
東北芸術工科大学 紀要

BULLETIN OF TOHOKU UNIVERSITY OF ART AND DESIGN

第28号 2021年3月

人口予測における誤差から考察する予測方法・地域の影響

Influence of Regional and Methodological Variety from the Viewpoint of Errors of
Population Forecasting

古藤 浩 | KOTO Hiroshi

人口予測における誤差から考察する予測方法・地域の影響

Influence of Regional and Methodological Variety from the Viewpoint of Errors of Population Forecasting

古藤 浩 | KOTO Hiroshi

Population forecasting is the work to obtain the basic preconditions for socio-economic planning or city planning. However, there are always errors in the prediction results. In addition, although elaborate methods have been developed, it is inevitable that there will be some error as long as it is a prediction. Therefore, it is necessary to use the forecast result after understanding how much error can occur. It is expected that the error will increase if there occur a calamity. Since the target period of this study is the Great East Japan Earthquake, we will try to measure such impact. Based on the above, in this study, we will compare and analyze the errors caused by three forecasting methods from the results of population forecast of cities, wards, towns, and villages in Japan.

Three forecasting methods are discussed, that are trend forecast, cohort change rate method and cohort factor method. Given the data at two time points in 2005 and 2010, the former two prediction methods are applied. In addition, we will use the results of population estimation by the cohort factor method by the National Institute of Population and Social Security Research, which is used as the standard for population forecast in Japan. It will be the basic forecast verification of how much error will occur in the forecast results of the 2015 population by comparing with the realized values.

Keywords:

人口予測 トレンド予測 コーホート変化率法 コーホート要因法
市区町村

population forecast, trend forecasting, cohort change ratio method,
cohort factor method, municipalities

1. 研究の目的

人口予測は社会経済計画や都市計画の基本的前提条件を得るための作業といえる。しかし、予測結果には常に誤差がつきまとう。また、精緻な手法も開発されているが、予測である以上、ある程度の誤差があるのは免れない。そのため、どの程度の誤差が出うるかを理解した上で予測結果を用いる必要がある。また、予測には、単純なトレンド予測から最も信頼性が高いと言われるコーホート要因法まで様々な手法がある。必要とする精度に対応してどのような予測をするのがよいか、作業の労力との関係も含めて考察・選択することが望ましい。さらに、震災など大きな事件があると誤差は大きくなることが予想される。本研究の対象期間は東日本大震災を挟んでいるのでそのような場合のインパクトの計量も試みる。

以上をふまえ、本研究では3つの予測手法による誤差を日本の市区町村別人口予測の結果から比較考察する。

予測手法は、2時点データからのトレンド予測とコーホート変化率法、さらにコーホート要因法を取り上げる。2005年・2010年の2時点データを所与として前者2つの予測手法を適用する。また、日本で人口予測の標準として用いられている国立社会保障・人口問題研究所によるコーホート要因法での人口推計結果を3番目の予測結果として用いる。短期間予測であるが、2015年人口の予測結果はどの程度の誤差が出るかを、実現値からの比較によって検証することは、予測検証の基礎となろう。

人口推計による誤差・推計精度について、J. Tayman et.al (1988)、S. K. Smith and M. Shahidullah (1995)、T. Wilson, F. Rowe (2011)、山内他(2015)、阿藤他(1987)、山田(2011)など国内外で研究され、誤差の発生理由や統計的な精度について議論されている。また都市計画等での人口推計の必要から相(2017、2018)、大澤他(2008)、川崎他(2018)、古藤・三浦(2012)では小地域に注目した人口推計とその精度の検証がおこなわれている。これらでの分析指標の主流は、絶対誤差率(MAPE)および誤差の平均率(MALPE)である。本研究でも、この2指標に注目する。また、J. Tayman et.al (1988)では誤差率と人口規模の関係に注目しており、人口が多いほど指数的に誤差率が小さくなることを指摘している。本研究の対象地域は数百人から数十万人まで人口規模が様々なので、この研究も参考に分析をすすめる。

さらに本研究は東日本大震災の起きた2011年を挟んだ期間を対象期間内とするので、このような大災害が人口推計の結果にどのように表れるかにも注目する。

2. データの概要と3つの予測モデル

2.1 データの概要

2005年国勢調査および2010年国勢調査の市区町村別5歳階級人口を本研究で使う予測の説明用データとする。そして、予測結果と2015年国勢調査の結果から求められる誤差を研究対象とする。また、コーホート要因法による予測結果は国立社会保障・人口問題研究所による結果(2013)を利用する。

