

ETCデータを用いた旅行時間信頼性の予測方法に関する研究

福田 大輔¹・伊藤 愛実²・

¹正会員 東京工業大学大学院准教授 理工学研究所土木工学専攻 (〒152-8552 目黒区大岡山 2-12-1)

E-mail: fukuda@plan.cv.titech.ac.jp

²非会員 興和不動産株式会社 [元東京工業大学学生]

本研究では、旅行時間信頼性指標の予測方法の確立に資するべく、東名高速道路における大量のETCデータを用いて、既存モデルの改善と新たなモデル構築という二つの視点から分析を行った。旅行時間信頼性の規定要因を探るための回帰分析の結果、旅行時間信頼性に影響を及ぼす幾つかの要因の存在を明らかにした。次に、隣接リンク間での旅行時間信頼性指標の統合の可能性について検討するため、旅行時間共分散の規定要因に関する検討を行い、主として連続する上下流の交通量の大小関係に存在している可能性があることを明らかにした。

Key Words : Travel Time Reliability, Travel Time Variability, Forecasting, Regression Analysis, ETC Data

1. はじめに

近年、経済活動の高度化や生活水準の向上に伴い、移動においても単なる速達性のみならず、旅行時間の正確さが求められるようになってきている。道路投資の費用対効果分析においても、旅行時間の短縮のみを考慮するのではなく、旅行時間の定刻性、すなわち、旅行時間信頼性を考慮した検討の重要性が増すと考えられる。しかし、我が国の費用対効果分析では、平均時間短縮便益、走行費用節約便益、交通事故削減便益の三項目のみが便益として計上されている。旅行時間信頼性を考慮する事により、遅刻リスクを小さくするための安全余裕時間の減少や、早着・遅刻による損失ペナルティの減少効果が便益として評価できるようになる。

このような旅行時間信頼性の経済評価を行うためには、限界の価値である旅行時間変動価値の推計^{1),2)}に加えて、事業実施に伴って将来の旅行時間信頼性がどのように変化するかを予測する手法の確立が必要である。道路条件、環境条件、交通条件等に応じて旅行時間信頼性を予測することができれば、旅行時間信頼性便益導入の実適用に資すると期待される。旅行時間信頼性の将来予測に関しては、標準偏差を旅行時間信頼性指標(非説明変数)として、説明変数に曜日や季節、天気、平均旅行時間を用いて旅行時間信頼性の予測式を構築している研究³⁾や、変動係数を被説明変数として平均旅行時間、自由旅行時間を用いた予測モデル⁴⁾⁻⁶⁾など、いくつかの経験式が考察されている。しかしながら、一般的法則の解明のためにはさらなる分析の蓄積が必要である。

また、分散や標準偏差などの旅行時間信頼性指標を求めるためには、旅行時間情報を長期にわたって収集する必要があるが、一般に、対象とするトリップ長が長くなるほど指標算出に必要な旅行時間データのサンプル数は少なくなる。このような場合に、比較的サンプル数の多いリンクの足し合わせによってより長いトリップでの旅行時間信頼性の評価が可能となれば、実適用の可能性が高まることが期待される^{7),8)}。

そこで本研究では、ETC (Electric Toll Collection) のデータを用いて、旅行時間信頼性の規定要因を統計学的に検証し、その将来予測モデルの構築を行うことを第一の目的とする。次に、推計された区間ごとの旅行時間信頼性の指標から対象区間全体の旅行時間信頼性を推計するための方法論の検討を行うことを第二の目的とする。

2. 用いるデータの概要

(1) ETC データ

本研究で用いるデータは、ETC を搭載した車が高速道路のためインターチェンジ (IC) に流入・流出した時に記録される情報である。対象道路区間は東名高速道路の東京 IC から三ヶ日 IC までとする。この区間の総 IC 数は 25 個であり、入口 IC と出口 IC の組み合わせは ${}_{25}C_2 \times 2 = 600$ 通りとなる。対象期間は 2008 年 4 月 1 日から 2010 年 3 月 31 日までの休日・祝日の延べ 238 日分である。条件抽出やクレンジングを行う前のフルデータにおける総サンプル数は 54,672,465 である。

