25th Australasian Transport Research Forum
Incorporating the BTRE Transport Policy Colloquium
Canberra 2-4 October 2002

Planning and modelling of high capacity transit systems

Eric C Keys, Charles Waingold and Tarek Salameh
Booz Allen Hamilton

Abstract

Transit systems throughout the world are under increasing pressure to challenge system capacity to increase passenger throughput while improving system safety, speed, frequency and reliability.

This paper will review the application of operations modelling as part of the system design process for new and existing high capacity urban transit systems. System design involves analysis of passenger demand, development of service plans and the specification of infrastructure requirements. These elements together define a system operations plan. To ensure the reliability and feasibility of these concept plans they are tested and evaluated using a system simulation model that incorporates disruptions to the schedule services to model actual operating conditions. The outcome of these evaluations highlights areas for improvement and refinement within an iterative process of system optimisation.

Contact author

Eric Keys
Booz Allen Hamilton
Associate,
Level 53, 101 Collins Street
Melbourne VIC Australia
Telephone: +61 3 9221 1900
Fax: +61 3 9221 1980
Email: Keys_Eric@Bah.Com

GO BACK TO START SCREEN **GO BACK TO INDEX**

Introduction

Transit systems throughout the world are under increasing pressure to challenge system capacity and to increase passenger throughput while improving system safety, speed, frequency and reliability.

This paper reviews the application of operations modelling, as part of the system design process for new and existing high capacity urban transit systems. System design involves analysis of passenger demand, development of service plans and the specification of infrastructure requirements. These elements together define a system operations plan. To ensure the reliability and feasibility of a concept plan it needs to be tested and evaluated using a system simulation that incorporates random disruptions to scheduled services, to model actual operating conditions. The outcome of these evaluations highlights areas for improvement and refinement within an iterative process of system optimisation.

Pressures to Develop High Capacity Transit Systems

Many cities are endeavouring to improve the efficiency and effectiveness of their transport systems to optimise their competitive position. This is being done to attract business and investment, while also improving the urban environment. Excessive use of private vehicles has been identified, in a number of metropolitan urban planning and environmental studies, as an inhibitor to the achievement of these objectives. To reduce private vehicle dependency, many urban strategies are planning a sustained increase in public transport patronage, in the future.

A number of cities are planning to increase the role for public transport to meet urban travel needs. Singapore's recently released Concept 21 Masterplan calls for further intensification of residential, industrial and business areas around Mass Rapid Transit (MRT) Stations (Urban Redevelopment Authority, 2002). The Land Transport Authority's (Singapore) objective is to increase public transport's share of daily passenger trips from current levels of around 63% to 75% by 2030 (Land Transport Authority, 2001).

A key target to be contained in the Growing Victoria Together Statement is to increase Melbourne's metropolitan public transport mode share from the current level of 9% to 20%, by the year 2020. This increase is planned in order to meet the Statement's economic, social and environmental objectives. A package of measures to achieve this target includes the development of high - density transit cities around major existing public transport nodes (Department of Infrastructure, 2002).

Perth's Metropolitan Transport Strategy seeks to arrest the decline in public transport mode share for all daily trips, by almost doubling the 1991 mode share of 6.4% to a targeted 12.5% by 2029 (Department of Transport, 1995).

A number of cities are also proposing new high capacity transit lines to serve new areas or to relieve capacity, despite the high cost of creating new corridors within the existing urban fabric. Singapore's MRT network is proposed to expand from 93km to 500km (Urban Redevelopment Authority, 2002), whilst new rail lines have been announced in Perth and Sydney. Melbourne is about to embark on a planning study of its rail network.

Creating transit systems attractive to patrons

Operational attributes of a high capacity transit system that make it attractive to patrons are:

- faster travel times, with high yet comfortable acceleration and braking;
- short dwell times;
- a high frequency service (2 – 6 minutes); and
- high levels of reliability and punctuality.

