The pursuit of statistical validity in analysing rail network enhancements, and the consequential development of the Traxim rail network modelling tool.

Derek Harris[1], Stuart Mau[1]
[1]Ovalovo Pty Ltd
E-mail: derek.andrew.harris@outlook.com, stuart@deauca.com

Abstract
The challenge of planning for growth and performance improvement on the Australian interstate and Hunter Valley coal networks catalysed a decades long pursuit of a practical tool to support quantitative analysis of competing rail network enhancement options. While inductive approaches to analysing performance, such as the International Union of Railways capacity calculation methods, offer considerable insight and provide a platform for hypothesis testing, they are unable to adequately synthesise the many complex interactions present in a large, heterogeneous rail network. The desirable complement to the inductive approach is to generate a population of perturbed timetables to give statistically significant performance outcomes for competing options for future infrastructure configuration and operational practices. However, the time involved in manually generating timetables, and the risk of systematic bias in a manual approach, demands a tool that can automate and objectify the process. Such a tool needs to be sufficiently macroscopic that it can generate viable timetables in a pragmatically short timeframe, while being sufficiently microscopic that it enables efficient and realistic replication of the many input variables that impact real-world operational performance. The authors have now advanced such a tool, Traxim, to a mature and commercially viable state.

Keywords
Timetabling problem, Timetable deconfliction, Automated timetable generation, Timetabling optimisation, Statistical rail network analysis.

1 Introduction

This paper will outline the challenge of strategic planning for a large, complex, heterogeneous rail network, and describe a tool that, inspired by those challenges, has now been matured to a commercially viable state to support the statistically valid quantification of potential infrastructure enhancements. The purpose is to give insight into some real-world issues at a strategic planning level, and to introduce Traxim as a solution for enhancing the robustness of the options assessment process.

Section 2 of this paper will describe the Australian interstate and Hunter Valley coal network, which were the incubator for the development of the tool, and give some insights into the nature of the challenge in robustly analysing competing investment options.

Section 3 will then describe the Traxim tool, with particular focus on the functional choices that have been made in developing the software and the logic for those choices.

Section 4 is a discussion of some of the conceptual issues encountered in network modelling, and observations on the way in which Traxim approaches these issues.

Section 5 provides a short summary and conclusion.

2 Strategic Analysis of a Large, Complex, Heterogeneous Rail Network

2.1 The Australian Interstate and Hunter Valley Coal Rail Networks

The majority of the Australian interstate rail network, and the entirety of the Hunter Valley coal network, is controlled and managed by the Australian Rail Track Corporation (ARTC), where the authors previously worked in the strategy division. ARTC assumed control of the network progressively between 1998 and 2010.

The interstate network can be characterised as a long thin network. It is approximately 8,500 kilometres, but with a relatively low train frequency. It is dominated by interstate intermodal freight, where the overarching strategic imperative is to simultaneously reduce both transit time and train cost structure. Most of the network is single track with crossing loops at relatively long spacing. Growth over the past 20 years has been underwhelming despite extensive upgrading of the capacity and performance of the network. While there has been some loss of market share to road, particularly on the shorter corridors, a significant contributor to the weak growth has been the decline of the Australian manufacturing industry. Domestic manufactures are the primary driver of interstate freight volumes. Arguably, in the absence of the enhancement initiatives undertaken over the last 20 years, road would have eroded rail volumes more than it has. While interstate freight dominates, the network is predominantly a heterogeneous mixture of passenger, bulk and intermodal trains. In some areas there is a differentiation between intermodal trains, with relatively faster and slower trains as well as the typical train.

In contrast, the Hunter Valley coal network saw a dramatic surge in demand over the 2004 – 2016 period, with tonnage doubling and tonne kilometres tripling as coal was increasingly sourced from mines significantly further from the port at Newcastle. This network consists of a 128 km core with double, triple and quadruple track sections, and two single track branch lines of 252 km and 170 km respectively. While the network is dominated by largely homogenous coal trains, there is enough non-coal traffic to create significant operational complexity. The primary objective on the Hunter Valley network has always been to keep capacity ahead of demand, while ensuring efficient train flow.

2.2 The Role of Theoretical Timetabling

Good decision making about rail infrastructure investment requires an analytical foundation, desirably including accurate quantification of the operational consequences of different infrastructure options.

