

幹線鉄道の輸送計画策定支援に向けた 旅客需要波動の予測手法

松本 涼佑* 奥田 大樹* 深澤 紀子*

Forecasting Method of Railway Passengers' Demand Fluctuations for Long-Distance Rail Service Planning

Ryosuke MATSUMOTO Daiki OKUDA Noriko FUKASAWA

For efficiently planning long-distance rail services, the appropriate forecasting of passengers' demand fluctuation based on the historical ridership record is necessary. However, its forecast is difficult because it is a mixture of passengers' demand variations. An effective approach to an appropriate forecast should be to decompose it into several independent demand variation components and to forecast each of them. This study applies the independent component analysis to decompose the fluctuation into several additive variation components. Then, a forecasting method of the passengers' demand fluctuation is developed with the actual ridership data and the calendar structure.

キーワード：独立成分分析，時系列分析，需要波動，需要予測，輸送計画，幹線鉄道

1. はじめに

質の高い幹線鉄道の輸送計画を策定するためには、日々変化する旅客需要波動（時系列的な旅客数の変動）を、過去のデータに基づき高い精度で予測する必要がある。しかし、それらの波動は、様々な旅客の需要の波動により構成され、また暦配列等の様々な要因が複雑に絡み合うことによって変動するため、予測することは難しい。このような複雑な波動を予測するためには、いくつかの独立した波動に分解し、独立したそれぞれの波動について予測を行うことが有効と考えられる。

複数の未知の波動が混合した波動を、複数の独立した波動に分離する手法として、独立成分分析（Independent Component Analysis・ICA）がある。この手法は、大勢の人が同時に話しているときにその中から特定の人の声を抽出する、脳派を外部で観測して脳の内部に発生している信号を分離する、また画像からノイズを消去し画質を改善するなど、様々な活用方法がある¹⁾。

独立成分分析を旅客需要波動に適用した先行研究は、わが国においても複数ある^{2) 3) 4)}。それらは、広島空港に流入する自動車の定点観測データに対して、独立成分分析を適用し、欠航日特有の需要波動を抽出することによって、異常日の需要波動を明らかにすることを目標としている。しかし、幹線鉄道の利用実績データに対して独立成分分析を適用した先行研究は、筆者らの知る限り公表されていない。

そこで本研究では、ある特定の発駅（Origin）と着駅（Destination）の組合せ（以下、OD）における幹線鉄道の利用実績データから需要波動を算出し、それらに独立

成分分析を適用した。そして、分析結果に暦配列情報を組み合わせることによって、将来のある1日における需要波動を予測する手法を構築した。

2. 分析手法

2.1 独立成分分析の概要

独立成分分析は、複数の未知の波動が混合した波動を、複数の独立した波動に分離する手法であり、多変量のデータから隠された因子や成分を見つけ出すための一手法である。

標本数（観測日数） I の観測系列のうち、 i 番目（第 i 日）の標本の、 t 番目（第 t 時間帯）における観測系列を $X_i(t)$ とする。ここで $X_i(t)$ が、いくつかの独立成分系列 $S_j(t)$ と、 t によらない混合係数 A_{ij} の積和によって観測される時、全ての i について式(1)が成り立つ。なお、独立成分分析のイメージ図を図1に示す。

$$X_i(t) = \sum_j S_j(t) A_{ij} \tag{1}$$

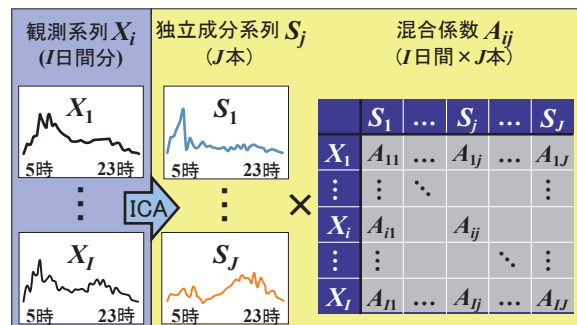


図1 独立成分分析のイメージ図

* 信号・情報技術研究部 交通計画研究室

特集：輸送・交通計画技術

独立成分分析は各*i*の観測系列 $X_i(t)$ に独立性を仮定し、 $X_i(t)$ のみの情報から、独立成分系列 $S_j(t)$ と混合係数 A_{ij} を同時推計する手法である。

