

# 交通性能照査のための混合正規分布モデル によるトリップ長構成の推定

山本 真誉<sup>1</sup>・柳沼 秀樹<sup>2</sup>・寺部 慎太郎<sup>3</sup>・康 楠<sup>4</sup>

<sup>1</sup>学生非会員 東京理科大学大学院 理工学研究科土木工学専攻修士課程

(〒278-8510 千葉県野田市山崎2641) E-mail:7616632@ed.tus.ac.jp

<sup>2</sup>正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: yaginuma@rs.noda.tus.ac.jp

<sup>3</sup>正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: terabe@rs.tus.ac.jp

<sup>4</sup>正会員 東京理科大学 理工学部土木工学科 (〒278-8510 千葉県野田市山崎2641)

E-mail: kangnan@rs.tus.ac.jp

性能照査型道路設計では、道路の機能区分に応じて担保すべきトリップ長が分離されていることが望ましい。しかしながら、実際の幹線道路では様々な距離のトリップが混在しており、トラフィック機能とアクセス機能が曖昧になっていることが多い。トリップが混在する要因として、規定速度に関係なく沿道施設が点在していることが挙げられているが、沿道状況がトリップ長構成にどれほどの影響を与えるかを定量的に評価する手法は存在しない。本研究では、埼玉県春日部市から千葉県白井市間の国道16号をケーススタディとした1ヶ月分のプローブデータにより、トリップ長構成を混合正規分布モデルで推定する。

**Key Words :** traffic performance, hierarchical road, probe data, trip length, gaussian mixture model

## 1. はじめに

現在の道路設計では、年間8760時間の時間交通量の上位30位を設計時の交通需要に用いている。しかしながら、交通需要は季節や曜日、時刻といった時間的な要因により変動するため、観光地では大型連休で交通量が大きく増加するといった時系列因果が考慮できないなどの問題点がある。これらを踏まえて、道路の求められる性能を保証する性能照査型道路設計への転換が検討されている<sup>1)</sup>。性能照査型道路設計では、一般的な時間交通量の変動状況からピーク時の時間交通量を交通需要として、実現する旅行速度などの交通現象を加味した性能照査を行う<sup>2)</sup>。

性能照査型道路設計では、道路の機能区分を明確にすることが求められている。短距離トリップが施設に進入する際の減速がトラフィック機能の低下を招くため、トラフィック機能重視の道路では長距離トリップ、アクセス機能重視の道路では短距離トリップを走行させるなど、道路の機能区分に応じて分担させることが望ましい。しかしながら、多くの幹線道路では、沿道施設が点在することにより長距離トリップと短距離トリップが混在している。また、短距離トリップは、通勤・退勤により朝・夕方にピークをむかえ、ピーク時の終盤から運送業務により長

距離トリップが発生していると思われる。このように、異なるトリップ長が重なりあった状態で観測されたピーク時の交通需要をもとに設計時の性能照査が行われると、前述の短距離トリップによる阻害により、実際の旅行速度が設計時より低くなる。そのため、交通需要を推定する際には、トリップ長の構成を考慮する必要がある。

本論文では、埼玉県春日部市～千葉県白井市を含む5枚の2次メッシュの内の国道16号をケーススタディとし、2015年10月分のプローブデータから混合正規分布モデルを用いたトリップ長構成の推定を試みる。

## 2. 分析データの概要

### (1) 利用データ

本研究では、(株)ナビタイムジャパンから購入した携帯カーナビプローブデータを使用した。このデータは、同社が提供しているアプリケーション NAVITIMEドライブサポーターもしくはCAR NAVITIME for Smartphoneが起動しているときに記録され、走行車がどの道路にいつ進入し、退出したかが記録されている。運転者ごとに日別ユーザーIDと経路IDが

付与されており、アプリが終了もしくは同じ地点に15分以上停車するとデータの記録が打ち切れ、再出発時に新たに異なる経路IDがふり直される。データ取得期間は、2015年10月の1ヶ月間である。

また、プローブ車両の走行道路は道路ネットワークデータのリンク・ノードデータにマッチングしている。道路ネットワークデータには、走行道路の道路規格など道路情報も含まれている。

