Original Article

Journal of RAIL AND RAPID TRANSIT

Proc IMechE Part F: J Rail and Rapid Transit
227(2) 140–147

DOI: 10.1177/0954409712456832
pif.sagepub.com

Mitigating train-type heterogeneity on a single-track line

Mark H Dingler[1], Yung-Cheng (Rex) Lai[2] and Christopher P L Barkan[3]

Abstract

North American railroads are experiencing growth in traffic demand and increasingly need to expand capacity on their lines; this will require changes in the mode of operation and infrastructure. Infrastructure expansion requires long lead times and is capital-intensive. Alternatively, some additional capacity may be achieved by alteration of the mode of operation which is often less expensive and faster to implement. In this paper dispatch simulation software was used to analyze a model of a single-track signalized line with characteristics typical of a North American railroad subdivision. The objective was to determine the effects of various operational and infrastructure changes in terms of reducing train delay. For each scenario a reduction in the delay was considered to be a benefit of the project, while any increases in delay, additional locomotives, increased fuel consumption, increased track maintenance costs and additional infrastructure were considered the expenses of the project. A cost–benefit analysis was conducted in order to determine which operational or infrastructure changes are the most cost-effective for each scenario.

Keywords

Capacity, heterogeneity, single track, freight traffic

Date received: 1 December 2011; accepted: 27 June 2012

Introduction

The ability of railroads to efficiently move goods is vital to the North American economy. However, freight railroads are expected to face increasing capacity constraints due to projected substantial long-term growth in demand.[1,2] In order to accommodate this new traffic, railroads will have to expand infrastructure and/or change their mode of operation. Infrastructure expansion requires long lead times and is capital-intensive. Alternatively, a less expensive and faster means of creating additional capacity is through changes in the mode of operation.[3] Previous work has shown the impact of train-type heterogeneity on capacity and train delay and identified the key characteristics that cause these delays.[4] While this impact of train-type heterogeneity is well known[5–8], none of the previous research has compared the benefits of reducing this heterogeneity through changing operational characteristics with other ways of improving capacity. In this paper, various methods to reduce train delay and increase capacity are considered for their effectiveness and economic benefit. Simulation software is used to investigate the impact of various operational changes with different traffic mixes and volumes on a hypothetical, single-track signalized rail line.

Methodology

Multiple operational and infrastructure scenarios are considered in order to calculate the effectiveness of various methods of reducing delay. Dispatch simulation software was used to simulate multiple traffic scenarios, with train delay being the primary metric to measure capacity and the cost of train operations. Base simulations and subsequent operational changes were made on a route with 20 mile (32 kilometers) distances between sidings. These changes were then compared with the improvements obtained when the siding spacing is reduced or a second track is added.

[1]CSX Transportation, Jacksonville, FL, USA
[2]Department of Civil Engineering, National Taiwan University, Taiwan
[3]Department of Civil and Environmental Engineering, University of Illinois at Urbana- Champaign, USA

Corresponding author:
Yung-Cheng Lai, Department of Civil Engineering, National Taiwan University, Room 313, Civil Engineering Building, 1 Roosevelt Road, Section 4, Taipei 10617, Taiwan.
Email: yclai@ntu.edu.tw

Representative rail lines

Specific characteristics of individual rail lines are unique and route characteristics influence the study of railroad operations. For this research three hypothetical rail lines were used in order to represent the characteristics of typical North American midwestern mainline subdivisions (Table 1).

Although the attributes are somewhat idealized, they provide a consistent basis for relative comparison of different scenarios of interest under a reasonably realistic set of operating conditions. Track maintenance, service failures and other factors that can affect capacity are not considered and there is no intent to imply that the results presented here predict absolute measurements for a particular set of conditions.

Train types

While each individual train is different, the attributes for each train type were selected to match their average characteristics (Table 2). The numbers of cars and units were obtained from the Cambridge Systematics National Rail Freight Infrastructure Capacity and Investment Study[1] conducted for the Association of American Railroads. Tonnages and lengths were based on averages for each car type. The power-to-ton ratios were based on experience and information from the TRB Workshop on Railroad Capacity and Corridor Planning.[9] As is typical, intermodal trains were assigned the higher priority, as is typical in most railroad operations.

