Chapter X

Using Optimisation and Machine Learning to Validate the Value of Infrastructure Investments

García-Flores, Rodolfo*
CSIRO Data 61
Private Bag 10, Clayton South, VIC 3168, Australia

Banerjee, Soumya
Mathews, George
CSIRO Data61
Australian Technology Park
Level 5, 13 Garden Street, Eveleigh, NSW 2015

Key Words: Infrastructure Value, Port Logistics, Infrastructure Planning, Rail Scheduling, Strategic Planning **AMS Subject Classification:** 90-00, 90-02, 90B80.

*E-mail address: Rodolfo.Garcia-Flores@csiro.au

Abstract

When stakeholders commit to building infrastructure as part of strategic, long-term planning, the final facilities are not normally amenable to modification after completion. A consequence of this is that users are forced to operate within the original specifications for, at least, as long as it takes to carry out major refurbishments or retrofitting, and even then, the constraints imposed by the original layout may be inescapable.

On one hand, the original infrastructure plans enhance (or limit) the users' ability to operate efficiently for years to come. As time passes and the payback period approaches, changing operating conditions and unforeseen bottlenecks in the original blueprint can, at best, affect the economic returns and, at worst, defeat the purpose of the whole project (see, for example, Castellón airport in Spain, which was built but is grossly underutilised), producing unanticipated economical, social and political repercussions. On the other hand, managers and operators (that is, those living with the consequences of the strategic planning) have some leeway to compensate for miscalculations by means of their tactical and operational planning.

In this chapter, we explore the use of quantitative techniques to, first, amend bottlenecks and uncertain market and operating conditions that affect the performance of infrastructure investments (the *tactic* and *operational* levels), and second, validate the effectiveness of the original infrastructure design (the *strategic* level) under these changing conditions.

More specifically, we present a rail scheduling case study where we combine demand forecasting using Machine Learning techniques and formal Operations Research methods to assess and maximise the value of already-existing infrastructure. Rail scheduling is a typical optimisation problem popular in the literature, but its potential value is bounded not only by its technical properties and specifications ("how good the algorithm is") but also by the accuracy of data feeding the algorithm. Such data is critical in specifying the demand that a facility will experience in the future, and the costs that will be incurred to operate it. The use of intensive data analytics and appropriate Machine Learning techniques can resolve this and provide a substantial competitive edge for investors and operators of rail inter-modal terminals.

We anticipate that Machine Learning algorithms that predict future demand, coupled with optimisation techniques that streamline operations of facilities, can be integrated to create tools that help policy makers and terminal operators maximise the value of their current infrastructure, while meeting ever-changing demand.

1. INTRODUCTION

When governments or industry commit to building infrastructure, the resulting facilities are not normally amenable to modification after completion. Since these facilities are the product of strategic planning, they normally require very large investments and are meant to operate continuously for many years in the future, possibly with a few major refurbishments or retrofitting during the course of their lifetimes. This means that, despite having little flexibility to accommodate for changes in market conditions or in the natural environment, they must still face the constraints and bottlenecks produced by these changes, and which may not have been

foreseen originally. This can produce unanticipated economical, social and political repercussions. A clear example is Castellón airport in Spain, which is currently in use but underutilised [1].

Fortunately, managers and operators have some margin to compensate for miscalculations by playing with their operational and tactical plans. Managers and operators must take advantage of the infrastructure available to them and "juggle" their resources in order to maximise their productivity, even in cases where the extant facilities represent more a limitation (or, in extreme situations, a liability) than an asset. In this chapter, we centre the discussion on quantitative approaches to assess the performance of large infrastructure investments when there are unforeseen changes in economic and environmental conditions, not taken into account at the strategical planning stage. These changes may be caused by:

1. *Variability in demand and supply.* Supply and demand are the most commonly used design variables for determining the capacity of new infrastructure investments, and their variability impacts directly on payback periods and facility utilisation.

2. *The economic environment.* In addition to supply and demand, other economic variables affect substantially the performance in time of large infrastructure facilities. These include fluctuations in market prices, labour costs, debts and subsidies, interest rates, insurance, and country risks.

3. *Social conditions.* Political will is not only the trigger to develop large construction projects, but often is also the engine that keeps them in operation, shuts them down, or carries out extensive refurbishment or auditing. Decision makers often know from the outset that some projects are not profitable, but they are realised nonetheless because they are indispensable to serve the population's needs. This is commonly the case with water networks and some transport infrastructure.

4. *The natural environment.* Increasingly, the environment is a crucial concern not only because of the need to preserve and maintain sustainable ecosystems, but also because of the massive changes that human activity is producing on global natural equilibria. Projects must now be designed to cope with increasing temperatures, rising sea levels, drought, floods, and higher population densities in most cities.

5. *Disruptions caused by market conditions.* In addition to the social and economic factors listed in items 1 to 3 above, less predictable situations can also affect projected operation plans. These include economic crises, conflicts, shifts in public attitudes, bankruptcies and disruptions caused by new technologies.

6. *Accidents and natural disasters.* Accidents and natural disasters can seriously affect the integrity of physical infrastructure, but even in cases where the effects

[1] `https://en.wikipedia.org/wiki/Castell%C3%B3n%E2%80%93Costa_Azahar_Airport`, accessed on the 11 of April 2016

of these events on operations are short-term, they may derail strategic and tactic objectives and thus compromise the long-term viability of the facilities.

Often, these changes occur simultaneously and it becomes difficult to calculate the magnitude of their individual contributions.

Adopting a quantitative approach for infrastructure assessment has many advantages, besides the objective comparison of scenarios: by verbalising their problem, the stakeholders are forced to reflect on the rules and constraints that actually define their operations; they are compelled to adjust their expectations and articulate trade-offs explicitly; they can visualise alternatives of what is viable and what is not by examining exact or approximate solutions. Finally, delivering software to automatically solve the problem repeatedly allows operators to test what-if scenarios and can release valuable staff time for use in other business priorities.

The purpose of this chapter is to demonstrate that the combination of Operations Research (OR) and Machine Learning (ML) is an appropriate methodology to assess the value of existing infrastructure: if, on one hand, by using the best possible operation schedules calculated with plausible scenarios, the existing infrastructure and resources can cope with demand and avoid bottlenecks, the infrastructure is valuable and adequate. If, on the other hand, only uncertain forecasts and poor infrastructure are available, the calculation of a good schedule can still provide insights on where resources should be invested in the future and what aspects of the operation require improved collection of data.

Figure 1 is a schematic of the main quantitative techniques reported in the literature that have been used to validate the value of infrastructure investments. The focus of the reviewed projects, which use a variety of analytical tools (simulation, optimisation and various heuristics), is normally on assessing the economics, the value to society, or the effect of disruptions over the normal operations of the infrastructure.

The structure of this chapter is as follows. Section 2. presents a non-exhaustive literature review of papers organised by the quantitative technique used, and grouped by area of application, using as a rough guide the schematic in Figure 1. Section 3. introduces the rail scheduling case study where we combine demand forecasting using ML techniques and formal OR methods to assess and maximise the value of already-existing infrastructure. Rail scheduling problems are defined not only by the system's own resources and constraints, but also by the accuracy of the parameters used, which are critical to correctly assess the value of the infrastructure. The discussion in this section and of the results in Section 4. will centre on the combined use of OR and ML techniques. Finally, Section 5. rounds up the discussion.

2. PREVIOUS WORK

The following review is organised around quantitative techniques to assess the value of infrastructure, and within the techniques, by the domain of application of the infrastructure being assessed. Many other concepts could have been used to organise

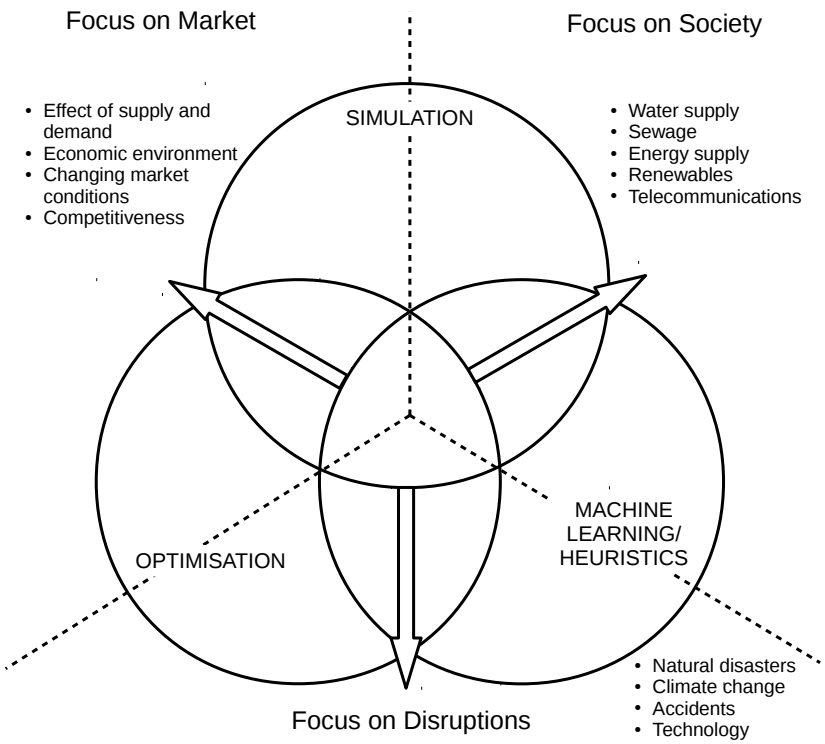

Figure 1. The main quantitative techniques used to validate the value of infrastructure investments. The area of application of the validation tools is normally the economics, the value to society or the effect of disruptions over the normal operations of the infrastructure.

the existing bibliography, as for example motivations (e.g., climate change mitigation, sector the infrastructure is used in, understanding economic cycles), cause of uncertainty (planned or unplanned), and solution approaches (empirical, quantitative, network-based, statistical, survey-based), all of which are equally valid. For a review on actions to mitigate climate change, we refer the reader to Trück et al. (2010), who presented an extensive study of the alternatives at the local level on the cost of mitigation and adaptation to climate change effects. For a complete review for managing disruptions in the abstract using OR techniques, please refer to Snyder et al. (2016), and for a similar review in the context of disaster management, to Galindo and Batta (2013).

2.1. Simulation

Simulation is the most popular approach to assess the performance of existing infrastructure. It works on the principle that the system's behaviour emerges from replicating its underlying structure, and has the advantage over analytical optimisation methods that it can describe large and complex systems. The following review illustrates major areas on which simulation has been used to validate the value of existing and planned infrastructure investments: services (such as water networks and energy), risk assessment (i.e., natural disasters and accidents), transportation and logistics, energy and industrial facilities, and is by no means exhaustive.

