

データフュージョンによる広域的な交通の 把握手法とその活用方法の検討

北海道開発局 建設部 道路計画課

○山崎 敦広
本田 卓己
草野 真一

近年、ETC2.0が本格運用開始されるなど、官民の双方で急速に交通系ビッグデータの蓄積が進んでおり、データの活用による、より定量的な交通施策の展開が期待されている。本稿では、既存の交通量常時観測データと交通系ビッグデータの一つである民間プローブデータを活用し、データフュージョンによる交通量の面的な把握手法の検討を行う。また、将来的に面的交通量データが実用可能となった場合の、施策への活用可能性について論じる。

キーワード：ビッグデータ、交通量、ETC2.0

1. はじめに

交通系ビッグデータは、鉄道・バス・ETC等のICカードの利用データや、カーナビゲーションシステム・携帯電話・ETC2.0・貨物車の電子タコグラフデータ等のGPSによる移動データまで多岐に渡るが、これらのデータはICTの進展とともに急速に収集・蓄積がなされ、社会現象の観測に資するデータセットとして注目されている。またビッグデータは、蓄積データを活用し、社会や地域の課題抽出だけでなく、課題解決策を導き出し、我が国の成長につなげていく政府の成長戦略¹⁾にも位置づけられている。本稿では、ビッグデータの活用が大きな注目を浴びている現状認識のもと、主に道路交通に関するビッグデータを用いて地域の現状の把握や課題の抽出を行うための基礎データとなる交通量を、広域的かつ面的に把握する手法の構築と、推計した交通データの施策への活用方策について論じることとする。

本稿の構成は次の通りである。2章で道路交通に関するビッグデータを概説し、3章にて本稿にて扱う分析対象データの特徴を述べる。4章では広域的な交通量の推計モデルの構築を行い、5章にて平成22年度道路交通センサスの交通量を使用して、モデルの精度の検証を行う。6章にて広域的な交通量データの活用シーンの検討を行い、7章にてまとめと今後の課題について整理を行う。

2. 道路交通に関するビッグデータ

交通量は、地域特性・地域課題・経済状況などを反映する基本的なデータセットである。交通量や車両の移動情報等データを取得する手法は様々であるが、我が国での代表的なデータとしては、古くは全国道路・街路交通

表-1 道路交通に関する代表的なデータの特徴

	対象車両	空間特性	時間特性	主な記録データ例
常観データ	道路上に設置したカウンターを通過した全車両	地点データ	1時間単位	集計値 時間帯別の車両の台数・速度を取得
民間プローブデータ	サンプル 会員登録された車両(約200万台)	面的データ	15分単位	DRM単位の所要時間、情報取得件数
ETC2.0データ	ETC2.0車載器がセットアップされた車両(約91万台)		秒単位	個車の時刻、位置情報(非集計値)、リンク単位の所要時間、台数(集計値)

情勢調査(いわゆる道路交通センサス)より始まり、その後、定点の交通量や速度を自動的に取得・蓄積する常時観測データ(以降、常観データ)が活用されてきた²⁾。近年では、カーテレマティクスの進展によりカーナビのGPS情報に基づくプローブデータ(以降、民間プローブデータ)の蓄積・活用が全国的に行われている³⁾。また、直近の大きな動向としては、ETC2.0車載器と道路上に設置された対応路側機が路車間通信して車両に蓄積された移動データを収集し、道路情報の提供を行うETC2.0サービスがスタートし、データの活用に期待が寄せられている^{4) 5)}。表-1は、道路交通に関する代表的なデータの特徴を整理したものであるが、特徴を考察すると、地点データであるが全車を取り扱うデータと、面的にデータを取得するがサンプルを取得するデータに大分される。またデータの時間的な分解能も秒単位～時間単位とそれぞれで異なり、取得データの記録に関しても、集計値として記録されたものと、個別データが記録されたものから