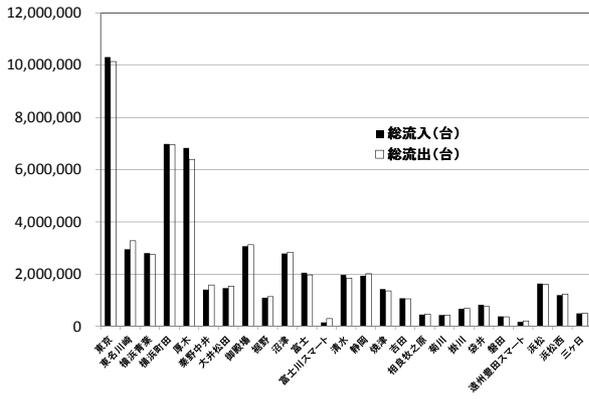


図 - 1 インターチェンジ別の流入人数

ICごとの流入台数をまとめたグラフが図 - 1である。基礎分析を踏まえて、分析対象区間を富士川スマートICを除く東京IC-静岡IC間とし、分析対象車種を軽自動車等・普通車に限定した。また、分析時間帯を朝5時から夜の11時までとした。

(2) データのクレンジング方法

本研究ではETCデータの流入時刻と流出時刻の差分により各車両の旅行時間を算出しているため、SAやPAでの休憩や混雑時の路肩走行によって、旅行時間の傾向が他車とは明らかに異なる外れ値が含まれる。図 - 2に2009年11月21日の東京IC-厚木IC間の流入時刻(時)と旅行時間(分)の散布図を例として示すが、多くの外れ値の存在がこの図からも伺える。特に、標準偏差は外れ値の存在に対して頑健ではないため、旅行時間信頼性の分析に先立ってデータのクレンジングが必要であることが確認される。ただし、単純にデータ獲得期間である約2年分のデータ全体の標準偏差や平均旅行時間を使って外れ値をクレンジングすると、本来残すべき渋滞による旅行時間の遅れのデータも取り除かれてしまうことが危惧される。そこで山崎⁹⁾を参考に、日ごとの旅行時間移動平均と移動標準偏差を用いて、外れ値のクレンジングを試みる。

まず、各1日24時間のデータを用いて算出した平均 μ と標準偏差 σ を用いて、 $\mu + 3\sigma$ 以上の値を外れ値として取り除いた後に、改めて移動平均 μ' と移動標準偏差 σ' を求め、 $\mu' - \sigma'$ 以下、または $\mu' + 2\sigma'$ 以上の値を外れ値として取り除く手順をとった。この方法では、あらかじめ $\mu + 3\sigma$ 以上の値を取り除いておくことで、移動平均と移動標準偏差の変動を小さくし、より適切に値が取り除けるように配慮したものである。図 - 3はクレンジング後の散布図である。図 - 2で存在していた外れ値が適切に取り除かれていることがわかる。こ

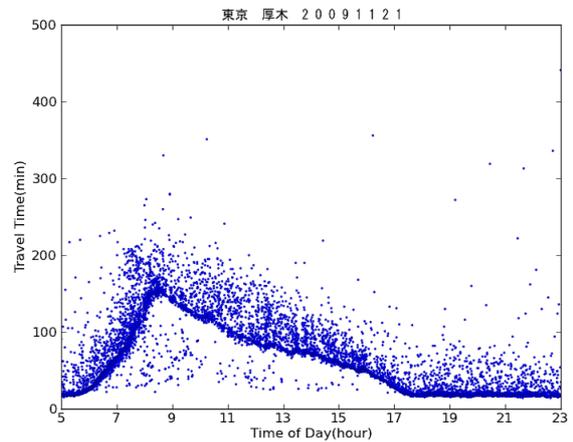


図 - 2 流入時刻-旅行時間散布図(東京IC-厚木IC間, クレンジング前)

