Hindawi
Advances in Civil Engineering
Volume 2021, Article ID 9144215, 10 pages
https://doi.org/10.1155/2021/9144215

Research Article

A Novel Approach for HST Delays Using Pythagorean Fuzzy AHP and Regression Analysis

Hümeyra Bolakar Tosun

Department of Civil Engineering, Faculty of Engineering, Aksaray University, Aksaray 68100, Turkey

Correspondence should be addressed to Hümeyra Bolakar Tosun; bolakarhumeyra@gmail.com

Received 23 September 2021; Revised 7 December 2021; Accepted 9 December 2021; Published 30 December 2021

Academic Editor: Yu Qian

The increased need for transportation worldwide has led to intense competition among several transportation types. Thus, considering factors affecting the choice of transportation means of passengers is essential. In many countries, railways have been losing market share in both freight and passenger transport, especially against highways. Railway transport systems must regain their declining share for the sake of the economy and sustainability. For this reason, many studies have been conducted to eliminate delays in high-speed trains, the speed of which is the most important criterion for preference. This study determines the reasons for train delays in a bid to make the high-speed train project successful in Turkey and for trains to serve better. Furthermore, regression analysis and the Pythagorean fuzzy analytic hierarchy process (AHP) analysis were performed according to the most effective criteria. The most effective criteria were determined as maintenance, repair, and closure due to renewal. Additionally, various suggestions regarding the effect of the obtained causes on train delays were put forward.

1. Introduction

Recently, state railways of the Republic of Turkey have taken serious steps to provide more efficient and faster passenger and freight transportation, and the most important step is high-speed train projects. According to the International Union of Railways (UIC), there are 594 km high-speed train lines operating in Turkey, which is among the top six countries with the largest high-speed train (HST) network in Europe and the top nine countries in the world.

Transportation models have varied according to different periods [1]. Transportation in Europe and North America in the 19th century was mostly by railways. However, toward the 20th century, highway transportation gained importance and was used more. In the 21st century, an increase in the income of people increased travel [2]. With advances in technology, the modes of transport that provides superiority are preferred by people. The ever-increasing supply for rail transport has demanded people work at maximum capacity, targeting the punctuality of service. However, even in developed rail networks with state-of-the-art communication facilities, problems can occur during operations. High-speed railways play a critical role in transportation and transit systems. Thus, recently, high-speed passenger railways have been developed around the world, especially in Europe and China [3]. Hundreds of literatures have evaluated the travel time savings. Passengers want to reach their destination as timely as possible to continue their activities. The reduction in travel time is also viewed favorably from the perspective of the whole society for several reasons. First, such reductions can be transformed into productive activities, resulting in a potential increase in the gross domestic product. Despite advanced communication, delays in train operations are inevitable due to unexpected disruptions, such as poor weather conditions, power outages, and facility failures [4]. The train delays cause significant losses for both railroad operators and passengers. The National Audit Office (NAO), UK, reported approximately 800,000 delays in the British national rail network between 2006 and 2007. This caused a delay of approximately 14 million train minutes lost time for the passengers and up to 1 billion pounds in financial loss [5, 6]. Train delays are categorized into two: primary and secondary delays. Primary delays are related to the train and can lead to other delays, while secondary delays are complex and depend completely on the network [7]. Despite the enforcement of buffer times, train delays are inevitable. People, vehicles, infrastructure, and complex stochastic interactions between them all contribute to the delays [8]. Variables such as late-arriving trains, delays at train stations due to overstayed waiting times, differences in arrival and departure times, and late adjustment of departing train routes due to connecting and overlapping trains also contribute to train delays.

The goal of a rail system line is to transport a passenger load from one route point to another as quickly as possible. To realize this goal, train businesses and services must work efficiently and effectively [9]. There are many literatures on train delays and modeling. Some of these are presented in Table 1.

Many studies have been conducted on evaluating time-saving, which is the main reason why people prefer HST. This study evaluates the effect ratios of independent variables causing delays in HSTs using mathematical methods, such as regression and AHP, and proffer improvement solutions.

The main contribution to academia of this article is to show how to use the AHP and regression models for solving delay on a rail network. This article also provides practical insights by highlighting the datasets available to applications of the feature railway engineering required. This article presented common reasons to minimize the delay with the accuracy and effective results of the methods used. With these results, it will make a great contribution to world politics and academia and will offer solutions for the problems that will arise.

