

合成人口データにおける常住地人口規模を考慮した 就業者への従業地の割当て

○岩瀬大樹 村田忠彦（関西大学） 原田拓弥（青山学院大学）

Assignment of Workplaces to Employed Persons Considering the Population Size of Residential Municipality in Synthetic Population Data

* D. Iwase and T. Murata (Kansai University) and T. Harada (Aoyama Gakuin University)

概要— 実社会を対象にした分析や社会シミュレーションにおいて、個人の行動の軌跡を推定できる汎用的に利用可能なデータを整備することで、日常生活の中で発生する社会現象モデルの構築が容易になる。国勢調査の常住人口に基づいて合成された仮想個票である合成人口データには、現状、個人の移動を規定するような地理情報は含まれていない。そこで本研究では、静的な合成人口データを動的なデータにするための端緒を開くことを目的に、公開されている公的統計データを用いて合成人口データの個々の就業者に従業地を割り当てる手法を提案する。

キーワード: 合成人口データ、従業地推定、公的統計、人口分布

1 はじめに

災害発生時における人的被害の想定や、感染症の伝播シミュレーションなど、人間行動を含む社会現象では、モデルを構築する際に人々の日常の生活行動を表現することが求められる。たとえば、感染症シミュレーションにおいては、その感染機会や感染経路を考える上で人々が所属しているコミュニティの情報や移動の状況をモデルに組み込む必要がある。このように、実社会を対象にした分析や社会シミュレーションにおいて、一人一人の行動の軌跡を推定し、汎用的に利用可能なデータを構築することで、日常生活の中で発生する社会現象モデルの構築が容易になると考えられる。

著者らは、実際の規模の人口データを用いた社会シミュレーションであるリアルスケール社会シミュレーションの研究基盤を構築することを目的に、日本全国の自治体の合成人口データを合成しており^{1,2)}、研究者向けに公開している。合成人口データとは、公開されている統計データに基づいて合成された世帯単位の仮想個票データである。合成人口データは、世帯や個人の現実の統計情報と合致するように合成されており、性別、年齢、就労している産業分類などの個人の属性に加え、個々の世帯が所在する都道府県、市区町村、町丁目、建物ごとに割り当てられた位置情報が含まれている。しかし、現状の合成人口データに含まれる地理情報は、基本的には国勢調査に基づく常住地、つまり夜間の人口を表した静的な位置情報のみであり、個人の空間的な移動を規定するような情報は含まれていない。そこで、昼間の滞在場所である従業地・通学地といった第二の地理情報を新たに加えることで、従来の静的な合成人口データを動的なデータにするための端緒を開くことができる。本研究では、その第一段階として、就業者を対象に、従業地属性を付加し、常住地と従業地をつなぐデータの構築手法を提案する。

2 関連研究

人々の移動行動を理解する試みは移動需要の予測を主な目的として多く行われている。特に近年は携帯電話・スマートフォンなどのモバイル端末の利用者数拡大に伴い、位置情報データの活用が盛んになっている。位置情報データには、代表的なものとして携帯電話

基地局データやGPSデータなどがある³⁾。携帯電話基地局データは、携帯電話が定期的に基地局と交信している履歴をもとに、人々の広域的な移動を把握するビッグデータである。移動体通信事業者は、携帯電話の交信が記録された基地局の位置や時刻から、自社ユーザーの滞在エリアや移動を把握することができる。また、携帯電話の契約者情報と紐付けることで、性別や年代、居住地といった正確な属性情報と併せて把握できる。一方、GPSデータは、GPS機能を用いて、許諾を得た特定のアプリの利用者から送信される位置情報をもとに、通信事業者やアプリ提供事業者が収集・蓄積している人流データである。基地局データに比べ取得できるサンプル数は少ないものの、人々の移動を緯度経度単位で、かつ高頻度で把握できるため、交通手段や勤務地などを推定することも可能である。小林⁴⁾はGPSの人流データから得られる時間帯および滞在時間データを元に、職業分類を考慮した居住地および通勤通学地の推定手法を構築している。また、大佛⁵⁾は複数の位置情報データを組み合わせることで都市内滞留者・移動者の時空間分布を推定する手法を提案している。

こうしたモバイル端末から得られる位置情報データは、拡大推計などの集計処理や秘匿処理が行われた上で申請内容に応じて提供され、都市・交通計画や防災計画をはじめ、観光客分析や商圈分析など、官民を問わず幅広い分野で活用されている。位置情報データには、買い物客や観光客等の非定期的な人流も反映されるため、時々刻々と変化する人口動態を即時的に分析できる点においてその有用性は明らかであるが、プライバシー保護の観点から集計可能な属性が限定的であるほか、都市部と郊外では推計値に信頼がおける空間解像度が異なる。また民間企業が保有するデータであるという特性上、原則として一般には公開されておらず、取得には金銭的なコストが発生する。中には、自由に再利用・再配布が可能なオープンデータとして公開されているものも存在する⁶⁾が、人々がどこからどこへ移動したのかといった詳細な移動履歴を得るには粒度が十分ではない。以上の理由から、現状の位置情報データは、分析の用途や対象とする地域に合わせて取得することでその効果が発揮されるものであり、汎用的に利用できる動的な合成人口データを構築する上

でこれらを利用することは難しいと判断した。

一方、公的機関が公開している統計データを用いたアプローチとして、日高ら⁷⁾は個人情報を含まない集計データを統合することで擬似的な人の移動行動データの生成を行なっている。移動行動生成モデルは、国勢調査に基づく性別、年齢層、居住地(1kmメッシュ)の属性をもつエージェントに対し、社会生活基本調査⁸⁾を元に1日の各時間帯の行為系列を与え、各行為に応じた目的地を選択するものである。同モデルを愛知県豊田市の市街地を中心とするエリアに適用し、携帯電話基地局データから作成される人口統計情報と比較検証を行っている。しかし、対象としたエリアで移動が閉じているという仮定のもとで推計が行われており、対象エリア内外の人々の往来は考慮されていない。

