National University Rail Center - NURail
US DOT OST-R Tier 1 University Transportation Center

NURAIL2013-MTU-R07

RESCHEDULING/TIMETABLE OPTIMIZATION OF TRAINS ALONG THE U.S. SHARED-USE CORRIDORS –

Development of the Hybrid Optimization of Train Schedules (HOTS) Model

By

Pasi Lautala
Director, Rail Transportation Program
Michigan Tech Transportation Institute (MTTI)
Michigan Technological University
ptlautal@mtu.edu

Hamed Pouryousef
PhD Candidate at Rail Transportation Program
: Civil and Environmental Engineering
Michigan Technological University
hpouryou@mtu.edu

11-23-2016

Grant Number: DTRT12-G-UTC18

DISCLAIMER

Funding for this research was provided by the NURail Center, University of Illinois at Urbana - Champaign under Grant No. DTRT12-G-UTC18 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research & Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

National University Rail Center - NURail
US DOT OST-R Tier 1 University Transportation Center

TECHNICAL SUMMARY

Title

Rescheduling/Timetable Optimization of Trains Along the U.S. Shared-Use Corridors: *Development of the Hybrid Optimization of Train Schedules (HOTS) Model*

Introduction

A growing demand for passenger and freight transportation, combined with limited capital to expand the United States (U.S.) rail infrastructure, are creating pressure for a more efficient use of the current line capacity. This is further exacerbated by the fact that most passenger rail services operate on corridors that are shared with freight traffic. Tools and methodologies for capacity analysis are one approach to investigate the situation. As the U.S. continues to develop higher speed passenger services with similar characteristics to those in European shared-use lines, understanding how such analysis are done in both continents grows in relevance. A detailed investigation was done to understand how each continent approaches capacity analysis, and whether any benefits could be gained from cross-pollination. It was found that there was no major divergence between approaches or criteria used for capacity. However, there are differences in the tools used in these two regions, as the tool designs follow the main operational philosophy of each region (timetable based in Europe vs. non-timetable based in the U.S.).

Timetable management is one of the operational methodologies commonly applied in the highly structured European rail system to improve the capacity utilization while maintaining acceptable level of service (LOS) parameters. The potential benefits of using similar methodologies to benefit the less structured U.S. system were studied in this research. A Hybrid Simulation approach was developed as part of the research to investigate the use of timetable management features and to analyze the trade-off between LOS parameters and capacity utilization in the U.S. In the method, the output from Rail Traffic Controller (RTC), a simulation tool commonly used in the U.S., was used as an input for timetable compression by RailSys (a simulation tool developed in Europe).

The results of applying such hybrid simulation approach suggested that timetable compression technique can be applied in the U.S., if an appropriate model and algorithm are developed to address the respective network and operational characteristics of the U.S. rail environment. While the hybrid simulation approach (developed in the project) proved to be successful and provided credible results, it was also extremely time-consuming, reducing its applicability to industry. In addition, the bi-

directional operations pattern commonly used for double/multiple-track corridors in the U.S. limited RailSys' capabilities to provide automatically compressed timetable on such corridors.

To address some of these limitations, a new multi-objective linear programming model called "Hybrid Optimization of Train Schedules" (HOTS) was developed. HOTS model works together with commercial rail simulation tools, extending their capabilities to improve the capacity utilization or the LOS parameters in a conflict-free and compressed timetable. The model was tested on one single-track corridor and one multiple-track corridor (North East Corridor or NEC) for different scenarios. The case study of the NEC between Baltimore and Washington, D.C. was chosen because it is one of the most congested and complicated corridors in the U.S. rail network.

1- Capacity Definition and Analysis Methodologies:

Typically, the capacity of a rail corridor is defined as the number of trains that can safely pass a given segment within a period of time. The capacity is affected by variations in system configurations, such as track infrastructure, the signaling system, operation philosophy, and rolling stock.

The capacity analysis methods are commonly divided into analytical and simulation methods, but a third, "combined simulation-analytical" category was used in this study.

Analytical Approach:

The analytical approach typically uses several steps of data processing through mathematical equations or algebraic expressions, mainly to determine theoretical capacity of the segment/corridor. The outcomes vary based on the level of complexity of the scenario. They may be as simple as the number of trains per day, or a combination of several performance indicators, such as timetable, track occupancy chart, fuel consumption, and speed diagrams. Analytical methods can be conducted without software developed for railroad applications, such as Microsoft Excel, but there are also specialized tools for rail applications. In some cases, analytical models take advantage of different optimization of parametric modeling features, such as probabilistic distribution or timetable optimization. Timetable compression is one of the main analytical approaches to improve the capacity levels in Europe, especially on the corridors with pre-determined timetables (structured operation pattern). A majority of techniques and tools for such applications in Europe are at least partly developed based on timetable compression.

Simulation Approach:

Simulation is an imitation of a system's operation often resembling its real-world equivalent as closely as possible. Generally, the process of simulation is repeated several times until the software achieves an acceptable result. The data needs for the simulation are similar to the analytical methods, but often require a higher level of detail. Simulation process utilizes computer tools to handle sophisticated computations and stochastic models in a fast and efficient way. Simulation use either general simulation tools, such as AweSim, Minitab, and Arena; or commercial software specifically designed for rail transportation, such as RTC, MultiRail, RAILSIM, OpenTrack, RailSys, and CMS.

Simulation Methods: Timetable Based vs. Non-timetable Based

The commercial rail transportation simulation software can be classified in two groups; non-timetable based or timetable based. The non-timetable based simulations are typically utilized for rail systems that use improvised (unstructured) operation pattern without a regular timetable, such as the majority of the U.S. rail network. In this type of simulation, after loading the input data in the software, the train dispatching simulation process creates the train movements, starting with the departure times from the starting station provided as part of the input data. The software may encounter a problem to assign all trains and request assistance from the user to resolve the issue through manual adjustment of the train data, or through modification of the schedule constraints. The Rail Traffic Controller (RTC), developed by Berkeley Simulation Software is the most common software in this category, used extensively in North America.

The simulation procedure in timetable based software (commonly used in Europe) is based on the initial timetable of trains (typically a conflict-free schedule), and the objective is to automatically improve the timetable as much as possible. The UIC's capacity method is often one of the main theories behind the timetable-based simulation approach. The simulation process in this methodology begins with creating an initial timetable for each train. If schedule conflict arises between the trains, the user must adjust the timetable until a feasible schedule is achieved. However, the user actions are more structured compared to the improvised method, and are implemented as part of the simulation process. There are several common software tools in this category, such as MultiRail (U.S.), RAILSIM (U.S.), OpenTrack (Switzerland), SIMONE (the Netherlands), RailSys (Germany), DEMIURGE (France), RAILCAP (Belgium), and CMS (UK).

Combined Analytical-Simulation Approach:

In addition to the analytical and simulation approaches, a combined analytical-simulation method can also be used to investigate the rail capacity. A combined methodology takes advantage of both methodologies' techniques and benefits, and the process can be repeated until an acceptable set of outputs and alternatives is found (Figure 1). Parametric and heuristic modeling (analytical approaches) are more flexible when creating new aspects and rules for the analysis. On the other hand, updating the railroad component input data and criteria tends to be easier in the simulation approach, and the process of running new simulated scenarios is generally faster, although simulation may place some limitations when adjusting the characteristics of signaling or operation rules.

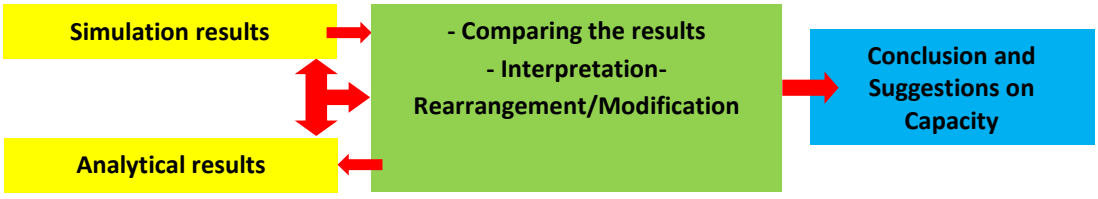

Figure* 1- *Basic diagram of combined analytical-simulation approach for capacity analysis

2- Hybrid Simulation Approach for Improving Railway Capacity and Train Schedules

"Hybrid Simulation Approach" differs from traditional analysis, as it combines the benefits of commercial non-timetable and timetable based software, using output from one software as input in another. The software used in the study included RTC as the non-timetable based simulation tool and RailSys as the timetable-based tool. Figure 2 presents capabilities of each simulation package used in the hybrid approach. RTC has the capability to use preferred departure times, train dispatching simulation process, and automatic train conflict resolution to develop the initial timetable (stringline). RTC uses a decision support core, called "meet-pass N-train logic" to dispatch all trains while avoiding conflicts and minimizing the overall delays and total operating costs of trains. In contrast, RailSys requires an initial timetable (typically conflict-free) for its simulation and uses a timetable compression technique to adjust/improve the initial timetable for more efficient capacity utilization or for improving the LOS parameters.

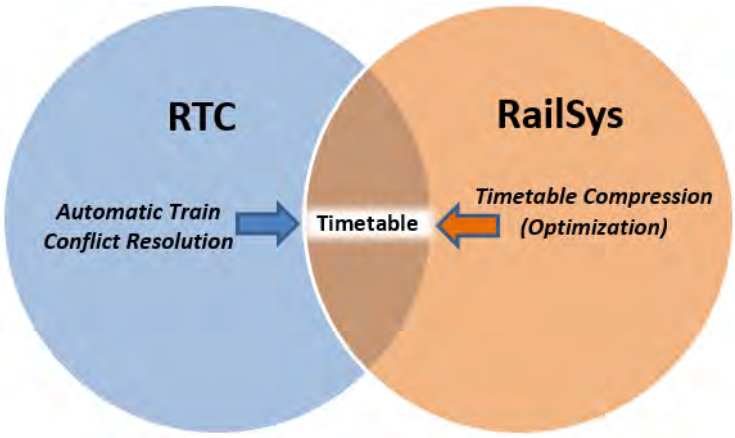

Figure 2- The main features of RTC and RailSys for timetable development

The hybrid simulation approach takes advantage of the initial timetable developed in the RTC as the input for RailSys and then applies the RailSys timetable compression technique to investigate the trade-off between capacity utilization/LOS by adjusting the initial timetable. The adjusted timetable developed by RailSys is then imported back to RTC as input, so the results can be validated. Figure 3 illustrates the steps of hybrid simulation approach in further detail.

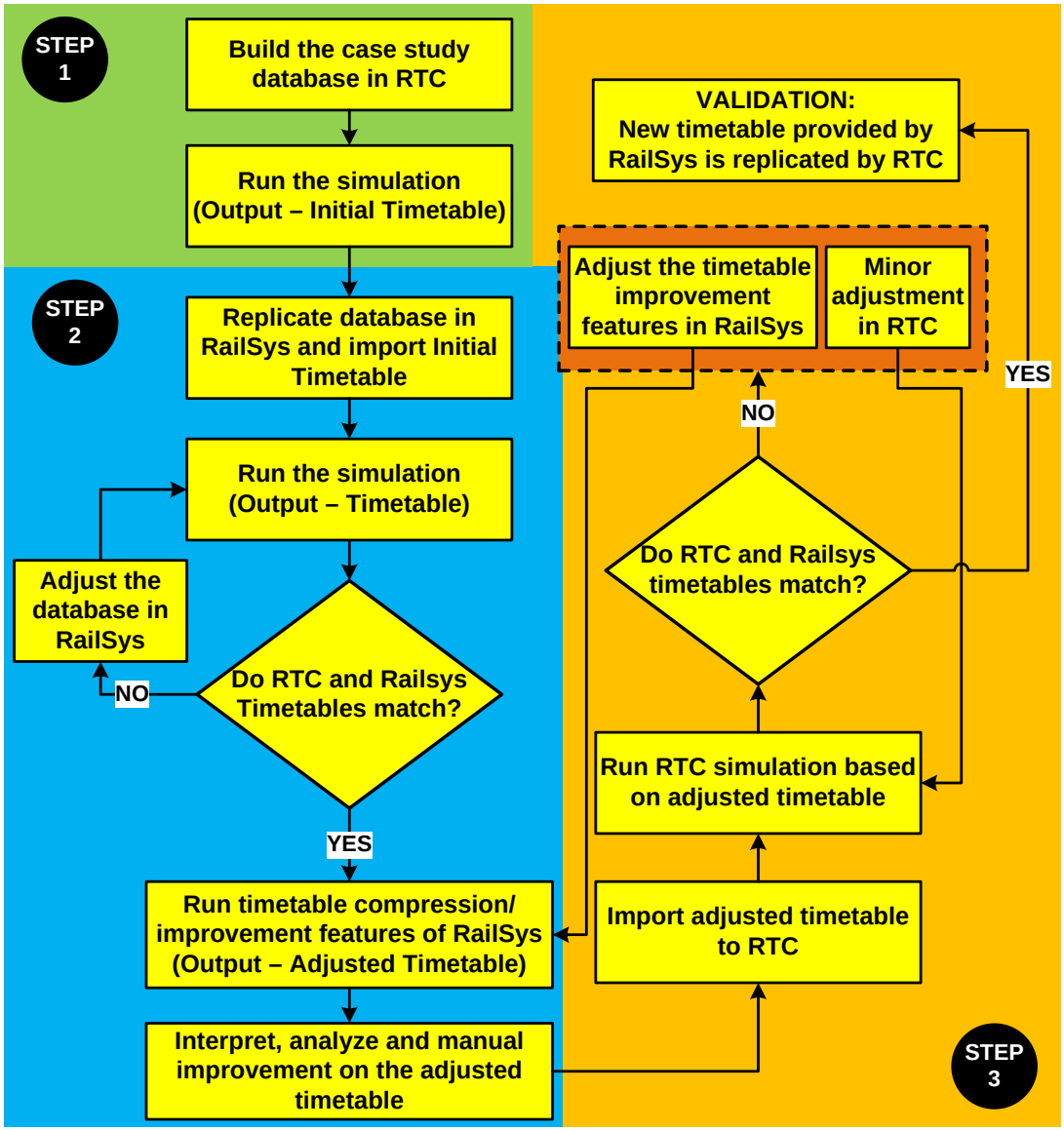

Figure 1- Flowchart of hybrid simulation steps (RTC-RailSys-RTC)

3- Development of Hybrid Optimization of Train Schedules (HOTS) Model for Railway Corridors

As briefly discussed in previous sections, various commercial simulation and timetable management tools can be used to evaluate and improve the corridor operations, but many of them are limited to specific corridor configurations. A new model called “Hybrid Optimization of Train Schedules” (HOTS), was developed as part of this research. The HOTS is designed as a standalone analytical model that works together with all simulation/ timetable management tools. It has no rail infrastructure limitations allowing research on single, double and multiple-track corridors with both directional and

non-directional operational patterns. It uses an initial timetable and user-defined parameters to provide a “Conflict-Free” and compressed schedule.

The HOTS model is formulated as a multi-objective linear programming (LP) model that attempts to minimize two separate parameters, departure times and deviation of dwell times. The model tries to compress the train schedules as much as possible by allowing flexible dwell times for trains during meet-pass and stop events and by departing trains as early as possible, based on the defined priority, allowed flexibilities, and requested departure times. The priority level is defined by user, but in general higher priority trains are expected to be departed earlier and they may have lower dwell time flexibility than the trains with lower priority. Figure 4 demonstrates the main inputs, optimization objectives and outputs of the model.

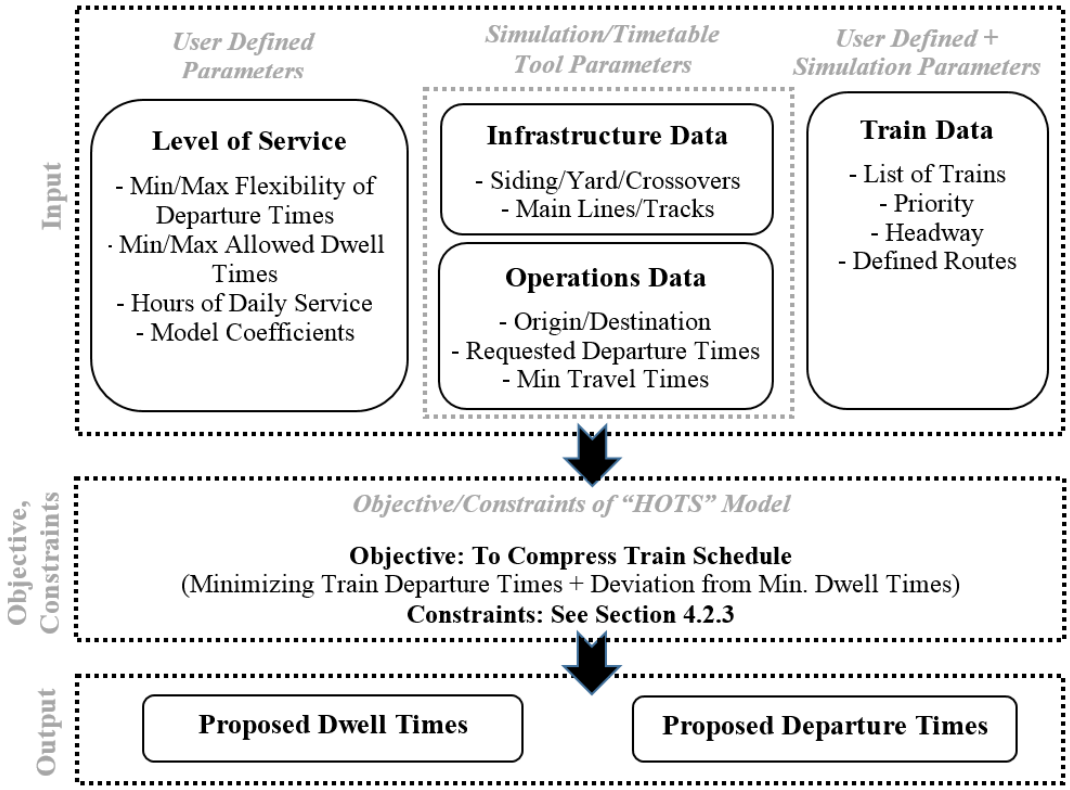

Figure 4- HOTS model input and output

Two categories of model input, “Infrastructure data” and “Operations data”, are extracted from simulation/ timetable management tools. The “Level of service” (LOS) parameters are defined by the user and can be adjusted (calibrated) in the model, as necessary. “Train data” is developed jointly from simulation/ timetable management information and user preferences. All model inputs (parameters) are used by the optimization part of HOTS model with an objective of to compress train schedules, or more specifically to minimize trains departure times and the deviation between adjusted dwell times and respective minimum values. The two main model outputs (variables) include; “proposed dwell times” and “proposed departure times”.

Testing HOTS Model in Single and Multiple Track Situations

Various single and multiple-track case study scenarios were developed to test the HOTS model performance and capabilities. All scenarios were initially developed in RailSys and the final results after HOTS optimization were validated in RailSys. The input parameters for the single-track test case study with mixed passenger/freight traffic are provided in Table 1.

Table 1- Single track case study parameters

Segment Length	30 miles, single track
Sidings/yards	2 sidings + 1 yard
Trains	11 (east) + 11 (west)
Traffic type	Mixed traffic (passenger, commuter, freight, intermodal)

HOTS performance was tested on three scenarios. Scenario 1 used a timetable with several conflicts as an initial schedule (timetable) and HOTS was used to both solve the conflicts and to compress the timetable. Scenario 2 used RTC simulation over the initial schedule from Scenario 1 to develop a conflict-free timetable, which then was compressed by HOTS model (Figure 5). Scenario 3 compared the performance of HOTS compression with that of RailSys. As such, timetable compression technique in RailSys was used to adjust the conflict-free timetable obtained from RTC (the output timetable of Scenario 2) and the resulting RailSys timetable was compared with the HOTS results over the same RTC timetable. Finally, an additional attempt was also made to further compress the RailSys timetable in HOTS as a way to compare their compression algorithm performance.

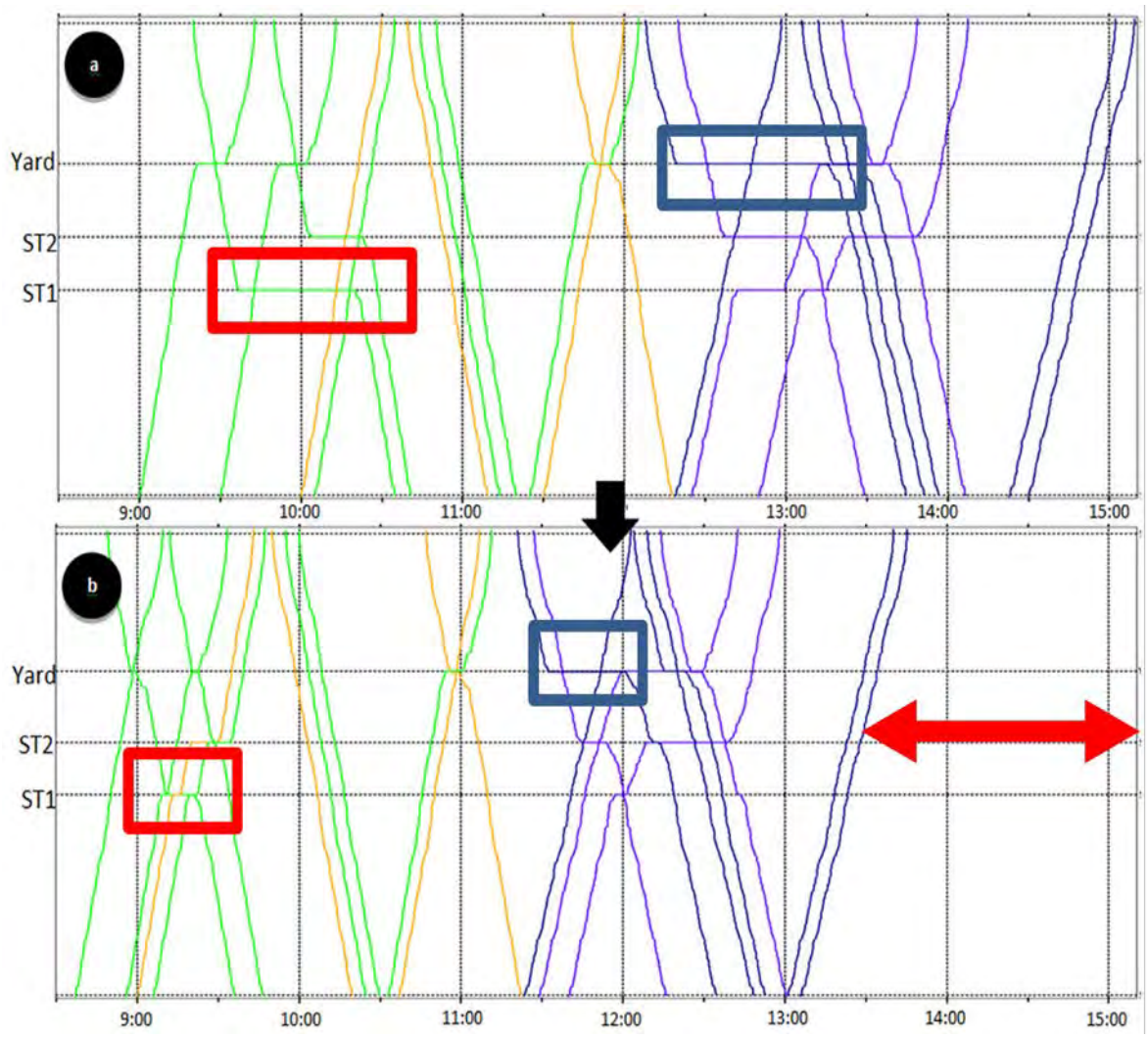

Figure 5- (a) The initial timetable (b) timetable improved by HOTS model

A segment of North-East Corridor (NEC) between Washington, D.C. and Baltimore, MD was used to test the HOTS model performance in a multiple track (N-track) environment (Table 2). The study used the Washington, D.C. – Baltimore, MD segment of the NEC as a stand-alone segment and did not examine continuation of routes on either end. The objective was not to evaluate or recommend any changes to current NEC operations, but rather to better understand the impacts of different operation philosophies on multiple-track corridors. Since the case study did not consider the movement of trains beyond the study limits, none of the suggested modifications are implementable without further study over the entire length of the corridor.

Table -2 Details of multiple-track case study (Baltimore- Washington, D.C.)

Corridor Length	40.60 miles
Length of double track	1.48 miles
Length of triple track	33.94 miles
Length of quadruple track	5.18 miles
Sidings/yards	2 main yards + 8 station/crossover
Number of trains	68 (North) + 68 (South)
Operation pattern	Non-directional
Number of different routes	28 routes

An RTC database that included infrastructure, signaling, rolling stock and operational characteristics was received from Amtrak and later replicated in RailSys for research purposes. Similar to the single-track case, three scenarios were created for the N-track study to evaluate the capabilities of HOTS. Scenario 1 used a conflict-free schedule under bi-directional operation pattern. For Scenario 2, the HOTS model results from Scenario-1 were used as the initial timetable, but adifferent routing alternative was developed for a selected train to examine the capability of HOTS model to provide a conflict-free schedule after new routing (Figure 6). Scenario 3 used a conflict-free schedule from Scenario 1 to evaluate HOTS rescheduling and compression capability under directional operations.

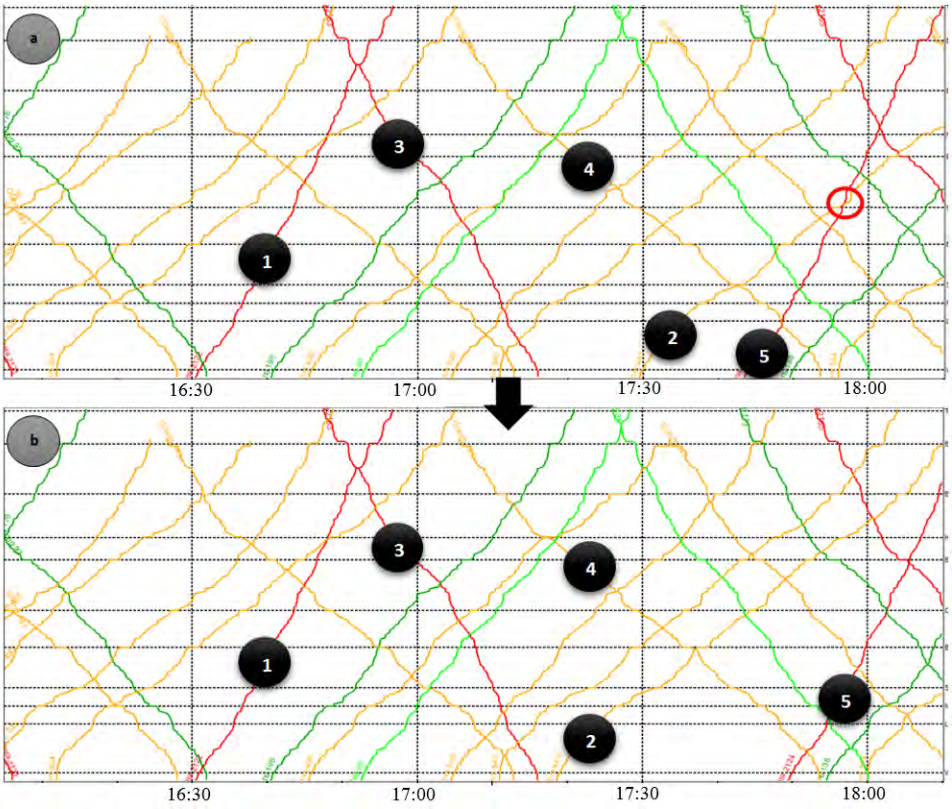

Figure 6- (a)A timetable developed in Scenario 1 of multiple-track case study (red circle indicated schedule conflict for Train #2 (b) Rescheduled timetable by the HOTS model to resolve conflict from Train #2 rerouting

Findings

There were two primary outcomes from the research: 1) development of "Hybrid Simulation Approach for Improving Railway Capacity & Train Schedules" and 2) the "Hybrid Optimization of Train Schedules" (HOTS) model. The Hybrid Simulation approach combines the strengths of two commercial rail simulation packages (RTC and RailSys) to analyze the trade-off between LOS parameters and capacity utilization in the U.S. environment. The results of this approach showed substantial improvements in Level of Service (LOS) parameters, such as reduction in total stops and dwell times. They also confirmed the reverse relationship between LOS criteria and capacity utilization levels i.e. if LOS is improved, the timetable tends to be stretched and capacity utilization may be degraded and vice versa. While the hybrid simulation approach proved to be successful and provided credible results, it was also extremely time-consuming, reducing its applicability to industry.

"Hybrid Optimization of Train Schedules" (HOTS) model was developed to allow the application of European-based scheduling approach over the North American operational environment. HOTS is a multi-objective linear programming (LP) model that works together with commercial rail simulation tools to improve capacity utilization or LOS metrics. Based on the single-track and multiple-track case study tests, the HOTS model was successful in rescheduling and compressing the timetable in multiple

corridor (single, double and multiple track corridors) and operational environments (directional and bi-directional). The model was also capable of resolving conflicts and could reschedule the trains whose routing scenarios were changed. The HOTS model was capable of rescheduling trains under both same-order and order-free scheduling approaches, based on user-defined model inputs. Finally, the HOTS model results were successfully validated in a commercial rail simulation tool, which increases the credibility of the results.

While HOTS model provided satisfactory performance under testing, it is recognized that it currently has certain limitations, such as the lack of automated evaluation of the station capacity limitations, and sensitivity to the requested departure times, flexibility parameters of departure times (FDB and FDA), and the minimum and maximum dwell times of trains. Since train acceleration and deceleration parameters are not included in the HOTS model formulation, minor deviations may appear between departure times suggested by the HOTS model versus those obtained from the simulation package. Finally, HOTS model structure does not offer simultaneous rerouting and rescheduling features, but new train routing alternatives must be defined by the user in the input, and then rescheduled by HOTS to provide a conflict-free timetable based on new routing alternatives.

Conclusions

This research investigated the concept of capacity methodologies and their relationship with the capacity utilization and/or level of service (LOS) metrics. While the objective of capacity analysis is common, there are several differences between the U.S. and European rail systems that affect the approaches, tools and outcomes of analysis. For instance, the research team learned that European capacity analysis tends to be linked to the UIC 406 method, while the U.S. does not seem to have as extensive principles as the European case studies, but the methodologies vary more from one study to another.

As part of this research, a hybrid simulation methodology was developed and tested to learn more about two common simulation tools/methodologies in the U.S. and Europe. In this hybrid simulation approach, an initial timetable was developed in the RTC (a U.S. based simulation software) as input for RailSys, (a European simulation software); and then the RailSys timetable compression technique was applied to investigate the trade-off between capacity utilization/LOS. The adjusted timetable developed by RailSys was then imported back to RTC as input, so the results could be validated. Although the results of this hybrid simulation approach was promising in terms of applying timetable compression technique over a U.S. single track case study; the procedure was time consuming and required significant level of expertise.

To address the limitations of the hybrid simulation approach, a new analytical model, “Hybrid Optimization of Train Schedules” (HOTS) was developed. HOTS model is a multi-objective linear programming (LP) model that works together with commercial rail simulation tools to improve capacity utilization and/or LOS metrics. The HOTS model can develop a conflict-free and compressed timetable of trains under both same-order and order-free scheduling approaches for different

infrastructure orientation (single and multiple-track corridors under directional or bi-directional operation approach), based on user-defined flexibility parameters. The model was tested by developing several scenarios over single and multiple-track case studies and the results were promising and comparable with commercial rail simulation tools.

Related Publications (attached in Appendix)

1. Pouryousef, H, P. Lautala, and D. Watkins; (2016) *Development of Hybrid Optimization of Train Schedules Model for N-track Rail Corridors*, Transportation Research Part C: Emerging Technologies, Vol. 67, pp 169-192, http://dx.doi.org/10.1016/j.trc.2016.02.008
2. Pouryousef, H, Lautala, P.; *Hybrid Simulation;* (2015) *Approach for Improving Railway Capacity and Train Schedules*, Journal of Rail Transport Planning & Management, Vol. 5, pp 211-224, http://dx.doi.org/10.1016/j.jrtpm.2015.10.001
3. Pouryousef, H, Lautala, P., White, T.; (2015) *Railroad Capacity Tools and Methodologies in the U.S. and Europe*; Journal of Modern Transportation- Springer, 23 (1):30-42. doi:10.1007/s40534-015-0069-z, USA
4. Pouryousef, H. and P. Lautala. *Evaluating Two Capacity Simulation Tools on Shared-use U.S. Rail Corridor*. in *2014 Annual Meeting of the Transportation Research Board*. 2014. Washington, DC
5. Pouryousef, H and P. Lautala. *Evaluating the Results and Features of Two Capacity Simulation Tools on the Shared-use Corridors*. in *ASME/ASCE/IEEE 2013 Joint Rail Conference (JRC2013)*. 2013. Knoxville,TN, USA.
6. Pouryousef, Hamed, Pasi Lautala, and Thomas White. *Review of Capacity Measurement Methodologies; Similarities and Differences in the U.S. and European Railroads*. in *2013 TRB Annual Meeting* 2013. Washington-DC.
7. Pouryousef, Hamed, *Timetable Management Technique in Railway Capacity Analysis: Development of the Hybrid Optimization of Train Schedules (HOTS) Model*, Dissertation for the Degree of PhD, Michigan Technological University, 2015 (not attached).

Primary Contact

Name: Pasi Lautala, Ph.D, P.E.
Title: Director, Rail Transportation Program
Department: Michigan Tech Transportation Institute (MTTI)
University: Michigan Technological University
Phone number: 906-487-3547
Email address: ptlautal@mtu.edu

Principal Investigators

Name: Hamed Pouryousef
Title: PhD Candidate at Rail Transportation Program
Department: Civil and Environmental Engineering
University: Michigan Technological University

Name: Pasi Lautala, Ph.D, P.E.
Title: Director, Rail Transportation Program
Department: Michigan Tech Transportation Institute (MTTI)

University: Michigan Technological University
Phone number: 906-487-3547
Email address: ptlautal@mtu.edu

Name:

NURail Center
217-244-4999
nurail@illinois.edu
http://www.nurailcenter.org/

Transportation Research Part C 67 (2016) 169–192

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Development of hybrid optimization of train schedules model for N-track rail corridors

Hamed Pouryousef [a,*], Pasi Lautala [a,1], David Watkins [b,2]

[a] Rail Transportation Program, Department of Civil and Environmental Engineering, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, United States
[b] Department of Civil and Environmental Engineering, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, United States

ARTICLE INFO

Article history:
Received 28 June 2015
Received in revised form 15 February 2016
Accepted 15 February 2016

Keywords:
Railway optimization
Train scheduling
Hybrid optimization
Railway capacity
Timetable compression technique

ABSTRACT

From a capacity perspective, efficient utilization of a railway corridor has two main objectives; avoidance of schedule conflicts, and finding a proper balance between capacity utilization and level of service (LOS). There are several timetable tools and commercial rail simulation packages available to assist in reaching these objectives, but few of them offer both automatic train conflict resolution and automatic timetable management features for the different types of corridor configurations. This research presents a new rescheduling model to address some of the current limitations. The multi-objective linear programming (LP) model is called "Hybrid Optimization of Train Schedules" (HOTS), and it works together with commercial rail simulation tools to improve capacity utilization or LOS metrics. The HOTS model uses both conflict resolution and timetable compression techniques and is applicable to single-, double-, and multiple-track corridors (N-track networks), using both directional and bi-directional operations. This paper presents the approach, formulation and data requirements for the HOTS model. Single and multi-track case studies test and demonstrate the model's train conflict resolution and timetable compression capabilities, and the model's results are validated by using RailSys simulation package. The HOTS model performs well in each tested scenario, providing comparable results (either improved or similar) to the commercial packages.

1. Introduction

Efficient use of railway capacity includes maximizing the number of trains that travel through a corridor while maintaining a predefined level of service (LOS). Modeling can be used to analyze the capacity, although configuration differences between different rail systems, such as the infrastructure ownership and the operations philosophy may lead to the use of different methodologies, techniques, and tools for the analysis (Pouryousef et al., 2013). There are two general approaches to improve the capacity of a rail corridor, either by adding new capital investments in infrastructure or by improving the operation of the rail services. The prevalent use of non-timetable based operating principles in the U.S. has steered the majority of past capacity analysis work toward identifying/evaluating infrastructure improvements that secure conflict free movements. Potential benefits of operational changes are more commonly evaluated for European structured

* Corresponding author. Tel.: +1 (906) 231 2320.
E-mail addresses: hpouryou@mtu.edu (H. Pouryousef), ptlautal@mtu.edu (P. Lautala), dwatkins@mtu.edu (D. Watkins).
[1] Tel.: +1 (906) 487 3547.
[2] Tel.: +1 (906) 487 1640.

http://dx.doi.org/10.1016/j.trc.2016.02.008
0968-090X/

(timetable based) operations, typically in the form of rescheduling and timetable management methods (Landex et al., 2006). As the U.S. develops higher speed passenger rail services with characteristics similar to those found in European shared-use lines (e.g. Northeast corridor and accelerated Michigan corridor), some of the operational differences may diminish, making operational capacity analysis more applicable in the U.S.

There are several methods and tools to evaluate potential improvements in railway capacity, including analytical methods applied by experienced rail personnel, simulation tools (Lai and Barkan, 2011), and combined analytical/simulation models (Pouryousef et al., 2013). In this paper, the "rail simulation" or "simulation" tools refer to the "commercial rail simulation" packages. Timetable management techniques, such as train scheduling, rescheduling, and timetable compression, can be applied for any corridor type, but the operations complexity of a shared-use corridor (where different types of trains share the same track infrastructure) is higher than in corridors with homogenous traffic. A common rescheduling objective is to evaluate the potential capacity for future traffic or to develop a higher quality of LOS for the existing traffic. There are several timetable tools and rail simulation packages that can be used for rescheduling, but according to the previous studies conducted by the authors (Pouryousef and Lautala, 2013, 2014; Pouryousef et al., 2015; Pouryousef and Lautala, 2015a), no commercial rail simulation could be identified with both (1) automatic train conflict resolution and (2) automatic timetable compression features, for various infrastructure configuration and operation pattern. Since the tools that target the U.S. rail environment tend to concentrate on train conflict resolution, timetable management techniques (e.g. timetable compression) and/or optimization models for rescheduling and timetable improvement are limited in most of them (Pouryousef et al., 2013). To address the situation, Pouryousef and Lautala (Pouryousef and Lautala, 2013) presented a hybrid approach where a U.S. simulation software, Rail Transport Controller (RTC) was first used to perform automatic train conflict resolution and initial timetable creation. Then a European software package, RailSys, was used to improve the timetable through automatic compression technique. This method provided good results, but was extremely time-consuming, as it required the construction of matching databases in each simulation package. It was recognized that no timetable compression model is available for the U.S. rail environment similar to the European models. In addition, the bi-directional operations pattern commonly used for double/multiple-track corridors in the U.S. limited RailSys capabilities to provide automatically compressed timetable on those corridors (Pouryousef and Lautala, 2014, 2015a).

The motivation for this study builds on addressing the rail simulation tool limitations identified in the previous studies. The Hybrid Optimization of Train Schedules (HOTS) model presented in this paper is a multi-objective linear programming (LP) model for rescheduling purposes in strategic and tactical planning levels. The modeling approach is called "hybrid", as it works together with existing rail simulation tools, extending their capabilities to improve the capacity utilization (by applying timetable compression technique) or the LOS (by adjusting the train schedule parameters) of a given rail corridor. The primary contributions of HOTS model are summarized as:

(1) The model simultaneously resolves the conflicts and compresses the initial timetable.
(2) The model is applicable for N-track network topologies (single-, double-, and multiple-track) and operational patterns (directional and bi-directional).
(3) The model incorporates various flexibility parameters for rescheduling, such as min/max allowed dwell time and early/late departure time deviation.
(4) The model includes two patterns of rescheduling: "same-order" and "order-free" approaches.

This paper begins with a brief literature review on scheduling and timetable management techniques in the rail industry (Section 2), followed by explanation of HOTS model, including the model concept, application steps, and mathematical formulation (Section 3). Two case studies, each with several scenarios, are used to test the functionality of HOTS model (Section 4). Finally, a discussion of results and conclusions of the research are presented in Section 5, as well as a discussion of future research topics.

2. Literature review

Train scheduling/timetable management has been practiced ever since the rail transportation industry started to develop in early 19th century. Operating rules and time schedules provide logical progression of trains along rail corridors and avoid conflicting movements between trains. Today, computerized timetable management tools and simulation techniques can help rail planners and dispatchers to be more efficient in train scheduling and operation management (Hansen and Pachl, 2008; White, 2005). Typically, evaluating a train's operational features is done either analytically or through simulation, but a combined approach that uses both analytical and simulation methods is also used (Schlechte et al., 2011; Cambridge Systematics, 2006; Pouryousef et al., 2015).

2.1. Analytical-based applications

In the analytical approach, timetable management and train scheduling/optimization are accomplished through mathematical equations or algebraic expressions to determine the optimal solution for the problem (Abril et al., 2007). Several analytical techniques and optimization models have been developed, mostly by academic researchers. The first analytical

models were developed by Amit and Goldfarb (1971) and Szpigel (1973). A train scheduling problem can be developed as a linear programming (LP) model; however, a mixed integer programing (MIP) model is a more common approach, since the number of trains or time periods should be considered in the model as integer values. Examples of MIP models include Kraay et al. (1991) and Carey and Lockwood (1995). More information on optimization models and techniques for train scheduling can be found in sources such as Ghoseiri et al. (2004), Train Scheduling (Chapter 1) by Patty (2015), Railway Applications Section of INFORMS (Ras, 2012), and Harrod (2012).

The following is a review of some of the most important and recently developed optimization models for train scheduling, rescheduling and timetable management, in chronological order. The review briefly explains the structure of the models/application, the approach for solving the models, and study conclusions and relevancy to scheduling, rescheduling or timetable compression applications.

Higgins et al. (1996) developed an optimization model of train scheduling for single track corridors based on each train's earliest departure time from the origin and planned arrival time to the destination. Directional traffic was used for any double-track segments, and the model accounted for scheduled stops and headways. The model variables were defined as optimum departure and arrival times of each train from each station, and the objective was to minimize the train delay at the destination, as well as the train operating costs.

Carey and Carville (2003) developed a train scheduling and platforming (assignment of a train to a particular platform at a given station) optimization model for busy/complex train stations to ensure no conflicts exist between trains. They used a heuristic method and defined an eight-step algorithm of track/platform assignment for each train to find the best platforming option. The objective of model was to minimize the deviation from the desired platforms/tracks as well as minimizing the deviation from the desired headway, turnaround time, and dwell time of each train.

Ghoseiri et al. (2004) introduced a multi-objective train scheduling model of passenger trains along single and multiple track corridors to minimize the fuel consumption and optimize the total time that passengers spent in a train. Zhou and Zhong (2005) developed a bi-criteria train scheduling model for high-speed train applications in China. The model objective was based on minimizing the expected waiting time and the total travel time of trains, assuming practical priority rules for different types of trains using a branch-and-bound algorithm. Other research conducted by Zhou and Zhong (2007) introduced a single-track train timetabling model based on a three-step approach with a branch-and-bound algorithm to find a feasible schedule and resolve all train conflicts. Burdett and Kozan (2006) developed analytical techniques and models to estimate the theoretical capacity of a line based on several criteria, such as traffic mix, directional operation pattern, location of crossings (crossovers, junctions, sidings) and intermediate signals, length of the trains, and dwell time of trains at sidings or stations. Tornquist and Persson (2007) developed a rescheduling model for service interruptions along single- or multiple-track corridors. Their model was applied to a given network with different numbers of parallel tracks and solved using a heuristic approach to find an optimal or near-optimal solution for train rescheduling.

D'ariano et al. (2008) evaluated the concept of timetable flexibility in real-time traffic management to improve the punctuality of service without decreasing the capacity usage of the lines. The foundation of their research was based on focusing more on inter-train conflict resolution during operations by providing a larger degree of freedom (more flexibility) reserved for real-time management to have more chance of recovering from service disturbances. Lindner (2011) reviewed the applicability of timetable compression technique (UIC approach), to evaluate the line and station capacity and concluded that UIC code 406 performs well when evaluating the main line capacity, but it may encounter some difficulties when evaluating node (station) capacity. Corman et al. (2011) developed an innovative approach of optimizing a multi-class rescheduling problem. The problem focused on train scheduling of multiple priority classes in several steps, using a branch-and-bound algorithm. In other research conducted by Corman et al. (2012) a bi-objective problem of conflict resolution was introduced to minimize the train delays, particularly in the case of service interruption, as well as to minimize the missed connections when the rest of train schedule had to be recovered. They used a detailed alternative graph model to ensure schedule feasibility of the case study and developed two heuristic algorithms to find the alternative schedule. Harrod (2012) reviewed the role of railway timetables relative to all other scheduling activities and described the four fundamental timetable formulations available for optimization purposes. Dündar and Şahin (2013) developed a rescheduling model for single-track corridors based on a genetic algorithm (GA) and artificial neural networks (ANN). Based on the model runs over the Turkish State Railway case study, they concluded that GA algorithm performed better in comparison to the ANN in terms of total conflict resolution delay.

In another study conducted by Canca et al. (2014), a nonlinear integer programming model was used for timetable development to adjust the arrival/departure times of trains based on a dynamic behavior of demand. The developed timetable could be used to evaluate the train service quality. Sun et al. (2014) developed a multi-objective optimization model of the train routing problem, combined with train scheduling over a high-speed rail network in China. They used an improved genetic algorithm for this problem, considering the average travel time of trains, energy consumption and user satisfaction parameters. In more recent research conducted by Meng and Zhou (2014), a simultaneous rerouting and rescheduling model was introduced for a given N-track network. The model uses an integer programming approach based on the big-M method and decomposed into a sequence of single train optimization sub-problems for faster solution. In their model, several flexibility parameters were assumed, such as the min/max dwell times and predetermined earliest starting time of a train.

2.2. Simulation-based applications

The simulation methods utilize either general simulation tools or commercial railway simulation software specifically designed for railway capacity analysis. The commercial railway simulation software can be classified as either non-timetable or timetable based (Abril et al., 2007; Khadem-Sameni et al., 2011). Both types of simulation incorporate two main components: Train movement simulation, to calculate the train speed along the track, and Train dispatching simulation, to emulate the actions of the actual dispatcher as closely as possible (White, 2005). The main objective of non-timetable-based simulation is to automatically resolve the train conflicts. They are typically used by railways that operate based on an unstructured operation pattern without detailed long-term timetables, such as found in the majority of the U.S. rail networks. Typically, the timetable based simulation software packages have limited or no capabilities for automatic train conflict resolution, instead they use timetable management features, such as compression technique, to automatically adjust/improve the initial conflict-free timetable/schedule. There are numerous software tools available in each category, but this paper uses a non-timetable based, Rail Traffic Controller (RTC) developed by "Berkeley Simulation Software, LLC, U.S.", and a timetable based, RailSys, developed by "Rail Management Consultants GmbH, Germany", simulation packages. More information on these two types of simulation and related software is provided by Pouryousef and Lautala (Pouryousef et al., 2015; Pouryousef and Lautala, 2015b).

Several recent studies have been conducted using commercial simulation tools to evaluate rail operations and capacity features. Sogin et al. (2012) used RTC to analyze the delay status of freight trains in double-track case studies. They applied various speed scenarios and passenger/freight train volumes and concluded that running faster passenger trains on a double track corridor can reduce the total capacity of the corridor and increase the overall delay. On the other hand, an equal priority scenario for all types of trains can reduce the overall delay. In another study by Sogin et al. (2013), RTC simulation and delay analysis were used to compare train performance on single- and double-track corridors. In this study, Sogin et al. developed and tested alternative scenarios by changing traffic volume, passenger train speed and the heterogeneity level of freight and passenger trains and concluded that increasing passenger train speed can reduce the travel time, but it may also reduce the reliability of trains. They took advantage of automatic train conflict resolution and randomization features of RTC, mainly to analyze the delay and speed metrics of different scenarios. Train scheduling and timetable management aspects (e.g. rescheduling and timetable compression) were not included in the studies.

Most timetable-based simulation research has been concentrated in Europe. The Swedish National Rail Administration (Banverket) carried out a research project in 2005 to evaluate the application of the UIC capacity methodology (timetable compression) for the Swedish rail network. RailSys software was used for the simulations. The research confirmed the validity of the UIC's approach for the Swedish rail network, but the team also concluded that buffer times are necessary for service recovery, and without them, service punctuality can be significantly degraded due to increased capacity consumption (Banverket, 2005). In another study, Schlechte et al. (2011) used the European rail simulation software OpenTrack to obtain microscopic level results of simulated runs, and then converted the results to a macroscopic level for further timetable development/improvement by an analytical algorithm. The improved timetable was returned to the simulation for further analysis. Gille and Siefer (2013) used RailSys in a three-step application to analyze the capacity improvement of a case study that included obtaining maximum level of track occupancy, running the simulation to determine the service quality, and adjusting the maximum level of track occupancy. Goverde et al. (2014) applied ROMA simulation package on Dutch railway corridors to analyze various signaling and traffic conditions. The analysis included timetable compression for unscheduled (disturbed) traffic conditions and Monte Carlo simulation.

