4d-GIS と機械学習を用いた 建築物ストックの解体現象のモデル化 -福岡県北九州市を対象としたケーススタディ-

森田 大登[1]・深堀 秀敏[2]・郭静[3]・白川博章[3]・谷川寛樹[3]

[1]学生会員　名古屋大学大学院　環境学研究科（〒464-8601 名古屋市千種区不老町 D2-1）
E-mail: morita.masato@j.mbox.nagoya-u.ac.jp

[2]非会員　北九州市立大学　国際環境工学部

[3]正会員　名古屋大学大学院　環境学研究科（〒464-8601 名古屋市千種区不老町 D2-1）

既存の蓄積された物質資源を長期的且つ効率的に利用するストック型社会への移行が望まれる．建築物ストックに対して解体現象をモデル化する事で，地理的･社会的な因子を考慮した高解像度の滞留年数推計が可能となり，ストックの長期的利用の検討に有用である．本研究では，4d-GIS と教師有り機械学習の一つである CART 法を用いて，地理的･社会的要因を考慮した建築物の解体確率モデルを構築した．また，汎化性能及び感度の観点から，既存研究の一手法であるロジスティック回帰に基づくモデルとの比較を行った．未学習データに対して CART 法の AUC は 0.956，ロジスティック回帰の AUC は 0.896 となり，CART 法によるモデルの方がより汎化性能が高い結果となった．また，築年数が大きくなるほど解体確率は増加し，高齢者が多い地域ほど解体確率が増加しづらい傾向にあった．

Key Words: *Urban Metabolism, Material Stock, Building lifespan, Geographic Information System, Machine Learning, Classification and Regression Tree*

1. はじめに

環境負荷を最小限に抑える為に，既存の蓄積された物質資源(以下，ストック)を長期的且つ効率的に利用する事で資源の新規投入を抑制するストック型社会への移行が望まれる [1]．ストックの長期利用に関して，ストックが社会に投入されてから排出されるまでの年数を意味する"滞留年数"が一つの評価指標となり得る．ストック型資源の一つである建築物ストックに対して解体現象をモデル化する事で，地理的･社会的な因子を考慮した高解像度の滞留年数推計が可能となり，ストックの長期利用の検討に有用である．加えて，都市における建設系資源排出のタイミングの予測が可能となり，資源循環計画の観点からも重要性は高い． また、建築物は必ずしも使用後適切に解体されるのではなく、老朽空き家化して社会に滞留し続けるケースも存在し、退蔵ストックと呼ばれる [2]。退蔵ストックは未利用の物質資源であり潜在的な廃棄物でもある為、"解体されない"ケースも考慮した分析が資源の効率的利用の観点から重要である。

建築物の滞留年数に関する既存研究として，小松ら [3] は，区間残存率推計法を用いてサンプルデータを生成し，建物工法と建物用途毎の残存率曲線から滞留年数推計を行ったが，地理的･社会的な要因について議論し，ストックの長期的利用にむけた施策を具体的に検討するには解像度が不十分である．奥岡ら [4]は，名古屋市中心部における 4d-GIS を用いてサンプルデータを生成し，尤度最大化に基づくロジスティック回帰を用いて解体確率関数を構築し、地理的･社会的要因を考慮した高解像度の建築物滞留年数推計を行った。しかし、ロジスティック回帰モデル以外は検討されておらず、ストックの退蔵を考慮する為にも採用するモデル構造に議論の余地がある．加えて上記の研究では，構築したモデルのパラメータの統計学的な有意性や各説明変数の感度に関しては議論されているが，モデルの汎化性能に関しては検討されていない．

本研究の目的は，建築物ストックの解体現象を地理的･社会的要因を考慮した上でモデル化し，高解像度な滞留年数推計や建築物廃棄の将来予測に応用可能な手法を検討する事である．そのために 4d-GIS と教師有り機械学習を用いて建築物ストックの解体確率モデルを