また木全ら⁹⁾は、防災計画での活用を想定し、同じく国勢調査や社会生活基本調査といった公的統計を主たる入力データとして用いた就業者行動モデルを構築している。同モデルにより、京阪神大都市圏における就業者分布の時刻変動を推定し、大規模地震発生時における潜在的な帰宅困難者の発生状況を分析している。なお、木全らは常住地、従業地、産業分類、職業分類、男女といった属性をもつ就業者のマイクロデータ(個票データ)を作成するにあたって、国勢調査から得られる多重クロス集計表を用い、さらに市区町村単位である常住地と従業地については、国勢調査の小地域集計および事業所・企業統計調査を用いて確率的に小地域へ割り当てている。ただし、これらの属性の組合せによる多重クロス集計表は公開されているものではないため、オーダーメイド集計¹⁰⁾を利用して取得している。オーダーメイド集計とは、政府統計の総合窓口(e-Stat)¹¹⁾に掲載されていない統計表の作成を所管省庁に委託できる制度である。しかし、これらの属性の組合せによる集計表は、個人情報の特定を避けるため、抽出詳細集計からしか得られない。抽出詳細集計とは、国勢調査における集計体系の一種であり、就業者の産業分類や職業分類に関する詳細な結果が得られる代わりに、統計的手法により抽出された一部の調査票を用いて集計したものである。そのため、集計結果は標本データから得られる推計値であり、さらに1の位を四捨五入した値となっている。母集団の概観は掴めるものの、標本誤差や丸め処理による誤差が含まれており、悉皆調査である国勢調査の全数データから得られるはずの詳細な特徴を十分に反映できていない可能性がある。

3 従業地の割当て手法

3.1 手法概要

本研究では、木全らの手法を参考に、e-Statに掲載されている公的統計データのみを用いて合成人口データの個々の就業者に従業地属性を割り当てる。就業者のマイクロデータとして、すでに常住地や性別、産業分類といった属性をもつ合成人口データを用いることで、前述の問題を回避することができる。本稿では、事例として大阪府高槻市を取り上げ、その就業者の合成人口データに対し提案手法を適用する。対象となるのは、平成27年(2015年)国勢調査¹²⁾に基づいて合成された高槻市における合成人口データ 330,437人のうち、就業者 150,529人である。なお、従業地の割当ては、従業地(市区町村)の割当てと、その従業地(市区町村)内における従業地(小地域)の割当ての二つのプロセスに分けて行う。Fig. 1はその概念図であり、高槻市を常住地とする就業者が、従業地として隣接する茨木市内のある小地域に割り当てられる様子を表している。

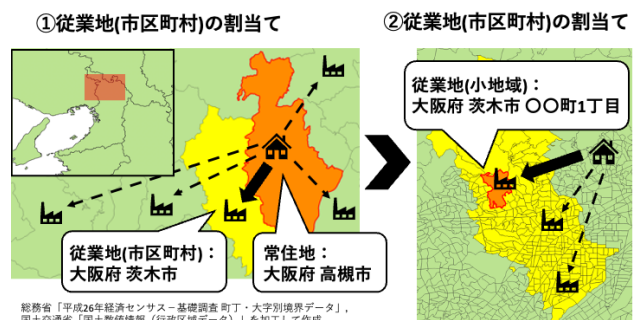


Fig. 1: 従業地の割当て手法の概要

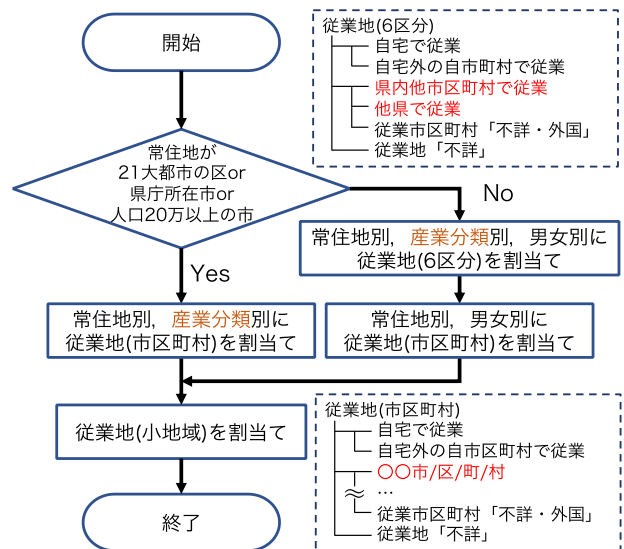


Fig. 2: 従業地の割当てフロー

Table 1: 割り当てる属性と使用する統計データ

| 割り当てる属性 | 使用する統計名 | 公開されている対象 | 表題/境界名称 | 統計表 |
|------------------|--|--|--|-----|
| 従業地(市区町村) | 平成27年国勢調査 / 従業地・通学地による人口・就業状態等集計(人口、就業者の産業(大分類)・職業(大分類)など) | 21大都市(東京都特別区部および政令指定都市)とその各区、県庁所在市、人口20万以上の市 | 常住地による従業市区町村、産業(大分類)別15歳以上就業者数 | A |
| | | 全市区町村 | 常住地又は従業地(9区分)による雇用者(3区分)、産業(大分類)、男女別15歳以上就業者数 | B |
| | | 全市町村 | 常住地による従業・通学市区町村、男女別15歳以上就業者数及び15歳以上通学者数(15歳未満通学者を含む通学者一特掲) | C |
| 従業地(小地域) | 平成26年経済センサス-基礎調査 / 町丁・大字別集計 | 全市区町村 | 経営組織(2区分)、産業(中分類)・従業者規模(6区分)別全事業所数及び男女別従業者数-市区町村、町丁・大字 | D |
| 従業地(小地域の図形中心点座標) | | | 平成26年経済センサス-基礎調査 町丁・大字別境界データ | - |

