

コーホート法に基づく小地域における人口推計

瀬谷 創¹・山形 与志樹²・堤 盛人³

¹学生会員 国立環境研究所 地球環境研究センター (〒305-8506 つくば市小野川16-2)
E-mail: seya.hajime@nies.go.jp

²非会員 国立環境研究所 地球環境研究センター (〒305-8506 つくば市小野川16-2)
E-mail: yamagata@nies.go.jp

³正会員 筑波大学准教授 システム情報系 (〒305-8573 つくば市天王台1-1-1)
E-mail: tsutsumi@sk.tsukuba.ac.jp

小地域における人口推計はデータの秘匿や人口の少なさに起因する推計値の精度の低さ等の問題に直面するため容易ではなく、方法論が確立されているとは言い難い状況にある。現在我が国が直面する人口減少・超高齢社会における各種の課題に計画的に対処するためにも、BAUシナリオにおける将来人口を精緻に予測する技術を確認することが急務であると言える。本研究では、最も標準的な人口推計法であるコーホート法（要因・変化率法）に着目し、まず、小地域人口の推計に関する既往研究を空間統計学の視点から体系的に整理することを試みる。次に、小地域推定の方法が一つの有用な方法になり得ることを指摘する。最後に、これらをつくば市の人口推計に適用し、手法の相対的な優劣に関する実証的な考察を行う。

Key Words: *The ratio change method, small area population projection, small area estimation, spatial statistics*

1. はじめに

我が国は人口減少・超高齢社会を迎え、今後高度成長期に集中的に整備された社会資本が一斉に更新期を迎える等の背景もあり、戦略的な「撤退の計画」の重要性が指摘されはじめている（例えば、谷口ら, 2007）。撤退の計画においては、安心・安全、利便性、低炭素化といった様々なQOL指標を勘案しながら、撤退地域を選定していく必要がある。そのためには、小地域（町丁目、3・4次メッシュ）における年齢階層別の将来人口を精緻に求める必要があることは言うまでもない。

また、温暖化がこれらのQOL値を変化させる可能性があることを認識することは重要である。例えば、東京都市圏では海拔0メートル地帯に住宅や事業所が集積し、洪水・高潮等・津波等の水害リスクに対して極めて脆弱であるが、さらにいくつかの研究において、温暖化が水害リスクを高める可能性があることが指摘されている（例えば、東ら, 2006）。気候変動に対しては、緩和・適応策を同時に進めていくことの重要であり（三村, 2006）、自治体レベルにおいても適応策策定の試みが始まりつつある（白井・田中, 2011）。無論、自治体が適応策の検討を行う上でも、暴露人口としての小地域単位における将来人口を精緻に推計する必要がある。

しかしながら、小地域における人口推計は、後述するデータの秘匿や人口の少なさに起因する推計値の精度の低さ等の問題に直面するため容易ではなく、方法論が確認されているとは言い難い状況にある。そこで本研究では、最も標準的な人口推計法であるコーホート法（要因法・変化率法）に着目し、次のような方法で、小地域における人口推計研究に貢献することを試みる。

まず、コーホート法を用いた小地域における人口推計に関する既往研究の体系的な整理を試みる。小地域における人口推計に関する研究はそれ自体非常に少ないが、特に都市計画や地理情報科学の分野においていくつかの興味深い研究が蓄積されている。それらは、空間統計学（spatial statistics）の一部とも位置付けられる空間疫学（spatial epidemiology）の観点から眺めると、分かりやすく整理することが可能である。続いて、これらの手法をつくば市の人口推計に適用し、手法の相対的な優劣に関しての実証比較分析を試みる。このような試みは筆者のレビューした範囲では行われておらず、実証研究として有益であると考えられる。

2. 人口推計の方法

人口推計は、人口統計学の分野を中心に発展し、邦書

では山口 (1990), 岡崎 (1999), 和田 (2006) 等で基礎的な概念が整理されている. 地域レベルの人口推計については, 濱・山口編 (1997) にまとめられている. ここで, 濱・山口編 (1997) における地域とは, 主として都道府県に対応するレベルであり, 市区町村以下の領域は, 「小地域」と呼ばれている. 地域という概念の曖昧さ故, 分析においては明確な定義を与える必要がある. 本稿では, 市区町村を地域, 市区町村内の町丁目, ないしは3次 (約1km²)・4次 (約0.5km²) メッシュを小地域と定義し, 後者 (小地域) における人口推計を対象とする. 地域 (市区町村) 程度の空間解像度における人口推計については, Rees et al. (2004) がイギリスの事例を元にレビューを行っており, 我が国では, 石川 (1993) の『市町村人口推計マニュアル』に具体的な推計方法が紹介されている. 石川 (1993) は, 移動マトリクス法と呼ばれる, 性・年齢別に地域間の移動マトリクスを求め, それを適用する方法の有用性を指摘している. また, 小池 (2008a, b) は, 都道府県の人口推計にこの方法を利用している. しかしながら, 市区町村以下ではデータの制約から移動マトリクス法の適用は困難である.

現在, 将来人口推計のために最も用いられている方法は, 移動マトリクス法の簡便版といえるコーホート法である. コーホート法では, 人口移動の仮定を, 純移動率, あるいはコーホート変化率で捉えるため, 必要とするデータが非常に少ない. コーホート法は, 大きくコーホート要因法とコーホート変化率法に分類できる. コーホート変化率法は2時点におけるコーホート変化率を用いるのに対し, コーホート要因法はコーホートの変化分を死亡数, 人口移動数, 出生数に分離して予測を行う (石川, 1993). このように自然増加と社会増加を明示的に分離できるという点に, コーホート要因法の利点がある.

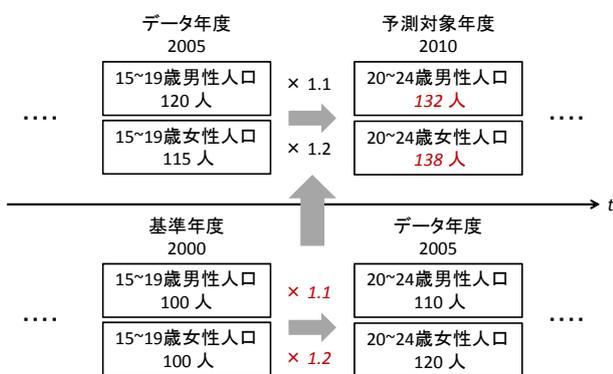


図1: コーホート変化率法概念図
(赤斜線は推計値)

コーホート変化率法は, 星田ら (2011) の表現 (記号) を借りれば, 次のように定式化できる.