なお、抽出する独立成分系列の数は、観測系列に主成分分析を適用し、分散に対する寄与率を基に観測系列の次元を縮約することによって限定することができる。

2.2 本研究での独立成分分析の適用

本節では、旅客需要波動を複数の波動に分離する手法として独立成分分析を適用する理由、および独立成分分析の計算に用いるアルゴリズムについて記述する。

観測された信号から源信号の抽出を行うブラインド信号源分離の主な手法として、独立成分分析と主成分分析がある。主成分分析では観測された信号が正規分布にしたがう仮定の下、源信号の抽出を行うが、独立成分分析では正規性の仮定を置かない。ここで、暦配列等の要因が複雑に絡み合うことによって変動する幹線鉄道の旅客需要波動が、正規分布にしたがう保証はないことが、本研究で独立成分分析を適用する理由である。

独立成分系列と混合係数を同時推計するアルゴリズムは複数種類あるが、本研究では交通量データに独立成分分析を適用した先行研究^{2) 3) 4)}に習い、Hyvärinenら¹⁾によって開発されたfastICAアルゴリズムを適用する。なお、fastICAアルゴリズムでは、前処理として観測系列の平均値を0に基準化するため、本研究における観測系列は平均値を0に基準化した後の観測系列を示す。

2.3 混合係数予測モデル

各独立成分系列に対する日付別の混合係数に関して、暦配列情報をダミー変数化したものを説明変数とし、ステップワイズ変数選択法による重回帰分析を行うことによって、混合係数予測モデルを構築する。第*i*日に関する第*k*暦配列ダミーを $C_k(i)$ とすると、第*j*独立成分系列に対する第*i*日の混合係数の予測値 A_{ij} は式(2)で表わせる。

$$A_{ij} = \alpha_j + \sum_k \beta_{jk} C_k(i) \quad (2)$$

α : 定数項
 β : 暦配列情報のパラメータ

本研究の暦配列に関する用語の定義を表1に示す。なお、混合係数予測モデルで用いた説明変数は、混合係数予測モデルのパラメータ推計結果を示す表2の1列目にある。ただし、曜日ダミーは水曜を基準としている。

3. 使用データ

3.1 概要

本研究では、ある地方都市の中心駅（以下、A 駅）から、

表1 暦配列用語の定義

用語	定義
休日	土曜・日曜・祝日・多客期間（事業会社定義によるGW・お盆・年末年始期間）
連休	3日以上休日が連続する日
連休前休日	連休ではない休日で、1週間後までに連休がある休日
連休後休日	連休ではない休日で、1週間前までに連休がある休日
単休日	前日、翌日がともに平日の休日
平日	休日以外の日
連休明け平日	前日が連休の平日
単平日	前日、翌日がともに休日の平日

ある大都市の中心駅（以下、B 駅）への幹線鉄道の利用実績データ（以下、A 駅発データ）、およびその逆方向のデータ（以下、B 駅発データ）を用いる。使用データは、2006年4月1日から2015年3月31日までの期間における旅客数を30分単位（5:00～5:29、…、23:00～23:29）で集計したデータである。

なお、当該OD間における幹線鉄道を利用した一般的な所要時間は50分である。

3.2 月次成分分解

まず、利用実績データの傾向を把握するために、日・時間帯別データを月次データへと集計し、月次成分分解を行った。月次成分分解では、月次データの原系列を、Clevelandら⁵⁾によって開発されたSTLアルゴリズムにしたがいトレンド成分、月次周期成分、ランダム成分へと分解する。トレンド成分は長期間における傾向、月次周期成分は1年周期で各月に現れる傾向、ランダム成分は原系列からトレンド成分と月次周期成分を差し引いた残差を表わす。

トレンド成分、月次周期成分をそれぞれ図2、図3に示す。なお、図2のトレンド成分の値は、2006年4月の値を1に基準化している。また、内閣府の設定⁶⁾による景気拡張期（2002年1月から2008年2月、2009年3月から2012年3月）を赤色網掛けで、景気後退期（2008年2月から2009年3月、2012年3月から2012年11月）を青色網掛けで示す。これら景気局面の設定は事後的な検証に基づいており、2012年11月以降の景気局面は公表されていない。

トレンド成分の動向について、必ずしも景気局面には左右はされず、リーマンショック（2008年9月）の前後を除いては、概ね緩やかな上昇傾向が見られる。また、A 駅発とB 駅発で概ね同じ動向であるが、B 駅発の方がやや上昇傾向にある。

