

# 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>・有村 幹治<sup>2</sup>・浅田 拓海<sup>3</sup>

<sup>1</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: 19041016@mmm.muroran-it.ac.jp

<sup>2</sup>正会員 室蘭工業大学准教授 大学院工学研究科 もの創造系領域 (〒050-8585 室蘭市水元町27-1)

E-mail: arimura@mmm.muroran-it.ac.jp (Corresponding Author)

<sup>3</sup>正会員 室蘭工業大学助教 大学院工学研究科 もの創造系領域 (〒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観測、予測しつつ、都市施設誘導や交通整備を進める必要がある。本研究の目的は、近年整備が進む各種のマイクロジオデータを用いて、メッシュ間OD交通量を推定することにある。本研究では、都市計画基礎調査からメッシュ毎の建物用途別延床面積データ、携帯電話網の運用データであるモバイル空間統計からメッシュ滞在人口データ及びメッシュ間OD交通量データ、また各種オープンデータからメッシュ間移動時間データを取得した。そのうえで、ランダムフォレスト法を用いて、各説明変数の寄与度を確認し、ニューラルネットワークを用いたOD交通量の推定を試み、推定精度の検証した。

**Key Words :** *neural network, random forest, mobile spatial statistics, compact city plus network*

## 1. はじめに

我が国では人口減少や街なかの活性化、また低炭素社会づくりの視点から、居住地と都市機能を集約し、公共交通で拠点間を連結するコンパクト・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、どのエリアに誘導させるべきかを考える必要がある。また都市内の移動環境の改善が、どの程度の交通量に影響を与えるのかといった都市施設誘導の効果や、交通アクセス改善効果の評価が必要となる。また、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。しかしその一方で、依然として郊外部での住宅数は増加傾向にある<sup>2)</sup>。

このような都市施設や交通整備の影響を評価するために、土木計画では現在までに多くの土地利用・交通モデルに関する多くの研究蓄積がなされてきた<sup>3) 4)</sup>。土地利用・交通モデルの重要性が現在も変わらないことは言うまでもない。しかし、例えばモデルのベースとなるパーソントリップ調査の実施間隔は、概ね約10年と長く、その間の都市施設整備や交通環境整備効果を観測するには個別の調査を実施する必要があった。そのため、近年では諸都市の状況に応じた計画情報を得るため、ミクロスケールの様々な時空間データを総称するマイクロジオデータ<sup>5)</sup>の活用に期待が寄せられている。著者らは先行研究<sup>6)</sup>において、都市施設の変更に伴う滞在人口変化を推定するために、札幌市を対象として、モバイル空間統計<sup>7)</sup>によるメッシュ別滞在人口を出力変数、都市計画基礎調査から得られる全34種類の建物用途別床面積を入力変数として、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ネットワークを適用することにより、メッシュごとの対象時間帯の入込人口の推計を試みた。また併せて、より詳細なメッシュでの滞在人口を推計するダウンスケーリング手法を開発した。しかし、この研究はメッシュ間の移動量の推定は対象としてい

かった．OD交通量推定にマイクロデータを用いた事例としては、今井ら<sup>8)</sup>の研究が挙げられる．この研究では、モバイル空間統計の人口流動統計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交通量の推計が行われている．しかし都市内の建物データは扱われておらず、都市施設整備に伴う交通量の変化については評価することができなかった．

本研究では、建物用途別延床面積と滞在人口、移動時間からメッシュ間OD交通量を推定する手法を開発する．これにより、交通整備による移動時間短縮や建物用途の変更・誘導がOD交通量に与える影響を簡便に予測することができる．そのために本研究では、モバイル空間統計に、現在入手可能なオープンデータや商用データを組み合わせたデータセットを構築した上で、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を確認し、このデータセットを組み替え、それにニューラルネットワークを適用することで、精度の高いモデルの構築を試みた．

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．

## 2. データの概要

本研究では、建物用途別延床面積や滞在人口、移動時間を考慮したOD交通量の分析を行う．使用したデータは、都市計画基礎調査とモバイル空間統計であり、移動時間については、駅すばあとWebサービス<sup>9)</sup>およびGoogle Maps Direction API<sup>10)</sup>を用いて収集した．これらの概要を以下に示す．

