This is author version of article published as:

Burdett, Robert L. and Kozan, Erhan (2006) Techniques for absolute capacity determination in railways. *Transportation Research Part B: Methodological* 40(8):pp. 616-632.

TECHNIQUES FOR ABSOLUTE CAPACITY DETERMINATION IN RAILWAYS

R. L. Burdett and E. Kozan
School of Mathematical Sciences, Queensland University of Technology,
PO Box 2434, QLD 4001, Australia

Abstract - Capacity analysis techniques and methodologies for estimating the absolute traffic carrying ability of facilities over a wide range of defined operational conditions are developed for railway systems and infrastructure in this paper. In particular we tackle numerous aspects on which the absolute carrying capacity depends that have not been previously and or significantly addressed before. These include the proportional mix of trains that travel through the system and the direction in which these trains travel, the length of trains, the planned dwell times of trains, the presence of crossing loops and intermediate signals in corridors and networks. The approaches are then illustrated in a case study.

Keywords: Capacity analysis, railways

1. INTRODUCTION

This paper which extends the work in Kozan and Burdett (2004 and 2005) is concerned with the quantification of capacity in more complex railway lines and networks. More specifically a general approach for determining *absolute capacity* is developed that is suitable for any railway system. For example the approach can be used to analyse complex railway networks consisting of uni-directional rail with high volumes of passenger trains that is typical of European railways. It can also be used to analyse long distance freight networks of bi-directional rail consisting of crossing (passing) loops and intermediate signals typical in North America or Australia. Any combination of these features can also be accommodated.

Primary motivation for this research occurs as a result of the additional demands that are to be placed upon railways in Australia to provide third party usage of infrastructure. The complete separation of railway infrastructure ownership and operation is also a potential reality in the near future which would result in even greater competition for railway infrastructure (see Ferreira (1997) for more information). To ensure fair and impartial access under these new regimes, railway operators must be able to clearly differentiate between free and used capacity. For example the railway operators will typically need to answer the question “does capacity exist in order to add an additional service to the existing timetable?” How much capacity exists (which is the focus of this paper) however must firstly be known if this is to be accomplished.

Railway capacity however is an elusive concept that is not easily defined or quantified. Difficulties include the numerous interacting/interrelated factors, the complex structure of the railway layout, and the magnitude of terminology required. In this paper *capacity* is defined as the maximum number of trains that can traverse the entire railway or certain critical (bottleneck) section(s) in a given duration of time. The term *absolute capacity* is defined and refers to a theoretical value (overestimation) of capacity that is realised when only critical section(s) are saturated (i.e. continuously occupied). It is an ideal case that is not necessarily possible in reality due to the collision conflicts that are caused on non-critical sections as a consequence of insufficient passing (overtaking) facilities. Alternatively, *actual* (*sustainable*) *capacity* is the amount that occurs when interference delays are incorporated on the critical section(s). Interference delays are enforced idle time in the system that results when collision conflicts are resolved. Both actual and absolute measures of capacity are required in practice for a variety of reasons. Fore mostly is that they allow traffic congestion to be quantified (for example as the relative difference between the two). Traffic congestion in any transportation system is an important issue, particularly its minimisation. Alternatively as an upper bound on capacity, the absolute measure can also be used for planning purposes. In some ways it is more robust than actual capacity approximations which are more statistical in nature.

Capacity is also not a unique value in railway scenarios. It may be different for each proportional mix of trains (train types). Similarly absolute capacity as defined in this paper may also be different for each proportional mix of trains. Hence capacity assessment can only be performed to decide whether the infrastructure can handle the intended traffic load. If there is sufficient capacity then a suitable train schedule/timetable may be developed otherwise the effort is unwarranted.

Capacity analysis is therefore an iterative process in which the modification of the train mix is a key step. For example, the planner would initially propose (or be given) proportions of each type of train in each direction for every section of rail. After evaluation the value of capacity and the actual numbers of each type of train are determined. These values are inspected and if satisfactory the process is complete. If they are not satisfactory then the process continues by evaluating mix alterations. A railway capacity analysis software tool (RCAT) has been developed to perform this process and is based upon the techniques developed in this paper.

Alternatively if there is insufficient capacity then additional infrastructure may be required and the capacity analysis techniques can then be used to determine where infrastructure should be placed to achieve the desired result.

We now review how capacity has been previously defined and quantified. The simplest definition and the most prevalent encountered in the literature is that the capacity of a single line is the total number of standard train paths that can be accommodated across a critical section in a given time period (i.e. the time period duration divided by the trains sectional running time), where a standard train is defined as the most prevalent type to traverse the corridor. This implies that a single bottleneck section limits the total flow of trains throughout the entire corridor and consequently the analysis is called a bottleneck approach. As an estimation of capacity, this approach is especially useful as an indicator of where additional infrastructure could best be placed. It is also useful because of its mathematical simplicity. Nonetheless, in practice rail corridors must accommodate trains of varying type and number, which travel in each direction in varying numbers. This definition also assumes that there is a prevalent type and that any train type can be converted to a standard train type without affecting the overall capacity level. It also assumes that sectional running times in each direction are equal and on bi-directional lines the traffic flow is equal in each direction. This bottleneck analysis also does not address the many other operational factors, such as the stopping protocols, the lengths of trains, the dwell times. Nor does it address the determination of capacity in more complex railway systems which contain interrelated lines. In contrast the approach developed in this paper does include each of these aspects.

However, the schedule that actually gives such an output in total trains if at all "possible" is not known from such a bottleneck analysis and must be solved for separately. It should also be noted that this capacity approach does not indicate how efficiently the track infrastructure is utilised under a given schedule at each instant of time nor does it indicate how to place additional trains into the schedule.

The capacity of a rail line has also been evaluated using the delays encountered by trains under different operating assumptions (Assad 1980). Petersen (1974) for example develops expressions for average train delay, which are then incorporated into an expression for mean sectional running times. The usual bottleneck analysis can then be performed to give an alternative measure of capacity. The approach for modelling delays however assumes that trains in each class are uniformly distributed over the time-period and that there is equal traffic in each direction, which again is unrealistic. Delays caused by multiple train interactions are also not handled. The delays encountered in railways have also been addressed by other researchers using queuing theory and statistical techniques and examples include Petersen (1975), Kraft (1983), Hall (1987), Greenberg (1988), Carey and Kwiecinski (1994), Higgins, Kozan and Ferreira (1995), and Higgins and Kozan (1998). However, no simple analytic expressions result from these researches. The numerical investigations for these approaches are also generally compared to the results provided from simulation. Comparisons among the different approaches have not been performed to the author's knowledge and hence it is not known which of these approaches is superior.

More recently De Kort et al (2003) provided a new approach for determining railway infrastructure capacity. Their approach is based upon the analysis of a generic infra element (building block), which allows many different combinations of infrastructure to be analysed. The approach however can not explicitly distinguish between different train types, although it can incorporate this aspect by considering the distribution of travel/release times weighted by the probability that a train is of a given type. It is also based upon the bottleneck analysis approach. The approach was illustrated on a high speed railway line being built in the Netherlands.

Other definitions of capacity exist, however these definitions have nothing to do with traffic volume. They are based upon the carrying capacity of trains in terms of passengers and freight, or the ability of the corridor to contain as many trains as possible at any moment of time.

Because railway terminology varies in each country, the following paragraph provides an explanation of the various basic railway components that are referred to in this paper so that any misconceptions that may occur are minimised. For readers that are unfamiliar with railways, this paragraph also provides a basic introduction. Other important entities and components will be defined within the paper as they are required.

A *railway corridor* is generally a single serial line (track) that is made up of one or more sections of specific length that are sequentially traversed. A *section* (segment) is the length of rail between two locations and a *location* is any fixed point of reference along the rail such as the end of a section of rail, the start or end of a crossing loop, crossover or junction, or the position of a signal device. Traffic movement on each line (track) can potentially be in one or both directions, that is, traffic is *uni-directional* or *bi-directional*. One train only however may generally occupy a section of rail at any given time (*section occupation condition* (SOC)) for safety reasons however it is physically possible for a train to follow another train (albeit separated by a headway interval) onto a section if it is long enough. Technological improvements (whatever they may be) may allow the SOC to be relaxed in the future. A *railway system* is a single corridor or a collection of separate and or interrelated corridors. A railway line that branches outwards and hence inwards in the opposite direction is referred to as *non-serial*. Hence a network of interrelated corridors is a non-serial track. It should also be noted that parallel tracks with the same topography (i.e. sections are placed in the same position on each track) have also been referred to as corridors. Train types that use the railway are electric or diesel powered and may be used for moving passengers or freight. The total journey time for a train is referred to as it *transit time*, and the time it takes to traverse a given section is the *sectional running time* (SRT). The total time spent by a train on a section however, is the *section occupation time* (SOT) which may include pre planned dwell (*dwell*) time and scheduled/unscheduled delays. A *trains path* is a passage through the system from one *input-output* (IO) point to another. A group of trains that travel in the same direction (sequentially) on the railway without being disturbed by traffic in the other direction is called a *fleet*.

