

解釈可能な機械学習モデルによる追従モデル パラメータと CD 現象の関係分析

山田 圭祐¹・甲斐 慎一郎²・和田 健太郎³

¹非会員 筑波大学大学院 理工情報生命学術院 システム情報工学研究群
(〒 305-8573 茨城県つくば市天王台 1-1-1)

E-mail: s2420483@u.tsukuba.ac.jp

²正会員 株式会社アイ・トランスポート・ラボ (〒 101-0052 東京都千代田区神田小川町 3-10)

E-mail: kai@i-transportlab.jp

³正会員 筑波大学准教授 システム情報系 (〒 305-8573 茨城県つくば市天王台 1-1-1)

E-mail: wadaken@sk.tsukuba.ac.jp (Corresponding Author)

本研究は、高速道路サグ部を対象とした追従モデルの入力パラメータとシミュレーションの結果得られるマクロな交通流特性の関係性を解釈可能な機械学習モデルを用いて分析する。具体的には、多数車両の追従挙動の集積として生じる代表的な集団現象である Capacity Drop (CD) 現象を分析する。本研究ではまず、Intelligent Driver Model (IDM) の改良版である IDM+に勾配の影響を加味したモデルのシミュレーションを網羅的なパラメータ条件下で実施する。次に、そこで得られたデータを学習データとして、入力パラメータから捌け交通量を予測する機械学習モデル(サロゲートモデル)を構築する。最後に構築した機械学習モデルを SHAP などの手法で解釈することにより、入力と出力の関係性を考察する。

Key Words: capacity drop, car following model, sag, machine learning

1. はじめに

高速道路サグ部において渋滞の発生後、捌け交通量が低下する Capacity Drop (CD) 現象が知られている(越ら¹⁾)。CD は、渋滞継続時間を長引かせ、結果として通行する車両群全体に多大な遅延をもたらす得る。

この現象に対し、Jin²⁾ および Wada et al.³⁾ は「地点依存の速度-車頭時間関係」と「有界加速度 (BA: Bounded Acceleration) モデル」を考慮した連続体交通流モデルによる説明を行っている。また、このモデルでは所与とする入力パラメータに対して CD の発生有無、捌け交通量が解析的に把握することができる。

しかしながら、このモデルはあくまで交通流を流体と捉え、モデル化したいわばマクロ交通流モデルに分類されるモデルであり、個別車両の走行挙動を表現することができない。また、個別車両の走行挙動を表現するマイクロモデルである追従モデルは、CD 現象は表現可能であるものの(例えば、Treiber et al.⁴⁾)。モデルパラメータと結果として実現する渋滞中交通流率(捌け交通流率)の関係は十分明らかになっていない。

本研究では、複雑な変数間関係性を自動で捕捉することの可能な機械学習アプローチを採用し、マイクロシミュレーションの車両の走行特性と CD 現象の関係

性を明らかにする。機械学習アプローチを採用する理由は、追従モデルを用いたマイクロシミュレーションでは各パラメータの関係性が非常に複雑となり、解析的にその関係性を捉えることが難しいためである。

本研究は以下に述べる三段階からなる。まず、甲斐ら⁵⁾ の提案した Intelligent Driver Model +g (IDM+g) を用いて構築した交通シミュレータから複数パターンの交通状況を設定し、シミュレーションデータを生成する。次にシミュレーションデータから時系列クラスタリングによって現実的に実現しうる CD 現象のサンプルを抽出する。最後に、抽出したシミュレーションデータからアンサンブル回帰木により高速な予測代替モデル(サロゲートモデル)を構築し、各特徴量の寄与度、解の予測精度から各パラメータがモデルの挙動にどのような影響を及ぼしているかを検討する。

2. データセットの構築

(1) ミクロシミュレーションモデル

まず、予測モデルの学習用データを生成するためのマイクロシミュレーションについて説明する。本研究では、全長 10,500m の高速道路単路部のうち、5,000m 地点から 6,500m 地点まで最大値 θ まで線形的に傾斜角度



