A Decision Support Screening Approach for Railway Infrastructure Capacity Planning on Single-track Lines

Mei-Cheng Shih[a], C. Tyler Dick[a,1], Christopher P.L. Barkan[a,2]
[a] Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign
Newmark Civil Engineering Laboratory, MC-250
205 North Mathews Ave, Urbana, IL 61801-2352
E-mail: mshih2@illinois.edu,
[1] E-mail: ctdick@illinois.edu,
[2] E-mail: cbarkan@illinois.edu

Abstract
Long-term growth in demand for freight rail transportation and expansion of passenger train operations on freight corridors is predicted to cause congestion on the North American rail network. Recent changes in rail traffic patterns and commodity flows have increased the rail traffic volume and train-type heterogeneity on the primarily single-track railway network. Under these circumstances, capital investments to expand track infrastructure capacity will be necessary to maintain the desired level of service for trains. However, the flexible (unscheduled) operating style adopted by North American freight railroads makes it difficult to identify an optimal infrastructure investment strategy without relying on detailed simulation models. To reduce reliance on simulation, this research developed an analytical decision support approach to identify and screen infrastructure expansion projects on single-track lines under flexible operation and heterogeneous traffic. Based on an initial train plan and given levels of train departure and trip time flexibility, the developed approach calculates the expected conflict density in each defined zone of the study corridor. The conflict densities are used to prioritize the order of zones along the route for infrastructure investment projects. The final output of the decision support approach is an ordered list of potential infrastructure projects along the single-track study corridor. In a case study, a capital investment strategy for a hypothetical single-track mainline is developed with the capacity screening approach and compared to a simulation-based zonal delay method to demonstrate the applicability of the developed approach to practical scenarios.

Keywords
Rail Capacity Planning ; Analytical Model ; Monte Carlo Process

1 Introduction

Recent traffic trends on the North American rail network include a general long-term growth of freight rail traffic volume. There also continues to be interest in expanding intercity passenger and commuter rail services in North America, including increased train service frequency and speeds on existing freight corridors (Bing et al., 2010, Martland, 2010). With its different operating speed and level-of-service requirements, the introduction of passenger service on a freight corridor adds extra traffic heterogeneity to the rail operations, decreasing the available time and space for operation of freight trains

(Sogin et al., 2013; Shih et al., 2015) and overall mainline capacity (Shih et al., 2015). These trends are compounded by the limited capacity of the single-track lines that comprise most corridors in the North American rail network. Additional passing sidings (passing loops) and sections of second main track are required to increase the capacity of these corridors.

With most of the North American track infrastructure owned by private for-profit freight railroads, infrastructure expansion projects to add capacity must be selected to maximize return on investment subject to capital budget constraints (Cambridge Systematics, 2007). In the field of transportation, models for efficient terminal layout, network design, or capacity planning are usually based on optimization techniques (Lai & Shih, 2013; Uchida et al., 2015; Sun, & Schonfeld, 2016). Outside of North America, or where train operations adhere to pre-planned schedules, various optimization techniques are used to directly solve for the best infrastructure expansion projects (Higgins et al., 1997; Shih et al., 2014). However, since North American train schedules are flexible and not planned in advance, optimization approaches that evaluate a fixed timetable are not appropriate (Pouryousef et al., 2015). Typically, to account for the lack of a pre-planned timetable, infrastructure expansion projects are prioritized and selected with the aid of detailed simulation models. Each possible infrastructure project is simulated under different combinations of rail traffic to assess its performance in reducing train delay to acceptable levels. The process can be tedious as it takes time to construct a separate detailed model of the route for each combination of possible infrastructure projects and then each of these models must be run through a computationally-intensive and time-consuming simulation. The effort required to undertake this process can limit the number of scenarios a planning group can study, potentially leading to sub-optimal decisions.

To reduce reliance on detailed simulation and consider more project options while better utilizing railway planning resources, there is a need for a tool capable of quickly evaluating and screening potential infrastructure projects to increase railway capacity. To meet this need, this research developed an analytical decision support approach to identify and screen infrastructure expansion projects on single-track lines under flexible operation and heterogeneous traffic.

2 Background

2.1 Rail Traffic Heterogeneity

Dingler et al. (2009) defined the differences between the speed, weight, length, acceleration and braking characteristics of trains that serve the domestic intermodal, bulk freight and passenger rail markets as "traffic heterogeneity". They also used Rail Traffic Controller (RTC) simulation to show how heterogeneous rail traffic consumes more capacity than homogeneous operations with a single train type, resulting in a lower level of service. Other studies also support this conclusion (Vromans et al., 2004; Landex, 2008; Dingler et al., 2013).

In addition to differences in speed, priority and other train characteristics, introduction of passenger and other premium freight services can lead to differing schedule and level-of-service requirements for each type of train operating on a shared corridor. To meet these requirements for the limiting train type, railroads may adopt different operating styles that limit flexibility through adherence to stricter train schedules. However, the previous studies mentioned above only considered the impact of priority and speed variation; they did not consider these effects in combination with changes in the operating style as quantified by train operating schedule and level-of-service requirements. To adequately consider shared-corridor operations, a capacity screening tool must consider the combined impact of operating style, priority and speed variation on train performance and line capacity (Figure 1). A more comprehensive definition of operating style and schedule flexibility will be introduced in the next paragraphs.

