Passenger Trains Delays in UK: A possible explanatory model
MRes Spatial Data Science & Visualization

Quantitative Methods final coursework

Introduction

Train delays
Passenger train delays cause remarkable losses to private operators, public companies and, especially to passengers (Marković et al. 2015). In 2006-07, 14 million train-minutes of delays, costed the passengers about £1 billion in lost time (Preston et al. 2009). Therefore, train delays are used as a parameter for organizing timetables and traffic operations. Studying how to reduce this phenomena, allow planners and political entity to evaluate how the system affect delays, preventing dysfunctional services and economical losses (Milinković et al. 2013; Marković et al. 2015).
Generally, two kinds of delays have been identified (Milinković et al. 2013; Higgins & Kozan 1998)

a. *Primary*, or *source delays*: delays that first occurs, as a consequence of external or internal disturbance.
b. *Knock-on delays*: delays caused by a primary delayed convoy (that may prevent other trains from passing it or crossing it).

Knock-on delays have been tackled with stochastic delay propagation models, for evaluating the robustness of timetables and, eventually, for optimizing them (Yuan 2006; Berger & Gebhardt 2011), while primary delays are predicted using artificial neural networks (ANN) (Yaghini et al. 2013), along with simulation models (Milinković et al. 2013). At the same time, other researches have been trying to identify the influencing factors of train delays. Olsson and Haugland (2004) have found that punctuality of trains is related with depart number of passengers, occupancy ratio and departure punctuality. Wiggenraad (2001) has evaluated alighting and boarding times procedures.

In this paper, we will examine train delays in UK territory, considering the performances of the Train Operating Companies (TOC) that are running passenger services, analysing delays and punctuality indexes globally.

UK, data and performance measures
In 1993 the British Conservative Government decided to begin a gradual process of privatization of the railways, that lead to the suspension of British Rail operations in 1997 (Poole 1997). Its services started to be managed by 25 private franchised companies. Since privatization, the organization of services and of the industry has changed several times. During 2015-16, twenty-three companies run passenger services (National Trend Portals Data), while the infrastructure is managed by Network Rail, a public entity of the Department of Transport. As a regulator, the *Office of Rail and Road* monitors the performances of TOC and of Network Rail employing two indexes of punctuality and reliability (Office of Rail and Road 2016):

- Public Performance Measure (PPM): the percentage of trains that arrived at their final destination within 5 minutes of their scheduled arrival time (within 10 minutes for Long Distance services). A higher score is better.
- Cancellations and Significant Lateness (CaSL): the percentage of trains that have been cancelled and/or arrived at their final destination late by more than 30 minutes. A lower score is better.

There are a number of issues with these indexes, since they are weighted by trains and not by passengers. Moreover, they measure delays only at the final destination of a train it does not measure lateness to the passenger's destination, but only the train's final destination (Office of Rail Regulation 2013). The *National Rail Trends Portal* make available data regarding TOCs performances since 1997 until today. We will use the PPM indexes along with minute of delays.

Research Hypothesis
The purpose of this paper is to use the indexes presented above in context and to examine possible causes of delays. What are the influencing factors of bad TOC performances? We will try to identify possible explanatory variables amongst TOCs and infrastructure factors. In the first group we will consider the age of rolling stock. As infrastructure variables, we will consider two features of the tracks: line speed limits and electrification capability. We believe that lower performances might be partially due by a combination of old rolling stock and more critical infrastructure characteristics.

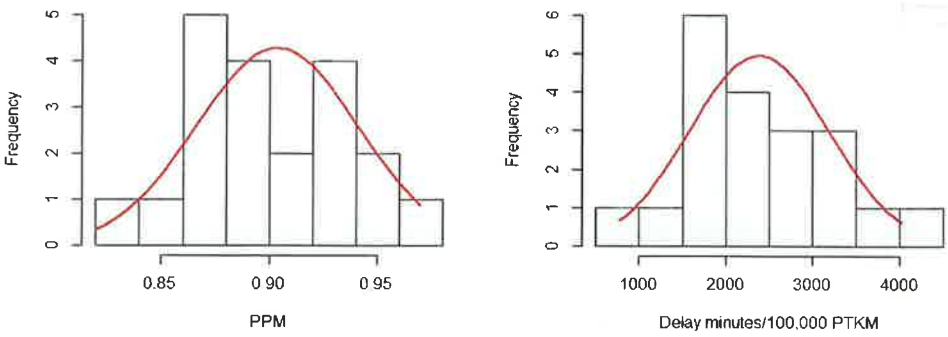

Figure 1. Histograms and distribution shapes of PPM and delay minutes/100,000 PTKM

Analysis
Our analysis will proceed as follow: we will perform a single-linear regression to check the relation between punctuality (minutes of delays) and the reliability index (PPM – mean and median: 0.903) as dependent variables with the average of the rolling stock. Afterwards, multiple regression analyses will be run to study the combined effect of several factors. Finally, we will perform a cluster analysis on 55 sub-operators.

