Enhancement of Blocking-time Theory to Represent Future Interlocking Architectures

Thorsten Büker [a], Thomas Graffagnino [b],
Eike Hennig [a 1], Alexander Kuckelberg [a]
[a] VIA Consulting & Development GmbH
Römerstr. 50, 52064 Aachen, Germany
[1] E-mail: e.hennig@via-con.de, Phone: +49 (241) 463 662-28
[b] Schweizerische Bundesbahnen SBB
Hilfikerstr. 3, 3000 Bern 65

Abstract
Infrastructure managers around Europe are facing two major topics within the next years: The large-scale renewal of command and control systems and the need to increase capacity. Both topics can be addressed with the further development of ETCS Level 3 and ATO in combination with the introduction of new types of interlockings.
Accompanying the further development of command and control systems there is a need to enhance the blocking-time theory by defining new time components. With this development, effects on capacity can be identified and feasible capacity gains can be evaluated. The enhanced blocking-time theory needs to be implemented into standard railway software tools, which can be quite challenging, due to a shift in paradigms compared to all current command and control systems.
Within this paper, experiences gained in previous studies regarding the necessary blocking-time theory enhancements, the implementation challenges and exemplary capacity gains are outlined. Based on this topics for further research and standardization are defined.

Keywords
Blocking-time theory, ETCS Level 3, ATO, Capacity assessment

1 Introduction

The most common means to express the capacity consumption per train movement is the blocking-time approach. It has been introduced in the 50[ies] and has been standardised for a broader set of applicants at the beginning of the 21[st] century. With the emergence of ETCS, efforts have been made for an appropriate extension of the blocking-time model. Ensuring a precise representation of the capacity consumption per occupation element (usually track-clearance section) the model enables conflict-free timetabling for all types of (mixed) signalling system as well as various forms of capacity assessments and simulations studies.

Recently, railway-infrastructure managers and the supplier industries have launched major programs with the aim to revise – or even reinvent – the overall interlocking architectures plus adjacent systems and operational principles. Representatives of such initiatives are “smartrail 4.0” (Website smartrail) in Switzerland and “Digitale Schiene Deutschland” (Website DSD) in Germany. Those programs are backed-up by a set of motivations:

- Existing command and control technology is overaged or becomes outdated,
- Skills to maintain technology get lost due to demographic aging,

- Applied technology is expensive and does not allow any further capacity gain.
- System conception and architecture don't allow to make full use of actual technical capabilities

While those programs stated above imply a severe redesign of architecture in a mid-term horizon, nationwide efforts to rollout ETCS and upgrade interlockings as well as traffic-management systems (TMS) can be considered as intermediate step in a shorter horizon. For instance, the Norwegian approach (Website "Norwegian ERTMS Program") can be considered as an example. In all cases, we see a common set of actions with varying emphases:

- Replacement of (relay-) interlockings,
- Clear separation of interlocking and traffic-management layers, implying a reallocation of safe and unsafe properties,
- Revised principles to prove operational safety,
- Introduction of ETCS, usually beyond standard capabilities of Level 2,
- Usage of Automatic Train Operation (ATO) in different grades of automation,
- Partial shift from railway-specific solutions to industry standards (e. g. GSM, GPS).

Since capacity improvement is a core target of all programs, there is a severe need to express the capacity impact of the related system architecture. Such quantifications serve the broad portfolio from political decision processes ("what will be the gain?") to detailed requirement specifications ("how shall it ideally look like?").

This article contributes to the enhancement of blocking-time theory with the aim of representing the impact of the aforementioned technologies in existing principles of railway operations research (e. g. simulation, queueing). The text is setup as follows: Paragraph 2 gives a brief summary of the development of blocking-time theory so far, before paragraph 3 describes its enhancement to cope with future situations. Since related computations are usually performed by specific tools, paragraph 4 spots on implementation aspects of those necessary enhancements. Afterwards we raise attention and give insights on chances and obstacles to be taken into account when introducing future interlocking architectures in paragraph 5. In paragraph 6 conclusions are drawn and we summarize where further work needs to be done within the research community.

2 Blocking-time Theory so far

The blocking-time model has been introduced in the 50ies by HAPPEL (Happel (1959)) and in independently by ADLER a couple of years later. Implementation of the model in practical railway scheduling required several decades to pass, until computer-based scheduling systems became available. Standardisation for a broad audience took place at latest with (Hansen et al (2008)). With the emergence of ETCS, efforts have been made for an appropriate extension of the blocking-time model (Büker and Kuckelberg (2013)). Hereafter, all basics already described in (Hansen et al (2008)) are only outlined when required and the focus is laid on the evolution of the model to meet new/future architecture needs.

The blocking time is the total elapsed time a section of track, which is allocated exclusively to a train movement (and thus blocked for any other train movements). The track section may correspond to a whole block section but it may also be a subpart, for instance a route portion requested by two crossing routes.

The blocking time starts as soon as the preparations to issue a train its Movement Authority (MA) demand for exclusive occupation of a route's element. The MA must be issued before the train reaches a location at which a missing MA might cause a deviation from its scheduled train path since braking is triggered. (This corresponds to the principle of conflict-free timetabling.) The blocking time ends after the train has completely left the section and all signalling components have been reset to normal position, if needed, so that another MA with their involvement can be issued. Thus, the blocking time of a track section is usually much longer than the time the train occupies the section.

The blocking time does not embrace the time-demand to process a route request being issued by the train either via trackside equipment or by radio, since this time span does not yet require an exclusive occupation of track sections. Hindrances to train movements due to a wrong processing and provision order of route requests have to be handled outside of the blocking-time model.

Hereafter we denote such supervision systems as Automatic Train Control (ATC), which ensure continuous speed/distance supervision and provide continuous data transmission. All other supervision systems are classified as Automatic Train Protection (ATP). In this metric, ETCS Level 2/3 belongs to ATC while ETCS Level 1 (FS and LS) is an ATP.

Furthermore, we differentiate between safety distance and overlap beyond a signal. The safety distance exists physically and is bordered by an insulated train-joint or an axle counter (the danger point may be located at this border or even further away from the signal). It has to be cleared before a route to the signal can be setup. Often, but not always, the end of the safety distance corresponds to the Supervised Location (SVL). An overlap is limited by the same means but is longer than the safety distance and merely exists temporarily. Overlaps are usually installed in combination to route towards exit signals.

In the following paragraphs, we introduce aspects of the blocking time components with differentiation by signalling systems. In contrast to signalling-system specific definitions, the wording is chosen in a generic manner to allow usage for all types of usage. Whenever purposeful and known, practical implications beyond standard literature are mentioned.

