

幹線道路リンクごとに集計されたプローブカー旅行時間データの特徴分析*

Analysis of Characteristic of Travel Time Using Arterial-link-based Probe-Car Data*

李 強**・姜 美蘭***・森川 高行****

By Qiang Li**・Meilan Jiang***・Takayuki MORIKAWA****

1. はじめに

近年、プローブカーデータを活用した旅行時間予測に関する研究が注目されている。研究の予測対象としては、特定高速道路区間や都市内任意トリップなどがある。その中で都市内トリップの旅行時間予測に主な役割をする蓄積データは、一般的に(財)日本デジタル道路地図協会(DRMA)のデジタル道路地図(以下 DRM)リンクごとに集計されている¹⁾。DRM リンクは短いため、集計された旅行時間データは、信号の影響を強く受け、赤信号で止まったデータと青信号でスムーズに通過したデータに明確に分けられている。即ち、DRM リンクごとに集計されたデータは正規分布に当てはまらない。従って、蓄積データを利用して区間旅行時間を予測する時、旅行時間の区間推定が不可能になっている。蓄積データを、もっと長い距離、例えば幹線道路リンクごとに集計すると信号の影響を減少し、データが正規分布に近づくと考えられる。

本研究では、プローブカーデータを、DRM リンクより長い幹線道路リンクごとに集計し、正規性の検討と時系列モデリングの適用性を検討する。

2. データの概要

本研究で集計単位とした幹線リンクは、名古屋市の桜山交差点～御器所交差点(長さ:956 m)と御

*キーワード: プローブデータ、幹線道路、旅行時間

**学生員、工修、名古屋大学大学院工学研究科

***正員、工博、名古屋大学環境学研究科

****正員、Ph.D、名古屋大学環境学研究科

(名古屋市千種区不老町、

TEL052-789-3730、FAX052-789-3565)

器所交差点～吹上交差点(長さ:948m)の二つのリンクである(図 1)。



図 - 1 研究対象幹線リンク

プローブカーデータは、名古屋都市圏において行われた「インターネット ITS プロジェクト」の実証実験により取得された 2002 年 10 月～2003 年 3 月のプローブタクシーデータである²⁾。一般的にタクシーの行動は、乗客を乗せた場合(実車)と乗せていない場合(空車)は異なるため、本研究では実車のデータのみトリップとして分析を行う。抽出されたトリップの概要(表-1)を見ると、幹線道路リンクを完全通過するトリップ数は、部分通過するトリップ数の1/2で、幹線道路リンクごとに集計してもサンプル数の減少は大きくない。

表 - 1 トリップの概要

区 間		トリップ数
桜山～御器所	完全通過	9609
	部分通過	16885
御器所～吹上	完全通過	12033
	部分通過	20231
桜山～御器所と 御器所～吹上	完全通過	7083
	部分通過	23903

データクリーニングとして、ルートマッチング(Route Matching)が間違ったトリップを取除いて

いる。取除いたトリップは2167個である。

3. 幹線リンクごと集計されたデータの統計分析

(1) リンクコストテーブル

従来、都市内任意トリップの旅行時間を推計するために、リンクコストテーブルという蓄積手法を利用していた。リンクコストテーブルは、DRMリンクの旅行時間を曜日別、時間帯別(5分間隔)に集計し、集計したサンプルから平均と分散を計算して作成された。しかしサンプルの分布が分からなく、分散の信頼性が非常に低いため、トリップの旅行時間予測には、平均のみ利用し、区間予測はできなかった。

