A simulation-based approach for evaluating train operating costs under different signalling systems

G. Corapi, D. Sanzari, V. De Martinis, L. D’Acierno & B. Montella
Department of Civil, Architectural and Environmental Engineering, ‘Federico II’ University of Naples, Italy

Abstract

In this paper, we propose a model framework for providing optimal driving strategies and related speed profiles which minimise the energy consumption of rail convoys. Previous models were extended and applied to evaluate the effects of different signalling systems upon rail operating costs. The proposed method was tested on a real rail network, the Cumana suburban railway (Italy).

Keywords: signalling systems, energy consumption, rail operating costs.

1 Introduction

In recent decades, increasing public interest in environmental issues and energy consumption has drawn attention to the rail transportation system as one of the most suitable solutions for sustainable mobility especially in urban or suburban contexts where residential densities can reach considerable values. Since even a small reduction in energy consumption due, for instance, to a better speed profile could generate significant rail operating economies, in the literature there has been great interest in automatic train control systems. Indeed, improvements in operation performances can be obtained by means of technologies such as Driving Assistance Management Systems (DAMS), as shown by Albrecht [1] and Hansen and Pachl [2], which help train drivers follow optimal speed profiles, or new communication systems (e.g. the GSM-R frequency used in the ETCSs) with localization systems [3], which allow the continuous updating of speed profiles from the real position of convoys. Moreover, many improvements have been proposed for modelling energy consumption as a function of the train running process [4, 5], for optimising jointly energy consumption and travel time.
delays by modifying train running times [6], or for minimising energy use by means of control strategies [7]. Likewise, Xuan [8] proposed a perturbation analysis to provide necessary conditions for defining optimal driving strategies in some specific infrastructure conditions (such as slope sections).

Detailed analysis on specific railway systems are also given, for example, by Lukaszewicz [9] on freight train operation and by Ke and Chen [10] on planning a mass rapid transit system. The former analyses the energy consumption trend and its relationship with the maximum traction ratio, maximum braking ratio, upper and lower restrictions of speed, and pre-braking coasting distance. Likewise, the latter provides a tool for the optimisation of the block layout and the running speed to achieve minimum energy consumption with maximum train capacity.

Bocharnikov et al. [11] proposed a genetic algorithm for optimising the coasting phase (i.e. minimising energy consumption) by modifying acceleration and braking rates. Likewise, in terms of energy saving control, Gu et al. [12] adopted non-linear programming methods for defining optimal speed profiles in the case of a moving block signalling system, and Ding [13] compared optimal train driver behaviours in the case of fixed and mobile block systems. Finally, Bai et al. [14] developed a real-time control tool to simulate the interaction between a real-time information system and train operations in the case of different infrastructure conditions and/or speed restrictions.

In terms of methodology, it is worth nothing that, as shown by [15–17], the performance of rail systems and their related capacities have mainly been analysed by neglecting effects on travel demand. Indeed, the first papers to consider that the main purpose of a rail system is actually to satisfy traveller requirements were those of Cascetta and Carteni [18], Hamdouch et al. [19], Kanai et al. [20] and Zheng et al. [21]. Moreover, neglecting rail vehicle capacity (in terms of maximum number of passengers per vehicle), Mazzeo et al. [22] and Quaglietta et al. [23] proposed a system of integrated models able to jointly calculate rail performance (i.e. rail enterprise efficiency) and related effects on users (i.e. service effectiveness and quality) in the case of rail systems failures. D’Acierno et al. [24] proposed an extension of [22] and [23] by introducing capacity constraints of rail vehicles in order to provide more realistic simulated effects. Finally, Gallo et al. [25] proposed a multimodal approach for defining optimal rail frequencies (in terms of trains per hour) by taking into account effects on travel demand, on other transportation systems, and on external costs (such as environmental pollution or energy consumption).

This paper proposes a simulation-based approach for evaluating rail operating costs under different signalling systems by implementing energy-efficient strategies. Traditional models proposed in the literature for simulating railway systems are extended by means of a model for estimating energy consumption and used as sub-routines in an optimisation model for determining optimal driving strategies which minimise energy consumption.

