

# データ融合による日別・旅行目的別都市間旅客流動量推計

近年活用事例が増えつつある集計携帯電話基地局情報等のビッグデータは、都市・地域間の旅客流動を年間を通じて容易に把握できる新たなデータソースとして期待されている。しかし、こうしたデータだけでは、アンケート調査のように旅行者の質的情報 (e.g. 旅行目的等) を把握することは難しい。本研究では、相互補完関係にあるこれら2つのデータ (ビッグデータとアンケートデータ) を融合して、多時点に渡る日別・旅行目的別の都市間旅客流動量を推計する手法を提案した。モバイル空間統計と全国幹線旅客純流動調査に対して提案手法を適用し、推計結果の解釈、別の統計調査から得られた流動量との比較、モデルの残差分析等を行ったところ、次のような主要な知見が得られた：(1) 年間での旅行目的構成比は、別統計調査の結果に概ね合致する、(2) 日単位での旅行目的別流動量に関しては、幹線調査のサンプル数が少ない地方発の都市間流動において不自然な推計結果が生じやすい。

キーワード | 都市間交通調査, 携帯電話基地局情報データ, データ融合, 旅行目的推定, 季節変動

**鈴木 新**  
SUZUKI, Arata

東京工業大学環境・社会理工学院修士課程

**山口裕通**  
YAMAGUCHI, Hiromichi

博士(工学) 金沢大学理工研究域助教

**福田大輔**  
FUKUDA, Daisuke

博士(工学) 東京工業大学環境・社会理工学院准教授

## 1—はじめに

我が国では、新幹線・空港等の都市間旅客交通網の整備が進められ、都市間移動の速達性や利便性の向上に大きく貢献してきた。その結果、2018年時点において、推計人口が40万人以上の47都市のうち中心部から1時間圏内に新幹線駅が存在しない都市は、札幌市・松山市・大分市・長崎市・高松市の5つのみとなっている。人口規模の大きい都市間については、高速・大容量の旅客交通サービスが既に整備されている状況であり、今後は、規模の小さな地方都市をそのネットワークに組み込むような都市間交通網の整備が検討されるようになるであろう。そのような場合には、都市間交通サービスの需要予測や事業評価において、より詳細かつ高精度の分析が求められるようになると考えられる。

しかし、従来の都市間旅客交通の分析においては、都市間旅行行動の時間的 (季節的) な変化に関する詳細な検討・分析は十分になされていない。その一因は、基礎となる実態調査である「全国幹線旅客純流動調査」(以降、幹線調査) が5年に一度しか実施されず、かつ、秋季2日分 (平日と休日) の調査しかなされていないことにあると考えられる。そのため、調査を実施していない季節に見られる都市間旅行行動 (e.g. 年末年始の帰省行動, ゴールデンウィーク (GW) の行楽行動等) は、需要分析において適切に考慮されていない。この結果、従来の需要分析では、例

えば「お盆の時期に多くの地方への帰省行動が殺到する」といったような、特定時期に偏在している都市間旅行パターンを過小評価している可能性が考えられる。幹線調査が実施されている秋季の平日・休日に卓越すると考えられるビジネス目的旅行や大都市近郊への観光旅行が過大に推計され、逆に、特定時期に偏在的に生じる地方部への帰省等の私用旅行が過小推計されると、需要予測や事業評価において規模の小さい地方都市間を接続する都市間交通サービスが過小評価されてしまう可能性がある。今後人口減少が進展すると、財源制約から維持すべき都市間交通サービスを絞り込む必要に迫られると考えられるが、特定の調査期間に偏ったデータのみでは適切な意思決定を支援できない可能性が危惧される。

こうした課題を解決するためにまず考えられるのが、季節変動等を考慮可能な時間解像度の高いビッグデータの活用である。近年は携帯電話基地局情報等を用いることで、24時間365日での流動把握も可能になりつつある。その代表例であるNTT Docomoのモバイル空間統計<sup>1)</sup>は、携帯電話基地局情報を集計して作成されており、最小1時間単位で当該空間 (地域メッシュ等) に滞在する人数 (滞在人口) の推計値を提供している。さらにモバイル空間統計では、滞在者の携帯電話の登録地、性、年齢階層等のセグメント毎の滞在人口の推計値も提供されており、滞在地と登録地の情報から都市間旅客流動量の情報を得ることもできる。

しかし、こうしたビッグデータには、交通の分析において重要な質的情報である「旅行目的」は含まれていない。都市間交通需要の将来予測を行うにあたり、サービス水準の変化に対する感度は旅行目的毎に大きく異なることが通常である<sup>2)</sup>。つまり、時間解像度の高いビッグデータを単純に活用するだけでは、上述のような詳細な都市間旅行行動の分析・需要予測に対して必ずしも十分ではないと考えられる。

本研究では、幹線調査とモバイル空間統計データを融合させるアプローチについて検討する。幹線調査が（調査時点が限定的でありながらも）旅行目的別の都市間流動を直接的に捉えているのに対し、モバイル空間統計は（旅行目的等の質的情報は分からないながらも）任意の日における居住地別・旅行先別滞在人口を把握できる。すなわち、両者は相互補完的な関係にあり、適切なデータ融合を行うことで、旅行目的等の質的情報を加味した上で都市間旅客流動量の季節変動特性を把握することが可能になると期待される。そこで本研究では、幹線調査とモバイル空間統計データを融合して、日別・旅行目的別の都市間旅客流動量を推計する手法を提案し、特に以下の2点について検討を行うことを目的とする。

- (1) 幹線調査とモバイル空間統計の融合により得られる日別・旅行目的別旅客流動量推計値には、どのような特徴があるのか？
- (2) どのような状況で、推計の妥当性が低下するのか？

これらの問いに答えるべく、本研究では、915日分のモバイル空間統計データと2010年幹線調査データの融合により日別・旅行目的別旅客流動量を推計し、提案手法より得られる推計結果の基本特性を明らかにする。

## 2——データの概要

### 2.1 モバイル空間統計

NTT Docomoによって作成されているモバイル空間統計<sup>1)</sup>は、同社の携帯電話基地局に集められた携帯電話情報を集計・データ化し、地域の滞在総人口を長期的・広域的に推計しているパッシブ型の位置情報ビッグデータである（表—1）。本研究で用いるデータの仕様は以下のとおりである。

■表—1 モバイル空間統計のデータフォーマット

| 日付        | 時刻   | 滞在地 | 居住地 | 滞在人数 |
|-----------|------|-----|-----|------|
| ..年..月..日 | AM4時 | 北海道 | 青森  | ～人   |
| ..年..月..日 | AM4時 | 北海道 | 宮城  | ～人   |
| ..年..月..日 | AM4時 | 北海道 | 岐阜  | ～人   |
| ..年..月..日 | AM4時 | 北海道 | 長崎  | ～人   |
| ..年..月..日 | AM4時 | 北海道 | 沖縄  | ～人   |

- ・後述の幹線調査のゾーニング条件に合わせて、分析空間単位を都道府県単位とする。
- ・対象とする期間は、手元に用意可能な2014年3月1日～2016年8月31日の915日間とする。
- ・居住都道府県以外へ旅行する「宿泊旅客」のみを対象とする。逆に言えば、同一都道府県内だけで完結する旅行や、日帰りで居住地に帰ってくる旅行は分析対象としない。これは、モバイル空間統計の「滞在人口統計」の利用を前提とした場合には、日帰り旅行の判定が難しいという理由による。本研究では、モバイル空間統計において旅客の移動が最も少ないと考えられている時間帯である午前4時台の滞在人口データの情報を、宿泊旅客数と見立てることとする。
- ・後述の幹線調査と同様、首都圏（埼玉県、千葉県、東京都、神奈川県）、中京圏（愛知県、岐阜県、三重県）、近畿圏（大阪府、京都府、兵庫県、奈良県）の各大都市圏内で完結する都市内流動については分析対象としない。

### 2.2 全国幹線旅客純流動調査

本研究では、2010年の平日・休日各1日を対象に実施された第5回全国幹線旅客純流動調査<sup>3)</sup>の都道府県間流動データを使用する（表—2）。旅行目的として、(1) 業務目的、(2) 観光目的、(3) 私用目的（帰省等）、の3つを対象とする。分析単位は都道府県（先述の三大都市圏を除く）とする。また、都道府県を越えて宿泊する旅客のみを対象とし、日帰り旅行者は分析対象外とする。なお、モバイル空間統計の「滞在地」に対応する情報として、幹線調査の「旅行先」の情報を用いる。すなわち本研究では、出発地から到着地へのトリップ情報ではなく、旅行者の居住地とその滞在地（旅行先）の間の旅行情報を分析する。

