**Proceedings of the 2013 Joint Rail Conference**
**JRC2013**
**April 15-18, 2013, Knoxville, Tennessee, USA**

**JRC2013-2446**

# ANALYZING THE PROGRESSION FROM SINGLE TO DOUBLE TRACK NETWORKS

**Samuel L. Sogin**
**University of Illinois**
**Urbana, IL, USA**

**C. Tyler Dick PE**
**University of Illinois**
**Urbana, IL, USA**

**Yung-Cheng Lai. PhD**
**National Taiwan University**
**Taipei, Taiwan**

**Christopher P.L. Barkan. PhD**
**University of Illinois**
**Urbana, IL, USA**

## ABSTRACT

Long term demand for rail transportation in North America is projected to increase considerably in the coming decades. A significant portion of the routes in the United States are single track with passing sidings. Eventually, the second mainline track will become necessary to maintain network fluidity. However, the full funding for the second track may not be available all at once; subsequently the track can be phased in over time creating a hybrid track configuration. Depending upon the traffic characteristics, traffic will transition from a delay characteristic of single track to a delay characteristic of double track. A response surface model was developed that tested various factors including the amount of second main track added, traffic volume, traffic composition, and the speed differential between train types. Design of experiments software (JMP) was used in conjunction with railway simulation software (Rail Traffic Controller) to conduct the analysis. The benefit of full double track can be realized for high priority trains with partial double-track. However, the low priority traffic may not experience double-track-like performance until nearly the entire second mainline track is installed. The results suggest a linear relationship between miles of second mainline track added and reduction in train delay. The maximum speed of the freight train has a great impact on train delays in a congested network. These results further the understandings of key mainline interactions between passenger and freight trains. In addition the models presented will facilitate the development of an optimal incremental upgrade model for capacity expansion. Also, the methodology presented can be adopted to analyze the progression from double to triple track.

## INTRODUCTION

Most of the railway traffic in the United States operates on routes that feature single track with passing sidings. Roughly 37% of mainlines with 10 million gross tons or more are multiple-mainline-track territory [1]. With simultaneous increasing demands for freight and passenger rail service, many of these single-track lines may need to be upgraded. Some intermediate upgrade solutions include extending sidings to accommodate longer trains as well as adding additional sidings near the midpoints of long single-track sections. After these intermediate solutions are implemented, double track may become necessary in order to handle the additional traffic. This second mainline track can be phased in over time to improve capacity with the amount of double track installed matched to the expected increases in rail traffic. These intermediate phases may have characteristics of both single and double track operations and will be referred to as hybrid track configurations. The subsequent analysis will focus on the capacity benefits of a single-track route as it transitions into a full Two-Main Track route.

There are many factors that may determine how trains perform over a railway network. Railway simulation software continues to grow in sophistication in order to better emulate operations. The purpose of this study is to focus on a subset of these factors to determine key fundamental relationships that can further the understanding of railway performance. The results presented in this paper are not intended to represent absolute predictive measurements for a particular set of conditions. Rather, they are meant to illustrate comparative effects under different conditions.

### Background

To investigate the properties of hybrid track configurations, design of experiments (DOE) software is used to create a subset of hybrid track scenarios. Rail Traffic Controller (RTC) simulates these scenarios and multivariate regression techniques are used to quantify main and interaction effects.

Previous research has shown that double and single track railway lines behave differently. Single-track railways operate at a significantly lower capacity than double-track railway lines. The primary reason for this reduction in capacity is due to trains from opposite directions having to alternate use of single-track-bottleneck sections. On double track mainlines, the theoretical capacity is likely related to the following distance between trains moving in the same direction. The theoretical capacity of double track is decreased if there are speed differentials, overtakes, or traffic traveling against the current on the track used to move traffic in the opposite direction [2–4].

Previous studies on single track determined that adding a high priority train to a freight network will increase average train delays more so than simply adding another freight train [5], [6]. The speed of the high priority train showed very small correlations between train delay and maximum train speed. The delay distributions are often characterized by being skewed to the right with none of the trains performing close to the minimum run time. Meets at sidings are cited as a primary delay mechanism [7]. Double track configurations are very sensitive to speed differentials. A faster-high-priority train may need to use the second track in the opposing direction to overtake a slower train. Double track configurations have delay distributions similar to exponential distributions with many trains operating close to the minimum run time [8]. Because single and double track configurations have different delay response models, there must be a potential transition function to describe how a single-track line may transition into a double-track line. Five hypothetical transition functions are shown in Figure 1. The shape of these curves may be different for different performance metrics.

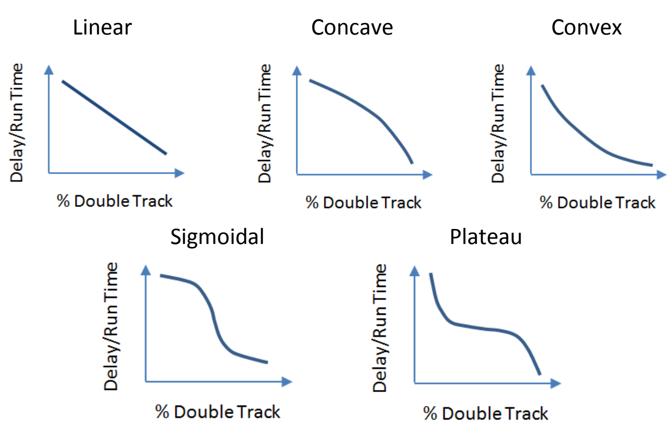

Figure 1: Possible double track transition functions

Regression modeling of train delays has been used in the past to quantify effects on train delay for various operational factors. Prokopy and Rubin used single track simulation results to develop a multivariate regression model [9]. Kruger used a similar approach using an updated simulation model and also summarized the data through multivariate regression [10]. Both models were developed by only varying one parameter at a time. Mitra et al. developed an 8-variable regression model for single track derived from simulation results. This particular model did not consider interaction effects between variables [11]. Lai and Huang used regression and neural networks to model Rail Traffic Controller (RTC) simulation results from both a single and double track network [12]. For both the single and double track models, Lai and Huang used a full-factorial experiment design analyzing five factors at three different levels.

