
ビッグデータ活用による旅客流動分析実証実験事業

都市全体の人流・交通動態の
動的シミュレーションモデルの構築

株式会社 瀬戸内

事業の概要

●目的

携帯電話の位置情報(ODデータ)を活用し、都市全体のシミュレーションモデルを構築する

現在、大都市だけでなく、多くの地方都市においても、朝夕の通勤時や連休の際に、慢性的な渋滞が発生している。また、特に地方都市においては、自動車「生活の足」という側面が強く、自動車を中心とした生活スタイルが築かれており、大都市圏と比較した際に、あまり公共交通機関の利用が活発とは言えない。

しかし、こういった現状は、地方都市居住者であれば、肌感覚としては理解しているものの、正確に現状を把握できているとは言い難い。そこで、地方都市の1日における“人”と“交通”の動きを、携帯電話の位置情報(ODデータ)と日本道路交通情報センター：JARTICの断面交通量情報(以降、トラカンデータ)にて把握し、さらに、機械学習(ML)を用いることにより都市全体のシミュレーションモデルを構築する。

●対象地域

愛媛県内(松山市、東温市、松前町、砥部町、伊予市)

構築したモデルを利用して、以下の課題解決の一助とすることを目的とする。

公共交通指向型の都市開発(TOD)の推進

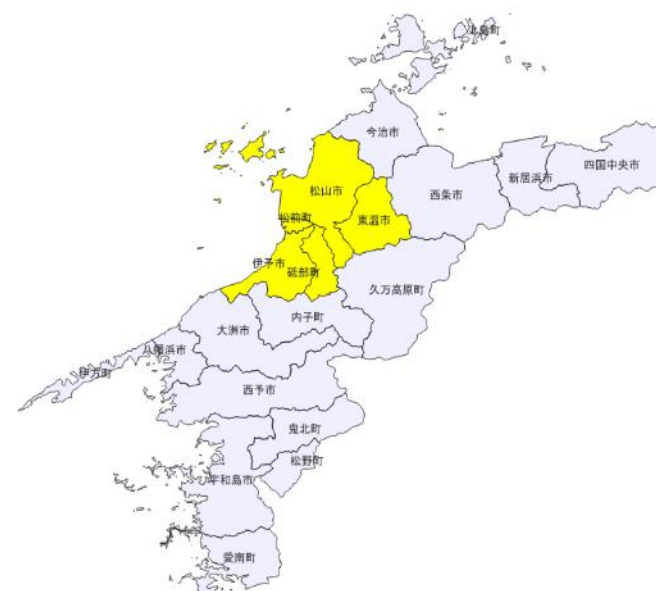
→自動車(特に自家用車)中心型の都市からの転換を促し、公共交通指向型にむけた地球環境に優しい街づくりを目指す

日々の生活における交通手段や利用ルート最適化

→交通が混雑しているルートや輻輳する箇所では、交通事故が発生しやすい

人流予測を用いた地域産業の活性化

→時間帯毎の動的な人流予測を用いることにより、地域住民および観光客に最適なアプローチを図る



解決を目指す課題の概要

●課題の概要

交通量把握の課題

トラカンデータ等にて断面交通量を実測している路線では交通量の把握や時間帯ごとの予測はできるが、実測していない路線については交通量の把握は困難である。

実測していない路線について以下のような手順で確認しており、交通量に対する情報が不足している状態で予測している。

①実測値からの経験則

実測値から見て交通の流れから概算で交通量を求める。

②過去からの踏襲

過去データをそのまま踏襲し、変更を加えないようにする。

③交通量予測の公式による一意の係数での計算

交通量の予測をする公式が存在するため、そちらに係数を加えて計算した結果をそのまま採用する。



そこで、KDDI Location Data(OD Data)を利用して“交通”の動きを把握できるようにすることで、交通量に対する情報が不足していた路線について予測できるようにする。

交通量を予測する機械学習モデル(ML)を構築する。



- ● ● 主要路線
- ● ● 実測箇所

実証実験の取組内容

●使用データの概要（特徴、採用理由）

(1) KDDI Location Data (OD Data)