構成され、全数かサンプルか、定点か面的か、個別値か集計値か、という観点から相互にデータを補完しあう関係が成立している。したがって、これらの特性に着目すれば、データフュージョン（データの組み合わせ）により時間的に、空間的に高密度な交通データを推計する可能性があると推察される。また、民間プローブデータとETC2.0データのデータ量を比較すると、後述する通り、現段階で民間プローブデータの方が多いが、将来的には現行のETCがETC2.0車載器に置き換わる事を想定すると、ETC2.0データを活用すれば、民間プローブより高精度に交通量の推計が可能となる事も示唆される。

3. 分析対象データの特徴

本分析では、蓄積データ量の関係から、主に常観データと民間プローブデータを用い分析を進めるが、将来的にETC2.0車載器が普及し、ETC2.0データに置き換わる事も想定し、常観データ、民間プローブデータに加えてETC2.0データの取得状況についても整理を行い、ETC2.0の普及見込みと現段階におけるデータの取得状況についても把握する。なお、広域的な交通量推計については、平成27年10月のデータを用い分析を行う事とする。

1) 各データの取得体制と収集状況について

常観データとETC2.0データは、定点でデータを取得するが、平成27年12月現在で常時観測ポイントは全道に214箇所設置されている。図-1は常時観測ポイントの配置状況である。また表-2は、常時観測ポイントの設置箇所数を開発建設部別に集計したものである。直轄



図-1. 北海道内常時観測ポイント配置図

表-2. 地域別常時観測ポイントの設置箇所数

開発建設部	直轄国道延長 [km]	常時観測箇所数	10km当たり箇所数
札幌	999	39	0.39
函館	693	20	0.29
小樽	488	18	0.37
旭川	760	31	0.41
室蘭	672	23	0.34
釧路	842	22	0.26
帯広	735	22	0.30
網走	885	23	0.26
留萌	268	8	0.30
稚内	284	8	0.28
全道	6,627	214	0.32

国道延長 10km 当たりの設置数に着目すると、0.26～0.41 箇所と幅を持っているが、概ね国道延長 30km に対し、1 箇所設置されている状況である。

また、ETC2.0 プローブ情報を収集可能な路側機が北海道管内の高速道路上及び一般国道上に設置されている。図-2 に、北海道内における ETC2.0 プローブ情報の取得状況に関するデータ取得状況（平成27年10月）を示したが、北海道内の高速道路・主要国道の殆どでデータ取得が可能な状況であることが確認出来る。

図-3 は北海道内の ETC2.0 と ETC の各対応車載器の経年のセットアップ台数であるが、現段階において ETC2.0 は 1.8 万台、ETC は 183 万台という状況である。ETC2.0 は、プローブデータの収集だけではなく、現行の ETC 車載器としての機能、情報配信等の機能を有している。ETC2.0 の普及率向上は今後の課題であるが、将来的に現行の ETC 全台数が置き換わると仮定すると、少なくとも現在の 100 倍程度のデータの収集・蓄積が期待される。

2) 分析対象データの蓄積状況

ここでは、平成27年10月の民間プローブデータと ETC2.0 の走行台キロ集計値と平成22年度全国道路・街路交通情勢調査（以降、センサス）の走行台キロを比較しデータの蓄積状況を把握する。結果を表-3 に示す。

センサスの走行台キロは、箇所別基本表より集計した平日を対象としたものであるため、民間プローブデータと ETC2.0 の走行台キロについても、10月の平日平均値を記載している。箇所別基本表には生活道路が含