## FACTOR SELECTION

Six factors that may have a large influence on train delays in a hybrid configuration are identified in Table 1. The different permutations of these six factors represent different shared corridor conditions, with three levels for each factor selected accordingly. Larger ranges of these factor levels will yield better parameter estimates. However, smaller factor ranges will yield higher resolution over a smaller region of the true response surface [13]. The levels for traffic *volume*, measured by total trains per day (TPD), are all indicative of a saturated or congested network. At these traffic levels, an additional train will add substantial delays to all other trains. *Mixture* reflects the traffic composition of the line, expressed as the percent of trains on the network that are freight trains. A value of 25% indicates a Passenger Dominated Corridor (PDC) where 75% of the total traffic is passenger and 25% of the traffic is freight. A composition of 75% is a Freight Dominated Corridor (FDC). A value of 50% is evenly split between the two types of traffic. The *percent double track* refers the amount of the route that has two tracks. This calculation includes both track miles of siding and track miles of second mainline track. The range for *maximum freight speed* are chosen to cover typical maximum freight speeds in North America. 30 mph is representative of local or low speed bulk trains. A 50 mph maximum train speed can represent manifest or bulk trains. 70 mph is representative of high speed intermodal train speeds. Without advanced signaling systems, the *maximum speed of a passenger train* is limited to 79 mph. Potential maximum speed upgrades on developing shared corridors across North America are 90 mph and 110 mph.

There are various strategies for how the second mainline track could be constructed in phases and distributed across a corridor. There will most likely be sections of the route that cost more to construct than others. Based on a strategy of minimizing capital investment per mile of double track, the inexpensive sections of double track would be expected to be constructed before the expensive sections. However if cost was not a factor in the decision making process, or if the cost of many double track segments were relatively equal, then there might be an optimal distribution of double track from an operating perspective. Four grouping strategies are illustrated in Figure 2: *extend*, *alternate*, *split*, and *group*.

Table 1: Factors studied in capacity analysis

| Factor | Low | Medium | High |
|---|---|---|---|
| Total Trains Per Day | 40 | 48 | 56 |
| Mixture | 25% | 50% | 75% |
| Percent Double Track | 32.5% | 52.8% | 73.0% |
| Freight Speed (mph) | 30 | 50 | 70 |
| Passenger Speed (mph) | 79 | 90 | 110 |
| Double Track Grouping | Siding Extension→Alternate → →Split→Group | | |

One extreme could be that the additional double track is consolidated in the center of the subdivision (*Group*). A rationale for this approach is to perhaps provide priority trains enough second mainline track in one continuous segment to execute an overtake maneuver and reach maximum speeds after slowing down for a turnout. On the opposite extreme, the double- track resource is allocated to each siding by extending one of the control points further into each of the single-track-bottleneck sections (*Extend*). This strategy shortens the single-track bottleneck sections, thus decreasing the average dwell time at sidings, and increasing the likelihood for a flying meet. This type of long-siding approach has been suggested in the past as a method of accommodating high speed passenger trains in a single track corridor [14–16]. The *extend* strategy requires reconfiguring half the CTC control points each time track is added to the baseline configuration. A different strategy might be to pick four to six points on the subdivision and build-out in both directions from these midpoints (*Alternate*). Another hybrid strategy is to split the double-track resource between the two terminals on the subdivision and build towards the midpoint (*Split*). This strategy has the benefit of addressing potential bottleneck constraints at terminals. In order of increasing continuous double-track units, the ordinal set is extend, alternate, split, and group.

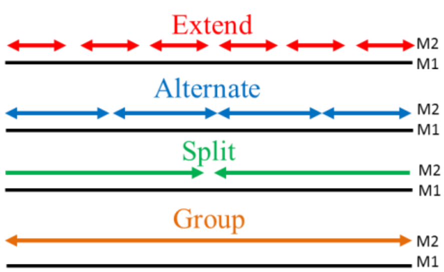


Figure 2: Different strategies on how to allocate sections of second-mainline track to a single track line

The double track percentage and the grouping strategy are the infrastructure related variables. The traffic related variables are the traffic volume, speeds of the trains, and the heterogeneity. Traffic related variables have been studied in previous work [6], [12]. The main effects of the infrastructure variables and their interactions with the traffic variables are a main focus of this study.

There are 972 different permutations of the six factors identified in Table 1. Simulating all 972 scenarios requires extensive computing time. Design of Experiments (DOE) can reduce the amount of scenarios to simulate by intelligently selecting a subset of the full factorial experiment design. Consequently, estimates of key effects and parameters can be made with tight confidence intervals over a smaller dataset. JMP is a commercial statistical analysis software package that is used to create partial factorial designs [17]. In order to create a response surface model for the six factors of interest, JMP suggested a design matrix with 46 different scenarios. The design matrix is summarized in the Annex in Table 9.

