

岡山市における地価公示価格の推定とその表示システムについて

中尾一生* 垂水共之**

Estimate of the posted land price in Okayama city and its visualizing system

Kazuo Nakao* and Tomoyuki Tarumi**

The posted land price is a declared land price from Land Appraisal Committee, and current market price is selling price of the market.

In this paper, we investigate the methods to estimate land price from the posted land price in Okayama city, and create the visualizing system of the price on.

key words: Land price, Google maps, GIS, Ajax

1 はじめに

現在地図情報を扱うソフトウェアは様々あるが、市販されているソフトウェアは高価で気軽に利用することができない。Web サイトで提供されている地図情報サイトは地図情報を調べることはできるが、地図に情報を加えたり編集したりすることはできない。これに対して、2005年にGoogle社はGoogle Mapsを自分のWebサイトに埋め込んで使用できるGoogle Maps APIを公開した。

それによって自分独自の地図を作ることができるようになった。しかしGoogle Maps APIはJavaScriptとXMLによって提供されているので、プログラミングを理解していないと利用は困難である。

我々が開発した「ポイントデータの簡易表示システム」では利用者がhtmlを作る必要がない。点の位置(緯度、経度等)を示すポイントデータをテキストエリアに貼り付けて、簡単な指示で地図に情報を載せることができるようになっていた。これにより誰でも容易に地図を利用することが可能となっている。この「ポイントデータの簡易表示システム」を利用して、岡山市の土地価格の推定とその表示システムを作成した。

土地の価格には地価公示価格、実勢価格等がある。地価公示価格とは土地鑑定委員会が毎年1回、1月1日時点の標準的な土地についての正常な価格として公表された土地の価格である。地価公示価格は土地の取引

価格に対して指標を与えるとともに、公共事業用地の取得価格算定の規準となる。実勢価格とは不動産の時価で、実際の取引で付けられた値段である。しかし実勢価格においては、一度も取引されたことが無い土地、または昔に取引された土地などが存在する場合は、周辺の土地から価格を推定するのが一般的である。

土地価格の推定において、正常な価格とされる地価公示価格を用いて岡山市の土地価格を推定した。

2 開発したシステムについて

2.1 Google Maps APIとは

Google Maps API (Application Programming Interface)とは、Google社が提供しているGoogle Mapsという地図検索システムから直接地図を取得できるシステムである。このシステムを使えば自分のWebサイトにGoogle Mapsを埋め込むことができる。座標を指示することにより、地図上にマーカーやラインなどを書き込むことができる。マーカーをクリックするとウィンドウが地図上に表示され、その地点に利用者が与えた情報を表示させることができる。現在はGoogleアカウントを取得することで、無料で使用することが出来る。

地図の情報はAjaxという技術を使用している。Ajaxとは動的にWebの表示を変化させるJavascriptと、表示内容をサーバーから非同期通信を介して取得するXMLHttpRequestを利用した技術である。これによって通信待ちを感じない滑らかな表示となっている。

*岡山大学大学院環境学研究科

**岡山大学アドミッションセンター

2.2 表示システムの概要

システムの流れは入力フォームにデータを入力して CGI に送信し、その送信されたデータを元に UTF-8 形式の html ファイルを出力する。

Google Maps API を使用してシステムを作成するためには、通常 Google アカウントを取得しなければいけない。しかし今回作成したシステムを使用すれば、アカウントを取得しなくても地図を自由に作成できる。またデータを地図上に載せるには Javascript 等によって GoogleMap の制御を行う必要があるが、本システムでは CGI で自動的に html を生成するので、Javascript 等の知識を必要としない。CGI は Perl で作成した。

GoogleMap は地図上に位置情報となるマーカーを載せることができる。地図上に表示することのできるデータは自分が用意した緯度、経度の座標を持つデータと、地価公示データベース(土地情報ライブラリー)による地価公示価格を表示することができる。

今回作成したシステムでは簡単な入力力で地図上に複数のマーカーを載せることができ、地図上のマーカーをクリックすることでその位置の情報を見ることができる。なお表示するマーカーには Google Mapki を利用した。

