

# 機械学習による建物用途を考慮した メッシュ入込人口のダウンスケーリング推計

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 もの環境系領域 (〒050-8585 室蘭市水元町27-1)

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

<sup>3</sup>正会員 室蘭工業大学助教 大学院工学研究科 もの環境系領域 (〒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い、地域の賑わいを評価する必要がある。本研究の目的は、近年整備が進む都市地域におけるマイクロデータを用いて、メッシュ毎の建物種類別延床面積から、そこに訪れる人口を時間別に推定することにある。さらに、その対象エリアに対してより細微なエリアでの人口を推計するダウンスケーリングを試みた。建物情報は建物ゼンリンポイントデータ、滞在人口は携帯電話網の運用データから集計されるモバイル空間統計を用いた。推定モデルの構築にはニューラルネットワークを使用し 4 分割の交差検証を行うことで、予測値の精度を確認した。その値に関して、メッシュ毎に単純に延床面積按分した場合との誤差を確認することで本研究の有用性を示す。

**Key Words :** *Neural Network, Building points data, 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、人口の減少や高齢化に伴い、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。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、本研究で対象としている室蘭市においては、人口減少が顕著な問題として挙げられている<sup>1)</sup>。また地方都市では、病院の統廃合などが検討されている実例が見受けられる。また、市内における大型商業施設の移転を検討されており、今後、人の賑わいに大きな変化をもたらすことが予想される<sup>2)</sup>。このように多くの地方都市では、人口減少や建物の老朽化に対応しつつ、より暮らしやすい都市施設の配置が必要とされている。

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、マイクロデータの適用が期待されている。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データを用いた研究事例が複数報告されている。小澤ら<sup>3)</sup>は建物ポイントデータの商業と医療の立地を活用して、都市における核間公共交通水準を加味した拠点の評価を全国の79都市で分析を行った。松中ら<sup>4)</sup>は都市中心部の歩行者空間の賑わいについて公共交通の到達圏人口との関連をGoogleEarthのストリートビューを用いて分析した。著者らによる先行研究<sup>5)</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た。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、この既往研究では、都市計画基礎調査のデータを用いていたことから、対象エリアは市街化区域に限られていた。また平日・休日における建物用途毎の入込人口に与えるインパクトの評価は行われていなかった。

本研究の目的は、各エリアの建物分類別延床面積から、任意の時間における入込人口を推定し、既往研究よりも空間分解能を向上させることにある。単純な建物延床面積の按分による入込人口推定と比較し、本手法の有効性を確認する。また、携帯電話網の位置情報から取得した人口統計値であっても、プライバシー保護の観点から秘匿処理され、従来把握することが困難であった、より細かなエリアの入込人口についても推定する。本研究では、ニューラルネットワークを適用し、より高い予測精度のモデルを構築する。これにより、建物の誘致や用途の変更から入込人口の変化、所謂その地域の賑わい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や都市計画の政策を補助するもの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と精度の検証方法について示す。第4章では4次メッシュにおける精度と5次メッシュへダウンスケーリングした結果を示す。第5章では本研究の結論を述べる。

## 2. 使用データ

本研究では、表-1のようなデータセットを作成し、建物分類別延床面積と滞在人口判定の時間、曜日と特異的なメッシュにおけるダミー変数の18種類の説明変数から入込人口の推定分析を行う。それぞれのデータについてはゼンリン住宅データ、モバイル空間統計を用いた。使用したデータの概要を以下に示す。

### (1) 建物ポイントデータ

ゼンリンが提供する建物ポイントデータ<sup>6)</sup>は全国約3,800万棟の建物情報を収録しており、項目では個々の建物に関する分類、階数、延床面積、空き家などがまとめられている。

本研究では室蘭市と登別市全域の計52,358棟についての建物ポイントデータ(2019年)を使用した。各市と建物種類別延床面積の内訳を表-2に示す。その全ての建物について図-1でGIS上に表示させた。次に種類別に4次メッシュ(約500m四方)毎と5次メッシュ(約250m四方)の延床面積を集計した。

### (2) 滞在人口

滞在人口にはモバイル空間統計<sup>7)</sup>と呼ばれる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を活用した。法人を除く契約者情報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

