実用性を考慮したローカルな不動産価格指数の理論と実証

松尾　和史[1]・村上　大輔[2]・堤　盛人[3]・今関　豊和[4]

[1]学生会員　筑波大学大学院　システム情報工学研究群（〒305-8573 茨城県つくば市天王台 1-1-1）
E-mail: matsuo.kazushi.sa@alumni.tsukuba.ac.jp (Corresponding Author)

[2]正会員　統計数理研究所助教　データ科学研究系（〒190-0014 東京都立川市緑町 10-3）
E-mail: dmuraka@ism.ac.jp

[3]正会員　筑波大学教授　システム情報系（〒305-8573 茨城県つくば市天王台 1-1-1）
E-mail: tsutsumi@sk.tsukuba.ac.jp

[4]非会員　株式会社オフィスビル総合研究所（〒104-0061 東京都中央区銀座 4-6-1）
E-mail: toyokazu_imazeki@sanko-e.co.jp

近年，場所ごとに異なる値を与えるローカルな不動産価格指数への注目が高まっている．しかし，既存の推定法は，計算効率や指数の安定性の面で課題があり，実用化には至っていない．そこで本研究では，既存の推定法の課題を克服する新たなローカルな不動産価格指数の推定法を提案する．実証分析の結果，重複期間法とモラン固有ベクトルに基づいた空間可変係数モデルを組み合わせた提案手法は計算効率と指数の安定性の面で既存手法に優る一方で，推定精度や捉えられる不動産価格の空間的異質性の綿密さの面では既存手法に劣ることが明らかになった．これらの結果は，提案手法がローカルな不動産価格の実用性の向上に寄与することを示している．但し，既存手法に比べ劣る点もあり，提案手法に改善の余地があることも示している．

Key Words: *local real estate price index, office market, practicality, rolling window, spatially varying coefficients model*

1.　はじめに

不動産は他の一般的な財と異なり，同質の財が存在せず，財の品質が経年的に変化し，取引の頻度も少ない．そのため，不動産市場の挙動を捉える場合は，時点ごとに取引された財の違いを考慮した代表値，すなわち不動産価格指数を定めたうえで，時点間の指数の変化や地域間の値の差異を比較する（Zhu et al., 2022[1)]）．一般的な不動産価格指数は，基準とする時点を 100 とした指数で表され，地域や財の規模などによって分割されたサブ市場ごとに 1 つの時系列が与えられる．この時，指数を定めた地域内においても立地によって不動産価格の分布は不均一であることが知られており，このような空間的異質性をどのように捉えるかについて，様々な議論がなされている（e.g., Pace et al., 1998[2)]）．

特に近年の議論では，ノンパラメトリックもしくはセミパラメトリックな手法を用いて，任意の地理座標・時点ごとに異なる不動産価格指数を付与する方法（ローカルな不動産価格指数）に関する研究に注目が寄せられている（e.g., Agarwal et al., 2021[3)]）．このようなローカルな不動産価格指数の大きな特徴は，都市内の小地域ごとの時間変化を詳細に追えることであり，複雑な不動産市場の構造の解明に寄与することである．また，地域区分を定める必要がないため，可変単位地区問題（MAUP: Modifiable Areal Unit Problem）も回避することができる（Clapp, 2004[4)]）．国全体などのよりグローバルな不動産価格指数を推定する場合においても，小さな地域区分（もしくは不動産種別）ごとの不動産価格指数に基づいて算出することで，集計バイアスを削減することができ（Malone & Redfearn, 2022[5)]），よりローカルな視点で不動産市場を捉える重要性が増している．

しかし，既存のローカルな不動産価格指数の推定法は計算効率や指数の安定性の面で課題があり，実用的には用いられていない．

そこで，本研究では，実用性を考慮したローカルな不動産価格指数の推定方法を提示し，2000 年 1 月から 2021 年 12 月までの東京オフィス市場のデータを用い

て，実証分析を行うことで，提案手法の実用性を検証する．

以下の本文では，第２章において，既存の不動産価格指数の推定手法について整理し，第３章において提案手法の概要を示す．第４章において，既存手法と提案手法を実データに適用し，特徴の違いと提案手法の実用性について議論する．そして，最後に第５章において，結論と課題を述べる．

2. 既存研究

(1) 一般的な不動産価格指数の推定手法

現在，多くの国の公的な不動産価格指数の推定において用いられているヘドニックアプローチ（Rosen, 1974[6]）は，不動産価格を不動産が有する様々な属性の束として捉える方法論である．ここで不動産の属性には，不動産の構造や規模，設備などの建物要因や，周辺環境などの立地属性，取引に関する情報や，経済状況などの情報が用いられ，一般的に式１のような形で定式化される．

$$lnP_{i,t} = \beta_0 + \sum_{j=1}^{J} X_{j,i,t}\,\beta_j + \delta_t + \varepsilon_{i,t},\quad \varepsilon_{i,t} \sim N(0,\sigma^2) \tag{1}$$

ここで，$lnP_{i,t}$は時点tにおける不動産iの価格（賃料）を表し，$X_{j,i,t}$は時点tにおける不動産iのj番目の属性値を，$\varepsilon_{i,t}$は期待値 0，分散σ^2の正規分布に従う誤差項を表す．また，δ_tは時間効果を表し，時間ダミーを説明変数として加え，その係数を推定することで得られる．この時間効果δ_tは，不動産の様々な属性による価格の差異をコントロールした上で得られる基準時点（$t = 1$）と比較した時点tの平均的な価格上昇率と解釈することができる．そのため，時間ダミーの係数に基づいた不動産価格指数は「ヘドニック時間ダミー価格指数（hedonic time dummy index; HDI）」と呼ばれ，基準時点を 100 として，式２で表される．

$$HDI_t = \exp(\delta_t) \times 100,\ \ \delta_1 = 0 \tag{2}$$

なお、式１の推定には，最小二乗法（Ordinary Least Squares regression; OLS）が用いられることが多い．

但し，前述の方法論は，様々な属性の価値が時間的に不変であることや，実務上は新たな時点のデータが追加される度に，モデルの係数が変化するため，過去の指数が一定に定まらないという問題がある．そのため，日本や欧州の公的な不動産価格指数の推定には，データを一定期間ごとにプールし，プールされた期間ごとにヘドニック回帰を行い，時点ダミーの係数を結合させせることで，前述の２つの課題を克服する重複期間型時間ダミー指数（Rolling time Dummy Index; RDI）が用いられている（Eurostat, 2017[7]）．

RDI の構築には，はじめに，全期間$1,2,\cdots$T期のデータからγ期からはじまるτ期間[$\gamma, \gamma+\tau-1$]のデータをプールし，τ期間を１期ずつずらしながら式１のモデルを推定する．ここで，$t+\tau-1$期から$t+\tau$期までの不動産価格の変化は式３で表される．

$$\frac{P_{t+\tau}}{P_{t+\tau-1}} = \frac{\exp\left(\hat{\delta}_{t+\tau}\right)}{\exp\left(\hat{\delta}_{t+\tau-1}\right)} \tag{3}$$