図-2. 北海道内 ETC2.0 データ収集状況（平成27年10月）

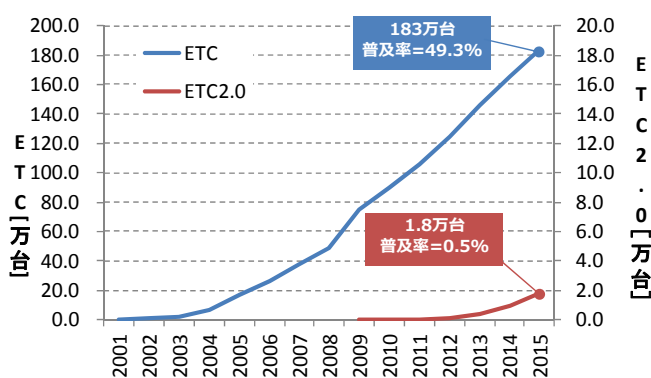


図-3. 北海道内の ETC・ETC2.0 セットアップ台数

表-3. 分析対象データの蓄積状況

開発建設部	センサス [万台キロ]	民間プローブデータ [万台キロ]	ETC2.0データ [万台キロ]
札幌	901.00	11.92	2.84
函館	187.68	1.31	0.49
小樽	181.58	1.48	0.47
旭川	308.97	2.19	0.44
室蘭	377.90	2.15	0.96
釧路	81.82	1.01	0.27
帯広	177.35	1.33	0.48
網走	95.98	1.06	0.27
留萌	3.43	0.11	0.09
稚内	21.67	0.11	0.04
全道	2,337.38	22.67	6.36

まれている点、及びセンサスは平成 22 年のデータである点には留意が必要であるが、センサスに比して台キロベースで民間プローブデータは、概ね 1/100、ETC2.0 は概ね 1/370 という蓄積状況である。したがって、民間プローブデータと ETC2.0 のどちらを分析に用いるべきか、という点については、将来的な ETC2.0 データの蓄積に期待しつつ、現段階では、民間プローブデータを用い、広域的な交通量の把握手法の構築を試みる方針が妥当と判断される。

4. 広域的な交通量の推計モデル構築

1) 把握手法の基本コンセプト

広域的な交通量の把握手法の概念図を図-4 に示す。現状では交通量を常時高精度に把握可能な地点は常時観測地点周辺道路のみだが、本手法ではプローブデータを活用することで広域的かつ面的に高精度な交通量を把握しようとするものである。推計手法構築にあたっての方針は、実務面に配慮し、入手しやすいデータから容易に交通量を予測ができることとした。

本稿では、2 つのステップから推計手法の構築・検証を行う。第 1 ステップは「予測モデルの構築」であり、常時観測機器の設置個所に対応する民間プローブデータを抽出・集計し、常観データの交通量との回帰分析を行う事で民間プローブデータの交通量を拡大し、常観データを推計するモデルを構築する。第 2 ステップは、「広域的な交通量の予測・検証」であり、回帰分析によるパラメータを用いて、全道の道路網を対象にした交通量を

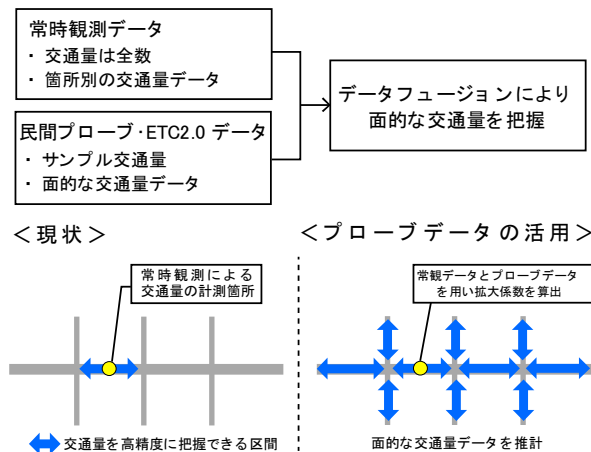


図-4. 本推計手法の概念図

推計し、センサス交通量との比較検証を行う。本提案手法の優位性としては、民間プローブデータ（将来的には ETC2.0 データ）と予めパラメータ推計を行った交通量の拡大係数を用い、広域的な交通量を容易に推計し、実務面で活用可能な点が挙げられる。