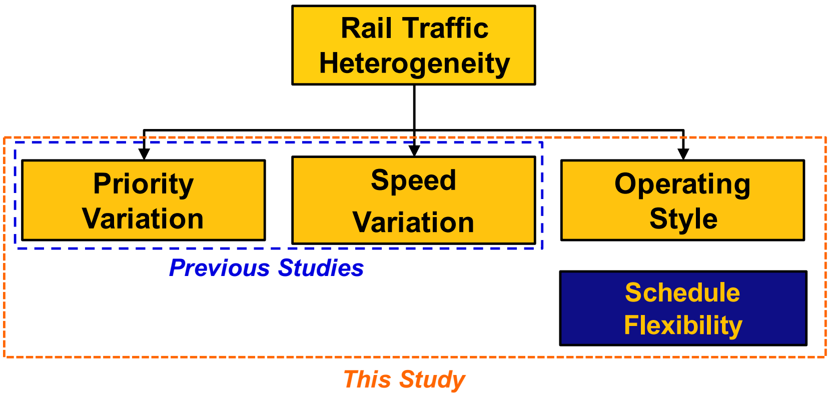

Figure 1: Factors considered in the previous study of rail traffic heterogeneity and in this study

2.2 Schedule Flexibility and Operating style

In this study, the schedule flexibility of a train is defined by its departure time and trip time flexibility (Figure 2). For a given train, departure time flexibility is defined as the potential time period for its departure from an initial terminal or the end of a route segment under study. Trip time flexibility is the distribution of train travel times to reach its final terminal or opposite end of the route segment under study. For each departure time in the period of departure flexibility, there is a corresponding value of trip time flexibility (blue area in Figure 2 is an example). Combined, departure time flexibility and trip time flexibility describe schedule flexibility, or the theoretical space formed by all of the potential train paths for a given train (yellow area in Figure 2). Increasing either departure time flexibility or trip time flexibility will also increase schedule flexibility. Since higher schedule flexibility leads to greater uncertainty in train arrival time, it is inversely related to the provided Level of Service (LOS) (Figure 3).

Operating style refers to the variation in schedule flexibility observed across the individual trains operating on a mainline during a given period. Rail systems adopt different operating styles according to their business needs and objectives (Figure 4). For freight railroads in North America, the business objectives of carload freight service require field personnel to dynamically adjust predefined train plans. This flexibility is often used to address random disruptions caused by events such as mechanical failures, signal failures, or temporary slow orders. This operating style was named “improvised operation” by Martland (2010) and is termed “flexible operation” in this study. The opposite operating style, where the operators aim to follow planned train paths, meet locations, dwells and routes exactly is termed “structured operation” (Martland, 2010). In general, passenger and transit systems follow this more structured operating style. Under structured operations, field personnel have little flexibility; even their responses to schedule disruptions are usually prescribed by some emergency handling procedures or a pre-set rescheduled timetable (Norio et al., 2005; Luethi et al., 2007). The operating style adopted by shared corridors in the US is a mixture of these two operating styles. On these routes, passenger trains follow structured operations and freight service follows flexible operations while sharing the same track infrastructure. This study refers to the operating style on these corridors as "mixed operation".

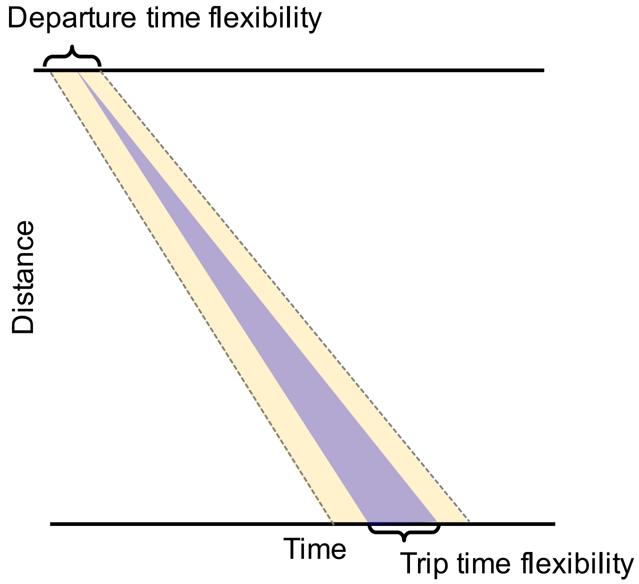

Figure 2: Departure and trip time flexibility

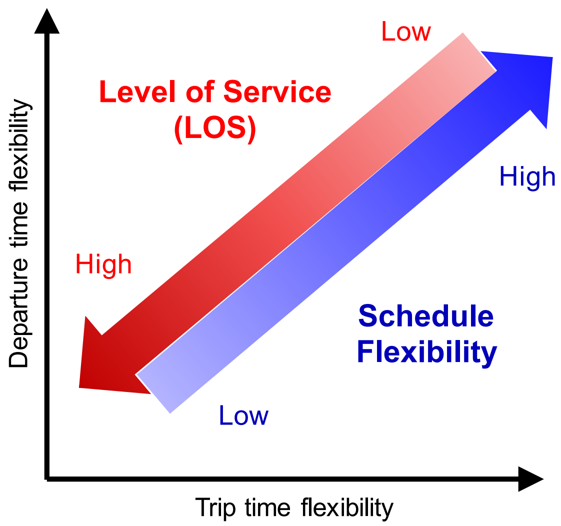

Figure 3: Relationship between departure and trip time flexibility, schedule flexibility, and Level of Service (LOS)

Operating style impacts train delay and line capacity. Each of the three time-distance diagrams in Figure 4 contains four train paths under a different operating style. The schedule flexibilities of each train follow the characteristics of the corresponding operating style. Focusing on the traffic conflicts encountered by the blue train path or band, the black dot, area, and line in the diagrams represent the different number and locations of potential rail traffic conflicts encountered by that train. Based on single-track train operation mechanisms, the number and location of conflicts affects the appropriate locations for capacity expansion projects on single-track lines. A different value of line capacity can be defined for the same single-track line due to variations in the number and location of rail traffic conflicts created by each operating style.

