

# マイクロジオデータと機械学習手法を用いた メッシュ間OD交通量の推定

奥村 航太<sup>1</sup>・生越 拓実<sup>2</sup>・有村 幹治<sup>3</sup>・浅田 拓海<sup>4</sup>

<sup>1</sup>学生会員 室蘭工業大学 建築社会基盤系学科 (〒050-8585 室蘭市水元町27-1)

E-mail: 15021021@mmm.muroran-it.ac.jp

<sup>2</sup>学生会員 室蘭工業大学 大学院工学研究科 環境創生工学系専攻 (〒050-8585 室蘭市水元町27-1)

E-mail: 16041014@mmm.muroran-it.ac.jp

<sup>3</sup>正会員 室蘭工業大学准教授 大学院工学研究科 暮らし環境系領域 (〒050-8585 室蘭市水元町27-1)

E-mail: arimura@mmm.muroran-it.ac.jp

<sup>4</sup>正会員 室蘭工業大学助教 大学院工学研究科 暮らし環境系領域 (〒050-8585 室蘭市水元町27-1)

E-mail: asada@mmm.muroran-it.ac.jp

我が国におけるコンパクト・プラス・ネットワーク施策の展開においては、建物立地状況や移動環境の変化に伴う人流の変化をより簡便に予測する必要がある。本研究の目的は、近年整備が進む都市マイクロジオデータを用いて、メッシュ毎の建物の用途別延床面積と滞在人口、移動時間から、メッシュ間OD交通量を推定することにある。建物情報は都市計画基礎調査、滞在人口とOD交通量は携帯電話網の運用データであるモバイル空間統計を用いた。推定モデルには4分割の交差検証を行うことで、予測値の精度を計測し、各説明変数がOD交通量に与える影響を確認した。

**Key Words :** *Neural Network, Random Forest, Mobile Spatial Statistics, Compact city plus network*

## 1. はじめに

我が国は戦後70年が過ぎ、高度成長期に建設された建物の老朽化が問題として挙げられ、更新時期が迫っている。加えて、人口の減少により、住居地と都市機能を集約し公共交通でそれらを連絡するコンパクト・プラス・ネットワーク施策の展開が急務となっている。しかし、コンパクト・プラス・ネットワーク施策の展開を検討する際には、例えば、どの種別の都市施設を、どの規模で、どの地区に集約させるべきかを考える必要がある。また都市内のアクセスの改善が、どの程度のOD交通量に影響を与えるのかといった、都市施設誘導や交通アクセスに関する諸課題が挙げられる。一概にコンパクト・プラス・ネットワーク施策を掲げても、対象となる都市の特性に応じて、施策展開の方向性は異なる。例えば、本研究で対象とする札幌市においては、2018年10月に超高層建築物である複合施設の「さっぽろ創生スクエア」がオープンし、大規模な都心部のリノベーションが見られた。また、国道5号線の創生川通については都心アクセス道路を新たに建設することが計画されており、観光やビジネス、物流などの改善が期待されている<sup>1)</sup>。

一方、近年では諸都市の状況に応じた計画情報を得るため、マイクロジオデータの適用が期待されている。例えば、コンパクト・プラス・ネットワーク政策の立案支援のためにマイクロジオデータを用いた研究事例として今井ら<sup>2)</sup>は人口流動統計であるモバイル空間統計<sup>3)</sup>やパーソントリップ調査、携帯電話のGPSデータを組み合わせ、自動車OD交通量を推計した。著者らによる先行研究<sup>4)</sup>では、札幌市を対象に、モバイル空間統計データと都市計画基礎調査から得られる全34種類の建物用途別変数を活用するため、非線形近似が可能なニューラルネットワークモデルの一種であるRBFネットワークを用いてメッシュごとの対象時間帯の入込人口を推計している。また、より詳細なメッシュに空間分解能をダウンスケールする手法を開発している。しかしながら、この研究は各メッシュに注目し、建物用途と規模とそのメッシュの滞在人口を機械学習手法を用いて推定するものであり、メッシュ間の移動時間を用いた分析ではなかった。

