EasyChair Preprint

№ 5181

An Integrated and Dynamic Commuter Flow Forecasting System for Railways

Goutham Reddy, Yaswanth Reddy and Bevish Jinila

March 20, 2021

An Integrated and Dynamic Commuter Flow Forecasting System for Railways

[1]D. Goutham Reddy, [1]G. Yaswanth Reddy and [2]Y. Bevish Jinila

[1] Student, [2]Associate Professor,
School of Computing
Sathyabama Institute of Science and Technology, Chennai, India
Corresponding Author : bevish.jinila@gmail.com

Abstract

Uncertain and instable passenger flow in Urban Metro Transit is a growing concern in the recent rail transport system. It is vital to forecast the passenger flow, in-order to provide a reliable daily operation and management. Short-term forecasting has become the most important component for an efficient rail management system. Existing literatures on passenger flow forecasting is based on Extreme Kernel approach that learns and forecasts signals with different frequencies. These approaches are not able to train and remember over a long time due to issues of backpropagated error. By addressing this problem, holt-winters forecasting algorithm is used. Experimental discussion shows that the holt algorithm provides better efficiency based on metrics including accuracy and F-measure.

Keywords: Rail transit, forecasting, short term, passenger flow, prediction

1 Introduction

Developing cities metro electric rail transport is one of the most important part of the urban infrastructure. It is the most commonly used transportation in metro cities [1]. It helps in reducing air pollution as most of these transport systems uses environment friendly fuel sources and also reduces traffic congestion. It is important to predict the passenger flow correctly to avoid the wastage of resources. And, in addition this data can be used by the commuters to precisely schedule their travel on time. Further, improving the operations of existing rail system can reduce the energy consumption and CO_2 emission. The energy consumption can be reduced by efficiently planning the number of trips scheduled, lighting, and ventilation. The travel forecast in metropolitan area is modeled and studied [2]. The inference of this work presented the total travel demand of a particular region.

The metro transportation system of the several cities in India is doing well and is also placing more and more resources in the sector for passenger's safety and comfort. This transportation system is also using resources uniformly resulting in using more number of resources in places required less resources. This results in the overflow of resources and manpower in places not necessary. The timing of the trains in the metro transport services are not according to the passenger crowd. This results in running the trains with less number of passengers in each route which often results in loss of operational costs which is not profitable or barely covers the operational costs. And other key issues is the availability according to the timing results in different routes are more frequent in less rush periods which results in exactly in the same situation as addressed above.

The less investment in research and development means sticking with the older equipment for an extended period of time. The older equipment is not compatible with the modern needs of the passengers. The emission of the pollution from the older vehicles is higher

than that of the modern vehicles this results in the emission of gases that contains more or higher carbon levels. This is quite common problem in the developing nations. This problem can be achieved through modern technology usage. There are number of useful ways that costs less and arrange this system to gain more profits through the usage of technology in the service sector.

There are a variety of methods used for forecasting the passenger flow in a rail transit system namely neural network models like Support Vector Machine (SVM), Convolution Neural Network (CNN), graphical CNN, Recurrent Neural Networks (RNN) and hybrid models, time series models, frequency based models, elasticity based models [3].

The recent existing method is a hybrid model which includes the combination of wavelet transform and machine learning. The idea behind this model is classifying the passenger flow datasets into lower and higher level sequences and then learn and forecast different frequencies using machine learning method. Finally, different prediction sequences are reconstructed using wavelet transform. Due to the back propagation error, they are unable to train and memorize data over a long time. To address this issue, Holt-Winters forecasting model is proposed. The inputs of the model are abnormal features, which consist of the recent time-series data.

Section II elaborately describes about the related work. Section III elates on the proposed methodology of forecasting the passenger flow. Section IV discusses on the performance analysis and discussion. Section V discusses the conclusion of the work and briefs on the future enhancements.

2 Related Work

This section briefly shows the related literatures on the prediction of passenger flow in rail transport system. As per the directions of the metro train systems the resources like fuel are managed and distance travelled the fuel is refilled through gas stations [4].This notes and referred how to save the resource and the searching of the nearby stations and gas refilling stations [5]. Those are used to maintain the traffic flow prediction [6].The improvements done through the traffic networks through volume of traffic. The routes should be used to compare the destinations and the passenger data sets [7].

