

空間的相関と異質性から見た日本の経済集積の 探索的空間データ分析

爲季 和樹¹・堤 盛人²

¹学生非会員 筑波大学 大学院システム情報工学研究科 (〒305-8573 茨城県つくば市天王台1-1-1)

E-mail: tamesue.kazuki@sk.tsukuba.ac.jp

²正会員 筑波大学教授 システム情報系社会工学域 (〒305-8573 茨城県つくば市天王台1-1-1)

E-mail: tsutsumi@sk.tsukuba.ac.jp

集積の経済の分析において、古くから立地ジニ係数や特化係数など、経済集積を測る直接的な指標を用いた研究が行われてきた。また近年では、新経済地理学や空間計量経済学の台頭により、Moran's *I* 等の空間統計量を活用して集積の空間的な側面を捉えようとする分析も注目を集めているが、我が国における適用例はほぼ皆無である。本研究では、従来の空間的自己相関を検定する局所空間統計量のほかに、Ord and Getis (2012) によって提案された空間的異質性を検定する統計量や二変量間の空間的自己相関を検定する統計量を、我が国の工業統計調査の従業者数データに適用することで、日本の経済集積の探索的空間データ分析を市町村レベルで行う。

Key Words : agglomeration economy, exploratory spatial data analysis, spatial heterogeneity

1. 序論

経済集積の計量的な実証分析では、集積の程度を測る指標が用いられることが一般的であり、例えば特化係数や立地ジニ係数などの伝統的な指標が古くから使用されてきた。集積指標は、単に値の大小によって集積の度合いを比較考察するだけのみならず、生産分析において被説明変数として用いられる等集積分析において重要な役割を担っている。しかしながら従来の集積指標は、Arbia (2001) が指摘するように、地理的情報を全く考慮していないために空間的な側面からの集積を捉えることができないという欠点を持つ。即ち、従来の指標では地域ごとに集積の程度を測るに留まるため、集積が地域単位を越えた空間的な範囲で存在していても、それを抽出することは不可能である。ある程度地理的に広域な範囲で分析するには、ある産業に特化した地域が空間的に集中しているのか分散しているのかを把握することは非常に重要であると言えよう。ところで、集積を測る伝統的な指標として Ellison-Glaeser の集積指数も挙げられるが、個々の工場の規模分布データを必要とし、このような非集計データは分析者が容易に入手することができないのが現状である。Ellison and Glaeser (1997) も雇用者数のデータは地域単位であり、事業所の規模分布に関しても別のデータソースを用いて近似した値を使用している。した

がって本研究においては、地域単位で集計されたデータを基に集積を測る指標に焦点を当てる。

上記の問題を克服し集積の空間的側面に着目した方法論として、データの空間的自己相関を検出する空間統計量を用いて集積を空間的側面から捉える実証研究が、Arbia (2001) を契機にフランス (Guillain and Gallo, 2010) やドイツ (Rusche et al., 2011) などを対象に行われているが、我が国を対象とした分析は筆者らの知る限り皆無である。さらに、上記のような既往研究で用いられる局所空間統計量は、Moran's *I* 等代表的なものが一般的に用いられているが、近年、空間的異質性の検定や、二変量間の空間的自己相関の検定を行う局所空間統計量の開発がされている。そこで本研究では、我が国を対象として、近年開発された局所空間統計量を含む探索的空間データ分析を行うことで、我が国における空間的側面に着目した経済集積の実証研究の一端を担うとともに、局所空間統計量を用いた集積の分析の可能性を検討する。

2. 予備的分析

(1) 用いるデータ

本研究では、経済産業省による工業統計調査のデータ (2010年度) を対象に分析を行う。工業統計調査では製造業に属する全事業所を対象に、西暦末尾が0,3,5,8年は

表1 産業中分類番号と名称

産業中分類番号	産業名称
09	食料品製造業
10	飲料・たばこ・飼料製造業
11	繊維工業
12	木材・木製品製造業
13	家具・装備品製造業
14	パルプ・紙・紙加工品製造業
15	印刷・同関連業
16	化学工業
17	石油製品・石炭製品製造業
18	プラスチック製品製造業
19	ゴム製品製造業
20	なめし革・同製品・毛皮製造業
21	窯業・土石製品製造業
22	鋼鉄業
23	非鉄金属製造業
24	金属製品製造業
25	はん用機械器具製造業
26	生産用機械器具製造業
27	業務用機械器具製造業
28	電子部品・デバイス・電子回路製造業
29	電気機械器具製造業
30	情報通信機械器具製造業
31	輸送用機械器具製造業
32	その他の製造業