2. Materials and Methods

The statistical data of cities in Turkey were compiled from the State Railways of the Republic of Turkey. Correlation analysis was conducted to determine the relationships between train delays and factors responsible for these delays. Regression analysis was performed to model the relationship between the related factors. In this study, p values less than 0.05 were considered statistically significant ($\alpha = 0.05$). Additionally, independent variables were evaluated as the main criteria and added to the scope of decision-making problems. Numerical values of the main criteria were defined as train delay per minute. These numerical data were evaluated using three decision-makers and their corresponding arithmetic means. The AHP method, a multi-criterion decision-making method based on the dual comparison of the main criteria, was applied in a fuzzy environment. Pythagorean fuzzy sets were integrated into the AHP method to eliminate ambiguity. The steps of the Pythagorean fuzzy AHP method are given below.

Step 1. Construct the compromised pairwise comparison matrix $R = (\text{rik})\ m \times m$ with respect to the opinions of the experts using Table 2.

Step 2. Find the differences matrix $D = (d_{ik})_{m\times m}$ between the lower and upper values of the membership and non-membership functions using the following equations:

$$d_{ikL} = \mu_{ikL}^2 - \nu_{ikU}^2, \tag{1}$$

$$d_{ikU} = \mu_{ikU}^2 - \nu_{ikL}^2. \tag{2}$$

Step 3. Find the interval multiplicative matrix $S = (s_{ik})_{m\times m}$ using the following equations:

$$s_{ik_L} = \sqrt{1000^{\,d_L}}, \tag{3}$$

$$s_{ik_U} = \sqrt{1000^{\,d_U}}. \tag{4}$$

Step 4. Calculate the determinacy value $\tau = (\tau_{ik})_{m\times m}$ of the risk using the following equation:

$$\tau_{ik} = 1 - \left(\mu_{ikU}^2 - \mu_{ikL}^2\right) - \left(\nu_{ikU}^2 - \nu_{ikL}^2\right). \tag{5}$$

Step 5. Multiply the determinacy degrees with the $(S = (s_{ik})_{m\times m})$ matrix to obtain the matrix of the weights, $(T = (t_{ik})_{m\times m})$, before normalization them using the following equation:

$$t_{ik} = \left(\frac{s_{ik_L} + s_{ik_U}}{2}\right)\tau_{ik}. \tag{6}$$

Step 6. Find the normalized priority weights, w_i, using the following equation:

$$w_i = \frac{\sum_{k=1}^{m} \text{tik}}{\sum_{i=1}^{m}\sum_{k=1}^{m} \text{tik}}. \tag{7}$$

The data used in the study are shown in Table 3.

Active HST lines, construction HST lines, and HST connected bus are shown in Figure 1.

2.1. Method

2.1.1. Regression Analysis. Multiple linear regression analysis has two purposes:

(1) Estimating the independent variable through the variables determined to affect the dependent variable

(2) Determining which of the independent variables affecting the dependent variable has a high impact rate [25]

Multiple linear regression analysis requires at least two independent variables. The relationship model between the Y dependent variable and p independent variables is expressed as follows:

$$Y = b0 + b1Xi1 + b2Xi2 + \cdots + bpXip + ei, \quad (i = 1, 2, \&, n), \tag{8}$$

Table 1: Studies in the literature.

