

継続的観測データを活用した 行動モデルパラメータの更新可能性

澤田 茜¹・佐々木 邦明²

¹正会員 パシフィックコンサルタンツ株式会社（〒101-8462 東京都千代田区神田錦町三丁目22番地）

E-mail: akane.sawada@tk.pacific.co.jp

²正会員 山梨大学教授 大学院総合研究部（〒400-8511 山梨県甲府市武田4-3-11）

E-mail:sasaki@yamanashi.ac.jp

本研究では、継続的に入手可能である観測データを用いて、アクティビティシミュレーションモデルに用いられるパラメータの年次更新を行った。本来、シミュレーションの予測を更新する際には、行動モデルパラメータは固定され、説明変数等の更新を行ってきたが、これらの変数の更新頻度は高くなく、土地利用等の変化の激しい変数については、統計データに現れるデータはタイムラグや多くの誤差を含む可能性がある。一方、いわゆるビッグデータに代表される観測データは継続的に取得され、そのような時間的な変動が反映されると考えられる。そこで、継続観測の情報をモデルパラメータに反映することで、モデルパラメータの更新を行うアプローチによってパラメータの更新を行い、新たなデータを用いたモデルとデータの関係についての考察を行うものとする。

Key Words: activity micro simulation, Bayesian update state-space model

1. はじめに

近年、政策などをよりミクロに質的評価することが求められており、個人の移動や移動の変化についてより詳細に把握する必要性が高まっている。そこで本研究では、アクティビティシミュレーションによる個人の生活パターンの予測の精度を、観測データと融合することで改善することを提案・検証することを目的とする。著者らによる既存研究¹⁾では、ツアーベースのモデルを用いて予測を行い、目的地選択の精度の課題を確認している。その後、予測値に対して事後的な観測データ同化し、滞在人数分布を状態変数として個人の再配置を行い、観測値が得られた状態での予測状態の修正が行われている。この個人を再配置した状態から次期の予測を行うため、再配置した情報をアクティビティシミュレーターのパラメータを更新し、観測データにあわせて再配置された状態をモデルパラメータ更新に反映する枠組みを提案した²⁾。本研究では、この手法を用いながら、継続的に入手可能な観測データを用いて年度ごとの継続更新を行い、更新されたモデルによる精度検証を行う。

本来、アクティビティシミュレーションの予測は土地利用データやLOSデータなどの説明変数等を基準年におけるデータを用いてパラメータを推計し、予測年次の外生変数を用いて予測することになるが、

モデルパラメータ推定のための変数の更新頻度はそれほど高くなく、土地利用等の変化の激しい変数については、モデル推定の時点で実態と合わなくなっている可能性もある。一方、ICTによる継続観測は、ほぼ同一条件での時系列変化を内包するため、これらの変化に対応した観測値となっている。そこで、ICTによる継続観測をモデルパラメータに反映することは、これらによる誤差を考慮できることを意味する。

そこで、本研究では、特に時系列的な安定性に課題のあるモデル定数項を再配置に応じて再推定し、更新した定数項を用いたアクティビティシミュレーションを行うことを提案し、それによる精度改善の検証を目的とする。

2. 研究の位置付け

著者らの先行研究¹⁾では、ツアーベースの行動モデルを用いてシミュレーションにより滞在人数分布を推計し、予測した滞在人数分布の観測データとしてモバイル空間統計による人数分布を用いて同化を行った。同化されることにより観測データに応じて滞在人数分布の再現性は向上していると確認された。この補正された滞在人数分布を用い、定数項を再推定するというアクティビティシミュレーションの更

新手法を提案した²⁾。そこで、本研究では、複数の観測データが利用可能な状況で、継続的に定数項パラメータの更新を行い、更新を行ったモデルと、行わないモデルでの予測精度の検証を行った。また、この定数項パラメータの差の検定を年度ごとに行い、継続的に入手可能な滞在人数分布の観測データによる更新の可能性を説明する。