### (1) 都市計画基礎調査データ(建物用途別延床面積)

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．

本研究では 2019 年度の札幌市「都市計画基礎調査」を GIS 上で利用できるデータを使用した．このデータには、個々の建物に関する位置情報や建物用途、階数、構造、延床面積などが収録されている．本研究では、この内、建物用途(表-1)と延床面積のデータを用いる．全 435,728 棟の建物用途(大分類)で色分けして可視化した結果を図-1 に示す．都心部は商業系が集まっており、郊外部は住宅系が支配的であるが商業系や工業系も分布しているエリアがあることがわかる．

表-1 都市計画基礎調査の建物用途区分

大分類	小分類
商業	業務施設、集合販売施設、専用店舗施設、宿泊施設、興業施設、風俗営業施設、遊戯施設、スポーツ施設、地方国家施設、自治体施設
住宅	専用住宅、共同住宅、一般店舗併用住宅、飲食店併用住宅、作業所併用住宅、事務所併用住宅
文教厚生	教育施設、研究施設、文化施設、宗教施設、記念施設、医療施設、運動施設、社会福祉施設、厚生施設
工業	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、供給処理施設、運輸倉庫施設、通信施設
その他	漁業施設、農業施設

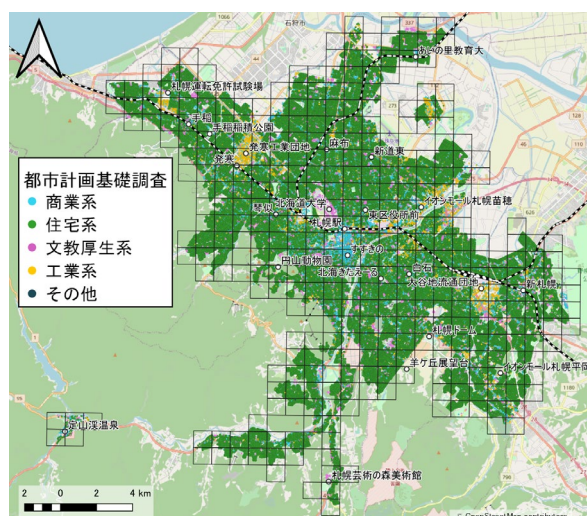


図-1 大分類による建物用途別の分布

### (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から拡大推計された人口統計であり、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において、情報の取得が可能である．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報告が複数挙げられている<sup>11), 12), 13)</sup>．本研究においては、各エリアに分布している滞在人口と、ある2つのエリア間の流動人口(OD 交通量)を活用した．これらのデータに関する属性等を表-2 に示す．プライバシーの観点から秘匿処理によってデータ量の減少を避けるため、滞在人口と流動人口ともに年齢と居住地の属性は考慮せず、OD 交通量に関しては 7 時から 9 時台の 3 時間について、完結するトリップにおいて集計することとした．また、OD交通量に関しての判定は、更に半径 1 km 以上の移動で該当メッシュ内に1時間以上滞在した場合に集計される．つまり、あるトリップに関して滞在時間が

表-2 モバイル空間統計の集計条件

	滞在人口	OD交通量
日付	2020年 1月20～24日	2019年 10月17日
曜日	平日	木曜日
時刻	7, 8, 9時	7～9時台
集計規模	500mメッシュ	1kmメッシュ
集計方法	各時間の平均値	総和
単位	人	trip

1時間未満である場合、そのトリップは立ち寄りとなり、本データのOD交通量としては集計対象外となる。メッシュペアは発生・集中エリアともに札幌市内のみを抽出し、その結果としてメッシュ間の移動は4,863ペア(合計361,527トリップ)となった。図-2, 3から札幌の中心部に流れる移動が多いことがわかる。また、札幌市内の該当するメッシュに対して滞在人口は8時で1,455,171人(図-4)、3時で1,451,582人となった。札幌市における平成27年度の国勢調査<sup>14)</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%以上の捕捉率であるが、説明変数として扱う滞在人口には札幌市外の居住地人数も集計されているため、推計としてはやや少なく見積もられていると考えられる。