Writer	Objective	Methods	Results
Schön and König (2018) [10]	Solving the Bellman equations recursively to minimize the total delay passengers experience at their final stations	We offer stochastic dynamic programming (SDP)	SDP approach outperforms other approaches for a reasonable time resolution to delay.
Rückert et al. (2017) [11]	Time delay information and passenger flow estimates, evaluating impacts of waiting decisions on passenger arrival delays at their final destination	Introduce a web-based simulation tool for dispatchers, called PANDA	Waiting or not waiting at a critical transfer station is based on a majority rule that considers 8 criteria defined as the delay distribution at the passenger's final destination.
Zhu (2011) [12]	To minimize train delays	Scenario-based route choice model	The model was built mainly to deal with predictably large passenger flow events but failed to forecast the network passenger flow distribution under unexpected events.
Berger et al. (2011) [13]	Proposed a delay propagation in large transportation networks, suited to process massive streams of real-time data	Stochastic model	A stochastic model was used for delay.
Corman et al. (2012) [14]	To minimize train delays and missed connections due to disturbances	Applied two heuristic algorithms to select the connections to be enforced	Study shows that good coordination of connected train services is important to achieve real-time efficiency of railway services since the management of connections can heavily affect train punctuality. The two algorithms accurately approximate the Pareto front in a limited computation time.
Krasemann (2012) [15]	Proposed a greedy algorithm that effectively delivers good solutions within the permitted time	Branch-and-cut approach	Detects that for certain scenarios, it is difficult to find good solutions within seconds using a branch-and-cut approach.
Almodovar and Ródenas (2013) [16]	Proposed a model for timetable rescheduling in emergency cases, reallocating trains/buses in real time to other service lines	An optimization approach	This model assumes passengers use travel strategies and waiting passengers are loaded at trains/buses on a first-come-first-served basis. The infrastructure restrictions are not considered by the model.
Milinkovic et al. (2010) [17]	Calculate train delay	Simulation model	Results of simulation are exported to a database for additional data mining and comparative analysis. Model is tested on a part of Belgrade Railway Node. Train delays are calculated in a simulation model using a Fuzzy Petri Net subsystem.
Markovic et al. (2015) [18]	This paper proposed machine-learning models that capture the relationship between passenger train-arrival delays and various characteristics of a railway system	Support vector regression and artificial neural network	Statistical comparison of the two models indicates that the support vector regression outperforms the artificial neural network.
Wang et al. (2021) [19]	The purpose of this study was to investigate how the winter weather precipitation affect the occurrence of primary delays and the transitions between delayed and nondelayed states.	Cox proportional hazard model and Markov chain model	Markov chain model to the train operation data is more reasonable, since it is strict to assume the transition intensity does not change over time in reality
Hou et al. (2020) [20]	This paper proposed to determine the effects of two train operation adjustment actions on train delay recovery were explored using train operation records from scheduled and actual train timetables.	Gradient-boosted regression tree (GBRT) machine-learning model	A comparison of the prediction results of the GBRT model with those of a random forest model confirmed the better performance of the GBRT prediction model.

Table 1: Continued.

Writer	Objective	Methods	Results
Jiang et al. (2019) [21]	This paper aimed to develop primary delay recovery (PDR) predictor model using train operation records from Wuhan-Guangzhou (W-G) high-speed railway.	Random forest regression (RFR) model, multiple linear regression (MLR), support vector machine (SVM), and artificial neural networks (ANN)	RFR model can achieve up to 80.4% of prediction accuracy, while the accuracy level is 44.4%, 78.5%, and 78.5% for MLR, SVM, and ANN models, respectively.
Wang and Zang (2019) [22]	The aim of study to determine patterns of train delays and to predict train delay time	Machine-learning model	The prediction model is useful not only for passengers wishing to plan their journeys more reliably, but also for railway operators developing more efficient train schedules and more reasonable pricing plans.
Corman and Kecman (2018) [23]	This paper aimed to present a stochastic model for predicting the propagation of train delays based on Bayesian networks.	Bayesian networks	The presented method is important for making better predictions for train traffic that are not only based on static, offline collected data, but are able to positively include the dynamic characteristics of the continuously changing delays.

Table 2: Pythagorean fuzzy AHP weighting scale [24].

	Linguistic term	Pythagoras fuzzy numbers			
		μ_L	μ_U	ν_L	ν_U
ALI	Absolutely low importance	0.00	0.00	0.90	1.00
LI	Low importance	0.20	0.35	0.65	0.80
BMI	Below moderately important	0.35	0.45	0.55	0.65
MI	Moderately importance	0.45	0.55	0.45	0.55
AMI	Above moderately importance	0.55	0.65	0.35	0.45
HI	High importance	0.65	0.80	0.20	0.35
VHI	Very high importance	0.80	0.90	0.10	0.20
AHI	Absolutely high importance	0.90	1.00	0.00	0.00
E	Equal	0.1965	0.1965	0.1965	0.1965

Table 3: Database of the study.

Code	
D1	Maintenance facilities
D2	Maintenance road
D3	Transport traction
D4	Transport passenger
D5	Transport freight
D6	Meeting-getting ahead-accident-incident
D7	Reasons out of organization

where b_0, b_1,..., b_p unknowns are partial regression coefficients.

In multiple linear regression analysis, multiple correlation coefficients show the strength of the relationship between the dependent and independent variables. An unlimited number of independent variables that explain the dependent variable can exist [26]. These situations were expressed with "p" values in this study. The correlation coefficient "r" is the coefficient that indicates of the relationship between the independent variables. This coefficient takes a value between (−1) and (+1). Positive values indicate direct linear relationship; negative values indicate an inverse linear relationship.

