



日本銀行ワーキングペーパーシリーズ

位置情報データによる 経済活動のナウキャストイング

王悠介*

yuusuke.ou@boj.or.jp

須合智広*

tomohiro.sugou@boj.or.jp

高橋耕史*

kouji.takahashi-2@boj.or.jp

松村浩平*

kouhei.matsumura@boj.or.jp

No.21-J-2
2021年3月

日本銀行
〒103-8660 日本郵便（株）日本橋郵便局私書箱 30号

* 調査統計局

日本銀行ワーキングペーパーシリーズは、日本銀行員および外部研究者の研究成果をとりまとめたもので、内外の研究機関、研究者等の有識者から幅広くコメントを頂戴することを意図しています。ただし、論文の中で示された内容や意見は、日本銀行の公式見解を示すものではありません。

なお、ワーキングペーパーシリーズに対するご意見・ご質問や、掲載ファイルに関するお問い合わせは、執筆者までお寄せ下さい。

商用目的で転載・複製を行う場合は、予め日本銀行情報サービス局 (post.prd8@boj.or.jp) までご相談下さい。転載・複製を行う場合は、出所を明記して下さい。

位置情報データによる経済活動のナウキャストイング*

王 悠介[†] 須合 智広[‡] 高橋 耕史[§] 松村 浩平^{**}

2021年3月

【要旨】

本稿では、高頻度のビッグデータのひとつである携帯電話の位置情報データを用いて、各地点の滞在人口から、サービス業の売上動向や製造業の生産活動をナウキャストイングする手法を提示する。

具体的には、小売・娯楽業などのサービス業については、各地点の来店客数は滞在人口で概ね近似可能である点に着目し、遊園地、ショッピングセンター、飲食業に関する活動指標を作成する。滞在人口にノイズが多く含まれる場合でも、クラスタリングといった統計的手法を適切に組み合わせることで、既存統計では迅速な把握が難しいサービス業の活動についても、高い精度でナウキャストイングできることを示す。

次に、製造業については、経済センサス活動調査の事業所データや昼間人口比率などの時間帯別の滞在人口分布を利用して、比較的大規模な工場敷地を特定したうえで、その地点の滞在人口をベースに生産活動をナウキャストイングする指標を作成する。業種別にみると、輸送機械や生産用機械といった労働集約的な業種の生産については、工場敷地の滞在人口によりかなり高い精度でナウキャストイングできることを示す。

本稿の分析結果は、位置情報データが、マクロ的な経済活動をナウキャストイングするツールとして有効である可能性を示唆している。

JEL 分類番号 : C49、E23、E27

キーワード : 位置情報データ、ナウキャストイング、クラスタリング

* 本稿の作成に当たり、日本銀行の亀田制作氏、佐久田健司氏、川本卓司氏、神山一成氏、桜健一氏、中島上智氏、飯田智之氏、東京大学金融教育研究センター・日本銀行調査統計局共催ビッグデータフォーラムの参加者から有益なコメントを頂いた。また、加来和佳子氏からは、計数作成においてご協力を頂いた。記して感謝の意を表したい。ただし、あり得べき誤りは筆者個人に属する。本稿で示されている見解は、日本銀行の公式見解を示すものではない。

[†] 日本銀行調査統計局 (yuusuke.ou@boj.or.jp)

[‡] 日本銀行調査統計局 (tomohiro.sugou@boj.or.jp)

[§] 日本銀行調査統計局 (kouji.takahashi-2@boj.or.jp)

^{**} 日本銀行調査統計局 (kouhei.matsumura@boj.or.jp)

1. イントロダクション

2010年代以降、急速な情報通信技術の進展を背景に、伝統的な統計調査やアンケート調査では不可能であった大規模なデータの収集が可能となり、実務的にもその活用が進んでいる。とりわけ、マーケティングの分野などでは、顧客の属性に応じた消費行動を把握する手段として、スマートフォンから得られる位置情報などのビッグデータの利用が拡大してきた。この間、経済学、とくにマクロ経済の分野では、1990年代以降、ビッグデータのひとつであるPOSデータを利用した、スーパーの売上高や価格の分析は進んでいたが、2020年初の新型コロナウイルス感染症の流行以降、位置情報データを中心に、ビッグデータの活用は一段と拡大している。

すなわち、感染症の拡大とそれに伴うロックダウン・緊急事態宣言の影響などにより、各国の経済活動は短期間で大きな変動を示し、政策当局による迅速な景気判断の重要性は飛躍的に高まった。しかしながら、従来、中央銀行が頼ってきた伝統的な統計調査やアンケート調査では、データの収集や集計・加工といった作業に相応の時間を要するため、公表までに数週間から1~2か月を要することが多い。こうした課題に対応するため、中央銀行の間では、企業へのヒアリングといったマイクロ調査とともに、ビッグデータ、とりわけ速報性の高い位置情報等の高頻度データを活用する動きが、急速に広がっている¹。

本稿では、こうした問題意識のもと、ビッグデータの中でも、位置情報データを用いて、マクロ的ないし産業レベルでみた経済活動をリアルタイムにナウキャストリングする手法を提示する。具体的には、位置情報データと施設情報等を組み合わせ、どの産業・セクターについて、高頻度かつ高精度に経済活動をナウキャストリングできるのかを包括的に検討した。本稿の分析結果によれば、位置情報データから、特定の経済活動と結びつけられるような人の動きをうまく抽出することによって、家計の消費活動や企業の生産動向を精緻にナウキャストリングすることができることが分かった。こうした結果は、マーケティングなどでは利用されていた位置情報というビッグデータが、マクロ経済の情勢判断にも資することを示唆している。

本稿の構成は以下のとおりである。第2節では、先行研究を紹介したうえで、本分析で使用するデータについて解説する。第3節では、位置情報データを用いてサービス業の動向を分析したうえで、第4節では、鉱工業生産の動向を分

