MATEA MIKULČIĆ, Ph.D. student[1]
(Corresponding author)
E-mail: matea.mikulcic@fpz.unizg.hr
TOMISLAV JOSIP MLINARIĆ, Ph.D.[1]
E-mail: tomislav.mlinaric@fpz.unizg.hr
[1] University of Zagreb
  Faculty of Transport and Traffic Sciences
  Department of Railway Transport
  Vukelićeva 4, 10000 Zagreb, Croatia

RAILWAY CAPACITY ENHANCEMENT WITH MODERN SIGNALLING SYSTEMS – A LITERATURE REVIEW

ABSTRACT

In times of ever stronger awareness of environmental protection and potentiation of a beneficial modal split, the railway sector with efficient asset utilization and proper investment planning has the highest chance of meeting customer expectations and attracting new users more effectively. Continuous increase in railway demand leads to an increase in the utilization of railway infrastructure, and the inevitable lack of capacity, a burning problem that many national railways are continually facing. To address it more effectively, this paper reviews available methodologies for railway capacity determination and techniques for its enhancement in the recent scientific literature. Particular focus is given to the possibility of increasing railway capacity through signalling systems and installing the European Train Control System (ETCS). The most important relationships with segments of existing research have been identified, and in line with this, the directions for a potential continuation of research are suggested.

KEYWORDS

signalling systems; railway capacity; capacity determination methodologies; capacity enhancement.

1. INTRODUCTION

Numerous predictions of traffic demand trends over the next 30 years show a promising mobility future. Estimates for the rail sector suggest an increase of 76% in passenger transport and 84% in freight across the European Union (EU) [1]. However, traffic increase for railways with traditional signalling systems often leads to degeneration of performance with associated capacity problems and, eventually, a failure to meet the social needs. The latest rail market monitoring report showed that the North-West European railways have up to 70% more network utilization than the EU average. The total length of congested track was of almost 3,000 km in 2016, including 1,000 km of rail freight corridors [2]. This requires an urgent search for a solution to improve the overall quality of the railway service and attract more potential users.

A contribution to safer rail transport opportunities outside mostly closed national frameworks is offered with the European Rail Traffic Management System (ERTMS). Advanced ERTMS functionalities are primarily intended for significant corridors across the EU to replace different incompatible national signalling and train protection systems. While its functional subsystem, the European Train Control System (ETCS), provides necessary signalling and train protection, the Global System for Mobile Communications for Railways (GSM-R) is a second-generation mobile communication technology that allows data and voice exchange. Generally, the railway signalling system is one of many interconnected subsystems in the overall complex railway system responsible for accident prevention and safe control and regulation of rail traffic operations. In maintaining safety, signalling relies on track detection systems that monitor the occupation of tracks, interlocking for technical protection in disabling conflicts, Automatic Train Protection (ATP) in charge of rail traffic supervision and control, and appropriate communication connection [3].

In this paper, the ETCS signalling system influence will be observed from the functional point of view as well as its impact on the railway capacity. More precisely, the aim is to narrow the research area concerned with increasing the railway infrastructure capacity by observing different configurations of the

modern ETCS signalling system at the functional level of operation. For this purpose, understanding the concept of capacity is essential since it cannot be interpreted unequivocally. It is commonly defined as a measure of the ability to move a specific amount of traffic over a defined rail line with a given set of resources under a specific service plan [4]. Accordingly, all capacity-related situations inevitably are linked with the rolling stock, infrastructure which they operate and specific operating plan [5]. Numerous studies agree that connections between these elements generate differences in capacity levels. Different authors bring various parameter choices and categories containing parameters of similar characteristics [6–8]. The most frequently encountered categorization covers *infrastructure*, *traffic* or *operational* groups of parameters [5, 9, 10]. Signalling systems are part of the first category described with a number of signalling aspects, train protection, and block lengths.

To separate train operations safely, the signalling systems characteristically divide the track into blocks by block signals. There are fixed block and moving block signalling systems whose work is related to the blocking theory [11]. In the first case, each section bounded by two blocking signals can be occupied by only one train at the same time, for the whole occupation time. This block occupancy is checked with the components of a track detection system. Graphic representation of all blocking times for the corresponding line forms the blocking time stairway, whose reallocation influences the line capacity. Trains are mutually minimally separated by maximum blocking time for the critical block section. Minimal time between two successive trains (headway) is conditioned by technical constraints from the signalling system and train characteristics in terms of (1) minimum headway time; and (2) the scheduling reliability reserves. Thereby, (1) is influenced by order of the trains, running time differences, including intermediate stops and the speed profiles [12]. On the other hand, (2) improve stability of scheduled operations with reducing the impact of delays due to maintenance (recovery time), avoiding propagation of traffic disturbances (buffer time) and meeting operational needs in stations (waiting time added to dwell times spent in a station for boarding, alighting and transfer of passengers) [11]. The higher the amounts of (1) and (2), the lower the line capacity.

The conducted literature review is divided into five sections. The presented results are based on the comparison and compilation of scientific knowledge from relevant research in the scientific literature related to this subject. From other research areas, those segments that are considered to have an intense impact on the expected results of capacity increase, are included in the analysis with reference to the appropriate literature. Considering the mentioned complexities by which signalling systems affect the railway capacity, the second Section summarizes the available methodologies for rail capacity determination, as the first step towards effective railway capacity management. The third Section brings a review of specific techniques for increasing the capacity from the relevant literature and practice, while the fourth one contains a more detailed review and analysis of the literature addressing capacity enhancement opportunities using ETCS signalling systems. The concluding Section highlights the resulting observations and suggestions for the future research.

2. CAPACITY DETERMINATION METHODOLOGIES

The type and complexity of the methodology used for capacity assessment relate to quantity and detail of information available and specifics of capacity-related parameters. As a result, different techniques provide slightly different accuracy and precision of the evaluated capacity whose interpretation is consistent with the project evaluation in infrastructure planning, traffic scheduling or other operation-related angles of observation [9, 13]. Nevertheless, these techniques often serve as a base for various models that expand the railway capacity, which is why knowing their features is essential.