構築し，汎化性能や要因についても検討する．

2. 研究手法

本研究では，福岡県北九州市をケーススタディ対象地域として時系列の GIS データを用いて建築物 4d-GIS データベースを構築し，連続する年代間(5 年毎)において各構造物の「解体」及び「残存」の情報を取得した．建築物解体をベルヌーイ分布に従う確率的事象とみなし，説明変数に建物年齢等の建物情報及び地理的・社会的要因を採用した解体確率モデルを構築した．モデルの構築には，教師有り機械学習の一つである決定木学習に基づく CART 法を用いた．また，既往研究の一手法である尤度最大化に基づくロジスティック回帰によるモデルも同様の条件で構築し，未知のデータに対する予測性能(汎化性能)及び感度の観点から，比較を行った．

(1) ケーススタディ対象地域及び 4d-GIS データ

本研究ではケーススタディ対象として，GIS データ等の整備状況を考慮し，福岡県北九州市を選定した．本市は，九州の最北端に位置し，面積約 492km2，人口約 95 万人の政令指定都市である．本市は全国的に人口減少・高齢化が顕著な都市である．また，都市のコンパクト化政策として 2017 年から立地適正化計画が施行されており，公共通利便性の高い都心・副都心・地域拠点(計 12 箇所)を都市機能誘導区域とし，及びその周辺に居住誘導区域が設定されている．少子高齢化により，日本の他の大都市でも将来同様の推移をする事が予想される為，本市での分析を他の都市のモデルとして利用可能であると考えられる．

また本研究では，建築物 3d-GIS データを時系列に重ねた 1985 年から 2019 年までの 7 年代の 4d-GIS データベースを構築し，連続する年代間(5 年毎)で各建築物に「解体」もしくは「残存」のラベルを付与した．

図 2 ケーススタディ対象地域

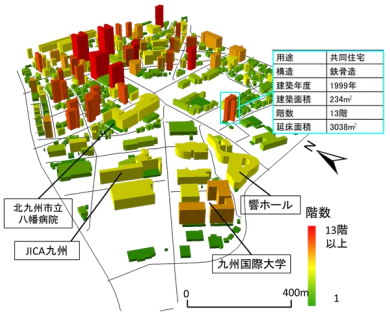

図 1 建築物 GIS データ(八幡駅周辺を抽出, 2018 年)

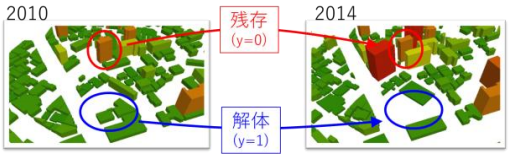

図 3 「解体」及び「残存」ラベルの付与

(3) 解体確率モデルの構築方法

解体/残存」のラベルが付与された建築物 GIS データを用いて，解体確率モデルを構築する．なお，データの整備状況や欠損を考慮し 2010 年から 2014 年における 202448 レコードをサンプルデータとする(**図 4** 参照)．本研究における解体確率モデルとは，「ある建築物が次年代(5 年後)までに解体されるかどうか」をベルヌーイ分布に従う確率的事象 y とみなし，ベルヌーイ分布のパラメータ，すなわち解体確率 λ を予測する数理モデルの事である．

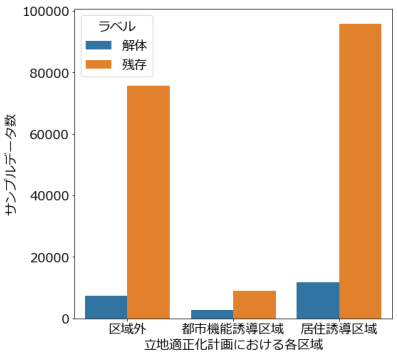

図 4 サンプルデータ数の内訳

解体確率 λ に関して，一般に建築物には耐用年数が設定されている事から，建物年齢が大きくなるにつれて増加すると予想される．また，建築物解体の傾向は建物年齢を含む建築物自体の特性に加えて，周辺環境の地理的・社会的要因によって異なると考えられる．従って，解体確率 λ に影響を与えると考えられる様々な要因を説明変数 X とし，λ を X(建物年齢，地理的条件，…，etc.)の関数(式 1)と仮定した．解体確率モデルで使用した説

