Journal of Construction Engineering, Management & Innovation Volume 9 Issue 1 (2026) 622

Contents lists available at www.goldenlightpublish.com/journals/jcemi/2-toc

Journal of Construction Engineering, Management & Innovation

Journal homepage: www.goldenlightpublish.com/journals/2/jcemi

Evaluating rail transit system delays: A statistical evaluation of variable changes

Yunus Dinç, Başak Varli Bingöl*

Cankırı Karatekin University, Department of Civil Engineering, Cankırı, Türkiye

* Corresponding author: B. Varli Bingöl (basakvarlibingol@karatekin.edu.tr)

https://doi.org/10.31462/jcemi.2026.622

Received 30 November 2025; Revised 11 January 2026; Accepted 13 February 2026; Available online 26 March 2026

Keywords

Thermal comfort prediction
Service delays
Statistical analyses
Welch's anova
Railway system
Transportation efficiency

Abstract

This study proposes to investigate the effects of service delays from various causes on the rail system throughout 2021. The study analyzed the frequency of delays from five distinct sources: train-related delays, passenger-related delays, operational delays, equipment-related delays, and system-related delays. Statistically significant differences among these sources were evaluated using Welch's One-Way ANOVA and Games-Howell post-hoc analyses. To quantify the magnitude of impact, η^2 and ω^2 effect sizes were calculated." The findings indicate that train-related delays constitute the predominant source of disruption within the transportation system. Train-related delay, greatly exceeding those from other sources, and notably affected the total delay performance of the transportation system. Delays associated with passengers exhibited notable discrepancies solely in relation to train delays and non-system delays, but not in connection with operational ones or delays connected to system equipment. Operational lags were markedly distinct solely in relation to train-related delays, as opposed to the other categories. Retardation attributable to system equipment exhibited a notable disparity in comparison to train-related delays. Ultimately, delays attributed to non-system sources markedly differed from those caused by trains and passengers, although were not dissimilar to operational or system equipment delays. The findings indicate that minimizing train-related delays is the paramount factor for enhancing the overall efficacy of the transportation system. Future research should ascertain the underlying reasons of train delays and devise appropriate methods to alleviate this issue.

1. Introduction

Rail networks are regarded as a crucial element of sustainable transportation policy. The environmental, economic, social, and safety benefits of rail systems are becoming more prominent, especially in urban transportation, which confronts issues including rising vehicle counts, fuel consumption, and air pollution. A multitude of research in the literature illustrates that rail systems provide substantial advantages for energy efficiency and social well-being [1–6].

From an environmental aspect standpoint, the primary benefit of rail systems is their minimal emissions and superior energy efficiency. Xue et al. [1] determined that rail transit generates fewer carbon emissions per passenger compared to automobile transport and offers considerable energy savings. Modern metro and high-speed train systems with elevated electrification levels utilize regenerative braking technologies for energy recovery, hence diminishing operational expenses. Research conducted by Yang et al. [2] demonstrated that energy efficiency in urban rail transit remains consistent despite fluctuations in weather conditions and passenger volume. These findings affirm that rail systems are a transportation modality aligned with long-term climate objectives. Table 1 summarizes the advantages of rail transit systems as; environmental, economic, social, and safety considerations.

A significant benefit of rail lines is their capacity to alleviate urban traffic congestion due to their substantial carrying capacity. Rail lines may transport a greater number of passengers during the same time-frame compared to cars with high-frequency services on congested corridors. A thorough systematic evaluation of "Transportation Rail Networks" [3] demonstrated that rail networks mitigate traffic congestion and decrease travel duration by enhancing transportation efficiency in urban areas. This outcome illustrates that rail investments are crucial for both environmental and operational efficiency. Rail systems demonstrate a distinct benefit regarding safety. Rail systems, exhibiting significantly lower accident rates than road traffic, enhance passenger safety and contribute to public health.

Table 1. Comprehensive advantages of sustainable rail transit

Classification	Benefits	Reference
Environmental	Minimal pollution, high energy efficiency.	Xue et al. [1]; Yang et al. [2]
Operational	High carrying capacity reduces urban traffic time.	Systematic evaluation [3]
Safety & Health	Accident rates are far lower than those on the road.	Kalhoro et al. [4]
Economic	Figuring out the significance of real estate	"High-Speed Rail Systems" [5]
Social	Better accessibility and social inclusion for non-drivers.	Liu and Nie [6]

Kalhoro et al. [4] asserted that public transportation systems confer indirect health advantages by enhancing individual physical activity levels and markedly decreasing accident and injury rates. In this perspective, train systems are both environmentally sustainable and a method of transit that encourages healthy lifestyles.

From an economic and spatial standpoint, rail networks seem to generate favorable economic externalization for urban development. Numerous studies indicate that real estate values rise in proximity to rail lines, enhance commercial activity, and broaden employment prospects. A thorough analysis entitled "High-Speed Rail Systems Worldwide" [5] highlighted that investments in high-speed rail and metro systems foster long-term regional economic development, facilitating urban expansion and spatial equilibrium. These findings illustrate the strategic significance of rail lines, serving both as a transportation medium and as an instrument for urban planning and economic advancement. From a societal standpoint, rail transportation is crucial in enhancing accessibility and promoting equity in transit. Liu and Nie [6] asserted that urban rail transit enhances access to economic possibilities for persons lacking private vehicles and promotes social inclusion. In this context, rail systems are regarded as instruments that advance transportation equity and mitigate socioeconomic disparities within the urban environment.

Rail systems, especially in metropolitan areas, have emerged as a crucial component of contemporary transportation networks due to their substantial passenger capacity, little emissions, and support for sustainable transportation objectives. The efficacy of these systems is intricately connected not just to technical infrastructure and energy efficiency but also to timely operational performance, namely punctuality rates. Recent studies indicate that delays elevate operational expenses and adversely affect passenger happiness and perceptions of system reliability [7-8]. The escalation of line density, reduced headway, insufficient maintenance planning, and external influences have rendered delays a multifaceted engineering and operational challenge. This has initiated a comprehensive analysis of the causes of delays in the literature and the formulation of solution-oriented models.

