Impact of Operational Practices on Rail Line Capacity: A Simulation Analysis

Mark Dingler[1*], Yung-Cheng (Rex) Lai[2], Christopher P.L. Barkan[1]

[1] *Railroad Engineering Program, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Newmark Civil Engineering Laboratory, 205 N. Mathews Ave., Urbana IL, 61801 USA*
[2] *Department of Civil Engineering, National Taiwan University, Rm 313, Civil Engineering Bldg, No 1, Roosevelt Rd, Sec 4, Taipei, 10617 Taiwan*

e-mail: dingler2@illinois.edu, yclai@ntu.edu.tw, cbarkan@illinois.edu

ABSTRACT

Long-term demand for freight and passenger rail traffic in North America is expected to expand increasingly leading to capacity constraints. Both infrastructure investment and operational changes can relieve congestion. However, given the high cost to build and maintain infrastructure, careful consideration of how operational practices can affect or mitigate demand is critical for cost-effective planning of new capacity. A key aspect of this is the effect of heterogeneous traffic characteristics on capacity. Different freight and passenger trains have substantially different operating characteristics including: speed, acceleration, braking and dispatching priorities. Greater heterogeneity on a line increases interference between trains and creates more delays than if all trains have similar characteristics. Train dispatching simulation software was used to analyze the impact of heterogeneity under a range of realistic infrastructure configurations and operational conditions. Analyses were conducted to evaluate the effectiveness of various operational changes to reduce delays. The trade-off between reduced congestion and increased cost was considered for each change. A benefit of operational changes is that they can be more rapidly implemented and offer greater flexibility than capital infrastructure investments if future traffic patterns change.

Submitted for the Proceedings of the 2009 Annual AREMA Conference

INTRODUCTION

Long-term demand for North American railroad freight traffic is projected to increase. The American Association of State Highway and Transportation Officials (AASHTO) predicts that freight rail service will increase 84% based on ton-miles by 2035 *(1)*. In addition, recently proposed expansion in intercity passenger rail service *(2)* will increase the number of passenger trains on certain freight rail corridors, as will commuter rail service in some metropolitan areas, thereby placing further demand on capacity on the affected routes. In order to accommodate this new traffic, various changes in both operations and infrastructure will be required. Infrastructure expansion is capital intensive and the need for it is affected by operational practices. Furthermore, some additional capacity may be achieved by alteration of operations. Efficient planning and investment in capacity projects requires understanding how operational aspects of new traffic will affect capacity, as well as how operational changes can mitigate the demand *(3)*.

A key factor affecting rail capacity is the interaction of different train types. Heterogeneity in train characteristics causes greater delays than a corresponding number of homogeneous trains would. In North America intermodal, manifest, unit and local trains may all share trackage. Some lines also have intercity passenger trains, and in metropolitan regions, commuter trains. Each of these train types can have considerably different characteristics and even trains of the same class may have varying weights and lengths. This heterogeneity has a substantial effect on rail line capacity *(4,5,6)*.

Previous work has investigated several of the causes of train type heterogeneity and train delay. In Spain, Abril et al *(7)* considered trains operating at two speeds, "normal" and 50% of normal on single- and double-track lines. In North America, Bronzini and Clarke *(8)* used a single-track simulation model to compare the delay-volume curves of different mixtures of intermodal and unit trains. Harrod *(9)* compared the differing impact of faster and slower non-conforming trains and found that the slower the non-conforming train, the greater the impact on the network. Gorman *(10)* used actual traffic data to statistically estimate delay. He found that the most useful measures of train speed heterogeneity for predicting congestion delay are meets, passes and overtakes.

By changing operations to reduce the impact of the factors that contribute to train delay additional capacity can be gained without additional infrastructure. However some operational changes increase costs and these factors need to be considered when deciding if methods to reduce the impacts of train type heterogeneity are beneficial.

Dispatching simulation software was used to investigate the impact of various operational changes on differing levels of heterogeneity on a hypothetical signalized, single-track rail line. Changes considered included: increase in train headway, equalizing maximum speeds, additional power and removing dispatching priorities. Delay was used as the principal metric to assess capacity impacts under different scenarios. The objective of this work is to provide insight into which aspects of train type heterogeneity have the greatest impact on delay and investigate the potential use of different operational changes that may reduce this delay.

METHODOLOGY

Railroad capacity is influenced by a complex relationship of infrastructure and operational parameters. Operational factors include: average and variability in operating speed, traffic volume, stability, terminal efficiency and heterogeneity in various train characteristics. These are interrelated with, and further influenced by, infrastructure characteristics such as: siding length and spacing, crossover spacing, number of tracks, signal and traffic control system, grade, and curvature. For our work we studied the impact of operational factors, with a specific focus on train type heterogeneity. Dispatch simulation software was used to model and simulate multiple traffic scenarios.

Capacity Metric

Delay was used as the principal metric for capacity comparisons in this study. We define delay as the difference between the minimum, or unopposed, run time, and the actual run time required to traverse the route. This includes the time spent stopped for meets and passes, along with the time for braking, and to accelerate from stops. The total delay was divided by the train miles for a normalized value of delay per 100 train miles. There has been some discussion about the use of delay as a metric of capacity *(11)*, however for the types of comparisons and circumstances addressed in our study, delay is a generally satisfactory measure and is used throughout this paper.