表-1 変数状況

| 変数   | 変数名       | 種類 | 単位             |
|------|-----------|----|----------------|
| 説明変数 | 建物分類別延床面積 | 9  | m <sup>2</sup> |
|      | 時間断面      | 3  | 時              |
|      | 平日曜日ダミー   | 5  | 無次元            |
|      | 室蘭工業大学ダミー | 1  | 無次元            |
| 目的変数 | 入込人口      | 1  | 人              |

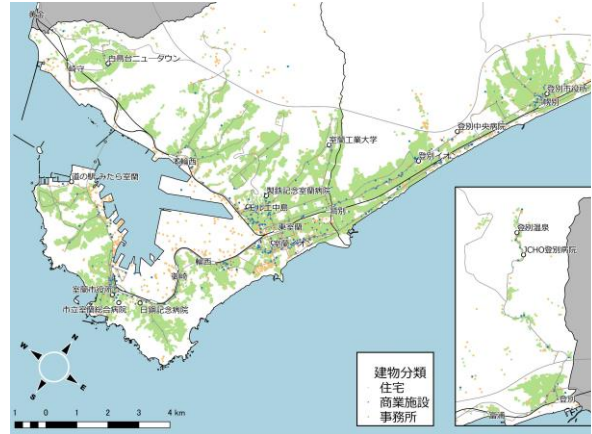


図-1 建物ポイントデータの分布

表-2 建物分類

| ゼンリンデータの分類名   | 解析分類 |
|---|------|
| 個人の家屋, 事業所兼住宅   | 戸建住宅 |
| マンション, アパート, 団地, 寮・社宅, 住宅系建物  | 集合住宅 |
| 飲食, 物販(食品), 物販(衣料), 物販(日用雑貨), サービス(レンタル), サービス(冠婚葬祭), サービス(生活関連), サービス(自動車), サービス(その他), 量販店, 商業複合系建物, 商業系建物 | 商業施設 |
| ホテル・旅館  | 宿泊施設 |
| スポーツ施設, 娯楽, 公共, 宗教関連  | 文化施設 |
| 教育  | 教育施設 |
| 医療・福祉   | 医療施設 |
| 金融・保険, 不動産, インフラ, 専門職, 宅配・引越・郵便, 運輸, 建設・設備, 自動車関連, 協同組合, 一般業, オフィス複合系建物, オフィス系建物                            | 事務所  |
| その他   | その他  |

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、帰宅困難者数の推計などあらゆる分野から報告が挙げられる。

滞在人口で用いたデータは滞在者の居住地別4次メッシュ分布となる。表-3で示すように、室蘭市内の該当するメッシュ滞在人口と同市が提供している人口統計表<sup>8)</sup>を比較すると5%ほどモバイル空間統計の方が多く人口を把握していることが分かった。これについて本研究で

活用したモバイル空間統計は居住地設定が北海道全域であることや室蘭市に在住しながらも住民票などの手続きを行っていないことなどがこの差分と考えられる。

(3) 入込人口の定義

本研究において、夜間人口に対して各メッシュかつ特定の時間に流出入人口を評価するために入込人口を定義する。具体的には、表-4で示すデータを活用し、各メッシュの3:00における滞在人口から対象の時間断面(10:00, 13:00, 16:00)の滞在人口との差分を求めた。数値が正の値をとると、メッシュに人口が流れてきたことを意味する。例として2016年11月8日13:00の4次メッシュ入込人口を図-2に示す。

3. 分析方法

本章では、入込人口の推定とダウンスケーリング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-3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>10)11)</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500回とし、中間層の1層あたりのニューロン数は50, 100, 150で行った。

従来の重回帰分析では、多重共線性の影響を考慮する必要があり点である。今回のデータセットにおいても、重回帰分析を適用する場合は説明変数間の相関が高い分類別延床面積に関しては片方の変数を棄却する必要がある。一方、NNモデルを扱った非線形的予測には説明変数間の相関を考える必要がないことから、重回帰分析と比較しても、全ての分類別延床面積を扱うことができる点がNNの利点としてある。モデル式の汎化性を確認するため、解析の対象となる入込人口の7,478パターンを

表-3 人口統計の把握

| データ      | 時期                            | 集計対象居住地 | 人口    |
|----------|-------------------------------|---------|-------|
| 年齢別人口統計表 | 2016年10月末                     | 室蘭市     | 68650 |
| モバイル空間統計 | 2016年11月7~13日<br>3:00 (7日間平均) | 北海道     | 71846 |

