DOI: https://doi.org/10.53486/dri2024.39
UDC: 519.23:656.7

CHARACTERISTICS OF THE APPLICATION OF ARIMA-SVM METHODS IN THE FORECASTING OF NON-SCHEDULED PASSENGER AIR TRANSPORTATION

Nazarli DASHQIN [50] PhD student
ORCID number (https://orcid.org/): 0009-0001-8063-3670

Abstract: *The article examines the independent application of ARIMA (Auto Regressive Integrated Moving Average) and SVM (Support Vector Machine) methods for forecasting non-scheduled passenger air transportation. Based on the SVM model, the data is classified based on different kernel functions, and the best prediction results are determined. The autoregression (ARIMA) model is applied to identify linear trends and regularities within time series data. Based on the analysis, the results show that ARIMA and SVM models offer superior forecasting accuracy and reliability. Nevertheless, the relative error of the prediction results compared to the actual indicators is smaller in the SVM model. This also shows that the identification of non-linear relationships between data in non-scheduled passenger air transportation makes forecasting results more effective and optimal. The obtained results will serve as an effective tool in forecasting the demand for non-scheduled passenger air transportation. As a result, substantial support will be observed in the planning of operations in the mentioned field, preparation of the existing infrastructure according to the demand, etc.*

Key words: *forecasting, statistical methods, autoregressive method, time-series analysis, statistical analysis, air passenger demand.*

JEL: *R41*

1. Introduction.

Unlike scheduled air transportation, the demand for non-scheduled air transportation is usually formed on the basis of orders and outside the schedule. It is clear from this that a large number of internal and external factors affect the mentioned air transportation. This leads to the creation of ambiguous relationships between the data. For this reason, when forecasting models are examined, factors that may have an active influence on the model should be taken into account. Each of the ARIMA and SVM models was applied in both independent and hybrid form in forecasting the demand for air transportation. Both models have negative and positive aspects depending on the field of application and the number of influencing factors. The ARIMA model is well regarded for its ability to model and predict linear patterns among input data. This makes it a valuable tool for understanding major trends and anomalies. The SVM model is a machine learning technique capable of identifying complex, non-linear relationships between data [1]. This paper examines the individual characteristics and applications of ARIMA and SVM methods in the context of non-scheduled passenger air transportation forecasting. By applying the mentioned methods independently, a thorough evaluation of their respective strengths and weaknesses, irregularities, and suitability for this forecasting problem was provided. With the application of ARIMA and SVM models in regular air passenger and freight transportation, it was investigated how appropriate their predictions are for the mentioned area. [2-3]

[50] dnazarli.32073@naa.edu.az, National Aviation Academy of Azerbaijan Airlines

2. Related works.

The issue of air transport demand forecasting has been investigated by various researchers. Forecasting indicators are given in the short and long term, depending on the field of application. When forecasting models are built, one or more variables are considered. In the context of statistical forecasting of regular passenger and cargo air transportation, the main issue to be considered is the verification of stationarity between the data. In order to build an autoregression model, the autocorrelation relationship between the data must not change over time. The models applied based on this method can be ARIMA, which takes into account trends and other factors, and SARIMA models, which take into account seasonal factors. [1-3]

Forecasting of non-scheduled passenger air traffic by the SVM method was investigated by researchers. In this study, calculations were made based on different kernel functions, and short-term forecast results were obtained. [4]

Another study considered the forecasting of air cargo transportation based on the SVM method. The statistics covered the volume of cargo transported from Beijing to Shanghai. The obtained prediction results were compared with the Brown cubic exponential smoothing method. It was found that the results based on the SVM method have more effective prediction accuracy. [5]

The SVM method has also been applied to predict the demand for passenger air transportation. [6-7]

In another study, combined Bootstrap aggregation (Bagging) and Holt Winters methods were used to more effectively forecast air transportation demand. During the application of the methods, trend, seasonal, and other components were taken into account as basis. The Holt Winters method was applied for time series modelling, and the final forecast results were obtained. Errors were identified by comparing the forecast results with the actual data of different countries. It was determined that the forecast results obtained by combined Bagging Holt Winters methods are more optimal and efficient. [8-10]

3. Problem statement.

Construction of forecasting models of non-scheduled passenger air transportation based on ARIMA-SVM methods and comparative analysis of the obtained forecasting results.

