

マイクロシミュレーション都市モデルのための非集計世帯・人口データの作成に関する研究*

Validity of Methods of Disaggregate data for Micro-simulation Urban Model*

杉木直**・宮本和明***・北詰恵一****

By Nao SUGIKI**・Kazuaki MIYAMOTO***・Keiichi KITAZUME****

1. はじめに

近年、わが国では都市の成熟化とともに、少子高齢化に伴う都市サービス需要の地域的な偏りや、都市の郊外化による公共交通の衰退等の課題が顕在化する一方で、自治体財政の逼迫が危惧されており、量的な人口分布のみならず、年齢や所属する世帯の属性を考慮した将来人口分布予測に基づいた、効率的な交通計画、施設配置計画が求められている。

人口・世帯分布の将来予測に基づいた交通施設整備の評価、交通サービスや公共施設配置等の都市サービスの需要予測を目的として、土地利用と交通の相互作用に着目しこれらの変化を記述する都市モデルは、わが国においても多くの研究がなされてきたが、このように計画課題が量的な予測から質的な予測にシフトする状況下において、多様な属性を考慮した人口・世帯の将来分布予測が必要とされている。そのための手法として、欧米諸国では近年マイクロシミュレーションを用いた都市モデル開発に対する関心が高まっており、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¹⁾。しかしながらわが国では、交通施設整備効果の計測を除くと、実用都市モデルの総合的な計画へ適用事例は非常に少なく、マイクロシミュレーション型都市モデルについては実際の都市を対象とした開発や適用がなされていないのが現状である²⁾。

マイクロシミュレーション都市モデルの適用における課題の一つとして、ベースデータとして多属性をもつ非集計の世帯・人口データの作成が挙げられる。このような非集計データは、対象データすべてを収集することは費用面や調査量の観点から不可能であり、また個人情報保護の観点から住民基本台帳等の直接的な利用も難しく、サンプリングや周辺分布に基づいた作成が必要とされる。しかしながら、これらの設定手法が母集団の精度に与え

*キーワード：土地利用・交通・環境統合モデル、人口分布、地域計画

**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、武蔵工業大学環境情報学部

****正員、博士(工学)、関西大学工学部

る影響や妥当性については十分な検討がなされていない。

そこで本研究では、マイクロシミュレーション都市モデルの実際の都市への適用を目標とした上で、その第一段階として、初期時点の非集計世帯・人口ベースデータ作成に関して、IPF法やモンテカルロサンプリングによる既存手法を整理した上で、モンテカルロサンプリング手法に関しては、わが国のデータ整備状況を踏まえた非集計データの作成手法を構築し、各推定手法やサンプリング率の設定による母集団データの精度に関して分析を行う。具体的には仙台都市圏、道央都市圏のパーソナリティ調査によって取得された世帯、個人およびその属性に関する非集計データを母集団とし、周辺分布やランダム抽出した個票データを用いてマイクロデータを作成し、母集団推定の精度を比較する。また、検討目的の視点から手法やサンプリング率の設定の妥当性について考察する。

2. 初期データ作成に関する手法

(1) 世帯・人口分布推定手法に関する既存研究

複数の属性に基づく世帯分布予測を目的とした研究としては、同時確率最大化法によって全世帯を世帯タイプと住宅タイプに分類して現在分布を推定する宮本ら³⁾、IPF法⁴⁾を用いて長期交通需要予測に必要な世帯分布を推定している西田ら⁵⁾の研究があるが、比較的大きなゾーンを対象としたものであり、推定ゾーンの粗さに限界がある。RAS法を用いて国勢調査区毎の世帯の現在分布を推定している青木ら⁶⁾、多次元に拡大したIPF法により世帯の現在分布を推定している五十嵐ら⁷⁾の研究では、比較的详细なゾーンを対象としているものの、取り扱うことが可能な属性数に限界がある。また、以上の既存研究は、西田らの研究では将来分布に関して世帯構成員ベースの予測がなされている事例はあるが、現在分布に関しては属性別世帯数の推計にとどまっており、属性別人口の推定を対象としたものではない。