$$\bullet 5\sim 84\text{歳人口} : {}^{t+5}P_i^{g,x+5\sim x+9}$$

$$= {}^tP_i^{g,x\sim x+4} \cdot \frac{{}^tP_i^{g,x+5\sim x+9}}{{}^{t-5}P_i^{g,x\sim x+4}} = {}^tP_i^{g,x\sim x+4} \cdot R_i^{g,x\sim x+4}, \quad (1)$$

$$\bullet 85\text{歳以上人口} : {}^{t+5}P_i^{g,85\sim}$$

$$= {}^tP_i^{g,80\sim} \cdot \frac{{}^tP_i^{g,85\sim}}{{}^{t-5}P_i^{g,80\sim}} = {}^tP_i^{g,80\sim} \cdot R_i^{g,80\sim}, \quad (2)$$

$$\bullet 0\sim 4\text{歳人口} : {}^{t+5}P_i^{f+m,0\sim 4}$$

$$= {}^{t+5}P_i^{f,25\sim 34} \cdot \frac{{}^tP_i^{f,0\sim 4}}{{}^tP_i^{f,25\sim 34}} = {}^{t+5}P_i^{f,25\sim 34} \cdot r_i^{f,25\sim 34}. \quad (3)$$

ここで, ${}^tP_i^{g,x\sim x+4}$ は, t 年における (小) 地域 i の $x\sim x+4$ 歳の性別 $g=\{f,m\}$ の人口を示す. R はコーホート変化率, r は婦人・子供比率と呼ばれる原単位である. 後者に関して, 0~4歳の人口は, その母親となり得る15~49歳の人口との比率 (婦人・子供比) で求めることが多い. ただし実際には, 25~34歳に出生率が集中しているため, 推計のバイアスを避けるために本研究ではこちらを採用することとする. 図1にコーホート変化率による人口推計の概念を簡単に示している.

上式では, コーホート変化率を時間不変なものとして与えているが, 時系列的な変化情報があれば, コーホート変化率の将来値を外挿によって求めることも可能である (石川, 1993). 実際, 奥村 (2005) は時系列モデルによる推計を試みている. このようにコーホート変化率では, 自然増加と社会増加を区別せず, コーホート変化率の情報のみで将来人口の推計を行う手法である. しかしながら, 小地域においてはこれらを精度よく分離することは容易ではないため, コーホート変化率法が用いられることも多い. コーホート要因法については, 石川 (1993) 等を参照されたい.

3. コーホート法を用いた小地域における人口推計

(1) コーホート法を用いた小地域における人口推計における課題の整理

古藤 (2008) は, 小地域の人口推計の問題を, ①値の小ささ, ②短期間のデータ, ③人口移動の不明, の3つに分類している. ①は, 人口が少ないために, コーホート変化率法におけるコーホート変化率や, コーホート要因法における社会増減が安定しないという問題である. 変化率の例でいえば, R の分母が小さいため, 変化率が5を大きく超えるような例である. 例えば, 人口が非常に

表1：我が国におけるコーホート法を用いた小地域の人口推計に関する既往研究

論文	方法	対象地域	空間解像度	既往研究の小地域への対応
奥村(2005)	要因法	広島都市圏	4次メッシュ	・全国平均生存率(生命表)を使用。 ・社会増減を因子分析で集約化。
沢田・羽根(2005)	要因法	中央区, 目黒区, 足立区	町丁目	・市区町村別の生存率(生命表)を使用。
小池(2010)	要因法	首都圏の一部	3次メッシュ	・都道府県別の生存率(生命表)を使用。
有賀ら(2011)	要因法	全国	3次メッシュ	・市区町村別の生存率(生命表)を使用。 ・メッシュ人口>3000人となるようグルーピングし, 配分。
(財)統計情報研究開発センター(2003)	変化率法	首都圏の一部	3次メッシュ	・自メッシュから7km以内のメッシュの平均変化率を使用。 ・人口が0・秘匿が60%以上ある場合は計算対象から除外。
土屋・室町(2005)	変化率法	全国	3次メッシュ	・人口が300人未満の場合, 年齢帯で分けずに推計。 ・自メッシュを中心とした7x7の49区画の平均変化率を使用。
江崎(2006)	変化率法	首都圏の一部	3次メッシュ	・1000人未満のメッシュは計算対象から除外。
古藤(2008)	変化率法	山形市の一部	町丁目	・コーホート変化率を因子分析で集約化。
星田ら(2011)	変化率法	全国の20%	4次メッシュ	・距離の逆数ベースの重み関数でスムージング。

少ない小地域において大規模な住宅地開発が行われた場合, このような値が発生し得る。しかし, このような特異な変化率が次の5年間も持続するとは考えにくい。①についてはまた, 分母が0のとき, 変化率が定義できないという点も実用上大きな課題である。

②は, 長期間の連続的なデータを手に入れることが難しいという問題である(合併による区域の変更等)。これへの対応として既往研究では3次・4次メッシュが分析に用いられるケースが多い。しかしながらこの場合, メッシュへの按分誤差の推定値への混入が避けられない。

③は, 区域が細かいため, 人口移動率のデータを手に入れることが不可能であるという問題である。したがって, 上位ゾーンの値を借用するか, 代わりにコーホート変化率法が用いられる。

その他にも, 秘匿データの問題がある。近年のGISの発展により, 小地域における人口分析の可能性は飛躍的に増大した(武者ら, 2002)。e-Stat(統計GIS)では, 国勢調査小地域, 3次メッシュでの人口が, 無償で公開されている。しかしながらこれらのデータでは, プライバシー保護の観点から, ある小地域の世帯数が6以下の場合, その小地域は秘匿対象とされ, 総人口・世帯数以外の情報は秘匿されている。性・年齢別の値は近い位置の合算地域メッシュに合算されて示される(古藤・三浦, 2012)。したがって秘匿データを復元する際には, これを単純に欠損値と扱って復元するのではなく, 復元値の合計が合算地域メッシュの値と一致するような制約を考慮しなければならず, 面補間(areal interpolation)の問題となる(例えば, Yoo and Kyriakidis, 2006)。

本稿は, 上記のうち, ①の「値の小ささ」の問題に取り組むものである。表1に, 既往研究における①の問題への対処法を整理した。これらは, 大きく説得力の借用(borrowing strength)型と, 情報抽出型に分類することができよう。前者は, 隣接エリア, あるいは上位ゾーンの情報を使って, コーホート変化率や社会増減数を安定化させようとするアプローチであり, 一方, 後者は因子分析で有用な情報を取り出す, あるいは異常値を省くと言

った作業でこれらを安定化させようとするアプローチである。

(2) 空間統計学・空間疫学の視点からの整理

(1) で述べた観測データの少なさに起因する観測値(比率)の不安定さの問題は, 空間疫学の分野において, small number problemとして知られ(Haining et al., 2010), 特に説得力の借用型の対処法に対する知見が蓄積されている。Anselin et al. (2006)は, それらを大きく, 1) 平均・中央値ベースのスムージング(mean and median based smoothing), 2) ノンパラメトリック・スムージング(nonparametric smoothing), 3) 領域化(regionalization), 4) ベイズスムージング(Bayesian smoothing)の4つに分類している。このうち, ベイズスムージングについては, 後述する通り空間疫学のモデルがほとんどポアソン分布に依拠している一方で, コーホート変化率はカウントデータとは見做せないで, そのままの適用は難しい。代わりに本稿では, 小地域推定の方法が有用な代替案と成り得ることを示す。以下のまとめは, Anselin et al. (2006)を参考にしているが, 彼らの論文は空間疫学への適用を前提としているため, ここでは人口推計に関連するアプローチのみを取り出し, 再構成している。また以下, 性・年齢別を示す添え字は省略し, コーホート変化率を簡単に R_i と記すこととする。

