

経路検索データを用いた 鉄道路線の需要予測モデルの提案

細江 美欧¹・桑野 将司²・森山 卓³

¹正会員 鳥取大学助教 工学研究科社会基盤工学専攻 (〒680-8552 鳥取県鳥取市湖山町南 4-101)
E-mail: mhosoe@tottori-u.ac.jp

²正会員 鳥取大学教授 工学研究科社会基盤工学専攻 (〒680-8552 鳥取県鳥取市湖山町南 4-101)
E-mail: kuwano@tottori-u.ac.jp

³正会員 鳥取大学助教 工学研究科社会基盤工学専攻 (〒680-8552 鳥取県鳥取市湖山町南 4-101)
E-mail: moriyama@tottori-u.ac.jp

公共交通の経路検索システムに蓄積される経路検索データには、検索者の「いつ、どこから、どこへ移動したい」という将来の移動意向が記録されている。本研究は、経路検索システムに現れる潜在需要が公共交通の需要変動の先行指標になる可能性に着目し、経路検索数から鉄道利用者数を予測するモデルの提案を目的とする。実証分析に際して、香川県で運行される鉄道路線「ことでん」を分析対象として、経路検索データと交通系 IC カードデータを用いた。2変量状態空間モデルと重み付き回帰モデルを組み合わせたアプローチにより、一見関連がないように変動する 2 つのデータに潜む関係性を抽出し、行動論的に関連のある部分をモデル化した。その結果、経路検索データを用いることによって、日々変動する交通需要を事前に予測できることを示した。

Key Words: route-search system, RP/SP model, state-space model, weighted regression model, big data

1. はじめに

公共交通の交通需要は交通事故や天候変化などの外的要因と、急な予定変更などの人的内的要因によって、経時的に変化する。そのため、交通需要と変動は切り離すことはできない。ここで、変動とは確率過程的な現象の移ろいと定義する¹⁾。変動のタイプには、偶発的な事象や特定が困難な事象による現象の移ろいに加え、周期をもって繰り返す現象の移ろいも含むとする¹⁾。

公共交通のダイヤや経路などの運行計画は、日々変更されるものではないため、平均的な交通需要予測に基づいて運用を計画することが望ましい。しかし、需要応答型交通や MaaS (Mobility as a Service) など、新しい交通サービスの運用においては、日々変動する交通需要の予測が必要である。特に、突発的に交通需要が増減する場合は、サービス利用者の満足度の低下や交通事業者の収益機会の損失を招く可能性があるため、事前に変動を予測し、対応策を講じることが重要といえる。従前、突発的な需要変動は天候やイベントなどの変動を引き起こす要因を外生変数とした仮説的因果構造分析によってモデル化されてきたが、その外生変数の予測が難しく、過去の

現象の記述に留まっている²⁾。

イベント発生による急な乗客増加に伴う運行遅れや、天候悪化による運休が生じた場合あるいは生じる可能性のある場合に、公共交通は定時性が保持されない。このとき、公共交通利用者はいつ車両が来るのか、そもそも自身の移動手段は確保されるかを不安に思う。この不安を解消させる手段に経路検索システムがある。経路検索システムを用いることによって、活動目的を達成するための経路を検索し、普段とは違う状況における移動に対する不安を解消できる。つまり、経路検索システムに蓄積される検索履歴データ（以降、経路検索データ）には、平常時とは異なる状況で、あらゆる要因が関係しあった結果、突発的に発生する需要が反映されていると考えられる。そのため、突発的な交通需要の発生に影響する要因について議論せずとも、経路検索データを活用することで、将来発生しうる突発的な交通需要を予測できる可能性がある。そこで、本研究では、経路検索データを用いて、突発的に発生する交通需要を含む日々の交通需要を予測するためのモデルの提案を目的とする。

2. 本研究の位置付け

四段階推定法のような伝統的な交通需要予測モデルでは、パーソントリップデータが用いられてきた。パーソントリップ調査は代表的な1日の行動を観測したものであるため、交通需要の変動に関する情報は含まれていない。一方で、ビッグデータは長期的に継続して収集されることから、現象の移ろいが観測できる。データ量が膨大であることはビッグデータの特徴であるが、データ量の多さに限れば、その特性を持つデータはビッグデータが目される以前から存在する。例えば、先に挙げたパーソントリップデータも何千人、何万人の回答にもなれば、膨大なデータ量を持つ。しかし、ビッグデータの本質はデータ量の多さではなく、時々刻々と変化する現象が記録されている点、すなわち、変動性を有するところにあると考える。

ビッグデータを用いた交通需要予測では過去の実績に基づく RP (Revealed Preferences) ビッグデータが用いられており、定常状態を予測するモデルが多い¹⁰⁾。近年は、非定常状態、すなわち突発的に変動する需要の予測を主眼とする研究も蓄積されている。交通需要の急激な増減を予測するために、既往研究では、時間的あるいは空間的に予測範囲を絞り込むことが多い。時間的に予測範囲を絞り込むとは、1時点先の変動を予測する場合に、現時点と1時点先の時間間隔を短くすることである。既往研究では、データの過去の変動傾向から将来の変動を予測するモデルが提案されている⁴⁾¹¹⁾。将来の変動は直前の変動に最も影響を受け、変化すると考えられるので、時間間隔は短い方が予測精度は高い¹²⁾。しかし、突発的な需要を正確に予測できたとしても、予測時点までの時間間隔が短いほど、交通事業者は需要の変化に対応するための時間を持たず、実用的な交通需要予測モデルとはいえない。

空間的に予測範囲を絞り込むとは、分析対象を対象地域全体ではなく、ある特定の範囲内やある路線というように分析対象を細分化し、限定することである。空間的に範囲を絞り込めば変動要因となりうる事象(例えば、天候やイベント)に関するデータを入手しやすい。そのため、既往研究においては、分析対象範囲を絞り、天候変化やイベントによる突発的な交通需要の発生を予測するモデルが提案されている²⁹⁾。しかし、天候変化の程度やイベントの規模は多様であり、毎回内容が同じであるとは考えられない。つまり、過去の天候変化や特定のイベントと突発的な交通需要の関係性を記述できても、将来の天候変化やイベント開催による突発的な交通需要の発生は予測できない。したがって、RP ビッグデータを用いた突発的な交通需要予測のための有効なモデルは確立されているとはいえない。

