

特集 新しい不動産情報の登場と社会課題解決への応用

自治体データ・民間データを活用した空き家分布推定手法の開発

東京都市大学 建築都市デザイン学部 准教授 秋山 祐樹
 東京大学 空間情報科学研究センター 客員研究員 あきやま ゆうき

1. はじめに

近年、我が国では少子高齢化とそれに伴う人口減少、都市への人口移動、産業構造の変化などを理由に、全国で空き家が増加し続けている。図1に総務省統計局（2019）による最新の住宅・土地統計調査（2018年）に基づく、日本全国の空き家数及び空き家率の推移を示す。この結果によると、最新の調査である2018年の空き家数は全国で約846万戸、空き家率は13.6%に達しているものと推計されている。また、空き家数、空き家率ともに増加の一途を辿っている。さらに今後も空き家

は増加し続けるものと予測されており、例えば今井ほか(2015)による悲観的なシナリオによると、2033年には日本全国の空き家数が2,150万戸と2018年調査の846万戸と比べて約2.5倍に増加すると推定されている。

なお、住宅・土地統計調査では、その管理の状態によって空き家がいくつかの種類に分類されている。すなわち、「①二次的住宅（別荘、その他（たまに宿泊する人がいる住宅）」、「②賃貸用又は売却用の住宅」、そして「③その他の住宅（上記以外の人が住んでいない住宅）」である。これらの

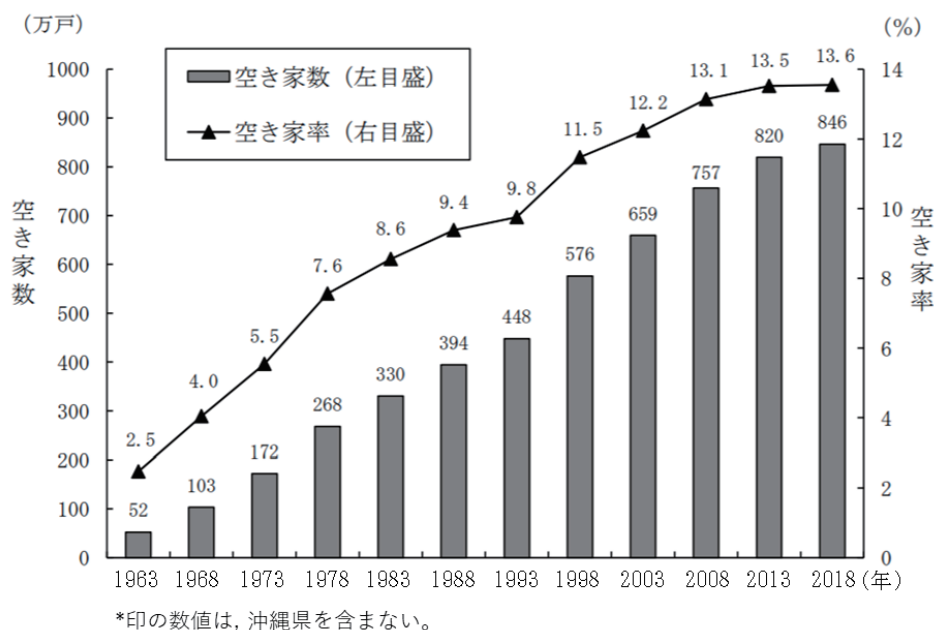


図1 日本全国の空き家数及び空き家率の推移（住宅・土地統計調査、2018年）

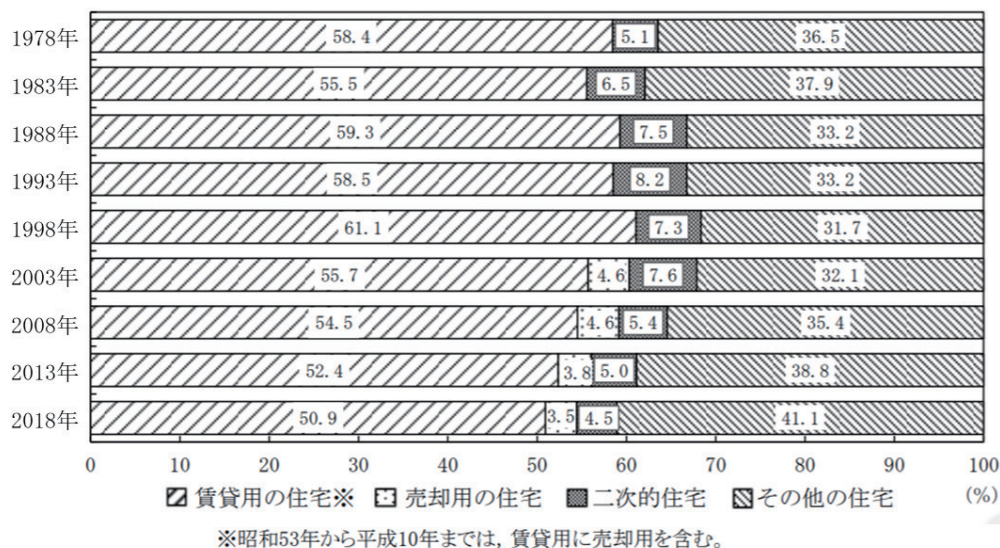


図2 日本全国の空き家の種類別割合の推移（住宅・土地統計調査，2018年）

中でも特に、「③その他の住宅」は、例えば転勤や入院のため居住世帯が長期不在となる住宅や、建替のため取り壊される住宅などが含まれており、他の区分に比べると管理が不十分になりやすいとされている。そして図2は総務省統計局（2019）による住宅・土地統計調査（2018年）に基づく、日本全国の空き家の種類別割合の推移を示したものであるが、特にその他の住宅の割合が増加し続けていることが分かる。さらに、その他の空き家は今後も増加することが予測されており、例えば秋山・柴崎（2015）は、2013年から2040年にかけて、その他の住宅が日本全国で318万戸から417万戸に増加すると推計している。

空き家の増加、特に管理が不十分な空き家が増加することにより、どのような問題がおこるであろうか。図3は近畿圏の市町村へのアンケート調査に基づく、空き家における問題事象の発生状況をまとめたものである（鈴木・伊藤，2012）。図3が示すように、空き家の老朽化による倒壊事故や強風等による落下・飛散事故の発生、放火や不審者侵入などによる防犯性の低下、景観への悪影響、更には地域の活性化への障害になるなど、外部不経済の状態を発生させていると言える。同様の影響については既存研究でも指摘されているところである（浅見，2014；Baba and Asami，2017）。ま

た、空き家の中でも特に「その他の住宅」は、管理が不十分なことによる倒壊危険性が特に高いことや、害獣・害虫の繁殖による衛生上の問題を有すること、さらに積雪寒冷地の場合は近隣への雪落の危険など、近隣住民や地域全体に深刻な影響を及ぼす可能性が高いとされている（浅見，2014；日本弁護士連合会法律サービス展開本部自治体等連携センター，2018）。

