Metadata of the chapter that will be visualized in SpringerLink

Book Title	Smart Technologies in Urban Engineering	
Series Title		
Chapter Title	Intelligent Passenger Frequency Prediction System for Transportation Sustainability Using Convolutional Neural Network and Kalman Filter Algorithm	
Copyright Year	2024	
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG	

Author	Family Name	**Jimoh**
	Particle	
	Given Name	**Onemayin David**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Federal University of Technology Minna, Niger State
	Address	PMB 65, Minna, Nigeria
	Email	
	ORCID	http://orcid.org/0009-0009-4898-5092
Corresponding Author	Family Name	**Ajao**
	Particle	
	Given Name	**Lukman Adewale**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Federal University of Technology Minna, Niger State
	Address	PMB 65, Minna, Nigeria
	Email	ajao.wale@futminna.edu.ng
	ORCID	http://orcid.org/0000-0003-1255-752X
Author	Family Name	**Adeleke**
	Particle	
	Given Name	**Oluwafemi Oyetunde**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Federal University of Technology Minna, Niger State
	Address	PMB 65, Minna, Nigeria
	Email	
	ORCID	http://orcid.org/0000-0001-8272-7219
Author	Family Name	**Kolo**
	Particle	

	Given Name	**Stephen Sunday**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Federal University of Technology Minna, Niger State
	Address	PMB 65, Minna, Nigeria
	Email	
Author	Family Name	**Olarinoye**
	Particle	
	Given Name	**Oyedeji Abdulwaheed**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Federal University of Technology Minna, Niger State
	Address	PMB 65, Minna, Nigeria
	Email	
	ORCID	http://orcid.org/0000-0003-0932-0302

Abstract	The prediction of passenger flow operation is very significant to study due to the challenges of student transportation between inter-campuses of the Federal University of Technology Minna (FUTMinna), Nigeria. However, the prevailing technique of passenger flow estimation is non-parametric which depends on the fixed planning and is easily affected by noise. In this research, we proposed the use of a Convolutional Neural Network and Kalman Filter (CNN-KF) with an Auto-Regressive Integrated Moving Average (ARIMA) model for learning and prediction purposes of the passenger flow frequency on the inter-campuses arterial route. The passengers' frequency of arrival at the bus terminals are obtained and enumerated through the closed-circuit television (CCTV) and demonstrated using the Markovian Queueing Systems Model (MQSM). The autocorrelation coefficient functions (ACF) and partial autocorrelation coefficient functions (PACF) are used to examine the stationary data with different features. The performance of the models was analyzed and evaluated in describing the passenger flow frequency at each terminal using the Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) values. The CNN-Kalman-filter model was fitted into the series and the MAPE values are below 10%, more than 80% percent of the time reflecting the abnormal fluctuations of passenger flow accuracy than ARIMA. The Mean Square Error (MSE) shows that the CNN-Kalman Filter model has the overall best performance with 83.33% of the time better than the ARIMA model and provides high accuracy in forecasting.

Keywords (separated by '-')	ARIMA - Convolutional neural network - Kalman filter - Short-term prediction - Transportation sustainability

Intelligent Passenger Frequency Prediction System for Transportation Sustainability Using Convolutional Neural Network and Kalman Filter Algorithm

Onemayin David Jimoh, Lukman Adewale Ajao(✉), Oluwafemi Oyetunde Adeleke, Stephen Sunday Kolo, and Oyedeji Abdulwaheed Olarinoye

Federal University of Technology Minna, Niger State, PMB 65, Minna, Nigeria
ajao.wale@futminna.edu.ng

Abstract. The prediction of passenger flow operation is very significant to study due to the challenges of student transportation between inter-campuses of the Federal University of Technology Minna (FUTMinna), Nigeria. However, the prevailing technique of passenger flow estimation is non-parametric which depends on the fixed planning and is easily affected by noise. In this research, we proposed the use of a Convolutional Neural Network and Kalman Filter (CNN-KF) with an Auto-Regressive Integrated Moving Average (ARIMA) model for learning and prediction purposes of the passenger flow frequency on the inter-campuses arterial route. The passengers' frequency of arrival at the bus terminals are obtained and enumerated through the closed-circuit television (CCTV) and demonstrated using the Markovian Queueing Systems Model (MQSM). The autocorrelation coefficient functions (ACF) and partial autocorrelation coefficient functions (PACF) are used to examine the stationary data with different features. The performance of the models was analyzed and evaluated in describing the passenger flow frequency at each terminal using the Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) values. The CNN-Kalman-filter model was fitted into the series and the MAPE values are below 10%, more than 80% percent of the time reflecting the abnormal fluctuations of passenger flow accuracy than ARIMA. The Mean Square Error (MSE) shows that the CNN-Kalman Filter model has the overall best performance with 83.33% of the time better than the ARIMA model and provides high accuracy in forecasting.