Dispatch Simulation Software

We used Rail Traffic Controller (RTC) from Berkeley Simulation Software for our analyses. RTC is a sophisticated software program designed to realistically simulate both freight and passenger operations over a railroad network *(12,13)*. The software uses infrastructure and traffic inputs specified by the user to resolve multi-train conflicts in a manner intended to mimic decisions of an actual railroad dispatcher. We used RTC because its flexibility permits rapid evaluation of a variety of different scenarios, and because of its widespread acceptance and use by the North American railroad industry.

Representative Rail Line

Specific characteristics of individual rail lines are unique and route characteristics influence the study of railroad operations. For our research we developed a hypothetical rail line intended to represent the characteristics of a typical midwestern North America, single-track mainline subdivision with the following attributes:

- 124 miles long
- 10 miles between ends of sidings
- 8,000 ft signaled sidings with #24 powered turnouts
- 2.5-mile signal spacing
- 3-block, 4-aspect signaling
- 0% grade and curvature

Although the attributes are somewhat idealized, the purpose is to provide a consistent basis for relative comparison of different scenarios of interest in this research under a reasonably realistic set of operating conditions. However, there is no intent to imply that the results presented here represent absolute predictive measurements for a particular set of conditions.

Train Types

Based on our previous research we selected the combination of intermodal and unit trains to investigate because we found that this scenario resulted in the greatest delays due to heterogeneity. The “intermodal” trains represent freight trains with the highest maximum speeds, power to ton ratios and dispatching priorities, while the “unit” trains represent those with the lowest speeds, power to ton ratios and dispatching priorities. Although we refer to them using the terms “intermodal” and “unit” for convenience, what is actually of significance in all our analyses are their specific operating characteristics, not the particular type of consist.

The TRB Workshop on Railroad Capacity and Corridor Planning *(14)* provided typical weights, lengths and horsepower to trailing ton ratios (HPTT) for various train types. We used this information as the principal basis for the physical characteristics of the train types used in this analysis (Table 1). The non-physical characteristic of each train is the priority assigned to it by the dispatcher. When two trains meet, priority is one factor the dispatcher will take into consideration when determining how to resolve the conflict. Generally dispatchers will try to minimize the total cost of delay *(12)*, this means that the trains carrying lower value, less time-sensitive freight will have lower priority and enter the siding, while the higher priority train holds the main and proceeds with little or no delay. In our study intermodal trains were assigned the higher priority, as is typical in most railroad operations.

Simulations

A base case of 46 trains per day was used comprised of various percentages of intermodal and unit trains. This is a theoretical traffic volume based on an equal temporal distribution of trains in each direction over a 24 hour period. It is not intended to represent practical, sustained operation, which would require inspection and maintenance windows, but rather to provide a basis for relative comparison of the effect of various factors of interest. The results are therefore more characteristic of the spacing or headway between trains than the actual volume. The ratio of train types was experimentally altered by varying the percentage of each type from 0% to 100%. The ratios and traffic pattern were the same for trains traveling in both directions. For each configuration a series of five simulations were performed with the departure time of each train randomized over a 20-minute interval, 10 minutes before or after the scheduled time for that train.

IMPACTS OF TRAIN TYPE HETEROGENEITY

With heterogeneous traffic, delay is also caused by conflicts that occur as a result of differences in train characteristics, some of which also increase meet frequency and duration. Additional sources of delay with heterogeneous traffic include:

- Train delayed by a slower train
- Train delayed by a train with slower acceleration
- Trains experience longer meets waiting for higher priority trains
- Train delayed waiting for another train to pass
- Trains experience more meets due to lower average speed, which can be caused by lower speed, lower power and/or lower priority

The magnitude of these delays is dependent on the specific train mix, volume and amount of heterogeneity *(6)*. Operational changes were considered as methods to reduce train delay due to heterogeneity.

Reducing Peak Traffic Levels

Traffic density and higher delays due to heterogeneity are directly related. If traffic can be more evenly distributed throughout the day total train delay can be reduced. Two headways were considered, the base scenario of 62 minutes between each train, corresponding to 46 trains per day if distributed over 24 hours and headways of 72 minutes, corresponding to 40 trains per day.

Delays were reduced in all traffic configurations (Figure 1a) and with both train types (Figure 1b). The unit train traffic had a larger reduction than the intermodal traffic but as the percentage of unit trains increase the reduction in delay declined.

There is typically an exponential relationship between delay and volume *(8, 15)* and higher volumes correspond to smaller headways. Due to the exponential nature of this relationship an increase in headway may result in a disproportionately larger reduction in delay. Smaller headways result in more meets and increase the possibility of a following train being delayed by a preceding train.

Reducing Heterogeneity in Speed

Heterogeneity in train speed leads to additional conflicts between traffic. To test the influence of heterogeneous speeds on delay the maximum speed of the intermodal traffic was reduced from 70 mph to 50 mph, while all other parameters were held constant. Increasing the maximum speed of the unit train traffic was not considered because with the power and tonnage configuration used for the unit trains they were unable to reach 60 mph.

The average delay decreased only in the scenarios with the highest levels of heterogeneity (Figure 2a). When this delay is separated by train type it reveals that the intermodal trains experienced little to no change in delay, while the unit trains' delay was reduced when the majority of traffic was intermodal (0% to 50% unit trains) (Figure 2c).

