

【原著論文】

メッシュデータを用いた人口増減地区の抽出方法

奥貫 圭一・草野 邦明

地理情報科学研究室

A method for detecting districts with significant population changes

Kei-ichi Okunuki and Kuniaki Kusano

Geographic Information Science

Abstract

This paper discusses a method of detecting districts in which the population is increasing or decreasing. Population change ratios are usually used to account for population change. However, when we try to find the districts with population increase or decrease by drawing a choropleth map showing the ratios of population changes, the results are able to vary depending on the reader's ability of map reading. Furthermore, the resulting districts are able to have the same tendency of population increase/decrease as their neighboring areas, and the increase/decrease should not be considered to be that of the districts locally, but to be that of the neighboring areas entirely. Therefore, we propose a method applying GWR: Geographic Weighted Regression model, which enables us to consider local population changes, to mechanically find the population increase/decrease areas without the need for human map reading. An example of the proposed method is shown, where a set of 1km mesh data of the census 2000, 2010 and 2020 in the case of Gunma prefecture is analyzed. Through the example, it is shown that the proposed method can be practically carried out under the use of GIS application software such as ArcGIS.

キーワード : 局所的人口増減, メッシュデータ, 地理的加重回帰

1. はじめに

本稿では、人口の空間的分布に着目して、人口の増減が顕著な地区を抽出する方法について論じる。わが国においては、長年にわたって少子高齢化が続き、人口減少の局面にある。この人口減少に起因するさまざまな社会的な問題が懸念されるとともに、諸富（2018）のように「人口減少時代」の社会

のあり方を論じたものも多く見られるようになっている。このように「人口減少時代」や「人口減少社会」といった言葉を目にすると、一般の人々は、あたかも日本国内のあらゆるところで人口が減少するかのように受けとめるかもしれない。しかし、将来において日本の総人口が減少していくとしても、国内のそれぞれの地域で人口の空間的分布を見ていけば、人口減少の地区がある一方で、人口増加の地区もあるかもしれない。地域全体の人口増減をとらえたとしても、地域内の各地区における局所的な人口増減が同じ傾向であるとは限らない。これからの人口減少の時代に生ずるであろう社会的な問題に対応していくためには、地域内のどこで人口が減少し、どこで人口が増加するのか（もしくは人口が横ばいであるか）、局所的な人口増減を把握することが大切であろう。

局所的な人口増減を把握しようとするとき、難しい課題がある。すなわち、地区の人口増減がその地区独自の動向とその地区を含む周辺地域全体の動向とを複合的に反映していることである。人口の増減を論じる際、人口増減率を指標とすることが一般的であり^{注1)}、ある地区の人口増減率が100%を超えていれば増加、100%を下回っていれば減少、と判断すれば良いように思えるかもしれない。しかし、地区の人口増減率から地区独自の動向が反映した部分だけを抜き出して、局所的な増加や減少を判断することは簡単でない。たとえば、ある地域で全体として人口増減率が100%を大きく上回っていたとして、その地域の中の各地区の人口増減率が100%を上回っていたとしても、それは、地域全体の人口増減傾向を反映しているだけのことで、それぞれの地区が特に顕著な人口増減を示しているとは判断できないかもしれない。ある地区の人口増減について、地区独自の動向（すなわち、その周辺地域全体の動向を除いたもの）を局所的な人口増減とみなすこととすれば、この局所的な人口増減をとらえるための方法は確立されていない。そのような方法があれば、地域の中で地区ごとに抱える社会的問題を考える上でも有用と考える。たとえば、ある都市で政策を検討するとき、ある都市のある地区では高齢化が進み人口減少が顕著であり、同じ都市の別の地区では新たな都市開発がなされ人口増加が顕著であるとすれば、同じ都市内でも異なった政策が求められるであろう。そこで本稿では、局所的な人口増減をとらえる簡便な方法を考えてみたい。

