24230	0.95170

RMSE = root mean squared error; MAE = mean absolute error; SR = sales ratio.

결과가 도출되었다. 용도지역 중 공업지역을 제외한 대부분의 변수가 유의하다고 나왔으며 Adjusted- R^2 가 0.814로 설명력이 우수하였다. 또한 *ncvTest*와 *shapiro-test* 결과 등분산성 및 정규성도 만족하였고 공간구조가 있는 자료로서 독립성은 만족하지 않았으나 크리깅 분석을 통하여 독립성 불만족의 문제는 해결하도록 한다. 비도시지역은 고밀도 개발이 제한된 지역으로 농지 및 임야의 거래가 대부분이라 용도지역에 따른 차이가 미미할 것으로 예상되었으며, 모형적합 분석 결과 실제로 용도지역에 따른 차이가 유의성이 없는 것으로 나타나 모형에서 제외하였다. 이용상황, 도로접면, 경사도의 결과치 대부분이 거래관행과 일치하게 나타났으며, 도시지역과 달리 상관과의 거리, 도로 및 철도와의 거리 등이 모형 설명변수에서 제외되었고, 고도 등이 유의미한 변수로 추가되었다. 대부분 변수가 유의하게 나왔고, Adjusted- R^2 는 0.584로 도시지역에 비하여 낮은 수준으로 나타났으며, 등분산성 및 정규성을 만족하였으며 다중공선성은 없는 것으로 확인되었다. 공간회귀 모형에서는 도시지역과 비도시지역에 대하여 각각 상기 회귀모형에서 산출된 잔차를 기준으로 검증 데이터의 잔차값을 크리깅하여 최종 추정값을 도출하였다. 회귀 분석 후 잔차를 Matérn함수 ($\nu = 0.5$)를 이용하여 적합시킨 결과로 높은 공간적 상관성이 존재함을 확인할 수 있었다. 지역별 베리오그램 모수 추정값은 Table 5.2와 같이 나타났다.

참고로 nugget, partial sill, 그리고 range는 베리오그램 함수의 모수들인데 nugget은 두 지점의 거리가 거의 0에 가까울 때 두 지점 사이의 베리오그램 값이고 partial sill은 베리오그램 문턱값에서 nugget을 뺀 값이며 range는 두 지점의 베리오그램 값이 문턱값에 이를 때 두 지점의 거리를 의미한다 (Table 5.2).

도출된 모형의 성능을 검증 및 비교하기 위하여 검증 데이터를 활용하여 토지 시세가격을 예측하였다. 예측의 정확도 확인을 위하여 가격비율(sales ratio; SR = 예측가격/실거래가격), 평균제곱근오차(root mean squared error; RMSE), 평균절대오차(mean absolute error; MAE)를 지표로 활용하였는데, SR은 1.0이 넘을 경우 과다추정, 1.0 미만인 경우 과소추정된다고 볼 수 있으며, RMSE는 사용한 모형이 실제값을 얼마나 유사하게 예측할 수 있는지를 나타내고 MAE는 실제값과 예측값의 차이를 나타내는 값으로 둘 다 수치가 작을수록 예측값이 우수하다고 볼 수 있다. Table 5.3은 도시지역과 비도시지역의 모형 성능을 각각 보여준다. 도시지역의 경우, RMSE와 MAE 모두 공간회귀의 결과치가 낮았으며, SR은 1을 웃돌아 예측가격이 과다추정 되었다고 볼 수 있으며, 평균과 중위수 모두 공간회귀 모형이 1에 더 근접하여 우수한 것으로 나타났다. 비도시지역 역시 공간회귀 모형의 결과값이 더 우수한 것으로 나타났으며, 도시지역과 달리 SR 중위수가 1미만으로 나타나 상대적으로 비도시지역에서 예측값이 과소추정되는 것을 보여준다. 전체적으로 공간회귀 모형이 개선되었으나 공간회귀 모형의 SR의 경