Classifications and diversity of individual approaches for capacity assessment have been studied extensively [5, 9, 10, 14, 15]. According to [16, 17] capacity utilization, i.e., describing the state of the network, should be considered at both the micro and macro levels and defined by the multiple metrics to avoid neglecting uninvolved effects of only one observation level. The metrics for measuring how well the infrastructure is utilized may differ from *throughput* (such as number of trains, number of passengers, tons, train-km, ton-km) and the *quality of service* (station dwell, delay, speed, percentage of punctuality) to *asset utilization* (speed, infrastructure occupation time, total or in a given

period, percentage). The broadest division subdivides more deterministic *synthetic (static) methods, analytical (dynamic) methods* relying on probabilities and *analogical methods* with asynchronous and synchronous subcategories. In practice, many studies simplify the use of analytical or simulation approaches [7, 11, 17–19]. Recently, a new category of the *graphical method* has been introduced [20]. To demonstrate similarities and differences in application, the essential characteristics of each category and some of their most prominent representatives are briefly presented in continuation.

2.1 Analytical methods

Analytical methods describe the movement of trains by different mathematical formulas and algebraic expressions. Among them, the stochastic and queueing-based techniques are often used to represent the system behaviour or some of its components. Such modelled queueing systems have some parameters fixed and others random or uncertain [21, 22]. Uncertainty is implied under the assumption of the known probability distributions. This was found to be appropriate for addressing the issues with not enough information or to model the parameters related to the quality of traffic, like an estimation of delays, priority assignment, track allocation, track occupation times, or waiting probabilities on the platform tracks [14, 23]. However, because of their simplicity in some cases, they are recommended for initial reference values. These values are often theoretical, i.e. the ideal highest value with the maximum number of trains separated by the minimum headway allowed by signalling system and block length while omitting the characteristic of heterogeneity and disturbances specific for rail traffic [5]. More realistic practical capacity is obtained using appropriate percentages of eligibility for theoretical capacity (usually 60-75%) or appropriate margins [13].

Commonly used analytical methods are a former standard for line capacity determination based on the identification of the most critical track section (bottleneck), the so-called UIC (Union Internationale des Chemins de Fer) 405 method [6], method typically used by the German railway company Deutsche Bahn [24] and Capacity Utilization Index (CUI), a normative method used in Great Britain. Their main disadvantage is the impossibility of capacity calculation for stations. Station capacity is important as often the busiest station in

the double-track line is the bottleneck of the whole system due to the frequent stops and station dwell time [25]. Additionally, CUI does not consider a specific infrastructure occupation based on blocking times, so the results are less precise than the current standard adopted by the UIC, UIC 406 method [16].

2.2 Optimization methods

Optimization methods represent an asynchronous category of methods. Although some authors consider them analytical methods, others classify them separately to obtain more accurate results in time-consuming and computationally challenging problems. Commonly, they include mathematical formulations like integer programming, linear programming, mixed-integer linear programming, nonlinear programming (especially iterative methods), dynamic programming, stochastic programming, robust programming, constraint programming, heuristics and metaheuristics [8]. The addressed problems comprise the network optimization, optimization of punctuality or reliability, different assignments, routing, scheduling, and other planning strategies. The optimization of certain elements usually leads to the search for a minimum value (e.g. delay, travel time, cost), a maximum value (e.g. a number of trains, service reliability, or quality) or finding a satisfactory (robust) solution for all possible scenarios.

One popular optimization method to measure the railway capacity consumption is the UIC 406 method [13]. It is based on a graphical compression of train paths in the preconstructed timetable for a given line section or node. During this rescheduling of train paths with minimum headway times and without buffer times, the train order, running times, running time supplement (5% of journey time), dwell times, or block occupation times must not change [26]. The track occupation ratio is obtained from the comparison of compensated and non-compensated time consumed. It is recommended to be between 75 and 85% in the peak hour and 60 to 70% in a daily period to ensure service stability. Capacity dependence on the choice of section limits the application of the method to smaller parts of the network, due to more straightforward compression and smaller size of the timetable and less impact of other line sections on the network. Besides, the capacity utilization is represented in the capacity balance of the number of trains, the average speed, the traffic heterogeneity, and stability [13].

2.3 Parametric method

To complement the analytical and simulation models, the literature also mentions parametric models for capacity evaluation using line disassembling. Studies that address parametric modelling are mainly conducted in the United States (US) [4, 18, 27, 28]. To determine a network location with problematic capacity (bottleneck), some critical parameters of the railway line are used. According to [4], the capacity of a railway line is modelled by parametric relationships between infrastructure, traffic, and operating parameters. These parameters are presented as capacity curves of subdivisions to predict train delays. In [28] parametric method of practical capacity measurement has been developed based on the section capacity. Because of the resistance to dynamic movement of railway traffic, the parametric models are suitable for strategic capacity planning and decision support tools [18, 28].

2.4 Simulation methods

Simulation methods enable creating cost-effective models true to real systems with different levels of details: microscopic, mesoscopic, and macroscopic. These methods, as well as optimization, do not offer a general solution, as they must be individually adjusted to solve the problem. Similarly, simulation methods and models are mainly applied for capacity research, evaluation, and optimization of the network for scheduling, dispatching, determination of conflicts and railway planning. They use relevant infrastructure, rolling stock and timetable elements, and differential equations for their interactions in two separate execution approaches [5]. First one is more commonly practised in Europe and based on a predefined timetable. The second is non-timetable, based on a timetable defined according to user instructions. It is applicable for railways with unstructured operating philosophy, like in the US. [15]. Unlike with optimization methods, computer-based tools are necessary here. This often complicates solving capacity problems at the level of the entire network but allows safe determination of the future effects on the rail infrastructure or service characteristics. According to literature [15], the existing computer-based solutions for simulations in railways are further divided into *commercial railway simulation tools*, such as OpenTrack, Railsys and SIMONE in Europe and RTC, MultiRail and RailSim in the US, or *general simulation tools* like

Arena, AweSim and Minitab. Both groups have their advantages and disadvantages, regarding the level of complexity, necessary foreknowledge, user-friendliness, fee, and flexibility, but all link the line capacity with the safety of traffic operation. For an extensive overview of the rail optimization and simulation models, see [29].