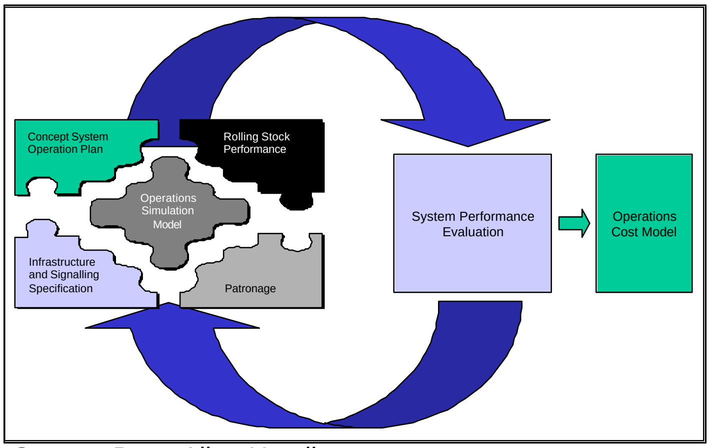
Passengers value high frequencies that enable users to virtually "walk" up to a service and not have to plan the timing of their trip. Similarly, passengers are attracted to services that are on time and reliable. Of the service attributes measured in Melbourne's Customer Satisfaction Survey, on time running has the highest weighting to the overall Customer Satisfaction Index. This index includes comfort on the train, ticketing, safety, and other attributes (Department of Infrastructure, 2000). This is reinforced in stated preference surveys, which have estimated that the time lost due to late running transport is up to six times the value of in-vehicle time. This compares with valuations of around two times the value for expected waiting and walk time (Booz Allen Hamilton, unpublished).

High capacity transit systems are becoming more prevalent due to the increasing recognition of the need to reduce car dependency. In order to attract and sustain patronage growth, these systems need to provide a high quality of service. In evaluating options to enhance existing and future high capacity transit corridors, the methodology must verify that an operating plan is workable and robust.

Operations Planning

The operations planning process involves planning, analysing and optimising rail infrastructure, rolling stock and operations, to meet demand. This process is iterative, with each cycle involving refinement towards identifying the optimised outcome that makes up the final operating plan. The impact on patronage may also require reassessing different levels of service and service patterns, compared to that of the initial patronage modelling stage. Finally, the outputs of the operations planning process are provided for eventual incorporation into the economic and financial evaluation.

Figure 1:Elements of the Operations Planning Process



Source: Booz Allen Hamilton

Elements of the Operations Planning Process

In this section, the elements of the operations planning process (as outlined in Figure 1), are illustrated through a case study of operations planning for a hypothetical new high frequency metro service. The case study focuses on how accommodating forecast variation could be handled.

The case study scenario for the hypothetical new high frequency metro service is outlined below:

A *new 35 km metro service is proposed to penetrate a high density urban area and provide through running services to the central business district, serving a number of regional centers en route.*

Element 1 - Patronage

An understanding of the nature and pattern of patronage demand is critical in determining the needs of the network. Peak passenger throughput is a key determinant in governing peak service and rolling stock requirements. The peak loading point may only occur at a specific location for a short period of time. Therefore, an understanding of the peak loading patterns is important in optimising system design.

An important consideration is an understanding of the operational implications of the patronage forecasts. The ramp up period, forecast variation and sensitivity, and how demand growth beyond the forecasting horizon is accommodated are aspects that may need to be incorporated into the operations planning process.

For example, the acceptable variation for validating forecasting models for the majority of public transport schemes in the United Kingdom is within 25%. Superimposing errors attributable to growth factors and sensitivity tests increases the bounds of potential patronage outcomes (Department of Transport, 2002). Therefore, providing flexibility in the design of operations and infrastructure needs to be considered to account for the range of possible market outcomes.

In the case study scenario, patronage forecasts indicate that during the peak two hour period an estimated 28,000 passengers will use the metro at the critical line section, and it is assumed that the passenger arrival pattern will be evenly distributed throughout the period. Due to the proposed metro serving a highly dense urban environment, there are a number of points along the route where patronage levels close to the critical line section is forecast to occur. The model used to make these forecasts has been validated to within 25%, implying that at the very extremes, the patronage could be as high as 35,000 or as low as 19,000 passengers in the first year of opening.

It has also been estimated that there will be a consistent annual patronage growth factor of 0.5% per annum, over the 30 year project evaluation period. This is equivalent to the final year having a 13% higher patronage than the first year of opening given a 5 year construction period. This will result in a final year patronage of 32,000. Sensitivity tests conducted on the growth rates of 0.75% per annum to 0.25% per annum suggest that at the extremes, final year patronage could range from 20,000 to 42,000 passengers per year.

