Linear regression models with autoregressive integrated moving average errors for measurements from real time kinematics-global navigation satellite system during dynamic test

Kok Mun Ng[1], Ravenny Sandin Nahar[1], Mamun IbneReaz[2]

[1]Vehicles Intelligence and Telematics Lab, School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
[2]Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Article Info	**ABSTRACT**

The autoregressive integrated moving average (ARIMA) method has been used to model global navigation satellite systems (GNSS) measurement errors. Most ARIMA error models describe time series data of static GNSS receivers. Its application for modeling of GNSS under dynamic tests is not evident. In this paper, we aim to describe real time kinematic-GNSS (RTK-GNSS) errors during dynamic tests using linear regression with ARIMA errors to establish a proof of concept via simulation that measurement errors along a trajectory logged by the RTK-GNSS can be "filtered", which will result in improved positioning accuracy. Three sets of trajectory data of an RTK-GNSS logged in a multipath location were collected. Preliminary analysis on the data reveals the inability of the RTK-GNSS to achieve fixed integer solution most of the time, along with the presence of correlated noise in the error residuals. The best linear regression models with ARIMA errors for each data set were identified using the Akaike information criterion (AIC). The models were implemented via simulations to predict improved coordinate points. Evaluation on model residuals using autocorrelation, partial correlation, scatter plot, quantile-quantile (QQ) plot and histogram indicated that the models fitted the data well. Mean absolute errors were improved by up to 57.35% using the developed models.

Corresponding Author:

Kok Mun Ng
Vehicle Intelligence and Telematics Lab, School of Electrical Engineering, College of Engineering,
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
Email: ngkokmun@uitm.edu.my

1. INTRODUCTION

Autonomous vehicles and robots rely on localization technologies to determine their current location and pose. This will subsequently aid the vehicle or robot path planning and navigation. Localization technologies such as fused measurements from sensors, cooperative positioning and range based methods have been developed with the aim to achieve high level of positioning accuracy [1]–[3]. However, the sensors employed in these localization systems produced measurement errors such as offset drifts in inertial sensors and single-point solution in global positioning system (GPS).

In this paper, we focus on the GPS, which often suffer from various noise sources that degrades its positioning accuracy. To enhance positioning accuracy, the development of real time kinematics (RTK) and the global navigation satellite system (GNSS) resulted in various RTK-GNSS devices that could provide

centimeters positioning [4], [5]. Centimeter positioning (between 1 to 5 centimeters) in RTK-GNSS is known as fixed-integer solution, which indicates that the receiver calculated a correct solution with minimal dilution. However, performance evaluation on five different brands of low-cost RTK-GNSS, tested in rural, sub-urban and urban landscapes revealed that they could not hold fixed-integer solution for any significant time in dynamic applications [5].

The RTK-GNSS may be subjected to noisy sources such as atmospheric and ionospheric disturbances [6], multipath errors [7], [8], canopy trees [9], and other noise sources that may be hard to distinguish. These noise components are usually time-correlated and inherit unique stochastic attributes. In recent years, error variation in GNSS data caused by weather conditions provided useful instability indices for weather forecasting such as rainfall and thunder storms [10]–[14]. GNSS error modeling was conducted with various machine learning techniques such as neural network [10], [11], classification functions [10], random forest [13], [14] and support vector machines [14] to forecast weather conditions. Hence, robust machine learning and signal processing techniques with noise models that can describe these time-correlated and stochastic characteristics are needed to process GNSS data for their related applications.

Throughout the decades, the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) have been used to model the time series of position changes at GPS static stations, which were used for monitoring of structural health and seismic activities. ARIMA and ARMA models could provide robust estimation of the time correlation in the noise compared to white noise model which may leads to biased in parameter estimates. Barba *et al.* [15] stated that ARIMA produces smoother curves and is more effective in forecasting GNSS time series compared to ARMA and wavelet methods. Hohensinn *et al.* [16] described the random noise of static GNSS time series data by mean of an ARIMA model to perform pre-filtering of GNSS noise. Subsequently, periodograms were used to detect the presence of harmonics in the data to assist detection of vibration in structures. Špoljar *et al.* [17] attempted to model multipath errors in GNSS time series positions from a stationary GNSS receiver in Huawei P20 Mate smartphone using autoregressive (AR) models. However, the effects caused by multipath were not modeled effectively. The noise may need to be further investigated using ARMA. In another work, Tao and Bonnifait [18] modelled the errors of a standalone L1-GPS used in autonomous vehicle. An autoregressive AR process was used to describe the errors. The Burg method was used to estimate the AR model parameters to realize a shaping filter for the GPS errors. The positioning errors are further improved by integrating front-view camera and map matching method.

On another note, ARMA and ARIMA have been used to model the overall stationary GNSS time series to assist in prediction and forecasting. Zhang *et al.* [19] fitted an ARMA model on GNSS time series data of a static GNSS station. The absolute outliers (AO) in the ARMA model were used for the detection of cycle slips in the GNSS carrier-phase observations. The authors developed a Bayesian method to detect AOs in the ARMA model that resulted in cycle slips detection in GNSS carrier-phase observations. Kaloop *et al.* [20] fitted an ARMA model to the time series measurements of a geodetic monitoring GPS system which was used to assess the safety behavior of a long span bridge. The coefficients and model errors predicted from the ARMA model were used to evaluate the semi-static and dynamic movement of the bridge. The model errors and parameters were used to study the safety of the bridge during GPS measurements. Recently, ARIMA was integrated with other machine learning methods for forecasting in [21]–[23]. Salma *et al.* [24] combines variational mode decomposition techniques with ARMA on GNSS time series data to forecast ionospheric delay. In another work, Xin *et al.* [25] combines Kalman filter, ARIMA and GARCH to process measured sensor data from GNSS attached to a bridge structure. The linear recursive ARIMA model was established to predict the bridge structure deformation.