### (3) 移動時間の推計

本研究ではメッシュの中心点間を移動すると仮定し、公共交通と自動車を使った2通りの移動時間を求める。

初めに公共交通を用いた移動については、駅すばあとWebサービスを活用した。これは経路探索や交通費積算などの機能をスマホアプリやWebサイトなどに導入するような活用が想定されているAPIである。本研究においては、モバイル空間統計のOD交通量の集計日時と合わせるとともに、通勤通学におけるピークを考慮し、9時には移動が完了していることを仮定し、2019年10月17日8時30分に到着を条件とした移動時間を抽出した。公共交通の中には、JRをはじめ、札幌市営地下鉄、民間バスのダイヤも組み込まれている。

次に自動車での移動時間については、Google Maps Direction API を活用した。これ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。また、出発日時については過去の時空間で取得することができないため、交通状況を考慮しない移動時間を説明変数として扱った。

本研究では、これらの両サービスを用いてメッシュの中心点における緯度経度座標から移動時間を算出し、説明変数として用いた。

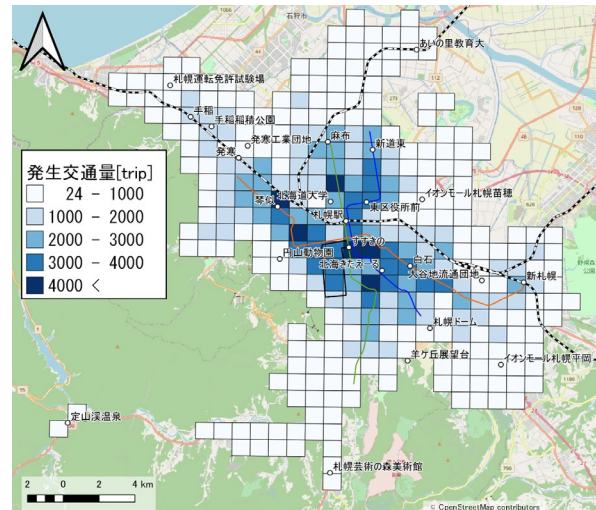


図-2 発生交通量の分布

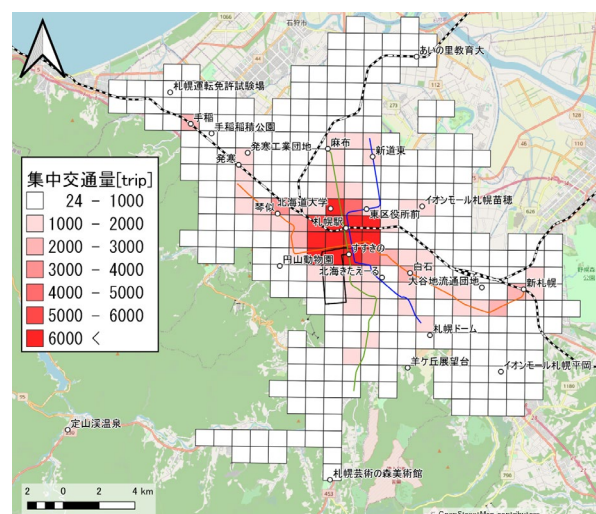


図-3 集中交通量の分布

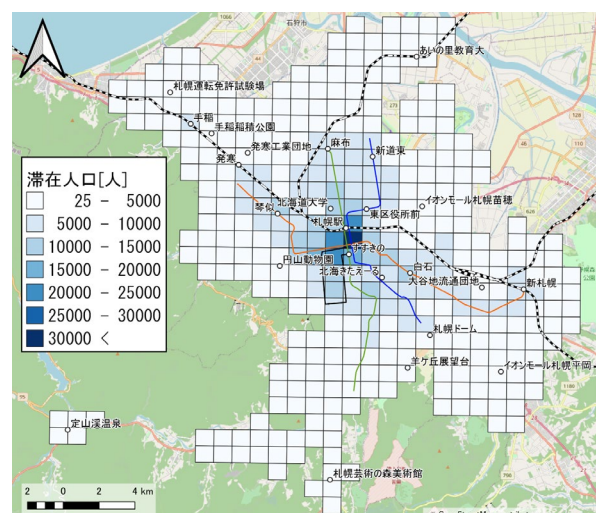


図-4 平日 8時の平均滞在人口



表-3 各変数の代表値

	業務施設	集合販売施設	専用店舗施設	宿泊施設	興業施設	風俗営業施設	遊技施設	スポーツ施設	地方国家施設	自治体施設
平均値	20131.7	4731.9	13297.3	5435.5	167.5	606.2	2777.8	513.8	947.0	1627.6
標準偏差	85137.1	26847.7	30181.8	32726.3	2273.0	10942.3	7727.5	2443.2	7395.6	9278.7
中央値	4245.0	0.0	5202.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	専用住宅	共同住宅	一般店舗併用住宅	飲食店併用住宅	事務所併用住宅	作業所併用住宅	教育施設	研究施設	文化施設	宗教施設
平均値	105356	128435.1	3301.9	216.8	2806.7	272.7	12811.8	370.3	1657.6	1572.3
標準偏差	70806.1	172564.4	4048.8	347.8	3087.6	404.5	29798.8	2370.7	7428.3	3071.7
中央値	101693.0	52125.0	1958.0	97.0	1754.0	69.0	7757.0	0.0	435.0	365.0
	記念施設	医療施設	運動施設	社会福祉施設	厚生施設	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	供給処理施設