Element 2 - Concept System Operations Plan

An initial concept system operations plan should broadly detail the specification of a new service or modifications to be made to an existing service, from an understanding of the patronage demand. Aspects of the plan should include service characteristics such as stopping patterns, opportunities for potential synergies if there are multiple lines using the same corridor, system flexibility and depot and stabling arrangements. System wide performance evaluation criteria should also be incorporated.

A number of service characteristics should be clearly defined in the plan. These include stopping patterns, frequency by time period, service span and rolling stock passenger capacity. A wide range of stopping patterns might be considered including short running, express, all - stopper and skip - stop services. Each stopping pattern should have unique train path requirements due to differences in speed and stop spacing. Ensuring the system has the capacity to accommodate the specified range and intensity of train paths without conflicts, is a key output of the exercise. The passenger capacity of the proposed train consists should also influence service frequency. In this case study, as there are a number of critical line sections along the route where close to the maximum patronage is achieved, it is proposed that a consistent level of service is provided throughout the entire length in both directions with stops at all stations. A two minute headway would be desirable to enable passengers to be able to “walk up” to a service to minimise wait time.

Should the corridor serve multiple lines, potential operating and infrastructure synergies should also be identified for further investigation. This could include identifying existing and proposed services on different lines that have similar stopping patterns and could share the same track; whether services are timed to facilitate passenger interchange between trains or common station platforms. Identifying these potential synergies would assist to maximise the value of the corridor and minimise capital costs by reducing land take and infrastructure requirements.

The approach to accommodate forecast variation and longer term growth should be also considered as part of concept development. This would assist to determine the amount and type of in-built flexibility required of the system at critical points to ensure that the system, and particularly the infrastructure, is able to accommodate the range of likely potential loadings. Types of measures that could be undertaken include incorporating flexibility into station and rolling stock design, to accommodate additional carriages in a train consist, and reserving land adjacent to sections of existing and proposed corridors.

For the case study, two approaches could be taken to determine the most desirable strategy to build flexibility into the system. In the first instance, consideration would be given to improving service frequency, and secondly to increasing passenger capacity on each rolling stock set.

A key aspect is the consideration of depots and stabling locations. Approaches used to build up services range from:

- launching services from one to two major locations; or
- launching services from multiple stabling locations sited along the length of the corridor.

Finally, system performance objectives and evaluation criteria should be defined. These criteria could comprise the percentage of trains arriving within a predetermined number of minutes within schedule, maximum acceptable delays and minimising dead running time from depots. Inputs into the cost model and evaluation should also be specified.

In the case study scenario, the key performance indicator that needs to be met by the service would be that 98% of services must arrive within two minutes of schedule. The measuring points for these services are at a midpoint city station (Station Q), and at the termini Station AH (for up services) and Station A (for down services). The concept system operations plan should be continuously refined throughout the life of any transit project. The case study infrastructure consists of two single line sections operating in an up and down direction.

Element 3 - Infrastructure and Signaling

In the first instance the process to develop infrastructure and signaling options involves undertaking a preliminary capacity analysis to determine the track and signaling layout options.

The infrastructure catering for the ultimate capacity requirements should form the master plan to guide the corridor's infrastructure and signaling development over the project life. This involves designating train paths and services to various tracks and determining the infrastructure and signaling requirements. For example, up and down all stopper services could be designated to outer tracks and the express services to middle tracks. Opportunities for synergies across services that were identified in the concept system operations plan should be considered at this stage.

Detailed track layouts would need to be specified in terms of junction configuration, grades, curves and station locations. The type and layout of the signaling system would need to be identified, including whether it is wayside or automatic train control and moving block. Signal locations if applicable and safety margins would also need to be identified. Infrastructure and signaling should be specified to achieve system performance criteria.

As this case study involves services operating with the same service pattern but at very high frequency, a conventional two uni-directional track system with a moving block system is to be adopted.

Element 4 - Rolling Stock

Decisions made on a network wide basis usually drive rolling stock selection. However, their capacity and train consists can vary depending on patronage and service level trade offs. For example lower capacity, but more frequent services, provide a more attractive service to patrons but could potentially place more stress on the system and prove costly (more drivers etc). These trade offs need to be determined when developing the system's operating plan.

In Singapore, it is planned that the City Circle Line will have train consists of three carriages that have a capacity of about 670 passengers operating at very high frequency (30 trains per hour). Seating will be limited to bench seating alongside the walls, enabling a maximum amount of area for standing. In contrast, Australian transit systems tend to have longer train consists with higher overall passenger capacity but lower passenger capacity per carriage due to more generous seating.