2.5 Graphical method

While respecting the stability of the timetable, a graphical method for easier capacity estimation is based on the existing analytical evaluation methodology applied in the Slovak railways and supplement limitations of the UIC 406 method. In this case, the capacity consumption can be expressed in precise absolute terms considering the traffic diversity, calculation periods, and line sections. Travel time and operational time intervals (station and track intervals) of each blocking section on a single or double railway line are considered part of the occupation time. Their surface approximation can be transferred to the time, separated by free space between their graphical approximation in a timetable as the free time between adjacent train paths. In this case, performing compression by simulation tools is not required [30, 31].

3. RAILWAY CAPACITY IMPROVEMENT OVERVIEW

According to [32], a good capacity planning process should be able to generate and evaluate the possible expansion alternatives with an appropriate capacity evaluation model and to suggest an optimal network capacity expansion plan. Planning of capacity enhancement varies between strategic, tactical, and operational levels depending on time consumption, complexity, and related costs. The railway capacity improvement methods and the derived expansion models are applied to solve the lack of capacity of identified critical locations or to design a completely new railway network. This necessarily involves various analyses, like the determination of the existing capacity, demand analysis and cost-benefit analysis. In fact, according to [29] the reconstruction, estimation and prediction of travel demand are the key factors that should be taken into account when assessing the transport systems. If a certain predictive capacity or forecast of transportation volumes are required for further planning, an aggregated or disaggregated

travel demand estimation should be considered. The cost-benefit analysis allows the identification and monetization of overall impacts when it comes to deciding on the desirability of the project realization regarding future advantages and disadvantages. The improvement measures implemented in these projects range from more direct, low-cost and short-term *soft measures* to longer-term, expensive and more complex *hard measures* [33]. Another frequent division of measures consists of upgrading or expanding the infrastructure through construction works, and others enforceable at the operational level [8, 27]. In such possibilities, organizational changes are suggested for the sake of economy and faster implementation.

3.1 Infrastructural approaches for capacity improvement

Some of the measures which imply changes in infrastructure are modifications in track and station layouts, such as construction or increase in the number or length of platforms and sidings, then duplication of the existing sections, changes in the track gradient and horizontal curves, upgrading tracks to accommodate higher axle loads, increasing line speed, changes in the signalling arrangement or installation of modern signalling systems, building flyovers, tunnels and other capacity enhancing structures [5, 8, 27, 33].

The literature on capacity enhancement with infrastructure changes offers various methods and models to support the strategic decision-making. They are intended mainly to increase the theoretical or operational capacity of individual infrastructure segments (single and/or double lines or stations) and collectively the entire networks. It is most often a question of looking for a more suitable variant of increasing the capacity or positioning. The analytical approach developed in [34] allows the assessment of alternative station layouts based on their quantified track complexities. The most suitable layout choice is one that maximizes the theoretical capacity of the station, regardless of train sequence or timetable, thus favouring the planning of stations of different sizes. In [35] the author compared the effects of track duplications and sub-divisions. The track duplication proved to be better and cheaper for the capacity increase. The capacity expansion in [36] involves static and time-varying models to identify the best sections to duplicate based on maximum capacity and to plan these activities over

time. Additional sensitivity analysis helps to determine the maximum number of duplications in each section. The improvement alternatives in [32] have been evaluated for each section based on the reduction of the block length and changes in the station layout or both. Afterwards, the investment selection model determines where and how to upgrade the infrastructure respecting the budget and total costs for each scenario and future demand. This model has been adopted by the Taiwan Railways Administration (TRA) for capacity evaluation and planning. Very similar framework [27] assists the North American Freight Railroads with their optimal capacity upgrade projects. The capacity improvement of identified subdivisions was determined using capital investment, delay, and operating costs as criteria. Moreover, prioritization of different capacity improvement solutions in [37] is carried through the amount of individually absorbed network demand, capacity increase, and overall benefits. The construction of new infrastructure with further capacity improvements proved to be a better option than without additional improvements. As seen, in addition to the amount of increased capacity, the supplementary selection criteria to identify the optimal choice are usually demand, operation and cost-oriented. Cost segments are related to capital investment and/or operating expenses, neglecting maintenance.

3.2 Maintenance activities for capacity improvement

Capacity planning is strongly associated with the maintenance planning. Continuous maintenance activities contribute to the improvement of railway infrastructure performance, provided with time supplements. While predictiveness of *preventive maintenance* enables its planned realization, this is not the case with *corrective maintenance* with a repair philosophy after the occurrence of the problems. This can disrupt the regularity of traffic flow and stability of a timetable if the predicted time reserves are exceeded.

One example of railway capacity enhancement as a consequence of the maintenance improvement framework is given in [38]. The capacity situation on the network can be improved by reducing the track possession times on detected bottlenecks due to the planned and unplanned maintenance activities with appropriate intervention measures. A mixed-integer optimization model suggested in [39] integrates the

scheduling of network maintenance and rail traffic, making it suitable for long-term tactical capacity planning. The authors in [40] proposed a data-driven model for maintenance decision support and validated it in a case study of railway signalling systems. The model integrates various parameters of corrective maintenance data to identify the systems that are more likely to fail, the causes of their failure and the most common corrective actions. Continuous improvement of operational capacity and service quality on the considered line section of the Swedish network and proposing prioritisations in the future maintenance planning are addressed in [41]. The authors explored a maintenance analysis method (risk matrix) for the identification of critical bottleneck zones of operational capacity. They adapted the criticality analysis method to create a hierarchical improvement list.

3.3 Operational arrangements for capacity improvement

Operational arrangements intended to increase the capacity include closing down less busy stations or opening new ones, rolling stock changes (modern vehicles, increase in length), improving logistics (sharing maintenance facilities among operators, expanding the track yard and marshalling facilities), modifying the timetable or the operating procedures (optimal train scheduling and rescheduling), intelligent traffic management and several others [5, 8, 27, 33]. The most frequently used operational measures assume the arrangements in timetable and signalling. The adjustment of train service to match the demand through a timetable directly affects the traffic operations and interactions in the railway network. Due to this strong correlation between the design of the railway infrastructure and the demand for expected traffic, timetabling often precedes the infrastructure planning and underlies numerous capacity assessment methods. Understanding the timetable structures is necessary to obtain a more extensive impression of the capacity fluctuations.

