

東京大学大学院新領域創成科学研究科
社会文化環境学専攻

2016 年度
修 士 論 文

携帯電話 GPS と鉄道ネットワークデータを用いた
東京都市圏における鉄道利用状況の推計
Estimation of Railway Passenger Volume in Tokyo Metropolis
Using Mobile Phone GPS Data and Railway Network Data

2017 年 1 月 23 日提出
指導教員 柴崎 亮介 教授

池 澤 俊
Ikezawa, Shun

目次

1	序論	3
1.1	研究背景	3
1.2	研究の目的	6
2	既往研究	7
3	鉄道利用路線の推定手法	10
3.1	手法の概要	10
3.2	本研究に用いたマップマッチング手法	10
4	鉄道ネットワークデータ	13
4.1	使用したデータセット	13
4.2	鉄道ネットワークデータの作成	17
4.3	マップマッチング及び経路探索の検証	25
5	鉄道利用路線の推定	35
5.1	使用したデータセット	35
5.2	推定結果	35
5.2.1	鉄道利用者数の時間的推移	35
5.2.2	駅別乗降客数	37
5.2.3	駅間移動人員数	44
6	結論	58
6.1	本研究の成果	58
6.2	本研究の課題と展望	59
	参考文献	61
	謝辞	63

1 序論

1.1 研究背景

鉄道利用者数を把握することは都市計画における交通状況把握，短期的なイベントによる鉄道需要の推定，災害や事故発生時の影響評価や被害推計といった社会的なニーズがある．民間業者が公共交通を運用している日本では Suica や Pasma といった各種交通系電子マネーの普及により，鉄道利用者の乗降駅の情報を得ることが可能になった．

しかし，異なる事業者間を跨る乗換といった途中の利用路線までを把握することは困難である．各鉄道会社では，車両の重量を量ることで車両ごとの利用人数はおおよそ把握することが可能[1]であるが，それは自身が運営している路線に限られる．また，鉄道事業者は各ユーザーの定期乗車券の区間を見ることで，多くのユーザーの利用路線をおおよそ把握することは可能であるが，それは通勤・通学の時間帯の移動経路に限られてしまう．このように鉄道事業者が各ユーザーの利用路線を把握することは難しい．

途中の経路や利用路線を知る手立てとして，パーソントリップ調査[2]といったアンケート調査がある．これは東京都市圏（東京都・神奈川県・埼玉県・千葉県・茨城県南部）に居住する約 1,600 万世帯のうち、無作為で選ばれた約 140 万世帯の方（満 5 歳以上）を対象に行うアンケート調査である．しかし，アンケート調査ということで実施者と被験者の負担が大きいため，継続的に実施することは難しく，利用路線まで細かく回答することは回答者の負担がさらに大きくなってしまう．また，国土交通省による大都市交通センサス[3]には鉄道利用者数の推計を行っている．（図 1.1.1, 1.1.2 参照）この大都市交通センサスでは、調査票配布期間中，調査票配布対象駅において、降車する旅客に対して調査票を配布することで調査を行っている．調査は首都圏、中京圏、近畿圏の三大都市圏で行われており、各都市圏は以下の条件により設定している．

- ・首都圏は東京駅、中京圏は名古屋駅、近畿圏は大阪駅までの鉄道所要時間が 2 時間以内（中京圏は、1 時間 30 分以内）を満たす市区町村
- ・首都圏は東京都 23 区、中京圏は名古屋市、近畿圏は大阪市への通勤・通学者数比率が 3%以上かつ 500 人以上を満たす市区町村

国土交通省 大都市交通センサスの調査概要[4]より引用

地下鉄における推計も行われている．しかし，この推計もアンケート調査であるため，継続的な実施が難しく，現在は 5 年ごとに実施されている．また，この推計は発着地の情報が取得できる事業者及び自動改札機が導入されている事業者・駅に限られている．

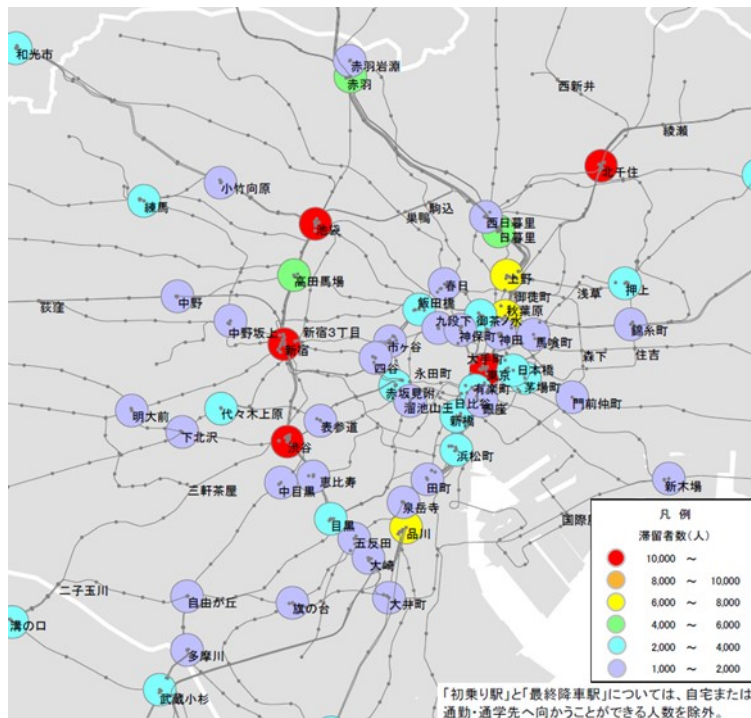


図 1.1.1 大都市交通センサス(2010 年度)による駅別乗降客数の推計

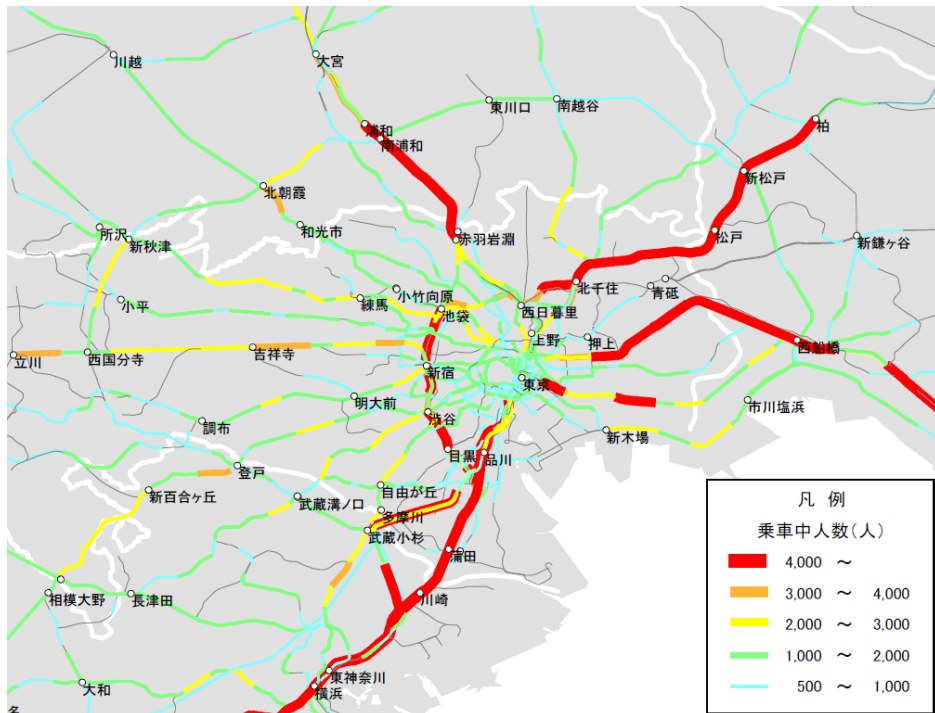


図 1.1.2 大都市交通センサス(2010 年度)による線別駅間移動人員数の推計

一方、近年では大規模に普及しているスマートフォン等の携帯電話に搭載されている GPS 機能から収集出来る GPS データを用いることで、広域かつ継続的に大量の人々の移動状況を把握することも可能になりつつある（関本ほか（2011）[5]）。また、このように GPS 機能は自動的にデータを取得することが可能であるため、利用者の負担も小さく、継続的な調査が可能となる。ただし、地下鉄のような GPS が機能しない場所では移動軌跡を取得す

ることは難しい。したがって、GPS データを用いて鉄道利用者の行動を把握するには地下区間のようなデータの欠損部分を補間する必要がある。

GPS データの欠損部分を補間する手段として経路探索を用いた経路補間があるが、経路補間には鉄道の路線・駅の接続情報すなわち位相情報を含んだ鉄道ネットワークデータが必要である。また、GPS データには誤差も含まれるため、これを修正する代表的な方法としてマップマッチングが挙げられるが、マップマッチングを行うためには鉄道の形状すなわちジオメトリ情報を含んでいる必要がある。現在、容易に入手が可能な鉄道データとして、国土交通省が運営するオープンデータサービスの国土数値情報ダウンロードサービス[6]に鉄道データがあるが、これは駅の位置情報・鉄道路線の位置及び形状のみのデータであり、各路線区間が実際に運行している路線のサービスと一致しない部分が多く存在している。また、駅データ.jp[7]では路線の運営区間、駅の配列及び位置情報は実際のサービスに非常に近い形の情報を含んだデータのほとんどを無料で提供しているが、路線の形状のようなジオメトリの情報は含んでいない。経路探索及びマップマッチングの両方を行うためには鉄道の幾何情報（ジオメトリ）と位相情報（トポロジ）の両方を含んでいる必要があるが、現状入手可能なデータにはこのような構造を持った鉄道ネットワークデータは存在していない。

1.2 研究の目的

これまで、GPS データを用いた鉄道利用者の行動把握に関する研究は多く行われてきた。しかし、これらはある特定の区間に絞ってデータを収集し、解析及び分析を行ってきたものであり、大規模な GPS データを用いた横断的な鉄道利用者の行動把握までは行えていない。現在では GPS 機能を搭載した携帯電話が広く普及しているため、今後 GPS データを用いたユーザーの行動把握の需要は増えると予想される。さらに、鉄道利用者が世界においても非常に多い日本においては、特に鉄道利用者の行動を横断的に把握できることは非常に重要である。鉄道事業者は他事業者の路線を含めた鉄道利用者の途中利用区間を正確に把握する手段がないことから GPS データを用いた鉄道利用者の行動把握を行うことが出来れば鉄道事業者はこれまで明らかにすることが難しい、もしくは手間のかかっていた人の流れを把握することができ、電車の混雑に対する新しいアプローチや鉄道運行本数の調整といったことに利用できる。しかし、これまで GPS データを用いた鉄道利用者の行動把握は横断的には行われておらず、まずはその定量的な精度を明らかにする必要がある。

そこで本研究では、携帯電話の GPS 機能から取得・蓄積をした GPS データを用いてユーザーの鉄道利用情報推定を行い、大都市交通センサスといった推計との比較を行うことで GPS データを用いた横断的な鉄道利用者数把握について定量的な精度を明らかにする。本研究において提案する手法では GPS データで途中区間を取得可能であるという特徴を活かすために GPS データの誤差・欠損区間をマップマッチング及び経路探索を行うことで補正・補間を行う。しかし、これらの補正・補間手法を用いることに適した鉄道ネットワークは存在していないため、まず既存の鉄道データを組み合わせることで現状のサービスに近い形かつマップマッチング及び経路探索が可能な鉄道ネットワークデータを作成する。そして、GPS 観測困難な地下鉄区間も含めた鉄道のネットワークデータを用いて途中区間を推計し時間連続な位置データとすることで鉄道利用者の利用実態把握を行う。さらに、推計結果から各駅における乗降人数及び各路線の利用人数の集計を行うことで異なる鉄道事業者を跨いだ複数路線について横断的な利用状況把握の可能性を検討する。

2 既往研究

電車内の混雑把握サービス

電車内の混雑を把握することが出来るサービスとして、「JR 東日本アプリ」[8]がある。このアプリは JR 東日本の路線のみではなく首都圏の様々な路線の運行情報を把握でき、各列車の位置情報を表示する機能も備わっている。このようにユーザーに対して鉄道運行に関する情報を提供している中に「山手線トレインネット」というサービスがある。これは、山手線に限られているが、運行している列車の各車両の混雑度、温度といった情報を確認することが出来る。（図 2.1 参照）JR 東日本では、車両の重量を量ることで各車両に何名のユーザーが乗車しているかを推定し、混雑度を算出している[2]。



図 2.1 山手線トレインネットの利用画面
(JR 東日本アプリより引用)

このようなサービスに対して本研究の手法を用いることが出来れば、山手線に限らず、全路線について乗車人数を把握することが可能となる。このように本研究の手法は特別な機材も使わずに GPS 機能を搭載した携帯電話さえあればデータ取得・蓄積し、混雑を推定することも可能となると考えている。

移動軌跡データを活用した鉄道利用者の交通行動把握

移動軌跡データを用いた交通行動把握は GPS 機能を搭載した携帯電話が普及し始めた頃から注目されており，その中でも，2009 年に三菱総合研究所が行った「移動軌跡データを活用した鉄道利用者の交通行動把握」[9]は携帯電話の GPS 機能を用いて取得した移動軌跡データを用いて鉄道利用者の交通行動把握の可能性を検討・検証した．この研究では，調査モニターは携帯電話にアプリケーションをインストールし，鉄道への乗車時，乗り換え時，降車時にアプリケーションを操作することで交通行動を詳細に収集した．（図 2.2 参照）



図 2.2 取得した移動軌跡データ例

その結果において，ユーザーの混雑回避行動の把握や輸送障害時の迂回路把握（図 2.3 参照）といった鉄道を利用した移動の把握から，駅までに向かう行動，駅から目的地までの行動分析も行っていた．



図 2.3 輸送障害時の迂回行動

最終的に交通行動実態把握に対して GPS データを用いることは有効であると結論づけているが，GPS が取得できない地下区間のような場所についての移動経路把握は課題としている．このように GPS データを用いた交通行動把握は有効であるとされているが，やはり地下区間の移動経路は課題となってしまうことがわかる．しかし，本研究の手法では地下区間のような GPS 情報を取得しづらい場所の行動については鉄道ネットワークデータを用い

てマップマッチング・経路探索を行うことで対応をする．さらに，この研究では鉄道を用いた移動を目的としている被験者に協力してもらって収集したデータを用いていたが，本研究では目的を設定していないただの GPS のトリップデータを用いて鉄道利用者の行動把握を行う点は，より一般性がある点であると考えている．

携帯端末を利用した鉄道利用者数推定に関する研究

携帯端末を利用した鉄道利用者数推定の先行研究として，中野・沼野らが行った「無線 LAN アクセスポイントへの検索要求を利用した鉄道車内混雑推定」[10]がある．彼らは無線 LAN を搭載した携帯端末は周辺のアクセスポイントを検索するためにプローブ要求を発信していることに着目し，このプローブ要求を用いて鉄道車内の混雑推定手法の提案を行った．

彼らの提案手法では，携帯電話から発信される端末の MAC アドレスのハッシュ値を ID として用いて，鉄道車内にある携帯電話と鉄道車外にある携帯電話の分類を行うことで鉄道利用者が乗車していた区間を推定するというものであった．この手法を用いた評価実験では，ユーザーが乗車していた区間の一部を推定する成果はあったが，ユーザーの乗車駅と降車駅の推定まで至ることはできなかった．また，ID として用いた MAC アドレスは鉄道車内でのみ一意な ID としていた．すなわち，ユーザーが列車から降車してしまうと ID が変わってしまうため，列車から降車したユーザーが駅から出たか，列車の乗り換えを行ったかなどの乗車前や降車後の行動を把握するには至らなかった．

本研究では，携帯電話の GPS のログデータを用いることでユーザーの行動は鉄道路線やその鉄道事業者によらず把握することが可能である．この点が本研究における大きな利点であると考えられる．

3 鉄道利用路線の推定手法

3.1 手法の概要

本研究で取得を想定する GPS データにはユーザーID, 時刻, 座標のみが含まれているため, まずは各観測点が滞在中・移動中に観測されたものなのか, また, どのような交通手段を用いていたかを推定する必要がある。

本研究では, まずは大野らの手法[11]を用いて GPS データに対して滞在・移動と交通手段の推定を行う。大野らの手法において交通手段の推定は平均速度, 移動距離をパラメータとして推定を行った。特に交通手段が「鉄道」となる移動の推定を行う際には鉄道のジオメトリデータを用いて GPS の観測点が鉄道路線とどの程度の割合で近いかを基準に推定を行った。このようにしてトリップ分割が済んだ GPS データから交通手段が「鉄道」であるトリップデータを用いる。

ここまでの処理では鉄道を利用しているユーザーを抽出したため, 次は実際に使用したと思われる鉄道路線, 駅を推定する。本研究ではユーザーの途中経路における GPS 観測点も利用することで GPS が観測できている部分に関してユーザーの鉄道利用情報の推定をより実態に近づけるため, 「スパースマップマッチング」[12]を用いて実際に利用した鉄道路線と駅の推定を行った。また, GPS が観測できない地下区間については経路探索を行うことでデータの欠損区間の補正を行った。そして, 利用した鉄道路線と駅の推定が完了したデータを鉄道利用者の移動データとした。さらに, 本研究の手法を実際に GPS のログデータに対して適用した際の推定精度について検証を行った。また, 鉄道利用者の移動データを駅ごと, 駅間について集計を行い, 国土交通省の算出した推計との比較を行うことで鉄道利用状況把握への GPS データの利用可能性について検討を行った。

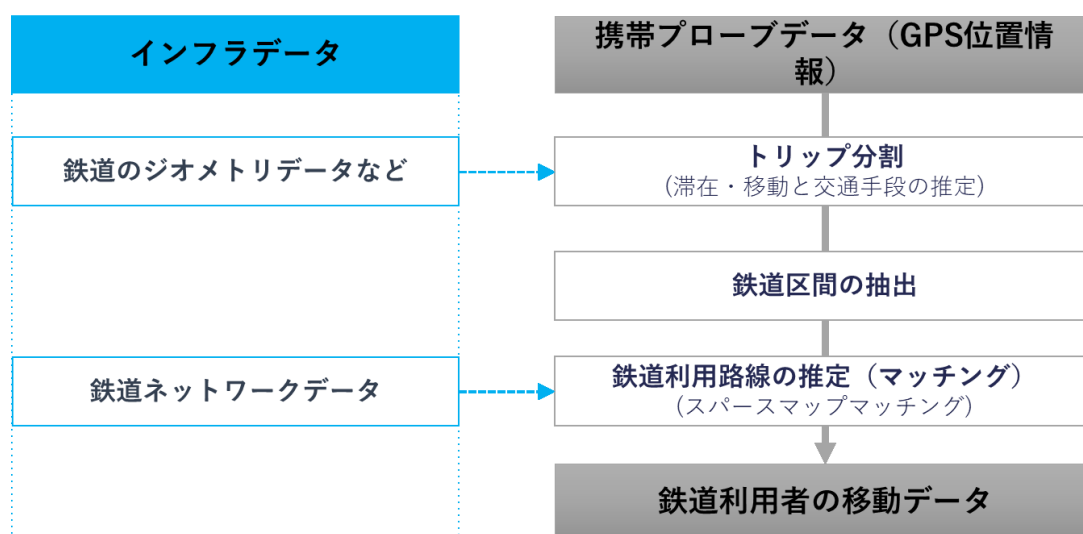


図 3.1.1 本研究の作業工程の概要

3.2 本研究に用いたマップマッチング手法

鉄道利用路線の推定では, 本研究で用いた GPS データは時間分解能が粗いため, 原らの提案したスパースマップマッチング[12]を用いた。元のスパースマップマッチングでは GPS 観測点のマッチング先を最近隣のノードとしていたが, 本研究で使用するデータでは駅がノードとなり解像度が低くなるため, 最近隣のリンク(鉄道路線)上に近傍点を取り, その近傍点をリンク補間点として探索を行うことで推定誤差を抑制する。

GPS のログデータを大野らの手法によりトリップ分割したデータを入力し、各観測点より最も近い鉄道ネットワークへのマッチングを行い、マッチングした鉄道ネットワーク上の観測点から垂線を下した位置に近傍点を取る。次にマッチングした鉄道ネットワークを用いて Dijkstra 法を用いた最短経路探索を行い、途中区間の補間を行った。(図 3.2.1 参照) また、探索の条件は鉄道を 40km/h、徒歩(乗り換え)を 3km/h、乗り換えをする際は所要時間を 5 分足すこととし、所要時間をコストにすることで探索を行った。

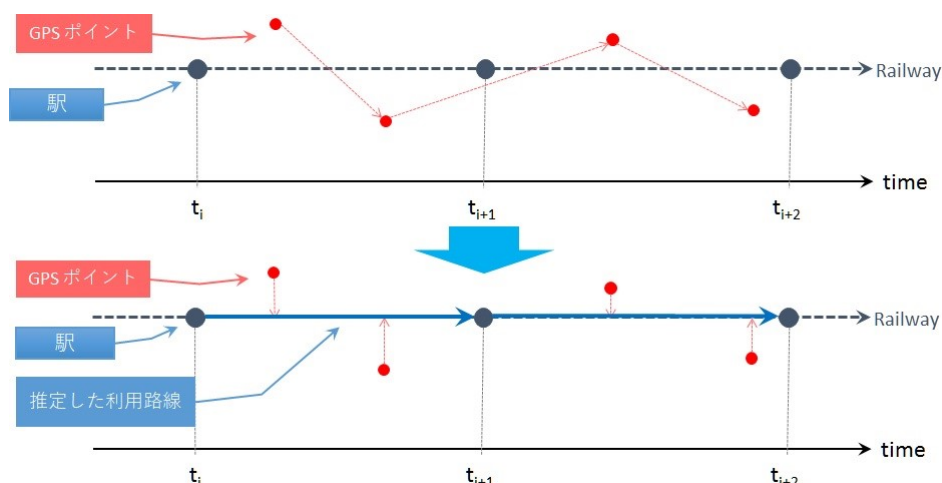


図 3.2.1 鉄道利用路線推定イメージ

次に駅間リンクにマッチングした位置でリンクを区切り、マッチング位置での時刻を GPS 観測時刻として各区間の速度を算出し、各駅における発着時刻を算出した(図 3.2.2 参照)。そのため、列車の停車時間は考慮していない。そして、発着時刻は GPS 観測点がマッチングした区間の平均速度から、各駅ノードにおける時刻を算出した。

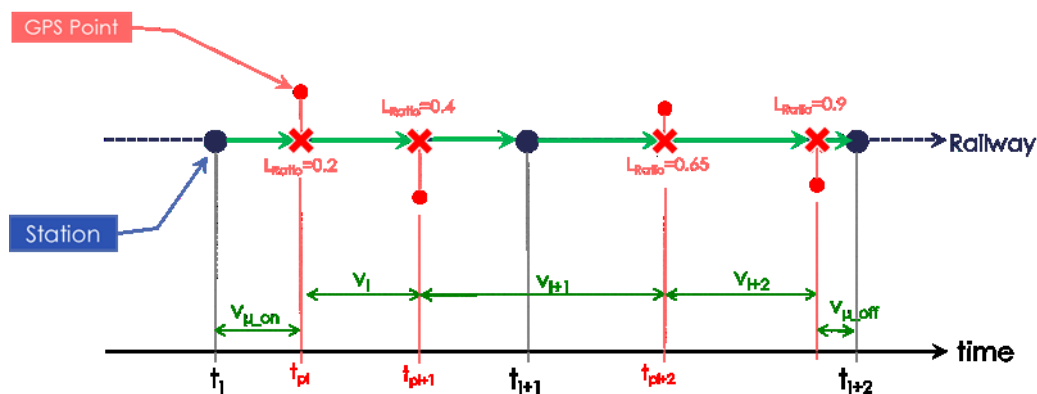


図 3.2.2 各区間での駅通過時刻推定イメージ

さらに、各ユーザーのトリップについて各ノード(駅)での通過時間及び使用路線の情報を記録したものを作成した。

鉄道利用路線の推定にあたり、他の手法として実際に鉄道の利用状況を教師データとして収集し機械学習を行うことで利用路線の推定・補間を行う方法があげられる。しかし、個別の鉄道利用者から教師データとして具体的な利用状況を効率的かつ一定数以上の規模で収集することは、スマートフォンアプリ等を利用したとしても容易ではない[18]。一方で、

