

時系列変化を考慮した道の駅の レジカウト人数と twitter 件数の関係

羽生 健士郎¹・寺部 慎太郎²・柳沼 秀樹³・海野 遥香⁴・鈴木 雄⁵

¹非会員 東京理科大学大学院 理工学研究科土木工学専攻 (〒278-8501 千葉県野田市山崎2641)
E-mail:7621522@ed.tus.ac.jp (Corresponding Author)

²正会員 東京理科大学教授 理工学部土木工学科 (〒278-8501 千葉県野田市山崎2641)
E-mail:terabe@rs.tus.ac.jp

³正会員 東京理科大学准教授 理工学部土木工学科 (〒278-8501 千葉県野田市山崎2641)
E-mail:yaginuma@rs.tus.ac.jp

⁴正会員 東京理科大学助教 理工学部土木工学科 (〒278-8501 千葉県野田市山崎2641)
E-mail:unoharuka@rs.tus.ac.jp

⁵正会員 東京理科大学助教 理工学部土木工学科 (〒278-8501 千葉県野田市山崎2641)
E-mail:yusuzuki@rs.tus.ac.jp

多様化した情報社会である現在において、SNS が観光行動に与える影響は無視できない。そこで、多くの自治体や道の駅は PR や更なる発展を目指し、SNS を積極的に活用するようになった。しかし、SNS の投稿がどれくらい集客に影響するかについて研究の余地がある。そこで本研究では、3 年強と 1 年分のデータ提供を頂いた 2 つの道の駅を対象に、道の駅の買物来訪者数を目的変数に twitter や新聞記事データなどを説明変数とした時系列回帰分析を行った。その結果、新聞掲載数では有意にはならなかったものの、twitter 投稿数では 2 つの道の駅ともに 1% 有意で、ツイートの件数が 1 増加すると買物来訪者数がそれぞれ約 1 人増加、約 6 人増加という結果が得られた。

Key Words: SNS, Social Networking Service Regression analysis, Roadside Station, Regional development

1. 序論

(1) 研究の背景

ICT 総研¹⁾による SNS (ソーシャル・ネットワーキング・サービス) の利用者数と普及率によると、日本国内における SNS の利用者は年々増加しており、2020 年度には 7975 万人に達し、SNS の利用者は 80.3% と高い結果となっている。このソーシャルメディアユーザーの増加により、SNS 上のクチコミがマーケティングに大きな影響を及ぼすようになってきている。そのため企業は、自社商品に関するクチコミを最大化させ、より多くのユーザーの目に触れさせることで購買意欲を掻き立てている。しかし、クチコミ数の最大化を目指す上でも、再現性のある施策展開を行わなければ一過性の成果に止まってしまう。そのため、クチコミ数を計測しながらどんな施策が有効だったかを仮説検証し、有効性のある施策を打ち出していくことが、マーケティングにおいて重要になっ

てきている。

次に SNS (ソーシャル・ネットワーキング・サービス) が主流になった昨今では従来とは異なり、自治体やメディアが一方向的に情報を発信するのではなく、誰もが情報を発信・受信できるようになった。SNS の利用者数や普及率の増加により、クチコミがマスメディアと同じくらいの影響力を持つようになり、多様化する情報社会において、インターネット上の情報を加味した分析が重要となっている。

さて、全国で 2022 年 8 月 5 日までに 1198 駅が登録されている道の駅には「休憩機能」に加え、道路情報、地域の観光情報、緊急医療情報などの提供を行う「情報発信機能」と文化教養施設、観光レクリエーション施設などの地域振興施設などの側面としての「地域連携機能」が求められている。したがって、道の駅には、当該地域の観光拠点として期待が寄せられている。

しかし、筆者らがいくつかの道の駅の担当者にイン

タビュールしたところ、少数ではあるが、道の駅の情報発信に SNS を重視していない、あるいは、来訪者数増減に SNS が影響していないという返答があった。

そこで本研究は、道の駅の来訪者数に、人々の興味・関心としての SNS 投稿数や新聞記事が少なからず影響していると考え、これを定量的に示すことを目的とし、時系列分析を行う^{2,3,4)}。さらに、筆者らによる先行研究⁹⁾を踏まえ、新聞報道件数の影響と、興味・関心が累積することの影響とを考慮した分析を行う。