The European railway traffic always fallows predefined conflict-free timetables. However, practical railway operation is prone to numerous initiators of delay, train conflicts and other disruptions whose prolongation is rarely associated with the initial location. To counterbalance, the literature suggests timetable rescheduling or more reliable management like predictions and prevention of delay propagation, changes in time supplements, train

allocation and many other delay absorbers. Still, these actions mainly increase the reliability in train services, the robustness of the scheduled traffic flow and capacity utilization, which is why creating the right balance is essential. For example, an increase of the network capacity in [42] is achieved by aligning the real-time rescheduling with train traffic control. This way, schedule reserve times are decreased to a maximum with the same level of stability. The authors in [7] analyzed the correlations between the topology of service, timetable and performance of signalling systems. The empirical results showed that the reduction of dwell times to homogenize train speeds, forming a group of trains with similar speeds and implementation of shorter block lengths contribute to capacity increase. The shorter block sections are mostly related to headway reduction and changes in the signalling systems.

4. SIGNALLING AND CAPACITY CORRELATION

Signalling changes are often considered the most effective way of improving line capacity because they are more straightforward and less expensive than improving other aspects of infrastructure. This improvement can be performed as a combination of hard and soft capacity improvement measures at a tactical or strategic level, highly related to a long lifespan of such systems [33]. As mentioned previously, the length of block sections, a number of signalling aspects and the possibility of train protection characterize the functional performance of the signalling systems.

Among the three operational levels within the ERTMS/ETCS configurations, ETCS Level 1 (L1) and Level 2 (L2) share the philosophy for conventional signalling with trackside signals. ETCS L1 offers overlay for the existing signalling with additional protection from ATP in discrete points. On the other hand, ETCS L2 represents a cab signalling system with wireless transmission of movement authority, continuous speed supervision linked with a dynamic braking system and an optional dependence on external signals. ETCS Level 3 (L3) is a moving block signalling system without classical blocks and aspects, which combines the benefits of continuous speed supervision and train integrity monitoring system. The minimal distance between the two trains depends on the speed and braking distance of the second train [3]. Today, the hybrid ETCS L3 concept with virtual blocks is also being

increasingly mentioned. Tracking the position of trains relies on a smaller amount of track detection equipment but it is still in the testing phase without more concrete evidence of capacity effects than the classic ETCS L3 type.

One way to meet the demand and transport more passengers and goods is by increasing the frequency of trains, which is directly dependent on the headway limitations of the signalling systems. As this leads to more trains on the same line, their running times, energy consumption and braking distances increase while operating speed declines. Therefore, when designing the signalling systems, the headway and average train speed should be adequately adjusted [43]. Meeting the target traffic volume under different configurations of signalling systems requires tuning the traffic mix with infrastructural characteristics. For the same level of capacity on a corridor with heterogeneous train operations, the moving block technology requires fewer double-track sections than three-aspect and four-aspect signalling systems, providing better resistance to train delays [44]. The capacitive advantages of the moving block are continuously growing with the increase in the amount of double-track railway sections but to the detriment of infrastructure savings. So, in addition to affecting rail performance, signalling planning requires the consideration of time and financial complexity. Still, when determining the investment strategies for railway signalling, relevant importance of benefits or lifecycle costs rests on the decision-makers and varies by concrete investment environment [45]. The overall costs of modern signalling components are related to the development, investment, operation, and maintenance, which in addition to infrastructure need to be incorporated into vehicles. Particularly modern cab signalling compared with conventional signalling has lower costs because of the reduction in trackside equipment. There are very few data about ERTMS costs with no proper estimation. The practice has shown that retrospective installation of ERTMS on the existing lines, although somewhat more complicated, is more expensive than implementing on the new ones [46]. In addition to the adaptation of interlockings and other ETCS-related costs, there are also costs associated with the installation and maintenance of GSM-R [46, 47]. This implies their importance in selecting the lines to be equipped when planning for the implementation of cab ETCS signalling, which requires a GSM-R connection.

4.1 GSM-R capacity impacts

Although it is communication technology, GSM-R also indirectly affects the capacity. Many authors pointed out its obsolescence and potential constraints in use [48] and showed a significant impact of communication delays and packet losses on the possible track utilization under ETCS L3 [47]. Others demonstrated that GSM-R communication failures do not have a severe impact on the capacity of a line with ETCS L3 [49]. According to them, this is appropriate for ETCS L2 because a braking phase on the track determines the minimum headway time. Still, encouraged by rapid progress in commercial mobile communication technology, a recently conducted research [50–54] has changed the focus on new potential technologies that could replace GSM-R in the future. Currently, a Future Railway Mobile Communication System (FRMCS) based on 5G technology has the highest potential to succeed the GSM-R.

4.2 ETCS capacity impacts

With all that said, it is not surprising that there is a lot of literature on the effects of ETCS levels on capacity, from which the most relevant is shown here. Overlaying the ETCS L2 system over the existing conventional configurations for four-aspect signalling, as the best practice system, on the sub-urban section in the United Kingdom (UK) showed minor capacity gains [55]. Similarly, the research on a double-track bottleneck connecting the Stockholm Central Station to the South Station in Sweden showed no effects on the capacity level for ETCS L2 in comparison with the current signalling if calculated by simulation, as opposed to the used analytical approaches. Nevertheless, all methods showed an increase compared to ETCS L3 [56]. The difference often comes because ETCS configurations have higher safety requirements to protect high-speed train traffic than the existing signalling systems, with defined block sizes following the national operating rules of the countries mentioned. Also, the capacity differences between the two systems may not be noticeable if they share similar characteristics, such as continuous data transmission or cab signalling. For the expected impact on the capacity increase, new blocking lengths for ETCS configurations must be appropriately adjusted. Regarding the importance of track layout sizing in stations when upgrading signalling, in [57], the

authors evaluated the capacity benefits for hybrid (i.e. ETCS L2) and moving-block (i.e. ETCS L3) systems over the existing three-aspect signalling system with fixed blocks. Four-station operational scenarios were investigated for the considered signalling options on the busiest double-track corridor with mixed traffic in the Taiwan Railways Administration. At first, the moving-block system offered, on average, 80% more benefits than the current signalling but showed to be significantly constrained from the station track layouts. Sensitivity analysis confirmed that the resulting capacity benefits additionally decrease with an increase in the braking safety factor and communication delay.