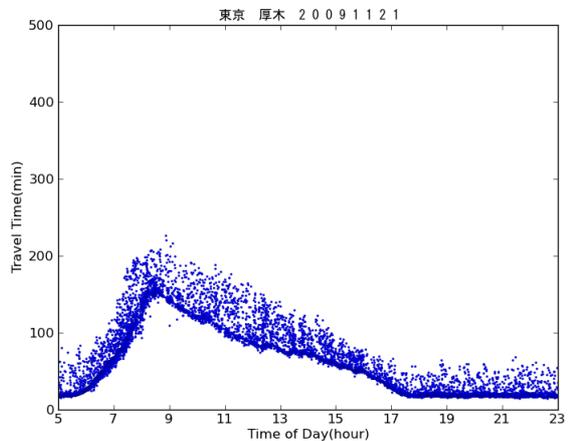


図 - 3 流入時刻-旅行時間散布図(東京IC-厚木IC間, クレンジング後)

の方法を、全ICペア全日数に対して適用し、ETCデータのクレンジングを行った。その結果、元々のサンプル数54,672,465(100.0%)に対し、以降の分析で用いるサンプル数は30,051,311(55.0%)となった。

3. 旅行時間信頼性の予測モデルの構築

旅行時間信頼性の予測モデルの構築には重回帰分析を用い、モデルのパラメータ推定においては通常最小二乗法を用いる。推定には、クレンジング済みのデータを用いてICペア毎に旅行時間信頼性指標を算出し、被説明変数として用いる。その際、IC間の距離が短く、妥当な値を示していないと判断された東京IC-東名川崎IC間のデータは除くこととした。したがって重回帰分析のサンプルとなるICペアの数は $({}_{13}C_2 - 1) \times 2 = 154$ となる。被説明変数には標準偏差を平均旅行時間で割っ

た値である変動係数を用い、説明変数には SA・PA の数、平均旅行時間、交通量など、旅行時間変動を規定する要因として考えられるものを探索的に検討する。

(1) モデル A

はじめに、式 (1a) の両辺の対数をとった予測式 (1b) をモデル A として検証する。

$$CV_t = \alpha CI_t^\beta d^\gamma \quad (1a)$$

$$\log(CV_t) = \log(\alpha) + \beta \log(CI_t) + \gamma \log(d) \quad (1b)$$

CV_t : 時間帯 t における変動係数

CI_t : 時間帯 t における渋滞指数

d : OD 間距離 (km)

α, β, γ : 未知パラメータ

式 (1a) はイギリスの 10 大都市地域の 34 ルート (複数リンクで構成される OD ペア) を対象に、プローブデータから算定した旅行時間を用いて構築されたものに準拠している⁵⁾。説明変数として渋滞指数と OD 間距離を用いているが、ここで渋滞指数とは、平均旅行時間/自由旅行時間で定義される無次元指標である。本研究において、自由旅行時間は IC 間距離 × 法定速度より算出した。

まず、式 (1b) が本研究のデータに対しても有効であるか検証した上で、必要があればさらに説明変数を加えて重回帰分析を行い、予測式の精度向上を試みる。本研究のデータを用いて最終的に得られた推定結果が表 - 1 である。なお、以降の推定結果において、決定係数とは全て自由度調整済み決定係数を指す。式 (1b) をそのまま用いた場合の決定係数が 0.609 であったのに対して、規定要因を加えたモデル A の決定係数は 0.646 となっており、大幅な改善がみられる。トンネルや法定速度などに関連する変数は有意とはならなかったが単位距離あたりの SA・PA の数は t 値が 2.73 となり、説明変数の中で最も有意となった。