(2) 対象地域

研究対象地域は、埼玉県春日部市から千葉県白井市を含む5枚の2次メッシュとして、国道16号をケーススタディ路線とした。その理由として、住宅開発や大規模店舗の開発により、交通が増加することで渋滞が促進されていること<sup>3)</sup>、開発により異なるトリップ長の混在が想定されることなどが挙げられる。

国道 16 号の対象区間で代表的な交差点を春日部市、野田市、柏市、白井市から小湊、野田市駅前、呼塚、白井を選出した。交差点は、様々なトリップ長が混在するように市役所付近かつ一般国道または主要地方道と接続する交差点を選出した。

さらに、交差する道路の種別が異なる環境でのトリップ長の混在の違いを見るために、柏市内で主要地方道と交わる若柴を選出した。

(3) 土地利用現況図

沿道に商業用地、住宅用地が点在する場合、私事による短距離トリップ、工業用地が業務による長距離トリップが発生すると想定され、トリップ構成には沿道土地利用が関係していると思われる。そこで、沿道土地利用のデータとして土地利用現況図を用いる。土地利用現況図とは、都市における人口、産業、土地利用、交通などの現況及び将来の見通しを定期的に把握するために行われる都市計画基礎調査の調査項目の1つである。本研究では、埼玉県、千葉県、茨城県の3県の平成22年度土地利用現況図のデータを用いる。

3. トリップ長構成の実態分析

(1) トリップ長構成

呼塚、白井の交差点のトリップ長構成を見ていく。トリップ長構成は、各距離帯のトリップ数を各交差点の全トリップ数で除した値である。呼塚と白井のトリップ長構成を図-1ならびに図-2に示す。

両者ともに正規分布のような形状とならず偏在していることが確認できる。呼塚の分布では、40km以降で構成率が減少するが様に分布しているため、トリップが混在していると考えられる。また、白井の分布では25km付近にピークが存在していることから、呼塚ほどトリップの混在が起きていないことがわかる。