Critical rolling stock parameters that are required to determine a service's train path time requirements include acceleration and braking rate profiles, and physical dimensions. A critical determinant of dwell time is the number of doors per carriage. The number of doors per car side is higher for Singapore's City Circle Line compared to Melbourne, to facilitate higher number of boardings and alightings at each station without prolonged dwell time.

Table 1: Comparison of Rolling Stock Proposed for Singapore and Melbourne (Peak Periods)

Parameter	Singapore City Circle Line	Melbourne Connex Network
Model	Alstom Metropolis	Alstom X'Trapolis
Passengers seated	150	548
Passengers standing	520	254
Total passengers	670	802
No. cars per set	3	6
Length over train set	69.25m	145.2m
No. doors per side/ car	4	3
Door opening	1500mm	1300mm
Max operating speed	100km/h	130km/h
Gauge	1435mm	1600mm
Max Acceleration	$1.3ms^{-2}$	$1.2ms^{-2}$
Max Deceleration	$1.3ms^{-2}$	$1.2ms^{-2}$

Source: Hammond (2002); Binnie Black Veatch et al (unpublished)

In the case study scenario each carriage has been set to carry 160 seating and standing passengers. It is standard that each train set comprises of three car sets. In relation to this fact, three car sets operating with a capacity of 480 passengers per set at 120 second headways should meet the critical demand given the forecast of 28,000 passengers over the two hour peak period in the final year of the project evaluation. However, the approach to cater for the potential range of demand needs to be considered. The rolling stock sets should be specified to be readily extended to provide four car sets, that provide a capacity of 640 passengers per set.

Element 5 - Operations Simulation

The first step in the operations simulation is to use the patronage, concept plan, rolling stock and track layouts to develop an initial timetable. Secondly, timetable and system robustness tests that are to be applied will need to be considered and designed. Finally, this timetable and the physical system attributes are then modeled to test for timetable proving and robustness. The system is then refined if required.

The scenarios that are to be tested to assist in determining the best approach to handle forecast variation, are shown in Table 2. The upper forecast range is 42,000 passengers over the peak period in one direction.

Table 2: Scenarios to be Simulated

Scenario	No. Cars	Headway (secs)	Two Hour Passenger Capacity (One Way)
3.1	3	120	28,800
3.2	3	90	38,400
3.3	3	60	57,600
4.1	4	120	38,400
4.2	4	90	51,200
4.3	4	60	76,800

Timetable Development

The first step in developing the timetable involves establishing the train path requirement for each stopping pattern. The train path is determined from the time needed for travel, station dwell, and recovery and layover times across the corridor.

Travel times are determined by running test trains in the simulation model to ascertain minimum travel times between stations and stopping points such as layovers.

Dwell times can be determined from empirical data or theoretically determined. Variations in passenger boardings, alightings and level of crowding at each station result in actual dwell times fluctuating by service throughout the day. In considering infrastructure and rolling stock design, factors affecting dwell time include: door opening and closing times; the number, width and spacing of doors; platform circulation; whether the platform configuration allows for single or dual platform; and unloading and disability access arrangements (Transportation Research Board, 1996).

In the case study scenario, the stations serving key city locations and regional centers have a distribution of dwell times with a base of 30 seconds, all other stations utilised the empirical distribution with a base dwell time of 20 seconds.

Recovery or coasting times are added to minimum ‘station-to-station’ times to account for driver variability and delays that may occur at stations. Recovery time can be distributed in a number of ways, including distributing it evenly across the corridor or biasing the distribution towards layovers and congested areas. One of the key outputs of the simulation testing is to determine whether the recovery time is sufficient, or whether additional recovery time needs to be built into the timetable. Another option is to enhance the infrastructure. The amount of recovery time allocated should not be excessive during passenger operations, such that the service quality is significantly compromised due to unnecessary idling. For loop services, it is advisable to distribute recovery time evenly across the route, to ensure that there is no prolonged dwell at a particular station. A coasting time of 5% is included in the system, and this is evenly distributed throughout the line section.

Should there be turn backs or any other ‘non – running’ activity, layover time needs to be incorporated into the timetable development phase.

The total journey time would then need to be validated. An important consideration during validation is to ensure that train priority and control rules are properly reflected in the model, and if they are not related, the underlying assumptions that would affect the outcomes.

The total journey time in the case study scenario was 55 minutes across a 35km route length with 36 stations.