Moreover, by observing the operational performance during a traffic disturbance, the homogeneity in the timetable contributes to the capacity increase. In [12], the authors investigated several combined effects of optimal train scheduling for ETCS L2 and the Dutch legacy signalling system in the scheduled and disturbed conditions. For the same timetable order, the dynamic infrastructure occupation of ETCS L2 was lower than the scheduled one because of the more homogenized speed profiles of rescheduling solutions. Additionally, the performance of ETCS L2 with optimization-based rescheduling was much higher than with simple traffic management rules. The performance results of ETCS L2 with a fixed block and ETCS L3 with a moving block on three routes with different traffic densities and train characteristics showed small differences during disturbance [58]. Still, ETCS L2 performed slightly better than ETCS L3 if both use fixed blocks due to the differences in route and train characteristics. This led to the suggestion that ETCS L2 system should be for station areas and ETCS L3 for free tracks.

In addition to comparing the capacity benefits with performance in different conditions and infrastructure occupations, some studies consider additional indicators to justify the ETCS technology against the existing ones. This is the case with the Signalling Equivalent Units (SEU) in [59]. It contributes to evaluating the detected benefits and potential risks of upgrading from current multi-aspect signalling to ETCS L2 and L3 in the UK. In the performed analysis, a wide range of possible implementation options were included, and what is of interest is the application on regional lines. This differs from other research mainly oriented towards major corridors or detected bottlenecks on them.

Regional routes bounded by long sections of absolute block showed a noticeable increase of capacity with ETCS L3.

Other frequently present criteria for the evaluation of different signalling systems are energy efficiency and operating costs. In [60], the authors used the principle of optimal driving strategy to maximize the energy savings and estimate the capacity for faster and easier deciding whether to apply 500 m long sections or moving blocks. According to them, if some signalling option affects the increase of journey time for passengers, this upgrade should be evaluated against the reduction in costs associated with energy-saving. A similar evaluation based on railway operating costs related to optimal driving strategies with minimal energy consumption under both disturbed and undisturbed traffic conditions was addressed in [61]. The performance was additionally evaluated by variation in added rolling stock, in the crew, in generalized cost, in terms of waiting times, in train running times, in energy consumption and variation in energy consumption in disturbed conditions. The energy savings were considered a part of the running time reserve derived from the running time difference between the simulated ETCS levels. ETCS Level 0 (lines without ETCS) and L1 with stochastic disturbances had higher energy consumption without resistance to disturbance, unlike ETCS L2 and L3 driving profiles with minimized increase. Additionally, to optimize the signalling layouts in [62], the considered option effects were evaluated with primary causes of train conflicts, capacity reduction, energy consumption and total costs over the entire system lifecycle. In that way the optimization of signalling layouts or other parameters related to infrastructural and operational components can also be addressed from an economic point of view while meeting the technical requirements for capacity, safety, and topology. More recent research [63] used backwards calculation to compute the optimal length of block sections for ETCS L2 while minimizing total costs. The signalling layout is iteratively adjusted according to the rules of signal engineers, the expected contracted headway, and the behaviour results of different trains in real operational conditions obtained from the simulation. This ensures the lowest amount of signalling equipment required for the desired operational performance, thus minimizing the overall costs.

When analyzing the methodologies used for capacity determination (*Table 1*), different microscopic modelling environments are the dominant choice. Detail representation of real specific infrastructure, traffic and signalling specifics using simulation methods offers accurate impacts on the overall system performances in normal and perturbed conditions. Nevertheless, the sensitivity of simulation approaches to the quantity and quality of input data can be highlighted as their greatest danger. Detailed input data are often difficult to obtain, are of questionable accuracy or in various formats, which complicates the subsequent processing. If not used independently, optimization methods of operational aspects on the network-level further shape the results to increase the capacity and simultaneously minimize costs, or to optimally adjust speed and headway on track level for better energy efficiency of traffic. Thus, in the foreground is no longer the realization of the highest capacitive gains for a given signalling system, but rather optimal results to the additionally set goals.

A comparison of the capacity differences is performed according to the performance of the existing signalling systems or performance among other ETCS configurations. The most evident reasons for capacity gains arise from the changes in *transmission of signal terms to the driver* (trackside or directly in the cab) and the *type of blocking system* (fixed or moving) *with optional length adjusting*. The objects of research of the presented studies are mainly double-track suburban and urban railway lines. Their configuration and operational characteristics are

significantly different from the regional lines that also face a lack of capacity, especially in the conditions of two-way traffic on the same track. Furthermore, the effects of traffic demand on the capacity oscillations with various ETCS configurations has not been sufficiently valorized. Equally, no precise recommendations are mentioned as to what amount of the capacity gain justifies the installation of ETCS technology on a line or part of a network. Such identified need to increase the capacity based on actual traffic demand, traffic conditions or age of the equipment on these lines would facilitate evaluation and balance with the economic and energetic perspective of the ETCS installation.

5. CONCLUSION

Increased transport demand and safety, environmental sustainability, ageing and other significant challenges impose clear guidelines for shaping the future rail mobility while competing with other modes of transport. One of the burning constraints to the smooth functioning of the rail traffic is the lack of available capacity of the existing infrastructure. Solutions can be seen in proper capacity planning by monitoring the utilization of railway capacity and better channelling investments at targeted infrastructure segments through the combined effects of operational or infrastructure arrangements.