Reducing the speed of the intermodal traffic eliminates the speed difference among the train types thereby making them more homogeneous and reducing delay. When trains travel at different speeds both the faster and the slower trains may be delayed. The faster train will be delayed when it approaches a slower train and must reduce speed to maintain a safe headway until it reaches a siding and can pass. The slower train will be delayed while waiting to be passed and due to the more frequent meets at higher volumes, these trains may never be able regain top speed between sidings. Therefore reducing their maximum speed will have little effect on the observed run time and greater homogeneity in speed will not affect capacity. In our simulations there was little change in delay for the intermodal trains. When combined with a decrease in maximum speed this results in increased run times and thus reduced capacity.

Increasing Horsepower to Trailing Ton Ratio

If the power to ton ratio of trains can be increased the delay due to meets can be reduced because the higher powered trains will accelerate faster. To test the effectiveness of increasing the power on the lower-powered, lower-priority unit trains, the unit trains were given an additional locomotive, increasing the HPTT from 0.78 to 1.05.

As the percentage of unit trains increased, the reduction in delay increased (Figure 3a). The reduction in delay for the unit trains was relatively constant over all train mixes (Figure 3b), and therefore as the percentage of unit trains increased the average delay increased correspondingly. However, the intermodal traffic experienced no change in delay.

During a meet a train must slow to a stop and then accelerate back to its normal speed. The time spent on the mainline accelerating has the potential to delay following trains if they are following closely enough behind. An additional operating challenge is that the lower priority trains will often have a lower HPTT ratio, thus the impact of lower priority trains is compounded because they have more stops and slower acceleration from these stops.

Removing Priorities

The last characteristic considered was the dispatching priority assigned to trains. Intermodal trains with their higher value merchandise and greater customer demand for fast, reliable service are typically given higher priority by railroad dispatchers. To test the effect of differential train priority on capacity, the dispatching priorities were removed for the two train types.

In the baseline scenario, delay increased with increased heterogeneity *(6)*. When individual train-type priorities were removed the delay increased as the percentage of unit trains increased (Figure 4a) effectively eliminating heterogeneity-related delays. The reduction in delay due to removing dispatching priorities was greatest when the traffic was mostly the higher priority intermodal traffic and decreased as the percentage of unit trains increased (Figure 4b). When there was a higher percentage of unit trains the delay to the intermodal trains increased when their priority was removed.

Priority is a characteristic given to particular trains in order to decrease their delay and increase the service quality for the higher valued freight. However not surprisingly, this priority increases the delay for the lower priority traffic, causing a higher overall average delay. When there are no priorities the first train to reach a siding will enter it for the meet. However with differential dispatching priorities the lower priority train may enter an earlier siding in order to prevent delays to the higher priority on-coming train. This will result in greater dwell times for the lower priority train and increase overall delay.

MITIGATION OF TRAIN TYPE HETEROGNEITY

One of the implications of understanding the operational factors that influence delay is that this knowledge can be used to change operations to reduce delay. The amount of heterogeneity on a route may be difficult to change due to various external factors; however, certain train characteristics may be more flexible. On a single-track route the main source of delay and principal consumer of capacity are meets and passes; therefore, to improve capacity the number of, or delay during, meets or passes must be reduced. Increasing headways and removing dispatching priorities reduces the number of meets, equalizing train speeds eliminates passes, and increased power on lower priority trains shortens the time lost in meets. Each may provide benefits due to reduced delays and increased capacity; however, they come with additional costs and this tradeoff needs to be considered.

Reducing Peak Traffic Levels

Increasing headways provides benefits to all train types under all traffic mixes and will increase reliability, by increasing the ability to absorb delays due to typical operational incidents *(7)*. To increase headways trains must be rescheduled in order to more evenly distribute traffic over the day. This requires planning and may require changing yard operations and other factors affecting scheduling in order to release trains at different times. Reducing peak traffic levels to reduce delay due to heterogeneity has benefits under all traffic conditions and should be among the first options considered as a way to increase capacity.

Reducing Heterogeneity in Speed

Heterogeneity in speed creates the need for overtakes and passes. However, the results of our simulations indicated that reducing this heterogeneity had little or no benefit, and might even reduce overall line capacity due to lower speeds. It is unlikely that the benefits of reduced delays will offset the increased run times. Additionally, the reduction in speed may not change operations if, as happens with high volumes, trains never get to their top speed due to the many meets. However there may be additional, non-capacity benefits including reduced fuel consumption. If train speeds can be increased instead of reduced, the benefits will be greater; however, it may not be possible without additional locomotives.

Increasing Power to Ton Ratio

Adding locomotives to lower priority unit trains reduced the delay on these trains in a similar manner for all train scenarios. However, additional locomotives are a major capital investment and will also increase maintenance and fuel costs. The increased power reduces time lost in meets for the unit trains, thereby reducing run times; however, if delays to following trains are not also reduced there will be no improvement in capacity. If additional locomotives can be combined with increased speed the benefits will be enhanced. The effectiveness of this method of increasing capacity is dependent on reducing the delays experienced by following trains due to the slow acceleration of a preceding train. Urban and other high density scenarios are the most likely situations where this method would be beneficial.

Removing Priorities

Removing differential train priorities was the most effective method of reducing average train delay. The largest benefit was when heterogeneity was greatest. However the individual benefits to each train decreased as the percentage of unit trains increased. Removing dispatching priorities has no additional cost but the increased run times for the higher priority traffic may be unacceptable to customers. This method is most effective when the higher priority trains are a majority. In this situation removing priorities, prevents lower priority trains from having excessive delays due to more and longer meets and passes.

DISCUSSION

The increased delays due to heterogeneity are caused by a combination of factors including headway, difference in speeds, power to ton ratios and priorities. Changing any of these can reduce delays and potentially increase capacity. However operational changes can also result in higher costs.

