

シミュレーテッドアニーリングによるエージェント属性決定手法を用いた人口推計モデルの評価

○福田純也 喜多一 (京都大学)

Evaluation of Population Estimation Model using a Method for Deciding Agent's Attributes with Simulated Annealing

*J. Fukuta and H. Kita (Kyoto University)

Abstract— In agent-based simulation using actual social statistics, attributes of all the agents have to be decided to meet the statistics. However, they can't be decided uniquely because statistics gives only limited number of aggregated values compared with the freedom of agents' attributes. To this problem, Ikeda et al. and Hara et al. propose a method of formulating the problem as a constrained optimization problem, and solving it using simulated annealing. They applied it to estimation of future population. In this paper, the authors discuss convergence of the optimization process and evaluate computational load of their method.

Key Words: Agent-Based Simulation, Simulated Annealing, Population Estimation

1 はじめに

社会をシミュレーションする手法にはさまざまなものがある¹⁾が、人や組織をソフトウェアエージェントとして記述し、社会構造をエージェント間の相互作用としてモデル化してボトムアップにシミュレーションを行うエージェントベースドシミュレーション (Agent-Based Simulation, 以下「ABS」) は他の手法では困難な政策評価などを含めてさまざまな応用が期待されている。

ABS はその解像度から、社会現象の原理的理解を目的に少数のパラメータでシンプルなモデルとして記述する Abstract Model, 対象をより詳細に記述する Middle Range Model, 特定の状況に限定された社会システムを可能な限り忠実に表現する Facsimile Model に分類される²⁾が、現実の問題を定量的に取扱い、具体的な政策の評価などを行うためには Facsimile Model の構築や評価のための手法の検討が重要である。

Facsimile Model 型の ABS を実行するためには、個々のエージェントの属性に関心ある事象に関して現実を一定程度再現できるように決定しなければならない。しかしながら、社会調査の個票に基づいてエージェントを設定できるような場合を除けば、多くの実社会のデータは統計として集計されたものであり、また個々の統計は個別に調査対象が設定されているため、統計相互に必ずしも整合している訳ではない。ABS の実施にあたっては、大規模なエージェントの属性をこのような統計への適合を考慮して決定しなければならない。

この問題に対し、池田ら³⁾、Hara ら⁴⁾ は人口推計モデルを例にエージェントの属性決定問題を拘束条件付最適化として定式化し、その解のランダムサンプリングをシミュレーテッドアニーリング法 (Simulated Annealing, 以下「SA」⁵⁾) により求める手法を提案している。

Hara らは上記手法を用いた人口推計モデルを提案しているが、SA 法により拘束条件充足問題がどの程度、良好に解かれているのか、計算量の多い同手法が実際のどの程度の計算時間を要し、どの程度の規模までの人口推計が現実的に扱えるのかは検討されていない。そこで本研究では Hara らのモデルについて、最適化計算の収束状況および計算負荷を評価する。

2 エージェント属性決定手法

ABS を用いた Facsimile Model 型の社会シミュレーションでは、対象となる社会現象をより忠実にモデル化するために、エージェントの属性を既存のデータに適合させる必要がある。統計調査の個票を用いる場合や、山本らの研究⁶⁾のように実態の把握しやすい小規模な地域を対象にする場合を除けば、適合させる対象は主に統計の集計量となる。しかしながら集計量からエージェントの属性を決定する場合、エージェントの属性の自由度のほうが高いため、これらを一意には決定できない。そこで池田ら、Hara らは人口推計モデルを例として、仮想社会の住民であるエージェントの属性決定問題を拘束条件付最適化問題として定式化し、SA を用いて解を求めるとともに^{3, 4)}、一意性のない属性を利用した ABS について複数回の試行を行うアンサンブルシミュレーションを提案している⁴⁾。

2.1 拘束条件付最適化

池田らの提案した拘束条件付最適化は、対象地域についての統計から個々のエージェントの属性を逆推定する手法である。各種の統計データをエージェントの属性を集計した値についての拘束条件とし、エージェントの属性決定問題を拘束条件をペナルティ関数とした最適化問題として定式化し、最適解をランダムにサンプリングして、拘束条件を満たす範囲で発生頻度の高い状態を属性値として採用しようというものである。