This paper provides an extensive survey of the five most prominent types of methodologies for capacity determination and applicable approaches

Table 1 – Methodologies used in the presented research

Reference	Analytical method	Optimization method	Simulation method
Barter [55]	/	/	Unnamed tool
Nelldal et al. [56]	Strele method	UIC 406	RailSys
Lai and Wang [57]	TRA capacity model, sensitivity analysis	/	/
Goverde et al. [12]	/	UIC 406, AGLIBRARY software	Train dispatching system ROMA, Monte Carlo simulation setup
Koning [58]	/	/	Unnamed tool
Ramdas et al. [59]	/	/	Parsons' Railway Integrated Modelling Environment (PRIME)
Dunbar et al. [60]	/	Bruce Force algorithm, fitness function, constrained solution method	Multi Train Simulator (MTS), BRaVE
Corapi et al. [61]	/	/	Unnamed tool
Quaglietta [62]	/	Heuristic "black-box" optimization algorithm OptQuest/Multi-Start	EGTRAIN
Vignali et al. [63]	Blocking Time Model	/	OpenTrack

from the literature for its improvement. The focus has been placed on the capacity enhancement through signalling systems and today the world dominant configurations of the ETCS system. It contributes to an increase of safety, interoperability, energy efficiency and traffic quality beyond national borders, offering better performance in both normal and disturbed operating conditions. Planning the railway capacity enhancement with ETCS technology on the existing lines proved to be requiring additional changes like resizing track layouts with fixed or moving operational features. To access these effects, the literature in this field mostly suggests different simulation approaches combined with optimization algorithms. However, if this is possible, it is better to strive for a model with a lower level of detail and a smaller set of input data to make the solutions more widely applicable, faster to calculate, and less susceptible to the classical dangers of simulation methods. Valorizing the capacity improvements with ETCS technology at such combined micro-, meso- and macro-levels of observation requires further research on the possibilities of more appropriate integrations among the available methodologies for addressing the railway capacity. The preferred solutions are supplementary defined as the least financially demanding, the most energy-efficient and/or the most resistant to delays. This proves the change of motivation in capacity management models that result in an optimal capacity increase within the mentioned constraints. Likewise, more attention should be devoted to exploring the conditions for achieving the capacity enhancement with ETCS on regional lines. Equipping them with ETCS usually has no priority, but due to similar issues also requires concrete solutions, an incentive in migration and adequate preparation for future challenges.

MATEA MIKULČIĆ, mag.ing.traff.[1]
E-mail: matea.mikulcic@fpz.unizg.hr
Prof. dr. sc. **TOMISLAV JOSIP MLINARIĆ**[1]
E-mail: tomislav.mlinaric@fpz.unizg.hr
[1] Sveučilište u Zagrebu, Fakultet prometnih znanosti
Zavod za željeznički promet
Vukelićeva 4, 10000 Zagreb, Hrvatska

UNAPREĐENJE ŽELJEZNIČKOG KAPACITETA MODERNIM SIGNALNIM SUSTAVIMA – PREGLED LITERATURE

SAŽETAK

U vremenu sve veće osviještenosti o zaštiti okoliša i potenciranja prednosti modalne raspodjele, željeznički sektor s učinkovitom upotrebom imovine te odgovarajućim investicijskim planiranjem ima najveće šanse udovoljiti korisničkim očekivanjima i djelotvornije privući nove korisnike. Kontinuirano povećanje željezničke potražnje dovodi do povećanog iskorištenja željezničke infrastrukture i neizbježnog nedostatka kapaciteta, gorućeg problema s kojim se stalno suočavaju mnoge nacionalne željeznice. Kako bi se navedeno učinkovitije adresiralo, u ovome se radu daje osvrt na dostupne metodologije za utvrđivanje željezničkog kapaciteta i tehnike za njegovo unapređenje iz recentne znanstvene literature. Posebna se pozornost daje mogućnosti povećanja željezničkog kapaciteta primjenom signalnih sustava i uvođenju Europskog sustava kontrole vlakova (ETCS). Identificirane su najbitnije poveznice sa segmentima postojećih istraživanja te u skladu s tim sugerirane smjernice za potencijalan nastavak istraživanja.

KLJUČNE RIJEČI

signalni sustavi; željeznički kapacitet; metodologije za utvrđivanje kapaciteta; unapređenje kapaciteta.

REFERENCES

[1] European Environment Agency. *Transitions towards a more sustainable mobility system, TERM 2016: Transport indicators tracking progress towards environmental targets in Europe.* EEA Report: No. 34/2016, 2016.

[2] European Commission. *Report from the Commission to the European Parliament and the Council, Sixth report on monitoring development of the rail market.* Brussels; 2019. Available from: https://ec.europa.eu/transport/sites/transport/files/staff_working_document_-_6th_rmms_report.pdf [Accessed 18 April 2020].

[3] Theeg G, Vlasenko S. *Railway Signalling & Interlocking – international compendium.* Hamburg, Germany: Eurail press; 2009.

[4] Krueger H. Parametric modelling in rail capacity planning. *Proceedings of the 1999 Winter Simulation Conference Simulation – A Bridge to the Future, 5-8 Dec 1999, Phoenix, AZ, USA*; 1999. p. 1194-1200. DOI: 10.1109/WSC.1999.816840

[5] Abril M, Barber F, Ingolotti L, Salido MA, Tormos P, Lova A. An assessment of railway capacity. *Transportation Research Part E: Logistics and Transportation Review.* 2008;44(5): 774-806. DOI: 10.1016/j.tre.2007.04.001

[6] Union Internationale des Chemins de Fer. *Liens entre capacite des infrastructures ferroviaires et qualite de l'exploitation.* Fiche 405 R., 1996.

[7] Dicembre A, Ricci S. Railway traffic on high density urban corridors: Capacity, signalling and timetable. *Journal of Rail Transport Planning & Management.* 2011;1(2): 59-68. DOI: 10.1016/j.jrtpm.2011.11.001

[8] Lindfeldt A. *Railway capacity analysis: Methods for simulation and evaluation of timetables, delays and infrastructure.* PhD thesis. KTH Royal Institute of Technology; 2015.

[9] Kontaxi E, Ricci S. *Techniques and methodologies for*

carrying capacity evaluation: Comparative analysis and integration perspectives. Ingegneria Ferroviaria, ed. CIFI; 2009. p. 1051-1080.

[10] Kontaxi E, Ricci S. Railway capacity analysis: Methodological framework and harmonization perspectives. In: Viegas J, Macario R. (Eds.) Proceedings of the 12th World Conference on Transportation Research. World Conference on Transport Research Society, 11-15 July 2010, Lisbon, Portugal; 2010. p. 1-21.

[11] Pachl J. Railway Operation and Control. Mountlake Terrace, WA: VTD Rail Publishing; 2002.