一方、交通需要予測では SP (Stated Preferences) データと RP データそれぞれを単体で用いた分析よりも、互いに補完的な性質を持つ両データを用いた分析の方が予測精度は高いといわれており、従前より SP データと RP データを用いた方法論が提案されている¹³⁾¹⁴⁾。既往研究において、RP/SP モデルの有効性は十分に実証されているが、多くはアンケート調査に基づくものである。これは将来の表明選好に関する交通系ビッグデータがほとんど存在しないことに起因する。そのような中で、唯一、交通系ビッグデータの中で SP ビッグデータに位置付けられるのが経路検索データである。しかし、既往研究において経路検索データを活用した交通需要予測モデルの議論は筆者らの知る限り存在しない。

経路検索データが RP データでは観測できない潜在的な交通需要を表す唯一の SP ビッグデータであることを踏まえると、交通需要予測の重要な情報源としてその活用可能性を検討することの意義は高い。また、RP ビッグデータに加えて、SP ビッグデータを用いることで、交通系ビッグデータを活用した交通需要予測においても予測精度の向上が期待される。

本研究では、公共交通の実際の利用者数を表す RP ビッグデータとして交通系 IC カードの利用履歴データ(以降、交通系 IC カードデータ)を、SP ビッグデータとして経路検索データを用いる。そして、突発的な需要の発生要因の解明に注力するのではなく、すべての要因が集約された代理変数が経路検索データであると仮定し、これによる新たな交通需要予測モデルを提案する。提案モデルによって交通需要の変動を捉えることができれば、変容する交通需要に応じた事前準備ができ、実用的意義は大きいと考える。

3. 経路検索と公共交通利用を関連づけるための仮説

経路検索データはあくまで Web 上に表明される潜在需要であり、実際にどの程度顕在化するのかが明らかではない。また、経路検索システムが平常時とは異なる移動の際に利用されることを考えると、経路検索データですべての交通需要を予測できるわけではない。

そこで、経路検索データで予測可能な交通需要を表-1に基づいて整理する。ここで、個人レベルでの経路検索行動と公共交通利用行動を直接紐づけるデータは存在しないため、分析対象とする交通機関で集計した経路検索数と交通系 IC カード利用数を用いて分析する。本研究では、個人レベルの「行動」は、集計レベルでは「需要」として観測され、経路検索データと交通系 IC カードデータには、日々安定的に発生する「定常的需要」と平常

表-1 経路検索データと交通系 IC カードデータが表す行動と需要

経路検索データ		交通系 IC カードデータ	
個人レベル	集計レベル	個人レベル	集計レベル
非定期的行動 例) 非定期的買い物 非定期的通院 観光 出張	定期的需要	定期的行動 例) 通勤・通学 定期的買い物 定期的通院 非定期的行動 例) 非定期的買い物 非定期的通院 観光 出張	定期的需要
非定期的行動 例) イベント参加 悪天候回避	突発的需要	非定期的行動 例) イベント参加 悪天候回避	突発的需要

時と比べて非周期的に増減する「突発的需要」が異なる割合でそれぞれ包含されるとする。

表-1 に示す通り、個人レベルでの交通系 IC カード利用には定期的行動と非定期的行動によるものがあり、集計レベルでの需要の大部分は利用者の定期的行動の積み重ねである定期的需要と考える。また、個人レベルでの交通系 IC カード利用の非定期的行動も日々一定数発生し、集計レベルでは定期的需要として観測されるといえる。一方、経路検索システムは個人レベルでの非定期的行動が行われるときに利用されるものの、これもまた集計レベルでは定期的需要として一定数観測され、交通系 IC カードデータと同様に、需要の大部分は定期的需要が占めると考えられる。ただし、両データに内在する定期的需要は、集計された個人レベルの行動に相違があるため関連づけられない。

しかし、突発的需要については、経路検索データも交通系 IC カードデータも、イベント参加や悪天候回避などが共通要因となる非定期的行動を集計したものと考えられることから関連づけられる可能性がある。本研究では、経路検索データと交通系 IC カードデータに内在する突発的需要部分をモデル化し、経路検索数から交通系 IC カード利用数の変動を予測する。そのために、2 変量状態空間モデルを用いて、経路検索データと交通系 IC カードデータそれぞれについて定期的需要と突発的需要を統計的に分解する。そして、関連のある需要部分を重み付き回帰モデルによってモデル化する。

4. 鉄道路線「ことでん」と使用データの概要

(1) ことでんの概要

本研究では、香川県内を運行する「高松琴平電鉄（以下、ことでん）」を分析対象とする。ことでんは高松琴平電気鉄道株式会社によって運行されている地方私鉄である。図-1 に示すように琴平線、長尾線、志度線の 3 路線全 52 駅が存在し、香川県高松市の中心部から放射状に路線が広がっている。高松築港駅から瓦町駅の区間は琴平線と長尾線の 2 路線が共用しており、ことでんの中で最も運行本数が多い区間である。この区間は高松市の中心市街地内に存在する。高松築港駅の近くには JR 高松駅や高松港が位置しており、他交通機関への乗り換えが可能である。片原町駅周辺には百貨店や丸亀町商店街などが存在し、高松市の商業の中心地である。瓦町駅は 3 路線の結節点となっており、3 路線すべてに乗り換え可能である。中心部から離れた駅周辺には、住宅地や観光地が広がる。