2005年国勢調査時の市区町村人口に注目し、2015年まで合併や吸収が一度もなかった市区町村を対象とする。分析対象となるのは1598市区町村数となった。

2015年人口で見ると、その中での最大値はさいたま市の40歳～44歳人口の109412人、最小値は東京都利島村・青ヶ島村の15歳～19歳人口の0人である。また対象総人口は1億1161万人で日本国の総人口の87.8%となる。

また、この期間の人口増加率は0.4%の微増傾向で、対象データでの人口増加率最大は東京都中央区の43%、最小は宮城県女川町の-41%であった。中央区の増加は近年のタワーマンションの増加によるものであろうか。また、女川町の減少については東日本大震災の影響がうかがわれる。

2.2 人口予測モデル

本研究では3種類、トレンド予測、コーホート変化率法、コーホート要因法の人口予測方法を適用する。まず下のよう

P_a^y : y 年の年齢階級 a の人口
ただし、 a : a 歳～ $a+4$ 歳 とする。

T_a^y : y 年の年齢階級 a の人口のトレンド予測値

R_a^y : y 年の年齢階級 a の人口のコーホート変化率法による予測値

F_a^y : y 年の年齢階級 a の人口のコーホート要因率法による予測値

年齢階級 a は $a = 0, 5, 10, \dots, 85$ (ただし $a = 85$ の場合は 85 歳以上) の値をとり、 a 歳 $\sim a + 4$ 歳を示すとする。また、予測値の評価はコーホート変化率法適用時の都合によっての値とした。

2015 年の年齢階級 a の人口のトレンド予測による予測値は式 (1)

$$\begin{aligned} T_a^{2015} &= P_a^{2010} + (P_a^{2010} - P_a^{2005}) \\ &= 2P_a^{2010} - P_a^{2005} \end{aligned} \quad (1)$$

と書くことができる。トレンド予測は過去の動態が線形にそのまま続くと考えて延長する手法で、人口予測に限らず様々な事物の短期的な予測において使われる。

次に 2015 年の年齢階級 a の人口のコーホート変化率法による予測値は式 (2)

$$R_a^{2015} = (P_a^{2010} / P_{a-5}^{2005}) P_{a-5}^{2010} \quad (2)$$

によって定義される。式 (2) は 5 年後の人口を考える場合は現在の一階級若い人口がベースとなり、そこに過去の増減率がかかって与えられると考える手法である。この手法は小地域の予測など後で述べるコーホート要因法が使いにくい場合などにコーホート予測の簡便法として使われることがある。

なお、コーホート変化率法で 85 歳以上人口の予測場合は年齢の上限がないことから式 (3) によってなされる。

$$R_{85}^{2015} = \left\{ P_{85}^{2010} / (P_{80}^{2010} + P_{85}^{2005}) \right\} (P_{80}^{2010} + P_{85}^{2010}) \quad (3)$$

また、 $a = 0$ の場合は合計特殊出生率や子ども女性比などより複雑な式によって求められるが、本研究は予測手法そのものの研究ではないので分析対象から除いた。

コーホート変化率法に移動率など様々な情報を加えたのがコーホート要因法である。式 (2) は当該自治体の過年度の値だけを使うが、コーホート要因法ではある程度の期間と、より信頼性があると考えられる広域の生存率 (式 (2)

の括弧内に相当) を用い、さらに地域によって異なる社会移動率を要因として与えて予測する。詳細については国立社会保障・人口問題研究所 (2011) などを参照されたい。

3. 全市区町村による分析結果

3.1 MAPE と MALPE

ここでは、まず、3 種類の人口予測の精度を全体から分析する。人口予測の評価に使われる主要指標として絶対誤差率 (MAPE: The mean absolute percent error) と平均誤差率 (MALPE: The mean algebraic percent error) があげられる。これらはある n 個のデータの (地域または年齢層など) の集団を考え、そこで、MAPE は誤差の平均的な大きさを与えるが誤差の方向性は考慮しない。一方、MALPE は誤差の方向性、バイアスを考慮した平均値を与える。トレンド予測を例にとると、MAPE、MALPE は次のように定義される。

$$\begin{aligned} MAPE &= \frac{\sum |T_a^{2015} - p_a^{2015}| / p_a^{2015}}{n} \\ MALPE &= \frac{\sum (T_a^{2015} - p_a^{2015}) / p_a^{2015}}{n} \end{aligned}$$

他の推計方法についての MAPE、MALPE は T_a^y を他の推計値に置き換えれば得られる。また、この計算において、実現人口が 0 人の場合 (2 節で紹介した二村 15 歳 $\sim$ 19 歳人口) のデータは計算不能なので除いた。