Once the train path requirements for each service are established, the timetable can be fully developed, in accordance with the requirements of the Concept System Operations Plan. If there is a wide mix of service patterns, scheduling train paths can potentially become a complex task. Timetable proving is therefore necessary to ensure there are no conflicts, particularly at junctions, and that signal headways are adhered to. Different signaling systems, such as wayside and automatic train control systems could be tested with different signal spacing.

Timetable and System Robustness

A high capacity transit system’s performance is extremely vulnerable to delays and incidents that could have wider "knock-on" effects, impacting on a number of following services due to the narrow headways. As well as proving, the timetable should be verified for robustness.

Dwell time variability could occur as a result of surges in boarding and alighting due to abnormal events, such as special events. The other disruption that can be applied is the delays to the start of a particular service. The variable time can be pre-determined to create a 'what–if' scenario or drawn from on-time running probability density functions. These functions can be applied to individual or groups of services with similar stopping patterns across a nominated time period, and can be derived from empirical evidence. For example, probability distributions of arrival delays for express trains would have very different profiles compared to all stopper trains.

In the simulation, dwell time extension probability distributions shown in Figures 3 and 4 were added to the base dwell times to reflect variability.

Figure 3: 20 Second Base Dwell Time Extension Probability Distribution

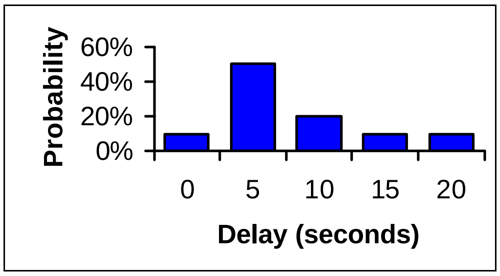
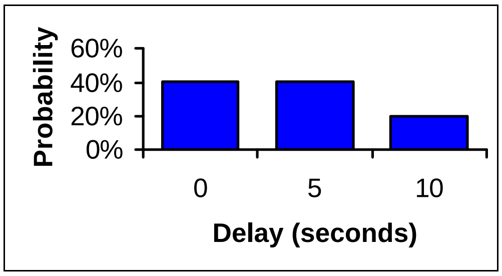


Figure 4: 30 Second Base Dwell Time Extension Probability Distribution



Other approaches that can be used to verify system robustness, include the blockage of tracks that could reflect a train break down or a failure in the track. This enables the system to be tested for its ability to cope with track blockages, as a result enhancing system stability.

Modelling

Timetable proving and robustness needs to be determined through modelling. This is usually undertaken with the use of specialist 'off-the-shelf' packages and, if required, integrated with tailored developed modules.

Specialist packages that are available include Simu++, M Train and Vista. Following the input of the timetable and physical characteristics, and the calibration of the model in the specialist package, timetable proving is undertaken. If necessary, the timetable or physical system characteristics is refined until the timetable is proven.

Timetable robustness tests are applied to the proven timetable. The ability of specialist packages to undertake robustness tests varies markedly. Therefore, specialist modules may need to be developed externally to the specialist packages, in order to complete the robustness tests. For example, some specialist packages cannot apply delay probability density functions to the timetable. These delays would need to be applied to the timetable prior to inputting the timetable into the specialist package. A sufficient number of tests would need to be undertaken to ensure that the results are statistically robust.

In the case study scenario, five disrupted simulations were undertaken for each of the six scenarios.

Element 5 - System Evaluation

Output from the simulation modelling exercise needs to be compared to the criteria established in the concept operating plan. This could include a percentage of trains running on time, conflict identification, and the average and distribution of delays by train type and period of the day (Radtke and Bendfeldt, undated).

Depending on the performance of the tested system against the criteria, elements of the operating system may need to be refined. Remedial measures could include adjusting the amount of recovery time, re-timetabling services, revising track layouts and adjusting the rolling stock fleet size. This re-iterative process should facilitate the optimisation of the proposed operating system plan.

To identify the most favorable option, comparisons should be made. This involves determining whether different options, or incremental enhancements to an existing option, result in significant improvements to the overall system performance.

Once the optimised system has been identified, a range of outputs including rolling stock fleet size, performance specifications, track layout and signaling requirements can be made to the model to further the design specification.