各モデルやシミュレーション内容や詳しい推定結果は先行研究¹⁾を確認していただきたい。

3. 事例研究

(1) 分析対象データ

モデルの構築、分析に用いたデータは以下の通りである。

a) 東京都市圏パーソントリップ調査

平成20年に行われた第5回東京都市圏パーソントリップ調査の全サンプル数は73万3873人である。調査の対象地域は、東京都、神奈川県、埼玉県、千葉県、茨城県南部（東京都島嶼部を除く）である。本研究では、計算の負荷を低減するため、東京都市圏全域に居住する人で、都心三区（港区、千代田区、中央区）を主要なトリップの目的地にした個人を抽出して分析を行った。モデルに用いた最終的なサンプル数は4万8683人となった。

この調査を用いて、ツアーベースの行動モデルを構築し、マイクロシミュレーションを行った。滞在人数を推定した時刻、ゾーンは表-2と図-1（表-3）の4つの時間帯、都心三区14ゾーンである。この区分での滞在人数を状態変数として推定を行っている。

b) モバイル空間統計データ³⁾

モバイル空間統計データは、ドコモインサイトマーケティング社が提供するNTTドコモの携帯電話基地局データを使用して作成される人口の統計情報である。モバイル空間統計は、非識別化処理、集計処理、秘匿処理の3段階のプロセスを経て作成される。運用データからのモバイル空間統計の集計処理は、以下の3段階の推計プロセスから構成される。

① 在圏数推計処理

非識別化された運用データに基づき、運用データのばらつきなどを考慮しつつ、基地局エリアごとの携帯電話の在圏数を推計する。

② 拡大推計処理

在圏数推計処理の結果である基地局エリアごとの携帯電話の在圏数に基づき、ドコモ携帯電話の普及率や偏りや携帯電話の電源断の影響などを考慮しつつ、基地局エリアごとの人口へと拡大推計する。

③ エリア変換処理

拡大推計処理の結果である基地局エリアごとの人口に基づき、基地局エリア単位に推計された人口をメッシュや行政区画単位など、応用分野で活用しやすい集計単位へと変換する。

本研究で利用するデータは、500mメッシュ単位の都心三区の人口流入分析である。

表-1 モバイル空間統計データ概要

対象地域	千代田区, 中央区, 港区
対象期間	平成26年, 平成27年, 平成28年 各10月
時間区分	1時間ごとに観測
粒度	500mメッシュ
データ概要	対象期間の時間帯別平均値

表-2 時間区分の定義

時間区分	
午前ピーク	(AM 6:30~AM 9:29)
日中	(AM 9:30~PM 3:59)
午後ピーク	(PM 4:00~PM 6:59)
その他	(PM 7:00~AM 6:29)

