

交通流データの高解像度化と分類

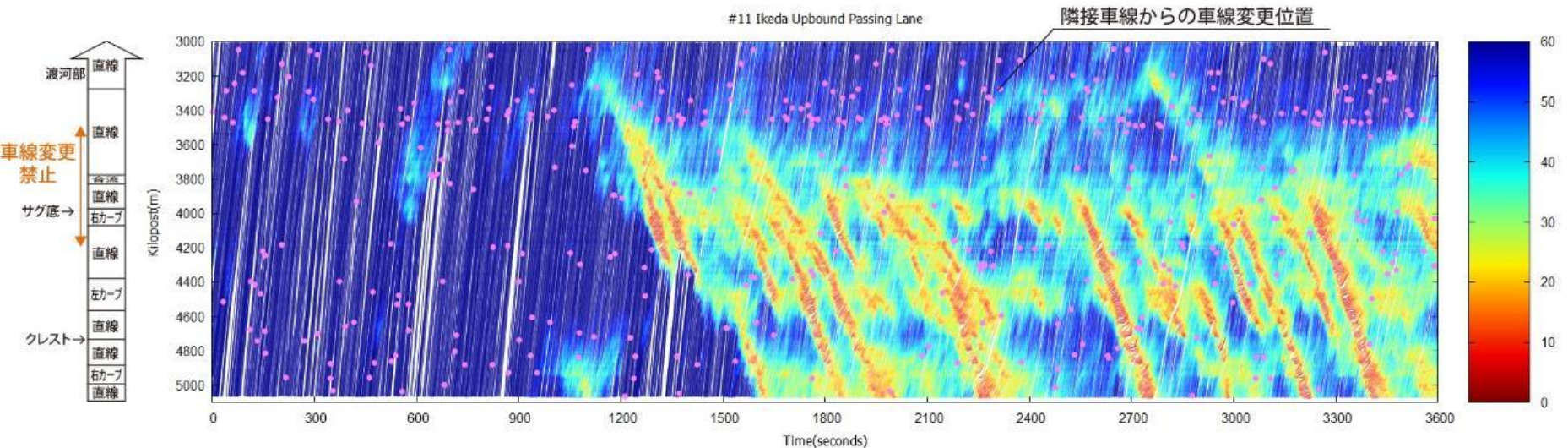
2023/09/18 行動モデル夏の学校 @ 本郷

金沢大学 中西 航

nakanishi @ se.kanazawa-u.ac.jp

■ 全車両軌跡データ

- 道路上のすべての自動車の時々刻々の移動(位置)の記録



出所：阪神高速

■ このような詳細(=高解像度)データによって

- できるようになったこと
- 困るようになったこと

1. 区間別の交通流FD推定

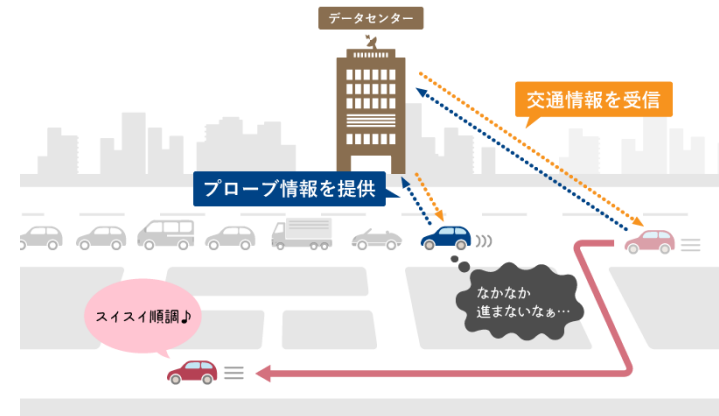
(中西2022)

マクロ交通流モデル

- 交通流を流体と捉えて記述
- 道路ネットワーク上で時々刻々と変化する**交通状態変数 (速度, 密度, 流率)**を知りたい
- 時空間的に網羅的に観測することは難しい
 - カウンタは設置地点のみで流率(+速度)を観測
 - プローブは自車の速度(+前車との間隔)を観測

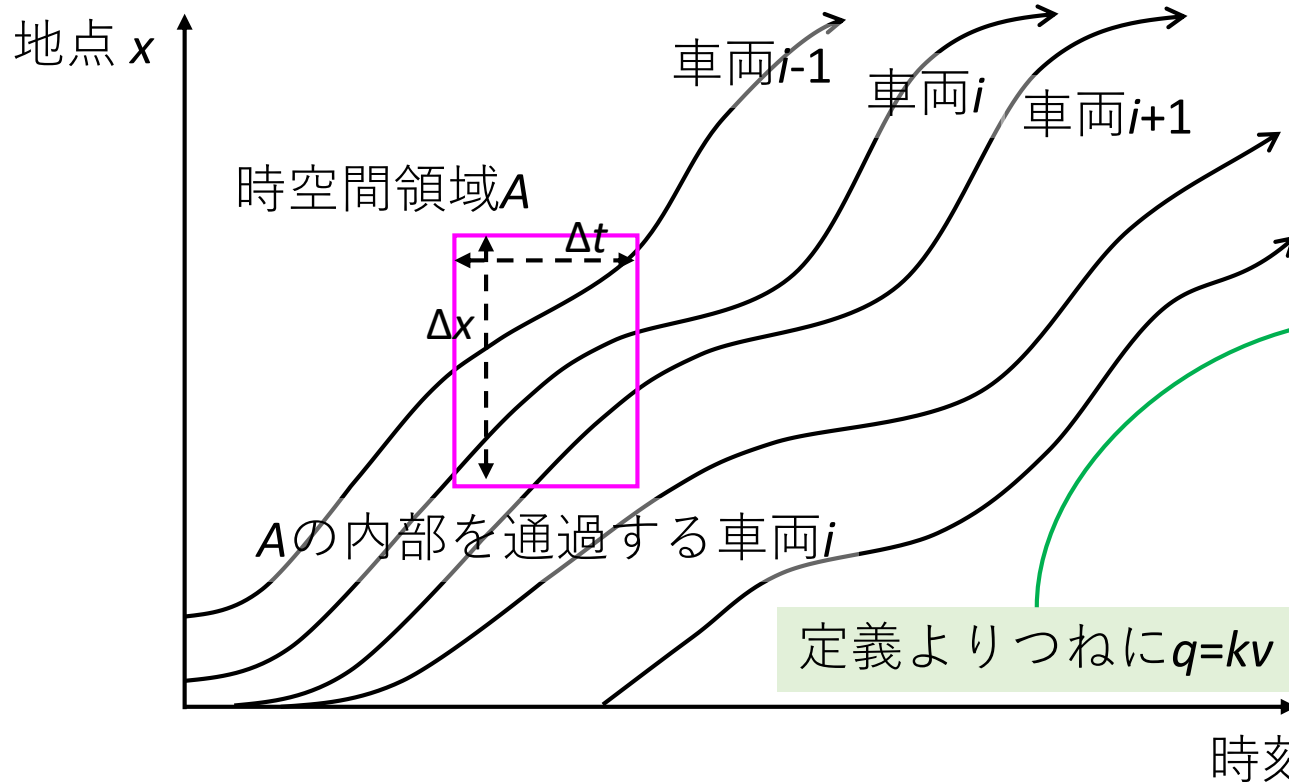


<http://www.crp.co.jp/business/system/>



<https://aichi-its.jp/knowledge/glossary/probe/>

■ ある時空間領域の流率 q , 密度 k , 速度 v



Edie (1963) の定義

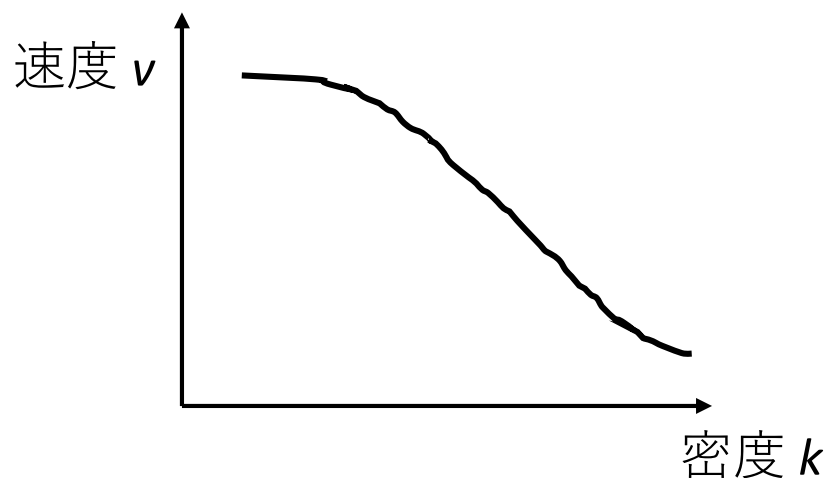
- 流率 $q(A) = \frac{\sum_{i=1}^n d_i(A)}{|A|}$
- 密度 $k(A) = \frac{\sum_{i=1}^n t_i(A)}{|A|}$
- 速度 $v(A) = \frac{\sum_{i=1}^n d_i(A)}{\sum_{i=1}^n t_i(A)}$
- $t_i(A)$: 領域 A 内での車両 i の移動時間
- $d_i(A)$: 領域 A 内での車両 i の移動距離
- $|A| = \Delta t \Delta x$: 領域 A の面積