1) 平均・中央値ベースのスムージング

R_i を用いて, 小地域*i*における平均ベースのスムージング推定値は, 他の小地域における観測値の重みつき平均として

$$\tilde{R}_i = \sum_j \tilde{w}_{ij} R_j, \quad (4)$$

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}},$$

と与えることができる。この方法は, 重みの与え方で様々に構成することが可能である。例えば, 小地域*i, j*

間の距離を d_{ij} 、距離の閾値を δ としたとき、 $w_{ij}=1, \text{ for } d_{ij} < \delta$ とする、あるいは w_{ij} を直接 d_{ij} の関数として与えるといった方法が考えられる。空間計量経済学（例えば、Anselin, 1988）では慣習的に $w_{ij}=0$ とおくことが多いが、ここでは自身の値も重要になるため、 $w_{ij}=1$ とすべきである点に注意が必要である。ただしこの場合、空間計量経済学でしばしば用いられる $(1/d_{ij})^\alpha$ (α はパラメータ) という距離の逆数ベースの重みを用いた場合、距離が 0 の自地域で重みが定義できなくなるという問題が発生する。本稿では簡便な解決法として、栗田・腰塚 (1988) による対面の領域間平均移動距離（一次近似）を用いる。これは各小地域が真の面積に等しい円であり、かつ人口移動者の発着地が小地域に様に分布していると仮定したときの距離である。このアプローチによれば、小地域 i, j の面積を a_i, a_j としたとき、内々距離・内外距離をそれぞれ、

$$\text{内々距離} : \frac{128\sqrt{z_i}}{45\pi^{1.5}}, \text{ 内外距離} : d_{ij} + \frac{z_i + z_j}{8\pi d_{ij}}, \quad (5)$$

と与えることができる。

さて、上記の手法を用いれば、5 以上のような極端なコーホート変化率をスムージングし、安定化させることができるが、逆にこのような特異値が周囲の安定的な値にまで影響を与えてしまうという問題が起こる。したがって、Anselin et al. (2006) は、小地域人口で重みをつける方法も提示している。

$$\tilde{R}_i = \sum_j \tilde{w}_{ij} R_j, \quad (6)$$

$$\tilde{w}_{ij} = \frac{\hat{w}_{ij}}{\sum_j \hat{w}_{ij}}, \hat{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}} \cdot \frac{P_j}{\sum_j P_j},$$

ここで、 P_i は性・年齢別人口であるが、 R_i に合わせて添え字は省略している。この方法を用いれば、異常値が周囲の値に与える影響を緩和することができると考えられる。似たような方法に、地球統計学の分野のポアソングリギング (poisson kriging) がある (例えば、Goovaerts, 2005)。

Anselin et al. (2006) は、平均値だけでなく、中央値ベースのスムージング方法も提案している。これは例えば、小地域 i から閾値 δ の正円を想定し、円内の小地域の中央値を推定値として採用する方法である。この方法は平均値ベースと比較して外れ値に頑健であるものの、推定値に自地域の観測値が反映されないという点で解釈に難があるといえる。

2) ノンパラメトリック・スムージング

ノンパラメトリック・スムージングは、いくつかの手法に分類されるが、このうち特に人口推計に有用な空間

比率 smoother (spatial rate smoother) は、比率 (コーホート変化率) を構成する分母と分子それぞれを重み付き平均するアプローチである。あらかじめ比率を計算してしまう 1) のアプローチと比べて、より特異値が発生しにくい方法であると考えられる。星田ら (2011) は、本手法において距離の関数として、bisquare weights function を用い、重みを構成している。(財) 統計情報研究開発センター (2003)、土屋・室町 (2005)、有賀ら (2011) の手法も、近隣集合内の観測値に様な重みを与える方法であり、空間比率 smoother の一種と理解することが可能である。

3) 領域化

領域化は、何らかのゾーンに集計して値を安定させる方法である。必ずしも、空間的な相関関係を利用して集計するわけではない点が 1) や 2) と異なる。Anselin et al. (2006) では、グリッド化して再集計する方法、クラスタリングによりグループ化し、グループ毎に平均を求める方法、値が均一になるように面積を調整するカルトグラムに関連したアプローチ (density equalizing map projections) 等が紹介されている。また、大澤ら (2008) は、小地域における人口推計を明示的な研究対象としているわけではないが、領域化による集計誤差について理論的な検討をおこなっている。

4) ベイズスムージング

空間疫学では、疾病地図を作成することが一つの重要なテーマとなっている。小地域においては、死亡数/人口、という比率が人口が少ない場合不安定となる。これに対して空間疫学では、上位ゾーンにおける値を事前情報としてベイズ的にモデルに取り入れることでスムージングするという方法がとられる (例えば、Waller and Gotway, 2004; 丹後ら, 2007; 高橋ら, 2008)。疾病による死亡数は人口に対して非常に少ないため、死亡数はポアソン分布 (や負の二項分布) に従うと仮定することが妥当である。しかしながら、比率データ R_i はポアソン分布に従うとはみなせないため、本研究では小地域推定の方法を用いることを提案する。

小地域推定の代表的なモデルである Fay and Herriot (1979) のモデル (Fay-Herriot モデル) は、

$$R_i = \theta_i + e_i, e_i \sim N(0, \sigma_e^2 / P_i) \quad (7-a)$$

$$\theta_i = \mathbf{x}_i' \boldsymbol{\beta} + v_i, v_i \sim N(0, \sigma_v^2) \quad (7-b)$$

と与えられる (例えば、久保川, 2007)。ここで、 σ_e^2 は後述する既知の定数である。データ R_i の分散は、人口が少ない小地域で大きく、多い小地域で小さいため、Fay-Herriot モデルは人口データの特徴をよく表したモデルであると考える。

式(7)において、 \mathbf{x}_i は説明変数ベクトル、 $\boldsymbol{\beta}$ は対応するパラメータベクトルであるが、これらによって構成される平均(トレンド)成分は、小地域によらず一定の値をとると仮定する($\mathbf{x}_i'\boldsymbol{\beta} = \phi$)。このとき、 $\hat{\sigma}_{ei}^2 \equiv \hat{\sigma}_e^2 / P(i)$ とおけば、次式が成り立つ。

$$\theta_i = \gamma_i R_i + (1 - \gamma_i)\phi, \quad (8-a)$$

$$\gamma_i = \sigma_v^2 / (\sigma_v^2 + \hat{\sigma}_{ei}^2), \quad (8-b)$$

すなわち、Fay-Herriotモデルによるコーホート変化率のEBLUP推定値である θ_i は、実際のデータと、グローバルな平均成分である ϕ の重みつき平均として与えられる。ここで重みは、グローバルな平均成分の分散 σ_v^2 と、観測データの既知の分散 $\hat{\sigma}_{ei}^2$ の比率として与えられ、分散の逆数が精度を表すと考えれば、データ精度の高い小地域は自身のコーホート変化率に近づき、データ精度の低い小地域はグローバルな平均に近づく。これによって、コーホート変化率の自然な安定化ができると期待される。