In addition to the timetable and non-timetable based simulation approaches, a new "Web-based Screening Tool for Shared-Use Rail Corridors", RailEval, was developed in the U.S. by Brod and Metcalf (2014) to perform a preliminary feasibility screening on proposed shared-use rail corridor projects. The outcomes of RailEval can be used to identify projects that warrant further investigation by applying more detailed analytical/simulation tools. It is based on a simplified simulation technique which does not provide optimization features or complex simulation algorithms. It requires development of basic levels of infrastructure, rolling stock and operation rules (train schedule) of the given corridor, and a conflict identifier that can help the user determine where a siding or yard extension is needed to resolve existing or future conflicts along the corridor.

2.3. Timetable compression technique

The timetable compression technique is a general method for rescheduling which can be completed in both analytical and simulation approaches. The method readjusts the operational characteristics of train service and is especially applicable for corridors with pre-scheduled timetables of all daily trains (structured operation pattern). A majority of European techniques and tools rely, at least partially, on a timetable compression technique. The UIC's standard for evaluating and improving capacity (UIC leaflet 406, updated in 2013) is also based on a timetable compression technique (Landex et al., 2006; Prinz and Hollmuller, 2005; Banverket, 2005; Uic, 2004; Khadem-Sameni et al., 2010).

In the initial UIC approach, the pre-scheduled timetable is modified by rescheduling trains to follow each other as closely as possible. Changes in the infrastructure or rolling stock specifications are not allowed during the process, and neither are modifications of the travel times, crossing and/or station locations, or commercial stops. Potential new slots on the timetable

which are generated through compression can be dedicated for additional train service or for maintenance activities (Abril et al., 2007). The basic steps of UIC methodology are presented in Fig. 1.

Typically, there are two approaches of rescheduling and compressing a timetable. "Same-order" approach maintains the train order based on the initially requested departure times, but the train order at arrival may differ from the initial schedule due to the compression and potential adjustments in stop patterns. "Order-free" (shuffle) approach departs trains based on user preferences, such as earliest possible departure times of trains. Train order may be changed in both departure and arrival locations.

Simulation and timetable management tools equipped with timetable compression techniques usually follow one of the two above mentioned approaches of rescheduling/compression. The UIC compression technique (2004 edition) is normally based on same-order approach, such as the timetable compression technique available in RailSys (Rmcon, 2010).

2.4. Literature and HOTS model

As previously mentioned, HOTS model is a multi-objective linear programming (LP) model for rescheduling purposes. There are similarities and differences between HOTS model and past studies, as described in Table 1.

For instance, similar to D'ariano et al. (2008) and Meng and Zhou (2014), the HOTS model considers several flexibility parameters, such as min/max dwell time at each stop point, but HOTS also considers other flexibility parameters which are associated with the maximum deviation allowed on the train departure times. The HOTS model applies a timetable compression technique derived from UIC 406 code, similar to other studies for improving the LOS and capacity utilization, such as Banverket (2005), D'ariano et al. (2008), Lindner (2011), and Goverde et al. (2014). However, the timetable compression techniques used in each model (including HOTS model) differ in the parameters and sub-algorithms used. In addition, HOTS is capable of applying timetable compression for both same-order and order-free patterns. Automatic conflict resolution, similar to Zhou and Zhong (2007) and Corman et al. (2012), is another feature included in the HOTS model. The HOTS model is applicable to either single-track or multiple-track cases (N-track networks), similar to the rescheduling models proposed by Ghoseiri et al. (2004), Tornquist and Persson (2007), and Meng and Zhou (2014). However, the HOTS model structure has been designed to handle both directional and bi-directional patterns of operation for multiple-track corridors, a characteristic not commonly available in other optimization models (e.g., Higgins et al., 1996; Burdett and Kozan, 2006). Finally, the rerouting option addressed by Meng and Zhou (2014) and Sun et al. (2014) is available in HOTS, but the re-routing decision is not represented as an optimization variable. Instead, the re-routing option for train optimization is defined by the user in HOTS model.

3. Overview of HOTS model

3.1. Conceptual design and methodology of HOTS model

The Hybrid Optimization of Train Schedules (HOTS) presented in this paper is a multi-objective linear programming (LP) model for train rescheduling at strategic and tactical planning levels. It works together with existing rail simulation tools, extending their capabilities to improve the capacity utilization or the LOS of a given rail corridor by applying timetable compression technique (adjusting the train schedule parameters). It should be noted that capacity utilization can be increased either by operating more trains in the same time period, or by reducing the time period (timetable duration) while maintaining the number of trains (timetable compression technique). Since the optimization concept of the HOTS model is derived from the timetable compression technique introduced by the International Union of Railways (UIC), capacity utilization is represented in this paper by the timetable duration parameter. Thus, HOTS keeps the same number of trains while adjusting the timetable duration of train schedules.

The HOTS model applies user-defined parameters, such as the flexibility of each train's departure time and allowable dwell time at each stop point. It tries to minimize the departure time of trains as well as the deviation between proposed dwell time and the allowable minimum values. Additionally, the model can reschedule different trains based on user-defined routing scenarios, instead of the current routes obtained from the simulation package. The model outputs include proposed train departure and dwell times. The outputs and suggested changes to train routings (provided manually by the user in the model input as necessary) can be exported to a commercial simulation tool to perform further analysis; or to simply verify the results.

Fig. 2 presents the cyclical process for applying the HOTS model, including:

(1) Extracting the initial (requested) timetable from a simulation or timetable management tool (A).
(2) Developing the respective datasets in a tabular format, based on initial timetable and user-defined parameters (such as min/max. flexibility of departure and dwell times, and train routing) (B).
(3) Running the optimization part of HOTS in a solver, such as CPLEX (by IBM), Gurobi (by Gurobi Optimization), or LINGO (by LINDO Systems, Inc.). The outputs of optimization include train departure and dwell times within the defined limits (C).

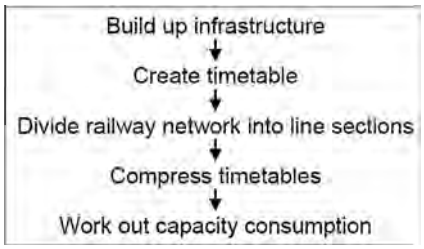

Fig. 1. Main steps of timetable compression for the same-order scheduling approach by UIC 406 method, including potential schedule conflict resolution and adjustment of the initial timetable (adapted from Landex et al. (2006).

Table 1
Similarities and differences between HOTS model and other available models.

Model/Publication	Similarities with HOTS model	Different in HOTS model
D'ariano et al. (2008) and Meng and Zhou (2014)	Both include several flexibility parameters, such as min/max dwell time	Includes additional flexibility parameters, such as maximum deviation (flexibility) before and after the requested departure time
Banverket (2005), D'ariano et al. (2008), Lindner (2011) and Goverde et al. (2014)	All use timetable compression technique	– Different parameters and sub-algorithms of timetable compression technique – HOTS includes both same-order and order-free patterns
Zhou and Zhong (2007) and Corman et al. (2012)	Automatic conflict resolution capability	Minor conflicts due to lack of acceleration/deceleration parameters
Ghoseiri et al. (2004), Tornquist and Persson (2007) and Meng and Zhou (2014)	All applicable to either single-track or multiple-track corridors (N-track networks)	HOTS includes both directional and bi-directional patterns of multiple-track corridors (N-track networks)
Meng and Zhou (2014) and Sun et al. (2014)	Re-routing capabilities	New routings determined by the user

(4) Updating the departure and dwell times, as well as the potential new routings (as defined by the user) in the tabular datasheets (D).
(5) Validating the new departure, dwell times (and new routes) in a simulation or timetable management tool, performing further analysis and/or starting a new iteration, as desired (A).

It should be noted that Step A (commercial rail simulation/timetable management tool) is not required to obtain a solution from the optimization part of HOTS (Step C), but it facilitates the procedure of data extraction and validation of proposed results. Also, Steps B and D are ordinary database management steps which are typically conducted in any optimization modeling study, either by using internal features of the solver or by applying other external database management tools.

The main contribution of this research is HOTS optimization (Step C of Fig. 2). The optimization includes three main components, described in more detail in Fig. 3:

(1) Model data and parameters (inputs).
(2) Objective functions and constraints (limitations/expectation).
(3) Decision variables (outputs).

The level of service (LOS) parameters used in the analysis depend on the perspective for the study. For example, meaningful parameters to evaluate the desired level of rail customer/clients' satisfaction may be quite different from parameters that evaluate the operational efficiencies from an infrastructure manager's or operator's standpoint. In the HOTS model, the LOS parameters (listed in Fig. 3) are defined from a scheduling standpoint and can be adjusted by the user, as necessary. Two categories of model input – infrastructure and operations data – are extracted from simulation/timetable management tools, and train data is developed jointly from simulation/timetable management tool and user preferences. All model inputs (parameters) are used by the HOTS decision core, with an objective to simultaneously resolve all potential conflicts and compress the initial timetable (Fig. 4). Thus, the initial timetable is always under pressure from both sides of decision core to provide a conflict-free and compressed timetable as the outcome of the rescheduling problem. More specifically, HOTS attempts to minimize a weighted sum of proposed departure times (output) and the deviation between proposed dwell times (output) and minimum allowed dwell times.

3.2. Mathematical formulation of the HOTS model

The optimization part of HOTS (Component "C" in Fig. 2) is formulated as a multi-objective Linear Programming (LP) model, and it can be solved by using either simplex or dual-simplex algorithms. The mathematical formulation is described in the following sections.

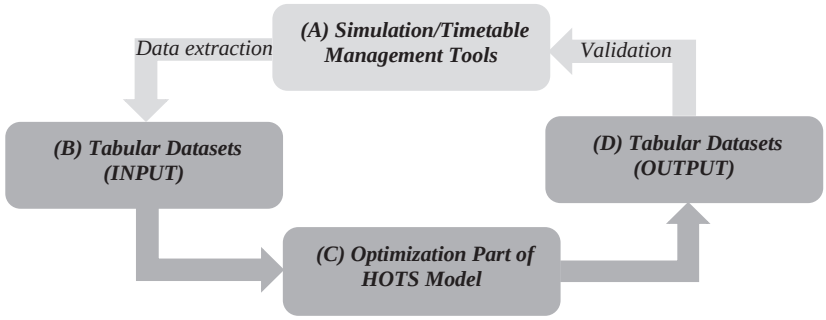

Fig. 2. Main steps of HOTS model operation.

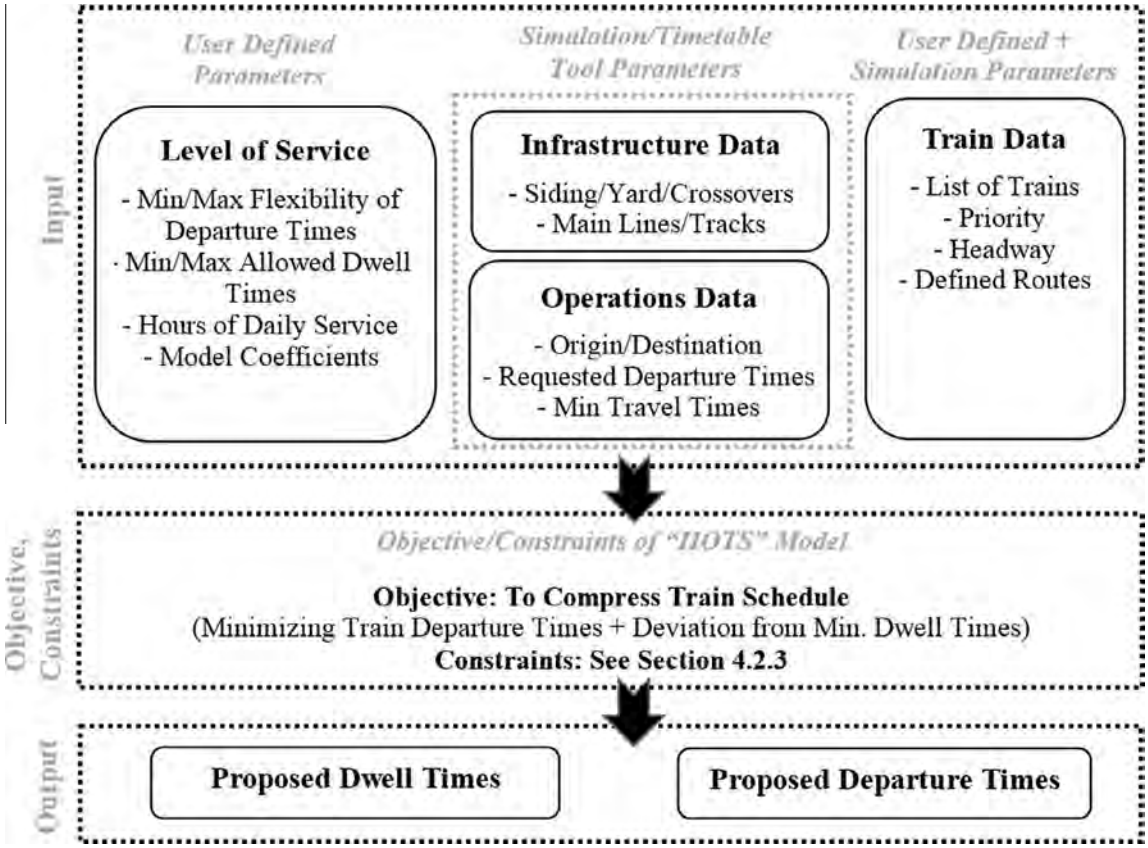

Fig. 3. HOTS model optimization components including Input, Objective/Constraints and Output.

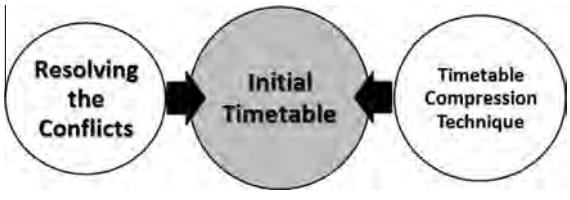

Fig. 4. Main decision core of HOTS model.

3.2.1. Model parameters and variables

Table 2 summarizes the HOTS model data and parameters (Input) and variables (Output).

3.2.2. Model objective

The HOTS model attempts to minimize two separate values – departure time and deviation of dwell time. The model compresses the train schedules as much as possible by allowing flexible dwell times for meet-pass and stop purposes, and by scheduling departures as early as possible, based on the defined priority, allowed flexibilities, and requested departure times. The priority level is defined by the user, but in general higher priority trains are expected to depart earlier and have less dwell time deviation than trains with lower priority. The objective function is presented in Eq. (1).

$$Objective: \ MIN \ \alpha_1 \times \sum_t \sum_i \left(XW_t^i - LW_t^i\right) \times R_t + \alpha_2 \times \sum_t \sum_i XDT_t^i \times R_t \tag{1}$$

Eq. (1): In this equation, α_1 and α_2 are weighting coefficients that indicate the relative importance of dwell time versus departure time, respectively. As the numeric values of dwell time deviation (first part of the function) are much smaller than train departure times (second part), the user can scale these two parameters depending on weighting preferences. Increasing α_1 allows user to prioritize the preservation of desired dwell times over the compression of the new rescheduled timetable, and vice versa, as discussed later in a sensitivity analysis of the α_1, α_2 coefficients (Section 4.3).

3.2.3. Model constraints

The HOTS model has several constraints which can be applied to both same-order and order-free rescheduling/compression approaches. The following sections provide a detailed description of the model constraints in each approach.

3.2.3.1. Model constraints under same-order approach. Eqs. (2)–(11) represent the constraints for the same-order rescheduling approach.

$$XDT_t^i \geqslant DT_t^i - FDB_t^i \qquad \forall t \in T, \quad \forall i \in S \tag{2}$$

Eq. (2): Departure time of each train from each stop point (left hand side) should be no less than the earliest possible departure time allowed for the given train (right hand side).

$$XDT_t^i \leqslant DT_t^i + FDA_t^i \qquad \forall t \in T, \quad \forall i \in S \tag{3}$$

Eq. (3): Departure time of each train from each stop point/station (left hand side) should be no greater than the latest possible departure time allowed for the given train (right hand side).

$$LW_t^i \leqslant XW_t^i \leqslant UW_t^i \qquad \forall t \in T, \quad \forall i \in S \tag{4}$$

Eq. (4): The dwell time of each train should be maintained between minimum and maximum allowed dwell time allowed at each stop point/station. It should be noted that a train will not be able to stop at a given stop point $\left(XW_t^i = 0\right)$ if both minimum and maximum dwell times are set to zero.

$$XDT_t^d - XDT_t^o = \sum_j \sum_i TR_t^{ij} + \sum_j XW_t^j \qquad \forall t \in T, \quad \forall i,j \in S, \quad |i-j| = 1, \quad d \in D_t, \quad o \in O_t \tag{5}$$

Eq. (5): Total travel time of each train (left hand side) is equal to the sum of route travel times between origin/destination plus the sum of all dwell times in the stop points/stations.

$$XDT_t^j = XDT_t^i + TR_t^{ij} + XW_t^j \qquad \forall t \in T, \quad \forall i,j \in S, \quad |i-j| = 1 \tag{6}$$

Eq. (6): Train departure time from each stop point/station (left hand side) is equal to the departure time of the previous stop point/station, plus the travel time of the previous section of route and the dwell time of the current stop point/station.

$$XDT_t^i - XDT_p^i \geqslant H(T_p) + H(T_t) + \left(TR_p^{ij} - TR_t^{ij}\right) \quad where \ (U_t \times U_p = 1) \cap \left(DT_t^i > DT_p^i\right) \cap \left(TR_p^{ij} \geqslant TR_t^{ij}\right) \cap \left(MR_p^{ij} = MR_t^{ij}\right), \quad \forall t,p \in T, \quad t \neq p, \quad \forall i,j \in S, \quad |i-j| = 1 \tag{7}$$

$$XDT_t^i - XDT_p^i \geqslant H(T_p) \quad where \ (U_t \times U_p = 1) \cap \left(DT_t^i > DT_p^i\right) \cap \left(TR_p^{ij} < TR_t^{ij}\right) \cap \left(MR_p^{ij} = MR_t^{ij}\right), \quad \forall t, \quad p \in T, t \neq p, \quad \forall i,j \in S, \quad |i-j| = 1 \tag{8}$$

Table 2
List of optimization parameters and variables of the HOTS Model.

Parameter	Description
T	Set of all trains "t" (or "p") $\forall t, p \in T$
R_t	Priority of train "t" $R_t \in \{1, 2, 3, \ldots\}$ (should be determined based on the importance of the train service quality and schedule of trains. The higher the priority of train, the higher value of R_t)
$H(T_t)$	A minimum headway of train "t" (departure headway) before dispatching another train on the same track. (min or sec)
SH	Maximum duration of timetable (converted to min or sec) $SH > 0$
S	Set of stop locations "i" (e.g. station, siding, yard, crossover) $\forall i \in S$
α_1, α_2	Weighting coefficients of dwell (α_1) and departure times (α_2) $\alpha_1, \alpha_2 > 0$
O_t	The origin sub-set of train "t" out of set "S" $\forall O_t \in S, o \in O_t, \forall t \in T$
D_t	The destination sub-set of train "t" out of set "S" $\forall D_t \in S, d \in D_t, \forall t \in T$
U_t	Direction of train "t" $\begin{cases} 1 & \text{if } O_t < D_t \ (Northbound) \\ -1 & \text{if } O_t > D_t \ (Southbound) \end{cases}$
DT_t^i	Requested departure time (daily clock time) of train "t" from stop point "i" (min or sec) $\forall i \in S, \ \forall t \in T$
FDB_t^i	Maximum deviation (flexibility) of departing train "t" before the requested time from stop point (station) "i" (min or sec) $\forall i \in S, \forall t \in T$
FDA_t^i	Maximum deviation (flexibility) of departing train "t" after the requested time from stop point (station) "i" (min or sec) $\forall i \in S, \forall t \in T$
LW_t^i	Minimum allowed dwell time of train "t" at stop point "i" (min or sec) $\forall i \in S, \forall t \in T$
UW_t^i	Maximum allowed dwell time of train "t" at stop point "i" (min or sec) $\forall i \in S, \forall t \in T$
TR_t^{ij}	Travel time of train "t" on allocated route between each two consecutive stop points "i–j" (min or sec) $\forall i, j \in S, \lvert i - j \rvert = 1 \quad \forall t \in T$
MR_t^{ij}	Matrix of track number which is assigned to each train "t" running along two consecutive stop points "i–j" (based on existing patterns from simulation tool or defined by user) $\forall i, j \in S, \lvert i - j \rvert = 1 \quad \forall t \in T$
XDT_t^i	Proposed departure times of train "t" from each stop point "i" (min or sec) **(VARIABLE)**
XW_t^i	Proposed dwell time of train "t" at each stop point "i" (min or sec) **(VARIABLE)**

Eqs. (7) and (8): There should be a minimum headway, or buffer time (right hand side), between departures of two consecutive trains (left hand side) considering multiple constraints. These constraints are written for a sub-set of trains defined by the following conditions, as specified by model data:

(1) Trains operate in the same direction $(U_t \times U_p = 1)$.
(2) The same order of trains is maintained $\left(DT_t^i > DT_p^i\right)$.
(3) Train t is faster than train p $\left(TR_p^{ij} \geqslant TR_t^{ij}\right)$; Eq. (7); or vise versa $\left(TR_p^{ij} < TR_t^{ij}\right)$; Eq. (8).
(4) Trains share the same route $\left(MR_p^{ij} = MR_t^{ij}\right)$.

Eqs. (7) and (8) differ in the order of slower and faster trains. Eq. (7) represents the scenarios where the faster train is following a slower one $\left(TR_p^{ij} \geq TR_t^{ij}\right)$. Therefore, Eq. (7) has an extra term on the right hand side of the constraint which represents an additional buffer time, calculated based on the minimum headway of the faster train $(H(T_t))$ and the speed gap between the trains $\left(TR_p^{ij} - TR_t^{ij}\right)$. Since faster and slower trains are determined by Eqs. (7) and (8), the headway defined in Table 2, $H(T_t)$, only considers the train dispatched earlier (previous train).

$$\begin{aligned} &XDT_t^i \geqslant XDT_p^j + TR_p^{ji} + H(T_p) \\ &where \ (U_t \times U_p = -1) \cap \left(DT_t^i \geqslant DT_p^j\right) \cap \left(MR_p^{ji} = MR_t^{ij}\right), \\ &\forall t, \quad p \in T, t \neq p, \quad \forall i, j \in S, \quad |i - j| = 1 \end{aligned} \tag{9}$$

Eq. (9): No train can depart (left hand side) until the previous train in the opposite direction has arrived to the given station (first and second parts of the right hand side) plus minimum headway between these two trains (third part of the right hand side). The following conditions define the sub-set of trains to which these constraints are applied:

(1) Trains operate in opposite direction $(U_t \times U_p = -1)$.
(2) The same order of trains is maintained $\left(DT_t^i \geqslant DT_p^j\right)$.
(3) Trains share the same route $\left(MR_p^{ji} = MR_t^{ij}\right)$.

In these three constraints (Eqs. (7)–(9)), the MR_t^{ij} and MR_p^{ji} parameters have been incorporated to check the assigned tracks for each pair of individual trains. If trains share the same track, the conflict resolution is activated. This feature, together with evaluation of train direction ($U_t \times U_p = 1$ and $U_t \times U_p = -1$) makes the model applicable to N-track network configurations (single, double, and multiple-track), as well as various operational patterns (directional or bi-directional).

It should be noted that the order of trains is inherently maintained in these three constraints, as one of the conditions was to check the initial order of trains $\left(DT_t^i > DT_p^i\right)$. If this condition is met (together with the other conditions explained above), the new departure times will be proposed while maintaining the initial order of trains.

$$XDT_t^d - XDT_p^o \leqslant SH \qquad \forall t,p \in T, \quad d \in D_t, \quad o \in O_t \tag{10}$$

Eq. (10): Timetable duration (left hand side) should be no greater than the maximum service hours defined by the user.

$$XDT_t^i \geqslant 0, \quad XDT_t^i \in real, \quad XW_t^i \geqslant 0, \quad XW_t^i \in real \tag{11}$$

Eq. (11): Proposed departure times and dwell times variable are defined as non-negative real values to provide faster and more reliable solution. Although in theory both dwell time and departure time are not required to be defined as integer values, we can assure that these variables take on integer values by specifying integer values for requested departure times, travel times and min/max allowed dwell times in the model, due to the structure of constraints defined by Eqs. (5) and (6).

3.2.3.2. Model constraints under order-free approach. In the order-free approach of the HOTS model, trains depart based on the earliest possible departure times, as determined based on allowed flexibility parameter (FDB in the model). All variables, parameters and constraints of order-free approach are the same as in the same-order approach, except constraints presented in Eqs. (7)–(9) which should be replaced by the following modified constraints (7.a)–(9.a), respectively:

$$\begin{aligned} &XDT_t^i \geqslant XDT_p^i + H(T_p) + H(T_t) + \left(TR_p^{ij} - TR_t^{ij}\right) \\ &where\ (U_t \times U_p = 1) \cap \boxed{\left(DT_t^i - FDB_t^i > DT_p^i - FDB_p^i\right)} \cap \left(TR_p^{ij} \geqslant TR_t^{ij}\right) \cap \left(MR_p^{ij} = MR_t^{ij}\right), \\ &\forall t, \quad p \in T, \quad t \neq p, \quad \forall i,j \in S, \quad |i-j| = 1 \end{aligned} \tag{7.a}$$

$$\begin{aligned} &XDT_t^i \geqslant XDT_p^i + H(T_p) \\ &where\ (U_t \times U_p = 1) \cap \boxed{\left(DT_t^i - FDB_t^i > DT_p^i - FDB_p^i\right)} \cap \left(TR_p^{ij} < TR_t^{ij}\right) \cap \left(MR_p^{ij} = MR_t^{ij}\right), \\ &\forall t, \quad p \in T, \quad t \neq p, \quad \forall i,j \in S, \quad |i-j| = 1 \end{aligned} \tag{8.a}$$

$$\begin{aligned} &XDT_t^i \geqslant XDT_p^j + H(T_p) + TR_p^{ji} \\ &where\ (U_t \times U_p = -1) \cap \boxed{\left(DT_t^i - FDB_t^i \geq DT_p^j - FDB_p^j\right)} \cap \left(MR_p^{ji} = MR_t^{ij}\right), \\ &\forall t, \quad p \in T, \quad t \neq p, \quad \forall i,j \in S, \quad |i-j| = 1 \end{aligned} \tag{9.a}$$

The constraints used for order-free approach (7.a)–(9.a) are similar to those of the same-order approach (7)–(9), presented previously, but the flexibility of early departure times (FDB) is incorporated in the constraints (highlighted by boxes above) to identify the train that is more likely to depart earlier. These constraints allow for changing the departure order of trains as part of the solution for the order-free approach, based on the user defined flexibility value (FDB) incorporated in the constraints.

The ability to modify the order of trains may allow a higher compression level, although the new schedule may also face a station capacity shortage if too many trains try to pass or stop at the same time. In current model, these situations must be reviewed manually.

4. Testing HOTS model in different applications

Various single and multiple-track case study scenarios were developed to test the HOTS model performance and capabilities. All scenarios were modeled using same-order approach. Order-free approach was only performed on selected scenarios. All scenarios were initially developed in RailSys and the final results after HOTS optimization were validated in RailSys. The databases for all scenarios were developed in Microsoft Excel®. LINGO 14® was used as the optimization solver. As the objective was to provide approximately equal weights for the dwell time and departure time coefficients (α_1 and α_2), an iterative calibration process was used to assign the values "50" and "1", respectively.

4.1. Single track case study

A single-track test case study used an actual rail line in the U.S., currently used for excursion passenger trains. The modeled track mimicked the existing infrastructure, but more complicated train and signal parameters were created for the case study. The case study includes a 30-mile long single-track segment with two sidings and a yard for meet/pass and stop purposes. Four types of trains were considered in the case study: intercity passenger (4 daily pairs or northbound/southbound), commuter (2 daily pairs), merchandise freight (2 daily pairs) and intermodal freight trains (3 daily pairs). There were no planned stops for any trains, but trains were allowed to stop at the sidings/yard for train meets/passes, as necessary. There

were no predefined arrival/departure timetables in the case study, although some preferred departure times were defined for each scenario. Table 3 summarizes the case study parameters.

HOTS performance was tested on three single-track track scenarios (Table 3). Scenario 1 used a timetable with several conflicts as an initial schedule (timetable) and HOTS was used to both solve the conflicts and to compress the timetable. Scenario 2 ran RTC simulation over the initial schedule from Scenario 1 to get a conflict-free timetable, which then compressed by HOTS model. Scenario 3 compares the performance of HOTS compression with that of RailSys. Thus, RailSys was used to adjust the conflict-free timetable obtained from RTC (the input timetable of Scenario 2) by using its timetable compression technique. Then, the resulting RailSys timetable was compared with the HOTS results over the same RTC timetable, but an additional attempt was also made to further compress the RailSys timetable in HOTS to compare their compression algorithms. All scenarios used the same-order pattern, except Scenario 1, which used both same-order and order-free patterns. FDB, FDA and maximum allowed dwell time flexibility parameter values differed between scenarios, based on the objective of each scenario, while the rest of the flexibility parameters (e.g. minimum allowed dwell time) remained the same. More details on each scenario are provided in Table 4.

4.1.1. Scenario 1: Compressing a timetable with several conflicts

The first scenario tested HOTS' performance in solving the conflicts and compressing the timetable over an initial schedule (timetable) with several conflicts. Table 5 summarizes the user-defined model parameters for the scenario. The flexibility parameters, such as FDB and FDA were considered equal at all stations.

The adjusted timetable departure and dwell times were generated by LINGO (using a PC, Intel Core 2 Duo, 2 GB RAM) in less than four seconds for both the same-order (4114 constraints, 7984 non-zero parameters, 220 variables, 271 solver iterations) and order-free (4115 constraints, 7986 non-zero parameters, 220 variables, 282 solver iterations) approaches. Then the output from LINGO was validated via RailSys simulation. Fig. 5 shows the RailSys stringlines (graphic time-distance representation of timetable information) from the initial timetable with conflicts in Fig. 5a, and the validated HOTS model results for both same-order, Fig. 5b, and order-free, Fig. 5c, approaches. More than 25 initial schedule conflicts were resolved in both same-order and order-free approaches.

In the same-order approach all commuter (orange[3]), intermodal (dark blue) and freight trains (blue) departed after the first passenger train (yellow), with FDB set at zero, although these trains could have departed earlier. In order-free approach, the FDB parameter was assumed to be zero for the passenger trains, while commuter, intermodal and freight trains were allowed to depart up to 90 min earlier than the initial schedule with no dependency on the passenger train schedule. These changes on FDB values in order-free approach led to changes in the order of some trains. For instance, as highlighted in Fig. 5, in order-free approach passenger trains (yellow) were moved to depart after two commuter trains (Trains #1 and #2). The timetable duration in the order-free approach was approximately 30 min shorter than in same-order approach, but with additional stops.

4.1.2. Scenario 2: Compressing a conflict-free schedule developed by RTC

In Scenario 2, the initial schedule (timetable) of Scenario 1 was first simulated in RTC to provide a conflict-free timetable and HOTS model was then used to perform the compression on the conflict-free timetable. The resulting timetable had a poor level of service (e.g. long waiting time at certain stations). No manual intervention was done during the RTC simulation to improve its conflict resolution. HOTS optimization followed the same-order steps described in Scenario 1 approach. The flexibility parameters were changed in this scenario to provide a better LOS. Table 6 summarizes the user-defined parameters of the HOTS model for Scenario 2.

Fig. 6 presents the results of the conflict-free timetable developed by RTC (replicated in RailSys for consistent graphical representation) and the compressed timetable by the HOTS model. The HOTS model compressed the timetable by approximately one hour and reduced the maximum dwell time at stations from 61 to 30 min and total dwell time from 271 to 166 min.

To evaluate the station capacity limitations of the HOTS model, station "ST2" was allowed to receive only two trains at the same time. As highlighted in Fig. 6b, three trains either pass or stop at "ST2" around 9:30 am, which exceeds the capacity of the station. The capacity issue was solved by departing the third train (train "A") after train "B", and modified input was used to rerun the HOTS model and update the timetable. Fig. 6c presents the second round of the HOTS model results, with changes on the stop patterns (trains "A" and "C") highlighted. The capacity shortage at station "ST2" was resolved in the second round, while stop patterns and departure order were maintained for all other trains. The overall timetable duration was increased by approximately 20 min from the schedule before evaluating station capacity limits (Fig. 6b), due to the fact that trains "A", "C" and all trains after "C" were departed 20 min later to address the station capacity shortage. However, the overall timetable duration of the new schedule (Fig. 6c) was still approximately 45 min shorter than the RTC timetable (Fig. 6a).

Table 7 provides a comparison of results after the HOTS model application. According to Table 7, the HOTS model could reduce the total and maximum dwell times. It also decreased the duration of the timetable, thus providing better capacity utilization.

[3] For interpretation of color in Figs. 5–7 and 9–12 , the reader is referred to the web version of this article.

Table 3
Details of case study infrastructure.

Segment length	30 miles, single track
Sidings/yards	2 sidings + 1 yard
Trains	11 (north) + 11 (south)
Traffic type	Mixed traffic (passenger, commuter, freight, intermodal)

Table 4
Single-track case study scenarios.

Scenario	Initial timetable used in each Scenario	Functionality tested	Rescheduling pattern	Level of flexibility parameters assigned
1	Initial schedule with several conflicts (Worst case)	Train conflict resolution and compression	Same-order & Order-free	High
2	A conflict-free schedule developed by RTC (a functional schedule)	– Timetable compression technique – Station capacity limitation of HOTS	Same-order	Low
3	RTC schedule adjusted by RailSys (an optimized schedule)	Timetable compression technique" performance between RailSys and HOTS model	Same-order	Same as RailSys

Table 5
Parameters for the HOTS model in Scenario 1 of single-track (timetable with conflicts).

Parameter	Passenger	Commuter	Intermodal	Freight
Min. allowed dwell time (min)	0	0	0	0
Max. allowed dwell time (min)	10	5	20	60
FDB (min)	0	300	300	300
FDA (min)	240	240	240	240
Headway (min)	2	2	2	2
Priority of train	3	4	2	1

4.1.3. Scenario 3: Comparing the timetable compression technique of RailSys and HOTS

Scenario 3 was developed to compare the performance of HOTS compression technique with that of RailSys. For the comparison, the UIC timetable compression technique, under the Austrian Rail Network operator (OBB) algorithm, and a 10-min maximum allowed dwell time were used to adjust the RTC timetable (the input timetable of Scenario 2) in RailSys. Overtaking was allowed at stations (Fig. 6a). More details on RailSys compression steps and results can be found in a paper by Pouryousef and Lautala (Pouryousef and Lautala, 2013). The compression was repeated in the HOTS model with the same parameter settings, although in RailSys the dwell time flexibility and departure flexibility parameters could not be defined with as much detail as in HOTS model. Table 8 summarizes the parameters used in Scenario 3. As shown, the departure time flexibility values (FDA and FDB) varied between different train types. However, no departure time flexibility was allowed for trains at the origin stations.

Fig. 7 presents the RTC timetable (Fig. 7a) and the compressed timetables developed by RailSys and the HOTS model (Fig. 7b and c). The main differences were related to stop patterns and total dwell times of trains. According to Table 9, the HOTS model was able to provide a timetable duration approximately 36 min shorter than RailSys, but with slightly more stops (11 vs. 9). The results also show that both HOTS and RailSys could significantly improve the LOS parameters from the RTC timetable, mainly in the form of a sizeable reduction of maximum and total dwell times. However, HOTS and RailSys were unable to reduce the duration of RTC timetable in Scenario 3, as the RTC timetable had a congested schedule and the maximum allowed dwell times in Scenario 3 were lower than in the (successfully compressed) Scenario 2.

In addition to direct comparison of compression results, an additional step was taken to compress the RailSys adjusted timetable (Fig. 7b) in HOTS. The objective was to determine whether HOTS compression technique could provide any further improvement; however, no further improvement was identified, as the results were nearly identical to the RailSys timetable.

4.2. Multiple-track case study

A segment of North-East Corridor (NEC) between Washington, D.C. and Baltimore, MD was used to test the HOTS model performance in an N-track network (double and multiple-track situations). The case study is comprised of a 40.6-mile long multiple-track segment with several stop points and crossovers. It is currently operated based on a bi-directional pattern where trains use all tracks in both directions as necessary. It should be noted that the study uses the Washington, D.C. – Baltimore, MD segment of the NEC as a stand-alone segment and does not examine continuation of routes on either end. The objective of this analysis was not to evaluate or recommend any changes to current NEC operations, but rather to better understand the impacts of different operation philosophies along an example multiple-track corridor. Since the case study did not consider the movement of trains beyond the study limits, none of the suggested modifications are implementable

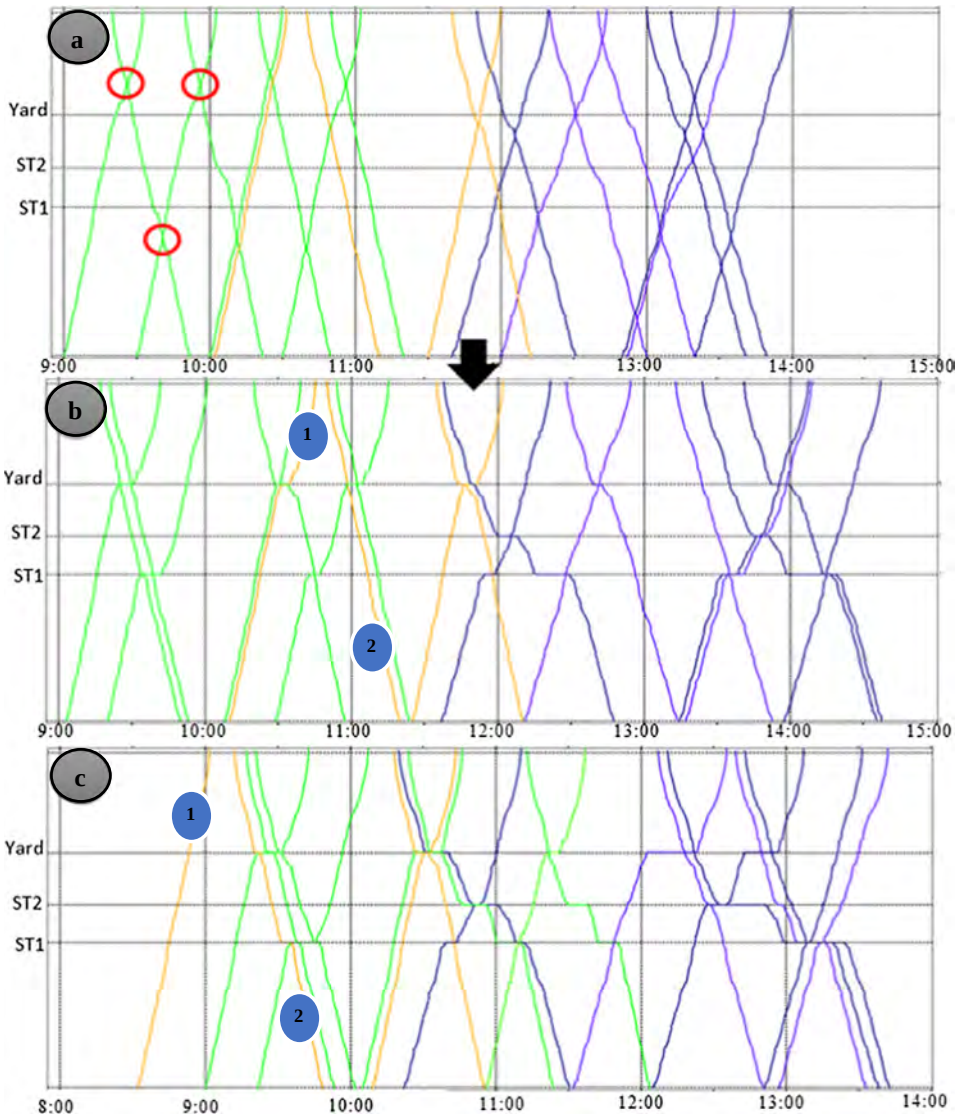

Fig. 5. Results of Scenario 1 of single-track case study. (a) The initial timetable (stringline) with several schedule conflicts (three marked examples), compressed timetable after the HOTS optimization: (b) same-order approach and (c) order-free approach.

Table 6
Parameters for the HOTS model in Scenario 2 of single-track (RTC timetable with no conflict).

Parameter	Passenger	Commuter	Intermodal	Freight
Min. allowed dwell time (min)	0	0	0	0
Max. allowed dwell time (min)	10	10	30	30
FDB (min)	60	60	180	180
FDA (min)	240	300	300	300
Headway (min)	2	2	2	2
Priority of train	3	4	2	1

without further study that evaluates the impacts and challenges over the entire length of the corridor. More details about capacity and allocation issues of NEC can be found in papers developed by Pouryousef and Lautala (2015a) and Pena-Alcaraz et al. (2015).

A track schematic of the case study infrastructure, including the main track, platforms, turnouts and crossovers, is shown in Fig. 8.

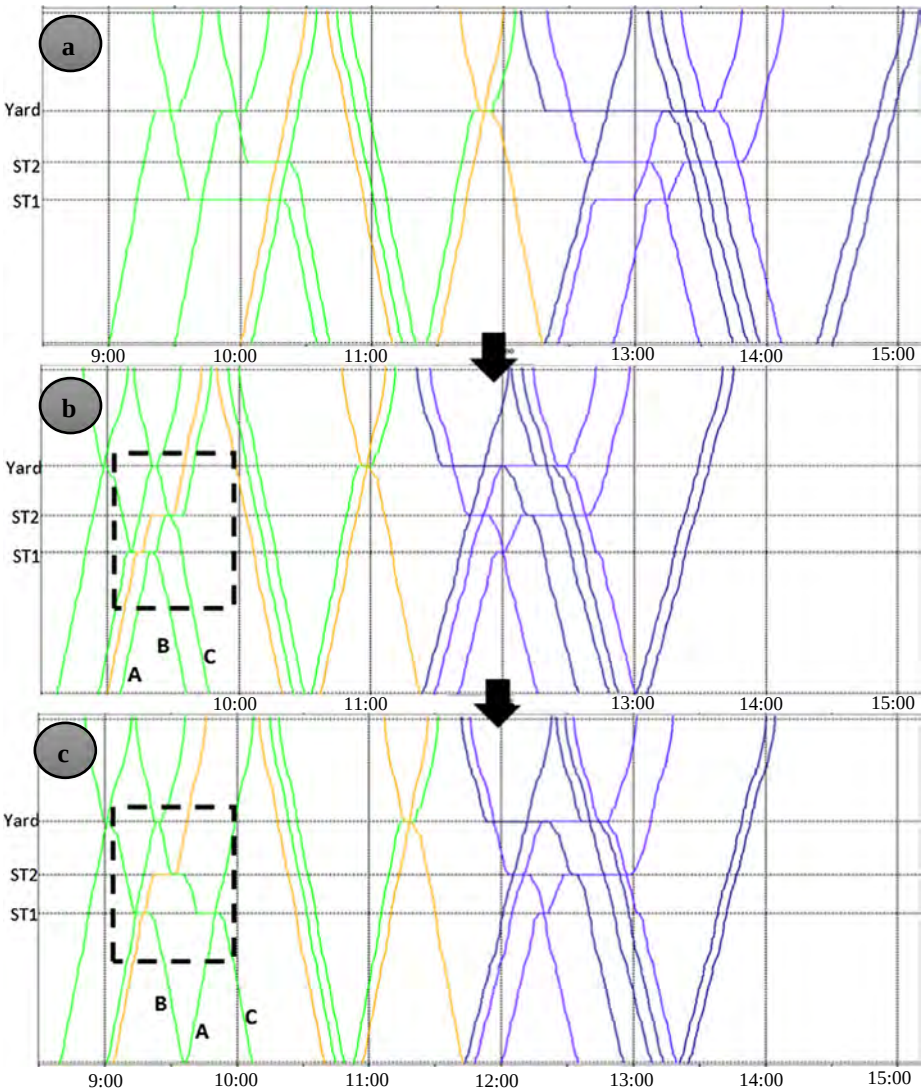

Fig. 6. Results of Scenario 2 of single-track case study, (a) The conflict-free timetable (stringline) developed in RTC with no manual improvement, (b) compressed timetable using the same-order approach in the HOTS model, (c) readjusted timetable after running the HOTS model for a second time to address the assumed station capacity limitations in ST2 siding.

Table 7
Comparison between initial and compressed timetables developed by the HOTS model in Scenario 2 of single track case study (same-order approach).

Parameter		Scenario 2		
		Initial timetable	HOTS timetable before readjustment to address station capacity issue	HOTS timetable after readjustment to address station capacity issue
LOS	Number of stops	14	20	19
	Min. dwell time (min)	0	0	0
	Max. dwell time (min)	61	30	30
	Total dwell time (min)	271	152	166
Capacity	Timetable duration	6 h 10 min	5 h 05 min	5 h 25 min
	Timetable Compression Level	–	65 min	45 min
		–	17%	12%

The segment is one of the most congested and complicated corridors in the U.S. rail network. Four types of trains (total of 136) operate in the corridor, consisting of Acela Express (16 daily pairs), commuter (28 daily pairs), long-distance Amtrak (7 daily pairs) and regional Amtrak trains (17 daily pairs). Since trains are operated under a bi-directional operation pattern,

Table 8
Parameters for HOTS model in Scenario 3 of single-track case study.

Parameter	Passenger	Commuter	Intermodal	Freight
Min. allowed dwell time (min)	0	0	0	0
Max. allowed dwell time (min)	10	10	10	10
FDB – origin station (min)	0	0	0	0
FDA – origin station (min)	0	0	0	0
FDB – other stations (min)	60	60	180	180
FDA – other stations (min)	240	300	300	300
Headway (min)	2	2	2	2
Priority of train	3	4	2	1

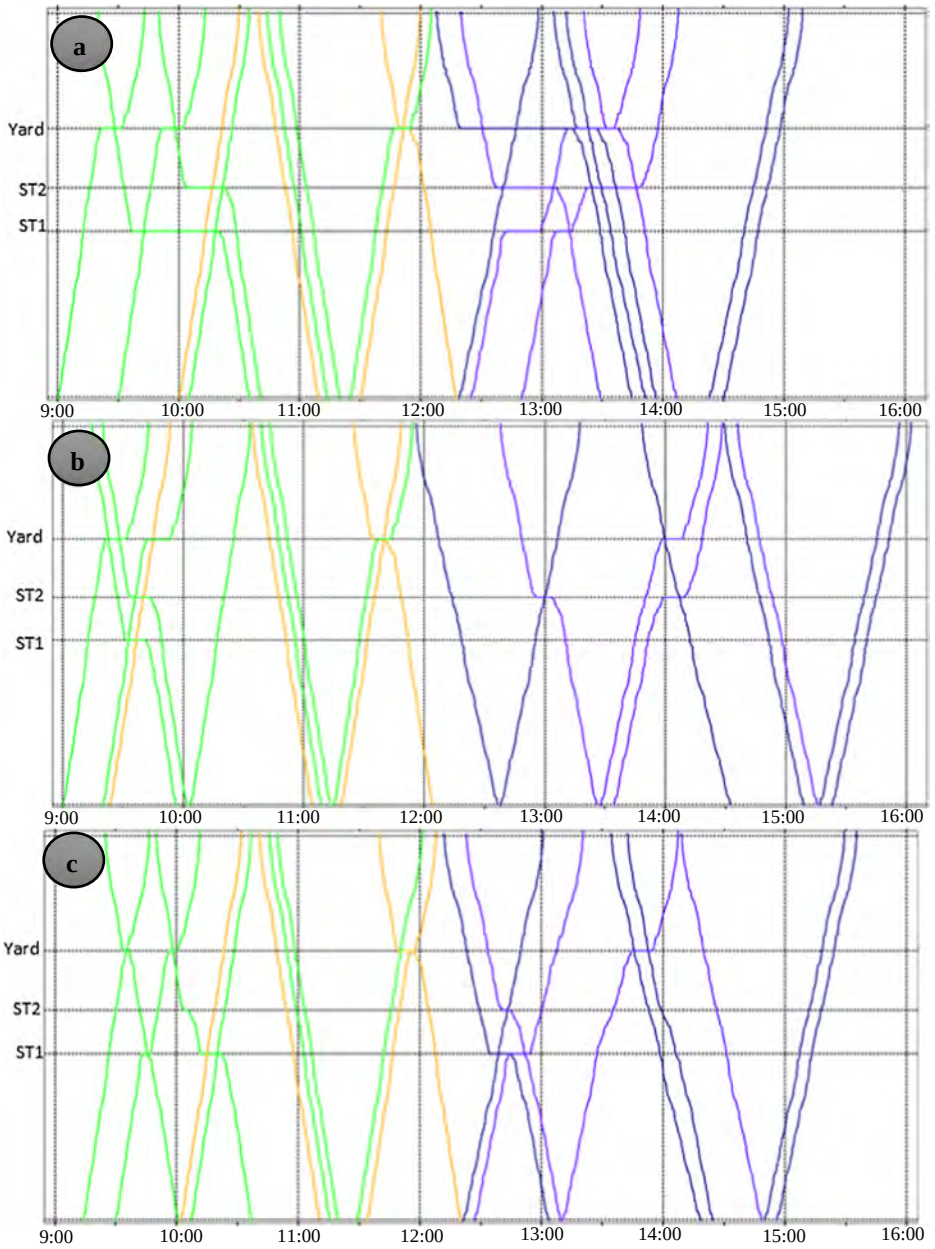

Fig. 7. Results of Scenario 3 of single-track case study, (a) The conflict-free timetable (stringline) from RTC, (b) adjusted by RailSys and (c) adjusted by HOTS model.