2.2 SA を用いたランダムサンプリング

SA は、高温にした金属を徐冷することで低エネルギー状態を作り出す「焼きなまし」を模した最適化手法である⁵⁾。SA は通常、最適化手法として捉えられているが、理論的に複数の最小解を等確率で見出すことが知られている。このことから、拘束条件をペナルティ関数として扱うことにより SA を拘束条件下でのランダムサンプリング手法として用いることができる。なお、データに適合する ABS のパラメータを推定する手法としては進化計算などを利用する逆シミュレーション⁷⁾が知られているが Hara らの研究では上記の特性に加え、大規模な問題でも取扱いの比較的容易な手法として SA が採用されている。

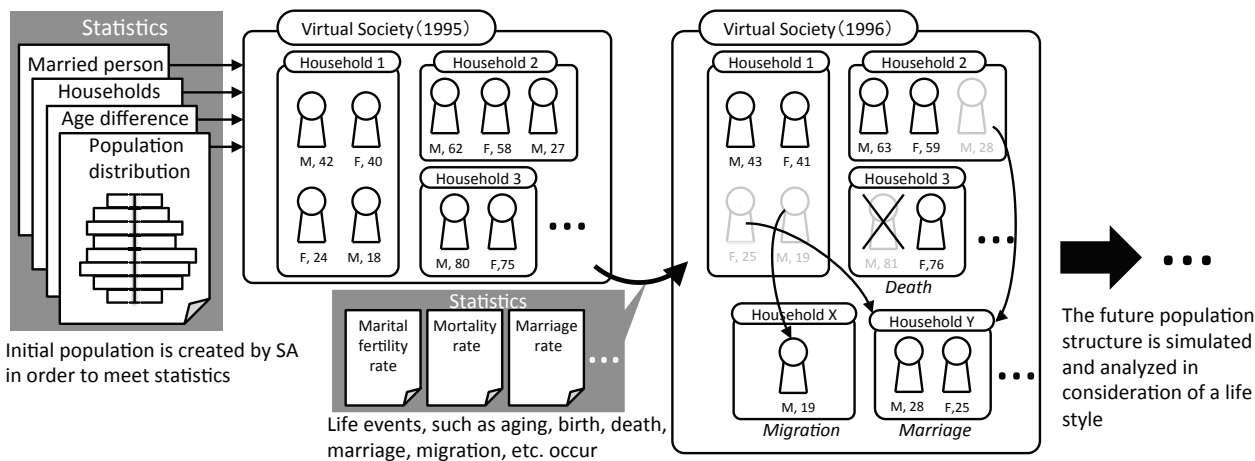


Fig. 1: Simulation model

2.3 アンサンブルシミュレーション

上記の手法を用いてエージェントの属性を決定しても、拘束条件と比べてエージェントの属性値の自由度は圧倒的に高く、属性値は一意に決まらない。そこで、「ランダムなサンプリングを行い、属性値を決定する」という工程を複数回繰り返し、エージェントが構成する仮想社会のアンサンブルを得る。そして、各仮想社会でシミュレーションを行い、これにより得られた将来の各仮想社会の分布から、将来予測値を考察する。この手法をアンサンブルシミュレーションと呼ぶ⁴⁾。

3 シミュレーションモデル

本研究では、Hara らによって構築されたモデル（以下、「原モデル」）をベースとしている。

3.1 モデルの概要

シミュレーションのサイクルは1年としており、シミュレーション開始時にはまず、初期エージェントが生成される。その後、サイクルごとに加齢、婚姻、離婚、出生、死亡、世帯分離・統合の順にライフイベントを所与のパラメータに適合するように確率的に発生させることで人口の推計を行う。Fig. 1 にモデルの概略図を示す。このうち、初期エージェントの生成に加えて、個人に対して外生的な確率で適用する出生と死亡を除く、婚姻、離婚、世帯分離・統合の各ライフイベントも拘束条件付最適化によりモデル化している。

3.2 初期エージェント生成

拘束条件付最適化を用いて統計データを充足するように個票を生成することで初期エージェントを生成する。

3.2.1 拘束条件

拘束条件として、以下の八つの統計データ（国立社会保障・人口問題研究所の人口統計資料集）を用いた。

1. 年齢別人口分布
2. ある年齢・性別の人が単身世帯にいる割合
3. ある年齢・性別の人が夫婦のみ世帯にいる割合
4. 父と子の年齢差の分布
5. 母と子の年齢差の分布
6. 夫婦の年齢差の分布
7. 女性の年齢階級別有配偶率
8. 男性の年齢階級別有配偶率（Table 1 参照）

