

# 同一便と同一バス停における乗客数の変動特性を考慮した LSTMによる路線バス乗客数予測

山村 竜也<sup>1</sup> 新井 イスマイル<sup>2</sup> 遠藤 新<sup>2</sup> 垣内 正年<sup>2</sup> 藤川 和利<sup>2</sup>

**概要:** 代表的な公共交通機関の1つとして、路線バスがある。利用者がバスに感じる不満として、車内が混雑していることが挙げられる。また、COVID-19の防止策として、混雑環境を避けることが利用者に求められている。このような背景からバスの車内混雑状況を予測し、利用者に提供する必要がある。先行研究では、Random Forestを用いて次のバス停出発時の車内混雑度の予測が行われた。しかしながら、特徴量とモデルが十分ではないため十分な精度が出ない結果となっている。そこで本研究では、特徴量を先行研究で使用されていたものだけでなく、N便前までのバス停出発時の車内人数も用い、モデルをRandom ForestからLong Short-Term Memoryに変更して車内人数を予測した。先行研究であるArabghaliziらの手法をRandom Forestの回帰によって、バス停出発時の車内人数を予測する手法に変更したものと提案手法を、Root Mean Square Errorを評価指標として比較した。結果、既存手法よりも提案手法の方が精度が最大で29%向上した。

## LSTM based bus ridership prediction considering the transition characteristics of passengers grouped by a trip and a bus stop

TATSUYA YAMAMURA<sup>1</sup> ISMAIL ARAI<sup>2</sup> ARATA ENDO<sup>2</sup> MASATOSHI KAKIUCHI<sup>2</sup>  
KAZUTOSHI FUJIKAWA<sup>2</sup>

### 1. はじめに

代表的な公共交通機関の1つとして、路線バス（以下、バス）が挙げられる。2016年に内閣府が行った公共交通に関する世論調査によると、バスを利用する際に「不満に思うことがある」と回答した人の割合は68.3%で、その中でも「車内が混雑している」ことに最も不満を感じると回答した人の割合は15.7%だった[1]。また、COVID-19の防止策として、混雑環境を避けることが利用者に求められている。さらに、利用者が混雑環境を避けることを促進するサービスも登場している。例としては、Google社の提供するGoogle Mapが挙げられる。各公共交通機関は世界共通の公共交通機関向けのオープンフォーマットであるGeneral Transit Feed Specification Realtime (GTFS Realtime) [2]を利用して、Google Mapを通じて現在のバスの車内混雑情報をバス利用者に提供できる。こう

した背景から、現在のバスの車内人数や、未来のバスの車内人数を予測し、利用者に提示する必要がある。

バスの車内混雑状況を予測する既存研究としては、Arabghaliziらの研究[3]がある。Arabghaliziらの研究では、同一便における過去のバス停出発時の車内人数を特徴量として利用して、次のバス停出発時の車内人数を予測する手法が提案された。この既存研究から着想を得て、我々は、同一バス停出発時における過去の車内人数の推移も特徴量として利用することで、次のバス停出発時の車内人数をより高い精度で予測できるという仮説を立てた。

そこで本研究では、まず、この仮説の事前検証として、同一バス停におけるバス停出発時の車内人数の時系列データにおける自己相関を、実際の路線バスから収集したデータを利用して調査する。その後、検証の結果に基づいて、同一便と同一バス停におけるバス停出発時の車内人数の時系列データ(図1)を利用して、より高い精度で車内人数を予測する手法の実現を目指す。

本論文の構成は以下の通りである。第2章では、現在の

<sup>1</sup> 奈良先端科学技術大学院大学 先端科学技術研究科  
<sup>2</sup> 奈良先端科学技術大学院大学 総合情報基盤センター

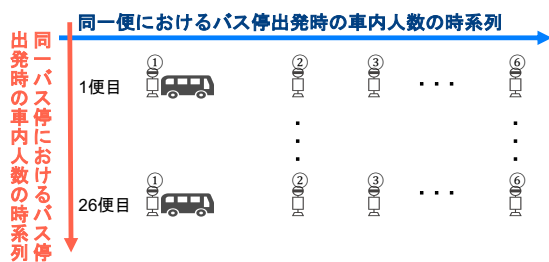


図 1: 同一便と同一バス停におけるバス停出発時の車内人数の時系列

バスの車内人数を推定する研究, 未来のバスの車内人数や混雑度を予測する研究を紹介し, その後, 提案手法の着想を得た関連研究を説明する. 第 3 章では, 本研究で利用するデータについて述べ, 仮説に関する事前検証をする. 第 4 章では本研究の提案手法について述べる. 第 5 章では, 提案手法によるバス車内人数の予測精度について評価するために実施した実験について説明する. 第 6 章では, 評価結果とその結果に対する考察や課題を述べる. 最後に第 7 章では, 本論文の結論と今後の展望について述べる.