Simulation is widely used to validate investments in water infrastructure, as population grows and cities continue to develop while governments' resources become increasingly limited. Aleisa et al. (2011) presented a simulation model to make projections of future demand for Kuwait's water network, including the outputs of four existing waste water treatment plants, and concluded that only one of the plants will require a significant upgrade. Etchells and Malano (2005) focused on the increasing uncertainty in water supplies, analysed the sources of uncertainty in water allocation models, and reviewed a few existing models. They emphasised the current deficiencies of software models to validate existing infrastructure, given the uncertainty of all the factors involved. Harris et al. (2009) presented an extensive validation of a model that considers existing infrastructure at Cockburn Sound, Western Australia. The model took into account the capacity and influence of a desalination plant, the effect of harbour modifications and the development of quays, although the objective of this study was to simplify the environmental approval process, rather than assess performance of existing infrastructure. Smajgl et al. (2013) discussed an agent-based model to assess the impact of mainstream dams in the Mekong river area, land use changes, large-scale irrigation, and sea level rise. The model was useful to understand unintended side-effects of new investments (hydro-power, rubber plantations, irrigation schemes for food and energy crops, and mining), paying especial attention on their effect in achieving poverty alleviation targets. Tjandraatmadja et al. (2013) presented a modelling framework for Makassar City, Indonesia, that assesses the suitability of existing and proposed infrastructure to ensure fresh water supply. The framework considered the effects of population growth and climate change and proposes alternatives to enhance the city's water

security.

Power load balance is a common theme of simulation projects dedicated to assess the value of electricity network infrastructure. To cite a few examples, Quezada et al. (2014) present an ambitious agent-based system to assess the effect of socio-technical factors that cause stress in the electric distribution network in southeast Queensland, including climate change and investment inertia, and propose adaptation strategies. Köpp et al. (2010) developed a demand and supply forecast method based on neural networks to help cope with the unavoidable intermittency of supply by solar and wind electricity generators. Additionally, they used a simulation to demonstrate that a smooth load balance is possible by using adequate control technology on both the consumer and the producer sides. Huang et al. (2015) combined simulation and optimisation to smooth the power load in electricity distribution networks. In this paper, time-varying supply and demand information is used to obtain smooth schedules with minimal peak power and generation costs by using a non-linear model, and the results are validated via simulation.

Some simulation-based assessments of private industrial premises exist, although they are not as numerous as government-funded energy and water projects. This is partly because they tend to be not as capital-intensive, and partly because they are not open to public scrutiny, and thus they are not often reported in the literature. For example, Berends and Romme (2001) used simulation to understand the causes and effects that economic cycles have on capital-intensive industries, and in particular in the paper industry. Like many activities that require large investments, the paper industry is exposed to economic downturns, economies of scale, the need to "keep the machines running", the incentives to keep investing large amounts of capital in order to remain competitive (often in detriment of market prices), and the lag between investment decisions and the moment the new capacity is available for actual production[2]. Berends and Romme validated the model for price and capacity, and demonstrated that cyclicality is to a large extent endogenous (e.g., produced by adoption of new technology) and not only dependent on external market conditions.

Regarding the assessment of logistics and transportation infrastructure, we notice that much recent work uses newly collected data sets that were not available until recently, thanks to recent advances in on-site sensing and data collection. Thekdi and Lambert (2015) introduced the Corridor Trace Analysis tool for assessing the impact of developments on adjacent land on road transport networks. The tool prioritised corridor segments that are vulnerable to adjacent land development. Thekdi and Lambert made use of a wealth of data not previously available, thanks to increased digitisation of transport operations, advances in satellite imagery, increased data storage capacities, and increased access to public data resources. Similarly, Higgins et al. (2013) and García-Flores et al. (2014) used recently collected data sets related to cattle movements to assess the robustness of road and cattle-producing infrastructure for the northern states of Australia, which is prone to disruption due to environmental change. The three models presented are strategic and operational, and cover simulation and optimisation aspects in order to provide recommendations

[2]Like in many industries, these circumstances foster consolidation into a few major players.

on infrastructure repairs and new investments. Tsekeris (2014) used a database provided by the Greek government concerning all public investment projects funded by the European Commission. The database comprises road, railways, airports, seaports and urban public transport, and was used in an ambitious project to assess public expenditure inter-dependencies.

Some simulation studies are aimed specifically to understanding the disruptive effect of accidents and natural disasters. For example, Bruzzone et al. (2000) presented a simulation model of harbour and maritime environments with the aim of designing harbour and maritime infrastructures in order to determine the resources, structures and services needed to face possible emergencies. A case study of an oil spill was used to determine the amount of oil that could reach the coast of Genoa. Ferrario and Zio (2014) proposed an assessment framework to study the safety of a nuclear power plant in case of earthquake, and used Monte Carlo simulation to calculate the probability that the plant enters into an unsafe state. This model enable operators to determine the likelihood of certain parts of the plant recovering earlier than others.

2.2. Mathematical programming

Simulation can model large and complex systems, but has the disadvantage of being only a tool for investigating a system's behaviour, unable by itself to suggest 'good' solutions; simulation is a *descriptive* modelling approach. By contrast, Mathematical optimisation aims at finding the 'best' possible solution. This solution is the 'best' in the sense that it is the most profitable set of actions a decision maker can possibly take, among a very large number of possible combinations of actions: optimisation is a *prescriptive* modelling approach. Because the number of possible combinations of actions is so large, optimisation models tend to be not as complex and detailed as simulation models. Uncertainty can be incorporated in these models by using different methodologies, which have been widely applied to infrastructure value assessment. Some of these include

1. *Stochastic programming.* Aims at producing a solution that, although may not be as good as a deterministic, optimal solution of a problem for which all the parameters are known with certainty, is far from being the worst given the possible uncertain scenarios (Kall and Wallace, 1994).

2. *Approximate dynamic programming.* Based on the idea that decisions should be made using estimates of the value of the states to which an action can take us, in contrast to 'myopic' policies that depend only on what is known at every time step (Powell, 2009).

3. *Robust optimisation.* Uses a measure of 'robustness' in the face of uncertainty, which is represented as deterministic variability in the value of the parameters (Bertsimas and Sim, 2004, Gabrel et al., 2014).

Detailed discussion of mathematical programming and related methodologies is beyond the scope of this chapter, and the interested reader is referred to the included

references. We next review some of the applications of these techniques to infrastructure assessment projects. As in the case of simulation above, the following review covers industry, services, risk assessment and transportation, and is by no means exhaustive.

Regarding the assessment of logistics and transportation infrastructure, Burdett et al. (2015) considers the problem of earthwork planning, that is, the problem of strategically moving earth material from one place to another, a necessary task in any infrastructure project. Burdett et al. propose accurate mixed-integer program (MIP) strategic models for linear infrastructure projects that explicitly incorporates fuel consumption and terrain gradients. Mishalani and Koutsopoulos (2002) proposed a general methodology, based on dynamic programming, for modelling the spatial variation of causal variables (such as traffic, soil conditions, pavement design characteristics, weather) and the identification of regions of physical infrastructure that deteriorate uniformly over time. The methodology is useful to assign maintenance work to regions of similar deterioration. The model was validated satisfactorily using detailed data from three roadway facilities.

Regarding the design and operation of water networks, D'Ambrosio et al. (2015) presents a review on the use of mathematical programming techniques in fresh water supply and distribution. Among this paper's findings is that design, operation, containment detection, and water quality management are the main areas of application of mathematical optimisation. Optimal operation of water networks is intrinsically related to giving existing infrastructure the best possible use. Projects of this type are normally difficult to solve and require simplifications, such as aggregating in the time dimension or linear approximation of nonlinear functions. The interested reader should also refer to Martin et al. (2012), which is a review that is broader in scope than D'Ambrosio et al.'s: it considers mostly mathematical programming models, but multi-objective optimisation and optimal control models are reviewed as well. Applications to optimise the operation of fresh water in industry are also common, such as Arzate et al. (2012), who presented a methodology to perform sensitivity analysis on the costs of investment required to upgrade treatment plants in water networks of refineries.

Regarding energy and electricity networks, a review is provided by Froger et al. (2016) in the context of maintenance scheduling, both corrective and preventative. Froger et al. note that there are differences between the needs of networks in regulated and unregulated markets: deregulated markets often present conflicts of interest between generation and transmission companies, whereas regulated markets focus mainly on reliability and costs. Uncertainty deserves a special mention in this review, and includes papers that use stochastic programming and heuristics.

Many OR papers dedicated to disaster management are related to the maintenance of system flow in emergency situations. For example, Matisziw and Murray (2009) proposed a novel constraint structure for network flow optimisation, and tested in Ohio's road network. The aim of the problem they solved was to identify network facilities most vital to network flow, more specifically, by identifying nodes and arcs associated to worst impact to system flow, given restrictions on the number of facilities damaged.

Approximate solutions for optimisation problems that are too large and difficult can be obtained by using *heuristics*. These are briefly reviewed in the next subsection.

2.3. Heuristics and metaheuristics

Heuristics provide fast ways of solving problems approximately, but they do not provide any proof of optimality or give an estimation of the quality of a solution. Efficient heuristics for certain problems use information that is specific for that type of problem, so that the heuristic can take advantage of the mathematical structure of the problem's search space. This was proved in the *no free lunch* theorem (Wolpert and McReady, 1997). The implication for us as practitioners is that the knowledge about the class of problems that are better suited for the optimisation heuristic method of choice must also hold for the practical problem at hand we are trying to solve. Rothlauf (2011), distinguish three types of heuristic methods: heuristics (for construction or improvement of solutions), approximation algorithms and modern heuristics. Modern heuristics are general, problem-invariant and widely applicable search strategies, and are often called *metaheuristics*. some applications related to infrastructure validation are reviewed next.

Won et al. (2012) introduce a heuristic based on genetic algorithms and rule extraction to, through future dividend policy, determine an optimal portfolio of investments; dividend policy is understood as the decisions about the relative proportion of dividends out of earnings over time. The proposed algorithm was used to evolve rules that represent the policy. It refines the multiple rules extracted through rule-based algorithms from dividend data sets using a genetic algorithm. Benchmarking on test data sets shows better results that when using rule extraction algorithms alone. Closely related to infrastructure, Zanakis and Becerra-Fernandez (2005) focused on the use of data mining techniques to identify important factors associated with determining a country's competitiveness and the development of knowledge-based models to predict a country's competitiveness score. Many of the factors Zanakis and Becerra-Fernandez used to train neural networks and regression tree models were related to the countries' existing infrastructure, including telecommunications, water networks, and percentage of urban population.