(2) モデル B

次に、モデル A 同様に式 (2a) の両辺の対数をとった式 (2b) をモデル B として分析を行う。式 (2a) はストックホルム中心部周辺のいくつかの一般道路を対象に、道路の各リンクで流入出した車のナンバープレート認識によって算出された旅行時間を用いて構築された変動係数の増加・減少の程度 of 非線形性を考慮した予測式に準拠したものである⁴⁾。本研究で用いるデータに対して、このモデル B を用いた推定結果が表 - 2 である。

$$CV_t = \exp(\alpha + \beta(CI_t - 1) + \gamma(CI_t - 1)^3) \quad (2a)$$

$$\log(CV_t) = \alpha + \beta(CI_t - 1) + \gamma(CI_t - 1)^3 \quad (2b)$$

決定係数は 0.449 であり、モデル A ほどの適合度は得られていない。また、モデル A の場合と同様に、説明変数を様々な形で加えたが、決定係数が 0.5 を越えることはなかった。この理由としては、本研究が都市間高速道路の比較的長距離な区間を対象としているのに対して、モデル B はストックホルムの比較的狭いエリアを分析対象としていることが考えられる。

(3) モデル C

モデル A、モデル B の結果を踏まえ、道路特性を考慮した線形のモデル C を新たに構築する。式 (3) は旅行時間変動を規定する要因として考えられるものを探索的に適用した中で最も精度が良い予測式であり、推定結果を表 - 3 に示す。ここで平均遅れ割合とは“1 - 1/渋滞指数”で表される無次元指数である。

$$CV_t = \alpha N + \beta V + \gamma E + \delta P \quad (3)$$

N : SA・PA の数 (個/km)

V : 交通量 (台/日・km)

E : 緊急交通規制回数 (回/年・km)

P : 平均遅れ割合

$\alpha, \beta, \gamma, \delta$: 推定パラメータ

(4) 総合考察

全てのモデルを通じて、トンネルの数やその延長は統計的に有意でなかった。その理由としては、含まれる区間が少ないことや、都市近郊の混雑エリアにトンネルが含まれるエリアがないことが考えられる (OD 間にトンネルが含まれるか否のダミー変数も有意ではなかった)。また、車線数が有意とならない理由は、交通量が多い区間ほど車線数は多く整備されており、車線数が少ないからといって混雑を引き起こす原因とはならないためであると考えられる。さらに法定速度に関しては、法定速度が低めに設定されている箇所は首都圏ではなく、郊外部で比較的スムーズかつ定常に交通が流れている箇所になる。

モデル全体の適合度を見ると、モデル C の決定係数は 0.936 となっており、3 つのモデルの中で最も精度の高い値となっている。特に、平均遅れ割合の t 値が高くなっており、自由旅行時間に対する平均旅行時間の遅れの割合が、旅行時間信頼性に与える影響が大きいことが示唆される。本研究のように高速道路を分析対象としたデータは、外部からの影響が少ないため、交通渋滞が旅行時間信頼性に与える影響が、一般道路を対象とした場合に比べて大きいと考えられる。したがって、モデル C のような平均遅れ割合、交通量などを線形的に考慮した予測式の精度が比較的良くなったものと考えられる。

表 - 1 モデル A の推定結果

	係数	標準偏差	t 値
$\log(\alpha)$	-1.179	0.191	-6.15
β	1.895	0.146	12.97
γ	-0.103	0.052	-1.97
SA・PA の数 (個/km)	1.953	0.715	2.73
交通規制数 (回/年・km)	0.003	0.001	1.81
交通量 (台/日・km)	0.007	0.004	1.66
サンプル数	154	決定係数	0.646

表 - 2 モデル B の推定結果

	係数	標準偏差	t 値
α	-1.551	0.062	-25.12
β	1.447	0.146	9.88
γ	-0.178	0.045	-3.89
サンプル数	154	決定係数	0.449

