

## アクティビティシミュレーションとデータ同化方法の改善に関する研究

山梨大学 工学部土木環境工学科 学生会員 ○小原 拓也  
山梨大学 大学院総合研究部 正会員 佐々木 邦明

### 1. 研究背景・目的

アクティビティシミュレーションによる交通需要・状態の予測は多くの特質を持つ一方、多くの選択を重層的に用いるため、その誤差の積み重ねが課題であった。一方、情報通信技術(ICT)の進展により収集・蓄積が可能となった多種多量のデータ、いわゆるビッグデータの活用が様々な分野で行われている。これらの観測データの活用法として「データ同化」という手法が注目を集めている。データ同化とはシミュレーションモデルに観測データを融合させることであり、主なメリットとしては以下のことが挙げられる。

- 1) 観測データを用いて数値シミュレーションモデルの精度や性能の改善が可能となる
- 2) 観測の不足分の補完や観測誤差の修正が可能となる

データ同化を行うことで、観測データとシミュレーションモデルの相互の欠点を補完することが可能となり、現実とよく一致する、尤もらしい状態の再現が可能になることが期待されている。

同研究室の先行研究<sup>1)</sup>においては、アクティビティモデルを用いた交通状態予測にモバイル空間統計のゾーン滞在人数データを同化させることでシミュレーションモデルの予測精度改善の可能性について検討した。本研究では、先行研究<sup>1)</sup>で構築された手法の改善を通じて、より精度の高いシミュレーションを構築することを目的としている。

### 2. シミュレーション

先行研究<sup>1)</sup>では構築したツアーベースモデルを用いて、マイクロシミュレーションを行っている。マイクロシミュレーションとは、ツアーベースモデルの効用関数を用いて選択を再現し、それらの積み上げの結果として都市全体の活動・移動を再現することである。

#### 2-1. シミュレーションの修正点

##### 1) 拡大係数の処理

先行研究<sup>1)</sup>では、対象となるサンプル 48,683 人に対してシミュレーションプログラムを適用したのちに、パーソントリップ調査で定められた拡大係数を乗じていた。その結果を集計して時間帯別のゾーン滞在人数として扱っていたが、シミュレーションを行った個人の結果を拡大係数倍しているため、シミュレーション誤差が大きくなってしまった。そこで、対象となるサンプルを拡大係数と同じ回数シミュレーションを行うことで、対象となる人全体をシミュレートした。それにより、同一属性においても行動結果にばらつきが入ることとなり、シミュレーション誤差を減じることができ。

##### 2) 二次ツアーの時間帯

ツアーベースモデルのツアーや活動の時刻選択では、①オンピーク②オフピーク③オン/オフピークの3つの選択肢の中から選択されるが、時間帯別の滞在人数等の導出には、①午前ピーク②日中③午後ピーク④その他の4つの時間区分で求める必要がある。先行研究ではシミュレーションで選択された3つの時間帯を、パーソントリップ調査から求めたそれぞれの比率にもとづいて4つの時間帯に配分していた。しかし、主要ツアーと二次ツアーの組み合わせについては、順序や時間帯考慮

キーワード：アクティビティモデル、データ同化、ビッグデータ

山梨大学土木環境工学科交通工学講座

(〒400-8511 甲府市武田 4-3-11 Tell/Fax 055-220-8671)

E-mail : t13ce023@yamanashi.ac.jp

していないため、主要ツアーと二次ツアーの目的地選択で不可能な組み合わせが存在してしまっていた。そこで、パーソントリップ調査から、可能であると考えられる 55 パターンの組み合わせにシミュレーションによって割り付けた。表-1 は主要ツアーでオンピークを選択した場合に、実際に可能な 10 つの組み合わせパターンである。オフピーク、オン/オフピークも同様に 19 パターン、26 パターンにシミュレーションを基に割り付けを行っている。

