

# 土地利用マイクロシミュレーションにおける初期マイクロ世帯データの推定手法\*

An Estimation Method of Household Micro-Data for the Base Year in Land-Use Micro Simulation \*

杉木直\*\*・宮本和明\*\*\*・Varameth VICHENSAN\*\*\*\*

By Nao SUGIKI\*\*・Kazuaki MIYAMOTO\*\*\*・Varameth VICHENSAN\*\*\*\*

## 1. はじめに

都市モデル分野においては、土地利用と交通の詳細な変化の記述するマイクロシミュレーションへの関心が高まっており、欧米諸国を中心として、複数の研究グループによつての研究事例および実際の都市への適用事例の蓄積が進められている<sup>1)</sup>。

居住立地モデルのような世帯を対象としたマイクロシミュレーションモデルの場合、各世帯には世帯収入、世帯人数、各世帯構成員の年齢、自動車保有、居住地、住宅タイプ等の多くの属性が定義されるが、シミュレーションを実行するためには、全ての世帯に対してこれらの属性を定義したシミュレーション初期年次データを用意する必要がある。しかし、住民基本台帳などから個人や個別世帯に関するデータを入手することは一般的に困難であり、またプライバシー保護の観点からも望ましくない。従つて、マイクロシミュレーションモデルでは、国勢調査などの入手可能な集計データと、個別世帯の属性情報を追加的に提供するサンプル調査を組み合わせ、人口データ推計を実施する必要がある。

人口データの作成手法としては、世帯タイプを設定した上で、IPF法によりタイプ別世帯数を推定する方法が多く用いられている。しかしIPF法では、モデルにおいて多数の世帯属性が取り扱われる場合、推定が困難になるという問題がある。また、これらとは異なる手法としては、モンテカルロサンプリング等を用いて個別世帯に対して複数の属性の組み合わせを設定したデータ（以降、マイクロデータ）を作成する手法が提案されている。

本研究は、マイクロデータを推定もしくは作成するための体系的な手法を構築することを目的としている。本稿では、まず初めに人口データ推計に関する既存研究のレビューを行い、既存データ推計手法の課題を整理する。そのうえで、サンプルマイクロデータによつて得られる  
\*キーワード：マイクロシミュレーション、マイクロデータ、初期データ推定、世帯属性

\*\*正員、修士(情報科学)、(株)ドーコン総合計画部  
(〒004-8585札幌市厚別区厚別中央1条5丁目4-1、

TEL011-801-1555、FAX011-801-1556)

\*\*\*フェロー、工博、東京都市大学環境情報学部

\*\*\*\*正員、博士(工学)、Kasetsart University

各属性間の相関性に関する情報、および既存統計による各ゾーンのコントロールトータル情報を利用して、シミュレーション初期時点のマイクロデータを推定するシステムを構築する。このシステムは、各世帯のマイクロな属性データを決定するための複数のアプローチによつて構成するものである。また、構築されたシステムを、道央都市圏パーソントリップ調査で取得したデータに対して適用し、その有効性を検討する。

## 2. 人口・世帯データ推計に関する既存研究

### (1) IPF法およびその拡張

Demingら<sup>2)</sup>によつて提案され、Beckmanら<sup>3)</sup>によつて人口データ推計問題<sup>4)</sup>への適用がなされたIPF法は、人口データ作成における一般的な手法である。同様の手法は、都市圏における世帯データ作成を対象として宮本ら<sup>5)</sup>によつて提案されており、周辺分布を制約とした同時確率最大化問題として記述されているが、解法の点でこれらは基本的にIPF法と同一である。

IPF法では、多次元のテーブルにおける各セルの世帯数が生成される。次元は、限定的な数の世帯属性によつて定義される。次元や属性間の相関性を考慮するためにサンプルデータセットが用いられる。低い次元への分配には国勢調査等によつて与えられる周辺分布に一致することが条件となる。IPF法によつて推定されたデータはマイクロシミュレーションにおいても十分活用可能であるが、個々の世帯別のマイクロデータではなく、タイプ別の世帯数である。

Guoら<sup>6)</sup>はゼロセル値問題、および世帯および個人レベルの属性の両方に関する統計的分布のコントロールの無能力さに関して緩和し、IPF法を改良した。

また、Pritchardら<sup>6)</sup>も、モンテカルロシミュレーションを用いて各エージェントにより多くの属性を考慮する機能を追加し、IPF法を改良した。これらは非常に有用な拡張であるが、IPF法を用いる限りセルベースのアプローチであることには変わらない。