## 2. 関連研究

バスの車内人数や混雑度に関する研究は, 大別すると, 現在のバスの車内人数を推定する研究 [4][5][6][7][8][9] と, 未来のバスの車内人数や混雑度を予測する研究 [3][10][11][12] である. 現在のバスの車内人数を推定する研究として, 中島らは, ドライブレコーダに記録された映像データ, バスに搭載されている GPS モジュールから得られるバスの位置データを用いて, バスの車内人数を推定した [4]. 映像データのみを用いて YOLOv3 で推定し, その後, バスの位置データを組み合わせて Random Forest (RF) で推定した. また, 金光らは, バス車内の乗客の所有する携帯端末から発せられる BLE 信号に含まれるデータを用い, XGBoost で車内人数を推定した [5]. 未来のバスの車内人数や混雑度を予測する研究として, 青地らは, 曜日, 路線番号, 時間帯, 降雨の有無, 降水量, 乗車人数, 降車人数を特徴量として, バス停毎に RF を用いて予測した [10][11]. また, Wood らは, アメリカを走るループ状の路線を走行するバスにおいて, ヘッドウェイ, 予測したい時点におけるバス車内の混雑度, 天気, 時間帯を特徴量として, 線形回帰と RF を用いて予測した [12]. また, Arabghalizi らは, 月, 曜日, 気温, 降水量, 降雪量, 何便目か, バスの種類, 同一便における  $N$  個前までのバス停出発時の車内人数を特徴量として, バス停毎に RF を用いて予測した [3].

提案手法の着想を得た研究は, この Arabghalizi らの研究である. Arabghalizi らが複数路線で実施した評価結果によると, Arabghalizi らの手法の F1-score は, 5 クラス分類という車内人数の予測と比べて低難易度のタスクにも関わらず, 平均 85.9% となり, 100% に達しなかった. この

結果の要因の 1 つとしては, 同一便における  $N$  個前までのバス停出発時の車内人数を特徴量として利用することで, 同一便におけるバス停出発時の車内人数のデータに変動特性があることは考慮しているが, 同一バス停における  $N$  便前までのバス停出発時の車内人数を特徴量として利用していないことが考えられる. このことから, 同一バス停における  $N$  便前までのバス停出発時の車内人数も有効な特徴量になり得るのではという仮説を立てた.

## 3. 本研究で利用するバス車内人数データ

本章では, バス停出発時の車内人数を予測するモデルを作成する際に必要なバス車内人数データの作成と基本統計量について述べる. 次に, そのデータの前処理について述べ, 最後に, そのバス車内人数データを利用して, 仮説に対する事前検証をする.

### 3.1 データの作成と基本統計量

本研究では, みなと観光バス株式会社が運営している 21 系統, 神戸国際大学前発の神戸行きで, 2021 年 10 月 1 日から 2022 年 1 月 31 日まで収集したデータを実験に用いた. この系統は, 6 つのバス停で構成されていて, 1 便の運行時間は 35 分である. 全員がバスを降車する終点を除く, 表 1 に示す 5 つのバス停出発時の車内人数を予測対象とした. 本研究では, 運行中のバスから表 2 に示すデータを 0.5 秒間隔で収集した. これらのデータから, 表 3 に示すような本研究で扱うバス車内人数データを作成した. このデータの作成では, 日時のデータからその日の何便目のバスかを表す便数を求め, 緯度と経度のデータからバス停名を求め, そのバス停を出発する時点における累積乗車人数と累積降車人数のデータから求めた車内人数を, そのバス停におけるバスの車内人数とした.

作成したバスの車内人数データの統計的な特徴を確認するため, 図 2 に示すように, バス停毎に平均値, 最大値, 最小値, 第一四分位数, 中央値, 第三四分位数の 6 つの基本統計量を示す箱ひげ図を作成した. 図 2 において, 最小値が負の値となるものが存在する理由は, データ収集の際に使用した精度 95% の乗降センサーが 1 人の降車を 2 人とカウントすることがあるためと考える. この図より, バス停 4 の神戸ベイシェラトンホテルにおけるバスの車内人数が平均的に多いことがわかる. 対照的に, バス停 1 の神戸国際大学前におけるバスの車内人数が平均的に少ないことがわかる.

### 3.2 データの前処理

収集したデータからバス車内人数データを作成したところ, 車内人数が負の値となっているノイズのあるデータや, 値が存在しない欠損しているデータを確認した. これは, データの収集に用いた乗降センサーの精度が 95% であるこ

表 1: バス停一覧

本研究での略称	バス停名
バス停 1	神戸国際大学前
バス停 2	ウエストコート 7 番街前
バス停 3	六甲アイランド病院前
バス停 4	神戸ベイシェラトンホテル
バス停 5	神戸・三宮