時間効果δ_tを，r期を用いて，$\hat{\delta}_r^{(r)}, \hat{\delta}_{r+1}^{(r)}, \cdots, \hat{\delta}_{r+\tau-1}^{(r)}$と定義し，全期間のうち最初の[$1,\tau$]期間の価格指数を，$RDI_1 = 100, RDI_2 = \exp\ (\hat{\delta}_2^{(1)}),\ \cdots, RDI_\tau = \exp\ (\hat{\delta}_\tau^{(1)})$とすると，$\tau+1$期の価格指数$\mathrm{RDI}_{\tau+1}$は式４で表される．

$$\begin{aligned} RDI_{\tau+1} &= RDI_\tau \times \frac{P_{\tau+1}}{P_\tau} \\ &= \exp(\hat{\delta}_\tau^{(1)}) \times \frac{\exp\left(\hat{\delta}_\tau^{(2)}\right)}{\exp\left(\hat{\delta}_{\tau-1}^{(2)}\right)} \end{aligned} \tag{4}$$

したがって，$\tau+1$期以降のk期の価格指数は，$\mathrm{k}-1$期の価格指数$\mathrm{RDI}_{\mathrm{k}-1}$とk期から逆算してτ期間[$k-\tau-1, \mathrm{k}$]のデータをプールし，推定した結果得られた時間ダミーの係数$\hat{\delta}_{\mathrm{k}-1}^{(k-\tau-1)}$，$\hat{\delta}_{\mathrm{k}}^{(k-\tau-1)}$を用いて得ることができる．

なお，RDI の構築では事前にデータをプールするτ期間を定める必要がある．一般にこの期間は Window と呼ばれる．Window を長くすれば，指数の堅牢性が向上し，データが少ない場合に有益である．逆に，Window が短い場合，直近の市場との関連性が向上する．RDI を構築する際にはこれらのトレードオフの中で Window の長さτを定める必要がある．四半期の価格指数では２四半期（フランス，ポルトガル）や４四半期（キプロス，クロアチア）が，月次の価格指数では 12 か月（アイルランド，日本）が採用されることが多い．近年では，適切な Window の長さを定める手法に関する研究もされている（Hill et al., 2022[8]）．

(2) ローカルな不動産価格指数の推定手法

近年，関心の高まりを見せているローカルな不動産価格指数の推定には，Clapp (2004)[4]が提案した Local Regression Model (LRM)が広く用いられている．時空間グリッド上に柔軟に不動産価格指数を定めることができる LRM は，パラメトリックな部分とノンパラメトリックな部分に分けて説明される（式５）．

$$lnP_{i,t} = \sum_{j=1}^{J} X_{j,i,t}\,\beta_j + f\left(\boldsymbol{Z}_{i,t}\right) + \varepsilon_{i,t},\quad \varepsilon_{i,t} \sim N(0,\sigma^2) \tag{5}$$

ここで，$\sum_j^J X_{j,i,t}\,\beta_j$は式 1 と同様に，不動産の有する様々な属性の価値の束であり，パラメトリックに扱われる．それに対して，$f\left(\boldsymbol{Z}_{i,t}\right)$は緯度・経度・時間で表される 3 次元ベクトル$\boldsymbol{Z}_{i,t}$に基づいて定められる時空間の

滑らかな関数であり，パラメトリックな部分では説明できない，時間的・空間的な不動産価格の変化を表す．

LRM の推定は大きく 3 つにステップに分けられる．

Step1.

はじめに，標準的なノンパラメトリック回帰手法で$E[lnP_{i,t}|\boldsymbol{Z}_{i,t}]$と$E[X_{j,i,t}|\boldsymbol{Z}_{i,t}]$を推定する．ここで，ノンパラメトリック回帰の手法には，スプライン（Hill & Scholz, 2018[9]）や高次の局所多項式回帰（Clapp, 2004[4]），Nadaraya-Watson カーネル回帰（Zhu et al., 2022[1]）などが用いられる．

Step2.

次に，Step1 で推定したノンパラメトリック回帰の残差（$lnP^*_{i,t}$と$X^*_{1,i,t},\dots,X^*_{J,i,t}$）を用いて，式 6 を OLS で推定し，$\beta_j$を導出する．

$$
\begin{gathered}
lnP^*_{i,t} = \beta_0 + \sum_{j=1}^{J} X^*_{j,i,t}\,\beta_j + \varepsilon_{i,t} \\
lnP^*_{i,t} = lnP_{i,t} - E\left[lnP_{i,t}\middle|\boldsymbol{Z}_{i,t}\right] \\
X^*_{j,i,t} = X_{j,i,t} - E[X_{j,i,t}|\boldsymbol{Z}_{i,t}]
\end{gathered} \tag{6}
$$

Step3.

最後に，推定された$\hat{\beta}_j$を用いて，$\varphi_{i,t}$を下記式で算出し，$\varphi_{i,t}$の期待値をノンパラメトリック回帰で推定することで，$f(Z_{i,t})$を得る（式 7）．

$$
\begin{gathered}
\varphi_{i,t} = lnP_{i,t} - \sum_{j=1}^{J} X_{j,i,t}\,\hat{\beta}_j \\
f(\boldsymbol{Z}_{i,t}) = E[\varphi_{i,t}|\boldsymbol{Z}_{i,t}]
\end{gathered} \tag{7}
$$

ここで，式4における定数項β_0は$f(\boldsymbol{Z}_{i,t})$に含まれている点に注意されたい．また，HDI や RDI と同様に基準時点を 100 とした指数に変換するため，$t = 1$における平均値$\overline{f(\boldsymbol{Z}_{\iota,1})}$を基準とした変化率の値に変換し，LRM に基づいたローカルな不動産価格指数（LRM based Local Index）$LLI_{\mathrm{i,t}}$とする(式 8).

$$
LLI_{\mathrm{i,t}} = \frac{\exp(f(\boldsymbol{Z}_{i,t}))}{\exp(\overline{f(\boldsymbol{Z}_{\iota,1})})} \tag{8}
$$

LRM の大きな特徴は，一般的なノンパラメトリック回帰と OLS で推定できるため，推定が容易であることである．但し，ノンパラメトリック回帰の推定に必要なパラメータ（バンド幅等）の決定において繰り返し処理を要することや，ノンパラメトリック回帰を各変数に対して個別に行う必要があるため，計算時間がサンプルサイズやモデルに含める属性（説明変数）の数に大きく依存するという課題もある．また，実用性の面では，式 1 で示した一般的なヘドニックアプローチに基づいた HDI と同様に，データの更新による指数の可変性も問題となる．LLI は HDI と異なり，時間方向の連続性を仮定しているため，データを一定期間ごとにプールして推定する重複期間法（Rolling Window approach）は適さないため，これらの課題の克服には異なるアプローチでの改善が必要である．