以上のように空き家の適正な管理の必要性が高まりつつある中、2015年5月より「空家等対策の推進に関する特別措置法」（通称：空家対策特別措置法）が施行された。同法が施行される以前は、自治体が独自の空き家対策に関する条例を作るなどして、空き家の活用や除却などの対策が行われていたが、法的効力を有しなかったため、空き家に関する最終的な判断は空家の所有者に委ねられていた。そのため空き家への対策が充分に進まなかったという問題があった。一方、同法が施行されたことにより、管理が不十分だと判断される空き家に対して、自治体の調査に基づき問題があると判断される空き家を「特定空き家」に指定し、所有者への指導や改善の促進を行うことが可能となった（宮路ほか，2015）。

しかし、自治体がかような行動を迅速かつ効果的に行うためには、まず自らの自治体内における

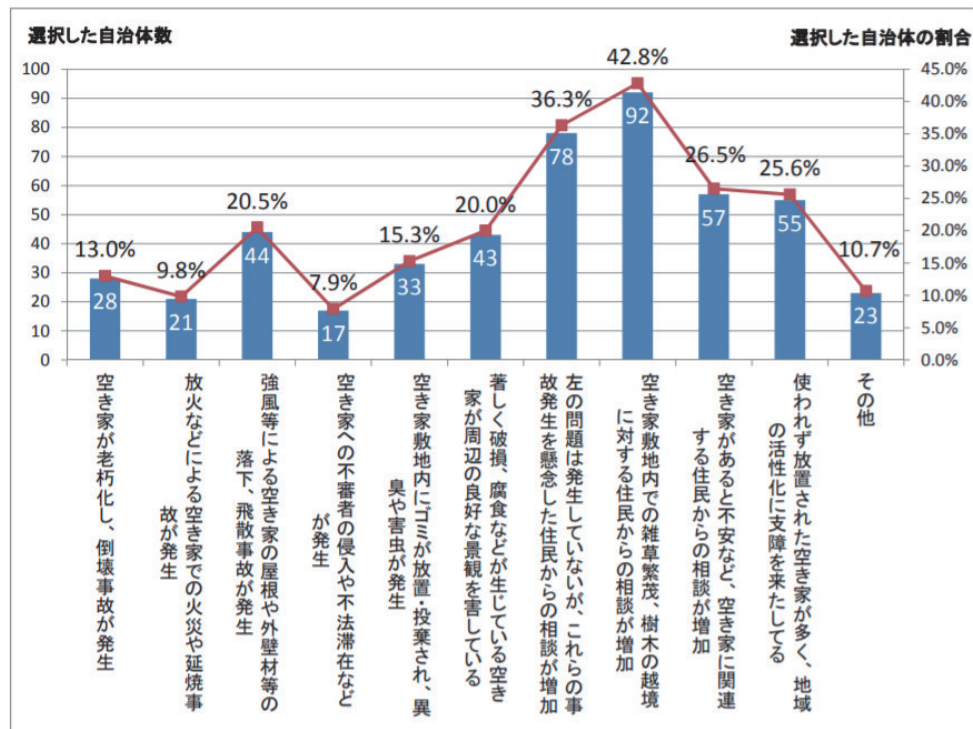


図3 空き家における問題事象の発生状況（鈴木・伊藤，2012）

空き家の空間分布の状況を把握する必要がある。これまでに空き家の空間分布を把握するための研究は数多く見られるが（益田・秋山，2020）、その多くは現地調査による個別目視（加藤ほか，2009；久保・益田，2015 など）や、地域住民などを対象とした聞き取り調査（佐々木ほか，2010；篠部・占部，2014 など）が中心である。また、自治体を実施する空き家分布調査の手法も、多くの場合は一棟一棟の個別訪問による外観目視という手法が中心となっているため、自治体による今日の空き家の分布調査は、多くの時間、費用、労力を要している。これが自治体において広域を対象とした空き家対策を進めていく上で大きな負担となっている。

2. 著者らによるこれまでの取り組み

以上の課題に対して、著者らはこれまでに自治体と民間企業が保有する様々なデータを活用して、自治体における空き家の分布調査を、可能な限り迅速かつ安価に実施できる手法の開発を行ってきた。著者らによる手法の特徴は、空き家の空間分

布を迅速かつ安価に調査するために、既存の手法に見られる対象地域全域を現地調査する手法は採用せず、調査対象地域の一部においてサンプル的な現地調査を実施し、その結果を空き家の教師データとするとともに、自治体と民間企業が保有する様々なデータを組み合わせることで、自治体全域の空き家の空間分布を推定するという点にある。また、ここでいう自治体が保有するデータ（以下、「自治体データ」とは、自治体外部の者でも研究・分析等の用途で自由に二次利用が可能なデータ、すなわち「オープンデータ」だけではなく、一般的には外部公開されていない、自治体内で利活用されるクローズドなデータ、例えば住民基本台帳のようなデータも含む。これらのデータを利用する理由は、自治体が既に保有しており、また継続的に更新・管理が行われているデータを可能な限り活用することで、自治体にとっての空き家分布調査のコストとハードルをできるだけ縮減するためである。

また本研究で対象とする空き家は、戸建て住宅の空き家とした。これは戸建て住宅と集合住宅と

の間には、空き家の分布を把握するための手法において、様々な点で相違があると考えられるためである。さらに前述した空き家による外部不経済の解消や、特定空き家となりうる空き家の分布把握という需要を考慮すると、戸建て住宅を対象とする方が合理的であると考えられるためである。

以下、3節では Akiyama et al. (2020) による鹿児島県鹿児島市および福岡県朝倉市を対象とした、自治体データと民間データを活用した空き家分布推定の研究を紹介する。なお3節は秋山ほか(2019)、秋山(2019) および、Akiyama et al. (2020) を本稿向けに再編集および再校正した内容である。

3. 自治体データ・民間データを活用した空き家分布推定の事例

本節では Akiyama et al. (2020) による鹿児島県鹿児島市および福岡県朝倉市を対象とした、自治体データと民間データを活用した空き家分布推定の研究を紹介する。

3.1 対象地域

対象地域は鹿児島県鹿児島市(以下、鹿児島市)

と福岡県朝倉市(以下、朝倉市)とした。これらの自治体では、比較的広域を対象に現地調査を実施すること、また自治体データの提供について、自治体より協力を得ることができた。また、鹿児島市は高齢化率、平均世帯人員、平均世帯収入、空き家率といった指標が、何れも全国平均値と近い値であることから、全国の自治体の中でも代表性が高い自治体と言える。さらに、何れの自治体も住宅地のほか、住宅と商業地域・工業地域が混在する地域も含まれ、加えて平坦地と傾斜地が共に含まれる地域である。このように多様な地域特性を持つ地域を対象とする理由は秋山ほか(2018)が指摘するように、地域の地理的特性が空き家の多寡に影響を及ぼしている可能性を考慮するためである。以上の理由より、鹿児島市と朝倉市を本研究の対象地域として選定した。

3.2 現地調査の実施

自治体と民間の各種データを用いて、空き家の空間的分布を把握・推定するためには、空き家と非空き家の特徴を、これらのデータに基づいて把握できるようにする必要がある。そのためには一定数の実際に空き家である建物の情報が必要とな

