

機関番号：14301
 研究種目：若手研究（B）
 研究期間：2008～2010
 課題番号：20780172
 研究課題名（和文） 空間データマイニングによるソーシャルキャピタルの局地的差異と
 要因分析
 研究課題名（英文） Factor analysis on spatial and regional variation of social
 capital by the spatial data mining methods
 研究代表者
 西前 出（SAIZEN IZURU）
 京都大学・地球環境学堂・准教授
 研究者番号：80346098

研究成果の概要（和文）：本研究は、ソーシャルキャピタルの尺度として提唱されている ISCI（Integrated Social Capital Index）の得点の高い高 SC 層の空間分布特性を空間データマイニング手法を用いて明らかにした。大規模なアンケートデータを元に京都府 A 市の 151 集落ごとの ISCI から高 SC 層の割合（P6）を算出し、回答者数の偏りによって生じる影響を経験ベイズ法によって補正して空間分析を行った。経験ベイズモランの I 統計量を用いて、定量的に P6 の空間分布特性を抽出し、集落単位の P6 には正の空間的自己相関が存在することを明らかにした。

研究成果の概要（英文）：This study aims to clarify the characteristics of spatial distribution of the percentage of the person who have high score of ISCI (Integrated Social Capital Index), which has been propounded as the measurement of Social Capital. We have calculated the percentage (P6) values of 151 settlements based on the database collected by the questionnaire survey in A city, Kyoto prefecture and adjust the undesirable effects caused by bias of the number of samples using Empirical Bayes approach. It has been made clear that the positive spatial auto correlation of P6's spatial distribution is detected by Empirical Bayes Index.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008 年度	2,200,000	660,000	2,860,000
2009 年度	500,000	150,000	650,000
2010 年度	500,000	150,000	650,000
年度			
年度			
総計	3,200,000	960,000	4,160,000

研究代表者の専門分野：地域計画学

科研費の分科・細目：農業工学・農業土木学／農村計画学

キーワード：ソーシャルキャピタル，空間データマイニング，空間分析

1. 研究開始当初の背景

日本の農山村は、高度経済成長と共に急速に進んだ地域の過疎・高齢化による地域活力の低下・衰退という深刻な社会問題を抱えている。この問題解決に取り組まれてきた方策は、ハード面の整備が主流であったが、期待に沿った効

果は挙げられなかった。現在はソフト面での方策に重点が置かれつつあるが、地域によってその効果の違いがあることが指摘されている。このことから近年「ソーシャルキャピタル(Social Capital, 以下 SC)」という概念に注目が集められている。これは分かりやすい日本語で言い換えるならば「地域の底力」と表現することもで

き、「人々の協調行動を活発にして、社会の効率性を高めることのできる社会特性」と定義できる。たとえば、「ある地域に強いリーダーシップを取る人間がいるのか」、「公民館等の既往施設が有効利用できるか、あるいははされているか」、「自治会の結束力は強いのか」、「地理的な条件（道路整備・都市への距離等）はどうか」、といった要素が、住民のSCを形成させ、ある方策が十分に機能するかどうかに影響を与えると示唆されている。こうした地域ごとの無形の力をSCと言い、これを定量化し、考慮することで、低コスト高パフォーマンスの農村振興の方策をより効果的・戦略的に決定することが期待されている。

研究開始当初は、SCに関する研究は発展途上にあり、SCを定量化する方法、底上げする方法、SCを高める方法などの研究が農村計画分野の最前線にあった。いずれの場合もSCを高める（あるいは低くする）要因を明確する必要があるが、SCは無形で、ぼんやりとした概念であるため既存のデータとの関連性がわかりにくい。こうしたはっきりとしないものの要因を抽出する方法として空間データマイニングがある。空間データマイニングは、空間データの局地的な特性・関連を様々な手法で把握する方法であり、以前には知られていない（想定できない）潜在的に有用な情報をデータベースから引き出す方法である。

本研究では、SCの局地的な差異がどのように形成されたのか、そしてどういった要素が影響を与えているのか、あるいは関連があるのかを解明することを空間データマイニングでとらえることを試みている。分析単位としての「空間」単位は集落・旧村・市町村、個人、あるいはデータをカテゴリー化することで得られる境界など様々なスケールで想定することができる。このことによりSCの変動とこれと関連する要因を把握することで、より効率的かつ戦略的な方策を打ち立てる一助となると考えられる。

