

200万人都市を対象とした 初期世帯マイクロデータ生成の試行

阪田 知彦¹・鈴木 温²・杉木 直³・正木 俊行⁴・田 寛之⁵

¹正会員 国立研究開発法人建築研究所住宅・都市研究グループ（〒305-0802 茨城県つくば市立原1）

E-mail: sakata@kenken.go.jp (Corresponding Author)

²正会員 名城大学理工学部社会基盤デザイン工学科（〒468-8502 名古屋市天白区塩釜口1丁目501番地）

E-mail: atsuzuki@meijo-u.ac.jp

³正会員 豊橋技術科学大学建築・都市システム学系（〒441-8580 豊橋市天伯町雲雀ヶ丘1-1）

E-mail: sugiki@ace.tut.ac.jp

⁴非会員 アカデミックエクスプレス株式会社（〒305-0047 茨城県つくば市千現2-1-6）

E-mail: masaki@academic-express.com

⁵非会員 アカデミックエクスプレス株式会社（〒305-0047 茨城県つくば市千現2-1-6）

E-mail: den@academic-express.com

本研究は、世帯マイクロシミュレーションに基づいた将来都市構造の予測技術の開発のうち、初期世帯マイクロデータ生成に関する実用化の一環として、前報での計算時間を大幅に短縮するためのアルゴリズムを元にした初期世帯マイクロデータの推計方法の実用性の拡張を目的として、大都市での試行に関するものである。まず、この初期世帯マイクロデータ生成の試行にあたり、一部のアルゴリズムや変数の取り扱いなどについて改良を行った。その改良プログラムを元に、北海道札幌市を対象として、オープンデータを用いたケーススタディを実施した。その結果、誤差関数値=0の初期世帯マイクロデータの生成ができることや、予想よりも計算時間を大幅に減少できることを確認した。

Key Words : *householdbased micro-simulation, population synthesis, open data, simulated annealing*

1. はじめに

本報では、初期世帯マイクロデータの推計方法の実用性の拡張のために、大都市での適用を主眼とした検討を行ったので、その概要と改良点、北海道札幌市を対象とした試行結果について報告する。

本格的な人口減少期に入った我が国の今後の都市構造を考えるには、多様な観点からの検討が必要であるが、その1つとして、将来の住宅立地や世帯分布の動向把握が不可欠であると考えている。それを支える技術は様々あるが、その1つに世帯マイクロシミュレーション（HUMS:Household based Urban Micro-Simulation model）に基づいた将来予測¹が提案されている。世帯マイクロシミュレーションを構成する大きな2つの要素として、①現況の個々人の個人属性と世帯構成を再現する初期世帯マイクロデータの生成と、②個々のライフイベント変化をモデル化した将来予測、がある。そのうち、本研究では、①の初期世帯マイクロデータ生成を対象としている。初期世帯マイクロデータの推計は、得られている統計データの集計値の組み合わせにより、個々のマイクロデータ

の属性と分布を推定しようとするものである。入手可能な統計データから初期世帯マイクロデータを推計する方法は、これまでに様々な方法が提案されている（詳しくは前報² 2章を参照されたい）。このうち、水流ら³は、国勢調査等、一般に公表されているオープンデータのみを用いて、鈴木ら¹の開発した世帯マイクロシミュレーションモデルの初期世帯マイクロデータを推計する方法論を提案している。水流ら³の方法は、オープンデータのみから、ある程度精度の高い初期世帯マイクロデータの推計が可能な方法であるが、誤差を減少させるために、長い計算時間を要していることが課題であった。今後の実用を考えた場合、より短時間で計算ができることは、マイクロシミュレーションに立脚した予測技術の活用場面を広げる上で、有益であろうと考えられる。

本研究は、水流ら³の方法をベースとしつつ、誤差修正の計算時間を大幅に短縮し、かつ既存研究よりもさらに精度の高い推計が可能な初期世帯マイクロデータの推計方法を実用化することを目的としている。前報²では、初期世帯マイクロデータを高速に生成するための計算アルゴリズムを検討・実装し（図-1）、その有効性を検証

