

# 長期人口統計における 巨大災害の影響抽出の試み

古橋 峻<sup>1</sup>・奥村 誠<sup>2</sup>

<sup>1</sup>学生会員 東北大学大学院 工学研究科博士後期課程  
(〒980-8577 仙台市青葉区片平2-1-1 通研2号館 地域計画学研究室)  
E-mail: rjfsighu@yahoo.co.jp

<sup>2</sup>正会員 東北大学教授 東北アジア研究センター  
(〒980-8577 仙台市青葉区片平2-1-1 通研2号館 地域計画学研究室)  
E-mail: mokmr@m.tohoku.ac.jp

東日本大震災などの巨大災害の復旧、復興のための政策に必要な財源を、どのように調達するかは大きな課題となっている。特に、一時的な増税によるか、将来世代の負担に頼る国債等の発行によるかの判断には、巨大災害の影響がどのような規模に広がり、何年ぐらい継続するのかという見通しが必要となる。本研究は、上記のような見通しを迅速に立てることができるように、社会経済統計における過去の巨大災害の影響の抽出方法を提案しようとするものである。具体的には、近年信号処理の分野などでの応用が進みつつある独立成分分析 (ICA) 手法を取り上げ、1920年から2010年の国勢調査都道府県別人口統計データへの適用を試みる。

**Key Words :** *huge scale disaster, population statistics, independent component analysis (ICA)*

## 1. はじめに

今年の3月の東日本大震災は、死者・行方不明者合わせて23,700名あまりの人的被害を出した。この巨大災害からの復旧・復興にはさまざまな政策が必要となるが、その財源をどのように調達するかが大きな課題となっている。特に、一時的な増税によるか、将来世代の負担に頼る国債等の発行によるかの判断には、災害の影響がどのような規模に広がり、何年ぐらい継続するのかという見通しが必要となる。本研究は、上記のような見通しを迅速に立てることができるように、社会経済時系列統計における過去の巨大災害の影響の抽出方法を提案することを目的とする。

巨大災害が社会経済現象にもたらす影響を評価に関する研究として、河田・柄谷(2000)は、わが国の災害による被害額が、人的損失の指標である平均寿命の減少分と強い相関関係にあることを示し、人口にもたらした影響を用いて経済的な影響の大きさを把握できる<sup>1)</sup>としている。中島ら(2005)は都道府県別の自然災害リスクと被害額の関係を分析し、ハザードに対する被害の強度を脆弱性と考へて、ハザードの大小と脆弱性の大小により都道府県のカテゴリを試みている<sup>2)</sup>。さらに外谷 (2009) は、この脆

弱性の程度を事前の防災投資額と消防などの緊急対応力によって説明することを試みている<sup>3)</sup>。以上のように、社会経済への災害の影響を人的被害を通して把握し、それを防災対策の評価に用いようとする試みは進められているが、被害の時間的、空間的な広がりや把握を試みた分析は行われていない。

本研究では、長期にわたり同一の信頼性で全国的に調査されている国勢調査に着目し、1920年から2010年までの5年ごとの10月1日現在の都道府県別人口のデータ<sup>4)</sup>を用いる。ただし1945年は11月1日に沖縄県以外で実施された人口調査に基づく値であり、以下では沖縄県を除く46都道府県を対象とする。複数の時系列データの中に存在する変動パターンを抽出するための手法としては、信号処理の分野で広く応用が進んでいる独立成分分析 (ICA : Independent Component Analysis) を用いる<sup>5)</sup>。

## 2. 独立成分分析の適用方法

### (1) 独立成分分析の概要

独立成分分析は多変量解析手法の1つであり、主成分分析や因子分析と同様に、変数を内生変数/外生変数に

分類することなく、データの挙動をより少数の変数に縮約して表現するための統計手法である。すなわち、多変量時系列として時刻  $t$  ごとに観測される  $x_i(t)$  が、相互に独立な複数の原系列  $s_j(t)$  が混合係数  $a_{ij}$  を介して重ね合わされたものであると考えると、原系列  $s_j(t)$  と混合係数  $a_{ij}$  を同時推定する方法である。因子分析では、原系列が相互に無相関であると仮定するのに対して、独立成分分析では相互に独立であると仮定するところに特徴がある。