本研究では、抽出したプローブカートリップから得られた幹線道路リンク旅行時間を曜日別、時間帯別(15分間隔、一日96区分)に集計している。祝日は日曜日としている。

図-2は、火曜日に桜山~御器所を完全通過するトリップから計算した時間帯別旅行時間と上下側4分位の差である。

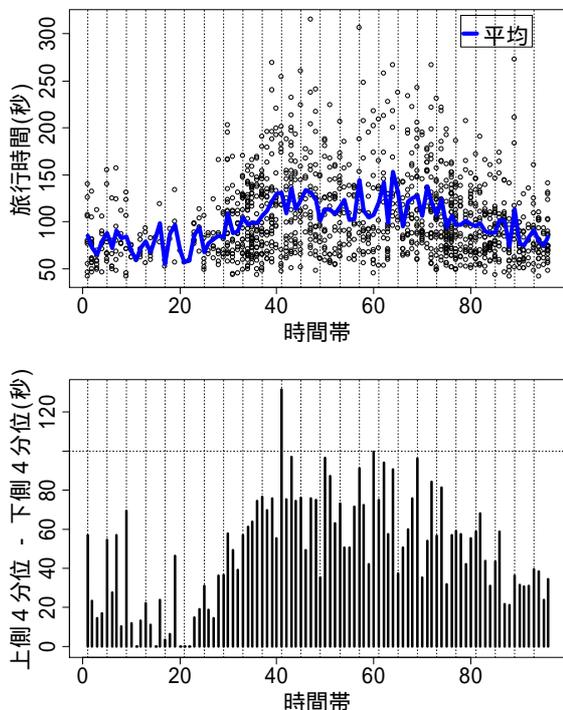


図-2 桜山~御器所時間帯別旅行時間(上)
上下側4分位差(下)(火曜日)

図-2を見ると、ピーク時間帯にも約50秒(速度70km/h)の短い旅行時間がある。これは、同じ曜日に

も、普段より非常に低い交通量を持つピーク時間帯が存在することを示している。即ち、交通状況の週変動はかなり大きいことを示している。サンプル平均値から推定すると、朝ピーク時は10:00~11:00(時間帯:41~44)、夕方ピーク時は16:00~17:00(時間帯:65~68)になり、サンプル平均値は交通状況の時間帯変化を正確に反映しているとは言えない。また上下側4分位差は80~100秒程度であり、約100秒の平均値に比べると大きすぎる。リンク旅行時間は走行時間と停止時間に分ける。その中、走行時間は交通状況の影響を大きく受けるが、停止時間は渋滞の時のみ交通状況の影響を大きく受ける。即ち、旅行時間は、大きいばらつきを持つ停止時間の影響で、交通状況の変動パターンを正確に反映することができない。

幹線道路リンクごとに集計する手法のデメリットは、各リンク各時間帯のサンプル数が少なくなることである。本研究の対象リンクでは、7:00~22:00までの各時間帯では(時間帯:29~88)サンプル数が10個以上で、ピーク時間帯には更に多くなっている(図-3)。そのため、分析に十分なサンプル数が得られている。

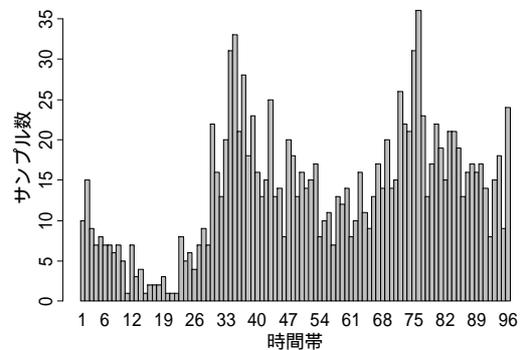


図-3 桜山~御器所時間帯別サンプル数(火)

(2) 正規性の検討

図-4、5は、火曜日桜山~御器所の時間帯32(7:45)と35(8:45)の旅行時間分布である。青い線は、サンプルと同じ平均と分散を持つ正規分布の確率密度関数のグラフである。図-4の旅行時間分布は正規分布に近づいているが、図-5は、二つのピークがある分布である。別の時間帯にもサンプル数が多く二つのピーク分布になる傾向が多い。これは交通信号の影響が主な原因であると考えられる。