3 種類の予測について年齢別に MAPE を見たのが図 1 である。ここからわかるとおり、トレンド予測は誤差が大きく、全年齢層で 10% 以上のずれとなった。2 つのコーホート法を用いた予測結果は 15 歳 $\sim$ 29 歳の一般に社会移動が多く予測が難しい年代で 10% 前後の誤差が出ているが、それ以外では前後の年齢層と 85 歳以上を除き 5% 以下の誤差となった。また、一般的にはコーホート要因法の適用が最良と言われているが、15 歳 $\sim$ 29 歳、85 歳以上ではコーホート変化率法の精度の方が高いという結果となった。単純に直近の傾向から予測したコーホート変化率法の方が正しい予測となるということになった。

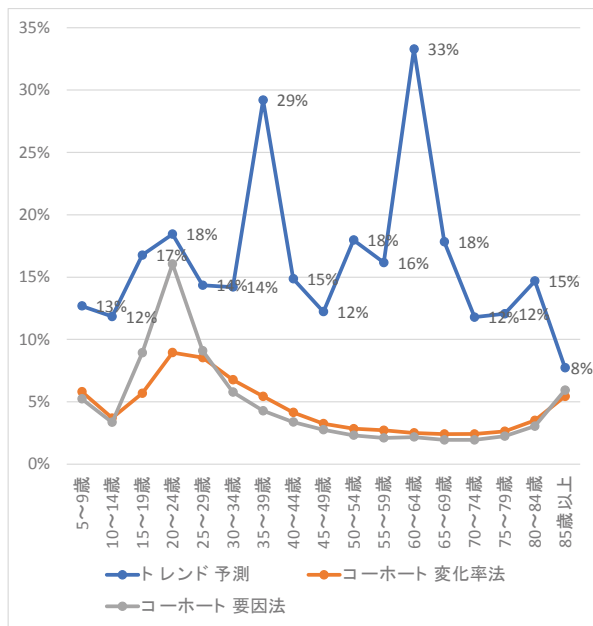


図1 3種類の予測手法によるMAPE

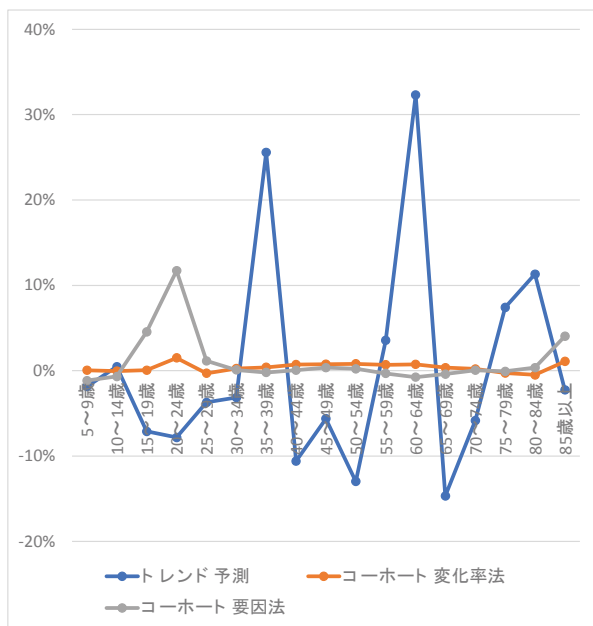


図2 3種類の予測手法によるMALPE

コーホート要因法では移動率を社会情勢から設定するが、その難しさを意味しているように思われる。ただ、10年後、15年後と先の予測を考えるとどちらがより精度が上がるかは別の問題があるように考えられる。ただし、移動率は過去の社会情勢（経済、環境など）によりきまり、未来の社会情勢は不確定のため反映できないとされている（文献

16）。

一方、方向性もふまえた誤差評価、MALPEを年齢別に見たのが図2である。ここでもトレンド予測の誤差が大きく、短期間であっても人口予測では適さないことがはっきりわかった。誤差の方向性についても一貫性を感じられない。

MALPEでもコーホート変化率法の方がコーホート要因法よりも方向性をふまえても精度が高い。コーホート要因法でのMALPEで20歳～24歳は11.7%なので、全体に過大推計をしていることになる。

以上から、トレンド予測は人口予測に全く適さないことがわかったので、以降ではコーホート変化率法とコーホート要因法に絞って結果を示す。

3.2 誤差における人口規模の影響

人口規模別でMAPEを求めると、人口が少ないほどMAPEが大きい傾向がわかる。たしかに、総人口が少なければ率で測るMAPEは大きな値が出やすい。いくつかの年齢層のコーホート変化率法によるMAPEについて、人口規模との関係を見たのが図3である。

