DOI: 10.26650/JGEOG2024-1426209
COĞRAFYA DERGİSİ
JOURNAL OF GEOGRAPHY
2024, (49)
https://iupress.istanbul.edu.tr/en/journal/jgeography/home

Analyzing Regional Factors Influencing Passenger, Aircraft Demand, and Freight Demand Using Penalized Geographically Weighted Regression Models

Cezalandırılmış Coğrafi Ağırlıklı Regresyon Modellerini Kullanarak Yolcu, Uçak Talebi ve Yük Miktarını Etkileyen Bölgesel Faktörlerin Analizi

Semra TÜRKAN[1]

[1]Hacettepe University, Department of Statistics, Ankara, Türkiye

ORCID: S.T.0000-0002-4236-2021

ABSTRACT
Airline passenger demand, aircraft demand, and cargo volume are among the most critical factors in the economic decision-making processes of airlines and airport management. Regional factors affecting airline passenger demand, aircraft demand and cargo volume can be divided into two main groups: socio-economic factors and air transport-related factors. This study has two main objectives: First, it emphasizes that when dealing with regional data and encountering local multicollinearity between variables, penalized models are more appropriate; second, to investigate whether the number of passengers, number of aircraft, and cargo volume are related to regional socioeconomic indicators. For this purpose, regional indicators from 48 provinces were obtained from TÜİK (Turkish Statistical Institute), including statistics on air transport and economic indicators. Based on the model performance criteria, Geographically Weighted Lasso Regression was determined as the most suitable model for data analysis. The findings reveal that the most important factor affecting passenger, aircraft, and cargo demand is exports, which is an indicator of regional economic growth.
Keywords: Penalized geographically weighted regression, local multicollinearity, airline statistics

ÖZ
Havayolu yolcu talebi, uçak talebi ve yük hacmi, havayolu şirketlerinin ve havalimanı yönetiminin ekonomik karar alma süreçlerinde en kritik faktörler arasında yer almaktadır. Havayolu yolcu talebini, uçak talebini ve yük hacmini etkileyen bölgesel faktörler iki ana gruba ayrılabilir: sosyo-ekonomik faktörler ve havayolu taşımacılığı faktörleri. Bu çalışmanın iki temel amacı vardır: Birincisi, bölgesel veriler olduğunda ve değişkenler arasında local çoklu bağlantı ile karşılaşıldığında cezalandırılmış modellerin kullanılmasının daha uygun olduğunu vurgulamak; ikinci olarak yolcu sayısı, uçak sayısı ve yük hacminin bölgesel sosyo-ekonomik göstergelerle ilişkili olup olmadığını araştırmak. Bu amaçla TÜİK'ten bölgesel göstergeler kullanılarak 48 şehire ilişkin havayolu istatistikleri ve ekonomik göstergeler elde edilmiştir. Model performans kriterlerine göre Coğrafi Ağırlıklı Lasso Regresyonu, verilerin analizi için en uygun model olarak belirlenmiştir. Bu çalışmanın bulguları, yolcu talebini, uçak talebini ve yük hacmini etkileyen en önemli faktörün bölgesel ekonomik büyümenin bir göstergesi olan ihracaatın olduğunu ortaya koymaktadır.
Anahtar kelimeler: Cezalandırılmış coğrafi ağırlıklı regresyon, lokal çoklu bağlantı, havayolu istatistikleri

Submitted/Başvuru: 26. 01.2024 • **Accepted/Kabul:** 01.10.2024

Corresponding author/Sorumlu yazar: Semra TÜRKAN / sturkan@hacettepe.edu.tr

Citation/Atıf: Türkan, S. (2024). Analyzing regional factors influencing passenger, aircraft demand, and freight demand using penalized geographically weighted regression models. *Cografya Dergisi, 49*, 19-30. https://doi.org/10.26650/JGEOG2024-1426209

1. INTRODUCTION

The common goal of all countries in the world is to prevent poverty by developing economies, industries, and trade. In order to achieve this, one of the most important tools that initiate common commercial relations and bring cultures together is the transportation sector, especially air transportation. By combining different geographies, air transportation facilitates the transportation of people and products and develops the regions in which the airports are located in terms of import and export and adds dynamism to those regions (Airline Industry Report, 2019). Hence, developments in airport service networks contribute to countries' economic growth and regional development. This creates direct and indirect economic effects in many sectors. Increasing employment, leading development plans, providing an important tax resource to the state, attracting foreign investors, and supporting sustainable economic growth are among the economic effects of the sector that can be considered important (Altuntaş and Kılıç, 2021).

In the aviation industry, as in many other industries, it is of great importance to read and analyze the data and, as a result, to predict how such data will be used in the future. Aviation authorities and operators attach great importance to measuring demand in terms of passengers and traffic or preventing capacity inadequacies through forecasts and simulations. Because demand for airline passengers is one of the most important factors in economic decision-making processes in airlines and airport management. Factors affecting airline passenger demand can be handled under two main headings: socioeconomic variables (population, GDP, per capita income, social structure, education, being a center of attraction, political events, government regulations) and airline transport variables (ticket price, flight time, comfort, confidence) (Taneja, 1971). It is important for airports to determine how there will be a change in passenger and freight transportation in line with socioeconomic data over the years (Tanyel et al., 2010). Socioeconomic development causes changes in individuals' transportation habits.

This study aimed to investigate how socio-economic indicators affect the number of passengers, aircraft, and freight in Turkey. For this purpose, data on passengers, aircraft, and freight and selected socioeconomic indicators related to provinces with active airports in Turkey were obtained from the Turkstat website. While selecting socioeconomic indicators, previous studies in this field were also taken into account. Studies on airports generally focus on efficiency and productivity, and the role of location information is also crucial in airport studies. Therefore, the location and spatial effects of airports need to be considered when examining airport benchmarking and development policies. When analyzing data using geographical units (cities, countries, states, regions, etc.), it should be investigated whether neighboring units have spatial effects on each other. It is assumed that geographical units that are close to each other will affect each other. When analyzing data in geographical units, spatial models that take into account spatial effects are used (Lewandowska, 2018). Geographically Weighted Regression (GWR) is a commonly used spatial statistical method that explores spatial diversity by establishing different regression models for each observation location.