The analytical capacity project screening approach developed in this study is based on the concept of rail traffic conflict analysis, which is also a component of the Root Cause Analysis approach proposed by White (2005). In adopting this approach, the screening tool tracks the distribution of potential rail traffic conflicts along the route under study given the overall operating style and variation in schedule flexibility, speed and priority of different trains. In this manner, the proposed screening tool can consider traffic heterogeneity, schedule flexibility and operating style in suggesting the locations of potential capacity expansion projects without the need for detailed simulations or a pre-planned timetable.

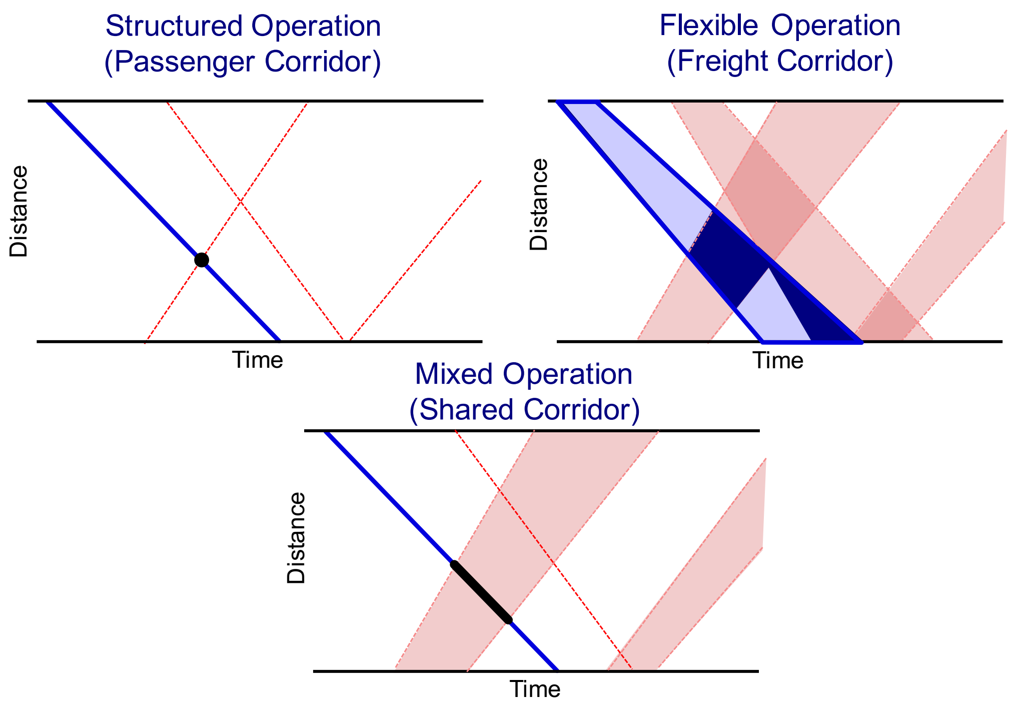

Figure 4: Examples of different railway operating styles

3 Capacity Screening Approach and Zonal Delay Approach

In the typical simulation approach to evaluating the performance of different infrastructure expansion projects, dispatching logic within the simulation resolves train conflicts by holding trains at passing sidings or on sections of double track until opposing trains have cleared the single-track bottleneck section (Figure 5a). When a train is held in this manner, it accumulates train delay that is tracked by the simulation model and reported as output. When a planner examines the simulation output, they will look at single-track sections adjacent to locations where large amount of train delays accumulate as potential locations for capacity expansion projects. Long routes will often be divided into zones and the zones with the largest cumulative train delay will become the focus of infrastructure expansion efforts. A potential drawback to this approach is that decisions made on the basis of reducing the train delay symptom of congestion may not adequately address the actual train conflicts that lead to the congestion and delays.

An alternative is the Root Cause Analysis proposed by White (2005), where the locations of unresolved rail traffic conflicts are used as a direct indicator of congestion and potential delays (Figure 5b). Unlike simulation, in this approach, train conflicts are not resolved and there is no train delay metric. Instead, the natural train meeting points are determined based on the departure time and operating speed of each train. These meeting points, or train conflict locations, may fall at passing sidings or in the middle of single-track bottleneck sections. A cluster of a large number of train conflicts directly indicates a location where additional parallel track infrastructure is needed to resolve the train conflicts. If the infrastructure does not already exist where the cluster occurs, planners should consider that location for a capacity expansion project. This method for identifying capacity constraints is used as the basic concept for the developed capacity screening approach.