Table 1: The rate of married male by age group (1995)

Condition X	Condition Y	Rate
Male, 15-19	Married	0.260
Male, 20-24	Married	6.483
Male, 25-29	Married	31.617
...
Male, 80-84	Married	74.843
Male, 85-	Married	56.433

例として、8番目の統計データを Table 1 に示す。同表の条件 X はある年齢階級に該当する男性であるかどうか、条件 Y はその年齢階級の男性エージェントが有配偶の状態にあるかどうかを意味している。上記の八つの統計データは、同様に「条件 X を満たす人（組）の数」に対する「条件 X, Y を満たす人（組）の数」の割合 (Rate) として表現できる。このことを利用して後述のように一般性をもった評価関数の設計が行える。

3.2.2 評価関数の設計

初期エージェント生成では、まず人口統計資料集で公表している「世帯人員別一般世帯数」および「家族類型別一般世帯数」を利用する。前者は世帯人員を 1, 2, ..., 7 人以上と分けた場合の各世帯数とその割合で、後者は家族類型（単身世帯、夫婦のみ、夫婦と子どもなど）別の各世帯数とその割合である。総世帯数を所与として、これらの統計に合うように、総エージェント数 (N とする) および世帯ごとの人数と役割（夫、妻、子どもなど）を伴った世帯構成を決める。

次に予め世帯の構成が割り付けられているエージェントの年齢、性別のセットを決定変数と考え、解空間を A として $\vec{a} = \{(age_i, gender_i)\}_{i=1, \dots, N}, \vec{a} \in A$ とする。そして、八つの統計データそれぞれの \vec{a} に対する誤差を表す関数を評価関数として $f_1^l(\vec{a}), f_2^l(\vec{a}), \dots, f_8^l(\vec{a})$ と定義し、それらを目的関数として定式化する。具体的には次のように導入している。

- 評価したいパラメータセット \vec{a} と、統計 l が与えられ、統計データ中の条件 X_l についての分類階級数を G_l とする¹⁾。
- 統計データ中の階級 $j = 1, \dots, G_l$ について、条件 X_l を X_{lj} 、条件 Y_l を Y_{lj} と表すものとし、統計値

¹⁾Table 1 のデータは、実際には 15 階級である

として示されている割合を r_{lj} とする。また、 \bar{a} について、「条件 X_{lj} を満たす人 (組) の数」を m_{lj} とし、「条件 X_{lj}, Y_{lj} を満たす人 (組) の数」を c_{lj} とする。統計上の割合が r_{lj} であるため、 c_{lj} の理想の数は $m_{lj}r_{lj}$ となる。

- これらにより、誤差関数を以下のように定義する。

$$f_l^I(\bar{a}) = \frac{4}{G_l} \sum_{j=1}^{G_l} (c_{lj} - m_{lj}r_{lj})^2 \quad (1)$$

この指標では、「実際の人 (組) の数 c_{lj} 」と「理想の人 (組) の数 $m_{lj}r_{lj}$ 」との差の二乗和を最小化しようとする。 c_{lj} が整数であり端数が生じることから各項目で 0.5 程度の差は問題なく、仮に全ての項目で 0.5 の差があった場合に $f_l^I(\bar{a}) = 1.0$ と正規化されるように係数 $4/G_l$ を掛ける形で (1) 式は定義されている。すなわち、この指標が 1 程度であれば従わせるべき統計データとほぼ正確に一致しているといえる。

次に、複数の評価指標を考慮するため、 $f_1^I(\bar{a}), f_2^I(\bar{a}), \dots, f_8^I(\bar{a})$ を線形和する。また、重視したい統計データに重みをつけるため係数 $\omega_1^I, \omega_2^I, \dots, \omega_8^I$ を用いて、次の最小化問題を SA で解く。

$$\min_{\bar{a} \in A} \sum_{l=1}^8 \omega_l^I f_l^I(\bar{a}) \quad (2)$$

各世帯を構成するエージェントの年齢、性別の決定後、後述の世帯分離・統合イベントを発生させるため、異なる世帯間で親世帯・子世帯の親子関係が設定される。ここでは、人口統計資料集中の「女性の年齢別出生数」および「女性の年齢別死亡率」を利用し、ある世帯主のエージェントにおいて、母親の年齢と生存を決定し、生存している場合に「当該年齢女性エージェントの選択」を確率的に行うことで、世帯間の親子関係が設定される。