Table 9
Comparison between RTC timetables and adjusted timetables by RailSys and HOTS model in Scenario 3 of single track case study (same-order approach).

Parameter		Scenario 3		
		Initial timetable	Adjusted by RailSys	Adjusted by HOTS
LOS	Number of stops	14	9	11
	Min. dwell time (min)	0	0	0
	Max. dwell time (min)	61	10	10
	Total dwell times (min)	271	80	66
Capacity	Timetable duration	6 h 10 min	7 h 04 min	6 h 28 min
	Timetable Compression Degree	–	–	36 min
			–	8%

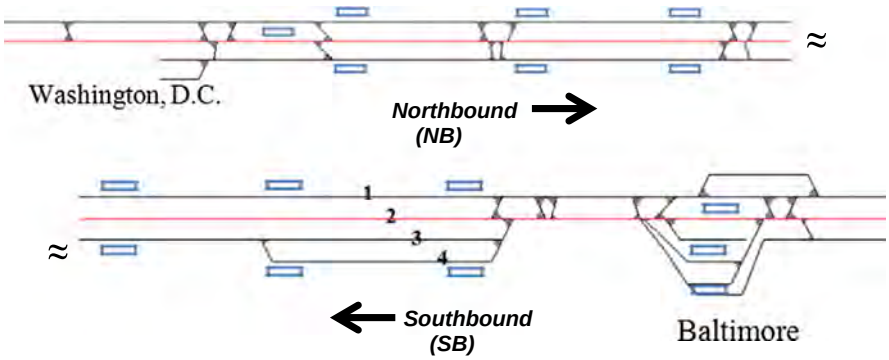

Fig. 8. Case study infrastructure between Washington, D.C. – Baltimore including the tracks, platforms, and crossovers along the corridor.

trains regularly switch between tracks via crossovers, creating 28 different route configurations for the case study. Table 10 presents a summary of the case study's characteristics.

An RTC database that included infrastructure, signaling, rolling stock and operation characteristics was received from Amtrak and later replicated in RailSys for analysis and comparison (Pouryousef and Lautala, 2014, 2015a). Similar to the single-track case study, three scenarios were created for the multiple-track case study.

Scenario 1 used a conflict-free schedule under bi-directional operation pattern to evaluate HOTS model capability to reschedule and compress an initial timetable of a multiple-track corridor with bi-directional operation philosophy. For Scenario 2, the HOTS model results from Scenario 1 were used as the initial timetable, but with different routing alternative for a selected train to examine the capability of HOTS model to provide a conflict-free schedule after new routing. Scenario 3 used a conflict-free schedule of the same case study to evaluate HOTS rescheduling and compression capability, but under a directional operation pattern. Additional details about each scenario are provided in Table 11. It should be noted that only flexibility parameter values (FDB and FDA) and maximum allowed dwell time differed between scenarios. The rest of the flexibility parameters (e.g. minimum allowed dwell time) remained the same.

4.2.1. Scenario 1: Compressing a conflict-free schedule with bi-directional pattern

The purpose of Scenario 1 was to evaluate whether the HOTS model is capable of compressing an initial timetable through rescheduling in a multiple-track case study with bi-directional routing patterns. Table 12 presents main parameters of the HOTS model defined for the scenario. Stop patterns and the minimum allowed dwell time of trains were maintained identical with the initial timetable, and unplanned stops were not allowed (flexibility parameters were assigned "zero" at unplanned stop points). Flexibility parameters (FDB and FDA) were assumed to be the same at all stations, but the FDB parameter was zero for the first train at the origin station. This allowed direct comparison between the schedules, as they all started at the same time. Also, the Acela and Commuter trains could be departed up to 30 min earlier, while regional and long-distance trains could be departed up to 90 min earlier in the same-order approach.

The adjusted timetable was generated by LINGO in less than one minute for the same-order approach model (231,579 constraints, 460,300 non-zero parameters, 2720 variables, and 302 solver iterations). The results obtained from LINGO were validated in RailSys, using the same validation process as in the single-track case study. A two-hour segment of the initial timetable is presented in Fig. 9a, and the rescheduled timetable obtained from the HOTS model is shown in Fig. 9b. Since Acela trains (Red) had higher priority, the model attempted to first reschedule them as early as possible (up to 30 min earlier), and then other trains were rescheduled to follow Acela trains while maintaining their initial order. Selected trains are identified in Fig. 9 to demonstrate the train order and the level of timetable compression. Overall, the HOTS model was able to compress the initial timetable by 48 min, while maintaining the initial departure order, routings, stop patterns, and minimum allowed dwell times of all trains.

Table 10
Details of multiple-track case study (Baltimore – Washington, D.C.).

Corridor length	40.60 miles
Length of double track	1.48 miles
Length of triple track	33.94 miles
Length of quadruple track	5.18 miles
Sidings/yards	2 main yards + 8 station/crossover
Number of trains	68 (North) + 68 (South)
Operation pattern	Bi-directional
Number of different routes	28 routes

Table 11
Details of the multiple-track case study scenarios.

Scenario	Initial timetable used in each Scenario	Functionality tested	Rescheduling pattern	Level of flexibility parameters assigned
1	A conflict-free schedule (*bi-directional pattern*)	Rescheduling and compressing the timetable	Same-order	Varied for each type of train
2	HOTS model schedule (*Results of Scenario1*) with new routing	Resolving any potential conflict after assigning a new routing alternative	Same-order	Same as Scenario1, (except for the rerouted train)
3	A conflict-free schedule (*directional pattern*)	Rescheduling and compressing the timetable	Same-order & Order-free	– "Same-Order": Same as Scenario1, – "Order-Free": Higher

Table 12
Parameters for the HOTS model in Scenario 1 of multiple-track case study.

Parameter	Acela	Commuter	Long-distance	Regional
Min. allowed dwell time (min)	1	1	1	1
Max. allowed dwell time (min)	2	2	2	2
FDB (min)	30	30	90	90
FDA (min)	30	30	90	90
Headway (min)	2	3	3	3
Priority of train	4	2	1	1

4.2.2. Scenario 2: Resolving the conflict after rerouting a train

The purpose of this scenario was to examine the capability of the HOTS model to provide a new conflict-free schedule while allowing new routing for a selected train (or several trains). Train rerouting is a common practice on double and multiple track corridors, but introducing one or more new routes may cause schedule conflicts, potentially making rerouting a complex and laborious process. Although HOTS model cannot provide simultaneous rerouting and rescheduling, it is still capable to reschedule the trains, if the routing information is changed in the model input. In this scenario, Train #2 was randomly selected from the adjusted timetable (Fig. 9b) and rerouted to match the route of Train #5. As highlighted in Fig. 10a, if both trains maintain their current schedule, there will be a conflict along the main line. The situation was resolved by defining a new route and higher departure flexibility for Train #2 while the remaining parameters of HOTS model were considered the same as in the Scenario 1. The HOTS model was able to provide a conflict-free timetable with new routing defined for Train #2, while maintaining the schedule before Train #2 (Fig. 10b). In addition to removing conflict between Train #2 and Train #5, six other trains (all departing after Train #2) were rescheduled by the model. The overall duration of timetable was the same as in Scenario 1.

4.2.3. Scenario 3: Compressing a conflict-free schedule with directional pattern

The purpose of this scenario was to evaluate the HOTS model performance in rescheduling and compressing timetable of a multiple-track corridor with directional operation pattern. As mentioned earlier, the NEC corridor is operated bi-directionally. Therefore, the initial train schedule was first converted to a fully directional operation on two tracks (each track was respectively assigned to only one direction of operation). More details about the heuristic methodology and conversion outcomes are provided in a paper by Pouryousef and Lautala (2015a).

Same parameters as in Scenario 1 (Table 12) were used. Fig. 11a presents a two-hour segment of the initial timetable under directional operation pattern and Fig. 11b depicts the compressed version using same-order approach of HOTS model. Selected trains are numbered in Fig. 11 and before/after comparison reveals that all trains have been shifted to the left (compressed) while maintaining the same train departure order. Overall, the HOTS model was able to compress the initial timetable under directional operation pattern by 50 min, while maintaining the initial departure order, stop patterns, and minimum allowed dwell times of all trains.

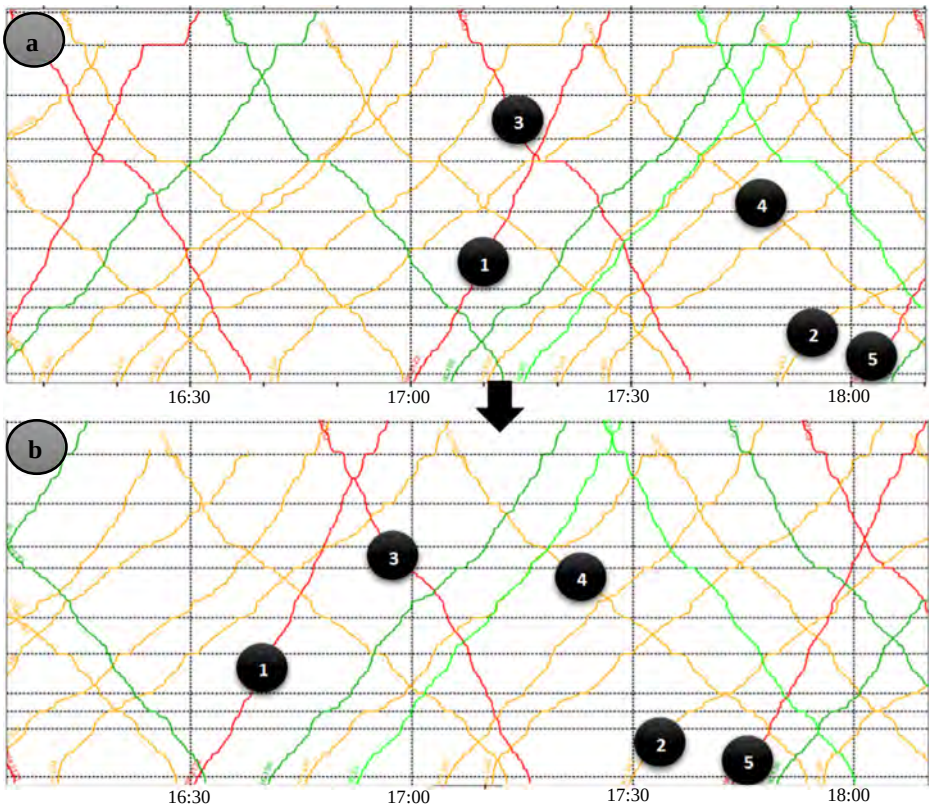

Fig. 9. Results of Scenario 1 of multiple-track case study (a) Initial and (b) rescheduled timetable (stringline) of NEC corridor based on same-order approach (specific trains are labeled in both figures for comparison purposes. Red: Acela, Orange: Commuter, Green: Long-distance and Regional trains).

In second part of this scenario, the initial timetable was rescheduled under directional pattern by using the order-free approach of HOTS model. The adjusted timetable was generated by LINGO (238,660 constraints, 474,462 non-zero parameters, 2720 variables, and 615 solver iterations). FDB and FDA for Acela trains had departure time deviation of maximum 15-min from their initial departure times, while other trains had higher flexibility values (for instance up to 90-min for long distance trains). After HOTS model was not able to find a feasible solution under directional operation pattern with two-minute maximum allowed dwell time (as assumed in first part of the scenario), the maximum allowed dwell time was increased to 10-min. Several trains have been rescheduled and reordered, as shown in Fig. 12. For instance, the schedule and the order of Trains #2 and #5 have been switched due to the higher values of FDB for these two trains; while other highlighted trains (mainly commuter and Acela trains with low FDB values) experienced only minor schedule changes.

As marked in Fig. 12 by the rectangles, the 10-min dwell time allowed the HOTS model to provide an alternative for some of the faster trains (mainly Acela) to overtake the slower trains (mainly commuter) at the stations. Overall, the new directional operation schedule developed by the HOTS model exhibits less compression than the previous scenario (same-order approach). This was caused by train shuffling based on a broad range of FDB values that increased the heterogeneity level.

The order-free approach of the HOTS model was also tested on the initial schedule of Scenario 1. All 28 initial routing alternatives of trains were updated and redefined in the model database to provide a fully directional pattern by using only one routing alternative for northbound and southbound trains, respectively. Similar to Scenario 2, the new routing caused serious schedule conflicts between trains which were resolved by applying the HOTS model. The outcomes of this scenario after rescheduling and resolving all potential conflicts were similar to Scenario 3, part 2.

Table 13 summarizes the outcomes of each tested scenario in the multiple-track case study. For the scenarios under bi-directional operation pattern, the HOTS model was able to reduce the total dwell time and timetable duration while maintaining the number of stops and min/max allowed dwell times. In the scenarios under directional operation pattern (Scenario 3, parts 1 and 2), the HOTS model was able to compress the timetable under the same-order approach, while the order-free approach produced a lower degree of compression, as the heterogeneity level of the train schedule was increased after shuffling the trains. The results highlight that in the order-free approach of rescheduling (as discussed in Scenario 3 as well), choosing right values of the flexibility parameter is an important task that can substantially change the heterogeneity level of the train schedules.

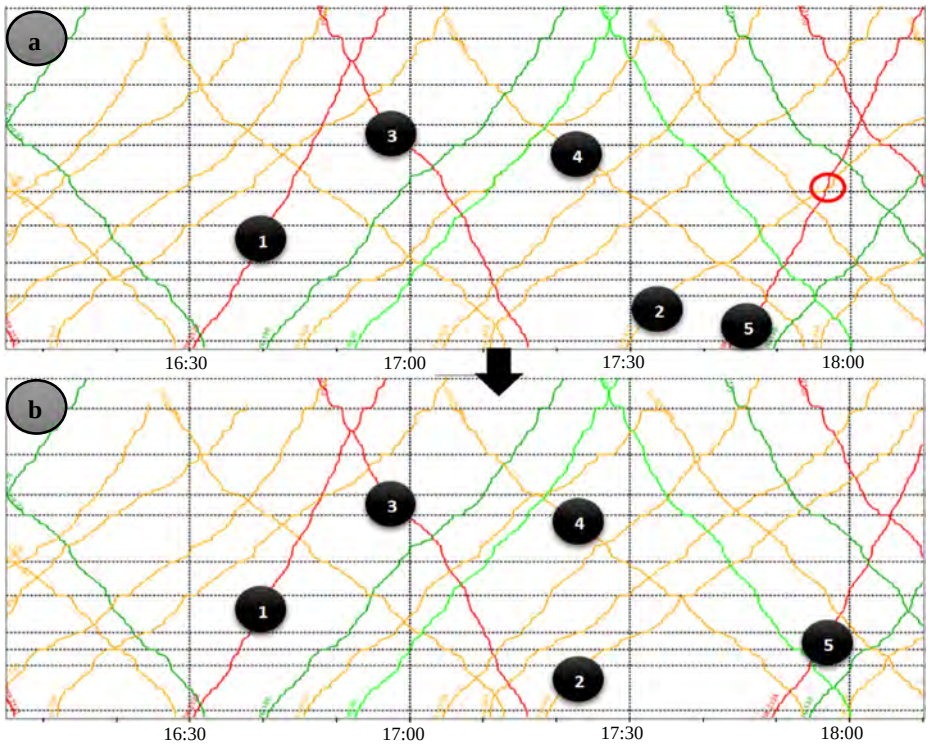

Fig. 10. Results of Scenario 2 of multiple-track case study (a) Timetable (Stringline) developed in Scenario 1, was rescheduled by the HOTS model (b) to address the new route defined for Train #2.

4.3. Sensitivity analysis of alpha coefficients

As pointed out in Section 3.2.2, there is a sizable difference in the numeric values for the two variables minimized by the objective function. The numeric value of dwell time deviation (first part of the objective function) is much smaller than the numeric value of train departure time (second part). The α_1 and α_2 coefficient parameters were included in the model objective to indicate the importance of dwell time versus departure time, respectively, and to allow the user to adjust the weighting preferences by changing these coefficients. To evaluate the weighting impact of α_1 and α_2 on the model results, an iterative sensitivity analysis was performed for the first scenario of the single-track case study (both same-order and order-free approaches). In the analysis, all values of α_1 and α_2 were normalized to be between 0 and 1 ($\alpha_1 + \alpha_2 = 1$), and seven different combinations were assigned to compare the main outputs of the HOTS model (number of stops, max. dwell time, total dwell time, and timetable duration). Tables 14 and 15 summarize the outcomes of sensitivity analysis.

Tables reveal that changing the coefficient values did not affect the maximum dwell time obtained from the model (20 min in all cases, although for freight trains it could have been increased to max. 60 min; see Table 5). The timetable duration had minimal variation for the same-order approach (values between 334 and 338 min), and remained constant (410 min) for the order-free approach. We hypothesize that the type of initial timetable (the congestion level, and initial departure time of trains), and the flexibility parameters to dispatch trains earlier and later (FDB and FDA), have greater impact on the timetable duration and maximum dwell time than changing the Alpha coefficients, but testing the hypothesis is beyond the scope of this research.

On the other hand, Tables 14 and 15 also reveal that both number of stops and total dwell time parameters were sensitive to changes in the Alpha coefficient values. Fig. 13 highlights the impact of Alpha coefficients on these two parameters for both same-order and order-free approaches.

As shown in Fig. 13, both the number of stops and total dwell time are non-sensitive when α_1 is much smaller than α_2 ($\alpha_1 < 0.25, \alpha_2 > 0.75$). However, increasing the value of α_1 above 0.25 gradually reduces both the number of stops and total dwell time. When considering the fact that timetable duration and max dwell time were nearly constant for all cases (Tables 14 and 15), it can be concluded that to minimize the number of stops and total dwell time in the single-track case study, the value of normalized α_1 should be between 0.75 and 0.9 (i.e., $0.1 \leqslant \alpha_2 \leqslant 0.25$). However, this "preferred" range may not be consistent across other scenarios, because results will likely depend on other parameters such as number of trains, initial departure times, initial stop pattern (number of stops and duration), and diversity and priority of trains to be rescheduled. It should be noted that, excluding this sensitivity analysis, all case study evaluations presented in this paper used coefficient

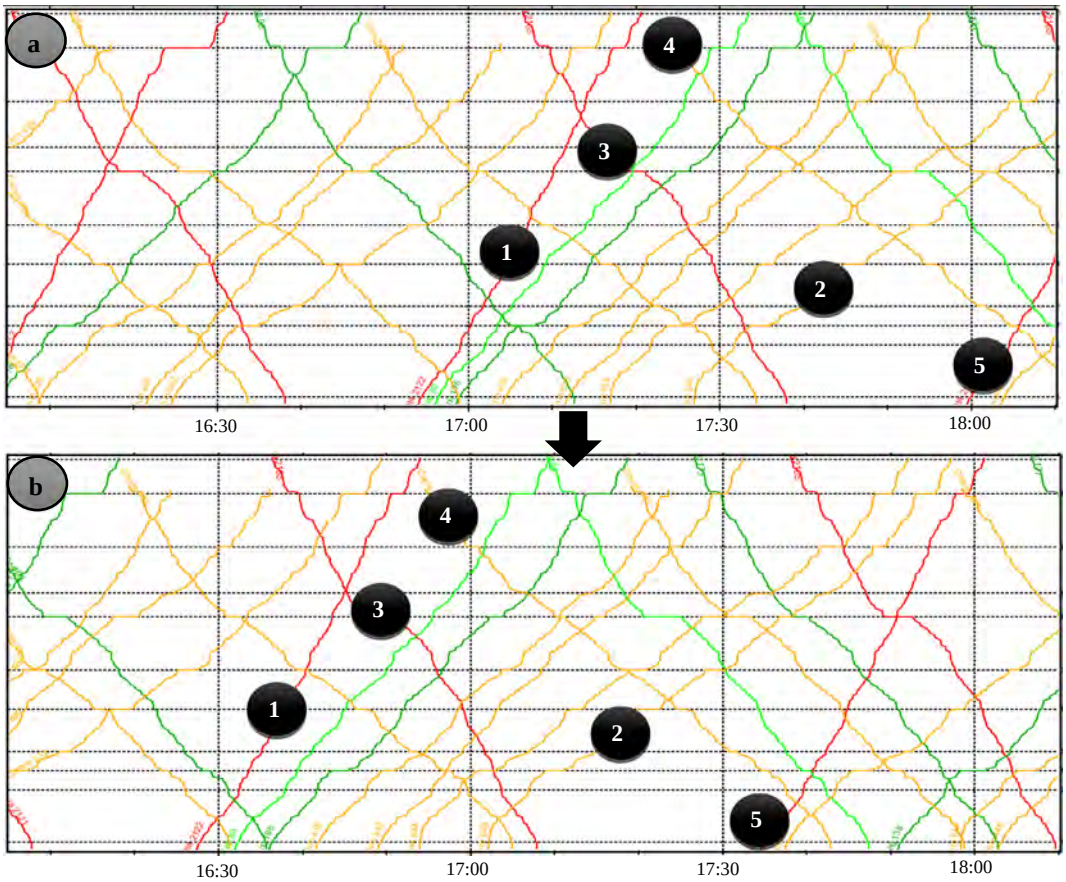

Fig. 11. Results of Scenario 3, same-order approach, of multiple-track case study (a) Initial timetable (stringline) with directional pattern and (b) compressed timetable by HOTS based on same-order approach (selected trains are labeled in both figures for comparison purposes).

values of $\alpha_1 = 50$ and $\alpha_2 = 1$ (or $\alpha_1 = 0.98$, $\alpha_2 = 0.02$ when normalized), which align closely with the preferred values found in the single-track sensitivity analysis.

4.4. Discussion of results

Based on the single-track and multiple-track case study tests, the HOTS model was successful in rescheduling and compressing the timetable of different train types on N-track networks (single, double and multiple track corridors) under both directional and bi-directional operation patterns. The model also provided a "Conflict-Free" train schedules, even if the initially requested schedule had serious conflicts between trains. In addition, the HOTS model could reschedule the trains that were assigned by new train routing scenarios (input of model) for double-, and multiple-track corridors. Finally, the HOTS model was capable of rescheduling trains under both same-order and order-free scheduling approaches, based on user-defined model inputs.

Overall, the hybrid simulation/optimization approach offers several benefits. For example, it allows a simplified dataset for the optimization by partially using the train performance calculator (TPC) outputs of the rail simulation packages. This can facilitate and expedite the process of obtaining the optimum solution by the model. In addition, the hybrid approach allows the use of better data for the optimization, as the extracted data from rail simulation tools offers accurate train travel times, based on TPC outputs. Such detailed train acceleration/deceleration analyses are rarely considered in train scheduling optimization models (including HOTS). Finally, the HOTS model allows for validation of results by commercial rail simulation tool, which increases the credibility of optimization results.

While HOTS model provided satisfactory performance under testing, it is recognized that it currently has certain limitations, such as the lack of automatically evaluating the station capacity limitations, as each stop point/station is considered a single node in the model. This leads to a risk of allowing a train arrival at a station even if all tracks are occupied, which is a specific concern in the order-free approach when a broad range of departure flexibility (FDB) is allowed. Such capacity shortages should become evident during the validation process in simulation/timetable management tools. The HOTS model is also sensitive to the requested departure times, flexibility parameters of departure times (FDB and FDA), and the minimum

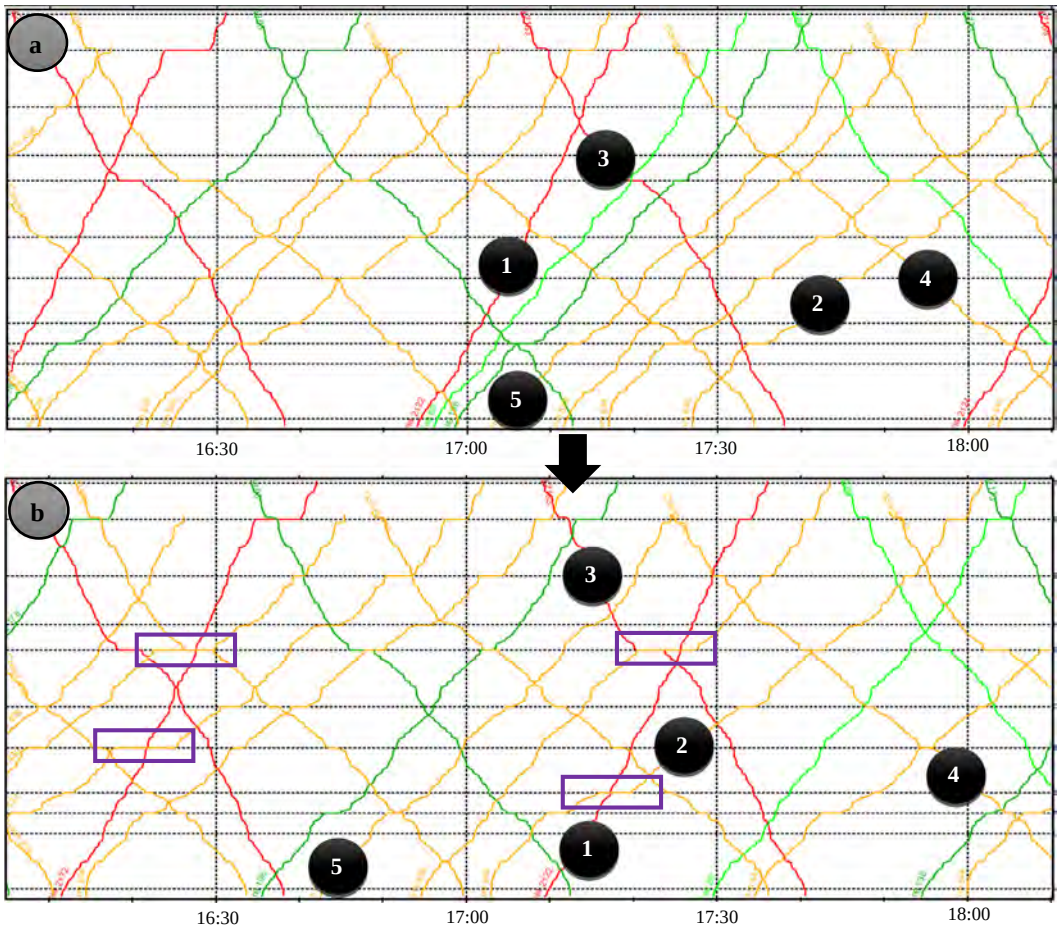

Fig. 12. Results of Scenario 3, order-free approach, of multiple-track case study (a) Initial timetable (stringline) with directional pattern and (b) compressed timetable by HOTS based on order-free approach (specific trains are labeled in both figures for comparison purposes. The rectangles represent the overtaking alternative for faster trains at stations).

Table 13
Comparison between initial timetable and rescheduled timetable developed by the HOTS model in the NEC multiple-track case study.

Parameter		Bi-directional operation pattern			Directional operation pattern		
		Initial timetable	Rescheduled by HOTS (Scenario 1)	Rescheduled by HOTS based on new route (Scenario 2)	Initial timetable	Rescheduled by HOTS (same-order) (Scenario 3) Part 1	Rescheduled by HOTS (order-free) (Scenario 3) Part 2
LOS	Number of stops	402	402	402	402	402	402
	Min. dwell time (min)	1	1	1	1	1	1
	Max. dwell time (min)	3	2	2	2	2	10
	Total dwell times (min)	557	405	405	437	437	581
Capacity	Timetable duration	23 h 46 min	22 h 58 min	22 h 58 min	23 h 47 min	22 h 57 min	23 h 28 min
	Timetable compression degree	– –	48 min 3.3%	48 min 3.3%	–	50 min 3.5%	19 min 1.2%

and maximum dwell times of trains. Reducing the flexibility may prevent the optimization part of HOTS model from finding a feasible solution for all trains. Also, since train acceleration and deceleration parameters are not included in the HOTS model formulation, minor deviations may appear between departure times suggested by the HOTS model versus those

Table 14
Sensitivity analysis between Alpha coefficients and HOTS model results; same-order approach.

α_1, α_2	Number of stops	Max. dwell time (min)	Total dwell time (min)	Timetable duration (min)
0.0001, 0.9999	26	20	227	337
0.1, 0.9	26	20	227	337
0.25, 0.75	26	20	227	337
0.5, 0.5	26	20	216	337
0.75, 0.25	24	20	161	336
0.9, 0.1	24	20	155	334
0.9999, 0.0001	22	20	144	338

Table 15
Sensitivity analysis between Alpha coefficients and HOTS model results; order-free approach.

α_1, α_2	Number of stops	Max. dwell time (min)	Total dwell time (min)	Timetable duration (min)
0.0001, 0.9999	22	20	223	410
0.1, 0.9	22	20	223	410
0.25, 0.75	22	20	223	410
0.5, 0.5	21	20	200	410
0.75, 0.25	19	20	149	410
0.9, 0.1	20	20	141	410
0.9999, 0.0001	21	20	134	410

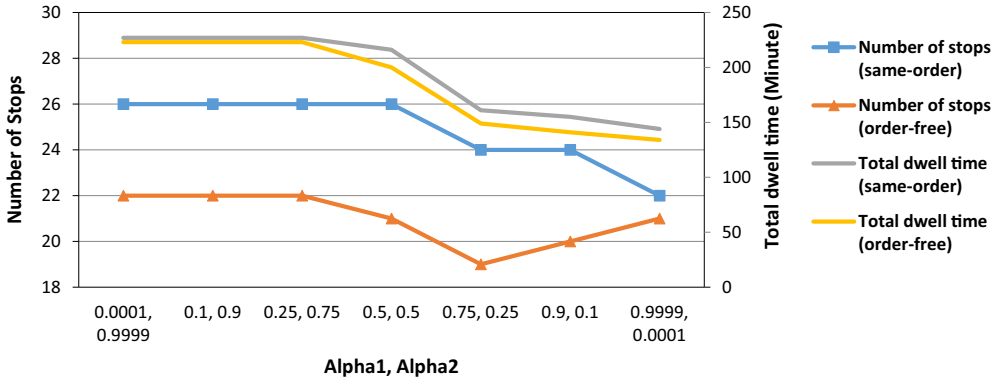

Fig. 13. Sensitivity analysis of Alpha coefficients on the number of stops and total dwell time.

obtained from the simulation package. To minimize this variation, it is important to consider proper train types and characteristics when determining minimum headways in the HOTS model. Finally, HOTS model structure does not offer simultaneous rerouting and rescheduling features, but new train routing alternatives can be defined by the user in the input, and HOTS can reschedule/provide a conflict-free timetable based on new routing alternatives.

5. Summary and conclusions

Rescheduling is one of the main methods to improve the capacity utilization or LOS characteristics of a rail corridor. There are several timetable tools and rail simulation packages that can be used for rescheduling, but no commercial rail simulation could be identified with both (1) automatic train conflict resolution and (2) automatic timetable management features. This is especially the case in tools that target the U.S. rail environment that uses non-timetable based operating principles.

To address some of these limitations, a new rescheduling model (multi-objective linear programming) called "Hybrid Optimization of Train Schedules" (HOTS) was introduced in this paper. HOTS model works together with existing rail simulation tools, extending their capabilities to improve the capacity utilization or the LOS parameters to provide a conflict-free and compressed timetable for N-track rail networks, including single and multiple-track corridors. The optimization part of HOTS model receives several rescheduling parameters from the rail simulation/timetable management tool, in addition to user-defined parameters such as min/max allowed dwell time and train departure flexibility parameters. The objective or function of the HOTS model is derived from a timetable compression technique introduced by UIC. It attempts to compress the train schedules as much as possible by minimizing the departure time and the deviation from minimum allowed dwell time while maintaining a conflict-free schedule. The HOTS model generates two separate output variables called the

proposed departure time and proposed dwell time, which can be validated in the rail simulation/timetable management tools. The HOTS model can be applied for both same-order and order-free rescheduling approaches.

Single-track and multiple-track case studies with six different scenarios were used to examine the capabilities of the HOTS model. In each scenario the HOTS model was able either to compress the timetable and/or reduce the dwell times, or to maintain the performance of initial timetable. In a single-track case study, the HOTS model first resolved the schedule conflicts of an initial timetable using both same-order and order-free rescheduling approaches (Scenario 1), and then compressed a conflict-free timetable, after manual adjustments were made to address station capacity limitations (Scenario 2). For the same case, a direct comparison of the compression techniques of HOTS model and RailSys showed that HOTS model provided comparable stop patterns of train schedules, even though the compression techniques differ between the models. The HOTS model could not further compress an already compressed RailSys timetable (Scenario 3).

In a multiple-track corridor with a bi-directional operation pattern, the HOTS model successfully compressed the initial timetable while maintaining the train routings, order and stop patterns (Scenario 1). In Scenario 2, HOTS provided a conflict-free and compressed schedule after new routes were defined for a selected train. In Scenario 3, HOTS model was able to reschedule and compress an initial timetable under directional operation pattern by using both same-order and order-free approaches. Overall, the same-order approach provided higher capacity utilization, as this approach maintained the homogeneity level.

Finally, a sensitivity analysis was conducted to evaluate the impact of varying the α_1 and α_2 coefficients, used as weights in the objective function of the HOTS model. The analysis demonstrated that these coefficients have a high impact on the number of stops and total dwell time output by HOTS, and for the single-track case study the best results (for the specified criteria) can be obtained when the normalized values of α_1 and α_2 are assigned between ($\alpha_1 = 0.75–0.9$, $\alpha_2 = 0.1–0.25$).

Although the HOTS model performed well, there are current limitations that should be addressed in future research. The exclusion of station capacity limits can be addressed through incorporation of a station capacity constraint. This would make the model more user-friendly and allow it to reach the final solution with a single run. Alternatively, the model can be upgraded from a node-based approach to a link-based approach to reflect the actual track/switch arrangements at stations. The optimization part of the HOTS model has currently been developed based on minimizing a weighted sum of the train departure times and the deviation from minimum allowed dwell time, which forces the trains to be departed as early as possible. In practice, a preference might be to reschedule selected trains to depart as early as possible, while others (e.g. freight trains) might prefer a late departure. This would provide more capacity in the middle of the timetable and could be addressed through a dual-objective algorithm for minimizing the departure time of selected trains while maximizing the departure time of others.

Acknowledgments

This research was supported by National University Rail (NURail) Center, a US DOT-OST Tier 1 University Transportation Center. The authors are grateful for the assistance by Amtrak in providing information and the NEC database for this research. Special thanks for the assistance and feedback by Davis Dure in Amtrak. The authors would also like to thank Eric Wilson (Berkeley Simulation Software, LLC), Daniel Huerlimann (OpenTrack), Sonja Perkuhn (RMCon), and Gabriele Löber (RMCon) for providing academic licenses to simulation packages, RTC, OpenTrack and RailSys, respectively, and for their support for the research team at Michigan Tech. Finally, authors appreciate Mark Wiley (Lindo Systems, Inc.) for providing the academic license of LINGO for the research.

References

Abril, M., Barber, F., Ingolotti, L., Salido, M.A., Tormos, P., Lova, A., 2007. An Assessment of Railway Capacity. Technical University of Valencia, Valencia, Spain.
Amit, I., Goldfarb, D., 1971. The timetable problem for railways. Develop. Oper. Res. 2, 379–387.
Banverket, 2005. Application of the UIC Capacity leaflet at Banverket. Banverket, Swedish National Rail Administration.
Brod, D., Metcalf, A.E., 2014. Web-Based Screening Tool for Shared-Use Rail Corridors (NCFRP Report 27). National Corporative Freight Research Program-Transportation Research Board (TRB).
Burdett, R.L., Kozan, E., 2006. Techniques for absolute capacity determination in railways. Transport. Res. Part B 40, 616–632.
Cambridge Systematics, I., 2006. Statewide Rail Capacity and System Needs Study, Final report. State of Washington-Transportation Commission, Olympia, WA.
Canca, D., Barrena, E., Algaba, E., Zarzo, A., 2014. Design and analysis of demand-adapted railway timetables. J. Adv. Transport. 48, 119–137 (John Wiley & Sons, Ltd.).
Carey, M., Carville, S., 2003. Scheduling and platforming trains at busy complex stations. Transport. Res. Part A 37, 195–224.
Carey, M., Lockwood, D., 1995. A model, algorithms and strategy for train pathing. Oper. Res. Soc. 46, 988–1005.
Corman, F., D'Ariano, A., Hansen, I.A., Pacciarelli, D., 2011. Optimal multi-class rescheduling of railway traffic. J. Rail Transport Plan. Manage. 1, 14–24, Elsevier.
Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M., 2012. Bi-objective conflict detection and resolution in railway traffic management. Transport. Res. Part C: Emerg. Technol. 20, 79–94.
D'Ariano, A., Pacciarelli, D., Pranzo, M., 2008. Assessment of flexible timetables in real-time traffic management of a railway bottleneck. Transport. Res. Part C: Emerg. Technol. 16, 232–245.
Dündar, S., Şahin, İ., 2013. Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways. Transport. Res. Part C: Emerg. Technol. 27, 1–15.
Ghoseiri, K., Szidarovszky, F., Asgharpour, M.J., 2004. A multi-objective train scheduling model and solution. Transport. Res. Part B 38, 927–952, Elsevier.
Gille, A., Siefer, T., 2013. Sophisticated capacity determination using simulation. In: Transportation Research Board 2013 Annual Meeting, Washington, DC, USA.

Goverde, R.M.P., Corman, F., D'Ariano, A., 2014. Investigating on the capacity consumption at Detch railways for varoius signaling technologies and traffic conditions. In: TRB 2014 Annual Meeting, Washington, DC, USA.
Hansen, I.A., Pachl, J., 2008. Railway Timetable & Traffic, Analysis, Modeling, Simulation. Eurailpress, Hamburg, Germany.
Harrod, S.S., 2012. A tutorial on fundamental model structures for railway timetable optimization. Surv. Oper. Res. Manage. Sci. 17, 85–96.
Higgins, A., Kozan, E., Ferreira, L., 1996. Optimal scheduling of trains on a single line track. Transport. Res. Part B 30B, 147–161.
Khadem-Sameni, M., Dingler, M., Preston, J.M., 2010. Revising the UIC 406 method: revenue generating capacity. In: Joint Rail Conference (JRC) 2010, Urbana, IL, USA.
Khadem-Sameni, M., Dingler, M., Preston, J.M., Barkan, C.P.L., 2011. Profit-generating capacity for a freight railroad. In: TRB 90th Annual Meeting, Washington, DC.
Kraay, D., Harker, P.T., Chen, B., 1991. Optimal pacing of trains in freight railroads: model formulation and solution. Oper. Res. 39, 82–99.
Lai, Y.-C., Barkan, C., 2011. A comprehensive decision support framework for strategic railway capacity planning. J. Transport. Eng. – ASCE 137, 738–749.
Landex, A., Kaas, A. H., Schittenhelm, B., Schneider-Tilli, J., 2006. Evaluation of railway capacity. In: Annual Transport Conference. Aalborg University, Denmark.
Lindner, T., 2011. Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. J. Rail Transport Plan. Manage. 1, 49–57, Elsevier.
Meng, L., Zhou, X., 2014. Simultaneous train rerouting and rescheduling on an N-track network: a model reformulation with network-based cumulative flow variables. Transport. Res. Part B: Methodol. 67, 208–234.
Patty, B.W., 2015. Handbook of Operations Research Applications at Railroads. Springer, US.
Pena-Alcaraz, M., Sussman, J.M., Webster, M., 2015. Analysis of Capacity Pricing and Allocation Mechanisms in Shared Railway Systems: Lessons for the Northeast Corridor. Massachusetts Institute of Technology, Engineering Systems Division.
Pouryousef, H., Lautala, P., 2013. Evaluating the results and features of two capacity simulation tools on the shared-use corridors. In: ASME/ASCE/IEEE 2013 Joint Rail Conference (JRC2013), Knoxville, TN, USA.
Pouryousef, H., Lautala, P., 2014. Evaluating two capacity simulation tools on shared-use U.S. Rail Corridor. In: 2014 Annual Meeting of the Transportation Research Board. TRB, Washington, DC.
Pouryousef, H., Lautala, P., 2015a. Capacity evaluation of directional and non-directional operational scenarios along a multiple-track U.S. Corridor. In: TRB Annual Conference of 2015 Washington-DC. Transportation Research Board (TRB).
Pouryousef, H., Lautala, P., 2015b. Hybrid simulation approach for improving railway capacity and train schedules. J. Rail Transport Plan. Manage.
Pouryousef, H., Lautala, P., White, T., 2013. Review of capacity measurement methodologies; similarities and differences in the U.S. and European Railroads. In: 2013 TRB Annual Meeting. Washington-DC.
Pouryousef, H., Lautala, P., White, T., 2015. Railroad capacity tools and methodologies in the U.S. and Europe. J. Modern Transport. 23, 30–42.
Prinz, R., Hollmuller, J., 2005. Implementation of UIC 406 Capacity Calculation Method at Austrian Railway (OBB). Austrian Railway (OBB), Vienna, Austria.
RAS, 2012. 2012 RAS Problem Solving Competition, Movement Planner Algorithm Design For Dispatching On Multi-Track Territories. Railway Applications Section – INFORMS. <https://www.informs.org/Community/RAS/Problem-Solving-Competition/2012-RAS-Problem-Solving-Competition2015>.
RMCon, 2010. RailSys User Manual. Rail Management Consultants GmbH (RMCon), Hannover, Germany.
Schlechte, T., Borndörfer, R., Erol, B., Graffagnino, T., Swarat, E., 2011. Micro-macro transformation of railway networks. J. Rail Transport Plan. Manage. 1, 38–48, Elsevier.
Sogin, S.L., Barkan, C.P.L., Lai, Y.-C., Saat, M.R., 2012. Impact of passenger trains in double track networks. In: 2012 Joint Rail Conference. Philadelphia, Pennsylvania.
Sogin, S.L., Lai, Y.-C.R., Dick, C.T., Barkan, C.P.L., 2013. Comparison of the capacity of single and double track rail lines using simulation analyses. In: Transportation Research Board 92nd Annual Meeting. Washington, DC, USA.
Sun, Y., Cao, C., Wu, C., 2014. Multi-objective optimization of train routing problem combined with train scheduling on a high-speed railway network. Transport. Res. Part C: Emerg. Technol. 44, 1–20.
Szpigel, B., 1973. Optimal train scheduling on a single track railway. Oper. Res. 72, 343–352.
Tornquist, J., Persson, J.A., 2007. N-tracked rail traffic re-scheduling during disturbances. Transport. Res. Part B 41, 342–362.
UIC, 2004. UIC Code 406R-Capacity-1st Edition. International Union of Railways (UIC), Paris, France (June).
White, T., 2005. Alternatives for railroad traffic simulation analysis. J. Transport. Res. Board (TRB) 1916, 34–41.
Zhou, X., Zhong, M., 2005. Bicriteria train scheduling for high-speed passenger railroad planning applications. Eur. J. Oper. Res. 167, 752–771.
Zhou, X., Zhong, M., 2007. Single-track train timetabling with guaranteed optimality: branch-and-bound algorithms with enhanced lower bounds. Transport. Res. Part B: Methodol. 41, 320–341.

Journal of Rail Transport Planning & Management xxx (2015) 1–14

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

journal homepage: www.elsevier.com/locate/jrtpm

Hybrid simulation approach for improving railway capacity and train schedules

Hamed Pouryousef [a, *], Pasi Lautala [b]

[a] Civil and Environmental Engineering Department, Michigan Tech. University, 1400 Townsend Drive, Houghton, MI 49931, USA
[b] Michigan Tech Transportation Institute (MTTI), Michigan Tech. University, 1400 Townsend Drive, Houghton, MI 49931, USA

ARTICLE INFO

Article history:
Received 1 June 2015
Received in revised form 30 September 2015
Accepted 2 October 2015
Available online xxx

Keywords:
Railway simulation
Timetable management
Level of service
Railway capacity
Timetable compression technique

ABSTRACT

Timetable management is one of the operational methodologies commonly applied in the highly structured European rail system to improve the capacity utilization while maintaining acceptable level of service (LOS) parameters; but their potential benefits to the less structured U.S. system have received little attention. The objective of this study was to investigate the use of timetable management features to analyze the trade-off between LOS parameters and capacity utilization in the U.S. The research applies a hybrid simulation approach, where output from RTC, a simulation tool developed in the U.S., was used as an input for timetable compression by RailSys, a simulation tool developed in Europe. 28 scenarios were developed in RailSys to identify a preferred scenario with reasonable LOS parameters while maintaining the capacity utilization under the recommended threshold, and the selected scenario of RailSys was then validated in RTC. The results of the study revealed that 10-min maximum allowed dwell time provided the best corridor capacity utilization. Also, the LOS parameters were significantly improved for total number of stops (55% reduction), total dwell times (80% reduction) and average dwell time (65% reduction); while the timetable duration was increased (capacity utilization was degraded) by 18% compared to the initial schedule.

1. Introduction

The majority of passenger rail services in the United States (U.S.) operate on shared-use corridors with substantial freight rail services. Passenger/freight traffic may each operate on dedicated tracks, but in most cases, all trains share the same track infrastructure. The European passenger rail services also operate on shared-use corridors, but the infrastructure conditions and the operating priorities and patterns are different, typically favoring passenger operations (FRA, 2009; Cambridge Systematics and Inc., 2007). Recently, the increasing demand for train traffic (passenger and freight) is creating pressure to add capacity in the U.S. either through the construction of new tracks and lines, or through improved operational strategies.

Capacity analysis, at the network, main line/corridor, or terminal/yard level is one of the tools used to evaluate the benefits and costs of capacity improvement alternatives. Although the concept of capacity and the objective to achieve a high

* Corresponding author.
E-mail addresses: hpouryou@mtu.edu (H. Pouryousef), ptlautal@mtu.edu (P. Lautala).

http://dx.doi.org/10.1016/j.jrtpm.2015.10.001
2210-9706/

utilization while maintaining sufficient level of service (LOS)[1] for trains is global, the configuration differences between the European and the U.S. rail systems (such as the infrastructure ownership and the operating philosophy) leads to the use of different methodologies, techniques, and tools for capacity evaluation. More information on these differences and how they affect the capacity studies is provided in Pouryousef et al., 2013 (Pouryousef et al., 2013).

This paper focuses on rail line/corridor level analysis. It provides a brief synopsis of methods and tools to evaluate main line capacity and the LOS parameters, but the main objective of the study was to investigate the use of timetable management features common in European rail environment to analyze the trade-off between LOS parameters and capacity utilization in the U.S. rail environment. The methodology included development of a "Hybrid Simulation Approach". This differs from traditional analysis, as it takes advantage of the complementary features of non-timetable and timetable based simulation software and uses output from one software as input in another. A single-track case study is used to demonstrate the approach, the outcomes and the challenges.

2. Capacity analysis

There is no standard definition for railway capacity, but one definition that is used is the number of trains that can safely pass over a given segment of the line within a selected time period (UIC, June 2004). Various definitions, metrics, methodologies and tools are applied for evaluating the capacity in Europe and North America, due to the differences of rail network characteristics between the two continents (Pouryousef et al., 2013). Three critical differences between Europe and the U.S. are the ownership of infrastructure, the predominant traffic type (freight vs. passenger), and operating philosophies. In the U.S. more than 90% of the infrastructure is owned and managed by private freight railroads (Thomas, 2005), while in Europe infrastructure is almost completely owned and managed by governments or public agencies. The U.S. operations are predominantly for freight transportation and the prevailing operating philosophy for the majority of freight trains and even some passenger and commuter services is based on the improvised pattern that has no repeatable dispatching plan on over extended time period. In Europe, passenger trains dominate the corridors and almost all trains (freight and passenger) follow structured operations with a regular schedule that is developed months in advance (Thomas, 2005). The reasons noted above, combined with variations in other characteristics, such as rolling stock and signaling systems, all affect capacity, as well as tools and techniques used for capacity analysis.