海外で開発が進められているマイクロシミュレーションモデルにおいても、公表されていない非集計世帯・人口データの生成に関する事例が見られる。現在の主流となっているのは2003年のTRANSIAMプロジェクトにおいて交通と大気汚染予測モデルで適用されている

Beckmanら⁸⁾によるIPF法であり、年齢・性別・収入・職業別の人口推計がなされている。また、Waddleら⁹⁾による土地利用・交通マイクロシミュレーションであるUrbanSimや、Huntら¹⁰⁾によるTLUMIP2 (2nd Generation Oregon Statewide Model) ではBeckmanのIPF法の変形型が初期の世帯および人口分布の推定に適用されている。

一方、モンテカルロサンプリングを用いた研究としては、Moeckelら¹¹⁾が土地利用・交通モデルプロジェクトILUMASS (Integrated Land-Use Modeling and Transportation System Simulation) のための初期世帯・人口分布作成を目的として構築された事例がある。モンテカルロサンプリングは、多属性の世帯・人口分布作成手法としてIPF等の既存手法に対して優位であると考えられるが、現段階では研究事例が少なく、またその精度や有効性について十分な検討がなされていない。

以下、既存手法として多く活用されているIPF法、およびマイクロシミュレーションのための非集計データの作成手法として活用が期待されるモンテカルロサンプリングについて概要を示す。

(2) IPF法

IPF(Iterative Proportional Fitting)法は、図-1のように各属性の周辺分布が既知である時、周辺分布を制約としてサンプルデータとして得られた初期分布に近い同時分布を見つけ出し、未知である分割表の中身を推定する手法である。二次元の場合は、産業連関表に用いるRAS法と数学的には同値である。この手法ではKullback-Leibler情報量の最小化を行っており、初期分布を π 、推定する同時分布を p^* 、凸集合 K 、属性を i, j とすると式(1)のように表現される。

$$I(p^*, \pi) = \min_{p \in K} I(p, \pi) = \min_{p \in K} \sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}}{\pi_{ij}} \right) \quad (1)$$

同時確率最大化法の解法としてその近似式にStirlingの公式を使う限りにおいてはIPF法と全く同じ目的関数になることから、最も確率的に生じ易い分布を推定していることと数学的に同値であるということも出来る。またIPF法は拡大後も属性間のオッズ比が等しく、母集団の再現性が高い。

一方で、多次元IPF法においても、次元数が増加するに従って制約条件の数が増加し、分割表の中身である要素数が増加するために初期分布において個々の要素の信頼性が低下するため、多属性の初期分布推定手法としては限界があることが確認されている⁷⁾。

(3) モンテカルロサンプリング

モンテカルロサンプリングは、統計データ等による1次元属性の周辺分布(世帯主年齢別世帯数、年齢階層別人口等)を、モンテカルロ法を用いて発生乱数に基づいて個別の非集計データに配分し、段階的に各データの複