■ 流率 q , 密度 k , 速度 v が満たす関係

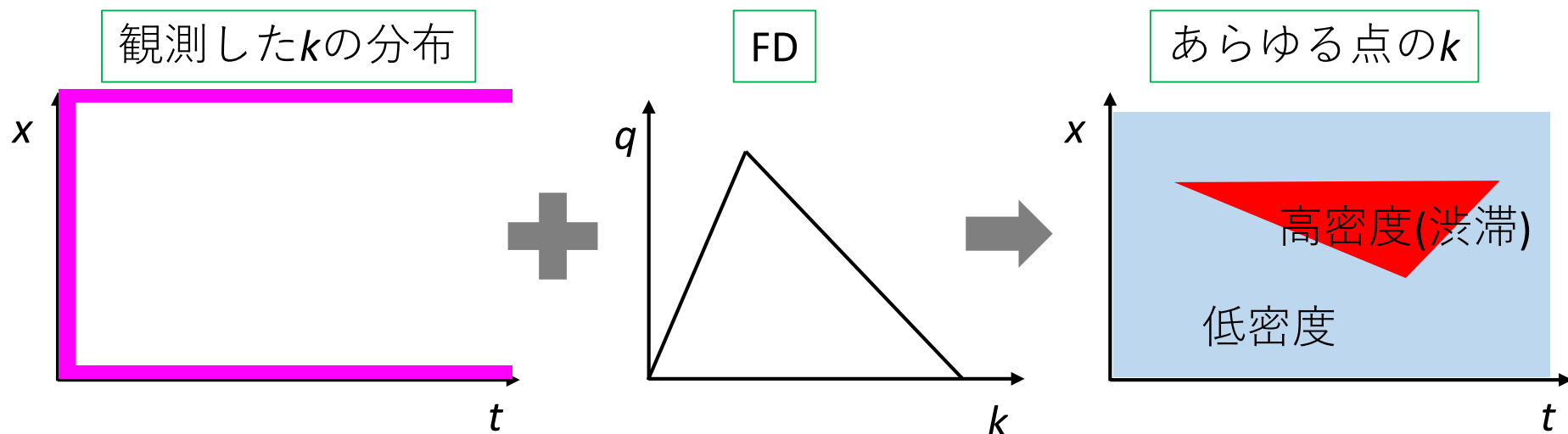
1. 常に成立する関係: 定義より $q = kv$

2. 経験的な関係

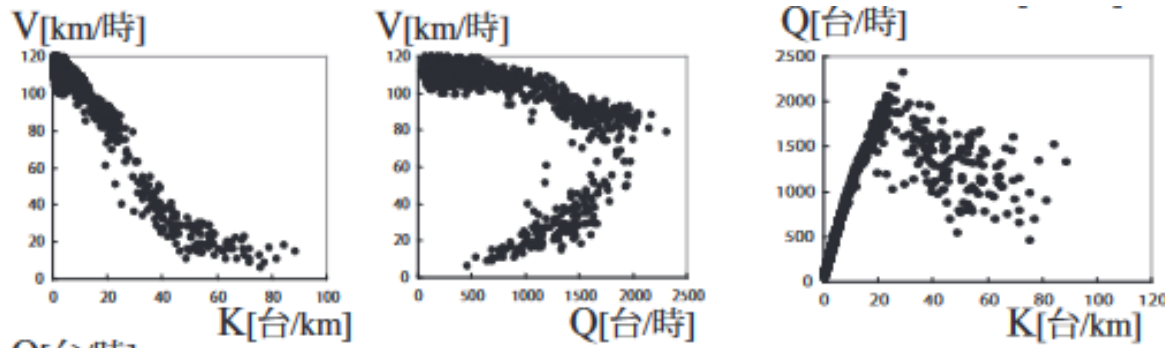
- kv 平面に観測データをプロットすると, ほぼ必ず(広義の)右肩下がりである
- 速度密度関係(kv 関係)



- いつ・どこで渋滞しているか?? 渋滞しそうか??
 - タイムスペース図の任意地点・時点の交通状態を知りたい
 - あらゆる地点・時点の**密度**がわかれば良い
- すべてを観測するのは当然不可能だが、FDを使って密度が伝播していく様子を考えることができる
= 観測せずに予測することもできる
- 全体イメージ



- 交通状態変数(流率, 密度, 速度)の関係を表す関数
 - 経験的に一定の関係があることが知られている



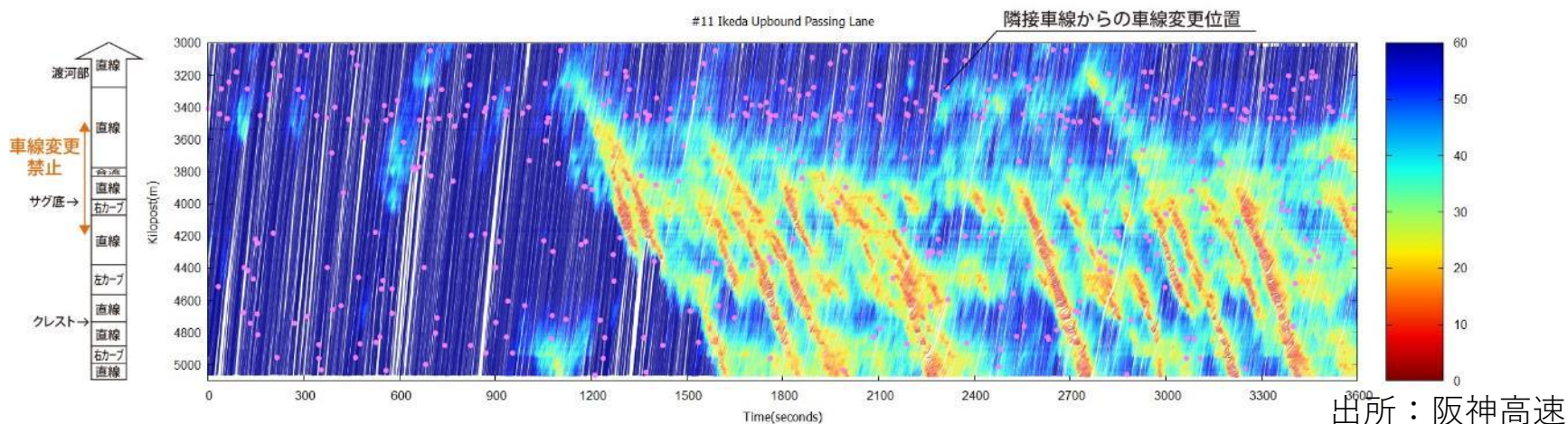
出所: https://www.iatss.or.jp/common/pdf/publication/commemorative-publication/iatss40_theory_04.pdf

- FD推定: この関数(のパラメータ)を求めること
 - 道路ネットワークの性能評価にあたって重要
 - 実際に任意の時点・地点について推定することは難しい
 - カウンターがないところでは何台どのくらいの速度で通ったかわからない
 - 場所だけでなく本来は人や時間帯によっても違う

全車両軌跡データ

- 文字通り，すべての車両の位置情報が記録されたデータ
- 代表例はNGSIM@アメリカ，近年阪神高速がZTDを公開
 - ZTDは世界最大規模のデータセット

```
1 0,073000600,1,44.1,2,135.4588401,34.5827436,6899.4,5.4,1
2 0,073000700,1,44.0,2,135.458847,34.5827526,6898.2,5.4,1
3 0,073000800,1,44.0,2,135.4588538,34.5827616,6896.9,5.4,1
4 0,073000900,1,44.0,2,135.4588608,34.5827707,6895.7,5.4,1
5 0,073001000,1,44.0,2,135.4588677,34.5827799,6894.5,5.4,1
6 0,073001100,1,44.0,2,135.4588748,34.5827893,6893.3,5.4,1
7 0,073001200,1,44.0,2,135.458882,34.582799,6892.0,5.4,1
8 0,073001300,1,43.9,2,135.4588894,34.5828089,6890.8,5.4,1
```



- ネットワーク上の分析に用いることを考えると、本来の意味で地点ごとに異なるFDを推定したい。

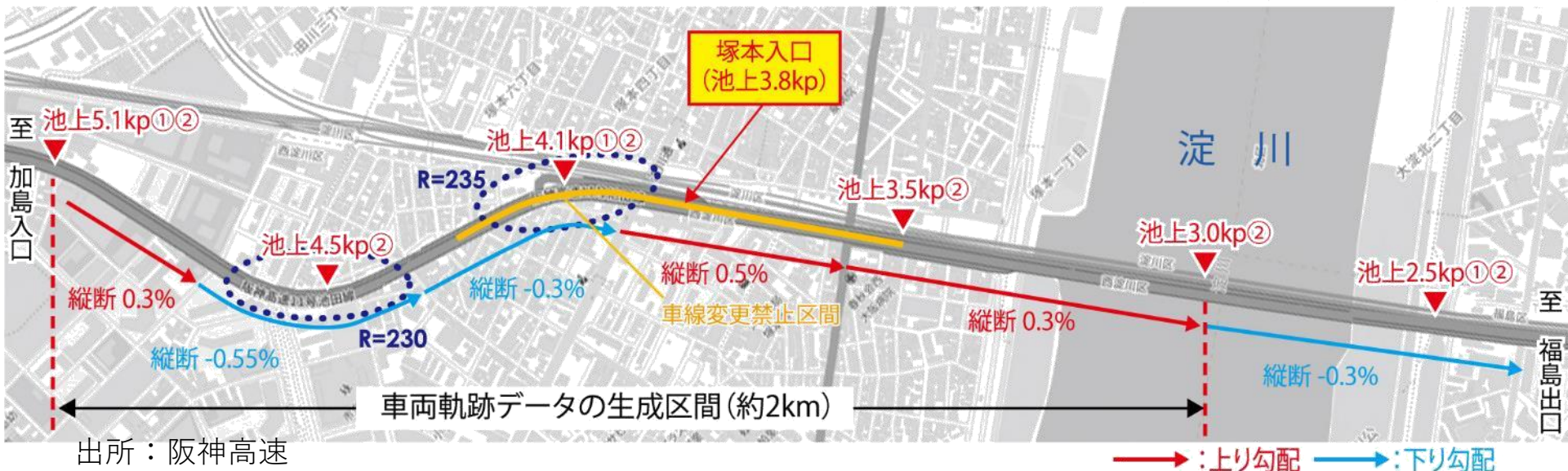
- 直線と曲線，上りと下り，知りたい任意の小さい区間

データを分類(分割)したい

- 全車両軌跡データがあればできるのでは？

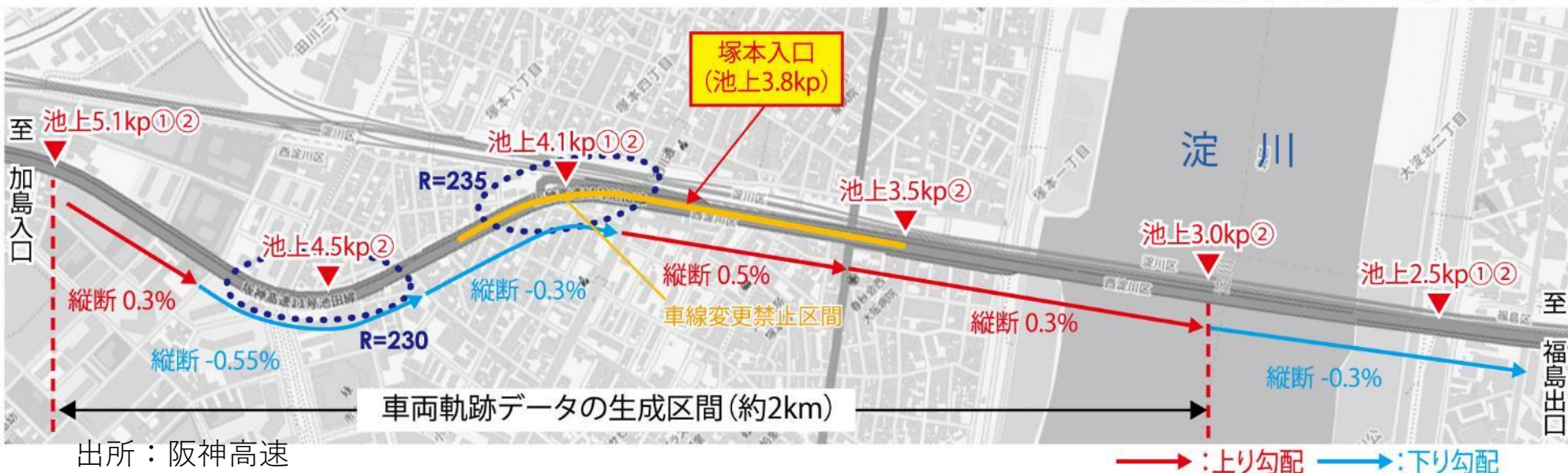
- 現状：感知器データをそのリンク全体のFDとみなす

▼：車両検知器設置位置(番号は設置車線)