Keywords: ARIMA · Convolutional neural network · Kalman filter · Short-term prediction · Transportation sustainability

1 Introduction

The current policy of the Nigeria Government discourages students' and workers' residency on campuses. This increases the population of both students and workers leaving off-campuses and far distance to the university location. Consequently, the population

O. Arsenyeva et al. (Eds.): STUE 2023, LNNS 807, pp. 1–13, 2024.
https://doi.org/10.1007/978-3-031-46874-2_39

of students who reside off-campus are become popular and is unlikely to be sustained in the future of Nigeria's higher institutions learning. However, university inter-campuses require suitable transport scheduling with a high concentration of management on the trip during the peak periods [1, 2]. Although, some universities globally are progressively promoting the implementation of sustainable transportation strategies at their campuses AQ1 [3].

This transportation sustainability requires adequate planning for the traffic flow in urban arterials, where bus shuttles are used as a public means of conveying passengers from one location to another through the road network [4, 5]. The passenger traffic flow is described as prevalent interactions between the travelers (passengers and vehicles) and the infrastructure (road networks, traffic control devices, or signage), for the optimal management of traffic congestion problems [6]. This traffic flow management can be achieved efficiently through timetable scheduling for passenger movement. This process can enhance the technical support and decision-making of the transport manager regarding the flow of passengers and control the bus dispatch system efficiently [7]. Understanding the travel behavior characteristics of a university community will assist greatly in the smooth running of inter-campus transportation sustainability, mobility AQ2 factors, and control of the traffic flow.

However, the randomness variation of short-term passenger flow makes it tough to analyze all its features with a single time-series model [8, 9]. This challenge leads to the use of combining different models that are linear and non-linear for better performances [10]. Therefore, this research justifies the adoption of a combined Convolution Neural Network (CNN) and Kalman Filter Algorithm (KFA) as a learning model to reduce the error rate and improve the speed during the prediction processes. This self-motivated learning model distinguishes between phenomena, and nominal to deduce the state of knowledge for nominal from the phenomena. This research contributes to knowledge by:

i. The CNN is used to extract, preprocess, and training of the data obtained through the closed-circuit television (CCTV) at the bus terminals for simplicity and to reduce the error rate.
ii. Analyzing the passenger flow frequency obtained for prediction using ARIMA and combined CNN-KFA.
iii. The ACF and PACF are used to investigate the stationary data with different features.
iv. The prediction model performance is evaluated using MAPE and MSE values.

The remaining part of this research is presented as follows. The related work is presented in Sect. 2. Section 3 presented the methodology, data collection process from the site survey, and prediction model of passenger frequency flow using the ARIMA model, Kalman filter, and Convolution Neural Network technique are presented in 3.1 and 3.2 respectively. Section 4 discussed results obtained from the study, and summarized the performance analysis using MAE, MAPE, and others as evaluation metrics. Section 5 conclude the research with future research recommendation.

2 Literature Review

Several efforts have been put together to optimize the stochasticity of traffic flows condition, passenger operations demand, and timetable schedule. Liang et al. [11] proposed ARIMA model for the prediction of passenger frequency in the Shanghai station. The observed historical events were taken within five weeks with a sampling interval of 10 min. It was reported that the time series has a cyclical and slow attenuation trend. The cyclical and trend phenomena were removed to obtain stationary series. The selected ARIMA (2,1) is used to predict the frequency of passengers at the station with an accuracy of 80%. Guo *et al.* developed a passenger frequency prediction model using a merging of support vector regression (SVR) and long short-term memory (LSTM) neural network techniques [12]. The passenger frequency information training was conducted using the LSTM model to train and show the large unsteadiness of irregular flow and its approximation. A fusion method based on the real-time prediction errors of the SVR outputs and LSTM is combined into the final computation of the prediction model. The hybrid model performance was evaluated using MAPE, RMSE, and MAE with average values of (12.59%, 68.54, and 40.06) respectively for a one-step-ahead process. The two-step ahead processes evaluation was recorded as (14.84%, 80.12, and 47.58) respectively. The comparative performance analysis shows that the SVR-LSTM model precisely mirrors the irregular unsteadiness of passenger frequencies, which performs well and produces greater forecast precision than single models.

Liu et al. proposed the hybridization of the Kernel Extreme Learning Machine (KELM) with wavelet transform (WT) for the development of a passenger flow prediction system [13]. The system achieved an appropriate breakdown of data about passenger flow in both high and low-frequency sequences with an MSE of 2982.277.

Zhang et al. presented the combination of a Residual Network (ResNet) for spatial correlation capture and temporal correlations using LSTM and Graph Convolutional Network (GCN) [14]. This approach was used for the extraction of network topology information in the prediction of passenger flow based on the short-term. The system achieved a Mean-Absolute Error (MAE) of 16.6318 but has poor interpretability. The adoption of a 3-Dimension CNN with a Graph Convolutional Network (GCN) for the extraction of spatiotemporal passenger flow features is developed. The performance of the system is evaluated using root means squared (RMS) error of 9.402%.

