Comparison of Time Series Methods for Post-COVID Transit Ridership Forecasting*

*Note: Title based on the content of the analyzed article

1st Carlos Castro
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Freddy Valenzuela
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This document presents the problem statement and literature review based on the study by Hightower et al. (2024), focusing on the obsolescence of traditional transit forecasting models in the face of COVID-19 disruptions and the evaluation of advanced algorithms to mitigate uncertainty. Our work is explicitly grounded in the published article *"A comparison of time series methods for post-COVID transit ridership forecasting"* by Hightower et al. (2024), from which we adopt both the dataset and the experimental design as a benchmark setting. The dataset consists of system-level, monthly unlinked passenger trips for heavy rail agencies in the continental United States from 2002 to 2023, segmented into pre-COVID, full-series, and post-COVID periods. Building on their findings, our project proposes contrasting those results with deep learning models (a CNN, an LSTM, and an Artificial Neural Network, ANN) that were not implemented in the original study, in order to explore their potential for post-COVID ridership forecasting.**

***Index Terms*—Forecasting, Transit Ridership, Time Series, COVID-19, Heavy Rail**

I. Related Literature Review

Recent literature on public transit demand forecasting has undergone significant changes due to the disruptions caused by the COVID-19 pandemic. Studies prior to 2020 focused on traditional methods such as ARIMA or exponential smoothing under stable conditions; however, research published in the last five years highlights the insufficiency of these approaches in the presence of abrupt structural changes [1].

Table I summarizes the most relevant studies analyzed, highlighting their methodologies and contributions to forecasting in dynamic environments. In particular, several recent studies have incorporated deep learning techniques, although the article by Hightower et al. (2024) limits its comparison to seven classical and semi-modern time series methods (ETS, ARIMA, STL-ETS, STL-ARIMA, TBATS, NNET, and a hybrid model), without implementing deep architectures such as CNN or LSTM in their own experiments. In this project, we explicitly build on that study: we use the same system-level heavy rail ridership dataset and the same temporal segmentation (pre-COVID, full series, and post-COVID) as a reference framework for our deep learning experiments.

Multiple authors have explored the integration of machine learning methods and advanced statistical models to reduce uncertainty. For example, Chen et al. (2020) demonstrated the usefulness of LSTM networks combined with seasonal decomposition to improve short-term metro ridership prediction [2]. Similarly, Egu and Bonnel (2021) examined medium-term forecasting, emphasizing the need to understand the underlying data-generating mechanisms rather than relying solely on historical extrapolation [3].

Following the onset of the pandemic, the literature shifted toward understanding the new dynamics. Azimian and Jiao (2021) used random forest models to predict daily ridership in Chicago, finding that accuracy declined as the forecast horizon increased and that variables such as temperature had a significant impact [4]. Likewise, Moghimi et al. (2022) addressed non-stationarity in the New York City subway using piecewise ARIMA models, demonstrating that explicitly modeling breakpoints improved performance relative to single-model approaches [5].

Comparisons between classical and modern methods have been central. Gao et al. (2023) compared ARIMA with Facebook's Prophet model in five major U.S. cities, finding that the lack of clear seasonality hindered Prophet's calibration, resulting in errors between 6% and 12% [6]. Meanwhile, Caicedo et al. (2023) evaluated deep learning models (CNN, LSTM) in Bogotá, concluding that under highly dynamic conditions (COVID and protests), deep learning models required up to 1.5 months to learn new demand patterns [7].

Other approaches have sought to optimize neural network architecture. Ayman et al. (2022) proposed automated hyperparameter search for deep networks, achieving significant improvements by customizing models for each route [8]. Finally, Ziedan et al. (2023) expanded the analysis to external factors, identifying telework and unemployment as major contributors to the nationwide decline in transit ridership [10].

Overall, the literature shows that deep architectures (CNN, LSTM) have demonstrated potential in dynamic contexts but remain underutilized in systemwide heavy rail ridership forecasting, reinforcing the motivation for our project to focus specifically on these models within the same dataset and scenario analyzed by Hightower et al. (2024).