表 2: バスから収集したデータの項目

データ名
日時
緯度
経度
累積乗車人数
累積降車人数

表 3: バス車内人数データの項目

データ名
日時
便数
バス停名
車内人数

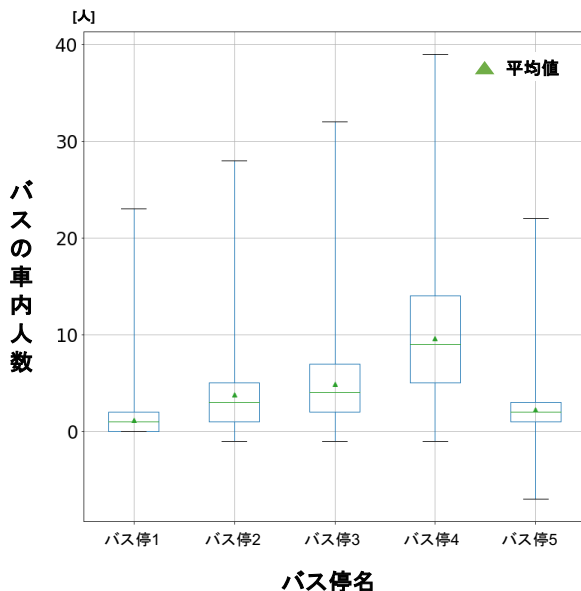


図 2: 5つのバス停におけるバスの車内人数の箱ひげ図

とに起因すると考える。本研究では、2021年10月1日から2022年1月31日までの1日に26便運行している、終点を除いた5つのバス停の車内人数データを用いる。そのため、欠損が発生しなかった場合の、表3に示すデータのタプル数（以下、データ数）が15,990個であるのに対して、表3のノイズのあるデータのタプル数（以下、ノイズデータ）は232個、表3の欠損しているデータのタプル数（以下、欠損データ）は433個であり、割合としてはそれぞれ1.45%と2.71%である。

一般的に、ノイズデータは補正した方が予測時の精度向上が見込める。そのため、ノイズデータの補正を前処理として実施した。また、欠損データは連続した期間で存在するわけではなく、点在していた。本研究では、バスの車内人数データを時系列データとして扱うので、点在する欠損データを許容しない。そのため、欠損データの補完を前処理として実施した。まず、ノイズデータの補正では、負の値となっているデータに対して、0からの差分を計算し、そのバス停以降のバス停で差分を加算する前処理を適用した。次に、欠損データの補完では、テストデータの目的変数として利用する2022年1月12日から2022年1月31日までのデータを除いたデータである、2021年10月1日から2022年1月11日までの期間のバスの車内人数の値をバス停と便数

毎に分類してから平均値を算出した値を適用した。このような補完をした理由は、時系列データの場合、テストデータの目的変数を欠損値データの補完に利用すると、データのリーケージとなるからである。なお、テストデータの目的変数として利用する期間の詳細は、第5章で述べる。

### 3.3 仮説に関する事前検証

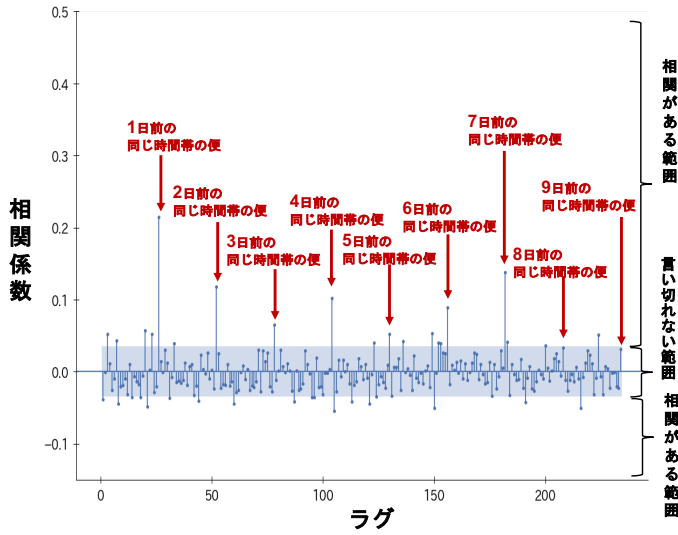
仮説に関する事前検証をする。そのため、5つのバス停それぞれで、同一バス停のN便前のバス停におけるバスの車内人数の変動特性を調査した。

自己相関は、N個前のデータとの相関を求める際に、N個前までに含まれるデータの影響を受けるため、N個前のデータのみとの相関関係を表すことができない。そのため、今回の事前検証では、N個前のデータのみとの相関を求めることができる偏自己相関をバス車内人数の変動特性の調査に用いた。

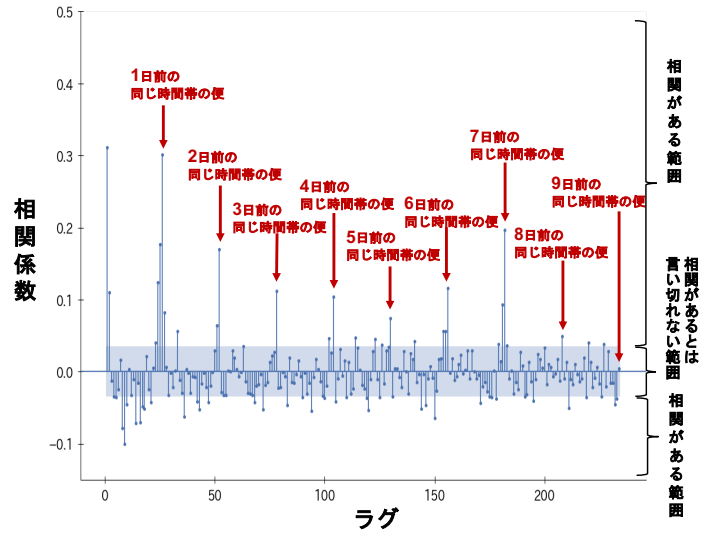
5つのバス停それぞれに偏自己相関を求めた結果を図3に示す。図中のラグNは、N便前のバス停におけるバスの車内人数との偏自己相関を表す。本研究で利用するバス路線の便数は1日あたり26便であるため、ラグ26が1日前の同じ時間帯の便を表す。また、図中に、Ljung-Box検定[13]によって求めた相関がある範囲と、相関があるとは言いきれない範囲を示した。これらの結果から、各バス停において、相関がある範囲にいくつかの値が存在することが確認できる。さらに、変動特性として、同じ時間帯の便において、相関係数の値が顕著に大きく、それは7、8日前まで継続することがわかる。そのため、同一バス停におけるN便前までの車内人数は有効な特徴量になり得ると言える。