これにより、今まででなされなかった SNS やメディア上での道の駅の注目度合いや SNS の投稿数の影響度を測ることが出来、それらが利用者数にどのような影響を与えているのかを示すことが期待できる。

2. 分析手法の概要

本研究の分析手法は時系列分析で、道の駅買物来訪者数と twitter 投稿数、新聞掲載数の関係を回帰分析する。

まず、モデル構築までは、単位根検定と自己相関に関する検定が必要である^{2,3)}。

単位根検定は、時系列データが定常過程かどうかを調べるものである。時点によって平均などの確率分布が変化しない確率過程を定常過程という。定常過程であれば、多くの時系列分析手法が適用できる。単位根過程の変数をそれとは独立の単位根過程の変数に回帰すると、ほぼ確実に有意な関係があらわれ回帰の説明量が高いように見える「見せかけの回帰」と呼ばれる現象があるため注意が必要である。単位根検定は、ADF 検定（拡張ディッキー・フラー検定）が最も一般的で、検定統計量を算出し、単位根過程であるとする帰無仮説が棄却できるかどうかを調べる。

自己相関に関する検定は、回帰分析の際に、通常最小二乗法を用いるか、一般化最小二乗法を用いるかを判断するために必要である。通常最小二乗法は、残差の共分散が 0 であることを前提としているため、残差に系列相関がある場合には、用いることができない。そのため、自己相関があるかどうかを判定する必要がある。その検定は、DW 検定（ダービン・ワトソン検定）が使われ、DW 比（ダービン・ワトソン比）を検定統計量とする。DW 比は 0 から 4 の値を取り、2 であれば系列相関がないと判定される。

以上の検定を終え、時系列データの様子が確認できれば、回帰分析を行う。検定の結果に応じて、通常最小二乗法を用いるか、一般化最小二乗法を用いるかを判断する以外は、一般的な回帰分析と同様である。

上記の流れに沿って同様に単位根検定を行った結果、道の駅 A、B の両駅において、累積 3 カ月をしたデータが単位根であると判定された。単位根と判定された場合

は、対数を取るか差分を取るかが一般的である。本研究では対数を取ったデータでは単位根であり、差分を取ったデータであると単位根ではないと判定された。そこで整合性を取るために、全ての回帰分析において 1 階差分を取ったデータを用いることとする。次に DW 検定を行った結果、自己相関はないと判断された。以上のことから本研究では、通常最小二乗法を用いた回帰分析 (OLS) を用いた結果を掲載する。

3. データ説明

(1) 目的変数

分析に用いるデータは、まず 2 つの道の駅（名称は非公表）から提供いただいたレジ通過人数である。レジ通過人数とは、道の駅の物販施設の会計レジに記録される買い物客数のことで、1 回の会計を 1 人と計数している。道の駅の来訪者は物販施設で買い物をしない場合もあるし、複数人で買い物をしても 1 回の会計を 1 人とみなしているため、厳密には道の駅の来訪者数＝道の駅の買物来訪者数＝レジ通過人数とは言えないものの、道の駅の来訪者数の代理指標として使われることが多い。本研究では、レジ通過人数を道の駅の買物来訪者数とみなす。また、準備したデータの概要を表-1 に示す。

(2) 説明変数

本研究で使用した説明変数について以下に記述する。

表-1 データの概要

項目		道の駅 A	道の駅 B
対象期間	開始日	2018/5/1	2020/4/1
	終了日	2021/7/26	2021/3/31
	日数	1183	365
レジ通過人数(人)	平均	619.3	1418.2
	標準偏差	211.0	573.7
	最大	1503	3081
	最小	0	0
Twitter 投稿数(件)	平均	6.9	5.2
	標準偏差	12.2	6.1
	最大	188	48
	最小	0	0
新聞記事件数(件)	平均	0.048	0.19
	標準偏差	0.0071	0.026
	最大	3	4
	最小	0	0

a) Twitter 件数 (累積なし, 1 週間, 3 カ月を準備)

