

長期・大規模なGPSデータから観光滞在判定 を行う閾値の感度分析

生形 嘉良¹・関本 義秀²・Teerayut Horanont³

¹正会員 日本工営株式会社 コンサルタント国内事業本部 (〒102-8539 東京都千代田区麹町5-4)
E-mail:a4882@n-koei.co.jp

²正会員 東京大学生産技術研究所 人間・社会系部門 准教授 (〒153-8505 東京都目黒区駒場4-6-1)
E-mail: sekimoto@iis.u-tokyo.ac.jp

³非会員 東京大学生産技術研究所 (〒153-8505 東京都目黒区駒場4-6-1)
E-mail: teerayut@iis.u-tokyo.ac.jp

本研究は、観光事業への投資が積極化する中でその投資効果の評価や新たな施策立案のために従来型のアンケート等による来訪客状況調査だけでなく、大規模で長期間に継続的にデータが取得でき、かつ比較的迅速にデータが収集できるGPS付き携帯電話による位置情報マーケティングデータを活用した観光実態把握を行うためのデータ処理方法について感度分析したものである。具体的には、観光実態（ここでは日常的な活動を除いた来訪者の行動）の定量的処理を行うための各種閾値について、観光統計で対象とする宿泊者に関して宿泊統計、旅行・観光消費動向調査による宿泊者数との比較、滞在判定を行うための滞在時間処理と滞在日数の関係について評価した。この結果、データ処理を行うための閾値の設定方法および統計処理上の留意点の考察ができた。

Key Words : GPS, GPS付携帯電話, 非集計GPSデータ, 観光, 非日常行動, 観光統計

1. はじめに

観光は裾野が広い産業と言われており、旅行消費による経済波及効果は49.4兆円と国民経済計算産出額の5.5%、雇用は6.6%を占める規模と報告されている¹⁾。平成24年に閣議決定された「観光立国推進基本計画」の中では2016年目標値として日本人国内宿泊旅行および日帰り旅行消費額を2009年比約7%増、雇用効果は同約17%増加と今後も観光産業の成長に期待がされているとともに地域主体の産業としても期待されているところである²⁾。このように、観光は地域活性化の核として捉えられている一方、地域ごとに観光のスタイルが異なるために画一的な調査がしにくいことや観光施設、宿泊業の大半が中小事業者であることから地域一体となった観光実態の把握やマーケティング活動への人的、資金的余力が少ないことも現実である。このような中で、各種取組みの効果を評価し、次の取組みに展開するための基礎となる観光流動実態の把握が求められている。

現在全国的な観光流動実態を把握できる調査には、旅行者に調査する「旅行・観光消費動向調査」や宿泊施設に調査する「宿泊旅行統計調査」、共通基準化が進められた「観光入込客統計」がある。これらは消費者へのア

ンケート調査や各宿泊、観光施設管理者およびイベント主催者への調査であることなどからその調査量には回収率の他、実施調査数の面から費用的な限度がある。例えば、「旅行宿泊統計調査」の平成22および24年10~12月調査³⁾では従業者数10人以上の対象施設数約8,000~9,000施設に対して回収率約65~70%であった。「旅行・観光消費動向調査」⁴⁾は年間25,000人を対象に調査し、国内旅行実人数約95百万人、延べ国内旅行者数約638百万人の規模まで拡大処理をして把握している。

一方、直接的なアンケートや実態把握を行わずに広範囲のメッシュ単位の滞在人口や流動を把握する手法としてGPS付き端末や携帯電話等を使用した分析事例も増えてきている。全国規模の流動を把握する技術として株式会社エヌ・ティ・ティ・ドコモ社が提供する携帯電話ネットワークの基地局情報を使用した「モバイル空間統計」がある⁵⁾。24時間・365日、日本国内を広くカバーしており、数日の解析で1時間単位の人口統計情報が提供できるまでに至っており、数年間隔で実施される統計調査を補完する利用が期待されている他、首都直下型地震が発生した場合の帰宅困難者の推計など予測分野の利用も報告されている。また、長期的な携帯電話基地局情報からユーザーの行動パターンを解析する方法⁶⁾や100万以

上の個人情報が秘匿された携帯電話ユーザーの基地局履歴から行動パターンを分析してユニークユーザーとして統計処理する方法なども報告されている^{7) 8)}。

既存のアンケート調査方式の統計では、観光や業務などの目的が直接的に詳細に把握できるが、大規模で長期間の履歴が蓄積されたGPSデータを基本として推定する場合は、どのような行動を観光滞在として判定していくかについて統計的な処理方法を確立していく必要がある。

従って、本研究は、全国に利用者が分布する大規模で約1年間という長期間のGPSデータを用いて、観光実態を把握するためにどのような抽出条件を設定し、閾値を設けることによって推定できるかを実施したものである。集計した結果が既存統計調査とどの程度整合がとれるものであるか、また閾値の変化による感度はどの程度あるものかを把握したものである。

