

地価公示における 標準地数削減に関する空間統計学的考察

友成 将¹・村上 大輔²・堤 盛人³

¹筑波大学大学院 システム情報工学研究科 (〒305-8573 茨城県つくば市天王台 1-1-1)

E-mail : s1620481@sk.tsukuba.ac.jp

²国立環境研究所 地球環境研究センター (〒305-8506 茨城県つくば市小野川 16-2)

E-mail : murakami.daisuke@nies.go.jp

³筑波大学 システム情報系 社会工学域 教授 (〒305-8573 茨城県つくば市天王台 1-1-1)

E-mail : tsutsumi@sk.tsukuba.ac.jp

地価公示における標準地の数は平成17年から平成27年まで減少傾向にあり、特に平成26年には大幅な削減が行われたが、筆者ら知る限り、それについての統計学的な観点からの議論や分析はほとんど無い。本研究では、標準地の実際の削減に関して、地球統計学に基づくサンプリングデザインの最適化を行った村上・堤 (2015) の方法を用いて実際の削減地数と同数削減し、比較・考察を行った。それにより公的地価調査の調査地点数削減業務に村上・堤 (2015) の手法を活用できる可能性を明らかにした。

Key Words : *officially assessed land price, geostatistics, sampling design, assessed sites*

1. はじめに

地価公示制度は、その公示価格が課税や土地取引の参考になるなど重要な役割を担っている。その一方で、財政の逼迫を契機にその縮小が余儀なくされている。近年、地価公示の調査地点である標準地数は図-1に示すように平成15年までは増加傾向にあったが、平成15年、平成16年の標準地数31,866地数をピークに平成17年以降は減少傾向にある。実際に平成24年度国土交通省行政事業レビューシート (事業番号 : 0117 事業名 : 地価公示) ¹⁾ では、「他の土地評価制度との関係を整理した上で、標準地の地点数の絞り込みを行い、より効率的に事業を執行すべき」との抜本的改善を指摘され、平成25年から平成26年に掛けて約10%の地点数の削減が行われ、ピーク時のおよそ4分の3の23,380地点にまで減少した。

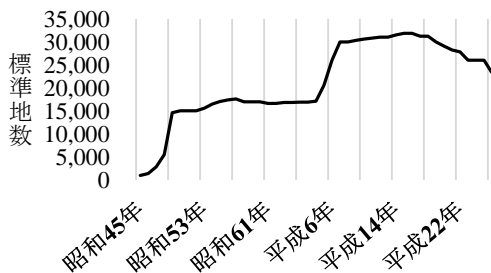


図-1 標準地数の推移

標準地の削減方法を議論した研究は筆者らの知る限り村上・堤 (2015) ²⁾ のみである。村上・堤 (2015) は、地球統計学 (e. g., Cressie (1993) ³⁾) を用いたサンプリングデザイン最適化のための手法 (e. g., Zimmerman (2006) ⁴⁾) を応用することで標準地及び都道府県地価調査の基準地の配置問題を考察した。

公的地価調査の調査地点数の削減に、村上・堤 (2015) を応用できる可能性があることを踏まえ、本研究では、統計学的観点に基づき、平成25年から26年の標準地数の大幅な削減を例に、実証的に考察を行う。第2章では、まず、本研究で用いる、村上・堤 (2015) の手法を概説する。続く第3章では、本研究が分析の対象とする標準地に関するデータについて記述する。第4章では、実際にサンプリング・デザインの手法を適用した結果について考察し、第5章において本研究を総括する。

2. 手法

(1) 地球統計モデル

地球統計学では連続空間上に分布するデータ (サンプルサイズ : n) を式 (1) でモデル化する。

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{C}) \quad (1)$$

\mathbf{z} は $n \times 1$ の被説明変数ベクトル, \mathbf{X} は $n \times K$ (定数項を含む) の説明変数行列, $\boldsymbol{\beta}$ は $K \times 1$ のパラメータベクトル, $\boldsymbol{\varepsilon}$ は $n \times 1$ の誤差項ベクトル, \mathbf{C} は $n \times n$ の共分散行列であり, その要素を共分散関数と呼ばれる距離 h の減衰関数で与えることで, 「近接したデータが類似した傾向を示す」という空間的相関を考慮する²⁾. 共分散関数には球型モデルやガウシアンモデルなど, これまでに数多くの共分散関数が提案されてきた. 本研究で用いる指数形モデルは式 (2) に示す通りである.

