

Analysis of the interaction between travel demand and rail capacity constraints

L. D'Acerno¹, M. Gallo², B. Montella¹ & A. Placido¹

¹*Department of Transportation Engineering,
'Federico II' University of Naples, Italy*

²*Department of Engineering, University of Sannio (Benevento), Italy*

Abstract

In urban contexts, the adoption of policies to promote the use of public transport systems represents a useful tool for decision-makers to reduce the environmental impact of private car use. Especially in high-density contexts most travel demand can be satisfied efficiently by means of high-quality rail systems. However, in the event of breakdowns, since faulty trains cannot usually be overtaken and their removal could pose extreme difficulties especially in metropolitan systems with two separate tunnels, re-establishing the regular service could involve inconveniently long travel times. Hence, emergency management has to take into account effects on travel demand. In this framework, we analyse such effects for different levels of degraded services in order to define the best strategy to adopt to minimise user discomfort. We extend ideas proposed elsewhere in the literature by introducing capacity constraints of rail vehicles in order to provide more realistic simulated effects. Finally, we describe the application of the proposed approach in the case of the Naples metro system.

Keywords: travel demand analysis, traffic assignment models, capacity constraints, rail passenger systems, public transport management, microsimulation approach.

1 Introduction

In any context, the management of transportation systems is a key issue which can affect both life quality and economic development. In large urban agglomerations, an efficient public transport system can help abate the negative externalities of private car use (such as congestion, air and noise pollution,



accidents, fuel consumption) without excessively penalising user travel times or zone accessibility.

Moreover, high-density contexts represent the ideal framework in which to adopt rail systems. Although they require greater construction, operating and maintenance costs than other public transport systems (such as buses, trolleybuses and taxis), high performance stemming from the use of exclusive lanes, the constrained drive and the signalling systems allows rail systems to achieve lower unit costs per seat-km (i.e. vehicular capacity multiplied by travel distances) or per carried passenger-km (i.e. travel demand multiplied by travel distances). Likewise, in the case of rail systems, externalities such as pollution or fuel consumption are also lower than those of other public transport systems.

Positive performance in terms of maximum travel speed or reduced headway between two successive convoys is partly offset by greater vulnerability to system failure. Indeed, in the case of breakdowns, since the faulty train cannot generally be overtaken and could be extremely hard to remove especially in metropolitan systems with two separate tunnels, re-establishing regular rail schedules could entail very substantial passenger delays. Hence, in dealing with emergencies, rail network managers have to make due allowances for any effects on travel demand.

As shown by Abril *et al.* [1], Gibson [2] and Lindner [3], the performance of rail systems and their related capacities has mainly been analysed by neglecting effects on travel demand: an extensive state-of-the-art analysis of rail capacity was proposed by [1], [2] and [3] examined in detail the problem of determining rail capacity as the maximum number of trains which can travel along a line or in a network.

Among the first papers to actually consider that the main purpose of a rail system is to satisfy traveller requirements were the contributions of Hamdouch *et al.* [4], Zheng *et al.* [5] and Kanai *et al.* [6]. In particular, [4] proposed an assignment model that differentiates the discomfort level experienced by sitting and standing passengers in the case of public transport systems, [5] provided the definition of capacity reliability of a rail network and developed a model for calculating it, while [6] proposed an algorithm for reducing user waiting times in the case of a rail (multi-line) network by modifying train timetables.

Recently, Mazzeo *et al.* [7] and Quaglietta *et al.* [8], neglecting capacity constraints, 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 failures of rail systems. Although these contributions might be considered an innovative approach to rail system analysis because they adopt a multi-objective approach in evaluating operational strategies, their capacity assumptions allow each user to board the first arriving train. They overlook the fact that in failure contexts some trains might not have enough space to accommodate all those boarding. Therefore, simulated failure effects tend to be calmed more rapidly than in real cases.