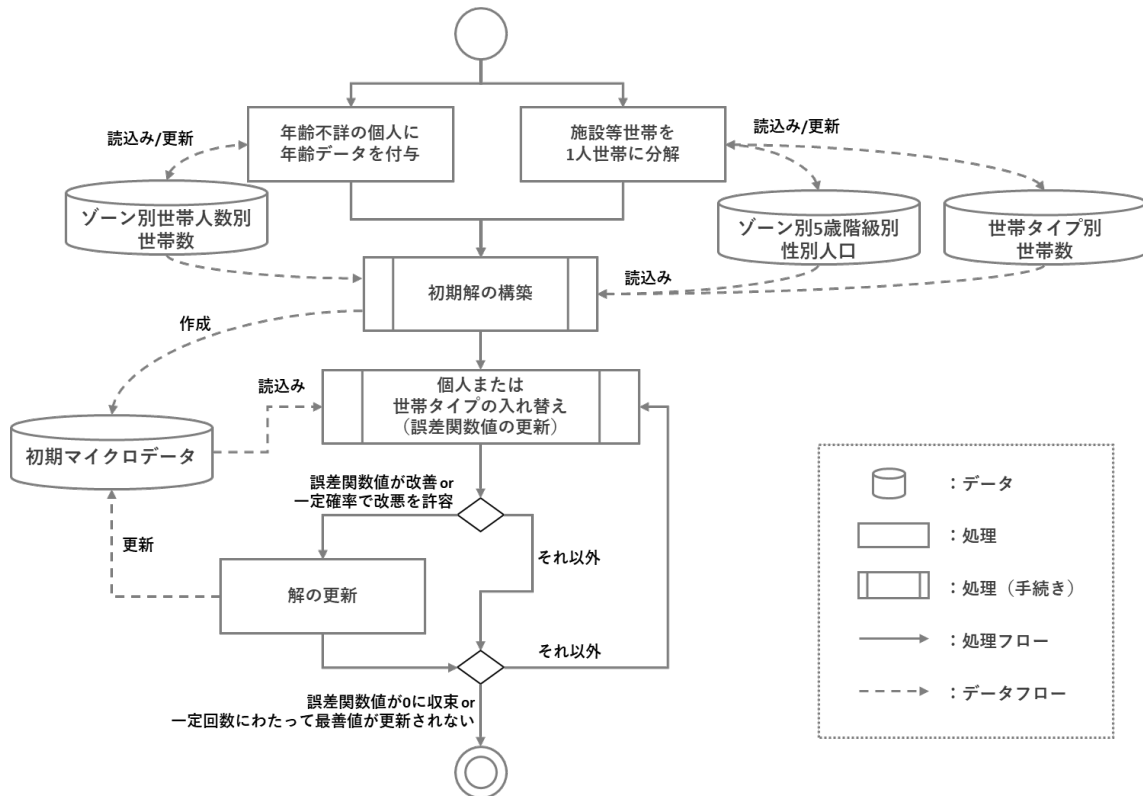


図-1 初期世帯マイクロデータ生成処理の全体フロー

するため、人口規模が約40万人の富山県富山市を対象としてオープンデータを用いたケーススタディを実施した。結果、平均17分程度で誤差関数値=0の初期世帯マイクロデータを生成できることを確認できた。これにより、高速化については、一定の目処があったと言える。一方で、都市の規模を大きくした場合の計算時間についての適用性や計算時間についての検証は今後の展開としてあった。