Many aspects of rail operations can be modelled using a spreadsheet and relatively simple formulas. For instance, the UIC406 compression approach to track utilisation can be generated in a spreadsheet. The advantages of this approach are that it is simple, relatively quick, and highly transparent. This approach was the foundation of the network modelling of the Hunter Valley coal network over its period of growth and provided a solid foundation for the investment strategy that was successful in keeping capacity ahead of the surging demand

However, such “rule-of-thumb” tools necessarily risk oversimplifying complex problems. They struggle to deal with varying loop lengths, different train speeds and train priorities, peaking, and the effects of following conflicts.

They can only deal with relatively plain track – it is hard to assess junctions, intermediate signals and so on. Some essential parameters, such as a threshold for practical capacity utilisation, are arbitrary even though they may be derived from extensive practical experience.

They can only deal easily with a line section - there is no validation of interactions between sections.

As such, they are best at allowing an analyst to gain a sense of how performance is likely to respond under different scenarios, and they allow rough “order of magnitude” quantification.

This style of analysis could be considered as essentially an inductive reasoning process. It allows an analyst to build a hypothesis about how the infrastructure might perform under different configurations or operational patterns.

As with any hypothesis though, it is highly desirable that it be validated through a controlled experiment. Typically, this is done by creating sample timetables based on given infrastructure and operational scenarios.

Rail infrastructure analysts may also adopt a deductive reasoning approach to infrastructure and operational issues. That is, they draw conclusions based on observation. This “observation” might be of current actual operations, or it might be insights gained in attempting to produce timetables for future scenarios. For example, an analyst might sum all crossing delays in a timetable to identify where on the network the greatest congestion is occurring.

Whether it is for validating a hypothesis or to gain insights by observation, the process of preparing and analysing timetables has two major shortcomings: It is highly labour intensive, and it lacks statistical validity.

The primary objective of the Traxim simulation software is to mitigate these shortcomings by automating the timetable generation process. It also creates a wealth of data that can be deployed to gain deeper insight into effects and their underlying causes.

2.3 Example Scenarios

To give a practical perspective on the challenges facing the analyst, and provide context for where a tool such as Traxim can support the analytical process, following are two case studies drawing from actual scenarios.

The first example, drawn from the interstate network, is the effect of the introduction of a new bulk traffic, say four trains per day in each direction carrying iron ore, over a sub-section of the network, say 500 km. While it would be largely straightforward to assess whether there is adequate capacity for the additional traffic using the compression method, in this case the key issue was the impact that the additional traffic would have on intermodal transit times. From a commercial perspective, the infrastructure owner may take the perspective that the introduction of the new traffic should not degrade intermodal transit times, since the market places a high value on delivery speed, and rail already struggles to compete with road transit times.

One approach to mitigation would be to give intermodal trains absolute priority over iron ore trains. This may lead to very long transit times for the iron ore trains, but might be the lowest overall cost solution. Alternatively, the iron ore trains could offset the extra crossing dwell by using a higher power-to-weight ratio. This has a direct capital and fuel cost, but might better meet the customer's overall business needs.

An alternative approach would be to only give intermodal trains a modestly higher priority over the iron ore trains, and to invest in enhancements that enabled intermodal trains to maintain their transit times despite the increase in train numbers. Such enhancements might be in the form of additional crossing infrastructure, but could also be a range of other interventions such as improved signalling systems, curve easings, or deviations. Those enhancements may not necessarily be on that part of the network being used by the proposed new traffic. If the objective is to maintain transit time, it would make sense to invest in enhancements wherever it is along the route that achieves a target time saving at the lowest cost.

To properly assess the range of options, it is necessary to quantify transit time for both the intermodal trains and the iron ore trains under each scenario. This can be done using theoretical models of transit time, with dwell being calculated using a probabilistic approach. However, it is inherently difficult, and sometimes impossible, to have regard to all of the interactions that impact on the transit time outcome when using a spreadsheet model. Even to the extent that it is possible to theorise an outcome, there is considerable room for conceptual error, and a properly constructed experiment will serve to validate the theoretical approach.

