

特集 I : 地域別将来人口推計

地域別将来推計人口の精度評価

一回帰モデルによる推計誤差の空間的特性の検証—

鎌田健司・小池司朗・菅桂太・山内昌和*

本稿は国立社会保障・人口問題研究所が実施・公表する「地域別将来推計人口」の精度評価として、推計結果とその後の実績値との差（推計誤差）に関する2種類の誤差率（APE, ALPE）を用いて、これまでの傾向ならびに回帰モデルによる推計誤差の空間的特性の検証を行った。分析対象は、都道府県は1985～2010年、市区町村は2000～2010年を基準とする推計の結果である。

推計誤差のこれまでの傾向は、近年になるほど少子高齢化の影響を受けて推計誤差が縮小傾向にあること、基本的属性からみると人口規模が小さくなるほど誤差率が高くなる傾向であること、東京圏を中心とする大都市圏では誤差率は小さいものの過小推計になる傾向が観察され、奈良県などの特定の県の中山間地域など小規模な地域自治体では過大推計になることが明らかとなった。

推計誤差の空間的特性の検証を行った結果、(1)推計誤差の空間的自己相関については、都道府県では推計期間が短い場合に検出され、市区町村別では明確な空間的従属性が存在することが明らかとなった。したがって市区町村別の推計誤差を回帰モデルで評価する場合は、空間誤差モデルなど空間計量経済モデルを用いる必要がある。(2)空間的異質性については、独立変数の偏回帰係数が地域によって異なるといった傾向は検出されなかったため、ローカル推定を行う必要性はないことがわかった。

【キーワード】地域別将来推計人口、推計誤差、空間的自己相関、空間的異質性

I. はじめに

一般に、将来推計人口は過去の趨勢を将来に投影（projection）するものであるため、推計結果とその後の実績値との差（推計誤差）は「誤り」となるわけではない。しかし、地方自治体を対象とした国立社会保障・人口問題研究所（以下、社人研）の「地域別将来推計人口」（以下、地域推計）は、都市計画、地方人口ビジョンなど各種計画の基となる予測値（forecast）として利用されており、推計誤差の情報それ自体が利用者にとっては有用な情報になりえる。これらを踏まえるならば、推計誤差について検討することは有意義な試みであるといえる（山内・小池 2015）。

従来地域別にみた将来推計人口の推計誤差に関する研究は二つに分けることができる。一つは、推計値と実績値の誤差率をもとに、性別や年齢、人口規模、推計期間等を考慮し

* 早稲田大学教育・総合科学学術院

つつ推計誤差率の統計的特性（平均・中央値・10パーセンタイル，90パーセンタイルなど）について検討する研究である（Wilson 2012, Smith et al. 2013, Statistics New Zealand 2016, Yamauchi et al. 2016など）。このうち山内・小池（2015）と Yamauchi et al.（2016）では，平成17（2005）年までの国勢調査をもとにした社人研の地域推計の誤差の傾向の整理及び諸外国の推計結果との比較を行い，近年の推計になるほど推計誤差が縮小傾向に推移していることや諸外国に比べて社人研の地域推計の推計誤差は大きくないことなどを示している。

もう一つは，本稿が関心を寄せる回帰モデルを用いて推計誤差率の規定要因を検討する研究である（Tayman and Swanson 1996, Tayman et al. 2011など）。たとえば Tayman et al.（2011）は，推計誤差率を従属変数とし，推計対象地域の人口規模，人口増加率，推計期間などを独立変数とした最小二乗法（OLS）によるモデル推定を行うことによって各独立変数の影響の評価を行っている。これらの研究を通じて，他の変数の影響を統制した上での独立変数の効果を定量的に評価できる点や独立変数間の影響の強弱に関する情報を得ることができる。ただし，OLSによるモデル推定を地域データに適用する場合，空間的な特性を適切に考慮できなければモデル推定の結果を歪めることになりかねないにもかかわらず，既存研究はその点を十分に考慮できていなかった。

OLSに代表される回帰モデルを地域データに適用してモデル推定を行う場合，回帰モデルの誤差項に独立変数との空間的自己相関が生じてしまい，推定値にバイアスを生じさせることがある。ここで空間的自己相関とは，地域間の距離が近いほど似たような値になるという空間的特性の一種である。空間的自己相関がモデルの誤差項と独立変数の間に観察されると，OLSが仮定する誤差項が独立かつ同一の分布に従うという仮定（i.i.d.）が満たされないことから，推定値にバイアスを生じさせてしまう。このような地域データ特有の問題を処理するためには，空間計量経済学的手法を導入する必要がある（Anselin 2005）。一般に，人口移動の発生頻度は移動距離と逆相関し，特定地域に人口集中が起こるとき，その近隣地域で移動率が高くなることが考えられる。また，人口移動は必ずしも行政区域単位で単一で起こるわけではなく，連続的に変化していると考えられる。そのため，行政区域単位で測定した移動率の誤差には空間的な相関等の非独立性が生じている可能性があり，そのような影響が推計誤差の空間分布に及ぼされることが考えられる。

そこで本稿では，回帰モデルを用いた推計誤差の評価研究に手法的改善を提案することを目的として，回帰モデルを用いた地域推計の推計誤差の評価を行う際に課題となる以下の二つの空間的特性に着目し，推計誤差の空間的特性の有無及びその対処方法について検討する。

空間的特性の一つ目は空間的自己相関（Spatial Autocorrelation）である。空間的自己相関は，上述したように地域の距離が近いほど似たような属性値を示す正の空間的自己相関と距離が近いほど異なる属性値を示す負の空間的自己相関に代表される空間的特性である（Anselin and Bera 1998）。正の空間的自己相関は Tobler（1970）の地理学における第一法則として知られており，空間分布の最も基本的な特性といえる。空間的自己相関

のあるデータを用いて回帰モデルを行う場合の空間計量経済モデルによる対処には、誤差項に空間的隣接行列による加重を用いて空間的自己相関を調整する空間誤差モデル (SEM : Spatial Error Model) がある。

もう一つの空間的特性は空間的異質性 (Spatial Heterogeneity) である。空間的異質性は、主に統計モデルとして評価する場合に推定値の空間分布が安定せず、モデルの誤差項の分散が空間的に不均一となる特性である (Anselin 1988, 瀬谷・堤 2014)。すなわち OLS によって推定された偏回帰係数が地域によって異なるという空間的非定常性 (spatial non-stationarity) が存在することを意味し、偏回帰係数の空間的分布を許容するローカル推定を行う地理的加重回帰モデル (GWR : Geographically Weighted Regression) によってその検証が可能である (Brunsdon et al. 1996, 1999, Fotheringham et al. 2002)。

本稿の構成は下記の通りである。Ⅱでは社人研が実施する地域別将来推計人口の概要と推計誤差の傾向、推計誤差の精度評価の先行研究について説明する。Ⅲでは分析計画として、空間的特性を検証するための統計的検定指標、統計モデルを説明し、Ⅳでは分析結果を示す。Ⅴでは全体のまとめと今後の課題を示す。

Ⅱ. 社人研が実施する地域別将来推計人口の概要と推計誤差の傾向

1. 地域別将来推計人口の概要

社人研の地域推計は、旧厚生省人口問題研究所が実施した1985年の国勢調査を基準人口とした「昭和62年1月推計」から開始され、2015年の国勢調査を基準人口とした「平成30年推計」で都道府県推計としては7回目となる。市区町村推計については2000年の国勢調査を基準人口とした「平成15年12月推計」から開始され、「平成30年推計」で4回実施されている (表1)。

市区町村推計の対象地域数は、2000年の国勢調査時点では3,245市区町村あったものが、2015年の国勢調査では2005年前後の平成の大合併の影響もあり1,798市区町村 (東京23区 (特別区) および12政令市の128区, 766市, 713町, 168村) と推計年次によって大きく異なる。2010年の国勢調査を基準とした「平成25年3月推計」及び2015年の国勢調査を基準とした「平成30年推計」では市区町村別の推計結果を都道府県ごとに積み上げた結果を都道府県推計結果と同時に公表した。ただし、両推計では、福島県は2011年の東日本大震災の原子力発電所事故による避難指示区域の設定や人口移動傾向を見通すことが困難であるため市町村推計を行わず県全体のみの公表である。

推計手法はいずれの推計もコーホート要因法である。出生や死亡の将来の動向については、全国と地域の相対的較差を用いて全国の動向と連動した仮定値が設定されている。出生仮定は、1985～2005年の国勢調査を基準とした都道府県推計では年齢別出生率を用い、2010・2015年の国勢調査を基準とした市区町村別推計では、子ども女性比を用いた推計を行った上で、市区町村別人口の合計を都道府県推計結果としている。そのため、当該年に

については都道府県の年齢別出生率の作成は行っていない。移動仮定は、原則として過去5年間の動向を基に平成25年3月推計までは純移動率モデルが用いられ、平成30年推計では多地域モデルの1つである「プールモデル」(小池 2015, 2016, 国立社会保障・人口問題研究所 2018) による推計がなされた。推計の結果は、基本仮定に沿った推計結果(本推計)と、純移動率が0とする封鎖人口を仮定した推計結果が公表されている。その他、地域推計実施の経緯や詳細は山内・小池(2014)を参照されたい。

表1 厚生省人口問題研究所・社人研による地域別将来推計人口の概要

実施機関	公表年月	対象地域	地域数	基準年	推計期間
厚生省人口問題研究所	1987年1月	都道府県	47	1985年	40年
厚生省人口問題研究所	1992年10月	都道府県	47	1990年	20年
国立社会保障・人口問題研究所	1997年5月	都道府県	47	1995年	30年
国立社会保障・人口問題研究所	2002年3月	都道府県	47	2000年	30年
国立社会保障・人口問題研究所	2007年5月	都道府県	47	2005年	30年
国立社会保障・人口問題研究所	2013年3月	都道府県	47	2010年	30年
国立社会保障・人口問題研究所	2018年3月	都道府県	47	2015年	30年
国立社会保障・人口問題研究所	2003年12月	市区町村	3,245	2000年	30年
国立社会保障・人口問題研究所	2008年12月	市区町村	1,805	2005年	30年
国立社会保障・人口問題研究所	2013年3月	市区町村	1,799	2010年	30年
国立社会保障・人口問題研究所	2018年3月	市区町村	1,798	2015年	30年

資料：厚生省人口問題研究所(1987; 1992)。

国立社会保障・人口問題研究所(1997; 2002; 2003; 2007; 2008; 2013; 2018)

注：市区町村推計において、基準年2010年及び2015年の地域数には、12政令市(北海道札幌市、宮城県仙台市、千葉県千葉市、神奈川県横浜市、神奈川県川崎市、愛知県名古屋市、京都府京都市、大阪府大阪市、兵庫県神戸市、広島県広島市、福岡県北九州市、福岡県福岡市の128区)の区別の推計結果が含まれる。