道の駅名が含まれた twitter の投稿数はレジ通過人数に影響を与えていると仮定し, 1 日ごとの twitter の投稿件数を変数とした。さらに, 数日後に twitter を見て来訪するというような時間差が生じることが考えられるため, 本研究では前 1 週間(7 日間)と前 3 カ月(90 日間)の件数を累積したデータを加え, tweet の累積件数も考慮した。tweet の変数は, twitter API⁹⁾を用いてデータを抽出したものである。

b) 新聞掲載数 (累積なし, 1 週間, 3 カ月を準備)

道の駅名が掲載された新聞記事が人々の目に留まることでレジ通過人数に影響を与えていると仮定し, 変数に加える。新聞を直接読むことがなくても, 何らかの話題になる=新聞で報道される, と考え, 1 日ごとの新聞記

事の件数を変数にした。この変数も twitter と同様, 新聞掲載日と来訪日には時間差があり得ると考え, 前 1 週間(7 日間)と前 3 カ月(90 日間)の件数を累積したデータも用いている。なお, この変数は, 新聞記事データベース G-search⁷⁾を用いて全国紙, 地方紙問わず日毎の新聞記事数を変数としている。なお, データベースで検索する際には, 道の駅の正式名称を用いた。

c) 全国紙ダミー

全国紙・地方紙ではレジ通過人数に与える影響を同等に考えることが出来ないと判断し, 朝日新聞・毎日新聞・読売新聞の 3 社が全国向けに発信した新聞記事があった日をダミー変数として加えている。

d) 写真有りダミー

投稿されている新聞記事に写真が掲載されているかど

表-2 道の駅 A の最小二乗法による回帰分析の結果

説明変数	累積無し			累積 1 週間			累積 3 カ月		
	係数	P 値	判定	係数	P 値	判定	係数	P 値	判定
切片	-0.268	0.924		-0.0900	0.973		0.0322	0.991	
Twitter	1.04	0.000	***	-0.158	0.334		-0.232	0.155	
新聞記事	-11.7	0.181		0.804	0.922		-1.33	0.884	
写真ダミー	15.6	0.646		9.63	0.764		23.4	0.500	
全国紙ダミー	7.90	0.744		-1.04	0.962		-9.20	0.699	
土曜ダミー	225	0.000	***	228	0.000	***	226	0.000	***
日曜ダミー	350	0.000	***	352	0.000	***	351	0.000	***
祝日ダミー	235	0.000	***	225	0.000	***	223	0.000	***
定休日ダミー	-536	0.000	***	-567	0.000	***	-571	0.000	***
決定係数	0.813			0.827			0.822		

***1%有意

表-3 道の駅 B の最小二乗法による回帰分析の結果

説明変数	累積無し			累積 1 週間			累積 3 カ月		
	係数	P 値	判定	係数	P 値	判定	係数	P 値	判定
切片	1.85	0.909		1.48	0.929		3.63	0.850	
Twitter	6.01	0.00682	***	-1.69	0.412		-1.59	0.462	
新聞記事	7.58	0.781		9.01	0.710		30.4	0.287	
写真ダミー	-161	0.298		-144	0.359		-87.6	0.632	
全国紙ダミー	40.6	0.752		45.8	0.721		1.56	0.992	
土曜ダミー	615	0.000	***	635	0.000	***	697	0.000	***
日曜ダミー	944	0.000	***	972	0.000	***	1051	0.000	***
祝日ダミー	597	0.000	***	615	0.000	***	727	0.000	***
定休日ダミー	-869	0.000	***	-854	0.000	***	-885	0.000	***
決定係数	0.756			0.752			0.776		

***1%有意

うかによって新聞記事の注目度が変化し、それにより来訪者数が変化すると考え、変数として加えている。

e) 土日祝日ダミー

土日祝日の固定効果を表すために、この変数を用いる。また、土曜・日曜・祝日に分けてダミー化することで、曜日ごとの来訪者数が統計的にどの程度影響を与えているかを測ることが可能であると考えている。

f) 定休日ダミー

日次のデータを用いているため、道の駅の定休日と営業日が同一で考えられてしまう。そこで、営業日と定休日を区別するために、この変数を用いる。

4. 最小二乗法による回帰分析の結果

3章で記述した目的変数・説明変数を用いた重回帰分析の結果を表-2、表-3に示す。

まず、道の駅 A, B に関して累積をしていない twitter 件数によって前者は約 1 人増加、後者は約 6 人増加するという結果が統計的に有意であった。しかし、累積 1 週間・累積 3 カ月の結果から、来訪者は twitter 投稿を見てから時間差で来訪するという仮説は、統計的に有意でないという結果が得られた。