It was shown in previous works the application of ARMA and ARIMA that could describe the correlated noise of the GNSS measurements [15]–[18] and could be used as forecasting methods [19], [20], [24], [25]. Soundy *et al.* [26] discovered that AR and ARMA have better performance compared to Gaussian white noise model, low-order MA process and the Ornstein-Uhlenbeck model when evaluated using the Akaike information criterion (AIC). The authors stated that "Directly estimating statistical parameters of the GPS time series by explicitly taking into account the heteroscedastic space and time-correlated error structure of the observations could result in improved accuracy at minimal computational cost" [26]. This demonstrated the usefulness of the AR, ARMA and the ARIMA method in describing the stochastic behavior in GNSS measurements. Recent developments, observed the application of ARIMA to predict clock biases and ionospheric errors in GNSS [27], [28].

This paper addressed the need to model RTK-GNSS dynamic measurement errors which were contributed by various noise sources especially during dynamic applications [4], [5]. Though, existing ARMA and ARIMA methods were successful in modelling GNSS time series errors; the applications were mainly focus on static GNSS stations [15]–[20], [24], [25]. To our knowledge, research using purely the regression with ARIMA errors to describe trajectory mapped by RTK-GNSS during dynamic tests has not been conducted. Most works employed state estimation methods such as Kalman filter based on

measurements from various sensors assuming certain noise models [29]–[33]. In this work, we aim to describe noise in RTK-GNSS during dynamic test or when it is used for applications with dynamic movement and measurements such as that in autonomous vehicle/robot using linear regression model with autoregressive integrated moving average (ARIMA) errors. This work also contributes to an experimental proof of concept via simulations that noisy RTK-GNSS measurements along a trajectory can be "filtered" using our proposed method. This will lead to improved accuracy in positioning. Initial performance of the fitted model was evaluated by comparison with accuracy of positioning before model implementation. Section 2 outlines the methods employed in this work starting with data collection and preliminary analysis on the data and, further describe model development of linear regression models with ARIMA errors on RTK-GNSS time series obtained under dynamic tests. The methods to evaluate model fitting and performance are also elaborated. Section 3 presents the results and discussion. Concluding remarks which include future studies were given in section 4.

2. METHOD
2.1. Data collection and initial analysis

This work employs the EMLID Reach RTK-GNSS as shown in Figure 1(a) to implement dynamic data logging. One unit of the EMLID Reach was set up as the base station whereas another unit was set up as the rover, attached to the roof top of a moving vehicle as shown in Figure 1(b). The base station sends phase carrier corrections continuously to the rover via LoRa radio frequency in the range of 868-1,100 MHz and air data rate of 0.81-18.23 kb/s. By using the *ReachView* app, radio frequency was set at 912 MHz and satellite system such as GPS, GLONASS, BeiDou and Galileo were selected to obtain the best solution. The rover receives the correction from the base and calculates its best position.
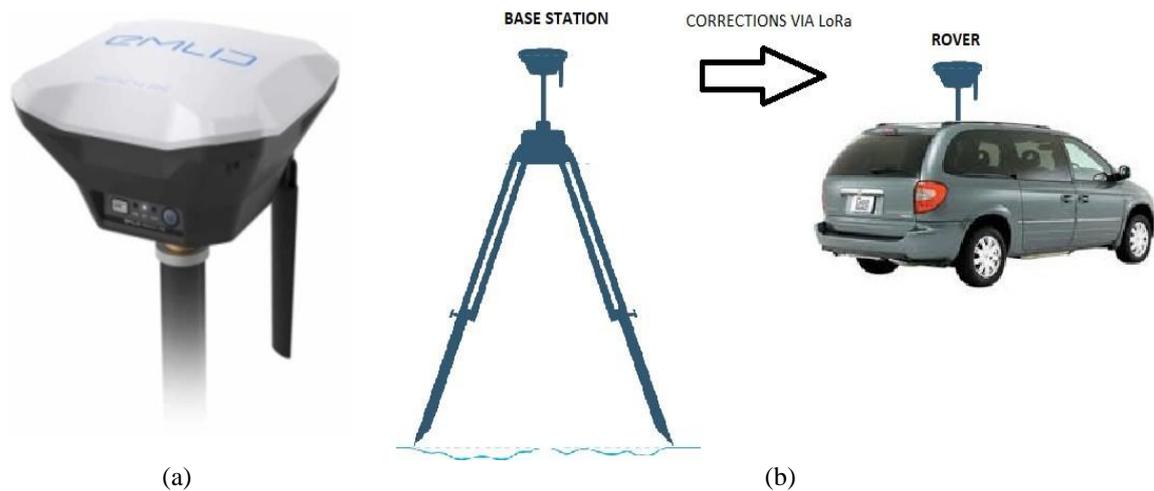

(a) (b)

Figure 1. RTK positioning (a) EMLID Reach RS and (b) setting up base and rover units

The EMLID Reach RTK-GNSS was deployed at a sub-urban location in the city of Shah Alam, Malaysia. Figure 2 shows the trajectory (red path) travelled by the rover for dynamic measurements in this location. It encircled some residential houses, trees and structures that may contribute multipath effects to the RTK-GNSS. The rover received carrier phase corrections from the base station (marked 'X') which was placed at a nearby football field. The baseline of the rover to the base station ranges from 60-200 meters. Data logging was conducted during fine weather and clear sky condition. Three sets of data were collected on August 8, 2020. It took approximately 2 minutes to complete data collection for the whole trajectory at a sampling rate of 0.5 Hz.

A geo-referenced map was developed for this location using quantum geographic information system (QGIS) software Desktop version 3.16.2. Fourteen fixed coordinate points were collected using EMLID RTK-GNSS as geo-referenced points to construct an accurate map for the area. Ground truth trajectory was derived using the geo-referenced map. The data collected was compared with ground truth trajectory to determine mean absolute error (MAE) and root mean square error (RMSE) of positioning. Autocorrelation, partial correlation, and histogram of the residuals were analyzed.

Figure 2. Trajectory of the rover

2.2. Model development using linear regression model with ARIMA errors

Figure 3 illustrates the methodology to perform linear regression with ARIMA errors on the trajectory geodetic points. The RTK-GNSS measurements and the ground truth which contains geodetic coordinates were initially converted to north-east-down (NED) coordinates (Y_{north}-X_{east}-Z_{down}). This conversion is necessary so that model fitting can be divided to two dimensions (i.e., the x and y axes respectively). A linear regression relationship between logged X_{east} points with the X_{east} ground truth was established for the three data sets collected. Similar linear regression was also established for logged Y_{north} points with the Y_{north} ground truth. Residuals of these regression models were analyzed in terms of autocorrelation function (ACF) and partial correlation function (PACF) plots.