局所的な人口増減を考えると、すぐに思いつく素朴な方法は、人口増減率の分布を図化することであろう。空間的分布を図化するときは、多くの場合、階級区分図が採用されるので、人口増減率の階級区分図を描いて人口増減の空間的分布を論じれば良さそうである。人口の空間的分布のデータは、国勢調査の小地域（町丁・字等別）や1kmメッシュ（3次メッシュ）を空間単位とするものが複数時点について公開されており、「e-Stat 政府統計の総合窓口」(<https://www.e-stat.go.jp/>)から入手できる。このデータは、一般のGIS（地理情報システム）へ容易に読みこむことができるので、GISを活用して階級区分図を描けば良い。ところが、描いた階級区分図を解釈する段階になると、「このあたりは人口が顕著に減少しているようだ」とか、「増加しているようだ」などと、図を読み取る人の主観的な判断に委ねることになり、その人の読図能力によってさまざまな解釈がなされ得る。元来、階級区分図には、階級の分け方などによって、地図から伝わる情報が違ってしまうという問題があり（若林(2022)）、この「さまざまな解釈」を避けることは難しい。階級区分図を読み取ることなく、局所的な人口増減

をとらえる方法があると良い。

以上を踏まえて、以下では、階級区分図を採用することなく、できるだけ周辺地域全体の人口増減傾向から切り離して、局所的に顕著な人口増減を示す地区を抽出する方法を提案してみたい。

2. 分布図による人口増減の判断

ここでは、まず、分布図から現象を解釈することの難しさを確認してみたい。例として、「e-Stat 政府統計の総合窓口」からダウンロードできる国勢調査 1km メッシュデータを利用して各メッシュの人口を図化する。図 1 は、群馬県の、2000 年、2010 年、2020 年の 3 時点の人口分布図をそれぞれ描い