ここで表—2の幹線調査集計フォーマットに記載の拡大人数について説明する。幹線調査では、アンケートに回答した旅行者の人数をそのまま記載するのではなく、回答サンプルと同一の行動をした人数も考慮するために、各種交通機関の1日利用者数と調査結果を用いてそのサンプルの拡大係数を算出している。このサンプル数に拡大係数を乗じた人数を、幹線調査では流動の推計値としており、「拡大人数」と称することとする。

■表—2 全国幹線旅客純流動調査の集計フォーマット

| 滞在地 | 居住地 | 旅行目的 | 拡大人数 |
|-----|-----|------|------|
| 北海道 | 青森  | 私用   | ～人   |
| 北海道 | 青森  | 観光   | ～人   |
| 北海道 | 青森  | 業務   | ～人   |
| 北海道 | 宮城  | 業務   | ～人   |
| 北海道 | 宮城  | 観光   | ～人   |

## 3.1 データ融合に関する既往研究

都市間交通のような長距離の移動は、個人単位では低頻度でしか行われぬ。また、旅行頻度そのものの個人差も非常に大きく、その全容を正確に把握することが困難であることは、これまでも指摘されてきた<sup>4)</sup>。一方、携帯電話位置情報をはじめとするパッシブ型の位置情報ビッグデータは、大量のサンプル（携帯電話ユーザー）の位置情報を広範囲かつ高頻度で取得したデータである。そのため、長距離流動に関しても、その実態を俯瞰的かつ時間的に詳細に把握できる可能性が高いと考えられている。例えば室井ら<sup>5)</sup>は、パッシブ型の位置情報ビッグデータのみを用いて、人為的な回答に依存しない都市間旅客交通の新たな統計データの構築方法について検討を行っている。

しかし、携帯電話位置情報データには、旅行先・旅行目的等の質的情報が含まれておらず、この情報を補完する方法に関する研究が進展している。例えば、都市内の日常的移動（通勤・通学等）を対象とする場合、滞在自宅・勤務地等の同じ場所に繰り返し訪れるという人の行動の周期性を活用し、トリップ・ツアーの抽出とその目的を推定するアルゴリズムが検討されている<sup>6)~9)</sup>。Alexander et al.<sup>8)</sup>では、観測頻度・時間情報からユーザーの自宅と勤務地を推定するアルゴリズムを用い、携帯電話のCall Detail Record (CDR) データから通勤、帰宅、その他という目的区分でトリップを抽出した上で、その発生時間分布・距離分布・空間分布等を従来型のアンケート調査と比較し、概ね旅行目的別の流動を良好に推定できることを確認している。また、Ge and Fukuda<sup>9)</sup>では、大規模交通調査と携帯電話位置情報を融合して、都市内の交通流動表の時点更新を行う方法論を開発している。これらはいずれも都市内交通を対象としている。

一方、都市間交通のような長距離の移動を対象とする場合、業務・観光・私用といった旅行目的別の移動は、個人毎に見た場合、それぞれが低頻度で行われるため、観測頻度や時間情報のみから旅行目的を判定することは一般に困難となる。これに対しJanzen et al.<sup>10)</sup>では、非集計の位置情報データから得ることができるツアーの距離・期間・目的地・季節といった情報と、従来型のアンケート調査の情報を併用して旅行目的を付与する方法を提案している。具体的には、ランダムフォレスト法によって従来型のアンケート調査を学習させ、旅行目的を推定する方法をCDRsに適用し、旅行目的情報を付与した結果が集計的な目的構成比に合致することを確認している。しかしこの方法では、個人情報保護の観点から非常に入手が困難である非集計情報が必須となる上、個々のツアーに対してその情報

の導出と旅行目的の判定を行う必要があり、計算負荷が大きくなるという欠点がある。個人情報保護のための秘匿処理が実施済みであり、入手も比較的容易なモバイル空間統計のような集計加工済の携帯基地局情報データの活用を念頭に置いた場合には、Janzen et al.<sup>10)</sup>の方法の適用は難しい。

本研究で提案するデータ融合方法は、入手が比較的容易な携帯電話位置情報による集計データと、幹線調査を相互補完するものであり、これにより、日別・旅行目的別の都市間旅客流動量を容易に推計することが可能になると期待される。

## 3.2 本研究の基本的考え方

モバイル空間統計より、長期間に渡る居住地別・旅行先別滞在人口を得ることができるが、旅行目的等の質的情報は含まれていない。一方、幹線調査からは、秋の平日・休日に限定されるものの旅行目的別の都市間旅客流動量を得ることができる。そこで、ある特定の日におけるモバイル空間統計の居住地別・旅行先別滞在人口に、幹線調査から得られる旅行目的別トリップ構成比を適切に組み合わせることで、旅行目的情報を付与した滞在人口データを構築する。

本研究では、旅行目的別トリップ構成比として、旅行目的別宿泊先選択比率 ( $R$ ) を採用する (図—1)。これは旅行目的と居住地都道府県の組み合わせ毎に、宿泊する都道府県の比率を算出したものである。この  $R$  を用いてデータ融合を行い、後に別の統計調査との比較を行うことによって、本研究のアプローチの妥当性を検討する。

## 3.3 データ融合モデル

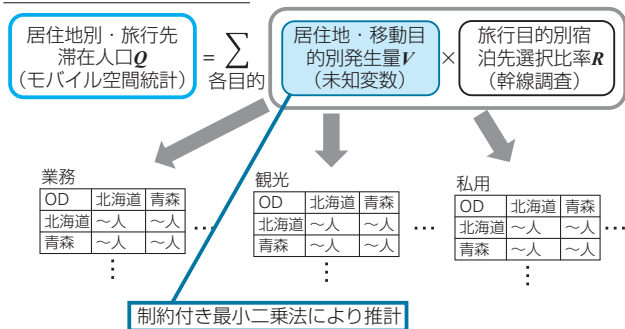
幹線調査データを用いて、以下の式 (1) に基づき旅行目的別宿泊先選択比率 ( $R$ ) を算出する。

$$R_{h,s,p} = \frac{J_{h,s,p}}{\sum_{S \in S} J_{h,s,p}} \quad (1)$$

ここで、 $h$ : 居住地都道府県、 $s$ : 宿泊先都道府県、 $S$ : 全ての宿泊先都道府県集合、 $p$ : 旅行目的（業務、観光、私用）、 $J$ : 幹線調査拡大人数である。なお  $R$  の値は、幹線調査より平日・休日別に求められることから、モバイル空間統計より得られるデータの観測日が平日か休日かによって、適用する  $R$  の値を使い分けることとする。

本研究では、「モバイル空間統計より得られる居住地別の旅行先滞在人数 ( $Q$ ) が、旅行目的別宿泊先選択比率 ( $R$ ) と居住地別・旅行目的別発生量 ( $V$ ) を掛け合わせたものを、全旅行目的に対して総和をとることによって与えられる」と考える (図—1)。このとき、旅行目的  $p$  の都道府県間流動量は、行列  $\{R_{h,s,p} \times V_{h,p,d}\}$  によって推計することが

○年×月×日の地域間流動量



■図—1 目的別都市間流動推計のイメージ

できる(図—1グレー枠内部)。このとき、データ融合の統計モデルは、式(2)のように定式化される。

$$Q_{h,s,d,t=4} = \sum_{p \in P} (R_{h,s,p} V_{h,p,d}) + \varepsilon_{h,s,d} \quad (2)$$

ここで、 $Q$ :旅行先滞在人数(モバイル空間統計)、 $R$ :旅行目的別宿泊先選択比率(幹線調査)、 $V$ :居住地別・旅行目的別発生量(未知変数)、 $d$ :日付ラベル(モバイル空間統計)、 $t$ :観測時刻(モバイル空間統計、本研究では午前4時に固定)、 $\varepsilon$ :モデルの予測値と実績値との乖離を表す誤差項である。

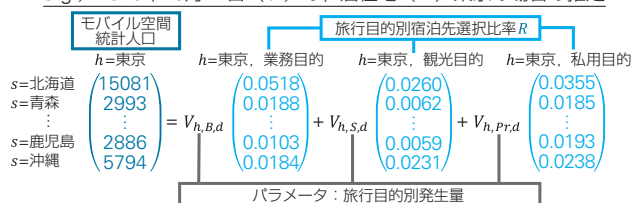
このモデル<sup>注1)</sup>の未知変数は居住地別・旅行目的別発生量( $V$ )であり、旅行先滞在人数( $Q$ )を従属変数、旅行目的別宿泊先選択比率( $R$ )を独立変数と見立てた定数項無しの線形重回帰モデルを非負制約付き( $V \geq 0$ )最小二乗法を用いて推計することで求めることができる。推計の具体的なイメージを図—2に示す。すなわち本モデルでは、日別・旅行目的別の都市間流動表を直接求めるのではなく、流動表を推計する際に必要となる居住地別・旅行目的別発生量( $V$ )をまず推計した上で、日別・旅行目的別の都市間流動量を推計する構造となっている。その際「幹線調査データから得られた旅行目的別、平日・休日別の宿泊先選択比率( $R$ )が時間的に(全季節で)一定である」という仮定を置くことにより、携帯電話基地局情報から得られる居住地—旅行先分布表を旅行目的別に分解している。この仮定からの乖離が小さい限り、本研究の方法を用いることで各日(季節別)における旅行目的別の都市間流動表を精度よく推計することができると考えられる。

