
**平成 24 年度
共同研究報告書**

**官庁統計とモバイル空間統計に
基づく新たな統計の創出に関する
共同研究**

2013 年 3 月

独立行政法人 統計センター

株式会社 NTTドコモ

目次

1 はじめに.....	3
1.1 まえがき.....	3
1.2 本報告書の構成と概要.....	3
2 モバイル空間統計について.....	5
2.1 概要.....	5
2.2 モバイル空間統計の特性.....	7
2.3 作成方法.....	7
集計処理.....	9
2.4 応用分野における有用性検証.....	15
まちづくり分野への応用.....	15
防災計画分野への応用.....	16
3 新しい統計に期待される要件の抽出.....	18
3.1 インタビューの概要.....	18
3.2 インタビュー結果.....	19
3.3 得られた知見.....	21
3.4 検討方針と課題.....	22
4. モバイル空間統計の信頼性検証.....	23
4.1 2010 地域メッシュ統計について.....	24
地域メッシュ統計とは.....	24
国勢調査地域メッシュ統計の作成方法.....	25
2010 年国勢調査の地域メッシュ統計.....	26
地域メッシュ統計利用上の留意点.....	26
4.2 検証手法.....	26
評価データについて.....	27
評価指標について.....	27
評価内容について.....	28
4.3 検証結果.....	29
空間解像度に対する全体傾向.....	29
3次・4次メッシュに関する地理的な傾向.....	30
3次・4次メッシュに関する推計人口と信頼性との関係.....	32
4.4 考察.....	34
5. ベンチマーク人口の評価と選定.....	35
5.1 住民基本台帳人口について.....	35

5.2 推計人口について	35
5.3 検証手法	35
5.4 検証結果	36
住民基本台帳人口と推計人口の傾向	36
差率による比較	37
5.5 考察	39
6 おわりに	40

1 はじめに

1.1 まえがき

情報通信技術の発展、特に情報機器の機能の急速な進歩とインターネットの普及に伴い、統計情報をより高度に活用するための環境が整備されてきた。そのため、統計のユーザーからはより詳細で、より多様な統計情報の提供が求められるようになった。そうした状況に対応するため、独立行政法人統計センターでは政府統計の総合窓口、e-Stat の運用を通じて広範かつ詳細な政府統計を各種の利用形態に応じて利用のしやすい形態で提供してきた。また、マイクロデータやオーダーメイド集計の提供を促進することにより、統計情報のより有効な活用の支援を行ってきた。しかしながら、これまでにない新たな統計情報を提供するために調査統計を拡充することには、費用負担、国民負担の面から極めて厳しい限界があるため、調査統計以外の情報を有効に活用することが求められるようになった。

他方、情報通信技術の発展は各分野で大量の電子データの蓄積を促すとともに、その大量の情報を処理することを可能とした。株式会社NTTドコモの「モバイル空間統計」もNTTドコモの基地局が把握している大量の情報を処理することで作成が可能となったものである。モバイル空間統計はこれまでの統計情報では得られない新たな情報を提供できることから、統計ユーザーの情報ニーズを満たすものとして、将来的に極めて有望な情報であると考えられた。

本研究は、このような背景のもと、時代の変遷に伴って発生しつつある新たな統計ニーズの高まりに対応するために、国勢調査などをはじめとする各種の官庁統計とモバイル空間統計を組み合わせることにより、新たに有用な統計情報を創出することを目的として、「モバイル空間統計」の新たな統計情報への活用方策を検討するために統計センターとNTTドコモが共同で実施したものである。

1.2 本報告書の構成と概要

本報告書は全6章から構成されている。本章「1 はじめに」では本研究の背景と目的を説明する。

「2 モバイル空間統計について」では、モバイル空間統計がどのようなものであるか、その特性、作成方法及び応用事例の紹介を行っている。

「3 新しい統計に期待される要件の抽出」では、本研究のはじめに行ったモバイル空間統計のニーズ把握のためのインタビューの結果について、インタビューのために訪問した6団体から聴取した具体的な内容とそれにより得られた知見を解説している。インタビュー結果は、総じて極めて好意的なものであり、モバイル空間統計に対し公共分野・産業分野・学術分野のいずれにおいても高い潜在需要があることが示唆された。また、時間解像度、空間解像度などの現在のモバイル空間統計の仕様は概ね妥当であるとの感触を得るとともに、信頼性要件に関してはいずれのインタビュー対象からも具体的な数値を伴う定量的な精度要求は得られなかった。

「4. モバイル空間統計の信頼性検証」では、2010年国勢調査メッシュ統計とモバイル空間統計を比較することにより行ったモバイル空間統計の信頼性検証の本年度の成果を説明している。国勢調査結果とモバイル空間統計の2010年10月1日の午後4時台の人口の差異について偏差率という指標を用いて評価を行っ

た。得られた知見としては、3次メッシュ（約1km四方）では人口集中地区と同程度の人口密度を持つ地域では高い信頼性が得られ、4次メッシュ（約500m四方）では人口集中地区の2倍程度の人口密度を持つ地域を対象とすることが望ましいという結果となった。また、国勢調査結果と差異の大きな都心部ではモバイル空間統計に居住者でない流入人口が多く含まれることも確認できた。

「5. ベンチマーク人口の評価と選定」では、モバイル空間統計の推計に用いるベンチマーク人口として利用可能な住民基本台帳人口と推計人口のどちらがベンチマークとしてより適切であるかを検証した。2010年国勢調査の結果に対して両統計の直近の結果の比較を行った結果、国勢調査結果からの乖離の傾向は類似しておりどちらの統計も不適切とはいえないものの、平均的な差率の大きさや差率の分布を勘案すると推計人口の仕様がより望ましいという結果となった。

「6 おわりに」においては、本年度の共同研究を総括するとともに、次年度以降における本研究の方向性について論じている。

2 モバイル空間統計について

地域ごとの人の集まり方は時々刻々と変化する。たとえば住宅地であれば、住民は朝になると通勤や通学、もしくは買い物などのために出かけていき、夕方や夜に帰宅する。その反対に、昼間には人で賑わうオフィス街は、夜には静まりかえる。

また、人の集まり方は毎日同じではなく、たとえば曜日によっても変化する。休日昼間のオフィス街は平日と比べて人出が少なくなり、その代わりにショッピングセンターなどの商業地やテーマパークなどの行楽地に人は集まるようになる。また、花火やお祭り、コンサートなどのイベント開催日には、開催地域やその周辺での人の集まり方は、普段と比べて大きく異なることになるだろう。

これらの人の増減の傾向は、我々の日常生活の中でも普段から感じられることである。しかし、それでは実際のところ、それぞれの地域で、どの時間にはどのような人がどのくらい集まっているのだろうか。

このような「それぞれの場所に実際にいた人の数」に基づく人口は「現在人口（現地主義に基づく人口）」とも呼ばれる。しかし、時々刻々と変化する現在人口の推移を、時間単位で継続的に、かつ広範囲で把握することは今まで容易ではなかった。

日本の人口に関する統計としては、5年ごとに実施される国勢調査によるもの（夜間人口）が最も網羅的かつ信用がおけるものである。しかし、国勢調査による夜間人口は、「それぞれの人が住んでいる場所」に基づいて作られた「常住人口（常住地主義に基づく人口）」であり、対象とする人口の定義が異なる。すなわち、日々の生活における、人々の移動に伴う人口の変化を反映するものではない。

また、交通量調査などのように、調査員が人手でカウントすることにより現在人口を測定することも困難である。たとえば、「ある日・ある時間帯において東京都新宿区に居た人の数」をこのような手段で測定することは、必要となる調査員の数などを考えても現実的ではない。

モバイル空間統計は、24時間365日、日本中のほぼどこでも携帯電話が繋がる仕組みを支える携帯電話ネットワークの運用データに基づき、これらの時間とともに変動する地域ごとの現在人口、すなわち「それぞれの場所に実際にいる人の数」に基づいた人口を、日本全国で継続的に推計する新しい人口統計である。

2.1 概要

モバイル空間統計は、携帯電話ネットワークの運用データを統計的に処理することにより推計された、現在人口の推計値である。

より具体的には、携帯電話ネットワークが「いつでもどこでも」携帯電話を繋げる仕組みを維持するための、携帯電話基地局ごとの携帯電話の在圏状況に関するデータである在圏データと、携帯電話ユーザの属性に関するデータである属性データの2種類の運用データに基づいて、携帯電話の普及率などを加味して統計処理することにより、携帯電話ユーザーではない人も含めた人口を推計する。

これにより、モバイル空間統計は、(1) ほぼ日本全国を対象範囲として、(2) 1時間単位で継続的に、(3) 標準地域メッシュごとや市区町村ごとなどのさまざまな地理的区分ごとに、(4) 性別・年齢層別・居住エリア別ごとに分類した上で、時々刻々と変動する現在人口の推移が推計できるという特徴を持つ。

図2.1にモバイル空間統計の仕組みと、モバイル空間統計によって提供される3種類の人口に関する推計値、すなわち人口分布、人口推移、人口構成について、それぞれのイメージを示す。

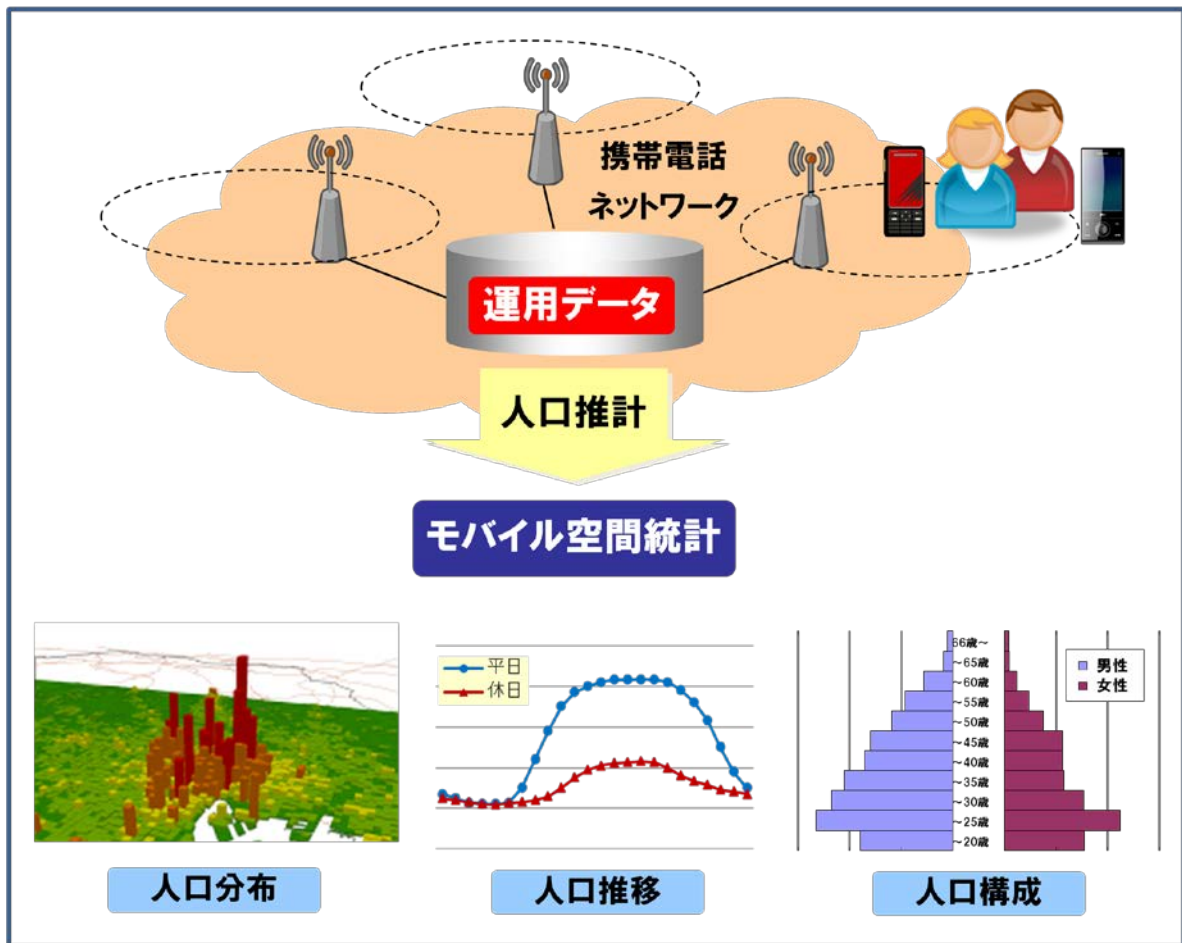


図 2.1 モバイル空間統計の仕組みと人口推計値

- **人口分布**：ある日のある時間帯における人口の地理的な分布である。モバイル空間統計は，平日・休日などを問わず，原則として任意の日および時間帯における人口分布を推計することができる。そのため，人口分布を分析することにより，都市近郊における昼夜での人の集まり方の違いや，平日と休日との人の集まり方の違い，もしくは大型連休やお盆などの特定期間における人の集まり方などを把握することができる。
 - **人口推移**：ある地域（市区町村やメッシュなど）における現在人口の変化の時間推移である。人口推移を分析することにより，ある地域で最も人口が多くなる時間帯は何時頃になるか，もしくは観光地における季節ごとの観光客数の変動などの，時間の経過に伴う人口の変動を把握することもできる。
 - **人口構成**：ある地域・ある時間帯における，性別・年齢別の人口や居住エリア別などの，属性ごとの人口の構成である。人口構成を分析することにより，たとえば若い女性が集まる傾向があるエリアや，あるオフィス街にはどの地域に居住している人が多く通勤しているかなどを把握することができる。
- このように，モバイル空間統計は，時々刻々と変動する現在人口を月，曜日，時間，性別，年齢，居住エリアごとに継続して推計し，時間や地域による推移や変化を調べるための手段を提供する。これにより，例えば国や自治体によるまちづくりや防災計画，地域活性化など，人口の分布や推移，構成に関する実態の把

握が重要となる分野において、企画・立案や施策効果の検証などにおける合理的な判断を支援するとともに、これらを通じた社会や産業の発展に寄与していくことが期待される。

2.2 モバイル空間統計の特性

モバイル空間統計は携帯電話ネットワークの運用データを利用して人口を推計することから、その特性は携帯電話ネットワークの性質に大きく依存する。以下、モバイル空間統計の人口統計としての特性について、対象地域、対象人口、空間解像度、時間解像度、推計精度のそれぞれの観点ごとに述べる。