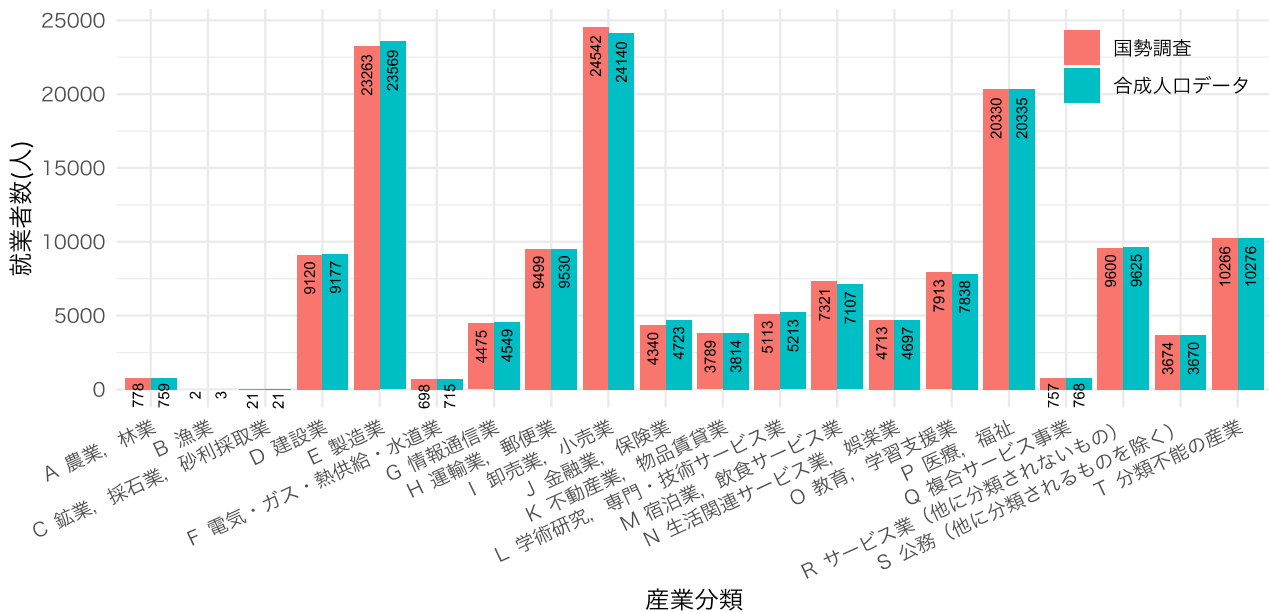


Fig. 3 : 国勢調査と合成人口データの産業分類別就業者数の差異 (平成 27 年, 高槻市)

Table 2 : (統計表 A) 産業分類別, 常驻地 (市区) による従業地 (市区町村) 別就業者数

表章項目: 15歳以上就業者数【人】 常驻地: 高槻市

| | A 農業, 林業 | B 漁業 | C 鉱業, 採石業, 砂利採取業 | D 建設業 | E 製造業 | F 電気・ガス・熱供給・水道業 | G 情報通信業 | H 運輸業, 郵便業 | I 卸売業, 小売業 | J 金融業, 保険業 | K 不動産業, 物品賃貸業 | L 学術研究, 専門・技術サービス業 | M 宿泊業, 飲食サービス業 | N 生活関連サービス業, 娯楽業 | O 教育, 学習支援業 | P 医療, 福祉 | Q 複合サービス事業 | R サービス業 (他に分類されないもの) | S 公務 (他に分類されるものを除く) |
|---------------|----------|------|------------------|-------|-------|-----------------|---------|------------|------------|------------|---------------|--------------------|----------------|------------------|-------------|----------|------------|----------------------|---------------------|
| 自宅で従業 | 423 | - | 1 | 1,233 | 540 | 1 | 241 | 161 | 1,165 | 97 | 711 | 814 | 340 | 560 | 372 | 464 | 1 | 552 | 2 |
| 自宅外の市区町村で従業 | 246 | - | 13 | 2,945 | 7,441 | 179 | 246 | 3,437 | 10,157 | 628 | 1,166 | 822 | 4,143 | 2,117 | 3,104 | 13,001 | 497 | 3,992 | 1,356 |
| 札幌市 中央区 | - | - | - | 1 | 1 | - | - | - | 1 | - | - | - | - | - | - | - | - | - | - |
| 与那国町 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 従業市区町村「不詳・外国」 | 4 | - | - | 184 | 165 | 1 | 36 | 98 | 158 | 30 | 24 | 39 | 49 | 24 | 27 | 40 | - | 83 | 8 |
| 従業地「不詳」 | 2 | - | - | 144 | 137 | 2 | 21 | 72 | 187 | 15 | 33 | 22 | 80 | 28 | 24 | 82 | 2 | 81 | 6 |

る。Fig. 2 にその具体的なフローを、割り当てる属性および使用する統計データの出典を Table 1 に示す。

一般に、事業所・企業の立地因子や立地条件はその産業分類と密接に関係していると考えられるため、従業地の割当ては可能な限り産業分類ごとに行う。なお、本研究で用いる平成 27 年国勢調査および合成人口データにおける産業分類とは、総務省が定義する日本標準産業分類 (平成 25 年) における 20 の産業分類 (大分類) に準じている。この産業分類には、「T 分類不能の産業」の項目があるが、3.4 従業地 (小地域) の割当てにおいて用いる平成 26 年経済センサス-基礎調査¹³⁾ の産業分類には存在せず、すべての事業所・企業は必ず 19 の産業分類 A~S のいずれかに分類されている。したがって、合成人口データにおいて産業分類が「T 分類不能の産業」である就業者 10,266 人を、男女別、産業分類別の構成比に応じて他の産業分類へ按分した上で、3.2 と 3.3 の従業地 (市区町村) の割当ておよび 3.4 従業地 (小地域) の割当てを実施する。

また、高槻市における、平成 27 年国勢調査と合成人

口データの就業者数を比較したところ、Fig. 3 に示すようなわずかな違いがあることがわかった。これは、合成人口データにおいて、就業者に一般労働者と短時間労働者の属性を割り当てる際に用いられた賃金構造基本統計調査が、悉皆調査ではなく標本調査であるために就業者数の推計が行われ、国勢調査の実数との差異が発生することに起因している¹⁴⁾。

3.2 従業地 (市区町村) の割当て (21 大都市の各区、県庁所在地, 人口 20 万以上の市)

国勢調査の「従業地・通学地による人口・就業状態等集計」からは、ある市区町村に常住する就業者がどの市区町村に何人勤めているかといった就業者の流出先の構成を属性ごとに得ることができる。これを常住地による従業地といい、この就業者数の分布比率に基づいて合成人口データに各従業地 (市区町村) (Fig. 1 参照) を割り当てる。前節で示したように、元々存在していたわずかな人数の差異に加え、合成人口データにおいて「T 分類不能の産業」の就業者を他の産業分