表-1 主要ツアーと二次ツアーの組合せ

| 主要ツアーを行う<br>時間帯 |    |    | 二次ツアーを行う<br>時間帯 |     |     | 比率<br>(%) |
|-----------------|----|----|-----------------|-----|-----|-----------|
|                 | 開始 | 終了 |                 | 開始  | 終了  |           |
| on              | 午前 | 午前 | on              | 午前  | 午前  | 0.63      |
|                 | 午前 | 午前 |                 | 午後  | 午後  | 3.45      |
|                 | 午前 | 午後 |                 | 午前  | 午前  | 1.88      |
|                 | 午前 | 午後 |                 | 午後  | 午後  | 13.79     |
|                 | 午前 | 午前 | off             | 日中  | 日中  | 8.46      |
|                 | 午前 | 午前 |                 | その他 | その他 | 2.19      |
|                 | 午前 | 午後 | on/off          | その他 | その他 | 42.95     |
|                 | 午前 | 午前 |                 | 日中  | 午後  | 2.82      |
|                 | 午前 | 午前 |                 | 午後  | その他 | 1.25      |
|                 | 午前 | 午後 |                 | 午後  | その他 | 22.57     |

### 3) 中間滞在の時間帯

先行研究<sup>1)</sup>で構築されたツアーベースモデルでは、ツアーを主要ツアーと二次ツアーに分類している。また、東京都市圏の主要ツアーでは主要な目的地以外に立ち寄る中間滞在地も多いと考え、中間滞在の目的地選択も考慮している。先行研究<sup>1)</sup>では、中間滞在を主要目的を行う時間帯と同じ時間帯に配分を行っていた。そのため、ゾーンごとの滞在人数を導出する際に、本来よりも中間滞在数が多くカウントされてしまっていた。

そこで、パーソントリップ調査の比率をもとに、中間滞在を主要目的を行う時間帯の前後に配分を行った。

## 2-2. シミュレーション結果の比較

修正前に比べて修正後では、時間帯別・ゾーン別に比較を行っても、滞在人数が少なくなっていることがわかる。(表-2, 図-1)。要因としては、個人の一日の行動における主要ツアーと二次ツアーの組み合わせとして、不可能なものを除くことができたためであると考えられる。また、中間滞在进行時間帯をパーソントリップ調査の比率を基に特定したことで、本来よりも余分に数えられていたものがなくなったためであると考えられる。

表-2 ゾーンごとの一日の滞在人数

| ゾーン | 修正前     |      | 修正後     |      |
|-----|---------|------|---------|------|
|     | 人数      | %    | 人数      | %    |
| 1   | 526090  | 7.27 | 478531  | 7.43 |
| 2   | 564586  | 7.81 | 487111  | 7.56 |
| 3   | 554835  | 7.67 | 488519  | 7.59 |
| 4   | 493725  | 6.83 | 459419  | 7.13 |
| 5   | 519528  | 7.18 | 455262  | 7.07 |
| 6   | 493477  | 6.82 | 445743  | 6.92 |
| 7   | 506380  | 7.00 | 449620  | 6.98 |
| 8   | 506249  | 7.00 | 449832  | 6.99 |
| 9   | 501064  | 6.93 | 447958  | 6.96 |
| 10  | 520872  | 7.20 | 468982  | 7.28 |
| 11  | 517689  | 7.16 | 460697  | 7.15 |
| 12  | 505815  | 6.99 | 452393  | 7.03 |
| 13  | 517260  | 7.15 | 453422  | 7.04 |
| 14  | 503929  | 6.97 | 442103  | 6.87 |
| 計   | 7231502 | 100  | 6439592 | 100  |

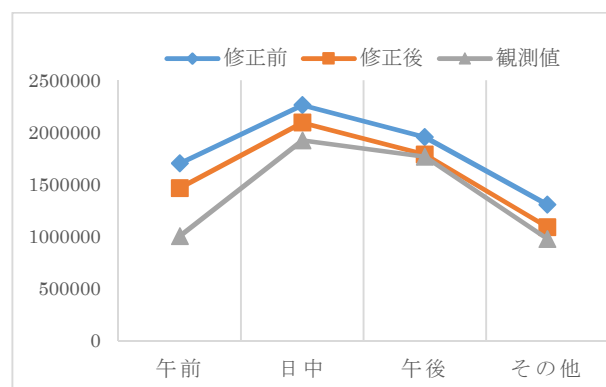


図-1 時間帯別の一日の滞在人数分布

### 3. データ同化

先行研究<sup>1)</sup>では、シミュレーション結果から高い再現性が得られなかったため、滞在人数分布を状態変数として、逐次同化により状態の予測精度の向上を図っている。その際、同化アルゴリズムの計算に時間がかかること、処理時の使用メモリ数が多いことなどの問題があった。