観測系列  $x_i(t)$ 、原系列  $s_j(t)$  および混合係数  $a_{ij}$  について、式(1)が成り立つ。

$$x_i(t) = \sum_{j=1}^k a_{ij} s_j(t), \quad (i=1, \dots, n) \quad (1)$$

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (2)$$

ただし、 $\mathbf{x}$ 、 $\mathbf{A}$ 、 $\mathbf{s}$  はそれぞれ観測系列  $x_i(t)$ 、混合係数  $a_{ij}$ 、および原系列  $s_j(t)$  のベクトル表示であり、 $\mathbf{A}$  を混合行列と呼ぶ。

Hyvarinenら<sup>5)</sup>は、「 $n$ 本の独立な確率変数(原系列)の和によって定義される確率変数(混合系列)の確率密度分布は、混合前の確率変数の確率密度分布よりも正規分布に近い」という性質に着目して、式(3)に示す4次のモーメントである尖度を系列間の独立性の指標として、尖度の絶対値を最大化するFastICAアルゴリズムを開発した。

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (3)$$

FastICAアルゴリズムはすでに各種の統計解析ソフトに実装されており、本研究ではフリーの統計解析ソフトR用に用意されたプログラムを用いる。

なお、FastICAアルゴリズムでは、最初に主成分分析を実施して観測系列の白色化を行い、各行が無相関で分散が1の行列とした上で分析を進める。このとき観測データ系列の分散に対する寄与率を基に白色化ベクトル(主成分) $\mathbf{z}$ の次元を縮約することで、原系列 $\mathbf{s}$ の抽出数を限定することが行われる。FastICAアルゴリズムでは、抽出数を変えると、異なる形の原系列が抽出されてくることが知られており、抽出された原系列や混合係数が無理なく意味づけができるように、抽出数を調整していく必要がある。

## (2) 独立成分分析の適用事例

筆者らのグループは、近年利用可能となった自動観測交通量データを対象として、独立成分分析の適用を試みてきている。加藤ら(2008)では、航空便の発着時刻の遅れ時間の1年間の時系列データに独立成分分析を適用し、降雪、台風、混雑空港などを原因とする遅延がそれぞれ異なるパターンで複数の便に波及していく様子を分析し、先行する便の遅延に基づいて、後続便の遅延の発生を推測する方法について考察した<sup>6)</sup>。塚井ら(2008)は地方

空港に流入する車両の連続観測データに独立成分分析を適用し、特定の曜日や特定の時間帯に見られる交通量の集中パターンを把握した<sup>7)</sup>。井上ら(2009)はさらに、そのようにして抽出された変動パターンのいくつかを、航空便の欠航や遅延に伴って発生していることを明らかにした<sup>8)</sup>。以上の分析では、十分なサイズの観測データが存在すれば、その中から非常時のみに現れてくる変動パターンを抽出できることを示したものである。一方、フォンセカら(2009)は、路線バスの停留所ごとの乗車人数の係数データに独立成分分析を適用して多くの日に共通して観測される変動パターンを抽出し、それらの合成により計算される期待乗車人数と現実の観測人数との差を用いて、車内混雑などの理由により潜在化した需要の大きさを推定する方法を提案した<sup>9)</sup>。この研究は、十分なサイズの観測データがないときには異常時の影響を原系列として抽出することは困難であるが、観測系列と混合系列との差異を用いて考察できる可能性を示している。

## (3) 人口における災害の影響分析の方法

### a) 人口時系列を系列とみなす方法

多時点の人口データが都道府県ごとに観測されているから、最も素直な方法は、観測時点  $t$  とし、都道府県のインデックスを  $i$  とし、人口の時系列を観測系列  $x_i(t)$  とみなすことである。この際、独立成分分析により推定される原系列  $s_j(t)$  は、戦争による大量の死亡や、ベビーブーム世代とその子世代の出生というように、歴史の中で起こる社会的な動きを表わしており、それらの動きがどの地域で発生するかという強度が混合係数  $a_{ij}$  を用いて表現される。

