Shipping Data Generation for the Hunter Valley Coal Chain*

Natashia Boland, Martin Savelsbergh, and Hamish Waterer
School of Mathematical and Physical Sciences, University of Newcastle, Australia

Abstract

Strategic capacity planning is a core activity for the Hunter Valley Coal Chain Coordinator as demand for coal is expected to double in the next decade. Optimization and simulation models are used to suggest and evaluate infrastructure expansions and operating policy changes. These models require input data in the form of shipping stems, which are arrival streams of ships at the port, together with their cargo types and composition. Creating shipping stems that accurately represent future demand scenarios has been a time-consuming and daunting challenge. We describe a multi-phase optimization-based framework that facilitates and enhances this process, and which has become an integral part of the work flow. The framework embeds sampling to allow for the generation of multiple shipping stems for a single demand scenario, employs targets, and desirable and permissable ranges to specify and control the characteristics of the shipping stems, and uses integer programming in a hierarchical fashion to generate a shipping stem that best meets the set goals.

1 Introduction

The Hunter Valley Coal Chain (HVCC) refers to the inland portion of the coal export supply chain in the Hunter Valley, New South Wales, Australia. The HVCC essentially follows the path of the Hunter River traveling south-east from the mining areas in the Hunter Valley to Newcastle. The Port of Newcastle is the world's largest coal export port. In 2008 port throughput was around 92 million tonnes, or more than 10 percent of the world's total trade in coal.

Most of the coal mines in the Hunter Valley are open pit mines. The coal is mined and stored either at a railway siding located at the mine or at a coal loading facility (used by several mines). The coal is then transported to one of the terminals at the Port of Newcastle, almost exclusively by rail. (Some coal is transported to the port by truck.) The coal is dumped and stacked at a terminal to form stockpiles. Coal brands are a blended product, with coal from different mines having different characteristics "mixed" in a stockpile to meet the specifications of the customer. Blends, and hence stockpiles, are most often make-to-order. However, the use of dedicated stockpiles for high volume products is increasing. Once the vessel for which the coal is destined arrives at a berth at the terminal, the coal is loaded onto the vessel. The vessel then transports the coal to its destination. The structure of the inland portion of a coal export chain is illustrated in Figure 1.

In 2003, the Hunter Valley Coal Chain Logistics Team (HVCCLT) was established to improve the movement of coal from Hunter Valley mines to the port's coal loaders and then to markets across the globe. HVCCLT pools the resources of port operators, railway operators, and railway infrastructure managers into one logistics team. In 2009, when the HVCC went through a major

*This research was supported by the ARC Linkage Grant no. LP0990739.

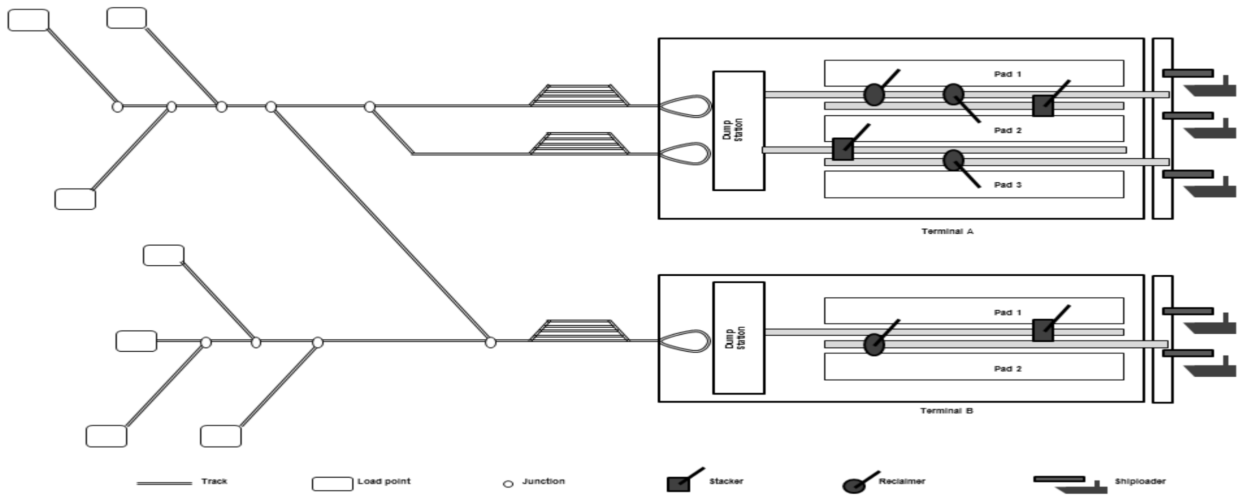

Figure 1: Schematic illustrating the inland portion of a coal export chain.

restructuring, the Hunter Valley Coal Chain Coordinator Limited (HVCCC) was incorporated as a new legal entity and formally replaced the HVCCLT (See `http://www.hvccc.com.au` and Vandervoort [20] for more information on the HVCCC). The HVCCC's mission is to plan and coordinate the cooperative daily operation and long term capacity alignment of the HVCC. Its strategic objectives include, among others:

- To plan and schedule the movement of coal through the HVCC in accordance with the agreed collective needs and contractual obligations of producers and service providers;
- To ensure minimum logistics cost and maximum throughput through the provision of appropriate analysis and advice on capacity constraints (whether physical, operational or commercial) affecting the efficient operation of the HVCC; and
- To advocate positions to other stakeholders and governments on issues relevant to efficient operation, in order to maximize opportunities for improved coordination and/or further expansion of the coal chain.

Accommodating future growth is one of the most pressing challenges, if not the most pressing challenge, facing the HVCC. Demand for coal is expected to more than double in the next decade. Increasing the annual throughput will be accomplished in two ways: (1) by expanding the infrastructure, and (2) by improving the operational efficiency. The HVCCC uses both opimization and simulations models in their planning efforts. The optimization model suggests a set of infrastructure expansions to accommodate a particular future demand scenario (see Singh et al. [19] for more details), and the simulation model analyzes the throughput for a given infrastructure, a given set of operational procedures, and a particular future demand scenario (see Welgama and Oyston [21] for more details). The typical analysis involves a future demand scenario that covers a period of a year. To be able to assess the achievable annual throughput with confidence, the HVCC is modeled at a fairly detailed level, i.e., at a daily operational level.