そこで本研究では、建物用途別延床面積と滞在人口、移動時間からOD交通量を推定する手法を開発する。複数の説明変数について、ランダムフォレスト法を用いて寄与度が高くなるようにデータセットを組み替え、ニュー

ーラルネットワークを適用することで、より高い予測精度のモデルを構築する。移動時間短縮や建物用途の変更や誘導に応じたOD交通量の変化を簡便に予測することができる。また、将来を見据えた効率の良い社会基盤への投資に繋がるものと考えられる。

本研究の構成は、以下第2章において、使用したデータの概要を説明する。第3章では機械学習による分析方法を示す。第4章では目的である各説明変数の寄与度を算出し、用途の集計方法別によるOD交通量の推定精度について検証する。第5章では本研究の結論を述べる。

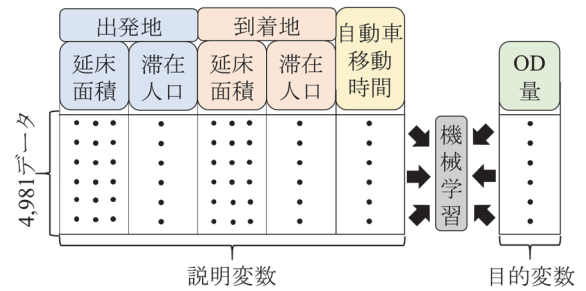


図-1 本研究におけるデータセット

## 2. 使用データ

本研究では、図-1のようなデータセットを作成し、建物用途別延床面積、滞在人口、移動時間からOD交通量の分析を行う。それぞれのデータについては都市計画基礎調査、モバイル空間統計、Google Distance Matrix API<sup>5)</sup>を用いる。使用したデータの概要を以下に示す。

### (1) 都市計画基礎調査データ(用途別延床面積)

都市計画基礎調査とは人口、産業、土地利用、建築物、交通などのデータを記録し、都市計画の運用を行うための基礎となるものである。建物利用状況の項目では個々の建物に関する用途、階数、構造、延床面積、築年数などがまとめられている。

本研究では平成 29 年度の札幌都市圏「都市計画基礎調査」を使用した。まずは全 430,884 棟の建物について図-2 のように GIS 上に表示させた。次に表-1 のおける用途別に 3 次メッシュ(約 1km<sup>2</sup>)毎の延床面積を集計した。

### (2) モバイル空間統計(滞在人口・OD交通量)

モバイル空間統計とは、NTT Docomo の携帯電話ネットワークの運用データを用いた人口統計である。国内約 7,600 万台と訪日外国人約 750 万台<sup>6)</sup>から拡大推計されたものであり、ある特定日時や性別、年代、調査エリアの規模などの様々な条件の元、情報を取得可能である。従って、観光客分析やまちづくり、防災計画などあらゆる分野から注目が集まっている。本研究では、あるエリアに滞在している人口(以下、滞在人口)や、ある 2 つのエリア間の OD 交通量(単位:trip)を解析で用いた。本研究におけるデータの条件は表-2 に示す。

OD 交通量に関しては、更に半径 1km 以上の移動でトリップ数で札幌市内で完結する移動のみを抽出した。結果としてメッシュ間の移動は 4,891 パターン(合計 362,557 トリップ)となった。図-3, 4 から札幌の中心部に流れる OD が多いことがわかる。また、札幌市内の該当するメ

表-1 都市計画基礎調査の建物用途区分

大分類	中分類	小分類
商業	官公省施設	地方国家施設, 自治体施設
	専用商業施設	業務施設, 集合販売施設, 宿泊施設
	娯楽施設	興業施設, 風俗営業施設, 遊戯施設, スポーツ施設
	店舗施設	専用店舗施設
住宅	住居施設	専用住宅, 共同住宅, 作業所併用住宅, 事務所併用住宅, 飲食店併用住宅, 一般店舗併用住宅
	文教厚生	文教施設
工業	厚生施設	医療施設, 運動施設, 社会福祉施設, 厚生施設
	工業施設	重化学工業施設, 軽工業施設, サービス工業施設, 家内工業施設
	都市運営施設	供給処理施設, 運輸倉庫施設, 通信施設
他	農業施設	農業施設, 漁業施設

