# 時空間における地価形成要因影響度の変遷・差異を考慮した普遍クリギングによる地価内挿の実証的検討

川松　祐太[1]・井上　亮[2]

[1]正会員　株式会社福山コンサルタント 東北事業部 (〒980-0820 仙台市青葉区二日町13-17)
E-mail:  y.kawamatsu@fukuyamaconsul.co.jp

[2]正会員　東北大学准教授　大学院情報科学研究科人間社会情報科学専攻　(〒980-8577 仙台市青葉区片平2-1-1)
E-mail: rinoue@plan.civil.tohoku.ac.jp

本研究は，不動産市場透明性向上の一手法として，任意地点・時点における地価情報提供を目指し，時空間で蓄積された公示地価を活用した地価内挿に着目する．地価は一般に，近隣地点や周辺時期のデータに対して相関を有し，また地価形成要因の場所・時期による相違の影響を受ける．これらの性質を持つデータの分析法として，空間データに対しては地理的加重回帰(GWR)と普遍クリギング(UK)を統合したGWRKが提案されている．本研究ではこのGWRKを時空間へ拡張した上で，時空間GWR, UKによる公示地価内挿の精度を比較し，GWRKは精度の高い公示地価内挿を実現することを確認した．

***Key Words :*** *land price, spatio-temporal interpolation, geographically weighted regression, universal kriging,*

## 1. はじめに

近年，我が国では不動産市場の活性化と透明性向上を目指し不動産価格に関わる情報の整備と公開が進められてきた．しかし不動産投資家や個人買主・売主などからは未だに「不動産市場の動向を把握する為の良質な情報が不足している」という, 質的な問題が指摘されている．

現在，一般の不動産市場参加者が利用可能な，国が公開する不動産価格情報には, 公示地価などの「公的地価指標」と「不動産取引価格情報」がある. しかし現在，公開されている情報だけから不動産市場の動向を把握する事は難しい．公的地価指標は，標準的地価を提供しているが, 公表地点・時点数が限られ, 短期的な価格変動を伝えられない点が問題とされている．また，不動産取引価格情報は，短期的な市場動向を提供するが，取引点の偏在や標準的な価格水準の把握が困難である事が問題として挙げられる．しかし，この２種類の不動産価格情報をもとに，すべての市場参加者が任意地点の標準的な価格を知ると同時に，その近隣の短期的な市場価格変動を確認できる環境が整備されれば，不動産市場の透明性向上に大きく寄与するであろう．

そこで本研究では，時空間で蓄積された公的地価指標を利用し，任意の地点・時点での標準的地価を情報提供する不動産価格情報の提供へ向け，時空間内挿の精度を向上する事を目指す．

ところで，これまで地価分析には回帰分析が多く用いられてきた．しかし通常の回帰分析では，説明変数の影響力が行政区など地域によって異なるという性質（以下，地域特性と呼ぶ）や，地価関数の攪乱項に生じる空間相関を考慮した分析ができない[1),2)]．

計量地理学では，地域特性の分析モデルとしてGeographically Weighted Regression(GWR)が提案されている[3)]が，GWR ではモデルの攪乱項の空間相関は解消できない．また，空間統計学では，空間相関を持つデータの分析手法として普遍クリギング(UK: Universal Kriging)が体系化されているが, 公示地価データに対して UK を適用すると，残差の分散が地域により大きく異なる事が確認されている[4)]．

Fotheringham らは，GWR, UK の統合モデル GWRK を提案している[5)]．Harris らは，GWR, UK, GWRK の空間内挿精度を，空間相関・地域特性のあるシミュレーションデータを用いて比較し，対象点近隣にデータが少なく，モデル推定に広域・離散的に点在するデータ群を用いる必要がある場合は GWRK が比較的高精度の内挿が可能であると示している[6)]．

公的地価指標は，地域特性や空間相関を持つデータであり，また，地点の間隔が広い“粗”な情報である. この地価情報を用いて高精度の内挿を行うには, 地域特性や空間相関を考慮できる GWRK が適当であろう．また，公的地価指標は時系列に退位しても

相関を有するため，多時点のデータを用いると高精度の内挿が可能になると期待される．ただし，時間によっても説明変数の影響度は変わる(時間特性を持つ)ため，地域特性と同様に対処が必要である．

そこで本研究は，地域特性や空間相関を考慮した内挿を行う GWRK に着目し, 時空間への拡張を行う. さらに公示地価データを用いた GWR, UK, GWRK の内挿精度比較を通して, 地価データへの GWRK の適用可能性を実証的に検討する.