## 4. 同一便と同一バス停におけるバス車内人数の時系列データの変動特性を考慮したLSTMによる車内人数予測手法

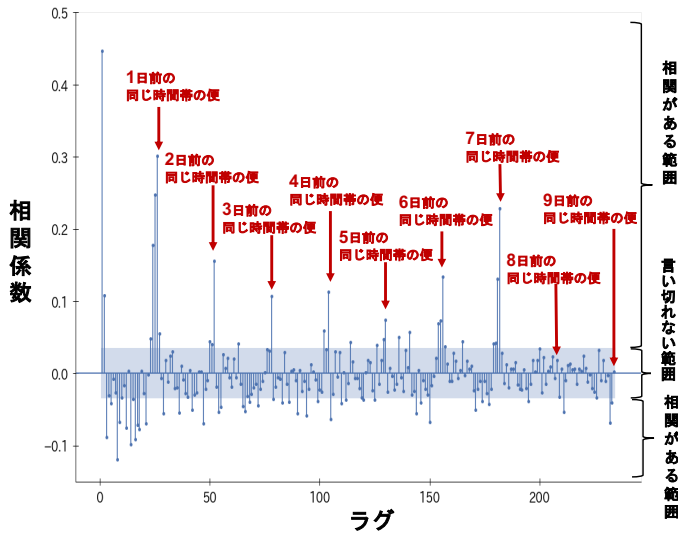
提案手法の概要を説明する。本研究では、同一便と同一バス停におけるバス車内人数の時系列データの変動特性を考慮したLong Short-Term Memory (LSTM) [14]によって、次のバス停の車内人数を予測する手法を提案する。LSTMは過去の時系列データから学習する。そこで、提案手法で採用する説明変数は、事前検証の結果に基づいて、



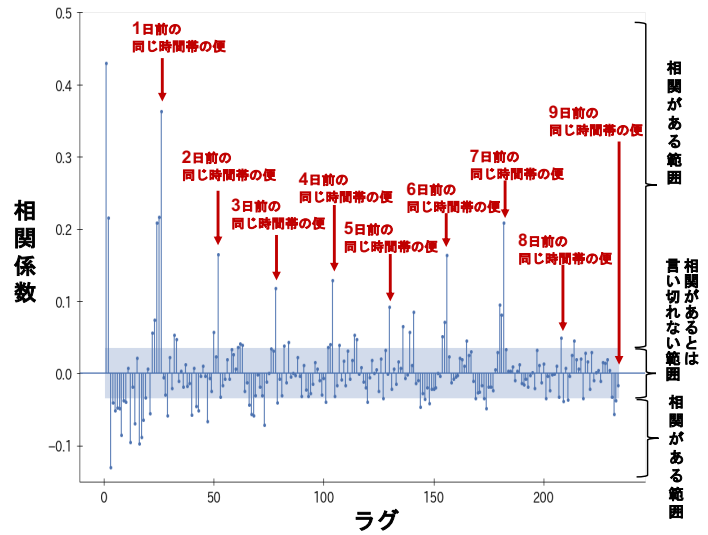
(a) バス停 1



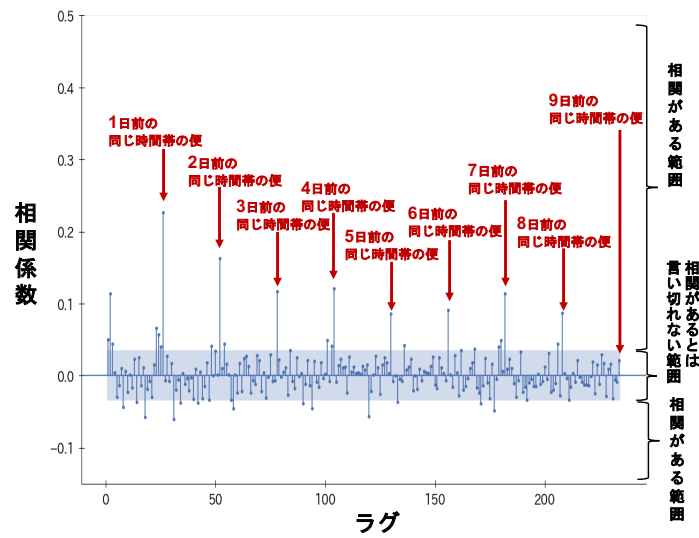
(b) バス停 2



(c) バス停 3



(d) バス停 4



(e) バス停 5

図 3: 5つのバス停におけるバスの車内人数の偏自己相関

過去 1 週間分の対象路線の各バス停における車内人数とした。このようにすることで、同一便と同一バス停における車内人数以外のデータも含むため、ノイズとなるデータも含まれるが、同一便と同一バス停におけるバス車内人数の時系列データの変動特性を考慮できる。