In each scenario the unit trains had the greatest reduction in delay. Unit trains often transport lower value, less urgent commodities, while intermodal traffic is high value, time-sensitive traffic. The trade off between practices that reduce delay to unit trains but increase delays to intermodal traffic needs to be carefully considered.

These methods have the potential to help in niche environments and to relieve short-term capacity constraints. Operational changes also have the advantage that they can be implemented more rapidly than infrastructure changes. Such changes also offer more flexibility in response to changing traffic levels and patterns and may provide relief during short periods of high traffic volumes, or as an interim measure while additional infrastructure is built.

Heterogeneity is a major source of delay and understanding how the characteristics of new traffic will affect capacity is vital to capacity planning. The possibility of operational changes to reduce the demands on infrastructure should always be considered as an option due to the increased flexibility, shorter times to implement and generally lower capital investment needed.

CONCLUSIONS AND FUTURE WORK

The projected, long-term demand for rail freight transportation and expanded rail passenger service on North American railroads will require considerable capital investment in new infrastructure. Investing this capital efficiently requires understanding the different operational characteristics of the intended traffic and consideration of possible operational changes that will enable more efficient use of infrastructure. We performed analyses using dispatch simulation software to determine the impacts of various aspects of train type heterogeneity on freight traffic. The train and operational characteristics studied included: increased headways, reduced train speed heterogeneity, increased power and equalizing dispatching priorities. Each contributed differently to train type heterogeneity and their benefits and additional costs were evaluated.

Future work will include cost benefit analyses and a case study using a actual railroad route and traffic volumes. Additionally, the work to date has only considered a single track route. In future work we intend to investigate the different characteristics that impact train type heterogeneity on a double track line.

ACKNOWLEDGEMENTS

The authors are grateful to Eric Wilson from Berkeley Simulation Software for his assistance in this research. Support for the first author's graduate study has been from a CN Railroad Engineering Research Fellowship at the University of Illinois at Urbana-Champaign.

REFERENCES

1. American Association of State Highway and Transportation Officials (AASHTO) Transportation - Invest in Our Future: America's Freight Challenge, AASHTO, Washington, DC., 2007.
2. Federal Railroad Administration – U.S. Department of Transportation. Vision for High-Speed Rail in America: High-Speed Rail Strategic Plan - The American Recovery and Reinvestment Act, http://www.fra.dot.gov/Downloads/Final%20FRA%20HSR%20 Strat%20Plan.pdf (downloaded May 2009).
3. Lai, Y.C. and C.P.L. Barkan. An Enhanced Parametric Railway Capacity Evaluation Tool (RCET). (In Press) *Transportation Research Record: Journal of the Transportation Research Board*, TRB, National Research Council, Washington, D.C., 2009.
4. Pachl, J. Railroad Operating and Control. VTD Rail Publishing, Mountlake Terrace, 2002.
5. International Union of Railways. UIC leaflet 406, UIC International Union of Railways, France, 2005.
6. Dingler, M.H., Y-C. Lai, and C.P.L Barkan. 2009. Impact of Train Type Heterogeneity on Single-Track Railway Capacity. (In Press) *Transportation Research Record - Journal of the Transportation Research Board,* TRB, National Research Council, Washington, D.C., 2009.
7. Abril, M., Barber F., Ingolotti L., Salido M.A., Tormos P., and Lova A. An Assessment of Railway Capacity, *Transportation Research Part E*, Vol. 44, No. 5, 2008, pp. 774-806.
8. Bronzini M.S. and D.B. Clarke. Estimating Rail Line Capacity and Delay by Computer Simulation, Tribune des Transports, Vol. 2, No. 1, 1985, pp. 5-11.
9. Harrod, S. Rail Capacity and Management Planning, Ph.D. Dissertation, Department of Quantitative Analysis and Operations Management, University of Cincinnati, Cincinnati, OH, 2007.
10. Gorman, M.F. Statistical Estimation of Railroad Congestion Delay, *Transportation Research Part E*, 2008, doi:10.1016/j.tre.2008.08.004
11. White, T. Examination of the use of Delay as a Standard Measurement of Railroad Capacity and Operation, Presented at the 85th Annual Meeting of the Transportation Research Board, Washington, DC., 2006.
12. Washington Group International, Inc. RTC Simulations – LOSSAN North Railroad Capacity and Performance Analysis. LOSSAN Rail Corridor Agency and IBI Group, www.sbcag.org/Meetings/SCSPC/2007/February/Item5LOSSANSR.pdf. Accessed March 15, 2009.
13. Parsons Brinckerhoff Quade & Douglas, Inc. The Long Term Financial Feasibility of the Northwestern Pacific Railroad, Draft Final. 2002. www.northcoastrailroad.org/Acrobat/FinancialFeasibility/NWP%20Appendix%20I.pdf Accessed March 15, 2009.

14. Transportation Research Board. Proceedings from Transportation Research Board Workshop on Railroad Capacity and Corridor Planning, Washington D.C. January 13, 2002.
15. Krueger, H. Parametric Modeling in Rail Capacity Planning, Proceedings of Winter Simulation Conference, Phoenix, AZ. 1999.