Pratesi and Salvati (2008) は、空間計量経済学の手法を援用し、グローバルな誤差 v_i に空間的な自己相関が存在すると仮定したモデル化を行っている。具体的には、誤差項が

$$v_i = \lambda \sum_j w_{ij} v_j + u_j, \quad u_i \sim N(0, \sigma_u^2) \quad (9)$$

に従うとするアプローチである。空間的自己相関の導入により、さらなるスムージング効果が期待されるため、本研究ではこのモデルも比較対象とする。

さて、Fay-Herriotモデルでは、 $\hat{\sigma}_e^2$ は既知の定数とされるが、この項は観測(サンプリング)誤差であり、データから推定することはできない。したがって、何らかの事前情報を元に分析者が設定する必要がある。このパラメータを既知とすれば、他のパラメータは制限付き最尤法(REML) (Pratesi and Salvati, 2008) や、ベイズ推定(You and Zhou, 2011)によって容易に求めることができ

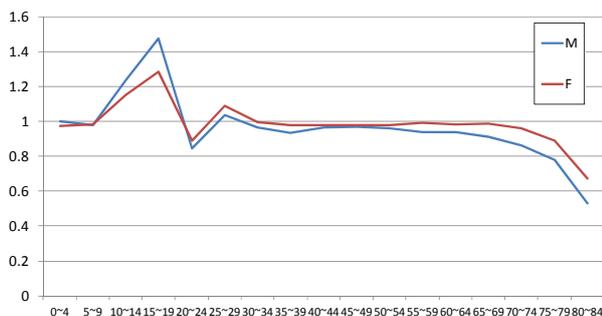


図2：つくば市における性・年齢別コーホート変化率(2000~2005年)

縦軸：コーホート変化率、横軸：コーホート

る。しかしながら、ベイズ推定は通常MCMC法等のシミュレーションに依拠するため、計算時間がかかる。ここでは自治体における人口推計と言う目的を想定して、より計算が簡便なREMLを用いる。なお、REMLを用いた場合、 θ_i はEBLUPとして求めることができるため、以下では式(7)、(9)のモデルをそれぞれEBLUP、Spatial EBLUP(S-EBLUP)と呼称することとする。

4. つくば市の人口推計への適用

(1) 対象地域

本研究では、3章でまとめた手法の実証比較分析を試みる。対象地域はつくば市、モデルの構築単位は町丁目(275)とし、2010年の人口を2000~2005年のコーホート変化率を用いて推計する。なお、人口データのソースは国勢調査である(株式会社日本統計センターのNSCデータベースより)。つくば市は、筑波大学を有する学園都市であるため、一部の小地域で、15~19歳において高い社会増加数(コーホート変化率)を示し、20~24歳において社会減少を示すという特徴がある。図2に、市平均の性・年齢別コーホート変化率を示すが、ここでもこのような傾向が見て取れる。つくば市では、2005年につくばエクスプレス(TX)が開業し、地域構造が大きく変化している。したがってこのような地域へのコーホート変化率法の適用性を見るという意味でも興味深い対象地域といえる。なお、小地域の人口推計では年齢不詳人口の扱いが問題となるが、本対象期間においては0であった。また、秘匿データについては、男女別人口総数の割合で割り振っている。期首人口が0にも関わらず、期末人口が0以外の小地域が275中3地域存在したが、これらは分析に与える影響が少ないため、00-05年の増加数と同じだけ05-10年も増加するとした。

(2) 対象とする手法

本研究の比較対象は、町丁目のコーホート変化率をそのまま適用する方法(micro direct)に加えて、以下とする。なお、説明の右に青下線でモデル名称を示している。

- (a) 平均ベースのスムージング
 - (人口重み無, $w_j=(1/d_{ij})^\alpha$ or $w_j=knn\#$) [dist \$\alpha\$](#) or [knn#](#)
 - (人口重み有, $w_j=(1/d_{ij})^\alpha$ or $w_j=knn\#$) [p_dist \$\alpha\$](#) or [p_knn#](#)
- (b) 空間比率スムージング
 - ($w_j=(1/d_{ij})^\alpha$ or $w_j=km$) [p_dist \$\alpha\$ s](#) or [p_knn# s](#)
- (c) 領域化
 - (市区町村の平均的な値を使用) [city](#)
 - (k-mean 法でクラスタリング) [k-mean](#)
- (d) ベイズスムージング
 - (空間的自己相関考慮無) [eblup](#)
 - (空間的自己相関考慮, $w_j=knn4$) [s-eblup](#)

ここで, $knn\#$ は k 近傍法を表し, 最近隣 # 小地域に重み 1 を与える方法である. 本研究では, $k=4, 6, 8$ の 3 ケースを考慮した. また, 距離の逆数ベースの重みについては, $\alpha=0.5, 1, 2$ の 3 パターンについて計算を行った. $\alpha=0.5$ のときは, 距離に対して重みの減衰が緩やかであるため, 比較的多くの観測点に重みが与えられる一方で, $\alpha=2$ のとき減衰は急であり, 近隣の観測値に大きな重みが割り振られる. 領域化として, 本研究では市区町村の平均的な値を用いるケースと, k-mean 法によって, 2 グループに分割し, それぞれについてコーホト変化率の平均値を算出するケースを想定した. ただし, 後者については, 特異な値があると, (a) 特異な値を持つ小地域と (b) その他という 2 つのクラスに分割されてしまうため, ここでは少数の小地域が孤立しないような基準として, 3 以上の変化率を 3 で置き換えるという方法をとった. その後平均を算出する段階では, オリジナルの値を用いている.

また, 小地域推定においては, 前述の通り各小地域の分散の情報を与える必要がある. ここでは簡単にデータの標本分散で与えることとした. この場合, $\hat{\sigma}_v^2$ と $\hat{\sigma}_e^2$

の値が似通った値をとることとなり, 例えば性・年齢別人口が 1 人であれば $\hat{\sigma}_e^2/1 \approx \hat{\sigma}_v^2$ となるため, 人口が 1 人の町丁目を市平均と同等程度に重視した重みを与える方法となる. ただし, 各小地域の標本分散を計算する際に, 分散が必要以上に大きくなってしまふことを防ぐために, k-mean と同様に値 3 以上に対して置き換えを行った.

(3) 推計結果

図 3 に, 各種法によってスムージングされたコーホト変化率のうち, 特に地域差の激しいと考えられる 15~19 歳の結果を示す. ただし, 男女に大きな違いがみられなかったため, ここでは男性の値のみ示している.