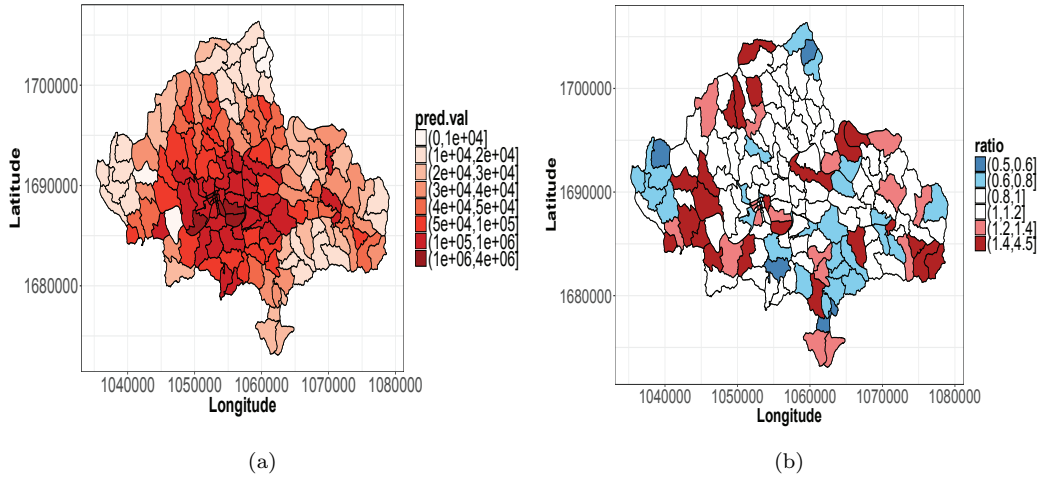


Figure 5.1. Areal average of sales ratio values (a) and predicted sales price (b).

우에도 인접지역에서 과다 또는 과소추정 되는 유사한 경향을 보여주는데, 이는 공간회귀 모형에 반영되지 못하는 공간구조가 일부 남아있는 것으로 보여진다.

Figure 5.1는 토지 시세가격에 대하여 회귀 모형 잔차에 대한 크리깅 값과 회귀분석을 통해 얻은 예측값을 합하여 구한 최종 토지 시세가격 예측값을 나타낸 그림으로, 읍면동 단위로 값의 크기에 따라 색을 달리함으로써 위치에 따른 토지 시세가격의 수준을 나타냈다. 값이 낮을수록 흰색으로, 높을수록 빨간색으로 나타났다. 최종 토지 시세가격 예측값은 강을 따라 진주시 중심부의 도시지역 가격수준은 높게 예측된 반면, 외곽 농업산간지대의 예측 가격은 낮은 수준으로 나타났음을 알 수 있다. 특히 오른쪽 외곽의 농업산간지대 중 가격이 높게 나타난 일부 지역은 외곽의 비도시지역 공간구조가 크리깅으로 반영된 결과로 볼 수 있으며 이러한 결과는 앞서 살펴보았듯이 회귀모형에 비하여 개선된 효과를 나타낸다.

6. 결론

본 연구에서는 경상남도 진주시를 대상으로 실거래 신고가격을 활용하여 회귀모형과 공간회귀 모형으로 토지 시세가격을 추정하고, 그 결과를 실증적으로 분석함으로써 토지 시세가격 산정 방법으로서의 공간회귀 모형 적용가능성을 살펴보았다. 부동산 실거래신고제도에 의하여 수집된 9천여 개의 사례를 활용하여 회귀 모형을 적용하였으며, 토지 특성 등 속성정보 외에 위치정보에 따라 토지가격이 달라지는 점을 고려하여 공간구조가 반영될 수 있도록 잔차크리깅을 통해 공간회귀 기법에 의한 토지 시세가격을 추정하였고 그 결과 공간회귀를 적용했을 때 실거래가격 대비 예측가격의 비율인 SR이 1에 더 근접하게 나타나 실거래가격수준의 예측력이 높아졌으며, 예측오차에 대한 공간분포도 개선되는 것으로 나타났다. 즉, 회귀모형에 남아있던 공간구조를 크리깅함으로써 공간의존성이 상당부분 해소되었음을 알 수 있다. 또한 본 연구에서 활용된 실거래의 가격 대비 공시지가의 비율 (현실화율) 중위수는 약 0.58에 그치는데 반해 실거래가격 대비 시세 추정가격 중위수는 약 0.99 수준으로 나타나 현실화를 측면에서도 우수한 것으로 나타나, 최근 정부에서 논의되고 있는 실거래가 기반 가격공시제도의 한 축을 담당할 수 있을 것이라 생각된다.

