

都市計画基礎調査データを用いた
立地履歴を考慮した住宅出現数の推定

Estimation of housing appearance considering location history by using basic survey of city planning data

室蘭工業大学建築社会基盤系学科 ○学生員 本田和樹 (Kazuki Honda)
室蘭工業大学大学院工学研究科 学生員 森本拓磨 (Takuma Morimoto)
室蘭工業大学大学院工学研究科 正会員 浅田拓海 (Takumi Asada)
室蘭工業大学大学院工学研究科 正会員 有村幹治 (Mikiharu Arimura)

1. はじめに

近年、我が国では、少子高齢化や人口減少の進展が問題視されており、多面的かつ早急な対策が求められている。また、各都市のスプロール化^{2,3)}や空き家問題⁴⁾への対策など、将来の都市構造のあり方についても様々な議論がなされている。しかし、このような議論を行うためには、第一に、現在の都市構造を把握し、将来の都市構造を定量的に分析、可視化することが必要となる。

一方、昨今では、住宅立地に関する資料のデータ化が進んでいる。都市計画に関する基礎調査（都市計画基礎調査）においても、各自治体においてデータ化が進められており、GISによる都市構造の可視化などが行われている。さらに、このデータには、建物個々の位置情報に加え、種別、建築年、延べ床面積などの詳細な属性情報が収録されており、都市構造分析や住宅立地の将来予測などの多様な解析に活用されることが期待されている。

この都市計画基礎調査データによる住宅立地分析の事例としては、浅田ら^{5,6)}の研究がある。この研究では、札幌市北区を対象に、住宅の「出現・非出現」と「消失・非消失」の判別モデルを構築し、住宅立地分布の中期的な予測を試みている。しかしながら、「出現・非出現」の判定では約60%と低い精度に留まっている。この理由としては、分析単位が住宅個々であり情報量が少ないこと、住宅系以外の建物情報を用いていないこと、過去の立地履歴を考慮していないことが挙げられる。

そこで、本研究では、札幌市の都市計画基礎調査データを用いて、上記の問題を踏まえた形で、住宅出現数推定モデルを開発する。具体的には、分析単位を4次メッシュとし、過去の住宅立地（出現・除却）履歴やメッシュ内の全建物種別の情報を目的変数として、次年度の住宅出現数を推定するモデルについて検討する。

2. データ

本研究では、札幌市の都市計画基礎調査データのうち、建物に関するデータを用い、札幌市全行政区を対象に分析を行う。このデータには、建築年に加え、表-1に示すような建物属性情報が収録されており、平成17年度から19年度、21年度、23年度、25年度、27年度まで2年間隔で構築、更新されている。これら6断面のデータを分析に用い、住宅の出現、除却、残存の3パターンを建物IDの年度間比較により判定した。図-1に示すよう

表-1 メッシュ集計値

除却住宅 (t-2年)	出現住宅 (t年)	全建物 (t年)
住宅(数)	住宅(数)	各施設(数)
築年数	-	築年数
実容積率%	実容積率%	実容積率%
実建蔽率%	実建蔽率%	実建蔽率%
延床面積m ²	延床面積m ²	延床面積m ²
敷地面積m ²	敷地面積m ²	敷地面積m ²
建築面積m ²	建築面積m ²	建築面積m ²
道路幅員m ²	道路幅員m ²	道路幅員m ²
駐車面積m ²	駐車面積m ²	駐車面積m ²
各構造(率)	各構造(率)	各構造(率)
地上階数	地上階数	地上階数

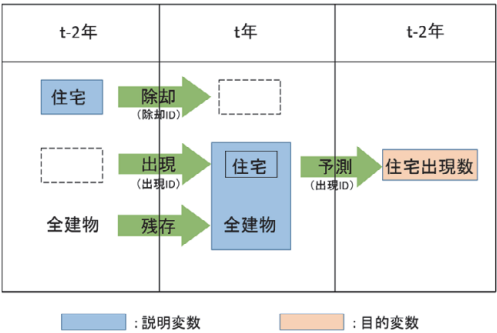


図-1 住宅立地パターンの判定

に、「出現」は次年度に住宅IDが出現する場合、「除却」は次年度に住宅IDが消失する場合とし、変化がない場合は「残存」とした。

図-2に各年度の残存住宅及び除却住宅の築年数分布を示す。残存住宅に関しては、分布の形状に大きな違いは無いが、右側にシフトしていることから、着々と高齢住宅が増えていることが分かる。例えば、築年数40年に着目すると、H19では1%程度であるが、H27では約3%にまで増加している。一方、除却住宅に関しては、全年度において築年数40年前後をピークとした正規分布になっていることが確認できる。また、このピーク値は、徐々に右側にシフトしており、除却の対象となる住宅では、古いものが増加している傾向を示している。

