

令和 2 年 5 月 6 日現在

機関番号：12601

研究種目：挑戦的研究（萌芽）

研究期間：2018～2019

課題番号：18K18535

研究課題名（和文）多様な地理パターンの定量的自動抽出手法の開発

研究課題名（英文）A general procedure for detecting geographical patterns

研究代表者

貞広 幸雄（Sadahiro, Yukio）

東京大学・大学院情報学環・学際情報学府・教授

研究者番号：10240722

交付決定額（研究期間全体）：（直接経費） 4,800,000円

研究成果の概要（和文）：特徴的な地理パターンの抽出は、地理学における重要な研究手法の一つである。そのため典型的地理パターンについては、効率的に抽出・検定する定量的手法が数多く開発されてきている。しかし、地理学で扱う地理パターンは遙かに多様であり、既存手法で抽出できる地理パターンは限定的である。そこで本研究では、より多様な地理パターンを抽出する新たな定量的手法を開発した。統計学的枠組みを利用し、様々な確率現象を仮定することで、既存手法にはない拡張性を確保した。

研究成果の学術的意義や社会的意義

本研究で開発した地理パターンの抽出方法は、極めて広範な地理解析に適用可能であり、既存手法の改良や新たな手法の開発などに大きく寄与することが出来る。また、地理解析は地理現象を対象とする様々な研究分野（例えば疫学や林学など）の研究者が行うことから、研究領域という意味においても広範な貢献が期待できると考えている。

研究成果の概要（英文）：Detection of spatial patterns is an important method in quantitative geography. There have been developed numerous methods for detecting spatial patterns efficiently and objectively. There exist, however, a wide variety of spatial patterns, not all of which can be captured by existing methods. To resolve the problem, this study proposed a new general procedure for detecting spatial patterns. The procedure is based on a statistical framework, which enables us to treat a wide variety of spatial patterns.

研究分野：空間情報科学

キーワード：地理パターン 統計的手法

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

特徴的な地理パターンの抽出は、地理学における重要な研究手法の一つである。地理パターンの背後にある要因を分析することで、空間現象の構造を解明し、その現象の引き起こす問題の解決へと繋げることができる。

地理パターン検出の最も基本的な方法は目視である。しかし目視には時間と労力を要し、また客観性にも欠けるという問題がある。そのため典型的な地理パターンについては、効率的に抽出・検定する定量的手法が数多く開発されてきている。Moran's I, LISA, Ripley の K 関数などはそれらの代表的な例である。しかし、地理学で扱う地理パターンは遙かに多様であり、既存手法で抽出できる地理パターンは限定的である。

2. 研究の目的

以上をふまえ本研究では、より多様な地理パターンを抽出する新たな定量的手法を開発する。目視による抽出の不安定さとそれに要する時間や労力を是正し、定量的パターン抽出手法を開発することで、分析の効率化と客観性の担保を目指す。

3. 研究の方法

本研究ではまず、既存手法で扱われている点パターンの種類について、詳細なレビューを行った。その結果、最も典型的なパターンである空間クラスター以外に、その時間次元への拡張である時空間クラスター、点の混合パターン、点クラスターの移動などの事例が見られた。これらについて可能な限り広範に適用可能な分析手法を検討し、その結果として空間スキャン統計量の考え方を応用することが最も適切であると判断し、研究を実施した。

4. 研究成果

本研究では、空間スキャン統計量の背景にある考え方を拡張し、さらに多様な地理パターンの分析手法を開発した。以下、その具体例の一つとして、点の生成・消滅パターンの分析手法を示す。

いま、領域 S_0 において、期間 T の間に N 個の点が存在したものとす。点の一部は全期間を通じて存在し、その他は一部期間のみ存在している。領域 S_0 内に中心 \mathbf{x} 、半径 r の円 $Z_0(\mathbf{x}, r)$ を取り、円柱 Z を $Z_0(\mathbf{x}, r) \otimes T$ 、円柱 S を $S_0 \otimes T$ とそれぞれ定義する。

次に、 Z 及び $S-Z$ において、生成が確率的に発生するモデルを考える。それぞれの発生確率を p_0, q_0 と表記し、その値は実際のデータから最尤法によって推定する。推定された p_0 が q_0 よりも大きければ、 Z では生成が相対的に頻繁であり、 p_0 が小さければ、生成が少ないということになる。