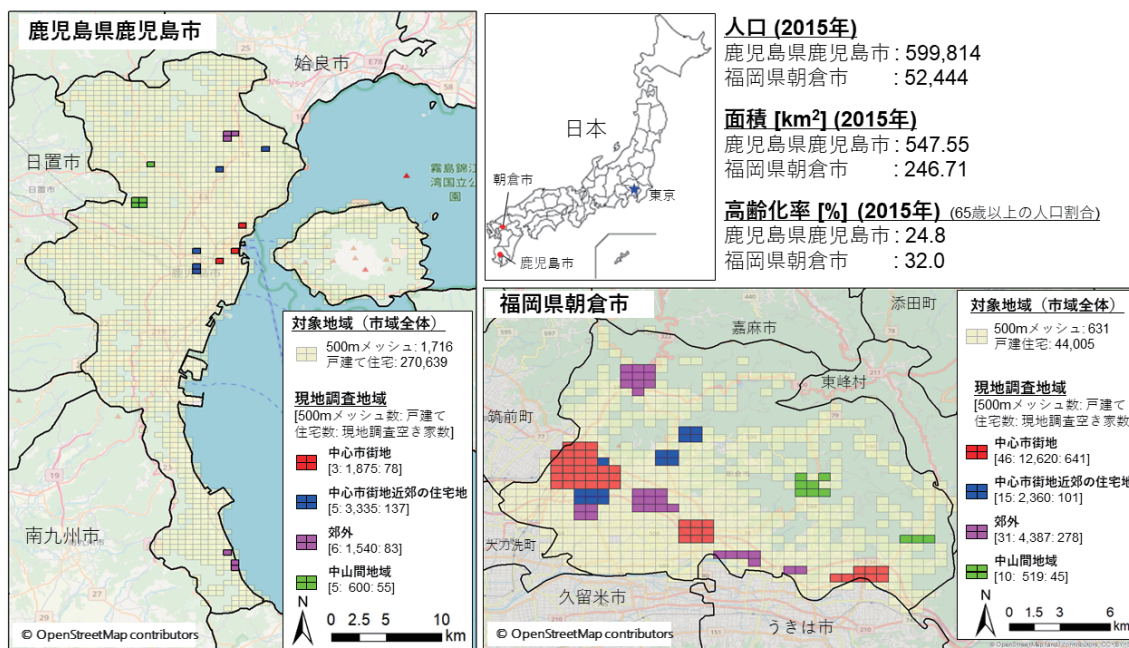


図4 研究対象地域および現地調査地域

現地調査の様子



空き家と判定された建物の例



図5 現地調査の様子と空き家と判定された建物の例（いずれも鹿児島市）

表1 現地調査実施日、調査対象建物数および現地調査で得られた空き家数

都市名	現地調査実施日	調査対象建物数	現地調査 空き家数	空き家率 [%]
鹿児島市	2016/08/22～26	7,350	353	4.80
朝倉市	2016/10/17～21	19,886	1,065	5.36

る。そこで研究対象地域である鹿児島市および朝倉市において、一定数の実際に空き家である建物の情報を収集するためのサンプル地域を設定し、これらの地区を対象に戸別目視による現地調査を行うことで、実際に空き家である建物に関する情報を収集した。図4に研究対象地域および、現地調査地域を示す。研究対象地域は市域全体を500mメッシュで分割し、メッシュ内に戸建て住宅が1棟以上存在するメッシュ全てを対象とした。また現地調査地域は、自治体へのヒアリングの結果、現地調査の実施が許可された地域であることに加えて、中心市街地、中心市街地周辺の住宅地、郊外、中山間地域がバランス良く含まれるように設定した。図5に現地調査の様子と空き家と判定された建物の例を、表1に現地調査実施日、調査対象建物数および現地調査で得られた空き家数を示す。

3.3 空き家分布推定のための空き家データベースの開発

現地調査で収集した空き家（本研究では「空き家真値」と呼ぶ）の空間分布と、自治体データから得られる建物ごとの様々な情報を、GISを用いてデジタル住宅地図（株式会社ゼンリン）の建物ポリゴンデータに結合することで、空き家の特性を分析するためのデータベースである「空き家データベース」を開発した。表2に空き家データベースの属性を示す。「① 住宅地図」からは建物の用途や面積を取得した。なお住宅地図は民間データであるが、多くの自治体で広く導入されているデータであり、自治体にとって入手する上での障壁が少ないものと考えられるため利用した。「② 現地調査」は表1で取得した現地調査結果が該当する。「③ 住民基本台帳」、「④ 水道使用量情報」、「⑤ 建物登記情報」が、本研究において自治体か

ら取得した自治体保有のデータである。「③ 住民基本台帳」は自治体に居住する全ての住民に関する情報である。「④ 水道使用量情報」は自治体の水道局が管理する水道栓ごとの水使用量のデータベースであり、閉栓状態の水道栓の分布も把握できる。そして「⑤ 建物登記情報」は自治体及び法務局が管理する、個々の土地・建物に関する情報である。本研究ではこれらのうち建物に関する情報を利用した。最後に「⑥ 国土数値情報」は、国土交通省が公開している各種 GIS データでありオープンデータである。本研究では国土数値情報のうち、全国の自治体が作成した都市計画総括図を基に作成された用途地域データを利用した。

以上の②～⑥のデータを①住宅地図に結合するためには、定量的な位置情報、すなわち座標情報（経度、緯度）が必要になる。②は現地調査により作成した座標情報付きの情報であり、また⑥は GIS データ（シェープファイル）であるため、GIS により①と直ちに空間結合させることが出来る。一方、③～⑤には座標情報は含まれていないが住所情報を持つため、アドレスマッチングにより座標情報に置換することが出来る。その結果、③～⑤も①と空間結合することが可能になる。

以上のうち、③④⑤は自治体データであり、自治体内で行う業務においては利用可能であるが、自治体の外の者が利用することは原則できないデータである。本研究の場合、調査対象地域となる鹿児島市、朝倉市の個人情報保護審査会を通して研究利用の承認を得るとともに、住民の名前といった個人情報を著者らが直接知ることができないように秘匿処理を施されたデータを用いた。

なお、全ての住宅地図の建物に自治体データが紐付くとは限らない点に注意が必要である。この原因としては、まず住所表記によってはその表記が不完全、あるいはアドレスマッチング側の住所データの表記揺れなどの問題で、アドレスマッチングの精度が低くなってしまったことが挙げられる。また④水道使用量情報の場合、特に中山間地域の住宅を中心に依然として井戸水が使用されていることにより、④の情報が紐付かなかった建物

も存在した。さらに、現地調査により空き家と判定された住宅における自治体データの結合率は、非空き家に対する結合率よりも低かった。すなわちこの結果は自治体データが空き家か否かの判定を行う上で有用であることを示唆している。

3.4 空き家分布推定のためのクロス集計表の開発

本研究では現地調査地域の全ての建物に与えられた表 2 に示す属性を説明変数とした複数のクロス集計を行い、組み合わせごとの空き家率を算出した。そして複数のクロス集計表から得られた値を平均した結果を対象地域全域、すなわち市域全体の建物に割り当てることにより、対象地域全体の建物毎の空き家率を推定した。本研究ではこのクロス集計表の各組み合わせのことを「セル」と呼ぶ。