Regarding transportation infrastructure, Chou (2009) presented a case-based reasoning expert prototype system to determine preliminary project and maintenance costs, using existing information from previous experience of pavement maintenance and construction. The system was used to assist decision makers in project screening and budget allocation, reusing existing project management information and reducing the impact of subsequent cost changes. In a similar vein, Deng et al. (2011) solved the problem of predicting passenger volume in Chinese highways to better estimate investment, management and maintenance decisions by using a combination of rough set theory and neural networks. Their combined rough set theory and neural network approach appeared to be more robust and stable than previously reported approaches. For risk assessment in port infrastructure, Mokhtari et al. (2012) present a decision support framework based in fuzzy set theory and

evidential reasoning. The framework defined a hierarchy of risks and was applied to three Iranian ports. The results were tested using sensitivity analysis and the authors report that the methodology is being adapted to other engineering applications.

Regarding urban water networks, Marlow et al. (2015) noted that water distribution infrastructure is undervalued because it is buried and out of sight. This poses an additional challenge on managing the pipe networks, and Marlow et al. proposed a rule-based expert system to provide suggestions about the technical and economic risks related to pipeline repair and maintenance. The paper discussed at length the trade-offs between renovation, replacement and rehabilitation of water distribution infrastructure.

2.4. Machine Learning

Machine learning and statistical data analytics techniques play an important role in the valuation process by extracting from data the critical information needed by a simulation or mathematical programming model. This may be just a single nominal value, or a full probability distribution that defines the uncertainty, in a model parameter that defines how the infrastructure may perform in the future. This includes two key problems: predicting the future demand for the infrastructure, and predicting the useful life left in the existing infrastructure.

The work of Li et al. (2014) developed an approach to failure prediction for pipes within underground water distribution networks. The approach was data driven and combined information regarding the pipe's age, diameter, depth, construction material, protective coating, internal water pressure, surrounding soil type, and others, with historical failure records to predict the probability of a pipe failing in the future. This is critical in determining when pipe renewal should occur and minimising the overall cost incurred by preventative and reactive maintenance.

Assessing damage in civil structures such as bridges is also an important application of machine learning. For instance the construction of bridges require significant upfront investment and any unplanned closures can cause significant disruption to the wider economy of the region. There are a number of approaches to detecting the early onset of damage to such structures that allow preventative maintenance to be scheduled in advance to avoid disruption at peak times. The work of Gul and Catbas (2009) provides an example of how anomaly detection and time series analysis techniques can be used to detect structural health. In Diez et al. (2016) similar unsupervised data driven techniques are being applied to detect the early onset of damage to the iconic Sydney harbour bridge.

Ports represent a critical piece of infrastructure for most economies. The ability to forecast port throughput enables stakeholders to make efficient decisions that not only covers infrastructure investments, but includes management of port development, operational restructuring and tariffs policy. There exist may studies that have developed forecasting methods and range from simple univariate techniques, for example auto-regressive integrated moving average Kim et al. (2011) and exponential smoothing Abraham and Ledolter (2009) models, to more complex multivariate

methods that model the interdependencies between a broader set of predictor variables such as socio-economic indicators, gross domestic products, commodity prices, etc. These methods include multivariable adaptive regression splines, dynamic factor models, vector autogreressive and auto-regressive integrated moving average with exogenous variables (Geng et al., 2015, Angelopoulos and Chlomoudis, 2015, Intihar et al., 2015). All of these methods are capable of extracting the models from data and capturing the uncertainty in how the demand will evolve in the future.

2.5. Hybrid and other approaches

Combinations of solution methodologies are common and some examples in the literature have already been mentioned; a review with a taxonomy of hybrid approaches can be found in Talbi (2013). In this subsection we briefly review *matheuristics* and other methodologies that have been used to validate the value of infrastructure.

2.5.1. Matheuristics

Matheuristics are a special type of hybrid optimisation approach where mathematical programming algorithms interact with metaheuristics. For an algorithm to be considered a matheuristic, there must be a point in the algorithm where the solution strategy takes advantage of the mathematical structure of the problem or sub-problem being solved. These are also known as "model-based heuristics". One example in the assessment of energy networks is Fischetti et al. (2015), who present a new matheuristic that combines MIP optimisation and a greedy algorithm to solve the smart grid energy management problem. The proposed application addresses the demand side energy management problem by solving a scheduling problem involving multiple appliances with different operational constraints, user preferences, renewable energy sources, and batteries. Matheuristics are currently a very active area of research.

2.5.2. Other approaches

Finally, some papers adopt other quantitative approaches to validate infrastructure. For example, Tsekeris (2014) presented an ambitious regression model to assess inter-dependencies among infrastructure investments, and assessed the effect of rail, road, air and sea transport investments on each other. Among other findings, the analysis demonstrated that increased relative growth share of maritime transport expenditure can stimulate the investment activity in other non-road transport expenditure categories. Also focusing on inter-dependencies, Ouyang (2016) presented a literature review of existing models for assessing the inter-dependencies of critical infrastructure systems, including empirical and agent-based approaches, which are beyond the scope of our discussion. Vandermeulen et al. (2011) used traditional cost-benefit analysis to assess the economic value of green infrastructure, which is normally not included in land use plans, but which has important social and environmental benefits.

Figure 2 shows some of the most relevant papers reviewed in this Section in the context of their focus and the techniques used.

3. CASE STUDY: STRATEGIC RAIL PLANNING

The following case study considers containerised freight transport to and from Sydney's Port Botany. Containers can be transported directly between the premises of the freight owners and the port with trucks via the road network. Alternatively, they can be transported via an intermediate intermodal rail terminal. The case study considers the Greater Sydney metropolitan region that extends out 50km from the port. Critical to this study is an assessment of whether the existing network of railway lines and intermodal terminals is adequate to support future growth. This section addresses this issue by combining detailed demand forecasting methods with rail scheduling techniques to assess capacities and determine bottlenecks.

The system comprises the following components

1. **Production Areas**. These areas contain producers that send containers with goods for export and also contain the end customers for imports. The areas aggregate the total supply of containers for export as well as demand of containers imported, to the suburb level. Customers within each suburb area determine their transport choice by selecting the service provider that best satisfies their price and service time requirements. The case study considers ten production areas.

2. **Inter-modal terminals (IMT)** An IMT is a transfer facility with road and rail access, and on-site warehousing. In the model presented here, each IMTs run its own rail assets (wagons/locomotives) and container handling assets (forklifts or cranes); in the real world, trains are operated by train companies that may or not own an IMT. For example, some logistics companies own trains and an IMT, but Cooks River IMT does not own any trains. Import containers change from trains to trucks at the IMT facilities on their way to the production areas, and export containers are placed into train consists from trucks on their way to the port terminals for shipping. The rate at which containers can be processed by the facility is determined by the number of handling assets, while the rate at which containers can be transported to and from the port is determined by the rail assets. Six IMTs are considered.

3. **Port terminals**. Terminals are export-shipping points and entry points for imports, and also have warehousing facilities available. Two terminals are included in the case study.

A simulation model of this system is described in Banerjee et al. (2016), with further details found in Chi Thai (2011). It is worth remarking here that, unlike the simulation model or real life operations, the optimisation model presented in this chapter seeks the greater benefit of all participants in the system. The model suggests to producers the amounts and routes to use to transport their containers

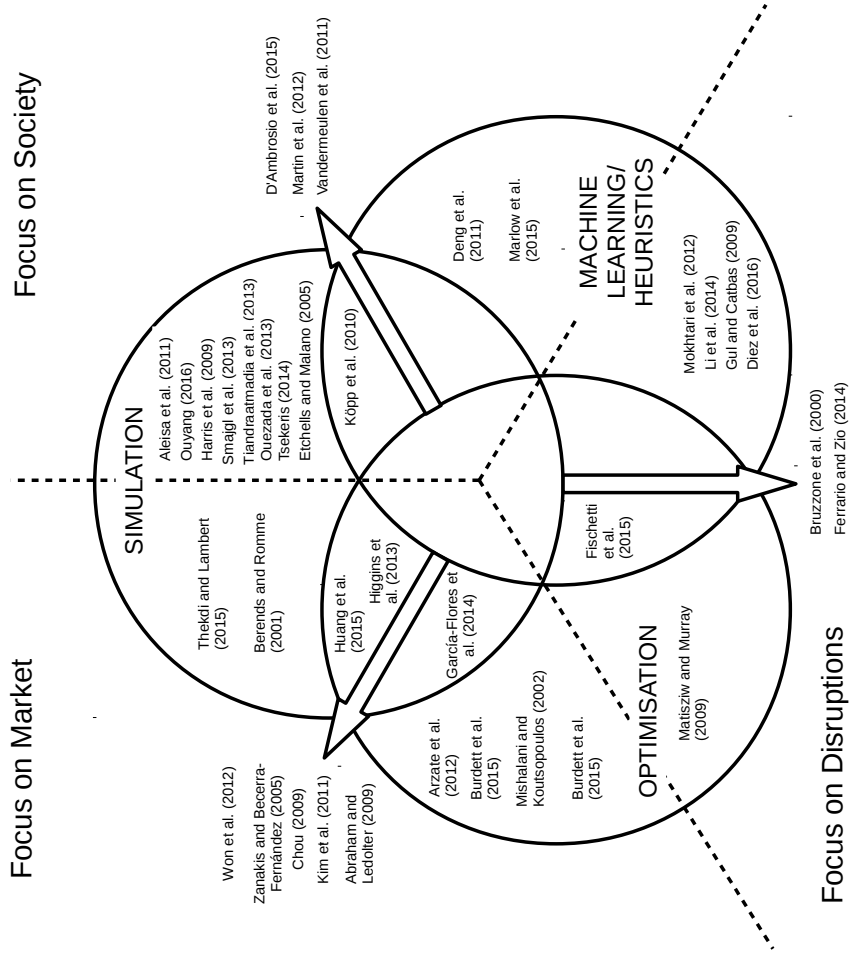

Figure 2. The (non-exhaustive) review of previous work organised by quantitative technique and application area. Simulation is widely used to assess the value of existing infrastructure. All techniques are normally used in all three application areas.

to maximise global benefit, and individual IMTs, who in real life compete with each other, have to follow these transportation plans, which may not strictly be reflected in reality.

A schematic of the system is depicted in Figure 3, and a map with the physical location of all the sites is shown in Figure 4. Producers have the option to send containers for export and receive imports directly from terminals by truck, or to use IMTs. If this is the case, containers are transferred from trains to trucks and vice versa in the IMTs for imports and exports, respectively. Storage facilities exist at terminals and warehouses.

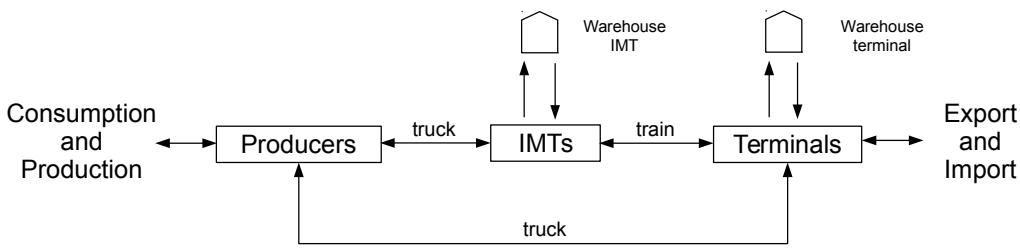

Figure 3. Schematic of the strategic rail tool. Containers are transported by truck between producers and IMTs and between producers and terminals. Transport by rail occurs only between IMTs and terminals. IMTs and terminals count with warehousing space to store containers.