## SIMULATION METHODOLOGY

Rail Traffic Controller (RTC) is used to simulate the scenarios suggested by JMP. RTC is the de facto standard for railway simulation analyses in the United States. Users include all Class I Railroads, Amtrak, BART, and many others. An RTC simulation run for a particular scenario analyzes five days of traffic in the corridor [18]. Each simulation run is then repeated six times to yield performance statistics for 30 days of railway traffic for each of the 46 scenarios in the design matrix. If a particular randomized simulation run seed was infeasible for RTC to dispatch, than it is likely that one or more of the six replicates was feasible and the data point can still be used in the final analysis. While the infrastructure is changing between the cases to reflect the varying amount of double track, the configurations follow a set of infrastructure standards summarized in Table 2. For the *alternate*, *split* and *group* strategies, at higher double track levels where existing sidings become incorporated into long double track segments, the control point at one end of each siding is converted to a universal crossover while the other control point is removed. This result in crossovers that are spaced at roughly 10-mile intervals on the double track segments.

Table 2: Route parameters guidelines

| Parameter | Value |
|---|---|
| Length | 240 miles |
| Bottleneck Length* | 8 miles |
| Siding length | 10,000 feet |
| Diverging turnout speed | 45 mph |
| Traffic control system | 2-Block, 3-Aspect CTC |
| Average signal spacing | 2.0 miles |

*Bottleneck length (end-to-end siding spacing) decreases in cases when the double track is distributed by the *extend* strategy

In addition to the physical and operating characteristics identified in Table 1, the schedule of traffic affects corridor performance. Knowing and controlling for the quality of a train schedule prior to simulation is difficult. Consequently, assuming a certain schedule for any given simulation may result in high experimental error which will lead to bias in the results. In order to counteract error due to schedule bias, each train’s departure time is determined through a random uniform distribution over a 24 hour period. The randomized scheduling process is expected to create a range of schedules such that “good” low-delay schedules are averaged against “bad” high-delay schedules over the set of simulation replicates for each scenario. Stable averages can be achieved by averaging train performance over multiple days.

Each train in the simulation is based on the characteristics specified in Table 3. The freight train characteristics are based on the Cambridge Systematics National Rail Freight Infrastructure Capacity and Investment Study (2007) conducted for the Association of American Railroads (AAR) [19]. Freight car tonnages and lengths are based on averages for each car type. The power-to-ton ratios are based on experience and information from the Transportation Research Board Workshop on Railroad Capacity and Corridor Planning (2002) [6].

The passenger train is based on the Midwest 110 mph operation between Chicago and Detroit. The acceleration curves from the simulation model were matched to GPS coordinate data from Amtrak's geometry car. The passenger train stops were spaced at 32.4 mile intervals based on the current Amtrak station spacing on routes in California, Illinois, Washington, and Wisconsin.

Table 3: Train Parameters for Simulation Model

| | Freight Train | Passenger Train |
|---|---|---|
| Locomotives | x3 SD70 | x2 P42 |
| No. of Cars | 115 hopper cars | 8 Single level cars |
| Length (ft.) | 6,325 | 500 |
| Weight (tons) | 16,445 | 800 |
| HP/TT | 0.78 | 15.4 |
| | | 32.4 miles between stops |

Train delay is used as a proxy for capacity in this paper. While train delay is not a flow measurement, high train delays are indicative of a congested or saturated network. Train delays also have different consequences in terms of distance loss if the speeds of the trains are changing. Using a regression analysis of pure run-time will yield the speed of the trains as the most dominant factor. The run-time regression model will not be able to distinguish between an increase in run-time simply due to the slower speed of the trains or an increase in run-time as a result of increasing network congestion due to the presence of the slower trains. This analysis will focus on freight train delays only because these delays are most responsive to the identified factors. Passenger trains are shielded from experiencing very high delays by being the high priority trains. Regression models on passenger train delays will be developed in the future.

## RESULTS

After the 46 scenarios are simulated, a multivariable regression model is built from the results. These models take the input parameters of Table 1 and construct relationships to train delay. Given that six replicates of the simulation are run for each of the 46 scenarios in the design matrix, there are 276 data-points available for the regression analysis. Additionally, the replication allows for "Lack of Fit" analysis to score how a response surface passes through the mean of the replication points. RTC was not able to find a feasible solution for cases that involved 30 mph freight trains and a traffic level of 56 trains per day. These cases were modified by increasing the freight trains speed to 40 mph and were then successfully re-simulated.

The average train delay of a simulation replicate is the response or Y-variable. The train delay is defined to be the difference between the minimum-run-time (MRT) and the actual run time. Where MRT is defined as the fastest a train can traverse across a route without any interference from other trains, slow orders, or other external factors. The delays presented in the subsequent analysis refer to the average extra time to traverse the 240 mile route.

### Model Selection

There are four classes of variables considered for building a multi-variable regression model over the simulated results: *main effects, curvature, continuous interactions,* and *nominal-interactions.* Main effects are the corresponding first order variables representing the factor levels in Table 1. The grouping strategy is considered an ordinal set and has three dummy-indicator variables to represent this factor. Curvature terms are only incorporated for the continuous factors in Table 1. A positive curvature term is indicating that the relationship between train delay and the continuous factor is convex. This results in greater increases in train delays at a high level than a lower level. A negative curvature term is indicative of a concave relationship where larger increases in train delays occur towards the lower end of the factor level. Only second order two-way interaction parameters are considered in this analysis. An additive interaction occurs when an increase in two variables corresponds to a change in train delay greater than what the main effects and curvature terms would independently estimate. An interference interaction occurs when two variables move in opposite directions and the corresponding change in train delay is greater than what the main effects and curvature terms would indicate. A nominal interaction refers to an interaction effect between a particular grouping strategy and one of the continuous variables.

