

新型コロナウイルス感染症の時空間集積性と それに基づく実効再生産数推定精度の向上

—東京都を例にして—

石岡 文生¹・椿 広計²・久保田 貴文³・鈴木 和幸⁴

(受付 2021 年 6 月 30 日；改訂 2022 年 2 月 21 日；採択 2 月 22 日)

要 旨

2019 年 12 月に中国の武漢で最初に報告された新型コロナウイルス感染症(通称: COVID-19)の脅威に対し、学術的な見地に基づいた様々な調査研究が、今まさに世界中で盛んに進められている。現在、様々な国では、時空間集積性を評価する試みが精力的に行われているが、日本国内での長期に亘る日々のホットスポットの動向についての検証は、いまだ為されていないのが実情である。本研究では、東京都が公表している 2020 年 3 月 31 日から 2021 年 5 月 31 日までの全 427 日間における市区町村別の累積陽性者数を用い、prospective-cylindrical scan 法に基づいた空間スキャン検定を毎日行うことで「どの地域」に「いつから」ホットスポットが存在していたのかを明らかにする。さらに、ホットスポットの発現とその要因との関連性、ならびに地域性や人流とホットスポットとの関連性について考察する。解析に当たっては、市区町村別居住人口に対し、地域の従業員数と人流を加味したものとなっている。加えて本稿では、ホットスポット情報を用い東京都の市区町村別実効再生産数を推定するとともに、これにより、この推定方法の精度がより高いことを示す。なお、これは人口が一定規模以下の自治体の実効再生産数の推定においてもある程度実用可能である。

キーワード: COVID-19, 時空間集積性, 空間スキャン検定, 実効再生産数。

1. はじめに

2019 年 12 月に中国の武漢で最初に報告された新型コロナウイルス感染症(通称: COVID-19)は、全世界で約 1 億 7500 万人が陽性と判定され、その死者は約 380 万人に達するなど(いずれも本稿執筆時の 2021 年 6 月 15 日時点での状況)、今なお世界で猛威を振るっている(World Health Organization, 2021)。国内に目を向けると、2020 年 1 月に最初の COVID-19 陽性者が確認されて以降、感染終息の目途がつかない状態が続いており、2021 年 4 月の中頃には、全国の 1 日の新規陽性者数が初めて 5,000 人を超えたという報道がなされている。本稿の執筆時、国内のワクチン接種がようやく起動に乗り始めたという明るい話題がある一方で、従来のもの

¹ 岡山大学 学術研究院環境生命科学学域: 〒700-8530 岡山県岡山市北区津島中 3-1-1

² 統計数理研究所: 〒190-8562 東京都立川市緑町 10-3

³ 多摩大学 経営情報学部: 〒206-0022 東京都多摩市聖ヶ丘 4-1-1

⁴ 電気通信大学大学院 情報理工学研究所: 〒182-8585 東京都調布市調布ヶ丘 1-5-1

より感染力が強いとされる変異株の流行が懸念されており、さらには開催を来月に控えた東京五輪とも相まって、依然先行き不透明な状況にある。

この COVID-19 の脅威に対し、学術的な見地に基づいた様々な調査研究が、今まさに世界中で盛んに進められている。これには、感染症疫学はもちろんのこと、社会学的・数学的な枠組みでの分析も活発に行われており、例えば因果分析、人口変動分析、感染症数理モデル、時空間解析等といった取り組みがなされている。そうした中で、本研究は東京都における COVID-19 の空間集積性について焦点を当てる。空間集積性とは、平面上の二次元空間データにおいて局所的に高い(または低い)観測値を示す場所を統計的根拠に基づき評価する科学で、疫学、犯罪学、環境学などの様々な分野で活用されている。高い観測値からなる集積地域を、特にホットスポット (hotspot cluster) などと呼ぶこともある。

メディアなどでよく目にする累積陽性者の色分け地図は、観測値の絶対量を単純にいくつかの段階に分けて視覚化しているに過ぎず、その地図上に塗られた色の濃淡を見て、ある領域群をホットスポットと断言するのは説得力に欠ける。また、地図から受ける印象は、色分けの階級区分の決め方次第で変化するだろうし、もしかしたら偶然変動の範囲でたまたま高い、または低い値が観測されただけかもしれない。従って、「どこかにホットスポットはあるのか?」「ホットスポットが存在しているとしたら、どの範囲までがそうなのか?」の判断を統計的根拠(データ)に基づいて客観的に行うことが重要となる。

Kulldorff (1997) が提唱した空間スキャン検定は、空間集積性の存在の有無について検定すると同時に、その場所をも同定する手法である。これは、フリーソフトウェア SaTScan™ (Kulldorff et al., 2021) によって実践可能で、さらに最近では関連する R のライブラリーも複数公開されるなど (Chen et al., 2018; French, 2020; Ishioka, 2020; Otani and Takahashi, 2020)、近年の主流な空間集積性評価法として広く利用されている。空間スキャン検定のアイデアを COVID-19 関連のデータに適用し、その集積性について評価する試みは世界で精力的に行われており、米国 (Amin et al., 2020; Hohl et al., 2020)、ニューヨーク (Cordes and Castro, 2020; Greene et al., 2020)、中国 (Chow et al., 2020)、韓国 (Kim and Castro, 2020)、バングラディッシュ (Masrur et al., 2020)、メキシコ (Benita and Gasca-Sanchez, 2020)、クウェート (Alkhamis et al., 2020)、ネパール (Acharya et al., 2020) 等といった国や地域を対象としたものが報告されている。日本国内を対象にしたものでは、Yoneoka et al. (2020) などの報告があるものの、長期にわたる日々のホットスポットの動向の検証についてはまだ行われていないのが実情である。

解析に当たり、我々は東京都が公表している 2020 年 3 月 31 日から 2021 年 5 月 31 日までの全 427 日間における東京都市区町村別の COVID-19 累積陽性者数を収集した。加えて、実際にその日・その地域で活動していた人の数を反映させることを目的に、市区町村別に得られる居住人口に対し、その従業者数や日々の人流で調整を施したものを解析に取り入れる。そしてこれら 1 年以上に渡って集めた日々のデータに対し、第一の解析として、prospective-cylindrical scan 法に基づいた空間スキャン検定を毎日行うことで「どの地域」に「いつから」ホットスポットが存在していたのかを明らかにし、そこからホットスポットの発現とその要因との関連性、ならびに地域性や人流とホットスポットとの関連性について考察する。さらに第二の解析として、東京都のホットスポット地域における実効再生産数についても着目する。1 人の感染者が平均で何人に感染させるかを示す指標として広く利用されている実効再生産数は、これまで国内の様々な行政単位に対しその推定が行われているが、我々は新たな試みとして検出されたホットスポット地域でクラスタリングを行い、それによって実効再生産数の推定精度が改善されるかを検証する。

本論文の構成は以下の通りである。第 2 節において本研究で用いるデータについて説明する。第 3 節では空間スキャン検定の概要を述べ、日々の COVID-19 陽性者ホットスポットの検

出を試みる。第4節では本研究における実効再生産数の推定法について述べるとともに、ホットスポット情報に基づいた実効再生産数の推定を試みる。第5節で種々の解析結果に対し時系列に沿って考察を述べ、第6節でまとめと今後の課題について触れる。

2. 使用データについて

2.1 東京都 COVID-19 陽性者データ

東京都は、2020年4月1日より、日々、前日までの市区町村別の累積陽性者数の公表を開始した(東京都, 2020; 東京都福祉保健局, 2021)。本研究では、2020年3月31日から2021年5月31日までの全427日間にわたり、東京都福祉保健局報道発表による東京都の62の市区町村別に集計された累積陽性者数のデータを収集した。なお、PCR検査等で発見されていない「潜在的な感染者を含む感染者数」と区分する際に、本研究では「感染者数」ではなく「陽性者数」にて記述する。本データで注意すべき点は、下記の通りである；

- 陽性者は感染した地域ではなく、患者の居住地にてカウントされる。
- 2020年5月10日までの「調査中」の患者552名の内の428名が、5月11日から14日の4日間に該当する各市区町村へ割り振られ、陽性者数が一気に増加している。
- 2021年2月14日に2020年11月18日～2021年1月31日における葛飾区、豊島区、渋谷区での報告漏れによる情報の滞留分が公表されている。特に豊島区では陽性数170名と一気に増加し、これまでの累積者数が一気に2,830名に増加している。

また、東京都の緊急事態宣言・まん延防止等重点措置期間とその対象地域は下記の通りである；

2020年4月7日～2020年5月31日 緊急事態宣言

2021年1月8日～2021年2月7日 緊急事態宣言

2021年2月8日～2021年3月7日 緊急事態宣言(延長)

2021年3月8日～2021年3月21日 緊急事態宣言(延長)

2021年4月25日～2021年5月11日 まん延防止等重点措置期間(対象地域：23区、八王子市、立川市、武蔵野市、府中市、調布市、町田市)

2021年5月12日～2021年5月31日 緊急事態宣言(延長)

2021年6月1日～2021年6月20日 緊急事態宣言(延長)

2021年6月21日～2021年7月11日 まん延防止等重点措置期間(対象地域：23区及び檜原村、奥多摩町を除く多摩地域の市町)

2.2 データの前処理

本研究では、地域間距離等の問題から空間的な扱いや解釈が困難な島しょ部(9町村)を除外し、この結果、東京都区部と多摩地域を合わせた全53市区町村を解析対象地域と定めた。なお、島しょ部の累積陽性者数は全体のわずか0.03%に過ぎず、大勢に影響はないと考える。期間全体の累積陽性者数は147,433人となっており、最も多かった地域は世田谷区の11,629人、次いで新宿区の8,733人であった。また、最も少なかったのは檜原村の10人であった。

公表されているデータは累積の陽性者数であるため、本研究では次のように日々の新規陽性者数に変換する。いま、地域*i*における観測日*t*の累積陽性者数を $c_{i,t}$ とする。このとき、新規陽性者数(これ以降、新規陽性者数を「陽性者数」にて表記する) $o_{i,t}$ を、

$$(2.1) \quad o_{i,t} = c_{i,t} - c_{i,t-1}, \quad i = 1, 2, \dots, 53; \quad t = 1, 2, \dots, 426$$

で与える。ここに、($t = 0$): 2020年3月31日、($t = 1$): 2020年4月1日、($t = 2$): 2020年4

表 1. 陽性者数 $o_{i,t}$ が負の値となった地域と観測日.

地域	観測日	陽性者数	地域	観測日	陽性者数
羽村	2020/5/3	-1	小金井	2020/11/1	-1
中野	2020/5/16	-3	武蔵野	2020/11/8	-1
港	2020/6/21	-17	西東京	2020/11/9	-1
小平	2020/6/22	-1	東村山	2020/12/21	-1
あきる野	2020/10/14	-1	あきる野	2021/5/10	-1
あきる野	2020/10/18	-1			

月 2 日, ..., ($t = 426$): 2021 年 5 月 31 日である. なお, 累積陽性者数 $c_{i,t}$ は, 後日に報告修正が行われることがあり, そのためごくまれに $o_{i,t} < 0$ となる場合が存在する. 実際, 今回収集したデータで $o_{i,t}$ を算出したところ, 表 1 に示した 11 の地域と観測日で陽性者数が負の値になる ($o_{i,t} < 0$) ことが確認されたが, これらについては 0 とすることで対処した.