以上の目的・背景より、本報では、大都市での適用を主眼とした適用における改良点を述べ、北海道札幌市を対象とした試行結果について報告する。

本研究は、建築研究所研究課題「将来都市構造の予測・評価手法の高度化による目標管理・推進評価技術の開発（平成28～令和3年度）」の一環として実施した。

2. 大都市での適用に向けた改良

この章では、大都市での適用に際して行ったプログラムの改良について述べる。

一概に大都市といっても、その定義は様々であるが、ここでは人口規模が大きい自治体ということで捉えることにした。わが国で最も多くの人口を抱える自治体（表-1）は、神奈川県横浜市（約 370 万人）であるが、今回の改良前のプログラムの処理限界を 100 万人として設計していた関係で、まずはその処理限界の 2 倍程度の人口を有する都市の 1 つである北海道札幌市を対象として行うことにした。

表-1 政令指定都市の基本指標

道府県	市名	人口【人】	世帯数【世帯】	面積【平方km】	2010年～2015年の人口増減率【%】
北海道	札幌市	1,952,356	921,837	1,121	2.0
宮城県	仙台市	1,082,159	498,953	786	3.5
埼玉県	さいたま市	1,263,979	533,209	217	3.4
千葉県	千葉市	971,882	417,857	272	1.1
	横浜市	3,724,844	1,645,618	437	1.0
神奈川県	川崎市	1,475,213	691,837	143	3.5
	相模原市	720,780	311,188	329	0.5
新潟県	新潟市	810,157	321,511	726	-0.2
静岡県	静岡市	704,989	286,013	1,412	-1.6
	浜松市	797,980	309,227	1,558	-0.4
愛知県	名古屋市	2,295,638	1,058,497	326	1.4
京都府	京都市	1,475,183	705,874	828	0.1
大阪府	大阪市	2,691,185	1,354,793	225	1.0
	堺市	839,310	350,301	150	-0.3
兵庫県	神戸市	1,537,272	705,459	557	-0.4
岡山県	岡山市	719,474	309,409	790	1.4
広島県	広島市	1,194,034	531,605	907	1.7
福岡県	北九州市	961,286	426,325	492	-1.6
	福岡市	1,538,681	764,820	343	5.1
熊本県	熊本市	740,822	315,456	390	0.9

（全て 2015 年国勢調査より著者作成）

北海道札幌市は、2015 年段階の人口約 195 万人であり、富山県富山市に対し、人口規模が約 5 倍となっているだけでなく、

- ・ ゾーン数が約 9 倍となり、対象都市の規模そのものが著しく増大している
- ・ ゾーンの粒度が相対的に小さくなっている
- ・ ゾーン内人口のほとんどを施設等世帯が占めるといった極端な特性をもつゾーンが存在する