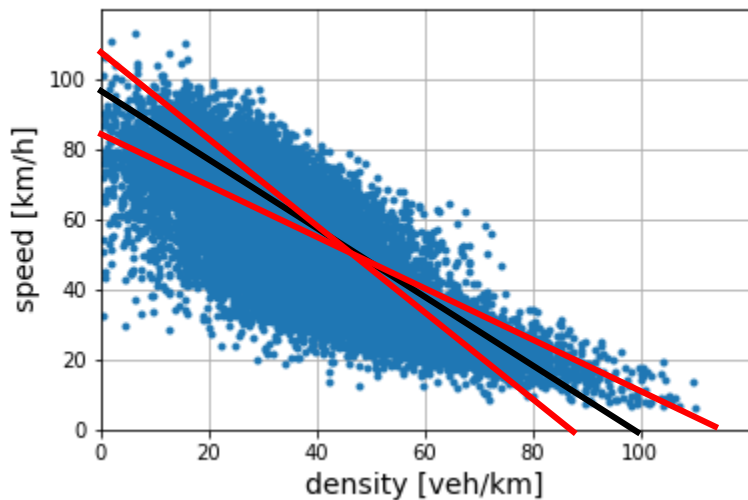


- 「どこが同一のFDに従う区間か」の特定が必要
 - 先験的知識には限界がある & あまり頼りたくない
- スペースモデリングを用いて以下を同時に実現：
 - 同一のFDに従う区間の特定
 - それらの区間におけるFDのパラメータ推定

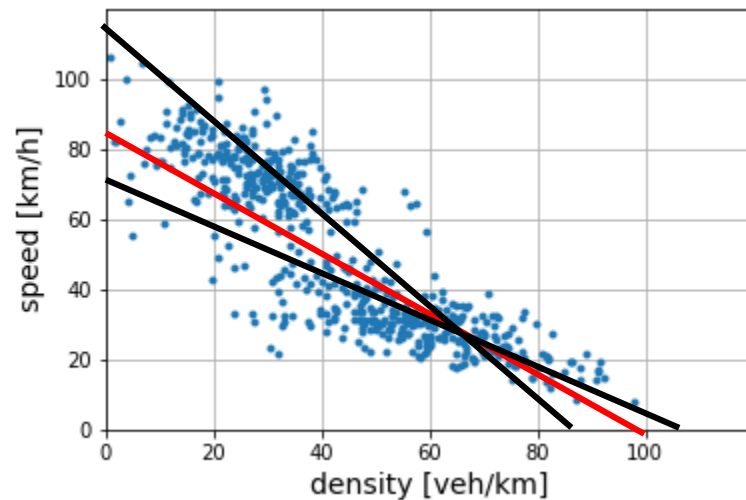
▼：車両検知器設置位置 (番号は設置車線)



■ 広く取り過ぎると



■ 狭く取り過ぎると



■ 異なるFDに従う区間を同一FDとみなす

- 根本的に正しくないFD
- 豊富な情報を無駄にする

■ 同一FDに従う区間に異なるFDが推定される

- ノイズを拾っている
- ネットワーク分析が無意味に複雑化

■ 対象時空間

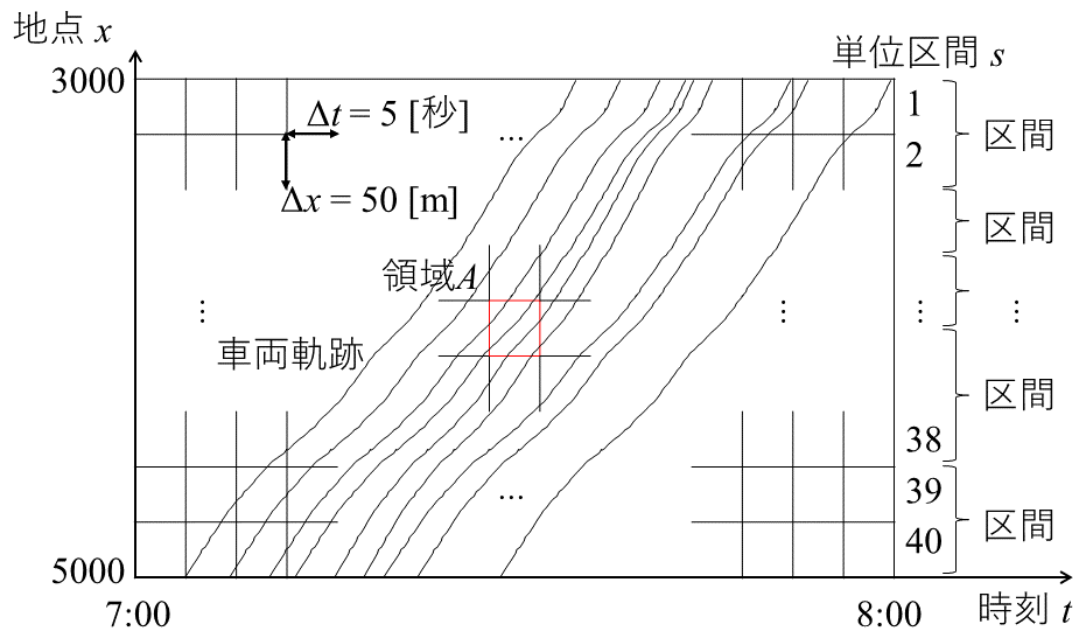
- 午前7-8時, 追越車線
- 位置 x
 - 下流端 $x=3000$
 - 上流端 $x=5000$

■ 交通状態

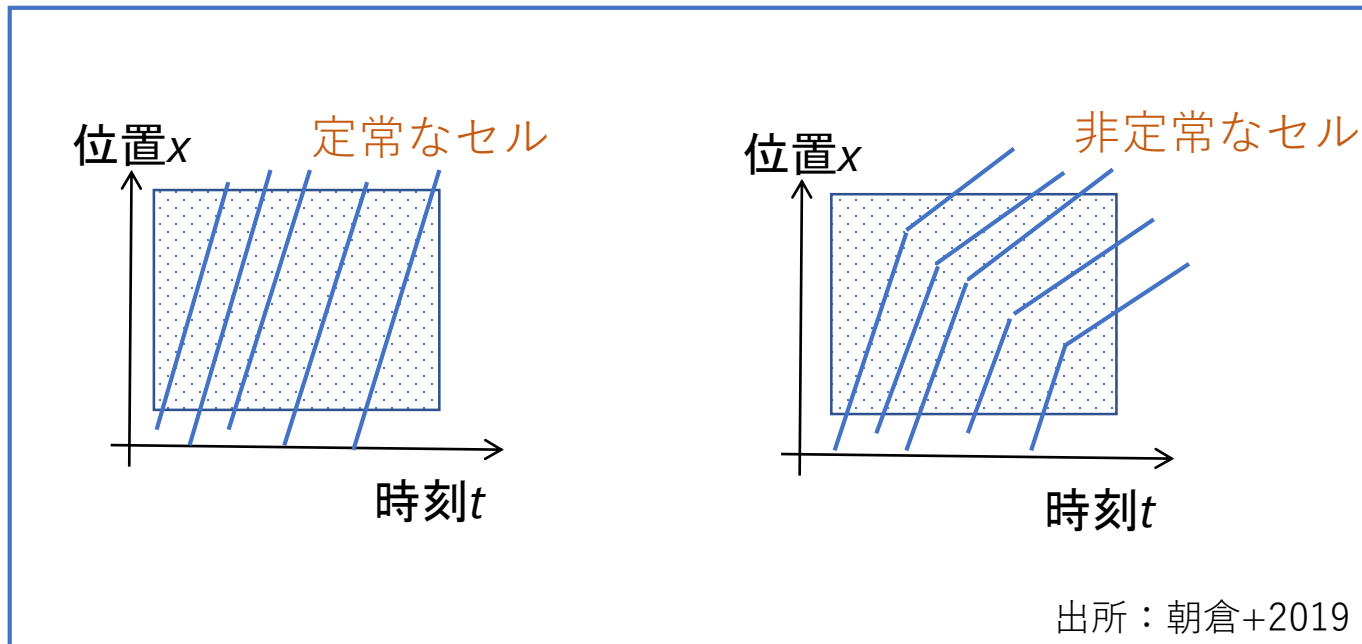
- 速度 v , 密度 k
- Edieの定義で算出, 矩形サイズは適宜 (今回は5s, 50m)

■ 区間別FDの「区間」

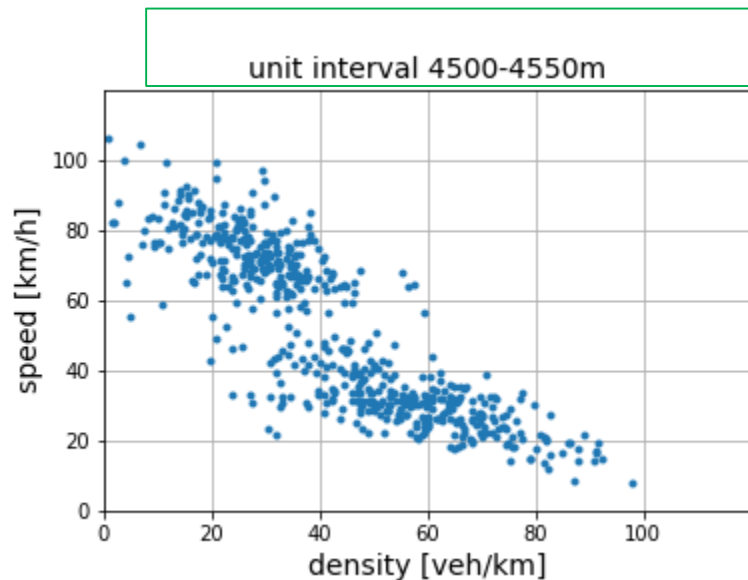
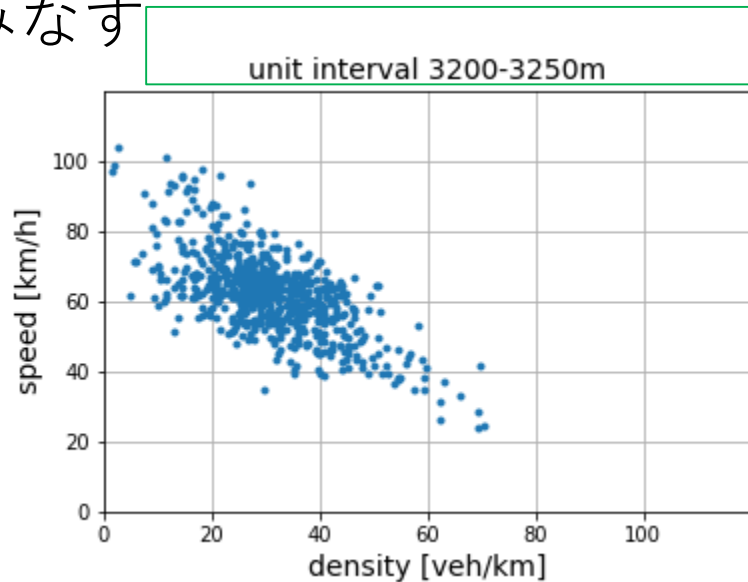
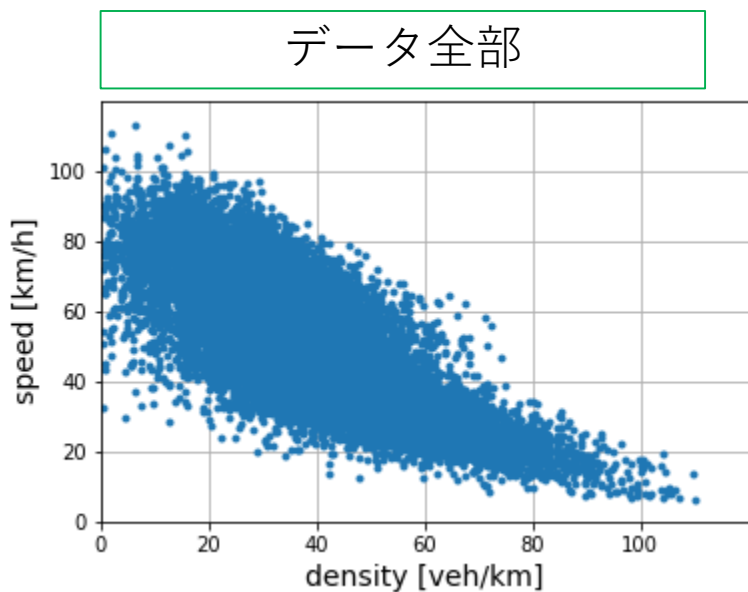
- 交通状態算出に用いる距離を単位区間 s と呼ぶ
- いくつかの連続する単位区間で, 同一のFDに従うとみなせる部分を区間と呼ぶ