Five different stepwise model creation techniques are used to select an elite group of candidate models. These models are scored against each other based on model simplicity and fit of the data. In general, a simpler model will have tighter confidence intervals of the parameters than a more complicated model, while a more complicated model may fit the data better. The candidate models are summarized in Table 4 in the Annex. In these five models, most of the variation in the dataset is explained by a few parameters. These models differ in which minor two-way interaction parameters, the selection criteria picked to use in the model. Model #1 will be used in the subsequent analysis. This model was agreed upon by two different variable selection protocols and is the simplest. Additionally, the simpler model still has nearly a good of fit as the more complicated models. The model fit statistics are summarized in Tables 6-8 in the Annex.

### Main Effects

The model parameters are indicators of how train delay might change for a unit increase of an independent variable. The sign of the parameter estimates of the continuous variable follow conventional wisdom. Adding more double track will reduce delay. Adding more trains will increase delay. Faster freight trains will incur less delay. Freight trains have lower delays in freight dominated corridors. Higher passenger train speeds can lead to very small reduction in freight train delays. However, this decrease in freight train delay can be overpowered by interactions between maximum passenger train speed and other variables.

The ranges of the independent variables represent a wide range of operating conditions but not necessarily an equal comparison between each other. For example, a change in delay over a change in traffic volume is a very different rate measurement than a change in freight delay per mile of double track. In order to correct for this unit-bias, point-elasticity (Equation 1) is used to show how the response surface reacts to normalized unit changes to a particular factor from a base case scenario. Equation 1 calculates into a dimensionless parameter so this estimate is independent of the units of Y and X variables. For this analysis, the base case is when all the factors in Table 1 are set to their medium levels. The elasticity for an individual factor is calculated twice. The first calculation is to see the maximum delay increase from the base case by manipulating only one variable. This calculation is then repeated by looking at the change in delay when the factor changes from its medium level to the level that yields the greatest delay reduction. For example, a change from 52.8% double track to 32.5% double track will correspond to a large delay increase. An increase in double track to 73% will show a large reduction in train delay.

**Equation 1)**

$$e = \frac{\%\Delta Y}{\%\Delta X_i} \cdot \frac{X_{i_o}}{Y_o}$$

$$\begin{aligned} e &= \text{point elasticity} \\ Y &= \text{dependent variable} \\ X_i &= \text{independent variable } i \\ (X_{i_o}, Y_o) &= \text{Base case conditions} \end{aligned}$$

This particular model will have non-constant elasticities due to the intercept, curvature, and interaction terms. The small changes in elasticity as a function of the independent variables are assumed to be smaller than relative differences between variables. This bias is also reduced by calculating the elasticities in different directions from the base case. The elasticities for the continuous variables are shown in Figure 3. The calculated elasticiteis are grouped by maximum delay reduction and maximum delay increases.

Freight train delays are most sensitive to the speed of the freight train. For a 1% increase in freight train speed, the delay is expected to decrease 2%. Alternatively, a 1% reduction in freight train speed is expected to correspond to a 2.5% increase in average freight train delay. An increase in train speed will result in a reduction in the amount of time that a freight train requires to clear a single-track bottle neck section. This time savings is analogous to reducing the average siding spacing of the route. With a faster train, a meet conflict in time between two opposing trains is more likely to occur near a siding or section of double track. The next most responsive main effect variable is the traffic level of the line. More operating traffic will results in higher train delays. Once a line is congested, railway traffic will experience more conflicts with other trains and incur delays. The amount of double track shows a constant 1% decrease in train delay per 1% increase in double track mileage. Traffic mixture and passenger train maximum speed both showed very small effects on train delays.

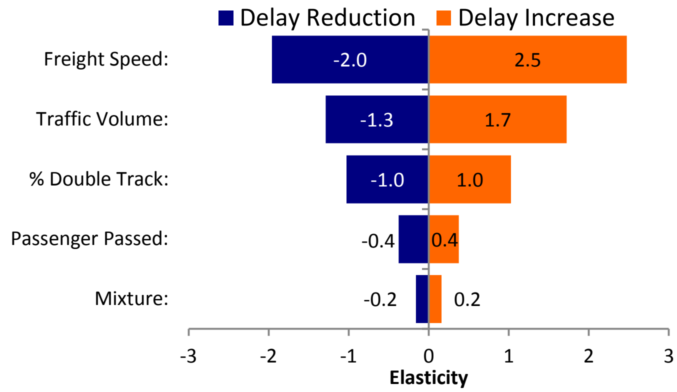


Figure 3: Point elasticities for the five continuous variables where only 1 variable is being changed from a base case to achieve a maximum delay reduction and maximum delay increase

**Interaction Effects**

The final model did include two-way interaction terms. There are greater changes in train delay by manipulating two variables simultaneous instead of one at a time as shown in the previous sections. The analysis in subsequent sections will vary two out of the six factors of Table 1 and keep the other four constant.