Table 3 : (統計表 B) 産業分類別, 男女別, 常住地 (市町村) による従業地 (6 区分) 別就業者数

表章項目: 15歳以上就業者数【人】 常住地: 高槻市

| | | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|---|---------------|-----------|----|----------------------|-------|-------|-------------------|-------|-------------|-------------|-------------|----------------|---------------------|-----------------|-------------------|--------------|-----------|----------|-----------------------|----------------------|
| | | 農業、 林業 | 漁業 | 鉱業、 採石業、 砂利採取業 | 建設業 | 製造業 | 電気・ガス・ 熱供給・水道業 | 情報通信業 | 運輸業、 郵便業 | 卸売業、 小売業 | 金融業、 保険業 | 不動産業、 物品賃貸業 | 学術研究、 専門・技術サービス業 | 宿泊業、 飲食サービス業 | 生活関連サービス業、 娯楽業 | 教育、 学習支援業 | 医療、 福祉 | 複合サービス事業 | サービス業 (他に分類されないもの) | 公務 (他に分類されるものを除く) |
| 男 | 自宅で従業 | 286 | - | - | 923 | 285 | - | 141 | 117 | 633 | 61 | 373 | 545 | 136 | 228 | 73 | 205 | - | 326 | 2 |
| | 自宅外の自市町村で従業 | 199 | - | 13 | 2,443 | 4,158 | 122 | 156 | 2,375 | 3,181 | 125 | 663 | 471 | 1,223 | 813 | 1,026 | 2,428 | 274 | 2,059 | 727 |
| | 県内他市区町村で従業 | 52 | - | 1 | 3,094 | 7,725 | 339 | 2,447 | 3,848 | 3,848 | 1,202 | 1,063 | 1,651 | 1,034 | 716 | 1,529 | 1,849 | 111 | 2,659 | 1,201 |
| | 他県で従業 | 29 | 2 | 6 | 974 | 4,238 | 129 | 684 | 1,032 | 1,950 | 561 | 294 | 708 | 440 | 204 | 670 | 541 | 63 | 746 | 470 |
| | 従業市区町村「不詳・外国」 | 3 | - | - | 178 | 144 | 1 | 32 | 91 | 93 | 18 | 15 | 33 | 35 | 14 | 5 | 13 | - | 59 | 7 |
| | 従業地「不詳」 | 2 | - | - | 131 | 95 | 1 | 16 | 61 | 98 | 6 | 26 | 14 | 35 | 13 | 8 | 24 | 2 | 50 | 5 |
| | 女 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |

Table 4 : (統計表 C) 男女別, 常住地 (市町村) による従業地 (市区町村) 別就業者数

表章項目: 15歳以上就業者数【人】 常住地: 高槻市

| | 男 | 女 |
|---------------|--------|--------|
| 自宅で従業 | 4,576 | 3,787 |
| 自宅外の自市町村で従業 | 23,206 | 33,937 |
| 札幌市 中央区 | 2 | 1 |
| ~ | | |
| 与那国町 | - | - |
| 従業市区町村「不詳・外国」 | 938 | 322 |
| 従業地「不詳」 | 3,861 | 2,596 |

類へ按分したことにより, 合成人口データと国勢調査から得られる統計表の就業者数は一致しない. 統計表の実数をそのまま用いると過不足が発生するため, 合成人口データの就業者数にしたがって次のように各従業地に配分する人数を調整する. ここで, 常住地 i において, n 個の属性の各属性値 (a_1, a_2, \dots, a_n) により構成される属性ベクトル \mathbf{a}_k ($k = 1, 2, \dots, K$; K は各属性の属性値数の組合せ総数) を考える. ただし, \mathbf{a}_k に適合する合成人口データの就業者数を $s(\mathbf{a}_k)$, 各従業地には $1, \dots, N$ と番号をつけ, 統計表から得られる常住地 i による従業地 j の就業者数は p_{ij} であるとする.

- p_{ij} の構成比に応じた合成人口データの人数 $q_{ij} = s(\mathbf{a}_k) \cdot p_{ij} / \sum_{j=1}^N p_{ij}$ を計算する.
- 従業地 j に $\lfloor q_{ij} \rfloor$ を配分する ($\lfloor x \rfloor$ は実数 x の整数部分を表す).
2. で配分した人数の合計が $s(\mathbf{a}_k)$ に満たない場合, $q_{ij} - \lfloor q_{ij} \rfloor$ の値が大きい従業地から順に配分する人数を 1 人ずつ加算する. ただし, $q_{ij} - \lfloor q_{ij} \rfloor$ が等しい従業地が複数存在する場合, 加算の優先順位は一様乱数によりランダムに決定される.

Table 1 における統計表 A からは, 産業分類別, 常住地による従業地 (市区町村) 別就業者数が得られる (Table 2 参照). 常住地 i における統計表 A を用いて, 産業分類ごとに従業地 (市区町村) の割当てを行う. この場合, $\mathbf{a}_k = (a_1)$, $a_1 =$ 産業分類であり, 上記の

手順によって算出された人数を \mathbf{a}_k に適合する合成人口データの中から無作為に抽出し, 従業地 j を割り当てる.

3.3 従業地 (市区町村) の割当て (人口 20 万未満の市町村)

統計表 A は, 個人情報保護の観点から, 公開の対象が 21 大都市 (東京都区部および政令指定都市) の各区 (特別区および行政区), 県庁所在市, 人口 20 万以上の市に限られている. なお, 統計表 A が公開されている対象は, 全国の 1,896 市区町村 (東京都区部は特別区単位, 政令指定都市は行政区単位) のうち, 290 市区であり, 平成 27 年国勢調査において, これらの市区の就業者の合計は全国におけるすべての就業者の約 52.2% に相当する. 本稿で取り上げている高槻市は, 公開対象に該当するため統計表 A を用いた割り当てが可能であるが, これらの条件に適合しない人口 20 万未満の市町村においては, 全市町村を対象に公開されている二つの統計表を用いて段階的に割り当てる手法を提案する. なお, 各従業地に配分する人数の調整方法は前節と同様である.

Table 1 における統計表 B からは, 産業分類別, 男女別, 常住地 (市町村) による従業地 (6 区分) 別就業者数が得られる (Table 3 参照). 産業分類や男女といった属性を考慮することができるが, 従業地に関しては「自宅で従業」, 「自宅外自市区町村で従業」, 「県内他市区町村で従業」, 「他県で従業」, 「従業市区町村「不詳・外国」」, 「従業地「不詳」」といった概略的な 6 区分 (Fig. 1 参照) になっている. 一方, Table 1 における統計表 C からは, 男女別, 常住地 (市町村) による従業地 (市区町村) 別就業者数が得られる (Table 4 参照). 産業分類の属性がない代わりに, 具体的な従業地 (市区町村) の構成を知ることができる. そこで, 統計表 B および C を用いて次のように従業地 (市区町村) を割り当てる.

