

SOM と GIS を用いた社会地域分析の試み

-別府市を例として-

Social Area Analysis Using SOM and GIS

-A Case Study on Beppu City

李 燕*・シャンムガナタン スバナ**

Yan LI*, Subana Shanmuganathan**

Social Area Analysis is an important and helpful tool in understanding a city's socio-spatial attributes. However, its applications in urban planning or other related practices are scarce. The first reason for this is that the linear correlation methods implemented in the past years oversimplified the complexities of the reality, and the results often bring a pattern irrelevant to the real world. Another reason is that visualization of spatial patterns requires such a vast amount of labor that sometimes it is impracticable. This paper aims to propose a non-linear correlation method using Self Organizing Map (SOM) techniques to investigate the social areas within a city and then to project the data on a GIS for visualization. We select Beppu City, Oita Prefecture as a testing area, because there are not many studies on social area analysis of Japanese cities, and after this, we intend to expand our study to larger cities of this country.

Keywords: Self Organization Map (SOM), Geographic Information System, social area analysis, population census
自己組織化マップ (SOM), 地理情報システム (GIS), 社会地域分析, 国勢調査

1. 研究の背景と目的

社会地域 (Social Area) とは、都市において、異なる社会特性をもつ居住者による住み分けによって形成される空間的構造であり、居住地域 (Residential Area) と呼ばれる。スラムやゲットー、チャイナタウンなどがその分かりやすい例であるが、高級住宅街や職人の町なども社会地域の一種である。

人種が多く、社会格差の激しい欧米では、社会地域は普遍的な現象として存在することが多くの研究で確認されている。1920年代のシカゴ学派は、シカゴの社会空間構造が同心円状であると分析し、その背後にある社会過程を競争 (competition)、闘争 (struggle)、応化 (accommodation)、同化 (assimilation) の四つに分けてとらえた¹⁾。1950年代においては、Shevky and Bell はサンフランシスコを例として、アメリカにおける居住分離は「社会地位」、「アーバンニズム」(ライフスタイルや家族構成など) および「民族」(エスニック・ステータス) の3つの要因に影響されると指摘し²⁾、その後、社会地域分析 (Social Area Analysis) が欧米諸国そして発展途上国においても多くの研究が行われた。手法論から見れば、Shevky and Bell の当時は、社会変動の理論から社会地域形成される要因を仮定し実証するという演繹的な方法を取っていたが、その後は、数理統計、特に因子分析を用いた帰納的な手法が主となった³⁾。

日本においては、外国人が少ないため、人種的・民族的住み分けについて、大都市以外の研究は報告されていないが、人口密度から見た「都市空間構造」や住宅の質や環境を用いた「居住地域」の研究が欧米の「社会地域分析」とほぼ同時代、すなわち1950年代から1970年代を中心に、多くなされている⁴⁾。

都市計画の実務においても、以上のような人口の社会的特徴を詳細に把握することが重要であるが、社会地域分析手法の利用が必ずしも多くない。その原因として、まずは、線形的な多変量解析に基づく従来の社会地域分析手法は、多様な現実社会を捉えにくい場合が多いことが考えられる。もう一つの原因は、実際の都市の社会地域を扱う際、仮に社会地域の特徴が統計的に分かっても、その空間分布を可視化することがたいへん苦勞を要するので、躊躇される。

近年、情報技術の飛躍的な進歩によって、社会経済に関する多くの情報がデータベース化されつつだけでなく、データ分析および可視化の手法も飛躍的な発展を遂げている。本研究は、非線形的なデータマイニング手法の1種である自己組織化マップ (SOM: Self Organization Map) を用いて社会地域の分析を試み、その可視化に地理情報システムを用いることにした。ケーススタディを通じて、提案する方法の有用性を示し、今後の都市計画・分析に資したい。

2. 自己組織化マップ (SOM: Self Organization Map)

SOM は、大脳皮質の感覚野に見られる機能地図つまり特徴地図が、知覚経験に基づくある種の学習によって形成されていくありさまを説明する数理モデルである。1980年代 Kohonen による提案され⁶⁾、日本においては、1990年代後半から、経済学、土木工学、情報工学、商業分野などの実用例が報告されるようになった⁷⁾。