#### 4——データ融合結果とその考察

##### 4.1 目的別の生成旅行量比較

推計した旅行目的別の宿泊先選択比率( $R$ )を用いて、居住地別・旅行目的別発生量( $V$ )を日別に推計した。この推計は、43,005回(915日×47居住地都道府県)実施しており、このうち34,964回の推計において非負制約に抵触し、

e.g.) 2015年10月21日( $d$ )の、居住地( $h$ )東京の場合の推定



「日数×居住地都道府県数」の数だけパラメータ推定を行う。

■図—2 パラメータ推定のイメージ

目的別の旅行者数の推計値が一部でゼロになる結果となっている。

それら全ての推計結果を集計して得られる日本全体の日別・旅行目的別生成交通量の推計値の推移を図—3に示す。グラフは縦軸が一日毎の目的別生成交通量を、横軸が日付を表している。GWやお盆のような特別な長期休暇期間中では、観光目的や私用目的の流動量が多いのに対して業務目的の流動量は少なくなっており、概ね予想された結果が得られている。

また、日別の旅客流動の変動を詳しく見るため、例として2015年5月一ヶ月間の旅客流動量の推計結果を図—4に示す(図—3のグレーハッチ部に相当)。観光目的や私用目的の旅客流動量は休日に増加し、逆に業務目的の旅客数は減少している。一方、平日には業務目的の旅客流動量が増加して観光目的や私用目的の旅客数が減少する結果となっている。

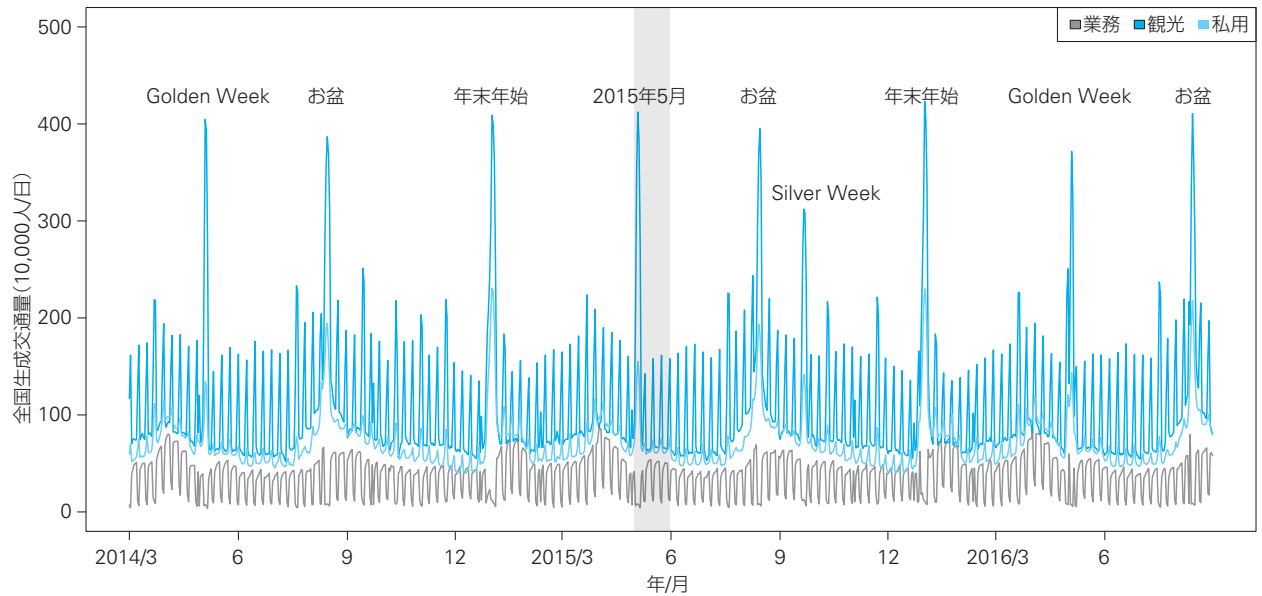
なお、重回帰モデルの決定係数は全期間を通じて0.52~0.94の範囲に収まっている(図—5)が、特にお盆等の長期休暇期間中に低い値を取る傾向があることが確認された。本モデルでは旅行目的別宿泊先選択比率が分析期間中を通じて一定という仮定を置いている。この値は幹線調査の調査日である秋の特定平休日のデータのみの情報から算出されており、特にお盆等長期休暇の期間中において、この仮定からの乖離が生じている可能性が高いことが示唆される。

##### 4.2 目的別の都道府県間旅行量比較

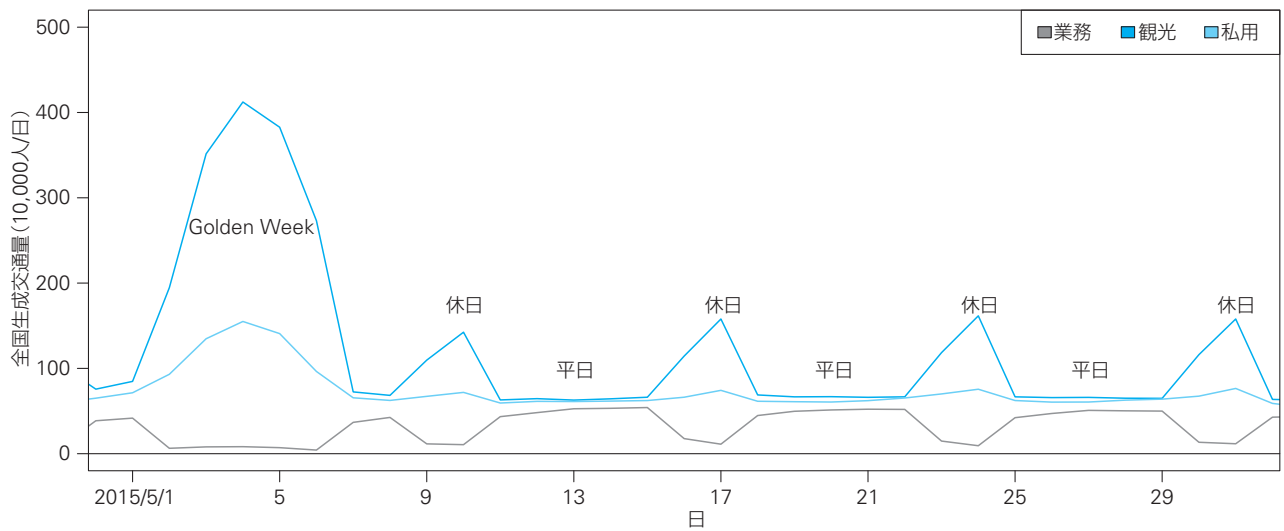
得られた居住地別・旅行目的別発生量( $V$ )を用いて宿泊旅客の日別・旅行目的別地域間流動量を推計する。実際には915日分の47×47の要素数を持つ旅客流動表が推計されており、分析結果は膨大になる。紙面の都合上、本研究では、居住地(出発地)が東京都もしくは青森県の場合を選定し、さらに、日付についても2つの異なる特徴的な日(正月:2015年1月1日(祝日)、秋の平日:2015年10月21日(水)[2015年幹線調査の平日調査実施日])の結果のみを例示する。