Because the Root Cause Analysis approach examines the cause of a conflict while the simulation approach examines the symptom of train delay, they may lead to different infrastructure solutions. Even in a simple example with two train and one meet, the conflict point and location where train delay is accumulated are different (Figure 5). While the delay-based simulation approach suggests the single-track section adjacent to the siding is in general the capacity constraint, the root cause analysis identifies the specific conflict point as the exact location of constrained capacity within this section. While this trivial example essential yields the same result, for more complex scenarios with many more train conflicts and additional existing infrastructure, the capacity constraints identified by the each of the two methods may be quite different.

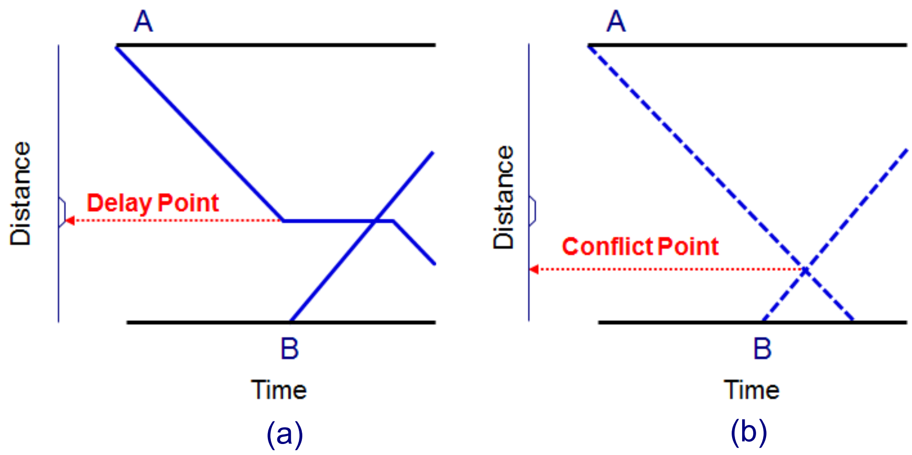

Figure 5. A possible scenario which biases the identification of a capacity constraint point (a) cumulative delay analysis (b) traffic conflict analysis

The two methods for identifying capacity constraints described above lead to two different general approaches to capacity planning: the zonal delay approach based on simulation and delay analysis, and the capacity screening approach based on conflict analysis (Figure 6). Based on simulation output, the zonal delay approach prioritizes the zones with higher cumulative traffic delay for capacity expansion projects. The screening approach developed in this paper determines the order of projects based on the traffic conflict density within their related zone. The core of the screening approach to calculating conflict density is described in the remainder of this paper. The case study section of this paper will compare the form and performance of the optimal capacity expansion strategy suggested by the screening tool to that of the common simulation-based approach when applied to the same route and rail traffic conditions.

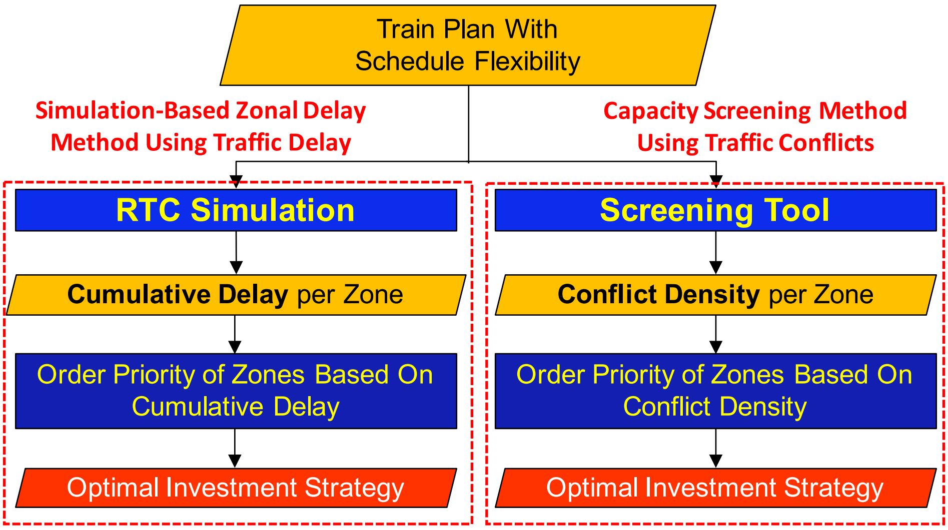

Figure 6: Two methods based on delay and conflict analysis to determine optimal investment strategies

4 Development of Screening Tool

To develop a screening tool that calculates conflict density under the North American operating condition of flexible schedules, it is necessary for this tool to consider the potential variation in departure and trip time flexibility under a mixed operating style. For mixed operations, the departure time and trip time flexibility for each train may not follow parametric distributions. Thus it is not practical to develop the screening tool using a direct mathematical approach. Instead, a process based on the Monte Carlo concept (Mooney, 1997; Chen et al., 2013; Khoo & Teoh, 2014) is proposed for the screening tool framework (Figure 7) with each iteration considering a different combination of train paths.

For each iteration, the train schedule generator creates a set of train paths within the bands defined by the given train departure and trip time flexibility. The conflicts between train paths are left unresolved to reveal the natural conflict locations. The projection process calculates the expected number of conflicts along each zone of the mainline based on the set of generated train paths and the current infrastructure layout. This process is repeated for a different set of random train paths within each train band until the desired number of iterations has been reached. The final output is the mean number train conflicts in each zone on the mainline.