While the chosen attributes approximate actual characteristics, at the most basic level the "intermodal" trains represent freight trains with the highest maximum speeds, power-to-ton ratios and dispatching priorities, while the "bulk" trains represent those with the lowest speeds, power-to-ton ratios and dispatching priorities. Although referred to as "intermodal" and "bulk" for convenience, what is actually of significance in the analyses are their specific operating characteristics, not the particular type of train.

Simulations

Simulations were conducted at different traffic volumes and levels of heterogeneity. Since 40 trains per day is the practical capacity limit for single-track operation, 11 different volumes were simulated ranging from eight to 40 trains per day, simulated in increments of four. The simulations at each volume used a uniform average temporal distribution of trains in each direction over a 24-h period. The results are intended to provide a basis for relative comparison of the effect of various factors of interest; however, inspection, maintenance, local traffic and other factors that affect capacity were not considered. Consequently, the results are more representative of the possible spacing or headway between trains than the actual train volume on a line. Seven different levels of heterogeneity were tested based on the percentage of the different train types. The tests were done with 12.5, 25, 50 and 100% of each train type. The level of heterogeneity corresponds to the percentage of one train type relative to the other. The train sequence was directly proportional to the percentage of each train type; for instance, when the traffic is 12.5% intermodal, each intermodal train would be followed by seven bulk trains. The ratios and traffic patterns

Table 2. Train composition characteristics used in simulations.

Intermodal	Bulk
16 three-pack spine cars	115 loaded hopper cars
Nine five-pack well cars	
5659 ft (1725 m)	6325 ft (1928 m)
5900 tons	16,445 tons
3.64 HPTT	0.78 HPTT
Five 4300 HP Locomotives	Three 4300 HP Locomotives
Maximum speed: 70 miles/h (112 km/h)	Maximum speed: 50 miles/h (80 km/h)

Table 1. Routes used in analysis.

Single track with 20 mile siding spacing	Single track with 10 mile siding spacing	Two tracks with 10 mile crossover spacing
262 miles (422 km) long	262 miles (422 km) long	260 miles (418 km) long
20 miles (32 km) between siding centers	10 miles (16 km) between siding centers	10 miles (16 km) between siding centers
8700 ft (2651 m) signaled sidings with 24 powered turnouts	8700 ft (2651 m) signaled sidings with 24 powered turnouts	Universal crossovers with 24 powered turnouts
2.55 mile (4.02 km) signal spacing	2.75 mile (4.43 km) signal spacing	2.75 mile (4.43 km) signal spacing
Two-block, three-aspect signaling	two-block, three-aspect signaling	two-block, three-aspect signaling
0% grade and curvature	0% grade and curvature	0% grade and curvature

were the same for trains traveling in both directions. For each configuration a series of 25 simulations was performed with the exact departure time of each train randomized according to a uniform distribution over a 30-min interval, 15 min before or after the scheduled time for that train.

Train delay, equipment and infrastructure costs

Train delay cost

Train delays affect the railroads through additional expenses and lost revenue. Each hour a train is delayed represents at least some degree of lost opportunity to transport cargo and increase revenue. If a railroad can reduce train delay on its network, average train velocities will increase, resulting in potentially significant savings for the railroad through improved utilization of railcars, locomotives and crews. Schafer and Barkan[10] calculated train delay using four components:

- car/equipment cost;
- unproductive locomotive cost;
- idling fuel cost;
- crew cost.

However, these factors do not include costs incurred due to delays to the lading. When a railroad's cars and locomotives are fully utilized, each delay causes the cycle time of a train to increase, and the potential revenue from additional shipments is lost. The shipment will be moved by a different railroad or mode of transportation. Fuel costs are considered separately and are not considered in this train delay calculation. For the current analysis, the train type specific delay cost has four components:

- car cost;
- locomotive cost;
- crew costs;
- lading costs.