まず常住地 i における統計表 B を用いて, 産業分類ごと, 男女ごとに従業地 (6 区分) の割当てを行う ($\mathbf{a}_k = (a_1, a_2)$, $a_1 =$ 産業分類, $a_2 =$ 男女). このうち, 「県内他市区町村で従業」または「他県で従業」が割り当てられた就業者に対し, 常住地 i における統計表 C を用いて男女ごとに具体的な従業地 (市区町村) の割当

Table 5 : (統計表 D) 産業分類別, 市区町村・小地域別従業者数

表章項目：従業者数【人】 市区町村：茨木市

| | A 農業, 林業 | B 漁業 | C 鉱業, 採石業, 砂利採取業 | D 建設業 | E 製造業 | F 電気・ガス・熱供給・水道業 | G 情報通信業 | H 運輸業, 郵便業 | I 卸売業, 小売業 | J 金融業, 保険業 | K 不動産業, 物品賃貸業 | L 学術研究, 専門・技術サービス業 | M 宿泊業, 飲食サービス業 | N 生活関連サービス業, 娯楽業 | O 教育, 学習支援業 | P 医療, 福祉 | Q 複合サービス事業 | R サービス業 (他に分類されないもの) | S 公務 (他に分類されるものを除く) |
|------------|----------------|---------|------------------------|----------|----------|--------------------|------------|---------------|---------------|---------------|------------------|--------------------------|-------------------|---------------------|----------------|-------------|---------------|----------------------------|---------------------------|
| 茨木市 三咲町 | - | - | - | 18 | 21 | - | - | 138 | 19 | - | 4 | - | 13 | 5 | - | 18 | - | 26 | - |
| 茨木市 三島丘1丁目 | - | - | - | 11 | 181 | - | - | 13 | 97 | 18 | - | - | 20 | 1 | - | - | - | 3 | - |
| 茨木市 三島丘2丁目 | - | - | - | 6 | 2 | - | 5 | 3 | 18 | - | 8 | 2 | 2 | 5 | 1 | 50 | 5 | - | - |
| 茨木市 三島町 | - | - | - | 4 | - | - | 2 | - | 23 | - | - | 5 | 6 | 32 | 40 | 56 | - | 4 | - |
| 茨木市 五日市1丁目 | - | - | - | 12 | 210 | - | - | 227 | 97 | - | 7 | 26 | 307 | - | - | - | - | - | - |
| 茨木市 野々宮2丁目 | 4 | - | - | 56 | 97 | 15 | - | 127 | 67 | 4 | 25 | 1 | 10 | - | - | - | - | 25 | - |

Table 6 : 従業地 (小地域) の割当てが行えなかった人数とその要因の内訳

| | 表Aを用いた場合 | | 表Bおよび表Cを用いた場合 | |
|--------------------------------|----------|------|---------------|------|
| | 10セット平均 | 標準偏差 | 10セット平均 | 標準偏差 |
| 従業地 (小地域) の割当てが行えなかった就業者の合計(人) | 2070.90 | 0.83 | 2074.90 | 4.76 |
| 従業市区町村「不詳・外国」である就業者 | 1048.20 | 0.40 | 1040.00 | 0.00 |
| 従業地「不詳」である就業者 | 1007.50 | 0.50 | 1010.00 | 0.00 |
| その他の要因による就業者 | 15.20 | 0.75 | 24.90 | 4.76 |

てを行う ($a_k = (a_2)$, $a_2 =$ 男女). ただし, 統計表 C (「自宅に従業」, 「自宅外の自市町村に従業」, 「従業市区町村「不詳・外国」, 「従業地「不詳」」を除く) を, 常住地 i が属する都道府県の従業地 (市区町村) か否かで二つの表に分割し, 「県内他市区町村で従業」の就業者には常住している都道府県内の市区町村に, 「他県で従業」の就業者には常住している都道府県以外の市区町村にそれぞれ割当てを行う.

3.4 従業地 (小地域) の割当て

3.2 および 3.3 で割り当てられた従業地 (市区町村) が「自宅」である就業者に関しては, 合成人口データにおける町丁目および位置情報属性をそのまま従業地 (小地域) とする. ここでは, それ以外の従業地 (市区町村) に対し, さらに詳細な小地域へ割り当てる.

国勢調査から得られる従業地は市区町村単位であるため, 小地域の割当てには経済センサス-基礎調査を活用する. 経済センサス-基礎調査とは, 総務省により, 全国すべての事業所・企業を対象に産業分類や従業者規模などが調査された統計である. ここでの従業者とは, 当該事業所に所属して働いているすべての人を指す. なお, 経済センサス-基礎調査は, 国勢調査と同時期に調査されているものが存在しない. よって, 今回用いる国勢調査及び合成人口データに最も調査時期が近い平成 26 年 (2014 年) の統計を用いる. ここでは Table 1 における統計表 D から, 全国すべての市区町村・小地域 (町丁・大字) 別, 産業分類 (大分類) 別従業者数表 (Table 5 参照) を作成し, これを利用して確率的に小地域を選択する. 合成人口データにおいて, 従業地 (市区町村) が w , 産業分類が d であるとき, 従業地 (小地域) が c_k である確率 $f_{w, d}(c_k)$ は, 式(1)のよ

うに計算される.

$$f_{w, d}(c_k) = \frac{n(w, d, c_k)}{\sum_{k=1}^{N_w} n(w, d, c_k)}, \quad (k = 1, \dots, N_w) \quad (1)$$

ここで, $n(w, d, c_k)$ は市区町村 w , 小地域 c_k で従業している産業分類 d の従業者数, N_w は市区町村 w 内の小地域の数であり, いずれも統計表 D から得られる.

なお, 国勢調査で用いられている小地域 (町丁・字等) および合成人口データの町丁目属性の区域と, 経済センサス-基礎調査で用いられている小地域 (町丁・大字) の区域には相違がある. たとえば, 「高槻市 黄金の里」という地域は住所上, 1 丁目までしか存在しないが, 国勢調査では単に「黄金の里」, 経済センサス-基礎調査では「黄金の里 1 丁目」となっているなどの小地域の名称が一致しない場合や, 国勢調査では, 住所上単一の大字を同じ名称で複数の小地域に分割している場合などがある. このように, 両統計データの小地域の名称やその境界線は一対一に対応しているわけではなく, 各小地域を識別する ID 体系も異なっているため, 名寄せを行うことは容易ではない. そこで, 小地域の名称だけでなく, 割り当てられた小地域内における従業地の代表地点として, 経済センサス-基礎調査 町丁・大字別境界データに含まれる各小地域境界の図形中心点座標 (世界測地系緯度経度) を統計表 D に結合し, この属性も同時に割り当てる. ただし, 図形中心点座標が欠損値となっている小地域は除外し, 割当てが行われないようにする. また, 飛び地などにより同一の ID をもつ小地域境界が複数個存在する場合は, 代表となっている境界 (原則として面積が最大の境界