The airport data were previously analyzed by different models in the literature. There are several spatial studies on airport data and its determinants. Tanyel et al. (2010) examined the relation between passenger and freight demand and the change in the number of aircraft to Adnan Menderes Airport with various socioeconomic factors, such as the number of cars, number of trucks, length of highway, import, export, electricity consumption, number of hospital beds, number of students, number of accidents, and number of injured in accidents. Baikgaki and Daw (2013) modeled the determinants (income, population, GDP, household consumption, expenditures, ticket prices, crude oil prices, employment) of passenger demand in South Africa by multiple regression analysis. Valdes (2015) examined airline demand in 32 middle-income countries with panel models; used GDP, income, net foreign direct investment, consumer price index, transportation fees, real exchange rate, jet fuel prices, and total number of seats offered by low-cost carriers. Efendigil and Eminler (2017) analyzed the number of passengers arriving at Istanbul Atatürk Airport by considering 20 destinations and 16 airlines by using regression and artificial neural networks. Factors affecting the number of passengers are import, export, national income per capita, number of seats, flight frequency, distance, population, airline type, airline characteristics, and tourism statistics. Chen et al. (2017) used spatial econometric models to analyze the cost function using the latitude and longitude data of Chinese airports between 2002 and 2012. Choo (2018) investigated the effect of immigration on demand for air transport in Canada. The variables included in the analysis in this study are population, GDP per capita, distance, and visa requirements. Panel data models revealed that immigration was the main determinant of arrivals in Canada. Maheshwari et al. (2018) conducted demand modeling using machine learning techniques for the first 30 airports in the US domestic air transport network. The variables used in this study are the distance between airports, the population of the cities with the relevant airports, and income per capita in the relevant cities. Zhang et al. (2019) investigated the influences of urbanization and other factors on airport CO2 emissions in China using LR and GWR models. Kiracı and Yasar (2020) analyzed the factors that determine the operational performance of airline companies between 1990 and 2017 using a panel data model. He et al. (2021) used GWR, GWL, and Ada-GWL to detect the relationship between station ridership and factors such as the number of restaurants, hotels, etc. within 500 m Pedestrian Catchment Areas (PCA), residents, and percentage of households with 2 or more vehicles. Alnıpak and Kale (2021) investigated the socio-economic factors of airline passenger demand in 23 European countries using a two-stage system generalized moment estimation estimator. Tirtha et al. (2022) used a linear mixed model to analyze the impact of COVID-19 on monthly air passenger departures. Das et al. (2022) used multiple regression models to determine factors such as population, economic status, distance, and travel time, existing major airports, and tertiary education hubs affecting demand for routes connecting individual airport pairs.

In this study, spatial effects are considered because the data of provinces with airports are examined. In addition, socioeconomic variables can often be related to each other. In this case, multicollinearity arises. Hence, penalized GWR models are chosen to detect the relationship between socioeconomic indicators and the number of passengers, aircraft, and amount of freight. Regarding the literature, no previous study has used penalized geographically weighted models to detect relations between socioeconomic variables and the number of passengers, aircraft, and amount of freight in Turkey. This study has two main aims: Its primary purpose is to emphasize that it is more appropriate to use penalized models when there are regional data and multicollinearity between variables; The second main purpose is to investigate whether the number of passengers, aircraft, and the amount of freight have a relationship with regional socioeconomic indicators. In this way, it will be possible to identify which socioeconomic indicators are more effective in different regions. Hence, spatial regression models, such as Geographically Weighted Ridge Regression (GWRR) and Geographically Weighted Lasso Regression (GWL), are used to detect the socioeconomic factors affecting the number of passengers, aircraft, and amount of freight. This study was conducted using 2019 statistical data for 48 airports in Turkey.

2. Data Descrıptıon and mEthodology

2.1. Aviation Industry in Turkey

In Turkey, during 2003, civil air transportation activities have developed very rapidly due to liberal aviation policies. In the last ten years (2010-2019), Flight traffic has grown by a factor of 1.68, the passenger number has increased by a factor of 2.02, and the cargo volume has surged by a factor of 2.81. In this study, 2019 aviation data were used.

In 2019, domestic traffic decreased by 11.5% to 99 946 572, and international traffic increased by 11.1% to 108 427 124 passengers. In this way, the total passenger traffic in 2019 was

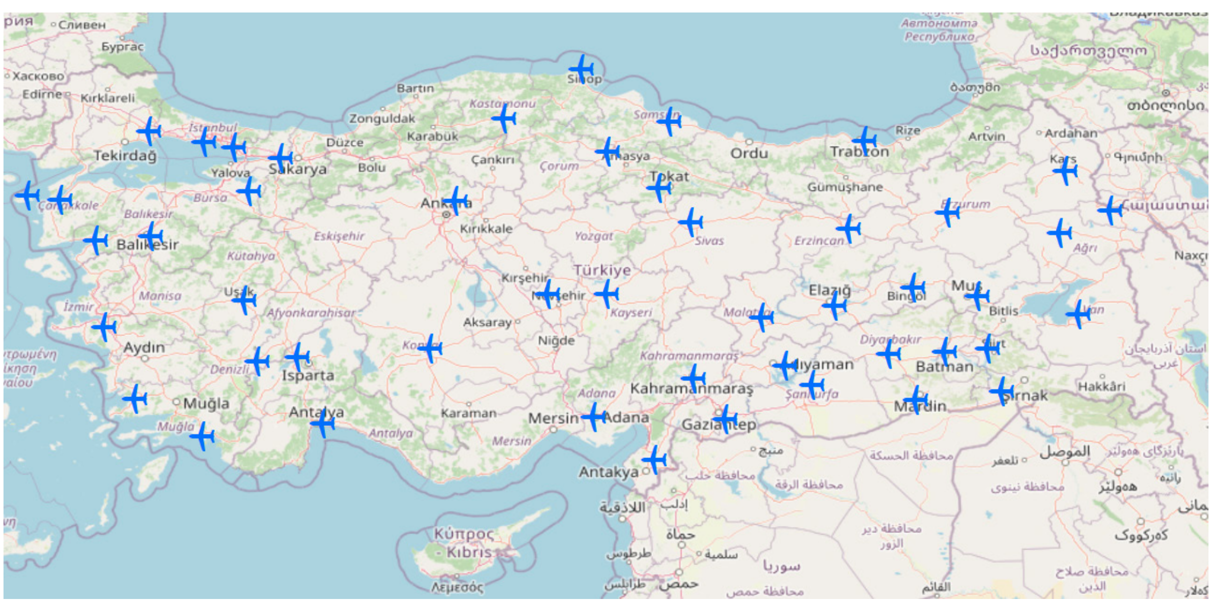

Figure 1. Airports Open to Civil Air Traffic in Turkey (Source: www.dhmi.gov.tr)

208 911 338 passengers, including direct transit (537,642) (Airline Industry Report, 2019).