It is worth pointing out that in any design or real-time management problem it is necessary to have simulation models that allow network performances and

features to be defined for each alternative project or each management strategy. These models can be classified into macroscopic, mesoscopic or microscopic according to the assumption on the level of detail considered. Macroscopic simulation models (see, for instance, Kettner and Sewcyk [9] or Prinz *et al.* [10]) 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). Likewise, mesoscopic simulation models (as shown in Marinov and Viegas [11]) 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 (as described in Nash and Huerlimann [12] or Siefer and Radtke [13]) represent the system elements (such as signalling systems, radiuses 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 the adoption of a microscopic model requires the solution of a system of differential equations, for our purposes we may adopt a numerical approach. In this regard, Nash and Huerlimann [12] adopt OPENTRACK [14], developed by the Swiss Federal Institute of Technology (ETH) of Zurich, as micro-simulation software.

Our proposal is to analyse effects on travel demand of different levels of degraded services in the case of rail system failure in order to define the best strategy to adopt in order to minimise user discomfort. Indeed, the only reason which justifies (or at least should justify) the construction of a rail system is to satisfy travel demand with a high level of service. In particular, our proposal represents an extension of papers [7] and [8] by introducing capacity constraints of rail vehicles in order to provide more realistic simulated effects. Finally, we propose to adopt OPENTRACK [14] for integrating the system of differential equations in the case of a microscopic simulation approach.

The paper is organised as follows: Section 2 describes the framework of the proposed approach for defining the best strategy in the case of rail system failure; Section 3 applies the strategy in a real case, namely Line 1 of the Naples metro system; finally, conclusions are drawn and research prospects summarised in Section 4.

2 Framework of the proposed approach

Simulation of a rail system in the case of failure can be obtained through the interaction of four kinds of simulation models: failure, service, supply and demand.

The *failure simulation model* provides performance of the rail systems related to each possible breakdown. Output of this model may consist, for instance, in reduction in maximum speed or the unavailability of a train or a track section. This model is based on the cause-effect relation between the faulty element and the operation of all systems. Details on the management of breakdowns are

analysed by RAMS (Reliability, Availability, Maintainability and Safety) procedures as shown by Mazzeo *et al.* [7] and Quaglietta *et al.* [8].

The *service simulation model* describes rail system performance depending on rail infrastructures, signalling system, rolling stock, timetable and travel demand. Under the assumption of a microsimulation approach, this model, whose implementation is conditioned by outputs from the failure simulation model, can be solved through a system of differential equations whose numerical solution can be tackled by means of suitable commercial software. In our case, we propose to adopt OPENTRACK [14] by adding an external tool for simulating the extension of dwelling time at stations due to an increase in travel demand.

The *supply model* provides user generalised costs on all transportation systems in the analysed area depending on the outputs from service and travel demand models. Details on these kinds of models can be found in Cascetta [15] and D'Acerno *et al.* [16].

Finally, the *travel demand model* imitates the user behaviour conditioned by the supply and service simulation models. In particular, in the case of failure of the rail system with the assumption of capacity constraint there are two kinds of travel demand models: a *pre-platform model* and a *platform model*. The former describes user choices in the case of regular service. Generally, as shown by Cascetta [15], this model can be formulated as the product of four sub-models:

- an *emission model*, which simulates the choice of whether or not to travel in a time period;
- a *distribution model*, which provides the choice of going towards a generic place;
- a *mode choice model*, which provides the choice of travelling by using a generic means of transport such as private car, public transport or simply the pedestrian system;
- a *path choice model*, which provides the route choice and, in the case of public transport, also line choice and related boarding/alighting stops/stations.

The results of these four models can be obtained by solving a fixed-point model as shown by Cantarella [17].

It is worth noting that the pre-platform model provides user flows on each rail platform. However, when a fault occurs, it is necessary to integrate the above four-stage model with another model which takes the non-ordinary condition into account. This model, indicated as *platform model*, analyses for each approaching train if the residual capacity (which is equal to the train capacity minus the on-board passengers plus the alighting passengers) is greater than the number of boarding passengers. If this condition is not satisfied, only a portion of travel demand, equal to the residual capacity, is able to board the train with a *First In – First Out* (FIFO) approach, while the surplus has to wait for the following train. Obviously, the FIFO approach has to take into account that on the same platform passengers may have different destinations and, therefore, different alighting stations. However, the problem can be solved by adopting an

extension of Cantarella's fixed-point model [17]. Therefore, the travel demand model can be synthetically formulated as a double fixed-point problem.