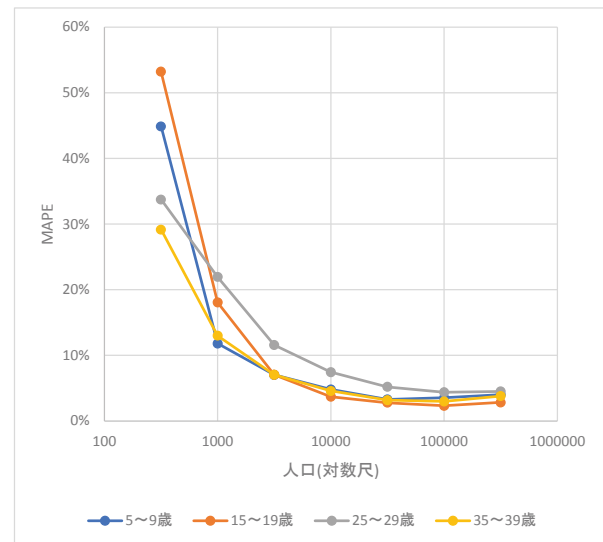


図3 人口規模別MAPE(横軸は対数尺)

図3の人口は対数値で示される人口以下の市区町村のグループの平均で与えているが、どの年齢層でも人口が多いほどMAPEが小さいことがわかる。この傾向はJ. Tayman et.al. (1988)やH. Koto (2019)でも指摘しており、またべき関数減衰で近似できることを提示している。す

なわち、ある係数 s, u によって、

$$MAPE = sP^u \quad (4)$$

(本節では記号の簡単化のため、 $p = p_a^{2015}$ とする)

という関数で大まかな傾向が示される。図3のデータの傾向も同様と見なせるので、コーホート変化率法での4つの年齢層に共通な係数を図4のように両対数平面上の直線回帰によって求めると、

$$MAPE = 1.696P^{-0.454}$$

となった(決定係数は0.779)。コーホート要因法でも同じ4つの年齢層に共通な係数を両対数平面上の直線回帰によって求めると、

$$MAPE = 1.795P^{-0.450}$$

となり(決定係数は0.697)。ほぼ同様の傾向となる。係数 u が -0.45 ほどなので -0.5 に近いと考え近似して考えれば、

$$MAPE \doteq \frac{1.696}{\sqrt{P}}$$

となるので、予測結果の人口の影響をMAPEで見る場合、概ね人口の平方根に反比例するといえる。

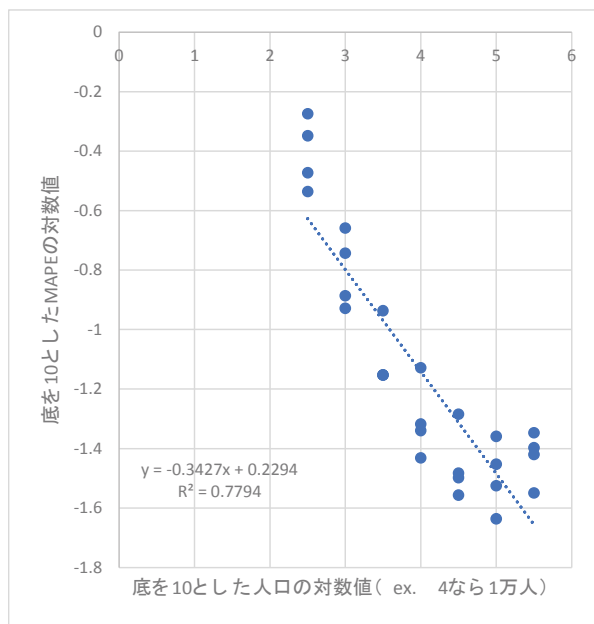


図4 両対数平面での人口規模別補正MAPEと回帰結果

4. 都道府県別の誤差の傾向

本節では、以上をふまえて都道府県別データでのMAPE、MALPEを検証する。2010年～2015年の期間には東日本大震災をはじめとした日本人口にも影響を与えるように思われる大災害が起きている。この影響は予測誤差に現れているか検証したい。ただ、国立社会保障・人口問題研究所では福島県の2015年市町村別人口予測をおこなっていない。これは震災の影響の推定が難しいこと、2015年現在では福島第一原子力発電所事故の影響で全町避難中の町もあったためと思われる。以上をふまえ、ここでは東北地方と関東地方の各都県に注目し、コーホート法の予測結果での誤差を調べる。