The prediction of passenger frequency is optimized through the learning network with Adam's long and short-term memory network (Adam-LSTM) in urban rail transit. The system ensured optimal prediction with a performance of 7.14% Mean Relative Error (MRE). Li et al. [10] developed a forecasting model for a multi-source data set of rail transit in an urban city of Beijing using SARIMA and SVM for the training and prediction analysis model to establish the traffic flow. The model utilizes an intelligent data acquisition system integrated with sensors and networks-based Internet of Things (IoT) for remote data collection on a large scale with precise passenger frequency data. The model is suitable for complex data, nonlinearity in nature, and periodic in urban rail transit. The obtained prediction outcome fits well with the measured data and is suitable for short-term passenger frequency forecasting.

AQ3

3 Research Methodology

This research focuses on passenger flow monitoring and prediction analysis to manage the inter-campuses bus shuttle and the passenger queue system. The model uses a convolutional neural network as a learning stage and a Kalman filter algorithm for the prediction resolutions. This process is achieved through the analysis monitoring of passenger flow during the week (Monday-Friday) and between 6:30 am to 6:30 pm of the days. The information is collected through the mounted CCTV camera installation at the bus terminals, and the historical data clips of the passenger arrival frequency were collected for the extraction, processing, and training using a CNN.

The CNN algorithm is used for the extraction of passenger headcount, and training, to learn the hidden pattern for the simplified input and output parameters relationship in the model. The output (preceding information) of learning or called supervising model is fed directly into the prediction model (Kalman Filter) for better prediction accuracy (Fig. 1). This prediction model is controlled by Kalman gain for tunning the parameters and updated immediately after each iteration using covariance matrix (Z) and error estimation. Although, the CNN is programmed with R language for data extraction, processing, and training to prevent multi-dimensional objectives in this concept. The future variations were studied and simulated in MATLAB using the CNN-Kalman Filter and ARIMA model.

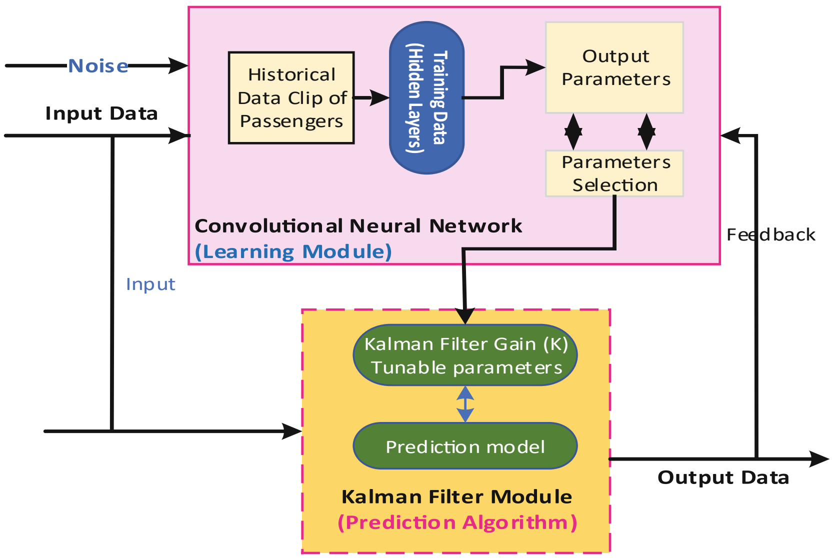


Fig. 1 Hybrid architecture for the passenger frequency prediction

3.1 Data Collection Process in the Site Survey

A site survey was carried out to determine the traffic flow at the geo-location of inter-campuses bus terminals by considering the vehicular speed, density, and passenger flow.

The transit route of FUTMinna (Bosso–Kpakungu–Gidan Kwano) has been surveyed and digitized to capture potential bus parks, markets, and public facilities within the corridor. The facilities required at each terminal include an in-vehicle intelligent transportation system, CCTV monitoring, and recording system, a mounted visual display unit powered with solar energy, and many others. Real-time passenger flow frequency data of both terminals (Bosso and Gidan-Kwano) of the inter-campus university route was collected using a CCTV recording system installed at the stations. The data recording was analyzed to estimate the frequency of passengers arriving at the terminals for the intervals of 15 min from 6.00 am to 6.00 pm for both stations (see Fig. 2). This data recording process covered two weeks between 25th September and 4th October 2019.