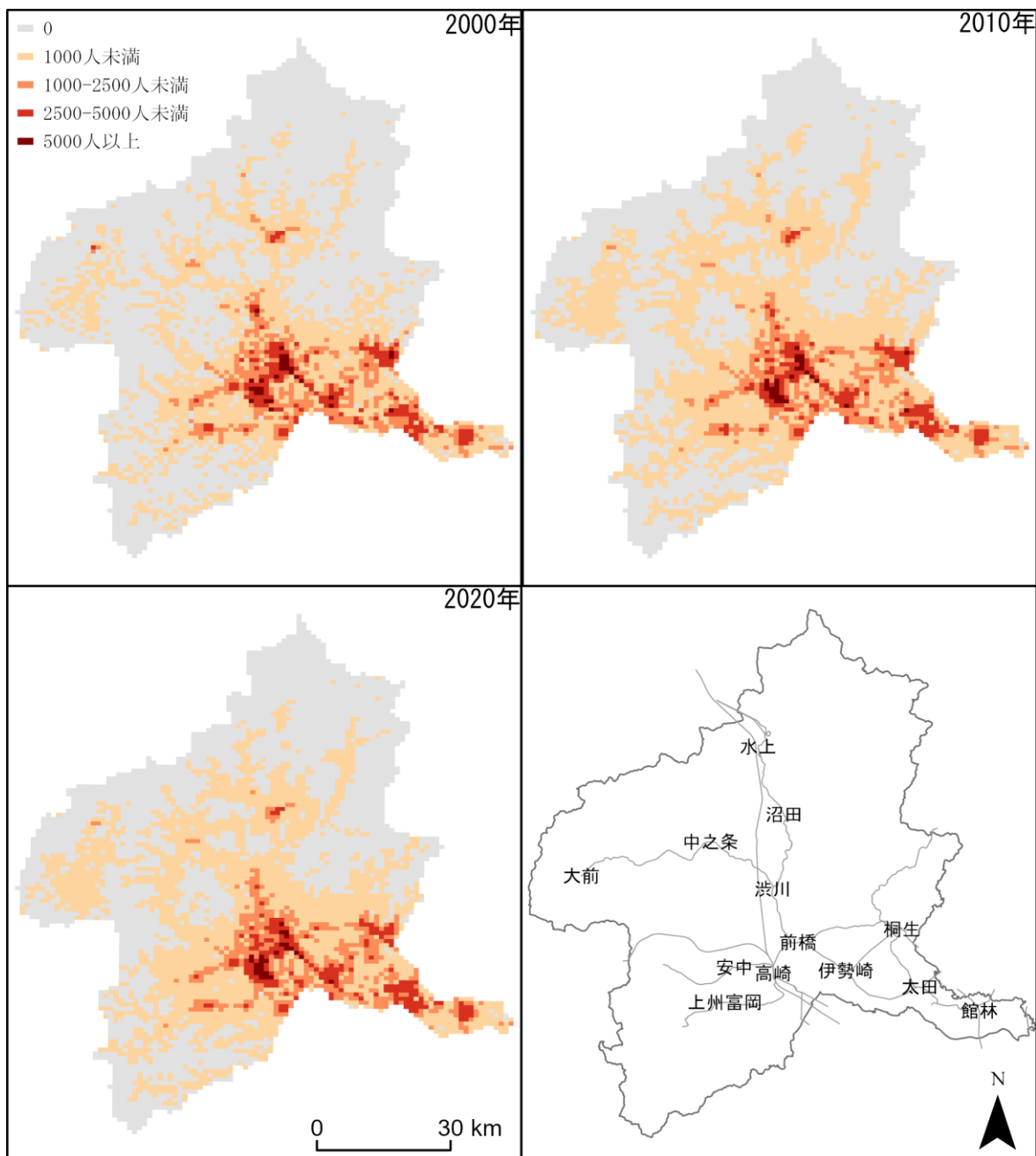


図 1. 2000 年、2010 年、2020 年の各時点における群馬県内人口分布
(国勢調査 1km メッシュデータを図化したもの)、鉄道網と主要駅

たものである。それぞれの図で県内各都市の位置がわかるように、鉄道網と主要駅を示す図も示した（図1右下）。描画にあたっては、ArcGIS Pro 3.0を用いた。この3時点の県内総人口を集計してみると、2000年に1,953,989、2010年に1,941,331、2020年に1,876,988となっており、2000年から2010年の人口がほぼ横ばいであるのに対して、2010年から2020年の期間では、人口増減率がおおよそ0.97とやや人口減少の傾向である。一方、図1から、人口分布の傾向は各時点で大きな違いがないように見える。わずかな違いとして、2000年と2010年の分布図を見比べてみると、この2時点間で、（グレーで色づけされた）人口が0の地区がわずかに減り、（ベージュで色づけされた）人口が1000人未満の地区になっているようである^{注2)}。2000年以後は、すでに都市の縮退への対策が論じられ始めていた（たとえば、海道（2007））ものの、群馬県では、2000年を過ぎても、農村部での市街化が進んでいた可能性がある。このことについて、野澤（2016）は、都市中心部とその隣接地域との間の都市計画規制の違いが影響し、都市中心部で人口が減少していても、開発規制の緩い隣接地域へ人口や開発需要が流出したと指摘している。

さて、この図1は、序章で述べた目的、すなわち、群馬県内の局所的な人口増減を把握するために十分な情報を提供できているであろうか。上記のとおり、3時点で、人口分布の傾向（どこで人口が相対的に多く、どこで相対的に少ないのか）はおおよそ違いがないように見える。しかし、その判断と解釈は、この図を見る人に依存しており、客観的に結論を得ることは難しい。

人口の増減を論じる際には、一般に、人口増減率を指標とする。いま、ある時点 t におけるメッシュ i の人口を $p(t)_i$ とすると、基準とする時点 t_0 に対する（時点 t_0 以後の）任意の時点 t のメッシュ i の人口増減率：

$$p_r(t_0, t)_i = \frac{p(t)_i}{p(t_0)_i} \quad (1)$$

と書ける。ここでは、図1に示した2000年、2010年、2020年の3時点のデータを用いて、図1の各メッシュについて、 $p_r(2000, 2010)_i$ 、 $p_r(2010, 2020)_i$ の人口増減率を算出し、それぞれの分布図を描くことにする。図2がその分布図である。図2の左の図が人口増減率 $p_r(2000, 2010)_i$ の分布図、右の図が人口増減率 $p_r(2010, 2020)_i$ の分布図である。ここでは、人口増減率の値が1以上である1kmメッシュを赤系統の色をつけて示し、0.85に満たない1kmメッシュを青系統の色をつけて示してある。この図2を見ると、いずれの時点においても、都市部（図1右下の主要鉄道駅のあたり）で赤系統の色が多く見られ、そうでないところで青系統の色が多いようである。2010年代には、2000年代と比べると、人口増減率が1.15以上の値を示す（濃い赤で色づけした）メッシュが減っているようにも見えるものの、近年においても、都市部において人口増加地区があると分かる。このように分布図を描けば、赤で描かれたメッシュの地区では人口が増加していたと判断することができるし、同様に、青で描かれたメッシュの地区では人口が減少していたと判断することができる。この手順であれば、この分布図を見る人によらず、おおよそ同じ判断がなされるかもしれない（その詳細な検討は後述する）。