図-1: シミュレーション状況

表-1: 各パラメータ

パラメータ名	説明
a	最大加速度 (m/s^2)
b	希望減速度 (m/s^2)
τ	安全車間時間 (s)
v	車両速度 (m/s)
u	希望速度 (m/s)
s	車間距離 (m)
s_0	最小車間距離 (m)
l	車長 (m)

が増加する勾配をもつサグ部が存在する状況を想定し、後述する IDM+g を用いて個別車両を動かすモデルを構築する。構築したモデルを用いて、渋滞の発生と個別車両のデータ取得のため、先頭車両がボトルネックを抜けた後 60 分間にわたってシミュレーションを行う。なお、本研究では、サグ部が車両にもたらす減速度とそれに伴った純粋な車両追従挙動を検討するため、車線変更や後続車の追突等は考慮しないものとする。

IDM+g のモデル式を以下の式 (1) に示す。

$$\frac{dv}{dt} = a \cdot \min \left[1 - \left(\frac{v}{v_0} \right)^4, 1 - \left(\frac{s^*(v, \Delta v)}{s} \right)^2 \right] - g\theta \quad (1)$$

$$s^*(v, \Delta v) = s_0 + v\tau + \frac{v\Delta v}{2\sqrt{ab}}$$

ここで、加速度 $dv/dt = 0$ 、相対速度 $\Delta v = 0$ の定常状態を仮定すると、図-2 のような FD (Fundamental Diagram) にを導出することができる。また、その頂点として以下のようなボトルネック容量を求めることができる。

$$Q_c = \frac{u\gamma\sqrt{\gamma}}{\tau u\sqrt{\gamma} + s_0 + l\gamma} \quad (2)$$

$$\gamma = \sqrt{\frac{1-g\theta}{a}}$$

(2) 学習データの生成

次に、構築したシミュレータを用いて、表-2 に示したパラメータを変化させてシミュレーションを行い、学習データを生成する。生成方法は、モデルパラメータ (a, b, τ, θ) および Demand (需要) をそれぞれ指定した範囲中で連続一様分布によって与え、10,000 回のシミュレーションを行い、3 地点の交通流率 (3,000m 地

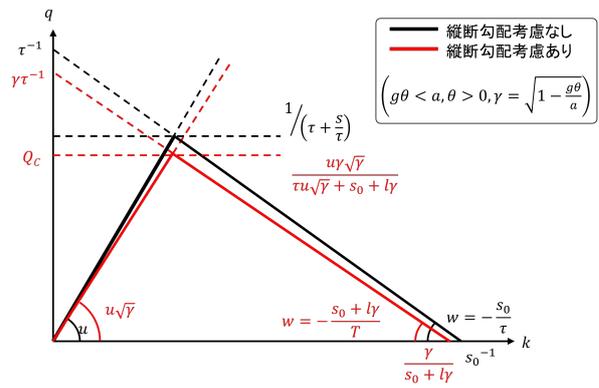


図-2: IDM+g 式の FD

表-2: データの生成範囲

パラメータ名	操作範囲
a	0.6~1.5
b	1.0~3.0
τ	1.25~2.25
θ	0.025~0.035
Demand	ボトルネック容量の 1.025 倍 ~1.1 倍

点, 5,000m 地点, 6,500m 地点) を、各地点とも最初の車両通過後の 60 分間計測することにより生成する。3 地点で交通流率を計測する理由は、次章で行う時系列クラスタリングを行う際、ボトルネックの下流端だけでなく、上流端やさらに手前の地点における交通状況も反映させることで、渋滞発生の有無や Stop and Go の挙動をクラスタリングに反映させる狙いがある。