2. 地域別将来推計人口の推計誤差の傾向

推計誤差の評価に用いる指標は誤差率 ALPE (Algebraic Percent Error) と誤差率の絶対値 APE (Absolute Percent Error) である。ALPE は推計値と実績値の差を実績値で除した誤差率 (%) であり、「バイアス」(bias) の指標、APE は ALPE を絶対値で示した「正確性」(precision) の指標として以下のように定義される (Tayman et al. 2011)。これらの指標は Smith (1987) や Smith and Shincich (1988) をはじめとする推計精度の評価研究において一般的に用いられる指標である。ALPE は正負の符号を持ち、正の推計誤差は推計値が実績値を上回る過大推計を、負の推計誤差は推計値が実績値を下回る過小推計を示す。

$$ALPE_{x,i} = \frac{e_{P_{x,i}} - a_{P_{x,i}}}{a_{P_{x,i}}} \times 100$$

$$APE_{x,i} = \frac{|e_{P_{x,i}} - a_{P_{x,i}}|}{a_{P_{x,i}}} \times 100$$

P : 総人口, e : 推計値, a : 実績値, x : 年齢, i : 都道府県もしくは市区町村

本稿では、集計及び分析対象となる推計は、都道府県は基準年を1985～2010年とするもの（以下、基準年で区別してX年基準推計と呼ぶ）、市区町村は2000～2010年である。市区町村推計における対象地域数は、統計量の集計と回帰モデルで一部異なる。詳細は各図表にて明示する。

表2は基準年別、推計期間別の都道府県 APE について、中央値と90パーセンタイルの値を示している。APE の傾向は推計期間が短いほど誤差率は小さく、さらに近年の推計になるほど小さい。地域推計の結果は移動率による変動が大きいため、原則的に直近5年間の移動傾向を将来に投影している社人研の推計においては、移動傾向に変化が生じると推計期間が長くなるに従って誤差率が上昇する。また、近年ほど誤差率が低い点については、少子高齢化によって誤差率の高い0～4歳人口及び10代後半から20代前半の人口が減少傾向にあり、総人口に及ぼす影響が縮小しているためと考えられる（山内・小池 2015）。

都道府県別にみた APE の範囲は、推計期間5年で0.3%～1.4%、推計期間10年では1.0%～2.1%、推計期間15年では2.1%～2.8%、推計期間20年では2.9%～4.5%、推計期間25年は6.6%、推計期間30年は8.5%である。

1990年基準年推計はバブル経済の影響で国内人口移動の傾向がそれ以前と異なっていたこともあり推計期間を20年とし、年齢別純移動率の水準を長期的に低下させる仮定を導入している（山内・小池 2014）。その結果、短期的な誤差率が高いものの、長期的には1985年・1995年と比べて誤差率の中央値は低い水準である。ただし、90パーセンタイルの数値についてみると1990年基準推計の誤差率は他の推計に比べてやや高めであり、一部の地域では大きな誤差が生じており誤差分布のばらつきが大きい。

表3には市区町村の APE の中央値・90パーセンタイル値の範囲について、基準人口の規模を1万人未満、1万人～10万人未満、10万人以上の3区分で示している。市区町村の APE の中央値は推計期間5年で1.2%～1.7%、推計期間10年で2.4%～3.5%、推計期間15年は5.7%である。

市区町村の APE の中央値を基準人口の規模別にみると、人口規模が小さくなるほど APE は大きくなる傾向がみられる。推計期間5年では基準人口の規模が1万人未満では1.9～2.1%、1万～10万人未満では1.0%～1.5%、10万人以上では0.9%～1.3%、推計期間10年では基準人口の規模が1万人未満では4%程度、1万～10万人未満では2.2%～3.1%、10万人以上では1.8%～2.7%、推計期間15年では、1万人未満では6.8%、1万～10万人未満では5.2%、10万人以上では4.0%であった。

表2 基準年別，推計期間年別，都道府県別にみた APE の中央値・90パーセンタイル値

基準年	中央値 (APE, %)						90パーセンタイル (APE, %)					
	推計期間 (年)						推計期間 (年)					
	5年	10年	15年	20年	25年	30年	5年	10年	15年	20年	25年	30年
1985年	0.9	1.5	2.8	4.5	6.6	8.5	1.9	3.0	5.8	9.5	11.4	15.1
1990年	1.4	2.1	2.5	2.9			2.3	3.9	8.1	10.2		
1995年	0.6	1.5	3.1	4.4			1.9	4.5	6.7	10.0		
2000年	0.6	1.3	2.1				1.6	3.4	6.0			
2005年	0.5	1.0					1.2	2.3				
2010年	0.3						1.1					

表3 基準年別，推計期間年別，人口規模別，市区町村別にみた APE の中央値・90パーセンタイル値

基準年	地域数	推計期間 (年)	中央値 (APE, %)				90パーセンタイル (APE, %)			
			基準人口の規模 (人)				基準人口の規模 (人)			
			総数	1万人未満	1万～10万人未満	10万人以上	総数	1万人未満	1万～10万人未満	10万人以上
2000年	3,244	5	1.7	1.9	1.5	1.3	4.8	5.5	4.2	3.6
2005年	1,805	5	1.3	2.1	1.2	1.0	4.1	6.8	3.3	3.0
2010年	1,682	5	1.2	2.0	1.0	0.9	3.8	5.7	3.1	2.9
2000年	3,244	10	3.5	4.0	3.1	2.7	9.4	10.9	8.1	7.5
2005年	1,796	10	2.4	4.1	2.2	1.8	7.9	11.6	6.7	5.6
2000年	3,232	15	5.7	6.8	5.2	4.0	15.2	17.9	12.9	10.9

注：基準年2000年の新潟県山古志村は2004年の中越地震により多大な被害を受けたため，集計から除外した。
 基準年2010年で区別に推計されている12政令市は市単位で集計した。
 基準年2010年（推計期間5年）では，福島県内市町村の推計が行われていない。
 基準年2005年（推計期間10年）では，原発事故による避難準備区域等による指定を受けた福島県9町村（広野町，楡葉町，富岡町，川内村，大熊町，双葉町，浪江町，梶尾村，飯館村），3地区（南相馬市鹿島地区・小高地区，田村市都路地区）を除いて集計した。
 基準年2000年（推計期間15年）では，上記の福島県9町村を除いて集計した。

表4には基準年別，推計期間別，都道府県別の将来推計人口のALPEを示している。特徴的な結果としては，東京都における過小推計（推計人口<実績人口），奈良県における過大推計（推計人口>実績人口）が続いている点がある。市区町村別にみると東京都は特別区（千代田区，中央区，港区など）において過小推計となっている他，奈良県では南部の中山間地域（上北山村，下市町，黒滝村など）において過大推計となっている傾向が観察される。茨城県，埼玉県，千葉県，神奈川県といった東京周辺県は推計期間が15年より短い場合に過小推計となっており，近年の東京圏一極集中の傾向が2010～15年でより一層進んだ結果を反映している一方で，推計期間が20年以上では過大推計の傾向がみられ，1990年代における郊外化の傾向が2000年代に入り落ち着き，一転，都心回帰の傾向を示していることによる影響がうかがえる。

1990年基準推計の結果は他の推計結果と異なる誤差率の傾向がみられ，首都圏や愛知県，兵庫県などで過大推計，その他の道府県は過小推計であり，バブル経済とその崩壊後に国内の人口移動傾向が変化したことの影響がでているものと考えられる。また，2010年基準

表4 基準年別、推計期間別にみた都道府県別 ALPE

(%)