## 2. 地域･時間特性や時空間相関を考慮した内挿

時間・地域特性や時空間相関を持った情報の内挿手法であるGWR, UK, GWRKについて記す.

### (1) GWR

GWR は，ある点 $i$ の近傍に重み付けしたモデルを用いた通常最小二乗推定(OLS)を行い，点 $i$ 独自の地域特性を表すパラメータを推定する手法である．いま，内挿点 $i$ を中心としたモデルによる時空間内挿を考える．ここで，第 $j$ 対角要素 $w_{ij}$ を，点 $ij$ 間の距離 $d_{ij}$・時間差 $t_{ij}$ を用いてガウス型カーネル関数

$$w_{ij} = \exp\left(-\left(d_{ij}/\delta_d\right)^2 - \left(t_{ij}/\delta_t\right)^2\right) \tag{1}$$

$\delta_d$: 空間バンド幅, $\delta_t$: 時間バンド幅

で与える対角重み行列 $\mathbf{W_i}$ を用い, モデルを設定する.

$$\mathbf{W_i}^{1/2}\mathbf{y} = \mathbf{W_i}^{1/2}\mathbf{X}\boldsymbol{\beta}_\mathbf{i} + \mathbf{u_i} \quad \mathbf{u_i} \sim N\left(0, \sigma^2\mathbf{I}\right) \tag{2}$$

$\mathbf{W_i}^{1/2} = diag\left[w_{i1}^{1/2}, ..., w_{in}^{1/2}\right]$，$\mathbf{y}$: 被説明変数ベクトル, $\mathbf{X}$: 説明変数行列, $\boldsymbol{\beta}$: パラメータベクトル, $\mathbf{u}$: 攪乱項ベクトル, $\sigma^2$: 分散

このとき，内挿点 $i$ の最良線形不偏予測量(BLUP)は

$$\hat{y}_i = \mathbf{x_i}'\left(\mathbf{X}'\mathbf{W_i}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W_i}\mathbf{y} \tag{3}$$

と求まる．ただし，最適な時空間バンド幅は交差検定を通して推定する [6]．なお本研究では，空間・時間のバンド幅は，交互に交差検定を行い推定する．

$$\hat{\delta} = \arg\min_{\delta}\sum_{j=1}^{n}\left[y_j - \hat{y}_{(j)}\left(\delta\right)\right]^2 \tag{4}$$

$\hat{y}_{(j)}\left(\delta\right)$ ：観測点$j$を除くデータを用いた$y_j$の内挿値

### (2) UK

UKは，空間確率場に二次定常性を仮定し，モデルの攪乱項間の共分散を距離の関数として空間相関を構造化し，任意地点での確率場の値を予測する手法である．時空間相関を扱う際には，時間と距離の関数を用いて構造化する．本研究では，空間には等方性を仮定した上で，異方性パラメータ $\phi$ を用いて時間軸を空間軸の一つと考える簡易的な指数型時空間共分散関数(式(6))を用いる.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad \mathbf{u} \sim N\left(0, \mathbf{V}\right) \tag{5}$$

$$V_{ij}\left(h_{ij}\right) = \begin{cases} \theta_0 + \theta_1 & \text{if } h_{ij} = 0 \\ \theta_1\left(\exp\left(\dfrac{-h_{ij}^2}{\theta_2^2}\right)\right) & \text{if } h_{ij} > 0 \end{cases} \tag{6}$$

$\theta_0 + \theta_1$: sill, $\theta_0$: nugget, $\theta_2$: range, $h_{ij} = \sqrt{d_{ij}^2 + \phi^2\tau_{ij}^2}$

UKでは，内挿点$i$のBLUPは式(7)で表される.

$$\begin{aligned} \hat{y}_i &= \mathbf{x_i}'\hat{\boldsymbol{\beta}} + \mathbf{c}'\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right) \\ \hat{\boldsymbol{\beta}} &= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \end{aligned} \tag{7}$$