2つのモデルを用いて、 Z と $S-Z$ で生成確率が異なるかどうかを考える。 $p_0=q_0$ 及び $p_0 \neq q_0$ をそれぞれ帰無仮説、対立仮説とし、最尤法によって導出した尤度関数の最大値をそれぞれ $L_0(Z)$ 、 $L_1(Z)$ と書く。このとき対数尤度比は

$$\lambda_A(Z) = \log L_1(Z) - \log L_0(Z) \quad (1)$$

で与えられる。そして、空間スキャン統計と同様に、各地点 \mathbf{x} において Z の半径を徐々に拡大し、予め定めた最大値までの間で $\lambda_A(Z)$ が最大となる円を抽出する。そのときの $\lambda_A(Z)$ を生成頻度指標 $\alpha_A(\mathbf{x})$ と呼ぶ。

$$\alpha_A(\mathbf{x}) = \max_{r \leq r_{\max}} \lambda_A(Z) \quad (2)$$

$p_0 > q_0$ のとき、 Z では生成の頻度が相対的に高く、 $p_0 < q_0$ のときには頻度が相対的に低いことになる。以下、これらを $\{A+, A-\}$ と表記する（消滅の場合は $\{D+, D-\}$ と書く）。この2つを異なる色相で区別し、 $\alpha_A(\mathbf{x})$ を明度或いは彩度で表現することで、生成の相対的頻度を図化し、その空間パターンを把握することが可能となる。

次に、相対的頻度が他と比べて統計的に有意に異なる地域を抽出する。まず S_0 に格子網を重ね、各格子点で $\alpha_A(\mathbf{x})$ を計算し、降順に並べる。次に、大きな値の $\alpha_A(\mathbf{x})$ に対応する円から順に、互いに重ならないものを予め定めた個数になるまで選択する。そして、FDR(False Discovery Rate) を制御することで、多重検定の問題を回避しながら統計的に有意な円を抽出する。

以上の手順は、点の消滅に関しても同様に行われる。消滅の場合の $\lambda_A(Z)$ の最大値をここでは消滅頻度指標 $\alpha_D(\mathbf{x})$ と呼ぶ。

手法の有効性を検証するために、次に NTT タウンページの電話帳データを利用し、2001~2016年の渋谷区における商業施設分布の分析を行う。紙面の都合上、ここでは衣料品店と飲食店に関する分析結果のみを示す。図1及び図2はそれぞれ2016年における衣料品店と飲食店の分布である。