This implies that yearly mine production forecasts (provided by the producers) need to be converted to daily demands on the HVCC. As the HVCC can be viewed as operating as a pull-system, in which the nomination of a vessel, i.e., the announcement of the vessel's pending arrival at the port two to three weeks in advance, triggers the production, transportation, storage, and ultimately the loading of coal onto the vessel, the mine production forecast needs to be mapped into a stream of vessel arrivals at the port. Such a stream of vessel arrivals is referred to as a shipping stem. Each vessel arrival, referred to as a *trip*, is characterized by an arrival time, the terminal at which the vessel is to be loaded, a cargo-profile, which specifies the various brands of coal and their tonnage that constitute the vessel's cargo, the associated brand-recipes, which specify the various coal components, and thus the mines, that make up a brand (or blend) and their tonnage.

The importance of generating shipping stems that are representative of what the future may bring cannot be overemphasized. Furthermore, generating such shipping stems is not simply a matter of scaling a historical stem, or "more of the same". The shipping stem has to reflect that the HVCC is changing; new mines are brought on line, existing mines are (temporarily) shut down, new brands and new brand-recipes are introduced, new terminals start their operations, and demand management will change the arrival patterns of vessels.

Shipping stems are currently produced manually, which is extremely time-consuming; creating a single shipping stem can take up to three weeks. We have developed a multi-phase approach for generating shipping stems that relies on integer programming and sampling and produces a shipping stem in a matter of hours. It allows for the generation of multiple shipping stems for the same forecast mine production. The generated shipping stems have been validated by the HVCCC and our shipping stem generator is now an integral part of their analysis framework.

The challenges associated with generating appropriate data sets for computational studies are, of course, well-known and recognized by the optimization community, but relatively few articles focus exclusively on issues surrounding data generation. An exception is the recent article by Hall and Posner [10] which succinctly captures the importance of data generation: "In many experiments, the methods chosen to generate synthetic data can significantly affect the results of an experiment." A different, but equally important perspective is provided by Reilly [18] who emphasizes the importance of understanding and properly inducing correlations between characteristics of data, which is one of the major drivers and, at the same time, one of the major challenges in our stem generation research.

The impact of ship arrival patterns on the productivity of a port has been acknowledged widely and port infrastructure designs and port operating procedures should ideally be tailored to the anticipated ship arrival patterns. Van Asperen et al. [1] study the impact of three different ship arrival patterns on port efficiency. The first assumes Poisson interarrival times and represents an "uncontrolled" arrival process. The second assumes equidistant arrivals, in which ships arrive perfectly smoothly during the year, at equally spaced intervals. The third assumes scheduled arrivals, in which ships arrival times are coordinated with activities at the port. The latter two represent "controlled" arrival processes. Small random perturbations are introduced in the controlled arrival process to add more realism. Van Asperen et al. simulate port operations using a priority queueing model and assess its performance according to several metrics, including queue length and ship waiting time. Their results show that the ship arrival pattern has a significant impact on these metrics: the performance is far better for the controlled arrival processes than for the uncontrolled arrival process. A clear indication that careful modeling of the ship arrival process is crucial for the validity of any conclusions drawn from computational studies involving ship arrival patterns. It also points to the potential benefits of demand management strategies. Lang and Veenstra [13] demonstrate the potential impact of demand management strategies for container port operations by studying the use of port congestion information to alter a vessel's speed in approaching the port.

Laih, Lin and Chen [14] investigate port queue pricing as a mechanism for manipulating container ship arrival patterns. We note that this stream of research focuses purely on arrival times of ships and does not consider the varying nature of demand on the ships, which is a crucial element in the stem generation research reported on in this paper.

Fioroni et al. [8] discuss the use of simulation studies in the design of the infrastructure of an iron ore port and provide convincing and persuasive arguments that the generation of appropriate data is essential. Noteworthy is the fact that they observe that product-mix requirements significantly complicate matters and that their model is adequate in their environment, which has to distinguish only two products, but would not be appropriate in richer product environments. The HVCC provides such an environment with more than 80 coal brands blended from coal from over 30 mines.

The discussion above highlights the fact that generating appropriate shipping stems for use by the HVCCC in the analysis of the benefits of potential infrastructure investments worth hundreds of millions of dollars is both extremely challenging and extremely important.

Our work offers contributions in a number of areas. Firstly, at the highest level, we introduce a successful, optimization-based approach to a complex forecasting problem that cannot be handled effectively with traditional forecasting techniques. Traditional forecasting methods rely on the assumption that the future will resemble the past. Unfortunately, the environment in the HVCC is changing so rapidly that that is no longer a realistic assumption. The demand for coal in the next decades will not resemble the demand for coal in the past decades. In the resources boom that Australia is experiencing, demand for coal is increasing at rates that have never been seen before. And it is not just the volume of coal that is increasing, the number of coal mines operating in the Hunter Valley is increasing, with a commensurate change in coal blends. Furthermore, the supporting infrastructure is changing, with new terminals starting operations and demand management changing the arrival pattern and cargo-profiles of vessels. Secondly, the need to be able to generate multiple shipping stems with similar characteristics is accommodated by using a two-phase approach in which likely trip profiles are created in the first phase, and complete trip details are determined for a sampled set of trips in the second phase. This combination of optimization and sampling to obtain possible futures is, to our knowledge, quite noel. Thirdly, an innovative approach to generating vessel arrival patterns that satisfy desired smoothness characteristics in terms of weekly variations and spacing during the year has been developed. Controlling or influencing the smoothness of demand is at the heart of many supply chain optimization techniques. Being able to generate demand patterns with specific smoothness characteristics enables study of a system's behavior under different circumstances and assessment of the potential value of such optimization techniques. Finally, the design of a simple, consistent, easy-to-use mechanism for planners to control the shipping stem characteristics, which is based on importance hierarchies, and targets and desirable and permissable ranges, and to express their knowledge (or expectations) of future environment characteristics, which is based on exploratory trip profiles is of great value to the HVCCC.