Subsequently, linear regression models with ARIMA errors for each trajectory data set were estimated from the residuals ACF and PACF plots based on unique sets of ACF and PACF relations outlined in [34]. The linear regression model with ARIMA errors is defined in (1) where y_t is the response series; x_t is a series of predictor data; β is the regression coefficient; c is the regression intercept and μ_t is the ARIMA errors disturbance series at t=1, ..., N, and N the number of samples. The series μ_t comprised of autoregressive coefficients $\alpha_1, ..., \alpha_p$ up to degree p and moving average coefficients $\emptyset_1, ..., \emptyset_p$ up to degree q, where ε_t is the white noise series and $L^j f_t = f_{t-j}$. The $(1 - L)^D$ is the degree D integration polynomial.

$$y_t = \beta x_t + c + \mu_t$$
$$\left(1 - \alpha_1 L - \cdots - \alpha_p L^p\right)(1 - L)^D \mu_t = (1 + \emptyset_1 L + .. + \emptyset_q L^q)\varepsilon_t \qquad (1)$$

The logged points for x-axis (i.e., X_{east} coordinates) and y-axis (i.e. Y_{north} coordinates) respectively were used as predictors and ground truth data was used as response series to estimate the models. Best fitted linear regression models with ARIMA errors were identified for both x and y axes respectively using the Akaike information criterion (AIC). Based on (1), linear regression with ARIMA errors models were developed for x-axis coordinates as shown in (2) and y-axis coordinates as shown (3), where X_{opt} and Y_{opt} depicts the optimized and improved coordinates.

$$X_{opt} = \beta X_{east} + c + \mu_t$$
$$\left(1 - \alpha_1 L - \cdots - \alpha_p L^p\right)(1 - L)^D \mu_t = (1 + \emptyset_1 L + .. + \emptyset_q L^q)\varepsilon_t \qquad (2)$$

$$Y_{opt} = \beta Y_{north} + c + \mu_t$$
$$\left(1 - \alpha_1 L - \cdots - \alpha_p L^p\right)(1 - L)^D \mu_t = (1 + \emptyset_1 L + .. + \emptyset_q L^q)\varepsilon_t \qquad (3)$$

The best fitted models were applied on the data sets to improve positioning errors in the respective axes. The residuals of the fitted model were analyzed by mean of ACF, PACF, histogram, scatter plot and quartile-quantile plot (QQ-Plot) to ascertain whether the model fit the data well. Simulation of each axis points from the fitted models resulted in improved geodetic points along the trajectory travelled. These estimated points are evaluated in terms of the MAE and RMSE errors of positioning by comparison with

positioning errors of the original trajectory. The model fitting and evaluation outlined in Figure 3 were implemented in MATLAB 2017b with the related toolboxes.
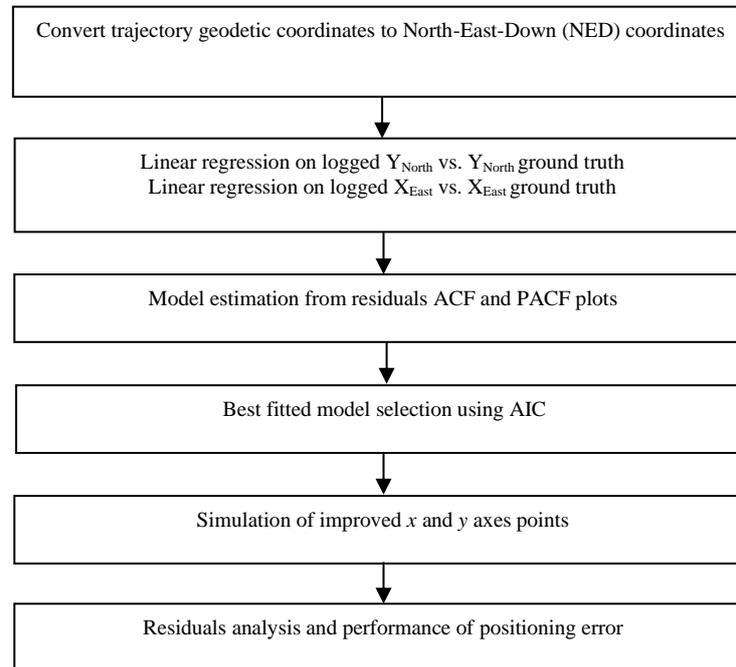

Figure 3. Model fitting and evaluation

3. RESULTS AND DISCUSSION

3.1. Results of initial analysis of logged trajectory

Table 1 shows the results of positioning errors and positioning availability on three sets of data collected on 8 August 2020 at the location depicted in Figure 2. The positioning errors were calculated in terms of MAE and root mean square error (RMSE). Positioning availability column in Table 1 presents the percentage ratio of data points collected with fixed solution (between 1 to 5 cm), float solution (above 5 to 50 cm) and single point solution (above 50 cm). Data set 1 recorded the lowest MAE at 0.2981 m whereas data set 3 has the highest MAE at 0.3905 m. It is obvious that the rover was not able to achieve fixed integer positioning all of the time. The positioning availability in Table 1 reveals that fixed integer positioning ranges from 11.48% to 22.03% out of the overall geodetic points logged along the trajectory. The rest of the geodetic points were mainly floating-point solutions (up to 86%) and small percentages of single point solutions. The results indicated multipath effects along the trajectory travelled by the rover that degrades the positioning accuracy.