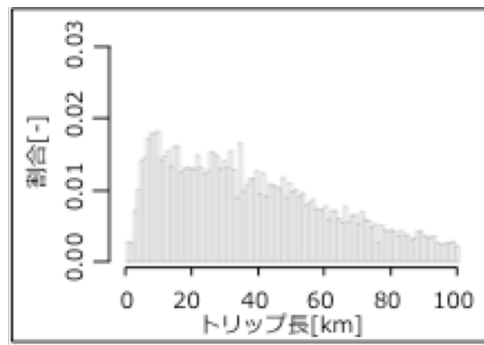


図-1 呼塚のトリップ長構成

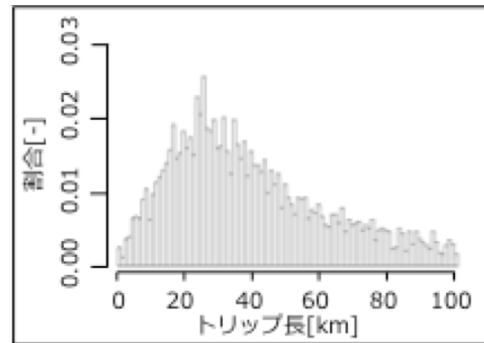


図-2 白井のトリップ長構成

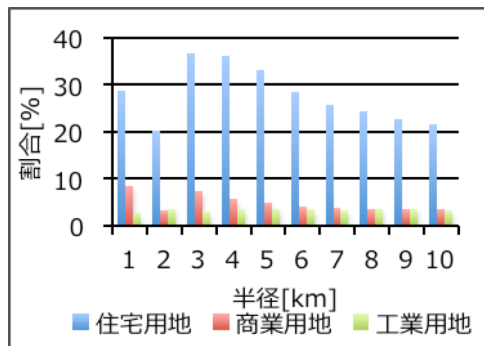


図-3 呼塚の現況土地利用

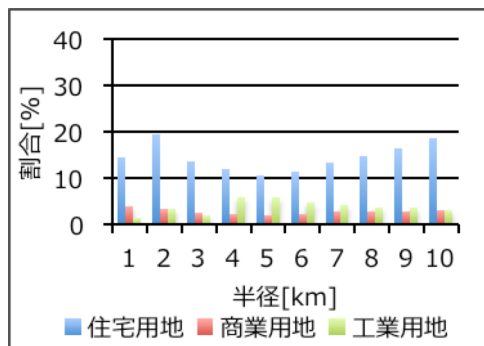


図-4 白井の現況土地利用

(2) トリップ長の実態分析

土地利用現況図からトリップ長構成に関係すると思われる住宅用地、商業用地、工業用地について集計する。集計方法は交差点を中心に半径 1km から

10km の範囲で集計した。呼塚と白井の土地利用の集計結果を図-3, ならびに図-4 に示す。

呼塚では、白井と比較して住宅用地と商業用地が多く、これにより様々な目的のトリップが発生することでトリップ長が混在していると推察される。また、白井では呼塚より住宅用地と商業用地の割合が少なく半径 5km 以上から工業用地が商業用地よりも多くなっていることから、短距離トリップが抑えられて 25km 前後のトリップが多くなっていると考えられる。

#### 4. トリップ長構成の推定

##### (1) 混合正規分布モデル

交差点毎のトリップ長構成の推定を行う。前節にて示したトリップ長分布が多峰型分布であると考えられるため、混合正規分布モデルによる推定を試みた。混合正規分布モデルは、単純な分布(要素モデル)を混合して複雑な分布を記述する混合分布の一種であり、要素モデルに正規分布を仮定したものである。

4) 正規分布の確率密度関数は以下のとおりである。

$$\phi(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

ここで、 $x$  はデータ、 $\mu$  平均、 $\sigma^2$  は分散である  $M$  個の正規分布を要素モデルとする確率密度関数は以下のようになる。

$$p(x; \theta) = \sum_{m=1}^M \pi_m \phi(x; \mu_m, \sigma_m) \quad (2)$$

パラメータ  $\pi_m$  は要素モデルがデータを生成するために選ばれる確率(混合比)を表しており、以下の制約を満たす。

$$\pi_m \geq 0 (m = 1, \dots, M), \quad \sum_{m=1}^M \pi_m = 1 \quad (3)$$

混合分布はどの要素モデルから生成されたかという情報に関して式の同時分布を周辺化したものと考えられることができる。ベイズの定理を用いれば、データ  $x$  が観測されたとき、それが要素モデル  $m$  から生成された事後確率は以下の式によって計算できる。

$$p(m|x; \theta) = \frac{p(x, m; \theta)}{p(x; \theta)} = \frac{\pi_m \phi(x; \mu_m, \sigma_m)}{\sum_{m=1}^M \pi_m \phi(x; \mu_m, \sigma_m)} \quad (4)$$

この事後確率を用いてデータが生成されたと考えられる要素モデルを決定する。

$n$  個のデータが観測された時、混合正規分布モデルの最尤推定量は以下の式で定義される。

$$\begin{aligned} \hat{\theta} &= \operatorname{argmax}_{\theta} \sum_{i=1}^n \log p(x_i; \theta) \\ &= \operatorname{argmax}_{\theta} \sum_{i=1}^n \log \sum_{m=1}^M \pi_m \phi(x; \mu_m, \sigma_m) \end{aligned} \quad (5)$$

式(5)では対数の中に和があるため推定量は陽に推定できない。もし、データ  $x_i$  を生成した要素モデル  $m_i$  がわかっていると、最尤推定量は以下のように陽に記述できる。

$$\begin{aligned} \hat{\theta} &= \operatorname{argmax}_{\theta} \sum_{i=1}^n \log p(x, m; \theta) \\ &= \operatorname{argmax}_{\theta} \sum_{i=1}^n \log \pi_{m_i} \phi(x; \mu_{m_i}, \sigma_{m_i}) \end{aligned} \quad (6)$$

しかし、入力されるデータには生成した要素モデルを示すラベル  $m_i$  が欠損しているため、式を適用できない。そこで、欠損している不完全データを完全データのフレームワークで逐次的に解決する EM アルゴリズムを用いる。

##### (2) EM アルゴリズム

EM アルゴリズムは、期待値を計算する E-step と最大化を行う M-step で構成されており、この 2 つを反復してパラメータを推定する。まず、暫定的に初期値  $\theta^{(0)}$  を与える。E-step では現在のパラメータ  $\theta^{(0)}$  のもとで尤度関数の条件付き確率に関する期待値である Q 関数を計算する。最期に M-step では E-step で計算した Q 関数を最大化するパラメータ  $\theta^{(1)}$  を非線形最適化等を用いて求める。その上で、初期値  $\theta^{(0)}$  を  $\theta^{(1)}$  に更新し、収束条件を満たすまで 2 つの step を反復することで、 $\theta$  を推定する。