The second example is the situation that was regularly encountered in modelling coal volume growth on the Hunter Valley network, where a single-track line section (that is, a section between two crossing loops) is forecast to have its capacity exceeded. This can be determined by applying the UIC406 compression method using a spreadsheet. This involves extensive mechanical calculation to make allowance for such things as train cancellations, maintenance and surge capacity, as well as needing quality information for actual sectional running times and signal clearance times. (For more information, the annual Hunter Valley Corridor Capacity Strategy available on the ARTC website (https://www.artc.com.au/projects/hv-strategy/) provides a detailed explanation of these issues.)

There is also a need to set a threshold for the theoretical limit of capacity. Historically this was set at 65% for the relevant network, but has been uplifted as high as 73.5% as confidence has grown in the ability of the system to function effectively at a higher level of intensity.

In a situation where a single-track section is at risk of exceeding its capacity limit, the default solution is construction of a new crossing loop. This can be easily modelled with the compression method by inserting a new loop and theorising the new section running times and signal clearance times.

However, there are other potential capacity enhancement solutions. An intermediate signal could be installed at approximately the midway point of the section to allow following trains to simultaneously occupy the section. Track speed could be increased by allowing an increase in the permissible speed for the maximum axle load. This can be particularly useful on the approach to ruling gradients. The Australian rail system generally prohibits the simultaneous entry of trains at a crossing loop, that is, a train isn't allowed to have only one signal separation from the opposing train as it enters the loop. Hence another option to increase capacity is to reconfigure the loop with an intermediate signal to allow simultaneous entry.

Each of these options raises complex analytical questions.

Furthermore, while the compression method provides a good rule-of-thumb, the ability of any individual track section to operate effectively is partly dependent on the utilisation levels of the adjacent track sections. If those sections are also close to capacity, it is likely that some of the capacity on the capacity limiting section will be sterilised by gaps in the presentation of trains. This is not easy to model theoretically.

The utility of an intermediate signal will depend on the proportionality of following and opposing train movements. This can be hypothesised and theoretically modelled, but there is considerable latitude for error in the logic, and there are many secondary effects that will inevitably mean confidence levels in the theoretical analysis will be low.

Increasing train speed is somewhat analogous to constructing an additional intermediate loop insofar as it acts to reduce the capacity limiting section time. There is an additional complexity here though in that, ideally, capacity analysis will be based on actual observed train performance, but for an increase in train speed it is necessary to rely on simulation. To maintain internal validity, the simulation model needs to be calibrated using actual data, or, all analysis needs to be undertaken using the same simulation tool.

Finally, a simultaneous entry configuration at a loop rases similar issues as for an intermediate signal. In this case though the issue is the probability that two trains will arrive at a loop simultaneously. While this can, again, be hypothesised, there is considerable room for theoretical error.

To address the multiple dimensions of the options and assess them on a consistent basis the solution that presents itself is to undertake a statistical analysis of performance outcomes for each enhancement option using a population of deconflicted timetables under an appropriate sample input scenario.

These types of problems were the inspiration for the development of Traxim and drove the design choices discussed in the next section.

3 The Traxim Tool

3.1 Key Analytical Functions of Traxim

The functional outputs of the Traxim tool can be generalised to three key metrics: Assessment of track utilisation; identification of bottlenecks, and; estimation of transit time.

Track utilisation rates consistent with the compression method are automatically generated by Traxim. However, while a useful output, this doesn’t in itself reveal anything about the viability of the modelled scenario. For instance, if the defined section is between two loops and there is an intermediate signal, it is theoretically possible to exceed 100% utilisation. This may not, however be desirable.

Summing dwell at each location, and smoothing it to account for different track features, gives immediate insight into the parts of the network with the highest gross delay. This is highly useful for targeting locations that will give the largest transit time benefit for a given investment. It can also be useful as a metric to compare the options in our sample scenarios – dwell corresponds to congestion and congestion is a consequence of capacity utilisation. Hence it is an indirect measure of the capacity benefit of different solutions.

The estimation of transit time follows from the calculation of dwell, but at a train rather than geographic level. Transit time also integrates deceleration and acceleration, and hence is a somewhat more pure metric of the overall performance of different solutions. As already noted, transit time and congestion are interlinked, especially on single track lines, so it is also effectively an indirect measure of capacity utilisation.

3.2 Functional specification

To generate the analytical outputs required, Traxim was designed with a number of important functional elements.