Rail system delays are often classified into three fundamental groups. Primary delays arise when trains diverge from their designated departure or arrival schedules. Cascading delays arise when the delay of one train affects subsequent trains owing to signaling and capacity limitations. Disruption delays, although infrequent, have a substantial impact and are triggered by exceptional occurrences like as signal failures, accidents, or emergencies. The data from UIC (Union Internationale des Chemins de fer) and the European Railway Agency (ERA) advocate for the utilization of On-Time Performance (OTP), delay minutes per train-kilometer, and headway regularity metrics in performance assessments [9]. As rail systems provide considerable benefits, their effectiveness is compromised by multiple types of delays, classified by their source and systemic influence in Table 2.

Literature indicates that the causes of delays possess a multifaceted nature. Operational and scheduling factors are critical, especially concerning insufficient buffer times and cumulative delays at regular turning places [10]. Research indicates that even minor main delays can escalate across the system when lines near capacity limitations [11]. The limitations of typical forecasting models in capturing complex spatial-temporal interactions mean that estimating the total cumulative delay effect at a station over a specified time period is more operational and beneficial for emergency planning than estimating the precise delay of a single train [12]. Additionally, meteorological conditions, variations in passenger demand, and security issues are regarded as exogenous factors influencing delay performance [13].

Research examining delay behavior reveals two primary methodological trends. Event-based models replicate delay propagation throughout an event graph, considering train block and station transition duration. These models analyze the system's response in the time domain by evaluating the relationships among signal blocks and train sequencing constraints [14].

Table 2. Categorization of delays in rail systems

Delay Category	Definition and Cause	Systemic Impact
Primary Delays	When trains miss their departure or arrival times	Change in the schedule
Cascading Delays	When signaling and capacity limits cause one train's delay to affect others	Spread around the network
Disruption Delays	Set off by signal failures, accidents	Impact is not often but significant

Data-driven methodologies have gained prominence with the advancement of machine learning techniques, especially post-2020. Long Short-Term Memory (LSTM) models have demonstrated great accuracy in short-term predictions, but Graph Neural Network (GNN) and Graph Attention Network (GAT) models have effectively predicted delay propagation at the network level by analyzing interactions across trains [15]. These methods offer a notable advantage by enabling the consideration of delay as both a time series and a spatial, network-dependent phenomenon. Moreover, the application of Explainable AI (XAI) methodologies has enhanced the comprehension of model outputs by decision-makers [16].

A notable technique that has garnered interest in recent literature is the concept of robust scheduling. This methodology seeks to create schedules that are adaptable and resilient to delays, thereby minimizing the effects of minor disruptions on system performance. Cacchiani and colleagues [17] shown that both capacity efficiency and punctuality can be enhanced through the intentional allocation of time buffers to crucial nodes and turning locations, rather than through random distribution. The UIC 406 approach additionally recommends examining the correlation between line capacity consumption rate and timeliness performance concurrently [9].

Disruption management is notably significant in the literature on delays. Studies indicate a bifurcated approach for proficiently addressing disturbances. Initially, preventative planning must be executed, ensuring the availability of spare vehicles and workers at essential locations. During the second phase, prompt rescheduling must be executed upon the occurrence of an incident, and, if required, bus bridging should be enacted [18]. These strategies enhance system resilience and mitigate adverse effects on passengers.

Recent study indicates that performance measurement ought to prioritize passengers rather than exclusively concentrating on trains. Beyond the standard OTP statistic, passenger-weighted delay (minutes) and perceived delay (minutes) indicators are indicated as more accurate metrics representing customer experience [19]. These indicators illustrate that two delays of identical time might provide varying effects on passengers; thus, planning must account for these discrepancies.

Recent investigations indicate three significant results. The predominant source of delays is propagation, which can be mitigated by effective scheduling and capacity management. Secondly, AI-driven forecasting models yield highly precise outcomes in the short term; however, enhancements are needed for explainable and data integrity. Third, robust scheduling and passenger-eccentric performance measures offer considerable potential for enhancing both operational efficiency and user happiness. Consequently, minimizing delays in rail systems transcends a technological challenge; it necessitates a multidisciplinary approach that effectively integrates data management, human factors, and decision support systems.

Deep learning methodologies that simultaneously represent spatial and temporal interdependence have surfaced in research analyzing the temporal progression of railway track layout. Research employing hybrid CNN–LSTM architectures rectified positional inaccuracies in measurement data collected at various intervals and forecasted future degradation patterns of track geometry parameters [20]. A separate study detected unsupervised deterioration areas in track geometry data by the application of auto encoders and clustering techniques [21]. Additionally, an analysis of long-term track measurement data indicated that statistical models like ARIMA, Weibull, and Gamma distributions are progressively being supplanted by artificial neural networks and hybrid systems [22].

Recent research on turnout systems have demonstrated good accuracy rates in failure prediction by the analysis of current and vibration signals derived from field data [23]. Classification studies utilizing tree-based algorithms have demonstrated superior performance in feature selection and model optimization relative to conventional methods [24]. Semi-supervised methodologies aimed at mitigating the scarcity of labeled data have enabled the utilization of unlabeled data through matrix-format current signals [25]. In a separate investigation employing a mix of wavelet transform and CNN-LSTM architecture, the accuracy rate for fault diagnosis surpassed 95% [26].

The application of statistical modeling in assessing dependability and maintenance methods inside rail systems is significant. Models designed for predicting remaining service life have estimated the likelihood of component failure by autonomously extracting features from time series data [27]. Analyses utilizing actual field data modeled the time between failures of turnouts and connectors with the Weibull distribution, revealing the impact of environmental conditions on failure likelihood [28]. In reliability studies including environmental variables, novel maintenance planning methodologies have been introduced by incorporating Weibull-based assessments into the life cycle cost model [29]. Hybrid models integrating statistical and machine learning techniques have been established for the prediction of ballast settlement and track deformations in long-term performance evaluations [30].

Based on another study [31], analyzes delays occurring in high-speed rail networks using statistical methods and makes recommendations for improving temporal planning. Another study [32] evaluates delays in arrivals and departures to find the optimal probability distributions for train delays. The researchers discovered that the models exhibited heavy-tailed characteristics, specifically Lognormal or Weibull distributions. In recent years, there has been a rise in the use of statistical analytic methods in rail systems that people want system to be reliable, easy to maintain, and always running [33-36]. Studies published after 2020 have demonstrated the increasing application of statistical and machine learning techniques in forecasting track geometry degradation, evaluating faults in switch systems, and determining equipment reliability.