- **対象地域**：モバイル空間統計は携帯電話ネットワークの運用データに基づいて推計されるため、その推計対象となる地理的範囲は携帯電話のサービスエリアとほとんど等しいものとなる。本稿執筆時点で、NTTドコモの携帯電話のサービスエリアは日本全国の市区町村役場を100%カバーするため、モバイル空間統計もほぼ同等の範囲を推計対象とする。
- **対象人口**：モバイル空間統計が推計対象とする人口は、携帯電話の利用人口に依存する。すなわち、携帯電話の普及率が極端に低い年齢層の人口を推計することはできない。具体的には、80歳以上や、携帯電話サービスの契約を行えない14歳以下の年齢層に関する人口は、現状ではモバイル空間統計の推計対象外である。したがって、モバイル空間統計の推計対象人口は年齢層が15-79歳の人口となる。
- **空間解像度**：モバイル空間統計の空間解像度、すなわちモバイル空間統計がどのくらいの大きさのエリアを単位として人口を推計できるかは、携帯電話基地局の設置間隔に依存する。すなわち、都市部などの人が多く集まるエリアでは基地局が密に設置されているが、郊外などでは基地局の設置間隔はより広いものとなるため、空間解像度は全国一律ではない。そのため、空間解像度の値を一概に示すことは困難であるが、およその目安として、東京23区内ではほぼ500mメッシュ単位で、郊外では数kmメッシュ単位程度の解像度で人口を推計することができる。
- **時間解像度**：モバイル空間統計の時間解像度、すなわちモバイル空間統計においてどのくらいの時間単位で人口を推計できるかは、携帯電話基地局がエリア内に在圏する携帯電話を把握する頻度に依存する。これは、通常はほぼ1時間程度の周期で行なわれるため、モバイル空間統計では、1時間単位を基本として人口を推計する。
- **推計精度**：モバイル空間統計は、国勢調査などの人口に関する悉皆調査と異なり、NTTドコモの携帯電話ユーザーを対象とする携帯電話ネットワークの運用データから統計的な推定を介して作成されることから、推計された人口には推計誤差が発生しうる。具体的な推計精度は地域やメッシュの大きさ、人口の多寡などに依存する。モバイル空間統計の具体的な推計精度については第4章で議論する。

2.3 作成方法

モバイル空間統計を作成するために用いられる携帯電話ネットワークの運用データであるそれぞれの在圏データと属性データは、本来の目的から携帯電話ユーザーのプライバシーに関する機微な情報を含みうる。

したがって、その扱いにあたってはプライバシーの保護に十分に留意する必要がある。そこで、モバイル空間統計の作成にあたっては、図 2.2 に示す非識別化処理，集計処理，秘匿処理の 3 段階の処理を通じてユーザーのプライバシーを慎重に保護している。

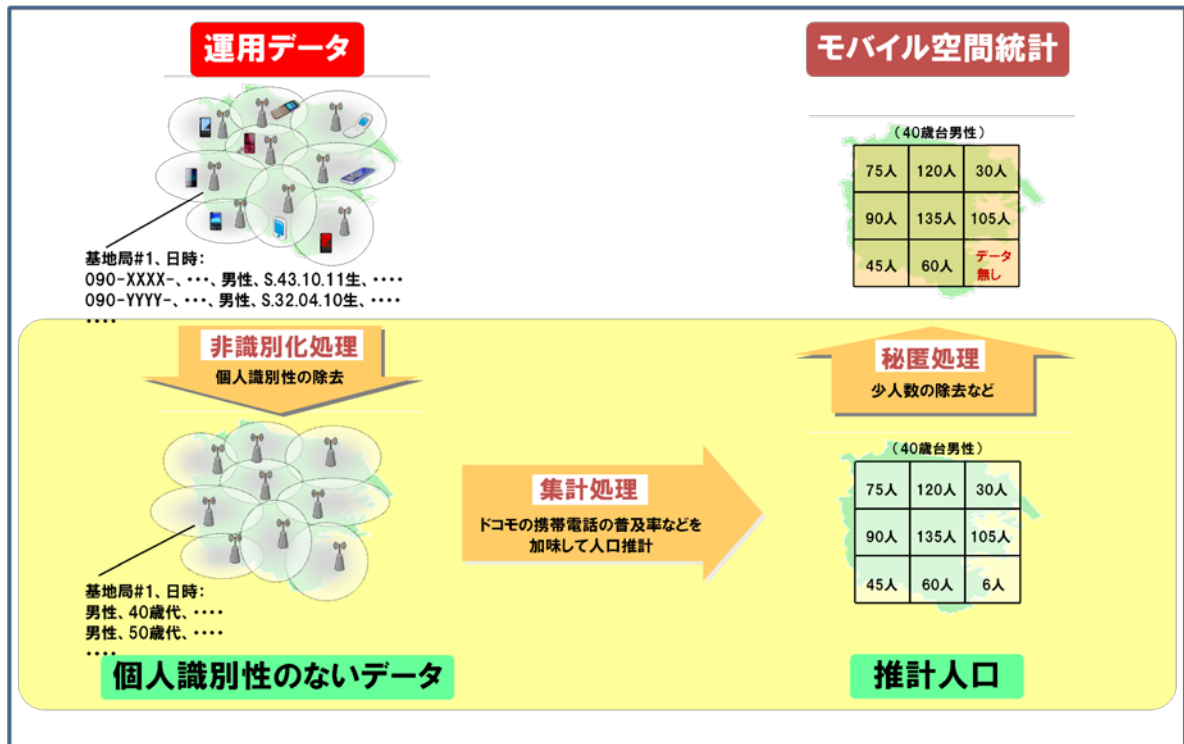


図 2.21 モバイル空間統計の作成方法

1. **非識別化処理**：まず、人口の推計に先立って、運用データから、人口の推計には不必要である個人を特定する情報を取り除く。モバイル空間統計は人口分布や人口推移、人口構成などの人口に関する統計であるため、それぞれの運用データが誰についてのものであるかを特定する必要はない。すなわち、時間ごと、エリアごと、そして性別や年齢層などの個人を識別できない属性ごとに推計値を作成できれば十分である。そこで、実際に人口の推計を行う集計処理に先立ち、運用データから電話番号や名前など個人を識別する情報を削除し、生年月日や住所などは年齢層や行政界コードなどに変換する。
2. **集計処理**：次に、非識別化処理された運用データに基づいて、基地局の通信圏内（基地局エリア）ごとの携帯電話台数をユーザーの属性別に推計し、携帯電話の普及率や基地局のカバーエリアに関する情報を加味することにより、NTT ドコモの携帯電話ユーザー以外も含めた人口を推計する。集計処理は、在圏数推計，拡大推計，エリア変換の 3 つのステップから構成される。それぞれのステップの詳細については後述する。

3. **秘匿処理**：最後に，人口がごく少人数になるエリアなど，極端な条件下などにおいてもユーザーのプライバシーを保護するために，集計処理の出力結果を補正する．秘匿処理は，国などが作成する公的統計の公開においても必要に応じて実施されている（「統計的開示制御（statistical disclosure control）」とも呼ばれる）．モバイル空間統計における秘匿処理は，公的統計で適用される手法に準じて，国内外の技術開発動向を踏まえて実施されている．この秘匿処理により補正されたデータが，最終的にモバイル空間統計となる．

なお，これらのモバイル空間統計の作成手順は，モバイル空間統計を作成・提供する際に遵守する基本事項をまとめたガイドラインとして公表されている．

集計処理

モバイル空間統計は携帯電話ネットワークの運用データを用いて現在人口の統計を作成する．運用データからのモバイル空間統計の作成は，ごく単純化すれば以下の手順により行なわれる．

- 1) 運用データに基づいて，それぞれの基地局エリアに存在する携帯電話の台数を集計する．
- 2) 人口あたりのドコモの携帯電話の普及率に基づいて，携帯電話の台数を人口に拡大する．
- 3) 上記で得られた基地局エリアごとの人口を，メッシュや市区町村などの地理区分ごとの人口へと再集計する．

しかし，携帯電話ネットワークはそもそも人口統計を生成することを目的として設計されていないため，実際には上記のような単純な手順で運用データからモバイル空間統計を作成することはできない．

たとえば，携帯電話の基地局は，着信制御のためにその基地局エリアに所在する携帯電話の存在をほぼ定期的（約1時間ごと）に確認している．運用データの一部である携帯電話の所在データは，具体的にはこの所在確認の状況を集約したデータである．しかし，所在確認のタイミングは携帯電話ごとにばらばらであり，また，ある時点で基地局エリアに存在していることが確認された携帯電話が，次の確認時点までずっとその基地局エリアに存在しているとは限らない．つまり，携帯電話ユーザーの移動に伴って他の基地局エリアに移っているかもしれないが，ある確認時点から次の確認時点まで，その携帯電話が実際にはどの基地局エリアに存在していたかについては，携帯電話ネットワークは知ることはできない．さらに，実際の携帯電話ネットワークでは，この確認の間隔は必ずしも一定ではなく，携帯電話の移動や利用の状況などによって大きく変動することがある．したがって，「ある時刻に何台の携帯電話が基地局エリアに存在していたか」という情報を携帯電話ネットワークは直接知ることはできず，それぞれの基地局における所在確認の状況から統計的に推定する必要がある．

また，基地局エリアごとの携帯電話の台数（在圏数）が推定できたとしても，これは当然ながらドコモの携帯電話ユーザー以外も含めた人口とは異なる値である．一見すると，日本の人口一人あたりのドコモ携帯電話の普及率がわかれば，在圏数を普及率で単純に「割り戻す」ことにより人口を推計できるように考えられる．たとえば，日本の人口が1億2000万人，ドコモの携帯電話サービスの契約数が6000万契約であるとする．このとき，人口2人に対して1台の割合（普及率=0.5）でドコモの携帯電話が普及していることになるので，基地局エリアの在圏数を0.5で割り戻す（2倍する）ことにより，その基地局エリアに存在する人口を算

出できる。しかし、実際には地域や年代・性別ごとにドコモの携帯電話の普及率は異なるため、このような単純な割り戻し処理では、地域や年代・性別によって大きく偏りが出てしまい、人口統計としては不適切なものとなる。さらに、携帯電話の電源は、常にオンの状態にあるとは限らない。たとえば、電源が切れている携帯電話がどの基地局エリアに存在するかについて、携帯電話ネットワークが知ることはできない。そこで、これらの要因を考慮した上で、携帯電話の在圏数を人口へと適切に拡大推計する必要がある。

最後に、これらの課題を解決して基地局エリアごとの人口を正しく推計できたとしても、それだけでは人口統計として活用することは困難である。たとえば、「基地局 X のエリアの人口は Y 人である」という単位で人口を提供されても、そもそも基地局 X というのがどこにあって、どのくらいの広がりを持つエリアなのかがわからなければ、この人口データに使い道はない。さらに、携帯電話の基地局には屋内局/屋外局、利用する無線周波数帯など様々なバリエーションがあり、それぞれの基地局でエリアカバー範囲も大きく異なる。そのため、基地局エリア単位で推定された人口を、都市計画や防災計画など、人口統計の活用分野で良く使われる単位である、メッシュや行政区画(市区町村界など)単位の人口へと変換する過程が必要となる。

以上をまとめると、集計処理は以下の 3 段階の推計プロセスから構成される。

1. **在圏数推計**：(非識別化処理された運用データに含まれる) 基地局ごとの在圏データに基づいて、在圏データの発生頻度の偏りによる影響を補正しつつ、それぞれの基地局エリアに在圏する携帯電話の台数(在圏数)を推計する。
2. **拡大推計**：基地局エリアごとに推計された在圏数を、携帯電話の普及率(何人に1台の割合でNTTドコモの携帯電話が利用されているか)に基づいて、基地局エリアごとの人口に拡大する。携帯電話の普及率は、実際には地域や年齢層、性別などの属性ごとに差異があり、また携帯電話の電源 ON/OFF の割合などによっても影響を受けるため、これらの影響による偏りを避けるために、属性ごとに普及率を調整しながら拡大推計を実施する。
3. **エリア変換**：基地局エリアごとに推計された人口を、それぞれの基地局の地理的なカバー範囲に基づいて、メッシュや市区町村などの地理区分ごとの人口に変換する。このエリア変換の結果が集計処理の出力結果となる。

以下、それぞれのステップにおける基推計の考え方と推計方式を示す。

在圏数推計

在圏数推計処理の目的は、非識別化された運用データに基づき、ある時間帯にそれぞれの基地局エリアに何台の携帯電話が在圏していたか、その在圏数を推計することである。

前述の通り、携帯電話ネットワークは「ある時点で、どの基地局に何台の携帯電話が在圏しているか」を正確に把握しているわけではない。そこで、1時間ごとなどのある一定期間(以下、観測期間 T と呼ぶ)に収集された所在データに基づき、その期間の平均在圏数を求める。図 2.3 に在圏数推計のイメージ図を示す。観測期間 T における、基地局エリア C の平均在圏数を m_C^T とする。在圏数推計は、それぞれの基地局エリアごとに m_C^T を推計する。

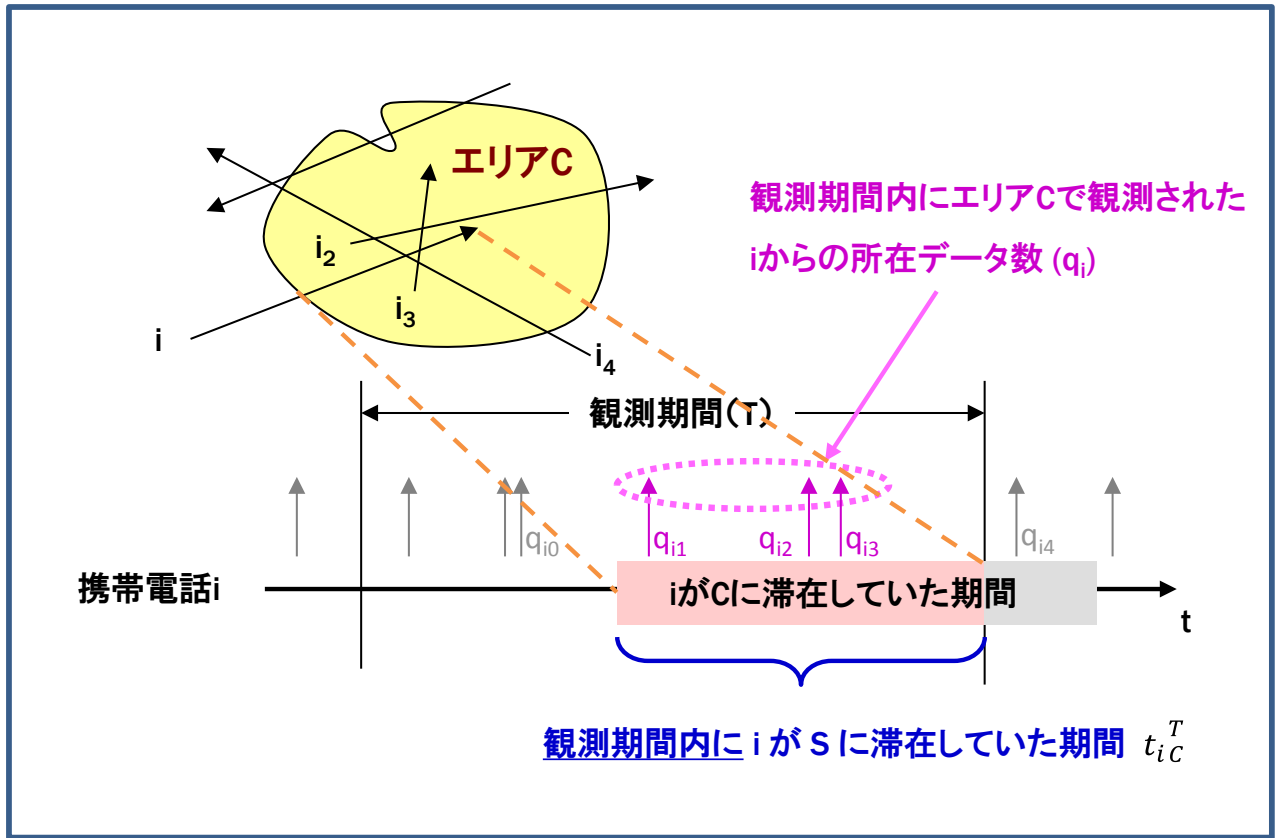


図 2.32 在圏数推計

在圏数 m_C^T は，観測期間 T における，基地局エリア C での携帯電話の滞在時間の総和を，観測期間の長さで除したものと計算できる．すなわち，観測期間 T ，基地局エリア C における携帯電話 i の滞在時間を t_{iC}^T とすれば，

$$m_C^T = \sum_i t_{iC}^T / |T|$$

として在圏数を求めることができる．

しかし，携帯電話ネットワークは前述の通り常に携帯電話の在圏状況を把握しているわけではないため，それぞれの携帯電話の正確なエリア滞在時間 t_{iC}^T はわからない．そこで，所在データの生成状況から統計的にこれを推定する．