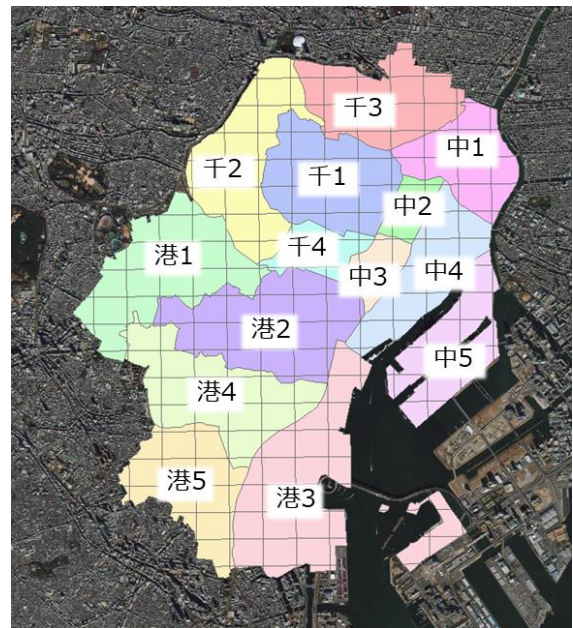


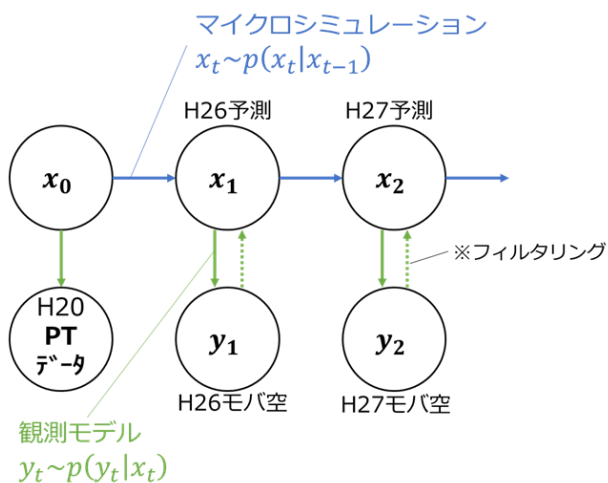
図-1 対象ゾーン

表-3 時間区分の定義

ゾーン	区名	地名
1 (千1)	千代田区	大手町・丸の内等
2 (千2)		永田町・飯田橋等
3 (千3)		神田等
4 (千4)		有楽町・霞ヶ関等
5 (中1)	中央区	日本橋等
6 (中2)		日本橋・八重洲等
7 (中3)		銀座等
8 (中4)		築地・明石町等
9 (中5)		勝どき・晴海等
10 (港1)	港区	赤坂・青山等
11 (港2)		六本木・新橋等
12 (港3)		芝浦・台場等
13 (港4)		麻布・芝等
14 (港5)		高輪・白金等

(2) 先行研究でのモデルフレーム

先行研究では、ツアーベースの非集計モデルを用いてシミュレーションによる予測を行い、モバイル空間統計を観測値として利用することを前提として、図-2に示すような状態空間モデルのフレームを用いる。状態空間モデルは図-2に示すような状態ベクトル x_t と観測ベクトル y_t から構成される。時刻 t における状態ベクトル x_t は、直接観測できない、または推定したい変数のベクトルである。観測ベクトル y_t は、状態変数の観測である。本研究では、状態ベクトル x_t をシミュレーションにより得られた滞在人数分布とし、計算負荷を低減するために1次元の分布として用いる。観測ベクトル y_t はモバイル空間統計から得られるゾーン別の滞在人数観測データとする。ここで時刻 t は、ツアーベースモデルが一日単位の再現を行うことから、逐次的なものでなく、年次レベルの間隔を想定している。ここで、年次の予測更新時はマイクロシミュレーションを用いるが、このモデルが年次ごとに変化することを想定している。大枠での流れは以下の通りである（図-2参照）。



1. PT調査データを時点 $t = 0$ として、アクティビティモデルのパラメータの推定を行い、アクティビティモデルを構築する。
2. 1.で構築したアクティビティシミュレーションに時点1のLOS等を代入し、時点1の滞留人口分布を求め、それを状態ベクトル x_1 とする。
3. 時刻1における観測ベクトル y_1 を取得し、この滞在人口分布を用いたフィルタリングを行い、状態 x_1 を推定する。
4. 新たに推定した状態 x_1 を再現する定数項を推定し、これを更新されたアクティビティシミュレーションの定数項とする。このモデルと次期のLOS等のデータを用いて次期の予測分布 x_2 を得る。

これを繰り返すことによって、新たな観測データ

が得られた場合に、その情報を用いて新たな状態を推定し、その状態をシミュレーションモデルに取り込んで次期の予測に用いる。時間軸に沿ってシミュレーションを逐次的に行うモデルの場合には、新たな配置状態から次のシミュレーションを実施することで、観測データの情報を取り込むことができる⁴⁾が、今回のツアーベースモデルは一日単位の再現で、代表的な一日の再現となるため、更新された状態変数（配置状態）をモデルのパラメータに情報を変換し、シミュレーションモデル取り込んだ上で、次期の予測に活用することにした。本方式では定数項の再推定などの手数は増えることになるが、新たな交通状態に関する観測データの情報が取り込まれることにより、次期の滞在人数分布の予測精度の改善を図れると考えている。