提案手法で LSTM を採用した理由を説明する。時系列予測では通常、変数間の関係を扱う必要がある。つまり、過去の各時点での値を無関係なものとして扱うのではなく、過去の様々な時点での値の関係性から次の時点での値を予測する。RF の元となったアルゴリズムである決定木という手法は 1 回の分岐に 1 つの変数しか見ないので、複数の値を同時に見る必要のある時系列予測には不向きである [15]。そのため、本研究では時系列予測のために開発されたアルゴリズムを用いる。

時系列予測のアルゴリズムとして Recurrent Neural Network (RNN) が注目されている [16]。しかし、RNN では数十ステップの短期依存には対応できるが、1000 ステップのような長期的な時系列データを学習できない。本研究では、1000 ステップのような長期的な時系列データを説明変数とする。そのため、長期的な時系列データを学習できるアルゴリズムである LSTM を採用した。また、損失関数としては、回帰のタスクを扱う際に一般的に用いられる Mean Squared Error (MSE) を用いた。

1 週間前までのデータを説明変数として設計した理由を説明する。図 3(a) より、バス停 1 において N 日前的の同じ時間帯の便にあたるラグを確認すると、ラグが 26, 52, 78, 104, 130, 156, 182 の時は相関がある範囲に値が存在するが、208 以降の同じ時間帯の便にあたるラグは相関があるとは言い切れない範囲に値が存在する。同様に図 3(b), 図 3(c), 図 3(d), 図 3(e) より、バス停 2 では 208 まで、バス停 3 では 182 まで、バス停 4 では 208 まで、バス停 5 では 208 までである。今回は不要なデータを含まないようにするため、ラグが 182 までのデータである 1 週間前までのデータを説明変数とした。さらに、予測したい値は、次のバス停におけるバスの車内人数という 1 次元データであるため、日、便、バス停という 3 次元データを図 4 に示すように 1 次元データに変換し、学習と予測をした。なお、LSTM の入力の制約上 3 次元データを 1 次元データに変換した。

## 5. 評価実験

本章では、提案手法によるバス車内人数の予測精度について評価するために実施した実験について述べる。

### 5.1 評価方法

本実験では、提案手法と既存手法の予測精度を Root Mean Square Error (RMSE) [17] を使って比較する。そのために、みなと観光バス株式会社が運営している 21 系統、

表 4: データ分割

データ名	目的変数の期間	割合 [%]	日数	便数	データ数
学習データ	2021/10/8 ~2022/1/1	74.1	86	2,236	11,180
検証データ	2022/1/2 ~2022/1/11	8.6	10	260	1,300
テストデータ	2022/1/12 ~2022/1/31	17.2	20	520	2,600

神戸国際大学前発の新神戸行きのバスより収集したデータから作成したバス車内人数データを使った、各手法によるバス車内人数の予測実験をする。

本研究で扱うデータは時系列データであるため、Train-Test Split を用いてデータを分割した。一般的に学習データ、検証データ、テストデータそれぞれの目的変数を分割する時は、学習データを 70~80%、検証データを 10%程度、テストデータを 10~20%とする。そのため、表 4 に示すようにデータを分割した。なお、第 4 章で説明した通り、説明変数を過去 1 週間分のデータと設計したため、説明変数として過去 1 週間分のデータが利用できる 2021 年 10 月 8 日の 1 便目のバス停 1 以降のデータを目的変数とした。2021 年 10 月 1 日~2021 年 10 月 7 日のデータは、説明変数として過去 1 週間分のデータが利用できないので、目的変数としては利用できない。

本研究では、提案手法に対する比較対象として、既存手法、統計的手法を採用した。既存手法として、Arabghalizi らの手法を RF 回帰によって、バス停における車内人数を予測する手法に変更したものを用いた。なお、本実験は人数予測であり、既存手法の混雑度予測とは精度の比較ができないため、既存手法を回帰にすることで人数予測にした。また、テストデータの目的変数以外の期間のデータである 2021 年 10 月 1 日から 2022 年 1 月 11 日までのバス停毎で分類した平均値を予測値として用いた手法を統計的手法 1、2021 年 10 月 1 日から 2022 年 1 月 11 日までのバス停と便数毎で分類した平均値を予測値として用いた手法を統計的手法 2 とした。