2. 分析に使用した「混雑統計⑧」

本研究に用いた「混雑統計⑧」データはゼンリンデータコム社の協力により提供されたものである。「混雑統計⑧」データは、許諾を得て取得された位置情報を個人特定出来ないよう統計処理した人の流れデータである。元となる位置情報は、7~80万人程度と言われていることから概ね日本人の200人程度に1人の割合、すなわち日本人の約0.5%をカバーしていることになる。元データの特徴として、「利用者の承諾を得て取得」の他、「最短約5分間隔で取得」、「個人情報との関連性を保持しないIDが付与されている」ことなどがあげられる。本研究ではゼンリンデータコム社の協力のもと、前述した特徴と秘匿処理を行った非集計データを用いている。さらに、1年間の行動履歴を分析するためにデータの解析ができた約19万ユーザーの2012年2月から2013年1月までの1年間データを対象とした。

3. 観光行動者抽出の前提条件の設定

観光入込客統計は、遊園地、公園、イベント会場など観光施設やイベントなどの比較的小さな単位をもとに市町村の集合体として集計されていることが一般的である。有料施設であれば入場者数は比較的容易の集計ができるが、海水浴場や花火大会など入場料を必要としないイベントは大規模な集客があるが実態把握は難しいものもある。このような状況のもと、今回の集計では、俯瞰的で長期的な視点の実態把握を目的として市町村単位などの比較的大きな集計単位で入込客数を把握する場合の滞在者判定を行うための閾値を設定した。

(1) 分析対象地域

分析対象は東北地方とし、具体的には表-1に示す29市町を対象とした。東北地方は2012年3月から約1年間、観光庁と地域が連携した広域的な取組みである「東北観光博」が実施された地域である。また、2011年6月に世界文化遺産に登録された中尊寺、金色堂などがある平泉町もある地域である。

(2) 観光行動と推定するための最低滞在時間的判定条件

滞在者を判定するにあたり排除する必要がある行動として通過交通がある。このため、Google Map, YAHOO! 路線情報を用いて対象市域を通過する一般国道、鉄道（在来線、有料特急）の所要時間を調査し、通過するだけでも最大30~60分程度は必要な状況であった。さらに、観光ツアーパンフレットで観光地の周遊所要時間を調査したところ2~3時間以上のプランが多くみられたこととあわせて、観光行動者の滞在時間として市町域に最短3時間は必要であると仮定した。

(3) 観光行動と推定するための空間的判定条件

観光行動の空間的条件として、滞在者が観光地を訪れた履歴の有無で分類することとした。対象とした観光地は国土数値情報（観光資源）に掲載されている観光資源を基にした（表-2）。ただし、市域の河川全長が対象となっている資源、年中行事のみの資源は除き、通年で観光資源として明確に認知されている場所のみとした。なお、観光地を訪れた履歴の有無は当該観光資源の4次メッシュ単位での滞在時間を30分以上を閾値とした。

表-1 分析対象地域

都道府県名	市町名
青森県	青森市、弘前市、八戸市、むつ市
岩手県	盛岡市、北上市、花巻市、八幡平市、一関市、奥州市、平泉町
秋田県	秋田市、仙北市、大仙市、横手市、鹿角市
山形県	山形市、鶴岡市、米沢市、酒田市、天童市
宮城県	仙台市、大崎市、松島町
福島県	福島市、郡山市、いわき市、会津若松市、猪苗代町

(4) 居住者および就業者の取扱い

市町単位での観光行動を分析するにあたり、当該市町にとっての日帰り観光と宿泊観光がある。後の比較検証を「旅行宿泊統計調査」と行うことを考慮して本研究では当該市町の居住者および就業者は市内観光時に宿泊を伴う観光は少ないと考えて分析対象外とした。なお、居住地および就業地については前年の2011年1年間のGPS履歴から日中および深夜時間帯の滞在履歴場所および頻度をもとに4次メッシュ単位で推定した。

以上の観光行動者抽出条件を整理すると図-1に示す抽出フローになる。

4. 観光行動者と非観光行動者の特徴

(1) 観光地滞在／非滞在者数の構成

観光行動を推定するための観光地滞在履歴者と非滞在履歴者数の構成比を集計した。この結果、観光地滞在履歴者率が大きく異なる結果となった(図-2)。観光地滞在履歴者率が20%を超える弘前市、平泉町、仙北市、松島町、会津若松市は観光入込客数の規模の大きい観光資

表-2 主な観光資源(抜粋)

市町名	主な観光資源
青森県弘前市	弘前城, 岩木山, 岩木高原
岩手県平泉町	中尊寺, 毛越寺庭園, 無量光院跡
秋田県仙北市	角館武家屋敷, 田沢湖
宮城県仙台市	青葉城跡, 秋保大滝, 輪王寺
宮城県松島町	松島観瀾亭, 瑞巖寺, 五大堂
福島県会津若松市	若松城跡(鶴ヶ城跡), 飯盛山白虎隊墓所