The importance of understanding this system resides in the fact that, on one hand, IMTs reduce road congestion and exploit economies of scale by pooling demand from surrounding areas and using rail to transport containers to and from ports. Additionally, it may be in the interest of local government authorities to encourage rail transport in order to reduce congestion and extend the life of valuable road infrastructure. On the other hand, rail requires additional handling of containers (lift on and lift off), while direct transport from and to ports by truck may be more convenient and save the need for extra-handling in IMTs. In other words, there is a trade-off to be analysed in the attractiveness of truck versus rail that depends on a number of variables such as cost, total travel time, frequency of services, risk etc.

In Banerjee et al. (2016), we simulated this system with emphasis on IMTs by capturing costs, capacities and service times for different asset mixes, a demand forecast model, and competition from other offerings. By contrast, the present model, introduced in detail in Appendix A, assesses the value of existing infrastructure, namely the rail capacity, road usage and warehousing capacities at IMTs and terminals and provides a means for comparing quantitatively rail and road transportation modes by combining optimisation and the forecasting models developed in Banerjee et al. (2016). To be more precise, the aim of the present problem is to minimise the total costs of operating the entire supply chain, which comprise transportation costs, the costs of violating the soft inventory limits at warehouses, the costs of

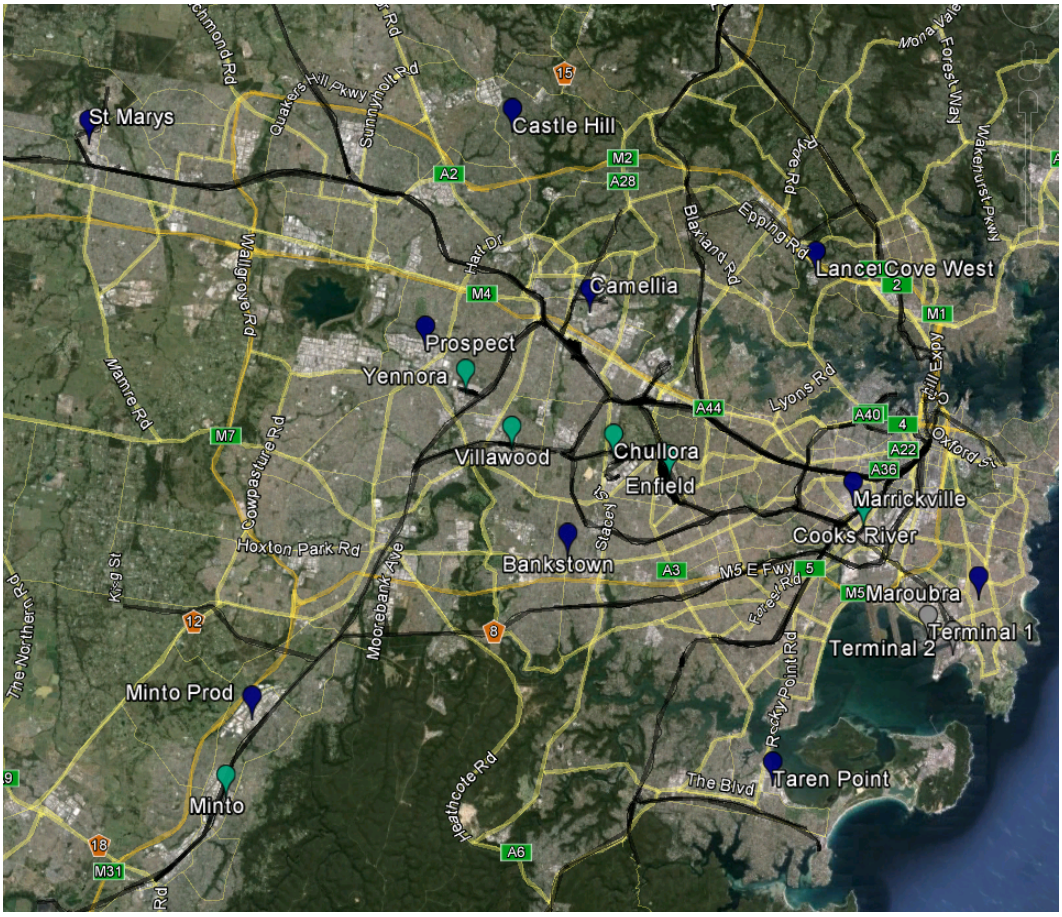

Figure 4. Location of all sites that form the case study. Producers are marked in blue, IMTs in green and terminals in gray. The model considers that all producers are connected to all terminals and to all IMTs by road (light green lines), and that all IMTs are connected to all terminals by rail (black lines); these links are not shown explicitly in the map.

moving containers to and from warehouses, and the costs of hiring extra trains and container-handling equipment. The model is useful to assess the value of infrastructure because the magnitude of all these costs indicates if the available resources are sufficient or not.

We should point out that the optimisation model presented in Appendix A seeks the greater benefit for all participants in the system, although this may not be strictly true for individual IMTs, who compete with each other and try to maximise their own gains. However, the model is still useful to assess economic trade-offs and infrastructure value and use, and can be adapted to assess trade-offs for individual IMTs.

3.1. Facility Capacities

The basic set of parameters used for sites that hold inventories (IMTs and terminals) is shown in Table 1. The model considers that there is a fleet of 50 trains available, each able to carry 32 wagons. In turn, each wagon can accommodate three TEUs slots, which is the same as three one-TEU containers, or two one-TEU container and one two-TEU container. The road transport cost is AUD $360.00 plus AUD $30.00 per additional kilometre, and the rail transport cost is fixed at AUD $200.00. These values reflect the market conditions discussed in Piyapatroomi et al. (2006), Chi Thai (2011).

Table 1. Parameters of IMTs and terminals in containers

Name	Type	Soft inventory limit	Hard inventory limit	Container handling capacity
Minto	IMT	4000	5000	2000
Yennora	IMT	10000	13000	3000
Chullora	IMT	5000	6000	2000
Enfield	IMT	4000	5000	2000
Cooks River	IMT	4000	5000	2000
Villawood	IMT	4000	5000	2000
Terminal 1	Terminal	80000	100000	4000
Terminal 2	Terminal	80000	100000	4000

3.2. Predicted Demand

The figures for supply (at producers/terminals) and consumption (at terminals/producers) of the exports/imports, respectively, are the main parameters we vary to assess the value of infrastructure. Figure 5 displays the production of export containers, specified in terms of a twenty-foot equivalent unit (TEU), for a base case scenario. This corresponds to a representative 60 day window. In addition to this base case, a more variable demand scenario is also considered. This scenario is displayed in Figure 6. Similar graphs for imports are not shown due to space limitations.

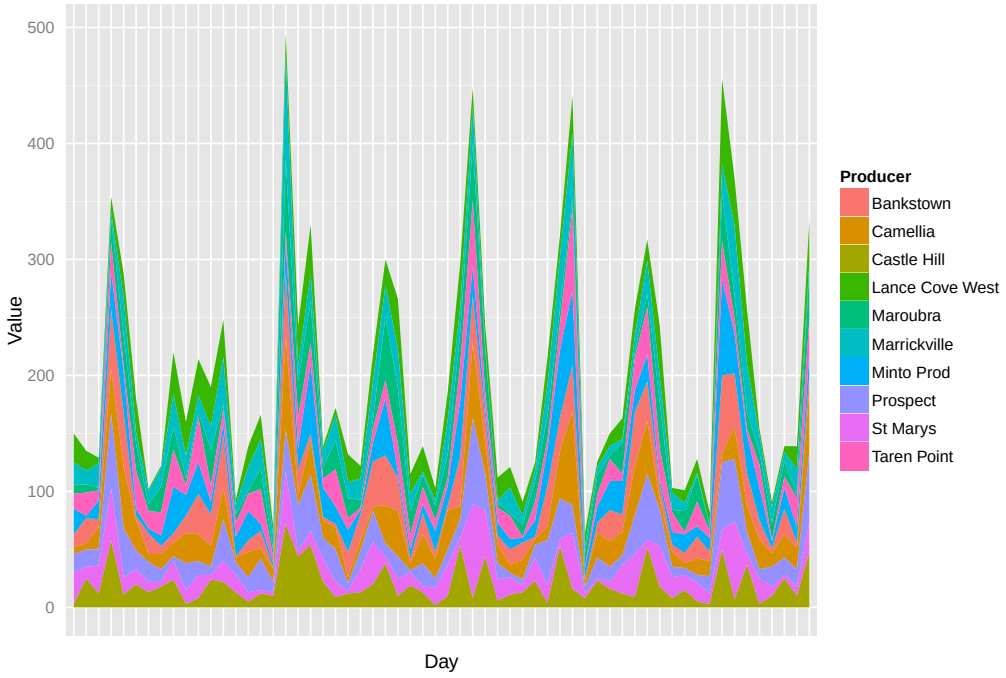

Figure 5. Base case scenario for production of one TEU containers for export by producer along the time horizon of 60 days.

The two cases have been determined by analysing the statistical properties of throughputs of forecasted import and exports. The base case figures were obtained from publicly available data sources and are approximately correct. The variable demand scenario was calculated using the same mean but with a higher standard deviation than the observed data, representing higher future volatility in the throughput volumes. This represents an extreme but plausible scenario to assess the value and effectiveness of existing infrastructure. The two scenarios we analyse in Section 4. are the following: *a*) effect of reducing the number of total available trains, *b*) comparing the base case scenario with a peaky demand scenario.

4. RESULTS

The problem was implemented in version 1.5 of the Clojure[3] language with an Excel interface. We obtained all the following results using version 12.4 of the CPLEX[4] optimiser in a 64-bit Intel Xeon CPU with two processors of eight cores (2.27 GHz) each and 48 GB of RAM. The problem has 32220 integer variables and 55352 constraints. A typical run takes around ten minutes to complete.

[3]http://clojure.org/, accessed on the 8 of April 2013.

[4]http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/index.html, accessed on the 16 of May 2016.