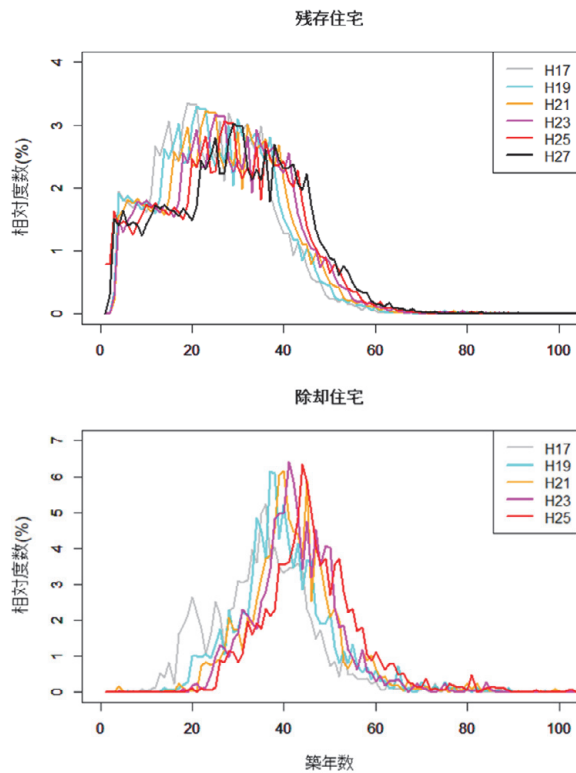


図-2 残存住宅および除却住宅の築年数分布

3. 重回帰分析による住宅出現数推定

3.1 モデルの概要

本研究では、推定精度向上を目的に、空間的な分析単位を広げ、4次メッシュ（500m×500m）を用いて、メッシュ毎に建物属性情報の集計を行った（表-1）。これらの情報を用いて、重回帰分析により次年度の出現住宅数の推定を試みる。具体的には、図-1に示したように、 $t+2$ 年度に出現する住宅数が目的変数となる。また、説明変数には、 t 年度に現存する建物種別（住宅系、商業系、文教厚生系、工業系）の建物属性平均値、さらに、本研究の特徴となる過去の立地履歴情報として、 $t-2$ から t 年度に出現、除却する住宅の建物属性平均値を加えた。なお、分析対象年度 t は平成23年度とし、他の年度については今後の課題とする。

分析対象範囲は、市街化区域であり、建物が存在する4次メッシュとする（図-3）。ただし、メッシュ内に住宅以外の敷地、施設が大部分を占める場合は、モデルの精度を低下させること、住宅立地推定の目的とならないことから、以下の方法で分析から除外した。メッシュ内における全建物延床面積合計値に対する住宅の延床面積合計値の割合（住宅占有率）を求めた。図-3に示すように、無色のメッシュは住宅占有率が50%未満であり、それ以上は10%刻みで色の濃淡で表現している。50%未満のメッシュでは、大型の駅や商業施設、公園、河川敷などが大部分であることが確認できる。したがって、住宅占有率50%未満のメッシュは、本研究の目的に合わないものと判断し、以後の分析から除外した。

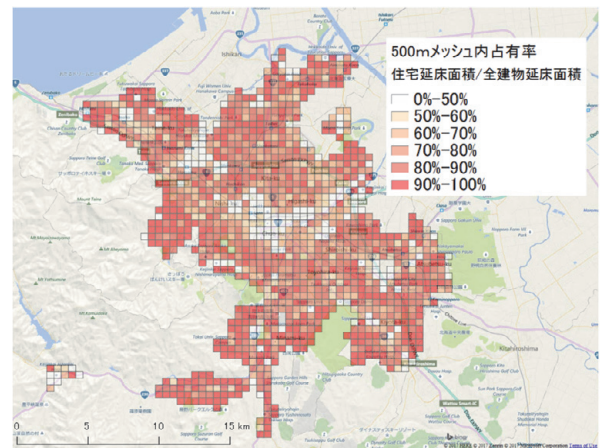


図-3 分析対象エリアとメッシュ別住宅占有率

表-2 重回帰分析に用いる変数

種類	変数名	単相関係数
目的変数	$\log(t+2 \text{ 出現住宅数})$	—
以下、説明変数		
t 年度 出現住宅	t 出現住宅木造住宅率	0.36
	$\log(t \text{ 出現住宅数})$	0.60
$t-1$ 年度 除却住宅	$\log(t-1 \text{ 除却住宅数})$	0.33
	$\log(t-1 \text{ 除却住宅平均建築面積})$	0.31
	$\log(t-1 \text{ 除却住宅平均道路幅員})$	-0.31
t 年度 現存建物	t 現存住宅建物数	0.39
	$\log(t \text{ 現存住宅建物数})$	0.54
	t 現存文教厚生平均実建蔽率	0.32