By combining an intervention strategy (such as driving the faulty train on a maintenance track) with the four simulation models we can determine, for each failure scenario, both rail system performance and passenger generalised costs. Therefore, by analysing for each failure scenario all possible intervention strategies, which are generally limited in number, we may select the best for minimising impacts.

3 Application on a real scale network

The proposed methodology was applied to Line 1 of the Naples metro system in southern Italy. This line, operated by METRONAPOLI [18], consists of 15 stations, as shown in Figure 1. Importantly, the line in question consists of two services:

- a metro service between Piscinola and Dante (indicated in Figure 1 with the continuous line);
- a shuttle service between Dante and Università (indicated in Figure 1 with the dotted line), which is provided by using a single track and a single train.

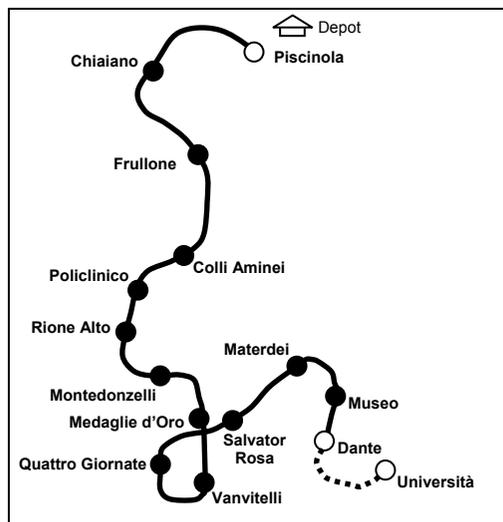


Figure 1: Line 1 of the Naples metro system (Italy).

Therefore, since the two services are completely independent and the train depot is located next to Piscinola, we analyse only the metro service, neglecting the shuttle.

In particular, we analysed failure scenarios during the morning peak-hour (i.e. 7.00–9.00) for simulating when discomfort effects are greatest. Obviously we considered a wider time period for analysing network loading (people and trains

on the network at 7.00 generally started before) and discomfort duration (failure effects could last also beyond the peak-hour). We thus considered the time period between 6.00 am and 2.00 pm.

Each train consists of 3 electric multiple units with a total capacity of 1251 passengers/train. The line extension is about 14 kms, and the timetable in terms of headways is:

- 12 minutes (i.e. 5 trains/hour) between 6.00 am and 7.00 am;
- 7 minutes (i.e. 8.6 train/hour) between 7.00 am and 9.00 am;
- 10 minutes (i.e. 6 trains/hour) between 9.00 am and 2.00 pm.

In order to implement the operational strategy analysis it is first necessary to calibrate the supply and demand models.

3.1 Supply model definition

This phase consisted in defining for each section the maximum speed between two successive stations, and was implemented by adopting the following steps:

- Step 1: survey on the field of travel times between two successive stations for different runs. Therefore, for the entire network, we obtain the matrix \mathbf{T} , of dimensions $(n_{TrackSections} \times n_{ObservedRuns})$ where $n_{TrackSections}$ is the number of track sections (i.e. the section of track between two successive stations) and $n_{ObservedRuns}$ is the number of observed runs. A generic element of matrix \mathbf{T} , indicated as T_{ij} , represents the travel time on the i -th track section at the j -th run;
- Step 2: calculation of average travel times between two successive stations, indicated as vector \mathbf{Tav} , of dimensions $(n_{TrackSections} \times 1)$, being calculated as follows:

$$\mathbf{Tav} = \frac{\mathbf{T} \cdot \mathbf{1}}{n_{ObservedRuns}} \quad (1)$$

where $\mathbf{1}$ is a vector of dimensions $(n_{ObservedRuns} \times 1)$, all of whose generic elements are equal to 1;