表 - 3 モデル C の推定結果

	係数	標準偏差	t 値
SA・PA の数 (個/km)	0.703	0.351	2.00
交通量 (台/日・km)	0.064	0.018	3.52
交通規制数 (回/年・km)	0.074	0.040	1.83
平均遅れ割合	0.911	0.112	8.12
サンプル数	154	決定係数	0.936

4. 旅行時間信頼性指標のリンク間での統合の可能性

(1) 旅行時間のリンク間相関

例として、ここでは図 - 4 のような直列 2 区間の道路ネットワークを考える。A, B, C は IC の位置に相当する。上流側 AB 間の旅行時間と、下流側 BC 間の旅行時間との間に相関関係があると仮定すると、AC 間を通しての総旅行時間の分散は、定義より式 (4) のように表される。

$$\sigma_{AC}^2 = \sigma_{AB}^2 + \sigma_{BC}^2 + 2Cov(T_{AB}, T_{BC}) \quad (4)$$

$\sigma_{AC}^2, \sigma_{AB}^2, \sigma_{BC}^2$: 各区間の旅行時間の分散

T_{AB}, T_{BC} : 各区間の旅行時間

Cov : 旅行時間の共分散

式 (4) を共分散について解くと式 (5) のようになる。

$$Cov(T_{AB}, T_{BC}) = \frac{\sigma_{AC}^2 - \sigma_{AB}^2 - \sigma_{BC}^2}{2} \quad (5)$$

旅行時間の観測を比較的容易に行えるのは、データの蓄積が豊富である AB 間, BC 間のような短い区間である。一方 AC 間のような比較的長い区間は、サン

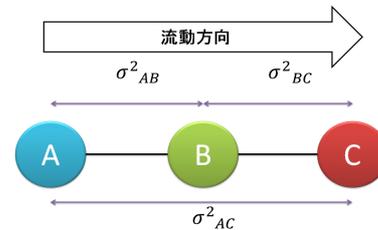


図 - 4 直列道路ネットワークと旅行時間の共変動

表 - 4 旅行時間共分散を被説明変数とした重回帰分析結果

	係数	標準偏差	t 値
交通量 a	-13.509	5.578	-2.42
交通量 a × 交通量 b	3.472	1.537	2.26
サンプル数	22	決定係数	0.165

ル数が少ないなどの理由により高い精度で旅行時間信頼性指標 (標準偏差や分散) を得ることが困難になる場合が多い。したがって、共分散 $Cov(T_{AB}, T_{BC})$ を何らかの間接的な方法で推計することが可能となれば、式 (4) より、既知の $\sigma_{AB}^2, \sigma_{BC}^2$ 並びに $Cov(T_{AB}, T_{BC})$ の推計値を用いて σ_{AC}^2 を推計することが可能となる。このことは、より一般的には、リンク単位での旅行時間信頼性指標を統合して、経路単位での指標へと拡大できる可能性を示唆している。

(2) 旅行時間共分散の規定要因の統計的検証

東名川崎 IC から静岡 IC までの 11 区間において、隣接する二区間上下両方向での旅行時間共分散を被説明変数、各区間の交通量を説明変数とした重回帰分析によって、旅行時間共分散の規定要因の解明を試みる。

表 - 4 に適合度が最も高い重回帰分析の結果を示す。ここで交通量 a, 交通量 b とは、それぞれ、上流側, 下流側における交通量 (万台/日) を指す。

決定係数は 0.165 であり、また、サンプル数も少ないことから、将来予測に耐えうるだけの精度を持ったモデルであるとは言えないが、各変数は有意となっており、隣接した道路リンク間での旅行時間の共変動の規定要因を検討するには十分である。