もし，所在データの生成契機である在圏確認の間隔がある周期 ω で一定であれば，エリア滞在時間が ω 増えるごとに，そのエリアで観測される所在データの数が1つ増えることが確率的に期待される．したがって，基地局エリア C において，期間 T に観測された所在データの集合を L_C^T とすると，エリア滞在時間の総和 $\sum_i t_{iC}^T$ は以下のように推定できる．

$$E\left(\sum_i t_{iC}^T\right) = \omega |L_C^T|$$

しかし、前節で述べたように、実際の携帯電話ネットワークでは、在圏確認の間隔は一定でなく、携帯電話の利用や移動の状況などの様々な要因により動的に変化する。すなわち、個々の所在データごとに ω が変動する（所在データの時間密度が変化する）。そこで、これらの所在データの時間密度のばらつきを考慮し、それぞれの所在データに対して周辺の信号間隔に応じた特徴量（重み）を与える。ある所在データに対する特徴量は、典型的には前後の所在データとの間隔の平均として与えることができる。ある所在データ j の特徴量を ω_j とすると、エリア滞在時間の総和 $\sum_i t_{iC}^T$ は以下の式で推定できる。

$$E\left(\sum_i t_{iC}^T\right) = \sum_{j \in L_C^T} \omega_j$$

したがって、観測期間 T における、基地局エリア C の在圏数 m_C^T は、当該期間およびエリアで観測された所在データ L_C^T とその特徴量により、以下の式を用いて推計される。

$$E(m_C^T) = \sum_{j \in L_C^T} \omega_j / |T|$$

さて、ここまでの議論では、人口構成や移動人口の推計に必要となる、属性（年齢層・性別・居住エリア）ごとの在圏数推計には言及してこなかった。ただし、これらの属性ごとの在圏数も、上記の属性なしの式に対する簡単な変更のみで推計することができる。

ある（年齢層，性別，居住エリア）の組からなる属性条件を A とする。たとえば、東京都在住の 30 代男性を表す属性は、 $A = (30 \text{ 代}, \text{男性}, \text{東京都})$ と表される。このとき、前述の t_{iC}^T のうち、属性 A の携帯電話 i の滞在時間を $t_{iC,A}^T$ とすれば ($t_{iC,A}^T \subseteq t_{iC}^T$)、

$$m_{C,A}^T = \sum_i t_{iC,A}^T / |T|$$

として、東京都在住 30 代男性の、上記エリア・時間帯における在圏数を求めることができる。したがって、 $L_{C,A}^T$ を所在データ L_C^T のうち属性 A の携帯電話に関する所在データとすれば、これは以下の式で推計できる。

$$E(m_{C,A}^T) = \sum_{j \in L_{C,A}^T} \omega_j / |T|$$

なお、所在データ L_C^T から属性 A の携帯電話のみに関する所在データ $L_{C,A}^T$ への絞りこみは、それぞれの所在データに対応する契約データを参照することにより容易に実現することができる。

拡大推計

拡大推計処理では、在圏数推計処理で求めた在圏数（携帯電話数）に基づいて人口を推計する。

前述の通り、もしドコモの携帯電話の人口あたり普及率が、地域や年齢層・性別などによらず一様であり、かつ全ての契約済み携帯電話の電源が入っておりどこかの基地局エリアに必ず在圏しているという前提を置くことができれば、拡大推計処理では単に推計された在圏数に対して、普及率の逆数を乗ずる（割り戻す）のみで人口を推計することができる。

しかし，現実には，年齢や性別，居住エリア域によってドコモの携帯電話の普及率は大きく異なる．この普及率の偏りを正しく反映することなしには，推計された人口も大きく偏りを生じることになる（たとえば，実際の普及率が高い地域では人口が過大に，普及率が低い地域では過少にそれぞれ偏って推計される）．

さらに，人々が就寝する夜間帯などで携帯電話の電源が切られているとすれば，その間は携帯電話と携帯電話ネットワークの間のやりとりは行われなため，これらの携帯電話はどの基地局エリアにも在圏数として表れない．したがって，単に契約状況だけに基づいて拡大の乗数（拡大率と呼ぶ）を決定すると，実際の人口に対して，特に在圏率が小さい（電源 OFF の端末が多いなど）時間帯において，実際よりも過少に人口が推計されることになる．

そこで，拡大推計では性別・年齢，居住エリアなどの属性によって変わる普及率を正しく人口推計値に反映させるために，属性ごとに拡大推計処理を行なう．また，拡大率の算出にあたっては，単に普及率を用いるのではなく，携帯電話の在圏状況を加味した在圏率を用いる．図 2.4 に拡大推計処理のイメージ図を示す．

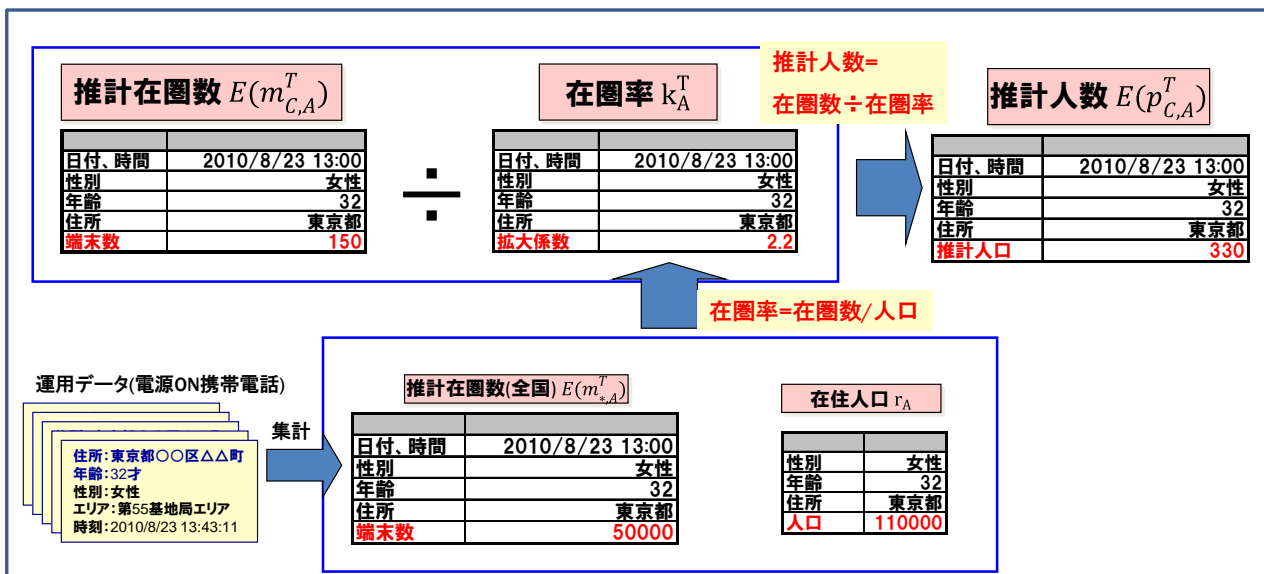


図 2.4 拡大推計

具体的には，まず，属性ごとの普及率の偏りを正しく反映させるために，住民基本台帳人口による在住人口を属性ごとに集計する．この属性 A の在住人口を r_A とする．属性の分割単位は，本稿執筆時点では，年齢 5 歳区分，性別ごと，都道府県ごととしている．次に，電源 OFF などによる非在圏端末の影響をなくすため，観測期間 T において，それぞれの属性に相当する総在圏数を（全国の基地局エリアから集約して）求める．観測期間 T における属性 A の総在圏数 $m_{*,A}^T$ は，以下の式により推計できる．

$$E(m_{*,A}^T) = \sum_C E(m_{C,A}^T)$$

このとき、観測期間 T における属性 A の人口について、人口 1 人あたりの在圏携帯電話数、すなわち在圏率 k_A^T は、以下の式で求められる。

$$k_A^T = E(m_{*,A}^T) / r_A$$

このとき、観測期間 T、基地局エリア C における属性 A の人口 $p_{C,A}^T$ は、以下の式によって推計される。

$$E(p_{C,A}^T) = E(m_{C,A}^T) / k_A^T$$

上記観測期間、基地局エリアにおける人口 p_C^T は、当該期間・エリアの属性ごと人口の総和として求められる。

$$E(p_C^T) = \sum_A E(p_{C,A}^T)$$

以上のように、拡大推計処理では、性別・年齢・地域でのドコモ携帯電話の普及率の偏りや、携帯電話の電源状態などによる在圏数の変動による影響なく、基地局エリアごとの携帯電話在圏数から人口を算出する。

エリア変換

エリア変換処理は、拡大推計処理までによって推計された基地局エリアごとの人口を、500m/1km などの各サイズのメッシュごと、もしくは市区町村などの行政界ごとなどの各種の応用分析で一般に用いられる地域区分ごとの人口に変換する。図 2.5 にエリア変換処理のイメージ図を示す。

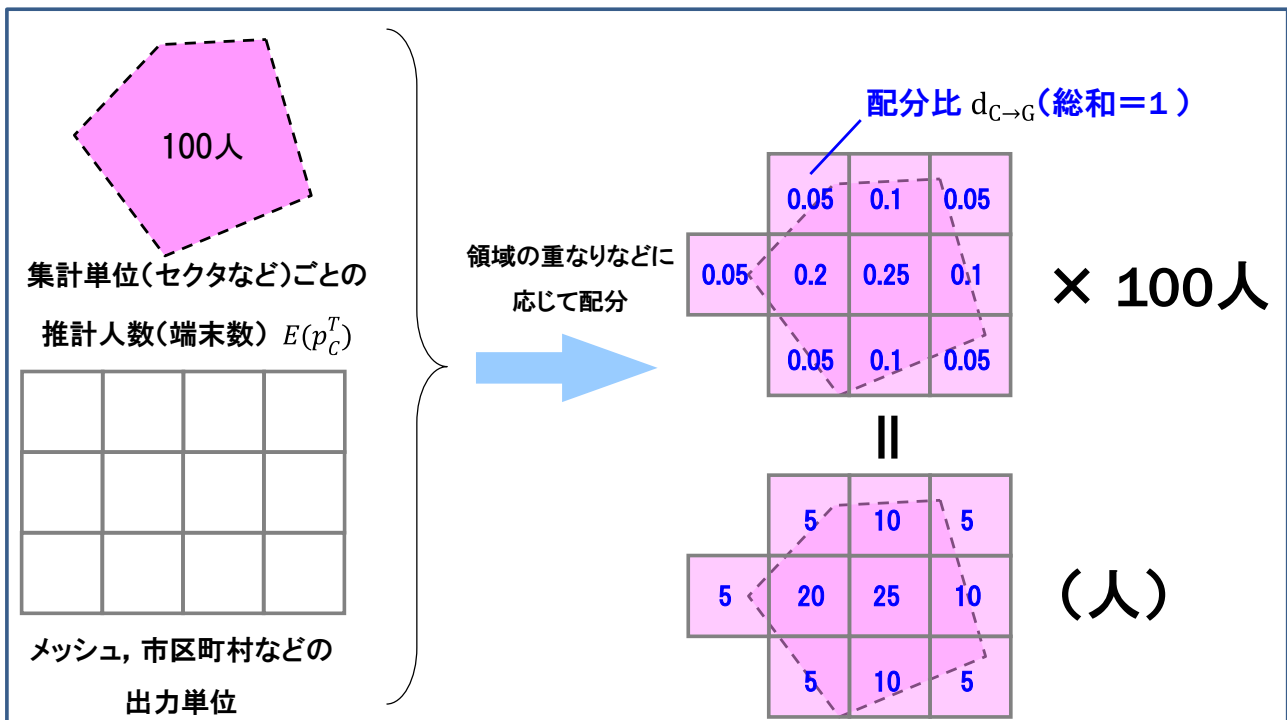


図 2.5 エリア変換

エリア変換処理において、観測期間 T における集計単位 G の人口 p_G^T は、以下のように推計される。

$$E(p_G^T) = \sum_C d_{C \rightarrow G} E(p_C^T)$$

ここで、 $d_{C \rightarrow G}$ は基地局エリアCから集計単位Gへの人口の配分比であり、基地局エリアごとにその総和は1となる。すなわち、 $\sum_C d_{C \rightarrow G} = 1$ が成立する。

配分比 $d_{C \rightarrow G}$ は、具体的には基地局エリアと各集計単位との地理的な「重なり」の大きさに基づいて求められる（面積按分と呼ばれる）。基地局エリアCの面積を $|C|$ 、基地局エリアCと集計単位Gとが地理的に重なる面積を $|C \cap G|$ とする。このとき、配分比 $d_{C \rightarrow G}$ は以下の式により求められる。

$$d_{C \rightarrow G} = \frac{|C \cap G|}{|C|}$$

2.4 応用分野における有用性検証

時々刻々と変化する人口の実態である現在人口を継続的に推計することができるモバイル空間統計は、公共施設や商業施設の立地選定などのまちづくりや、大規模災害に供えた避難所・避難物資の配置などの防災計画の立案など、市民の利便性や安全・安心を向上させるための施策について、企画の立案や効果の検証などを行うにあたって有用なものになることが期待される。

NTTドコモでは、大学との共同研究を通じ、これらの応用分野における専門家と連携してモバイル空間統計の有用性の検証を進めている。以下では、まちづくり分野への応用において東京大学（新領域創成科学研究科 清家剛准教授）と、および防災計画分野への応用において工学院大学（建築学部 村上正浩准教授）と、それぞれ共同研究を通じて進めた有用性検証の結果について概要を述べる。

まちづくり分野への応用

都市内部の人口構造が大きく変化しつつある現代において、きめ細やかな施策を行うためには、ダイナミックに変化する人口の分布やその変動を逐次把握することが望ましい。しかし、これまでは国勢調査などの静的な人口統計に頼らざるを得ないのが現状であった。そこで本研究では、柏市をフィールドとしてケーススタディを行い、モバイル空間統計のまちづくりへの活用可能性を示した。以下に、ケーススタディの観点と結果について概要を示す。

(1)土地利用 –時間ごとの人口推移に着目したエリア特性の把握–

柏市域を拠点エリア、住居エリア、田園エリアの3つに区分し、それぞれのエリアでの一日を通じた人口推移をモバイル空間統計で推計した。

その結果、一日の人口推移が拠点エリアでは凸型（昼間人口 > 夜間人口）、住居エリアでは凹型（昼間人口 < 夜間人口）、田園エリアでは直線型（昼間人口 \approx 夜間人口）となる傾向が、ほとんどの地区で当てはまることを確認した。このように、モバイル空間統計の一日の人口変化をみることで、そのエリアが拠点性をもつのか、住宅街であるのか、人口の流入が少くない田園系のエリアであるのかといった地域の特性を容易に把握することができた。なお、モバイル空間統計による人口推移が凸型となる田園系エリアも存在したが、そこに教育施設が存在する、あるいは鉄道が縦断するなど、特殊な傾向を示す要因も特定できた。

(2)公園・緑地 –実利用時間帯の人口に着目した公園の配置状況の評価–

モバイル空間統計により、公園・緑地の利用が多くなる昼間の時間帯における人口分布に基づいて、一人あたりの公園面積を 1km メッシュ単位で算出し、それを夜間人口に基づいて同様に算出したものと比較した。