予想される通り, [micro_direct](#) においては, 筑波大学周辺部 (図 5 のつくば駅北部) で高い値を示し, 郊外部では 1 未満を示す町丁目が多い. これに対して, [city](#) では,

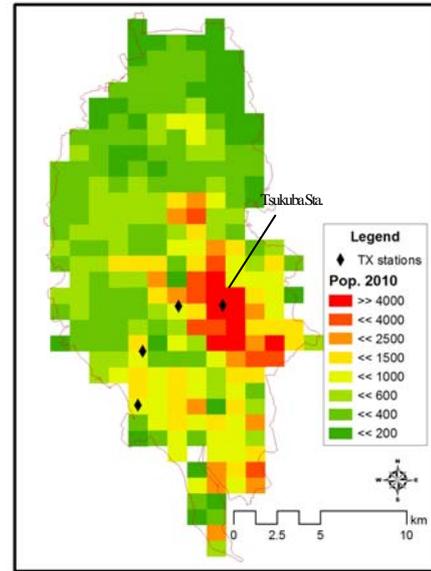


図 5 : つくば市の2010年の人口分布

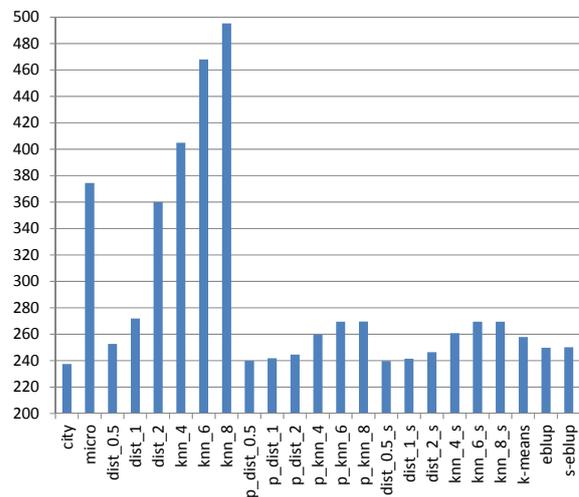


図 6 : 予測におけるRMSE (人)

全小地域が平均の 1.474 であるため, 郊外部で過大推計を起こす可能性があると考えられる.

[dist](#) は, 距離でスムージングする方法であるが, 実際には郊外の変化率の低い小地域を高く設定してしまう等, 必ずしも実態を反映しているとはいえないことが分かる. [knn](#) は, 郊外部で非所に小さな変化率が広がっており, 明らかに奇異である. しかしながら, これは人口で重みづけを行うことで解消できることが分かる ([p_dist](#), [p_knn](#)). 平均ベースのスムージングと空間比率スムージングは, 方法は似通っているものの, 結果のコーホト変化率の分布図は大きく異なっている点は興味深い. k-mean では, コーホト変化率が高い小地域, 低い小地域に適切に分けられていることが分かる. 無論これらの領域は必ずしも空間的に連続するものとはならない. 本研究で人口予測手法として提案した小地域推定

を用いる方法は、`micro_direct` における分布特性を保存しつつも、全体の平均値を取り入れながら適切にスムージングされていることが見て取れる。

図4に、2010年の人口を予測した際の予測誤差（推計値／実測値）の分布を示す。ただし、TXの開業によって町丁目形状が大きく変化しているため、ここではメッシュに面積按分して比較している（2010年の人口分布自体については、図5に示している）。予想される通り、TX駅付近では開業以降の住宅地開発により、人口が過小推計されている。`micro_direct`の結果を見ると、TXつくば駅北側の筑波大学周辺において、人口が過大に推計されていることが分かる。一方で`city`の結果では、コーホート変化率の地域差を無視した結果として、市の北側の人口が少ない小地域において過大推計の傾向が見られる。各手法によって予測誤差の分布傾向は大きく異なることが分かる。小地域推定を用いる方法では、`city`における北部の過大評価の傾向に苦しまず、`micro_direct`の不安定性（つくば駅北部の過大推定）を解消している点は興味深い。

図6に、予測におけるRMSEを計算した結果を示す。様々なスムージング手法が存在するにも関わらず、`city`の結果が最もよいという結果となった。これは小地域における人口予測の実務に示唆を与える結果といえるかもしれない。また、平均ベースのスムージング、特に`knn`においては、人口で重みを付けることが不可欠であることが示唆された。距離ベースの重みは、`knn`に比べて全ケースにおいて予測誤差が小さい。また、 $\alpha=0.5$ で最も精度が良く、薄く広く各観測値に重みを与えることが重要であるという結果が得られた。

5. おわりに

本研究では、最も標準的な人口推計法であるコーホート法（要因・変化率法）に着目し、まず、小地域人口の推計に関する既往研究を空間統計学の視点から体系的に整理することを試みた。

得られた結論は以下のとおりである。

平均ベースのスムージングを用いる場合、距離と人口で重みをつけるべきである。しかしながら、距離ベースの重みは、実際の小地域におけるコーホート変化率の分布特性を必ずしも保持するものではない。小地域推定の手法を用いれば、小地域における分布特性を保持しつつ、精度よく予測が可能である。しかしながら実際には、市レベルの平均を用いれば十分である可能性がある。

今後の課題としては、人口や特性の異なる他の様々な地域への適用によって、結果の安定性について議論することが挙げられる。また、人口推移は住み替えや住宅のライフサイクルと密接に絡むことから、動的な不動産モ

デルと組み合わせた都市シミュレーションモデルに昇華することが長期的課題である。このような取りくみとして、近藤・吉川(2007)を指摘することができる。また、コーホート法と住宅立地モデルを結合した取り組みとして、宮澤ら(2006)、Kanaroglou et al.(2009)、小池ら(2010)等が挙げられる。

謝辞

本研究は、文部科学省の気候変動適応戦略イニシアチブプロジェクトの助成を受けて実施したものである。

参考文献

- 1) 有賀敏典, 松橋啓介, 米澤健一: 自然増減と社会増減を明示的に考慮した地域内人口分布の変化—1980年から2005年までの全国国勢調査・基準地域メッシュデータを用いて—, 都市計画論文集, 46(3), 847-852.
- 2) 石川晃: 市町村人口推計マニュアル, 古今書院, 1993.
- 3) 江崎雄治: 首都圏人口の将来像—都心と郊外の人口地理学—, 専修大学出版会, 2006.
- 4) 大澤義明, 小野田竜巳, 小林隆史: コーホート変化率法による地域別人口予測の集計誤差, 日本建築学会計画系論文集, 73(634), 2605-2612.
- 5) 岡崎陽一: 人口統計学(増補改訂版), 古今書院, 1999.
- 6) 奥村誠: 国勢調査メッシュデータに基づく地区の将来人口構成予測手法, 都市計画論文集, 40(3), 193-198, 2005.
- 7) 久保川達也: 線形混合モデルと小地域の推定, 応用統計学, 35(3), 139-161, 2007.
- 8) 栗田治, 腰塚武志: 領域間平均距離の近似理論とその応用, 都市計画論文集, 23, 43-48, 1988.
- 9) 小池淳司・漆谷敏和・山崎清: 世帯別立地行動を考慮した都市経済モデルの開発, 土木計画学研究・講演集, 41, CD-DOM(391), 2010.
- 10) 小池司朗: 地域別将来人口推計における人口移動モデルの比較研究, 人口問題研究, 64(3), 87-111, 2008a.
- 11) 小池司朗: 地域別将来人口推計における純移動率モデルの改良について, 人口問題研究, 64(1), 21-38, 2008b.
- 12) 小池司朗: 首都圏における時空間的人口変化—地域メッシュ統計を活用した人口動態分析—, 人口問題研究, 66(2), 26-47, 2010.
- 13) 古藤浩: 小地区短期間多地域データからの地区成分解析—山形市町丁目人口データを対象とした分析と人口予測—都市計画論文集, 43(3), 61-66, 2008.
- 14) 古藤浩, 三浦英俊: メッシュデータによる低密度地域の人口推計, GIS-理論と応用, 20(1), 71-80, 2012.
- 15) 近藤玲, 吉川徹: 住宅ストックと世帯構成に着目した都市シミュレーション—東京都を対象として—, 都