表 1　解体確率モデルで用いた説明変数

要因グループ	説明変数名	説明変数の概要
建物の属性情報	建物年齢(年)	対象建築物の年齢
	RC 造ダミー変数(0 or 1)	建物工法が RC 造の建築物=1/それ以外=0
	木造ダミー変数(0 or 1)	建物工法が W 造の建築物=1/それ以外=0
	建築面積(m^2)	対象建築物の建築面積
地理的条件	傾斜角度(°)	10m メッシュの DEM(数値標高モデル)データを元に作成された 10m メッシュ毎の傾斜角度データ
社会的要因	市街地密度 (棟数/10000m^2)	対象建築物が所属する 100m メッシュに含まれる全建築物の棟数
	総人口(人/1km2)	対象建築物が所属する 1km メッシュ内の総人口(国勢調査より)
	高齢者人口(人/1km2)	対象建築物が所属する 1km メッシュ内の 65 歳以上の人口(国勢調査より)
公共交通利便性	駅からの距離(m)	対象建築物と最も近い駅までの距離
	バス停からの距離(m)	対象建築物と最も近いバス停までの距離
都市計画	立地適正化誘導区域 ダミー変数(0 or 1)	立地適正化計画における都市機能誘導区域もしくは居住誘導区域内の建築物=1/それ以外=0

明変数 X を表 1 に示す．本研究では，教師あり機械学習の一手法である CART 法 (Classification and Regression Trees)を用いてλと X を関連づけ，様々な空間情報(説明変数 X)から建築物の解体確率λを予測するモデルを構築した．CART 法による二分木の分岐基準にはジニ不純度を用いた．

$$\boldsymbol{y \in \{0,1\}}, \quad \boldsymbol{Pr(y|\lambda) = Bern_y(\lambda)}$$
$$\boldsymbol{\lambda \in [0,1]}, \quad \boldsymbol{\lambda = f(X) = f(x_1, x_2, \ldots, x_i)} \qquad (1)$$

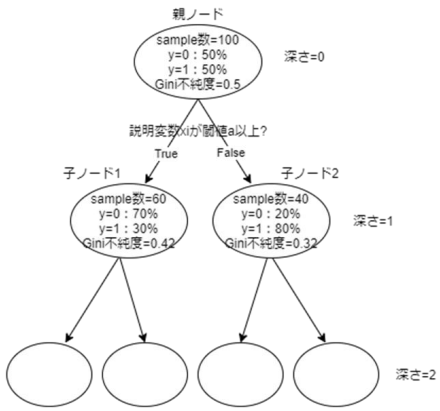

図 5　CART 法によるツリーモデル概念図

また，汎化性能及び感度の観点から既存研究の手法との比較を行う為，本研究では既存研究の一手法であるロジスティック回帰による解体確率モデルを同一のサンプルデータ及び説明変数を用いて構築した．

(4)　モデルの汎化性能の評価方法

モデルの汎化性能の評価方法にはホールドアウト法を用いる．サン プルデータを学習用と検証用に分け，学習用データを用いてモデルを構築し，検証用データを予測させて観測値と比較する事で，未知のデータに対するモデルの予測精度，すなわち汎化性能を評価する．本研究では，サンプルデータの 80%を学習用データ，20%を検証用データとして，解体/残存ラベル y の比率が等しくなるようにランダムに分割した．

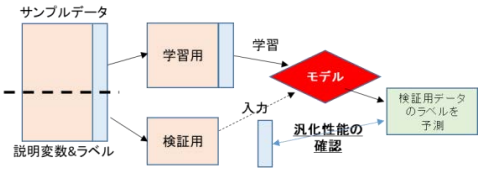

図 6　ホールドアウト法による学習と検証

また本研究では，AUC と呼ばれる指標を用いてモデルの予測精度を評価する．閾値による二値分類問題において，実際に陽性(y=1)であるサンプルのうち正しく陽性と予測されたサンプル数の割合を真陽性率(True Positive Rate)，実際に陰性(y=0)であるサンプルのうち誤って陽性と予測されたサンプル数の割合を偽陽性率(False Positive Rate)という．閾値を 0 から 1 の範囲で変化させた時の両値の変動を描いた図を ROC 曲線という． ROC 曲線から下側の面積である AUC(Area under the curve)が予測精度の評価指標として利用され，1 に近づくほど予測精度が高いと判断できる．