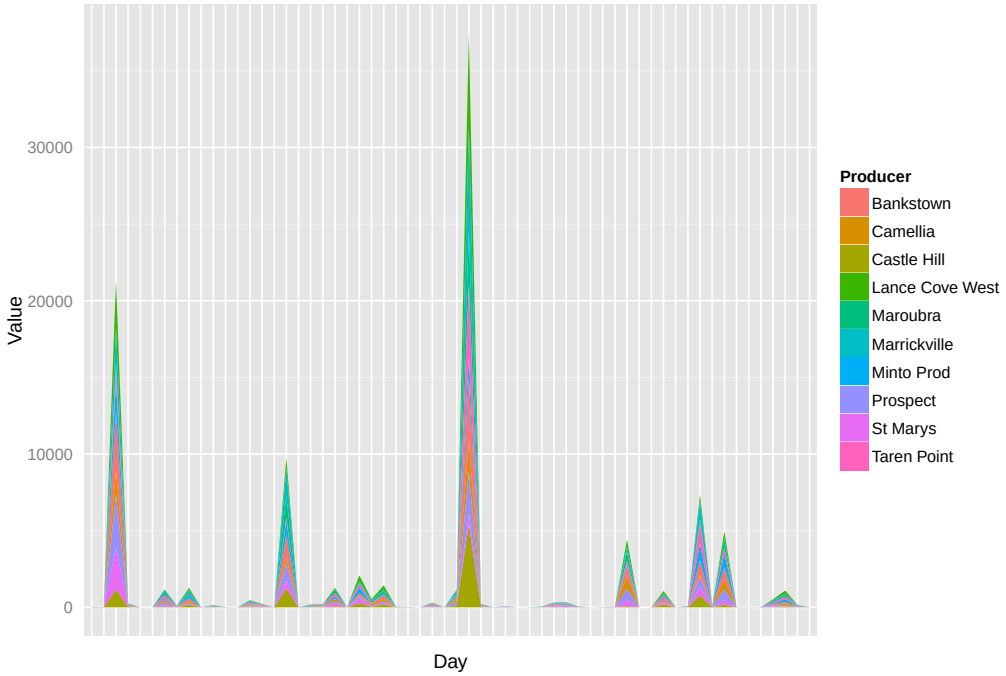

Figure 6. Variable scenario for production of one TEU containers for export by producer along the time horizon of 60 days.

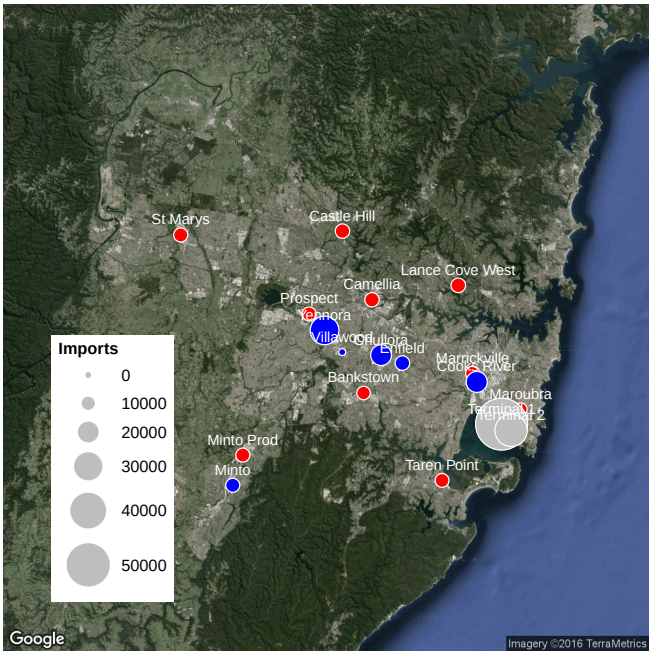

Figure 7. The total number of import containers handled by each site in the optimised bi-modal supply chain along the total time horizon. Producers are marked in red, IMTs in blue and terminals in gray. The size of the circle is proportional to the number of import containers handled. A similar map can be produced for exports (not shown).

For the base case, the flow of one and two TEU containers for import per site, as calculated by the optimiser, is shown on Figure 7. Similar maps can be produced for exports, but these are not shown due to space restrictions. The actual amounts of exports and imports and percentage delivered directly to and from each source, or via an IMT are shown in Tables 2 and 3. The optimal solution for the base case does not recommend transfer of containers in either direction at Villawood IMT, while Yennora takes the largest proportion of transported containers both for imports and exports, followed by Chullora and Cooks River. Most of the transport to terminals should be done by rail, as only Maroubra is close enough to send and receive containers exclusively by truck. Figure 8 shows the road infrastructure that actually get used at some moment of the 60-day time horizon. Increasing the cost of rail transport does little to modify this road layout, except for encouraging Taren Point to use roads directly to Terminal 2 instead of sending containers via Cooks River IMT (map not shown).

Table 4 shows the effect of reducing the size of the fleet, scenario a). It is clear that the system collapses below a fleet size of thirty trains, as the number of trains needed and the number of days when these extra trains are needed also increase substantially. The variable demand scenario also wreaks havoc in the system, with 204 extra trains needed distributed in 34 days. Figure 9 show the result of having

Table 2. Total number of export containers and percentage received at terminals from IMTs or directly from producers for base case.

	Terminal 1	Terminal 2	% Terminal 1	% Terminal 2
Minto	819	977	7.72	14.89
Yennora	3266	1884	30.77	28.71
Chullora	2084	1584	19.64	24.14
Enfield	1066	704	10.04	10.73
Cooks River	1835	1414	17.29	21.55
Villawood	0	0	0.00	0.00
Lance Cove West	0	0	0.00	0.00
Marrickville	0	0	0.00	0.00
Maroubra	1543	0	14.54	0.00
Taren Point	0	0	0.00	0.00
Minto Prod	0	0	0.00	0.00
Bankstown	0	0	0.00	0.00
Camellia	0	0	0.00	0.00
Prospect	0	0	0.00	0.00
St Marys	0	0	0.00	0.00
Castle Hill	0	0	0.00	0.00
TOTAL:	10613	6563	100.00	100.00

a more variable demand, scenario *b*). Under this scenario, some producers use multiple IMTs at different times to avoid the storage and handling bottlenecks that are present. For example, in the map we can observe that Camellia uses Yennora, Villawood and Chullora IMTs at different times during the 60-day horizon. Selecting these transport scheduling decisions is not trivial and requires a quantitative analysis tool like the model we presented. This analysis shows that the existing infrastructure can handle the two scenarios considered, provided that adequate decision support is available to the participants.

5. CONCLUDING REMARKS

When large infrastructure projects are built, they are not normally amenable to modification after completion, but instead they must operate continuously for many years in the future, possibly with only a few refurbishments during the course of their lifetimes. However, they must still face the constraints and bottlenecks produced by market changes and other disruptions, most of which may not have been foreseen originally and which can have unanticipated repercussions. Fortunately, managers and operators can compensate for miscalculations by playing with operational plans.

In this chapter, we adopted the point of view of an operator who must make the most out of the existing infrastructure through optimal operation plans and accurate estimates of parameter data. The combination of OR and ML is an appropriate way to assess the value of existing infrastructure: if, on one hand, by

Figure 8. The optimised bi-modal supply chain of the base case. In this map, producers are marked with a blue marker, IMTs are indicated with orange, terminals with green, road links are shown in orange and rail links in black.

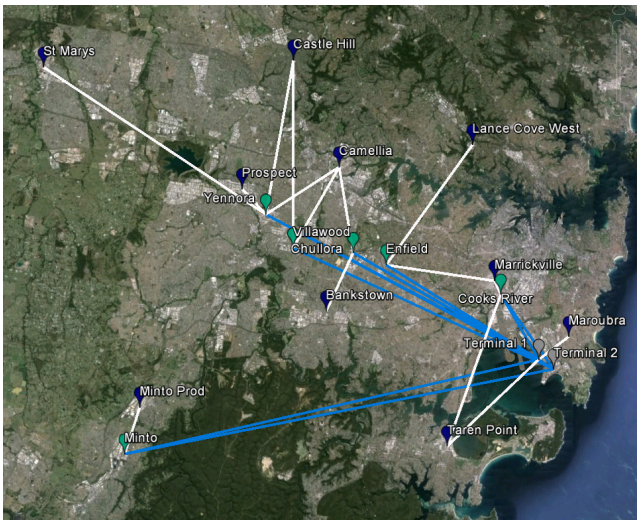

Figure 9. The optimised bi-modal supply chain of a variable demand and supply case scenario. In this map, producers are marked with a blue marker, IMTs are indicated with light blue, terminals with gray, road links are shown in white and rail links in light blue.

Table 3. Total number of import containers and percentage sent from terminals to IMTs or directly to producers for base case

	Terminal 1	Terminal 2	% Terminal 1	% Terminal 2
Minto	5018	4489	8.37	13.29
Yennora	15308	13643	25.54	40.39
Chullora	11014	7488	18.37	22.17
Enfield	6256	3882	10.44	11.49
Cooks River	14531	4274	24.24	12.65
Villawood	0	0	0.00	0.00
Lance Cove West	0	0	0.00	0.00
Marrickville	0	0	0.00	0.00
Maroubra	7818	0	13.04	0.00
Taren Point	0	0	0.00	0.00
Minto Prod	0	0	0.00	0.00
Bankstown	0	0	0.00	0.00
Camellia	0	0	0.00	0.00
Prospect	0	0	0.00	0.00
St Marys	0	0	0.00	0.00
Castle Hill	0	0	0.00	0.00
TOTAL:	59945	33776	100.00	100.00

using the best possible operation schedule calculated with plausible scenarios, the existing infrastructure and resources can cope with demand and avoid bottlenecks, the infrastructure is valuable and adequate. If, on the other hand, only uncertain forecasts and poor infrastructure are available, the calculation of a good schedule will provide insights of where resources should be invested in future.

Our literature review shows that simulation, optimisation and machine learning are widely used, alone or in combination, to study the economic and social value of existing infrastructure, and to analyse the impact of disruptions. The application areas span services (water and electricity networks), industry (mostly capital intensive, such as paper, oil and energy), and sudden, unforeseen events (e.g., climate change, technological innovations and disaster management). We also presented a case study case study based on the Greater Sydney rail system and studied a scenario with reduced number of trains, and a scenario with variable supply and demand. The results show that Villawood IMT may be an unnecessary facility for the base case from a global point of view, and that Yennora, Chullora and Cooks River take the largest proportion of transported containers both for imports and exports. We also demonstrated that a fleet size of less than 30 trains would struggle to cope with the volume of containers that need to be transported.

The quantitative model presented in Appendix A combines OR and ML, formalises operation rules in the supply chain, and facilitates comparison of scenarios. However, we should note that this model seeks the benefit of all participants in the system, although this may not strictly be in the interest of individual IMTs, who actually compete with each other. In any case, the model is still useful to assess

Table 4. Effect of fleet size. Between 20 and 30 trains the fleet cannot cope with demand. The cost of additional trains is set to AUD $ 100,000 per train.