なお、現状の実務でも、常時観測交通量から面的な交通量の推計を行っているが⁶⁾、現状の手法ではある一時点での網羅的な交通量観測結果（例えば道路交通センサス）を必要としている。本手法では、プローブデータの取得件数を利用して、網羅的な交通量観測結果を必要とせず、広域的な交通量推計を可能とするため、本手法が実用化された後には、現状の手法に比べ観測コスト面での優位性が期待される。

2) 広域的な交通量推計モデルの構築方針

図-5 は、平成 27 年 10 月における、常観交通量と対応する箇所の民間プローブデータのデータ件数の散布図である。両者には相関関係があると判断され（相関係数 0.864）、両者間の拡大係数を意味する回帰直線の傾きは 114 倍程度となり、台キロベースで 1/100 程度のデータ量である事も整合する。一方で、各地点のデータは回帰直線に対してばらつきがある事も確認される。ばらつきを表す指標として平均誤差を表す RMSE（平均二乗誤差）があるが、RMSE は 4,300 台程度となっている。常時観測地点の平均交通量は 8,700 台程度であるが、RMSE を平均交通量で除した %RMSE を計算すると、約 50% と予測値が約 5 割の誤差率を持っている状況である。ばら

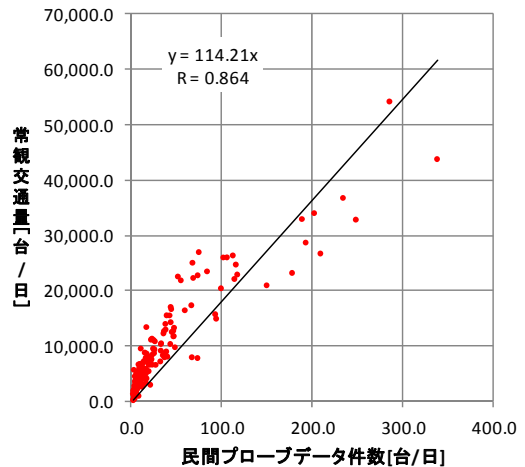


図-5. 民間プローブと常時観測交通量の散布図

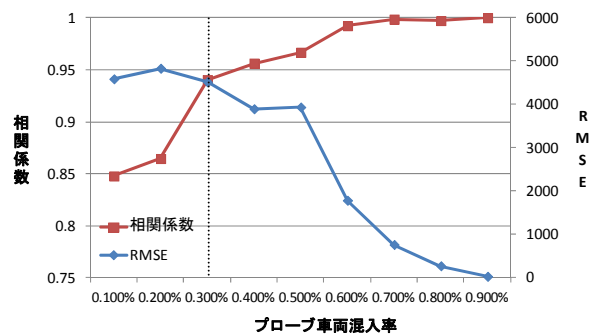


図-6. プローブ混入率と相関係数・RMSE の関係

つきの要因を考察すると、まだ十分な情報件数が無い事、常時観測地点ごとに民間プローブデータのプローブ車両の混入率が異なる事、民間プローブデータは乗用車のみのデータであるため大型車混入率の違いによりばらつきが発生する事も想定される。

図-6 は、プローブ車両の混入率別に、常観交通量と民間プローブの情報件数の相関分析を行った結果であるが、プローブ車両の混入率が高くなれば、相関係数は向上し、RMSE も低下する傾向にあることがわかる。現段階における常時観測地点におけるプローブ車両の平均混入率は 0.3%程度であるが、例えば、RMSE を 2,000 台以下（誤差率約 20%以下）で、交通量を予測するためには、概ね 0.6%（現状の 2 倍）のプローブ車両の混入がないと、単純な手法での高精度な予測は難しい事を意味している。本稿では、民間プローブデータの情報件数では説明しきれない要素を説明する変数を導入し、ばらつきを緩和する事を基本方針として以降の分析を行う。