2.1.2. Pythagorean Fuzzy AHP. Objective and subjective criteria can be compared using the AHP method by considering some specific criteria [27, 28]. Although the AHP method receives information from experts, it does not reflect people's thoughts. Therefore, fuzzy AHP is achieved by combining AHP with fuzzy logic. AHP methodology determines the weight of any qualitative criteria (inputs or outputs). This is quite important for systems where some of their performance measures are qualitative, such as railway and production systems [29].

3. Results and Discussion

The regression analysis results (Table 4) show that no significant relationship exists between the D1 variable and HST delays ($r = 0.39$, p value $= 0.43$, p value >0.05). Summarily, the delays were not affected by the total time for the D1 variable.

Upon examining the D2 variable, there was a strong and significant positive relationship with the dependent variable. We can state that HST delays were affected by the total time spent on maintenance and road, and an increase in this time can increase the delay times ($r = 0.92$, p value $= 0.01$, p value <0.05).

Figure 1: Destinations of the HSTs from Ankara.

Table 4: Investigation of relationships between variables affecting delay.

Variables	HST delay	
	r	p value
D1	0.39	0.43
D2	0.92	0.01
D3	0.88	0.02
D4	0.76	0.07
D5	0.01	0.99
D6	0.71	0.11
D7	0.90	0.01

Moreover, there was a strong and significant positive relationship between the D3 independent variable and the delays ($r = 0.88$, p value $= 0.01$, p value <0.05). Thus, the increase in this period increased the HST delays.

Upon examining the D4 variable, no significant relationship between the HSR delays and the D4 variable was observed. The delays were not affected by the total time spent for passenger transportation ($r = 0.76$, p value $= 0.07$, p value >0.05).

Similarly, examining the D5 variable, there was no significant relationship between the dependent variable and the HST delays ($r = 0.01$, p value $= 0.99$, p value >0.05). The delays were not significantly affected by this variable.

Additionally, no significant relationship exists between the D6 variable and HST delay when the relevant data were examined ($r = 0.71$, p value $= 0.11$, p value >0.05).

Correlation analysis shows a significant relationship.

Furthermore, there was a strong and significant positive relationship between the D7 independent variable and HSR delay time. The effect of the D7 variable on the delays was considerably high ($r = 0.90$, p value $= 0.01$, p value <0.05).

3.1. Multiple and Linear Modeling of the Relationships between the Variables Affecting Delays. The regression model modeled the relationship between the variables at multiple levels. The model was presented after verifying whether the dependent variables related to the independent variables were related on multiple levels by testing the significance of the determined model (F), the explanation of the independent variables (R^2) (the variables represent D1–D7), and the significance of the coefficients (t). Meeting these three conditions showed that the model was statistically valid.

From Table 5, a significant relationship between D2, D3, and D6 can be seen. No significant relationship between the other variables and HST delays ($F = 13.57$, p value $= 0.01$, p value <0.05) is noticed.

The explanation percentage of the model was 74% ($R^2 = 0.74$) and considered high.

Also, the coefficients of the D2, D3, and D6 variables were significant. (p value $= 0.01$, p value <0.05). The Durbin Watson test was conducted to examine the presence of autocorrelation in the model, and the results showed that there was no autocorrelation (D. W. = 1.83). Thus, the model was found to be significant.

From the results, the most significant variable affecting the HST delays was D2. The effect of the D6 variable was close to that of D2 but at a lower and negative level.

From the study, one unit increase in total time spent for the D2 variable caused an increase in the HST delay time by 0.62 units, while a unit increase in the D3 variable increased the dependent variable delay time by 0.40 units. However,

Table 5: Multiple and linear modeling of the relationships between variables affecting the delay.

The dependent variable	Independent variables D2 (β)	D3 (β)	D6 (β)	F model	R^2
HST delay (Y)	0.62	0.40	−0.58	$F=13.57$	0.74

Regression analysis applied, D.W. = 1.83.

Table 6: Pairwise comparison of subcriteria according to D2 main criterion.