In 2019, the number of aircraft was 716 523, an increase of 10% compared to 2018. The number of commercial aircraft was 1,339,301 in 2018. In 2019, it was 1 308 970 with a decrease of 2.3%. The amount of domestic commercial air was 718 482 in 2018. In 2019, the number was 623,580, representing a decrease in 13.2% (Airline Industry Report, 2019).

The increase in passenger traffic in our country has begun to reflect the total cargo traffic in the last few years. In 2019, freight traffic; 833 768 tons in domestic lines and 3 256 399 tons in international lines, 4 090 168 tons, an increase of 6,1% compared to 2018. In 2019, 65 667 tons in domestic lines and 1 456 737 tons in international lines, a total of 1 522 404 tons, an increase of 9.6% compared to 2018 (Airline Industry Report, 2019).

Turkey's map regarding Turkish Civil Air Traffic Open Airports was obtained, as shown in Figure 1:

2.2. Data Description

Passenger demand is a crucial driving factor behind the advancement of air transportation. From this point of view, modeling passenger demand holds utmost significance as a core function of airline management and represents a critical concern within the air-transport industry (Srisaeng et al., 2015). Demand modeling is generally conducted at airport or regional levels. The factors affecting demand in the modeling studies include population, education ect. as socio-demographic factors, income, unemployment rate, GDP etc. as socio-economic factors, and the number of trade centers, tourist attractions etc. as built environment factors (Tirtha et al., 2022).

In this study, airport passenger demand is handled at the airport level. The data on airports actively operating in Turkey were examined. The socio-economic indicators of the provinces where the airports are located are considered as factors affecting the number of passengers, number of aircraft, and amount of freight. The socio-economic indicators related to the provinces were determined according to the literature.

The data used in this study were obtained from the Turkish Statistical Institute (TURKSTAT). In this study, the data from 48 airports in Turkey for 2019 were analyzed using penalized GWR models. The variables used in the analysis, variable abbreviations, and institutions from which the variables were obtained are given in Table 1:

Table 1. Variables, variable abbreviations and sources of data

Dependent Variables	Variable Abbreviations	Source of Data
Number of Passengers	PSN	TURKSTAT
Number of Planes	PLN	TURKSTAT
Freight (Tonne)	F	TURKSTAT
Independent Variables		
Imports (1000 TL)	IM	TURKSTAT
Exports (1000 TL)	EX	TURKSTAT
Number of Highway Accidents	HAN	TURKSTAT
Death Toll in Accidents	DTN	TURKSTAT
Number of Cars	CN	TURKSTAT

The descriptive statistics for all variables are presented in Table 2.

Table 2. Descriptive statistics of variables

Dependent Variables	Min.	Mean	Median	Max.	Standart Deviation
PSN	27475	2912430	522761	38315932	7739842
PLN	279	35453	3120	598254	119483.4
F	198	24584	4683	334207	67048.49
Independent Variables					
IM	575	14112313	135051	419732000	62757491
EX	1127	4927214	271877	85636418	17046311
HAN	222	3515	2489	16737	3622
DTN	18	72.82	71	161	36
CN	7	120	131	264	67

This summary statistics is based on the variables' values before pre-processing. The preprocessing step of variables included scaling because of the different scales of variables. Scaling variables helps in making an explicit analysis and improving the stability and performance of regression models (Pourmohammadi et al., 2021).

2.3. Methodology

Linear regression (LR) is used to describe the relationship between dependent and independent(s) variables. The regression model is defined as follows:

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \varepsilon_i;\ i = 1,\ldots,n \quad k = 1,\ldots,p \tag{1}$$

where yi is value of the i-th response variable, x_{ik} is the value of the k-th independent variable for the i-th observation, β_0 is constant term and β_k is the k-th estimated regression coefficient and ε_i is error term with $\varepsilon_i \sim N(0,\sigma^2)$.The regression model in (1) is expressed in matrix form as follows:

$$y = X\beta + \varepsilon \tag{2}$$

where y is an nx1 vector of the dependent variable, X is an nxp matrix of independent variables, β is px1 vector of unknown parameters and ε is nx1 vector of errors.

In practice, the relationships between variables may vary geographically. Local linear regression, known as Geographically Weighted Regression (GWR), is used to examine the relationships between variables that vary geographically. GWR models use a weight matrix that depends on the proximity between observation regions. This means that the closer a region then the weight will be even greater. Unlike global regression models, GWR enables local variations in the estimation of coefficients (Lewandowska, 2018; Millo and Piras, 2012). The GWR model is expressed as follows:

$$y_i = \beta_{0(u_i,v_i)} + \sum_{k=1}^{p} \beta_{k(u_i,v_i)} x_{ik} + \varepsilon_i \tag{3}$$

where (u_i,v_i) is the coordinate location of i. The estimator for each location is

$$\beta = \left[X'W_{(u_i,v_i)}X\right]^{-1} X'W_{(u_i,v_i)}y$$

where

$$W_{(u_i,v_i)} = \operatorname{diag}\left[W_{1(u_i,v_i)},\ldots,W_{n(u_i,v_i)}\right]$$

where is the weight matrix for each location using an exponential kernel function with

$$w_{j(u_i,v_i)} = \exp\left(\frac{d_{j(u_i,v_i)}}{b}\right).$$

where $d_{j\,(u_i,v_i)}$ is Euclidian the distance between location i and location j, and b is the optimal bandwidth.

Before applying GWR, Moran's I index, which is an indicator of whether or not spatial effects exist in the data, should be calculated. The combination of statistically significant p-values and positive z-values indicates a clustered spatial pattern, whereas significant p-values and negative z-values indicate a scattered spatial pattern (Pan et al., 2019; Chioni et al., 2020). Moran's I index is stated below:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

Multicollinearity is a situation where one or more linear correlations between the variables of a regression model. In the presence of multicollinearity, the parameter estimation of linear regression (LR) exhibits a large variance. This problem is also valid for the GWR model. Hence, the model is less appropriate for local multicollinearity (Saputro et al., 2021). To detect local multicollinarity in GWR, the local variance inflation factor (VIF) and local condition index (CI) are used. The local VIF and CI values are

$$VIF_k(u_i,v_i) = \frac{1}{1-R_k^2(u_i,v_i)}$$

$$CI_i = \sqrt{\frac{\lambda_{max}}{\lambda_i}}$$

where R_k (u_i,v_i) is coefficient of determination of x_k and λ_{max} is the largest eigenvalue of k variables and λ_i is the eigenvalue of the ith variable (Wheeler, 2007). Local VIF values indicate local multicollinearity values greater than 5 or 10. The variables with a local CI greater than 30 have local multicollinearity.
To solve the local multicollinearity problem, penalized GWR regression, such as geographically weighted ridge regression (GWRR) or geographically weighted lasso regression (GWL), might be used. The formula used to estimate the GWRR coefficients is

$$\beta = \left[X'W_{(u_i,v_i)}X + \lambda I\right]^{-1} X'W_{(u_i,v_i)}y$$

The ridge parameter λ is determined using generalized cross validation.