その結果、現状の公園配置は、夜間人口に基づく算出結果より、昼間の時間帯における人口に基づく算出結果で見たほうが、一人あたり公園面積のバランスが取れている傾向が得られた。このように、従来夜間人口に基づいて検討されてきた都市施設の再配置計画を、実際の利用が多い時間帯の人口に基づいて検討することにより、実ニーズとマッチした施設配置を行うことができるようになることが期待される。

(3)交通 -居住エリア別人口を用いた公共交通ニーズの把握-

モバイル空間統計の人口構成に基づき、中心市街地滞在者の居住エリア別の人口を時間帯ごとに推計した。この推計結果と中心市街地にアクセスするバス停別のバス便数のデータを同一地図上にプロットすることで、中心市街地への移動人口が多いにも関わらず、中心市街地に直接アクセスするバス便数の少ない地区を抽出することができた。しかし、モバイル空間統計では実際の移動手段までは把握できないため、これらの結果を実際のニーズとしてとらえるためには、他の交通機関の状況の考慮など、さらなる分析を行う必要がある。

(4)拠点振興 -属性別・曜日別の来街状況把握-

モバイル空間統計の人口推移と人口構成を組み合わせることにより、柏市外の各市町村から中心市街地へ流入する人口（来街者数）の変動と、年齢層・性別ごとの来街者数の変動を、平日と休日のそれぞれで推計した。

その結果、松戸市・我孫子市からの来街者数は平日と休日で大きな変化がないのに対して、流山市からの休日の来街者数は平日に比べて大きく落ち込む傾向が得られた。この結果から、流山市からの平日の通勤、通学者に比べて、休日の買物客を取り込めていないのではないかという推察が可能となる。このように、個々の来街者の実際の目的はモバイル空間統計では把握できないながらも、平休日の人口を比較することで、買物人数の傾向を推察することが可能であることがわかった。

また、年齢層・性別ごとの推計結果からは、20代女性の来街者数は平日より休日に大幅な増加を示しているのに対し、60代女性は対照的に大きく減少している傾向が得られた。このように、属性別・曜日別の特性を把握することにより、地域や年代、性別に応じたきめ細かな市街地活性化施策が検討できるようになると期待される。

防災計画分野への応用

2011年3月11日の東日本大震災に起因した、首都圏での大量な帰宅困難者の発生は記憶に新しい。首都圏では、首都直下地震に起因する帰宅困難者問題に対して活発な議論が進められている。本研究では、モバイル空間統計を用いて首都直下地震時に東京都で発生が予想される帰宅困難者や徒歩帰宅者に関する状況を推計し、防災計画分野における有用性を検証した。

なお、本研究での帰宅困難者の定義としては、内閣府や東京都が利用している定義と同じものを用いた。すなわち、地震発生時に滞留する人口を、居住エリアまでの距離を基準として、(1) 20km を越える滞留者は

全員帰宅困難者，(2) 10km 未満は全員帰宅可能，(3) 10km-20km までは 1km を越えるごとに，個々人の体力などの影響で帰宅困難者になる割合が 10%増加する，との仮定に基づいて推計を実施した。

(1) 地域ごとの帰宅困難者数

まず，モバイル空間統計により，東京都内の各市区町村（滞在エリア）における平日・休日の時間帯ごとの人口を，居住する市区町村（居住エリア）別，性別，年齢層別に推計した。この滞在エリアごとに推計された，居住エリア別，性別，年齢層別の人口を滞留人口と呼ぶものとする。その後，前述の帰宅困難者の定義に従い，それぞれの滞在エリアから居住エリアまでの距離に基づき，滞留人口から帰宅困難者数を算出した。

その結果，帰宅困難者数の東京都全域での総和が最大となるのは平日の 15 時台と推定され，その数は約 425 万人と推計された。本結果を活用することにより，帰宅困難者を一時待機させる施設の必要数などの検討が可能となる。特に，既存の被害想定では分からなかった性別・年齢層別の推計値，さらには近隣県以外からの外来者数の推計値により，属性を考慮した支援策など，対策の具体化が期待できる。

(2) 外出先で帰宅困難となる住民数

前述の帰宅困難者数を，居住エリアごとに集約することにより，各市区町村の住民がどの地域でどのくらいの数，帰宅困難となるのかを把握することができる。本研究では，国道 20 号線沿いの 11 市区（新宿区，渋谷区，杉並区，世田谷区，三鷹市，調布市，府中市，国立市，立川市，日野市，八王子市）について，それぞれの市区における住民がどこでどのくらい帰宅困難となるのかを算出した。

その結果，区部では全住民数と比較して帰宅困難となる住民数は相対的に少なく，都心部から遠ざかるほど相対的に多くなる傾向や，それぞれの市区の住民がどの滞在エリアで帰宅困難となるかに関する傾向が定量的に得られた。既存の被害想定は，どの市区で帰宅困難者が何名発生するのかという分析に限られていたが，新たに外出先で帰宅困難者となる住民の人数を算出することで，住民を支援するための施策（バスなどの代替輸送手段の提供など）の検討への活用が期待できる。

(3) 徒歩帰宅者数の時間推移

前述の滞留人口について，滞在エリアから居住エリアまで徒歩で帰宅した場合の移動状況をシミュレーションすることにより，発災後に各市区町村に流入する徒歩帰宅者数の時間推移を算出した。徒歩帰宅のシミュレーションにあたっては，それぞれの滞留者は，滞在エリアの面積重心から居住エリアの面積重心までを，最短経路となる一般道を経由して帰宅するモデルとし，徒歩速度は時速 4km とした。

その結果，各区市を通過する徒歩帰宅者数の総数に加え，発災後のピーク時間帯が把握でき，一部区市では発災直後よりも数時間経過した後の方が，徒歩帰宅者の流入数が大きくなることも分かった。区市別の徒歩帰宅者数はこれまで分析されておらず，本結果は，徒歩帰宅者を支援する災害時帰宅支援ステーションの整備の過不足評価などへの応用が期待できる。

3 新しい統計に期待される要件の抽出

時代の変遷に伴って発生しつつある新たな統計ニーズを調査するため、国勢調査などの公的統計を先進的に活用しているユーザーに対してインタビューを実施し、モバイル空間統計を活用した新しい統計に期待される要件を抽出した。

3.1 インタビューの概要

インタビューは、現在国勢調査などの統計を活用している4企業、1大学、1自治体に対して実施した。2012年6月から7月にかけて、各団体を訪問してモバイル空間統計の概要を説明し、現在の統計情報の利用状況、モバイル空間統計の活用の可能性、要求・要望などをヒアリングした。具体的には下記の4つの質問を中心に対話し、統計へのニーズを洗い出した。

- (1) モバイル空間統計を自社の業務（研究分野）で活用できると思うか。活用できるとすれば、どのように使用するか。
- (2) 自社の業務（研究分野）以外で、このモバイル空間統計が活用できそうな業務（分野）はあると思うか。それはどのような業務（分野）か。
- (3) モバイル空間統計の提供形態として、どのような形態を希望するか。
- (4) モバイル空間統計をより有効活用するために、さらに必要と思われる情報：機能はあるか。それはどのようなものか。

インタビューを実施した各団体の概要を下記に紹介する。

(1) 株式会社 J P S (Japan Planning Systems)

GIS マーケティングの会社であり、マーケティング諸活動支援、コンサルティング業務、ソフトウェア販売、統計データや地図データの販売などを行っている。

(2) 株式会社野村総合研究所

大手総合コンサルティング、IT ソリューション会社である。ビッグデータの利活用を専門とするコンサルタントから意見を伺った。

(3) 工学院大学

防災対策を専門としている准教授にインタビューを実施した。准教授は、すでにモバイル空間統計を利用した防災計画について検討の経験があり、実際の活用経験を踏まえての意見を伺った。

(4) 株式会社カーブスジャパン

40～70歳代をターゲットとした女性専用のフィットネスクラブを運営している。コンサル会社から派生した会社であり、同業他社と比較して出店計画などに統計データを先進的に活用している。

(5) 三井不動産株式会社

オフィスビル事業、住宅事業など不動産事業を行っている。三井不動産の社内シンクタンクとしての役割を担っている S&E 総合研究所に対してインタビューを実施した。

(6) 柏市

人口約 40 万人で、中核市、業務核都市に指定されている。商工振興課と防災安全課に対してインタビューを実施した。

3.2 インタビュー結果

各団体からのインタビュー結果を紹介する。

(1) 株式会社 J P S

マーケティングにおける具体的な使い方として、スーパーなどの新聞広告で早朝にチラシを見て、何時ごろ人が動くかを見て特売の時間を決める事や、車の出張展示などを 30～40 代男性の多いところで行うような事が挙げられた。これらの用途では、「特定の日」の人口はどうであったかということとは重要ではなく、ある「時期」ではどれほどの人がいることが期待されるかが重要となる。つまり、1 時間毎の人口より、「昼夜」や「休日・平日」、「春夏秋冬」といった時期毎の人口のほうがより求められるとの意見が寄せられた。また、年齢区分はそれほど細かい区分への要求はなく、10 歳区分であれば十分であるとのことであった。居住エリア別人口については、当該店舗に多くの人 coming しているエリアに集中的に広告をすることを想定し、当該店舗から見てどちらの方向から人が来ているか、あるいはどの沿線の人が多く来店しているかなどを判別可能な解像度が要望された。具体的には、市区町村単位よりも細かい解像度、たとえば 1km メッシュ、5km メッシュ相当の大きさで居住エリアを区分できれば有用ということであった。

(2) 株式会社野村総合研究所

株式会社野村総合研究所とのインタビューにおいては、多くの分野についての活用可能性が示された。具体的には、商圈分析、プロモーション、イベントランキング、防災計画などが挙げられた。商圈分析については、JPS からのヒアリングと共通して、商圈の把握、休日・平日の違いが求められるのに加えて、大型商業施設が建つ前と建ったあとでどのように商圈が変わったかがわかることに価値があるということであった。プロモーションについては、例えば、空港から水族館に向かう人へのプロモーションなど、「多くの人を通る」経路に広告を出して、特定の場所に誘導するような応用ができるのでは、ということであった。イベントランキングについては、たとえば花火大会など全国各地で開催されているイベントに対して集客人数によってランキングした結果などの公表が、潜在ユーザーの興味を集める手段として有用なのではないか、ということであった。現在、各イベントの集客人数は主催者等により公表されているが、決まった測定方法が存在せず、各々異なった方法で集客人数をカウントしており、異なるイベントの集客人数を比較することができないという課題を持つため、もしイベントによらず同一の基準で各イベントの集客数をカウントすることが可能になるのであれば、いままで困難であった各イベントに対する「横並びの比較」ができるようになるだろう、とのことであった。また、各地のイベントを防災の観点から考えると、イベントによる実際の参加人数に合わせた物資の確保という活用方法が考えられるだろう、という意見が寄せられた。イベント実施時に地震などが発災した場合、イベントの集客数を考慮した数の物資を常備しておくような防災計画を立てればよいが、各地において大きなイベントは年に数回程度であり、通常期では過剰な量となる。そこで、モバイル空間統計で全国のイベントの例年の集客人数を把握しておき、イベント規模に合わせて開催

場所に必要な場所に物資を移動させていくなど、災害対策物資確保の最適化ができるのでは、ということであった。

(3) 工学院大学

モバイル空間統計を利用した防災計画の検討に関する経験を踏まえた活用方法や今後の要望を聞くことができた。防災分野では、帰宅困難者数を把握することが重要視されている。モバイル空間統計の居住エリア別人口を用いることにより、これらの帰宅困難者がどこからどれくらい来ているかを推計できる。これにより、発災時に帰宅困難者をどの方向に帰すようにすればよいかを把握できること、また、発災時刻はいつか分からないため、24時間の時系列で帰宅困難者数を把握できることなどは帰宅困難者対策の検討において有用であるとのことであった。一方、要望として、より細かい地理的解像度や、低年齢層・高年齢層の推計対象への追加があげられた。地理的解像度については、よりきめ細かい防災計画を立てることを考えた場合、建物単位に計画を立てるのが理想であり、その場合は建物単位の解像度が得られることが望ましいということであった。年齢区分については、災害弱者が多い子供や高齢者を把握することの重要性が挙げられた。また、提供形態について、「簡単にすぐ使える」ことが重要との観点から、オンラインでの提供や、データだけでなく「アプリケーションとのセット」で提供することが要望された。精度に関する要望は特に明確なものはなく、ある程度の精度があれば十分というものであった。

(4) 株式会社カーブスジャパン

カーブスでは小さなフィットネスクラブを数多く出店しており、その出店計画に国勢調査のデータを積極的に活用している。40～70歳代女性が主要ターゲットであるため、スーパーマーケットの近くに出店する戦略をとっているとのことであった。そのため、スーパーマーケットに集まっている人を取り込めているか判断をするために、スーパーマーケットに集まっている人数とカーブスに来ている人数の相関を取りたいという要望があった。また、過疎地への出店判断のため、該当地域がどの程度の速度で過疎化しているかを把握したいという要望があった。これらの用途においては、顧客のカーブスへの滞在時間が30～40分程度であるため、30分間隔の時間解像度が望ましいとのことであった。また、年齢区分は年代によって販売促進戦略が異なるので5歳～10歳区分が望ましいこと、商圈把握のために居住エリア別人口は町字の粒度が望ましいことなどが意見として挙げられた。

(5) 三井不動産株式会社

不動産に関する多くの利用用途がヒアリングできた。貸ビル事業において、テナントへの情報提供により、サービスを差別化していきたいというニーズがあるとのことであった。例えば、オフィスビルにどのような人が来ているかということや、災害対策として、ビルにどの程度の帰宅困難者数が発生し得るかという情報を提供することなどが活用方法として挙げられた。また、今後退職していくであろうシニア層の住宅需要を把握するため、現在どこからどこへ通勤しているか(可能であれば経路も含めて)把握できれば有用ということであった。さらに、地価のデータと組み合わせることにより、何時にどのような人がいるところの地価が上がっているかというような、地価と人口の相関をとり、町のポテンシャルを測ることができれば有用だということであった。地理的解像度については、貸ビル事業でのテナントへの情報提供を想定すると、ビル単位での情報提供が可能となるようなために解像

度が必要だろう、とのことであった。さらに、ビルの何階に何人いるかというところまでわかると理想的であるとの意見が寄せられた。また、シニア層の住宅需要については、駅勢圏単位で人口がわかる程度の解像度で十分であるが、移動経路の把握を考慮すると道路単位で人口が把握できることが望ましい、とのことであった。また、鉄道・バスの移動経路の把握という観点では、15分単位程度の時間解像度が必要であろう、との意見であった。

(6) 柏市

防災の課題として、震災時において鉄道が停止したときにどこのルートに代替輸送手段を設定するかということがあげられた。すなわち、柏市にどこからどれだけ人が来ているかや、柏市民がどこに行っているかを把握することができれば、効果的に代替輸送手段を設定できるため有用であるとのことであった。一方、災害弱者の多い低年齢層・高齢性の人口を知りたいという要望が寄せられた。精度については、最大の被害が想定できればよいため、厳密な値は必要なく、最大でどれくらいの人口なのかがわかれば十分ということであった。商業振興の観点からは、類似都市の政策の横展開に有用であろうとの期待が寄せられた。現在、施策の効果検証に関して共通基準がなく、各自治体が独自に実施している状態であるため、他の自治体で効果があったという施策が自分の自治体でも有効かどうか判断することができないという課題があるとのことである。モバイル空間統計のような人口データがあれば、これらを共通の指標で比較することができ、施策効果の検証や他自治体を参考にした政策立案において有用であろう、とのことであった。なお、その際には、人口だけでなく通過者数(入込数)が把握できると、従来の駅乗降客数などとの比較ができるために望ましいとの意見があった。