基本的な SOM は2次元に配列されたノード (神経細胞) からなっている (図1の出力スペース)。この出力スペースは必要に応じて1次元や3次元などにすることができる。各ノードには、参照ベクトルと呼ばれるn次元ベクトルが初期値として与えられている。

*正会員、立命館アジア太平洋大学 (Ritsumeikan Asia Pacific University)

**正会員、立命館アジア太平洋大学 (Ritsumeikan Asia Pacific University)

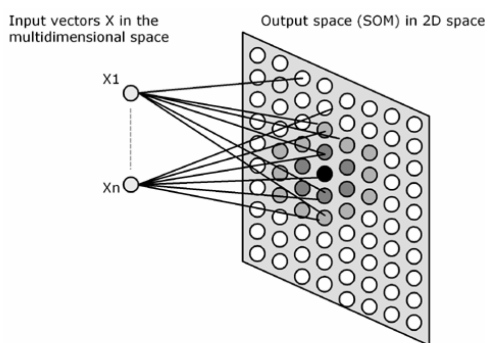


図1 SOMの基本概念⁸⁾

同じく n 次元の入力ベクトル X から入力 x_1 を与えられると、 x_1 に最も合致する参照ベクトルをもつノードが「勝者」となり、そのノードおよびその周囲のノードがこの入力を学習する。入力 x_1 に対する「学習」は、入力に合致する方向へすべてのノードの参照ベクトルを変更することでなされるが、ノード配列上のユークリッド距離に影響されるため、主として勝者の周囲で起こる。その結果、後で再び x_1 が入力されたときや x_1 に類似する入力が入力されたときには、このときの勝者がまた勝者になって学習を起しやすい傾向が生じる。同様に、ほかの入力が与えられたときもそれぞれの勝者を決め、学習を起す。このような X からの入力、勝者としてのノードの決定、周囲ノードの学習が繰り返し行われ、その結果、もともと n 次元の入力ベクトルのパターンが、2次元の出力スペースに現れる。このように、SOMは高次元のデータや、ベクトル空間上にないデータを、2次元の平面上などのより低次元で容易に観察できる空間に写像する（次元削減する）ことでデータの可視化を実現する。

次元削減によって可視化を行う手法としては他に主成分解析などがあるが、曲面上に分布している場合は主成分解析ではうまく削減できない。SOMなら、高次元空間上のニューロンの配置が曲面にフィットするよう変形するので、表示用の空間を有効に利用できる。データセットの次元を減らし、通常的手法で見つからない知識を発見すること、すなわちデータをマイニングすることができるので、情報の急増している現代社会において、非常に有用であると考えられる⁹⁾。本研究は、データマイニングの有力のソフトの1つである Viscovery® SOMine を用いて、標準に設定されているパラメータを一切変更しないで SOM を行った。

3. 研究対象エリアおよびデータセット

本研究はケーススタディとして、大分県別府市（人口12万人）を対象に、2000年に行われた国勢調査の小地域集計から、次の86個の変数を抽出した。

- (1) 総人口、男性数、女性数、世帯数、人口密度。
- (2) 5歳間隔の男女別年齢。男女それぞれ0歳から99歳まで5歳間隔の20グループおよび100歳以上の1グループに分けられ、トータル42個の変数になる。本文の図表では、

F(0-4)、M(5-9)のような記号で表すが、F(0-4)は0から4歳の女性の数、M(5-9)は5から9歳の男性の数を指す。

- (3) 外国人男性・女性の数 (M_Foreigner、F_Foreigner)
- (4) 住宅の形態。一戸建て、長屋、1-2階建ての共同住宅、3-5階建ての共同住宅、6-10階建ての共同住宅、11階建て以上の共同住宅の6種類あるが、それぞれ持家、借家、間借りに区別されているので、全体で18個の変数になる。記号では、「O」で始まるものは持家のことを現し、「R」で始まるものは借家のことを表し、「R_r」で始まるものは間借りを表す。

(5) 居住期間。男女別にそれぞれ1年以下、1-4年、5-9年、10-19年、20年以上、出生時からの6区分が設けられているので、全体では、12個の変数になる。記号は、例えば、D_F_14は居住期間が1-4年の女性の数を表す。