といった違いがある。

また、基準とする 2015 年段階で、人口が単調減少傾

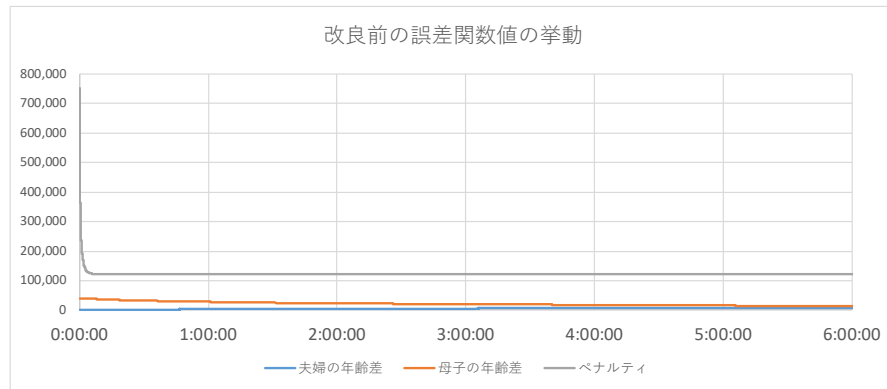


図-2 改良前のプログラムによる札幌市での誤差関数値の挙動

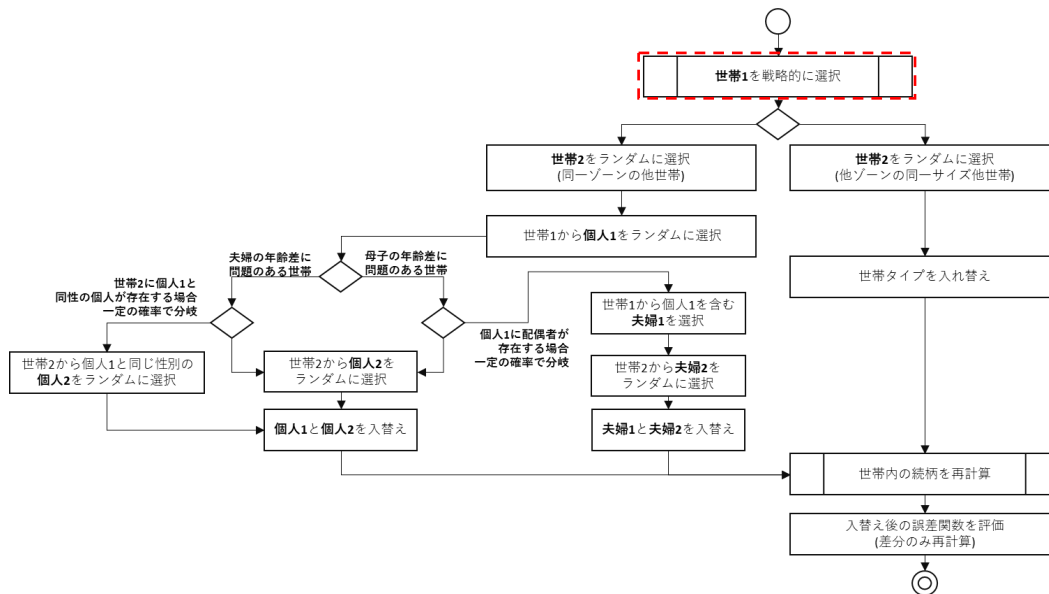


図-3 SA法の処理フロー (1反復分)

向に入った富山県富山市と異なり、まだ減少傾向局面に入っていない北海道札幌市を対象とすることで、人口推移傾向の異なる都市に対する挙動を確認することも可能になる。

まず、図-1に基づいて作成したプログラム（以下、改良前のプログラムとする）をそのまま北海道札幌市に適用したが、次のような問題が明らかとなった。

- I. SA法開始直後は誤差関数のペナルティ項が急速に下降するものの、約30分が経過した時点で完全に下げ止まり、開始から7時間が経過しても一切改善が見られない（図-2）
- II. この状態で探索を打ち切ると、初期世帯マイクロデータ内に続柄不明の個人、未成年のみの世帯などが大量に生成されてしまい、以降の将来予測シミュレーションを適切に実施できなくなる危惧が生じた

これらに対する検討と試行錯誤の結果、次の2点の改良を加えることで、北海道札幌市に対しても誤差関数値

= 0の初期世帯マイクロデータを生成できることが判明した。ただし、ここでの改良は、図-1に示す初期世帯マイクロデータの生成処理フロー自体を変更するものではない。

(1) 入れ替え対象となる世帯の選択基準の変更

まず、I.については、下記のような変更を加えることにした。

変更前：続柄不明等のエラーを含む世帯が存在する場合、必ずその中から世帯1を選択する。
 変更後：続柄不明等のエラーを含む世帯が存在する場合、一定の確率(パラメータ)でその中から世帯1を選択する。

そもそもSA法は一時的に解の改悪を許容することで局所解から脱し、より良い解を探索するアルゴリズムである。変更前のアルゴリズム（図-3）では、エラー世帯が発生した場合、エラーが解消されるまで必ずそのエラー世帯が交換対象になり続けるため（図-3中の点線囲