鉄道の具体的な利用区間を整理するためにも、駅間の接続情報を含んだ鉄道のネットワークデータが不可欠となる。従って本研究では、他の研究手法においても共通して利用可能な基盤データとして鉄道ネットワークデータを整備し、探索結果の妥当性が鉄道ネットワークの正確さに影響される受けるスパースマップマッチングを用いて鉄道利用路線の推定を行った。

4 鉄道ネットワークデータ

4.1 使用したデータセット

本研究で用いた鉄道ネットワークデータの作成にあたり，2 つのデータセットを用いた．

鉄道路線形状データ

1 つ目は鉄道路線・駅の形状データとして国土数値情報ダウンロードサービスが提供する鉄道データセット(平成 27 年)[13]を用いた．本データは国土地理院「数値地図 25000(空間データ基盤)」，「電子地形図(タイル)」，(株)電気車研究会・鉄道図書刊行会「鉄道要覧(国土交通省鉄道局監修)」，各鉄道事業者の公式 HP 等を原典資料とし，全国の旅客鉄道・軌道の路線や駅について形状や属性を整備したデータセットである．本データの属性情報をそれぞれ表 4.1.1，4.1.2 に示す．

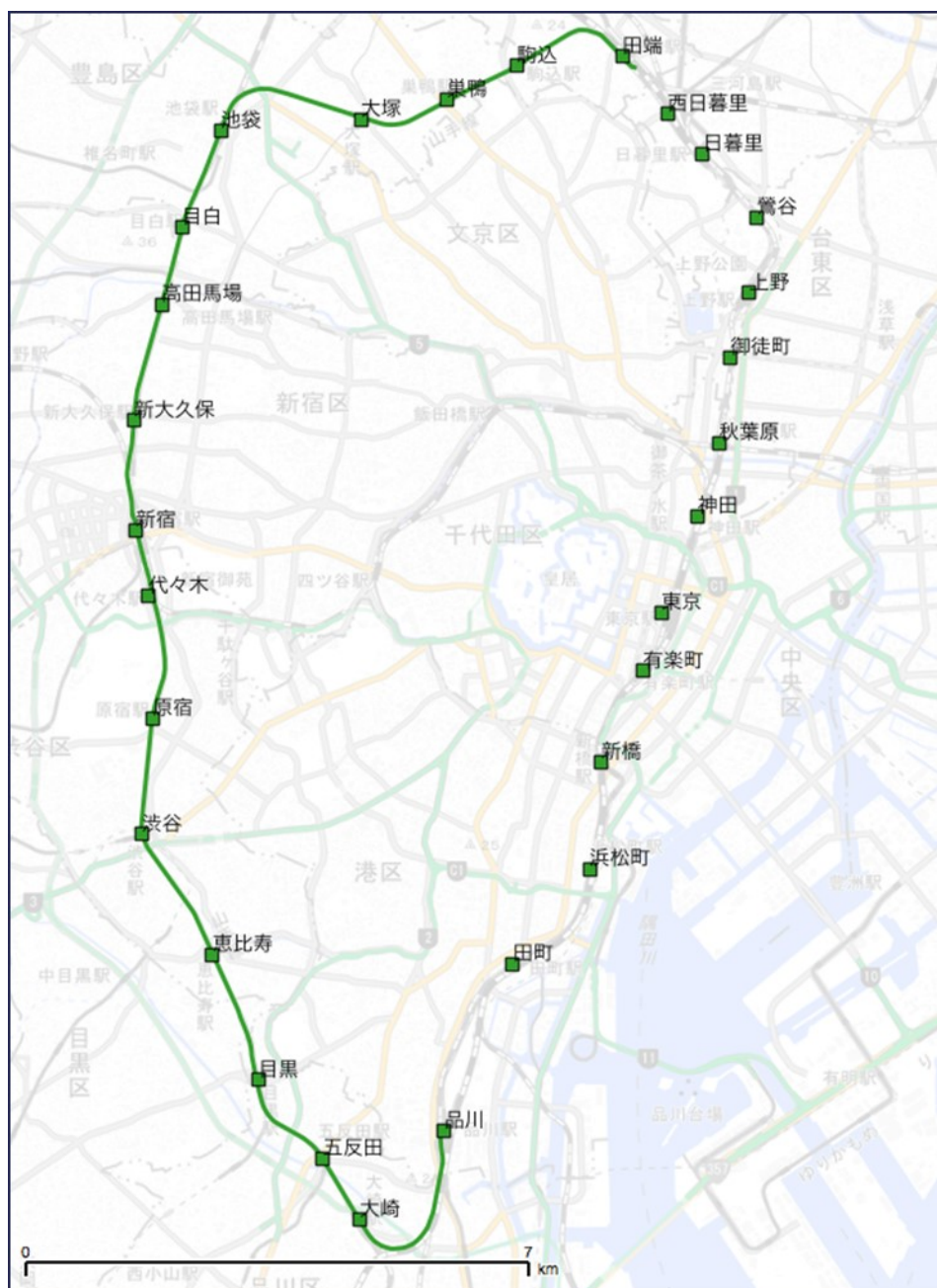
表 4.1.1 鉄道駅データの属性

1. 形状情報
2. 鉄道区分
3. 事業者種別
4. 路線名
5. 運営会社名
6. 駅名
7. 駅に関連している鉄道路線

表 4.1.2 鉄道路線データの属性

1. 形状情報
2. 鉄道区分
3. 事業者種別
4. 路線名
5. 運営会社名
6. 路線に関連している駅

しかし，本データは鉄道路線ごとの形状や属性を記録しているが，複数の路線で共有している区間があり，この区間に関しては元の路線の情報しか含まれていない．即ち，実際に運行している路線と一致しない区間が存在している．山手線を例に見てみると，実際のサービスでは山手線は環状線になっているが，本データにおいては山手線の東側にあたる田端駅から品川駅までの路線形状は含まれていない．(図 4.1.3 参照)



これは国土数値情報の鉄道データは帰属路線の属性についてのデータであるために発生していると考えられる. このように本データは実際の運行区間とは異なる箇所が多いため, 運行区間のデータとしては利用することはできない.

さらに、本データは隣接駅間におけるジオメトリのみのデータであるため、隣接駅間におけるジオメトリ同士、もしくは駅同士の接続情報は無い。本研究においてはマップマッチングを行った後に経路探索を行うため、路線・駅の接続情報（トポロジ）も必要となる。

そして、このような鉄道運営区間のままのネットワークデータで経路探索を行って経路補間を行うと余計な乗り換えが発生するといった実態に沿わない可能性がある。そのため、本データの鉄道区間をより実際のサービスに近い形で作成し直す必要がある。

駅データ

2 つ目は、鉄道路線の運営区間のデータとして、駅データ.jp が提供する駅の配列データを用いた。本データは各路線の駅の配列（各路線の停車駅）・接続情報と各駅の座標が記録されているデータセットであり、各路線の駅配列は国土数値情報ダウンロードサービスの鉄道データよりも実際のサービスに近い。本データは基本無料であるが、一部区間（新幹線）は有料となっているが、今回は両方を用いた。

しかし、駅データ.jp の駅データの駅の座標は路線ごとに分かれておらず、1 つの駅に対して1つの座標しか持っていない。例えば、東京駅には東京メトロの東京駅と JR の東京駅があるが、JR の東京駅のうち京葉線のホームのみ離れてしまっているが駅データ.jp における東京駅は JR で1つにまとまってしまっているため京葉線のホーム付近に駅が存在していない。（図 4.1.4 参照）このため、京葉線東京駅のような同じ駅であっても少し離れた位置にホームが存在する駅については実態に沿わない結果を生んでしまう。

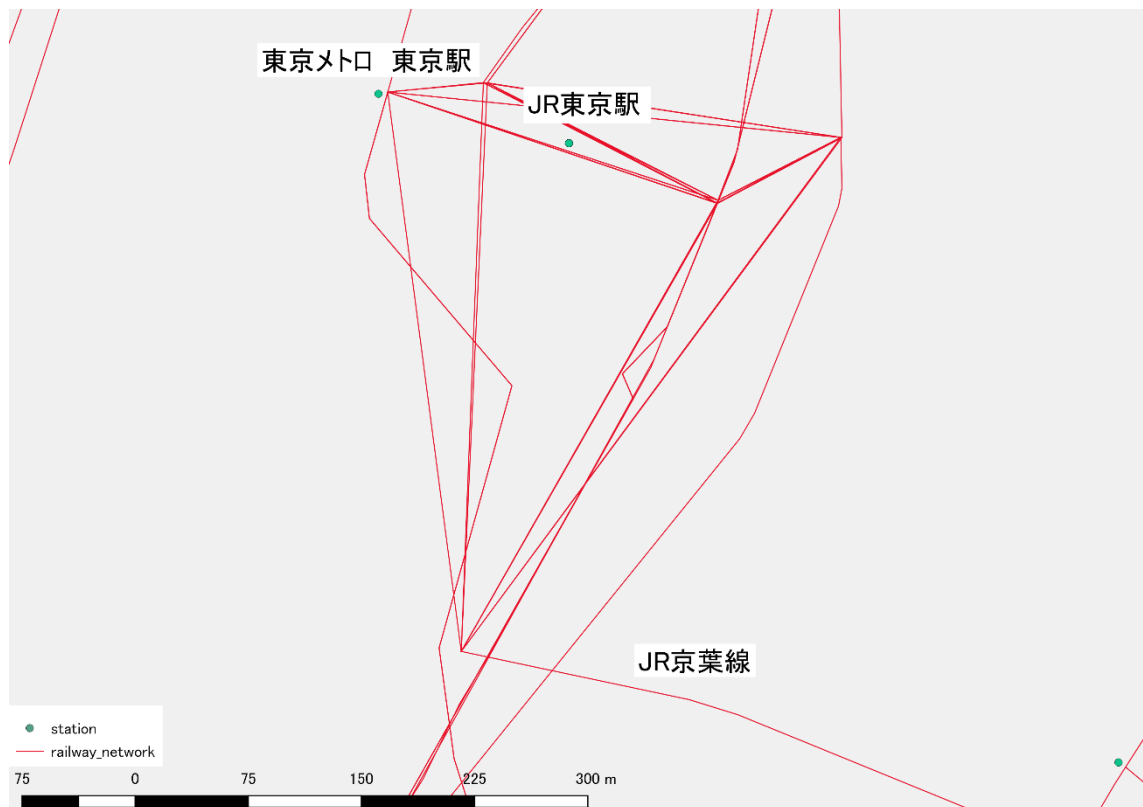


図 4.1.4 駅データ.jp の駅データにおける東京駅

しかし、実態は京葉線のホームのみ離れているため、これを図 4.1.5 のように修正しないとスパースマッチングの結果が実態に沿わない結果を生む可能性がある。

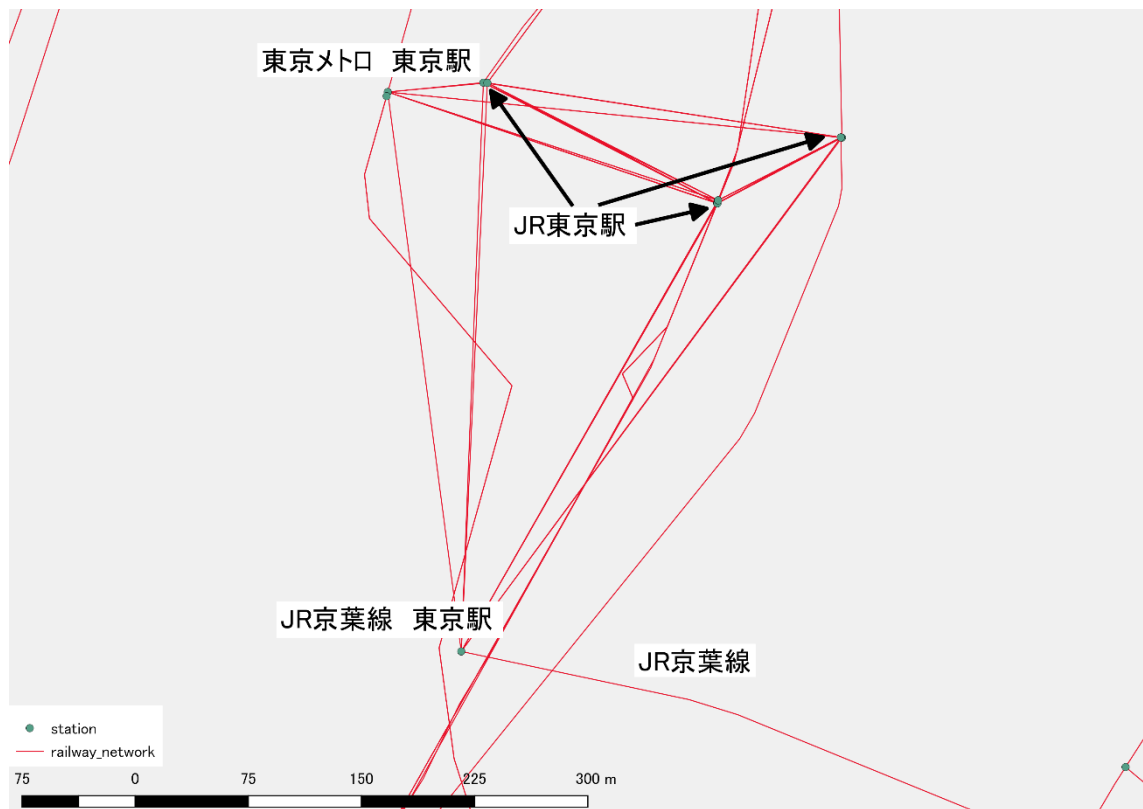


図 4.1.5 実態に沿わせた東京駅

また、本データは駅の情報のみであるため、路線の形状の情報は含まれていないため、やはり本データのみでも本研究の処理を行うには足りない。さらに、駅の座標も国土数値情報のものに比べて低精度であるという特徴もある。

以上より、2つの鉄道データについて以下のことがわかった。

- ・ 国土数値情報の鉄道データには路線形状が含まれるが、路線や駅の接続情報はない。
- ・ 駅データ.jp の駅データは路線や駅の接続情報はあるが、路線形状はない。

そして、路線形状は国土数値情報の鉄道路線データを用い、路線・駅の接続情報は駅データ.jp の駅データを用いることで各路線を帰属路線から運行路線として区間を整備し直し、乗り換えが可能な駅同士を単なる近傍駅間の組み合わせではなく駅データ.jp の駅データの接情報を用いて繋ぐことで新たな鉄道ネットワークデータを作成した。したがって、前述した山手線の東側のような駅データ.jp の駅データには路線があるが、国土数値情報の路線データには存在していなかった路線の区間が新たに生成される。

4.2 鉄道ネットワークデータの作成

作成方法の概要

図 4.2.1 に鉄道ネットワークの作成フローの概要を示す。

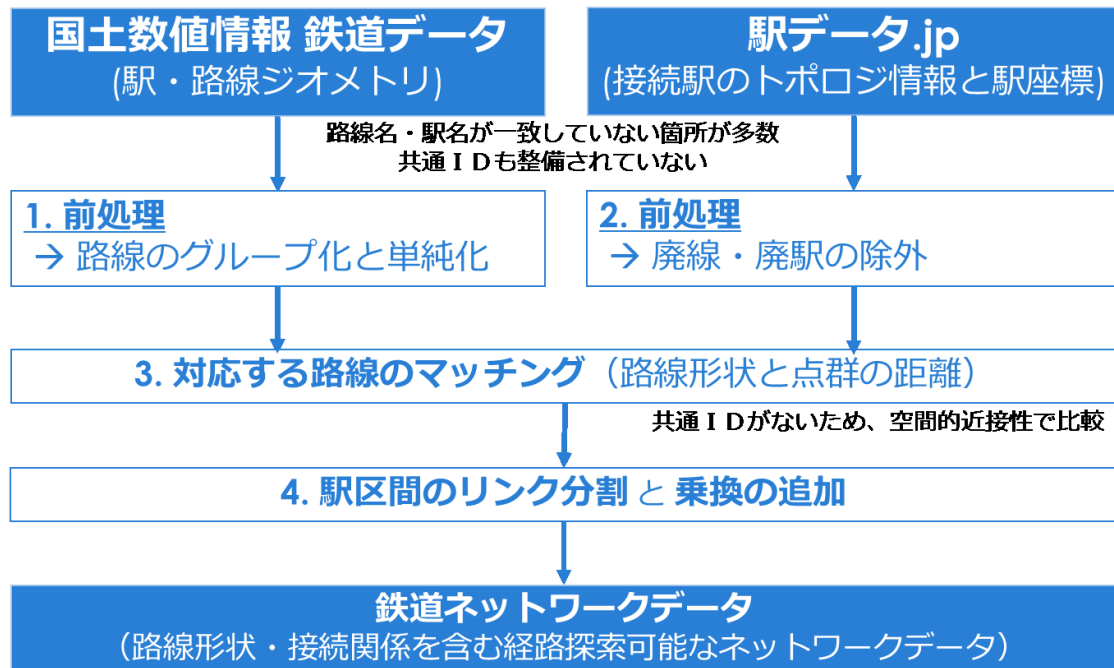


図 4.2.1 鉄道ネットワークデータ作成フローの概要

本研究における鉄道ネットワークデータは、本研究においてマップマッチングと経路探索を行う必要があるため、以下の2つの条件を満たすことを目的とした。

- ・ 駅データ.jp の駅配列データの駅配列・接続情報を持つ
- ・ 国土数値情報ダウンロードサービスの鉄道データの路線形状情報を持つ

鉄道ネットワークデータ作成は、まずは駅データ.jp における駅データからある路線の駅配列を取り出し、駅座標を用いて国土数値情報のジオメトリから候補となる路線を抽出し、駅データ.jp の駅データの駅座標から候補となったジオメトリ上に近傍点を取り新たな駅座標とする。そして、候補となったジオメトリを新たな駅座標で区切るといった流れで作成した。

次に各データに対する前処理について示す。国土数値情報の鉄道データにおける千代田線のラインを確認してみると、千代田線の路線データは単一のラインではなく43本のラインで構成されていることがわかる。（図 4.2.2 参照）

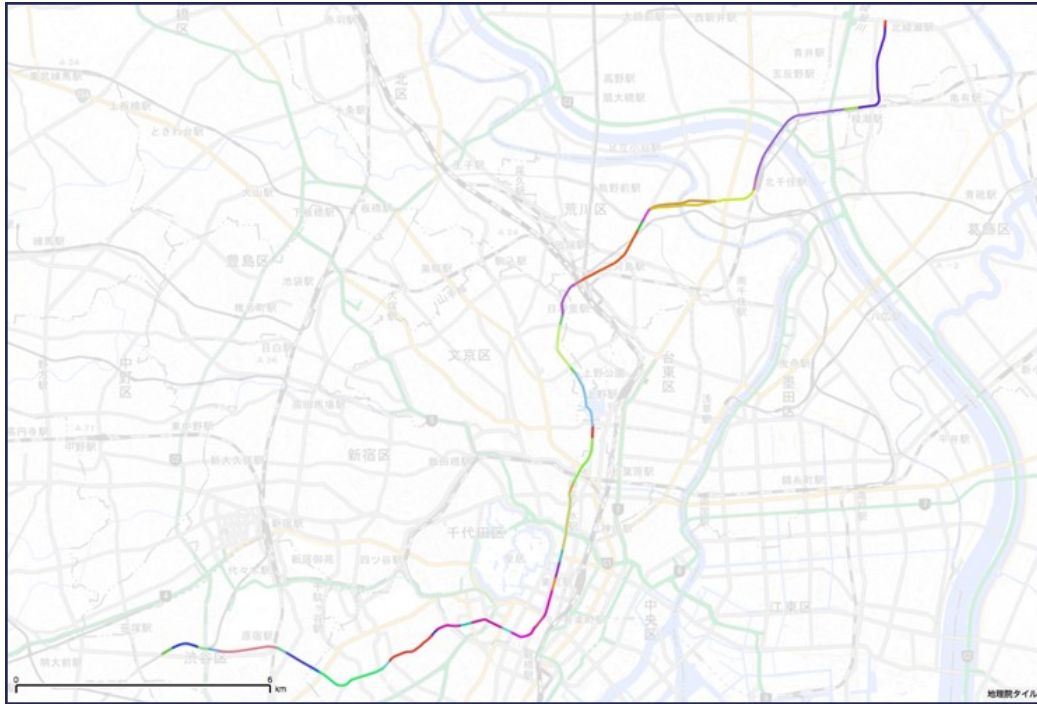


図 4.2.2 国土数値情報における千代田線

このように国土数値情報の鉄道データは、路線データが隣接駅間ごとにラインが記録されている。鉄道ネットワークデータの作成における候補路線の抽出では路線ごとの駅座標の座標配列を用いて最も形状に近いジオメトリを候補とするため、用いるジオメトリデータは路線ごとにグループ化されている必要がある。そこで、図 4.2.3 のように路線ごとにラインをグループ化し、単一のラインに簡略化を行った。

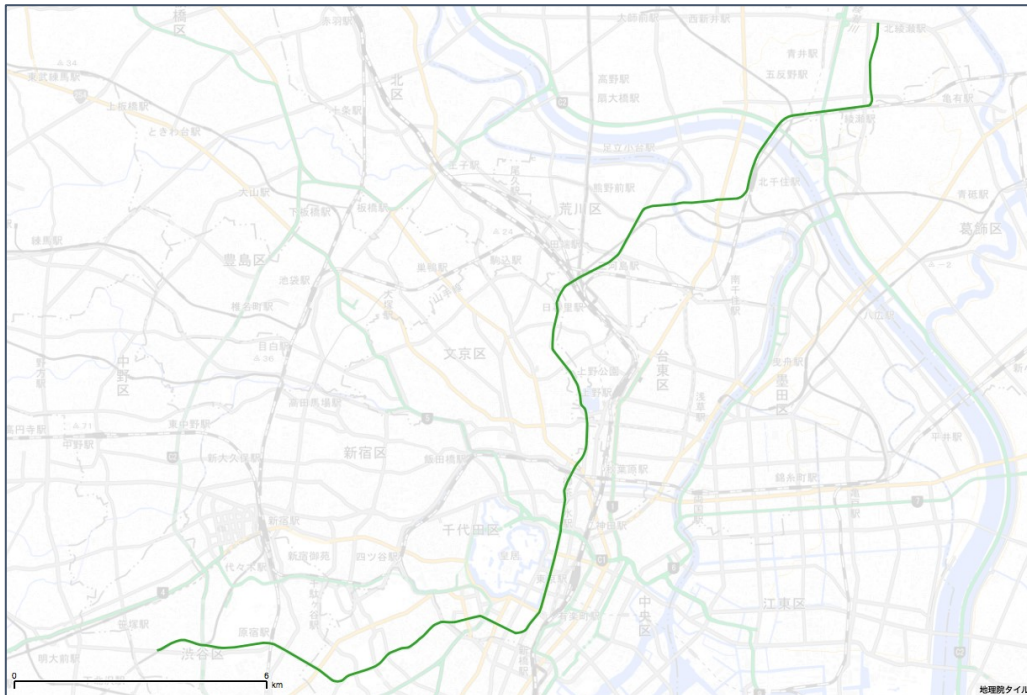


図 4.2.3 前処理後の千代田線のジオメトリ

次に、駅データ.jp の駅データに対する前処理として、廃線・廃駅の除外を行った。国土数値情報の鉄道データには該当区間の路線がないため、マッチング時の誤差原因になってしまうため本処理を行う必要があった。

さらに、国土数値情報の鉄道路線データは1つの路線で形状が複数の線に分かれている場合があるため、これを単線化する。これは複線の部分が存在していると駅データの駅座標をマッチングさせる際にどの路線にマッチングするかわからなくなってしまう、駅が路線上に乗らない可能性が出てくるという問題がある。路線の単線化は複線の部分について路線からバッファを発生させ、その中心を新たな路線形状とした。(図 4.2.4 参照)