As a preliminary step, we have estimated, for each TOC, the number of Passenger Train Kilometres (PTKM, since expressed in millions) per route and in all track conditions, starting from aggregated information regarding the route length partition. The PTKM consider not only the number of trains and the length of the travel, but their daily frequencies. In figure 1, the frequency distribution of the dependent variables is displayed. The PPM mean and median are 0.903; minute of delays every 100,000 PTKM has a mean of 2383.597 and median 2388.065.

Age of rolling stock
The first explanatory variable examined is the average age of the rolling stock. We believe that older fleets, especially if used for commuter and high-frequency services, may be related with lower performances. For examining this relationship, we have plotted PPM and minutes of delays (every 100,000 PTKM), showing their variation in relation with the age of rolling stock. Especially in the figures b), the observations show a quite random distribution.

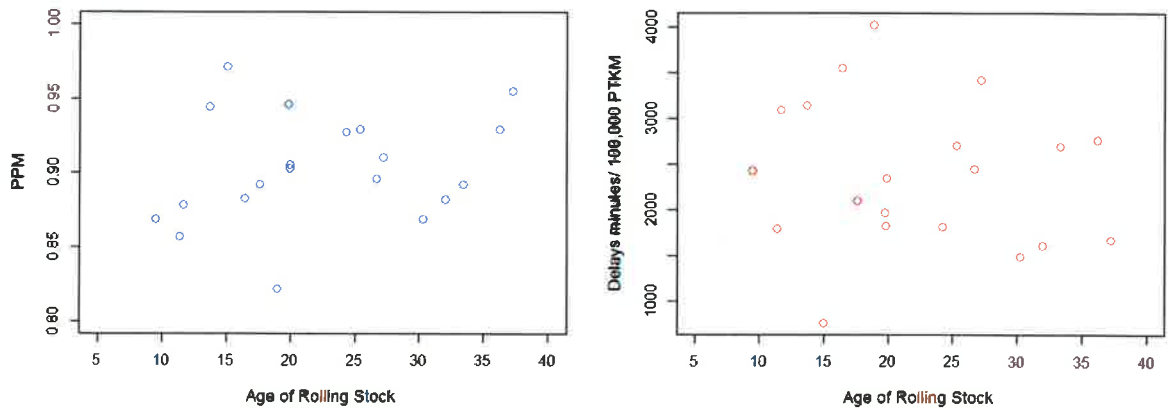

Figure 2. a) Scatter plot PPM vs age of rolling stock. b) Delay minutes vs age of rolling stock

Performing two single linear regressions, we have obtained the following results:

Residuals:	*Min*	*1Q*	*Median*	*3Q*	*Max*
	-0.077563	-0.021355	-0.002	0.022	0.0751
Coefficients:	*Estimate Std.*	*Error*	*T value*	*Pr(>|t|)*	*Sign*
intercept	0.880	0.0236490	37.241	0.00	0
slope	0.000	0.000	- 1.003	0.329	/

Table 1. Summary linear regression analysis: x=Age of rolling stock, y=PPM

PPM (y) and age of rolling stock (x) have a *Pearson's Correlation Coefficient* (R) of 0.229 and a *Coefficient of Determination* (R^2) of 0.052. The slope is not significantly different form 0 and changes in x, so in the age of rolling stock, do not predict changes in PPM. The proportion of variation that can be explained by the regression line is close to zero.

Residuals	*Min*	*1Q*	*Median*	*3Q*	*Max*
	-1706.15	-554.19	-86.23	556.06	1599.51
Coefficients	*Estimate*	*Std. Error*	*T value*	*Prob. (>|t|)*	*Significance*
Intercept	2649.16	529.05	5.007	0.00	0
Slope	-11.89	22.22	-0.535	0.599	/

Table 2. Summary linear regression analysis: x=Age of rolling stock, y=Delay minutes

Using delays minute every 100,000 PTKM as dependent variable, the results are similar: again the slope is not significantly different from 0 and R and R^2 are -0.125 and 0.0156. Operating a data transformation, as *Ln*, we have not obtained better results, as proof of the fact that, at least in this context, the age of rolling stock has no relation with punctuality and reliability of passenger trains.