The double track distribution had a modest effect on train delay as shown in Figure 4. The grouping strategies are compared against each other where most of the other factors are held constant at their medium levels. Grouping strategy has a low interaction effect with the amount of double track added. Most of the changes in delay are due to the increase double track miles rather than the distribution strategy. At 32% double track, the maximum difference between an *extend* strategy and a *split* strategy is 33 minutes and this differences decreases to 26 minutes at 72% double track. There is a much stronger interaction between the grouping strategy and the maximum freight train speed. With 30 mph freight trains, the difference between *extend* and *split* is 57 minutes. When the freight train speed is 70 mph, the maximum difference between strategies is 4.7 minutes between *extend* and *alternate* strategies. The *extend* strategy appears to be the best performer by a small to modest amount across the board. This might because the *extend* strategy uniformly decreases the length of single track bottleneck sections across the entire route while the other strategies preserve some of the original single track bottlenecks. However, the *extend* configuration is the most difficult to implement over time. There could be very high costs to re-extending every siding and reconfiguring the control points to effectively add more second track each time there is an increase in traffic. In the other strategies, once a section of track is double-tracked, that particular section is not reconfigured again. Thus the *alternate*, *split* and *group* strategies present much more practical methods for the phased addition of second track as traffic gradually increases.

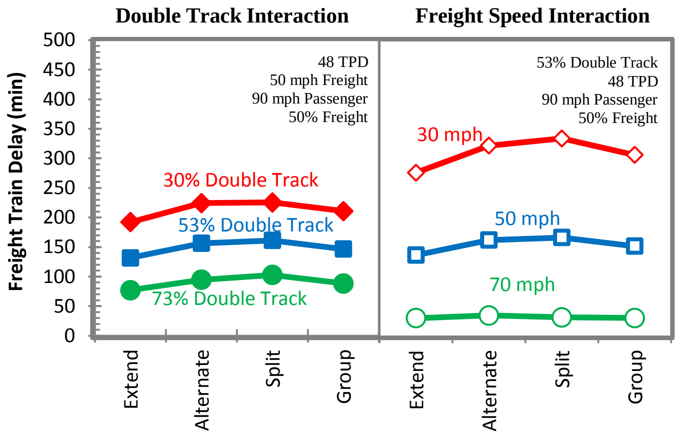


Figure 4: Delay sensitivity to grouping strategy (*left*) when freight train speed is held constant and the amount of double track is varying and (*right*) when the double track s held constant and the freight train speed is held constant. All other factors are set to the medium levels of Table 1.

Beyond the double track grouping effects there were significant interactions between certain factors. Figure 5 shows the interaction effect between amount of double track and the maximum speed of the freight train as well as the interaction between amount of double track and traffic volume. As more miles of double track are installed, the train delay is expected to decrease in a linear manner. There are greater reductions in train delay when the freight speed is 30 mph compared to a gradual reduction in delay when freight speed is 70 mph. A similar interaction occurs between double track and volume. When more double track sections are added, delay decreases linearly. There are greater reductions in train delays when the traffic level is 56 trains per day compared with 40 trains per day. The interaction between freight train speed and the amount of double track is larger than the interaction between the traffic volume and the amount of double track.

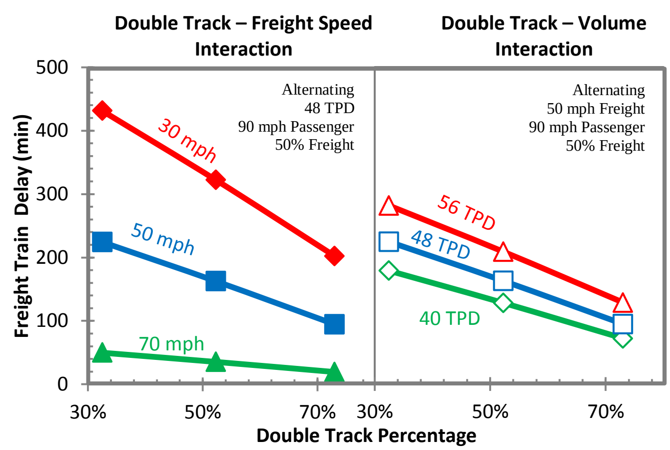


Figure 5: Interaction effects with double track and (*left*) freight speed and (*right*) double track traffic volume.

The maximum speed of the freight train has so far shown to have significant interaction effects with the amount of double track in the route and grouping strategy. Two more significant interactions are shown in Figure 5. Slower 30 mph freight trains are more sensitive to traffic volume increases compared with 70 mph freight trains. Additionally, the effect of heterogeneity diminishes with faster freight trains. Slower 30 mph freight trains perform worse than 70 mph freight trains in passenger dominated corridors, presumably because the slower freight trains exhibit the greatest disparity in performance characteristics compared to passenger trains.

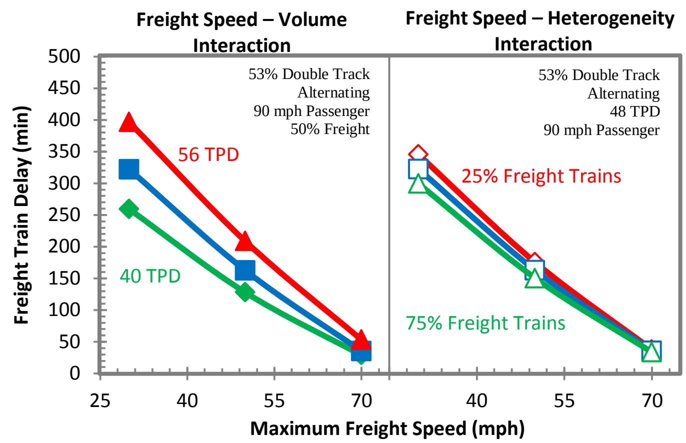


Figure 6: Interaction effects between (*left*) maximum freight speed and traffic volume and (*right*) maximum freight speed and heterogeneity measured by the percent of freight trains on the network.