全数調査、それ以外では従業者4人以上の事業所を対象としており、したがって本データは全数調査を行ったものとなっている。なお、2011年以降は経済センサスの活動調査に含まれることとなっているが、産業別・市区町村別に集計され公開されているものとしては本データが現時点で最新のものである。

製造業内の産業分類は、表1に示されるように24業種に分類されており、各産業に対して産業中分類（二桁）番号がつけられている。

(2) 特化係数とモラン統計量

本研究では、従業者数のデータをそのまま用いるのではなく、以下に示す特化係数に加工した後分析を行う。

まず、産業*i*の地域*j*における対全国シェア比を

$$s_{ij} = \frac{e_{ij}}{\sum_j e_{ij}} \quad (1)$$

と定義する。ここで*e*は就業者数である。*s_{ij}*の値が高ければ、産業*i*は地域*j*に相対的に集中していると解釈することができる。さらにこれを、全国全産業の雇用者合計に対する地域*j*の雇用者数の割合で除することで、式(1)の地域シェアを全国の産業構成比で基準化した Kim (1995)の特化係数 (location quotient; LQ) となり、

$$LQ_{ij} = \frac{s_{ij}}{s_{*j}} \quad (2)$$

と表すことができる。ただし、

$$s_{*j} = \frac{\sum_i e_{ij}}{\sum_i \sum_j e_{ij}} \quad (3)$$

である。したがって、地域の産業構成比が全国と等しければLQは1をとり、LQが1より大きければ地域*j*の産業*i*は全国水準を上回る規模であると解釈できる。

この特化係数を基に、空間的自己相関を検出する古典的な統計量のひとつであるMoran's *I*を各産業ごとに求めた。

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j W_{ij} (x_i - \bar{x}) - (x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (4)$$

ここで*n*は地域数、 \bar{x} は特化係数の平均値、*x_i*と*x_j*はそれぞれ地域*i*と地域*j*の特化係数である。*W_{ij}*は地域*i, j*間の近接性を表す空間重み行列 **W** の要素であり、本研究では*i, j*間のユークリッド距離の逆数の二乗によって定義する。また、*S₀*は空間重み行列の全ての要素の和である。

図1は式(3)と式(4)で求めた産業ごとLQの市区町村間の平均値およびMoran's *I*統計量を示したレーダー図である(分類番号は表1を参照)。LQ平均値が最も高いのは木材・木製品製造業の0.746で、最も低いのは輸送用機械器具製造業の0.319となっている。中村(2008)