Car and locomotive delay costs were estimated using a time-based metric for the cost of their ownership. Car delay cost was calculated using the average car hire cost and locomotive costs were calculated from its depreciation over its economic life. Labor costs were calculated using average straight time and fringe benefit costs for a two-man crew. Lastly, the shipment delay cost was calculated based on the average revenue during the cycle of a train. The average revenue was calculated using the revenue per car, in other words the typical commodities carried in each train type divided by the ratio of total-to- loaded miles. This value was divided by the average cycle time and the availability of the cars to obtain the time-based value of the shipment.

Summing the four components yielded a total delay cost of $1392 for intermodal and $586 for bulk (Table 3). Although this provides a good estimate it does not account for all the costs of delays. Extra crew costs due to hours of service limitations, fuel costs or possible additional maintenance costs were not considered in this calculation. The lading delay is an opportunity cost, based on the assumption that if a train is delayed, the ability to move more cargo is lost. This assumes full utilization of cars and locomotives and that any excess equipment could be used to move more cargo. If equipment is not fully utilized, then the lading delay cost should be reduced or omitted from the delay costs. The opportunity cost of a shipment is better represented by the profit of a shipment and not the revenue. However, since this data is unavailable, revenue will be used to determine the shipment delay cost.

Infrastructure and equipment costs

Using information obtained from multiple sources, estimates for the cost to build a siding or a second track were calculated. The calculation included the cost of the civil engineering work, track, signals, design and additional fees. It was assumed that no additional right-of-way acquisition was required, although in some circumstances this may be necessary The construction cost for a new 8750 ft (2667 m) siding was estimated to be $6,500,000 and for a new main track $2,750,000 per mile (about $1,709,000 per kilometer). Additional infrastructure implies increased maintenance costs. Zarembski et al.[11] provide estimates on the maintenance costs for various tonnages and track classes. Using this information the maintenance costs were estimated to be $50,000 per mile (about $31,000 per kilometer) of siding and $70,000 per mile (about $43,000 per kilometer) of mainline track.

The purchase cost of a new locomotive varies by type and extra features included. For this analysis the additional cost of a new locomotive was assumed to be $1,750,000.[12,13] Fuel consumption calculations from the dispatch simulation software and a cost of $3.13 per gallon ($0.83 per liter) of fuel were used.[14] Locomotive maintenance costs were not considered due to lack of data.

Table 3. Train delay costs per train-hour.

	Intermodal	Bulk
Freight car ownership cost ($)	84.90	57.54
Locomotive ownership cost ($)	87.52	52.51
Crew cost ($)	66.64	66.64
Shipment delay cost ($)	1153.38	409.67
Total train delay per hour($)	1392.43	586.36

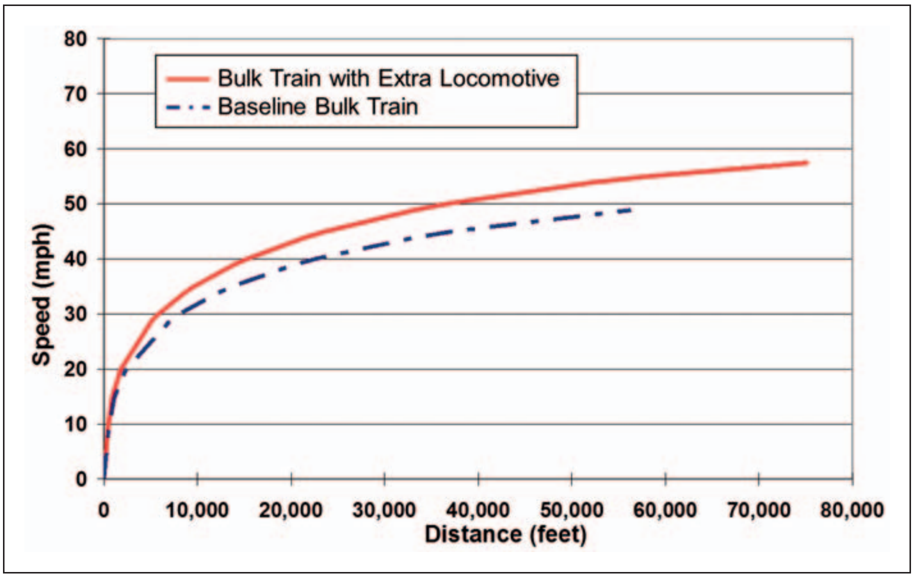

Figure 1. Acceleration curve of baseline bulk trains and with an additional locomotive with the maximum speed increased to 60 mile/h.