3.3 婚姻イベント

婚姻イベントによって独身の男女 (男性 18 歳以上、女性 16 歳以上) が婚姻状態となり、新たな世帯を構成する。その際、拘束条件付最適化を用いて婚姻に関わる統計データを充足するような組み合わせを決定することで、エージェント同士を婚姻させる。

3.3.1 拘束条件

拘束条件として、厚生労働省が公表している人口動態統計特殊報告および人口動態調査から、以下の三つの統計データを用いた。なお、拘束条件の設定に際しては初婚と再婚の区別をせずに用いた。

1. 初婚夫妻の年齢差別婚姻率 (男性年齢 - 女性年齢)
2. 男性の年齢別婚姻率 (人口千対)
3. 女性の年齢別婚姻率 (人口千対) (Table 2 参照)

例として、3 番目の統計データを Table 2 に示す。条件 Y はその時点で婚姻の組み合わせとして選ばれているか、ということの意味している。

Table 2: Age-specific marriage rate of females (1995)

Condition X	Condition Y	Rate ²
Female, 16	Couples	1.0
...
Female, 25	Couples	96.3
Female, 26	Couples	95.0
...
Female, 74	Couples	0.1
Female, 75-	Couples	0.5

² Number of marriages per 1,000 females

3.3.2 評価関数の設計

婚姻件数 N_m は Table 2 の統計データと年齢別の女性エージェント数から次式のように計算できる。なお、当該統計データの項目数は 60 である。

$$N_m = \frac{1}{1000} \sum_{j=1}^{60} m_{3j} r_{3j} \quad (3)$$

決定変数は男女二人のエージェントの組である婚姻関係であり、解空間を M として $\vec{m} = \{(M_i, F_i)\}_{i=1, \dots, N_m} \in M$ と書くことにする。ここで (M_i, F_i) は婚姻関係にある男性エージェント M_i と女性エージェント F_i である。そして、三つの統計データそれぞれの \vec{m} に対する誤差を表す関数を評価関数として $f_1^M(\vec{m}), f_2^M(\vec{m}), f_3^M(\vec{m})$ とする。

- 評価したいパラメータセット \vec{m} と、統計データ l が与えられ、統計データ中の項目数を G_l とする。
- 統計データ中の項目 $j = 1, \dots, G_l$ について、条件 X_l を X_{lj} 、条件 Y_l を Y_{lj} 、割合を r_{lj} とする。また、 \vec{m} について、「条件 X_{lj} を満たす人 (組) の数」を m_{lj} とし、「条件 X_{lj}, Y_{lj} を満たす人 (組) の数」を c_{lj} とする。統計上の割合が r_{lj} であるため、理想の数は $m_{lj}r_{lj}$ となる。
- これらにより、誤差関数を次のように定義する。

$$f_l^M(\vec{m}) = \frac{4}{G_l} \sum_{j=1}^{G_l} (c_{lj} - m_{lj}r_{lj})^2 \quad (4)$$

そして各統計データに対する誤差関数を重み ω_l^M で荷重して目的関数を次のように定式化し、SA で解くことでエージェントの婚姻の組み合わせを決定する。

$$\min_{\vec{m} \in M} \sum_{l=1}^3 \omega_l^M f_l^M(\vec{m}) \quad (5)$$

3.4 離婚イベント

夫婦が存在する世帯に対して離婚が発生する。夫はそのままその世帯に残り、妻は新世帯を構成するものとする。また、どの世帯の夫婦を離婚させるか決めるだけでなく、子どもの親権をどちらに与えるかも決定変数に加えている。

3.4.1 拘束条件

拘束条件として、厚生労働省の人口動態統計特殊報告で公表されている以下の三つの統計データを用いる。

1. 夫の年齢階級別離婚率 (人口千対)
2. 妻の年齢階級別離婚率 (人口千対)
3. 夫婦の親権を行う者別離婚率

3.4.2 評価関数の設計

1年間の離婚件数 N_d は厚生労働省の人口動態統計で公表されている離婚率（人口千対）を用いて、 $N_d = \text{全エージェント数} \times \text{離婚率}/1000$ と計算される。決定変数は、 N_d 組の夫婦および親権の場合分けのセットであり、解空間を D と表すものとし、 $\vec{d} = \{(H_i, W_i, C_i)\}_{i=1, \dots, N_d} \in D$ と書く。 H_i は夫の役割を担うエージェント、 W_i は妻の役割を担うエージェント、 C_i は4種類の親権の場合分けである。

