



## I. 서론

농지는 토지정책수단의 실질적인 집행대상임에도 불구하고 토지정책수립, 집행, 평가에 있어서 가장 중요한 기준이 되는 농지에 대한 공시지가는 실제가격과의 괴리가 심하다. 그 결과 공영개발에 따른 토지보상민원이 급증하고 개별공시지가에 대한 이의신청이 증가하고 있다. 이는 농지에 대한 현행 개별공시지가 산정 시 전통적인 농업지대론 중심의 토지특성항목 구성과 감정평가에 의한 표준지 공시지가 산정후 가장 가까운 비교표준지 1개를 기준으로 비준표를 이용한 지가산정방식에 따른 지가산정오류에 기인한 것으로 판단된다.<sup>1)</sup> 또한 도시인근지역의 경우에는 도시와의 거리를 고려한 도시지대론이 고려되지 못하고 미래의 전용가능성에 따른 농지의 옵션가치가 미고려되었기 때문이라고 추정된다. 나아가 대량평가 시 사용되는 전통적인 헤도닉가격모형(hedonic price model)에 의한 지가산정 시 토지가격이 가지는 공간중속성(spatial dependence) 또는 공간이질성(spatial heterogeneity)을 모형에서 고려하지 않을 경우 문제점이 발생하기 때문이다.

우리나라에서도 도시화의 속도가 둔화되면서 최근 들어 대규모 개발위주에서 토지은행과 같은 필지단위 관리중심으로 토지정책의 방향이 전환되고 있다. 본 논문에서는 토지정책수단의 세부집행방안 수립의 근거자료로서 농지가격 결정요인 식별을 연구의 주요 목적으로 한다. 세부적으로는 도시화에정용지로서 농지의 가격평가기준을 마련하기 위해서 개발가능성과 접근성 등 도시지대론을 고려하고, 공간적 의존성과 이질성을 고려한 필지단위 농지특성가격모형 도출을 목적으로 한다.

이를 위해 본 논문에서는 헤도닉 가격모형에서 공간효과를 반영하는 방법으로서 공간계량경제(spatial econometrics) 접근방법인 공간자기회귀모형(spatial autoregressive model: SAR), 공간오차모형(spatial error model: SEM), 일반공간모형(general spatial model: SAC)을 적용하였다. 또한 공간계량경제 접근방법과 가중최소자승(Weighted Least Square: WLS) 개념을 접목시킨 비모수적 추정방법의 일종인 공간가중회귀모형(Geographically Weighted Regression: GWR)을 적용하였다. 실증분석을 위해 2006년을 기준으로 수도권 근교지역인 경기도 광주시 농지의 실거래가격자료와 토지특성조사표 및 GIS분석 자료를 이용하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 공간계량경제 접근방법을 적용하여 농지가격결정요인을 분석한 선행연구를 검토하였고 제3장에서는 공간계량경제모형과 공간가중회귀모형에 대해 살펴보았다. 제4장에서는 실증분석에 사용된 자료와 기초통계 그리고 추정결과를 제시하였고, 제5장에서는 결론을 내렸다.

## II. 선행연구 검토

부동산가격에 관한 대부분의 실증연구에서 헤도닉 가격모형은 회귀분석에 의한 모수추정이 주류를 이루고 있으며, 자료들이 지니고 있는 공간적 특성을 대부분 고려하고 있지 않다. Gillen et al.(2001)은 공간자료가 가지는 공간효과를 고려하지 않고 부동산가격을 추정하게 되면 오차항(error terms)에 공간적 자기상관이 나타나게 되므로 추정된 모수의 표준오차가 커지게 될 뿐만 아니라 이에 따른 통계적 검정

1) 공익사업에 의한 토지보상이 아닌 개발제한구역, 수변구역 토지매수와 사유림 매수, 농지은행 농지매수 및 국공유지 매각의 경우에도 실무적으로 비교표준지 1개를 기준으로 개별요인, 지역요인, 기타요인으로 보정하는 방식인 토지보상평가지침에 따라 예정가격을 산정하고 있으며, 실거래가격과의 차이를 보정하기 위해서 인근 평가 선례 등을 참고하여 기타요인으로 보정하고 있음.

에서도 편향된 결과를 가져 온다고 주장하였다.

Dubin(1988, 1992)은 공간자기상관의 존재는 위치와 관련하여 주변지역의 특성과 접근성에 의해 유발되는 것으로 인식하고 헤도닉 가격모형에서 중요한 입지와 관련한 속성들을 모형에서 고려하지 않고 부동산가격을 추정할 경우 잔차(residuals)들은 공간적으로 자기상관(spatially autocorrelated)된다고 주장하였다.

이와 같은 공간효과를 고려하기 위하여 공간계량경제 접근방법을 적용하여 농지가격결정요인을 분석한 선행연구는 사례가 많지 않고 OLS모형에 의하여 지가를 추정한 경우가 대부분이다. 또한 필지단위가 아닌 시·군·구 단위의 집계자료를 이용한 연구가 대부분이고 특히 국내의 경우 도시토지에 대한 분석이 주류를 이루고 있는 실정이다.

Plantinga et al.(2002)은 미국 48개 주 2,955개 County 단위로 1997년 농지가격을 종속변수로 하여 OLS모형으로 지가를 추정하였다. 그 결과 인구밀도, 농지밀도, 고속도로밀도 변수가 5% 수준에서 유의한 것으로 나타났다. Kumar et al.(2005)도 마찬가지로 방법으로 인구밀도, 농업생산량, 타운과의 거리가 농지가격에 유의하게 영향을 미치는 것으로 분석하였다.

Patton et al.(2003)은 북아일랜드 농지거래 자료를 대상으로 OLS, 하위시장 더미변수를 추가한 OLS, 그리고 SAR모형을 각각 적용하여 농지가격예측력을 비교분석하였다. 전체와 하위시장을 구분한 모형에서 Spatial Chow, LM(lag), LM(error) 검정을 실시하여 공간 lag 종속성과 이질성이 존재함을 확인하였고, 거래 사례가 적은 지역의 경우 감정평가 시 일정지역 범위 내 토지가격의 영향력을 고려해야 함을 주장하였다.

Cotteleer et al.(2008)은 농지매수자의 주소지와 매수 농지위치 간의 거리를 근거리 반경 6.7km를 해당농지에 대한 농업용 토지시장으로 정의하고, 이 지역 내 매수자와 매도자 수를 이용하여 Market Power개념을 적용하여 네덜란드에서 2003년 농민 간에 거래된 건

물이 없는 947필지의 토지를 대상으로 헤도닉 특성이 격함수로 지가를 추정하였다. 다양한 변수를 포함하여 도시형, 도농형, 농촌형으로 지역을 구분하여 추정한 결과, 농가 및 인구 특성, 농지토양유형 등의 변수는 유의성이 없는 것으로 나타났다. 주거지와와의 거리 변수가 도농형과 농촌형에서 고속도로와의 거리가 농촌형지역에서 1% 수준에서 유의하였으며, 도시형의 경우 Reilly Index변수가 1% 수준에서 유의하였다.

국내의 경우, 박현수(2000)는 안양시 주거지역 표준지를 대상으로 OLS모형과 국지가중회귀모형(Locally Weighted Regressive Model: LWR)을 적용하여 OLS방법과 인근 1개 비교표준지선정에 의한 개별공시지가산정방법보다 LWR방법이 RMSE가 각각 35%, 59% 이상 낮아짐을 실증적으로 분석하였다. 김익희(2005)는 서울 서초, 강남, 송파구의 주거 및 상업업무용 표준지를 대상으로 OLS와 GWR 모형적합도를 비교하였는데, 입지특성과 지가의 관계에 일정한 패턴이 없었고 같은 동 내부에서도 이질적으로 분포함을 확인하였다. 최열 외(2006)는 창원시 주거지역 개별공시지가를 사용하여 주변 용도지역과의 접근성변수를 고려하여 OLS, SAR, SEM 모형의 적합도를 비교분석하였다. 특히 주거용토지, 중심상업지역과의 거리, 공업지역과의 거리에서 OLS에서는 음의 상관관계가 나타났으나 SAR, SEM에서는 양의 상관관계가 나타남을 확인하였다.

서교(2005)는 전국 농촌지역의 용도별 표준지공시지가 median을 종속변수로 삼고 주요통계지표를 독립변수로 하여 OLS, SAR, SEM, SAC 모형으로 추정하였으며, 인구밀도가 모든 용도의 토지가격에 공통적으로 영향을 미친다고 주장하였다. Kang et al.(2010)은 전국 시·군별 비도시지역 평균공시지가를 종속변수로 삼고 인구밀도변화율, 토지거래량, 경제개발예산 등을 독립변수로 하여 OLS, SAR, SEM, SAC, GWR 그리고 Mixed GWR을 각각 적용하여 모

형적합도를 비교하였다.

이상과 같이 토지에 대한 전통적 헤도닉 특성가격 함수에 의한 가격추정 시 발생하는 공간효과 미고려의 문제점을 보완하기 위하여 공간계량경제모형과 공간가중회귀분석과 같은 비모수적 접근방법을 적용한 선행연구를 종합해보면 다음과 같다. 첫째, OLS모형에 비해 공간효과를 고려한 SAR, SEM, SAC, GWR 모형의 적합도가 전반적으로 높은 것으로 나타났다. 둘째, 인구 및 접근성 관련 변수의 유의성이 높은 것으로 나타났다. 셋째, 박현수(2000), 김익희(2005), 최열 외(2006)를 제외하면 대부분의 연구가 시·군·구 단위 등의 집계자료를 이용하여 Robinson(1950)이 제시한 생태학적 오류(ecological fallacy) 내재가능성의 문제가 있다.