この分析にあたり、膨大な量の空間情報を必要とするが、そのプラットフォームは近年、かなりの程度整備されている。平成19年に「地理空間情報活用推進基本法」が国会で成立した。この基本法では、地理空間情報を「場所や位置に結び付けられた情報全て」と定義しており、これまでの「地図」情報というイメージを一新している。これまでも地理情報システム（Geographic Information System、以下GIS）のアクションプログラムは存在し、国レベルでの整合性のとれたGIS利用・普及活動は行われてきたが、法律の裏付けができたことになる。地方公共団体は国や関連する民間事業者と協力して「どこでも誰でも位置・場所の容易にわかる環境」を構築する努力を行うことが義務となっている。農業・農村分野においても、この基本法を待たずして、農業振興地域のデジタルオルソ画像や日本水土図鑑GIS（農地情報整備促

進事業）に挙げられるように多くの地理空間情報が整備されてきた。しかしながら、農村部では、これらの空間データ整備に比べて、その利用方法、農村・地域計画への適用方法についてはいまだ発展途上であり、多くの問題点が残されている。

2. 研究の目的

本研究において明らかにすべき点は、①わが国の都市・農村の形成過程を考慮し、地域の無形の力であるSCを定量化し、その要因を解明していく。その際に一つのSCを形成する地理空間を規定する。これは様々な空間単位があるという仮定に基づき研究を進めていくが、その場合の如何に関わらず空間単位について科学的に検証示す。②導き出した空間単位でのSCをその規定要因に着目して定量的に算出する。空間データマイニングによる要因として考えられる空間情報の抽出および、地域の特性を考慮した空間データマイニングの際の新しい数学的手法の開発（データ抽出方法）。④SCの規定要因を考慮したその地域に活力を与えるための施策の提案と空間情報を最大限に利用できる結果表示システムの構築。といった点に集約される。

これらにより、最終的には対象地域におけるSC形成のメカニズム、およびその地域の特性に配慮した地域資源計画を空間的に提示することが可能となると考えられる。

3. 研究の方法

(1) アンケート調査の概要

本研究は、ソーシャルキャピタルの定量化、および空間データマイニングの手法論を、それぞれ、同時並行で実施した。前者は、既に大規模なアンケート調査を実施していた京都府北部のA市におけるSCデータを利用している。集落数は151であり、4,519票のデータを用いた。

(2) SCの定量化指標

ソーシャルキャピタルの定量化指標として、ISCI（Integrated Social Capital Index）を使用した。ISCIは6つの質問項目から構成されている。この質問項目のうち、「財布が返る確率」と「地域内信頼感」はイギリスの統計局のGHS（General Household Survey）、「争い解決力」と「地域貢献への態度」は世界銀行の作成したSOCAT（Social Capital Assessment Tool）、「つきあいの程度」と「つきあいの人数」は内閣府国民生活局が平成15年と17年に行った調査票を参考にして作成したものである。また、「地域内信頼感」「財布が返る確率」は信頼に、「地域貢献への態度」は規範に、「付きあいの程度」「つきあいの人数」はネットワークに対応しており、日本国内で一般性のある尺度を作成する目的で提唱されている。

SC の尺度 ISCI は 6 つの質問項目の比率の単純合計であり、地域 i の ISCI の値を $ISCI_i$ と記述し、(1)式で定義される。

$$ISCI_i = \frac{r1_i + r2_i + r3_i + r4_i + r5_i + r6_i}{m_i} \quad (1)$$

ただし、

- i : 地域の番号
- $r1_i \sim r6_i$: 地域 i における表 1 に示す変数
- m_i : 地域 i の回答者数

なお、ISCI は SC の尺度であるから、この値が大きくなると SC が高いと評価できる。6 つの質問項目の比率の和であるから、0 から 6 までの値を取り、小集団でも SC の尺度として評価することができる。本研究では、(2)式に示すように特に個人の ISCI が最大の 6 となる SC の高い高 SC 層の割合を $P6$ と表し、(2)式のように定義した。

$$P6_i = l6_i / m_i \quad (2)$$

ただし、

- $P6_i$: 地域 i における ISCI の値が 6 となった人の割合
- $l6_i$: 地域 i における ISCI の値が 6 となった人数

農山村では、リーダーの不在による地域活力の低下が叫ばれており、高 SC 中心層がリーダーシップを取ることで地域活力が高まる可能性が指摘されている。本研究では個人の ISCI の値が 6 となる高 SC 層の割合の空間分布特性を探ることとした。