The novelty of this study lies in its rigorous methodological approach to transit reliability. While previous research has identified various sources of delays, this work is among the first to apply Welch's ANOVA and Effect Size quantification to monthly rail transit logs. By accounting for heteroscedasticity and preserving high-impact outliers, this study provides a more accurate and robust estimation of operational risk than traditional parametric analyses. The resulting significance hierarchy offers a data-driven framework for transit authorities to prioritize maintenance and investment strategies based on the quantified magnitude of each disruption source.

2. Research Methodology

Research includes use of a methodology, designed to investigate service delays that occur within the Ankara urban rail network. The methodology is intended to facilitate the transfer from raw data that have been statistically confirmed with relation to the causes of delays.

The impact of service delays was tested across five different categories using a comprehensive statistical procedure. To verify the validity of the compared outcomes, the data were analyzed firstly, by diagnostic testing for the major assumptions of ANOVA. The Shapiro–Wilk test was applied to control normality, that is suitable for the sample size in study (N=12 each group). Also, Levene's test was used for evaluate the homogeneity of variances. Due to the presence of heteroscedasticity in the data (Levene's $p < 0.001$), a standard One-Way ANOVA was considered inaccurate. Thus, Welch's ANOVA was applied to figure out if the variations in delay were statistically significant. The Games-Howell post-hoc test was used for comparing pairs because to its robustness against unequal variances and non-normality, resulting im significant results. The magnitude of these differences was measured using η^2 and ω^2 impact sizes. All statistical analyses were conducted at a 95% confidence level ($\alpha = 0.05$). The study analyzed monthly aggregated data to correlate with the transportation authority's performance reporting cycles and to identify seasonal trends while mitigating daily random variations. During the data cleansing phase, category names were standardized to guarantee longitudinal consistency. An examination of the raw logs verified that the dataset was entirely complete, with no absent values over the research duration.

Throughout the year 2021, this study conducted an analysis of service delays that occurred on the Ankara M1 (Batıkent–Kızılay), M2 (Kızılay-Çayyolu), M3 (Batıkent-Sincan), M4 (Keçiören-Kızılay) metro lines. The information was collected from the records of the EGO General Directorate. Total of 2,433 delays were obtained from the official records of the EGO General Directorate. These events were classified into five sources to function as the independent variables for statistical analysis. Following the cleansing and normalization of the data, the delays were categorized according to their causes.

The M1 line spans from Batıkent to Kızılay and comprises 12 stops. The line connects with the M2, M3, and M4 lines in Kızılay (Fig. 1). The M2 line commences in the Çayyolu neighborhood and extends to the Kızılay center. The route comprises stations at Bilkent, METU, and the National Library (Fig. 2). The M3 line extends from the Batıkent station of the M1 line to Sincan, linking the western residential districts to the city center (Fig. 3). The M4 line commences at Keçiören and extends to Kızılay, facilitating access from the northeastern sector of the city to the central area (Fig. 4).