(2) 使用データの概要

本研究では、経路検索システム「NAVITIME」の経路検索データに加えて、鉄道利用者数を表すデータとして、交通系 IC カード「InuCa」の乗降履歴データを使用する。

a) 「NAVITIME」の経路検索データ

NAVITIME を用いて、起終点にことでんの駅を指定し、利用指定日を 2014 年 1 月 1 日～2014 年 12 月 31 日の 1 年以内とした経路検索データ 195,448 件を用いる。

図-2 に 2014 年における日別経路検索数の推移を示す。1 日あたりの平均経路検索数は約 535 件であり、最も経路検索数が多かったのは 8 月 13 日の 1,503 件、最も少なかったのは 10 月 13 日の 188 件であった。

図-3 には曜日別平均経路検索数を示す。曜日別平均経路検索数は月曜日から木曜日にかけて減少し、木曜日に最も少ない。そして、金曜日と土曜日に増加し、土曜日に最も多いが、日曜日に再び減少する。金曜日から日曜日の平均経路検索数が他の曜日よりも多いことから、平

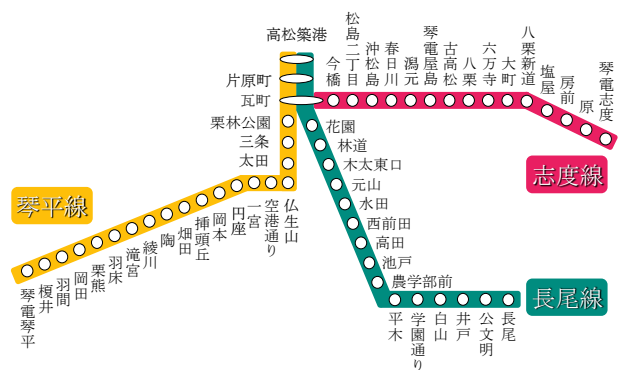


図-1 ことでん路線図¹⁹⁾

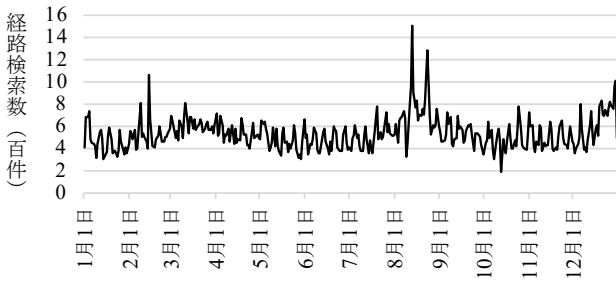


図-2 日別経路検索数の推移

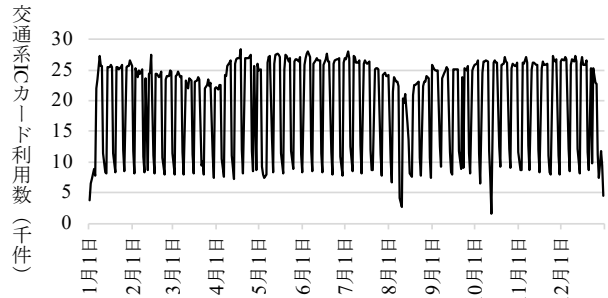


図-5 日別交通系 IC カード利用数の推移

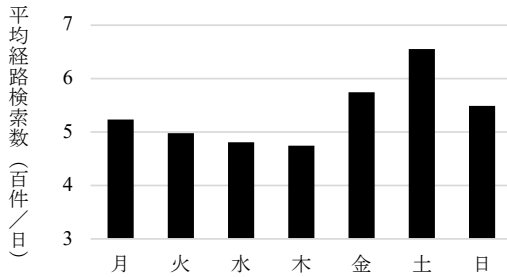


図-3 曜日別平均経路検索数

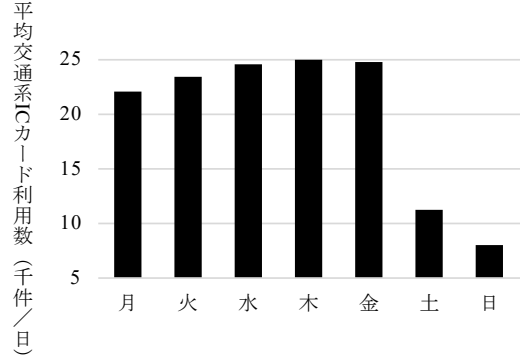


図-6 曜日別平均交通系 IC カード利用数

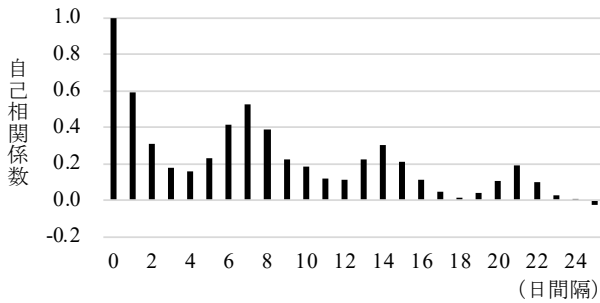


図-4 経路検索数のコレログラム

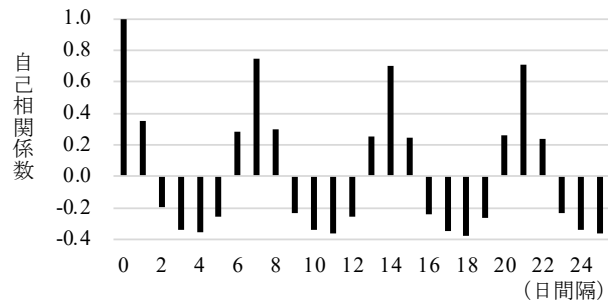


図-7 交通系 IC カード利用数のコレログラム

日よりも休日の方が検索されやすいといえる。さらに、図-4 の経路検索数のコレログラムからは、7 日の周期で強い自己相関があるとわかり、曜日による周期性の存在が示唆された。

b) 「InuCa」の利用履歴データ

ことでは独自の交通系 IC カード「InuCa」が導入されている。本研究では経路検索データと同様の利用期間である 2014 年の 1 年間に収集された利用履歴データ 7,265,029 件を用いる。