(6) 主な家計収入の種類。賃金・給料のみ (Sal_Only)、恩給・年金のみ (Pen_Only)、賃金・給料以外の収入もある (Sal+)、恩給・年金以外にも収入ある (Pen+)、仕送り (Allow)、農業 (Agri)、農業以外の事業収入 (non_agriB)、内職 (inh_work) とその他 (Others) の9種類に分けられる。

なお、別府市は180個の国勢調査基本単位があるが、人口のほとんどない山間部の17個を除いたため、入力ベクトルは86次元の163個の要素が含まれる。

4. 別府市の社会地域分析

上述の人口・社会に関するデータを用いて、自己組織を行わせ、その写像を2つのクラスター（グループ、分類）から徐々に増やし、最後は8クラスターまで分けた。それぞれの自己組織写像は図2に示す。

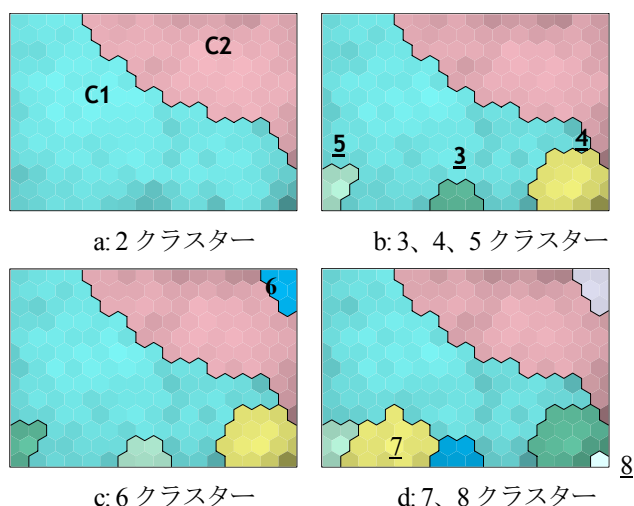


図2 2から8クラスターに形成させたときの SOM

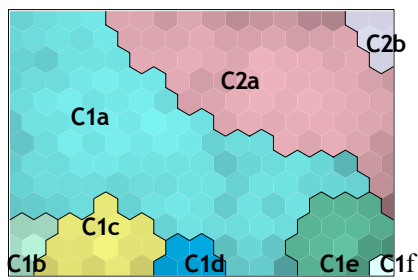


図3 各クラスターの名前

この図から分かるように、自己組織写像はまず2つ(C1、C2)に分かれる(図2a)。さらにC1から3つの新しいクラスターが形成され(図2b)、6クラスターに分けると、新たに1つのクラスターがC2から形成される(図2c)。7クラスターでは、再びC1が細分化される。8クラスターはクラスター4から細分化されたことが図2dで分かる。便利のため、図3のように各クラスターに名前をつけた。

4.1 2クラスターSOM

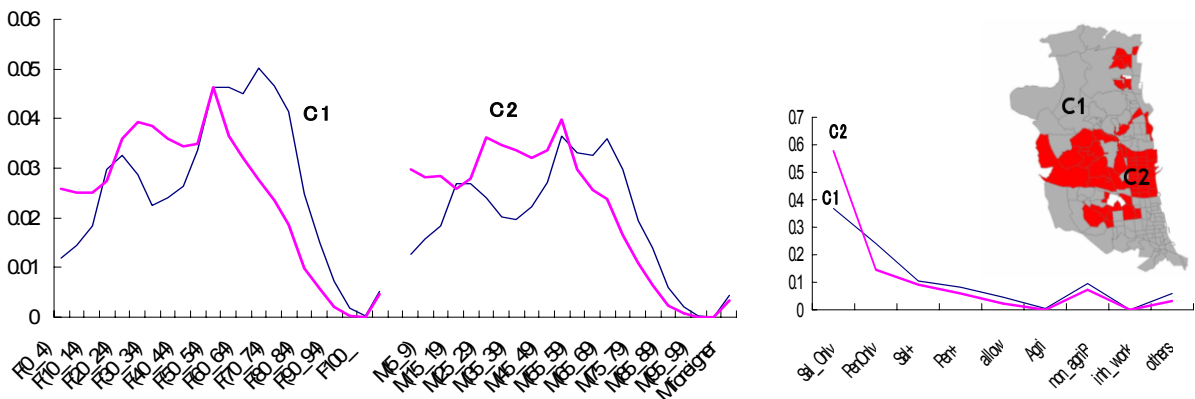
2クラスターSOMの空間分布および各変数の比較を示したのが図4である。この図から、次のことが分かる。