All trains are dynamically simulated based on detailed track geography, respecting the lesser of track speed and train speed, and incorporating turnout speed constraints. All stops are dynamically simulated to ensure accuracy. Signal clearance time can be set either generically, or as a bespoke value for individual turnouts and signals.

The infrastructure is defined in terms of turnouts and intermediate signals. To simplify the infrastructure model construction, turnouts are assumed to have a generic signalling configuration. Each track section can be defined as bi-directional or uni-directional. Turnouts are defined with a default branch to support rational pathing of trains.

Trains are defined in terms of length, trailing weight, maximum train speed and locomotive configuration. The dynamic simulation is based on the sum of the tractive effort curves for each locomotive. Driving method, in terms of acceleration and braking rates, can be user defined, and practical maximum speed can be adjusted as a proportion of theoretical maximum speed to give a further dimension when calibrating to observed train performance.

Train pathing through the network is automatic, based on a principle of minimising the number non-default branches a train takes through turnouts. As previously noted, each turnout can be defined with a default path, either “through”, “diverge”, or “neutral”, that gives the user the ability to nudge the train onto a preferred route, which is particularly useful on multiple track sections. A train can also be forced to pass through, or not pass through, specified nodes, giving direct control over train pathing if required. Rules for stopping and dwelling at a node, or not departing a node before an earliest time, can also be set.

Trains are given an individual priority, which is unbounded. Trains can also be specified to have absolute priority, which guarantees that in conflicts they will only yield to other trains with absolute priority.

Trains can also be linked.

For the purposes of producing a clean data set, Traxim has a user defined run-in period, to ensure that the resolved timetable has reached a steady state. It uses a user-defined window period to ensure that the data collected on metrics such as dwell and distance relates only to the timetable while it is in a steady state, that is, it excludes both the run-in and run-out periods.

Importantly, the software is quite forgiving in terms of the inputs. In most cases it will compensate for poor, missing or inconsistent data rather than fail. This can be helpful for early testing of concepts, or for simplifying the analytical task where some detail may not be important to the rigour of the output.

This set of functional specifications was considered to offer a good balance between precision of the simulation on the one hand, and simplicity and ease of product use on the other.

3.3 Analytical Outputs

Traxim produces three sets of outputs: train graphs; tabular data, and; geographically referenced KML (Google Earth) files.

Train Graph

Firstly, it generates a train graph (or stringline) in pdf form for each perturbation, including full blocking information, and with train paths colour coded by train group. This is useful to review to ensure that the inputs are generating the form of output that was expected. It can also give visual clues as to patterns of system behaviour. An excerpt from a sample train graph is shown in figure 1.