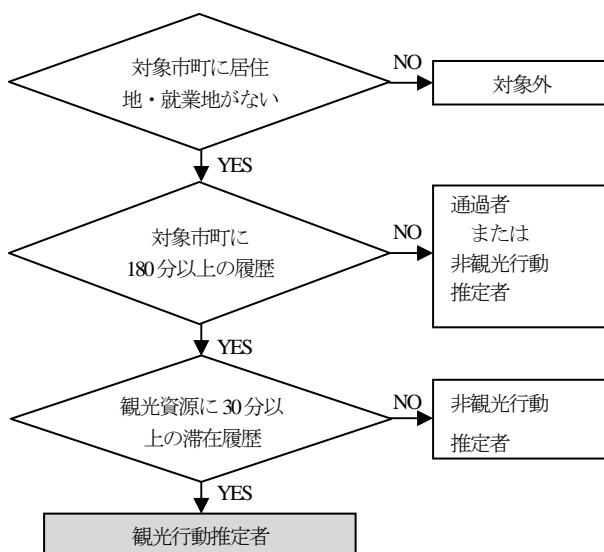


図-1 観光行動者抽出フロー

源が備わっている市町であり、これらは全国的に著名な資源であるとも言える。特に平泉町、松島町の観光地滞在履歴者はそれぞれ62%、44%を占めており他の市町に比べてもその比率が高いことがわかる。ただし、この時点では集客規模が大きい観光地では業務目的で観光地に来訪する人も多いと考えられるため来訪頻度によるリピート特性とあわせて評価する必要がある。

(2) 分析対象市町と分析対象月の設定

当該市町村居住者、就業者を除いた市域滞在者には観光行動の他、業務や私用など様々な行動が含まれる。さらに、業務の中でもルートセールスのようにウィークデーはほぼ毎日のような高頻度で来訪する場合や低頻度であるが定期的に支社や工場に来訪する場合も考えられる。

このような来訪目的の違いを区分するために月間の滞在日数を分析した。この分析には、前項で集計した結果を用いて、観光地滞在履歴者と非滞在履歴者の違いが明確であり、観光行動の特性が見つけやすいと考えられる観光地滞在履歴者率20%以上の5市町を中心に分析した。

集計対象とした月は、観光行動の滞在日数の長短に影響すると考えられる月間の休暇特性を考慮し、滞在が長期化する可能性のあるGWを含む5月および祝祭日が無い6月を対象に集計した。

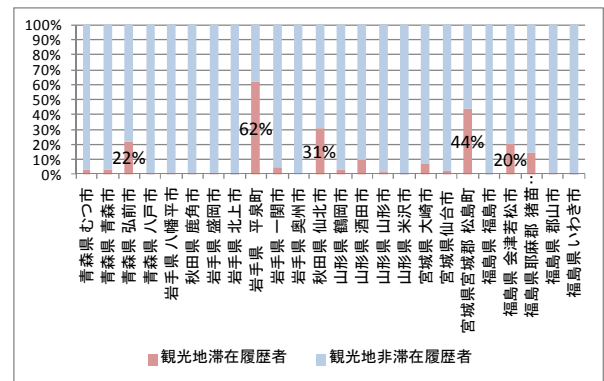


図-2 観光地滞在／非滞在者数の構成

(3) 月間滞在日数でみる観光行動者の特徴

観光行動以外の行動が主と考えられる来訪頻度が比較的高い行動者がどの程度存在するかについて観光地滞在履歴の有無別に集計した(図-3)。

この結果、観光地滞在履歴の有無によらず滞在日数1日が最も多いことがわかる。滞在日数1日の滞在者数に比べて2日の滞在者数は大きく減少している傾向がみられる。さらに滞在日数3日になるとさらに大きく減少し、以降は分析した3カ月でそれほど大きな変動がなく安定的に滞在者が存在する状況であった。

なお、滞在日数1日は当該市町には宿泊しない行動形態であり、2日は日帰りで2回滞在した場合と1泊2日の行動が含まれる。3日以降も同様である。21~31日の滞在履歴者もいるがこれらの多くは居住地、就業地を評価した2011年以降にそれぞれの場所が変更になっていた例が多く含まれていたものである。

また、滞在日数1日に比べて2日以降の滞在者数が大きく落ち込んでいる傾向が顕著にみられた市町として平泉町があり、観光資源に来訪する人は月間に1日しか来訪しない行動が主であることがうかがえる(図-4)。

滞在者数のボリュームが大きく月間の変動も大きい滞在日数1日および2日の滞在者数に占める比率をみると観光地滞在履歴者の7~8割を占め、観光地の滞在履歴無しも含めた総滞在者数の6~7割を占める状況である。

(4) 年間の来訪頻度でみる観光行動者の特徴

同一月間での滞在日数の他、月間滞在日数は低頻度であるが年間を通じて多頻度に来訪する滞在者の状況を分析した。観光地滞在履歴者をベースに月間1日の滞在者の他の月の滞在状況をみると他の月には来訪しない、すなわち年間に1回しか滞在しない人が約7~8割を占めた(図-5,7)。月間2日の場合は他の月に来訪しない人は各市町で10ポイント程度低い値であり、月間3日までは他の月には来訪しない人の比率が高かった。