[12] Goverde RMP, Corman F, Ariano AD. Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. Journal of Rail Transport Planning & Management. 2013;3(3): p. 78-94. DOI: 10.1016/j.jrtpm.2013.12.001

[13] International Union of Railways. Code 406R. (2nd edition); 2013.

[14] Hansen I, Pachl J. Railway Timetabling and Operations. Analysis, Modelling, Optimization, Simulation, Performance Evaluation. 2nd ed. Eurailpress; 2014.

[15] Pouryousef H, Lautala P, White T. Railroad capacity tools and methodologies in the U.S. and Europe. Journal of Modern Transportation. 2015;23(1): 30-42. DOI: 10.1007/s40534-015-0069-z

[16] Sameni MK, Landex A, Preston JM. Developing the UIC 406 method for capacity analysis. Proceedings for 4th International Seminar on Railway Operations Research. Rome, Italy; 2011. p. 1-19.

[17] Sameni MK, Dingler MH, Preston JM, Barkan CPL. Profit-generating capacity for a freight railroad. Transportation Research Board 90th Annual Meeting, Washington, DC; 2011.

[18] Lai Y-C, Barkan CPL. An enhanced parametric railway capacity evaluation tool (RCET). Journal of the Transportation Research Board. 2117;2009: 33-40. DOI: 10.3141/2117-05

[19] Sogin SL, Barkan CP. Railroad capacity analysis. In: Lautala P, Nelson D. (eds.) Railroad Engineering Education Symposium (REES), 11-13 June 2012, Overland Park, KS; 2012.

[20] Vakhtel S. Rechnerunterstützte analytische Ermittlung der Kapazität von Eisenbahnnetzen. Rheinisch-Westfalische Technische Hochschule Aachen; 2002. German

[21] Weik N, Niebel N, Nießen N. Capacity analysis of railway lines in Germany: A rigorous discussion of the queueing based approach. Journal of Rail Transport Planning & Management. 2016;9: 99-115. DOI: 10.1016/j.jrtpm.2016.06.001

[22] Schwanhäußer W. Die Bemessung der Pufferzeiten im Fahrplangefüge der Eisenbahn. RWTH Aachen University; 1974. German

[23] Weik N, Nießen N. A quasi-birth-and-death process approach for integrated capacity and reliability modeling of railway systems. Journal of Rail Transport Planning & Management. 2017;7(3): 114-26. DOI: 10.1016/j.jrtpm.2017.06.001

[24] Deutsche Bundesbahn. Richtlinien fur die Ermittlung der Leistungsfahigkeit von Fahrstrassenknoten. 1979. German

[25] Rotoli F, Malavasi G, Stefano R. Complex railway systems: capacity and utilization of interconnected networks. European Transport Research Review. 2016;8(4). DOI: 10.1007/s12544-016-0216-6

[26] Landex A, Kaas AH, Schittenhelm BH, Schneider-Tilli J. Evaluation of railway capacity. Proceedings of Traffic Days. 2006. p. 1-23.

[27] Lai Y-C, Barkan CPL. Comprehensive Decision Support Framework for Strategic Railway Capacity Planning. Journal of Transportation Engineering. 2011;137(10): 738-749. DOI: 10.1061/(ASCE)TE.1943-5436.0000248

[28] Mitra S, Tolliver D, Mitra S, Bachkar K, Kayabas P. Estimation of Railroad Capacity Using Parametric Methods. Journal of the Transportation Research Forum. 2010;49(2): 111-126. DOI: 10.5399/osu/jtrf.49.2.4503

[29] Botte M, D'Acierno L. Dispatching and Rescheduling Tasks and Their Interactions with Travel Demand and the Energy Domain: Models and Algorithms. Urban Rail Transit. 2018;4(4): 163-197. DOI: 10.1007/s40864-018-0090-8

[30] Gašparík J, Abramović B, Halás M. New graphical approach to railway infrastructure capacity analysis. Promet – Traffic&Transportation. 2015;27(4): 283-290.

[31] Gašparík J, Abramović B, Zitrický V. Research on Dependences of Railway Infrastructure Capacity. Tehnicki vjesnik - Technical Gazette. 2018;25(4).

[32] Lai YC, Shih MC, Jong JC. Railway capacity model and decision support process for strategic capacity planning. Transportation Research Record: Journal of the Transportation Research Board. 2010;2197: 18-28. DOI: 10.3141/2197-03

[33] Sameni MK, Preston JM, Armstrong J. Capacity Assessment in Railway Networks. In: Proceedings of the 2010 Joint Rail Conference (JRC), 27-29 April 2010, Urbana, Illinois, USA; 2010.

[34] Jovanovic P, Pavlović N, Belošević I, Milinković S. Graph Coloring-Based Approach for Railway Station Design Analysis and Capacity Determination. European Journal of Operational Research. 2020;287(1): 348-360. DOI: 10.1016/j.ejor.2020.04.057

[35] Burdett RL. Optimization models for expanding a railway's theoretical capacity. European Journal of Operational Research. 2016;251(3): 783-797. DOI: 10.1016/j.ejor.2015.12.033

[36] Bevrani B, Burdett RL, Yarlagadda P. A case study of the Iranian national railway and its absolute capacity expansion using analytical models. Transport. 2017;32(4): 398-414. DOI: 10.3846/16484142.2015.1099053

[37] Shafiepour M, Tamannaei M, Abtahi SM. A Methodology to Prioritize the Construction Projects of New Railway Infrastructures for Privatization in Railway Networks (Case Study: Iran). International Journal of Transportation Engineering. 2018;6(2).

[38] Famurewa SM, Asplund M, Rantatalo M, Kumar U. Maintenance improvement: An opportunity for railway infrastructure capacity enhancement; 2015. Available from: http://fudinfo.trafikverket.se/fudinfoexternwebb/Publikationer/Publikationer_002001_002100/Publikation_002086/IHHA2013_Final.pdf [Accessed 20 July 2019].