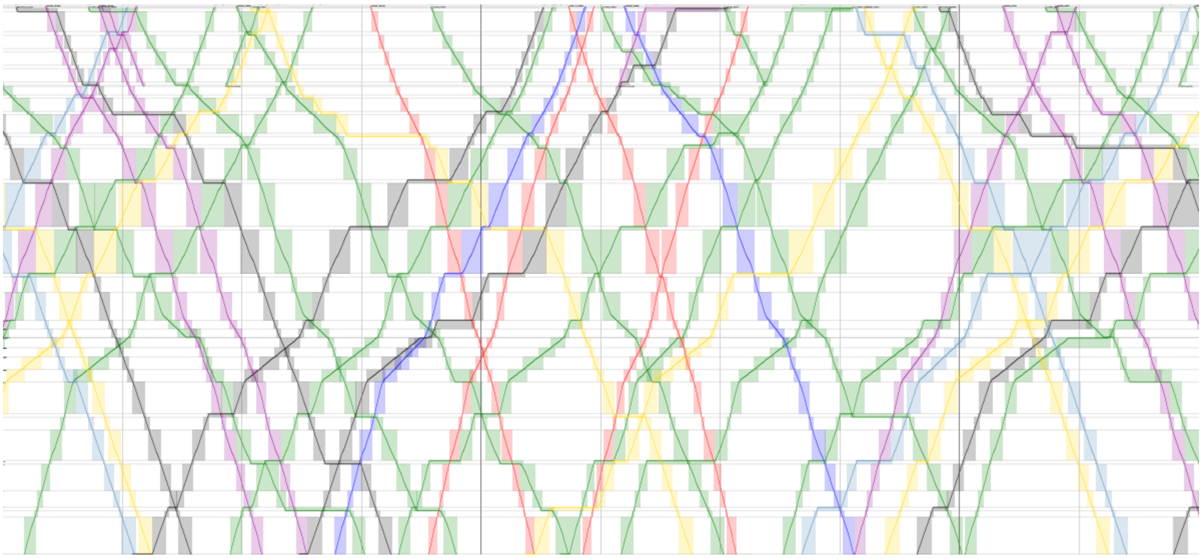

Figure 1: Example Train Graph Excerpt

Tabular Data

A wide array of data for both each individual perturbation, and for a population of timetables for a given scenario, is generated and recorded. This includes section utilisation, transit time, dwell and distance travelled. Data can be automatically aggregated by user defined train groups and regions. All of the data is easily manipulated in Excel, allowing extensive analysis and graphing of results.

Google Earth KML

A kml file showing delay by location is generated for the average of all timetables in a scenario. The kml file can be opened in Google Earth to get an immediate visual analysis of where the greatest congestion on the network is. Multiple kml's can be opened to allow a visual comparison of competing solutions.

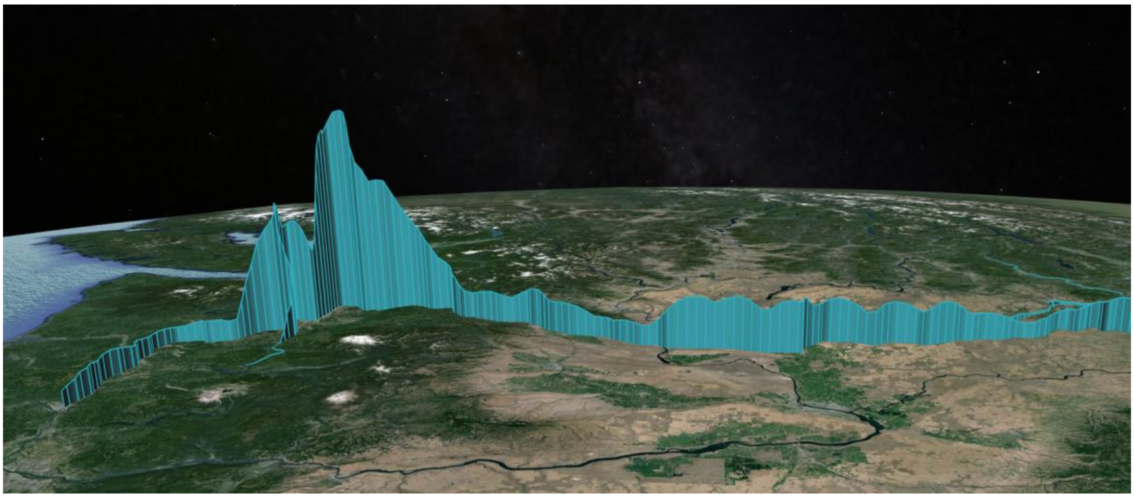

Figure 2: Example Google Earth Graphic

3.4 Technical description of Traxim

Traxim evolved organically to meet a practical strategic need. As such it can't claim have a strong theoretical foundation. However, to place Traxim within a theoretical framework, it could be considered to have the following characteristics as defined by Medeossi and de Fabris (2018):

Firstly, it has both continuous and discrete elements. Traxim dynamically simulates train performance continuously, but conflicts and their resolution are discrete events.

Secondly, it is stochastic, though in a mild sense. Timetables are generated with randomisation around train entry times (within user defined bounds) and crossing decisions (within a choice framework where the utility of a crossing option is weighting based on train priority). Both are discrete uniform distributions. All other elements are deterministic. Randomisation of train entry times is used to reflect real-world external uncertainty. Randomisation of crossing decision making is used to ensure that a viable solution is found to all conflicts (within the capacity limits of the network) while recognising that in a non-homogeneous network, crossing choice based purely on priority weighted localised dwell won't necessarily result in an optimised outcome. These randomisation elements mean that when seeded differently, markedly different timetables are generated for the same infrastructure and train plan scenario, which provides the basis for statistically robust results when assessing the performance of a given configuration.