The content of the remainder of the paper is as follows. Techniques for simpler railway lines are firstly developed in section 3 after necessary terminology has been given in section 2. In section 4, the approaches in section 3 are extended for general railway networks. Following this a case study which illustrates the techniques proposed in this paper is provided. Lastly our results and accomplishments are summarised in the conclusions.

2. NOMENCLATURE

i: Train type index. The set of trains is $I = \{1, 2, ...\}$.

T: Time period duration

l,k: Location index.

$\rightarrow$: This symbol signifies uni-directional travel between two locations, for example $l \rightarrow k$. In contrast "$-$" is used to signify a non directional or generic link as occurs when defining a corridor.

Φ: The set of input/output (IO) points which are locations where trains may enter and leave the network. Any location may be defined as an IO point.

□ : The set of valid sections in the railway network.

□ : The set of valid corridors in the railway network. Note: $\square = \{(l,k) : l,k \in \Phi\}$

CS_{l-k}: The critical section of the corridor that lies between *l* and *k* (i.e. corridor $l-k$).

C_{abs}^{l-k}: The absolute capacity of corridor $l-k$. Alternatively it is the capacity of section $l-k$ because the smallest corridor is a single section. To avoid confusion however the absolute capacity of a section is redefined as $C_{abs}^{(l,k)}$. Therefore, $C_{abs}^{l-k} = \min_{(l',k') \in S_{l,k}} \left(C_{abs}^{(l',k')} \right)$.

$\Delta_i^{l \rightarrow k}$: Total dwelling time of train type *i* on corridor $l-k$ when travelling in the direction of *k*. Note: $\Delta_i^{l \rightarrow k} = \sum_{\forall l' \in P_{l-k}} \left(\Delta_{i,l'}^{l \rightarrow k} \right)$ where $\Delta_{i,l'}^{l \rightarrow k}$ is the specific dwelling time at location l'.

$J_i^{l \rightarrow k}$: Total transit time of train type *i* on corridor $l-k$ when travelling in the direction of location *k*. Note: $J_i^{l \rightarrow k} = \sum_{(l',k') \in S_{l-k}} \left(\overline{SRT}_i^{l' \rightarrow k'} \right)$

η_i^{l-k}: Proportion of train type *i* on corridor $l-k$ where $(l,k) \in \square$.

$\mu_i^{l \rightarrow k}$: The proportion of train type *i* in the direction of *k* on corridor $l-k$. If $(l,k) \in \square$, then $\mu_i^{l \rightarrow k} + \mu_i^{k \rightarrow l} = 1$. Note that $\mu_i^{l \rightarrow k} = 0$ if movement from *l* to *k* is restricted (for example as occurs in uni-directional systems).

$L_i^{\text{train}}, L_{(l,k)}^{\text{section}}, L_{l-k}^{\text{corr}}$: Respectively the length of train *i*, section (*l*,*k*) and corridor $l-k$.

$SRT_i^{l\to k}$: The sectional running time of train *i* when travelling on section (l,k) in the direction of location *k* including the time lag required by the rear to pass the section boundary (i.e. $\left(1+L_i^{\text{train}} \big/ L_{(l,k)}^{\text{section}}\right)$ times standard SRT). Note that the time lag may be greater if dwell times exist and the length of the train is longer than the section. Techniques for calculating the correct departure time for the rear and associated time lag have been developed however are not addressed in this paper.

P_{l-k}, S_{l-k} : Set of locations and sections traversed on corridor $l-k$ respectively.

X_i^{l-k} : Number of trains of type *i* that use corridor $l-k$. Note: $X_i^{l-k} = x_i^{l\to k} + x_i^{k\to l}$

$x_i^{l\to k}$: Number of trains of type *i* that travel from location *l* to *k*. Note: $x_i^{l\to k} = 0$ if movement in the given direction is restricted.

In summary corridors and sections are referenced by two values namely the boundary locations. The ordering of these two values is used to define direction of travel which is not possible if a single index/label is used. If desired the $\to$ and $-$ notation may be replaced with a simple comma and brackets notation instead, for example (l,k). It does however explain the directional aspects more easily.

3. TECHNIQUES FOR RAILWAY LINES

In this section we extend the bottleneck approach (as described in the introduction) by incorporating a variety of additional factors. The improvements in this section relate to railway lines while more complex techniques for networks are presented in the next section. It should also be noted that unless otherwise mentioned all quantities are assumed to be real valued. Quantities such as the absolute capacity and the number of trains of each type can be real valued because partial traversal of a corridor is still utilisable capacity (i.e. it is not important that whole trains reach their destination in a given duration of time). The time it takes for trains to reach a specific position prior to the time period is also of no concern (i.e. a “steady state” is assumed.)

3.1. Train Mixes

The bottleneck analysis is usually performed in relation to a single train type. However in reality, the section will have to accommodate a variety of different trains with varying speeds. To remedy this, the concept of a *percentage train mix* is introduced, from which an *actual train mix* can be determined. Consequently the percentage of total traffic that consists of each train type is defined as the *proportional distribution*, while the percentages for travel in each direction are defined as the *directional distribution*. For every distinct proportional and directional distribution (i.e. percentage train mix), the capacity of the corridor will be potentially different. For a line with boundaries (*l*,*k*) and particular proportional and directional distributions, the proportion of traffic in each direction respectively is determined from the balance equation $\sum_i \left(\eta_i^{l-k}\left(\mu_i^{l\to k}+\mu_i^{k\to l}\right)\right)=1$ as $\sum_i \left(\eta_i^{l-k}\mu_i^{l\to k}\right)$ and $\sum_i \left(\eta_i^{l-k}\mu_i^{k\to l}\right)$ respectively.

A section of rail that is fully utilised is referred to as saturated, and this occurs when it is fully occupied. Under these conditions and for feasibility, the time-period must be greater than or equal to the section occupancy time (i.e. number of trains in each direction multiplied by their respective SRT in that direction). That is, $\sum_{\forall i}\left(x_i^{l\to k}SRT_i^{l\to k}+x_i^{k\to l}SRT_i^{k\to l}\right)\le T$. If the number of type *i* trains in each direction respectively is $\left(\eta_i^{l-k}\mu_i^{l\to k}\right)C_{abs}^{(l,k)}$ and $\left(\eta_i^{l-k}\mu_i^{k\to l}\right)C_{abs}^{(l,k)}$, then the substitution of these terms and the re-arrangement of the above equation gives the following expression for absolute capacity.

$$C_{abs}^{(l,k)}=\frac{T}{\sum_{\forall i}\eta_i^{l-k}\left(\mu_i^{l\to k}SRT_i^{l\to k}+\mu_i^{k\to l}SRT_i^{k\to l}\right)} \tag{1}$$

The denominator is the weighted average sectional running time and is defined as $\overline{SRT}^{l-k}$. It gives the average performance of trains in the given mix. When capacity is not fully utilised, free capacity (i.e. $F_i^{l\to k}$) in terms of a particular train type and for a particular direction can then be determined by subtracting the current occupancy level $\sum_{\forall i'}\left(x_{i'}^{l\to k}SRT_{i'}^{l\to k}+x_{i'}^{k\to l}SRT_{i'}^{k\to l}\right)$ from T and then dividing by $SRT_i^{l\to k}$. This allows the rail operator a "what if" capability to see whether capacity exists for additional trains to be run.