3.3 得られた知見

本インタビューにおけるインタビュー対象からの反応は、いずれも極めて好意的なものであり、モバイル空間統計のような時間単位での人口推移を推計する人口統計に対し、公共分野・産業分野・学術分野のいずれにおいても高い潜在需要があることが示唆された。また、6対象中5対象からデータの利用可否や本共同研究への協力・参画などが打診されるなど、将来的な期待に留まらず、直近の課題解決に用いたいとの強いニーズも見られた。

また、年齢別・性別、あるいは居住エリア別の人口構成が併せて得られることに対して、応用領域を大きく広げる重要な特性であるとの高い期待が寄せられた。実際に、インタビューの席においても具体的な想定応用について様々なアイデアがインタビュー対象から挙げられた。その一部を下記に示す。

- 帰宅困難者把握
- 災害時の交通手段情報の提供
- 商圈調査や出店判断
- 地価データとの組み合わせによる町のポテンシャルの計測
- 類似する他都市との比較による政策立案

どのくらいの時間解像度、すなわち「時間的な細かさ」を持つことが望ましいかについては、ほぼ現行のモバイル空間統計が提供する「一時間ごと」の時間解像度で十分との意見が多かった。ただし、一部には、より細かい時間間隔（15～30分ごと）で必要という意見や、その逆に（高い解像度のデータは作成・利用

コストが高額になるだろうとの懸念から）朝・昼・夕・深夜などの4区分でも十分に活用可能であるなどの意見も得られた。また、どのくらいの空間解像度、すなわち「地理的細かさ」が望まれるかについては、現行のモバイル空間統計の空間解像度（東京23区内で500mメッシュ程度など）で十分に活用可能であるとの意見が多い一方で、より細かい空間解像度が得られればさらに応用領域が広がるであろうとの期待も寄せられた。

それに対し、どのくらいの「正確さ」を備えるか、すなわち信頼性に対する要件については、いずれのインタビュー対象からも具体的な数値を伴う定量的な精度要求は得られなかった。その背景としては、いままでに無い種類の新しい人口統計であることから、具体的な必要精度についての「相場観」的なものが醸成されていないために責任を持った回答が困難であろうということ、および信頼性要件については実際に「使っていきながら」手探りで検討していくことになるのではないかとということ、などが一部から言及された。

その他、特筆すべき要望・意見として以下のようなものが挙げられた。

- 地方自治体における政策立案の観点からは、他の自治体と比較可能な共通基準として利用できることが望ましいこと。
- 再開発における開発効果の事前見積りなど、過去のデータから効果を「予測」するような応用への期待が高まるであろうことを考慮すべき。そのためには、過去の類似事例から効果が適切に予測できるよう、継続的に蓄積し、比較・照合可能な形で提供していくことが望ましいこと。
- 日本全国における、1時間ごと、かつ属性別の人口推移というのは、莫大なデータ量になる。紙ベースでの提供はもはや有りえず、オンライン提供は必須となる。また、その際に、そのような大量のデータを適切に扱えるようなリテラシーを持つ人は限られてしまうので、データを容易に利用可能にするための「アプリケーション」付きで提供することが望ましいこと。

これらを踏まえると、限られた対象からのインタビュー結果ではありながらも、本共同研究が狙いとする「官庁統計とモバイル空間統計に基づく新たな統計」に対する期待や需要は高いであろうと結論づけられる。また、必要となる特性や解像度についても、インタビューの対象により多少の振れ幅はありながらも、現行のモバイル空間統計が備える特性や解像度と大きくかけ離れた要求はなかった。その一方で、統計としての信頼性要件については、現時点で利活用サイドから定量的な基準を示すことは困難であり、提供サイドからデータ自体と併せてその信頼性の目安を示していくことが必要であろうということが示唆された。

3.4 検討方針と課題

前節までに示したインタビュー結果を踏まえ、本共同研究における検討方針を議論するとともに、その課題を抽出する。

まず、従来の公的統計を補完する「新たな統計」としてどのような統計が求められるかについて議論する。インタビューの結果として、モバイル空間統計が持つ「1時間ごとの人口推移」について「年齢層別、性別、および居住エリア別の人口構成」を得られるという特性について、高い評価を得た。したがって、それらの特性を有する人口統計、すなわち現在人口の時間ごとの推移を属性別に推計した人口統計を「新たな統計」として作成するものとする。

ここで、必要とする解像度について、一部のインタビュー対象からは現行のモバイル空間統計より細かい空間解像度・時間解像度への期待も寄せられたが、統計の悪用やプライバシーへの懸念があること、および「大量の細かい数字」を提供しても利用サイドで使いこなすのが難しいとの懸念が指摘されたことなどから、解像度の向上については今後の課題とし、本共同研究では扱わないものとする。

信頼性については、データ自体と併せてその信頼性の目安を利用者にとってわかりやすい形で示す必要がある。ここで留意すべき点として、現在日本国内で広く使われている人口に関する指標は、5年に1度実施される国勢調査に基づく夜間人口・昼間人口であり、提供時の混乱を防ぐ観点からこれらとの整合性について十分に説明可能であるよう配慮する必要があることが挙げられる。

そのため、本共同研究において、国勢調査による夜間人口とモバイル空間統計により推計された人口の差異について議論するとともに、これらを比較照合することにより信頼性検証を行ない、モバイル空間統計の信頼性の目安となる基準を導出する。ここで、モバイル空間統計の信頼性は、携帯電話基地局の設置密度に大きく依存するが、一般の利用者にとって基地局の設置密度を把握することは困難であるため、これを基準として信頼性の目安を示すことは不適切である。そこで、基地局の設置傾向（通信負荷の平準化のため、人口密度が高いエリアでは多くの基地局を設置し、人口密度が低いエリアでは少ない基地局を設置する傾向がある）に着目し、人口密度を基準とした、利用者にとってもわかりやすい「信頼性の目安」を導出する。

また、現在のモバイル空間統計は推計に利用する常住人口として住民基本台帳に基づく人口を利用しているが、これは住民票の移転漏れなどの理由から国勢調査による夜間人口と乖離がある。しかし、国勢調査は5年に1度しか実施されないため、モバイル空間統計の推計に国勢調査による夜間人口をそのまま用いると、国勢調査の実施年から離れるに従って増加する経年変化が無視できなくなる。そこで、国勢調査による人口を基礎として、その後の人口動向を「人口動態統計」、「出入国管理統計」、「住民基本台帳人口移動報告」などの他の人口関連資料から得て算出される「推計人口」を、住民基本台帳に基づく人口に替えて用いることが望ましいと考えられる。

その他の課題としては、データの提供形態についての指摘がインタビューで挙げられたが、これについては今後の課題とし、本共同研究では扱わない。また、技術面だけでなく制度面でもいくつかの課題が考えられるが、これらについても今後の課題とする。

4. モバイル空間統計の信頼性検証

モバイル空間統計は、携帯電話ネットワークの運用データに基づいて、あるエリア・ある時間にいる実際の人口（以降、現在人口と呼ぶ）を推計する新しい人口統計である。ここで、モバイル空間統計の統計としての信頼性には携帯電話ネットワークの特性に起因する様々な理由により制約がある。

たとえば、モバイル空間統計をどのくらいの地理的な細かさ（地理的解像度）で推計できるかについても、基地局の設置間隔に強く依存する。基地局の設置状況は地域によって大きく異なるため、同じ空間解像度であっても地域によってはモバイル空間統計の信頼性は大きく異なる。

同様に、モバイル空間統計をどのくらい細かい時間単位（時間的解像度）で推計できるかは、携帯電話ネットワークの運用データの更新頻度の偏りに強く依存する。そのため、時間解像度が1分単位と1時間単位とではモバイル空間統計の信頼性が大きく異なる。

モバイル空間統計の信頼性を検証するためには、上記時間解像度と空間解像度に着目して、現在人口とモバイル空間統計のずれを評価することが必要である。ここで、時間解像度と空間解像度の評価観点および評価に必要な評価用人口データ（現在人口を正しく反映している人口統計）の持つべき特徴について述べる。

- **空間解像度の観点**：行政界（都道府県や市区町村など）や標準地域メッシュ（1kmメッシュや500mメッシュなど）の地域区分で正しい人口分布が推計できるかを評価する。本評価を実施するために用いる正解データは、ある時点における行政界や標準地域メッシュ単位の現在人口を正確に捉えていることが必要である。
- **時間解像度の観点**：一日を通じた人口推移の把握に有用となる、1時間ごとの時間単位で正しい人口分布の推計ができるかを評価する。本評価を実施するために用いる正解データは、時間ごとに変化する現在人口の推移を正確に捉えていることが必要である。

本章では上記評価観点のうち、空間解像度の観点に対する信頼性評価の結果を議論する。評価にあたり、時々刻々と変化する現在人口の「正解データ」を準備することは非常に困難であるが、深夜帯においては現在人口と常住人口がほぼ一致することが期待される。そこで、本評価においては、2010年国勢調査に基づく地域メッシュ統計を比較対象とし、空間解像度に着目したモバイル空間統計の信頼性評価方法やその評価結果、評価結果から得られた知見についてまとめる。

4.1 2010 地域メッシュ統計について

地域メッシュ統計とは

地域メッシュ統計とは、緯度・経度に基づき地域をすき間なく網の目（メッシュ）の区域に分けて、約1km四方、約500m四方などに区切ったそれぞれの区域に関する統計データを編成したものである。地域メッシュ統計の特性としては、①地域メッシュは、ほぼ同一の大きさ及び形状の区画を単位として区分されているので、地域メッシュ相互間の事象の計量的比較が容易なこと、②地域メッシュは、その位置や区画が固定されていることから、市町村などの行政区域の境域変更や地形、地物の変化による調査区の設定変更などの影響を受けることがなく、地域事象の時系列的比較が容易なこと、③任意の地域について、その地域内の地域メッシュのデータを合算することにより、必要な地域のデータを容易に入手できること、④地域メッシュは、緯度・経度に基づき区画されたほぼ正方形の形状であることから、位置の表示が明確で簡便にできるので、距離に関連した分析・計算・比較が容易なことなどがあげられる。

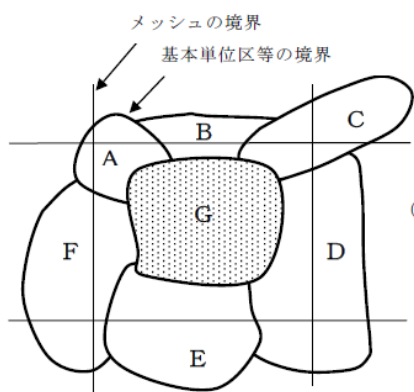
地域メッシュの区分方法としては、昭和48年7月12日 行政管理庁告示第143号「統計に用いる標準地域メッシュ及び標準地域メッシュコード」を用いている。この告示は昭和51年1月に日本工業規格（JIS）として制定（JIS X 0410）され、規定されている区画の種類は次の表のとおりである。

区画の種類	区分方法	緯度の 間隔	経度の 間隔	一辺の 長さ	地図との関係
第1次地域区画	全国の地域を偶数緯度及びその間隔(120分)を3等分した緯度における緯線並びに1度ごとの経線とによって分割してできる区域	40分	1度	約89km	20万分の1地勢図の1図葉の区画
第2次地域区画	第1次地域区画を緯線方向及び経線方向に8等分してできる区域	5分	7分30秒	約10km	2万5千分の1地勢図の1図葉の区画
基準地域メッシュ (第3次地域区画)	第2次地域区画を緯線方向及び経線方向に10等分してできる区域	30秒	45秒	約1km	
2分の1地域メッシュ	基準地域メッシュ(第3次地域区画)を緯線方向、経線方向に2等分してできる区域	15秒	22.5秒	約500m	
4分の1地域メッシュ	2分の1地域メッシュを緯線方向、経線方向に2等分してできる区域	7.5秒	11.25秒	約125m	

総務省統計局では、国勢調査と経済センサス（旧事業所統計調査）の結果に基づく地域メッシュ統計を作成しており、国勢調査は昭和40年調査から、事業所統計調査は昭和41年調査から地域メッシュ統計を作成している。

国勢調査地域メッシュ統計の作成方法

国勢調査のメッシュ統計は、国勢調査結果としては最も小さい地域の集計結果である基本単位区別集計結果から作成される。国勢調査の基本単位区とは「街区又は街区方式に準じ、道路や河川、鉄道など地理的に明瞭で恒久的な施設等により区分された区域の区画」である。住居表示実施区域では原則として街区で、1街区の面積は3,000～5,000m²、20～30住戸程度を含む区域であり、住居表示未実施地域では街区に相当する区画を設定している。なお、基本単位区は国勢調査の実施において、1人の調査員が調査を担当する区域である調査区を設定するための基本的な区域になる。



調査単位区の境界と地域メッシュの境界は一致していないので、複数の地域メッシュにまたがった調査単位区の調査結果をそれぞれの地域メッシュに配分することにより地域メッシュの統計数値を作成する。配分の方法は以下のような6種類の方法を順次採用して地域メッシュ統計を作成する。

- ① 基本単位区全域が一つの地域メッシュに含まれるものは、全ての結果数値を当該地域メッシュに配分する。

- ② ①以外の基本単位区のうち面積が0.05km²未満のものは、分割された各区域の面積割合に応じて配分する。
- ③ ①及び②以外の基本単位区のうち住宅建物がある場合、電子地図を利用し、それぞれの地域メッシュに含まれる戸数に応じて配分する。
- ④ ①～③以外の基本単位区のうち人口分布点を持つ基本単位区については、その人口分布点が含まれる地域メッシュに配分する。なお、人口分布点とは過去の地域メッシュ統計作成過程において基本単位区内に位置づけられた居住中心点である。
- ⑤ ①～④以外の基本単位区のうち事業所建物がある場合、電子地図を利用し、それぞれの地域メッシュに含まれる事業所数に応じて配分する。
- ⑥ ①～⑤以外の基本単位区については、基本単位区境界の幾何学的重心（図心）を求め、その図心を含む地域メッシュに配分する。

2010年国勢調査の地域メッシュ統計

国勢調査の地域メッシュ統計の作成は、基本単位区別集計結果の集計作業とともに基本単位区の境界データの確定を待って行われるため、公表は調査実施からおおよそ2年後以降となる。この研究に用いられている2010年国勢調査の地域メッシュ統計は2012年11月20日に公表されたものである。そのため、研究の初期においては、2005年国勢調査のメッシュ統計を用いて試験的に検証作業を行い、2010年国勢調査の地域メッシュ公表後にデータを差し替えて検証作業を行った。

地域メッシュ統計利用上の留意点

地域メッシュ統計を利用する上で留意すべきこととしては、地域メッシュが緯度・経度をもとに設定されているため、地球上のどこにあるかによって大きさが異なることである。基準地域メッシュの形は一辺がほぼ1kmの正方形といわれているが、我が国は南北に長い国であるため、例えば北海道庁の所在する札幌市の基準地域メッシュの横の長さが1,018mであるのに対し、沖縄県庁が所在する那覇市の基準地域メッシュの横の長さは1,249mで231mの違いがあり、面積にして那覇市の基準地域メッシュは札幌市の基準メッシュに比べ約1.2倍大きくなる。このように、地域メッシュは、区分が同じであっても位置によって大きさが異なるため、特に遠距離間で地域メッシュ統計を比較する際には、形状と大きさに注意を払う必要がある。