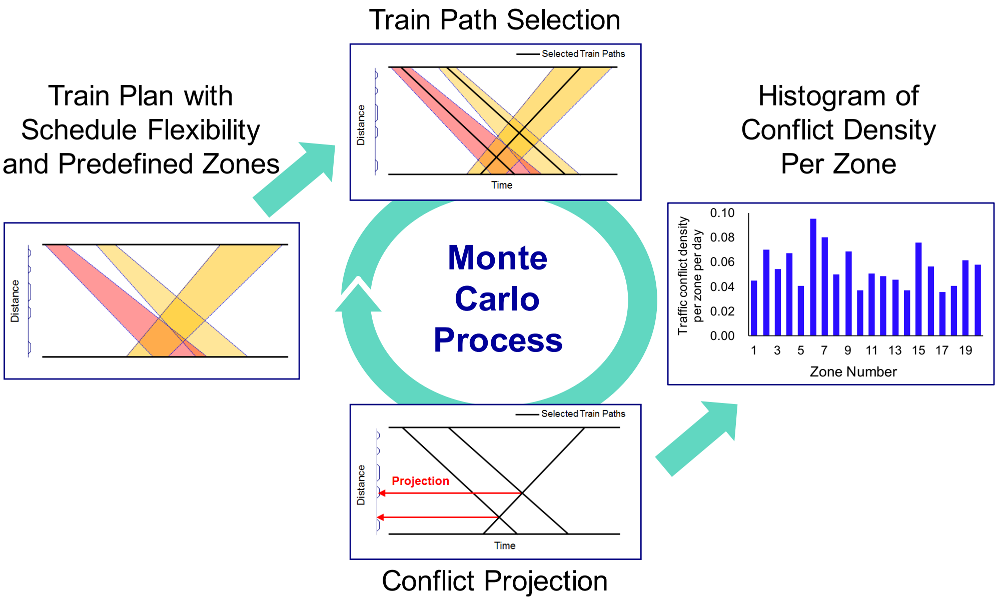

Figure 7: Framework of the screening tool

For the initialization process (Figure 8a), a pre-determined train plan with associated schedule flexibility is specified by the user. As operations become more structured and the width of each train path narrows, the locations of train conflicts become more certain. The conflict density distribution created by the screening tool is likely to show a limited number of peaks to be addressed by infrastructure expansion projects. The line under study is also divided into zones (Figure 8b), just like the evenly divided mainline in the example. The zones can be defined randomly, based on user needs, or to correspond to the location of known potential capacity expansion projects.

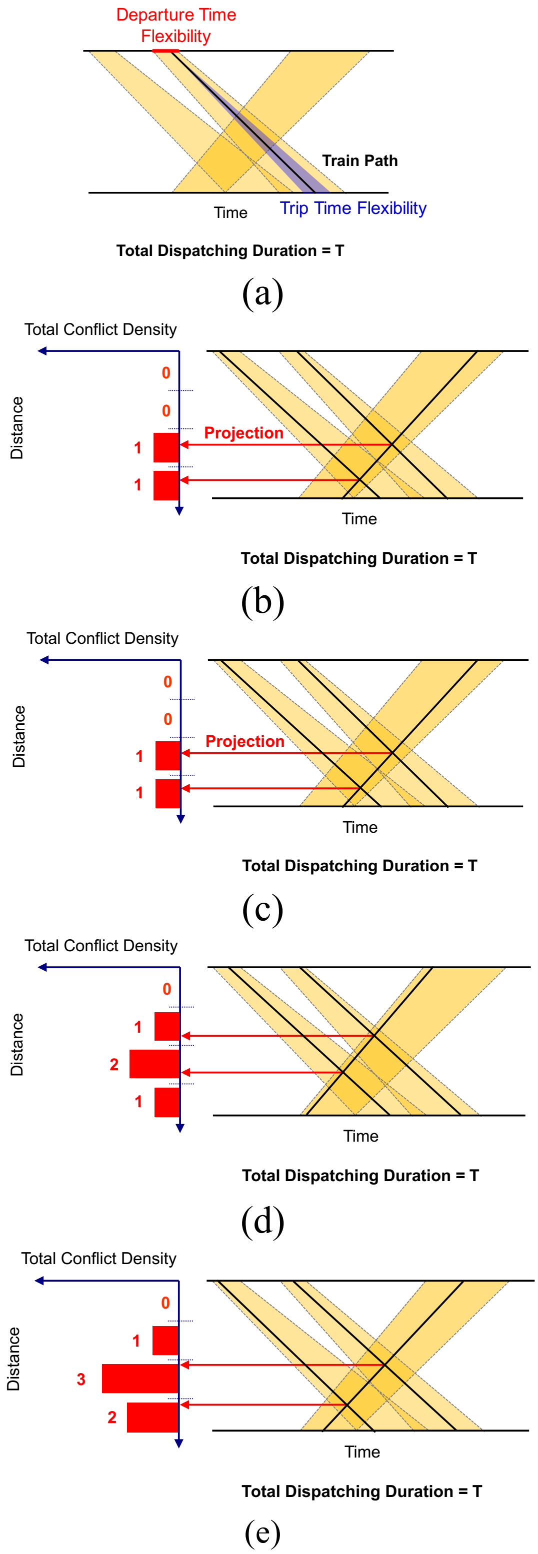

Figure 8. Example of the Two Initialization Steps and Three Iterations of the Screening Tool (a) Initialization of train plan with schedule flexibility (b) Dividing mainline into zones (c) First Iteration (d) Second Iteration (e) Third Iteration

Three iterations are shown to demonstrate the function of the screening tool (Figure 8c to 8e). For every iteration of the process, the number of traffic conflicts within each predefined zone is added to the cumulative total number of conflicts in that zone obtained from the previous iteration. The final output is the total number of projected conflicts in each zone. Since the train paths were generated randomly within each train band based on the stochastic properties of the train operation, dividing the total number of projected conflicts in each zone by the number of iterations is equivalent to calculating the expected number of traffic conflicts along the mainline.