	D21	D22	D23	D24	D25
D21	0.19	0.9	0.8666	0.9	0.7
	0.19	1	0.9666	1	0.8333
	0.19	0	0.0333	0	0.1666
	0.19	0	0.0666	0	0.3
D22	0	0.19	0.3	0.45	0.2
	0	0.19	0.4166	0.55	0.35
	0.9	0.19	0.5833	0.45	0.65
	1	0.19	0.7	0.55	0.8
D23	0.0333	0.5833	0.19	0.5833	0.3583
	0.0666	0.7	0.19	0.7	0.4166
	0.8666	0.3	0.19	0.3	0.5277
	0.9666	0.4166	0.19	0.4166	0.4722
D24	0	0.45	0.3	0.19	0.2
	0	0.55	0.4166	0.19	0.35
	0.9	0.45	0.5833	0.19	0.65
	1	0.55	0.7	0.19	0.8
D25	0.1666	0.65	0.5277	0.65	0.19
	0.3	0.8	0.4722	0.8	0.19
	0.7	0.2	0.3583	0.2	0.19
	0.8333	0.35	0.4166	0.35	0.19

the D6 independent variable negatively affected delays and its level corresponding to one unit was 0.58.

3.2. Pythagorean Fuzzy AHP. Since knowledge can be expressed in a more natural by using fuzzy sets, many engineering and decision problems can be easily. Decision-makers usually find that it is more confident to give interval judgments than fixed-value judgments. This is because generally he/she is unable to explicit about his/her preferences due to the fuzzy nature of the comparison process. In the study, seven main criteria and 27 subcriteria that cause train delays were determined. As the initial stage, 3 decision-makers evaluated the main criteria and subcriteria using pairwise comparison matrices. Then, the effect weights of the main criteria and subcriteria were calculated using the Pythagoras fuzzy clusters in the fuzzy environment of the AHP method.

The pairwise comparison matrix of the five subcriteria determined in the problem from the D2 main criterion was created in Table 6 using the weighting scale provided in Table 2.

Then, the difference matrix between lower and upper values of the membership and nonmembership functions is created using equations (1) and (2), which are given in Table 7. Table 8 showed the interval multiplicative matrix that obtained using equations (3) and (4). The determinacy value that is calculated with the help of equation (5) is given in Table 9. Finally, the weights matrix and the normalized priority weights given in Table 10 are computed using equations (6) and (7).

The interval multiplicative matrix was created by using the difference matrices of subcriteria equations (3) and (4) and given in Table 8.

The determinacy value was created using equation (5) and provided in Table 9.

Unnormalized weights were calculated for each subcriterion of the D2 main criterion using equations (6) and (7) provided in Table 10.

Using the comparison values provided by the 3 decision-makers using Table 2, the pairwise comparison matrix of the main criteria was created as in Table 11.

The normalized weights of each of the main criteria using the Pythagorean Fuzzy AHP method are given in Tables 12 and 13.

According to the results of Table 12, the most important criterion was determined as D2 with a rate of %36.52. This was followed by %25.61 D3, %12.28 D1, %9.87 D5, %7.24 D4, %6.84 D6, and %1.62 D7. As in the results, it was determined that the most important for this difference was the D2 criterion.

In this study, I have investigated the determination of weight of criteria method in a decision-making process under Pythagorean fuzzy sets and proposed Pythagorean

Table 7: Difference matrix.

	D21		D22		D23		D24		D25	
D21	1	1	16.4059	31.62278	25.11886	16.4059	16.4059	31.62278	3.981072	10
D22	0.031623	0.060954	1	1	0.251189	0.562341	0.707946	1.412538	0.125893	0.354813
D23	0.039811	0.075858	1.778279	1	1	1	1.778279	3.981072	0.721296	0.695973
D24	0.031623	0.060954	0.707946	1.412538	0.251189	0.562341	1	1	0.125893	0.354813
D25	0.1	0.251189	2.818383	7.943282	1.436837	1.386394	2.818383	7.943282	1	1

Table 8: Interval multiplicative matrix.

	D21	D22	D23	D24	D25
D21	1	0.81	0.8133	0.81	0.7333
D22	0.81	1	0.7666	0.8	0.7
D23	0.8133	0.7666	1	0.7666	1.0103
D24	0.81	0.8	0.7666	1	0.7
D25	0.7333	0.7	1.0103	0.7	1

Table 9: Determinacy value matrix.

	D21	D22	D23	D24	D25	Wİ
D21	1	19.4516	16.8867	19.4516	5.1263	0.7571
D22	0.0374	1	0.3118	0.8481	0.1682	0.0289
D23	0.0470	1.0650	1	2.2077	0.7159	0.0615
D24	0.0374	0.8481	0.3118	1	0.1682	0.0289
D25	0.1287	3.7665	1.4262	3.7665	1	0.1233

Table 10: Weight matrix before normalization.