表示する地図は、マーカーの分布の中心が、地図の真ん中に来るように自動的に設定している。さらに、マーカー全体が表示できるように、地図の縮尺を調整している。

Fig.1 は入力から表示までのシステムの流れである。入力フォームからサーバーにある googlemaps.cgi へ利用者が与えたデータと表示の指定を送信する。googlemaps.cgi で html を作成し、地図を表示する。

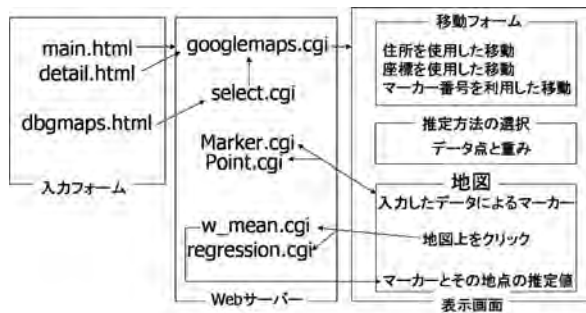


Fig. 1 システムの流れ

2.3 マーカーの色の变化について

表示するマーカーの色は入力フォームで条件指定と自動色階調によって変化させることができる。自動色階調とは入力フォームにおいて数値データを持つ列を

指定することによって、地図上に表示されるマーカーの色を変化させることができる。自動色階調は 3 種類あり、次に詳細を示す。

2.3.1 default

default は指定した列の値を元にマーカーの色を変化させている。推定に使用する値を x_i としたとき、色を決める変数 c_i が式 (1) から得られる。

$$c_i = \text{int}\left(\frac{1024 \times (x_i - x_{\min})}{x_{\max} - x_{\min}}\right) \quad (1)$$

式 (1) から得られた値 c_i は 0 から 1024 の値となる。そして c_i を Fig.2 に当てはめることによってマーカーの色を決める。色は大きい値だと暖色となり、小さい値だと寒色となる。

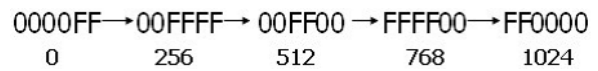


Fig. 2 色の变化

2.3.2 5%トリム

default では全てのデータを対象に c_i を計算する。しかし、外れ値が存在することにより c_i が偏ることがある。このため両端 5% (片側 2.5%) のデータを取り除き、残った値を式 (1) により c_i を求める。取り除いた値は上側を $c_i = 1024$ 、下側を $c_i = 0$ とする。

2.3.3 8等分

この方法では指定した変数を 8 等分する。

$$c_i = \text{int}\left(\frac{8 \times j}{n}\right) \quad (2)$$

ここに、 j は i 番目データの降順の順位であり、 n はデータ数である。 c_i は 0 から 7 の値を取り、マーカーの色は (3) から得られた値を取る。

$$\left\{ \begin{array}{ll} 0000FF & (c_i = 0) \\ 00FFFF & (c_i = 1) \\ 00FF80 & (c_i = 2) \\ 00FF00 & (c_i = 3) \\ 80FF00 & (c_i = 4) \\ FFFF00 & (c_i = 5) \\ FF8000 & (c_i = 6) \\ FF0000 & (c_i = 7) \end{array} \right. \quad (3)$$