¹ ここでは、①頻繁に更新される「高頻度データ」、②個別企業間の取引のような「高粒度データ」、③SNSの書き込みのような「テキストデータ」をビッグデータと呼ぶことにする。

析する手法を提示する。第 5 節は、結びと位置情報データを利用する際の留意点について述べる。

2. 先行研究と本分析で使用するデータ

ここでは、本稿に関連する先行研究を紹介したうえで、本分析で使用するデータについて解説する。

2. 1 先行研究

本稿は、主に二つの分野の先行研究に深く関連する。一つ目は、スマートフォンの位置情報データを利用して、人出の状況を分析している研究である。特に、新型コロナウイルスの感染拡大以降は、ロックダウンの影響を分析したものが多い。例えば、米国については、Couture et al. (2021)が、スマートフォンの位置情報データを用いて、各地域間の移動指標 (location exposure index) や、対面接触に関する指標 (device exposure index) を開発している。また、Coven and Gupta (2020)は、スマートフォンの位置情報データを用いて、感染症の拡大以降、高所得者は、都市部から避難する傾向がみられる一方、低所得者は引き続き外出し、元々の職場での労働に従事している傾向が強いことを指摘している²。中国については、Fang et al. (2020)が、スマートフォンの位置情報を利用してロックダウンの影響を検証している。日本に関する位置情報データの分析例としては、Watanabe and Yabu (2020)が挙げられる。彼らは、携帯電話の位置情報から作成した「stay-at-home 指標」を用いて、日本人の自粛行動が、①政府による緊急事態宣言の直接的な結果 (intervention effect) によるものか、②感染者数の報道なども含めたアナウンスメントの影響 (information effect) によるものかを、定量的に推計している。もっとも、これらの分析事例は、人の移動を対象とした分析であり、経済活動との明確な関連性については分析されていない。

本稿が関連する二つ目の先行研究は、ナウキャスト指標の作成に関する研究である。FRB エコノミストである Cajner et al. (2019)は、民間企業である Automatic Data Processing (ADP) のデータを用いて、労働市場を捕捉する指標 (ADP-FRB active employment) を作成した。FRB は、ごく足もとの雇用・賃金動向を把握するために、この指標を利用している。その他、決済情報を用いたナウキャストの研究も近年増加しており、例えば、Aprigliano et al. (2019)は、企業間や企業・消費者間の決済情報が GDP や設備投資・個人消費のナウキャスト

² Chen and Pope (2020)は、2016 年におけるスマートフォンの位置情報データを用いて、高所得者の方が、低所得者よりも移動量および訪問場所が多いことを指摘している。

ティングの精度を高めることを示した。加えて、Galbraith and Tkacz (2018)は、デビットカードや小切手の決済情報が GDP や個人消費のナウキャストリングに有用であることを示したほか、Aladangady et al. (2019)は、クレジット・デビットカードの決済情報を用いて消費動向をリアルタイムに捕捉する指標を開発した。また、新型感染症拡大以降の研究としては、Chetty et al. (2020)が、民間データを用いて高頻度の経済指標を開発した。Chetty らは、この指標をもとに、高所得者層による消費抑制が、そうしたサービスを提供する低所得者層の所得・消費減をもたらしていると分析している。こうした分析手法は、公的統計に先行して経済情勢を把握することを可能にすることから、実務的にも有用性が高い。もともと、これらの分析は、経済活動（消費、雇用・賃金）と明示的にリンクした高頻度データを用いた分析であり、位置情報データは利用されていない。

位置情報を利用したナウキャストリングに関する研究としては、中国における携帯電話の GPS (Global Positioning System) データを用いて、企業の売上げや一部のサービスセクターの消費動向をナウキャストリングする指標を開発した Dong et al. (2017)が挙げられる。この他にも、Arslanalp et al. (2019)や Cerdeiro et al. (2020)は、船舶の位置情報を用いて、世界の貿易量を捕捉する指標を作成しているが、こうした先行研究はこれまでのところ非常に少ない。

本稿は、こうした 2 つの先行研究の流れを組み合わせたものであり、本稿の貢献は、位置情報データをもとに、高頻度かつ高精度に経済活動をナウキャストリングできる指標を開発し、マクロ経済分析に位置情報データが有用である可能性を示した点にある。

2. 2 データ

本稿では、位置情報データとして、株式会社 Agoop が提供する 2017 年 1 月～2020 年 3 月の滞在人口データを用いる。本データは、日本全国を 100m 四方のメッシュ——全体では約 2000 万メッシュ³——に区切ったうえで、各メッシュにおける 1 時間ごとの滞在人口を推計したものであり、この推計値は、スマートフォン向けアプリケーションにおいて、同社がユーザーから許諾を得て取得した GPS 位置情報データに基づいている。

もともと、本稿で用いたデータでは、メッシュの緯度・経度情報と時間ごとの滞在人口が利用可能であり、そのメッシュにどのような施設が存在するかという情報は含まれていないため、滞在人口の推移は把握できても、このデータから

³ 山岳地帯など、平均的な滞在人口がゼロに近いメッシュは予め除外されている。

個人消費や生産動向などの経済活動を直接的には計測できない。そこで、本稿では、各メッシュにおける滞在人口がどのような経済活動を表しているかを捉えるため、①経済センサス活動調査（経済産業省）の個票データ、②国土数値情報（国土交通省）、および、③商業施設関連の API (Application Programming Interface) などと組み合わせて、各メッシュに経済的な意味付けを行い、消費や生産動向を捉えることができる指標への変換を試みている⁴。

具体的には、各メッシュに含まれる施設・建物を特定し、そのメッシュにおける主たる経済活動が、消費関連（小売・娯楽・飲食）であるのか、または生産活動（工場）であるのか、という紐づけを行う。このような作業を行うことで、人出データの動きから経済活動を捕捉することが可能となる。

2. 3 指数の作成方法

本稿では、上述の滞在人口データを用いて、サービス業などの経済活動を計測する指標（Economic Indicator from GPS Data: EIG）を開発する。具体的には以下のステップにより作成する。