月次周期成分の動向について、お盆期間が含まれ、かつ31日間ある8月に最も大きく、28日間しかなく、か

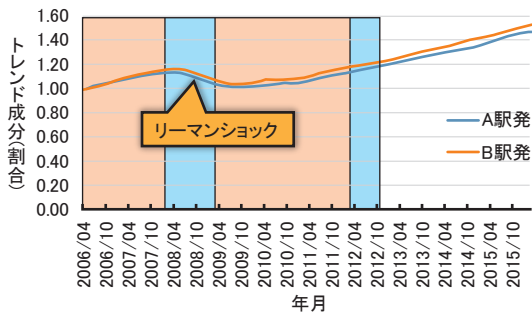


図2 トレンド成分

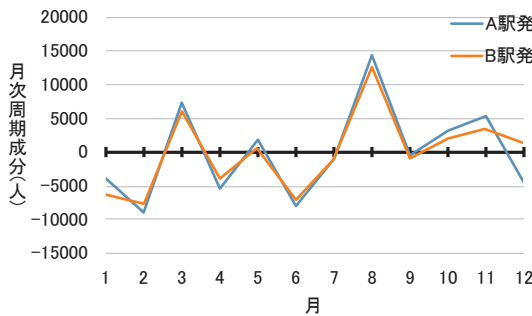


図3 月次周期成分

つ祝日が建国記念日のみの2月に最も小さい。また、A 駅発と B 駅発で概ね同じ動向であるが、年末12月は B 駅発が多く、年始1月は A 駅発が多い。これは大都市から地方都市への帰省需要が年末に大きく、地方都市から大都市への U ターン需要が年始に大きいことと整合している。

3.3 独立成分分析を適用するデータ

独立成分分析を適用するデータの期間は、最新のデータから2年間遡った、2014年4月1日から2016年3月31日とする。これはトレンド成分の上昇傾向に概ね変化がない期間であり、また3連休の構成曜日が全て同じ(土・日・月)であることから設定した期間である。

そのうち、定時運行率が90%以下であった日、輸送障害があった日、列車運休があった日、事業会社の定義によるGW・お盆・年末年始の多客期間、および2015年9月19日から23日の5連休期間を除外した599日間分のデータを対象とする。

なお、3.2のトレンド成分、および月次周期成分を用いて、観測系列に対してトレンド調整と月次周期調整を行っている。月次周期成分は、当該月の日数で割ることによって、各月の1日あたりの成分を算出できる。各日付における観測系列から、当該の月次周期成分を差し引くことによって月次周期調整を行っている。また、各観測系列を、2006年4月を1としたトレンド成分の割合で除することによってトレンド調整を行っている。

4. 分析結果

4.1 独立成分系列の抽出

A 駅発、B 駅発の観測系列に主成分分析を適用することにより、それぞれ6本、7本の独立成分系列で、全変動の80%以上を説明できることが分かった。本研究の独立成分分析では、A 駅発、B 駅発の観測系列からそれぞれ6本 ($S_{A1} \sim S_{A6}$)、7本 ($S_{B1} \sim S_{B7}$) の独立成分系列を抽出する。結果は4.3節で逐次示す。

4.2 混合係数予測モデルのパラメータ推計

各独立成分系列に対する日付別の混合係数に関して、ステップワイズ変数選択法を用いた重回帰分析により得られたパラメータ推計結果を表2に示す。なお、視認性を高めるために、一部の独立成分系列には-1の値を乗じており、対応する混合係数予測モデルの各パラメータにも-1を乗じている。

ここで、混合係数予測モデルのパラメータを解釈する方法について記述する。定数項は、全ての日付にかかる値である。基準である通常平日の水曜には、定数項の値のみがかかる。その他の曜日にかかる値は、定数項に各曜日パラメータを加えた値となる。また、対象期間における3連休の構成曜日は全て、土・日・月であることから、各3連休構成日(1・2・3日目)にかかる値は、定数項に対応する3連休構成日のパラメータと、それぞれ土・日・月のパラメータを加えることによって算出できる。