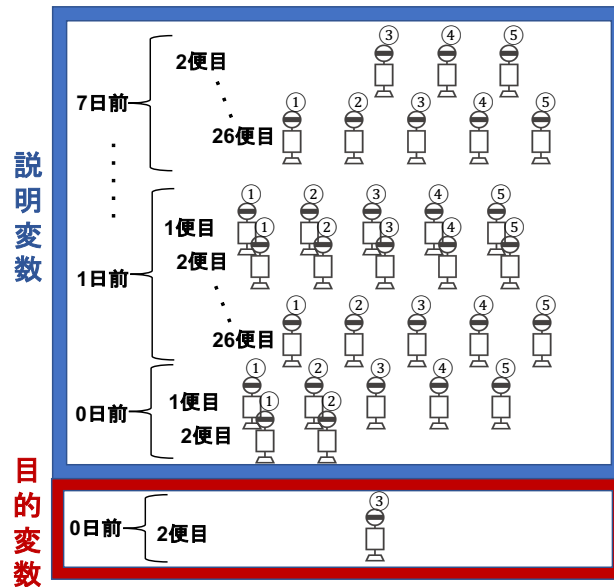
### 5.2 パラメータチューニング

提案手法と既存手法、それぞれに対してグリッドサーチを用いてパラメータチューニングを行った。なお、RF は、一般的にパラメータチューニングを行わなくても過学習を起しにくいアルゴリズムになっている [18]。しかし、本研究では提案手法との公平性の観点からパラメータチューニングを行った。

#### 5.2.1 提案手法のパラメータチューニングの結果

表 5 に提案手法のパラメータチューニングの結果を示す。それぞれのパラメータの候補値は、一般的に LSTM のグリッドサーチの際に使用される値を設定した。表中の候補の項目に記載の値に対してグリッドサーチを行った結果、





説明変数 5 バス停 × 26 便数 × 7 日 = 910

目的変数 1



図 4: 3次元データを1次元データに変換

エポック数が 53, 学習率が 0.001, バッチサイズが 64, 中間層のノード数が 80, LSTM 層の数が 3 の時に最も損失関数の値が小さくなった. なお, 最も損失関数の値が小さくなった時を, 最も精度が良いものとする.

図 5 に, パラメータチューニング後における LSTM のニューラルネットワークの構造を示す. 入力層が最初に存在し, 次に LSTM 層が 3 層存在し, 最後に Dense 層を用いた. 最終的な出力値がバスの車内人数という線形的な値であるため, 最後の層には Dense 層を用いた.

図 6 は, 最も精度が良い時の値のパラメータにおいて, 学習データと検証データを用いて学習した際の損失値の推移を示したグラフである. 損失値の結果より, エポック数が 53 で停止する際には, 学習データと検証データの損失値が同程度となっているので, 過学習や未学習を起さずに学習できていると言える.

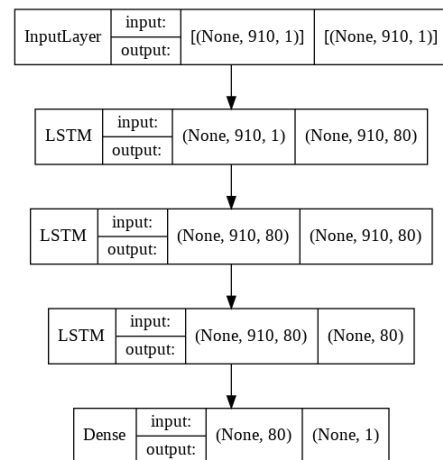


図 5: 提案手法のチューニング後のニューラルネットワーク構造

表 5: LSTM のパラメータチューニングの結果

パラメーター	候補	最も精度が良い時の値
エポック数	[1~200]	53
学習率	[0.01, 0.001, 0.0001]	0.001
バッチサイズ	[2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048]	64
中間層のノード数	[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200]	80
LSTM 層の数	[1, 2, 3, 4]	3

### 5.2.2 既存手法のパラメータチューニングの結果

既存手法はバス停毎にモデルを作成するため, パラメータチューニングもバス停毎に行った. 表 6 にパラメータチューニングの結果を示す. 全てのバス停において, 最も精度が良い時の決定木の数, 最小サンプルサイズ, 決定木の深さの最大値のパラメータは, 候補の中でも比較的小さな値となった.