3.2 データセット

本モデルの説明変数として用意した情報（表-1）の中には、目的変数への影響が小さいものも含まれている可能性がある。そこで、変数選択の手順として、まず、単相関係数の確認を行った。具体的には、目的変数、説明変数全てを対数（自然対数）で変換した値も用意し、単相関係数0.3未満の変数を除外した。次に、多重共線性発生への対応として、説明変数間の相関係数を求め、相関係数0.80以上の組み合わせがある場合は、単相関係数の低い方を除外した。以上の手順で残ったものを説明変数（表-2）として採用し、重回帰分析を行った。

3.3 住宅出現数の要因分析

重回帰分析の結果を表-3に示す。決定係数は、0.43となり、精度は低いため、推定精度の向上には至らなかった。今後、説明変数やモデルそのものについて改善を行うことが課題となる。本モデルにおいては、幾つかの有意な説明変数が得られたことから、今後の精度向上を目的に、住宅出現数の要因、さらに、予測誤差の原因について調べることにした。

表-3に示したように、有意な説明変数としては、 t 出現住宅木造住宅数、 $\log(t \text{ 出現住宅数})$ 、 $\log(t-1 \text{ 除却住宅数})$ 、 t 現存文教厚生平均実建蔽率が得られた。まず、

$\log(t)$ 出現住宅数)の標準偏回帰係数は 0.41 であり、前年度から対象年度に出現した住宅数が多いエリアでは、次年度も住宅の出現傾向が強いことを意味している。つまり、このような傾向のあるメッシュは、住宅立地が盛んであり、その流れから以降も住宅が出現する可能性が高いと言える。 t 出現住宅木造住宅率の標準偏回帰係数は 0.09 であり、特に木造住宅の出現が多いと、次年度に住宅が多く建設されると推測される。一方、 $\log(t-1)$ 除却住宅数)の標準偏回帰係数は 0.14 を示し、古い住宅の除却が多いエリアでは、敷地に余裕ができ、新しい住宅の立地につながるものと読み取れる。以上から、住宅立地の新陳代謝が良好なエリアは、住宅の増加に進む傾向があると言える。また、今回新たな情報として取り入れた各種建物データも変数として選出されており、 t 現存文教厚生系平均実建蔽率の標準偏回帰係数が 0.09 を示している。これは、文教厚生施設の中でも比較的建蔽率の大きい病院などの厚生施設が多いエリアでは、次年度における住宅の出現が増加することを意味している。このことから、厚生施設が存在する地域は、比較的生活水準の高い整備が行われているため、新たに住宅の整備もされやすいということが読み取れる。

以上のように、標準偏回帰係数は「 t 住宅出現数」と「 $t-1$ 住宅消失数」が最も大きいことから、過去の立地履歴はその後の住宅出現に大きな影響を及ぼすことがわかった。また、住宅以外の建物種別の情報も住宅出現に寄与することがわかり、今後のモデル改善につながる結果が得られた。以下では、誤差の大きいエリアの特性について調べ、今後、採用すべき変数や、分析対象のスクリーニング方法について考察を行う。

3.4. 誤差の原因

今回のモデルでは、先行研究で作成した住宅立地判別モデルに、新たにデータや分析範囲の拡大を行うことで、住宅立地判別モデルの「出現」に関する精度の向上を予測し分析を行った。しかし、今回使用したモデルを評価する決定係数は「0.43」と予想より低い信頼性を示す結果となった。住宅出現数の予測値と実測値の関係を図-4に示す。決定係数が低くなった理由については、今回使用した重回帰分析と分析対象範囲の設定に原因があると考えられる。

重回帰分析では、使用するデータが正規分布となっていることが前提となっている。しかし、本研究では目的変数、説明変数で対数変換を行ったが、それでも正規分布にならず、モデルへの適用が的確ではなかったと考えられる。住宅出現に影響を与えている可能性のある変数が他にも存在しており、それらを変数として抽出出来なかった可能性が考えられるが、今回使用した重回帰分析では、変数に制限をかける処理を行うため、そのような変数をすべて抽出することは出来ない。そのため、モデルの精度向上には、ニューラルネットワークやディープラーニングのような、ノンパラメトリックな手法を用いて分析を行う必要がある。例えば、生越ら⁷⁾は、RBF ネットワークを用いて、本研究と同じような建物種別属性値からメッシュ内の滞在人口を推定するモデルを構築している。