また、観光地滞在履歴の有無の違いにより月間1回の滞在者の他月の滞在頻度の分布をみると観光地非滞在履歴者は当月のみが半数を下回る構成比であり、さらに他に6カ月以上にわたって滞在した人が約1~2割を占める状況であった。(図-6,8)。これらの結果から月間のみ来た場合には暦の休暇日数の特性に関わらず月間4日以上滞在履歴者は半数以上が年間のうち6カ月以上にわたり滞在しているケースであり、滞在日数1日および2日の行動とは異なる行動形態ではないかと考えられる。例えば、観光地滞在履歴者で高頻度に滞在するケースでは同一観光資源(溪谷)に滞在しているサンプルもみられ、業務、私用の他、趣味に関するレジャー関連行動などが想定される。

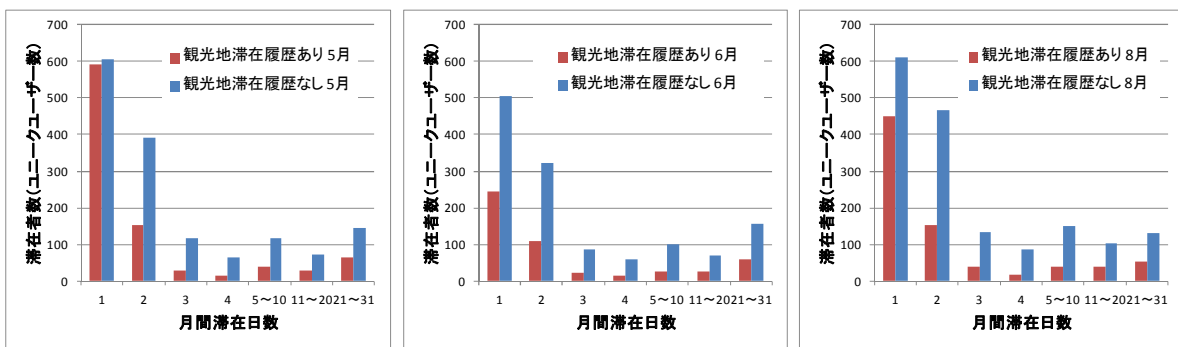


図-3 月間滞在日数別滞在者数(観光地滞在履歴の有無別) 観光地滞在履歴者率20%以上の5市町の合計

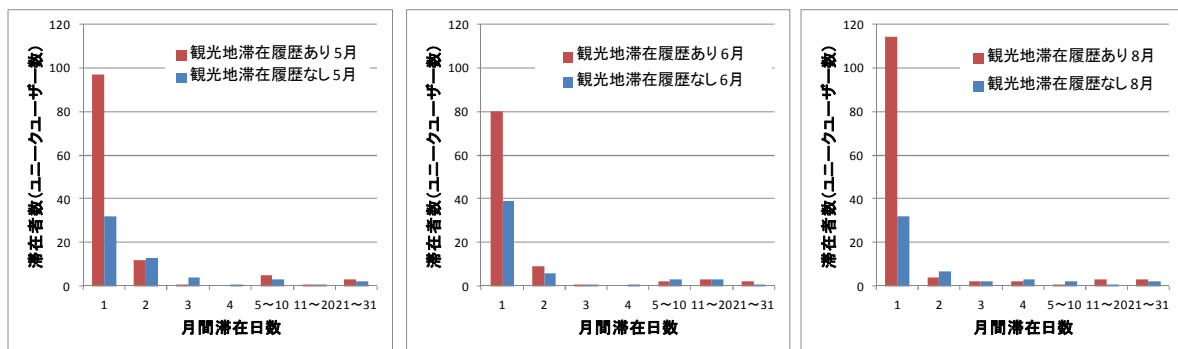


図-4 平泉町の月間滞在日数別滞在者数(観光地滞在履歴の有無別) 観光地滞在履歴者率20%以上の5市町の合計

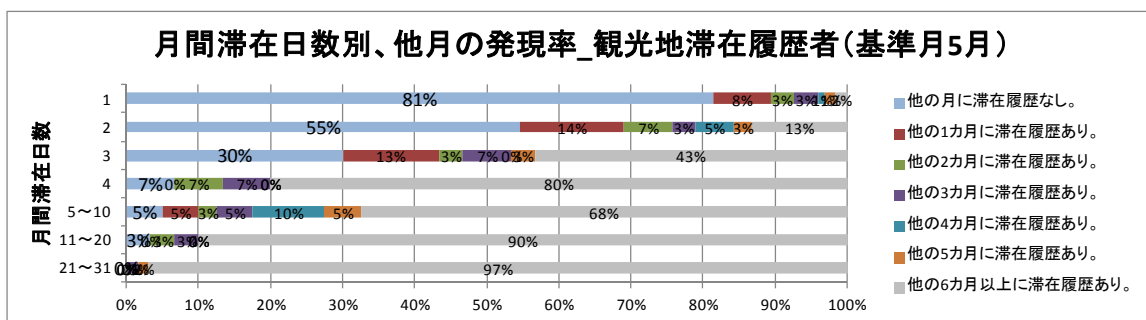


図-5 2012年5月の月間滞在日数別にみた他の月の滞在履歴状況(観光地滞在履歴者)