The literature mainly divides capacity analysis approaches into analytical and simulation methods (Pachl, 2002; Abril et al., 2007; Murali et al., 2009; Khadem Sameni et al., 2011; Sogin and Barkan, 2012; Lai and Barkan, 2009). A combined analytical-simulation approach that takes advantage of both analytical and simulation methods has also been used (Schlechte et al, 2011; Cambridge Systematics and Inc., 2006). The simulation methods typically utilize either general simulation tools or commercial railway simulation software that has been specifically designed for rail transportation (Abril et al., 2007; Khadem Sameni et al., 2011). The commercial railway simulation software can be divided into two major categories: Non-timetable based and Timetable based software. Both incorporate two components: "Train movement simulation" to calculate the train speed along the track, and "Train dispatching simulation" to emulate the actions of the actual dispatcher as closely as possible (Thomas, 2005). The non-timetable based simulations are typically used in railways which are operated based on unstructured operation pattern without initial timetable, such as the majority of the U.S. rail network. The primary objective of this type of simulation tools is to automatically resolve the train conflicts. The Rail Traffic Controller (RTC), developed by "Berkeley Simulation Software, LLC" is the most common software in this category and is used extensively by the U.S. rail industry (Thomas, 2005; Khadem Sameni et al., 2011). The simulation procedure of timetable based software, commonly found in Europe, is based on the initial timetable of trains (often a conflict-free timetable is required). The software can identify the train conflicts, but in most cases have limited capabilities to resolve all conflicts without user intervention. The software typically include features to automatically adjust/improve the initial timetable and are equipped with other timetable management features, such as timetable compression technique. RailSys, developed by Rail Management Consultants GmbH in Germany, is one example of a timetable-based simulation package, and details of the different simulation tools has been provided by Pouryousef et al., 2015 (Pouryousef et al., 2015).

Table 1 provides a sample of recently published capacity studies in the U.S. and Europe, and shows the difference between tools commonly used for analysis. RTC has been the software of choice for all U.S. studies while several timetable-based packages have been used in Europe.

In addition to the software packages highlighted in the Table 1, there are other simulation tools used in the U.S., by rail transit and commuter services (e.g. MultiRail, RailSim), and in Europe (e.g. OpenTrack, Viriato, SLS, RAILCAP, CMS). A review of Table 1 indicates that train delay analysis is a common performance metric for capacity evaluation in the U.S. and one that is recommended by the Federal Railroad Administration (Tolliver, 2010). Europeans have a variety of different methodologies to evaluate the railway performance, but most of them utilize timetable management techniques. For instance, timetable compression technique used in this research, was developed by International Union of Railways (UIC) to improve the capacity utilization or LOS by adjusting operational characteristics, such as dwell times, stop patterns, train departure times and/or the

[1] Level of service (LOS) may include various parameters to evaluate the desire level of rail customer/clients' satisfaction. In the U.S., common parameters used are various types of train delays, but in this research, the LOS parameters are taken into account from the timetable and scheduling standpoint. Common parameters include number of stops (unplanned or meet-pass stops), average dwell time, maximum allowed dwell time and total dwell time.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Table 1
Review of selected capacity simulation studies (academic research) conducted in the U.S. and Europe.

Authors		Simulation package	Applied technique/method through simulation
The U.S. (5 studies)	Khadem Sameni et al. (2011)	RTC	Evaluated a new metric of capacity (profit-generating capacity) for the intermodal and bulk train services in the U.S. by applying different heterogeneity scenarios between these two trains
	Sogin et al. (2012)	RTC	Delay analysis of freight trains along a double-track case study based on applying various speed scenarios and number of passenger/freight trains
	Sogin et al. (2013)	RTC	Compared single and double track performance (train delay analysis) by changing traffic volume, passenger train speed and heterogeneity level of freight and passenger trains
	Atanassov and TylerBarkan (2014)	RTC	Evaluated the additional capacity of different scenarios of adding double track segments to the existing single track, based on delay analysis of freight trains
	Shih et al. (2014)	RTC	Compared different scenarios of single track lines with sparse siding options, in terms of freight train delay
Europe (5 studies)	Schlechte et al. (2011)	OpenTrack	Used simulation package to obtain microscopic level results and to convert the results to macroscopic level for further timetable development by using a specific algorithm, and then the new timetable was retransformed again to the simulation for further analysis
	Gille and Siefer (2013)	RailSys	Used simulation package through a 3-step method of capacity improvement: 1- obtaining max. level of occupancy, 2- running the simulation and determining the service quality, 3- adjustment of max. level of occupancy
	Medeossi and Longo (2013)	OpenTrack	Developed an approach of estimating the stochastic inputs of simulation to be more practical for generating realistic simulation scenarios
	Sipila (2014)	RailSys	Applied simulation package to evaluate different train run time scenarios (vs. minimum run times) based on delay analysis
	Goverde et al. (2014)	ROMA	Used timetable compression technique (UIC method) for computing capacity of corridors with scheduled trains, while for unscheduled (disturbed) traffic conditions, Monte Carlo simulation technique was used for the analysis. (Both applied via ROMA which combines alternative graphs of train-paths)

order of trains. This technique is most applicable on corridors with pre-scheduled timetables and predetermined routes for daily trains (structured operating philosophy) and its objective is to modify the pre-determined timetable by rescheduling the trains as close as possible to each other (UIC, June 2004; Landex et al, 2006; Prinz and Hollmuller, 2005; Banverket, 2005; Khadem Sameni et al., 2010). While U.S. shared corridors rarely operate under structured operating philosophy, the daily schedules for passenger trains rarely change, making the shared use corridors with regular passenger traffic more applicable for the technique.

3. Hybrid simulation approach

"Hybrid Simulation Approach" differs from traditional analysis, as it takes advantage of the complementary features of non-timetable and timetable based software and uses output from one software as input in another. The tools used in the study included RTC as the non-timetable based simulation tool and RailSys as the timetable-based tool. Fig. 1 presents key features of each simulation package. RTC has the capability to use preferred departure times, train dispatching simulation process, and its automatic train conflict resolution to develop the initial timetable (stringline). RTC does not consider a specific total timetable duration, but rather uses a decision support core, called "meet-pass N-train logic" to dispatch all trains while avoiding conflicts and minimizing the overall delays and total operating costs of trains (Pouryousef et al., 2015). In contrast, RailSys requires a specific total timetable duration and an initial timetable (typically a conflict-free schedule) for its simulation and uses a timetable compression technique (based on UIC code 406) to adjust/improve the initial timetable for more efficient capacity utilization or for improving the LOS parameters.

The hybrid approach uses the initial timetable developed in the RTC as input for RailSys and then applies the RailSys timetable compression technique to investigate the trade-off between capacity utilization/LOS by adjusting the initial timetable. The adjusted timetable developed by RailSys is then imported back to RTC as input, so the results can be validated in the U.S. rail environment (Fig. 2).

Fig. 3 illustrates additional details on the approach in a step-by-step basis. Step 1 represents the development of the initial timetable using RTC. Step 2 adjusts/improves the RTC timetable through RailSys compression techniques, and Step 3 validates the new timetable in the RTC.

As presented in Fig. 3, the hybrid approach requires conversion of the data from RTC to RailSys and then checking that the key simulation outcomes match with each other. There are four categories in the database and the level of conversion criteria and level of difficulty vary (Table 2). The conversion of infrastructure and operating rules is straightforward and consists mainly of unit conversion (English to metric). The conversion of train and signaling characteristics is more complicated and may require specific adjustments in individual parameters, as the train performance calculator (TPC) and signal system emulator of RailSys (and many other European-based simulation tools) are less sophisticated and less tuned to the U.S. operations than the ones in RTC, which have been customized for the U.S. rail environment.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

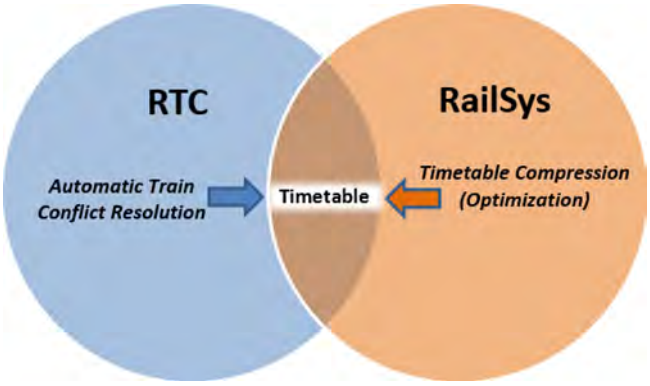

Fig. 1. The main features of RTC and RailSys for timetable development.

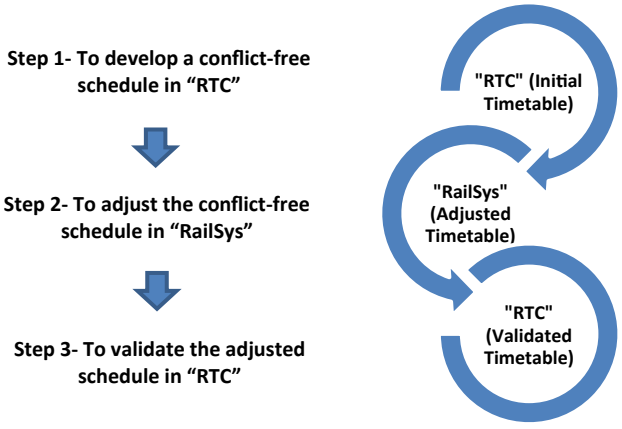

Fig. 2. Main steps (left) and outputs (right) of a "Hybrid Simulation Approach".

The validation process depends on the parameters that need to be matched. In the case study, the main objective was to maintain the same schedule and run time of trains, as well as to check that there were no deviations in train routings. Any deviations in these parameters were used to determine necessary adjustments.

4. Case study

A case study was developed as part of the research to demonstrate the hybrid approach. The case study used a rail line in the U.S. that is currently used for excursion passenger trains, but train and signaling parameters were hypothetical. The input data was developed for each simulation package and included four database categories "operating rules, trains, signaling, and infrastructure".

The line is a 30-mile single track corridor with three sidings/yards for meet/pass and stop purposes. (Fig. 4) The vertical track profile and locations of the sidings were derived from an existing corridor data. The horizontal curves were not included as their impact on the train speed was not considered essential for the simulation results. Table 3 summarizes the infrastructure parameters for the case study.

The signaling system was an absolute permissive block (APB) for single track operation with four-aspect signaling along the main line. The length of blocks varied from 1.2 to 2.5 miles and all sidings/yard tracks were equipped with controlled interlocking systems.

Four types of trains were considered in the case study: intercity passenger (4 daily pairs), commuter passenger (2 daily pairs), merchandise freight (2 daily pairs) and intermodal freight trains (3 daily pairs). It was assumed that the characteristic and configuration of trains was uniform in each specific category and each train was operated in both westbound and eastbound directions. All passenger and commuter trains were propelled by a single diesel-electric locomotive and all freight trains were loaded in both directions. Since the type and configuration of locomotives were different in the RTC and RailSys database, some of the characteristics of selected locomotives in RTC (such as power, weight, length, axle load, acceleration/deceleration rate, resistance) were imposed and adjusted over the existing rolling stock database of RailSys as a new type of locomotive.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

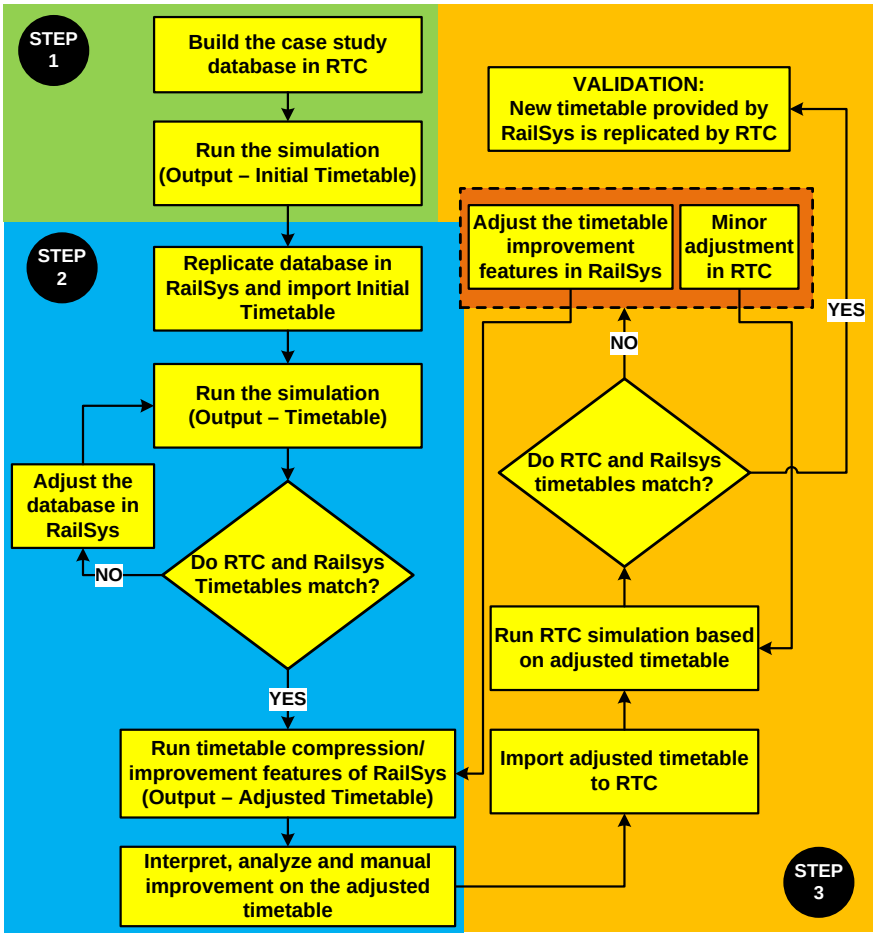

Fig. 3. Flowchart of hybrid simulation steps (RTC-RailSys-RTC).

Table 2
Summary of data conversion from RTC to RailSys.

Category	Conversion criteria	Difficulty level	Main adjustments
Operating rules	Match	Straightforward	Unit conversion
Trains	Maintain run times	Complicated	Train consist, power, max speed, train resistance
Signaling	Maintain routes and run times	Complicated	Signal features, interlocking, blocks
Infrastructure	Match	Straightforward	Unit conversion

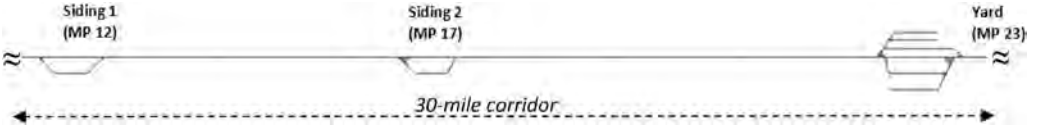

Fig. 4. Case study corridor schematic.

The scenario developed for initial RTC simulation included several relevant operational rules, such as the train priority, speed limits, stop patterns, and preferred time and order of train departures. The train priority (in descending order) was commuter trains, passenger trains, intermodal, and merchandise trains. Due to short train lengths and type of signaling system, the minimum headway between trains was 3 min. The maximum speed of passenger/commuter trains was 60 mph, and for freight trains was 50 mph. In addition, the initial speed of all trains was 30 mph until they reached the track segment that started the simulation process. A predefined timetable was not used, but requested departure times were developed with a goal to have a congested and homogenous dispatching pattern of trains. The congestion aspect of train schedule (with over

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Table 3
Details of case study infrastructure.

Corridor length	30 miles, single track
Sidings/yards	2 sidings + 1 yard
Length of siding track	0.40 mi
Length of yard tracks	Minimum 0.43 mi
Length of freight trains	0.3 mi
Length of intermodal trains	0.42 mi
Max. grade	1.78%
Curvature	Horizontal curves ignored
Turnout#	#11

25 conflicts) allowed evaluation of rail simulation package capabilities in resolving the potential conflicts. The homogenous pattern of train schedule mimicked a common trend in operating the shared-use corridors in the U.S. where the passenger and commuter trains are often separated from the freight services, particularly during peak hours. No consideration was given for real market conditions or for potential interdependencies between the trains and there were no planned stops for any trains. However, passenger, commuter or merchandise trains were allowed to stop at the sidings due to the meet-pass logic. The intermodal freight trains were allowed to have a meet-pass stop only in the yard tracks, as siding lengths could not accommodate these trains (Table 3).

5. Outcomes and discussion

5.1. Replicating initial timetable

Fig. 5 presents the initial simulation results obtained from RTC simulation with no manual adjustments in distance–time diagram format (string-line). There were no planned stops for the trains, but several stops were suggested by RTC for meet-passes in the sidings to resolve train conflicts. The simulated arrival/departure times showed a deviation from the preferred departure times requested, as RTC's automatic decision making features resolved the conflicts between trains and dispatched them to minimize train delay/operational costs.

Certain trains witnessed a notable deviation between the requested and actual departure times when they reach the point that starts simulation process. These "limbo times" are identified by red boxes in Fig. 5 and by "*" and "**" in Table 4. Limbo times are caused by conflicts between the trains and result in several dispatching delays, and in some cases a change in the dispatching order of trains.

In addition to presenting the limbo times, Table 4 compares the requested departure times and simulated departure times and the order of trains by RTC in both westbound and eastbound directions.

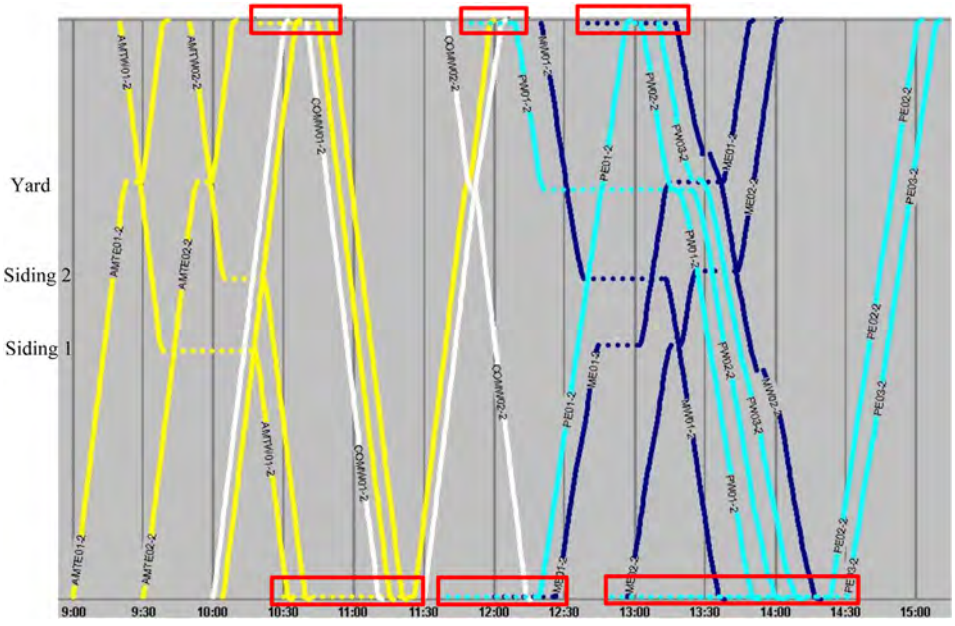

Fig. 5. Simulated train timetable (stringline) in RTC (Commuter: White, Passenger: Yellow, Intermodal: Blue, freight: Navy blue) *(Note: Trains with limbo times are identified by boxes).* (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Table 4
Comparison between requested and simulated departure times (HH:mm) and order of trains (X) in RTC.

Train	Requested departure time, (order of train) – eastbound	Simulated departure time, (order of train) – eastbound	Requested departure time, (order of train) – westbound	Simulated departure time, (order of train) – westbound
Pass1	9:00, (1)	9:00, (1)	9:20, (1)	9:20, (1)
Pass2	9:30, (2)	9:30, (2)	9:50, (2)	9:50, (2)
Pass3	10:00, (4)	10:03*, (4)	10:20, (3)	10:45**, (4)
Pass4	10:30, (5)	11:27*, (5)	10:50, (5)	10:50, (5)
Comm1	10:00, (3)	10:00, (3)	10:40, (4)	10:40, (3)
Comm2	11:30, (6)	11:30, (6)	11:40, (6)	11:40, (6)
Interm1	11:40, (7)	12:20*, (7)	11:50, (7)	12:08**, (7)
Interm2	12:50, (9)	14:23*, (10)	13:00, (10)	13:02**, (9)
Interm3	13:20, (11)	14:30*, (11)	13:10, (11)	13:10, (10)
Freight1	12:00, (8)	12:25*, (8)	12:20, (8)	12:20, (8)
Freight2	12:50, (10)	12:55*, (9)	12:40, (9)	13:15**, (11)

As presented in Table 4, trains with higher priority (commuter and passenger) had lower deviation between their requested and simulated departure times. There was also conflict between requested departure time of passenger 3 and commuter 1 (eastbound direction), as both trains were requested to depart at 10:00. RTC solved the time conflict by maintaining the initial schedule of commuter train (with higher priority) and delaying passenger train for 3 min at the entry point of the line. Similar situation occurred between intermodal 2 and freight 2 in eastbound direction (both were planned to depart at 12:50). RTC changed departure times of both trains to facilitate necessary meet-pass events. After the changes, the high priority commuter trains had one short stop in a siding due to the meet-pass enforcement, while passenger trains faced more frequent and longer delays in the sidings and on the entry points of the line. (Fig. 5) The same trend was noticed for freight and intermodal train schedules with even more delays and longer meet-pass time in the sidings, since the priority of these two types of trains was lower than passenger and commuter trains. However, the merchandise freight trains had lower delays in comparison to the intermodal freight trains, although the priority of intermodal trains was slightly higher than merchandise train. This may be due to the fact that merchandise trains had more flexibility for meet-pass stop locations, while intermodal trains were limited to stopping in the yards with sufficient siding lengths to fit the full train. In conclusion, the majority of trains (particularly commuter trains with highest dispatching priority) maintained the order initially requested, while dispatching order for some of the other trains was changed. It should be noted that the dispatching order is very dependent on the gap between the priority values for different train types. Therefore, changing the priority order of trains, or increasing the gap in the priority values causes the RTC to propose a different schedule and dispatching order of trains from those presented in Fig. 5 and Table 4.The output from the RTC simulation (simulated eastbound and westbound departure times) was used as the actual terminal departure times in the RailSys simulation (Fig. 6). Since the actual departure times from RTC output equaled the preferred departure times for RailSys, there were no “limbo times” identified in the RailSys simulation. However, there were some minor deviations between arrival/departure times in RailSys and RTC, due to differences between rolling stock and signaling features/equations of each simulation package, such as tractive effort of engines, acceleration, deceleration, and braking diagram. Despite these differences, approximately 96% of timetable parameters (order of trains, stop patterns, departure/arrival times) were identical in RailSys when compared to the initial timetable obtained from RTC.

5.2. Timetable adjustment/improvement

Once the accuracy of the converted database was verified, RailSys capabilities were used to adjust and compress the initial timetable for LOS and capacity utilization analysis. RailSys uses UIC 406 compression technique with predefined patterns and algorithms to automatically adjust the initial timetable, resulting in changes to LOS/capacity utilization. Several criteria have to be defined in RailSys prior to the automatic timetable adjustment/compression, such as:

- The initial timetable (RTC output was used as initial timetable of RailSys)
- Selection between compression technique (Austrian method, OBB, or German method, DB)
- Overtaking option in the sidings/stations
- Timetable duration (the portion of timetable which is planned to be adjusted)
- Maximum allowed dwell time of trains in the sidings

The case study used the OBB compression algorithm and allowed an overtaking option at a maximum of two stations. OBB was selected over DB algorithm, as it maintained the number of simulated trains extracted from RTC results and under this algorithm RailSys does not allow for changes in the departure order of trains. In single track scenarios like ours, it is recommended to allow overtaking option at meet/pass locations. The locations (sidings/stations) and the trains for meets/passes and any additional stops are determined automatically by the RailSys, based on its own decision algorithm of compression technique.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

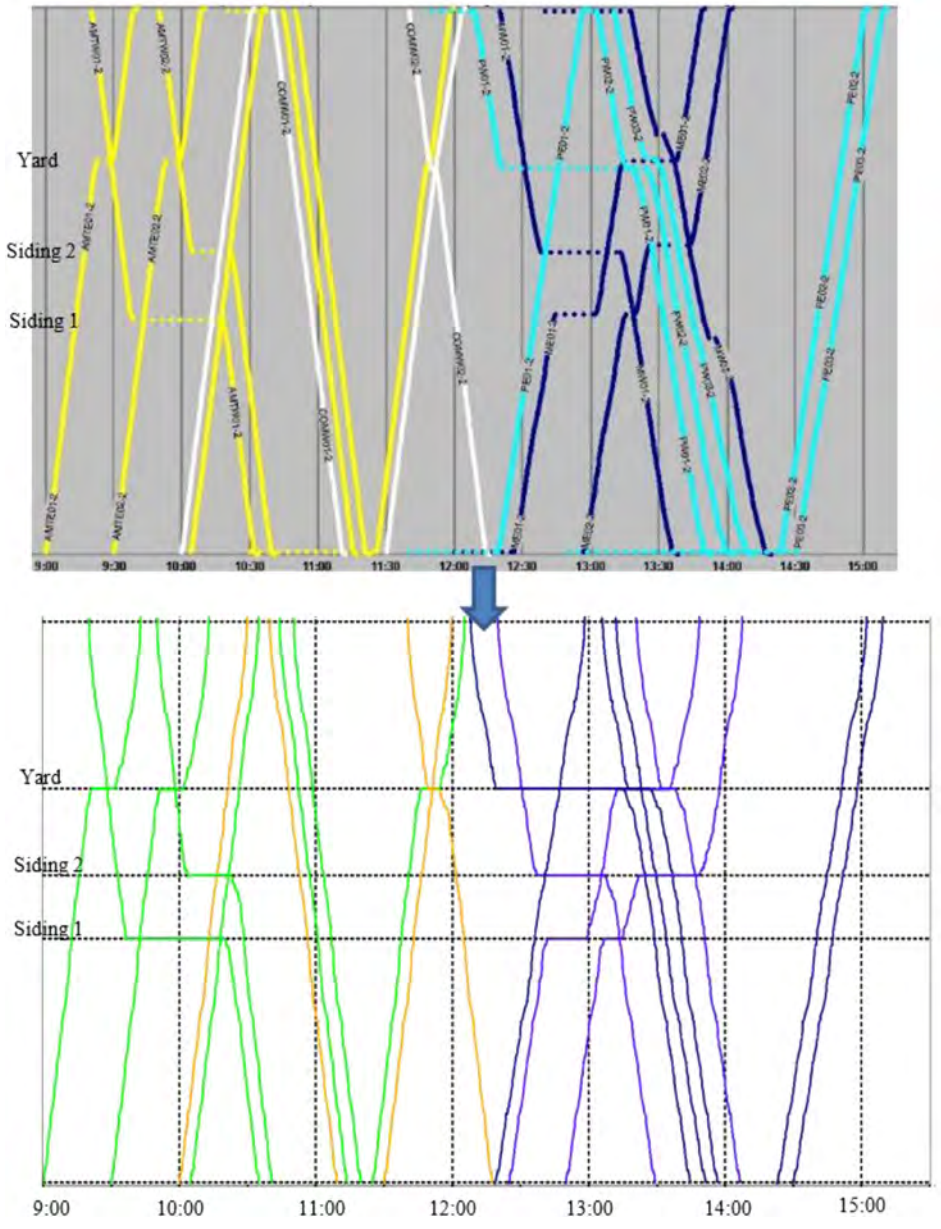

Fig. 6. The output of RTC stringline (top) as replicated in RailSys as input (bottom).

As noted earlier, RTC does not use a specific timetable duration when developing the initial timetable. However, in RailSys, timetable duration and maximum allowed dwell time of trains are critical parameters. In the base scenario developed to obtain a maximum LOS for the corridor, both train stops and maximum allowed dwell times were assigned as zero. This matched the initial requests in RTC before simulation (no stop pattern, no dwell time). Fig. 7 demonstrates the adjusted timetable developed by RailSys for maximized LOS values. In this scenario, the total duration of the timetable was almost 2 h longer than in the results obtained from the RTC, as the overall capacity utilization was decreased because all meet-pass options were eliminated. Therefore, a more realistic timetable duration/maximum allowed dwell time combination was investigated that would meet recommended capacity utilization, while improving LOS of the initial timetable.

A total of 28 timetable duration/maximum allowed dwell time combinations were applied to identify a duration/dwell time combination that aligns closely with the 70% practical capacity utilization threshold recommended in railway literature (Pouryousef et al., 2013; UIC, June 2004; Pachl, 2002). Fig. 8 presents the capacity utilization percentage of all simulated scenarios and illustrates that the timetable compression with "10-h timetable duration and 10-min maximum allowed dwell

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

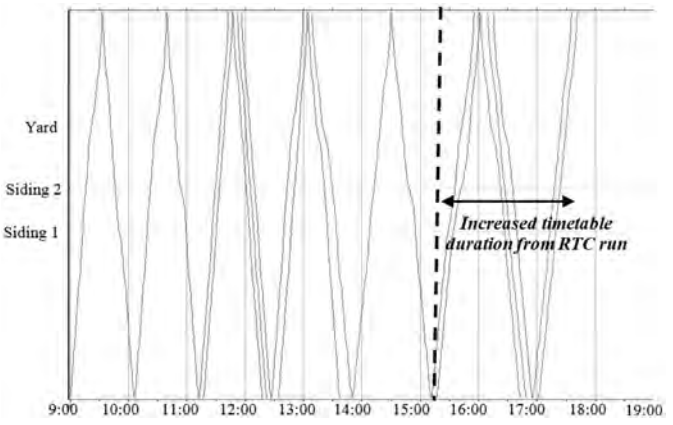

Fig. 7. Adjusted timetable (stringline) with maximum LOS in RailSys based on "no dwell time" scenario.

time" keep the capacity utilization under the 70% threshold. Any other combination (e.g. 11-hours timetable duration and 20-min max. allowed dwell time) would either unnecessarily extend the maximum allowed dwell time, or increase the capacity utilization above recommended threshold.

Table 5 compares the differences in train departure times and dispatching order between initial and "10-min maximum allowed dwell time" scenarios in RailSys. As illustrated in the table, the dispatching order remained the same during the compression for both eastbound and westbound directions, as it was in the initial schedule.

Fig. 9 presents the timetable with the preferred "10-min maximum allowed dwell time" scenario. The timetable has been compressed by approximately 60 min from the maximum LOS option (Fig. 7). The adjusted timetable is still 57 min longer than the initial timetable of RTC, but the LOS parameters have been substantially improved. This demonstrates the trade-off between LOS parameters and total timetable duration (or capacity utilization).

Fig. 9 revealed a few occasions where trains were stopped for a siding without reason (Fig. 10-top). It was speculated that RailSys maintained the unnecessary stops, as they were needed to resolve the train conflicts in the initial schedule. Manual timetable adjustments were made to eliminate the unnecessary stops. The new adjusted timetable (RailSys compression technique + manual adjustments) reduced the overall duration of timetable by approximately 25 min as illustrated in Fig. 10-bottom, making it 32 min longer (instead of 57 min) than the initial RTC timetable, but with improved LOS parameters.

5.3. Validation of adjusted timetable in RTC

As illustrated in the research flow diagram (Fig. 2), the final step of the hybrid process was to validate the new timetable developed in RailSys by running it through RTC. Fig. 11 shows the final timetable by RailSys and the validated timetable in RTC.

As illustrated in Fig. 11, all trains were successfully dispatched in RTC with the same order and same stop patterns. However, the signaling and rolling stock differences between RailSys and RTC caused the similar deviations of arrival/departure times and dwell times (approx. 1–4 min deviation), as were witnessed during the RTC/RailSys conversion. These deviations caused 40 min longer timetable duration in RTC (Fig. 11-bottom) when compared to the RailSys results (Fig. 11-top).

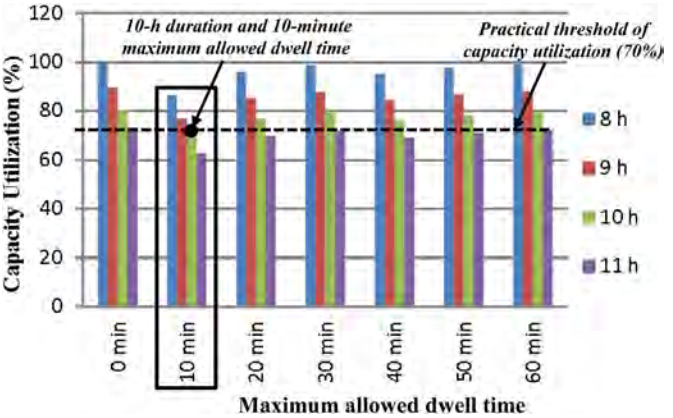

Fig. 8. Capacity utilization based on different timetable durations and maximum allowed dwell times.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Table 5
Train departure times (HH:mm) and order of trains (X) for initial and adjusted timetable developed by RailSys (10-min maximum allowed dwell time scenario).

Train	Initial departure time, (order of train) – eastbound	Adjusted departure time, (order of train) – eastbound	Initial departure time, (order of train) – westbound	Adjusted departure time, (order of train) – westbound
Pass1	9:00, (1)	9:00, (1)	9:34, (1)	9:20, (1)
Pass2	9:30, (2)	10:07, (2)	10:40, (2)	9:50, (2)
Pass3	10:05, (4)	11:18, (4)	11:53, (4)	10:44, (4)
Pass4	11:25, (5)	12:28, (5)	11:57, (5)	10:50, (5)
Comm1	10:00, (3)	11:12, (3)	11:48, (3)	10:40, (3)
Comm2	11:30, (6)	12:35, (6)	13:05, (6)	11:40, (6)
Interm1	12:20, (7)	13:51, (7)	13:10, (7)	12:08, (7)
Interm2	14:23, (10)	16:17, (10)	16:03, (9)	13:06, (9)
Interm3	14:30, (11)	17:03, (11)	16:10, (10)	13:12, (10)
Freight1	12:25, (8)	15:12, (8)	14:31, (8)	12:20, (8)
Freight2	12:50, (9)	15:18, (9)	16:16, (11)	13:21, (11)

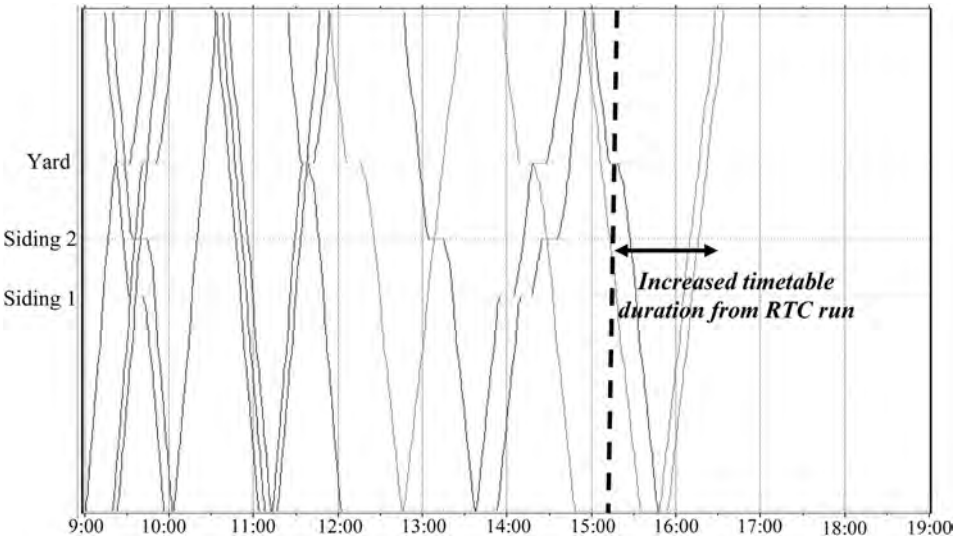

Fig. 9. Adjusted timetable (stringline) in RailSys with 10-min maximum allowed dwell time.

In addition to comparing overall duration, this step of research compared the order of trains, stop patterns, and departure/arrival times. The results showed approximately 92% match between RTC and RailSys in validation process. Table 6 presents a breakdown of the matching levels for individual parameters.

The first column of Table 6 (Matching Parameters) describes the parameters which were included for evaluating the validation step. The second column of the table presents the observed deviations between the RTC and RailSys outputs for each respective parameter. The third column (Matching %) uses the observed deviations to calculate the matching percentage of each respective parameter. For instance, in "Overall Duration" (the first parameter), the value of deviation (40) was divided by the overall timetable duration of RTC (462) and then deducted from "1" to calculate the "Matching %" (92%).

The importance of each parameter for evaluating the validation step is shown in the fourth column (Impact Factor). For this study, a higher impact factor of "0.2" is assigned for "Overall Duration", "Order of Trains", and "Stop patterns", as these parameters are more critical in defining a timetable. The impact factor of the remaining parameters is "0.1".

The "Matching %" of each parameter was multiplied with the respective "Impact Factor" to compute the "Normalized Matching %". Then, the "Normalized Matching %" were aggregated to calculate the "Overall Matching %" between RTC and RailSys outputs.

5.4. Discussion of results

Table 7 summarizes the timetable characteristics derived from outcomes of the hybrid simulation approach for 10-h timetable duration and 10-min maximum allowed dwell time. The table shows substantial improvements in LOS parameters, including 55% reduction in total stops, over 80% reduction in total dwell times and over 65% reduction in average dwell time with low standard deviation. These improvements are countered by an 18% increase in the timetable duration from the

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

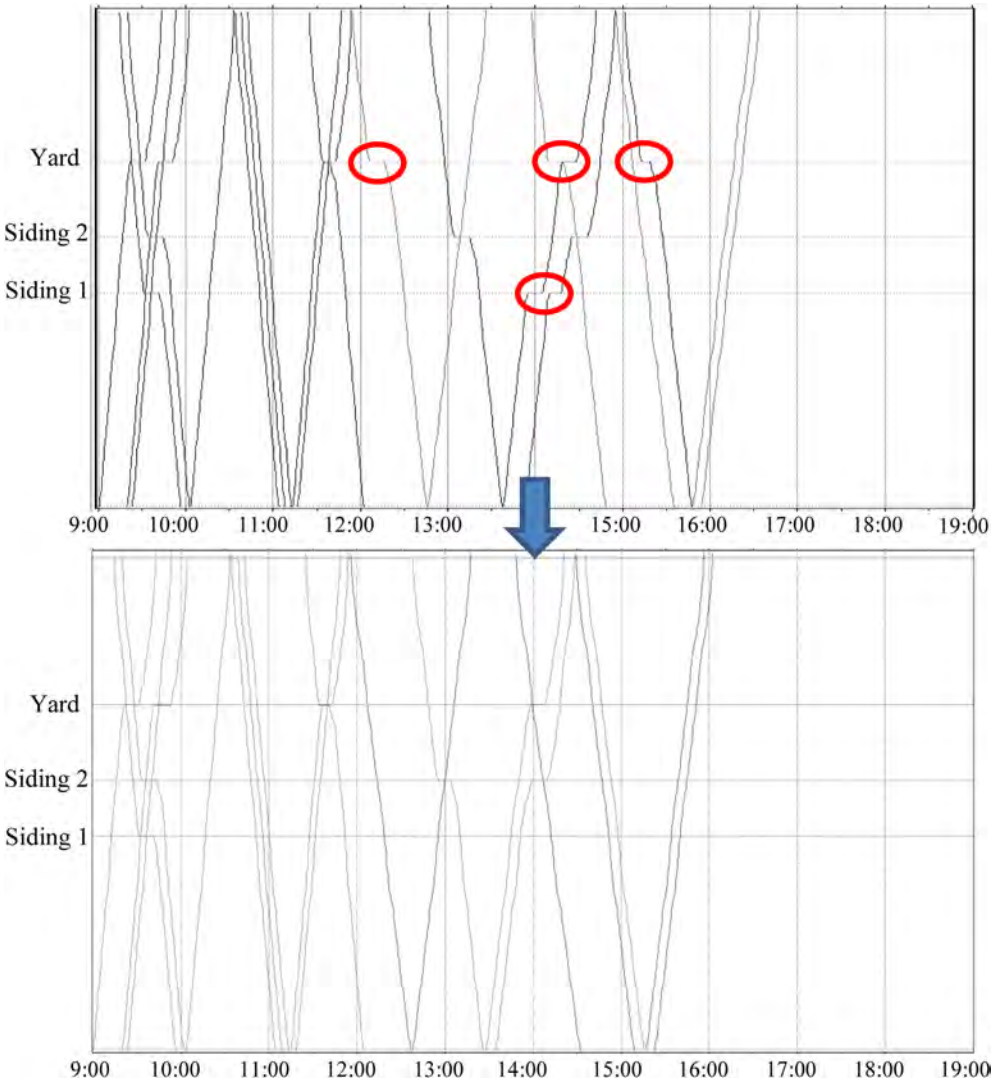

Fig. 10. The unnecessary stops (red circles) in adjusted timetable (top) were manually removed (bottom). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

initial timetable (after RTC validation). The results highlight timetable management's potential to increase the LOS, but also confirms the reverse relationship between LOS criteria and capacity utilization levels. As we see, if LOS is improved, the timetable tends to be stretched and capacity utilization may be degraded and vice versa.

6. Summary and conclusions

This paper started with an introduction to the railway capacity and LOS analysis, and briefly discussed the use of commercial railway simulation software. The paper also introduced a hybrid simulation approach that uses the capabilities of timetable based (RailSys) and non-timetable based software (Rail Traffic Controller or RTC) to investigate the trade-off between timetable duration (capacity utilization) and LOS parameters. The hybrid simulation approach used the output of RTC as input in RailSys, and the timetable compression technique offered by RailSys was applied to adjust the initial timetable. The adjusted RailSys timetable was then validated through RTC simulation to confirm its repeatability in the U.S. based software.

The hybrid simulation approach was successfully completed to turn a train dispatching request with numerous conflicts first to a conflict-free initial timetable by RTC, and then to a modified timetable with maximized LOS (no stop pattern, no dwell time) in RailSys. 28 scenarios were then developed in RailSys to identify timetable duration/maximum allowed dwell time combination that provided reasonable LOS parameters while maintaining the capacity utilization under the recommended threshold of 70%. Finally, the selected scenario of RailSys was validated in RTC with over 90% match between the simulation results.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

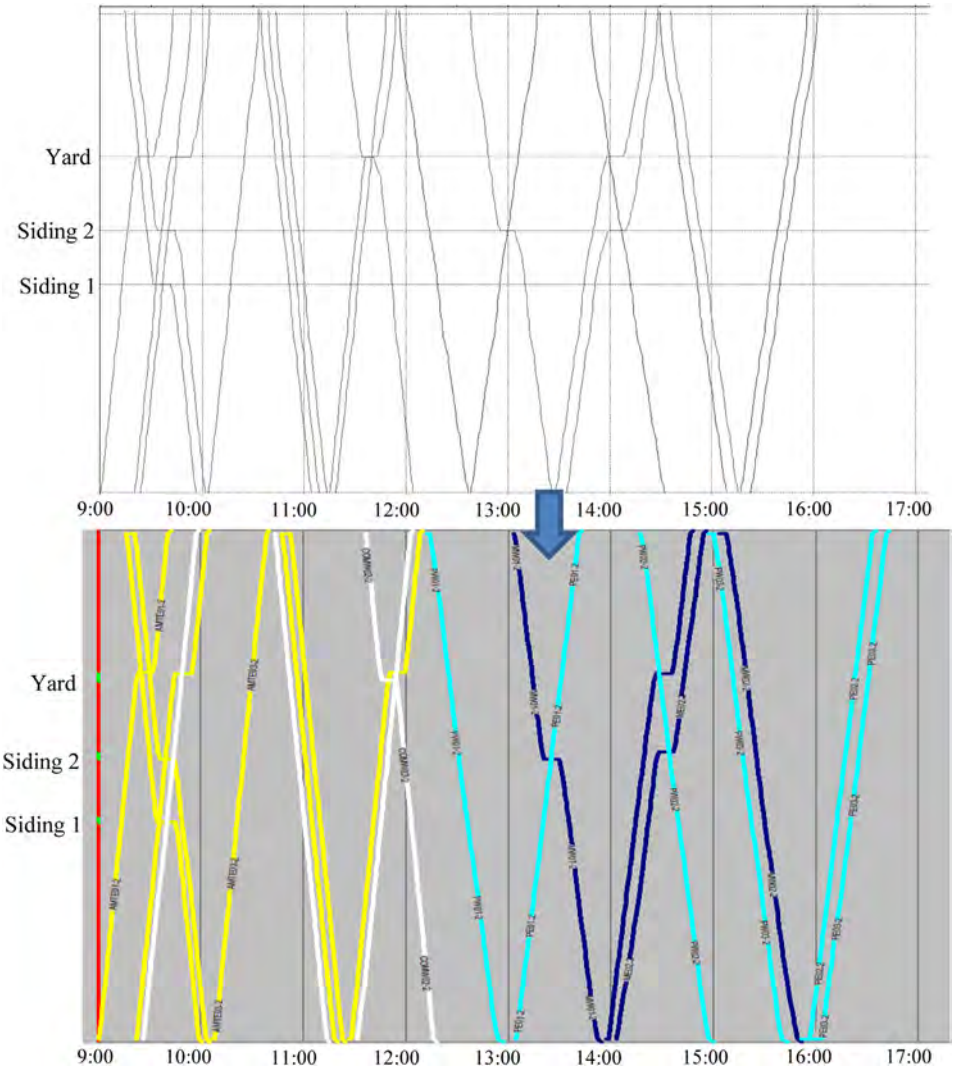

Fig. 11. The RailSys final stringline (top), validated stringline in RTC (bottom).

Based on the results, the scenario of 10-h timetable duration and 10-min maximum allowed dwell time was identified as the preferred scenario for evaluating the trade-off between LOS parameters and capacity utilization. When comparing with the initial timetable, the unnecessary stops were reduced by 55%, delays reduced by 85%, and average dwell times were reduced over 70% with low standard deviation. As a trade-off, the total timetable duration was increased by 72 min (18%). The

Table 6
Validation of the RailSys timetable in RTC.

Matching parameters	Deviations of RTC vs. RailSys	Matching %	Impact factor ($0 < X < 1$)	Normalized matching %
Overall duration	40-min. longer	$\left(1-\frac{40}{462}\right)\times 100 = 92\%$	0.2	18.4%
Order of trains	Same	100%	0.2	20%
Stop patterns	Same	100%	0.2	20%
Departure/arrival time deviation[a]	11 deviations	$\left(1-\frac{11}{22\times 5}\right)\times 100 = 90\%$	0.1	9%
Max. allowed dwell time	2-min. longer	$\left(1-\frac{2}{12}\right)\times 100 = 84\%$	0.1	8.4%
Total dwell time	21-min. longer	$\left(1-\frac{21}{105}\right)\times 100 = 80\%$	0.1	8%
Average dwell time per stop	2.4-min. longer	$\left(1-\frac{2.4}{11.7}\right)\times 100 = 80\%$	0.1	8%
Overall matching %				**91.8 ≈ 92%**

[a] For 22 trains departed from/arrived to five different stop points.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Table 7
Comparison between initial and adjusted timetables (10-min maximum allowed dwell time) through hybrid simulation approach.

Criteria		Initial timetable		Adjusted timetable	
		Developed by RTC	Replicated in RailSys	Developed by RailSys	Validated in RTC
LOS	Max allowed dwell time	61′	60′	10′	12′
	Number of stops	20	20	9	9
	Total dwell time- delay	702′	685′	84′	105′
	Average dwell time- delay	35.1′	34.2′	9.3′	11.7′
	Standard deviation of dwell time	17.79	17.70	0.67	0.47
Capacity Utilization	Timetable duration (simulated)	390′	390′	422′	462′
Matching % with original timetable		–	96%	–	92%

results highlight the timetable management's potential to improve the LOS, but also confirm the reverse relationship between LOS criteria and capacity utilization levels.

Overall, the study suggests that timetable compression technique can be applied in the U.S., if an appropriate model and algorithm are developed to address the respective network and operational characteristics of the U.S. rail environment. While the hybrid simulation approach as developed for the research proved to be successful and provided credible results, it was also extremely time-consuming, which reduces its applicability to industry applications. The fact that RailSys is developed in Europe also made the conversion to North American rolling stock and signaling systems challenging and caused minor differences between the simulation outcomes. Future research could investigate whether a different commercial rail simulation package (such as OpenTrack, Schedulemiser, Trapeze) can produce the outcomes of hybrid simulation approach in single software, or if they face similar challenges as software used in the study. Alternatively, one of the tested software packages could either be extended to complete all the steps, or a standalone model could be developed to replace the need for a second simulation software.

Acknowledgements

This research was supported by National University Rail (NURail) Center, a US DOT-OST Tier 1 University Transportation Center. The authors would like to thank Eric Wilson (Berkeley Simulation Software, LLC), Sonja Perkuhn (RMCon), and Gabriele Löber (RMCon) for providing academic licenses to simulation packages, RTC and RailSys respectively, and for their support to the Michigan Tech research team.