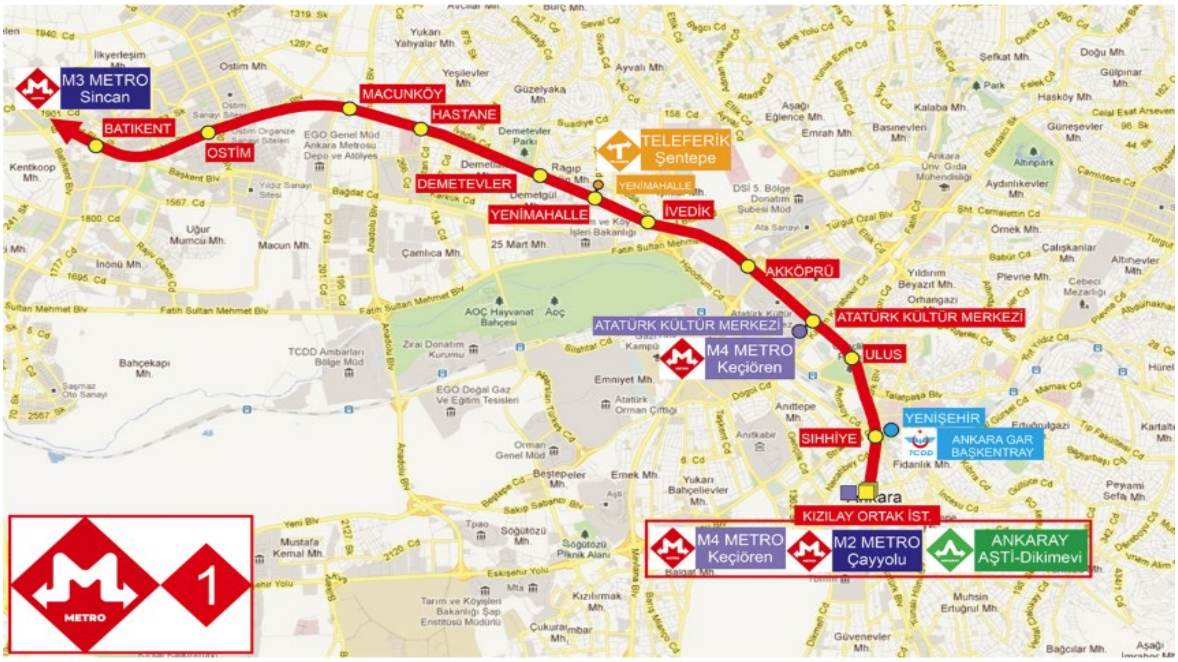

Fig. 1. M1 Ankara line rail system map

Fig. 2. M2 Ankara line rail system map

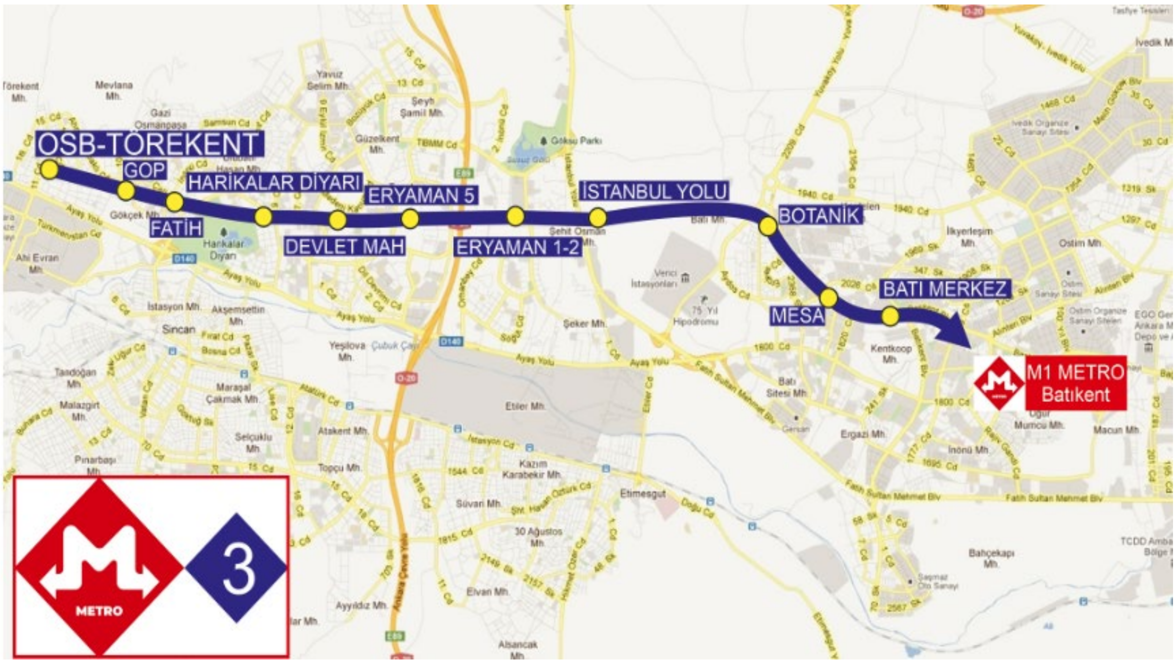

Fig. 3. M3 Ankara line rail system map

It can be clearly seen the comparison of metro lines in Fig. 5. M2 line is the longest metro line, measuring 16,590 meters while M4 line is the shortest, measuring 12,500 meters. M1 and M4 lines each comprise 12 stations, resulting in an identical quantity of stations for both routes. The M2 line possesses the greatest length, featuring an average station interval of 1.51 km.

Fig. 6 illustrates the variation in various forms of delays over time and the cumulative duration of these delays within a transportation system in 2021. The graph was utilized to analyze the delays encountered within a particular time-frame with greater specificity. The forms of delays encompass train-related delays, passenger-related delays, operational delays, equipment-related delays, and system-related delays. The graph shows each month of 2021 along the x-axis, with stacked bars depicting the delay duration for each month. Figure features a black line on the right y-axis, enabling the monitoring of the cumulative total delays for each month. This line illustrates the variation in total delay duration over time, whilst the stacked bars indicate the contribution of each delay type. This dual-axis methodology enables a comprehensive understanding of both the overall delays and the temporal distribution of each delay cause. The graph indicates that train-related delays were significantly elevated in the onset of 2021, particularly in January.