- 定常状態:
時空間領域内の全車両が、車間距離と速度を変化させる動機を持たない状態
- Fundamental Diagram は定常な交通流の下で成立する関係



- 速度の変動係数 $\sigma < 0.15$ を定常とみなす
- いくつかのk-v図

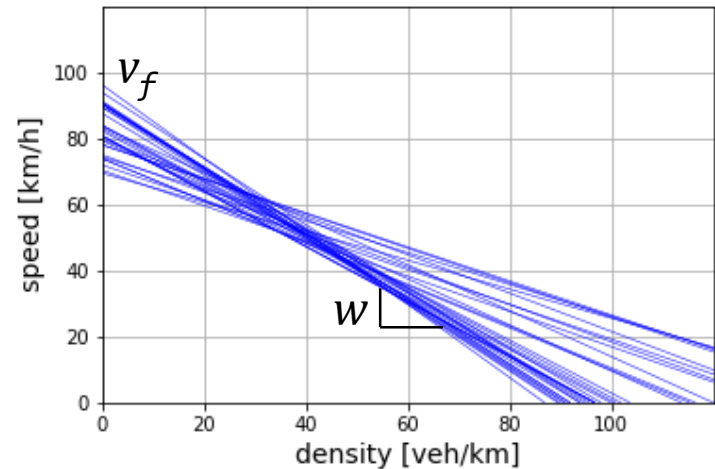


■ 仮定：

- 定常な交通状態は Greenshields FD に従う：

$$v = v_f + wk$$

- 異なる区間別FDでは 上式のパラメータ(v_f と w)が異なる
- 隣接単位区間どうしで、パラメータが大きく異なるようなデータのと看両単位区間が異なる区間別FDに従う。そうでない場合、同一のFDに従う区間となる。



■ 留意点：

- そもそも Greenshields で良いかは議論していない
- ボトルネックを明示的に考慮していない

- スパースモデリングの代表的手法：回帰モデルの損失関数にパラメータの L_1 ノルムを正則化項(罰則項)として加える

- 変数選択 + 過学習の抑制を期待

- モデル式が $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ のとき、

最小化する式は $\min_{\boldsymbol{\beta}} \left[\frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \right]$

普通の最小二乗回帰

罰則項

変数の定義：

\mathbf{y} : 非説明変数ベクトル

\mathbf{X} : 説明変数行列

$\boldsymbol{\beta}$: パラメータベクトル

$\boldsymbol{\varepsilon}$: 誤差項ベクトル

λ : 正則化項の係数

- 正則化項の係数 λ の値は適宜定める

- 情報量規準最小化，選ばれる変数の実用的な制御，など
- $\lambda \rightarrow 0$ のとき普通の最小二乗回帰
- λ が十分大きいときすべての係数 $\rightarrow 0$ (または定数項のみ)

■ Lassoの様々な発展系: 正則化項の様々なバリエーション

- Ridge, Elastic net, Group Lasso, ...

■ Generalized Fused Lassoは

隣接パラメータが同じ値を

取りやすくするように、以下を最小化

変数の定義：
 y_i : \mathbf{y} の*i*行
 X_{ij} : \mathbf{X} の*i*行*j*列
 β_j : $\boldsymbol{\beta}$ の*j*行
 \mathcal{C} : 隣接関係の集合
 λ, γ : 正則化項の係数

$$\min_{\boldsymbol{\beta}} \left[\frac{1}{2} \sum_{i=1}^N \left(y_i - \sum_{j=1}^K \beta_j X_{ij} \right)^2 + \lambda \sum_{(\mu, \nu) \in \mathcal{C}} |\beta_{\mu} - \beta_{\nu}| + \gamma \lambda \sum_{j=1}^K |\beta_j| \right]$$

普通の最小二乗回帰

正則化項1:
隣接パラメータの差

正則化項2:
通常のLassoと一緒に

■ γ は正則化項同士の重みをコントロール，決め方は λ 同様

■ この手法は従来空間統計分野で活用されている

- 前提条件から，区間別FDは

$$v_i = \sum_{s \in S} I_{is} (v_{fs} + w_s k_i) + \varepsilon_i$$

サンプル*i*が単位区間*s*から
得られたとき1

変数の定義：

v_{fs} : 単位区間*s*の v_f
 w_s : 単位区間*s*の w
 S : 単位区間の集合

- このとき変数候補は v_{fs} と w_s の80個

- 総当たりで変数選択を行うと 2^{78} 通り，当然計算不可

- Generalized Fused Lassoで定式化すると，

$$\min_{v_f, w} \left[\frac{1}{2} \sum_{i=1}^N \left(v_i - \sum_{s \in S} I_{is} (v_{fs} + w_s k_i) \right)^2 + \lambda \sum_{(\mu, \nu) \in C} (|v_{f\mu} - v_{f\nu}| + |w_\mu - w_\nu|) + \gamma \lambda \sum_{s \in S} (|v_{fs}| + |w_s|) \right]$$

普通の最小二乗回帰

正則化項1:
隣接パラメータの差

正則化項2:
通常のLassoと一緒に

- Generalized Fused Lassoで定式化すると,

$$\min_{v_f, w} \left[\frac{1}{2} \sum_{i=1}^N \left(v_i - \sum_{s \in S} I_{is} (v_{fs} + w_s k_i) \right)^2 + \lambda \sum_{(\mu, \nu) \in C} (|v_{f\mu} - v_{f\nu}| + |w_\mu - w_\nu|) + \gamma \lambda \sum_{s \in S} (|v_{fs}| + |w_s|) \right]$$

普通の最小二乗回帰

正則化項1:
隣接パラメータの差

正則化項2:
通常のLassoと一緒

- 本研究では $\gamma=0$ に固定

- 特定の v_{fs} や w_s が0になることは考えにくい
- 複雑なFD式ならこの項も必要

- この定式化は, Special caseとして下記を含む

- $\lambda=0$ のとき全単位区間に個別のFDを推定する場合
- $\lambda \gg 0$ のとき区間全体で単一のFDを推定する場合

適用結果

■ 今回のpracticalな推定手順

- $\lambda \gg 0$ から $\lambda=0$ まで徐々に変化させる
- 選択される変数が80個から2個までの全区間設定を得る
- 選択された変数のみを用いて区間別FDを再推定

■ 結果

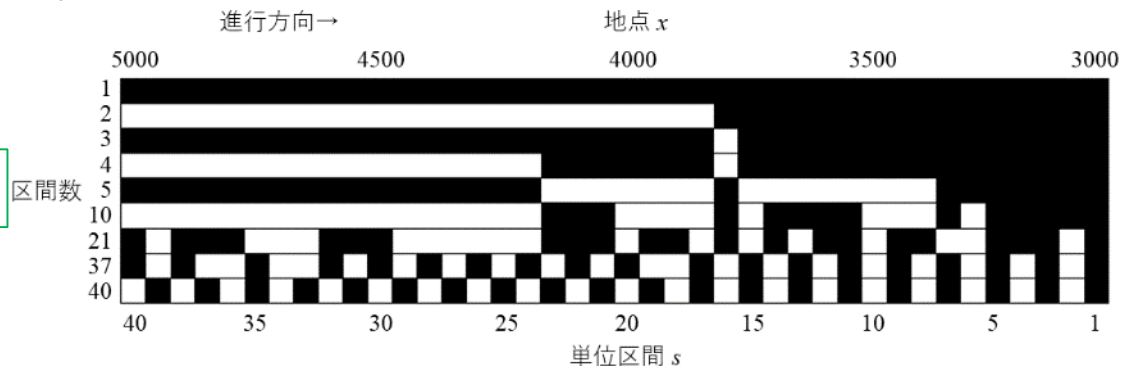
区間数	区間構成 (単位区間 s の組み合わせ)	λ	変数の数	AIC	BIC
1	1-40	18500.292	2	190884.3	190908.7
2	1-16,17-40	16674.702	3	190850.8	190883.3
3	1-15,16,17-40	13765.022	4	190786.2	190826.9
4	1-15,16,17-23,24-40	11382.255	5	190739.3	190788.1
5	1-7,8-15,16,17-23,24-40	10881.135	6	190586.5	190643.4
10	1-5,6,7,8-10,11-14,15,16,17-20,21-23,24-40	5889.954	11	189630.7	189728.2
21	1,2,3-5,6-7,8-9,10,11-12,13,14,...,17,18-19,20, 21-23,24-29,30-32,33-35,36-38,39,40	950.489	26	188858.7	189078.3
37	1,2,...,17,18-19,20,21,...,32,33-34,35,36-37,38,39,40	163.36	61	188637.6	189141.8
40	1,2,...,40	0	80	188640.7	189299.5

区間数 21 のとき BIC 最小, 37 のとき AIC 最小である.