表-3 重回帰分析の結果

説明変数	標準偏回帰係数	偏回帰係数	P値	有意性
切片	0.000	0.608	1.000	
t 出現住宅木造住宅率	0.093	0.251	0.014	*
$\log(t)$ 出現住宅数	0.411	0.429	0.000	***
$\log(t-1)$ 除却住宅平均道路幅員	0.051	0.026	0.387	
$\log(t-1)$ 除却住宅平均建築面積	-0.003	-0.001	0.961	
$\log(t-1)$ 除却住宅数	0.141	0.186	0.000	***
t 現存住宅建物数	0.042	0.004	0.201	
$\log(t)$ 現存住宅建物数	0.071	0.024	0.168	
t 現存文教厚生平均実建蔽率	0.091	0.005	0.000	***
決定係数	0.43			

P 値 : *** : 0% ** : 0.001%以下 * : 0.01%以下

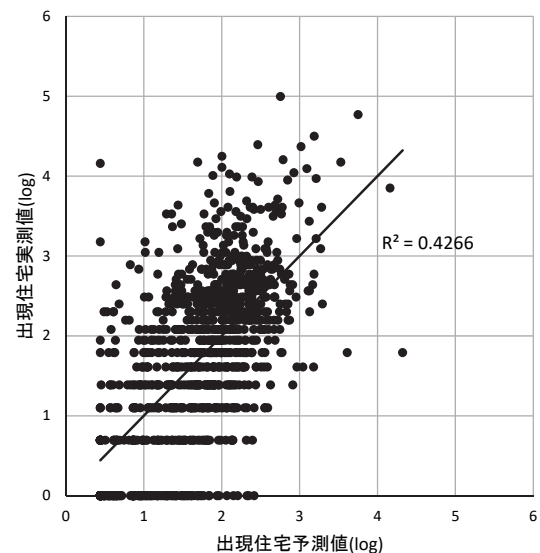


図-4 実測値と予測値

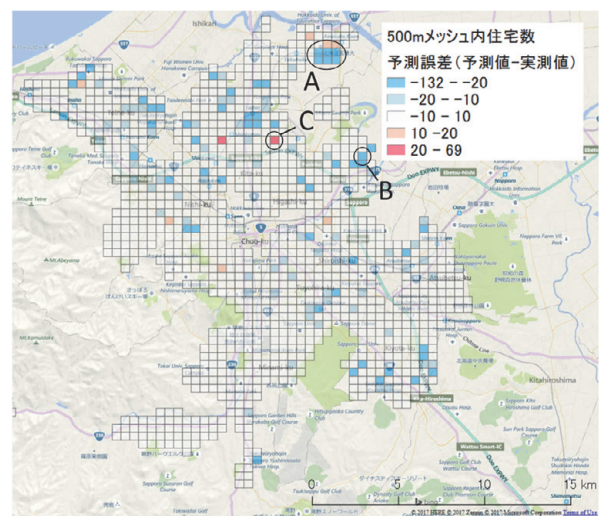


図-5 予測誤差の空間的分布

次に、分析対象範囲の設定について考察する。本研究では、住宅占有率から分析対象メッシュを絞り込んだが、これらの中にはモデルで表現できない外れ値を含むメッシュも存在すると考えられる。そこで、住宅出現数の予測値と実測値の差分（予測誤差）を空間的に可視化し、それが大きいエリアの特性を明らかにする。多くのメッシュでは予測誤差 ± 10 棟となっており、一定の成果が得られている。しかし、郊外部の地域では、予測値が実測値を下回るメッシュ（図中青）が顕著であり、逆に予測値が過剰であるメッシュ（図中赤）も幾つか存在する。例として、図-5 に示した A のエリアでは、予測値が過小となるメッシュが集中している。表-4 に示すように、このエリアでは、前年度除却住宅と現存の文教厚生施設が 0 に等しく、予測値が過小となる傾向にある。また、前年度の住宅出現数及び残存住宅数が比較的大きい。実施にこのエリアを調べると、新しい宅地開発地域であり、毎年、住宅が出現し続けている地域であることが分かった。また、実測値が最も上回ったエリア B においては、除却住宅、公共施設などがあまり見られず、公園や未利用地が比較的多く存在している。このような、公園や川、未利用地などが多く存在する住宅数の少ないメッシュでは、正しい予測が難しいことが分かった。対象に、予測値が最も上回ったエリア C では、比較的公共施設などが多くみられ、残存住宅も多い地域であり、次年度出現する条件に当てはまるが、実際には住宅の出現はほぼ見られなかった。ここは、対象年度より以前に住宅が同時期に出現し、比較的新しい住宅が存在しているため、残存住宅が多くなり、出現住宅が少なくなったと考えられる。