Absolute utilisation levels (not incorporating congestion and interaction effects) can also be computed by equation (2) and is the percentage time that an actual mix of trains utilises the section.

$$U_{abs}^{(l,k)}=\frac{\sum_{\forall i}\left(x_i^{l\to k}SRT_i^{l\to k}+x_i^{k\to l}SRT_i^{k\to l}\right)}{T} \tag{2}$$

Consequently $\left(1-U_{abs}^{(l,k)}\right)$ is the percentage capacity that is theoretically free on section (l,k). The overall utilisation of the corridor can then be calculated as the minimum utilisation level of any section, i.e. $U_{abs}^{l-k}=\min_{\forall(l',k')\in S_{l-k}}\left(U_{abs}^{(l',k')}\right)$.

3.2. *Signals and Reference Locations*

Equation (1) however assumes that crossing loops occur at the section boundaries thus allowing traffic to pass each other. However in practice this is not always the case. For example one or both section boundaries may have signal devices. Signal devices allow throughput in one direction to be increased, because it allows additional trains to safely occupy a section. However utilisation of capacity can be lost when consecutive trains travel in opposite directions as occurs when bi-directional flow is allowed. This is because trains can not pass each other at the section boundary. They can only pass each other at the nearest crossing loop or passing facility if one exists. Consequently *enforced headways* on that section must be incurred and these may be of considerable duration in some railways. An example of this is shown in Fig 1 on a time versus distance chart. The grey blocks are enforced headways. The train velocities are 100 km/h.

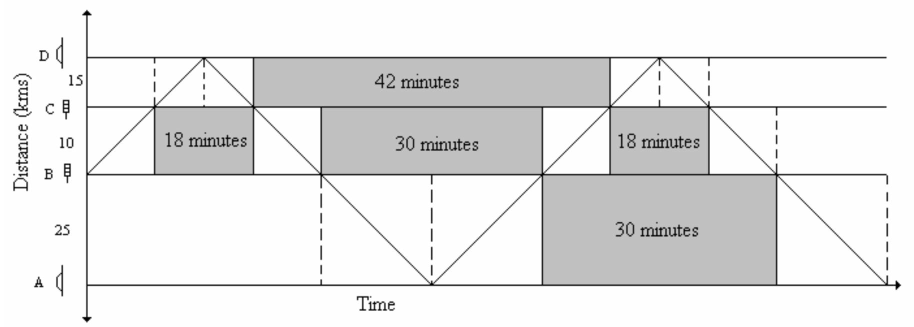


Figure 1. Enforced headways on a corridor containing signals

Therefore previously proposed capacity calculations may overestimate capacity levels on sections that are not bounded by crossing loops and a new approach is necessary. The approach taken in this paper in particular is based upon the observation that the number of enforced headways is proportional to the number of pairs of alternating trains. This is dictated by a particular sequence (i.e. train schedule), and is unknown from a capacity analysis viewpoint. Therefore, we can only define a range in which the absolute capacity will lie and depending on the level of fleeting (see introduction for definition), this value will be closer to one bound than the other. In particular, the absolute capacity will be greatest (i.e. the upper bound) when there are two fleets, one in each direction. Absolute capacity will be smallest when there are many fleets of one train (i.e. trains commonly alternate in the sequence in each direction).

The lower bound value is determined by noting that the worst sequence containing alternating trains is of length 2*z*+1 where $z = \min(\#up, \#dn)$. The remaining trains form a fleet. For example the two possible resulting sequences have the following format: (*d u d u d … d d …. d*) or (*u d u d u …. u u ….u*) where *u* and *d* refer to an up or down train respectively. Consequently there will be at most 2*z* enforced headways that must be included. The upper bound is determined by choosing the permutation of the two fleets (i.e. all up and then all down, or all down and then all up) that has the smallest enforced headway between them.

We now define $h^{(l,k)}$ as the enforced headway time on section (*l*,*k*) between traffic moving in opposite directions when there is no crossing loop at location *k*. Similarly, $h^{(k,l)}$ is the enforced headway on the same section but at the opposite section boundary. In particular the headway between an $l \to k$ and a $k \to l$ train is $h^{(l,k)}$, and the headway between a $k \to l$ and an $l \to k$ train is $h^{(k,l)}$. It should be noted that $h^{(k,l)}$ and $h^{(l,k)}$ are zero if location *l* and *k* respectively are crossing loops. The enforced headways are calculated as the sum of two weighted average travelling times, i.e. $h^{(k,l)} = \overline{SRT}^{l \to l'} + \overline{SRT}^{l' \to l}$.The weighted average travelling times to and from the nearest crossing loop location l' may be different depending on the particular proportional and directional distributions. Otherwise this value would just be twice the weighted average travelling time. The weighted average sectional running time in a particular direction is calculated as follows:

$$\overline{SRT}^{l \to k} = \left(\frac{1}{\sum_{\forall i} \left(\eta_i^{l-k} \mu_i^{l \to k} \right)} \right) \sum_{\forall i} \left(\eta_i^{l-k} \mu_i^{l \to k} SRT_i^{l \to k} \right) \tag{3}$$

The second part of this equation is the weighted average sectional running time for trains travelling in one direction only however since the sum of these percentages does not add to 100 percent, the value must be scaled by the bracketed value in the first part of the equation. The first part in particular is the inverse of the percentage traffic in the given direction. It should also be noted that:

$$\overline{SRT}^{\,l-k} = \sum_{\forall i}\left(\eta_i^{l-k}\mu_i^{l\to k}\right)\overline{SRT}^{\,l\to k} + \sum_{\forall i}\left(\eta_i^{l-k}\mu_i^{k\to l}\right)\overline{SRT}^{\,k\to l} \tag{4}$$

To derive equations for the lower and upper bounds on the absolute section capacity we note that for feasibility, the following inequalities (which represent section saturation conditions) must be satisfied for the lower and upper bound cases respectively:

$$\sum_{\forall i}\left(SRT_i^{l\to k}x_i^{l\to k} + SRT_i^{k\to l}x_i^{k\to l}\right) + \left(h^{(l,k)} + h^{(k,l)}\right)z \le T \quad \text{(LB only)} \tag{5}$$

$$\sum_{\forall i}\left(SRT_i^{l\to k}x_i^{l\to k} + SRT_i^{k\to l}x_i^{k\to l}\right) + \min\left(h^{(l,k)}, h^{(k,l)}\right) \le T \quad \text{(UB only)} \tag{6}$$

where z is the number of trains travelling in the direction with the least number trains and is computed as $z = \sum_{\forall i}\left(x_i^{l\to k}\right)$ if $\left(\sum_{\forall i}\left(\eta_i\mu_i^{l\to k}\right) \le \sum_{\forall i}\left(\eta_i\mu_i^{k\to l}\right)\right)$ and $\mathrm{z} = \sum_{\forall i}\left(x_i^{k\to l}\right)$ otherwise. We also know that $x_i^{l\to k} = \eta_i^{l-k}\mu_i^{l\to k}C_{abs}^{(l,k)}$ and $x_i^{k\to l} = \eta_i^{l-k}\mu_i^{k\to l}C_{abs}^{(l,k)}$ from section 3.1. After substitution of these values the resulting intermediate equations are as follows:

$$C_{abs}^{(l,k)}\sum_{\forall i}\left(SRT_i^{l\to k}\eta_i^{l-k}\mu_i^{l\to k} + SRT_i^{k\to l}\eta_i^{l-k}\mu_i^{k\to l}\right) + \left(h^{(l,k)} + h^{(k,l)}\right)z \le T \tag{7}$$

$$C_{abs}^{(l,k)}\sum_{\forall i}\left(SRT_i^{l\to k}\eta_i^{l-k}\mu_i^{l\to k} + SRT_i^{k\to l}\eta_i^{l-k}\mu_i^{k\to l}\right) + \min\left(h^{(l,k)}, h^{(k,l)}\right) \le T \tag{8}$$

The summation term is actually the weighted average sectional running time $\overline{SRT}^{\,l-k}$ which can be substituted. Simple rearrangement in terms of $C_{abs}^{(l,k)}$ and the replacement of the inequality with an equality sign gives the following equations (Note: *LB* and *UB* replace $C_{abs}^{(l,k)}$ to avoid confusion).