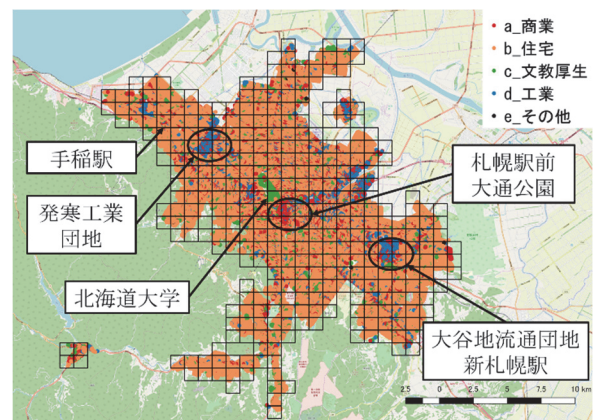


図-2 大分類による建物用途別の分布



ッシュに対して滞在人口は 8 時で 1,472,129 人(図-5), 3 時で 1,529,644 人となった。札幌市における平成 27 年度の国勢調査<sup>9)</sup>によると 15~79 歳の人口は 1,580,054 人である。ゆえに、モバイル空間統計の滞在人口は約 90% 以上の捕捉率となる。

### (3) 移動時間の推計

ここでは、札幌市内 369 メッシュを対象とし、全メッシュの相互間(136,161 通り)の移動時間を推計する。解析の対象地域はグリッド型都市の札幌市であるため、マンハッタン距離を指標として推計を簡易化した。

初めに、メッシュの中心座標の緯度経度の差分から全相互間のマンハッタン距離を算出した。次に、無作為に 2,000 通りの移動パターンを抽出し、その間の移動距離と自動車の移動時間を Google Distance Matrix API(以下、Google API)を用いて求めた。ここで図-6 で示すように、マンハッタン距離と Google API による移動距離の関係から一次近似式を導出した。この近似式を用いてマンハッタン距離から全相互間の移動距離を推定した。最後に、図-7 より Google API から取得した移動距離と移動時間の関係から一次近似式を導出し、先に求めた推定の移動距離から推定の移動時間を求めた。この値を本研究で用いる移動時間とする。なお、Google API による移動手段は全て自動車であり、高速自動車国道を通過しない移動時間である。

## 3. 分析方法

本章では、OD交通量の推定と各説明変数の寄与度を得る手法について焦点を当てる。具体的には機械学習に分類されるニューラルネットワークとランダムフォレストを概要を述べる。

### (1) ニューラルネットワークの概要

ニューラルネットワーク(Neural Network, 以下NN)とは機械学習の代表例であり、非線形なデータにも柔軟に対応することが可能である。構成については図-8のように入力層、中間層、出力層の3層以上であり、パターン認識や回帰分析などのあらゆる問題に適用されている<sup>78)</sup>。