【特徴】

KDDI社の携帯GPS位置情報を基に125mメッシュ単位での移動人口を算出し、独自のカスタマイズを施した統計データ

【採用理由】

125mメッシュ単位という詳細なODデータにて移動内容が把握できる

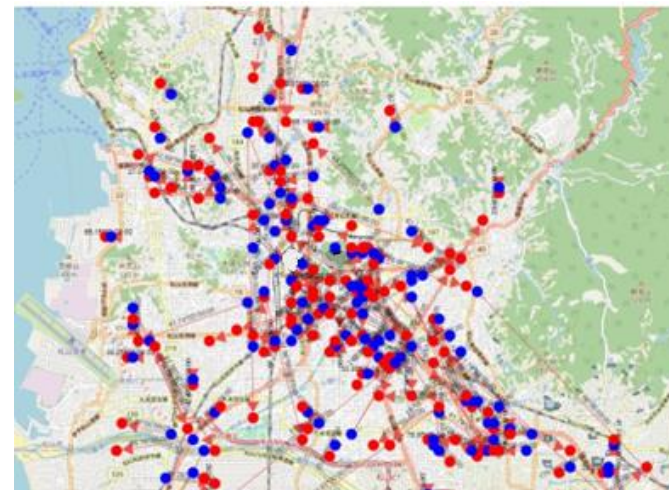
(2) 日本道路交通情報センター：JARTICの断面交通量（トラカンデータ）

【特徴】

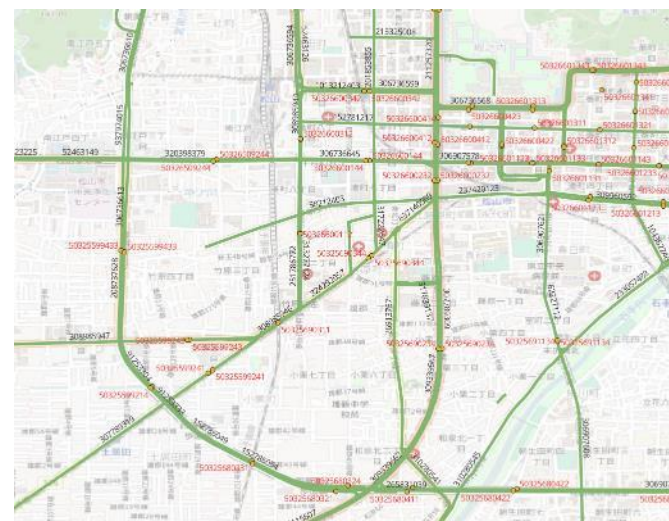
各都道府県警察が車両感知器などの計測機器で収集した断面交通量に関する情報を警察庁においてとりまとめたデータ

【採用理由】

実測値として比較的多数の箇所にて交通量が確認できる。



●●● 出発点
●●● 到着点
対象データ「S：自動車」



実証実験の取組内容

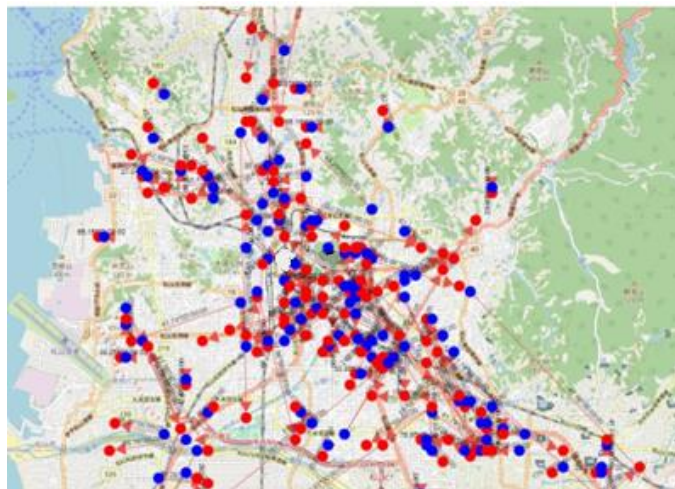
●使用データの概要（OD Data）

KDDI Location Data（OD Data）は
右の収集条件を指定し、個別に依頼した。

しかしながら、当初ご提供いただいたKDDI Location Data（OD Data）では滞在と判定されたデータの割合が多く、移動と判定されたデータが少なかったためデータ解析には不十分な状態であった。

そのため、KDDI様と協議を行い調整していただいたことで、移動と判定されたデータが大幅に増加した。

ODデータの分布は2022/04/01 00:15:00で以下となった



●●●● 出発点
●●●● 到着点
対象データ「S：自動車」

■データ定義

【対象エリア】 松山市・東温市・松前町・砥部町・伊予市
【対象期間】 2021年10月,2022年4月(2か月分のみ)
【集計日単位】 日別
【メッシュ粒度】 125m
【集計時間単位】 15分
【移動手段】 あり
【性別】 なし
【年代】 なし
【居住地/勤務地】 なし
【滞在判定時間】 あり(15分)

KDDI_LocationData

2021年10月分 & 2022年4月分混在

種類	年月	旧件数	新規受領分	増加(倍)
区分なし				
H				25
R	202110		59539	4.929
S	202110	779401	3358376	4.309
W	202110	157167	716722	4.560
K	202110	126477	329070	2.602
区分なし			177096	
H	202204	2197	31424	14.303
R	202204	12192	60816	4.988
S	202204	655890	2966097	4.522
W	202204	127717	591831	4.634
K	202204	101912	259084	2.542

※Kはメッシュ移動ありのみ

事業の概要

●活用手法

携帯電話の位置情報としてKDDI Location Data (OD Data) データを利用し、以下のように構築処理を実施した。

① データ解析、加工

以下の3つの手順にてデータのクリーニングおよび解析処理を実施

1. 位置情報の付与

KDDI Location Dataには、位置情報は含まれていない。そこで、Location Dataの125mメッシュに対し、位置情報を付与した。

2. 移動手段および平休日で分割

移動手段や平休日毎でデータ特性があるため、まず、Location Dataを移動手段(「S:自動車」)に絞り込み、さらに平休日で分割した

3. データクリーニング

主にトラカンデータにて降雪や事故などの特殊な渋滞状況を省くため、各メッシュにおける集計単位(15分)毎の外れ値を除去した。
(Z-scoreを算出し、一般的な外れ値とされる3以上の値をもつものを除去)

② 機械学習(ML)によるモデル作成

「行=時間」、「列=到着メッシュエリア」の行列(matrix)データから、LSTM AutoEncoderを使い特徴量を抽出し、シミュレーションモデルを作成した。

③ 路線データ割り付け

KDDI Location Dataとトラカンデータを関連付けするため、断面交通量を主要路線ごとに125mメッシュ区切りで距離案分して計算し、都市全体のシミュレーションモデルに対して割り付けとシミュレーションモデル再作成を実施した。