クロス集計表を作成するためには、定量的変数に適切な閾値を与えて各変数をセル分割する必要がある。そこで本研究では、まず鹿児島市と朝倉市の現地調査地域における空き家データベースの各定量的変数の分布をグラフ化して把握することで、サンプル数が非常に少ない、あるいは空き家率が類似したセル同士を結合することで、閾値を決定した。以上の方法で設定した閾値に基づいて、表 2 に示す全ての変数を用いたクロス集計表を作成した。その結果、このクロス集計表による現地調査地域における空き家数の推定結果は、5 次メッシュ（250m 四方）ごとの空き家真値と推定値による相関分析で相関係数が 0.8956 となり、空き家真値と推定値の間に非常に強い相関が見られることが明らかになった。

なお、以上のような閾値設定におけるより定量的な手法としては、カイ二乗適合度検定や赤池情報量規準(AIC)を用いた方法などが挙げられる(北川ほか, 1983)。ただし本研究は自治体においてモデルをできるだけ簡便に構築し、現場において迅速に展開できるという実用性を重視したため、敢えて閾値を恣意的に決定する方法を採用した。

続いて、現地調査地域の一部のデータで現地調査地域全体をどの程度の信頼性で推定出来るか、

表2 空き家データベースの属性一覧

データソース	データ 作成時期 ^a	データ件数 ^a	属性	データ タイプ ^b
① 住宅地図	K: 2016 年 A: 2016 年	K: 285,917 A: 44,692	建物用途	C
			建物面積	V
② 現地調査	表 1 参照	表 1 参照	空き家判定	C
③ 住民基本台帳	K: 2016 年 6 月 A: 2016 年 9 月	K: 292,285 A: 120,326	①の結合判定 (結合、非結合)	C
			世帯数	V
			居住者数	V
			居住者の平均年齢	V
			居住者の最若年齢	V
			居住者の最高年齢	V
			居住者の最短居住年数	V
			居住者の最長居住年数	V
④ 水道使用量情報	K: 2016 年 6 月 A: 2016 年 10 月	K: 270,218 A: 167,530	①との結合判定 (結合+開栓、結合+閉 栓、非結合)	C
			閉栓年数 (閉栓の場合)	V
			年間水道使用量 (開栓の 場合)	V
⑤ 建物登記情報	K: 2016 年 6 月 A: 2016 年 6 月	K: 253,511 A: 33,804	①の結合判定 (結合、非結合)	C
			築年数	V
			建物構造	C
			建物用途	C
⑥ 国土数値情報	K: 2011 年 A: 2011 年	—	用途地域	C

^a ”K”は鹿児島市、”A”は朝倉市を意味する。

^b ”C”は定性的情報、”V”は定量的情報を意味する。

ということを検証した。これは本研究では対象地域全域（市域全体）の空き家分布推定を、対象地域全体から見れば一部地域となる現地調査地域に基づいて推定（外挿）を行うためである。現地調査地域から 50%をランダムに抽出してクロス集計表を作成し、両市において同クロス集計表を用いた推定結果と真値を比較することで、同手法の

信頼性を検証した。表 3 は 5 次メッシュ（250m 四方）ごとの真値と推定値の相関分析の結果である。現地調査地域全体の相関係数は 0.8598 であり、現地調査結果を 100%使用して作成したクロス集計表による推定結果の相関係数である 0.8956 とほぼ変わらない精度となった。また鹿児島市のみ、あるいは朝倉市のみ結果もほぼ同程度の相関係

数となっており、平均絶対誤差（MAE）も低い水準となった。

さらに同様のランダムサンプリングを 1000 回繰り返し、全てのサンプルを用いて 1000 パターンのクロス集計表を作成した。そして両市においてこれらのクロス集計表を用いた推定結果と真値を先程と同じく 5 次メッシュごとの真値と推定値の相関分析により比較することで、信頼性の変動を検証した。図 6 に 1000 回全ての相関係数を、図 7 に MAE を示す。相関係数は何れの地域も 1000 回の平均が約 0.84 であった。またサンプル間で相関係数に大幅な違いは見られなかった。一方、MAE は鹿児島市の平均値が 2.09、朝倉市が 0.57、全体が

0.81 であった。またサンプル間で MAE に大幅な違いは見られなかった。ただし鹿児島市の場合、中心市街地や大規模な住宅団地など戸建て住宅の数が他のメッシュと比較して非常に多くなるメッシュが存在するため、メッシュごとの建物数を用いた外れ値検定を行い、平均値±2 標準偏差以上のメッシュを外れ値として外した上で、再度相関分析を行ったところ、1000 回の MAE の平均値は 0.94 となった。これらの MAE の結果が示すことは、鹿児島市、朝倉市何れの市においても 5 次メッシュごとに得られる推定空き家数は真値と比較して平均してその誤差が 1 棟未満となるということである。ただし鹿児島市の相関係数と MAE は他の結果と比

表3 50%をランダムに抽出した結果を用いて作成したクロス集計表による信頼性検証結果

	対象とする現地調査地域		
	全域 (鹿児島市+朝倉市)	鹿児島市	朝倉市
相関係数	0.8598	0.8520	0.8583
補正決定係数	0.7377	0.7259	0.7332
MAE	0.6433	1.7876	0.4310
p 値	4.25e-135	2.23e-26	1.99e-120

P-values of all correlation analyses satisfy < .0001.

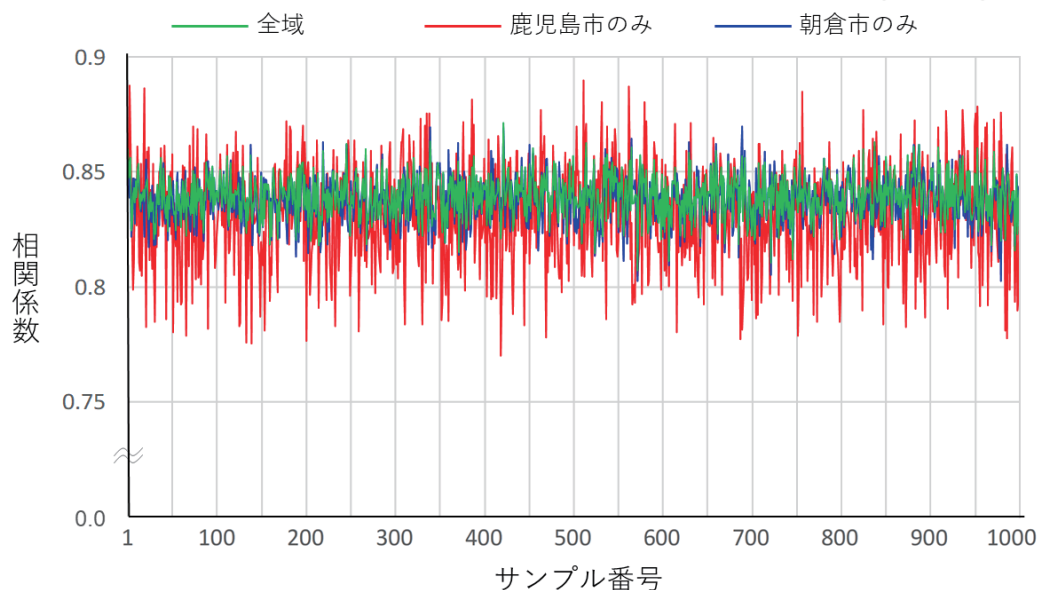


図6 現地調査地域の全域（鹿児島市+朝倉市）、鹿児島市、朝倉市それぞれにおける1000回のランダムサンプリングによるクロス集計表で得られた相関係数（5次メッシュ集計）