따라서 본 논문에서는 필지단위의 micro level 데이터를 이용하여 기존 개별공시지가 산정 시 토지 특성에 포함되지 않는 인구 및 접근성 변수를 포함하였다. OLS와 함께 공간효과를 고려하기 위하여 SAR, SEM, SAC, GWR 모형을 각각 적용하여 농지 가격결정요인에 대하여 비교분석하였다.

### III. 모형: 공간계량경제 접근방법

#### 1. 공간계량경제모형: SAR, SEM, SAC

일반적으로 토지가격 추정 시 대상지역 내 모든 토지를 서로 독립적으로 보고 모수를 추정하는 헤도닉 가격함수를 이용하여 추정하게 된다. 그러나 토지는 위치의 고정성과 함께 외부효과에 의해 큰 영향을 받는 특성 때문에 인접한 토지의 가격 및 특성에 의해 공간종속성(Spatial Dependency)과 공간이질성(Spatial Heterogeneity)을 가지게 된다. 여기서 공간종속성이란 종속변수끼리 상관관계를 가지게 되는 경우를 말하고 공간이질성은 공간이 종속변수의 결

정에 미치는 영향이 균일하게 나타나지 않음을 의미한다. 지가추정 시 사용하게 되는 통상적인 OLS추정을 통한 선형회귀모형의 경우 종속변수는 독립변수에 따라 구성되는 함수여야 하고, 동일한 설명변수에 대하여 오차범위 내에서 비슷한 경향으로 결정되는 변수여야 한다. 그러나 공간종속성이 종속변수나 오차항에 존재할 경우 OLS추정은 선형성, 정규성, 동분산성의 기본기능을 위반하게 되어 모형의 잘못된 해석을 유도할 수 있다. 나아가 공간이질성으로 인해 공간자체가 일종의 확률변수를 포함하는 형태로 구성되어 있기 때문에, 매우 유사한 설명변수를 가진 지역이라 할지라도 지역에 따라 다른 종속변수의 값을 가지게 된다(이성우 외, 2006). 따라서 모든 토지가 서로 독립적이라는 가정은 성립될 수 없고 OLS모형을 통한 지가추정에 한계가 존재한다.

공간계량경제모형은 공간종속성과 공간이질성 문제 해결을 목적으로 공간을 설명변수로 포함하기 위하여 공간가중치행렬이 일종의 설명변수로 포함되어 있으므로 WLS와 같이 OLS모형의 확장된 형태로 볼 수 있다. 이와 같이 공간자기상관을 고려하여 공간가중치행렬을 구축하고 공간정보를 회귀분석으로 통합한 공간계량경제모형에는 대표적으로 SAR, SEM, SAC가 있다(Anselin, 1988).

Anselin(1988)에 따라 개별 토지가격은 다음의 공간특성가격함수(spatial hedonic price function model)에 의해 결정된다고 하면 다음과 같다.

$$Y = \rho W_1 Y + X\beta + \epsilon \quad <식 1>$$

$$\epsilon = \lambda W_2 \epsilon + v \quad <식 2>$$

여기서  $Y$ 는 토지가격을 나타내는 종속변수로  $N \times 1$  벡터이며,  $X$ 는 토지가격에 영향을 미치는 독립변수로  $N \times k$  벡터이다.  $\rho$ 와  $\lambda$ 는 공간자기회귀계수(spatial autoregressive coefficient)로 스칼라이

다.  $W_1$ 과  $W_2$ 는  $N \times N$  공간가중치행렬(spatial weight matrix)이다.  $v$ 는 평균이 0이고 분산이  $\delta^2$ 인 서로 독립적이고 분포가 동일한(i.i.d.) 정규분포를 하는 오차를 나타내는  $N \times 1$  벡터이다.

첫 번째로  $W_2=0$ 인 경우를 공간자기회귀모형(Spatial Autoregressive Model: SAR) 또는 공간시차 모형(Spatial Lag Model: SLM)이라고 한다.

$$Y = \rho WY + X\beta + v \quad \text{<식 3>}$$

$v$ 는 평균이 0이고 분산이  $\delta^2$ 인 서로 독립적이고 분포가 동일한(i.i.d.) 정규분포를 따른다. 위의 <식 3>을 정리하면 <식 4>의  $(I - \rho W)^{-1}$ 는 공간승수효과(spatial multiplier)를 나타낸다.

$$Y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} v \quad \text{<식 4>}$$

이러한 공간승수는 공간적 상호작용에 대한 간접효과 또는 전체 외부효과(global externality or spillover)를 의미하며, 하나의 시스템에서 모든 지점이 서로 연관되어 있다(Anselin, 1988)는 의미로, 한 지점의 토지가격은 자신이 속하는 지역의 주변 특성뿐만 아니라 공간가중치행렬을 통해 다른 지역의 주변 특성의 변화에도 영향을 받는다(박헌수 외, 2003).

두 번째로  $W_1=0$ 인 경우를 공간오차모형(Spatial Error Model: SEM)이라고 정의한다.

$$\begin{aligned} Y &= X\beta + \lambda W\epsilon + v \\ Y &= X\beta + \lambda W(Y - X\beta) + v \\ (I - \lambda W)Y &= (I - \lambda W)X\beta + v \\ Y &= X\beta + (I - \lambda W)^{-1}v \end{aligned} \quad \text{<식 5>}$$

SEM에서는 모형에서 고려하지 못한 변수들 사이에 공간적 종속성이 있다고 가정한다. 이는 모형에

서 고려하지 못한 독립변수가 주변지역의 독립변수에 의해 영향을 받게 되고 이는 공간가중치행렬을 통해 모형에 반영된다. <식 5>를 보면 SAR에서는 공간승수효과가 전체 변수에 미치는 점과는 달리 SEM에서는 공간승수효과가 오차항에만 영향을 미치는 것을 알 수 있다.

세 번째로  $W_1, W_2 \neq 0$ 인 경우를 일반공간모형(General Spatial Model: SAC)이라고 정의한다. 여기에서  $W_1$ 과  $W_2$ 는 각각 공간가중치행렬을 나타내고  $\rho$ 와  $\lambda$ 는 각각 공간가중치행렬의 스칼라이므로 공간종속성과 공간이분산성의 척도를 추정하는 역할을 수행하게 된다. 종속변수의 공간자기상관의 정보가 독립변수로 설정되는 동시에 오차항에 대한 공간자기상관도 통제하게 된다. SAC를 사용할 때는  $W_1$ 과  $W_2$ 를 선택해주는 일종의 경험적 직관이 있어야 하며 자료의 성격에 따라 달라질 수 있다(이성우 외, 2006).

이와 같은 공간계량경제모형은 OLS를 사용하게 되면 추정량에 편기가 생긴다. 이를 보정하기 위하여 간접최소자승법(Instrumental Variables Estimation: IVE), 최우추정법(Maximum Likelihood Estimation: MLE), 2단계 최소자승법(Two Stage Least Square Estimation: 2SLS), 적률추정법(Method of Moment) 등으로 추정할 수 있으며, 본 논문에서는 최우추정법을 사용하기로 한다.

## 2. 공간가중회귀모형: GWR

앞서 설명한 SAR, SEM, SAC와 같은 공간계량경제모형이 전역적인 모형이라면 공간가중회귀모형(Geographically Weeighted Regression: GWR)은 국지적인 공간상관성을 이용하는 국지적 모형으로 공간확장모형(spatial expansion model)의 일종이다. 이는 가중최소제곱(Weighted Least Square: WLS)

모형의 원리를 2차원 공간으로 확장시킨 것으로 WLS란 일반적으로 이분산성이 관측되는 선형회귀모형의 효율성을 높이기 위한 추정법이다. 이분산성을 가지는 회귀식을 <식 6>과 같이 구성하였을 때,

$$Y_{ij} = \sum_{j=1}^k \beta_{ij} X_{ij} + \epsilon_i$$

$$\Rightarrow \frac{Y_i}{\sigma_i} = \frac{1}{\sigma_i} \sum_{j=1}^K \beta_{ij} X_{ij} + \frac{\epsilon_i}{\sigma_i} \quad \text{<식 6>}$$

이고, 지역변수에 의해 발생하는 이분산성을 해결하는 가중치를  $W_i^{1/2}$ 라고 하면

$$W_i^{1/2} y = W_i^{1/2} X \beta_i + W_i^{1/2} \epsilon_i$$

$$\Rightarrow y_i^* = X^* \beta_i + \epsilon_i^* \quad \text{<식 7>}$$

이다. <식 7>의 추정회귀계수는

$$\hat{\beta}_i = (X^{*'} X^*)^{-1} X^{*'} y^*$$

$$= (X' W_i^{1/2} W_i^{1/2} X)^{-1} X' W_i^{1/2} W_i^{1/2} y$$

$$= (X' W_i X)^{-1} X' W_i y \quad \text{<식 8>}$$

$$W_i = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & & \\ \vdots & & \ddots & \\ 0 & & & w_{in} \end{bmatrix} \quad \text{<식 9>}$$

이다. <식 7>은 <식 6>에서 이분산성 문제를 해결하기 위해 관찰치의 표준편차를 활용하였던 것을 공간의 특성을 반영하여 처리하여 주기 위해  $W_i^{1/2}$ 라는 가중치행렬로 대체하여 사용한 것이다. <식 7>을 최소제곱법에 의해 회귀계수  $\hat{\beta}_i$ 를 추정하면 <식 8>과 같이 얻을 수 있다. <식 7>과 <식 8>은 관찰지역 모두에 대해 추정을 실행하여 주는 방법이므로 실제적으로는 관찰지역의 수

만큼 회귀분석을 실행해 주는 모형이다(이성우 외, 2006: 268-269). <식 9>에서 일정한 상수로 구성된 가중치행렬을 사용하는 WLS와 달리, GWR에서의 가중치는 각각의 관측치(i)마다 다르기 때문에 모든 관측치별 회귀계수( $\hat{\beta}_i$ )를 추정할 때 그것과 가까운 이웃관측치에 보다 많은 가중치를 부여하는 방법으로 가중치행렬이 작성된다(Fotheringham et al. 2002: 54).