## Model Validation

Regression models tend to have a good fit to the data that created the model but have limited utility unless they also have a good fit to "unseen" data. To validate the response surface model, 22 additional scenarios were simulated that were not selected in the original partial-factorial design. The response surface model should be able to predict these cases with the same accuracy that it predicts the original data used in fitting the model. The mean square prediction error (MSPR) should be about the same magnitude as the mean square error (MSE) of the regression model. This is an indication that the model is unbiased and can have reasonable predictive ability. The MSPR for the validation set is 226.1 and the MSE of the response surface model is 204.2. Since the MSE and MSPR are the same magnitude, the model is valid [13].

**Equation 2)**

$$MSPR = \frac{\sum_{i=1}^{n^*}\left(Y_i - \hat{Y}_i\right)^2}{n^*}$$

$Y_i =$ value of the response variable at the $i^{th}$ case in the validation set

$\hat{Y}_i =$ predicted value of the response variable at the $i^{th}$ case in the model building set

$n^* =$ size of the validation set

## FUTURE WORK

The factor ranges in this analysis may have significant impacts on the relative sizes of the parameter estimates. For example, traffic volumes of 40 and 56 trains per day give high resolution of a congested network but do not indicate performance of a free-flow network. If the range for traffic volume was 8 to 56 trains per day, then traffic volume may have shown a larger relative effect on train delays. Additionally, as demonstrated in previous work, a 100% double track line with homogenous traffic should have no delays under the traffic levels used for this analysis [8]. By only looking in the range of 25% to 75% freight trains, the effect of heterogeneity may be underestimated. This range limitation also applies to the amount of double track added. Figure 5 shows a linear relationship between double track and train delays in the range of 30% to 70% double track. This linear trend may not continue in the range of 70% to 100% double track. The design matrix will be augmented in the future to incorporate new ranges of the identified factors.

Regarding grouping strategy, one of the stated reasons for pursuing a *split* strategy is to add double track near terminals where congestion often occurs. The RTC simulation model has a limited ability to account for yard congestion and the ensuing delay when trains are held outside terminals. A simulation model that specifically addresses mainline-yard interaction may reveal additional delay reduction benefits of double track near terminals. These delay reduction benefits that are not captured in the current mainline RTC model may place the split strategy in a more favorable position relative to the other strategies.

## CONCLUSION

Partial-factorial and response surface modeling can be used to quickly compare relative effects in railway simulation analysis. If a single-track corridor is being upgraded to a double track corridor, major factors influencing the amount of delay include the volume, the amount of double track installed, and the speed of the freight train. The evidence of this study suggests that train delays are expected to reduce in a linear manner per mile of double track installed. The distribution of the double track can have a modest effect on train delays. Minor effects include the passenger train speed and the heterogeneity of the line.


## ACKNOWLEDGMENTS

- Support from CN Research Fellowship in Railroad Engineering
- National University Rail Center (NuRail)
- Eric Wilson of Berkeley Simulation Software
- Samantha G. Chadwick, Graduate Research Assistant (University of Illinois at Urbana-Champaign)
- Mark Dingler of CSX Transportation
- Undergraduate research assistants:
  - Scott Schmidt
  - Ivan Atanassov


## REFERENCES


[1] C. W. Richards and T. Cobb, “Multipe-Track Main Lines,” *Trains*, Jan-2006.

[2] A. M. Kahn, “Railway Capacity Analysis and Related Methodology,” Ottawa, 1979.

[3] J. Pachl, *Railway Operation and Control*, 2nd ed. Mountlake Terrace: VTD Rail Publishing, 2009.

[4] M. Abril, F. Barber, L. Ingolotti, and M. Salido, “An Assessment of Railway Capacity,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 5, pp. 774–806, 2008.

[5] S. L. Sogin, C. P. L. Barkan, and M. R. Saat, “Simulating the Effects of Higher Speed Passenger Trains in Single Track Freight Networks,” in *Proceedings of the 2011 Winter Simulation Conference*, 2011, pp. 3679–3687.

[6] M. Dingler, Y.-C. Lai, and C. P. L. Barkan, “Impact of Train Type Heterogeneity on Single-Track Railway Capacity,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 640, no. 2117, pp. 41–49, 2009.

[7] M. Dingler, A. Koenig, S. L. Sogin, and C. P. L. Barkan, “Determining the Causes of Train Delay,” in *AREMA Annual Conference Proceedings*, 2010.

[8] S. L. Sogin, C. P. L. Barkan, Y.-C. Lai, and M. R. Saat, “Impact of Passenger Trains in Double Track Networks,” in *Proceedings of the 2012 Joint Rail Conference*, 2012.

[9] J. C. Prokopy and R. B. Rubin, “Parametric Analysis of Railway Line Capacity,” Federal Railway Administration, Washington, D.C., 1975.

[10] H. Krueger, “Parametric Modeling in Rail Capacity Planning,” in *Proceedings of the 1999 Winter Simulation Conference*, 1999, pp. 1194–1200.

[11] S. Mitra, D. Tolliver, and S. Mitra, “Estimation of Railroad Capacity Using Parametric Methods,” vol. 49, no. 2, 2010.

[12] Y.-C. Lai and Y.-A. Huang, “Estimation of Single and Double-Track Capacity with Parametric Models,” in *Transportation Research Board: 91st Annual Meeting*, 2012, pp. 1–22.

[13] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th ed. Irwin, 2004.

[14] E. Petersen and A. Taylor, “Design of Single-Track Rail Line for High-Speed Trains,” *Transportation Research Part A: General*, vol. 21, no. 1, 1987.