Microscopic (up to a point): All track and signalling is modelled in Traxim, but signals are only modelled as two-aspect, and signal locations at turnouts are simplified. There is only limited control of available routes through interlockings. Traxim was developed primarily for longer-distance, thinly utilised, mixed-use rail networks rather than compact, intensive, passenger dominated networks. As such, natural variability (due to driver behaviour, equipment condition and weather) creates levels of variance that would generally overwhelm any differences in performance that would be apparent as a result of micro-simulation of interlockings. As such, the chosen trade-off between precision and simplicity is considered appropriate.

Synchronous: The Traxim algorithm resolves each conflict sequentially. While it provides for a category of train with absolute priority, which will only be delayed when in conflict with another absolute priority train, this prioritisation is applied as part of the sequential process. All crossing decisions are made having regard to the relative amounts of delay, and the relative priorities, of the trains in conflict.

Non-timetable based: Traxim is primarily directed toward simulation of networks with an unstructured operating pattern. As such, the timetable is appropriately an output rather than an input to the simulation.

4 Observations on Traxim's approach to dealing with conceptual modelling challenges

4.1 Randomness

Traxim uses randomness as a mechanism to generate unique timetables.

The randomness is generated in two ways. It is primarily generated through perturbation of train departures. Train departure times are defined as a window with a user defined tolerance, with actual departure being randomly determined as a time within the nominated window with a uniform distribution, though this window can be set to zero. Secondly, there is also an element of potential randomness generated by the deconfliction process. This is discussed in more detail below, but it is worth noting that even if every train is set to a departure window with zero tolerance, the timetable generated will still be somewhat randomised.

The concept of randomness is fundamental to the underlying purpose of Traxim, that is, to provide a statistically significant quantification of the performance of different scenarios. As such, it is important that the user provide sufficient latitude in setting the departure windows that there isn't systematic bias in the results. At the same time, the departure windows need to be meaningful in the context of the intended operational mode of the railway. For instance, in the Australian intermodal market, customers prefer late evening train departures. It would be meaningless then to create a scenario where intermodal trains were artificially randomised into windows that were not market attractive. This is ultimately a matter for user judgement. Fortunately, experience suggests that the "butterfly effect" is very strong, especially with single track networks, and hence modest variation in train entry times is sufficient to create high levels of randomisation.

4.2 Optimisation

The approach to conflict resolution in Traxim is driven by a hierarchy of decision making that balances minimisation of weighted delay within the discrete event against efficiently finding a viable resolution to the conflict. Specifically, where a conflict exists in isolation, the algorithm will preference minimising priority weighted delay in the conflict solution. If that conflict starts to interact with other conflicts though, it will relax this preference to enable identification of a viable resolved timetable.

Hence, Traxim is not an optimisation tool as such, though it's found solutions lean toward good local optimisation. Due to the inbuilt randomness in the timetable generation process, some timetables will achieve greater optimisation (in terms of less weighted delay) than others.

For the purposes of infrastructure analysis and planning though, it is arguably desirable to avoid excessive optimisation. The level of optimisation of the network in real time is purely dependent on the performance of individual train controllers. Even where support tools like Movement Planner or Nitro are used, these tools don't provide any sort of global optimisation. Their train plans are only optimised to a similar level as those of Traxim. While train controllers do their best and generally have extensive experience to draw on in making their decisions, the practical reality of the real-time environment and limits of human performance, and the current limits of computing, mean that it is unlikely that live-run performance will be materially better than an average of timetables generated by Traxim.

For this reason, in using Traxim the preference is to use the whole population of timetables generated rather than preferencing those timetables that are relatively optimised.

4.3 Relationship with Network Reliability

An important question for railway planners and operators is 'how reliable will my performance be?' Traxim seeks to achieve a performance level equivalent to how a predominantly freight railway would operate in live-run. The reality is that such railways essentially operate in a permanently disrupted state due to natural variation in train entry times and on-track train performance. The results Traxim produces are therefore, in a sense, the 'unreliable' state.

The extent to which a railway is perceived to be unreliable will depend on the difference in delay between the planned timetable and the typical disrupted state. That is, if a railway adopts a highly optimised Master Train Plan, with well-designed crosses and minimal delay, it would be expected to be relatively ‘unreliable’ when subjected to disruption. If the timetable is inefficient, with poorly designed crosses, lots of additional crossing allowances, and scheduled recovery time, it could be expected to be highly reliable.