The high performance of the system in the case study scenario is due to moving block, and the train characteristics that have high acceleration and deceleration rates. Placing an extra car onto the train does not affect the performance of the system. It is not until the headway is decreased to 60 seconds that the system fails to meet the key performance benchmark. Delays for 60 second headway cases increase substantially over the change from 3 to 4 car sets per train, and from 120 second to 90 second headways. This finding is consistent with transit design practice, with 90 seconds being the lowest headway that can be currently catered for.

Delay distribution profiles at departure for the 6 scenarios are shown in figures 6 through 11.

Figure 6: 120 second Headway, 3 car train

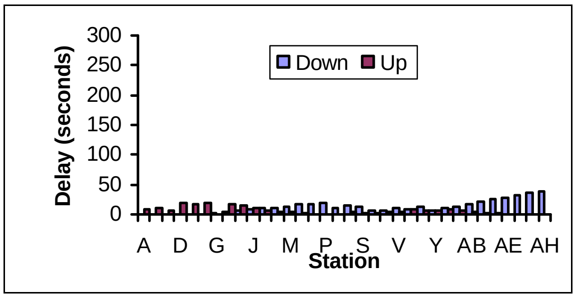


Figure 7: 120 second headway, 4 car train

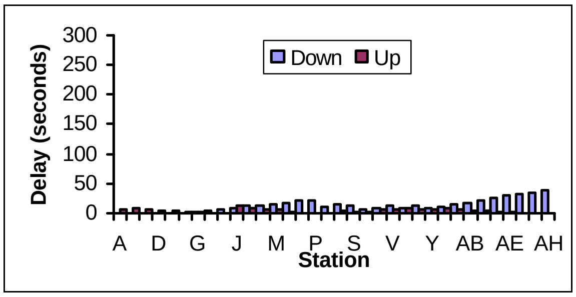


Figure 8: 90 second Headway, 3 car train

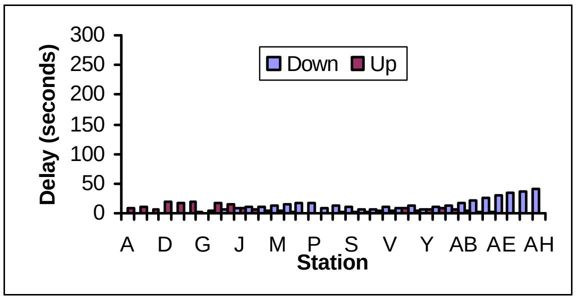


Figure 9: 90 second headway, 4 car train

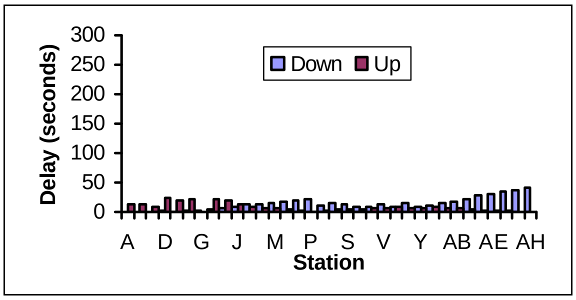


Figure 10: 60 second Headway, 3 car train

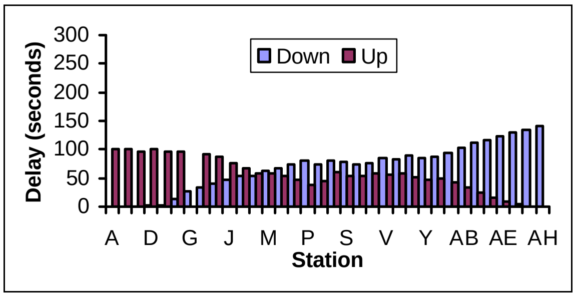
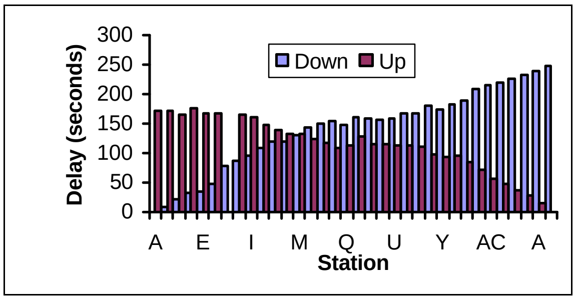


Figure 11: 60 second headway, 4 car train



It is clear in the distributions that the entrance to the system is at station A and AH for the down and up scenario respectively. From station A, delays gradually accumulate until the train leaves the system, and vice versa for the up scenario.