また、国勢調査地域メッシュ統計は個々の世帯の位置に基づいて作成されるものではなく、前述のように国勢調査の最小集計地域である基本単位区別の集計結果を地域メッシュに配分することによって作成されているため、配分のしかたによって精度が低下していることに注意が必要となる。

4.2 検証手法

本評価では、都道府県や市区町村などの行政界単位や、1次メッシュ(おおよそ80km四方)～4次メッシュ(おおよそ500m四方)単位など、応用分野で用いられる様々な空間解像度のモバイル空間統計について、その信頼性の程度を定量的に明らかにする。

評価データについて

モバイル空間統計の信頼性を評価するにあたっては、本来は「正解データ」を準備し、その正解データの値とモバイル空間統計による推計値との差異を分析し、評価・議論することが望ましい。ここで、モバイル空間統計における正解データとは、真の現在人口、すなわち、あるモバイル空間統計の推計対象エリアにおいて、推計対象となった時間帯に「実際にそのエリア内にいた人数」が相当する。

しかし、実際には正解データとして正確な「真の現在人口」を得ることは非常に困難である。たとえば、このようなデータを得るための手段としては、人手による実測値を用いることが考えられる。しかし、あるエリア・ある時間帯のモバイル空間統計の信頼性を評価するために、実際にそのエリアで当該時間帯に存在する人数を全員分正確に数えることは、たとえそのエリアが(本評価における最小単位である) 500m メッシュであったとしても現実的ではない。

そこで、本評価では、モバイル空間統計が推計の対象とする現在人口は、ほとんどの人が帰宅している時間である深夜帯においては常住人口(「どのエリアにどのくらいの人が住んでいるか」を表す人口)と近い値になることに着目し、現在人口の近似値として、常住人口に関する統計として信頼性が高い国勢調査による夜間人口を用いる。具体的には、直近の国勢調査である 2010 年国勢調査に基づく都道府県/市区町村ごと夜間人口および小地域メッシュ統計と、国勢調査の実施日である 2010 年 10 月 1 日におけるモバイル空間統計による現在人口の推計値を比較評価する。なお、モバイル空間統計による推計値は時刻により変動するが、本評価では、常住人口と最も近い値をとると思われる、午前 4 時台の人口を用いる。

ここで、モバイル空間統計の信頼性評価のための比較対象データとして国勢調査に基づく常住人口を用いるにあたっては、現在人口と常住人口の差異を考慮する必要がある。すなわち、ほとんどの人が帰宅する深夜帯においては現在人口と常住人口の値は近いものとなるが、完全に等しくなることは限らない。たとえば、泊まりがけの出張や旅行、深夜残業の人などの存在により、実際には両者には差が生じる。

したがって、国勢調査の人口を用いてモバイル空間統計の信頼性を定量的に評価するにあたっては、評価の結果として表れる差異は、真の現在人口に対する差異に比較して、より大きなものとして見えることに留意が必要である。具体的な影響の例については、各評価結果の中で説明する。

評価指標について

前述のとおり、本評価では国勢調査による常住人口(以下、国調人口と呼ぶ)とモバイル空間統計により推計された深夜帯の現在人口(以下、推計現在人口と呼ぶ)を比較し、その差異の大きさにより空間解像度の信頼性を評価する。ここで、「差異の大きさ」を定量的に評価するために、指標として「偏差率」を導入する。

偏差率とは、あるエリアに関する 2 種類の人口(国調人口と推計現在人口)の関係が、理想的な状態(両者の値が一致する)からどれだけずれているかを表す評価指標である(図 4.1)。すなわち、エリア i の国調人口を k_i 、推計現在人口を m_i とし、それらの平均を μ_i とする ($\mu_i = (k_i + m_i)/2$)。エリア i の偏差率 δ_i とは、推計現在人口 m_i と平均 μ_i の差(平均偏差)の、 μ_i に対する比として定義される ($\delta_i = (m_i - \mu_i)/\mu_i$)。ここで、 $\mu_i = (k_i + m_i)/2$ を代入すると、偏差率 δ_i は以下の式で表される。

$$\delta_i = \frac{m_i - k_i}{k_i + m_i}$$

直感的な理解としては、図 4.1 に示す通り、偏差率 δ_i とは座標 (k_i, m_i) から直線 $k_i = m_i$ に対して垂線を下ろしたときに、その垂線の長さ b の、直線 $k_i = m_i$ と垂線との交点から原点までの距離 a の比に相当する。

定義式からわかるように偏差率 δ_i は-1以上1以下の数値をとり、0に近いほど両者の値に差異が小さく、-1または1に近づくにつれて差異が大きくなることを表わす。また、偏差率が正の場合は推計現在人口が国調人口に対して過大推計されており、負の場合は逆に過小推計されていることを示す。

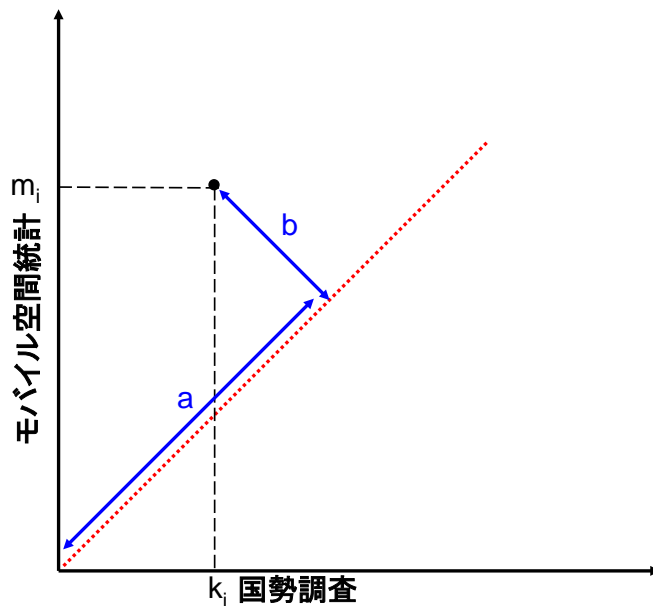


図 4.1 偏差率

評価内容について

本評価では、空間解像度とモバイル空間統計人口の信頼性の関係を明らかにするために、以下の観点から分析を行う。

- 1) 空間解像度に対する全体傾向の評価
- 2) 3次・4次メッシュに関する地理的な傾向の評価
- 3) 3次・4次メッシュに関する推計人口と信頼性との関係の評価

まず、1) 空間解像度ごとの信頼性に対する全体傾向の評価において、モバイル空間統計の応用分野でよく用いられる以下の空間解像度について、それぞれ偏差率がどのような分布を示すかを評価する。

- a) 行政界単位(都道府県単位，市区町村単位)
- b) 標準メッシュ単位(1次メッシュ単位～4次メッシュ単位)

なお、1次～4次メッシュ単位は、それぞれ以下の大きさのメッシュ単位である。

- 1次メッシュ：およそ 80km 四方
- 2次メッシュ：およそ 10km 四方
- 3次メッシュ：およそ 1km 四方

- 4次メッシュ：およそ 500m 四方

なお、3次メッシュと4次メッシュの評価にあたっては、秘匿処理による影響を排除する観点と、山間地など明らかに3次・4次メッシュでの人口推計が困難なエリアによる影響を排除する観点から、モバイル空間統計1人以上かつ、推計現在人口が500人以上となるメッシュを評価の対象とした。

次に、2) 3次・4次メッシュに関する地理的な傾向の評価において、上記のうちで細かい空間解像度を持つ単位である3次メッシュ、4次メッシュ単位でのモバイル空間統計人口について、信頼性に関する地理的な傾向を評価する。

最後に、3) 3次・4次メッシュに関するモバイル空間統計人口と信頼性との関係の評価で、同じく3次・4次メッシュを対象として、モバイル空間統計人口がある値をとったときに、それがどの程度の信頼性を持つかを評価する。

4.3 検証結果

以下、前述の1)~3)の各評価について、それぞれの評価結果を示す。

空間解像度に対する全体傾向

前節で述べたそれぞれの空間解像度ごとに、国調人口と推計現在人口との偏差率を計算した。まず全体的な傾向を把握するために、各空間解像度で偏差率±20%以下となったエリアの割合、および±10%以下となったエリアの割合を表4.1に示す。

表 4.1 空間解像度ごとの偏差率の傾向

エリアの単位	エリアの大きさ	対象エリアの総数	偏差率±20%以下の エリアの割合	偏差率±10%以下の エリアの割合
都道府県界	各都道府県の面積	47	100.0%	100.0%
市区町村界	各市区町村の面積	1901	95.9%	80.0%
1次メッシュ	およそ80km四方	136	83.1%	60.0%
2次メッシュ	およそ10km四方	3130	64.6%	50.0%
3次メッシュ	およそ1km四方	69531	56.9%	30.0%
4次メッシュ	およそ500m四方	137926	50.1%	30.0%

表 4.1 からわかるように、偏差率±20%を基準として見た場合、全ての空間解像度において評価対象エリアの50%以上が基準内となっている。特に、行政界単位(都道府県および市区町村単位)では、島嶼部や過疎地域、および後述の経年変化による影響が著しい一部の市区町村を除き、ほぼ全てのエリアが基準を満たす。

また、都道府県単位では偏差率が±10%以下のエリアの割合も100%であったが、他の空間解像度では、解像度の細かさに応じて割合に落ち込みが見られる。

偏差率の大きさと、その偏差率以下となったエリアの割合について、それぞれの空間解像度ごとにグラフで表したもの(x軸を偏差率の絶対値、y軸を累積確率とした、偏差率の絶対値の累積確率分布)を図4.2に示す。表4.1における偏差率±20%以下のエリアの割合は、図4.2において、各空間解像度のグラフと直線 $x = 0.2$ の交点の値に相当する。この結果から、市区町村や一次・二次メッシュは偏差率±20% ($x=0.2$) 付近、

三次・四次メッシュは偏差率±30% ($x=0.3$)付近までエリアの割合が急激に増加するが、それ以降はなだらかに増加していく傾向がわかる。すなわち、偏差率を±30%とした場合、その基準を満たすエリアの割合は、3次メッシュ・4次メッシュを含め、評価対象とした全ての空間解像度で5割以上となることがわかる。

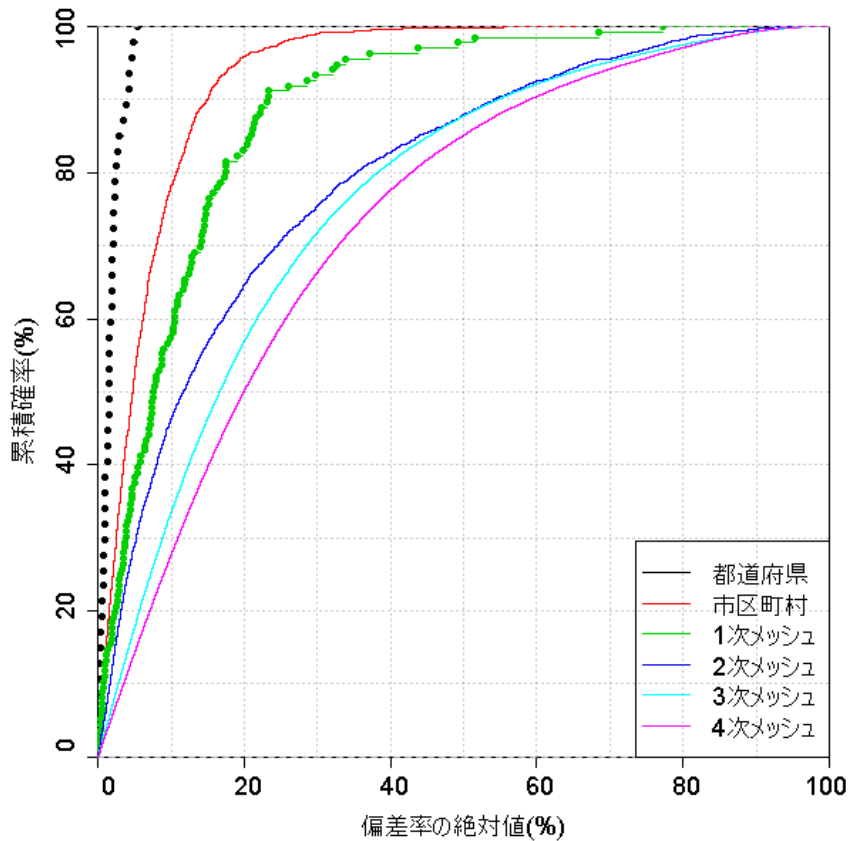


図 4.2 偏差率の絶対値の累積分布

3次・4次メッシュに関する地理的な傾向

次に、細かい空間解像度である3次メッシュと4次メッシュ単位の推計現在人口に着目し、各エリアにおける偏差率の地理的な傾向について評価する。

図 4.3(a)は3次メッシュにおける各メッシュの偏差率について、その分布を地図上に東京近郊を中心として可視化したものである。また、図 4.3(b)は同様に4次メッシュについて可視化したものである。

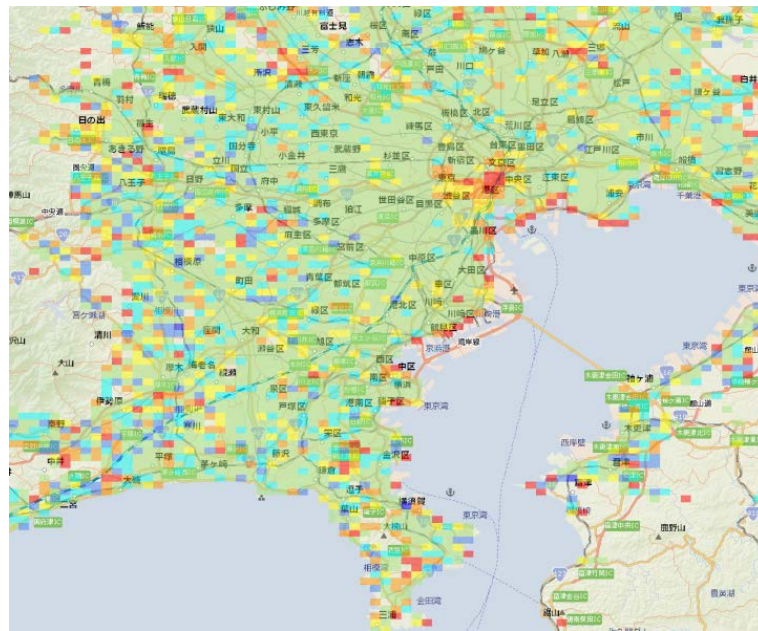


図 4.3(a) 偏差率の地理的分布(3次メッシュ)