図-5 に 2014 年における日別交通系 IC カード利用数の推移を示す。1 日あたりの平均交通系 IC カード利用数は約 19,904 件であり、最も交通系 IC カード利用数が多かったのは 4 月 18 日の 28,378 件、最も少なかったのは 10 月 13 日の 1,623 件であった。交通系 IC カード利用数の最少日は経路検索数の最少日と一致し、この日は台風によることでの運休や外出の見合わせのために、利用が減ったと考えられる。

図-6 には曜日別平均交通系 IC カード利用数を示す。曜日別平均交通系 IC カード利用数は月曜日から木曜日にかけて増加し、金曜日から日曜日では減少している。

また、平日の平均交通系 IC カード利用数 23,988 件は休日の平均値 9,649 件に対して 2.5 倍であり、ことでは平日よりも平日に利用される。さらに、図-7 の交通系 IC カード利用数のコレログラムからは、7 日の周期で強い自己相関があるとわかり、経路検索数と同様に交通系 IC カード利用数にも曜日による周期性が確認された。

c) 経路検索数と交通系 IC カード利用数の相関分析

経路検索データを用いて、日々の交通系 IC カード利用数を予測するあたり、両者の単純な関係性を明らかにする。図-8 に日別経路検索数と日別交通系 IC カード利用数の散布図を示す。

全体の相関係数を算出したところ、-0.35 と負の相関があることがわかった。一方で、平日と休日別に相関係数を算出すると、平日では-0.52 と負の相関に、休日では 0.50 と正の相関があることが明らかとなった。

平日においては、通勤・通学、定期的な買い物や通院に関連する定期的行動が多く行われると考えられる。このとき、ことでは利用者は普段行う移動のために、経路検

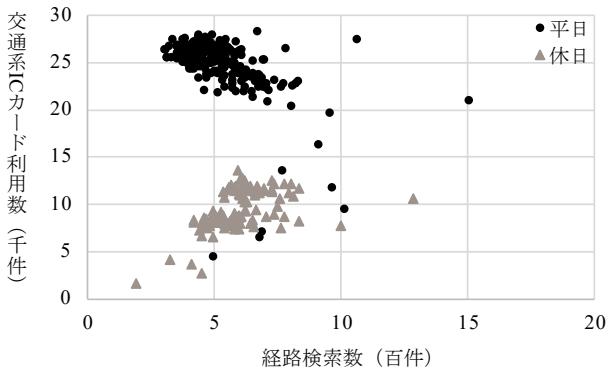


図8 日別経路検索数と日別交通系 IC カード利用数の散布図

索システムを利用しない。そのため、交通系 IC カード利用数は多くなるが、経路検索数は少なくなり、両者の相関は負となると推察される。一方、休日においては、定期的行動はあまり行われず、非定期的の買い物や通院、観光や出張といった非定期的行動が多くなる。普段とは異なる移動の際に経路検索システムが利用されることから、休日は経路検索数が多くなり、非定期的行動による交通系 IC カード利用数も多くなることで、両者は正の相関になると考えられる。

以上より、経路検索システムと交通系 IC カードは利用されやすい状況が異なるため、両者は直接的には結びつけられないといえる。そこで、次章では、経路検索数と交通系 IC カード利用数を関連づける方法論を提案し、予測モデルを構築する。

5. 経路検索数に基づく交通系 IC カード利用数予測モデルの構築

本研究では、 t 時点までに収集した日別経路検索数と日別交通系 IC カード利用数を用いて、 $t + 1$ 時点の交通系 IC カード利用数を予測するモデルを提案する。提案モデルは 6 つの手順で構成される。図-9 に提案モデルの手順を示す。

経路検索データと交通系 IC カードデータには定常的需要と突発的需要が異なる割合で混在する。さらに、2 種類の需要はデータ間で互いに影響を及ぼしながら変動すると考えられる。そこで、Step1 では、線形・ガウス型の 2 変量状態空間モデル¹⁶⁾²¹⁾を用いて、経路検索データと交通系 IC カードデータの両時系列データに含まれる変動成分を分解する。

状態空間モデルは時系列の変動をいくつかの成分に分解でき、多様なバリエーションのモデルがある。本研究では、4.(2)節のコログラムの結果を踏まえ、確定的周期要素のある確率的ローカルレベルモデルを適用する。つまり、日別経路検索数と日別交通系 IC カード利用数

Step1	2変量状態空間モデルによる 経路検索数と交通系ICカード利用数の成分分解
Step2	重み付き回帰モデルによる 経路検索数と交通系ICカード利用数の誤差成分の モデル化
Step3	$t + 1$ 時点の経路検索数と交通系ICカード利用数の レベル成分、回帰成分の予測
Step4	$t + 1$ 時点の経路検索数の誤差成分の予測
Step5	$t + 1$ 時点の交通系ICカード利用数の誤差成分の 予測
Step6	$t + 1$ 時点の交通系ICカード利用数の予測

図-9 提案モデルの予測手順

の変動には、確率的レベル成分、周期成分（曜日成分と祝日成分）、誤差成分が存在すると仮定する。2 変量状態空間モデルの観測方程式と状態方程式はそれぞれ式(1)と式(2)のようになる。

$$\begin{bmatrix} y_S \\ y_I \end{bmatrix}_t = \begin{bmatrix} x_S \\ x_I \end{bmatrix}_t + \begin{bmatrix} D_{S,Mon} \cdots D_{S,Sun} & D_{S,PH} \\ D_{I,Mon} \cdots D_{I,Sun} & D_{I,PH} \end{bmatrix} \begin{bmatrix} d_{Mon} \\ \vdots \\ d_{Sun} \\ d_{PH} \end{bmatrix} + \begin{bmatrix} v_S \\ v_I \end{bmatrix}_t \quad (1)$$

$$\begin{bmatrix} v_S \\ v_I \end{bmatrix}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R}_t)$$

$$\begin{bmatrix} x_S \\ x_I \end{bmatrix}_t = \begin{bmatrix} x_S \\ x_I \end{bmatrix}_{t-1} + \begin{bmatrix} w_S \\ w_I \end{bmatrix}_t \quad (2)$$

$$\begin{bmatrix} w_S \\ w_I \end{bmatrix}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q}_t)$$