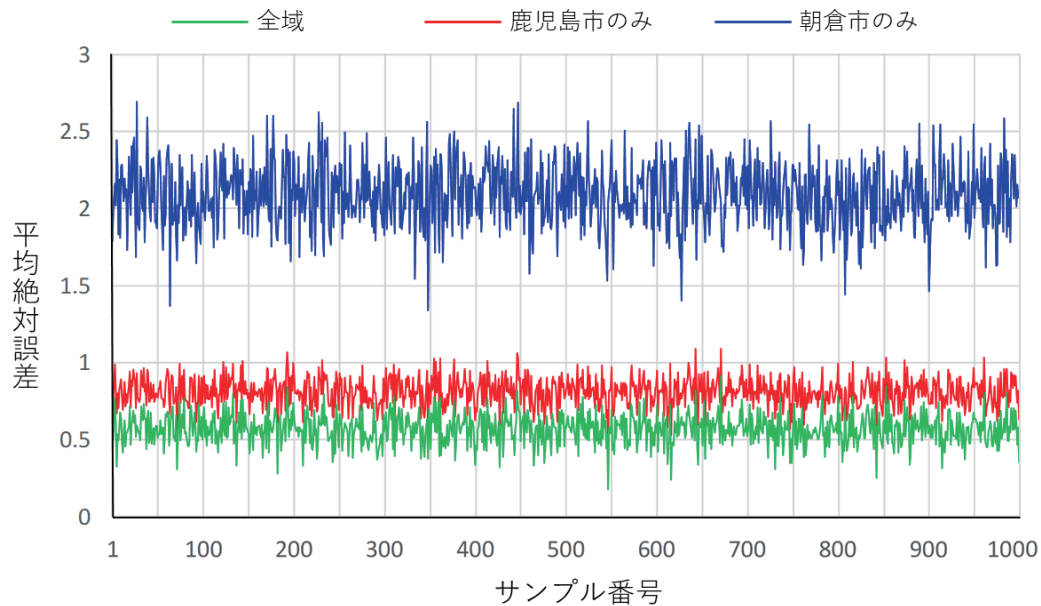


図7 現地調査地域の全域（鹿児島市+朝倉市）、鹿児島市、朝倉市それぞれにおける1000回のランダムサンプリングによるクロス集計表で得られた平均絶対誤差（MAE）（5次メッシュ集計）

べるとばらつきが大きくなった。これは鹿児島市が比較的大規模な地方都市の中心市街地から中山間地域まで多様性に富む地域を含む一方で、朝倉市は地方の小規模都市であり鹿児島市に比べると地域特性の多様性が小さいためであると考えられる。この課題は今後、さらに他の都市の様々な地域特性の現地調査データを入手し、空き家データベースを充実させていくことで改善できるものと期待される。

3.5 複数のクロス集計表を用いた任意領域における空き家数および空き家率の推定

最後に3.4で構築した1000種類のクロス集計表を用いて、任意の領域（例えば町丁目単位や地域メッシュ単位）の空き家数および空き家率を推定する。まず1000種類のクロス集計表を用いて建物ごとに1000種類の空き家率を算出し、それらの平均値を最終的な空き家率とする（図8）。これは過学習による汎化性能の低下、すなわち局所的な精度の低下を回避するための措置である。そして最後にある任意の領域Aにおける空き家数および空き家率を推定する。ある領域Aの推定空き家数 V_A

は式1で推定することが可能となる。

$$V_A = \sum_{k=1}^n r_k \quad (1)$$

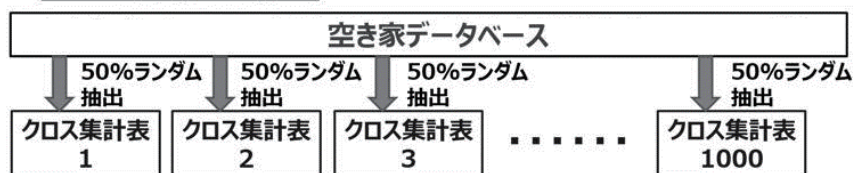
n は領域A内の戸建住宅数、 r_k は戸建住宅 k の推定空き家率である。また領域Aにおける推定空き家率 R_A は式2で推定することが可能となる。

$$R_A = \frac{V_A}{n} \quad (2)$$

3.6 空き家分布推定の結果

図9および図10に本研究の手法で推定した鹿児島市全域および朝倉市全域の4次メッシュ（500mメッシュ）ごとの推定空き家数と推定空き家率を示す。まず、鹿児島市では中心市街地（図9中A）周辺の住宅街や、吉野町（図9中B周辺）、喜入中名町（図9中C周辺）、桜島地区の錦江湾沿岸地域（図9中E周辺）などで特に推定空き家数が多くなっていることが分かった。中心市街地は建物数が高いため、居住地としての快適性が近郊と比較

◆クロス集計表の作成



◆建物ごとの空き家率の推定

- 空き家率 (クロス集計表1による): R_{A1}
- 空き家率 (クロス集計表2による): R_{A2}
- ⋮
- 空き家率 (クロス集計表1000による): R_{A1000}



建物Aの推定空き家率

$$R_A = \Sigma R_{An} / 1000 \text{ (} R_{An} \text{の平均値)}$$

図8 1000種類のクロス集計表を用いた建物ごとの空き家率の推定方法

して相対的に劣ることなどから空き家数が多くなっているものと考えられる。また吉野町は鹿児島市で最も人口が多い町丁目であり、該当するメッシュ周辺は台地上に拓かれた住宅団地であるため、中心市街地と同様に建物数が多い地域となってい

る。一方、喜入中名町や桜島地区では高齢化と人口減少が進んでいることから、空き家が増加し続けているものと考えられる。また空き家率は市北部や西部の中山間地域（図9中F周辺）で特に高い値となっていることが分かった。さらに喜入生