Most importantly, however, our stem generation tool allows the planning team at the HVCCC to continue to do what it is charged to do. The scenarios that the planning team is currently investigating, with a throughput of more than 250 Mtpa, can no longer be generated manually. The complexity of preparing a shipping stem exhibiting all the desired characteristics for that level of demand has become too much for the human brain to accomplish, even when given weeks to do so. It has become too difficult to keep track of the intricate relationships between the various characteristics of a shipping stem and to anticipate the consequences of a change to one characteristic of the shipping stem on its other characteristics. The manual construction process simply did not converge anymore. The only way forward is the use of optimization-based stem generation technology; without the results of the research, the HVCCC would not be able to do what it is charged to do.

Our work is part of a growing body of research on various aspects of mining operations. For an overview of that research see Newman et al. [17] and for an example of the impact of that research see Epstein et al. [7].

The remainder of the paper is organized as follows. In Section 2, we introduce relevant terminology and discuss the results of an analysis of historic shipping stems. In Section 3, we motivate and present a fully automated approach for shipping stem generation. In Section 4, we review the results of an extensive computational study. In Section 5, we comment on the implementation of the technology and the impact the technology has on the HVCCC. In Section 6, we conclude with some final remarks.

2 Terminology and Historical Data Analysis

In this section, we introduce terminology and discuss the findings of an analysis of historical shipping stems, which motivates our approach to shipping stem generation.

Coal *mines* are owned and operated by a number of *producers*. Each producer has individual agreements with various *customers* to supply coal. The vast majority of coal is supplied to fulfill long-term on-going contracts. Coal is transported to the customers using *vessels*. Each vessel belongs to a *vessel class*. A vessel berths at a coal *terminal*. An arriving vessel's *cargo-profile* specifies the quantities of one or more coal *brands* that need to be loaded. A brand consists of one or more coal *components* which are blended according to one of the brand's *recipes* in prescribed fractions. As the above suggests, a brand may have more than one recipe. This flexibility is essential for the producers as coal characteristics for a mine change as coal is extracted. Thus to be able satisfy customer specifications, which are typically part of long-term agreements, it may be necessary to change recipes. An example of a possible cargo-profile for a vessel is shown in Figure 2. A coal component is sourced from a unique mine and is railed by one or more *consists* (a consist

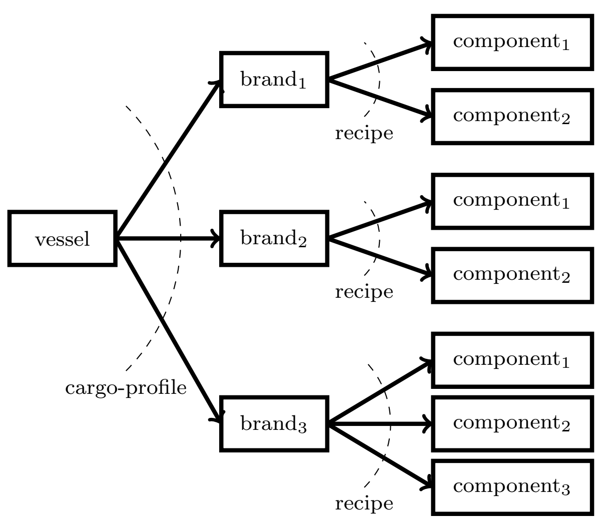

Figure 2: Example of a cargo-profile of a vessel

is the group of rail vehicles which make up a train) from a *load point* to a *stockyard* at a terminal where it is stockpiled. Each *stockpile* in the stockyard forms a single brand. The combination of a vessel, a terminal, a cargo-profile, and an arrival time is known as a *trip*. An example of a trip is given in Table 1. Each producer has individual agreements with the terminal *operators* to export

Table 1: An example of a trip

arrival time	11 October 2008 13:40
terminal id	3
vessel id	438
vessel class	Panamax
vessel tonnage	90,502
brand id	42
brand tonnage	15,086
component id	25
component tonnage	15,086
brand id	62
brand tonnage	36,433
component id	2
component tonnage	36,433
brand id	74
brand tonnage	38,983
component id	25
component tonnage	5,423
component id	37
component tonnage	33,560

a certain amount of coal through one or more of their terminals. An example of a set of possible relations between producers and terminal operators based on individual agreements is shown in Figure 3. In fact, the individual agreement between a producer and a terminal operator does not only specify the amount of coal that will exported through each of their terminals, it specifies the amount of coal of each brand that will be exported through each of their terminals.

Data from 2008 was analyzed to gain insight into the HVCC operations. Many statistics about the operations are readily available or can easily be derived, such as the total tonnage handled, the tonnage handled by each terminal, the tonnage produced by each of the mines, and the fraction of vessels that berthed at each of the terminals. These statistics provide high-level insight into the size and structure of the operations. As our technology has to provide a detailed shipping stem, gaining a low-level insight and understanding of the features and characteristics of cargo-profiles and brand recipes is as important, if not more important. The discussion that follows focuses on these aspects of the operations. In excess of 92 million tonnes of coal, sourced from 31 mines, and blended into 79 different brands, was loaded via two terminals aboard 492 unique vessels for a total of 1063 trips. There were 271 unique brand combinations observed on the trips. The information can be further categorized by vessel class. The 120 trips made by Handimax vessels loaded 139 brands totaling 5.6 million tonnes, the 662 trips made by Panamax vessels loaded 835 brands totaling 50.1 million tonnes, and the 281 trips made by Cape vessels loaded 577 brands totaling 36.7 million tonnes. The average number of brands on the 1063 trips was 1.46 with a minimum of 1 and a maximum of 6. An in-depth analysis reveals that the different brands are loaded fairly smoothly throughout the year in terms of frequency and tonnage, but that for many brands the recipes vary considerably in terms of the coal components that constitute the recipe and their relative fractions.