3. 予測モデルの構築

(1) DTW による交通状態の分類

生成した学習データ用のシミュレーション結果について、自由流走行や不自然な停発進状態となるパラメータの組合せを除外するため、動的時間伸縮法 (DTW: Dynamic Time Warping) を用いた時系列クラスタリングで、交通状態を分類する⁵⁾。平均的な交通流率の大小がクラスタリングの支配的要因にならないよう、シミュレータによって計測された交通流率をそのままクラスタリングに適用するのではなく、ボトルネック交通容量 Q_c で交通流率を正規化した値をクラスタリングに適用する。また、3 地点間データの連結に際して、時間的、空間的な補正は行わないものとする。DTW のクラスタ数は、エルボー法による誤差平方和 (SSE) の減少傾向 (図-3) と、クラスタリング後の各交通状態の確認の結果、5 クラスタに設定した。

5 クラスタに分類後の交通状態のうち、自由流走行や

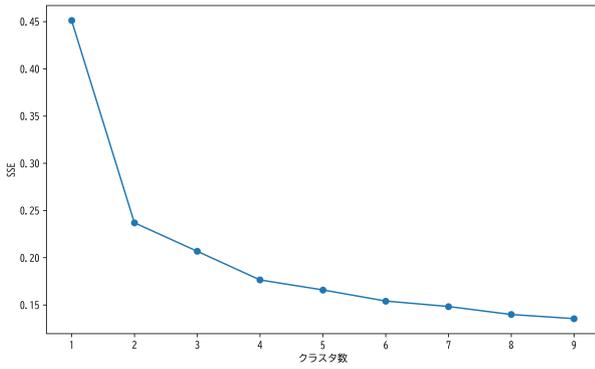


図-3: クラスター数の違いによる誤差平方和の変化 (エルボー法)

不自然な停滞状態を除いた 3 クラスターを予測モデル構築に用いることとした。クラスター 1~3 に分類されたパラメータでシミュレーションを行った場合の典型的な結果を図-4~図-6 に示す。抽出した 3 つすべてのクラスターで Capacity Drop の発生が確認され、Cluster1 から Cluster3 にかけて、Stop and Go の発生が強まっていく傾向がみられる。

表-3: 記述統計 (上段: 平均値, 下段: 標準偏差)

クラスター	<i>a</i>	<i>b</i>	τ	<i>Demand</i>	θ
Cluster1	1.164	1.828	1.876	1.065	0.027
	0.195	0.503	0.255	0.020	0.001
Cluster2	1.039	2.373	1.511	1.076	0.027
	0.167	0.464	0.170	0.017	0.001
Cluster3	1.146	2.375	1.715	1.067	0.027
	0.187	0.433	0.258	0.020	0.001

(2) サロゲートモデルの構築

本研究では予測器として勾配ブースティング決定木の一種である LightGBM(Ke et al.⁶⁾) を用い、作成したデータセットを学習用と検証用で 8:2 に分割 (Hold out 法) することによる予測精度の検証を行った。また、ハイパーパラメータ自動最適化フレームワークである Optuna(Akiba et al.⁷⁾) を用い、ベイズ最適化 (TPE:Tree-structured Parzen Estimator) によって予測値の平均絶対値誤差 (MAE) が最小となるよう 100 回の探索を行い、ハイパーパラメータを自動最適化した。

構築したモデルを用いて、捌け交通量を予測した結果が表-5 である。予測精度の評価指標としては二乗平均平方根誤差 (RMSE)、平均絶対値誤差 (MAE)、平均絶対パーセント誤差 (MAPE) の三種類を用いた。この表の結果から、Cluster1, Cluster3, および全てのデータか

表-4: 操作したハイパーパラメータ

パラメータ名	説明	操作範囲
num_leaves	葉の最大数	2~256
min_data_in_leaf	葉内の最低サンプル数	3~100
max_depth	木の最大深さ	3~8
learning_rate	学習率	1e-8~1.0
n_estimators	木の本数	3~2000
subsample_freq	バギングの頻度	0~100
colsample_bytree	特徴量の選択割合	0.7

表-5: 各予測指標の結果

クラスター名	RMSE	MAE	MAPE	サンプル数
Cluster1	9.03	6.25	0.45	2606
Cluster2	39.3	30.1	1.92	838
Cluster3	22.5	14.5	0.98	1585
All	21.55	12.53	0.85	5029