References

Abril, M., Barber, F., Ingolotti, L., Salido, M.A., Tormos, P., Lova, A., 2007. An Assessment of Railway Capacity. Technical University of Valencia, Valencia, Spain.
Atanassov, Ivan, Tyler, Dick C., Barkan, Christopher P.L., 2014. Siding spacing and the incremental capacity of the transition from single to double track. In: 2014 Joint Rail Conference (JRC). Colorado Springs, CO, USA.
Banverket, 2005. Application of the UIC Capacity Leaflet at Banverket. Banverket, Swedish National Rail Administration.
Cambridge Systematics, Inc, 2006. Statewide Rail Capacity and System Needs Study, Final Report. State of Washington-Transportation Commission, Olympia, WA.
Cambridge Systematics, Inc, 2007. National Rail Freight Infrastructure Capacity and Investment Study. Association of American Railroads (AAR), Cambridge, Massachusetts.
FRA, 2009. Vision for High-speed Rail in America, U. DOT, Editor. Federal Railroad Administration (FRA), Washington, DC.
Gille, Andreas, Siefer, Thomas, 2013. Sophisticated capacity determination using simulation. In: Transportation Research Board 2013 Annual Meeting. Washington, DC, USA.
Goverde, Rob M.P., Corman, Francesco, D'Ariano, Andrea, 2014. Investigating on the capacity consumption at Detch railways for varoius signaling technologies and traffic conditions. In: TRB 2014 Annual Meeting. Washington, DC, USA.
Khadem Sameni, Melody, Dingler, Mark, Preston, John M., 2010. Revising the UIC 406 method: revenue generating capacity. In: Joint Rail Conference (JRC) 2010 (Urbana, IL, USA).
Khadem Sameni, Melody, Dingler, Mark, Preston, John M., Barkan, Christopher P.L., 2011. Profit-generating capacity for a freight railroad. In: TRB 90th Annual Meeting. TRB, Washington, DC.
Lai, Y., Barkan, C., 2009. Enhanced Parametric Railway Capacity Evaluation Tool, vol. 2117. Journal of the Transportation Research Board, National Academies, Washington, D.C, pp. 33–40.
Landex, Alex, et al., 2006. Evaluation of railway capacity. In: Annual Transport Conference. Aalborg University, Denmark.
Medeossi, Giorgio, Longo, Giovanni, 2013. An approach for calibrating and validating the simulation of complex rail networks. In: TRB 2013 Annual Meeting. Washington, DC, USA.
Murali, Pavankumar, Dessouky, Maged M., Ordonez, Fernando, Palmer, Kurt, 2009. A Delay Estimation Technique for Single and Double-track Railroads. Dept. of Industrial & Systems Engineering, University of Southern California, Los Angeles, CA.
Pachl, Joern, 2002. Railway Operation and Control. VTD Rail Publishing- USA, Mountlake Terrace-WA.
Pouryousef, Hamed, Lautala, Pasi, Thomas, White, 2013. Review of capacity measurement methodologies; similarities and differences in the U.S. and European railroads. In: 2013 TRB Annual Meeting (Washington-DC).
Pouryousef, Hamed, Lautala, Pasi, Thomas, White, 2015. Railroad capacity tools and methodologies in the U.S. and Europe. J. Mod. Transp. 23 (1), 30–42.
Prinz, Robert, Hollmuller, Josef, 2005. Implementation of UIC 406 Capacity Calculation Method at Austrian Railway (OBB). Austrian Railway (OBB), Vienna, Austria.
Shih, Mei-Cheng, et al., 2014. Comparison of capacity expansion strategies for single-track railway lines with sparse sidings. In: TRB 2014 Annual Meeting. Washington, DC, USA.
Sipila, Hans, 2014. Evaluation of single track timetables using simulation. In: 2014 Joint Rail Conference (JRC). Colorado Springs, CO, USA.
Sogin, Samuel L., Barkan, Christopher P.L., 2012. Railroad capacity analysis. In: Railroad Engineering Education Symposium (REES) by AREMA. Kansas City.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

Sogin, Samuel L., et al., 2013. Comparison of the capacity of single and double track rail lines using simulation analyses. In: Transportation Research Board 92nd Annual Meeting. Washington, DC, USA.
Sogin, Samuel L., Barkan, Christopher P.L., Lai, Yung-Cheng, Saat, Mohd Rapik, 2012. Impact of passenger trains in double track networks. In: 2012 Joint Rail Conference (Philadelphia, Pennsylvania).
Thomas, White, 2005. Alternatives for Railroad Traffic Simulation Analysis, vol. 1916. Journal of the Transportation Research Board (TRB), Washington, DC, pp. 34–41.
Tolliver, Denver, 2010. Railroad Planning & Design, CE 456/656 Lecture Notes. North Dakota State University.
Schlechte, Thomas, et al., 2011. Micro–macro transformation of railway networks. Elsevier J. Rail Transp. Plan. Manag. 1, 38–48.
UIC, June 2004. UIC Code 406R-capacity-1st Edition. International Union of Railways (UIC), Paris, France.

Please cite this article in press as: Pouryousef, H., Lautala, P., Hybrid simulation approach for improving railway capacity and train schedules, Journal of Rail Transport Planning & Management (2015), http://dx.doi.org/10.1016/j.jrtpm.2015.10.001

J. Mod. Transport. (2015) 23(1):30–42
DOI 10.1007/s40534-015-0069-z

JMT
www.springer.com/40534

Railroad capacity tools and methodologies in the U.S. and Europe

Hamed Pouryousef · Pasi Lautala · Thomas White

Received: 5 May 2014 / Revised: 17 February 2015 / Accepted: 18 February 2015 / Published online: 3 March 2015

Abstract A growing demand for passenger and freight transportation, combined with limited capital to expand the United States (U.S.) rail infrastructure, is creating pressure for a more efficient use of the current line capacity. This is further exacerbated by the fact that most passenger rail services operate on corridors that are shared with freight traffic. A capacity analysis is one alternative to address the situation and there are various approaches, tools, and methodologies available for application. As the U.S. continues to develop higher speed passenger services with similar characteristics to those in European shared-use lines, understanding the common methods and tools used on both continents grows in relevance. There has not as yet been a detailed investigation as to how each continent approaches capacity analysis, and whether any benefits could be gained from cross-pollination. This paper utilizes more than 50 past capacity studies from the U.S. and Europe to describe the different railroad capacity definitions and approaches, and then categorizes them, based on each approach. The capacity methods are commonly divided into analytical and simulation methods, but this paper also introduces a third, "combined simulation–analytical" category. The paper concludes that European rail studies are more unified in terms of capacity, concepts, and techniques, while the U.S. studies represent a greater variation in methods, tools, and objectives. The majority of studies on both continents use either simulation or a combined simulation–analytical approach. However, due to the significant differences between operating philosophy and network characteristics of these two rail systems, European studies tend to use timetable-based simulation tools as opposed to the non-timetable-based tools commonly used in the U.S. rail networks. It was also found that validation of studies against actual operations was not typically completed or was limited to comparisons with a base model.

Keywords Railroad capacity · Simulation · Railroad operation · The U.S. and European railway characteristics

H. Pouryousef (✉) · P. Lautala
Civil and Environmental Engineering Department, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, USA
e-mail: hpouryou@mtu.edu

P. Lautala
e-mail: ptlautal@mtu.edu

T. White
Transit Safety Management, 3604 220th Pl SW, Mountlake Terrace, WA 98043, USA
e-mail: taw@vtd.net

1 Introduction

Typically, the capacity of a rail corridor is defined as the number of trains that can safely pass a given segment within a period of time. The capacity is affected by variations in system configurations, such as track infrastructure, signaling system, operation philosophy, and rolling stock.

The configuration differences between European and the U.S. rail systems may lead to different methodologies, techniques, and tools to measure and evaluate the capacity levels. There are high utilization corridors in Europe where intercity passenger, commuter, freight, and even high-speed passenger services operate on shared tracks, and all train movements follow their predefined schedule in highly structured daily timetables that may be planned for a full year in advance. On the contrary, the prevalent operations

pattern on current shared corridors in the U.S. follows unstructured (improvised) philosophy, where schedules and routings (especially for freight trains) are often adjusted on a daily or weekly basis. Recently, the U.S. has placed an increasing emphasis on either the development of new higher speed passenger services, or to incrementally increase the speeds of current passenger services on selected shared corridors [1]. At the same time, the slower speed freight rail transportation volumes are also expected to increase [2]. These increases in volumes and operational heterogeneity can be expected to add pressure for higher capacity utilization of the U.S. shared-use corridors. Capacity measurement and analysis approaches (and their methods and tools) will play a crucial part in preparing the U.S. network for these changes. To maximize the efficiency of future improvements, such as new passenger and high-speed rail services, the accuracy and applicability of capacity tools and methods in the U.S. environment need to be carefully evaluated. Whether the analytical and operational approaches utilized in Europe would provide any benefits for the U.S. shared-use corridors should also be reviewed.

This paper starts by identifying the various definitions of capacity and by discussing the similarities and differences between the U.S. and European rail systems that may affect both the methods and outcomes of capacity analysis. It will also identify different approaches to conduct the analysis and concludes with an examination of several past capacity studies from both continents.

2 What is capacity?

2.1 Capacity concept and definitions

The definition used for rail capacity in the literature varies based on the techniques and objectives of the specific study. For instance, Barkan and Lai [3] defined capacity as "a measure of the ability to move a specific amount of traffic over a defined rail corridor in the U.S. rail environment with a given set of resources under a specific service plan, known as level of service (LOS)". They listed several infrastructure and operational characteristics which affect capacity levels, including length of subdivision, siding length and spacing, intermediate signal spacing, percentage of number of tracks (single, double, and multi-tracks), and heterogeneity in train types (train length, power-to-weight ratios). In another paper, Tolliver [4] introduced freight rail capacity as the number of trains per day for typical track configurations depending on several factors, such as track segment length, train speed, signal aspects and signal block length, directional traffic balance, and peaking characteristics. The American Railroad

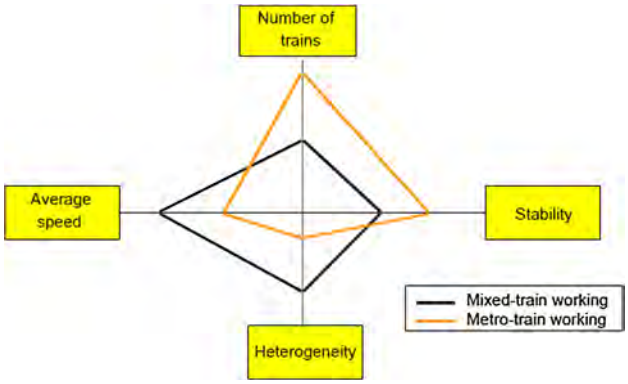

Fig. 1 Capacity balance according to UIC code 406 definition [7]

Engineering and Maintenance-of-Way Association (AREMA) offers a simplified approach for line capacity that estimates practical capacity by multiplying theoretical capacity (C_t) and dispatching efficiency (E) of the line ($C = C_t \times E$). AREMA's method for calculating theoretical capacity and dispatching efficiency requires consideration of various factors, such as number of tracks, the operations rules (single or bi-direction operation), stopping distance between trains (or headway), alignment specifications (grade, curves, sidings, etc.), trains specifications (type of train, length, weight, etc.), maintenance activities requirements, and the signaling and train control systems [5]. A capacity modeling guidebook for the U.S. shared-use corridors, released by the Transportation Research Board (TRB), defines capacity as "the capability of a given set of facilities, along with their related management and support systems, to deliver acceptable levels of service for each category of use." Similar to the other capacity definitions, TRB notes that different parameters and variables should be considered in the capacity analysis, such as train dispatching patterns, train type and consist, signaling system, infrastructure, track maintenance system, etc. [6].

In Europe, the most common method for capacity analysis is provided by the International Union of Railways (UIC) code 406. According to UIC 406, there is no single way to define capacity, and the concerns and expectations vary between different points of view by railroad customers, infrastructure and timetable planners, and railroad operators. UIC also emphasizes that the capacity is affected by interdependencies and the interrelationships between the four major elements of railway capacity including average speed, stability,[1] number of trains, and heterogeneity,[2] as shown in Fig. 1 [7]. According to the figure, a rail

[1] The state of keeping the same train schedule by providing time margins/buffers between trains arrival/departures; despite of minor delay which may occur during operation.

[2] Diversity level of train types which are in operation along a shared-use corridor.

line with various types of trains on the same track (mixed traffic operations or shared-use corridor) has a higher heterogeneity level compared to the urban metro (subway) system with dedicated right-of-way and homogeneous operations. While the average speed of a mixed traffic corridor might be higher than a dedicated metro line, the various train types reduce the stability of train schedules, as well as the total number of trains that can operate on the corridor, due to increased headway requirements.

According to UIC, the absolute maximum capacity, or "Theoretical Capacity", is almost impossible to achieve in practice, and it is subject to:

- Absolute train-path harmony (the same parameters for majority of trains)
- Minimum headway (shortest possible spacing between all trains)
- Providing best quality of service [7].

In addition to the UIC literature, research conducted as part of European Commission's "Improve Rail" project produced a definition of ultimate capacity that was similar to the UIC's theoretical capacity definition, but placed higher emphasis on the train schedules and running time [8].

2.2 Capacity metrics

The literature categorizes the main metrics of capacity level measurements into three groups: throughput (such as number of trains, tons, and train-miles), LOS (terminal/station dwell, punctuality/reliability factor, and delay), and asset utilization (velocity, infrastructure occupation time, or percentage) [9]. In 1975, The Federal Railroad Administration (FRA) introduced a parametric approach developed by "Peat, Marwick, Mitchell and Company" to measure capacity in the U.S. rail network based on delay units (hours per 100 train-miles) [4]. The European rail operators typically use throughput metrics (number of trains per day or hours) to measure the capacity levels, although punctuality and asset utilization metrics are also applied as secondary units [8, 10].

3 Differences between the U.S. and European rail systems

The U.S. and European rail networks have several similarities, such as mixed operations on shared-use corridors, and using modern signaling and traffic control systems (e.g., developing ETCS in Europe and PTC in the U.S.). On the other hand, significant differences also exist and they may change the preferred methodologies, tools, and the outcomes of capacity analysis. Figure 2 and the following discussion uses the literature review to highlight several key differences between infrastructure, signaling, operations, and rolling stock in Europe and the U.S.

3.1 Infrastructure characteristics

- *Public versus private ownership of infrastructure* The ownership of rail infrastructure is one of the important differences between Europe and the U.S. rail networks.

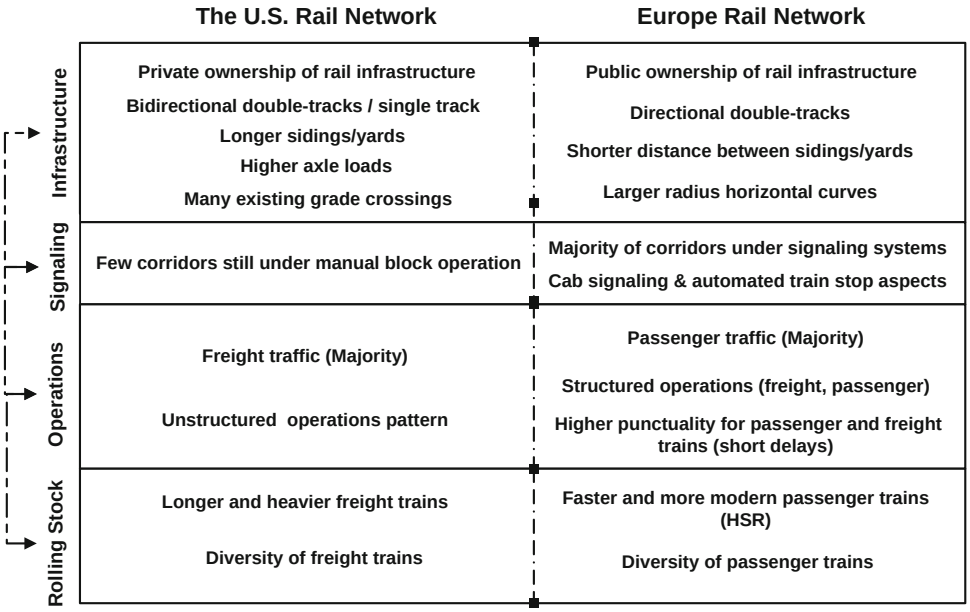

Fig. 2 The main differences in the U.S. and Europe rail systems

More than 90 % of the infrastructure is owned and managed by private freight railroads in the U.S., while in Europe almost all infrastructure is owned and managed by governments or public agencies. In addition, operations and infrastructure are vertically separated in Europe, while in the U.S. the majority of operations (mainly freight) are controlled by the same corporations who own the infrastructure. The ownership and vertical separation have wide impact in the railroad system. Perhaps the greatest effect is on the prioritization of operations and accessibility for operating companies, but other aspects, such as operations philosophy, maintenance strategy and practices, signaling and train control systems, rolling stock configuration, and capital investment strategies are also affected [4, 11].

- *Single versus double track* More than 46 % of rail corridors in Europe are at least double track [12, 13], while approximately 80 % of the U.S. rail corridors are single track [2, 4].
- *Directional versus bidirectional* Most of the U.S. double tracks operate in bidirectional fashion and use crossovers along the corridor, while directional operation with intermediate sidings and stations is the common approach in Europe [4].
- *Distance between sidings* The distances between stations and sidings in the European rail network are generally shorter than in the U.S. The average distance between sidings/stations throughout the European network (total route mileage vs. number of freight and passenger stations) is approximately four miles between sidings/stations in both UK and Germany [13, 14]. In the U.S., the distance between sidings varies greatly between corridors. On double track sections, passing sidings are typically further apart than in Europe, often more than twice the average European distance [11, 15].
- *Siding length* Siding/yard tracks in the U.S. are typically longer than the European rail network, but in many cases are still not sufficient for the longest freight trains operating today [11, 16].
- *Track conditions* Typically, railroad structure in the U.S. is designed for higher axle loads, but has tighter horizontal curves (smaller radius) and lower maximum speed operations than the European rail network [11, 16].
- *Grade crossings* There are approximately 227,000 active grade crossings along the main tracks in the U.S. [17, 18], while there are few grade crossings on the main corridors in Europe, partially due to higher train speeds. High frequency of grade crossings and difficulty of their elimination cause operational and safety challenges for increased train speeds in the U.S. [19].

3.2 Signaling characteristics

- *Manual blocking versus signaling systems* Manual blocking is absent on main passenger corridors in the U.S. today, but relatively common on lower density branch ones, including some of the lines proposed for passenger corridors. In Europe, most shared-use corridors are equipped with one of the common signaling systems [20].
- *Cab signaling* A more significant difference is the extensive use of cab signaling and enforced signal systems, such as ETMS and ATS in Europe. Implementation of automatic systems is limited in the U.S., despite the current effort to introduce the positive train control (PTC) on a large portion of corridors [11].

3.3 Operation characteristics

- *Improvised versus structured operation* While some specific freight trains (mainly intermodal) have tight schedules, the U.S. operations philosophy is based on the improvised pattern with no long-term timetable or dispatching plan. On the passenger side, the daily operation patterns of many Amtrak and commuter trains are also developed without details, anticipating improvised resolution of conflicts among the passenger trains, or between passenger and freight trains. In Europe, almost all freight and passenger trains have a regular schedule developed well in advance, known as structured operations [21].
- *Freight versus passenger traffic* The majority of the U.S. rail traffic is freight, while the majority of European rail traffic is passenger rail [4, 22].
- *Delay versus waiting time* Delay (deviation of train arrival/departure time from what was predicted/planned) and waiting time (scheduled time spent at stations for passing or meeting another train) are two fundamental concepts in the railroad operations. The waiting time concept is typically used in Europe to manage rail operations, due to the structured operations pattern with strict timetables. Delay is more commonly used in the U.S. capacity analysis as the main performance metric, while it is limited in Europe to the events that are not predictable in advance [21].
- *Punctuality* The punctuality criteria of trains are quite different in the U.S. and Europe. Amtrak's trains are considered on-time if they arrive within 15 min of a scheduled timetable for short-distance journeys (less than 500 miles) or within 30 min for long-distance trains (over 500 miles). In 2011, Amtrak's train punctuality was 77 % for long-distance trains, 84 % for short-distance trains, and 92 % for Acela trains on Northeast Corridor. According to Amtrak, more than 70 % of passenger train delays were caused either by

the freight trains performance or infrastructure failure [23]. The passenger trains in Europe have shorter average delay per train. For instance, Network Rail in the UK reported that approximately 90 % of all short-distance passenger trains had less than 5 min deviation from planned timetable, while for long-distance trains, the same was true for deviation less than 10 min [24]. In Switzerland, more than 95 % of all passenger trains are punctual with an arrival delay of 5 min or less [25]. The punctuality of European freight trains in 2003 was reported to be approximately 70 % [26].

3.4 Rolling stock characteristics

- *Train configuration* (*length and speed*) Typically freight trains in the U.S. are longer and heavier than freight trains in Europe. Based on the Association of American Railroads (AAR), the typical number of cars in a U.S. freight train varies between 63–164 cars in the West and 57–110 cars in the East, while the typical number in Europe is 25–40. From speed perspective, the average speed of intercity passenger trains in Europe is significantly faster than in the U.S. [2, 11, 16]. Freight trains also typically operate at higher speeds and with less variability in Europe.
- *Diversity of freight versus passenger trains* The U.S. rail transportation is more concentrated on the freight trains than Europe, and there is a great diversity between the types, lengths, etc., of freight trains. On the passenger side, Europe has more diverse configurations (such as speed, propulsion, train type, power assignment, HSR services, diesel, and electric multiple unit (EMU) trains) in comparison to the U.S. [2, 20].

While the principles of rail capacity remain the same in all rail networks, the above characteristics have an effect on capacity and its utilization. What remains unclear is how these differences have been considered in various capacity analysis tools and methodologies used and how much they limit the applicability of the U.S. tools in the European environment and vice versa. This paper introduces some of the common tools and methodologies, including examples of their use in past studies, but excludes any direct comparisons between the capabilities of individual tools. A more detailed (case study based) comparative analysis of selected U.S. and European simulation tools and methodologies is provided by the authors in separate papers [27, 28].

4 Capacity measurement, analytical, simulation, and combined approaches

Generally speaking, there are two main approaches to improve the capacity levels: either by applying new capital investment toward upgraded or expanded infrastructure or by improving operational characteristics and parameters of the rail services [29]. In either approach, it is necessary to assess and analyze the benefits, limitations, and challenges of the approach, often done through capacity analysis. The literature classifies capacity analysis approaches and methodologies in several different ways. Although the approaches differ, the input typically includes infrastructure and rolling stock data, operating rules, and signaling features. Abril et al. [30] classified the capacity methodologies as analytical methods, optimization methods, and simulation methods. Pachl [31] divided the capacity methodologies into two major classes: analytic and simulation. Similar categorization was used in research conducted by Murali on delay estimation technique [32]. Khadem Sameni, and Preston et al. [9] categorized capacity methods to timetable-based and non-timetable-based approaches. The capacity guidebook developed by TRB also divides capacity evaluation methods into two approaches: simple analysis, and complex simulation modeling [6]. Finally, in research conducted at the University of Illinois, Sogin, Barkan et al. [3, 33] divided capacity methods to theoretical (analytical), parametric, and simulation methods. Overall, the analytical and simulation methods are the most common methods found in the literature. For our review, we divided methods into three groups: analytical, simulation, and combined. Although the term "combined methodology" was not used commonly in the reviewed literature, it was added as a new class to address the fact that many reviewed studies took advantage of both analytical and simulation methods.

4.1 Analytical approach

The *analytical approach* typically uses several steps of data processing through mathematical equations or algebraic expressions and is often used to determine theoretical capacity of the segment/corridor. The outcomes vary based on the level of complexity of the scenario and may be as simple as the number of trains per day, or a combination of several performance indicators, such as timetable, track occupancy chart, fuel consumption, speed diagrams, etc. Analytical methods can be conducted without software developed for railroad applications, such as Microsoft Excel, but there are also analytical capacity tools specifically developed for rail applications. One example is SLS PLUS in Germany, which is used in the German rail network (DB Netz AG) for capacity estimation through analytical determination of the performance, asynchronous simulation, and manual timetable construction [34]. Figure 3 presents the different levels of analytical approach and how complexity can be added to the process to provide more detailed results. In some cases, analytical models are

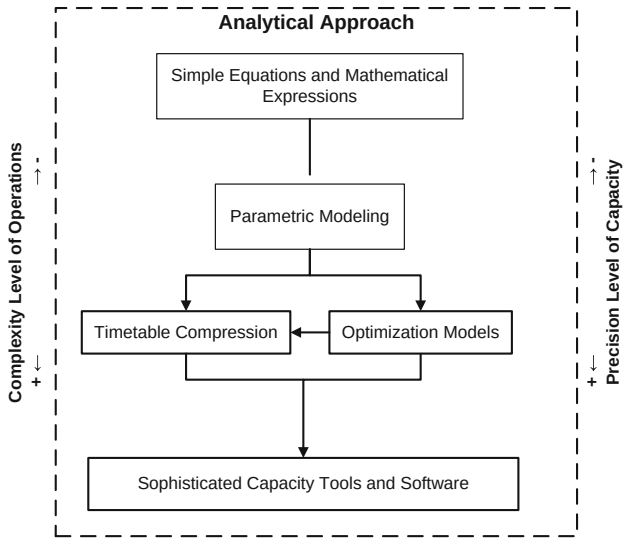

Fig. 3 Levels of analytical approaches for capacity analysis

called optimization methods or parametric models, taking advantage of different modeling features, such as probabilistic distribution or timetable optimization. The latter method, timetable optimization, is typically achieved using specialized software or simulation tools [30, 31].

Timetable compression method is one of the main analytical approaches in Europe to improve the capacity levels, especially on the corridors with pre-determined timetables (structured operation pattern). A majority of techniques and tools for improving the capacity utilization in Europe, including the UIC method (leaflet 406), are partly developed based on timetable compression [7, 10, 35–37]. The UIC's method modifies the pre-determined timetable and reschedules the trains as close as possible to each other [30]. Figure 4 provides an example of the methodology where a given timetable along a corridor with quadruple tracks (Scenario a) is first modified by compressing the timetable (Scenario b) and then further improved by optimizing the order of trains (Scenario c). As demonstrated in the figure, the third scenario could provide a higher level of theoretical capacity in comparison to the Scenarios a and b [10]. It should be emphasized that due to the unstructured nature of the U.S. rail operation philosophy, timetable compression technique has not been practically applied yet in the U.S. rail environment.

4.2 Simulation approach

Simulation is an imitation of a system's operation which should be as close as possible to its real-world equivalent [30]. In this approach, the process of simulation is repeated several times until an acceptable result is achieved by the software. The data needed for the simulation are similar to the analytical methods, but typically at a higher level of detail. The simulation practices in rail industry started in the early 1980s through the development of models and techniques, such as dynamic programming and branch-and-bound, proposed by Petersen, as well as heuristic methods developed by Welch and Gussow [30]. Today, the simulation process utilizes computer tools to handle sophisticated computations and stochastic models in a faster and more efficient way. The simulation approaches use either *general simulation* tools, such as AweSim, Minitab, and Arena [32, 38]; or *commercial railroad simulation* software specifically designed for rail transportation, such as RTC, MultiRail, RAILSIM, OpenTrack, RailSys, and CMS [9, 30]. The use of general simulation tools requires the user to develop all models, equations, and constraints step by step (often manually). This requires more expertise, creativity, and effort, but it can also offer more flexible and customization when it comes to results and outputs. The commercial railroad simulation tools offer an easier path toward development of different scenarios, in addition to providing a variety of outputs in a user-friendly way, but the core decision models and processes are not easily

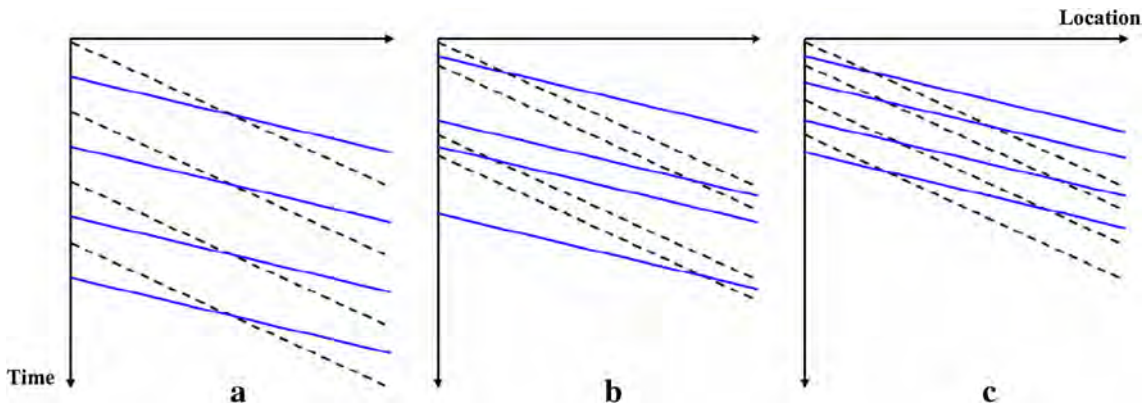

Fig. 4 Actual timetable for a quadruple-track corridor (**a**) compressed timetable with train order maintained (**b**) compressed timetable with optimized train order (**c**) (*Note* chart layout follows typical European presentation and *solid and dot lines* represent different types of trains) [10]

customizable or reviewable, which may reduce the flexibility of applying these tools.

The commercial railroad simulation software typically revolves around two key simulation components: (1) train movement and (2) train dispatching. The first component uses railroad system component data provided as an input, such as track and infrastructure characteristics (curvature and grades), station and yard layout, signaling system, and rolling stock characteristics, to calculate the train speed along the track. Train dynamics is typically determined based on train resistance formulas, such as Davis equation and train power/traction. The dispatching simulation component typically emulates (or attempts to emulate) the action of the dispatcher in traffic management, but in some cases, it can be also used as part of a traffic management software to help traffic dispatchers to manage and organize the daily train schedules (Fig. 5) [21].

According to Pachl, the simulation method can also be divided into asynchronous and synchronous methods. Asynchronous simulation software is able to consider stochastically generated train paths within a timetable, following the scheduling rules and the train priorities. In synchronous simulation, the process of rail operations is

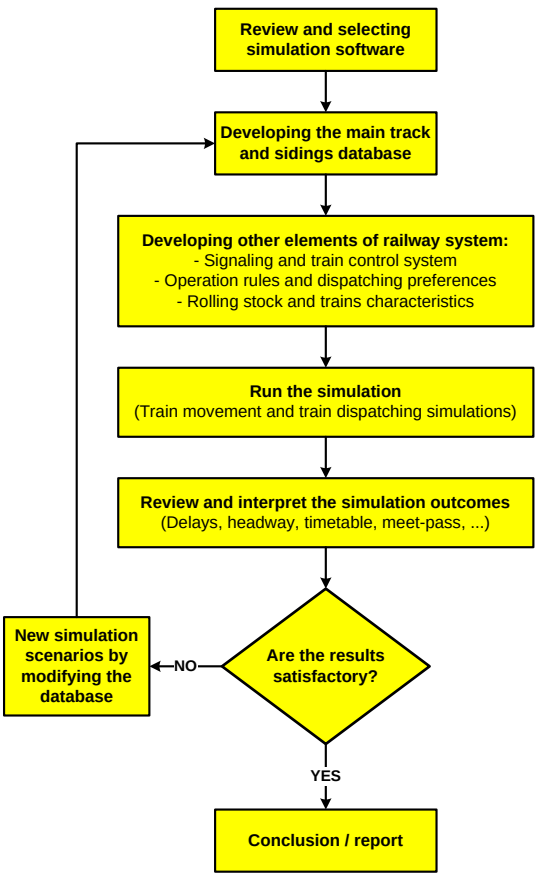

Fig. 5 Steps for railway capacity analysis in commercial simulation approach

followed in real-time sequences, and the results are expected to be closely aligned with real operations. In contrast to the asynchronous method, synchronous methods cannot directly simulate the scheduling, or develop a timetable, without use of additional computer tools and programs to create a timetable [31]. The outputs of simulation software typically include several parameters such as delay, dwell time, waiting time, elapsed time (all travel time), transit time (time between scheduled stops), trains speed, and fuel consumption of trains [21, 30].

4.2.1 Simulation methods: timetable based versus non-timetable based

The commercial railroad simulation software can be classified in two groups: non-timetable based and timetable based. The non-timetable-based simulations are typically utilized by railroads that use the improvised (unstructured) operation pattern without an initial timetable, such as the majority of the U.S. rail networks. In this type of simulation, after loading the input data in the software, the train dispatching simulation process uses the departure times from the initial station that are provided as part of the input data. The software may encounter a problem to assign all trains and request assistance from the user to resolve the issue by manually adjusting the train data, or by modifying the schedule constraints [9, 21].

The simulation procedure in timetable-based software (typically used in Europe) is based on the initial timetable of trains and the objective is to improve the timetable as much as possible. The UIC's capacity approach is often one of the main theories behind the timetable-based simulation approach. The simulation process in this methodology begins with creating a timetable for each train. In the case of schedule conflict between the trains, the user must adjust the timetable until a feasible schedule is achieved. However, the user actions are more structured compared to the improvised method, and is implemented as part of the simulation process [21]. There are several common software tools in this category, such as MultiRail (U.S.), RAILSIM (U.S.), OpenTrack (Switzerland), SIMONE (The Netherlands), RailSys (Germany), DEMIURGE (France), RAILCAP (Belgium), and CMS (UK) [9, 30]. A comprehensive capability review of various simulation tools is outside the scope of this paper, but three simulation packages (RTC, RailSys, and OpenTrack) are briefly introduced to demonstrate some key differences between non-timetable-based and timetable-based software.

The rail traffic controller (RTC), developed by Berkeley Simulation Software, is the most common software in the non-timetable-based category, used extensively by the U.S. rail industry [9]. RTC was launched in the U.S. (and North American) rail market in 1995 and has since been

continuously developed and upgraded. Since majority of the U.S. train services (particularly freight trains) have frequent adjustments in their daily schedules, RTC has several features and tools for simulating the rail operations in non-scheduled environment, including train movement animation, automated train conflict resolution, and randomization of train schedule. The dispatching simulation component of RTC is based on a decision support core, called "meet-pass N-train logic". For any dispatching simulation practice, "meet-pass N-train logic" will decide when the given trains should exactly arrive and depart from different sidings, based on the defined train priorities and preferred times of departure. The simulation outcomes may include variation between the simulated departure times and preferred times [39]. Besides its decision core fitting the U.S. operational philosophy, RTC has other system characteristics, such as attention to grade crossing, that make it well suited to the U.S. market.

RailSys, developed by Rail Management Consultants GmbH (RMCon) in Germany, is an operation management software package that includes features, such as timetable construction/slot management, track possession planning, and simulation. It has been in the market since 2000 and it is one of the commonly used timetable-based simulation software in Europe. The capacity feature of RailSys uses the UIC code 406 which is based on the timetable compression technique [40, 41]. OpenTrack is another common simulation package in Europe. It was initially developed by Swiss Federal Institute of Technology-Zurich (ETH-Zurich) and has since 2006 been supplied by OpenTrack Railway Technology Ltd. OpenTrack is also a timetable-based simulation tool with several features, such as automatic conflict resolution based on train priority, routing options and delay probabilistic functions, as well as several outputs and reporting options, such as train diagram, timetable and delay statistics, station statistics, and speed/time diagram [42, 43].

4.3 Combined analytical–simulation approach

In addition to the analytical and simulation approaches, a *combined analytical–simulation method* can also be used to investigate the rail capacity. Parametric and heuristic modelings (in analytical approach) are more flexible when creating new aspects and rules for the analysis. On the other hand, updating the railroad component input data and criteria tends to be easier in the simulation approach, and the process of running the new scenarios is generally faster, although simulation may place some limitations when adjusting the characteristics of signaling or operation rules. A combined simulation–analytical methodology takes advantage of both methodologies' techniques and benefits, and the process can be repeated until an acceptable set of outputs and alternatives is found (Fig. 6). There are several ways to combine analytical and simulation tools. For instance, finding a basic and reasonable schedule of trains through simulation, followed by analytical schedule can be considered as one example of combined analytical–simulation approach. Another example would be application of a simplistic analytical model to provide the basic inputs, such as determining the type of signaling system, or developing train schedule, followed by more extensive and detailed analysis in commercial rail simulation tools.

5 Review of capacity studies in the U.S. and Europe

The approaches, methodologies, and tools highlighted in previous section have been applied in numerous U.S. and European capacity studies. The team reviewed 51 total studies using all three approaches (17 analytical studies, 22 simulation studies, and 12 combined simulation–analytical approaches). Then, 25 of them that had sufficient details of the study approach and respective results were used to conduct a detailed assessment of studies conducted in Europe versus in the U.S.

5.1 Studies with analytical approach

One of the first analytical models was developed by Frank [44] in 1966 by studying the delay levels along a single track corridor considering both directional and bidirectional scenarios. He used one train running between two consecutive sidings (using manual blocking system) and a single average speed for each train to calculate the number of possible trains (theoretical capacity) on the given segment. Petersen [45] expanded Frank's idea in 1974 by considering two different speeds, independent departure

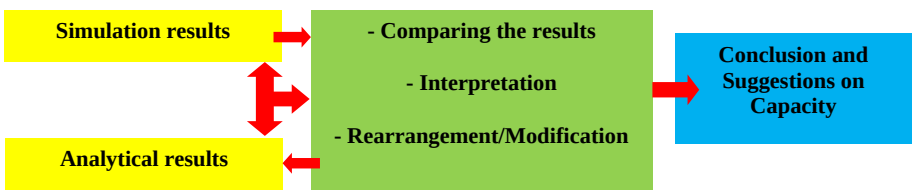

Fig. 6 Basic diagram of combined analytical–simulation approach for capacity analysis

times, equal spacing between sidings, and constant delays between two trains. Higgins et al. [46] developed a model in 1998 for urban rail networks to evaluate the delays of trains by considering different factors such as trains' schedule, track links, sidings, crossings, and the directional/bidirectional operation patterns throughout the network.

De Kort et al. [47] analyzed the capacity of new corridors in 2003 by applying an optimization method and considering uncertainty of demand levels on the planned route. Ghoseiri et al. [48] introduced a multi-objective train scheduling model of passenger trains along single and multiple tracks of rail network, based on minimizing the fuel consumption cost as well as minimizing the total passenger time of trains. Burdett and Kozan [49] developed analytical techniques and models in 2006 to estimate the theoretical capacity of a corridor based on several criteria, such as mixed traffic, directional operation pattern, crossings and intermediate signals along the track, length of the trains, and dwell time of trains at sidings or stations. Wendler [50] used queuing theory and the semi-Markov chains in 2007 to provide a technique of predicting the waiting times of trains based on the arrival times, minimum headway of trains, and the theory of blockings. Lai and Barkan [3] introduced an enhanced technique of capacity evaluation tools in 2009 based on the parametric modeling of capacity evaluation, which was initially developed by CN Railroad. The railroad capacity evaluation tool (RCET), developed by Lai and Barkan, can evaluate the expansion scenarios of network by estimating the line capacity and investment costs, based on the future demand and available budget.

Lindner [51], recently, reviewed the applicability of timetable compression technique, UIC code 406, to evaluate the corridor and station capacity. He used several case studies and examples to conclude that UIC code 406 is a good methodology for evaluating the main corridor capacity, but it may encounter difficulties with node (station) capacity evaluation. Corman et al. [52] conducted another study in 2011 to analyze an innovative approach of optimization of multi-class rescheduling problem. The problem focused on train scheduling with multiple priority classes in different steps, using the branch-and-bound algorithm.

In addition to specific studies on railroad capacity, a book edited by Hansen and Pachl [53], containing several articles and sections conducted by different railroad studies mostly by European universities and academic centers, was released in 2008 as one of the latest resources of timetable optimization and train rescheduling problem. The book covers articles on various topics, such as cyclic timetabling, robust timetabling, use of simulation for timetable construction, statistical analysis of train delays, rescheduling, and performance evaluation.

5.2 Studies with simulation and combined approach

Studies using analytical approach preceded simulation and combined approaches. One of the first general simulation studies was conducted by Petersen et al. in 1982 by dividing a given corridor into different track segments where each segment represented the distance between two siding/switches [54]. Kaas [55] developed another general simulation model in 1991, called "Strategic Capacity Analysis for Network" (SCAN), by defining different factors of simulation which could determine the rail network capacity. In another study, Dessouky et al. (1995) [56] used a general simulation model for analyzing the track capacity and train delay throughout a rail network. Their model included both single and double track corridors, as well as other network parameters, such as trains length, speed limits, and train headways. Sogin et al. [57] recently used RTC to simulate several case studies at University of Illinois, Urbana-Champaign. One of their studies evaluated the impact of passenger trains along U.S. shared-use double track corridors, considering different speed scenarios. They concluded that increasing speed gap between the trains can result in higher delays.

The Missouri DOT used the combined analytical–simulation approach in 2007 to analyze the rail capacity on the Union Pacific (UP) corridor between St. Louis to Kansas City to improve the passenger train service reliability and to reduce the freight train delay. Six different alternatives were generated based on a theory of constraints (TOC) analysis[3] and then compared with each other using the Arena simulation method. A set of recommendations and capital investment for each proposed alternative were proposed with respect to delay reduction [38].

In another project, Washington DOT (WSDOT) conducted a master plan in 2006 to provide a detailed operation and capital plan for the intercity passenger rail program along Amtrak Cascades route. The capacity of the corridor was also evaluated using the combined simulation–analytical approach. First, analytical methods were used to determine the proposed infrastructure. Then the proposed traffic and infrastructure were simulated with RTC software to test the proposed infrastructure and operational results. After running simulation on RTC software, a heuristic (analytical) method, called root cause analysis (RCA), was applied to evaluate the simulation output. The objective of RCA method was to identify the real reason of a delay along the rail corridor by comparing

[3] TOC is a management technique that focuses on each system constraints based on five-step approach to identify the constraints and restructure the rest of the system around it. These steps are: 1) identify the constraints, 2) decision on how to exploit the constraints, 3) subordinate everything around the above decision, 4) elevate the system's constraints, and 5) feedback, back to step 1.

the output reports of each delayed train with other train services and to re-adjust the simulation outputs to be more accurate, in addition to locating infrastructure bottlenecks which caused the capacity issues and delays [58].

The Swedish National Rail Administration (Banverket) carried out a research project in 2005 to evaluate the application of the UIC capacity methodology (timetable compression) for the Swedish rail network. RailSys software was used for the simulations and the research team analytically evaluated the capacity consumption, its relationship with time supplements (or buffer times), and the service punctuality. The research concluded that the buffer times are absolutely necessary for the service recovery, in case of operation interruption. When there is no buffer time, the service punctuality can be significantly degraded due to increased capacity consumption. Banverket also confirmed the validity of the framework and the results of the UIC's approach and asked their experts and consultants to implement this analytical approach in their network [36].

In research conducted through combined analytical–simulation approach, Medeossi et al. [59] applied stochastic approach on blocking times of trains to improve the timetable planning using OpenTrack simulation software. They redefined timetable conflicts by considering a probability for each train conflict as a function of process time variability. The method repeatedly simulated individual train runs on a given infrastructure model to show the occupation staircase of trains in different color spectrums, while each color represents the probability of trains' conflict which should be resolved.

Recently, a new "Web-based Screening Tool for Shared-Use Rail Corridors" was developed in the U.S. by Brod and Metcalf [60] to perform a preliminary feasibility screening of proposed shared-use passenger and freight rail corridor projects. The outcomes can be used to either reject projects or move them to more detailed analytical/simulation investigations. The concept behind the tool is based on a simplified simulation technique which does not provide optimization features or complex simulation algorithms. The tool requires development of basic levels of infrastructure, rolling stock, and operation rules (trains schedule) of the given corridor; and a conflict identifier assists the user in identifying locations for a siding or yard extension needed to resolve the conflict between existing and future train services.

5.3 Detailed assessment of selected studies

Only a subsection of reviewed studies offered sufficiently detailed explanation of the study approach and respective results. These studies were broken into several categories and subcategories for a comparison between the studies conducted in Europe versus in the U.S. Table 1 and the following discussion summarize the approach, tools used,

Table 1 Category/subcategory breakdown of 25 selected studies in the U.S. and Europe [2, 3, 9, 10, 21, 25, 29, 32, 35, 36, 38, 57–70]

Category/subcategory		The U.S. (14 studies)	Europe (11 studies)
Capacity approach	Analytical	4 studies [2, 3, 29, 32]	–
	Simulation	5 studies [57, 61, 62, 64, 65]	5 studies [10, 25, 36, 67, 69]
	Combined analytical–simulation	5 studies [9, 21, 38, 58, 60]	6 studies [35, 59, 63, 66, 68, 70]
Tools/software	Only mathematical/parametric modeling	3 studies [2, 3, 29]	–
	General simulation software	3 studies [32, 38, 60]	–
	Timetable-based simulation software	–	11 studies [10, 25, 35, 36, 59, 63, 66–70]
	Non-timetable-based simulation software	8 studies [9, 21, 57, 58, 61, 62, 64, 65]	–
Purpose of research	New methodology development/ methodology approval	5 studies [3, 9, 21, 29, 60]	7 studies [10, 25, 35, 36, 59, 66, 68]
	Master plan/capacity analysis	3 studies [2, 38, 58]	–
	Academic research/project	6 studies [32, 57, 61, 62, 64, 65]	4 studies [63, 67, 69, 70]
Type of outcomes/solutions	Delay analysis/improvement	3 studies [32, 57, 61]	1 study [69]
	Infrastructure development,	1 study [2]	–
	Rescheduling/operation changes	2 studies [21, 62]	4 studies [25, 35, 63, 67]
	Combination of above solutions	4 studies [38, 58, 64, 65]	2 studies [68, 70]
	New tools/methodology approval	4 studies [3, 9, 29, 60]	4 studies [10, 36, 59, 66]
Validation of simulation results	No comparison	6 studies [2, 9, 29, 57, 61, 62]	3 studies [25, 35, 36]
	Base model	7 studies [3, 21, 38, 58, 60, 64, 65]	7 studies [10, 59, 63, 66, 67, 69, 70]
	Base and alternative results	1 study [32]	1 study [68]

study purpose, types of outcomes, and validation methods of the 25 studies selected for more detailed comparison.

Approach Most studies used either simulation or combined analytical–simulation approaches. However, research conducted by AAR [2], University of Illinois at Urbana-Champaign (UIUC) [3, 29], and University of Southern California (USC) [32] applied analytical-only methodologies.

Tools and software All European studies used timetable-based simulation software (e.g., RailSys, OpenTrack, ROMA), while the U.S. studies relied on other tools like optimization/parametric modeling (UIUC and USC) [2, 3, 29], general simulation software (e.g., Arena) [38], web-based screening tools [60], and non-timetable-based rail simulation software (RTC).

Purpose of Research Three main purposes were identified for studies: (1) introducing new methodology for capacity evaluation, (2) evaluating the capacity status of a given corridor as part of a corridor master plan development, and (3) academic research on various capacity issues. The majority of European studies (Denmark, Austria, Germany, the Netherlands, and Sweden) were conducted by industry or academic research teams to justify and evaluate the UIC's approach (UIC code 406) for capacity evaluation [10, 35, 36, 67, 70], while the objectives of the U.S. studies included all three subcategories.

Type of outcomes or solutions The outcomes and solutions obtained from the U.S. studies included variety of different types such as delay analysis (UIUC by using RTC and USC by using Awesim/Minitab), rescheduling and recommendations related to current operations (UIUC and White) [21, 62], infrastructure development, and combination of all outcomes mentioned above (typically as part of a master plan). In addition, new tools and parametric models were also developed as the final outcome of three U.S. studies (mainly by UIUC). The outcomes of European studies were not so diverse, as they either approved the application of UIC's capacity methodology to be used on their network [10, 36], or suggested network rescheduling and operational changes (the timetable compression concept) [25, 35, 63, 67, 70]. One of the common conclusions of various studies was the identification of operational heterogeneity as a major reason of delay, especially in the U.S. rail network with unstructured operation pattern.

Validation of simulation results None of the studies using analytical method compared the results to a real-life scenario, but some of the simulation-based studies validated the results with one of the following three types of comparisons:

- *No comparison* No specific information or comparison was provided between simulated results and actual practices. As presented in Table 1, approximately one-third of the studies (9 out of 25) did not validate the simulation results, either because the study was not based on actual operational data, or comparison was not conducted as part of the research.
- *Base model* Only the results of a base model were compared with the real data. More than half of the studies (14 out of 25) compared the simulation results only with the base model.
- *Base and alternative results* In addition to base model comparison, the alternative outcomes were compared with the real data. Only two studies belonged to this category.

6 Summary and conclusions

This paper has provided an overview of capacity definitions, alternative analysis approaches, and tools available to evaluate capacity. It has also highlighted the key similarities and differences between the U.S. and European rail systems and how they affect related capacity analysis. Finally, the paper has reviewed over 50 past capacity studies and selected 25 of them for more detailed investigation,

The review revealed no single definition of railroad capacity. Rather, the definition varies based on the techniques and objectives of the specific study. The capacity analysis approaches and methodologies can also be classified in several ways, but are most commonly divided into analytical and simulation methods. This paper also introduced a third "combined" approach that uses both analytical and simulation approaches.

While the objective of capacity analysis is common, there are several differences between the U.S. and European rail systems that affect the approaches, tools, and outcomes of analysis. Europe tends to use a structured operations philosophy and thus uses often timetable-based simulation approaches for analysis, while the improvised U.S. operations warrant non-timetable-based analysis. Other factors, such as differences in ownership, type and extent of double track network, distance between and length of sidings, punctuality of service, dominating type of traffic (passenger vs. freight), and train configuration also affect the analysis methods and tools.