みの箇所) , 結果的にエラー世帯の存在が許容されないことになる。この結果, SA 法の利点を損ない, 局所解から脱出できず膠着状態に陥ってしまう。

変更後のアルゴリズムでは, 初期世帯マイクロデータ内にエラー世帯が存在しても, 一定の確率でしかその世帯が存在しないようになる。これにより, エラー世帯の存在が一定期間に渡って許容されるため, 反復の中で局所解を脱出できる可能性が生じる。また, この変更であれば, アルゴリズム自体の大きがかりな変更は回避できる。

このことをふまえ, 試行錯誤の結果, エラー世帯の選択確率を 50%とすることで, 北海道札幌市でも誤差関数値=0 の初期世帯マイクロデータが得られた。

またこの改良でエラー世帯の選択確率を減少させたことで誤差関数値の収束が遅れるもしくは 0 にならない懸念があったことから, 確認のために富山県富山市でも計算を実施した。結果, 誤差関数値=0 の初期世帯マイクロデータが得られ, その SA 法の収束に要する時間は, 改良前が 17 分であったのに対して, 改良後が 16 分と, 遜色ない結果であった。

(2) 施設等世帯の扱いの変更

続いて II. のような状況が発生した理由として, 改良前のプログラムでは, 施設等世帯を分解したことによって生じた見かけ上単身世帯のように振る舞う施設等世帯の取り扱いを明確化していなかったという可能性が考えられた。

そこで, II. に対しては, 下記のような改良を加えることにした。

変更前: 施設等世帯に属する個人を割り当てるための単身世帯を当該ゾーンに追加する。
変更後: 施設等世帯に属する個人を割り当てるための未成年のみで構成されることを許容する特殊な単身世帯を当該ゾーンに追加する。

前述の通り, 北海道札幌市では施設等世帯がゾーンの大多数を占めるゾーンが存在する。具体的には, 人口の 50%以上が施設等世帯に所属するゾーンが 19 ゾーンある。一例としては, 「北海道札幌市清田区真栄」では, 一般世帯人数 121 人に対して施設等世帯人数が 1360 人。計 1481 人のうち未成年 (5 最階級で 14 歳以下) が 154 人おり, 施設等世帯を全て通常の単身世帯に変換してしまうと, 成人の絶対数が不足してしまうことが判明した。このことは富山県富山市では見られない状況であった。

具体的な対応策として, 施設等世帯を分解したことによって生じた単身世帯, およびそこに属する個人に対して, 以下のように扱うこととした。

- 初期世帯マイクロデータにおいて世帯タイプを「0」とすることで, 単身世帯と区別する。
- 未成年のみで単身世帯を構成することを許容し, 誤

差関数値のペナルティ項の集計から除外する。

- 個人に付与するオプション属性は, 就業状態は「無職」運転免許は「未保有」, 住宅タイプは「未定(空白)」とする。

- ただし, 性別年齢別人口等, 個人を対象とする人口集計には, 施設等世帯に属する個人を計上する。

ただし, 初期世帯マイクロデータ生成処理においては, 一般の単身世帯と施設等世帯由来の単身世帯の違いは「未成年を割り当て可能か否か」のみである。仮に, 同一ゾーン内で「一般単身世帯の個人」と「施設等世帯の個人(成人)」をどれだけ入れ替えても, 誤差関数は一切変化しない。これは, SA 法から出力された単身世帯と施設等世帯の年齢分布が一切保証されないことを意味する。例えば, 養護施設等の存在を考慮すると, 施設等世帯においては, 一般の単身世帯と比較して未成年と高齢者の割合が高くなることが予想される。しかし, 初期世帯マイクロデータに何の補正も加えなかった場合, 特定の年齢で施設等世帯の割合が不自然に高くなる等の事象が生じる。

こうした問題に対して, 同一ゾーン内で「一般単身世帯の個人」と「施設等世帯の個人(成人)」をどれだけ入れ替えても誤差関数は一切変化しないことを逆にとり, SA 法により出力された初期世帯マイクロデータに対して一般単身世帯の成人と施設等世帯の成人を同一ゾーン内で入れ替えることで, 単身世帯と施設等世帯の年齢分布を補正することにした。実装した補正アルゴリズムは次の通りである。