ら構築したモデルについて 1 %未満の誤差で精度高く予測できていることが見て取れる。このことは機械学習モデルによりシミュレータを十分に代替できる可能性を示唆している。一方で捌け交通量の時系列的変動が激しい Cluster2 ではやや精度が落ちる結果となった。

4. 予測モデルの解釈

(1) モデル解釈の概要

本章では構築した機械学習モデルを解釈することにより、シミュレーションパラメータと捌け交通量の関係性について考察を行う。具体的にはモデルの予測に対し、特徴量の寄与度を定量的に算出することのできる SHAP(SHapley Additive exPlanations)⁸⁾ を用い、各パラメータが予測値に及ぼしている影響及び関数形やパラメータ間の交互作用を、線形性を仮定せず近似的に算出する。以降、SHAP を用いたモデルの解釈を行っていく。

(2) モデルパラメータと CD 強度の関係性

図-7 から、図-9 の Violin Plot では横軸が SHAP の貢献度、縦軸が変数の重要度、プロットの色は赤色に近づくほどパラメータが大きいサンプルであることを示し、青色に近づくほど小さいサンプルであることを示す。図-7 から、図-9 にかけて、構築したすべてのモデルで最大加速度 *a* と安全車間時間 τ が捌け交通量への影響に対して支配的な役割を果たしており、それぞれ捌け交通量に対して正の相関と負の相関を持つことを示している。この点は、Wada et al.²⁾ が示した連続体交通流理論に定性的一致がある。図-8、図-10、図-9 を順に比較す