平均値	19.6	7718.2	563.5	4927.1	153.3	28.9	4453.5	1557.4	406.1	1829.6
標準偏差	267.8	14467.9	2410.0	5873.2	1068.7	182	15257.2	3964.2	1146.9	8247.2
中央値	0.0	2465.0	0.0	2727.0	0.0	0.0	49.0	0.0	0.0	354.0
	運輸倉庫施設	通信施設	漁業施設	農業施設	滞在人口7時	滞在人口8時	滞在人口9時	公共交通移動時間	自動車移動時間	OD交通量
平均値	9722.2	1283.5	1.0	138.4	3793.1	3779.9	3759.7	22.6	11.5	74.3
標準偏差	24902.9	6765.7	19.0	302.7	3328.3	3847.8	4550.9	11.8	8.5	79.5
中央値	1173.0	0.0	0.0	0.0	2980.8	2797.8	2776.8	20.0	8.8	49.0

#### (4) データセットの概要

本研究のデータセット作成において、各変数の代表値(平均値、標準偏差、中央値)を表-3で示す。なお、建物用途別延床面積と滞在人口は369メッシュ分、移動時間は4,836ペア分が集計対象である代表値である。また、中央値が0の場合は、その用途施設の建物がそもそも少ないことや建物の空間分布に偏りがあることがいえる。これらの変数を用いて、表-4に示すように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、その間の移動時間とし、目的変数はそれに対応する流動人口量をOD交通量としてモデルを構築する。モデル式の汎化性を確認するため、分析の対象となる移動の4,863ペアを無作為に10分割することで、9/10を学習データ、1/10をテストデータとして、それぞれ入れ替えて学習させる。また、その学習データの内、10%を検証データとして用い、それに対する損失関数が小さくなるよう学習を進める。いわゆる、10分割交差検証(10-fold cross-validation)でモデルの汎化性能を確認する。

なお、各データ $a$ のスケールを揃える必要があるため、式(1)を用いて各変数別に、その平均値と標準偏差を用いて正規化を行う。したがって、分析にはその値 $x$ を用いることとする。

$$x = \frac{a - \bar{a}}{a_{std}} \quad (1)$$

$a_{std}$  : その変数における標準偏差

$\bar{a}$  : その変数における平均値

表-4 各変数の条件

変数	変数名	種類	単位
説明変数	出発地 建物用途別延床面積	34	m <sup>2</sup>
	到着地 建物用途別延床面積	34	m <sup>2</sup>
	出発地 滞在人口 (7,8,9時)	3	人
	到着地 滞在人口 (7,8,9時)	3	人
	自動車移動時間	1	分
	公共交通移動時間	1	分
目的変数	メッシュ間 OD交通量	1	trip