なお、 $P6$ の値と集落数との関係については、最小値は 0.0、最大値は 0.8 となる。0.0 から 0.1 の区間の値を取る集落が最大で 56 集落あり、次いで 0.1 から 0.2 の値を取る集落が 50 集落となった。

地理空間上に分布する変数の空間的な相関を計る代表的な指標としてモランの I が挙げられる。この指標により地理空間上に分布する変数の空間パターンのクラスタリングの程度を定量的に検討できる。モランの I は、(3)式で表される。

$$I = \frac{n}{\sum_{ij} w_{ij}} \frac{\sum_{ij} w_{ij} (P6_i - \overline{P6})(P6_j - \overline{P6})}{\sum_i (P6_i - \overline{P6})^2} \quad (3)$$

ただし、

- n : 分析する地域の数
- w_{ij} : 地域 i と j の距離によって定義される重み
- $\overline{P6}$: 対象地域での $P6$ の平均値

モランの I では、空間的自己相関がなければ、この期待値は小さな負の値 $-1/(n-1)$ となる。モランの I がこの期待値より大きければ正の空間的

空間的自己相関があることを、小さければ負の空間的自己相関が存在することを意味する。このモランの I は、対象地域における全体的な空間的自己相関の程度を表す指標である。

(3) 経験ベイズのモランの I の応用

本研究では、集落といった小さな空間単位を分析単位として用いているため、各集落の回答者の数にはある一定量の偏りがある。回答者数が 10 人以下の集落も 40 集落あり、200 人以上の回答のあった集落も 2 集落あることがわかる。一般に人口規模が小さい地域ほどそれに伴う地域内の回答者も減少し、このため空間的な特性を抽出する場合に好ましくない影響を与えてしまう可能性がある。本ケースに限らず、地域分析では地域ごとのサンプル数に偏りがあるとすると、サンプル数の少ない地域の数字・変動が全体の結果に実際よりも大きな影響を与えるという問題を引き起こすことがある。こうした地域は分析対象から外すことが一つの解決であるが、一方で、過疎地域などでは、地域の住民自体の数が少なく、回答者数が少なくてもそれが実態を反映している場合もあり、統計上の処理を理由に分析対象から排除すると地域の実情を逆に無視する恐れも生じる。さらには空間を連続の面として捉えるためには、サンプル数が少ない地域だとしても対象地域から外すことは極力避けなければならない。

したがって、このような場合には統計的な分析のために偏りを平滑化する処理が必要となる。また、分散が不均一であるデータにモランの I をそのまま適用すると、空間的自己相関の検出力が低下することが知られており、回答者数の少ない集落で集計された $P6$ が空間的自己相関の傾向を乱すことも考えられる。

以上のような知見から、本研究では、上述のようなすべての要請を満たすことのできる経験ベイズモランの I (Empirical Bayes Moran's I , 以下 EBI) 統計量を用いた。EBI は(4)式のように定義される。

$$EBI = \frac{n}{\sum w_{ij}} \frac{\sum w_{ij} z_i z_j}{\sum (z_i - \bar{z})^2} \quad (4)$$

ここで、 z_i は i 地区の標準化得点と呼ばれ、(5)式のように定義される。

$$z_i = \frac{6_i - \beta/\alpha}{\sqrt{\beta/\alpha^2 + \beta/(cm_i)}} \quad (5)$$