LIST OF TABLES AND FIGURES

TABLE 1: Train types used in Simulations

Intermodal	**Unit**
90 cars	115 cars
6,300 ft	6,325 feet
8,100 tons	16,445 tons
2.12 HP/Trailing Ton	0.78 HP/Trailing Ton
4 SD70 4,300 HP Locomotives	3 SD70 4,300 HP Locomotives
Maximum Speed: 70 mph	Maximum Speed: 50 mph

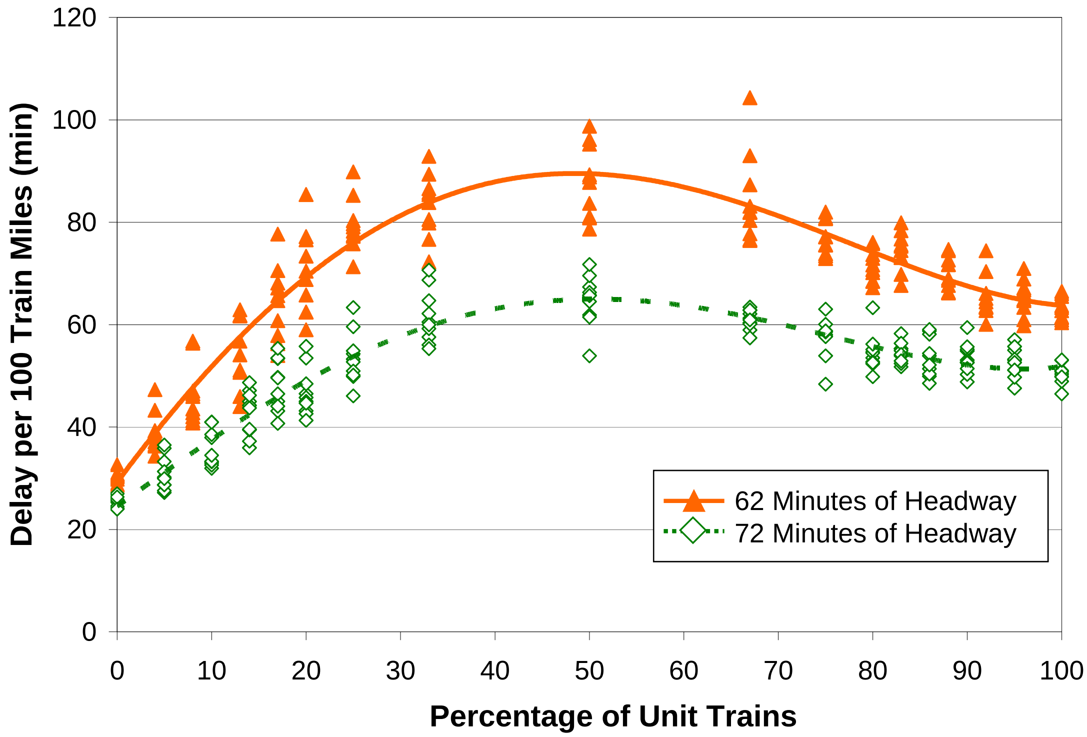

FIGURE 1a: Effect of Train Headway on Average Delay