Analysis of net present value of infrastructure and operational changes

Previous work has shown that on a single-track route, meets are the principal source of delay and user of capacity; therefore, reducing the number and duration of meets will have the largest effect on capacity.[15] Several methods are possible to reduce delays that are a result of meets. Increased power (e.g. using additional locomotives) increases the acceleration rate thereby reducing the time to leave a siding after a meet, increased train speeds reduce the dwell time of a train waiting in a siding to meet other trains, and removing dispatching priorities reduces the number of meets. Each method may provide benefits due to reduced delays and increased capacity. Infrastructure expansion will also reduce delay time and improve operations. If sidings are more closely spaced, the time trains must wait for oncoming traffic is reduced, and adding a second track totally eliminates meets. Although infrastructure expansion is an effective method of increasing capacity, it is also expensive. One of the objectives of the analyses described here was to develop a better understanding of how operational changes compare to infrastructure in terms of capacity. Both operational and infrastructure changes were simulated and compared to the base scenario with 20 miles (32 kilometers) between sidings.

In each of the scenarios considered, the reduction in delay is due to an operational change or an investment that incurs some additional expenditure. Deciding which, if any, of these approaches is most appropriate requires understanding both the costs and benefits. Consequently, this research developed a framework to consider each of the scenarios. For each scenario, the reduction in delay was considered to be the benefit of the project, while any increases in delay, additional locomotives, increased fuel consumption, increased track maintenance costs and additional infrastructure were considered to be expenses. These benefits and expenses were used to calculate the net present value (NPV) of each alternative per 100 miles (161 kilometers) of track using a study period of 10 years and a discount rate of 7%.

Additional locomotive for each bulk train

The first alternative considered for its potential to reduce delays is to add an extra locomotive to each bulk train. The additional locomotive increases the horsepower-to-trailing-ton (HPTT) ratio for the bulk trains from 0.78 to 1.05, thereby reducing acceleration distance (Figure 1). Previous work has shown that acceleration distance has a significant impact on train delay for heterogeneous traffic.[16] Faster acceleration reduces the time that trains occupy the mainline while traveling below normal maximum speed. This reduces the time lost in meets, thereby reducing overall run time. However, additional locomotives require major capital investment as well as additional maintenance and fuel costs.

Adding locomotives to each bulk train is not a cost-effective solution to improve operations (Figure 2). Although the reduction in delay increases with a higher number of bulk trains, the cost of additional locomotives outweighs the benefits of the reduction in delay. This results in a zero or negative NPV for all traffic volumes and mixes, with losses increasing as the percentage of bulk trains increases.

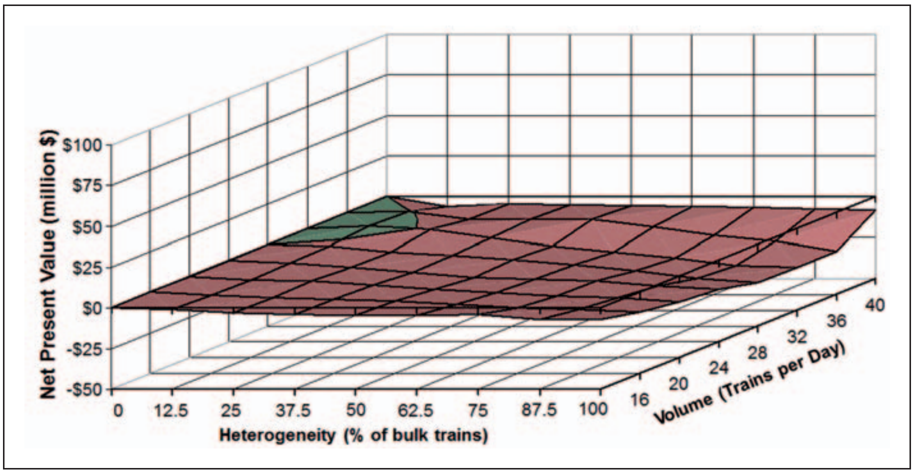

Figure 2. NPV of additional locomotives for each bulk train.