そして、婚姻イベントと同様に3種の統計データそれぞれの \vec{d} に対する誤差を表す関数 $f_1^D(\vec{d})$, $f_2^D(\vec{d})$, $f_3^D(\vec{d})$ と定義し、重み ω_i^D で荷重和したものを目的関数として最小化問題を解くことで離婚する夫婦を決定する。

3.5 出生イベント

夫婦が存在する世帯では、女性エージェントの属性値に応じた確率で出生し、0歳のエージェントをその世帯に追加する。ただし、一人の女性に何度も出生イベントが発生しないように、世帯に所属する子どもの数に応じて出生確率を変化させる。利用する統計データは、人口動態統計特殊報告の第1子～第3子の「年齢階級別有配偶出生率（有配偶女性人口千対）」である。

3.6 死亡イベント

エージェントの年齢・性別に応じた確率に従ってエージェントが死亡する。エージェントが従う死亡確率は、国勢調査による5歳階級別男女人口および厚生労働省の人口動態調査による5歳階級別男女死亡数の二つの統計データから算出した、5歳階級別男女死亡率である。

3.7 世帯分離・統合イベント

1年間で世帯構成の分布に大きな変化はないと仮定し、介護や独立によって世帯構成が変化することで先の婚姻、離婚で生じた世帯構成の偏りを調整するイベントである。世帯分離では独立などを理由に若い世代が新規世帯を生成するものとし、世帯統合では介護などを理由に子世帯が親世帯に加わるものとする。

3.7.1 拘束条件

以下の二つの世帯構成の分布に関して1年目は統計データへの一致を、2年目以降は前年との一致を拘束条件とした。

1. 家族類型別世帯数の割合
2. 世帯人員別世帯数の割合

3.7.2 評価関数の設計

決定変数は世帯であり、すべての世帯が評価対象である。仮想社会の全世帯数を N_h とし世帯を H_i ($i = 1, \dots, N_h$) とする。そして解空間を X と表すものとし、決定変数を $\vec{h} = \{H_i\}_{i=1, \dots, N_h} \in X$ と書くこととする。婚姻イベントなどと同様に、2つの統計データそれぞれについて \vec{h} の誤差を表す関数 $f_1^H(\vec{h})$, $f_2^H(\vec{h})$ を定義し、それらを重み ω_1^H , ω_2^H で荷重和して目的関数として定式化する。この最小化問題を解くことで、世帯分離・統合が発生する世帯群を決定する。

4 評価実験

本稿の実験では、人口推計モデルにおけるエージェント属性決定手法の評価を目的として、日本全体を対象に将来人口推計を行っている。推計期間は1995年から

2045年までの50年間としている。本実験では、1995年をシミュレーション開始年としているため、利用する統計データも基本的には1995年のものとしている。ただし、1995年のデータが存在しない統計に関しては、近い年のデータを利用した。

本実験の目的は、1) SAを用いた最適化プロセスの収束状況の検証(4.2節)、2) 実行する計算機上でのメモリの使用量と所用計算時間(4.3節)、3) 人口推計の結果とコホート要因法による推計の比較(4.4節)の3点である。

4.1 仮定値

50年間のシミュレーションを行うにあたり、シミュレーションサイクルを構成する各イベントのためのパラメータを設定する必要がある。本モデルでは設定すべきパラメータはシミュレーションにより変化する人口分布そのものにはあまり依存しないものが選ばれている。婚姻率、死亡率などについてはシミュレーション開始時の統計データから得られるパラメータが将来も一定であると仮定した。

4.2 最適化状況の分析

最適化問題として定式化しているのは、初期エージェント生成、婚姻イベント、離婚イベント、世帯分離・統合イベントの四つである。ここでは代表として、初期エージェント生成および1年目の婚姻イベントの最適化に関して、世帯数5,000の場合について分析する。なお、(2)式および(5)式で示した各統計データの重み ω_i^k , $k = I, M$ はすべて1を標準として実施した。

複数の統計データに関する評価指標である誤差関数値について、SAの探索につれてその値がどのように推移・収束するかを観察した。SAのアニーリングスケジュールとしては事前の予備実験での検討の上、初期エージェント生成では、初期温度は1000、探索回数は1900×全人口(≈1.33×10⁷回)で最終温度0.00171まで、婚姻イベントでは、初期温度は100、探索回数は500×男女の独身人口(≈2.08×10⁶回)で最終温度0.00298まで一定比率で温度を低下させるものとした。