図 1 は, 2020 年 4 月 1 日から 2021 年 5 月 31 日までの市区町村別の陽性者数 $o_{i,t}$ を色の濃淡によって視覚的に表現しており, 色が濃いほどその日の陽性者数が多いことを意味する. 併せて, 全体の日次陽性者数の推移を棒グラフ, 市区町村別の日次陽性者数の分布を箱ひげ図で示した. 図より, 全体を通して多摩地域よりも東京都区部の方が色が濃いことが見て取れる. 特に 2021 年 1 月はそれが際立っており, 対応する棒グラフからその時期には連日 1,000 人を超える陽性者数が観測されるなど, 感染が爆発的に拡大していたことが伺える. 地域毎の陽性者数に目を向けると, 東京都区部以外では, 八王子市, 府中市, 町田市で比較的多く観測されている.

2.3 調整人口

COVID-19 のように人から人へと伝播する感染症を扱う場合に, 単にその地域の居住人口のみを用いて調整するのではなく, その地域で実際に活動している人, すなわち「従業者数」も考慮に入れた調整が適切と考えられる. また, 特に人の往来が激しい東京都区部においては, そこに人の動き(人流)も深く関わってくると思われる. そこで本研究では, ①居住人口として東京都公開のもの(東京都総務局統計部, 2021), ②従業者数として経済センサス-基礎調査(総務省統計局)にて公開のもの(統計データ分析コンペティション, 2021), ③人流データとして, それぞれ新型コロナウイルス感染症対策サイト(内閣官房)から著者の鈴木が収集・整理したもの(～2020 年 5 月 20 日)と NTT ドコモのモバイル空間統計(2021)にて公開のもの(2020 年 5 月 21 日～), に基づいて, 新たに地域毎に日次の「調整人口」を定義する. ①～③の各データの詳細を表 2 に示す. また, 地域毎の①居住人口と②従業者数を図 2 に, 日毎の③人流の推移を図 3(点線)にそれぞれ示す.

これらを用いて, 調整人口 $n_{i,t}$ を次のように算出した.

$$(2.2) \quad \text{東京都区部: } n_{i,t} = \frac{\mathbf{N}^{(r)} + \mathbf{n}^{(b)} \cdot \mathbf{n}^{(m)\top}}{2} \quad i = 1, 2, \dots, 23; \quad t = 1, 2, \dots, 426$$

$$(2.3) \quad \text{多摩地域: } n_{i,t} = \frac{\mathbf{N}^{(r)} + \mathbf{n}^{(b)} \cdot \mathbf{1}^{(1 \times 426)}}{2} \quad i = 24, 25, \dots, 53; \quad t = 1, 2, \dots, 426$$

ここに, $\mathbf{1}^{(1 \times 426)}$ は全ての要素が 1 で構成された 1×426 の行ベクトルである.

図 2 は, 地域によって居住人口と従業者数に大きな差があることを示しており, 地域によっては, 実際にそこで活動している人の数を居住人口のみで表現することは必ずしも妥当でないと思われる. また, このグラフはその地域が居住地域か商業地域かといった地域の特性, ならびにその規模を測るおおよその目安となる. さらに, 図 3(点線)からは, 緊急事態宣言の有無

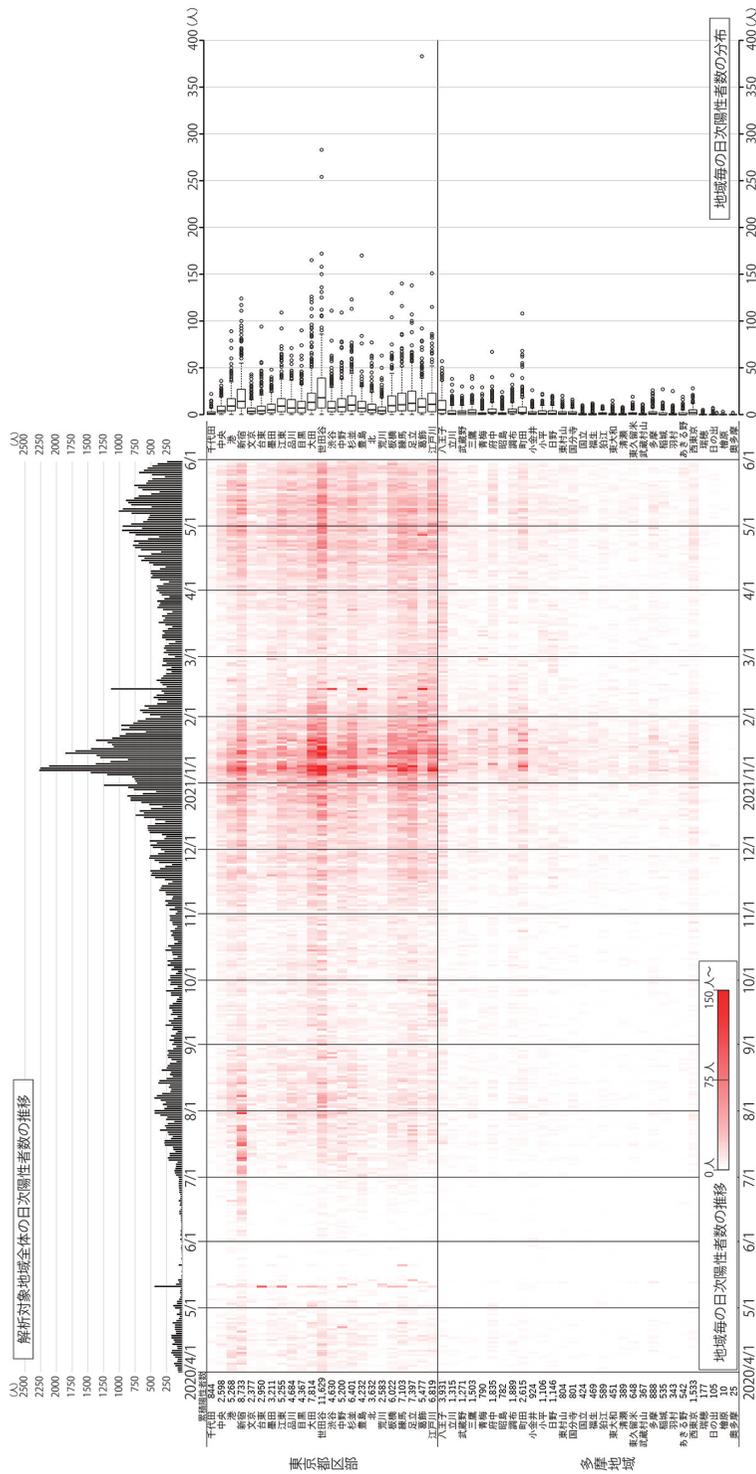


図 1. 陽性者数の推移.

表 2. 本研究で用いる人口情報.

取得形態	表記	出典
①居住人口	地域毎・日毎 $\mathbf{N}^{(r)} = [n_1^{(r)}, n_2^{(r)}, \dots, n_{426}^{(r)}]$	平成 27 年国勢調査人口 (確報値) に基づく人口 (推計) $\mathbf{n}_t^{(r)} = [n_1^{(r)}, n_2^{(r)}, \dots, n_{53}^{(r)}]$ は毎月更新
②従業者数	地域毎 $\mathbf{n}^{(b)} = [n_1^{(b)}, n_2^{(b)}, \dots, n_{53}^{(b)}]^T$	SSDSE 2020 年版 (2020 年 6 月 26 日公開)
③人流データ	日毎 $\mathbf{n}^{(m)} = [n_1^{(m)}, n_2^{(m)}, \dots, n_{426}^{(m)}]^T$	JR 山手線主要 5 駅と東京メトロ全駅における対前年比の 改札通過人数を基に算出した平均値 (~2020 年 5 月 20 日) 東京都区部 9 カ所 15 時時点における前年同月比の 増減率の平均値 (2020 年 5 月 21 日~)

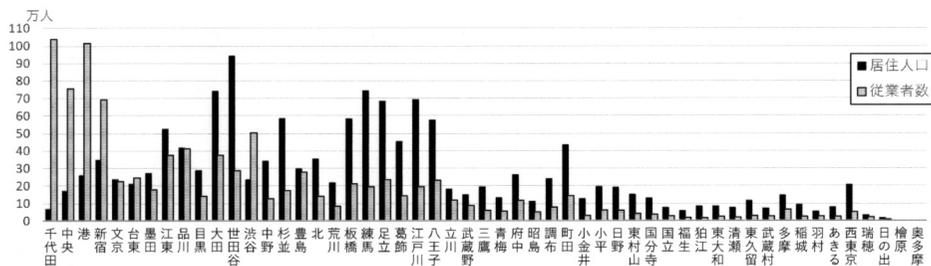


図 2. 居住人口と従業者数. なお, 居住人口数は毎月更新しているため, この図中の居住人口は地域毎に平均したものになっている.

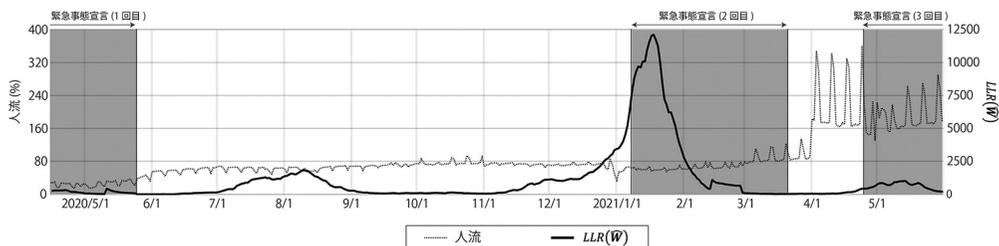


図 3. 人流と $LLR(\hat{W})$ の推移.

や年末年始, 年度変わり等の節目節目における人流の変化が見て取れる. 陽性者数の大小に加えて, 「居住人口」と「日々の人流の影響で重み付けした従業者数」を平均した「調整人口」をも組み入れて検出されるホットスポットは, 実際の人の活動や地域特性がより色濃く反映されたものになると考える.