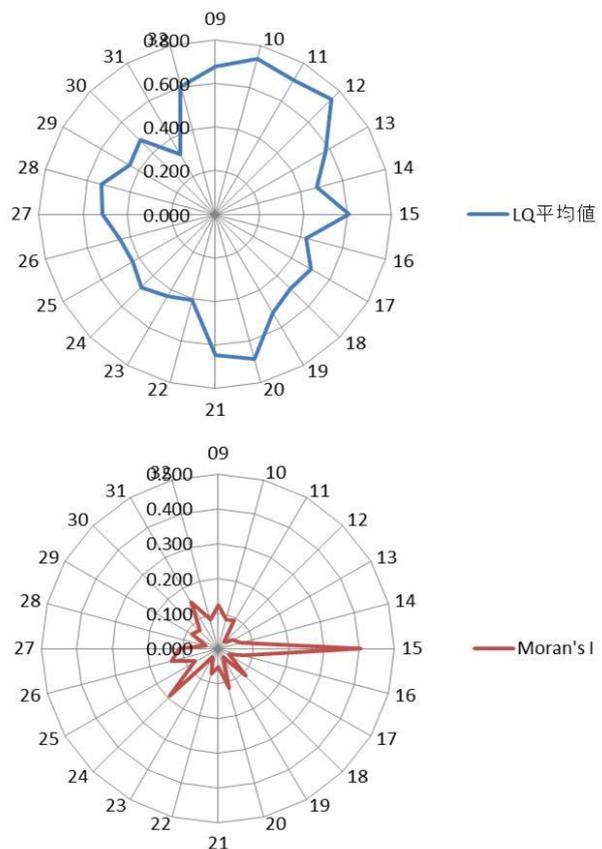


図1 LQ平均値及びMoran's I

では、この LQ 平均値は「伝統的な経済地理学の分野において、各産業の地域的特化度を示す指標」とされていることから、値の高い木材・木製品製造業などは相対的に地理的に集中していると見ることができる。しかし、前章で述べた様に LQ 値では境界を越えた地域の集中度を伺い知ることはできない。一方、Moran's I の値を見ると、印刷・同関連業が 0.406 と突出しており、そのほかの産業は概ね 0.1 程度となっている。いずれも空間的自己相関は無いという帰無仮説が 1% 水準で棄却されていることから、地域間の

3. 地理的集中と地理的特化地域の抽出

(1) Moran scatterplot と local Moran による分析

Anselin (1996) は、近隣地域との関係性を視覚化する方法として、Moran scatterplot を提案した。Moran scatterplot は、観測地域における標準化した値を x 軸、観測地域の周辺地域における平均値（標準化）を y 軸にプロットしたものであり、周辺地域の重みは空間重み行列によって定義される。

x 軸及び y 軸の値がともに標準化されているため、原点が平均産業規模水準となり、それ以上であれば大規模水準、それ以下であれば小規模水準と考えることができる。このことから Moran scatterplot は各々の軸によって分割された四つの象限によって、周辺地域との関係性を考慮したクラス分類が可能となる。

通常、Moran scatterplot を用いた分析を行った既往研究は主に第一象限（ホットスポット）や第三象限（クールスポット）といった周辺地域と類似性の高い地域の抽出に焦点が当てられる。しかし、集積の観点から見た場合、クールスポットは産業規模が周辺とともに小さい（またはそもそもその産業が存在しない）地域のクラスである。ここで重要となるのは、周辺地域とともに産業規模が大きい地域のクラス（地理的集中）と、周辺地域の産業規模は小さいが自地域は大きいクラス（地理的特化）である。

さらに、局所空間統計量（local indicator of spatial association; LISA）である local Moran's I (Anselin, 1995) を用いることで、各クラスに分類された市区町村の中から、統計学的に有意な市区町村を抽出することが可能となる。

$$I_i = \frac{n}{S_0} \frac{\sum_j W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (5)$$

表2はMoran scatterplotによって分類したクラスに含まれる市区町村の割合と、LISAによる検定で1%有意となった市区町村の割合を表している。HH, LH, LL, HL はそれぞれ順に第一象限から第四象限に対応している。まず Moran scatterplot でのクラス分類の結果に着目してみ

表2 各分類クラスに占める市区町村の割合 (%)

分類 番号	Moran scatterplot				LISA			
	HH	LH	LL	HL	HH	LH	LL	HL
09	18.67	14.63	46.12	20.59	9.47	0.00	10.16	0.00
10	8.56	12.87	58.56	20.00	4.73	0.00	11.86	0.00
11	15.69	8.67	52.29	23.35	9.10	0.00	19.47	0.00
12	12.02	9.04	51.06	27.87	4.95	0.00	16.38	0.00
13	14.31	9.26	50.11	26.33	7.66	0.00	10.00	0.00
14	16.49	9.04	51.22	23.24	10.59	0.00	15.32	0.00
15	12.93	9.15	65.74	12.18	10.27	0.05	16.06	0.00
16	15.53	8.14	52.02	24.31	9.41	0.00	15.21	0.00
17	6.70	7.61	63.56	22.13	3.99	0.00	11.22	0.00
18	21.91	9.36	44.63	24.10	12.34	0.00	13.09	0.00
19	10.80	5.32	57.66	26.22	6.44	0.00	16.91	0.00
20	7.02	4.20	70.00	18.78	5.27	0.11	15.05	0.00
21	15.00	13.09	48.24	23.67	5.43	0.00	13.03	0.00
22	10.16	7.71	60.90	21.22	6.60	0.00	19.15	0.00
23	12.61	6.01	50.27	31.12	7.07	0.00	17.18	0.00
24	22.66	12.87	46.01	18.46	13.78	0.05	12.18	0.00
25	17.61	7.07	50.37	24.95	11.49	0.00	19.36	0.00
26	22.23	9.84	46.38	21.54	13.72	0.00	13.67	0.00
27	15.48	7.45	53.88	23.19	10.69	0.00	25.53	0.00
28	12.55	8.72	49.63	29.10	6.91	0.00	12.77	0.00
29	17.55	9.68	48.51	24.26	10.21	0.00	13.40	0.00
30	10.53	4.95	56.44	28.09	7.61	0.05	27.98	0.00
31	15.37	6.76	55.37	22.50	9.63	0.00	19.52	0.00
32	15.53	9.47	52.77	22.23	10.59	0.00	14.41	0.00