$$C(h) = \begin{cases} \sigma^2 \exp\left(-\frac{\|h\|}{r}\right) & \text{if } \|h\| > 0 \\ \tau^2 + \sigma^2 & \text{otherwise} \end{cases} \quad (2)$$

τ^2 , σ^2 , r はそれぞれナゲット, パーシャルシル, レンジと呼ばれるパラメータであり, それぞれ空間的相関で説明できない誤差項の大きさ, 空間的相関で説明される誤差項の大きさ, 空間的相関の及ぶ距離を表す.

共分散関数は leave-one-out 交差検証法 (間瀬, 2010)⁵⁾ を用いて選択することができる. この手法は, 「 i 番目の標本の値を残りの全標本で予測する」という操作を全ての i について繰り返し, 平均二乗予測誤差 (RMSE) を求めることでモデルの精度を評価する. RMSE は式 (3) のように記述される.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |z(s_i) - \hat{z}(s_i)|^2} \quad (3)$$

空間統計モデルを用いることで, 任意地点 s_0 におけるデータ $z(s_0)$ の予測値やその期待二乗誤差を評価できる. いま, $z(s_0)$ が式 (4) に従うとする.

$$z(s_0) = \mathbf{x}'(s_0)\boldsymbol{\beta} + \varepsilon(s_0) \quad (4)$$

「 $'$ 」は転置の演算子, $\mathbf{x}(s_0)$, $\varepsilon(s_0)$ はそれぞれ地点 s_0 における説明変数ベクトル, 誤差項である. 不偏かつ線形の条件のもとで $z(s_0)$ の期待二乗誤差を最小にすることで, $z(s_0)$ の最良線形不偏予測量 (式 (5)) とその期待二乗誤差 (式 (6)) を導出できる.

$$\hat{z}(s_0) = \mathbf{x}'(s_0)\hat{\boldsymbol{\beta}} + \mathbf{c}'\mathbf{C}^{-1}(\mathbf{z} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (5)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{z}$$

$$E\left\{z(s_0) - \hat{z}(s_0)\right\}^2 = \sigma^2 + \tau^2 - \mathbf{c}'\mathbf{C}^{-1}\mathbf{c} + (\mathbf{x}(s_0) - \mathbf{c}'\mathbf{C}^{-1}\mathbf{X})(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})(\mathbf{x}(s_0) - \mathbf{c}'\mathbf{C}^{-1}\mathbf{X})' \quad (6)$$

ここで \mathbf{c} は \mathbf{z} と $z(s_0)$ との間の共分散ベクトルを表す.

(2) サンプリングデザイン

a) 目的関数

サンプリングデザインの最適化は, 重み付けられた期待二乗誤差の平均値・最大値を目的関数 (例えば式 (7), (8)) として定め (Zhu and Stein, 2005)⁶⁾, それらを最小化することで最適な配置を探索する.

$$\min_{S_i \in S} [C_1(S_i)] = \min_{S_i \in S} \left[\sum_m w_m c_m(S_i) \right] \quad (7)$$

$$\min_{S_i \in S} [C_2(S_i)] = \min_{S_i \in S} \left[\max_m \{w_m c_m(S_i)\} \right] \quad (8)$$

ここで S は配置の集合, $C(S_i)$ は配置 S_i における目的関数の値を表す. 式 (7), (8) は対象地域上に稠密に配置された点群 (添字: m) 上で予測誤差分散 $c_m(S_i)$ を評価することで, 対象地域全域における配置の予測精度を定量化しようというものである. 式 (7) は重みと期待二乗誤差の積の平均値を最小化するミニサム型の計算式であり, 式 (8) は重みと期待二乗誤差の最大値を最小化するミニマックス型の計算式である. ミニサム型の計算式は全体としての精度を向上させるのに優れ, ミニマックス型の計算式は特異な精度の悪化を防ぐ. w_m は地点 m に対する重みである. 例えばこれを (周辺) 世帯数で与えることで, 世帯数の多い地域の地価をより精度よく説明するような地価調査地点の配置が特定できる.