The developed tool can be used for identifying capacity constraints and planning of infrastructure projects. The current version is only applicable to single-track lines. Possible future directions to improve the screening tool will be discussed in the conclusion section.

5 Case Study

To evaluate the effectiveness of the developed screening approach, a case study comparison between the screening tool and delay-based simulation method was conducted. Both methods were used to identify an optimal capacity expansion strategy for a hypothetical single-track line (Figure 9). Each of the strategies contains three construction periods and for each construction period, two passing siding projects can be selected for implementation based on conflict density (for the screening approach) or cumulative delay (for the delay-based simulation approach).

- **Parameters:**
  - 3 train-types with priority high, medium, and low
  - 26 trains per day
  - Variable maximum train speed: 30-80 mph
  - 19 zones based on the existing siding spacing
  - 2-mile siding length and 45 mph siding and turnout speed restriction

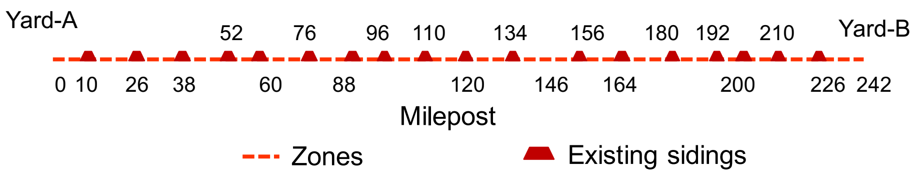

Figure 9: Infrastructure layout of the hypothetical mainline and some related parameters used for analysis

To implement the zonal delay approach using Rail Traffic Controller (RTC) simulation software commonly used for infrastructure planning in North America, the initial infrastructure layout was input to RTC and simulated to generate the histogram of cumulative train delay in each zone (Figure 10a). The two zones with highest cumulative train delay were then selected as the locations for the first two passing siding construction siding projects during the first year of the three-year capacity expansion plan. The infrastructure layout in RTC is then modified to include the two additional sidings implemented during construction period I. This new layout is simulated with RTC again to generate a new histogram of cumulative train delay by zone. The two zones with highest cumulative train delay were then selected as the locations for the second pair of passing siding construction siding projects (year two of the three-year capacity expansion plan). The infrastructure layout is modified again to include the two additional passing sidings implemented during construction period II. This updated layout is simulated with RTC to generate an updated zonal delay histogram and the two zones with the most delay are selected for the final two passing siding projects. The resulting set of infrastructure projects is defined as the optimal infrastructure investment strategy in this study, and the final condition is simulated in RTC to obtain train delay values after construction of all six passing siding projects. In total, four different RTC infrastructure input models must be constructed by the user and four sets of RTC simulation runs must be conducted.

To implement the capacity screening approach, the train plan and associated schedule flexibility was used as the input for the screening tool to generate a single histogram of conflict density along the route (Figure 10b). The histogram is used to identify the six zones with highest conflict density. These zones are selected as the locations of the six passing siding projects to be implemented two at a time over the three construction periods. In generating the capacity expansion plan, the screening tool is only run once and there is no need to develop a detailed model of the existing track infrastructure. The zonal delay approach requires repetitive simulation runs. This difference in effort and total processing time between the two approaches is revisited later.