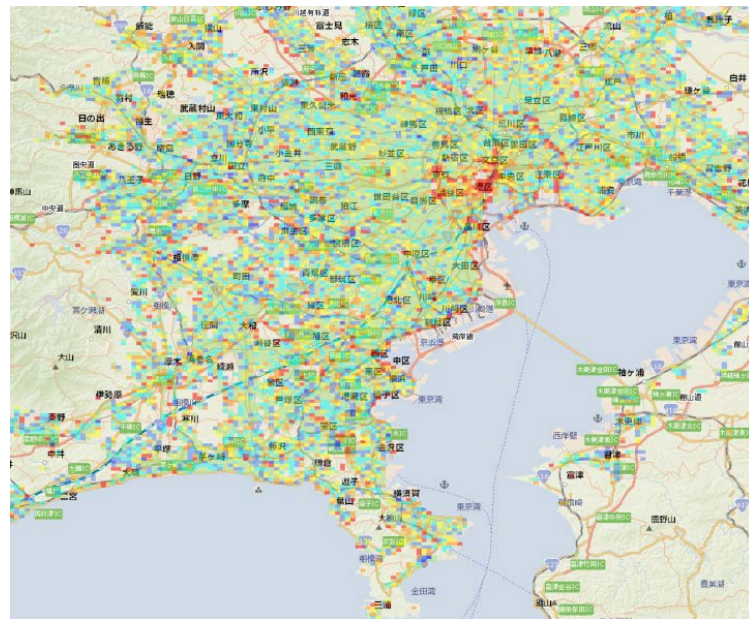


図 4.3(b) 偏差率の地理的分布(4次メッシュ)

図 4.3(a)を見ると、平野部では一部の例外を除き、ほとんどのエリアにおいて偏差率が $\pm 10\%$ の間の値を取ることがわかる。また、山間部や沿岸部などにおいて、偏差率のばらつきが大きくなる傾向が見てとれる。図 4.3(b)の傾向もほぼ図 4.3(a)と同じではあるが、全体的に図 4.3(a)と比較してばらつきが大きい傾向にある。また、図 4.3(a)ではあまり強く見られなかった、荒川・多摩川などの広い河川敷を持つ河川沿いで偏差率のばらつきが見られる。

また、特徴的な現象として、図 4.3(a),(b)のいずれにおいても、東京の都心部(千代田区，中央区付近)で偏差率が著しく大きい地域が見られることがわかる。すなわち、国調人口に対しモバイル空間統計人口が大幅に上回る値をとっている。この現象は、主に常住人口と現在人口の定義の違いによる影響と考えられる。

前述の通り、国勢調査の夜間人口が常住人口を表したものであるのに対して、現在人口の推計値であるモバイル空間統計は、居住しているかいないかにかかわらず、調査の対象時点でそのエリアに滞在している人口を推計する。千代田区などではドーナツ化現象により常住人口が少なめであるのに対し、オフィスや店舗、ホテルなどの非居住施設が密集しており、定常的に出張者や観光客、深夜労働者などの夜間滞在者が多く存在する。

具体例として、図 4.4 に大手町を含むエリアでの推計現在人口の、居住エリア別の人口構成を示す。同エリアは東京都千代田区と中央区とに跨るエリアであるが、両区を居住エリアとする人口は、全体の 1/4 に満たない(23%)。すなわち、3/4 以上が千代田区・中央区以外から同エリアに訪問・滞在している人であると推定される。この結果から、図 4.3(a)および図 4.3(b)に見られる都心部における偏差率の増加は、これらのエリアにおいては住民に対して夜間滞在者が多く存在するという状況が反映された結果であると考えられる。

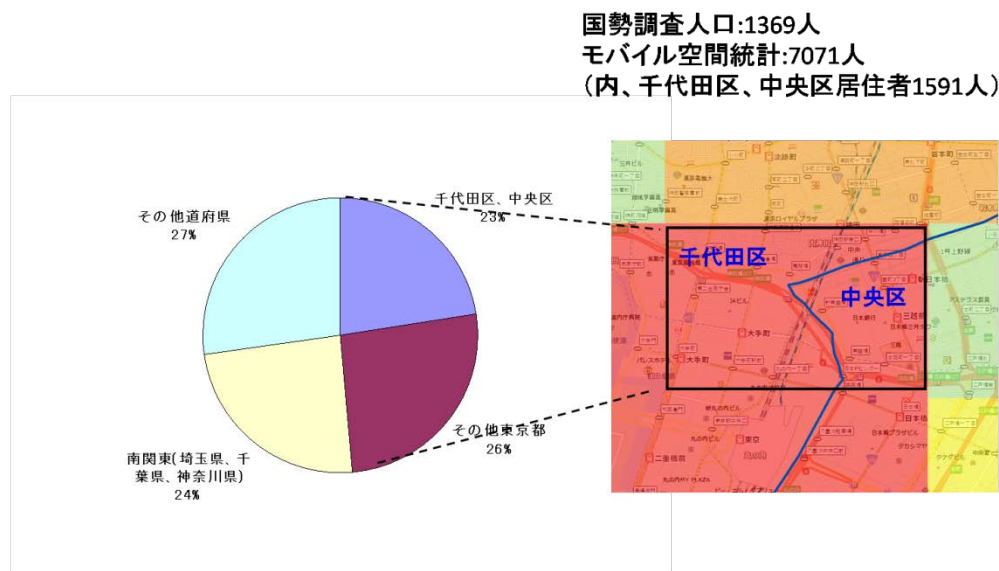


図 4.4 大手町を含むメッシュの居住エリア別人口構成

3次・4次メッシュに関する推計人口と信頼性との関係

最後に、モバイル空間統計による推計現在人口と信頼性との関係进行评估する。評価対象は、地理的な傾向の評価と同様に、3次メッシュ・4次メッシュとする。

図 4.5(a)は、横軸に推計現在人口、縦軸に偏差率をとり、3次メッシュにおける推計現在人口と偏差率の関係を図示したものである。灰色の各点がそれぞれ1つのメッシュに対応している。また、偏差率の分布に関する両側%分位点を曲線近似した結果を併せて図 4.5(a)に表す。この曲線を見ることで人口密度と偏差率のばらつきの関係を把握できる。図 4.5(b)は同様に、4次メッシュを対象として図示したものである。

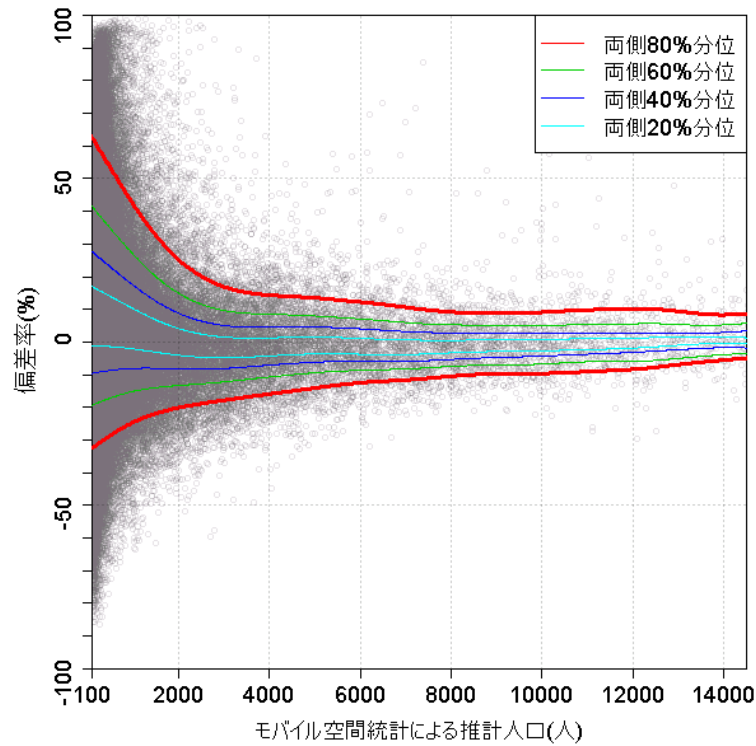


図 4. 5(a) 推計現在人口と偏差率の関係(3次メッシュ)

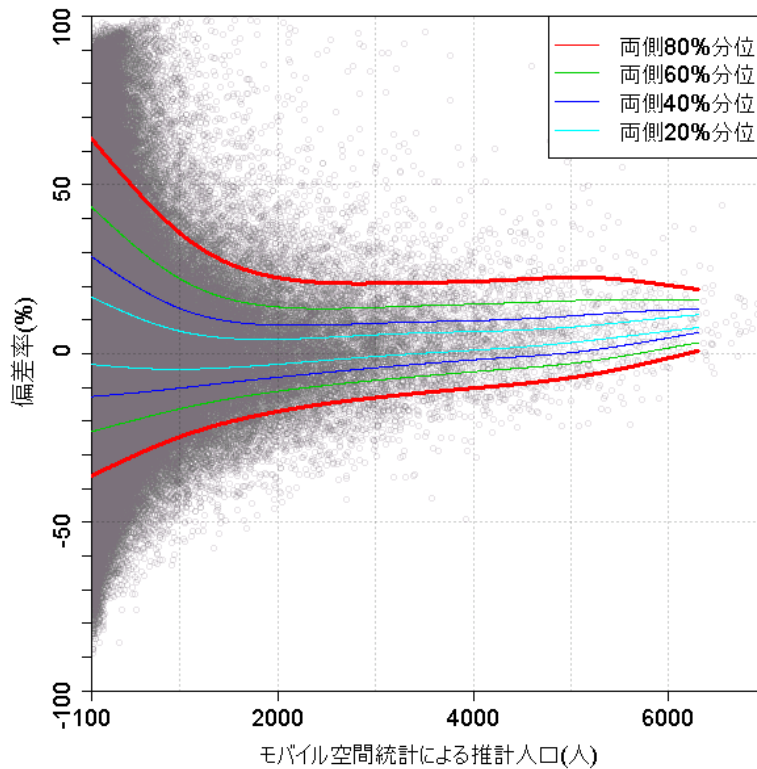


図 4. 5 (b) 推計現在人口と偏差率の関係(4次メッシュ)

図 4.5(a), (b)を見ると、偏差率の分位点を結んだ曲線が、正領域と負領域ともに人口の増加に伴い偏差率 0 に近づくことから、人口の増加に伴って偏差率のばらつきが減少することがわかる。すなわち、人口が多い(人口密度が高い)エリアほど、モバイル空間統計はより正確に人口を推計できる。これは、人口が密集している地域ほど、通信負荷の平滑化のために基地局をより多く設置する必要があり、その結果としてモバイル空間統計の空間解像度が向上することを反映している。

また、図 4.5(a)によれば、三次メッシュに関しては推計現在人口が 4,000 人程度以下の区間においては、人口が増えるにつれ偏差率は急速に 0 に近づいており、それ以上のエリアでは、8 割のエリアがおおよそ偏差率±20%以内になることがわかる。すなわち、たとえば人口集中地区(人口密度 4,000 人/km²が連続して 5,000 人以上の地域と定義される)において、3 次メッシュ単位でのモバイル空間統計は高い信頼性を持つと言える。

その一方、図 4.5(b)によると、4 次メッシュにおいては推計現在人口の増加に対して偏差率のばらつきがなだらかに減少しており、2,000 人程度以上のエリアにおいて、8 割のエリアがおおよそ偏差率±20%以内となる。なお、人口集中地区の基準である人口密度 4,000 人/km²は、4 次メッシュ(およそ 500m 四方)において人口約 1,000 人のエリアに相当する。そのため、利用目的にもよるが、4 次メッシュでのモバイル空間統計の利用は、人口集中地区より高い人口密度、具体的には人口集中地区の二倍程度の人口密度を持つ地域を対象とすることが望ましいと考えられる。

4.4 考察

モバイル空間統計の信頼性について、空間解像度との関係を中心に評価を行なった。

その結果、以下の知見が得られたと考えられる。

- モバイル空間統計は、都道府県・市区町村などの行政界単位、また 1 次メッシュ、2 次メッシュなどの大きなメッシュ単位では高い信頼性を持つと考えられる。
- 3 次メッシュ、4 次メッシュでは、東京近郊の平野部など、ある程度以上に人口が集中している地域では高い信頼性を持つ。ただし、沿岸部や山間部、河川敷周辺など、地形の変化や人口のばらつきが大きい地域では、推計値の扱いに留意が必要となる可能性がある。
- 特に 3 次メッシュでは、人口集中地区と定義される地域においては高い信頼性を持つ。
- 4 次メッシュも人口が集中している地域ほど高い信頼性を持つが、人口集中地区の 2 倍程度の人口密度を持つ地域を対象とすることが望ましいと考えられる。

ただし、前述の通り、常住人口である国勢調査に基づく人口を比較対象として評価を行った関係から、本評価の結果は実際よりも悲観的なものとなっている可能性がある。特に、都心部などにおいては常住人口と現在人口の定義の違いによる影響による事象が強く観察された。今後、これらの影響を除いたより正確な信頼性の評価方法を検討していくとともに、本評価による知見を活かしてより信頼性が高いモバイル空間の推計方式の検討へと繋げていくことが期待される。

5. ベンチマーク人口の評価と選定

日本全国を網羅する人口を推計するためのベンチマークとなりうる統計には、総務省自治行政局が住民基本台帳に基づいて算出する「住民基本台帳人口」と総務省統計局が直近の国勢調査結果を基に推計して算出する「推計人口」がある。ここでは、モバイル空間統計の作成においてどちらの人口をベンチマークとすることがより適切であるかを検証した。

5.1 住民基本台帳人口について

住民基本台帳人口は、総務省自治行政局が住民基本台帳法（昭和42年法律第81号）に基づき、毎年3月31日現在の住民票に記載されている者の数について、市区町村長からの報告に基づきとりまとめたものであり、毎年8月上旬に公表されている。

住民基本台帳は住民の届出によって作成されるものであるため、地方から大都市に出て下宿している大学生が現住所を親元のままにしてあるケースなど、実際の居住地と異なる場合がある。

5.2 推計人口について

推計人口は、総務省統計局が5年ごとに実施される国勢調査による人口を基礎として、その後の人口動向を「人口動態統計」、「出入国管理統計」、「住民基本台帳人口移動報告」などの他の人口関連資料から得て、毎月1日現在の人口を算出している。都道府県、男女、年齢階級（5歳階級）別の人口は各年10月1日現在の人口のみが翌年4月に公表されている。

なお、住民基本台帳人口が人単位で公表されるのに対し、推計人口の公表値は千人単位になっている。

また、国勢調査の人口を基礎としてその後の5年間で推計していくため、国勢調査の実施年から離れるほど推計誤差が増加する傾向がある。

5.3 検証手法

検証は2010年国勢調査結果と、住民基本台帳人口と推計人口の数値との差の大きさを比較することとし、モバイル空間統計を作成する際に用いるベンチマークの区分である都道府県、男女、年齢階級（5歳階級）別の数値の乖離の大きさを比較することとした。

この3つの統計の時点は、国勢調査が2010年10月1日であるのに対し、住民基本台帳人口では利用できる直近の公表値である2010年3月31日現在の人口を用いることにした。

推計人口は毎年10月1日現在で公表されるが、国勢調査の実施年には推計値は公表されないことから、直近の推計値として2009年10月1日現在の推計人口を用いることとした。

また、2010年国勢調査結果には年齢不詳の人口が含まれていることから、比較の対象としては推計人口の基準として使用するために、2010年国勢調査人口に含まれる国籍及び年齢不詳人口をあん分して作成された2010年国勢調査による基準人口（平成22年10月1日現在）を使用した。

しかし、年齢階級別（5歳階級）の人口を比較する場合、各年齢階級（コーホート）間の人口の違いにより、調査時点の違いが大きく影響を与える。例えば、2010年国勢調査の各歳別結果によると60歳人口は2,066,423であるのに対し、65歳人口は1,426,865となっており、約60万の差がある。これはその後1年

間に人口変動（自然増減及び社会増減）がなかったとしても、60～64歳人口は2011年では約60万人増加することとなるため、単純な比較は不可能である。

このコーホート間の人口差を調整するため、2010年国勢調査の各歳別人口を用いて、国勢調査の結果を比較する統計の時点にあわせる処理を行った。具体的には、国勢調査人口と住民基本台帳人口の比較は2010年3月1日現在、国勢調査人口と推計人口の比較は2009年10月1日現在で行うために、国勢調査人口を次のように調整した。