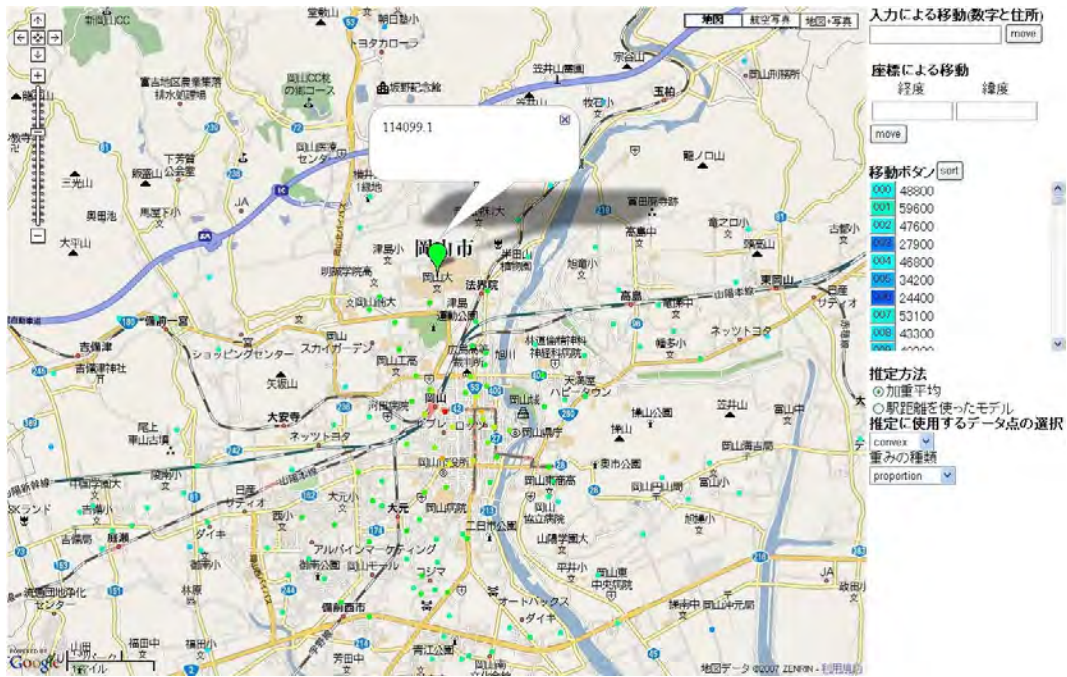


Fig. 3 地図の表示

2.3.4 対数

「default」と「5% トリム」は指定した列の値に比例して色が変わった。この機能では値の対数を取ることで、元の値では綺麗に表示できない場合にも対応している。

$$c_i = \text{int}\left(\frac{1024 \times (\log(x_i - x_{\min} + 1))}{\log(x_{\max} - x_{\min} + 1)}\right) \quad (4)$$

式 (4) から得られた値 c_i は 0 から 1024 の値を取り、Fig.2 に当てはめることによってマーカーの色を決める。

2.4 推定システムについて

地図上をクリックすることにより Google Maps API の関数で座標を取得する。取得した座標と推定方法をサーバーに送信して統計解析ソフトウェア R によって推定値を計算する。そして計算が終わったら推定値を受信し、地図上にマーカーとウィンドウを表示する。そのウィンドウの中に推定値を表示する。これらを非同期通信によって行っている。Fig.3 は地図上に推定値を表示させたものである。マーカーは一度に一つ表示される。2ヶ所目をクリックすると、マーカーが消えて新たな推定地点にマーカーが表示される。推定方法は右下のボタンによって選択することができる。

3 加重平均による推定

地価公示データには緯度、経度の他に地積、容積率、などがあるが、本節の推定には緯度、経度だけを使用した。地価公示価格のような空間データの場合、距離が近い所では似た値となりやすいという空間従属性が認められる。そのため、地価公示価格が未知の地点の価格は、その地点からデータ点までの距離を重みとする。よって加重平均を推定値とした。

加重平均を使用するにあたって、推定に使用するデータ点の選択と距離による重みの方法を考えた。これらを組み合わせた加重平均を計算する。

3.1 データ点の選択方法

3.1.1 三角形を形成するデータ

使用するデータ点は

条件 A

推定地点を内点とし、与えられたデータ点が内点とならない組み合わせ

を満たす三角形の組み合わせの中で推定地点から最も近い点を使う組み合わせを使用する。Fig.4 は推定地点とデータ点の散布図である。

データ点のラベルは推定地点からの距離の近い点の順位をつけている。最も近い点 3 個で三角形を構成し、条件 A を満たすか判定する。条件を満たせば推定に使

用するデータ点とする。条件を満たさなければ三角形を構成する点の中で最も遠い点を取り除く。そして、次に近い点を使用して三角形を構成し条件 A を満たすか判定をする。