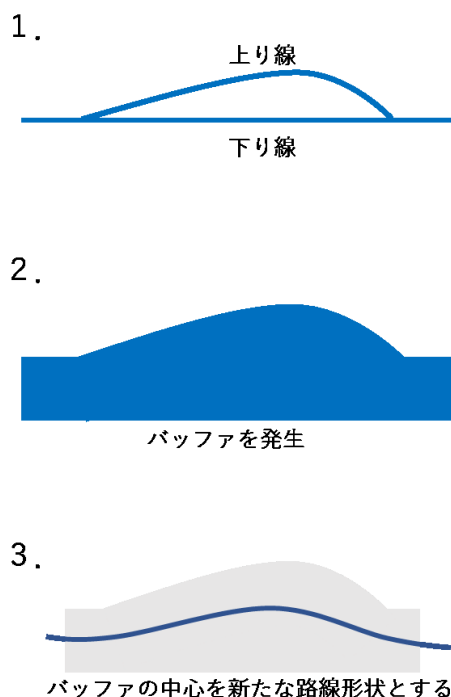
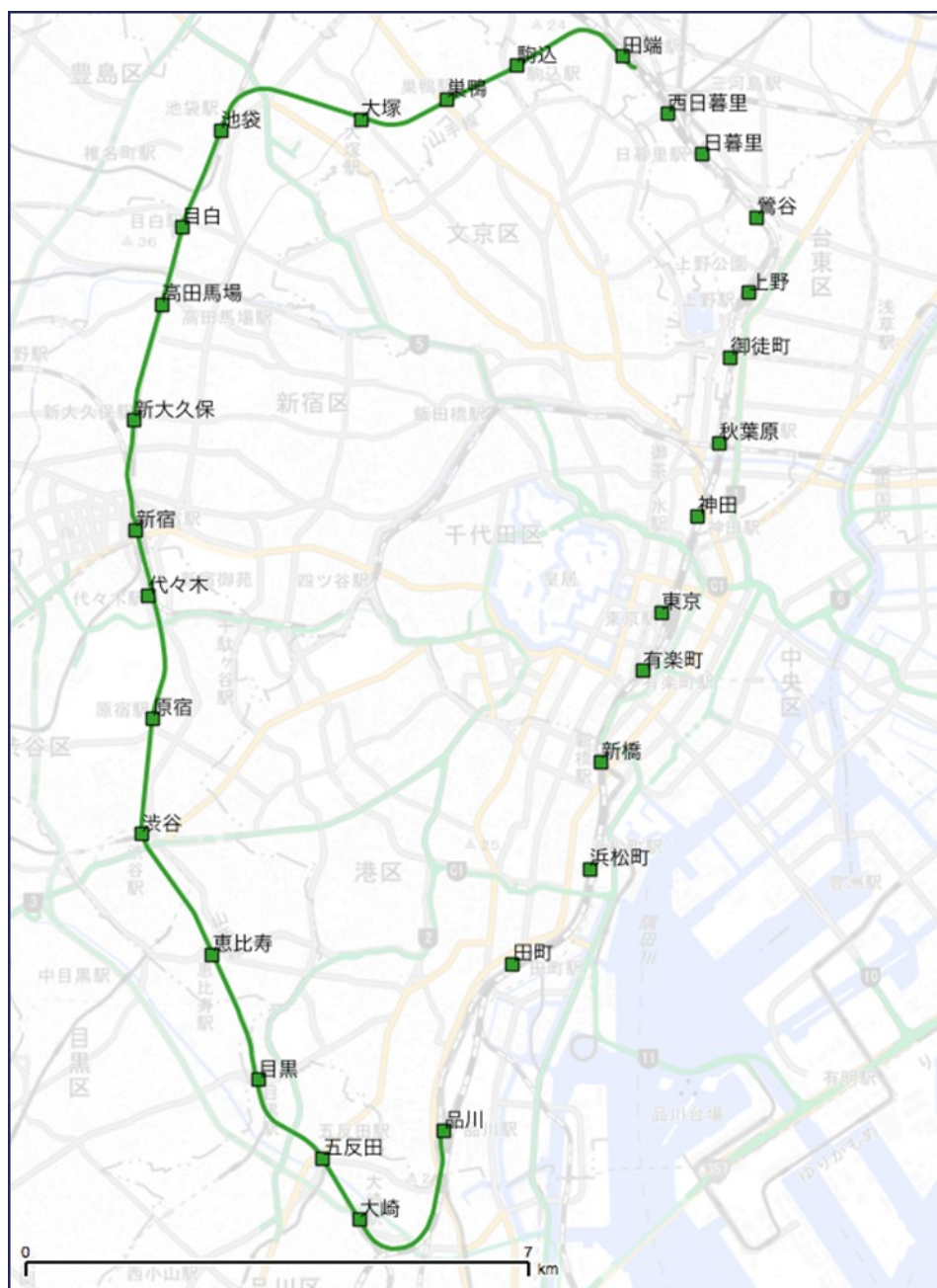


図 4.2.4 複線部分の単線化

次に駅データの駅座標とのマッチングを行う際に、まずは路線ごとに駅データ.jp の駅データから駅の配列を抽出し、駅の配列に対して候補となる路線形状を国土数値情報の路線データより抽出する。

候補の抽出はまず、ある路線の駅配列のそれぞれの駅座標からの距離の合計が最も近いものを候補として抽出する。(図 4.2.5 参照)



次に候補が見つかった駅を駅の配列から除外し、残っている駅配列で同様に候補を探す。
(図 4.2.6 参照)

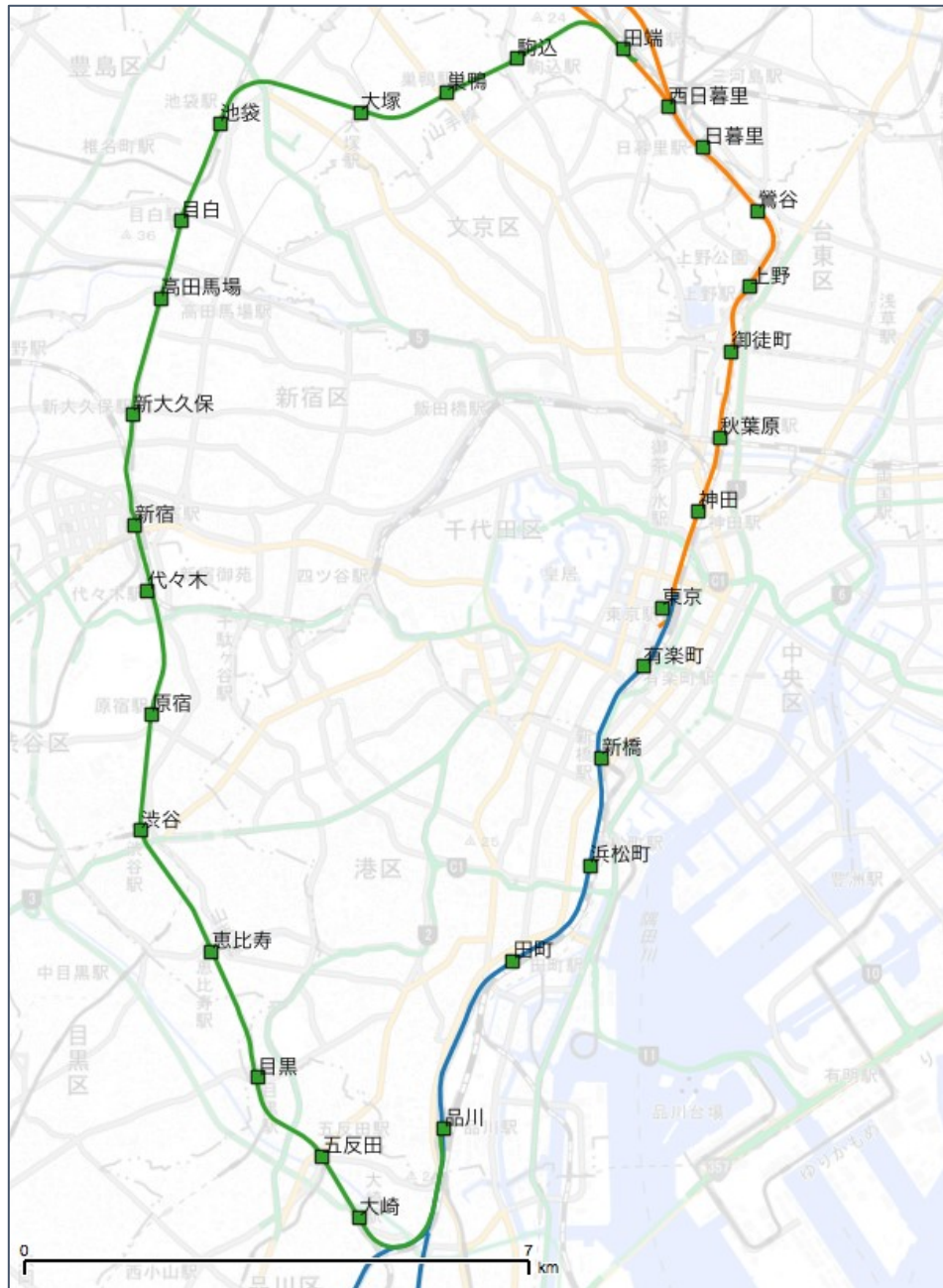


図 4.2.6 駅の座標配列からの候補路線抽出例 2

そして全ての駅配列から候補が見つかったところで、最後に候補同士をつなげるために、各候補路線よりバッファを発生させバッファの中心を通る路線形状を目的駅配列の路線形状の候補とする。(図 4.2.7 参照)



次に路線のジオメトリ候補に駅データの各駅を再近隣の座標に位置する路線上にマッチングさせる．そして，路線にある近傍点を新たな駅座標とし，新たな駅座標で路線の形状を区切ることにより，各駅間のネットワークを作成した．(図 4.2.8 参照)



図 4.2.8 各駅間のネットワーク作成

また、駅のジオメトリを区切る際に路線が分岐することがあり、その場合は分岐している点でネットワークが区切られてしまうため、経路探索を行って尤もらしいラインを取る必要がある。今回は Dijkstra 法を用いた最短経路探索を行った。

さらに、各路線リンクにはリンク ID・事業者コード・路線コード・リンクの各末端駅の駅コードを割り振っており、各駅ノードには事業者コード・路線コード・駅コード・駅グループコードを割り振った。駅ノード・駅グループコードとは乗り換えることが可能な駅同士において同一のコードを持つように割り振っており、乗り換えが可能かどうかは駅データ.jp の駅データの駅接続情報を元に判断した。ただし、駅ノード・路線リンクは各路線ごとに作成した。

以上の工程により、新たに経路探索・マップマッチング可能な鉄道ネットワークを作成した。また、作成した鉄道ネットワークデータにおける関東圏の各路線について抜け漏れなく繋がっているかを各路線の末端となっている駅を抽出することで確認を行ったところ関東圏においては各路線でネットワークが切れてしまっている部分は存在しなかった。

さらに、以前まで本研究で用いていた金杉らの作成した鉄道ネットワークデータ[14]と実際に経路探索を行うことで改善点の比較を行う。金杉らの手法を用いて作成した鉄道ネットワークデータは国土交通省から公開されている国土数値情報の 2011 年の鉄道のデータを元に作成した日本全国の鉄道のデータセットである。このデータセットは各駅をノード、各駅間の接続関係をリンクとした位相情報と、各駅間の路線形状を幾何情報として保有している。経路探索を行う区間は山手線の駒込駅から品川駅までとし、JR の路線を用いて探索する。この区間を選択した理由は国土数値情報の路線データには山手線の東側すなわち田端駅から品川駅の区間において山手線のデータが存在していないため、新たな鉄道ネットワークデータを用いた経路探索においてこの区間を山手線を用いた経路が選択されていれば国土数値情報の鉄道データのみを用いて作成した鉄道ネットワークデータよりも実態に沿えているといえる。

まず、金杉らの手法を用いて作成した鉄道ネットワークデータを用いて経路探索を行った結果を図 4.2.9 に示す。

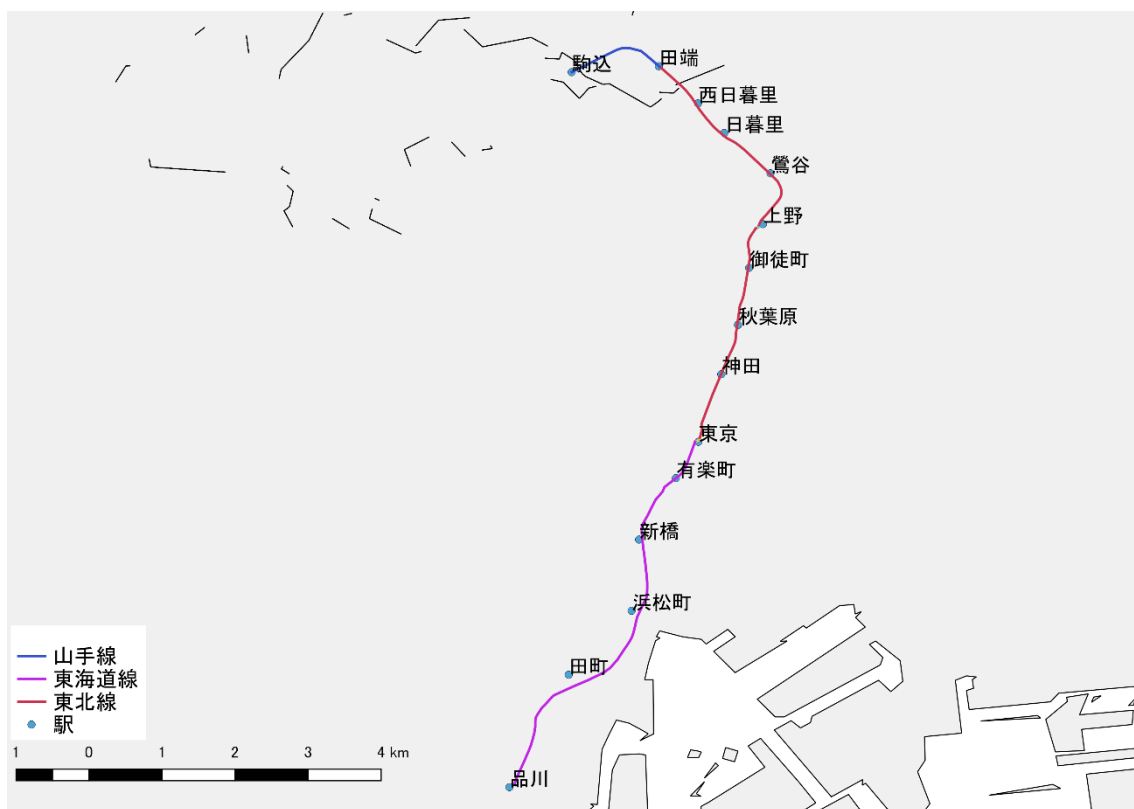


図 4.2.9 金杉らの手法を用いて作成した鉄道ネットワークデータを用いた経路探索結果

図 4.2.9 より，やはり山手線のネットワークが存在していない区間に関しては別の路線を用いた経路が選択されている．そのため，実際は発生しづらいような乗り換えが発生してしまっている．

次に，新たな鉄道ネットワークデータを用いて経路探索を行った結果を図 4.2.10 に示す．

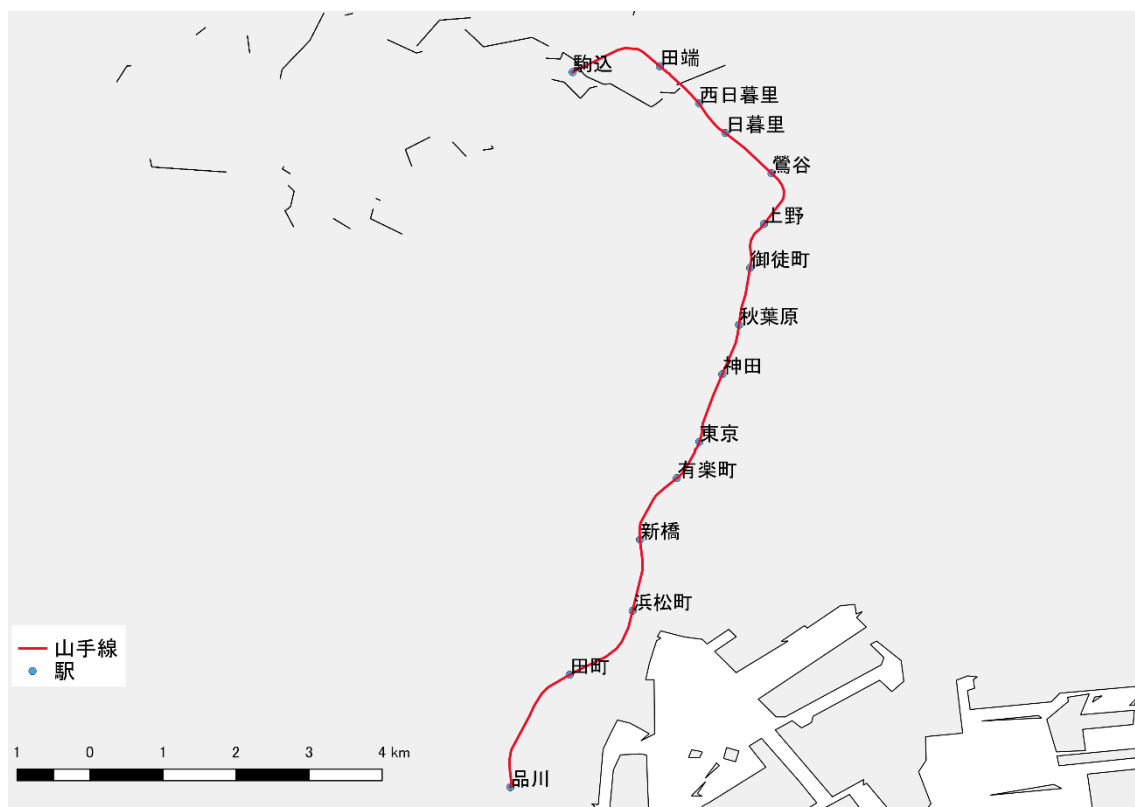


図 4.2.10 より，新たな鉄道ネットワークデータを用いた経路探索では山手線のみを利用した経路が選択されている．したがって，以前まで用いていた鉄道ネットワークデータには存在していなかった山手線のネットワークがしっかりと作成されていることがわかる．したがって，新たな鉄道ネットワークデータを用いて経路探索を行うことに，以前の鉄道ネットワークデータを用いて経路探索を行うよりも，実態に沿わないような乗り換えが減ると考えられる．

4.3 マップマッチング及び経路探索の検証

ここで新たな鉄道ネットワークを用いて本研究における手法を適用し、本研究の手法及び鉄道ネットワークデータの精度を検証する。データは情報銀行[15]が提供している GPS ロガーアプリを用いて蓄積した自身の移動履歴から鉄道乗車区間のみを抽出したものを 3 種類用意し検証に用いた。この GPS ロガーアプリでは最短 1 分で位置情報を記録している。検証においては地上区間の移動の推定精度を検証するための地上区間のみの移動、地下区間の推定精度を検証するために地下区間を含む移動、そして本研究の手法では平行する路線が複数ある場合に推定精度が落ちる可能性があるためこれを検証するため並行して運行する路線が複数存在する区間の移動の 3 点について検証を行った。

データ 1 : 地上区間のみの移動

1 つ目のデータの乗車区間は京王井の頭線北沢駅から京王相模原線南大沢駅までであり，明大前駅において京王井の頭線から京王線への乗り換えを行っている．また，図 4.3.1

に蓄積された GPS データのうち鉄道乗車区間のみを抽出し, GPS 観測点を地図上に表示したものを示す.