b) アニーリング法

配置探索の際に, 全てのパターンについて目的関数の値の算出を行うとすると膨大が時間がかかる. そこで, 現実的な時間内でかつ局所的な最適解に陥らないようにするアルゴリズムを用いる. アニーリング法は, 改悪方向にも遷移する確率を与えることにより局所最適解からの脱出を図る方法である (柳浦・茨木, 2001)⁷⁾. 改悪方向への遷移確率はパラメータ T で与える. 次に示す手順で配置の探索を行う.

step1. 初期配置を生成する. 本研究では標準地数削減後の最適配置を求めるため, 初期配置は削減後の地点数をランダムに配置したものとなる. この初期配置を S_0 とする. さらに T は改悪確率を定めるパラメータである.

step2. 以下 (ア) ~ (ウ) を k 回繰り返す.

(ア) 配置 S_0 の地点のうち 1 地点をランダムに変更し, その配置を S とする.

(イ) 配置 S_0 の時の目的関数を $C(S_0)$, 配置 S の時の目的関数を $C(S)$ とし, $D = C(S_0) - C(S)$ とする.

(ウ) D が 0 以上の時は確率 1 で S_0 を S で置き換える. そうでない場合は確率 $\exp(-D/T)$ で S_0 を S で置き換える.

step3. 反復条件を満たした場合, 計算終了. そうでない場合, T を pT で置き換え step2 に戻る. ただし $p(0 < p < 1)$ は T の減衰率である.

p は1に極限まで近い値を与えると、解が理論的には大域的最適解に一致する (Kirkpatrick *et al.*, 1983)⁸⁾。本研究ではミニマックス型の探索を行う際は $T=200$, $p=0.95$ とし、ミニサム型の探索を行う際は $T=200$, $p=0.92$ とした。

3. 対象地及びデータの説明

本研究では、茨城県における平成25年から平成26年にかけての削減 (693地点→647地点) を分析する。具体的には、平成25年地価公示データに2章の手法を適用することで削除が推奨される地点を特定し、この削除推奨地点と平成26年の実際の削減地点を比較する。それにより「削減地点の選択がどの程度適正であったか」を定量化する。なお、平成25年から平成26年にかけての茨城県における削減数・削減率は46地点・約6.64%であった。本研究で用いる説明変数は表-1に記述したとおりである。

地価公示 (標準地) の空間分布は図-2 (平成25年) と図-3 (平成26年) に示すとおりである。水戸・土浦・つくばといったような茨城の中心に多く標準地点が存在し、それは鉄道沿線であるとも言える。削減前後の配置は目視ではあまりその差がわからない。続いて図-4に平成25, 26年の公示価格及びその間の削減地点の地価の箱ひげ図を示す。同図より、平成25, 26年地価公示の四分位範囲と削減地点の四分位範囲に大差はなく、価格が集中している地点から削減が行われていることが伺える。さらに、高価格帯の標準地、具体的には地点10万円/m²を超える価格帯の標準地は一切削減されていないことがわかる。

表-1 変数の説明

変数	説明
最寄り駅までの距離 [m]	最寄りの鉄道駅までの直線距離 (出典：国土数値情報：平成25年)
最寄りバス停までの距離 [m]	最寄りのバス停までの直線距離 (出典：国土数値情報：平成25年)
つくば駅・水戸駅・土浦駅のいずれかまでの距離 [m]	つくば駅もしくは水戸駅もしくは土浦駅まで直線距離で最短の距離 (出典：国土数値情報：平成25年)
土地利用の面積 [m ² / 500m ²]	土地利用細分メッシュデータより、建物・鉄道用途以外の半径500m以内における各土地利用面積 (出典：国土数値情報：平成21年)
DIDダミー	人口集中地区に含まれるかどうかのダミー変数 含まれる場合:1, 含まれない場合:0 (出典：国土数値情報：平成25年)
世帯数	町丁目別世帯数 (出典：国税調査：平成22年)