具体的には、E-step は式(7)、M-step は式(8)のようになる。

$$Q(\theta, \theta^{(t)}) = \sum_{i=1}^n \sum_{m=1}^M p(m|x_i; \theta^{(t)}) \log p(x_i, m; \theta) \quad (7)$$

$$\theta^{(t+1)} = \operatorname{argmax}_{\theta} Q(\theta, \theta^{(t)}) \quad (8)$$

なお、EM step が 1 回終了するごとに  $t=t+1$  とする。

##### (3) 混合正規分布モデルの留意点

EM アルゴリズムの収束値は極大値であるが、全域的な最大値となるとは限らないため、異なる初期値から幾つかの推定値を求め、最も尤度が高いものを用いる。また、混合正規分布モデルを推定する際は、用いる要素モデルの数(混合数)が問題となる。要素モデルの数が多ければ、データへの当てはまりはよくなるが、over fit が懸念される。分布の特徴を十分に捉えられる範囲で可能な限り少ない混合数で推定できることが望ましい。

本研究では、初期値  $\theta^{(0)}$  を平均が 1 から 100 の間での無作為抽出し、分散を全て要素モデルに対して 1、混合比は 1 を要素モデルの数で除した数値として、合計 100 通りの初期値で計算を行い、最も尤度が高い  $\theta$  を採用した。また、混合数が 1 から 10 の各ケースでそれぞれモデルを推定し、最も BIC(ベイズ情報量規準)が低い混合数を採用している。

(4) 推定結果と考察

本研究では、100km 以上のトリップ数が全体の 10% 以下だったため、推定対象から除外して 100km 未満のトリップで推定を行っている。表-1 に推定したパラメータを示す。全ての交差点で相関係数が 0.9 以上と当てはまりは良好である。混合数は白井, 小淵, 野田市駅入口で 6 となり、柏市内にある呼塚, 若柴は 7 となった。また、混合正規分布モデルを用いて推定した呼塚と白井のトリップ長構成分布を図-5 ならびに図-6 に示す。呼塚と白井の両者ともに相関係数が 0.9 以上と当てはまりは良好だが、呼塚では 35km, 白井では 24km から 26km で見られるような急激な変化は正確には推定できてはいない。

表-1 から白井, 小淵, 野田市駅入口の 5, 6 番目の要素モデルと呼塚, 若柴の 6, 7 番目の要素モデルでは、パラメータに差が見られず、トリップ長が 90km 以降のトリップは全ての交差点を通過していると推察され、沿道土地利用にあまり影響されないと考えられる。また、柏市内にある呼塚, 若柴の混合数が 7 つであるのは、交差点間の距離が近いため土地利用の変化があまり起きていないからと推察される。また、図-5 と図-6 より、0km から 40km の範囲では、呼塚は 4 つの要素モデルで表現しているが、白井が 3 つで表現しており、土地利用の影響を受けているトリップ長の範囲は 0km から 40km と考えられる。表-1 より 2 つの 1 番目の要素モデルを比較すると、混合比  $\pi$  が大きく異なっているため、0km から 10km の範囲では混合比  $\pi$  が土地利用に影響されると思われる。