II. Identification of the Object of Study

A. Problem Identification

The object of study is the inability of historical forecasting methods to accurately predict passenger ridership in heavy rail systems across the United States following the structural break caused by the COVID-19 pandemic. Seasonal patterns and stable trends that existed before March 2020 have disappeared or changed drastically, making models trained on historical data less reliable for the current "new normal."

B. Justification of Importance

Addressing this issue is critical from both administrative and financial perspectives. Public transit agencies depend on accurate systemwide forecasts for operational budgeting, long-term investment prioritization, and service frequency planning. Without reliable post-COVID forecasting models capable of handling uncertainty, agencies risk severe budget deficits or mismatches between service supply and actual demand, ultimately affecting operational efficiency and user experience.

C. Nature of the Dataset

This project relies directly on the dataset used and documented by Hightower et al. (2024). The data consist of system-level, monthly heavy rail ridership in terms of unlinked passenger trips (UPT) for fourteen heavy rail agencies in the continental United States, as reported in the National Transit Database. The time span ranges from January 2002 to December 2023, providing a long historical record that includes both pre-pandemic and post-pandemic conditions.

Following Hightower et al. (2024), the series are segmented into three distinct analysis periods: (i) a pre-COVID period used to establish a baseline under stable conditions, (ii) a full-series period that incorporates both pre- and post-COVID observations, and (iii) a post-COVID period that isolates the "new normal." For each period, the last twelve months are held out as a test set. By adopting the same dataset and temporal segmentation, our deep learning models (CNN, LSTM, ANN) can be compared directly against the time series methods evaluated in the original study.

III. Problem Statement

The central problem lies in the obsolescence of traditional forecasting models when facing the volatility introduced by the pandemic, which generates uncertainty in transit planning. Technically, this is a problem of **univariate time series analysis and forecasting**, in which the data's stationarity and periodicity have been compromised.

In the study by Hightower et al. (2024), this complexity is addressed through seven algorithms ranging from statistical to light machine learning approaches. Classical methods such as **ETS** (Exponential Smoothing) and **ARIMA** are evaluated, along with their decomposed variants (**STL-ETS** and **STL-ARIMA**). To capture nonlinear patterns and complex seasonality, the authors use the **TBATS** model (Box-Cox transformation and Fourier terms) and Neural Network Autoregression (**NNET**). Finally, a **Hybrid** model that combines the forecasts of the previous methods is proposed to increase robustness under uncertainty.

In our project, instead of replicating all these algorithms, the work of Hightower et al. serves as a reference and comparison baseline. Building on this, we propose evaluating three deep learning models: a one-dimensional *Convolutional Neural Network* (CNN), a *Long Short-Term Memory* (LSTM) network, and an *Artificial Neural Network* (ANN), with the objective of determining whether these architectures better capture post-COVID patterns in heavy rail ridership. Our models are trained and evaluated on the same dataset and time periods defined by Hightower et al. (2024), ensuring direct comparability between classical time series methods and deep learning approaches.

The justification for adopting an analytical approach involving deep neural networks, in addition to the classical methods reported in the original study, is based on the need to capture nonlinear relationships and complex seasonal structures that purely linear models (such as simple ARIMA) may fail to detect in the post-COVID environment. Since the data-generating mechanisms have changed, it is essential to empirically evaluate which mathematical architecture best adapts to the "new normal," without assuming the superiority of classical models a priori.

Project Objectives

This document seeks to compare and contextualize various time series forecasting approaches to predict heavy rail ridership across the continental United States, specifically addressing the challenges posed by the post-COVID era.

- **Evaluate the accuracy of three deep learning models** (CNN, LSTM, and ANN) by computing and comparing error metrics (MAPE and MASE) on heavy rail agency data, contrasting their performance with the results reported by Hightower et al. (2024) using classical methods (ETS, ARIMA, STL, TBATS, NNET, and hybrid).
- **Analyze the impact of the training period on forecasting quality** by segmenting the data into three distinct periods (pre-COVID, full series, and post-COVID only), in order to determine whether including historical data prior to 2020 improves or worsens prediction accuracy when using deep learning models.
- **Investigate the persistence or alteration of ridership patterns** through residual analysis and resulting trends, to identify whether pre-pandemic seasonal patterns have re-emerged and how this influences future demand forecasting difficulty.