(5)　モデルの感度解析方法

本研究では，解体確率モデルにおいて説明変数 X の各要因が出力λにどのような影響を与えるかを把握する為に，PDP(Partial Dependence Plot，部分従属プロット)と Sobol'法を用いた感度解析を行う．PDP とは，構築した機械学習モデルの解釈可能性を得る為の 1 手法であり，各説明変数とモデル出力値がどのような関係にあるかを視覚的に把握する事ができる．また，Sobol'法は分散に基づく大域的感度解析の代表的手法であり，どの説明変数の不確実さがモデル出力の不確実さに寄与するかを，交互効果も含めて分析する事ができる．

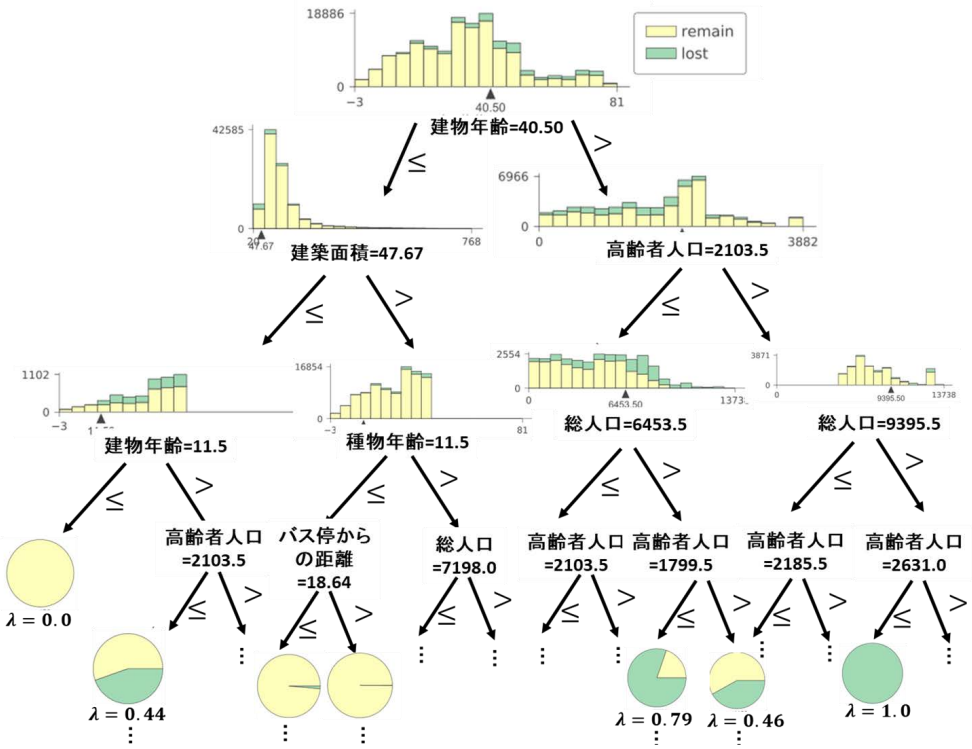

図 7　CART 法による解体確率モデルの二分木(15 層のうち，上から 4 層目まで)

3.　結果と考察

(1)　解体確率モデルの構築

学習用データ(200000 レコード)を基に，CART 法を用いて構築した解体確率モデルの二分木を**図 7** に示す．なお，本研究では二分木の成長制限として深さ 15 層，葉を構成する最小サンプル数 20，分岐に必要な最小サンプル数 2000 とした．一層目の分岐条件は建物年齢 ≤ 40.5 年であり，築 40 年以下の建物では，建築面積や人口情報，公共交通利便性等によって分岐される．一方，築 40 年以上の建物は周辺地域の人口情報に応じて解体確率が異なる傾向が示唆された．