表-4 解析対象

| 期間                | 時間                     | 規模     | 居住地 | 年代            |
|-------------------|------------------------|--------|-----|---------------|
| 2016年<br>11月7~11日 | 10:00, 13:00,<br>16:00 | 4次メッシュ | 北海道 | 15~79歳の<br>総和 |

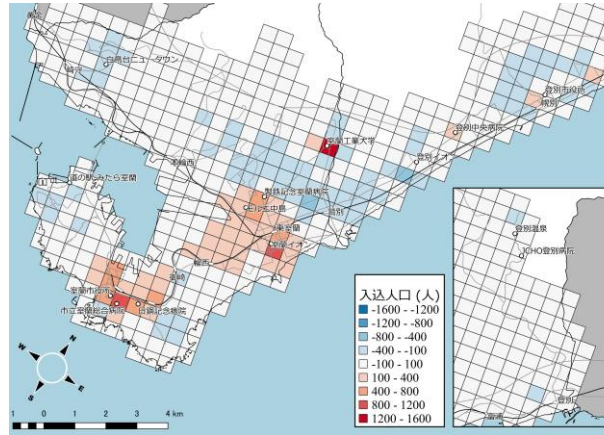


図-2 2016年11月8日(火)13時の入込人口

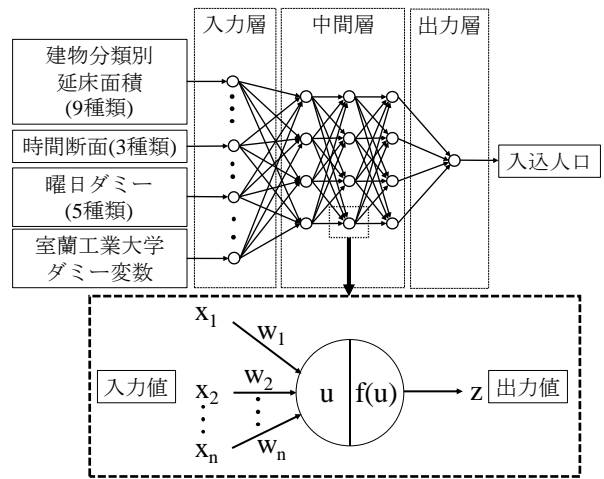


図-3 ニューラルネットワークの構造

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。そのテストデータとなる該当の日時かつメッシュにおいては5次メッシュで集計したエリアの情報を適用することで図-4のようにダウンスケーリングを行った。