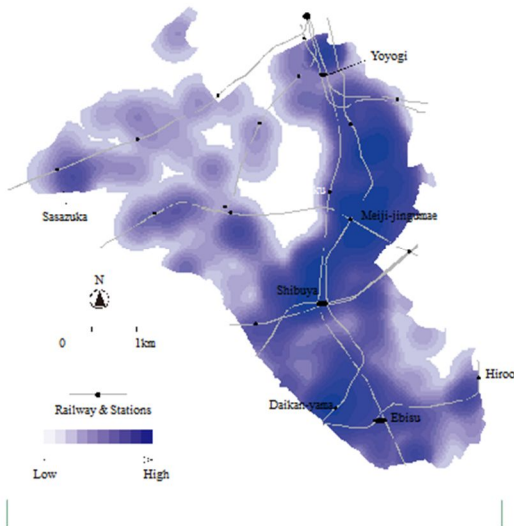


図1 2016年における衣料品店の分布

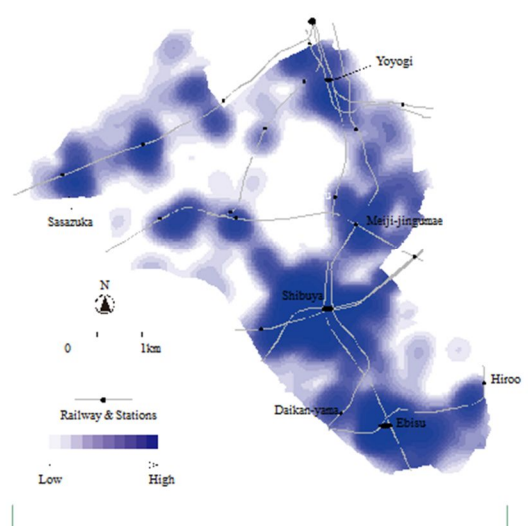


図2 2016年における飲食店の分布

図3は衣料品店の $\alpha_A(x)$ と $\alpha_D(x)$ の分布を表す。紫色が濃い地域では開閉店が頻繁であり、店舗間競争が厳しいことを表す。開閉店は渋谷駅北や代官山駅周辺で最も頻繁であり、原宿駅東や恵比寿駅南がそれに続く。濃い緑色は開閉店が少ないことを表しており、笹塚・恵比寿・広尾駅周辺、渋谷駅南などでは衣料品店間の競争がそれほど厳しくないと考えられる。図3は図1と類似しており、一般的に、店舗が多い地域ほど競争が厳しく、少ない地域ほど競争がないといえる。渋谷駅周辺では、北と南で開閉店の頻度が大きく異なる。前者では若者向けの様々な種類の衣料品店が多く競争も厳しいが、後者では古着やアメカジの店が長く存続していることがその要因と考えられる。

図4は飲食店の $\alpha_A(x)$ と $\alpha_D(x)$ の分布を表す。衣料品店とは異なり、図4は必ずしも図2とは類似していない。例えば渋谷駅周辺には飲食店が多いが、開閉店はそれほど頻繁ではない。渋谷駅近隣には老舗の飲食店が多く、昔からの顧客が根強く支持していることがその要因と思われる。恵比寿駅周辺にも飲食店は多いが、図4を見ると閉店が少なく開店が多い、即ち、飲食店が増加している。この地域では商業集積が西に拡大しつつあることが背景にあるものと考えられる。

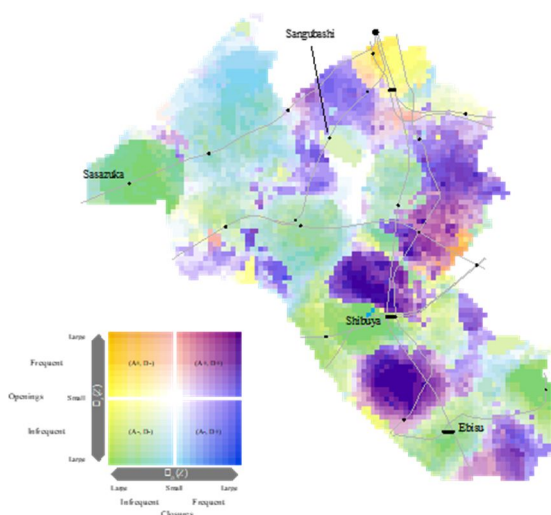


図3 衣料品店の $\alpha_A(x)$ と $\alpha_D(x)$ の分布

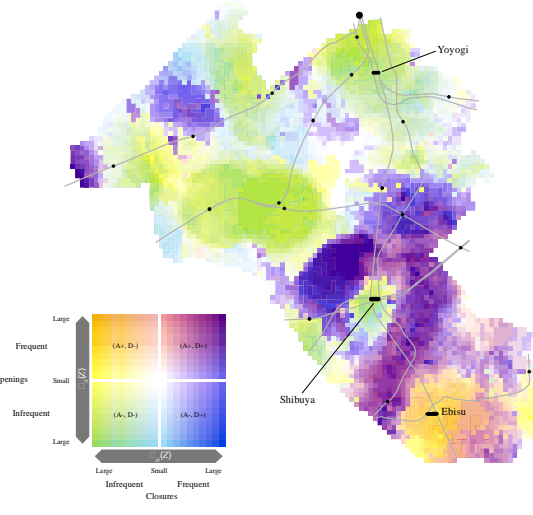


図4 飲食店の $\alpha_A(x)$ と $\alpha_D(x)$ の分布

図5は開閉店が統計的に有意に多い或いは少ない地域を示している。概ね図3・4と合致しているが、統計的検定によって $\alpha_A(x)$ や $\alpha_D(x)$ に対する評価がより細分化されることがわかる。例えば図3では渋谷、恵比寿、広尾駅周辺は同様の色に見えるが、図5aでは渋谷駅周辺が1・3位、広尾駅周辺が5位、恵比寿駅周辺は11位以下となっており、その差異が明確である。