経験ベイズ法により補正された $P6$ を $P6(EB)$ と表記すると α 、 β は、その補正式となる (6)式で定義される

$$P6(EB)_i = \frac{16_i + \beta}{m_i + \alpha} \quad (6)$$

この補正式は、ベイズ統計学の枠組みの中で示されたものであり、 α 、 β は正のパラメータでデータより経験ベイズ推定法により算出される。この定数を用いることで、地域の6つの質問項目の数、および回答者数が α 、 β よりも十分に大きい場合には α 、 β の補正はほとんど受けることはなく、逆に回答者数が少なければ(6)式の α 、 β 以外の数値は相対的に小さくなり、補正値は α/β で示される水準に近づくこととなる。つまり、集落内の回答者数が少なく比率の信頼性が乏しいほど、平均的な比率に近づくようにP6の値を補正することができ、また、可能な限り実データの情報も損なわないで分析に用いることができるようになる。したがって、EBIを用いることで、回答者数の少ない集落の情報も考慮して、かつ統計上の不安定さを排除することができ、対象地域全体のP6の空間分布特性を表すことが可能となる。

(4) 経験ベイズ推定法の解釈

本研究の場合、この経験ベイズ推定法の考え方は以下のように解釈することができる。対象地域には「P6の真の値」が存在する。しかし、実際の集計値には、「P6の真の値の地域差」と「アンケート調査によって生じるP6の確率的な誤差」の2つの成分があると仮定できる。この仮定を基にして、それぞれの成分の確率分布にガンマ分布とポアソン分布を適用してベイズの定理から得られる「各集落の真のP6の値」を予測する。また、P6の確率分布にポアソン分布を近似する上で χ^2 乗検定を行ったところ、5%有意で近似が可能であった。以上の解釈により予測式(5)式が得られた。

なお、EBIの値の解釈は、先述したモランのIによる空間的自己相関の検出と同様であり、空間的自己相関がなければ、この期待値は小さな負の値 $-1/(n-1)$ となり、これより大きければ正の空間的自己相関が、小さければ負の空間的自己相関が存在することを示す。

4. 研究成果

本研究では、アンケート調査で得た、対象地域のP6について回答者数の少ない地域の値を補正してEBIを算出した。重み w については、対象とする地域から近い順に集落を選択する方法である最近隣による近傍地区を用いた。対象地域の集落界の形状が大小様々であり均一でないことから、空間特性を把握するために1から10個の数の隣接集落を取ることとした。隣接集落を10集落とった場合、図1に示すような部分集合を取ることとなる。集落界ごとに重心をとり、10集落の重心が含まれるまで中心集落からの同心円を広げていき、部分集合を作成してい

る。近傍集落数がこれより少ない場合も同様である。(6)式で示される経験ベイズ法によるP6の補正により、空間的な平滑化をおこなったP6(EB)の値を図2に示す。なお、地図表示による地域の類推を極力避けるため、P6を変数として比例させて表示した円カルトグラムを採用している。回答者数の少ない集落でのP6の値は経験ベイズ法により補正されており、バイアスは可能な限り取り除かれている。しかしながら、南西部に、P6(EB)の値の低い集落が集まっているものの、中心部や東部にも点在し、視覚だけでは明確な空間的な特徴は見出せないことがわかる。地図表示だけで空間パターンの検出を判断すると恣意的な要素が加わるため、空間的自己相関の分布傾向を定量的かつ客観的に把握するための方法としてモラン散布図を適用した。経験ベイズ法により修正されたP6(EB)の値をモラン散布図と同様の方法で表示した経験ベイズモラン散布図を図3に示す。なお、図3の近傍集落数は10の場合である。横軸にはP6(EB)を標準化した値を、縦軸には10次の最近隣で選択した近傍10集落のP6(EB)の平均値を記してある。各集落と近隣集落のP6(EB)の空間的な相関をこの散布図の4つの象限をみることで客観的に検証できる。第I象限は、P6(EB)が正であり、かつ近隣の10集落のP6(EB)の平均値も同様に正となることを意味し、第III象限ではそれぞれがいずれも負の値をとることを意味する。つまり、P6(EB)が同じ傾向を示す集落が連続していることとなり、正の空間的自己相関があることを示す。逆に第II、IV象限では周辺のP6(EB)の正負が逆となり、負の空間的自己相関が存在することを示す。図3に示された回帰直線の傾きはEBIに相当する。EBIの値は、0.213を示し、正の空間的自己相関があると推定できる。