(2)　解体確率モデルの汎化性能評価

作成した解体確率モデルを用いて学習用及び検証用データを予測させた時の ROC 曲線を**図 8** に示す．検証用データに対して CART 法の AUC は 0.953，ロジスティック回帰の AUC は 0.896 となり，CART 法によるモデルの方がより未知のデータに対する予測精度が高い結果となった．また，両モデルにおいて学習用データと検証用データの結果に大きな差が無いことから，学習データに過学習していない事が確認された．

なお，本研究において解体確率モデルの作成に使用したサンプルデータは「解体」と「残存」ラベルの割合が不均衡なデータとなっており，このようなデータを用いてモデルを構築する場合，閾値による分類問題において小数側の判定数が小さくなる傾向にある．従って本研究ではサンプル数に応じてラベル毎の重みを調整した．

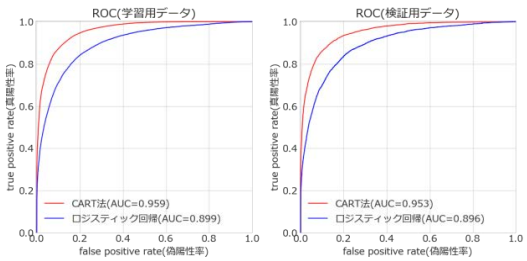

図 8　解体確率モデルにおける ROC 曲線

(3)　解体確率モデルの感度解析結果

解体確率モデルにおける各説明変数の PDP を**図 9** に示す．横軸は対象の説明変数の値，縦軸は対象の説明変数の値を変動させた時のモデル出力 λ (解体確率)を表す．両モデルにおいて解体確率を大きく変動させる説明変数は共通しており，建物年齢が大きくなるほど解体確率が増加し，また総人口が多く且つ高齢者人口が少ない地域の建築物ほど解体確率が大きくなる結果となった．一方で，RC 造ダミー変数及び木造ダミー変数は解体確率に大きな変化を与えない結果となり，建物工法の違いは建