Fleet size	Additional trains	Cost of additional trains	No. of days when trains are needed
50	2	200000.00	1
40	8	800000.00	2
30	6	600000.00	4
20	184	18400000.00	28

economic trade-offs and infrastructure value and use, and can be modified to analyse the situation of individual participants in the suppy chain.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Drs Thomas Vitsounis and Yuriy Tyshetskiy for their valuable input, and Ms Blandine Vacher for her contribution in coding the problem.

A MATHEMATICAL MODEL

A1. Nomenclature

Sets

$\mathcal{K}^{\texttt{EXP}}$ set of commodities (containers) that are exported

$\mathcal{K}^{\texttt{IMP}}$ set of commodities (containers) that are imported

$\mathcal{K}$ set of commodities (containers)

$\mathcal{L}_P^{\texttt{EXP}}$ set of valid links where containers for export are transported by rail (i.e., from IMTs to terminals)

$\mathcal{L}_R^{\texttt{EXP}}$ set of valid links where containers for export are transported by truck (i.e., from producers to IMTs and from producers to terminals)

$\mathcal{L}_P^{\texttt{IMP}}$ set of valid links where containers for import are transported by rail (i.e., from terminals to IMTs)

$\mathcal{L}_R^{\texttt{IMP}}$ set of valid links where containers for import are transported by truck (i.e., from IMTs to producers and from terminals to producers)

$\mathcal{L}_P$ set of valid links where containers are transported by rail (i.e., from IMTs to terminals and vice versa)

$\mathcal{L}_R$ set of valid links where containers are transported by truck (i.e., from producers to IMTs and vice versa, and from producers to terminals and vice versa)

$\mathcal{P}$ set of terminals

$\mathcal{R}$ set of IMTs

$\mathcal{S}$ set of sources

$\mathcal{T}$ total number of planning periods (weeks) in the model

Decision variables

$\alpha_{ikt}^{\downarrow}$ number of containers of type $k \in \mathcal{K}$ short of the preferred lower inventory level at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$\alpha_{ikt}^{\uparrow}$ number of containers of type $k \in \mathcal{K}$ in excess of the preferred upper inventory level at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

δ_{it} number of additional container-handling equipment (cranes or forklifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

γ_t number of additional trains needed at time $t \in \mathcal{T}$

n_{ijt} number of trains running in rail track segment $(i,j) \in \mathcal{L}_P$ at time $t \in \mathcal{T}$

u^+_{ikt} number of containers of type $k \in \mathcal{K}$ that go into the warehouse of IMT $i \in \mathcal{R}$ at time $t \in \mathcal{T}$

u^-_{ikt} number of containers of type $k \in \mathcal{K}$ that go out of the warehouse of IMT $i \in \mathcal{R}$ at time $t \in \mathcal{T}$

v^+_{ikt} number of containers of type $k \in \mathcal{K}$ that go into the warehouse of terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

v^-_{ikt} number of containers of type $k \in \mathcal{K}$ that go out of the warehouse of terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

w_{ikt} number of containers of type $k \in \mathcal{K}$ that are stored in the warehouse of site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$x^{\texttt{EXP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\texttt{EXP}}$ that travel by truck through link $(i,j) \in \mathcal{L}^{\texttt{EXP}}_R$ at time $t \in \mathcal{T}$

$x^{\texttt{IMP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\texttt{IMP}}$ that travel by truck through link $(i,j) \in \mathcal{L}^{\texttt{IMP}}_R$ at time $t \in \mathcal{T}$

$y^{\texttt{EXP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\texttt{EXP}}$ that travel by rail through link $(i,j) \in \mathcal{L}^{\texttt{EXP}}_P$ at time $t \in \mathcal{T}$

$y^{\texttt{IMP}}_{ijkt}$ number of containers of type $k \in \mathcal{K}^{\texttt{IMP}}$ that travel by rail through link $(i,j) \in \mathcal{L}^{\texttt{IMP}}_P$ at time $t \in \mathcal{T}$

$z^{\texttt{EXP}}_{ikt}$ number of containers of type $k \in \mathcal{K}$ that go to export from terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

$z^{\texttt{IMP}}_{ikt}$ number of containers of type $k \in \mathcal{K}$ that come as import to producer $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

Parameters

CC_i cost of additional container-handling equipment (cranes or forklifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

CM_{ikt} cost of moving containers at site $i \in \mathcal{R} \cup \mathcal{P}$ of commodity $k \in \mathcal{K}$ at time $t \in \mathcal{T}$

$CONS_{ikt}$ number of containers of type $k \in \mathcal{K}$ that are imported to producer $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

EXP_{ikt} number of containers of type $k \in \mathcal{K}$ that are exported from terminal $i \in \mathcal{P}$ at time $t \in \mathcal{T}$

FC_t cost of setting up additional trains at time $t \in \mathcal{T}$

$MAXNT$ maximum number of trains available at any moment

N number of wagons in a train (normally 32)

NS_k number of slots taken by container of type $k \in \mathcal{K}$

$PROD_{ikt}$ number of containers of type $k \in \mathcal{K}$ produced at $i \in \mathcal{S}$ at time $t \in \mathcal{T}$

PTC_{ijkt} cost of transporting commodity $k \in \mathcal{K}$ at time $t \in \mathcal{T}$ from site $i \in \mathcal{R} \cup \mathcal{P}$ to site $j \in \mathcal{R} \cup \mathcal{P}, i \neq j$ by rail

Q_{ij} maximum capacity of road segment $(i, j) \in \mathcal{L}_R$

RTC_{ijkt} cost of transporting commodity $k \in \mathcal{K}$ at time $t \in \mathcal{T}$ from site $i \in \mathcal{S} \cup \mathcal{R}$ to site $j \in \mathcal{S} \cup \mathcal{P}, i \neq j$ by truck

SP_{kt} selling price of container of type $k \in \mathcal{K}$ at time $t \in \mathcal{T}$

SVC_{it} cost of violating the soft inventory limits of site $i \in \mathcal{R} \cup \mathcal{P}$ at time $t \in \mathcal{T}$

$W_i^{\mathtt{max}}$ maximum number of containers that cranes (or forklifts) at site $i \in \mathcal{R} \cup \mathcal{P}$ can move at any time

$WDES_i^{\mathtt{max}}$ desirable upper limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

$WDES_i^{\mathtt{min}}$ desirable lower limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

$WLIM_i$ maximum limit on storage capacity at site $i \in \mathcal{R} \cup \mathcal{P}$ at any time

Sub-indexes

i, j site

k commodity (i.e., type of container)

p terminal

t planning period

A2. Formulation

The aim of the problem is to minimise the total costs of operating the entire supply chain, which comprise transportation costs, the costs of violating the soft inventory limits at warehouses, the costs of moving containers to and from warehouses, and the costs of hiring extra trains and container-handling equipment if the available resources are not sufficient.

Let us define the decision variables $x_{ijkt}^{\mathtt{EXP}}$ and $y_{ijkt}^{\mathtt{EXP}}$ as the containers for export that travel between sites i and j carrying commodity k at period t by truck and

by rail, respectively; $x^{\texttt{IMP}}_{ijkt}$ and $y^{\texttt{IMP}}_{ijkt}$ as the containers for import that travel between sites i and j carrying commodity k at period t by truck and by rail, respectively; $z^{\texttt{EXP}}_{ikt}$ and $z^{\texttt{IMP}}_{jkt}$ the containers that go to export and that come as import at terminal i, respectively; u^{+}_{ikt} and u^{-}_{ikt} the containers that are moved into and out of the IMTs' warehouses, respectively; v^{+}_{ikt} and v^{-}_{ikt} the containers that are moved into and out of the terminals' warehouses, respectively; w_{ikt} and w_{ikt} the containers that are stored in the IMT and terminal warehouses, respectively; $\alpha^{\uparrow}_{ikt}$ and $\alpha^{\downarrow}_{ikt}$ the amount by which the minimum and maximum desired inventory levels at sites with warehouses (i.e., IMTs and terminals) are violated, respectively; γ_t the number of additional trains needed at time t and δ_{it} the number of additional container-handling equipment at IMT i at time t. We can now define the objective function as

$$\begin{aligned}
\text{Minimise} \quad & \sum_{(i,j)\in\mathcal{L}^{\texttt{EXP}}_R}\sum_{k\in\mathcal{K}^{\texttt{EXP}}}\sum_{t\in\mathcal{T}} RTC_{ijkt}\; x^{\texttt{EXP}}_{ijkt} + \sum_{(i,j)\in\mathcal{L}^{\texttt{IMP}}_R}\sum_{k\in\mathcal{K}^{\texttt{IMP}}}\sum_{t\in\mathcal{T}} RTC_{ijkt}\; x^{\texttt{IMP}}_{ijkt} + \\
& \sum_{(i,j)\in\mathcal{L}^{\texttt{EXP}}_P}\sum_{k\in\mathcal{K}^{\texttt{EXP}}}\sum_{t\in\mathcal{T}} PTC_{ijkt}\; y^{\texttt{EXP}}_{ijkt} + \sum_{(i,j)\in\mathcal{L}^{\texttt{IMP}}_P}\sum_{k\in\mathcal{K}^{\texttt{IMP}}}\sum_{t\in\mathcal{T}} PTC_{ijkt}\; y^{\texttt{IMP}}_{ijkt} + \\
& \sum_{i\in\mathcal{R}\cup\mathcal{P}}\sum_{t\in\mathcal{T}} SVC_{it}\sum_{k\in\mathcal{K}}(\alpha^{\uparrow}_{ikt}+\alpha^{\downarrow}_{ikt}) + \\
& \sum_{i\in\mathcal{R}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} CM_{ikt}\left[u^{+}_{ikt}+u^{-}_{ikt}+\sum_{j\in\mathcal{S}}(x^{\texttt{EXP}}_{jikt}+x^{\texttt{IMP}}_{ijkt})+\sum_{j\in\mathcal{P}}(y^{\texttt{EXP}}_{ijkt}+y^{\texttt{IMP}}_{jikt})\right] + \\
& \sum_{i\in\mathcal{P}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} CM_{ikt}\left[v^{+}_{ikt}+v^{-}_{ikt}+\sum_{j\in\mathcal{S}}(x^{\texttt{EXP}}_{jikt}+x^{\texttt{IMP}}_{ijkt})+\sum_{j\in\mathcal{R}}(y^{\texttt{EXP}}_{jikt}+y^{\texttt{IMP}}_{ijkt})+z^{\texttt{EXP}}_{ikt}+z^{\texttt{IMP}}_{ikt}\right] + \\
& \sum_{t\in\mathcal{T}} FC_t\gamma_t + \sum_{i\in\mathcal{R}} CC_i\sum_{t}\delta_{it}\ . \qquad (1)
\end{aligned}$$

The problem is subject to the following constraints:

1. *Soft inventory capacity constraints.* There is a penalty for exceeding the desirable limits to the amount of containers that may be stored in the warehouses. For both IMTs and terminals:

$$WDES^{\texttt{min}}_i - \sum_{k\in\mathcal{K}}\alpha^{\downarrow}_{ikt} \le \sum_{k\in\mathcal{K}} w_{ikt} \le WDES^{\texttt{max}}_i + \sum_{k\in\mathcal{K}}\alpha^{\uparrow}_{ikt} \qquad \forall i\in\mathcal{R}\cup\mathcal{P} \quad \forall t\in\mathcal{T}\ . \qquad (2)$$

2. *Hard inventory capacity constraints.* There is a maximum number of containers that can be stored at sites with warehousing capability:

$$WDES^{\texttt{max}}_i + \sum_{k\in\mathcal{K}}\alpha^{\uparrow}_{ikt} \le WLIM_i \qquad \forall i\in\mathcal{R}\cup\mathcal{P} \quad \forall t\in\mathcal{T}\ . \qquad (3)$$

3. *Crane capacity constraints.* There is a limit on how many containers the cranes (or forklifts) can move at any time period, and a penalty if the available container moving resources are not sufficient. For IMTs, containers must be

moved in and out of the warehouses, as well as to trucks or trains if they are going to be brought into and sent out of the facility at the current planning period:

$$\sum_{k\in\mathcal{K}}(u^{+}_{ikt}+u^{-}_{ikt}+\sum_{j\in\mathcal{S}}x^{\texttt{EXP}}_{jikt}+\sum_{j\in\mathcal{S}}x^{\texttt{IMP}}_{ijkt} + \sum_{j\in\mathcal{P}}y^{\texttt{EXP}}_{ijkt}+\sum_{j\in\mathcal{P}}y^{\texttt{IMP}}_{jikt})-\delta_{it}\leq W^{\texttt{max}}_{i} \quad \forall i\in\mathcal{R},\ \forall t\in\mathcal{T}\ . \tag{4}$$

For terminals, in addition to all the containers moved into and out of the warehouses and the train and truck loading and unloading operations, we need to consider that the containers for export must be loaded into the ships, and containers for import must be unloaded from the ships:

$$\sum_{k\in\mathcal{K}}(v^{+}_{ikt}+v^{-}_{ikt}+\sum_{j\in\mathcal{S}}x^{\texttt{EXP}}_{jikt}+\sum_{j\in\mathcal{S}}x^{\texttt{IMP}}_{ijkt}+\sum_{j\in\mathcal{R}}y^{\texttt{EXP}}_{jikt}+\sum_{j\in\mathcal{R}}y^{\texttt{IMP}}_{ijkt}+ z^{\texttt{EXP}}_{ikt}+z^{\texttt{IMP}}_{ikt})-\delta_{it}\leq W^{\texttt{max}}_{i} \quad \forall i\in\mathcal{P},\ \forall t\in\mathcal{T}\ , \tag{5}$$

where δ_{it} is the number of additional container-handling equipment needed at site i at time t and

$$z^{\texttt{EXP}}_{ikt}=EXP_{ikt} \qquad \forall i\in\mathcal{P},\ \forall k\in\mathcal{K}^{\texttt{EXP}},\ \forall t\in\mathcal{T}\ , \tag{6}$$

$$z^{\texttt{IMP}}_{ikt}=IMP_{ikt} \qquad \forall i\in\mathcal{P},\ \forall k\in\mathcal{K}^{\texttt{IMP}},\ \forall t\in\mathcal{T}\ . \tag{7}$$

4. *Conservation constraints at IMTs.* The expressions that define the flow of containers at IMTs and their warehouses are, respectively,

$$\sum_{j\in\mathcal{S}}x^{\texttt{EXP}}_{jikt}+\sum_{j\in\mathcal{P}}y^{\texttt{IMP}}_{jikt}+u^{-}_{ikt}=\sum_{j'\in\mathcal{S}}x^{\texttt{IMP}}_{ij'kt}+\sum_{j'\in\mathcal{P}}y^{\texttt{EXP}}_{ij'kt}+u^{+}_{ikt} \qquad \forall i\in\mathcal{R},\ \forall k\in\mathcal{K},\ \forall t\in\mathcal{T},\ \text{and} \tag{8}$$

$$w_{ik,t+1}=w_{ikt}+u^{+}_{ikt}-u^{-}_{ikt} \qquad \forall i\in\mathcal{R},\ \forall k\in\mathcal{K},\ \forall t\in\mathcal{T}\ . \tag{9}$$

We consider that the initial inventory level at all warehouses is zero.

5. *Conservation constraints at terminals.* The expressions that define the flow of containers at terminals and their warehouses are, respectively,

$$\sum_{j\in\mathcal{S}}x^{\texttt{EXP}}_{jikt}+\sum_{j\in\mathcal{R}}y^{\texttt{EXP}}_{jikt}+v^{-}_{ikt}+z^{\texttt{IMP}}_{ikt}=\sum_{j'\in\mathcal{S}}x^{\texttt{IMP}}_{ij'kt}+\sum_{j'\in\mathcal{R}}y^{\texttt{IMP}}_{ij'kt}+v^{+}_{ikt}+z^{\texttt{EXP}}_{ikt} \qquad \forall i\in\mathcal{P},\ \forall k\in\mathcal{K},\ \forall t\in\mathcal{T},\ \text{and} \tag{10}$$

$$w_{ik,t+1}=w_{ikt}+v^{+}_{ikt}-v^{-}_{ikt} \qquad \forall i\in\mathcal{P},\ \forall k\in\mathcal{K},\ \forall t\in\mathcal{T}\ . \tag{11}$$

6. *Conservation constraints at producers.* It is assumed that the producers do not hold inventory and that all containers produced must enter the supply chain. The expressions that define the flow of containers in these sites are

$$\sum_{j\in\mathcal{R}\cup\mathcal{P}}x^{\texttt{IMP}}_{jikt}=CONS_{ikt} \qquad \forall i\in\mathcal{S},\ \forall k\in\mathcal{K}^{\texttt{IMP}},\ \forall t\in\mathcal{T}\ , \tag{12}$$

$$\sum_{j\in\mathcal{R}\cup\mathcal{P}}x^{\texttt{EXP}}_{ijkt}=PROD_{ikt} \qquad \forall i\in\mathcal{S},\ \forall k\in\mathcal{K}^{\texttt{EXP}},\ \forall t\in\mathcal{T}\ , \tag{13}$$

where $PROD_{ikt}$ and $CONS_{ikt}$ are the amounts of produced and consumed containers of export and import commodities, respectively, in producer i.

7. *Constraints for consist assembly.* The space for containers in the trains is limited. We know that wagons have room for three slots; a slot is more properly known as a twenty-foot equivalent unit (TEU). We also know that containers can be 40 foot and occupy two TEUs, or 20-foot and occupy one TEU. For the time being, the model only distinguishes between containers for export or import. Thus, we have four commodities, $k = \{\texttt{20ft-EXP}, \texttt{40ft-EXP}, \texttt{20ft-IMP}, \texttt{40ft-IMP}\}$. Let n be the number of wagons in a consist (normally 32), n_{ijt} the number of trains that travel between sites i and j, and γ_t the number of additional trains needed in excess of the total available, $MAXNT$. To accommodate containers into trains, we need to define additional constraints. First, knapsack-like constraints so that the number of containers transported by rail in a road segment is not more than the capacity of all the trains travelling in that road segment at any given period:

$$\sum_{k \in \mathcal{K}^{\texttt{IMP}}} NS_k \; y^{\texttt{IMP}}_{ijkt} - 3 \; N \; n_{ijt} \leq 0 \qquad \forall (i,j) \in \mathcal{L}_P^{\texttt{IMP}}, \; \forall t \in \mathcal{T} \; , \text{ and} \tag{14}$$

$$\sum_{k \in \mathcal{K}^{\texttt{EXP}}} NS_k \; y^{\texttt{EXP}}_{ijkt} - 3 \; N \; n_{ijt} \leq 0 \qquad \forall (i,j) \in \mathcal{L}_P^{\texttt{EXP}}, \; \forall t \in \mathcal{T} \; , \tag{15}$$

where NS_k is the number of slots a container of type k takes. The total number of additional trains needed is

$$\sum_{(i,j) \in \mathcal{L}_P} n_{ijt} - \gamma_t \leq MAXNT \qquad \forall t \in \mathcal{T} \; . \tag{16}$$

The model assumes that every truck can only carry one container.

8. *Road capacity constraints.* We assume that roads have a fixed capacity:

$$\sum_{k \in \mathcal{K}} x^{\texttt{EXP}}_{ijkt} \leq Q_{ij} \qquad \forall (i,j) \in \mathcal{L}_R^{\texttt{EXP}}, \; \forall t \in \mathcal{T} \; , \tag{17}$$

$$\sum_{k \in \mathcal{K}} x^{\texttt{IMP}}_{ijkt} \leq Q_{ij} \qquad \forall (i,j) \in \mathcal{L}_R^{\texttt{IMP}}, \; \forall t \in \mathcal{T} \; , \tag{18}$$

where Q_{ij} is the capacity of the road segment (i,j).

9. *Upper bounds.* Finally, the upper bounds of the amounts transferred to and from warehouses are:

$$u^+_{ikt} \leq W_i^{\texttt{max}} \qquad \forall i \in \mathcal{R}, \; \forall k \in \mathcal{K}, \; \forall t \in \mathcal{T} \; , \tag{19}$$

$$u^-_{ikt} \leq W_i^{\texttt{max}} \qquad \forall i \in \mathcal{R}, \; \forall k \in \mathcal{K}, \; \forall t \in \mathcal{T} \; , \tag{20}$$

$$v^+_{ikt} \leq W_i^{\texttt{max}} \qquad \forall i \in \mathcal{P}, \; \forall k \in \mathcal{K}, \; \forall t \in \mathcal{T} \; , \tag{21}$$

$$v^-_{ikt} \leq W_i^{\texttt{max}} \qquad \forall i \in \mathcal{P}, \; \forall k \in \mathcal{K}, \; \forall t \in \mathcal{T} \; . \tag{22}$$

References

Abraham, B. and J. Ledolter (2009). Statistical methods for forecasting. In *Statistical methods for forecasting*. John Wiley and Sons.

Aleisa, A., O. Alkassar, A. Al-Jadi, S. Al-Sabah, and R. Hishmi (2011, December). A simulation-based evaluation for a strategic residential wastewater master plan. In S. Jain, R. Creasey, J. Himmelspach, K. White, and M. Fu (Eds.), *Proceedings of the 2011 Winter Simulation Conference*, Winter Simulation Conference Proceedings, Phoenix, AZ, pp. 953–962. IEEE.

Angelopoulos, J. and K. Chlomoudis (2015, August). Towards a port services indicator- the case of us ports. In *Proceedings of the Annual conference of the International Association of Maritime Economists (IAME)*, Hamburg, Germany.

