

パーソントリップ調査のデータを用いた時刻別滞留人口推計に向けた基礎的研究

小磯 和紀¹・岡本 直久²・石田 東生³

¹学生会員 筑波大学大学院 システム情報工学研究科社会工学専攻 (〒305-0006)
茨城県つくば市天王台1-1-1
E-mail:s1420475@sk.tsukuba.ac.jp

²正会員 筑波大学 システム情報系 社会工学域准教授 (〒305-0006) 茨城県つくば市天王台1-1-1

³正会員 筑波大学 システム情報系 社会工学域教授 (〒305-0006) 茨城県つくば市天王台1-1-1

パーソントリップ調査の抱える課題を、ビッグデータを組み合わせる事によって克服出来ないかが議論され始めている。交通関連ビッグデータの一例としてGPS記録携帯電話等から連続的に収集した位置情報から、時刻毎にメッシュ毎の滞留人口を推計したデータである滞留人口データを、PT調査の補完データとして活用する上で、双方のデータの形を揃えるためには、時刻別の移動・滞留人口の算出精度を把握しておく必要がある。

本研究ではPT調査から算出される時刻別移動・滞留人口とメッシュ単位の滞留人口情報との比較する上で生じる、PT調査の時間記録の実態に関する課題を把握する事を目的とする。

Key Words : *person trip survey, big data, data fusion*

1. 研究の背景と目的

(1) 研究の背景

都市圏パーソントリップ調査(以下PT調査)は人々の移動(トリップ)を総合的に把握することを目的として全国の都市圏等において実施され、各都市交通計画の策定に活用されてきた。

しかし、PT調査は以下のような問題点が指摘できる。

- ・高コスト構造となっている
 - ・調査頻度が低い、得られるデータは秋の平日1日のみ
 - ・アンケート調査であるが故の誤差や記入漏れ等の存在
- 一方、近年交通関係のビッグデータが注目を集めている。携帯電話などのGPS情報、ETCやICカード乗車券の利用記録、カーナビゲーションの走行記録、交通量トラフィックカウンターのデータ等、多種多様なビッグデータが存在する。これらのデータは、高精度なデータが長期的かつ継続的に記録されているという特徴を持つ。GPS技術の進展により、携帯電話等の位置情報を集計しメッシュ毎の人口を推計している事例もある。

現在、PT調査の課題や弱点を克服する為の手段として、高精度かつ継続的に記録されたビッグデータを組み合わせ活用することが出来ないか議論されている。¹⁾

なお、本研究で対象とする移動・滞留人口を以下の

ように定義する。

移動人口：ある時刻において移動中であるサンプル(人)をゾーン単位で集計したもの

滞留人口：ある時刻において非移動中であるサンプル(人)をゾーン単位で集計したもの

移動・滞留人口：上記の移動人口と滞留人口の和

(2) 本研究の目的

PT調査とビッグデータを組み合わせて課題を克服するため、GPS記録によるメッシュ毎の時刻別滞留人口情報が利用可能であるかを検討する。そのためには、それぞれのデータの精度、特性を把握しておく必要がある。本研究ではPT調査から時刻別移動人口・滞留人口データを作成する上で議論すべき課題であるPT調査の時刻記録の実態を把握する事を目的とする。

(3) 既往研究と本研究の位置づけ

PT調査のマスターデータから時刻別移動・滞留人口を算出する試みは、既に中野ら(1994)²⁾(2008)³⁾により行われている。中野らは高知都市圏を対象にPT調査から時刻別移動・滞留人口を推計し、災害発生時の被害想定や帰宅困難者の推計を行っている。この研究においては、滞留人口はある時刻において移動中でない人口をゾーン

別に集計したもの、移動人口は、移動中のトリップについて、ある時刻までの発地からの所要時間を考慮し、交通手段別にネットワーク上のある地点に割り当てたものをゾーン別に集計したものと定義されている。

また、戸田ら(1990)⁵⁾、谷口ら(1992)⁶⁾はPT調査のデータを用いて非移動中の時間(総滞留時間)に着目し、京阪神都市圏において小ゾーン単位で人々の滞留行動を分析している。