独立成分分析では混合係数が一定であると仮定するので、例えばベビーブーム世代とその子世代の出生地の空間分布が安定的であればこの方法が適用できる。しかし、巨大災害の発生場所は一定ではないので、災害の影響を原系列として抽出することは困難であり、観測系列と混合系列との差異に注目することが必要であると考えられる。

### b) 各時点の人口の空間分布を系列とみなす方法

人口の変動パターンを解釈する上では、人口データを転置し、都道府県を  $t$  とし、観測時点  $i$  とし、時点ごとの人口の空間分布を観測系列  $x_i(t)$  とみなした上で独立成分分析を適用することが考えられる。この際推定される原系列  $s_j(t)$  は、標準的な人口の地域分布、地方圏から三大都市圏への移動に伴う人口の差分、大都市圏内部のドーナツ化による郊外と周辺の人口の差分、東京圏への一極集中のパターンなどの、人口の空間分布の標準的なパターンを意味する。一方混合係数  $a_{ij}$  は、それらの空間パターンがいつの時点から強まってきたのかを

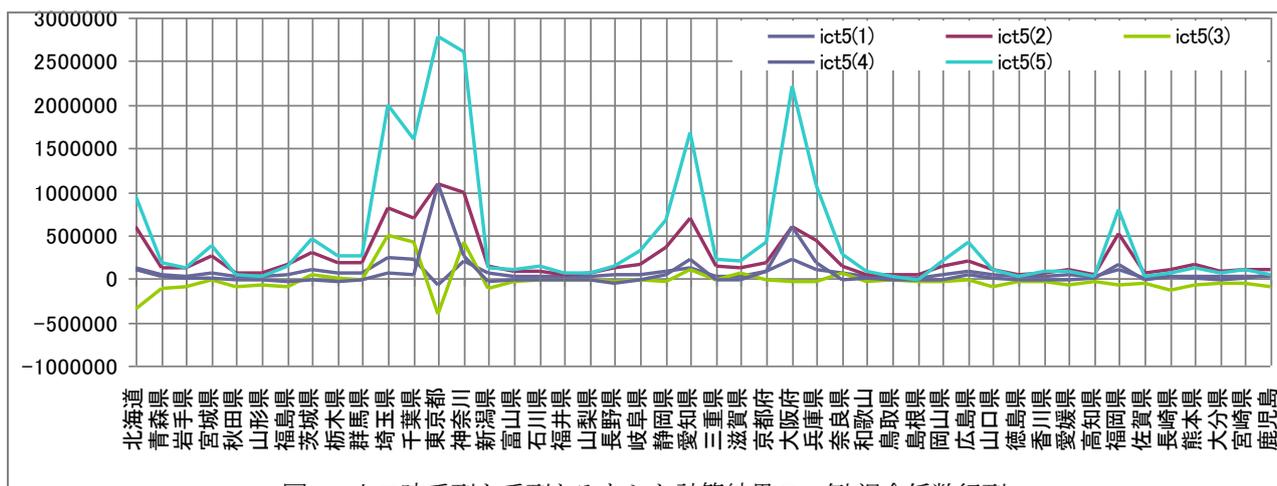


図-1 人口時系列を系列とみなした計算結果の一例(混合係数行列)

表わす。

この方法では、巨大な災害の影響が一つの原因として抽出できる可能性がある。すなわち、当該災害の人的被害の空間分布に類似する原系列が抽出され、混合係数行列の対応する列を見ることによって影響の時間的な変化を理解できる可能性がある。ただし、関東大震災の影響と東京一極集中に伴う社会移動の影響のように、類似の空間パターンを持つ減少が同じ原型列の中に含まれてしまう危険性が高い。また、災害の空間的波及が大きくなければ、災害の影響は原系列として抽出できないため、観測系列と混合系列との差異を精査して影響を考察する必要がある。