Table 1. Positioning errors and availability

Data Set	Positioning Errors		Positioning Availability		
	MAE (m)	RMSE (m)	Float (%)	Fixed (%)	Single (%)
1	0.2981	0.3022	77.65	22.03	0.32
2	0.3359	0.3382	74.87	23.35	1.78
3	0.3905	0.3931	86.34	11.48	2.18

Analysis on the residuals between logged and actual ground truth were conducted to investigate ACF, PACF and the normality of the residuals. For brevity, only analyses on residuals of data set 1 is shown in this paper. Figure 4 shows the ACF and PACF plots of the residuals of data set 1 in Figures 4(a) and 4(b) respectively, which were correlated over time as spikes in the plots exceeded the upper and lower bound line. This implies the presence of correlated noise in the measurements. In addition, the histogram plot in Figure 4(c) did not exhibit normality in the data and was slightly right skewed with non-zero mean. It is obvious that the measurements were affected by noises contributed by multipath effects.
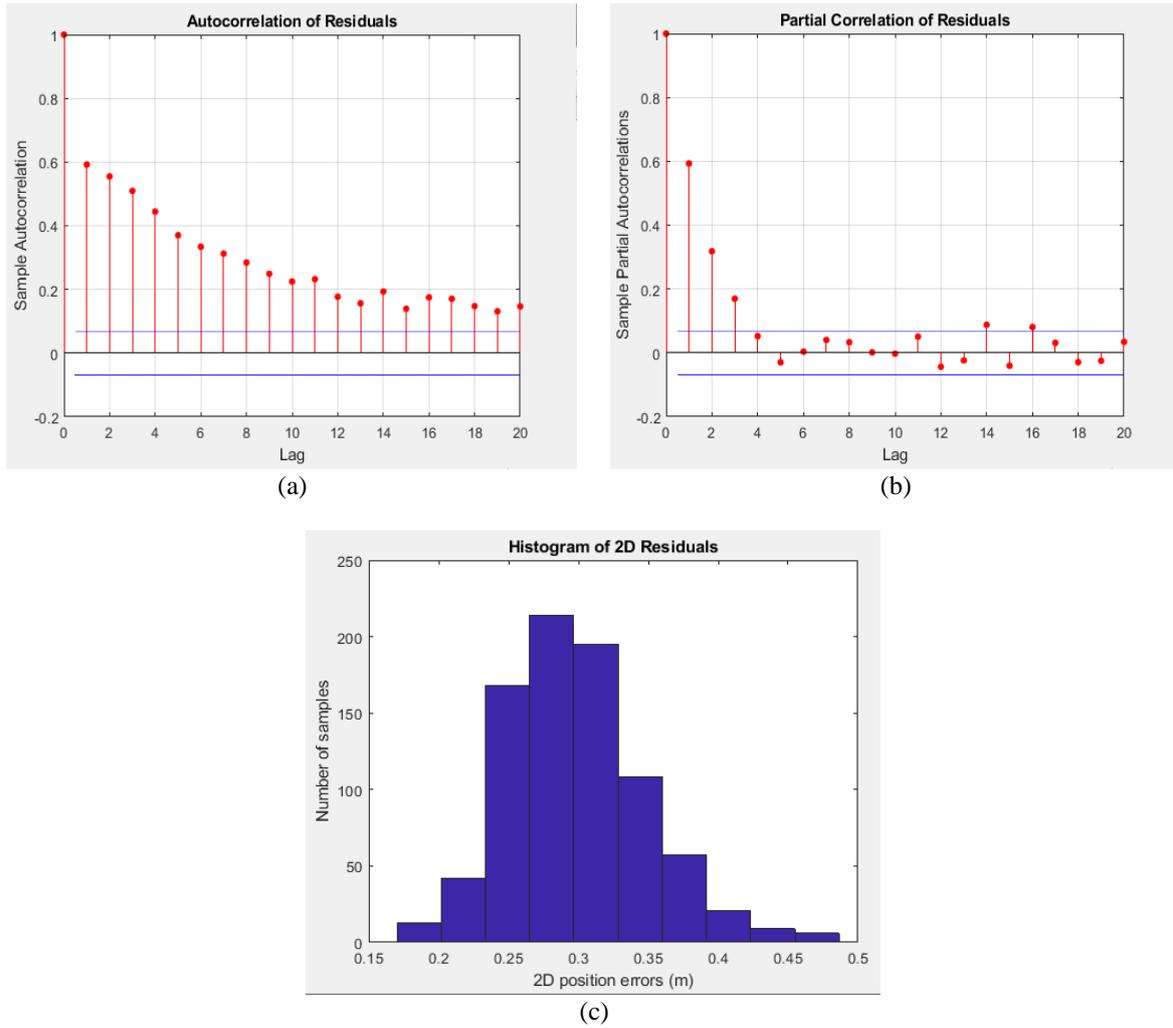
(a)

(b)

(c)

Figure 4. Residuals analysis of data set 1 in (a) autocorrelation plot, (b) partial correlation plot,
and (c) histogram

3.2. Results of fitted linear regression models with ARIMA errors

Linear regressions on all three data sets were implemented in MATLAB. The linear regression models $X_{opt}=\beta X_{east}+c$ and $Y_{opt}=\beta Y_{north}+c$ was derived for both x and y axes respectively and tabulated in Table 2. The value in all coefficients of determination R^2 for all regressions indicated high degree of correspondence between predictor and response data. Though the R^2 is high, it could not determine whether the predictions are biased, or the regression model provides an adequate fit to the data. Hence, the residual plots need to be assessed.

Table 2. Linear regression models

Data set	X_{East} Regression model	R^2	Y_{North} Regression model	R^2
1	$X_{opt}=0.995X_{east} - 0.053$	0.9942	$Y_{opt}=0.999Y_{north}+0.049$	0.9927
2	$X_{opt}=1.005X_{east}+0.095$	0.9931	$Y_{opt}=1.001Y_{north} - 0.061$	0.9918
3	$X_{opt}=1.003X_{east}+0.031$	0.9936	$Y_{opt}=1.001Y_{north} - 0.083$	0.9926

