Compositional modeling of railway Virtual Coupling with Stochastic Activity Networks

Francesco Flammini [1,2], Stefano Marrone[3], Roberto Nardone[4] and Valeria Vittorini[5]

[1]Mälardalen University, Eskilstuna, Sweden
[2]Linnaeus University, Växjö, Sweden
[3]University of Campania "Luigi Vanvitelli", Caserta, Italy
[4]University of Naples "Parthenope", Naples, Italy
[5]University of Naples Federico II, Naples, Italy

Abstract. The current travel demand in railways requires the adoption of novel approaches and technologies in order to increase network capacity. Virtual Coupling is considered one of the most innovative solutions to increase railway capacity by drastically reducing train headway. The aim of this paper is to provide an approach to investigate the potential of Virtual Coupling in railways by composing stochastic activity networks model templates. The paper starts describing the Virtual Coupling paradigm with a focus on standard European railway traffic controllers. Based on stochastic activity network model templates, we provide an approach to perform quantitative evaluation of capacity increase in reference Virtual Coupling scenarios. The approach can be used to estimate system capacity over a modelled track portion, accounting for the scheduled service as well as possible failures. Due to its modularity, the approach can be extended towards the inclusion of safety model components. The contribution of this paper is a preliminary result of the PERFORMINGRAIL (PERformance-based Formal modelling and Optimal tRaffic Management for movING-block RAILway signalling) project funded by the European Shift2Rail Joint Undertaking.

Keywords: Railway; Performance Modelling; Stochastic Activity Networks; Moving block; Virtual Coupling

1. Introduction

The ever-higher speeds reached by current trains and the consequently reduced travel times attract an increasing number of users and in turn, require increased capacities to the railway network. That is especially critical to avoid bottlenecks in specific railway corridors where multiple lines converge. Railway operators are experimenting with different techniques and solutions to reduce the spaces between two subsequent trains. Among the different available and envisioned solutions to increase efficiency and capacity, cooperative driving seems an extremely promising option for smarter railway transportation. Similar solutions have been successfully experimented in the automotive sector showing many advantages and opportunities in terms of accident reduction, due to reduced

Correspondence to: Francesco Flammini, e-mail: francesco.flammini@mdh.se, francesco.flammini@lnu.se

human control, and infrastructure utilization, due to reduced traffic congestion. The improvement of infrastructure utilization and the consequent increase in network capacity are also two primary objectives addressed by the Shift2Rail Joint Undertaking (JU) for future railways [Shi15]. The European Shift2Rail Joint Undertaking (JU) has been established in 2014 to manage and coordinate all rail-focused research and innovation activities funded under the EU Horizon 2020 framework program. A possible solution to the aforementioned issues addressed in Shift2Rail is the so-called Virtual Coupling that allows trains to run as they were physically joined due to the considerable reduction of train headway.

Virtual Coupling offers a really innovative and somehow visionary scenario for future railways, although several fundamental aspects have to be studied before it can be actually introduced in real systems. With reference to the advantages of Virtual Coupling, it is rather intuitive that closer trains can have a great impact on network capacity. On the other hand, the possible advantages of Virtual Coupling have to be analysed with respect to specific topologies and lines to concretely measure the positive effects on network capacity.

A previous work published in reference [FMN$^+$18] describes a solution to integrate Virtual Coupling in modern and future railway signalling systems, and proposes some directions for ground-breaking research with the ultimate goal to enable safe train platooning. The work presented in this paper develops that research direction and introduces a compositional model-based approach for the quantitative evaluation of network capacity increase considering specific network topologies. The proposed approach is based on Stochastic Activity Networks (SAN) [MMS85], which is a modeling formalism similar to Petri Nets. SANs allow engineers to use C-code in their models. The proposed approach aims at modeling the service level over a given track portion, in order to evaluate the real benefits of Virtual Coupling when used in specific scenarios. The main benefit of the proposed approach is to provide engineers with an effective tool to evaluate the effects of virtual coupling in real-world conditions (e.g., track configuration, traffic volumes, types of rolling stock).

The rest of the paper is organized as follows. Section 2 introduces the necessary background that includes the concepts related to the train control systems and the basics on the SAN formalism. Section 3 introduces Virtual Coupling as a novel operating mode for future railway controllers. Section 4 describes the details about the SAN model, including the model of operating modes of the railway controller as well as an approach to represent the service over a specific railway line. Section 5 describes the results of the analysis conducted over a simple yet realistic case-study, which are useful to validate the model and to evaluate the benefits introduced by Virtual Coupling. Section 6 discusses the related works, while Sect. 7 ends the paper by giving some closing remarks.

2. Background

In this section, we describe the background needed by a non-expert reader to understand how the reference European railway control systems works, together with the basics on the SAN formalism that will be used to build the models.

2.1. ERTMS/ETCS

The European Rail Traffic Management System/European Train Control System (ERTMS/ETCS) is the standard for signalling and interoperability in modern European railways. ERTMS/ETCS ensures both technological compatibility among trans-European railway networks and allows integration of the new signalling systems within existing legacy interlocking systems. A reference architecture for ERTMS/ETCS systems consists of three main subsystems: the onboard system is the core of the control activities located on the train, the line-side system is responsible for providing geographical position information to the on-board subsystem, and the trackside system is in charge of monitoring the movement and spacing of the trains. The ERTMS/ETCS specification identifies three functional levels of growing complexity, named Level 1, Level 2 and Level 3. The specific level depends on the equipment of the line and on how the information is transmitted between the onboard and the line-side systems.

In Level 1, the movement authority (MA) and the corresponding information on the route are transmitted to the train and displayed in the cabin to the driver (cab signalling) in a discontinuous way by means of electronic beacons, named balises, placed between the rails.
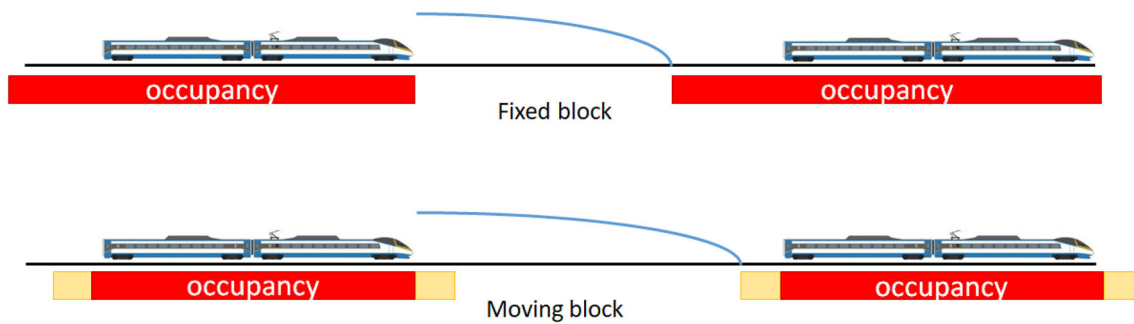
Fig. 1. Railway fixed block and moving block signalling

The task of the balises is to enable a self-localization of a train and to transmit dynamic information about the status of the track ahead. Level 2 allows the management of train distancing through a central controller named Radio Block Center (RBC). The RBC receives information about the status of the line and the trains in order to send updated information to trains regarding their MA (i.e., distance ahead and maximum speed profile). Each train can define its specific speed profile on the basis of its own weight and braking characteristics. ERTMS/ETCS Level 2 is based on the usage of Global System for Mobile Communications—Railway (GSM-R) for both train-to-ground and ground-to-train communications. It uses fixed block signalling with trackside train location equipment such as track circuits for safe train separation. ERTMS can be considered as the combination of ETCS with GSM-R, hence from now on in this paper, we can refer to ERTMS only.

Level 3, which introduces the so-called moving block signalling, is still at an early stage of development, also due to safety validation concerns. In Level 3, lineside equipment such as train detection devices based on track circuits can be removed. It hence overcomes the concept of fixed block sections by introducing the concept of moving block sections representing the actual distance between trains plus some safety margins. The main advantage of Level 3 consists of reducing the distance between trains to the minimum.

Further development of Level 3, with the integration of new wireless communication technologies is supported by the Innovation Program 2 (IP2) of the Shift2Rail Joint Undertaking [Shi15], which aims at a line capacity increase by leveraging on moving block signalling. Such innovation has to ensure "continuity and backward compatibility with the current signalling and supervision systems through ERTMS standards". IP2 also opens the possibility of introducing Virtual Coupling, considered as "a new way of operating trains" by creating multi-train convoys. In Virtual Coupling, trains can be coupled and decoupled dynamically, according to the service needs and in the respect of the safety requirements.

Virtual Coupling represents a further optimization of moving block signalling and hence it makes sense as an add-on of ERTMS Level 3. As shown in Fig. 1, in ERTMS Level 2 (upper part), the train safe headways are computed by the RBC on the basis of the occupation status of the track, i.e., MAs must include one or more free block sections, in turn including one or more track circuits (*fixed block*); in ERTMS Level 3 (lower part), RBC uses the position and integrity information reported by the train to compute MAs (*moving block*). Virtual Coupling requires a highly dependable positioning and communication system due to the stricter time requirements resulting from shorter headways, which also stresses the importance of research on new generation communication and positioning systems.