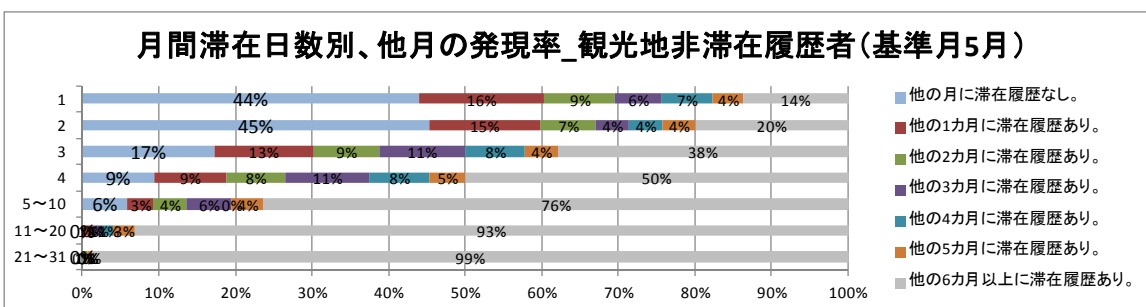


図-6 2012年5月の月間滞在日数別にみた他の月の滞在履歴状況(観光地滞非在履歴者)

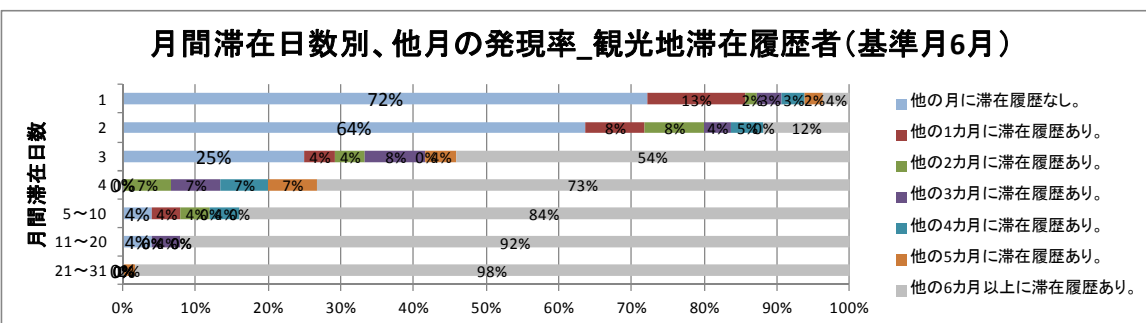


図-7 2012年6月の月間滞在日数別にみた他の月の滞在履歴状況(観光地滞在履歴者)

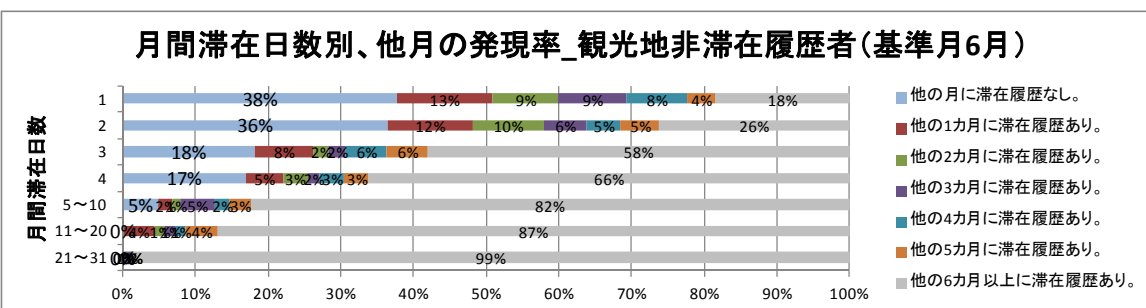


図-8 2012年6月の月間滞在日数別にみた他の月の滞在履歴状況(観光地滞非在履歴者)

図-5~8は観光地滞在履歴者率20%以上の5市町の合計

5. 観光行動者の閾値感度分析と統計データとの比較

集計の量的信頼性を確認するために既存統計データとの比較を行った。閾値として月間滞在日数2日から10日までの滞在者を対象に28市町の宿泊日数を集計し、感度分析を行った(表-1に示した29市町のうち平泉町は市町別宿泊旅行統計調査結果が無いため除外し28市町とした)。