住民基本台帳人口との比較（2010年3月1日現在）

$$P_{ij} = P_{ij} - \frac{P_i}{2} + \frac{P_{j+1}}{2}$$

推計人口との比較（2009年10月1日現在）

$$P_{ij} = P_{ij} - P_i + P_{j+1}$$

P_{ij} : 年齢階級 i 歳～ j 歳の人口

P_i : i 歳の人口

P_{j+1} : $j+1$ 歳の人口

5.4 検証結果

住民基本台帳人口と推計人口の傾向

住民基本台帳人口（2010年3月1日現在）と推計人口（2009年10月1日現在）を国勢調査結果（2010年10月1日現在）と直接比較してみたそれぞれの統計の傾向は以下のとおりである。

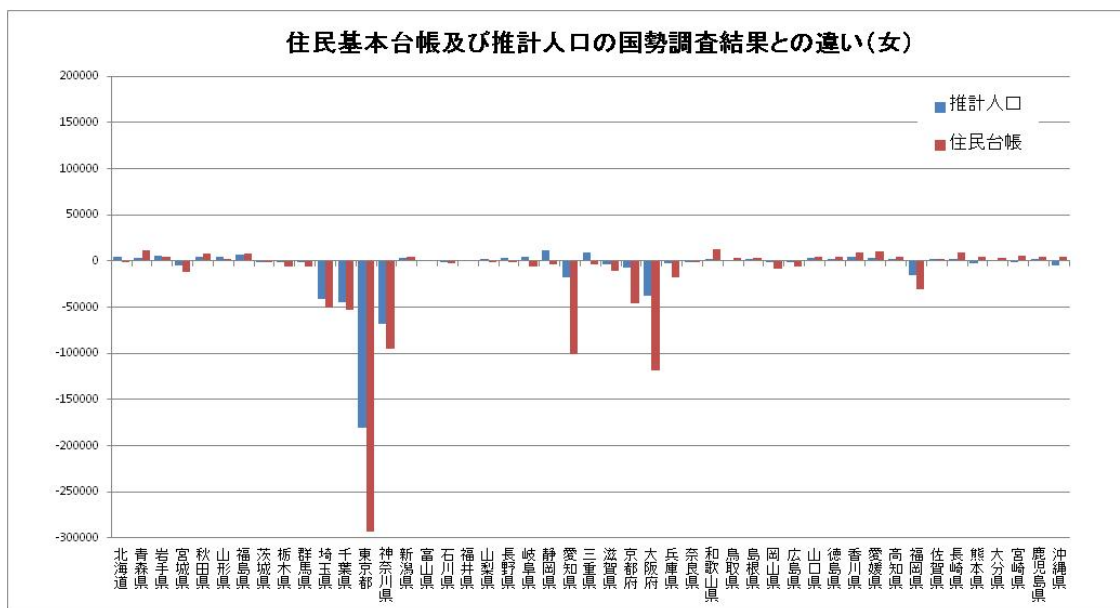
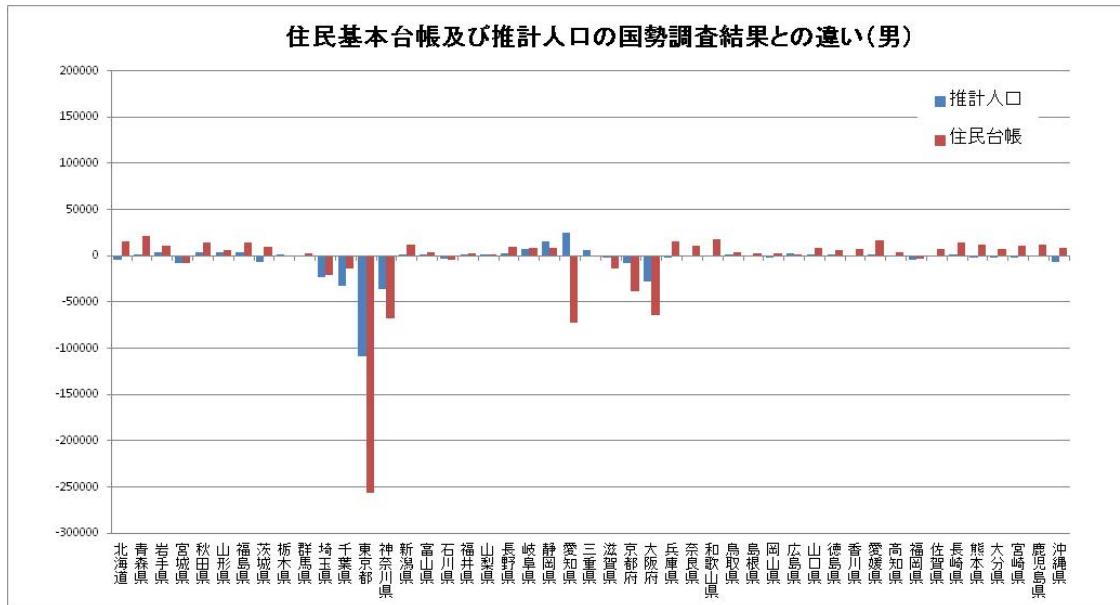
- ・ 総人口及び男女別人口

住民基本台帳人口、推計人口ともに総人口、男女別人口が国勢調査結果より小さくなっている。総人口で住民基本台帳が約100万、推計人口が約50万過小となっており、住民基本台帳人口の差が大きい要因として女性の差が大きくなっている。

		国勢調査 2010. 10. 1	住民基本台帳人口 2010. 3. 1	推計人口 2009. 10. 1
人口	総数	128,057,352	127,057,860	127,510,000
	男	62,327,737	62,080,435	62,130,000
	女	65,729,615	64,977,425	65,380,000
国勢調査 との差	総数	-	-999,492	-547,352
	男	-	-247,302	-197,737
	女	-	-752,190	-349,615

・ 都道府県別人口

都道府県別の人口を比較すると、住民基本台帳、推計人口とも大都市を擁する都道府県では過小になっており、人口の少ない県では過大となっている。推計人口の都道府県間転出入者の推計に住民基本台帳の移動報告が用いられていることから、両統計とも国勢調査との乖離傾向は同様となっている。



差率による比較

住民基本台帳人口と推計人口の国勢調査結果からの乖離度を比較するため、都道府県、男女、年齢階級（5歳階級）別に国勢調査結果（時点調整済みの基準人口）との差の絶対値を国勢調査結果で除した差率によって比較した（下式）。全体としての比較は差率の単純平均を使用した。

なお、推計人口の推計誤差による乖離は年毎に均一に増加するものと考え、比較に用いた推計人口が国勢調査から4年目に推計された数値であることから平均的な乖離はその1/2として計算した。

$$(\text{差率}) = \frac{x_{ijk} - y_{ijk}}{y_{ijk}}$$

$$(\text{平均差率}) = \frac{\sum_{ijk} \frac{x_{ijk} - y_{ijk}}{y_{ijk}}}{47 \times 2 \times 13}$$

x_{ijk} : i 都道府県、 j 男女、 k 年齢階級（5歳階級）別住民基本台帳人口
又は推計人口

y_{ijk} : i 都道府県、 j 男女、 k 年齢階級（5歳階級）別国勢調査基準人口

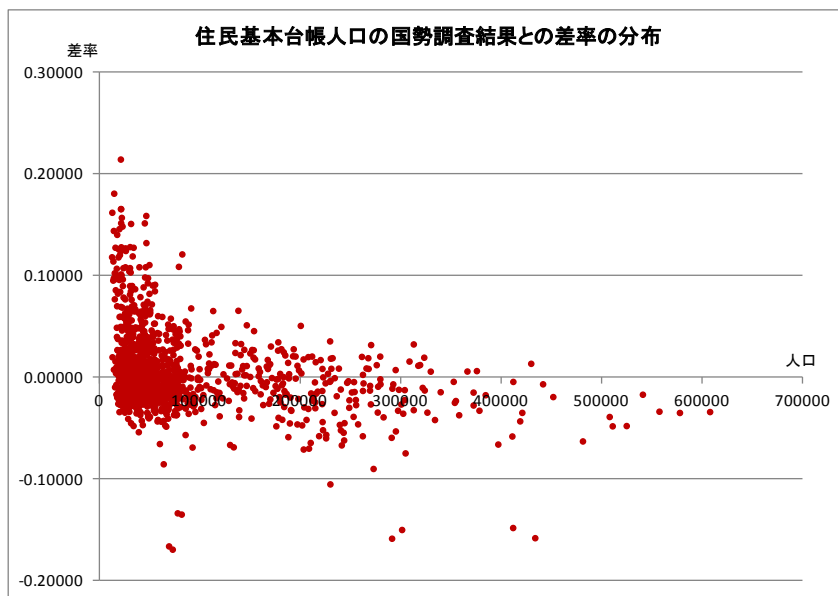
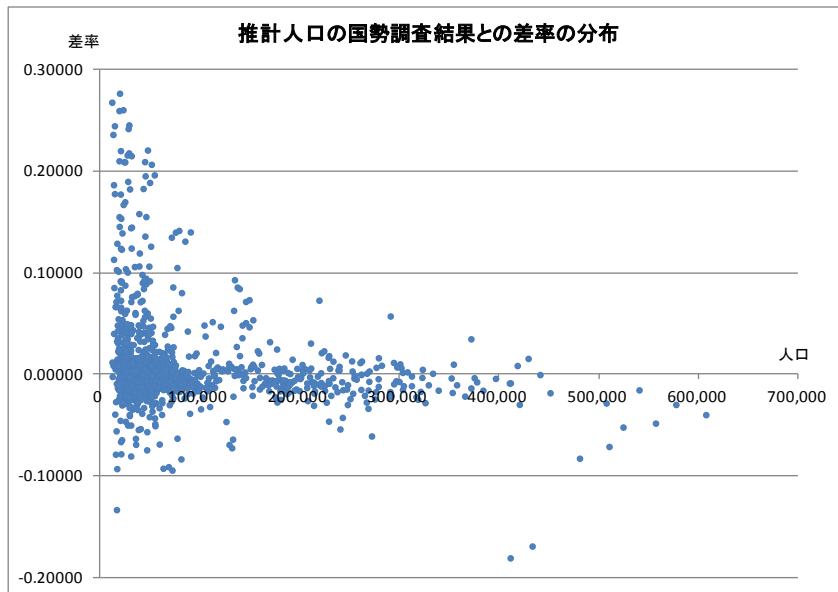
全体の平均差率は、住民基本台帳人口が0.01503、推計人口が0.01242となっている。年齢階級別には20代以下の年代では推計人口の平均差率が高く、それ以外の年代ではほとんど住民基本台帳人口が高くなっている。

国勢調査結果との平均差率(絶対値)

		推計人口	住民台帳
総数		0.01242	0.01503
男	総数	0.01407	0.01593
	15～19歳	0.02617	0.02101
	20～24	0.06665	0.01647
	25～29	0.01790	0.01193
	30～34	0.00763	0.01917
	35～39	0.00688	0.00434
	40～44	0.00675	0.00627
	45～49	0.00442	0.00967
	50～54	0.00460	0.00970
	55～59	0.00428	0.02406
	60～64	0.00539	0.03705
	65～69	0.00463	0.00441
	70～74	0.00931	0.01463
75～79	0.01834	0.02845	
女	総数	0.01077	0.01412
	15～19歳	0.02308	0.01702
	20～24	0.05748	0.01588
	25～29	0.01442	0.01176
	30～34	0.00715	0.01934
	35～39	0.00625	0.00365
	40～44	0.00469	0.00566
	45～49	0.00453	0.00946
	50～54	0.00369	0.00926
	55～59	0.00330	0.02533
	60～64	0.00393	0.03378
	65～69	0.00327	0.00855
	70～74	0.00292	0.00854
75～79	0.00532	0.01532	

横軸に国勢調査の基準人口、縦軸に差率をとって都道府県、男女、年齢階級（5歳階級）別の分布をみると、住民基本台帳人口、推計人口とも同じような分布形となっているが、人口規模が10～30万程度の区分では人口推計の差率が小さくなっているのに対し、住民基本台帳人口の差率はバラツキが大きくなっている。

また、人口規模の小さい区分で推計人口の差率に大きなものがあるのは地方の若年層での差が大きくなっていることを反映していると思われる。



5.5 考察

住民基本台帳人口と推計人口の国勢調査結果からの乖離の傾向は類似しており、モバイル空間統計のベンチマークとしてはどちらの統計も不適切ではないと考えられるが、平均差率や差率の分布を勘案すると推計人口を使用することがより望ましいと考えられる。

6 おわりに

この共同研究の今年度の目的は、これまでに得られない新たな統計情報である「モバイル空間統計」を活用した新たな統計の創出可能性について、ニーズを抽出して作成方法を検討するとともに、その精度を検証することであった。

ニーズヒアリングの結果としては、各方面から極めて高い期待が示されるとともに、各種の活用方法について多くの示唆を得ることができた。しかし、本年度の研究においては、具体的な形態や提供方法については十分な検討を行うことが出来なかったため、今後は今年度のニーズヒアリングの結果を生かし、より具体的な活用方策を検討することが必要である。

また、モバイル空間統計の精度については、統計理論に基づく定量的な検証方法は確立されていないため、平成22年国勢調査メッシュ統計結果との比較において偏差率という評価指標を設定して評価を行った。その結果として、一般的に利用されると考えられる3次メッシュ、4次メッシュではある程度以上に人口が集中している地域では高い信頼性を持つことが示された。また、都心部において示される高い偏差率のようなモバイル空間統計と国勢調査の作成方法の違いに基づくモバイル空間統計の特徴も明らかにされた。

上述のように、今回の共同研究を通じてモバイル空間統計の無限の可能性の片鱗を垣間見ることができた。一方、「モバイル空間統計」は多くの統計と同様、実際に利用されて初めてその存在に価値が生まれるものである。そしてこの新しい統計が利用されるためには、よりその精度自体を高めていくことは無論のこと、精度を含めた統計が保有する特徴を明らかにすることも必要不可欠である。統計が持つ特徴を分かりやすく提示する一つの方法は比較である。今後、精度を高める研究を進めるとともにモバイル空間統計が持つ特徴を別の統計と比較検証してその特徴を明らかにしていく努力が必要である。

さらに、利用をしてもらうためには、モバイル空間統計が社会に認知されることも必要である。しかしながら、ビッグデータ時代の到来とともに、情報の保護とデータ利用の両立にますます高い関心が寄せられており、モバイル空間統計がプライバシーの侵害に直結すると誤解されないようにこの統計を社会に浸透させていく十分な工夫が必要である。

その工夫として、例えば、既に公表され広く市民権を得ている「国勢調査メッシュ統計」を補完する統計としてモバイル空間統計に基づく「国勢調査“延長”メッシュ統計」を作成・提供していくことなどが考えられる。また、提供の形態としては、ニーズヒアリングでの意見として挙げられたように、一定の地域と一定の時間区間を統計利用者が指定することで、その地域全体における時間区間の人口の推移を時系列でオンデマンド集計して提供することも検討すべきであろう。ここで、提供の初期段階にあたっては、市町村間の移動レベルまでマクロ化したもののみを提供するなど、社会受容性や情報保護との両立を図る配慮のあり方も検討が必要になると考えられる。

今回の共同研究の先にある、究極の目的は「モバイル空間統計」を世に送り出し、十分に利用してもらうことで、より良い社会を実現することにある。この究極の目的を目指し、更なる研究の進展が望まれる。