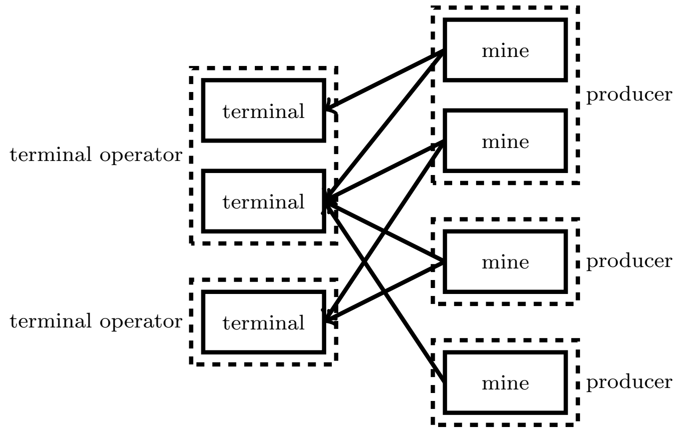

Figure 3: Example of a set of possible relations between producers and terminal operators

The fact that brands are loaded fairly smoothly throughout the year is not unexpected as mining companies tend to have long-term contracts with their customers. The variation in brand recipes is not unexpected either as it is a response to changes in the quality of coal extracted from mines and fluctuations in production and transportation capacity. Because the producers only provide mine production forecasts and do not provide detailed information about the cargo-profiles and brand recipes to be used in the future, the historical cargo-profiles and brand recipes form the main source of information for gaining insight into the combinations of brands that customers purchase and the recipes used to produce these brands.

3 Methodology

As observed above, historical cargo-profiles are the only source of information regarding the combinations of brands that customers purchase and regarding the recipes used to produce these brands. Since producers and customers typically sign long-term contractual agreements, it seems reasonable to assume that there will not be significant changes to the brand combinations typically seen in customers' orders. Furthermore, since brand recipes determine the chemical properties of the coal, it is also unlikely that brand recipes change significantly. Consequently, the key idea underlying our methodology for converting mine production forecasts into a shipping stem is to re-use the cargo-profiles and the brand-recipes seen in the historical trips, but to allow the component quantities to change (within pre-specified bounds).

A natural way to convert that idea into an integer programming formulation (see and Nemhauser and Wolsey [16] and Wolsey [22] for in-depth coverage of integer programming) is to introduce two sets of decision variables: a set of decision variables to determine how to adjust the recipes for the brands in the cargo-profile of a historical trip, i.e., to determine how much each mine contributes to a brand, and a set of decision variables to determine how often modified historical trips appear in the shipping stem so as to satisfy forecast mine productions. The latter, however, is not as straightforward as it may seem at first, since we want to allow the recipes of brands on a historical trip to be adjusted in different ways. To accommodate that flexibility, multiple copies of a historical trip need to be present in the model, and how to determine upfront how many copies of a particular trip are needed is not obvious. A pragmatic approach would be to simply have k copies of each historical trip, which allows the optimization to create k variations of the trip, i.e., k different sets of recipes for the brands in the cargo-profile of the trip, for some fixed number k. Unfortunately, that pragmatic approach leads to integer programs of enormous size that cannot be handled by state-of-the-art commercial integer programming solvers.

Therefore, we have decided to developed a multi-phase approach, where we first estimate for each historical trip the likelihood that it will be used to meet forecast mine productions, and then use these likelihoods to decide the number of copies of the historical trips to include in the integer program that ultimately determines the characteristics of the trips that make up the shipping stem. More specifically, the trip likelihoods are converted to probability distributions and sampling is used to select the trips to include in the integer program. The added benefit of this approach is that we are able to generate several different stems, but with similar characteristics, for the same mine production forecasts, which is valuable for the simulation studies that HVCCC wants to perform.

The overall approach is summarized below:

1. *Data Cleansing.* This module performs some basic validation and manipulation of the historical trip information to detect and correct any inconsistencies.

2. *Trip Assessment.* This module determines the likelihood that a particular historical trip will be used to meet the mine production forecasts.

3. *Shipping Stem Generation*

   a. *Trip Sampling.*

      This module samples historical trips based on the likelihood information determined in the trip assessment module. Trips are repeatedly sampled until the tonnage sourced from each mine exceeds some multiple of the mine's production forecast and the tonnage flowing through a terminal exceeds some multiple of the terminal's throughput target. (That is we over-sample to introduce some additional flexibility.)

   b. *Trip Generation.*

      This module determines which of the sampled trips will make up the shipping stem and the component tonnages that make up the cargo-profiles of each of these trips.

   c. *Trip Balancing and Timing.*

      This module distributes the trips over a set of time buckets, e.g., 52 weeks, so as to smooth the demand for coal sourced from mines. This in turn smooths vessel arrivals and load point activity over the planning period. Then arrival times are generated for the trips in the time buckets.

The *Data Cleansing* and *Trip Assessment* phases are executed only once for a given scenario, i.e., a set of mine production forecasts and a set of desired shipping stem characteristics, and the *Shipping Stem Generation* phase can be executed as often as desired.

As mentioned above, shipping stems are used primarily in the strategic capacity planning efforts at HVCCC. As these efforts typically involve a number of what-if analyses, there is a need to be able to control and direct the characteristics (or attributes) of the shipping stem generated. We provide this control by allowing a user to provide permissible and desirable ranges about a target value for each characteristic. The target value a^t of characteristic a has to lie within a desirable (soft) range $[\underline{a}^s, \overline{a}^s]$, which in turn has to fall inside a permissible (hard) range $[\underline{a}^h, \overline{a}^h]$, i.e., $\underline{a}^h \leq \underline{a}^s \leq a^t \leq \overline{a}^s \leq \overline{a}^h$. Our optimization guarantees that the value of a characteristic is within its permissible range and penalizes being outside the desirable range, thereby hoping to guide the value towards the desirable range.

To handle a desirable range $[\underline{a}^s, \overline{a}^s]$ for a characteristic a in an optimization model, we introduce an absolute deviation variable $\delta \geq 0$, introduce constraints $\delta \geq v - \overline{a}^s$ and $\delta \geq \underline{a}^s - v$, where v is the value of the characteristic, and minimize either the sum of the quadratic terms $(\delta/a^t)^2$ or the linear terms (δ/a^t). Handling a desirable range $[\underline{a}^s, \overline{a}^s]$ for a characteristic that is specified as a fraction or ratio of values v/w is more complicated. (An example of such a characteristic is the fraction of vessels of a certain vessel class berthing at a terminal.) We introduce an absolute deviation variable $\delta \geq 0$ and, as before, introduce constraints $\delta \geq v/w - \overline{a}^s$ and $\delta \geq \underline{a}^s - v/w$. However, since these constraints are nonlinear, we approximate them with linear constraints as follows. Let $\hat{w}$ be a constant that approximates the value of w, then we use constraints $\hat{w}\delta \geq v - \overline{a}^s w$ and $\hat{w}\delta \geq \underline{a}^s w - v$ and minimize either the sum of the quadratic terms $(\delta/a^t)^2$ or the linear terms (δ/a^t).