This paper is organised as follows: Section 2 describes the system simulation model and optimal driving strategies; Section 3 defines the performance indexes
for evaluating different signalling systems; Section 4 proposes a numerical application; finally, Section 5 concludes and provides research prospects.

2 System simulation and definition of optimal driving strategies

In design problems or in (real-time or off-line) management of transportation systems, it is necessary to have system simulation models that allow exploration of a large number of alternative projects or simulate beforehand effects of a strategy in terms of future network conditions. In this context, given the complexity of rail systems due to the interaction among infrastructure, signalling systems, rolling stock, timetable and travel demand (users or freight), as shown by D’Acierno et al. [24], several simulation approaches have been proposed that can be classified, according to the assumption of the level of detail considered, as follows: macroscopic, mesoscopic or microscopic. Macroscopic simulation models adopt a high abstraction level of railway infrastructure and operations. They are mainly adopted in long-term planning to determine at a macro level some network or service features (such as the number of stations, number of lines, average service frequencies, average speed or required rolling stock) without providing detailed information of node management. Likewise, mesoscopic simulation models are able to simulate a simplified system by means of a multi-scale framework consisting of both macroscopic and microscopic elements. Finally, microscopic simulation models represent the system elements (such as signalling systems, radii of curvature, slopes, timetables, locomotive types, number of passenger cars, number of freight cars or adhesion values) in order to provide a more precise description of rail operations.

Since our purpose is to analyse effects of different signalling systems on train performance in terms of energy consumption and travel times (i.e. maximum frequencies, headways, waiting times for passengers, etc.), we have to adopt a microscopic approach. This approach, although it requires high calculation times, can be implemented by means of off-line procedures in order to define predetermined driving strategies to be adopted. Moreover, since adoption of a microscopic model requires the solution of a system of differential equations, we adopt the microscopic model proposed by Quaglietta et al. [23] for integrating the system of differential equations and extend it by adding an energy efficiency module (based on the calculation of energy consumption and on the removal of a time optimal driving strategy in order to allow the system to adopt any generic driving strategy) for optimising travel speed profiles.

Analytically, a system simulation model can be formulated as:

\[ SP = SSM(IN, RS, SS, TM, DS(y)) \]  
\[ \text{with: } \quad y = \{v, a, d, j\} \]

where \( SP \) is the vector of speed profiles; \( IN \) is the vector of infrastructure parameters, such as track lengths, gradients, radii, etc.; \( RS \) is the vector of rolling stock parameters, such as vehicle characteristics, resistance models, etc.; \( SS \) is the vector of signalling system parameters, such as block section definition, type
of signalling system, etc.; $TM$ is the vector of timetable parameters, such as reserve time on the section, minimum running time, etc.; $DS$ is the vector of the driving strategy which depends on values of vector $y$; $y$ is the vector of driving strategy parameters; $v$ is the vector of maximum speeds for each track section; $a$ is the vector of maximum acceleration for each track section; $d$ is the vector of maximum deceleration for each track section; $j$ is the vector of maximum jerks for each track section; $SSM$ is the system simulation model which can be formulated as a function of $IN$, $RS$, $SS$, $TM$ and $DS$ which provides vector $SP$ as output.

In general, a driving strategy consists of a pre-determined combination of moving phases that can be referred to four kinds of regimes: acceleration, deceleration, cruising and coasting. Acceleration and deceleration regimes are used for transition phases, typically in those cases when the maximum allowable track speed changes from one section to another (undisturbed condition), when an unplanned stop happens (disturbed conditions) or when the convoy approaches a station. The cruising regime consists of maintaining a constant speed on the track. Finally, the coasting regime consists of moving forward using only kinetic energy of a convoy, avoiding the application of any tractive or braking effort and hence without any consumption of energy for traction. Hence, components of vector $y$ specify target values of acceleration ($a$), deceleration ($d$), cruising ($v$) regimes and their transitions ($j$), while the coasting regime is introduced for energy consumption reduction if a running time reserve is available. Details of moving phases are shown in fig. 1, in the case of a speed profile with a given vector $y$ and a pre-fixed coasting regime. Dotted lines represent the speed profile with the same vector $y$ without the coasting phase.