④ 実地検証

都市全体の交通の流れを把握するため、携帯GPS位置情報を基にシミュレーションモデルの構築を行い、モデルによる予測結果の整合性の確認再作成を実施した。

分析手法詳細と分析結果

●分析手法

① データ解析、加工

以下の3つの工程を実施した。

2-1. 位置情報の付与

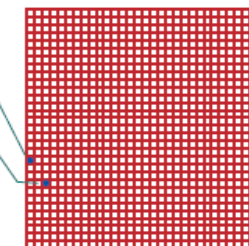
KDDILocation Dataには、位置情報は含まれていない。

そこで、まず、Location Dataに対し、位置情報を付与する必要がある。

今回、利用するデータの粒度が125m×125mであるため、メッシュ作成アドオンにて、8分の1地域メッシュ(第6次メッシュ)を作成し、その中心点をLocation Dataの位置情報とした。

項目名	内容	例
yyyyymm	年月	202204
mesh_size	メッシュサイズ	125
yyyyymmdd	年月日(yyyyMMdd)	20220401
start_mesh_area	出発メッシュエリア	533925463
start_city	出発市区町村コード(総務省 全国地方公共団体コード6桁の下1桁を含まない5桁)	13111
start_time	出発時間帯 (h:mm)	9:15
end_mesh_area	到着メッシュエリア	533925465
end_city	到着市区町村コード(総務省 全国地方公共団体コード6桁の下1桁を含まない5桁)	13111
end_time	到着時間帯 (h:mm)	9:30
ODRouteType	移動手段 (S:自動車)	S
population	推計人数(15分)	900,63

KDDILocation Data 仕様



8分の1地域メッシュ (125mメッシュ)

2-2. 移動手段で分割

移動手段でデータ特性があるため、Location Dataを移動手段(「S:自動車」)に絞り込み分割した。

2-3. データクリーニング

主にトラカンデータにて降雪や事故などの特殊な渋滞状況を省くため、各メッシュにおける集計単位(15分)毎の外れ値を除去した。

(Z-scoreを算出し、一般的な外れ値とされる3以上の値をもつものを除去)

移動手段	加工前	z-score除去	除去率
W=歩行	加工前	284881 行	
	z-score除去	279161 行	2.007856 %除去
S=自動車	加工前	1435270 行	
	z-score除去	1399727 行	2.4764 %除去
K=不明	加工前	228389 行	
	z-score除去	224228 行	1.821892 %除去

時刻	メッシュコード		外れ値
	50000000001	50000000002 ..	
2021/10/01 9:00	25	10	
2021/10/02 9:00	201	3	外れ値
:			
2021/10/31 9:00	37	110	

分析手法詳細と分析結果

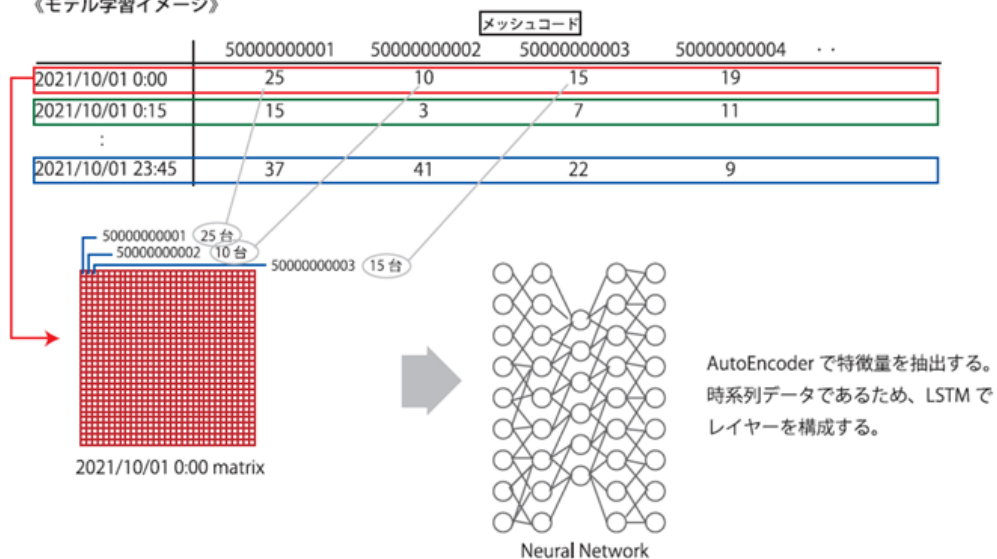
●分析手法

② 機械学習(ML)によるモデル作成

モデルはLSTM AutoEncoderの機能を使用する。
LSTMは別途洪水予測モデルにも活用例があり、交通量の予測にも応用できるのではないかと考えた。

「時間毎の交通量の特徴」および「各メッシュ間の相関関係」といった特徴から少ない計測箇所の計測値からエリア全体の交通量を推測することができる。
(解析するためのNeuralNetworkレイヤーをメッシュ数および時間数分新しく構築し、それぞれのつながりを全パターン網羅して実施する形式)

《モデル学習イメージ》



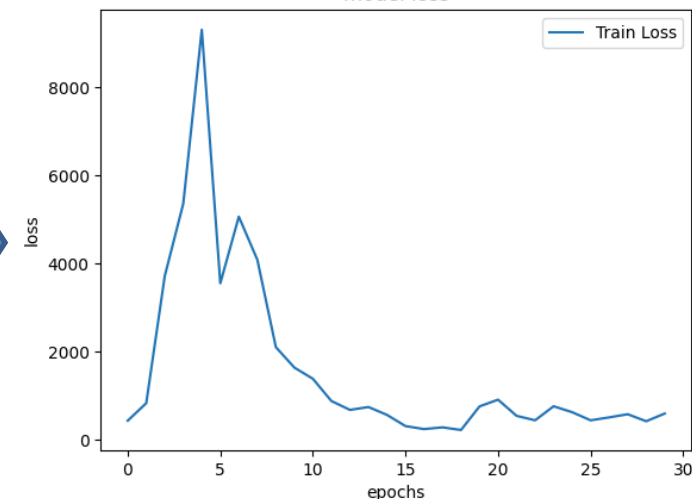
Model: "sequential_5"