以上の結果より、本モデルは住宅の出現に関して、予測していた精度よりは低いものとなったが、要因として抽出された 7 つの変数、特に t 出現住宅木造住宅数、 $\log(t)$ 出現住宅数、 $\log(t-1)$ 除却住宅数、t 現存文教厚生平均実建蔽率の 4 つの変数は、P 値からも有意性のある変数であることが示されているので、これらの変数は出現する住宅に対して少なからず影響力を持っているということが確認できた。また、モデル自体の信頼性としては低いものとなっているが、抽出された変数に関しては、今後の分析の際、指標として使用することが可能であると考えられる。

4. まとめ

本研究では、札幌市の都市計画基礎調査データのうち建物に関するデータと全建物種別データ（住宅系、商業系、文教厚生系、工業系）を平成 17 年度から平成 27 年度の 6 断面のデータを用いて、先行研究で提案した住宅立地判別モデルの再構築を行い、「出現」に関する精度の向上を図った。

モデル構築に関しては、全ての年度間で分析を行った結果、最も関係性が強く見られた平成 21 年度、23 年度、25 年度の 3 年度を使用した。基準年度を平成 23 年度、予測年度は平成 25 年度とした。初めにモデルの精度低下を防ぐため住宅以外の建物が多いメッシュは除外した。次に、前年度 H21 から H23 にかけて「出現」、「消失」した住宅の各属性値、H23 における建物種別データの 3

表-4 予測誤差が大きいメッシュの特徴

場所	誤差 (予測値 - 実測値)	説明変数			
		出現 木造住宅率	$\log(\text{出現住宅数})$	$\log(\text{除却住宅数})$	文教厚生 平均実建蔽率
A	-28	0.93	4.09	0.00	53.00
	-20	0.80	3.40	0.00	43.50
	-75	0.91	5.08	0.00	0.00
	-60	1.00	1.79	0.00	0.00
	-69	0.96	3.33	0.00	0.00
	-62	0.00	0.00	0.00	0.00
B	-132	0.96	3.83	0.00	0.00
C	69	0.92	4.71	2.30	69.67

青 : 誤差要因(予測値減少)

赤 : 誤差要因(予測値増加)

つの種類のデータを説明変数とし、その中から変数除去を行い、影響度の高い変数を抽出した。この説明変数を用いて H25 における住宅の出現数を判別するモデルを重回帰分析により作成した。本モデルによる予測結果では一定の結果を得られたが、重回帰モデルの信頼性である決定係数は 0.43 となった。

今後は、得られた成果と課題を踏まえつつ、重回帰分析では拾いきることの出来なかった変数を取り扱えるように、ニューラルネットワークやディープラーニングなどノンパラメトリックな分析手法を使用し、モデルの精度向上を行い、考察を行う。

参考文献

- 厚生労働省：平成 28 年人口動態統計の年間推計，
<http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suisei16/index.html>，2017.10 閲覧
- 谷口守：リバース・スプロールを考える：人口減少期を迎えたスプロール市街地が抱える課題，都市住宅学，61 号，pp. 28-33, 2008.
- 植村哲士，宇都正哲，水石仁，榊原渉，安田純子：人口減少時代の住宅・土地利用・社会資本管理の問題とその解決に向けて（上）（中）（下），知的資産創造，8 月号・9 月号・10 月号，pp. 6-23・pp. 62-75・pp. 60-77, 2009.
- 福田健志：空き家問題の現状と対策，国立国会図書館，調査と情報－ISSUE BRIEF－，No. 791, 2013.
- 森本拓磨，浅田拓海，有村幹治：都市計画基礎調査データを用いた居住地分布の逐次の中期予測手法の開発，土木計画学研究・講演集，Vol.54，CD-ROM，2016.
- 浅田拓海，生富直孝，有村幹治：SVM を用いた空間的立地パターン判別による将来居住分布の推定，土木学会論文集 D3（土木計画学），Vol.71，No.5（土木計画学研究・論文集 第 32 巻），I_221-I_228，2015.
- 生越拓実，有村幹治，浅田拓海：RBF ネットワークを用いた時間帯別入込人口推計のダウンスケーリング手法の開発，土木学会論文集 D3（土木計画学），Vol.73，No.5（土木計画学研究・論文集 第 34 巻），2017.（登載決定）