Figure 6.1에서 볼 수 있듯이 지역적으로 공시가격 현실화율의 정도가 상이한바 현재 공시가격의 수준

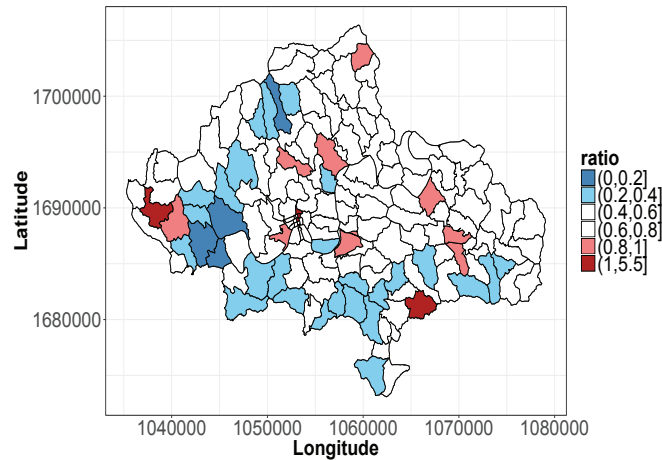


Figure 6.1. Areal average of realization rate (disclosure price/real price).

은 과세형평성이라는 측면에서도 미진한 점이 존재하고 있다. 반면, 기존 방식에 의한 공시지가에 비하여 공간회귀 모형에 의한 추정가격은 조사자 편의가 적고, 실제 거래가격에 기초한 값을 추정함으로써 가격에 대한 설득력이 높아 공시제도의 보완을 통하여 공시지가 산정 모형으로서의 활용방안을 논의해 볼 수 있을 것이다. 물론 경상남도 진주시 1개의 지역만을 대상으로 하여 해당 모형을 일반화하여 전국에 적용할 수는 없을 것이나, 개별토지에 대한 공시지가는 해당 시·군·구청장이 산정 및 공시하도록 하고 있어 산정단위가 시·군·구임을 고려할 경우 시·군·구별 적합모형을 도출하여 적정시세가격을 도출하고, 이를 활용하여 공시지가로 활용하는 것을 고려해 볼 수 있을 것이다. 또한 공시지가를 대신하여 RTMS 가격적정성 진단모형으로서 공간회귀를 검토해 볼 수도 있을 것이다. 다만, 본 연구는 실거래가격의 적정성을 담보하는 것이 중요하나 일부 지역에서 추정오차가 상대적으로 크게 나타나고, 활용 데이터를 실거래 신고자료에만 의존할 경우 저가·고가 신고된 거래사례 또는 거래당사자간의 사정이 개입된 가격이 데이터로 활용되는 경우 토지 시세가격 추정값의 정확도가 저해될 수 있는 한계점이 존재한다. 또한 변수 선정에 있어 토지가격 형성에 반영되는 다양한 변수를 포함하지 못하는 점 등이 여전히 한계로 존재하는바 향후 추가 연구를 통하여 변수 및 모형에 대한 검증 및 보완이 요구된다.