#### 3-1. 同化手法

同化手法として用いたのは、多様な分布に対応可能なパーティクルフィルタである。本研究では、ゾーンへの時間帯別滞在人数の分布を状態変数として、PT 調査からパラメータを推定したツアーベースモデルをシステムモデルとした。そのうえで、モバイル空間統計から得られる時間帯別のゾーン滞在人数を状態の観測として、フィルタリングを行う。特にパーティクルフィルタとは、状態をシステムモデルによって多数のパーティクルによってその分布を再現し、得られた事前分布から観測データに基づいた尤度に比例して、パーティクルの再サンプリングを行う。本研究では状態変数が分布であることから、ツアーベースモデルによるマイクロシミュレーションで再現し、それをパーティクルと見立てている。観測にあう（＝尤度の高い）個人はより多くリサンプリングされ、低いものは破棄される。これにより、事後分布を求めることができる。

#### 3-2. 改善方法

##### 1) 同化アルゴリズムの概要

先行研究<sup>1)</sup>で用いられている同化アルゴリズムの概要は以下のとおりである。

##### 1. 予測： $x_{t|t-1}^i \sim p(x_t|x_{t-1})$

モンテカルロ法により予測分布 $p(x_t|x_{t-1})$ に従う粒子 $x_{t|t-1}^i$ を発生させる。ここでは、粒子 $x_{t|t-1}^i$ はマイクロシミュレーション結果を用いている。

##### 2. 尤度計算： $w_t^i = p(y_t|x_{t|t-1}^i)$

観測値 $y_t$ 、粒子 $x_{t|t-1}^i$ を用いて、重み（尤度） $w_t^i$ を

計算する。

##### 3. 重み $w_t^i$ , $i = \{1, \dots, M\}$ を所与とする

##### 4. ランダム・リサンプリングを行う

##### ① $u_t^i = \text{runif}(n, \min = 0, \max = \sum_{l=1}^n w_t^l)$

ここで、 $\text{runif}(n, \min = 0, \max = \sum_{l=1}^n w_t^l)$ は0以上 $\sum_{l=1}^n w_t^l$ 以下の一様乱数である。

##### ② $\sum_{l=1}^{j-1} w_t^l < u_t^i \leq \sum_{l=1}^j w_t^l$ を満たす $j$ を探す。

##### ③ その $j$ に該当する $x_{t|t-1}^j$ をリサンプリング後の粒子 $x_t^i$ として採用する

##### 2) アルゴリズムの問題点

ランダム・リサンプリングの際に、(2)  $\sum_{l=1}^{j-1} w_t^l < u_t^i \leq \sum_{l=1}^j w_t^l$ を行う場合、繰り返し処理を行うため、処理時間が長くなってしまう。使用しているプログラミング言語のRは、繰り返し処理の実行速度が遅いため、処理時間を短縮するためには繰り返し処理を避ける必要がある。また、粒子の数が多くなるほど、処理時間は長くなる。

##### 3) 改善方法

上記の問題点を解決するために、4.の②の計算の行列化を行った。行列を用いた計算方法に変更することで、一括での計算が可能になるため、繰り返し処理を避けることが可能となる。

#### 3-3. 改善結果

先行研究<sup>1)</sup>では1日を4つの時間帯に分けた「その他の時間帯」で同化を行っていた。その際のプログラムの処理には約30時間を要していた。改善後のプログラムでは、同じ「その他の時間帯」で同化を行う場合、約1分で処理を終えることができるようになり、処理時間を約1/1800とすることが可能となった。

また、計算アルゴリズムの改善を行うことで、4つの時間帯を一括した一日での同化を行うことが可能となった。同化を行う際の処理時間は約3時間である。同化を行った結果は図-2、表-3である。表-3から、同化を行うことでゾーンごとの滞在人数の比率を観測値に近づけられることが確認でき

た．また，今回は，修正したシミュレーション結果を予測値として用いているため，拡大係数の扱い方の変更などにより，シミュレーション誤差を減じた予測値での同化を行うことが可能になった．そのため，同化後の分布が観測値の分布に対して当てはまりが良くなっていると考えられる．滞在人数の総計は観測値に近づけられてはいないが，ゾーン間の滞在人数の差を表現できている．