$$u = \theta + \sum_{i=1}^K w_i x_i \quad (1)$$

$$z = f(u) \quad (2)$$

ただし、 $u$ : ニューロンへの入力値、 $x$ : 入力値、 $w$ : 重み、 $\theta$ : バイアス、 $K$ : 直前の階層におけるニューロ

表-2 モバイル空間統計の集計条件

データ	滞在人口	OD交通量
日付	2016年9月5日(水)	
時間帯	8時	7~9時台の合計
規模	3次メッシュ	
年代	15~79歳の総和	

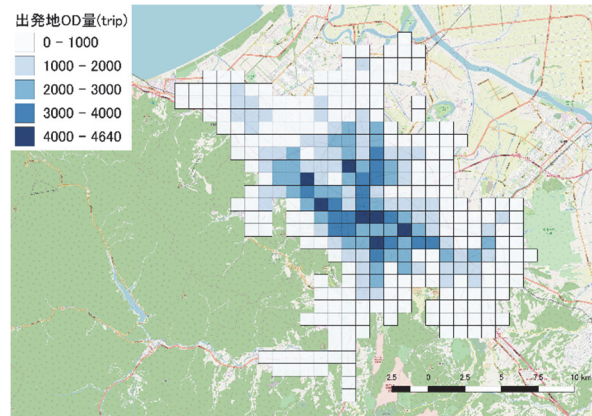


図-3 出発地における OD 交通量

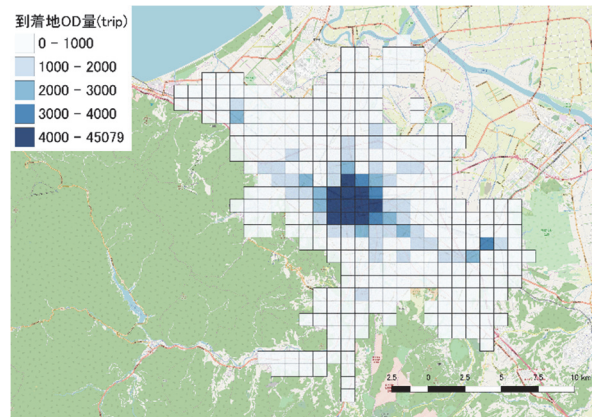


図-4 到着地における OD 交通量

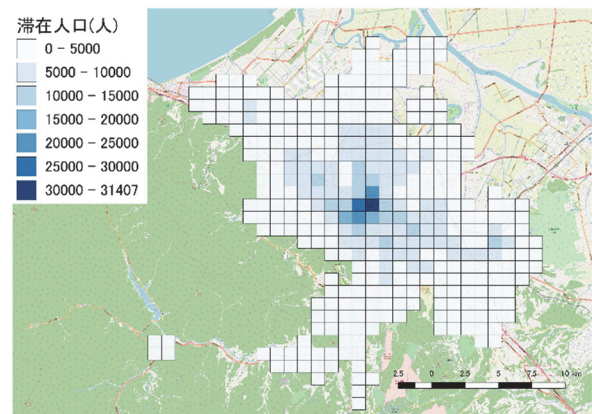


図-5 平日 8:00 の滞在人口

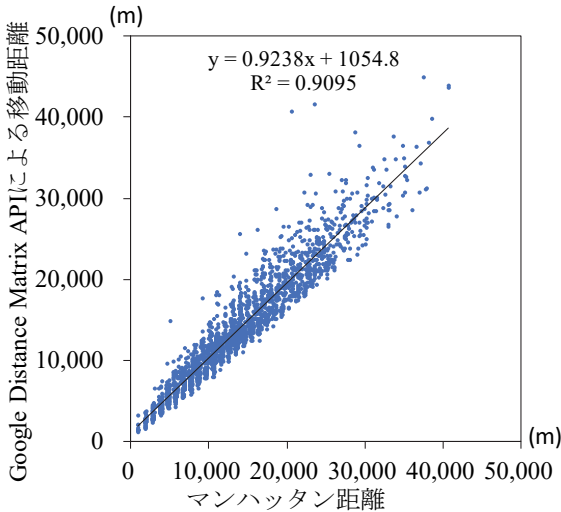


図-6 距離に関する相関

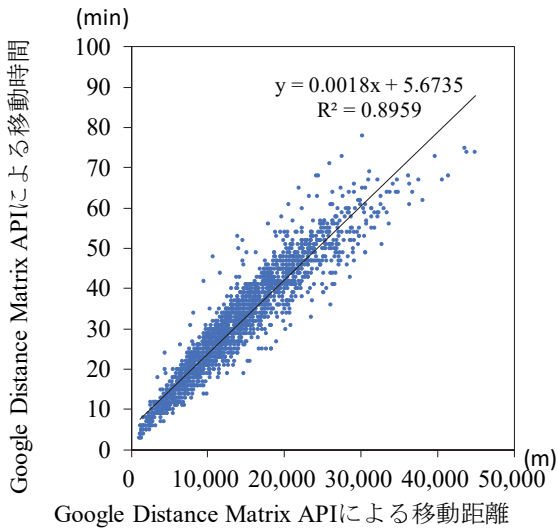


図-7 移動時間と移動距離の関係

ン数,  $f(u)$ : 活性化関数,  $z$ : 出力値である.