基準年 推計期間	1985年	1990年	1995年	2000年	2005年	2010年	1985年	1990年	1995年	2000年	2005年
	5年	5年	5年	5年	5年	5年	10年	10年	10年	10年	10年
北海道	1.6	-2.2	0.3	0.1	0.1	-0.4	1.5	-3.4	0.8	0.6	-0.4
青森県	2.0	-3.5	-0.5	2.0	0.9	-0.2	1.0	-6.8	0.8	5.2	1.6
岩手県	1.0	-2.1	-0.3	1.7	0.9	-1.0	0.2	-3.9	0.9	4.5	0.9
宮城県	0.1	-1.3	1.2	1.6	-0.6	-1.2	-0.2	-0.8	3.8	2.8	-1.8
秋田県	0.8	-2.1	0.0	1.3	0.7	0.0	-0.2	-3.6	1.1	3.5	1.4
山形県	0.0	-1.4	0.2	0.8	0.8	-0.7	-0.7	-2.1	1.2	2.6	0.9
福島県	0.1	-1.4	1.0	1.0	0.5	-0.1	-0.6	-1.1	3.0	2.7	3.2
茨城県	0.7	-0.5	2.4	1.1	-1.2	0.2	1.9	2.0	5.9	1.2	-1.5
栃木県	-0.7	0.2	1.0	0.2	-0.1	0.0	-0.5	2.1	2.2	0.7	0.2
群馬県	0.5	-0.2	0.5	0.4	-0.3	-0.1	1.5	0.6	1.7	0.7	-0.6
埼玉県	-2.1	3.0	2.3	0.6	-1.6	-0.8	-0.2	9.5	5.4	-0.1	-3.2
千葉県	-0.5	2.8	1.6	-0.3	-1.7	-0.5	3.0	8.1	2.9	-2.0	-2.2
東京都	1.2	0.5	-4.2	-2.2	-1.9	-1.2	3.1	-3.1	-10.4	-5.5	-3.4
神奈川県	-0.6	3.4	-0.4	-1.1	-1.0	0.2	3.2	6.9	-1.9	-2.6	-1.2
新潟県	0.2	-1.5	0.6	0.9	-0.4	-0.3	-0.6	-2.0	2.1	1.7	-0.8
富山県	0.1	-0.8	0.3	0.2	-0.3	-0.2	0.0	-1.1	0.7	0.4	-0.8
石川県	0.9	-0.8	0.8	0.0	-1.3	-0.1	1.6	-0.5	1.8	-1.1	-2.3
福井県	1.4	-0.5	-0.2	0.6	0.1	-0.2	3.0	-0.7	0.1	1.3	0.1
山梨県	0.6	-1.8	2.3	0.5	1.0	0.3	0.3	-0.8	5.2	2.3	2.2
長野県	0.7	-1.4	0.6	1.1	0.1	-0.4	0.5	-1.7	2.7	2.5	-0.2
岐阜県	0.8	-0.5	0.9	-0.1	0.1	0.2	1.8	0.3	1.8	0.1	0.4
静岡県	0.0	0.1	0.6	-0.5	0.2	-0.1	0.9	1.5	0.9	-0.5	0.3
愛知県	-0.4	0.3	-0.5	-1.3	-0.6	-0.2	0.4	0.8	-2.1	-2.8	-1.2
三重県	0.2	-0.9	1.6	-0.3	0.0	0.3	0.1	0.1	3.0	-0.3	0.4
滋賀県	0.1	-0.4	0.6	1.1	-0.7	0.5	0.8	0.4	2.6	2.2	-0.5
京都府	1.0	-0.7	-0.4	0.3	-0.3	0.2	1.6	-1.1	-0.7	0.4	-0.8
大阪府	0.9	-0.4	-1.5	-0.4	-1.5	-0.4	2.1	-0.2	-3.4	-2.2	-2.9
兵庫県	-0.5	1.9	0.6	0.5	-0.4	0.0	1.8	1.2	2.2	1.0	-0.9
奈良県	1.3	0.2	2.5	2.1	-0.8	0.4	4.1	3.7	7.3	3.4	-1.2
和歌山県	0.4	-2.2	1.0	1.8	-0.8	-0.2	-0.8	-2.9	4.0	2.9	-1.6
鳥取県	0.9	-1.0	-0.5	0.6	1.2	-1.1	1.4	-1.6	-0.5	2.7	1.2
島根県	1.9	-1.4	-0.5	1.0	-0.1	-1.0	3.0	-2.9	0.1	2.2	-0.9
岡山県	1.2	-1.5	0.9	-0.5	-0.2	-0.4	1.5	-1.6	1.3	-0.7	-0.6
広島県	0.9	-0.7	0.8	-0.3	-0.7	-0.7	2.1	0.2	1.2	-0.9	-2.1
山口県	1.8	-1.7	0.1	0.4	-0.5	-0.4	2.7	-2.5	0.3	0.5	-1.3
徳島県	0.7	-1.4	0.6	0.5	0.4	0.0	0.6	-1.9	1.5	1.6	0.8
香川県	1.1	-1.2	0.4	0.3	-0.5	-0.7	1.8	-1.5	1.2	0.4	-1.4
愛媛県	1.6	-1.3	-0.1	0.5	-0.2	-0.1	2.5	-2.1	0.3	1.2	-0.4
高知県	1.7	-1.7	-1.1	1.6	0.9	0.2	2.6	-3.9	-0.8	4.5	1.9
福岡県	0.8	-1.5	0.1	0.8	-0.8	-1.1	1.0	-2.1	1.0	1.3	-2.4
佐賀県	1.3	-2.0	0.9	0.0	0.0	-0.6	1.2	-2.6	1.8	0.3	-0.4
長崎県	1.3	-1.7	-0.1	0.3	0.3	-0.5	1.6	-2.6	0.3	1.0	0.1
熊本県	1.7	-1.9	0.3	0.7	-0.4	-0.6	2.2	-3.0	1.1	1.2	-1.1
大分県	1.8	-1.5	-0.2	-0.2	-0.9	0.3	2.7	-2.8	-0.7	-1.0	-1.0
宮崎県	1.3	-2.4	0.5	0.8	-0.7	0.3	1.0	-3.6	1.6	1.1	-0.8
鹿児島県	2.3	-2.0	-0.7	1.2	0.1	0.1	3.3	-3.9	-0.3	3.0	0.5
沖縄県	2.1	-1.6	0.0	-0.5	0.1	-1.6	4.0	-2.1	0.0	-0.5	-1.2

表4 基準年別、推計期間別にみた都道府県別 ALPE (つづき) (%)

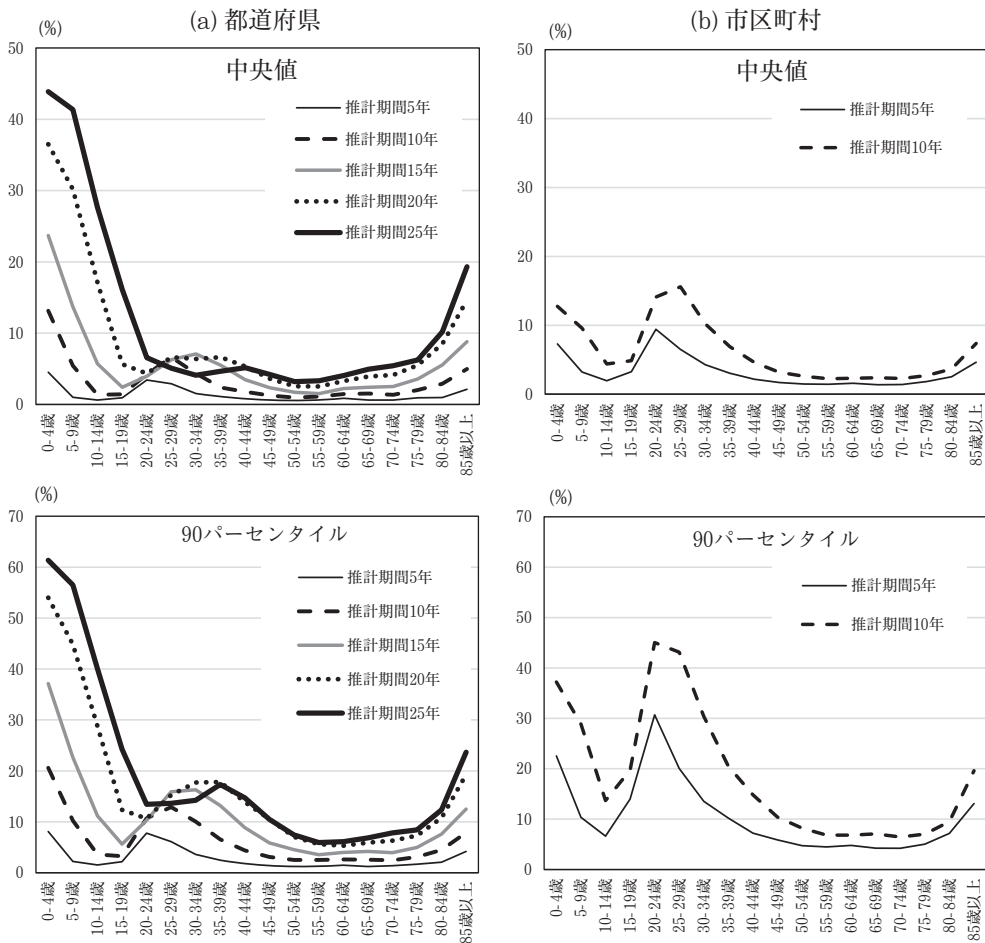
基準年 推計期間	1985年	1990年	1995年	2000年	1985年	1990年	1995年	1985年	1985年
	15年	15年	15年	15年	20年	20年	20年	25年	30年
北海道	2.6	-3.8	1.9	0.4	4.0	-3.6	2.0	5.5	5.9
青森県	0.2	-7.8	3.4	8.1	1.0	-7.6	5.4	2.6	3.3
岩手県	-0.3	-3.7	3.4	6.6	0.6	-2.3	4.9	2.2	2.5
宮城県	1.4	1.4	6.3	3.4	4.5	3.4	8.1	7.0	8.6
秋田県	-0.8	-3.8	3.3	5.5	-0.2	-2.8	5.2	0.9	1.6
山形県	-0.8	-1.6	3.3	3.7	-0.1	-0.1	4.7	1.4	1.8
福島県	0.3	0.5	5.9	6.8	2.0	3.0	11.1	4.3	8.5
茨城県	6.6	6.2	8.8	2.2	12.7	9.5	12.5	17.6	23.2
栃木県	1.6	4.6	3.8	1.7	3.7	7.6	5.8	5.9	8.1
群馬県	3.9	2.7	3.0	1.1	7.0	4.7	4.4	9.8	12.2
埼玉県	5.2	17.2	7.3	-0.7	11.5	23.5	9.0	15.7	19.4
千葉県	9.5	13.5	2.9	-2.1	15.8	17.2	4.4	20.2	25.8
東京都	0.7	-8.6	-17.1	-7.7	-4.1	-14.7	-22.5	-9.9	-14.5
神奈川県	7.5	9.4	-3.7	-2.8	10.4	11.2	-4.5	12.3	15.3
新潟県	-0.3	-1.1	3.7	2.4	0.9	-0.1	5.1	1.6	1.7
富山県	0.9	-0.7	1.3	0.4	1.9	-0.4	1.4	2.5	2.7
石川県	4.0	0.7	1.8	-2.1	6.8	0.7	2.0	8.2	9.4
福井県	5.1	0.1	0.9	1.9	7.8	1.3	1.2	10.6	12.9
山梨県	2.8	1.2	9.8	4.3	6.0	4.8	14.9	10.3	14.9
長野県	1.6	-0.3	5.5	3.7	4.1	1.8	7.9	6.6	8.4
岐阜県	4.6	1.6	3.1	0.3	7.2	3.1	4.4	9.9	12.6
静岡県	3.3	2.9	1.8	-0.5	5.2	4.8	2.6	7.3	9.0
愛知県	1.4	0.5	-3.7	-3.9	1.3	0.0	-5.2	0.6	-0.1
三重県	2.3	1.3	5.0	0.4	4.4	2.9	7.5	6.6	9.1
滋賀県	2.9	2.9	4.5	4.6	6.5	5.2	7.8	9.6	13.7
京都府	2.3	-1.1	-0.9	0.4	3.1	-1.1	-1.5	3.3	3.2
大阪府	3.7	-0.2	-6.3	-3.8	4.3	-1.6	-9.3	3.0	1.5
兵庫県	1.7	2.5	3.3	1.6	3.0	3.8	4.3	3.8	4.4
奈良県	10.9	9.9	11.6	5.1	20.5	15.6	16.5	29.4	39.0
和歌山県	-0.2	-1.4	6.3	3.8	2.1	-0.4	8.3	3.4	4.1
鳥取県	2.1	-1.6	1.0	3.8	3.3	0.1	1.7	5.8	6.7
島根県	4.2	-2.7	1.1	2.7	6.3	-2.2	1.0	8.4	9.0
岡山県	3.7	-1.8	2.0	-1.1	5.1	-1.9	2.4	6.3	7.0
広島県	5.1	1.0	1.4	-2.2	7.7	1.6	0.6	9.4	9.9
山口県	4.5	-2.7	0.4	0.2	6.6	-3.1	-0.1	8.0	8.5
徳島県	1.6	-1.9	3.2	2.8	2.8	-0.8	4.8	4.5	6.0
香川県	3.7	-0.9	1.8	0.0	6.0	-0.6	1.9	7.7	8.6
愛媛県	4.1	-2.1	0.8	1.8	6.1	-1.9	1.0	7.8	9.1
高知県	3.2	-4.3	0.9	7.5	5.4	-3.4	2.4	8.6	11.5
福岡県	2.3	-1.7	1.5	0.9	4.2	-1.7	1.2	5.5	5.5
佐賀県	2.8	-2.4	3.1	0.0	4.2	-1.9	4.0	5.6	6.0
長崎県	2.5	-2.7	1.2	1.0	3.6	-2.0	1.4	4.7	4.6
熊本県	3.6	-3.3	1.9	1.5	5.4	-3.5	2.3	6.7	7.6
大分県	4.1	-3.9	-1.5	-1.1	5.3	-5.2	-1.7	5.6	6.3
宮崎県	1.8	-3.6	2.6	1.9	3.4	-3.9	3.9	4.1	4.9
鹿児島県	4.3	-4.3	0.7	5.1	6.3	-4.1	1.9	8.4	10.4
沖縄県	6.7	-2.6	0.6	-1.7	8.9	-2.8	0.0	10.9	10.9