- (1) 右上の分布図から見ると、縦長い対象エリアにおいて、C2が中間の部分に分布し、C1が南北両端(南:別府地区、北:亀川地区)に分布している。この分布は、

現在の別府市は1933年において別府と亀川によって合併されたという発展過程に合致している。したがって、C2は比較的に新しい地域である。また、亀川地域においてもC2に属される地域があるが、都市再開発が盛んに行われている。

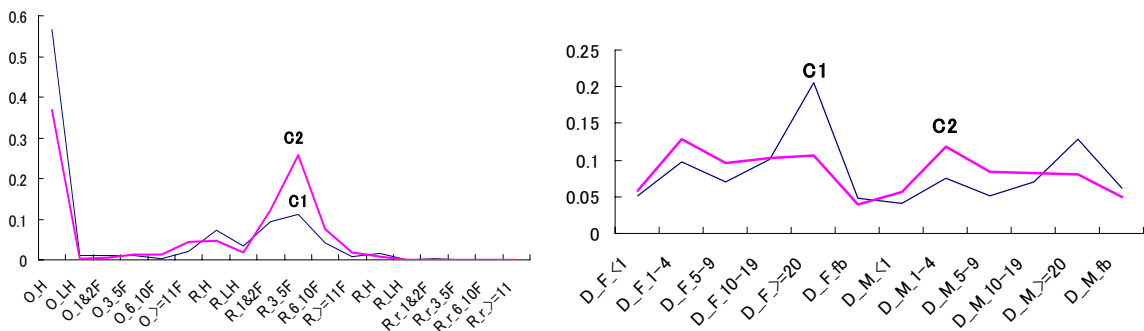
- (2) 年齢層の人口割合(図4a)を見ると、C1は55歳以上の割合がC2より大きくなっているのが目立つ。また、C2は24から45歳および0から14歳の年齢層がより多い。このような傾向は男女ともに現れている。
- (3) 収入の種類を見る(図4b)と、C1では年金のみ、C2ではサラリーのみ世帯の割合が相対的に高いことが分かる。
- (4) 住宅のタイプ(図4c)をみると、C1では持家の割合が高く、マンション借家の割合が低い特徴がある。C2はその反対で、低持家・高マンション賃貸の特徴になる。
- (5) 居住年数を見ると、C1では20年以上その地域に住み続いた人の割合が非常に高く、C2では1年から9年の割合が高くなっている。

以上の特徴を総合すると、C2は賃貸マンション住まいの若いサラリーマンの核家族が典型であるが、C1は持家の一戸建て住まいの中高年者が目立つことが言える。



a: 各年齢層の割合 (左: 女性、右: 男性)

b: 収入の種類



c: 住宅の種類

d: 居住年数

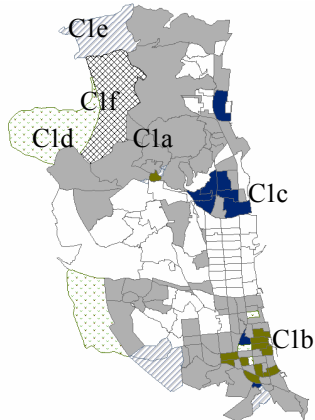
図4 C1、C2の空間分布(右上の図)および社会特性(図aからd)

4.2 クラスタC1の細分化(C1aからC1f)の分析

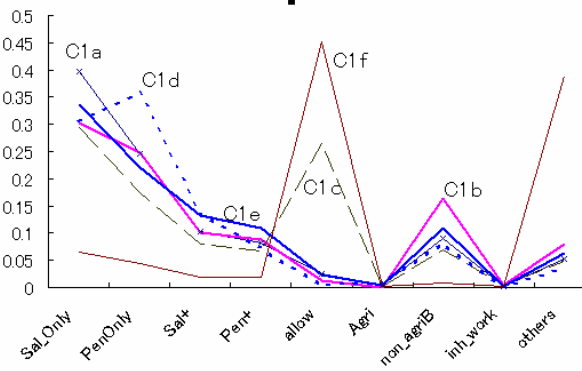
C1から細かく分けると、C1aからC1fの社会地域に分けることができる。その空間分布および各変数の比較を示したのが図5である。この図から、次のことが分かる。