#### (4) 人口変化における災害の影響分析の方法

##### a) 人口の階差時系列を系列とみなす方法

災害の影響は、各時点の人口そのものではなく、むしろ死亡や社会移動に伴う人口の変化に作用すると考えると、人口の階差を取って5年間の人口の増減量のデータを作成し、独立成分分析を適用する方法が考えられる。すなわち、2時点間の変化期間のインデックスを  $t$  とし、都道府県のインデックスを  $j$  とし、各都道府県ごとの人口増減量の時系列を観測系列として扱う。

独立成分分析により推定される原系列  $s_j(t)$  は、戦争による死亡による人口減少とそのリバウンドとしてのベビーブーム世代の出生というように、歴史の中で起こる社会的変化の時間変化パターンを表わしており、それらの変化がどの地域で発生するかという強度が混合係数  $a_{ij}$  を用いて表現される。

この方法では災害の影響が原系列の形で抽出される可能性は小さく、観測系列と混合系列との差異を考察する必要がある。

##### b) 各期間の人口増減の空間分布を系列とみなす方法

人口増減量データを転置し、都道府県を  $t$  とし、観測時点間の期間のインデックスを  $j$  とし、期間ごとの人

口増減の空間分布を観測系列  $x_i(t)$  とみなした上で独立成分分析を適用する方法も考えられる。この際推定される原系列  $s_j(t)$  は、標準的な人口増加の地域分布、地方圏から三大都市圏への移動、大都市圏内部のドーナツ化による郊外から周辺の移動、東京圏への一極集中的な移動などの、人口移動の空間分布パターンを意味する。一方混合係数  $a_{ij}$  は、それらの人口移動の空間パターンがいつの時点から強まってきたのかを表わす。

この方法でも、巨大災害の影響が一つの原因として抽出できる可能性がある。また、災害の空間的波及が大きくなければ、災害の影響は原系列として抽出できないため、観測系列と混合系列との差異を考察する必要がある。

### 3. 計算結果と考察

#### (1) 人口時系列を系列とみなした計算結果

独立成分分析の特徴として、抽出する成分数を増やしていった場合、すでに抽出されていた成分が必ず抽出するとは限らない。ここでは一例として、5成分の抽出を行った場合の混合係数行列を図-1に示す。

このように、都道府県別の混合比率が、多くの成分について類似するパターンを持っており、この方法では明瞭に解釈できる人口パターンを抽出できないことがわかった。

#### (2) 人口の空間分布を系列とみなした計算結果

一例として、5成分の抽出を行ったときの独立成分を図-2に示す。5つの成分は順に、東京、大阪、工業圏、東京圏郊外、北海道と東北・静岡への人口集中パターンに対応している。図-3の混合行列から、はじめの2つの成分は戦前から値が高かったが戦災の影響で負に転じている時期があることがわかる。工業化の成分に比べて、東京圏のドーナツ化は1960年代以降に急速に進んできた