공간계량경제모형과 마찬가지로 GWR모형 역시 가중치를 어떻게 산정하느냐가 매우 중요하다. 공간 가중치행렬의 각 원소  $w_{ij}$ 는 다양한 형태의 가중치가 사용될 수 있는데 일반적으로 다음과 같이 Exponential, Tricube, Gaussian 가중치 중에서 하나를 사용하게 된다.

$$\text{Exponential : } w_{ij} = \sqrt{\exp\left(\frac{-d_{ij}}{d_i}\right)}$$

$$\text{Tricube : } w_{ij} = \begin{cases} 1 - \left(\frac{d_{ij}}{d_i}\right)^3 & d_{ij} < d_i \\ 0 & d_{ij} > d_i \end{cases}$$

$$\text{Gaussian : } w_{ij} = \phi\left(\frac{d_{ij}}{\sigma d_i}\right) \quad \text{<식 10>}$$

단,  $\phi(\cdot)$ 는 표준정규분포함수의 P.D.F.

여기서  $d_{ij}$ 는 회귀분석의 중심점  $i$ 에서 각 지역  $j$ 까지의 거리이며  $d_i$ 는 bandwidth를 의미한다. Exponential과 Gaussian 가중치의 경우 bandwidth는 회귀분석의 중심점  $i$ 에서 대상지역 내 관측치와의 최대거리를 의미하며, Tricube 가중치의 경우 인접한  $q$ 개 관측치와의 거리를 의미한다. 여기서 가중치를 만드는 bandwidth가 고정된 fixed kernel 방식과 Tricube 가중치와 같이 대상지역에 포함되는 관측치( $q$ ) 수를 정하여 가변적인 거리를

bandwidth로 사용하는 adaptive kernel 방식이 있다.

#### IV. 실증분석

##### 1. 자료 및 변수선정

연구지역의 범위는 서울 근교에 위치한 경기도 광주시의 비시가화지역 농지를 대상으로 하였다. 비시가화지역은 도시지역 중 개발제한구역과 관리지역, 농림지역, 자연환경보전지역을 대상으로 하였으며, 농지는 지목이 전, 답이고 토지특성조사표상 토지 이용상황이 전, 답인 토지를 모두 충족하면서 2006년 1년 동안 1회만 거래된 732필지를 표본으로 선정하였다.<sup>2)</sup>

앞서 선행연구 검토에서 제시했듯이, 본 논문에서는 도시인근 지역 농지의 특성을 고려하여 개별공시지가 산정 시 사용되는 토지특성조사항목과 도시지대론에서 주요지가형성요인으로 제시하는 인구 및 접근성 항목을 변수로 추가하여 농지가격을 추정하였다. 본 연구의 표본자료에 대한 대상변수의 기초 통계량은 <표 1>과 같다.

모든 변수를 독립변수로 추정할 경우, 다중공선성(multicollinearity)의 문제가 발생하여 회귀모형 추정을 곤란하게 할 가능성이 크며, 편의(bias)의 양은 최소화할 수 있으나 반대로 분산이 커지므로 평균평방오차(Mean Square Error: MSE)가 최소가 되는 것은 아니기 때문에 추정에 어려움이 많이 발생하게 된다. 이와 같은 관례로 통계학에서는 축소된 모형으로 종속변수를 추정하는 방법을 많이 이용하는데, 신뢰성이 높은 축소모형(reduced model)을 추정하는 방법으로 변수증가법(forward selection method), 변수

감소법(backward elimination method), 변수증감법(stepwise method)이 있다(이용범, 1998: 63). 본 연구에서는 변수증가법<sup>3)</sup>에 의하여 모형에 포함될 변수를 선정하였다.

먼저 종속변수는 2006년 경기도 광주시 농지  $m^2$  당 실거래가격으로 하였다. 독립변수로는 토지특성조사항목 중 면적, 토지 이용상황, 토지형상, 도로접면, 개발제한구역, 농업진흥지역을 변수로 포함하였으며, 접근성변수로는 2차선도로와의 거리, 4차선도로와의 거리, 도시와의 거리, 주거지역과의 거리, 공업지역과의 거리를 포함하였다. 아울러 인구변수로 농지 반경 5km 내 및 2km 내 인구 수를 포함하였고 개발압력 관련 변수로 2006년 1년 동안 반경 2km 내 토지거래량을 포함하였다. 인구와 접근성 관련 변수를 생성하기 위해서 새주소DB의 용도별 건물 및 도로 레이어, KLSI의 용도지역도, 동리별 주민등록인구통계를 이용하여 GIS작업을 수행하였다. 고전적 농업지대론에서 지가결정요인으로 주장하는 비옥도와 경지정리 항목의 경우 총 732필지 중 556필지의 경우 입력값이 없었기 때문에 변수에 포함시키지 못했다.

적절한 함수형태를 선정하기 위해서 <표 2>와 같이 linear, semi-log, log-log, Box-Cox1(semi-log), Box-Cox1(log-log), Box-Cox2(semi-log), Box-Cox2(log-log) 등 총 7개 함수형태로 추정하였다.<sup>4)</sup> 추정결과 semi-log 함수형태가 다른 함수형태에 비해 AIC, SC, RMSE가 가장 낮았으며,  $R^2$ 와 log likelihood값이 가장 높은 것으로 나타났다. 아울러 Box-Cox 검정절차에 의거하여 선형형태의 종속변수의 기하평균을 구하여 종속변수를 기하평균으로 나눈 후 회귀방정식

2) 공부상지목과 농지법상 농지개념이 동시에 충족되는 토지로 한함. 국토해양부 내부자료.

3) 본 연구에서는 SC(Schwarz criterion)을 이용하였음.

4) Box-Cox1은 y변수만 변환시킨 것이고, Box-Cox2는 y변수와 x변수를 모두 변환시킨 것임.

을 추정한 결과 선형함수의 잔차제곱합(RSS)은 654.73, semi-log함수는 214.64 그리고 log-log함수는 225.37로 나타났다.5) 따라서 본 논문에서는 semi-log

함수형태를 사용하였다. 다시 말해, 토지가격은 자연 대수를 취하고 접근성, 인구 및 개발 관련 변수는 모두 연속변수(continuous variable)로서 로그변환을 취하

표 1\_ 연구대상지역 표본자료 기초통계

분류		필지 수	평균	표준편차	Median	Min	Max	단위	비고			
중속변수		732	149110.5	131083.3	112490.5	913462	4995	원/m <sup>2</sup>	2006년 실거래가			
독립 변수	구조적 특성	면적		732	972.7145	1065.185	645	7320	3	m <sup>2</sup>		
		토지 이용	전	389	0.531421	0.499353					dummy	
			답@	343								기준변수
		도로 접면	소로이상	60	0.081967	0.274502						dummy
			세로가	219	0.29918	0.458212						
			세로불 맹지@	73	0.099727	0.29984						
		토지 형상	정형	98	0.13388	0.340756						dummy
			부정형@	634								기준변수
		고저	평지	344	0.469945	0.499437						dummy
			평지 이외@	388								기준변수
	용도 지역 지구	개발제한구역		160	0.218579	0.413565					dummy	
		개발제한구역 외@		572								기준변수
		관리지역		432	0.590164	0.49214						dummy
		관리지역 외@		300								기준변수
		농업진흥지역		581	0.793716	0.404914						dummy
	농업진흥지역 외@		151								기준변수	
	주변 지역 특성	2차로거리		732	84.02596	107.7926	54	938	6		m	KLIS, 새주소DB ArcGIS 직선거리
		4차로거리		732	5200.799	2784.889	5130.5	12465	58			
도시와의거리		732	5771.025	2303.737	5460.5	11691	448					
주거지와외거리		732	15041.23	5984.429	15045	27352	2313					
상업지와외거리		732	2779.689	1841.635	2311	8553	89					
공업지와외거리		732	3998.492	1868.341	3695.5	9607	485					
인구 및 개발 특성	주변 5km 내 인구수		732	8992.044	3154.478	9177	14886	1759		명	새주소DB 주거용 건물 1동당 인구 수 이용	
	주변 2km 내 인구수		732	29696.03	29685.35	16832	126688	914				
	주변 1km 내 인구수		732	5168.414	6026.265	2499.5	42942	113				
	주변 2km 내 토지거래량		732	1529.6	2015.379	683.5	13114	0		건	2006년 경기도 광주 시 토지거래량	
	주변 2km 내 지목변경 수		732	126.1243	88.46773	114	677	7			KLIS	

주: @는 기준변수를 의미함.

5) 자세한 것은 University of London의 Jonathan Wadsworth 교수의 웹사이트(<http://personal.rhul.ac.uk/uhte/006/ec2203/index2.html>)와 (서경천 외, 2001: 9)를 참조.