4. Method and methodology.

The differential autoregressive moving average model (ARIMA) is an important method for studying time series. In ARIMA (p, d, q), p is the number of autoregressive items, q is the moving average item number, and d is the number of differences made to make it a stationary sequence. The ARIMA (p, d, q) model is an extension of the ARMA (p, q) model. [5]

The ARIMA model in the following form:

$$Y_t = y_t + z_t \tag{1}$$

where, y_t is trend and z_t is errors.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} , \tag{2}$$

$$z_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} , \tag{3}$$

We apply the method of least squares to find the unknown coefficients. For this, the following issue should be resolved:

$$\sum_{t=1}^{N}[(\overline{Y}_t - Y_t)]^2 \dot{\rightarrow} min \quad (4)$$

The solution to problem (4) is reduced to the following matrix equation:

$$A\varphi = B \quad (5)$$

The SVM kernel is considered a function that takes low-dimensional input space and transforms it into higher-dimensional space, usually it converts non-separable problems to separable problems. It is mostly useful in non-linear separation problems. Consider the following formulas:

- Linear: $K(w, b) = w^T x + b$ (6)
- Polynomial: $K(w, x) = (\gamma w^T x + b)^N$ (7)
- Gaussian RBF: $K(w, x) = exp(-\gamma \|x_i - x_j\|^n)$ (8)
- Sigmoid: $K\left(x_i, x_j = tanh(\alpha x_i^T x_j + b)\right)$ (9)

5. Experimental results.

Statistical data covering the years 2020-2023 of non-scheduled passenger air transportation were collected to build the calculation model. Statistical indicators are given in Fig 1. The forecast results of each of the ARIMA and SVM models for 2023 will be calculated based on the indicators of 2020-2022. In order to check the accuracy of the forecast results, the actual indicators of 2023 were not included in the calculation models. This information will be used to determine the error of the forecast results based on the actual indicators.

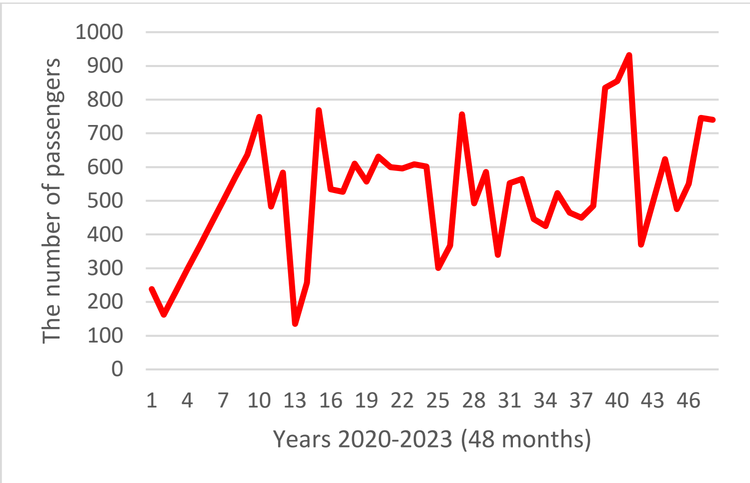

Fig 1. Monthly statistics of non-scheduled passenger air transportation for 2020–2023

Source: From author own investigation

Fig 2 shows the forecast results of ARIMA and SVM models for 2023 based on the statistical indicators of 2020-2022. As observed from Fig 2, the results of the SVM model provide linear solutions. As we mentioned earlier, this model is designed for classification. Our analyses show that this model creates limitations in expressing the process of non-scheduled air transportation in all cases. The results of the ARIMA model indicate an unscheduled passenger air transportation process, but there are anomalous deviations at several points. The main reason for this is that non-scheduled passenger air transport is subject to random fluctuations. Various smoothing methods are usually applied to avoid such errors. Data are smoothed, and calculations are performed before entering the model. In our calculation, these cases did not have a serious effect on the final results.

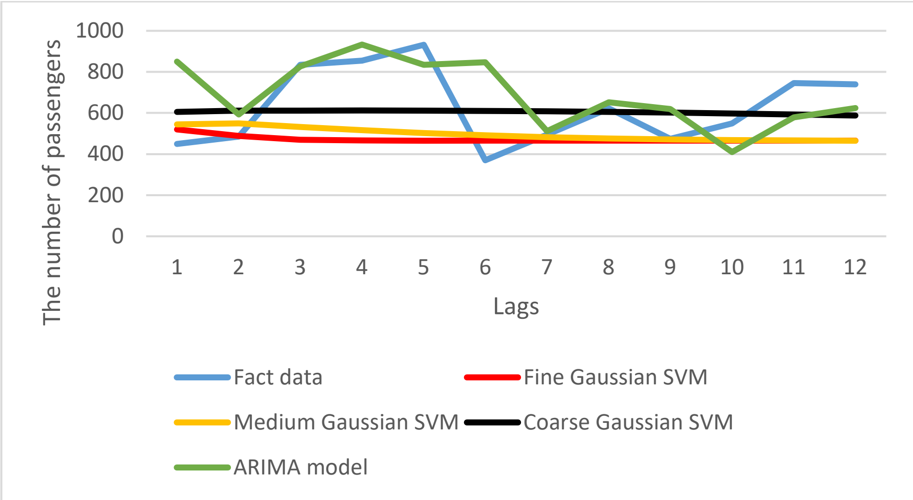

Fig 2. Forecasting results of non-scheduled passenger air transportation based on ARIMA-SVM models