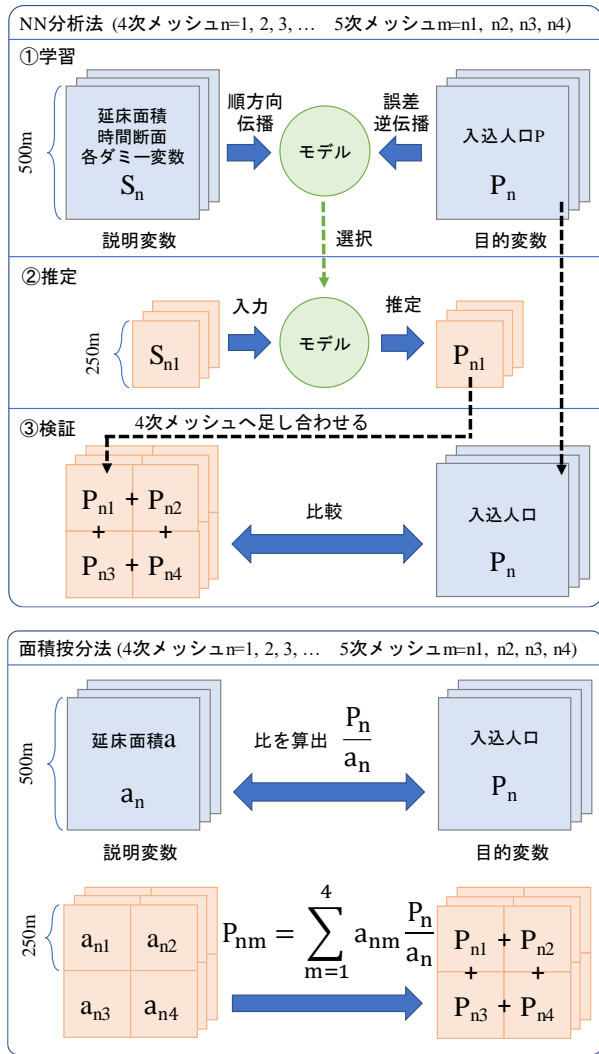


図-4 ダウンスケーリング方法

(2) データセットの概要

本研究のデータセットは表-1で示した内容である。説明変数は、該当する各メッシュの情報、建物分類別延床面積と時間断面、室蘭工業大学が立地されるメッシュへダミー変数を付与した。目的変数はそれに対応する入居人口としてモデルを構築する。また、延床面積の分類に関しては、建物ポイントデータの分類を広義にとらえた解析用分類において作成した。なお、各項目の説明変数 $a$ は式(3)を用いて正規化を行い、分析にはその値 $x(0 \leq x \leq 1)$ を用いる。

$$x = \frac{a - a_{min}}{a_{max} - a_{min}} \quad (3)$$

表-5 解析対象

| ニューロン数 | R <sup>2</sup> | MAE   | RMSE  |
|--------|----------------|-------|-------|
| 50     | 0.769          | 63.57 | 88.99 |
| 100    | 0.759          | 64.16 | 90.24 |
| 150    | 0.785          | 61.75 | 85.39 |

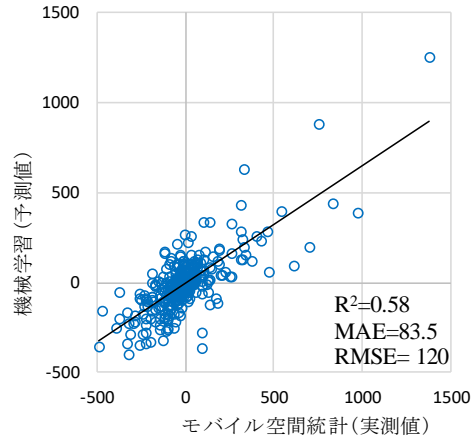


図-5 ダウンスケーリング適用後の比較

4. 分析結果

本章ではNNによる入居人口の推定結果を示し、単純に延床面積と按分した数値と比較する。

(1) NNのパラメータチューニング

交差検証を行い、最もRMSE(平方二乗誤差)が小さいモデルを選択した。その結果、表-5で示すように中間層の構成は、3層で、1層あたりのニューロン数は150、活性化関数はReLUのときにRMSE=85.39と最小となった。従って、以後の分析結果この中間層構成に基づき、例として2016年11月8日(火)13時の入居人口を示す。

(2) ダウンスケーリングの推定結果

推定結果は図-5にモバイル空間統計とNNモデルで算出した散布図を示す。5次メッシュにダウンスケーリングを行うことで、元の4次メッシュ入居人口と比較すると統計的指標の減少が確認される。しかしながら、図-6で示すように病院や大型商業施設には人の集まりは確認できる。

5. 結論

本研究は、建物用途別延床面積と移動断面、特異的な建物によるダミー変数から入居人口の推定を行った。今回作成したモデルから建物の用途や規模の変更により将来的な入居人口の増減を数値で算出することは可能であ



る。現在の日本では、人口減少や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やコンパクト・プラス・ネットワークの施策が起こった際、各地域の入込人口や賑わいを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。

本研究では建物ポイントデータとモバイル空間統計の2つマイクロデータと機械学習のNNを用いて、入込人口を推定するモデルを構築した。その結果として得られたことを次に示す。

- NNモデルの適用により、5次メッシュへのダウンスケーリングの有用性を示した。
- 入込人口を日付別で見た際、大きな差異がないため、非常に精度がよいモデルとなった。
- 室蘭工業大学や各病院が立地されているエリアに焦点を当てると、モバイル空間統計で増加の入込人口を示しているメッシュにおいてもダウンスケーリング後減少を示している区画があることが確認できた。
- 単純な延床面積による入込人口の按分より建物用途を考慮したNNモデルは大学や商業施設に大きく人が流れていることが推定できた。

今後の課題として、入込人口の属性を追加し、性別や年齢別等の考慮が挙げられる。また、本研究においてはある平日1週間のデータセットを作成し交差検証を行ってきた。そのデータセット内では、非常に高い予測精度のモデルを構築することが出来た。しかし、休日や人口分布や異なる都市の規模や時期などを考慮は未達である。上記を確認し、将来の都市構造を見据えた入込人口の推計を行う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。