Fig. 4. M4 Ankara line rail system map

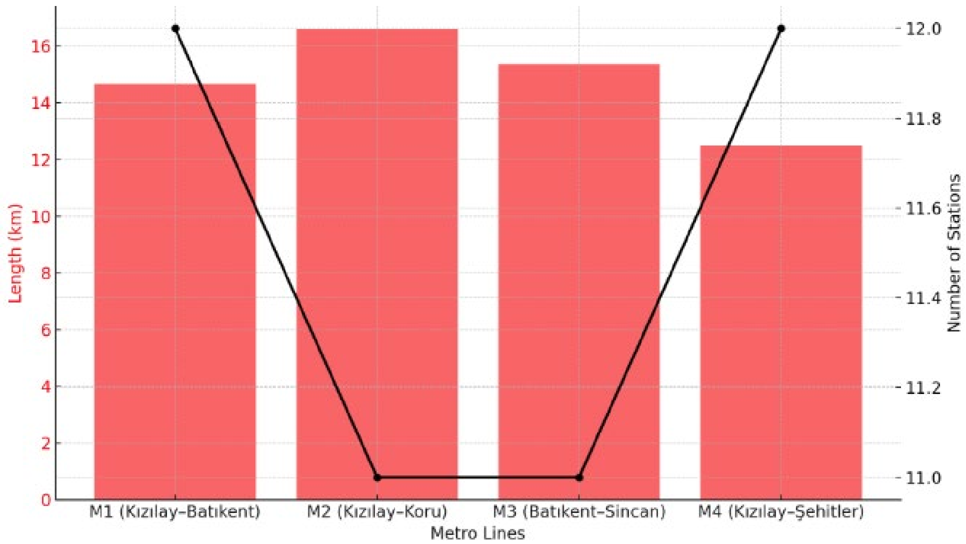

Fig. 5. Length and number of stations of Ankara metro lines

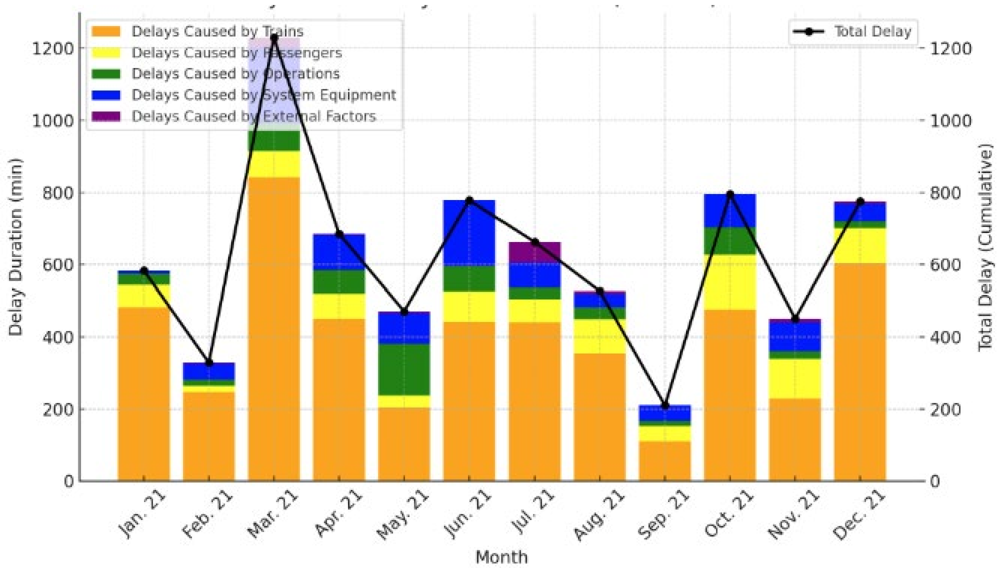

Fig. 6. Monthly service delay periods for all M1-M2-M3-M4 lines

This may pertain to operational challenges or delays induced by external events in the initial month of the year. A notable rise in overall delays is evident throughout the summer months, namely in June, July, and August. This rise may be attributable to the effects of passenger-related delays. Delays are likely to escalate during busy seasons due to an increased volume of passengers and external influences.

Total delays exhibit greater fluctuations in the final quarter of the year, specifically from October to December. A notable rise in delays due to external sources is noticed throughout this era. This indicates that external factors, including meteorological conditions or infrastructure-related problems, exert a more significant influence on delays.

Fig. 7 shows service delay duration by month for all lines. March 2021 is distinguished by the darkest tone on the graph, signifying a cumulative delay of 1230 minutes for that month. This significant delay may suggest that the transportation system was encountering substantial issues. The unusual delays in March were probably attributable to external circumstances, such as inclement weather, elevated passenger volume, or operational disturbances. Conversely, August 2021 is depicted by the lightest hue on the graph, indicating a delay of merely 210 minutes. This indicates that the transportation system encountered markedly less delays compared to other months during that time frame and generally functioned more efficiently. The reduced delay in August is presumably attributable to diminished congestion or enhancements in operations. Ultimately, September 2021 exhibits a somewhat increased delay time relative to August, totaling 527 minutes. Although this signifies a minor growth, it remains depicted in the lighter hues on the graph. The rise in September may have resulted from seasonal fluctuations or operational factors; however, the overall delay duration remained minimal in comparison to other months.

A primary observation from the Fig. 8 is that train-related delays often constitute a substantial fraction of the year. This delay, peaking in March, may indicate significant issues within the transportation infrastructure. Moreover, months characterized by significant delays due to passengers and external influences signify elements that can affect the effectiveness of the transportation system, especially during peak times. This graph elucidates delay trends over time and identifies the most significant elements affecting the transportation system. It also furnishes critical information for operational enhancements and peak-season formulating strategies. Identifying the most common types of delays by month offers information into the areas requiring enhancement within transportation systems.

Fig. 9 presents the causes of service delays for the M1, M2, M3, and M4 metro lines. March 2021 stands out as the month with the most amount of delays for the year. This month, there were 292 incidents. This signifies that disturbances in the transportation system have reached their zenith, resulting in a substantial issue inside the system. The elevated figure in March may result from external influences, congestion, or operational disturbances. Conversely, May 2021 shows the fewest delays on the chart. This month, only 103 events were documented. Minimal delays signify that the transportation system is functioning efficiently or seeing less congestion. June and July 2021 rank as the second and third biggest months for delays, with 186 and 182 delays, respectively, following March. Both months encountered substantial delays, indicating that the transportation system was impacted by increased congestion. The minimum number on the chart is May 2021, with 103 instances recorded. This month signifies the interval during which the transportation system saw minimal disturbance.