3. 提案手法

本研究では LRM とは異なるセミパラメトリック法を適用し，重複機関法と組合せることで，既存手法の課題の解決を図る．

不動産価格などの空間データの分析において，空間的異質性を捉える方法論として，カーネル回帰を援用した GWR (Geographically Weighted Regression; Brunsdon et al., 1996[10]; Fotheringham et al., 2002[11])や LRM の他に，M-SVC (Moran’s eigenvector-based Spatially Varying Coefficients) model がある．同手法は，固有ベクトル空間フィルタリング (Eigenvector Spatial Filtering; ESF; Griffith, 1996[12])を拡張したものであり，GWR に比べ係数が柔軟に定式化でき，計算時間がサンプルサイズに依存せず高速計算が可能であるという特徴を有している（Murakami et al., 2017[13]）．本研究では，M-SVC をローカルな不動産価格指数の推定に適用することを提案する．

本研究で提案する M-SVC を用いたローカルな不動産価格指数は式 9 で表される．

$$
\begin{gathered}
lnP_{i,t} = g_{MC,0}(\boldsymbol{s}_i) + \sum_{j}^{J} X_{j,i,t}\,\beta_j + \delta_{i,t} + \varepsilon_{i,t}, \\
\delta_{i,t} = b_t + g_{MC,t}(\boldsymbol{s}_i), \qquad \varepsilon_{i,t} \sim N(0,\sigma^2)
\end{gathered} \tag{9}
$$

ここで，$g_{MC,0}(\boldsymbol{s}_i)$は ESF に基づいて得られる地点ごとに異なる係数（空間可変係数）であり，定数項と見なされる．また，時間ダミーの係数として得られる時間効果$\delta_{i,t}$は基準時点（$t = 1$）と比較した時点tの平均的な価格上昇率b_tと各地点における平均的な価格上昇率からのズレ$g_{MC,t}(\boldsymbol{s}_i)$によって表される．空間的に可変な定数項と時間効果を用いることで，LRM と類似した形で M-SVC に基づいたローカルな不動産価格指数（M-SVC based Local Index）$MLI_{\mathrm{i,t}}$を導出することができる（式 10）．

$$
\begin{gathered}
f(\boldsymbol{Z}_{i,t}) = g_{MC,0}(\boldsymbol{s}_i) + \delta_{i,t} \\
MLI_{\mathrm{i,t}} = \frac{\exp(f(\boldsymbol{Z}_{i,t}))}{\exp(\overline{f(\boldsymbol{Z}_{\iota,1})})}
\end{gathered} \tag{10}
$$

空間可変係数の推定に用いる ESF は，空間的自己相関の代表的な検定統計量である Moran’s I 統計量（Moran, 1948[14]）における近接性の定義を用いて，変数の空間的自己相関を考慮するアプローチである．

Moran’s I 統計量I_{moran}はサンプルサイズをn，$n \times 1$の変数のベクトルを$\boldsymbol{y}$，$n \times n$の単位ベクトルを$\boldsymbol{I}$，1 を各要素に持つ$n \times 1$のベクトルを$\boldsymbol{1}$，$n \times n$の隣接行列を$\boldsymbol{C}$として，式 11 で表される．

$$I_{moran} = \frac{N}{\mathbf{1}'\mathbf{C}\mathbf{1}} \frac{\boldsymbol{y}'\left(\boldsymbol{I}-\frac{\mathbf{1}\mathbf{1}'}{n}\right)\boldsymbol{C}\left(\boldsymbol{I}-\frac{\mathbf{1}\mathbf{1}'}{n}\right)\boldsymbol{y}}{\boldsymbol{y}'\left(\boldsymbol{I}-\frac{\mathbf{1}\mathbf{1}'}{n}\right)\boldsymbol{y}} \tag{11}$$

Moran's I 統計量が，正に有意な値を取るとき$\boldsymbol{y}$は正の自己相関を，負に有意な値を取るとき$\boldsymbol{y}$は負の自己相関を持つとされる．また，Moran's I 統計量は$(\boldsymbol{I}-\mathbf{1}\mathbf{1}'/n)\boldsymbol{C}(\boldsymbol{I}-\mathbf{1}\mathbf{1}'/n)$を用いて近接性を定義した統計量とみなすことができる．

ESF では，モラン I 統計量で用いられる近接行列の固有ベクトルを用いる（式 12）．

$$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n\} = evec[(\boldsymbol{I}-\mathbf{1}\mathbf{1}'/n)\boldsymbol{C}(\boldsymbol{I}-\mathbf{1}\mathbf{1}'/n)] \tag{12}$$

ここで$evec[\ ]$は括弧内の行列の固有ベクトルを与える演算子である．$\boldsymbol{e}_1$は最大の固有値を持つ固有ベクトルであり，Moran's I 統計量で説明可能な最も大域的な空間パターンを表す．$\boldsymbol{e}_2$は 2 番目に大きな固有値を持つ固有ベクトルであり，$\boldsymbol{e}_1$と無相関な成分のうち，Moran's I 統計量で説明可能な最も大域的な空間パターンを表す．同様に$\boldsymbol{e}_n$は$n-1$番目に大きな固有値を持つ固有ベクトルであり，$n-1$番目までの全ての固有ベクトルと無相関な成分のうち，Moran's I 統計量で説明可能な最も大域的な空間パターンを表す．

空間可変係数$f_{MC,0}(\boldsymbol{s}_i)$は正の空間相関パターンを説明する$p$個の固有ベクトル$\boldsymbol{E}_p = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_p]$と$p \times 1$の係数ベクトル$\boldsymbol{\gamma}$の線形結合によって表される（式 13）．

$$g_{MC,0}(\boldsymbol{s}_i) = \boldsymbol{E}_p \boldsymbol{\gamma} \tag{13}$$

このモデルでは，サンプルサイズnに依存する行列（被説明変数や説明変数を含む$n \times m$の行列）を内積計算によって$m \times m$の行列に置き換えることで，大規模なデータに対しても，効率的に推定が可能である（Murakami & Griffith, 2019[15]; Murakami et al., 2020[16]）．

但し，LRM とは大きく異なる点として，時間に関する変数を離散的に扱っていることがあげられる．LRM は時間を位置座標と同様に連続的に扱っているが，M-SVC では時間ダミーを用いているため，OLS に基づいた一般的な不動産価格指数と同様に，時間を離散的な変数として用いている．

また，MLI は HDI や LLI と同様に，過去の指数の安定性の面で課題がある．しかし，LLI とは異なり，時間方向の連続性を考慮していないため，重複期間法の適用が可能である．そこで本研究では，重複期間法と M-SVC を組合わせたローカルな不動産価格指数も提案する．