Layer (type)	Output Shape	Param #
lstm_16 (LSTM)	(None, 24, 128)	11546112
dropout_4 (Dropout)	(None, 24, 128)	0
lstm_17 (LSTM)	(None, 64)	49408
repeat_vector_4 (RepeatVector)	(None, 24, 64)	0
lstm_18 (LSTM)	(None, 24, 64)	33024
lstm_19 (LSTM)	(None, 24, 128)	98816
time_distributed_4 (TimeDistributed)	(None, 24, 22422)	2892438

Total params: 14,619,798
Trainable params: 14,619,798
Non-trainable params: 0

Lossが減少し、安定してモデル作成されている

model loss



分析手法詳細と分析結果

●分析手法

③ 路線データ割り付け(割り付け前モデル作成)

○KDDI Location Data (OD Data) のみによる交通量の把握の問題点
移動手段が「自動車」であるKDDI Location Data (OD Data) のみを対象とした場合、携帯電話利用者の端末情報であるため推定台数は以下のように不確定な数値である。

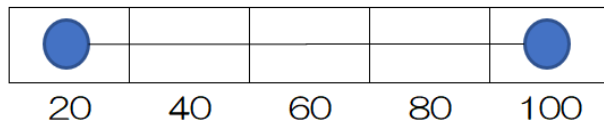
- ① 複数人が1台に乗車して移動していることもあるため、端末情報＝交通量ではない。
- ② GPSの移動速度から判断されており正確な台数ではない。

そのため、以下のようにKDDI Location Dataに実測値であるトラカンデータを組み合わせることで解決する。

①KDDI Location Dataのシミュレーションモデル作成（前述3.）
対象年月のKDDI Location Dataにてモデル構築を実施し、125mメッシュで区切られた都市全体の推定人数シミュレーションモデルを作成する。

②主要路線について手動で160箇所ほど指定したトラカンデータの測定箇所間について、125mメッシュ区切りで2点間の断面交通量の距離案分を行う。案分された値をそれぞれメッシュ単位で保管する。

例) 一点が20、もう一点が100とした場合の案分



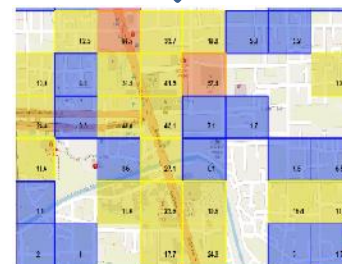
③ ②の距離案分結果を①のシミュレーションモデルで解析した結果に上書き反映する。

④ ③の結果を再度シミュレーションモデルにて解析する。

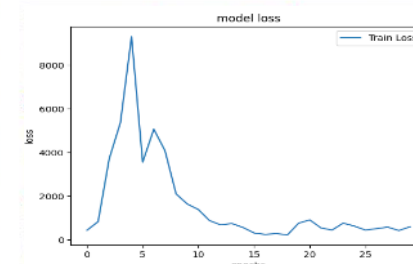
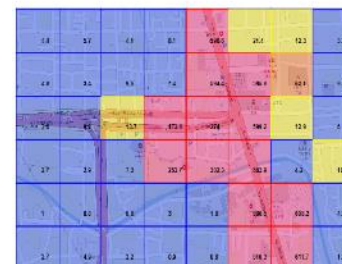
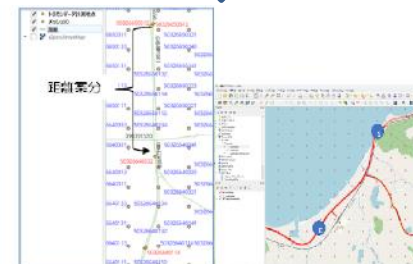