なお、宿泊の集計の定義は、時間的には午前零時を含む連続7時間の滞在とし、空間的にはその滞在時間が一つの市町内であることとした。

(1) 宿泊旅行統計調査結果による平均宿泊日数

はじめに比較した統計資料は「宿泊旅行統計調査」であり、宿泊場所の対象はホテル、旅館、ペンション・民宿・貸別荘の宿泊施設に限定されており実家、知人宅や

ウィークリーマンションといった施設は調査対象外となっている調査である。これらの施設の東北6県合計の平均宿泊日数（延べ宿泊者数/実宿泊者数）は全数で1.34泊/回（2012年4-6月）であった。また宿泊目的の違いを考慮すると観光目的が主体の施設では1.20泊/回、観光目的以外が主体の場合は1.45泊/回と差があるが平均的には1~2泊が主体であることが想定される。

(2) 市町単位の宿泊者数の推定処理

「宿泊旅行統計調査」は個別宿泊施設の回答サンプルのみを市町単位で集計し、都道府県単位の推計値としているため市町単位の総数としての推計値は公表されていない。分析では市町別での相対的なバランスも把握するため宿泊旅行統計調査結果から市町単位の実宿泊人数を次のように推定した(1a)。この結果を都道府県に含まれる市町の回収サンプルのみの宿泊実人数の合計と県単位に拡大された宿泊実人数との倍率1.69~2.29を各市町に一律に乗じて市町別想定宿泊実人数とした(表-3)。

$$\text{市町の推定実宿泊人数} = \frac{\text{県単位の実宿泊者数}}{\text{県内市町村のサンプル実宿泊者数}} \quad (1a)$$

(3) 月間滞在日数の違いによる宿泊者数の感度分析

表-1に示した東北地方の28市町の宿泊者数の合計について2012年4~6月の3カ月について全体傾向をみた。この結果、市町別想定実宿泊人数と近くなる月間滞在日数は3日となった。実際の行動としては2泊3日までの行動

が対象となる結果であった。これは、延べ宿泊人数でも同様であった(図-10, 11)。ただし、集計対象とする月間滞在日数を拡大した場合、宿泊日数が多い対象が含まれるため宿泊延べ人数は実人数に比べて統計値との乖離が大きくなる傾向にある。

また、市町別の実宿泊人数について宿泊実人数の規模と推計値の差異の大きさについて考察した(図-12)。この結果、規模が小さい市町ほど乖離が大きい傾向にあり、このような市町を個別に取り扱う場合には留意が必要と考えられる(図-13)。

(4) 全ての宿泊場所を考慮した場合の感度分析

先述したとおり「宿泊旅行統計調査」は宿泊場所として実家や知人宅への宿泊は調査対象としていない。本研究に用いたGPSデータでは詳細な宿泊場所までは分析していないため宿泊場所も対象を広げて評価する必要がある。宿泊場所を限定しない宿泊者数が推計されている「旅行・観光消費動向調査」を用いて分析した。延べ宿泊数の宿泊施設内訳から推定すると宿泊旅行統計で集計対象としているホテルおよび旅館等の延べ宿泊者数に対して全宿泊場所の延べ宿泊数はその約1.7倍(2012年4-6月)である。

このような差の要因を把握するために宿泊旅行目的や宿泊場所別の延べ宿泊者数の内訳をみると、帰省・知人訪問等の私用が37%を占め(図-14)、この目的の宿泊場所は他の2つも目的と異なりホテルや旅館等以外の場

表-3 県別の拡大率（2012年4-6月）

都道府県名	4月	5月	6月
青森県	1.86	2.02	1.94
岩手県	1.84	1.92	1.82
秋田県	1.69	1.74	1.78
山形県	2.08	2.25	2.27
宮城県	1.85	1.89	1.86
福島県	2.19	2.25	2.29