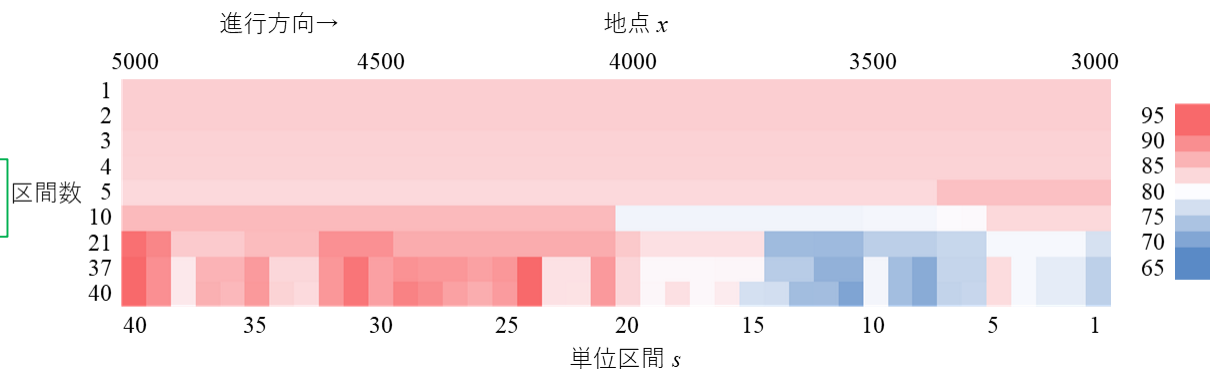
統計的には、闇雲に単位区間の長さを小さく取り、それぞれ個別にFDを推定すれば良いわけではない(もちろん実用的でもない)

■ 区間の分かれ方の可視化

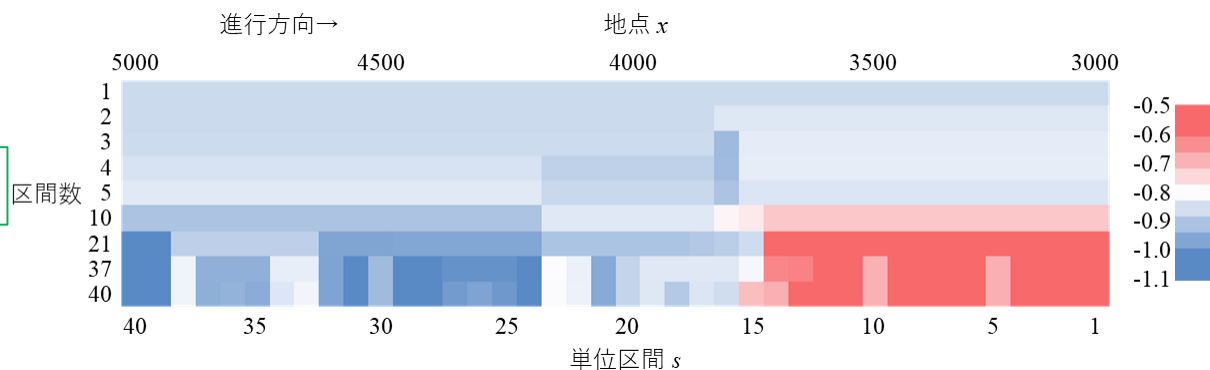
分割状況



切片



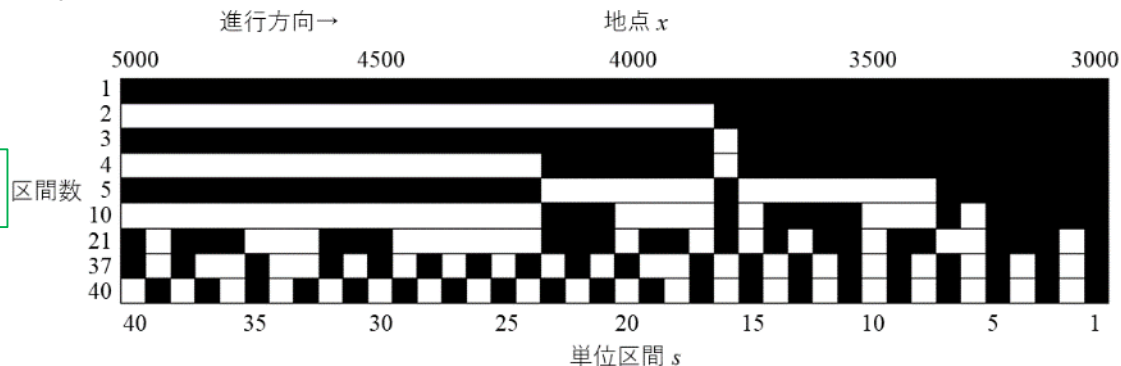
傾き



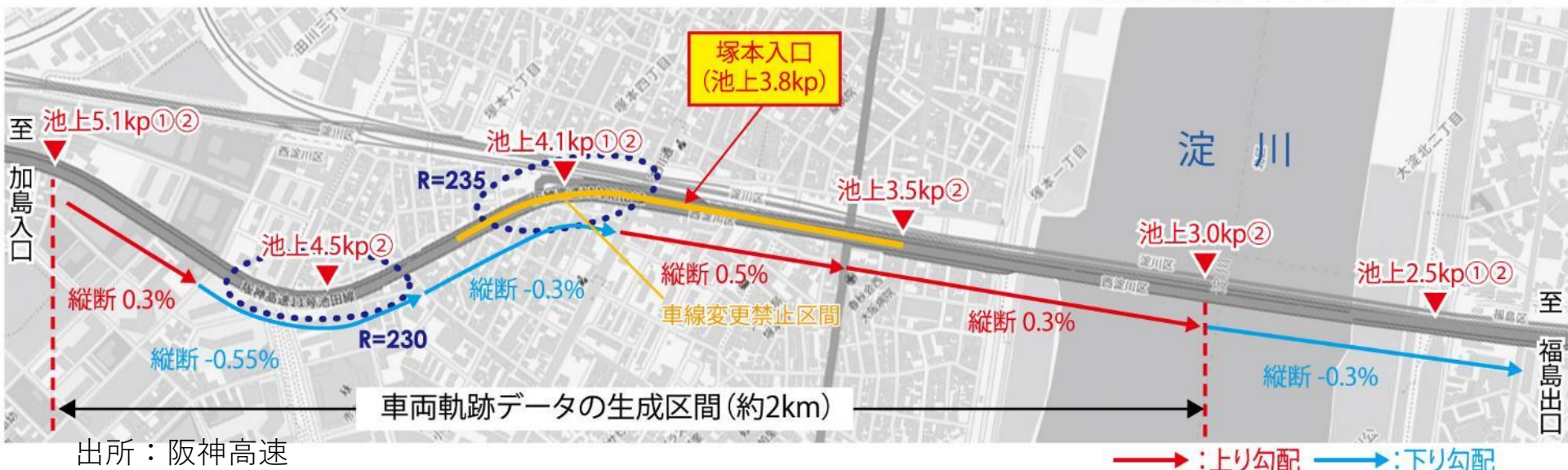
- 区間数2で地点 $x=3800$ が境界，区間数3で地点 $x=3750$ も境界
- 以降のおおむねの傾向：
 - 上流側から直線下り勾配の $x=4150$ までの区間
 - その後 $x=3800$ までの右カーブ下り勾配からサグを経て直線上り勾配の区間
 - $x=3800$ 付近のみの区間
 - その後の橋梁手前 $x=3350$ までの直線上り勾配区間
 - その後の上り勾配区間
 - ただし，区間数を増やしていくと， $x=4000$ 付近のサグより下流側が細かい分割となりやすい。

■ 区間の分かれ方の可視化

分割状況



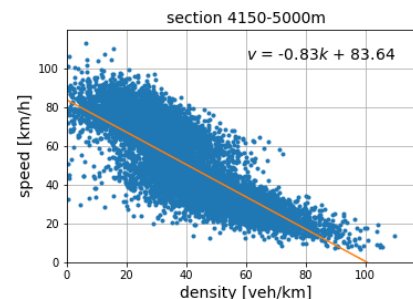
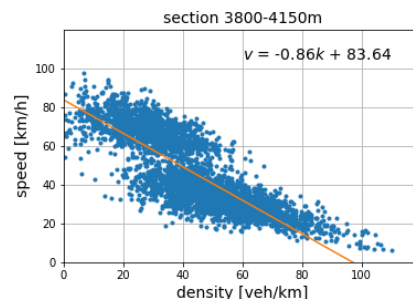
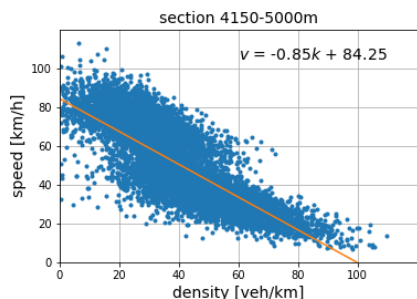
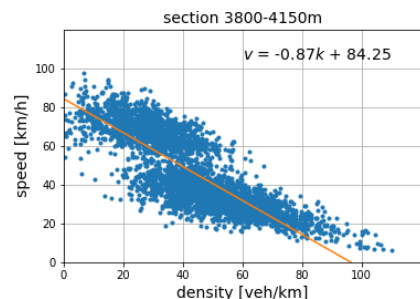
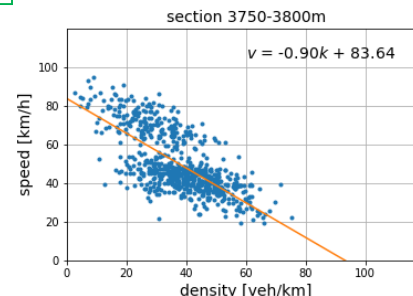
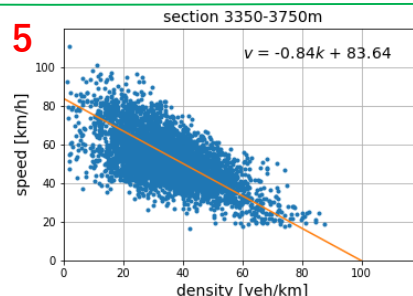
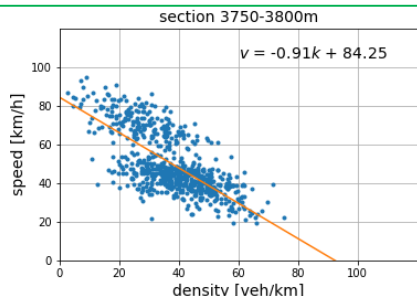
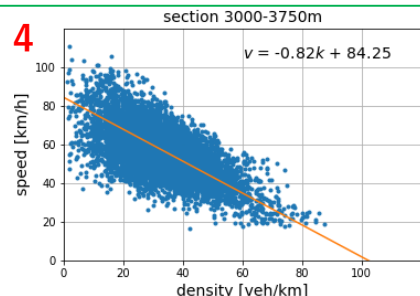
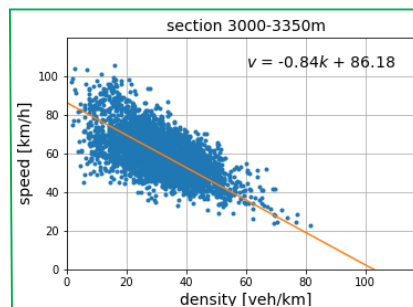
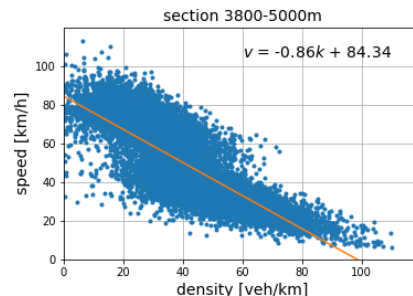
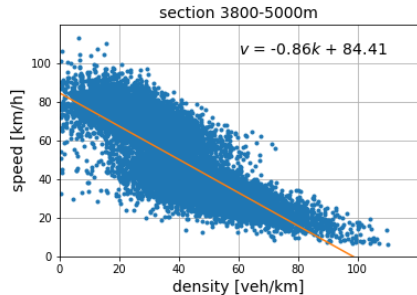
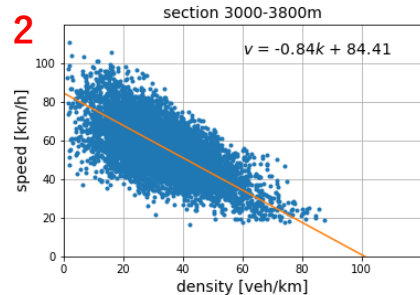
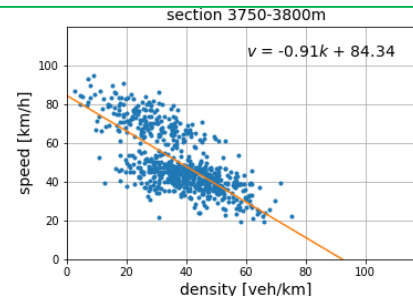
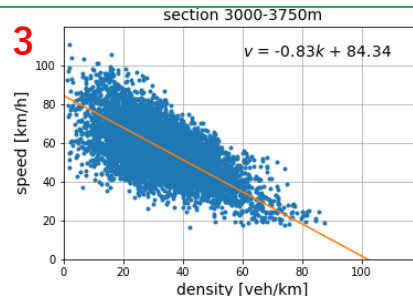
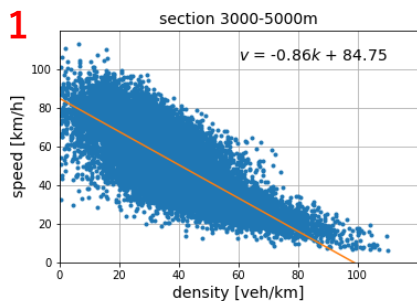
▼：車両検知器設置位置 (番号は設置車線)