MAPEについての東北地方5県の予測結果は図5・図6に示される。コーホート変化率法・コーホート要因法いずれで見ても青森県・宮城県・岩手県の誤差が大きめなのは震災の影響と思われる。また、コーホート要因法でのMAPEは3.1節で見たのと同様に20歳～24歳での誤差が大きい。社会移動が予測以上に多かった・または少なかったことがわかる。ただ、他の年齢層ではコーホート要因法の誤差は小さい。

関東地方の都県のコーホート法によるMAPEを図7・図8に示す。比較的大きな誤差が25歳～34歳で発生している。東京都は全年齢でどちらの予測方法によっても3%を超えており、東北地方のどの県よりも誤差が大きい。若い年齢層が東京に入っていく傾向による誤差は震災の影響よりも大きいように思われる。

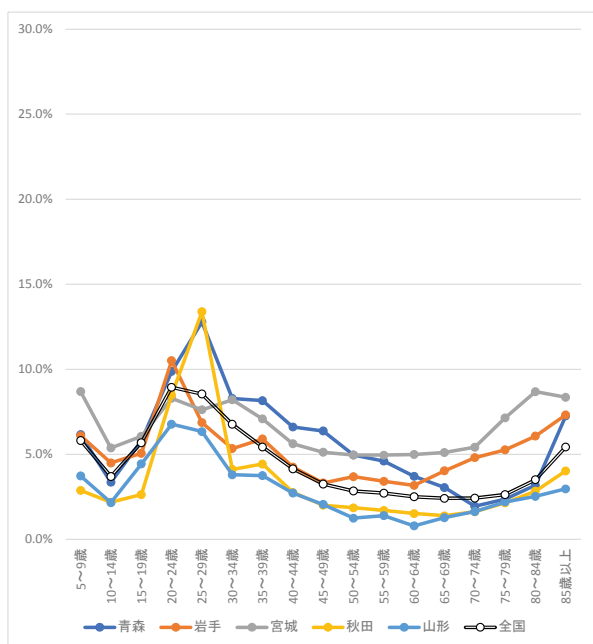


図5 東北地方各県のMAPE(コホート変化率法)

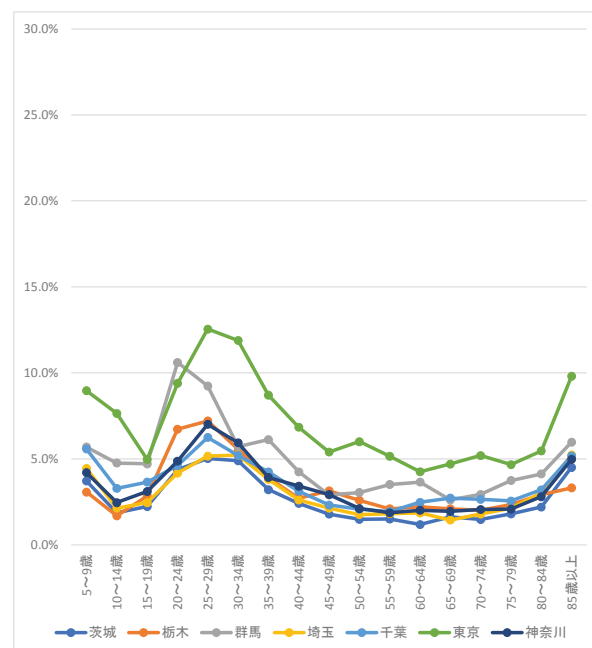


図7 関東地方各都県のMAPE(コホート変化率法)

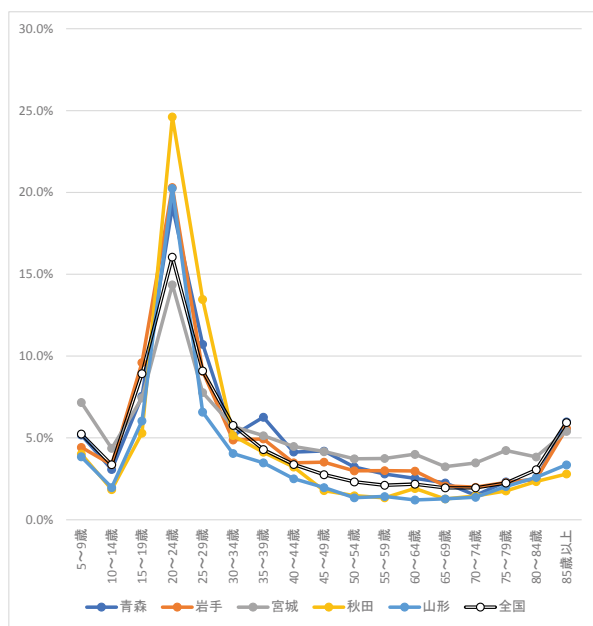


図6 東北地方各県のMAPE(コホート要因法)