The review of over 50 past studies revealed that a majority of analyses (approximately 65 % of studies) utilized either simulation or combined simulation–analytical methods, while the remainder relied on analytical methods. Although the general simulation tools and modeling approaches have been used, most studies use commercial simulation software either in the U.S. (non-timetable based) or in Europe (timetable based). Based on the more detailed review of 25 of the studies, European capacity

analysis tends to be linked to the UIC 406 method, while the U.S. does not seem to have as extensive principles as the European case studies, but the methodologies vary more from one study to another. The outcomes of European studies were also less diverse than in the U.S., and commonly suggested rescheduling and operation changes as the solutions for capacity improvement. Also the studies showed limited effort in comparing the simulation results to the actual conditions (the validation step), especially after recommended improvements were implemented. Only two studies did the full validation, 14 out of 25 only compared the results with the base model, and the remaining one-third of the studies had no validation process. Overall, it was found that there was no major divergence between approaches or criteria used for capacity evaluation in the U.S. and Europe. However, there are differences in the tools used in these two regions, as the tool designs follow the main operational philosophy of each region (timetable vs. non-timetable) and include features that concentrate on other rail network characteristics for the particular region.

References

1. FRA (2009) Vision for high-speed rail in America. Federal Railroad Administration (FRA), Washington, DC
2. Cambridge Systematics I (2007) National rail freight infrastructure capacity and investment study. Association of American Railroads (AAR), Cambridge
3. Lai Y-C, Barkan C (2009) Enhanced parametric railway capacity evaluation tool. J Transp Res Board 2117:33–40. doi:10.3141/2117-05 (National Academies, Washington, DC)
4. Tolliver D (2010) Railroad planning & design, CE 456/656 Lecture Notes. North Dakota State University
5. AREMA (2006) Manual for railway engineering. Volume 4-systems management. American Railway Engineering and Maintenance-of-Way Association, Lanham
6. Fox J, Hirsch P, Kanike O, Simpson DP, Cebula AJ, Bing AJ, Horowitz ESB (2014) Capacity modeling guidebook for shared-use passenger and freight rail operations, NCHRP Report 773. NCHRP, Washington, DC
7. UIC (2004) UIC CODE 406 R-capacity, 1st edn. International Union of Railways (UIC), Paris
8. IMPROVERAIL (2003) IMPROVEd tools for RAILway capacity and access management-deliverable 6 methods for capacity and resource management. Competitive and Sustainable Growth (GROWTH) Programme—European Commission, Brussels-Belgium
9. Sameni MK, Dingler M, Preston JM, Barkan CPL (2011) Profit-generating capacity for a freight railroad. In: TRB 90th Annual Meeting, TRB, Washington, DC
10. Alex Landex EA (2006) Evaluation of railway capacity. In: Annual Transport Conference, Aalborg University, Denmark
11. White T (2012) Comparison of U.S. and European railroads, Lecture notes for CE-4490. Michigan Tech. University-CEE Department, Houghton
12. Railway-Technical (2008) Railway statistics for Britain. http://www.railway-technical.com/statistics.shtml. Accessed May 2012
13. ERRAC (2004) A comparison of member state public research programmes with the ERRAC SRRA 2020. European Rail Research Advisory Council (ERRAC), Brussels
14. DB-AG (2011) Station categories 2012. DB-AG, Berlin-Germany (in German)
15. Amtrak (2011) Annual report fiscal year 2011. Amtrak, Washington, DC
16. Teixeira P (2009) Railway engineering principles, Technical University of Lisbon, CTIS Master's course in MIT-Portugal Program, Lisbon
17. FRA (2008) Highway-rail grade crossing safety fact sheet. FRA Public Affairs, Washington, DC
18. DOT (2011) National transportation statistics. U.S. Department of Transportation, Washington, DC
19. NetworkRail (2010) Curtail crazy driving at level crossings or risk more lives, says rail chief. http://www.networkrailmediacentre.co.uk/Press-Releases/CURTAIL-CRAZY-DRIVING-AT-LEVEL-CROSSINGS-OR-RISK-MORE-LIVES-SAYS-RAIL-CHIEF-13c1/SearchCategoryID-2.aspx.Accessed on 2012
20. Verkehrswissenschaftliches (2008) Influence of ETCS on the line capacity. International Union of Railways (UIC), Paris
21. White T (2005) Alternatives for railroad traffic simulation analysis. J Transp Res Board 1916:34–41 (TRB-Washington, DC)
22. Systra (2007) Technical classification and main technologies for the HSR systems. Complementary studies of Tehran-Qom-Isfahan HSR, Paris
23. Amtrak (2012) Monthly performance report for December 2011. AMTRAK, Washington, DC
24. Wearden G (2007) UK train times back on track. The Guardian, Friday 23 November 2007
25. Marco Luethi AN, Weidmann U (2007) Increasing railway capacity and reliability through integrated real-time rescheduling. Swiss Federal Railways, SBB AG, Zurich
26. CER (2004) Rail freight quality: meeting the challenge, A Report on the first year of the CER-UIC-CIT Freight Quality Charter. Community of European Railway and Infrastructure Companies (CER), Bruxelle, Belgium
27. Pouryousef H, Lautala P (2013) Evaluating the results and features of two capacity simulation tools on the SHARED-use corridors. In: ASME/ASCE/IEEE 2013 Joint Rail Conference (JRC2013), Knoxville, TN, USA
28. Pouryousef H, Lautala P (2014) Evaluating two capacity simulation tools on shared-use U.S. rail corridor. In: 2014 Annual Meeting of the Transportation Research Board, TRB, Washington, DC
29. Lai Y-C, Barkan C (2011) A comprehensive decision support framework for strategic railway capacity planning. J Transp Eng 137:738–749. doi:10.1061/(ASCE)TE.1943-5436.0000248 (ASCE 137 (October 2011))
30. Abril M, Barber F, Ingolotti L, Salido MA, Tormos P, Lova A (2007) An assessment of railway capacity. Technical University of Valencia, Valencia, Spain
31. Pachl J (2002) Railway operation and control. VTD rail publishing, Mountlake Terrace
32. Pavankumar Murali MMD, Ordonez F, Palmer K (2009) A delay estimation technique for single and double-track railroads. Department of Industrial & Systems Engineering, University of Southern California, Los Angeles, CA
33. Sogin SL, Barkan CPL (2012) Railroad capacity analysis. In: Railroad engineering education symposium (REES) by AREMA, Kansas City

34. Schultze GI (2012) SLS PLUS (Software Manual). Schultze + Gast Ingenieure, Berlin
35. Robert Prinz JH (2005) Implementation of UIC 406 capacity calculation method at Austrian Railway (OBB). Austrian Railway (OBB), Vienna
36. Banverket (2005) Application of the UIC capacity leaflet at Banverket. Banverket, Swedish National Rail Administration
37. Melody Khadem Sameni MD, Preston JM (2010) Revising the UIC 406 method: revenue generating capacity- JRC2010-36281. In: Joint Rail Conference (JRC) 2010, Urbana, IL, USA, 2010
38. Noble JS, Charles N (2007) Missouri freight and passenger rail capacity analysis. Missouri Department of Transportation (MoDOT), Jefferson City
39. RTC (2011) RTC online help (Manual). Berkeley Simulation Software, CA, USA
40. RMCon (2012) RailSys 8 classic, Railsys 8 Enterprise, network-wide timetable and infrastructure management. Rail management consultants GmbH (RMCon), Hannover
41. RMCon (2010) RailSys user manual. Rail management consultants GmbH (RMCon), Hannover
42. Huerlimann D, Nash AB (2012) Manual of opentrack-simulation of railway networks, Version 1.6. OpenTrack Railway Technology Ltd. and ETH Zurich, Zurich, Switzerland
43. Huerlimann D (2012) OpenTrack—simulation of railway networks. OpenTrack Railway Technology Ltd., Zurich
44. Frank O (1966) Two-way traffic in a single line of railway. Oper Res 14:801–811
45. Petersen ER (1974) Over the road transit time for a single track railway. Transp Sci 8:65–74
46. Higgins A, Kozan E (1998) Modeling train delays in urban networks. Transp Sci 32(4):251–356
47. de Kort AF, Heidergott B, Ayhan H (2003) A probabilistic approach for determining railway infrastructure capacity. Eur J Oper Res 148:644–661
48. Ghoseiri K, Szidarovszky F, Asgharpour MJ (2004) A multi-objective train scheduling model and solution. Transp Res Part B 38:927–952
49. Burdett RL, Kozan E (2006) Techniques for absolute capacity determination in railways. Transp Res Part B 40:616–632
50. Wendler E (2007) The scheduled waiting time on railway lines. Transp Res Part B 41:148–158
51. Lindner T (2011) Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. J Rail Transp Plan Manag 1:49–57
52. Corman F, D'Ariano A, Hansen IA, Pacciarelli D (2011) Optimal multi-class rescheduling of railway traffic. J Rail Transp Plan Manag 1:14–24
53. Hansen IA, Pachl J (2008) Railway timetable & traffic, analysis, modeling, simulation, 1st edn. Eurailpress, Hamburg
54. Petersen ER, Taylor AJ (1982) A structured model for rail line simulation and optimization. Transp Sci 16:192–206
55. Kaas AH (1991) Strategic capacity analysis of networks: developing and practical use of capacity model for railway networks. ScanRail consult, Technical University of Denmark
56. Dessouky MM, Leachman RC (1995) A simulation modeling methodology for analyzing large complex rail networks. Simulation 65(2):131–142
57. Sogin SL, Barken CPL, Lai Y-C, Saat MR (2012) Impact of passenger trains in double track networks. In: 2012 Joint rail conference, Philadelphia, Pennsylvania
58. Transit Safety Management Inc.; HDR Engineering (2006) Amtrak cascades operating and infrastructure plan—technical report. Vol 1. Washington State Department of Transportation-Freight Systems Division
59. Medeossi G, Longo G, Fabris Sd (2011) A method for using stochastic blocking times to improve timetable planning. J Rail Transp Plan Manag 1:1–13
60. Brod D, Metcalf AE (2014) Web-based screening tool for shared-use rail corridors (NCFRP REPORT 27). National cooperative freight research program-Transportation research board (TRB), Washington, DC
61. Sogin SL, Barkan CPL, Lai Y-C, Saat MR (2012) Measuring the impact of additional rail traffic using highway & railroad metrics. In: 2012 Joint rail conference (JRC), Philadelphia, Pennsylvania
62. Dingler M, Lai Y-C, Barkan CPL (2009) Impact of operational practices on rail line capacity: a simulation analysis. In: 2009 Annual AREMA conference, Chicago, IL
63. Siggel S (2006) Analysis of capacity and optimization of infrastructure of the line minot (North Dakota)—Havre (Montana). Braunschweig Technical University, Braunschweig-Germany
64. Atanassov I, Dick CT, Barkan CPL (2014) Siding spacing and the incremental capacity of the transition from single to double track. In: 2014 Joint rail conference (JRC), Colorado Springs, CO, USA
65. Shih M-C, Dick CT, Sogin SL, Barkan CPL (2014) Comparison of capacity expansion strategies for single-track railway lines with sparse sidings. In: TRB 2014 Annual meeting, Washington, DC, USA
66. Schlechte T, Borndörfer R, Erol B, Graffagnino T, Swarat E (2011) Micro–macro transformation of railway networks. J Rail Transp Plan Manag 1:38–48
67. Gille A, Siefer T (2013) Sophisticated capacity determination using simulation. In: Transportation research board 2013 Annual meeting, Washington, DC, USA
68. Medeossi G, Longo G (2013) An approach for calibrating and validating the simulation of complex rail networks. In: TRB 2013 Annual meeting, Washington, DC, USA
69. Sipila H (2014) Evaluation of single track timetables using simulation. In: 2014 Joint rail conference (JRC), Colorado Springs, CO, USA
70. Goverde RMP, Corman F, D'Ariano (2014) A investigation on the capacity consumption at Dutch railways for various signalling technologies and traffic conditions. In: TRB 2014 Annual Meeting, Washington, DC, USA

EVALUATING TWO CAPACITY SIMULATION TOOLS ON SHARED-USE U.S. RAIL CORRIDOR

Hamed Pouryousef*

PhD Candidate, Rail Transportation Program,
Civil and Environmental Engineering Department,
Michigan Technological University,
1400 Townsend Drive, Houghton, MI 49931
phone: (906) 231-2320
hpouryou@mtu.edu

Pasi Lautala, Ph.D, P.E.

Assistant Professor
Director, Rail Transportation Program,
Department of Civil and Environmental Engineering,
Michigan Technological University,
1400 Townsend Drive, Houghton, MI 49931
phone: (906) 487-3547
ptlautal@mtu.edu

*** Corresponding Author**

Submitted for Presentation at the 2014 Annual Meeting of the Transportation Research Board
August 1, 2013

Total words = 4733 + 250*11 (8 Figures + 3 Tables) = 7483

ABSTRACT

Most passenger rail services in the United States (U.S.) operate on corridors that are shared with freight traffic. As the demand for passenger and freight transportation grows and emphasis is placed on increased speed and on-time performance of passenger services, the available capacity becomes even more consumed. When higher speed passenger trains are mixed with freight, the increased heterogeneity from expanding speed differential creates further challenges for reliable operations. Based on the experiences in the other parts of the world (particularly in European rail corridors), the required density and reliability is typically secured through structured/planned/scheduled operations instead of the unstructured, or improvised, operations philosophy that is currently prevalent in the U.S.

There are several tools and methodologies available in both the European and U.S. rail environments that utilize user defined infrastructure specifications, operational rules, signaling systems and rolling stock characteristics to evaluate capacity. This paper introduces the main components of two simulation software packages, U.S. developed Rail Traffic Controller (RTC) and European RailSys, and applies them both to a shared-use case study corridor in the U.S. The outputs from each package are compared and the non-timetable based software output (RTC) is applied in the timetable based software (RailSys) as input to form a hybrid model that allows the utilization of timetable compression techniques.

The research revealed that simulation outputs from both software packages are very similar, if the trains can be operated according to initial arrival/departure times on the corridor. However, RTC's database and timetable parameters are easier to implement, while RailSys has more timetable management features and options that can be used to improve an existing timetable when introducing new trains running along the corridor.

INTRODUCTION

Railway capacity is a complex concept and rail organizations around the world use various definitions for capacity. One simple definition of capacity is the number of trains that can safely pass along a given segment through a period of time. Capacity is affected by different system configurations, such as: 1) Track infrastructure; 2) Signaling system; 3) Operations philosophy; and 4) Rolling stock. Differences between the U.S. and European rail systems, such as system ownership and type and extent of double track network, also affect capacity and its utilization.[1] Simulation software is commonly applied to evaluate the capacity utilization, but the characteristics and features of each package must be adjusted to meet the characteristics of the specific network being investigated. The configurations and parameters mentioned above may be considered at various level of detail, mainly based on the region where the software is used. The same is true for the logic behind core decisions made by simulation software and how much detail is included when building the required database of a given case study.

A review of capacity simulation tools commonly used in the U.S. and Europe can help researchers to evaluate the potential advantages and challenges of expanding the application of these tools to the other side of Atlantic. Since some of the software packages are based on timetables and some are not, there is also a potential to utilize these tools collaboratively in a hybrid approach where initial simulation results on non-timetable software can be used as inputs on the timetable based software to investigate further improvements in capacity utilization and timetable development.

This paper focuses on two major simulation tools from the U.S. and Europe, RTC and RailSys, respectively, to evaluate the use of a hybrid approach on a real-life case study in the U.S. In the first part of the paper, different tools and methodologies for capacity analysis will be briefly reviewed in both the European and U.S. rail environments. The case study used for this research (section of Northeast

Corridor) will be briefly introduced in the second part of the paper including review of inputs; infrastructure, signaling, rolling stock, and operations characteristics. The research presented in this paper considered the selected section as a stand-alone piece of infrastructure, neglecting any continuation of routes in either end. The objective of the research was not to evaluate or recommend any changes to current NEC operations, but rather to use actual infrastructure and train operational data to understand the capabilities of simulation tools and theories behind them in larger context.

The third part of the paper provides an overview on the main features and components of RTC and RailSys, as well as explanation of different scenarios applied in the capacity analysis on the case study. It also reviews the outcomes of using both simulation tools on the given case study. Finally, the conclusions and next steps of the research are briefly summarized in the last part of the paper.

REVIEW OF CAPACITY METHODOLOGIES AND TOOLS

Several methodologies and tools can be used to evaluate the capacity utilization of any rail corridor or system. Typically, methodologies can be classified in three main approaches; analytical, simulation and combined analytical-simulation. Analytical and simulation approaches are most commonly found in the literature,[2-5] but there are also several examples of the combined approach that requires the use of both analytical and simulation tools. More details regarding capacity methods have been explained by Pouryousef, et al.[1]

There are several parameters which affect the capacity utilization and different tools place varying weight on individual parameters and attributes, mainly based on the network and operating characteristics of the region they were designed for. Although the U.S. and European rail networks have several similarities, the differences between these two regions affect the selection of capacity tools and methodologies and how they incorporate infrastructure, signaling, operation rules and rolling stock specifications. More detailed description about key differences between network characteristics in Europe and the U.S and their impact on capacity are discussed by Pouryousef, et al.[1] and 2010 Sameni, et al.[6]

The **commercial rail simulation packages,** such as RTC, Railsim, OpenTrack, RailSys, and CMS [3, 5], are commonly used tools to evaluate capacity and rail operations features in many rail networks including Europe and North America. They are typically divided to two major groups; 1) **non-timetable based** vs. 2) **timetable based** software. The non-timetable based simulations are typically applied in railways which are operated based on the unstructured or improvised operation pattern and may have no initial train timetable, such as the majority of the U.S. rail network. The simulation procedure in the timetable based software (typically used in Europe) uses the initial timetable provided for each specific train in the beginning of simulation to improve the capacity utilization and level of service attributes of the original timetable. In the case of schedule conflict between trains, the user must change the timetable until the feasible schedule is achieved; however, the user interference is not arbitrary as in the improvised method, but it is implemented as part of the simulation process.[7] More details on these two types of simulations are explained in a separate paper by Pouryousef and Lautala. [8]

CASE STUDY OF THE U.S. SHARED-USE CORRIDOR

Objective

While several simulation tools are used in both the U.S. and European rail networks, the impact of tool selection on the outcomes has rarely been researched. In addition, the potential to combine the strengths of two separate tools might offer benefits over a single tool, even though the increased input effort may limit such use in industry applications. To address these issues, the study was conducted with

an objective to 1) run two simulation tools on a single U.S shared-use corridor case study and highlight the advantages and challenges of using each tool and 2) apply a hybrid approach (combining the input/output of these two packages) to improve the outcomes of one or both simulations.

Review of Case Study Characteristics

The case study selected for the research was a short segment of the Northeast Corridor (NEC) between Baltimore and Washington, DC. The selected segment is one of the most congested and complicated corridors in the U.S. rail network, in terms of:

- Number of trains per day,
- Diversity of train types,
- Operation of the only high speed train service in the U.S. (Acela Express),
- Complexity of signaling systems (both wayside and cab signaling systems), and
- Number of tracks along the corridor (Sections with triple and quadruple tracks).

The research used all existing tracks, sidings, crossovers and signaling systems along the section. All existing passenger and commuter trains running along this segment of the corridor (141 daily trains in both directions) have been considered, although the initial analysis presented in this paper used 40 randomly selected trains, to reduce the complexity and research time required during the first phase of the study. The objective is to replicate the study with full schedule of 141 daily trains in the next phase. Courtesy of Amtrak, the researchers were able to secure a complete RTC database as input, which was also used to develop the four database categories in the RailSys simulation software.

Infrastructure Characteristics

The case study's infrastructure contains 40.6 miles of triple track, (about 5 miles of quadruple and about 1.5 miles of double track rail) with several crossovers and intermediate stations/ platforms along the corridor (FIGURE 1). Horizontal and vertical alignments were accurately developed for both RTC and RailSys input database and are summarized in TABLE 1.

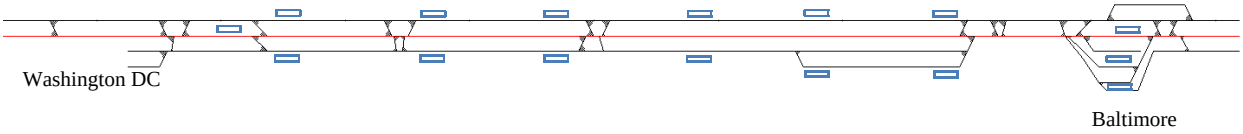

FIGURE 1- Snapshot of the case study infrastructure between Washington DC- Baltimore

TABLE 1- Details of case study infrastructure

Corridor Length	40.6 miles
Sidings/yards	2 main yards + 7 station platforms
Max. vertical grade	2.12%
Curvature	0.01 - 7.27 degrees
Length of double track	1.48 miles
Length of triple track	33.94 miles
Length of quadruple track	5.18 miles
Turnout #s	# 32.5, # 15 (one crossover)

Signaling Characteristics

The signaling system included a wayside system of automatic permissive block (APB) under CTC control system, together with a cab signaling system. These two systems have been integrated and work in unison to improve the capacity and safety levels of the corridor. All trains running through NEC are required to be equipped with working cab signals. In case of failure of the cab signals en route, the dispatcher grants permission for movement in the absolute block between each interlocking, with a 79 mph speed limit.

Rolling Stock Characteristics

Four types of trains have been considered in the case study; Long-distance passenger, commuter, Regional Amtrak, and high speed trains (Acela). The characteristics of each train type have been closely derived from the actual configurations of current rail services along the corridor. It should be pointed out that NEC (including the Baltimore-Washington, DC section) is one of the few electrified corridors in the U.S. Therefore, some of the trains considered in this case study (including Acela trains) are electrified and use overhead power supply system. Since the type and configuration of pre-programmed locomotives are fairly different in the RTC and RailSys database, some of the main characteristics of locomotives (such as power, weight, length, axle load, acceleration/ deceleration rate, and resistance) were included in the RailSys database as new locomotive type. The main characteristics of rolling stock used in the case study are presented in TABLE 2.

TABLE 2- Main features of case study's trains

Train	Daily trains (pairs)	# of cars	Trailing weight (ton)	Trailing length (feet)
Acela	10	6	378	649
Long-distance Amtrak	10	9	450	816
Regional Amtrak	10	7	385	744
Commuter	10	5	175	483

Operation Rules

There are several operation rules for simulation, such as the train's priority, speed limits, stopping patterns, and preferred time and order of train departures. The priority of different types of trains in diminishing order was Acela Express, commuter trains, Regional and long-distance passenger trains. In the case study, the maximum speed of Acela trains was 137 mph, but its practical speed was calculated by the software based on the track profile and reduced speed limits along the track, e.g. due to crossovers. Intercity passenger trains were limited to 110 mph; while commuter trains were limited to 90 mph. The initial speed of all trains from Washington, DC toward Baltimore (Northbound direction) was 30 mph when they reached the track segment starting the simulation process. For the southbound direction, the initial speed of trains had to be maintained in 30 mph for approximately 1.2 miles, due to the technical requirements along Baltimore- Bridge interlocking section. There are various stop patterns by different trains, but all trains stop at Baltimore and DC. For example, some Acela trains have no other planned stops at the intermediate sidings/platforms. The predefined arrival/departure times and preferred priority of trains have been considered for all trains according to daily operation practices.

CAPACITY ANALYSIS, REVIEW OF TRAIN TIMETABLE

Brief Introduction of Applied Tools

RTC and RailSys used in the research are two well-established commercial railway capacity analysis tools. TABLE 3 provides a comparison of some of the features and characteristics of RTC and RailSys.

TABLE 3- Comparison between RTC and RailSys[8]

Criteria	RTC	RailSys
Developer	Berkeley Simulation Software, LLC, (USA)	Rail Management Consultants GmbH (RMCon) (Germany)
Features and Modules	- Animation of traffic flow - Time-distance diagrams (stringline) - TPC profile - Track occupancy chart - Detailed train status - Timetable at various level of detail - Operating statistics at the individual train level or summarized by train type or at a system-wide level - Graphical network interface	-Infrastructure manager -Timetable construction -Capacity Management (UIC code 406) - Track Possession planning -Simulation Manager -Rolling stock circulation planning - Graphical Timetable -Platform and track occupation diagram - Graphical network interface -Delay statistics
Simulation Category	Non-timetable based simulation	Timetable based simulation (UIC code 406)
Capacity Metrics	Delay statistics, Track occupation time, time-distance diagram	Delay statistics, infrastructure occupation time, optimized timetable
Example Users	Class 1 RRs: (UPRR, BNSF, CSX, NS, KCS, CN, CP, Amtrak), U.S. railway consultants, urban rail transit agencies	Many European rail operators and consultants, international rail companies

RTC was launched in the North America's rail market in 1995 and has since been continuously developed and upgraded for a variety of simulation practices. RTC can be categorized as non-timetable based simulation software used predominantly for improvised operation philosophy conditions (the dominant operations approach in the U.S. rail environment). It is developed by Berkeley Simulation Software and it is the most common package in this category, used extensively by the U.S. rail industry. In this type of simulation, after loading the input data in the software, the train dispatching simulation process improvises the train departure times from the originating station provided as part of the input data. However, it can also receive the preferred, or scheduled, arrival and departure times of different trains for the simulation process through user input. The dispatching simulation component of RTC is based on a decision support core, called "meet-pass N-train logic". For any dispatching simulation practice, "meet-pass N-train logic" will decide when the given trains should exactly arrive and depart from different sidings, based on the defined train priorities and preferred times of departure. The simulation outcomes may include variation between the simulated departure times and preferred times. [9]

RailSys developed by Rail Management Consultants GmbH (RMCon) in Germany, is an operation management software package tool that includes infrastructure data management, timetable construction/slot management, track possession planning, and simulation features. It has been in the market since 2000 and it is one of the most common timetable-based simulation software used in Europe. The capacity feature of RailSys uses the UIC code 406 which is based on the timetable compression technique. Given train timetables, a segment of the route is selected to automatically compress the utilized train-paths, while considering the minimum headways and acceptable buffer times between the trains.

The compression technique always begins at the start of the calculation period and ends after the calculation period is fully occupied by the last possible train. The remaining usable level of capacity is identified by the number of new train-paths available, until the given time period is saturated by the train-paths and buffer times.[10, 11]

Outcomes of RTC Simulation

The case study simulation results obtained from RTC are presented in distance-time diagram format (train string-line) (FIGURE 2). Since the RTC database and schedule were prepared by Amtrak authorities, there is no deviation between the simulated arrival/departure times and the requested times (the initial departure/arrival times requested by software user) in RTC's database.

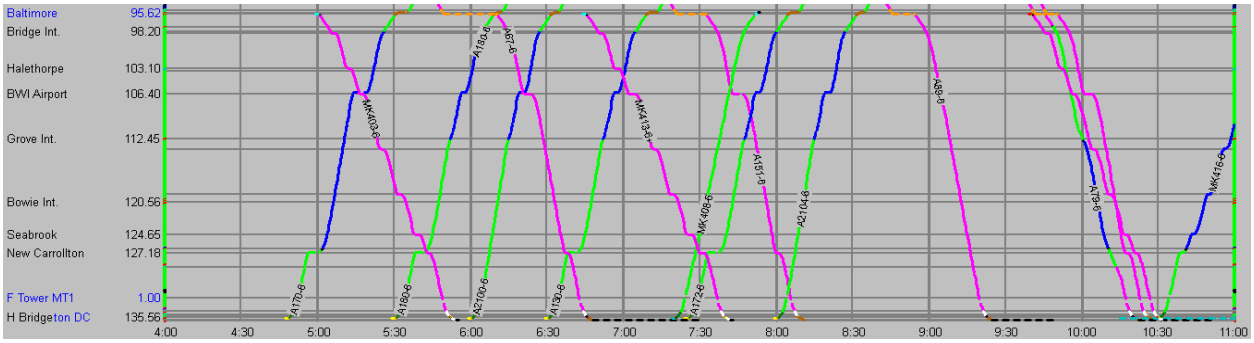

FIGURE 2- Simulated train string-line schedule in RTC, 4 am -11 am, (each color represents different track)

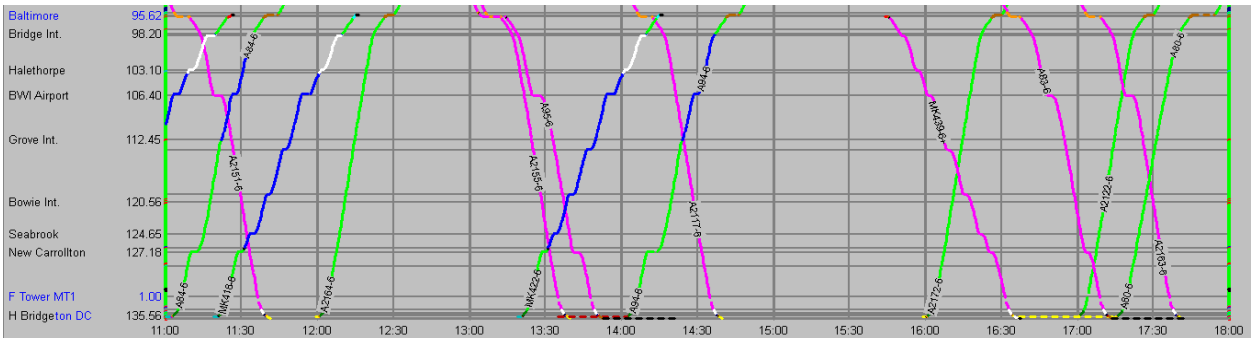

FIGURE 2 - (Continued) Simulated train string-line schedule in RTC, 11 am -6 pm

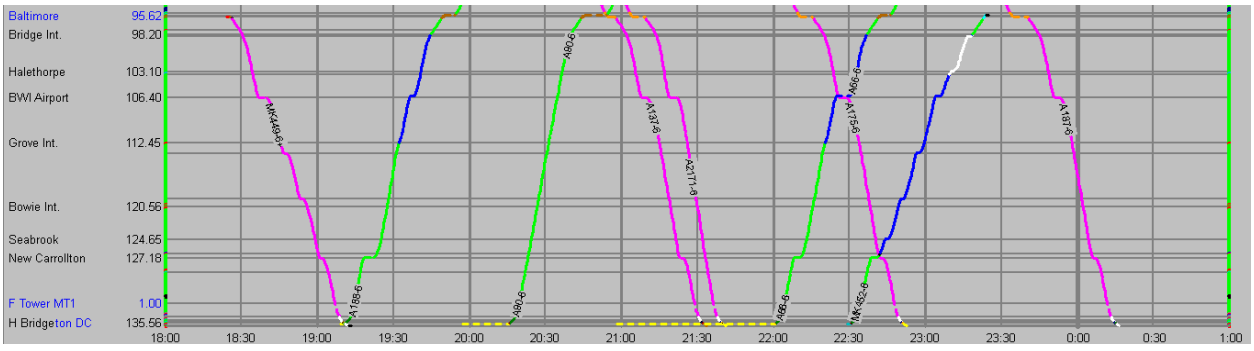

FIGURE 2 - (Continued) Simulated train string-line schedule in RTC, 6 pm -12 am

Although the schedule of all simulated trains in RTC followed precisely the requested times of the input, RTC does not have a complete package of tools to determine a schedule conflict between two or more trains before running the simulation. During the simulation, the software automatically resolves each schedule conflict as a dispatcher would resolve them, based on "meet-pass N-train logic", and

displays the impacts of the conflict resolutions both graphically and in terms of run times. In case a conflict between trains is identified by the software, a user intervention is needed to modify the schedule of trains and avoid the conflict. Such interventions are facilitated by the user-friendly animation tools of RTC which can help the software users to understand and analyze updates on train routing and signaling features, as necessary.

Outcomes of RailSys Simulation

The infrastructure characteristics (including main lines, gradient, curvatures, crossovers and sidings), rolling stock (type and number of trains), signaling systems (both permissive and cab-signaling systems) and operation rules (preferred timetable of trains, stop patterns of each train, speed limit along crossovers, train priorities) were developed in RailSys based on the database and network characteristics obtained from RTC simulation software. RailSys implementation required certain conversions, such as conversion of track curvatures from degree to radius and adjustment of rolling stock characteristics to SI units. FIGURE 3 shows the string-line train schedule of simulated trains in RailSys.

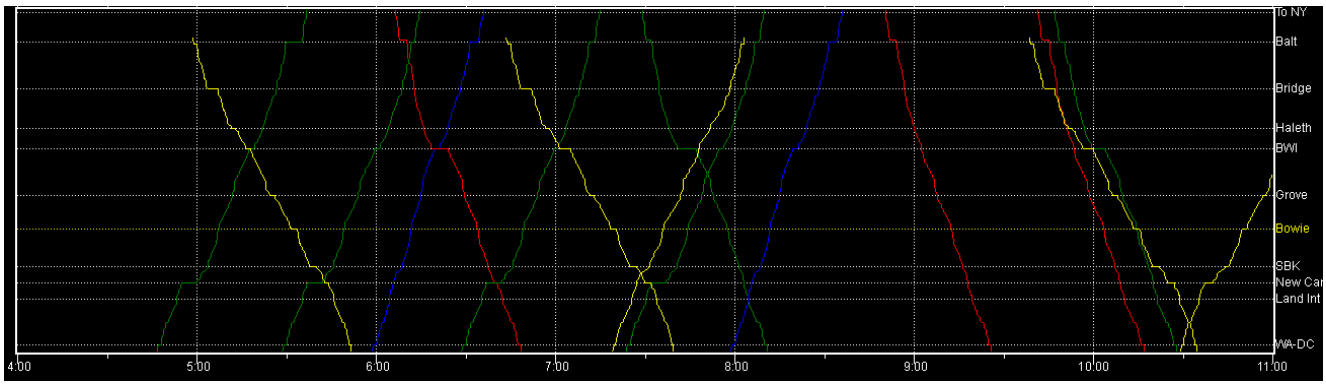

FIGURE 3- Simulated train string-line schedule in RailSys, 4am -11 am, (Green: Regional Amtrak, Red: Long-distance Amtrak, Blue: Acela, Yellow: Commuter)

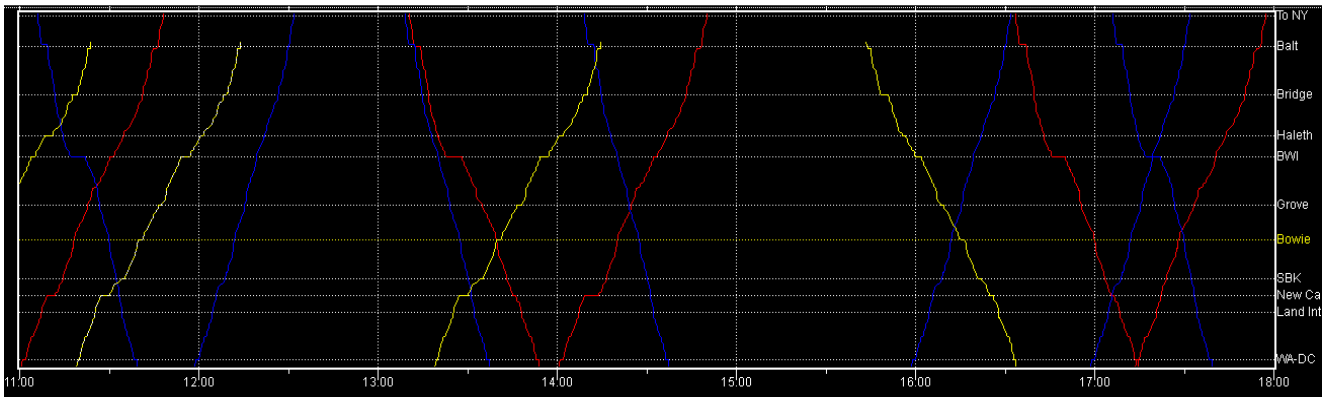

FIGURE 3 (Continued) Simulated train string-line schedule in RailSys, 11am -6 pm

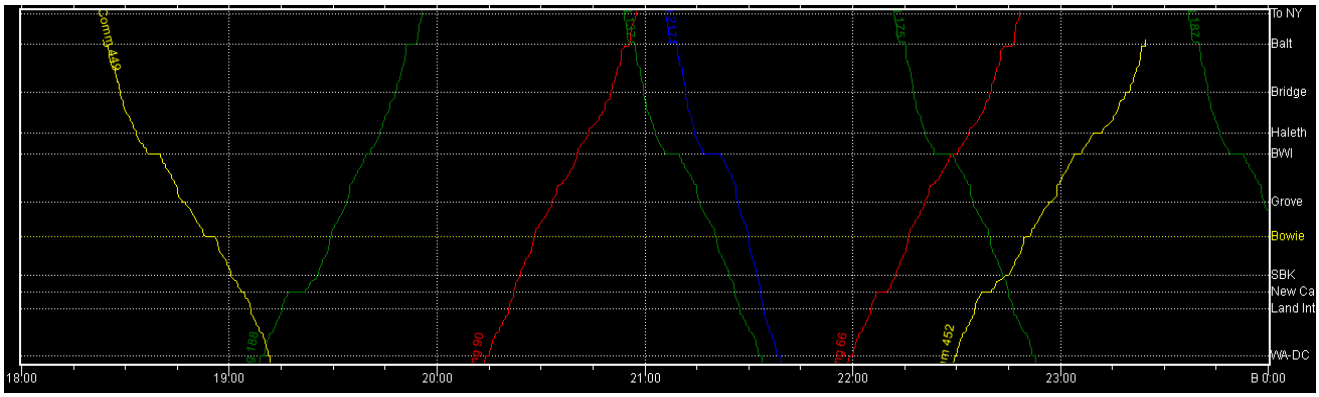

FIGURE 3 (Continued) Simulated train string-line schedule in RailSys, 6 pm -12 am

As presented in FIGURE 3, train schedules in RailSys match the same arrival and departure times as in RTC with some minor deviations between arrival/departure times (from couple of seconds up to approximately two minutes). The deviations were caused by variations of simulated train running times along the corridor, mainly due to minor differences between rolling stock and signaling features/equations in RailSys vs. RTC (such as tractive effort of engines, acceleration, deceleration, braking diagram, etc.). Overall, the simulated outcomes obtained from RailSys matched almost 90% of the requested departure/arrival times.

In some cases the tracks used by each train in RailSys differed from those in RTC, as the train routing in multiple-track corridors is dependent on user decisions. The general principle of train routing in RailSys was to allocate the first track for southbound trains (Baltimore to DC) and use the second, third and fourth tracks for northbound trains (DC to Baltimore). The second track was also used for non-stop trains (Acela or long-distance Amtrak trains) in both directions. There were significant differences how trains were routed through the stations. For example, at Baltimore, all 40 trains used in the research were routed along tracks 1 through 4, while tracks 5-7 saw no activity. On the other hand, at BWI all tracks were utilized by trains, since they were in reality extensions of the main line tracks. There were also significant differences between percentages of occupation of each track. The average percentage a track was occupied varied between 1.42% and 7.28%, during whole operation hours, and between 3.55% and 25.48% per hour during peak times.

Capacity Analysis and Applying Timetable Compression Technique on the Case Study

Since the requested times for trains were already developed by Amtrak, both RTC and RailSys successfully used the input to develop train schedules/timetables. To analyze the capability of selected tools to address a revision to daily operations, two different scenarios were introduced:

- **Scenario 1:** A new freight train with potential conflict with other train schedules
- **Scenario 2:** Evaluating the timetable compression technique of the existing schedule (RailSys only)

Scenario 1: New Freight Train

A demand frequently arises to run a new freight/passenger service along the existing tracks, in addition to the current trains. In this scenario, a new southbound freight train was introduced to depart around 9:50 am from Baltimore to Washington, DC. There were no requested intermediate stops and its departure time could be changed, if there were any schedule conflict. As shown in Figure 4, RTC dispatched the freight train after all other current passenger and commuter trains, (around 12:10 am next day instead of 9:50 am) due to the fact that the priority of this train was much lower than other trains and

earlier dispatch would have introduced conflicts between the schedules of new and current trains (assuming no change in train priority). However, RTC could resolve the conflict differently, if train priorities were manually adjusted.

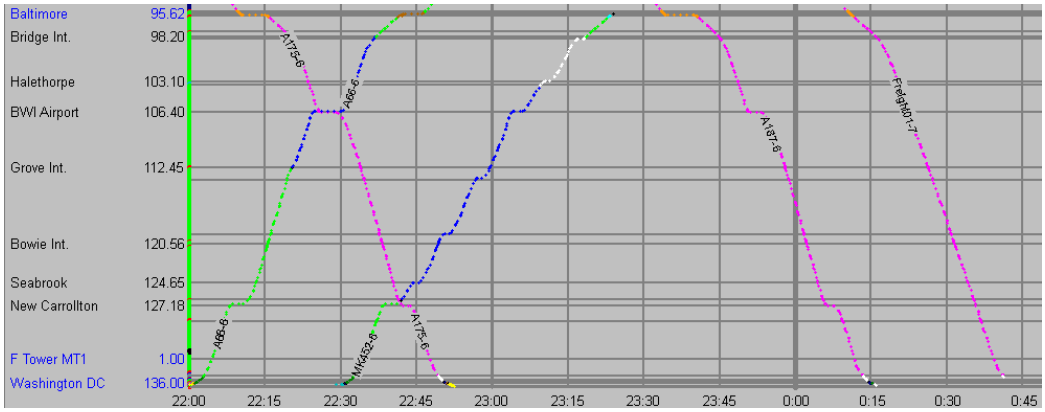

FIGURE 4- The simulated freight train was dispatched in RTC after all other trains, despite its initial requested departure time assumed at 9:50 am

RailSys recognized the conflicts between the new and current trains as well and used its supportive features of train conflict management to identify (graphically and in table-based format) where these conflicts took place (FIGURE 5).

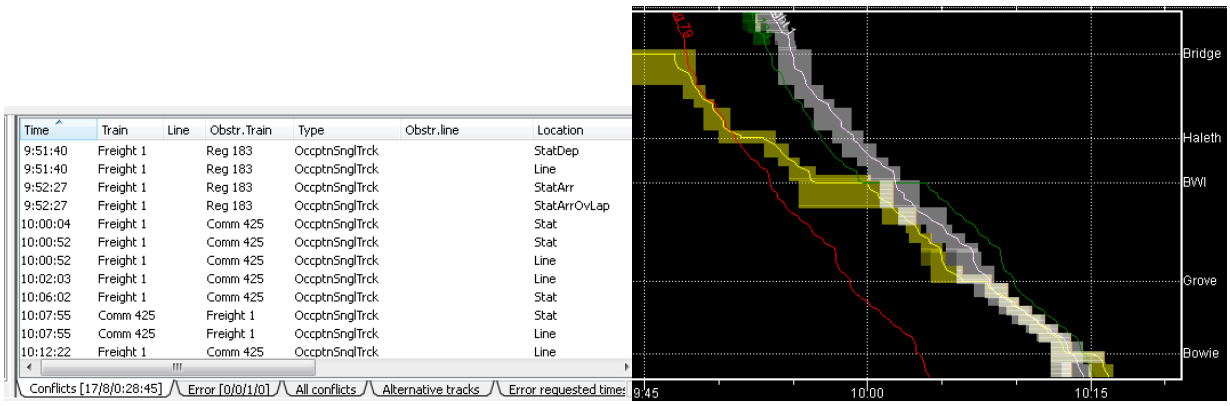

FIGURE 5- RailSys train conflict management tool output (graphical and tabular formats)

The software allows user to resolve the conflicts by adjusting the departure/arrival times, by rerouting the trains, and/or by considering any conditional stop in the sidings to provide any meet-pass opportunity. As depicted in FIGURE 6, the freight train was successfully dispatched in RailSys by adjusting the departure time of freight train to 10 am instead of 9:50 am, and by rerouting some of the other trains via crossovers along different segments of the corridor.

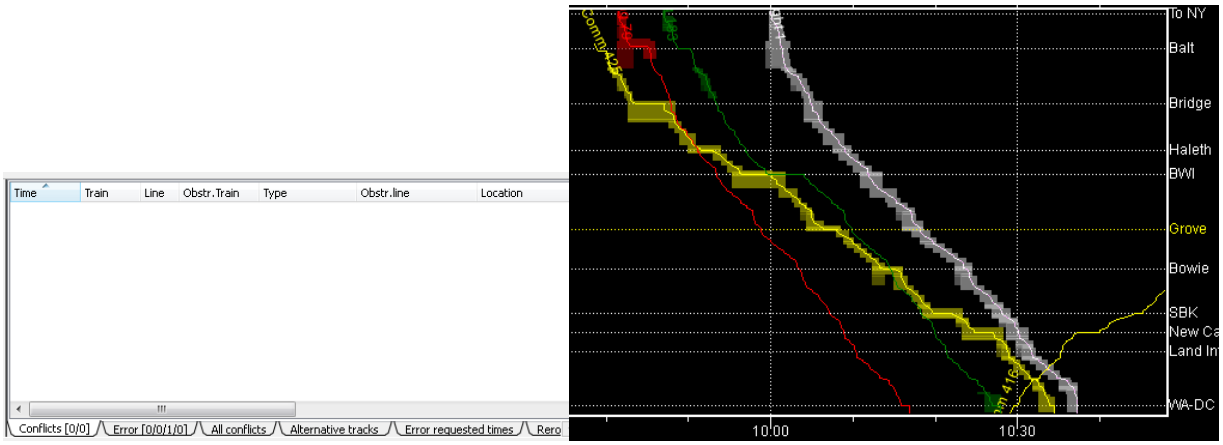

FIGURE 6- The resolution in RailSys by adjusting the departure time of freight train to 10 am and rerouting other trains in some segments of corridor

Scenario 2: Timetable Compression Technique

As discussed before, one of the techniques of improving capacity utilization and level of service used in Europe is timetable compression. RailSys uses a compression technique (UIC 406) to optimize a feasible timetable and to improve the capacity utilization levels. There are several factors which should be defined prior to the capacity optimization (timetable compression), such as:

- Overtaking option in the sidings
- Maximum dwell time of trains in the sidings
- Using initial timetable as input data
- The compression technique (Austrian method, OBB, or German method, DB)
- Timetable duration (the portion of timetable which is planned to be optimized)
- Directional or bidirectional operations
- The route option (The tracks or platforms numbers which are going to be used in the analysis)

In this research, we applied the compression procedure of RailSys to the current timetable and considered overtaking option in maximum two stations based on OBB compression algorithm. DB algorithm wasn't used in this study, as it considers one of the trains as a "dummy" train for the purpose of the compression technique, causing the number of simulated trains to deviate from RTC results. Other major differences between OBB and DB methods are related to the way occupation time of trains along the corridor is calculated, as well as the criteria and steps of compressing the first and last trains of the service within the compression period. FIGURE 7 presents the final results of compressed timetable by using UIC 406 compression approach. Railsys organized routing of train operations in directional manner with southbound trains using the first track and northbound trains the second track and used maximum of two minutes dwell time in sidings/yards. After timetable compression, the homogeneity indicator of operations (an index showing the similarity between trains speed and characteristics) was approximately 97.4% and 97.8%, respectively for southbound and northbound directions. This reveals that the trains operating in the study scenario had high level of homogeneity, making their operational characteristics consistent with each other and easier to reach higher levels of capacity utilization (the percentage of capacity consumption out of available capacity for each line). The utilization after compression was estimated as 13.2% and 12.5% for respective directions, which is fairly low for homogeneous train operations. However, these values should be used cautiously, as they may change significantly, once all 141 trains are considered in the next phase of analysis.

Railsys provides the compressed timetables (FIGURE 7) separately for each track/route and direction of operations, since European operations of multi-track corridors are typically directionally oriented. It is not possible to automatically combine both compressed timetables in a single stringline diagram in RailSys, except for single track operations.