Additional locomotive for each bulk train and 10 mile/h (16 km/h) speed increase

Additional locomotives not only improve acceleration performance but also permit trains to reach higher top speeds. Consequently, the effect of a 10 mile/h (16 km/h) speed increase for bulk trains was tested. Increasing the speed of slower-moving bulk trains reduces conflicts due to the speed difference between train types. More importantly, increased speeds reduce the time to travel between sidings, thereby improving capacity. However, an increase in train speed increases train stopping distance and track maintenance costs.

Adding additional locomotives in order to increase the speed of the bulk trains has a negative NPV for almost all volumes and traffic compositions (Figure 3). As the percentage of bulk trains increases the NPV decreases. As more bulk trains are added the costs of locomotives and fuel increase and the resultant reduction in delay does not offset these additional costs.

Equalization of priorities

Previous work identified priority as a significant factor influencing delay on a single-track rail line.[4] Consequently, manipulating the equalizing priority was analyzed as a potential operational method to reduce delay. In typical operations, the lower-priority train will stop and wait at an earlier siding in order to prevent delays to the higher-priority train. Consequently, heterogeneous priorities increase both the number and duration of meets. When traffic has the same priority during a meet the first train to arrive at a siding will enter the siding and wait for the oncoming train. In the preceding examples the higher-priority trains also had a higher maximum speed, but in this analysis the faster trains have the same priority as the slower ones. Therefore, when a faster train wants to overtake a slower train it will not be able to do so unless the preceding train has already stopped in a siding in preparation for the meet.

One of the advantages of equalizing priorities is the limited additional cost, since no additional infrastructure or equipment is required. However, equalization of priorities increases delays to the train type that previously received the higher priority. Thus, there is a trade-off between greater delays to the trains whose customers expect higher service quality versus reduction in delays to trains with less-demanding schedule requirements.

Removal of heterogeneity in priorities is cost-effective only when intermodal trains are a majority at the highest traffic volumes (Figure 4). When traffic is homogenous the delays are unaffected and therefore there are no additional costs or benefits of removing priorities. Removing the priorities reduces the delays to bulk trains but increases the delays to intermodal trains. This trade-off is most beneficial when the traffic volumes are high with mostly intermodal trains. However, if the traffic consists solely of intermodal trains then equalizing priorities has no effect.

Additional sidings

Reducing siding spacing on a single-track line increases capacity because it allows there to be more meets and passes.[17] In order to understand the benefit of adding sidings, the distance between sidings was reduced from 20 to 10 miles (32 to 16 kilometers). This permits a comparison of the relative benefits of operational changes to infrastructure expansion. However, a closer siding spacing also reduces delay because it allows for meets or passes to be resolved at more ideal locations, reducing the time trains must wait in sidings. The complexity of meets is also reduced because it is possible for a train to advance to a siding instead of waiting for another train to meet or pass at the current siding.