- Step 3: use of the uniform acceleration assumption for calculating the initial maximum speed for each section between two successive stations, indicated as $V_{max,i}^0$, such that the travel time is equal to the average value determined in the previous step. Indeed, the velocity-time graph with the assumption of uniform acceleration, as shown in Figure 2, can be expressed by means of a trapezium (as shown in Figure 2a) or a triangle (as shown in Figure 2b) according to trip duration and the space to be covered. In particular, the slope of these graphs is the acceleration and the area under the curve is the space. Therefore, by fixing the starting and the braking accelerations of considered trains, the maximum speed

of a section between two successive stations may be calculated by solving the following equation:

$$D = \int_0^{T_{av}} \mathbf{v}(V_{max}^0, \mathbf{x}) \cdot d\mathbf{x} \tag{2}$$

where \mathbf{D} is the vector of distances, of dimensions $(n_{TrackSections} \times 1)$, whose generic element D_i is the length of the i -th section; $\mathbf{v}(\cdot)$ is the velocity vector function, of dimensions $(n_{TrackSections} \times 1)$, whose generic element $v_i(\cdot)$ is the velocity function associated to the i -th section whose shape is indicated in Figure 2; V_{max}^0 is the vector of the initial maximum speeds, of dimensions $(n_{TrackSections} \times 1)$, whose generic element $V_{max,i}^0$ is the initial maximum speed in the i -th section; \mathbf{x} is the vector of integration variables, of dimensions $(n_{TrackSections} \times 1)$, whose generic element x_i is the integration variable in the case of the i -th section. Since the above assumptions allow function $v_i(\cdot)$ to depend only on variables related to the same track section, eqn. (2) can be decomposed into $n_{TrackSections}$ independent equations, that is:

$$D_i = \int_0^{T_{av_i}} v_i(V_{max,i}^0, x_i) \cdot dx_i \quad \forall i \tag{3}$$

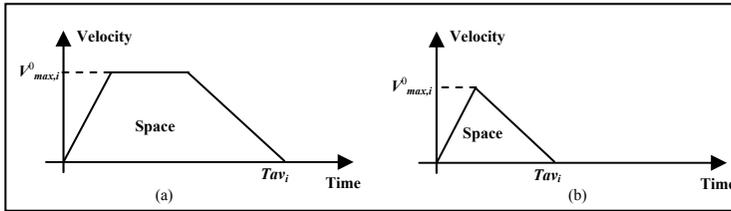


Figure 2: Velocity-time graphs with the uniform acceleration assumption.

Step 4: having defined the initial maximum speed vector (V_{max}^0), it is necessary to calculate the ‘real’ maximum speed vector, indicated as \hat{V}_{max} , so that it is closest to the initial speed vector V_{max}^0 and, jointly, generate by means of the service simulation model ‘real’ travel times closest to average travel times calculated at Step 2, that is:

$$\hat{V}_{max} = \arg \min_{V_{max} > 0} (Z_1(V_{max}, V_{max}^0) + Z_2(T^*, T_{av})) \tag{4}$$

subject to:

$$T^* = A(V_{max}) \tag{5}$$

where V_{max} is the maximum speed vector, whose optimal value is \hat{V}_{max} , T^* is the vector of ‘real’ travel times, being calculated by means of the

service simulation model synthetically described by function A in eqn. (5); Z_1 and Z_2 are scalar functions that can be considered respectively as ‘distance’ measures of the maximum speed vector (V_{max}) from the initial speed vector (V_{max}^0) and of the ‘real’ travel time vector (T^*) from the average travel times (Tav).

Problem (4) can be easily solved by adopting a projected gradient algorithm (see details in Cascetta [15]). The system simulation model, i.e. eqn. 5, can be solved by using OPENTRACK [14].

3.2 Demand model definition

This phase consisted of generating the pre-platform model so that the related platform model could be implemented for each failure scenario. This stage was implemented first of all by counting passenger flows at stations, and then by implementing the well-known procedure of travel demand flows’ estimation using traffic counts. Details of the procedure can be found in Cascetta [15].