[15] S. Harrod, “Capacity Factors of a Mixed Speed Railway Network,” *Transportation Research Part E*, vol. 45, no. 5, pp. 830–841, 2009.

[16] S. P. S. Pawar, “An Analysis of Single Track High Speed Rail Operation,” University of Birmingham, 2011.

[17] “JMP.” SAS Institute INC., Cary, NC.

[18] E. Willson, “Rail Traffic Controller (RTC).” Berkeley Simulation Software, 2012.

[19] Cambridge Systematics, “National Rail Freight Infrastructure Capacity and Investment Study,” 2007.

## ANNEX

Table 4: Candidate Response Surface Models

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Terms | 27 | 27 | 28 | 26 | 28 |
| Main Effects | 8 | 8 | 8 | 8 | 8 |
| Curvature | 2 | 2 | 2 | 2 | 2 |
| Interactions (continuous) | 8 | 8 | 8 | 8 | 8 |
| Interactions (grouping) | 9 | 9 | 10 | 8 | 10 |
| RMSE | 14.240 | 14.240 | 14.230 | 14.289 | 14.230 |
| R-Square | 0.98037 | 0.98037 | 0.98048 | 0.98015 | 0.98048 |
| R-Square Adjusted | 0.97806 | 0.97806 | 0.97809 | 0.97791 | 0.97809 |
| Model Statistic | 423.575 | 423.575 | 409.099 | 436.773 | 409.099 |
| AIC | 2,130.582 | 2,130.582 | 2,131.642 | 2,130.914 | 2,131.642 |
| Cp | 22.649 | 22.649 | 23.343 | 23.165 | 23.343 |
| Lack of Fit | 1.4200 | 1.4200 | 1.4222 | 1.4853 | 1.4222 |
| Selection Technique | Forward AIC | Backward AIC | Forward P-Value | Backward P-Value | Mixed P-Value |

Table 5: Regression ANOVA Table for Model #1

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Ratio** |
|---|---|---|---|---|
| Model | 27 | 2,319,211.3 | 85,896.7 | 423.5745 |
| Error | 229 | 46,438.9 | 202.8 | **Prob > F** |
| C. Total | 256 | 2,365,650.2 | | <.0001* |

Table 6: ANOVA Table for Lack of Fit Analysis for Model #1

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Ratio** |
|---|---|---|---|---|
| Lack Of Fit | 18 | 5,017.5 | 278.8 | 1.42 |
| Pure Error | 211 | 41,421.4 | 196.3 | **Prob > F** |
| Total Error | 229 | 46,438.9 | | 0.1242 |

Table 7: Parameters Used for Model Selection

| | **Terms** | |
|---|---|---|
| **Main Effects** | DT, VOL, F_SPD, P_SPD, MIX, STRAT | |
| **Curvature** | $DT^2$, $VOL^2$, $F\_SPD^2$ $P\_SPD^2$, $MIX^2$ | |
| **Interactions** | (DT·VOL)<br>(DT·F_SPD)<br>(DT·P_SPD)<br>(DT·MIX)<br>(DT·STRAT)<br>(F_SPD·P_SPD)<br>(F_SPD·MIX)<br>(F_SPD·STRAT) | (VOL·F_SPD)<br>(VOL·P_SPD)<br>(VOL·MIX)<br>(VOL·STRAT)<br>(P_SPD·MIX)<br>(P_SPD·STRAT)<br>(MIX·STRAT) |

DT = Percentage Double Track,
VOL = Trains Per Day,
F_SPD = Freight Train Speed,
P_SPD = Passenger Train Speed,
MIX = Freight Train Percentage,
STRAT = Double Track Grouping Strategy

Table 8: Parameter Estimates from Model #1

| **Term** | **Estimate** | **Std Error** | **t Ratio** | **Prob>\|t\|** |
|---|---|---|---|---|
| Freight Speed | -136.3976 | 1.937148 | -70.41 | <.0001* |
| % Double Track | -59.85581 | 1.450053 | -41.28 | <.0001* |
| % Double Track*Freight Speed | 49.629093 | 1.71898 | 28.87 | <.0001* |
| Traffic Volume(40,56) | 34.913622 | 1.561362 | 22.36 | <.0001* |
| Traffic Volume*Freight Speed | -28.19366 | 1.964621 | -14.35 | <.0001* |
| Traffic Mixture(0.25,0.75) | -16.70285 | 1.440257 | -11.60 | <.0001* |
| % Double Track*Traffic Volume | -11.53321 | 1.628588 | -7.08 | <.0001* |
| Δ(Extend-Alternate) | -10.10322 | 1.612663 | -6.26 | <.0001* |
| Freight Speed*Traffic Mixture | 10.431946 | 1.683469 | 6.20 | <.0001* |
| Freight Speed*Freight Speed | 16.188184 | 2.678693 | 6.04 | <.0001* |
| Δ(Alternate-Split) | -8.563164 | 1.421265 | -6.03 | <.0001* |
| Δ(Alternate-Split)*Freight Speed | 9.0019488 | 1.803419 | 4.99 | <.0001* |
| Δ(Extend-Alternate)*Freight Speed | 10.252475 | 2.091186 | 4.90 | <.0001* |
| Δ(Extend-Alternate)*Passenger Speed | 8.0138568 | 1.905612 | 4.21 | <.0001* |
| Passenger Speed*Traffic Mixture | 6.3562952 | 1.562024 | 4.07 | <.0001* |
| Δ(Alternate-Split)*Traffic Mixture | 5.4047383 | 1.563343 | 3.46 | 0.0007* |
| % Double Track*Passenger Speed | 4.7102704 | 1.429266 | 3.30 | 0.0011* |
| (ΔSplit-group)*traffic mixture | -4.519154 | 1.403029 | -3.22 | 0.0015* |
| (ΔSplit-group)*traffic volume | 3.7871897 | 1.253949 | 3.02 | 0.0028* |
| Δ(Extend-Alternate)*Traffic Mixture | -5.194092 | 1.899609 | -2.73 | 0.0067* |
| Traffic Volume*Passenger Speed | -3.841643 | 1.484569 | -2.59 | 0.0103* |
| Δ(Split-Group) | 3.0179989 | 1.219342 | 2.48 | 0.0140* |
| Traffic Volume*Traffic Volume | 5.6915348 | 2.381097 | 2.39 | 0.0176* |
| Freight Speed*Passenger Speed | 3.9830354 | 1.679401 | 2.37 | 0.0185* |
| Δ(Extend-Alternate)*% Double Track | 3.6126354 | 1.98485 | 1.82 | 0.0700 |
| Passenger Speed(79,110) | -2.287555 | 1.316719 | -1.74 | 0.0837 |
| (ΔSplit-Group)*Freight Speed | -2.22681 | 1.387371 | -1.61 | 0.1099 |