Conventional Lineside Signalling

In Figure 1 the components of the blocking time in case of conventional lineside signalling are visualised. In the example, the brake initiation point to ensure standstill at the main signal is located ahead of the distant signal. In consequence, also the approaching time starts before the train passes the distant signal. Depending on the local configuration, the start of the approaching time and passage of the distant signal may also be at the same location.

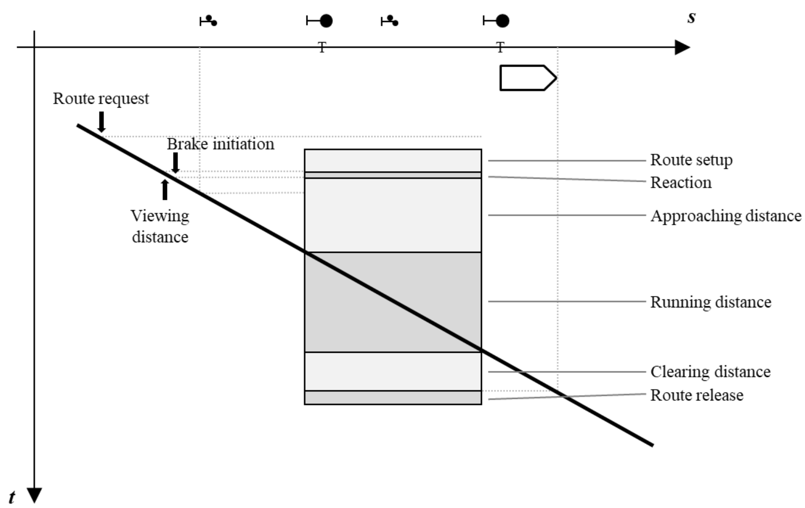

Figure 1: Blocking time for a block section

In Table 1 additional remarks on the components are stated if necessary.

Table 1: Selected time components in one/two-section signalling

Time component	Remark
Route setup	Preparation of the MA has to start sufficiently early that the signal aspect changes at latest with the start of the reaction time. The preparation covers moving switches, locking route elements, commanding the signal aspect. If multiple interlockings (IXL) are involved, their synchronisation cycles have to be taken into account, too.
Reaction	A reaction offset to interpret the distant signal aspect is granted to the train driver. It may be defined as a time or as a distance (in correspondence to minimum sighting distances). In case of a scheduled stop ahead of the track section the reaction time has to be replaced by the time demand for the departure process, which may usually be triggered just after the opening of the exit signal.
Approaching distance	May also start ahead of distant signal at brake initiation point

In layouts as sketched above, the approaching time is always related to the last distant signal in approach to the section. It may be either a separate distant signal (one-section signalling) or a combined main/distant signal (two-section signalling). If in two-section signalling the braking distance takes more than one block section, three-section signalling has to be applied granting two block sections for braking.

There are various principles to realise multi-section signalling as stated by (Pachl

(2005)). From the viewpoint of the blocking-time model, all have in common that a train-specific relationship between the start of the block section and the related distant signal has to be setup. Table 2 gives an overview of the impact on blocking-time theory.

Table 2: Specific time components in three-section signalling

Time component	**Remark**
Route setup	If signalling principle requires commanding signals to black aspect, feedback time to IXL needs to be taken into account.
Approaching distance	Starts multiple blocks ahead of the investigated track section, computation requires a logic to denote the relevant start

ATP Cab Signalling

In case of cab signalling the overall principles remain very similar. Merely the approaching time is determined by the time the train runs through the indication distance that is signalled by the cab signal system. The start of indication distance goes along with a change in the driver-machine display (DMI) indicating to expect the end of the current MA. In Figure 2 the indication distance in case of ETCS Target-Speed Monitoring (TSM) is illustrated. As soon as the train reaches the speed-distance function denoted by „I“ the Movement Authority has to be extended to guarantee hindrance-free operation. The diagram is based on UNISIG Subset 026-3 but enriched by a visualisation of SBD/SBI1 principles. In the given parametrisation, SBI2 is decisive for Warning, Permitted and Indication Curve, anyway. A separate Guidance Curve replaces the Permitted Curve if not inhibited by National Values.

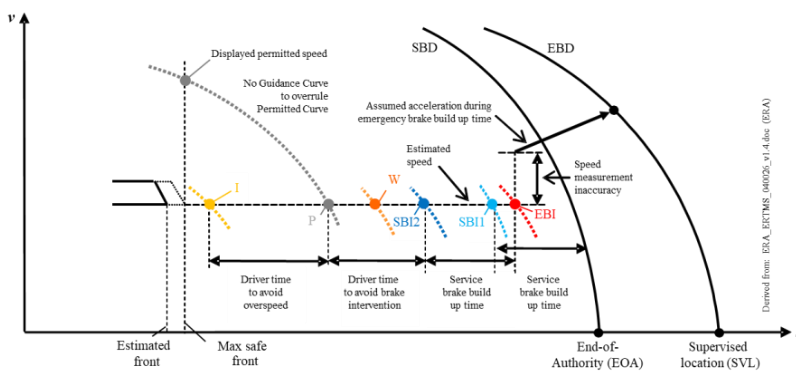

Figure 2: Indication distance in case of ETCS Full Supervision

Since ATP systems rely on dis-continuous data transmission, the MA has to be updated at latest at the transmission spot (e. g. balise group) which is closest to the start of the indication distance. Table 3 summarizes the major differences to the previously discussed principles with the nomenclature being related to ETCS. The mechanism and time components introduced above are applicable to comparable Class B systems (e. g. EBICAB, TBL2).

Table 3: Specific time components in ATP cab signalling (here ETCS Level 1)

Time component	**Remark**
Route setup	If cab signalling is operated in combination to dark signals, feed-back time to IXL after command dark aspect needs to be taken into account.
MA creation	MA is derived either by Lineside Electronic Unit (LEU) from signal aspect or, in case of centralised LEU, from state of route elements and provided to trackside balise group (BG).
MA transmission	Usually via air-gap between BG and trainborne balise antenna
MA interpretation	By ETCS on-board unit (ETCS-OB)
Reaction	Indication distance usually covers a reaction time. Nonetheless, an additional in-advance reaction time may be granted to the train driver.
Odometer error front	Display of indication curve is triggered under consideration of the distance run since the last reset of the train-borne odometry, which requires a continuous adaption of the EBI curve.
Approaching distance	Starts at the transmission spot which is closest to the start of the indication distance

ATC Cab Signalling with Fixed Blocks

For continuous data transmission in case of ATC cab signalling, there is no need to consider the transmission spot closest to the start of the indication distance. (Ideally, most trackside balise are of passive nature.) Instead, the start of the indication distance equals the start of approaching distance. This way, balise engineering is simplified and approaching times are minimal per train movement. In contrast to the ATP case, additional blocking-time components have to be accounted for. They are listed in Table 4.