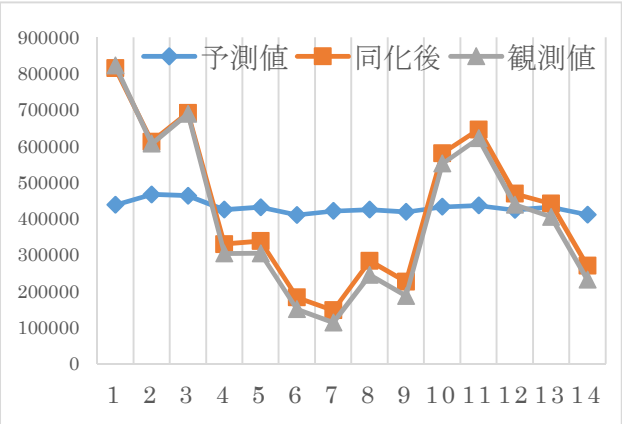


図-2 ゾーンごとの一日の滞在人数分布

表-3 滞在人数の比較

| ゾーン | 予測値     |     | 同化後     |      | 観測値     |      |
|-----|---------|-----|---------|------|---------|------|
|     | 滞在人数    | %   | 滞在人数    | %    | 滞在人数    | %    |
| 1   | 437894  | 7.3 | 814895  | 13.5 | 822086  | 14.5 |
| 2   | 466513  | 7.7 | 612447  | 10.1 | 606504  | 10.7 |
| 3   | 462945  | 7.7 | 691636  | 11.5 | 688933  | 12.1 |
| 4   | 424785  | 7.0 | 329825  | 5.5  | 304000  | 5.4  |
| 5   | 431400  | 7.1 | 338375  | 5.6  | 304342  | 5.4  |
| 6   | 410114  | 6.8 | 183274  | 3.0  | 150462  | 2.7  |
| 7   | 421392  | 7.0 | 147630  | 2.4  | 114041  | 2.0  |
| 8   | 424585  | 7.0 | 283542  | 4.7  | 244758  | 4.3  |
| 9   | 418687  | 6.9 | 226093  | 3.7  | 187054  | 3.3  |
| 10  | 432679  | 7.2 | 580345  | 9.6  | 551942  | 9.7  |
| 11  | 436429  | 7.2 | 645845  | 10.7 | 621716  | 11.0 |
| 12  | 423967  | 7.0 | 469102  | 7.8  | 438549  | 7.7  |
| 13  | 431889  | 7.2 | 441347  | 7.3  | 405070  | 7.1  |
| 14  | 411269  | 6.8 | 270193  | 4.5  | 232090  | 4.1  |
| 計   | 6034548 | 100 | 6034548 | 100  | 5671546 | 100  |

4.まとめ

先行研究<sup>1)</sup>で行われたシミュレーションプログラムの改善を行った．個人を複数回シミュレートする方法に変更したため，同一属性においても行動結果にばらつきが入ることとなり先行研究<sup>1)</sup>のシミュレーション結果に含まれていたシミュレーション誤差を減じることができた．

また，同化アルゴリズムの改善を行うことで，先行研究<sup>1)</sup>で計算処理にかかっていた時間を短縮することができた．処理時間が膨大であったため，先行研究では「その他の時間帯」でのみ同化を行っていたが，改善したプログラムを用いて一日での同化を行い，シミュレーションの事前分布が観測データの滞在人数分布に近づくことが確認できた．

今後の方針としては，図-2からもモデルの再現性が高くないことがわかるため，同化を行った結果を用いながら，目的地選択モデルの定数項の推定を行う．発表時にはその推定結果についても示したい．

5.参考文献

- 1) 澤田茜：一日の行動を再現する大都市圏非集計需要予測モデル，2016
- 2) 北村隆一・森川高行・佐々木邦明・藤井聡・山本俊行：交通行動の分析とモデリング，技報堂出版，2002
- 3) 樋口知之：予測にいかす統計モデリングの基本，株式会社講談社，2011
- 4) 間瀬茂：R プログラミングマニュアル，数理工学社，2007
- 5) 細川敬祐：データ同化の考え方とその方法，電気通信大学