推計ではその直後に発生した東日本大震災の発生とその後の復興需要の影響により岩手県・宮城県で過小推計となっている（ただし、宮城県南三陸町24.8%，女川町17.9%では人口減少が進み過大推計となっている）。福島県の誤差率は-0.1%であった。

図1には推計期間別，5歳階級別，都道府県／市区町村別に見たAPEの中央値・90パーセンタイル値を示している。推計期間が5年から15年では0～4歳のAPEが最も大きく，次に20～30代となる。30代以上では，APEは低下するものの，80代以降に再び誤差が上昇する。都道府県推計の0～4歳のAPEの中央値は推計期間5年では4.5%，10年13.1%，15年23.7%，20年36.5%，25年43.9%と上昇する。0～4歳は出生仮定に関する誤差や推計時に置いた移動仮定と実際の移動状況との差によってもたらされるほか，郊外化や都心回帰といった地理空間的な移動トレンドの変化の影響も考えられる。

また，推計期間が長くなると，それ以前の0～4歳の誤差の影響が蓄積されることにより，19歳以下の誤差が20～30代の誤差よりも高くなる他，20～30代の誤差のピークが高齢

図1 推計期間別，5歳階級別，都道府県／市区町村別に見たAPEの中央値・90パーセンタイル値



注：市区町村の対象地域は表3を参照。

にシフトするといった傾向も読み取れる。20代の APE の中央値の傾向は推計期間 5 年から 25 年の変化は約 3% から約 6% の変化であり、高齢部分では、推計期間が大きくなるほど APE の中央値の上昇が大きくなり、85 歳以上では推計期間 5 年では 2.1% であったものが 25 年では 19.3% まで上昇し、近年の高齢者の移動傾向の変化等を反映しているといえる。

市区町村の推計誤差の年齢別の傾向は都道府県の傾向とおおむね同様ではあるが、0～4 歳の誤差よりも 20 代の誤差の方が高い等、移動が原因と考えられる推計誤差の影響が顕著にみられる。また、人口規模が小さいほど推計誤差が大きくなり、総じて都道府県推計よりも誤差は大きい。0～4 歳の APE の中央値は推計期間 5 年で 7.3%、推計期間 10 年では 12.8% である。25～29 歳では推計期間 5 年 6.5% から推計期間 10 年では 15.6% と APE の中央値が上昇する。

3. 推計誤差の精度評価に関する先行研究

Smith et al. (2013) では推計誤差の特徴として、1) 推計期間が長い、2) 対象となる地域単位が小さい、3) 人口変化が激しい地域、4) 年齢別にみると出生や人口移動が大きく影響する若年層において推計誤差が大きくなる、5) より新しい時期に実施された将来人口推計の方が、誤差が小さくなりやすいとは一概にはいえない、ということが指摘されている。

1) 推計期間については、基本的な知見として多くの研究で指摘されており、将来推計人口の基本的な構造が投影であることを考えるとその後の社会経済の変化が生じることで推計期間が長くなるほど実績値と過去の推計結果の乖離で測った推計誤差が大きくなりうることは想像に難くない（例えば、Murdock et al. 1984, Smith 1987, Smith and Sincich 1988, 1991, Tayman and Swanson 1996, Tayman et al. 2011, Lenze 2000, Rayer 2008, 山内・小池 2015 など）。

2) 対象地域の単位については、人口規模が小さくなるほど推計誤差が大きくなる言い換えることも可能であり 1) と同様多くの知見が存在する（より詳細には Smith and Tayman 2003, Smith et al. 2013 など）。人口規模が小さい地域は、小さい変動であっても誤差率が大きくなりやすく、移動傾向が変化することによって性・年齢別人口の構成が変化したときに大きな誤差が生じる。

3) 人口変化については、人口増加率が高い地域ほど推計誤差が大きくなるとする先行研究がある（Smith 1987, Smith and Shahidullah 1995, Tayman et al. 2011 など）。Tayman et al. (2011) は推計誤差と人口増加率の関係は U 字型の関係（二次関数）が観察され、人口増加率が低い地域と高い地域で推計誤差が大きくなることを明らかにしている。本分析においては人口増加率ではなく、より地域間の人口変動に直接的に影響すると考えられる直近 5 年間の純移動率を用い、推計誤差との関係性について検証する。

4) 年齢別にみた推計誤差については、山内・小池（2015）では、0～4 歳人口や 20～30 代で推計誤差が大きくなることが指摘されている。過去のトレンドに概ねしたがってスムーズに変化する死亡と比べ出生や人口移動は期間変動が著しく、ファミリー層や若者の

移動傾向の変化に関する誤差が反映されていると考えることができる。

5) 推計年次については、推計手法の違い、対象となる地域の人口構造（高齢化の度合い）や社会経済的環境（好景気・不況）、推計基準年におけるイレギュラーな事象の発生（経済的恐慌、自然災害、戦争・紛争等）、さらに本質的に将来に対する不確実性によって生じるため、その要因を明確にすることには困難が生じることが指摘されている（Chi 2009, Smith et al. 2013）。ただし、社人研の推計については、近年になるほど推計誤差は小さくなる傾向がみられており、少子高齢化による若者人口の減少に伴う移動者の減少がその要因であると考えられる（山内・小池 2015）。

Ⅲ. 分析計画：データと分析課題

1. データ

本分析において用いるデータは、総務省統計局『国勢調査報告』における実績値および旧厚生省人口問題研究所及び社人研の『地域別将来推計人口』の将来推計値である。都道府県は1985～2010年基準推計、推計期間は20年までを分析対象とする¹⁾。市区町村については2000～2010年基準推計、推計期間は15年までを分析対象とする。東京都特別区は区ごと、政令市は区を合算したものをを用いる。市区町村の対象地域数は国勢調査実施時点ではなく、推計結果の公表時点の数である。対象地域数は2000年3,244地域、2005年1,805地域、2010年1,746地域である。ただし、手法によって分析可能な地域数には若干の変動がある。詳細は各分析結果の注に記載した。

2. 分析課題

(1) 推計誤差の空間的自己相関の検証

本稿では上述の通り、ALPE と APE というやや性格の異なる指標を用いて複合的な推計誤差の評価を行う。ALPE と APE について同様の方法で推計誤差の評価を行う。その方法を説明する際の数式表記の都合上、両値とも y_i と表記する。推計誤差の空間的特性を検証するために二つの分析課題を設定した。第一の課題は、推計誤差の空間的自己相関の検証であり、第二の課題は推計誤差の空間的異質性の検証である。

推計誤差の空間的自己相関の検証は、はじめに推計誤差の空間的自己相関を定量化するモラン I の統計量を算出し、次に OLS モデルの推定を行い誤差項と独立変数の空間的自己相関の検出を行う。最後に空間的誤差モデルを推定することによって空間的自己相関の有無を評価し、赤池情報量規準 AIC によるモデルフィットの改善の有無を検証する。

空間的自己相関を定量化するモランの I 統計量は、 -1 から 1 までの範囲をとり、正の空間的自己相関は隣接する地域間の数値が近くなる傾向を示し、 0 は空間的にランダムに分布、負の空間的自己相関は隣接地域ほど値が異なる傾向を示し、下記のように定義され

1) 推計期間25～30年は1985年基準推計のみ推計誤差が算出可能であるが、回帰モデル分析において推計期間25～30年が一時点のみの観察となりパラメータを識別できないため、分析から除外した。

る (Moran 1950).

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

N : 都道府県/市区町村の地域数, w_{ij} : 空間加重行列

モランの I 統計量の統計的検定には, 一般的に漸近正規性を仮定した Z 検定がある (Cliff and Ord 1981). モランの I 統計量の期待値 $E(I)$, 分散 $Var(I)$ を標準化すると Z 値は下記のように定義される.

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}}, \quad E(I) = \frac{-1}{N-1}, \quad Var(I) = E(I^2) - E(I)^2$$

Z は漸近的に標準正規分布 $N(0,1)$ に従うため, 行標準化された空間行列下において空間的自己相関はないとする帰無仮説の検定を行うことができる. I が $E(I)$ より十分に大きい (小さい) とき正 (負) の空間的自己相関が観察される.

第一の分析課題の検証は空間計量経済学モデルの一つである空間誤差モデル (Spatial Error Model: SEM) による推定によって行う. 空間誤差モデルは誤差項どうしの空間的自己相関をモデル化するものであり, モデルの誤差項が独立変数と空間的自己相関が生じる場合, 誤差項が独立かつ同一の分布に従うという仮定 (i.i.d.) が担保されないという推定上の問題点に対処することが可能となる (Anselin 2005, 瀬谷・堤 2014).

$$\begin{aligned} Y &= X\beta + u \\ u &= \lambda Wu + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned}$$

空間誤差モデルでは, 誤差項 u に空間的自己相関構造をもつ λWu を導入している. λ は空間パラメータであり, 隣接行列を行基準化した場合, $1/\omega_{\min} \sim 1$ の範囲の値をとる. W は $n \times n$ の都道府県・市区町村間の隣接行列であり行和で標準化している. 隣接行列は queen 型 (全方位型) のラグ 1 (直接隣接する地域がある場合に 1, そうではない場合は 0) で都道府県・市区町村ともに設定した (図 2)²⁾. 北海道, 沖縄県や市区町村の島嶼部については, 必ずしも最近傍の地域と属性が近いとはいえ, 最近傍地域との距離もそれぞれ異なることから, 上記地域は独立した分析対象と仮定し, 隣接地域は無しとして隣接行列を作成した. ε は空間的自己相関構造をもたない誤差項である.

2) 隣接行列の型には, bishop 型 (斜め方位型), rook 型 (上下方位型), queen 型 (全方位) があり, 一般的に queen 型が用いられる. また, ラグ (隣接地の数) については, 直接的に隣接するラグ 1 のほかにラグ 2 (直接隣接地の隣接地を含む隣接行列) ~ ラグ 5 までについて検討した結果, ラグが増えるに従い空間的自己相関が低下するため, ここではラグ 1 (直接隣接) とした.