まず、ターゲットとなる業種 J を決定し、時点 t において、その業種と関連するメッシュの集合 I_t^J を抽出する。そして、選ばれたメッシュ i の滞在人口 (Pop_{it}) を、経済センサス活動調査など他のデータを用いて算出したウエイト (w_{it}) に基づき集計することにより、全体の滞在人口 (TotalPop_t^J) を以下のように算出する。

$$\text{TotalPop}_t^J = \sum_{i \in I_t^J} w_{it} \times \text{Pop}_{it}$$

また、多くの業種において、全ての企業や施設をカバーするメッシュが集合 I_t^J に含まれるわけではないことなどから、以下のように、集計された滞在人口を基準化した指数を EIG として定義する。

$$\text{EIG}_t^J = \frac{\text{TotalPop}_t^J}{\text{TotalPop}_s^J} \times 100$$

⁴ 経済センサス活動調査や商業施設関連の API を利用することにより、各メッシュに存在する企業の業態データなどを取得することができる。また、国土数値情報は、土地の利用用途に関する情報が取得できることから、滞在人口のデータと経済活動に関する関連付けを行うことが可能となる。

ここで、 $TotalPop_s^j$ は、基準時点 s の滞在人口を表す。なお、上で説明したように、滞在人口は1時間毎のデータが利用可能である。このため、伝統的なデータと同様の月次等の周期からみた分析だけでなく、時間毎や週毎のデータを用いた柔軟な分析が可能になる。

3. 位置情報データとサービス業

本節では、位置情報データを用いて、サービス業の経済活動をナウキャストリングする指標を作成する。サービス業については、それぞれの施設における滞在人口が、来店客数と相関している可能性が高いと考えられる。すなわち、施設における滞在人口を正確にカウントすることができれば、来店客数または売上高をナウキャストリングすることが可能となる。

もっとも、あるメッシュの中で、特定のサービス業の活動だけが行われているというケースは稀である。例えば、メッシュの中に、百貨店やスーパーもあれば、レストランも存在するというケースは多くみられる。そうしたケースについて、飲食サービスの動向を把握するためには、百貨店やスーパーへの人出の影響を強く受けるメッシュを除去することが、高い精度でナウキャストリングするための重要な要素となる。

こうした問題意識のもと、以下では、ノイズを除去するための統計的手法なども用いながら、①遊園地（娯楽業）、②ショッピングセンター、③飲食業の活動をナウキャストリングする指標を開発し、既存の経済指標と比較することにより、その妥当性を検証する⁵。

3. 1 遊園地（娯楽業）

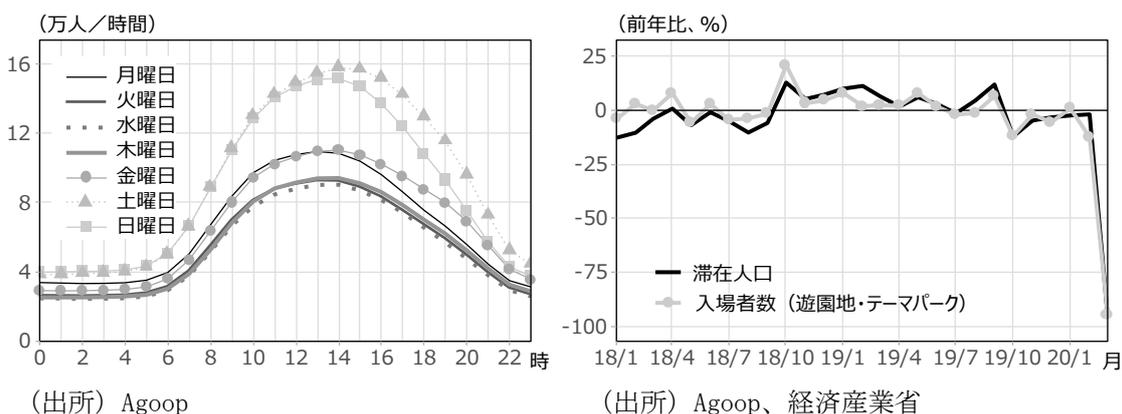
遊園地は、滞在人口が入場者数を近似していると考えられるほか、主要な施設数が少ないため、カバレッジを高めることが比較的容易なセクターと考えられる。ここでは、日本における主要なテーマパークの全時間帯の合計滞在者数を計算した。

まず、曜日・時間帯別の滞在人口をみると、土日昼間や、平日のなかでも週末の前後となる月・金曜日に当該メッシュを訪れる人が多いことがわかる（図表1）。これは、遊園地の一般的な特性と一致する。次に、滞在人口（前年比）の推

⁵ Chetty et al. (2020) が指摘しているように、オルタナティブデータを経済分析に利用する際には、データ提供元のシステム変更等から生じる段差等を補正する必要がある。本稿でも、こうした先行研究に倣い、データ提供元のデータ集計方法の変更等から生じるバイアスを調整した。

移をみると、特定サービス産業動態統計調査の「入場者数（遊園地・テーマパーク）」と概ね同様の動きとなっており、位置情報データで既存の公的統計を概ねトレースすることができる（図表2）。