### 3. データの概要

#### (1) ランダムフォレストの概要

ランダムフォレスト(以下、RF : Random Forest)はアンサンブル学習における機械学習アルゴリズムの一種である。RFでは、ランダムに説明変数を抽出し、複数の決定木から出力結果を多数決や平均化することで、判別や回帰を行うことができる<sup>15, 16</sup>。また、目的変数に対する各説明変数の寄与度を算出することが可能であり、そのアプローチを用いた分析事例も見受けられる<sup>17</sup>。

本研究では、各建物用途の延床面積がOD交通量に与える影響について、RFによって得られる寄与度から評価した。また、その寄与度を、次節で説明するニューラルネットワークによる予測を行う上での変数選択の指標とした。決定木の本数や深さなどのハイパーパラメータについては、平方二乗誤差(以下、MSE : Mean Square Error)を基準とし、汎化性能が最も高い組み合わせとなるようにチューニングした。なお、使用した計算ライブラリは、Pythonのsklearnに含まれるGridSearchCVである。

## (2) 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(以下、NN : Neural Network)とは、機械学習の代表的なモデルの一つであり、非線形なデータにも柔軟に対応することが可能である。近年では、パターン認識や非線形の回帰分析などのあらゆる問題に適用されている<sup>18)</sup>。モデルの構成は、図-5に示すように、入力層、中間層、出力層から成る。NNは、高い予測精度が期待できるものの、中間層の複雑な構造から、説明変数の寄与度については判断することができない、いわゆるブラックボックスであることが欠点として挙げられる。各ニューロンにおける計算は以下のようになる。図-5の $u$ は直前の層の出力値である $x_i$ と重み係数 $w_i$ の積の総和であり、バイアス $\theta$ を用いて式(2)のように示す。

$$u = \theta + \sum_{i=1}^K w_i x_i \quad (2)$$

NNの学習に関しては、一般に、中間層の最適な構成はデータセットにより異なり、明確な定め方は確立されていない。そこで、本研究では、様々な中間層の構成において、学習データと検証データの誤差曲線を確認しながら決定した。活性化関数にはRelu関数を採用し、中間層層数は3層と固定した。連続量であるOD交通量を推定する回帰モデルを構成するため、出力層の活性化関数は恒等関数にした。最適化アルゴリズムにはAdam<sup>19)</sup>を採用し、出力信号と教師信号のMSEが最小となるようにパラメータの最適化を行った。さらに、過学習を防ぐ対策として、ドロップアウト層<sup>20)</sup>を追加し、繰り返し学習回数が20回を超えてもMSEが降下しない場合に学習を打ち切ることとした。チューニングするハイパーパラメータは、各中間層のニューロン数とドロップアウト率で行い、MSEで比較しモデルを決定する。ニューロン数のパターンは、 $2^n$  (ただし、 $n \in (2, 3, 4, \dots, 8)$ )の7種類、ドロップアウト率のパターンは0, 0.1, 0.3, 0.5の4種類とし、これらの組み合わせで、計28パターンで計算を行った。