### (2) モンテカルロサンプリング

Moeckelら<sup>7)</sup>はエージェントベースの手法を採用している。モンテカルロサンプリング手法が用いられており、図-1に示すようにマイクロシミュレーションモデルの

実行に必要な多くの属性を、世帯の各構成員ごとに考慮することが可能である。この場合、考慮可能な属性数は選択された属性間に有意義な関係が存在するかのみによって制限される。

本研究の推定手法は、これらのアプローチを発展させるものであり、より理論的な手法によって選択された属性間の関係や相関性を決定する。

### (3) 既存データ推計手法の課題

より複雑なマイクロデータを作成する場合、IPF法には限界がある。行列のセルに対する初期データの信頼性の観点において非常に高い基準を必要とする。IPF法では、0のセルを0.1もしくは0.01に設定するため、これらの設定は確率に影響を与えるが、理論的な根拠はない<sup>7)</sup>。Moeckelらによって指摘されているこのようなセル要素の問題に加え、改良されたIPF法を含むこれらのセルベースのアプローチには次のような問題がある。

- 属性およびそれらのカテゴリーを任意に設定することができない。
- 年齢や収入などは本来連続変数であるが、カテゴリーを離散的に設定する必要があり、これらの設定はゾーニングにおけるMAUP (Modifiable Area Unit Problem) と同様の問題を引き起こす。
- 属性数が増えるほどゼロセル問題が顕著になり、信頼性が低下する。

IPF法は、限られた数の属性を設定することによって、コンピュータの計算負荷量を削減することを目的として考案されたものであるという側面もあるが、属性カテゴリーの組み合わせが膨大になってしまう傾向があり、その一方で大部分のセルが0となってしまう。

## 3. 初期マイクロ世帯データ推定手法の構築

### (1) 前提条件の設定

コンピュータの計算容量を考慮すると、計算量やメモリへの負荷に関する問題が解決されることから、人口データ推定問題においてはエージェントベースのアプローチのほうが拡張性が高いと考えられる。そこで本研究では、Moeckelら<sup>7)</sup>によって構築されたエージェントベースの手法を拡張したアプローチにより、データ推計手法を構築する。

既存研究との相違を明示化するため、以下のように限定的な人口データ推定問題を設定する。

- 対象となるエージェントは世帯とする。
- 対象とする世帯タイプを、世帯主、妻および2人の子供によって構成される4人世帯に限定する。その際、子供の性別は考慮しない。
- 対象とする属性は各世帯構成員の年齢とする。
- 対象地域においては、5歳年齢階層別の人口データが国勢調査より入手可能であるものとする。

- すべての世帯人員の年齢情報を含む限定的な数の世帯サンプルが入手可能であるものとする。

これらの前提条件のもとで、対象地域におけるすべての世帯の世帯構成員の年齢を推定生成することが目的となる。以上の設定は非常に限定的なものであるが、世帯の相違を取り扱うという点で既存手法を拡張するものであると考えられる。

### (2) 推定手法

分析の基本的な考え方は次のとおりである。

- 属性（各世帯構成員の年齢）間の相関性を考慮する。
- 人口データ推計においては、周辺分布（5歳年齢階層別人口）に一致するように調整を行う。
- 推定、データ生成、調整はすべてモンテカルロ法等を用いて確率的に行う。

本推定手法の最も特徴的な点は、属性間の相関性の処理方法である。ここでは、連続変数である世帯構成員の年齢のみを属性として考慮する。

まずサンプル世帯データの属性変数  $(X = (x_1, x_2, x_3, x_4))$  を主成分分析を用いて無相関変数  $(P = (p_1, p_2, p_3, p_4))$  に変換する。