次項では上記4番の通り状態 x_1 に対応するパラメータを求め、再度シミュレーションによる予測を行うことを検証する。

(3) アクティビティシミュレーションの逐次更新

実際に本研究で適用した更新の流れを図-3に基づきながら、簡潔に以下で述べる。

- H20 H20のPT調査を基にモデルとアクティビティシミュレーションシステムを構築する。
- H26 シミュレーションにH26のLOSを代入し、H26の予測結果を事前分布として得る。H26の観測データが入手できたら、H26の予測結果を観測データに同化させ、事後分布を得る。これを基にモデル定数項を更新する。
- H27 更新した定数項を用いたシミュレーションを基に、H27のLOSを代入し、**H27の予測結果**を事前分布として得る。H27の観測データが入手できたら、H27の予測結果を観測データに同化させ、事後分布を得る。これを基にH27の定数項を更新する。
- H28 H27の定数項を更新したシミュレーションを基に、H28のLOSを代入し、**H28の予測結果**を事前分布として得る。H28の観測データが入手できたら、H28の予測結果に観測データを同化させ、事後分布を得る。これを基にH28の定数項を更新する。

このように LOS 等の説明変数を更新していただくだけでなく、予測値の観測が得られた時に、その状態に合わせて定数項のパラメータを更新する。

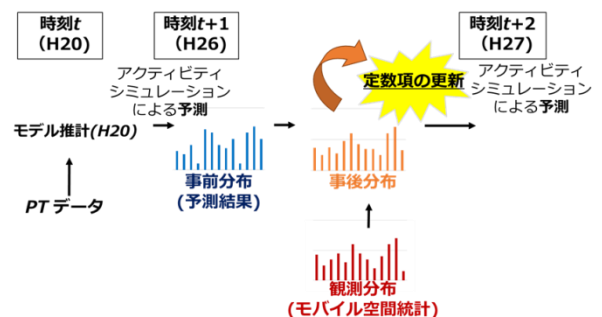


図-3 更新の流れ

提案する方式は、各年次などの単位で一日のアクティビティシミュレーションを行い、観測データを得て個人の滞在場所等をフィルタリングによって更新し、その情報をモデルに取り込んで、次期の予測を行うことになる。このことは、様々な要因によって変化する人の動きに対して、それら全体を包括した動きの情報がモデルに含まれることになる。

一方、モデルは説明変数に用いる様々な変数に含まれる誤差によって影響される。例えば、目的地選択モデルに用いる変数は、人口や従業者数、事業者数、公共施設数、集客施設数、容積率など様々なものが挙げられるが、この多くは5年に1回程度調査されるものや、10年以上調査されていないもの、1年分のデータしか入手できないものも多く、年次更新レベルでは変化のないまま予測を行うことになる。それらはモデルの誤差として扱われ、特に定数項にそれらの期待値が集約される。そこで、観測データを得るたびに定数項の補正を行うことは、説明変数の観測誤差などの影響を考慮できることになる。

本研究で検証を行うのは、H27の予測と H28の予測である。いずれも比較は、定数項を更新しない予測と、前期で定数項の更新を行ってそれに基づいて予測を行ったものである。

前節で示したように、更新された配置の情報を次期予測に反映するために、目的地選択の定数項の再推定を行った。推定方法は、観測データに同化するために、個人の観測データにおける出現確率を重みとしたインポートランスサンプリングで各個人をリサンプリングし、リサンプルされた個人を推定のためのサンプルとして、定数項の最尤推定を行った。ここでは、定数項以外のパラメータは固定して、定数項のパラメータのみを再推定した。これによって、ゾーンの滞在人数の分布は、同化後の分布におおむね一致する。このプロセスで再推定した定数項を表-4、表-5に示した。この定数項パラメータの再推定した結果をアクティビティシミュレーションモデルの定数項として使用して時期を予測する。以下の図-4で H27の、図-5で H28の結果を示し、精度検証を表-6で行う。