표 2\_ 다양한 함수변환 결과 비교

구분	OLS (linear)	OLS (semi-log)	OLS (log-log)	Box-Cox1 (semi-log)	Box-Cox1 (log-log)	Box-Cox2 (semi-log)	Box-Cox2 (log-log)
R <sup>2</sup>	0.4159	0.6259	0.6072	0.6113	0.5526	0.3222	0.5012
log likelihood	-9467.0	-589.7	-607.5	-603.7	-655.2	-807.2	-695.0
AIC	18970.0	1215.3	1251.0	1245.4	1348.4	1652.4	1427.9
SC	19053.0	1298.0	1333.7	1332.8	1435.7	1739.7	1515.2
RMSE	100120.0000	0.5415	0.5549	0.5520	0.5922	0.7289	0.6253

지 않고 실제거리(m) 및 수량을 사용하였다. 경제학적인 관점에서 semi-log 함수형태는 특정 토지특성 변화에 따른 화폐가치를 나타낼 수 있고 통계적인 관점에서 보면 예측오차는 토지가격이 커질수록 절대값이 커지는 경향이 있기 때문에 semi-log 함수형태는 토지가격과 잔차와의 이분산성을 교정할 수 있다(박헌수, 2000).

2. 이분산성, 비정규성, 공간중속성 검증

본 논문에서는 공간계량경제모형 접근방법을 적용하여 공간적 효과를 반영하고 최적의 농지특성가격모형을 제시하기에 앞서 OLS모형에 대하여 오차항의 이분산성, 비정규성 및 공간중속성에 대한 검증을 실시하여 이들의 존재여부를 확인하였다.

먼저 OLS모형에서 오차항의 이분산성을 검증하기 위하여 Breusch-Pagan(Koenker) 검정법을, 비정규성을 검증하기 위하여 Jarque-Bera 검정법을 실시하였다. Breusch-Pagan 검정결과는 67.444548(p=0.000000), Jarque-Bera 검정결과는 28.55748(p=0.0000006)으로 분석되어 OLS모형에서 오차항의 비정규성과 이분산성이 존재하지 않는다는 귀무가설을 기각하게 되어 오차항에서 이분산성과 비정규성의 존재를 확인하였다(<표 3> 참조).

앞서 공간계량경제모형에 대한 설명에서 언급한 바와 같이 공간중속성을 검증하기 위해서는 공간가

표 3\_ 이분산성, 비정규성 검증결과

검정	DF	Value	Prob
Breusch-Pagan(Koenker)	16	67.444548	0.000000
Jarque-Bera	2	28.55748	0.0000006

중치행렬을 미리 설정해야 하며, 이는 공간중속성을 모형 내에서 구조화하는 방법으로 공간계량경제모형에서 가장 중요한 비중을 차지하게 된다. 공간가중치행렬을 설정하는 방법에는 인접성(contiguity) 또는 좌표 간 거리를 이용하는 방법이 있다. 본 논문에서는 각 지점의 위치를 알고 있는 경우에는 거리를 이용한 가중치행렬이 인접성을 이용한 가중치행렬보다 바람직하기 때문에 거리가중치행렬을 작성하였다(Anselin, 1988: 20).

거리를 이용한 가중치 행렬에서 가중치의 설정은 일반적으로 역거리 또는 역거리 제곱 가중함수와 같은 직접적 거리측정이 사용된다(서경천 외, 2001).  $d_{ij}$ 를  $i$ 와  $j$ 간의 거리,  $h$ 를 한계이웃관측치의 수라고 하면,  $j$ 번째 토지가  $i$ 번째 토지의  $h$ 번째 이웃관측치보다 가까이 위치하는 경우 가중치  $w_{ij}$ 는  $1/d_{ij}^\alpha$ 로 부여하고 나머지는 0으로 부여하게 된다. 여기서  $\alpha$ 가 2 이상일 경우 즉, 가중치를 거리제곱의 역수 이상으로 설정할 경우 각 가중치의 값이 0에 근사하여 거리에 따른 가중치가 큰 차이를 나타내지 못하기 때문에 본 논문에서는 거리역수를 가중치로 설정하였다(김진 외, 2009).<sup>6)</sup>

본 논문에서는 실제 모형에 적용될 거리가중치 행렬의 가중치를 작성하는 방법으로 한계거리 대신 한계이웃관측치수(h)를 적용하여 자신을 제외한 가장 가까운 이웃관측치수를 1개부터 731개까지 늘려 가면서 Moran's I 통계량을 분석하였다. 그 결과 <그림 1>과 같이 표본전체의 이웃관측치인 731개가 될 때까지 Moran's I 통계량이 계속해서 낮아지는 것으로 나타났다. 이에 따라 본 연구에서는 SAR, SEM, SAC모형 추정 시 한계이웃관측치수를 자기 자신을 제외한 731개 모든 관측치와의 거리역수를 적용한 거리가중치행렬을 대입하였다.

그림 1\_ 한계이웃관측치(h)별 Moran's I 통계량 변화

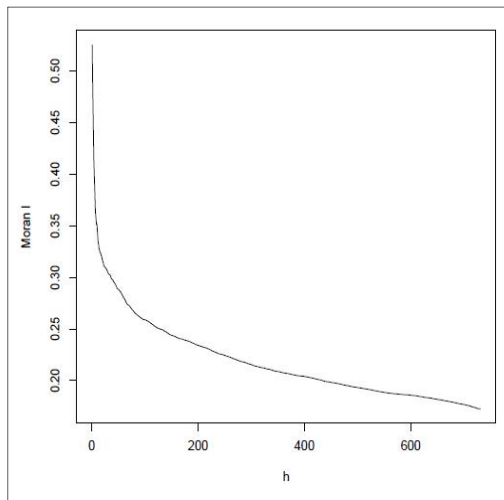


표 4\_ OLS모형 공간종속성 검증결과

threshold	Moran's I		LM error	LM lag
	MI	Value		
h=731	0.4253501	12.86686***	197.4371***	203.509***

\*\*\* 유의수준 0.01, \*\* 유의수준 0.05, \* 유의수준 0.1

이상과 같은 거리가중치행렬을 이용하여 본 연구에서는 공간종속성을 검증하기 위하여 대표적인 검증방법인 Moran'I와 LM(Lagrange Multiplier) error, LM lag 검정을 실시하였고 그 결과 <표 4>와 같이 공간종속성이 존재함을 확인하였다.

### 3. 분석결과

#### 1) OLS모형 추정결과

OLS모형에 의한 지가추정결과 공업지와외의 거리와 2km 내 토지거래량 변수를 제외하고는 유의수준 10%에서 모든 변수가 농지가격에 영향을 주는 요인으로 나타나고 있다. 다른 변수가 일정하고, 즉 다른 변수들이 평균적인 값을 취하고 토지 이용만 변화한다고 가정했을 때, 토지 이용이 전인 경우는 답인 경우에 비해 농지가격이 약 10.8% 상승한 수준으로 결정된다고 볼 수 있다. 이는 답보다 전이 일반적으로 대지, 공장, 창고 등 도시화용지로의 토지형질변경이 상대적으로 용이하기 때문이라고 할 수 있다.

접면도로에 대해서는 계수의 크기는 소로이상, 세로불, 세로가의 순으로 나타나고 있는데 맹지를 기준으로 소로이상은 약 57%, 세로불은 21.9%, 세로가는 19.3% 정도 증가되는 것으로 파악된다. 토지형상에 대해서는 정형은 부정형에 비해 약 17%, 용도지역지구구역에 대해서는 개발제한구역과 농업진흥지역의 경우 그 이외의 지역과 비교했을 때 각각 -90%와 26% 수준으로 지가가 결정되는 것으로 나타났다.

6) 이는 Tobler(1970)의 지리학의 제1법칙에 보다 부합하는 것으로 판단됨.

7) Semi-log 함수형태 모형결과를 해석할 때 다른 변수가 평균일 경우를 가정하여 터미변수의 경우 기준변수와 비교했을 때 “회귀계수×100”% 수준으로 종속변수가 변하는 것으로 해석하며, 연속형 변수의 경우 연속형 변수 1단위가 변함에 따라 “회귀계수×100”% 수준으로 종속변수가 변하는 것으로 해석함.

접근성변수에 대해서 살펴보면, 2차로와 거리가 1m 멀어질수록 농지가격은 평균보다 0.2% 감소하고 4차로와의 거리가 1m 멀어질수록 농지가격은 평균보다 0.002% 감소하는 것으로 추정된다. 또한 도시와의 거리와 주거지와의 거리가 1m 멀어질수록 농지가격은 평균보다 각각 약 0.002%와 0.005%씩 감소하는 것으로 추정된다.

## 2) SAR, SEM, SAC 모형 추정결과

본 논문에서는 공간종속성을 감안한 SAR, SEM, SAC모형에서 공간가중치행렬의 한계거리 및 한계 이웃관측치의 수를 고려하지 않고 전체 이웃관측치와의 역거리를 가중치로 한 공간가중치행렬을 적용하여 모형을 추정하였다. 아울러 SAC모형에서는  $W_1$  과  $W_2$  에 동일한 가중치행렬을 적용하였다.

공간가중치행렬을 이용하여 공간자기회귀 및 공간오차 효과를 고려한 공간계량경제모형은 MLE의 방법을 통해 추정되기 때문에 OLS추정 모형과 비교할 만한 동일한 통계적 기준은 엄밀한 의미에서 존재하지 않는다(이성우, 2006). 공간계량경제모형에서는  $R^2$ 이 아닌 로그우도(Log-likelihood), AIC(Akaike Information Criterion), SC(Schwarz Criterion)로 모형의 적합성을 검정하며, 일반적으로 로그우도가 증가하고 AIC 및 SC가 감소하면 모형적합성의 향상을 나타낸다(Anselin, 2005).