ると、全ての産業において最も市区町村の割合が多いのが LL クラスとなっており、周辺地域とともに産業規模の大きい地域とされる HH クラスにおける割合はどの産業でも小さい。

local Moran による有意な市区町村の割合を見ると、ほぼ HH と LL のみとなっていることが分かる。本研究において HH クラスと共に興味のある、地理的特化を表す HL クラスは、全ての産業で有意な市区町村が検出できないという結果となっているが、これは local Moran が空間的自己相関を検出するのに対し、HL や LH は周辺地域と比べて空間的に異質であることが原因であると考えられる。

図2は、印刷・同関連業を例に、各市区町村を Moran scatterplot によって分類したクラスごとに色で分けその地理的分布を表したものと、local Moran による有意な地域の分布を表している。HH クラスに属する地域は東京特別区と札幌・仙台・新潟・名古屋・大阪・広島・福岡といった政令指定都市であることが分かる。また、図2下図からは、これらの都市は local Moran での検定でも有意であり、印刷・同関連業は政令指定都市を中心とした地域において地理的に集中したクラスターを形成していることが示唆された。

(2) G と H 統計量による分析

Local Moran よりも以前に提案された LISA として、次式で与えられる Getis's G^* 統計量 (Getis and Ord, 1992) が有名である。

$$\tilde{x}_j = \frac{\sum_j w_{ij} x_j}{\sum_j w_{ij}} \quad (8)$$

と定義し、周辺地域における局所的な残差を

$$u_j = x_j - \tilde{x}_j \quad (9)$$

と定義したとき、 H 統計量は

$$H_i = \frac{\sum_j w_{ij} |u_j|^\alpha}{\sum_j w_{ij}} \quad (10)$$

で与えられる。 $\alpha = 1$ ならば絶対値測度であり、また $\alpha = 2$ ならば分散測度となる。

この H 統計量と G^* 統計量の絶対値の大小の関係性によって、対象を四つの特性に分類することができる (Ord and Getis, 2012を参照)。特に、 $|G^*|$ が大きく H_i も大きい場合は、ホットスポットであり且つ異質性が大きな地域となるため、空間的な外れ値をもつ地域を抽出することが可能である。Ord and Getis は上記二つの統計量の大小の組み合わせでの解釈が可能、とだけ述べそれ以上の議論を行っていないが、本研究はこれをモラン散布図と同様に、 $|G^*|$ が大きく H_i をそれぞれ標準化したものを縦軸と横軸にとりプロットし、その象限でクラス分類を行う。すなわち、第一象限はホットスポット・異質性大、第二象限はホットスポットであり周辺と類似した傾向、というように各クラスの特徴は Ord and Getis (2012) に合わせて分類する。

表3は上述の方法によって分類したクラスに含まれる市区町村の割合と、 G^* と H 統計量両方による検定で1%有意となった市区町村の割合を表している。ただし全ての産業において、1%有意であったのは第一象限のクラスに属する地区町村のみであったため、その他の象限に含まれる市区町村の割合の掲載はしない。

まずクラス分類の結果に着目してみると、全ての産業において最も市区町村の割合が多いのは第三象限クラスとなっており、これは Moran scatterplot でのクラス分類の結果と同様である。周辺地域と比べて産業規模の大きい地域とされる第一象限クラスにおける割合は、はん用機械器具製造業が21.12%と最も高く、パルプ・紙・紙加工品製造業が8.88%と最も低い。

G^* と H 統計量による有意な市区町村は、上で述べたとおり第一象限クラスでのみ検出されたが、ここでははん用機械器具製造業を抑えて印刷・同関連業が10.74%と最も高い。どの産業でも有意となった市区町村の割合が低いのは、 G^* と H 統計量両方で1%有意となった市区町村のみが抽出されるためである。