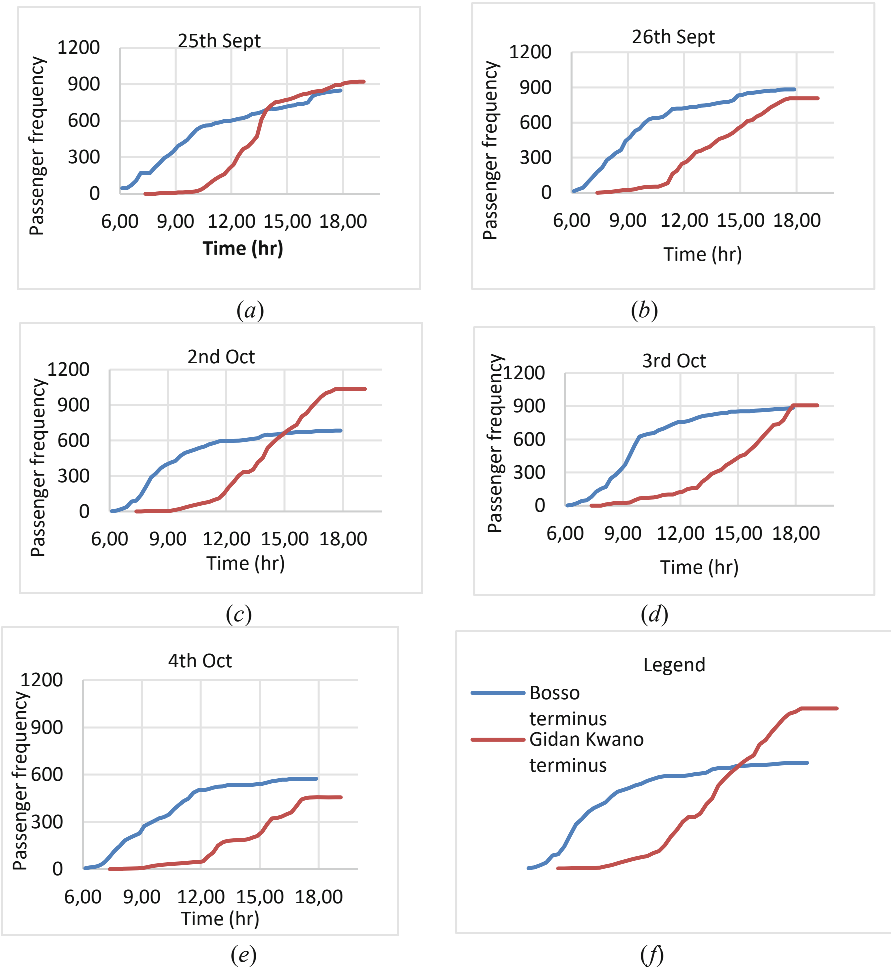


Fig. 2 Frequency curve of passenger traffic flow between 25th to 4th October 2019

However, the passenger transitions that occur between the arrival rate λ and departure rate μ at the bus terminal of k are duly considered by assuming the average arrival as $(\lambda = \lambda_1, \lambda_2, \lambda_3, \lambda_4, \ldots \lambda_k)$, and departure as $\mu = \mu_1, \mu_2, \mu_3, \mu_4, \ldots \mu_k)$ per unit time to the queue. Therefore, the state equation for the passenger arrival time and departure time can be monitored using probability in state n. Therefore, the state probability of arrival and departure rate can be established as in Eq. (1), when the condition is tending to infinity is expressed as (2). The adjustment of complete state probability of the passenger flow (arrival and departure) is given in (3) when $\mathrm{P}_n(n \geq 1)$

$$\mathrm{P}_n = \frac{\lambda_{1-1}\lambda_{n-2}\lambda_{n-3}\ldots\lambda_0}{\mu_n\mu_{n-1}\mu_{n-1}\ldots\mu_1}\mathrm{P}_0 = \mathrm{P}_0\prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}}, \tag{1}$$

$$\sum_{n=0}^{\infty}\mathrm{P}_n = \mathrm{P}_0 + \mathrm{P}_0\sum_{n=0}^{\infty}\prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}} = 1, \tag{2}$$

$$\mathrm{P}_0 = \frac{1}{1 + \sum_{n=0}^{\infty}\prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}}}. \tag{3}$$

We assume that, if the arrival mean rate is λ then $\mathbb{C}_\theta^2 = \varrho_\theta^2/(1/\lambda)^2$, and the service rate is μ, then $\mathbb{C}_x^2 = \varrho_x^2/(1/\mu)^2$. The waiting time of the passenger on the queue can be perfectly approximate as expressed given in Eqs. (4, 5, 6).

$$\delta_{nq} \approx \frac{\varrho^2(\mathbb{C}_\theta^2 + \mathbb{C}_x^2)}{2(1-\varrho)}g, \tag{4}$$

where,

$$g = exp\left(\frac{2(1-\varrho)\left(1+\varrho^2\mathbb{C}_\theta^2\right)^2}{3\varrho\left(\mathbb{C}_\theta^2 + \mathbb{C}_x^2\right)}\right), when\mathbb{C}_\theta^2 \leq 1; \tag{5}$$

$$g = exp\left(\frac{(1-\varrho)\left(1-\mathbb{C}_\theta^2\right)^2}{\left(\mathbb{C}_\theta^2 + 4\mathbb{C}_x^2\right)}\right), when\mathbb{C}_\theta^2 > 1. \tag{6}$$