4.3 独立成分系列の解釈

4.1節、および4.2節の分析結果を基に、A 駅発、B 駅発の各独立成分系列の解釈を行う。

A 駅発の朝方に需要のピークをむかえる独立成分系列 $S_{A2} \cdot S_{A5} \cdot S_{A6}$ と、B 駅発の昼間以降に需要のピークをむかえる独立成分系列 $S_{B1} \cdot S_{B3} \cdot S_{B4} \cdot S_{B6} \cdot S_{B7}$ をそれぞれ図4、図5に示す。

A 駅発の S_{A2} 、 S_{A5} 、 S_{A6} はそれぞれ7:30、9:00、8:00に需要のピークをむかえる波動である。いずれの波動でも、曜日別では土曜に最も大きく見られ、日曜にも大きく見られる。また、3連休の1日目・2日目にも見られることから、これらは観光・行楽需要が反映されていると考えられる。加えて、 S_{A2} 、 S_{A6} は月曜から金曜においても見られるが、比較して S_{A5} で見られる絶対値の大きさは小さく、 S_{A5} は火曜から金曜において負の方向で見られる。これは早朝時間帯にピークをむかえる S_{A2} 、 S_{A6} は、観光・行楽需要に加えて、通勤や出張の需要を兼ねているが、9:00にピークをむかえる S_{A5} にはそのような需要は反映されていないと考えられる。

B 駅発の S_{B1} 、 S_{B3} 、 S_{B4} 、 S_{B6} 、 S_{B7} はそれぞれ22:00、

表2 混合係数予測モデルのパラメータ推計結果

説明変数等	A 駅発						B 駅発						
	S_{A1}	S_{A2}	S_{A3}	S_{A4}	S_{A5}	S_{A6}	S_{B1}	S_{B2}	S_{B3}	S_{B4}	S_{B5}	S_{B6}	S_{B7}
定数項	11.55	6.11	18.68	20.01	-1.43	6.56	-6.10	10.32	4.03	12.03	21.47	-4.17	-1.50
月	-4.87	4.65	-5.32	-5.44	1.69	1.19	-1.56	-0.88	-2.63	-0.97	-4.01		-0.73
火	-1.21		-1.10	-1.29				0.47	-0.76	-0.75			
木	1.05		0.83					0.83			0.59		
金	6.39	-3.29	8.64	-2.37	-3.13	-2.58	1.23		7.13	1.96	-3.43	1.41	3.32
土	7.13	11.21	-17.05	6.79	19.09	8.27	5.12	2.44	2.83	0.92	-10.27	6.89	4.28
日		2.92		-2.09	8.08	5.35	4.26	-10.35	10.54	16.74	-22.61	11.05	11.28
3 連休 1 日目				5.77	2.36		-8.19		-4.02		2.20		
3 連休 2 日目	8.67	8.59	-4.22	5.21	8.24	7.42		8.34		-2.71	8.51		-2.13
3 連休 3 日目		-8.83	13.90	6.05	4.13		3.93	-12.53	9.79	20.35	-20.31	7.22	13.50
連休前土曜			-2.17	-1.83		1.37	3.51						
連休前日曜			-1.94					1.19		-1.87			
連休後土曜		1.98		-1.33			3.98			1.26		2.32	
連休後日曜	1.26	2.40	-4.48		4.26	3.00	3.28	1.35		-1.22		2.42	3.01
単休日		4.50	-11.72	-3.16	11.99	6.97	4.62	-4.09	6.12	1.51	-16.20	9.98	10.42
連休明け平日	-3.69	4.71	-4.54	-2.86	1.81				-2.17	-1.91	-2.30		
単平日	-2.56			-5.56					-2.62	-2.96			
補正決定係数	64%	53%	82%	52%	82%	62%	29%	69%	73%	83%	86%	53%	68%