The traffic flow prediction can be achieved through seasonal ARIMA process. This model requires the limited input data through data sets [8]. The traffic flow on the rails can also be predicted by combining Kohonen map with ARIMA. Through this model requires time series models to forecast the traffic flow and get to the destinations through minimum time delays [9]. The resources also should be utilized through different processes and the requirements through modeling the outcome of the resources shared on the efficiency of supply chain [10].

A hybrid optimization of computation of intelligence techniques for highway passenger flow or concentration prediction is conducted and recorded as the data [11]. The neural network based on radial basis function is used for freeway traffic volume forecasting [12]. The traffic flow is also studied in [13-17].

3 Proposed Methodology

The methodology that is used in the prediction is based on Holt-Winters time-series forecasting model and SARIMA. The extreme machine learning used to study the relevant data sets which are used to make prediction of the future of the data.

The statistical representation is done through many graph structures which will be shown in the figures. The immediate requirement is to out write a code that will analyze the data and implement the prediction required. The main way of the representation is to through short long term memory. The basic idea is to analyze the weekly data and in special days through long term memory. That means we are not only using regular day schedules but also using the data which is recorded in the previous events.

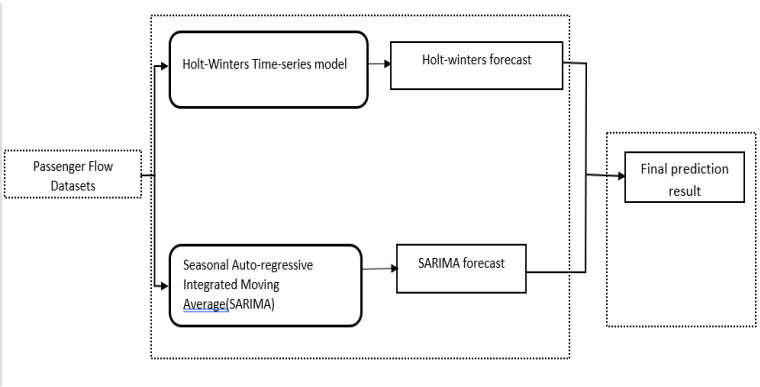


Fig. 1 Proposed system architecture

3.1 Holt-Winters Forecasting Model:

Holt-Winters forecasting model is a time series approach that helps to predict the behavior over time. Holt-Winters is one of the models of time-series behavior. This model consists of three features namely, average, slope and seasonality. This model works based on the computation of central value and further addition of the slope and seasonality

3.2 Seasonal Autoregressive Integrated Moving Average (SARIMA):

The SARIMA model process requires three new seasonal parameters which are hyper active, these parameters are stated as auto-regression, differencing and average. There is an additional parameter included and is called as period of seasonality. SARIMA is used in python using three steps:

1. Define the model
2. Fit the defined model
3. Make the prediction with the fit model

The following topics will be covered while executing this process:

1. Stationarity (Differencing and Augmented Dickey-Fuller Test)
2. ACF and PACF Plots
3. Grid search and Akaike Information criterion (AIC)
4. Walk Forward Validation
5. Mean Absolute Percentage Error (MAPE)
6. Exogenous Variable

The above mentioned processes the SARIMA will be processed step by step until the each step thar are mentioned above are responsible for the end result. The effectiveness of the SARIMA will be measured using mean absolute percentage error. These results are plotted on a graph to compare the results got out if the process that is accurate enough to go through the next procedures that is present in the prediction sector**.**

4. Experimental Discussion

This section shows the experimental results obtained in the implementation of the forecasting algorithm. To implement this approach, the dataset of traffic flow is taken from the Kaggle dataset. Fig.2 shows the statistical figure that represents the number of passengers getting in and out of the train at the destinations of the entire year. It clearly shows the usage of the railway passengers over a period of time.