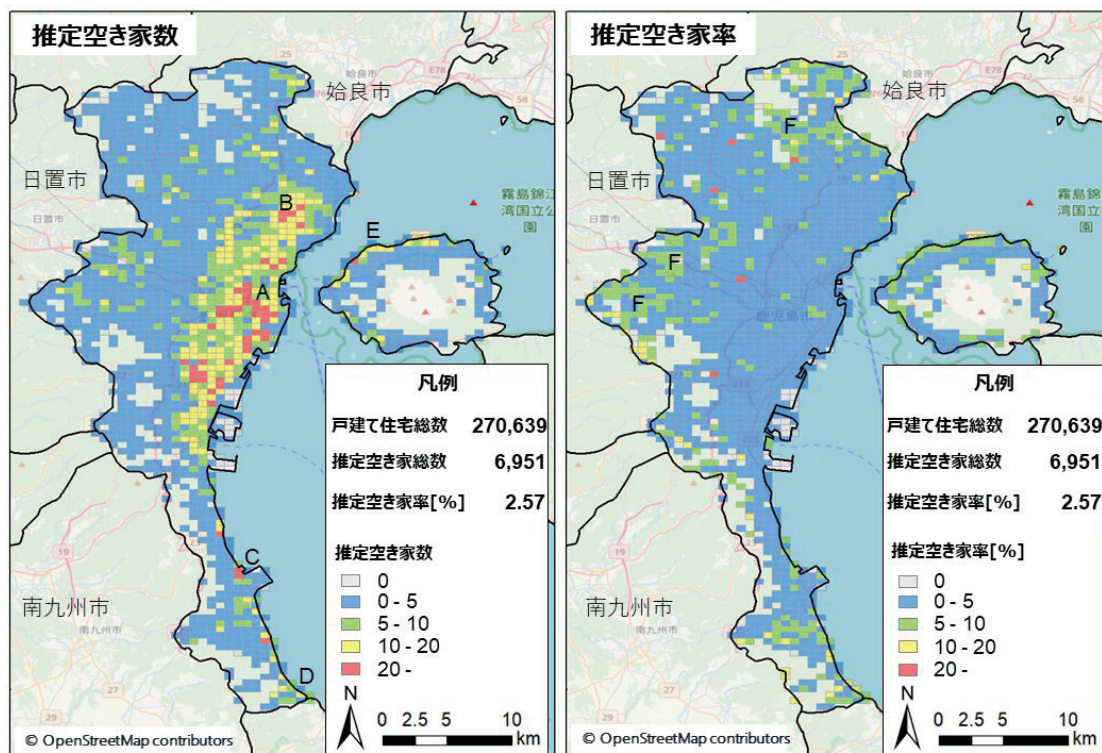


図9 鹿児島市全域における推定空き家数と推定空き家率（4次メッシュ集計）

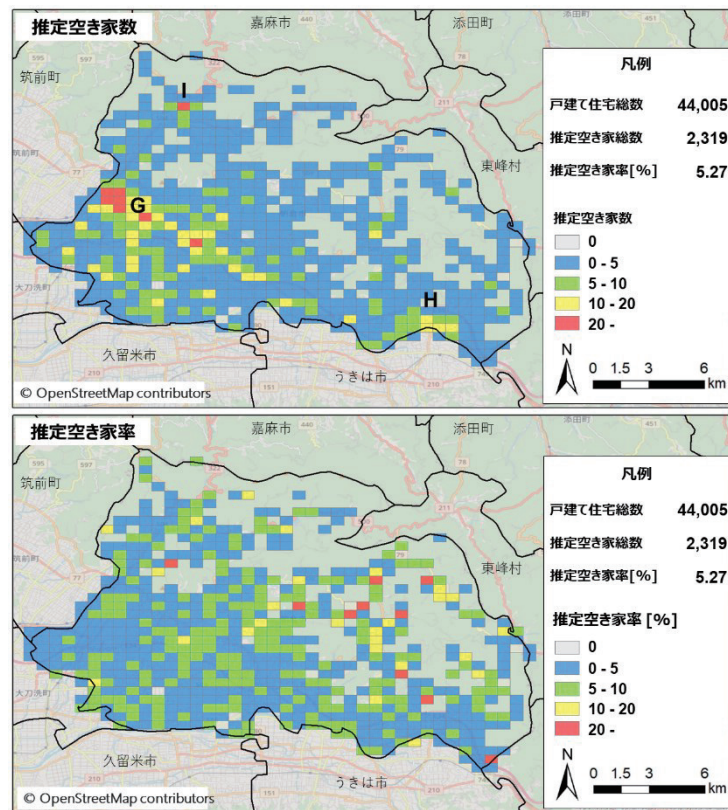


図10 朝倉市全域における推定空き家数と推定空き家率（4次メッシュ集計）

見町（図9中D周辺）では空き家数、空き家率ともに高い値となった。喜入生見町は鹿児島市南端に位置する農漁村的な性質が強い地域であり、高齢化が進みつつある地域でもあるため、このような結果となったものと考えられる。

一方、朝倉市は2006年に複数の市町村が合併して新設された市であり、合併前の旧中心市街地（図10中G周辺、H周辺）で空き家数が多くなっていることが分かった。また、旧市街地である城下町（図10中I周辺）でも空き家数が多いことが分かった。一方、空き家率は市南部の平野部の値は概ね一様であり、鹿児島市ほど地域間が顕著ではなかった。また空き家率が特に高かった地域は、市北部の中山間地域であった。

3.7 小括

本節では鹿児島市全域と朝倉市全域を対象に、自治体と民間企業が保有するデータを活用して、自治体における空き家の分布調査を可能な限り迅

速かつ安価に実施するための手法開発に関する研究について紹介した。本研究ではまず、鹿児島市と朝倉市の一部地域を対象に現地調査を行い、現地調査結果と複数の民間および自治体のデータを組み合わせることで空き家データベースを構築した。続いて空き家データベースから複数種類のクロス集計表を作成し、それらを用いて過学習を回避しながら各建物の空き家率を推定した。さらにその結果を町丁目やメッシュなど任意の単位で集計化することで、鹿児島市および朝倉市全域を対象に空き家数や空き家率を推定することが可能になった。

本研究の課題は、まずこの方法をより多くの自治体に適用しながら、その可能性や課題、また限界を解明していく点が挙げられる。また、空き家に関する現地調査データを、数多くの自治体を対象に収集したり現地調査を実施したりすることで、空き家データベースを強化し、様々な地域特性に対応できるデータベースを実現していくことも重要である。さらに、本研究で活用した民間や自治

体データ以外にも、まだ著者らが把握しきれていない様々な種類のデータが存在するものと期待される。そこで、今後は空き家分布推定に有用なデータの調査も進めていく必要があると考えている。

4. 他の自治体における空き家分布推定の事例

3 節では鹿児島市および朝倉市における空き家分布推定の事例を紹介したが、著者らはすでに他の自治体においても民間および自治体のデータを活用した空き家分布推定の取り組みを進めている。本節では和歌山県和歌山市（以下、和歌山市）と群馬県前橋市（以下、前橋市）における取り組みの事例を紹介する。

4.1 和歌山市における事例

著者らは総務省統計局統計データ活用センター、和歌山県、和歌山市と連携し、和歌山市が保有する自治体データ（住民基本台帳、水道使用量情報、建物登記情報、空き家調査データ）を活用することで、和歌山市全域の空き家分布状況を迅速、安価に推定するモデルを構築するとともに、同モデルの信頼性の検証を行う研究を実施している（秋山ほか、2020a）。同研究では前述の鹿児島市の手法を更に発展させ、建物ごとの空き家率の推定を複数のクロス集計表を用いた手法ではなく、機械学習（XGBoost）を適用することで、クロス集計表の閾値を定性的に決定せざるを得なかったと