$\mathbf{c}$：内挿点と観測点の攪乱項間共分散ベクトル

共分散関数のパラメータ$\boldsymbol{\theta}(\theta_0, \theta_1, \theta_2)$は，式(8)の重み付き最小二乗基準(WRSS)を用いて, 式(9)の経験セミバリオグラム $\hat{\gamma}\left(h\right)$ に式(7)の理論セミバリオグラム $\gamma\left(h;\boldsymbol{\theta}\right)$ が適合するように推定する.

$$WRSS = \sum\frac{N\left(h\right)}{\hat{\gamma}\left(h\right)^2}\left(\gamma\left(h;\boldsymbol{\theta}\right) - \hat{\gamma}\left(h\right)\right)^2 \tag{8}$$

$$\hat{\gamma}\left(h\right) = \frac{1}{2N\left(h\right)}\sum_{N(h)}\left(e_i - e_j\right)^2 \tag{9}$$

$$\gamma(h;\boldsymbol{\theta}) = \begin{cases} \theta_0 + \theta_1\left[1 - \exp\left(-h^2/\theta_2^2\right)\right] & h > 0 \\ 0 & h = 0 \end{cases} \tag{10}$$

$N(h)$: 距離$h$の内挿・観測点数，$e_i$: $i$点の残差

### (3) GWRK

GWRKは, GWRと同様に対角重み行列 $\mathbf{W_i}$ を用いて点$i$の近傍に重み付けしたモデルの攪乱項に対して, UKと同様の二次定常性を仮定し内挿を行う手法である．内挿点$i$を中心としたモデルは

$$\mathbf{W_i}^{1/2}\mathbf{y} = \mathbf{W_i}^{1/2}\mathbf{X}\boldsymbol{\beta}_\mathbf{i} + \mathbf{u_i} \quad \mathbf{u_i} \sim N\left(0, \mathbf{V_i}\right) \tag{11}$$

と表され，内挿点の BLUP は以下のように求まる.

$$\begin{aligned} \hat{y}_i &= \mathbf{x_i}'\hat{\boldsymbol{\beta}}_\mathbf{i} + \mathbf{c_i}\mathbf{V_i}^{-1}\mathbf{W_i}^{1/2}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_\mathbf{i}\right) \\ \hat{\boldsymbol{\beta}}_\mathbf{i} &= \left(\mathbf{X}'\mathbf{W_i}^{1/2}\mathbf{V_i}^{-1}\mathbf{W_i}^{1/2}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W_i}^{1/2}\mathbf{V_i}^{-1}\mathbf{W_i}^{1/2}\mathbf{y} \end{aligned} \tag{12}$$

なお先行研究[6]では，GWRによるパラメータ推定と, 通常クリギング(Ordinary Kriging)によるその残差の内挿結果を足した値を，GWRKによる内挿値としている．この方法は，パラメータ推定では攪乱項の共分散0を仮定し, その残差には二次定常性を仮定するという極めて不合理な推定を行っており，問題が多いことをここで指摘しておく.

## 3. 公示地価内挿精度の検証

本章では， GWR, UK, GWRKの時空間モデルに公示地価データを適用し，内挿精度の比較を行う.

使用データは，2000年から2010年までの，東京都区部と武蔵野市，三鷹市の住居系用途地域内，4927点の公示地価情報である．地価関数の被説明変数は地価の対数値とし，説明変数は，最寄り駅までの距離(m)，地積($m^2$)，前面道路幅員(m)，容積率(%)，前面道路方位ダミー(南東・南・南西を1，それ以外を0)，最寄り駅から都内主要駅(東京・新宿・池袋・渋谷・上野)までの平均鉄道所要時間(分)，日経平均株価前年平均(円)，住宅ローン金利(%)を用いる．

### (1) パラメータの設定

共分散関数の時間軸・空間軸の異方性パラメータと, 時間・空間バンド幅の設定方法について述べる．

#### a) 異方性パラメータの設定

本研究では，時間差無限の観測点における攪乱項間の分散と，空間距離無限の観測点の攪乱項の分散が等しいと仮定して，設定する．

まず，通常最小二乗法(OLS)の残差を利用し，同一時点のデータのみを用いた経験セミバリオグラム，及び，同一地点のデータのみを用いた経験セミバリオグラムを作成する．その上で，それぞれの理論セミバリオグラムを求め，rangeが等しくなる時間軸と空間軸の比率を異方性パラメータとする．