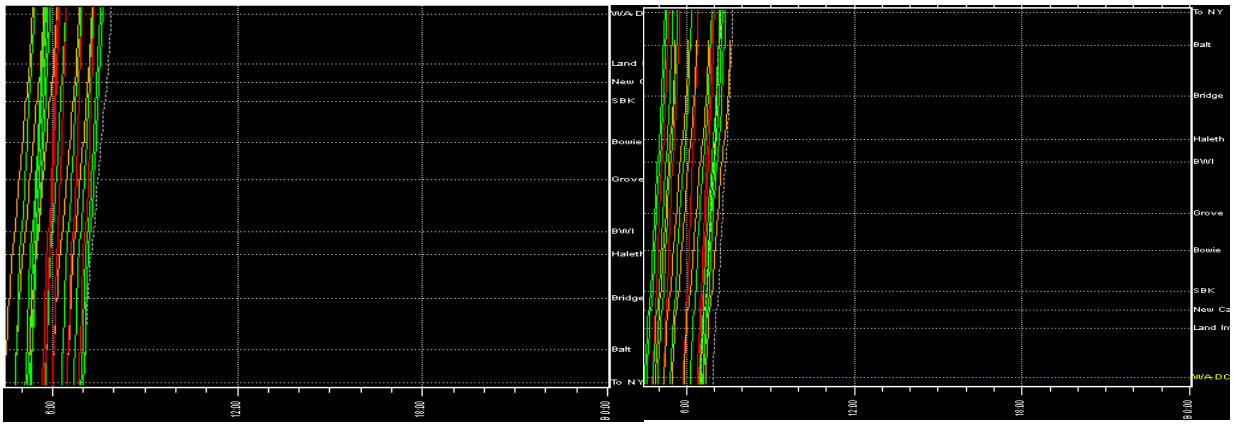

FIGURE 7- Compressed stringline of trains in both directions, 4 am -12 am (Left: Southbound, Right: Northbound)

In addition to directional considerations, several other observations were made during the application of compression technique:

- The order of trains in the optimized timetables of RailSys was exactly the same as defined in the input timetable, but the optimized arrival/departure times were different. It was not clear whether RailSys optimization technique used the preferred departure times from input timetable.
- The maximum dwell time at stations considered by RailSys was the same for all trains and at all stations, while it might be variable in real practices. Consideration of an individual dwell time for each train or each station might improve the outcomes of timetable compression technique.
- In addition to compressing the existing timetable, new trains that possess the same or different operational characteristics (speed, stop patterns, type of trains, etc.) can be introduced in between the existing trains. FIGURE 8 shows new trains that could run along southbound direction of the case study, considering the existing train schedules. According to RailSys, there is, theoretically, an option of running 353 new trains during the 19.5 hours of operations until 96.5% of capacity utilization indicator (traffic saturation factor) is reached.

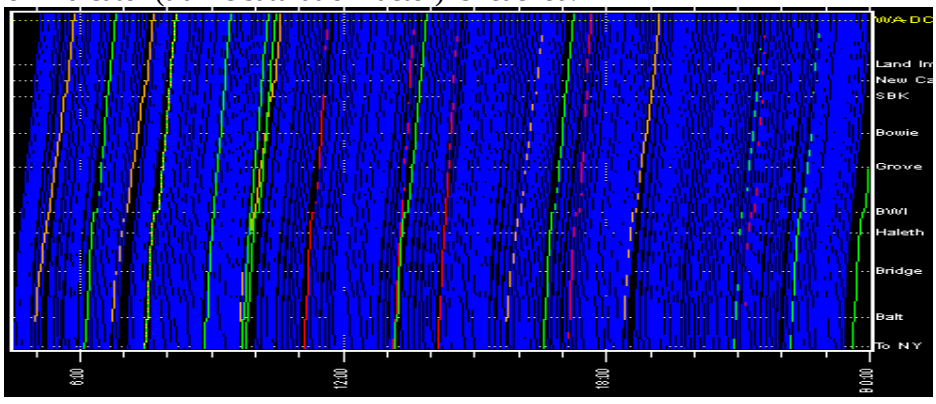

FIGURE 8- 353 new trains (shown in blue lines) literally inserted within the existing schedule of trains along southbound direction of case study

CONCLUSIONS AND NEXT STEPS OF RESEARCH

This paper introduced two commercial railway simulation tools available in the market for evaluating the capacity levels and train operations. Rail Traffic Controller (RTC) is non-timetable based simulation software, typically used predominantly for improvised operation philosophy conditions (the dominant operations approach in the U.S. rail environment). On the other hand, RailSys, is a timetable-based simulation software commonly used in Europe which includes infrastructure data management, timetable construction/slot management, track possession planning, and simulation features.

To compare the similarities and differences of RTC and RailSys software, a short segment of the Northeast Corridor (NEC) between Baltimore and Washington, DC was selected as a shared-use corridor case study and applied in both simulation packages. The comparison of the simulation procedure and outcomes led to the following observations and conclusions:

1. Both RTC and RailSys software are powerful tools for operations simulation, but the procedure and steps of developing the operations rules and dispatching system for improvised operation philosophy with no predefined schedule (preferred departure times only as input) is easier to implement in RTC. RTC can dispatch a predefined schedule of trains, but specific timetable management should be conducted manually by the user, as necessary.
2. RailSys requires more steps and details when developing the network and original timetables, but also possesses more versatile features and tools for identifying train conflicts and rerouting trains when considering new trains or improving existing timetable. RTC suggests reroutes as a function of its dispatching capability, if tracks are not assigned or if alternate nodes are allowed. In RailSys, rerouting should be set up by the user, based on the assistance provided by the timetable and network graphical and tabular features.
3. Solutions to train conflicts in RTC are automatically suggested and tested during the RTC simulation. They can then be manually hardcoded into the schedule and used iteratively in new simulation runs, until the schedule is optimized. The train conflicts in Railsys must be manually resolved, but there are several features and graphical and tabular tools provided by the Railsys to assist the user in gradually resolving the conflict.
4. Since RailSys is originally developed in Europe, the procedure of developing North American rolling stock and signaling features is relatively challenging in RailSys, as default database and information use European characteristics rather than North American ones.
5. Several factors should be defined in the capacity optimization tools of RailSys but overtaking scenario, the selected route, directional and bidirectional operations, the amount of dwell time and the algorithm used for timetable compression (OBB vs. DB pattern) seem to impact the final results of the optimized timetable the most.
6. RailSys timetable compression technique maintains the order of trains through the optimized timetable option (as defined in the input timetable), but it doesn't keep the preferred departure times. RailSys can also impose new trains within the current trains schedule, but it only considers one direction of operations, instead of both directions.
7. The timetable compression technique of RailSys may not be an ideal solution for double and multi-track operations in the U.S., as the outcome of compressed timetables in both directions can't be automatically combined to a single diagram. The separate presentation of the compressed timetable is especially challenging at station exit and entrance sections if there is an option of using crossovers or bi-directional operations.

The next step of the research is to evaluate the use of timetable management modeling approaches, such as timetable compression techniques to improve the train timetable, capacity utilization, of a given case study in the U.S. shared-use corridor. The main objective of next step of research will be to identify the key modeling parameters for operational management techniques and how they can be

implemented using current simulation tools and features. It will also expand the use of hybrid approach by returning the compressed timetable to RTC for validation process.

ACKNOWLEDGMENTS

The authors are grateful for the assistance provided by Amtrak, including respective information and database of NEC provided for this academic research. Special thanks for the assistance and feedback received from Davis Dure. The authors would also like to thank Eric Wilson (Berkeley Simulation Software, LLC), Sonja Perkuhn (RMCon), and Gabriele Löber (RMCon) for providing academic licenses to simulation packages, RTC and Railsys respectively, and for their great support for the research team at Michigan Tech.

REFERENCES

1. Pouryousef, Hamed, Pasi Lautala, and Thomas White. *Review of Capacity Measurement Methodologies; Similarities and Differences in the U.S. and European Railroads*. in *2013 TRB Annual Meeting* 2013. Washington-DC.
2. Pachl, Joern, *Railway Operation and Control*2002, Mountlake Terrace-WA VTD Rail Publishing-USA.
3. M. Abril, F. Barber, L. Ingolotti, M.A. Salido, P. Tormos, A. Lova *An Assessment of Railway Capacity*, 2007, Technical University of Valencia: Valencia, SPAIN.
4. Pavankumar Murali, Maged M. Dessouky, Fernando Ordonez, and Kurt Palmer, *A Delay Estimation Technique for Single and Double-track Railroads*, 2009, Dept. of Industrial & Systems Engineering, University of Southern California: Los Angeles, CA.
5. Melody Khadem Sameni, Mark Dingler, John M. Preston, Christopher P.L. Barkan. *Profit-Generating Capacity for a Freight Railroad*. in *TRB 90th Annual Meeting*. 2011. Washington, DC: TRB.
6. Melody Khadem Sameni, Mark Dingler, John M. Preston. *REVISING THE UIC 406 METHOD: REVENUE GENERATING CAPACITY- JRC2010-36281*. in *Joint Rail Conference (JRC) 2010*. 2010. Urbana, IL, USA.
7. White, Thomas, *Alternatives for Railroad Traffic Simulation Analysis.* Journal of the Transportation Research Board (TRB)- Washington, DC, 2005. **1916** p. 34-41.
8. Pouryousef, H and P. Lautala. *Evaluating the Results and Features of Two Capacity Simulation Tools on the Shared-use Corridors*. in *ASME/ASCE/IEEE 2013 Joint Rail Conference (JRC2013)*. 2013. Knoxville,TX, USA.
9. RTC, *RTC Online Help*, B.S. Software, Editor 2011: USA.
10. RMCon, *RailSys 8 Classic, RailSys 8 Enterprise, Network-Wide Timetable and Infrastructure Management*, R.M.C. GmbH, Editor 2012: Hannover, Germany.
11. RMCon, *RailSys User Manual*, R.M.C. GmbH, Editor 2010: Hannover, Germany.

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

JRC2013-2441

Evaluating the Results and Features of Two Capacity Simulation Tools on the Shared-use Corridors

Authors:

Pasi Lautala, PhD, P.E.
Assistant Professor
Director, Rail Transportation Program
Civil and Environmental Engineering Department,
Michigan Tech. University
1400 Townsend Drive, Houghton, MI 49931
phone: (906)-487-3547
ptlautal@mtu.edu

Hamed Pouryousef
PhD Candidate at Rail Transportation Program
Civil and Environmental Engineering Department,
Michigan Tech. University
1400 Townsend Drive, Houghton, MI 49931
phone: (906)-231-2320
hpouryou@mtu.edu

ABSTRACT

The majority of passenger rail services in the United States (U.S.) operate on the shared-use corridors with freight rail. These types of operations tend to be challenging due to high heterogeneity, particularly in terms of reliability of service and capacity availability. The projected growth in demand for rail transportation is likely to exacerbate the situation. Similar to the U.S., the European passenger rail services are generally operated on shared-use corridors, but the infrastructure conditions and the operational priorities and patterns typically allow more reliable and higher speed passenger operations in comparison to the U.S. trains.

Both continents use capacity and simulation software to analyze capacity allocations and operational limitations. However, the effects of the software selection haven't been investigated. This research reviews two common simulation tools developed in the U.S. and Europe, Rail Traffic Controller (RTC) and Railsys, respectively. The paper reviews the structure and the main components of these two simulation tools. It will also present the outcomes of running RTC and Railsys through a given segment of shared-use track based on the same rolling stock, operation and signaling characteristics and analyze the similarities and differences between the outcomes of RTC and Railsys.

Key Words: *Rail Capacity, Rail Simulation, Shared-use corridors*

1- INTRODUCTION

Railroad capacity is a complicated concept of rail transportation planning and operations with several factors that should be considered through the analysis. Typically, the capacity of rail line is defined as the number of trains that can safely pass along a given segment through a period of time based on particular level of service (LOS). The capacity is affected by different system configurations, such as: Track infrastructure, Signaling system, Operations philosophy, and Rolling stock.

The differences in railway operations between Europe and North America may lead to use of different methodologies, techniques, and tools to evaluate and measure the capacity levels. In Europe, there are high utilization corridors where intercity passenger, commuter, freight, and even high speed services operate on the shared tracks with highly structured timetables and schedules. In the U.S., the improvised operational philosophy is the dominant pattern on current shared

corridors where almost all freight and some passenger trains are assigned their slots in the network on a daily basis (although passenger trains do have their regular schedules). [1] As the U.S. continues to develop higher speed passenger service with similar characteristics to those in European shared-use lines, the accuracy of capacity analysis methods becomes more important. In addition, the increases in volumes and operational heterogeneity can be expected to add pressure for higher capacity utilization of the U.S. shared-use corridors. A comparison of two major railway simulation tools in the U.S. and Europe can help researchers to evaluate the structure, components and different features of these tools. Besides that, it can highlight the main challenges and similarities or differences between the capacity analysis outcomes of these tools based on similar scenarios and practices.

This paper focuses on two major simulation tools in the U.S. and Europe, RTC and Railsys, respectively, to address the above mentioned goals and concerns. In the first part of the paper, different tools and methodologies of capacity analysis will be briefly reviewed in both European and the U.S. rail environments. Second, the main features and components of RTC and Railsys will be explained regarding the capacity analysis, followed by a general case study of capacity analysis conducted in both RTC and Railsys. The outcomes of the case study will be reviewed and any particular similarities and differences between these two software's results are highlighted. Finally, the conclusion and next steps of the research will be briefly summarized in the last part of the paper.

2- CAPACITY TOOLS AND METHODOLOGIES

2-1- General Discussion

The U.S. and European rail networks have several similarities, such as using modern signaling and traffic control systems (e.g., developing ETCS in Europe and PTC in the U.S.), while significant differences also exist. These characteristic differences between these two rail environments may change the preferred methodologies and the outcomes of capacity analysis. Figure 1 briefly highlights some of the key differences between infrastructure, signaling, operations and rolling stock in Europe and the U.S. which may affect the capacity utilization.

The ownership of rail infrastructure is one of the main differences between Europe and the U.S. rail networks which affect the capacity tools and methodologies. More than 90% of the infrastructure is owned and managed by private freight railroads in the U.S.; while almost all infrastructures is owned and managed by governments or public agencies in Europe. In addition, operations and infrastructure are vertically separated in Europe; while the majority of operations in the U.S. (mainly freight) are handled by the same corporations who own the infrastructure.

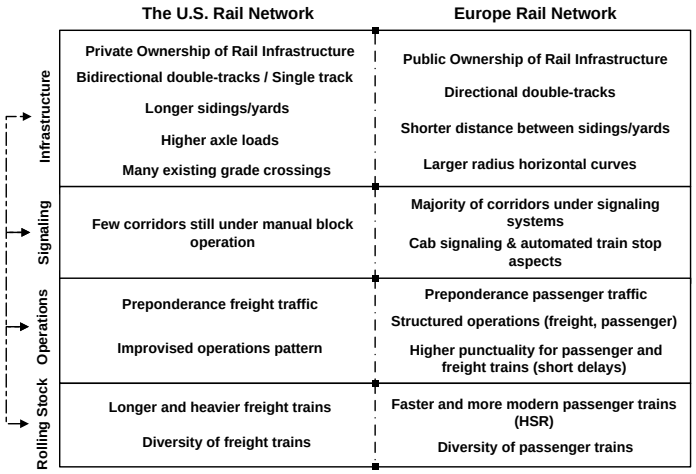

Figure 1- The main differences in the U.S. and Europe rail systems which may affect the capacity [1]

The operations philosophy is another major characteristic which affect the tools and methodologies for analyzing and measuring the capacity levels. The U.S. operations philosophy follows improvised pattern for almost all freight trains except some intermodal trains. It means there is no repeatable dispatching plan for these trains in-advance. Indeed, the improvisation is accomplished by railway dispatchers who resolve conflicts between trains as they occur, while in Europe these train interactions are pre-planned by the timetable. On the passenger side, many Amtrak and commuter train daily operation patterns are also planned without any particular details, anticipating improvised resolution of conflict among the passenger trains, or between passenger and freight trains. In Europe, almost all freight and passenger trains have a regular schedule developed well in advance, known as structured operations. Indeed, one of the reasons for philosophical differences is the higher variability between freight trains in the U.S., while in Europe passenger train configurations are more diverse [2]

The ownership status of railway network and operation philosophy (improvised vs. structured operations) has wide impact on the capacity tools and methodologies in Europe and the U.S. In addition, differences in rail network characteristics may lead to dissimilar techniques and methodology of capacity analysis. More details of difference between the U.S. and European rail configuration is explained in a TRB paper by Pouryousef and Lautala, 2013. [1]

2-2- Analytical vs. Simulation Approach

The techniques and approaches for measuring and analyzing the capacity utilization have different names,

but the literature commonly divides them to two main categories, analytical and simulation approaches. For instance, Abril, et al., classified the capacity methodologies to analytical methods, optimization methods, and simulation methods. [3] Joern Pachl divided the capacity methodologies to two major classes: analytic and simulation methods. [4] Similar categorization was used on research conducted by Murali on delay estimation technique [5] and a research conducted at the University of Illinois Urbana Champaign by Sogin, Barkan, et al., who classified capacity methods as theoretical (analytical), parametric, and simulation methods. [6, 7]

2-2-1- Analytical Approach

The analytical approach typically uses several steps of data processing through mathematical equations or algebraic expressions to determine a solution for the problem (theoretical capacity). [3] The outcomes vary based on the level of complexity of the scenario and may be as simple as number of trains per day, or include a combination of several performance indicators, such as timetable, track occupancy chart, fuel consumption, speed diagrams, etc. Analytical methods can be conducted with or without specific software.

Figure *2* presents the different methodologies that can be used in the analytical approach of capacity evaluation and how complexity, such as optimization and timetable compression methods, can be added to provide more detailed results.

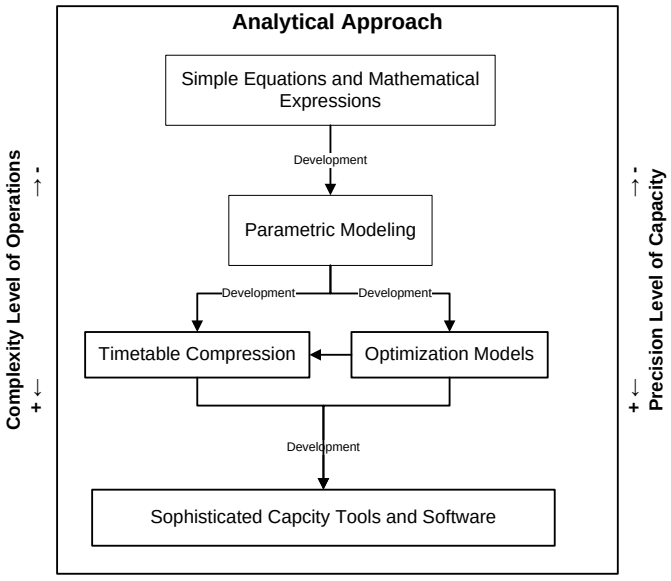

Figure 2- The relationship between operations complexity, precision level of capacity and different methodologies of analytical approach[1]

2-2-2- Simulation Approach

Simulation is an emulation of a system's operation which should be as close as possible to its real-world equivalent. [3] The process of simulation is repeated several times until an acceptable result is achieved by the software (Heuristic approach). There are several simulation tools that can provide different perspectives of simulated results. They can be classified as either general simulation tools, such as AweSim, Minitab, and Arena [5, 8]; or commercial simulation software specifically designed for rail transportation, such as RTC, MultiRail, RAILSIM, OpenTrack, RailSys, and CMS [3, 9]. General simulation software is typically used for limited purposes, such as estimating the train delay, or measuring the level of service or reliability of current or future train services. A commercial rail simulation software include customized tools to process several activities related to the rail transportation including estimating travel time, fuel consumption, train-paths, train speed, time-distance diagram, capacity analysis, etc. The rail simulation software typically needs more detailed database of information than the general simulation tools, but on the other hand, it can provide more versatile outcomes and analysis.

The commercial railroad simulation software is typically developed based on two major components; 1) Train movement simulation, and 2) Train dispatching simulation. The first component calculates the train speed along the track by using the train resistance formula (like Davis equation) and train traction power. The dispatching simulation component typically emulates (or attempts to emulate) the action of the dispatcher in improvising traffic management, but in some cases, it can be used as part of a traffic management software to help traffic dispatchers to manage and organize the daily trains' schedules. [2]

The software is typically divided to two major groups; 1) non-timetable based vs. 2) timetable based. The non-timetable based simulations are typically applied in railways that operate based on the improvised operation pattern without initial timetable, such as the majority of the U.S. rail network. In this type of simulation, after loading the input data in the software, the train dispatching simulation process improvises the departure times from the initial station that are included in the input data. The software may encounter a problem to assign all trains and request assistance from the software user to resolve the issue by adjusting the train data, or by modifying the schedule constraints. [2, 9] The rail traffic controller (RTC), developed by Berkeley Simulation Software is the most common software in this category, used extensively by the U.S. rail industry.

The simulation procedure in the timetable based software (typically used in Europe) is based on the initial timetable of trains and uses software tools to improve the timetable as much as possible. The

simulation process in this methodology begins with creating a timetable for each particular train. In the case of schedule conflict between trains, the user must change the timetable until the feasible schedule is achieved; however, the user interference is not arbitrary as in the improvised method, but it is implemented as part of the simulation process. [2] RailSys developed by Rail Management Consultants GmbH in Germany, and OpenTrack developed by OpenTrack Railway Technology Ltd. in Switzerland, are two common examples of several timetable-based simulation software in Europe.

The actual simulation steps of capacity analysis for a given project are summarized in Figure 3. If the capacity outcomes of the simulation are not satisfactory for the capacity planners, new scenarios of simulation will be developed, typically by adjusting the operation rules, signaling components, train configurations or by upgrading/rearranging track components including main track, sidings and crossovers.

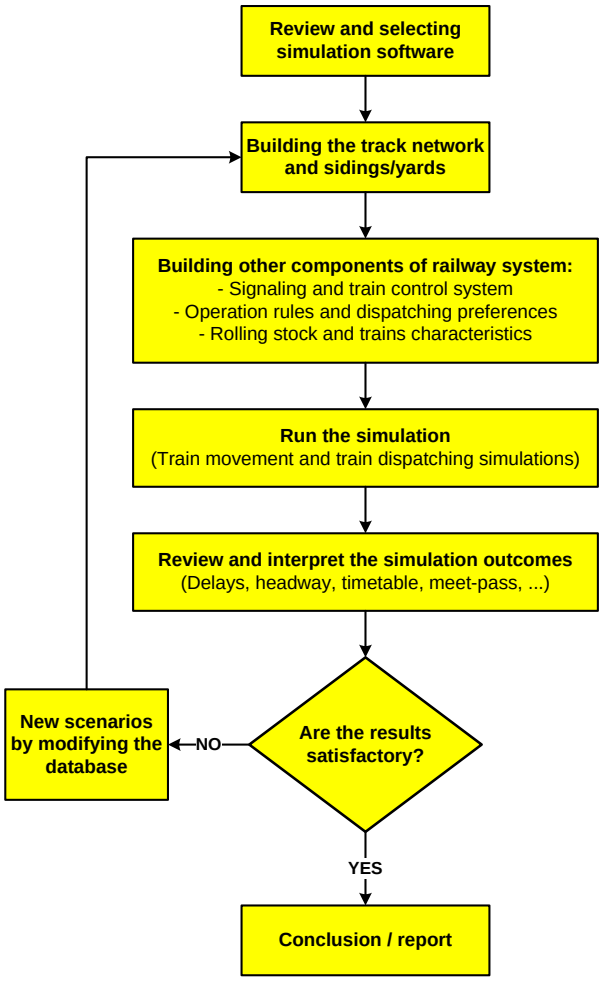

Figure 3- The basic steps of simulation procedure for the capacity analysis

2-2-3-Combined Analytical-Simulation Method

In addition to the analytical and simulation approaches, combined analytical-simulation method can also be applied to investigate the rail capacity. Although the term "combined methodology" is not commonly used in the literature, many studies have taken advantage of both analytical and simulation methods. [1] While updating the capacity factors and criteria tends to be easier and the process of running the new scenarios is generally faster in simulation approach, parametric and heuristic modeling is more flexible in analytical approach in terms of creating new aspects and rules through the analysis. The simulation approach may place some limitations when adjusting the characteristics of signaling or operation rules. Thus, combined simulation-analytical methodology can take advantage of both methodologies' techniques and benefits.

In the combined approach, simulation tools are used to evaluate and understand the capacity bottlenecks through the corridor, and analytical methods are developed to improve the capacity utilization levels. The process of applying simulation-analytical practices may be repeated by the research team until an acceptable set of outputs and alternatives is found. (Figure 4)

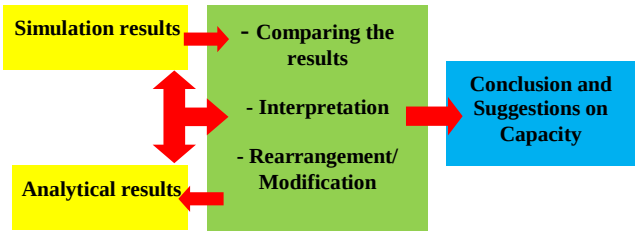

Figure 4 - Basic diagram of combined analytical-simulation method to evaluate capacity [1]

3- FEATURES AND COMPONENTS OF RTC AND RAILSYS

Rail Traffic Controller (RTC) and RailSys are two well-established commercial software reviewed as part of the research. Table 1 provides a comparison of some of the features and characteristics of RTC and RailSys.

Table 1- Comparison between RTC and RailSys

Criteria	RTC	RailSys
Developer	Berkeley Simulation Software, LLC, (USA)	Rail Management Consultants GmbH (RMCon) (Germany)
Features and Modules	- Animation of traffic flow - Time-distance diagrams (stringlines) - TPC profiles - Track occupancy chart - Detailed train status - Timetables at various levels of detail - Operating statistics at the individual train level or summarized by train type or at a system-wide level - Graphical network interface	-Infrastructure manager -Timetable construction -Capacity Management (UIC code 406) - Track Possession planning -Simulation Manager -Rolling stock circulation planning - Graphical Timetable -Platform and track occupation diagrams - Graphical network interface -Delay statistics
Simulation Category	Non-timetable based simulation software	Timetable based simulation software (UIC code 406)
Capacity Metrics	Delay statistics, Track occupancy chart, time-distance diagrams	Delay statistics, infrastructure occupation time, optimized timetable
Major Users	Class 1 RRs: (UPRR, BNSF, CSX, NS, KCS, CN, CP, Amtrak), U.S. railway consultants, urban rail transit agencies	Many European rail operators and consultants, international rail companies

RTC was launched in the North America's rail market in 2001 and since then it has been continuously developed and upgraded for variety of simulation practices particularly for improvised operation philosophy (the dominant operations approach in the U.S. rail environment). RTC can be categorized as non-timetable based simulation software, although it can receive the preferred arrival and departure times of different trains for the simulation process. The dispatching simulation component of RTC is based on a decision support core, called "meet-pass N-train logic". For any dispatching simulation practice, "meet-pass N-train logic" will eventually decide when the given trains should exactly arrive and depart from different sidings. This procedure is based on the defined train priorities and preferred times of departure, even though the simulated departure times may not match the preferred times initially defined in the simulation process. If the simulation process can't be accomplished due to the train conflicts, the user should either change the preferred times of the departure/arrivals, or change the priorities of trains to allow more delays for particular trains, or to provide the meet-pass opportunity for other trains.

The decisions of RTC are based on the minimum cost of train operations, as defined by user, and it contains only direct train costs. Train conflicts are typically resolved by giving preference to the trains with higher operation cost. The crew requirements and their operational hour limits can significantly change the arrival/departure times of trains. For instance, the trains with crews that are approaching their maximum service hours are preferred to be departed sooner than the other trains running behind their schedule considering the delay costs factor. On the other hand, trains which are ahead of their schedule may be treated with less preference, even if they had initially higher priorities. In summary, RTC runs the simulation either to minimize the trains' delay or to minimize the direct operating costs of trains. [10]

RailSys is an operation management software package tool that includes infrastructure data management, timetable construction/slot management, track possession planning, and simulation features. It has been in the market since 2000. The capacity feature of RailSys uses the UIC code 406 approach which is based on the timetable compression technique. Given train timetables, a segment of the route is selected to automatically compress the utilized train-paths, while considering the minimum headways and acceptable buffer times between the trains. The compression technique always begins at the start of calculation period and it stops after the calculation period is fully occupied by the last possible train. The remaining usable level of capacity is identified by the number of new train-paths available, until the given time period is saturated by the train-paths and buffer times. For the future trains which may have no particular timetables, Railsys can also evaluate the future network capacity utilization, using a RailSys model called "NEMO" (Network Evaluation Model). This model can be applied based on the infrastructure information and train origin-destination matrix. [11, 12]

4- CASE STUDY: AN IMAGINATIVE SEGMENT OF SHARED-USE CORRIDOR

For testing purposes, a single case study was conducted as part of the research to evaluate the outcomes, similarities, difficulties and challenges of each software. The case study was developed on an existing rail line in the U.S., but the train and signaling parameters were developed by the researchers for this specific test. The four database categories commonly

required as input by simulation software (track and infrastructure, signaling and train control system, rolling stock, and operation rules) were developed for the case study and applied in each software. More details of each part of case study are explained in the following sections.

4-1- Infrastructure Characteristics of Case Study

The case study's infrastructure contains a 30 mile long single track segment and three sidings/yards along the route for any meet/pass and stop purposes. Horizontal curves were not considered along the line to simplify the infrastructure, but the vertical track profile and locations of the sidings were precisely derived from an existing rail line mainly used for excursion passenger trains. It should be pointed out that the horizontal curve impact on the train speed is not typically as significant as the grade's impact on the speed profile, especially for speeds under 50 mph. Table 2 summarized the infrastructure parameters for the case study.

Table 2- Details of case study infrastructure

Segment Length	30 miles, single track
Sidings/yards	2 sidings + 1 yard
Max. grade	1.78%
Curvature	No curves
Length of siding track	0.34 - 0.42 miles
Turnout #	# 11

4-2- Signaling Characteristics of Case Study

The case study's signaling system uses automatic permissive block (APB) for single track operation under CTC control system with four-aspect signaling along the main blocks. The length of blocks varies between 1.2 and 2.5 miles and all sidings/ yard tracks are equipped with CTC controlled interlocking systems.

4-3- Rolling Stock Characteristics of Case Study

Four types of trains have been considered in the case study; intercity passenger, commuter passenger, merchandise freight and intermodal freight trains. It is assumed that the characteristic and configuration of each train in a specific category is uniform and each train is operated in both westbound and eastbound directions. All passenger and commuter trains are propelled by a single diesel-electric locomotive and all freight trains are loaded in both directions. Since the type and configuration of locomotives are relatively different in the RTC and Railsys database, some of the main characteristics of selected locomotives in RTC (such as power, weight, length, axle load, acceleration/ deceleration rate, resistance) are imposed and adjusted in the Railsys database as new type of locomotive. Some of the main characteristics of used rolling stock are explained in Table 3.

Table 3- Main features of case study's trains

Train	Daily trains (pairs)	# of cars	Trailing weight (ton)	Trailing length (feet)
Intermodal	3	35	1850	2100
Merchandise	2	19	1330	1140
Passenger	4	10	420	700
Commuter	2	4	170	280

4-4- Operation Rules of Case Study

There are several relevant operation rules for simulation, such as the train's priority, speed limits, stop patterns, and preferred time and order of train departures. The priority of different types of trains in diminishing order was commuter trains, passenger trains, intermodal, and merchandise trains. In the case study, the speed of passenger/commuter trains is limited to 60 mph, while freight trains are limited to 50 mph. In addition, the initial speed of all trains is 30 mph when they reach the track segment that starts the simulation process. There are no planned stops for any trains, but passenger, commuter or merchandise trains may have to stop at the sidings due to the meet-pass logic. The intermodal freight trains may have meet-pass stop only in the yard tracks since the length of this type of trains is longer than the siding lengths. In the case study, there were no predefined arrival/departure timetables, but some preferred departure times were considered, as explained in the next section.

5- SIMULATION RESULTS AND ANALYSIS

This section will summarize and compare the case study outcomes of RTC and Railsys with each other. Since there was no initial timetable for the case study, the preferred departure times were inserted into RTC and then compared with the simulated train timetable results provided by both RTC and RailSys software.

5-1- Outcomes of RTC Simulation

The case study simulation results obtained from RTC are shown in Figure 5 in distance-time diagram format (timetable string-line). As noted earlier, there were no planned stops for the trains, but several stops were required for meet-pass in the sidings. The simulated arrival/departure times have also some deviation in comparison to the initial preferred times of train dispatching schedule, due to simulation software solving conflicts between trains included in the initial plan.

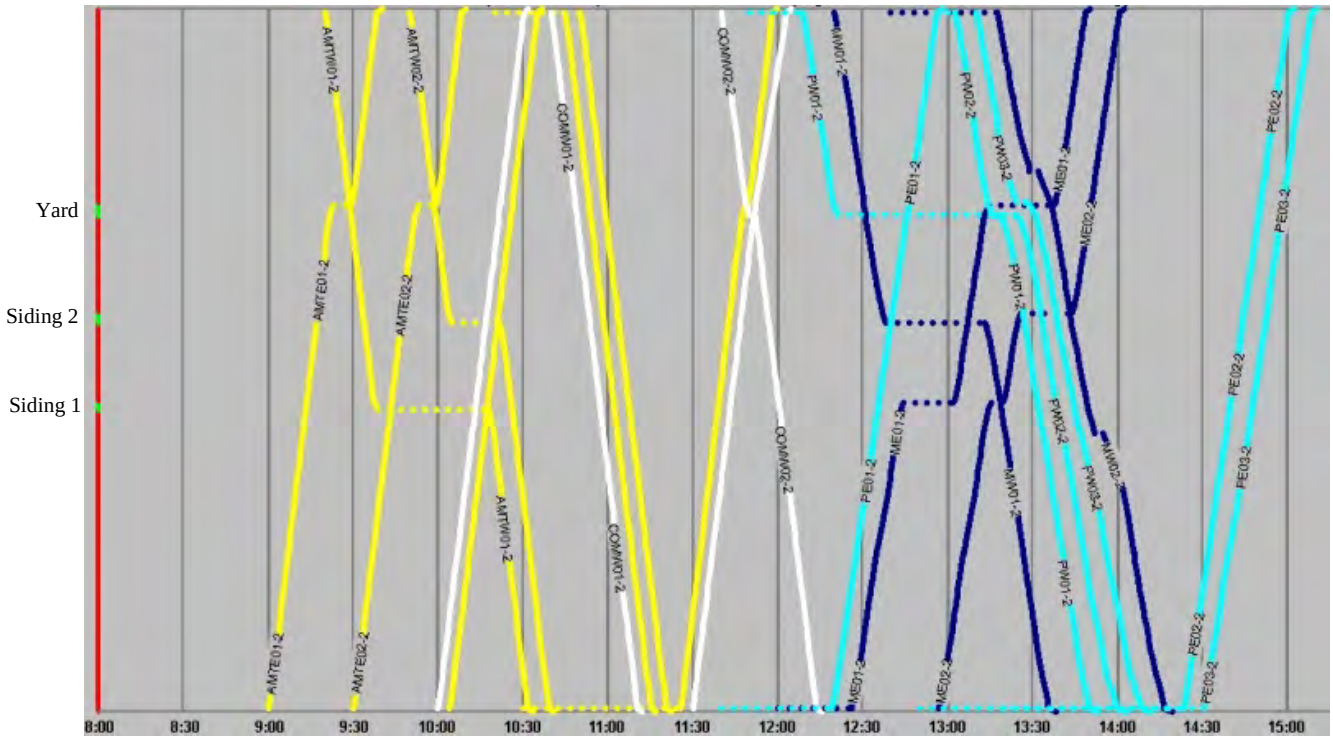

Figure 5- Simulated timetable string-line in RTC (Commuter: White, Passenger: Yellow, Intermodal: Blue, freight: Navy blue)

The initial preferred train departure times did not consider all factors of scheduling, so some times were adjusted by the RTC. (Table 4) As presented in Table 4, trains with higher priority (commuter and passenger) had lower deviation between their requested and simulated departure times. The departures with deviated time have been highlighted in table cells by yellow (eastbound) and green (westbound) colors. There was also conflict between requested departure time of passenger 3 and commuter 1 (eastbound direction), as both trains were requested to depart at 10:00 which is practically impossible. RTC solved the time conflict by maintaining the initial schedule of commuter train (with higher priority) and delaying passenger train for 3 minutes at the entry point of the line. Similar situation occurred between Intermodal 2 and Freight 2 (both were planned to depart at 12:50). RTC changed departure times of both trains to facilitate essential meet-pass events. As presented in Figure 5, commuter trains stopped in a siding only once and for a short time due to the meet-pass enforcement, while passenger trains had more and longer delays in the sidings and in the entry points of the line, since their priority was lower than commuter trains. The same trend can also be noticed for freight and intermodal train schedules with even more delays and longer meet-pass time in the sidings, since the priority of these two types of trains was lower than passenger and commuter trains. However, the merchandise freight trains were simulated with lower delays in comparison to the intermodal trains, although the priority of intermodal trains was slightly higher than merchandise train. The reason of such higher delay for intermodal train may be hidden in the fact that merchandise trains had more flexibility for meet-pass stops in any siding/yard, while intermodal trains could only stop in the yard. The summary of trains' performance (average speed) and delays extracted from RTC simulation are represented in Table 5.

Table 4- Comparison between planned and simulated departure times in RTC

Train	**Planned departure-Eastbound**	**Simulated departure-Eastbound**	**Planned departure-Westbound**	**Simulated departure-Westbound**
Pass1	9:00	9:00	9:20	9:20
Pass2	9:30	9:30	9:50	9:50
Pass3	***10:00***	10:03	10:20	10:45
Pass4	10:30	11:27	10:50	10:50
Comm1	***10:00***	10:00	10:40	10:40
Comm2	11:30	11:30	11:40	11:40
Interm1	11:40	12:20	11:50	12:08
Interm2	**12:50**	14:23	13:00	13:02
Interm3	13:20	14:30	13:10	13:10
Freight1	12:00	12:25	12:20	12:20
Freight2	**12:50**	12:55	12:40	13:15

Table 5- Summary of trains' performance in RTC

Train	Ave speed (mph)	Train-Miles	Delay (minutes per 100 trains-miles)
Passenger-commuter	40.4	372.8	15.02
Intermodal	19.7	186.4	57.71
Freight	20.7	124.3	54.06

5-2- Outcomes of RailSys Simulation

Since Railsys is timetable-based software, it needs a feasible and reasonable timetable for further steps of capacity analysis, although it can also provide an optimum timetable based on just number of trains and infrastructure characteristics. In our case study, the preferred departure time of trains had many conflicts and thus could not be used as a feasible timetable. To make the results of RailSys more comparable with RTC's results, the output of RTC simulation was inserted in the Railsys as input timetable. As presented in Figure 6 there were some minor deviations and differences between arrival/departure times in RaiSys and RTC. These deviations occurred due to the fact that simulated running time of trains along the corridor had slight variations between software, since there were minor differences between rolling stocks and signaling features/equations (such as tractive effort of engines, acceleration, deceleration, braking diagram, etc.) in Railsys vs. RTC. However, almost 95% of departure/arrival times of input timetable in RailSys were exactly the same as in the output obtained from RTC.

One of the main advantages of timetable-based simulation software, such as RailSys, is the possibility of improving and optimizing the initial timetable based on predefined patterns and algorithms. RailSys takes advantage of UIC 406 compression technique to optimize a feasible timetable and improve the capacity utilization levels. There are several factors which should be defined prior to the capacity optimization (timetable compression), such as:

- Overtaking option in the sidings
- Maximum dwell time of trains in the sidings
- Using initial timetable as input data
- The compression technique (Austrian method, OBB, or German method, DB)
- Timetable duration (the portion of timetable which is planned to be optimized)

In this research, we evaluated the initial timetable (developed based on RTC simulation output) through

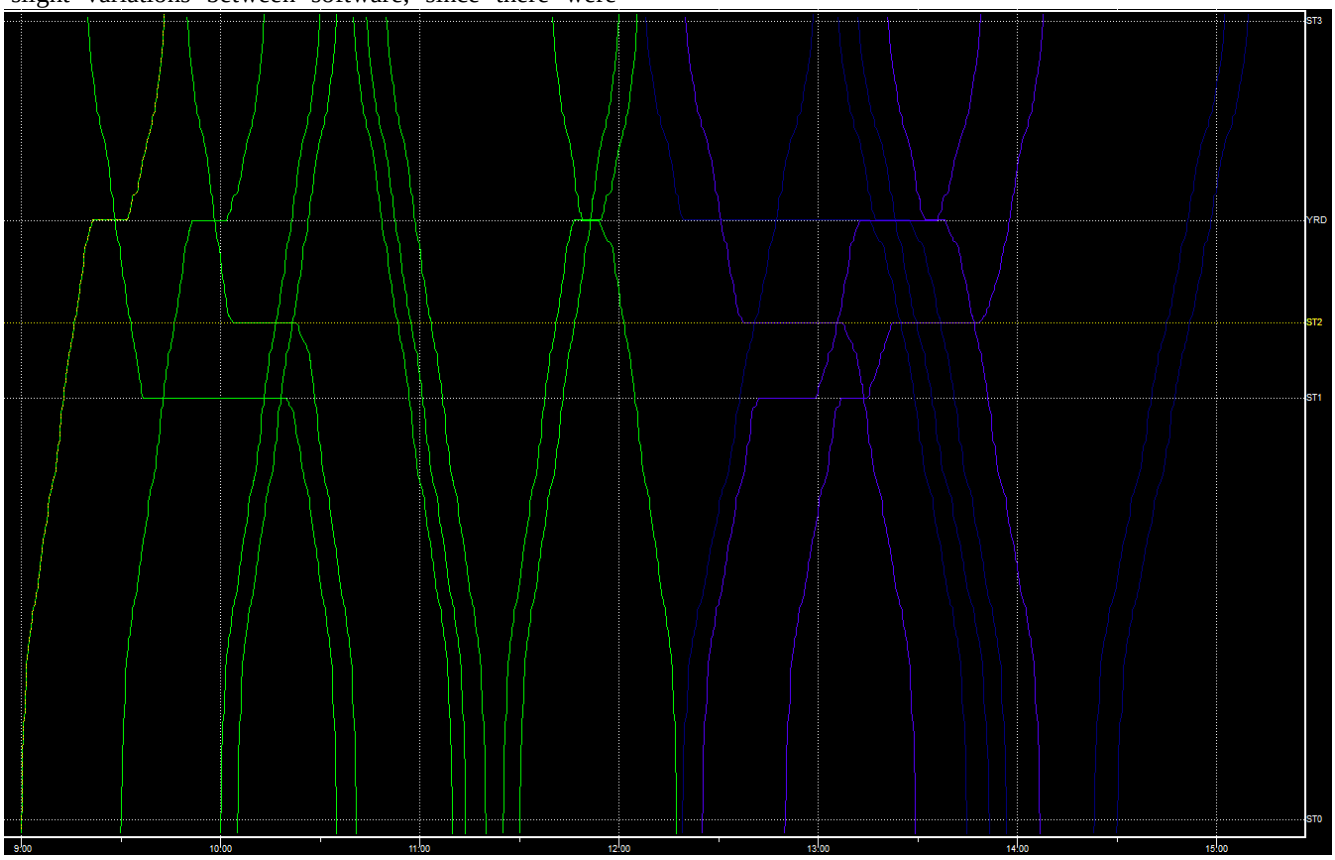

Figure 6- The initial timetable in Railsys, developed based on RTC simulation output)

the compression procedure of RailSys with considering overtaking option in maximum two stations based on OBB compression algorithm. DB algorithm wasn't used in this study, as it considers one of the trains as a dummy train for the purpose of compression technique, causing the number of simulated trains to deviate from RTC result. Figure 7 presents the final results of optimized timetable by using UIC 406 compression approach, based on different duration scenarios and dwell times in sidings/yard.

As presented in Figure 7, by increasing the timetable duration, the capacity utilization of current trains is reduced, opening up capacity for new trains. In addition, considering dwell time can change the capacity utilization. For instance, timetable with no dwell time in the stations will allow all trains to pass through the sidings without any stop. This approach consumes more capacity levels (Figure 8), while 10 minutes of dwell time, according to Figure 7, can consume less capacity between scenarios. However, dwell times of more than 30 minutes utilizes more capacity in Railsys compression technique (based on initial timetable information), in comparison to the 10 and 20 minutes scenarios, since the software may stop some trains unnecessarily in the sidings which could have been passed or stopped with shorter dwell times. Thus, considering a new option of dwell time for each train or certain categories of trains through RailSys capacity optimization features, may improve the outcomes of timetable compression in RailSys.

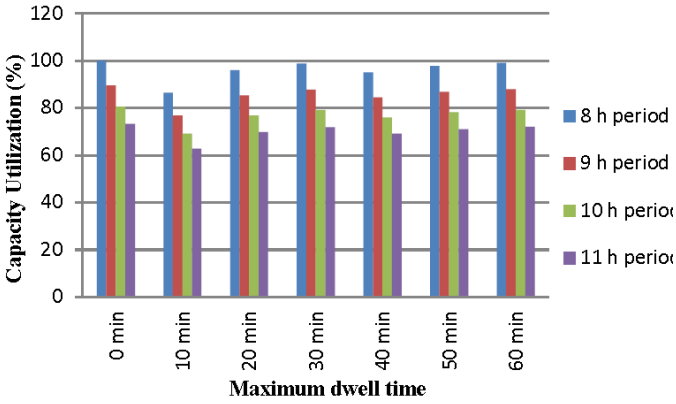

Figure 7- Capacity utilization in RailSys optimization feature, based on different scenarios of timetable duration and maximum dwell times

As presented in Figure 8 and Figure 9, the optimized timetables by RailSys have limited delay or meet-pass stops in the sidings while the initial timetable inserted to the software as input data (Figure 6) had several delays with even more than 90 minutes of duration. The order of trains in the optimized timetables of RailSys was exactly the same as defined in the input timetable. Although Railsys maintained the same order of trains, the optimized arrival/departure times were different and it was not evident that RailSys optimization technique used the preferred departure times from input timetable. (Table 6)

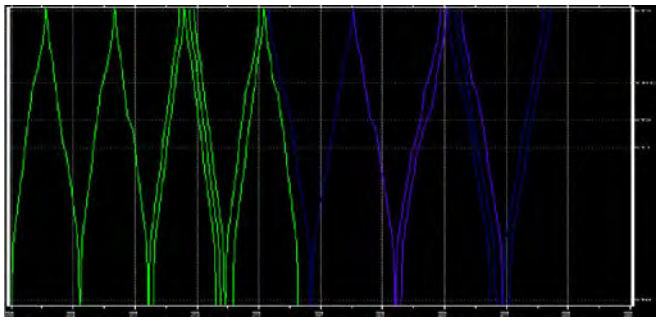

Figure 8- Optimized timetable in RailSys based on no dwell time scenario

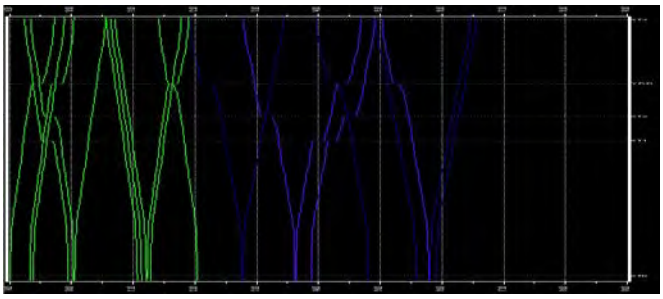

Figure 9- Optimized timetable in RailSys based on maximum10 minutes dwell time scenario

Table 6-Comparison between initial and optimized timetable (departure times) in RailSys (in 10 minutes dwell time scenario)

Train	Initial departure-Eastbound	Optimized departure-Eastbound	Initial departure-Westbound	Optimized departure-Westbound
Pass1	9:00	9:00	9:34	9:20
Pass2	9:30	10:07	10:40	9:50
Pass3	10:05	11:18	11:53	10:44
Pass4	11:25	12:28	11:57	10:50
Comm1	10:00	11:12	11:48	10:40
Comm2	11:30	12:35	13:05	11:40
Interm1	12:19	13:51	13:10	12:08
Interm2	14:23	16:17	16:03	13:06
Interm3	14:30	17:03	16:10	13:12
Freight1	12:25	15:12	14:31	12:20
Freight2	12:50	15:18	16:16	13:21

6- SUMMARY AND CONCLUSIONS

There are several capacity software available in the market for analyzing and evaluating capacity levels of railway network, either as general simulation or commercial railway-specific simulation software. The current research reviewed and compared two common capacity simulation software typically applied in the U.S. (RTC) and Europe (RailSys). RTC is non-timetable based simulation software, and it can be easily applied in railway with improvised operation philosophy to simulate and create timetable. On the other hand, RailSys, is a timetable-based simulation software which needs an initial and feasible timetable as input to provide further analysis and evaluations including capacity optimization option.

To compare the similarities and differences of RTC and RailSys software, a case study was developed and applied in both simulation packages. The comparison of the simulation procedure and outcomes led to the following conclusions:

1. Both software are powerful tools for operations simulation, but the procedure and steps of developing the operations rules and dispatching system in RTC is easier for a given improvised operation philosophy which doesn't have any particular predefined schedule, except the number of daily trains and some preferred departure times. However, the RTC output (developed timetable) may not be optimized and may have several delays and long meet-pass stops which should be manually improved by the user.
2. RailSys can optimize the timetable and provide more capacity levels for a given case study, if there is a feasible timetable of trains inserted in the software as input data.
3. The procedure of developing North American rolling stock and signaling features is relatively challenging in RailSys, since all default database and information have been designed based on European characteristics that may not match the North American railway characteristics. The adjustment and calibration of these parameters to match the desired characteristics and specifications is necessary and time consuming.
4. The timetable development in RTC is based on maintaining the preferred departure times of trains as much as possible, regardless how much delay and duration of meet-pass may occur through the simulation. The RTC user should intervene and manually modify the preferred departure times or assign new meet-pass events to improve the first results of simulated timetables. On the other hand, RailSys compression technique maintains the order of trains through the optimized timetable option (as defined in the input timetable), but doesn't keep the preferred departure times of input timetable through the optimized timetable.
5. Several factors can be defined in the capacity optimization plan of RailSys but overtaking scenario, the amount of dwell time and the algorithm used for timetable compression (OBB vs. DB pattern) seem to have higher impacts on the final results of optimized timetable.
6. It has been derived from the case study run by RailSys that considering maximum 10 minutes of dwell time in the sidings for the timetable compression technique can provide more capacity levels in comparison to the other scenarios. However, considering dedicated dwell time for each train or certain categories of trains, (instead of unique dwell time for every train), may improve the outcomes of timetable compression technique in RailSys.