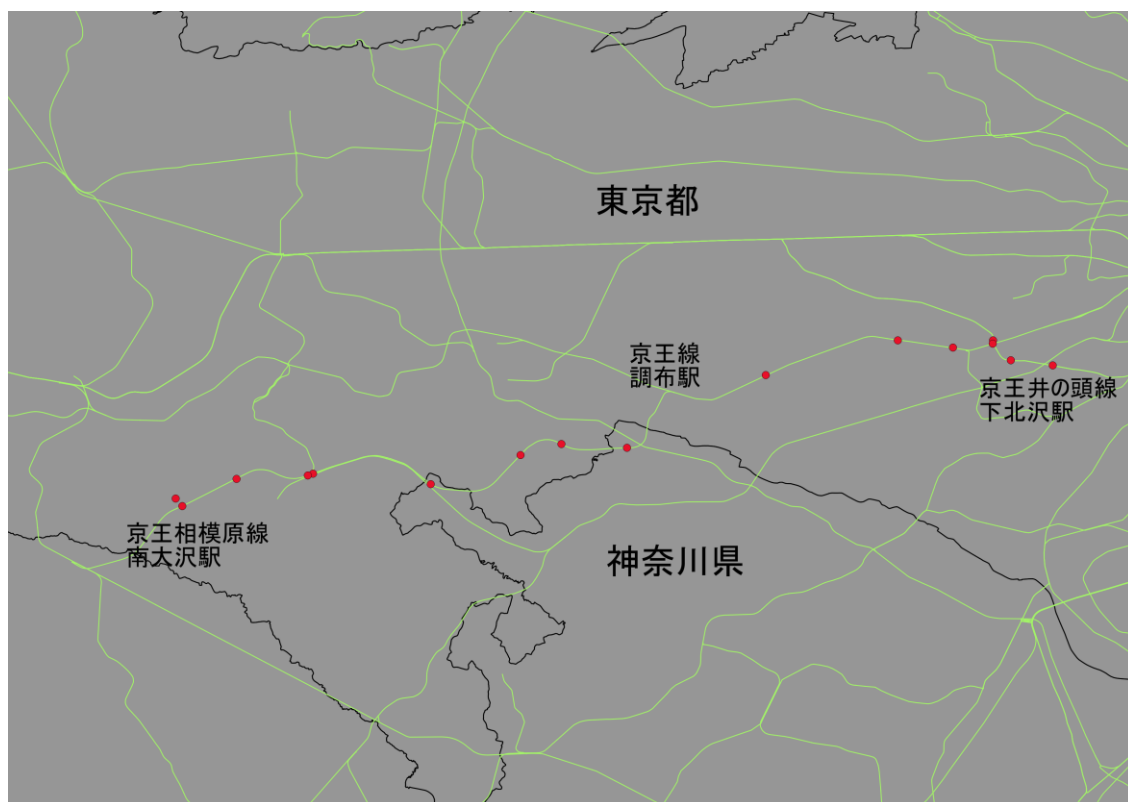


図 4.3.1 鉄道乗車区間のみ抽出した GPS データー1

次に図 4.3.2 に本研究の手法を適用し探索された経路における駅ノードを地図上に表示したものを示す。

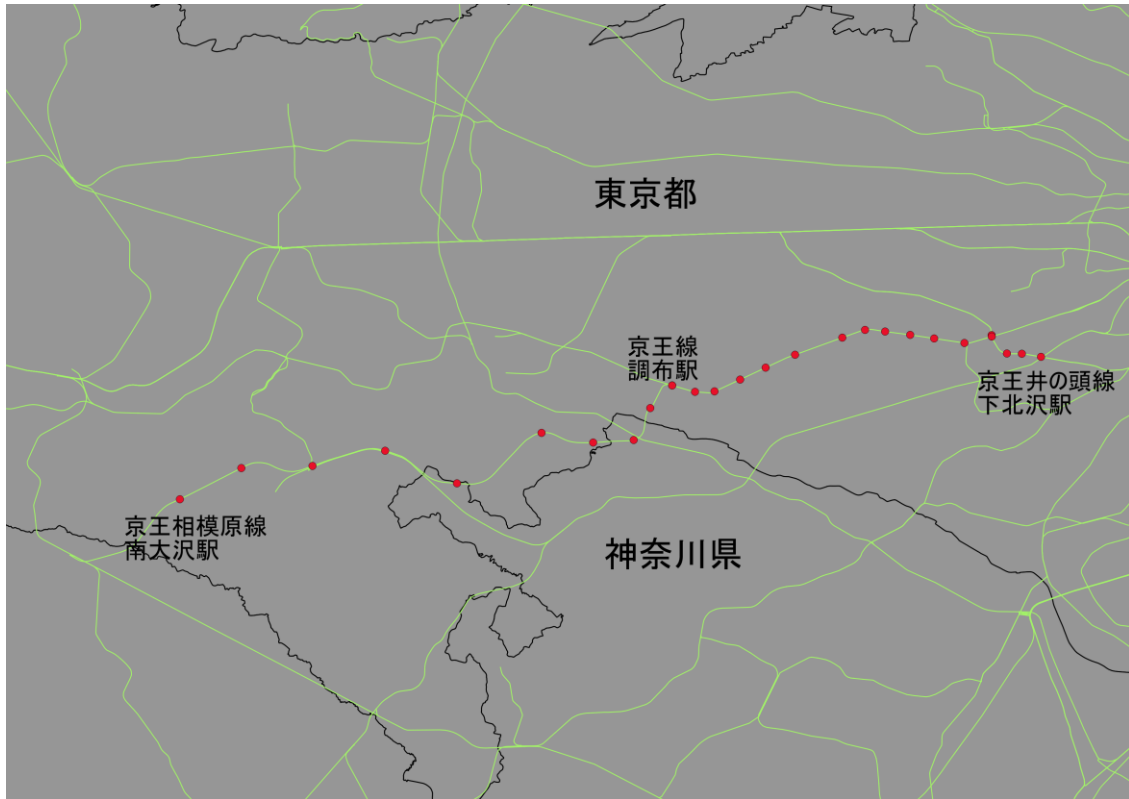


図 4.3.2 本研究の手法の適用結果－1

図 4.3.2 より、推定結果は実際の経路と一致しており、GPS データ（図 4.3.1）では欠損していた調布駅の区間も補完し経路を推定出来ていることがわかる。

また、表 4.3.3 にマップマッチングを行い、推定された利用路線・利用駅を示す。

表 4.3.3 推定された利用路線・利用駅

事業者 コード	事業者名	路線 コード	路線名	駅 コード	駅 グループコード	駅名
24	京王電鉄	24006	京王井の頭線	2400605	2400605	下北沢
24	京王電鉄	24006	京王井の頭線	2400606	2400606	新代田
24	京王電鉄	24006	京王井の頭線	2400607	2400607	東松原
24	京王電鉄	24006	京王井の頭線	2400608	2400106	明大前
24	京王電鉄	24001	京王線	2400106	2400106	明大前
24	京王電鉄	24001	京王線	2400107	2400107	下高井戸
24	京王電鉄	24001	京王線	2400108	2400108	桜上水
24	京王電鉄	24001	京王線	2400109	2400109	上北沢
24	京王電鉄	24001	京王線	2400110	2400110	八幡山
24	京王電鉄	24001	京王線	2400111	2400111	芦花公園
24	京王電鉄	24001	京王線	2400112	2400112	千歳烏山
24	京王電鉄	24001	京王線	2400113	2400113	仙川
24	京王電鉄	24001	京王線	2400114	2400114	つつじヶ丘
24	京王電鉄	24001	京王線	2400115	2400115	柴崎
24	京王電鉄	24001	京王線	2400116	2400116	国領
24	京王電鉄	24001	京王線	2400117	2400117	布田
24	京王電鉄	24001	京王線	2400118	2400118	調布
24	京王電鉄	24002	京王相模原線	2400201	2400118	調布
24	京王電鉄	24002	京王相模原線	2400202	2400202	京王多摩川
24	京王電鉄	24002	京王相模原線	2400203	2400203	京王稲田堤
24	京王電鉄	24002	京王相模原線	2400204	2400204	京王よみうりランド
24	京王電鉄	24002	京王相模原線	2400205	2400205	稲城
24	京王電鉄	24002	京王相模原線	2400206	2400206	若葉台
24	京王電鉄	24002	京王相模原線	2400207	2400207	京王永山
24	京王電鉄	24002	京王相模原線	2400208	2400208	京王多摩センター
24	京王電鉄	24002	京王相模原線	2400209	2400209	京王堀之内
24	京王電鉄	24002	京王相模原線	2400210	2400210	南大沢

表 4.3.3 より、乗車駅、降車駅、乗換駅が実際の経路一致していることがわかる。推定結果では調布駅において乗り換えを行っていることになっているが、これは駅データ.jp の駅データでは京王線内の直通運転がカバーしきれていないために発生している。データ 1 では経路ほとんどが地上区間であったが、推定は上手くいっていると考えられる。

データ 2：地下区間を含む移動

2 つ目のデータの乗車区間は埼京線武蔵浦和駅から有楽町線飯田橋駅までであり，池袋駅において埼京線から丸ノ内線への乗り換えを行っている．また，図 4.3.4 に蓄積された GPS データのうち鉄道乗車区間のみを抽出し，GPS 観測点を地図上に表示したもの示す．

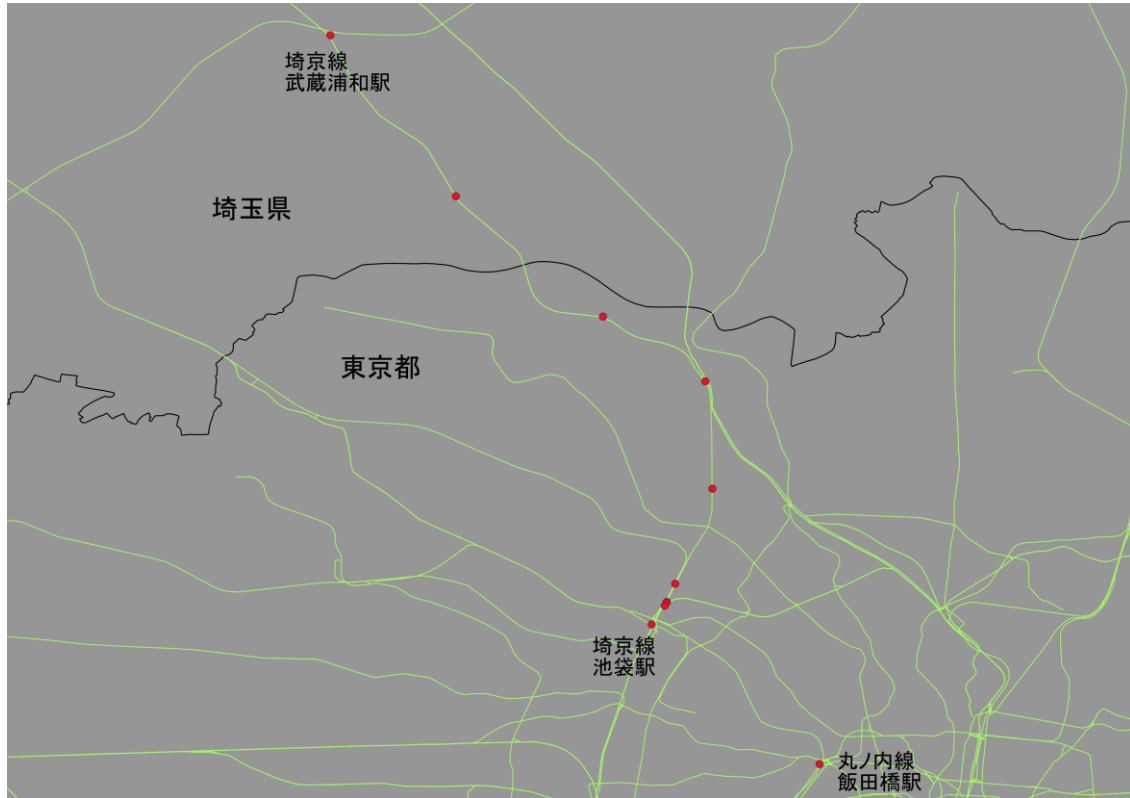


図 4.3.4 鉄道乗車区間のみ抽出した GPS データ-2

次に図 4.3.5 に本研究の手法を適用し探索された経路における駅ノードを地図上に表示したものを示す．

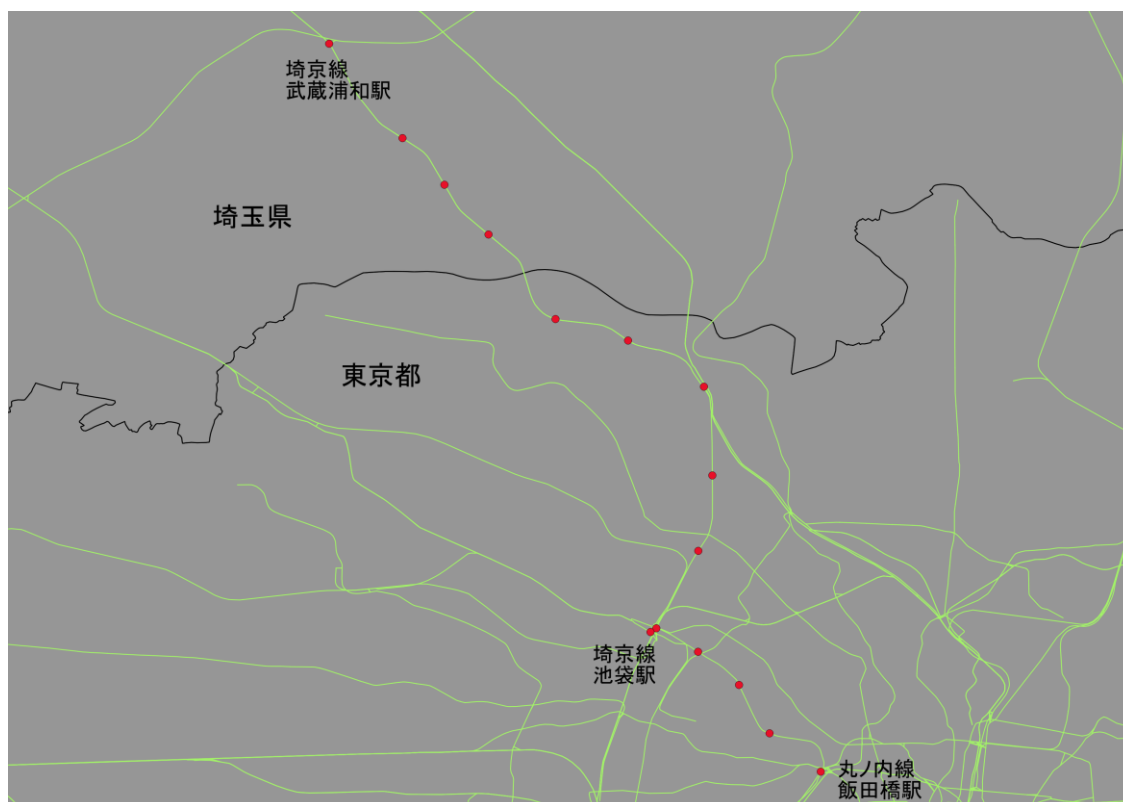


図 4.3.5 本研究の手法の適用結果－2

図 4.3.5 より，推定結果は実際の経路と一致していることがわかる．また，地下鉄の区間である池袋駅から飯田橋駅間において GPS データが欠損していたが補間することが出来ている．

また，表 4.3.6 にマップマッチングを行い，推定された利用路線・利用駅を示す．

表 4.3.6 推定された利用路線・利用駅

事業者 コード	事業者名	路線 コード	路線名	駅 コード	駅 グループコード	駅名
11	JR	11321	JR 埼京線	1132114	1130510	武蔵浦和
11	JR	11321	JR 埼京線	1132113	1132113	北戸田
11	JR	11321	JR 埼京線	1132112	1132112	戸田
11	JR	11321	JR 埼京線	1132111	1132111	戸田公園
11	JR	11321	JR 埼京線	1132110	1132110	浮間舟渡
11	JR	11321	JR 埼京線	1132109	1132109	北赤羽
11	JR	11321	JR 埼京線	1132108	1131903	赤羽
11	JR	11321	JR 埼京線	1132107	1132107	十条
11	JR	11321	JR 埼京線	1132106	1132106	板橋
11	JR	11321	JR 埼京線	1132105	1130212	池袋
28	東京メトロ	28006	東京メトロ有楽町線	2800609	1130212	池袋
28	東京メトロ	28006	東京メトロ有楽町線	2800610	2800610	東池袋
28	東京メトロ	28006	東京メトロ有楽町線	2800611	2800611	護国寺
28	東京メトロ	28006	東京メトロ有楽町線	2800612	2800612	江戸川橋
28	東京メトロ	28006	東京メトロ有楽町線	2800613	1131205	飯田橋

表 4.3.6 より，乗車駅，降車駅，乗換駅が実際の経路と一致していることがわかる．データ 2 では経路に地下区間を含んでいたが，推定は上手くいっている．

データ 3：並行して運行する路線が複数存在する区間の移動

3 つ目のデータの乗車区間は埼京線武蔵浦和駅から埼京線新宿駅までであり、乗り換えは行っていない。また、図 4.3.7 に蓄積された GPS データのうち鉄道乗車区間のみを抽出し、GPS 観測点を地図上に表示したもの示す。



図 4.3.7 鉄道乗車区間のみ抽出した GPS データー2

次に図 4.3.8 に本研究の手法を適用し探索された経路における駅ノードを地図上に表示したものを示す。

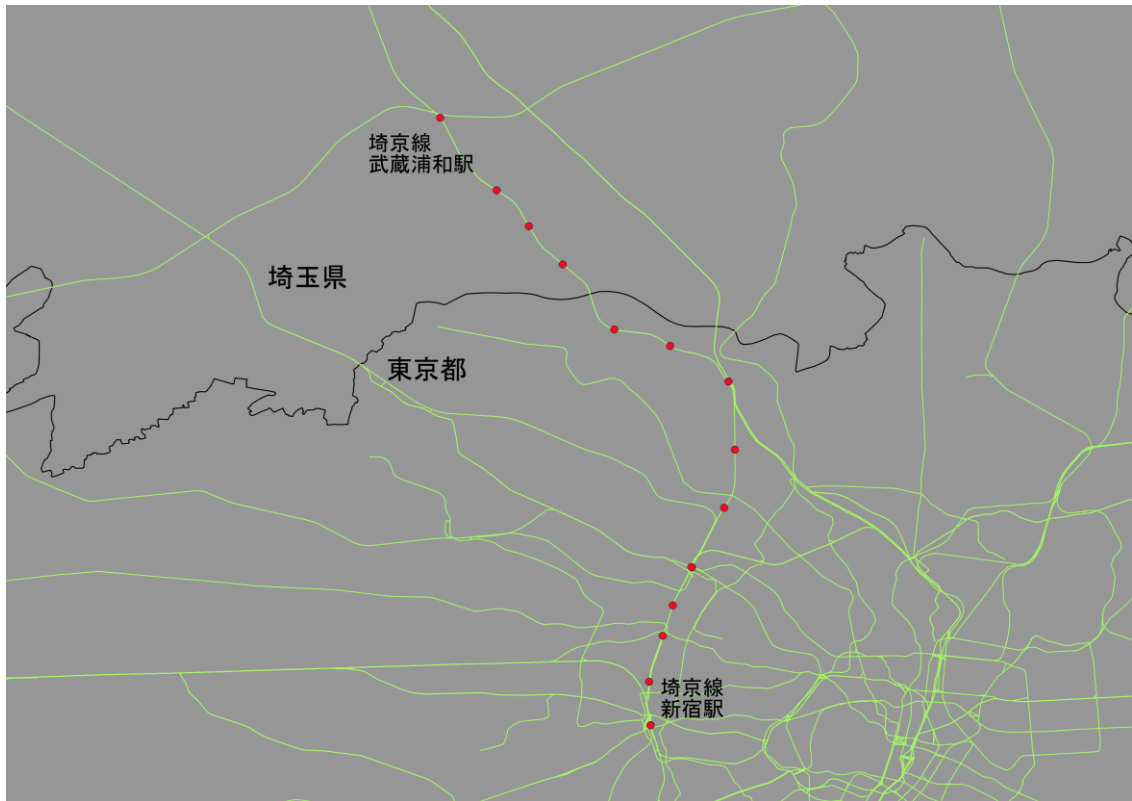


図 4.3.8 本研究の手法の適用結果－2

図 4.3.8 より，推定結果は実際の経路と一致していることがわかる．
また，表 4.3.9 にマップマッチングを行い，推定された利用路線・利用駅を示す．

表 4.3.9 推定された利用路線・利用駅

事業者 コード	事業者名	路線 コード	路線名	駅 コード	駅 グループコード	駅名
11	JR	11321	JR 埼京線	1132114	1130510	武蔵浦和
11	JR	11321	JR 埼京線	1132113	1132113	北戸田
11	JR	11321	JR 埼京線	1132112	1132112	戸田
11	JR	11321	JR 埼京線	1132111	1132111	戸田公園
11	JR	11321	JR 埼京線	1132110	1132110	浮間舟渡
11	JR	11321	JR 埼京線	1132109	1132109	北赤羽
11	JR	11321	JR 埼京線	1132108	1131903	赤羽
11	JR	11321	JR 埼京線	1132107	1132107	十条
11	JR	11321	JR 埼京線	1132106	1132106	板橋
11	JR	11321	JR 埼京線	1132105	1130212	池袋
11	JR	11302	JR 山手線	1130212	1130212	池袋
11	JR	11302	JR 山手線	1130211	1130211	目白
11	JR	11302	JR 山手線	1130210	1130210	高田馬場
11	JR	11302	JR 山手線	1130209	1130209	新大久保
11	JR	11302	JR 山手線	1130208	1130208	新宿

表 4.3.9 より、乗車駅、降車駅、乗換駅が実際の経路一致していることがわかる。しかし、実際にあれば、埼京線はほとんどの時間帯で新宿駅まで運行し、他の路線と比べて池袋駅から新宿駅までの所要時間が短いため、武蔵浦和駅から新宿駅向かう際には埼京線から乗り換えることは少ないことが考えられる。このような探索結果になってしまったのは、池袋駅—新宿駅間において埼京線のラインと山手線のラインが重なっている部分が存在し、マッチングの際にどちらの路線が選ばれるかわからないことが原因であると考えられる。

検証結果

検証の結果、本研究における鉄道利用路線推定手法及び作成した鉄道ネットワークデータを用いた鉄道利用路線の推定は地上においても、地下においても極端に実際とは異なるような経路を選択していないため、妥当な推定結果が得られた。しかし、平行して多くの路線が運行している区間では実際と異なる推定結果になってしまう可能性を有している。

5 鉄道利用路線の推定

5.1 使用したデータセット

「混雑統計®」(株式会社ゼンリンデータコム)は、株式会社 NTT ドコモが提供する「ドコモ地図ナビ」において、利用者の承諾を得て最短5分間隔で取得・蓄積された位置情報データを集計し、人々の流動パターンを定量化・可視化した統計データである。なお、「混雑統計®」で用いられるデータは、株式会社ゼンリンデータコムの依頼により株式会社 NTT ドコモにて個人が特定されないよう総体的かつ統計的に加工したもの(※以下同)であり、性別・年齢等の属性は含まれない。さらに本研究では同データから大野の手法により移動手段の推定値を与え、交通モードが「鉄道」と推定されたものを抽出して用いた。なお混雑統計®に対する同処理は株式会社ゼンリンデータコムが行った。これ以降の混雑統計®に関するデータの総体的処理も、株式会社ゼンリンデータコムが行った。このデータのうち、本研究では、2013年7月22日から2013年7月28日の期間で関東圏(東京都、埼玉県、神奈川県、千葉県の一都三県)を通過するものを利用した。また、この期間における鉄道利用者数のID数とトリップ数は図5.1.1のようになっている。

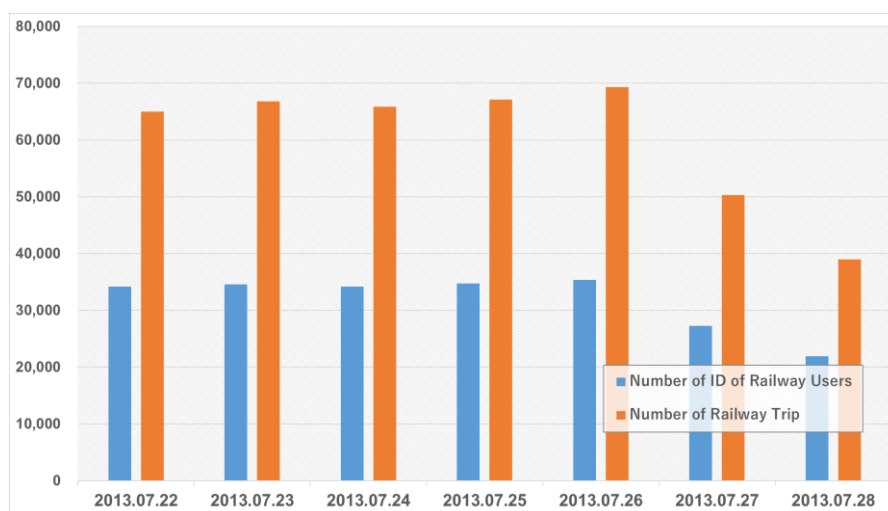


図 5.1.1 鉄道利用者数のID数とトリップ数の推移

5.2 推定結果

5.2.1 鉄道利用者数の時間的推移

図 5.2.1.1 に時間帯別の鉄道の利用状況を示す。図 5.2.1.1 より、鉄道利用状況のピークは7時台、8時台及び18時台であることがわかる。これは朝通勤や通学によって多くのユーザーが鉄道を利用し、夕方には帰宅のために同じく多くのユーザーが鉄道を利用するためにこのような時間帯にピークがあると考えられる。また、朝のピークが7時と8時に分かれているのは都心の多くの企業や学校の開始時間が8時頃から9時頃であることが原因であり、夕方のピークが朝のピークより少ないことは学生と社会人の帰宅時間が多様であるからであると考えられる。

また、本研究では休日の利用者数の時刻分布も算出可能であるため、大都市交通センサス(2010年度)における鉄道利用者数の推計や鉄道事業者提供資料を元に算出した国土数値情報ダウンロードサービスにおける駅別乗降客数[16](2013年度)からは見られなかった平日と休日利用者数の変化を見ることが可能である。やはり、休日は多くの企業や学校が休みになるため、平日よりも大幅に減っていることがわかる。そのため、一日の中での利用者数のピークは休日では見られない。

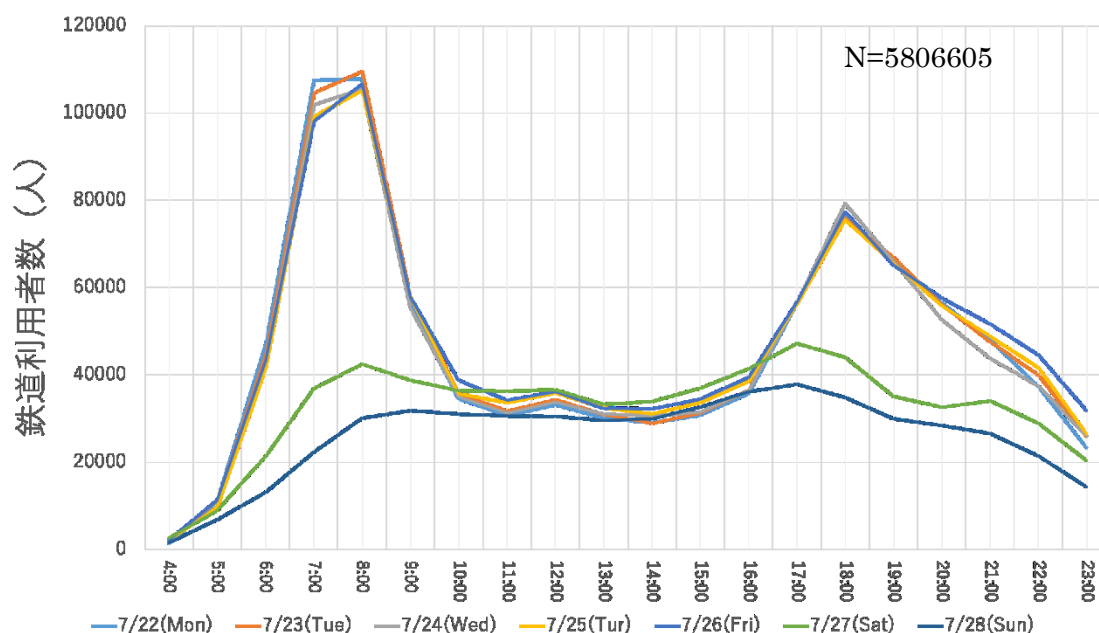


図 5.2.1.1 鉄道利用者の時刻分布

図 5.2.1.2 に大都市交通センサスより抜粋した首都圏における平日の鉄道利用者による駅・列車内滞留者数の時刻分布を示す。大都市交通センサスの推計が平日の鉄道利用者を対象としているため、図 5.2.1.1 における平日の時刻分布と図 5.2.1.2 の比較を行う。また、大都市交通センサスでは鉄道利用者の時刻分布のデータは公開していないため、数値的な比較は行えない。図 5.2.1.1 の昼間の利用者のピークは 7 時台及び 8 時台にあり、図 5.2.1.1 の昼間にピークは 8 時にある。また、図 5.2.1.2 における夕方のピークは 18 時台にあり、図 5.2.1.2 における夕方のピークは 18 時半にある。さらに昼間の 12 時頃ではどちらもピークに比べて利用者は少ない。以上のことから鉄道利用者数の時刻分布は非常に似ていることがわかる。