空間パラメータ λ の検定は尤度比 (Likelihood ratio: LR) 検定を行う。尤度比検定は下記のように定義される。

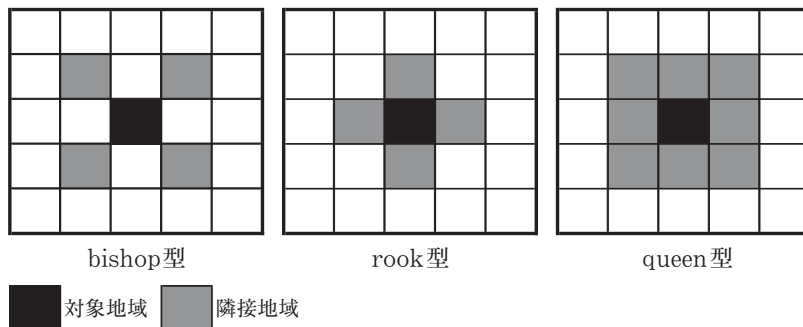
$$LR_\lambda = 2[\hat{l}_c - \tilde{l}_c]$$

\hat{l}_c は $\lambda = 0$ という制約がない状態での集約対数尤度、 \tilde{l}_c は $\lambda = 0$ という制約下での集約対数尤度である。 LR_λ が漸近的に自由度 1 の χ 二乗分布に従うことを利用する検定である。

回帰モデル全体の適合度を評価する赤池情報量規準 (Akaike's Information Criterion: AIC) は $AIC = -2\ln L + 2K$ によって定式化され (L は最大尤度, K はパラメータ数), AIC が小さいほどモデルの適合度が高いことを示す。

モデルの評価は OLS モデルで推定したときの誤差項の空間的自己相関の有無の検定及び、 λ の LR 検定結果, 赤池情報量規準 (AIC) による OLS モデルと SEM モデルのモデルフィットの比較, 推計誤差自体の空間的自己相関の有無によって行う³⁾。

図2 隣接行列型の種類



(2) 推計誤差の空間的異質性の検証

第二の課題は推計誤差の空間的異質性の検証である。ここで空間的異質性は、具体的には空間的非定常性 (Spatial Non-Stationarity) を対象とする。空間的非定常性とは、独立変数と従属変数の間の関係 (具体的には、回帰モデルの偏回帰係数) が地域によって異なることをいう。このような空間的非定常性を検証するためには、地域によって偏回帰係数が異なることを許容するローカル推定を行う必要があり、地理的加重回帰モデル (Geographically Weighted Regression, 以下 GWR) によってローカル推定ならびに空間的非定常性の検定が可能である⁴⁾。

3) 分析は R 3.5.3 を使用し、モランの I 統計量の算出・検定 (`moran.test`), SEM モデルの推定 (`errorsarlm`) ・検定は (`lm.morantest`) `spdep1.1-3` パッケージによって行った。

4) 地理的加重回帰モデルの推定 (`gwr.sel`, `gwr`), F 検定 (`LMZ.F1GWR.test`, `LMZ.F2GWR.test`, `LMZ.F3GWR.test`) は `spgwr0.6-32` パッケージによって行った。

GWRは、対象地域の緯度経度情報を用いて回帰モデルの係数の推定に空間的加重をかけることで係数の空間的なばらつきを表現するモデルである (Brunsdon et al. 1996, Fotheringham et al. 2002).

通常の回帰モデルは、すべての地点を含んで係数の推定を行うグローバル・モデルであり、推定された偏回帰係数はすべての地点で一定であるとする空間的定常性を仮定している。一方、GWRは係数が地点*i*ごとに計算されるローカル・モデルであるため、偏回帰係数が地点*i*ごとに異なるとする空間的非定常性を仮定している。

$$\text{通常の回帰モデル (OLS)} : y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i$$

$$\text{地理的加重回帰モデル} : y_i = \beta_0(i) + \beta_1(i)x_{1i} + \dots + \beta_p(i)x_{pi} + \varepsilon_i$$

以下は係数を行列形式で示したものである。通常の回帰モデルの係数の推定値に、位置*i*を中心に空間的に加重された $W(i)$ 関数を組み込んでいる。係数にかけられる空間的加重は距離が遠いほど空間的加重が小さくなる距離減衰関数である。このように、距離減衰関数によって加重された最小二乗法によって個々の推定値が求められる。

$$\text{通常の回帰モデルの係数} : \beta' = (X^T X)^{-1} X^T Y$$

$$\text{地理的加重回帰モデルの係数} : \beta'(i) = (X^T W(i) X)^{-1} X^T W(i) Y$$

Y は n 次元ベクトル、 X は $n \times p$ 行列、 $W(i)$ は $n \times n$ 対角行列であり、対角要素に i 地域からみた他の地域との距離に応じて減少する $w_j(i)$ を配置する空間的加重行列である。この空間的加重行列においては、対角成分は標本地点 i の空間的加重であり、その他の非対角成分は 0 である。GWR は空間的加重の分布にカーネル関数を用いていることに特徴がある。カーネル関数は距離が小さいとき空間的加重が大きくなる距離減衰関数を示すため、回帰地点から標本地点までの距離に応じて位置 i にかかる空間的加重が変化する。本分析ではカーネル関数のバンド幅の設定方法には固定型カーネル関数による推定を行った。また、カーネル関数にはガウス型を用いた。ガウス型カーネル関数は以下のように求められる。

$$w_j(i) = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right]$$

$w_j(i)$ は標本地点 i にかかる回帰地点 j からの距離減衰型の空間的加重、 d_{ij} は回帰地点 j と標本地点 i との距離、 h はカーネルのバンド幅を表す。最適なバンド幅を選択するために、モデル比較には CV (交差妥当化: Cross Validation) を用い、CV が最小となるバンド幅を最適であると判断する。CV は以下のように求められる。

$$CV = \sum_i [y_i - y_{\neq i}^*(h)]^2$$

$y_{\neq i}^*(h)$ はバンド幅 CV の範囲から除外された位置 i における予測値 y_i である。 CV が低いほど、よいモデルフィットであるといえる。

GWR によるモデル推定の評価には、ローカル係数の妥当性の検定があり、ローカル係数の t 検定を行うことができる。

GWR のモデル検定については、Leung et al. (2000) の $F(1)$ から $F(3)$ 検定がある。 Leung et al. (2000) では、GWR のモデル検定量として、OLS との比較を中心に、3 種類の F 検定による検定量を考案している。

$F(1)$ 検定は OLS の残差平方和と GWR の残差平方和 (RSS) の比が F 分布で近似されることを用い、 $F(1)$ 値が有意に小さければ GWR は OLS よりもモデルフィットがよいことを示す。

$$F(1) = \frac{RSS_{GWR}/\delta_1}{RSS_{OLS}/(n-p-1)}$$

n はデータ数、 p は推定するパラメータの数 $\delta_1 = \text{tr}[(\mathbf{I}-\mathbf{L})^T(\mathbf{I}-\mathbf{L})]$ であり、 \mathbf{I} は n 番目の推定値の単位行列、 \mathbf{L} は $n \times n$ の行列を示す⁵⁾。なお、 $\mathbf{L}\mathbf{Y}$ は GWR の線型予測を与える n 次元ベクトルに対応する。

$F(2)$ 検定は OLS の残差平方和と OLS から GWR へモデルを変更したときの改善度 ($DSS = RSS_{OLS} - RSS_{GWR}$) の比が F 分布で近似されることを用いる。 $F(2)$ 値が有意に小さければ OLS と GWR のデータとの適合度に統計的に有意な差がないことを示す。

$$F(2) = \frac{DSS/v_1}{RSS_{OLS}/(n-p-1)}$$

$v_1 = n - p - 1 - \delta_1$ を示す。

$F(3)$ 検定は、係数ごとに分散分析を行い、 $F(3)$ 値が大きければ、係数の地域差が統計的に有意であるということを示す⁶⁾。たとえば β_k については以下が F 分布で近似されることを用いる。

$$F(3) = \frac{v_k^2/\gamma_1}{\hat{\sigma}^2}$$

$$v_k^2 = \frac{1}{n} \sum_{i=1}^n \left(\hat{\beta}_{ik} - \frac{1}{n} \sum_{i=1}^n \hat{\beta}_{ik} \right), \gamma_1 = \text{tr} \left[\frac{1}{n} \mathbf{B}^T \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{B} \right], \hat{\sigma}^2 = \frac{RSS}{\delta_1}$$

\mathbf{J} はそれぞれの要素が 1 である $n \times n$ の行列を示す。なお、 $\mathbf{B}\mathbf{Y}$ が係数推定量を与える p 次ベクトルに対応する。

これらの 3 種類の F 検定によって、従来の回帰分析 (OLS) からの改善度を測定することができる。また、ローカルなモデルの当てはまり具合は測定するローカル R^2 値によって局地的な当てはまりを測定することができる。

5) \mathbf{L} は $\mathbf{x}_i^T(\mathbf{X}^T\mathbf{W}(i)\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}(i)$ $i = (1, 2, \dots, n)$ を i 行目に配置したもの。ここで \mathbf{x}_i は \mathbf{X} の i 行目に対応する n 次元ベクトルである。

6) \mathbf{B} は $\mathbf{e}_k^T(\mathbf{X}^T\mathbf{W}(i)\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}(i)$ $i = (1, 2, \dots, n)$ を i 行目に配置したもの。 \mathbf{e}_k は $k+1$ が 1、それ以外は 0 の列ベクトルを示す。

(3) 推定モデル

推定モデルは、(1)分析課題1の検証として都道府県／市区町村別に推計基準年、推計期間別に空間誤差モデルの推定を行う。(2)分析課題2の検証として市区町村別に推計基準年、推計期間別に地理的加重回帰モデルの推定を行う。

共変量は先行研究にならい以下のように設定した。1) 基準年の人口規模(対数化)について線形関数と二次関数を設定した。2) 基準年までの5年間の純移動率(絶対値)について線形関数と二次関数を設定した。3) 東京圏(東京都・埼玉県・千葉県・神奈川県)は純移動率が経時的に変動しやすく、近年は東京一極集中傾向にあり、時系列では過小推計になりやすい傾向にあるためダミー変数を用いる。

分析課題1は都道府県及び市区町村において各推計年次について空間的自己相関の検証を行う。また、分析課題2の空間的異質性の検証では市区町村の緯度経度情報を用いて偏回帰係数の空間的非定常性の検証を行う。なお、本稿の主目的は分析モデルの選択にあることや紙幅の関係から各分析結果表における共変量の推定値は示さない。推計誤差と共変量との関係は、II.2や3でみてきたように、推計期間が長くなるほど、人口規模が小さいほど、若年人口が多いほど、人口変化が大きいほど、推計誤差が大きくなるという傾向は変わらない。その上で、各統計量の検定を通じた空間的特性の検証を行っていききたい。