The absence of trackside train detection equipment has several advantages but also poses some challenges and risks that discourage railway operators to adopt ERTMS Level 3 [FvHAM17]. While virtual coupling could be—in theory—implemented on any moving block signalling systems, when associated to ERTMS it can boost Level 3 performance and therefore its possible adoption due to the full exploitation of line capacity. That is a huge advantage in terms of flexibility and efficiency of railway transportation against other transport modes, especially considering high traffic corridors that have currently reached their capacity limit. Furthermore, the ERTMS standard specification includes a clear reference architecture, solid working principles, advanced Automatic Train Control functionalities and a well-structured language for messages, packets and variables, so that implementation of Virtual Coupling over ERTMS appears more viable compared to completely new developments, as we will explain in more details in the following sections. This is why someone referred to Virtual Coupling over ERTMS as ERTMS Level 4 [Mea16].
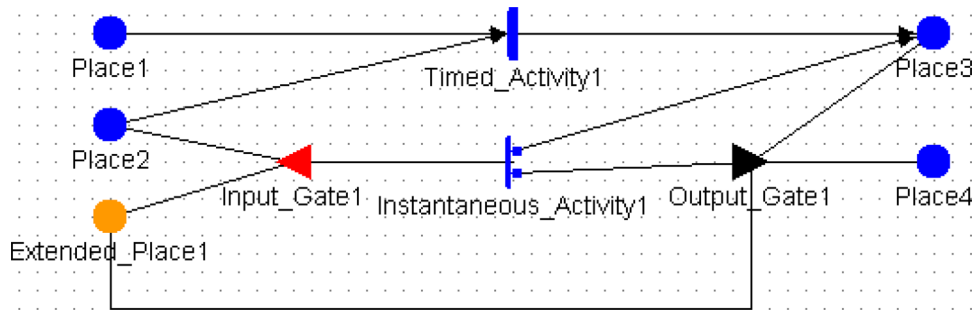
Fig. 2. SAN model—example

2.2. Stochastic Activity Networks

Stochastic Activity Networks (SANs) are a stochastic extension of Petri Nets, which has been introduced to analyze concurrency, timeliness, fault tolerance, and degradable performance of complex computing systems [MMS85]. In the following, we suppose that the reader is acquainted with the basics of Petri Nets.

A SAN model comprises four primitives which define its structural component: places, activities, input gates, and output gates. Places are used to describe possible local system states as in Petri Nets. Activities are called transitions in Petri Net terminology; they can be timed or instantaneous, and are used to describe events that can modify the system state. Gates are introduced in SANs to allow for greater flexibility in defining enabling and completion rules. Timed transitions model time-consuming activities whose completion induces a state change in the system; instantaneous activities model logical conditions, or activities requiring a negligible amount of time to be completed.

The expressive power of SANs is raised by the possibility to associate actions (called *cases*) with the activities that may be taken upon the completion of an event, an *output function* with the output gates, an *enabling predicate* and an *input function* with the input gates.

The temporal specification of SANs is stochastic, it is defined by associating an activity time distribution function with each timed activity and a probability distribution with each set of cases. Hence, cases permit to model the uncertainty about the next state assumed upon completion. Once activated in a specific marking, timed activities can be restarted, either with the same or different distribution, according to what specified by the associated *reactivation function*. The reward structures that relate the possible behaviors of the stochastic process to specified *performance variables* quantify benefits associated with activity completions and number of tokens in places. This means that the reward structure of a SAN allows for a variety of ways of defining different types of performance (reward) variables.

An example of SAN is provided and described in the remainder of this Section. A complete definition and discussion of SANs is given in [SM01].

SANs at a glance. Figure 2 shows the structure of a SAN. It includes: places, drawn as circles; instantaneous transitions, drawn as solid bars; timed transitions, represented by hollow ovals; input and output gates, drawn as triangles. Cases are drawn as small circles on one side of the associated activity.

Special places may be used, i.e., the *extended places*, which differ from standard places for the type of the tokens they may contain. Tokens in the extended places may represent atomic variables, data structures or arrays of primitive data types (i.e., *short*, *int*, *long*, *float*, *double*, *bool* and *char*).

The firing of timed activities is associated with general distributed random variables, such as *Exponential*, *Normal*, *Binomial*, whose parameters can be numeric constant or marking dependent; the probability distribution associated with each case could be specified as numerical constant or a function, as well. If no cases are present, the default is assumed with a probability equal to one.

Input and output gates can be used to control the "enabling" condition of an activity and to change the state of the system when the activity "fires". An activity is enabled when the predicates of all input gates connected to the activity are evaluated as true, and each place connected to incoming arcs contains at least one token. When an activity fires, the input and the output functions of input and output gates (respectively) are executed, while tokens of connected places are updated as in the Petri Net firings.

The tool used throughout the paper to edit and analyze SAN models is Möbius [CGK$^+$09], which also supports the compositional and hierarchical development of models. In Möbius all enabling predicates, input and output functions, parameters, types, and variables are expressed by C++ statements.

In Fig. 2 the timed activity `Timed_Activity1` is enabled by tokens in places `Place1` and `Place2`; when it fires, a new token is added in `Place3`. At the bottom of the figure, the instantaneous activity `Instantaneous_Activity1` is enabled by the predicate of the `Input_Gate1`, which in turn is evaluated on the marking of `Place2` and `Extended_Place1`. When the activity fires, two cases are possible. The first adds a token in `Place3`; the second enables the execution of the output gate `Output_Gate1`, which in turn updates the marking of `Place4` and `Extended_Place1` according to the output function associated with the activity.

3. Architecture and operating modes of future railways

As anticipated in the previous section, ERTMS Level 3 has the potential of increasing significantly the capacity of railway lines. Thanks to moving block signalling, the MA is dynamically computed in a continuous way thus enabling each train to move till the safe rear end of the preceding train. Such a potential represents the main improvement with respect to ERTMS Level 2, where the MA granted to trains has the granularity of the fixed block sections, with a length that can be several kilometres. Hence, ERTMS Level 3 allows more trains to run on a single track in order to increase line capacity. In order to reach its maximum performance while keeping the required level of safety, ERTMS Level 3 needs a dependable communication between the train and the RBC in order to ensure prompt hazard detection and safety monitoring, especially regarding train position and integrity status.

Virtual Coupling over ERTMS Level 3 represents a further optimization as long as each train safely receives sufficient information about the preceding train from the RBC or from the train itself by means of train-to-train communication channels, which is not yet standardized within ERTMS.

Train-to-train communication can be an essential mechanism to enable Virtual Coupling by reducing the risk of missing or delayed vital messages. Based on the experience gained within the automotive research with platooning of vehicles over highways [RG94], possible topologies for railway Virtual Coupling are: "fully connected", in which the RBC communicates with the entire platoon of trains, and each train, in turn, forwards some information to its neighbours; "chain-like", in which the RBC communicates with the first train of the platoon that, in turn, forwards the messages to the following train like in multi-hop/mesh routing. The first solution offers a higher availability with respect to temporary communication loss, either with the RBC or with the preceding train, since there is a second channel through which a train could receive vital information. Moreover, such solution can offer a backward compatibility with the existing control solutions, based on the communication between RBC and trains. Therefore, our proposal is based on such topology.

As shown in Fig. 3, the evolution towards ERTMS Level 3 with Virtual Coupling is possible by having the RBC evaluate position and integrity received by each train (green arrows) and compute the MA of the entire train fleet (blue arrows). In addition, each train can communicate its speed, acceleration and position to the following train (red arrows). The onboard unit can compute its own speed profile, according to the information received by both the RBC and preceding train.

Regarding possible real-world implementation, we assume the presence of an additional "operating mode" corresponding to Full Supervision with Virtual Coupling. The actual use of such operating mode is not essential for the results of the simulations because it can be thought of as a status only (sort of virtual operating mode). In fact, the ERTMS standard [UNI16] defines several operating modes for the onboard European Vital Computer (EVC) according to different conditions affecting the status of the trackside and of the train itself, in order to guarantee the required performance and safety levels. The main one is named *Full Supervision (FS)* mode, where the complete supervision of each train is ensured when all trains and track data required are available. Currently, in FS mode, the entire system is able to operate at maximum performance (i.e., train speed) and safety, thus also ensuring the highest capacity. For degraded operational situations, operating modes are *Shunting*, *On-Sight*, *Staff Responsible*, which rely much more on train drivers and offer limited maximum speed and thus a significantly lower capacity with respect to FS mode. For the sake of modelling abstraction, we group all of them into a single operating mode named *Partial Supervision (PS)*.
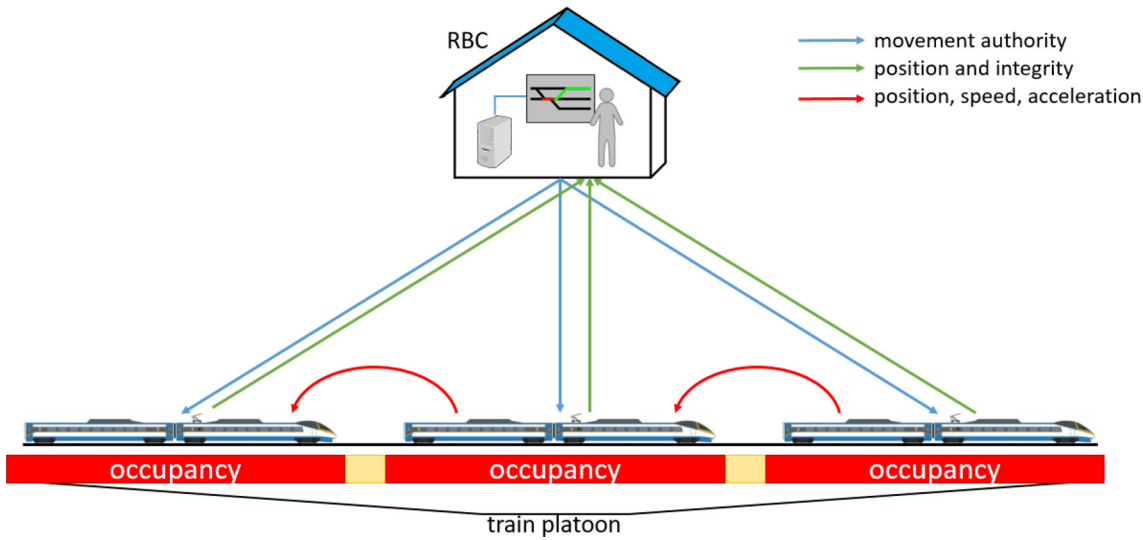
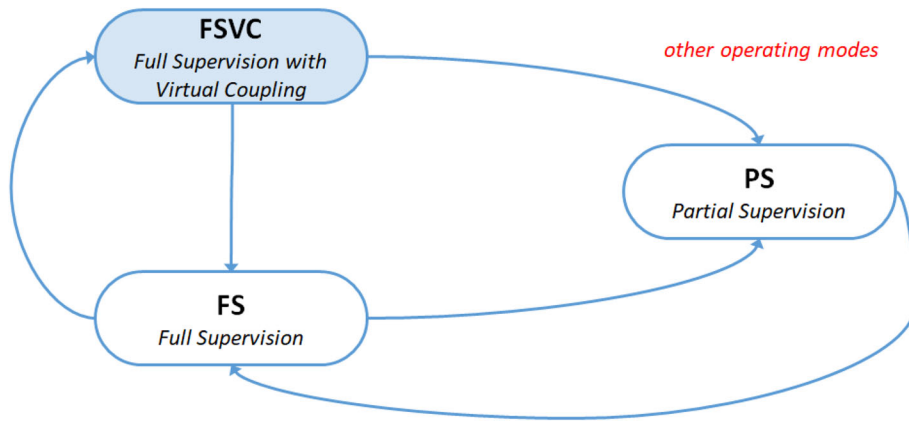**Fig. 3.** Communications needed for Virtual Coupling