Table 9: Experiment Design Matrix

| Double Track (%) | Grouping Strategy | Traffic Volume (TPD) | Freight Speed (mph) | Passenger Speed (mph) | % Freight Trains | Avg. Delay (min) |
|---|---|---|---|---|---|---|
| 32.5% | Split | 48 | 30 | 90 | 25% | 462.2 |
| 32.5% | Extend | 56 | 40 | 90 | 25% | 398.6 |
| 32.5% | Split | 40 | 30 | 79 | 75% | 327.4 |
| 32.5% | Group | 40 | 30 | 110 | 75% | 317.5 |
| 32.5% | Split | 56 | 50 | 110 | 50% | 293.3 |
| 52.8% | Alternate | 40 | 30 | 79 | 25% | 288.3 |
| 32.5% | Group | 56 | 50 | 79 | 25% | 283.2 |
| 52.8% | Group | 56 | 40 | 110 | 25% | 269.7 |
| 52.8% | Group | 56 | 40 | 79 | 75% | 260.4 |
| 52.8% | Split | 56 | 50 | 79 | 25% | 248.3 |
| 52.8% | Extend | 48 | 30 | 79 | 75% | 243.8 |
| 52.8% | Extend | 40 | 30 | 110 | 50% | 209.0 |
| 73.0% | Split | 40 | 30 | 110 | 25% | 203.8 |
| 73.0% | Alternate | 56 | 40 | 110 | 25% | 180.6 |
| 73.0% | Group | 40 | 30 | 79 | 25% | 174.2 |
| 32.5% | Extend | 48 | 50 | 90 | 75% | 172.0 |
| 32.5% | Extend | 40 | 50 | 79 | 25% | 168.5 |
| 32.5% | Alternate | 40 | 50 | 110 | 25% | 168.3 |
| 73.0% | Extend | 48 | 30 | 79 | 50% | 156.6 |
| 52.8% | Alternate | 48 | 50 | 90 | 50% | 154.5 |
| 52.8% | Group | 48 | 50 | 90 | 50% | 143.2 |
| 52.8% | Split | 48 | 50 | 90 | 75% | 131.1 |
| 73.0% | Alternate | 40 | 30 | 90 | 75% | 128.1 |
| 73.0% | Alternate | 56 | 50 | 79 | 75% | 115.9 |
| 73.0% | Extend | 48 | 50 | 90 | 25% | 101.5 |
| 73.0% | Extend | 56 | 50 | 110 | 75% | 100.0 |
| 73.0% | Group | 48 | 50 | 110 | 50% | 97.6 |
| 73.0% | Split | 40 | 50 | 79 | 50% | 71.9 |
| 32.5% | Alternate | 48 | 70 | 79 | 25% | 59.6 |
| 52.8% | Alternate | 56 | 70 | 90 | 25% | 56.1 |
| 32.5% | Group | 56 | 70 | 110 | 75% | 55.4 |
| 32.5% | Split | 56 | 70 | 79 | 75% | 50.7 |
| 32.5% | Extend | 48 | 70 | 110 | 25% | 50.0 |
| 32.5% | Split | 40 | 70 | 90 | 25% | 42.3 |
| 52.8% | Extend | 56 | 70 | 79 | 50% | 39.3 |
| 32.5% | Group | 40 | 70 | 79 | 50% | 36.2 |
| 73.0% | Split | 56 | 70 | 110 | 25% | 34.6 |
| 52.8% | Group | 40 | 70 | 110 | 25% | 32.9 |
| 32.5% | Alternate | 40 | 70 | 90 | 75% | 32.4 |
| 73.0% | Alternate | 40 | 70 | 79 | 25% | 26.7 |
| 73.0% | Group | 56 | 70 | 79 | 25% | 25.6 |
| 73.0% | Alternate | 48 | 70 | 110 | 75% | 24.8 |
| 73.0% | Split | 48 | 70 | 90 | 50% | 23.8 |
| 52.8% | Split | 40 | 70 | 110 | 75% | 23.0 |
| 73.0% | Extend | 40 | 70 | 90 | 75% | 18.1 |
| 73.0% | Group | 40 | 70 | 79 | 75% | 14.8 |