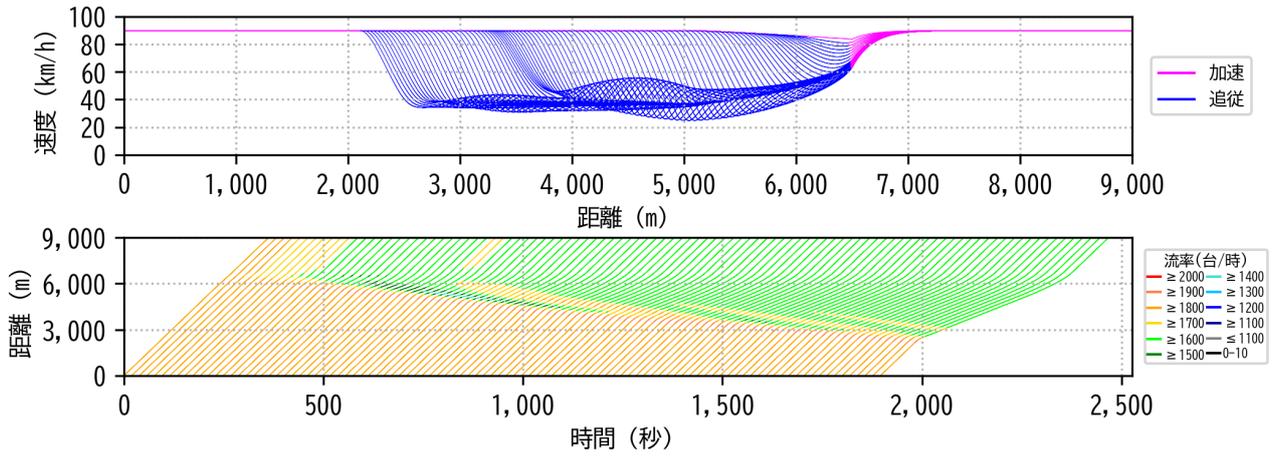


図-4: Cluster1 のシミュレーション結果

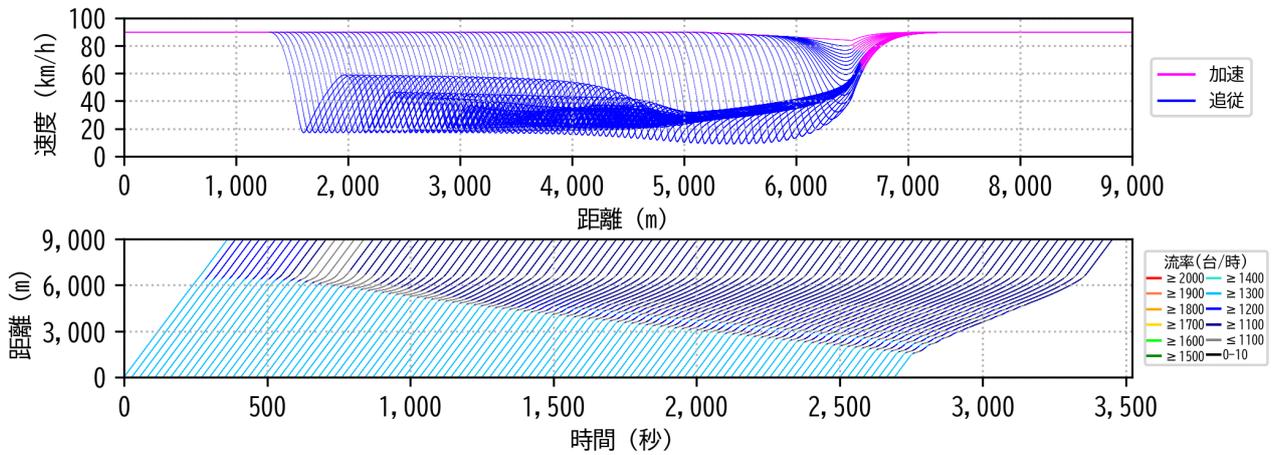


図-5: Cluster2 のシミュレーション結果

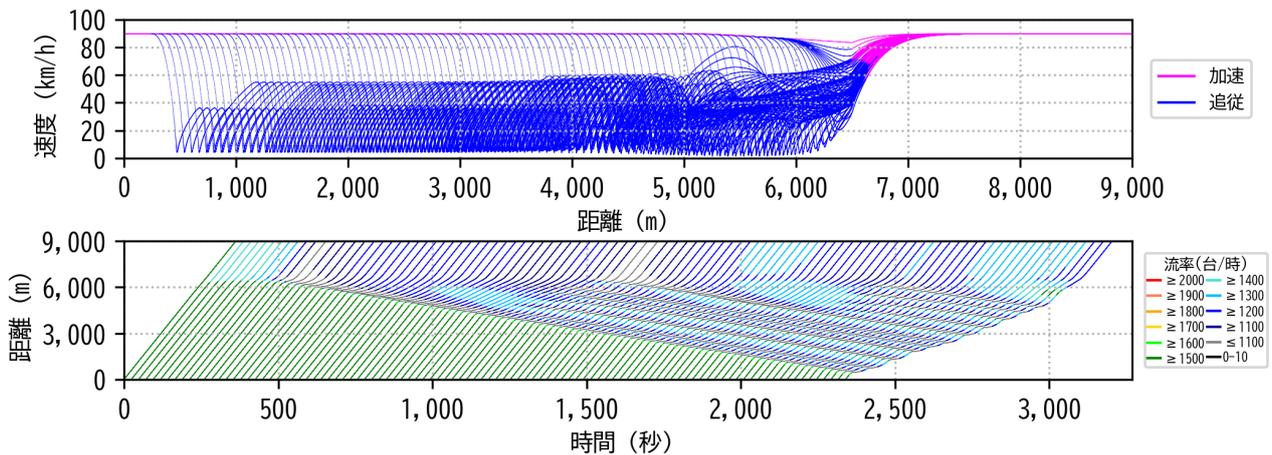


図-6: Cluster3 のシミュレーション結果

ると、マクロ交通流モデルでは表現できない Stop and Go が激しくなるにつれて希望減速度 b の影響度が低下しており、特徴量重要度が b と θ で入れ替わっていることが見て取れ、Stop and Go の発生に伴って車両減速度が CD に及ぼす影響が低下していく傾向がみられる。