Source: From author own investigation

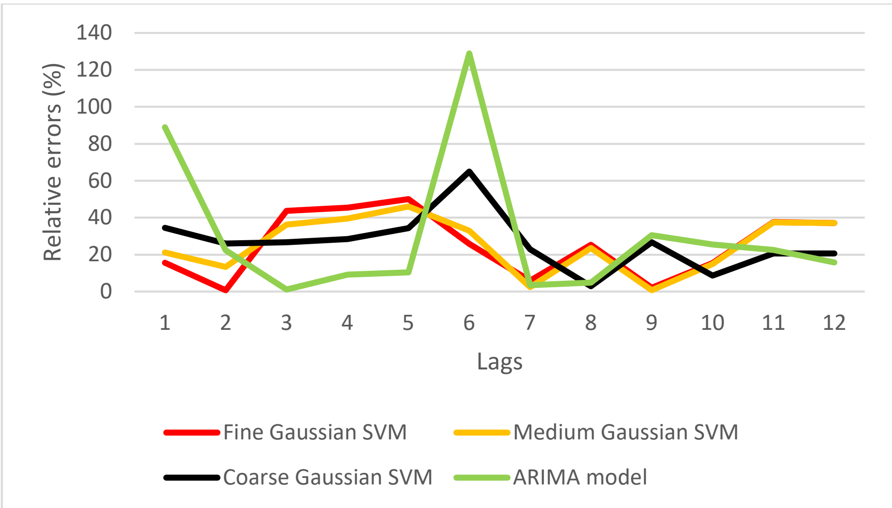

Fig 3. Relative error of ARIMA-SVM models forecasting results based on actual indicators

Source: From author own investigation

Fig 3 shows the variation of the relative errors of the forecasting results obtained based on the ARIMA-SVM models with respect to the actual indicators. As can be seen from Fig 3, the relative errors of the models are, respectively, ARIMA (19.8%), Fine Gaussian SVM (25.2%), Medium Gaussian SVM (24.7%), and Coarse Gaussian SVM (28.1%).

6. Conclusions.

To conclude, it should be noted that ARIMA and SVM (fine, medium, and coarse) methods have been applied to predict non-scheduled passenger air transportation. The results show that the ARIMA model represents the process better. The reason is that this model captures the process based on trend and autocorrelation relationships. In the SVM method, the relative error is greater. The forecast results obtained based on the ARIMA model are more optimal compared to the actual indicators, and the relative error is smaller than other models. When comparing the results of the ARIMA model, we observe that the model expresses the general trend of the actual indicators. It is recommended to apply smoothing methods to avoid sudden deviations in the results obtained during forecasting. The application of these methods varies depending on the characteristic features of the process. We can use the results of our independently applied models as a basis for obtaining more accurate and effective forecast results in the forecasting of non-regular passenger air transportation. This basic part is considered one of the important elements in building machine learning models.

Bibliographical references.

1. Kong, J.G. and Xiao L. (2014) Application of Wavelet Analysis Based on MATLAB in the Forecasting of Passenger Transport Capacity of Civil Aviation. China Science and Technology Information, 16, 40-41.
2. F. Al-Rukaibi and N. Al-Mutairi, Forecasting air travel demand of Kuwait: a comparison study by using regression vs. artificial intelligence, Journal of Engineering Research 1(1) (2013), 113-143.
3. Tang, X.X. and Deng, G.M. (2016) Prediction of Civil Aviation Passenger Transportation Based on ARIMA Model. Open Journal of Statistics, 6, 824-834. http://dx.doi.org/10.4236/ojs.2016.65068
4. Aghayev, N., Nazarli, D. (2024). Support vector machines for forecasting non-scheduled passenger air transportation. Problems of Information Technology, 15(1), 3–9. https://doi.org/10.25045/jpit.v15.i1.01
5. H. J. Heng, B. Z. Zheng, and Y. J. Li, "Study of SVM-based air-cargo demand forecast model," in *Computational Intelligence and Security, 2009. CIS'09. International Conference on*, 2009, vol. 2, pp. 53–55.
6. Xie G, Wang S, Lai KK. Short-term Forecasting of Air Passenger by Using Hybrid Seasonal Decomposition and Least Squares Support Vector Regression Approaches. Journal of Air Transport Management 2014;37(2014):20-26.
7. A. Shabri, "Least Square Support Vector Machines as an Alternative Method in Seasonal Time Series Forecasting," *Applied Mathematical Sciences. Sci.*, vol. 9, no. 124, pp. 6207–6216, 2015.
8. T. M. Dantas, F. L. C. Oliveira and H. M. V. Repolho, Air transportation demand forecast through bagging Holt-Winters methods, Journal of Air Transport Management 59 (2017), 116-123.
9. G. Ren and J. Gao, "Comparison of NARNN and ARIMA models for short-term metro passenger flow forecasting," in *Proceedings of the 19th COTA International Conference of Transportation Professionals g*, Nanjing, China, 2019.
10. Suryani E, Chou SY, Chen CH. Air Passenger Demand Forecasting and Passenger Terminal Capacity Expansion: A System Dynamics Framework. Expert Systems with Applications 2010;37:2324-2339.