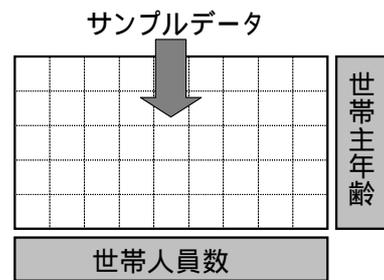


図-1 IPF法

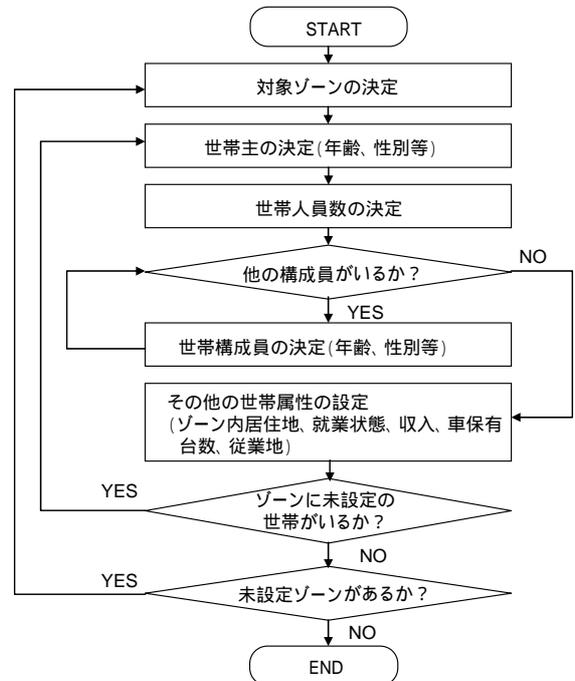


図-2 モンテカルロサンプリングによる初期世帯・人口分布データ作成フローの事例(ILUMASS)¹¹⁾

数の属性を決定しながら最終的に多属性の非集計データを構築する手法であり、理論的には無限の属性を考慮することが可能である。各非集計データに対して属性を決定する際に、サンプルデータによって得られるクロスされた属性に対する確率分布を援用することで、母集団分布の推定精度を向上させることが可能である。

図-2にILUMASSにおける初期世帯・人口データ作成フローの事例を示す。この事例では、世帯主年齢、世帯人員数、世帯内の各構成員の決定にモンテカルロサンプリングが用いられている。その他の世帯属性として、ゾーン内の詳細な居住場所、各世帯構成員の就業状態、世帯の収入、車保有台数、就業状態にある世帯構成員の従業地についても設定しているが、これらは周辺分布を配分するものではなく、サンプリングデータより確率的に設定されるものである。また、以上の手法では、各世帯の世帯構成員として年齢別人口を配分してゆくと、最終的に世帯主に成り得ない子供のデータのみが残る場合が存在するが、すでに決定されたデータより1世帯をランダムに選択し、世帯主として入れ替え可能な構成員と入

れ替えるという手法で対応している。

モンテカルロサンプリングは確率的な試行であるため、個別の非集計データとしては、複数回繰り返すと試行毎に結果が異なるが、作成された非集計データのばらつきが都市モデルの初期分布データとして用いることに対して妥当であるか否かの検証はされておらず、このことが本研究の問題意識の一つとなっている。

3. モンテカルロサンプリングによる非集計データ作成手法の構築

(1) 分析手法の概要

本研究では、ILUMASSにおける構築事例を参考に、わが国での統計データ整備状況を踏まえたモンテカルロサンプリングによる非集計世帯・人口データの作成手法を構築する。図-3に構築されたデータ作成フローを示す。各世帯の属性及び世帯構成員は、国勢調査によって提供されている世帯属性（世帯主年齢×世帯類型、世帯人員数、住宅の建て方、住宅の所有形態）別の世帯数、および性別年齢階層（5歳階級）別人口を周辺分布として各世帯に配分する形で決定される。また、対象地域の世帯属性、世帯構成員、居住する住宅属性に関する非集計サンプルデータが別途調査によって得られており、これらのデータから各属性の決定段階に必要なクロスされた世帯属性に対する確率分布が設定可能であることを前提とする。このような非集計サンプルは、アンケート調査等によって入手するのが一般的であるが、近年仙台都市圏や道央都市圏で行われたパーソントリップ調査においては居住に関する付帯調査が実施されていることから、既存データとして入手可能となっており、今後の交通計画が居住状況や都市構造の変化を考慮せずに策定すること

が困難になっていることを考えると、同様の調査データが提供される機会は増えるものと期待される。

既存手法と比較すると、周辺分布として世帯類型を用いることで世帯構成に関する精度が向上することが期待される。また、始めに世帯主を決定し、人口データより世帯主に相当する個人を配分した上で、その他の世帯構成員を配分する手法により、最後に子供のみが残る状況を回避した。車保有台数や収入などについては、モンテカルロサンプリング手法を用いる部分のみに限定したため本研究のデータ作成フローには含まれていないが、これらについては別途サンプル分布等によって付加可能である。以下、各段階の分析手法について詳細を記す。