しかしながら、図2に示したような人口増減率の分布図にも、いくつかの課題がある。そのひとつ

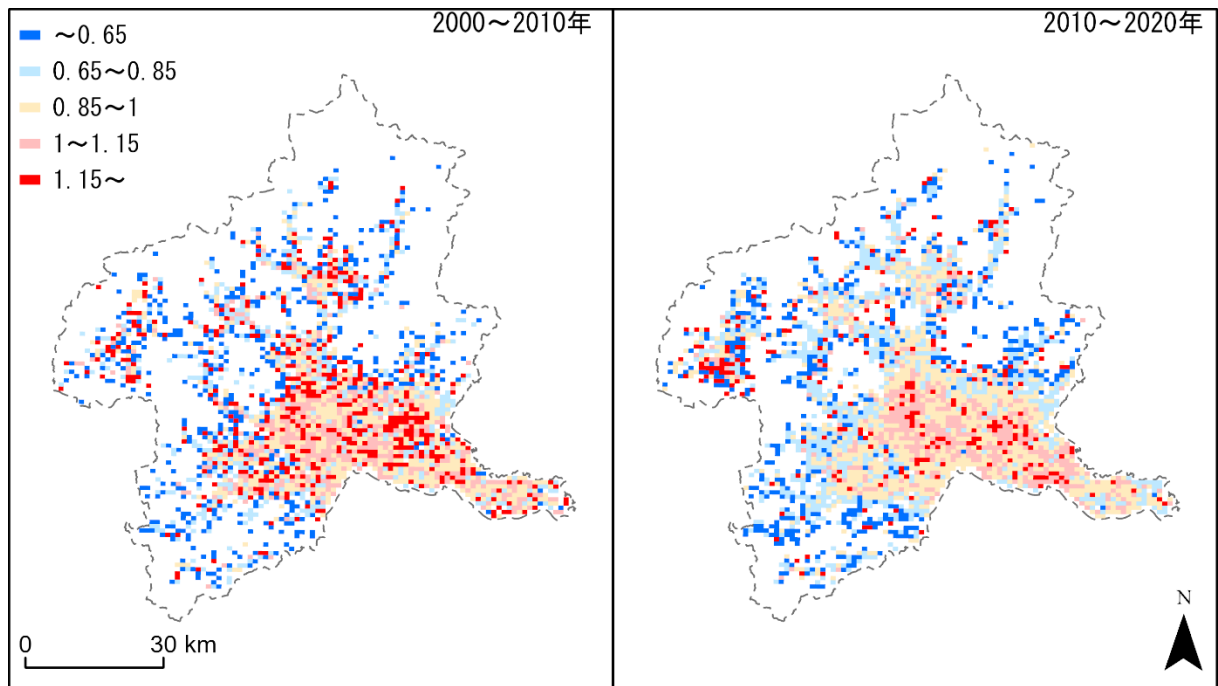


図 2. 2000 年から 2010 年, 2010 年から 2020 年の人口増減率の分布

は、比率を扱っていることである。図 2 の左右いずれの分布図においても、都市部でないはずの地区に赤系統の色が付けられたメッシュがある。式 (1) で示した通り、人口増減率は比率の指標であるので、式 (1) の分子の値が小さなものであったとしても、分母の値もまた小さな値であれば、人口増減率の値は 1 を上回ることがあり得る。ときには、1 を大きく上回ることもあられるかもしれない。たとえば、都市部の地区において人口が 100 人であったものが人口増加して 200 人となった場合と、農村部の集落において人口が 1 人から 2 人に増えた場合とで、いずれも人口増減率は 2 になるものの、それぞれの地区が抱えるはずの問題はまったく違うはずである。こうした違いを見出すためには、人口増減率によらない別の方法が必要であろう。とくに、それぞれのメッシュについて、そのメッシュが位置する地域（メッシュの周辺地域と言っても良い）に応じて、当該メッシュにおける人口増減を判断する方法が良いだろう。