2. 比較検討課題の整理

(1) PT調査の概要と特徴の整理

a) PT調査の概要

PT調査は1967年に広島都市圏で実施されて以来、交通の主体である人の動きを総合的に把握できる唯一の調査として、日本全国の様々な都市圏で調査が行われてきた。

都市交通計画などを検討する上では、地域の交通の課題を把握した上で目指すべき都市像に対応した交通政策・交通計画を立案・策定していくことが重要である。そのためには地域の交通の実態を総合的かつ定量的に把握する必要がある。PT調査の調査項目からは鉄道や自動車などの各交通手段の利用割合や、交通量なども定量的に把握する事が可能であり、交通計画において幅広く活用できる調査であると言える。

しかし、PT調査は以下のような課題を抱えている。

1点目は、PT調査が大標本調査を前提とし、高コスト構造であるという点である。調査協力率が低下する一方、コスト削減の動きが高まっており、従来の予算の確保や必要なサンプル数を集める事が難しくなっている。

2点目が、調査頻度が低いという点である。PTの調査頻度は東京都市圏でも10年に1度である。調査が1度行われたきりの都市圏もある。しかし、調査回数・頻度を増やすという事はコスト面から容易ではなく現実的ではない。

3点目は得られるデータが平日1日分のみであるという点である。PT調査で得られるデータは10年に1度の秋の平日1日分のデータのみである。調査ニーズも多様化する中、平日だけでなく休日・災害発生時などへのデータも求められている。

b) 東京都市圏PT調査の歴史

東京都市圏PT調査は昭和43年より10年毎に実施されている。平成20年に第5回目の調査を迎えた。

国土交通省関東地方整備局と1都4県5政令市と都市再生機構、東日本高速道路、首都高速道路、中日本高速道路で構成される東京都市圏交通計画協議会が昭和43年の発足し、東京都市圏PT調査が行われてきた。

c) 東京都市圏PT調査の概要

本研究では平成20年第5回東京都市圏PT調査のマスターデータを用いた。平成20年の調査の概要を以下に示す。

- ・調査時期：平成20年10月～12月の火曜日から木曜日までの平日の1日間(祝日・祝前後の日を除く)
- ・対象都市：東京を中心とする半径約80km圏域で、東京都、神奈川県、埼玉県、千葉県、茨城県(南部)の238市区町。
- ・対象者：対象者は東京都市圏に居住する約1600万世帯の内、無作為抽出した約140万世帯の約3,460万人(平成20年当時)の満5歳以上の人々。
- ・調査方法：対象世帯に郵送で調査票を配布・回収
- ・回収率：票数は約34万世帯票、東京都市圏の人口の約2%に相当する約73万人から調査票を回収。(有効回収率約24%)
- ・調査項目：世帯・個人属性及び移動状況を調査
 - 世帯・個人属性：居住地、性別、年齢、職業、勤務先・通学先、自動車・二輪車の保有台数運転免許の有無
 - 移動状況：出発地、到着地、移動目的、移動手段、所要時間(出発時刻・到着時刻)、車の運転有無、駐車・駐輪場所、有料道路利用有無

(2) PT調査とビッグデータを組み合わせる上での課題

PT調査とビッグデータを組み合わせる上での課題は大きく分けて2つ存在すると考える。

1つはゾーンサイズ・形の違いである。第5回東京都市圏PT調査は大きい単位から順に、大ゾーン、中ゾーン、市町村ゾーン、計画基本ゾーン、小ゾーンという独自の単位を用いて、都市圏の各エリアを区分している。各ゾーンの詳細は次の通りである。

- ①大ゾーン：地理的、歴史的な地域のまとまりを考慮しつつ、東京都市圏全域のマクロ的な分析、検討の単位となるゾーン
- ②中ゾーン：ほぼ市区町村を単位とするが、大都市では数個に分割し、周辺では市町村がいくつかまとまっている場合もあるゾーン
- ③市区町村ゾーン：市区町村の単位となるゾーン。
- ④計画基本ゾーン：小ゾーンを数個集めて構成し、広域における計画単位として、また、地域としてのまとまりのある交通計画の単位となるゾーン。
- ⑤小ゾーン：夜間人口約15,000人を目安とし、地区計画の単位となるゾーン

PT調査で扱うことの出来る最小地区単位は夜間人口15,000人を基準とした小ゾーンである。そのため、ゾーンの大きさは過疎地域では大きく、ベッドタウンや都

市部においては小さくなる傾向にある。ゾーンの形は市区町村や町丁目などの境界を基準に区切られており不定形である。

もし、PT調査から小ゾーン単位で時刻別移動・滞留人口を推計しても滞留人口の情報は単純には比較したり組み合わせたり出来ないという課題がある。

もう1つはPT調査の時刻の記録実態に関する課題である。ビッグデータでは、実際の時刻が記録されるものが多いが、PT調査は記述に依存しているため、どうしても正確な時刻が記入されていないという課題がある。