本研究では最大繰り返し学習回数を4,000回とし, 中間層のチューニングは3層を固定のニューロン数を $2^n$  ( $n=1,2,3,\dots,8$ )である8パターン行う. また, 活性化関数については一般的であるシグモイド関数(式(3))を用いた.

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (3)$$

## (2) ランダムフォレストの概要

ランダムフォレスト(Random Forest, 以下RF)は機械学習の一種であり, ランダムに説明変数を抽出し, 複数の決定木から目的変数を多数決や平均化することで, 分類や回帰分析を行うことができる<sup>9)</sup>. さらに, RFでは説明変数の寄与度を算出することが可能である.

本研究では, 各説明変数が目的変数にどれほどの影響があるのか, つまり寄与度を得ることができる点に注目

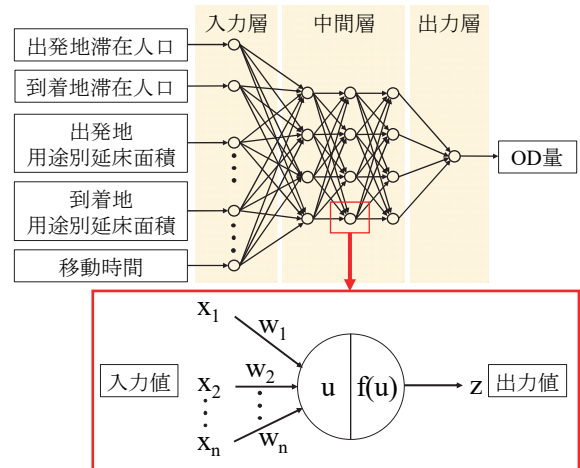


図-8 ニューラルネットワークの構成

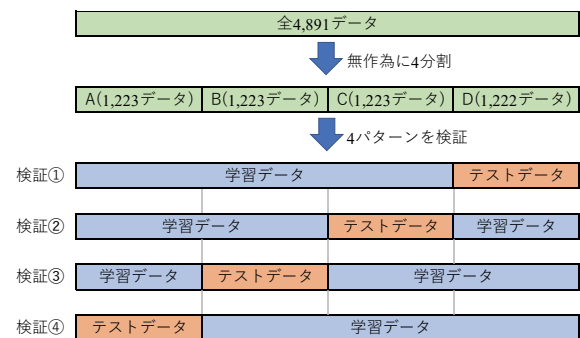


図-9 4分割交差検証

し, 本手法を導入した. 本研究では, 決定木の本数を1,000刻みで1,000~10,000本の10回, 寄与度を算出し, その平均値を最終的な値としている.

## (3) データセットの概要

本研究のデータセットは図-1で示した構造である. 説明変数は, 出発地と到着地に該当するメッシュの建物用途別延床面積と滞在人口, その間の自動車での移動時間であり, 目的変数はそれに対応するOD交通量としてモデルを構築する. また, 延床面積の用途に関しては, 小分類と中分類において作成した. モデル式の汎化性を確認するため, 解析の対象となる移動の4,891パターンを無作為に4分割することで, 3/4を学習データ, 1/4をテストデータとして, それぞれ入れ替えて学習させる. 即ち, 4分割交差検証(4-fold cross-validation)でモデルの精度を確認する. なお, 各データ $a_i$ は式(4)を用いて項目別に正規化を行い, 分析にはその値 $x(0 \leq x \leq 1)$ を用いる.

$$x = \frac{a}{a_{max}} \quad (4)$$

$a_{max}$ : その項目における最大値

## 4. 分析結果

本章ではRFによる各説明変数の重要度とNNによるOD交通量の推定結果を示す。

### (1) 各説明変数の寄与度

図-10, 11から建物用途別延床面積を小分類(34種類)と中分類(10種類)で区分した場合における説明変数について、寄与度をRFで算出した。どちらの区分においても滞在人口と移動時間がOD交通量を予測するために重要な説明変数であることがわかる。小分類の場合の建物用途別延床面積は大半が機能していない状態である。つまり、建物の用途や規模の変更が生じたとき、OD交通量への影響はかなり低いと考えられる。図-11は建物用途区分を中分類で行ったものである。やはり、滞在人口や移動時間には劣るが、図-10より建物がOD交通量に与える影響が大きいことが確認できる。