Fig. 4. Operating modes relevant for Virtual Coupling

As mentioned above, in this paper we assume a further operating mode, named *Full Supervision with Virtual Coupling (FSVC)*, on top of the FS mode. This operating mode is reachable from the FS mode if the system verifies all required operating conditions. The real operation can be as follows: when the system switches from FS to FSVC, the RBC sends a custom ERTMS message ("coupling" command) to the EVCs, starting the coupling procedure of vehicles in platoons; on receipt of such message, each EVC starts to exchange messages regarding position, speed and acceleration as previously described; the RBC computes the MAs for the entire train convoy and communicates them to the trains (see Fig. 3). In case of different routes, the RBC sends the "decoupling" command to revert back the system in the FS operating mode in order to allow a specific control on MAs. The switch from FSVC to FS might be started by a single EVC, e.g., in case of a temporary loss of a train-to-train communication. In case of major failures, the system could switch directly to PS operating modes. The resulting control state machine is depicted in Fig. 4, where the FSVC mode is highlighted in light blue, showing its connections with the operating modes currently defined by the ERTMS standard.
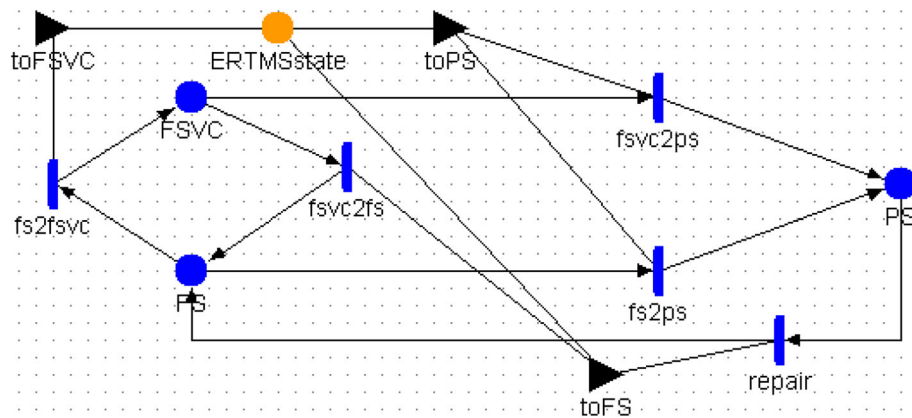
Fig. 5. SAN model for operating modes

4. The SAN model

In order to evaluate the actual advantages provided by Virtual Coupling in future railways, we built a SAN model with the aim of predicting the capacity increase on a specific railway track topology.

The model has been developed as a composition of reusable atomic SAN models. The rationale behind that was the ambition of building model libraries representing specific railway sections, with their own topology and operating conditions. The model library is a collection of basic models that can be manually composed according to different configurations; the current version is publicly available on GitHub.[1] It includes two classes of models: railway control system and railway line. The first class aims at representing the RBC and, specifically, the ERTMS operating modes. The second class represents the physical characteristics of the railroad, including switch points; in addition, there are models representing the incoming train flow of the modeled track portion and models representing the measurement terminal needed to evaluate the line capacity (i.e., endpoints of the modeled track where to measure the number of arriving trains). To analyze a specific railway network, the global model can be obtained by place superposition of atomic SAN models, i.e., by sharing state variables.

The modeling approach consists of splitting the considered line into a sequence of elementary track elements, connected to each other. The introduction of these track elements is only a ploy to model the continuous running of trains as discrete, but they are not related in any way to the track occupancy of trains (known as *virtual block*). Moreover, the shorter is the track element, the higher is the accuracy of the model since the discrete movement is closer to the continuous one. In the model, in fact, each train could occupy multiple elementary track elements at a specific time instant, but it would not be possible for two different trains to occupy the same block. The elementary track elements are connected by means of place superposition. Each train is identified by a non-zero identification number, stored in the extended places of the model. The hypothesis is that the track element is shorter than the minimum distance between two trains. The running of trains along the modeled line can be seen as the movement of a sequence of numbers, each equal to the train id over those track elements.

In the following, we describe atomic SAN models by detailing their parameters as well as their possible composition.

ERTMS operating modes. The transition between operating modes, as explained in the previous section, can be modeled by the SAN model shown in Fig. 5. The three places of the SAN model represent the corresponding operating modes that are: Partial Supervision (PS), Full Supervision (FS), and Full Supervision with Virtual Coupling (FSVC). Stochastic transitions connect those places. The assumption is that these transitions represent system failures (from higher to lower supervision modes) and corresponding restorations (or simply repairs), therefore we used exponential probability distributions. The specific rates are parameters of the model; hence we have four parameters named *FS2FSVC*, *FSVC2FS*, *FS2PS*, and *REPAIR*.
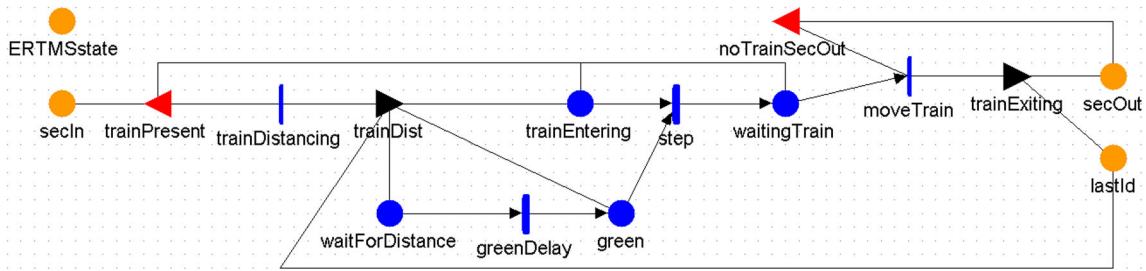
---

[1] https://github.com/RobertoNardone/VC-model.
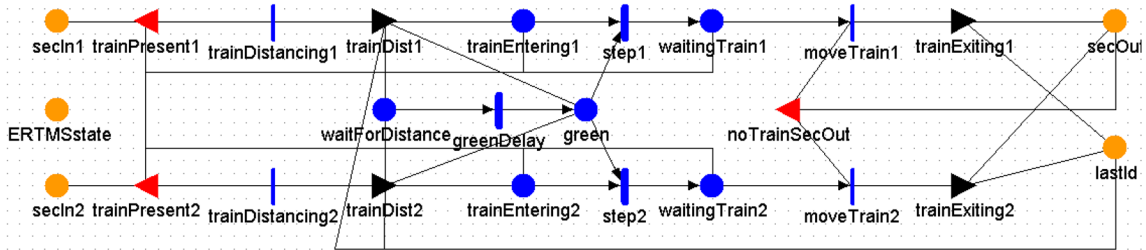
Fig. 6. SAN model for Line—straight

Fig. 7. SAN model for Line—point

The extended place `ERTMSstate` stores over its state variable (typed as short) the current ERTMS operating mode, encoded with the values 1, 2 and 3. This place is used for enabling a simple communication of the operating modes with the other submodels. The initialization code initialises the behaviour, putting a token in the place representing the initial state of the system and encoding the corresponding value in the extended place. At last, associated with the transitions, the different output gates update the value encoded in the extended place according to the entered state.