(2) 世帯タイプおよび世帯主年齢の設定

国勢調査の世帯主年齢階層別世帯類型別世帯数データより1世帯ずつ取り出し、世帯主の性別を設定する。性別の設定はサンプルデータにおける世帯類型別の性別年齢別世帯主の確率分布に基づいてモンテカルロサンプリングにより決定し、設定された世帯主に相当する人口データより1人減少させる。

(3) 世帯人員数の設定

各世帯の世帯人員数を、サンプルデータの世帯主・世帯類型別の人員数分布に基づいて、モンテカルロサンプリングにより決定し、該当する国勢調査の世帯人員数別世帯数データを配分する。世帯類型のうち、単身世帯や夫婦のみの世帯については、それぞれ世帯人員数が1人、2人であるため、これら以外の類型の世帯が対象となる。

(4) 世帯構成員の設定

決定された世帯人員数に応じて、世帯主以外の世帯構成員データを作成する。世帯主が除かれた人口データを、サンプルデータの世帯主・世帯類型別の世帯構成員の属

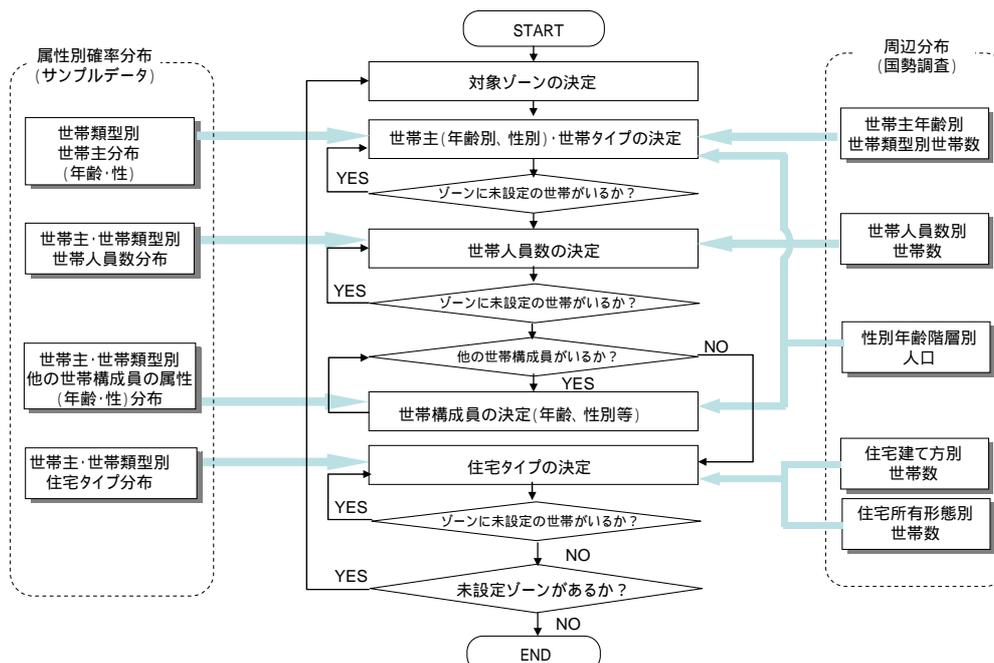


図-3 モンテカルロサンプリングによる非集計データ作成フロー

性（性・年齢）分布に基づいてモンテカルロサンプリングにより配分する。

（５）住宅タイプの設定

国勢調査の住宅の建て方、および住宅所有形態別世帯数を周辺分布として、各世帯の住宅タイプを設定する。世帯主・世帯類型別の住宅タイプ分布に従って、モンテカルロサンプリングを実行する。

４．利用データベースおよび属性の設定

（１）利用データ

本研究では、上記において構築されたモンテカルロサンプリング手法、およびIPF手法の検証を仙台都市圏および道央都市圏のパーソントリップ調査において得られた非集計データを対象として実施する。仙台都市圏では居住と意識に関する意識調査、道央都市圏ではライフスタイル調査がそれぞれ付帯調査として実施されており、住居タイプを含めた非集計データの利用が可能となっている。各調査におけるサンプル数は、仙台都市圏で10,812世帯、道央都市圏では19,394世帯である。