	D21		D22		D23		D24		D25	
D21	0	0	0.81	1	0.9333	0.81	0.81	1	0.4	0.6666
D22	−1	−0.81	0	0	−0.4	−0.1666	−0.1	0.1	−0.6	−0.3
D23	−0.9333	−0.7466	0.1666	0	0	0	0.1666	0.4	−0.0946	−0.1049
D24	−1	−0.81	−0.1	0.1	−0.4	−0.1666	0	0	−0.6	−0.3
D25	−0.6666	−0.4	0.3	0.6	0.1049	0.0946	0.3	0.6	0	0

Table 11: Pairwise comparison of main criteria.

	D1	D2	D3	D4	D5	D6	D7
D1	0.19	0.3	0.4166	0.5833	0.4833	0.5833	0.8333
	0.19	0.4166	0.5166	0.7	0.5833	0.7	0.9333
	0.19	0.5833	0.4833	0.3	0.4166	0.3	0.0666
	0.19	0.7	0.5833	0.4166	0.5166	0.4166	0.1333
D2	0.5833	0.19	0.4833	0.7	0.5833	0.7	0.9
	0.7	0.19	0.5833	0.8333	0.7	0.8333	1
	0.3	0.19	0.4166	0.1666	0.3	0.1666	0
	0.4166	0.19	0.5166	0.3	0.4166	0.3	0
D3	0.4833	0.4166	0.19	0.6166	0.5166	0.6166	0.8666
	0.5833	0.5166	0.19	0.75	0.6166	0.75	0.9666
	0.4166	0.4833	0.19	0.25	0.3833	0.25	0.0333
	0.5166	0.5833	0.19	0.3833	0.4833	0.3833	0.0666
D4	0.3	0.1666	0.25	0.19	0.4166	0.4833	0.65
	0.4166	0.3	0.3833	0.19	0.5166	0.5833	0.8
	0.5833	0.7	0.6166	0.19	0.4833	0.4166	0.2
	0.7	0.8333	0.75	0.19	0.5833	0.5166	0.35

Table 11: Continued.

	D1	D2	D3	D4	D5	D6	D7
D5	0.4166	0.3	0.3833	0.4833	0.19	0.4833	0.7
	0.5166	0.4166	0.4833	0.5833	0.19	0.5833	0.8333
	0.4833	0.5833	0.5166	0.4166	0.19	0.4166	0.1666
	0.5833	0.7	0.6166	0.5166	0.19	0.5166	0.3
D6	0.3	0.1666	0.25	0.4166	0.4166	0.19	0.65
	0.4166	0.3	0.3833	0.5166	0.5166	0.19	0.8
	0.5833	0.7	0.6166	0.4833	0.4833	0.19	0.2
	0.7	0.8333	0.75	0.5833	0.5833	0.19	0.35
D7	0.0666	0	0.0333	0.2	0.1666	0.2	0.19
	0.1333	0	0.0666	0.35	0.3	0.35	0.19
	0.8333	0.9	0.8666	0.65	0.7	0.65	0.19
	0.9333	1	0.9666	0.8	0.8333	0.8	0.19

Table 12: Weights of criteria.

	WI
D1	0.1228
D2	0.3652
D3	0.2561
D4	0.0724
D5	0.0987
D6	0.0684
D7	0.0162

Table 13: Normalized criteria weights for each main criterion and subcriteria.

D1	Due to on-board equipment failure	0.1228	0.1744
	Delay due to communication failure		0.0154
	Catenary failure and power cut		0.0415
	Due to roadside Automatic Train Stop (ATS) system failure		0.0665
	Due to roadside Europe Railway Train Management System (ERTMS) failure		0.2634
	Roadside signaling system failure		0.4387
D2	Shutdown due to maintenance/repair/renewal (D21)	0.3652	0.7571
	Delay due to crossings with guard barrier and other barriers (D22)		0.0289
	Delay as per 5588 model (D23)		0.0616
	Waiting as a result of rail break (D24)		0.0289
	Switch failure caused by road service (D25)		0.1233
D3	Delays due to meeting and getting ahead	0.2561	0.5211
	Waiting for security reasons		0.0698
	Lost cruise due to weather conditions		0.2849
	Transfer due to accident and incident		0.1061
	Due to natural disaster, snow, freezing, and falling stones		0.0179
D4	Effect of main or support locomotive/train set	0.0724	0.1467
	The locomotive's inability to pull its load		0.2104
	Train inspection of wagon technicians		0.0436
	Wagons and arrays/train sets		0.1265
	Lost cruise due to malfunction in the wagon		0.4726
D5	Due to traffic-related work and transactions	0.0987	0.7595
	Stop/wait as per written/oral order		0.2404
D6	Arriving passenger and waiting for train/bus	0.0684	0.1555
	Waiting for passenger getting on and off		0.8444
D7	Vandalism, stone throwing, glass breakage, etc.	0.0162	0.3461
	Stopping the train by passengers and other persons		0.6538