Two main driving strategies may be identified in the literature: the Time Optimal ($TO$) and the Energy Saving ($ES$) strategies. The first, used by the system simulation model proposed by Quaglietta et al. [23], is based on the assumption of adopting maximum values of components of vector $y$ and on the absence of any coasting phase. This strategy, which allows the minimum running time to be reached on tracks in undis turbed conditions, can be formulated by assuming in eqn (1) the following value for variable $y$:

$$y = \{v_{\text{max}}, a_{\text{max}}, d_{\text{max}}, j_{\text{max}}\}$$ (3)
where $v_{\text{max}}$, $a_{\text{max}}$, $d_{\text{max}}$ and $j_{\text{max}}$ are respectively the maximum speed, maximum acceleration, maximum deceleration and maximum jerk allowed by the infrastructure, by the rolling stock and/or passenger (or equivalently freight) comfort on each track section. Hence, the microsimulation (i.e. a simulation with a microscopic approach) model proposed by Quaglietta et al. [23] can be formulated as:

$$SP = SSM(IN, RS, SS, TM, DS(v_{\text{max}}, a_{\text{max}}, d_{\text{max}}, j_{\text{max}}))$$ (4)

where $SSM$ is a function, i.e. it provides a unique result in terms of vector $SP$ for each set of variables $IN$, $RS$, $SS$, $TM$, $v_{\text{max}}$, $a_{\text{max}}$, $d_{\text{max}}$ and $j_{\text{max}}$.

On the other hand, the $ES$ strategy also takes into account the coasting regime which is introduced, for instance, following the $ASAP$ (as soon as possible) rule that consists of starting the coasting phase so that the travel distance is covered in the planned timetable by using the running time reserve which represents a buffer time. Hence, in fig. 1, the continuous line represents the speed profile in the case of $ES$ strategy, while the dotted line represents the speed profile variation in the case of $TO$ strategy.

Energy consumption can be estimated by means of a model, termed Energy Consumption Estimation ($ECE$), which depends on infrastructure, rolling stock and speed profile parameters.

One of the aims of this paper is to develop an optimisation model to determine the optimal driving strategy which minimises energy consumption. Hence, it is necessary to formulate a bi-level multidimensional optimisation constrained problem where the upper level consists in minimising energy consumption and the lower level in calculating speed profiles, that is:

$$\hat{y} = \arg \min_{y \in S_y} EC(IN, RS, SP(y))$$ (5)

s.t.: $SP = SSM(IN, RS, SS, TM, DS(y))$ (6)

with:

$$S_y = \begin{cases}
0 \leq v \leq v_{\text{max}} & 0 \leq a \leq a_{\text{max}} & 0 \leq d \leq d_{\text{max}} \\
0 \leq j \leq j_{\text{max}} & j\cdot1\text{sec} \leq a & j\cdot1\text{sec} \leq d
\end{cases}$$ (7)

$$TM \supseteq A_{TM}(SP)$$ (8)

where $\hat{y}$ is the optimal value of driving strategy parameters; $EC$ is the objective function to be minimised which expresses energy consumption; $S_y$ is the feasibility set of vector $y$ whose elements have to satisfy conditions in eqn (7); $A_{TM}$ is a function providing the simulated timetable depending on the speed profile. In particular, eqn (8) imposes that the speed profile has to generate a simulated timetable which has to be included in the planned timetable (described by vector $TM$) which is an input of the problem. Indeed, neglecting the timetable constraint (i.e. eqn 8) yields as an optimal solution the stop of convoys which represents a trivial condition of zero energy consumption.

Figure 2 describes the optimisation problems in terms of data framework, where the Optimisation Algorithm ($OA$) represents the numerical procedure for determining iteratively optimal values of driving parameters. In particular, bold
lines highlight differences from the System Simulation Model (SSM), described by eqn (1), and the optimisation problem described by eqn (5).