3. 東京都における COVID-19 の時空間ホットスポットの検出

3.1 空間スキャン検定

空間スキャン検定 (Kulldorff, 1997) は, 空間集積性の有無を尤度に基づいて評価する手法で, 近年様々な分野において広く利用されている. 高い観測値からなる集積地域を特にホットスポット (hotspot cluster) などと呼ぶこともある. 本研究で扱う伝染性の感染症を対象とする場合, ホットスポットが「どこに」発生したのかに加え, それが「いつから」, または「いつまで」発生していたのかといった時間的な検証も重要となる. Kulldorff et al. (1998) は, このような時

空間集積性 (space-time cluster) を検出するための空間スキャン検定として cylindrical scan 法を提唱している。本節では、まず初めに空間スキャン検定ならびに cylindrical scan 法の概要について述べる。

いま、地域 $i (i = 1, 2, \dots, 53)$ における観測日 $t (t = 1, 2, \dots, 426)$ の陽性者数の確率変数 $O_{i,t}$ は、互いに独立に Poisson 分布に従うと仮定する。

$$O_{i,t} \sim \text{Poisson}(\mu_{i,t})$$

ここに、 $\mu_{i,t}$ は観測日 t における地域 i の期待陽性者数である。次に、「ある 1 つ以上の {地域, 観測日} = $\{i, t\}$ が連結してできる部分集合」をウィンドウとよび \mathbf{W} で表す。 $O(\mathbf{W})$ を \mathbf{W} 内の陽性者数の確率変数、 $\mu(\mathbf{W})$ を \mathbf{W} 内の期待陽性者数とすると、 \mathbf{W} がホットスポットとなるか否かは、

$$H_0 : E[O(\mathbf{W})] = \mu(\mathbf{W}) \quad \text{for } \forall \mathbf{W} \in \mathcal{W}$$

$$H_1 : E[O(\mathbf{W})] > \mu(\mathbf{W}) \quad \text{for } \exists \mathbf{W} \in \mathcal{W}$$

の仮説検定問題となる。ここに、 \mathcal{W} は \mathbf{W} の取り得る全体集合である。また、 \mathbf{W} の補集合を \mathbf{W}^c とするとき、 $O(\mathbf{W} \cup \mathbf{W}^c) = \mu(\mathbf{W} \cup \mathbf{W}^c)$ であるとする。

検定統計量はそれぞれ H_0 と H_1 の状況下の尤度関数の比で定義され、

$$(3.1) \quad LR(\mathbf{W}) = \begin{cases} \left(\frac{o(\mathbf{W})}{\mu(\mathbf{W})} \right)^{o(\mathbf{W})} \left(\frac{o(\mathbf{W}^c)}{\mu(\mathbf{W}^c)} \right)^{o(\mathbf{W}^c)}, & o(\mathbf{W}) > \mu(\mathbf{W}) \\ 1, & \text{その他} \end{cases}$$

で与えられる。ここに、 $o(\mathbf{W})$ は \mathbf{W} 内の陽性者数の実現値である。なお、計算を簡単にするために、実際には (3.1) 式を対数変換した $LLR(\mathbf{W}) = \ln LR(\mathbf{W})$ が用いられる。つづいて、ウィンドウ \mathbf{W} の全体集合 $\mathcal{W} = \{\mathbf{W}_1, \mathbf{W}_2, \dots\}$ の中から最大の対数尤度比をもつウィンドウ、すなわち、

$$(3.2) \quad \hat{\mathbf{W}} = \arg \max_{\mathbf{W} \in \mathcal{W}} LLR(\mathbf{W})$$

なるウィンドウ $\hat{\mathbf{W}}$ をホットスポットの候補と考える。 $\hat{\mathbf{W}}$ が統計的に有意なホットスポットであるかどうかを評価するためには、 H_0 の下での $\max_{\mathbf{W} \in \mathcal{W}} LLR(\mathbf{W})$ の分布が必要だが、これを一意に定めることは解析的に困難なため、モンテカルロ法によるシミュレーションで求めた p 値によって有意性を評価するのが通例となっている。

さてここで、 \mathcal{W} をどのように求める (これを “スキャンする” と言う) かが問題になる。Kulldorff et al. (1998) の cylindrical scan 法では、データを空間的な XY の二次元上に時間軸 T を加えた三次元空間と捉え、平面上を同心円状のウィンドウによってスキャンするとともに、 $T_1 \leq s_1 \leq s_2 \leq T_2$ なる時点 s_1 から時点 s_2 までを「高さ」と対応させた円柱状のウィンドウ \mathbf{W} を採用している。さらに、Kulldorff (2001) は、常に最近の時点が含まれるよう $T_2 = s_2$ に固定して cylindrical scan を行うことで、解析実行時点における現在進行中のホットスポット (alive cluster) の有無について評価する prospective-cylindrical scan 法 (前向き法) を提案している。この前向き法は、各種のサーベイランス問題の解析に応用されている (高橋・丹後, 2008)。本研究では、東京都の COVID-19 陽性者数ホットスポットの特徴を、この前向き法の観点から分析することを試みる。

3.2 時空間ホットスポットの検出

空間スキャン検定の適用にあたり、本研究では (2.2) 式と (2.3) 式の調整人口 $n_{i,t}$ と解析実行

時点(解析日) T_2 を用いて, H_0 の下で地域人口数から期待される陽性者数 $\mu_{i,t}$ を次の(3.3)式によって算出した.

$$(3.3) \quad \mu_{i,t} = n_{i,t} \times \sum_{i=1}^{53} \sum_{t=1}^{T_2} \frac{O_{i,t}}{n_{i,t}}$$

ここに, $\sum_{i=1}^{53} \sum_{t=1}^{T_2} O_{i,t} = \sum_{i=1}^{53} \sum_{t=1}^{T_2} \mu_{i,t}$ である. また, 本研究では次の設計の下で prospective-cylindrical scan 法を適用した. なお, 実際の解析にはソフトウェア SaTScan™ (Kulldorff et al., 2021), ならびに SaTScan™ を R 上で実行するための rsatscan ライブラリー (Kleinman, 2015)を用いている.

- ウィンドウの平面上の最大半径: 10km
- ウィンドウの時間軸上の最大日: 14日間
- 読み込むデータの開始日 T_1 : 2020年4月1日 ($t=1$)
- 解析日 T_2 : 2020年4月16日から2021年5月31日まで毎日解析 ($t=16, 17, \dots, 426$)
- モンテカルロシミュレーションの回数: 9999回
- 有意と判断する基準: $p < 0.0027$

本研究の設計の概略を図4に示す. このように T_2 を更新しながら解析することで, 日々の

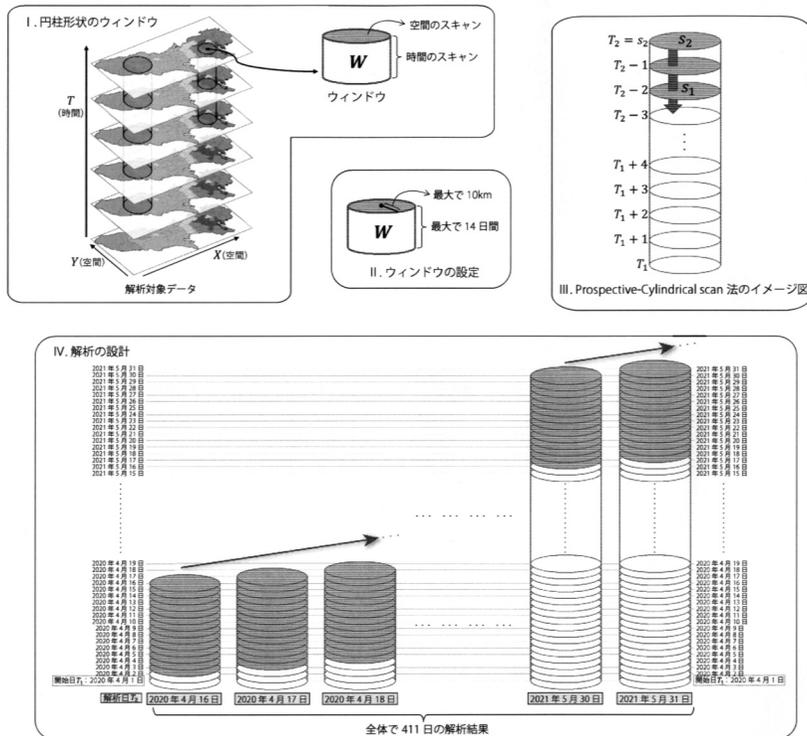


図4. I. Cylindrical scan 法のイメージ図. 平面を空間, 高さを時間に対応させた円柱形状のウィンドウを用いてスキャンを行う. II. 本研究で設定したウィンドウ. III. Prospective-cylindrical scan 法のイメージ図. 常に最近の時点が含まれるようにスキャンする. IV. 本研究における解析設計の概要. 読み込むデータの開始日は2020年4月1日に固定し, 解析日は2020年4月16日から2021年5月31日まで毎日行う.

サーベイランスの観点からの COVID-19 陽性者のホットスポットの動向に焦点をあてた解析となる。なおここに、地域間の距離を算出するために必要な地域 i の代表点の座標には、国土数値情報ダウンロードサービス（国土交通省, 2021）で公開されている行政区域（ポリゴン）のシェープファイルから算出される地域の重心座標（緯度・経度）を用いた。SaTScan™ において、これら緯度・経度に基づいた地球表面上の距離が使用される。また、ウィンドウの時間軸上の最大日を 14 日間としたのは、COVID-19 の潜伏期間が 1~14 日間ほどとされていることによる。さらに $p < 0.0027$ は、検出されたホットスポット候補が 1 年に 1 回の頻度よりもまれな現象であるか否か ($1/365 \approx 0.0027$) という意味合いを持つ。日々の解析を実施した結果、有意なホットスポットと認められた地域を図 5 にまとめている。この図は、縦方向に地域、横方向に解析日を示しており、各々の解析日において有意なホットスポットと検出された地域に色付けしている。ここに、色の濃淡は“解析終了日から遡って何日間継続中のホットスポットであるか”を表している。また、ホットスポットとして認められた回数を月毎に集計した結果をそれぞれ図 6 と図 7 に、 $LLR(\hat{W})$ の日次推移を図 3 (実線) に示す。

4. 東京都のホットスポット地域に対する実効再生産数の推定

4.1 対数感染者数変化率と実行再生産数の関係と推定に必要な仮定

Wallinga and Lipsitch (2007) は、再生方程式論を用いて、観測日 t の実効再生産数 R_t と、感染者数 y_t (ここでの「感染者数」は、PCR 検査等で発見されていない「潜在的な感染者を含む感染者数」という趣旨で用いており、前節で扱った「陽性者数」とは異なる) の観測日 t における対数感染者数変化率 $\lambda_t = \ln(y_t/y_{t-1})$ に、観測日 t 近傍で λ_t は一定との仮定の下で (4.1) 式が成立することを示した。

$$(4.1) \quad R_t = 1/M(-\lambda_t)$$

ここで、 $M(\theta)$ は、観測日 t で感染した者が、観測日 $s > t$ で他者に感染させる世代時間分布 $p(s|t)$ のモーメント母関数である。なお、世代時間分布は連続分布である。以下では、(2.1) 式の地域毎の日次陽性者数 $o_{i,t}$ を用いて、COVID-19 の地域対数感染者数変化率 $\lambda_{i,t}$ を推定し、(4.1) 式に基づいて実効再生産数を推定する。日々の期待陽性者数は、真の感染者数 $y_{i,t}$ の過小評価であり、そのバイアスも感染初期の検査強度が小さかった時期と PCR 検査が普及した 2021 年以降では異なると考えられる。一方、本節における分析の重要な仮定は、対数期待陽性者数の日次変化率の推定値を、対数感染者数日次変化率 $\lambda_{i,t} = \ln(y_{i,t}/y_{i,t-1})$ として利用可能とすることである。これは、感染者数の検査等による捕捉率 $q_{i,t} = o_{i,t}/y_{i,t}$ について、 $q_{i,t} = q_{i,t-1}$ との仮定を行うことに相当する。感染者数の変動よりは捕捉率の変動の方が緩慢であるという想定である。従って、推定に当たっては検査強度の曜日効果等について配慮する必要がある。