（図表1）遊園地の曜日・時間帯別滞在人口
（図表2）遊園地の滞在人口の推移



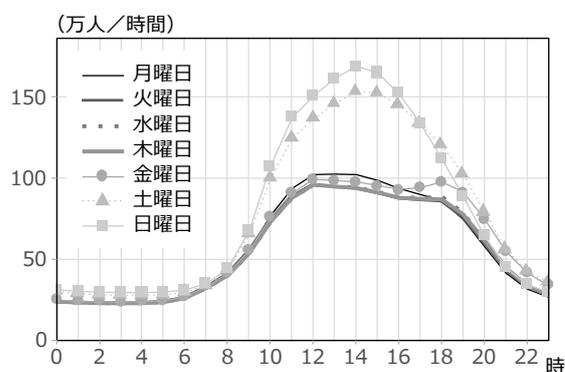
3. 2 ショッピングセンター

ショッピングセンター（SC）は、一定程度の敷地面積を有するため、他の経済活動の影響を受けるようなメッシュを除き、消費活動の中心となるメッシュを残すことが比較的容易であると考えられる。ここでは、①日本全国のショッピングセンターの住所に対応するメッシュを抽出（3,203メッシュ）したうえで、②駅隣接の施設を含むメッシュ、および土日の滞在人口の平均値が平日を下回るメッシュを除外した。駅隣接の施設を含むメッシュを除外するのは、駅が多様な経済活動の影響を受けるためであり、土日の滞在人口の平均値が平日を下回るメッシュを除外するのは、オフィスなどが入った複合施設の可能性を排除するためである。こうしたナウキャストニングの精度を低下させる要因を排除した結果、今回の分析で対象となるメッシュは、2,361メッシュまで絞り込まれる。

ショッピングセンターにおける曜日・時間帯別の滞在人口をみると（図表3）、土日の滞在人口は午後に大きく増加することや、金曜日は18時頃にピークをむかえることがわかり、一般的なショッピングセンターの人出を上手く捉えていると評価できる。実際に、10～20時の滞在人口（前年比）の推移をみると、業界統計であるショッピングセンター売上高と概ね似た動きとなっており、位置情報データで消費動向をナウキャストニングできることが分かる（図表4）。もっとも、仔細にみると、2019年10月の消費税率引き上げ前後の消費動向については、十分に捉えきれていない。これは、一人当たりの消費額が変化した場合に

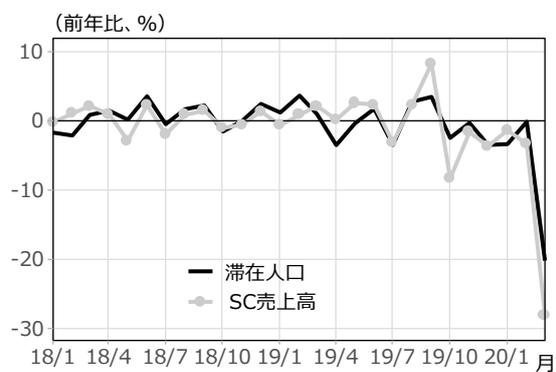
は、滞在人口だけでは消費動向を十分に捕捉できないことを示唆している。

(図表 3) ショッピングセンターの曜日・時間帯別滞在人口



(出所) Agoop、国土交通省、
日本ショッピングセンター協会

(図表 4) ショッピングセンターの滞在人口の推移



(出所) Agoop、国土交通省、
日本ショッピングセンター協会

3. 3 飲食業

飲食業では、敷地面積が小さい店舗の割合が高いと考えられるうえ、狭い範囲で複数の経済活動が行われている場所に立地することが多いため、メッシュの選別が難しい。ここでは、飲食業を代表するメッシュを特定する方法として以下の方法を採用した。具体的には、①東京都 23 区内から、水野他 (2020) に従い、「昼間人口 $>0.8 \times$ 夜間人口」という条件に適合する非住宅地のメッシュを抽出した (27,595 メッシュ) うえで、②中心から 300m 以内に 301 店舗以上の飲食店を含み⁶、駅を含まないメッシュを抽出した (758 メッシュ)。その後、③k-medoids 法⁷と呼ばれる統計的手法を用いて、5 グループにクラスタリングを行い、飲食業以外の活動の影響を強く受けていると想定されるメッシュを除外した。

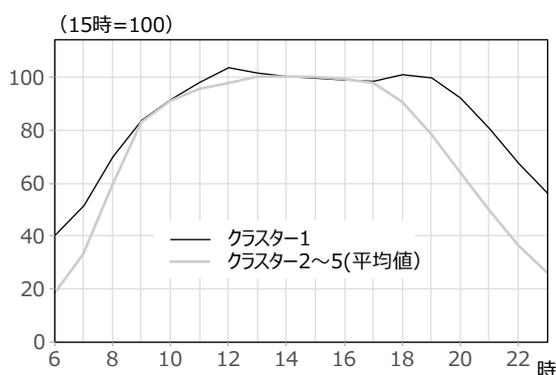
クラスタリング後の結果をみると (図表 5)、クラスター 1 (435 メッシュ) は平日の正午前後と 19 時頃にピークがあり、飲食関連の人出を捉えている可能性が示唆される一方、それ以外のクラスターでは 15 時頃にピークがきており、オフィス等の職場への人出を捉えていると考えられる。次に、クラスター 1 を JCB

⁶ 各メッシュの中心から 300m 以内に存在する店舗数データの取得については、ぐるなび API を用いた。

⁷ k-medoids 法とは、各クラスターを代表するデータ点 (medoid) をうまく選ぶことで、その medoid とクラスター内の各データ点との距離の総和を最小化するようにクラスタリングする方法である。

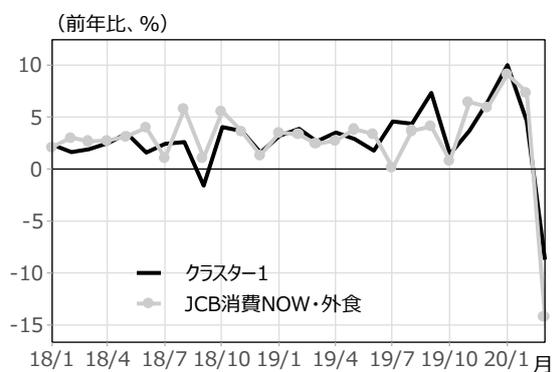
消費 NOW の「外食」の前年比と比較すると、両者は概ね似通った動きとなっている（図表 6）。相関係数をみても、JCB 消費 NOW の「外食」の前年比と高い相関にあり、クラスタリング前の指標よりも相関がより高くなっている（図表 7）。新型コロナウイルス感染症の影響で、営業時間の短縮などから大きく落ち込んだ 2020 年 3 月のデータを除いてみても、高い相関を示しており、今回開発した指標が、平時の飲食業の動向をうまく捉えていることが分かる。