仮に、この回帰分析結果を真と仮定して、共分散の正負の変換点を求める。回帰結果から構築される予測式より、下流側区間の交通量 b の値の大きさによってのみ共分散の正負が決まることがわかる。このときの交通量 b の境界値は、

$$\text{交通量 } b = 13.509 \div 3.472 \approx 3.89 \text{ (万台/日)}$$

となる。つまり、下流側の交通量 b が境界値である 3.89 (万台/日) という交通量より多い場合は共分散が正、

少ない場合は負の値をとることが示唆される．以上の結果を場合分けしてまとめると次のようになる．

a) 旅行時間共分散が正の場合

交通量 a , 交通量 b が増えるほど共分散の絶対値は大きくなる．つまり，もともと下流側リンクの交通量が相対的に大きく，下流側で発生した渋滞が，上流側にも影響を及ぼし，両区間で旅行時間が増加するような正の相関関係が発生していると解釈することができる．

b) 旅行時間共分散が負の場合

これは，下流側の交通量が相対的に小さい状況に相当し，このフェーズでは，交通量 a が増えるほど共分散の絶対値は大きくなる．つまり，上流側で発生した渋滞が下流側では解消し，上流側では旅行時間の増加が，一方で下流側では相対的に旅行時間が減少するというような負の相関関係が発生していると解釈することができる．また，交通量 a が少ない場合は，共分散の絶対値は小さな値をとる．つまり両区間は互いにあまり影響を与えず（相関の度合いが低い），円滑な交通流動が起こっていると解釈することができる．

5. おわりに

本研究では，旅行時間信頼性指標の予測方法の確立に資するべく，東名高速道路における大量の ETC データを用いて，既存モデルの改善と新たなモデル構築という二つの視点から分析を行った．まず，旅行時間信頼性の規定要因を探るための回帰分析を行った結果，旅行時間信頼性に影響を及ぼす幾つかの要因の存在を明らかにした．次に，隣接リンク間での旅行時間信頼性指標の統合の可能性について検討するため，旅行時間共分散の規定要因に関する検討を行い，主として連続する上下流の交通量の大小関係に存在している可能性があることを明らかにした．

今後は，クレンジング方法の妥当性の再検証，より精度の高く実用性のある旅行時間信頼性の予測モデルの構築，隣接する区間の旅行時間共分散をより高い精度で推計できるような方法論の開発等をさらに進める必要がある．

参考文献

- 1) 福田大輔: 旅行時間変動の価値付けに関する研究展望とプロジェクト評価への適用に向けた課題の整理, 土木計画学研究・論文集, Vol. 27, pp. 221–238, 2010a .
- 2) Fosgerau, M. and Fukuda, D.: Valuing travel time variability: Characteristics of the travel time distribution on an urban road, *Transportation Research Part*

C: Emerging Technologies, Vol. 21, No. 1, pp. 83–101, 2012.

- 3) Peer, S., Koopmans, C. C., and Verhoef, E. T.: Prediction of travel time variability for cost-benefit analysis, *Transportation Research Part A: Policy and Practice*, Vol. 46, No. 1, pp. 79–90, 2012.
- 4) Eliasson, J.: Forecasting travel time variability, In *European Transport Conference*, 2006.
- 5) U.K. DfT: Forecasting travel time variability in urban areas, *Technical Report*, 2007.
- 6) Eliasson, J.: A cost-benefit analysis of the Stockholm congestion charging system, *Transportation Research Part A: Policy and Practice*, Vol. 43, No. 4, pp. 468–480, 2009.
- 7) 福田大輔: 旅行時間変動のリンク間加法性に関する一考察, 土木計画学研究・講演集, Vol. 42, 2010b .
- 8) Engelson, L. and Fosgerau, M.: Additive measures of travel time variability, *Transportation Research Part B: Methodological*, Vol. 45, No. 10, pp. 1560–1571, 2011.
- 9) 山崎浩気: 変動を考慮した時間信頼性による高速道路サービス水準評価と利用者行動変化に関する研究, 博士論文, 京都大学, 2011 .