青太字：1%有意 太字：5%有意

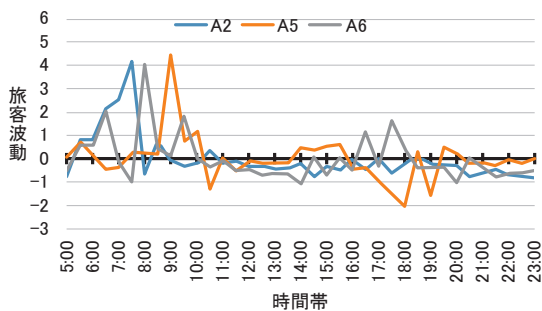


図4 A 駅発の独立成分系列 $S_{A2} \cdot S_{A5} \cdot S_{A6}$

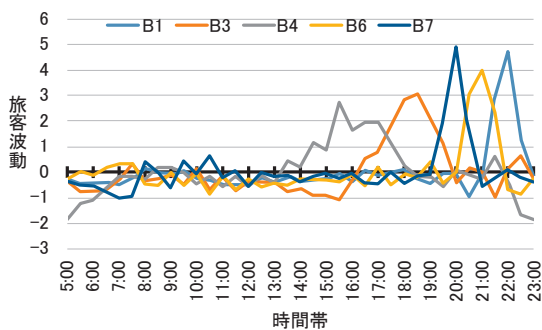


図5 B 駅発の独立成分系列 $S_{B1} \cdot S_{B3} \cdot S_{B4} \cdot S_{B6}$

18:30, 15:30, 21:00, 20:00 に需要のピークをむかえる波動である。ピークが 22:00 と最も遅い S_{B1} を除いては、曜日別では日曜に最も大きく見られる。また、同様に S_{B1} を除いては、3 連休 2 日目・3 日目にも見られることから、これらは観光・行楽の戻り需要が反映されている

と考えられる。一方で、 S_{B1} は混合係数予測モデルの決定係数の値が最も低いことから、暦配列以外の要因が影響していると考えられる。

次に B 駅発の朝方に需要のピークをむかえる独立成分系列 $S_{B2} \cdot S_{B5}$ と、A 駅発の昼間以降に需要が増加する独立成分系列 $S_{A1} \cdot S_{A3} \cdot S_{A4}$ をそれぞれ図 6, 図 7 に示す。

B 駅発の S_{B2} , S_{B5} はそれぞれ 9:00, 8:00 に需要のピークをむかえる波動である。これらの波動は、曜日別では月曜から土曜までは正の影響が見られるが、比較して日曜にはほぼ見られない。また、9:00 にピークをむかえる S_{B2} は月曜から金曜と比較して土曜にやや大きく見られ、8:00 にピークをむかえる S_{B5} は土曜と比較して月曜から金曜により大きく見られる。これらから平日における通勤や出張の需要が、土曜の観光・行楽需要よりも早い時間帯に多いことが影響していると考えられる。

A 駅発の S_{A3} は 15:00 から 20:00 の時間帯に需要が増加する波動である。曜日別では金曜に最も大きく見られ、月曜から木曜と日曜にも見られるが、土曜には見られない。金曜に特に大きく見られる逆方向 (B 駅発) の朝方や昼間の需要がないため、 S_{A3} が金曜に大きく見られるのは単身赴任者の居住地側への戻り需要や、泊りがけ出張の帰り需要が反映されていると考えられる。なお、月曜から木曜に見られるのは通勤や出張の帰り需要、日曜に見られるのは観光・行楽の帰り需要と考えられる。

A 駅発の S_{A1} は昼前と夕方以降の時間帯に需要が増加する波動であり、 S_{A4} は朝方から夕方までの時間帯に需要が増加する波動である。これらの波動が見られるのは、