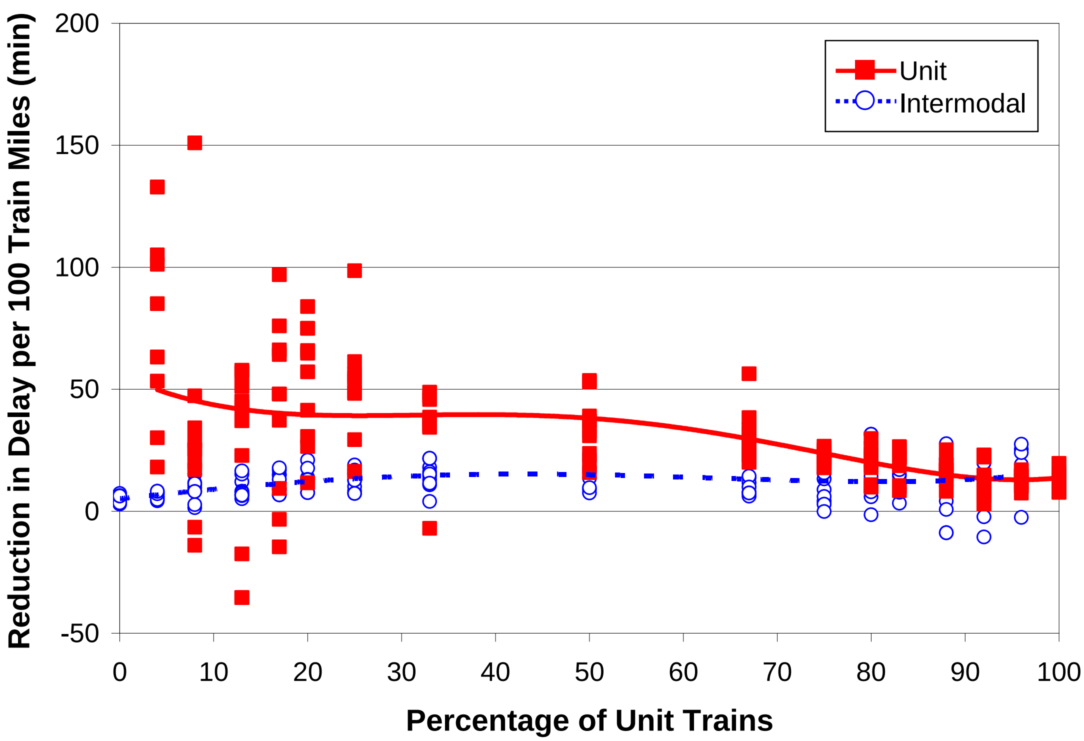

FIGURE 1b: Reduction in Delay for each Train Type due to Increased Headways

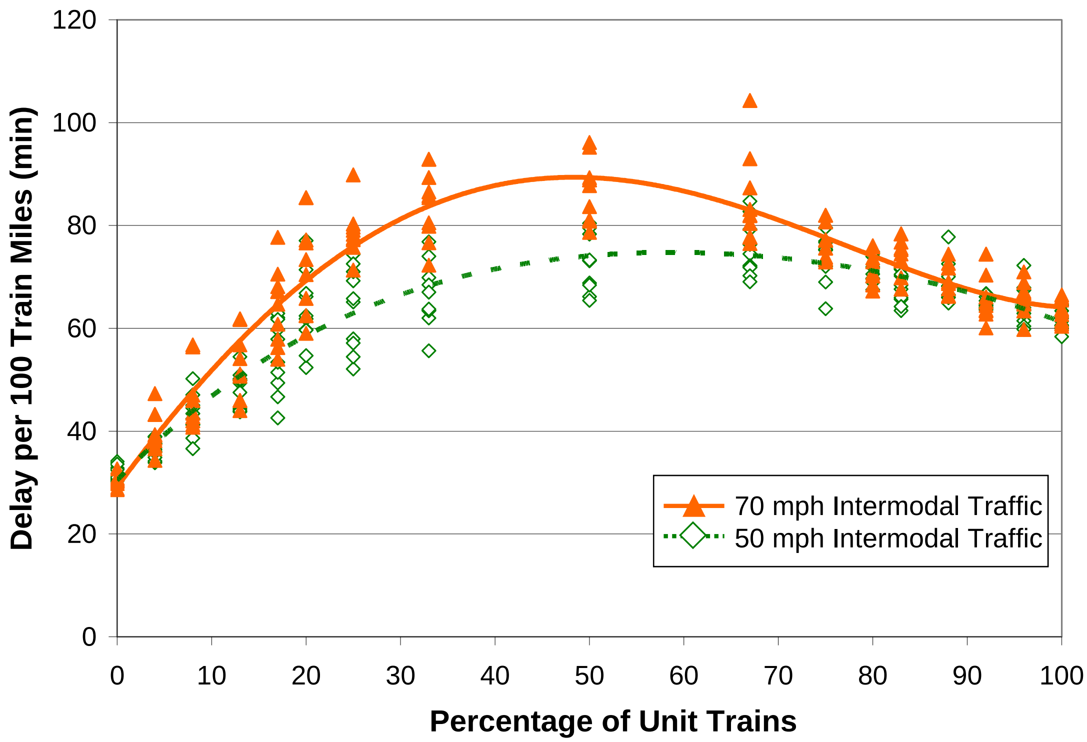

FIGURE 2a: Effect of Intermodal Train Speed on Average Delay

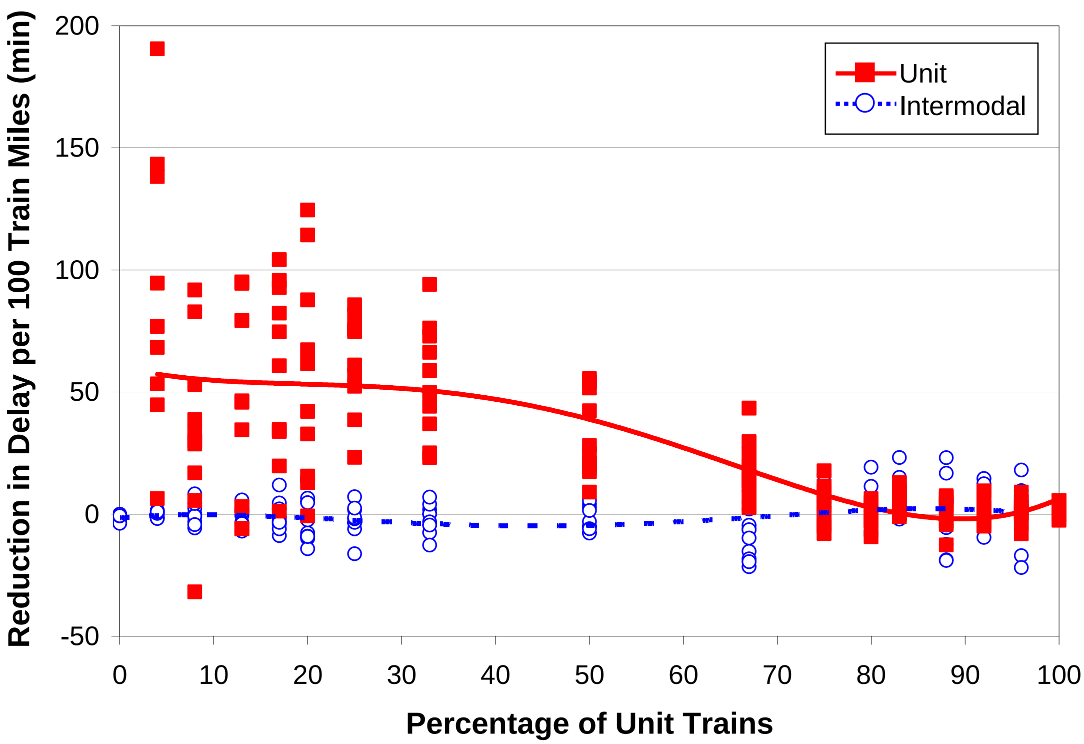

FIGURE 2b: Reduction in Delay for each Train Type due to reduced Intermodal Train Speed