### (2) OD交通量の推定結果

各分類において交差検証における最終的な中間層の構成はAIC(赤池情報量基準)を用いて決定した。表-3で示すように結本研究の結果は、中間層が3層で1層当たりニューロン数は小分類で4と中分類で8が適当であった。双方のR<sup>2</sup>値0.74程度、MAE値は23.1トリップ程度であった。モデルの精度については多少、中分類で作成したデータセットの方が良いが、大きな差は見られなかった。また、図-12, 13の解析結果において、トリップ数に大きな誤差を観測したところはメッシュ内や近接な間での移動であることが多かった。また森林等の市街地の領域が少ないメッシュで存在したことがわかった。

## 5. 結論

本研究は、建物用途別延床面積と滞在人口、移動時間からOD交通量の推定を行い、各説明変数の寄与度を示した。今回作成したモデル式から建物の用途や規模または人口の変化、移動時間の短縮などによりOD交通量の増減を数値で算出することは可能である。しかし、建物用途別延床面積の寄与度が低いことが問題として挙げられる。現在の日本では建物の老朽化は大きな課題とされているため、そのような背景で土地利用のリノベーションが起こった際、OD交通量の推測することは都市計画を考える上で極めて重要である。

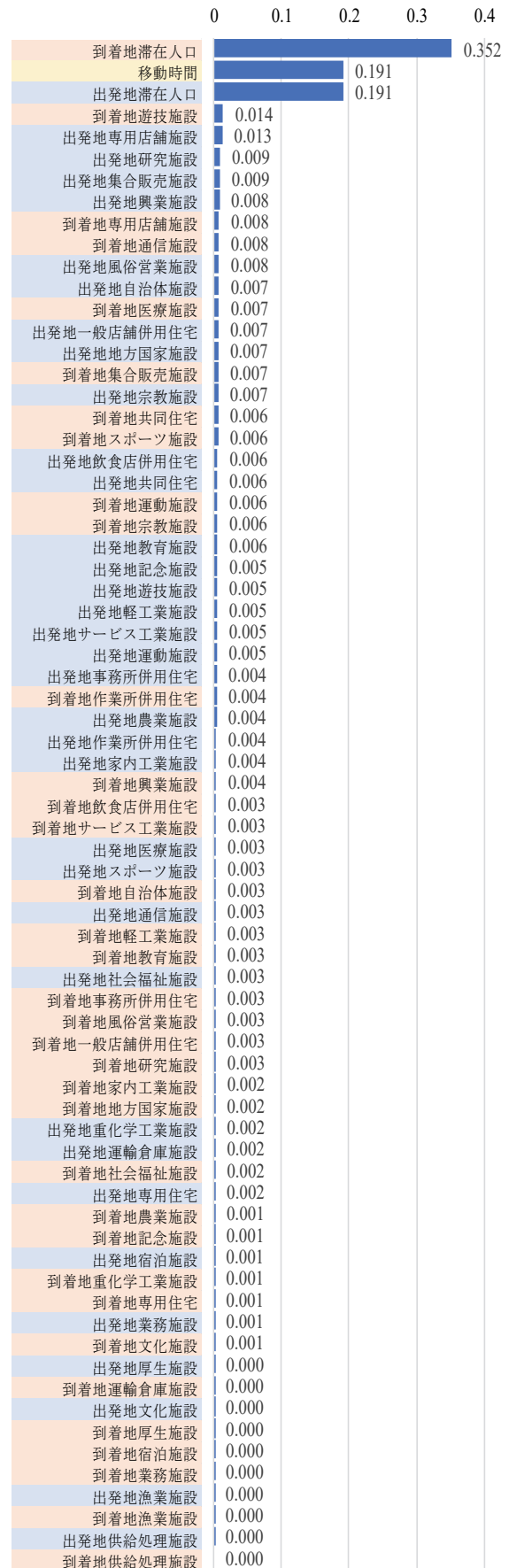


図-10 建物用途区分が小分類の場合における説明変数の寄与度



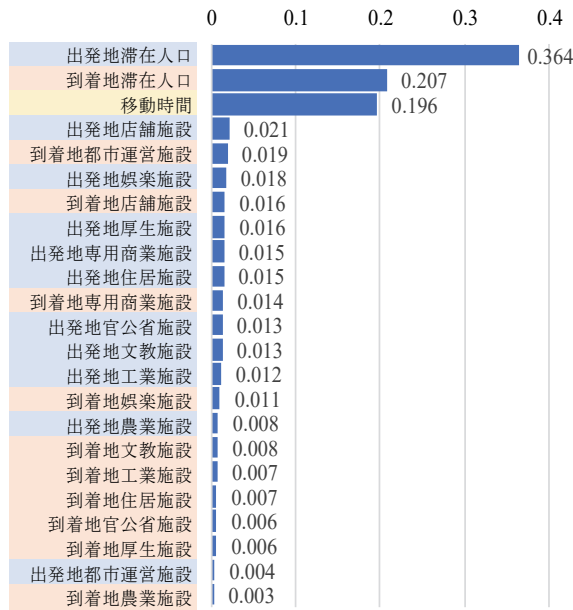


図-11 建物用途区分が中分類の場合における説明変数の寄与度

本研究では都市計画基礎調査とモバイル空間統計の2つマイクロジオデータと機械学習のNNを用いて、OD交通量を推定するモデルを構築した。また、RFによって各説明変数の寄与度を明記した。その結果として得られたことを次に示す。

- ・滞在人口とメッシュ間の位置関係がODに強い影響を与えていることが明らかになった。
- ・札幌市内の平日における通勤時間帯の移動に関して、実測値と予測値に強い相関がみられた。