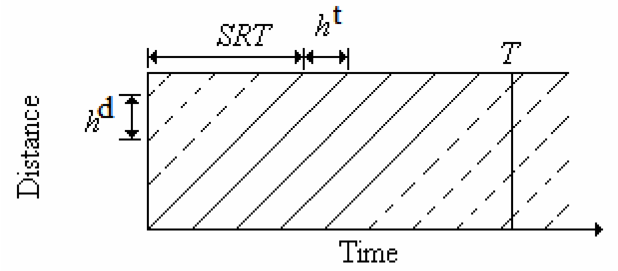
$$LB = \frac{T - z\left(h^{(l,k)} + h^{(k,l)}\right)}{\overline{SRT}^{\,l-k}}\ , \quad UB = \frac{T - \min\left(h^{(l,k)}, h^{(k,l)}\right)}{\overline{SRT}^{\,l-k}} \tag{9}$$

In practice, passing headways may also be imposed for safety reasons on sections bounded by crossing loops. Alternatively they may be added because crossing loops are not single points but small sections with their own length. Consequently a similar approach is taken for determining absolute capacity. This case is potentially simpler than the other case because the passing headway may be independent of train type and hence static. This is not the case however when the passing headway is proportional to a fixed distance, because trains have different velocities (i.e. they cover the same distance in different times).

If the section occupation condition was theoretically removed, and replaced with a minimum headway distance condition, the lower bound for absolute capacity would not be changed because the condition does not affect pairs of trains that travel in opposite directions.

The removal of this condition however significantly increases the throughput in one direction and hence the upper bound will be increased.

To calculate the upper bound for this scenario we need to be able to determine the capacity of a section with uni-directional traffic of one type only and minimum headways. A typical example is shown in Fig 2.



(a) **(b)**

Figure 2. Section saturation in one direction with minimum headways

Note that the trains given by the dashed lines are also included normally but are not in the upper bound calculations. This is because there are two distinct unidirectional fleets being considered. It should be noted that the maximum number of trains that can be on a section at any given time (length excluded) is $\left(L_{(l,k)}^{\text{section}} DIV\ h^{\text{d}}\right)+1$ given that h^{d} is the headway distance. Using this expression and Fig 2 the number of train paths that may be realised during a time period equal to the standard sectional running time (i.e. $SRT_i^{l\to k}$) is computed as: $1+2\sum_{j=1}^{\left(L_{(l,k)}^{\text{section}} DIV\ h^{\text{d}}\right)}\left(1-\left(\frac{j\times h^{\text{d}}}{L_{(l,k)}^{\text{section}}}\right)\right)$. This equation was derived by noting that there is at least one full path and a number of partial paths which are symmetrically positioned on either side. The summation term in particular computes the number of partial paths on one side only.

The following equations are proposed for determining the absolute capacity of a line that has unidirectional traffic of a single train type.

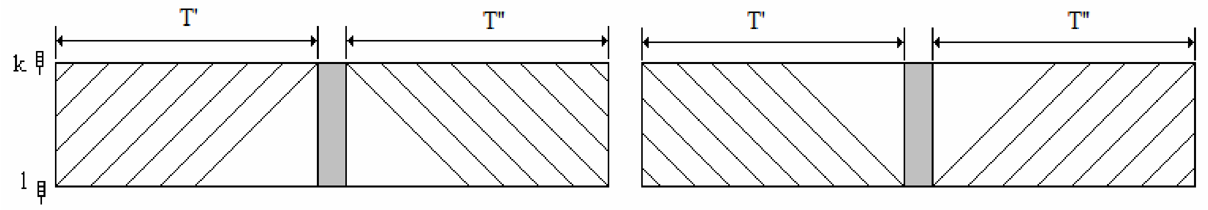
$$C_{abs,i}^{l\to k}=G\left(i,l,k,T,h,0\right)+G\left(i,l,k,SRT_i^{l\to k},h,1\right) \tag{10}$$

$$G(i,l,k,t,h,\alpha,\beta=\lfloor t/h\rfloor)=\sum_{i=\alpha}^{\beta}\left(\frac{\min\left(t-i\times h,SRT_i^{l\to k}\right)}{SRT_i^{l\to k}}\right) \tag{11}$$

Equation (10) consists of two parts and is based upon the observation (from Fig2(a)) that $\lfloor T/h^{\text{t}}\rfloor+\lfloor SRT_i^{l\to k}/h^{\text{t}}\rfloor+1$ trains (where $\lfloor\ \rfloor$ is the floor function) partially utilise the line (in the direction of *k* for example) during time period *T*. *The sum of the partial utilisations is the total number of train paths and hence the level of capacity*. The partial utilisation level for a particular train is determined by the bracketed term in function *G*. Function *G* requires parameters (*i*,*l*,*k*) and three additional values (t, h,α). The first two are the time period duration and the headway time. The third is a binary parameter that signifies whether the first train starting at time zero is included or not. Normally this is true, however when determining the upper bound this train would be included twice and hence a one is required in the second part of equation (10).

When there are multiple train types, one approach that may be taken is to assume that there is only one train type with SRT's given by the weighted average in one direction (as defined by equation (3)). We therefore replace $SRT_i^{l\to k}$ in equation (10) and (11) with □$SRT^{l\to k}$ and redefine $C_{abs,i}^{l\to k}$ as $C_{abs}^{l\to k}$. By doing this, the conflicts that occur between fast and slow trains can be ignored, although inaccuracies may consequently result.

Finally the capacity of two fleets should be calculated to determine the upper bound. The trains that are represented by dashed lines in Fig 2 should be removed. In particular the trains with dashed lines on the right should be removed for the first fleet and the trains with dashed lines on the left but in the opposite direction should be removed for the second fleet. This is shown clearly in the Fig 3(a) and (b).



(a) **(b)**

Figure 3. Upper bound calculation on signalled section

The following equations result. Note that index *i* is no longer required.

$$UB = G(l,k,T',h,0) + G(k,l,T'',h,0) \tag{12}$$

$$UB = G(k,l,T',h,0) + G(l,k,T'',h,0) \tag{13}$$

The individual uni-directional capacities may be determined by splitting time period T into two sub periods (i.e. $T = h^{(l,k)} + T' + T''$ or $T = h^{(k,l)} + T' + T''$ respectively). Because we can not explicitly define these, the following non-linear mathematical model (for the first case shown in Fig 3(a) for example) should be solved.

$$\textit{Maximise} \quad C_{abs}^{(l,k)} = G(l,k,T',h,0) + G(k,l,T'',h,0) \tag{14}$$

Subject to,

$$T = h^{(l,k)} + T' + T'' \tag{15}$$

$$G(l,k,T',h,0)\sum_{\forall i}\left(\eta_i \mu_i^{k\to l}\right) - G(k,l,T'',h,0)\sum_{\forall i}\left(\eta_i \mu_i^{l\to k}\right) = 0 \tag{16}$$

The decision variable is T' because T'' can be replaced after (15) is rearranged and substituted. Constraint (16) is a balance equation which enforces that the proportional and directional distributions are satisfied. This equation is derived by noting that traffic flow in each direction must satisfy $G(l,k,T',h,0) = C_{abs}^{(l,k)}\sum_{\forall i}\left(\eta_i \mu_i^{l\to k}\right)$ and $G(k,l,T'',h,0) = C_{abs}^{(l,k)}\sum_{\forall i}\left(\eta_i \mu_i^{k\to l}\right)$. Rearrangement of these expressions in terms of $C_{abs}^{(l,k)}$ and then setting them to be equal gives constraint (16).

In general when bi directional traffic is not split into two distinct uni-directional fleets (an example is shown in Fig 4) absolute capacity may be determined by the following equation.

$$C_{abs}^{(l,k)} = \sum_{\forall b}\left[\delta_b G(l,k,t_b,h,0) + (\delta_b - 1) G(k,l,t_b,h,0)\right] \tag{17}$$

The number of blocks (B) of alternating traffic (equivalently the number of fleets), the length of each block t_b, and the direction of travel for that block δ_b (where $\delta_k = 0$ or 1) might be chosen or solved for. However these variables should satisfy $\sum_b t_b + (B-1)h = T$ and $\delta_b \neq \delta_{b+1} \quad \forall b < B$.