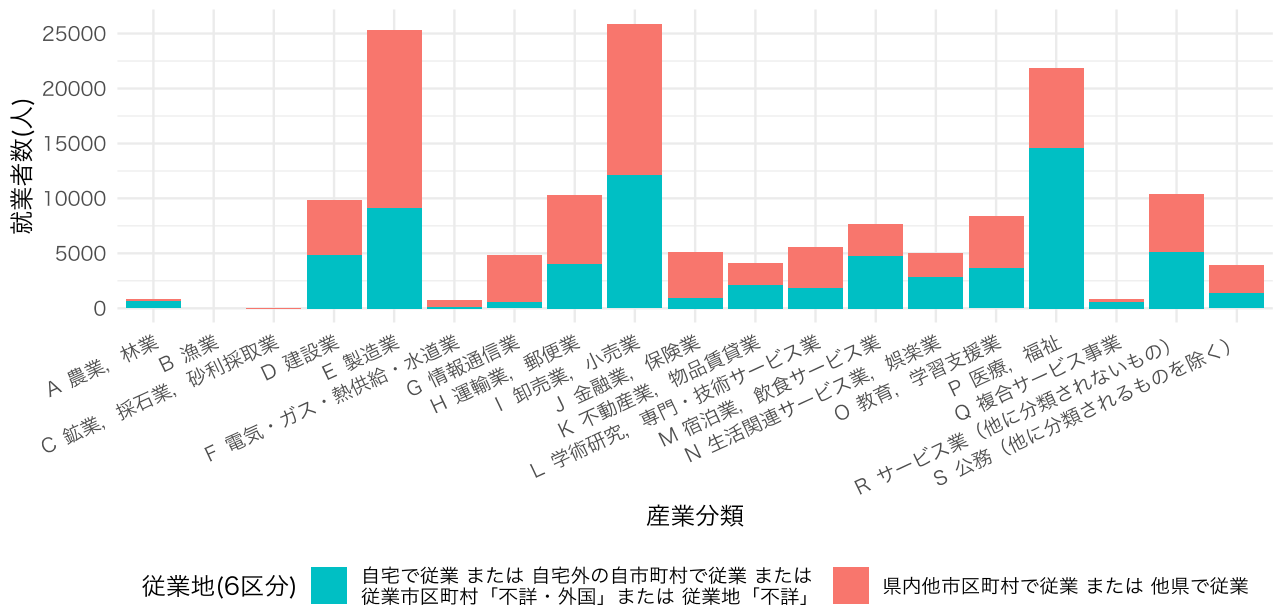


Fig. 4 : 従業地（6区分）に割当てを行った段階における産業分類別就業者数の従業地構成

であるが、島を含む場合は陸地の境界）を採用する。

4 従業地の割当て結果

3.2 および3.3の従業地（市区町村）の割当てにおいては、常住地によって異なる統計表を用いる手法を提案した。そこで、統計表Aを用いた場合と、統計表BおよびCを用いた場合とで、割当てられ方の違いを見るため、高槻市における10セットの合成人口データに両方の場合を適用させ、その結果を比較する。なお、疑似乱数のシード値はデータセットごとに変更するものとする。今回対象とした就業者150,529人を100%としたとき、従業地（小地域）まで割当てが行えた人数は、いずれの場合においても約98.62%となった。従業地（小地域）の割当てが行えなかった残りの約1.38%について、その要因の内訳をTable 6に示す。それぞれの場で「従業市区町村「不詳・外国」」および「従業地「不詳」」に割り当てられた人数が異なっているのは、統計表Bを用いて従業地（6区分）を割り当ての際に、統計表Aにはない男女属性を考慮したためである。また、その他の要因として、割り当てた従業地（市区町村）において、該当する産業分類の従業者が経済センサス-基礎調査上に存在しなかったため、小地域が割り当てられなかった就業者が存在した。これは、今回用いた国勢調査と経済センサス-基礎調査という二つの統計の調査対象や調査時期が異なることに起因する。この問題を回避するためには、3.2および3.3の従業地（市区町村）の割当てにおいて、経済センサス-基礎調査上に従業者が存在しない市区町村と産業分類の組合せの従業地（市区町村）には割り当てられないようにするといった工夫が必要である。統計表BおよびCを用いた場合の方が、このその他の要因に該当する人数が多いのは、統計表Cからは産業分類ごとの構成比を知ることができないため、国勢調査において本来存在しない産業分類と従業地（市区町村）の組合せとなる従業地（市区町村）が割り当てられる可能性があるためである。

統計表Cを用いて従業地（市区町村）の割り当てを

行う就業者の割合を産業分類ごとに見るため、統計表Bを用いて従業地（6区分）に割り当てた段階における産業分類別就業者数の従業地構成をFig. 4に示す。統計表Cを用いずに従業地（市区町村）の割当てが終了する「自宅で従業」または「自宅外の自市町村で従業」に該当する就業者は、統計表Aを用いた場合と同等の精度を保つことができるが、「県内他市区町村で従業」または「他県で従業」に該当する就業者は、実際とは異なる市区町村に割り当てられる可能性がある。提案手法（統計表Aを用いた場合）で小地域まで従業地の割当てを行った結果として、京阪神地区における、高槻市に常住する就業者（全産業分類）の従業地による分布をFig. 5に示す。なお、Fig. 5は10セットの合成人口データのうち、あるデータセットの結果である。また、色の塗り分けを行っている区域は平成26年経済センサス-基礎調査における小地域（町丁・大字）である。このようなコロプレスマップ（階級区分図）における統計値として、人口などの絶対的統計値を用いると各区域の面積の大小に影響されるため、実情とは異なった印象を与えることがある。そのため、ここでは相対的統計値である人口密度を用いて表す。