移動人数は2名であっても、車は1台

①KDDI ODモデル



②距離案分



③反映

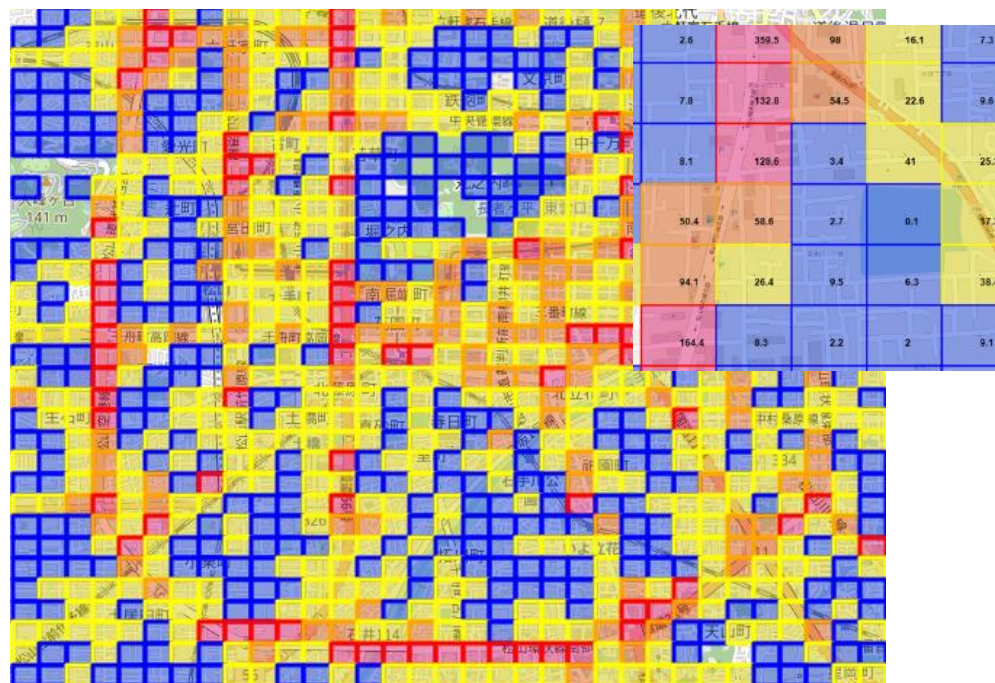
④再解析

分析手法詳細と分析結果

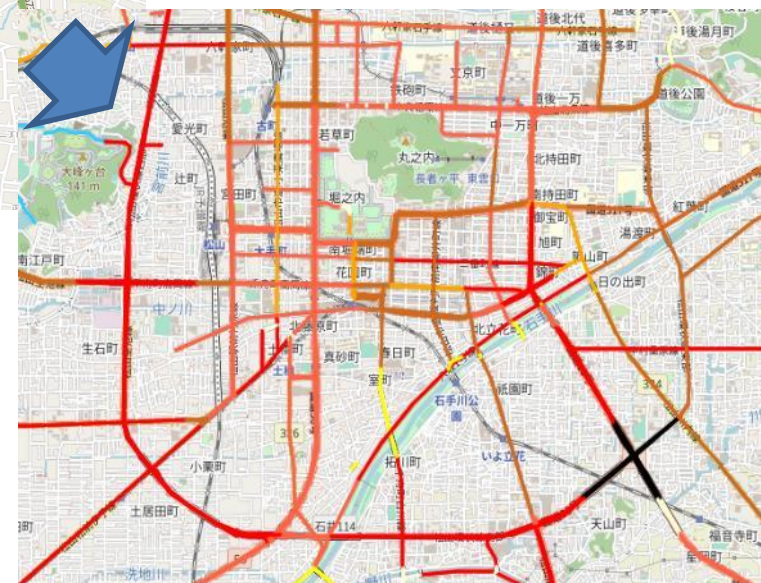
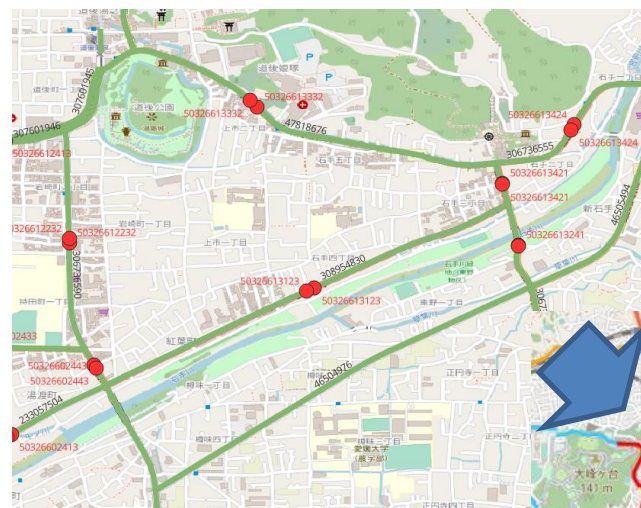
●分析手法

④ 路線データ割り付け(割り付け実行)