Then, the passenger waiting time and departure time determinants can be derived as expressed in "Eqs. (7, 8)"

$$\wp_0 = \frac{1}{\left[\sum_{\mathcal{M}=0}^{c-1}\frac{(c\varrho)^{\mathcal{M}}}{\mathcal{M}!} + \frac{(c\varrho)^c}{c!(1-\varrho)}\right]} \tag{7}$$

$$\varpi_q = \frac{\delta_{nq}}{\lambda}. \tag{8}$$

3.2 Prediction Model for the Passenger Frequency

A stochastic time series model $(y_1, y_2, y_3, \ldots, y_t)$ for ARIMA was used in the forecasting of passenger frequency information observation at the bus terminals of the university inter-campuses. This stochastic model varies every time and predicts future values with a confidence interval around the predictions. The autoregressive part model is p[AR], the degree of differencing order in the model is d[I], the order of the moving average part of the model is q[MA]. The difference of the part model is D, the autoregressive seasonal arrangement of the model is P, while Q is the order of the moving average seasonal part of the model, and *s* is the period of the model.

If $y_t = \{X_t\}$ is in series with mean μ, the series of ARIMA model is $(p, d, q)(P, D, Q)^s$, and can be expressed as given in "Eqs. (9, 10)". These ARIMA models of *(p, d, q)* $(P, D, Q)^s$ are autoregressive, differencing, and moving-average respectively, *s* is the number of the observation periods. The β represent backshift notation and can be expressed as $\beta^{\kappa} y_t = y_{t-\kappa}$, the estimation of ARIMA coefficient are $\phi(\beta) = \left(1 - \phi_1 \beta^s, \ldots, \phi_\rho \beta^{s*v}\right)$, and z_t is a noise process error assumed, and σ^2 is a covariance of the forecasting model of $y_t = \beta r_t + \eta_t$. The β represent regression coefficient, η_t represent the regression error in the ARIMA method of forecasting.

$$\begin{cases} Y_t = (1-\beta)^d (1-\beta^s)^D X_t - \mu \\ \phi(\beta)\Phi(\beta^s) Y_t = \theta(\beta)\theta(\beta^s) z_t, z_t \sim N\left(0, \sigma^2\right) \end{cases} \tag{9}$$

$$\begin{cases} \phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i, \theta(z) = 1 - \sum_{i=1}^{P} \theta_i z^i \\ \theta(z) = 1 - \sum_{i=1}^{q} \theta_i z^i, \theta(z) = 1 - \sum_{i=1}^{Q} \theta_i z^i \end{cases} \tag{10}$$

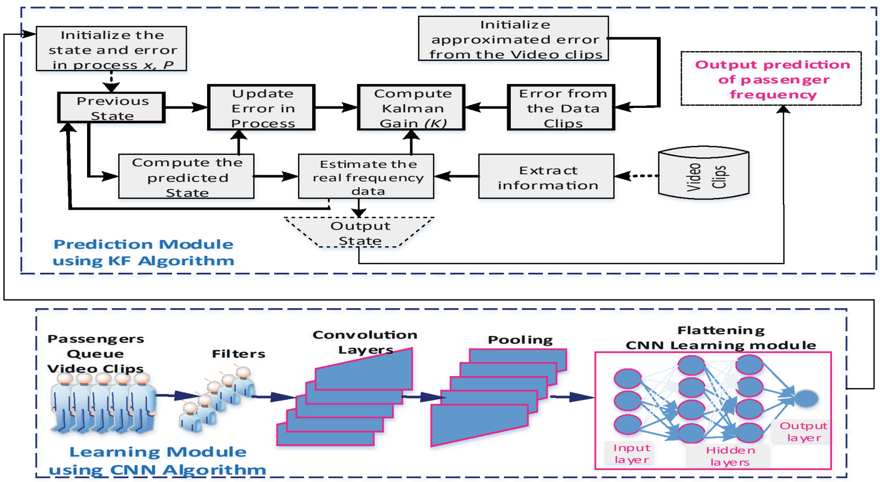


Fig. 3. Overview of the hybrid model for the passenger frequency prediction

Also, the time-series for the hybrid prediction model using CNN-KFA was demonstrated. The convolution neural network (CNN) acts as an intelligent learning module for the passenger features extraction and analysis. The extraction of the passenger's video clip is collected from the storage device which feeds into the R programming environment for the data cleaning (removing any outlier and filling of the missing parameters). After the preprocessing (data cleaning), the dataset is subjected to training and validation which was evaluated using MSE and MAE. While the modeling of the output dataset is fed into the KF algorithm for the prediction resolutions (see Fig. 3).

The state vector and its associated covariance are predicted by Kalman, while the state vector and its covariance are corrected using the measured data. Since we are modeling time series, the observation model was formulated as an autoregressive model of order 2 AR(2) and embedded in the KF model to resolve the non-stationary time series data as presented in Table 1.

This KFA model is a well-known recursive approach to the linear data challenges, and provides efficient computation to estimate the state processes, by minimizing the MSE. It helps in the estimation of $(mx1)$ state vector x_t, , which is usually a vector of parameters that are not directly observable. It does this by recursively updating the state of the dynamic systems (or state vector).