Line—straight. The first track element represents a straight element with an incoming and an outgoing connection (Fig. 6). Both are represented by extended places, named `secIn` and `secOut`, in order to encode in their state variables the identification number of trains entering and crossing the track element. In fact, when a train portion enters the straight track element, a non-zero id is written on the place `secIn`. This condition, together with the absence of trains (i.e., tokens) inside the element, enables the immediate transition `trainDistancing`, which executes the output gate `trainDist`, aimed at evaluating train separation. More specifically, the output gate `trainDist` evaluates if the incoming train id is different from the last id, in which case it puts a token in place `waitForDistance`, otherwise it puts a token in the place `green`. In the former case, the transition `greenDelay` evaluates the correct train separation, which depends on current operating mode obtained from place `ERTMSstate`, before putting a token in place `green`. In both cases, a token is put in place `trainEntering`. When the green light is obtained, the modelled behavior is the one of a train running along the track element. The firing of the transition `step` is given by a deterministic distribution function, its value being a model parameter.

When the train is ready to exit the track element, a token is set in the place `waitingTrain` in order to wait until the clearance from a preceding train (i.e., a zero value in place `secOut`). The output gate `trainExiting` moves the train id from the place `secIn` to the place `secOut` and updates the place `lastId` with the train identifier.

Line—point. This element represents the SAN model of a switch point, or simply point. As depicted in Fig. 7, this model shows two different incoming connections, named `secIn1` and `secIn2`, and one outgoing connection, named `secOut`. The behaviour of this model is similar to the one described above, with the only difference coming from place `green` that is shared by the two tracks. Hence, when a train enters the track element from one of the two incoming connections, it runs up to the end of this element. As in the previous model, the two values of the deterministic distribution functions associated with the transitions `step1` and `step2` are to be considered as model parameters.
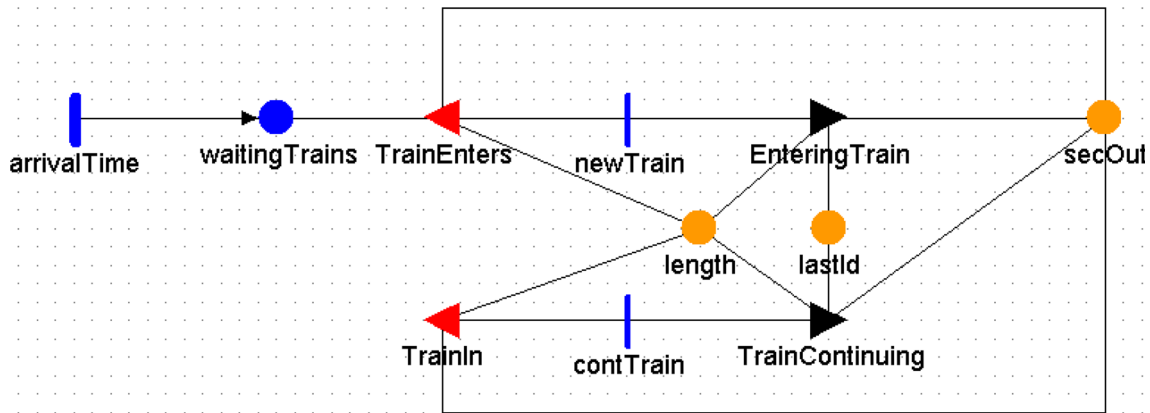
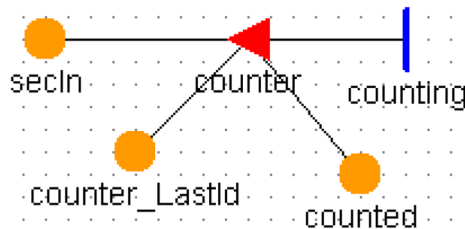**Fig. 8.** SAN model for Injector terminal

Fig. 9. SAN model for Measurement terminal

Line—train injector. The atomic SAN model in Fig. 8 represents the incoming train flow entering the modelled line portion. The transition `arrivalTime` represents the arrival of trains, which are buffered as tokens in the place `waitingTrains`. When the extended place `secOut` is free from trains, the instantaneous transition `newTrain` fires and the output gate `EnteringTrain` inserts a proper numeric identifier in the place.

The lower part of the model is useful for reproducing the last id in place `secOut` for a number of times equal to the length of the train divided by the length of the modeled track element. By doing so, as explained above, each train occupies different elements of the model at the same time, and its id is present in different places.

Line—measuring terminal. In order to evaluate the capacity of a line, we need a SAN model which is able to count the trains arrived at a specific terminal. This model has also to "consume" the numerical identifiers of the trains to enable their movement along the line. This model is depicted in Fig. 9; it has an incoming connection, the place `secIn`, connected by an input gate to the transition `counting`. The firing of this last transition, together with the capability to store the id of the last train arrived in the place `counter_LastId`, enables train counting. Specifically, the value in the place `counted` provides at every instant of time the effective number of trains arrived at the measurement terminal. Thus, the line capacity can be evaluated as the difference between the values stored in the place `counted` at two distinct time instants, divided by the time interval.

The models described in this section can be easily composed by using the join operator. The place superposition allows to combine the track elements in sequence in order to model a specific railway line. An injector terminal is needed at each entering point to model the train flow from that line, while a measuring terminal is needed at each final track element. The place `ERTMSstate` of each model in the line also needs to be shared with the corresponding place of the ERTMS model in order to have different train spacing and speed according to the specific operating mode.
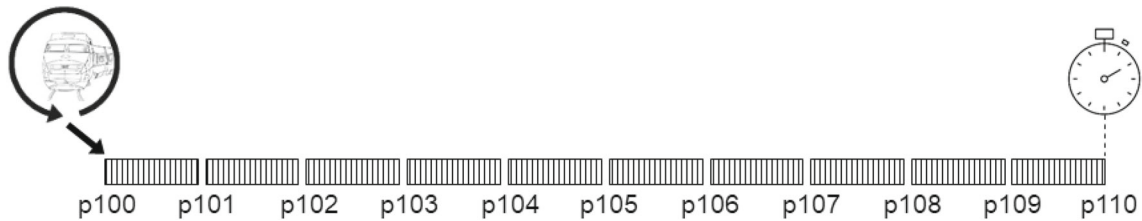
Fig. 10. Graphical representation of case study 1

5. Capacity analysis

This section describes the capacity analysis we performed by using different case-studies and by varying input parameters accordingly. The goal of the analysis is twofold: (1) to validate the proposed models and modelling approach, and (2) to show how to evaluate the benefits of Virtual Coupling in specific railway scenarios based on those models. SAN model evaluation has been performed by using the Möbius simulator [CGK+09], and more specifically Möbius version 2.5 running on a notebook equipped with an Intel(R) Core (TM) i7-10510U CPU @1.80 GHz, with 16.0GB RAM and Windows 10 Pro OS.[2] For all simulations, we set the confidence level to 0.95 and the relative confidence interval to 0.5%. After appropriate experimentation, we verified that these values represent a good compromise between accuracy and simulation time: all the estimated values converged within a maximum of 10.000 batches, with a maximum error of 0.5 every 100 trains. Moreover, for all the experiments described in the following of this paper, we measured a maximum elapsed running time lower than 75 s on the described notebook.

5.1. Straight Line

The first case study is depicted in Fig. 10 and represents a 2 km long straight line. The elementary track element length is set to 200 m, hence the line model is composed by 10 track elements. The connection points are numbered from $p100$ to $p110$; as depicted in the figure, the train injector is connected to point $p100$, while the measurement terminal is connected to $p110$.

We assume that the trains running on the line are composed by a FS E.402B locomotive [MCC02], which is around 19.5 m long, with 19 coaches, for a total length of around 400 m. That means that each train occupies 3 track elements. Trains run at 90 km/h, so they need 8 s to cover each track element. In PS operating mode, we assume a speed of 30 km/h, so a total of 24 s is needed to cover the track element. We assumed a minimum distance of 200 m, which is exactly one track element, between two subsequent (virtually coupled) trains in FSVC operating mode, while a distance of 600 m is realistic in FS mode. The final composed model in Möbius is depicted in Fig. 11. It is made up of 10 submodels *straight*, each one representing a track element; a submodel *entering* modeling the train injector at the point $p100$; a submodel *measuringTerminal* representing the measurement terminal at the point $p110$; and a submodel *ERTMSstates* modelling the operational modes. The latter shares the state variable *ERTMSstate* with all the other models, furthermore the input and the output connections of each submodel is shared with the models of the preceding and following track element. Starting from the atomic models described in the previous section, the creation of the joint model for this line is limited to the creation of 12 shared variables among models. Graphically, the join operator is represented as a box to which the joint models are connected. The creation of the shared variables is manually done, by the configuration popup window associated with this box. Möbius allows to select places of the joint submodels to be superimposed, so providing the actual connections (see bottom of Fig. 11).

Given the difficulty to compare evaluation results against real-world systems, since no Virtual Coupling implementation is currently available, this case study has been used to support model validation. To that aim, we considered a stable FSVC, FS, or PS operating mode for the entire simulated period and we enabled the simulation logs produced by the Möbius simulator.