The first step of the algorithm for estimating the GWL coefficients is to estimate ridge parameter λ and kernel weight bandwidth value b using generalized cross validation. After determining these values, for each location i, $W_{(ui,vi)}$ is calculated. Then, obtain $Xw = W^{1/2}_{(ui,vi)}$ X and $y_w = W^{1/2}(u_i,v_i)$ y. Finally, the least-angle regression (LARS) algorithm was used to solve LASSO. Detailed information about the procedure of the LARS algorithm can be found in the literature (Wheeler, 2009).

We compare the estimates of the LR, GWR, GWRR, and GWL models by evaluating the root mean square error (RMSE), root mean square prediction error (RMSPE), Akaike Information Criteria (AIC), Corrected Akaike Information Criteria (AICc), and coefficient of determination (R^2):

$$R^2 = \left[\frac{\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}} \right]^2$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2}$$

$$AIC = n\ln(MSE) + 2p$$

$$AICc = n\ln(MSE) + 2p + \frac{2(p+1)(p+2)}{(n-p-2)}$$

where p is the number of variables.

The most suitable model for the data had the highest R^2 and the lowest values for RMSE, RMSPE AIC, and AICc.

3. RESULTS

The global Moran's I is a measure of spatial autocorrelation. Moran's I value for the variables are given in Table 3. A Moran's I statistic greater than 0.3 indicates strong spatial correlation. According to these results, the DTN and CN variables exhibit strong spatial correlation. In addition, the variables that have the lowest (and statistically significant) Moran's I statistic values are PSN, PLN, F, IM, EX, and HAN. As a result, there are spatial relationships among airports (He et al., 2021).

Table 3. Values of Moran's I statistics

Dependent Variables	Moran's Statistic I:	p-value
PSN	0.112	0.039*
PLN	0.119	0.028*
F	0.144	0.007*
Independent Variables		
IM	0.064	0.026*
EX	0.253	0.000*
HAN	0.201	0.003*
DTN	0.365	0.000*
CN	0.812	0.000*

The local VIF and CI values were also used to detect the local multicollinearity in the GWR models. A summary of the local VIF values is presented in Table 4. When examining the local VIF values in Table 4, it is observed that the local VIF values are greater than 10 in five locations, indicating the presence of local multicollinearity.

Table 4. Summary of variance inflation factor (VIF) values

Variable	Min.	Median	Max.	VIF >10
IM	1.709	4.122	209.353	3
EX	1.068	1.093	169.390	2
HAN	1.305	2.829	8.645	0
DTN	1.746	2.497	2.807	0
CN	1.814	2.896	3.321	0

The map of local CI values is shown in Figure 2. As shown in Figure 2, for the 12 locations, the local CI values were greater than 30. This means that there is strong local multicollinearity between the variables. Therefore, penalized GWR models should be used to analyze the data.

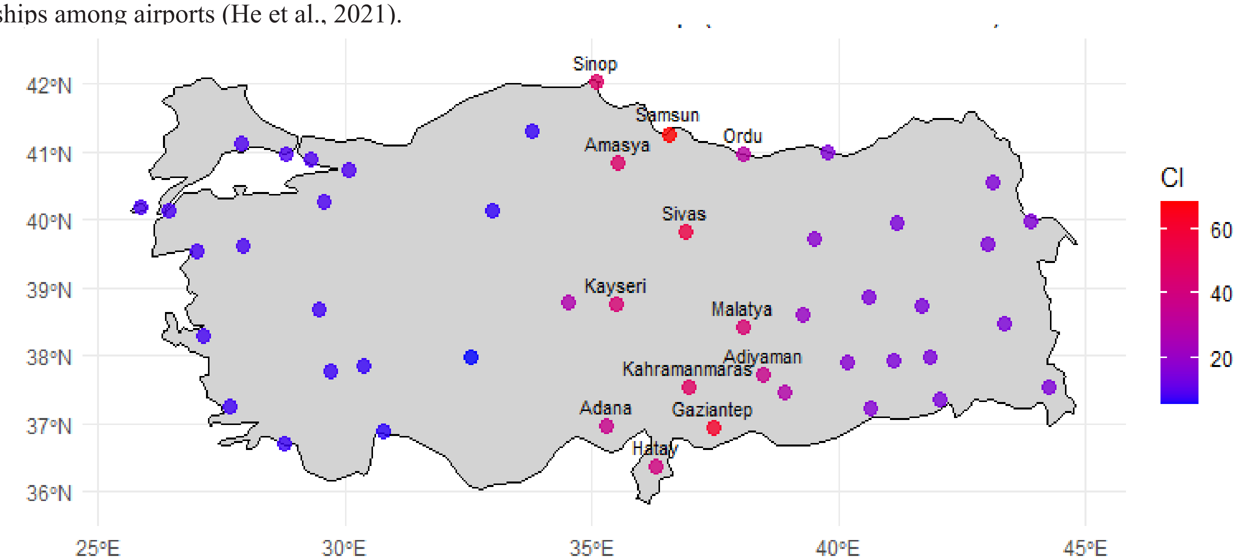

Figure 2. Map of Local Condition Index (CI) Values (Cities with CI>30)