#### 参考文献

- 1) 止まらぬ室蘭市の人口減少、年間1000～1500人ペース、室蘭民報、2020年1月26日(日)朝刊
- 2) イオン室蘭店 卸売市場跡に移転\_北海道建設新聞社 - e-kensin, 発行2019年02月15日
- 3) 小澤悠, 高見淳史, 原田昇: 都市計画マスタープランにみる多核連携型コンパクトシティの計画と現状に関する研究, 都市計画論文集, 52.1, 10-17, 2017
- 4) 松中亮治, 大庭哲治, 井手秀, 立花拓也: 地方都市における到達圏人口が都市中心部の歩行者空間の賑わいに及ぼ

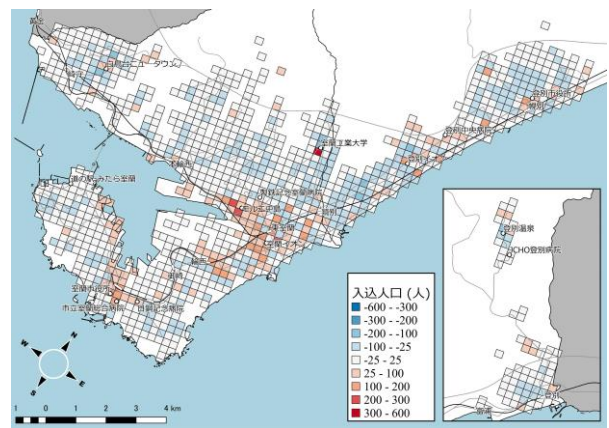


図-6 2016年11月8日(火)13時 入込人口  
NNモデル適用によるダウンスケーリング結果

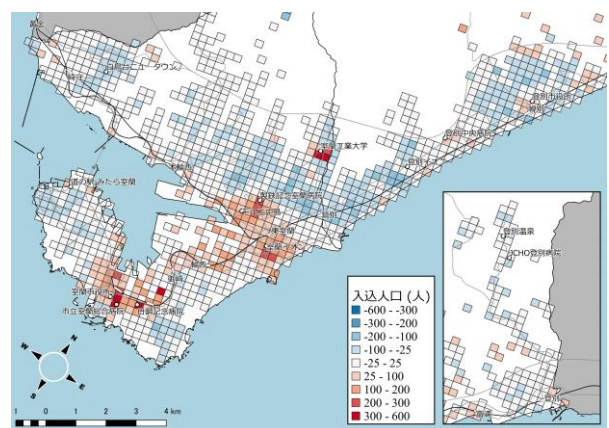


図-7 2016年11月8日(火)13時 入込人口  
延床面積按分によるダウンスケーリング結果

す影響分析, 都市計画論文集, 53(3), 573-580, 2018

- 5) 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, 1\_483-1\_491, 2017.
- 6) 建物ポイントデータ | 株式会社ゼンリン, <https://www.zenrin.co.jp/product/category/gis/contents/building-point/index.html> (2020年3月8日)
- 7) モバイル空間統計 | サービス・ソリューション | 法人のお客様 | NTTドコモ, <https://www.nttdocomo.co.jp/biz/service/spatial-statistics/function/> (閲覧日: 2020年3月8日)
- 8) 室蘭市/年齢別人口統計(平成28年10月末), <https://www.city.muroran.lg.jp/main/org3600/nenh2810.html> (閲覧日: 2019年3月9日)

- 9) 石河正寛, 松橋啓介, 有賀敏典: 建物ポイントデータの床面積補正を通じた民生部門エネルギー消費量の推計, 土木学会論文集 G, Vol.72, No.6, II\_87-II\_94, 2016.
- 10) 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.
- 11) 中谷祐介, 石崎裕大, 西田修三: 深層学習を用いた感潮河川の水質変動予測, 土木学会論文集 B1 (水工学), 73.4:I\_1141-I\_1146, 2017

(2020. 3. 8 受付)

## DOWNSCALING ESTIMATION OF POPULATION IN MESH CONSIDERING BUILDING USE BY MACHINE LEARNING

Kota OKUMURA, Mikiharu ARIMURA, Takumi ASADA