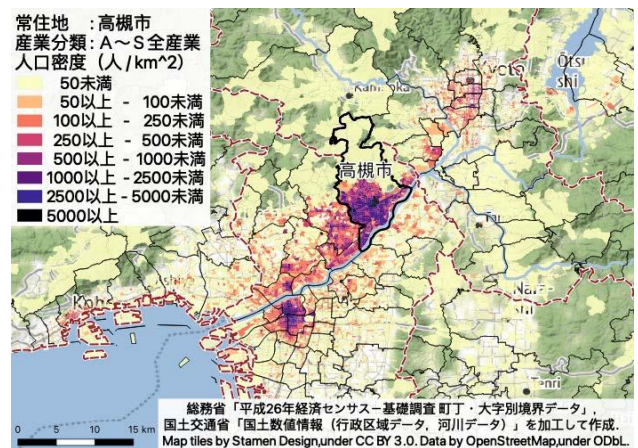


Fig. 5 : 就業者（全産業分類）の従業地による分布

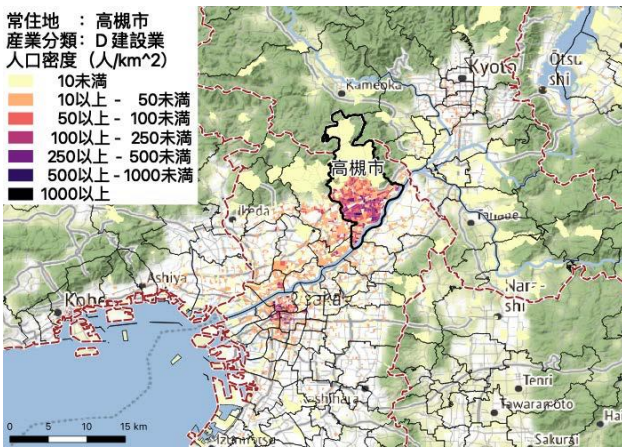


Fig. 6：就業者 (D 建設業) の従業地による分布

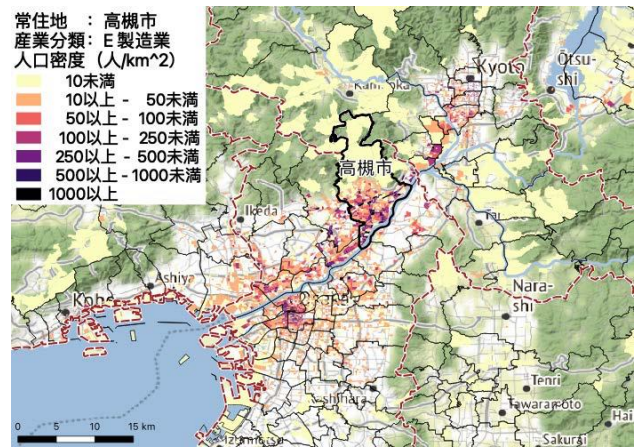


Fig. 7：就業者 (E 製造業) の従業地による分布



Fig. 8：就業者 (H 運輸業, 郵便業) の従業地による分布

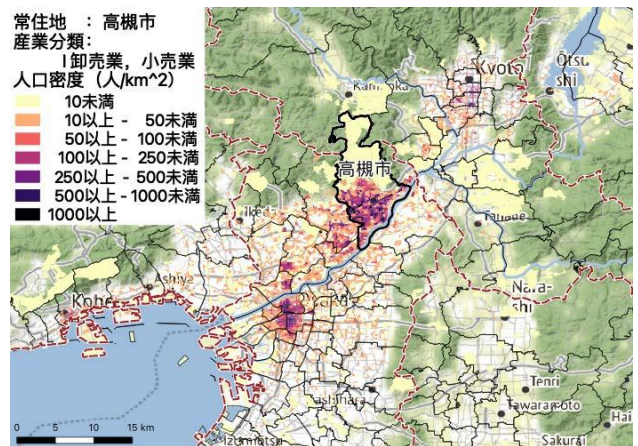


Fig. 9：就業者 (I 卸売業, 小売業) の従業地による分布

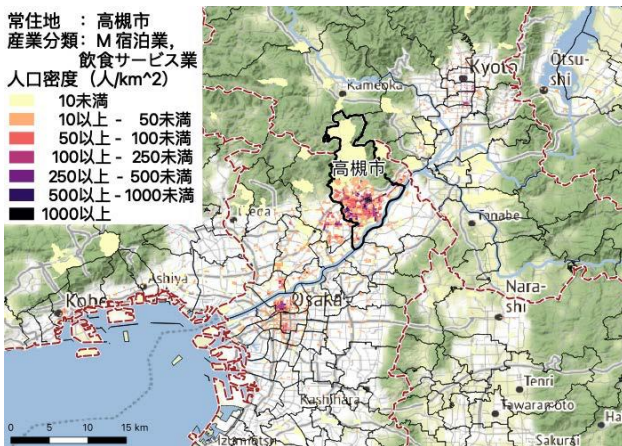


Fig. 10：就業者 (M 宿泊業, 飲食サービス業) の従業地による分布

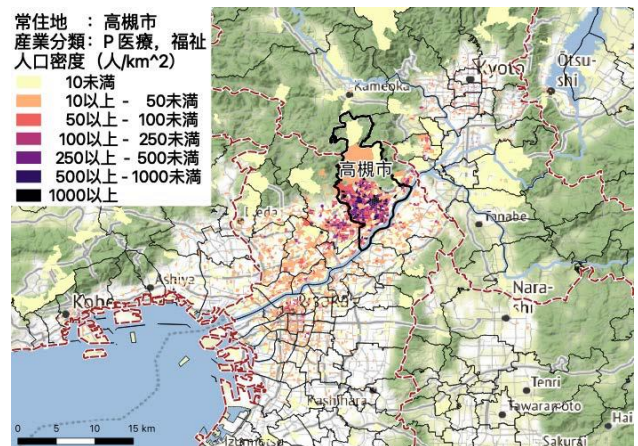


Fig. 11：就業者 (P 医療, 福祉) の従業地による分布

高槻市は大阪市と京都市のほぼ中間にあり、市域の南部を流れる淀川の右岸に位置している。Fig. 5 から、昼間において、高槻市に常住する就業者は、高槻市内などの北摂地域や大阪市内を中心に京阪神地区に広く分布していることがわかる。淀川の右岸と左岸とで分布の状況が大きく異なっているのは、鉄道や国道・主要地方道が淀川と並行して通っており、常住地からのアクセスが比較的容易な地域で従業している人が多いためであると考えられる。高槻市内においても、商業施設が集積している JR 高槻駅・阪急高槻市駅周辺の

中心市街地に、特に多くの就業者が集中していることが読み取れる。また、産業分類を考慮して従業地の割当てを行ったため、産業分類ごとにこの分布の傾向を把握することができる。ここでは、代表的な産業分類として、「D 建設業」、「E 製造業」、「H 運輸業, 郵便業」、「I 卸売業, 小売業」、「M 宿泊業, 飲食サービス業」、「P 医療, 福祉」を取り上げ、それぞれの分布を Fig. 6～Fig. 11 に示す。たとえば、Fig. 7 より、「E 製造業」に該当する工場などは、交通の利便性や広大な敷地を必要とする条件などが立地に関わっ