Table 4: Specific time components in ATC cab signalling (here ETCS Level 2)

Time component	**Remark**
MA creation	MA is derived by Radio Block Centre (RBC) from IXL data, usually status of switches and main signals.
MA request	Applicability depends on RBC architecture. While in early ETCS Level 2 implementations the RBC transmits the MA as soon as created ("push"), the RBC waits for the train's request in newer implementations ("pull"). Requesting starts in time offset T_MAR before reaching the indication distance and then hap-pens cyclic. In worst case, the whole cycle time has to be consid-ered ahead MA transmission.
MA transmission	Via GSM-R (today) from RBC to ETCS-OB
Approaching distance	Starts at the indication distance

Again, the mechanisms and time components described above are transferable to comparable class B systems (e. g. LZB, TVM). In recent ETCS Level 2 implementations, e. g. at Gotthard base tunnel, the Train Position Report (TPR) is used to facilitate the release of the safety distance – but not the block section – if certain conditions on the train speed are met. Additional blocking-time components have to be taken into account as enumerated in Table 5. Attention is drawn that an application of TPR is applicable only in combination to train integrity detection.

Table 5: Specific time components in recent ETCS Level 2 implementations

Time component	**Remark**
Clearing distance	The clearing distance (safety distance) does not necessarily be bordered by trackside equipment if TPR is applied to validate liberation of the safety distance.
Odometer error back	As for the Max Safe Front End in TSM, also the outermost (virtual) train end needs to be accounted for, once TPR is used.
Train position report	Position reports are sent out by the ETCS-OB in fixed cycles of T_CYCLOC or they are triggered periodically after passing D_CYCLOC (being related to the train front).

If TPR is in use, it may also serve the release of overlaps once the train has come to standstill (or at least underruns a threshold speed) instead of linking the release of overlaps to IXL-based section timers. If the IXL-based section timers are quite conservative, this may improve the capacity, since the overlap distance can be occupied by another route at an earlier moment.

For both applications of TPR we see the necessity to merge IXL and RBC functionality (or at least the need for a powerful bidirectional interface). Addressing future architectures in the next paragraph, the shift of responsibilities between IXL and RBC is intensified.

3 Enhancement of Blocking-time Theory for New Architectures

There are many different reasons to focus on the further development of IXL- and RBC-architectures. One reason mentioned already above is the necessity to combine IXL-, RBC- and TMS functionalities to guarantee a powerful interaction between these systems. Another reason comes from the economic view: the life-cycle costs of today's control command and signalling (CCS) systems are relatively high and therefore infrastructure managers' aim for their reduction. This shall be achieved by one main principal, namely limiting the outside CCS components with the use of ETCS Level 3. Dynamic block sectioning and ETCS Level 3 form a system where a train route can start and end anywhere with the use of cab-signalling and train integrity inspection. With these improvements neither outside signals nor outside track clearance equipment are mandatory. The related geometric interlockings (GIXL) will evaluate all safety relevant real-time data continuously. Thus, it will be possible to reduce double safety margins used today, to increase the system performance. If GIXLs are combined with a powerful TMS, many functionalities of todays' interlockings can be shifted to the TMS, so that the amount of "SIL 4" functions within the interlocking can be reduced.

One representative of the development of such architecture can be seen with the development of the ETCS-Interlocking (EI) with smartrail 4.0 in Switzerland. The idea behind an EI is to use digitalization to reduce the necessary outside CCS-components by up to 70 % (Grabowski and Schmidt (2018)). The trend to merge at least IXL and RBC can already be seen in ongoing L2 implementations. For the enhancement of blocking time theory, we assume the existence of GIXL like EI and refer to “smartrail 4.0” (Website smartrail).

3.1 New Blocking-time Components

Together with dynamic block sectioning, ETCS Level 3 can be seen as a time-discrete system instead of a distance-discrete system. This causes that the occupation for a train run comes close to a blocking-time band, whereas with all other system designs (such as ETCS Level 2), the occupation will always look like a blocking-time stairway.

In earlier studies, ETCS Level 3 systems have been modelled as blocking-time band (Büker et al (2010)). During further investigations, especially with smartrail 4.0, the assumption has been revised, because the subsystems (GIXL, RBC, TPR) work periodically. One can say that even an ETCS Level 3 system is still discrete, but the time-steps in the blocking-time get more regular, depending on the different system cycle-times. The blocking-time band looks as displayed in Figure 3.

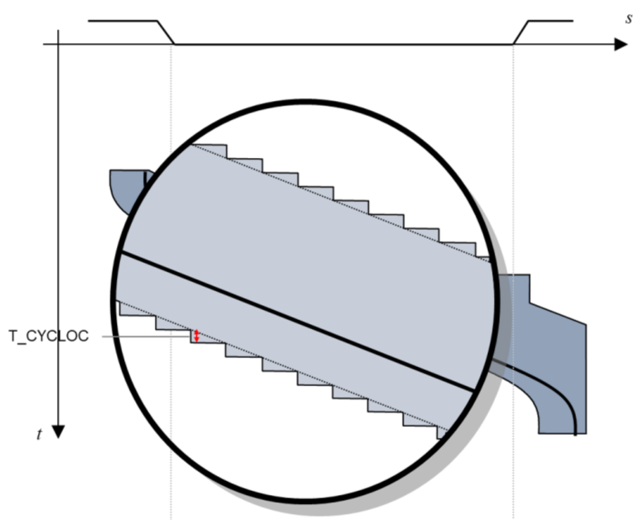

Figure 3: Discrete moving block outside fixed elements like points

In table 6 new time components of a time-discrete system are introduced.

Table 6: Specific blocking-time components in ETCS Level 3 signalling

Time component	Remark
MA creation	As above, but at least cycle time of RBC
MA preoccupation	Optional time/distance buffer to ensure smooth operation
Running distance	The track section melts down to (infinitesimal) short distance, if dynamic block sectioning is applied.
Clearing distance	Each train movements pushes its virtual safety distance ahead.

To provide the whole picture, Table 7 marks the applicable components of blocking times for the whole range of architectures (including CBTC architectures).