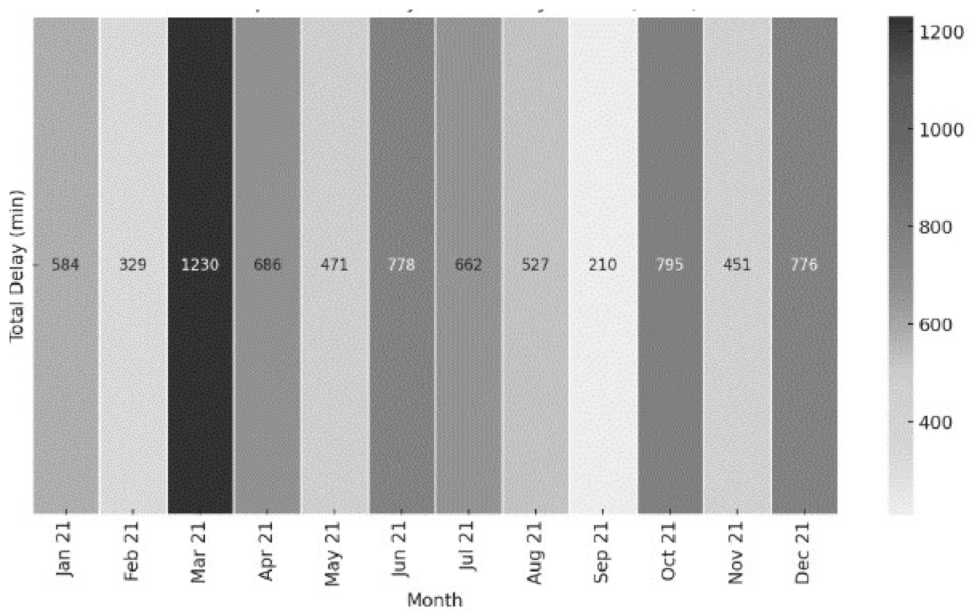

Fig. 7. Service delay duration by month for all M1-M2-M3-M4 lines

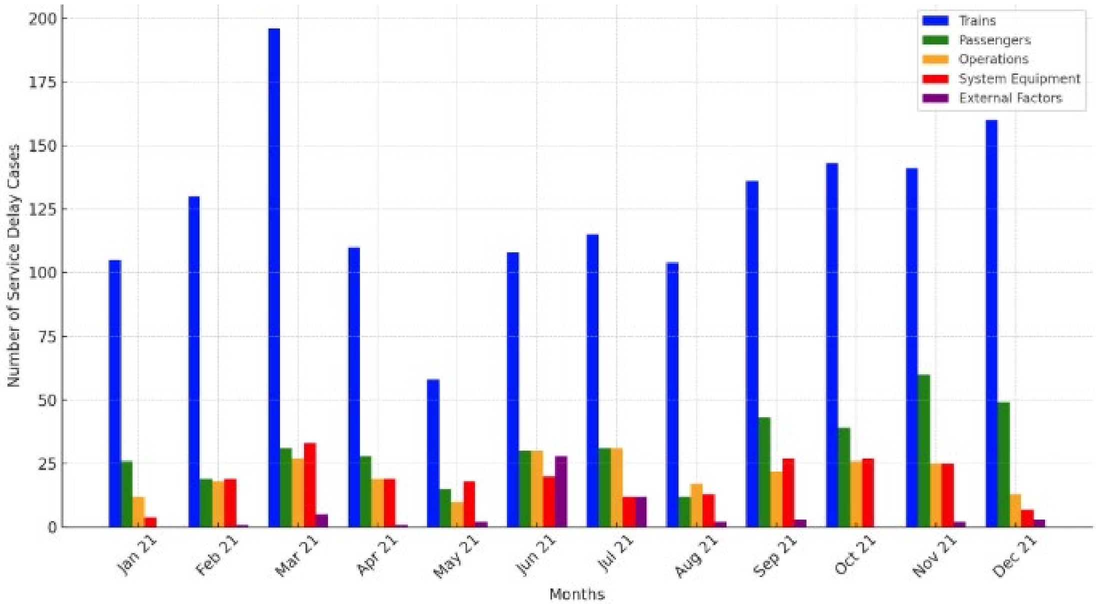

Fig. 8. Number of service delay cases for all M1-M2-M3-M4 lines

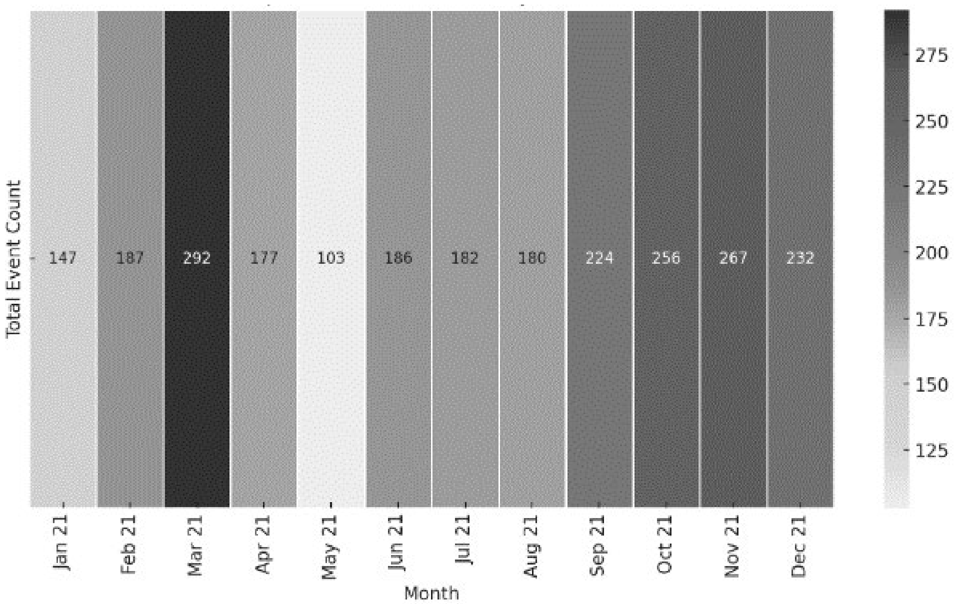

Fig. 9. Service delay event by month for all M1-M2-M3-M4 lines

3. Findings and Discussion

For the goal of this study, a statistical analysis was performed to investigate the factors that led to delays on the M1, M2, M3, and M4 metro lines in Ankara. This section of the study contains a comprehensive discussion of the findings that were obtained from the statistical analysis.

Table 3 shows an outline of the monthly delay frequencies and the results of the Shapiro–Wilk normality test. The mean values show that the impact is very different for various kinds of delays. Train-related delays (M = 125.50) are the most common, whereas External Sources (M = 2.42) are the least. The Shapiro–Wilk test showed that four of the five groups have a normal distribution ($p > 0.05$).

Table 3. Descriptive statistics and normality assessment

Delay Category	N	Mean	Std. Dev.	Shapiro-Wilk (W)	p-value (Normality)
Train-related	12	125.50	34.29	0.956	0.723
Passenger-related	12	37.08	17.09	0.926	0.339
Operations-related	12	18.42	5.84	0.938	0.468
System Equipment	12	19.33	9.02	0.956	0.729
External Sources	12	2.42	3.40	0.716	0.001*

The "External Sources" category, on the other hand, was very different from normal ($W = 0.716$, $p = 0.001$), being so many months with no delay. This first assessment led to the choice of strong statistical approaches in the next step.

To figure out the fit of a standard ANOVA, Levene's test was used to assess the homogeneity of variances. The test produced a significant outcome ($F(4, 55) = 7.482$, $p < 0.001$), demonstrating that the variances among the five delay groups are unequal, as illustrated in Table 4. The study deviated from the conventional Fisher's F-test due to the violation of the basic theory of homoscedasticity, choosing for a more robust analytical method. Since $p < 0.05$, the assumption of equal variances is rejected, necessitating the use of Welch's ANOVA.

Table 4. Test of homogeneity of variances

Test	Statistic (F)	df1	df2	p-value
Levene's Test	7.482	4	55	< 0.001*

Welch's ANOVA was conducted as a methodological adjustment for unequal variances. Table 5 demonstrates that variations among delay groups are statistically significant ($F(4, 25.25) = 60.90$, $p < 0.001$). Moreover, η^2=0.868 and ω^2 = 0.856 indicate an exceptionally large impact size. This, indicates that the source of delay is not only a significant factor but the predominant one, accounting for almost 86% of the entire variance in system delay frequency.