初期エージェント生成と婚姻イベントでの結果を、それぞれFig. 2, Fig. 3に示す。どちらも縦軸を対数表記としている。SAの探索終了時点では双方のどの統計データの誤差関数値も1前後に収束しており、誤差関数の構成法より満足できる値に低下している。特定の誤差関数のみ急激な低下をしている様子もみられず、各評価関数が適切に設計されているといえる。すなわち、本実験で設定したSAの探索回数であれば、拘束条件付最適化を用いてエージェントの属性決定に関する適切な解が得られているものと考えられる。

評価指標である誤差関数値の収束値について、初期エージェント生成および婚姻イベントの結果をそれぞれTable 3, Table 4に示す。婚姻イベントでは、三つの統計データとも誤差関数値がいずれも1以下に収束しており、婚姻に関するエージェントの属性を統計データとほぼ一致させることができたといえる。他方、初期エージェント生成では誤差関数値が1以上となる統計データが三つ存在している。これは、拘束条件としている統計データが八つと多く、さらには人口構成や世帯構成、婚姻関係といった多様な統計データを採用

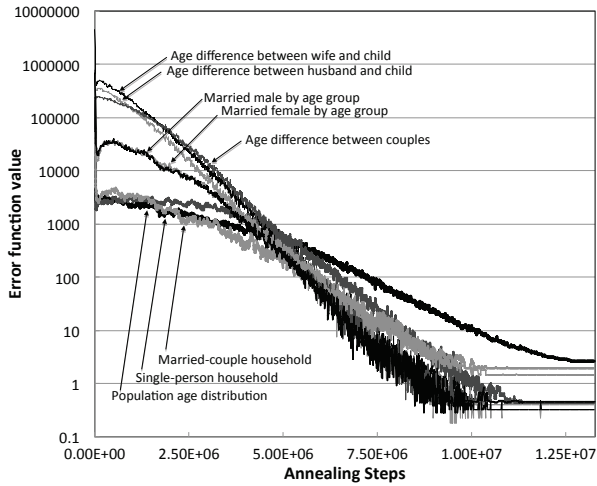


Fig. 2: Change of error function values in creation of initial population (5,000 households)

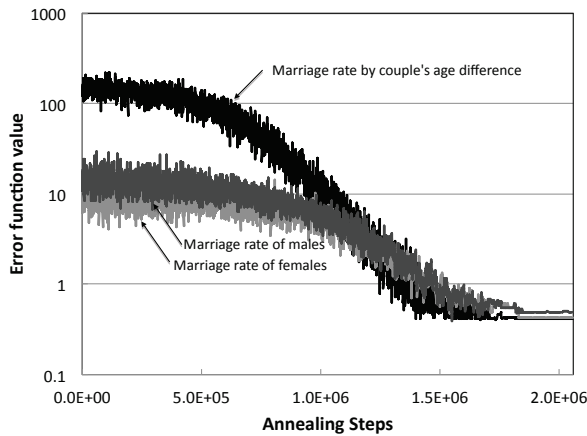


Fig. 3: Change of error function values in marriage event generation (5,000 households)

しているため、全ての統計データについてほぼ正確に一致させることが困難となっていると考えられる。

そこで、(2) 式で示した評価指標の重み ω_1^l に着目する。「年齢別人口分布 (Population age distribution)」の誤差関数値が 2.5871 で最大となっているが、この拘束条件は人口構成という重要な要素に関わることから、当該評価指標の重みを $\omega_1^l = 5.0$ とすることで他の統計データよりも重要視する。その結果、年齢別人口分布の誤差関数値は 0.9126 となり、1 以下まで低下した。他の統計データについての評価値は概ね増加したが、全体として誤差が均衡したといえる (Table 3 参照)。

4.3 計算量の評価

ランダムサンプリングの手法として用いている SA では、最適解を得るまでに膨大な計算量を必要とする。そこで本節では、空間計算量と時間計算量の二つの計算量から、SA を用いたエージェント属性決定手法が現実的に有用であるかの評価を行う。

なお、シミュレータは Java で実装されており、実行環境には Linux 上で稼働する JDK 1.6.0 を用いた。使用した計算機の仕様は以下の通りである。

- OS : Scientific Linux 6.2
- メモリー : 12GB (4GB × 3)
- CPU : Intel Core i7 3.07GHz (Quad Core)