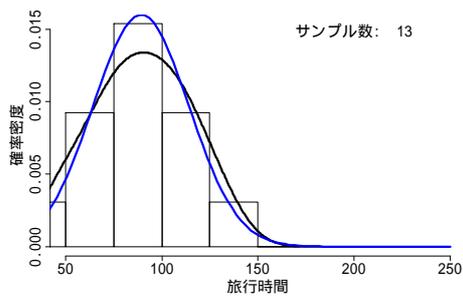


図 - 4 桜山～御器所旅行時間分布(時間帯：32、火)

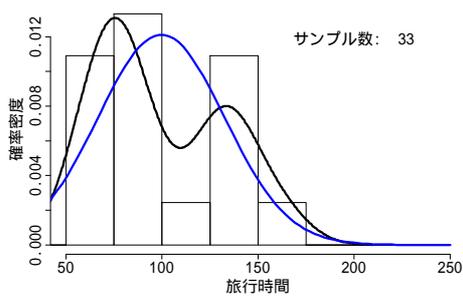


図 - 5 桜山～御器所旅行時間分布(時間帯：35、火)

(3) 隣接リンクの相関

図 - 6 は、二つの隣接幹線リンク（桜山～御器所と御器所～吹上）を完全に通過したデータ（火曜日、8時から9時まで）を利用した旅行時間プロット図である。隣接幹線リンクの旅行時間の相関係数は0.465で低い相関を示している。これは隣接幹線リンクの異なる交通状況と交通信号の共同影響であると思われる。

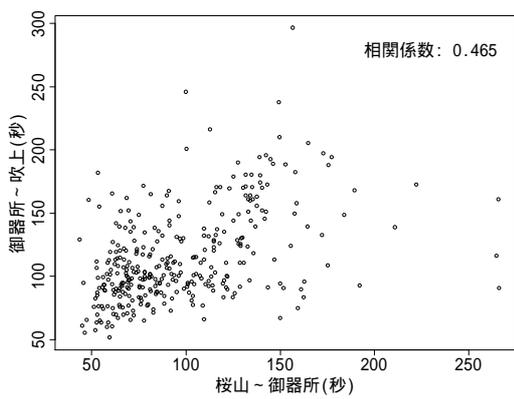


図 - 6 隣接リンクの相関(8:00～9:00、火)

4. 時系列モデルの適用性

(1) プローブ時系列

本研究では、平日（7時から22時）に桜山～御

器所を通過したプローブデータから、表 - 2 のようなプローブ時系列を作成する。同日、同時帯に1個以上のデータがある時平均値を取り、1個もない時欠損値として「NA」と表記する。

表 - 2 プローブ時系列

時間帯 日付	29	30	31	...	88
2002/10/15 (火)	NA	115.1	NA	...	NA
2002/10/16 (水)	92.9	NA	NA	...	83.4
2002/10/17 (木)	134.2	87.25	66.6	...	NA
2002/10/18 (金)	79.6	NA	79.1	...	100

プローブ時系列は、一般の時系列と比べると、重なる周期（年、月、週、日）、多い欠損値、祝日による周期の崩れなどがある。そして、本研究では欠損値が少ない平日の7時から22時のデータだけで時系列を構成している。

(2) 欠損値の処理

2002年10月から2002年3月までの平日121日の中、欠損値の数が15以下の日は32日である。この32日のデータを利用し、曜日変動は考えなく、時間帯変動だけに注目して時系列モデルの適用性を検討する。データの欠損値は、時間帯平均で補完している。

(3) 周期成分の除去

本研究では、連続していない32日のデータを分析しているため、トレンド成分はないと仮定すると、単純なモデル

$$X_t = s_t + Y_t \quad (1)$$

を用いることができる。ただし、 s_t は周期60（時間帯：29～88）を持つ t の周期的な関数として

$$s_t = a_0 + a_1 \sin((2p/60)t) + a_2 \cos((2p/60)t) + a_3 \sin(2*(2p/60)t) + a_4 \cos(2*(2p/60)t) \quad (2)$$