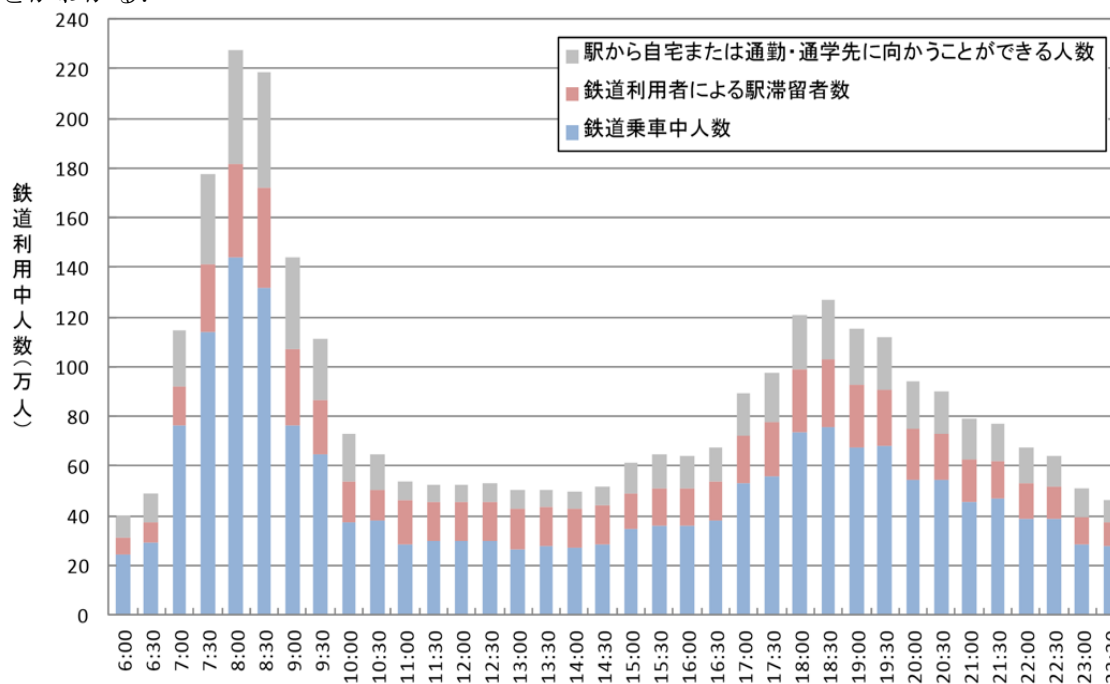


図 5.2.1.2 大都市交通センサスによる鉄道利用者の時刻分布

5.2.2 駅別乗降客数

駅別の乗降客数については推定結果のトリップデータから乗車, 乗換, 降車を行った駅ノードにおけるトリップのみを抽出し, 駅グループコードを用いて駅ごとに集計を行うことで算出した. また, 駅別乗降客数の集計は平日 1 日 (7 月 24 日水曜日), 休日 1 日 (7 月 28 日日曜日) の 2 日間について行った. まず, 本研究の手法を用いて算出した推計値について基本統計量を表 5.2.2.1, 表 5.2.2.2 に示す

表 5.2.2.1 平日の駅別乗降客数についての基本統計量 (有効数字 2 桁)

平均	275.79
標準誤差	25.209
中央値 (メジアン)	46
最頻値 (モード)	0
標準偏差	963.24
分散	927823.90
尖度	153.84
歪度	10.18
範囲	20075
最小	0
最大	20075
データの個数	1460

表 5.2.2.2 休日の駅別乗降客数についての基本統計量（有効数字 2 桁）

平均	147.32
標準誤差	13.74
中央値（メジアン）	26
最頻値（モード）	0
標準偏差	524.92
分散	275544.80
尖度	198.36
歪度	11.59
範囲	11759
最小	0
最大	11759
データの個数	1460

表 5.2.2.1 及び表 5.2.2.2 より，どちらも尖度及び歪度の値が大きいことから平日及び休日の駅別乗降客数の集計結果は非常に偏り，特に小さい値に偏っていることがわかる．

次に，どのような偏りか確認をするためにヒストグラムを作成した．図 5.2.2.3，5.2.2.4 に平日及び休日の駅別乗降客数のヒストグラムを示す．

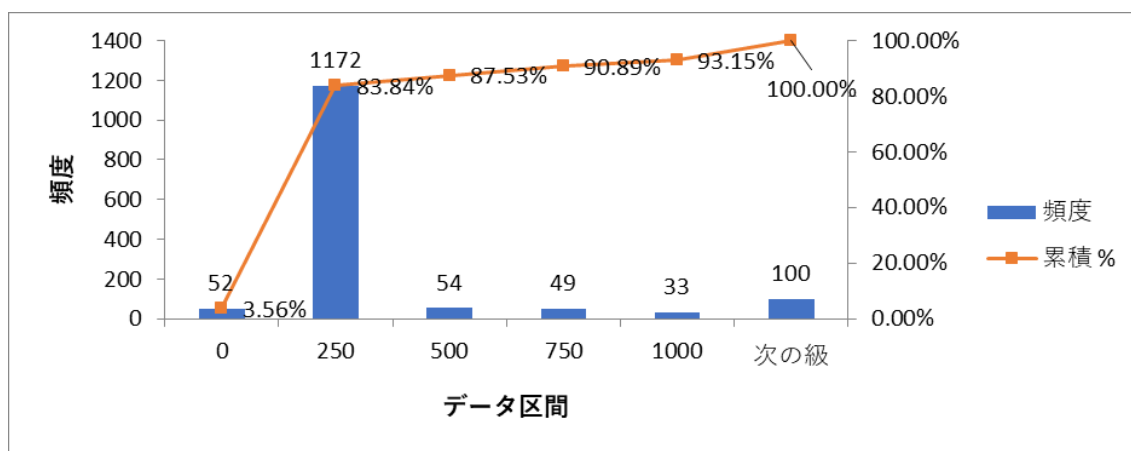


図 5.2.2.3 平日の駅別乗降客数についてのヒストグラム

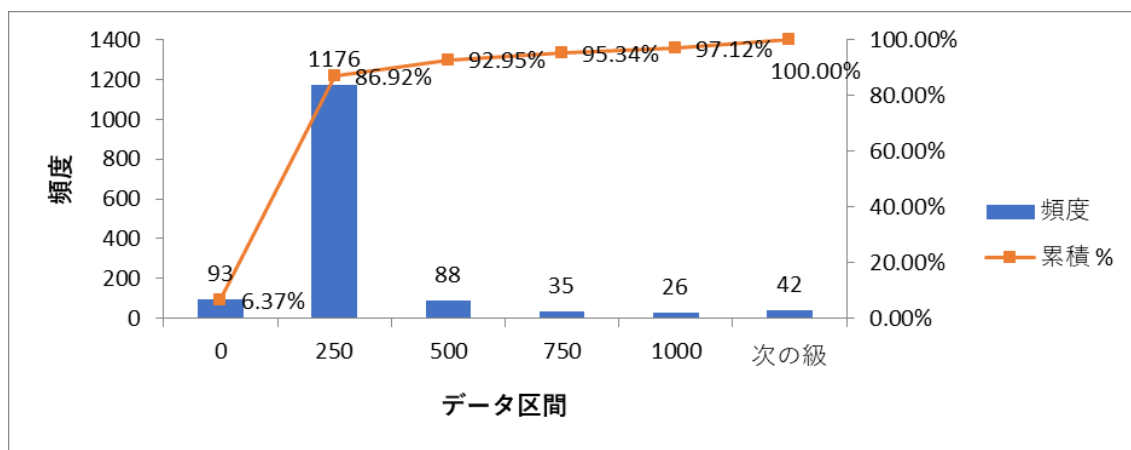


図 5.2.2.4 休日の駅別乗降客数についてのヒストグラム

図 5.2.2.3 及び図 5.2.2.4 より、非常に多くの駅の乗降客数は 0 人以上 250 人以下であることがわかった。

次に平日における駅別乗降客数の集計結果を地図上に表示したものを図 5.2.2.5 に示す。

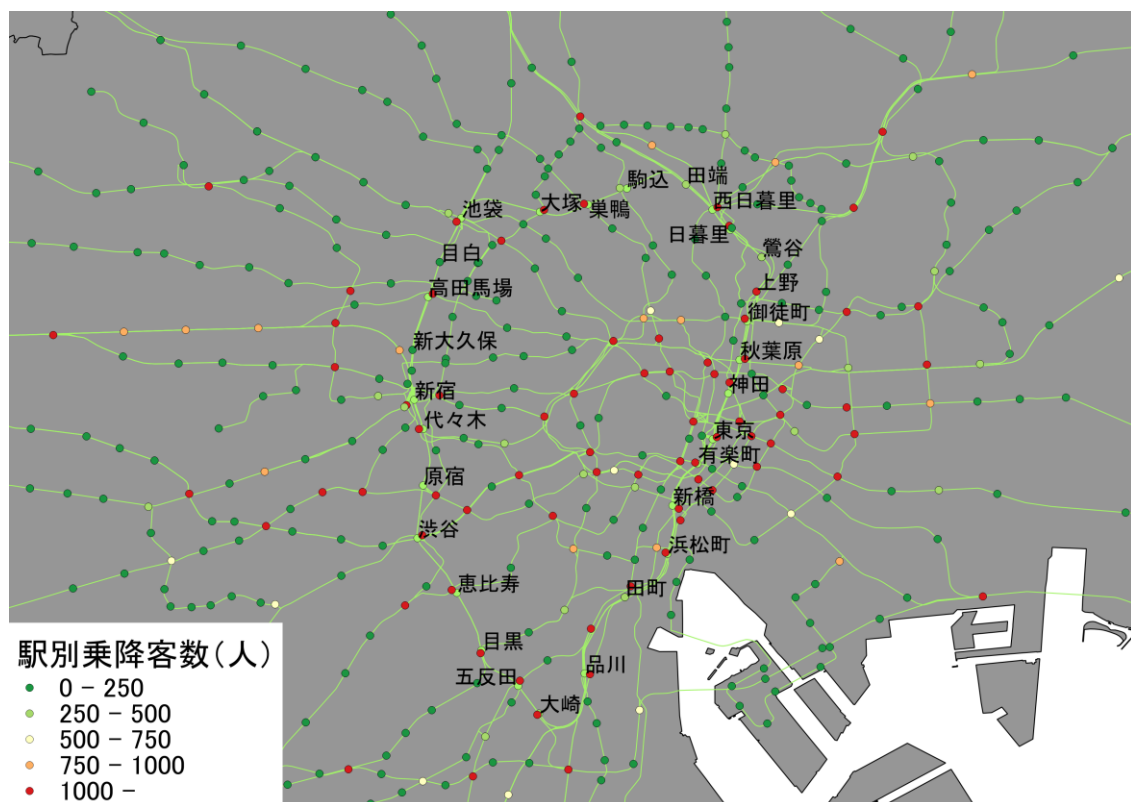


図 5.2.2.5 平日の駅別乗降客数の分布

図 5.2.2.5 より、乗降客数の非常に多い駅（乗降客数 1000 人以上、上位約 5%）は東京都内の中心に固まっていることがわかる。これは東京では中心に巨大なビジネス街や歓楽街が位置しているために多くのユーザーが東京の中心に向けて移動をしていることが原因であると考えられる。

次に休日における駅別乗降客数の集計結果を地図上に表示したものを図 5.2.2.6 に示す。

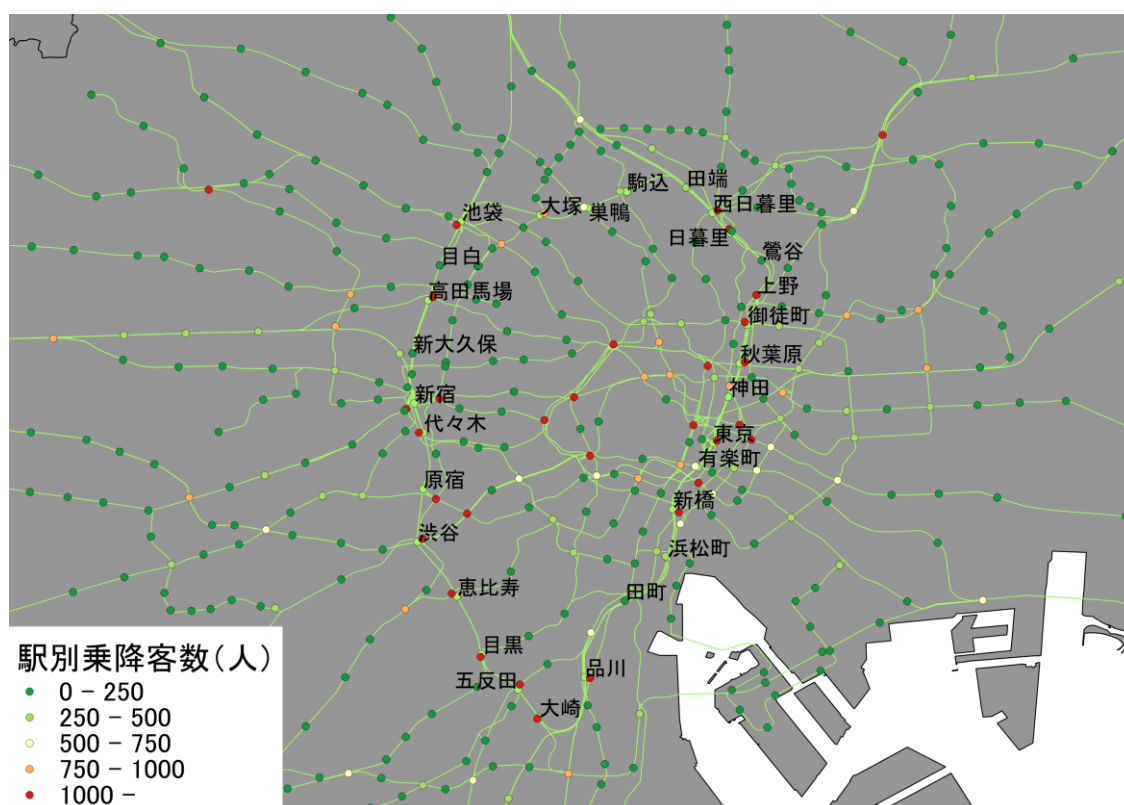


図 5.2.2.6 休日の駅別乗降客数の分布

図 5.2.2.6 より、平日に比べ乗降人数が非常に多い駅は少ないが、乗降客数が非常に多い駅は東京の中心に固まっていることがわかる。これは、休日は平日に比べ通勤、通学を行うユーザーが少ないため全体的に利用者は減るが、遊びなどを目的にして東京の中心地に移動するユーザーが多いためであると考えられる。

次に本研究における平日の推計値と国土数値情報ダウンロードサービスが提供する駅別乗降客数のデータとの比較を行う。国土数値情報の駅別乗降客数データは 1 日の駅ごとの乗降客数の集計値であり、今回は混雑統計の年度と合わせて 2013 年のデータを比較対象に選んだ。ただし、国土数値情報の駅別乗降客数のデータには駅ごとにコードが振り分けられておらず、乗降客数は各駅かつ路線ごとに記録されているため、駅名で集計を行うことにより比較用のデータとした。まず、比較対象の国土数値情報駅別乗降客数データの基本統計量とヒストグラムを表 5.2.2.7 及び図 5.2.2.8 に示す。

表 5.2.2.7 国土数値情報の駅別乗降客数データについての基本統計量（有効数字 2 桁）

平均	52461.47
標準誤差	4207.32
中央値（メジアン）	19621.50
最頻値（モード）	42453
標準偏差	160761.35
分散	25844211047
尖度	208.14
歪度	12.73
範囲	3295543
最小	4
最大	3295547
データの個数	1460

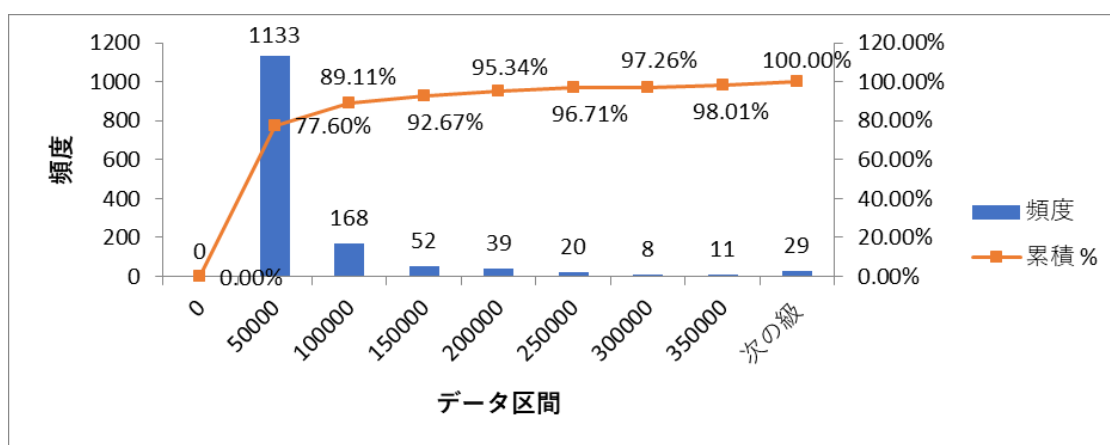


図 5.2.2.8 国土数値情報の駅別乗降客数データについての基本統計量

表 5.2.2.7 より、尖度及び歪度の値が高いため国土数値情報の駅別乗降客数データは推計値と同様に偏っており、乗降客数の少ない駅が多いことがわかる。また、図 5.2.2.8 より、ほとんどの駅の乗降客数は 50,000 人以下であることがわかる。

そして、本研究における乗降客数の推計値と国土数値情報の駅別乗降客数データを散布図にプロットしたものを図 5.2.2.9 に示す。ただし、国土数値情報の駅別乗降客数データでは乗降客数を公開していない駅の乗降客数は「0」となっているため、比較する際には国土数値情報の駅別乗降客数データにおいて乗降客数が 0 人である駅を除いて比較を行った。このデータは全国で 8,617 駅あり、そのうち乗降客数が 0 人であった駅は 1,958 駅であった。

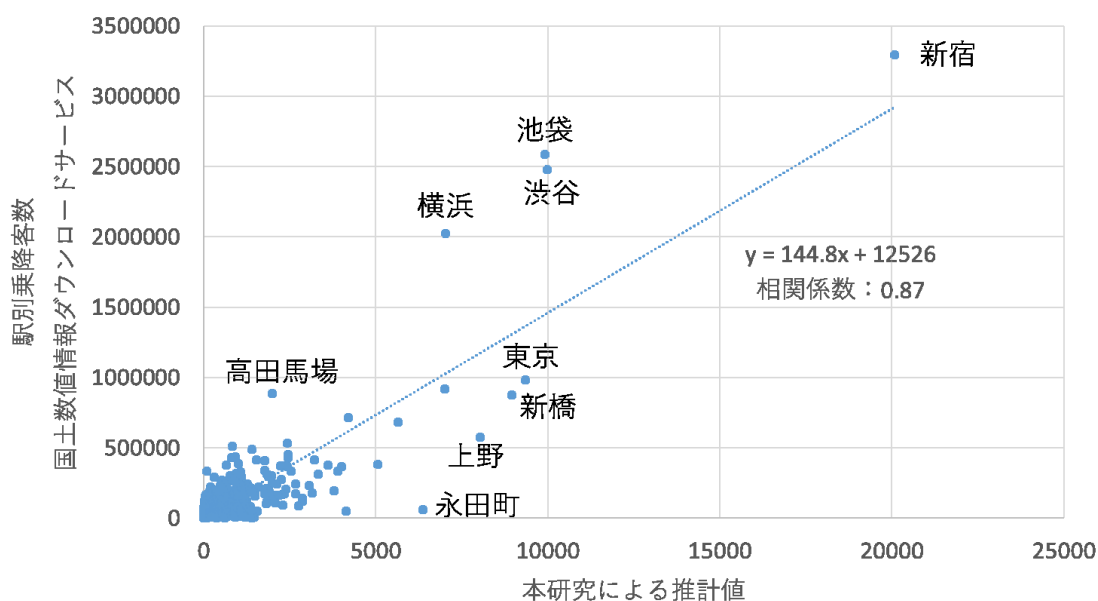


図 5.2.2.9 本研究による推計値と国土数値情報による駅別乗降客数データの比較結果 1

図 5.2.2.9 より、相関係数の値が非常に高いことから 2 つのデータの間には非常に強い相関があることがわかる。また、新宿・池袋・渋谷・横浜の 4 駅が集団から特に離れているが、これらの駅は世界の駅の中で最も乗降客数の多い駅であり、ビジネス及び歓楽の中心のものであることが原因であると考えられる。しかし、多くのデータが下に固まっているため、国土数値情報の駅別乗降客数が 500,000 人以下を抽出し、再度分析を行う。再度分析を行った結果を図 5.2.2.10 に示す。

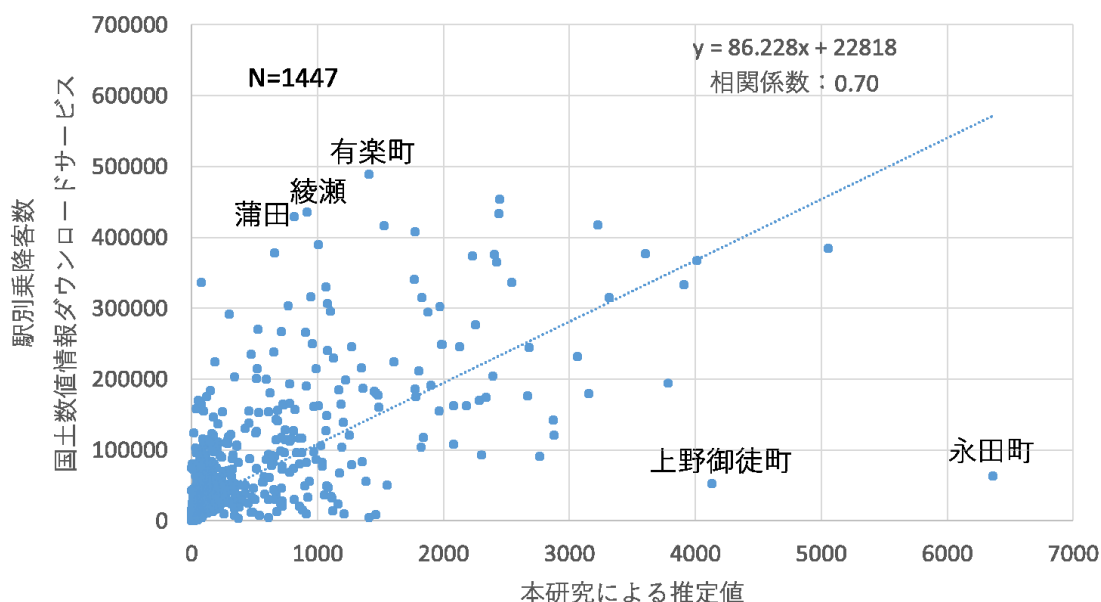


図 5.2.2.10 本研究による推計値と国土数値情報による駅別乗降客数データの比較結果 2

図 5.2.2.10 より、相関係数は高い方であることがわかる。また、永田町駅が特に離れているが、これは永田町駅付近には多くの駅が徒歩圏内に位置していることから降車駅のマッ