###### 4.2.1 居住地が東京都の場合

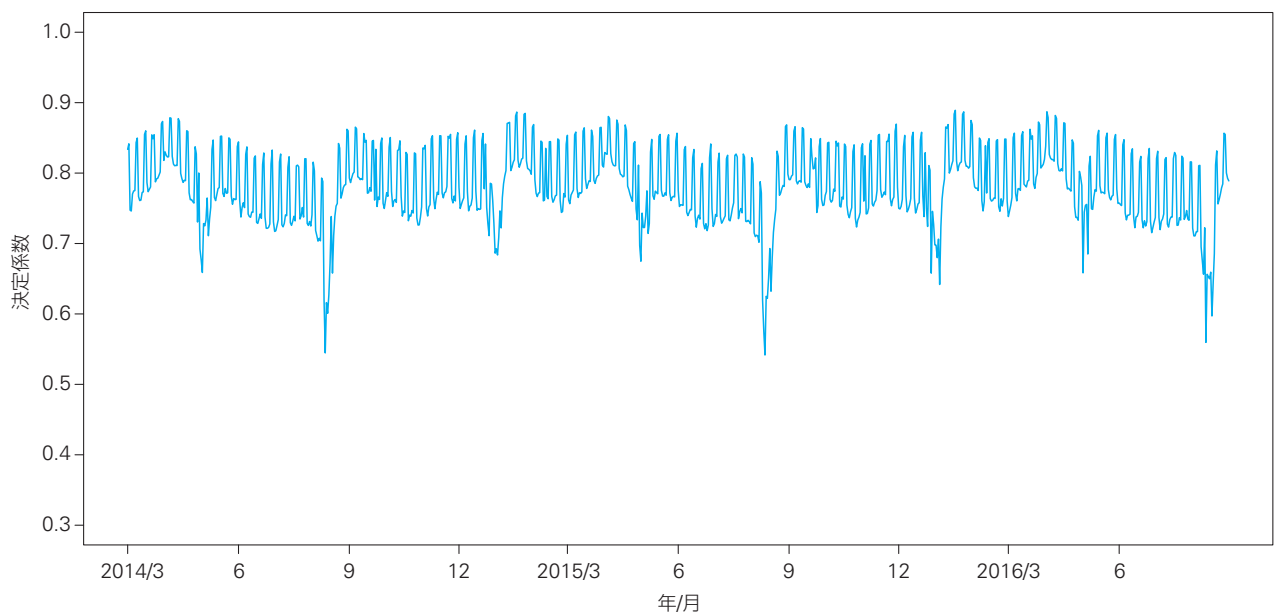
本節では、東京都居住者の各県への旅行目的別の旅行



■図-3 日別・旅行目的別生成交通量の推移 (2014年3月1日～2016年8月31日)



■図-4 日別・旅行目的別生成交通量の推移 (図-3中の2015年5月 [グレーハッチ部] を拡大表示)



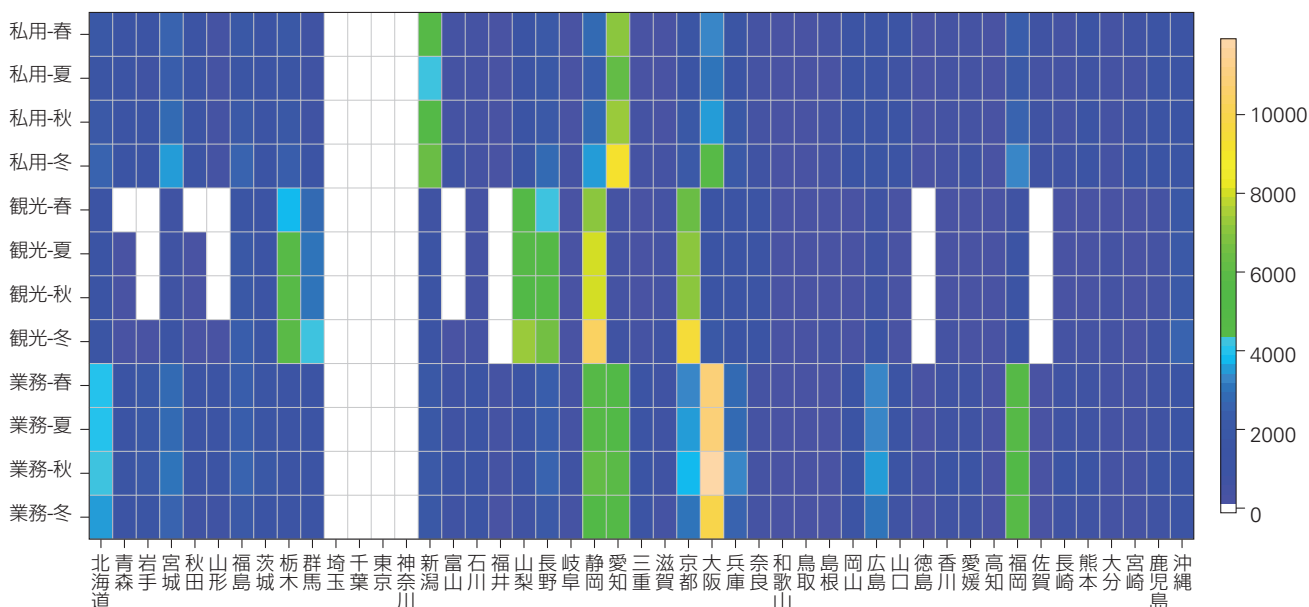
■図-5 モデルの決定係数の推移 (2014年3月1日～2016年8月31日)

数をいくつかの時点で比較していく。まず、図—6は、4つの季節における平日・水曜日（2015年2月25日：冬，5月27日：春，7月15日：夏，10月21日：秋）について、旅行目的別旅行者数の推計結果を示したものである。この図から、観光旅行は京都府や静岡県・山梨県等を旅行先とするものが多く、一方、業務目的では特に大阪府への流動量が多い等といった傾向を読み取ることができる。次に、季節の違いを見ると、概ね同じような傾向にあるものの、一方で冬（2月25日）に私用目的が多い等といった季節毎に固有の特徴が推計されていることも確認できる。

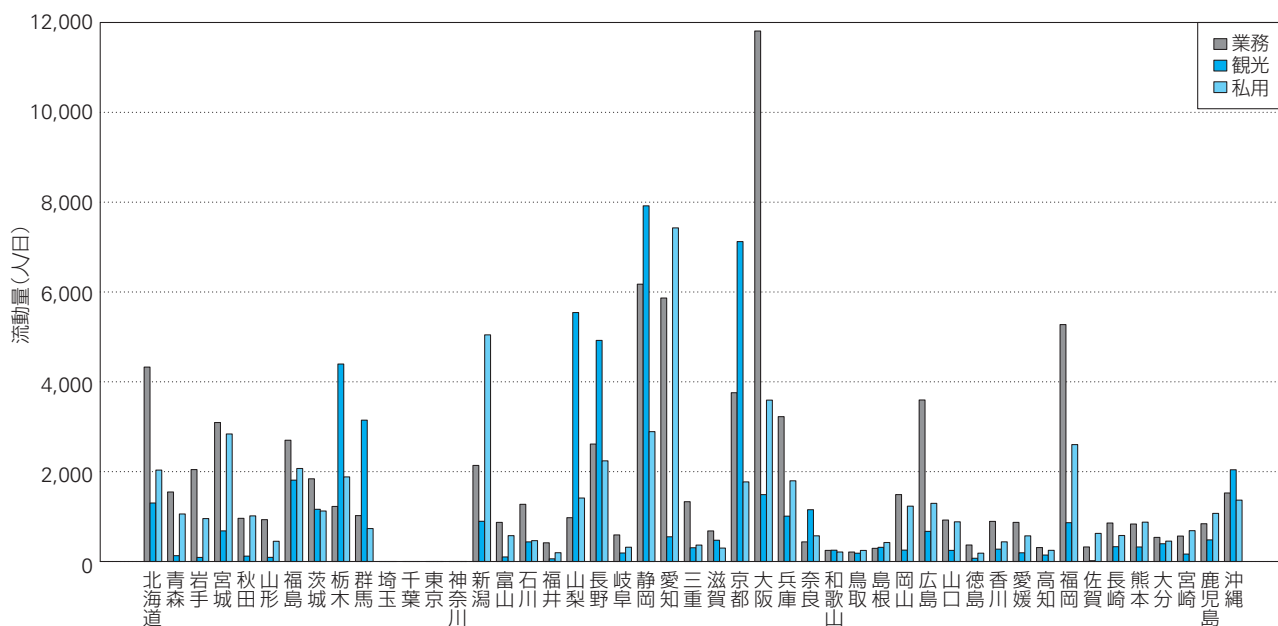
この推計結果を、さらに細かく見ていこう。図—7は、秋（10月21日）の東京都居住者による各都道府県への目的別旅行者数の推計結果を棒グラフで示したものである。この図から大阪府・愛知県・北海道・宮城県等の大都市がある

都道府県への業務目的流動量が特に多いことが分かる。このことは、平日は支社・本社機能等を多く有する地方の主要都市への業務旅行が大半を占めることを反映していると考えられる。他方、群馬県や栃木県等東京都近郊の県や、静岡県、京都府等有名観光地がある都道府県に対しては、平日にもかかわらず観光目的の旅客流動量が多いことも分かる。