3) 全道での交通量の推計モデルの構築

交通量の推計モデルは、本手法のコンセプトである「容易性」の観点から、民間プローブデータの情報件数とセンサスの道路状況調査で把握可能な車線数と信号交差点を用いて構築する事とした。交通量の推計に用いる式は（1）式の通りである。なお、説明変数の車線数と信号交差点密度はモデルの説明力を向上するために導入するものであり、車線数は道路の幹線性、信号交差点密度は民間プローブ車両が都市部に多いと考えられる事を踏まえ、都市部か郊外部かを表現する代理指標である。

$$Q_{24} = q_{24} \cdot \alpha + X_1 \cdot \beta + X_2 \cdot \gamma + \varepsilon \dots (1)$$

ここに、

- Q_{24} : 推計 24 時間交通量
- q_{24} : 民間プローブデータの 24 時間情報件数
- X_1 : 車線数ダミー (2 車線=0, 多車線=1)
- X_2 : 信号交差点密度 (箇所/km)
- α : パラメータ (拡大係数)
- β : ダミー変数 X_1 に対するパラメータ
- γ : ダミー変数 X_2 に対するパラメータ
- ε : 切片 (誤差項)

表-4 はパラメータ推定結果であるが、相関係数は 0.938 と向上し、各パラメータの t 値も十分に優位な結果を得た。RMSE についても約 3,000 台まで低下する結果となっている。図-7 は常時観測交通量と推計交通量の

表-4. パラメータ推定結果

R	0.938
RMSE	2988
α	122.83 (21.83)
β	5708.94 (7.40)
ε	3588.56 (13.96)

()内は t 値を表す

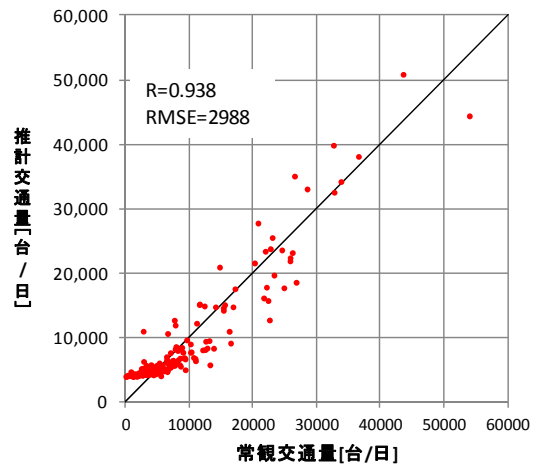


図-7. パラメータ推定結果と常観交通量の散布図

散布図であるが、概ね常観交通量 5,000 台以下では、切片パラメータ ε に大きな影響を受け、推計交通量が常観交通量と一致する 45 度線を逸脱する観測点が多いことがわかる。これは交通量が多い国道によりモデルが支配的に決定される推計モデルであることを意味している。

4) 地域別の交通量の推計モデルの構築

前述の通り、広域分散型の北海道においては、表-3 で示されるように各地域の交通量は大きく異なり、自動車の分担率やカーナビの普及率も、都市部・地方部で違いがあるものと想定される。したがって、地域別に交通量の推計モデルを構築すると推計精度の改善が期待される。地域別にパラメータを推計するためには、地域の設定を行う必要があるが、本稿では統計分析に必要となるサンプルを確保ができる北海道 6 圏域にデータを分け、地域別のパラメータを推計し、全道のパラメータを用いた交通量推計結果と精度の向上について比較検証を行うこととする。

ここでの分析モデルは、（1）式の通りとし、データを 6 圏域に分けてパラメータを推計する事とする。表-5 は地域別にパラメータ推計を行った結果であるが、各地域ともに相関係数が 0.9 以上と非常に高く、パラメータ値、t 値ともに概ね良好な結果を得た。拡大係数である α に着目すると、道央が最も小さく、十勝がもっとも大きく 3 倍程度となっていることから地域別に拡大係数が異なるという仮説は正しかったと判断される。各地域のモデルを用い、全道で常時観測データと推計交通量を比較すると、相関係数は、0.961 と全道一括推計の 0.938 に対して向上し、RMSE も約 3,000 台から約 2,400 台に減少しており、モデルの精度向上が読み取れる。