次に新聞記事件数では、係数に着目すると両駅共に、累積なしの結果より、累積したほうが係数が大きくなるという結果は得られたが 5% 有意を基準とすると統計的に有意ではなかった。

このように Twitter・新聞掲載数が仮説通りにならなかった原因としては、道の駅名が含まれた投稿として調査をしたため、投稿された内容まで把握しきれていないことが挙げられる。例えばイベントの開催やキャンペーンを報じるものであれば来訪者数が増加する要因になり得るが、当該道の駅近辺で事件が起こったなどの負の要素を取り切れていないことも原因の可能性がある。

土日祝日ダミーは全て統計的に有意であり、それぞれ土日祝であると、道の駅 A では、約 220 人～約 350 人増加し、道の駅 B では、約 600 人～約 1000 人程度増加するという結果が得られた。

5. カルマンフィルタを用いた時系列回帰モデル⁴⁾

(1) 状態空間モデル

前章では、データを全体的に見た平均の値での回帰モデルを行ったが、本章では、日ごとの twitter 件数がレジ通過人数に与える影響を視覚化することを目的とする。そこで、状態空間モデルを用いる。これは、観測値間の関連性を観測変数同士で直接表現する ARIMA モデル（自己回帰移動平均モデル）とは対照的なものである。

ARIMA モデルは観測値間の関連性を、基本的には観測値同士で直接表現するものである。

一方、状態空間モデルは、観測値間の関連性を状態経路で間接的に表現し、ある時点のデータと別時点のデータの間の関連性を生成する要因を分析者が極力想定して、潜在変数を考慮することで分析を行える。

つまり、状態空間モデルを用いることは、状態という潜在変数を導入したことによって、解釈に都合のいい状態を組み合わせて、複雑なモデルを構築することが容易になり、モデル化の柔軟性が向上する。

(2) カルマンフィルタリングと平滑化

カルマンフィルタとは、非常過程も問題なく扱うことが可能であり、線形・ガウス型状態空間モデルにおいて、推定すべきデータの真値と点推定値の間の平均 2 乗誤差を最小にする意味で最適な逐次推定法である。そして、フィルタリング分布は式(1)で表される。

$$p(x_t|y_{1:t}) = p(x_t|y_{1:t-1}) \frac{p(y_t|x_t)}{p(y_t|y_{1:t-1})} \quad (1)$$

次に平滑化分布を式(2)に示す。

$$p(x_t|y_{1:T}) = p(x_t|y_{1:t}) \int \frac{p(x_{t+1}|x_t)}{p(x_{t+1}|y_{1:t})} p(x_{t+1}|y_{1:T}) dx_{t+1} \quad (2)$$

これは、フィルタリング分布を補正したものであり、補正の程度は一期先における平滑化分布に基づく。ここで、時点 t での平滑化分布はフィルタリング分布に等しく、時点 t-1 での平滑化分布は、時点 t-1 でのフィルタリングを時点 t での平滑化分布を踏まえて補正することで得られる。

つまりカルマンフィルタの考え方は、当期の観測値を使って、当期の状態推定値を修正するというものであり、本論文では、カルマンフィルタに関するフィルタリング・平滑化を中心に行っていく。

6. 状態空間モデルによる分析の結果

道の駅 A, B を対象に 3 章で記述したレジ通過人数を目的変数に、twitter 件数と 4 章で有意な結果が得られた土日祝ダミー・定休日ダミーを説明変数として時系列重回帰分析を行った結果を図-1、図-2 に示す。この図は、twitter の回帰係数の遷移を示したものである。そのため、曲線が山になっている部分は twitter に影響されてレジ通過人数が増加した日、谷になっている部分は、twitter によってレジ通過人数が減少した日という解釈を行うことが出来る。

黒の実線はレジ通過人数と twitter 投稿件数で単回帰分析をしたものであり、青の実線はこれに土日祝ダミーと定休日ダミーを加えた重回帰分析の結果である。このグラフから、twitter 件数を累積していない左図から累積を