Permissible ranges simply restrict the set of feasible shipping stems. Desirable ranges, on the other hand, do not restrict the set of feasible shipping stems, but allow measurement of their attractiveness. Since there are several characteristics of interest, a mechanism needs to be introduced to determine the attractiveness of shipping stems when several characteristics are not within their desirable ranges. We have opted to do so through a ranking of the characteristics. The ranking is translated into a hierarchy of objectives, each associated with a characteristic, and solved in order. The hierarchy reflects a user's view on the importance of the different shipping stem characteristics. The optimal value of an objective is only allowed to deteriorate by a pre-specified percentage during an optimization at a lower level in the hierarchy. This approach is an example of lexicographic goal programming. (See Ehrgott [6], Barichard et al. [2], and Jones and Tamiz [12] for overviews of multicriteria optimization and goal programming.)

Next, before discussing each of these phases in our approach in more detail, we introduce notation.

The following sets and indices are used in the models:

- the set of mines M, mine $m \in M$,
- the set of terminal operators O, operator $o \in O$,
- the set of terminals L, terminal $l \in L$,
- the set of vessel classes V, vessel class $v \in V$,
- the set of brands B, brand $b \in B$,
- the set of historical trips K, and trip $k \in K$,

The set of trips may be partitioned in different ways depending on need. Thus, we will sometimes use K_l, the set of trips that berth at terminal $l \in L$, K_o the set of trips that berth at a terminal of operator $o \in O$, and K_v the set of trips that use a vessel in vessel class $v \in V$.

The following information concerning historical trips is used:

- the tonnage t_k (draught survey tonnage) of trip $k \in K$,
- the tonnage t_{mbk} of coal sourced from mine $m \in M$ as a component of brand $b \in B$ in the cargo-profile of trip $k \in K$,
- the vessel class $v_k \in V$ of the vessel that performed trip $k \in K$, and
- the terminal $l_k \in L$ at which trip $k \in K$ berthed.

The following information concerning future production is used:

- the forecast production tonnage F_m of mine $m \in M$, which implies that the forecast system-wide production F is $\sum_{m \in M} F_m$.

The following characteristics of a shipping stem can be controlled:

- T: the total tonnage flowing through the system (the target T^t is equal to the system-wide production forecast F);
- T_m: the total tonnage sourced from mine $m \in M$ (the target T_m^t is equal to the mine production forecast F_m);
- T_l: the total tonnage flowing through terminal $l \in L$;
- T_{mo}: the total tonnage sourced from mine $m \in M$ for operator $o \in O$;
- T_{bo}: the total tonnage of brand $b \in B$ for operator $o \in O$;
- f_{lv}: the fraction of trips berthing at terminal $l \in L$ performed by a vessel of class $v \in V$;
- t_l: the tonnage of a trip that berths at terminal $l \in L$;
- t_v: the tonnage of a trip performed by a vessel of class $v \in V$;
- t_{bv}: the tonnage of brand $b \in B$ of a trip performed by a vessel of class $v \in V$;
- t_{bl}: the tonnage of brand $b \in B$ of a trip that berths at terminal $l \in L$;
- t_m: the tonnage sourced from mine $m \in M$ for a brand of a trip;
- $t_{\overline{m}}$: the average tonnage sourced from mine $m \in M$ for a brand of a trip;

For each of these characteristics, we assume that a target, a desirable range, and a permissable range are specified, e.g., for system tonnage we assume we have values $\underline{T}^h, \underline{T}^s, T^t, \overline{T}^s$, and $\overline{T}^h$.

For convenience, we also introduce the following notation. For each operator $o \in O$, let $l(o)$ denote the set of terminals operated by the operator. For each trip $k \in K$, let $b(k)$ denote the set of brands in the cargo-profile and let $m(k)$ denote the set of mines from which coal was sourced. Furthermore, for each trip $k \in K$ and each brand $b \in b(k)$, let $m(b,k)$ denote the set of mines from which the components were sourced, and $m(k) = \bigcup_{b \in b(k)} m(b,k)$.

3.1 Data cleansing

This module performs some basic validation and manipulation of the historical trip information to detect and correct any inconsistencies.

The historical trip information is known to contain inaccuracies. The tonnage t_k (draught survey tonnage) is accurate, because it is the value measured to ensure that the vessel can sail on the tide. The component tonnages are less reliable. As a result, in almost all cases, the sum of component tonnages of a historical trip does not equal the trip tonnage, and brand tonnages and component tonnages are not always within their permissible ranges. Therefore, each historical trip $k \in K$ is "scaled" so as to ensure that the brand tonnages and the component tonnages sum up to the trip tonnage and that they are within their permissible ranges. The reasons for the permissible ranges vary. Take, for example, the range of the permissible tonnages that can be sourced from mine $m \in M$ for a particular brand on a trip, i.e., $[\underline{t}_m^h, \overline{t}_m^h]$. Coal is transported from a mine to a terminal using consists and the lower limit ensures a high utilization of the consists as it is set to the size of the smallest consist. At a mine, coal is loaded into the consists at a load point and the upper limit ensures that there is sufficient loading capacity at the load point. Next, consider the range of permissible tonnages of a brand in the cargo-profile of a trip performed by a vessel of class v, i.e., $[\underline{t}^h_{bv}, \bar{t}^h_{bv}]$. The reason for the minimum permissible brand tonnage is that a vessel consists of a number of cargo holds and a brand should at least take up one full cargo hold. The maximum permissible brand tonnage is typically set to the vessel capacity.