모형비교결과 SAC모형이 가장 높은  $R^2$ , log-likelihood 와 가장 낮은 AIC, SC, RMSE를 나타내고 있어 OLS모형은 물론이고 SAR, SEM모형보다도 모형 적합도가 높은 것으로 나타났다. 전반적으로도 OLS

모형에 비해 공간계량경제모형이  $R^2$ , log-likelihood 가 높고 AIC, SC, RMSE가 낮은 것으로 나타났다. 한계이웃관측치의 수(h)를 731로 설정한 공간가중치행렬을 통해 주변지역 종속변수의 공간시차(Spatial lag)와 모형 내 변수에 포함되지 않은 공간오차(Spatial error)를 변수로 포함시켰기 때문에 SAC모형이 가장 설명력이 높은 것으로 판단된다.

추정된 회귀계수를 살펴보면, 공간적 효과를 분리하였기 때문에 도로집면 세로불을 제외하고 대부분의 변수에서 OLS모형보다 낮은 계수값을 나타내고 있다.<sup>8)</sup> 구조적 특성변수 중 도로집면을 살펴보면 맹지와 비교했을 때 소로이상은 약 53%, 세로가는 약 19%, 세로불의 지가는 약 24% 증가하는 것으로 나타났다. 용도지역지구구역에 대해서는 개발제한구역과 농업진흥지역의 경우 그 이외의 지역과 비교했을 때 각각 -67%와 19% 수준으로 지가가 결정되는 것으로 나타났다. 접근성변수에 대해서 살펴보면, 2차로와 거리가 1m 멀어질수록 농지가격은 평균보다 0.24% 감소하는 것으로 추정된다. 도시와의 거리와 주거지와의 거리가 1m 멀어질수록 농지가격은 평균보다 각각 약 0.002%와 0.004%씩 감소하는 것으로 추정된다. 아울러 OLS모형에서는 4차로와의 거리 변수가 1% 수준에서 유의한 변수였지만, SAR, SEM, SAC모형에서는 10% 수준에서 유의한 것으로 나타났다. 반면 농업지와의 거리는 OLS모형에서 유의성이 없는 것으로 추정되었지만, SAC모형에서는 5% 수준에서 유의한 것으로 나타났다.

8) 본 논문의 심사의견에 따라 전미분하여 구한 semi-log함수의 탄성치는  $(1-\rho W)^{-1}\beta X$ 이며 해석 시 공간자기회귀계수( $\rho, \lambda$ )로 지가의 영향이 분산되었기 때문에 OLS보다 공간계량경제모형의 회귀계수는 작게 나타나야 함. 그러나 공간계량경제모형에서 행표준화된 공간가중치행렬을 사용하기 때문에 대상지역 전체에 대한 회귀계수 해석 시 공간가중치행렬  $W$ 를 1로 보고 해석하게 된다. 따라서 각 관측치별로 회귀계수를 해석할 경우 전역적으로 해석하는 경우보다 회귀계수의 공간적인 변이가 다양하게 나타날 수 있을 것임.

표 5\_ OLS, SAR, SEM, SAC 추정결과

변수	OLS(Semi-log)			SAR			SEM			SAC			
	Coeff	St. Error	t-stat	Coeff	St. Error	t-stat	Coeff	St. Error	t-stat	Coeff	St. Error	t-stat	
상수	12.24844***	0.14702	83.31332	1.67234***	0.58425	2.86236	11.72672***	0.41702	28.12046	5.52029***	1.83665	3.00563	
면적	-0.00006***	0.00002	3.42367	-0.00005***	0.00002	3.40963	-0.00005***	0.00002	3.04905	-0.00005***	0.00002	3.05125	
토지이용	전	0.10841***	0.04277	2.53511	0.07194	0.03702	1.94315	0.06903**	0.03869	1.78431	0.07152**	0.03823	1.87085
	답@												
도로접면	소로이상	0.5704***	0.07839	7.27616	0.53179***	0.06787	7.83556	0.53302***	0.07285	7.31719	0.53182***	0.07131	7.45775
	세로가	0.19354***	0.04904	3.94688	0.17924***	0.04250	4.21774	0.18699***	0.04555	4.10568	0.19131***	0.04468	4.28179
	세로불	0.21946***	0.07134	3.07634	0.22874***	0.06174	3.70513	0.24456***	0.06685	3.65831	0.24262***	0.06523	3.71952
	맹지@												
토지형상	정형	0.16935***	0.06193	2.73465	0.13002***	0.05361	2.42537	0.06801*	0.05624	1.20928	0.07552**	0.05550	1.36074
	부정형@												
개발제한구역	-0.89989***	0.06635	13.56377	-0.49246***	0.06327	7.78305	-0.86532***	0.08720	9.92289	-0.67237***	0.09871	6.81145	
개발제한구역 외@													
농업진흥지역	0.26587***	0.05699	4.66535	0.19527***	0.04938	3.95444	0.18828***	0.05923	3.17868	0.19022***	0.05616	3.38715	
농업진흥지역 외@													
2차로 거리	-0.00285***	0.00022	12.71126	-0.00207***	0.00020	10.37040	-0.00276***	0.00024	11.44749	-0.0024***	0.00025	9.49341	
4차로 거리	-0.00002***	0.00001	2.97774	0*	0.00001	0.92959	-0.00001*	0.00001	1.24116	-0.00001*	0.00001	1.13803	
도시와의 거리	-0.00002***	0.00001	4.95640	-0.00001***	0.00001	2.20718	-0.00003***	0.00001	3.24297	-0.00002***	0.00001	2.85470	
주거지와의 거리	-0.00005***	0.00002	3.24203	-0.00004***	0.00001	3.06605	-0.00006***	0.00002	2.80787	-0.00004***	0.00002	2.40107	
공업지와의 거리	0	0.00001	0.32461	0.00001*	0.00001	1.28016	0.00002*	0.00002	1.26869	0.00002**	0.00001	1.51780	
주변 5km 인구 수	0*	0.00000	1.12384	0**	0.00000	1.65933	0**	0.00000	1.63141	0**	0.00000	1.64872	
주변 2km 인구 수	0**	0.00001	1.54910	0	0.00001	0.63261	0*	0.00001	0.90031	0	0.00001	0.83590	
주변 2km 토지거래량	0.00012	0.00027	0.47524	0.0002*	0.00023	0.87910	0.00032*	0.00034	0.95248	0.00024	0.00030	0.83351	
rho	0			0.86334***	0.04694	18.39304	0			0.54403***	0.15178	3.58439	
lamda	0						0.95252***	0.02742	34.73196	0.69078***	0.15157	4.55760	
R <sup>2</sup>	0.62592			0.71318			0.7198			0.72191			
log-likelihood	-589.65202			-513.26512			-511.0706			-500.69837			
AIC	1215.30403			1022.86347			1005.78749			1002.24129			
SC	1298.02808			1110.1833			1093.10732			1094.1569			
RMSE	0.5415			0.47416			0.46866			0.46689			

주: 1) \*\*\* 유의수준 0.01, \*\* 유의수준 0.05, \* 유의수준 0.1  
 2) @는 기준변수를 의미함.

### 3) GWR 모형 추정결과

본 논문에서는 <표 6>과 같이 Gaussian, Exponential, Tricube 가중치의 GWR모형을 모두 추정하였다. CV 값을 최소로 하는 neighbor의 수 85를 bandwidth로 하고 Tricube 가중치를 적용하는 방식과 0.84071, 2.14492를 각각 bandwidth로 설정한 Gaussian 및 Exponential 가중치를 적용하는 방식을 각각 실시하였다. 그 결과 Tricube 방식의 GWR모형이 Gaussian 및 Exponential 가중치 방식보다 RMSE가 낮음은 물론이고 OLS, SAR, SEM, SAC 등과 비교하였을 때도 RMSE가 가장 낮은 것으로 나타났다. 구체적으로 보면, Tricube 방식 GWR모형은 다른 모든 모형과 비교했을 때 가장 높은 R<sup>2</sup>, Log-likelihood와 가장 낮은 AIC, SC, RMSE를 나타내고 있다. 특히 OLS모형과 비교했을 때, RMSE는 약 30% 정도 낮아지는 것으로 나타났다. Gaussian과 Exponential 방식의 GWR모형의 경우 R<sup>2</sup>, Log-likelihood, AIC, SC, RMSE가 OLS 모형과 거의 비슷하게 나타났다.<sup>9)</sup> 즉, 본 연구를 통해서 OLS, SAR, SEM, SAC모형에 비해 Tricube 가중치의 GWR모형이 수도권 근교농지의 가격추정에

가장 바람직한 모형이라 판단할 수 있을 것이다.

OLS, SAR, SEM, SAC와 같이 각 독립변수에 대한 회귀계수가 모형의 지역범위 내에서 동일한 값을 갖는다고 가정하는 모수적 접근방법과 달리 GWR모형은 비모수적 접근방법의 일종으로 관측치별로 모형을 추정하게 되어 앞서 제시한 이들 모형들과 변수의 영향력을 직접적으로 비교하는 데 어려움이 있다. 대신 GWR의 경우 <그림 2>와 같이 관측치별 회귀계수의 공간적 분포를 확인할 수 있으므로 모형의 대상지역을 상세화하여 각 독립변수가 지가에 미치는 영향력을 분석할 수 있다.