ただし、

$y_{S,t}, y_{I,t}$: 経路検索数と交通系 IC カード利用数の観測値

$x_{S,t}, x_{I,t}$: 経路検索数と交通系 IC カード利用数のレベル成分

$D_{S,Mon}, \dots, D_{S,Sun}, D_{S,PH}$: 経路検索数の曜日成分・祝日成分

$D_{I,Mon}, \dots, D_{I,Sun}, D_{I,PH}$: 交通系 IC カード利用数の曜日成分・祝日成分

$d_{Mon,t}, \dots, d_{Sun,t}, d_{PH,t}$: 曜日・祝日ダミー変数

$v_{S,t}$: 経路検索数の誤差成分 (観測誤差)

$v_{I,t}$: 交通系 IC カード利用数の誤差成分 (観測誤差)

$w_{S,t}$: 経路検索数の過程誤差

$w_{I,t}$: 交通系 IC カード利用数の過程誤差

\mathbf{R}_t : 観測誤差の分散共分散行列

\mathbf{Q}_t : 過程誤差の分散共分散行列

である。

時系列の変動をレベル成分、回帰成分、誤差成分に分けたときに、レベル成分と回帰成分の和で表現されるのが定常的需要、誤差成分で表現されるのが突発的需要と考える。そして、Step2 で、行動論的に関連があると考えられる突発的需要の関係をモデル化する。ここで、突発的需要の変動が大きい日の予測精度を向上させるため

に、重み付き回帰モデルを適用する。

Step3 では、Step1 で推定した t 時点までの経路検索数と交通系 IC カード利用数のレベル成分、回帰成分を用いて、次時点の各成分を予測する。 $t + 1$ 時点のレベル成分と回帰成分はそれぞれ式(3)、式(4)より求める。

$$\begin{aligned} \begin{bmatrix} \hat{x}_S \\ \hat{x}_I \end{bmatrix}_{t+1} &= \begin{bmatrix} x_S \\ x_I \end{bmatrix}_t + \begin{bmatrix} w_S \\ w_I \end{bmatrix}_{t+1} \\ &= \begin{bmatrix} x_S \\ x_I \end{bmatrix}_t + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x_S \\ x_I \end{bmatrix}_t \end{aligned} \quad (3)$$

$$\begin{aligned} \begin{bmatrix} \hat{D}_{S,Mon} \cdots \hat{D}_{S,Sun} \hat{D}_{S,PH} \\ \hat{D}_{I,Mon} \cdots \hat{D}_{I,Sun} \hat{D}_{I,PH} \end{bmatrix}_{t+1} \\ = \begin{bmatrix} D_{S,Mon} \cdots D_{S,Sun} D_{S,PH} \\ D_{I,Mon} \cdots D_{I,Sun} D_{I,PH} \end{bmatrix} \begin{bmatrix} d_{Mon} \\ \vdots \\ d_{Sun} \\ d_{PH} \end{bmatrix}_{t+1} \end{aligned} \quad (4)$$

式(3)は、 t 時点のレベル成分 $x_{S,t}$ 、 $x_{I,t}$ と $t + 1$ 時点の過程誤差 $w_{S,t+1}$ 、 $w_{I,t+1}$ で構成される。ここで、 $x_{S,t}$ 、 $x_{I,t}$ は 2 変量状態空間モデルの適用により推定された t 時点のレベル成分の推定値であるため予測可能である。しかし、過程誤差 $w_{S,t+1}$ 、 $w_{I,t+1}$ は次時点 $t + 1$ においてどのような値を取るかを事前に予測できない。そこで、予測にあたっては、 $w_{S,t}$ 、 $w_{I,t}$ を 0 とし、過程誤差の相関は捨象する。そのため、時点 $t + 1$ のレベル成分 $\hat{x}_{S,t+1}$ 、 $\hat{x}_{I,t+1}$ は近似値となる。

式(4)は、2 変量状態空間モデルより推定された回帰成分 \hat{D}_S 、 \hat{D}_I と $t + 1$ 時点の曜日・祝日ダミー変数 d_{t+1} からなり、これらの積で $t + 1$ 時点の回帰成分が求まる。

このとき、 $t + 1$ 時点の経路検索数の予測値 $\hat{y}_{S,t+1}$ は、式(1)に基づいて式(5)で与えられる。

$$\hat{y}_{S,t+1} = \hat{x}_{S,t+1} + \hat{D}_S d_{t+1} + \hat{v}_{S,t+1} \quad (5)$$

式(5)において、誤差成分 $\hat{v}_{S,t+1}$ のみ未知である。 $t + 1$ 時点の経路検索数 $y_{S,t+1}$ は事前に得られるとすると、経路検索数の誤差成分の近似値 $\hat{v}_{S,t+1}$ は式(6)で求まる (Step4)。なお、ここでも観測誤差の相関は捨象する。

$$\begin{aligned} \hat{v}_{S,t+1} &= y_{S,t+1} - \hat{y}_{S,t+1} \\ &= y_{S,t+1} - (\hat{x}_{S,t+1} + \hat{D}_S d_{t+1}) \end{aligned} \quad (6)$$

式(6)によって経路検索数の誤差成分の近似値 $\hat{v}_{S,t+1}$ が求まれば、Step3 で推定した重み付き回帰モデルに基づいて、式(7)によって、 $\hat{v}_{S,t+1}$ から交通系 IC カード利用数の誤差成分の近似値 $\hat{v}_{I,t+1}$ が算出される (Step5)。

$$\hat{v}_{I,t+1} = \hat{\alpha} \times \hat{v}_{S,t+1} + \hat{\beta} \quad (7)$$