4.2 対数期待陽性者数の推定とその対数変化率の推定

本節では、地域 i ($i = 1, 2, \dots, 53$) における観測日 t ($t = 1, 2, \dots, 426$) の対数期待陽性者数を新たに $\ln \mu'_{i,t}$ と与え、一般化加法モデルを当てはめる。ここに、(2.2) 式と (2.3) 式の調整人口 $n_{i,t}$ 当たりの陽性者数の変化を説明するため $\ln n_{i,t}$ をオフセット項として用いた。

$$(4.2) \quad \ln \mu'_{i,t} = \ln n_{i,t} + \alpha_{i,j} d_{j,t} + f_i(t) + \beta_{1,i} \ln(o_{i,t-1} + 0.5) + \beta_{2,i} \ln(o_{i,t-2} + 0.5)$$

ここで、 $d_{j,t}$ ($j = 1, 2, \dots, 7$) は曜日ダミー変数で、例えば $d_{1,t}$ は、観測日 t が月曜日ならば 1、そうでなければ 0 となる。 $\alpha_{i,j}, \beta_{1,i}, \beta_{2,i}$ は未知パラメータ、 $f_i(t)$ はノンパラメトリック項で 3 次スプライン関数を用い、全ての観測日 t をノードに指定している。すなわち、 $\ln \mu'_{i,t}$ の t に関する 2 階偏導関数、対数陽性者変化率の導関数の二乗をペナルティとする平滑化回帰を実施し

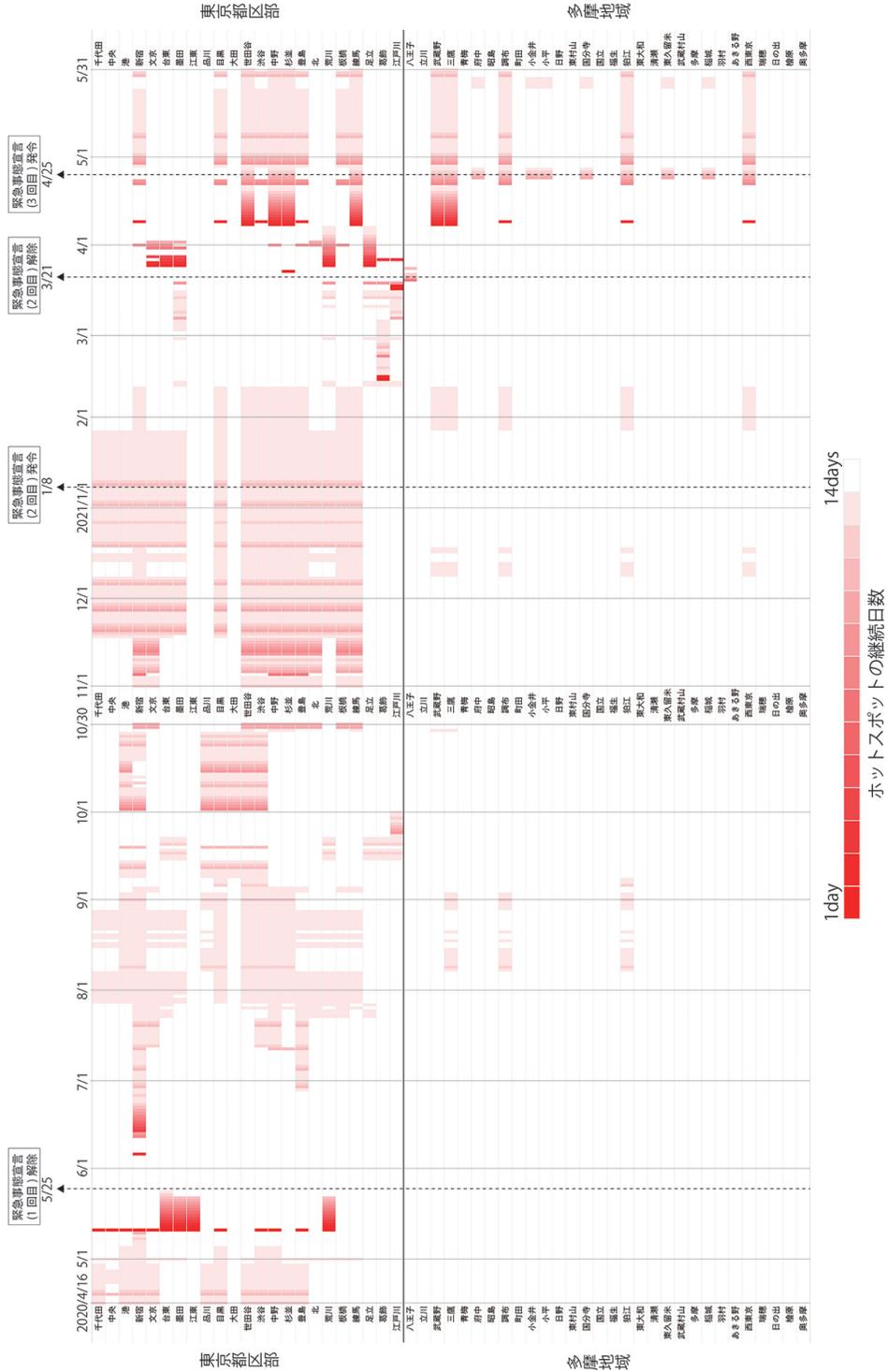


図 5. 検出された日々の時空間ホットスポット.

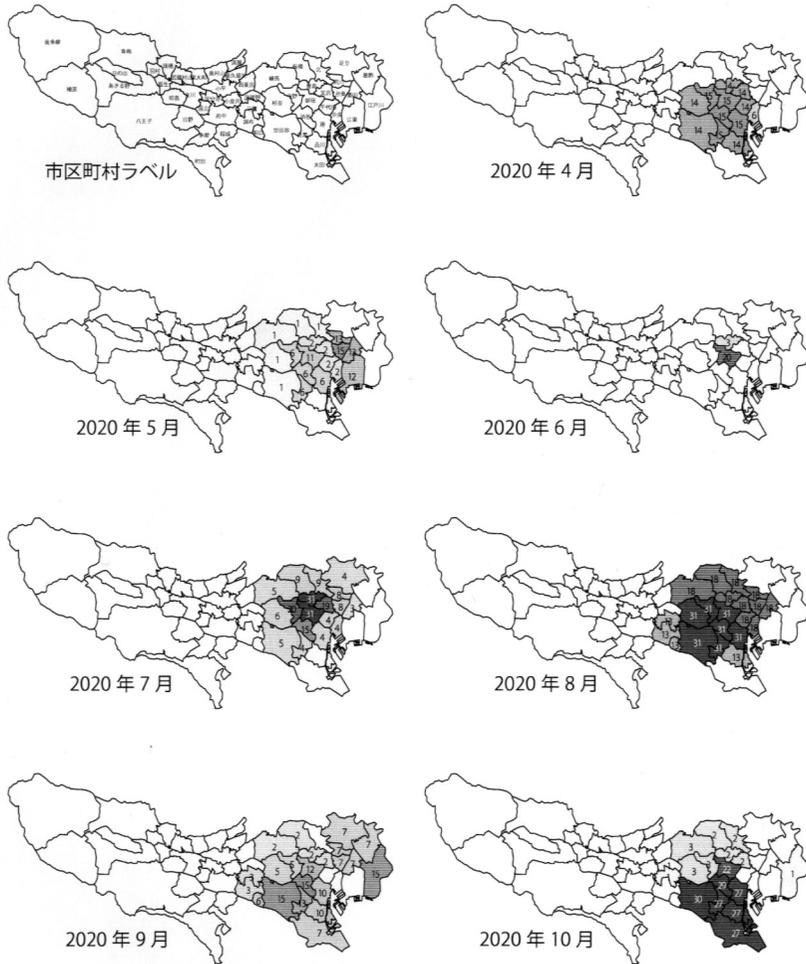


図 6. ホットスポットと認められた回数の月毎の集計結果(2020年4月~10月).

た. (4.2) 式のモデルを当てはめた上で, 本研究では対数感染者数変化率の推定値 $\hat{\lambda}_{i,t}$ を (4.3) 式で求めている.

$$(4.3) \quad \hat{\lambda}_{i,t} = \hat{f}_i(t) - \hat{f}_i(t-1)$$

また, (4.3) 式の標準誤差は, 厳密性を欠くが (4.3) 式の右辺第 1 項, 第 2 項を無相関と考え, 単純に標準誤差二乗和の平方根で評価している. なお, 一般化加法モデルの当てはめには, R の mgcv ライブラリー (Wood, 2021) の gam 関数を用いた.

4.3 実効再生産数の推定

(4.1) 式の $p(s|t)$ に従う世代時間確率変数 $S-t$ が期待値 $\nu(t)$ の確率分布に従うと仮定する. 更に分布形に指数分布を仮定すると, 地域 i の実効再生産数は, 単純に (4.4) 式で推定される.

$$(4.4) \quad \hat{R}_{i,t} = 1 + \nu(t)\hat{\lambda}_{i,t}$$

COVID-19 の世代時間分布について, Nishiura et al. (2020) では, 確定症例以外の症例を含

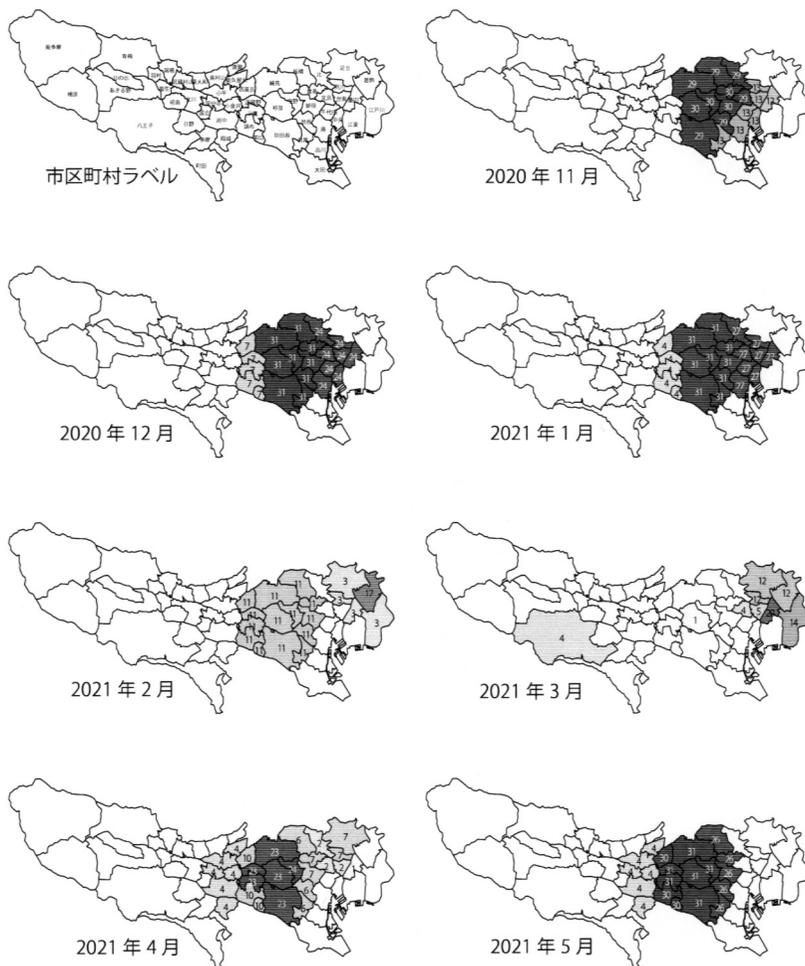


図7. ホットスポットと認められた回数の月毎の集計結果(2020年11月~2021年5月).