図-8 に推計交通量の結果について示した。民間プローブデータが取得出来ている区間において、交通量が広域的・面的に推計出来ていることが分かる。図-9 に地域別パラメータを用いた推計モデルの散布図を示した。全道一括推計の散布図である図-7 に比較して 5,000 台以下の交通量の精度向上が確認できる。

表-5. 地域別パラメータ推定結果

	地域別推計							全道一括推計
	道南	道央	十勝	釧根	オホーツク	道北	全道	
R	0.979	0.952	0.929	0.951	0.939	0.953	0.961	0.938
RMSE	1316	3467	1234	1724	1100	1427	2401	2988
平均交通量	10051	12959	5088	6750	5034	5391	8741	8741
%RMSE	13%	27%	24%	26%	22%	26%	27%	34%
α	277.89 *	107.14 *	337. *	318.73 *	186.63 *	208.36 *	-	122.83 *
β	-	7540.31 *	-	661.1	2853.91	-	-	5708.94 *
γ	1760.76 *	-	-	-	1605.38	310.91	-	-
ε	1591.26 *	3948.47 *	1126.85 *	1482.87 *	1905.26 *	1738.52 *	-	3588.56 *

*は5%有意なパラメータを示す

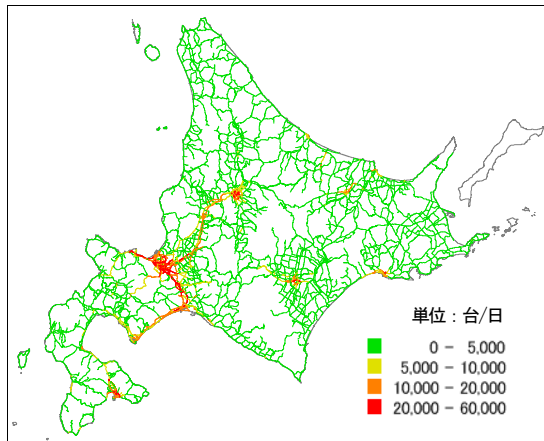


図-8. 北海道内の推計交通量の結果

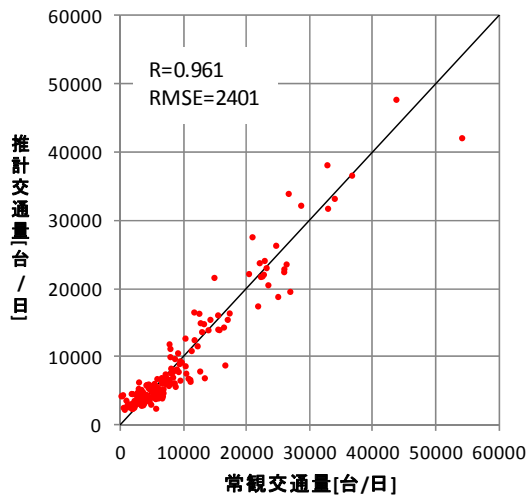


図-9. 常時観測交通量と推計交通量の散布図 (地域別パラメータ)

表-6. 推計交通量とセンサス交通量の相関分析結果

道路種別	R	RMSE
高速道路	0.962	3,723
国道	0.873	4,030
主要地方道(道)	0.870	3,675
主要地方道(市)	0.463	7,480
一般道道	0.660	3,800
指定市一般市道	0.817	6,313