指定年月のデータを前述4.にて作成したトラカンデータモデルしたモデルにて解析する。以下のような125mメッシュごとの交通量が予測できる。



路線とメッシュを重ね合わせ、メッシュの値を路線に反映することで、主要路線上に実測値を反映する。



案分の値を組み合わせたことにより、傾向として主要路線付近に交通量が集中している

目的は路線の交通量であるため、メッシュを路線に反映させる。

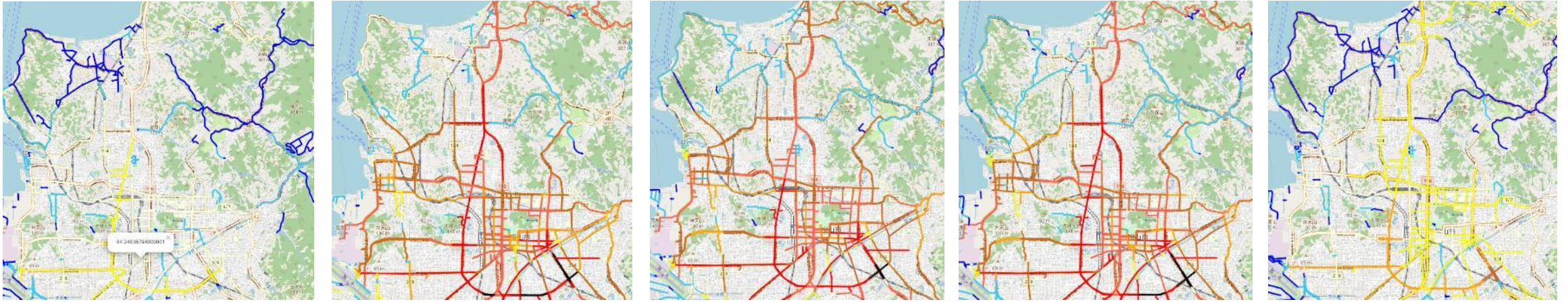
路線に交通量が反映される

分析手法詳細と分析結果

●分析結果

時間帯による交通量の変化

5:00、8:00、12:00、18:00、23:00の移り変わりを確認した



時間帯によって変化していることが確認でき、調査箇所1について近傍のトラカンデータで示した通り、深夜～早朝帯を除いて慢性的に交通量が多い状態が確認できた。

分析手法詳細と分析結果

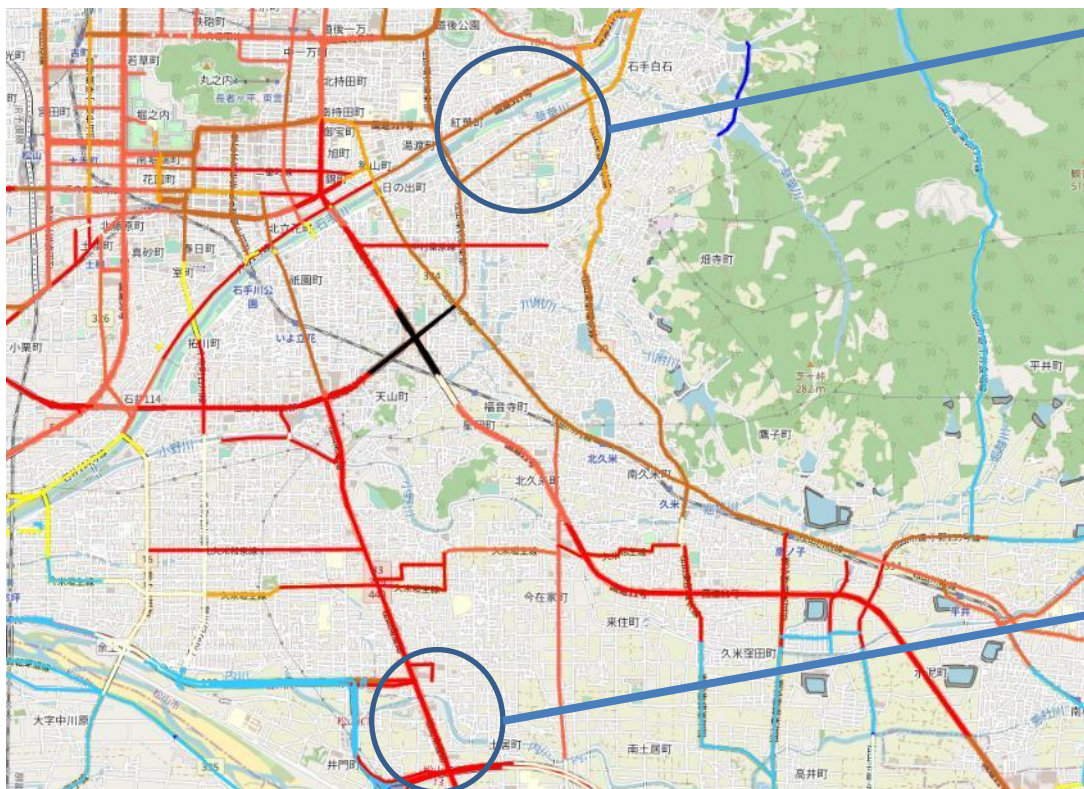
●分析結果

トラカンシミュレーションモデルの評価

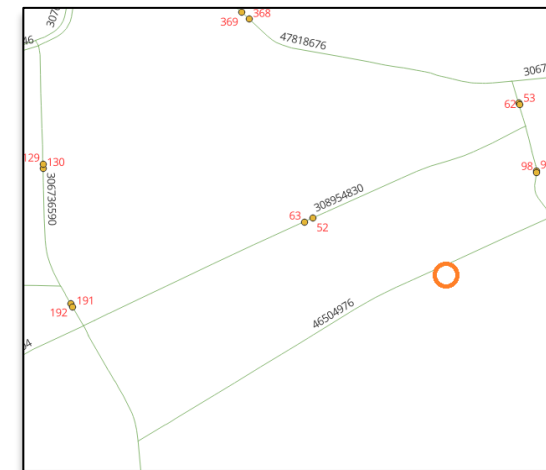
計測方法の異なる2拠点をとり出し、実地確認を実施した。

1箇所目は、トラカン計測の中間地点であり、路線の実測値の案分が正しく反映されているか。近傍のトラカン計測点の交通量と大きな差がでていないかを確認する。

2箇所目は、トラカンデータの実測を行っていない路線であり、交通量の多い箇所をピックアップした。



測定箇所2
(愛媛県松山市東野1丁目)



測定箇所1
(愛媛県松山市井門町)