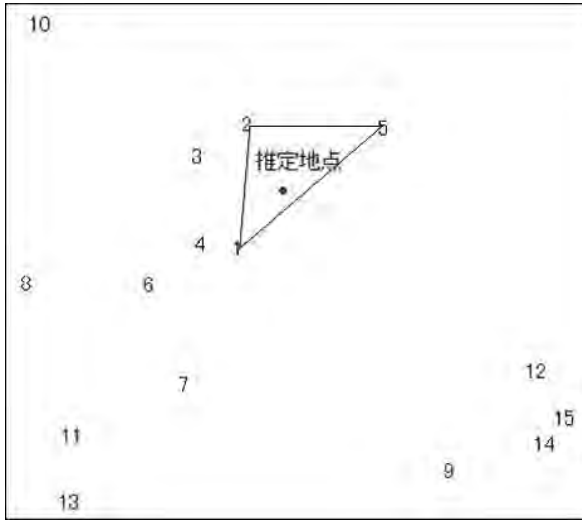


Fig. 4 三角形を構成するデータ点

3.1.2 凸多角形を形成するデータ点

前述の条件 A を満たす凸多角形を形成する。凸多角形を形成するアルゴリズムは以下の流れとなる。

凸多角形を形成するアルゴリズム

1. 条件 A の三角形を形成する
2. 残りの点の中で推定地点に最も近い点と三角形 (凸多角形) の点で凸多角形を形成する
3. 条件 A を満たすか判定
満たせば凸多角形を構成する点に組み込む
4. 2. に戻る .

Fig.5 は凸多角形を形成した図である。

3.1.3 q 個の近傍点

推定地点から近い q 個の地点のデータを使用して推定を行う。最適な個数 q はクロスバリデーションによって決めた。Fig.6 では q = 6 のとき使用する点を表している。円より内側の点を使用して推定を行う。

3.1.4 四方向のデータ点

推定地点を原点とし、データ点を第一象限から第四象限に分ける。各象限の最も近い点 q 個を推定に使用

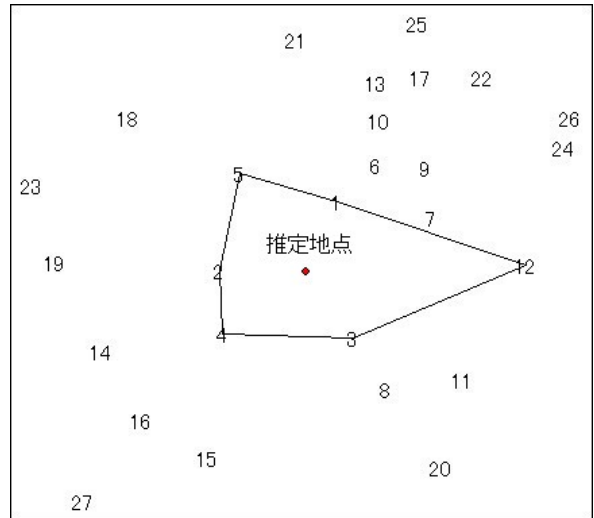


Fig. 5 凸多角形を構成するデータ点

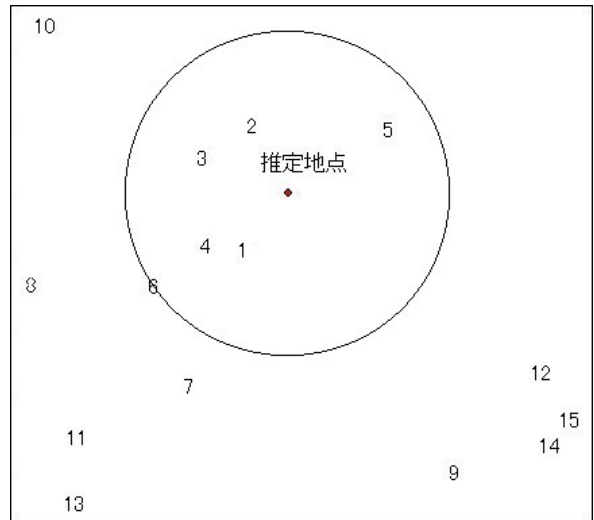


Fig. 6 q = 6 の時使用する点

する。最適な q はクロスバリデーションによって決めている。Fig.7 では q = 2 のときに使用する点を表している。中心の点が推定する地点であり、楕円で囲った点を使用して推定を行う。