次に、図—8は、幹線調査の実施された季節とは異なる、春の通常の平日（5月27日）の東京都発の旅行者数の推計値を示している。この日の推計結果は、図—7で示した秋期平日（10月21日）にかなり近い値となっていることが確認される。これは、旅行目的別の発生量（V）の推計結果が春季の平日と秋季の平日との間でかなり類似していることを示唆している。但し、本研究のモデルでは、目的別の宿



■図—6 旅行目的別旅客流動量の推計結果（居住地：東京都，4季節の水曜日，単位：人/日）



■図—7 旅行目的別旅客流動量の推計結果（居住地：東京都，2015年10月21日水曜日）

泊先割合(R)が年間通じて全ての平日で同一と仮定しており、これが推計結果が類似した要因である可能性も否定できないことには留意する必要がある。

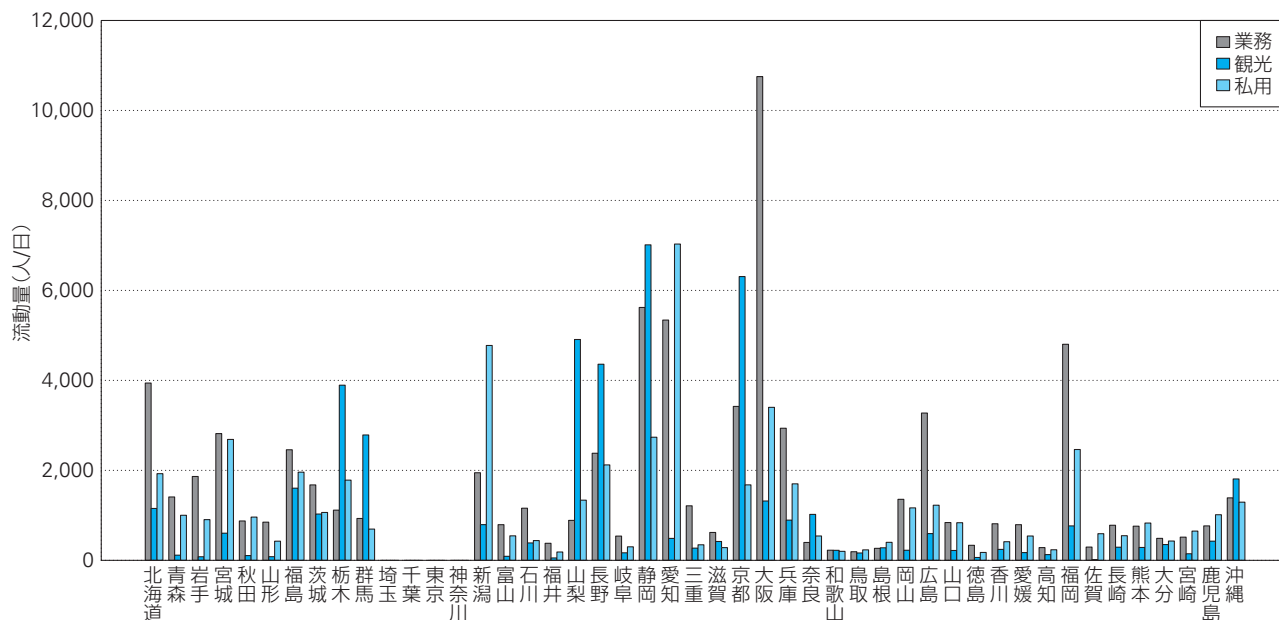
次に、正月やお盆のような特殊な休日の流動量推計結果について考察する。図一9は2015年1月1日の東京都居住者の都道府県別の旅行数推計結果を示している。業務目的の流動量はほぼゼロという結果となった。また、春や秋の推計結果(図一7, 8)では少なかった私用目的の流動量が多いことが分かる。旅行先について細かく見ると、茨城県や山梨県等の東京都近郊の県に加えて、大阪府や福岡県等といった大都市への流動量がかなり大きいことも分かる。これは、人口の多い地方大都市から東京に上京し居住している人々が年末年始で実家等に帰省しているような流動を表していると推察される。他方、静岡県や京都府等代

表的な観光地を有する都道府県への観光目的流動量も多くなっているが、これは年末年始の連休中に帰省等を行わず、観光旅行に行く人が一定数存在することを表している。2015年8月15日のお盆の流動量推計結果(図一10)も同様に解釈可能である。

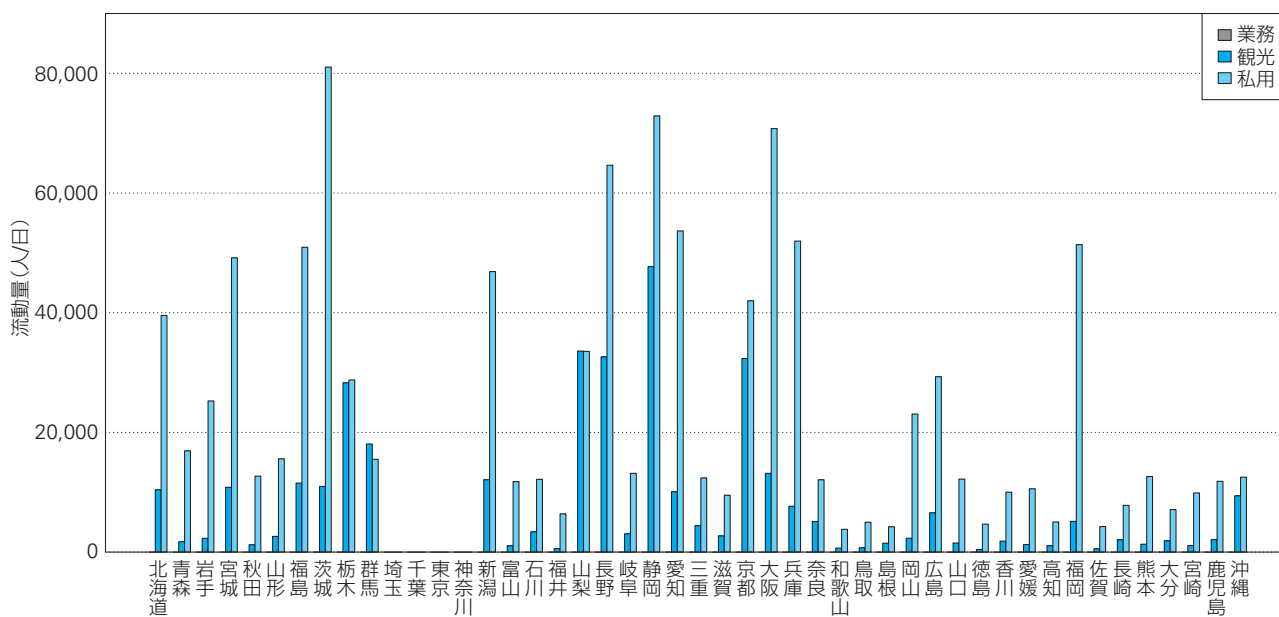
以上のように、東京都居住者の旅行目的別流動量の推計結果については、真値が存在しないため推計結果の妥当性を定量的に検証することはできないものの、当初予想された都市間旅行パターンと大きな乖離がない推計結果を得ることができたと考えている。

#### 4.2.2 居住地が青森県の場合

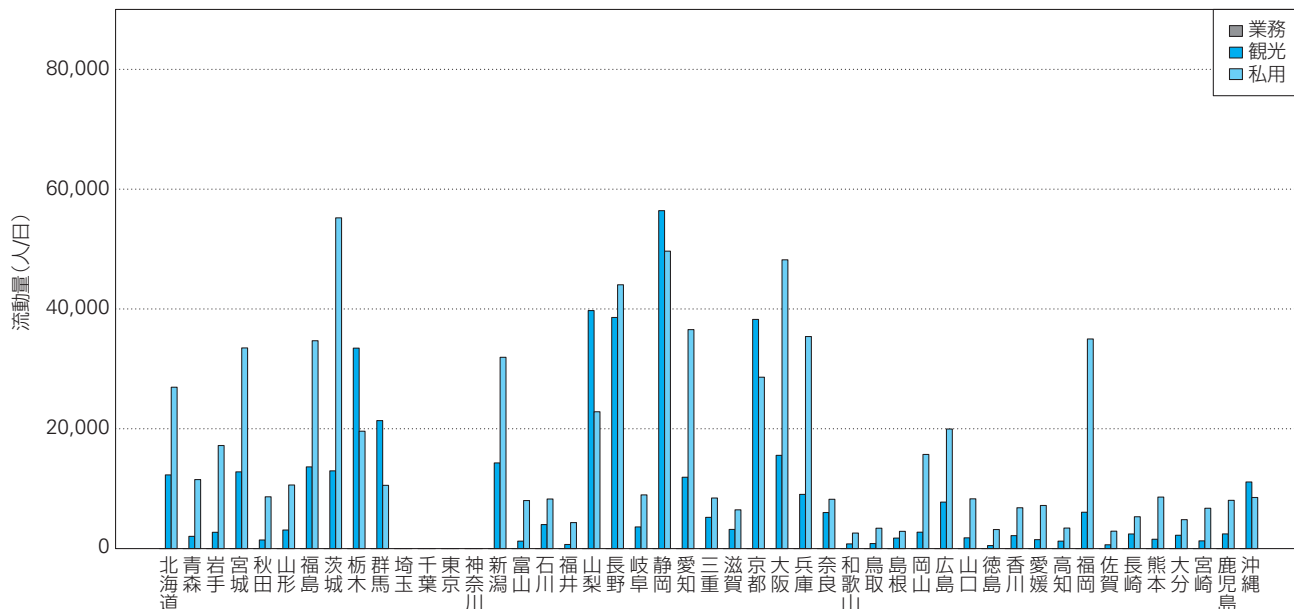
次に、地方部の推計結果の代表例として、青森県居住者の流動量について考察する。図一11に、4つの季節における平日・水曜日(図一6の分析と同じ日付)の流動推計