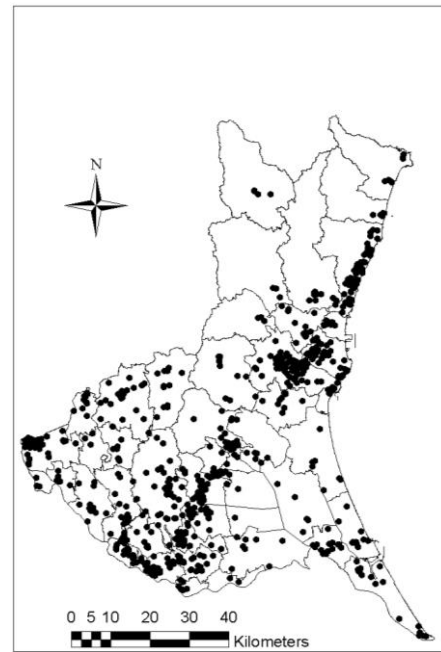


図-2 平成25年地価公示標準地の分布

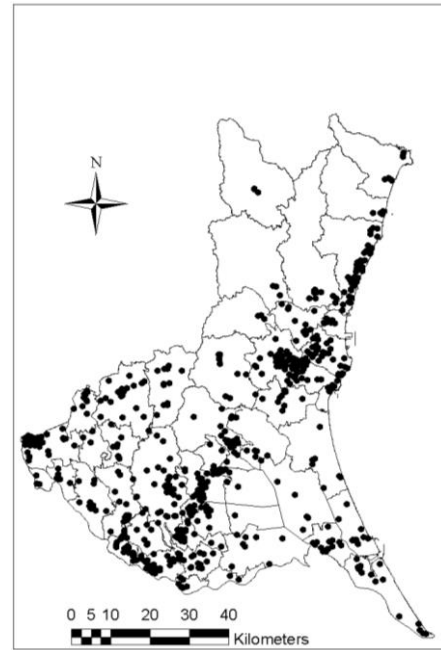


図-3 平成26年地価公示標準地の分布

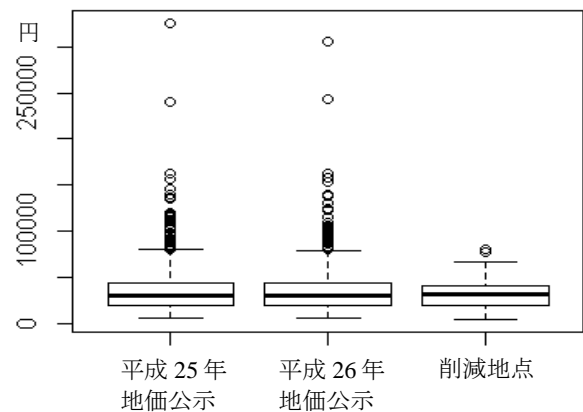


図-4 公示価格の箱ひげ図

4. 分析・結果

本章では、最初に地球統計モデルの推定結果を説明し、次に推定された地球統計モデルを用いたサンプリングデザインの最適化手法を標準地の削減問題に適用する。

(1) 地球統計モデルの推定

まずleave-one-out交差検証法を用いて共分散関数のためのモデルを選択した同交差検証で得られた各モデルのRMSEを表-2の2列目に示す。同表より、RMSEが最小である指数型モデルに決定した。

表-2 RMSEの比較

共分散関数モデル	RMSEの値
球型モデル	3.53×10^{-1}
指数型モデル	3.42×10^{-1}
ガウシアンモデル	3.75×10^{-1}
線形モデル	3.71×10^{-1}
円形モデル	3.56×10^{-1}
5次モデル	3.52×10^{-1}