---

[2] The Möbius project used in this paper as well as the evaluation output log files can be publicly accessed at https://github.com/RobertoNardone/VC-model.
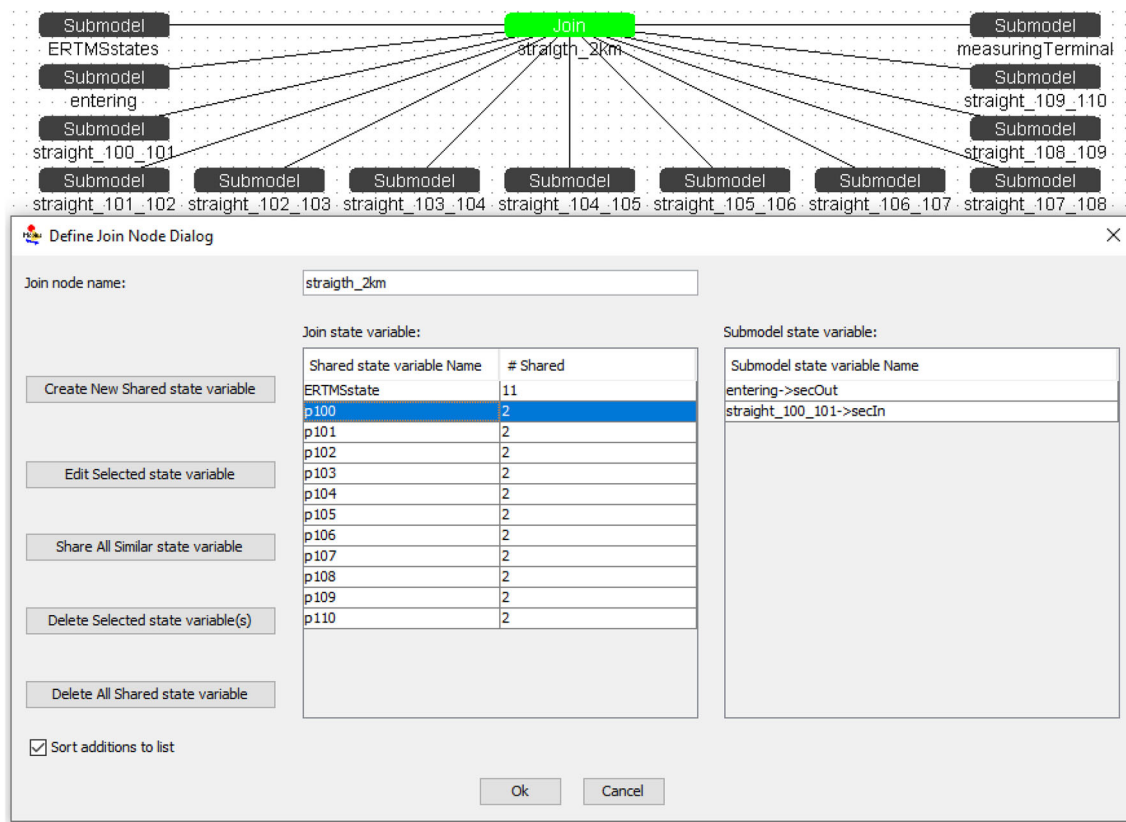
Fig. 11. Composed SAN model for case study 1

Then, at each firing of transitions, the state of the places before and after the firing have been compared with results obtained by an empirical evaluation, which has been manually executed.

We performed different capacity analysis by choosing reasonable values for the rates of transitions, thus varying possible input parameters as described in the following paragraphs.

Capacity versus operating modes. The first analysis aims at evaluating the impact on line capacity of the transition rates between operating modes. In the first experiment, we assume that the rates are those reported in Table 1 (please note that the table uses the notation $x : y : z$ for incremental ranges from x to z with a step y). Variations in the rate of the transition imply increased probabilities to be in a specific operating mode compared to another. In particular, the highest (resp. lowest) value for the rate of transition *FSVC2FS* corresponds to a reduced (resp. increased) probability to be in FSVC mode, and an increased (resp. reduced) probability to be in FS mode.

The plot in Fig. 12 reports simulation results for experiment 1. On the horizontal axis, the graph reports the rate of the transition *FSVC2FS* varying from $1.0E-02$ to $1.0E+00$, while on the vertical axis it reports line capacity in terms of trains/h. As expected, the maximum capacity can be obtained when *FSVC2FS* reaches its minimum value. The theoretical maximum capacity in full supervision is depicted with the red line in the plot and can be estimated at 75.0 trains/h, while the theoretical maximum capacity for the line using Virtual Coupling is depicted with the green line in the plot and can be estimated at 150.0 trains/h. These values have been evaluated by considering the system in a stable FS (resp. FSVC) mode, with no failures. Please, note that even if the distance between trains reduces from 600 to 200 m, with the used parameters VC is capable of placing exactly one train more every other train and the capacity of the line doubles.

Table 1. Parameters for experiment 1 of case study 1

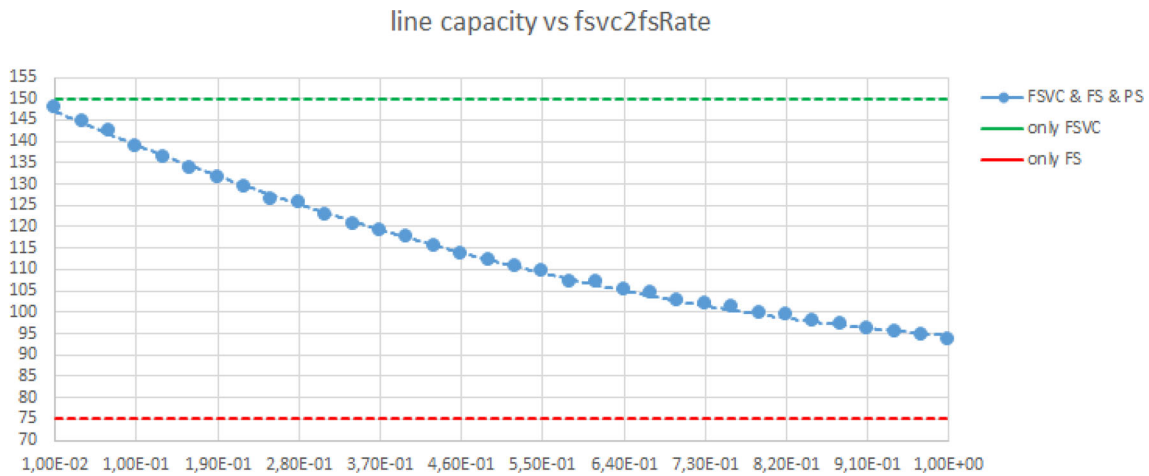| Transition	Rate [h^{-1}]
FS2FSVC	5.0E−02
	(50E−03)
FSVC2FS	1.0E−02 : 3.0E−02 : 1.0E+00
	(10E−03 : 30E−03 : 1000E−03)
FS2PS	1.0E−02
	(10E−03)
REPAIR	2.0E−01
	(200E−03)

Fig. 12. Impact of FSVC2FS rate on line capacity in case study 1

The analysis demonstrates that with a rate of the transition *FSVC2FS* equal to $1.0E-02$ h^{-1} (five times lower than the rate of the transition *FS2FSVC*), we can obtain 148.0 trains/h, that is very close to the theoretical maximum. With a rate of *FSVC2FS* equal to $1.0E+00$ h^{-1} the capacity is higher than the one obtained with full supervision only. In that case, we obtained a line capacity of 93.5 trains/h,[3] which is still higher than 75.0 trains/h that is the maximum capacity in full supervision. Such a result would demonstrate a potential benefit of Virtual Coupling in those reference—realistic—conditions.

In the second experiment, we kept the rate of transition *FSVC2FS* fixed, and only varied the rate of transition *FS2FSVC*, as reported in Table 2. The corresponding results are plotted in Fig. 13. The results show a slight capacity increase with the increase of the transition rate from full supervision to Virtual Coupling, but the effect is not as evident as in the previous experiment. In fact, if we neglect some small fluctuations due to the confidence interval of simulations, the estimated line capacity varies from 143.7 on the extreme left side of the graph to 146.9 on the extreme right.

From the analysis performed on the sample track, it emerged that the capacity impact of transitions between operating modes is not symmetrical. In fact, the negative effect of increasing the rate of the transition *FSVC2FS* has a much higher impact than the benefits of increasing the rate of the transition *FS2FSVC*. That means efforts should be concentrated on minimizing the risk of switching from FSVC to FS rather than improving the enabling conditions for FSVC.

---

[3] Such a results includes the possibility of switching to partial supervision.

Table 2. Parameters for experiment 2

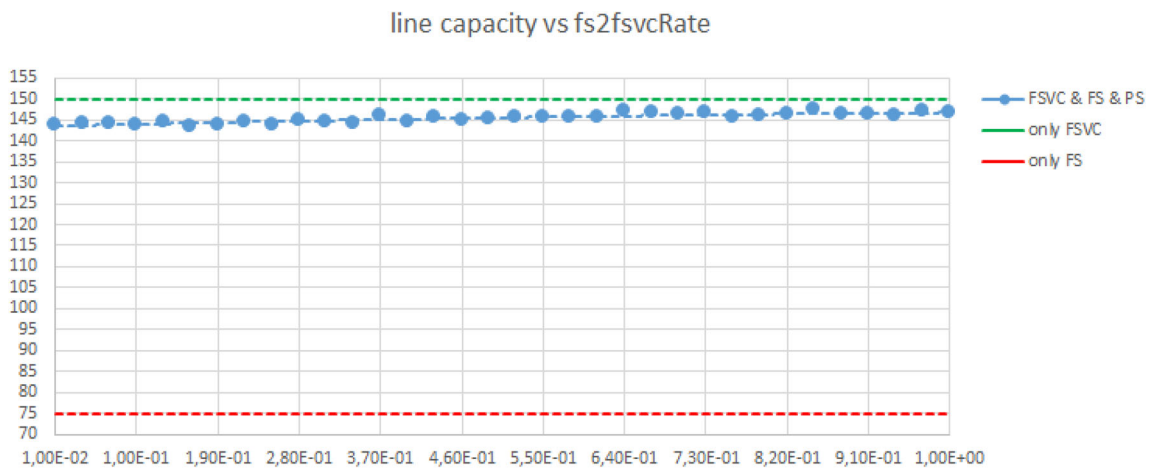| Transition	Rate [h^{-1}]
FS2FSVC	1.0E−02 : 3.0E−02 : 1.0E+00
	(10E−03 : 30E−03 : 1000E−03)
FSVC2FS	5.0E−02
	(50E−03)
FS2PS	1.0E−02
	(10E−03)
REPAIR	2.0E−01
	(200E−03)

Fig. 13. Impact of FS2FSVC rate on line capacity in case study 1

Capacity versus train length. In this third experiment, the analysis aims to evaluate the impact of train length on line capacity. We assumed the rates of transitions as reported in Table 3 and we varied the train length from 200 to 600 m.