The second aim of the paper is to compare different signalling systems in terms of energy consumption provided for each of them by means of the solution to problem (5). Indeed, the signalling system strongly influences the capacity on a railway infrastructure (as shown by Dicembre and Ricci [26]). Hence, variation in the signalling system on a given railway infrastructure may yield different reliability of a given timetable or headways, i.e. the elapsed time computed at a given section between the running of two following trains on the same line. In general, the use of different signalling systems yields a variation in information on track occupancy, which may vary from point-based (in the case of a simple fixed block with physical signals on the line) to continuous data (in the case of a moving block with satellite tracking data and radio communication centre). Moreover, an increase in available information during train operations allows an increase in timetable reliability by keeping headways between two following trains fixed or a reduction in headway values by keeping the timetable reliability fixed.

Signalling systems considered in the paper are the ETCS levels that are designed to standardise European national signalling systems [27].

![Diagram](Figure 2: Framework of the energy optimisation model.)
3 Evaluation of optimal signalling systems

Definition of the optimal driving strategy allows, for any considered signalling system, minimum energy consumption to be determined. However, in order to compare different signalling systems, we also need to take into account other indicators which express an increase or decrease in implementation costs. In particular, we propose to analyse the following indicators:

- **variation in added rolling stock.** The percentage increase in enterprise costs in terms of maintenance and personnel for maintenance can be estimated as:

  \[
  \Delta_{RS}^{ETCS^{-i}} = \left(\frac{N_{RS}^{ETCS^{-i}}}{N_{RS}^{ETCS^{-0}}} - 1\right) \times 100
  \]

  \[
  \Delta_{TD}^{ETCS^{-i}} = \left(\frac{N_{TD}^{ETCS^{-i}}}{N_{TD}^{ETCS^{-0}}} - 1\right) \times 100
  \]

  where \( \Delta_{RS}^{ETCS^{-i}} \) and \( \Delta_{TD}^{ETCS^{-i}} \) is the variation in rolling stock (RS) and train drivers (TD) in the case of signalling system ETCS Level \( i \) (with \( i \in \{0, 1, 2, 3\} \)); \( N_{RS}^{ETCS^{-i}} \) and \( N_{TD}^{ETCS^{-i}} \) is the number of convoys (RS) and train drivers (TD) with an ETCS Level \( i \). These quantities are estimated by means of the SSM in the case of strategy TO;

- **variation in train crew.** The variation in operating personnel (including train conductors) needed for train operation may vary for different signalling systems. It can be estimated as:

  \[
  \Delta_{TC}^{ETCS^{-i}} = \left(\frac{N_{TC}^{ETCS^{-i}}}{N_{TC}^{ETCS^{-0}}} - 1\right) \times 100
  \]

  where \( \Delta_{TC}^{ETCS^{-i}} \) is the variation in train crew (TC) with an ETCS Level \( i \). These quantities are depending on the SSM (for the TO strategy) and the European ETCS legislation [27];

- **variation in generalised costs.** The generalised cost of users can be defined as a weighted sum of travel times and monetary costs (see [28]). In particular, since monetary costs (i.e. tickets) are constant for a fixed trip (generally ticket value depends on the origin and destination stations) and the variation in time on board for different signalling systems is minimal compared to the average travel time, the real variations in generalised costs can be identified in terms of waiting times. Hence, this indicator can be calculated as:

  \[
  \Delta_{GC}^{ETCS^{-i}} = \text{VOT}_{w} \cdot \left(\frac{T_{w}^{ETCS^{-i}}}{T_{w}^{ETCS^{-0}}} - 1\right)
  \]

  where \( \Delta_{GC}^{ETCS^{-i}} \) is the variation in generalised costs (GC); \( \text{VOT}_{w} \) is the monetary value of waiting times; \( T_{w}^{ETCS^{-i}} \) is the average waiting time of travellers with an ETCS Level \( i \). These quantities are estimated by means of SSM (applied for the TO strategy);