#### b) バンド幅の設定

GWRでは交差検定を用いてバンド幅の推定を行う．最小二乗推定を観測点数だけ繰り返す必要があるが，計算負荷は大きくない．

しかし，GWRKで交差検定を用いたバンド幅推定することは現実的ではない. GWRKではUKのパラメータ推定が必要だが，セミバリオグラムの推定結果を用いて分散共分散行列を設定し，その後パラメータ推定を行う過程を踏むため，最小二乗推定と比較して多くの計算を必要とするからである(図-1)．

そこで本研究では，空間バンド幅を500~1800m・100m間隔，時間バンド幅を1～10年・1年間隔に設定して格子点探索を行い，バンド幅を推定した．

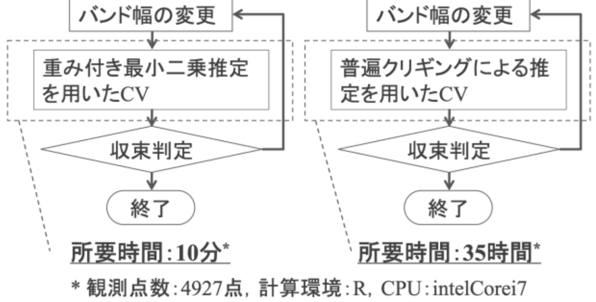


図-1 GWR, GWRKバンド幅推定手順

### (2) 内挿精度の検証

GWR, UK, GWRKを利用した内挿精度を検証する. なお, 計算にはデータ解析環境Rのgstatパッケージを使用した．gstatパッケージは，OLS推定の残差から作成した経験バリオグラムを用いて理論バリオグラムのパラメータ推定を行い，その結果から攪乱項の分散共分散行列を作成して時空間相関を考慮したモデルのパラメータ推定を行う．地価関数とバリオグラムのパラメータを同時には推定していないという限界を有することには注意が必要である．

精度は, 観測データセットから1データを順次除去し，除去したデータの値を残りの観測データを用いて内挿し比較するLeave-one-out Cross-validationを用いて検証する．また，内挿精度を表す指標として，内挿地価と検証用地価の平均二乗平方根誤差(RMSE)を用いる．

$$RMSE = \sqrt{\sum_{i=1}^{n}\left(y_i - \hat{y}_{(i)}\right)^2 / n} \tag{13}$$

$\hat{y}_{(i)}$ ：観測点$i$を除くデータを用いた$y_i$の推定値

#### a)全期間共通のバンド幅を用いた内挿精度比較

まず，全期間共通のバンド幅を用いたGWR, GWRKによる内挿について, UKの結果と合わせて図-2に年毎の内挿精度を示す. なお, 推定バンド幅は, GWRは413m, 0.5年，GWRKは1250m, 5年である．

まず, 全ての年でGWRKが高い内挿精度を示した. また，時間による違いに着目すると，全モデルで，2008年と対象期間端部付近で精度低下が見られた．

2008年の精度低下の原因は，前後と地価の変動傾向が大きく異なることにあると考えられる．経年的な地価変動を見ると，2007年まではいざなみ景気と呼ばれる景気拡大期間の影響によって地価上昇が続いたが，2009年は世界金融危機の影響により地価が急落した．GWR，UK, GWRKでは時間特性や時系列相関を考慮した内挿を行っているため，時系列で変化の傾向が安定している場合には高い精度が得られるが, 変動傾向が変化する場合, 精度低下が生じる．

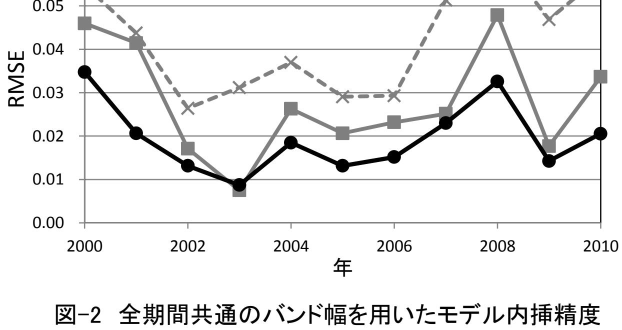


図-2 全期間共通のバンド幅を用いたモデル内挿精度

また，対象期間端部では，近隣の観測点が少なくなるEdge Effectにより精度が低下が見られる．

### b) 各年で最適なバンド幅を用いた内挿精度比較

各年で最も高い内挿程度を示したバンド幅（**表-1**）を用いてGWRKを通した内挿を行い，全期間共通のバンド幅を用いた結果と比較する（**図-3**）. 全体的にGWRKの内挿精度に改善が見られ，年により最適バンド幅が異なることが確認できる．その時々に適切なバンド幅を選択することにより，価格傾向変動の違いを考慮した内挿が可能であるといえる．