The next step of the research is to evaluate a real case study of a planned U.S. shared-use corridor through a similar process, (based on more complicated operation and infrastructure characteristics), considering delay and timetable management analysis. The outcomes will be more explicitly compared and analyzed especially considering mixed traffic scenarios of passenger trains with higher speed and heavier/longer freight train operations. The main objective is to evaluate whether the selection of simulation software has any meaningful effect to the outcomes of the capacity analysis, and thus to the recommendations for improvements in the evaluated corridor.

7- BIBLIOGRAPHY

1. *Pouryousef, Hamed, Pasi Lautala, and Thomas White. Review of Capacity Measurement Methodologies; Similarities and Differences in the U.S. and European Railroads. in 2013 TRB Annual Meeting 2013. Washington-DC.*
2. *White, Thomas, Alternatives for Railroad Traffic Simulation Analysis. Journal of the Transportation Research Board (TRB)- Washington, DC, 2005.* ***1916*** *p. 34-41.*
3. *M. Abril, F. Barber, L. Ingolotti, M.A. Salido, P. Tormos, A. Lova An Assessment of Railway Capacity, 2007, Technical University of Valencia: Valencia, SPAIN.*
4. *Pachl, Joern, Railway Operation and Control2002, Mountlake Terrace-WA VTD Rail Publishing- USA.*
5. *Pavankumar Murali, Maged M. Dessouky, Fernando Ordonez, and Kurt Palmer, A Delay Estimation Technique for Single and Double-track Railroads, 2009, Dept. of Industrial & Systems Engineering, University of Southern California: Los Angeles, CA.*
6. *Samuel L. Sogin, Christopher P.L. Barkan. Railroad Capacity Analysis. in Railroad Engineering Education Symposium (REES) by AREMA. 2012. Kansas City.*

7. *Y. Lai, C. Barkan, Enhanced Parametric Railway Capacity Evaluation Tool. Journal of the Transportation Research Board, National Academies, Washington, D.C., 2009.* ***2117****: p. 33-40.*
8. *James S. Noble, Charles Nemmers, Missouri Freight and Passenger Rail Capacity Analysis, 2007, Missouri Department of Transportation (MoDOT): Jefferson City, MO.*
9. *Melody Khadem Sameni, Mark Dingler, John M. Preston, Christopher P.L. Barkan. Profit-Generating Capacity for a Freight Railroad. in TRB 90th Annual Meeting. 2011. Washington, DC: TRB.*
10. *RTC, RTC Online Help, B.S. Software, Editor 2011: USA.*
11. *RMCon, RailSys 8 Classic, RailSys 8 Enterprise, Network-Wide Timetable and Infrastructure Management, R.M.C. GmbH, Editor 2012: Hannover, Germany.*
12. *RMCon, RailSys User Manual, R.M.C. GmbH, Editor 2010: Hannover, Germany.*

REVIEW OF CAPACITY MEASUREMENT METHODOLOGIES; SIMILARITIES AND DIFFERENCES IN THE U.S. AND EUROPEAN RAILROADS

Hamed Pouryousef*

PhD Candidate at Rail Transportation Program,
Civil and Environmental Engineering Department, Michigan Tech. University,
1400 Townsend Drive, Houghton, MI 49931
phone: (906)-231-2320
hpouryou@mtu.edu

Pasi Lautala, Ph.D, P.E.

Assistant Professor
Director, Rail Transportation Program,
Michigan Tech Transportation Institute (MTTI), Michigan Tech. University,
1400 Townsend Drive, Houghton, MI 49931
phone: (906)-487-3547
ptlautal@mtu.edu

Thomas White,

Senior Operations Specialist
Transit Safety Management
3604 220th Pl SW
Mountlake Terrace WA 98043
phone: 425-771-4289
taw@vtd.net

*** Corresponding Author**

Submitted for Presentation at the 2013 Annual Meeting of the Transportation Research Board
August 1, 2012

Total words = 5687 + 250*6 (5 Figures + 1 Tables) = 7,187

ABSTRACT

Most passenger rail services in the United States (U.S.) operate on corridors that are shared with freight traffic, creating more complicated operation practices. As the demand for passenger and freight transportation grows and emphasis is placed on increased speed and on-time performance of passenger services, the available capacity is further consumed.. Where higher speed passenger trains are mixed with freight, the increased heterogeneity from expanding speed differential creates further challenges for reliable operations. Based on the experiences in the other parts of the world, the required reliability is typically secured through structured/planned/scheduled operation. As the U.S. continues to develop higher speed passenger service with similar characteristics to those in European shared-use lines, the accuracy of capacity analysis methods becomes more important, and tools applied in Europe may become more applicable to the U.S. conditions as well. This paper presents the fundamental particulars on railway capacity obtained through the literature review. It will provide a brief review of capacity definitions used in both Europe and the U.S., followed by description of differences in the respective rail systems. The paper will also introduce the main methodologies of capacity measurement approaches, and highlights several capacity analysis case studies conducted in the U.S. and Europe.

INTRODUCTION

Typically, the capacity of rail line is defined as the number of trains that can safely pass along a given segment of the line through a period of time and is affected by different system configurations, such as: 1) Track infrastructure, 2) Signaling system, 3) Operations philosophy, and 4) Rolling stock.

The configuration differences between European and the U.S. rail systems may lead to different methodologies, techniques, and tools to evaluate and measure the capacity levels. There are high utilization corridors in Europe where intercity passenger, commuter, freight, and even high speed services operate on shared tracks and where all train movements follow highly structured timetables and schedules. In the U.S. the prevalent operations pattern on current shared corridors is improvised operational philosophy where some trains are assigned their slots in the network on a daily basis. Recently, the U.S. has placed an increasing emphasis to introduce new, or to incrementally increase the speeds of passenger services on selected shared corridors [1] while the slower speed freight rail transportation is also expected to increase [2]. These increases in volumes and operational heterogeneity can be expected to add pressure for higher capacity utilization of the U.S. shared-use corridors. Capacity measurement and analysis approaches, methods and tools play a crucial part in preparing the U.S. network for these changes. The accuracy and applicability of these in the U.S. environment should be carefully evaluated. It would also be beneficial to investigate whether the analysis and operations approaches utilized in Europe would provide any benefits for the U.S. application.

This paper focuses on reviewing the capacity analysis approaches and methodologies in the U.S. and Europe. The paper provides a brief review of different capacity definitions, identifies main differences between the U.S. and European rail systems, reviews the main methodologies of capacity measurement approaches, and highlights several capacity analysis case studies conducted in the U.S. and Europe.

WHAT IS CAPACITY?

Capacity Concept and Definitions

The definition used for rail capacity in the literature varies based on the techniques and objectives of the specific study. For instance, Barkan and Lai defined capacity as "a measure of the ability to move a specific amount of traffic over a defined rail line in the U.S. rail environment with a given set of resources

under a specific service plan, known as level of service (LOS)". They listed several infrastructure and operational characteristics which affect capacity levels, such as: length of subdivision, siding length and spacing, intermediate signal spacing, percentage of number of tracks (single, double and multi-tracks), heterogeneity in train types (train length, power-to-weight ratios). [3] In another piece of U.S. literature, Tolliver introduced freight rail capacity as the number of trains per day for typical track configurations depending on several factors, such as track segment length, train speed, signal aspects and signal block length, directional traffic balance, and peaking characteristics. [4] American Railway Engineering and Maintenance-of-way Association (AREMA) offers a simplified approach for line capacity that estimates practical capacity by multiplying theoretical capacity (Ct) and dispatching efficiency (E) of the line (C= Ct×E). AREMA's method for calculating theoretical capacity and dispatching efficiency require consideration of various factors, such as number of tracks, the operations rules (single or bi-direction operation), stopping distance between trains (or headway), alignment specifications (grade, curves, sidings, etc.), trains specifications (type of train, length, weight, etc.), maintenance activities requirements, and the signaling and train control systems. [5]

In Europe, the most common method for capacity analysis is provided by International Union of Railways (UIC) code 406. According to UIC 406, there is no solid definition of capacity and the rail infrastructure capacity concerns and expectations vary between different points of view by railroad customers, infrastructure planner, timetable planner, and railroad operators. UIC also emphasizes that the capacity is affected by interdependencies and the interrelationship between the four major elements of railroad as shown in Figure 1. [6]

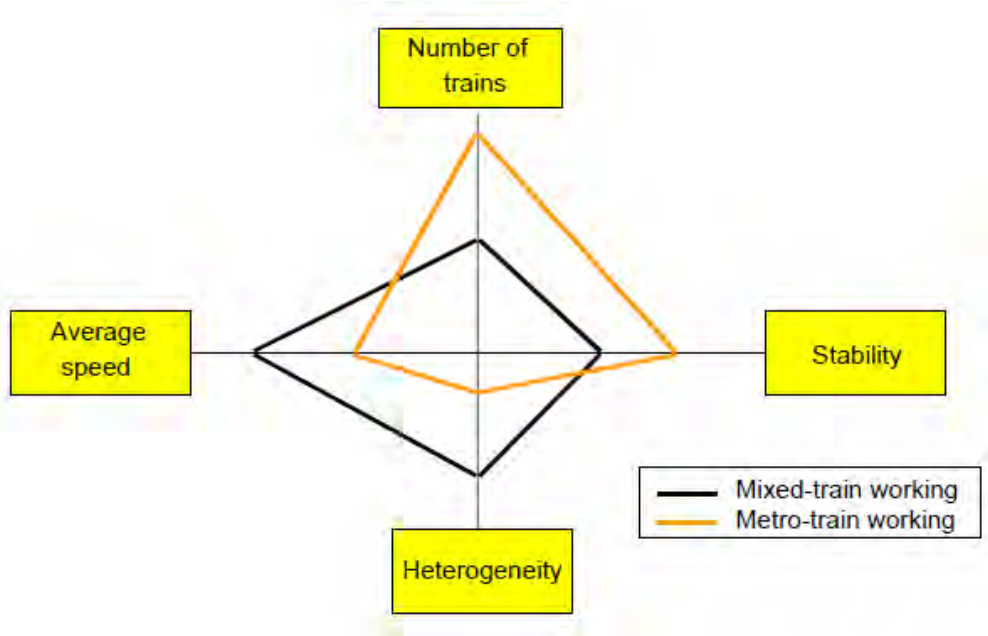

FIGURE 1- Capacity balance according to UIC code 406 definition [6]

According to UIC, the "Theoretical Capacity" is the absolute maximum capacity which can be achieved subject to:

- Absolute train-path harmony (the same parameters for majority of trains)
- Minimum headway (shortest possible spacing for all trains)
- Providing best quality of service

UIC also recognizes that it is almost impossible to achieve theoretical capacity in practice. [6]

Besides the UIC literature, research conducted as part of European Commission's Improve Rail project produced a definition of ultimate capacity that was similar to the UIC's theoretical capacity definition, but placed higher emphasis on the train schedules and running time. [7]

Capacity Metrics

The literature divides the main types of metrics to measure the capacity levels to three groups: throughput (such as number of trains, tons, train-miles), level of service (LOS) (terminal/station dwell, punctuality/reliability factor, and delay), and asset utilization (velocity, infrastructure occupation time or percentage). [8] In 1975, The Federal Railroad Administration (FRA) introduced a parametric approach developed by "Peat, Marwick, Mitchell and Co" to measure capacity in the U.S. rail network based on delay unit (hours per 100 train-miles). [4] The European rail operators typically use throughput metrics (number of trains per day or hours) to measure the capacity levels, although punctuality and asset utilization metrics are also applied as secondary units. [7, 9]

MAIN DIFFERENCES BETWEEN THE U.S. AND EUROPEAN RAIL SYSTEMS

The U.S. and European rail networks have several similarities, such as operating mixed traffic on shared-use corridors, and using modern signaling and traffic control systems (e.g., developing ETCS in Europe and PTC in the U.S.). On the other hand, significant differences between the U.S. and European networks also exist and they may change the preferred methodologies and the outcomes of capacity analysis. Figure 2 and the following discussion uses the literature review to highlight some of the key differences between infrastructure, signaling, operations and rolling stock in Europe and the U.S.

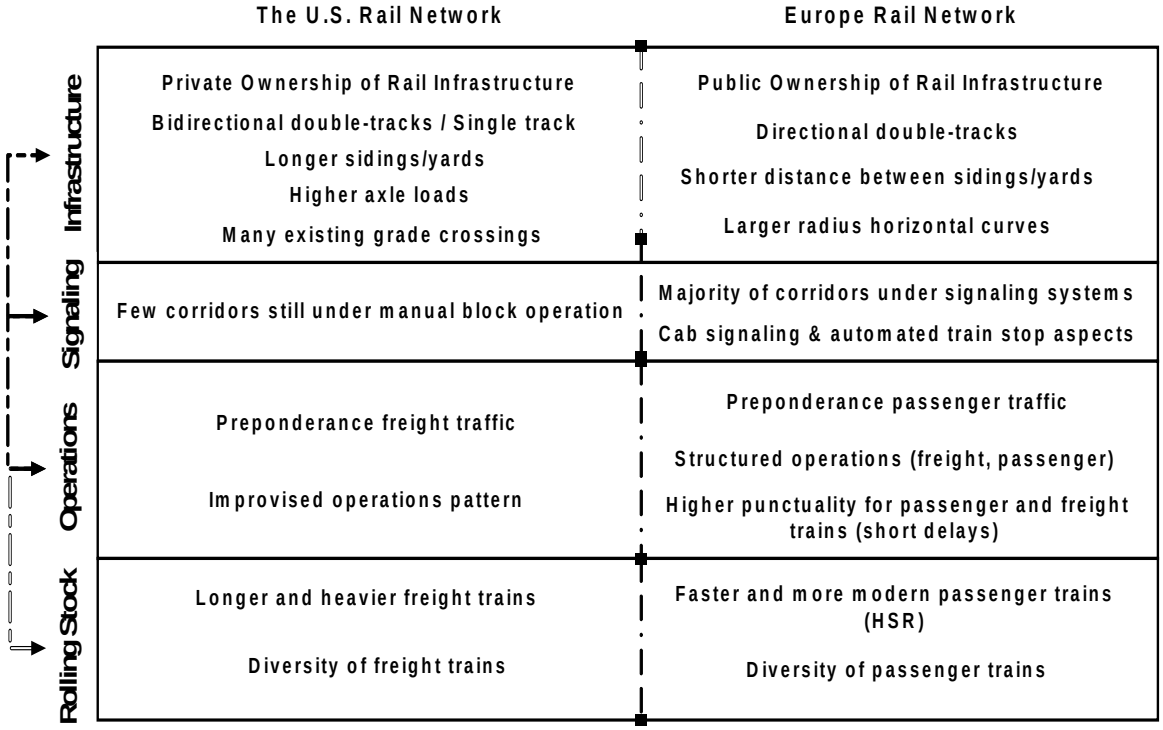

FIGURE 2- The main differences in the U.S. and Europe rail systems

Infrastructure Characteristics

- **Public vs. Private Ownership of Infrastructure**: The ownership of rail infrastructure is one of the main differences between Europe and the U.S. rail networks. More than 90% of the infrastructure is

owned and managed by private freight railroads in the U.S.; while almost all infrastructures are owned and managed by governments or public agencies in Europe. In addition, operations and infrastructure are vertically separated in Europe while in the U.S., the majority of operations (mainly freight) are handled by the some corporations who own the infrastructure. The ownership and vertical separation have wide impact in the railway system. Perhaps the greatest effect is on the prioritization of operations and accessibility for operating companies, but other aspects, such as operations philosophy, maintenance strategy and practices, signaling and train control systems, rolling stock configuration, and capital investment criteria are also affected. [4, 10]

- **Single vs. Double-Track:** More than 46% of rail corridors in Europe are at least double-track [11, 12], while approximately 80% of the U.S. rail lines are single-track. [2, 4]
- **Directional vs. Bidirectional:** Most of the U.S. double tracks operate in bidirectional fashion and use crossovers along the corridor, while directional operation with intermediate sidings and stations is the common approach in Europe. [4]
- **Distance between Sidings:** The distances between stations and sidings in the European rail network are generally shorter than the U.S. The siding distribution rate throughout the European network (total route mileage per number of stations, including freight and passenger services) is approximately four miles/station in both UK and Germany [12, 13]. In the U.S. the distance between sidings varies greatly and passing sidings on double-track sections are relatively far apart. [10, 14]
- **Siding Length:** Siding/yard tracks in the U.S. are typically longer than the European rail network, but in many cases are still not sufficient for the longest freight trains. [10, 15]
- **Track Conditions:** Typically, railroad structure in the U.S. is designed for higher axle loads, but tighter horizontal curves (shorter radius) and lower maximum speed operations, in comparison to the European rail network. [10, 15]
- **Grade Crossings:** There are approximately 227,000 grade-crossings in operation along the main lines in the U.S. [16, 17], while there are few grade-crossings along main corridors in Europe. High frequency of grade crossings and difficulty of elimination is an operational and safety challenge for increased train speeds. [18]

Signaling Characteristics

- **Manual blocking vs. signaling systems**: Manual blocking is absent on main passenger corridors in the U.S. today, but some of the planned passenger corridors are located along such lines. On the other hand, most shared-use corridors in Europe are equipped with one of the common blocking systems.[19]
- **Cab Signaling:** A more significant difference is the extensive use of cab signaling and enforced signal systems (among which are PTC systems such as ETMS) in Europe, while such implementation is limited in the U.S. [10]

Operation Characteristics

- **Improvised vs. Structured Operation:** The U.S. operations philosophy is based on the improvised pattern, (no repeatable dispatching plan in-advance) for almost all freight trains, except some intermodal trains. On the passenger side, many Amtrak and commuter train daily operation patterns are also developed without details, anticipating improvised resolution of conflict among the passenger trains, or between passenger and freight trains. In Europe, almost all freight and passenger trains have a regular schedule developed well in advance, known as structured operations. [20]
- **Preponderance Freight vs. Passenger Traffic:** The preponderance of U.S. rail traffic is freight while the preponderance of European rail traffic is passenger rail. [4, 21]
- **Delay vs. Waiting Time:** Delay (deviation of train arrival/departure time from what was predicted/planned) and waiting time (scheduled time spent at stations for passing or meeting another train) are two fundamental concepts in the railroad operations. The waiting time concept is typically used in the European rail operation management due to the structured operations pattern in Europe.

Delay is more used in the U.S. capacity analysis as the main capacity metric, while it is limited in Europe to the events that are not predictable in advance. [20]

- **Punctuality:** The punctuality of trains is quite different in the U.S and Europe. Amtrak's trains are considered on-time if they arrive within 15 minutes of a scheduled timetable for short distance journeys (less than 500 miles) or within 30 minutes for long distance trains (over 500 miles). In 2011, Amtrak trains' punctuality was 77% for long-distance trains, 84% for short-distance trains, and 92% for Acela. According to Amtrak, more than 70% of passenger train delays are caused either by the freight trains performance or infrastructure failure. [22] The passenger trains in Europe have shorter average delay per train. For instance, Network Rail in the UK reported that about 90% of all passenger trains were punctual with arrival time deviation within five minutes from planned timetable (short-distance trains) and 10 minutes (long-distance trains) [23]. In Switzerland, more than 95% of all passenger trains are punctual with an arrival delay of five minutes or less. [24] The punctuality of European freight trains in 2003 was reported to be approximately 70%. [25]

Rolling Stock Characteristics

- **Train configuration (length and speed):** Typically freight trains in the U.S. are longer and heavier than freight trains in Europe. Based on the Association of American Railroads (AAR), the number of cars in the average U.S. freight train varies between 63-164 in West and 57-110 in East, while the typical number in Europe is 25-40. In addition, the average speed of intercity passenger trains in Europe is faster than in the U.S. [2, 10, 15]. Freight trains also typically operate on higher speeds and with less variability in Europe.
- **Diversity of Freight vs. Passenger Trains:** The U.S. rail transportation is more concentrated on the freight trains than Europe, and there is a great diversity between the types, lengths, etc. of freight trains. On the passenger side, Europe has more diverse configurations (such as speed, propulsion, train type, power assignment, HSR services, diesel and electric multi-unit trains) in comparison to the U.S. [2, 19]

While the principles of rail capacity remain the same in all rail networks, the characteristics reviewed above all have an effect on capacity and its utilization. What remains unclear is the effect of these differences in various capacity analysis tools and methodologies used and whether they limit the applicability of the U.S. tools in the European environment and vice versa.

CAPACITY MEASUREMENT, ANALYTICAL, SIMULATION AND COMBINED APPROACHES

The literature classifies capacity analysis approaches and methodologies in several different ways. Even though the approaches differ, the input data for most of them is similar and includes infrastructure and rolling stock data, operation rules and signaling features. Abril, et al., classified the capacity methodologies as analytical methods, optimization methods, and simulation methods. [26] Joern Pachl divided the capacity methodologies into two major classes: analytic and simulation. [27] Similar categorization was used in research conducted by Murali on delay estimation technique. [28] Khadem Sameni, and Preston, et al., categorized capacity methods to timetable based and non-timetable based approaches. [8] Finally, research conducted at the University of Illinois, Sogin, Barkan, et al., classified capacity methods as theoretical (analytical), parametric, and simulation methods. [3, 29] The analytical and simulation methods are the most common methods found in the literature. For our review, we have divided methods into three groups; analytical, simulation, and combined. Although the term "combined methodology" is not a commonly used term in the reviewed literature, we added it as new class, because many studies take advantage of both analytical and simulation methods.

Analytical Approach

The analytical approach typically uses several steps of data processing through mathematical equations or algebraic expressions and is often used to determine a solution for the problem (theoretical capacity). [26] The outcomes vary based on the level of complexity of the scenario and may be as simple as the number of trains per day, or include a combination of several performance indicators, such as timetable, track occupancy chart, fuel consumption, speed diagrams, etc. Analytical methods can be conducted with or without specific software. One example of analytical capacity software is SLS PLUS in Germany. SLS PLUS is used in the German rail network (DB Netz AG) for estimating capacity through analytical determination of the performance, asynchronous simulation and manual timetable construction. [30] Figure 3 presents the different methodologies that can be used in the analytical approach and how complexity, such as optimization and timetable compression methods, can be added to provide more detailed results of capacity estimation. In some cases, analytical models are introduced under different names like optimization methods or parametric models, taking advantage of different modeling features, such as probabilistic distribution or timetable optimization. The latter method, timetable optimization, is typically achieved by using specific software, or specific simulation tools. [26, 27]

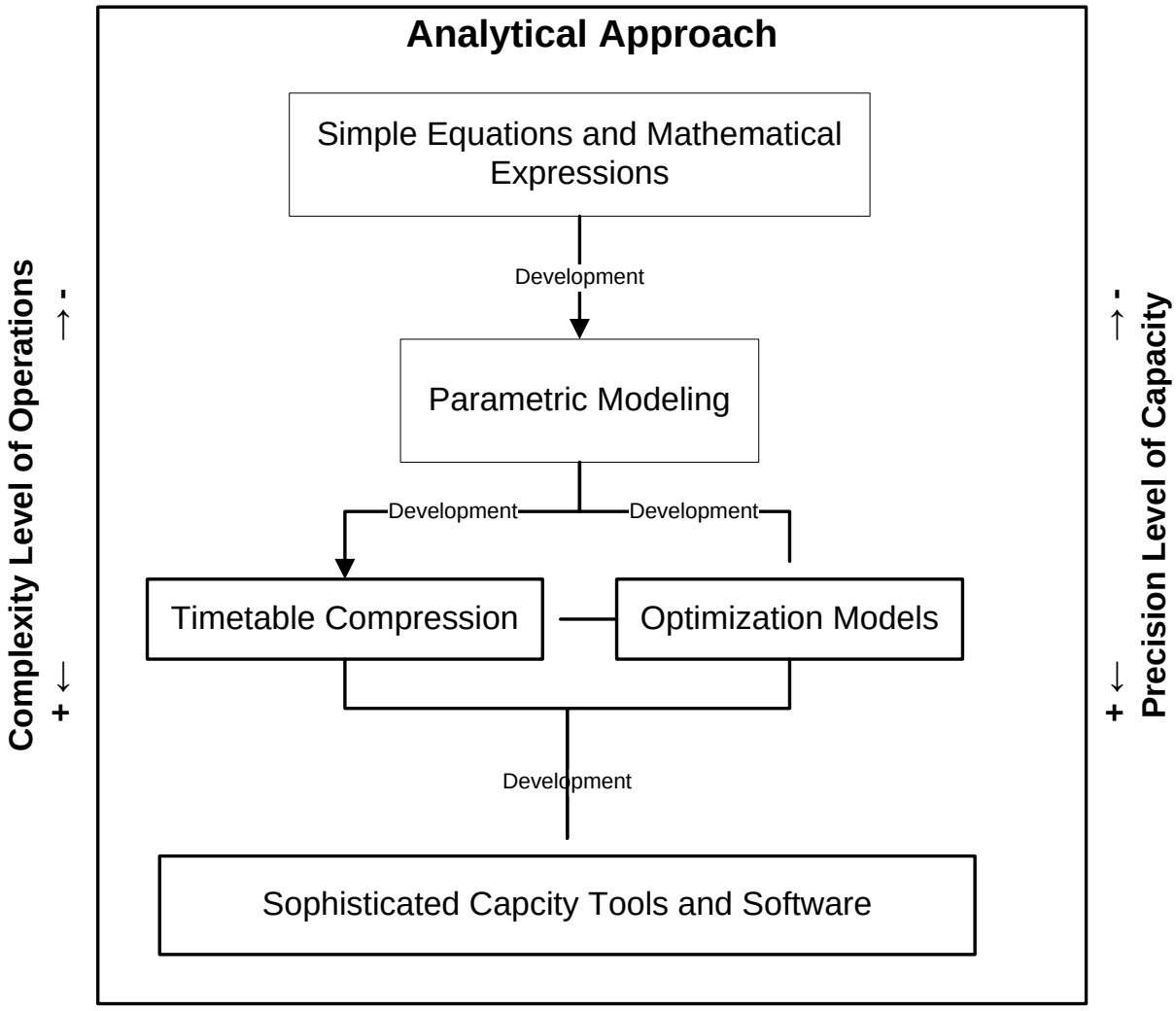

FIGURE 3- The relationship between operations complexity, precision level of capacity and different methodologies of analytical approach

Timetable compression method is one of the main analytical approaches in Europe to improve the capacity levels, especially for those corridors which have pre-scheduled timetables of all daily trains (structured operation pattern). A majority of techniques and tools for improving the capacity levels in Europe, including the UIC method (leaflet 406), are partly developed based on timetable compression. [6, 9, 31-33] The UIC's method modifies the pre-scheduled timetable and reschedules the trains as close as possible to each other. [26] Figure 4 provides an example of the methodology where a given timetable of trains along a quadruple segment of tracks (Scenario a) is first modified by compressing the timetable (Scenario b) and then further improved by optimizing the order of trains (Scenario c). As demonstrated, the third scenario could provide a higher level of theoretical capacity in comparison to the scenarios a, and b. [9]

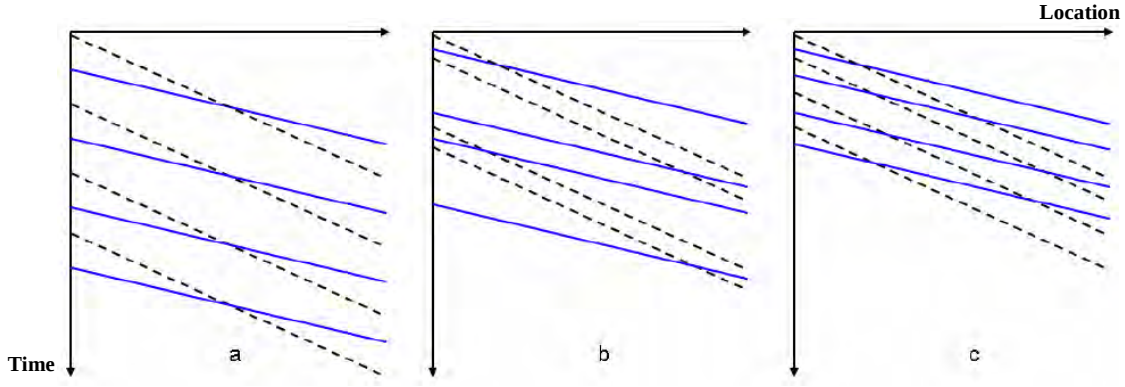

FIGURE 4 - Actual timetable for quadruple track (a) compressed timetable (b) compressed timetable with optimized train order (c) (chart layout follows typical European presentation) [9]

Simulation Approach

Simulation is an imitation of a system's operation which should be as close as possible to its real-world equivalent. [26] The process of simulation is repeated several times until an acceptable result is achieved by the software (Heuristic approach). The data needed for the simulation are similar to the analytical methods, but typically at a higher level of detail. The simulation practices in rail industry started in the early 1980s by developing models and techniques, such as dynamic programming and branch-and-bound, proposed by Petersen, as well as heuristic methods developed by Welch and Gussow in 1986. Today, the simulation process utilizes computer tools to handle sophisticated computations and the stochastic models in a faster and more efficient way. The commercial railroad simulation software is typically developed based on two major components; 1) Train movement simulation, and 2) Train dispatching simulation. The first component calculates the train speed along the track by using the train resistance formula (like Davis equation) and train traction power. The dispatching simulation component typically emulates (or attempts to emulate) the action of the dispatcher in improvising traffic management, but in some cases, it can be used as part of a traffic management software to help traffic dispatchers to manage and organize the daily trains' schedules. [20]

According to Pachl, the simulation method can also be divided into asynchronous and synchronous methods. Asynchronous simulation attempts to consider stochastically generated train paths within a timetable, following the scheduling rules and the train priorities. In synchronous simulation, the process of rail operations is followed in real time sequences and the results are expected to be closely aligned with data of real operations. In contrast to asynchronous method, it cannot directly simulate the scheduling, or develop a timetable, unless the simulation results are used by additional computer tools and programs to create a timetable. [27] The outputs of simulation software typically include several parameters such as delay, dwell time, waiting time, elapsed time (all travel time), transit time (time between scheduled stops), trains speed, and fuel consumption of trains. [20, 26]

Simulation Methods: Timetable Based vs. Non-timetable Based

The commercial railroad simulation software can be classified in two major groups; 1) non-timetable based vs. 2) timetable based. The non-timetable based simulations are typically applied in railways that operate based on the improvised operation pattern without initial timetable, such as the majority of the U.S. rail network. In this type of simulation, after loading the input data in the software, the train dispatching simulation process improvises the departure times from the initial station that are included in the input data. The software may encounter a problem to assign all trains and request assistance from the software user to resolve the issue by adjusting the train data, or by modifying the schedule constraints. [8, 20] The rail traffic controller (RTC), developed by Berkeley Simulation Software is the most common software in this category, used extensively by the U.S. rail industry. [8]

The simulation procedure in timetable based software (typically used in Europe) is based on the initial timetable of trains and uses software tools to improve the timetable as much as possible. The UIC's capacity approach would be one of the main theories behind timetable based simulation approach. The simulation process in this methodology begins with creating a timetable for each particular train. In the case of schedule conflict between trains, the user must change the timetable until the feasible schedule is achieved; however, the user interference is not arbitrary as in the improvised method, but it is implemented as part of the simulation process. [20] Common software used in this category are: MultiRail (U.S), RAILSIM (U.S), OpenTrack (Switzerland), SIMONE (Netherland), RailSys (Germany), DEMIURGE (France), RAILCAP (Belgium), and CMS (UK). [8, 26]

Combined Analytical-Simulation Approach

In the combined approach, simulation tools are used to evaluate and understand the capacity bottlenecks through the corridor, and analytical methods are developed to improve the capacity utilization levels. The process of applying simulation-analytical practices may be repeated by the research team until an acceptable set of outputs and alternatives is found. (Figure 5)

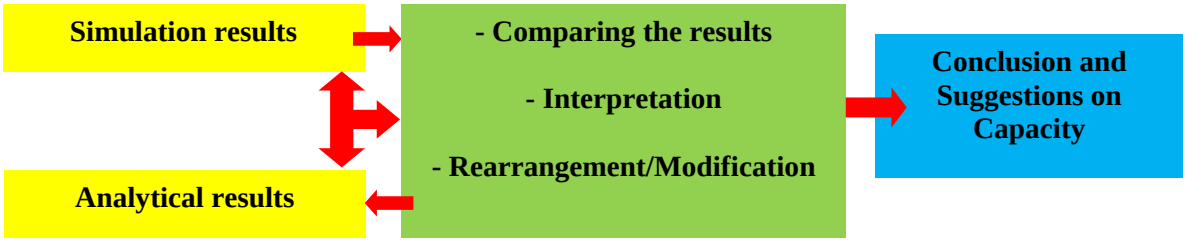

FIGURE 5 - Basic diagram of combined analytical-simulation approach to evaluate capacity

As an example of combined analytical-simulation approach, the Missouri DOT analyzed the rail capacity on the Union Pacific (UP) line between St. Louis to Kansas City in 2007 to improve the reliability of service for the passenger trains and to reduce the freight train delay. Six different alternatives were generated based on a Theory of Constraints analysis and then compared with each other using the Arena simulation method. Finally, a set of recommendations were proposed with respect to delay reduction and capital investment for each proposed alternative. [34]

In another project, Washington DOT (WSDOT) conducted a master plan in 2006 to provide a detailed operation and capital plan for the Amtrak Cascades intercity passenger rail program. The capacity of the corridor was also evaluated using the combined simulation-analytical approach. First, analytical methods were used to determine the proposed infrastructure. The proposed traffic and infrastructure were simulated with RTC software to test the infrastructure and operational results. After running simulation on RTC software, an analytical method, called Root Cause analysis, was applied to

evaluate the simulation output. The objective of Root Cause analysis method was to re-adjust the simulation outputs to be more accurate, in addition to locating infrastructure bottlenecks which caused the capacity issues and delays. [35]

The Swedish National Rail Administration (Banverket) carried out a research project in 2005 to evaluate the application of the UIC capacity methodology (timetable compression) for the Swedish rail network. Railsys software was used for the simulations and the research team analytically evaluated the capacity consumption, its relationship with time supplements and the traffic simulation punctuality. The research concluded that the time supplements are absolutely necessary for the recovery time. When there is no time supplement, the service punctuality can be significantly degraded by increasing capacity consumption. Banverket also confirmed the validity of the framework and the results of the UIC's approach in their network and asked their experts and consultants to implement this capacity approach when using different software such as Railsys, Simon, and OpenTrack. [32]

REVIEW OF CAPACITY CASE STUDIES IN THE U.S. AND EUROPE

Several capacity-related case studies (CS) have been conducted in the U.S. and Europe. The research team reviewed more than 40 studies and selected studies (16 of them) that included sufficiently detailed explanation of the used capacity analysis approach and respective results for further review and evaluation. Table 1 and the following discussion summarize the approach, tools, purpose, types of effort and outcomes, and accuracy assessment of these case studies.

TABLE 1- Review of 16 selected Case Studies (CS) in the U.S. and Europe [2, 3, 8, 9, 20, 24, 28, 31, 32, 34-40]

Criteria		North America	Europe
Capacity Approach [Ref.]	*Analytical*	4 Case Studies (4 CSs) [2, 3, 28, 36]	-
	Simulation	3 CSs [37-39]	3 CSs [9, 24, 32]
	Combined analytical-Simulation	4 CSs [8, 20, 34, 35]	2 CSs [31, 40]
Tools/ Software (name of the tools) [Ref.]	*Mathematical/ Parametric modeling*	3 CSs (Decision support tools, Waybill/FAF) [2, 3, 36]	-
	General Simulation software	2 CSs (AweSim/Minitab, Arena) [28, 34]	-
	Timetable based simulation software	-	5 CSs (Railsys, OpenTrack, SIMU) [9, 24, 31, 32, 40]
	Non-Timetable based simulation software	6 CSs (RTC) [8, 20, 35, 37-39]	-
Purpose of Research [Ref.]	*New methodology development/methodology approval*	4 CSs [3, 8, 20, 36]	4 CSs [9, 24, 31, 32]
	Master plan/capacity analysis	3 CSs [2, 34, 35]	-
	Academic research/project	4 CSs [28, 37-39]	1 CS [40]
Type of outcomes /solutions [Ref.]	*Delay analysis/improvement*	3 CSs [28, 37, 39]	-
	Infrastructure development,	1 CS [2]	-
	Rescheduling/ operation changes	2 CSs [20, 38]	3 CSs [24, 31, 40]
	Combination of above solutions	2 CSs [34, 35]	
	New Tools / methodology approval	3 CSs [3, 8, 36]	2 CSs [9, 32]
Accuracy of simulation results [Ref.]	*Base Model*	4 CSs [3, 20, 34, 35]	2 CSs [9, 40]
	Base and Alternative results	1 CS [28]	-
	No Comparison	6 CSs [2, 8, 36-39]	3 CSs [24, 31, 32]

Approach: Most case studies used either simulation or combined analytical-simulation approaches. Yet, research conducted by Association of American Railroads (AAR), University of Illinois at Urbana-Champaign (UIUC) and University of Southern California (USC), applied analytical-only methodologies.

Tools and Software: All European case studies used timetable based simulation software while the U.S. case studies relied on other tools like optimization/parametric modeling (applied by UIUC and USC), general simulation software (e.g., Arena) and non-timetable based rail capacity software (RTC).

Purpose of Research: Three different subcategories of research purposes were identified: 1) introducing new methodology for capacity evaluation, 2) evaluating the capacity status of a given corridor as part of a corridor master plan development, and 3) an academic research on different capacity issues. The majority of European case studies (Denmark, Austria, and Switzerland) were conducted by industry to justify and evaluate the UIC's approach (UIC code 406) for capacity methodology while the objectives of the U.S. case studies included all three subcategories.

Type of Outcomes or Solutions: The outcomes and solutions obtained from the U.S. case studies varied from delay analysis and suggested improvements (UIUC by using RTC and USC by using Awesim/Minitab), to rescheduling and recommendations related to current operations (UIUC and White), infrastructure development, and combination of all outcomes mentioned above (typically as part of the master plans). In addition, new tools and parametric models were also evaluated as the final outcome of three U.S. case studies (all by UIUC). The outcomes of European case studies were not as diverse, as they either approved the application of UIC's capacity methodology to be used on their network, or suggested network rescheduling and operational changes (the timetable compression concept). One of the common conclusions of various case studies was the identification of operational heterogeneity as a major reason of delay, especially in the U.S. rail network with improvised operation pattern.

Accuracy of Simulation Results: Some of the case studies assessed the accuracy of simulation results in comparison to the real practices. Three types of accuracy assessments were conducted:

- *Base Model*: Only the results of basic model were compared with the real data. Several case studies in both the U.S. and Europe regions used this type of assessment.
- *Base and alternative results*: In addition to basic model comparison, the alternative outcomes were compared with the real data. Only the USC case study can be considered in this category.
- *No comparison*: In the final category no specific information or comparison were provided between simulated results and real practices.

As presented in Table 1, majority of the case studies did not address the accuracy of simulation results, either because case study was not constructed based on real operational data, or simulation results were not compared with the real practices as part of research. The case studies that used general simulation software claimed that capacity delays derived from the modeling approach were close to the real operation practices.

CONCLUSIONS AND NEXT STEPS OF RESEARCH

This paper has used a literature review to provide an overview of the capacity definitions in the U.S. and Europe, to discuss the main similarities and differences between their respective rail systems and to introduce different approaches and methodologies for capacity analysis. The review revealed no single definition of rail capacity, but it can rather be interpreted in various ways based on different perspectives and tools and parameters applied. There are several differences between the U.S. and European rail systems that affect the capacity, such as ownership, type and extent of double track network, distance

between and length of sidings, operation philosophy, punctuality of services, preponderance passenger traffic, and train configurations, but the effect of these differences on capacity or capacity analysis hasn't been evaluated in detail.

The capacity analysis approaches and methodologies can also be classified in several different ways. The methods were typically divided into analytical and simulation methods, but this paper also offered an additional "combined" category. The case studies revealed that majority of analysis utilize simulation approaches, but analytical methods have also been used, either by themselves, or in combination with simulations. The European rail networks typically take advantage of several commercial simulation software available in Europe, which have been developed based on the timetable compression concept, while the U.S. railroads usually apply the non-timetable based simulation, in addition to the general analytical tools and modeling approaches. The accuracy of the simulation results is a major concern when conducting the analysis, but the case studies showed limited effort in comparing the simulation results to the actual conditions, especially after recommended improvements were implemented.

The literature review and case studies presented in this paper are part of an effort to develop a foundation for a more in-depth analysis of current capacity analysis tools and methodologies used in the U.S. and in Europe. As the U.S. continues developing its passenger traffic on shared corridors, the future operation patterns of shared corridors in the U.S. will likely have closer resemblance to the European shared-use lines. The objective of the next research steps is to apply both the U.S. and European based methods on selected U.S. corridors and evaluate the applicability and accuracy of both approaches and tools in the U.S. environment. An interesting additional research question is whether implementation of the structured operational approach that is currently prevalent in Europe would provide any benefits for the shared U.S. corridors and what are the roadblocks or obstructions for its implementation.

REFERENCES:

1. FRA, *Vision for High-Speed Rail in America*, U. DOT, Editor 2009, Federal Railroad Administration (FRA): Washington, DC.
2. Cambridge Systematics, Inc., *National Rail Freight Infrastructure Capacity and Investment Study*, 2007, Association of American Railroads (AAR): Cambridge, Massachusetts.
3. Y. Lai, C. Barkan, *Enhanced Parametric Railway Capacity Evaluation Tool.* Journal of the Transportation Research Board, National Academies, Washington, D.C., 2009. **2117**: p. 33-40.
4. Tolliver, Denver, *Railroad Planning & Design, CE 456/656 Lecture Notes*, 2010, North Dakota State University.
5. AREMA, *Manual for Railway Engineering*, in *Volume 4- Systems Management*2006, American Railway Engineering and Maintenance-of-Way Association: Lanham, MD.
6. UIC, *UIC CODE 406R-Capacity-1st edition*, June 2004, International Union of Railways (UIC): Paris, France.
7. IMPROVERAIL, *IMPROVEd tools for RAILway capacity and access management-Deliverable 6 Methods for Capacity and Resource Management*, 2003, COMPETITIVE AND SUSTAINABLE GROWTH (GROWTH) PROGRAMME - European Commission: Brussels-Belgium.
8. Melody Khadem Sameni, Mark Dingler, John M. Preston, Christopher P.L. Barkan. *Profit-Generating Capacity for a Freight Railroad*. in *TRB 90th Annual Meeting*. 2011. Washington, DC: TRB.
9. Alex Landex, et al. *Evaluation of Railway Capacity*. in *Annual Transport Conference*. 2006. Aalborg University, Denmark
10. White, Thomas, *Comparison of U.S. and European Railroads, Lecture notes for CE-4490*, 2012, Michigan Tech. University-CEE Department: Houghton, MI.

11. Railway-Technical. *Railway Statistics for Britain*. 2008 May 2012]; Available from: http://www.railway-technical.com/statistics.shtml.
12. ERRAC, *A comparison of Member State public research programmes with the ERRAC SRRA 2020*, 2004, European Rail Research Advisory Council (ERRAC): Brussels, Belgium.
13. DB-AG, *Station Categories 2012 (in German)*, 2011, DB-AG: Berlin-Germany.
14. Amtrak, *Annual Report Fiscal Year 2011*, 2011, Amtrak: Washington, DC.
15. Teixeira, Paulo, *Railway Engineering Principles,*, 2009, Technical University of Lisbon, CTIS Master's course in MIT-Portugal Program: Lisbon, Portugal.
16. FRA, *Highway-Rail Grade Crossing Safety Fact Sheet*, 2008, FRA Public Affairs: Washington, DC.
17. DOT, *National Transportation Statistics* 2011, U.S. Department of Transportation: Washington, DC.
18. NetworkRail. *CURTAIL CRAZY DRIVING AT LEVEL CROSSINGS OR RISK MORE LIVES, SAYS RAIL CHIEF*. 2010 [cited 2012; Available from: http://www.networkrailmediacentre.co.uk/Press-Releases/CURTAIL-CRAZY-DRIVING-AT-LEVEL-CROSSINGS-OR-RISK-MORE-LIVES-SAYS-RAIL-CHIEF-13c1/SearchCategoryID-2.aspx.
19. Verkehrswissenschaftliches, *Influence of ETCS on the line capacity*, 2008, International Union of Railways (UIC): Paris, France.
20. White, Thomas, *Alternatives for Railroad Traffic Simulation Analysis.* Journal of the Transportation Research Board (TRB)- Washington, DC, 2005. **1916** p. 34-41.
21. Systra, *Technical Classification and Main Technologies for the HSR Systems*, in *Complementary studies of Tehran-Qom-Isfahan HSR*June. 2007: Paris, France.
22. Amtrak, *Monthly Performance Report for December 2011*, 2012, AMTRAK: Washington, DC.
23. Wearden, Graeme, *UK Train Times Back on Track*, in *the Guardian*2007: UK.
24. Marco Luethi, Andrew Nash, Ulrich Weidmann, *Increasing Railway Capacity and Reliability through Integrated Real-Time Rescheduling*, 2007, Swiss Federal Railways, SBB AG: Zurich, Switzerland.
25. CER, *Rail freight quality: meeting the challenge, A Report on the first year of the CER-UIC-CIT Freight Quality Charter*, December 2004, COMMUNITY OF EUROPEAN RAILWAY AND INFRASTRUCTURE COMPANIES (CER): Bruxelle, Belgium.
26. M. Abril, F. Barber, L. Ingolotti, M.A. Salido, P. Tormos, A. Lova *An Assessment of Railway Capacity*, 2007, Technical University of Valencia: Valencia, SPAIN.
27. Pachl, Joern, *Railway Operation and Control*2002, Mountlake Terrace-WA VTD Rail Publishing-USA.
28. Pavankumar Murali, Maged M. Dessouky, Fernando Ordonez, and Kurt Palmer, *A Delay Estimation Technique for Single and Double-track Railroads*, 2009, Dept. of Industrial & Systems Engineering, University of Southern California: Los Angeles, CA.
29. Samuel L. Sogin, Christopher P.L. Barkan. *Railroad Capacity Analysis*. in *Railroad Engineering Education Symposium (REES) by AREMA*. 2012. Kansas City.
30. Schultze, Gast Ingenieure, *SLS PLUS*, S.G. Ingenieure, Editor 2012: Berlin- Germany.
31. Robert Prinz, Josef Hollmuller *Implementation of UIC 406 Capacity Calculation Method at Austrian Railway (OBB)*. 2005.
32. BANVERKET, *Applicationof the UIC Capacity leaflet at Banverket*, 2005, Banverket, Swedish National Rail Administration.
33. Melody Khadem Sameni, Mark Dingler, John M. Preston. *REVISING THE UIC 406 METHOD: REVENUE GENERATING CAPACITY- JRC2010-36281*. in *Joint Rail Conference (JRC) 2010*. 2010. Urbana, IL, USA.
34. James S. Noble, Charles Nemmers, *Missouri Freight and Passenger Rail Capacity Analysis*, 2007, Missouri Department of Transportation (MoDOT): Jefferson City, MO.

35. Transit Safety Management Inc.; HDR Engineering, *Amtrak Cascades Operating and Infrastructure Plan - Technical Report*, 2006, Washington State Department of Transportation-Freight Systems Division.
36. Yung-Cheng Lai, Christopher Barkan, *A Comprehensive Decision Support Framework for Strategic Railway Capacity Planning.* In Press - ASCE J. Transportation Engineering, 2011.
37. Samuel L. Sogin, Christopher P.L. Barkan, Yung-Cheng Lai, Mohd Rapik Saat. *MEASURING THE IMPACT OF ADDITIONAL RAIL TRAFFIC USING HIGHWAY & RAILROAD METRICS*. in *2012 Joint Rail Conference (JRC)*. 2012. Philadelphia, Pennsylvania.
38. Mark Dingler, Yung-Cheng (Rex) Lai, Christopher P.L. Barkan. *Impact of Operational Practices on Rail Line Capacity: A Simulation Analysis*. in *2009 Annual AREMA Conference*. 2009. Chicago, IL.
39. Samuel L. Sogin, Christopher P.L. Barkan, Yung-Cheng Lai, Mohd Rapik Saat. *IMPACT OF PASSENGER TRAINS IN DOUBLE TRACK NETWORKS*. in *2012 Joint Rail Conference*. 2012. Philadelphia, Pennsylvania.
40. Siggel, Stefan, *Analysis of Capacity and Optimization of Infrastructure of the Line Minot (North Dakota) – Havre (Montana)*, in *Institute of Railway Systems Engineering and Traffic Safety*2006, Braunschweig Technical University: Braunschweig-Germany.