分析手法詳細と分析結果

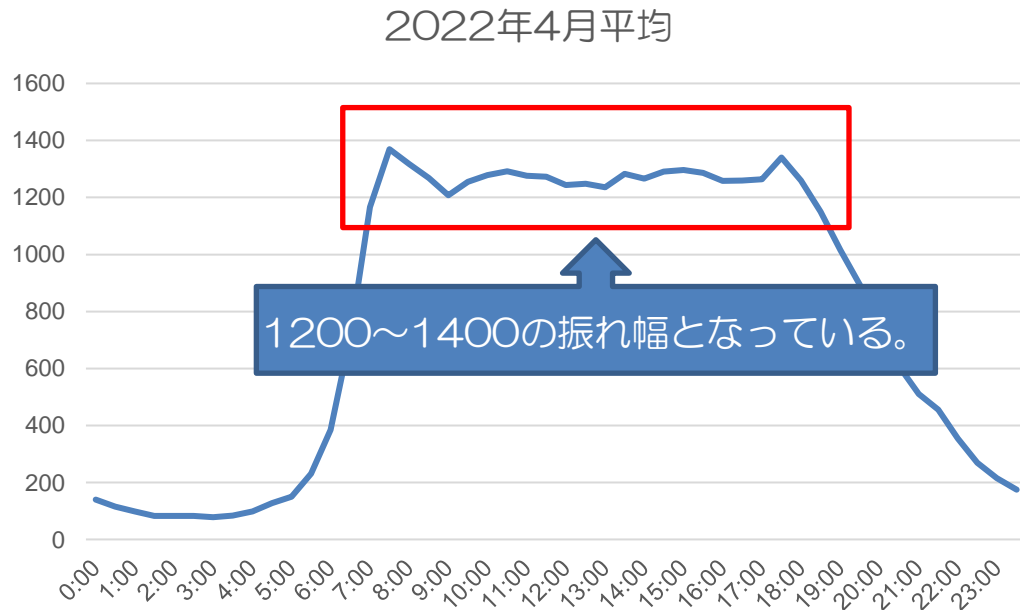
●分析結果

④実地検証（検証前モデル値の妥当性確認）

以下の2日間交通量を実測し、
2022年4月の交通量から作成したモデルと比較した。

- ①2023年2月28日（火） 晴れ
- ②2023年3月 2日（木） 午前中雨、午後晴れ

調査箇所1について近傍のトラカンデータを確認したところ、
2022年4月は平均して以下のように交通量が変化していた。



検証結果は以下の通りとなった。

調査箇所1	7:00	8:00	9:00	12:00	13:00	16:00	17:00	18:00
モデルによる予測値	1407	1457	1406	1463	1451	1464	1428	1437
実測値(2/28)	1467	1440	1288	1250	1270	1244	1314	1382
実測値(3/2)	1444	1406	1244	1185	1243	1197	1312	1301
2/28比							1.0868	1.0398

1400~1470の振れ幅となっている。

調査箇所2	7:00	8:00	9:00	12:00	13:00	16:00	17:00	18:00
モデルによる予測値	410	415	489	530	537	506	520	523
実測値(2/28)	302	390	343	436	399	437	402	445
実測値(3/2)	295	360	341	419	338	386	333	385
2/28比	1.3576	1.0641	1.4257	1.2156	1.3459	1.1579	1.2935	1.1753

もともとの平均値の振れ幅から推測して、
作成したモデルが指し示す予測値はほぼ誤差がないことから。。。

トラカンデータ実測値を持つ路線について、
モデルによる予測値は妥当な値を持っている
と推定できる。

分析手法詳細と分析結果

●分析結果

④実地検証（実測値の妥当性確認）

検証結果は以下の通りとなった。

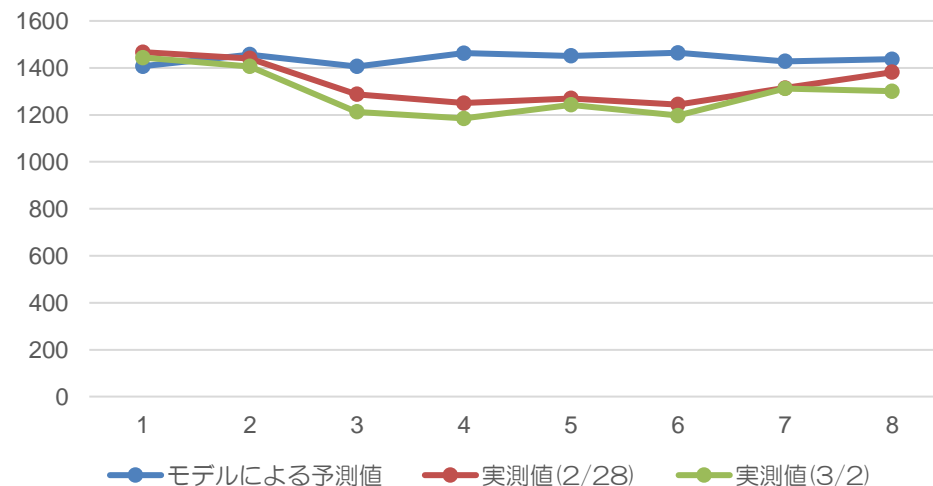
調査箇所 1	7:00	8:00	9:00	12:00	13:00	16:00	17:00	18:00
モデルによる予測値	1407	1457	1406	1463	1451	1464	1428	1437
実測値(2/28)	1467	1440	1288	1250	1270	1244	1314	1382
実測値(3/2)	1444	1406	1213	1185	1243	1197	1312	1301
2/28比	0.9591	1.0118	1.0916	1.1704	1.1425	1.1768	1.0868	1.0398

調査箇所 2	7:00	8:00	9:00	12:00	13:00	16:00	17:00	18:00
モデルによる予測値	410	415	489	530	537	506	520	523
実測値(2/28)	302	390	343	436	399	437	402	445
実測値(3/2)	295	360	341	419	338	386	333	385
2/28比	1.3576	1.0641	1.4257	1.2156	1.3459	1.1579	1.2935	1.1753