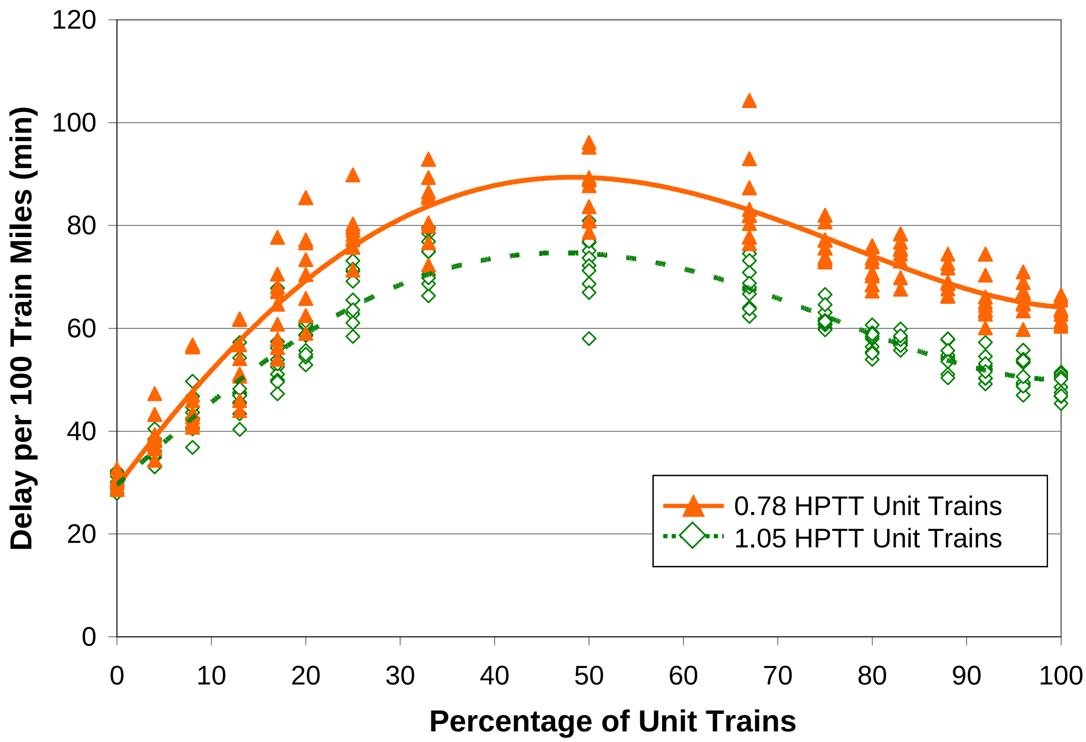

FIGURE 3a: Effect of Unit Train HPTT on Average Delay

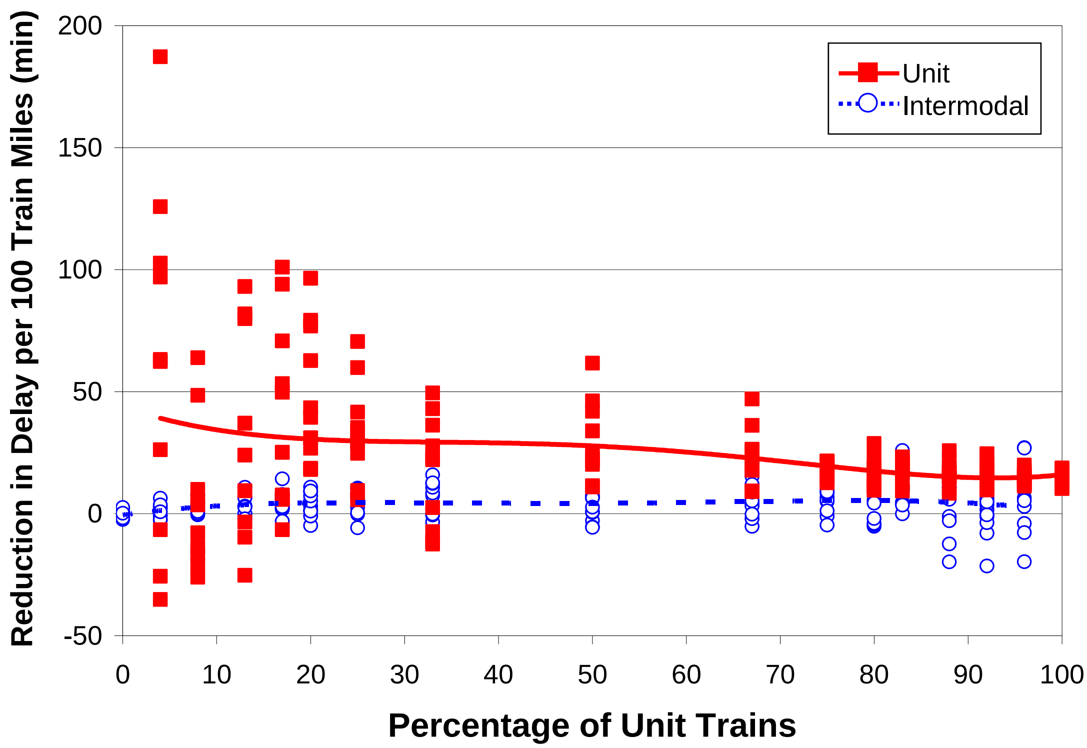

FIGURE 3b: Reduction in Delay for each Train Type due to increased Unit Train HPTT