Infrastructure variables

Since our first research hypothesis has been disconfirmed by the results, we will move our attention to the infrastructure wherein companies are operating. First the relationships between PPM and PTKM travelled per route is examined. In this case, the absolute values of PTKM are considered, as the PPM measure already takes into account number of passengers and planned trains. Observing performance separately amongst the routes, emerges that the variation of the amount of passengers transported in the routes London North Eastern and Sussex are negatively related (R -0.434 and R^2 0.188; 0.508 and R^2 0.258) with the variation of PPM. The variation of the minutes of delays minutes every 100,000 PTKM is moderately related with the proportion of PTKM travelled in Sussex, as for the PPM index, and more modestly with Kent (R 0.548, $R^2$0.300; R 0.389 and R^2 0.151). At this point it is difficult to clarify the direction of this relation. It might be possible that travelling in these regions causes negative performance or perhaps companies in charge of services are simply offering low quality services.

In this sense we have obtained two correlation matrixes for exploring the routes features and the relation with our variables of interest.

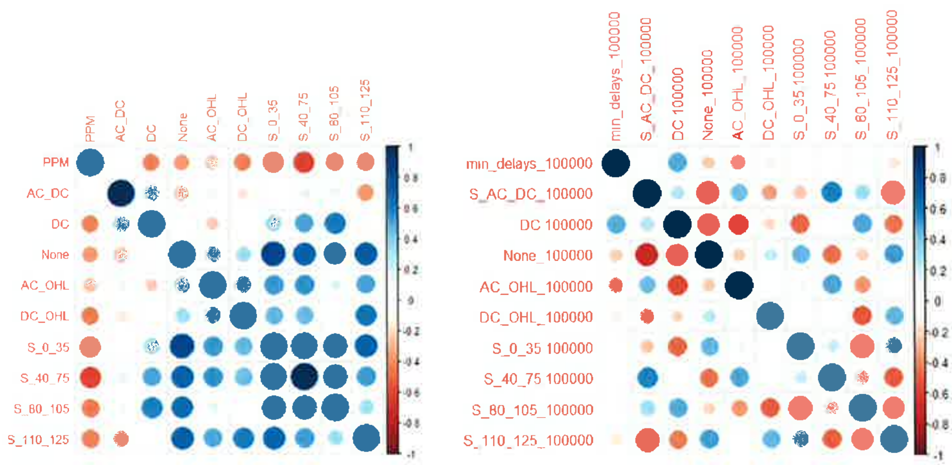

Figure 3. Correlation matrix: a) PPM vs km travelled in each track condition. b) Delay minutes vs km travelled in each track condition

Even though more PTKMs usually mean more delays, the first matrix shows that this relation is stronger in slow sectors of the track (between 0-35 mph and 40-75 mph): R -0.611 and -0.619. However, analysing the data regarding the minutes of delays, this relation disappears. The second matrix suggests that is not possible to predict from the proportion of PTKM travelled in each sector the variation of punctuality.

Since the first figure reveals an interesting relation, also considered that only 11% of the PTKM are travelled on 0-35 sectors, a simple multiple regression has been performed between PPM, as response variable, and amount of PKTM in 0-35 and 40-75 sections of the track, as predictor variables. The variables considered respect the assumption of normality, considering the limits of the *adjusted Fisher-Pearson coefficient of skewness* (G_1) for a 25 sample (±0.726). No outliers have been identified using the *IQR ±1.5* as limits. Table 2 shows the regression coefficients and intercept.

Whereas R for regression is significantly different from zero ($p< 0.05$), the regression coefficients do not contribute to the model. In addition, the partial correlation shows how much the value of R^2 would be reduced omitting one of the two variables. In both cases the amount is not particularly relevant.

Residuals	*Min*	*1Q*	*Median*	*3Q*	*Max*
	-0.056	-0024	0.002	0.027	0.040
Coefficients	*Estimate*	*Std. Error*	*T value*	*Prob. (>|t|)*	*Significance*
Intercept	0.040	0.0134	70.062	<0.000	0
S_0_35	0.000	0.000	-0338	0.739	
S_40_75	0.000	0.000	-0.607	0.552	/

	Correlations			
	Zero-order	*Partial*	*Part*	*Multiple R-squared: 0.387* *Adjusted R-squared: 0.315*
S_0_35	-0,612	-0,082	-0,064	*F-statistic: 5.379*
S_40_75	-0,619	-0,146	-0,115	*p-value: 0.015*

Table 2. Summary multi-linear regression analysis: x_1= S_0_35 x_2= S_40_75, y=PPM

Thus, trying to predict the variation of reliability based on the number of PTKM of route composed by slow sections (from 0 to 75 mph) we obtained a poor model. The two dependent variables, as discussed above, are more explicative if taken singularly, also due the insignificance of the coefficients in the multi regression model.