表-3 パラメータ推定結果

変数	推定値	t値	VIF
定数項	9.88	1.16×10^2 ***	
最寄り駅距離	5.25×10^{-6}	7.47×10^{-1}	1.32
最寄りバス停距離	-9.80×10^{-5}	-3.67 ***	1.05
三駅最短距離	-1.31×10^{-6}	-8.67×10^{-1}	1.11
田	1.04×10^{-8}	1.10×10^{-1}	1.18
その他農地	-1.03×10^{-8}	9.33×10^{-2}	1.31
森林	4.46×10^{-8}	3.08×10^{-1}	1.08
荒地	-1.92×10^{-8}	-3.19×10^{-2}	1.04
道路	2.20×10^{-6}	1.31	1.03
鉄道	-6.10×10^{-7}	-2.44×10^{-1}	1.06
その他用地	1.90×10^{-7}	8.04×10^{-1}	1.15
河川地・湖沼	1.22×10^{-7}	5.18×10^{-1}	1.06
海浜	-1.11×10^{-6}	-6.40×10^{-1}	1.20
海水域	6.65×10^{-7}	1.24	1.14
ゴルフ場	2.02×10^{-7}	3.41×10^{-1}	1.02
DIDダミー	4.07×10^{-1}	1.02×10 ***	1.06
ナゲット	6.72×10^{-2}		
パーシャルシル	1.95×10^{-1}		
レンジ(m)	5.59×10^3		

***は1%有意水準で有意であることを表す

パラメータの推定結果を表-3に要約した。推定されたパラメータのうち、その多くが有意とはならなかった。有意な結果が出たのは、定数項・最寄りバス停までの距離・DIDダミーのみである。バス停が近い地点、人口が集中している地点であるほうが地価が高い傾向にあるという解釈は直観に整合する。ナゲットとパーシャルシルの推定結果からは空間的相関を持つ分散成分が非空間の分散成分の約3倍であること、またレンジの推定結果からは茨城県内における空間的相関が及ぶ範囲が5.59km程度であることが、それぞれ確認できる。

(2) サンプリングデザインの最適化

サンプリングデザインの最適化を進めるにあたり、本研究では目的関数の重み w_m を各町丁目ごとの世帯数とした。これは一般に世帯数が多い地域ほど課税対象や土地の取引数が多く、より重点的に標準地を配置する必要があるためである。目的関数に関しては、全体の精度を上げるミニサム型(式6)・特異的な精度の低下を抑えるミニマックス型(式7)の双方を用いて探索を行う。

そして本研究では、アニーリング法による探索を行う際により効率的に統計的観点から望ましい結果を得るために、本来ランダムに生成すべき初期配置を実際の削減後の配置と指定した探索も行う。表-4に実際の削減地点と探索により得られた削減地点の一致率を示し、実際の削減地点及び配置探索後の結果を図-3, 4, 5, 6, 7に示す。さらに図-8に目的関数別に削減地点の価格帯の箱ひげ図を示し、算出された目的関数の値を表-5に示す。

計算には、intel(R) core(TM)i7 : メモリ32GB, windows64ビットOSの計算機上で、無償の統計ソフトRのバージョン3.2.2を用いた。計算に要した時間は、目的関数ごとにいずれもおよそ16時間であった。

表-4 削減地点の一致率

	実際の削減地点	ミニマックス型	(初期配置指定)ミニマックス型	ミニサム型
ミニマックス型	6.52%	-	-	-
ミニマックス型(初期配置指定)	6.52%	15.2%	-	-
ミニサム型	8.70%	6.52%	15.2%	-
ミニサム型(初期配置指定)	8.70%	4.35%	13.0%	80.4%