TABLE I
SUMMARY OF RELATED LITERATURE AND METHODOLOGIES

Authors	**Year**	**Methodology / Algorithm**	**Main Finding / Contribution**
Chen et al. [2]	2020	LSTM + Seasonal Decomposition	Combining LSTM networks with decomposition improves short-term metro ridership accuracy.
Azimian and Jiao [4]	2021	Random Forest	Accuracy decreases with forecast horizon; climate variables significantly affect daily prediction during COVID.
Egu and Bonnel [3]	2021	Time Series Analysis	Highlights the need to understand data-generating mechanisms rather than relying solely on extrapolation.
Ayman et al. [8]	2022	Neural Networks (NAS)	Automated hyperparameter search improves performance over generic network designs.
Moghimi et al. [5]	2022	Piecewise ARIMA	Explicitly modeling breakpoints outperforms single ARIMA models under non-stationarity.
Gao et al. [6]	2023	ARIMA vs. Prophet	Lack of clear seasonality post-COVID hinders Prophet's calibration compared to ARIMA.
Caicedo et al. [7]	2023	Deep Learning (CNN, LSTM)	Deep learning models require approx. 1.5 months to learn new demand patterns after a shock.
Ziedan et al. [10]	2023	Multiple Mediation Analysis	External factors (telework, unemployment) explain 13–38% of the ridership decline.

REFERENCES

[1] S. Washington, M. Karlaftis, F. Mannering, and P. Anastasopoulos, *Statistical and Econometric Methods for Transportation Data Analysis*. New York, NY: CRC Press, 2020.

[2] D. Chen, J. Zhang, and S. Jiang, "Forecasting the short-term metro ridership with seasonal and trend decomposition using loess and LSTM Neural Networks," *IEEE Access*, vol. 8, 2020.

[3] O. Egu and P. Bonnel, "Medium-term public transit route ridership forecasting: what, how and why? A Case Study in Lyon," *Transp. Policy*, vol. 105, pp. 124–133, 2021.

[4] A. Azimian and J. Jiao, "Prediction of 'L' Train's daily ridership in downtown Chicago during the COVID-19 Pandemic," *Findings*, 2021.

[5] B. Moghimi, C. Kamga, A. Safikhani, S. Mudigonda, and P. Vicuna, "Non-stationary time series model for station-based subway ridership during COVID-19 Pandemic: case study of New York City," *Transp. Res. Rec.*, vol. 2677, no. 4, 2022.

[6] Y. Gao, J. Li, J. Wang, and Z. Yang, "U.S. public transportation ridership analysis and prediction based on COVID-19," *Int. Soc. Opt. Eng.*, vol. 12596, 2023.

[7] J. D. Caicedo, M. C. González, and J. L. Walker, "Public Transit Demand Prediction During Highly Dynamic Conditions: A Meta-Analysis of State-of-the-ART Models and Open-Source Benchmarking Infrastructure," arXiv preprint, 2023.

[8] A. Ayman, J. Martinez, P. Pugliese, A. Dubey, and A. Laszka, "Neural Architecture and Feature Search for Predicting the Ridership of Public Transportation Routes," in *2022 IEEE International Conference on Smart Computing (SMARTCOMP)*, Helsinki, Finland, 2022.

[9] J. M. Hoque, I. Zhang, D. Schmitt, and G. D. Erhardt, "Are public transit investments based on accurate forecasts? An analysis of the improving trend of transit ridership forecasts in the United States," SSRN, 2023.

[10] A. Ziedan, L. Lima, and C. Brakewood, "A multiple mediation analysis to untangle the impacts of COVID-19 on nationwide bus ridership in the United States," *Transp. Res. Part A Policy Pract.*, vol. 173, 2023.

[11] X. Zheng and P. Wei, "Air transportation direct share time series forecasting: a hybrid model," *J. Adv. Transp.*, vol. 3, pp. 1–13, 2020.

[12] G. Jayanthi and P. Jothilakshmi, "Traffic time series forecasting on highways - a contemporary survey of models, methods and techniques," *Int. Journal Logist. Syst. Manag.*, vol. 39, no. 1, pp. 77–110, 2021.