<그림 2>를 구체적으로 살펴보면, 우선 동일한 범주로 구분한 OLS와 GWR모형의 잔차분포를 비교한 결과 OLS에서는 지역별 공간상관이 뚜렷한 반면에 GWR모형에서는 공간적 종속성이 상당히 완화되어 잔차의 분포가 랜덤한 경향을 보이는 것으로 나타났다. 도로와의 접근성에서는, 2차로와의 거리 변수의 회귀계수는 거의 모든 지역에서 상대적으로 큰 음의 값을 갖으며, 특히 광주시내에서 거리가 멀어질수록 음의 계수 값이 커지는 것으로 나타났다. 반면 4차로와의 거리는 광주시내 주변지역에서 양

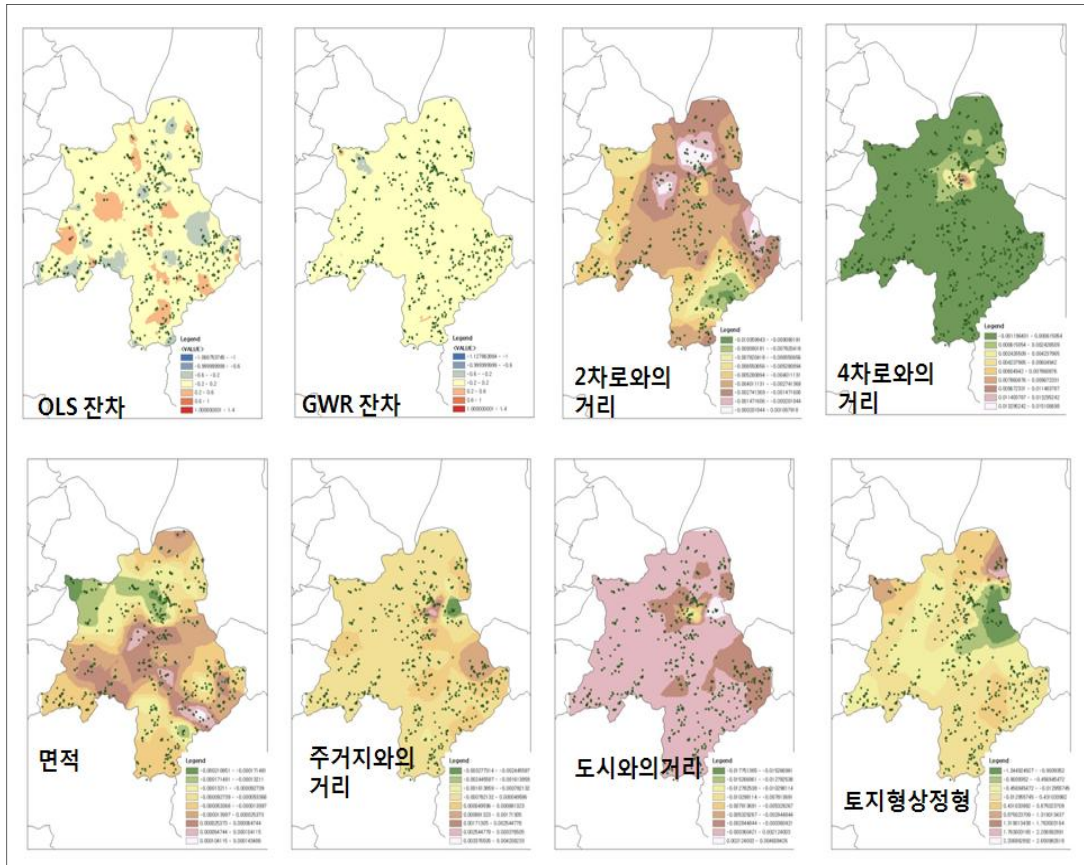
의 계수 값을 갖는데 반해 대부분의 지역에서 0과 비슷한 매우 작은 계수 값을 갖는 것으로 나타났다. 면적의 경우 기존 도시지역 인근지역의 회귀계수가 양의 값을 갖는 것으로 나타나 대규모개발의 상대적 용이함으로 인한 프리미엄으로 해석할 수 있을 것이다. 도시와의 거리의 경우 도시와 거리가 먼 지역일수록 음의 계수 값이 커지는 것으로

표 6 \_GWR - Gaussian, Exponential, Tricube 추정결과

구분	Gaussian	Exponential	Tricube
bandwidth	0.84071	2.14492	85
R <sup>2</sup>	0.6247758	0.6455406	0.8182815
Adj R <sup>2</sup>	0.61638	0.63761	0.81422
ML Variance	0.29412	0.27785	0.14244
Log likelihood	-590.76730	-569.93090	-325.39350
AIC	1217.53500	1175.86200	686.78700
SC	1300.25900	1258.58600	769.51110
RMSE	0.54233	0.52711	0.37741

9) 일반적으로 GWR모형은 OLS모형을 각 관측치별로 추정한 것으로 Least Square라는 동일한 추정방법을 사용하게 됨. 그러나 모형의 지역범위 내에서 1개의 모형을 기준으로 R<sup>2</sup>을 계산하는 OLS모형과 달리 GWR모형은 각 관측치별 개별모형에 대하여 R<sup>2</sup>을 계산하게 되어 R<sup>2</sup>을 기준으로 양자 간의 모형적합성을 절대적으로 비교하는 데 한계가 있음.

그림 2\_OLS와 GWR 잔차분포 비교 및 GWR모형(tricube) 회귀계수의 공간적 분포



나타났으며, 주거지와의 거리는 도시와의 거리가 가까운 지역일수록 음의 계수 값을 가지는 것으로 나타났다.<sup>10)</sup> 토지형상정형의 경우 대부분의 지역에서 지가영향력이 낮은 것으로 보였지만 남한산성 주변과 남한강 주변지역의 경우 양의 계수 값이 뚜렷하게 나타났다.

<표 7>은 OLS모형의 전역적인 계수추정치와 Tricube 가중치 GWR모형의 국지적인 계수추정치를 비교하기 위해서 관측치별 계수의 표준편차, 최댓값, 최솟값과 이들의 평균 및 평균의 표준편차를 나타내

고 있다. 이를 보다 시각적으로 표현하기 위해서 주요 계수추정치의 분포도를 그래프로 나타내면 <그림 3>과 같다. 여기서 특이한 점은 2차로와의 거리변수의 경우 GWR의 계수가 OLS모형의 계수보다 전반적으로 작게(-) 나타났으며, 이는 접근성이 지가에 미치는 영향이 GWR모형이 OLS모형보다 크게 나타나는 것으로 볼 수 있어 현재 OLS모형에 의한 비준표작성방식에서 접근성에 대한 가중치보정은 개선될 여지가 있다고 생각될 수 있다는 것이다.

GWR모형에서 유의할 점은 윈도우 크기 즉 bandwidth

10) 도시와의 거리변수의 회귀계수는 도시와의 거리가 멀거나 산악지역을 중심으로 도시와 1m 멀어질수록 농지가격이 약 0.03~0.12% 감소하는 지역들이 분포한 반면 주거지와의 거리변수는 도시와의 거리가 가까운 지역에서 주거지와 1m 멀어질수록 농지가격이 약 0.02~0.04% 감소하는 지역들이 분포하는 것으로 분석되었음.