IV. 分析結果

1. 推計誤差の空間的自己相関の検証(都道府県／市区町村)

分析課題1の検証として、都道府県の結果を表5、市区町村の結果を表6に示した。各表では、APE及びALPE別に、推計期間、基準年と地域数、最適なモデル選択についての結果(最小二乗法OLSか空間誤差モデルSEMか)、その根拠となる統計量(1)~(4)を示している。(1)は誤差率に空間的自己相関があるかどうかの統計量と検定結果であり、統計的に有意であるとAPE、ALPEの空間分布に空間的自己相関があるかがわかる。さらに、OLSモデルによって誤差率と独立変数との関係に空間的自己相関によるバイアスが生じているのかを検定したのが(2)OLS残差が独立変数との間で空間的な自己相関があるかどうかの検定結果である。その上で、(3)空間誤差モデルの推定結果では、空間パラメータの統計的検定を通じて空間的自己相関の有無を判断する。最後に、(4)赤池情報量基準AICによるモデルの適合度を示した。

最適なモデル選択においては、(1)~(3)の検定結果から、(2)の検定結果が統計的に有意ではない場合は、OLSモデルを最適モデルとし、(3)の検定結果が統計的に有意である場合にはSEMを最適モデルとした。(2)が統計的に有意であっても、(3)が統計的に有意ではない場合は、最適モデルは選択できないとした。

推計誤差分布の空間的自己相関の有無については、APE、ALPEともに都道府県では推計期間と基準年によってばらつきがあり(表5)、市区町村では概ね空間的自己相関が検出される結果となった(表6)。

都道府県の APE モデルでは（表 5），1985年基準推計・1990年基準推計について推計期間が 5 年・10 年については空間的自己相関が検出されるがそれ以外の基準年・推計期間では検出されないという結果である。ALPE については，推計期間が 5 年では 1995 年基準推計・2000 年基準推計を除き空間的自己相関が検出される他，1985 年基準推計・1990 年

表 5 都道府県別 APE, ALPE に対する空間誤差モデルによるモデル選択結果

推計期間 (年)	基準年 (地域数)	最適なモデル		(1)誤差率の 空間的自己相関	(2)OLS残差の 空間的自己相関	(3)空間誤差 モデル	(4)赤池情報量 基準 AIC		
		最小二乗法 OLS	空間誤差 モデル SEM	モランの I 統計量	モランの I 統計量	空間パラメータ λの LR 検定	OLS	SEM	
A P E	5	1985年 (47)	●	0.463 **	0.450 **	0.600 **	99.7	85.9	
		1990年 (47)	●	0.166 *	0.169 *	0.354 +	100.4	99.3	
		1995年 (47)	●		0.354 **	-0.119	-0.432	100.9	100.3
		2000年 (47)	●		0.060	0.084	0.141	91.5	92.9
		2005年 (47)	●		0.261 **	-0.021	-0.040	49.6	51.5
		2010年 (47)	●		0.046	0.069	0.156	48.8	50.3
	10	1985年 (47)		●	0.275 **	0.220 **	0.385 *	141.2	138.8
		1990年 (47)		●	0.256 **	0.250 **	0.458 *	124.9	120.7
		1995年 (47)	●		0.390 **	-0.001	-0.006	186.9	188.9
		2000年 (47)		●	0.138 +	0.197 *	0.262	169.3	168.7
		2005年 (47)	●		0.082	-0.054	-0.103	116.4	118.2
	15	1985年 (47)	●		0.105	0.004	0.008	188.2	190.2
		1990年 (47)	●		0.359 **	-0.066	-0.255	181.4	182.6
		1995年 (47)	●		0.350 **	0.045	0.171	234.4	236.0
		2000年 (47)		●	0.091	0.176 *	0.259	215.6	215.3
	20	1985年 (47)	●		0.172 *	0.048	0.118	225.2	226.9
		1990年 (47)	●		0.454 **	-0.022	-0.099	237.2	239.1
		1995年 (47)	●		0.360 *	0.107	0.368	270.0	269.9
A L P E	5	1985年 (47)	●	0.431 **	0.342 **	0.475 **	118.3	111.5	
		1990年 (47)		●	0.402 **	0.172 *	0.375 +	128.6	127.1
		1995年 (47)	●		-0.046	-0.076	-0.185	123.6	124.9
		2000年 (47)		●	0.203 *	0.120 +	0.193	109.9	110.8
		2005年 (47)		●	0.287 **	0.320 **	0.503 **	91.7	84.8
		2010年 (47)		●	0.140 +	0.207 *	0.438 *	75.0	72.1
	10	1985年 (47)		●	0.320 **	0.225 **	0.407 **	154.4	151.5
		1990年 (47)		●	0.369 **	0.181 *	0.330 +	214.1	213.0
		1995年 (47)	●		-0.034	-0.067	-0.196	197.8	199.2
		2000年 (47)		●	0.254 **	0.195 *	0.275	185.2	184.6
		2005年 (47)		●	0.115	0.165 *	0.294	159.3	158.9
	15	1985年 (47)	●		0.168 *	0.078	0.177	194.4	195.7
		1990年 (47)		●	0.273 **	0.188 *	0.330 +	263.5	262.4
		1995年 (47)	●		-0.023	-0.028	-0.084	240.8	242.7
		2000年 (47)		●	0.193	0.182 *	0.265	230.0	229.6
	20	1985年 (47)	●		0.034	0.014	0.033	233.4	235.3
		1990年 (47)		●	0.200 *	0.163 *	0.299	292.6	292.1
		1995年 (47)	●		-0.009	0.015	0.047	272.4	274.3

有意水準: 0.01 **, 0.05 *, 0.1 +

注: (3)空間誤差モデルの数値は空間パラメータλの値を示している。

基準推計で推計期間10年において空間的自己相関が検出される。その他では、1990年基準推計・推計期間15年では空間的自己相関が検出されるが、その他の基準年・推計期間では検出されていない。

市区町村の結果をみると（表6）、APEモデル、ALPEモデルともに空間的自己相関が検出される結果となっている。都道府県よりも小さいレベルでの観察においては、誤差の正確性（APE）、バイアス（ALPE）ともに推計誤差は連続的に分布していることがわかる。

このように都道府県の誤差について回帰モデルによって推定を行う場合、都道府県モデルでは推計期間が短い場合はOLSの誤差項が独立変数と空間的自己相関している可能性があるため、その検定を行った上でモデル選択を行う必要がある。市区町村モデルについてはOLSモデルを適用せずにSEMモデルによるモデル誤差項の空間的自己相関の調整を行った上で適切なパラメータの推定を行うことが可能となる。

表6 市区町村別 APE, ALPE に対する空間誤差モデルによるモデル選択結果

推計期間 (年)	基準年 (地域数)	最適なモデル		(1)誤差率の 空間的自己相関	(2)OLS残差の 空間的自己相関	(3)空間誤差 モデル	(4)赤池情報量 基準 AIC	
		最小二乗法 OLS	空間誤差 モデル SEM	モランの I 統計量	モランの I 統計量	空間パラメータ λ の LR 検定	OLS	SEM
A P E	5	2000年 (3,244)	●	0.074 **	0.045 **	0.099 **	15,148	15,136
		2005年 (1,804)	●	0.074 **	0.046 **	0.133 **	7,712	7,703
		2010年 (1,682)	●	0.108 **	0.092 **	0.214 **	7,008	6,976
	10	2000年 (3,244)	●	0.111 **	0.089 **	0.197 **	18,933	18,877
		2005年 (1,795)	●	0.103 **	0.095 **	0.232 **	9,866	9,827
		2000年 (3,232)	●	0.130 **	0.119 **	0.280 **	23,022	22,912
A L P E	5	2000年 (3,244)	●	0.093 **	0.081 **	0.172 **	17,067	17,023
		2005年 (1,804)	●	0.132 **	0.070 **	0.188 **	8,848	8,826
		2010年 (1,682)	●	0.133 **	0.101 **	0.241 **	8,124	8,084
	10	2000年 (3,244)	●	0.205 **	0.162 **	0.330 **	20,674	20,495
		2005年 (1,795)	●	0.204 **	0.122 **	0.292 **	10,874	10,809
		2000年 (3,232)	●	0.214 **	0.176 **	0.379 **	24,190	23,963

有意水準: 0.01 **, 0.05 *, 0.1 +

注: 対象地域については、表5を参照。

(3)空間誤差モデルの数値は空間パラメータ λ の値を示している。

2. 推計誤差の空間的異質性の検証（市区町村）

分析課題2は（表7）、推計誤差と独立変数との関係が地域によって異なるという空間的異質性（空間的非定常性）の検証が目的である。表7には、APE及びALPE別に、推計期間、基準年と地域数、最適なモデル選択についての結果（最小二乗法OLSか地理加重回帰モデルGWRか）、その根拠となる統計量(1)～(4)を示している。F検定は、(1)GWRモデルがOLSよりもモデルフィットがよいか、(2)OLSモデルからGWRモデルへの改善度がないかどうか、(3)共変量の偏回帰係数が地域によって異なるかどうかについての検定結果であり、(4)はOLSモデルとGWRモデルの赤池情報量基準によるモデルの適合度を示した。

最適なモデル選択では、F(3)検定結果を基に判断しており、空間的非定常性が認められない場合はOLSモデルが最適なモデルとなる。

GWRモデルによる推定を行った結果、APEモデルでは、2010年基準推計（推計期間5年）、2000年基準推計（推計期間10年、15年）においては、F(1)検定においてGWRモデルの方がOLSモデルよりもモデルが改善するという結果が得られているが、F(3)検定では地域によって偏回帰係数が異なるとする空間的異質性（空間的非定常性）は棄却されており、偏回帰係数の地域差は認められないという結果であった。

したがって、推計誤差の規定要因について回帰モデルを用いて推定する場合、地理加重回帰モデルのようなローカル推定ではなく、分析課題1にあるような空間誤差モデル等の空間隣接行列による調整を行った上でのグローバル推定が適切であるといえる。