空間的自己相関は、実際には相関は存在しないのに、たまたま偶発的に地図上の配置されたことにより観測される場合がある。つまり、集落のP6には実際には空間的な相関は存在しないのに、相関があるような配置が偶然にもなされていたという可能性がある。そこで、モンテカルロシミュレーションにより検定を行った。



図1 10次の最近隣による近傍の選択

モンテカルロシミュレーションによりランダムなデータセット、つまりランダムな地域の入替えを 1,000 回作成し、その各回について EBI を測定した。観測値以上の EBI 値が得られた確率 p 値は 0.1% となり、この P6 (EB) の分布には両側 1% 水準で有意な正の空間的自己相関が存在することが確認できた。

補正を行わなかった場合と比較するために、経験ベイズ法による補正を行わないで通常のコランの I を算出した(図 4)。モランの I の値は、0.193 (p 値 0.1%) となり、EBI に比べて小さな値となった。近傍を 1 から 10 集落まで変化させ、それぞれの値を表 1 に示した。いずれの場合も EBI の値が高い。また、別途集計したところ、モランの I ではどの近隣集落数の値の場合でも第 II・IV 象限にある集落のうちのおよそ 90% は、回答者の数が 10 人以下となっており、このサンプル数の少なさが空間的自己相関の値を乱しており、経験的ベイズ法を使うことが有効である可能性が示唆された。

本研究は、SC の地域的な差異がどのように形成されたのか、そしてどういった要素が SC を醸成させるのか、あるいは関連があるのかといった問題を解明する上で、高 SC 層の割合 P6 に着目し、その空間分布特性を抽出することを試みたものである。

対象地域でのアンケート調査を元に集計した P6 を集落界ごとに振り分けた GIS データを構築し、サンプル数の少ない地域の変動を補正して最小限に抑えることのできる経験ベイズモランの I 統計量を用いて分析を行った。その結果、集落ごとの P6 の明確な正の空間的自己相関を確認することが出来た。

一般に、個々の空間的事物は、空間座標とその属性によって捉えることができるが、多くの場合は事物間の位置関係に対応した規則的な影響を受けあって共変動関係を持つ。しかしながら、SC は、個人や集団の内面によって異なるものであり、その尺度としての ISCI が高い P6 についてそれぞれ空間的な影響や関係を集落間で持つと考えるのは妥当ではなく、歴史的な経緯で醸成されてきたと考えられる SC が、集落の枠

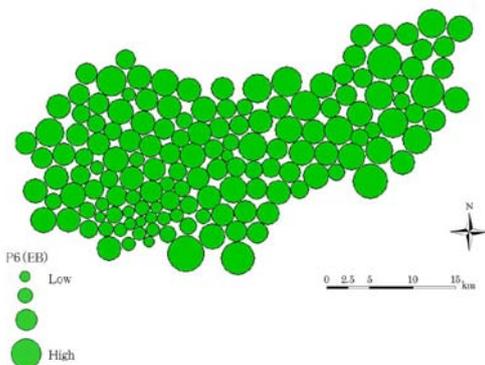


図 2 経験ベイズ法により補正された P6 の分布

組みを超えた空間単位で醸成されてきたため空間的自己相関が検出されたと考えられる。

近隣の集落では P6 の連続性が確認できたため、集落ごとに SC を考慮するよりも、施策の性質によっては、旧村や市町村などの、より大きな単位に同一性のある SC が存在するという空間特性を考慮すべき場合もあると示唆できる。また、ISCI は、SC を計る一つの尺度ではあるが、更なる検証も必要であり、必ずしもそれが唯一のものではない。しかし、ISCI 以外の指標を用いたとしても、人口の減りつつある日本の農山村で SC を計る様な場合にはサンプル数が極端に少なくなることが多々あり、統計上の不安定さが増すようなケースは少なくない。本研究で用いた経験ベイズの推定法によりこうした問題は解決が可能であることも実証的に示された。