図 2 のような分布図が抱えるもうひとつの課題は、階級の分け方によって解釈が変わり得ることである。この課題は、すでに述べたとおり、人口増減率に限らず、階級区分図について一般に共通するものであり、階級区分図を採用すれば避けて通れないものである。たとえば、図 2 では、人口増減率の値が 1.15 以上である 1km メッシュを濃い赤色をつけて示し、0.65 に満たない 1km メッシュを濃い青色をつけて示した。これを、1.05 以上を濃い赤色、0.85 未満を濃い青色で示すように階級を変えて分布図を描きなおしてみる。図 3 が、その結果として描かれた分布図である。これを見ると、2000 年代と 2010 年代とで、人口増減率の分布図には大きな変化がないようにも解釈できそうである。図 2 では、濃い赤で示された 1km メッシュが減っているようにも見たものが、図 3 になると、そのような傾向を見出せそうにない。同じデータから作図しているにもかかわらず、図 2 と図 3 との間で解釈が

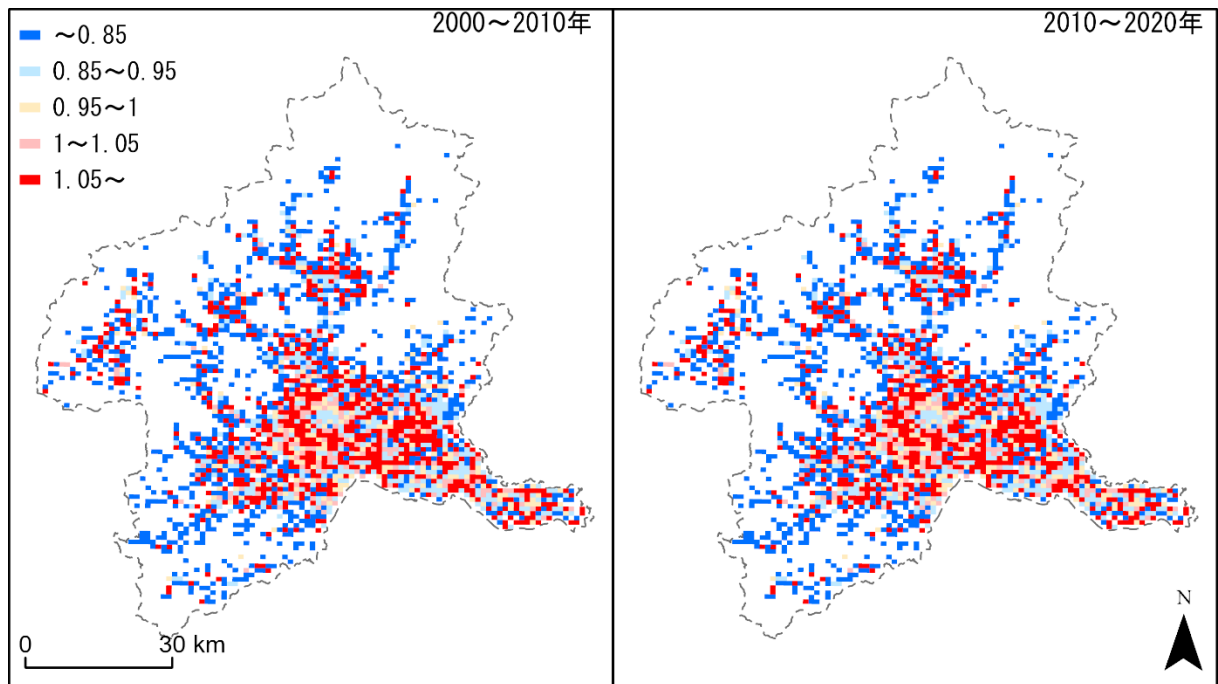


図 3. 2000 年から 2010 年, 2010 年から 2020 年の人口増減率の分布
(人口増減率 1.05 以上を濃い赤, 0.85 未満を濃い青で表示した場合)