The residuals from these linear regression models were further analyzed using the ACF and PACF. As the residuals were non-stationary, first order differencing was applied, followed by the ACF and PACF analyses. For brevity, the results of ACF and PACF of first order differenced residuals from x-axis are shown in Figures 5(a) and 5(b) respectively, whereas Figures 5(c) and 5(d) show the ACF and PACF respectively for the y-axis residuals. Based on the ACF and PACF relations by [34], the x-axis positioning may possess

integrated moving average (IMA) errors or integrated autoregressive errors (ARI) as both ACF and PACF plots displayed a cut off at a particular lag. The *y*-axis also displayed a similar characteristic as depicted by ACF and PACF plots in Figures 5(c) and 5(d) respectively. Hence, the best integrated MA or AR order needs to be estimated. Similar characteristics were also observed in data set 2 and 3. An algorithm written in MATLAB was implemented to determine the best IMA order and ARI order respectively for each data set with their respective AIC values. The results are shown in Table 3.
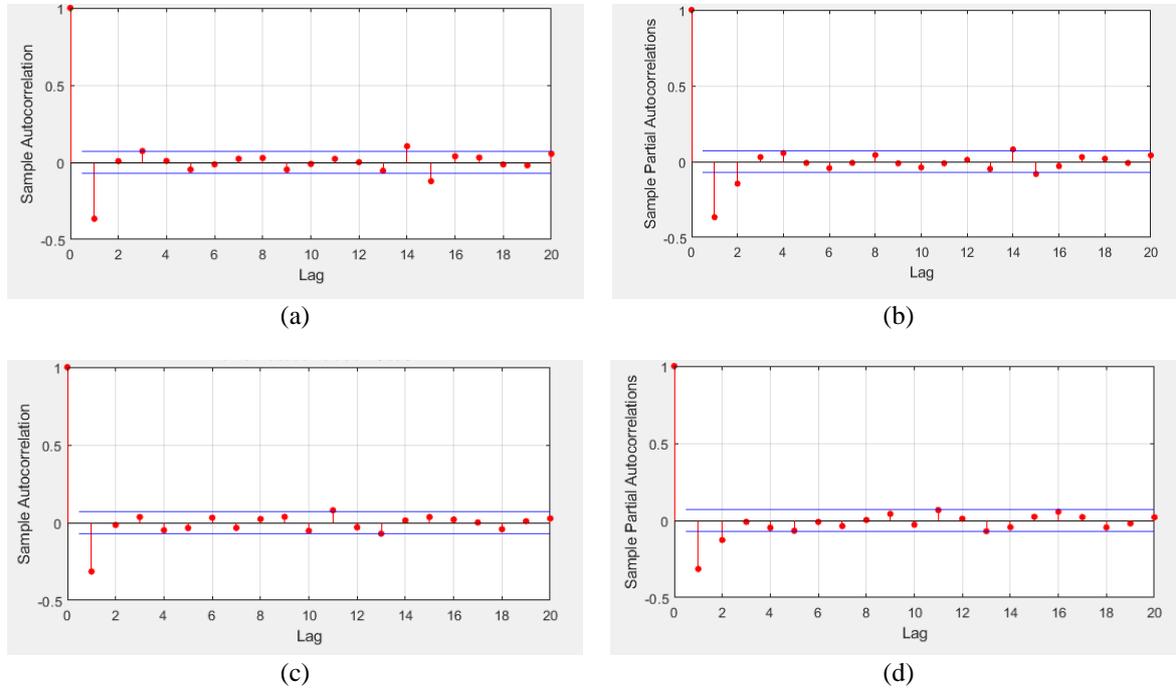
(a)

(b)

(c)

(d)

Figure 5. Residuals analysis of data set 1 in (a) ACF of *x*-axis residuals, (b) PACF of *x*-axis residuals, (c) ACF of *y*-axis residuals, and (d) PACF of *y*-axis residuals

Table 3. Akaike information criteria of ARI and IMA models

Data set	X_{East} error models				Y_{North} error models			
	ARI model	AIC	IMA model	AIC	ARI model	AIC	IMA model	AIC
1	**2,1,0**	**-3132.47**	0,1,2	-3126.70	2,1,0	-2783.29	**0,1,1**	**-2785.87**
2	5,1,0	-2922.53	**0,1,2**	**-2923.75**	3,10	-2594.33	**0,1,1**	**-2601.14**
3	5,1,0	-1952.24	**0,1,2**	**-1957.26**	6,1,0	-1039.83	**0,1,2**	**-1046.81**

Table 3 shows the best ARI and IMA order for the respective axis with their respective AIC values estimated in MATLAB. Further comparison between ARI and IMA models reveal that an ARI (2,1,0) was the appropriate model for *x*-axis errors in data set 1 as its AIC value (i.e., -3132.47) is smaller compared to the IMA model. On the other hand, an IMA (0,1,1) was identified for *y*-axis errors in data set 1. Error models with the smallest AIC were identified for the other data sets (shown in bold font in Table 3). Hence, data set 2 was best modelled with IMA (0,1,2) and (0,1,1) error models for *x*-axis and *y*-axis respectively, whereas an IMA (0,1,2) were identified for both axes in data set 3. The best fitted linear regression models with their estimated regression coefficient, intercept and AR or MA coefficients for each data set are shown in Table 4.

Table 4. Fitted linear regression models with ARIMA errors for *x* and *y* axes

Data set	X_{East}	Y_{North}
1	$X_{opt}=0.995X_{east}-0.053+\mu_t$	$Y_{opt}=0.999Y_{north}+0.049+\mu_t$
	$(1+0.432L-0.154L^2)(1-L)\mu_t = \varepsilon_t$	$(1-L)\mu_t = (1-0.365L)\varepsilon_t$
2	$X_{opt}=1.005X_{east}+0.095+\mu_t$	$Y_{opt}=1.001Y_{north} - 0.061+\mu_t$
	$(1-L)\mu_t = (1-0.410L-0.070L^2)\varepsilon_t$	$(1-L)\mu_t = (1-0.584L)\varepsilon_t$
3	$X_{opt}=1.003X_{east}+0.031+\mu_t$	$Y_{opt}=1.001Y_{north} - 0.083+\mu_t$
	$(1-L)\mu_t = (1-0.458L-0.256L^2)\varepsilon_t$	$(1-L)\mu_t = (1-0.396L-0.338L^2)\varepsilon_t$