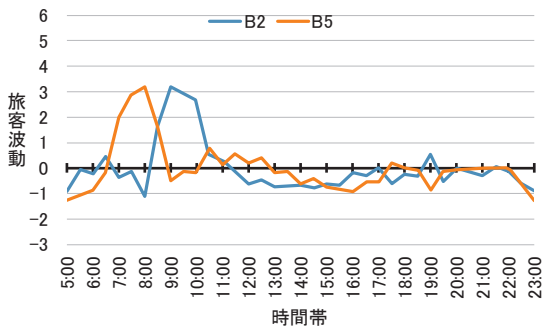


図6 B 駅発の独立成分系列 $S_{B2} \cdot S_{B5}$

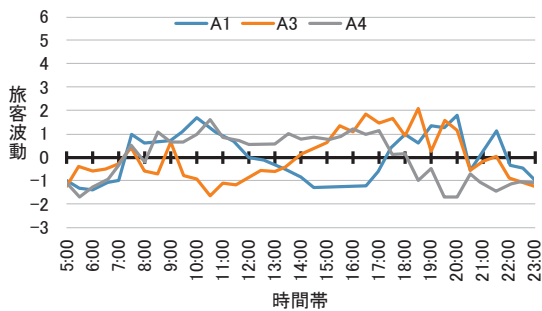


図7 A 駅発の独立成分系列 $S_{A1} \cdot S_{A3} \cdot S_{A5}$

曜日別に多少の大小はあるが、いずれの曜日においても正の影響が見られる。したがってこれらは、いずれの日においても見られる基礎的な波動であると考えられる。

5. 現況再現性

5.1 現況再現性の確認

現況再現性を確認するために、各日付における観測系列と、観測系列を再現した系列（以下、再現系列）の間に、2標本コルモゴロフ-スミノルフ検定（以下、2標本KS検定）、および相関分析を行う。2標本KS検定は、2つの母集団の確率密度分布が同じものであるという帰無仮説を、母集団の確率密度分布に仮定を設けずに検証するものである。ただし、KS検定では時間軸上で集計した確率密度分布の差を検定するため、検定する2つの系列について、ピークの発生時間が異なるが、その大きさや発生時間の長さが似ている場合、それらの系列に差はないと判断する恐れがある²⁾。このことから、本研究ではKS検定を行う2系列に、相関分析を併用することで、現況再現性の確認を行う。具体的には、KS検定ではピークの大きさの類似性を検定し、相関分析ではピークの発生時間帯の類似性を検定する。

KS検定結果のP値、および相関係数の度数分布表をそれぞれ表3、表4に示す。なお、全ての日付において、有意水準5%で、相関係数の値は有意であった。

KS検定結果から、有意水準5%で、A駅発、B駅発それぞれの約70%、約65%を再現できていることが分

かった。

相関分析結果から、A駅発、B駅発それぞれの約85%、約75%の日付で0.9以上、それぞれ99%、96%の日付で0.8以上の相関係数の値が得られていることが分かった。

以上から、暦配列に関する情報のみを加味したモデルを用いて、一定水準の再現性を確保できることが分かった。また、大都市発のB駅発と比較して、地方都市発のA駅発の方が、やや再現性が高いことが分かった。

それらに加えて、KS検定の結果よりも、相関分析の結果の方が概ね良好な値を得られていることから、暦配列に関する情報のみを用いたモデルでは、ピークの大きさの予測と比較して、ピークの発生時間帯の予測の精度の方が高いことが分かった。

5.2 再現誤差の考察

ある土曜の日付、第*i*日における、A駅発、B駅発の観測系列と再現系列を、それぞれ図8、図9に示す。なお、KS検定のP値はそれぞれ0.144、0.01未満であり、相関係数はそれぞれ0.948、0.808である。

A駅発では概ね再現できているが、B駅発では20:30から21:30の時間帯に再現できていない需要の増加が見られる。これはB駅付近の会場で17:00に開演した人気アイドルグループのイベントが影響したものと考えられる。

次に、第*i*+1日の日曜における、A駅発、B駅発の観測系列と再現系列を、それぞれ図10、図11に示す。なお、KS検定のP値は両方とも0.01未満であり、相関係数の値はそれぞれ0.795、0.952である。

第*i*+1日においても、同様の人気アイドルグループが、

表3 KS検定結果のP値の度数分布表

P値		A 駅発		B 駅発	
以下	以上	相対 度数	累積相対 度数	相対 度数	累積相対 度数
1.00	0.10	61%	61%	58%	58%
0.10	0.05	7%	69%	5%	64%
0.05	0.01	11%	80%	13%	76%
0.01	0.00	20%	100%	24%	100%

表4 相関係数の度数分布表

相関係数		A 駅発		B 駅発	
以下	以上	相対 度数	累積相対 度数	相対 度数	累積相対 度数
1.0	0.9	84%	84%	73%	73%
0.9	0.8	15%	99%	23%	96%
0.8	0.7	1%	100%	3%	99%
0.7	0.0	0%	100%	1%	100%