- 市計画報告集, 5, 127-133, 2007.
- 16) (財) 統計情報研究開発センター: 小地域統計・境界データの利用に関する研究II, 2003.
 - 17) 沢田英一, 羽根義: 小地域における人口予測手法に関する研究, 清水建設研究報告, 82, 67-72, 2005.
 - 18) 白井信雄, 田中充: 温暖化影響と適応策に関する地域研究の動向, 資源環境対策, 47 (2), 67-72, 2011.
 - 19) 高橋邦彦, 横山徹爾, 丹後俊郎: 疾病地図から疾病集積性へ, 保健医療科学, 57 (2), 86-92, 2008.
 - 20) 谷口守, 松中亮治, 妹尾一慶: 都市撤退 (リバーズ・スプロール) に関する計測手法の開発とその適用, 土木計画学研究・論文集, 24 (1), 183-190, 2007.
 - 21) 丹後俊郎, 高橋邦彦, 横山徹爾: 空間疫学への招待—疾病地図と疾病集積性を中心として (医学統計学シリーズ), 朝倉書店, 2007.
 - 22) 土屋貴佳, 室町泰徳: メッシュ単位の将来人口推計モデルの構築に関する研究, 土木計画学研究・講演集, 32, CD-ROM (77), 2005.
 - 23) 濱英彦, 山口喜一 (編): 地域人口分析の基礎, 古今書院, 1997.
 - 24) 東博紀, 大楽浩司, 松浦知徳: 地球温暖化による豪雨発生頻度の変化と洪水氾濫への影響評価, 水工学論文集, 50, 205-210, 2006.
 - 25) 星田侑久, 佐藤俊明, 荻野寛人, 浅田理恵, 岡部篤行: 500m メッシュ単位の男女別 5 歳階級別の将来人口の推計, GIS-理論と応用, 19 (1), 9-15, 2011.
 - 26) 三村信男: 地球温暖化対策における適応策の位置づけと課題, 地球環境, 1 (1), 103-110, 2006.
 - 27) 宮澤俊治・高木朗義・秋山孝正・大森貴仁: ファジィ推論とGISデータを用いた世代別立地均衡モデルの構築, 土木計画学研究・論文集, 23 (1), 211-220, 2006.
 - 28) 武者忠彦, 江崎雄治, 西岡八郎, 青木賢人, 小口 高: 人口分析における GIS の可能性, CSIS Discussion Paper #48, 2002.
 - 29) 山口喜一: 人口推計入門, 古今書院, 1990.
 - 30) 和田光平: Excelで学ぶ人口統計学, オーム社, 2006.
 - 31) Anselin, L. *Spatial Econometrics: Methods and Models*, Dordrecht, Kluwer Academic Publishers, 1988.
 - 32) Anselin, L., Lozano, N., and Koschinsky, J.: Rate Transformations and Smoothing, Spatial Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign, 85pp, 2006.
 - 33) Fay, R.E., and Herriot, R. Estimates of income for small places: An application of James-Stein procedures to census data, *Journal of the American Statistical Association*, 74 (366), 269-277, 1979.
 - 34) Goovaerts, P. Geostatistical analysis of disease data: Estimation of cancer mortality risk from empirical frequencies using Poisson kriging, *International Journal of Health Geographics*, 4 (31), online, 2005.
 - 35) Haining, R.P., Kerry, R., and Oliver, M.A.: Geography, spatial data analysis, and geostatistics: An overview, *Geographical Analysis*, 42 (1), 7-31, 2010.
 - 36) Kanaroglou, P.S., Maoh, H.F., Newbold, B., Scott, D.M., and Paez, A.: A demographic model for small area population projections: An application to the Census Metropolitan Area of Hamilton in Ontario, Canada, *Environment and Planning A*, 41(4), 964-979, 2009.
 - 37) Pratesi, M., and Salvati, N.: Small area estimation: the EBLUP estimator based on spatially correlated random area effects, *Statistical Methods and Applications*, 17 (1), 113-141, 2008.
 - 38) Rees, P., Norman, P., and Brown, D.: A framework for progressively improving small area population estimates, *Journal of the Royal Statistical Society A*, 167 (1), 5-36, 2004.
 - 39) Waller, L.A., and Gotway, C.A.: *Applied Spatial Statistics for Public Health Data*, New York, John Wiley & Sons, 2004.
 - 40) Yoo, E.-H., and Kyriakidis, P.: Area-to-point Kriging with inequality-type data, *Journal of Geographical Systems*, 8 (4), 357-390, 2006.
 - 41) You, Y., and Zhou, Q.M.: Hierarchical Bayes small area estimation under a spatial model with application to health survey data, *Survey Methodology*, 37 (1), 25-37, 2011.

(2012. 8. 2 受付)