## 4. 分析結果

本章では、RFによって得られる各説明変数の寄与度について考察し、さらに、寄与度の高い変数を用いたNNモデルによるOD交通量の推定結果を示す。

### (1) 各説明変数の寄与度

RFを用いて説明変数の寄与度を算出し、降順に並べた結果を図-6に示す。OD交通量に最も寄与する変数は、

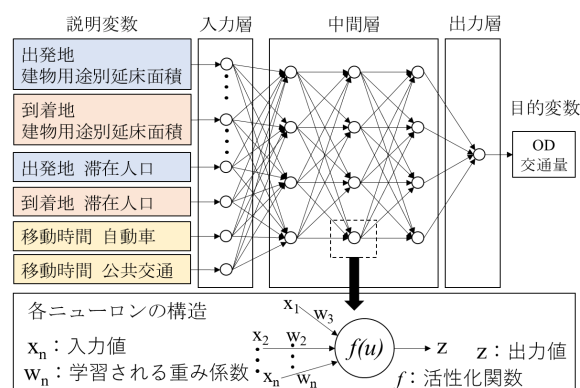


図-5 ニューラルネットワークの構成

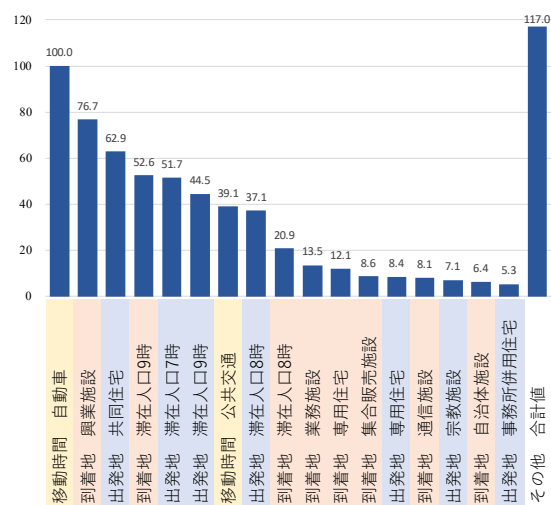


図-6 RFによる説明変数の寄与度

「移動時間 自動車」であることがわかった。さらに、移動時間や滞在人口のみならず、建物用途別延床面積の変数においても、映画館や劇場、競馬場などが含まれる商業施設や共同住宅、業務施設が上位に入っており、OD交通量を予測する重要な説明変数であることが確認できる。一方で、ここで挙げた説明変数以外については、過大な変動がない限り、市町村内のOD交通量に与える影響は小さいことが考えられる。

### (2) OD交通量の推定

次に、全説明変数を用いた場合と、RFで寄与度の高い説明変数だけを用いた場合でそれぞれNNモデルを構築し、比較検討を行った。後者については、最も寄与度の高い「移動時間 自動車」を100とした場合の5以上となる17の説明変数を用いる。全説明変数を用いたモデルAと、上記の17の説明変数を用いたモデルBのそれぞれで交差検証を行い、MSE (平方二乗誤差)が最小となるモデルを決定した。両モデルの予測精度を表-5に示す。モデルBにおいて、中間層が3層で1層当たりニューロン数

表-5 モデルAとBの予測精度

モデル	説明変数	各層のニューロン数	ドロップ率	RMSE	$R^2$
A	76	32	0	37.1	0.78
B	17	64	0.1	34.8	0.81

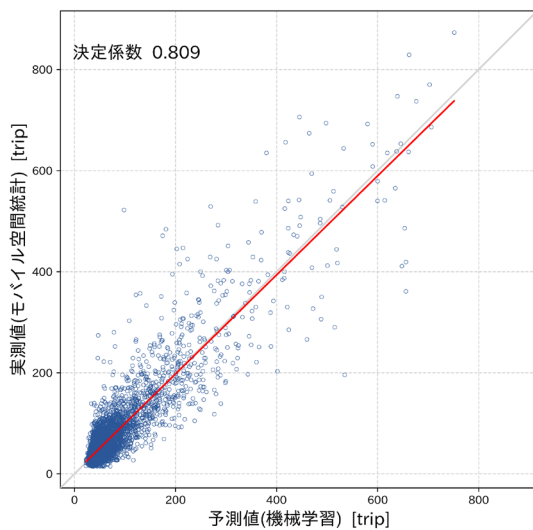


図-7 NNモデルの予測値と実測値の関係

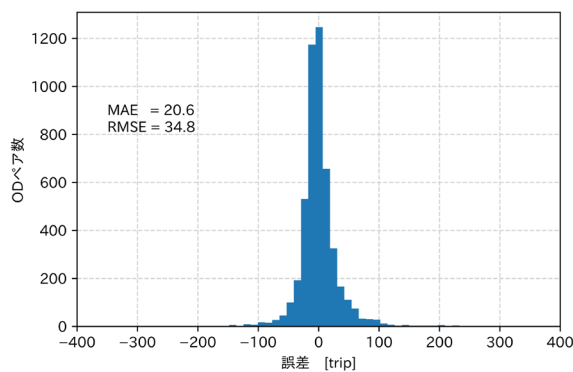


図-8 メッシュ別予測結果の誤差分布

は64のとき、 $MSE$ が最小となった。両者の $RMSE$ (二乗平均平方根誤差)および $R^2$ には大きな差はないが、過学習や学習時間の点から、RFで説明変数を減らしたモデルBの方が扱いやすい。