む 28 例に対して対数正規分布を当てはめた際の期待値は 4.7 日, 18 の確定症例データに対するワイブル分布の当てはめでは 4.8 日となっている. また, 世代時間分布の標準偏差は, それぞれ 2.9 日となっている. 実際には様々な世代時間あるいは感染可能時間研究において, 様々な分布が当てはめられている. 特に発症前の感染可能性のデータに基づく推定は, 困難であるが重要である. しかし, 対数感染者数変化率が 0 の近傍では, $R_{i,t} = 1 + \nu(t)\hat{\lambda}_{i,t} + O(\hat{\lambda}_{i,t}^2)$ となる. なお, ここでの O はランダウ記法によるオーダー評価を意味する. 少なくとも東京都の日々の対数感染者数変化率から実効再生産数を推定する場合には, 世代時間分布の期待値推定精度の確保が重要であり, 実行再生産数推定への世代時間分布形の寄与は顕著とならない.

(4.4)式の実効再生産数の推定精度について, δ 法を用いれば $\hat{R}_{i,t}$ の推定誤差分散は,

$$(4.5) \quad \text{Var}[\hat{R}_{i,t}] \sim \text{Var}[\nu(t)]\hat{\lambda}_{i,t}^2 + \nu(t)^2 \text{Var}[\hat{\lambda}_{i,t}]$$

と近似評価できる. ここで, $\nu(t)$ と $\hat{\lambda}_{i,t}$ は推定されたデータソースが異なるため統計的に独立とみなせる. また, Var は推定量の分散である. (4.5)式より, 陽性者数の少ない地域単独で推

定した $\hat{R}_{i,t}$ の標準誤差は大きくなることが予想される。従って、東京都の各々の地域に対して何らかの基準でクラスタリングを行い、その同一クラスター内の実効再生産数は等しいという仮定を置く。本研究では、3節で行った解析で得られたホットスポット地域内の実効再生産数は等しいと仮定する。

世代時間分布の期待値 $\nu(t)$ は、ウイルスの変異や公衆衛生政策(感染者の隔離等)に強く依存する概念だが、本研究では Nishiura et al. (2020) の 4.8 日を定数として用い、その推定誤差分散として、Nishiura et al. (2020) で推定された世代時間分布の分散を確定症例のサンプルサイズで除した $\text{Var}[\nu(t)] = 2.9^2/18 = 0.467$ を用いる。

また、ホットスポット地域の(4.4)式による推定値の加重平均値をホットスポットの実効再生産数の推定値とする。ホットスポットに属する地域の集合を J とすると、ホットスポットの実効再生産数の推定値 $\hat{R}_{J,t}$ は(4.6)式となる。

$$(4.6) \quad \hat{R}_{J,t} = \frac{\sum_{i \in J} I_{i,t} \cdot \hat{R}_{i,t}}{\sum_{i \in J} I_{i,t}}$$

ただし、 $I_{i,t}$ は推定誤差分散((4.5)式)の逆数であり、加重平均値の分散は

$$(4.7) \quad \text{Var}[\hat{R}_{J,t}] = \frac{1}{\sum_{i \in J} I_{i,t}}$$

と推定される。

4.4 対数陽性者変化率を用いた対数感染者変化率の推定

$LLR(\hat{W})$ が最大となる 2021 年 1 月 18 日のホットスポットを形成した千代田区、中央区、港区、新宿区、文京区、台東区、墨田区、目黒区、世田谷区、渋谷区、中野区、杉並区、豊島区、北区、荒川区、板橋区、練馬区の 17 区各々について、トレンド項 $f(t)$ の偏回帰プロットの推定値と、実効再生産数 $\hat{R}_{i,t} \pm$ 標準誤差を算出した結果を図 8 に示す。また、比較のために 2021 年 1 月 18 日にホットスポットというクラスターを形成していない 36 市区町村についても同様の計算を行ったものが図 9 である。

それに対し、これら 17 のホットスポット地域の範囲を 4.3 節で定義した地域の集合 J として、それぞれ実効再生産数を(4.6)式、その \pm 標準誤差を(4.7)式に基づいて算出し、その結果を図示したものが図 10(A)である。また、時空間的ホットスポットを形成していないため統計的に重み付き平均値の算出に意味があるとは言えないが、比較のため非ホットスポット 36 市区町村についても同様な加重平均を算出したのが図 10(B)となっている。

5. 考察

検出されたホットスポットと実効再生産数の推定結果に基づいた考察、ならびにこれら 2 つの分析結果の関連性について確認する。なお、前向き法によるホットスポットの詳細な結果については、紙面の制約上、それぞれの期間において主だった 10 日間のみをピックアップした。

5.1 2020 年 4 月～7 月中旬

図 5, 図 6 において、感染の第一波と言われる 2020 年 4 月に着目すると、東京都区部の都心から南西部にかけてホットスポットが認められ、5 月に入るとそれが徐々に縮小していく様子が見られる。なお、5 月 11 日に唐突に発生した広範囲に及ぶホットスポットは、これまで調査中となっていたものや都外データの多数が、台東区、江東区、墨田区、荒川区等に加算されたことに起因している。その後、1 回目の緊急事態宣言が解除された 5 月 25 日から翌 6 月初旬ま

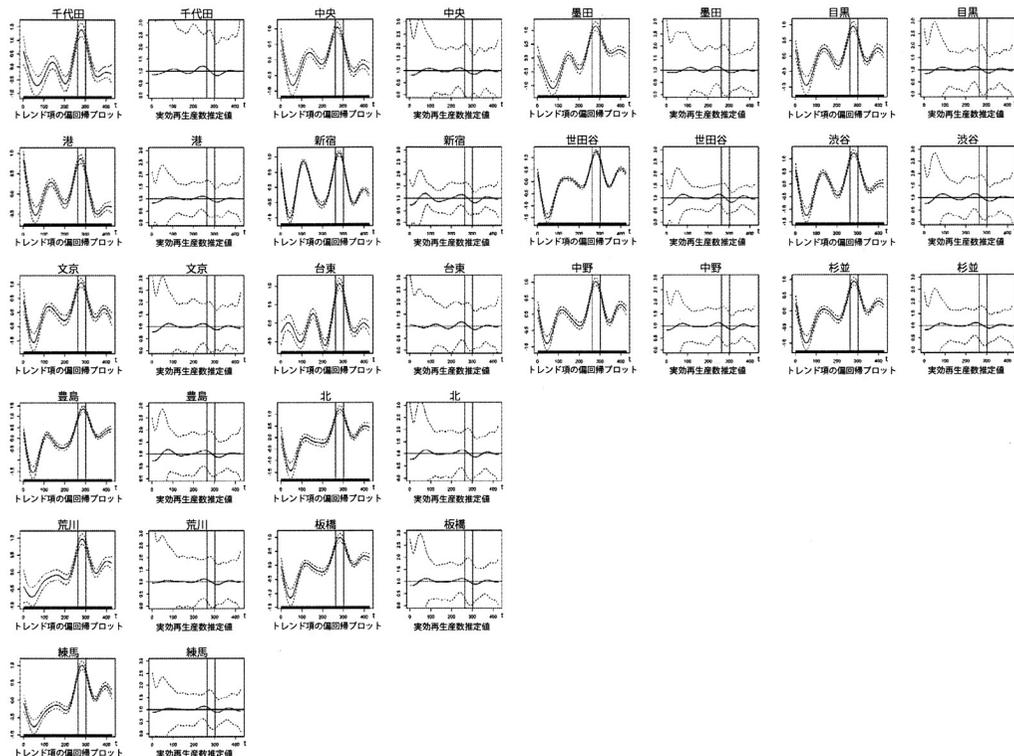


図 8. 「2020 年 12 月 19 日から 2021 年 1 月 27 日までのホットスポット」の 17 区の陽性者数平滑化トレンドと実効再生産数の推定値(それぞれ、実線：点推定値、破線：点推定値 ± 標準誤差、垂直線 2 本：2020 年 12 月 19 日と 2021 年 1 月 27 日、を示す)。

表 3. 2020 年 6 月 10 日から 6 月 19 日までの解析結果.

解析日	\hat{W}		$o(\hat{W})$	$\mu(\hat{W})$	$RR(\hat{W})$	$LLR(\hat{W})$	p 値
	地域	期間					
6/10(水)	新宿	6/4 - 6/10	36	18.813	1.914	6.206	0.1897
6/11(木)	新宿	6/4 - 6/11	45	21.364	2.106	9.945	0.0060
6/12(金)	新宿	6/4 - 6/12	57	23.869	2.388	16.600	0.0001
6/13(土)	新宿	6/4 - 6/13	62	26.172	2.369	17.777	0.0001
6/14(日)	新宿	6/11 - 6/14	48	10.303	4.659	36.309	0.0001
6/15(月)	新宿	6/11 - 6/15	68	12.959	5.247	57.992	0.0001
6/16(火)	新宿	6/11 - 6/16	77	15.531	4.958	62.188	0.0001
6/17(水)	新宿	6/11 - 6/17	79	18.064	4.373	56.005	0.0001
6/18(木)	新宿	6/11 - 6/18	88	20.635	4.265	60.720	0.0001
6/19(金)	新宿	6/11 - 6/19	98	23.121	4.239	67.213	0.0001

での期間は、本解析において唯一ホットスポットが検出されず、この頃には感染が収束しつつあったと言えるだろう。ところが、6月中旬に新宿区が再び有意なホットスポットとなり、以後しばらくの間「新宿区のみホットスポット」という状態が続く。