Table 7: Blocking-time components including recent/future architectures

Time component	System	Optical signalling, Level 1 LS	Level 1 FS	L2 „push“ with pure track sections	L2 “pull” with partial TPR	L3 „pull“
	Route setup	X	X	X	X	X
System Times Pre	**MA creation**		LEU	RBC	RBC	RBC
System Times Pre	**MA request**				X	X
System Times Pre	**MA preoccupation**					X
System Times Pre	**MA transmission**		Air gap	GSM-R	GSM-R	GSM-R / FRMCS
System Times Pre	**MA interpretation**		X	X	X	X
	Reaction	X	X	X	X	X
	Odometer error front		X	X	X	X
	Approaching distance	Distant signal	Last balise group for MA Extension	Indication Distance	Indication Distance	Indication Distance
	Running distance	X	X	X	X	
	Clearing distance	X	X	X	X	Virtual
	Odometer error back				X	X
System Times Post	**Train Position Report**				X	X
System Times Post	**Route release**	X	X	X	X	X

3.2 Interaction with Automatic Train Operation (ATO)

In the context of the IXL architectures discussed within this paper, Automatic Train Operation (ATO) is mainly envisaged to improve the system performance and reduce energy consumption. It does not primarily contribute to an increase of safety, since it is applied in combination with an Automatic Train Control (ATC). The most popular combination, even though not yet fully standardised, is ATO-over-ETCS. Potential applications of ATO in combination to non SIL-4 systems are out of the scope of this article. Since ATO-over-ETCS according to subsets 125, 126 and 130 seems on its way to industry standard, we hereafter refer to this specific system, whenever needed.

The on-board ATO (ATO-OB) maintains a train run within a defined tolerance of its path following a particular target function (for instance energy consumption). The system marginally adjusts operating parameters such as the ratio of power to coast when moving. Ideally, the path can be readjusted by the TMS to the current situation and transferred by the trackside ATO (ATO-TS) to the ATO-OB. This results in an outer control loop by the TMS and an inner control loop on the rolling stock, as described by (Weidmann et al (2014)), with both components following particular target functions. Communication from ATO-TS to ATO-OB is, in a simplified description, backed-up by static segment profiles and dynamic journey profiles with the latter ones covering the train-path specific timing points.

There are five Grades of Automation (GoA) of trains with GoA 0 being regular on-sight train operation. Right now, two variants are of highest interest:

- GoA 2 means semi-automatic train operation where starting and stopping is automated, but a driver operates the doors, drives the train if needed and handles emergencies.
- GoA 4 means unattended train operation, where starting and stopping, operation of doors and handling of emergencies are fully automated without on-train staff.

In the scope of timetable and capacity modelling, ATO impacts various aspects such as easily inserting new trains to react to unforeseen peak of demand. Some of them, as identified so far, are addressed in the following paragraphs. While the consequences for modelling are mostly elaborated, there is a lack of published study results or even of in-field experiences. At least the outcomes of a study on potential improvements on the suburban "S-Bahn" network in the Stuttgart area are available (Website VM Baden-Württemberg).

Reduction of Approaching Times

In Supervised Speed Envelope Management (SSEM), the on-board-ATO establishes the maximum speed the train can run without interfering with the ETCS speed limits. In TSM this means, the ATO-OB shall drive the train so as not to reach the EBI curve. For this purpose, the ATO-OB (re-)computes various speed-distance functions within the current MA using the information sent by the ETCS-OB. In particular the ATO-OB inhibits the service-brake command being triggered by overpassing an SBI supervision limit as well as any "Sinfo" sounds in relation to speed and distance monitoring. In consequence, the set of ETCS information/intervention curves (cf. Figure 2) melts down to the outermost EBI.

Instead, an appropriate representation of ATO-borne requires consideration of ATO-OB specific time components. As this part of the system behaviour is vendor-specific one has to make assumptions how to represent the properties in a generic manner, which allows fine-tuning as soon as in-field experiences have been collected. According to expert judgement, the following model seems promising:

- By means of maximum brake decelerations, a SBD-equivalent for ATO is defined.
- An ATO service brake build up time allows deriving an SBI-equivalent.
- To avoid ETCS emergency brake intervention, a "buffer time" ensures sufficient computation times for the ATO-OB control loop and missing synchronisation between ETCS-OB braking curves and their ATO-OB replica. Whenever the ATO-SBD injures the ETCS-EBI, it is replaced accordingly.

In Figure 4 the two additional ATO speed-distance functions and their interaction with

ETCS speed-distance functions are visualised. By comparison to Figure 2 it becomes evident which time components lose their relevance for the indication distance.

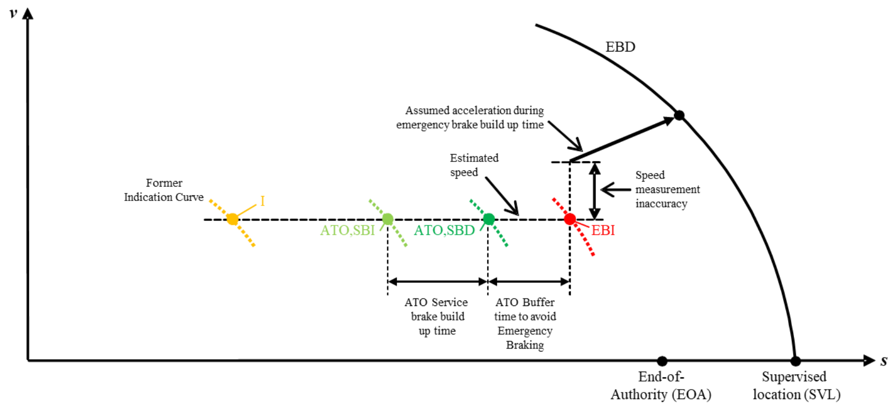

Figure 4: Indication distance in case of ATO-over-ETCS

If ETCS Level 1/2 is implemented as overlay to the signalling layout of a high-capacity Class B system, often an increase of indication distance and thus a loss of capacity emerges. This is due to the multiple convenience and safety margins, which also take effect if the system is configured without service-brake intervention in TSM. By means of ATO-over-ETCS, a considerable share of the convenience margins are suppressed and disadvantages from the ETCS implementation on capacity consumption are mitigated.

Reduction of Regular Supplements

Best practice amongst railway operators is to augment technical/minimum running times by supplements to compensate randomness of technical/minimum running times to an appropriate extent. Such randomness may arise from dwelling time delays, human driving behaviour, weather conditions, track-wheel adhesion or availability of full rolling-stock characteristics. (Running-time extensions because of track works are taken into account separately.) The margins are usually defined as a relative increase of technical/minimum running times or as an additional running time per running distance. In the first metric, values of 3 % to 7 % are common. The resulting running time is referred to as regular running time. It forms the basis for the timetable compilation. As well, the corresponding speed profile serves the computation of approaching times in blocking-time theory.