### (3) 発生・集中交通量の誤差

トリップ毎の予測値と実測値の関係を図-7で示すように、OD交通量の推定としては $RMSE=34.8$ 、 $R^2=0.81$ となり、実測値のモバイル空間統計と予測値のNNモデルは強い相関が現れた。OD交通量の予測値と実測値の誤差を調べたところ、図-8に示すように、誤差は、0を平均とした正規分布に近い分布を示した。そこで、メッシュ毎のOD交通量の予測値を出発地と到着地の合計として求め、発生と集中交通量のそれぞれについて予測精度を調べた。図-9、10に示すように、予測値と実測値の間には、非常

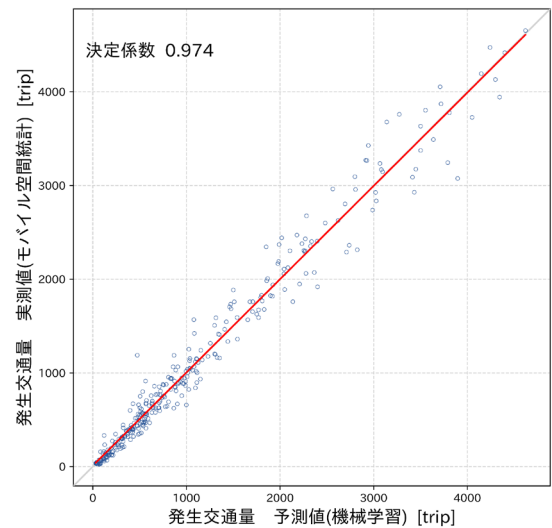


図-9 メッシュ別発生交通量の予測結果

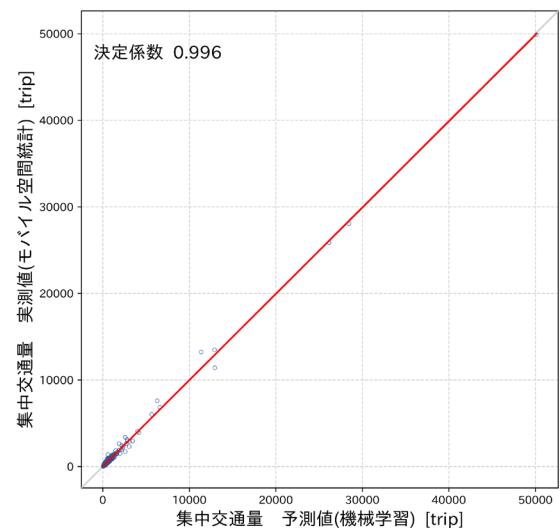


図-10 メッシュ別集中交通量の予測結果

に強い相関が見られた。したがって、発生交通量と集中交通量についても、高い精度で予測できることがわかった。

## 5. 分析結果

本研究では、高頻度で得られる各種マイクロジオデータにより得られるメッシュ内の建物用途別延床面積と滞在人口、メッシュ間の移動時間から、機械学習モデルの一種であるNNを用いてメッシュ間OD交通量の推定を行った。構築したモデルにより、建物の用途や規模の変更、時間帯別の滞在人口の変化、また、自動車交通や公共交通の移動時間の変化等によってOD交通量を推定することが可能となった。モデルから得られた知見を以下に示す。

- ・RFにより、滞在人口とメッシュ間の時間関係に加え、建物用途別延床面積がOD交通量に強い影響を与えていることが明らかになった。
- ・推定したOD交通量の予測結果について誤差は正規分布に近く、OD交通量をメッシュ毎に総和した発生・集中交通量は高い精度を示した。
- ・RFにより、説明変数選択を行ったモデルBが、より推定精度が高くなった。しかし、多くの建物用途別延床面積による変数が棄却されているため、棄却された変数が過度に変化した場合、将来的なOD交通量を表現できない可能性がある。

上記を踏まえ、説明変数の選択や組み方、OD交通量の推計誤差についてどれほど許容するのかは、課題である。今後の展望としては、他の都市または時間帯における精度検証や、プローブデータの活用等により実際の交通流を加味した移動時間を用いることで、より実用的なモデルの構築を進めたい。