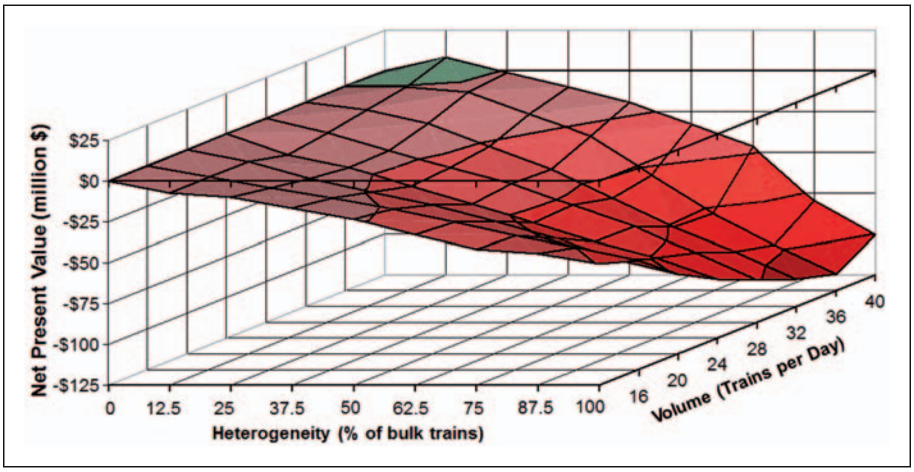

Figure 3. NPV of additional locomotives for each bulk train and a 10 mile/h speed increase.

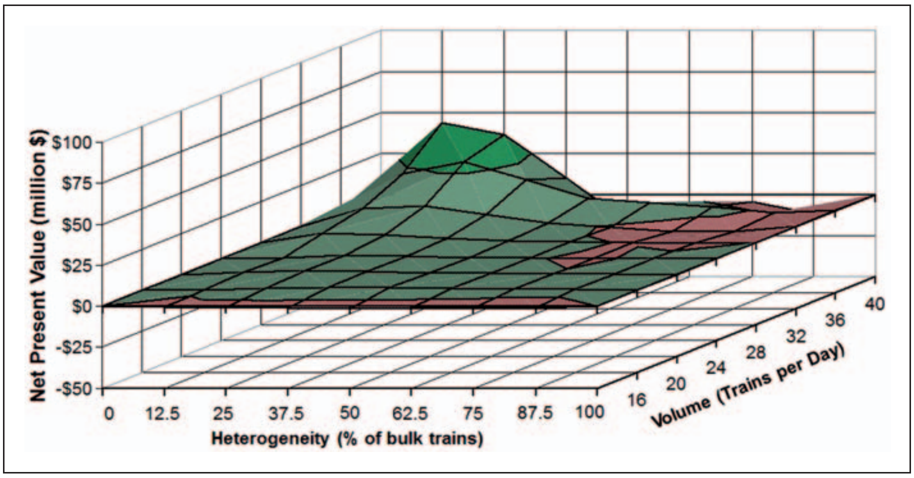

Figure 4. NPV when the priority are equalized for all traffic.

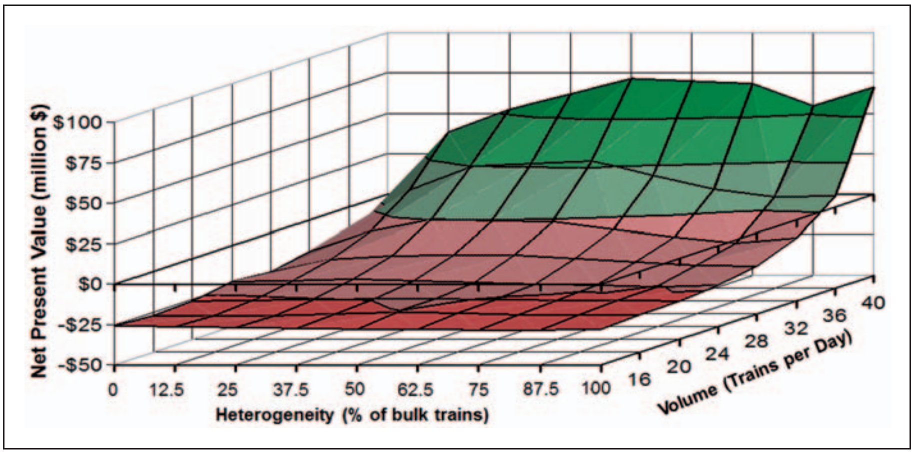

Figure 5. NPV when sidings are added.

The cost-effectiveness of adding sidings is directly related to traffic volume (Figure 5). Below 32 trains per day the NPV for all traffic mixes is negative, however, it becomes positive when the traffic volume increases above 36 trains per day. However, even at the highest volumes the NPV is negative when the traffic is all intermodal since homogenous intermodal traffic has a much higher capacity than heterogeneous or homogenous bulk train traffic.