By means of (at least) semi-automatic driving, at least the stochastic (human) impact of the driver is eliminated from acceleration, coasting and deceleration processes. As long as there is the same journey profile, same (version of the) on-board equipment and the same train/track conditions, the same trajectory shall result thanks to ATO-OB. Furthermore, an automatic train operation can regulate the jerk very accurately, which can reduce the running time thanks to too high acceleration and deceleration. This means less deviation of the actual running times. An example of human-driven trajectories versus ideal trajectories on the Belgian network is provided by (Bienfait et al (2012)). (Unfortunately, speed-distance diagrams of later test-runs on the Brussels – Leuven line with ATO in comparison to drivers are not published. As well, similar experiences from ATO operation on the RATP network are not publicly available.)

Following the logic of regular supplements to eliminate the impact of the largest share of randomness to the timetable, if the new technical systems does not need new time supplements, the magnitude of regular supplements could be reduced. This would allow to:

- Reduce scheduled running times and thus reduce travel times. (It may result in a little increase of capacity consumption because of higher approaching time and because of higher heterogeneity of train paths, anyway.)
- Keep scheduled running times and facilitate the gained supplements for delay reduction whenever needed. This way, stability of the overall timetable concept is increased.

While the mechanisms are known, knowledge on their magnitude still needs to be gathered as only very few main-line systems have been taken into operation with ATO. (For comparability, urban metro systems can only be compared with limitations, as they are often operation CBTC-borne from their early days.) In some countries a specific regular time supplement is added to compensate variations of adhesion and driving style. At a first glance, a reduction of the regular supplements by a third seems appropriate according to expert judgement.

Backward Compatibility
In tools of railway operation research, randomness is taken into account whenever simulations of operation are performed and is being represented. Usually, operation is disturbed by delays at entry and by primary delays at stops. Furthermore, running times may be increased randomly. To ensure a precise representation of knock-on delays, reaction times are incorporated. The general algorithmic representation of railway operations is quite deterministic, nonetheless. As we see above, ATO contributes to the reduction of a part of the stochasticity from railway operation. From system design one can conclude, that reactions times are replaced by transmission times with transmission times being relatively short as ATO is intended to be a non-safe system. Thus, reacceleration after an unscheduled standstill may start earlier than in today's operation.

In the close future, there will be various studies to assess the benefits of ATO on specific infrastructures. To ensure reliable outcomes and a precise differentiation between system behaviour with/without ATO, stochastic properties have to be represented properly. With regard to the mostly deterministic representation of the current non-ATO operation, the challenge within the tools may be rather to create a more random representation of the status quo operation while only fine-tuning to the aforementioned ATO aspects.

3.3 Constraints of Moving Block Application

Applying the moving block principle is subject to various constraints as it has already been shown in (Büker et al (2010)). Depending on the specific infrastructure design, certain sections have to be operated in (virtual) blocks in any case, as trains should not come to standstill for various reasons:

1. Moveable elements (switches, bridges) can only be occupied as a whole.
2. Overhead catenary design (OCS) design does not allow standstill.
3. Initial traction effort is too low to ensure reacceleration.
4. Maximum coupling forces may avoid reacceleration.

While the two latter constraints may rather happen on the open line along steep inclinations, the two first constraints usually take effect in station areas. In consequence, they frequently become decisive for minimum headway times and foil the benefits of the moving block principle as shown by Figure 5:

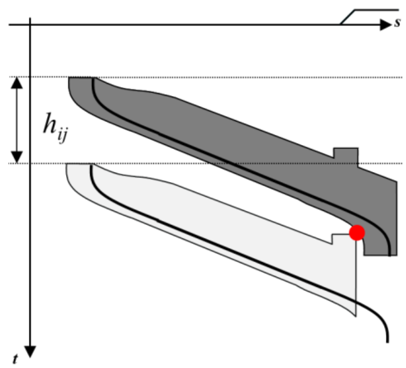

Figure 5: Moving block being interrupted in switchpoint

Virtual blocks may also be caused in station/bifurcation areas to avoid partial occupation of heavily loaded track sections: Let there be a train sequence using the same track beyond a bifurcation with the first train being a suburban train stopping in the vicinity of the bifurcation and the second train being a heavy freight service. With the first train moving on the moving block of the second train consequently occupies the bifurcation as a feature of geometric interlocking. Once the stop of the first train takes significantly longer than expected, any third movement along the bifurcation is excluded.

In conventional block sectioning this effect might be avoided by the static block logic not granting a route to the second train at all. In geometric signalling this – and the implication 2-4 stated above – have to be mitigated by the TMS layer instead.

4 Implementation Aspects

The implementation of ETCS toolboxes or modules for different levels, specifications and purposes itself is a challenge. Starting from scratch, the implementation of the aforementioned features might seem more or less straightforward. Anyway, integrating the new aspects into the traditional approaches for running-time and occupation calculation (including derived functionalities like conflict detection) has to consider existing data structures as well as integration and extension of data-exchange interfaces, backend databases or graphical interfaces and output graphics and diagrams.

4.1 Running Time Computation

The common understanding and basic approach for current microscopic tools realizing detailed running time computations is to start with a (static) speed profile and stopping policy for a single train run. Based on this input, the highest (technical/physical) train speed is determined and the shortest running time is calculated. Additional time margins, regular and specific supplements are applied afterwards, resulting in a (decreased) running speed corresponding to the new running times.

Within a forward-oriented loop, the train speed is increased as long as it remains within speed profiles. Due to traction forces resulting accelerations might be negative ("missing traction power"), but principally these phases intends to determine the highest physical running speed possible.

In contrast, the character of braking phases usually follows a backward-oriented approach: Within the forward oriented determination of running speed the change of profile speeds towards a lower value requires braking phases towards the (new) target speed. Typically, braking phases are determined by (fixed or value-dependent) braking acceleration values. With these values, one braking curve anchored at target speed and location can be computed. Following that curve, the intersection point with the former train run can be identified as braking point. Braking accelerations are usually predefined (theoretical) values representing the "usual" braking behaviour of trains respectively train drivers. Braking phases towards scheduled stops are treated equally with a target speed of zero. Moreover, a "green wave" is assumed usually implying that ATP/ATC influence remains inactive with respect to running time computation.

The introduction of ETCS, especially ETCS Level 3 plus dynamic/moving blocks affects the well-proven running time principles for some reasons:

- Braking acceleration values become much more dynamic due to braking models, national parameters and multiple dependencies from train and traction characteristics.
- A large set of probably overlapping braking curves is computed, the most restrictive curve has to be derived from that set and the resulting curve might be a section-wise partitioning of multiple curves.
- For lowering speed changes the most restrictive curves might become relevant, the static braking acceleration cannot be used any more.
- Most restrictive curve or any derived simplification like static braking accelerations underriding the most restrictive curve is much more complex than any other legacy supervision curve or braking model.

Moreover, especially for ETCS Level 3 that eliminates "traditional" spatial blocking, the semantic of braking processes has shifted towards a "forward oriented" braking computation, imitating the behaviour and decisions on a train driver with respect to DMI visualizations.