- ① ゾーン内の全ての「単身世帯の個人」と「施設等世帯の成人」から世帯情報を削除し, 「世帯未割り当て個人」とする
- ② 「施設等入居者の 5 歳階級分布」のうち, 「世帯未割り当て個人」に存在しない 5 歳階級の入居確率を 0 とした上で, 残りの 5 歳階級の確率の総和が 1 となるように再調整する。
- ③ 「施設等入居者の 5 歳階級分布」を用い, モンテカルロ法により, 施設等世帯に割り当てる個人の年齢を決定する。
- ④ で選択された年齢の個人を「世帯未割り当て個人」の中からランダムに選択し, 施設等世帯に割り当てる。
- ⑤ 施設等世帯の定員に達するまで②~④を繰り返す。
- ⑥ 残った「世帯未割り当て個人」を単身世帯に割り当てる。
- ⑦ 全ゾーンについて①~⑥を繰り返す。

なお, 今回の改良において一般世帯と施設等世帯間の入れ替えを明確化したことで, 今後世帯タイプの変更に際や新たな世帯タイプを設定する際にも応用が可能な仕組みを実装できた。

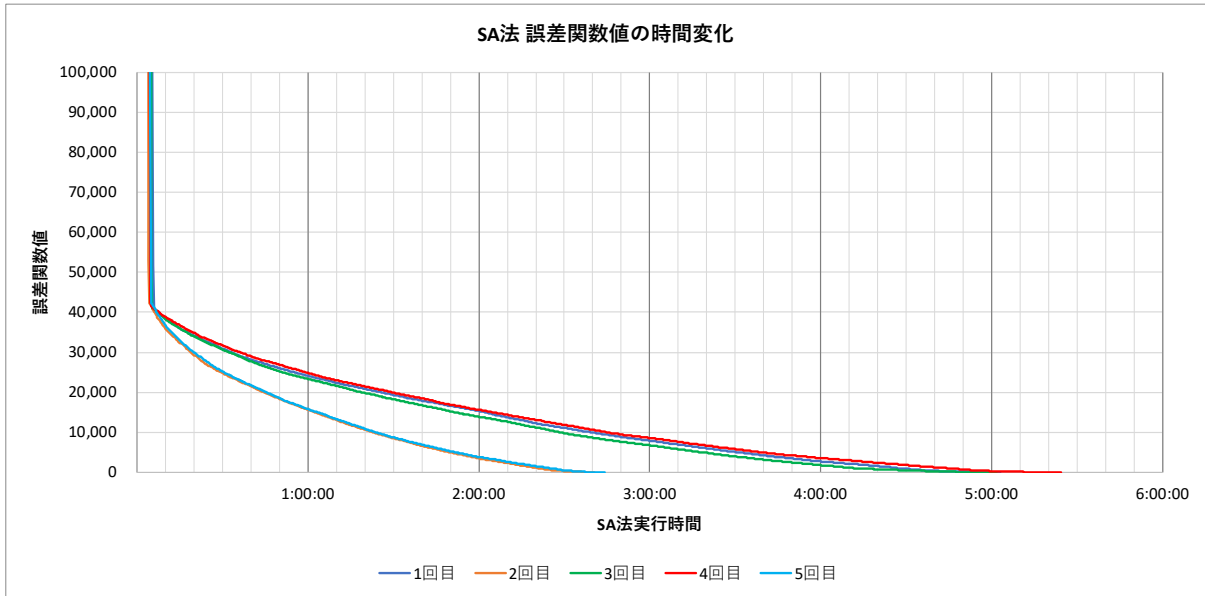


図-4 札幌市での誤差関数値の変化

3. 北海道札幌市での計算実行

以上の改良を加えた初期世帯マイクロデータ生成プログラムを、北海道札幌市のデータ（2015年国勢調査、人口約 195 万人、世帯数約 92 万）に対して表-2 の環境下で適用した。