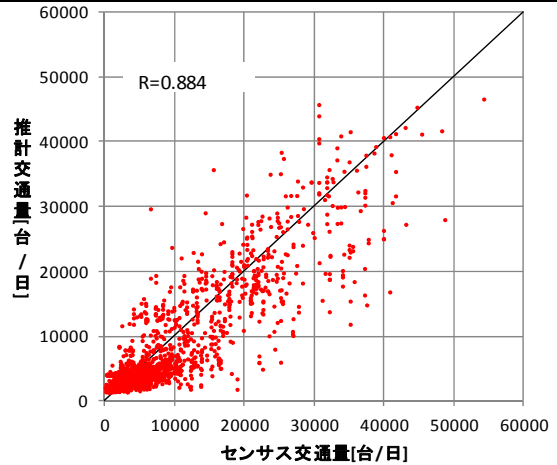


図-10. 推計交通量とセンサス交通量の比較 (高速道路・国道・主要地方道)

2) 推計モデルの検証と考察

表-6 は、道路種別ごとのセンサス交通量と予測モデルにより推計した推計交通量の相関分析結果である。高速道路、国道、主要道、指定市一般市道については相関係数が 0.8 以上と比較的良好な結果を示しているのに対し、一般道道・主要市道においては、0.8 を下回る結果となっている。図-10 に高速道路・国道・主要地方道について、センサス交通量と推計交通量の比較を示したが、全体としてややばらつきが見られる状況である。この要因としては、道路上の交通量に占めるプローブ車両の割合（混入率）が十分ではないことや道路種別により均一ではない事が指摘される。

すなわち、幹線性が高く交通量の多い道路についてはプローブ車両の混入率は高く、推計誤差も小さく、生活道路や補助幹線道路などにおいては混入率が低いため推計誤差も大きくなる事が原因と考えられる。今回の分析に用いた常時観測機器の設置箇所は幹線性の高い国道であり、国道で算出した拡大係数は、幹線性の低い道路と比較して小さい値であると考えられる。従って、幹線性

5. 広域的な交通量の推計モデルの検証

1) 推計モデルの検証方法

前章では、広域的な推計モデルとして、全道モデルと地域モデルを構築し、地域モデルの方が交通量の推計結果が良好である事を示した。ここでは、地域モデルを用いて、センサス交通量との比較分析を行うことで、広域的な交通量の推計精度を検証する。

また、予測モデルは直轄国道の常時観測交通量と民間プローブの情報件数の関係から構築されているため、国道以外の道路も含めた道路種別毎での、予測精度に関する考察も行う。

の低い道路においては拡大係数が低めに見積もられる結果となり、推計される交通量も過小になると想定される。

また 4 章の図-6 で示したとおり、プローブ車両混入率が現状の 2 倍程度になると飛躍的に誤差が減少することが示されており、今後の ETC2.0 データ等のデータ量の増大による推計精度の向上が期待される。

6. 広域的な交通データの活用シーンの検討

最後に、本手法にて推計可能な広域的な交通データの実務面における活用方策について検討を行う。

表-7 にデータの活用シーンを示す。なお、活用シーンの検討については、将来的に ETC2.0 データを活用することを想定し、本推計手法及びデータに関する下記特徴を踏まえたものとした。

- ①データの広域性：生活道路を含む広域的な交通量予測が可能
- ②予測の連続性：日々の面的な交通量を把握可能
- ③リアルタイム性：将来的にリアルタイムに予測が可能となり得る

具体例を挙げると、都市内での渋滞対策の一例として、各種施策（例：トランジットモール等）を実施した際の交通シミュレーションを行う際に、面的で詳細な交通量を活用することにより、格段の精度向上が期待出来る。また北海道内では、夏期と冬期の交通量変動が大きい傾向にあるが、予測の連続性を生かして、季節毎の面的な交通量を把握し、各時期での施策立案に活用することも、本手法の有用な活用先として考えられる。