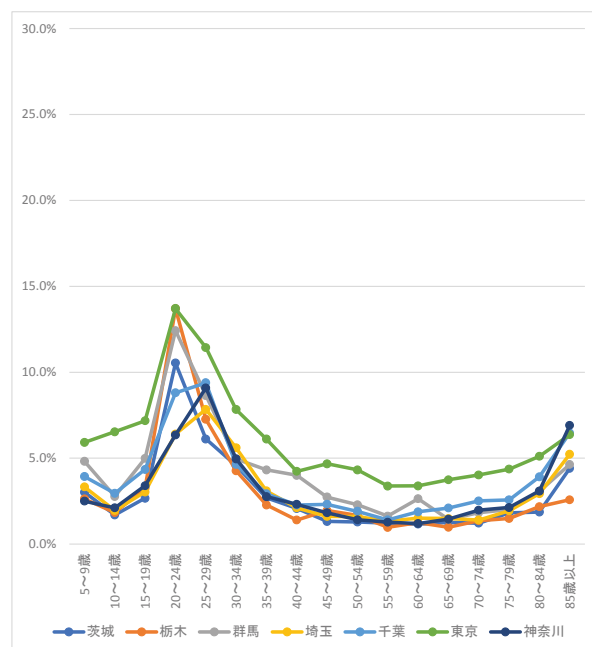


図8 関東地方各都県のMAPE(コホート要因法)

誤差について、市区町村の人口規模の影響があるので都道府県別の市区町村平均人口とMAPEの平均値の関係を図9に見る。すると、概ね人口に応じて減衰する3節で見たような傾向が現れているが、東京だけは特別であるとわかる。また、震災の被害が大きかった岩手県・宮城県はそれほど特異な状況に変化した(誤差が出た)というわけでもないとわかった。

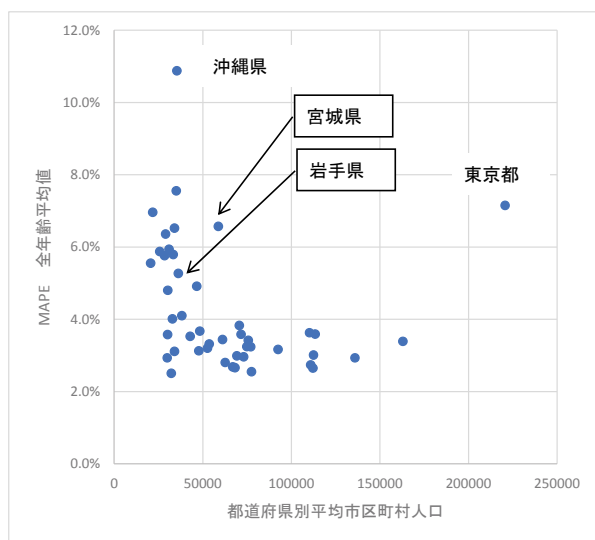


図9 都道府県別市区町村平均人口と全年齢平均MAPE

MALPEについての東北地方5県の予測結果は図10・図11に、関東地方各都県の予測結果は図12・図13に示される。コーホート変化率法・コーホート要因法いずれで見ても宮城県・岩手県の誤差が大きくでたといえず、むしろ青森県が目立つ結果となった。

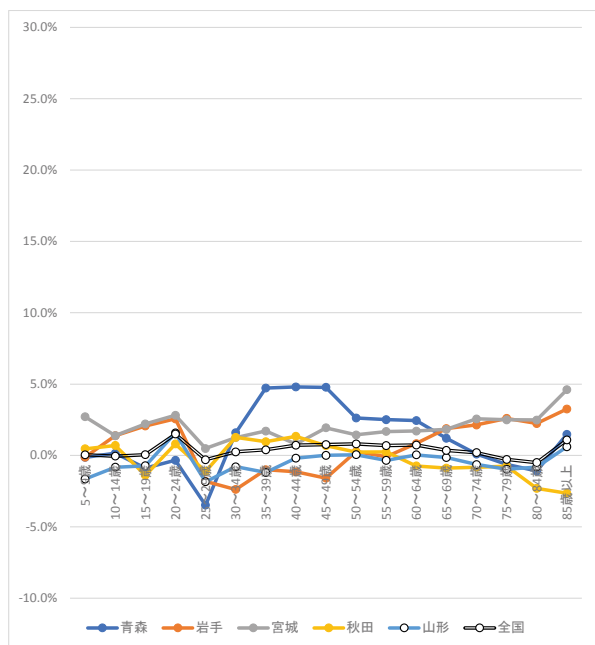


図10 東北地方各県のMALPE(コーホート変化率法)