- **variation in train running times.** The variation of running time of trains by comparing the Energy Saving (ES) with the Time Optimal (TO) strategy can be estimated as:

  \[
  \Delta_{RT}^{ETCS^{-i}} = \left(\frac{R_{ES}^{ETCS^{-i}}}{R_{TO}^{ETCS^{-0}}} \cdot R_{TD}^{ETCS^{-i}} - 1\right) \times 100
  \]

  where \( \Delta_{RT}^{ETCS^{-i}} \) is the variation in train running time in the case of signalling
system ETCS Level $i$; $RT_{ES}^{ETCS-i}$ [ $RT_{TO}^{ETCS-i}$ ] is the train running time in the case of $ES$ [$TO$] strategy with an ETCS level $i$. It is worth noting that, in general, the $RT_{TO}^{ETCS-i}$ does not depend on signalling system, i.e. $RT_{TO}^{ETCS-i} = RT_{TO} \forall i$. These quantities are estimated by means of the SSM;

− variation in energy consumption. The variation of energy consumption by comparing $ES$ and $TO$ strategy can be estimated as:

$$
\Delta_{E}^{ETCS-i} = \left( E_{ES}^{ETCS-i} - E_{TO}^{ETCS-i} \right) / E_{TO}^{ETCS-i}
$$

where $\Delta_{E}^{ETCS-i}$ is the variation in energy consumption with an ETCS Level $i$; $E_{ES}^{ETCS-i}$ [ $E_{TO}^{ETCS-i}$ ] is the train energy consumption in the case of $ES$ [$TO$] strategy with an ETCS level $i$. It is worth noting that, in general, the $E_{TO}^{ETCS-i}$ does not depend on signalling system, i.e. $E_{TO}^{ETCS-i} = E_{TO} \forall i$. These quantities are estimated by means of the Energy Consumption Estimation ($ECE$);

− variation in energy consumption in disturbed conditions. With the assumption of disturbed conditions, such as an unpredictable reduction of maximum speed, variation of energy consumption can be estimated as:

$$
\Delta_{E,\text{stoch}}^{ETCS-i} = \left( E_{ES,\text{stoch}}^{ETCS-i} - E_{ES}^{ETCS-i} \right) / E_{ES}^{ETCS-i}
$$

where $\Delta_{E,\text{stoch}}^{ETCS-i}$ is the variation [the value] in train energy in the case of disturbed condition once implemented a $ES$ strategy. These quantities are estimated by the $ECE$.

4 Numerical applications

The proposed method was applied in the case of the Cumana railway, a suburban railway that connects the centre of Naples (in southern Italy) with its western hinterland. Although the route between Bagnoli and Arco Felice stations is currently a single track line, since expansion to double track has been both planned and funded, all applications were implemented as if the whole line were double track. Currently, the signalling system is based on automatic blocking sections where a blocking section is the track between two successive stations. Moreover, stations are equipped with fixed signals for station entrance and for booking the next blocking section. The rolling stock consists of electric trains (ET-400) based on a two fixed-block composition.

Numerical applications consist in implementing model (5) subject to constraints (6), (7) and (8) for different signalling systems (from ETCS level 0 to level 3) by applying $TO$, $ES$ and $ES$ with disturbed condition strategies. For all signalling systems we calculated performance indexes described in the previous section. Moreover, by means of a travel demand model calibrated in the regional context of the suburban railway in question, the increase in passengers was calculated by means of consolidated estimation techniques [28]. Details of simulations are summarised in table 1; fig. 4 shows the traffic diagram with blocking time under $TO$ and $ES$ strategies in the case of ETCS level 0; fig. 5 provides speed profiles for different signalling systems.
In terms of simulations, it is worth noting that ETCS level 0 requires at least one free section between two occupied sections, while for ETCS levels 1 and 2 two consecutive sections may be occupied by two different convoys. Hence, ETCS level 2, despite having more information on board, permits the same frequency as level 1 since they have the same block section constraint (one train per block section). Finally, in the case of ETCS level 3, it is possible to increase service frequency considerably by overcoming the traditional framework of fixed block sections.