- ・到着地より出発地の建物用途別延床面積がOD交通量に大きく影響を与えていることを明らかにした。
- ・延床面積の集計に関して、用途の小分類と中分類の違いから、OD交通量の推定に大きな差は見られなかった。
- ・建物用途区分を小分類で構築したモデルを使用すると、土地利用の変化が生じるとOD交通量の推測は困難である可能性が高い。

上記を踏まえ、OD交通量の推計精度に大きな影響を与えずに建物用途に関する説明変数の寄与度を向上させることが今後の課題である。移動時間に動的な情報を持つプローブや、移動手段及びその選択率を組み合わせることにより、実用的なモデルの構築を進めたい。

謝辞：本研究は Japan Science and Technology Agency (JST), e-ASIA Joint Research Program の助成を受けた。ここに記して感謝の意を表する。

表-3 各条件による予測精度

用途	中間層	AIC	R <sup>2</sup>	MAE	MSE	RMSE
小分類	2	51375	0.612	27.3	2073	45.5
	4	49472	0.735	23.0	1405	37.5
	8	49751	0.725	23.0	1487	38.6
	16	49816	0.720	22.4	1507	38.8
	32	49730	0.727	22.7	1481	38.5
	64	50214	0.716	23.4	1635	40.4
	128	49866	0.722	23.1	1523	39.0
	256	50242	0.707	23.4	1644	40.5
	中分類	2	51683	0.584	28.1	2207
4		49455	0.736	22.8	1400	37.4
8		49243	0.748	23.2	1340	36.6
16		49977	0.708	22.5	1558	39.5
32		50012	0.718	23.4	1569	39.6
64		49676	0.723	22.5	1465	38.3
128		50231	0.704	24.0	1640	40.5
256		50016	0.703	22.8	1570	39.6