However, The Kalman observation model can be expressed as in "Eqs. (11, 12, 13, 14), where, (φ_1, φ_2) are the state vector of AR(2) parameters.

$$y_t = \varphi_1.y_{t-1} + \varphi_2.y_{t-2} + v, \tag{11}$$

$$H = \left[y_{t-1,}y_{t-2} \ldots y_{t-p}\right]^T, \tag{12}$$

$$H = \left[y_{t-1,}y_{t-2}\right]^T, \tag{13}$$

$$x_t = [\varphi_1, \varphi_2]. \tag{14}$$

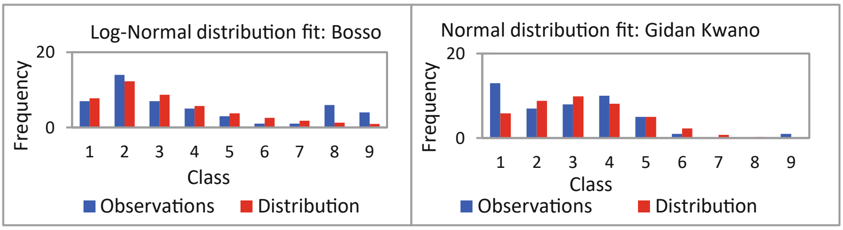
4 Results and Discussion

The passenger arrival frequency framework was evaluated under the conditions of characteristics distribution of Normal, Log-normal, Logistic, and Gamma functions respectively. The performance of the frequency passenger flow was assessed using a statistics p-value for the hypothesis testing. The p-value conveys information about the weight of evidence against the null hypothesis (H_0). The null hypothesis states that the frequency of passengers arriving at a bus terminal follows a normal distribution will reject the hypothesis if the weight of evidence (p-value) against H_0 is not significant. That is, if smaller than the significant level of 0.05 ($\alpha = 5\%$), and accept if the p-value is greater than the significant level. The smaller the P-value, the greater the evidence against H_0. . The frequency of passenger flow at both terminals is recorded, which varies with the days and time. The result analysis shows that Log-Normal distribution fits Bosso passenger flow, and Normal distribution fits Gidan-Kwano passenger flow (see Fig. 4). So, this

Table 1 Kalman filter algorithm model

Steps	Procedures
1	The Kalman system model is formulated with the dynamic and observation model thus: $x_t = F_x.x_{t-1} + F_n.n$ (Dynamic model) $y_t = H.x_t + v$ (Observation model) where x_t is the state vector of the current time step; x_{t-1} is the state vector of the previous time step; n is the perturbation vector (noise for the dynamic model with zero); y_t is the measurement vector; v is the measurement noise with zero mean; F_x is the transition matrix; F_n is the perturbation matrix; H is the measurement matrix; H_t is the transformation matrix into the measurement domain for the state vector parameters mapping n and v are white noise that independent of each other, and they follow normal probability distributions of $p(n) \sim N(0, Q)$, and $p(v) \sim N(0, R)$ where Q and R are covariance matrix processing noise
2	Initialize F_x, F_n, H, and x_t, then initialize covariance matrices P, Q, and R
3	Begin the time loop by a predict the state vector x_t with its covariance P, and correct the x_t with its covariance P by computing the following expressions: – Expectation, $e = H.x_t$ – Covariance of expectation, $E = H.P.H\prime$ – Innovation, $z = y - e$ – Covariance of innovation, $Z = R + E$ – Kalman gain, $K = P + H\prime * Z^{-1}$ $K = P + H\prime * Z^{-1}$ then, calculate the new x_t and P as expressed in "(15)–(16)" $x_t = x_t + K * P$ (15) $P = P - K * H * P$ (16) The standard Kalman filter algorithm model with the transformation matrix H_t to allow the prediction and measurement are summarized from Kalman gain in "(17)–(18)" $K = \frac{H\sigma_1^2}{H^2\sigma_1^2 + \sigma_2^2} \rightarrow K_t$ (17) $K = P_{t\|t-1} H_1^2 (H_t P_{t\|t-1} H_t^T + R_t)^{-1}$ (18) The conversion of the transformation matrix into the domain for the prediction and measurement is expressed as in "(19)–(20)" $\hat{x}_{t\|t} = \hat{x}_{t\|t-1} + K_t(z_t = H_t\hat{x}_{t\|t-1})$ (19) $P_{t\|t} = P_{t\|t-1} - K_t H_t P_{t\|t-1}$ (20)
4	End the time loop
5	Plot results

means that more commuters are arriving at Bosso terminal than Gidan-Kwano terminal within the period 5:00 GMT and 16:00 GMT. However, there are more commuters at Gidan Kwano terminal after 16 GMT hoping to return to their destinations. The passenger frequency at Bosso terminal is high as 250 by 6:00 GMT and rises to 650 by 13.00 GMT of the day. The situation at Gidan Kwano terminal differs from that of Bosso terminal where passenger flow of about 110 at 12.30 GMT with a peak of 900 at 17:30 GMT.