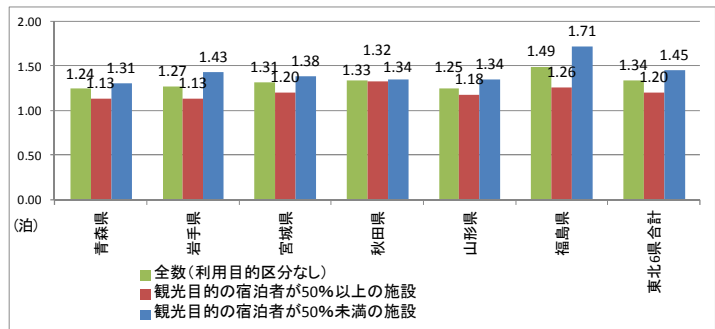


図-9 平均宿泊日数（2012年4-6月）（旅行宿泊統計調査から作成）

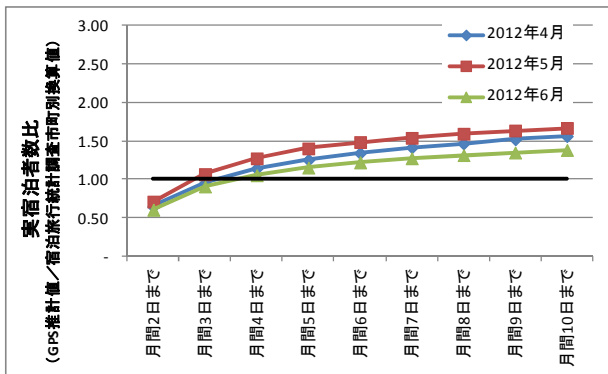


図-10 実宿泊人数と月間滞在日数の関係

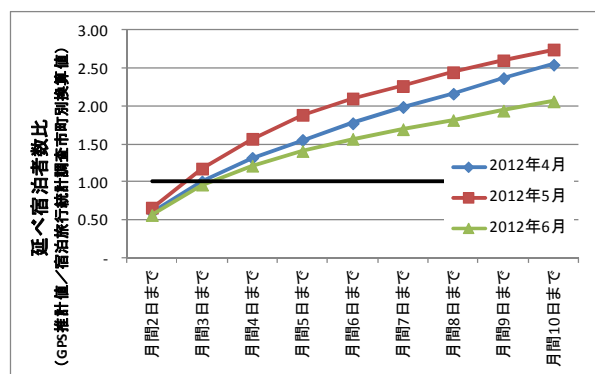


図-11 延べ宿泊人数と月間滞在日数の関係

所が約8割を占める状況であることがわかる(図-15)。また、2012年4-6月の3カ月間の平均宿泊日数は7.41泊(図-16)であり当該目的の旅行実施頻度1.5回/3カ月で1回あたりの宿泊日数を換算すると4.94泊/回と推定さ

れる。

先述の図-11に示した集計結果と照合すると月間滞在日数を5~6日まで拡大するとホテルや旅館等を対象とした延べ宿泊者数の約1.7倍になるという結果が得られた。

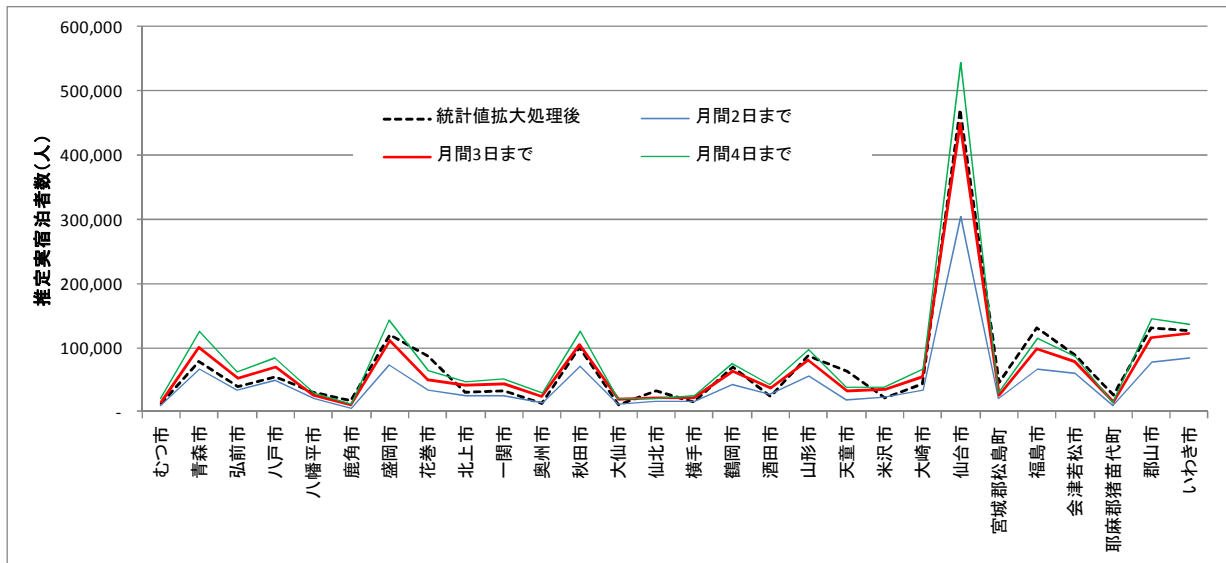


図-12 対象28市町の実宿泊者数の統計値(拡大処理後)とGPS推計値との比較(2012年4月)

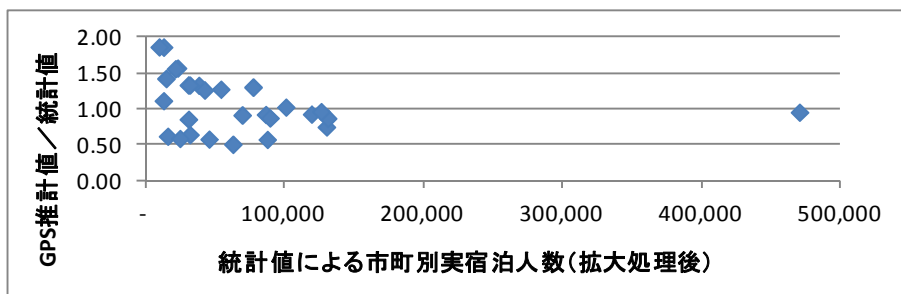


図-13 対象28市町の実宿泊者数の統計値とGPS推計値との差異(2012年4月)