(1) 図5aはC1の細分化の空間的分布を示している(C2

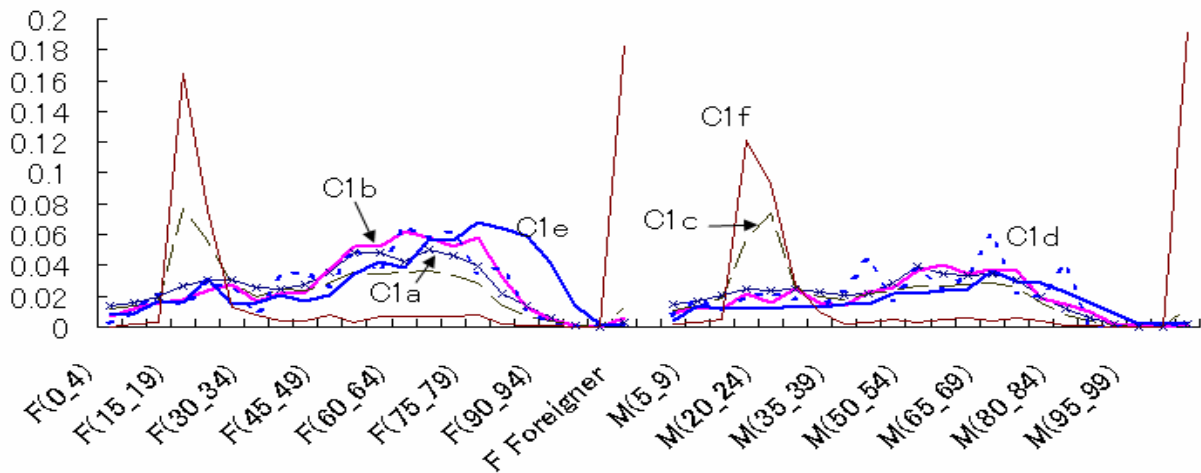
の部分)は白色)。この図を見ると、C1d、C1e、C1fは主に西側の山地に分布している(別府市の地勢は主として西側の山から東の海への斜面である)。主な市街地の地域では、C1cとC1bが周囲の地域(C1a)と異なる社会特徴を持っている。



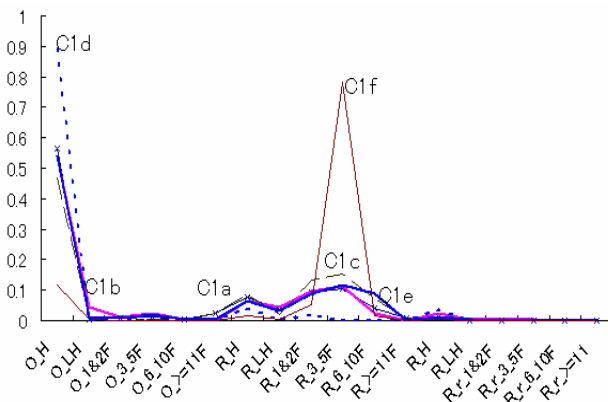
a: 空間的分布



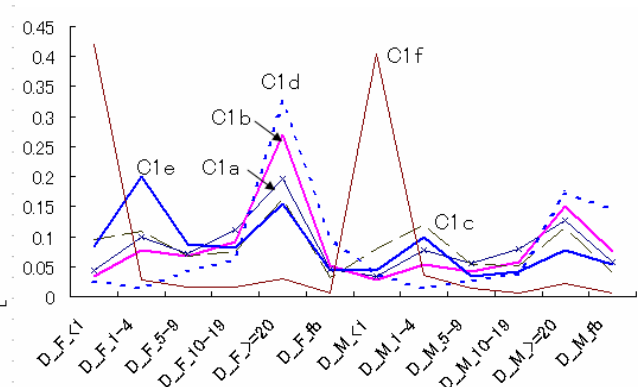
b: 収入の種類



c: 年齢層の割合 (左: 女性、右: 男性)