Table 5. Robust test of equality of means and effect sizes

Statistical Metric	Value	Result
Welch's F-statistic	60.90	p<0.001 (Significant)
Degrees of Freedom (df1,df2)	4, 25.25	—
η^2	0.868	Large Effect
ω^2	0.856	Large Effect

The outcomes of the comparisons in pairs, performed using the Games-Howell post-hoc approach, are displayed in the significance matrix (Table 6). The research shows that almost all delay groups are statistically different from each other ($p < 0.05$). Train-related delays showed the highest frequency and were determined distinct from all other categories ($p < 0.001$), proving as the principal problem in the transport system. Also, External Sources were markedly lower than all other categories ($p < 0.001$), indicating the least common reason of service outages. The one comparison that did not achieve statistical significance was between Operations-related delays and System Equipment delays ($p = 1.000$). This suggests that although these two causes are statistically more significant than external factors and less significant than passenger-related difficulties, they contribute to system delays at a nearly same rate, creating a common level of operational impact.

4. Conclusions

The purpose of this study was to investigate the influence that various sources of delay had on the transportation system during the year 2021 and to compare the number of delay events that were experienced by each source. The study focused the principal causes of service delays using a comprehensive statistical framework, guaranteeing that all foundational data assumptions were thoroughly evaluated. Analysis employing Shapiro–Wilk and Levene's tests revealed heteroscedasticity among groups which causes use of Welch's ANOVA. The results showed that there was a huge difference in the number of delays between the five groups that were looked at ($p < 0.001$). The most important result is that delays connected to trains were statistically shown to be the most common cause of system equipment. Conversely, external sources had the least effect. The computation of a substantial effect size ($\eta^2 = 0.868$) offers compelling evidence that the classification of delays used in this study accounts for approximately 87\% of the variance in performance, which means service reliability is not greatly affected by random noise from other sources, but rather by technical and operational issues. According to the results, authorities may focus on investing in train-related technical infrastructure and maintenance, as these sectors have the most potential to cut down on system delays and improve service quality.

The average delays that are caused by trains are much higher than those that occur in the other groups, resulting in a statistically significant difference at a significance level of $p<0.05$. It may be deduced from this that the most significant issue facing the transportation system is the delays that are caused by trains. The key finding is that train-related delays (N=1,506) form the main challenge that the system shows, exceeds all other types of delays.

Table 6. Significance matrix for pairwise comparisons (Games-Howell)

Category	Train	Passenger	Operations	System Eq.	External
Train-related	—	<0.001	<0.001	<0.001	<0.001
Passenger-related	<0.001	—	0.009	0.042	<0.001
Operations-related	<0.001	0.009	—	1.000(ns)	<0.001
System Equipment	<0.001	0.042	1.000(ns)	—	<0.001
External Sources	<0.001	<0.001	<0.001	<0.001	—

Delays releated to passengers (445 instances) were recognized as the second most determinable cause of disruption. Although these delays clearly differ from those associated with trains, they are not statistically different from operational or system equipment categories, indicates that the situation represents a re-curing difficulty. This means that passenger behavior is an important element, but it is not currently an important outlier in system performance compared to normal technical or operational concerns.

The study highlighted March 2021 as a significant temporal anomaly, as a peak of 292 distinct occurrences and 1,230 total delay minutes. The result indicates a "compound failure" that is a seasonal demand and operational pressures exceeds the system's recovery mechanisms.

In contrast, August 2021 was identified as the most efficient period, with only 210 minutes of delay. This corresponds with the minimal event count in May (103 occurrences), illustrating that reduced summer congestion enhances "buffer times," so keeping modest primary delays from increasing.

From the information, it can be seen that delays are caused by trains have the most substantial influence on the transportation system, surpassing all other types of delays by a significant margin. The delays that were caused by passengers and those that were not related to the system were responsible for considerable variations during particular time periods. On the other hand, the delays that were caused by system equipment and operational delays stayed at lower levels all throughout the period. With the use of this study, important issue areas within the transportation system are identified, which provides valuable insights that can be used to make modifications that are necessary to increase efficiency and successfully manage resources.

Declaration

Funding

This research received no external funding.

Author Contributions

Y. Dinç: Methodology, Literature Review, Writing–Original Draft, Validation. B. Varli Bingöl: Methodology, Formal Analysis, Data Processing, Writing– Review & Editing.

Acknowledgments

Not applicable.

Data Availability Statement

No new data were created or analyzed in this study.

Ethics Committee Permission

Not applicable.

Conflict of Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Use of generative AI and AI-assisted technologies

The authors used AI tools for language polishing and have reviewed and take full responsibility for the final content.

References

[1] Xue H, Yang P, Zhang H (2017) Research on comprehensive benefits of urban rail transit system based on the joint evaluation methods. In: Proceedings of International Conference on Geo-Spatial Knowledge and Intelligence. Springer, Singapore.

[2] Chen Q, Qin Y, Yao M, Zhang Y, Zhai Z (2024) The consistency of subjective and objective factors influencing walking path choice around rail transit stations. Buildings 14(7): 2225.

[3] Abd Majid N, Mohd ZN (2025) A systematic review of transportation rail networks: Impacts on urban development and sustainability. International Journal of Academic Research in Business and Social Sciences 15(2).

[4] Kalhoro M, Yong HNA, Kalhoro S, Ramendran SC (2024) Public transportation usage and health benefits: A literature review. International Journal of Environment and Health 11(2): 103–126.

[5] Momenitabar M, Ebrahimi ZD, Arani M (2020) A systematic and analytical review of the socioeconomic and environmental impact of the deployed high-speed rail (HSR) systems on the world. https://doi.org/10.48550/arXiv.2003.04452

[6] Silva Ó, Cordera R, González-González E, Nogués S (2022) Environmental impacts of autonomous vehicles: A review of the scientific literature. Science of the Total Environment 830: 154615.

[7] Chen E, Luo Q, Chen J, He Y (2023) Understanding passenger travel choice behaviours under train delays in urban rail transits: A data-driven approach. Transportmetrica B: Transport Dynamics 11(1): 1496–1524.