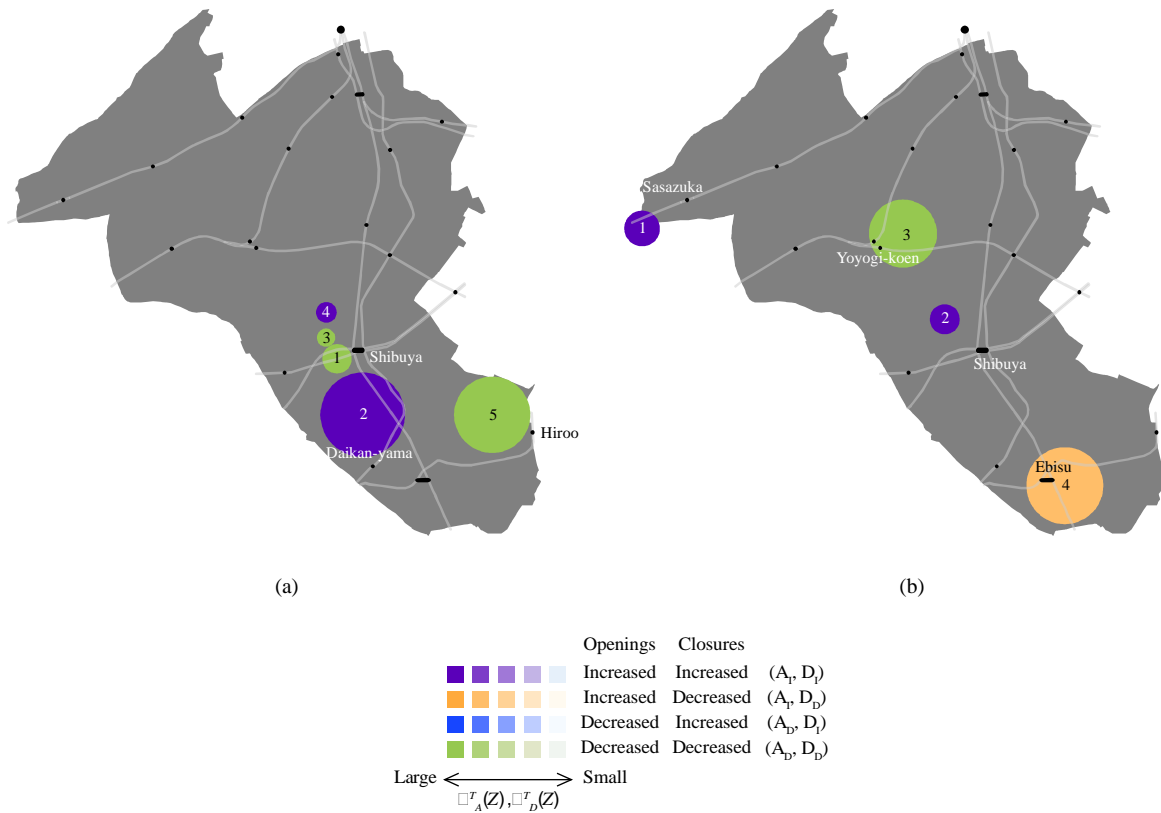


図5 開閉店の統計的に有意に多い或いは少ない地域。(a) 衣料品店, (b) 飲食店

上記の分析では, 点の生成と消滅を確率現象と見なし, その空間パターンを可視化, 特異地域の検出を行っている。ここで, 点の生成と消滅という確率現象を, 他の確率現象によって置き換えることで, 様々な地理パターンの評価が可能である。例えば Sadahiro (2019)では, 複数種の点分布が存在する状況において, 点の混合度を評価する手法を提案している。ここでは, 点が各種に割り当てられる現象を確率的に捉え, 上記と同様の枠組みを用いて分析を行っている。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 0件）

1. 著者名 Sadahiro, Y	4. 巻 33
2. 論文標題 Analysis of the appearance and disappearance of point objects over time	5. 発行年 2019年
3. 雑誌名 International Journal of Geographical Information Science,	6. 最初と最後の頁 215-239
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Sadahiro, Y & Y. Liu	4. 巻 -
2. 論文標題 A scale-sensitive approach for comparing and classifying point patterns	5. 発行年 2018年
3. 雑誌名 Journal of Spatial Science	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1080/14498596.2018.1492466 .	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計2件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 貞広幸雄
2. 発表標題 点分布における点の生成と消滅
3. 学会等名 第28回地理情報システム学会研究発表大会
4. 発表年 2019年

1. 発表者名 貞広幸雄
2. 発表標題 複数種の点分布における空間的混合
3. 学会等名 日本オペレーションズ・リサーチ学会2019年秋季研究発表会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----