チング先が他の駅に決まってしまう可能性が考えられる。

さらに、国土数値情報の駅別乗降客数が 10 万人以下である駅が非常に多いことから国土数値情報の駅別乗降客数が 10 万人以下の駅をさらに抽出して再度分析を行う。再度分析を行った結果を図 5.2.2.11 に示す。

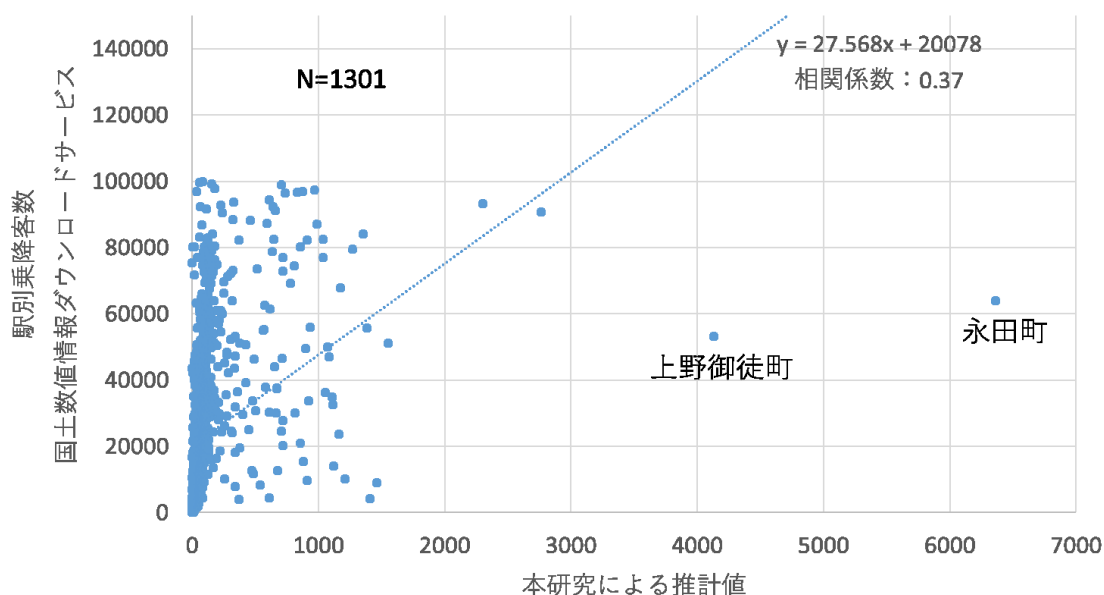


図 5.2.2.11 本研究による推計値と国土数値情報による駅別乗降客数データの比較結果 3

図 5.2.2.11 より、相関係数が低いためこれらの駅については相関が非常に弱いことがわかる。また、本研究による推計値が少ない駅が非常に多いことがわかるが、これらの多くは推計値が 150 人以下の駅である。

また、特に上野御徒町駅と永田町駅について推計過大になってしまっている。これは、今回の鉄道ネットワークデータ作成のために駅の接続情報として用いた駅データ.jp の駅データにおいて本来は違う駅であるにも関わらず、乗り換え可能とされていれば同一の駅グループコードが割り振られていることが原因であった永田町駅においては赤坂見附駅と乗り換え可能となっており、上野御徒町駅においては御徒町駅と乗り換え可能となっている。そして、これはどれも非常に乗降客数の多い駅であるため、これらの駅同士の乗降客数が 1 つに集計されてしまったためにこのような現象が起きたと考えられる。したがって、このような現象を回避するためには集計を行う際には駅グループコードが同じであっても駅名が異なる場合には集計を分ける必要がある。もしくは、東京駅と大手町駅といった駅名が別であり、一般的には乗り換え可能とされているが駅グループコードが異なっている駅も今回作成した鉄道ネットワークには含まれているため、駅グループコードの振り分けを見直すことも必要であると考えている。

次に、本研究による推計値が特に少ない駅の特徴を見るために推計値が 150 人以下の駅を抽出し、地図上にプロットした。プロットした結果を図 5.2.2.12 に示す。

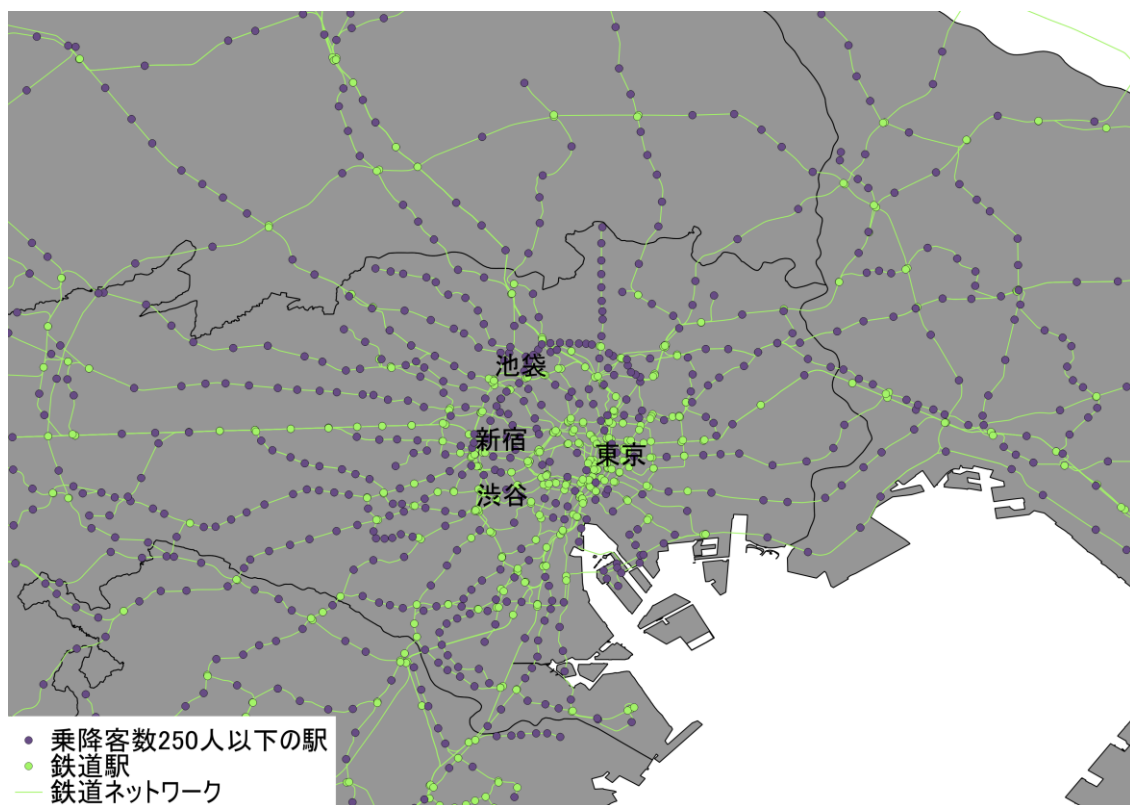


図 5.2.2.12 乗降客数の推計値が 150 人以下の駅の分布

図 5.2.2.12 より、乗降客数の推計値が 250 人以下の駅は広い範囲に分布していることがわかるが、東京の中心に向かうにつれて乗降客数の推計値が 250 人以下の駅は少ないことがわかる。また、ネットワークが交差している駅（乗換可能駅）も乗降客数が 250 人以下であることは少ない。これらの特徴から、ターミナルとなっている駅・ビジネス街に位置する駅、歓楽街に位置する駅ではなく、ベッドタウンとなっているような駅の乗降客数が少なくなっていると考えられる。ベッドタウンとなっている駅の乗降客数が少ないことから今回用いた GPS データのユーザー数では自宅の位置が分散してしまい、各ベッドタウンに位置する駅の乗降客数が非常に少なくなってしまった可能性が考えられる。

5.2.3 駅間移動人員数

駅間移動人員数については推定結果のトリップデータを各駅間（リンク間）で集計を行うことで算出した。また、駅間移動人員数の集計も駅別乗降客数の集計と同様に平日 1 日（7 月 24 日水曜日）、休日 1 日（7 月 28 日日曜日）の 2 日間について行った。さらに、集計は 1 時間ごと及び 1 日の合計の 2 種類で行った。まず、本研究の手法を用いて算出した推計値における 1 日の駅間移動人員数の合計について基本統計量を表 5.2.3.1、5.2.3.2 に示す。

表 5.2.3.1 平日の駅間移動人員数についての基本統計量（有効数字 2 桁）

平均	469.85
標準誤差	10.00
中央値（メジアン）	330
最頻値（モード）	20
標準偏差	418.85
分散	175431.90
尖度	0.77
歪度	1.06
範囲	2429
最小	1
最大	2430
合計	825053
データの個数	1756

表 5.2.3.2 休日の駅別乗降客数についての基本統計量（有効数字 2 桁）

平均	252.81
標準誤差	5.11
中央値（メジアン）	197
最頻値（モード）	11
標準偏差	214.35
分散	45946.50
尖度	1.032
歪度	1.06
範囲	1497
最小	1
最大	1498
合計	444187
データの個数	1757

表 5.2.3.1 及び表 5.2.3.2 より、どちらも尖度及び歪度の値が小さいことからデータの偏りは少なく、若干小さい値に偏っていることがわかる。ただし、尖度の値を比較すると平日のデータの偏りは休日のデータに比べて少ない。また、合計値を比較すると平日は休日比べ、移動人員数が約半分であることがわかる。

次に、どのような偏りか確認をするためにヒストグラムを作成した。図 5.2.3.3、図 5.2.3.4 に平日及び休日の駅別乗降客数のヒストグラムを示す。

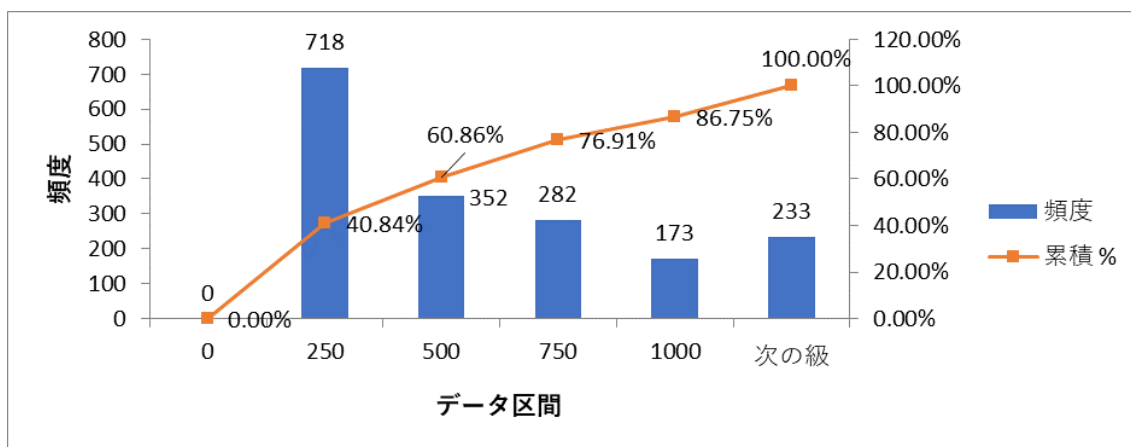


図 5.2.3.3 平日の駅間移動人員数についてのヒストグラム

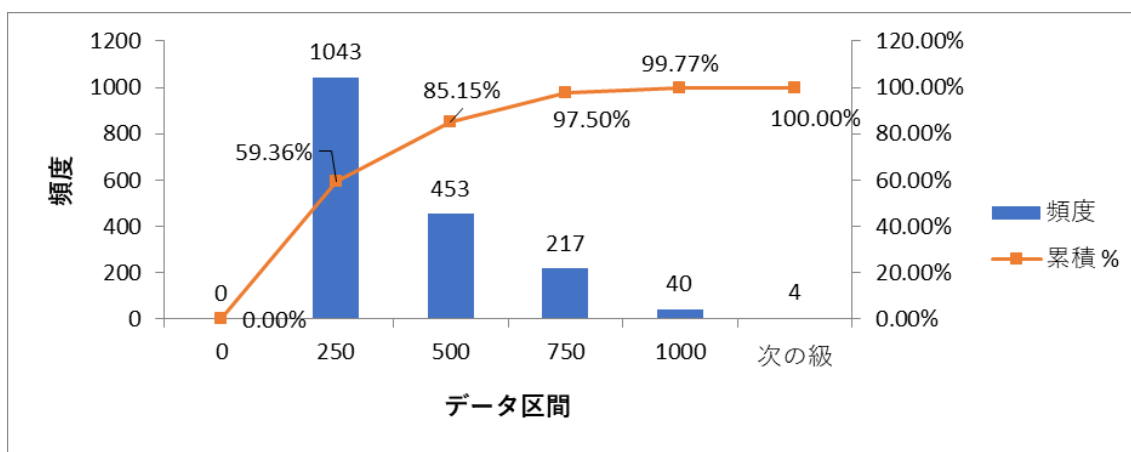


図 5.2.3.4 休日の駅間移動人員数についてのヒストグラム

図 5.2.3.3 及び図 5.2.3.4 より、非常に多くの駅の乗降客数は 0 人以上 250 人以下であることがわかった。しかし、平日の駅間移動人員数は休日の駅間移動人員数に比べて人員数にばらつきがあることがわかる。

次に平日における 1 日の駅間移動人数の合計を地図上に表示したものを図 5.2.3.5 に示す。

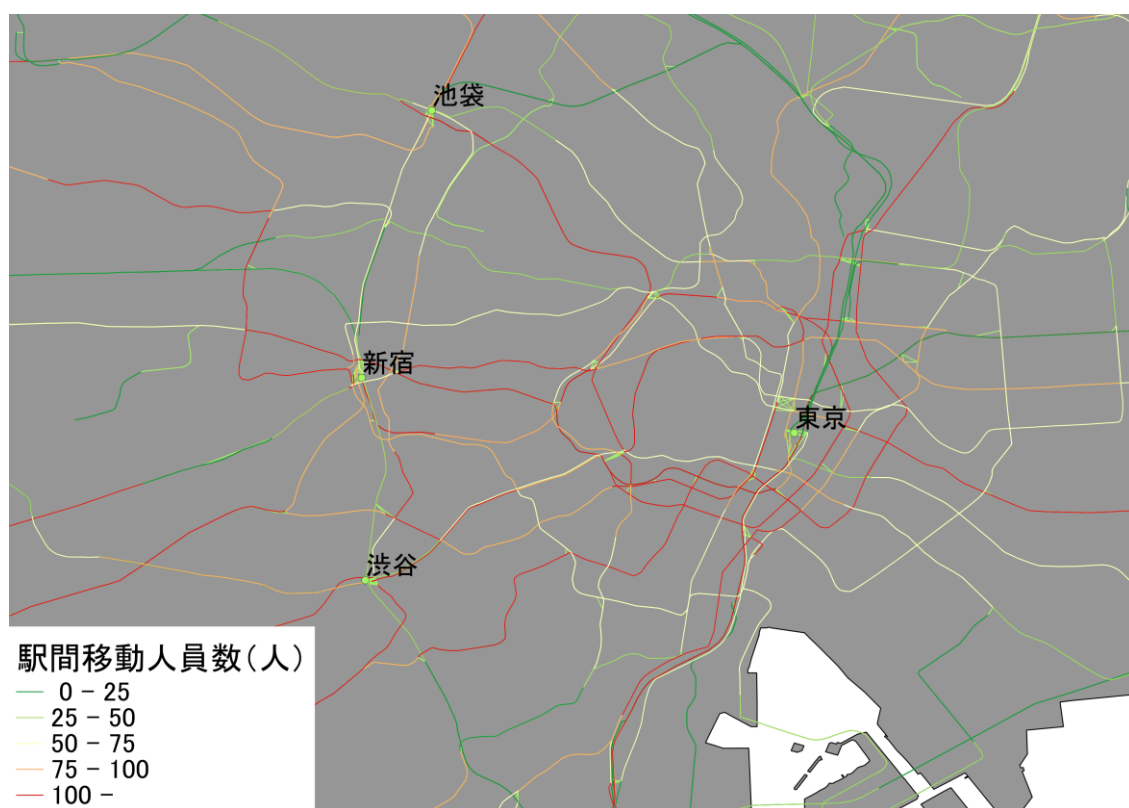


図 5.2.3.5 平日 1 日の駅間移動人員数の分布

図 5.2.3.5 より，全体的に移動人員数が非常に多いことがわかり，特に東京駅に近い東側の路線と渋谷駅及び新宿駅に接続している路線の移動人員数が非常に多いことがわかる．

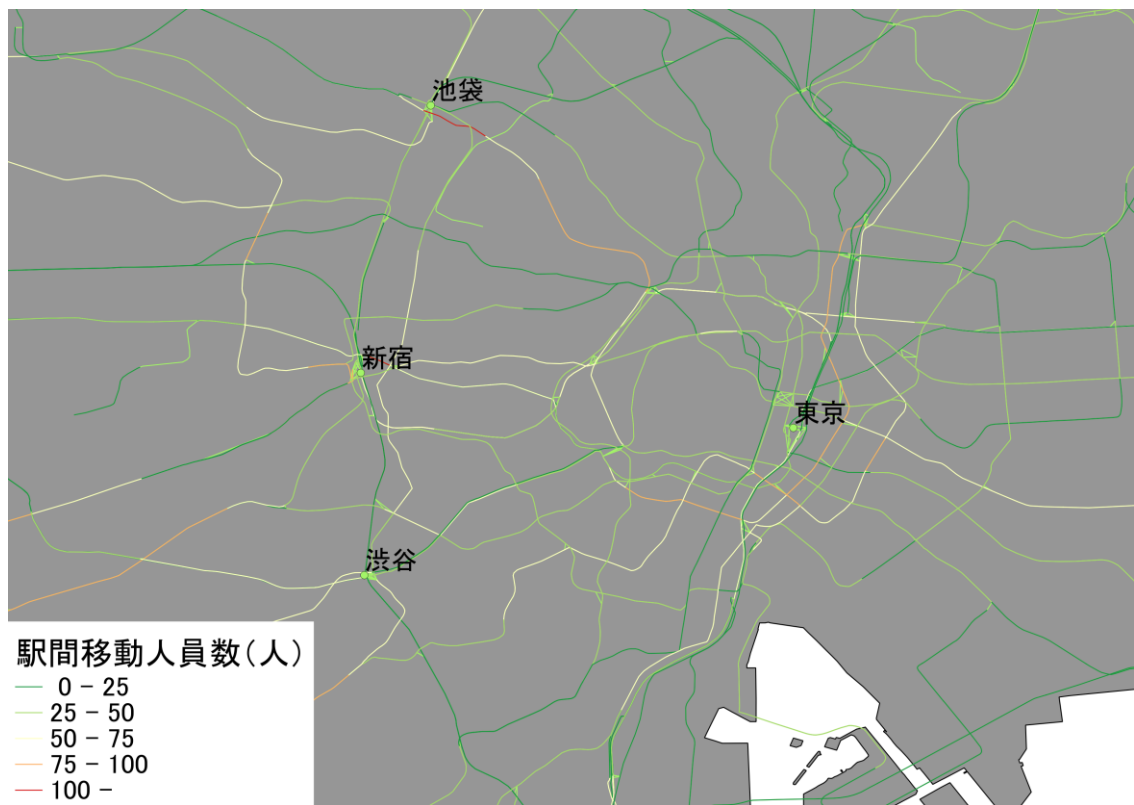


図 5.2.3.6 休日 1 日の駅間移動人員数の分布

図 5.2.3.6 より，平日に比べ乗降人数が非常に多い路線・区間は少なく，全体的に移動人員数の差が小さい。

次に平日 1 時間ごとの駅間移動人員数の変化を見る。駅間移動人員数を 7 時, 8 時, 9 時, 12 時, 18 時, 19 時, 20 時について地図上に表示したものを図 5.2.3.7~5.2.3.13 に示す。