変わってしまう。

ここまで述べてきた通り, 人口増減率の階級区分図を描くことで, 地域のどこで人口が増加し, どこで人口が減少しているのか, を判断して, 地域の様相を読み取ることは難しさがある。とくに, 序章でふれたような局所的に顕著な人口増減を見出すとなると, 上記とは別の方法を考える必要がある。そこで以下では, 階級区分図を描かずに, 局所的な人口増減を機械的に判断する方法を考えてみたい。

3. 地理的加重回帰モデルを応用した人口増減地区の抽出

局所的に顕著な人口増減を機械的に判断する簡単な方法として, ここでは, 回帰モデルの応用を考えてみる。すなわち, ある基準とする時点 t_0 におけるメッシュ i の人口 $p(t_0)_i$ を独立変数として, それ以後の任意の時点 t の人口 $p(t)_i$ を従属変数とする回帰モデル

$$p(t)_i = b_0 + b_1 p(t_0)_i + \varepsilon_i \quad (2)$$

を考える。ここで, b_0 は切片, b_1 は回帰係数, ε_i は誤差項である。この式に, 実際に入手できる 2 時点の人口データをあてはめてモデル式を推定すれば (切片と回帰係数の推定値が得られるので, それぞれ, \hat{b}_0 , \hat{b}_1 と書くと), 時点 t におけるメッシュ i の人口の推定値 $\hat{p}(t)_i$ は,

$$\hat{p}(t)_i = \hat{b}_0 + \hat{b}_1 p(t_0)_i \quad (3)$$

となる。すると, 実際に観測された人口の値 $p(t)_i$ と推定値 $\hat{p}(t)_i$ との残差 $r(t)_i$ は,

$$r(t)_i = p(t)_i - \hat{b}_0 - \hat{b}_1 p(t_0)_i \quad (4)$$

で得られる。通常、回帰分析の考え方にならうと、こうして得られるはずの残差の値が比較的大きければ、その値は異常値として疑う候補となる。残差が大きいメッシュでは、地域全体の人口増減の傾向と比べると、その増減の度合いが顕著に大きいと判断できる。この残差の大きさを判断するにあたって、誤差項を正規分布と仮定して、その標準偏差の値を手掛かりにすることが多く、判断の目安となる材料は得られる。こうして、式 (4) に示した残差によって、統計的に有意な人口増減を示すと判断できるメッシュが抽出される。

さて、式 (2) の回帰モデルでは、分析対象地域の各メッシュで、時点 t の人口が時点 t_0 の人口に比例し、係数 b の値が地域全体で同じであることを前提としている。たとえば、ある地区で人口がおおよそ 2 倍になっているとすれば、別の地区でも同様に人口がおおよそ 2 倍になっているだろう、と期待していることになる。しかし、都市域で人口が増加している、山間地域では過疎化が進んで人口が減少しているかもしれない。あるいは、ある都市で都心への人口回帰が進んで人口増加があったとしても、別の都市では郊外部への人口流出が進んでいて都心の人口減少が見られるかもしれない。このように、地域内の地点によって人口増減の傾向は異なることが予想できる。人口増減に限らず、地理的現象がどの地点でも同質とみなせないことを空間的異質性と呼び、地理的問題を考える上で着目すべきひとつとされる。式 (2) の回帰モデルは、空間的異質性を考慮したものでなく、群馬県の事例で言えば、高崎や前橋での人口増減の傾向と、草津や水上での人口増減の傾向とを同じものと期待して考えていることになる。そこで以下では、空間的異質性を考慮した回帰モデルである地理的加重回帰モデル (GWR : Geographically Weighted Regression model) を応用してみたい。

地理的加重回帰モデル (以下、GWR モデル) は、Fotheringham らが提案したものであり (Fotheringham *et al.* (2003)), ArcGIS などの GIS ソフトウェアにもその分析ツールが実装されている。フリーのソフトウェアツールも提供されている他、近年は、R を用いて実践することもできる (村上 (2022))。GWR モデルでは、式 (2) の切片と回帰係数を地点によって異なるものとする、すなわち、

$$p(t)_i = b_0(u_i, v_i) + b_1(u_i, v_i) p(t_0)_i + \varepsilon_i \quad (5)$$

であり、ここで、 (u_i, v_i) はメッシュ i の位置座標を示している。式 (5) のモデルを推定する実際の計算にあたっては、各メッシュについて、その周辺のメッシュの観測値 (本稿の例では人口) に重み付けする。このとき、「周辺」の定義にいくつかのものがあ、そのひとつは、バンド幅 h を与えて、そのバンド幅で覆われる範囲を「周辺」とみなすものである。さらに、重み付けにもいくつかのタイプがあり、メッシュ i との距離が遠くなるほど値が減衰する関数

$$w(d_{ij}) = \begin{cases} \left\{ 1 - \left(\frac{d_{ij}}{h} \right)^2 \right\}^2 & \text{if } d_{ij} < h \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