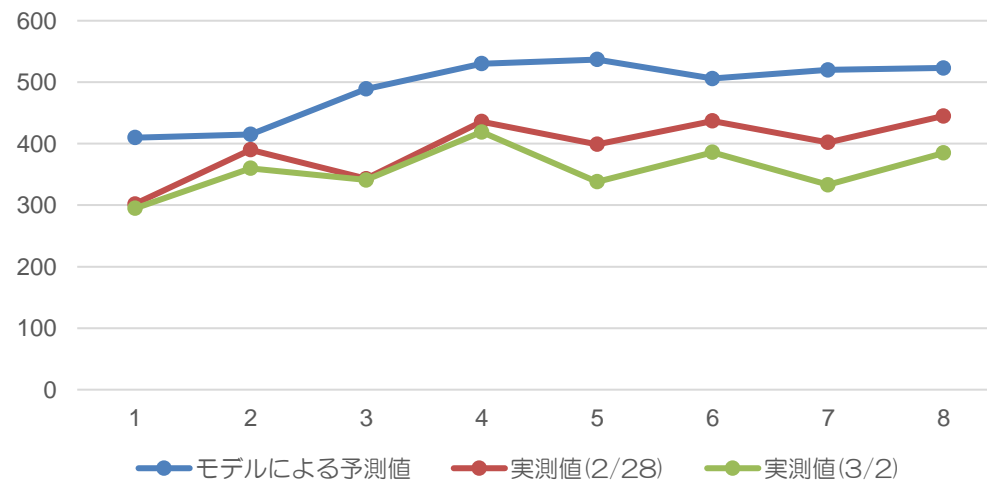
調査箇所 1 については実測値にかなり近づいた値となった。
一方、調査箇所 2 については誤差が生まれたが大きな乖離が発生していないことが確認できた。

トラカンデータ実測値を持たない路線について、
モデルによる予測値は妥当な値を持っていると
推定できる。

調査箇所 1



調査箇所 2



分析手法詳細と分析結果

●検証がうまくいった点

トラカンデータ実測値を持つ路線の交通量について想定通りの結果が出力された。
トラカンデータ実測値を持たない路線についても計測値と近い値が出力されることが確認できた。

●検証がうまくいかなかった点

トラカンデータ実測値を持たない路線について、実際の交通量と少々開きがあった。
周辺（山間）部等について交通量の調査・補完が必要と感じた。

●検証後に明らかになった課題

KDDI Location Data (OD Data) について2ヵ月間のデータで検証したため、
通年にわたる交通量データが推定できない問題があった。

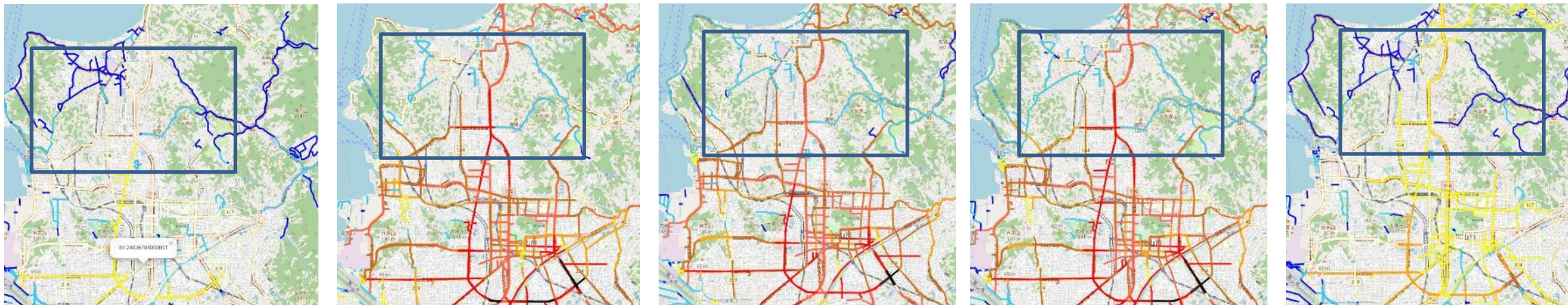
データ量を増やすなどを実施し、さらに交通量シミュレーションを研磨する
必要があると感じた。また、交通量に影響のある上り下りの判別を今回考慮していないため、
分類し傾向をつかむ必要があると感じている。
今後も引き続き精査を進めていきたい

分析手法詳細と分析結果

●補足

周辺（山間）部について交通量の調査・補完が必要と感じた理由

5:00、8:00、12:00、18:00、23:00の移り変わり



先程の交通量変化を再確認すると、トラカン観測点から大きく離れた箇所の交通量について体感より交通量が少なく表示されている。

主に山間部を越えるルートであり、地元では抜け道としてよく利用されているため交通量が多いが、それらが反映されていないように思われる。

実測値をとって確認し、トラカンデータの影響を調整してモデルを作成することで精度の改善が図れると思われる。

本事業の結果を踏まえた今後の活動予定について

●今後の活動予定

本実証実験事業で作成したモデルを活用して、団体（地方公共団体、商工団体、観光団体）へのさらなる適用を進めていくよう活動を行う。

具体的には以下の通り

- 1) 構築したモデルの活用に向けた取り組みとして、近日中（3月中旬）に開催する地域経済研究会時に参加者に対し本事業で作成したモデルを紹介後、アンケートを実施し定量的な評価指標を得るようにする。
- 2) 国土地理院による愛媛地域連携協議会において、発表の機会を頂きたいと思っている。
- 3) 構築したモデルのさらなるブラッシュアップとして、愛媛大学 社会共創学部とやり取りを行い、学生と連携してモデル改善を進めていけるよう調整を進めている。