3.2 重みの種類

3.2.1 距離に比例した値

推定地点からデータ点までの距離に比例した値を重みとする。

$$w_i = 1 - \frac{d_i}{\max(d)} \tag{5}$$

ここに d_i は推定地点から i 番目のデータ点までのユークリッド距離とする。推定地点に最も近い点に大きい重みが掛かり、最も離れた点の重みは 0 となる。

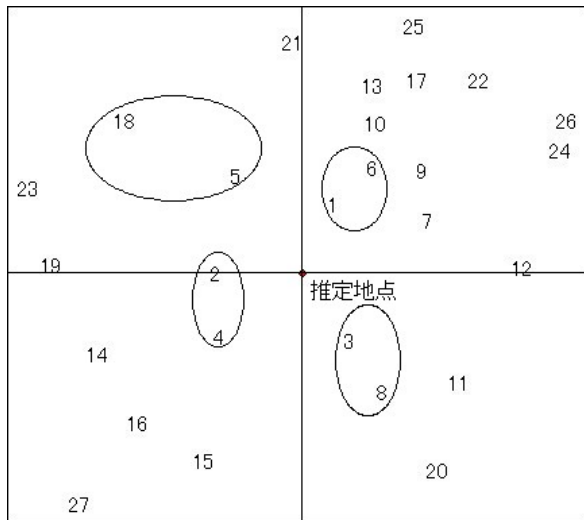


Fig. 7 $q=2$ のとき使用する点

3.2.2 逆距離加重法

推定地点からのユークリッド距離またはユークリッド二乗距離の逆数を重みとする。

$$w_i = \frac{1}{d_i} \tag{6}$$

3.2.3 gauss 型関数

$$w_i = \exp(-(d_i/\delta)^2) \tag{7}$$

δ はバンド幅であり δ が大きいほど、より広範囲の地点が大きい重みとなる。逆に δ が小さいと、狭い範囲に大きい重みがかかる (杉浦芳夫;2006)。最適バンド幅 δ は推定地点から最も離れた地点までの距離を 100 等分した値をバンド幅とし、クロスバリデーションを行って決めた。

3.2.4 tri-cube 関数

$$w_i = (1 - (d_i/d_q)^3)^3 I(d_i \leq d_q) \tag{8}$$

推定地点から q 番目より離れた地点の重みを 0 とし、 q 番目より近い点に重みを付ける。距離に比例して重みを付けたものと比較して、より近傍点に大きい重みがかかる (James P.LeSage;2001)。最適な q はクロスバリデーションによって求めた。

4 岡山駅からの距離を使った土地価格のモデル

前節では加重平均を使用して推定を行った。Table 1 から岡山駅からの距離と地価公示価格の相関関係が高いことがわかる。

Table 1 岡山市の地価公示価格の相関係数

	住宅地等	商業地
相関係数	-0.733	-0.560

本節では岡山駅からの距離に関する影響を回帰モデルで取り除き、残った残差の値 e_i に対し空間的従属性を考慮した加重平均による推定を行い、合わせたものをその地点の推定値とする。

モデルの作成手順

1. 岡山駅からの距離 d_i を説明変数とし、地価公示価格を目的変数とする回帰モデルを作成する。
- $$y = f(d) + e$$
2. 前節の加重平均と同様に推定に使用するデータ点と重みを選ぶ。
 3. データ点の 1. で求めた回帰モデルによる残差 e_i に重みを付ける。
 4. 推定地点から岡山駅までの距離 d_0 を回帰モデルに与えて、得られた値に 3. の残差の加重平均を加えた値を推定値とする。