Table 5. Results of the LR models

Variables	Estimate	Std. Error	t-statistics	p-value	VIF value
Passenger					
Constant	0.000	0.0312	0.0000	0.9999	
IM	-0.0306	0.0334	-0.9180	0.3638	1.1184
EX	0.7491	0.0573	13.0720	0.0000**	3.3034
HAN	0.2929	0.0697	4.2020	0.0001**	4.8873
DTN	0.0418	0.0505	0.8290	0.4118	2.5618
CN	-0.0114	0.0551	-0.2080	0.8364	3.0541
Airplane					
Constant	0.0000	0.0268	0.0000	0.9999	
IM	-0.0130	0.0286	-0.4550	0.6516	1.1184
EX	0.8941	0.0492	18.1420	0.0000**	3.3034
HAN	0.1592	0.0599	2.6550	0.0111**	4.8873
DTN	0.0812	0.0434	1.8730	0.0679*	2.5618
CN	-0.0583	0.0473	-1.2320	0.2247	3.0541
Freight					
Constant	0.0000	0.0297	0.0000	0.9999	
IM	-0.0279	0.0317	-0.8790	0.3845	1.1184
EX	0.7913	0.0545	14.4960	0.0000**	3.3034
HAN	0.2581	0.0663	3.8870	0.0003**	4.8873
DTN	0.0561	0.0481	1.1670	0.2495	2.5618
CN	-0.0270	0.0525	-0.5150	0.6094	3.0541

**Significant at 0.05; *Significant at 0.1.

The results of the LR model, along with the comparison between the LR and GWR models, as well as the results for the GWR, GWRR, and GWL models, are presented in Tables 5-7.

Table 6 presents a detailed comparison between the LR and GWR models for different sectors: passenger, aircraft, and freight using the Leung F-tests. The F(1) test compares the residual sum of squares to determine whether GWR offers improved model fit, the F(2) test adjusts the degrees of freedom to assess the improvement factor gained by moving from LR to GWR, and the F(3) test examines regional variations and the spatial distribution of coefficients. Additionally, the F(4) test assesses the overall model fit by comparing the significance of all variables across regions, further indicating whether the GWR model is more appropriate for capturing spatial non-stationarity in the data (Leung et. al., 2000; Yüzbaşı and Görür, 2023; Görür and Yüzbaşı, 2024). For the F(1) test, all sectors show significant results, although the F-values and p-values vary slightly: the Passenger sector is notably significant at p = 0.0026, while the airplane and freight sectors also display strong significance. The F(2) test provides F-values without corresponding p-values across all sectors, indicating that while it calculates the degree of improvement between the LR and GWR models, it does not directly assess statistical significance in the traditional hypothesis-testing sense. A lower F-value suggests a more substantial improvement in fit. F(2) measures the improvement factor through degrees of freedom adjustments; its primary goal is not to establish statistical significance via p-values but rather to quantify the degree of improvement. Thus, p-values are not always calculated, as the test is designed to provide a structural comparison between models rather than a direct significance outcome like in F(1) or F(3) tests. In the F(3) test, the Passenger sector reveals significant p-values for the variable IM at 0.0096 and the constant at 0.0161, these are influential predictors in this model. Conversely, the IM and EX variables show no significant p-values in the Airplane and Freight sectors. The F(4) test results reveal varying degrees of significance across sectors, with all achieving significant p-values but different F-values, indicating that the impact of variables may differ by sector.

Overall, the results in Table 6 indicate that GWR provides nuanced insights that differ significantly from LR, particularly regarding how specific variables influence model outcomes in different sectors. The table effectively highlights these differences, revealing the variable impacts and their statistical

Table 6. Comparison between the LR and GWR

Passenger		Airplane		Freight	
Leung et al. (2000) F(1) test					
F-value	p-value	F-value	p-value	F-value	p-value
0.5862	0.0026**	0.6907	0.0288**	0.6080	0.0048**
Leung et al. (2000) F(2) test					
F-value	p-value	F-value	p-value	F-value	p-value
2.4671	-	2.0965	-	2.3900	-
Leung et al. (2000) F(3) test					
	p-value		p value		p value
Constant	0.0161	Constant	0.3071	Constant	0.1382
IM	0.0096**	IM	0.3411	IM	0.1255
EX	0.7005	EX	0.1143	EX	0.7082
HAN	0.0003**	HAN	0.3869	HAN	0.0004**
DTN	0.0000**	DTN	0.0000**	DTN	0.0000**
CN	0.0000**	CN	0.0682*	CN	0.0141**
Leung et al. (2000) F(4) test					
F-value	p-value	F-value	p-value	F-value	p-value
0.4573	0.0105**	0.5388	0.0337**	0.4742	0.0137**

Table 7. Summary of standardized coefficients of GWR, GWRR and GWL

	GWR			GWRR			GWL		
Passenger									
Variable	Min.	Median	Max.	Min.	Median	Max.	Min.	Median	Max.
Constant	-0.0158	0.0019	0.0105	-0.0847	-0.0392	0.5641	-1.3105	0.0000	0.2990
IM	-0.0315	-0.0305	-0.0299	-0.6676	0.1750	0.8155	-0.0391	0.0000	0.0000
EX	0.7300	0.7459	0.7717	-0.3506	0.1513	0.6948	0.4222	0.7142	0.8280
HAN	0.2444	0.2908	0.3324	-0.3861	0.0676	0.5596	0.0000	0.2567	0.3276
DTN	0.0175	0.0359	0.0666	-0.4233	0.0144	0.4390	-0.0081	0.0000	0.0727
CN	-0.0298	-0.0044	0.0123	-0.2530	0.0248	0.5996	0.0000	0.0000	0.3341
Airplane									
Variable	Min.	Median	Max.	Min.	Median	Max.	Min.	Median	Max.
Constant	-0.0109	0.0006	0.0061	-0.0676	-0.0347	0.5812	-0.5701	0.0000	0.0420
IM	-0.0146	-0.0125	-0.0116	-0.4091	0.1548	0.8094	-0.0189	0.0000	0.0082
EX	0.8758	0.8920	0.9038	-0.4182	0.1295	0.2674	0.6649	0.8768	1.0069
HAN	0.1313	0.1576	0.1888	-0.3676	0.0612	0.4871	0.0000	0.0706	0.2997
DTN	0.0615	0.0753	0.1011	-0.4358	-0.0028	0.1998	0.0000	0.0095	0.1685
CN	-0.0711	-0.0542	-0.0435	-0.2187	0.0199	0.6175	-0.0434	0.0000	0.2889
Freight									
Variable	Min.	Median	Max.	Min.	Median	Max.	Min.	Median	Max.
Constant	-0.0147	0.0018	0.0093	-0.0825	-0.0383	0.5695	-0.7663	-0.0293	0.4814
IM	-0.0291	-0.0276	-0.0268	-0.5045	0.1715	0.8173	-3.5242	0.0000	0.7101
EX	0.7724	0.7915	0.8122	-0.3605	0.1475	0.5254	0.0000	0.7156	2.5780
HAN	0.2117	0.2549	0.2966	-0.3622	0.06358	0.5501	-0.2144	0.1159	0.5178
DTN	0.0318	0.0499	0.0806	-0.4273	0.0164	0.3521	-0.1255	0.0000	0.4521
CN	-0.0455	-0.0203	-0.0040	-0.2475	0.0239	0.6054	-0.2794	0.0208	0.6299

significance across different types of regression models and sectors.