本研究ではこの2つ目の課題、PT調査の時刻の記録実態に着目して分析を行った。

3. PT調査の時刻の記録実態に関する課題

この章では、PT調査の時刻の記録実態に関する課題を指摘し、混雑統計のようなメッシュデータと比較する上の課題について実際のデータを用いて検証する。

(1) 時刻別移動・滞留人口の推移

千代田区の小ゾーン番号00100地区を対象に1日の滞留人口（移動人口は除く）の推移を示した。（図-1）このゾーンは東京の都心部にあるため夜間人口に比べて昼間人口が極めて多い事が分かる。7:00から11:00頃にかけて滞留人口は一気に上昇し、午後になると14:00をピークに徐々に減少し始める。そして、電車が運行していない時間帯になると殆ど滞留人口は変化しなくなる。

続いて時刻毎（毎時00分）に、移動人口と滞留人口の割合を算出した。（図-2）

どの時刻においても滞留人口が全体に占める割合が大きい事が分かる。一方で、8:00や18:00といった通勤・帰宅時刻に移動人口の割合が高くなる事も分かる。12:00にも移動人口の割合が高くなっているが、これは昼食行動が影響していると思われる。

時刻が記入していないといった理由により、その時刻において移動していたのか滞留していたのが判別できなかった「不明」の割合が各時刻において10%近く存在

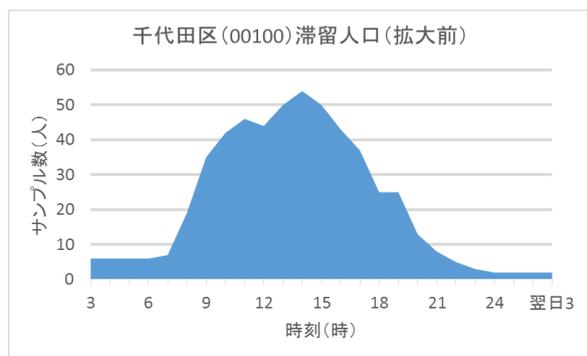


図-1 時刻別滞留人口

する事が分かる。PT調査から時刻別移動・滞留人口を推定する際には、所在地を推定できないサンプルが常に一定数存在する事に留意する必要がある

(3) 時刻データの記録実態

a) 時刻の記録実態

トリップの発着時刻情報は『午前午後』『時』『分』という3つの項目で記録されている。それぞれ時刻が「有」る（記録されている）サンプルと「不明」として記録されているサンプルがある。想定される「有」及び「不明」の組み合わせは9パターンであるが、実際には4パターンのみ存在した。（表-1）

午前午後・時・分が全て記録されているサンプルは全体の93.15%であった。次いで全ての時刻情報が「不明」と記録されたデータが3.91%であった。その次に午前・午後のみが『有』る（記録された）サンプルが全体の2.9%を占めた。

b) 発着時刻の記録実態

各サンプル（トリップ）の発時刻及び着時刻の実態を調べた。発時刻及び着時刻について、『時』や『分』に不明という記録が残されているものについて「不明」と

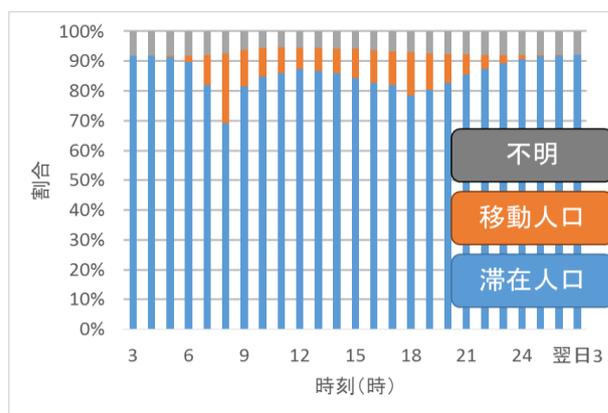


図-2 時刻別移動・滞留人口の推移

表-1 存在した時刻情報

午前午後	時	分	発時刻	着時刻	計(A)	(A)の割合
○	○	○	1,678,962	1,676,628	3,355,590	93.15%
○	○	×	694	657	1,351	0.04%
○	×	○	0	0	0	0.00%
×	○	○	0	0	0	0.00%
○	×	×	51,281	53,289	104,570	2.90%
×	○	×	0	0	0	0.00%
×	×	○	0	0	0	0.00%
×	×	×	70,321	70,684	141,005	3.91%