The plot in Fig. 14 reports the results of simulations for the experiment: capacity shows a trend slightly below the maximum obtainable from FSVC, but significantly above the one associated with FS. That means the overall capacity in realistic conditions is close to its theoretical maximum regardless of train length, and that result would further support general adoption of Virtual Coupling. In general, as expected, the benefits of virtual coupling decrease as train length increases, and the marginal benefit is lower for longer trains; however, longer trains also

Table 3. Parameters for experiment 3

Transition	Rate [h^{-1}]
FS2FSVC	5.0E−02
	(50E−03)
FSVC2FS	5.0E−02
	(50E−03)
FS2PS	1.0E−02
	(10E−03)
REPAIR	2.0E−01 (200E−03)

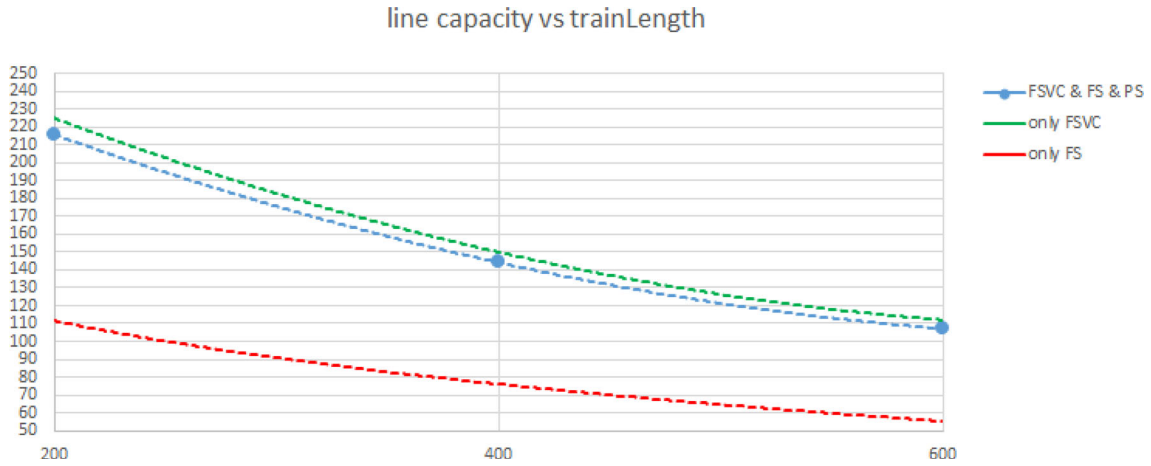line capacity vs trainLength

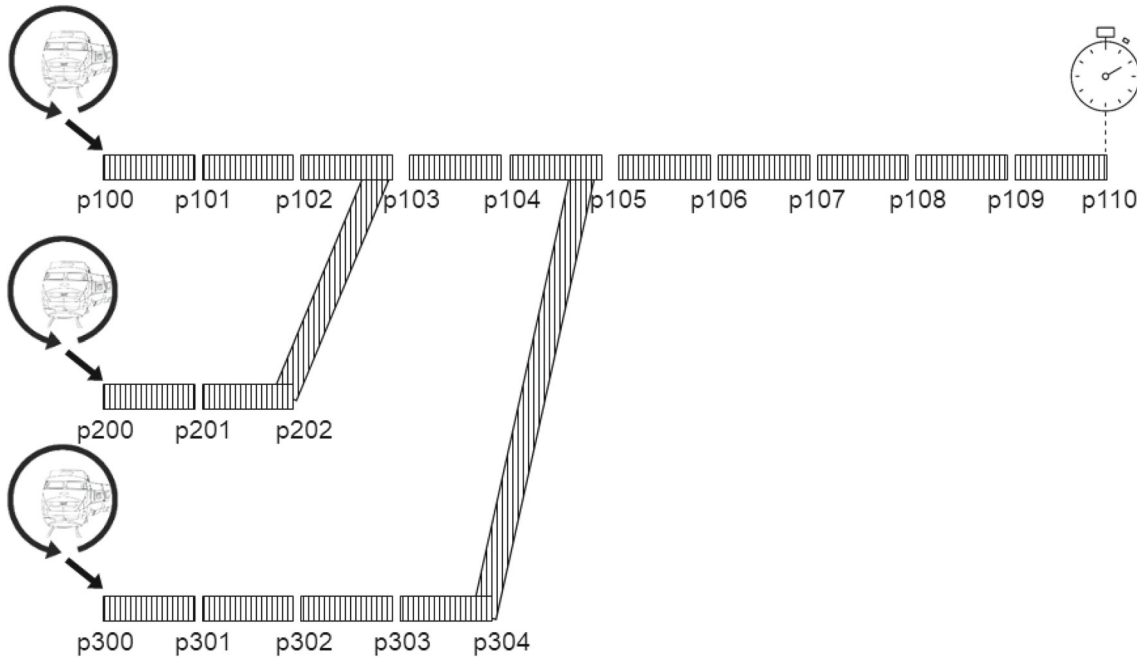**Fig. 14.** Impact of train length on the line capacity in case study 1

Fig. 15. Graphical representation of case study 2

reduce the overall impact of falling back to FS from FSVC due to lesser couplings needed. This is why the blue and red lines are closer when trains are longer.

5.2. Bottleneck Line

The second case study is depicted in Fig. 15 and represents a 2 km long line, where 3 tracks converge into a single one, creating a bottleneck. This is a situation which is not uncommon in central Europe high-traffic corridors, especially considering railway hubs such as Milan, Frankfurt or Brussels. As in previous case study, the length of the elementary track element is set to 200 m, the line model is composed by 18 track elements. Specifically, the two points connect $p102$ and $p202$ with $p103$, and $p104$, $p304$ with $p105$.
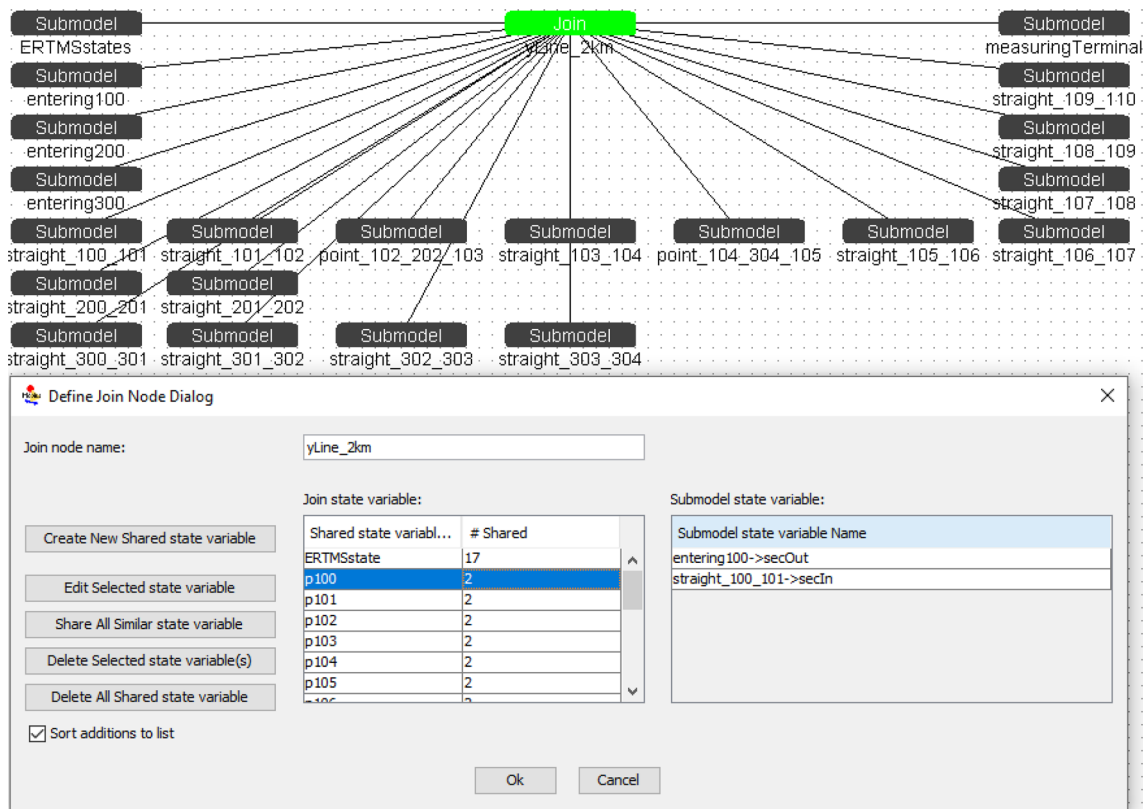
Fig. 16. Composed SAN model for case study 2

For the sake of simplicity, we assume the same train types as in previous case study, with similar characteristics in all the operating modes. The composed model in Möbius is depicted in Fig. 16. It is made up of 14 *straight* submodels; 3 *points* submodels; 3 *entering* submodels, modeling the train injectors at $p100$, $p200$ and $p300$; a *measuringTerminal* submodel, representing the measurement terminal at point $p110$; and a *ERTMSstates* submodel, modelling operational modes. The effort to create this model resides in the definition of the 20 state variables.