Table 3: The convergence values of error function in creation of initial population (5,000 households)

	$\omega_1^l = 1.0$	$\omega_1^l = 5.0$
Population age distribution	2.5871	0.9126
Single-person household	0.4366	0.5269
Married-couple household	1.9366	1.6776
Age difference between husband and child	0.4047	0.7029
Age difference between wife and child	0.3227	0.8227
Age difference between couples	0.4485	0.4485
Married female by age group	1.4559	1.9444
Married male by age group	0.4518	0.6874

Table 4: The convergence values of error function in marriage event generation (5,000 households)

Error Function	Value
Marriage rate by couple's age difference	0.4125
Marriage rate of males	0.4223
Marriage rate of females	0.4896

Table 5: Computational complexity according to the model scale

Households	5,000	50,000	500,000
Agents	13,970	140,124	1,401,786
Memory usage[MB]	308	923	1,746
CPU time[min]	4.8	54.2	1,278.8

空間計算量の評価としてはマシンのメモリー消費量の最大瞬間値を、時間計算量としては CPU 時間を対象とした。仮想社会の初期世帯数を 5,000, 50,000, 500,000 とし、50 年間のシミュレーションを通して各世帯数における計算量を観察した。観察には、JMX (Java Management Extensions) 使用に準拠した監視ツールである JConsole を用いた。

結果を Table 5 に示す。500,000 世帯でのメモリー使用量は 1,746 MB となり、この値ならば通常のデスクトップ PC で十分利用可能であるが、これ以上の規模でのシミュレーションの実施には注意を要する。

CPU 時間を見ると、50,000 世帯でおよそ 54 分であったものが、500,000 世帯で 21 時間強となり、24 倍程度となっている。処理ごとの詳細な CPU 時間を分析したところ、最も支配的であるのが、初期エージェント生成後に行われる世帯間の親子関係の構築であり、50,000 世帯で 69 秒、500,000 世帯では 9,081 秒 (≈151 分) かかっている。この処理で世帯数 N に対して CPU 時間が $O(N^2)$ となる処理をしていることが原因であり、実装上のさらなる工夫が必要である。

4.4 将来人口の推計

本モデルは日本全国を対象とした人口推計モデルであるが、先に評価したシミュレーションに要する計算量の視点から世帯数を 50,000 としシミュレーションを実施した。2045 年 (50 年後) 時点の 5 歳階級別人口分布について、乱数系列が異なる 10 試行のアンサンブルの結果を男女別にそれぞれ Fig. 4, Fig. 5 に示す。比較の対象として、国立社会保障・人口問題研究所 (National Institute of Population and Social Security Research, 以下「NIPSSR」) が平成 24 年 1 月に実施した将来人口推計⁸⁾の結果 (出生中位, 死亡高位仮定)

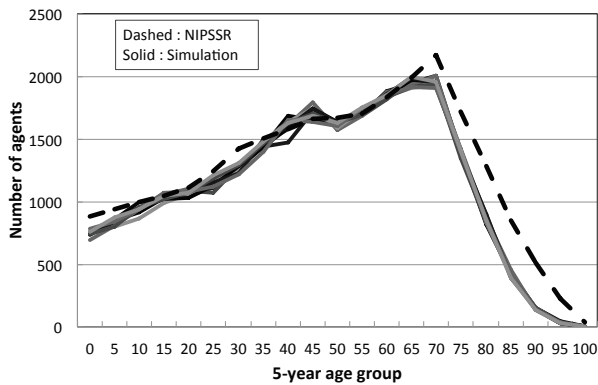


Fig. 4: Population distribution of males by 5-year age group (2045)

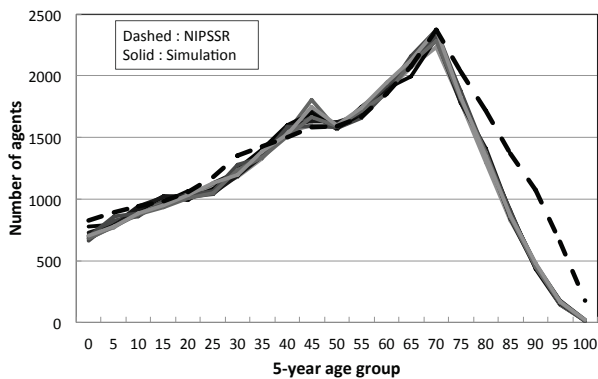


Fig. 5: Population distribution of females by 5-year age group (2045)