RDI と同様に，全期間$1,2,\cdots$ T期のデータからγ期からはじまるτ期間$[\gamma, \gamma+\tau-1]$のデータをプールし，τ期間を 1 期ずつずらしながら式 9 のモデルを推定する．

時間効果$\delta_{i,t}$を，r期を用いて，$\hat{\delta}_{i,r}^{(r)}, \hat{\delta}_{i,r+1}^{(r)}, \cdots, \hat{\delta}_{i,r+\tau-1}^{(r)}$と定義し，全期間のうち最初の$[1,\tau]$期間のセミパラメトリック項$f(Z_{i,t})$を，$f(\boldsymbol{Z}_{i,1}) = g_{MC,0}(\boldsymbol{s}_i), f(\boldsymbol{Z}_{i,2}) = g_{MC,0}(\boldsymbol{s}_i) + \hat{\delta}_{i,2}^{(1)}, \cdots, f(\boldsymbol{Z}_{i,\tau}) = g_{MC,0}(\boldsymbol{s}_i) + \hat{\delta}_{i,\tau}^{(1)}$とすると，$\tau+1$期のセミパラメトリック項$f(\boldsymbol{Z}_{i,\tau+1})$は式 4 と同様に求めることができ，RM-SVC に基づいた時空間不動産価格指数（RM-SVC based Local Index）$RMLI_{\mathrm{i,t}}$が導出される（式 14）．

$$f(\boldsymbol{Z}_{i,\tau+1}) = f(\boldsymbol{Z}_{i,\tau}) \times \frac{P_{i,\tau+1}}{P_{i,\tau}}$$
$$= \left(g_{MC,0}(\boldsymbol{s}_i) + \hat{\delta}_{\mathrm{i},\tau}^{(1)}\right) \times \frac{\exp\left(\hat{\delta}_{\mathrm{i},\tau}^{(2)}\right)}{\exp\left(\hat{\delta}_{\mathrm{i},\tau-1}^{(2)}\right)}, \tag{14}$$
$$RMLI_{\mathrm{i},\tau+1} = \frac{\exp(f(\boldsymbol{Z}_{i,\tau+1}))}{\exp(\overline{f(\boldsymbol{Z}_{\iota,1})})}$$

4. 実証分析

(1) データ

次に，2 章，3 章で示した手法について，実データに適用し，それぞれの特徴を整理する．本研究では，実証分析に用いるデータとして，三幸エステート株式会社から提供いただいたオフィス市場に関する業務データを使用する．このデータには，2000 年 1 月から 2021 年 12 月までにおける計 7,322,841 件の標準的な賃貸オフィスビルの位置情報や貸付床面積，地上階数，築年数等の基本的な情報が含まれており，そのうち，1,584,777 件については募集賃料に関する情報が含まれている．

ここで，標準的な賃貸オフィスビルとは事務所用途に限定されたものであり，自社保有のオフィスビルや，倉庫，店舗用途であるもの等はデータから除かれている．なお，オフィスビルのデータは，三幸エステート株式会社が把握しているものに限られている点に留意されたい．

なお，データの網羅率については，賃貸オフィスビルに関する公的統計が存在しないため，正確には把握することはできない．但し，本研究で扱うデータは他の民間企業が公開しているデータと比較し，同程度のオフィスビルを網羅していることが確認されている（松尾ら，2022[17]）．

賃料に関する情報は，1 棟ごとの標準的な募集賃料を用いる．これは，オフィスビル内における複数の募集情報のうち，店舗用途やシェアオフィスといった特殊事例を取り除き，標準的な事務所用途として募集された共益費込みの 1 坪当たりの賃料情報をそのオフィスビルの募集賃料の代表値として用いたものである．

LRM や M-SVC の推定では，地理的な変数の影響は全てノンパラメトリックな部分に含まれるため，**表-1** における建物属性を表す 5 つの変数を説明変数$X_{j,i,t}$として用いる．OLS については，前述の 5 つの変数に加え，地理

表-1　変数の概要と記述統計

	変数	単位	概要	最小	最大	平均	標準偏差
	募集賃料	対数・円/坪	１坪当たりの共益費込み募集賃料	8.24	11.08	9.58	0.34
建物属性	地上階数	対数・階	建物の地上階数	1.10	4.09	2.02	0.35
	1フロア面積	対数・坪	1フロアあたりの床面積	1.10	8.16	3.82	0.90
	築年数	年	竣工からの経過年数	0.00	94.00	24.93	11.66
	建物内空室率	%	建物内の空室率	0.00	1.00	0.23	0.26
	冷房	ダミー	冷房設備の有無	0.00	1.00	0.98	0.16
地理的属性	CBD	対数・m	東京駅までの直線距離	5.33	9.81	8.06	0.75
	周辺貸付面積	対数・坪	半径 500m 圏内のオフィスビルの貸付可能床面積	2.48	13.45	11.43	1.24
	周辺空室率	%	半径 500m 圏内のオフィスビルの平均空室率	0.00	1.00	0.06	0.04
	駅徒歩	対数・分	最寄駅からの駅徒歩時間	0.69	3.43	1.40	0.49
	建物密集度	対数・m	最寄り 20 棟までの平均距離	3.10	8.34	4.36	0.63
	周辺高さ	対数・階	半径 500m 圏内のオフィスビルの平均地上階数	1.10	3.29	1.93	0.14

的な属性を表す変数や，時点ダミー，地域ダミーを説明変数として加え，推定を行う．

(2) 手法間の比較

はじめに，既存手法と提案手法の比較を行うため，重複期間法を適用していない，HDI，LLI，MLI の 3 つの手法の比較を行う．また，LRM はサンプルサイズが大きい場合，膨大な計算時間を必要とするため，各年 12 月時点のデータを抽出し，無作為に抽出した 20,000 のデータに対して各手法を適用した．

表−2 は各モデルにおける説明変数の係数と決定係数を示したものである．モデル間で説明変数の有意性に差異はあるものの，係数の符号とその大きさはおおよそ一致していることが読み取れる．なお，LRM におけるノンパラメトリック回帰には，統計分析ソフト R の np パッケージで容易に実装可能な Nadaraya-Watson カーネル回帰を使用した．カーネル回帰におけるカーネル関数にはガウスカーネルを利用し，バンド幅は交差検証の結果，緯度方向は 0.000793 度，経度方向は 0.000706 度，時間方向は 3.53 年が選択された．以下の分析では，地理的属性を加えた**表−2** の列(2)の OLS と，列(3)，(4)に示されている LRM，M-SVC の 3 つのモデルの比較を行う．

a) 指数の推移と分布

次に，2000 年を基準とした各モデルに基づいたオフィス賃料指数の推移を**図-1** に示す．地点ごとに異なる値を付与する LLI と MLI は各時点におけるセミパラメトリック項 $f(Z_{i,t})$ の平均値をその時点における代表的な賃料として，HDI と同様に時点ごとに 1 つの指数を算出した．