References

- Choi, J. H. and Kim, B. J. (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**, 55–63.
- Jang, W. D. (2008). A study on ways to improve the efficiency of the real estate actual transaction price registration system (Master's thesis), Daegu University, Daegu.
- Jin, H. G., Kim, S. W., and Kim, Y. (2017). A spatial prediction for the flowering and autumnal dates in Korea, *Korean Journal of Applied Statistics*, **30**, 417–426.
- Kim, B., Ku, C. Y., and Choi, J. (2010). Population distribution estimation using regression-kriging model, *Journal of the Korean Geographical Society*, **45**, 806–819.
- Ku, B. and Shin, B. J. (2015). Using ridge regression to improve the accuracy and interpretation of the Hedonic pricing model: focusing on apartments in Guro-gu, Seoul, *Korean Journal of Construction Engineering and Management*, **16**, 77–85.
- Lee, C. R., Eum, Y. S., and Park, K. H. (2014). A Bayesian estimation of price for commercial property:

- using subjective priors and a kriging technique, *Journal of the Korean Geographical Society*, **49**, 761–778.
- Lee, C. R. and Park, K. H. (2013). The effects of neighborhood segmentation on the adequacy of a spatial regression model, *Journal of the Korean Geographical Society*, **48**, 978–993.
- Lee, G. and Kim, K. (2013). A study on the spatial mismatch between the assessed land value and housing market price: exploring the scale effect of the MAUP, *Journal of the Korean Geographical Society*, **48**, 879–896.
- Seo, S. B. and Kwak, S. N. (2014). A study on the adequacy of standard comparison table of land price by Hedonic price model, *Journal of Korea Planning Association*, **49**, 187–204.
- Sohn, H. and Lee, C. (2016). Estimating farmland prices using a regression-kriging model, *The Korea Spatial Planning Review*, **89**, 39–53.
- Sohn, J., Kim, K., and Kim, Y. (2003). A study on the forecasting model of real estate market: the case of Korea, *Housing Studies Review*, **11**, 49–75.
- Yoo, S. J., Lee, J. H., and Jung, H. N. (2014). Analysis on the equity of publicly announced individual land price by comparing sales price, *Journal of the Residential Environment Institute of Korea*, **12**, 39–56.

공간회귀모형을 이용한 토지시세가격 추정

최지혜^a · 진향곤^b · 김용구^{b,1}

^a한국감정원, ^b경북대학교 통계학과

(2018년 1월 16일 접수, 2018년 3월 13일 수정, 2018년 3월 15일 채택)

요약

부동산 투기근절, 공평과세 목적으로 부동산 실거래 신고제도가 도입된 이후, 정부에서 운영 중인 부동산거래관리시스템에는 연간 약 200만 건의 부동산 실거래 신고자료가 축적되고 있다. 인터넷이 발달하고 정보에 대한 접근성이 높아진 요즘, 부동산 투자에 대한 관심 증가로 부동산 가격정보에 대한 요구도 나날이 증가하고 있다. 하지만 이는 단순히 거래사례에 대한 정보만을 제공할 뿐이라 공동주택 실거래의 경우 동, 호수, 토지건물 실거래의 경우 지번을 개인정보보호 등의 이유로 공개하고 있지 않아 실거래의 위치별 정확한 데이터를 구득하기 어려운 실정이라서 정보의 비대칭성이 여전히 존재하고 이러한 부동산 정보의 특수성이 부동산시장에서의 투기가 근절되지 않는 이유 중 하나이다. 본 논문에서는 축적된 실거래 신고가격 데이터를 활용하여 실거래 미발생 지점에 대한 시세가격 추정 모형을 도출하는 것으로, 부동산 가격이 지리적 위치에 따라 결정되는 특수성을 가지는 것을 고려하여 공간구조가 반영될 수 있도록 공간회귀 모형을 통한 추정 토지 시세가격의 정확도를 살펴보았다.

주요용어: 공간예측, 공간회귀모형, 공시지가, 토지실거래가격

본 논문은 제1 저자인 최지혜의 석사학위 논문 중 일부를 정리한 것임.

¹교신저자: (41566) 대구광역시 북구 대학로 80, 경북대학교 통계학과. E-mail: kim.1252@knu.ac.kr