fuzzy sets to AHP to determine the weights of criteria. A numerical example is considered to illustrate the Pythagorean fuzzy number to the AHP method. The main contribution in this study is developing a new approach to find weights of criteria based on Pythagorean fuzzy numbers and applied to AHP. Then, providing the numerical examples to show the practicality and effectiveness of weight of criteria using Pythagorean fuzzy sets. Analytic hierarchy process has been widely used as a useful multiple-criterion decision-making tool in many areas, such as selection, evaluation, planning and development, decision-making, forecasting, and so on [30].

From the results, the D2 main criterion showed the highest impact, followed by D3, D1, D5, D4, D6, and D7, respectively. Among the subcriteria of the D2 main criterion, D21 had the highest impact value, followed by D25, D23, D22, and D24, respectively. These methods have been recently developed to use in many study. Academically, further research may be the application of these methods to the supplier selection problem and the comparison of the results.

4. Conclusions

This study analyzed the relationship between train delays and various characteristics of the railway system geared toward planning changes and investments to reduce delays. Accordingly, the most effective criteria highlighted were maintenance, repair, and closure due to renewal. Potential implementations arising from the variables considered were examined, and solutions were presented relative to the most affecting criteria. This study includes the following contributions:

(1) When an infrastructure-related issue is detected, the operator restricting to a temporary speed until the issue is resolved will prevent delays

(2) Maintenance and repair teams must be assigned at the right time intervals for various tasks depending on both traffic conditions and the priorities of the projects

(3) Establishing a functional relationship between train delays and the characteristics of the railway system will be useful for planning

(4) When estimating delays, considering interactions between trains, stations, and weather-related factors in terms of prediction accuracy is useful

Data Availability

All data generated or analysed during this study are included in this published article.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The author thanks State Railways of the Republic of Turkey for giving the data.

References

[1] D. Cook, P. Dixon, W. M. Duckworth et al., "Binary Response and Logistic Regression Analysis," 2021, http://www.public.iastate.

[2] A. Schafer and D. Victor, "The past and future of global mobility," *Scientific American*, vol. 277, no. 4, pp. 58–61, 1997.

[3] R. Vickerman, "High-speed rail and regional development: the case of intermediate stations," *Journal of Transport Geography*, vol. 42, pp. 157–165, 2015.

[4] H. Khadilkar, "Data-enabled stochastic modeling for evaluating schedule robustness of railway networks," *Transportation Science*, vol. 51, no. 4, pp. 1161–1176, 2017.

[5] T. Burr, S. Merrifield, D. Duffy, J. Griffiths, S. Wright, and G. Barker, *Reducing Passenger Rail Delays by Better Management of Incidents*, Stationary Office, London, England, UK, 2008.

[6] J. Preston, G. Wall, R. Batley, J. N. Ibáñez, and J. Shires, "Impact of delays on passenger train services," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2117, no. 1, pp. 14–23, 2009.

[7] L. E. Meester and S. Muns, "Stochastic delay propagation in railway networks and phase-type distributions," *Transportation Research Part B: Methodological*, vol. 41, no. 2, pp. 218–230, 2007.

[8] R. Goverde, I. A. Hansen, G. Hooghıemstra, and H. P. Lopuhaa, "Statistical analysis of train traffic: the Eindhoven case," *TRAIL studies in. Transportation Science*, Delft University Press, Delft, Netherlands, 2001.

[9] M. T. Akçay and İ. Kocaarslan, "Matlab/simulink yardımıyla çok-araçlı sinyalizasyon sisteminin simülasyonu ve tren zaman çizelgesinin tasarımı," *El-Cezeri Fen ve Mühendislik Dergisi*, vol. 6, pp. 799–807, 2019.

[10] C. Schön and E. König, "A stochastic dynamic programming approach for delay management of a single train line," *European Journal of Operational Research*, vol. 271, no. 2, pp. 501–518, 2018.