Some interesting characteristics can be noted in the delays over all stations, from which judgments regarding the infrastructure can be made. At station P, there is a rapid change in the magnitude of delays in both directions. This variation can be attributed to changes in the track gradient. The distributions show that the trains cannot accelerate up the gradient fast enough from a stop to reacquire lost time, *which results in cumulative delays over the increasing track gradient. This is clear with the 4 car set train on a 60 second headway, given that it must pull one third more weight than the 3 car set train.*

Table 3: Key Performance Indicators. Percentage of Trains Arriving Within 2 Minutes of Schedule

Scenario	Station A		Station U		Station AH	
	Up Case	Down Case	Up Case	Down Case	Up Case	Down Case
120 sec., 3 car	99.8%	100%	100%	100%	100%	99.2%
90 sec., 3 car	100%	100%	100%	100%	100%	99.6%
60 sec., 3 car	35.4%	100%	90.2%	74.8%	100%	8.2%
120 sec., 4 car	100%	100%	100%	100%	100%	99.6%
90 sec., 4 car	100%	100%	100%	100%	100%	99.6%
60 sec., 4 car	12%	100%	28%	13%	100%	4.8%

Clearly the cases involving 60 second headways did not meet the key performance indicator benchmark of 98% of services within 2 minutes of schedule.

Based on this analysis, the operating strategy is to initially provide three car sets at 120 second headways and gradually decrease headways commensurate with demand. It has been found that the 3 car set operating at 90 seconds should be able to accommodate the critical demand in the final year of operation. However, for patronage greater than 36,000 per direction over the two hour period, 4 car sets should be used instead of 3 car sets. This is because 3 car sets would have to operate headways of less than 90 seconds, beyond which reliability rapidly declines. At these close headways, passengers are unlikely to discern differences in waiting time, but would react more strongly to an unreliable and more unevenly loaded service.

Table 4: Service Pattern to Meet Various Demand Outcomes

Demand Case	Demand (No)	No. Cars	Headway (secs)	Two Hour Passenger Capacity (One Way)
Expected Open Year	28,000	3	120	28,800
High Open Year	35,000	3	90	38,400
Expected Final Year	32,000	3	105	32,914
High Open Year	42,000	4	105	43,885

Conclusion

High capacity transit systems are becoming more prevalent in many cities, to alleviate the dependency on personal vehicles, and as a means to accommodate public transport patronage growth. Increasing the capacity of existing heavy rail corridors to very high frequencies in built up areas maximises the value of these corridors, whilst potentially alleviating the need to construct a costly new corridor.

When experiencing major delays, a service on a high capacity transit system is likely to have 'knock – on' effects that impact on a number of following services, due to the close timetabled headways. An operations planning process has been presented that refines and optimises a proposed operating system, to ensure that it is workable and robust.

References

Binnie Black and Veatch, URS Consulting and Booz Allen Hamilton (unpublished) *Feasibility and Preliminary Engineering Study for the Proposed Eastern Region Line*, Singapore

Booz Allen Hamilton (unpublished) *Valuation of Public Transport Attributes*, Wellington, New Zealand

Department of Infrastructure (2000) *Track Record*, Australia

Department for Planning and Infrastructure (2002) *Perth Urban Rail Development* http://www.purd.transperth.wa.gov.au

Department of Premier and Cabinet (2002) *Growing Victoria Together* State Government of Victoria, Australia

Department of Transport (1995) *Perth Metropolitan Transport Strategy 1995 – 2029,* Department of Transport, Western Australia

Department of Transport (2002) *Major Scheme Appraisal in Local Transport Plans Part 3: Detailed Guidance on Forecasting Models for Major Transport Schemes - June 2002,* http://www.local-transport.dft.gov.uk/

Hammond,I. (2002) *First X'Trapolis Train Arrives* Transit Australia, Volume 57, No. 7 pp 175 – 178.

Land Transport Authority (2001) *Land Transport Directory of Singapore,* Land Transport Authority, Singapore.

Radtke, A and Bendfeldt, J.P. (Undated) *Handling of Railway Operation Problems with RailSys*, University of Hanover, Germany

Transportation Research Board (1996) *Rail Transit Capacity: Report 13*, National Academy Press, Washington D.C.

Urban Redevelopment Authority (2002) *Concept 21* http://www.ura.gov.sg, Singapore.