3.3. Results of evaluation of model fit and performance of positioning errors

The fitted models in Table 4 were implemented accordingly for each data set. The residuals of these fitted models from the respective axis's regressions with ARIMA errors were evaluated using ACF, PACF, histogram, scatter plot and QQ-Plot to ascertain whether the model fit the data well. Figure 6 shows the ACF, PACF and the scatter plot of the residuals from both axes for data set 1. The ACF and PACF did not indicate any correlations in the residuals. In addition, the scatter plot i.e., the residuals versus fitted values displayed symmetrically distributed points and tendency to cluster towards the zero mean. Figure 7 shows normally distributed histograms and QQ-Plots of the residuals. The histograms clearly displayed the residuals have zero mean for both axes. In short, these results indicated that the model fitted the data well.
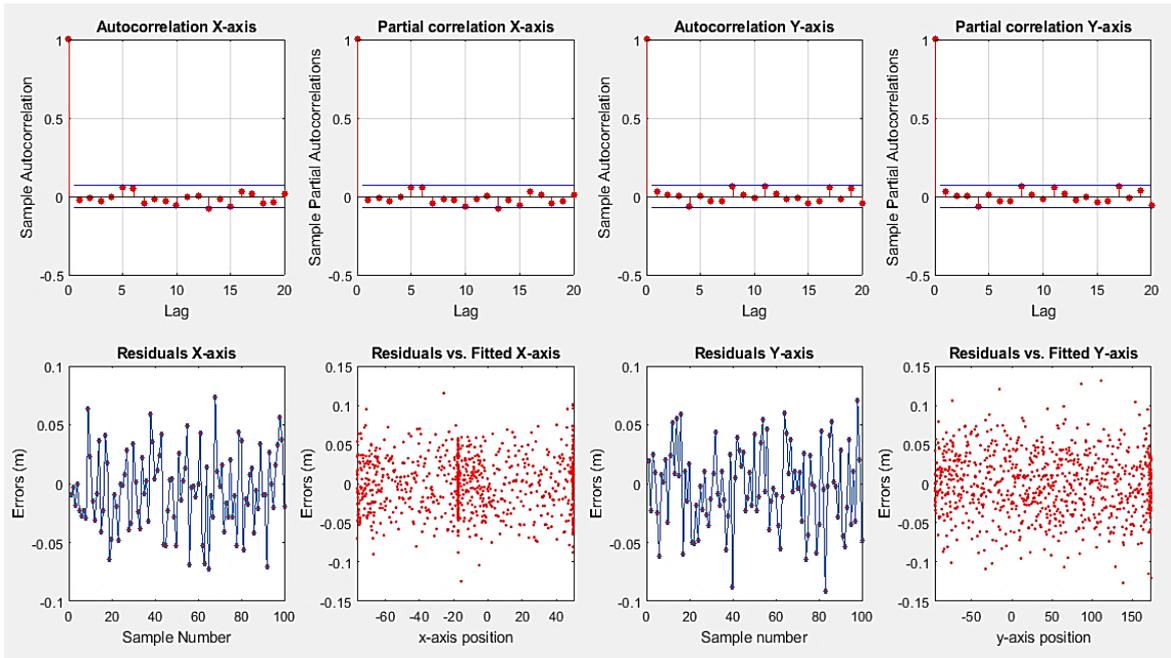

Figure 6. ACF, PACF and scatter plot of residuals from fitted models of data set 1
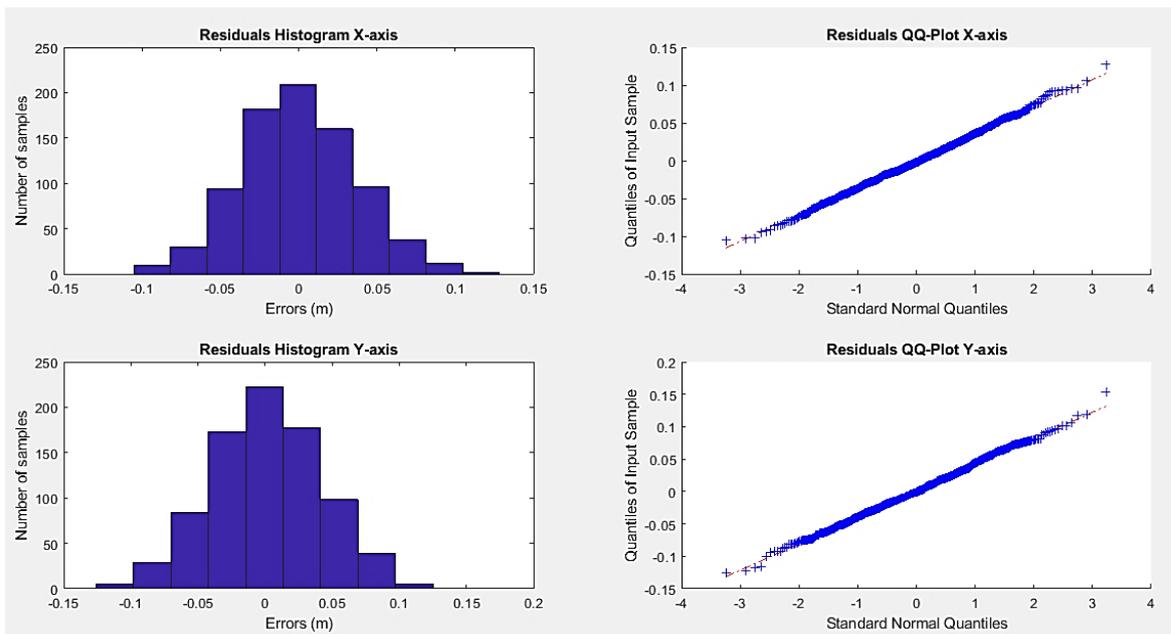

Figure 7. Histograms and QQ-Plots of residuals from fitted models of data set 1

Table 5 shows the MAEs and RMSEs of 2-dimension (*x-y* coordinates) positioning errors produced by the fitted models for each data set. These MAEs and RMSEs were compared with original MAEs and RMSEs of each data set to evaluate the improvement in positioning errors brought about by the models. It was observed that the fitted models improved MAE by 38.73%, 57.35% and 45.56% for data set 1, 2 and 3 respectively.