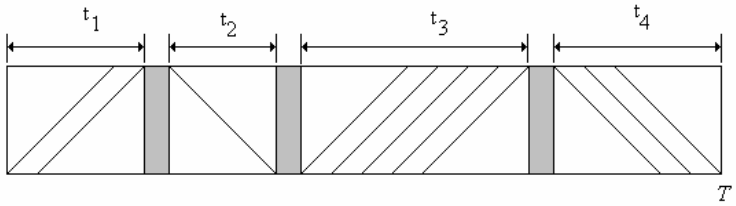


Figure 4. General case on signalled section

3.3. Dwell Times

In practice it is also common for trains to have pre planned dwell times at intermediate locations. Passenger set downs and pick ups, and loading and unloading of freight are normal examples. Dwell times significantly reduce the utilisation of standard line capacity because stationary trains do not utilise capacity very well. They utilise capacity most efficiently when they are moving, and in particular when they are travelling at their maximum velocity. No generalised measure to our knowledge has been proposed for incorporating this aspect. Although, incorporating dwell times at those locations without passing facilities (for example signals) is easily accomplished. The dwell time at the far boundary in the direction of travel is just added to the train's sectional running time. For example on section (l,k) the dwell time at *k* is added if the direction of travel is $l \rightarrow k$, and the dwell time at *l* is added if the direction of travel is $k \rightarrow l$. This is because trains stopped at these locations have not left the current section and hence other trains are not allowed to enter because of the section occupation condition.

Crossing loops however are different because a train that crosses a section boundary into a crossing loop is no longer on that section anymore, i.e. they are on another. Consequently the dwell time may not be added to the sectional running time as they previously were. However, as a crude first approach this is reasonable as an estimate. Dwell times also do not occur at the section boundaries but rather at the opposite ends of the crossing loops in the direction of travel. Because the crossing loop length is small, the absolute capacity of such a section will be very high even with dwell times included. Hence it is unlikely that such a section will become critical and thus dictate the overall "absolute" capacity for the corridor. Therefore, it is not necessarily appropriate to define the capacity of the entire corridor as the capacity of a single critical section.

We propose that a different approach be taken which is based upon the observation that dwell times affect the overall capacity of a railway line by reducing throughput of individual trains. The reduction in throughput is proportional to the total dwell time and the transit time. Therefore λ_i^{l-k} is defined as the reduction in throughput for a particular train of type *i* on corridor $l-k$ and λ^{l-k} is defined as the reduction in throughput for a mix of trains. Consequently absolute capacity calculation is proposed in the following way:

$$C_{abs}^{l-k} = \lambda^{l-k}\left(\frac{T}{\overline{SRT}^{CS_{l-k}}}\right) \quad (18)$$

One of the following approaches can be used for the calculation of the reduction factors.

Approach 1:

$$J_i^{l\to k} = \sum_{\forall (l',k')\in S_{l,k}} \left(SRT_i^{l'\to k'} \right),\ \bar{J}_i^{l\to k} = J_i^{l\to k} + \Delta_i^{l\to k},\ \lambda_i^{l\to k} = \left(\frac{J_i^{l\to k}}{\bar{J}_i^{l\to k}} \right),$$

$$\lambda_i^{l-k} = \mu_i^{l\to k}\lambda_i^{l\to k} + \mu_i^{k\to l}\lambda_i^{k\to l},\ \lambda^{l-k} = \sum_{\forall i}\left(\eta_i^{l-k}\lambda_i^{l-k} \right) \tag{19}$$

The reduction factor $\lambda_i^{l\to k}$ for a particular train type and direction is calculated as the ratio of the transit time $J_i^{l\to k}$ and the overall transit time $\bar{J}_i^{l\to k}$. The overall transit time includes dwell times whereas the standard measure does not. For each train type, the reduction factor λ_i^{l-k} is calculated by averaging the reduction in each direction by the directional distribution. Lastly the overall reduction factor λ^{l-k} is calculated as a weighted average value, i.e. the value is averaged with respect to the proportional distribution for the different train types.

Approach 2:

$$J_i^{l-k} = \mu_i^{l\to k} J_i^{l\to k} + \mu_i^{k\to l} J_i^{k\to l},\ \Delta_i^{l-k} = \mu_i^{l\to k}\Delta_i^{l\to k} + \mu_i^{k\to l}\Delta_i^{k\to l},$$

$$\tilde{J}^{l-k} = \sum_{\forall i}\left(\eta_i^{l-k} J_i^{l-k} \right),\ \tilde{\Delta}^{l-k} = \sum_{\forall i}\left(\eta_i^{l-k}\Delta_i^{l-k} \right),\ \lambda^{l-k} = \frac{\tilde{J}^{l-k}}{\tilde{J}^{l-k} + \tilde{\Delta}^{l-k}} \tag{20}$$

For each train type the weighted average journey time J_i^{l-k} and the weighted average total dwell time Δ_i^{l-k} are calculated. These values are obtained by averaging the values for a particular direction by the directional distriubtion. $\tilde{J}^{l-k}$ is then defined as the weighted average transit time and $\tilde{\Delta}^{l-k}$ is the weighted average total dwell time on corridor $l-k$ over all train types. The overall reduction factor is then calculated as the ratio between the weighted average transit time (no dwell time) and the weighted average total transit time (that includes dwell time).

Approach 3:

$$\lambda_i^{l'\to k'} = \frac{j_i^{l'\to k'}}{\hat{j}_i^{l'\to k'}}\ \forall l' \neq l, l' \neq k,\ \lambda_i^{l'\to k'} = \frac{SRT_i^{l'\to k'}}{SRT_i^{l'\to k'} + \Delta_{i,l'}^{l\to k}\left(\text{or } \Delta_{i,l'}^{k\to l}\right)} \qquad l' = l\left(\text{or } l' = k\right)$$

$$\lambda_i^{(l',k')} = \mu_i^{l\to k}\lambda_i^{l'\to k'} + \mu_i^{k\to l}\lambda_i^{k'\to l'},\ \lambda^{(l',k')} = \sum_{\forall i}\left(\eta_i \lambda_i^{(l',k')} \right),$$

$$C_{abs}^{(l',k')} = \lambda^{(l',k')}\left(\frac{T}{\tilde{SRT}^{(l',k')}} \right),\ C_{abs}^{l-k} = \min_{\forall (l',k')\in \square}\left(C_{abs}^{(l',k')} \right) \tag{21}$$

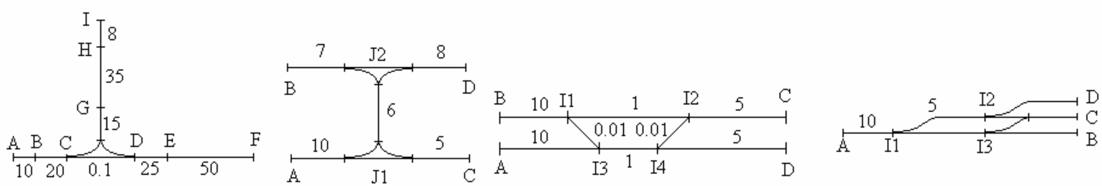
In this approach the reduction in throughput is assumed to be potentially different on each section. Consequently there is no single reduction factor for the entire corridor and the absolute capacity is determined as the capacity of the section with the smallest throughput.

Throughput on a particular section in particular is reduced by dwell times on prior sections in the direction of travel. The variables $j_i^{l'\to k'}$ and $\hat{j}_i^{l'\to k'}$ are introduced as the partial (cumulative) transit time and total transit time for a particular train type. More specifically they are the time to reach section boundary location l' (in direction of k') with dwell times not included and included respectively. The calculation of these quantities however is not shown above.

The reduction factors are calculated in a similar way to approaches 1 and 2, i.e. the values are averaged with respect to the directional and proportional distributions. The main difference in this approach is that the reduction factor (for each section) is proportional to the ratio of travelling time and total travelling time up to the current section and not including those sections afterwards in the direction of travel. It should be noted that at each corridor boundary, the ratio of the two partial transit times is zero because $j_i^{l' \to k'}$ is zero. This implies that there is no throughput in this direction which is clearly incorrect. Consequently an alternative measure is given for defining the reduction in throughput. This expression takes the ratio of the transit times to the opposite section boundary instead (i.e. k') without taking into account the dwell time at that location (as occurs normally).

4. TECHNIQUES FOR RAILWAY NETWORKS

In this section the approaches that were developed for railway lines in section 3 are modified and extended so that they are applicable to networks. Some examples /features of railway networks (displayed as line diagrams) that can be analysed by the approaches developed in this section are shown in Fig 5.