Company performances

Since it has been difficult to obtain a comprehensive model of the influencing factors, we will perform a cluster analysis to identify homogeneous groups of operators, depending on how well they perform. Two measures have been used: the already presented PPM, and the Right Time Measure that expects train to arrive at the final destination within 1 minute of the scheduled arrival time (Office of Rail Regulation 2013). A more detailed dataset, where punctuality is measured for each sub-operator (54 observations, see figure 4), will be analysed with the K-means clustering method, a non-hierarchical approach that selects K initial centroids randomly (MacQueen 1967).

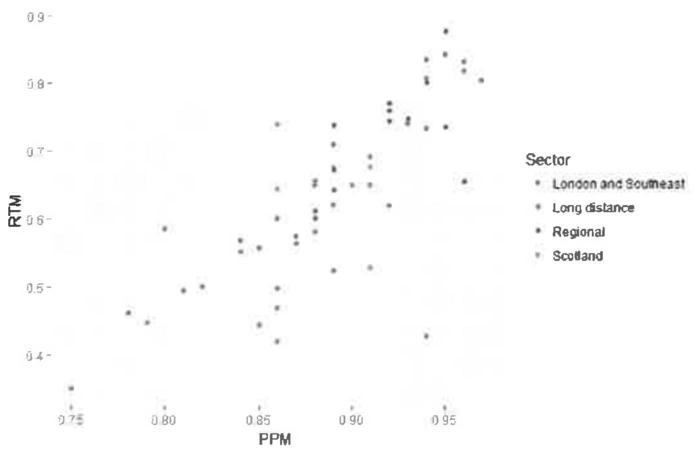

Figure 4. Sub-operator performances, PPM vs RTM, grouped by services.

The used algorithm will cluster the observations into k groups in a way that the sum of squares of the observations to their cluster centres is a minimum. To identify the ideal number of group we have computed the graphical *Dindex* and plotted the histogram (figure 5). The optimal number of groups is 3 (figure 5).

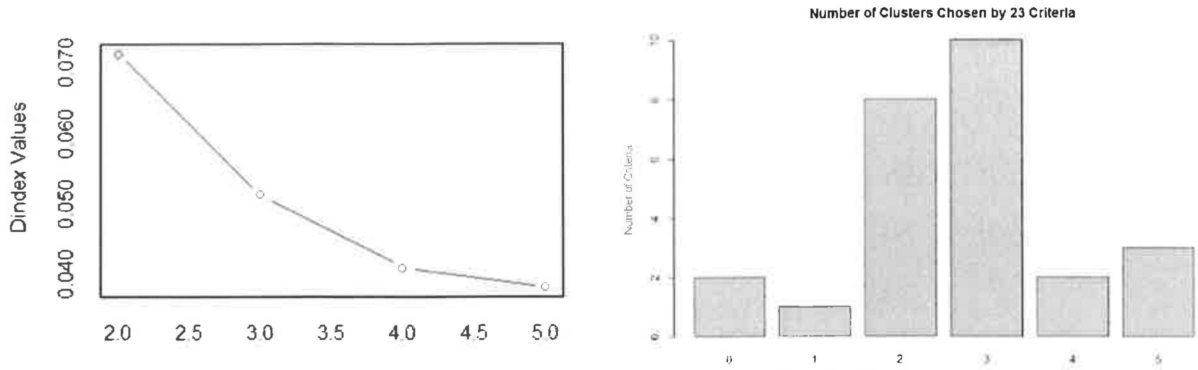

Figure 5. Dindex graph and histogram for determining optimal number of groups

A graph representing the structure obtained by the cluster analysis is showed in figure 6 along with the silhouette plot. The average silhouette width (ASW) is equal to 0.48, indicating that the groups may be slightly artificial and with a weak structure; in addition, three operators do not seem to belong to the cluster 3, that represent the low performance companies. The average distance between the groups are 0.053, 0.065 and 0.100. Even though, the analysis has been unable to return a strong structure, for our interests it could be used as an illustrative tool, as the cluster means are: PPM 0.933, 0.892 and 0.842; RTM 0.783, 0.643 and 0.494.