表 3 は、6 月 10 日から 6 月 19 日までに検出されたホットスポットの動向をまとめたもので

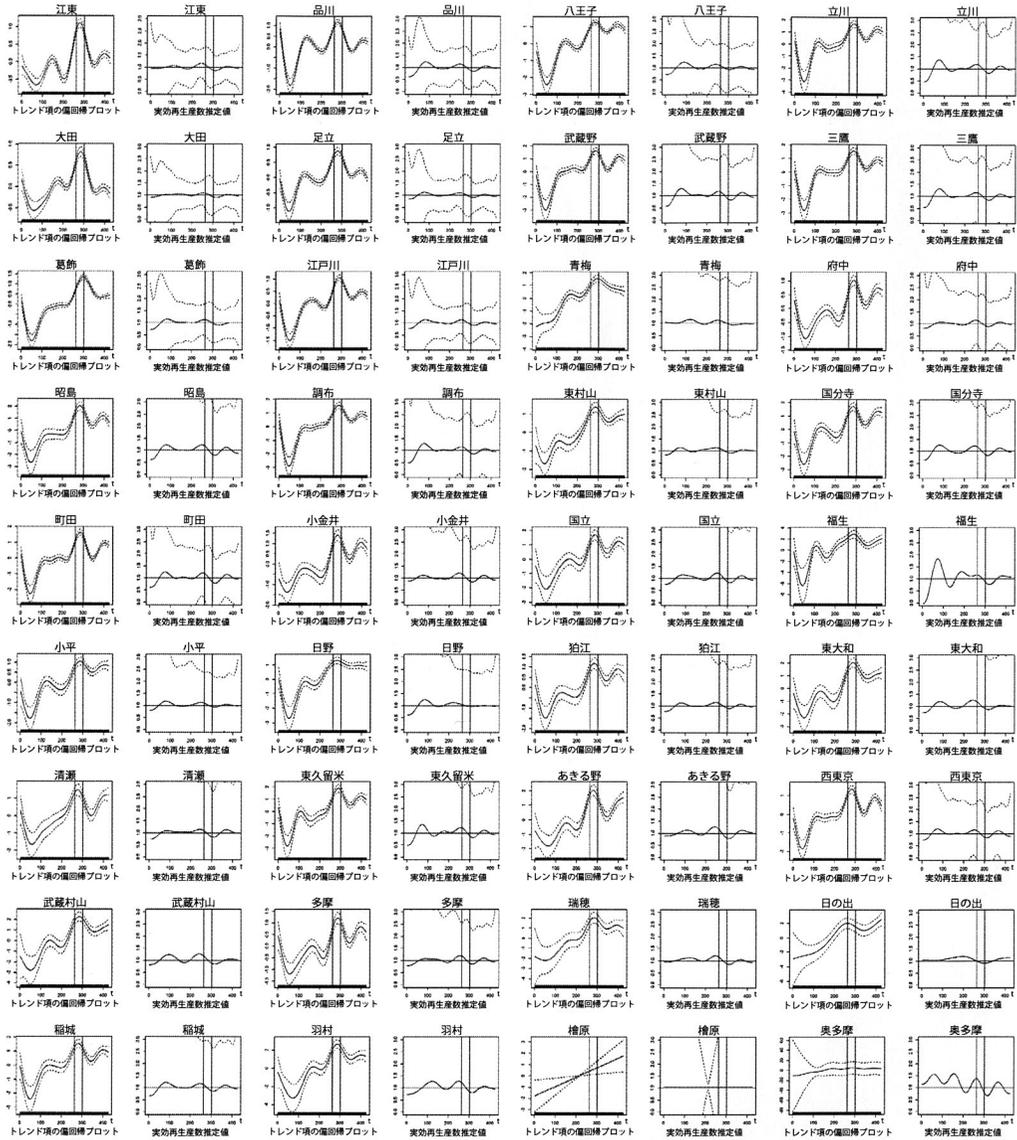


図 9. 「2020 年 12 月 19 日から 2021 年 1 月 27 日までのホットスポット」以外の 36 市区町村の陽性者数平滑化トレンドと実効再生産数の推定値(それぞれ、実線：点推定値、破線：点推定値 ± 標準誤差、垂直線 2 本：2020 年 12 月 19 日と 2021 年 1 月 27 日、を示す)。

ある。ここに、ホットスポット内の相対リスクを $RR(\hat{W}) = o(\hat{W})/\mu(\hat{W})$ で示している。

6 月 11 日の新宿区の $LLR(\hat{W})$ が前日のものから突如として約 1.5 倍に上昇し、翌 12 日以降から有意なホットスポットが検出されている。また、6 月 14 日の解析に着目すると、前日の結果に比べ $RR(\hat{W})$ と $LLR(\hat{W})$ が約 2 倍に上昇している。さらに、この 6 月 14 日以降に検出されるホットスポットは、皮肉にも東京アラートが解除された日にあたる 6 月 11 日を起点日

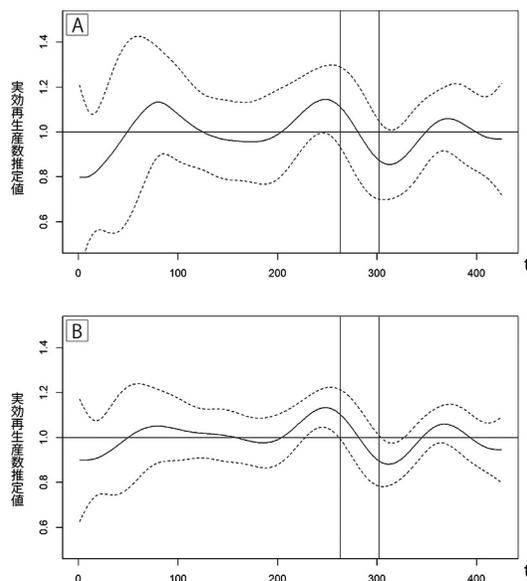


図 10. (A)ホットスポットの 17 区の実効再生産数の推定値の加重平均 $\hat{R}_{J,t}$ と $\hat{R}_{J,t} \pm$ 標準誤差および(B)非ホットスポットの 36 市区町村の実効再生産数の推定値の加重平均 $\hat{R}_{J,t}$ と $\hat{R}_{J,t} \pm$ 標準誤差.

としたものであった。また、この時期における新宿区の実効再生産数の推定値が 1 を大きく上回っていることから (図 8; $t = 71 \sim 80$)、その深刻さが伺える。

この頃は、主に「夜の街」関連の陽性者が新宿に集中しているとニュースで頻繁に報道されていた時期で、実際に陽性者の多くは、いわゆる「接待を伴う飲食店」に対する集団検査によって確認されている。東京アラート解除に伴うこれら飲食業に対する休業要請の段階的な緩和とそれに関連した人々の活動が、ホットスポットの結果に少なからず影響していた可能性が考えられる。その後、6月下旬から7月にかけて、新宿区に隣接する中野区や、商業活動の中心地で副都心と位置づけられている豊島区、渋谷区、文京区へとホットスポットが拡大している様子が図 5 から読み取れる。新宿区におけるこの時期の感染拡大の抑え込みの取り組みが、一つのターニングポイントであったことが示唆される結果となった。

また、図 8 の台東区と墨田区、図 9 の江東区の 3 地域の 7 月上旬辺り ($t = 100$ 周辺) の実効再生産数に着目すると、それが上昇傾向を示している様子が見て取れる。これは同時期の他地域では見られない動向であった。この頃は陽性者がある程度抑えられていたこともあり、これら互いに隣接しあう 3 地域がこの時期にホットスポットとはなっていないが、この後 7 月中旬から下旬にかけ、これら 3 地域に隣接する地域がホットスポットとなることは興味深い。

5.2 2020 年 7 月下旬～11 月上旬

図 5 より、第二波と言われる 7 月下旬から 8 月にかけては、主に世田谷区、杉並区、目黒区、港区を中心とする都心から西部において連日ホットスポットが認められた。加えて、世田谷区に隣接する三鷹市、調布市、狛江市といった多摩地域にまでホットスポットが広がっている様子が見て取れる。9 月に入るとホットスポットの範囲にやや縮小の様子が見られるが、これまで検出されることのなかった南部の大田区や、東部の江戸川区、葛飾区、足立区等といった居住地域の占める割合が多い地域で新たにホットスポットが認められるなど、以前とは異なる傾

表 4. 2020 年 7 月 22 日から 7 月 31 日までの解析結果.

解析日	\hat{W}		$o(\hat{W})$	$\mu(\hat{W})$	$RR(\hat{W})$	$LLR(\hat{W})$	p 値
	地域	期間					
7/22(水)	新宿, 文京, 渋谷, 中野, 豊島	7/9 - 7/22	1093	155.702	7.020	1242.975	0.0001
7/23(木)	新宿, 文京, 台東, 中野, 豊島, 北, 荒川, 板橋, 足立	7/10 - 7/23	1467	287.533	5.102	1289.652	0.0001
7/24(金)	新宿, 文京, 台東, 中野, 豊島, 北, 荒川, 板橋, 足立	7/11 - 7/24	1438	291.909	4.926	1218.846	0.0001
7/25(土)	新宿, 文京, 台東, 中野, 豊島, 北, 荒川, 板橋, 足立	7/12 - 7/25	1462	296.742	4.927	1238.462	0.0001
7/26(日)	新宿, 文京, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 板橋, 練馬	7/13 - 7/26	1654	414.240	3.993	1131.292	0.0001
7/27(月)	新宿, 文京, 台東, 中野, 豊島, 北, 荒川, 板橋, 足立	7/14 - 7/27	1402	301.188	4.655	1117.408	0.0001
7/28(火)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	7/15 - 7/28	2136	657.375	3.249	1153.200	0.0001
7/29(水)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	7/16 - 7/29	2172	665.147	3.265	1179.920	0.0001
7/30(木)	千代田, 中央, 港, 新宿, 文京, 台東, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	7/17 - 7/30	2147	653.822	3.284	1169.788	0.0001
7/31(金)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	7/18 - 7/31	2328	699.767	3.327	1296.765	0.0001

向が見られた。その後、10 月から 11 月にかけては、東京都区部の南西部から北西部へとホットスポットの地域が大きく移り変わる様子が見て取れた。また、この頃から $LLR(\hat{W})$ が再び上昇し始めていることがわかる(図 3)。なお、10 月 1 日には経済対策として国が実施していた Goto トラベルキャンペーンに東京都が追加されている。

7 月 22 日から 7 月 31 日までに検出されたホットスポットの動向を表 4 にまとめた。連日に渡り「14 日間のホットスポット」が検出され、その地域は目まぐるしく変化しながら徐々にその範囲を広げている。

5.3 2020 年 11 月中旬～2021 年 1 月

第三波と言われる 11 月中旬からは、ホットスポットは東京都区部のほとんどの地域とそこに隣接する多摩地域の一部まで含めた広範囲に及び、翌年の 1 月下旬までその状況にほとんど変化が見られなかった(図 5)。なお、1 月には東京都全体で連日 1,000 人を超える陽性者が報告され(図 1)、さらにこの頃のホットスポットは常に 10 日以上継続中の状態で滞留するなど、これまでになく深刻な状況であったことが伺える。

図 8 で示した都区部 17 地域は、2020 年 12 月 19 日から 2021 年 1 月 27 日までの 40 日間絶えずホットスポットとして維持され続けていたが、これらの偏回帰プロットの形状が近く、かつ地理的にも近い 17 地域が抽出されたと解釈できる。一方、対数陽性者変化率ならびにその線形変換である実効再生産数の点推定値自体は、これら 17 地域の感染拡大傾向の推移をよく表している。特にホットスポットとなっていた期間においては、感染拡大から縮小に移行する状況は 17 地域で共通している。この変動の類似性からも、ホットスポット地域でクラスタリングを施した実効再生産数の推定は有効と考えられる。なお、17 地域の中で単独での実効再生産数推定精度が高い世田谷区に比べても、図 10(A) に示す加重平均値の標準誤差は平均で 36.6% になっている。ただし、この方法でも実効再生産数が有意に 1 を超えているあるいは下回っているとするだけの推定精度は出ないが、空間クラスタリングによる小地域実効再生産数推定精度向上の効果は確認できたものと考えられる。

2021 年 1 月 9 日から 1 月 18 日までに検出されたホットスポットの動向をまとめたものを表 5 に示す。この期間は、 $RR(\hat{W})$ 、 $LLR(\hat{W})$ 共に非常に高い値で推移しており、1 月 18 日には本研究期間において最も高い $LLR(\hat{W})$ を示した。表 5 の 1 月 9 日から 1 月 11 日の結果に

表 5. 2021 年 1 月 9 日から 1 月 18 日までの解析結果.