(a) **(b)** **(c)** **(d)**

Figure 5. Line diagram examples of railway networks

4.1. Train Mixes

For railway lines the absolute capacity was the number of trains that could traverse the critical section. This is equivalent to the number of trains that travel in each direction (i.e. from both IO points). The absolute capacity of a network is similar. It is the total number of trains that traverse all corridors. Equivalently it is the number of trains that travel between each pair of IO locations, layout permitting.

Quantifying absolute capacity is however more difficult in networks because of the interaction of corridors. This occurs when corridors have common sections. Consequently, an optimisation model is required instead of single standalone equations. To determine the absolute capacity when train mixes are incorporated the following core model is used.

$$\textit{Maximise } C_{abs}^{network} = \sum_{\forall l,k \in \Phi} \left(\sum_{\forall i} \left(x_i^{l \to k} \right) \right) \tag{22}$$

Subject to,

$$x_i^{l \to k} + x_i^{k \to l} = \eta_i^{l-k} \sum_{\forall i'} \left(x_{i'}^{l \to k} + x_{i'}^{k \to l} \right) \quad \forall i \in I, \forall l,k \in \Phi \mid l < k \tag{23}$$

$$x_i^{l \to k} = \mu_i^{l \to k} \left(x_i^{l \to k} + x_i^{k \to l} \right) \quad \forall i \in I, \forall l,k \in \Phi \mid l \neq k \tag{24}$$

$$x_i^{l \to l} = 0 \quad \forall i \in I, \forall l \in \Phi \tag{25}$$

$$\sum_{\forall i} \left(SRT_i^{l \to k} x_i^{l \to k} + SRT_i^{k \to l} x_i^{k \to l} \right) \leq T \quad \forall (l,k) \in \square \tag{26}$$

$$X^{l-k} = \sigma^{l-k} C_{abs}^{network} \quad \forall (l,k) \in \square \tag{27}$$

Constraint (23) and (24) enforce that the proportional and directional distributions are obeyed on valid corridors of the network. They may be alternatively written as $X_i^{l-k} = \eta_i^{l-k} \sum_{\forall i'} \left(X_{i'}^{l-k} \right)$ and $x_i^{l \to k} = \mu_i^{l \to k} X_i^{l-k}$. Without these constraints, the solution would be one sided as it would only contain the fastest trains travelling in the directions with the shortest sectional running times.

Constraint (26) is the standard balance equation which ensures that the flow through each section of the network must be less than or equal to the saturation limit (i.e. capacity of the section). To ensure that traffic on each corridor (i.e. corridor usage) is a given percentage of the total network traffic, constraint (27) may be optionally added. σ^{l-k} is the proportion of traffic on corridor $l-k$ where $\sigma^{l-k} \in (0,1)$. This value gives the precentage usage of the corridor with respect to the other corridors of the network. The decision variables of the model however are the number of trains that travel to and from each IO point. The other variables signifying the number travelling across individual sections are therefore redundant. This means that the $x_i^{l \to k}$ variables in constraint (26) must be converted. To accomplish this we note that only corridors that contain the section contribute to the occupancy level of the section. An additional summation sign is hence added to the original constraint to loop through only the associated corridors. A binary variable $\Theta_{l',k',l,k}$ is introduced and signifies whether a section (*l*,*k*) is traversed in corridor $l'-k'$. It also signifies that when travelling from l' to k' if section (*l*,*k*) is traversed in the direction of *l* or *k*. For example, $\Theta_{l',k',l,k} = 1$ implies that location *l* is reached before location *k* when travelling from l' to k'. It should be noted that $\Theta_{l',k',l,k} = 1$ implies that $\Theta_{k',l',k,l} = 1$ and $\Theta_{l',k',k,l} = 0$. The following constraint can then be written after conversion.

$$\sum_{\forall l',k' \in \Phi \mid \Theta_{l',k',l,k}=1} \left(\sum_{\forall i} \left(SRT_i^{l \to k} x_i^{l' \to k'} + SRT_i^{k \to l} x_i^{k' \to l'} \right) \right) \le T \quad \forall (l,k) \in \square : l < k \tag{28}$$

This approach (i.e. for determining network capacity) also works for serial lines which have a number of internal IO points. Previously this would not have been possible, especially using the theory of section 3. This is because only one corridor and one associated proportional and directional distribution existed. The mix of trains on each section was also consequently equal. This is not true when there are internal IO points however.

In a network it may also be reasonable to determine the proportional and directional distributions on sections as well as corridors, particularly sections which are members of two or more corridors. To do this the following equations are used.

$$\eta_i^{(l,k)} = \frac{x_i^{(l,k)}}{\sum_{\forall i'} \left(x_{i'}^{(l,k)} \right)} \forall i; \forall (l,k) \in \square \ , \quad \mu_i^{l \to k} = \frac{x_i^{l \to k}}{X_i^{l-k}} \quad \forall i; \forall (l,k) \in \square \tag{29}$$

These expressions however can only be evaluated after the absolute capacity of the network has been determined. The proportional and directional distributions on each section do not appear to be directly calculable from the corridor proportional and directional distributions.

4.2. Signals

The incorporation of signals and other reference locations is more complex in networks. The added complexity results as a consequence of determining the enforced headway time on a section. Enforced headways in railway lines were calculated as the sum of two weighted average travelling times. The travelling times in particular were also based upon the time it takes each train to reach the nearest crossing loop or passing facility and or return. Because the line is serial, there is only one nearest crossing loop on each side of the section if any at all. In networks however, there may be many crossing loop or passing facility because a section may be common to several different corridors. The particular one used depends on the corridor that the train is specifically traversing. A weighted average enforced headway is therefore proposed that takes into account all the enforced headways that may be incurred when travelling on different corridors. Therefore the mathematical model presented in section 4.1 is modified by removing equation (28) and adding the following equations to determine the lower bound.

$$\sum_{\forall i}\left(SRT_i^{l\to k}y_i^{l\to k}+SRT_i^{k\to l}y_i^{k\to l}\right)+h^{(l,k)}\min\left(Y^{l\to k},Y^{k\to l}\right)\le T\quad \forall(l,k)\in\square\ |\, l<k \tag{30}$$

$$h^{(l,k)}=\sum_{\forall l',k'\in\Phi\,|\,\Theta_{l',k',l,k}=1}\left(\sum_{\forall i}\left(h_{i,l,k}^{l'-k'}\left(\frac{x_i^{l'\to k'}}{Y^{l\to k}}\right)+h_{i,k,l}^{l'-k'}\left(\frac{x_i^{k'\to l'}}{Y^{k\to l}}\right)\right)\right)\quad \forall(l,k)\in\square\ |\, l<k \tag{31}$$

$$y_i^{l\to k}=\sum_{\forall l',k'\in\Phi\,|\,\Theta_{l',k',l,k}=1}\left(x_i^{l'\to k'}\right)\text{ and }\ Y^{l\to k}=\sum_{\forall i}\left(y_i^{l\to k}\right) \tag{32}$$

In these equations $y_i^{l\to k}$ and $Y^{l\to k}$ are used to define the total number of trains of type *i* and the total number of trains respectively that traverse section (l,k) in the direction of *k*. These variables denote travel between sections while the original variables $x_i^{l\to k}$ and $X^{l\to k}$ denoted travel between IO points only. In essence the redundant variables in the original model are re-inserted in equation (30). They are however calculated separately by the equations at (32). Equation (30) is otherwise taken directly from section 3. The equations at (32) are also formulated to make equation (31) more understandable. Equation (31) is a new equation which determines the headway on a section as a weighted average value that is proportional to the current actual mix of trains. The percentage values are determined by terms $\left(\frac{x_i^{l'\to k'}}{Y^{l\to k}}\right)$ and $\left(\frac{x_i^{k'\to l'}}{Y^{k\to l}}\right)$. In this equation $h_{i,l,k}^{l'-k'}$ is the enforced headway time for a train of type *i* on corridor $l'-k'$ on section (l,k) in the direction of *k*. It is calculated with respect to the distance to the nearest crossing loop from section (l,k) on corridor $l'-k'$ in the direction of *k*.

The resulting model is non-linear however the non-linearity in constraint (30), that is, the min function, may be removed by introducing additional binary variables and constraints for each section. The upper bound is evaluated in the same way as previously described in section 3.