表-2　各モデルの推定結果の比較

model	(1) OLS coeff	t-value		(2) OLS coeff	t-value		(3) LRM coeff	t-value		(4) M-SVC coeff	t-value	
(Intercept)	8.747	539.63	***	8.477	97.54	***						
１フロア面積	0.113	57.30	***	0.119	59.08	***	0.149	33.80	***	0.112	61.44	***
地上階数	0.182	34.79	***	0.154	29.44	***	0.184	14.97	***	0.177	36.57	***
築年数	-0.007	-49.89	***	-0.008	-55.32	***	-0.011	-56.20	***	-0.008	-60.24	***
建物内空室率	-0.015	-2.46	**	-0.005	-0.81		-0.049	-11.22	***	-0.009	-1.60	
冷房	-0.080	-8.29	***	-0.076	-8.15	***	-0.007	-0.53		-0.071	-8.01	***
駅徒歩				-0.093	-29.08	***						
周辺貸付面積				0.049	12.45	***						
周辺空室率				-0.514	-10.24	***						
建物密集度				0.005	1.01							
周辺高さ				-0.067	-3.76	***						
CBD				0.014	1.94	*						
Year dummies	Yes			Yes			No			Yes		
District dummies	Yes			Yes			No			No		
Observations	20,000			20,000			20,000			20,000		
Adjusted-R^2	0.628			0.654			0.819			0.702		

*p<0.1; **p<0.05; ***p<0.01. District dummies は三幸エステート株式会社が定める地域区分に準拠している[注1)].

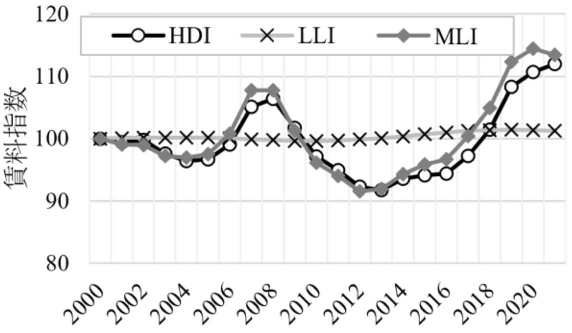


図-1 オフィス賃料指数の推移

指数の推移は，HDI と MLI が非常に類似した傾向を示している一方で，LLI は時点間の変動が小さく，極めて滑らかな推移を示していることが読み取れる．LLI は，時間方向の連続性を考慮しているため，時間方向のバンド幅が大きい場合，時点による局所的な変化を捉えにくいことが考えられる．

また，指数の空間分布を確認するために，指数推移のピーク（2008 年，2020 年）と底（2004 年，2012 年）における 250m メッシュごとの LLI と MLI を算出し，可視化した（**図-2**）．ここで，オフィスビルが立地しない地域を対象地域から除外するため，メッシュの中心点を基準として半径 250m 圏内にオフィスビルが立地するメッシュのみを対象とした．また，賃料指数の基準値は，**図-1** と同様に，2000 年における賃料分布の平均値を 100 としている．LLI と MLI の分布を比較すると，LLI の方が，空間的に局所的な変動を捉えていることが読み取れる．これは，LRM の推定において緯度・経度方向のバンド幅が短いことに起因していると考えられる．

b) 推定精度

次に，各モデルの推定精度の比較結果を示す．本研究では，予測精度を示す指標として，絶対平均誤差（MAE: Mean Absolute Error, 式 15）と対数平均二乗誤差（RMSLE: Root Mean Squared Logarithmic Error, 式 16）を用いる．

$$MAE = \frac{\sum |P_{i,t} - \hat{P}_{i,t}|}{n} \tag{15}$$

$$RMSLE = \sqrt{\frac{\sum (lnP_{i,t} - ln\hat{P}_{i,t})^2}{n}} \tag{16}$$

精度の算出には，モデル推定に用いた 2 万件のサンプル（in-sample）とモデル推定に用いていない約 11 万件のサンプル（out-of-sample）の両方を用いた．

その結果，いずれも LRM，M-SVC，OLS の順に精度

表-3 各モデルの推定精度

		MAE	RMSLE	n
in-sample	OLS	2385.35	0.203	20,000
	LRM	1684.01	0.147	
	M-SVC	2179.26	0.189	
out-of-sample	OLS	2393.68	0.204	108,720
	LRM	2199.87	0.190	
	M-SVC	2334.57	0.200	