いう課題を解決し、精度向上を図っている。

その結果、このモデルを用いることで和歌山市による全域調査のうち、95.4%の建物を空き家、あるいは非空き家と正しく判定することができた。図 11 は和歌山市が実施した現地調査による空き家分布調査の結果と、本研究のモデルで推定した空き家分布推定結果の比較である。空き家の数量、分布傾向ともに真値と類似した結果が得られることが分かった。

今後はモデルに投入する現地調査データをどの程度少なくしても十分な精度を担保できるのか、ということをはっきりとしたり、3 章の鹿児島市の例のように、モデルに投入する現地調査地域が限定的になる場合に、他の地域への外挿がどの程度可能なのか、といった点を明らかにしたりしていきたいと考えている。また、利用する自治体データを敢えて限定したり、あるいは様々な公的統計（国勢調査、住宅・土地統計調査など）を利用して特徴量を充実させたりすることで、和歌山市以外の自治体への拡張性を検討していきたい。

4.2 前橋市における事例

著者らは群馬県前橋市と研究協議会（超スマート自治体研究協議会）を立ち上げ、官民のビッグデータを活用した EBPM（エビデンスベースドポリシーメイキング）の推進に関する研究を行っている。その活動の一環として、民間と自治体のデー

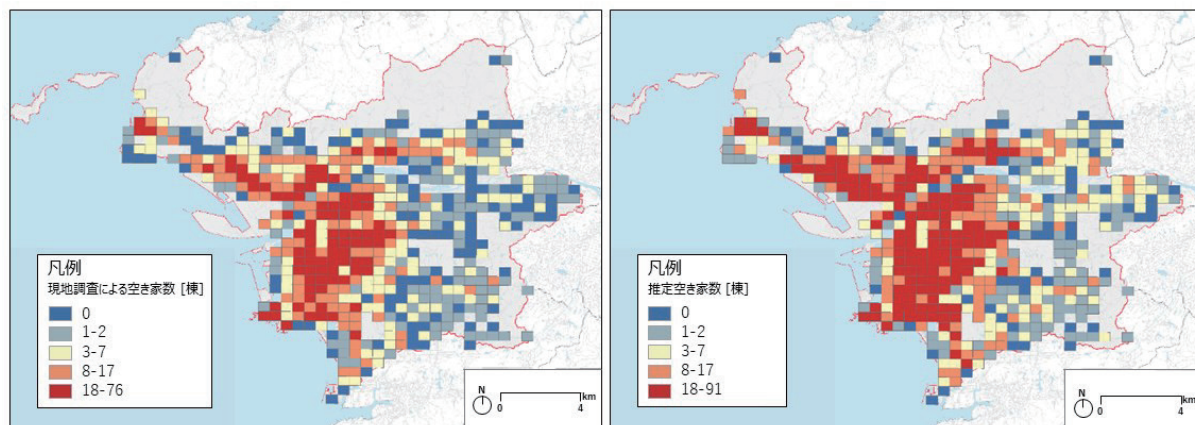


図11 和歌山市全域における空き家棟数の真値（市による現地調査結果）（左）と自治体データを活用した推定結果の比較（右）（4次メッシュ集計）

タを活用した空き家分布推定も実施している。特に前橋市では空き家分布の現状を把握するだけでなく、過去のデータの時系列変化の傾向から、空き家の将来分布を推定する研究にも取り組んでいることが特徴的な点と言える（秋山ほか，2020b；馬場ほか，2019）。なお前橋市の同研究では自治体データとして住民基本台帳、水道使用量データ、固定資産税台帳（家屋）を利用した。

同研究では、ある t 年の自治体データから得られる空き家の特徴量から α 年後の空き家分布を推定するモデルを構築し、 $t + \alpha$ 年の空き家調査結果との整合性を検証する。そしてこのモデルが十分な精度を有しており、また時間が経過しても空き家の発生、分布、存続などのメカニズムの傾向が大きく変わらないと仮定すれば、ある t 年から任意の β 年後、すなわち $t + \beta$ 年から α 年後の空き家分布

の推定も可能になると言える。

同研究では2014年（ $t = 2014$ ）の自治体データを用いて1年後（ $\alpha = 1$ ）の空き家分布の状況を予測し、実際に前橋市が実施した空き家分布の現地調査結果と比較して、その信頼性を検証した。モデルは非線形の機械学習モデル（XGBoost）を採用した。同研究では各建物に空き家確率が0～100%の範囲で付与されるため（図12）、空き家確率が50%以上の建物を空き家と仮定することで、正答率は約90%に達した。またこのモデルを2018年（ $\beta = 4$ ）の自治体データに適用し、2019年の現地調査結果と比較したところ、2015年の推定と同様に空き家確率が50%以上の建物を空き家と仮定した場合、正答率は72.5%となり、2015年の結果ほどでは無いが、それでもかなり高い信頼性を有するモデルが実現しつつある。また同モデルか

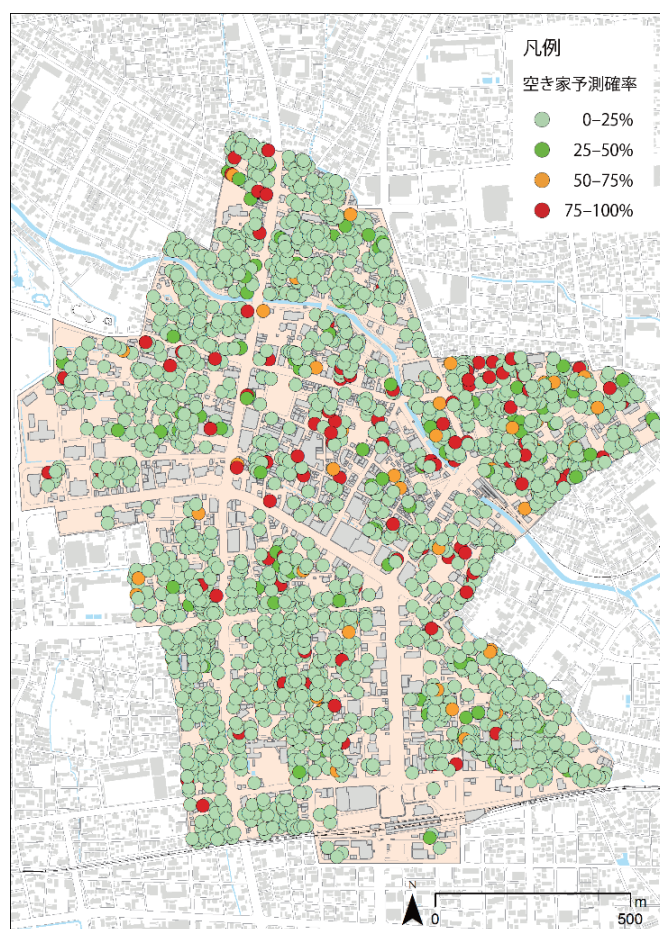


図12 前橋市中心市街地における戸建て住宅ごとの空き家確率の推定結果

ら得られる空き家の特徴量の重要度によると、建築年、建築面積、住宅種別などが重要な要素として挙げられており、これらは既往研究での知見に沿うものであることから、本手法の妥当性を支持するものと言えよう。