$$P = AX \tag{1}$$

世帯サンプルの無相関変数  $p_i$  の値に基づいて、図-1

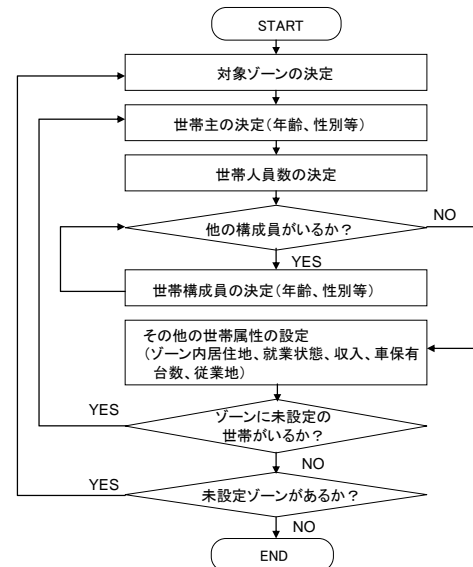


図-1 Moeckelら<sup>7)</sup>による人口・世帯データ生成手法

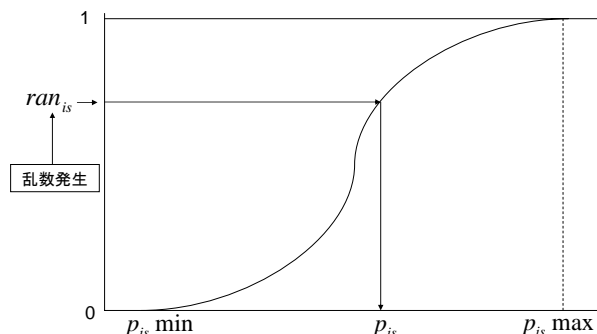


図-2 無相関変数を用いた相関の統合

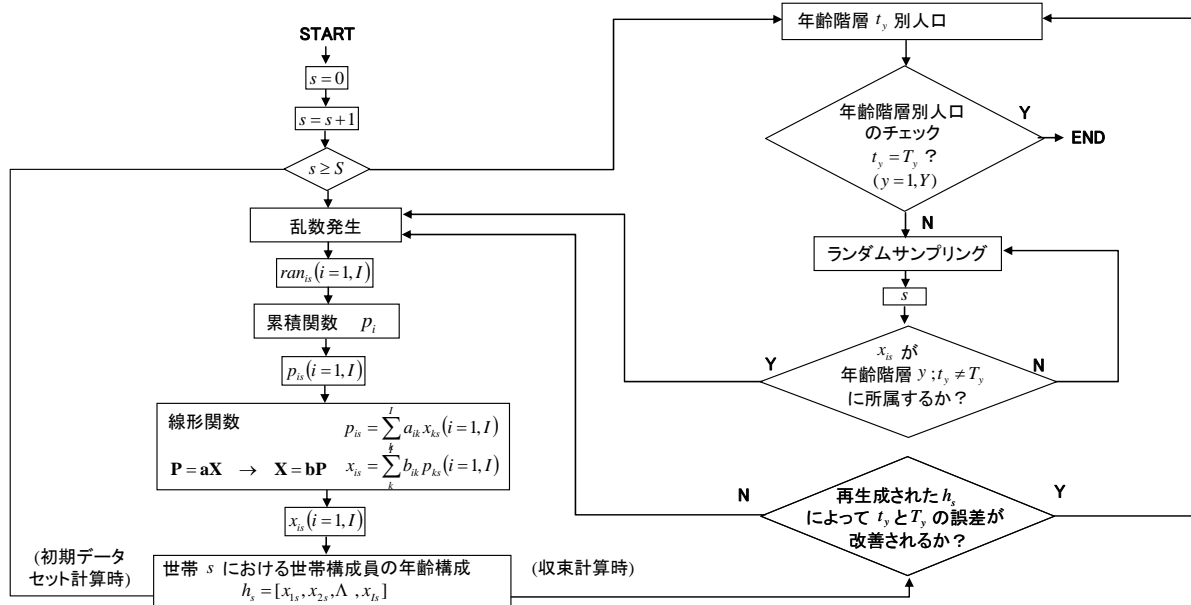


図-3 初期マイクロデータ推定フロー

に示すように、 $i=1,4$ について累積曲線を作成する。式(1)より次式が導出される。

$$X = A^{-1}P = BP \quad (2)$$

世帯を生成する際には属性*i*に対して乱数 $ran_i$ を発生させ、図-2の累積曲線より $i=1,4$ について生成世帯に対する

$p_i$ を求める。生成世帯の $x_i$ は式(2)より $i=1,4$ について

求められる。同様の処理を対象地域内のすべての世帯数に対して実行することで、世帯推定の初期データセットが作成される。この初期データセットは、周辺分布である5歳階級人口を満たさないため、モンテカルロ法を用いて世帯をランダム抽出し、世帯構成員の年齢が更新された新たなデータによって置き換える。このような調整処理を、周辺分布を満たすまで繰り返し実行する。本研究の初期マイクロデータ推定手法のフローを図-3に示す。