![Cumana suburban railway](image)

**Figure 3:** The Cumana suburban railway.

<table>
<thead>
<tr>
<th></th>
<th>ETCS level 0</th>
<th>ETCS level 1</th>
<th>ETCS level 2</th>
<th>ETCS level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum service frequency</td>
<td>15 trains/h</td>
<td>20 trains/h</td>
<td>20 trains/h</td>
<td>30 trains/h</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-0}}$</td>
<td>–</td>
<td>+ 31 %</td>
<td>+ 31 %</td>
<td>+ 100 %</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-1}}$</td>
<td>–</td>
<td>+ 32 %</td>
<td>+ 32 %</td>
<td>+ 98 %</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-2}}$</td>
<td>–</td>
<td>–</td>
<td>– 35 %</td>
<td>– 2 %</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-3}}$</td>
<td>–</td>
<td>–</td>
<td>– 0.04 €/trip</td>
<td>– 0.08 €/trip</td>
</tr>
<tr>
<td>Increase in travel demand</td>
<td>–</td>
<td>+ 0.6 %</td>
<td>+ 0.6 %</td>
<td>+ 1.8 %</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-0}}$</td>
<td>+ 3.1 %</td>
<td>+ 3.7 %</td>
<td>+ 4.3 %</td>
<td>+ 5.1 %</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-1}}$</td>
<td>– 4.6%</td>
<td>– 6.1%</td>
<td>– 9.4%</td>
<td>– 14.3%</td>
</tr>
<tr>
<td>$\Delta_{\text{TCS-3}}$</td>
<td>+ 7.4%</td>
<td>+ 6.8%</td>
<td>– 39.6%</td>
<td>– 39.6%</td>
</tr>
</tbody>
</table>

It should be pointed out that an increase in service frequency requires an increase in the number of convoys, hence in the number of train drivers, while the number of conductors depends both on the number of convoys (i.e. service frequency) and on the signalling system, since the more communication is available on board the fewer conductors are required. Moreover, an increase in service frequency generates a decrease in passenger waiting times at the stations and hence a decrease in perceived generalised costs.

In the definition of the timetable for a railway line, it is necessary to use a running time reserve in order to reduce timetable stochastic perturbations and the propagation of delay. European railways generally adopt a running time reserve amounting to 3–7% of the minimum running time. Hence, as shown in fig. 1, the planned running time is equal to the sum of the real travel time and the residual running time reserve.
Assuming for our application, in the case of ETCS level 0, a running time reserve of 7%, it is possible to consider a part of this time for implementing an \( ES \) strategy instead of a traditional \( TO \) strategy. Simulation results provided a higher value of \( \Delta^{ETCS-i}_{RT} \) in the case of ETCS level 3 because continued communication is possible and the management of delays is easier with respect to the ETCS with lower levels.

Finally, we applied stochastic disturbance by imposing a breakdown of a pre-fixed convoy consisting in a 70% reduction in maximum speed and analysing effects on the whole planned timetable. In this case the relevant result is that for ETCS levels 0 and 1, the breakdown of a single convoy will generate a great increase in energy consumption because following trains are unable to adapt their driving profile to the faulty train. On the other hand, in the case of ETCS levels 2 and 3, following trains are able to adapt their speed to the faulty train by minimising the increase in energy consumption.

![Figure 4](image1.png)  
**Figure 4:** Traffic diagram with blocking time under TO (thin) and ES (bold) strategies in the case of ETCS level 0.

![Figure 5](image2.png)  
**Figure 5:** Speed profiles for different signalling systems.

## 5 Conclusions and research prospects

Results obtained by applying the proposed approach in the case of a real rail network have shown the effects of the different signalling systems on user and train operator costs. In particular, in the case of a regular service, for a given track and a given timetable, the ‘real’ speed profile related to a generic driving strategy depends on the signalling system and driver’s attitude. Hence, the use of Automatic Train Control (\( ATO \)) systems provides the possibility to develop a
standard driving strategy to reduce energy consumption. Such aspects become more significant in the case of non-regular services when the transfer of information could allow the impacts of delays and energy consumption to be minimised. Indeed, as in the case of ETCS levels 2 and 3, the speed profile may be adapted in real time according to line conditions.

For future research, we propose to analyse the relation between timetable line capacity and terminal station capacities since both factors can determine the minimum headway. Indeed, part of the time spent at terminal stations could be used as additional running time, which obviously depends on the signalling systems used, to implement energy-saving strategies.

Acknowledgements

This paper is partially supported under research projects PON – SFERE grant no. PON01_00595 and PON – DIGITAL PATTERN grant no. PON01_01268.

References