Table 5. Evaluation of positioning errors

Data Set	Fitted Models		Original Data		% Improvement in MAE
	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	
1	0.2058	0.2154	0.3359	0.3382	38.73
2	0.1433	0.1634	0.3360	0.3391	57.35
3	0.1453	0.1632	0.2669	0.2861	45.56

4. CONCLUSION

This work successfully described measurement errors of RTK-GNSS along a trajectory point logged under multipath effects using linear regression model with ARIMA errors. The work has shown that regression with ARIMA errors can be extended to dynamic applications, compared to its conventional applications in static stations. As demonstrated via simulations, the models act as a filter for measurement noise along the trajectory resulted in improved positioning accuracy ranging from 38.73% to 57.35% compared to original measurements. As the regression model relied on historical data of logged GNSS points, this may not directly applicable for real-time implementation. Future studies may need to integrate other sensors such as inertial and odometry sensors to be coupled with RTK-GNSS measurements to estimate the models during real-time dynamic tests.

REFERENCES

[1] E. Saad, M. Elhosseini, and A. Y. Haikal, "Recent achievements in sensor localization algorithms," *Alexandria Engineering Journal*, vol. 57, no. 4, pp. 4219–4228, Dec. 2018, doi: 10.1016/j.aej.2018.11.008.
[2] A. Paul and T. Sato, "Localization in wireless sensor networks: a survey on algorithms, measurement techniques, applications and challenges," *Journal of Sensor and Actuator Networks*, vol. 6, no. 4, Oct. 2017, doi: 10.3390/jsan6040024.
[3] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, Apr. 2018, doi: 10.1109/JIOT.2018.2812300.
[4] Novatel, "OEMV-1DF^TM L1/L2 receiver provides RTK positioning for space-constrained applications." White Paper, NovAtel's OEMV-1DF Compact L1/L2 RTK Receiver, Accessed Dec. 14, 2021. [Online] Available: https://hexagondownloads.blob.core.windows.net/public/Novatel/assets/Documents/Papers/OEMV-1DF_White_Paper/OEMV-1DF_White_Paper.pdf (accessed Dec. 14, 2021)
[5] D. Gebre-Egziabher, J. Jackson, R. Saborio, S. A. Ghazanfar, and B. Davis, "Evaluation of low-cost, centimeter-level accuracy OEM GNSS receivers," Minnesota Department of Transportation, 2018.
[6] Y. Luo, Y. Yao, and L. Shan, "Analysis of ionospheric disturbances caused by the 2018 Bering Sea meteor explosion based on GPS observations," *Sensors*, vol. 20, no. 11, Jun. 2020, doi: 10.3390/s20113201.
[7] B. Bidikar, B. Prasad Chapa, M. Vinod Kumar, and G. Sasibhushana Rao, "GPS signal multipath error mitigation technique," in *Satellites Missions and Technologies for Geosciences*, IntechOpen, 2020, doi: 10.5772/intechopen.92295.
[8] M. A.-E.-A. Youssef and D. A. Hussein, "Impact of multipath on the distance between the buildings and GPS receiver," *Journal of Civil Engineering*, vol. 4, no. 2, 2018
[9] T. Feng, S. Chen, Z. Feng, C. Shen, and Y. Tian, "Effects canopy and multi-epoch observations on single-point positioning errors of a GNSS in coniferous and broadleaved forests," *Remote Sensing*, vol. 13, no. 12, Jun. 2021, doi: 10.3390/rs13122325.
[10] P. Benevides, J. Catalao, and G. Nico, "Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors," *Remote Sensing*, vol. 11, no. 8, Apr. 2019, doi: 10.3390/rs11080966.
[11] M. Sangiorgio *et al.*, "Improved extreme rainfall events forecasting using neural networks and water vapor measures," in *International Conference on Time Series and Forecasting*, 2019, pp. 820–826.
[12] G. Guerova, T. Dimitrova, and S. Georgiev, "Thunderstorm classification functions based on instability indices and GNSS IWV for the Sofia plain," *Remote Sensing*, vol. 11, no. 24, Dec. 2019, doi: 10.3390/rs11242988.
[13] M. Łoś, K. Smolak, G. Guerova, and W. Rohm, "GNSS-based machine learning storm nowcasting," *Remote Sensing*, vol. 12, no. 16, Aug. 2020, doi: 10.3390/RS12162536.
[14] M. Sangiorgio *et al.*, "A comparative study on machine learning techniques for intense convective rainfall events forecasting," in *Theory and Applications of Time Series Analysis*, Springer International Publishing, 2020, pp. 305–317, doi: 10.1007/978-3-030-56219-9_20.

[15]   P. Barba, B. Rosado, J. Ramírez-Zelaya, and M. Berrocoso, "Comparative analysis of statistical and analytical techniques for the study of GNSS geodetic time series," *The 7th International conference on Time Series and Forecasting*, Jun. 2021, doi: 10.3390/engproc2021005021.

[16]   R. Hohensinn, S. Häberling, and A. Geiger, "Dynamic displacements from high-rate GNSS: error modeling and vibration detection," *Measurement: Journal of the International Measurement Confederation*, vol. 157, Jun. 2020, doi: 10.1016/j.measurement.2020.107655.

[17]   D. Spoljar, N. Crnjaric-Zic, K. Lenac, and V. Perinovic, "Characterisation of multipath-caused commercial-grade GPS positioning error in intelligent transport systems (ITS)," in *International Symposium ELMAR*, Sep. 2019, pp. 27–30, doi: 10.1109/ELMAR.2019.8918652.

[18]   Z. Tao and P. Bonnifait, "Modeling L1-Gps errors for an enhanced data fusion with lane marking maps for road automated vehicles," *European Navigation Conference*, 2015.

[19]   G. Zhang, Q. Gui, S. Han, J. Zhao, and W. Huang, "A Bayesian method of GNSS cycle slips detection based on ARMA model," in *Forum on Cooperative Positioning and Service*, May 2017, pp. 219–222, doi: 10.1109/CPGPS.2017.8075128.

[20]   M. R. Kaloop, M. Hussan, and D. Kim, "Time-series analysis of GPS measurements for long-span bridge movements using wavelet and model prediction techniques," *Advances in Space Research*, vol. 63, no. 11, pp. 3505–3521, Jun. 2019, doi: 10.1016/j.asr.2019.02.027.

[21]   A. S. Arifin and M. I. Habibie, "The prediction of mobile data traffic based on the ARIMA model and disruptive formula in industry 4.0: a case study in Jakarta, Indonesia," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 907–918, Apr. 2020, doi: 10.12928/TELKOMNIKA.v18i2.12989.