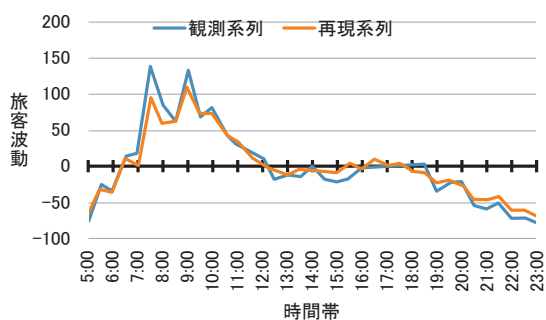


図8 第*i*日におけるA駅発の観測系列と再現系列

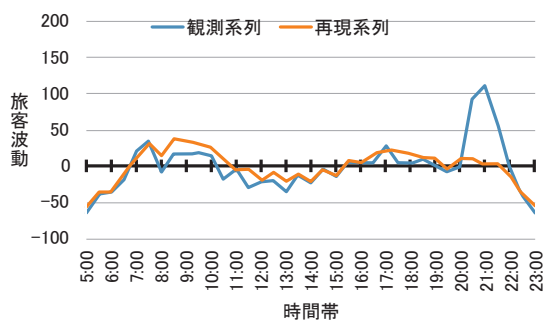


図9 第*i*日におけるB駅発の観測系列と再現系列

11:00開演と17:00開演のイベントを行っている。A駅発では、8:00前の早朝時間帯において、再現できていない需要の増加が見られるが、これらは11:00開演のイベントが影響していると考えられる。また、B駅発での14:30から17:00にかけての需要の増加と、20:30から21:30にかけての需要の増加も、人気アイドルグループのイベントの帰り需要と考えられる。

以上から、暦配列に加えた、イベントの開催日と時間の情報を加味することの必要性が示唆された。また、イベント内容によって、当然影響の大きさも変わると考えられる。そのため、イベント情報の分類をいかに行うかも今後の課題である。

6. おわりに

本研究では、ある地方都市の中心駅と、ある大都市の中心駅間のODを対象に、独立成分分析と重回帰分析に暦配列情報を組み合わせることによって、旅客需要波動を予測するモデルを構築した。本研究を通して、暦配列情報を基にしたモデルで、一定水準の旅客需要波動の再現性を確保できることが分かった。

一方で、暦配列情報以外のイベントに関する情報が、旅客需要波動に影響を与えていると考察できた。した

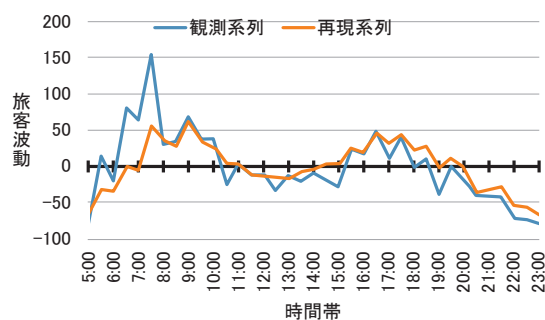


図10 第*i+1*日におけるA駅発の観測系列と再現系列

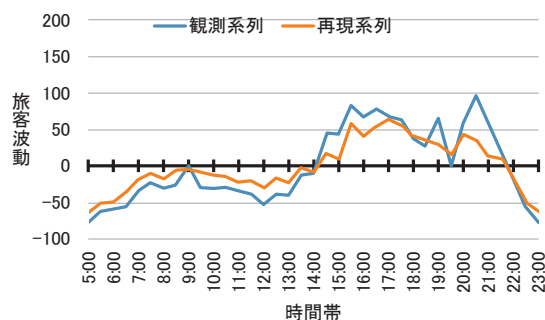


図11 第*i+1*日におけるB駅発の観測系列と再現系列

がって、暦配列情報に加えて、イベントの開催時間や観客動員数、駅からイベント会場までの距離等の情報を混合係数予測モデルに加えることが今後の課題である。

文献

- 1) Hyvärinen, A. et al., 根本他訳：【詳解】独立成分分析，東京電機大学出版局，2005
- 2) 井上真一他：独立成分分析による非日常時の地方空港利用交通特性の分析，土木計画学研究・論文集，Vol.26，No.3，pp.583-593，2009
- 3) 井上真一他：広島空港における連続観測断面交通量に基づく非日常交通特性の分析，土木計画学研究・講演集，Vol.37，No.54，2008
- 4) 塚井誠人他：データマイニングによる広島空港流入車両の時間分布パターン分析，交通工学研究発表会論文報告集，Vol.28，pp.269-272，2008
- 5) Cleveland, R. et al., "STL: A Seasonal-Trend Decomposition Procedure Based on Loess" *Journal of Official Statistics*, Vol.6, No.1, pp.3-73, 1990.
- 6) 内閣府：景気基準日付
<http://www.esri.cao.go.jp/jp/stat/di/150724hiduke.html>
 (閲覧日：2017年6月21日)