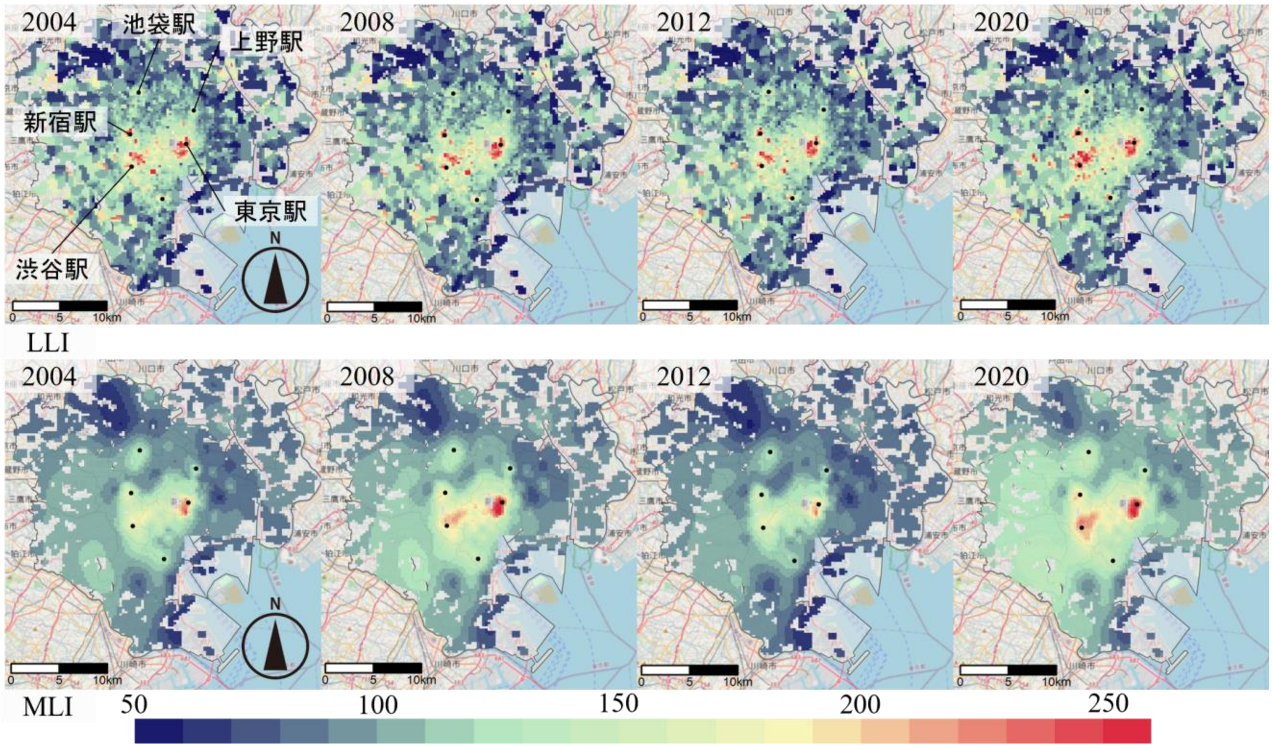


図-2 オフィス賃料指数の空間分布

が高いことが明らかになった（**表-3**）．M-SVC は近似を用いているため，全サンプルを用いて推定する LRM に精度の面で劣ることは妥当な結果である．

c) 残差の空間的自己相関

次に，各モデルがオフィス賃料の空間的な構造をどの程度捉えているのかを検証するために，各モデルの残差に対して Moran's I 統計量を用いて空間的自己相関の検定を行った（**図-3**）．その結果，OLS の残差はすべての時点において 5%水準で有意に空間的自己相関がある一方で，M-SVC は一部の時点で，LRM は全ての時点で，空間的自己相関がないという帰無仮説が 5%水準で棄却されないという結果が示された．この結果は，M-SVC や LRM は OLS では捉えられていない賃料の空間的異質性を捕捉していることを示唆している．但し，M-SVC が捉えている賃料の空間的異質性は部分的であり，LRM には及ばないことも読み取れる．

d) 計算効率

最後に，サンプルサイズや，対象期間の長短によって，計算効率がどのように変化するのかを検証する．サンプルサイズごとの比較では，サンプルサイズを 5,000 件から 20,000 件まで，対象期間による比較では，サンプルサイズを一定（20,000 件）に固定し，対象期間のみ 11 時点（2000 年～2010 年）から 22 時点（2000 年～2021 年）まで変化させて，各パターンにおける計算時間を計測した（**図-4**）．なお，本研究では DELL Precision (2.20GHz, 10-Core Intel Xeon Silver 4210 with 64GB memory)を使用する．また，モデルの推定には R(version; 4.0.3)を用いており，LRM の推定には np(version; 0.60-11)，M-SVC の推定には spmoran(version; 0.2.2.4)を使用する．

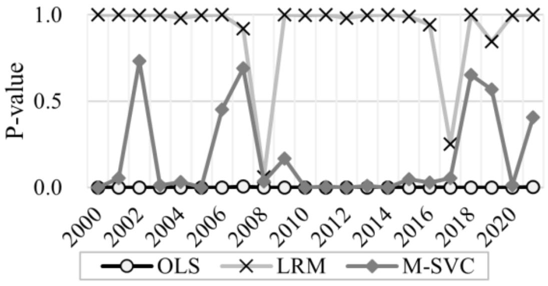


図-3 残差の空間的自己相関の推移

計測の結果，LRM はサンプルサイズが増えると計算時間が線形的に増加するが，M-SVC は計算時間がサンプルサイズに依存せず一定の水準であることが明らかになった．特に 2 万サンプルの場合，M-SVC の計算に要した時間は約 5.4 万秒（0.63 日）であるのに対して，LRM の計算に要した時間は約 27 万秒（3.13 日）であり，両者には 5 倍程度の差があることが読み取れる．但し，空間可変係数で捉える変数の数に対して，サンプルサイズが少ない場合，M-SVC では計算できないことも明らかになった．また，LRM の場合，データに含まれる時点数が変化しても計算時間はほぼ一定に保たれるが，M-SVC の場合，時点数が多いほど，指数関数的に計算時間が増加する傾向がみられた．

これらの結果は，サンプルサイズが大きいほど，LRM と M-SVC の計算効率の差は広がり，M-SVC の方が効率的に推定可能であることを示している．但し，時点数が多いデータについては，M-SVC の適用が困難であることも明らかになった．そのため，時点ごとのサンプルサイズが小さく，対象期間が長い場合には，LRM の方が効率的になることも示唆しており，データの特徴によって有効的なモデルが異なることを示している．

(3) RMLI の適用

前節の結果は，大規模なデータに対する計算効率の面で M-SVC が有効である一方で，時点数が多い場合には M-SVC が不適であることを示している．3 章で示した重複期間法を適用した RM-SVC は，時点数の増加に伴う計算効率の悪化という欠点を補うという点においても有益であると考えられる．さらに，実用上，重複期間法を適用することで，新たな時点が追加されるごとに求められる計算負荷を大幅に抑えることが可能になる．