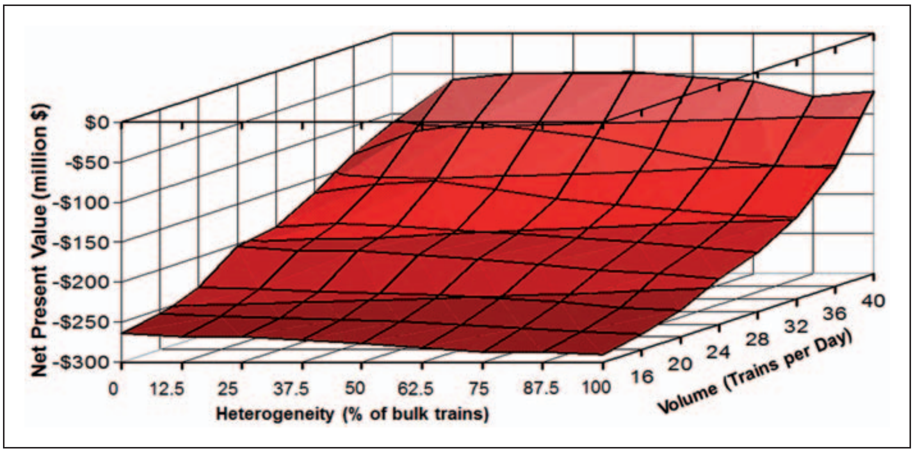

Figure 6. NPV when a second track is constructed.

Table 4. Best alternative to reduce train delay.

Percentage of bulk trains	Volume 12	16	20	24	28	32	36	40
0	–	–	–	–	–	–	–	–
12.5	–	=P	=P	=P	=P	=P	=P	=P
25	–	=P	=P	=P	=P	=P	=P	+SD
50	–	=P	=P	=P	=P	=P	+SD	+SD
75	–	=P	=P	–	–	–	+SD	+SD
87.5	–	=P	=P	=P	–	–	+SD	+SD
100	–	–	–	–	–	–	+SD	+SD

=P equalizing priorities.
+SD adding sidings.

Adding a second track

The second infrastructure change considered was the construction of a second main track. Operations are vastly different with multiple tracks than with a single track. With multiple tracks trains can be separated by direction, referred to as directional running, eliminating meets. It also reduces the impact of heterogeneous train speeds. If there is sufficient capacity, and a suitable traffic control system, a faster train can easily crossover to the other track to pass a slower train and then return to the original track at a later crossover point. This reduces delays to both the faster and slower trains. However, a second track requires significant construction and maintenance costs and crossovers require twice as many turnouts compared to passing sidings, further increasing capital and maintenance costs.

At the traffic levels considered, adding a second track is not a cost-effective way to reduce delay (Figure 6). While adding a second track nearly eliminates delay, the cost to build and maintain that track is substantial. The NPV increases with volume; however, even at the highest volumes the reduction in delay does not justify the expenditure required for a second track. However, at higher volumes, which are possible as incremental projects are completed, the NPV of constructing a second track will increase. Additionally a second track offers other benefits not taken into account in this analysis. For example, with two tracks, regular track maintenance will have a lower disruption on traffic and also pickups and set downs by local traffic will not block all traffic.

Best delay reduction strategy

Each volume and level of heterogeneity has a specific operational or infrastructure change that provides the best economic return. The NPV for each scenario was considered at various volumes and traffic mixes. The alternative with the best NPV in each condition is listed in Table 4. This study found that for all traffic mix and volume combinations one of three alternatives was best: no change, equalizing priorities or adding sidings. When the traffic was all intermodal there was sufficient capacity and therefore no changes were cost-justified. At moderate to high volumes, equalizing priorities was beneficial with heterogeneous traffic. At the highest volumes, and when bulk was a majority of the traffic, sidings were the alternative with the best NPV.

Discussion and conclusions

With increasing demand for freight rail services, railroads must evaluate the most economic methods to reduce train delay and increase capacity. Depending on the volume and specific traffic mix the best alternative may be infrastructure expansion, changes in the mode of operation or some combination.