According to the model performance criteria presented in Table 8, the best model is considered the GWL model for the passenger, aircraft, and freight dependent variables. The smallest RMSE, RMSPE, AIC, and AICc values and the highest R2 belong to the GWL among the other models.

The spatial distribution of the local coefficients in the estimated GWL model for passenger demand was obtained as in Figure 3-5.

Table 8. Comparison of models

Passenger				
	LR	**GWR**	**GWRR**	**GWL**
R2	0.9572	0.9619	0.9348	**0.9805**
RMSE	0.2089	0.1931	0.2525	**0.1381**
RMSPE	1.1197	1.2561	0.6401	**0.1476**
AIC	-138.3264	-145.8765	-120.1290	**-178.0586**
AICc	-136.2776	-143.8277	-118.0802	**-176.0098**
Airplane				
R2	0.9684	0.9712	0.9354	**0.9929**
RMSE	0.1796	0.1678	0.2514	**0.0832**
RMSPE	1.1728	2.0234	0.6134	**0.1419**
AIC	-152.8342	-159.3583	-120.5482	**-226.7048**
AICc	-150.7854	-157.3095	-118.4994	**-224.656**
Freight				
R2	0.9612	0.9654	0.9349	**0.9886**
RMSE	0.1990	0.1840	0.2523	**0.1054**
RMSPE	1.1630	1.5374	0.6328	**0.1261**
AIC	-142.9872	-150.5107	-120.2051	**-203.9993**
AICc	-140.9385	-148.4619	-118.1563	**-201.9505**

Figure 3 illustrates a series of maps showing the geographical distribution of significant coefficients for various variables related to passenger data. Each map indicates the locations where the coefficients are statistically significant (blue dots) or not significant (red dots). The maps cover variables such as Intercept, IM, EX, HAN, DTN, and CN. This visualization helps pinpoint where each variable has a significant impact on the passenger movement, providing a clear geographic pattern of influence. For example, the maps show a substantial number of significant blue dots for variables like EX and HAN, which these factors are critical drivers in specific locations. These spatial insights are crucial for understanding regional differences in factors affecting the number of passengers. In addition, when examining the coefficients listed in Table 7, it is observed that the EX and HAN coefficient values have a positive effect on the GWL model. Accordingly, as the EX and HAN increased, the number of passengers was positively influenced.

Figure 4 displays a series of geographic maps illustrating the significance of coefficients for several variables (Intercept, IM, EX, HAN, DTN, and CN) in a spatial analysis context. The maps reveal that most regions are dominated by non-significant coefficients (represented by red dots) for the Intercept, IM, HAN, DTN, and CN variables, suggesting that these variables have little to no impact on those areas. In contrast, the EX variable shows a notable concentration of significant coefficients (blue dots), particularly in central regions, indicating its strong