図3は、前節と同じく印刷・同関連業を例にとり、クラス分類と有意な市区町村の空間的な分布を表したもの

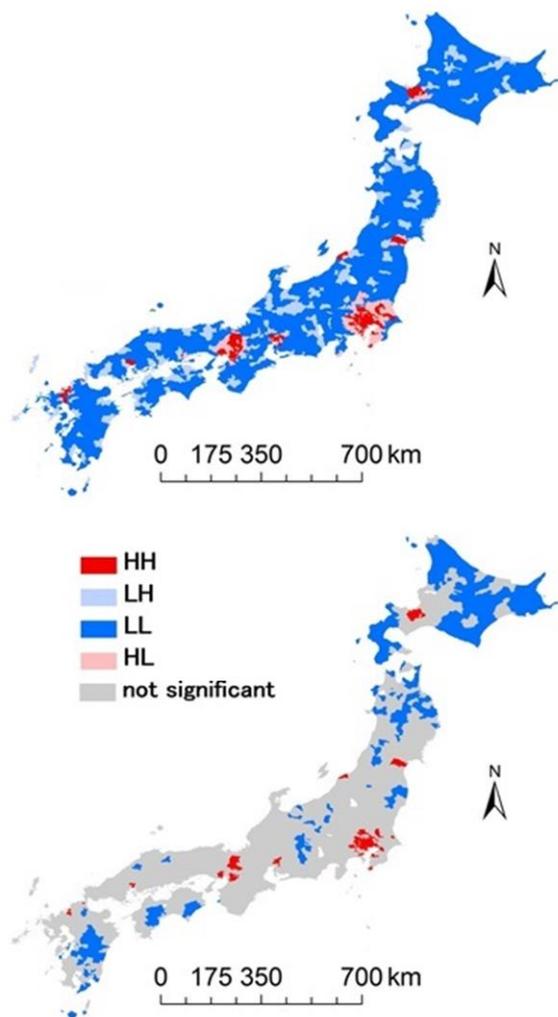


図2 Moran scatterplot によるクラス分類 (上) と有意な市区町村 (下) の空間分布

$$G_i^* = \frac{\sum_j w_{ij} x_j - W_i^* \bar{x}}{s\{[(nS_{ii}^*) - W_i^{*2}]/(n-1)\}^{1/2}} \quad (6)$$

ただし、

$$W_i^* = W_i + W_{ii}, \quad S_{ii}^* = \sum_{j=1}^n W_{ij}^2. \quad (7)$$

また、 s は x の標準偏差であり、 w_{ij} は式 (4) や式 (5) 等の通常の空間重み行列の要素とは異なり、自地域の重み $w_{ii} \neq 0$ である。

G^* 統計量は式 (6) における分子に示される通り、自地域と周辺地域の値の加重和によって表されるため、周辺地域との偏差を見る Moran 統計量のように周辺地域との関係性を表現するものではない。

近年、彼らによって、空間的な異質性を検定する H 統計量が開発された (Ord and Getis, 2012)。 \tilde{x} を

表3 G*とH統計量による分類クラスに占める

市区町村の割合 (%)

分類番号	クラス分類				G and H
	1	2	3	4	1
09	15.00	26.17	39.73	19.10	4.84
10	11.38	22.82	56.38	9.41	3.72
11	15.05	24.47	45.27	15.21	6.44
12	13.72	27.77	42.18	16.33	3.24
13	10.32	25.21	54.84	9.63	2.71
14	8.88	26.01	54.10	11.01	2.45
15	16.44	3.24	74.36	5.96	10.74
16	16.17	19.79	51.65	12.39	3.67
17	11.12	18.09	62.13	8.67	3.78
18	16.17	22.34	43.24	18.24	5.90
19	11.33	22.13	55.59	10.96	3.94
20	11.54	9.36	72.82	6.28	4.63
21	13.62	24.31	47.18	14.89	3.99
22	14.15	17.39	61.22	7.23	4.15
23	17.34	19.41	50.21	13.03	3.83
24	17.93	17.02	49.26	15.80	6.60
25	21.12	18.72	48.56	11.60	4.84
26	19.41	20.74	45.05	14.79	7.55
27	18.78	13.83	55.90	11.49	5.96
28	15.80	23.46	43.56	17.18	4.26
29	16.38	17.50	50.16	15.96	5.64
30	17.55	11.01	57.02	14.41	6.33
31	18.30	14.89	52.45	14.36	6.86
32	15.05	18.19	57.39	9.36	5.16