図 5.2.3.7 平日 7 時の駅間移動人員数の分布

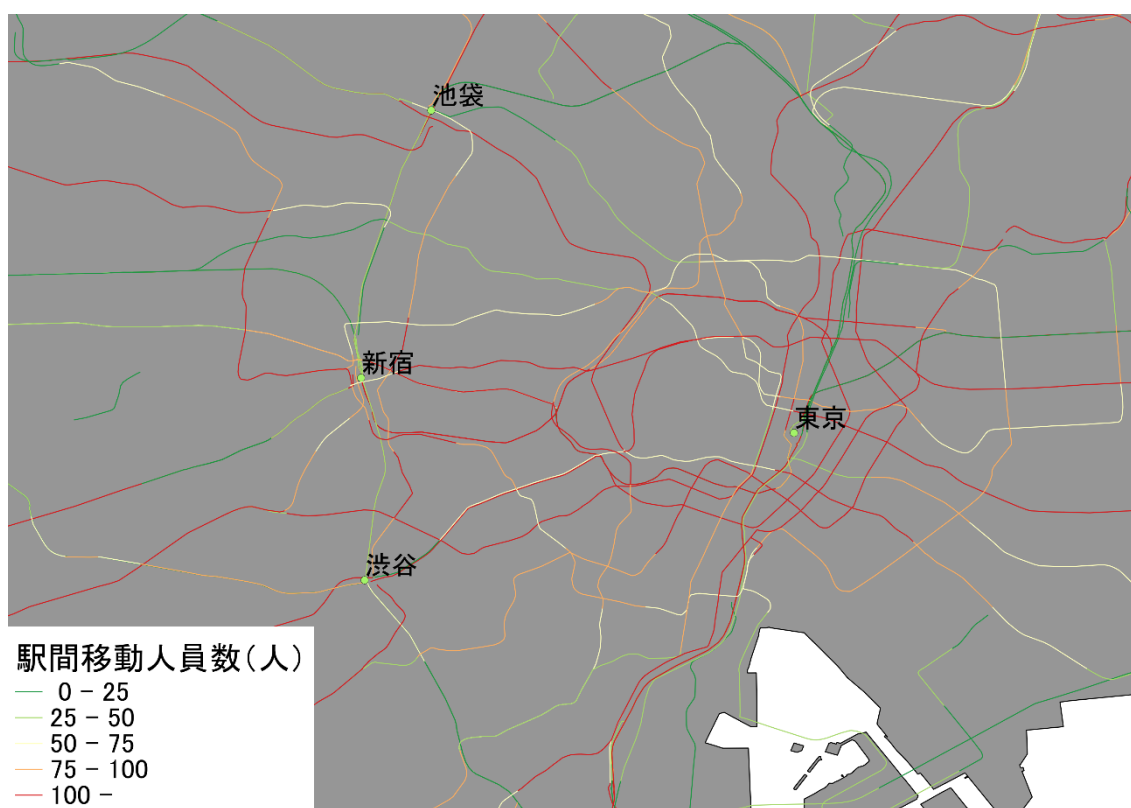


図 5.2.3.8 平日 8 時の駅間移動人員数の分布

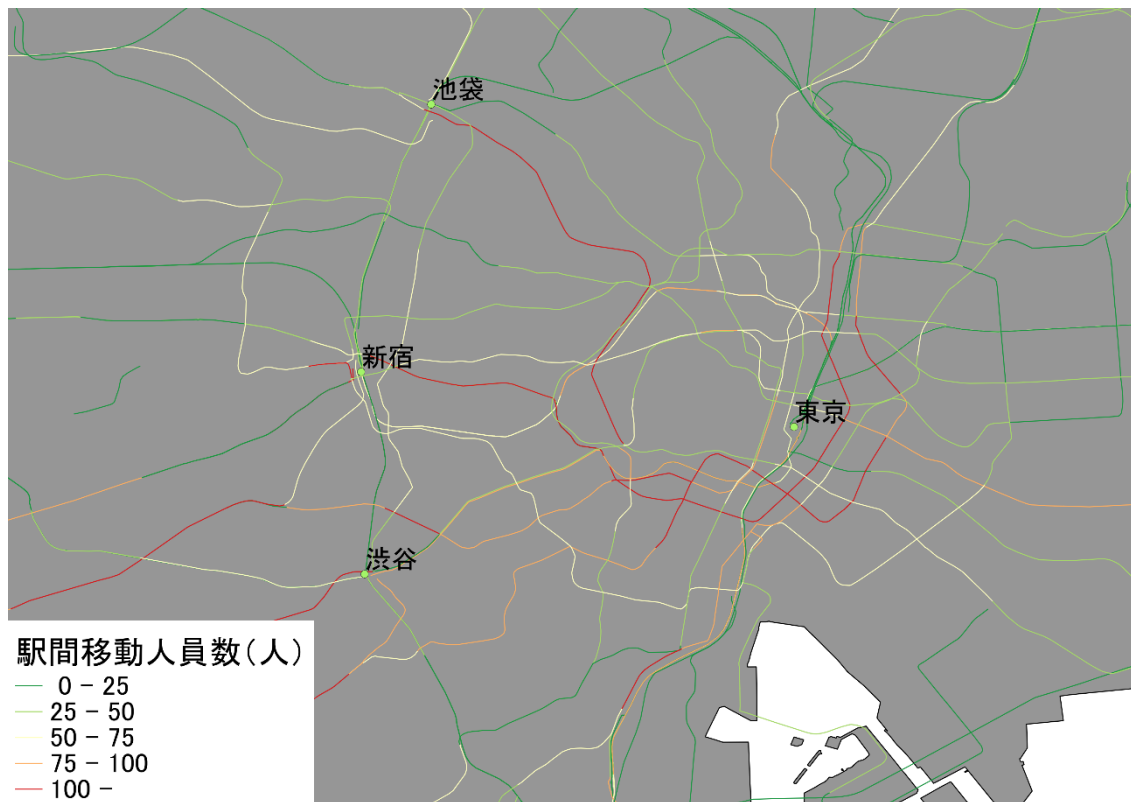


図 5.2.3.9 平日 9 時の駅間移動人員数の分布

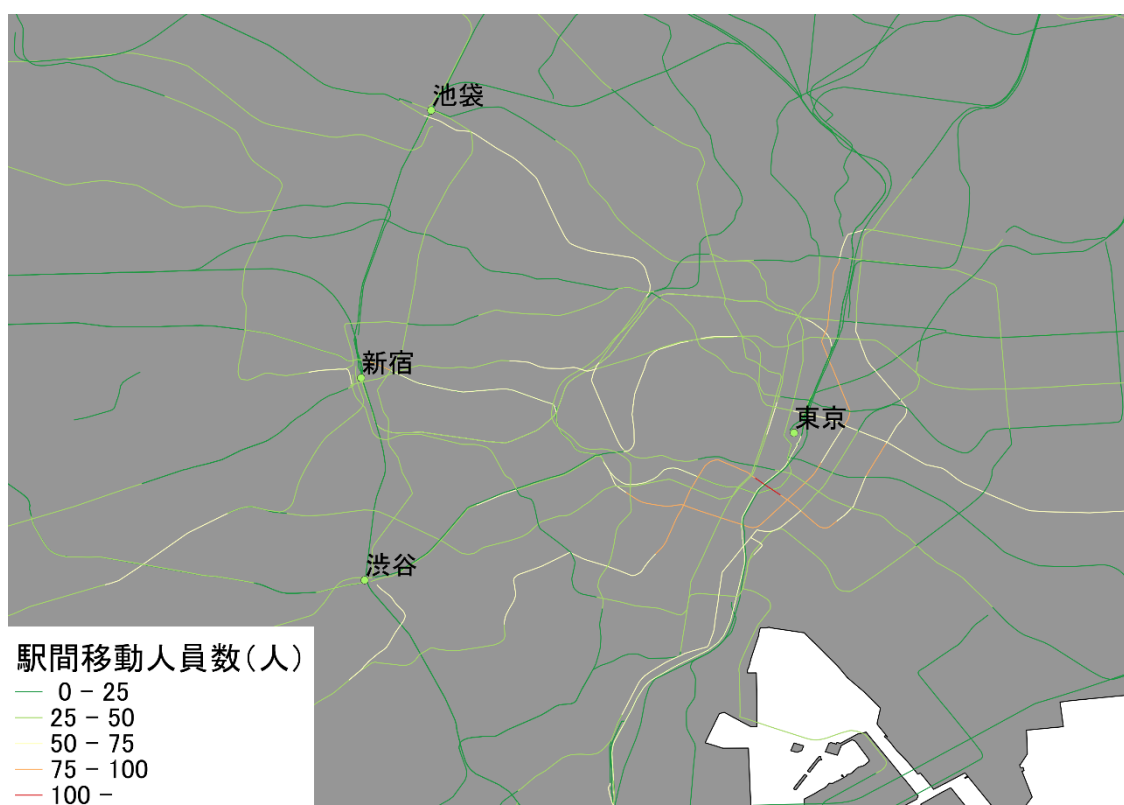


図 5.2.3.10 平日 12 時の駅間移動人員数の分布

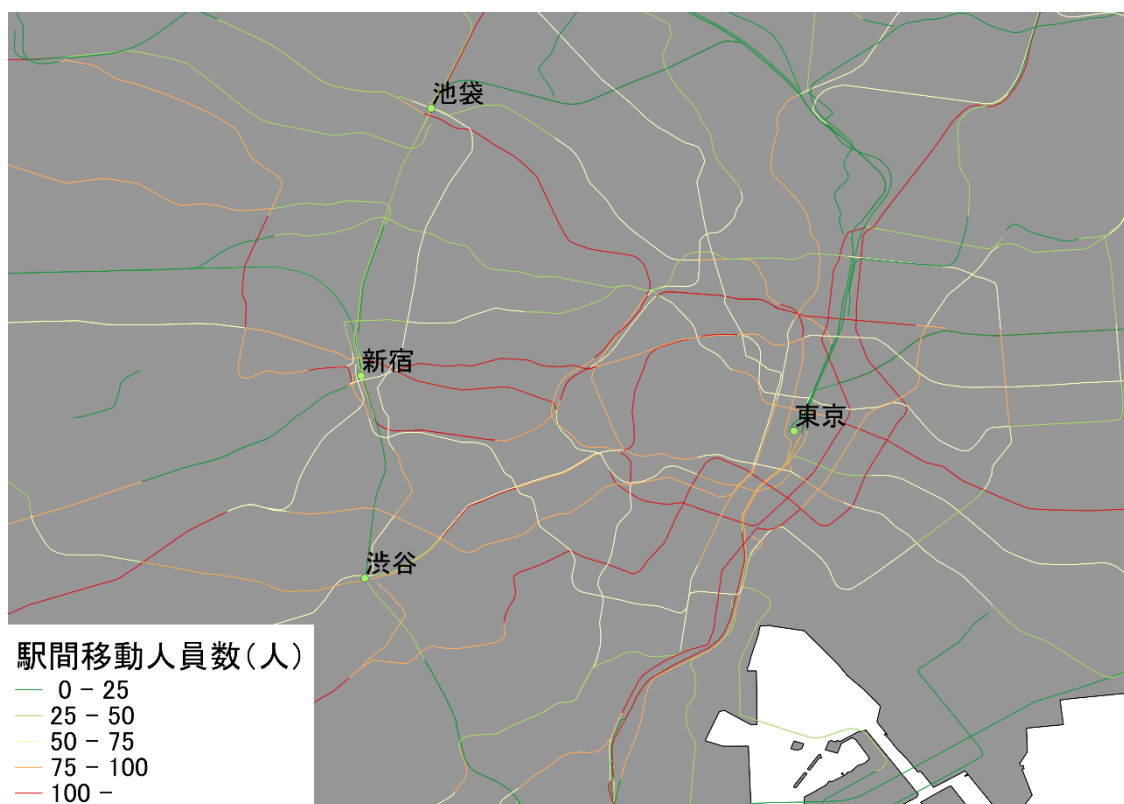


図 5.2.3.11 平日 18 時の駅間移動人員数の分布

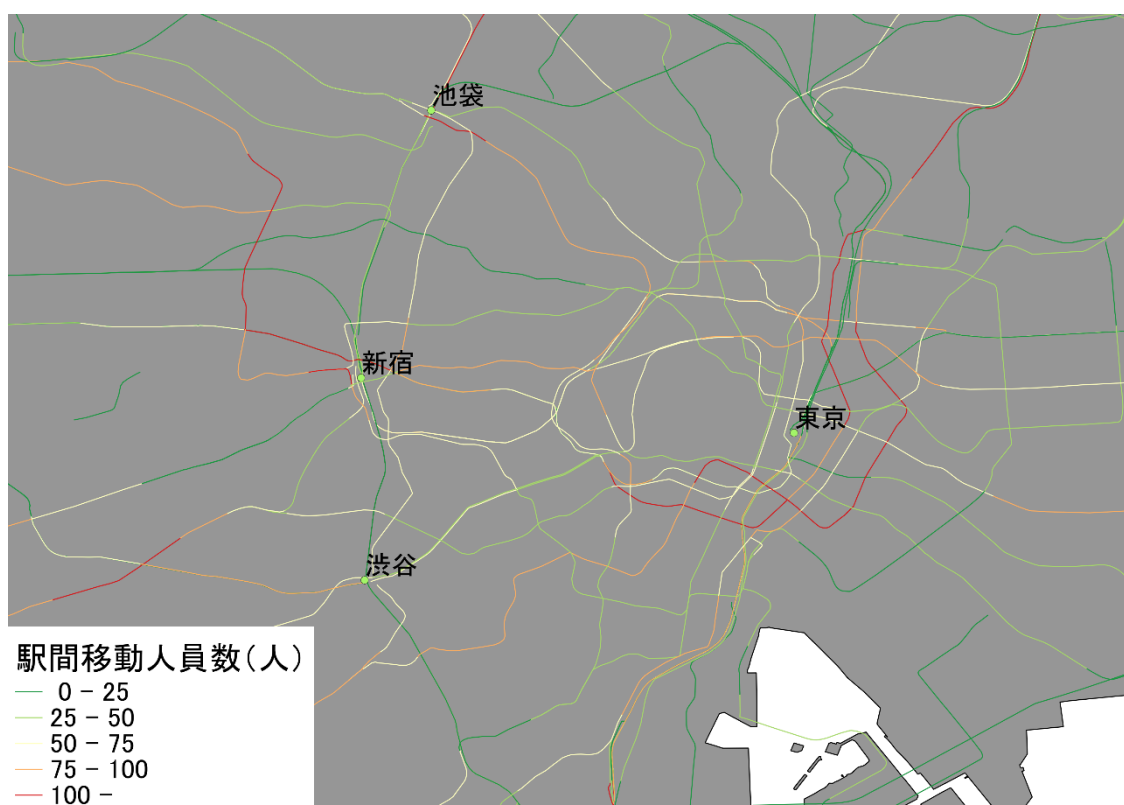


図 5.2.3.12 平日 19 時の駅間移動人員数の分布

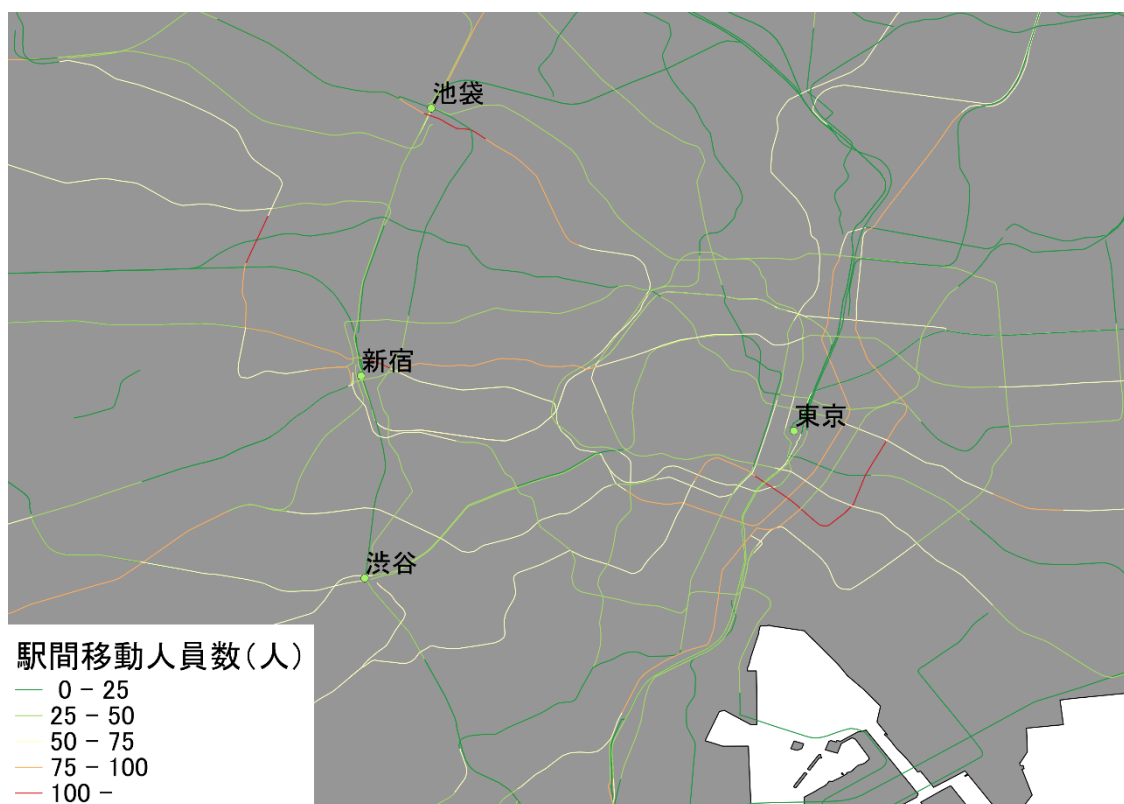


図 5.2.3.13 平日 20 時の駅間移動人員数の分布

図 5.2.3.7～5.2.3.9 より、7時から8時へかけて移動人員数が全体的に増えていき、9時には減り始めていることがわかる。これは、多くの企業や学校の始業時間が9時付近であり、8時台の移動が多いと考えると妥当な推移であると考えられる。図 5.2.3.10 より、12時台の移動人員数は7時や8時に比べ、全体的に少ないことがわかる。これは、平日の12時台は多くのユーザーが就業中であるため、12時台の移動人員数の分布も妥当であると考えられる。さらに、図 5.2.3.11～5.2.3.13 より、18時台の移動人員数が多く、20時に向けて徐々に減り始めていることがわかり、18時台の移動人員数は8時の移動人員数と比べ、少ないことがわかる。これは朝の通勤時間・通学時間は多くの企業・学校において同一であるが、帰宅の時間は企業・学校によってばらつきがあり直帰するとは限らないからであると考えられる。

次に休日1時間ごとの駅間移動人員数の変化を見る。駅間移動人員数を8時、12時、18時について地図上に表示したものを図 5.2.3.14～5.2.3.16 に示す。

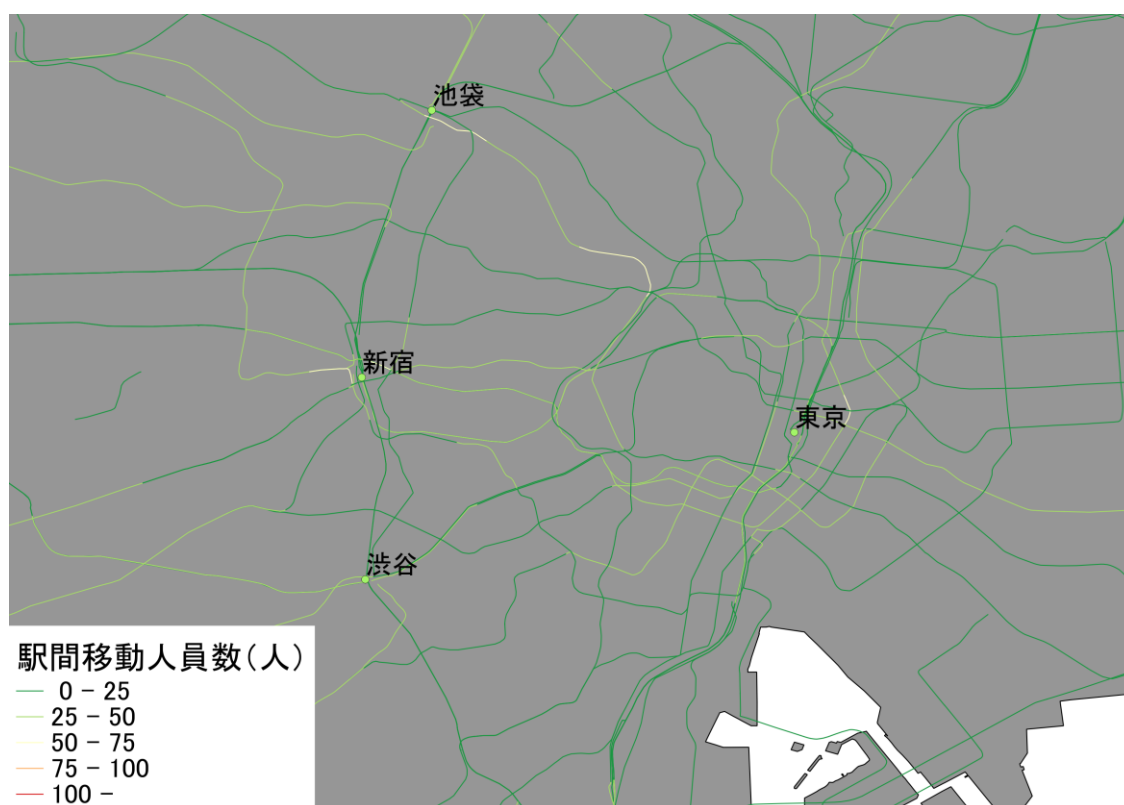


図 5.2.3.14 休日 8時の駅間移動人員数の分布

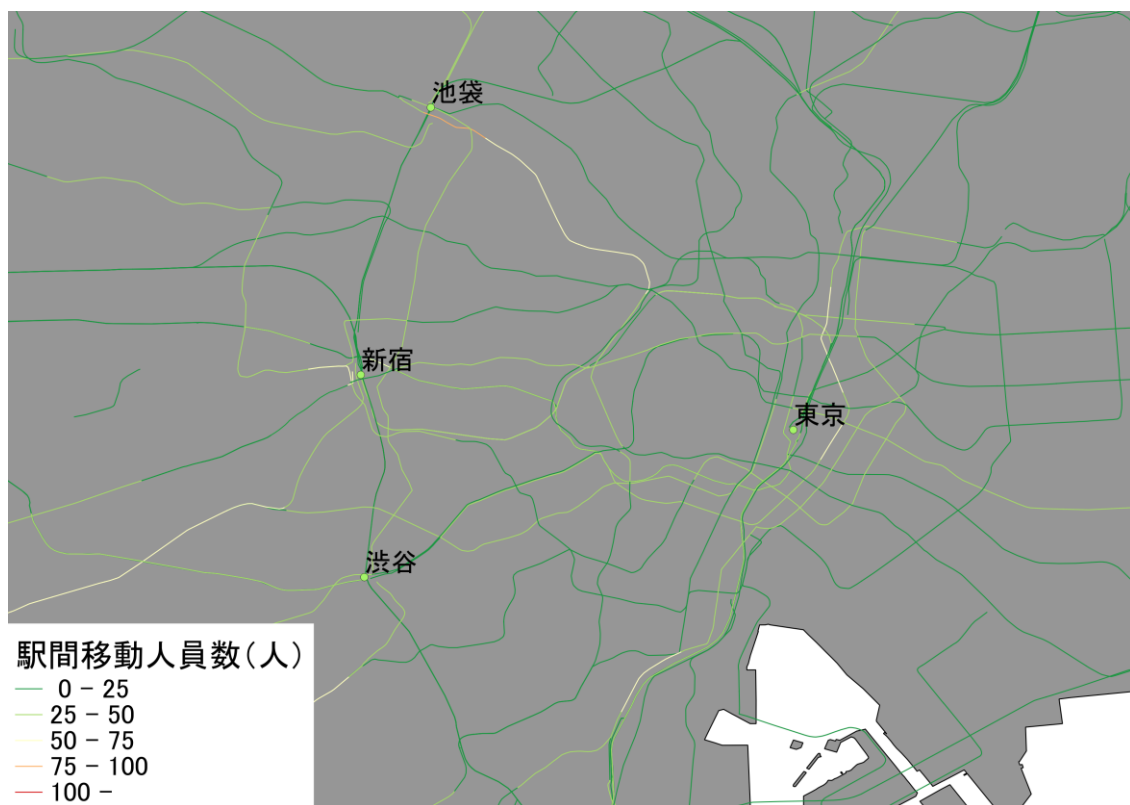


図 5.2.3.15 休日 12 時の駅間移動人員数の分布

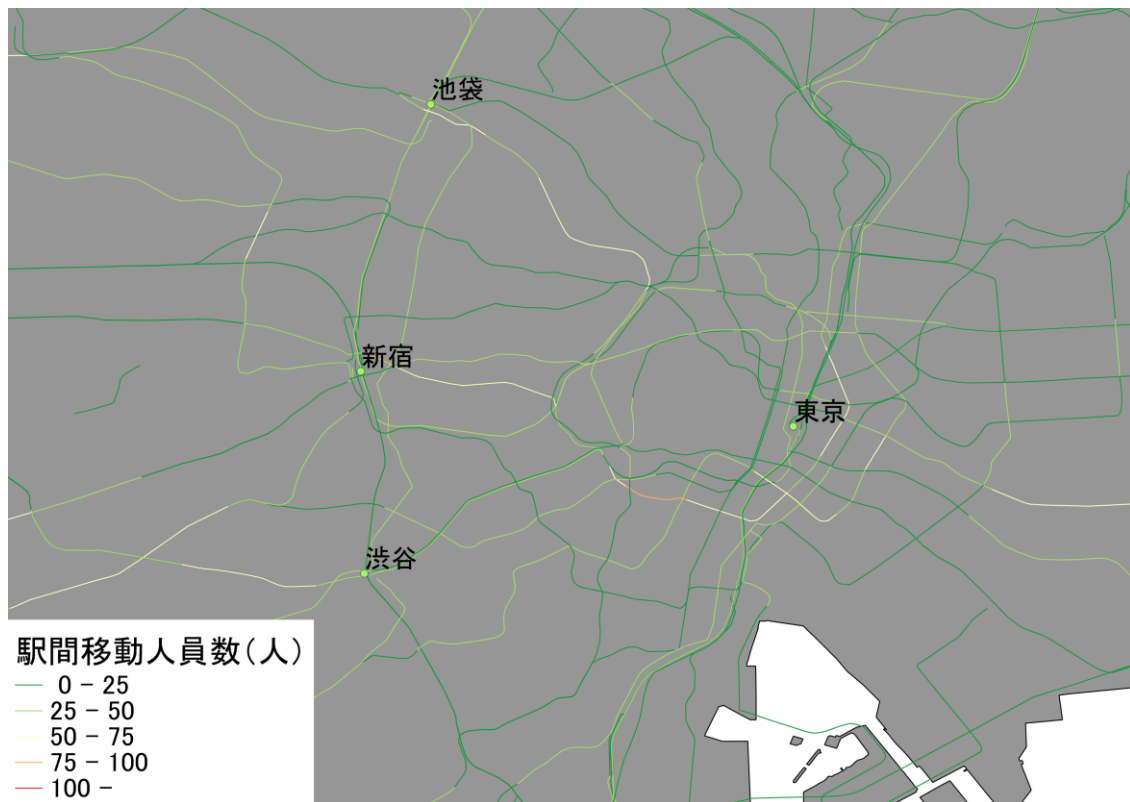


図 5.2.3.16 休日 18 時の駅間移動人員数の分布

図 5.2.3.14～5.2.3.16 より，やはり平日に比べ全体的に駅間移動人員数が少ないことがわかる．また，移動人員数は 8 時が最も少ないこともわかる．しかし，駅間移動人員数が少ない中でも池袋駅・新宿駅・渋谷駅・東京駅といったターミナル駅に接続している路線は休日の中でも移動人員数が多いことがわかった．