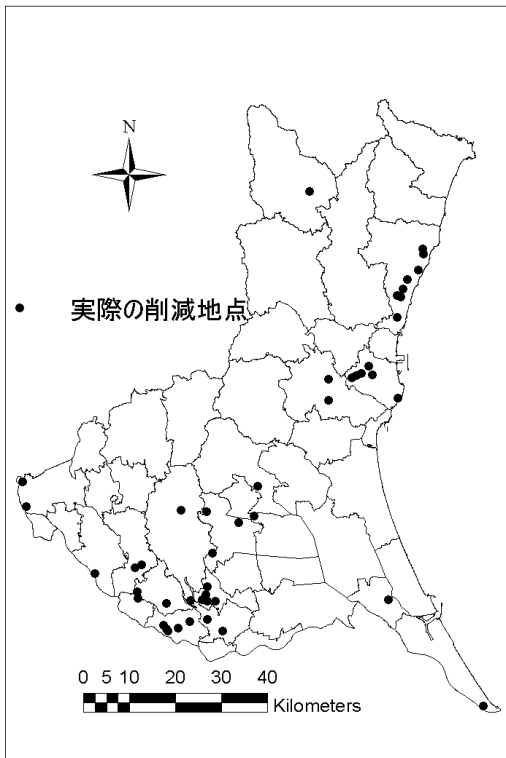


図-3 実際の削減地点

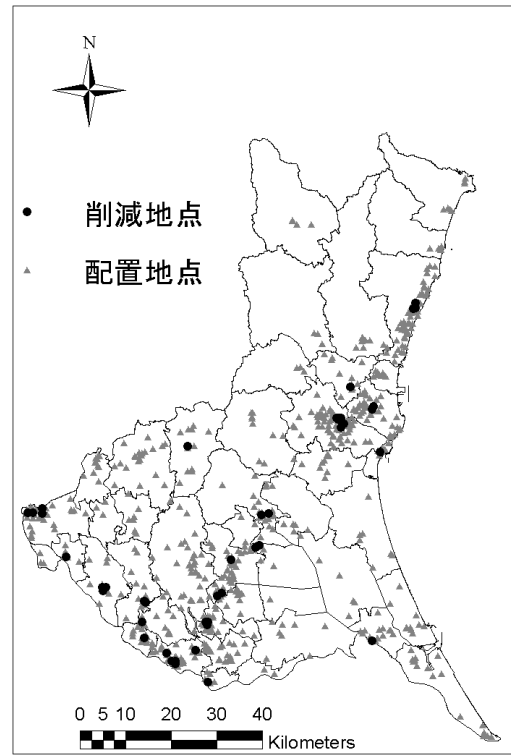


図-5 ミニマックス型の関数を与えて得られた削減地点及び削減後の配置 (初期配置指定)