まず、表-6の精度検証結果を見ると、観測値とのベクトル距離は、定数項を更新しないものよりも、定数項を更新した方が小さくなり、精度改善につながる事が示された。また、定数項を更新しないと H27よりも H28の方が、誤差が大きくなっていることが示されている。したがって、時間経過によりシミュレーション精度も落ちてくる事が分かる。図-4、図-5で示した通り、ゾーン別滞留人口比率を比較すると、おおそ観測分布に近づいていることが分かる。定数項更新の有無でそれぞれのグラフを比較すると、定数項を更新せずに、土地利用データ等の説明変数のみ更新を行うだけでは、精度が悪くなってしまいうゾーンが存在することが分かる。こういった説明変数で表現できない部分が定数項で表現可能になっていることが示唆される。

表-4 定数項再推定結果 (H26)

	パラメータ	t 値
定数項 zone1	0.0776	21.99
定数項 zone2	0.332	97.45
定数項 zone3	0.990	295.39
定数項 zone4	-2.45	-591.76
定数項 zone5	0.489	124.96
定数項 zone6	-0.661	-146.27
定数項 zone7	-0.782	-168.37
定数項 zone8	0.288	72.36
定数項 zone9	-0.508	-97.96
定数項 zone10	0.577	151.71
定数項 zone11	0.817	236.04
定数項 zone12	1.03	277.40
定数項 zone14	-0.422	-88.20
サンプル数		2194117
初期尤度		-6605375
最終尤度		-5435806
決定係数		0.177
修正済み決定係数		0.177

表-5 定数項再推定結果 (H27)

	パラメータ	t 値
定数項 zone1	0.376	153.63
定数項 zone2	0.922	377.47
定数項 zone3	-2.50	-738.84
定数項 zone4	0.414	129.51
定数項 zone5	-0.699	-182.64
定数項 zone6	-0.938	-232.68
定数項 zone7	0.266	82.83
定数項 zone8	-0.550	-118.46
定数項 zone9	0.541	178.96
定数項 zone10	0.806	315.17
定数項 zone11	1.11	396.09
定数項 zone12	-0.389	-95.14
定数項 zone14	0.376	153.63
サンプル数		2270474
初期尤度		-6889131
最終尤度		-5632532
決定係数		0.182
修正済み決定係数		0.182