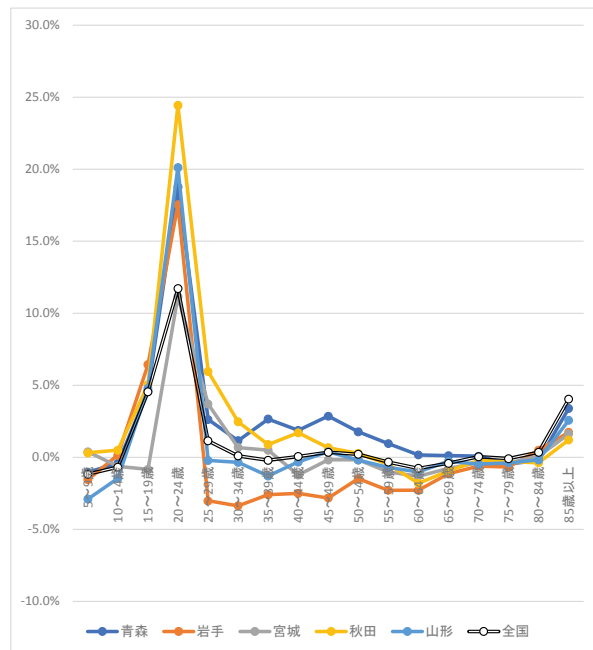


図11 東北地方各県のMALPE(コーホート要因法)

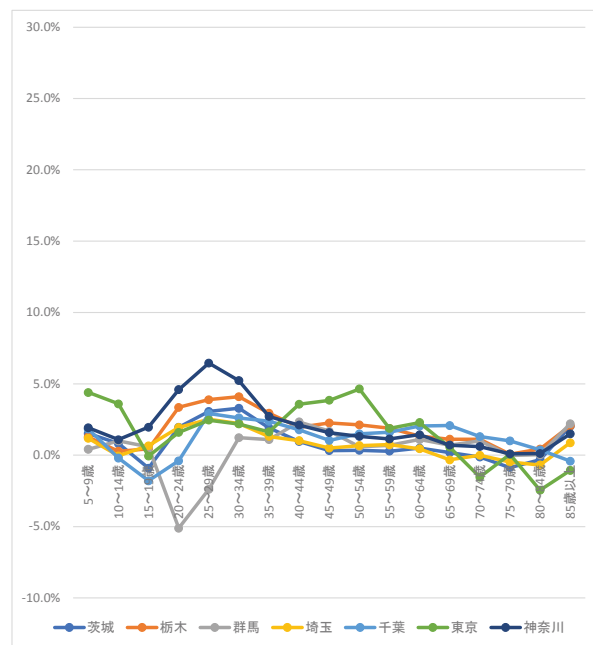


図12 関東地方各都県のMALPE(コーホート変化率法)

図11のコーホート要因法の結果では20-24歳が大幅な過剰予測となっていることがわかり、この世代が予測よりずっと多く流出したことがわかる。その分であろうか、図13では東京都・神奈川県で過小予測となっている。むしろ図12などで見ると関東地方の誤差が大きく出た。また、コーホート要因法の誤差は30歳～84歳まで県にかかわらず小さく、この範囲での予測説明力は高い。

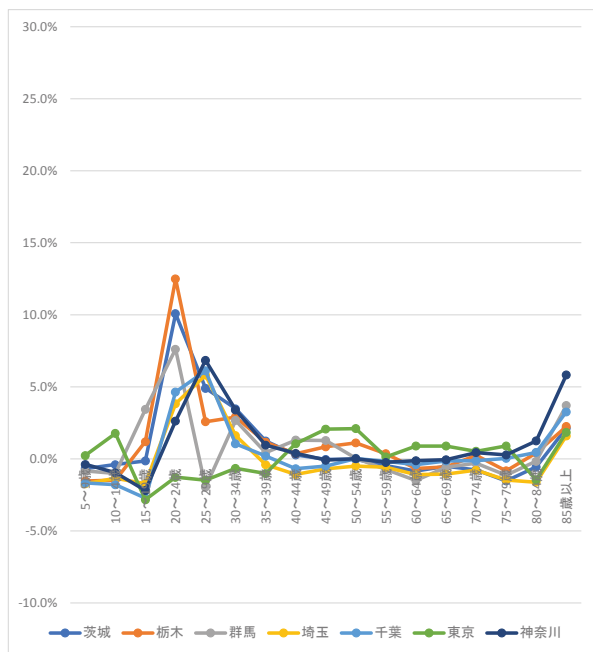


図13 関東地方各都県のMALPE(コホート要因法)