考察：区間数少ないとき

■ 得られたFD

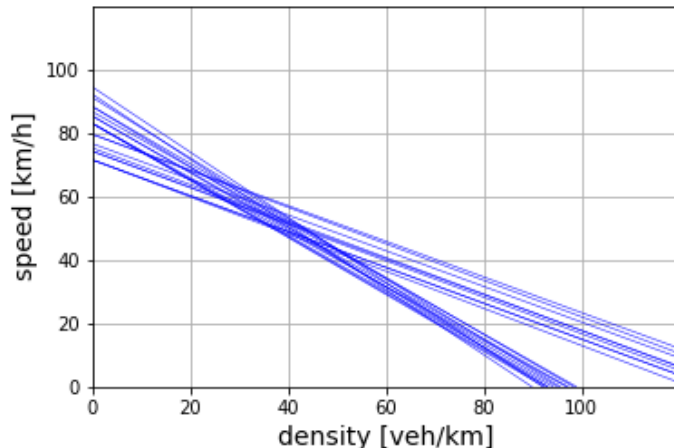
データが
分類された



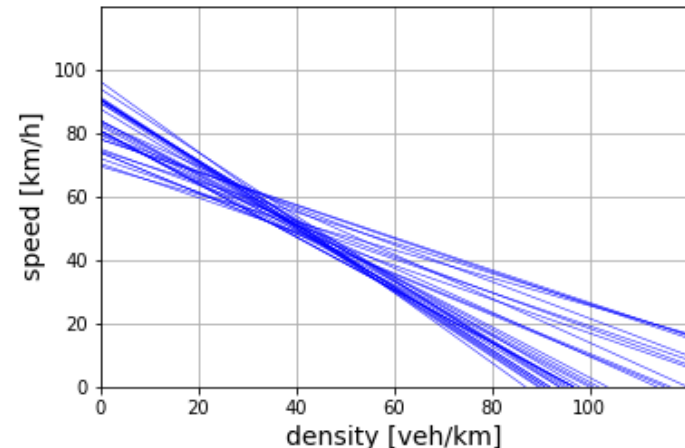
■ FDは2つに大別できそう

- 切片が下り勾配で大きく上り勾配で小さい
- 傾きは上流側で大きく下流側で小さい
- ただしこの境界は $x=3700$ くらいにある。区間数が少ないときとはやや異なる地点であり検討が必要

minimum BIC

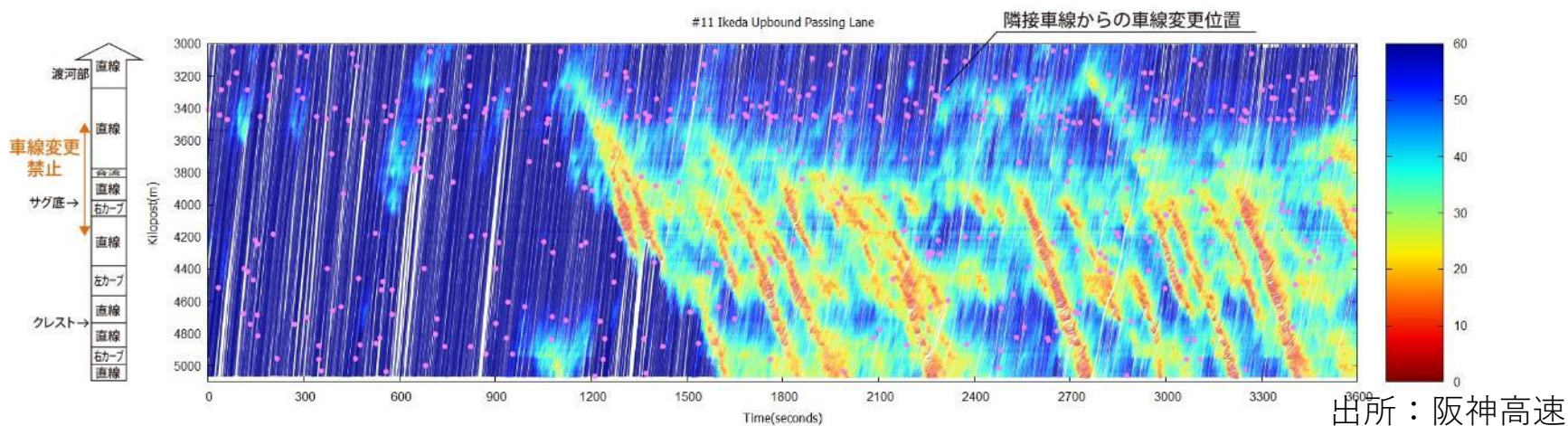


minimum AIC



■ タイムスペース図との比較

- ここまでに述べた傾向がおおむね表れている



■ 実交通量との比較

- $x=3800$ がボトルネックだと考えても不自然ではない

時間帯	$x = 3850$		$x = 3800$		$x = 3750$	
	交通量	1時間換算値	交通量	1時間換算値	交通量	1時間換算値
7:15:00-7:17:29	91	2184	80	1920	79	1896
7:17:30-7:19:59	98	2352	84	2016	86	2064
7:20:00-7:22:29	79	1896	65	1560	64	1536
7:22:30-7:24:59	80	1920	71	1704	70	1680

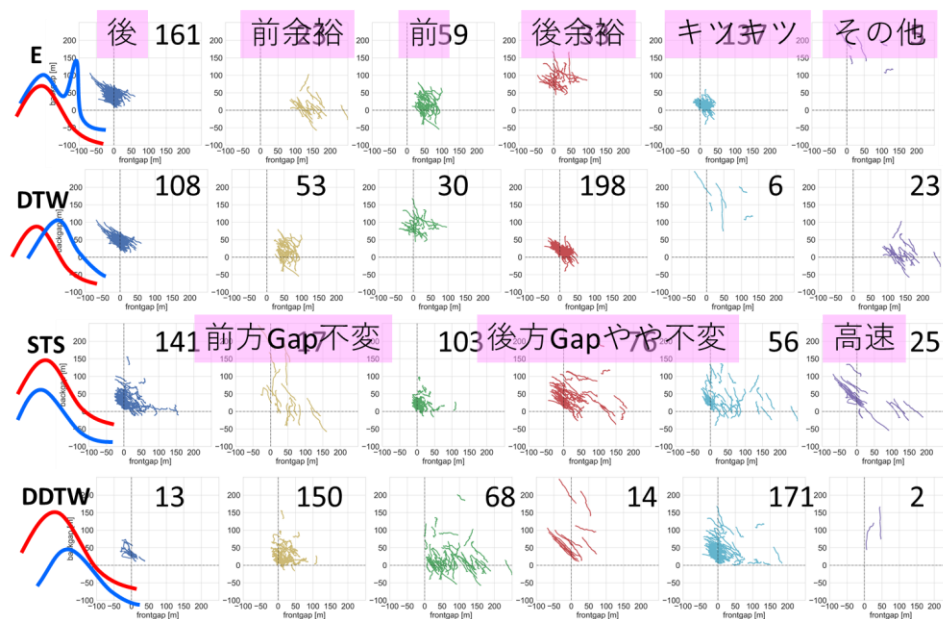
■ 区間別FD: スパースモデリングによるデータ分類

- プローブデータとの統合 (小田+2023)

■ 車線変更挙動分析

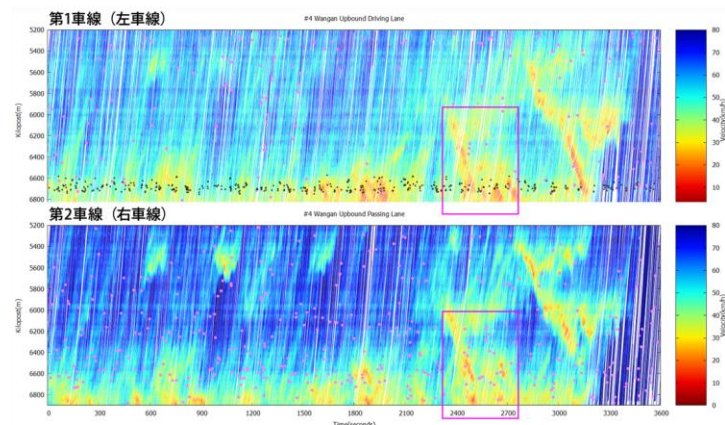
- 統計的クラスタリングによるデータ分類 (打越, 中西2023)

前方・後方Gap, クラスタ数6

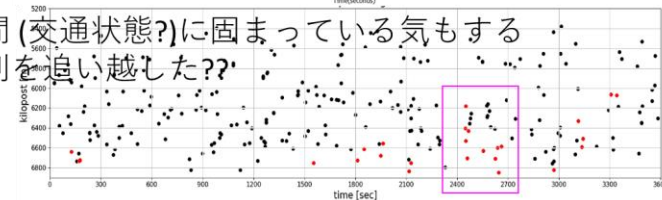


TS図上で見ると

- K=6,DTW のとき
- ある1つの cluster [前余裕]



特定の時空間(交通状態?)に固まっている気もする
頑張って車列を追い越した??