[8] Aantjes JC (2022) System Reliability Analysis of Interactions Between ETCS, Train Drivers and Dispatchers, Demonstrated by STPA. Master Dissertation, Delft University of Technology.

[9] Fitzová H (2022) Has European rail policy improved the efficiency of European railways? Journal of Transport Economics and Policy 56(4): 445–463.

[10] Rählmann C, Wagener F, Thonemann UW (2021) Robust tactical crew scheduling under uncertain demand. Transportation Science 55(6): 1392–1410.

[11] Scheepmaker G, Goverde R (2021) Multi-objective railway timetabling including energy-efficient train trajectory optimization. European Journal of Transport and Infrastructure Research 21(4): 1–42.
[12] Zhang D, Peng Y, Zhang Y, Wu D, Wang H, Zhang H (2021) Train time delay prediction for high-speed train dispatching based on spatio-temporal graph convolutional network. IEEE Transactions on Intelligent Transportation Systems 23(3): 2434–2444.
[13] Chen Q, Qin Y, Yao M, Zhang Y, Zhai Z (2024) The consistency of subjective and objective factors influencing walking path choice around rail transit stations. Buildings 14(7): 2225.
[14] Corman F, D'Ariano A, Hansen IA, Pacciarelli D (2011) Optimal multi-class rescheduling of railway traffic. Journal of Rail Transport Planning and Management 1(1): 14–24.
[15] Li H, Zhao Y, Mao Z, Qin Y, Xiao Z, Feng J, Zhang M (2024) Graph neural networks in intelligent transportation systems: Advances, applications and trends. https://doi.org/10.48550/arXiv.2401.00713
[16] Chen H, Zhao K, Zhang Z, Zhang H, Lu L (2024) Evaluating transit-oriented development performance: An integrated approach using multisource big data and interpretable machine learning. Journal of Advanced Transportation 2024(1): 7450495.
[17] Lou LX (2024) Passenger-Centric Robust Timetabling in Railways: A Case Study for The Eindhoven–Den Bosch–Tilburg Network. Master Dissertation, Delft University of Technology.
[18] Wen X, Si B, Wei Y, Cui H (2025) Resilience assessment of urban rail transit systems: A literature review. Public Transport 17(3): 657–681.
[19] Monsuur F, Enoch M, Quddus M, Meek S (2021) Modelling the impact of rail delays on passenger satisfaction. Transportation Research Part A: Policy and Practice 152: 19–35.
[20] Wang X, Bai Y, Liu X (2023) Prediction of railroad track geometry change using a hybrid CNN-LSTM spatial-temporal model. Advanced Engineering Informatics 58: 102235.
[21] Eltouny K, Gomaa M, Liang X (2023) Unsupervised learning methods for data-driven vibration-based structural health monitoring: A review. Sensors 23(6): 3290.
[22] Liao Y, Han L, Wang H, Zhang H (2022) Prediction models for railway track geometry degradation using machine learning methods: A review. Sensors 22(19): 7275.
[23] Gürsoy Özcan A, Yiğit E, Toksöz NY (2020) Transfer preslerde durum izlemesi ve bir kestirimci bakım sisteminin geliştirilmesi. In: Proceedings of 7th International Scientific Research Congress (UBAK). Ankara, Türkiye. (in Turkish)
[24] Lao Z, He D, Wei Z, Shang H, Jin Z, Miao J, Ren C (2023) Intelligent fault diagnosis for rail transit switch machine based on adaptive feature selection and improved LightGBM. Engineering Failure Analysis 148: 107219.
[25] Li W, Xu Z, Mei M, Lan M, Liu C, Gao X (2024) A semi-supervised adaptive matrix machine approach for fault diagnosis in railway switch machine. Sensors 24(13): 4402.
[26] Liu H, Yang S, Liu C, Liu S, Liu R (2023) Research on switch machine fault diagnosis based on wavelet transform and CNN-LSTM. In: Proceedings of International Conference on Electrical and Information Technologies for Rail Transportation. Springer Nature, Singapore.
[27] Liu J, Du D, He J, Zhang C (2024) Prediction of remaining useful life of railway tracks based on DMGDCC-GRU hybrid model and transfer learning. IEEE Transactions on Vehicular Technology 73(6): 7561–7575.
[28] Litherland J, Andrews J (2023) A reliability study of railway switch and crossing components. Proceedings of the Institution of Mechanical Engineers Part F: Journal of Rail and Rapid Transit 237(2): 205–217.
[29] Ji K, Choi I, Won J (2025) Reliability analysis-based life cycle assessment of railway components using long-term maintenance data. Developments in the Built Environment 21: 100629.
[30] Johansson F, Ansell A, Johansson D, Funehag J, Norrman J (Eds) (2025) Tunnelling Into A Sustainable Future – Methods and Technologies, First Edition. Proceedings of the ITA-AITES World Tunnel Congress 2025 (WTC 2025), Stockholm, Sweden. CRC Press. https://doi.org/10.1201/9781003559047
[31] Huang P, Wen C, Li J, Peng Q, Li Z, Fu Z (2019) Statistical analysis of train delay and delay propagation patterns in a high-speed railway system. In: Proceedings of 5th International Conference on Transportation Information and Safety (ICTIS).
[32] Yang Y, Huang P, Peng Q, Li J, Wen C (2019) Statistical delay distribution analysis on high-speed railway trains. Journal of Modern Transportation 27(3): 188–197.
[33] Wen C, Li Z, Lessan J, Fu L, Huang P, Jiang C (2017) Statistical investigation on train primary delay based on real records: evidence from Wuhan–Guangzhou HSR. International Journal of Rail Transportation 5(3): 170–189.
[34] Landex A (2008) Methods to Estimate Railway Capacity and Passenger Delays. PhD Dissertation, Technical University of Denmark.
[35] Yağmur E (2020) Elastic modulus prediction for fiber-reinforced concretes. Pamukkale University Journal of Engineering Sciences 26(6): 1098–1109.
[36] Wen C, Mou W, Huang P, Li Z (2020) A predictive model of train delays on a railway line. Journal of Forecasting 39(3): 470–488.