**表-1 GWRK 各年のバンド幅推定結果**

| 年 | バンド幅 空間(m) | バンド幅 時間(年) | 年 | バンド幅 空間(m) | バンド幅 時間(年) |
|---|---|---|---|---|---|
| 2000 | 1400 | 2 | 2006 | 800 | 3 |
| 2001 | 1000 | 2 | 2007 | 1200 | 5 |
| 2002 | 600 | 2 | 2008 | 1200 | 3 |
| 2003 | 700 | 5 | 2009 | 800 | 5 |
| 2004 | 700 | 2 | 2010 | 700 | 4 |
| 2005 | 1000 | 3 | | | |

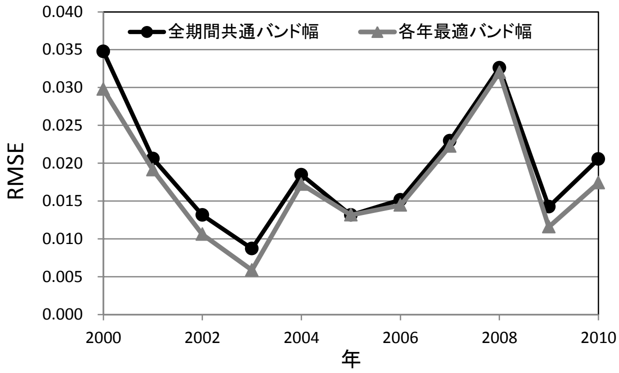


**図-3 バンド幅推定の違いによるGWRK内挿精度比較**

## 4. おわりに

本研究では，GWRKを地価データに適用し，内挿への適用可能性を実証的に検討した．GWR，UK，GWRKを地価データへ適用し，内挿精度の比較を行い，その結果，時間・地域特性や時空間相関を考慮したGWRKがGWR，UKよりも精度の高い内挿を実現することを示した．またGWRKにおいては，時系列で適切なバンド幅を設定する事により内挿精度を改善し，価格傾向の変動の捕捉が可能であることも確認された．以上によって，本研究ではGWRKの地価データへの適用可能性を示唆した．

しかし, 交差検定による GWRK のバンド幅推定の実行可能性に伴う限界のため，本研究では格子点探索を用いたことは課題であり，効率的なバンド幅推定手法の検討が必要である．

### 参考文献


1) 杉浦芳夫：地理空間分析, 朝倉書店, 2003.

2) 加藤尚史：不動産価格関数の推定と利用について, 日本統計学会誌, Vol34, pp.131–161、2005.

3) Fotheringham, A.S., Brundson, C. and Charlton, M. : Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environment and Planning. Vol30, pp.1905–1927, 1998.

4) 井上亮, 清水英範, 吉田雄太郎, 李勇鶴：時空間クリギングによる東京 23 区・全用途地域を対象とした公示地価の分布と変遷の視覚化, GIS－理論と応用, Vol.17, No.1, pp.13-24, 2009.

5) Fotheringham,A.S. Brunsdon,C. and Charlton,M.E.: *Geographically Weighted Regression : The analysis of spatially varying relationships*. John Wiley &Sons Ltd, 2002.

6) Harris P, Fotheringham AS, Crespo R. and Charlton M.：The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. *Mathematical Geosciences*, Vol.42, pp.657–680, 2010.


# An Empirical Study on Land Price Interpolation by GWR-Kriging Considering Spatio-Temporal Changes of Factors in Land Price Formation


Yuta KAWAMATSU and Ryo INOUE



One way to improve the transparency of real estate markets is to provide abundant price information to the market participants. A viable approach to the information provision is the interpolation of appraised land prices, which have been accumulated in space and time. Since land prices are affected by regional characteristics and economic environments of the time, they usually have strong spatio-temporal correlation and heterogeneity by its nature. In this study, we focus on GWR-Kriging, which is able to interpolate spatial data considering the spatial correlation and heterogeneity. We extend the model to examine spatio-temporal data, and analyze its applicability for spatio-temporal land price interpolation empirically.