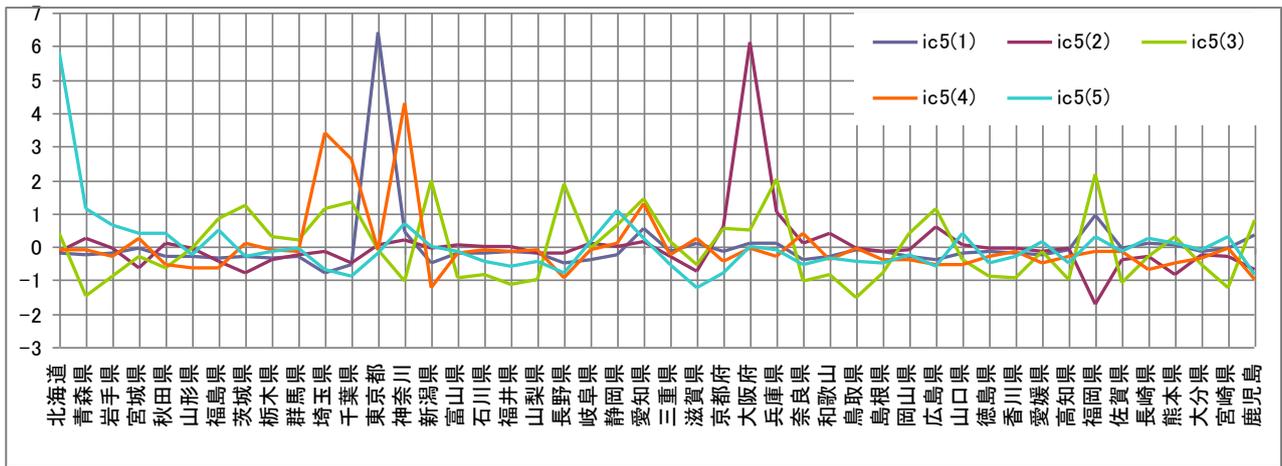


図2 人口の空間分布を系列とみなした計算結果の一例(独立成分)

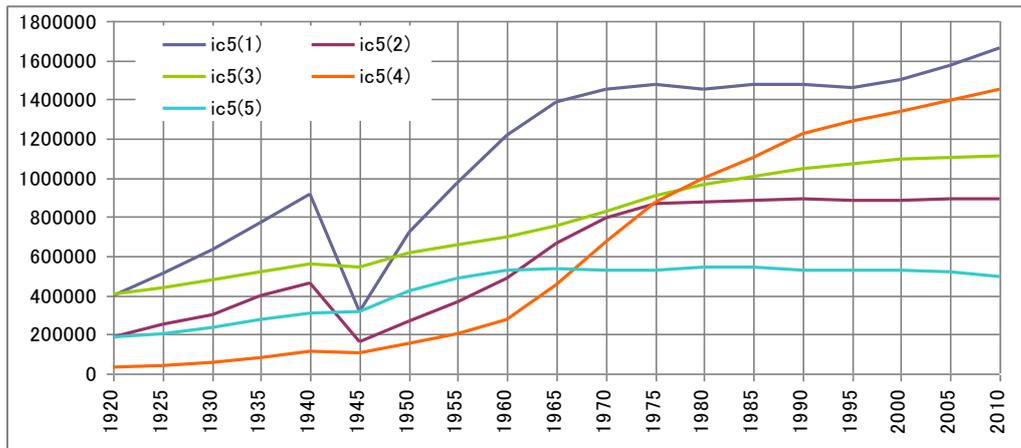


図3 人口の空間分布を系列とみなした計算結果の一例(混合係数)

ことが読み取れる。さらに抽出する成分数を増やしていくと、グループの中から個別の地域だけを切り離して説明する成分が得られるだけであり、当てはまりは改善されるがパターンの意味は面白みが失われる。

なお、成分を10個程度に増やしても、災害の影響が明瞭に見られる成分は抽出できなかった。そこで分析結果から混合率を用いて独立成分を再合成し、観測系列との差異を計算し、その中に自然災害との対応が見られるかどうかをチェックする。いま、10成分の計算値に基づく残差成分を図4に示す。これより確かに1995年において兵庫県(28)の人口観測値が過小となっており、阪神淡路大震災の影響が疑われる。また2000年以降は過大となっていることから、震災の影響は5年程度で解消していると解釈できる。ただし、1990年の愛知県の人口も過小であるが、これに対応する自然災害は記録されていない。