표 7\_OLS와 GWR모형의 계수추정 비교

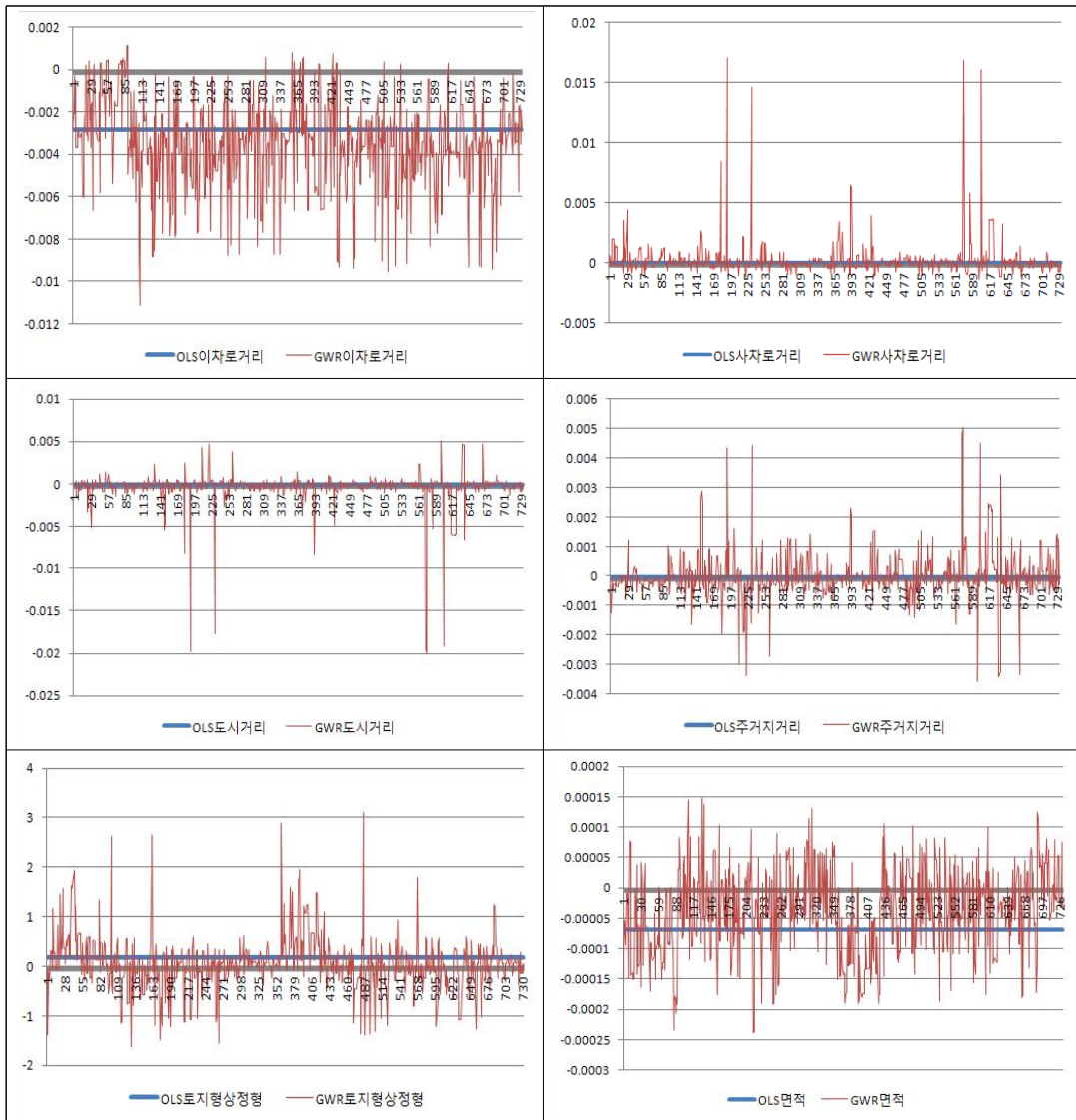
구분	OLS(semi-log)				GWR-Tricube				
	Coeff	Std.Dev	N	Mean	MeanStd	Std.Dev	Min	Max	
상수항	12.24844***	0.14702	732	16.33000	6.53600	19.86000	-57.52000	189.40000	
면적	-0.00006***	0.00002	732	-0.00004	0.00007	0.00007	-0.00024	0.00015	
토지 이용 전	0.10841***	0.04277	732	0.03304	0.12810	0.18950	-0.35570	0.71980	
도로 접면	도로접면 소로	0.5704***	0.07839	732	0.53390	0.30120	0.39030	-0.55290	2.12000
	접면세로가	0.19354***	0.04904	732	0.11740	0.15690	0.25970	-0.87480	0.61890
	접면세로불	0.21946***	0.07134	732	0.11530	0.57670	0.33760	-0.75020	0.94630
토지형상정형	0.16935***	0.06193	724	0.07717	0.53250	0.54770	-1.62800	3.11500	
개발제한구역	-0.89989***	0.06635	402	-0.34980	2.24900	0.70500	-1.85400	1.84500	
농업진흥지역	0.26587***	0.05699	699	0.33630	0.92910	0.52230	-1.22800	2.75400	
2차로거리	-0.00285***	0.00022	732	-0.00351	0.00126	0.00213	-0.01112	0.00114	
4차로거리	-0.00002***	0.00001	732	0.00023	0.00038	0.00160	-0.00135	0.01708	
도시거리	-0.00002***	0.00001	732	-0.00035	0.00044	0.00198	-0.02008	0.00508	
주거지거리	-0.00005***	0.00002	732	-0.00002	0.00033	0.00082	-0.00354	0.00505	
공업지거리	0	0.00001	732	-0.00014	0.00028	0.00057	-0.00452	0.00121	
주변5km인구	0*	0.00000	732	0.00000	0.00005	0.00012	-0.00078	0.00043	
주변2km인구	0**	0.00001	732	0.00002	0.00011	0.00020	-0.00056	0.00101	
토지거래량	0.00012	0.00027	732	-0.00282	0.00411	0.00799	-0.04230	0.02221	
Optimal bandwidth					85				
R-square	0.6259				0.8182815				
Adjusted R-square					0.81422				
ML Variance					0.14244				
Log likelihood	-589.7				-325.39350				
AIC	1215.3				686.78700				
SC	1298.0				769.51110				
RMSE	0.5415				0.37741				

주: \*\*\* 유의수준 0.01, \*\* 유의수준 0.05, \* 유의수준 0.1

내에 포함되는 표본의 더미변수 값이 모두 동일한 지역의 경우 더미변수와 상수항이 서로 선형관계를 갖기 때문에 다중공선성의 문제가 발생한다는 점이 다(박현수, 2000). 특히 공간적으로 덜 조밀(dense)하게 분포되고, bandwidth를 절대거리로 고정한 fixed kernel 방식이며, 폴리곤을 공간단위로 취합된 집계(aggregated)데이터인 경우가 그렇지 않은 경우에 비

해서 GWR모형 추정 시 extreme coefficients 현상이 더욱 심화되는 경향이 있다(Cho et al. 2009). 본 연구에서는 OLS, SAR, SEM, SAC, GWR모형의 적합성 비교를 위하여 동일한 변수를 사용하여 추정하였지만, Tricube 가중치 GWR모형의 경우 토지형상정형, 개발제한구역, 농업진흥지역 변수에서 공간적 분포가 균일하지 않아 이러한 다중공선성 문제가 발생하

그림 3\_GWR 계수추정치치의 분포도



여 이들 관측치에서는 다중공선성이 발생하는 변수를 제외하고 모형을 추정하였다.11)

### V. 결론

본 논문에서는 토지정책수단의 주요 대상이 되는 농

지에 대한 가격추정 시 세부지역별로 발생될 수 있는 공간적 효과를 고려하지 않는 전통적 회귀모형의 한계를 극복하고자 공간계량경제 접근방법의 일환으로 SAR, SEM, SAC, GWR모형으로 수도권 근교농지의 가격을 추정하였다. 경기도 광주시 비시가화지역 실거래농지 732필지에 대한 실증분석 결과 전통

11) 토지형상정형은 724개, 개발제한구역은 422개, 농업진흥지역은 699개 관측치에서만 이들 변수를 포함하여 모형을 추정하였음.

적인 모수적 추정방법인 OLS모형에 비해서 비모수적 추정방법인 GWR모형의 지가예측력이 가장 뛰어난 것으로 나타났다. OLS모형보다 Tricube 방식 GWR모형이 약 30% 정도 RMSE가 낮아지는 것으로 나타났다. 한편 종속변수와 오차의 공간효과를 고려하여 공간가중치행렬과 결합하여 회귀분석에 포함시킨 형태의 공간계량경제모형인 SAR, SEM, SAC모형 추정결과 RMSE는 각각 0.47416, 0.46866, 0.46689로 세 가지 모형 모두 OLS모형의 0.5415보다 낮은 것으로 나타났다.

GWR모형의 경우 공간중속성이 확인되었기 때문에 OLS모형에 비해 모형적합도가 개선된 것은 당연한 결과라고 판단되며, 이는 OLS와 GWR 모두 동일한 Linear Least Square Regression방식으로 추정하기 때문에 공간중속성이 존재하는 상황에서는 관측치별로 선형회귀식을 추정하는 GWR이 전체지역에 대해 하나의 선형회귀식을 추정하는 OLS보다 모형의 적합도가 높게 나타났다.

근래 도시화와 인구증가의 속도가 둔화되면서 토지정책의 방향이 「토지보상법」<sup>12)</sup>에 의거하여 강제토지수용이 수반되는 대규모 공영개발 중심의 직접적이고 단기적인 토지정책수단에서 토지은행, 농지은행 등 필지단위의 간접적이고 장기적인 토지정책수단 중심으로 변화하고 있는 상황이다. 단기적인 개발이익의 배제가 필요한 직접적인 토지정책수단과 달리 간접적인 토지정책수단은 장기적인 토지수급 및 관리가 목적이므로 매수 및 매각을 원활하게 추진하기 위해서는 현재시가를 기준으로 적정하게 가격이 산정되어야 할 것이다. 그러나 현재 토지은행, 농지은행, 개발제한구역, 수변구역, 사유림의 토지매수 및 국공유지 토지매각의 경우 실무적으로 주

변 1개의 표준지공시시가를 기준으로 하는 토지보상평가지침에 의거하여 예정가격을 결정하고 있다. 따라서 향후에는 이들 예정가격의 결정 및 가격사정 시 실거래가격자료를 이용하여 GWR모형을 적용하는 방안을 모색하는 것이 필요할 것으로 사료된다. 특히 현행 공적지가평가체계에 대한 안정성을 유지하면서 공시지와 실거래가 간의 괴리로 인하여 발생하는 사유재산권 침해, 조세형평성, 국가 및 지자체 자산가치 저평가 및 미활용 문제를 해결하기 위해서는 이를 감정평가 시 기타요인보정계수 산정의 근거자료 또는 참고자료로 활용하는 방안을 제안하고자 한다. 감정평가 시 표준지공시시가를 제외하고 기타요인의 감정가격에 대한 영향력이 가장 높으며, 다른 지목에 비해 임, 전, 답의 경우 더욱 영향력이 높기 때문에(오영선 외, 2009) 농지에 대한 공적지가 산정 시 실거래가를 이용한 GWR모형의 적용이 필요할 것이다. 나아가 장기적으로는 개별공시지가 산정 시 1개 이상의 복수 비교표준지를 선정하여 GWR모형 등을 적용하는 방안에 대한 타당성 검토 및 연구가 필요할 것이라고 생각된다.

본 연구에서는 OLS, SAR, SEM, SAC, GWR모형의 적합도를 비교하기 위해서 모든 모형에 동일한 변수를 포함하여 지가를 추정하였으나, 비모수적 방법인 GWR모형을 세부지역별로 유의성이 높은 변수만을 선정하여 지가를 추정할 경우 모형의 적합도가 더 높아지는 것으로 나타났다. 따라서 향후 연구에서는 보다 다양한 변수 중에서 세부지역별 또는 관측치별로 유의성이 높은 변수만을 포함하고 다중공선성이 발생하는 변수를 제외하는 등의 방식으로 GWR모형에 대한 다양한 반복 시뮬레이션이 필요할 것으로 판단된다.