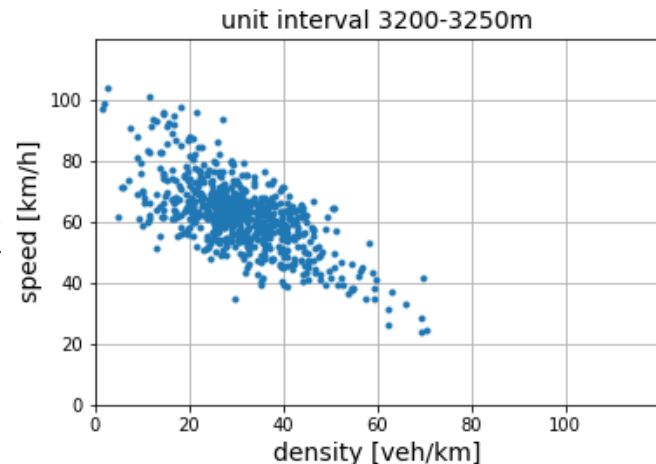


2. 新しい(?)課題

- データが高解像度になったので、より多くのことが分かるようになった！ とは一筋縄ではいかない
- いくつかの問題
 - (a) 観測データが高解像度 \neq 観測データが高精度:
ノイズだらけのデータになる可能性
 - (b) 観測データが高解像度 \neq 予測モデルが高解像度:
見かけの予測精度が悪化する可能性
 - (c) 観測データ量が増大 \Rightarrow 複雑な予測モデルを支持:
同じ現象なのにモデルが複雑化

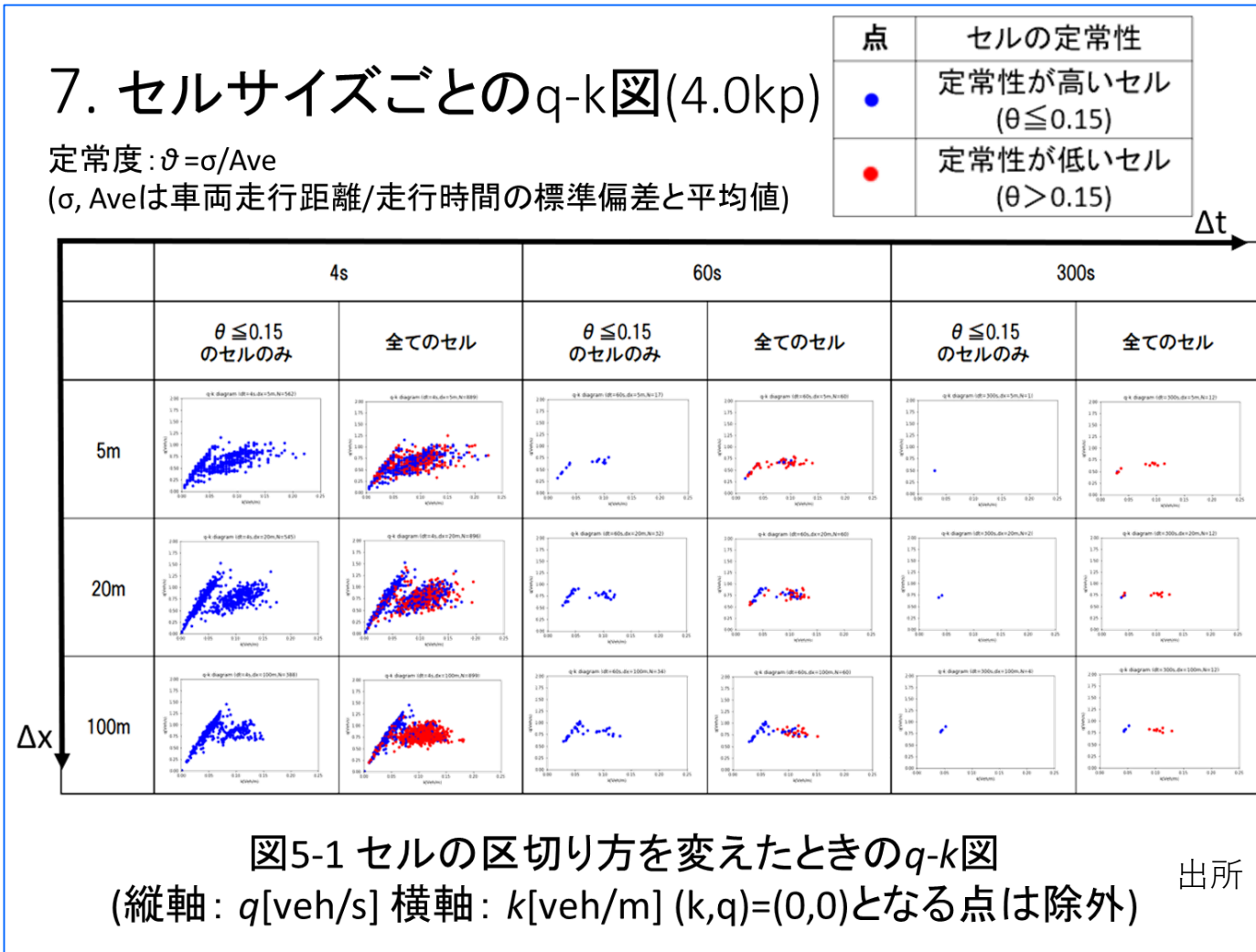
(a) 高解像度 ≠ 高精度

- 定常判定: これは定常状態?
- 一般に見かけるk-v図
=カウンタのデータで作成したk-v図は
もっときれい.
- つまり,
全車両軌跡・Edieの定義・小さい時空間(例: 5秒)は,
断面交通量・地点速度・長い時間集計単位(例: 1分)
よりもばらつきが大きくなってしまふ
- より厳密・正確と思われるデータ操作が, 低精度に見える
結果を生み出す場合がある
- 個別の問題としては, より正確な定常判定の追求(小暮2022)



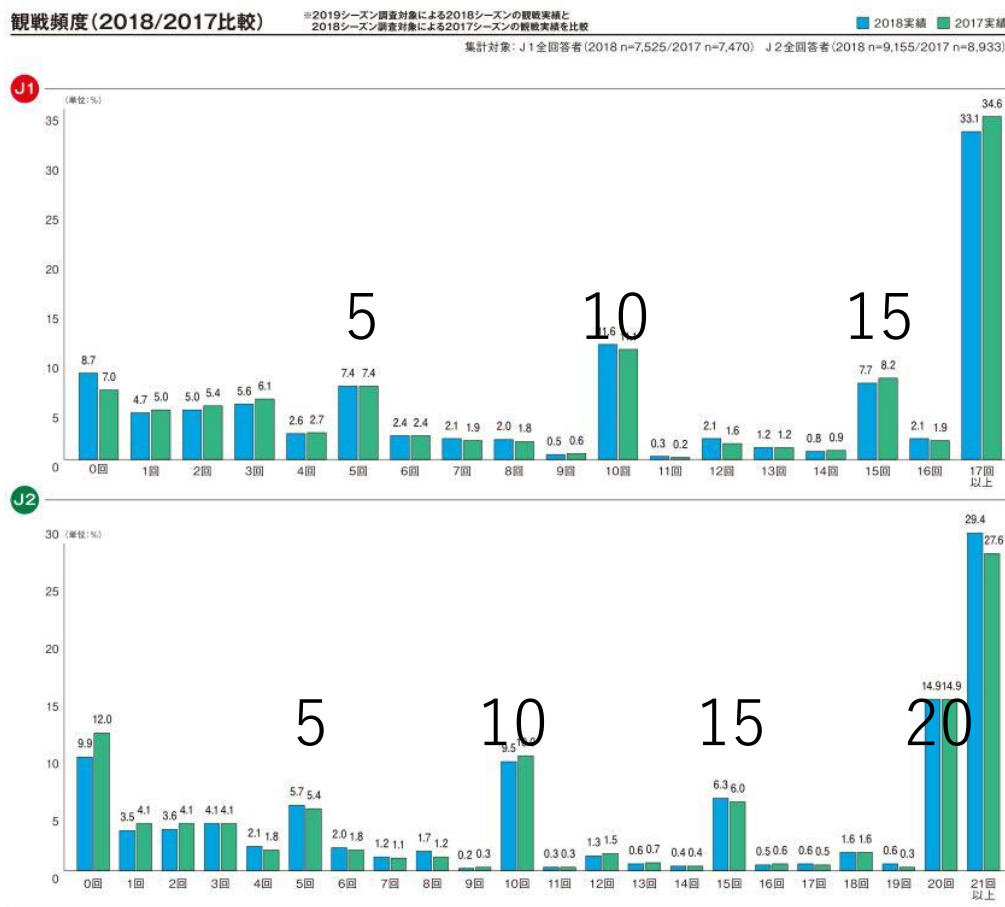
(a) 高解像度 ≠ 高精度

- ちなみに、時空間が大きければEdieの定義でも十分きれい



- こんなことなら粗いデータを使っていたら良かった...???

- 調査における出発時刻の回答 (Sato, Maruyama 2020)
- Jリーグ年間観戦回数



(b) 観測の解像度 ≠ 予測の精度

■ 先ほどの検証

- 区間別FDがあれば交通状態の波及が詳細に予測できるはずだが...

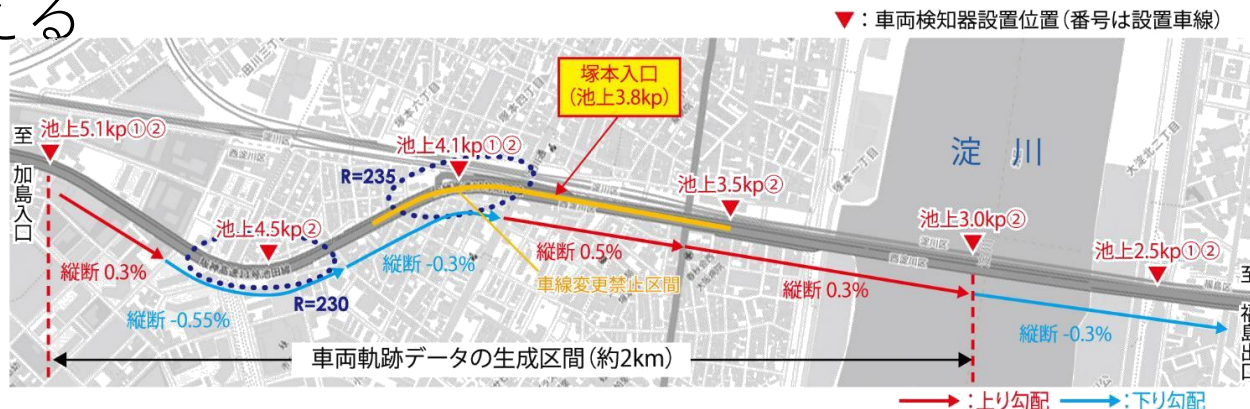
■ 実交通量との比較

- $x=3800$ がボトルネックだと考えても不自然ではない

時間帯	$x = 3850$		$x = 3800$		$x = 3750$	
	交通量	1時間換算値	交通量	1時間換算値	交通量	1時間換算値
7:15:00-7:17:29	91	2184	80	1920	79	1896
7:17:30-7:19:59	98	2352	84	2016	86	2064
7:20:00-7:22:29	79	1896	65	1560	64	1536
7:22:30-7:24:59	80	1920	71	1704	70	1680

■ 一方で、そもそも全車両軌跡でなければ発生しない問題

- 例: カウンタだけでFDを推定して交通量を予測すると、渋滞だけを議論するならばむしろ高精度
- 検証に使える観測データが増えたために予測精度が悪くみえる



出所: 阪神高速

たとえば出発時刻だと

- 1時間単位で観測して，1時間単位で予測するモデル

Sample #	Observed Dep. time	Estimated Dep. time	Residual [min.]
1	7時台	7時台	0
2	8時台	7時台	60
3	6時台	6時台	0
4	7時台	7時台	0
5	7時台	8時台	60
6	8時台	8時台	0
...