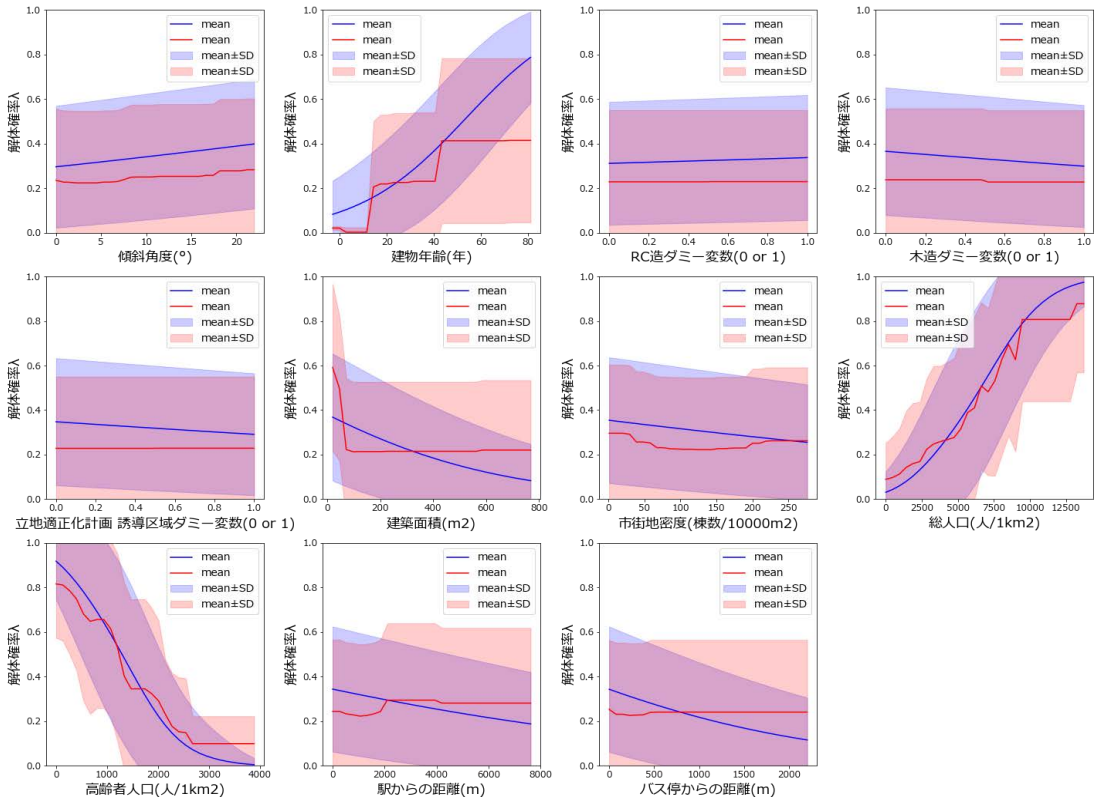

図 9　解体確率モデルにおける PDP(赤：CART 法，青：ロジスティック回帰)

築物の解体確率に，直接的な影響を与えない事が示唆された．また，立地適正化計画ダミー変数も解体確率には大きな影響を与えず，市の指定する誘導区域の内外において建築物の解体傾向に大きな変化は確認されない結果となった．建物年齢に関して，ロジスティック回帰では年齢の増加に伴って解体確率も増加するが，CART 法では築 40 年以上は解体確率は一定となる点が，両モデルの違いだといえる．

以下に Sobol'法の感度解析結果を示す．**図 10** は，算出された解体確率モデルにおける Sobol'法の一次感度指標 S_i，二次感度指標 S_{ij}，及び 3 次元以上の感度指標の内訳を示したものである．不確実性の約 9 割を 1･2 次元の効果が占めており，総人口･高齢者人口･建物年齢の主効果 S_i，及びそれら 3 つのいずれかを含む交互効果 S_{ij} が不確実性寄与分の上位を占める事が分かった．特にロジスティック回帰によるモデルでは，総人口，高齢者人口，及び両値の交互効果がモデル出力の不確実性の約 84%を占める結果となった．

また，**図 11** は Sobol'法の総感度指標 S_{Ti} を図示したものである．両モデルに共通して，高齢者人口が最も解体確率の不確実さに寄与し，次いで総人口，建物年齢という結果になった．ロジスティック回帰を用いたモデルでは高齢者人口と総人口の寄与分が他の説明変数と比較して極端に大きい事が，総感度指標からも明らかになった．一方で CART 法を用いたモデルでは，建物年齢の影響が大きくなり，公共交通利便性や建築面積等との交互作用も含めて再現可能なモデルである事が考察された．

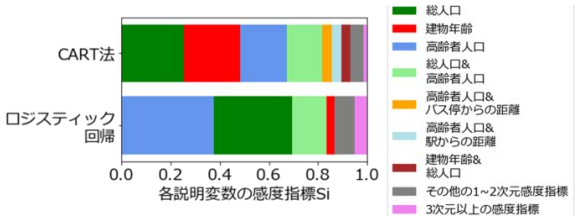

図 10　Sobol'法における各感度指標の内訳

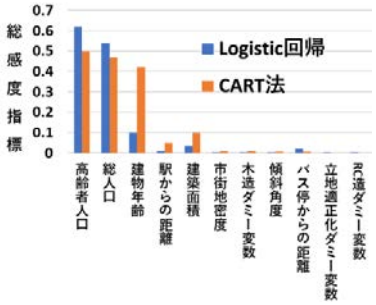

図 11　各説明変数の総感度指標 S_{Ti}

4. おわりに

本研究では，福岡県北九州市を対象に建築物 4d-GIS データベースを構築し，連続する年代間(5 年毎)において各構造物の「解体」及び「残存」の情報を取得した．建築物の解体を確率的事象とみなし，機械学習の 1 手法である CART 法を用いて，説明変数に築年数等の建物情報や地理的･社会的を採用した解体確率モデルを構築した．既存研究の 1 手法であるロジスティック回帰を用いたモデルも同様に構築し，汎化性能及び感度の観点から比較を行った．