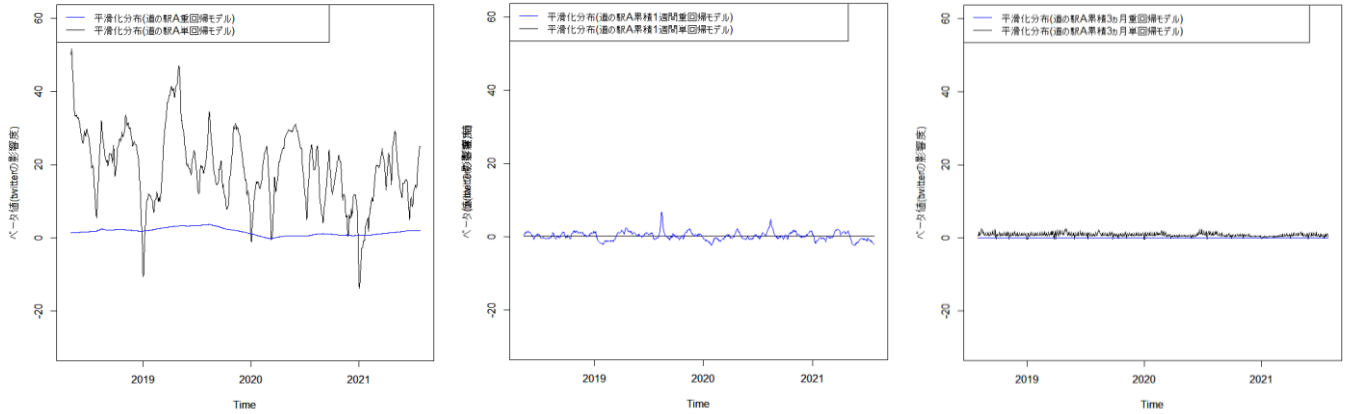


図-1 道の駅 A の twitter が来訪者数に与える影響

累積無し(左図), 累積 1 週間(中央図), 累積 3 カ月(右図)

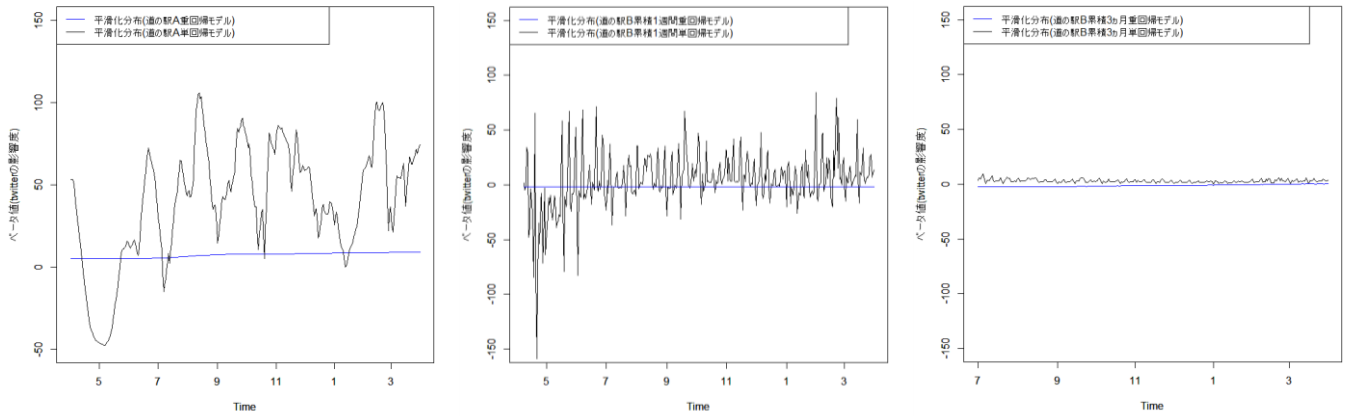


図-2 道の駅 B の twitter が来訪者数に与える影響

累積無し(左図), 累積 1 週間(中央図), 累積 3 カ月(右図)