により与えられる調和関数の和である。 Y_t は、周期成分を除去したあとの残差である。

図 - 7 は、最初の3日間の旅行時間と周期成分（赤線）である。横軸は3日間180個時間帯である。

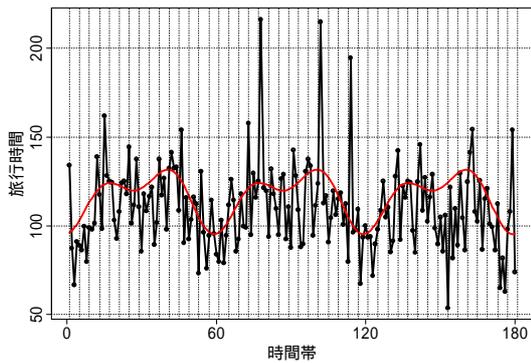


図 - 7 旅行時間と周期成分

(4) 残差と独立性

この3日間の残差と標本自己相関関数(ACF)(図-8)には、有意な依存性が見られない。

もし、残差の間に依存性がなければ、残差は独立な確率変数の観測値で、平均と分散を推定する以外にモデリングする必要がないと考えられる。

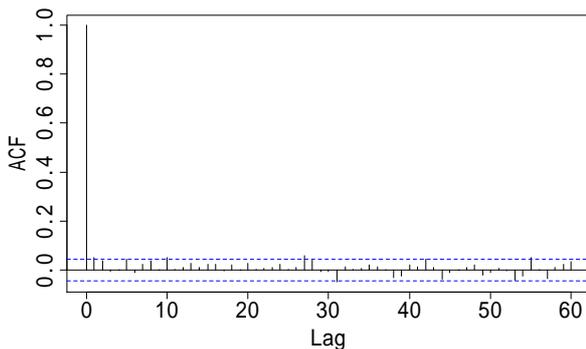
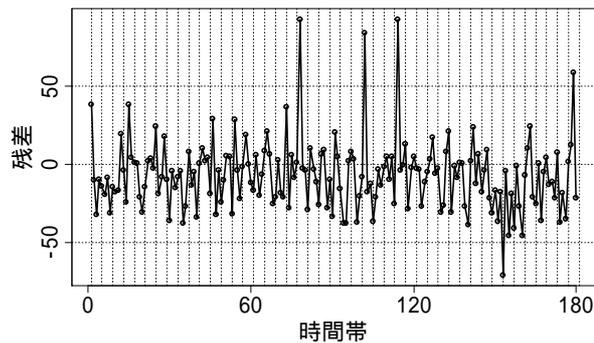


図 - 8 残差とACF

4. おわりに

本研究では、プローブカーから取得した旅行時間データを DRM リンクごとではなく、もっと長い幹線道路リンクごとに集計して分析した。各リンク各時間帯の旅行時間データは、長いリンクごとに集計しても信号の影響を受け、二つのピークの分布で正規

分布ではないことが分かった。また、プローブ時系列から調和回帰で周期成分を除去した残差は、有意な依存性がなく、複雑な時系列モデルで分析する必要がないことを示した。有意な依存性がない原因としては、交通状況変動の不安定性や信号の影響などが考えられる。今後の課題として、プローブカーデータによる旅行時間を交通状況の影響を大きく受ける走行時間と渋滞の時のみ交通状況の影響を受ける停止時間に分けて分析し、交通状況変動を正確に反映する要因の抽出などがあげられる。

参考文献

- 1) 三輪富生・倉内慎也・森川高行：意思決定タイミングを内生化した動的な経路選択モデルの構築，土木計画学研究・講演集，Vol.30，CD-ROM，2004
- 2) 三輪富生・境隆晃・森川高行：プローブカーデータを用いた経路特定手法と旅行時間推定に関する研究，第2回 ITS シンポジウム proceedings，pp.277-282，2003．