本節では，2000 年 1 月から 2021 年 12 月までの全期間の月次データに対して，RM-SVC を適用し，月次のローカルなオフィス賃料指数を推定する．

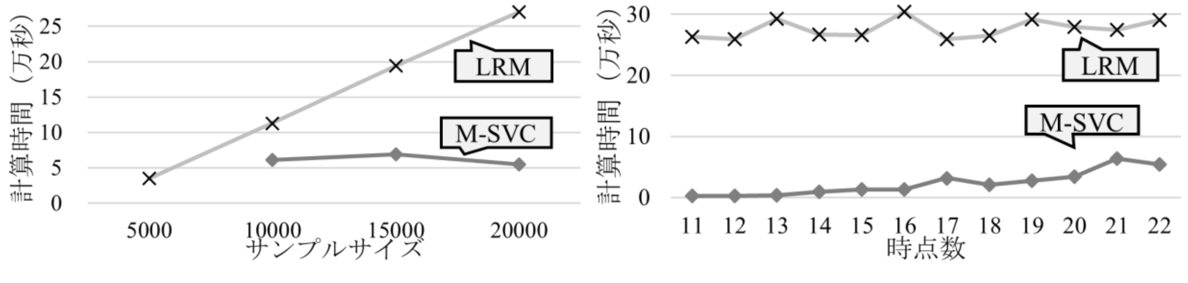
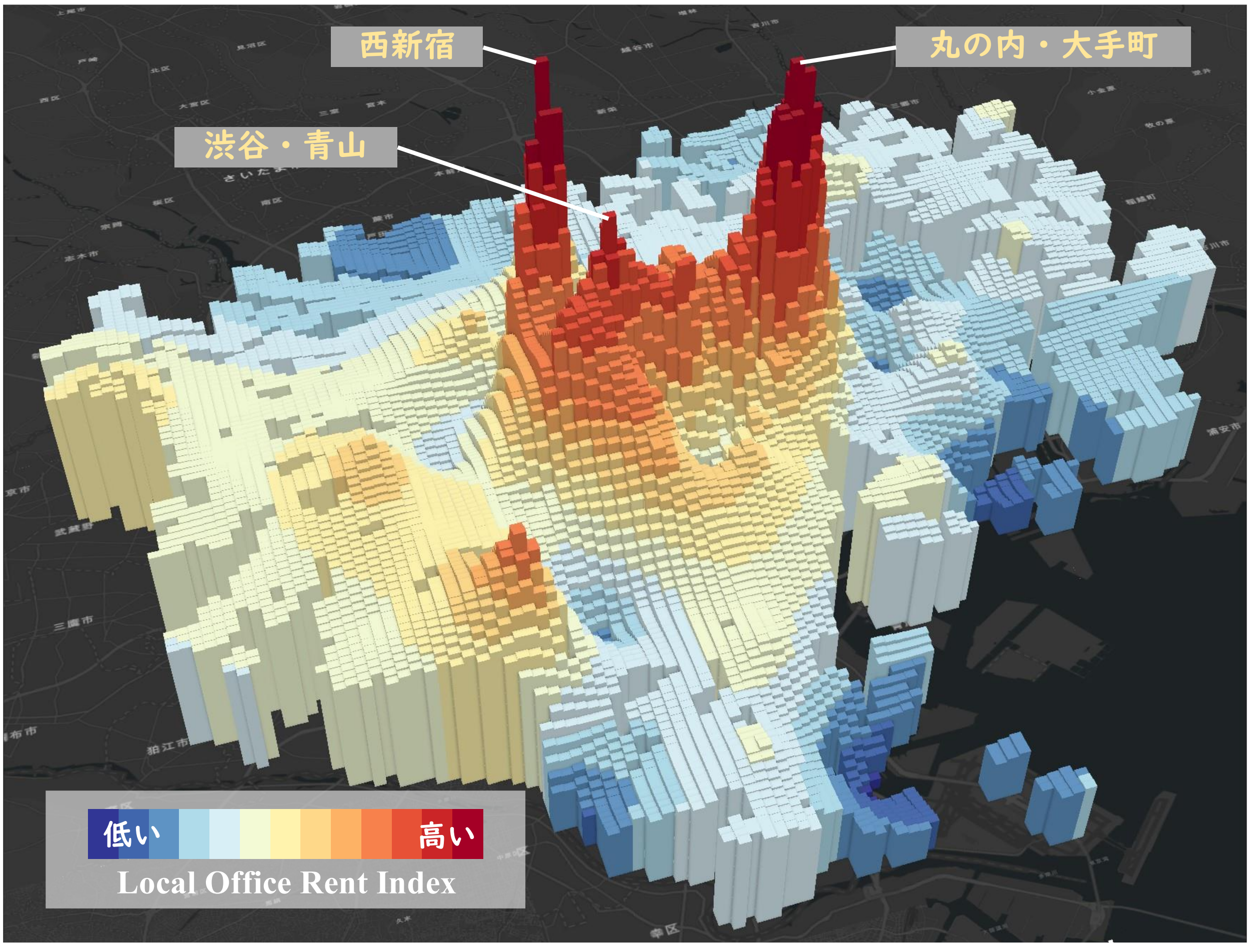


図-4 サンプルサイズや時点数による計算時間の差異



図–5 オフィス賃料指数の空間分布（2021 年 12 月時点）

a) RMLI の地域性

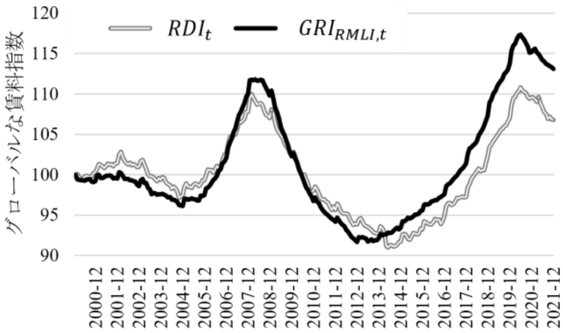
推定の結果，図–2 における LLI と同様に，滑らかな賃料の分布を表す指数が月次単位で得ることができた．また，本指数は図–5 に示すように 3 次元でより直感的に賃料の分布の地域性を表すことができる．また，RMLI の大きな特徴は，図–5 に示すような分布を月次単位で表すことができることであり，任意の地点の過去 20 年の指数の変動等も柔軟に示すことができる．

b) RDI と RMLI の比較

次に RDI と RMLI に基づいた都市単位のグローバルな賃料指数の比較を行う．RMLI は地点ごとに異なる値を与えるため，都市単位のグローバルな指数を構築するためには，平均値等の代表値を求める必要がある．本研究では，地域ごとのストックの偏りを考慮するため，RM-SVC の推定結果に基づいた RMLI をオフィスビルのストックデータに適用し，募集賃料が公表されていないビルも含め，地域内のすべてのオフィスビルiに対して，賃料指数$RMLI_{i,t}$を推計し，各オフィスビルの貸付床面積$S_{i,t}$に基づいた重み$W_{i,t}$を用いて加重平均を行うことで，RMLI に基づいたグローバルなオフィス賃料指数$GRI_{RMLI,t}$を導出する（式 17）．

$$GRI_{RMLI,t} = \sum_{i}^{n_t} W_{i,t} RMLI_{i,t} \,, \quad W_{i,t} = \frac{S_{i,t}}{\sum_{i}^{n_t} S_{i,t}} \tag{17}$$

図–6 は RDI と RM-SVC の両手法に基づいた東京 23 区におけるオフィス賃料指数の推移を表している．RM-SVC に基づいた指数GRI_{RMLI}の方が，RDI に基づいた指



図–6 東京 23 区におけるグローバルな
オフィス賃料指数の推移

数に比べ，指数の変動幅が大きく，リーマンショック後の下落局面からの回復が早いことが読み取れる．GRI_{RMLI}は，オフィスビルのストック量によって重みがつけられているため，規模の大きなビルの変化をより強く反映する．そのため，全地域のオフィスビルを均一に扱う RDI に比べ，変化の大きな都心部の影響が強く反映された指数となっていることが考えられる．

この結果は，提案手法である RMLI が，グローバルな指数の構築においても，地点ごとのストックの偏りを反映したより実態に即した指数を構築可能であるという点において有益であることを示している．また，この指標は，任意の空間的範囲ごとに柔軟に構築することができる．

5. 結論と課題

本研究では，計算効率や指数の安定性等の側面で実用上の課題があった従来のローカルな不動産価格指数の推定手法に対して，M-SVC と重複期間法を組合わせる方法論を提案し，東京オフィス市場の実データに適用することで，既存手法との比較および実用性の検証を行った．

分析の結果，

- M-SVC を用いたローカルな不動産価格指数である MLI は，計算時間がサンプルサイズにほとんど依存しないという点において有益であるが，推定精度や，捕捉できる空間的異質性の綿密さ，時点数の増加に伴う計算効率という点において，既存手法に劣ること
- M-SVC と重複期間法を組み合わせた RM-SVC は，時点数の多いデータに対しても効率的に計算でき，指数の安定性や構造変化も考慮できるため，ローカルな不動産価格指数の実用性を大きく改善させること
- RMLI とストックデータと組み合わせることで，賃料情報のない不動産も含めたストック量の空間的な偏りを考慮したグローバルな指数が構築可能であること
- RMLI は任意の地域区分ごとにグローバルな指数を構築することが可能であり，柔軟性に長けていること