凡例 ○：有 ×：不明

判定した。結果を(表-2)に示す。

総交通生成量のうち7.6%は時刻情報が不明である事が分かった。その内の殆どは発着時刻が共に不明であるサンプルであった。

c) 人に着目した1日のトリップの発着時刻の記録実態

人々の1日の1連のトリップに着目し、1日のトリップ数2~4のサンプルについて、発着時刻の「有」・「不明」の組み合わせのパターンを表に示し、その傾向を把握した。特に割合の高い上位10パターンについてに掲載する。(表-3)(表-4)(表-5)

トリップ数2, 3, 4の場合のいずれの場合についても全ての発着時刻が有るサンプルが最も多い。次に多いパ

表-2 存在した発着時刻情報

	サンプル数	構成比
交通生成量	1,674,612	100.0%
発着時刻有	1,547,966	92.4%
発着時刻不明	118,929	7.1%
発時刻のみ不明	2673	0.2%
着時刻のみ不明	5044	0.3%

表-3 1日のトリップ数が2の人の時刻記録実態

トリップ番号				人数	割合(%)
1		2			
発時刻	着時刻	発時刻	着時刻		
○	○	○	○	304,299	84.0
○	○	×	×	34,896	9.6
○	○	△	△	14,171	3.9
×	×	×	×	4,356	1.2
○	△	△	○	692	0.2
○	○	○	△	418	0.1
△	△	△	△	395	0.1
○	△	×	×	279	0.1
△	△	○	○	228	0.1
×	×	△	△	209	0.1

表-4 1日のトリップ数が3の人の時刻記録実態

トリップ番号						人数	割合(%)
1		2		3			
発	着	発	着	発	着		
○	○	○	○	○	○	78,340	78.2
○	○	○	○	△	△	11,678	10.6
○	○	○	○	×	×	5,712	5.7
○	○	×	×	×	×	987	0.9
×	×	×	×	×	×	689	0.7
○	○	△	△	△	△	378	0.3
○	○	×	×	○	○	289	0.3
○	○	△	△	○	○	184	0.2
○	○	○	○	△	○	162	0.2
○	○	○	○	○	△	149	0.1

ーンが1日の最後のトリップの発着時刻のみ全て不明、又は午前午後情報以外は不明でそれ以外の1日のトリップの発着時刻は記録されているサンプルであった。

1日の全てのトリップの時刻が記録されていたサンプルはトリップ数2の場合が84.0%、トリップ数3の場合が78.2%、トリップ数4の場合が85.2%である。午前午後のみ情報が有っても、時・分の記録がなければ時刻情報としては殆ど意味が無い。そこで、表中において△と×で示した結果を以下、同一のもののみとする。1日の最後のトリップのみ発着時刻不明(時・分の情報が不明であるもの)のサンプルが全体に占める割合はトリップ数2の場合が13.5%、トリップ数3の場合が16.3%、トリップ数4の場合が8.9%であった。1日の最後のトリップの時刻が不明であるサンプルは1日の全てのトリップの時刻が不明であったサンプルよりもはるかに多い事が伺える。

PT調査は1日の最後のトリップ、すなわち、帰宅するトリップの時間を記入し忘れる傾向があるのではないかと考えられる。1日のトリップ数2~4以外のトリップにも着目し、1日の最初と最後のトリップの出発時刻及び到着時刻の記録実態を表-6に示した。1日の最初のトリップの出発時刻に比べて1日の最後のトリップの着時刻は10倍以上の割合で不明となっている事が分かる。

表-5 1日のトリップ数が4の人の時刻記録実態

トリップ番号								人数	割合(%)
1		2		3		4			
発	着	発	着	発	着	発	着		
○	○	○	○	○	○	○	○	82,877	85.2
○	○	○	○	○	○	△	△	6,359	6.5
○	○	○	○	○	○	×	×	2,324	2.4
○	○	×	×	○	○	△	△	630	0.5
○	○	○	○	△	△	△	△	420	0.4
○	○	×	×	○	○	○	○	411	0.4
○	○	×	×	×	×	×	×	392	0.4
×	×	×	×	×	×	×	×	391	0.4
○	○	△	△	○	○	△	△	377	0.4
○	○	△	△	○	○	○	○	333	0.3