モデル式は (9) となる。

$$\hat{y}_0 = \hat{y}'_0 + \frac{\sum w_i e_i}{\sum w_i} \tag{9}$$

\hat{y}_0 : 地価公示価格の推定値

\hat{y}'_0 : 回帰モデルによる地価公示価格の推定値

e_i : データ点の回帰モデルの残差

w_i : 重み

5 推定方法の比較

5.1 比較に使用する値

本研究では 2006 年地価公示価格を用いて推定を行った。岡山市の地価公示のデータ点を一つ取り除き、その地点を推定する。比較に使用する地点の全てで行い、以下の 3 つの推定方法のよさを比較する値を用いる。絶対誤差

$$\sum |y_i - \hat{y}_i| \tag{10}$$

相対誤差

$$\sum \frac{|y_i - \hat{y}_i|}{y_i} \tag{11}$$

χ^2 分布の適合度検定を基にした値

$$\sum \frac{(y_i - \hat{y}_i)^2}{\hat{y}_i} \tag{12}$$

5.2 比較条件

Fig.8 は岡山市の地価公示価格を商業地と住宅地等に分けて箱髭図をプロットしたものである。

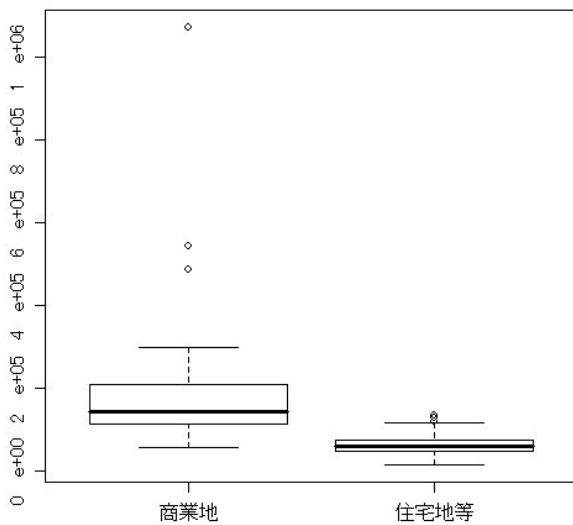


Fig. 8 boxplot

商業地の最小値は住宅地等における中央値とほぼ同じ値である。そして商業地においては外れ値がいくつか存在することがわかる。

そして商業地には土地価格が 1,070,000 円の土地が存在する。この土地はデータ点、推定地点に使用すると誤差が大きくなり、正しく比較できない。

以上のことから「商業地」とそれ以外の土地の「住宅地等」は土地価格の性質が異なっていることから分けて推定を行う。さらに、商業地の推定においては 1,070,000 円のデータを取り除いて推定を行う。

市の境付近の推定地点は市を越えた地域のデータ点が存在せず、ある方向のデータだけを使って推定を行うことになる。また、データ点の選択において「凸多角形」、「三角形」など推定できない地点が存在する。このことから市の境付近のデータは推定地点には使用せず、データ点としてのみ使用する。

データ点と重みの組み合わせにおいて適切ではない組み合わせが存在する。データの選択法である「全デー

タ」と「 q 個の近傍点」は、「距離の比例」、「距離の逆数」、「二乗距離の逆数」の重みを使って推定した場合には、「 q 個の近傍点」はデータ点の候補に「全データ」が含まれている。よって上記の三つの重みと「全データ」の組み合わせは、「 q 個の近傍点」を使った組み合わせの方が結果が良くなるのが明白なので、比較しない。そして「三角形を構成するデータ」において「距離の比例」、「tri-cube 関数」の重みは、3 個のデータ点の内、最も遠い点の重みが 0 となる。よってデータ点から得られる情報が大きく減少するので比較に個の組み合わせを使用しない。

岡山駅からの距離を使用したモデルでは「住宅地等」においては Table 1 から線形モデルを使用する。しかし、「商業地」においては「住宅地等」と比較して相関係数は低い。Fig.9 は縦軸に「地価公示価格」、横軸に「岡山駅からの距離」をとった散布図である。Fig.9 より高い土地価格の存在により、非線形モデルの方がモデルとして適切である可能性がある。