ていると考えられ、国道・主要地方道の沿道や河川敷周辺を中心に広範囲にわたって分布していることがわかる。一方、「P 医療 福祉」に該当する病院や保育所、老人ホームなどは、その役割から主に住宅地に立地していると考えられる。Fig. 11 より、「P 医療 福祉」では、高槻市内やその周辺の地域に多くの就業者が点在しており、職住接近の傾向が強いことが読み取れる。

5 おわりに

本研究では、合成人口データの就業者に対し、国勢調査と経済センサス-基礎調査を用いて従業地属性を付加する手法を提案した。提案手法により就業者に従業地を割り当てることで、従来の夜間人口に加え、就業者の昼間人口の空間分布を表現できるようになる。また、構築したアルゴリズムは、本稿で事例として取り上げた高槻市のみならず、すでに任意の市区町村の合成人口データに適用が可能であり、今後、全国の合成人口データに対し、同様のデータを標準的に整備していく必要がある。また、北下ら¹⁵⁾により、本手法で作成したデータを用いて、常住地-従業地間を結ぶ通勤経路および通勤所要時間を推定する研究が行われている。この研究を並行して進めることにより、災害発生時において特定の移動経路の被災が人々に与える影響や、人々が滞留する可能性のある場所の推定などが可能となり、合成人口データの活用機会が一層多様なものになることが期待できる。

提案手法で考慮できていない点としては、以下に挙げるものが考えられる。3.2 および 3.3 の従業地（市区町村）の割当てにおいては、ある属性の組合せに適合する合成人口データを無作為に抽出することで従業地（市区町村）を割り当てているが、その中で具体的にどの就業者を選択するかについては何も指示していない。実際には距離や地形、交通などの要因により、統計表からはわからない小地域レベルでの偏りが存在する可能性が考えられる。また、3.4 従業地（小地域）の割当てにおいては、従業地（市区町村）と産業分類が同じ就業者には、その常住地に関わらず同じ確率分布を用いている。そのため、ある従業地（市区町村）内に勤める就業者の常住地ごとの分布に偏りがあっても反映することはできない。さらに、位置情報について、常住地側（建物単位の座標）と従業地側（小地域の図形中心点座標）での空間解像度が異なっている点、企業規模や年齢などの考慮していない属性との整合性は保証できない点に留意が必要である。

今後取り組むべき課題としては、用いた統計データとの最適化を行い、確率的に推計した数値の整合性を図っていく必要があるほか、対象を非就業者にも拡大し、学生における通学地などの割当て手法を別途考案する必要がある。また、人々の1日の時間帯別の行動についてモデル化を行なった研究として、市川ら¹⁶⁾が、NHK 放送文化研究所によるデータブック国民生活調査¹⁷⁾を用いた標準生活行動モデルを提案しているほか、文献⁷⁾や文献⁹⁾においても生活時間調査を用いて1日の具体的なスケジュールを推定する試みが行われている。こうした生活時間に関する属性の割当ても併せて行っていくことで、公開されている統計データのみから、任意の時刻における時空間分布の推定などが行えるようになり、個人の生活行動を推定したより詳細な動的データを構築することが可能となる。

謝辞

本研究は、2020 年度関西大学研究拠点形成支援経費、JST 未来社会創造事業（JPMJMI20B3）、学際大規模情報基盤共同利用・共同研究拠点（jh210040-MDH）、HPCI システム利用研究課題による支援（hp200262）を受け、実施しました。

参考文献

- 1) T. Murata, T. Harada, D. Masui: Comparing transition procedures in modified simulated-annealing-based synthetic reconstruction method without samples, *SICE Journal of Control, Measurement, and System Integration*, **10-6**, 513/519 (2017)
- 2) T. Harada, T. Murata: Projecting household of synthetic population on buildings using fundamental geospatial data, *SICE Journal of Control, Measurement, and System Integration*, **10-6**, 505/512 (2017)
- 3) 国土交通省都市局 都市計画課都市計画調査室: 総合都市交通体系調査におけるビッグデータ活用の手引き (2018)
- 4) 小林: 職業属性を考慮した人流データの居住地及び通勤通学地推定, 東京大学大学院 新領域創成科学研究科 社会文化環境学専攻 2019 年度修士論文 (2020)
- 5) 大佛, 早坂: 携帯電話人口統計を用いた都市内移動者の時空間分布推定, 日本建築学会計画系論文集, **84-762**, 1853/1862 (2019)
- 6) 国土交通省: 全国の人流オープンデータ, <https://www.geospatial.jp/ckan/dataset/mlit-1km-fromto>, 最終閲覧日: 2021-01-25
- 7) 日高, 大野, 志賀: 集計データの統合による都市内の移動行動データ生成, 土木学会論文集 D3 (土木計画学), **72-4**, 324/343 (2016)
- 8) 総務省統計局: 平成 23 年社会生活基本調査, <http://www.stat.go.jp/data/shakai/2011/>, 最終閲覧日: 2021-01-25
- 9) 木全, 樋本, 西野, 田中: 全国的に整備される統計資料を利用した都市圏内就業者の時刻別空間分布の推定 - 平日の京阪神大都市圏を対象としたケーススタディ -, 日本建築学会計画系論文集, **78-686**, 891/898 (2013)
- 10) 独立行政法人 統計センター: オーダーメイド集計の利用, <https://www.nstac.go.jp/services/order.html>, 最終閲覧日: 2021-01-25
- 11) 政府統計の総合窓口 (e-Stat), <https://www.e-stat.go.jp/>, 最終閲覧日: 2021-01-25
- 12) 総務省統計局: 平成 27 年国勢調査, <https://www.stat.go.jp/data/kokusei/2015/index.html>, 最終閲覧日: 2021-01-25
- 13) 総務省統計局: 平成 26 年経済センサス-基礎調査, <https://www.stat.go.jp/data/e-census/2014/index.html>, 最終閲覧日: 2021-01-25
- 14) 杉浦, 村田, 原田: 賃金構造基本統計調査に基づく合成人口の労働者への就業属性別所得の割当て, システム制御情報学会論文誌, **32-2**, 69/78 (2019)
- 15) 北下, 村田: 合成人口データを活用した就業者の通勤経路・通勤所要時間の推定, 計測自動制御学会 システム・情報部門 学術講演会 (SSI2021), GS6-1-2, 291/294 (2021)
- 16) 市川, 小森: 人々の日常の標準生活モデルの提案に向けた統計情報分析, *SICE 社会システム部会 第 12 回社会システム部会研究会*, 188/194 (2017)
- 17) NHK 放送文化研究所: データブック国民生活時間調査 2015 (2015)