表-6 精度検証

	H27 観測値との ベクトル距離	H28 観測値との ベクトル距離
定数項更新なし	21.798	22.984
定数項更新あり	16.596	11.573

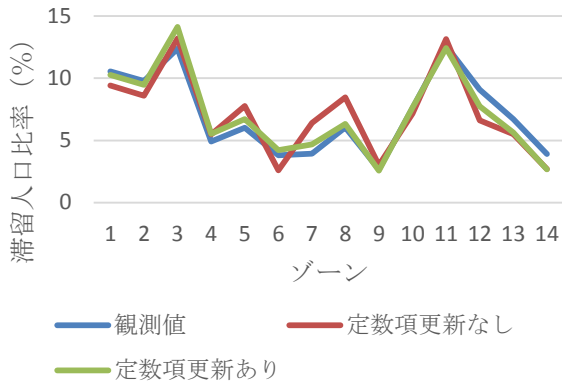


図-4 H27シミュレーション結果

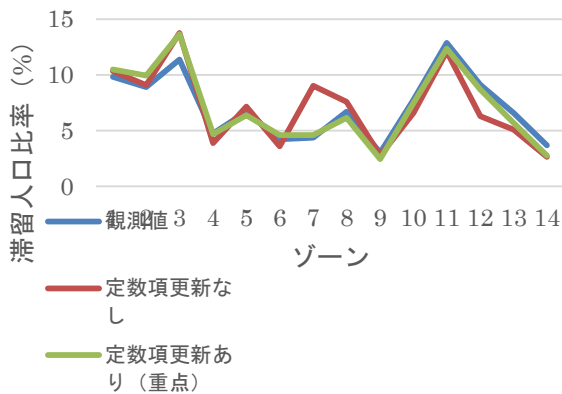


図-5 H28シミュレーション結果

(4) 各年度の定数項パラメータ有意差検定

ここで、H26と H27、H28の定数項推計結果を用いて、パラメータの差をみるために t 検定を行った結果を表-7に示す。ただし、前節で示した推定結果では、定数項の数に違いが生じてしまい、比較が難しくなることから、本節では各年度の定数項を11個で固定して推定し比較した。

t 検定を行った結果、H26と H27ではどのゾーンにおいても有意差はみられなかったが、H27と H28では、ゾーン5,8,9で有意差がみられ、H27と28の1年間で有意に定数項のパラメータが変化していることが確認された。これらから、説明変数で表現できないモデル構造の変化が示唆される。

観測データの年度別の滞留人口(図-6)をみても、H26と H27はほとんど変わらないが、H28には全体的に滞留人口の増加がみられる。特に有意差がみられたゾーン5(銀座)、8(築地等)ではH27からH28にかけて観測データでも増加が確認されているが、銀座地区では再開発が進んでいることや、中央区全体では住宅などの増加により、定住人口の増加が続いていることなどが考えられる。この結果により、本研究によるアプローチによって、土地利用データ等が更新されない場合でも、年々変化しつつある土地利用変化等による滞在人数分布の変化を考慮することが可能であると示唆された。

表-7 パラメータ差の t 検定

	H26と H27の パラメータ差の t 検定	H27と H28の パラメータ差の t 検定
定数項 zone2	-0.847	0.475
定数項 zone3	0.811	0.575
定数項 zone4	0.455	-0.816
定数項 zone5	0.783	-2.569
定数項 zone6	0.279	-0.865
定数項 zone7	1.617	-0.910
定数項 zone8	0.108	-2.458
定数項 zone9	0.299	-2.382
定数項 zone10	0.304	-1.366
定数項 zone11	-0.034	-1.278
定数項 zone14	-0.496	-0.262

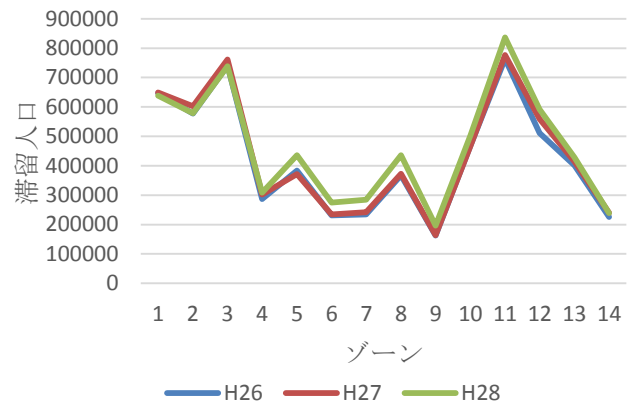


図-6 モバイル空間統計の年度別滞留人口

4. おわりに

本研究では、一日の行動を再現し、OD表などの予測を行うツアーベースモデルのシミュレーションの予測を、モバイル空間統計観測データに同化し、その結果を次期のシミュレーションへ反映し、シミュレーション結果改善を可能とする推計手法の提案を行った。特に予測精度の低い目的地選択の定数項のみを更新することで、簡易にモデル全体の精度を上げることができた。定数項は説明変数では説明しきれなかった誤差の期待値であるので、更新した配置状態をベースに定数項を更新することで、観測データに含まれる目的地選択に関する情報を反映させることができる。得られた更新定数項を用いることで、定数項を更新しない場合と比較して精度改善が確認できた。また年度ごとの定数項のパラメータの差の検定を行ったところ、有意にゾーンの変化がみられる箇所が存在した。これにより、説明変数で表現できない変化が存在し、有意に土地利用等の変化が起きていることが示唆される。