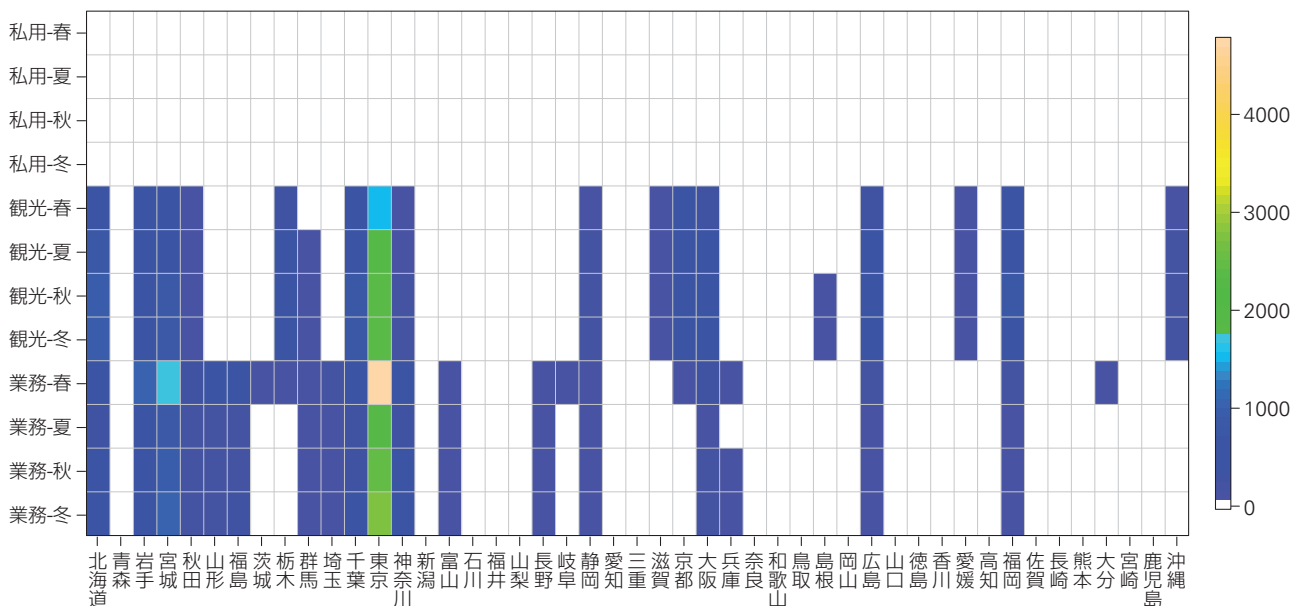
■図一8 旅行目的別旅客流動量の推計結果(居住地:東京都, 2015年5月27日水曜日)



■図一9 旅行目的別旅客流動量の推計結果(居住地:東京都, 2015年1月1日(元日))



■図一10 旅行目的別旅客流動量の推計結果 (居住地:東京都, 2015年8月15日 (お盆))



■図一11 旅行目的別旅客流動量の推計結果 (居住地:青森県, 4季節の水曜日, 単位:人/日)

結果を示す。一部の季節・旅行目的のもとでの東京都や宮城県等への流動を除き、多くの季節と旅行先の組み合わせに対し流動量の推計値は非常に小さい値もしくはゼロとなっている。このように、そもそも旅客需要が少ない地方部の居住地—地方部の旅行先ペアでは、ある特定の日に於いて目的別流動量がゼロになるという結果も十分起こり得ると考えられる。実際、幹線調査においても、都道府県間流動表の2,132個の起終点ペアのうち、例えば休日・観光目的243個のペア、平日・私用目的591個のペアにおいて旅客流動量はゼロとなっている。但し幹線調査においては、「ある目的の総発生流動量がゼロ」という状況は含まれておらず、図一11に示したような全ての季節で私用目的の総発生流動量がゼロとなった推計結果は、やはり不自然と言えるだろう。

このような結果が生じた理由として、幹線調査のサンプ

ル数の問題があると考えられる。幹線調査のサンプル抽出率は、全国を通して大きな差はなく、旅行数自体が多い東京都と比較すると、特に地方部においては、同じ抽出率では人口の分だけサンプル数が少なくなってしまう。その結果、青森県発の旅行目的別宿泊先選択比率の精度が低くなると同時に、場合によってはその値がゼロとなる旅行先都道府県が多くなってしまふ可能性が考えられる。例えば青森県居住者のデータを見ると、旅行目的別宿泊先選択比率は46都道府県中25県で観測数がゼロとなっている。その結果、図一11に示したように目的別の旅行者数も大半の場合でゼロもしくは非常に小さな値となっている。旅行目的を推計するにあたって利用可能な実績値として、秋にしか行われていない幹線調査の情報のみを用いている本提案手法の限界であると言えよう。この課題を解決するためには、今後の幹線調査等において、人口の少ない地



方部においては、旅行目的別の宿泊先選択比率 ( $R$ ) を、季節等による差異を考慮した上でより精度よく把握できるよう、サンプル抽出率を大きくするような対応をとることが考えられる。あるいは、季節毎に行われている他の統計調査の補完的な活用等も考えられる。

### 4.3 既存旅行統計の目的別旅行量との比較

本節では、旅行目的別宿泊先選択比率 ( $R$ ) とモバイル空間統計データ ( $Q$ ) から算出した目的別都市間の旅行発生量 ( $V$ ) に関して、別の既存統計あるいは幹線調査から得られる年単位の目的構成比との比較を通じて、提案手法の妥当性を検証する。

#### 4.3.1 年単位での検証

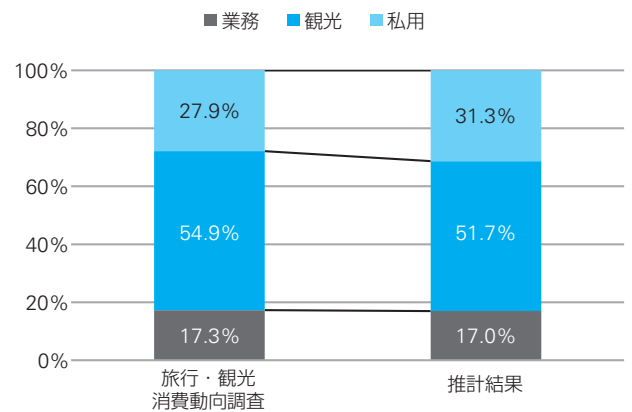
年単位での検証を行うにあたり、地域間旅客流動に関する第三のデータソースとして、観光庁が行っている旅行・観光消費動向調査<sup>11)</sup>を用いる。この調査は、過去三ヶ月間における旅行回数、旅行目的、旅行日数、目的地、居住地等の情報を、無作為のアンケート形式で収集する調査である。調査は三ヶ月毎に年間4回行われており、季節間の差異を把握することもできる。

本研究では、2015年の三ヶ月毎計4回の旅行・観光消費動向調査から、年間の目的別旅客流動量を算出し、その上で旅行目的構成比率を算出した。その結果を、本研究で提案したデータ融合手法を用いて算出した2015年の年間旅行目的別構成比率と比較した(図一12)。幹線調査とモバイル空間統計データを融合した推計結果と、旅行・観光消費動向調査から得られた推計結果が類似していることが確認される。これは、質的情報を持つ1日分の情報(幹線調査)と量的情報のみ1日毎1年分のデータ(モバイル空間統計)を融合することで、通年での旅行目的別旅客流動量の推計を適切に行うことができる可能性を示唆している。