表7 市区町村別APE、ALPEに対する地理加重回帰モデルによるモデル選択結果

推計期間 (年)	基準年 (地域数)	最適なモデル		F(1)検定 Leung et al. (2000)	F(2)検定 Leung et al. (2000)	F(3)検定 Leung et al. (2000)	(4)赤池情報量基準 AIC		
		OLS	GWR	GWR>OLS	GWR≠OLS	空間的 非定常性	OLS	GWR	
APE	5	2000年 (3,244)	●		1.000	2.869 *	No	15,148	15,140
		2005年 (1,804)	●		0.980	6.193 **	No	7,712	7,342
		2010年 (1,682)	●		0.916 *	4.344 **	No	7,008	6,842
	10	2000年 (3,244)	●		0.949 +	2.990 **	No	18,933	18,734
		2005年 (1,795)	●		0.976	4.619 **	No	9,866	9,811
	15	2000年 (3,231)	●		0.940 *	3.304 **	No	22,856	22,784
ALPE	5	2000年 (3,244)	●		1.000	2.587 *	No	17,067	17,059
		2005年 (1,804)	●		0.997	6.542 **	No	8,848	8,834
		2010年 (1,682)	●		0.970	5.496 **	No	8,124	8,061
	10	2000年 (3,244)	●		1.000	2.303 *	No	20,674	20,666
		2005年 (1,795)	●		0.991	4.722 **	No	10,874	10,847
	15	2000年 (3,231)	●		0.896 **	4.281 **	No	23,879	23,793

有意水準: 0.01 **, 0.05 *, 0.1 +

注: GWRモデルの推定は固定型カーネル加重による。

対象地域については、表5を参照。推計年2000年推計期間15年では、小笠原村は推定されたカーネル加重関数をもとにしたローカルパラメータが推定できないため、分析から除外した。

V. 結論と展望

本稿は地域推計の推計誤差に空間的自己相関及び空間的異質性の検証を行うことで、回帰モデルを用いた推計誤差の評価研究に手法的改善を提案することが目的であり、これまでの地域推計の誤差率の傾向について整理するとともに、推計誤差の空間的特性に着目した分析を行った。とくに推計誤差の空間的特性の評価を行うために、空間誤差モデル、地理的加重回帰モデルなど空間的特性に着目した空間計量経済学モデルを用いて検証を行った。

推計誤差の空間的特性に関する2つの分析課題を設定し統計的検定による分析を行った結果、(1)推計誤差の空間的自己相関は都道府県では特定の基準年と推計期間によっては部分的に検出されるに留まったが、市区町村では全ての基準年・推計期間において空間的自己相関が検出された。市区町村の推計誤差について回帰モデルで精度検証を行う場合、空間誤差モデルを用いるなど空間的隣接行列による推定値の調整が必要になることが明らかとなった。都道府県別の誤差分布をみると、東京都や奈良県など特異な誤差分布を持った地域があることで空間的自己相関の検出が推計基準年や推計期間によって異なり、とりわけ推計期間が長くなるにつれて東京都の過小傾向が拡大するなど特異な結果がみられることが影響しているものと考えられる。1990年推計基準でみたAPEの中央値は前後の推計に比べて大きく出やすいが、推計期間が長くなるにつれて誤差率が前後に比べて低くなるなど推計に用いた仮定の影響が空間的自己相関の検出方法にも影響していることが示唆される。

(2)市区町村の推計誤差と共変量との関係は地域によって異なるとする偏回帰係数の空間的異質性については、一部のモデルでモデルフィットが改善する場合があるが、偏回帰係数の地域分布は検出されなかったことから、ローカル推定を行う必要はないことがわかった。

社人研の地域推計の推計誤差の特徴は、基本的属性からみると人口規模が小さくなるほど誤差率が高くなる傾向であること、東京圏を中心とする大都市圏では誤差は小さいものの過小推計になる傾向が観察され、奈良県などの特定の県の中山間地域など小規模な地域では過大推計になることが明らかとなった。また推計誤差の空間的特性は、都道府県では推計期間が短い場合に検出され、市区町村別の誤差分布では明確な空間的自己相関が存在することが明らかとなった。

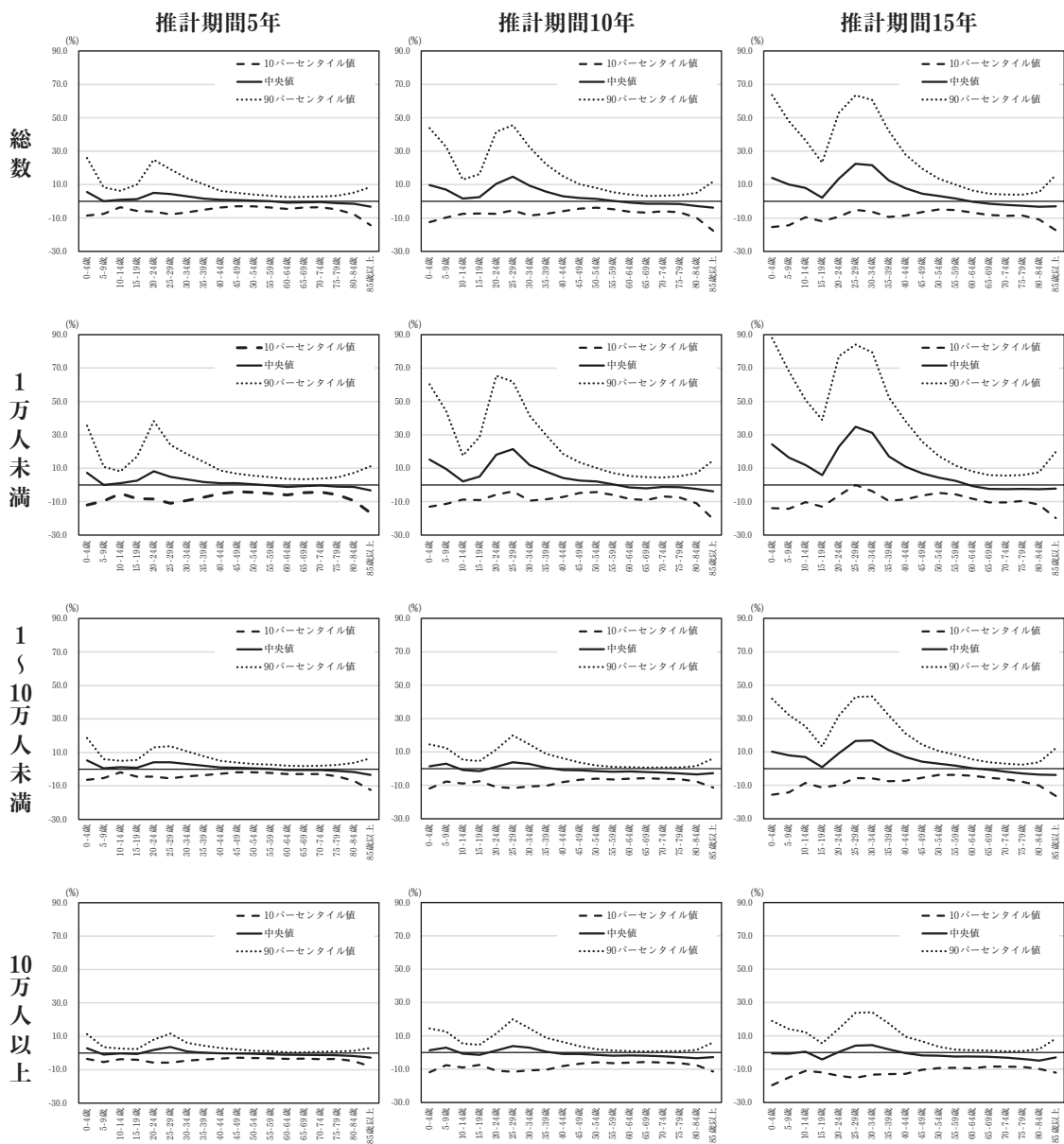
地域別将来推計人口の将来の仮定値について、出生率（子ども女性比）と死亡率（生残率）は全国値との相対的較差から全国将来推計人口の仮定値を用いて将来の仮定値を作成していることから、基準年時点の空間的特性を将来まで保持するような仮定値設定を行っている。したがって全国将来推計人口の仮定値の誤差部分と将来の仮定値に関する地域分布の変化分によって将来の誤差が生じるため、推計誤差の空間的特性にも影響を与える可能性がある。

社人研が平成30（2018）年に公表した「平成30年推計」では、移動率の仮定値設定が多地域モデル（プールモデル）に変更された。このような手法的な変更に伴い、これまで純移動率モデルにおいて設定されていた将来的に純移動率が縮小される仮定から、将来にわたり移動傾向が原則として一定となるような仮定に変更されている（国立社会保障・人口問題研究所 2018）。さらに、プールモデルにおいて転入数の推計に用いている配分率の仮定には、「他地域の人口分布の変化を考慮した係数」と「推計対象地域の人口が全国人口に占めるシェアの変化を考慮した係数」を導入しており、これまでみられた東京圏の過小推計や奈良県の過大推計傾向が縮小される可能性がある。誤差率の年齢別分布は移動率の年齢パターンに似た形状を示すことから、移動率の仮定値設定の考え方は誤差率の変動と大きく関係している。今回の推計結果が2020年以降の総人口の変動とどのように関わってくるのかについて継続して分析を行っていくことで、次回以降の地域推計の仮定値設定の在り方に対する知見を蓄積することが可能となるであろう。