#### 4. 初期マイクロ世帯データ推定手法の適用

##### (1) 利用データ

本研究では、第4回道央都市圏パーソントリップ調査データを用いて、構築された初期マイクロデータ推定手法に関するケーススタディを行う。第4回道央都市圏パーソントリップ調査では、19,394世帯について世帯構成員の詳細情報を含むマイクロデータが取得されている。このうち、世帯主と妻および2人の子供からなるサンプルは2,566世帯であり、これらのマイクロデータを母集団データセットとしてケーススタディを行う。

母集団に対してサンプリング調査が実施されたことを想定し、母集団データセットより100世帯サンプルを抽出し、式(2)におけるパラメータを推定する。主成分分析によるパラメータ行列Aの逆行列より導出されるパラメ

ータBの推定結果を、表-1に示す。乱数 $ran_i$ によって与えられる非相関変数 $p_i$ を、パラメータBに作用されることによって、世帯構成員の属性変数 $x_i$ が生成される。

##### (2) シミュレーション結果

図-1の初期データ作成フローに従って、ケーススタディに対する各世帯の構成員データの推定を行った。本研究のアプローチは既存研究の手法とは異なる新たなものであるため、これらを比較する適切な手法がない。また、このようなマイクロデータの適合度を評価する手法自体が存在せず、筆者らは新たな手法の構築に関する取り組みを現在行っているところである。

そこで、部分的な検証ではあるが、世帯構成員のタイプ(世帯主、妻、第1子、第2子)別に推定世帯データセットと母集団データセット(観測データ)を集計し、年齢階層別人口の再現状況を検証した。分析結果を図-4に示すが、本手法による世帯マイクロデータの推計は母集団の構造をよく再現しており、推定手法の妥当性が示された。

表-1 パラメータ推定結果

$b_{ik}$	$p_1$	$p_2$	$p_3$	$p_4$	$C$
$x_1$	0.541	-0.652	-0.531	-0.034	9.440
$x_2$	0.506	-0.248	0.809	0.170	9.140
$x_3$	0.490	0.454	-0.010	-0.744	3.640
$x_4$	0.460	0.554	-0.253	0.645	3.010