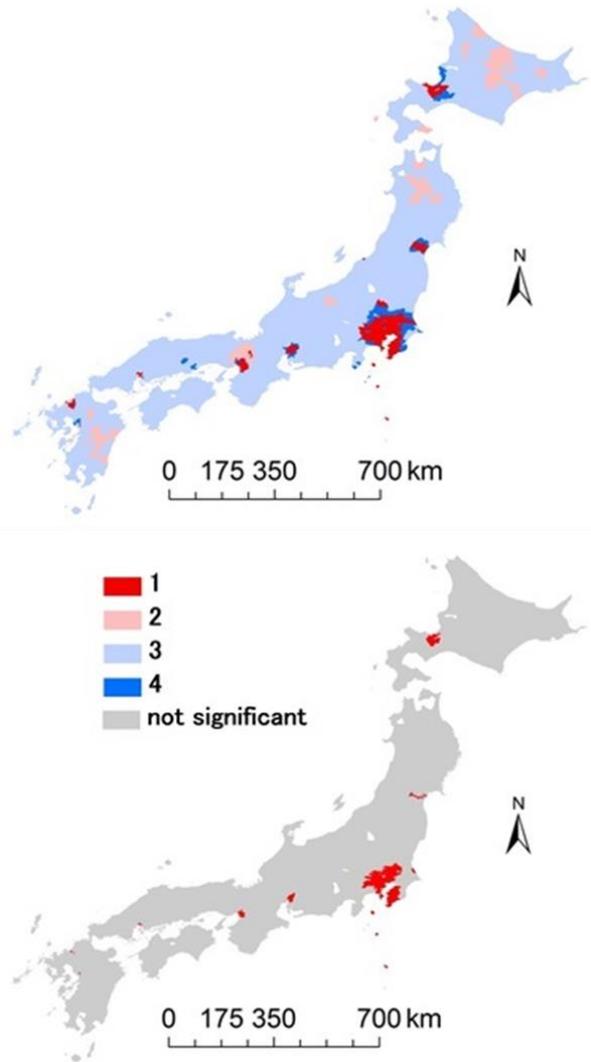


図3 G*とH統計量によるクラス分類（上）と有意な市区町村（下）の空間分布

である。第一象限クラスに属している市区町村は、Moran scatterplot による分類における HH クラスの空間分布と類似していることが見てとれる。G* と H 統計量による有意な市区町村もまた、local Moran での結果と類似しているように見えるが、その数は local Moran の時よりも少ない。この結果から、印刷・同関連業は東京都区部と政令指定都市において地理的に特化かつ集中して立地していると推察することができる。

4. 共集積地域の抽出

第3章で見てきたのは、あるひとつの産業に着目したときのその産業内での集積であった。しかし集積は、同業種間の集積だけではなく、異なる業種との間においても存在しうる。このような経済の多様性もまた重要なトピックのひとつである。

Rushce et al. (2011)では、このような産業間の共集積の分析として、bivariate local Moran（または local cross-Moran）統計量を用いている。bivariate local Moran は、産業 k の特化係数を x、産業 l の特化係数を y とすると、

$$bI_i^{kl} = \frac{n}{S_0} \frac{\sum_j W_{ij} (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (11)$$

として表せられ、二変量 x, y 間の空間的自己相関を検定することができる。しかしながら、式 () の分子が示すように、 $bI_i^{kl} \neq bI_i^{lk}$ であるため、産業 k, l 双方の関係を測るには二つの統計量を算出する必要がある。

一方、Moran統計量とピアソンの相関係数を組み合わせて導出されたLee (2001)によるL統計量のLISAである local Lee 統計量 (Lee, 2009) は次式によって表される。

$$L_i = \frac{n}{S_{00}} \frac{\left[\sum_j W_{ij} (x_j - \bar{x}) \right] \left[\sum_j W_{ij} (y_j - \bar{y}) \right]}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (12)$$