A hierarchical optimization model is used to scale a trip $k \in K$. Let α be the trip tonnage scaling factor and α_{mb} for $b \in b(k)$ and $m \in m(b,k)$ be the scaling factor for the component coming from mine m in brand b. The hierarchical objectives are (in order):

$$\min \sum_{b \in b(k)} \sum_{m \in m(b,k)} (\alpha - \alpha_{mb})^2$$

$$\min\ (1-\alpha)^2,$$

and the constraints are

$$\sum_{b \in b(k)} \sum_{m \in m(b,k)} t_{mbk}\alpha_{mb} = t_k \alpha \tag{1}$$

$$\max((1-\epsilon_t)t_k, \underline{t}^h_l, \underline{t}^h_{v_k}) \leq t_k\alpha \leq \min((1+\epsilon_t)t_k, \bar{t}^h_l, \bar{t}^h_{v_k}) \tag{2}$$

$$\max(\underline{t}^h_{bl}, \underline{t}^h_{bv_k}) \leq \sum_{m \in m(b,k)} t_{mbk}\alpha_{mb} \leq \min(\bar{t}^h_{bl}, \bar{t}^h_{bv_k}) \quad b \in b(k) \tag{3}$$

$$\underline{t}^h_m \leq t_{mbk}\alpha_{mb} \leq \bar{t}^h_m \quad b \in b(k), m \in m(b,k) \tag{4}$$

$$(f_{mbk} - \epsilon_f) \sum_{m' \in m(b,k)} t_{m'bk}\alpha_{m'b} \leq t_{mbk}\alpha_{mb} \leq (f_{mbk} + \epsilon_f) \sum_{m' \in m(b,k)} t_{m'bk}\alpha_{m'b} \quad b \in b(k), m \in m(b,k) \tag{5}$$

$$\alpha,\ \alpha_{mb} \geq 0 \quad b \in b(k), m \in m(b,k),$$

where f_{mbk} specifies the fraction, in terms of tonnage, that component m makes up of brand b in the original trip, and where tolerances ϵ_t and ϵ_f are specified by the user (default settings are 0.1 and 0.05, respectively).

Constraint (1) ensures that the scaled components and brand tonnages sum up to the scaled trip tonnage. Note that to ensure that the scaled component and brand tonnages are within their permissable ranges, it may be necessary to scale the trip tonnage as well. Constraint (2) ensures that if the trip tonnage is scaled, it remains within its permissable ranges, i.e., tonnage allowed at the terminal and tonnage allowed for the vessel class, and it does not deviate too much from its original value. Constraints (3) and (4) ensure that the scaled brand tonnages and the scaled component tonnages are within their respective permissable ranges. Constraints (5) ensure that any deviation from the original brand recipe is limited.

The first objective in the hierarchy focuses on scaling the component and brand tonnage values so that they are in their permissable ranges and sum up to the trip tonnage. The second objective in the hierarchy focuses on keeping the trip tonnage as close as possible to its original value.

Certain inconsistencies may cause trips to be infeasible: a trip contains component from a mine that does not have a contract with the terminal that the trip berths at, is of a vessel class that is not allowed at the terminal that the trip berths at. These inconsistencies are detected, an error is reported, and the trip is deleted from the set of historical trips.

3.2 Trip assessment

During trip assessment we establish the likelihood of using a historical trip to satisfy forecast mine production. Since the difference between the mine production forecast and the historical mine production is likely to be different for different mines, some historic trips may be more useful than others when it comes to satisfying forecast mine production.

The model has continuous variables $w_k \geq 0$ for all $k \in K$ denoting the likely number of required copies of trip k. These values will be scaled to produce likelihoods. The model has the following constraints to enforce the permissible ranges:

$$\underline{T}^h \leq \sum_{k \in K} t_k w_k \leq \overline{T}^h \tag{6}$$

$$\underline{T}_l^h \leq \sum_{k \in K_l} t_k w_k \leq \overline{T}_l^h, \qquad \forall l \in L \tag{7}$$

$$\underline{T}_m^h \leq \sum_{\substack{k \in K: \\ m \in m(k)}} \sum_{\substack{b \in b(k): \\ m \in m(b,k)}} t_{mbk} w_k \leq \overline{T}_m^h, \qquad \forall m \in M \tag{8}$$

$$\underline{T}_{bo}^h \leq \sum_{\substack{k \in K_o: \\ b \in b(k)}} \sum_{m \in m(b,k)} t_{mbk} w_k \leq \overline{T}_{bo}^h, \qquad \forall b \in B,\ \forall o \in O \tag{9}$$

$$\underline{T}_{mo}^h \leq \sum_{\substack{k \in K_o: \\ m \in m(k)}} \sum_{\substack{b \in b(k): \\ m \in m(b,k)}} t_{mbk} w_k \leq \overline{T}_{mo}^h, \qquad \forall m \in M,\ \forall o \in O \tag{10}$$

$$\lceil \frac{\underline{T}_m^h}{\overline{t}_m^h} \rceil \leq \sum_{\substack{k \in K: \\ m \in m(k)}} \sum_{\substack{b \in b(k): \\ m \in m(b,k)}} w_k \leq \lfloor \frac{\overline{T}_m^h}{\underline{t}_m^h} \rfloor, \qquad \forall m \in M \tag{11}$$

$$\lceil \frac{\underline{T}_l^h}{\overline{t}_m^h} \rceil \leq \sum_{\substack{k \in K_l: \\ m \in m(k)}} \sum_{\substack{b \in b(k): \\ m \in m(b,k)}} w_k \leq \lfloor \frac{\overline{T}_l^h}{\underline{t}_m^h} \rfloor, \qquad \forall m \in M,\ \forall l \in L \tag{12}$$

$$\lceil \frac{\underline{T}_{mo}^h}{\overline{t}_m^h} \rceil \leq \sum_{l \in l(o)} \sum_{\substack{k \in K_o: \\ m \in m(k)}} \sum_{\substack{b \in b(k): \\ m \in m(b,k)}} w_k \leq \lfloor \frac{\overline{T}_{mo}^h}{\underline{t}_m^h} \rfloor, \qquad \forall m \in M,\ \forall o \in O \tag{13}$$

$$\lceil \frac{\underline{T}_l^h}{\overline{t}_l^h} \rceil \leq \sum_{k \in K_l} w_k \leq \lfloor \frac{\overline{T}_l^h}{\underline{t}_l^h} \rfloor, \qquad \forall l \in L \tag{14}$$