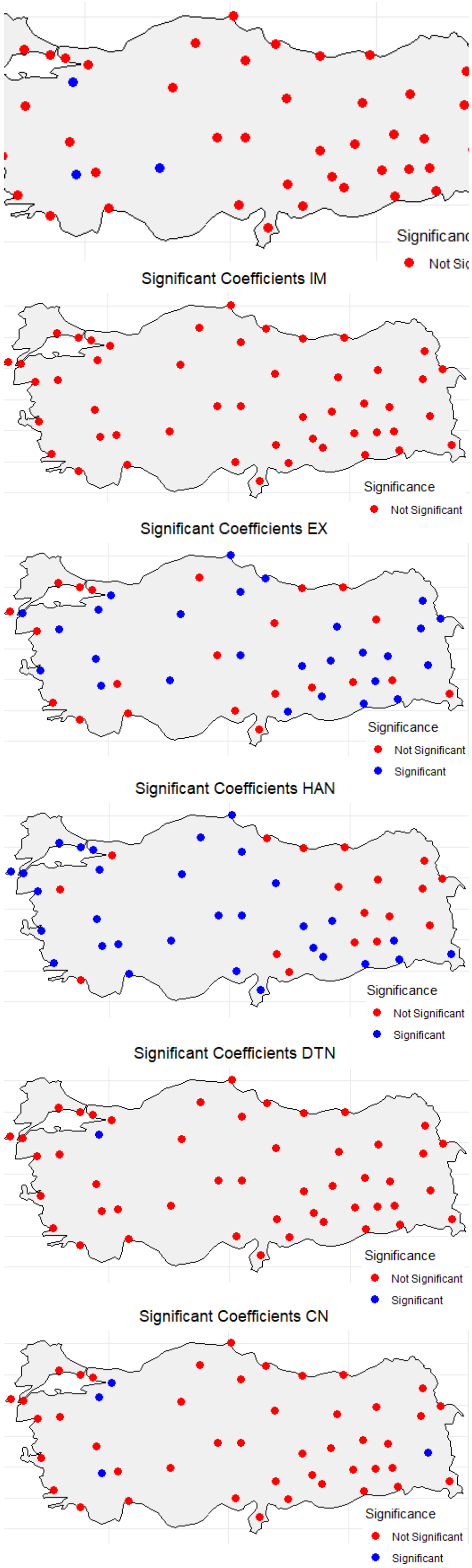

Figure 3. Map of significant coefficients for passenger number

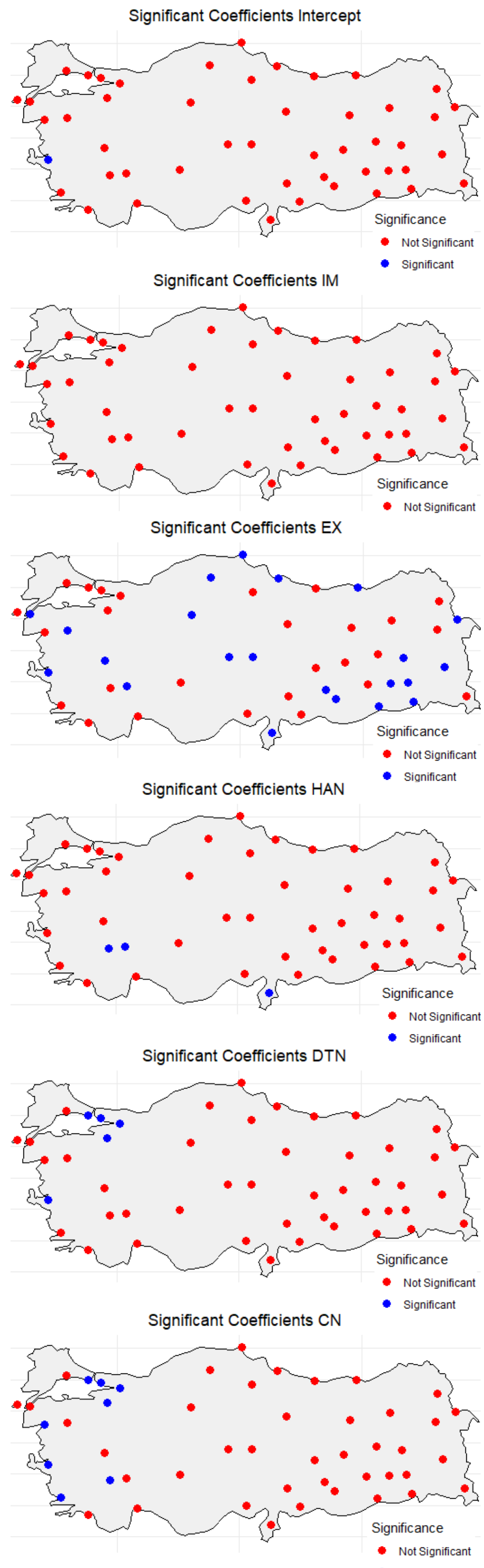

Figure 4. Map of significant coefficients for the number of aircraft

influence there. This spatial visualization clearly demonstrates how the impact of different variables varies across geographic locations, highlighting regions where specific factors play key or negligible roles. In addition, when examining the coefficients in

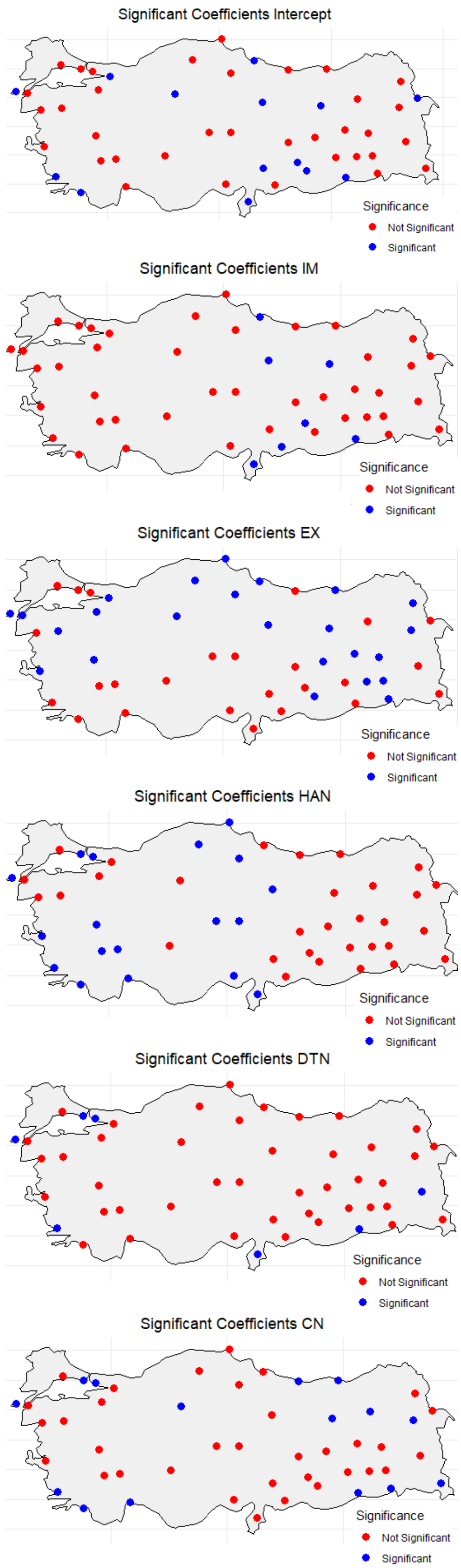

Figure 5. Map of significant coefficients for freight quantity

Table 7, it is observed that the EX coefficient values have a positive effect on the GWL model. Accordingly, as EX increases, the number of aircraft is positively influenced.

Figure 5 displays a series of maps that illustrate the geographical distribution of significant and non-significant coefficients for variables affecting freight volume, including Intercept, IM, EX, HAN, DTN, and CN. The maps show a high density of significant blue dots for variables such as EX and HAN across various regions. This visual analysis helps identify key drivers of the volume of freight between different geographical locations and guides targeted strategies for managing freight logistics. When examining Table 7, it can be seen that the coefficients of EX and HAN are positive. This indicates that as the EX and HAN increase, the freight volume also increases.

5. CONCLUSION

The results of this study demonstrate that there was local multicollinearity between social-economic factors in terms of local variance inflation factors and condition index. Therefore, penalized Geographically Weighted Regression models were used for data analysis. It can be said that the Geographically Weighted Lasso Regression (GWL) had a considerably better fit than the other commonly used penalized models, such as the Geographically Weighted Regression (GWR) and Geographically Weighted Ridge Regression (GWRR), according to the model selection criteria for the data.

The analysis across Figures 3-5 provides a detailed visualization of how different variables impact the geographical distribution of both passenger and freight volumes. In both figures, export (EX) and number of highway accidents (HAN) consistently show a high density of significant dots, indicating that these variables are strong predictors of transport volume across multiple regions. The factors such as EX and HAN are critical for understanding variations in the number of passengers, number of airplanes, and freight movements. Specifically, the EX variable was found to have a strong impact on both passenger and freight transportation, with a notable effect in central regions.

The positive relationship between the increase in exports, the number of highway accidents, and the increase in air passengers and cargo volumes is particularly notable in the Central Anatolia region. Concentration of industrial and commercial activities in this area increases exports and transportation demand. As trade volumes grow, the demand for road transportation rises, which in turn leads to higher traffic density and accident rates. At the same time, this road traffic congestion and risk pushes businesses to opt for air transport as a safer and faster alternative. As a result, air transportation has become more attractive for both passengers and cargo, especially in airports close to major trade hubs in Central Anatolia. This trend can also be attributed to the region's strategic geographic location and proximity to industrial centers.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The author has no conflict of interest to declare.
Grant Support: The author declared that this study has received no financial support.