3.3 Operational strategy analysis

In this phase we considered two failure scenarios whose direct effects are a degradation in train performance: at 7.00 in the Chiaiano station, the second station after the depot, a convoy experiences a breakdown which limits the maximum train speed in each scenario to 80% or 20%. Since there are only two maintenance tracks on the network (at Medaglie d’Oro and Dante) the analysed solution strategies are:

- the train continues the service until Medaglie d’Oro and then, after unloading passengers on the platform, it is driven onto the maintenance track;
- the train continues the service until the following terminus, i.e. Dante, and is then driven onto the maintenance track;
- the train continues the whole service until the depot, i.e. Piscinola.

In all cases, we assumed that, as no substitute vehicle was available, the timetable had to be revised.

The results of the six strategies are compared (Tables 1 and 2) with the regular service scenario. A common result is that when a train breaks down, it increases the headway with the preceding convoy because it travels with a lower speed than the previous one. Likewise, the faulty train decreases the headway with the following convoy because the following train initially travels with a higher speed than the faulty one. Signalling systems tend to ensure that the following train then travels at the same speed as the faulty train. The slowing-down wave spreads progressively over all following trains.

In terms of travel demand, an increase in headways provides an increase in boarding passengers and a possible exceeding of train capacity. Likewise, a decrease in headways provides a decrease in passengers boarding trains. Therefore, the faulty train tends to be saturated by passengers, while the

Table 1: Strategy analysis with a 20% reduction in train performance.

		Regular service scenario	Strategy 1 (Medaglie d'Oro station)	Strategy 2 (Dante station)	Strategy 3 (Piscinola station)
user generalised costs [k€]		80.942	88.432	85.823	86.214
peak-hour time period	minimum headway [minutes]	7.00	5.55	5.02	3.17
	maximum headway [minutes]	7.00	14.00	14.00	14.00
weak-hour time period	minimum headway [minutes]	10.00	10.00	10.00	10.00
	maximum headway [minutes]	10.00	10.00	10.00	10.00

Table 2: Strategy analysis with an 80% reduction in train performance.

		Regular service scenario	Strategy 1 (Medaglie d'Oro station)	Strategy 2 (Dante station)	Strategy 3 (Piscinola station)
user generalised costs [k€]		80.942	103.931	136.106	192.524
peak-hour time period	minimum headway [minutes]	7.00	1.28	20.00	20.00
	maximum headway [minutes]	7.00	20.20	33.87	51.47
weak-hour time period	minimum headway [minutes]	10.00	10.00	4.80	1.92
	maximum headway [minutes]	10.00	10.00	6.02	5.47

following convoys tend to be vacant except for people who were unable to board the previous train.

When the faulty train is eliminated by the system (i.e. it reaches the maintenance track), if there are passengers on board, they have to alight onto the appropriate platform to board the following train. Therefore, train elimination will increase network performance, since trains are not yet constrained by the signalling system and are able to reach their maximum speed, but will produce a combined deterioration in service quality since user waiting times and vehicle crowding will increase.

Analysis of the application results shows that in the first scenario (Table 1), performance reduction does not lead to substantial increases in user travel times. Hence the best strategy in this case is to complete the service until Dante station then drive the faulty train onto the maintenance track. By contrast, in the case of the second scenario (Table 2), performance reduction is considerable and large increases are generated in user travel times. Therefore, the best strategy consists in excluding the faulty train from the system as soon as possible.

4 Conclusions and research prospects

In this paper we proposed a *Decision System Support* (DSS) for defining the best operational strategy in the case of rail system failure. We showed that, although effects on travel demand have often been neglected in the literature, they can profoundly affect analytical results.



We expanded on contributions by Mazzeo *et al.* [7] and Quaglietta *et al.* [8] by introducing capacity constraints in order to develop a more realistic model which might allow for the fact that, in a failure scenario, passengers might not be able to board the first arriving train since the boarding flow may be higher than the convoy residual capacity. However, this assumption made it necessary to split the traditional demand model into two sub-models:

- a pre-platform model, for simulating user choices in terms of boarding and alighting stations;
- a platform model, for simulating the boarding process on train arrival, conditional upon residual capacities.

Application on a real dimension network showed that it cannot be stated *a priori* (i.e. without any model implementation) which is the best operational strategy because it is necessary to solve a system of differential equations for defining vehicle performance and, at the same time, a double fixed-point problem for evaluating user generalised costs.