が明らかになった．

本研究の貢献は，ローカルな不動産価格指数の実用性を向上させる方法論を提案し，それらの発展の可能性を示した点にある．地域性を有する不動産市場において，行政区域等の特定の空間単位に縛られず，情報の把握が可能なローカルな不動産価格指数は，不動産市場の空間的な構造とその変遷を理解するための有益な情報提供の一助になることが考えられる．

しかしながら，本研究にはいくつかの課題が残されている．一つ目は，モデルの改善に関するものである．M-SVC は，加法モデルに基づいた手法であるため，パラメトリックに推計する部分とノンパラメトリックに推計する部分を LRM よりも柔軟に指定することができ，グループ効果，非線形効果なども柔軟に加えることができるという特徴を有する．今後は，地理的な変数も加えることで，推定精度や，捉えられる空間的異質性の情報量を向上していくことが必要である．

二つ目の課題は，推定精度や捕捉可能な空間的異質性の情報量という面で有益な LRM の計算効率の向上である．LRM におけるノンパラメトリック回帰にカーネル回帰を用いる場合，バンド幅やカーネルの決定方法に，計算効率が大きく左右される．本研究では，カーネルを 1 つに定めた上で，バンド幅を交差検証によって内生的に定める手法を採用したが，事前にカーネルやバンド幅の候補を一定数に限定した上で，限定された候補の中から最適なものを採用するといった方法が用いられることもある（Zhu et al., 2022[1]）．今後の研究では，バンド幅やカーネルの与え方などの計算過程を工夫することで，計算効率を大きく改善することが求められる．

三つ目の課題には，本研究の結果は地域ごとの指数の推移の差を示したに留まっているという点が挙げられる．今後は，開発したローカルな時空間不動産価格指数に基づいて，不動産価格（賃料）の波及経路や，ローカルなダイナミクスを明らかにしていくことが求められる．

今後の研究ではこれらの課題の解決に取り組みながら，ローカルな時空間不動産価格指数の実用化を通じて，不動産市場における新たな情報資源の整備を進めていくことが重要である．

謝辞：本研究は JST 次世代研究者挑戦的研究プログラム JPMJS2124 の助成，及び一般社団法人三幸オフィス基金からの寄附を受けて実施した成果の一部である．ここに記して謝意を表する．

NOTES

注1) 三幸エステート株式会社「オフィスレントデータ 2023」 https://www.sanko-e.co.jp/data/rentdata/rentdata_2023/ （最終閲覧 2023 年 2 月）

REFERENCES

1) Zhu, E., Wu, J., Liu, H., and Li, X.: Within-City Spatial Distribution, Heterogeneity and Diffusion of House Price: Evidence from a Spatiotemporal Index for Beijing. *Real*

Estate Economics, 50(3), 621–655, 2022. https://doi.org/10.1111/1540-6229.12293

2) Pace, R. K., Barry, R., Clapp, J. M., and Rodriquez, M.: Spatiotemporal Autoregressive Models of Neighborhood Effects. *The Journal of Real Estate Finance and Economics*, 17(1), 15–33, 1998. https://doi.org/10.1023/A:1007799028599
3) Agarwal, S., Fan, Y., McMillen, D. P., and Sing, T. F.: Tracking the pulse of a city—3D real estate price heat maps. *Journal of Regional Science*, 61(3), 543–569, 2021. https://doi.org/10.1111/jors.12522
4) Clapp, J. M.: A Semiparametric Method for Estimating Local House Price Indices. *Real Estate Economics*, 32(1), 127–160, 2004. https://doi.org/10.1111/j.1080-8620.2004.00086.x
5) Malone, T., and Redfearn, C. L.: To measure globally, aggregate locally: Urban land, submarkets, and biased estimates of the housing stock. *Real Estate Economics*, 50(3), 656–671, 2022. https://doi.org/10.1111/1540-6229.12316
6) Rosen, S.: Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34–55, 1974. https://doi.org/10.1086/260169
7) Eurostat.: Technical manual on Owner-Occupied Housing and House Price Indices 2017, 2017. https://ec.europa.eu/eurostat/documents/7590317/0/Technical-Manual-OOH-HPI-2017/
8) Hill, R. J., Scholz, M., Shimizu, C., and Steurer, M.: Rolling-Time-Dummy House Price Indexes: Window Length, Linking and Options for Dealing with Low Transaction Volume. *Journal of Official Statistics*, 38(1), 127–151, 2022. https://doi.org/10.2478/JOS-2022-0007
9) Hill, R. J., and Scholz, M.: Can Geospatial Data Improve House Price Indexes? A Hedonic Imputation Approach with Splines. *Review of Income and Wealth*, 64(4), 737–756, 2018. https://doi.org/10.1111/roiw.12303
10) Brunsdon, C., Fotheringham, A. S., and Charlton, M. E.: Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298, 1996. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x
11) Fotheringham, S. A., Brunsdon, C., and Charlton, M.: *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* John Wiley & Sons, 2002.
12) Griffith, D. A.: Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying geo‐referenced data. *Canadian Geographer*, 40(4), 351‐367, 1996. https://doi.org/10.1111/J.1541-0064.1996.TB00462.X
13) Murakami, D., Yoshida, T., Seya, H., Griffith, D. A., & Yamagata, Y. (2017). A Moran coefficient-based mixed effects approach to investigate spatially varying relationships. Spatial Statistics, 19, 68–89. https://doi.org/10.1016/j.spasta.2016.12.001
14) Moran, P. A. P.: The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, 10(2), 243–251, 1948. https://doi.org/10.1111/j.2517-6161.1948.tb00012.x
15) Murakami, D., and Griffith, D. A.: Spatially varying coefficient modeling for large datasets: Eliminating N from spatial regressions. *Spatial Statistics*, 30, 39–64, 2019. https://doi.org/10.1016/j.spasta.2019.02.003
16) Murakami, D., Kajita, M., and Kajita, S.: Scalable Model Selection for Spatial Additive Mixed Modeling: Application to Crime Analysis. *ISPRS International Journal of Geo-Information*, 9(10), 577, 2020. https://doi.org/10.3390/IJGI9100577
17) 松尾和史，堤盛人，今関豊和：空室率データで見る東京オフィス市場の空間的特，*GIS -理論と応用-*, 30(1)，51‐63，2022. [Matsuo, K., Tsutsumi, M., and Imazeki, T.: Spatial characteristics of the Tokyo office market in terms of vacancy rates. *Theory and Applications of GIS*, 30(1), 51–63, 2022.]