（図表 5） 平日のクラスター・時間帯別滞在人口



（出所）Agoop、ぐるなび、国土交通省

（図表 6） クラスター1の滞在人口の推移



（出所）Agoop、JCB/ナウキャスト「JCB 消費 NOW」、ぐるなび、国土交通省

（図表 7） 飲食業指標との相関係数

	クラスタリング前		クラスター 1	
	～20/2月まで	～20/3月まで	～20/2月まで	～20/3月まで
外食	0.59	0.85	0.66	0.86
うち居酒屋	0.47	0.75	0.51	0.77

（出所）Agoop、JCB/ナウキャスト「JCB 消費 NOW」、ぐるなび、国土交通省

以上の 3 つの分析結果に対しては、「感染症の拡大以降、サービス関連の商業施設における滞在人口は大きく落ち込んでおり、そうしたデータを用いれば経済活動の落ち込みをナウキャストリングすることは容易である」との批判もありうる。もっとも、本稿で用いたデータは 2020 年 3 月までであり、経済が感染症の影響を強く受けた時期のデータはほとんど含まれていない。これは、滞在人口に基づく活動指標が、新型コロナウイルスの感染拡大により経済状況が急変しているときのみならず、平時の経済情勢下においても、タイムリーに経済活動

をナウキャストリングするツールとして機能することを示唆している。

4. 位置情報データと製造業

本節では、製造業の生産活動をナウキャストリングするための位置情報データの利用方法について説明する。企業の生産活動は一般的に、資本稼働率や技術進歩等、様々な要因によって決定され、その関係は業種ごとに区々であると考えられる。もっとも、この中でも労働投入は、概ねどの業種においても重要なファクターであると考えられる。

もし、①生産が労働投入の単純な関数形によって近似可能であり、さらに、②労働投入自体が滞在人口でうまく捉えられるとすれば、滞在人口データによる生産活動のナウキャストリングは景気判断において有用となることが期待される。

前節のサービス業と同様、製造業の場合は、工場における滞在人口が生産活動を捉える指標になるとの発想に基づき、指標を開発する。以下では、労働投入と生産活動とが線形関係となることを既存統計で確認したあと、位置情報データを用いて開発した活動指標について解説する。

4. 1 労働投入と生産活動の関係

ここでは、上記①の論点に基づき、滞在人口データと生産活動の関係を検証する前段階として、労働投入と生産の関係について確認する。毎月勤労統計を使って計算した労働投入（＝総実労働時間×雇用者数）と、鉱工業生産指数の前年比の相関係数をみると、業種によってばらつきがあるものの、生産用機械や輸送機械では高い相関を示している（図表 8）。こうした産業では、労働投入と生産の線形性が強く、滞在人口データを用いて労働投入の代理変数を作成できれば、これをナウキャストリングに用いることが可能と見込まれる。一方、電子部品・デバイス・電子回路や食料品などは、オートメーションが進んでいることもあって、この相関が殆どみられない。これらの産業では、生産の決定要因のうち労働の重要度は低く、滞在人口データを用いて生産活動をナウキャストリングすることは難しいと予想される。

(図表 8) 総労働投入と生産の相関係数 (前年比)

産業分類	相関係数
生産用機械器具	0.75
輸送用機械器具	0.72
金属製品	0.43
プラスチック製品	0.23
パルプ・紙・紙加工品	0.18
窯業・土石製品	0.15
電子部品・デバイス・電子回路	0.12
食料品、飲料・たばこ	0.05