解析日	\hat{W}		$\sigma(\hat{W})$	$\mu(\hat{W})$	$RR(\hat{W})$	$LLR(\hat{W})$	p 値
	地域	期間					
1/9(土)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2020/12/29 - 2021/1/9	8603	1407.635	6.112	8782.824	0.0001
1/10(日)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2020/12/29 - 2021/1/10	9276	1548.688	5.990	9336.537	0.0001
1/11(月)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2020/12/29 - 2021/1/11	9849	1690.925	5.825	9702.461	0.0001
1/12(火)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2020/12/30 - 2021/1/12	9793	1685.611	5.810	9616.546	0.0001
1/13(水)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2020/12/31 - 2021/1/13	10080	1701.554	5.924	10070.996	0.0001
1/14(木)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2021/1/1 - 1/14	10179	1736.318	5.862	10074.571	0.0001
1/15(金)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2021/1/2 - 1/15	10806	1795.362	6.019	10958.430	0.0001
1/16(土)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2021/1/3 - 1/16	11239	1840.339	6.107	11548.436	0.0001
1/17(日)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2021/1/4 - 1/17	11562	1869.034	6.186	12014.511	0.0001
1/18(月)	千代田, 中央, 港, 新宿, 文京, 台東, 墨田, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 北, 荒川, 板橋, 練馬	2021/1/5 - 1/18	11675	1890.928	6.174	12109.949	0.0001

着目すると, これらは 12 月 29 日がホットスポットの起点日となっており, 続く 1 月 12 日で「12 月 30 日を起点日とする 14 日間のホットスポット」となって以降, その「14 日間のホットスポット」の状態が日々検出され続けている. この時期に特に高尤度となった要因として, ホットスポット地域での陽性者数の急増とそれが長期間継続していたことに加えて, 多くの企業で仕事納めだった 12 月 28 日の翌 29 日以降, 図 3 で示される人流の減少が (2.2) 式における従業者数 ($n^{(b)}$) \times 人流 ($n^{(m)}$) の減少に繋がり, ひいてはホットスポット地域の中でもとりわけ従業者数が多い千代田区, 中央区, 港区, 新宿区等の調整人口を減少させ, それにより相対的に $LLR(\hat{W})$ が押し上げられたと考えられる.

なおこの期間に, 東部に位置する江東区, 足立区, 葛飾区, 江戸川区や, 南部の品川区, 大田区がホットスポットとならなかったのは, 設定したウィンドウサイズの制限に依るものと考えられ, これら地域の感染が抑えられていたと判断するものではないことは図 1 から明らかである. また, この時期の 1 月 8 日には 2 回目の緊急事態宣言が発令されているが, ホットスポットに縮小の傾向が見られたのは 1 月末になってからであった.

5.4 2021 年 2 月～2021 年 5 月

2 月から 3 月にかけては, 複数の地域がホットスポットとして点在するものの, その $RR(\hat{W})$ や $LLR(\hat{W})$ は第三波の頃と比較すると著しく減少しており, 感染状況にやや落ち着きが見える. ところが, 実効再生産数に着目すると, 2 回目の緊急事態宣言が明ける 3 月中旬頃から ($t = 350 \sim$), 多くの地域で 1 を上回り始める様子が見られる (図 8, 図 9).

2021 年 4 月 6 日から 4 月 14 日までに検出されたホットスポットの動向をまとめたものを表 6 に示す. この時期は, 世田谷区, 中野区, 杉並区, 練馬区, 武蔵野市, 三鷹市など, 主に居住人口の方が高い地域でホットスポットが検出されている. 2 回目の緊急事態宣言の解除と新年度を迎える時期とが重なったこと等が原因と考えられる急激な人流の増加 (図 3) によって, 従業者数の多い都心部で調整人口が増加したことも, これら居住人口の多い地域を中心にホットスポットが検出された一つの要因と考えられる.

表 6. 2021 年 4 月 6 日から 4 月 14 日までの解析結果.

解析日	\hat{W}		$o(\hat{W})$	$\mu(\hat{W})$	$RR(\hat{W})$	$LLR(\hat{W})$	p 値
	地域	期間					
4/6(火)	荒川, 足立	3/24 - 4/6	427	286.546	1.490	29.957	0.0001
4/7(水)	荒川, 足立	3/25 - 4/7	427	290.897	1.468	27.868	0.0001
4/8(木)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/8	255	131.861	1.934	45.104	0.0001
4/9(金)	新宿, 目黒, 世田谷, 渋谷, 中野, 杉並, 豊島, 練馬, 武蔵野, 三鷹, 調布, 狛江, 西東京	4/7 - 4/9	606	404.000	1.500	43.890	0.0001
4/10(土)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/10	475	283.220	1.677	53.998	0.0001
4/11(日)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/11	573	363.741	1.575	51.327	0.0001
4/12(月)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/12	645	428.389	1.506	47.538	0.0001
4/13(火)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/13	734	493.263	1.488	51.253	0.0001
4/14(水)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/14	859	558.385	1.538	69.762	0.0001
4/15(木)	世田谷, 中野, 杉並, 練馬, 武蔵野, 三鷹	4/7 - 4/15	1005	625.314	1.607	97.791	0.0001

その後の 4 月下旬から 5 月にかけて、 $LLR(\hat{W})$ を徐々に増加させながら、ホットスポットが再び都心部の地域にも拡大していく様子が見て取れる (図 5)。この頃には従来のものよりも感染力が強いとされる変異株の流行が確認され始めており、これも感染拡大の要因の一端を担っていると考えられる。それぞれ、4 月 12 日にはホットスポットを含む 23 区・6 市へのまん延防止等重点措置の適用、4 月 25 日には 3 回目の緊急事態宣言の発令、が行われたものの、5 月 31 日に至るまでホットスポットが縮小する様子は見られない。

6. おわりに

本研究では、東京都発表による 2020 年 3 月 31 日から 2021 年 5 月 31 日までの市区町村別集計された COVID-19 累積陽性者数に基づき算出した日次陽性者数に対し、まず第一の解析として、prospective-cylindrical scan 法に基づく空間スキャン検定を適用することにより、日々のホットスポット地域、及びそれがいつから発生しているのかを明らかにした。また、ホットスポットとして検出された地域の共通点や人的流動の観点からも解釈を試みた。その結果、経時的に検出されるホットスポットの起点日周辺と、政府・都の感染対策の動きや 1 年の節目の時期で生じたと思われる人流の変化との間に、少なからず関連性が見られた。また、地域的な観点では、従業者数の多い千代田区、中央区、港区の都心部が検出される時期、または居住人口の多い大田区、世田谷区、中野区、杉並区、板橋区、練馬区等の居住地域が多く検出される時期、といった具合に検出されるホットスポット地域に一定の共通点を見出すことができた。次に第二の解析として、日々の対数期待陽性者数とその対数変化率、及びそれらに基づく実効再生産数の推定法について述べると共に、新たな試みとして、第一の解析で得たホットスポット地域でクラスタリングを行うことによる実効再生産数の推定精度の改善を試みた。その結果、人口が一定規模以下の自治体の実効再生産数の区間推定を、ある程度実用的なものにできる可能性が示唆された。これら新たに確認できた種々の解析結果からも、東京都の陽性者数には地域性が存在することは明らかと言えるだろう。

東京都は、2020 年 9 月 15 日以降の週毎の濃厚接触者における感染経路別の割合を公開しており (東京都防災ホームページ, 2021)、それをまとめたものが図 11 である。これを図 1 や図 5 と対応させながら眺めてみると、「家庭内」の占める割合が高くなり始める 2020 年 10 月末頃を境に、ホットスポットは従業者数の多い都心を完全に外れ、比較的居住人口が多い地域で主に検出されている。陽性者が増加し始める 11 月中旬には都区部のほぼ全域がホットスポットとなるが、2021 年 2 月になりいったん感染状況が落ち着いた後、「家庭内」の占める割合が高くな

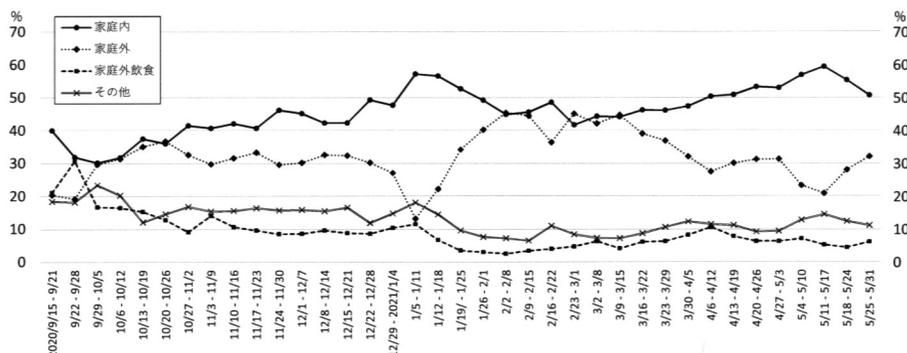


図 11. 東京都公表の陽性者数の濃厚接触者における感染経路別の割合の推移。公表資料における「同居」を家庭内、「施設等」と「職場」を合算したものを家庭外(施設等は、特別養護老人ホーム・介護老人保健施設・医療機関・保育園・学校等の教育施設等及び通所介護の施設をいう)、「会食」と「接待を伴う飲食」を合算したものを家庭外飲食としている。

る3月中旬頃から、再び居住地域を中心にホットスポットが広がり始める。推定された実効再生産数も上述の節目から上昇に転じている。本研究で得られたホットスポットの結果は、当初のCOVID-19関連の報道等で盛んに言われていたクラスター的な感染が、家庭内感染の占める割合の増加を起点として居住地域に移行し、その後その範囲を拡大していったことを示す一つの例証となるだろう。

以下に本研究の課題について述べる。一般的に、地域によって性・年齢・世代等の人口構成や陽性率は異なる。本来であれば、これらの要因を共変量と見なして調整を施した期待陽性者数を用いることが感染症や空間疫学的見解においては望ましい。本研究で用いたような従業者数や人流のデータに対し、このような各要因での経時的なCOVID-19陽性率の情報を完全に取得することは困難であるが、これらが考慮されたホットスポットについて検討することは重要である。また本研究では、シンプルにまずは「陽性者数」と「調整人口」から算出される日々の期待陽性者数に基づきホットスポット検出を行ったが、PCRの検査体制等で生じる陽性者数の曜日変動を考慮した期待陽性者数を用いる方法もあるだろう。なお、今回得られた結果では、特定の曜日に $LLR(\hat{W})$ や $RR(\hat{W})$ が高くなるといった特徴は見られなかった。一方で、後半の実効再生産数の推定では、日々の実際の感染者に対する検査陽性者という捕捉率の観点から、曜日変動等も取入れたより柔軟なモデルによって期待陽性者数の算出を試みている。実効再生産数の高低とホットスポット地域との間には必ずしも連動しない部分が存在したが、これは両解析におけるこれら期待陽性者数の与え方の違いも要因の一つと考えられる。モデルの整合性の確保については、引き続き検討すべき課題と考える。

また、今回は時空間ホットスポット検出において一般的によく利用される円柱形状のウィンドウによるスキャンを行ったが、これは空間的に円形状の範囲をとるホットスポットしか検出できない。商業地域や居住地域等といったホットスポット地域の特徴をより明確化するため、ウィンドウの形状には改善の余地が残されている。さらに本研究では最大対数尤度比となるウィンドウのみをホットスポットと見なして議論を進めたが、空間スキャン検定では2番目以降のホットスポットを検出するための方法もいくつか提案されている。同時期に検出される複数のホットスポットについても注視し、その動向を見極めることは、日々のサーベイランスの観点では重要になってくるであろう。

実効再生産数の推定においても、一般化加法モデルのベースとなる確率モデル、検査数等の

調整すべき他の入力変数の追加, 地域に共通な潜在パラメータの検討といった改善が考えられる. 引き続き, これら諸課題の改善に取り組むとともに, これら一連の解析を体系化した新たなサーベイランス手法の開発に繋げていきたい.