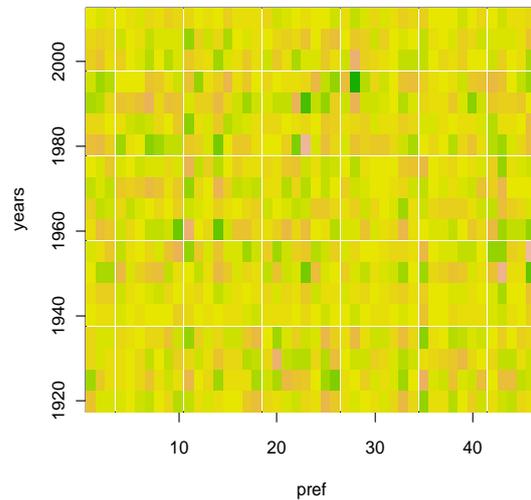


図4 人口の空間分布を系列とみなした場合の残差の分布(緑色が過小、赤色が過大)

### (3) 人口の階差時系列を系列とみなした計算結果

このケースも(1)と同じく、異なる時系列の混合係数の空間パターンが類似してしまい、明瞭に解釈できる結果が得られなかった。

### (4) 人口増減の空間分布を系列とみなした計算結果

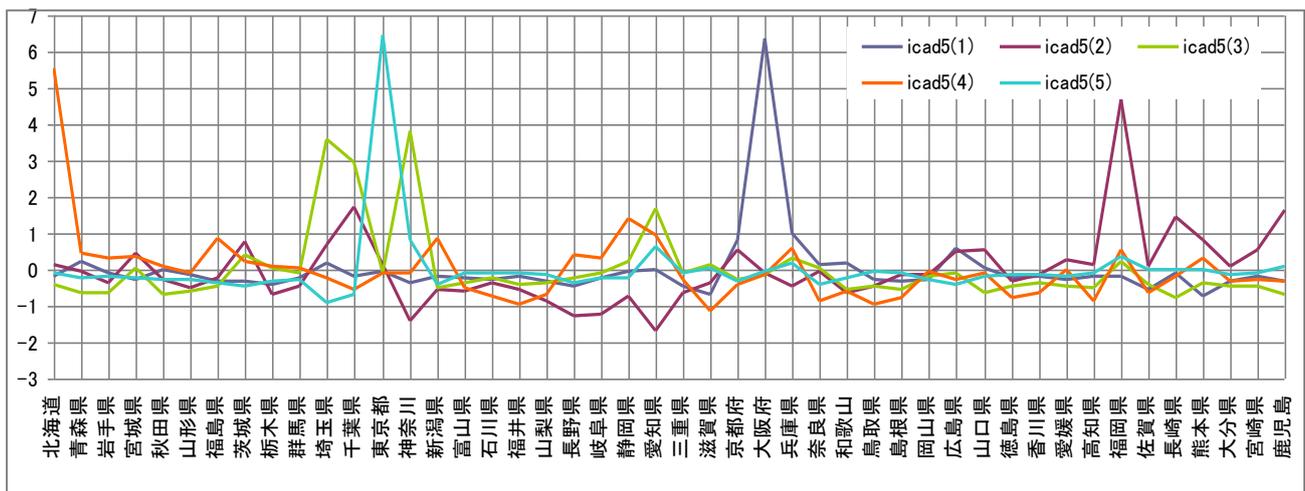


図-5 人口増減の空間パターンを系列としたときの分析結果の例（独立成分）

人口増減のパターンに着目することにより、人口が増加している地域と同時に、同時期に人口が減少している地域を抽出することにより、社会移動の大きな流れをつかむことが可能となっている。

例えば5成分を抽出したときの空間分布成分を図-5に示す。グラフの負の部分に着目することにより、大阪に集中する人口は三重・滋賀のほか九州北部の県からの移動が大きいことなどが推察できる。

この計算方法で10成分の分析を行ったものの、やはり災害の影響を直接表わすような成分は抽出されなかった。一方、混合率を用いて独立成分を再合成し、観測系列との差異を計算したが、(2)のケースに比べて残差が極めて大きく、自然災害との対応をつけるのが困難であることがわかった。