This specific case study demonstrates the flexibility of composable SAN models in representing railway tracks of increasing complexity, especially in presence of possible bottlenecks. As in previous example, we performed different capacity analysis by varying input parameters.

Capacity versus operating modes. The first analysis aims at evaluating the impact on line capacity of the transition rates between operating modes. In the first experiment, we assume the same rates reported in Table 1. Results are plotted in Fig. 17. As expected, the trend is similar to the one obtained in the previous example, with the difference that the range here is limited, varying from a minimum capacity of 129.7, with the highest rate of the transition *FSVC2FS*, to a maximum one of 159.9. According to expectations, the results clearly show the benefits of Virtual Coupling: a low rate of the transition *FSVC2FS* results in a high probability for the system to be in the FSVC mode, corresponding to the highest capacity, close to the theoretical maximum; if the rate of that transition increases, line capacity decreases due to the increased probability for the system to be in the FS mode. Overall, results confirm that, even in the worst case, Virtual Coupling increases line capacity of around 9%.

A similar result is obtained when fixing the rate of transition *FSVC2FS* and varying the rate of transition *FS2FSVC*, with parameters reported in Table 2. The graph in Fig. 18 shows a trend similar to the previous case study, from a minimum capacity of 157.3 to a maximum of 159.8.
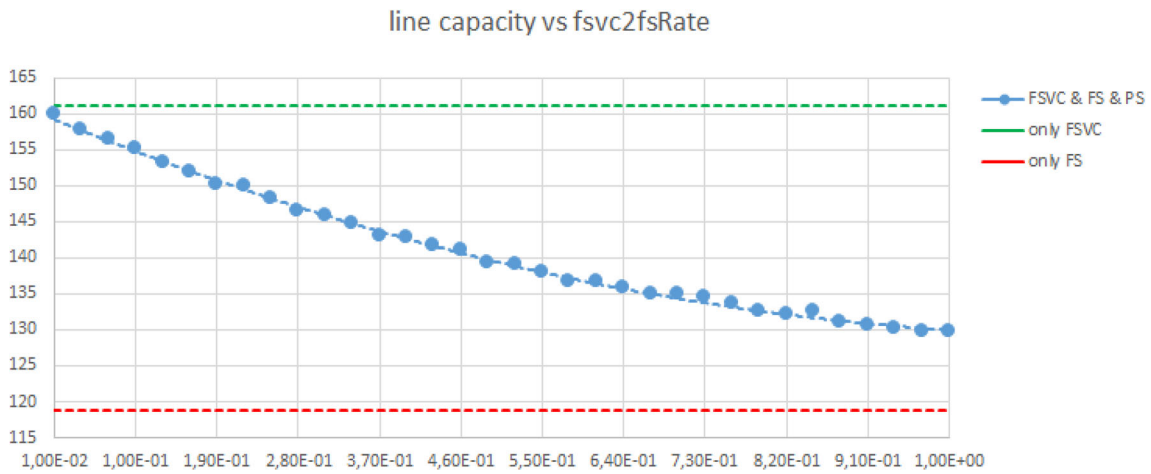
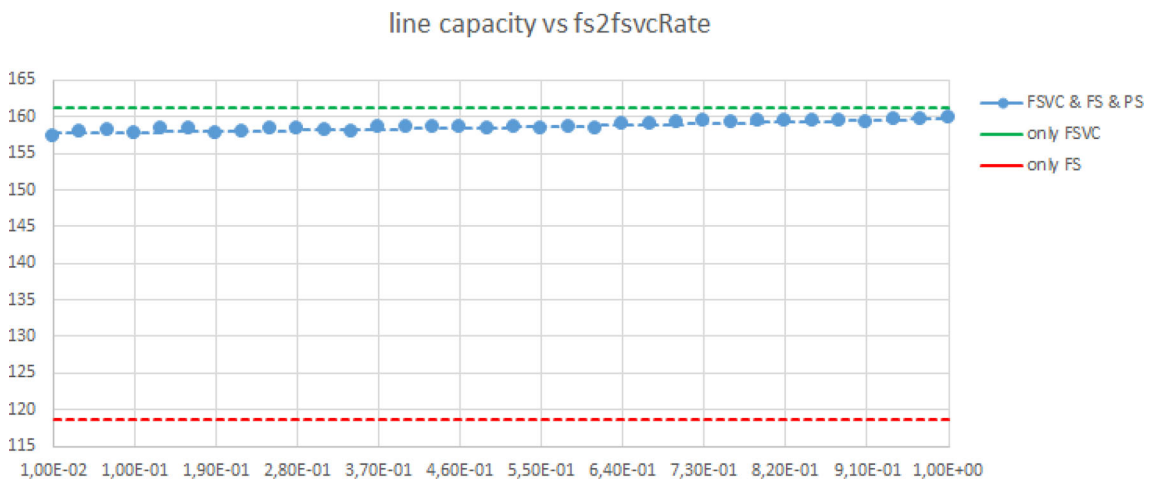**Fig. 17.** Impact of FSVC2FS rate on line capacity in case study 2

Fig. 18. Impact of FS2FSVC rate on line capacity in case study 2

Similarly to previous track configuration, the capacity impact of transitions between operating modes is not symmetrical. In fact, the negative effect of increasing the rate of the transition *FSVC2FS* has a much higher impact compared to the benefits of increasing the rate of the transition *FS2FSVC*. Such an effect relates to the choice of parameters and to the fact that, after a failure and the consequent activation of the PS modes, the system is recovered eventually to the FSVC mode; hence, it is important to avoid the occurrence of failures leading to an exit from this state.

Capacity versus train length. This analysis evaluates the impact of train length on line capacity, assuming the rates reported in Table 3, and varying train length from 200 to 600 m. The plot in Fig. 19 confirms the trend of actual capacity being slightly below its maximum in FSVC. The graph also suggests that train length has an even more limited impact compared to case study 1, perhaps due to synchronization phenomena at the junction (reduced probability of inter-train gaps with longer trains); however, that does not account for the possibly smoother and more uniform distribution of trains from the incoming lines in case of shorter trains.
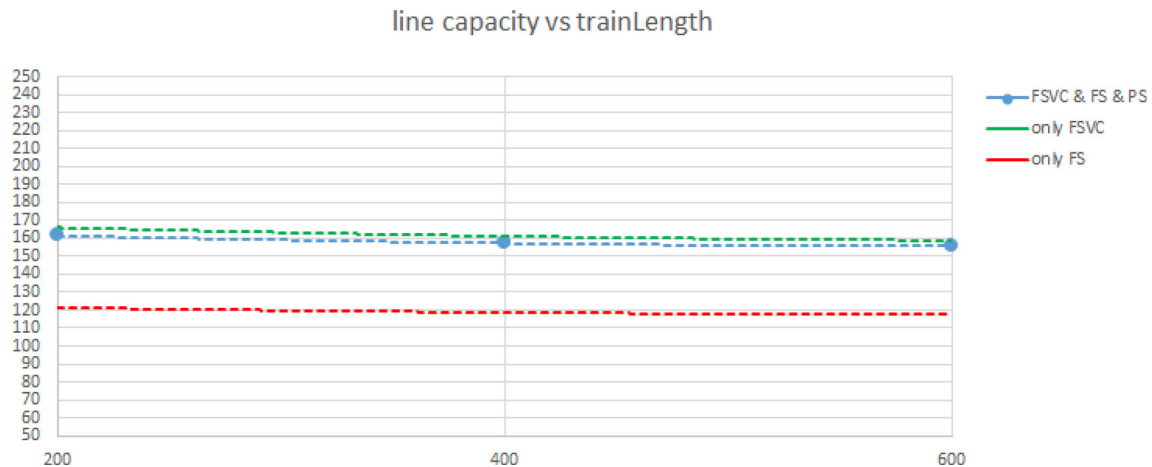
Fig. 19. Impact of train length on line capacity in case study 2

6. Related work

The work described in this paper progresses previous research aimed at performability evaluation of Virtual Coupling in modern and future railway signalling systems. Specifically, the previous work reported in reference [DMDVF⁺20] provides a proof-of-concept of Virtual Coupling, and defines a coupling control algorithm accounting for time-varying delays affecting the communication links. More specifically, the work in reference [FMN⁺18] evaluates a solution integrating Virtual Coupling within a specific operating mode using the ERTMS/ETCS standard specification, and proposes some directions for new research with the goal to enable efficient and safe train platooning. With respect to the previous works, in this paper we have introduced a model-based approach for quantitative evaluation of capacity increase considering specific railway network topologies, and a set of atomic models based on SAN [SM01], in order to build a library of reusable and composable models. The SAN formalism is a powerful modeling language inspired by Petri Nets. SANs have been widely used for performance and dependability evaluation; they have been applied to different domains and purposes, including performability analysis.