最後に, 本研究では2021年5月31日までのデータに対し解析を行ったが, それ以降も依然としてコロナ禍の完全終息は見通せない状況にある. 今後も継続的に解析を行い, 現状把握とその評価・検証に取り組んでいきたい.

謝 辞

本研究を進めるに当たり, 統計数理研究所「新型コロナウイルス対応プロジェクト」(代表: 松井知子教授)のメンバーの先生方からは様々な助言や指導等を頂きました. 心より深く感謝申し上げます.

参 考 文 献

- Acharya, B.K., Khanal, L., Mahyoub, A.S., Ruan, Z., Yang, Y., Adhikari, S.K., Pandit, S., Neupane, B.K., Paudel, B.K. and Lin, H. (2020). Execution of intervention matters more than strategy: A lesson from the spatiotemporal assessment of COVID-19 clusters in Nepal, medRxiv, doi: 10.1101/2020.11.07.20227520.
- Alkhamis, M.A., Al Youha, S., Khajah, M.M., Haider, N.B., Alhardan, S., Nabeel, A., Al Mazeedi, S. and Al-Sabah, S.K. (2020). Spatiotemporal dynamics of the COVID-19 pandemic in the State of Kuwait, *International Journal of Infectious Diseases*, **98**, 153–160, doi: 10.1016/j.ijid.2020.06.078.
- Amin, R., Hall, T., Church, J., Schlierf, D. and Kulldorff, M. (2020). Geographical surveillance of COVID-19: Diagnosed cases and death in the United States, medRxiv, doi: 10.1101/2020.05.22.20110155.
- Benita, F. and Gasca-Sanchez, F. (2020). On the main factors influencing COVID-19 spread and deaths in Mexico: A comparison between Phase I and II, medRxiv, doi: 10.1101/2020.12.22.20248716.
- Chen, C., Kim, A.Y., Ross, M. and Wakefield, J. (2018). *SpatialEpi v1.2.3: Methods and Data for Spatial Epidemiology*, <https://CRAN.R-project.org/package=SpatialEpi>.
- Chow, T.E., Choi, Y., Yang, M., Mills, D. and Yue, R. (2020). Geographic pattern of human mobility and COVID-19 before and after Hubei lockdown, *Annals of GIS*, **27**(2), 127–138, doi: 10.1080/19475683.2020.1841828.
- Cordes, J. and Castro, M.C. (2020). Spatial analysis of COVID-19 clusters and contextual factors in New York City, *Spatial and Spatio-temporal Epidemiology*, **34**, 100355, doi: 10.1016/j.sste.2020.100355.
- French, J. (2020). *Smerc v1.2: Statistical Methods for Regional Counts*, <https://CRAN.R-project.org/package=smerc>.
- Greene, S.K., Peterson, E.R., Balan, D., Jones, L., Culp, G.M. and Kulldorff, M. (2020). Detecting emerging COVID-19 community outbreaks at high spatiotemporal resolution, New York City, June–July 2020, medRxiv, doi: 10.3201/eid2705.203583.
- Hohl, A., Delmelle, E.M., Desjardins, M.R. and Lan, Y. (2020). Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States, *Spatial and Spatio-temporal Epidemiology*, **34**, 100354, doi: 10.1016/j.sste.2020.100354.
- Ishioka, F. (2020). *echelon v0.1.0: The Echelon Analysis and the Detection of Spatial Clusters using Echelon Scan Method*, <https://CRAN.R-project.org/package=echelon>.

- Kim, S. and Castro, M.C. (2020). Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020), *International Journal of Infectious Diseases*, **98**, 328–333, doi: 10.1016/j.ijid.2020.07.004.
- Kleinman, K. (2015). *rsatscan v0.3.9200: Tools, Classes, and Methods for Interfacing with SaTScan Stand-Alone Software*, <https://CRAN.R-project.org/package=rsatscan>.
- 国土交通省 (2021). 国土数値情報ダウンロードサービス, <https://nlftp.mlit.go.jp/ksj/> (最終閲覧日: 2021年6月30日).
- Kulldorff, M. (1997). A spatial scan statistic, *Communications in Statistics: Theory and Methods*, **26**(6), 1481–1496.
- Kulldorff, M. (2001). Prospective time periodic geographical disease surveillance using a scan statistic, *Journal of the Royal Statistical Society, Series A*, **164**(1), 61–72.
- Kulldorff, M., Feuer, E.J., Miller, B.A., Athas, W.F. and Key, C.R. (1998). Evaluating cluster alarms: A space-time scan statistic and brain cancer in Los Alamos, *American Journal of Public Health*, **88**(9), 1377–1380.
- Kulldorff, M. and Harvard Medical School, Boston and Information Management Services Inc. (2021). *SaTScan™ v9.7: Software for the Spatial and Space-time Scan Statistics*, <http://www.satscan.org>.
- Masrur, A., Yu, M., Luo, W. and Dewan, A. (2020). Space-time patterns, change, and propagation of COVID-19 risk relative to the intervention scenarios in Bangladesh, *International Journal of Environmental Research and Public Health*, **17**, 5911, doi: 10.3390/ijerph17165911.
- モバイル空間統計 (2021). 新型コロナウイルス感染症対策特設サイト, <https://mobaku.jp/covid-19/> (最終閲覧日: 2021年6月13日).
- Nishiura, H., Linton, N.M. and Akmetzhanov, A.R. (2020). Serial interval of novel coronavirus (COVID-19) infections, *International Journal of Infectious Diseases*, **93**, 284–286, doi: 10.1016/j.ijid.2020.02.060.
- Otani, T. and Takahashi, K. (2020). *rflexscan v0.3.1: The Flexible Spatial Scan Statistic*, <https://CRAN.R-project.org/package=rflexscan>.
- 高橋邦彦, 丹後俊郎 (2008). 疾病集積性の検定を用いた症候サーベイランス解析, *保健医療科学*, **57**(2), 122–129.
- 統計データ分析コンペティション (2021). SSDSE(教育用標準データセット), <https://www.nstac.go.jp/SSDSE/index.html> (最終閲覧日: 2021年6月13日).
- 東京都 (2020). 東京都区市町村別患者数(都内発生分)3月31日時点の累計値, https://www.metro.tokyo.lg.jp/tosei/hodohappyo/press/2020/04/02/documents/20_01a.pdf (最終閲覧日: 2020年4月1日).
- 東京都防災ホームページ (2021). 東京都新型コロナウイルス感染症モニタリング会議資料, <https://www.bousai.metro.tokyo.lg.jp/taisaku/saigai/index.html> (最終閲覧日: 2021年10月14日).
- 東京都福祉保健局 (2021). 報道発表 新型コロナウイルスに関連した患者の発生について(全バックナンバー: 2020年3月31日~2021年5月31日), <https://www.fukushihoken.metro.tokyo.lg.jp/hodo/index.html> (最終閲覧日: 2021年6月25日).
- 東京都総務局統計部 (2021). 東京都の統計, <https://www.toukei.metro.tokyo.lg.jp/jsuikai/js-index2.htm> (最終閲覧日: 2021年6月13日).
- Wallinga, J. and Lipsitch, M. (2007). How generation intervals shape the relationship between growth rates and reproductive numbers, *Proceedings of the Royal Society B: Biological Sciences*, **274**(1609), 599–604, doi: 10.1098/rspb.2006.3754.
- Wood, S. (2021). *mgcv v1.8-38: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*, <https://CRAN.R-project.org/package=mgcv>.

- World Health Organization (2021). WHO Coronavirus (COVID-19) Dashboard, <https://covid19.who.int/> (最終閲覧日：2021年6月13日).
- Yoneoka, D., Thoue, Y., Kawashima, T., Nomura, S., Shi, S., Eguchi, A., Ejima, K., Taniguchi, T., Sakamoto, H., Kunishima, H., Gilmour, S., Nishiura, H. and Miyata, H. (2020). Large-scale epidemiological monitoring of the COVID-19 epidemic in Tokyo, *The Lancet Regional Health—Western Pacific*, **3**, 100016, doi: 10.1016/j.lanwpc.2020.100016.

Space-time Hotspot Clusters of New Coronavirus and Improvement of Accuracy for Estimating Effective Reproduction Number Based on It —As an Example of Tokyo—

Fumio Ishioka¹, Hiroe Tsubaki², Takafumi Kubota³ and Kazuyuki Suzuki⁴

¹Graduate School of Environmental and Life Science, Okayama University

²The Institute of Statistical Mathematics

³School of Management and Information Sciences, Tama University

⁴Graduate School of Informatics and Engineering, The University of Electro-Communications

Coronavirus disease (COVID-19) has spread globally since the first case was confirmed in Wuhan, China, in December 2019. With the outbreak of coronavirus, various researches are being conducted all over the world including the evaluation of the data on hotspot clusters. Japan has not yet verified the domestic daily hotspot trends over a long period of time. This paper proposes a more accurate estimation method of effective reproduction number with hotspot information. Specifically, using the cumulative number of positive cases by municipalities for a total of 427 days from March 31, 2020 to May 31, 2021 that is publicized by the Tokyo Metropolitan Government, the spatial scan statistic based on prospective-cylindrical scan method is performed. As a result, when and where the hotspot existed each day are clarified. Moreover, from the obtained hotspot information, the occurrence of hotspot and its factors as well as the relationship between area / mobility and hotspot are detected. In the analysis, the number of employees and the mobility in a specific area are considered in addition to the resident population by municipalities. This paper examines the effective reproduction number by municipalities in Tokyo using the above hotspot information and shows that the method mentioned the above is more accurate. In addition, it is possibly to estimate to some extent even in the estimation of effective reproduction number of local governments with a small population.