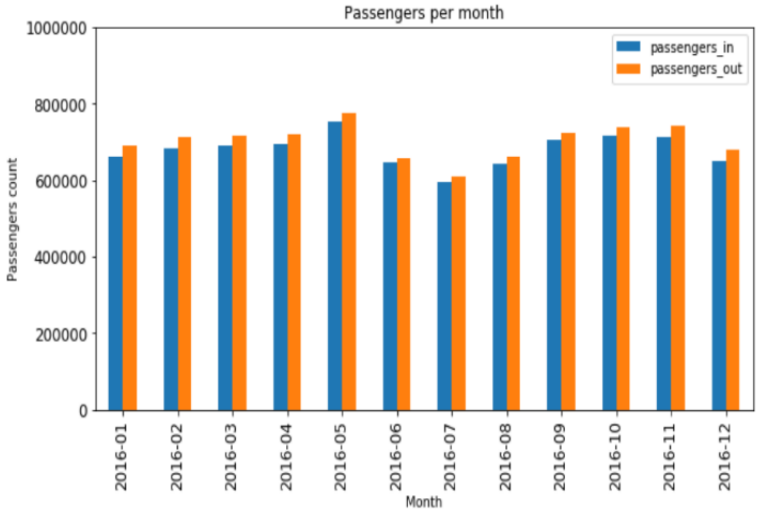


Fig. 2 Month vs passenger count

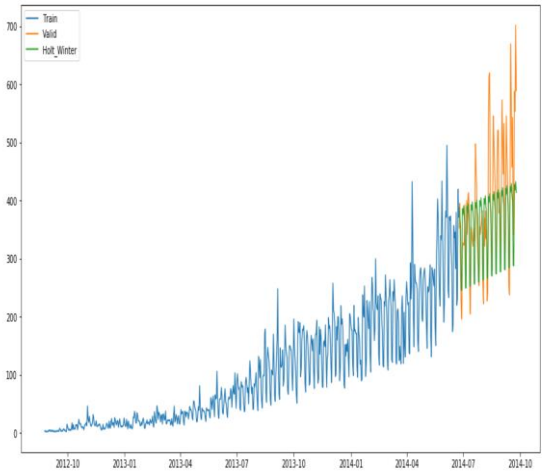
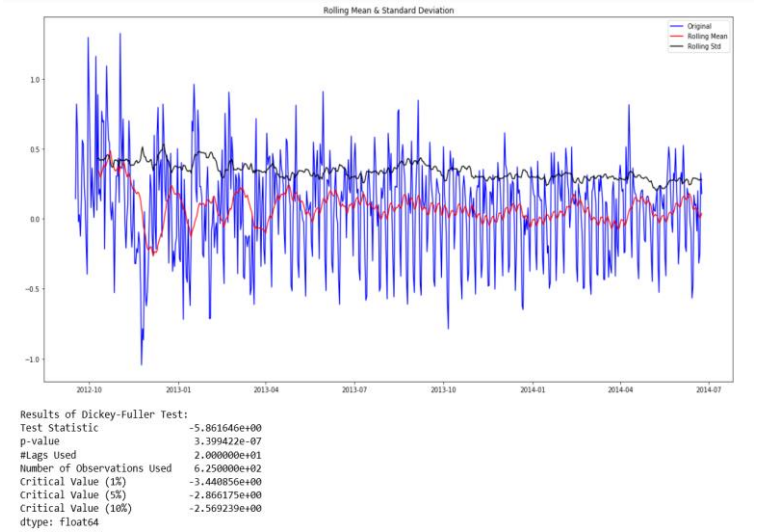


Fig.3 Graphical representation of the datasets with a linear base line using holt algorithm.

The data is analyzed with the time series holt algorithm. Fig. 3 shows the graphical representation of the data distribution. Fig. 4 shows the results obtained after presdiction using the rolling mean and standard deviation. The auto correlation and partial auto correlation is shown in fig.6.




```
Results of Dickey-Fuller Test:
Test Statistic                -5.861646e+00
p-value                        3.399422e-07
#Lags Used                     2.000000e+01
Number of Observations Used    6.250000e+02
Critical Value (1%)           -3.440856e+00
Critical Value (5%)           -2.866175e+00
Critical Value (10%)          -2.569239e+00
dtype: float64
```

Fig.5 The prediction results of rolling mean and standard deviation.