などがつかわれる。ここで、 d_{ij} は、メッシュ i とメッシュ j との間の距離である。なお、本稿では、メッシュ間の距離を測るにあたって、メッシュの代表点 (おおよそメッシュの中心) の間の距離を測っている。今、上記の式 (6) を重みづけの関数として採用するとして、重み行列

$$\mathbf{W}(u_i, v_i) = \begin{pmatrix} w_{i1} & 0 & 0 & \dots & 0 \\ 0 & w_{i2} & 0 & \dots & 0 \\ 0 & 0 & w_{i3} & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & w_{in} \end{pmatrix} \quad (7)$$

と書く。ここで、 n はメッシュの総数であり、 $w_{ij}=w(d_{ij})$ である。このとき、式(5)の独立変数の行列を \mathbf{X} ^{注3)}、従属変数の観測値のベクトルを \mathbf{y} 、GWRモデルの係数の推定値のベクトルを $\mathbf{b}(u_i, v_i)$ と書けば、その推定値は

$$\mathbf{b}(u_i, v_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \quad (8)$$

を算出することで得られる。

こうして式(5)のGWRモデルが推定されれば、回帰モデルの場合と同様に、各メッシュについて残差(式(4))を求めることで、人口増減が顕著なメッシュを抽出することができる。図4は、上記の手順の結果として、人口増減が顕著であったとされるメッシュを描いたものである。図4の左図は、