4.2 Curve Variations and Dependencies

Depending on the usage of ATO the braking curves might vary, time components have to be considered respectively ignored. The availability of ATO functionality is dependent from train characteristics as well as trackside equipment that moreover might be installed only partially.

Therefore the implementation has to extend the trains by appropriate properties but also the infrastructure model has to be enhanced by information about availability and type of ATO infrastructure, e.g. directed begin and end graph nodes for a microscopic network graph including technical ATO properties. These begin and end elements have to be evaluated for each train passing these elements and the validity and consequences for the train runs respectively running time and occupation computations have to be realized, resulting in dynamic property vectors that have to be considered by the ETCS curve computation. Finally, to ensure that existing ATO functionality is available, the complete braking curve from braking point to EOA has to be within such an ATO enabled area. This might imply a repeated computation if the assumption, that ATO is available fails while computing due to ATO area leaving.

4.3 Occupation-time Computation

The classic blocking-time theory determines braking points based on braking curves with static braking acceleration values. Depending on the train control system used, the point of MA submission is derived from braking points, system or transmitting times are added and begin time of succeeding block sections are derived from these indication points. Similar to running-time computation, the determination of ETCS braking curves has to happen for each possible EOA and for each change to a lower speed.

The quantity of braking curve computations raises in case of ETCS Level 3 with dynamic block sectioning, because EOA becomes approximately continuous which increases the more complex ETCS curve computation extensively. Therefore, the implementation considerations concerning shifting paradigms from backward- to forward-oriented braking computation become much more relevant. Finally, the integration of this reasonable paradigm shift and its integration into existing tools and applications are another implementation problem. Current railways operation functionalities add another implementation complexity dimension, e. g. tools simulating train operation have to handle dynamic aspects like occupied block sections etc. that disturb the ideal world of green wave planning. For train timing, red signal and knock-on-delays the management of complex ETCS braking curves additionally complicates the implementation.

With ETCS Level 3, one new aspect has to be incorporated: Cyclic system times and partial discrete occupation element. While cycle times resulting from system component frequencies discretize the theoretically continuous occupation band in time (cf. paragraph 5.1), e.g. technical times for switch changes moreover contradicts occupation band continuity which has to be considered by the implementation as a new challenge additionally.

Conflict detection has to be modified, too. While former conflict detection evaluates overlapping time ranges for single occupation elements, microscopic conflict detection for ETCS level 3 has to evaluate band overlaps instead, where upper and lower boundaries of occupation bands are derived as location-time regions. A regional overlapping has to be performed for conflict detection instead. The determination of e.g. minimum headway times also has to be adapted if implementing the new approaches. While former times are derived from time buffers between two occupation blocks, the determination of buffer times between occupation bands follows a continuous distance detection approach between two regions in space. From an implementation point of view it might be interesting, which algorithms and probably which granularity of discretization are used to detect overlapping regions.

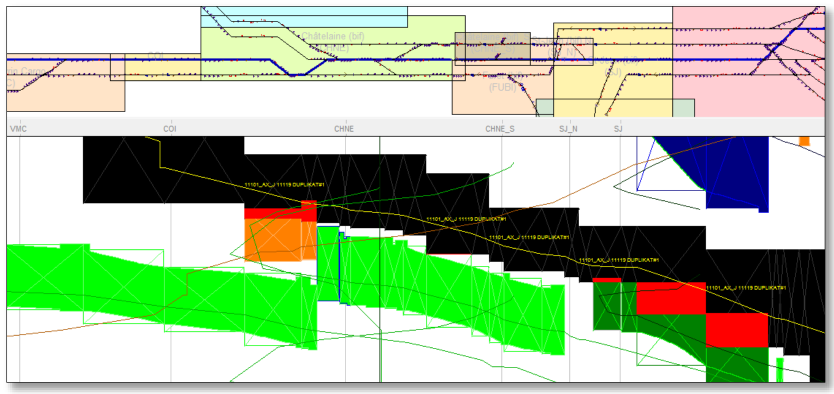

Figure 6: Sample implementation of ETCS Level 3 occupation bands (green & blue) and traditional occupation blocks (black) including conflict detection (red), sectional bounding boxes and discrete switch occupations.

4.4 Occupation Points, Occupation Areas and Model Synchronization

While the two previous paragraphs described pure implementation aspects of running-time and occupation-time computation, more complexity arises when synchronizing different occupation paradigms, e.g. mixing up classical, section based occupation graphs and occupation bands from ETCS Level 3. Integrating the continuous occupation bands into sectional occupation blocks requires some activities:

- (Spatial) segmentation of occupation bands due to corresponding occupation sections.
- Relating band segments as “virtual occupations” to occupation elements.
- Detecting the bounding box for each band segment and setting begin and end times of the associated virtual occupation accordingly.

When segmenting occupation bands, some interesting implementation aspects are:

- Occupation areas around switches are assigned to two occupation blocks, the switch itself and the preceding block.
- The discretization of occupation bands results in more or less detailed saw-tooth bands within occupation blocks.
- Switches have additional technical setup times, e.g. route setting times, that do not move in the same way the ETCS blocks do. Therefore switches can usually be identified easily due to the “blister” shown on top of the occupation band.

Mapping occupation bands in that way, it is possible to study and analyse “mixed scenarios”, e.g. ETCS Level 3 controlled trains and trains operated under conventional train control (fall back or mixture of differently equipped trains). This approach moreover directly fits into existing conflict detection and solving paradigms of succeeding tool functionalities like capacity assessment (analytical approaches, simulations, UIC 406 etc.), therefore the approached presented within this paper can completely be integrated into existing tool implementations and their functionalities.

5 Decisive Effects on Headway Times

As described above, ETCS Level 3 (with absolute braking distance) in combination with GIXL and TMS architecture requires an enhancement of the blocking-time theory taking into account all previously new defined time components. Resulting minimum headway times drop, as expected, but various effects have to be taken into account. For the discussion of decisive effects of an ETCS Level 3 system in combination with geometric interlockings we will focus on the minimum headway times as well as the maximum element occupation time of a set of trains depending on the used CCS.

5.1 Cycle Times

For the horizon of smartrail 4.0 it is assumed that for example driven by autonomous street cars, the localization technology will be affordable and precise enough to be used within the railway system. It is currently aimed for a localization accuracy of 1 m and a fail-safe data transmission (Website smartrail). Since in current ETCS level 2 implementations the TPR runs in cycles of around 5 seconds and needs to be evaluated by the RBC (working in cycles as well), this leads to the steps on the bottom of the moving block, which are shown in Figure 3. For the evaluation of minimum headway times it has to be assumed that in the worst case two trains follow each other with TPR cycles not being synchronized, which will increase the headway time by T_CYCLOC.