12) 「공익사업을 위한 토지 등의 취득 및 보상에 관한 법률」의 줄임말이며, 동법에 의하여 당해개발사업에 의한 지가상승분을 제외 토록 하는 개발이익배제원칙이 적용된 토지보상평가지침이 보상가격의 산정근거가 됨.

또한 연구대상지역에 서울근교지역 이외에 중소 도시근교지역과 순수농촌지역을 추가하여 지역별 모형적합도 비교도 필요할 것이다. 나아가 본 연구결과와 연계하여 실거래가를 이용한 토지가격지수산정방안을 수립함으로써 다양한 하위시장구분 등 토지시장분석의 준거틀을 제시하고 토지시장정보의 비대칭성을 완화함으로써 정상적인 거래활성화는 물론이고 정부의 효과적인 토지정책 수립이 가능하게 되어 토지시장안정화를 도모할 수 있으리라 기대된다.

참고문헌 •••••

김성우·정건섭. 2010. “공간계량경제모형에서의 실제거리를 반영한 공간가중행렬에 관한 연구. 부산아파트 실거래가를 중심으로”. 주택연구 제18권 제4호. pp59-80.

김익희. 2005. “GWR을 이용한 지가설명함수의 공간적 이질성 (Spatial Heterogeneity) 탐색: 서울시 서초구 강남구 송파구를 사례로”. 서울대학교 석사학위논문.

김진서·충원. 2009. “오피스 임대료 추정에 있어서 공간자기상관에 관한 연구”. 국토계획 제44권 제2호 pp95-110.

민태욱. 2005. “적정가격보상원칙과 시가보상론”. 부동산학연구 제11집 제2호 pp15-32.

박현수. 2000. “모수적방법과 비모수적방법에 의한 토지가격 추정에 관한 연구”. 한국지역개발학회지 제12권 제1호 pp101-109.

박현수·안지아. 2005. “공간종속성을 이용한 아파트가격의 공간효과에 관한 연구”. 부동산학연구 제11집 제2호 pp119-128.

박현수·정수연·노태욱. 2003. “공간계량경제모형을 이용한 아파트 가격과 공간효과분석”. 국토계획 제38권 제5호 pp115-125.

서경찬·이성호. 2001. “공간적 자기회귀모델과 토지시장분할에 의한 효율적 지가추정에 관한 연구”. 국토계획 제36권 제4호.

서교. 2005. “헤도닉분석기법과 공간계량경제모형을 이용한 농촌 지역 지가의 영향인자 분석”. 한국농촌계획학회지 제11권 제3호. pp11-17.

오영선·유선중. 2009. “감정평가 가격형성요인의 기타요인에 관한

연구”. 부동산연구 제19집 제1호. pp127-150.

이성우·윤성도·박지영·민성희. 2006. 공간계량모형응용. 서울 : 박영사.

이용범. 1998. “임지의 가격형성요인에 관한 연구”. 부동산학연구 제4집. pp57-71

이정진. 2008. 토지경제학. 서울 : 박영사

최열·이백호. 2006. “공간자기상관과 주변 용도지역에서 접근성을 고려한 주거지내 지가추정에 관한 연구”. 국토계획 제41권 제5호. pp45-60.

허윤경. 2007. “도시별 주택가격의 공간적 영향력 검증: 서울과 부산의 아파트가격을 중심으로”. 주택연구 제15권 제4호 pp5-23.

Anselin, L. 1988. *Spatial Econometrics: Method and Models*. Dordrecht : Kluwer Academic.

\_\_\_\_\_. 2005. *Exploring Spatial Data with Geoda: A Workbook*. Santa Barbara CA : Center for Spatially Integrated Social Science.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E. 1996. “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity”. *Geographical Analysis* vol.28, no.4. pp281-298.

Cho, S. H., Lambert, D. M., Kim, S. G. and Jung, S. H. 2009. “Extreme Coefficients in Geographically Weighted Regression and Their Effects on Mapping”. *Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting*. Milwaukee, Wisconsin. July 26-29. 2009

Cleveland, W. S. and Devlin, S. J., 1988. “Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting”. *Journal of the American Statistical Association*, vol. 83. no.403. pp596-610.

Cotteleer, G., Gardebreek, C. and Luitj, J. 2008. “Market Power in a GIS-Based Hedonic Price Model of Local Farmland Markets”. *Land Economics* vol.84. no.4. pp573-592.

Dubin, R. 1988. “Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms”. *Review of Economics and Statistics* vol.70. pp466-474.

\_\_\_\_\_. 1992. “Spatial Autocorrelation and Neighborhood Quality”. *Regional Science and Urban Economics* vol.22 pp433-452.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. E. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. UK : John Wiley & Sons.

- Gillen, K., T. G. Thibodeau, and S. Wachter. 2001. "Anisotropic Autocorrelation in Housing Prices". *Journal of Real Estate Finance and Econometrics* vol.23 no.1. pp5-30.
- Jeanty, P. W., Kraybill, D. S., Libby, L. W., and Sohngen, B. 2002. "Effects of Local Development Pressure on Land Prices: A Spatial Econometric Approach". *AAEA Meeting*. 2002.
- Kang, D. W., Kim, M. Y., Cho, D. H., and Lee, S. W. 2010. "The effects of Urban Development Pressure on Agricultural Land Price; Application of a Mixed GWR Model". *Journal of Rural Development* vol.33, no.4. pp63-83.
- Kumar, P., Pardhan, B. K. and Subramanian, A. 2005. "Farmland Prices in a Developing Economy: Some Stylised Facts and Determinants". *Journal of International and Area Studies* vol.12, no.2. pp93-113.
- Pace, K. 1997. "Performing Large Spatial Regressions and Autogressions". *Economics Letters* vol.54. pp283-291.
- Patton, M. and McErlean, S. 2003. "Spatial Effects within the Agricultural Land Market in Northern Ireland". *Journal of Agricultural Economics* vol.1, issue.1. pp35-54.
- Plantinga, A. J., Lubowski, R. N. and Stavins, R. N. 2002. "The Effects of Potential Land Development on Agricultural Land Prices". *Journal of Urban Economics* vol.52. pp561-581.
- Robinson, W.S. 1950. "Ecological Correlations and the Behavior of Individuals". *American Sociological Review* vol.15, no.3. p351-357.
- Tobler, W. 1970. "A Computer Movie Simulating Urban Growth in the Detroit region". *Economic Geography* vol.46, no.2 pp234-240.
- Yrigoyen, C. C., Rodriguez, I. G. and Otero, J. V. 2006. "Modelling Spatial Variations in Household Disposable Income with Geographically Weighted Regression". *Instituto L.R. Klein-Centro Gauss. U.A.M. D.T.* no.15.
- <http://personal.rhul.ac.uk/uhte/006/cc2203/index2.html>.

- 
- 논문 접수일: 2012.1.11
  - 심사 시작일: 2012.1.18
  - 심사 완료일: 2012. 3. 7

## A Study on the Estimation of Farmland Price Using Spatial Econometrics Approach : Focused on Urban Fringe in Seoul Metropolitan Area

**Keywords:** Spatial Econometrics, Geographically Weighted Regression,  
Farmland Price, Landbanking, Public Land Disposal

This study aims at identifying a parcel based land price estimation model, considered spatial dependency and heterogeneity, focused on the farmland which is the main target of land policy as a future urbanized land. In addition to ordinary least square estimation like hedonic price model, spatial econometrics approaches such as spatial auto regression model, spatial error model and geographically weighted regression model were used to estimate the farmland price at the urban fringe in Seoul Metropolitan Area and compared. The empirical results showed that GWR model weighted tricube adaptive kernel has about 30% lower RMSE than OLS model. In the future, it is strongly suggested to apply transaction price based GWR models for public purpose land price estimation such as landbanking and public land disposal, both indirect and long-term land policy, especially when calibrating other price factors.

### 공간계량경제 접근방법을 이용한 농지가격추정에 관한 연구 : 수도권 근교농지를 중심으로

**주제어:** 공간계량경제모형, 공간가중회귀모형, 농지가격, 토지은행, 국공유지 처분

전통적인 농업지대론 중심의 토지특성항목구성과 1개 비교표준지 기준의 지가산정방식으로 인하여 농지의 경우 다른 토지에 비해 공시지가와 실제가격간의 괴리가 특히 심하다. 본 연구는 도시화예정용지로서 직간접적 토지정책수단의 실질적인 집행대상인 농지에 대한 가격평가기준 마련을 목적으로 한다. 세부적으로는 개발가능성과 접근성 등 도시지대론을 고려하고, 공간적 의존성 및 이질성을 고려한 필지단위 토지특성가격모형을 도출하였다. 이를 위해 전통적 OLS 회귀모형과 함께 SAR, SEM, SAC, GWR과 같은 공간계량경제 접근방법으로 각각 수도권 근교 농지가격을 추정하여 그 결과를 비교하였다. 실증분석결과 OLS모형보다 Tricube 가중치 GWR모형은 약 30% 정도 RMSE가 낮아지는 것으로 나타났다. 현재 실무적으로 토지보상평가지침을 기준으로 예정가격이 결정되는 토지은행, 국·공유지 등 공공토지 매수·매각의 경우, 향후에는 정상적인 시장가치 반영을 위하여 감정평가시 실거래가격을 이용한 GWR모형을 기타요인보정계수 산정의 근거자료로 활용하는 방안을 제안하고자 한다.