d: 住宅の形態



e: 居住年数

図5 C1aからC1fの空間的分布およびその社会特徴の比較

- (2) 収入の種類を見ると(図5b)、C1fとC1cが特出して仕送りの割合が高い。これはそれぞれに立命館アジア太平洋大学および別府大学が立地しており、その周囲に学生が多く居住しているからと思われる。また、年齢層から見てもこの二つの地域が15から29歳の割合が高く(図5c)、住宅形態も主として賃貸マンションである(図5d)。居住年数も国勢調査年度設立された立命館アジア太平洋大学の所在地(C1f)では、人口の居住年数が一年未満がほとんどであり、一方、別府大学地域では1から4年の居住年数の割合が高くなっている。C1fの外国人数が非常に高いことは、この大学の学生は半数が外国人で、留学生用の寮がその地域にあるからである。
- (3) C1bは、農業以外の事業収入の割合が相対的に大きい(図5b)。年齢層から見ても、C1eに次ぐ2番目高齢者の割合になっている(図5c)。また、長屋の割合が相対的に高く(図5d)、居住年数もC1dに次ぐ2番目に長い(図5e)。この地域はJR別府駅の周辺であり、歴史が長く、温泉、観光、商店街など年を取った自営業者が多く居住している街である。
- (4) C1dの最も大きな特徴としては、年金収入の割合が高く(図b)、一戸建ての持家に長く住んでいることが特徴である(図dとe)。この地域は、若いときに農業を中心に営んだ高齢者が多く見られる。
- (5) C1eは居住年数が1から4年が多く(図5e)、最も高

齢者の多い地域(図5c)である。これらの地域において、老人ホームが多く分布されている。

- (6) C1bからC1fに比べると、C1aの特徴は最も広く分布しており、社会特徴も鮮明でない。これを次の節でC2の細分グループC2aとC2bと比べることにする。

4.3 クラスタ-C2の細分化(C2aとC2b)およびC1aとの比較分析

図6は、C2の細分化C2aとC2b、および特徴の鮮明でなかったC1の細分化C1aの空間分布および各変数の比較を示している。この図から、次のことが分かる。

- (1) 図6の右上の分布図を見ると、C2bは中部の東側に位置している。この地域においては1970年代から土地区画整理事業が行われ、賃貸マンションが多く立地する地域である。
- (2) 図6aの年齢層割合を見ると、C2bは突出して年齢層が若いことが分かる。C1aは高齢化が特徴であり、C2aはその中間にある。
- (3) C2b、C2a、C1aの順に、サラリーのみ世帯が減少し、年金のみ世帯が増える(図6b)。また、持家の一戸建ての割合も急増し、高層マンション住まいの世帯が減少する(図6c)。居住年数もC2bが最も短く、C1aが最も長い。