5. おわりに

以上で人口推計の結果に関する分析結果の説明を終える。一般にはコホート要因法は信頼性のある手法とされている。しかし、本研究はコホート変化率法の精度が結構よいという結果となった。これは5年間という比較的短い期間の予測結果であるためかもしれない。例えば石川(1993),p.53では、長期間の分析ではコホート要因法の精度が優位であることを述べている。

また、3.2節で示したように人口が少ない市町村をたくさんかえる県はMAPE等が大きい値となった可能性があることにも注意する必要がある。

一方、東京都のこの5年間の人口増加は予測結果との比較によってかなり特殊とわかった。これに関しては川崎他(2018)でも予測誤差の説明で東京都の湾岸の開発の影響が1%有意で影響力を持っていることを指摘していることに対応するであろう。

市区町村別の人口予測結果で手に入りやすい形で公開されているのは近年の結果に限られる。また平成の大合併もあったため、現状では長期の評価が難しい。2020年の国勢調査をはじめとするこれからの結果をふまえて長期的な予測結果の評価研究をしていきたい。

参考文献

- 1) J. Tayman, E. Schafer & L. Carter (1988), The Role of Population Size in the Determination and Prediction of Population Forecast Errors: An Evaluation Using Confidence Intervals for Subcounty Areas, Population Research and Policy Review, Volume 17, pages1-20.
- 2) S. K. Smith and M. Shahidullah (1995), An Evaluation of Population Projection Errors for Census Tracts, Journal of the American Statistical Association, Vol. 90, No. 429, pp. 64-71.
- 3) S. Rayer (2008), Population Forecast Errors, Journal of Planning Education and Research 27, pp.417-430.
- 4) T. Wilson, F. Rowe (2011), The Forecast Accuracy of Local Government Area Population Projections: A Case Study of Queensland, Australasian Journal of Regional Studies, Vol. 17, No. 2, pp. 204-243.
- 5) 相尚寿 (2017), 小地域単位での住環境得点による人口増加の再現性検証と改良の試み, 都市計画論文集, 52巻 3号, pp.1290-1297.
- 6) 相尚寿 (2018), 最新の人口動態を用いた小地域単位での住環境得点による人口増加の再現性検証—住環境得点と人口動態の地図化とクロス集計—, 都市計画論文集, 53巻 3号, pp. 625-631.
- 7) 阿藤誠・池ノ上正子 (1987), 将来人口推計の推計精度について, 人口問題研究, Vol.182, pp. 56-61.
- 8) 石川晃(1993)市町村人口推計マニュアル.古今書院
- 9) 大澤 義明, 小野田 竜巳, 小林 隆史(2008)コ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.Vol. 73, No. 634 pp. 2605-2612.
- 10) 川崎 薫・大橋 瑞生・谷口 守(2018) 人口減少予測にあらうための基礎分析 -地方自治体の推計人口と実人口の乖離に着目して-.都市計画論文集 / 53 巻3 号, pp. 1080-1086.
- 11) 古藤 浩・三浦英俊 (2012):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,vol.20, No.1, pp.71-80.
- 12) H. KOTO (2019) Estimating Errors According to Population Size in Multi-regional Population Projection. 2019 Asian Conference of Management Science & Applications. Penglai, China
- 13) 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, vol.71-3, pp. 216-240
- 14) 山田 茂(2016) 2015年国勢調査結果の精度について : 抽出速報集計を利用した暫定的考察.国土館大學政経論叢,28(4), pp.1-35
- 15) 山田 茂(2011) 抽出速報集計からみた2010年国勢調査結果の精度について.国土館大學政経論叢,23(4), pp.71-95.
- 16) 吉田・堤(2017) 人口構成比の観点からみた将来日本の縮図 : 組成データ解析の適用.GIS-理論と応用 25(2), pp.79-89, 2017-12
- 17) 国立社会保障・人口問題研究所(2011) 将来人口推計の方法と検証—平成18年推計の仕組みと評価—.第12回社会保障審

議会人口部会資料.

https://www.mhlw.go.jp/stf/shingi/shingi-hosho_126704.html

- 18) 国立社会保障・人口問題研究所(2013)将来推計人口・世帯数データアーカイブス.

<http://www.ipss.go.jp/syoushika/tohkei/Mainmenu.html>