している右図にかけて波の変動が小さくなっていることが分かる。このため、twitter 投稿件数は発信してから期間が短い程有効な手段である可能性が考えられる。次に、青の実線を見ると黒の単回帰分析でのグラフと比べてグラフの変動が少なく一定であることから、twitter 件数が来訪者数に与える影響は大きくないと考えられ、ほとんどは土日祝ダミー、定休日ダミーによって説明されているという結果となった。

7. 結論

本研究では、道の駅に関する twitter 件数や新聞掲載数が道の駅の買物来訪者数推移に与える影響を調べるため、日次データをもとに時系列回帰分析を行った。重回帰分析を行った結果から、twitter 件数による道の駅買い物客数は道の駅 A では約 1 名、道の駅 B では約 6 名増加するという結果が得られた。しかし、筆者らが実際に道の駅

の担当者に話を聞いた、道の駅のことがテレビや新聞で放送され、その内容が期間を空けて新聞や雑誌に掲載され、遅れて twitter 件数や問い合わせ数が増加し、買物来訪者数が増加するという時間差を考慮した結果をうまく表現することが出来なかった。

次にカルマン平滑化を用いて、日次ごとの回帰係数をグラフ化することで twitter 件数は、投稿から時間が経つにつれ、道の駅の買物来訪者数に与える影響が小さくなるという結果が得られた。また、図-1 道の駅 A 累積 1 週間重回帰分析結果グラフの 2019 年下旬のようにグラフが山や谷になっているような twitter の影響度が特に大きい日を視覚的に判断できるグラフを作成できた。そのため、日次ごとの特徴を把握することが出来たことは本研究の 1 つの成果である。

本研究の限界点は、データの都合上、twitter 投稿者の属性や掲載内容まで考慮できていないことが挙げられる。そのため、上記の限界点を考慮に入れた分析を行うこと

でさらに精度の高い分析を行うことが可能であると考え
る。

謝辞：本研究は、新道路技術会議平成 31 年度道路政策
の質の向上に資する技術研究開発（研究テーマ：交通・
物流・交流・防災拠点としての道の駅の性能照査と多目
的最適配置に関する研究）の研究の一環として実施され
たものである。

REFERENCES

- 1) ICT 総研：https://ictr.co.jp/report/20150729000088-2.html, 2022 年 9
月 24 日閲覧
- 2) 村尾博：R で学ぶ VAR 実証分析：時系列分析の基礎から予
測まで，オーム社，2019.
- 3) 小森政嗣：人の行動を捉える！—現代時系列分析入門
—，日本行動計量学会第 23 回春の合宿セミナー配布資
料，2021.
- 4) 萩原純一郎，瓜生真也，牧原幸史：基礎からわかる時系
列分析_R で実践するカルマンフィルタ・MCMC・粒子フ
ィルタ，pp.66-177，株式会社技術評論社，2018.
- 5) 大西一真，寺部慎太郎，柳沼秀樹，海野遥香（2022）：
道の駅買物来訪者数とインターネット検索数や SNS 投稿
数の関係，第 42 回交通工学研究発表会論文集，34.
- 6) Twitter: Twitter API Documentations, https://developer.twit-
ter.com/en/docs, 2022 年 1 月 30 日閲覧.
- 7) 新聞記事データベース G-search, https://db.g-search.or.jp/,
2022 年 8 月 30 日閲覧.

(?)

(?)

RELATIONSHIP BETWEEN THE NUMBER OF PEOPLE COUNTED AT THE CASH REGISTERS OF ROADSIDE STATIONS AND THE NUMBER OF TWITTER COUNTS, TAKING INTO ACCOUNT TIME SERIES CHANGES

Kenshiro HANYU, Shintaro TERABE, Hideki YAGINUMA,
Haruka UNO, and Yu Suzuki

In today's diversified information society, the influence of SNS on tourism behaviour cannot be ignored. Therefore, many municipalities and roadside stations have started to actively use SNS for publicity and further development. However, there is still room for research on the extent to which SNS postings influence the attraction of visitors. In this study, a time-series regression analysis was conducted using the number of visitors to a roadside station as the objective variable and twitter and newspaper article data as explanatory variables for two roadside stations that provided data for a little over three years and one year respectively. The results showed that although the number of newspaper articles did not reach significance, the number of twitter posts was significant at 1% for both of the two roadside stations, and an increase of one tweet led to an increase of approximately one and six shoppers respectively.