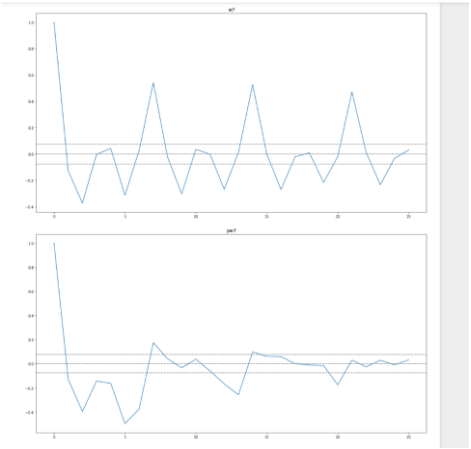


Fig.6 Representation of ACF and PACF.

6 Conclusion

Urban rail transport system is important in the construction of Urban Transport Infrastructure as rail transport is the most preferred method of transport in short distances. It is important to predict the passenger flow correctly because in some days passenger flow is very low like during festivals or public holidays. In this paper, a time series based algorithm on forecasting is suggested. The experimental results show that a better forecasting can be done prior to the travel.

References

1. Jianjun W, Huijun S, Ziyou G, Linghui H, Bingfeng S. Risk-based stochastic equilibrium assignment model in augmented urban railway network. Journal of Advanced Transportation 2014; 48(4): 332–347.

2. Mwakalonge J, Badoe D. Trip generation modeling using data collected in single and repeated cross-sectional surveys. Journal of Advanced Transportation 2014; 48(4): 318–331
3. Andreoni A, Postorino MN. A multivariate ARIMA model to forecast air transport demand. Proceedings of the Association for European Transport and Contributors, pages 1–14, 2006.
4. D. Gong, M. Tang, S. Liu, G. Xue, and L. Wang, ''Achieving sustainable transport through resource scheduling: A case study for electric vehicle charging stations,'' *Adv. Prod. Eng. Manage.*, vol. 14, no. 1, pp. 65–79, 2019.
5. M. Lippi, M. Bertini, and P. Frasconi, ''Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 871–882, Jun. 2013.
6. E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, ''Short-term traffic forecasting: Overview of objectives and methods,'' *Transp. Rev.*, vol. 24, no. 5, pp. 533–557, 2004.
7. M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said, ''Short-term prediction of traffic volume in urban arterials,'' *J. Transp. Eng.*, vol. 121, no. 3, pp. 249–254, 1995.
8. S. V. Kumar and L. Vanajakshi, ''Short-term traffic flow prediction using seasonal ARIMA model with limited input data,'' *Eur. Transp. Res. Rev.*, vol. 7, no. 3, pp. 1–9, Sep. 2015.
9. M. Van Der Voort, M. Dougherty, and S. Watson, ''Combining Kohonen maps with ARIMA time series models to forecast traffic flow,'' *Transp. Res. C, Emerg. Technol.*, vol. 4, no. 5, pp. 307–318, 1996.
10. D. Gong, S. Liu, and X. Lu, ''Modelling the impacts of resource sharing on supply chain efficiency,'' *Int. J. Simul. Model.*, vol. 14, no. 4, pp. 744–755, 2015.
11. S. Lee and D. B. Fambro, ''Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting,'' *Transp. Res. Rec.*, vol. 1678, no. 1, pp. 179–188, 1999.
12. B. Williams, ''Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling,'' *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1776, no. 2, pp. 194–200, 2001.
13. Sutar, S.H., Jinila, Y.B., "Congestion control for better performance of WSN based IoT ecosystem using KHA mechanism", International Journal of Recent Technology and Engineering, Volume-8, Issue-2S3, 2019
14. Jeba Blessingson, Y.Bevish Jinila, " Multi Utility / Tracing Kit for vehicles using RFID Technology", Recent Advances in Space Technology Services and Climate Change – 2010 (RSTS & CC-2010), PP. 273 – 276,2010
15. Y. Bevish Jinila, K. Komathy, "Rough Set Based Fuzzy Scheme for Clustering and Cluster Head Selection in VANET", ELEKTRONIKA IR ELEKTROTECHNIKA, Vol.21, No.1, pp.54-59, ISSN : 1392-1215, 2015
16. Kulkarni R V, Patil S H, Subhashini R, "An overview of learning in data streams with label scarcity, " Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016, 2017
17. Prayla Shyry S., Charan, V.S.K., Kumar, V.S," Spam Mail Detection and Prevention at Server Side", 2019 Innovations in Power and Advanced Computing Technologies (i-PACT),pp 1-6.