Since the TPR runs periodically and requires radio transmission, it has to be evaluated, if the TPR is sufficient to be used in heavily used track topologies for track-clearance detection. A way to avoid problems with the needed timespan would be to use track-clearance equipment (for example axle counters) in dense areas, since they might work faster and use the TPR for track-clearance on the line. To get any major benefits in terms of headway times a TPR-cycle would need to be reduced to about one second (or less), which would result in even smaller vertical steps in the running-time band (manufacturer survey).

5.2 Switchpoints

In legacy IXL a switchpoint is always covered by a signal. This causes, that in the blocking-time theory a switchpoint is already occupied as soon as the approaching distance of the corresponding signal is reached. In case of L3/GIXL, the occupation of switchpoints has to be taken into consideration separately. As stated above, switchable elements can only be occupied as a whole. Speaking in terms of blocking-time theory, the whole length of a switchpoint has to be preoccupied at once, taking into account that already before reaching the approaching distance a time span for the turnaround of a switchpoint is necessary.

In Figure 5 it is visible, that especially switchpoints may become the decisive element for determine headway times, since the occupation of a switchpoint starts earlier (due to switchpoint turnaround time) than the occupation of the track section directly in front of and behind the switchpoint. The capacity effects of a moving block are noticeable limited by switchpoints. The negative effect on capacity could be reduced by different approaches:

1. Use of smaller switches to reduce the occupation length
2. Investigation, if it is possible (and has benefits) to only preoccupy the moveable

parts of a switchpoint instead of the whole switchpoint

3. In advance swinging (by TMS) to avoid turnaround time affecting blocking time
4. Possibility of turning all switches simultaneously instead of one after one
5. Reduction of switchpoints

5.3 Speed Changes

One issue that currently reduces the positive effects of the moving block is the speed supervision in speed changes with ETCS as standardised with Baseline 3. For running-time calculations, the permitted curve (cf. grey curve Figure 2) is used, even though ATO is assumed. A yellow status of the DMI is accepted, thus. In Figure 7 we take a look at a train sequence of two (passenger) trains. It is visible, that the permitted curve reaches the target speed significantly earlier as the speed change is mandatory. This distance is mostly dependent of the trains' braking characteristics. If both trains are following each other with the minimum headway time needed at 140 km/h (in this example 70 seconds), the train sequence of these two trains is not conflict free in the speed change any longer, resulting in an occupation time conflict of 21 seconds. If the headway time is increased up to 91 seconds (second train shifted by 21 seconds), the train sequence would be conflict free again.

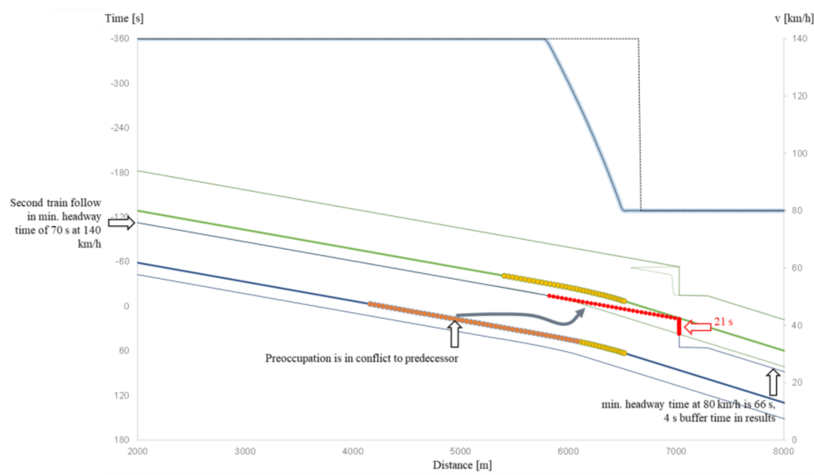

Figure 7: Permitted Curve towards v-Step

The negative influences on the minimum headway time in speed changes can be minimized, if a second speed step is implemented and the whole sequence is run at yellow DMI. This scenario is presented in Figure 8. This does increase the running time of the trains slightly, but reduces the minimum headway time at the same time.

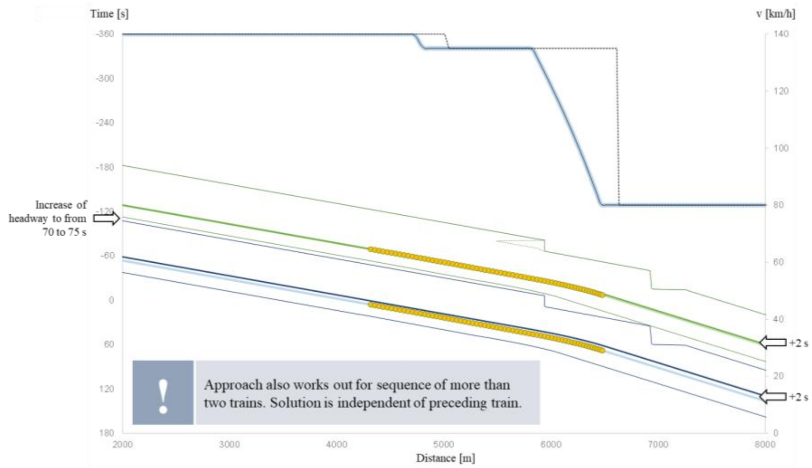

Figure 8: Optimized train sequence in v-Step

The benefit in using two speed changes instead of one is that we preoccupy the section between the two speed changes with the permitted distance and not with the indication distance, which results in a reduction in minimum headway time. In the given example, the minimum headway time between the two trains could be reduced by 16 seconds (from 91 to 75 seconds) with the introduction of the intermediate v-step. The impact of the intermediate v-step on the running time is comparable small (2 seconds) and should be accepted for a reduced minimum headway time. With this approach it is possible to increase capacity (and reduce the minimum headway time) if a speed change is the decisive section for the minimum headway time of two trains following each other directly. Adding more than one intermediate speed change does not reduce the minimum headway time any more since the benefits come mostly from the use of the permitted distance in between the two v-steps.

The issue of increased running time with ETCS Full Supervision (FS) due to more restrictive speed supervision in speed changes is valid for all current ETCS FS systems. It has to be discussed, whether an ATO has to follow the permitted curve in speed changes, or if it is possible to run beyond the permitted curve to reduce the running time again.