凡例

- ：時刻記入あり
- △：時刻記入なし、午前・午後の記入あり
- ×

表-6 1日の最初と最後のトリップの時刻記録実態

	該当者数(人)
トリップ数1以上のサンプル	629,098(100.0%)
1日の最初のトリップの発時刻が記録されたサンプル	619,645(98.5%)
1日の最初のトリップの発時刻が不明のサンプル	9,453(1.5%)
1日の最後のトリップの着時刻が記録されたサンプル	531,100(84.4%)
1日の最後のトリップの着時刻が不明のサンプル	97,998(15.6%)

PT調査から滞留人口を推計する上で、サンプル数を最大限に活用する必要がある場合、1日の最終トリップの発着時刻を推定することが出来れば、多くのサンプルが活用できるようになるのではないかと考えた。

(4) 時刻データ推定

サンプルを最大限に活用するために、『不明』と記録された不完全なデータの時刻を推定する手法を検討する。今回は特に不明データが多く見られた事から1日の最終トリップの発時刻及び着時刻を推定する。1日の最終トリップについて、(発時刻・着時刻)の組み合わせの中で不明が1つでも含まれる3つのパターン(発時刻・着時刻)が(不明・正常)(正常・不明)(不明・不明)について推定する。

そのために、以下の2つの推定手法を用いる。使用するデータは東京都市圏PT調査のデータの他、ゾーン間の距離を測るために小ゾーンの中心座標データを用いる。

【手法A】 1日の最後のトリップのうち帰宅目的の移動について自宅着時刻の分布から、他の時刻不明サンプルについて帰宅時刻を推定する

【手法B】 発着時刻の一方が不明な場合に移動したゾーン間の直線距離から代表交通手段毎に所要時間を推定し不明な時刻を推定する。

パターン(不明・正常)(正常・不明)については手法(B)を、パターン(不明・不明)には手法(A)と(B)の両方を用いて、最終トリップの時刻を推定する。

a)1日の最後のトリップの自宅到着時刻の推定(手法A)

手法(A)では、はじめに、1日の最後のトリップで帰宅目的の移動について、帰宅時刻が正常なサンプルを対象に自宅到着時刻の分布を1時間単位で集計し算出した。結果は以下の通りである。

図-3に示す自宅到着時刻分布をもとに、パターン(発時刻・着時刻)の組み合わせが(不明・不明)で、且つ、帰宅目的のトリップあった1日の最後のトリップについて、帰宅時刻を上記の分布に基づいて、乱数を発生し帰

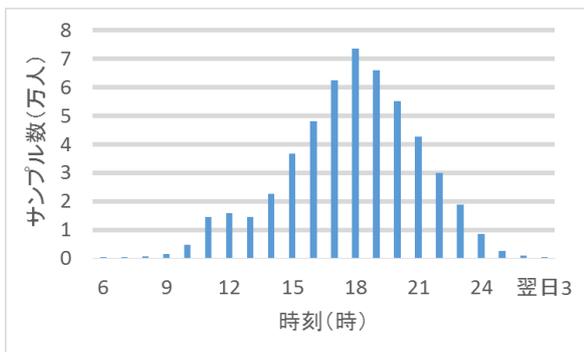


図-3 1日の最後の帰宅目的の移動の自宅到着時刻 (拡大前)

宅時刻を推定した。仮定した時刻が前のトリップの時刻と比較して矛盾が生じる場合は再度試行を行った。前のトリップが仮定した帰宅時刻を基に1日の最後のトリップの出発時刻を推定する為に手法(B)を用いた。

b)移動時刻の推定(手法B)

手法(B)では移動時刻の推定を行う。手法(A)で着時刻を仮定したサンプルの発時刻の推定や、1日の最後のトリップで発着時刻の一方が欠けているデータについて他方の時刻の推定を行う。手法は、時刻が正常に記入されているサンプルから代表交通手段毎に距離別の移動時間を算出し、時刻不明データを移動距離と代表交通手段情報と、仮定した着時刻や正常に記録されている一方の発着時刻情報から他方の時刻を推定する。

PT調査のデータの内、全ての発着時刻が正常なサンプルを対象に距離別・代表交通手段別で所要時間をプロットした。尚、小ゾーン間の距離は、ゾーン間中心座標を示した緯度・経度の情報から、全ての2ゾーン間の直線距離を算出した。この際に、同じゾーン内での移動である内々トリップについては、ゾーン間距離が0となるため、集計の対象外としている。