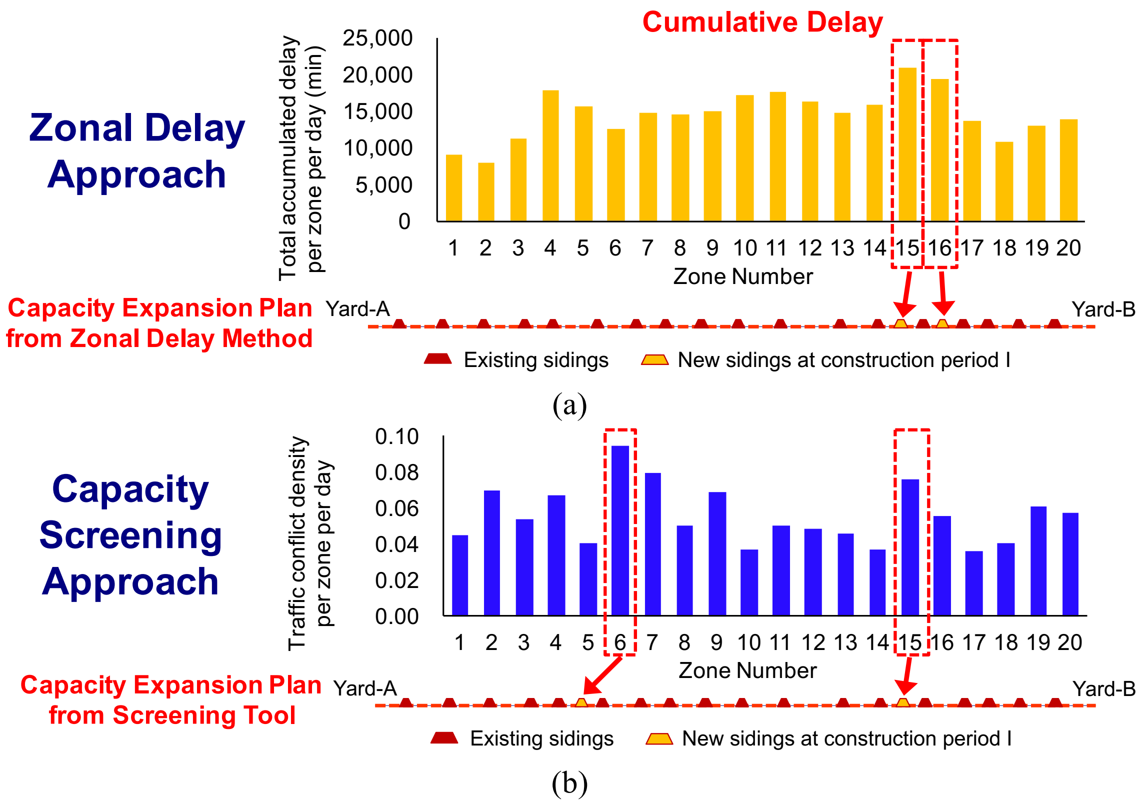

Figure 10: Example of (a) zonal delay (b) capacity screening approach at construction period I

To evaluate the performance of the capacity expansion plan generated by the screening tool, an RTC infrastructure input corresponding to the infrastructure layout at the end of each construction period was created based on the selected projects. The models were simulated with RTC to determine the average train delay after each stage of the capacity plan identified by the screening tool. These average train delay values were compared to equivalent values obtained from RTC simulation of the optimal infrastructure investment strategy determined by the delay-based simulation approach (Figure 11).

The result comparison indicates the capacity screening approach has an equivalent performance in terms of average train delay compared to the zonal delay approach using detailed simulation. However, the zonal delay approach requires human manipulation to construct multiple infrastructure models in RTC and the RTC simulation runs required by the zonal delay method are time consuming. Since the zonal delay approach repeats these two processes for each construction period, it is very time consuming compared to the proposed capacity screening tool. Using this case study as an example, the screening tool took 10 minutes to generate a capacity expansion plan while the zonal delay method requires 4 to 6 hours depending on the efficiency of human manipulation of the RTC files and complexity of the infrastructure model. Using the screening tool can greatly reduce the computational effort required to obtain an optimal infrastructure investment strategy that provides equivalent train-delay performance to plans developed using the zonal delay approach and detailed simulation models.

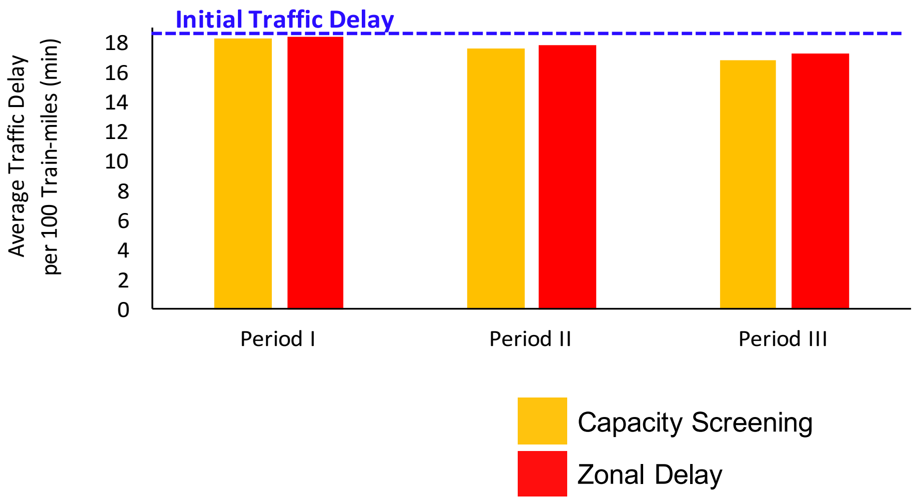

Figure 11: Comparison of performances of optimal infrastructure investment strategies from capacity screening and zonal delay approach

6 Conclusion and Future Study

In this study, a decision support capacity screening approach was developed to help identify the optimal infrastructure investment strategy of a single-track mainline under typical North American freight operations involving a mix of flexible schedules. The developed tool used the concept of identifying traffic conflicts instead of cumulative train delay from simulations. In the case study, an application of the developed approach was demonstrated, and a comparison between this approach and a zonal delay approach using cumulative delay and simulation analysis was made. The output showed that the capacity screening approach can identify an infrastructure investment strategy with equivalent performance to the zonal delay approach but with reduced computation time and effort from the railway infrastructure planner.

A potential improvement to the developed capacity screening approach is to modify the screening tool to consider the existing track infrastructure layout. In the case study, the only consideration given to the current infrastructure layout is in defining the predetermined zones for totalling train conflicts. A more sophisticated mechanism in the screening tool to consider the effect of existing infrastructure layout can help improve the adaptability of the model and the optimality of the outputs. Also, using conflict analysis to identify capacity constraints on multiple-track line is another possible direction for future model improvement. The concept of identifying traffic conflicts may have the potential to predict the delay performance and run-time variability of individual trains. This concept can be used to generate some alternative features to replace the previous ones used to predict train delay (Krueger, 1999; Mitra and Tolliver, 2010). A model constructed based on these features may provide a more accurate prediction of train performance, and thus improve the robustness of train plans.