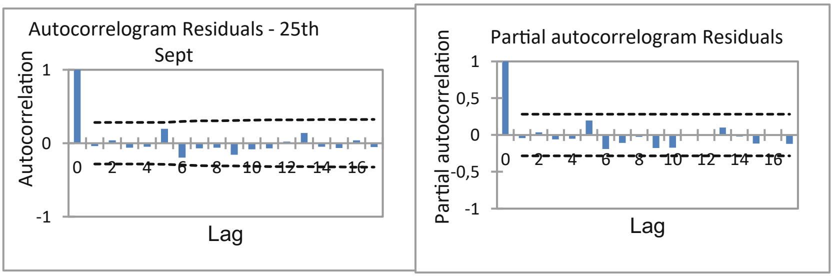
(*a*) (*b*)

Fig. 4 Log-normal fitted distribution and normal fitted distribution

However, the analysis of the time-domain is primarily used for passenger frequency prediction. A stationary time series model with autocorrelation is adopted in the data analysis of short-term passenger frequency. This includes autoregressive process with order p AR(p), moving average process with order q MA(q), and autoregressive moving average process with orders p and q (ARIMA (p, q)).

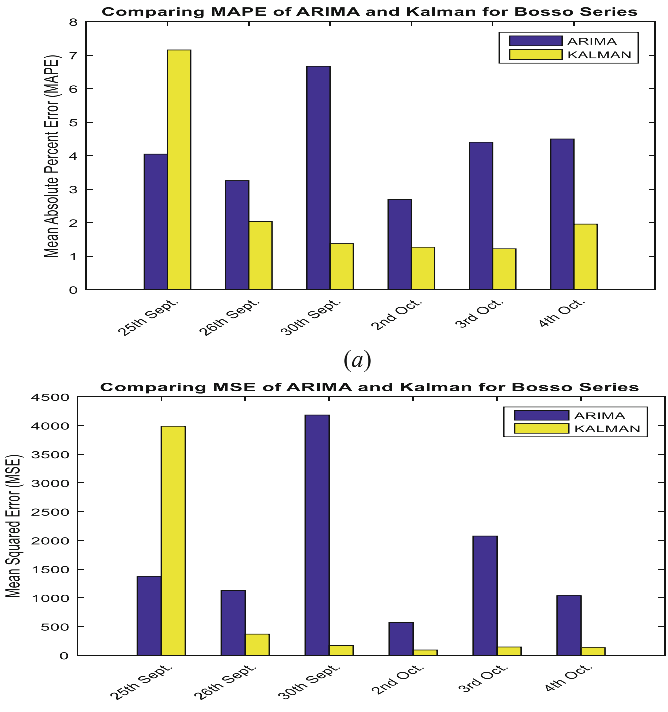
The parameter of the selected model is observed using the differential operation method to transform the non-stationary series to a stationary one, and measured with the ACF and PACF with different features. The prediction analysis of Bosso terminal shows the ARIMA models that performed best are ARIMA (2,1,1), ARIMA (3,1,1), ARIMA (3,1,3), ARIMA (1,1,4), and ARIMA (1,1,1) model for the respective days between 25th Sept., 26th Sept., 30th Sept., 2nd Oct., 3rd Oct. and 4th Oct. Similar analysis was carried out for Gidan-Kwano terminal and the best model is ARIMA (1,1,1), ARIMA (1,1,1), ARIMA (1,1,1), ARIMA (1,1,1), ARIMA (1,1,1), and ARIMA (2,1,1) for the respective days. The auto correlogram residuals and partial auto correlogram residuals of the series in Bosso is presented (see Fig. 5).

So, the observed time-series data were used to estimate the parameters for the ARIMA and CNN-KFM, these parameters are then used to generate predictions beyond the last instance of the observed time-series data. The prediction values of both ARIMA and CNN-KFM were compared with the test set values obtained from the original series using Mean Absolute Percent Error (MAPE) and the Mean Squared Error (MSE). The comparative performance of the proposed model for passenger flow prediction at the Bosso terminal and GK bus station using ARIMA and CNN-KFM techniques are graphically presented (see Fig. 6).



(*a*) (*b*)

Fig. 5 Time series analysis in the terminals using **a** ACR, and **b** PACR



(*b*)

Fig. 6. The passenger frequency comparative analysis using **a** MAPE, and **b** MSE

5 Conclusion

The passenger flow prediction for the transportation sustainability and planning of inter-campuses shuttle on the city arterial route was investigated and analyzed for the future preparation of time-table schedule. A Markova model was used in analyzing the arrival of the passenger flow frequency and queue system through the stationary CCTV position at the bus stations which help to manage the travel pattern and allow recognizing the population group, patronage period, and for the proper planning of the timetable. The Auto-Regressive Integrated Moving Average (ARIMA) and CNN-KFM were used for learning and prediction purposes. The performances are evaluated using MAPE and MSE statistical analysis. The MAPE results for the Bosso terminal using ARIMA and CNN-KFM model shows that CNN-KFM has better performance with 83.33% times than the ARIMA model. Similar results were obtained for the Gidan-Kwano terminal which shows that CNN-KFM has the overall best performance in terms of the Mean Absolute Percent Error (MAPE) with 66.66% times the ARIMA model. However, the two models are suitable for forecasting passenger frequencies at the Bosso and Gidan Kwano bus terminal, since 85.5% of the time, MAPE results for the two models are below 10% which showed high accuracy. However, the investigation will help to control the environmental pollution and other related factors that may result in mass travelers overstaying at the bus terminals. The future works will focus on the development of time-table scheduling for public transportation and, transport fare collection system using the genetic programming concept of a meta-heuristic algorithm to enhance the transportation sustainability prediction.