今後の課題としては、まず現状では前橋市の中心市街地のみを対象としているため、対象地域を拡張して十分な量のトレーニングデータを確保することで、モデルの信頼性を高めることを検討している。これは今後、前橋市の中心市街地以外にこのモデルを適用する上で不可欠な取り組みである。また、本研究ではいずれの自治体データも付与されない建物が若干数存在したため、自治体データが付与されていない建物が分析の対象外となる問題が明らかとなった。これは他の自治体データや民間データ、あるいは4.1でも触れた様々な公的統計などを組み合わせることで、解決を図っていきたいと考えている。

5. おわりに

本章では自治体データ・民間データを活用した空き家分布推定手法の開発に関する取り組みを、著者らが実施している研究を事例に紹介した。民間のデータだけでなく、自治体が保有するデータ、特にオープンデータではなく、自治体において通常は内部利用のみ許されている「クローズド」なデータを活用することで、空き家分布推定が高い信頼性を持って実現しつつある。

本章で紹介した研究成果は、自治体の空き家分布調査において非常に有用性が高いものになると考えている。まず、本章で紹介した成果は既存研究における課題、特に広域を対象とした空き家の空間分布を迅速に調査するという課題の解決に大いに貢献するものである。また、本章の一連の研究は自治体データを主に活用して実施している。自治体データは日本全国の自治体が通常業務として継続的に整備しているものであるため、本章で紹介した手法を自治体において活用したい際に、新たにデータを整備するという負担を強いることがない。さらに、本章で紹介した一連の研究で活

用している自治体データは、全国の自治体においてもほぼ同様のデータが整備されているため、全国のあらゆる自治体において汎用的に使える手法となるものと期待される。加えて、多くの自治体データはある特定の目的のために整備・運用されているものであるが、これらを複合的に活用することで、当初は想定もしていなかった価値のある結果が得られつつあることは、非常に興味深いことであると言えるのではないだろうか。

一方で、自治体データを様々な担当課を跨りながら収集し、活用する環境を整えることは、個人情報保護の観点や、データによっては目的外利用と判断される可能性も有するため容易ではない。とはいえ、以上の一連の研究を通して分かってきたことは、自治体データは継続的に維持・更新がされ続けている信頼性も更新頻度も高いデータであるため、都市・地域の様子を知る上で極めて有用性が高いということである。これらのデータをこれまで以上に高度利用する方法の研究、あるいは現時点では考えもしなかったような活用方法に関する研究、さらに自治体データを今まで以上にセキュアに管理・運用する方法の研究、自治体データ活用の際に適切な匿名化技術に関する研究などが今後数多く実施され、自治体データ活用のメリットを広く示すことができるようになることで、自治体データの利活用は今後ますます活発になるものと考えられる。そしてこれらの成果は、空き家問題を含む様々な地域課題の解決に貢献できるようになるものと期待している。

参考文献

- 秋山祐樹・柴崎亮介 (2015) マイクロ将来推計人口データを用いた将来の空き家分布推定. CSIS DAYS 2015 研究アブストラクト集, 41.
- 秋山祐樹・上田章紘・大野佳哉・高岡英生・木野裕一郎・久富宏大 (2018) 鹿児島県鹿児島市における公共データを活用した空き家の分布把握—自治体の公共データを活用した空き家の分布把握手法に関する研究(その1)—. 日本建築学会計画系論文集, 744, 275-283.
- 秋山祐樹・上田章紘・大内健太・伊藤夏樹・大野佳哉・高岡英生・久富宏大 (2019) 公共データを活用した空

- き家の分布把握手法の高度化—自治体の公共データを活用した空き家の分布把握手法に関する研究（その2）—。日本建築学会計画系論文集，764，2165-2174.
- 秋山祐樹（2019）ビッグデータは何を語るか？地理学×ビッグデータの可能性とその将来展望，地理空間，12（3），159-178.
- 秋山祐樹・馬場弘樹・徳富智哉（2020a）令和元年度和歌山県における空き家分布推定に関する研究成果報告書。和歌山県.
- 秋山祐樹・馬場弘樹・早川玲理・岡澤由季・六信孝則（2020b）令和元年度群馬県前橋市におけるEBPM推進に向けた官民ビッグデータの活用研究業務実績報告書。前橋市未来の芽創造課.
- 浅見泰司（2014）『都市の空閑地・空き家を考える』プロGRESS
- 今井絢・杉本慎弥・榊原渡・水石仁（2015）「空き家問題」の今後と中古住宅の活用可能性。知的資産創造，23（8），20-37.
- 加藤浩司・山本玲子・北島力・中島孝行・中島宏典（2009）伝健地区における空き家の発生要因と活用の仕組みに関する研究。住宅総合研究財団研究論文集，35，107-118.
- 北川敏男・坂元慶行・石黒真木夫・北川源四郎（1983）『情報量統計学』共立出版
- 久保倫子・益田理広（2015）岐阜市中心部における空き家増加の実態。日本地理学会発表要旨集 2015 年度日本地理学会春季学術大会，247-247.
- 佐々木貴生・佐野可寸志・川端光昭・梶雅弘（2010）農村地域における空き家提供意向の実態と提供促進政策の提案。農村計画学会誌，29，173-178.
- 篠部裕・占部智大（2014）空き家の適正管理条例の現状と課題 東日本の地方自治体を事例として。日本建築学会技術報告集，20（45），723-726.
- 鈴木圭一・伊藤伸一（2012）空き家の現状と対応方策の検討。JICE レポート，22，31-37.
- 総務省統計局（2019）平成 30 年住宅・土地統計調査 住宅数概数集計 結果の概要。
https://www.stat.go.jp/data/jyutaku/2018/pdf/g_gaiyou.pdf（最終閲覧日：2020 年 1 月 11 日）
- 日本弁護士連合会法律サービス展開本部自治体等連携センター（2018）『深刻化する「空き家」問題——全国実態調査からみた現状と対策』明石書店
- 馬場弘樹・秋山祐樹・谷内田修（2019）群馬県前橋市における公共データを活用した空き家分布推定手法の検討，第 28 回地理情報システム学会講演論文集，F-2-4，2019.
- 益田理広・秋山祐樹（2020）昨今の空き家研究の盛衰と手法について—日本における研究状況に着目して—。地理空間，13（1），1-26.
- 宮路和明・西村明宏・山下貴司（2015）『空家等対策特別措置法の解説』大成出版社
- Akiyama, Y., Ueda, A., Ouchi, K., Ito, N., Ono, Y., Takaoka, H. and Hisadomi, K. (2020) Estimating the Spatial Distribution of Vacant Houses using Public Municipal Data. Geospatial Technologies for Local and Regional Development, 165-183.
- Baba, H. and Asami, Y. (2017) Regional differences in the socio-economic and built-environment factors of vacant house ratio as a key indicator for spatial situation of shrinking cities. Proceedings of 2017 international conference of Asian-Pacific planning societies, #069.