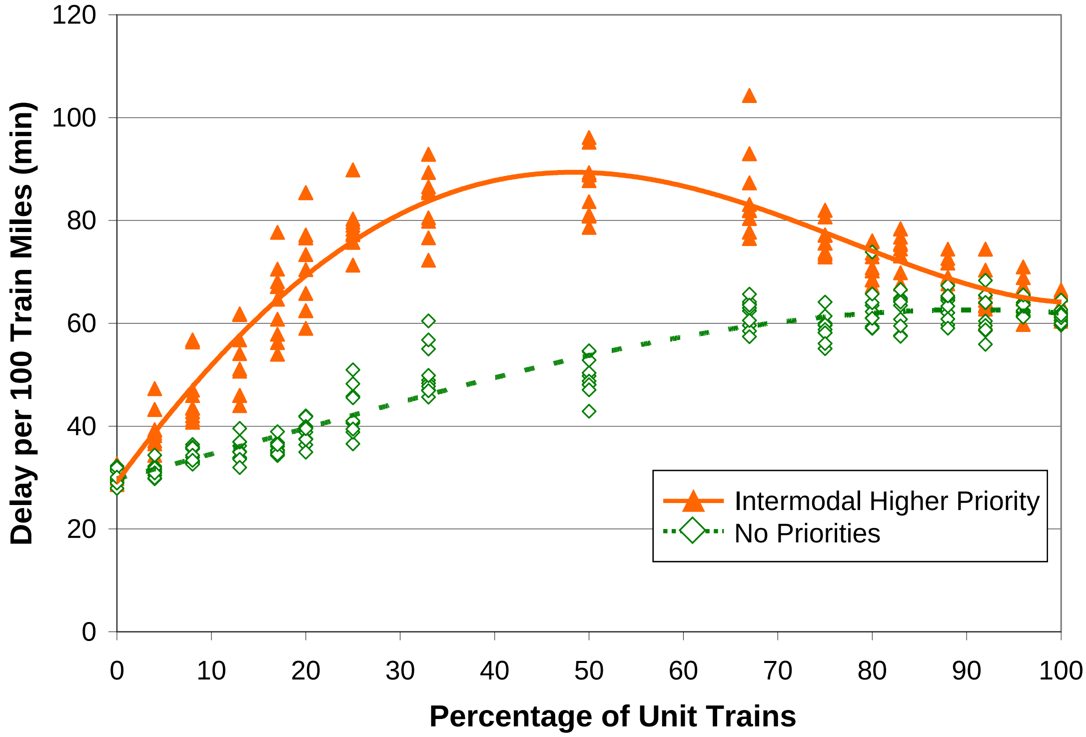

FIGURE 4a: Effect of Train Priorities on Average Delay

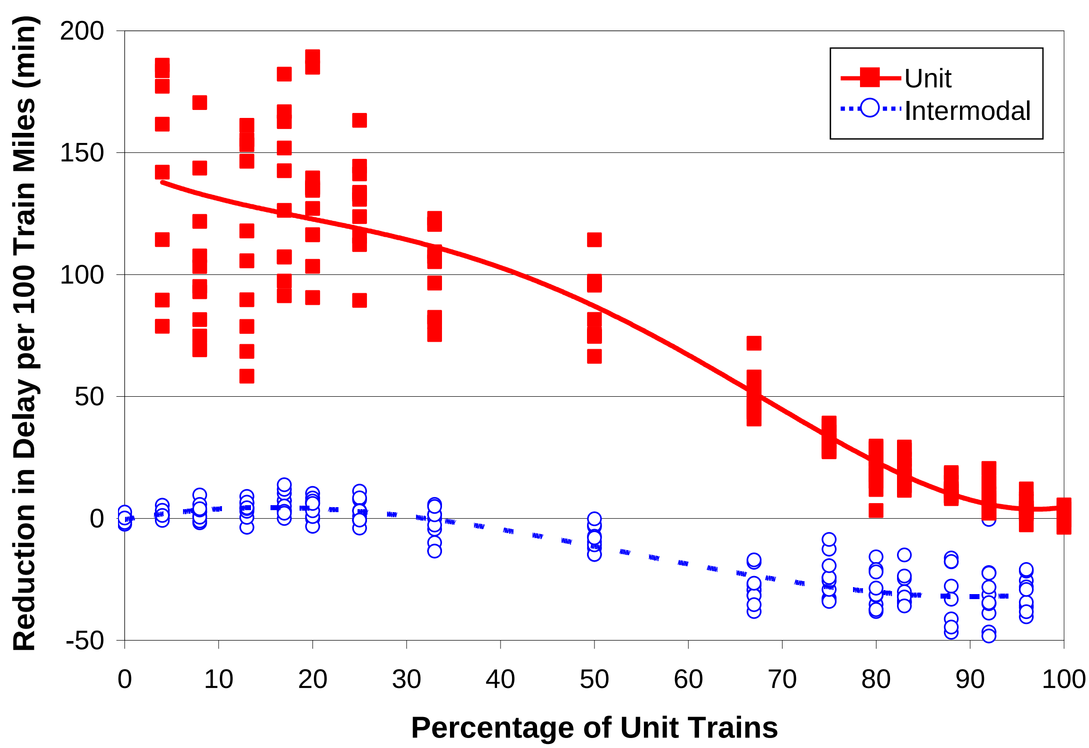

FIGURE 4b: Reduction in Delay for each Train Type due to removal of dispatching priorities