代表交通手段毎に移動距離と所要時間の関係から回帰直線を求めた。鉄道・地下鉄の場合、 $y(\text{所要時間(分)}) = 1.6142x(\text{移動距離(km)}) + 33.448$ という結果になったが決定係数R2乗値が0.4726と低かった。この結果は他の代表交通手段についても同様であり、決定係数の値が不十分であった。

そこで、代表交通手段毎に距離1km毎の移動にかかった所要時間の平均値を算出した。そして、代表交通手段毎に近似式を算出した。以下に鉄道・地下鉄の場合の結果と、近似式を示した。近似曲線は、対数近似等も試したが、最も決定係数の値が高くなったのは直線での近似であった。

結果(図-4)は $y = 1.5964x + 32.643$ 、R二乗値は 0.9759 と

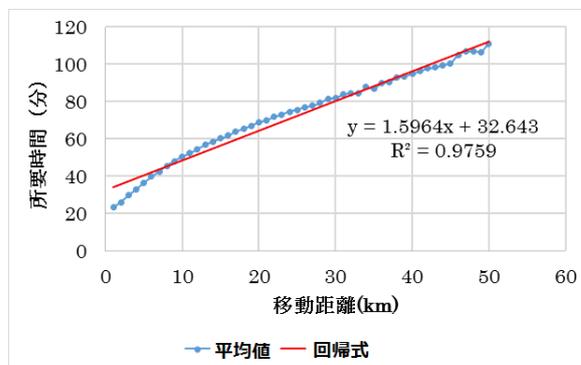


図-4 移動距離と所要時間(平均値)の関係 (代表交通手段:鉄道)

なった。各代表交通手段毎に作成した回帰式を元に時刻不明データの推定を行った。

1日の最後の発着時刻の少なくとも一方が不明であった99,978サンプル（全1,906,033サンプルの5.3%）のうち、48,469サンプル（全サンプルの2.5%、99,978サンプルの内の48.5%）について、不明時刻の推定が出来た。（図-5）

5. まとめ

本研究ではPT調査のデータ及び、混雑統計を例に携帯電話位置情報を活用した交通ビッグデータの特徴や課題を整理し、それぞれのデータを組み合わせて利用出来ないか比較検討をする上での課題を整理した。

また、PT調査のマスターデータの記録実態を様々な視点から把握し、時刻不明データの扱いに関して、対処の方向性の一例を示した。

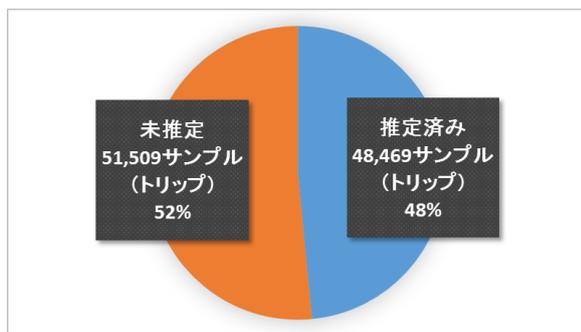


図-5 時刻不明データの推定

今後の課題として、PT調査の不明データの推定方法の改善の必要性が挙げられる。また、それぞれのデータの精度やゾーンの大きさや形の違いといった課題にも対応する必要もあるだろう。

最終的には、ビッグデータとPT調査を組み合わせる事で具体的に何が出来るのかを検討する必要がある。

参考文献

- 1) 国土交通省：第1回「情報技術を活用した公共交通活性化に関する調査」検討委員会議事概要，2013
- 2) 国土交通省：「道路分科会建議 中間とりまとめ」平成24年度フォローアップとりまとめ（案），2012
- 3) 中野敦：都市圏交通計画・調査の課題と改善方策に関する研究，2008.
- 4) 中野敦，本田武志：都市における時刻別地域別人口分布と震災被害に関する研究－PT調査データの分析に基づいて－，土木計画学研究委員会，阪神・淡路大震災調査研究論文，pp.75-82，1997.
- 5) 戸田常一，谷口守，秋本直人：都心地区における来街者の滞留行動に関する研究，第25回日本都市計画学会学術論文集，pp.79-84，1990.
- 6) 谷口守，秋本直人，天野光三：滞留時間分析システムを用いた滞留促進のための基盤整備に関する研究，土木計画学研究・論文集 No.10，pp.119-126，1992.

(? 受付)