(出所) 経済産業省、厚生労働省

(注) 期間は 14/1～19/12 月。総労働投入＝総実労働時間×常用雇用

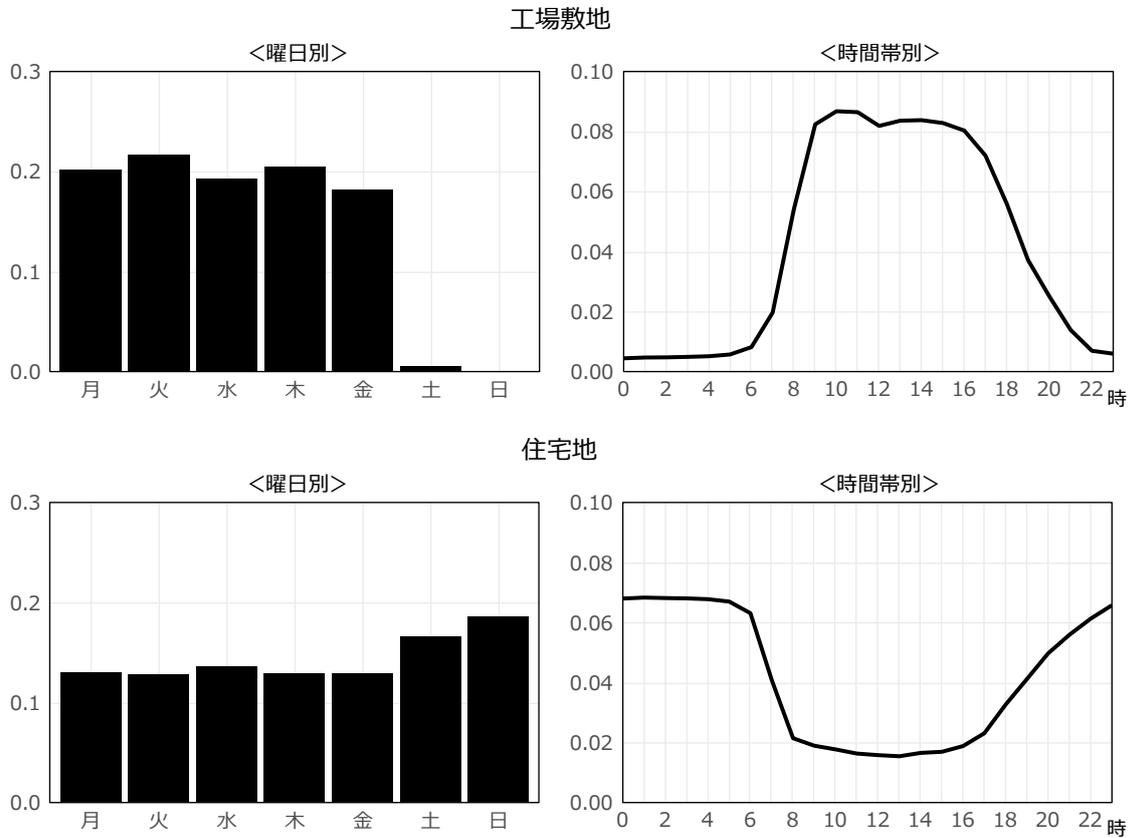
4. 2 工場地に対応するメッシュの特定

本稿では、経済センサス活動調査（経済産業省）の個票データを用いて、生産活動が行われているメッシュを特定した。より具体的には、経済センサス活動調査に含まれる工場の業種・住所・付加価値額等の情報をメッシュと紐づけ、それらのメッシュの滞在人口をカウントした。このとき、ノイズを除去するために、敷地面積が 1 万平方メートル以上の工場のうち、付加価値額の上位 1 万事業所（全産業ベース）を集計対象とした。

メッシュを特定する際の主な課題として、多くの工場の面積は、その記載された住所を含む単一メッシュ（100m 四方）より大きい。例えば、経済センサス活動調査からは、敷地の形状や具体的な境界についての情報は取得できないという点が挙げられる。しかしながら、高い精度で生産活動をナウキャストिंगするためには、記載住所に直接対応するメッシュ以外にも、工場を含むと想定されるメッシュを適切に選ぶ必要がある。こうした問題に対応するため、ヒューリスティックなアプローチとして、集計対象となるメッシュの候補を住所の周囲から広めに集め、その後に生産活動と関係の薄いメッシュを取り除くという方法を採用した。具体的には、①日曜比率、②昼間比率、③平均滞在人口、という 3 つの基準に照らして、メッシュの絞り込みを行った。まず、①日曜比率、および、②昼間比率に着目した理由としては、住宅地と工場地の間で両比率が大きく異なっており、工場地と住宅地を識別する際の有益な変数であると考えたためである。実際に、平均的な工場を含むメッシュの滞在人口のパターンをみると、日曜の比率が低く、昼間の比率が高い（図表 9 上段）。一方、典型的な住宅地を含むメッシュをみると、日曜の比率が高く、昼間の比率が低いことがわかる（図表 9 下段）。更に、③の平均滞在人口を用いることで、少数サンプルの問題から生じ

るノイズの影響を可能な限り小さくすることを試みた。

(図表 9) 工場敷地と住宅地における滞在人口分布 (曜日別・時間帯別)



(出所) Agoop

(注) 前後週および当週の休日が2日であるような週を対象に比率を計算。時間帯別の分布は平日のみを使用。平均的な工場敷地と住宅地を含むメッシュをそれぞれ一つ(100m 四方)選び、滞在人口の分布を示している。

本稿では、こうした比率が業種によって異なり得ることを勘案し、労働投入との相関が高くなるような組み合わせを業種ごとに決定する。例えば、輸送用機械では、候補メッシュの中で、日曜比率が0.5以下、昼間比率が0.6以上、平均滞在人口が10人以上であるようなメッシュの滞在人口を集計すると、労働投入との相関が0.86と最も高くなることから、このようなメッシュを工場として特定した(図表10)。このように特定したメッシュの滞在人口と労働投入との相関係数は、業種毎に大きく異なっているものの、製造業全体で見れば、滞在人口は労働投入の動きを捉えられることを示唆している(図表11)。

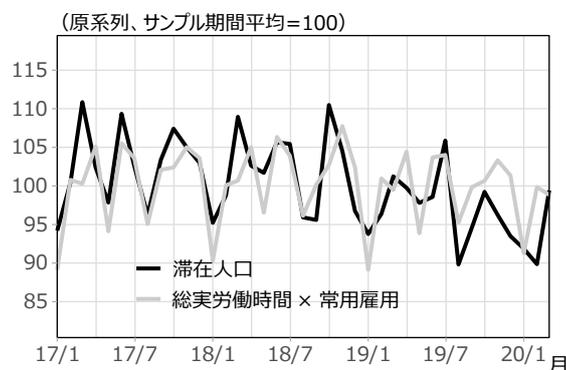
(図表 10) 日曜比率・昼間比率・滞在人口に関する絞り込み条件

業種	日曜比率 上限	昼間比率 下限	滞在人口 下限	相関係数
輸送用機械器具	0.5	0.6	10	0.86
生産用機械器具	0.1	2	40	0.69
情報通信機械器具	0.9	1	80	0.68
はん用機械器具	0.5	1.8	40	0.63
食料品、飲料・たばこ	0.1	1	60	0.60
業務用機械器具	0.5	2	10	0.54
非鉄金属	0.1	1.4	20	0.49
電気機械器具	0.1	0.8	10	0.42
電子部品・デバイス・電子回路	0.9	1.4	10	0.31
化学、石油・石炭	0.9	0.6	10	0.20

(出所) Agoop、厚生労働省

(注) 労働投入量＝総実労働時間×常用雇用。前後週および当週の休日が2日であるような週を対象に比率を計算。日曜比率＝(日曜日の平均人口) / (平日の平均人口)、昼間比率＝(9～16 時台の平均人口) / (0～4 時台の平均人口)。

(図表 11) 滞在人口データと企業の労働投入 (製造業)