CART 法で構築された二分木から，建物年齢 40 年の前後で解体の傾向が異なり，築 40 年以上の建物は周辺地域の人口情報に応じて解体確率が異なる事が明らかになった．モデルの汎化性能に関して，未学習データに対して CART 法の AUC は 0.956，ロジスティック回帰の AUC は 0.896 となり，CART 法によるモデルの方がより汎化性能が高い結果となった．モデルの感度解析に関して，PDP の結果，両モデルにおいて解体確率を大きく変動させる説明変数は共通しており，築年数が大きくなるほど解体確率は増加し，また総人口が多く高齢者人口が少ない地域ほど解体確率が大きくなる傾向にあった．Sobol 法の結果，ロジスティック回帰によるモデルでは高齢者人口と総人口の影響が全体の 84%を占め，CART 法によるモデルでは各説明変数の偏りが小さく，より多くの要因や交互効果を再現可能である事が示唆された．

今後の課題として，他地域のサンプルデータの拡充が挙げられる．本研究で構築した解体確率モデルは，北九州市の建築物解体に過学習している可能性を否定できない．他年代のデータや本市とは都市形成の背景が異なる地域のデータを拡充する事で，より多種の要因を考慮したモデルを構築できると考えられる．また，モデル構造に関しても，汎化性能と解釈可能性の両面から改善の余地がある．本研究で構築したモデルは，モンテカルロ法を用いて高解像度の滞留年数推計や建築物廃棄の将来推計に活用可能であると考えられる．

謝辞：本研究は，環境省環境研究総合推進費 (1-2003, s18 - 4 – 2, 2 - 2013)の助成を受けて行われたものである．ここに記して謝意を表する．

参考文献

1) 岡本久人. (2006). ストック型社会への転換. 鹿島出版社.
2) 谷川寛樹, 醍醐市朗, 小口正弘, 奥岡桂次郎, 高木重定. (2017). 物質ストック・フローに着目したストック型社会構築に向けた指標. 廃棄物資源循環学会誌, Vol. 28, No. 6, pp. 431-437
3) 小松幸夫, 加藤裕久, 吉田倬郎, 野城智也. (1992). わが国における各種住宅の寿命分布に関する調査報告. 日本建築学会計画 系論文報告集,第 439 号
4) 奥岡桂次郎, 野中一鴻, 谷川寛樹. (2018). 名古屋市中心部における 4d-GIS を用いた建築物滞留年推計モデルの構築と物質ストック分析. 環境システム研究論文集 Vol.46, p. II_267-II_273.
5) Alessio Miatto, Heinz Schandl, Hiroki Tanikawa. (2017). How important are realistic building lifespan assumptions for material stock and dmeolition waste accounts?. Reseources, Conservation and Recycling.Vol.122, 143-154,
6) Leo Breiman. (1984). Classification ans Regression Trees, Taylor & Francis
7) 北九州市立大学(2016):北九州市都市計画 GIS データ
8) 北九州市(2019)：北九州市立地適正化計画 計画書
9) 総務省統計局：平成 27 年，平成 22 年，平成 17 年，平成 12 年国勢調査結果
10) 国土地理院 数値標高モデル 10m メッシュ

(Received August 23, 2021)

Modeling of Building Demolition from Existing Stock Using 4d-GIS and Machine Learning – the Case Study of Kikakyusyu, Japan-

Masato MORITA, Hidetoshi HUKAHORI, Guo JING, Hiroaki SHIRAKAWA and Hiroki TANIKAWA

It is desirable to shift to a stock-based society in which existing accumulated material resources are used efficiently for a long period of time. By modeling the demolition phenomenon for the building stock, it is possible to estimate high-resolution building lifespan considering geographical and social factors, which is useful for studying the long-term use of the stock. In this study, we developed a building demolition probability model considering geographic and social factors by using 4d-GIS and supervised machine learning method – CART(classification and regression trees). For untrained data, the AUC(area under the curve) of the CART method was 0.956, and the AUC of logistic regression was 0.896, indicating that the model based on the CART method had higher prediction accuracy for unknown data. In addition, the probability of demolition increased as the building age increased, and it tended to be lower in areas where many elderly people lived.