#### 4. おわりに

本研究は、災害の長期的な影響の広がり把握することを目指して、長期人口時系列データに独立成分分析手法を適用した。通常的人口統計表を転置して、人口の空間分布パターンを系列と考えることで、現実的な解釈ができる成分を抽出することができた。また、人口増減の空間分布に同様に手法を用いると、人口の社会移動のパターンがより明確に抽出できるが、残差は大きくなることがわかった。残念ながら、抽出された成分には戦争の影響は強く現れているものの、他の自然災害の影響は見られない。一方1995年の兵庫県のよう、阪神大震災の影響が負の残差として抽出できる可能性があることがわかった。

今後の課題としては、データ上の改良と、手法上の改良がある。現在インターネットでは、1955年以降の時点については都道府県ごとに、5歳階級別人口が公開され

ている。これと生命表の情報を組み合わせれば、自然増減の部分の推定できるので、人口変化の中から社会増減部分を取り出して分析を行うことが可能となる。また総人口については、市町村別の人口推移のデータも入手可能である。以上のようなより詳細なデータを活用して分析を進めることが望まれる。

一方手法上では、制約条件をつけた独立成分分析手法<sup>1)</sup>の適用が考えられる。すなわち、災害の発生時刻あるいは発生場所を人口データと並べて入力すると共に、これらの変数の混合比を1に制約することにより、強制的にその動きに対応する成分を導入することができる。その成分に対応する実際の人口への混合比を確認すれば、災害の影響を把握できる可能性がある。

#### 参考文献

- 1) 河田恵昭・柄谷有香：大規模な人命の損失に伴う社会的価値の損失の評価，土木計画学研究・論文集，No.17,pp.393-400,2000.
- 2) 中島奈緒美，吉村美穂，目黒公郎：マルチハザードリスクから見た地域別潜在的災害リスク特性の評価，土木学会第60回年次学術講演会，4-276,2005.
- 3) 外谷英樹：防災政策による災害被害の軽減効果：都道府県別データを用いたパネル分析，内閣府経済社会総合研究所，経済学的視点を導入した災害政策体系のあり方に関する研究報告書，pp.67-89,2009.
- 4) 総理府統計局 HP：国勢調査長期人口データ，<http://www.e-stat.go.jp/SG1/estat/>（2011.7.1現在）
- 5) Aapo Hyvarinen, Juha Karhunen, Erkki Oja：詳解 独立成分分析，東京電気大学出版社，2005.
- 6) 加藤寛道・木村裕介・奥村誠・塚井誠人：仙台空港発着便における遅れ・欠航の分析，土木計画学研究・論文集，Vol.25, No.3, pp.723-729, 2008.
- 7) 塚井誠人・井上真一・奥村誠：データマイニングによる広島空港流入車両の時間分布パターンの分析，交通工学研究発表会論文報告集，No.28, pp.269-272, 2008.
- 8) 井上真一・塚井誠人・桑野将司・奥村誠：独立成分

- 分析による非日常時の地方空港利用交通の分析, 土木計画学研究・論文集, Vol.26, No.3, 583-593, 2009.
- 9) フォンセカ カルロス ナバ・奥村誠・塚井誠人: バス利用計数データによる潜在需要分析の試み, 都市計画論文集, Vol.44(3), pp.469-474, 2009.
- 10) 内閣府: 平成 23 年度防災白書, <http://www.bousai.go.jp/hakusho/hakusho.html> (2011.7.20 現在)
- 11) 梅山伸二・赤穂昭太郎・菅生康子: 教師付き独立成分分析の提案とその顔画像分析への応用について, 信学技法, NC-99-2, pp.9-16, 1999. (2011.8.5受付)

## EXTRACTING THE EFFECT OF HUGE SCALE NATURAL DISASTERS FROM A LONGITUDINAL POPULATION STATISTICS

Shun FURUSAHI and Makoto OKUMURA

In order to design the recovery and restoration plan and scheme for a huge scale natural disasters, spatial scope and time scale of the effects of that natural disaster should be foresighted. Especially, financial choice must be strongly dependent of time scale of the effects. This paper tries to propose a simple method to extract disasters' effects from a longitudinal socio-economic statistics, using independent component analysis. The ICA method is applied to the prefectural population time series by the Japanese Population Census, between 1920-2010.