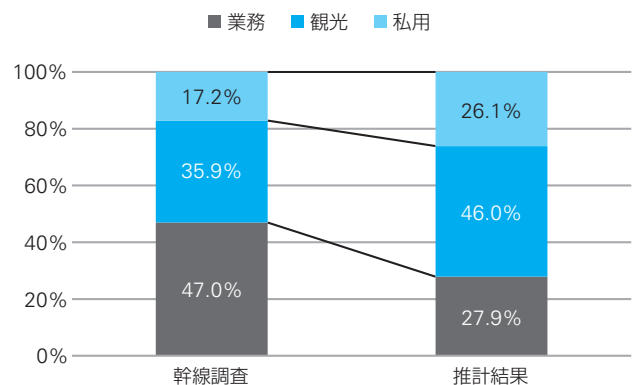
#### 4.3.2 日単位での検証

次に、日単位での検証として、2010年度の幹線調査実施日である2010年11月28日(日・休日)、及び、同年12月1日(水・平日)の両日における幹線調査の推計結果そのものと、本研究で提案するデータ融合手法による推計結果とを比較した。なお、モバイル空間統計では2010年のデータ提供はなされておらず、ここでの比較は2015年11月29日(日)、12月2日(水)両日を対象としたデータ融合結果を用いている。

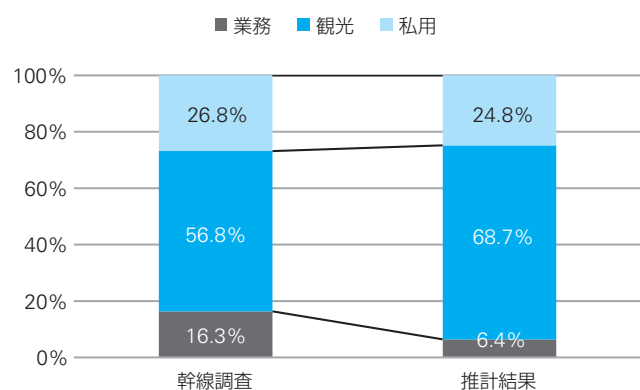
図一13に平日の比較結果を、図一14に休日の比較結果を示す。年単位の場合と異なり、平日・休日のいずれの場合においても、データ融合により推計された旅行目的別流動比率と幹線調査から得られた旅行目的別流動比率にはやや大きな乖離が見られることが分かる。具体的には、平



■図一12 旅行目的比率の比較(年単位)



■図一13 旅行目的比率の比較(日単位, 平日)



■図一14 旅行目的比率の比較(日単位, 休日)

日において業務目的が約20%程度過少推計になっており、休日においても業務目的が10%以上過少推計となっているのに対し、観光目的が過大に推計されている。この原因として、4.2.2で示したような地方部居住者の流動の推計精度が低いことや、2010年と2015年の間での旅行目的別宿泊先選択比率 ( $R$ ) が変わったことが考えられる。例えば、北陸新幹線開業前後のモバイル空間統計のデータを見ると、石川県と富山県への旅行量は明らかに増加しており、これより $R$ の値も大きく変わったことが推察されるが、本研究では2010年時点の幹線調査から得られた $R$ を用いざるを得なかった。より正確な原因については、2015年の幹線調査データ<sup>注2)</sup>が利用可能となった後にその時点のデータと比較することで明らかにすることができると期待さ

■表—3 モデル残差に基づく旅行日のクラスタリング結果 (3クラスター, 915日分)

| クラスター | 曜日毎の帰属日数 (日)            |     |     |     |     |     |     |    | 残差の範囲             |
|-------|-------------------------|-----|-----|-----|-----|-----|-----|----|-------------------|
| ID    | 月                       | 火   | 水   | 木   | 金   | 土   | 日   | 祝  | (単位: 人/日)         |
| 1     | 111                     | 118 | 120 | 123 | 123 | 0   | 0   | 0  | -25,664 ~ +34,861 |
| 2     | 1                       | 1   | 0   | 0   | 0   | 124 | 123 | 24 | -24,408 ~ +26,093 |
| 3     | 47 (GW, シルバーウィーク, 年末年始) |     |     |     |     |     |     |    | -45,781 ~ +49,121 |

れる。

以上のように、日別という細かい単位で見たとき、(別の時点ではあるものの類似度が高いと想定される) 観測値と提案手法による推計値との間で一定程度の乖離が存在することが確認された。一方、図—7~10で確認されたように、東京都のような大都市の居住者に対しては、(i) 平日において業務目的旅行の割合が大きくなる、(ii) 休日において観光目的旅行の割合が大きくなる、(iii) お盆や正月において私用目的旅行の割合が大きくなる、といった傾向は、提案手法による推計結果にも現れていることが分かる。

### 5——旅行実施日特性とモデル適合度との関連性

本章では、本研究で提案したデータ融合アプローチから得られる日別の推計モデルの残差 (式 (2) の  $\epsilon_{h,s,d}$  の推計値、すなわち、推計した日別OD旅行数の全目的合算値とモバイル空間統計による居住地・旅行先別推計人口の差) の分析を通じて、「どの時期であれば本提案手法の推計結果が信頼できるか?」を検証する。逆に言えば、残差の絶対値が大きい、あるいは、残差の傾向が他と大きく異なる時点、つまり「旅行目的別宿泊先選択比率 ( $R$ ) が一定である」という提案手法の仮定が成立しない条件を探る。仮定が成立し得ないと考えられる時点においては、適切なデータ融合を行うためには、旅行目的等に関する追加的調査等が今後必要になると考えられる。

本研究では、915日×47ゾーン×47ゾーンの組み合わせ毎に行った回帰分析より求められる膨大な量の残差データを取り扱う必要がある。そこで、包括的な評価を効率的に行うため、K-means法による誤差値の分類を行った。X-means法<sup>12)</sup>に基づいて最適なクラスター数を求めたところ2となったものの、意味解釈をより適切に行うことができることから、以下では次善の結果であるクラスター数が3の場合の適用結果を用いて考察する。

クラスタリングの結果を表—3に示す。クラスター分析の結果、通常の平日が全てクラスター1に、通常の休日が全てクラスター2に、そして、連休・長期休暇等の特殊な休日が全てクラスター3にそれぞれ正確に分類される結果となった。例として、2015年8月中の31日のクラスタリング結果を図—15に示す。各クラスターが、平日 (白, クラスター1)、休日 (薄青, クラスター2)、お盆 (濃青, クラスター3) に正確

| 月  | 火  | 水  | 木  | 金  | 土  | 日  |
|----|----|----|----|----|----|----|
|    |    |    |    |    | 1  | 2  |
| 3  | 4  | 5  | 6  | 7  | 8  | 9  |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| 31 |    |    |    |    |    |    |

(クラスターID □:1, ■:2, ■:3)

■図—15 残差による旅行日のクラスタリング結果の一部例 (2015年8月)

に分類されていることが確認される。また、表—3に示した残差の範囲の比較より、クラスター3に帰属する旅行日では特にモデルの残差が大きくなっていることが確認される。すなわち、本研究の提案手法は、GW、お盆、シルバーウィーク、年末年始のような特別な休日における旅客流動を把握する際には誤差が大きくなる可能性が示唆される。これは、特別な休日においては旅行目的別宿泊先選択比率 ( $R$ ) が一定という仮定が妥当でないことを示唆しており、より適切な  $R$  を適用する必要があると考えられる。一方、通常の平日と休日については、残差の範囲は同じような傾向にあることが確認される。

### 6——おわりに

#### 6.1 本研究の結論と今後の課題

本研究では、全国幹線旅客純流動調査とモバイル空間統計データを融合して、日別・旅行目的別の都市間流動量を推計する方法を提案し、ケース・スタディを通じて提案手法の妥当性の検討を行った。得られた主要な知見は、以下のとおりである。

- (1) 日別・旅行目的別生成交通量の推計結果の傾向 (図—3)、及び、旅行・観光消費動向調査との年間値の比較 (図—12) より、全国単位もしくは年単位での旅行目的構成比の推計結果は、概ね妥当なものとなっていることが確認された。
- (2) 地域間流動に関して、東京都居住者による日別・旅行目的別流動量の推計結果 (図—6~10) は、真値が不明であるために厳密に妥当性を示すことはできないものの、明らかに不自然な推計結果は得られなかった。また、紙面の都合上示すことができなかったが、出発地が三大都市圏のような大都市の場合においては、東京都の場合と同様、概ね妥当な推計結果が得られた。