次に駅間移動人員数の推計値と大都市交通センサスにおける線別駅間移動人員数[17]の比較を行う．大都市交通センサスの線別駅間移動人員数データは平日 1 日の駅間ごとの移動人員数の集計値であり，今回は平成 22 年度のデータを比較対象に選んだ．また，大都市交通センサスの線別駅間移動人員数データは乗車駅から降車駅，乗車駅から乗換駅，乗換駅から乗換駅，乗換駅から降車駅のペアごとに移動人員数が記録されているため，各駅間の移動人員とはなっていない．そのため，本研究における鉄道ネットワークの作成の際に用いた駅データ.jp の駅データの位相情報のみを利用して経路探索を行い，各駅間の移動人員のデータに変換した．まず，比較対象の大都市交通センサスの駅間移動人員数の基本統計量とヒストグラムを表 5.2.3.17 及び図 5.2.3.18 に示す

表 5.2.3.17 大都市交通センサスの駅間移動人員数データにおける基本統計量
(有効数字 2 桁)

平均	56364.13
標準誤差	1466.16
中央値 (メジアン)	36128.50
最頻値 (モード)	3507
標準偏差	61473.78
分散	3779025646
尖度	2.68
歪度	1.63
範囲	397003
最小	4
最大	397007
合計	99088143
データの個数	1758

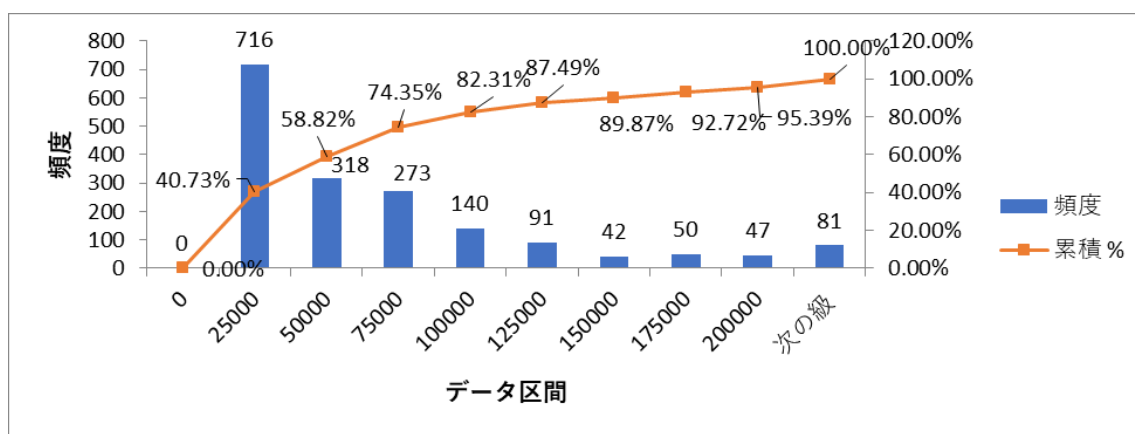


図 5.2.3.18 大都市交通センサスの駅間移動人員数データにおけるヒストグラム

表 5.2.3.17 及び図 5.2.3.18 より、尖度及び歪度の値が 0 を超えていることから偏りがあり、小さい値に偏っていることがわかる。また、その偏りは本研究による推計値よりも大きいものである。さらに駅間移動人員数が 25000 人以下の駅間が非常に多いこともわかる。

そして、本研究における駅間移動人員数の推計値と大都市交通センサスの駅間移動人員数データを散布図にプロットしたものを図 5.2.3.19 に示す。

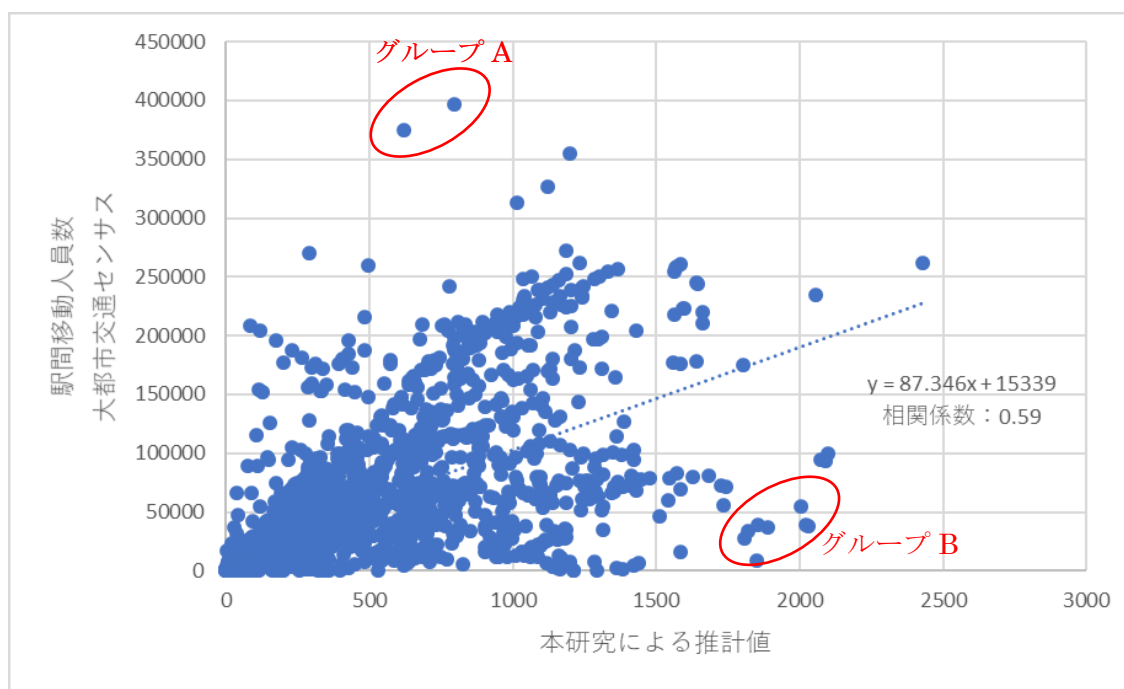


図 5.2.3.19 本研究による推計値と大都市交通センサスによる駅間移動人員数の比較

図 5.2.3.19 より、点の分布は非常にばらつきがあることがわかる。さらに、相関係数の値からあまり相関が強くないことがわかった。しかし、本研究における手法及び鉄道ネットワークデータにはまだ改善点が多く存在しているため、相関が必ずしも高くない可能性はある。特に図 5.2.3.19 において特に推計過小(グループ A)となってしまう区間は JR 東海道本線 品川駅-川崎駅間や中央線快速 新宿駅-四ツ谷駅間といった平行して運行する路線が存在する駅であるため、やはりこのような区間における推定精度はより実態に沿った乗り換えを行えるよう、さらなる検討の必要があるとわかった。

逆に推定過大(グループ B)となってしまう駅については、平行して運行する路線がない路線区間が多いが、推計過大となってしまう区間の多くは日比谷線及び銀座線の区間であり、特に銀座駅・日比谷駅付近における区間であった。しかし、現在まででは確実な原因が特定できなかったため、これらの部分に関して鉄道ネットワークデータの確認・修正、さらに経路選択の手法を再検討する必要があると考えた。また、比較に用いた大都市交通センサスの線別駅間移動人員数データも経路探索を行っているため、経路探索の精度に依存した比較データとなってしまう。さらに、駅別乗降客数データの比較において述べたが、今回用いた GPS データのユーザー数では都心の目的地が他のユーザーと被る可能性が大きいが帰宅する駅が被る可能性は少ないというデータ上の問題点も原因と考えられる。

そこで、全体の比較ではなく、東京の中心部の方であればデータが集まりやすいという特性を考慮して山手線の結果のみを抽出し、比較を行った。以下の図 5.2.3.20 に本研究にお

ける推計値と大都市交通センサスの線別駅間移動人員数データの比較結果を示す。

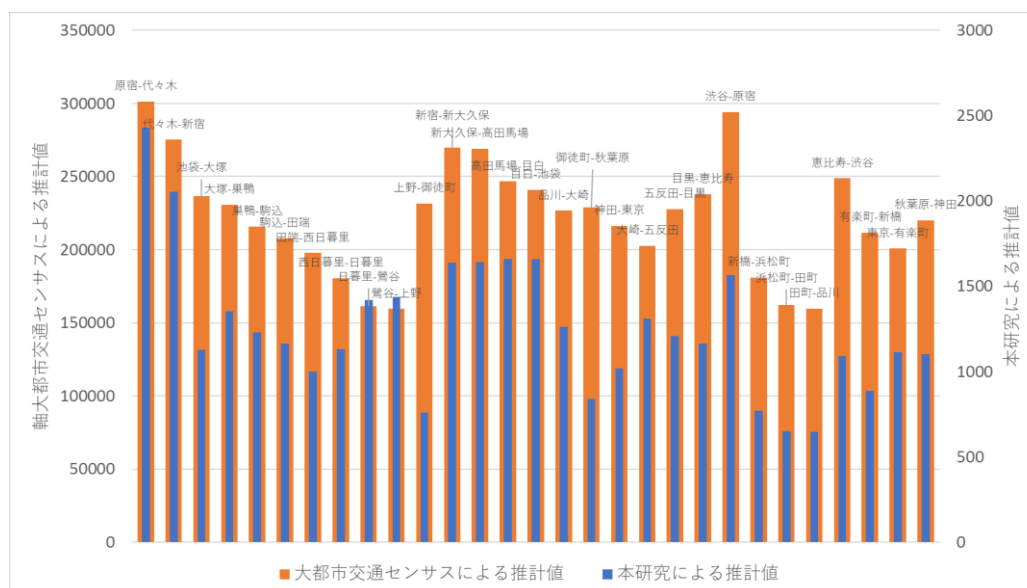


図 5.2.3.20 本研究による推計値と大都市交通センサスの山手線のデータとの比較

図 5.2.3.20 より、本研究の推計値は大都市交通センサスによる推計値よりも値が偏ってしまっていることがわかる。さらに、山手線のみでの比較だと相関係数は 0.67 であるため、全体での比較における相関よりも相関は強いことがわかる。駅別乗降客数の比較においてユーザーが集まりやすい駅では相関が強くなったにも関わらず、ユーザーが集まりやすい駅の多くで運行する山手線において相関は非常に強いとまではいかなかった。これは山手線には平行して運行する路線が存在する区間が多いため、本研究の手法の特性からユーザーが複数の路線に分散してしまったと考えられる。

6 結論

6.1 本研究の成果

本研究において新たに作成した鉄道ネットワークデータを用いた鉄道利用実態把握のために提案した本研究の手法を自身の移動履歴に適用した結果、地上経路のみの移動における本研究の手法の適用結果と地下経路を含む移動における本研究の手法の適用結果よりと平行して運行する路線がない場合にはある程度まで妥当な経路が導かれた。しかし、平行して運行する路線が複数存在する場合には本来ならば必要とはならないような乗り換えをしてしまうといった妥当ではない経路が導かれられない可能性も含んでいるため、本研究における鉄道ネットワークデータ及び利用路線推定手法を利用する際には平行して運行する路線が複数存在している区間に関しては注意が必要である。したがって、本研究の手法を用いた経路探索では以下のことがわかった。

- ・ 平行して運行する路線がない場合にはある程度まで妥当な探索結果が得られる
- ・ 平行して運行する路線がある場合には妥当な探索結果が得られない場合もある
- ・ 地下区間であっても降車駅付近で GPS データがあれば妥当な経路が選ばれる。
- ・ 平行して運行している路線が多い区間は実際の経路とは異なる可能性がある。

推計結果から作成した鉄道利用者の時刻分布と大都市交通センサスの鉄道利用者の時刻分布との比較から分布の分布が非常に似ていたことから鉄道利用者の時刻分布は妥当であると判断した。

推計結果を駅ごとに集計した駅別乗降客数の結果では、GPS データさえあれば容易に駅ごとの利用者数を把握することができ、これまで行われてこなかった休日との比較も可能になった。さらに国土数値情報ダウンロードサービスの駅別乗降客数データとの比較では多くのユーザーが集まってくるような都心の駅においてはある程度妥当と考えられる結果が見られた。しかし、ユーザーの帰着駅となるような駅はユーザーが集まりやすいような駅に比べ、種類が非常に豊富であることから今回用いた GPS データのユーザー数では各帰着駅にユーザーが分散してしまい、各帰着駅における利用者数が非常に少なくなってしまった。(図 6.1.1 参照)

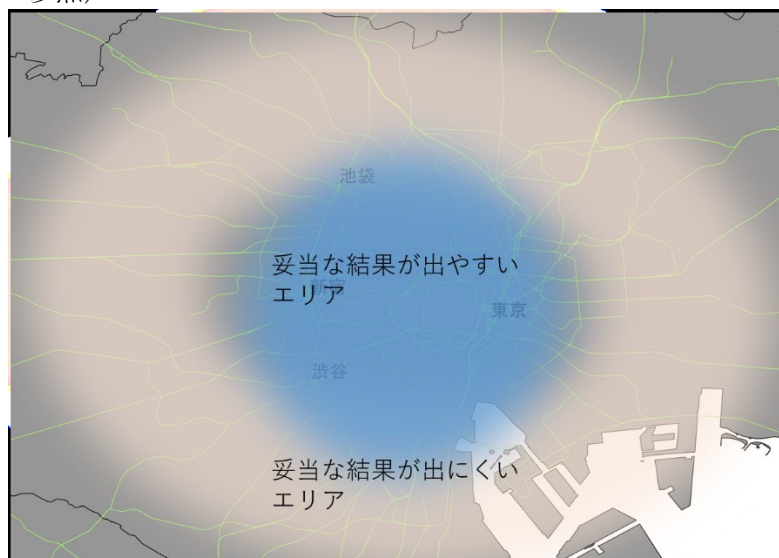


図 6.1.1 本研究において妥当な結果が出やすかったエリアと出にくかったエリア

そのため、帰着駅となるような駅における乗降客数をより正確に推定するにはさらに大規模な GPS データを用いることが出来れば帰着駅となるような駅の利用者も増えると考えられるため、今回用いたデータでは相関が弱かった駅においても相関が強まる可能性がある。が必要であることがわかった。しかし、駅別乗降客数全体の比較としては非常に高い相関係数が算出されたため、本研究における手法自体は有効であると考えている。

駅間移動人員数についての集計における路線ごとの利用状況の時間的推移の結果より、駅間の移動人員数もある程度妥当性のある結果が得られた。大都市交通センサスでは 5 年に 1 度アンケートで行っているため、非常に手間がかかり、アンケート調査であるためユーザーが回答することでユーザーへの負担も大きい、本研究の手法を用いることでこれらの手間や負担をかけずに調査を行える可能性を示した。

駅間移動人員数の比較では、大都市交通センサスの線別駅間移動人員数データを経路探索を用いて変換する必要があったことや駅別乗降客数の比較において明らかとなった今回の GPS データの特性もあり、全体では妥当な比較結果が得られなかった。しかし、今回の GPS データの特性を考慮して抽出した山手線の路線区間のみでの比較では一部を除き、非常に強い相関が見られ、本研究を利用した駅間移動人員数推定は有効である可能性が大きいと判断した。

以上より、本研究において提案した GSP データを用いた鉄道利用実態把握において乗降客数が 10 万人以上の駅では国土数値情報の駅別乗降客数のデータと強い相関が見られるといった 1 つの定量的な線引きを行うことが出来た。しかし、利用者が多い駅間における路線ごとの利用者数は大都市交通センサスの線別駅間移動人員数データとの間の相関はあまり強くなかったため、経路の選択手法について今後さらなる検討が必要である。

6.2 本研究の課題と展望

本研究における課題はまず、鉄道ネットワークデータについてであり、様々な経路探索を試すとやはりまだ駅同士・路線同士が上手く繋がっていないことで探索結果が妥当なものにならないことである。また、妥当な探索結果が出ない理由として今回は所要時間をコストとし乗り換えを行うことに重みをつけ、乗り換え数が減り妥当な結果が得られるようにしたが、乗り換え時間、列車の速度といったパラメータを利用していないことも原因であると考えられる。そして、これらのようなパラメータを利用することが出来れば、本研究における推定結果よりも妥当な経路探索結果が得られるであろう。

また、鉄道ネットワークデータの検証において発生した、平行して運行している路線へのマッチング方法の修正については特に重要である。東京圏においては路線同士が平行して運行している区間は都心にちかづくに連れて非常に多くなるため、これを修正することで都心における実態に沿わない乗り換えが減ると考えられる。

さらに、本研究において作成した鉄道ネットワークデータは直通運転、相互乗り入れといったさらに実態に沿った鉄道ネットワークに改良できる余地はあるが、これらに関しては直通運転区間、相互乗り入れ区間についての信頼のおけるデータが手に入らない限り、困難であると考えられる。

今後は鉄道ネットワークデータについて、現状ではまだ作成が上手くいっていない箇所もあるため、まずはこれらの修正を行い、可能な限り鉄道ネットワークデータを信頼できるものにしていく。そして、鉄道ネットワークの確立後は鉄道経路の選択手法の見直しを行っていく。

参考文献

- [1] NTT ソフトウェア株式会社. 導入事例 東日本旅客鉄道株式会社様,
最終アクセス日 2017-01-22
<https://www.ntts.co.jp/case/jreast1603/>
- [2] 東京都市圏交通計画協議会. パーソントリップ調査, 最終アクセス日 2017-01-22.
<http://www.tokyo-pt.jp/person/>
- [3] 国土交通省. 大都市交通センサス, 鉄道利用者に関する推計 (第 7 章),
最終アクセス日 2017-01-22
<http://www.mlit.go.jp/common/001001534.pdf>
<http://www.mlit.go.jp/common/001001538.pdf>
- [4] 国土交通省. 大都市交通センサスの調査概要
最終アクセス日 2017-01-22
http://www.mlit.go.jp/sogoseisaku/transport/sosei_transport_fr_000077.html
- [5] 関本 義秀, Teerayut Horanont, 柴崎 亮介,
“携帯電話を活用した人々の流動解析技術の潮流”, 2011, 情報処理, 52(12), pp.1522-1530.
- [6] 国土交通省. 国土数値情報ダウンロードサービス, 最終アクセス日 2017-01-22
<http://nlftp.mlit.go.jp/ksj/index.html>
- [7] 株式会社コードプラス. 駅データ.jp, 最終アクセス日 2017-01-22
<http://www.ekidata.jp/>
- [8] 東日本旅客鉄道株式会社. JR 東日本アプリ, 最終アクセス日 2017-01-22
<http://www.jreast-app.jp/>
- [9] 鯉渕 正裕, 加藤 勲, “移動軌跡データを活用した鉄道利用者の交通行動把握”, 2009,
Vol.2009-ITS-37 No.5
- [10] 中野 隆介, 沼尾 雅之, “無線 LAN アクセスポイントへの検索要求を利用した鉄道車
内混雑推定” 2012, DEIM Forum 2012, A10-1
- [11] 大野 夏海, “携帯電話による大規模・長期間の GPS データを用いた, 東京都市圏に
おける交通モードの推定およびモビリティの分析”, 2013, 東京大学, 新領域創成科学
研究科社会文化環境学専攻, 修士論文
- [12] 原 祐輔, 桑原 雅夫, “スパースなデータに対するマップマッチング手法と頻度変更型
測位に関する研究”, 2013, 第 33 回交通工学会論文集, pp. 415-418
- [13] 国土交通省. 国土数値ダウンロードサービス 鉄道データ,
最終アクセス日 2017-01-22
http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-N02-v2_3.html
- [14] 金杉 洋, 関本 義秀, 檜山 武浩, “人々の流動再現に向けた鉄道インフラデータの構
築”, 2013, 第 22 回地理情報システム学会, F-3-3

- [15] 情報銀行, 最終アクセス日 2017-01-22
<https://ibank.iis.u-tokyo.ac.jp/ibank>
- [16] 国土交通省. 国土数値情報ダウンロードサービス 駅別乗降客数データ,
最終アクセス日 2017-01-22
http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-S12-v2_2.html
- [17] 国土交通省. 第 11 回大都市交通センサス調査結果集計表,
最終アクセス日 2017-01-22
http://www.mlit.go.jp/sogoseisaku/transport/sosei_transport_tk_000035.html
- [18] 円山琢也, ”スマホ・アプリ配布型大規模交通調査の可能性”, 交通工学会, Vol.48,
No.01, pp.4-7, 2013

謝辞

本修士論文を書き上げるにあたり、多くの方からご指導、ご協力を頂きました。心より感謝申し上げます。

柴崎先生には、二度の国際学会を始め、二度の国内の学会への参加といった多くの挑戦の場を提供していただきました。私は発表があまり得意ではなかったため、学会前の発表練習では何度も厳しい言葉頂きました。しかし、先生の言葉を受け、「伝わりやすい」発表の大切さを実感し、練習を重ねました。その結果もあり、私はブルガリアで開催された国際学会にて新人賞を受賞することができました。受賞したときは他の学会参加者の方々にたくさん祝っていただき、今までがんばってきたことが報われました。このような経験が出来たこともすべて先生からのご指導のおかげだと思っております。心より感謝いたします。

副指導教員としてご指導いただいた日下部先生には研究全体について私が普段お世話になっている研究員の方とは少し違った視点で多くの意見を頂くことができとても良い刺激頂きました。しかし、私が病気を患ってしまったことで本来ならばさらに多くのご指導を頂くことができたと考ええると残念でなりません。それでも二度の日下部先生からのご指導は日下部先生の気さくな人柄もあり、とても楽しく刺激的で貴重な経験となりました。有り難うございました。

研究員の金杉さんには技術的な点で大変お世話になりました。私は、高度なプログラミング技術は持っていなかったため、非常にお忙しいにも関わらず金杉さんには何度もご指導いただきました。初めは金杉さんのおっしゃられることのほとんどが理解不能でしたが、今では少しは金杉さんの話を理解できるようになってきたことで自身の成長を感じました。また、金杉さんの指導がなければ私の研究を現在の段階まで進めることは到底不可能でした。本当に感謝いたします。

また、研究員の松原さんにも技術的な点でお世話になりました。松原さん、金杉さん、そして同期の和田君を含めた定例のミーティングは大変でしたが非常に楽しい時間でした。ミーティング中に松原さんから何度も的確なアドバイスをしていただき、有り難うございました。

秋山先生には研究の駆け出しの段階で右も左もわからない状態の私に優しく指導していただき、大変お世話になりました。また、韓国にて開催された国際学会では学会だけではなく、韓国のインチョン大学、韓国国土地理院において発表の場を与えて頂き、とても良い経験となりました。

研究員の小川さん、そして小野さんには主に研究の合間に何度も興味深い話を伺い、私の人生の糧となる知識を多く身につけることが出来ました。また、プライベートでもお世話になり、研究室外でも楽しい時間を過ごすことができました。

同期の和田君、西本さん、には公私ともにお世話になりました。特に和田君とは何度も行動を共にすることが増え、気軽に話せる間柄になり、互いに不満を打ち明け合ったりすることでお互いにモチベーションを高め合い、大変な研究生活を乗り切るパワーを得ることが出来ました。また、ブルガリアで開催された国際学会からの帰り道において飛行機のトラブルに遭った際は私 1 人では乗り切れなかったと思います。西本さんとは途中から関わりは薄くなってしまいましたが、西本さんの底抜けの明るさをみて研究生活で暗い気持ちになってもがんばる気力を起こすことができました。

そして、株式会社ゼンリンデータコムの皆様には本研究で用いた GPS データの提供及び、同データの相対的な集計化処理をして頂きました。普通であれば利用することが出来ないような貴重な GPS データを用いた研究を行えたことは他の学校や研究室では体験できないような大変貴重な経験をすることが出来たと痛感しております。本当に有り難うございました。

今回名前を挙げた以外にも多くの方々に研究生生活，プライベートの両面でお世話になりました．皆様のおかげで決して楽ではありませんでしたが，非常に充実した 2 年間を過ごすことができ，本趣旨論文を書き上げることが出来ました．最後に急遽本大学院に進学することになったにも関わらず私の学生生活を支え続けてくれた家族に感謝の意を表して，謝辞とさせていただきます．