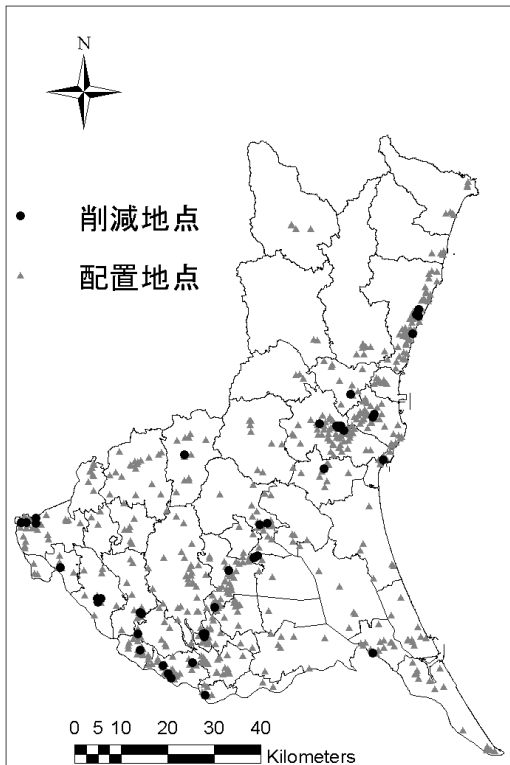


図-4 ミニマックス型の関数を与えて得られた削減地点及び削減後の配置

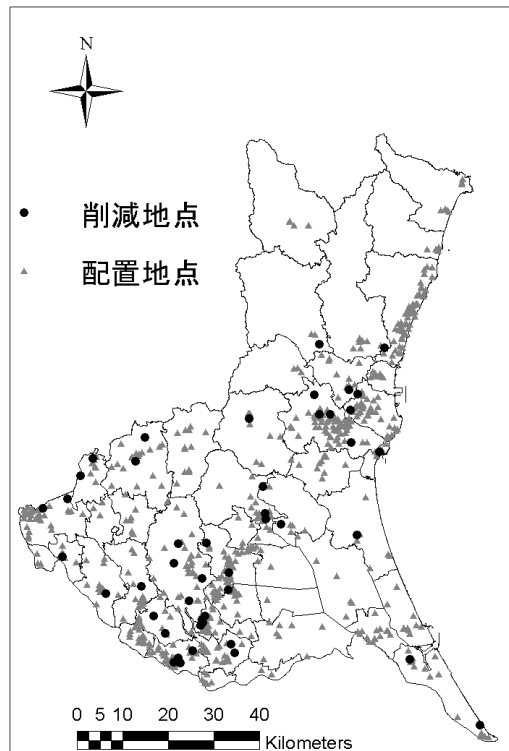


図-6 ミニサム型の関数を与えて得られた削減地点及び削減後の配置

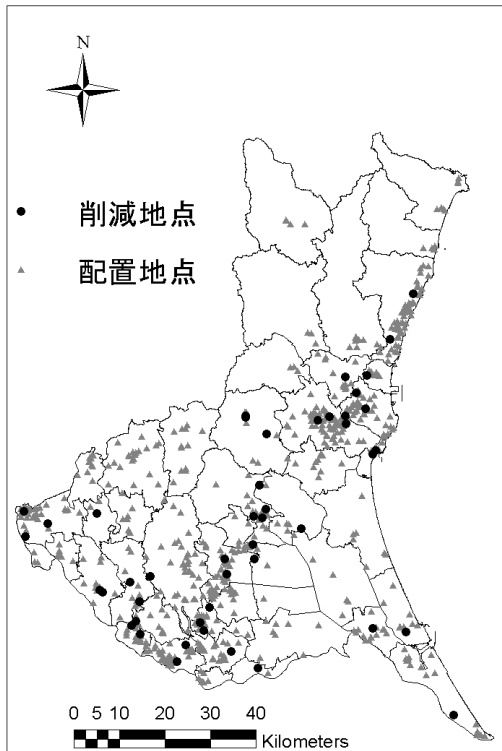


図-7 ミニサム型の関数を与えて得られた削減地点及び削減後の配置 (初期配置指定)

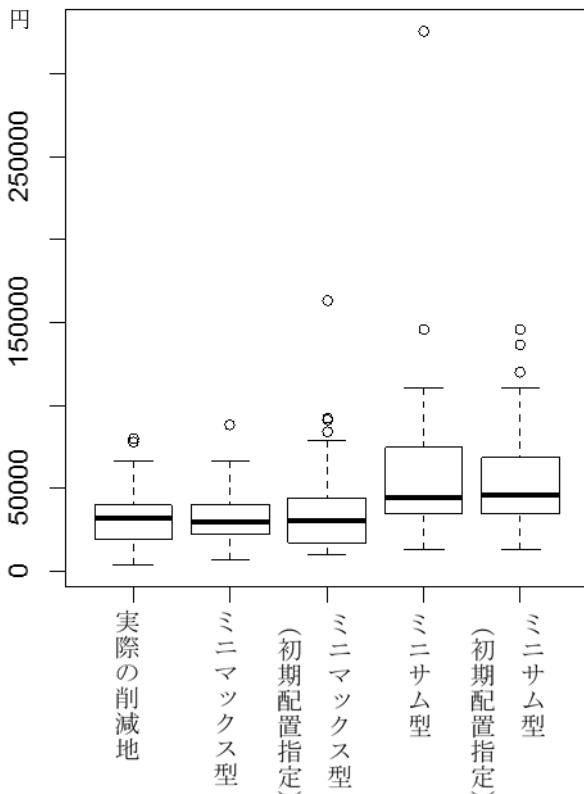


図-8 目的関数別の削減地点における地価の分布

表-5 目的関数の値の比較

	重みと期待二乗誤差の積の	
	最大値	合計値
実際の削減後の配置	4.62858×10^2	1.4494×10^5
ミニマックス型による探索結果	4.61557×10^2	-
ミニマックス型による探索結果 (初期配置指定)	4.61559×10^2	-
ミニサム型による探索結果	-	1.4388×10^5
ミニサム型による探索結果 (初期配置指定)	-	1.4389×10^5

注) 表-5における実際の削減後の配置は、平成25年の標準地の配置から削減された配置であり、選定替えの考慮をしていない。

ミニサム型の目的関数を最小化した場合は、ミニマックス型の目的関数を用いた場合よりも、実際の削減後の配置との一致率が高くなる傾向がみられた。他に興味深い点は、ミニサム型を用いた場合、初期配置指定の場合とそうでない場合の削減推奨地点の一致率 (80.4%) が突出して高いということである。ミニマックス型でも同様に一致率 (15.2%) は高い傾向にある。ミニマックス型とミニサム型にて求めた結果の一致率を比べると、初期配置を指定した場合、そうでない場合に比べて一致率が高い。初期配置が最終的な結果に影響を与えている可能性がある。