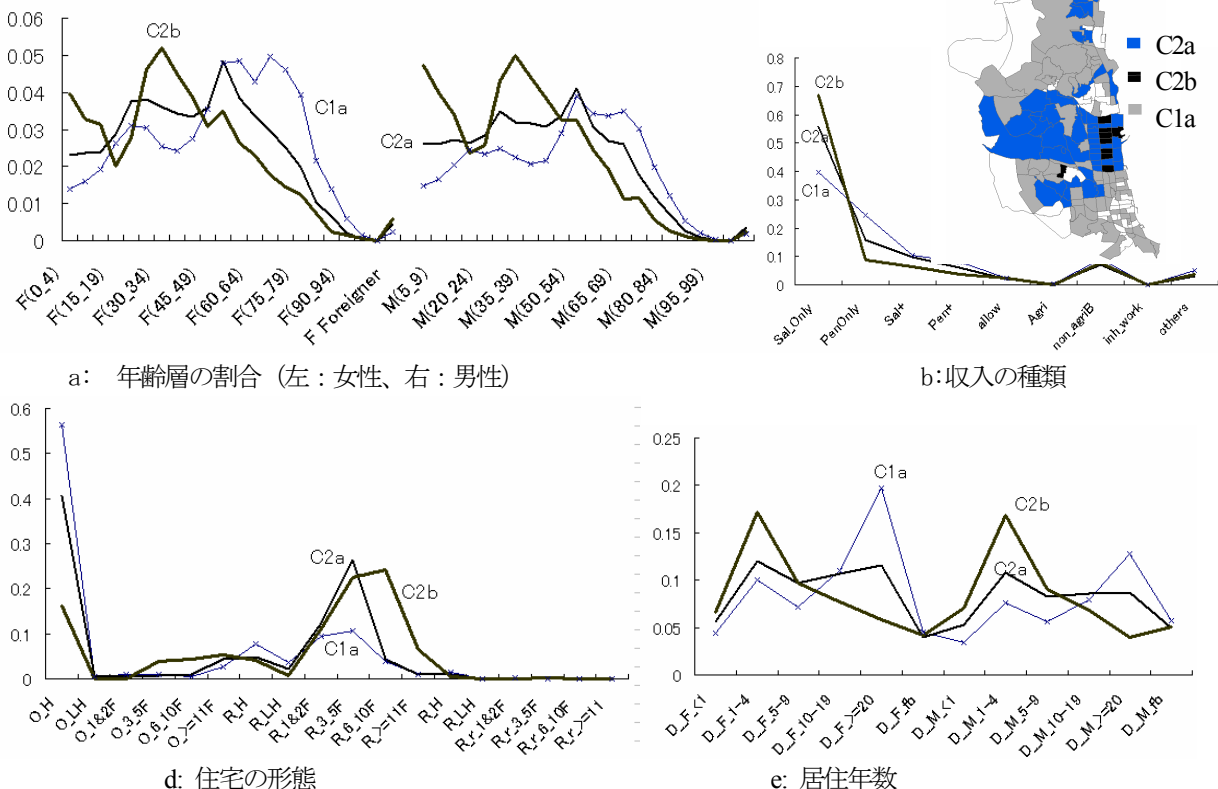


図6 C2a、C2bおよびC1aの空間的分布およびその社会特徴の比較

以上の分析から分かるように、C2aはC1aより年齢層が若く、収入がサラリーのみで、賃貸マンション住まい、居住期間が短い世帯が多いが、C2bはさらにその特徴が鮮明である。前述のように、C2エリアは歴史的新しい地区で、特にC2bは土地区画事業によって整理され、高層の賃貸マンションが多く立地されているからであると思われる。

5. 終わりに

以上は、別府市を例として、国勢調査のデータを用いて、SOMによる分類、GISによるビジュアル化の方法を用いて、社会地域の分析を行った。

まず、本研究で使用した変数は、人口の社会特性を描写するものだけで、変数の取捨はしていない。データの統計処理も標準化以外は一切行っておらず、パラメータも主観的な操作をしていない。したがって、データの選択・加工など高度な知識の要る統計手法に比べると、SOMによる分析は便利で、主観性が少ないという利点がある。

次に、同じデータセットを統計分析のクラスターリングを行ったが、有意な結果が得られていない。これも非線形形のデータマイニングによる手法の利点であると思われる。

また、変数においてお互いの位置関係を示すものを使用していないにもかかわらず、空間的に連続している社会地域の存在がGISの画面から確認できた。欧米では、社会地域の分布は社会地位、ライフスタイル、民族に影響されるとされているが、別府を通じて、日本では都市の発展や計画による特定の年齢構成の空間的集中、職業、すなわち社会地位とライフスタイルの相互作用が主な原因になっていると言える。

【参考文献】

- 1) Park, R. E. and Burgess, E. W. (1928): Introduction to the Science of Sociology, Chicago: University of Chicago Press.
- 2) Shevky, E. and Bell, W. (1955): Social Area Analysis. Stanford, CA: Stanford University Press.
- 3) Economic Geography, Vol. 47, Supplement: Comparative Factorial Ecology, Jun., 1971
- 4) 木内信蔵: 都市地理学原理、古今書院、1979。
- 5) 川上秀光、地方中心都市における密度構造の変容、都市計画論文集、1980、pp.73-78。
- 6) Kohonen, T. (1982). Self-organised formation of topographically correct feature maps. Biological Cybernetics, 43:59-69.
- 7) 徳高平蔵、藤村喜久郎、山川烈監修、自己組織化マップ応用事例集: SOMによる可視化情報処理、海文堂出版、2002.10.
- 8) Koua1, E. L. and Kraak1, M. J. (2004): Geovisualization to support the exploration of large health and demographic survey data, International Journal of Health Geographics 2004, 3:12

- 9) Amari, S. (1995). Foreword. Foundations of Neural Networks, Fuzzy systems and Knowledge Engineering. N. K. Kasabov, The MIT Press: xi