Acknowledgements

Partial support for this research was provided by the Association of American Railroads (AAR) and the National University Rail (NURail) Center, a USDOT-OST Tier 1 University Transportation Center. The lead author was also supported by the CN Research Fellowship in Railroad Engineering. The authors also thank Darkhan Mussanov, Graduate Research Assistant, University of Illinois at Urbana-Champaign for his assistance in running RTC simulations.

References

Bing, A.J., Beshers, E.W., Chazvez, M., Simpson, D.P., 2010. *Guidebook for Implementing Passenger Rail Service on Shared Passenger and Freight Corridors*, National Cooperative Highway Research Program, NCHRP Report 657, ISSN 0077-5614. Washington, DC, USA.

Cambridge Systematics, 2007. *National rail freight infrastructure capacity and investment study. Cambridge Systematics*, Cambridge Systematics, Cambridge, MA, USA.

Chen, A., Kasikitwiwat, P., Yang, C., 2013. "Alternate capacity reliability measures for transportation networks", J*ournal of Advanced Transportation*, vol. 47, issue 1, pp. 79-104.

Dingler, M.H., Lai, Y.C., Barkan, C.P.L., 2009. "Impact of train type heterogeneity on single-track railway capacity", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2117. pp. 41-49.

Dingler, M.H., Lai, Y.C., Barkan, C.P.L., 2013. "Mitigating train-type heterogeneity on a single-track line", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,* vol. 227, issue 2, pp. 140-147.

Higgins, A., Kozan, E., Ferreira, L., 1997. "Modelling the number and location of sidings on a single line railway", *Computers & operations research*, vol. 24, issue 3, pp. 209-220.

Khoo, H.L., Teoh, L.E., 2014. "An optimal aircraft fleet management decision model under uncertainty", *Journal of Advanced Transportation*, vol. 48, issue 7, pp. 798-820.

Krueger H., 1999. "Parametric modelling in rail capacity planning", *In: Proceedings of the 1999 Winter Simulation Conference*, pp. 1,194-1,200, Institute of Electrical and Electronics Engineers (IEEE), Phoenix, Arizona, USA.

Lai, Y.C., Shih, M.C., 2013. "A stochastic multi‐period investment selection model to optimize strategic railway capacity planning", *Journal of Advanced Transportation*, vol. 47, issue 3, pp. 281-296.

Landex, A., 2008. "Methods to Estimate Railway Capacity and Passenger Delays", *Doctoral Thesis*, Technical University of Denmark, Department of Transport, Lyngby, Denmark.

Luethi, M., Nash, A., Laube, F., Wuest, R., Weidmann, U., 2007. "Increasing Railway Capacity and reliability through integrated real-time rescheduling", URL https://www.researchgate.net/profile/Andrew_Nash/publication/268297119_Increasing_Railway_Capacity_and_Reliability_through_Integrated_Real-Time_Rescheduling/links/54d1ef5d0cf28370d0e171eb.pdf. Accessed: 2016-03-02.

Mooney, C.Z., 1997. *Monte Carlo Simulation*. SAGE Publications, Thousand Oaks, California, USA.

Mitra, S., Tolliver, D., 2010. "Estimation of railroad capacity using parametric methods", *Journal of the Transportation Research Forum*, vol. 49, issue 2, pp. 111-126.

Norio, T., Yoshiaki, T., Noriyuki, T., Chikara, H., Kunimitsu, M., 2005. "Train rescheduling algorithm which minimizes passengers' dissatisfaction", *Innovations in Applied Artificial Intelligence*, vol. 3533, pp. 829-838.

Pouryousef, H., Lautala, P. White, T., 2015. "Railroad capacity tools and methodologies in the U.S. and Europe", *Journal of Modern Transportation*, vol. 23, issue 9, pp. 30-42.

White, T., 2005. "Alternatives for railroad traffic simulation analysis", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1916, pp. 34-41.

Vromans, M., Dekker, R., Kroon, L.G., 2004. "Reliability and heterogeneity of railway services", *European Journal of Operational Research*, vol. 172, issue 2, pp. 647-665.

Martland, C.D., 2010. "Improving on-time performance for long-distance passenger trains operating on freight routes", *Journal of the Transportation Research Forum*, vol. 47, issue 4, pp. 63-80.

Sogin, S.L., Lai, Y.C., Dick, C.T., Barkan, C.P.L., 2013. "Comparison of capacity of single- and double-track rail lines", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2374, pp. 111-118.

Shih, M.C., Lai, Y.C., Dick, C., Wu, M.H., 2014b. "Optimization of siding location for single-track lines", *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2448, pp. 71-79.

Shih, M.C., Dick, C.T., Barkan, C.P.L., 2015. "Impact of passenger train capacity and level of service on shared rail corridors with multiple types of freight trains", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2475, pp. 63-71

Sun, Y., Schonfeld, P.M., 2016. "Capacity investment model for airport facilities under demand uncertainty", *Journal of Advanced Transportation*, online preview.

Uchida, K., Sumalee, A., Ho, H.W., 2015. "A stochastic multimodal reliable network design problem under adverse weather conditions", *Journal of advanced transportation*, vol. 49, issue 1, pp. 73-95.