以上より、様々な観測装置により継続的に観測さ

れるインデックスをモデルに反映することは、これらの誤差を考慮できる可能性を示した。この試みは、モデルを推定し、その説明変数の変化によって予測を行うという従来の手法論とは少し異なる。モデルによって予測される様々な状態変数に対して、それに対応した観測データを用いることで、各時点でモデルの内部構造を調整し、予測をさらに精度良く行うことになる。今回は単一の状態変数と観測値を用いたが、今後はより多くの観測値を活かしてモデルの更新を行いながら予測を行う手法が可能になってくる。特に短期的な予測については、様々な観測変数を用いることで、様々な面からの補正が可能になり、精度の高い短期予測が可能になると考えられる。

以上、観測データとモデルの新たな関係の一例として示したが、複数の異なる観測データの活用については、まだ多くの課題が残されており、今後とも継続的に研究を行うことが求められる。

謝辞：本研究は国土交通省「平成27年度、道路政策の質の向上に資する技術研究開発」の委託研究で実施し、数多くの有益なアドバイスを頂いた。ここに記して感謝の意を表します。

参考文献

- 1) 澤田茜, 川辺拓哉, 白須瑛紀, 佐々木邦明: アクティビティマイクロシミュレーションと観測データの融合による需要予測手法, 土木計画学研究発表会・講演集, Vol53, pp.2007-2015, 2016.
- 2) 澤田茜, 佐々木邦明: 観測データの融合によるアクティビティシミュレーションの更新手法, 土木計画学研究発表会・講演集, Vol56, P66, 2016.
- 3) 寺田雅之, 永田智大, 小林基成: モバイル空間統計における人口推計技術, NTT DOCOMO テクニカル・ジャーナル Vol.20 No.3
- 4) 原田遼: 詳細な交通行動推定のためのアクティビティシミュレーションと観測データの統合に関する研究, 東京大学社会基盤工学専攻修士論文, 2017. .
- 5) J.L. Bowman and M.E. Ben-Akiva: Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A* 35, pp.1-28, 2000.
- 6) Bradley, M, J. L. Bowman and B. Griesenbeck. SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution, *Journal of Choice Modeling*, Vol. 3 No. 1, 5-31, 2010.
- 7) Mohamed Omer, Kuniaki Sasaki and Kazuo Nishii, Tour-based Travel Demand Modeling using Person Trip Data and its Application for Advanced Policies, *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol.17, pp.112-127, 2009.
- 8) 鈴木宏典, 中辻隆: フィードバック原理に基づく交通状態推定手法を応用した高速道路上起終点旅行時間の推定, 土木学会論文集, No.695, pp.137-148, 2002.
- 9) Yang, S., Kalpakis, K. and Biem, A.: An adaptive observation site selection strategy for road traffic data assimilation, *IWCTS '12*, pp.2-7, 2012.
- 10) 長谷川瑤子, 関本義秀, 金杉 洋, 櫻山 武浩: 同化手法を用いたスパースな携帯基地局情報に基づく人の移動推定, 交通工学研究発表会論文集, No.54, pp.321-328, 2014.
- 11) 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, pp.126-130, 2002.
- 12) 杉木直, 宮本和明: 土地利用マイクロシミュレーションモデルにおける空間集計・主体集計の影響分析
- 13) 矢野浩一: 粒子フィルタの基礎と応用: フィルタ・平滑化・パラメータ推定, 日本統計学会誌, 第44巻, 第1号, pp.189-216, 2014.
- 14) 北川源四郎: モンテカルロ・フィルタおよび平滑化について, 統計数理, No.44, pp.31-48, 1996.

(???? 受付)

THE POSSIBILITY OF UPDATING MODEL PARAMETERS BY USING CONTINUOUS OBSERVATION DATA OF TRAFFIC STATE

Akane SAWADA, Kuniaki SASAKI