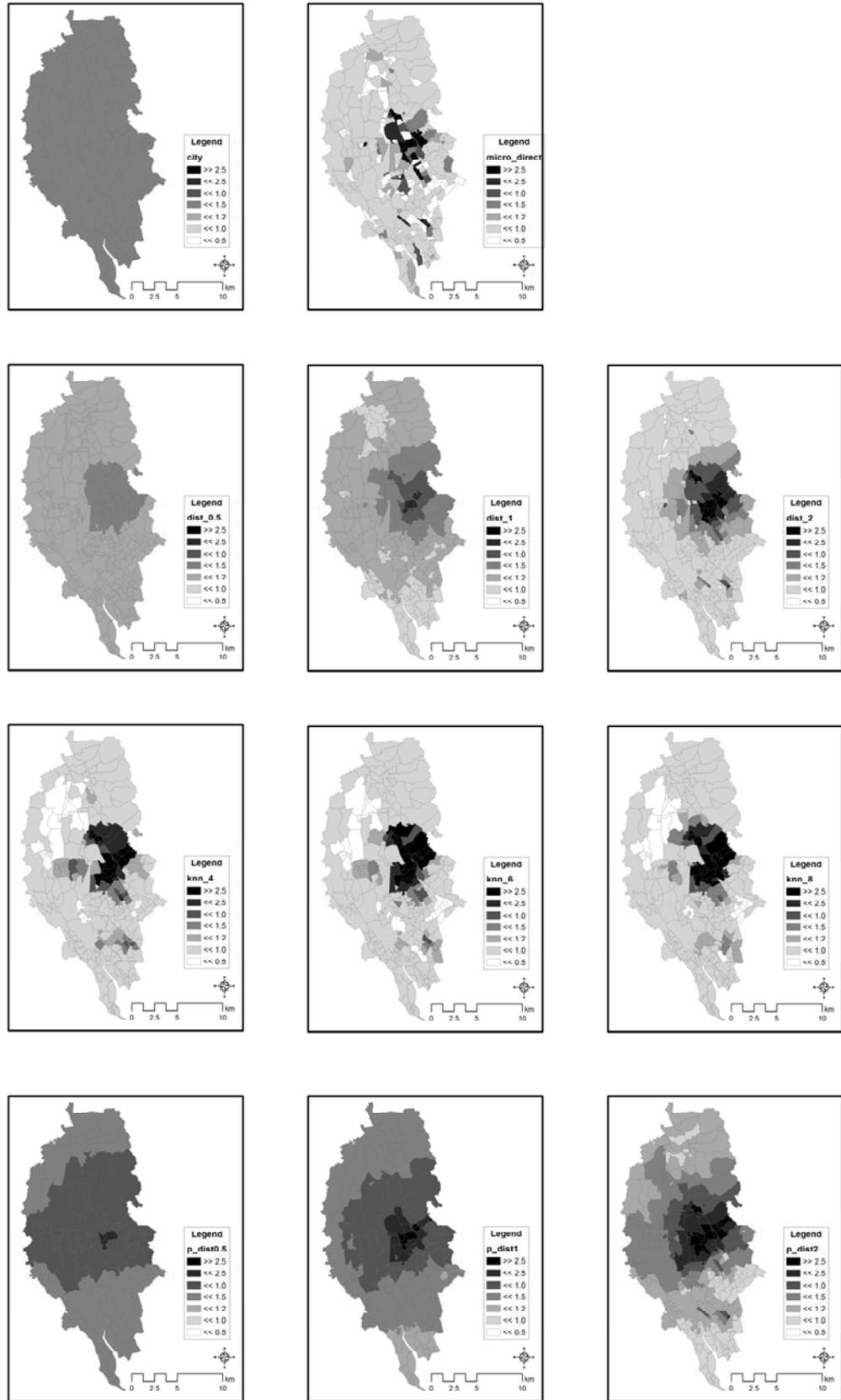


図3-1：地域別コーホート変化率

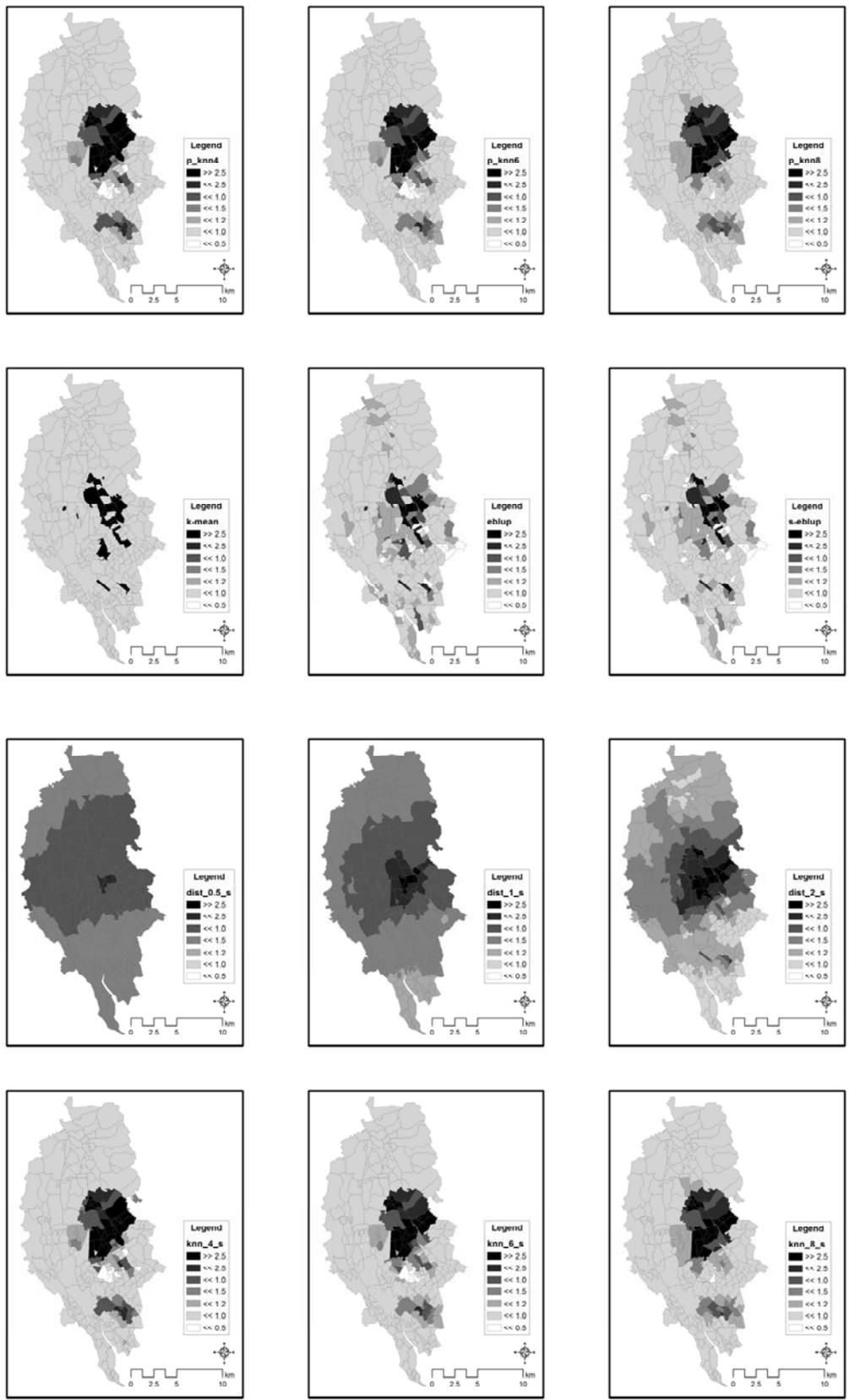


図 3 - 2 : 地域別コーホート変化率