Constraint (6) ensures that the total tonnage flowing through the system is in the permissible range, constraints (7) ensure that the tonnage flowing through a terminal is in the permissible range, constraints (8) ensure that the tonnage sourced from a mine is in the permissible range, constraints (9) ensure that the brand tonnages handled by and operator are in the the permissable range, and constraints (10) ensure that the component tonnages handled by an operator are in the permissable range. Constraints (11), (12), and (13) ensure that the limits on the component tonnage on a single trip are respected. For example, because $\sum_{k \in K:\, m \in m(k)} \sum_{b \in b(k):\, m \in m(b,k)} w_k$ can be thought of as representing the number of trips to which mine m contributes and the values $\lceil \frac{\underline{T}_m^h}{\overline{t}_m^h} \rceil$ and $\lfloor \frac{\overline{T}_m^h}{\underline{t}_m^h} \rfloor$ give lower and upper bounds on the number of trips, respectively, if the limits on the component that can be sourced from the mine m for a single trip are to be respected. Finally, constraints (14) bound the number of trips to a terminal so as to respect the limits on a trip's tonnage at a terminal.

We next discuss the objectives in the hierarchy. All objectives are based on minimizing the deviation from a target. The ranking of the objectives reflects the priorities of the HVCCC, which, to some extent, are aligned with the size of the entities considered, e.g., the total tonnage flowing through the system and the total tonnage flowing through terminals are at the top of the hierarchy, whereas per-trip tonnages appear further down the hierarchy.

To control the total tonnage flowing through the system, we define a variable $\delta \geq 0$ capturing the absolute deviation from the target total system-wide tonnage. More precisely:

$$\sum_{k \in K} \tilde{t}_k w_k - \overline{T}^s \leq \delta$$
$$\underline{T}^s - \sum_{k \in K} \tilde{t}_k w_k \leq \delta$$

We then minimize δ.

To control the total tonnage flowing through the terminals, we define variables $\delta_l \geq 0$ for each terminal $l \in L$ to capture the absolute deviation from the target total terminal tonnage. More precisely:

$$\sum_{k \in K_l} \tilde{t}_k w_k - \overline{T}_l^s \leq \delta_l, \quad l \in L$$
$$\underline{T}_l^s - \sum_{k \in K_l} \tilde{t}_k w_k \leq \delta_l, \quad l \in L$$

We then minimize $\sum_{l \in L} \frac{\overline{T}_l^s - \underline{T}_l^s}{T_l^t} \delta_l^2$. Observe that the coefficients in the objective function represent the width of the desirable range for the terminal divided by the target for the terminal. By dividing the absolute deviations by the (associated) target value, differences that arise due to differences in the size of the terminals are eliminated. By multiplying the absolute deviations with the width of the desirable range, the importance the user assigns to a terminal is reflected. Note too that the minimization involves the sum of the squares of the deviations. When using the sum of the deviations, a deviation of τ at, say, three terminals (giving an objective function value of 3τ) is equivalent to a violation of 3τ at one of the terminals (also giving an objective function value of 3τ), whereas when using the sum of the squares of the deviations, a deviation of τ at three terminals (giving an objective function value of $3\tau^2$) is not equivalent to a violation of 3τ at one of the terminals (as it gives an objective function value of $9\tau^2$). Consequently, minimizing the sum of the squares of the deviations tends to spread deviations across the terminals (as opposed to concentrating the deviation at a single terminal), which is preferred by the planners. In a similar way, we minimize the absolute deviation from the target total mine-operator tonnages, the target total brand-operator tonnages, and the target total mine tonnages.

This is followed by minimizing the deviation from the target vessel class fraction at a terminal and the target average mine component tonnage per trip. As these characteristics are ratios of variables, they require handling of nonlinear terms. In the case of the target vessel class fraction at a terminal, this involves approximating the number of trips that berth at terminal l, i.e., approximating $\sum_{k \in K_l} w_k$. This done as follows. First, we compute $\bar{t}_l$ the expected tonnage of a trip berthing at terminal l: $\bar{t}_l = \sum_{v \in V} f_{lv} \sum_{k \in K_v} t_k / |K_v|$. Then, we set the approximate number of trips berthing at terminal l to be $T_l^t / \bar{t}_l$.

Finally, for every historical trip, we want its relative importance to be as close as possible to its original relative importance, i.e., to $\frac{1}{|K|}$. We accomplish this by introducing non-negative variables δ_k for $k \in K$, constraints

$$-\delta_k \leq \sum_{k' \in K} w_{k'} - |K| w_k \leq \delta_k, \; \forall k \in K$$

and then minimize $\sum_{k \in K} \delta_k^2$.

At the end of the hierarchical optimization, we have a likely number of required copies for each historical trip $k \in K$, namely w_k.

3.3 Trip sampling

The goal of the sampling module is to generate a set of historic trips that can be selected and adjusted in the trip generation module to form the basis of the stem that is to be created.

The first step is to examine the likely number of required copies of each trip $k \in K$ determined in the trip assessment phase, i.e., w_k, and convert these into trip likelihoods or probabilities as follows. As the shipping stem to be created has to reflect the forecast mine production, we set the probability that copies of trip $k \in K$ are used in the shipping stem to satisfy the forecast mine production of mine $m \in M$ to $\frac{w_k}{\sum_{k' \in K : m \in m(k')} w_{k'}}$. Next, cumulative distribution functions (CDFs) for each of the mines are created using these trip probabilities. This allows us to sample a trip for a mine by drawing a uniform random number in the range $[0, 1)$ and using the mine's CDF to identify the corresponding trip.

The second step is to include $\lfloor w_k \rfloor$ copies of each trip $k \in K$ in the trip generation model. This reduces the amount of sampling that needs to be performed by ensuring that trips with a small likely number of required copies, but still greater than one, are represented in the model. This was found to be important to effectively handle mines with a small forecast production. Without the inclusion of these trips, it sometimes happened that even when sampling a large number of trips, there were no trips that visited such a mine.