[39] Lidén T, Joborn M. An optimization model for integrated planning of railway traffic and network maintenance. Transportation Research Part C:

Emerging Technologies. 2017;74: 327-347. DOI: 10.1016/ j.trc.2016.11.016

[40] Morant A, Larsson-Kråik PO, Kumar U. Data-driven model for maintenance decision support: A case study of railway signalling systems. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit.* 2016;230(1): 220-234. DOI: 10.1177/0954409714533680

[41] Famurewa SM, Asplund M, Rantatalo M, Parida A, Kumar U. Maintenance analysis for continuous improvement of railway infrastructure performance. *Structure and Infrastructure Engineering.Maintenance, Management, Life-Cycle Design and Performance.* 2015;11(7): 957-969. DOI: 10.1080/15732479.2014.921929

[42] Lüthi M, Nash A, Weidmann U, Laube F, Wüst R. Increasing Railway Capacity and Reliability through Integrated Real-Time Rescheduling. In: *Proceedings of the 11th World Conference on Transport Research, 24-28 June 2007, Berkeley, USA*; 2007.

[43] Ke BR, Lin CL, Lai CW. Optimization of train-speed trajectory and control for mass rapid transit systems. *Control Engineering Practice.* 2011;19(7): 675-687. DOI: 10.1016/j.conengprac.2011.03.003

[44] Dick CT, Mussanov D, Evans LE, Roscoe GS, Chang TY. Relative Capacity and Performance of Fixed- and Moving-Block Control Systems on North American Freight Railway Lines and Shared Passenger Corridors. *Transportation Research Record: Journal of 34 the Transportation Research Board.* 2019;2673(5): 250-261. DOI: 10.1177/0361198119841852

[45] Gutsche K, Scheier B, Jäger B. RAILONOMICS ®-determining investment strategies for railway signalling through simulation. *WIT Transactions on the Built Environment.* 2008;103: 65-74. DOI: 10.2495/CR080071

[46] Mendez-Villalon A, Greedy S, Thomas DWP. Robustness study of ZigBee networks in an EM environment for railway signalling systems. In: *2016 IEEE International Conference on Intelligent Rail Transportation, ICIRT 2016*; 2016. p. 55-60. DOI: 10.1109/ ICIRT.2016.7588550

[47] Zimmermann A. Towards modeling and evaluation of ETCS real-time communication and operation. *Journal of Systems and Software.* 2005;77: 47-54. DOI: 10.1016/j.jss.2003.12.039

[48] Thomas P, Fisher D, Sheikh F. Evaluation of the capacity limitations and suitability of the European Traffic Management System to support Automatic Train Operation on Main Line Applications. *WIT Transactions on the Built Environment.* 2008;103: 193-202. DOI: 10.2495/ CR080201

[49] Jansen DN, Klabes SG, Wendler E. The impact of GSM-R on railway capacity. WIT *Transactions on State of the Art in Science and Engineering.* 2010;46: 143-151. DOI: 10.2495/CR080241

[50] Sniady A, Soler J. LTE for railways: Impact on performance of ETCS railway signaling. *IEEE Vehicular Technology Magazine.* 2014;9(2): 69-77. DOI: 10.1109/ MVT.2014.2310572

[51] Sniady A, Soler J. Capacity gain with an alternative LTE railway communication network. In: *7th International Workshop on Communication Technologies for Vehicles,*

Nets4Cars-Fall 2014; 2014. p. 54-58. DOI: 10.1109/Net-s4CarsFall.2014.7000913

[52] Gonzalez-Plaza A, Garcia Loygorri J, Val I, Arriola A, Rodriguez PM, Jimenez F, Briso C. 5G communications in high speed and metropolitan railways. In: *11th European Conference on Antennas and Propagation, EUCAP 2017*; 2017. p. 658-660.

[53] Samra M, Sidahmed S, Greedy S, Mndez A. Feasibility analysis of wireless technologies for railway signalling systems. In: *2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT), 23-25 August 2016, Birmingham, UK*; IEEE; 2016. p. 1-6. DOI: 10.1109/ICIRT.2016.7725793

[54] Sniady A. *Communication Technologies Support to Railway Infrastructure and Operations.* PhD thesis. Technical University of Denmark; 2015. DOI: 10.11581/ DTU:00000010

[55] Barter WAM. ERTMS Level 2: Effect on capacity compared with "best practice" conventional signalling. *WIT Transactions on the Built Environment.* 2008; p. 213-222.

[56] Nelldal, BL, Magnarini, M, Bruno, DC. Evaluation of ETCS on railway capacity in congested area: A case study within the network of Stockholm. In: Impastato S, Ricci S. (eds.) *II congres on Rail safety and operation: Innovation and new challenges in rail systems, 18 February 2011, Rome, Italy*; 2011. p. 223-232.

[57] Lai Y-C, Wang S-H. Development of Analytical Capacity Models for Conventional Railways with Advanced Signaling Systems. *Journal of Transportation Engineering.* 2012;138(7): 961-974. DOI: 10.1061/(ASCE)TE.1943-5436.0000388

[58] Koning JA. Comparing the performance of ERTMS level 2 fixed block and ERTMS level 3 moving block signalling systems using simulation techniques. In: Allan J, Hill RJ, Brebbia CA, Sone S. (eds.) *Computers in Railways VIII.* 2002;61: 43-52.

[59] Ramdas V, Bradbury T, Denniss S, Chapman D, Bloomfiled R, Fisher D. *ERTMS Level 3 Risks and Benefits to UK Railways.* Final Report CPR798. Crowthorne; 2010.

[60] Dunbar R, Roberts C, Zhao N. A tool for the rapid selection of a railway signalling strategy to implement train control optimization for energy saving. *Journal of Rail Transport Planning and Management.* 2017;7(4): 224-244. DOI: 10.1016/j.jrtpm.2017.09.002

[61] Corapi G, Sanzari D, De Martinis V, D'Acierno L, Montella B. A simulation-based approach for evaluating train operating costs under different signalling systems. *WIT Transactions on the Built Environment.* 2013;130: 149-161. DOI: 10.2495/UT130121

[62] Quaglietta E. A simulation-based approach for the optimal design of signalling block layout in railway networks. *Simulation Modelling Practice and Theory.* 2014;46: 4-24. DOI: 10.1016/j.simpat.2013.11.006

[63] Vignali V, Cuppi F, Lantieri C, Dimola N, Galasso T, Rapagna L. A methodology for the design of sections block length on ETCS L2 railway networks. *Journal of Rail Transport Planning & Management.* 2020;13. DOI: 10.1016/j.jrtpm.2019.100160