(3) 関数形の捕捉

図-12 から図-15 にかけては、横軸と色が第一変数の値、縦軸がサンプルの SHAP 値を取った、Dependence Plot と呼ばれる図である。また、軸のスケールに差異があることについて留意が必要である。以降ではこれら

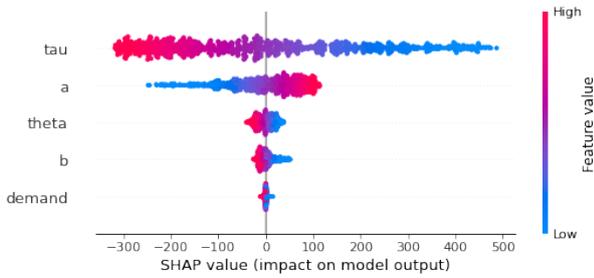


図-7: SHAP の分布 (All)

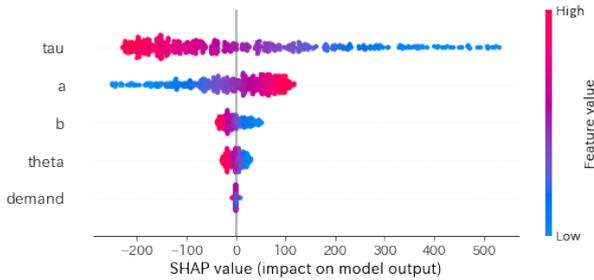


図-8: SHAP の分布 (Cluster1)

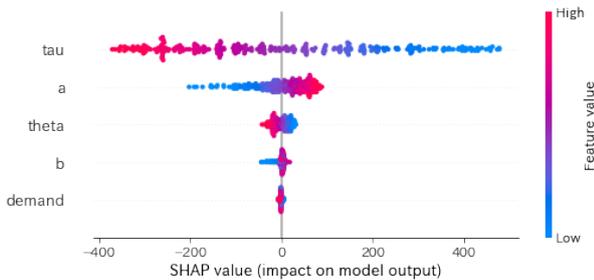


図-9: SHAP の分布 (Cluster2)

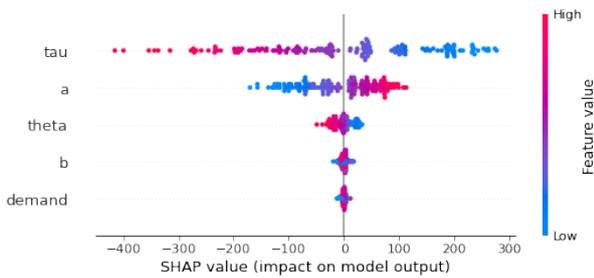


図-10: SHAP の分布 (Cluster3)

の図を用いて重要となるパラメータが CD の及ぼす影響の変化について考察を行っていく。また、理論式の構築に際して関数形の捕捉が重要であることから、全データを学習データとして用いたモデルを対象に考察を行う。また、特徴量重要度の高い順から考察を行っていく。