手法上の制約から、ISCI の値を 6 とそれ未満と区分けしたが、このことが地域活性化に役立つ地域リーダーの要件である高 SC 層とは必ずしも定義できず、更なる検証が必要となる。しかしながら、高 SC 層の分布特性を掌握することは SC を醸成する方策を練るためにも肝要であり、本研究の成果は今後の SC 研究において一助となると考える。

今後は、本論で得られた空間特性を基にして SC の醸成、高めるための方法、具体的な政策パフォーマンスへの応用などの具体的な研究も必

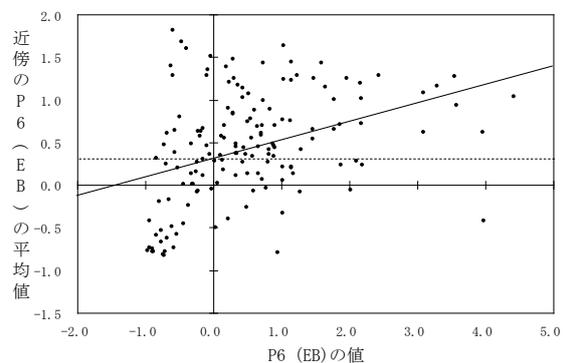


図 3 経験ベイズモランの I の散布図(近傍 10 集落)

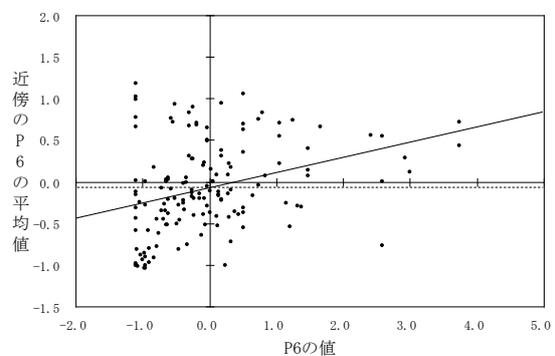


図 4 モランの I の散布図 (近傍 10 集落)

要となる。空間単位は集落・旧村・市町村、個人、あるいはデータをカテゴリー化することで得られる境界など様々なスケールで想定することができる。その都度、施策に応じた適正な空間規模を把握し、SCの変動とこれと関連する要因を明らかにすることができれば、より効率的かつ戦略的な方策を生み出すことが可能になる。

表2 近傍集落数別 EBI とモランの I の値

近傍集落数	EBI	モランの I
1	0.348	0.154
2	0.364	0.181
3	0.366	0.184
4	0.369	0.208
5	0.348	0.172
6	0.332	0.168
7	0.195	0.174
8	0.201	0.180
9	0.208	0.188
10	0.213	0.193

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① Saizen, I., A. Maekawa, N. Yamamura: Spatial analysis of time-series changes in livestock distribution by detection of local spatial associations in Mongolia. *Applied Geography* 30, 2010, 639-649. 査読あり
- ② 福島慎太郎, 吉川郷主, 市田行信, 西前 出, 小林慎太郎. 一般的信頼と地域住民に対する信頼の主観的健康感に対する影響の比較, *環境情報科学論文集*, vol.23, 2009, 269-274. 査読あり
- ③ 西前 出, 市田行信, 吉川郷主, 小林慎太郎. 経験ベイズ法を用いたソーシャルキャピタルの尺度の空間的自己相関分析, *環境情報科学論文集* vol.22, 2008, 315-320. 査読あり

[学会発表] (計3件)

- ① 福島慎太郎, 吉川郷主, 市田行信, 西前 出, 小林慎太郎. 一般的信頼と地域住民に対する信頼の主観的健康感に対する影響の比較, *環境研究発表会*, 2009年11月30日、日本大学(東京都)
- ② 西前 出, 市田行信, 吉川郷主, 小林慎太郎. 経験ベイズ法を用いたソーシャルキャピタルの尺度の空間的自己相関分析, *環境研究発表会*, 2008年11月25日、日本大学(東京都)
- ③ 西前 出, ソーシャルキャピタルを計る, *環境研究発表会*, 2008年11月25日、日本大学(東京都)

6. 研究組織

(1) 研究代表者

西前 出 (SAIZEN IZURU)
 京都大学・地球環境学堂・准教授
 研究者番号：80346098