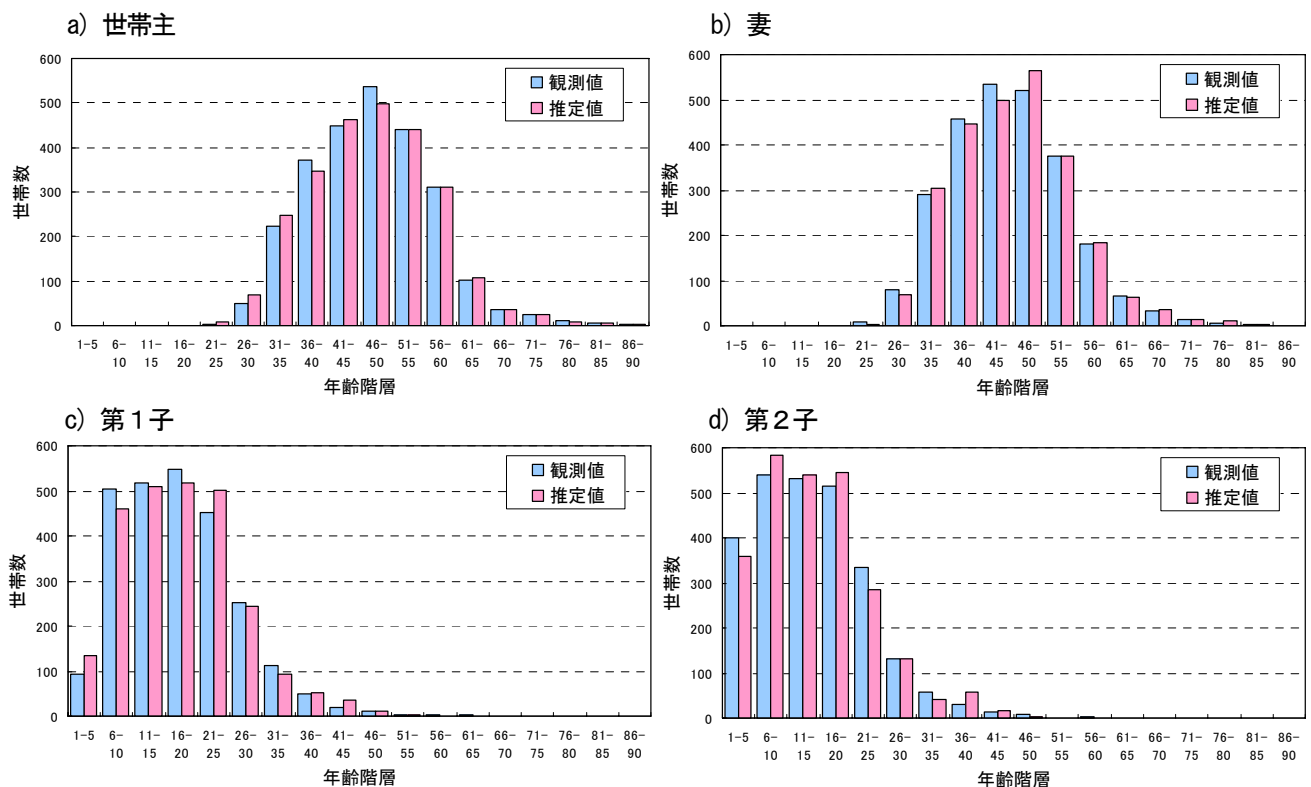


図-4 マイクロ世帯データの推計結果

## 5. おわりに

本研究では、新たなエージェントベースアプローチによる世帯マイクロデータ推定のためのシステムに関して主要部分の一部を構築した。実際のマイクロシミュレーションモデルへ適用可能なシステム構築のためには基本的な手法の構築を行ってゆく必要があるが、構築された部分の妥当性はケーススタディによってある程度確認された。このような問題に関する適合度計測手法の開発を用いて、システムの適用範囲を拡張して行く予定である。

なお本稿の内容に関しては11th International Conference on Computers in Urban Planning and Urban Management (CUPUM)において発表予定であることを付記する<sup>10)</sup>。

本論文は、平成20～21年度科学研究費補助金（基盤研究（B））、課題番号：20360232、研究課題名：詳細属性情報を含む世帯の空間分布予測のためのマイクロシミュレーションシステム）の研究成果の一部を取りまとめたものである。ここに記して感謝の意を表したい。

### 参考文献

- 1) Wegener, M.: Overview of Land-Use Transport Models, Proceedings of CUPUM '03, Sendai, CD-ROM, 2003.
- 2) Deming, W.E. and Stephan, F.F.: On a Least Squares Adjustment of a Sampled Frequency Table when the Expected Marginal Totals are known, Annals of Mathematical Statistics, Vol.11, pp.427-444, 1940.
- 3) Beckman, R. J., Baggerly, K. A. and McKay, M. D.: Creating Synthetic Baseline Populations, Transportation

Research A, Vol.30, No.6, pp.415-435, 1996.

- 4) Guo, J. Y. and Bhat, C. R.: Population Synthesis for Microsimulating Travel Behavior, Transportation Research Record: Journal of the Transportation Research Board, No.2014, pp.92-101, 2007.

- 5) 宮本和明, 安藤淳, 清水英範: 非集計行動分析に基づく都市圏住宅需要モデル, 土木学会論文集, No.365/IV-4, pp.79-88, 1986.

- 6) Pritchard, D. R. and Miller, E.J.: Advances in Agent Population Synthesis and Application in an Integrated Land Use / Transportation Model, 88th Annual Meeting Compendium of Papers, Transportation Research Board, DVD, 2009.

- 7) Moeckel, R., Spiekermann, K., and Wegener, M.: Creating a Synthetic Population, Proceedings of CUPUM '03, Sendai, CD-ROM, 2003.

- 8) Otani, N., Miyamoto, K., and Sugiki, N.: Goodness-of-Fit Evaluation Method between Observed and Estimated Sets of Micro-Data in Land-Use Micro-Simulation, Proceedings of CUPUM '09, Hong Kong, CD-Rom, 2009.

- 9) 宮本和明, 北詰恵一, 鈴木温: 世界における実用都市モデルの実態調査とその理論・機能と適用対象の体系化, 平成18年度～19年度科学研究費補助金（基盤研究（C））、課題番号：18560524）研究成果報告書, 2008.

- 10) Miyamoto, K., and Sugiki, N.: An Estimation Method of Household Micro-Data for the Base Year in Land-Use Micro Simulation, Proceedings of CUPUM '09, Hong Kong, CD-Rom, 2009.