**謝辞：**本研究はJSPS科研費20K04730、及び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。

#### 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：2020年6月19日)
- 2) 浅田拓海，田中優太，Woramol CHAOWARAT，有村幹治：都市計画基礎調査データの建物属性情報を用いた住宅寿命の要因分析，土木学会論文集 D3 (土木計画学)，Vol. 72，No. 5 (土木計画学研究・論文集第33巻)，pp. I\_269-I\_275，2016。
- 3) Wegener, M.; Overview of land-use transportation models, *The 8th International Conference on Computers in Urban Planning and Urban Management*, pp. 20-40, 2003.
- 4) 宮本和明，V. Vichiensan，杉木直，北詰恵一：先進諸国における都市モデルの適用状況—土地利用モデルを中心として—，土木計画学研究・講演集，No. 33，CD-ROM，2006。
- 5) マイクロジオデータ研究会，<http://microgeodata.jp/> (閲覧日：2020年6月19日)
- 6) 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，pp. I\_483-I\_491，2017。
- 7) 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：2020年6月19日)
- 8) 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，Vol. 53，pp. 619-627，2016。
- 9) 駅すばあと Web サービス フリープラン：<https://ekiworld.net/service/lp/websevice/> (閲覧日：2020年6月19日)
- 10) Developer Guide | Directions API | Google Developers：<https://developers.google.com/maps/documentation/directions/intro> (閲覧日：2020年6月19日)
- 11) 川上陸，宇野伸宏，中村俊之：モバイル空間統計のデータ特性を考慮した OD 推計手法：京都観光地間流動におけるケーススタディ，土木学会論文集 D3 (土木計画学)，Vol. 75，No. 6，pp. I\_379-I\_391，2020。
- 12) 菊池雅彦，岩館慶多，羽藤英二，茂木渉，森尾淳：全国 PT 調査データと携帯電話基地局データを用いた地方都市での OD 表の実務的推計，土木学会論文集 D3 (土木計画学)，Vol. 74，No. 5，pp. I\_677-I\_691，2018。
- 13) 松原徳和：モバイル空間統計による動態人口把握：観光防災，帰宅困難者対策の観点から，情報管理，60巻，7号，pp. 493-501，2017。
- 14) 札幌市国勢調査，平成 27 年国勢調査，第 8 表，<http://www.city.sapporo.jp/toukei/tokusyuu/documents/27kokutyoo8.pdf> (閲覧日：2020年6月19日)
- 15) Leo, B.: Random forest, *Machine Learning*, Vol. 45, No. 1, pp. 5-32, 2001.
- 16) 波部斉：ランダムフォレスト，研究報告コンピュータビジョンとイメージメディア(CVIM)，31号，pp. 1-8，2012。
- 17) 内閣府：付注 2-1 ランダムフォレスト・決定木分類について，[https://www5.cao.go.jp/keizai3/2017/0118nk/img/n17\\_5\\_b\\_2\\_01c.html](https://www5.cao.go.jp/keizai3/2017/0118nk/img/n17_5_b_2_01c.html) (閲覧日：2020年6月19日)
- 18) 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，pp. 22-33，2017。
- 19) Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980, 2014.
- 20) Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting, *The Journal of Machine Learning Research*, Vol. 15, pp. 1929-1958, 2014.

(Received June 19, 2020)

(Accepted December 28, 2020)

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

Kota OKUMURA, Mikiharu ARIMURA and Takumi ASADA

In the deployment of compact plus network policy in Japan, it is necessary to promote urban facility location guidance and arrangement of traffic facilities while observing and predicting changes in human flow due to changes in building location conditions and moving environments more easily. The purpose of this study is to estimate the OD traffic volume between meshes using various types of micro-geo data that have been developed in recent years. In this study, the total floor area data for each mesh by the basic urban planning survey, the mesh staying population data and the OD traffic data between meshes by the Mobile Space Statistics which is based on the operation data of the mobile phone network, and the travel time data by various open data were acquired. Next, we used the Random Forest method to check the contribution of each explanatory variable. In addition, we tried to verify the estimation accuracy by estimating the OD traffic volume using a neural network.