Regarding railway research, in the last years, a considerable effort has been spent in Europe to bring groundbreaking innovations in railways. The Shift2Rail Joint Undertaking was established under Horizon 2020 to foster coordination of research and innovation in the rail sector. Shift2Rail (S2R) has devoted an entire pillar of its programme (i.e., IP2, advanced Traffic Management and Control System) to enhance capacity and reliability through improved management of signalling and supervision systems. One important step in this direction is to increase rail capacity by introducing moving block signalling at the third level of ERTMS/ETCS together with the more visionary idea of Virtual Coupling, as investigated by the recent S2R research projects, from X2Rail-1 to X2Rail-5, AS-TRail, MOVINGRail, PERFORMINGRAIL.[4] Although ERTMS L3 has been widely investigated and expected to be operational in the near future, Virtual Coupling is considered by the International Union of Railways (UIC) still far away from technical maturity.[5] Recent reports from S2R projects and international railway organizations or professional institutions have introduced virtual coupling related concepts (e.g., [Shi15, Mea16]), but only a few works in the scientific literature address the challenges posed by virtually coupled trains. As of the time of writing, we were able to find 35 scientific papers, of which more than a half published in the years 2019 and 2020. The most recent works address virtual coupling operational scenarios [AQG20, QWG20], control approaches [SLY⁺20], coordination control models [LWZW19], collision avoidance and mitigation approaches [CWM21, CXL20]. Some of those works focus on the technologies required for implementation (e.g., [Goi16, FLFCC17]), in particular Vehicle to Vehicle (V2V) communications [ULdPM16, GMB⁺18], but just a few pioneering studies exist ad-

---

[4] S2R projects page: https://projects.shift2rail.org/.
[5] www.railway-energy.org/static/Virtually_coupled_trains_86.php.

dressing the evaluation of qualitative and quantitative aspects of Virtual Coupling. In reference [LLS15], the authors investigate the usage and expected performance of 5G technology for train-to-train communications in coupling/de-coupling and platooning scenarios, while the simulation described in [Sch17] is meant to study the expected capacity increase. These papers present some preliminary results, in particular in reference [Sch17], Virtual Coupling operational scenarios at low speed (e.g., in the proximity of a station) are simulated referring to an existing high-speed line in Japan.

In particular, the work presented in this paper takes advantage of some features of SANs which allow for a modular, compositional approach very suitable for the specific goal of developing models of railway lines.

7. Conclusions and future work

This paper deals with the problem of evaluating the capacity increase of future railways, where train spacing constraints can be reduced thanks to the adoption of ERTMS Level 3. With a particular reference to Virtual Coupling, we developed a compositional approach based on SAN models which allows engineers to simulate train traffic along a line and thus evaluate the subsequent capacity increase. The approach represents a mean for railway engineers to evaluate the capacity increase in specific scenarios, according to different operating conditions and traffic flows.

The SAN modeling formalism allowed us to start to build a library with basic templates to be composed into more complex models in order to get useful insights about Virtual Coupling scenarios. We obtained some preliminary results by evaluating those models on reference case studies in order to measure the expected capacity increase of Virtual Coupling. However, rather than discussing some specific results, our research focused on providing a general tool to build and compose models with an expressive power such to represent all performance and dependability aspects of interest for the analysis of railway Virtual Coupling in real-world scenarios.

The research described in this paper has been partly funded by the EU Shift2Rail PERFORMINGRAIL project. Coherently with PERFORMINGRAIL project objectives, future work will be aimed at extending the models in order to enable safety evaluation of moving block signalling and Virtual Coupling within future railway technologies. Furthermore, we will make an effort to develop a modeling environment which can be actually used by railway engineers in industry. That would leverage aspects such as multi-paradigm modeling (including abstraction, metamodeling and multiformalism), domain specific languages (including specific profiles of the Unified Modeling Language) together with model-to-model transformations [MFM+14]. The overall approach would be such to hide the complexity of low-level modeling artefacts—such as the SAN used in this paper—to the end users, who will be dealing with "bricks" coming from composable model libraries. That would allow us to achieve a trade-off between ease of use, expressive power and solving efficiency. Moreover, we plan to strengthen model validation by leveraging results obtained from additional models or diverse parallel modeling. Extensions to reliability, safety and security modeling based on (semi)formal methods are also planned as future developments [FFG14, Fla12].

Acknowledgements

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[AQG20] Aoun J, Quaglietta E, Goverde RMP (2020) Exploring demand trends and operational scenarios for Virtual Coupling railway signalling technology. In: 2020 IEEE 23rd International conference on intelligent transportation systems, ITSC 2020

[CGK+09] Courtney T, Gaonkar S, Keefe K, Rozier EWD, Sanders WH (2009) Möbius 2.3: an extensible tool for dependability, security, and performance evaluation of large and complex system models. In: 2009 IEEE/IFIP International conference on dependable systems and networks. IEEE, pp 353–358

[CWM21] Cao Y, Wen J, Ma L (2021) Tracking and collision avoidance of Virtual Coupling train control system. Alex Eng J 60(2):2115–2125

[CXL20] Chen M, Xun J, Liu Y (2020) A coordinated collision mitigation approach for Virtual Coupling trains by using model predictive control. In: 2020 IEEE 23rd International conference on intelligent transportation systems, ITSC 2020

[DMDVF+20] Di Meo C, Di Vaio M, Flammini F, Nardone R, Santini S, Vittorini V (2020) ERTMS/ETCS Virtual Coupling: proof of concept and numerical analysis. IEEE Trans Intell Transport Syst 21(6):2545–2556

[FFG14] Fantechi A, Flammini F, Gnesi S (2014) Formal methods for railway control systems. Int J Softw Tools Technol Transf 16(6):643–646

[Fla12] Flammini F (2012) Railway safety, reliability, and security: technologies and systems engineering. IGI Global, Hershey

[FLFCC17] Fraga-Lamas P, Fernández-Caramés TM, Castedo L (2017) Towards the internet of smart trains: a review on industrial IoT-connected railways. Sensors (Switzerland) 17(6)

[FMN+18] Flammini F, Marrone S, Nardone R, Petrillo A, Santini S, Vittorini V (2018) Towards railway Virtual Coupling. In: 2018 IEEE International conference on electrical systems for aircraft, railway, ship propulsion and road vehicles and international transportation electrification conference (ESARS-ITEC). IEEE, pp 1–6

[FvHAM17] Furness N, van Houten H, Arenas L, Maarten B (2017) ERTMS level 3: the game-changer

[GMB+18] Gomez AA, Mozo E, Bernado L, Zelenbaba S, Zemen T, Parrilla F, Alberdi A (2018) Performance analysis of ITS-G5 for smart train composition coupling. In: Proceedings of 2018 16th international conference on intelligent transport system telecommunications, ITST 2018

[Goi16] Goikoetxea J (2016) Roadmap towards the wireless virtual coupling of trains. In: International workshop on communication technologies for vehicles. Springer, Berlin, pp 3–9

[LLS15] Le HQ, Lehner A, Sand S (2015) Performance analysis of ITS-G5 for dynamic train coupling application. In: Kassab M, Berbineau M, Vinel A, Jonsson M, Garcia F, Soler J (eds) Communication technologies for vehicles. Springer International Publishing, Cham, pp 129–140

[LWZW19] Liu L, Wang P, Zhang B, Wei W (2019) Coordinated control method of virtually coupled train formation based on multi agent system. Smart Innov Syst Technol 129:225–233

[MCC02] Mingozzi E, Cau G, Cavaliere F (2002) The train communication network in the trains of FS fleet: optimisation, integration and interoperability of railway functionality. WIT Trans Built Environ 61

[Mea16] Mitchell I et al (2016) ERTMS level 4, train convoys or virtual coupling. IRSE International Technical Commitee

[MFM+14] Marrone S, Flammini F, Mazzocca N, Nardone R, Vittorini V (2014) Towards model-driven V&V assessment of railway control systems. Int J Softw Tools Technol Transf 16(6):669–683

[MMS85] Meyer JF, Movaghar A, Sanders WH (1985) Stochastic activity networks: structure, behavior, and application. In: International workshop on timed Petri Nets. IEEE Computer Society, Washington, DC, pp 106–115

[QWG20] Quaglietta E, Wang M, Goverde RMP (2020) A multi-state train-following model for the analysis of Virtual Coupling railway operations. J Rail Transp Plan Manag 15

[RG94] Ren W, Green D (1994) Continuous platooning: a new evolutionary operating concept for automated highway systems. In: Proceedings of 1994 American control conference-ACC'94, vol 1. IEEE, pp 21–25

[Sch17] Schumann T (2017) Increase of capacity on the Shinkansen high-speed line using Virtual Coupling. Int J Transp Dev Integr 1(4):666–676

[Shi15] Shift2Rail Joint Undertaking. Multi-annual action plan, Nov 2015

[SLY+20] She J, Li K, Yuan L, Zhou Y, Su S (2020) Cruising control approach for virtually coupled train set based on model predictive control. In: 2020 IEEE 23rd International conference on intelligent transportation systems, ITSC 2020

[SM01] Sanders WH, Meyer JF (2001) Stochastic Activity Networks: Formal definitions and concepts. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 2090, pp 315–343

[ULdPM16] Unterhuber P, Lehner A, de Ponte Müller F (2016) Measurement and analysis of ITS-G5 in railway environments. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 9669, pp 62–73

[UNI16] UNISIG (2016) ERTMS/ETCS: System requirements specification—SUBSET-026, issue 3.6.0