(3) 一方、青森県居住者の地域間流動の推計結果(図—11)は、妥当とは考えにくいものであった。これは、青森県では幹線調査のサンプル数が絶対的に不足しているため、分析に用いた旅行目的別宿泊先選択比率の精度がそもそも低いことに主に起因すると思われる。紙面の都合上示すことができなかったが、他の地方発の推計結果についても同様の傾向が見られることが確認されている。

(4) 旅行実施日特性とモデル適合度との関連性を見ると、特に、ゴールデンウィーク、お盆、シルバーウィーク、年末年始等の特殊な休日において、モデルの残差が相対的に大きくなっている(表—3)。これは、幹線調査では見られなかった旅行パターンが、この時期には多く含まれていることを示唆していると考えられる。

以上を総括すると、提案したデータ融合アプローチによる日別旅行目的別の都市間流動量推計方法に関しては、大都市を出発地とし、かつ、特殊でない平日・休日の旅行目的別流動量を推計する上では大きな問題は確認されなかった。一方、地方都市を出発地とする場合、あるいは、お盆や元日等の特殊な休日の場合には、流動量の再現性、あるいは、適合度の観点から課題が残されていることが明らかになった。

地方発の場合の推計結果の妥当性を向上させる方法を検討すること、日単位の旅行目的構成比の(既存統計との)乖離を小さくするためのモデルの改良を行うこと、第三のデータソースとの融合によって推計結果の妥当性の向上を図ること等が今後の主要な課題であると考えられる。

## 6.2 今後の都市間旅客流動調査に対する示唆

本研究を通じて、時間解像度の高いモバイルビッグデータと、時間解像度は必ずしも高くないが質的情報を得ることができるアンケート調査を適切に融合することにより、季節変動等を考慮した上で旅行目的別の都市間流動を“一定程度”適切に推計できる可能性が示唆された。公的交通調査のコスト削減が求められるようになりつつある昨今<sup>13)</sup>、都市間旅客流動調査の体系においても、幹線調査をはじめとする高コストのアンケート型調査は、量的情報が豊富なモバイル空間統計等の交通ビッグデータへの代替やそれとの融合を前提に、質的情報の調査に重点を置くようにシフトするものと考えられる。

そのような情勢の中、本研究の分析結果は、今後の都市間旅客流動調査が、旅行目的別宿泊先選択比率のような「質的情報」を高い精度で得ることへとシフトすることの必要性を示唆している。本研究のデータ融合手法の根幹である旅行目的別宿泊先選択比率を精度よく推計するためには、同じ交通条件の地域を出発するサンプルの「数」

が重要となるが、現状の幹線調査のように、人口の絶対数が少ない地方において大都市と大差ないサンプル抽出「率」で調査を行っている限り、分析に必要なサンプル数を確保することができない。実際、青森県居住者(図—11)等の場合では、旅行目的別宿泊先選択比率の精度が低いために、妥当とは考えにくい結果が得られている。つまり、データの融合を前提としたとき、重要なのはサンプル抽出「率」ではなく、同じような交通条件にある地域のサンプル「数」であると考えられる。そのためには、大都市部居住者のサンプル抽出率を小さくしてでも、地方部居住者のサンプル「数」を十分に確保できるよう、傾斜をかけたサンプル抽出の実施等が考えられるであろう。無論、幹線調査のどの部分を増加させるべきかや、それが実施主体にとって受容可能かどうかという点については、今後十分な検討が必要である。

以上のように、都市間交通調査に関しては、そのメリットを十分に活かすことができる質的情報に特に着目し、その推計精度の向上を図るような調査設計を行い、その上で、得られた結果を交通ビッグデータと適切に融合することにより、都市間旅客流動の全体像をより適切に把握できるようになると期待される。

**謝辞:** 本研究は、科学研究費補助金・基盤研究(B)「交通関連調査体系の再構築と政策評価への展開」(代表:石田東生, 課題番号:26289171), 同若手研究(B)「ビッグデータの時間情報が拓く、新たな都市間旅行需要予測手法の開発」(代表:山口裕通, 課題番号:17K14736)の一環として実施された。ここに記して感謝の意を表したい。

### 注

注1) ここに示した最終形のモデルに至る前に、独立変数の数を変えた、複数のモデル候補についての検討を行っている。具体的には、ベースである(i)業務・観光・私用の3旅行目的(3独立変数)での分析の他、(ii)業務・観光、観光・私用、業務・私用の2旅行目的(2独立変数)で分析を行い、残りの旅行目的の人数は残差から得るようにしたもの、(iii)平日業務目的、休日観光目的等、業務・観光・私用をさらに平日・休日の2パターンに分け、6旅行目的(6独立変数)で分析を行ったものである。これらの比較の結果、(i)の特定化の場合が最も良い結果を得られたため、本論文では旅行目的が3つの場合の結果のみを示している。

注2) 本論文の投稿時点において、2015年に行われた第6回全国幹線旅客純流動調査のデータはまだ公開されていない。そのため、2010年の第5回調査のデータを使用した。

### 参考文献

- 1) NTT-Docomo, “モバイル空間統計”, <https://www.mobaku.jp/>.
- 2) Kato, H., Kato, K., Endo, K., Kaneko, Y., and Shimizu, T. [2011], “Inter-regional travel demand analysis using integrated model for practical travel demand forecast”, *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 8, pp. 69–84.
- 3) 国土交通省総合政策局 [2010], “第5回幹線旅客純流動調査・幹線旅客流動の実態”, <http://www.mlit.go.jp/common/001005632.pdf>.
- 4) Axhausen, K.W., Madre, J.L., Polak, J.W., and Toint, P. (Eds.) [2003], *Capturing Long-Distance Travel*, Research Studies Press.

- 5) 室井寿明・磯野文暁・鈴木俊博 [2015], “モバイル・ビッグデータを用いた都市間旅客交通への活用に関する研究”, 「土木計画学研究・講演集」, Vol. 51, Paper No. 167.
- 6) Song, C., Koren, T., Wang, P., and Barabási, AL. [2010], “Modelling the scaling properties of human mobility”, *Nature Physics*, Vol. 6, pp. 818–823.
- 7) Hasan, S., Schneider, C. M., Ukkusuri, S. V., and González, M. C. [2013], “Spatiotemporal patterns of urban human mobility”, *Journal of Statistical Physics*, Vol. 151, No. 1-2, pp. 304–318.
- 8) Alexander, L., Jiang, S., Murga, M., and González, M. C. [2015], “Origin–destination trips by purpose and time of day inferred from mobile phone data”, *Transportation Research Part C: Emerging Technologies*, Vol. 58, pp. 240–250.
- 9) Ge, Q. and Fukuda, D. [2016], “Updating origin–destination matrices with aggregated data of GPS traces”, *Transportation Research Part C: Emerging Technologies*, Vol. 69, pp. 291–312.
- 10) Janzen, M., Vanhoof, M., Axhausen, K., and Smoreda, Z. [2016], “Estimating long-distance travel demand with mobile phone billing data”, *The 16th Swiss Transport Research Conference*.
- 11) 観光庁, “旅行・観光消費動向調査”, <http://www.mlit.go.jp/kankocho/siryou/toukei/shouhidoukou.html>.
- 12) Pelleg, D. and Moore, A. [2000], “X-means: Extending K-means with efficient estimation of the number of clusters”, *ICML*, Vol. 1, pp. 727–734.
- 13) 石田東生・小向太郎・渡田滋彦・中村正・谷島賢 [2015], “交通政策・運用・調査とビッグデータ”, 座談会, 「交通工学」, Vol. 50, No. 1, pp. 8–17.  
(原稿受付2018年10月2日, 受理2019年1月9日)

## Data-Fusion Approach for Estimating Inter-Regional Passenger Flow for Multiple Days across Different Trip Purposes

By Arata SUZUKI, Hiromichi YAMAGUCHI and Daisuke FUKUDA

In recent years, aggregate mobile phone data such as Mobile Spatial Statistics (MSS) has been increasingly applied for transport data collections. Though MSS enables to obtain the data of various days and seasons easily, it does not contain passengers ‘qualitative information’ such as trip-purposes. Thus we propose a novel data-fusion methodology of Inter-Regional Travel Survey (IRTS) and MSS to estimate inter-regional passenger flow by different trip purposes for multiple days and seasons. We found that: (1) the predicted annual passenger demand by each trip purpose is similar with the predicted one from another statistics; (2) the predicted daily passenger demand is reasonable for some origin prefectures with large cities; but, (3) the predicted daily passenger demand seems to be unreasonable for some origin prefectures in rural areas.

*Key Words* : **inter-regional travel survey, aggregate mobile phone data, data fusion, trip-purpose estimation, seasonal variation**