（２）世帯属性の設定

国勢調査および各都市圏におけるパーソントリップ調査より、初期分布の世帯属性として以下の4種を設定して分析を行う。

- ・世帯主年齢：5歳階級（15-19歳～85歳以上）
- ・世帯類型(6)：単身、夫婦のみ、夫婦と子供、その他親族世帯、非親族世帯
- ・世帯人員(7)：1人、2人、・・・6人、7人以上
- ・住宅タイプ(6)：持家戸建、持家集合、賃貸戸建、賃貸集合、官舎・社宅、その他

（３）分析内容

分析は、調査によって得られているデータを母集団として行う。周辺分布は調査データを属性別に集計した世帯数および人口であり、国勢調査によって与えられる属性別データに相当するものを作成する。また、属性別の確率分布を設定するためのサンプルデータは、サンプリング率を設定した上で調査データよりランダム抽出されたものを用いる。対象圏域の総人口、総世帯数に対する調査データの数の制約より、ゾーン分割は行わずに各都市圏を1つのゾーンと仮定して分析を行うが、各手法においてはゾーン別に周辺分布が与えられることが前提となっているため、母集団の推計精度を検証するという本研究の目的に対して問題がないものとする。

モンテカルロサンプリング手法については、推計データと母集団の比較により検証する。推計データは試行毎によって結果が異なるため、複数回試行による推計精度のばらつきを分析し、妥当性を検証する。また、多次元IPF法によって世帯属性（世帯主年齢、世帯タイプ、世帯人員数、住宅タイプ）別の世帯数を推定し、モンテカ

ルロサンプリングによる結果との比較より、各手法の有効性を検証する。

５．おわりに

本研究ではマイクロシミュレーション都市モデルのための、初期世帯・人口ベースデータ作成に関して、既存手法を整理し、わが国のデータ整備状況を踏まえた新たなモンテカルロサンプリング手法を構築した。仙台都市圏および道央都市圏のデータを用いたモンテカルロサンプリング手法の検証、および各手法の比較結果については講演時に報告予定である。

参考文献

- 1) Wegener, M.: Overview of Land-Use Transport Model, Proceedings of CUPUM '03, CD-ROM, 2003.
- 2) 杉木直・宮本和明・北詰恵一・Varameth VICHENSAN: 土地利用 - 交通モデルにおけるマイクロシミュレーションの適用状況, 土木計画学研究・講演集, Vol.35, CD-ROM, 2007.
- 3) 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, vol.365/ , pp.79-87, 1986.
- 4) Saburo Saito: Extensions of Iterative Proportional Fitting Procedure and I-projection Modeling, Kyushu University Press, 1998.
- 5) 西田悟史・山本俊行・藤井聡・北村隆一: 非集計交通需要分析のための将来世帯属性生成システムの構築, 土木計画学研究・論文集, No.17, pp.779-788, 2000.
- 6) 青木俊明・稲村肇・増田聡: 小地区単位における都市の居住特性の将来予測, 土木計画学研究・講演集, No.21(2), pp.523-526, 1998.
- 7) 五十嵐豪・徳永幸之・宮本和明・杉木直: 詳細な属性を考慮した都市圏世帯分布予測システムの構築, 都市計画論文集, Vol.40-3, pp.943-948, 2005.
- 8) Richard J. Beckman et al., Creating Synthetic Baseline Population, Transportation Research A, pp.415-429, 1996.
- 9) Waddell, P., Borning A., Noth, A., Freier N., Becke M. and Ulfarsson G.: Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim, Networks and Spatial Economics, Vol.3, No.1, pp.43-67, 2003.
- 10) Hunt, J.D., Donnelly, R., Abraham, J.E., Batten, C., Freedman J., Hicks, J., Costinett, P.J. and Upton, W.J.: Design of a Statewide Land Use Transport Interaction Model for Oregon, Proceedings of 9th World Conference on Transport Research, CD-ROM, 2001.
- 11) Moeckel, R., Spiekermann, K., and Wegener, M.: Creating a synthetic population, Proceedings of CUPUM '03, CD-ROM, 2003.