Arzate, E., P. Huitzil, A. González, and B. Martínez (2012, June). Automated optimization model to perform sensitivity analysis on cost of investment required to upgrade treatment plants in water networks. In I. Bogle and M. Fairweather (Eds.), *Proceedings of the 22 European Symposium on Computer Aided Process Engineering (ESCAPE)*, Volume 30 of *Computer-Aided Chemical Engineering*, University College London, London, UK, pp. 1063–1067. Elsevier Science BV.

Banerjee, S., G. Mathews, C. Aracena, B. Thorne, Y. Tyshetskiy, and T. Vitsounis (2016, October). Computationally simulating intermodal terminal attractiveness and demand. In *Proceedings of the 23 Intelligent Transport Systems World Congress*, Melbourne, Australia. Accepted.

Berends, P. and A. Romme (2001, December). Cyclicality of capital-intensive industries: a system dynamics simulation study of the paper industry. *Omega 29*(6), 543–552.

Bertsimas, D. and M. Sim (2004, January–February). The price of robustness. *Operations Research 52*(1), 35–53.

Bruzzone, A., R. Mosca, R. Revetria, and S. Rapallo (2000, Jun–Aug). Risk analysis in harbor environments using simulation. *Safety Science 35*(1–3), 75–86.

Burdett, R., E. Kozan, and R. Kenley (2015). Block models for improved earthwork allocation planning in linear infrastructure construction. *Engineering Optimization 47*(3), 347–369.

Chi Thai (2011). Metropolitan Intermodal Terminal Study 2011. Technical report, Shipping Australia Limited.

Chou, J. (2009, March). Web-based cbr system applied to early cost budgeting for pavement maintenance project. *Expert Systems with Applications 36*(2), 2947–2960.

D'Ambrosio, C., A. Lodi, S. Wiese, and C. Bragalli (2015, June). Mathematical programming techniques in water network optimization. *European Journal of Operational Research 243*(3), 774–778.

Deng, W., W. Li, and X. Yang (2011, April). A novel hybrid optimization algorithm of computational intelligence techniques for highway passenger volume prediction. *Expert Systems with Applications 38*(4), 4198–4205.

Diez, A., N. L. D. Khoa, M. Makki Alamdari, Y. Wang, F. Chen, and P. Runcie (2016). A clustering approach for structural health monitoring on bridges. *Journal of Civil Structural Health Monitoring*, 1–17.

Etchells, T. and H. Malano (2005, December). Identifying uncertainty in water allocation modelling. In A. Zerger and R. Argent (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM05)*, Melbourne, Australia, pp. 2484–2490. Modelling and Simulation Society of Australia and New Zealand: University of Western Australia, Nedlands, WA, Australia.

Ferrario, E. and E. Zio (2014, May). Assessing nuclear power plant safety and recovery from earthquakes using a system-of-systems approach. *Reliability Engineering and System Safety 125*, 103–116.

Fischetti, M., G. Sartor, and A. Zanette (2015, January). MIP-and-refine matheuristic for smart grid energy management. *International Transactions in Operations Research 22*(1), 49–59.

Froger, A., M. Gendrau, J. Mendoza, and E. Pinson (2016, June). Maintenance scheduling in the electricity industry: A literature review. *European Journal of Operational Research 251*(3), 695–706.

Gabrel, V., C. Murat, and A. Thiele (2014, June). Recent advances in robust optimization: An overview. *European Journal of Operational Research 235*(3), 471–483.

Galindo, G. and R. Batta (2013, October). Review of recent developments in or/ms research in disaster operations management. *European Journal of Operational Research 230*(2), 201–211.

García-Flores, R., A. Higgins, D. Prestwidge, and S. McFallan (2014, March). Optimal location of spelling yards for the northern australian beef supply chain. *Computers and Electronics in Agriculture 102*, 134–145.

Geng, J., M. Li, Z. Dong, and Y. Liao (2015, June). Port throughput forecasting by mars-rsvr with chaotic simulated annealing particle swarm optimization algorithm. *Neurocomputing 147*, 239–250.

Gul, M. and F. N. Catbas (2009). Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications. *Mechanical Systems and Signal Processing 23*(7), 2192 – 2204.

Harris, J., J. Antenucci, P. Okely, R. Mills, and P. Yeates (2009, July). Three-dimensional hydrodynamic modelling of a coastal embayment for multiple use industrial discharge and infrastructure development. In R. Anderssen, R. Braddock, and L. Newham (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM09)*, Cairns, Australia, pp. 4163–4169. Modelling and Simulation Society of Australia and New Zealand: University of Western Australia, Nedlands, WA, Australia.

Higgins, A., I. Watson, C. Chilcott, M. Zhou, R. García-Flores, S. Eady, S. McFallan, D. Prestwidge, and L. Laredo (2013). A framework for optimising capital investment and operations in livestock logistics. *The Rangelands Journal 35*(2), 181–191.

Huang, Y., S. Mao, and R. Nelms (2015, September). Smooth scheduling for electricity distribution in the smart grid. *IEEE Systems Journal 9*(3), 966–977.

Intihar, M., T. Krnberger, and D. Dragan (2015, August). The relationship between the economic indicators and the accuracy of container port throughput forecasting. In *Proceedings of the Annual conference of the International Association of Maritime Economists (IAME)*, Hamburg, Germany.

Kall, P. and S. Wallace (1994). *Stochastic programming.* New York: John Wiley and Sons.

Kim, J., J. Pak, Y. Wang, S. Park, and G. Yeo (2011, June). Study on forecasting container volume of port using sd and arima. *Journal of Korean navigation and port research 35*, 343–349.

Köpp, C., H. Mettenheim, M. Klages, and M. Breitner (2010, September). Analysis of electrical load balancing by simulation and neural network forecast. In B. Hu, K. Morasch, S. Pickl, and M. Siegle (Eds.), *Operations Research Proceedings 2010*, Operations Research Proceedings, Univ Bundeswehr Munchen, Neubiberg, Germany, pp. 519–524. German Operations Research Society (GOR): Springer-Verlag Berlin.

Li, Z., B. Zhang, Y. Wang, F. Chen, R. Taib, V. Whiffin, and Y. Wang (2014). Water pipe condition assessment: a hierarchical beta process approach for sparse incident data. *Machine Learning 95*(1), 11–26.

Marlow, D., S. Gould, and B. Lane (2015, December). An expert system for assessing the technical and economic risk of pipe rehabilitation options. *Expert Systems with Applications 42*(22), 8658–8668.

Martin, A., K. Klamroth, J. Lang, G. Leugering, A. Morsi, M. Oberlack, M. Ostrowski, and R. Rosen (Eds.) (2012). *Mathematical optimization of water networks*, Volume 162 of *International Series of Numerical Mathematics.* Basel: Springer.

Matisziw, T. and A. Murray (2009, January). Modeling s-t path availability to support disaster vulnerability assessment of network infrastructure. *Computers and Operations research 36*(1), 16–26.

Mishalani, R. and H. Koutsopoulos (2002, February). Modeling the spatial behavior of infrastructure condition. *Transportation Research Part B 36*(2), 171–194.

Mokhtari, K., J. Ren, C. Roberts, and J. Weng (2012, April). Decision support framework for risk management on sea ports and terminals using fuzzy set theory and evidential reasoning approach. *Expert Systems with Applications 39*(5), 5087–5103.

Ouyang, M. (2016, January). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering and System Safety 121*, 43–60.

Piyapatroomi, N., J. M. Bunker, and L. Ferreira (2006). Freight intermodal terminal systems for port of Brisbane, Melbourne and Sydney. Technical report.

Powell, W. (2009, February). What you should know about approximate dynamic programming. *Naval Research Logistics 56*(3), 239–249.

Quezada, G., G. Grozev, S. Seo, and C. Wang (2014, April). The challenge of adapting centralised electricity systems: peak demand and maladaptation in South East Queensland, Australia. *Regional Environmental Change 14*(2), 463–473.

Rothlauf, F. (2011). *Design of Modern Heuristics: Principles and application.* Natural Computing Series. Springer Berlin Heidelberg.

Smajgl, A., J. Ward, and S. Egan (2013, December). Validating simulations of development outcomes in the Mekong region. In J. Piantadosi, R. Anderssen, and J. Bowland (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM13)*, Adelaide, Australia, pp. 2096–2102. Modelling and Simulation Society of Australia and New Zealand: Modelling and Simulation Society of Australia and New Zealand Inc, MSSANZ, Christchurch, New Zealand.

Snyder, L., Z. Atan, P. Peng, Y. Rong, A. Schmitt, and B. Sinsoysal (2016, February). Or/ms models for supply chain disruptions: a review. *IIE Transactions 48*(2), 89–109.

Talbi, E. (2013, July). Combining metaheuristics with mathematical programming, constraint programming and machine learning. *4OR-A Quarterly Journal of Operations Research 11*(2), 101–150.

Thekdi, S. and J. Lambert (2015, June). Integrated risk management of safety and development on transportation corridors. *Reliability Engineering and System Safety 138*, 1–12.

Tjandraatmadja, G., D. Kirono, L. Neumann, S. Larson, S. Stone-Jovicich, R. Barkey, A. Amran, and M. Selintung (2013, December). Assessing urban water security and climate change adaptation in makassar, indonesia. In J. Piantadosi, R. Anderssen, and J. Bowland (Eds.), *Proceedings of the International Congress on Modelling and Simulation (MODSIM13)*, Adelaide, Australia, pp. 2235–2241. Modelling and Simulation Society of Australia and New Zealand: Modelling and Simulation Society of Australia and New Zealand Inc, MSSANZ, Christchurch, New Zealand.

Trück, S., W. Bradford, A. Henderson-Sellers, S. Matthew, J. Scott, M. Street, and R. Taplin (2010). *Climate alert: Climate change monitoring and strategy*, Chapter Assessing climate change adaptation options for local governments, pp. 362–399. Sydney, NSW, Australia: Sydney University Press.

Tsekeris, T. (2014, December). Multi-sectoral interdependencies of regional public infrastructure investments. *Socio-Economic Planning Sciences 48*(4), 263–272.

Vandermeulen, V., A. Verspecht, B. Vermeire, G. Huylenbroeck, and X. Gellynck (2011, November). The use of economic valuation to create public support for green infrastructure investments in urban areas. *Landscape and Urban Planning 103*(2), 198–206.

Wolpert, D. and W. McReady (1997, April). No free lunch theorems for optimisation. *IEEE Transactions on Evolutionary Computation 1*(1), 67–82.

Won, C., J. Kim, and J. Bae (2012, December). Using genetic algorithm based knowledge refinement model for dividend policy forecasting. *Expert Systems with Applications 39*(18), 13472–13479.

Zanakis, S. and I. Becerra-Fernandez (2005, October). Competitiveness of nations: A knowledge discovery examination. *European Journal of Operational Research 166*(1), 185–211.