## 6. 評価結果と考察

提案手法, 既存手法, 統計的手法 1, 統計的手法 2 の評

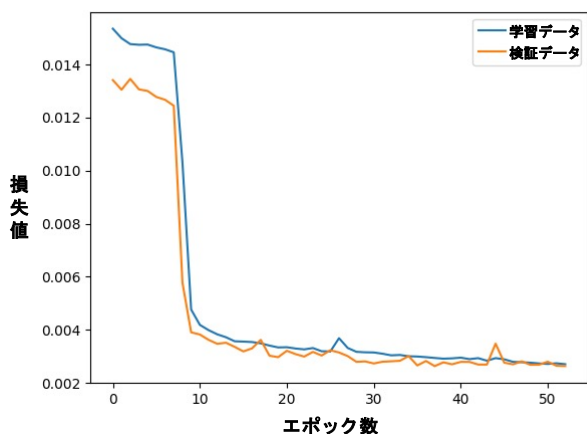


図 6: 学習データと検証データの損失値

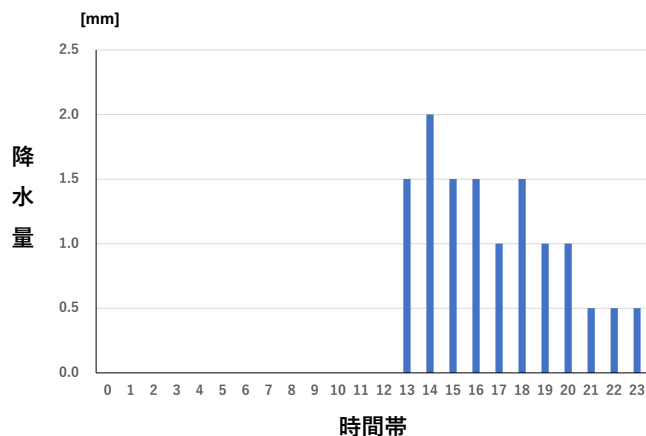


図 7: 2022 年 1 月 23 日の 1 時間毎の降水量

表 6: 既存手法におけるパラメータチューニングの結果

バス停	パラメータ	候補	最も精度が良い時の値
バス停 1	決定木の数	[5, 10, ~ 295, 300]	55
	最小サンプルサイズ	[5, 10, ~ 95, 100]	45
	決定木の深さの最大値	[5, 10, ~ 95, 100]	10
バス停 2	決定木の数	[5, 10, ~ 295, 300]	65
	最小サンプルサイズ	[5, 10, ~ 95, 100]	40
	決定木の深さの最大値	[5, 10, ~ 95, 100]	15
バス停 3	決定木の数	[5, 10, ~ 295, 300]	55
	最小サンプルサイズ	[5, 10, ~ 95, 100]	10
	決定木の深さの最大値	[5, 10, ~ 95, 100]	10
バス停 4	決定木の数	[5, 10, ~ 295, 300]	65
	最小サンプルサイズ	[5, 10, ~ 95, 100]	30
	決定木の深さの最大値	[5, 10, ~ 95, 100]	20
バス停 5	決定木の数	[5, 10, ~ 295, 300]	40
	最小サンプルサイズ	[5, 10, ~ 95, 100]	40
	決定木の深さの最大値	[5, 10, ~ 95, 100]	20

表 7: バス車内人数の予測結果

	RMSE [人]				
	バス停 1	バス停 2	バス停 3	バス停 4	バス停 5
提案手法	1.054	<b>1.304</b>	<b>0.837</b>	<b>1.853</b>	<b>1.229</b>
既存手法	1.189	1.764	1.247	2.607	1.581
統計的手法 1	0.992	2.411	3.056	4.668	1.601
統計的手法 2	<b>0.904</b>	1.736	2.060	3.259	1.540

価結果を表 7 に示す。なお、バス停毎に最も結果の良い値を太字で記載する。提案手法は既存手法と比べると、全てのバス停において RMSE の値が小さい結果となった。このことから、既存手法と比べて提案手法は、次のバス停における車内人数の予測精度が向上したと言える。しかし、バス停 1 においては、統計的手法 2 の RMSE の値が提案手法よりも小さい結果となった。

提案手法の予測精度が既存手法と比べて向上した結果は、提案手法で使用した特徴量とモデルの組み合わせが、既存手法で使用した組み合わせよりも有効であることを示すと考える。一方で、5 つのバス停間における RMSE の値に大きな差が見られない結果となったのは、本研究で用いた路線の車内人数のデータが、図 2 に示した通り、平均的に少なかったためであると考えられる。また、バス停 1 にお

て提案手法よりも統計的手法 1、統計的手法 2 の RMSE の値が小さい結果となった要因としては、バス停 1 が始点であるので、同一便のバス停におけるバスの車内人数のデータの特徴量として利用できないためであると考えられる。

一般的に、乗客数と天候には相関関係があることが示されている [19]。しかし、本実験では、降雨の有無がバス車内人数の予測精度に与える影響について評価できなかった。理由としては、本実験では 2022 年 1 月 12 日から 2022 年 1 月 31 日までをテストデータの目的変数として用いたが、この期間において 1 時間に 0.5mm 以上の降水があったのは、2022 年 1 月 23 日のみであり、降水があった日のデータが十分にはないからである。なお、2022 年 1 月 23 日の 1 時間毎の降水量を図 7 に示す。今回使用した降雨データは、気象庁が公開している兵庫県神戸市の気象台で観測されたデータで [20]、降雨量は 0.5mm 単位で表示されていた。

今後の課題は、平均車内人数が多い路線でデータを収集し、説明変数の種類とデータ数の妥当性を検証すること、降雨の有無がバス車内人数の予測精度に与える影響について評価することである。

## 7. おわりに

本研究では、同一便と同一バス停における車内人数の時系列データに変動特性として、偏自己相関を考慮したバス車内人数の予測手法を提案した。提案手法では、LSTM を用いて、過去 1 週間分の予測対象路線の各バス停における車内人数を説明変数として、次のバス停における車内人数を予測した。評価では、提案手法が既存手法よりも高い予測精度を持つことを示した。しかし、始点のバス停においては、提案手法よりも統計的手法の予測精度が最も高い結果となった。