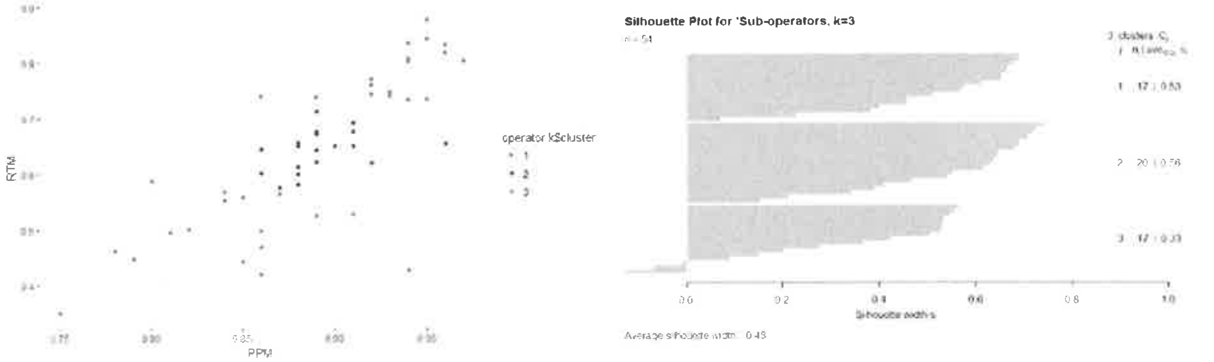

Figure 6. Cluster analysis structure and silhouette plot

This division, a part some peculiar observations, shows that several long distance operators belong to the third group, three Scotland services, out of four are in the second group and no regional services are in the third group: most of them are offering reliable services to passengers. Finally, amongst the top ten companies, six are operating in London and southeast services.

Conclusion and future directions

We have analysed several factors that according to hour hypothesis may be worth to consider for explaining delays and train performances in UK. In particular, we expected the average age of rolling stock to cause bad performances, at least in commuter and regional service. The absence of relation emerged might be due to other factors as good maintenance and reliability of the stock employed. However, it would be interesting to conduct the same analysis on a more detailed dataset.

Similarly, exploring the track features, the results have been partially controversial: the electrification capability of the route does not have particular effects on the performance, there is no difference in this sense between non-electrified tracks and electrified ones. It has emerged that, if taken singularly, the slowest sections of the track (0-35 and 40-70 mph sectors), when travelled, may tend to cause lower performances. This output is difficult to contextualize, especially considering that we have not found the same pattern looking at the proportions of the PTKM travelled on these part of the routes. Even though we can hypothesize that slow sectors partially influence train connections, it is difficult to translate this result in an operational recommendation, since usually speed reduction is due to infrastructure works.

Finally, the cluster analysis has divided in three groups 54 sub-operators, a weak structure that, nevertheless, could give partial indications about the type of services more affected by unreliability. In this sense, and for the reasons cited above, it is clear that collecting data at a more detailed level, especially concerning the routes and the partition of the services amongst the sub-operator, might permit to conduct more specific and relevant analysis, overcoming the limits of this piece of work.

References

Berger, A. & Gebhardt, A., 2011. Stochastic delay prediction in large train networks. In A. Caprara & S. Kontogiannis, eds. *11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems.* Dagstuhl Publishing, pp. 100–111.

Higgins, A. & Kozan, E., 1998. Modeling Train Delays in Urban Networks. *Transportation Science*, 32(4), pp.346–357.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In *Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California, pp. 281–297.

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

Milinković, S., Marković, M., Vesković, S., Ivić, M. & Pavlović, N., 2013. A fuzzy Petri net model to estimate train delays. *Simulation Modelling Practice and Theory*, 33, pp.144–157.

Office of Rail and Road, 2016. *Passenger and Freight Rail Performance 2015-16 Q4 Statistical Release*, Report.

Office of Rail Regulation, 2013. *Presenting Rigft-Time Performance Infromation to Rail Passengers*, Report.

Olsson, N.O.E. & Haugland, H., 2004. Influencing factors on train punctuality - Results from some Norwegian studies. *Transport Policy*, 11(4), pp.387–397.

Poole, F., 1997. *The Privatised Railway*, Report.

Preston, J., Wall, G., Batley, R., Ibáñez, J.N. & Shires, J., 2009. Impact of Delays on Passenger Train Services. *Transportation Research Record: Journal of the Transportation Research Board*, (2117), pp.14–23.

Wiggenraad, P.B.L., 2001. *Alighting and boarding times of passengers at Dutch railway stations*, Report.

Yaghini, M., Khoshraftar, M.M. & Seyedabadi, M., 2013. Railway passenger train delay prediction via neural network model. *Journal of Advanced Transportation*, 47(3), pp.355–368.

Yuan, J., 2006. *Stochastic Modelling of Train Delays and Delay Propagation in Station,* Phd Thesis, Netherlands TRAIL Research School