Finally, there appears scope for further research into applying the proposed DSS by considering a broad set of train breakdowns or other rail networks. Moreover, the method could be applied to draw up rail timetables by adopting a multimodal approach (Gallo *et al.* [19]).

References

- [1] Abril, M., Barber, F., Ingolotti, L., Salido, M.A., Tormos, P. & Lova, A., An assessment of railway capacity. *Transportation Research Part E*, **44(5)**, pp. 774–806, 2008.
- [2] Gibson, S., Allocation of capacity in the rail industry. *Utilities Policy*, **11(1)**, pp. 39–42, 2003.
- [3] Lindner, T., Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. *Journal of Rail Transport Planning & Management*, **1(1)**, pp. 49–57, 2011.
- [4] Hamdouch, Y., Ho, H.W., Sumalee, A. & Wang, G., Schedule-based transit assignment model with vehicle capacity and seat availability. *Transportation Research Part B*, **45(10)**, pp. 1805–1830, 2011.
- [5] Zheng, Y., Zhang, Z., Xu, B. & Wang, L., Carrying capacity reliability of railway networks. *Journal of Transportation Systems Engineering and Information Technology*, **11(4)**, pp. 16–21, 2011.
- [6] Kanai, S., Shiina, K., Harada, S. & Tomii, N., An optimal delay management algorithm from passengers' viewpoints considering the whole railway network. *Journal of Rail Transport Planning & Management*, **1(1)**, pp. 25–37, 2011.
- [7] Mazzeo, A., Mazzocca, N., Nardone, R., D'Acierno, L., Montella, B., Punzo, V., Quaglietta, E., Lamberti, I. & Marmo, P., An integrated approach for availability and QoS evaluation in railway systems. *Lecture Notes in Computer Science*, **6894**, pp. 171–184, 2011.



- [8] Quaglietta, E., D'Acerno, L., Punzo, V., Nardone, R. & Mazzocca, N., A simulation framework for supporting design and real-time decisional phases in railway systems. *Proc. of the 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington (D.C.), USA, art. no. 6082913, pp. 846–851, 2011.
- [9] Kettner, M. & Sewcyk, B., A model for transportation planning and railway network evaluation. *Proc. of the 9th World Congress on Intelligent Transport Systems*, Chicago (IL), USA, 2002.
- [10] Prinz, R., Sewcyk, B. & Kettner, M., NEMO: Network Evaluation Model for the Austrian railroad (ÖBB). *Eisenbahntechnische Rundschau*, **50(3)**, pp. 117–121, 2001.
- [11] Marinov, M. & Viegas, J., A mesoscopic simulation modelling methodology for analyzing and evaluating freight train operations in a rail network. *Simulation Modelling Practice and Theory*, **19(1)**, pp. 516–539.
- [12] Nash, A. & Huerlimann, D., Railroad simulation using OpenTrack. *Computers in Railways*, **9**, pp. 45–54, 2004.
- [13] Siefer, T. & Radtke, A., Railway simulation: key for better operation and optimal use of infrastructure. *Proc. of the 1st International Seminar on Railway Operations Modelling and Analysis*, Delft, The Netherlands, 2005.
- [14] OPENTRACK, <http://www.opentrack.ch>
- [15] Cascetta, E., Transportation systems analysis: models and applications. Springer: New York (NY), USA, 2009.
- [16] D'Acerno, L., Carteni, A. & Montella, B., Estimation of urban traffic conditions using an Automatic Vehicle Location (AVL) System. *European Journal of Operational Research*, **196(2)**, pp. 719–736, 2009.
- [17] Cantarella, G.E., A general fixed-point approach to multimodal multi-user equilibrium assignment with elastic demand. *Transportation Science*, **31(2)**, 107–128, 1997.
- [18] METRONAPOLI, <http://www.metro.na.it>
- [19] Gallo, M., D'Acerno, L. & Montella, B., The transit network design problem with elastic demand and internalisation of external costs: An application to rail frequency optimisation. *Transportation Research Part C*, **19(6)**, pp. 1276–1305, 2011.