ただし、

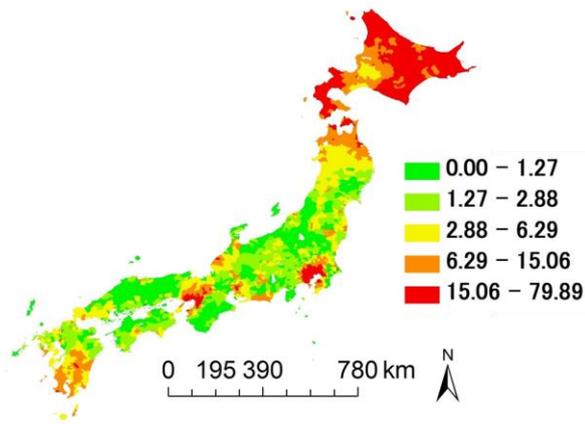


図4 有意なlocal Leeの合計値の空間分布

$$S_{00} = \sum_i \left(\sum_j W_{ij} \right)^2 \quad (13)$$

local Lee 統計量はピアソン相関係数から導出されたことから、bivariate local Moran と異なり x と y の順番を問わないという利点がある。本研究では、この local Lee 統計量を各産業間について求めた。

図4は、各産業間のlocal Lee の内、5%有意であった値のみに関して市区町村ごとに総和をとった値の空間的な分布を表している。東京都区部と大阪市周辺で高い値を示しているほか、九州南部も比較的に値が高い傾向にある。また、特に北海道に関してはほぼ全域にわたって高い傾向を示しており、北海道における製造業の多様性の高さが図4から示唆された。

5. 結論

本研究は、局所空間統計量を用いることで、従来の集積の程度を測る指標では捉えることのできなかった地理的集中および地理的特化地域の抽出を目的とした探索的空間データ分析を行った。Local Moran を用いることで地理的に集中している地域を抽出することは海外の既往研究で行われてきたが、空間的自己相関を検定する統計量 (local Moran) では、地理的に特化した地域を抽出することは困難であった。本研究では、空間的異質性を検定する H 統計量を用いることで、local Moran では検出できなかった地理的特化地域を抽出することが可能であることが示唆された。

また、産業間の集積を捉えることを目的に、二変量間の空間的自己相関を検定する local Lee 統計量を用いた分析も併せて行った。bivariate local Moran と異なり二変量の順番を問わないため利便性が高いが、産業分類が多いな

かでの効率的な視覚化は今後の課題とされる。また、同業種内での分析と同様に、産業間の地理的集中のみならず二変量間の空間的異質性の検定も検討していきたい。

謝辞

本研究は JSPS 特別研究員奨励費 251786 の助成を受けたものである。

参考文献

- 1) Arbia, G.: The role of spatial effects in the empirical analysis of regional concentration, *Journal of Geographical Systems*, Vol.3, pp.271-281, 2001.
- 2) Ellison, G. and Glaeser, E. I.: Geographic concentration in US manufacturing industries: a dartboard approach, *Journal of Political Economy*, Vol.105, pp.889-927, 1997.
- 3) Guillain, R. and Gallo, J. L.: Agglomeration and dispersion of economic activities in and around Paris: an exploratory spatial data analysis, *Environment and Planning B*, Vol.37, pp.961-981, 2010.
- 4) Rusche, K., Kies, U. and Schulte, A.: Measuring spatial co-agglomeration patterns by extending ESDA techniques, *Jahrbuch fur Regionalwissenschaft*, Vol.31, pp.11-25, 2011.
- 5) Kim, S.: Expansion of markets and the geographic distribution of economic activities: the trends in U.S. regional manufacturing structure, 1860-1987, *the Quarterly Journal of Economics*, Vol.110, No.4, pp.881-908, 1995.
- 6) 中村良平：都市・地域における経済集積の測度（上），岡山大学経済学会雑誌，Vol.39，No.4，pp.99-121，2008.
- 7) Anselin, L.: The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In Fischer, M.M., Scholten H.J. and Unwin D. (eds.), *Spatial analytical perspectives on GIS*. Taylor and Francis, London, pp.111-125, 1996.
- 8) Anselin, L.: Local indicators of spatial association – LISA, *Geographical Analysis*, Vol.27, No.2, pp.93-115, 1995.
- 9) Getis, A. and Ord, J.K.: The analysis of spatial association by use of distance statistics, *Geographical Analysis*, Vol. 24, No.3, pp.189-206, 1992.
- 10) Ord, J.K. and Getis, A.: Local spatial heteroscedasticity (LOSH), *The Annals of Regional Science*, Vol.48, No.s, 529-539, 2012.
- 11) Lee, S.: Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's I , *Journal of Geographical Systems*, Vol.3, pp.369-385, 2001.
- 12) Lee, S.: A generalized randomization approach to local measures of spatial association, *Geographical Analysis*, Vol.41, pp.221-248, 2009.

(2013.5.7 受付)