(出所) Agoop、厚生労働省

4. 3 滞在人口データから計測した活動指標

次に、特定したメッシュの滞在人口から、生産活動の代理指標を作成する方法について説明する。工場の滞在人口を集計する際に最も重要な点は、設備のメンテナンス等の生産活動と直接結びつかない活動が行われた場合でも、滞在人口に変化が生じるため、全ての曜日・時間帯のデータを使うことが必ずしも精度の高いナウキャストに繋がるわけではない、ということである。この点を踏まえて、本稿では、集計の際に、①休日を含めるか否か、②日中のどの時間帯を利用するか、という2点について、鉱工業生産指数の前年比との相関が、最も

高くなるような組み合わせを選択した。結果をみると、業種によって区々の動きとなっているものの、夕方の時間帯における滞在人口データが、生産活動を捉える目的に適していることが分かる（図表 12）。

（図表 12） メッシュの絞り込み条件

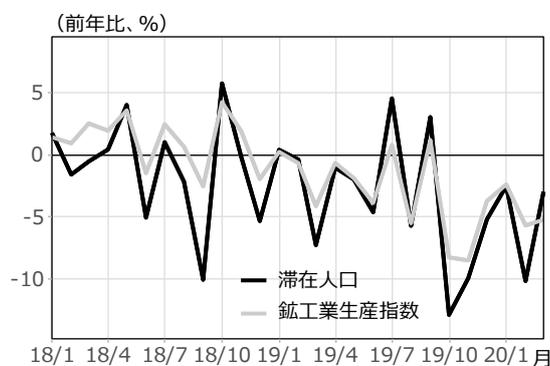
業種	絞り込み条件			相関係数
	休日含む	始点	終点	
輸送機械工業	○	17	18	0.85
生産用機械工業	○	18	21	0.82
鉄鋼・非鉄金属工業	○	8	12	0.78
電気・情報通信機械工業	×	16	17	0.72
金属製品工業	×	16	19	0.71
プラスチック製品工業	○	17	18	0.69
汎用・業務用機械工業	○	10	11	0.65
パルプ・紙・紙加工品工業	×	9	10	0.61
窯業・土石製品工業	○	14	15	0.55
電子部品・デバイス工業	○	8	9	0.36
化学工業	×	16	17	0.35
食料品・たばこ工業	×	22	23	0.30

（出所）Agoop、経済産業省

（注）例えば、始点=18、終点=21 は 18:00～21:59 を指す。

工場として特定されたメッシュにおいて、上記のように選ばれた時間帯における滞在人口を集計した指数をみると、鉱工業生産指数の前年比をうまく再現できている（図表 13）。特に、台風被害の影響が大きかった 2019 年 10 月については、滞在人口から計測した活動指標が大きく落ち込んでおり、鉱工業生産が落ち込んでいることを精緻にナウキャストリングすることができていることが分かる。

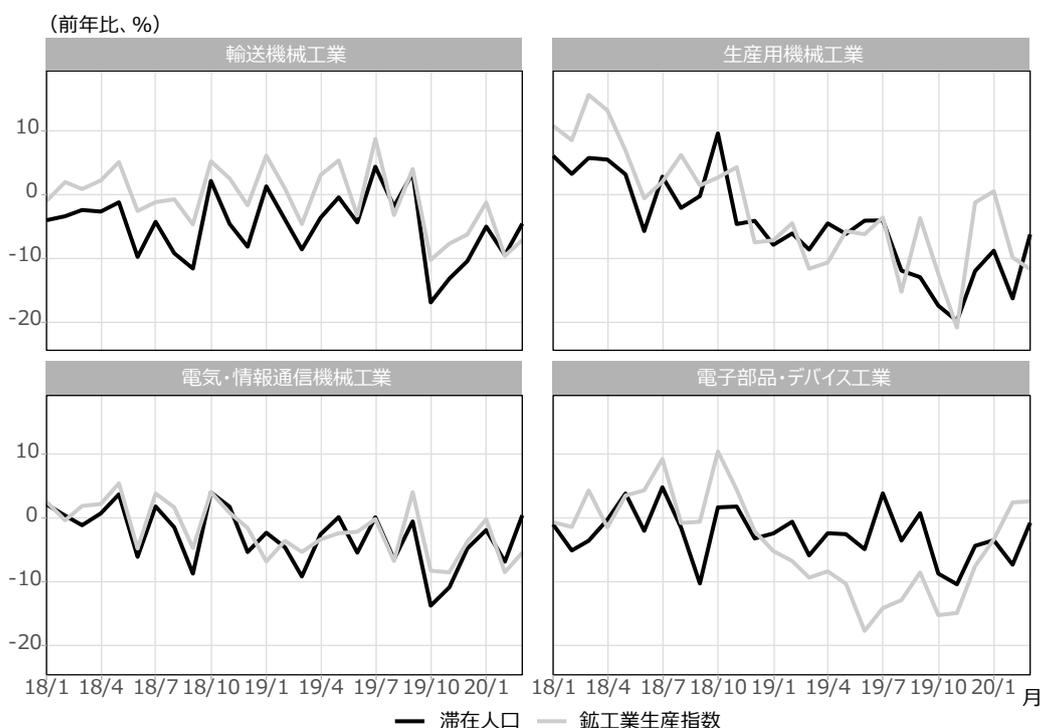
（図表 13） 滞在人口データと生産活動



（出所）Agoop、経済産業省

業種別に活動指標と鉱工業生産の動きをみると、輸送機械や生産用機械、電機・情報通信機械などの業種では、ナウキャストイングのパフォーマンスが良い（図表 14）。一方、電子部品・デバイスについては、滞在人口データによる活動指標が 2019 年中の生産の落ち込みを捉えきれていないなど、ナウキャストイングの精度は低いと評価できる。これは、輸送機械や生産用機械などは、労働集約的な業種であるため、労働投入が生産活動を十分説明でき、そのうえで滞在人口データが労働投入をよく再現できているためである。一方、電子部品・デバイスはそもそも資本集約的であるうえ、滞在人口データが労働投入をうまく捉えられていないことが、精度が低い要因であると考えられる。