4.3. Dwell Times

When incorporating dwell times in networks, absolute capacity is determined by the following equation which apart from the added reduction factor essentially the same as equation (22).

$$C_{abs}^{network} = \sum_{\forall(l,k)\in\square} \left(\lambda^{l-k} \sum_{\forall i} \left(x_i^{l\to k} + x_i^{k\to l} \right) \right) \tag{33}$$

The reduction factors are calculated for each corridor according to the first two approaches in section 3.3. The original model proposed in section 4.1 can be solved firstly and then equation (33) can be evaluated. Alternatively this equation can be used directly as the model objective function. Note that the equations presented in section 3.3 should be added as additional constraints if the second choice is taken. Otherwise the approaches developed in section 3.3 remain the same for networks.

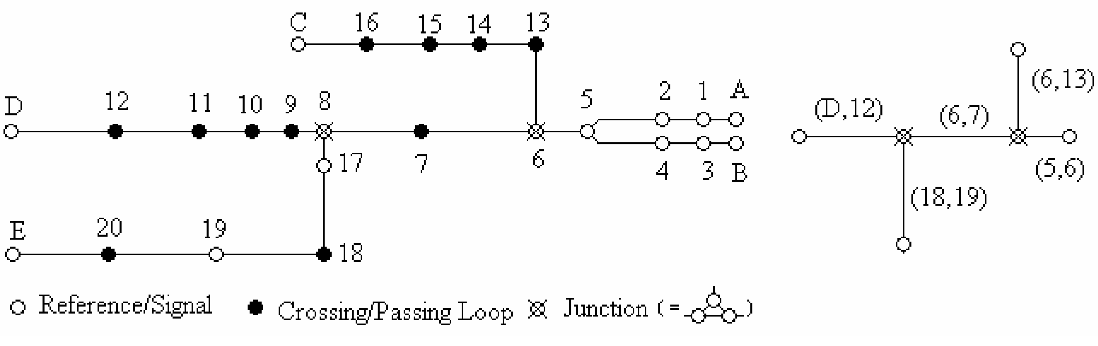
The third approach for incorporating dwell times from section 3.3 is more difficult for networks for the same reason as incorprating signals in section 4.2 was more difficult for networks. The reduction factor on a section must now take into account all the different corridors that contain the section. The following modifications are needed for networks. Firstly the right hand side of equation (29) must be multiplied by $\lambda^{(l,k)}$. Secondly we need to recalculate $\lambda^{(l,k)}$ as follows.

$$\lambda^{(l,k)} = \sum_{\forall l',k'\in\Phi \mid \Theta_{l',k',l,k}=1} \left(\sum_{\forall i} \left(\lambda_{i,l,k}^{l'-k'} \left(\frac{x_i^{l'\to k'}}{Y^{l-k}} \right) + \lambda_{i,k,l}^{l'-k'} \left(\frac{x_i^{k'\to l'}}{Y^{l-k}} \right) \right) \right) \quad \forall (l,k) \in \square : l < k \tag{34}$$

In this equation $\lambda_{i,l,k}^{l'-k'}$ is the reduction factor for train type i on section (l,k) of corridor $l'-k'$ in the direction of k. The equation has the same structure as (31).

5. CASE STUDY

To illustrate the proposed techniques the following network shown in Fig 6(a) is considered. It encompasses the most important, difficult and significant elements of theory developed in this paper. The network has nine corridors (i.e. $\square = \{(A,C),(A,D),(A,E),(B,C),(B,D),(B,E),(C,D),(C,E),(D,E)\}$) out of a possible ten. There are twenty four sections, with a total of 171.59 kilometres of track. The set of IO locations is $\Phi = \{A,B,C,D,E\}$. Three train types use this network with velocities respectively (in km/h) $V = (80,100,120)$ and equal in both directions.



(a) **(b)**

Figure 6. Network layout and equivalent reduced network (shown to scale)

It should be noted that in Fig 6(b) an equivalent but reduced network is shown. During our numerical investigations we found that many parts of the network are redundant and may be removed to simplify the analysis. The simplification involves reducing each serial line to one section which is a local bottleneck. Parallel sections can also be represented as single track with twice the capacity. The rest of the data is as follows.

Table 1. Network data

#	**Corridor**	**# of Sections**	**Length (km)**	**CS**	**#**	**Section**	**Length**	**#**	**Section**	**Length**
1	A-C	9	55.29	6-13	1	A-1	3.36	13	3-4	4.23
2	A-D	11	76.23	6-7	2	1-2	4.23	14	4-5	7.89
3	A-E	11	89.01	6-7	3	2-5	7.89	15	6-13	9.05
4	B-C	9	55.29	6-13	4	5-6	5.56	16	13-14	5.97
5	B-D	11	76.23	6-7	5	6-7	11.93	17	14-15	5.31
6	B-E	11	89.01	6-7	6	7-8	10.40	18	15-16	6.67
7	C-D	12	89.45	6-7	7	8-9	3.45	19	16-C	7.25
8	C-E	12	102.23	6-7	8	9-10	4.12	20	8-17	3.64
9	D-E	10	78.50	18-19	9	10-11	5.58	21	17-18	9.23
					10	11-12	8.80	22	18-19	11.50
					11	12-D	10.92	23	19-20	11.25
					12	B-3	3.36	24	20-E	10.01

Table 2. Proportional and directional distributions

η_i^{l-k}	Train Type			$\mu_i^{l\to k}, \mu_i^{k\to l}$	Train Type					
Corridor	1	2	3	Corridor	1		2		3	
A-C	0.00	0.44	0.56	A-C	-	-	0.47	0.53	0.63	0.37
A-D	0.31	0.45	0.24	A-D	0.26	0.74	0.33	0.67	0.50	0.50
A-E	0.86	0.14	0.00	A-E	0.92	0.08	0.23	0.77	-	-
B-C	0.33	0.32	0.35	B-C	0.04	0.96	0.35	0.65	0.32	0.68
B-D	0.74	0.03	0.23	B-D	0.54	0.46	0.51	0.49	0.13	0.87
B-E	0.29	0.00	0.71	B-E	0.16	0.84	-	-	0.24	0.76
C-D	0.87	0.10	0.03	C-D	0.19	0.81	0.33	0.67	0.82	0.18
C-E	0.39	0.26	0.35	C-E	0.55	0.45	0.76	0.24	0.61	0.39
D-E	0.49	0.36	0.15	D-E	0.05	0.95	0.99	0.01	0.49	0.51

Table 3. Sectional running times (equal in both directions)

SRT	Train Type				Train Type		
Section	1	2	3	Section	1	2	3
A-1	2.52	2.01	1.68	3-4	3.17	2.54	2.12
1-2	3.17	2.54	2.12	4-5	5.91	4.73	3.94
2-5	5.91	4.73	3.94	6-13	6.79	5.43	4.53
5-6	4.17	3.34	2.78	13-14	4.48	3.58	2.98
6-7	8.95	7.16	5.97	14-15	3.98	3.18	2.65
7-8	7.80	6.24	5.20	15-16	5.01	4.00	3.34
8-9	2.59	2.07	1.73	16-C	5.44	4.35	3.62
9-10	3.09	2.47	2.06	8-17	2.73	2.18	1.82
10-11	4.18	3.35	2.79	17-18	6.92	5.54	4.62
11-12	6.60	5.28	4.40	18-19	8.63	6.90	5.75
12-D	8.19	6.55	5.46	19-20	8.44	6.75	5.62
B-3	2.52	2.01	1.68	20-E	7.51	6.01	5.01

Table 4. Dwell times (in both directions)

Location	Train Type			Location	Type		
	1	2	3		1	2	3
1	0.00	0.00	0.00	11	1.71	4.82	2.80
2	0.00	0.00	0.00	12	2.05	0.35	0.85
3	0.00	0.00	0.00	13	1.39	2.97	2.11
4	0.00	0.00	0.00	14	2.82	0.65	0.08
5	0.00	0.00	0.00	15	4.87	2.12	4.67
6	0.00	0.00	0.00	16	4.39	1.63	1.45
7	2.33	1.93	4.89	17	0.00	0.00	0.00
8	0.00	0.00	0.00	18	4.16	0.10	2.28
9	2.33	0.65	1.77	19	0.00	0.00	0.00
10	3.59	1.70	1.40	20	0.14	3.01	1.17

The first part of the analysis determines the capacity of individual corridors as if they were independent. The results are shown in Table 5. The first column gives the results of the ideal situation which occurs when all sections are assumed to be bounded by passing facilities. The second column shows the results for the actual scenario (lower bound shown only). The remaining columns are the same except for the addition of dwell times. Approach 1 and 3 are shown also when dwell times are included. No significant difference between the first two dwell time approaches was found on this example and only one approach is therefore shown. This is not necessarily true of all problems however. The likely cause for this similarity is the assumption of equal travelling time in each direction in this case study.