ただし、

$\hat{\alpha}$: 回帰係数

$\hat{\beta}$: 定数項

である。

最後に、Step6 では、Step3 で求めた $t + 1$ 時点の交通系 IC カード利用数のレベル成分、回帰成分と、Step5 で求めた $t + 1$ 時点の誤差成分を用いて、式(8)より $t + 1$ 時点の交通系 IC カード利用数を求める。

$$\hat{y}_{I,t+1} = \hat{x}_{I,t+1} + \hat{D}_I d_{t+1} + \hat{v}_{I,t+1} \quad (8)$$

6. 経路検索数の事前収集時間別交通系 IC カード利用数の予測

(1) 交通系 IC カード利用数の事前予測のための方法

突発的需要を含む日々の交通需要の予測に際して、各日の終日までに収集された日別経路検索数では、その日の交通系 IC カード利用数を事前に予測できない。例えば、1月1日の交通系 IC カード利用数を予測する場合に、1月1日 24 時までに収集した経路検索数では事前予測は不可能である。そこで、本研究では、事前に収集可能な日別経路検索数に基づいて予測を行う。具体的には、1月1日の始発直前 (4 時台)、1月2日の始発直前、... というように、各日を指定した検索の検索日時 (検索者が経路検索した日時) がその日の a) 始発直前 (当日 4 時台) までの日別経路検索数を用いる。その他、各日を指定した検索で検索日時が b) 始発 6 時間前 (前日 23 時台) まで、c) 始発 12 時間前 (前日 17 時台) までの日別経路検索数を用いることとする。便宜上、a) から c) を事前収集時間と呼ぶ。モデル推計と予測は琴平線、長尾線、志度線の 3 つの路線ごとに行う。

モデルに投入する各路線の事前収集時間ごとの日別経路検索数の推移を図-10、図-11、図-12 に示す。図-10 より、琴平線について、事前収集時間を考慮すると、事前収集時間が長いほど徐々に経路検索数は少なくなることがわかる。長尾線と志度線についても同様である。これら事前収集時間ごとの日別経路検索数を用いて、各日のどのくらい前までに収集された経路検索データであれば、交通需要を事前に予測可能かを検討する。

日別交通系 IC カード利用数の予測と予測精度の検証は 2014 年 12 月の 1 ヶ月間を対象に行う。事前収集時間を a) 始発直前とし、2014 年 12 月 1 日の交通系 IC カード利用数を予測する場合、2 つのデータ収集期間は 2014 年 1 月 1 日から 11 月 30 日の始発直前までである。5 章で示した予測手順に基づいて、各日の 4 時台までに収集された経路検索数と、各日の交通系 IC カード利用数を用いて、パラメータを推定する (Step1 から Step2)。そし

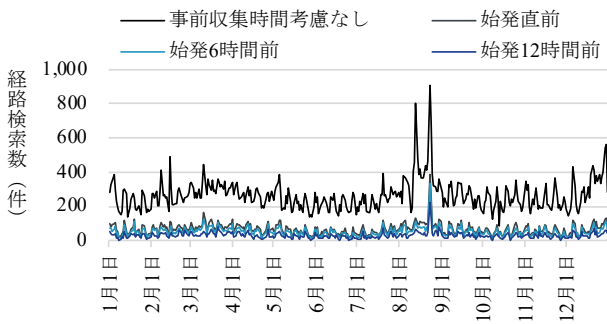


図-10 琴平線の事前収集時間ごとの日別経路検索数の推移

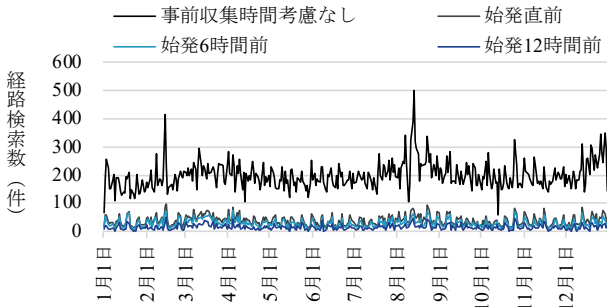


図-11 長尾線の事前収集時間ごとの日別経路検索数の推移

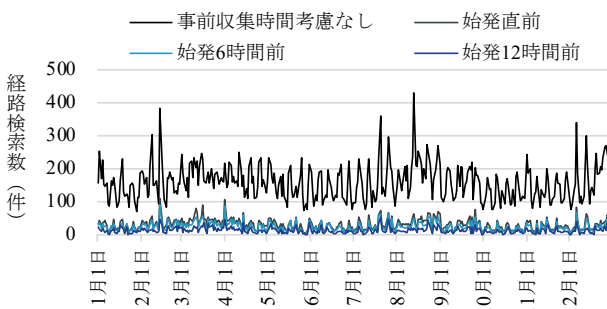


図-12 志度線の事前収集時間ごとの日別経路検索数の推移

て、12月1日4時台までに収集した経路検索数により、12月1日の交通系 IC カード利用数を予測する (Step3 から Step6)。

次に、12月2日の交通系 IC カード利用数を予測する場合、データ収集期間を1日増やし、2014年1月1日から12月1日とする。このとき、12月1日の交通系 IC カード利用数を予測したときと同様に、経路検索数については各日を指定した検索の検索時間がその日の4時台までのものである。そして、各日の4時台までに収集された経路検索数と、各日の交通系 IC カード利用数を用いてパラメータ推定を行い (Step1 から Step2)、12月2日4時台までに収集された経路検索数に基づいて12月2日の交通系 IC カード利用数を予測する (Step3 から Step6)。以後、次の時点の交通系 IC カード利用数を予測する場合、データの収集期間を1日ずつ増やしながらか、パラメータ推計を行い、予測日4時台までに収集された経路検索数に基づいて交通系 IC カード利用数を予測する。