Acknowledgment. This work was funded by Federal Republic of Nigeria under the National Research Fund of Tertiary Education Trust Fund (TETFund) DR&D/CE/NRF/STI/34/VOL1.

References

1. Rérat P (2021) A campus on the move: modal choices of students and staff at the University of Lausanne, Switzerland. Trans Res Inter Persp 12:100490. https://doi.org/10.1016/j.trip.2021.100490
2. Suman HK, Bolia NB, Tiwari G (2017) Comparing public bus transport service attributes in Delhi and Mumbai: policy implications for improving bus services in Delhi. Trans Pol 56:63–74. https://doi.org/10.1016/j.tranpol.2017.03.002
3. Hamad K, Htun PTT, Obaid L (2021) Characterization of travel behavior at a university campus: a case study of Sharjah University City, UAE. Trans Res Inter Persp 12:100488. https://doi.org/10.1016/j.trip.2021.100488
4. Alkhereibi AH, Tahmasseby S, Mohammed S, Muley D (2021) Blue collar laborers' travel pattern recognition: machine learning classifier approach. Transp Res Inter Persp 12:100506. https://doi.org/10.1016/j.trip.2021.100506
5. Ajao LA, Apeh ST (2023) Secure edge computing vulnerabilities in smart cities sustainability using petri net and genetic algorithm-based reinforcement learning. Int Sys App 200216. https://doi.org/10.1016/j.iswa.2023.200216
6. Ajao LA, et al (2020) In-vehicle traffic accident detection and alerting system using distance-time based parameters and radar range algorithm. In: 2020 IEEE PES/IAS PowerAfrica. IEEE, pp 1–5. https://doi.org/10.1109/PowerAfrica49420.2020.9219812
7. Jimoh OD, Ajao LA, Adeleke OO, Kolo SS (2020) A vehicle tracking system using greedy forwarding algorithms for public transportation in urban arterial. IEEE Access 8:191706–191725. https://doi.org/10.1109/ACCESS.2020.3031488
8. Wang B, Ye M, Zhu Z et al. (2020) Short-term passenger flow prediction for urban rail stations using learning network based on optimal passenger flow information input algorithm. IEEE Access 8:170742–170753. https://doi.org/10.1109/ACCESS.2020.3024224

9. Zhao J, Ye M, Yang Z et al (2021) Operation optimizing for minimizing passenger travel time cost and operating cost with time-dependent demand and skip-stop patterns: nonlinear integer programming model with linear constraints. Tra Res Int Pers 9:100309. https://doi.org/10.1016/j.trip.2021.100309
10. Li W, Sui L, Zhou M et al (2021) Short-term passenger flow forecast for urban rail transit based on multi-source data. J Wir Com Net 2021:9. https://doi.org/10.1186/s13638-020-01881-4
11. Liang S, Ma M, He S, Zhang H (2019) Short-term passenger flow prediction in urban public transport: Kalman filtering combined k-nearest neighbor approach. IEEE Access 7:120937–120949. https://doi.org/10.1109/ACCESS.2019.2937114
12. Guo J, Xie Z, Qin Y et al (2019) Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM. IEEE Access 7:42946–42955. https://doi.org/10.1109/ACCESS.2019.2907739
13. Liu R, Wang Y, Zhou H et al (2019) Short-term passenger flow prediction based on wavelet transform and kernel extreme learning machine. IEEE Access 7:158025–158034. https://doi.org/10.1109/ACCESS.2019.2950327
14. Zhang J, Chen F, Guo Y et al (2020) Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit. IET Intel Tran Sys 14(10):1210–1217. https://doi.org/10.1049/iet-its.2019.0873

Author Proof

Author Queries

Chapter 39

Query Refs.	Details Required	Author's response
AQ1	Please be aware that your name and affiliation and if applicable those of your co-author(s) will be published as presented in this proof. If you want to make any changes, please correct the details now. Please note that after publication corrections won't be possible. Due to data protection we standardly publish professional email addresses, but not private ones, even if they have been provided in the manuscript and are visible in this proof. If you or your co-author(s) have a different preference regarding the publication of your mail address(s) please indicate this clearly. If no changes are required for your name(s) and affiliation(s) and if you agree with the handling of email addresses, please respond with "Ok".	
AQ2	Please provide appropriate orcid "http://orcid.org/0000-0002-6495-27551".	
AQ3	As references [8] and [15] are the same, we have deleted the duplicate reference and renumbered accordingly. Please check and confirm.	