Hakem Değerlendirmesi: Dış bağımsız.
Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.
Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

REFERENCES

Altuntaş, M., and Kiliç, E. (2021). Havayolu Taşımacılığı Ile Ekonomik Büyüme Arasındaki İlişkinin İncelenmesi: Türkiye Örneği. *Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi* 23(1): 187-202

Alnipak, S. and Kale, S. (2021). Avrupa Bölgesinde Havayolu Yolcu Talebini Etkileyen Faktörler. *International Social Sciences Studies Journal* 7 (90): 4948-4957

Baikgaki, O.A. and Daw, O.D. (2013). The Determinants Of Domestic Air Passenger Demand In The Republic Of South Africa. *Mediterranean Journal Of Social Sciences* 4(13): 389-389

Chen, Z., Barros, C., & Yu, Y. (2017). Spatial Distribution Characteristics of Chinese Airports: A Spatial Cost Function Approach. *Journal Of Air Transport Management*: 59, 63- 70

Chioni, E., Iliopoulou, C., Milioti, C., & Kepaptsoglou, K. (2020). Factors Affecting Bus Bunching At The Stop Level: A Geographically Weighted Regression Approach. *The International Journal Of Transportation Science And Technology* 9(3): 207-217

Choo, Y.Y. (2018). Immigration And Inbound Air Travel Demand In Canada. *Journal Of Air Transport Management* 71: 153-159

Das, A.K., Bardhan, A.K. and Fageda, X. (2022). What Is Driving Passenger Demand on New Regional Air Routes In India: A Study Using The Gravity Model. *Case Study Transport Policy* 10(1): 637-646

Efendigil, T., and Eminler, Ö. (2017). The Importance Of Demand Estimation In The Aviation Sector: A Model To Estimate Airline Passenger Demand. *Journal Of Yasar University* 12: (14)

He, Y., Zhao, Y., & Tsui, K.L. (2021). An Adapted Geographically Weighted Lasso (Ada-Gwl) Model For Predicting Subway Ridership. *Transportation* 48(3).

Görür, Ç., Yüzbaşı, B. (2024). Study of factors influencing cultural similarity in the post-migration adaptation process in the Van Province, using the GWR method. *Cografya Dergisi*, 48, 137-154. https://doi.org/10.26650/JGEOG2024-1353398

Kiraci, K., & Yaşar, M. (2020). The Determinants Of Airline Operational Performance: An Empirical Study On Major World Airlines. *Sosyoekonomi* 28(43): 107-117.

Leung, Y., Mei, C., & Zhang, W. (2000). Statistical tests for spatial nonstationarity based on a geographically weighted regression model. Environment and Planning A, 32(1), 9-32.

Lewandowska, G. K. (2018). Geographically Weighted Regression In The Analysis Of Unemployment In Poland. *Isprs International Journal Of Geo-Information*: 7(1), 17.

Maheshwari, A., Davendralingam, N. and Delaurentis, D.A. (2018). A Comparative Study Of Machine Learning Techniques For Aviation Applications. 2018 Aviation Technology, Integration, And Operations Conference, June 25-29, 2018, Atlanta, Georgia.

Millo, G., & Piras, G. (2012). Splm: Spatial Panel Data Models In R. *Journal Of Statistical Software*, 47(1), 1-38.

Pan, Y., Chen, S., Li, T., Niu, S., & Tang, K. (2019). Exploring Spatial Variations in the Bus Stop Influence Zone With Multi-Source Data: A Case Study In Zhenjiang, China. *Journal Of Transport Policy* 76: 166-177.

Pourmohammadi, P., Strager, M.P., Dougherty, M.J., and Adjeroh, D.A. (2021). Analysis Of Land Development Drivers Using Geographically Weighted Ridge Regression. *Remote Sensing* 13(7): 1307.

Republic Of Turkey Ministry Of Transport And Infrastructure, General Directorate Of State Airports Authority, Airline Industry Report (2019) [Online].Available:Https://Www. Dhmi. Gov. Tr/Lists/ Havayolusektorraporlari/Attachments/ (Accessed December 1, 2020)

Saputro, D. R. S., Hastutik, R. D., &Widyaningsih, P. (2021). Modeling the Human Development Index (Hdi) In Papua—Indonesia Using Geographically Weighted Ridge Regression (Gwrr). Proceedings of the Aip Conference 2326(1), P 020025, Aip Publishing Ltd.

Srisaeng, P., Baxter, G., & Wild, G. (2015). An Artificial Neural Network Approach To Forecast Australia's Domestic Passenger Air Travel Demand. *World Review Of Intermodal Transportation Research* 5(3): 281-313.

Taneja, N.K. (1971). A Model For Forecasting Future Air Travel Demand On The North Atlantic. Cambridge, MA: North Atlantic. Massachusetts Institute Of Technology, Flight Transportation Laboratory.

Tanyel, S., Topal, A., Görkem, Ç., Şengöz, B., & Özuysal, M. (2010). Adnan Menderes Havaalani Yolcu Ve Yük Taleplerinin Değişimi Üzerine Bir İnceleme. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi*, 12(1): 19-33.

Tirtha, S.D., Bhowmik, T., and Eluru, N. (2022). An Airport Level Framework For Examining The Impact Of Covid-19 On Airline Demand. *Transport Research Part A: Policy And Practice* 159: 169-181.

Valdes, V. (2015). Determinants Of Air Travel Demand In Middle Income Countries. *Journal Of Air Transport Management* 42: 75-84.

Wheeler, D. C. (2007). Diagnostic Tools And A Remedial Method For Collinearity In Geographically Weighted Regression. *Environment Planning A* 39(10): 2464-2481.

Wheeler, D. C. (2009). Simultaneous Coefficient Penalization And Model Selection In Geographically Weighted Regression: The Geographically Weighted Lasso. *Environment Planing A* 41(3): 722-742.

Yüzbaşı, B. & Görür, Ç. (2023). Examining Job Role by Geographically Weighted Poisson Regression in Post-migration Adaptation Process: The Case of Van. *Van Yüzüncü Yıl University, Journal of Social Sciences Institute*, Republic Special Issue, 11-27.

Zhang, W., Jiang, L., Cui, Y., Xu, Y., Wang, C., Yu, J., Streets, D., & Lin, B. (2019). Effects Of Urbanization On Airport Co2 Emissions: A Geographically Weighted Approach Using Nighttime Light Data In China. *Resour Conserv Recy. Advances* 150: 104454.