[11] R. Rückert, M. Lemnian, C. Blendinger, S. Rechner, and M. H. Müller, "Panda: a software tool for improved train dispatching with focus on passenger flows," *Public Transport*, vol. 9, no. 1-2, pp. 307–324, 2017.

[12] W. Zhu, *Research on the Model and Algorithm of Mass Passenger Flow Distribution in Network for Urban Rail Transit*, (PhD Thesis), Tongji University, Shanghai, China, 2011.

[13] A. Berger, G. Andreas, M. H. Matthias, and O. Martin, "Stochastic delay prediction in large train networks," in *Proceedings of the 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, A. Caprara and S. Kontogiannis, Eds., pp. 100–111, Dagstuhl, Saarbrücken, Germany, September 2011.

[14] F. Corman, A. D'Ariano, D. Pacciarelli, and M. Pranzo, "Bi-objective conflict detection and resolution in railway traffic management," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 79–94, 2012.

[15] J. T. Krasemann, "Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 62–78, 2012.

[16] M. Almodóvar and R. G. Ródenas, "On-line reschedule optimization for passenger railways in case of emergencies," *Computers & Operations Research*, vol. 40, no. 3, pp. 725–736, 2013.

[17] S. Milinkovic, M. Markovic, S. Veskovic, M. Ivic, and N. Pavlovic, "A Fuzzy petri net model for estimation of train delays," in *Proceedings of the 7th EUROSIM congress on Modelling and Simulation*, Prague, Czech Republic, September 2010.

[18] N. Marković, S. Milinković, K. S. Tikhonov, and P. Schonfeld, "Analyzing passenger train arrival delays with support vector regression," *Transportation Research Part C: Emerging Technologies*, vol. 56, pp. 251–262, 2015.

[19] J. Wang, M. Granlöf, and J. Yu, "Effects of winter climate on delays of high speed passenger trains in Botnia-Atlantica region," *Journal of Rail Transport Planning & Management*, vol. 18, no. 1, Article ID 100251, 2021.

[20] Y. Hou, C. Wen, P. Huang, L. Fu, and C. Jiang, "Delay recovery model for high-speed trains with compressed train dwell time and running time," *Railway Engineering Science*, vol. 28, no. 4, pp. 424–434, 2020.

[21] C. Jiang, P. Huang, J. Lessan, L. Fu, and C. Wen, "Forecasting primary delay recovery of high-speed railway using multiple linear regression, supporting vector machine, artificial neural network, and random forest regression," *Canadian Journal of Civil Engineering*, vol. 46, no. 5, pp. 353–363, 2019.

[22] P Wang and Q. P. Zhang, "Train delay analysis and prediction based on big data fusion," *Transportation Safety and Environment*, vol. 1, no. 1, pp. 79–88, 2019.

[23] F. Corman and P. Kecman, "Stochastic prediction of train delays in real-time using Bayesian networks," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 599–615, 2018.

[24] E. Ilbahar, A. Karaşan, S. Cebi, and C. Kahraman, "A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system," *Safety Science*, vol. 103, pp. 124–136, 2018.

[25] R. Alpar, *Introduction to Applied Multivariate Statistical Methods I*, Nobel, Ankara, Turkey, 2003.

[26] Ö. Ünver and H. Gamgam, *Applied Basic Statistical Methods*, Seckin Bookstore, Ankara, Turkey, 2008.

[27] F. Çavdur, A. Sebatlı, and M. K. Küçük, "öğrenci-proje takımı oluşturma problemi için grup-karar verme ve hedef programlama temelli çözüm yaklaşımı," *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, vol. 34, no. 1, pp. 505–521, 2019.

[28] M. A. Y. Yerİikaya and F. A. Arikan, "KOBİ'lere sağlanan desteklerin performans etkinlik sıralarının Promethee ve Oreste yöntemleri ile belirlenmesi," *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, vol. 31, no. 4, pp. 1007–1016, 2016.

[29] A. Azadeh, S. F. Ghaderi, and H. Izadbakhsh, "Integration of DEA and AHP with computer simulation for railway system improvement and optimization," *Applied Mathematics and Computation*, vol. 195, no. 2, pp. 775–785, 2008.

[30] O. S. Vaidya and S. Kumar, "Analytic hierarchy process: an overview of applications," *European Journal of Operational Research*, vol. 169, no. 1, pp. 1–29, 2006.