図-3, 4, 5, 6, 7を比べると茨城県県北に位置する大子町において、実際には削減されているが、計算によって得られた結果では削減していないことがわかる。大子町周辺は標準地点が少なく、標準地削減による期待二乗誤差の上昇幅が他の地域よりも大きく出るものと考えられる。

それぞれ削減された地点の価格帯を確認すると、ミニマックス型の計算式により得られた削減地点の価格分布は、ミニサム型の計算式により得られた削減地点の価格分布よりも実際のそれに近いことがわかる。削減地点の一致率はミニサム型より得られる結果である場合が高いが、価格帯に関してはミニマックス型から得られる結果がより類似する結果となった。

算出された目的関数の値を比べると、いずれの結果においても実際の削減後の配置から得られる目的関数の値を下回っているものの、実際には最適解が得られているかどうか不明である。しかし、初期値は経験・試行錯誤により設定しており大域的な最適解に近い解であると考えられる。さらに興味深い結果として、初期配置をランダムに与えた場合、初期配置を指定した場合よりも良い結果を得られたことが挙げられる。また、計算の際にランダムに与えた

初期配置から得られた目的関数の値は、ミニサム型の計算を行う際には実際の削減後の配置から得られる値よりも大きい値を得たが、ミニマックス型の際には小さい値となっていた。

以上より、削減地点に関して違いはあるものの得られた目的関数の値に大きな違いはなく、実際の削減は統計的な観点からも効率的に行われている可能性が示唆される。一方で、村上・堤 (2015) の手法から得られた配置の目的関数値は全ケースで実際の配置よりも改善しており、地球統計モデルを活用することで、地価調査地点が効果的に削減することができるとの示唆を得た。そこで、今後削減作業を行う際の初期案として、村上・堤 (2015) の手法を活用する可能性が考えられる。例えば、目的関数の値に大きな差のない配置案を複数、目的関数ごとに用意し、その中から実務的観点を踏まえて削減を進めてゆくというような活用方法が考えられる。

6. 本研究の総括

本研究は、村上・堤 (2015) が示した、地価公示・都道府県地価調査における標準地・基準地削減に対する地球統計学的アプローチの有用性を踏まえ、実際に行われた削減との比較・考察をした。統計学的見地から、実際の削減パターンよりも良い削減パターンがあることを定量的に示し、村上・堤 (2015) が示した地価調査地点の配置問題に対する地球統計学的手法を用いたサンプリングデザインの最適化の有用性をより確かなものとした。

今後の課題として、計算アルゴリズムの改良をする必要がある。本研究における最終的な計算時間はおおよそ16時間であったが、その前段階でパラメータの設定などに16時間以上の時間をかけており、実用上に課題が残る。より短時間でより良い解にたどり着ける方法を模索する必要がある。

参考文献

- 1) 国土交通省：平成 24 年行政事業レビューシート，事業名：地価公示，事業番号：0117，入手先〈<http://www.mlit.go.jp/common/000223757.pdf>〉，（最終閲覧日：2016 年 4 月 22 日）。
- 2) 村上大輔，堤盛人：地価公示・都道府県地価調査における標準地・基準地点の配置問題への地球統計学の応用，土木学会論文集 D3（土木計画学），72(2)，31-43，2015。
- 3) Cressie, N. : Statistics for Spatial Data. Revised Edition, John Wiley & Sons, 1993.
- 4) Zimmerman, D. L. : Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction, *Environmetrics*, Vol. 17, pp. 635-652, 2006.
- 5) 間瀬茂：地球統計学とクリギング法 R と geoR によるデータ解析，オーム社，2010。

- 6) Zhu, Z. and Stein, M. L. : Spatial sampling design for parameter estimation of the covariance function, *Journal of Statistics Planning and Inference*, Vol. 134, No. 2, pp.583-603, 2005.
- 7) 柳浦睦憲・茨木俊秀：組み合わせ最適化-メタ戦略を中心として-，朝倉書店，2001。
- 8) Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. : Optimized by simulated annealing, *American Association for the Advancement of Science*, Vol. 220, No. 4598, pp. 671-680, 1983.

(2016. 4. 22 受付)