Table 5. Capacity levels on independent corridors (w.r.t. the given mix)

#	Corridor	No Dwell Times			Dwell (Approach 1)			Dwell (Approach 3)		
		All XL C_{abs}	Actual (LB) C_{abs}	Reduct -ion (%)	All XL C_{abs}	Actual (LB) C_{abs}	Reduct -ion (%)	All XL C_{abs}	Actual (LB) C_{abs}	Reduct-ion (%)
1	A-C	292.49	150.07	0.49	231.24	118.73	0.49	248.66	132.8	0.47
2	A-D	193.83	119.14	0.39	157.59	96.88	0.39	159.9	108.94	0.32
3	A-E	165.58	129.23	0.22	147.94	115.57	0.22	151.847	126.2	0.17
4	B-C	258.94	181.61	0.3	201.97	141.75	0.3	184.64	142.35	0.23
5	B-D	175.44	100.38	0.43	142.52	81.6	0.43	148.08	83.8	0.43
6	B-E	210.71	125.39	0.4	174.71	103.96	0.4	185.92	109.44	0.41
7	C-D	165.92	141.70	0.15	120.68	103.08	0.15	120.59	101.71	0.16
8	C-E	193.59	147.39	0.24	148.54	113.12	0.24	145.41	111.48	0.23
9	D-E	190.16	129.54	0.32	148.78	101.2	0.32	181.51	80.2	0.56
	Totals	1846.65	1224.18	0.34	1473.97	975.89	0.34	1526.56	996.92	0.35

From this table we can see that the dwell times cause roughly a 20% reduction in the utilisation of absolute capacity. The third approach for incorporating dwell times also gives a higher level of capacity than approach 1.

The second part of the analysis involves the determination of the capacity for the entire network. The results in Table 6 are obtained after the mathematical model was solved using GAMS. Because the model is non-linear the solutions are guaranteed to be only locally optimal to a given tolerance.

Table 6. Network capacity for various options

Train Type	No Dwell		Dwell (Approach 1)		Dwell (Approach 3)	
	All XL	**Actual**	**All XL**	**Actual**	**All XL**	**Actual**
1 (only)	450.13	389.53	359.14	303.27	349.28	278.52
2 (only)	562.66	407.34	460.29	316.16	410.1	234.08
3 (only)	675.19	471.45	518.84	339.8	495.2	296.63
Given mix	574.28	366.97	457.92	283.5	475.66	258.74

From Table 5 and 6 the utilisation level of each corridor is determined as a percentage of the total network capacity. Secondly the utilisation level of the corridor as a percentage of the unencumbered (i.e. independently used) corridor is determined. The effects of corridor interaction are calculated by comparing the absolute capacity of the network with the capacity of the network of independent corridors. For example for the given mix, the two values are 366.97 and 1224.178 respectively which implies a 70% reduction in the utilisation of network capacity. As expected, the more interacting corridors the less capacity that can be utilised.

Lastly it should be noted that the removal of the directional distributions and the subsequent replication of the analysis will also give upper and lower bounds that are valid for all mixes.

6. CONCLUSIONS

General approaches for determining absolute capacity in railways were developed in this paper and are based upon the logic of an existing bottleneck approach. The approaches are suitable for railway lines and networks with uni and or bi-directional traffic and do not require any major modifications when dealing with one scenario or another. Simplifications may however be made to reduce the computational burden in some circumstances. Our approaches are based upon an existing bottleneck approach however they are unique because many realistic aspects of railway operation that contribute to the determination of capacity were incorporated that to our knowledge are not normally included in analyses of capacity. These included train mixes, train lengths, dwell times, signals, section occupation conditions, and networks. Another significant factor in the approaches was the inclusion of headways of one particular type or another, for example, headway times and distances, enforced or passing headways, and safety headways. Because trains of different type use the railway, the most efficient way of performing the analysis was also to define a train with average performance or characteristics. Hence weighted average calculations also played a large part. For example these included the calculation of weighted average sectional running time, dwell time, transit time, and headways.

The main approach developed in this paper is a generic optimisation approach. Consequently absolute capacity is determined by solving an optimisation model. The objective in particular is the total throughput between all pairs of IO locations. Alternative stand alone equations were also developed that are only suitable for simpler serial scenarios.

The approaches developed can also be used to measure (quantify) the potential interaction affects that occur in railway networks between interrelated corridors. The case study demonstrated the steps required to accomplish this. From our case study in particular we have illustrated that corridors can not be efficiently utilised if they are part of a complex network consisting of interacting corridors. This conclusion can be reached using common sense and validates our approach.

Absolute capacity is a term that is defined for the first time to our knowledge in this paper. It is a value of capacity that can be used for several planning purposes. For particular input train mix proportions the approaches developed in this paper give a value for absolute capacity and an associated "actual" train mix. In practice however it is really the associated mix of trains that is most important when developing or modifying an actual timetable. Determination of absolute capacity values however depend on many conditions. This may be seen as a drawback, however to our knowledge it is unavoidable. There does not appear to be a definition of capacity that is unique for all situations, i.e. that is unconditional.

A value of absolute capacity may also be used in conjunction with an approximation of actual capacity to calculate the potential congestion that will occur on the railway. The approaches in this paper may also be used to calculate actual capacity as well. Actual capacity approaches have generally been statistical in nature. In particular they involve the calculation/approximation of interference delays but do not compute actual capacity directly. The interference delays must be added to the sectional running times and used in a bottleneck analysis (like our one) before actual capacity can be obtained.

Absolute capacity however is an ideal and not necessarily a reality. For this reason no simulation results and or other empirical proofs or comparisons were given. Simulation in particular is insufficient for any purpose associated with absolute capacity because only schedules that show actual operations can be obtained. Empirical proof of actual capacity through exact timetabling however is a possibility that is currently being further investigated.

ACKNOWLEDGEMENTS

The authors wish to thank two anonymous referees for their useful comments. The research underlying this paper was supported by Australian Research Council Linkage Grant LP0219595 with Queensland Rail.

REFERENCES

Assad A. A. (1980). Models for rail transportation. *Transportation Research A*, 14A, 205-220

Carey M. and Kwiecinski A. (1994). Stochastic approximation to the effects of headways on knock-on-delays of trains. *Transportation Research B,* 28B (4), 251-267.

Greenberg B. S. (1988). Predicting dispatching delays on a low speed single track railroad. *Transportation Science*, 22(1), 31-38.

Hall R. W. (1987). Passenger delay in a rapid transit station. *Transportation Science*, 21(4), 279-292.

Higgins A., Kozan E., and Ferreira L. (1995). Modelling delay risks associated with train schedules. *Transportation Planning and Technology*, 19, 89-108.

Higgins A. and Kozan E. (1998). Modelling train delays in urban networks. *Transportation Science*, 32 (4), 346-357.

De Kort A. F. , Heidergott B., Ayhan H. (2003). A probabilistic (max,+) approach for determining railway infrastructure capacity. *European Journal of Operational Research*, 148, 644-661.

Ferreira L. (1997). Rail track infrastructure ownership: investment and operational issues. *Transportation*, 24, 183-200.

Gibson S. (2003). Allocation of capacity in the rail industry. *Utilities Policy*, 11, 39-42.

Kozan E. and Burdett R. L. (2004). Capacity determination issues in railways. RTSA Conference on Railway Engineering (CORE), Darwin, Australia.

Kozan E. and Burdett R. L. (2005). A Railway Capacity Determination Model and Rail Access Charging Methodologies. *Transportation Planning and Technology*, 28(1), 27-45

Kraft E. R. (1987). A branch and bound procedure for optimal train dispatching. *Journal of Transportation Research Forum*, 28, 263-276.

Petersen E. R. (1974). Over the road transit time for a single track railway. *Transportation Science*, 8, 65-74.

Petersen E. R. (1975). Interference delays on a partially double tracked railway with intermediate signalling. *Journal of the Transportation Forum*, 16, 54-62.