も示している。ここでいくつかの仮定の中から死亡高位仮定を選択しているのは、本推計の死亡に関する仮定値に極力近づけるためである。

アンサンブルシミュレーションで得られた各人口分布を見ると、50歳未満の人口に比較的大きなばらつきが生じている。この世代は、シミュレーション開始後に新たに出生した世代であるため、初期エージェントの構成の影響を受けやすくエージェントの属性値の自由度の高さによりばらつきが生じたと考えられる。50歳以上は死亡が主な人口変動の要因であるためシミュレーションごとのばらつきはかなり少ない。シミュレーション結果については属性決定がランダムに行われることから、このように複数の試行を行いシミュレーションのアンサンブルとして結果を把握する必要がある。

次に、NIPSSRの推計結果と比較すると、70歳以上の高齢人口があまり一致していない。この高齢人口に関しては、本モデルとNIPSSRの推計とでの将来死亡率の仮定に差があることが原因と考えられる。最も死亡しやすい仮定である死亡高位仮定の推計結果を示しているが、本推計で将来一定と仮定して用いられている死亡率と比較すると、2010年時点の年齢別のデータでは全体的にNIPSSRの死亡率の方が低く、より死亡しにくい仮定となっている。特に高齢者に関してはその差がより大きく、例えば本モデルでは70-74歳の年齢階級における男性の死亡率が31.4（人口千対）であるのに対し、NIPSSRでは72歳男性の死亡率は22.2（人口千対）となっている。さらに、NIPSSRでは近年平均寿命が延びていることなどの傾向を反映させるた

め、年数が進むにつれて死亡率を低下させている。このような死亡率の仮定値の違いから、高齢者の人口に差が生じていると考えられる。

5 おわりに

本稿では、SAによって解を求めるエージェント属性決定手法を用いたエージェントベースの人口推計モデルを評価した。最適化の収束状況および最適解の評価指標値の結果から、統計データに従った属性決定が行われており、人口推計モデルとして適切な設計となっていることが確認できた。また、計算量の計測結果からは、属性決定に計算量が多いSAを用いているが、500,000世帯程度の規模であれば現実的な計算量で推計が可能であることが示された。得られた将来人口分布の結果では、コホート要因法を用いた国立社会保障・人口問題研究所の推計との比較で妥当な推計結果が得られている。現在のモデルは日本全国を対象としているが、地域を明示的に考慮し転入・転出を加えた地域別モデルに発展させる予定である。また、エージェントベースのモデルである利点を活かし、人口問題に対する政策を評価で可能にしてシナリオ分析を進めたい。

なお、本研究の成果は現在、システム制御情報学会論文誌に投稿中である。

参考文献

- 1) N. Gilbert and K. G. Troitzch: Simulation for the Social Scientists, 2nd. Ed., Open University Press (2005)
- 2) 高橋真吾：社会システムの研究動向3 – 評価・分析手法(1) – モデルの解像度と妥当性評価；計測と制御, Vol. 52, No. 7, pp. 582–587 (2013)
- 3) 池田心, 喜多一, 薄田昌広：地域人口動態シミュレーションのためのエージェント推計手法；計測自動制御学会 システム工学部研究会 (2010)
- 4) S. Hara, H. Kita, K. Ikeda, M. Susukita: Configuring Agents' Attributes with Simulated Annealing, Tadahiko Murata, Takao Terano, Shingo Takahashi (eds.), *Agent-Based Approaches in Economic and Social Complex Systems VII*, Agent-Based Social Systems Volume 10, pp. 45–59 (2013)
- 5) S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi: Optimization by simulated annealing, *Science*, Vol. 220, pp. 671–680 (1983)
- 6) 山本一徳, 小山友介, 出口弘: Soarsを用いた村落の人口減少についてのデモグラフィックシミュレーション；計測自動制御学会システム工学部研究会資料 (2005)
- 7) 倉橋節也：社会システムの研究動向4 – 評価・分析手法(2) – モデル推定と逆シミュレーション手法；計測と制御, Vol. 52, No. 7, pp. 588–594 (2013)
- 8) 国立社会保障・人口問題研究所：日本の将来推計人口（平成24年1月推計）(2012)