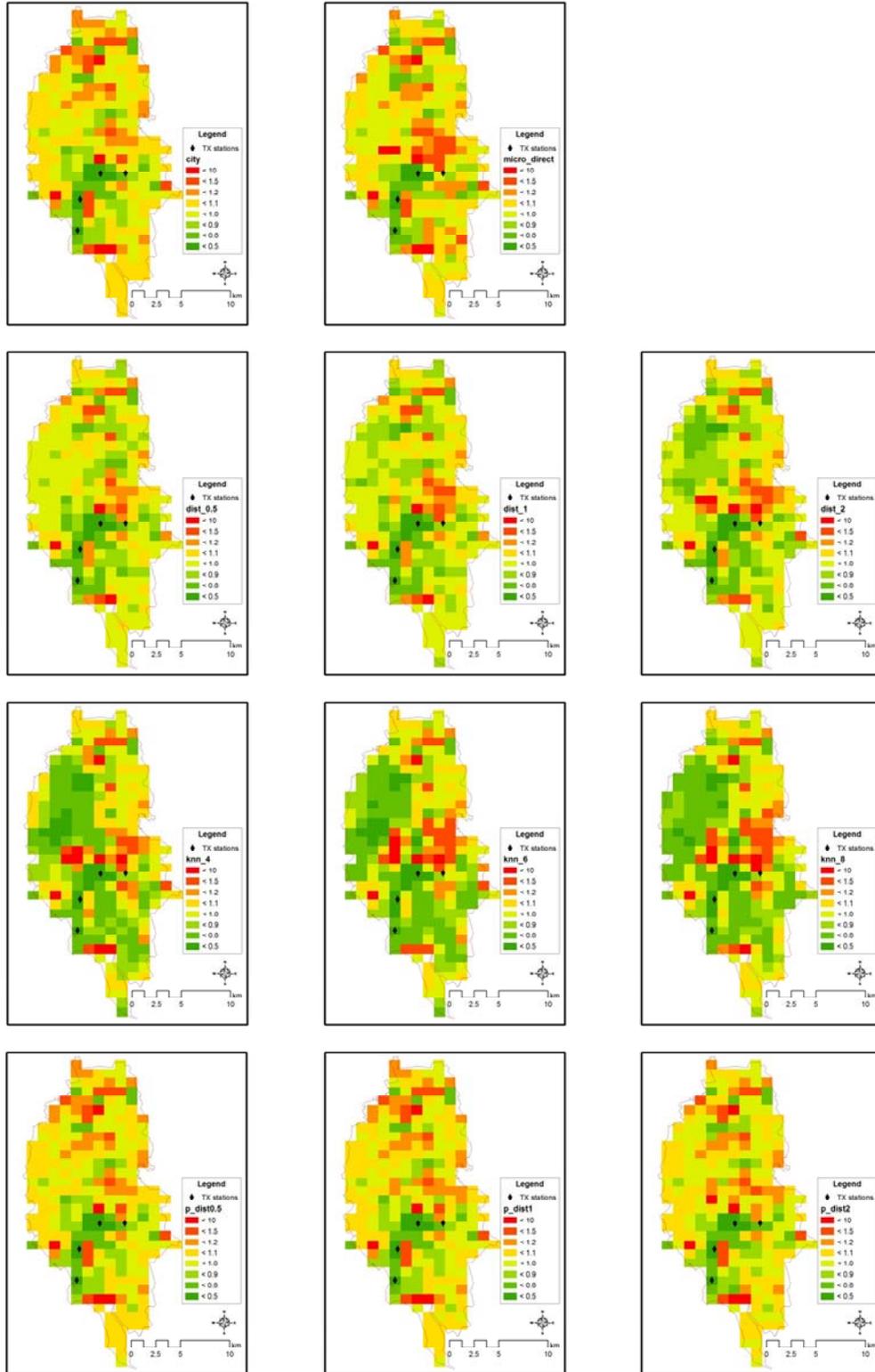


図 4-1 : 推計誤差 (推計値/実測値) の分布

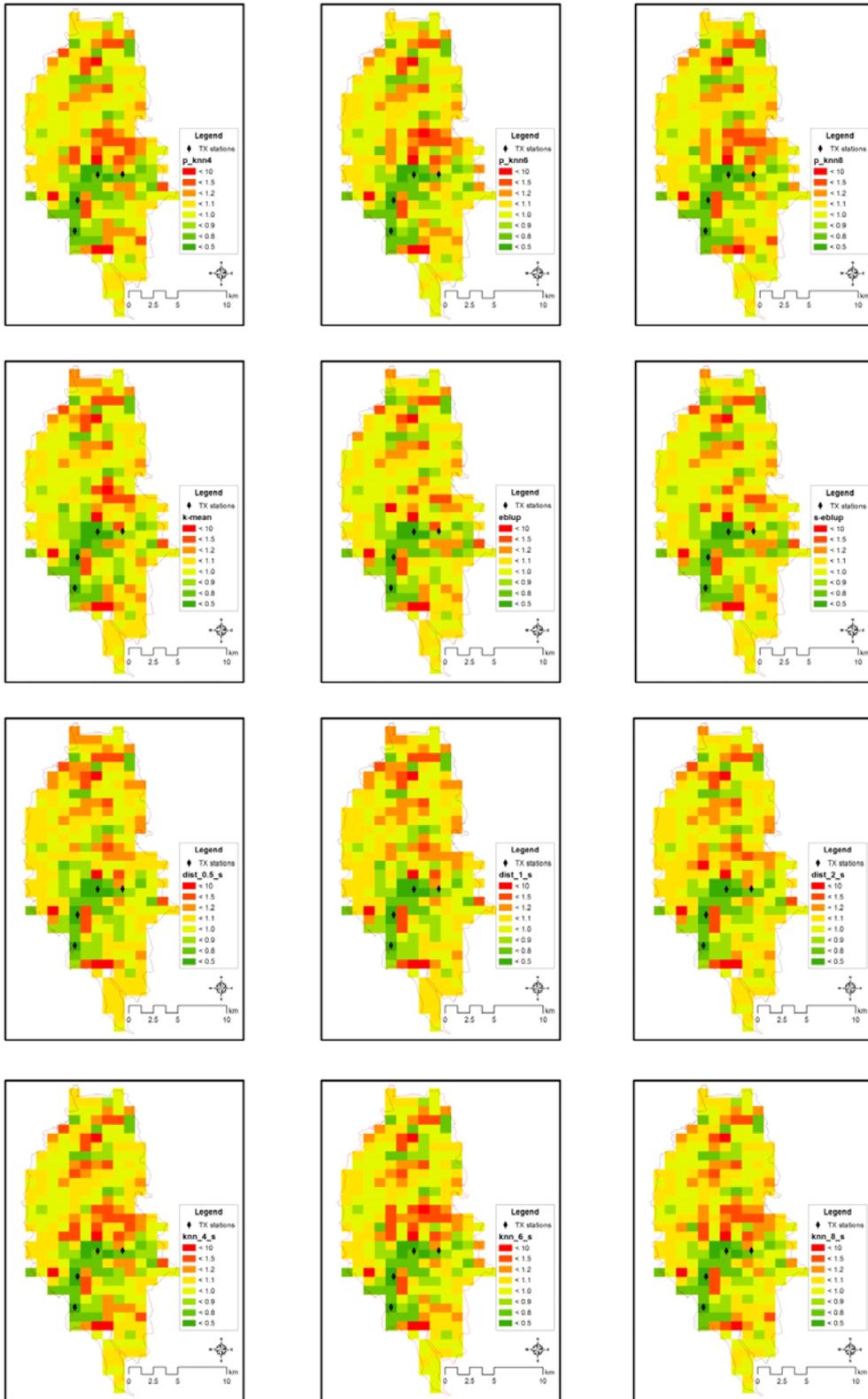


図 4 - 2 : 推計誤差 (推計値 / 実測値) の分布