さらにデータの広域性を生かし、今後、本手法の精度が十分に向上した際には、各種の大規模交通量調査の代替手段として活用することも考えられる。

2 章にて述べた通り、交通量は地域特性・地域課題・経済状況等を反映する基本的なデータセットであり、本

表-7. 想定される活用シーン

予測手法の特徴	施策イメージ	
データの広域性	大規模な交通量調査への活用	・道路事業の供用前後の影響・効果の把握 ・道路交通センサス、パーソントリップ調査等の大規模調査への活用
	渋滞対策への活用	・渋滞分析、渋滞対策への活用 ・混雑状況の情報配信 ・生活道路への交通流入等の分析
リアルタイム性	イベント交通への活用	・大規模イベント時の経路誘導等の交通マネジメントへの活用 ・渋滞状況のリアルタイム配信によるモビリティマネジメントへの活用
	救急搬送の効率化	・救急車両への混雑状況等の配信
	交通事故等の突発事象の活用	・交通量の急激な変化等を検出することによる異常事象のモニタリング
	防災・減災への活用	・線状降水帯のような広範囲かつ多量の降雨や降雪、吹雪時の交通モニタリング
予測の連続性	地域活力のモニタリング	・地域経済の広域的継続的なモニタリング
	維持管理コスト縮減	・道路の老朽化に伴う道路の維持管理コストの縮減のための基礎データ収集

手法にて提案した交通量が、既存の交通量に比べ、時間的・空間的の両面で豊富な情報量を持つことから、今後、地域の経済活力のモニタリングの一指標として、本手法にて提案した広域的・面的な交通量を活用することも、有力な活用方法として考えられる。なおその際には、多様なビッグデータが出現している昨今の情勢を踏まえ、異種データ（例えば、帝国データバンクの企業間取引データ⁷⁾等）とのデータフュージョンを行い、分析内容を拡張していくことが、望ましいと考えられる。

7. まとめ

本稿においては、自動車交通に関するビッグデータの特性や収集状況を整理するとともに、常時観測データとプローブデータの相互補完性に着目し、データフュージョンによる広域的な交通量の推計手法を提案した。交通量の推計手法については、常観データと民間プローブデータを用いて、日単位の広域的な交通量を推計するモデルの構築を行い、その推計精度の検証を行った。検証の結果、主要道道以上の交通量の推計は一定レベルで可能であり、今後は ETC2.0 データ等の蓄積により、プローブデータのサンプル数が増えることが期待され、細街路を含む交通量推計が可能になると考えられる。

少子高齢化時代を迎えた我が国においては、限られた人材、予算のなかで、よりきめ細やかな調査、計画、実効性のある施策立案等への迅速な対応が社会的な要請となっている。本手法はこれらの調査・計画において活用可能である事、データの収集やパラメータ推定を比較的容易に行う事が本稿において確認されたことから、今後、予測精度の更なる向上を目指すとともに、他地方整備局におけるの展開を目指すなど、幅広く実務面で活用されるよう、活用シーンの検討もあわせて深化させたい。

参考文献

- 1) 世界最先端 IT 国家創造宣言について、首相官邸 HP, http://www.kantei.go.jp/jp/singi/it2/pdf/it_kokkasouzousengen.pdf, (URL は平成 28 年 1 月時点)
- 2) 橋本, 山下, 高宮：常時観測交通量データを用いた交通量の変動特性分析, 第 30 回日本道路会議, CD-ROM, 2013.
- 3) 水木, 橋本ら：民間プローブデータを用いたボトルネック交差点とその影響範囲の特定方法に関する分析土木計画学研究・講演集 Vol. 47, CD-ROM.
- 4) 横井, 堺, 田代：交通ビッグデータからみた道路施策の面的評価, 第 31 回日本道路会議, CD-ROM, 2015.
- 5) 上原, 森山, 山崎：北海道におけるプローブ情報の活用に関する検討, 第 31 回日本道路会議, CD-ROM, 2015.
- 6) 橋本, 山崎, 上坂：交通量常時観測データを活用した交通量データ収集の効率化, 土木学会論文集 F3(土木情報学), Vol. 68, No. 2, pp. I_64-I_72, 2013.
- 7) 企業概要データベース COSMOS2, 帝国データバンク, <http://www.tdb.co.jp/lineup/cosmos2/>, (URL は平成 28 年 1 月時点)