5.4 Exemplary Capacity Gains by smartrail (EI)

To evaluate the capacity effects that can be gained in the smartrail setup, we have conducted various analysis, with examples being presented here. In a first analysis we calculate the occupation times for three different existing railway lines (one optimised for freight trains, one for long distance passenger trains and one for local passenger trains) for every main signal on these lines. One major result of the calculation of occupation times is, that with EI in combination with ETCS Level 3 the distribution of occupation times becomes more homogeneous and all in all the occupation times can be reduced. Details can be seen in table 8, being based on the following assumptions:

- ETCS braking model without Service Braking (SRS 3.6.0)
- Preoccupation starts with the indication distance

- Fixed odometry confidence interval
- Running time sufficiently long to pass a diverging switchpoint with the given speed

Table 8: Exemplary distribution of technical occupation times

	Conventional signalling [s]	**EI with L3 and ATO [s]**
5 %	78,9	68,8
25 %	99,7	74,9
50 %	123,7	88,8
75 %	150,1	141,3
95 %	193,1	163,2
Average	128,0	103,5

The enhancement of the blocking time theory is prototypical implemented in LUKS®, which is a software tool for the assessment of capacity on railway networks. With this tool, we are able to validate timetable concepts. For timetable compilation we use the following assumptions:

- Preoccupation: Indication distance
- Running time calculation: permitted curve
- Buffer for operational quality is included in the route clearance time, thus not visible

In Figure 9 a screenshot of a future timetable is given. On the top part, all trains use conventional signalling. There it is visible, that the desired train sequence is not possible without blocking-time conflicts (purple). In the bottom part, we see the same time-slot in new architecture. There it is obvious, that the same timetable does not have blocking-time conflicts anymore. This given example is taken from a current timetable validation project and the same issues can be seen in different locations and constellations.

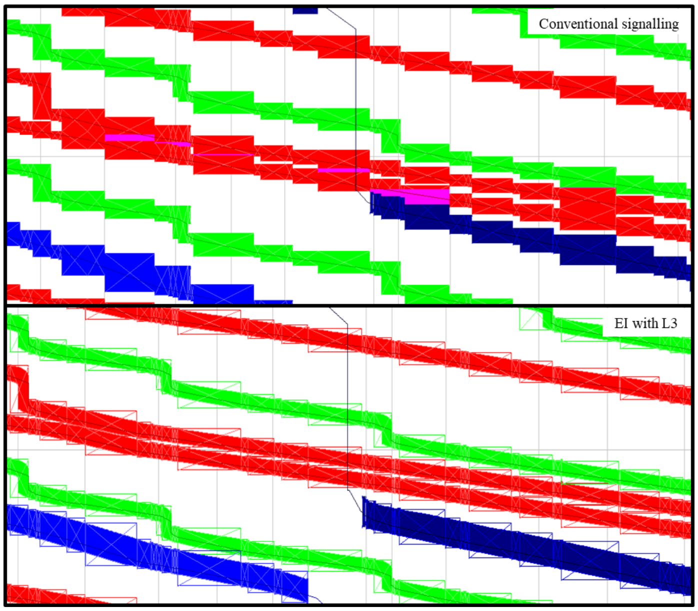

Figure 9: Exemplary time distance diagram with conventional IXL (top) and EI (bottom)

At large, it can be said that EI in combination with ETCS Level 3 can lead to a reduction of minimum headway times, which can increase the capacity of a railway infrastructure. Never the less there are more boundary conditions which determine the overall capacity of a railway infrastructure which cannot all be solved with this new architecture.

6 Conclusion and Future Challenges

Recent studies lead to an enhancement of the blocking-time model to represent future architectures (ATO, GIXL, RBC and TMS). With their introduction, a contribution towards capacity increases on today's railway network can be made – but mostly only, if all subsystems are introduced as a whole.

In case of ETCS, the safety level is adjusted by a combination of distance EOA-SVL, by National Values and by rules how to compute the train's brake capability. Even though the underlying Braking Curve Model is flexible, it cannot be adapted to every situation. This results in a safety surplus in certain situations. A continuous re-examination of the overall situation shall be part of geometric interlocking and the safety distance is virtual in case of ETCS Level 3. So far it has not yet been elaborated, to which extent a continuous adaptation of the safety distance at the desired safety level may serve a reduction of approaching time.

As popularity of ATO-over-ETCS grows, we see further needs for research and development. To our best knowledge, the following things should be addressed in the future:

- Behaviour of ATO-OB for timetable compilation tools and for TMS

- Influence of ATO on system times
- Influence of ATO on regular supplements (less stochastic influences)
- Legal aspects of the use of ATO

After the uncertainties regarding the ATO-specification and -implementation have been resolved and it becomes foreseeable to which extend the assumptions made have been met, it will be possible to integrate especially ATO in software tools for timetable compilation. Before that has been achieved, all assumptions on capacity will be preliminary and we strongly recommend to not overestimate the system by making unlikely assumptions.

References

Bienfait, B.; Zoetardt, P.; Barnard, B.: *Automatic Train Operation: The mandatory improvement for ETCS applications*. At: ASPECT IRSE, London 2012.

Büker, Th.; Kogel, B.; Nießen, N., 2010. *Influence of the European Train Control System (ETCS) on the capacity of nodes*, Paris, May 2010, ISBN 978-2-7461-1801-0.

Büker, Th.; Kuckelberg, A., 2013. *ETCS approximation by legacy ATP/ATC systems under railway operation research methods.* In: Signal+Draht 105(2013) issue 10.

Grabowski, D.; Schmidt, S.: *The "ETCS Interlocking".* Signal+Draht 110(2018) issue 10.

Happel, O., 1959. *Sperrzeiten als Grundlage für die Fahrplankonstruktion*. In: Eisenbahntechnische Rundschau (ETR) 8 (1959) 2, pp. 79-90.

Hansen, I.; Pachl, J., 2008. *Railway Timetable & Traffic: Analysis - Modelling – Simulation*, Eurailpress (2008).

https://www.smartrail40.ch (checked 19th Jan. 2019)

https://www.deutschebahn.com/de/presse/suche_Medienpakete/medienpaket_digitale_schiene_deutschland-1177310 (checked 19th Jan. 2019)

https://www.banenor.no/en/startpage1/Projects/ERTMS-National-Implementation (checked 19th Jan. 2019)

https://vm.baden-wuerttemberg.de/fileadmin/redaktion/m-mvi/intern/Dateien/Praesentationen/181024_Praesentation_Verkehrsausschuss_ETCS_DSTW_Pilotprojekt_Stuttgart.pdf (checked 19th Jan. 2019)

Pachl, J., 2005. *Modelling Specific Signalling Features in Computer-based Scheduling Systems*. In: Rail Delft 2005 Proceedings, International Association of Railway Operations Research.

Weidmann, U., Laumanns, M., Montigel, M., Rao, X, 2014. *Dynamische Kapazitätsoptimierung durch Automatisierung des Bahnbetriebs*. In: Eisenbahn-Revue 12/2014.