(2) 経路検索件数の事前収集時間別予測精度の検証

3路線ごとに、各事前収集時間で収集した日別経路検索数を用いて推定した12月の日別交通系 IC カード利用数の予測値と観測値を図-13、図-14、図-15に示す。また、3路線ごとの事前収集時間別 RMSE (Root Mean Square Error: 平均平方二乗誤差) を表-2に示す。表-2中の丸括弧内は路線ごとの12月の平均交通系 IC カード利用数に対する誤差を表す。

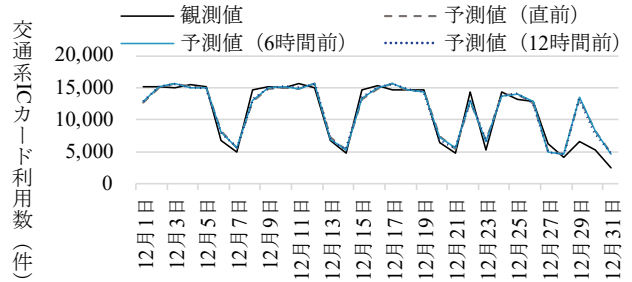


図-13 琴平線における12月の観測値と予測値

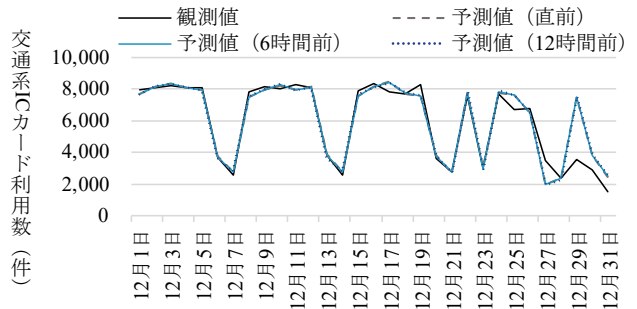


図-14 長尾線における12月の観測値と予測値

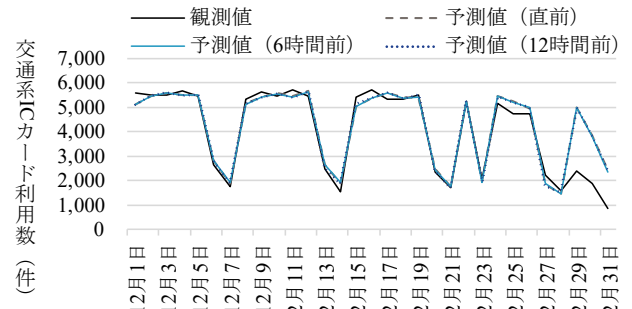


図-15 志度線における12月の観測値と予測値

表-2 事前収集時間別 RMSE (12月全日)

事前収集時間	RMSE (件)		
	琴平線	長尾線	志度線
直前	1,653 (15%)	843 (14%)	687 (17%)
3時間前	1,648 (15%)	844 (14%)	689 (17%)
6時間前	1,634 (15%)	842 (14%)	677 (17%)
12時間前	1,571 (14%)	855 (14%)	673 (17%)

表-3 事前収集時間別 RMSE (12月28日まで)

事前収集時間	RMSE (件)		
	琴平線	長尾線	志度線
直前	953 (9%)	411 (7%)	237 (6%)
3時間前	939 (8%)	410 (7%)	237 (6%)
6時間前	902 (8%)	413 (7%)	237 (6%)
12時間前	841 (8%)	417 (7%)	233 (6%)

表-2より、予測精度を表す RMSE は事前収集時間別に路線ごとで大差はない。各日の終日までに収集した経路検索数（事前収集時間を考慮しない場合の経路検索数）に基づくモデルの路線ごとの RMSE は、琴平線 1,460 件（誤差 13%）、長尾線 706 件（誤差 11%）、志度線 531 件（誤差 13%）であり、どの路線もいずれの事前収集時間の経路検索数を用いた場合でも、良好な予測精度を得られているといえる。ただし、図-13、図-14、図-15 に示すように、どの路線においても 12 月 29 日以降の予測値は観測値と大きくずれていた。そこで、28 日までを予測対象として RMSE を算出した結果を表-3 に示す。表-3 より、各路線の 12 月の平均交通系 IC カード利用数に対して 6~9% の誤差で予測できることが明らかとなった。

7. おわりに

経路検索データは公共交通の潜在需要を反映した交通ビッグデータとして注目されていたものの、交通需要予測において有用なデータであるのかはこれまで明示されてこなかった。さらに、SP ビッグデータとしての活用可能性についても議論が十分ではなかった。

本研究では、経路検索システム上に現れる潜在需要と実際に顕在化する公共交通の需要の関連のある部分をモデル化し、経路検索データを用いた新たな交通需要予測モデルを提案した。提案モデルは、経路検索データと交通系 IC カードデータの両方を用いて、将来の交通需要を予測するものである。実証分析の結果、予測日前日夕方までに収集した経路検索データであっても、全データ数を用いた予測（予測日終日までに収集した経路検索データでの予測）と同程度の精度となることがわかった。これより、提案モデルを用いることで、突発的需要を含む日々の交通系 IC カード利用数を事前に予測できることが示された。したがって、経路検索データによって、データ収集が困難なイベント情報や天候情報を用いずとも、日々変動する交通需要を予測できるといえる。

働き方や暮らし方の変化を通して人々の行動が多様化

している中で、これからの行動分析において求められることは、突発的変動を捉えることだと考える。突発的現象が起きたときには、日常とは異なる行動が発生するが、非日常の行動は稀にしか計測できないため、そのモニタリングは難しいとされてきた。一方、人々は非日常時の状態に置かれたときこそ、一層情報を求める。そのようなときに経路検索システムは使われ、公共交通における非日常時の行動をモニタリングできることが本研究成果より示された。今後、経路検索システムが単なる経路情報だけでなく、検索者の状況にあった付加情報が提供されるようになれば、経路検索データにはこれまで以上に多様な突発的現象に基づく変動が反映されるといえる。そこで、経路検索データから突発的発生の発生を予測することで、人々の急な行動変容に対応するための策を講じることや、非日常時に役立つ情報を提供することが可能になると考えられる。