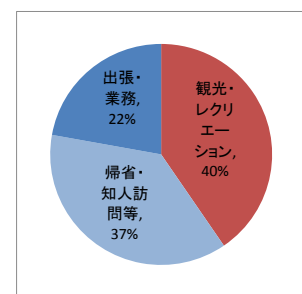


図-14 国内宿泊旅行の目的内訳(2012年4-6月)

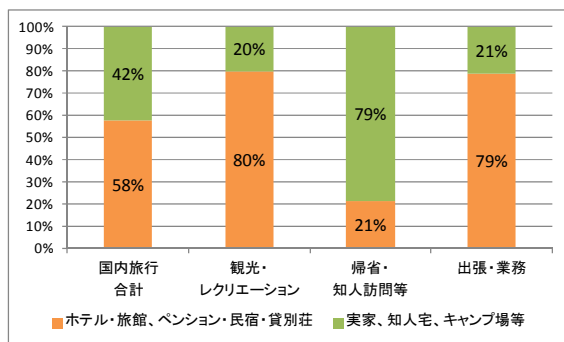


図-15 宿泊旅行目的別宿泊場所別延べ宿泊者数構成比

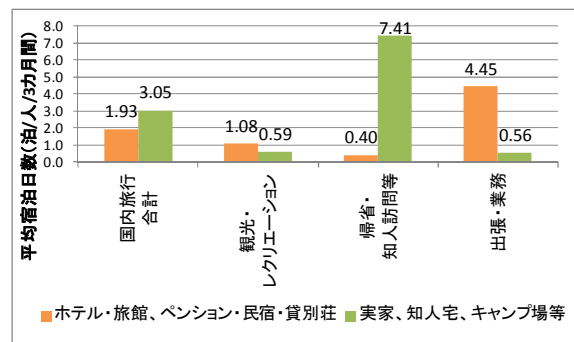


図-16 旅行目的別3カ月間の平均宿泊日数(2012年4-6月)

6.考察

本研究では、大規模で長期間に及ぶGPSデータを用いて観光、旅行といった非日常的行動に着目した分析を行った。

観光旅行という点では、著名な観光地の滞在履歴と年間の来訪頻度を用いて観光行動を抽出することが可能ではないかということがわかった。年間の来訪頻度を把握するためには長期間にわたり同レベルでのデータ収集技術が特に効果を発揮したといえる。対象を全ての旅行目的に広げる点では月間滞在日数を閾値として観光地滞在比率の高い市町で分析した月間滞在日数の閾値を用いて他市町にの宿泊者数を統計データと比較したが概ね整合することが確認できた。しかしながら、観光行動だけを抽出した集計までには至らなかったことおよび宿泊者数が少ない市町では乖離が大きい傾向であったことについては今後の検討課題としたい。

また、本研究では、来訪頻度の違いによる具体的な行動内容についてはサンプリングしたに留まっているが、今後は滞在場所の細分化やラベリングを行い、行動に基づいた対象ごとの集計を行うために研究を進めていきたい。

謝辞：本研究にあたり、分析用データを提供いただいた株式会社ゼンリンデータコム佐藤様、足立様、東京大学空間情報科学研究センター柴崎研究室の檜山武浩特任助

教、東京大学生産技術研究所の金杉洋特任研究員にはここに深謝の意を表す。

参考文献

- 1) 観光庁：「旅行・観光産業の経済効果に関する調査研究」の概要について 別紙, pp.1-2, 2012.
- 2) 観光庁：観光立国推進基本計画 参考資料, pp.3, 2012.
- 3) 観光庁：旅行宿泊統計調査 平成 22,24 年 10~12 月分, 参考第 5 表, 2011 および 2013.
- 4) 観光庁：2010 年旅行・観光消費動向調査 年報, 第 2, 3 表, 2011.
- 5) モバイル空間統計による社会・産業の発展に関する研究会：社会・産業の発展に寄与する「モバイル空間統計」利活用のあり方に関する報告書（最終案） pp.14-15, 2010.
- 6) Marta C. Gonzalez, Cesar A. Hidalgo & Albert-Laszlo Barabasi: Understanding individual human mobility patterns, NATURE, Vol 453, pp.779-782, 5 June 2008.
- 7) Santi Phithakkitnukoon, Zbigniew Smoreda, Patrick Olivier: Socio-Geography of Human Mobility: A Study Using Longitudinal Mobile Phone Data, PLoS ONE 7(6):e39253. doi:10.1371/journal.pone.0039253, pp.1-9, June 2012.
- 8) Yves-Alexandre de Montjoye^{1,2}, Cesar A. Hidalgo^{1,3,4}, Michel Verleysen² & Vincent D. Blondel: Unique in the Crowd: The privacy bounds of human mobility, Sci. Rep. 3, 1376; DOI:10.1038/srep01376 (2013).

(2013.5.7 受付)