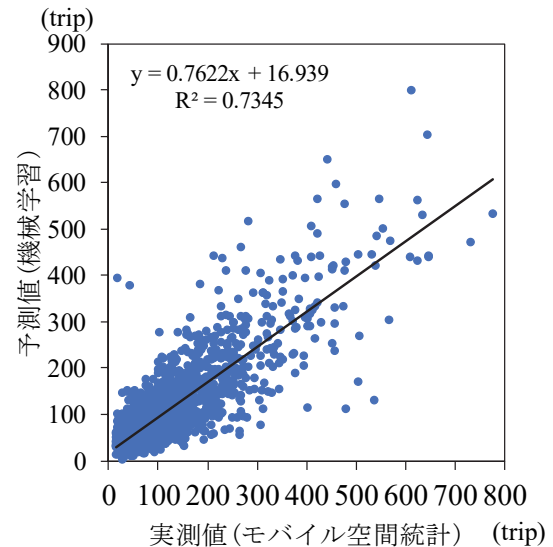


図-12 小分類における予測結果

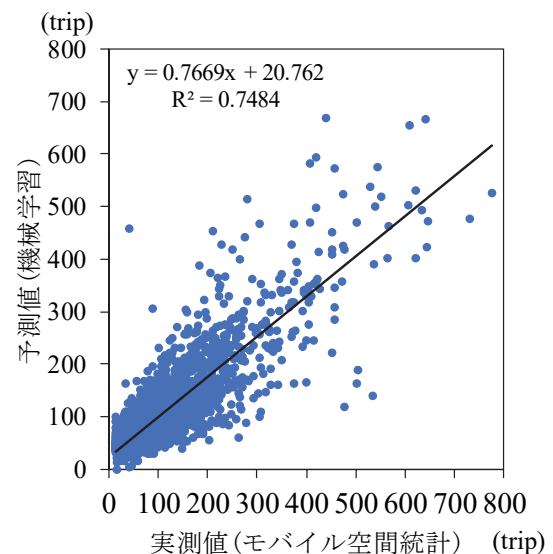


図-13 中分類における予測結果

## 参考文献

- 1) 札幌市「都心アクセス強化(創生川通の機能強化)／札幌市」<http://www.city.sapporo.jp/sogokotsu/toshin-access/index.html>, (閲覧日: 2019年3月9日)
- 2) 今井龍一, 池田大造, 永田智大, 福手亜弥, 金田穂高, 重高浩一, 鳥海大輔, 廣川和希: 携帯電話網の運用データを用いた人口流動系から算出した自動車OD交通量と道路交通センサスとの比較分析—道路交通分野へのモバイル空間統計の適用可能性—, 土木計画学研究・講演集, 53, 619-627.
- 3) NTTdocomo 「モバイル空間統計」のしくみ  
[https://www.nttdocomo.co.jp/biz/service/spatial\\_statistics/function/](https://www.nttdocomo.co.jp/biz/service/spatial_statistics/function/), (閲覧日: 2019年3月9日)
- 4) 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の開発, 土木学会論文集D3(土木計画学), 73巻5号, I\_483-I\_491, 2017.
- 5) Developer Guide | Distance Matrix API | Google Developers : <https://developers.google.com/maps/documentation/distance-matrix/intro>, (閲覧日: 2019年3月9日)
- 6) 札幌市国勢調査, 平成27年国勢調査第8表  
<http://www.city.sapporo.jp/toukei/tokusyu/documents/27kokutyoo08.pdf>, (閲覧日: 2019年3月9日)
- 7) 一言正之, 桜庭雅明: 深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法. 土木学会論文集B1(水工学), 73巻1号, 22-33, 2017.
- 8) 堀木賢一, 福田正.: ニューラルネットワークによる舗装のパフォーマンスモデル, 土木学会論文集, No.496, pp.99-102, 1994
- 9) 全邦釘, 井後敦史: Random forestによるコンクリート表面ひび割れの検出, 土木学会論文集F3(土木情報学), 71巻2号, I\_1-I\_8, 2015.

(2018. 3. 9 受付)

ESTIMATION OF INTER-MESH OD FLOW  
BY USING THE MICRO-GEO DATA AND MACHINE LEARNING

Kota OKUMURA, Takumi OGOSHI, Mikiharu ARIMURA, Takumi ASADA