- 「n時台」をn時30分で代表させるとすると，
残差二乗和 7200 min^2

- 観測の解像度が上がって1分単位になった場合

Sample #	Observed Dep. time	Estimated Dep. time	Residual [min.]
1	7:13	7時台	17
2	8:17	7時台	47
3	6:53	6時台	23
4	7:05	7時台	25
5	7:32	8時台	58
6	8:11	8時台	19
...

- 「n時台」をn時30分で代表させるとすると、
残差二乗和 $7377 \text{ min}^2 > 7200 \text{ min}^2$ (前ページ)
- 観測も1時間単位だった場合より精度が悪く見える

- 予測モデルも1分単位になった場合でも

Sample #	Observed Dep. time	Estimated Dep. time	Residual [min.]
1	7:13	7:40	27
2	8:17	7:34	43
3	6:53	6:25	28
4	7:05	7:27	22
5	7:32	8:24	52
6	8:11	8:38	27
...

- 残差二乗和 $7279 \text{ min}^2 > 7200 \text{ min}^2$ (前々ページ)
- 予測モデルに使える変数が増えない限り，予測精度が上がるとは限らない
- むしろ，もともとそんなに合っていなかったことが分かる

さらに...

- 本人が7時30分に出発したと思っているのに，7時33分と特定し，精度良く予測する意味は？
 - 交通需要予測:
 - とても重要，数分の違いで渋滞したりしなかったり
(注: 実際に分単位の精度で交通量予測できるかは別問題)
 - Well being 的なこと:
 - むしろ本人の認識が重要では？
- これまでは，予測精度に比べて観測解像度が低いことが多かったので，解像度を上げられるだけ上げたかった
- これからは，いくらでも解像度を高くすればよいわけではない(かもしれない)
 - 何ができるようになるかが重要(だと中西は思う)

(c) 複雑なモデルの支持

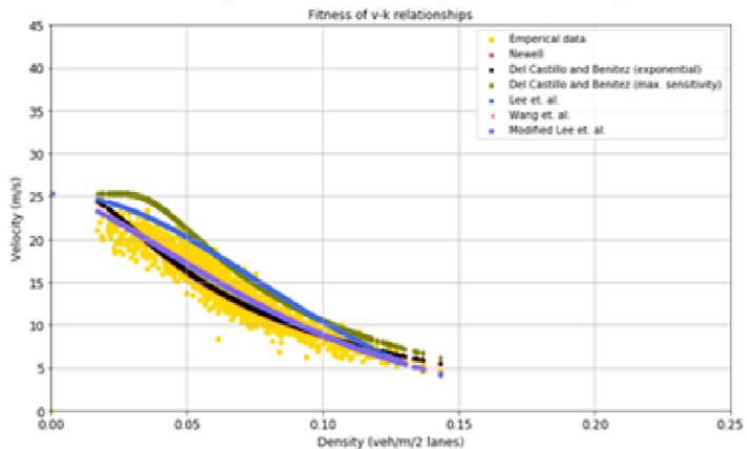
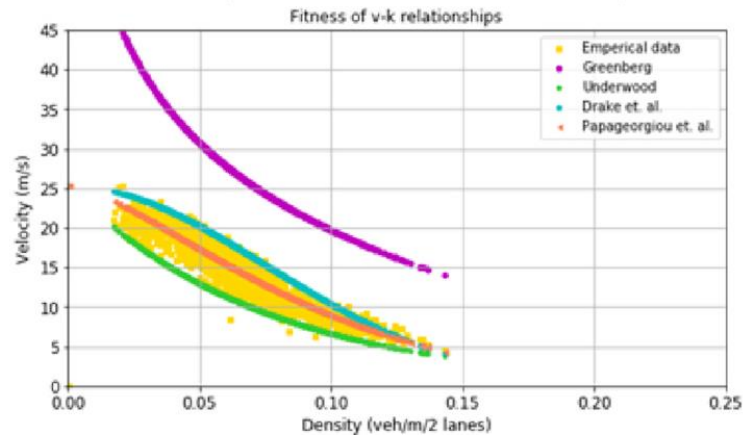
- 一般的な統計的検定(t 検定, AIC, BIC, ...)は, データ量が増えるにしたがって複雑なモデルを支持するようになる
- データが高解像度になる
 - データ量が増えることも多い
 - より複雑なFD関数を適用すれば当てはまりが良くなる
- 区間別FDのモデルはいくらでも複雑にできる
 - 区間ごとにFD関数形も異なる場合
 - 自由流と渋滞流で異なる関数形に従う場合
 - ...
- それで良い??

(c) 複雑なモデルの支持

■ 複雑なFD関数のほうが「良い」と判断されやすい

Table.1. Speed-density functional relationships

Author	Functional Form	Parameters
<i>Single-regime models</i>		
<i>Linear</i>		
Greenshields et al. (1935)	$v = v_f \left(1 - \frac{k}{k_j}\right)$	v_f, k_j
Drew (1968)	$v = v_f \left[1 - \left(\frac{k}{k_j}\right)^m\right]$	v_f, k_j, m
Pipes (1967)	$v = v_f \left(1 - \frac{k}{k_j}\right)^n$	v_f, k_j, n
May and Keller (1967)	$v = v_f \left[1 - \left(\frac{k}{k_j}\right)^{m,n}\right]$	v_f, k_j, m, n
<i>Logarithmic</i>		
Greenberg (1959)	$v = v_m \ln \frac{k_j}{k}$	v_m, k_j
<i>Exponential</i>		
Underwood (1961)	$v = v_f \exp\left(-\frac{k}{k_m}\right)$	v_f, k_m
Drake et al. (1967)	$v = v_f \exp\left[-\frac{1}{2}\left(\frac{k}{k_m}\right)^2\right]$	v_f, k_m
Papageorgiou et al. (1989)	$v = v_f \exp\left[-\frac{1}{a}\left(\frac{k}{k_m}\right)^a\right]$	v_f, k_m, a
<i>Complex</i>		
Newell (1961)	$v = v_f \left\{1 - \exp\left[-\frac{\lambda}{v_f} \left(\frac{1}{k} - \frac{1}{k_j}\right)\right]\right\}$	v_f, k_j, λ
Del Castillo and Benitez (1995)	<i>Exponential curve</i> $v = v_f \left\{1 - \exp\left[\frac{ C_j }{v_f} \left(1 - \frac{k_j}{k}\right)\right]\right\}$	v_f, k_j, C_j
	<i>Maximum sensitivity curve</i> $v = v_f \left\{1 - \exp\left[1 - \exp\left(\frac{ C_j }{v_f} \left(\frac{k_j}{k} - 1\right)\right)\right]\right\}$	
Lee et al. (1998)	$v = \frac{v_f \left(1 - \frac{k}{k_j}\right)}{1 - E \left(\frac{k}{k_j}\right)^\theta}$	v_f, k_j, E, θ
Wang et al. (2013)	$v(k, \theta) = v_b + \frac{v_f - v_b}{\left[1 + \exp\left(\frac{k - k_t}{\theta_1}\right)\right]^{\theta_2}}$	$v_f, k_t, v_b, \theta_1, \theta_2$
Modified Lee et al.'s model (2019)	$v = \frac{v_f \left[1 - \left(\frac{k}{k_j}\right)^\alpha\right]}{1 + E \left(\frac{k}{k_j}\right)^\theta}$	$v_f, k_j, E, \theta, \alpha$



■ それで良い??

- 統計的にはそれで良い(それを示す理論もある)
- 工学的には?

■ 問: 複雑なモデルで誰のどんな行動が当たるようになるか?

- 平均的な人の平均的な行動がたくさん当たれば、モデルの精度は上がったと判断されやすい
- このとき、高解像度に多様なデータを得た結果(情報量)は、主にボリュームゾーンの表現力向上に使われる
- この意味での「精度」が上がれば、珍しい人の行動は当たらなくて良い?

- Dahiya, G., Asakura, Y., Nakanishi, W.: A Study of Speed-Density Functional Relations for Varying Spatiotemporal Resolution Using Zen Traffic Data, IEEE 23rd International Conference on Intelligent Transportation Systems, Web conference, 2020.09.20-23.
- Dahiya, G., Asakura, Y., Nakanishi, W.: Analysis of the Single-Regime Speed-Density Fundamental Relationships for Varying Spatiotemporal Resolution Using Zen Traffic Data, Asian Transport Studies, Vol.8, 100066, 2022.
- Sato, Y., Maruyama, T.: Modeling the rounding of departure times in travel surveys: Comparing the effect of trip purposes and travel modes, Transportation Research Record, Vol. 2674 (10), pp.628-637, 2020.
- 朝倉康夫, 吉田光太郎, 中西航, 瀬尾亨: 時空間セルのサイズはプローブカーデータの集計結果にどのような影響を及ぼすか? -阪神高速道路ZTDデータからわかること-, 第11回高速道路の交通データ利用に関する勉強会, 北海道建設会館, 北海道札幌市, 2019.09.29.
- 打越天真, 中西航: 連続走行軌跡を用いた車線変更挙動の時系列クラスタリング, 第43回交通工学研究発表会論文集, CD-ROM, 2023.
- 小田紘生, 瀬尾亨, 中西航: コネクティッドカーデータとスパースモデリングに基づく区間別流率密度関係の推定, 第43回交通工学研究発表会論文集, CD-ROM, 2023.
- 小暮遼河: 全車両軌跡データを用いた交通流基本図のベイズ推定, 東京工業大学修士論文, 2022
- Jリーグ: スタジアム観戦者調査2019 サマリーレポート, <https://aboutj.jleague.jp/corporate/management/spectator-survey/>
- 中西航: 全車両軌跡データとスパースモデリングによる区間別Fundamental Diagramの推定, 土木学会論文集D3, Vol.78, No.1, pp.24-33, 2022.
- 阪神高速道路株式会社: Zen Traffic Data, <https://www.zen-traffic-data.net/>