[22]   M. A. Shaari, R. Samsudin, and A. S. Ilman, "Forecasting drought using modified empirical wavelet transform-ARIMA with fuzzy C-means clustering," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 11, no. 3, pp. 1152–1161, Sep. 2018, doi: 10.11591/ijeecs.v11.i3.pp1152-1161.

[23]   Y. S. Triana and A. Retnowardhani, "Enhance interval width of crime forecasting with ARIMA model-fuzzy alpha cut," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, no. 3, Jun. 2019, doi: 10.12928/telkomnika.v17i3.12233.

[24]   S. Salma, G. Sivavaraprasad, B. T. P. Madhav, and D. V. Ratnam, "Implementation of VARMA model for Ionospheric TEC forecast over an Indian GNSS station," in *ICDCS 2020 - 2020 5th International Conference on Devices, Circuits and Systems*, Mar. 2020, pp. 144–148, doi: 10.1109/ICDCS48716.2020.243568.

[25]   J. Xin, J. Zhou, S. Yang, X. Li, and Y. Wang, "Bridge structure deformation prediction based on GNSS data using Kalman-ARIMA-GARCH model," *Sensors*, vol. 18, no. 1, Jan. 2018, doi: 10.3390/s18010298.

[26]   A. W. R. Soundy, B. J. Panckhurst, P. Brown, A. Martin, T. C. A. Molteno, and D. Schumayer, "Comparison of enhanced noise model performance based on analysis of civilian GPS data," *Sensors*, vol. 20, no. 21, Oct. 2020, doi: 10.3390/s20216050.

[27]   H. Ge, B. Li, T. Wu, and S. Jiang, "Prediction models of GNSS satellite clock errors: evaluation and application in PPP," *Advances in Space Research*, vol. 68, no. 6, pp. 2470–2487, Sep. 2021, doi: 10.1016/j.asr.2021.05.025.

[28]   I. L. Mallika, D. V. Ratnam, S. Raman, and G. Sivavaraprasad, "A new ionospheric model for single frequency GNSS user applications using klobuchar model driven by auto regressive moving average (SAKARMA) method over Indian region,thod over Indian region," *IEEE Access*, vol. 8, pp. 54535–54553, 2020, doi: 10.1109/ACCESS.2020.2981365.

[29]   C. E. Magrin and E. Todt, "Multi-sensor fusion method based on artificial neural network for mobile robot self-localization," in *Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, Oct. 2019, pp. 138–143, doi: 10.1109/LARS-SBR-WRE48964.2019.00032.

[30]   X. Ruan, S. Liu, D. Ren, and X. Zhu, "Accurate 2D localization for mobile robot by multi-sensor fusion," in *IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Dec. 2018, pp. 839–843, doi: 10.1109/ITOEC.2018.8740490.

[31]   S. Erfani, A. Jafari, and A. Hajiahmad, "Comparison of two data fusion methods for localization of wheeled mobile robot in farm conditions," *Artificial Intelligence in Agriculture*, vol. 1, pp. 48–55, Mar. 2019, doi: 10.1016/j.aiia.2019.05.002.

[32]   M. Doumbia and X. Cheng, "State estimation and localization based on sensor fusion for autonomous robots in indoor environment," *Computers*, vol. 9, no. 4, Oct. 2020, doi: 10.3390/computers9040084.

[33]   U. Hani and L. Moin, "Realtime autonomous navigation in V-Rep based static and dynamic environment using EKF-SLAM," *IAES International Journal of Robotics and Automation (IJRA)*, vol. 10, no. 4, pp. 296–307, Dec. 2021, doi: 10.11591/ijra.v10i4.pp296-307.

[34]   Z. Liu, Z. Zhu, J. Gao, and C. Xu, "Forecast methods for time series data: a survey," *IEEE Access*, vol. 9, pp. 91896–91912, 2021, doi: 10.1109/ACCESS.2021.3091162.

BIOGRAPHIES OF AUTHORS

Kok Mun Ng received the B.Eng. and M.Eng. degrees in systems engineering from the Universiti Tun Hussein Onn, Malaysia, in 2002 and 2006 respectively. He obtained the Ph.D. degree in Systems Engineering from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 2016. From 2002 to 2003, he worked as a system engineer at Matromatics Handling Systems, Malaysia. He was with Stamford College, Malaysia from 2006 to 2008 as a lecturer before joining the School of Electrical Engineering, Universiti Teknologi MARA, where he works as a senior lecturer at the Systems Engineering department. He is the author of more than 30 articles and conference papers. His research interests include system modeling and optimization, artificial intelligence, intelligent transportation systems, mapping and localization. He is a sub-editor of the Journal of Electrical and Electronic Systems Research (JEESR) and a research fellow at the university vehicle telematics industrial research lab. He can be contacted at email: ngkokmun@uitm.edu.my.

Ravenny Sandin Nahar 🆔 ⑧ SC ◐ graduated with Bachelor of Engineering (Hons.) Electronic Engineering at the Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) in 2019. He is currently pursuing his post graduate studies in master's in engineering (Electrical Engineering) at the School of Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam. His research interest includes artificial intelligence and smart systems. He can be contacted at email: ravenhiccup297@gmail.com.

Mamun IbneReaz 🆔 ⑧ SC ◐ was born in Bangladesh in December 1963. He received the B.Sc. and M.Sc. degrees in applied physics and electronics, both from the University of Rajhashi, Rajshahi, Bangladesh, in 1985 and 1986, respectively. He received the D.Eng. degree in 2007 from Ibaraki University, Hitachi, Japan. He is currently a Professor with the Universiti Kebangsaan Malaysia, Bangi, Malaysia, where he is involved in teaching, research, and industrial consultation. Since 2008, he has been a Regular Associate with the Abdus Salam International Center for Theoretical Physics, Trieste, Italy. He has vast research experience in Norway, Ireland, and Malaysia. He has extensively published in the area of integrated circuit (IC) design and biomedical application ICs. He is the author or coauthor of more than 100 research articles in design automation and IC design. He can be contacted at email: mamun.reaz@gmail.com.