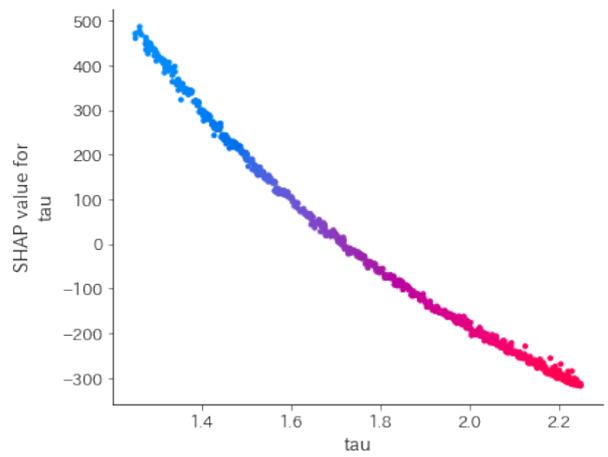


図-11: τ に関する SHAP の分布

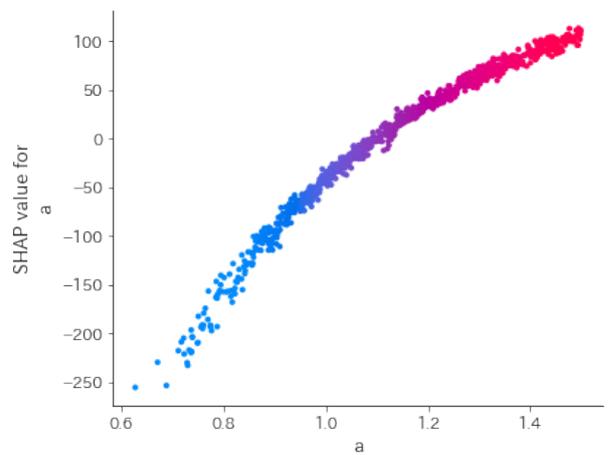


図-12: a に関する SHAP の分布

図-11 から、安全車間時間 τ は捌け交通量に対して反比例する傾向が見られ、関数形には理論式の分母に含まれる可能性が考えられる。

図-12 から、最大加速度 a が捌け交通量にもたらす影響は正の相関を持ちつつも影響度が逡減していく傾向がみられた。

図-13 からは、勾配の強さ θ は車両に対する加速度を直接的に低下させるパラメータであるため、線形に捌け交通量を低下させる関係性が見られた。

図-14 から、希望減速度 b は捌け交通量に対して反比例する傾向が見られ、 $b = 1.75$ 付近で傾向が変化していることが見られる。

図-15 からは、到着需要 Demand は渋滞の発生有無に影響するパラメータであると考えられることから、緩やかな負の相関が見られたものの、強い傾向は見られなかった。一方で不連続点がいくつか存在しており、到着需要の大きさによって何らかの交通状態の遷移が起きている可能性が指摘できる。

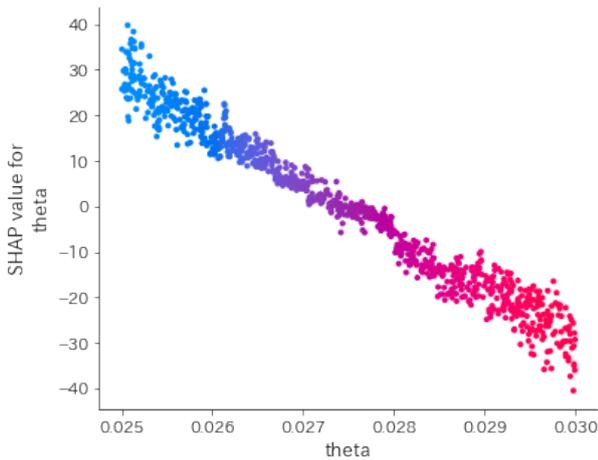
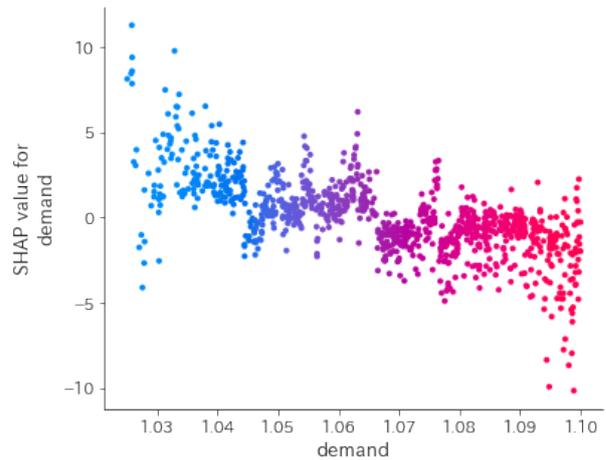
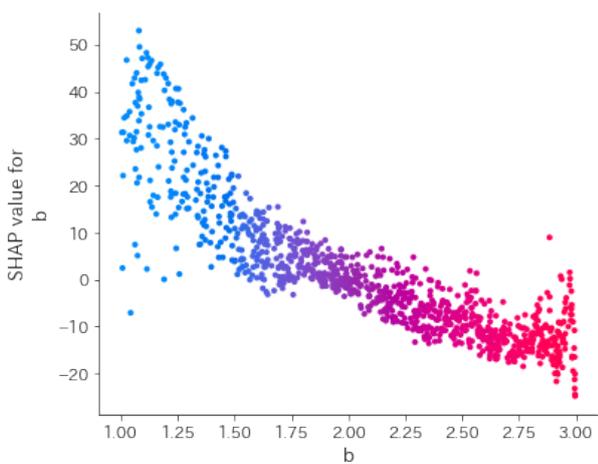
図-13: θ に関する SHAP の分布