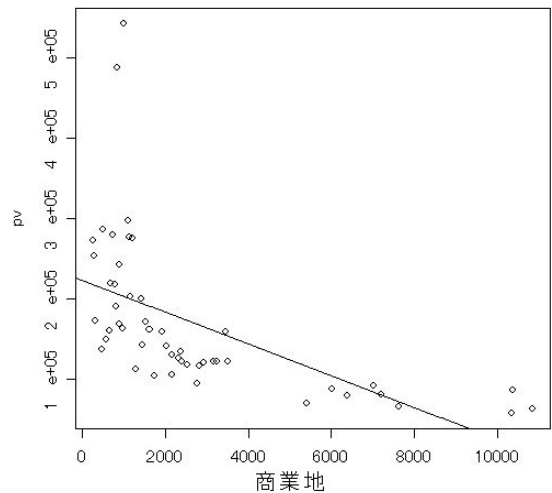


Fig. 9 商業地における地価公示の散布図

よって「岡山駅からの距離を使ったモデル」は商業地においては、線形モデルと非線形モデルを残差二乗和を用いて比較を行い、良い結果を使用した。比較結果は Table 2 となる。

Table 2 回帰モデルの比較結果

回帰モデル	$y = ax + b$	$y = a/x + b$
残差二乗和	3.06E+11	4.95E+10
$y = ax^b + c$	$y = a \log x + b$	$y = ax + bx^2 + c$
2.11E+10	1.13E+09	5.45E+09

以上の結果から商業地においては $y = a \log x + b$ を

使用する。

5.3 比較結果

以上の条件を元に比較を行った。Table 3 から Table 6 は「加重平均」と「岡山駅からの距離を使ったモデル」において、比較した結果が最も良かったデータの選択法と重みの組み合わせとその値である。

Table 3 住宅地等の加重平均による地価公示価格の推定値の比較

比較方法	データ点	重み	推定方法
絶対誤差	4 方向	gauss 型	1,232,108.11
相対誤差	4 方向	gauss 型	24.24
適合度検定	4 方向	距離の逆数	335,248.43

Table 4 商業地の加重平均による地価公示価格の推定値の比較

比較方法	データ点	重み	推定方法
絶対誤差	凸多角形	gauss 型関数	1,716,897.60
相対誤差	凸多角形	gauss 型関数	7.93
適合度検定	凸多角形	距離の逆数	1,078,661.45

Table 5 住宅地等の岡山駅からの距離を使ったモデルによる地価公示価格の推定値の比較

比較方法	データ点	重み	推定方法
絶対誤差	4 方向	二乗距離の逆数	1,216,093.36
相対誤差	4 方向	二乗距離の逆数	23.64
適合度検定	4 方向	二乗距離の逆数	329,158.25

6 まとめ

住宅地等において「加重平均」、「岡山駅からの距離を使ったモデル」ともデータ点の選択では、「四方向」からデータを取る方法がすべての比較方法において、最も良い結果となった。商業地においてはデータ点の選択では「凸多角形を構成するデータ点」を使用する方法が最も良い結果となった。これは商業地は岡山駅を中心に固まって分布しているのに対して、住宅地等は均等に分布している等のデータ点の分布の違いがデータ点の選択方法に現れたと考えられる。

Table 6 商業地の岡山駅からの距離を使ったモデルによる地価公示価格の推定値の比較

比較方法	データ点	重み	推定方法
絶対誤差	凸多角形	距離の比例	1,759,298.60
相対誤差	凸多角形	距離の比例	8.19
適合度検定	凸多角形	距離の逆数	1,001,936.44

今回の研究では位置情報だけを用いたが、地価公示価格のデータに含まれている他の変数を使うことにより、より良い推定が得られる可能性があり、地価を決める要因分析も可能となると思われる。

参考文献

Google Mapki

http://mapki.com/wiki/Main_Page

James P.LeSage 「A Family of Geographically Weighted Regression Models」・2001

杉浦芳夫

地理空間分析・株式会社 朝倉書店・2006

土地総合情報ライブラリー

<http://tochi.mlit.go.jp/>