（2020年1月12日査読終了）

参考図 基準年別，推計期間別，人口規模（3区分）別，5歳階級別，市区町村のALPEの10パーセンタイル値・中央値・90パーセンタイル値

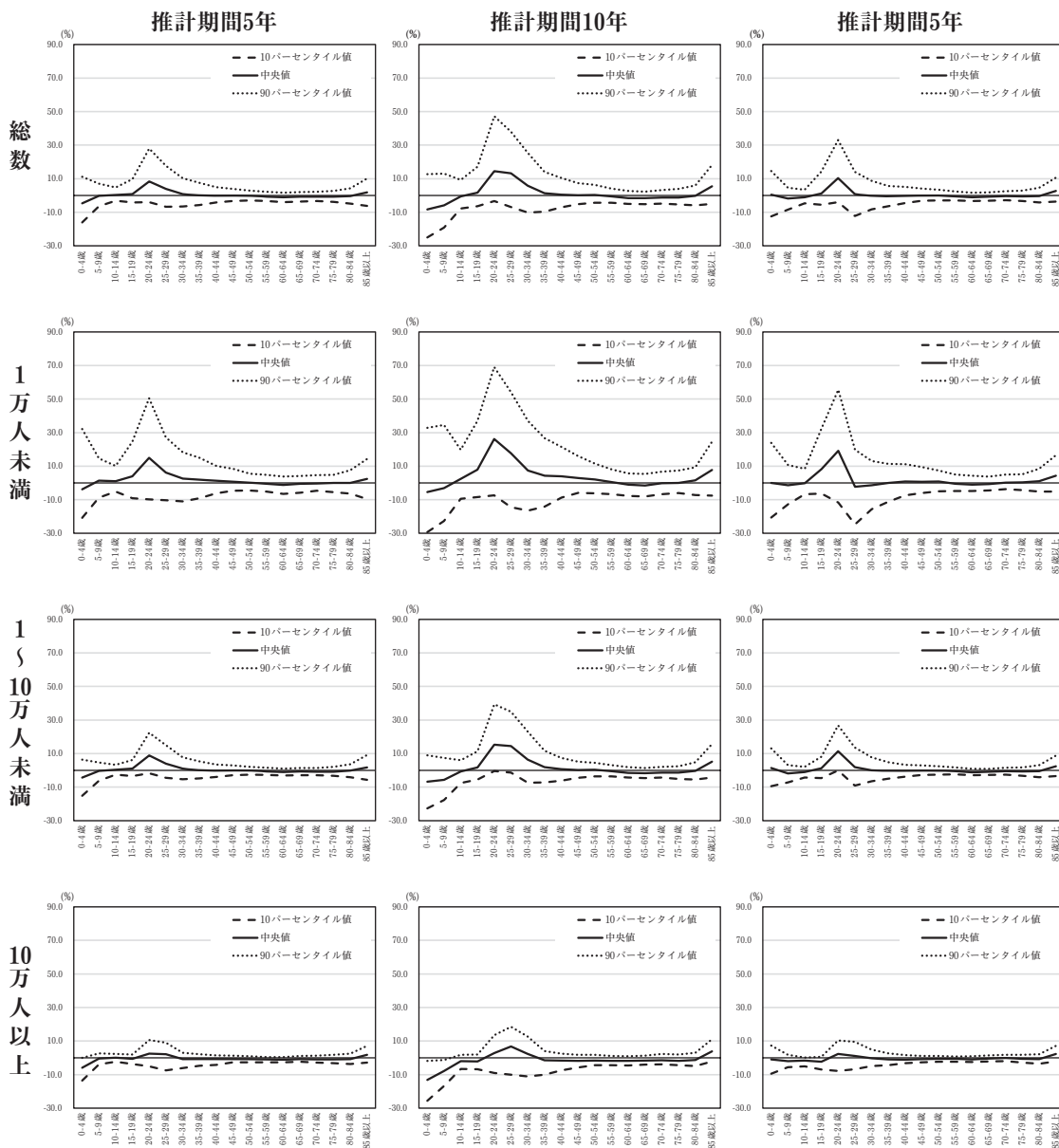
基準年2000年



参考図 基準年別，推計期間別，人口規模（3区分）別，5歳階級別，市区町村のALPEの10パーセンタイル値・中央値・90パーセンタイル値 つづき

基準年2005年

基準年2010年



謝辞

※本研究（の一部）は、厚生労働行政推進調査事業費補助金（政策科学総合研究事業（政策科学推進研究事業））「国際的・地域的視野から見た少子化・高齢化の新潮流に対応した人口分析・将来推計とその応用に関する研究（研究代表者石井太，課題番号（H29-政策-指定-003）」による助成を受けた。

参考文献

- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, Kluwer, Dordrecht.
- Anselin, L. (2005), *Exploring Spatial Data with GeoDaTM: A Workbook*, Center for Spatially Integrated Social Science.
- Anselin, L. and Bera, A. K. (1998) "Spatial dependence in linear regression models with an introduction to spatial econometrics", *Handbook of Applied Economic Statistics* (eds. Ullah, A. and Giles, D. E.), pp. 237-289, Marcel Dekker, New York.
- Brunsdon, C., Fotheringham, A. S. and Charlton, M., (1996), "Geographically weighted regression: a method for exploring spatial nonstationarity", *Geographical Analysis*, No.28, pp.281-298.
- Brunsdon, C., Fotheringham, A. S. and Charlton, M., (1999), "Some notes on parametric significance tests for geographically weighted regression", *Journal of Regional Science*, 39, pp.497-524.
- Chi, G. (2009), "Can Knowledge Improve Population Forecasts at Subcounty Levels?", *Demography*, 46: 405-427.
- Cliff, A. D. and Ord, J. K., (1981) *Spatial Processes: Methods and Applications*, Pion, London.
- Fotheringham, A. S., Brunsdon, C. and Charlton, M., (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley & Sons.
- Lenze, D. G. (2000), "Forecast accuracy and efficiency: an evaluation of ex ante substate long-term forecasts", *International Regional Science Review*, 23, pp.201-226.
- Leung, Y., Mei, C.-L. and Zhang, W.-X., (2000), "Statistical tests for spatial nonstationarity based on the geographically weighted regression model", *Environment and Planning A*, 32, pp.9-32.
- Moran, P.A.P., 1950, "Notes on Continuous Stochastic Phenomena," *Biometrika*, 37, pp.17-33.
- Murdoc, S. H., Leistritz, F., Hamm, R., Hwang, S., Parpia, B. (1984), "An Assessment of the Accuracy of a Regional Economic-demographic Projection Model", *Demography*, 21: 383-404.
- Statistics New Zealand (2016). How accurate are population estimates and projections? An evaluation of Statistics New Zealand population estimates and projections, 1996-2013. Retrieved from www.stats.govt.nz.
- Rayer, S. (2008), "Population forecast errors: A primer for planners", *Journal of Planning Education and Research*, 27, pp.417-430.
- Smith, S. (1987), "Tests of Forecast Accuracy and Bias for County Population Projections", *Journal of the American Statistical Association*, 82(400): 991-1003.
- Smith, S. and Shahindullah, M. (1995), "An Evaluation of Population Projection Errors for Census Tracts", *Journal of the American Statistical Association*, 90(429): 64-71.
- Smith, S. and Sincich, T. (1988), "Stability Over the Distribution of Population Forecast Error", *Demography*, 25:461-474.
- Smith, S. and Sincich, T. (1991), "An Empirical Analysis of the Effect of Length of Forecast Horizon on Population Forecast Errors", *Demography*, 28: 261-274.
- Smith, S. and Tayman, J. (2003), "An Evaluation of Population Projections by Age", *Demography*, 40: 741-757.
- Smith, S., Tayman, J., Swanson, D. (2013), *A practitioner's guide to state and local population projections*.

- Dordrecht, The Netherlands: Springer.
- Tayman, J. and Swanson, D. (1996), "On the Utility of Population Forecasts" *Demography*, 33: 523-528.
- Tayman, J., Smith, S., Rayer, S., (2011), "Evaluating Population Forecast Accuracy: A regression approach using county data". *Population Research and Policy Review*, 30: pp. 235-262.
- Tobler, W. R., (1970), "A Computer Movie Simulating Urban Growth in the Detroit Region", *Economic Geography*, 46, pp.234-240.
- Yamauchi M., Koike S. and Kamata K., (2016), "How Accurate Are Japan's Official Subnational Projections? Comparative Analysis of Projections in Japan, English-Speaking Countries and the EU", Swanson D. (ed.) *The Frontiers of Applied Demography*, pp.305-328.
- Wilson, T. (2012) "Forecast accuracy and uncertainty of Australian Bureau of Statistics State and Territory Population Projections", *International Journal of Population Research*, Vol. 2012, pp.1-16.
- 小池司朗 (2015) 「多地域モデルによる都道府県別将来人口推計の結果と考察」『人口問題研究』第71巻第4号, pp.351-371.
- 小池司朗 (2016) 「プールモデルの投影精度に関する研究」『人口問題研究』第72巻第3号, pp. 256-275.
- 厚生省人口問題研究所 (1987) 『都道府県別将来推計人口—昭和60年～100年間毎5年—(昭和62年1月推計)』研究資料第247号.
- 厚生省人口問題研究所 (1992) 『都道府県別将来推計人口—平成2年～22年間毎5年—(平成4年10月推計)』研究資料275号.
- 国立社会保障・人口問題研究所 (1997) 『都道府県別将来推計人口—平成7(1995)～37(2025)年—(平成9年5月推計)』研究資料第293号.
- 国立社会保障・人口問題研究所 (2002) 『都道府県別将来推計人口—平成12(2000)～42(2030)年—(平成14年3月推計)』人口問題研究資料第306号.
- 国立社会保障・人口問題研究所 (2003) 『日本の市区町村別将来推計人口—平成12(2000)～42(2030)年—(平成15年12月推計)』人口問題研究資料第310号.
- 国立社会保障・人口問題研究所 (2007) 『日本の都道府県別将来推計人口—平成17(2005)～47(2035)年—(平成19年5月推計)』人口問題研究資料第316号.
- 国立社会保障・人口問題研究所 (2008) 『日本の市区町村別将来推計人口—平成17(2005)～47(2035)年—(平成20年12月推計)』人口問題研究資料第321号.
- 国立社会保障・人口問題研究所 (2013) 『日本の地域別将来推計人口—平成22(2010)～52(2040)年—(平成25年3月推計)』人口問題研究資料第330号.
- 国立社会保障・人口問題研究所 (2018) 『日本の地域別将来推計人口—平成27(2015)～57(2045)年—(平成30年推計)』人口問題研究資料第340号.
- 瀬谷 創・堤 盛人 (2014) 『空間統計学 自然科学から人文・社会科学まで』, 朝倉書店.
- 山内昌和・小池司朗 (2014) 「地域人口推計」, 『人口問題研究』, 70-4, pp. 351-424.
- 山内昌和・小池司朗 (2015) 「英語圏諸国との比較からみた社人研の地域別将来推計人口の誤差」, 『人口問題研究』, 71-3, pp. 216-240.

An Evaluation of the Accuracy of Regional Population Projections: Investigation on the Spatial Characteristics in the Projection Error Rates using Regression Models

KAMATA Kenji, KOIKE Shiro, SUGA Keita and YAMAUCHI Masakazu

In this paper, we evaluate the accuracy of "The Regional Population Projections" conducted by the National Institute of Population and Social Security Research. The projections include two types of error rates (APE, ALPE) that are used to analyze spatial and time series characteristics of error rates using a regression method. In this paper, we analyze the projections conducted on base-populations from 1985 to 2010 for prefectures, and from 2000 to 2010 for municipalities.

We found that error rates have had the tendency to decrease in recent projections, due to declining birthrates and population aging, and that the error rates tend to increase for smaller base-populations. In metropolitan areas, the error rates tend to be under-projected, whereas they become over projected for small local governments in non-metropolitan prefectures such as the Nara Prefecture.

We found salient features in spatial characteristics of error rates in two aspects: (1) spatial autocorrelation of the error rates were detected in prefectural models when the projected period was short, which they were for municipal models in general. Therefore, it is necessary to use spatial econometric models, such as the spatial error component model, when evaluating the error rates at municipal levels in regression based evaluations. (2) We did not detect spatial heterogeneity in the regression coefficients. On account of this, there was no need to employ local estimation techniques in the evaluation of population projection errors for Japanese prefectures and municipalities.

【Key Words】 Regional Population Projections, Projection errors, Spatial Autocorrelation, Spatial Heterogeneity