図-15: Demandに関する SHAP の分布

図-14: b に関する SHAP の分布

5. まとめ

本研究では高速道路サグ部での CD 現象を勾配の影響を考慮したマイクロ交通流モデルによって再現した。その後、シミュレーション状況を変更し複数パターンにわたるデータを生成し、時系列クラスタリングによって現実に起きうる交通状態の抽出を行った。最後に、勾配ブースティング決定木によってパラメータから捌け交通量を予測するサロゲートモデルを構築することで、モデルの挙動の解釈可能性及び計算量の削減によるシミュレータの代替可能性について検証した。その結果、SHAP を用いた解釈により各パラメータと捌け交通量の関係性について関数形や変数重要度といった観点からシミュレータの特性の解明に示唆をもたらすことができた。また、1%未満での高精度な予測により、機械学習モデルがシミュレーションを代替できる可能性についても提示することができた。

今後の課題としては、本研究で得た示唆を元に実際に理論式を構築することや、サロゲートモデルの応用可能性について検討することが挙げられる。

謝辞: 本研究は JSPS 科研費・基盤研究 B (課題番号 23K26218) の助成金を受けた研究の一部である。ここに記し、感謝の意を表します。

参考文献

- 1) 越正毅, 1986. 高速道路のボトルネック容量, 土木学会論文集, IV-5 (371), 1 - 7. [Koshi, M., 1986. Capacity of Motorway Bottlenecks, *Transactions of the Japan Civil Engineering Society*, IV-5 (371), 1-7.]
- 2) Jin, W.-L.: Kinematic wave models of sag and tunnel bottlenecks. *Transportation Research Part B: Methodological*, Vol.107, pp.41 - 56, 2018.
- 3) Wada, K., Martainez, I., Jin, W.-L.: Continuum carfollowing model of CD at sag and tunnel bottlenecks, *Transportation Research Part C*, 113, 260 - 276, 2020.
- 4) Treiber, M., Kesting, A. *Traffic flow dynamics. Traffic Flow Dynamics: Data, Models and Simulation*, Springer-Verlag Berlin Heidelberg, 983-1000, 2013
- 5) 甲斐慎一郎, 和田健太郎, 堀口良太: 追従モデルによる高速道路サグ部の CD 現象の再現. 土木学会論文集 D3 (土木計画学), Vol.78, No.5, pp.I.963-I.971, 2023. [Kai, S., Wada, K. and Horiguchi, R.: Reproducing CD phenomenon by car following models, *Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management)*, Vol.78, No.5, pp.I.963-I.971, 2023]
- 6) Ke, Guolin, et al. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.
- 7) AKIBA, Takuya, et al. Optuna: A next-generation hyperparameter optimization framework. In: *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, p. 2623-2631.2019,
- 8) Lundberg, S. M., Lee, S. I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

(Received October 3, 2024)

(Accepted October 3, 2024)

ANALYSIS OF RELATIONSHIP BETWEEN CAR-FOLLOWING MODEL
PARAMETERS AND CAPACITY DROP USING EXPLANATORY TECHNIQUE
FOR MACHINE LEARNING MODELS

Keisuke YAMADA, Shin-ichiro KAI and Kentaro WADA