Operational changes are advantageous because they can be implemented more rapidly, are more flexible than infrastructure changes and may be less capital-intensive. Such changes enable a railroad to respond to changing traffic levels and patterns, provide relief during short periods of high traffic volumes or serve as an interim measure while additional infrastructure is built.

On a route without bottlenecks, one of the most effective operational changes is equalizing priorities. This is a rapid and flexible method that can be used to improve operations and the only cost is the additional delays to higher-valued traffic. During periods of high traffic levels, a dispatcher can choose to utilize equal priorities as an operational strategy, however, this can be difficult as the higher-priority traffic often has a set arrival time and slowing down a higher-priority train can make meeting the schedule difficult. Each railroad should determine routes that make sense to equalize priorities and adjust schedules accordingly.

Other operational changes that have the potential to improve capacity but are not considered in this paper include increasing maintenance to reduce slow orders and service failures, modifying local service to reduce mainline delay, adjusting yard operations to reduce switching on the mainline and reducing crew change times to increase throughput in terminal areas.

Funding

This research was supported by the Railroad Engineering Program at the University of Illinois at Urbana-Champaign. The first author was supported in part, by a CN Research Fellowship at the University of Illinois at Urbana-Champaign.

References

1. Cambridge Systematics. *National rail freight infrastructure capacity and investment study*. Cambridge, MA: Cambridge Systematics, 2007.
2. American Association of State Highway and Transportation Officials. *Transportation - invest in our future: America's freight challenge*. Washington, DC: American Association of State Highway and Transportation Officials, 2007.
3. Lai YC and Barkan CPL. An enhanced parametric railway capacity evaluation tool (RCET). *Transp Res Record: J Transp Res Board* 2009; 2117: 33–40.
4. Dingler MH, Lai Y-C and Barkan CPL. Impact of train type heterogeneity on single-track railway capacity. *Transp Res Record: J Transp Res Board* 2009; 2117: 41–49.
5. Bronzini MS and Clarke DB. Estimating rail line capacity and delay by computer simulation. *Trib Transp* 1985; 2(1): 5–11.
6. Pachl J. *Railroad operating and control*. Mountlake Terrace, WA: VTD Rail Publishing, 2002.
7. UIC Code 406: 2004. Capacity.
8. Abril M, Barber F, Ingolotti L, et al. An assessment of railway capacity. *Transp Res E, Logistics Transp* 2008; 44(5): 774–806.
9. Transportation Research Board. *Proceedings from transportation research board workshop on railroad capacity and corridor planning*. Washington, DC: Transportation Research Board, 2002.
10. Schafer D and Barkan CPL. A prediction model for broken rails and an analysis of their economic impact, In: *The AREMA annual conference*, Salt Lake City, UT, 21–24 September 2008. Available at: http://www.arema.org/proceedings/index.aspx.
11. Zarembski AM, Resor R and Cikota J. Technical monograph: estimating maintenance costs for mixed high-speed passenger and freight rail corridors. Unpublished report, FRA, Washington, DC, 2004.
12. Murray T. How much does it cost? *Trains Mag* 2008; 67(1): 34–43.
13. Railway Age. Locomotive leasing: what's power worth today? *Railway Age,* June 2008, pp. 40–43.
14. Association of American Railroads. *Analysis of class I railroads*. Washington, DC: Association of American Railroads, 2008.
15. Dingler MH, Koenig A, Sogin S and Barkan CPL. Determining the causes of train delay. In: *The AREMA annual conference*, Orlando, FL, 29 August–1 September, 2010. Available at: http://www.arema.org/proceedings/index.aspx.
16. Dingler MH. *The impact of operational strategies and new technologies on railroad capacity*. Masters Thesis, University of Illinois, Urbana, IL, 2010.
17. Krueger H. Parametric modeling in rail capacity planning. In: Farrington PA (ed.) *The winter simulation conference*, Phoenix, AZ, 5–8 December 1999, pp. 1194–1200. The University of Michigan: IEEE.