今後の課題は、次のバス停における車内人数だけでなく、次の便や翌日などのより遠い未来における車内人数を予測することがある。また、今回評価対象として用いた路線は平均車内人数が少なかったため、平均車内人数が多い

路線でデータを収集し、説明変数の種類とデータ数の妥当性の検証、降雨がある日で評価することも課題である。

## 参考文献

- [1] 内閣府: 平成 28 年度公共交通に関する世論調査, 内閣府世論調査サイト (オンライン), 入手先 <<https://survey.gov-online.go.jp/h28/h28-kotsu/2-1.html>> (参照 2022-05-17) .
- [2] Google: GTFS Realtime Overview, Google Developer サイト (オンライン), 入手先 <<https://developers.google.com/transit/gtfs-realtime>> (参照 2022-05-17) .
- [3] Arabghalizi, T. and Labrinidis, A.: How full will my next bus be? A Framework to Predict Bus Crowding Levels, *Proc. The 8th SIGKDD International Workshop on Urban Computing*, ACM, pp. 1-8 (2019).
- [4] Nakashima, H., Arai, I. and Fujikawa, K.: Passenger Counter Based on Random Forest Regressor Using Drive Recorder and Sensors in Buses, *Proc. 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, IEEE, pp. 561-566 (2019).
- [5] 金光勇慈, 田谷瑛悟, 立花巧樹, 中村優吾, 松田裕貴, 諏訪博彦, 安本慶一: BLE を用いた路線バスの混雑度推定, マルチメディア, 分散協調とモバイルシンポジウム 2021 論文集, Vol. 2021, No. 1, pp. 62-69 (2021).
- [6] 山田遊馬, 廣森聡仁, 山口弘純, 東野輝夫: 測域センサを利用した高精度な路線バス乗降計測システム, 第 25 回マルチメディア通信と分散処理ワークショップ論文集, Vol. 2017, pp. 24-32 (2017).
- [7] Mehmood, U., Moser, I., Jayaraman, P. P. and Banerjee, A.: Occupancy Estimation using WiFi: A Case Study for Counting Passengers on Busses, *Proc. 2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, IEEE, pp. 165-170 (2019).
- [8] Tang, C., Li, W., Vishwakarma, S., Chetty, K., Julier, S. and Woodbridge, K.: Occupancy Detection and People Counting Using WiFi Passive Radar, *Proc. 2020 IEEE Radar Conference (RadarConf20)*, IEEE, pp. 1-6 (2020).
- [9] Hsu, Y.-W., Chen, Y.-W. and Perng, J.-W.: Estimation of the Number of Passengers in a Bus Using Deep Learning, *Sensors*, Vol. 20, No. 8, pp. 2178-2196 (2020).
- [10] 青地美桜, 白石陽: 利用者の乗降データを用いたバス乗客人数予測モデルの評価, 研究報告高度交通システムとスマートコミュニティ論文集, Vol. 2020, No. 13, pp. 1-8 (2020).
- [11] 青地美桜, 白石陽: 停留所毎モデルを用いたバス乗降データによる乗客人数予測の評価, 研究報告マルチメディア通信と分散処理論文集, Vol. 2021, No. 12, pp. 1-6 (2021).
- [12] Wood, J., Yu, Z. and Gayah, V. V.: Development and evaluation of frameworks for real-time bus passenger occupancy prediction, *International Journal of Transportation Science and Technology*, Vol. 7, No. 3, pp. 169-228 (2022).
- [13] Ljung, G. M. and Box, G. E. P.: On a measure of lack of fit in time series models, *Biometrika*, Vol. 65, No. 2, pp. 297-303 (1978).
- [14] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. and Schmidhuber, J.: LSTM: A Search Space Odyssey, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28, No. 10, pp. 2222-2232 (2017).
- [15] Safavian, S. R. and Landgrebe, D.: A survey of decision tree classifier methodology, *IEEE transactions on systems, man, and cybernetics*, Vol. 21, No. 3, pp. 660-674 (1991).
- [16] Sherstinsky, A.: Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network, *Physica D: Nonlinear Phenomena*, Vol. 404, pp. 1-28 (2020).
- [17] Chai, T. and Draxler, R. R.: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, *Geoscientific Model Development*, Vol. 7, No. 3, pp. 1247-1250 (2014).
- [18] Liaw, A. and Wiener, M.: Classification and Regression by randomForest, *R news*, Vol. 2, No. 3, pp. 18-22 (2002).
- [19] Elnoshokaty, A., Arai, I., El-Tawab, S. and Salman, A.: Transit System Prediction for Real-time Weather Conditions: Fleet Management and Weather-related Ridership, *Proc. 2022 IEEE International Conference on Smart Mobility (SM)*, IEEE, pp. 14-20 (2022).
- [20] 国土交通省: 国土交通省気象庁, 過去の気象データ, 国土交通省気象庁サイト (オンライン), 入手先 <<https://www.data.jma.go.jp/obd/stats/etrn/index.php>>(参照 2022-05-17) .