In interpreting an analysis undertaken using Traxim, the typical timetabling practice of the modelled railway needs to be considered. To the extent that delay in the Master Train Plan is greater than predicted in an equivalent Traxim simulation, it would be expected that trains entering late, or experiencing problems enroute, would not incur any further delay and may even recover lost time. To the extent that the timetable has less delay than predicted in Traxim, trains that enter late or encounter problems are likely to, on average, lose further time.

4.4 Relationship to a Master Train Plan

A timetable as generated by Traxim is not expected to be strictly the same as a railway “working timetable” or “master train plan”.

The key difference is that a Traxim timetable does not “wrap” around the week. That is, the position of the trains at the start of the timetable is not identical to the position of the trains one week (or other recurrence frequency) later.

It would also be unusual for Traxim to produce a timetable with a train path pattern that repeated itself on multiple days, even though this is a reasonably common characteristic of master train plans.

Traxim does not allow for a variety of legitimate factors that may influence crossing decisions in a manual timetable production process. The major one, at least in the Australian context, is grandfathering of paths – that is, ensuring that long-standing services continue to achieve certain transit time and arrival time performance.

Each of these characteristics of a manual timetabling process is likely to introduce some “inefficiency” into the timetable, manifested as increased delay. To the extent that they do so, a Traxim based analysis will not pick this up.

4.5 Deadlocking and Unresolvable Scenarios

A fundamental problem of simulating train operations is the risk of deadlocking, that is, a situation arising where no train can move forward due to other trains blocking its path.

Traxim is designed such that it will not get into a deadlock situation.

However, where track utilisation exceeds the capacity of the infrastructure, Traxim may not find a viable solution. There is no precise way to define the limits of what can be resolved – it depends not just on the utilisation of individual sections, but the intensity of utilisation of the network as a whole. Typically though, it will become unviable to identify a solution where train entries exceed the capacity of a region of the network for more than around six hours.

This in itself though provides a useful metric of capacity. If it is accepted that Traxim is sophisticated and effective in generating timetables where it is viable to do so, exhausting its ability to identify a valid timetable solution is a clear indicator that the proposed combination of infrastructure and trains is not realistically viable from a capacity perspective.

4.6 Appropriate Uses and Limitations of Traxim

Traxim is most effective in analysing networks that are exclusively or predominantly for freight, or for passenger services that do not operate to a repeating ‘clockface’ pattern. While Traxim can preserve a preferred pattern of passenger operations by ascribing them absolute priority, a predominantly passenger rail network with a clockface pattern is likely to be better analysed in the context of an ‘intelligent design’ process.

Returning to the two case studies described in section 2, Traxim offers a useful complement or alternative to reliance on the manual methods of analysis discussed.

Its ability to quickly and reliably estimate transit time makes it an ideal tool for any scenario where the analyst is seeking to understand the transit time outcomes of different scenarios, such as in the intermodal network example discussed.

Testing network capacity is, necessarily, a more arbitrary exercise. Irrespective of the method used, the threshold for an acceptable level of capacity utilisation is ultimately a value judgement. Where Traxim can help in that process is by providing a consistent basis for examining the limits to which the network can be pushed. Specifically, in this case the task would be to identify the volume scenario at which Traxim ceased to be able to resolve timetables, and to compare the uplift in that threshold across different capacity enhancement options. This would, for instance, allow the quantification of an intermediate signal option relative to an intermediate loop option. The relative uplift can then be applied to a common baseline to infer the absolute level of capacity added.

5 Conclusions

This paper has described the context in which an automatic rail timetable generation tool, Traxim, has been developed, and the key features and conceptual modelling considerations associated with its use. The complexities of a large, heterogeneous rail network are such that it has historically been challenging to establish high levels of confidence in the benefit afforded by competing enhancement options. The ability of the Traxim tool to rapidly generate a population of viable, randomised, quasi-optimised timetables, appears to fill a gap within the range of existing products in the market, and provides a valuable complement or alternative to manual approaches to network analysis.

References

Medeossi, G., & de Fabris, S. (2018). Simulation of rail operations. Handbook of Optimization in the Railway Industry, 1-24.