本研究では、取得したデータの収集期間の都合上、1 年間分のサンプルを用いて分析を行ったが、図-13、図-14、図-15 に示すように、12 月 29 日から 31 日の 3 日間の交通系 IC カード利用数は高精度に予測できていない。12 月の最後 3 日間が当たらないのは、データの期間の最後に年末という特異日が来てしまったことが原因と考えられ、何もない平日や休日であれば予測精度が向上することが期待される。あるいは、データ期間を拡張し、2015 年 1 月 1 日以降のデータを含めた推計ができれば、年末という特異な動きを平滑化できる可能性がある。ただし、サンプル数が多くなるほど、パラメータ推計は安定するが、計算時間の増大やノイズが含まれる可能性が高まる。そこで、提案モデルにおいて、どの程度の期間のサンプル数を用いることで、高精度に予測できるかを今後検証する必要がある。

謝辞：本研究の遂行にあたり、香川高等専門学校建設環境工学科の宮崎耕輔教授、東京大学大学院情報理工学系研究科の伊藤昌毅准教授より、IruCa と NAVITIME のデータをご提供いただきました。記して感謝申し上げます。

参考文献

- 1) 北村隆一：変動についての試行的考察，土木計画学研究・論文集，Vol.20-1，pp.1-15，2003。
- 2) Guo, Z., Wilson, N., Rahbee, A.: Impact of weather on transit ridership in Chicago, Illinois, *Transportation Research Record: Journal of the Transportation Research Board*, Vol.2034, pp.3-10, 2007。
- 3) Kalkstein, A. J., Kuby, M., Gerrity, D., Clancy, J. J.: An analysis of air mass effects on rail ridership in three US cities, *Journal of transport geography*, Vol.17, pp.198-207, 2009。
- 4) Li, Y., Wang, X., Sun, S., Ma, X., Lu, G.: Forecasting short-term subway passenger flow under special events

- scenarios using multiscale radial basis function networks, *Transportation Research Part C*, Vol.77, pp.306-328, 2017.
- 5) Zhou, M., Wang, D., Li, Q., Yue, Y., Tu, W., Cao, R.: Impacts of weather on public transport ridership: Results from mining data from different sources, *Transportation Research Part C*, Vol.75, pp.17-29, 2017.
 - 6) 青地美桜, 白石陽: 交通系 IC カードの乗降履歴データを用いた路線バスの乗客人数予測手法の提案, 情報処理学会第 82 回全国講演論文集, Vol.2020, No.1, pp.353-354, 2020.
 - 7) Rodrigues, F., Borysov, S. S., Ribeiro, B., Pereira, F. C.: A bayesian additive model for understanding public transport usage in special events, *IEEE transactions on pattern analysis and machine intelligence*, Vol.39, pp.2113-2126, 2017.
 - 8) Pereira, F. C., Rodrigues, F., Ben-Akiva, M.: Using data from the web to predict public transport arrivals under special events scenarios, *Journal of Intelligent Transportation Systems*, Vol.19, pp.273-288, 2015.
 - 9) Ding, C., Wang, D., Ma, X., Li, H.: Predicting short-term subway ridership and prioritizing its influential factors using gradient boosting decision trees, *Sustainability*, Vol.8, 16 pages, 2016.
 - 10) 角野惇, 猪井博登, 土井健司: 交通系 IC カードデータを用いた路線バス需要予測, 第 50 回土木計画学研究・講演集, CD-ROM, 5 pages, 2014.
 - 11) Xue, R., Sun, D.J., Chen, S.: Short-term bus passenger demand prediction based on time series model and interactive multiple model approach, *Discrete Dynamics in Nature and Society*, 13 pages, 2015.
 - 12) Yang, X., Xue, Q., Ding, M., Wu, J., Gao, Z.: Short-term prediction of passenger volume for urban rail systems: A deep learning approach based on smart-card data, *International Journal of Production Economics*, Vol.231, 12 pages, 2021.
 - 13) Dissanayake, D. and Morikawa, T.: Travel Demand Models with the RP/SP Combining Technique for the Developing Countries, The International Conference CODATU IX, pp.103-107, 2000.
 - 14) 森川高行, 山田菊子: 系列相関を持つ RP データと SP データを同時に用いた離散型選択モデルの推定, 土木学会論文集, Vol.1993, No.476, pp.11-18, 1993.
 - 15) 高松琴平電気鉄道株式会社, 路線図, <https://www.kotoden.co.jp/publichtm/kotoden/time/>
 - 16) J.J.F. コマンダー, S.J. クープマン: 状態空間時系列入門, pp.9-127, シーエーピー出版, 2008.
 - 17) Holmes, E. E., Ward, E. J., Scheuerell, M. D.: Analysis of multivariate time-series using the MARSS package version 3.10.12, Northwest Fisheries Science Center NOAA Seattle, WA, USA, 2020.
 - 18) 石田基広, 萩原淳一郎, 瓜生真也, 牧山幸史: 基礎からわかる時系列分析, pp.37-38, pp.72-151, 技術評論社, 2018.
 - 19) 野村俊一: カルマンフィルタ -R を使った時系列予測と状態空間モデル-, pp.71-112, 共立出版, 2016.
 - 20) Holmes, E. E., Ward, E. J., Wills, K.: MARSS: Multivariate Autoregressive State-space Models for Analyzing Time-series Data, https://journal.r-project.org/archive/2012-1/RJournal_2012-1_Holmes-et-al.pdf
 - 21) Holmes, E. E., Ward, E. J., Wills, K.: MARSS: Multivariate Autoregressive State-space Models for Analyzing Time-series Data, *The R Journal*, Vol.4, 19 pages, 2012.

DEMAND FORECASTING MODEL FOR RAILROAD USING ROUTE-SEARCH HISTORY DATA

Mio HOSOE, Masashi KUWANO and Taku MORIYAMA