結果としては、3時間～6時間程度で誤差関数値=0の初期世帯マイクロデータの生成が完了した（図-4 及び表-3）。今回実行した5回の試行の中には、3時間を切るケースと5時間を超えるケースとに二極化しており、富山県富山市と比較してばらつきが大きい結果となった。これは、解空間が膨大で探索に時間を要することから、ある程度ランダムに生成される初期解の優劣が時間経過とともに拡大された可能性が考えられる一方で、誤差関

数値の初期値が大きい場合でも結果的に SA 法にかかる計算時間が短くなっている場合もあることなどから、改良により SA 法による誤差収束の効率化が図られた効果の方が大きいとも考えられる。

ここで計算量の増加を理論的に考えると、初期解生成は単純にデータサイズに比例するが、SA法の探索空間（=入れ替え可能な個人の組み合わせ）は入力データサイズが N 倍になると N^2 倍に拡大する。北海道札幌市の人口は富山県富山市の約5倍であることから、理論値による単純増加量は $5^2=25$ 倍であるが、計算処理時間を見ると、富山県富山市20分に対して北海道札幌市6時間と18倍程度であり、理論値による単純増加量を3割程度下回る結果となった。

4. まとめ

初期世帯マイクロシミュレーションを用いた将来予測手法の検討の一環として、大都市での適用のための改良と北海道札幌市を対象とした試算結果について述べた。今回は、前報³⁾のアルゴリズムの適用上のいくつかの変更点を示し、それに基づいた北海道札幌市を対象とした試行を行った結果、3時間～6時間と予想していた計算時間よりも短時間で誤差関数値=0の初期世帯マイクロデータを得ることに成功した。

これは単なる計算時間の短縮といった効果だけではなく、より広域的な地域での初期世帯マイクロシミュレーションの実行が可能であるといった適用可能性が広がったことを意味している。

今後の課題・展開としては、都市規模や都市の様相の異なるケーススタディのさらなる蓄積によるパラメータ

表-2 開発・計算環境

実装言語	Python 3.5.2
使用パッケージ (非標準のみ記載)	numpy 1.13.1 panads 0.20.3
実動作環境	Windows Subsystem for Linux (Windows10標準搭載のLinux仮想環境)
CPU	Intel® Core™ i7-7500U (ノートPC向け超低消費電力)
メモリ	8GB

表-3 札幌市での計算時間と誤差関数値

試行	処理時間			誤差関数値	
	初期解	SA法	計	初期値	収束値
1回目	0:15:19	5:05:33	5:20:52	6,282,225	0
2回目	0:14:54	2:43:14	2:58:08	3,988,591	0
3回目	0:15:18	5:04:04	5:19:22	6,016,257	0
4回目	0:15:10	5:26:55	5:42:05	4,074,572	0
5回目	0:15:14	2:46:51	3:02:05	4,523,787	0

のチューニングや、生成結果の精度検証等があげられる。これらについては、機会を改めたい。

参考文献

- 1) 鈴木温, 杉木直, 宮本和明: 空間的マイクロシミュレーションを用いた都市内人口分布の将来予測—人口 40 万人規模の富山市を対象として—, 都市計画論文集, Vol.51 No.3, 2016
- 2) 阪田知彦, 鈴木温, 杉木直, 正木俊行, 田寛之: 世帯を単位とした将来推計における初期世帯マイクロデータ生成の高速化, 土木計画学研究・講演集, Vol.61, CD-ROM, 2020
- 3) 水流風馬, 平野巧真, 鈴木温: オープンデータを用いた初期世帯マイクロデータ生成方法に関する研究, 土木計画学研究・講演集, Vol.60, CD-ROM, 2019

(Received, 2020)

A TRIAL OF PRODUCING INITIAL MICRO DATA OF HOUSEHOLDS TARGETED AT 2 MILLION CITY

Tomohiko SAKATA, Atsushi SUZUKI, Nao SIGIKI, Toshiyuki MASAKI and Hiroyuki Den