(図表 14) 滞在人口データと生産活動（業種別）



(出所) Agoop、経済産業省

4. 4 分析の拡張：データの高頻度化

4.3 節の一つの応用として、作成した月次の活動指標を、より高頻度化することが考えられる。集計に用いている位置情報データは、1 時間毎の高頻度データであることから、集計単位を週次や日次にまで短期化すれば、経済が急変する状況をよりタイムリーに捕捉することが可能となる。

ここでは、輸送機械工業について、滞在人口を週毎に集計したうえで、統計的手法により季節要因を取り除いた⁸（図表 15）。結果をみると、2019 年 10 月 20 日週に大きく落ち込んでおり、2019 年の台風 19 号が生産活動に与えた影響をタイムリーに捉えられることを示唆している。このように、位置情報データは、豪雨や台風といった自然災害時においても、景気のごく短期的な変動を分析する手段として、有用であることを示している。

（図表 15） 週次生産指標（輸送機械工業）



（出所）Agoop、経済産業省

5. まとめ

本稿では、高頻度のビッグデータのひとつである携帯電話の位置情報データを用いて、各地点の滞在人口から、サービス業の売上動向や製造業の生産活動をナウキャストする手法を提示した。

まず、サービス業については、各地点の来店客数は滞在人口で概ね近似可能である点に着目し、遊園地、ショッピングセンター、飲食業に関する活動指標を作成した。滞在人口にノイズが多く含まれる場合でも、クラスタリングといった統計的手法を適切に組み合わせることで、既存統計では迅速な把握が難しいサービス業の活動についても、高い精度でナウキャストできることを示した。次に、製造業については、経済センサス活動調査の事業所データや昼間人口比率などの時間帯別の滞在人口分布を利用して、比較的大規模な工場敷地を特定したうえで、その地点の滞在人口をベースに生産活動をナウキャストする指標を作成した。業種別にみると、輸送機械や生産用機械といった労働集約的な

⁸ フェイスブックが開発した「Prophet」と呼ばれる時系列予測のためのオープンソースソフトウェアライブラリを用いた。

業種の生産については、工場の滞在人口によりかなり高い精度で予測できることを示した。

本稿の分析結果は、位置情報データが、マクロ的な経済活動をナウキャストリングするツールとして有効である可能性を示唆している。

本稿で用いた位置情報データは、100m四方の「場所」を対象としたパネルデータであって、消費者・労働者の年齢や性別といった属性情報は含まない。このため、高齢者の消費行動など、属性別の行動様式を分析することはできない。また、本データはスマートフォン向けアプリケーションを通じて収集したデータであるため、スマートフォンユーザー以外の行動様式が十分に反映されていない可能性がある点にも留意が必要である。加えて、入手可能なデータが数年分であり、分析の妥当性の検証には更なるデータの蓄積が必要となる。本分析では、経済センサスなどと組み合わせたように、分析のテーマに適した他のデータを見つけてくることも重要となる。例えば、工場の移設等により、それまで集計対象としていたメッシュにおける滞在人口が大きく変動した場合には、滞在人口データ単体からは、その変動要因を見出すことはできなくなるため、こうした変化を捉えられる他のデータを見つけてくることが課題となる。また、高度にオートメーション化された生産活動を分析対象とする場合、滞在人口に加え、資本の稼働率や生産性の変動も重要な要因となる。こうした要因は、今回開発した指標の精度に悪影響を及ぼすが、この問題を克服していくことは、今後の研究課題としたい。

参考文献

水野貴之, 大西立頭, 渡辺努. (2020). "流動人口ビッグデータによる外出の自粛率の見える化," 人工知能, 35(5), 667-672.

Aladangady, A., Aron-Dine, S., Dunn, W., Feiveson, L., Lengermann, P., and Sahm, C. (2019). "From Transactions Data to Economic Statistics: Constructing Real-time, High-frequency, Geographic Measures of Consumer Spending," NBER Working Paper 26253.

Aprigliano, V., Ardizzi, G., and Monteforte, L. (2019). "Using Payment System Data to Forecast Economic Activity," International Journal of Central Banking, 15(4), 55-80.

Arslanalp, S., Marini, M., and Tumbarello, P. (2019). "Big Data on Vessel Traffic:

- Nowcasting Trade Flows in Real Time," International Monetary Fund Working Paper WP/19/275.
- Cajner, T., Crane, L. D., Decker, R. A., Hamins-Puertolas, A., and Kurz, C. (2019). "Improving the Accuracy of Economic Measurement with Multiple Data Sources: The Case of Payroll Employment Data," Finance and Economics Discussion Series 2019-065, Board of Governors of the Federal Reserve System.
- Cerdeiro, D. A., Komaromi, A., Liu, Y., and Saeed, M. (2020). "World Seaborne Trade in Real Time: A Proof of Concept for Building AIS-based Nowcasts from Scratch, " International Monetary Fund Working Paper WP/20/57.
- Chen, M. K., and Pope, D. G. (2020). "Geographic Mobility in America: Evidence from Cell Phone Data," NBER Working Paper 27072.
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., and the Opportunity Insights Team (2020). "How Did Covid-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data," NBER Working Paper 27431.
- Couture, V., Dingel, J. I., Green, A., Handbury, J., and Williams, K. R. (2021). "Measuring Movement and Social Contact with Smartphone Data: A Real-Time Application to Covid-19," Journal of Urban Economics, 103328.
- Coven, J., and Gupta, A. (2020). "Disparities in Mobility Responses to Covid-19," NYU Stern Working Paper.
- Dong, L., Chen, S., Cheng, Y., Wu, Z., Li, C., and Wu, H. (2017). "Measuring Economic Activity in China with Mobile Big Data," EPJ Data Science, 6, 1-17.
- Fang, H., Wang, L., and Yang, Y. (2020). "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China," Journal of Public Economics, 191, 104272.
- Galbraith, J. W., and Tkacz, G. (2018). "Nowcasting with Payments System Data," International Journal of Forecasting, 34(2), 366-376.
- Watanabe, T., and Yabu, T. (2020) "Japan's Voluntary Lockdown," CARF Working Paper.