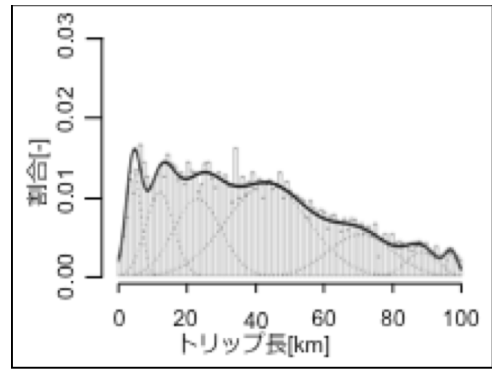


図-5 呼塚の推定トリップ長構成

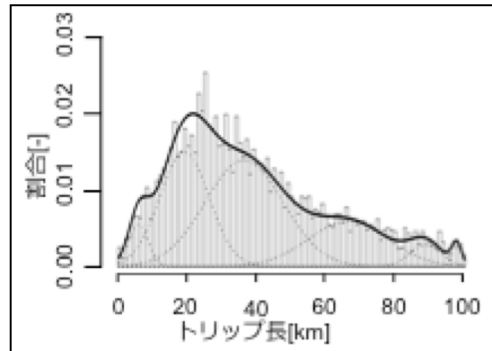


図-6 白井の推定トリップ長構成

5. おわりに

本研究の成果として、沿道土地利用がトリップ長構成に関係すること、混合正規分布モデルを用いれば高い精度でトリップ長構成の推定ができることが示された。さらに、各々の要素モデルのパラメータが推定できれば、より明確なトリップの属性分けが可能となる。今後は各々の要素モデルのパラメータを交差点付近の土地利用状況を変数とした構造化を検討したい。これにより、土地利用の改変施策が特定のトリップ長に増減に与える影響を評価すること

表-1 推定パラメータ

m	呼塚			白井			小淵			野田市駅入口			若柴		
	$\mu$	$\sigma$	$\pi$	$\mu$	$\sigma$	$\pi$	$\mu$	$\sigma$	$\pi$	$\mu$	$\sigma$	$\pi$	$\mu$	$\sigma$	$\pi$
1	5061	2304	0.093	6050	2828	0.042	4851	1469	0.028	4941	2255	0.057	4425	2193	0.084
2	11420	3920	0.134	19351	7197	0.294	13861	5318	0.216	13404	4673	0.155	12192	4141	0.124
3	23874	7251	0.238	36717	11337	0.429	30738	9629	0.349	27802	9232	0.291	23263	7040	0.194
4	44210	12031	0.351	67019	11637	0.188	57239	13709	0.335	55243	15529	0.419	43802	12129	0.397
5	71493	9596	0.133	89486	4663	0.036	85244	6531	0.058	87276	5918	0.062	71272	9548	0.144
6	90020	4599	0.040	98425	1450	0.011	97814	1812	0.014	97448	1879	0.016	89236	4694	0.041
7	98660	1461	0.010										97963	1837	0.014
相関係数	0.967			0.955			0.923			0.918			0.955		

が可能になると期待される。しかし、現時点ではデータ数が少なく、混合数も交差点によって異なるため構造化を行うためには工夫が必要である。

#### 参考文献

- 1) 中村 英樹,大口 敬:性能照査型道路計画設計の導入に向けて,土木学会論文集 D3(土木計画学), Vol.67, No.3, 195-202, 2011.
- 2) 内海 泰輔,中村 英樹,大口 敬:性能照査型道路計画設計における設計交通量の考え方,土木計画学研究・講演集 No.43, CD-ROM, 2011.
- 3) 国道 16 号や周辺道路の現状:  
[http://www.ktr.mlit.go.jp/chiba/kashiwa/chiba\\_kashiwa/pdf/v3chiikinogenjou1.pdf](http://www.ktr.mlit.go.jp/chiba/kashiwa/chiba_kashiwa/pdf/v3chiikinogenjou1.pdf)
- 4) 金森敬文, 竹之内高志, 村田昇: R で学ぶデータサイエンス 5 パターン認識, pp.36-49, 共立出版, 2009
- 5) 葛西 誠,小田 崇徳:一般道サービス水準の考え方と計測法に関する論点整理,土木計画学研究・講演集 No.47, CD-ROM, 2013.
- 6) 内海 泰輔,中村 英樹,磯和賢一,渡辺将光:機能に対応した道路計画設計のための交通量変動特性分析,土木計画学研究・講演集 No.33, CD-ROM, 2006.
- 7) 渡辺将光,内海 泰輔,中村 英樹:年間を通じた時間交通需要変動の再現手法の構築,土木計画学研究・講演集 No.34, CD-ROM, 2006.
- 8) 内海 泰輔,中村 英樹,大口 敬:性能照査型道路計画設計の考え方と検討課題,土木計画学研究・講演集 No.35, CD-ROM, 2007.

(2017. 4. 28 受付)

## ESTIMATION OF THE TRIP LENGTH CONSTITUTION BY THE GAUSSIAN MIXTURE MODEL FOR A TRAFFIC-PERFORMANCE EXAMINATION METHOD

Masataka YAMAMOTO, Hideki YAGINUMA, Shintaro TERABE,  
and Nan KANG