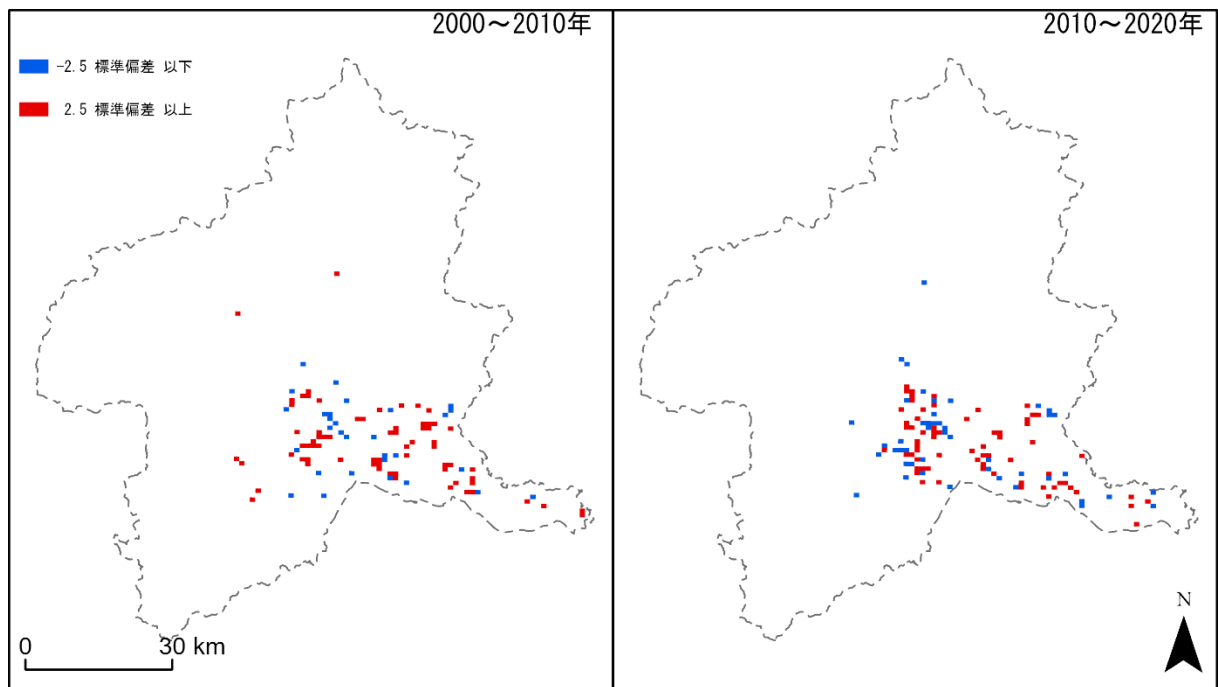


図4. 2000年から2010年、2010年から2020年のそれぞれの人口増減について
GWRモデル推定値からの残差が大きい1kmメッシュ

2010年の人口を従属変数とし、2000年の人口を説明変数としてGWRモデルを推定した上で、そのモデルからの残差の分布を描いたものである。図4右図は、同様に2020年の人口を従属変数、2010年の人口を説明変数とした場合の残差の分布図である。いずれの図でも、とくに残差が大きい値(具体的には、誤差項において標準偏差の2.5倍を超える値)を示すメッシュのみを色付けして示してある。赤色で示された1kmメッシュでは、GWRモデルの推定と比べて人口が顕著に増えていることになり、一方で、青色で示された1kmメッシュでは、GWRモデルの推定と比べて人口が顕著に減っていることになる。なお、ここで、誤差項 ε_i は正規分布に従うことを仮定しており、「周辺」を定義するため

のバンド幅 h は 12km とし、式 (6) の減衰関数を採用して計算している。これらの計算のすべてについて ArcGIS Pro 3.0 を用いた。

こうして得られた図 4 は、この図を読み取る人の読図能力に頼ることなく、局所的に顕著な人口増減地区を抽出したものとなっている。ただし、GWR モデルでの「周辺」の定義や重みづけの仕方など（パラメータにあたるもの）によって、図 4 に示した結果は変わり得る。同じパラメータを採用する限りにおいて、主観によらない結果が機械的に得られるということである。

4. おわりに

本稿では、ある地域の中で、人口の増減が局所的に顕著な地区を抽出する方法について論じてきた。複数時点の人口分布の階級区分図を見比べる方法（図 1）や人口増減率の階級区分図を読み取る方法（図 2, 3）では、図を読む人の判断に依存してしまう問題があることに加えて、地域全体の人口増減の傾向とは別に局所的に顕著な人口増減を見出すことが難しい。そこでここでは、地理的加重回帰（GWR）モデルを応用して、局所的に顕著な人口増減地区を抽出する方法を提案した（図 4）。GWR モデルの実際の計算は ArcGIS などのソフトウェアなどの他、最近では R でも実現できる。人口データもまた国勢調査の 1km メッシュのデータが全国について無償提供されているので、ここで述べた手順を実践できる環境は整っていると行ってよいだろう。

注

- 1) たとえば、国土交通省の人口関係参考資料 <https://www.mlit.go.jp/common/001046872.pdf> をはじめとする政府の検討資料などで採用されている。
- 2) 本稿で採用した 1km メッシュでは、秘匿地域のメッシュがあった場合、そのメッシュの値を別のいずれかのメッシュに合算している。年次によって合算される地域が異なるため、複数時点の間での各メッシュの増減は現実のものとは違いがある可能性がある。ここでは、その可能性があるものの、データをそのまま処理している。
- 3) 本稿の GWR モデルでは独立変数を 1 つとしているので、行列とせずにベクトルとして良い。しかし、一般には、独立変数が複数あるので、ここでも行列表記してある。

引用文献

- 海道清信, 2007. 「コンパクトシティの計画とデザイン」, 学芸出版社.
- 野澤千絵, 2016. 「老いる家崩れる街 住宅過剰社会の末路」, 講談社現代新書.
- 村上大輔, 2022. 「実践 Data Science シリーズ R ではじめる地理空間データの統計解析入門」, 講談社.
- 諸富 徹, 2018. 「人口減少時代の都市 成熟型のまちづくりへ」, 中央公論新社.
- 若林芳樹, 2022. 「デジタル社会の地図の読み方 作り方」, 筑摩書房.
- Fotheringham, A. S., Brunson, C. and Charlton, M. 2003. *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. John Wiley & Sons, West Sussex.