The final step is to sample trips until the production forecast of each mine and the target throughput of each terminal is covered by the sampled trips (including the trips set aside in the previous step), i.e., the sum of the tonnages of the selected trips is more than the forecast mine production and more than the target terminal throughput. More precisely, we maintain the set of mines ordered by remaining uncovered tonnage. For the mine with the largest remaining uncovered tonnage, we sample a trip for that mine and update the covered tonnage (and thus the remaining uncovered tonnage) for all the affected mines, i.e., the mines from which components are sourced in the trip. We repeatedly select trips until there are no mines left with remaining uncovered tonnage. (Note that the mines are reordered whenever appropriate.) Next, we ensure that the target throughput of each terminal is covered by the sampled trips in a similar way.

To create additional flexibility and reduce the chance for infeasibility during trip generation we over-sample, i.e., we continue sampling trips for mine m until ρT_m^t tonnage is covered and trips for terminal l until ρT_l^t tonnage is covered, where $\rho > 1$ is the over-sampling factor.

3.4 Trip generation

The trip generation model resembles the trip assessment model, but is distinctly different. In the trip assessment model, we used linear combinations of the historical trips. In the trip generation model, we no longer allow linear combinations of historic trips, but we allow the brand proportions in a cargo-profile and the component proportions in a brand recipe of a trip to change, within bounds, from the historical proportions. The bounds, i.e., the allowable deviations, are provided by the planners and reflect their predictions about the expected deviations in cargo-profiles and their knowledge about brand recipes. Note that we may have multiple copies of the same trip. In the description of the model below, each of these trips is considered separately, but for consistency with the trip assessment model, we still use K to represent the set of trips. The model has the following variables:

- $x_{mbk} \geq 0$ for all $k \in K, b \in b(k)$, and $m \in m(b,k)$ denotes the tonnage sourced from mine m for brand b in the cargo-profile of trip k.
- $z_k \in \{0,1\}$ for all $k \in K$ indicates whether trip k is selected.

Since the model is actually building trips, by determining the tonnage sourced from a mine for a brand in the cargo-profile of the trip, constraints are needed to enforce trip-related permissible ranges:

$$\max((1-\epsilon_t)t_k, \underline{t}^h_{l_k}, \underline{t}^h_{v_k})z_k \leq \sum_{b \in b(k)} \sum_{m \in m(b,k)} x_{mbk} \leq \min((1+\epsilon_t)t_k, \bar{t}^h_{l_k}, \bar{t}^h_{v_k})z_k \quad \forall k \in K \tag{15}$$

$$\max(\underline{t}^h_{bl_k}, \underline{t}^h_{bv_k})z_k \leq \sum_{m \in m(b,k)} x_{mbk} \leq \min(\bar{t}^h_{bl_k}, \bar{t}^h_{bv_k})z_k \quad \forall k \in K,\ \forall b \in b(k) \tag{16}$$

$$\underline{t}^h_m z_k \leq x_{mbk} \leq \bar{t}^h_m z_k \quad \forall k \in K,\ \forall b \in b(k),\ \forall m \in m(b,k) \tag{17}$$

$$(f_{mbk} - \epsilon_f) \sum_{m' \in m(b,k)} x_{m'bk} \leq x_{mbk} \leq (f_{mbk} + \epsilon_f) \sum_{m' \in m(b,k)} x_{m'bk} \quad \forall k \in K,\ \forall b \in b(k),\ \forall m \in m(b,k). \tag{18}$$

Constraints (15) ensure that the tonnage on a trip is within the permissable range for the trip, which depends on the terminal the trip is berthing at, the class of the vessel performing the trip, and its original trip tonnage. A slight, user controlled, deviation from the original trip tonnage is allowed. Constraints (16) ensure that the brand tonnage on a trip is within the permissible range for that trip, which depends on the terminal the trip is berthing at and the class of the vessel performing the trip. Constraints (17) ensure that the component tonnage on a trip is within the permissable range for the associated mine. Finally, constraints (18) ensure that brand recipe does not deviate too much from the original. The allowed deviation is controlled by the user. Observe that these constraints also ensure that when a trip is not selected, the x variables for that trip are set to zero. These constraints are in addition to constraints (6)–(10) of the trip assessment model (the constraints from the trip assessment model need to be modified slightly to use x_{mbk} instead of $t_{mbk}w_k$).

We use hierarchical optimization to handle desirable ranges. All objectives are to minimize the deviation from the desirable ranges. However, for computational efficiency, we use linear terms instead of quadratic terms for the deviations. The first objectives are similar to those in the trip assessment model, i.e., system tonnage, terminal tonnage, mine tonnage, brand-operator tonnage, mine-operator tonnage, terminal-vessel fractions, and average mine component tonnage per trip. This is followed by trip-related objectives, i.e., component tonnage, trip tonnage, and brand-recipe fractions. Recall that in the trip generation model, the tonnage sourced from a mine for a brand in the cargo-profile of a trip is a variable. Therefore, the trip tonnage and the mine component tonnages can deviate from the historical ones. A brand-recipe fraction refers to the fraction of a mine component in a brand. Again, because the tonnage sourced from a mine for a brand is a variable, the fraction can deviate from the historical one. The trip-related objectives try to minimize these deviations.

Let the final solution be given by (x^*, z^*). Then $\hat{K} = \{k \in K : z_k^* = 1\}$ is the set of trips that will make up the shipping stem. The characteristics of the trips in $\hat{K}$ is determined by x^*, e.g., the trip tonnage t_k of a trip $k \in \hat{K}$ is $\sum_{b \in b(k)} \sum_{m \in m(b,k)} x^*_{mbk}$.

3.5 Trip balancing and timing

At the completion of the trip generation model, we have a set of trips satisfying the mine production forecasts (within the permissable range) for the planning period. What remains is to assign trip arrival times to have a complete shipping stem. This is nontrivial, however, as the analysis of historical trips has shown that the arrival pattern is not random, but smoothed out over the year with respect to a number of characteristics:

- the coal production at each of the mines,
- the brand tonnages handled by operators,
- the tonnages flowing through the terminals, and
- the number of berths at a terminal.

This smoothing is not unexpected as producers and terminal operators try to ensure in their contractual agreements that demand materializes throughout the year so as to be able to create operational efficiencies.

To accomplish smoothing, we divide the planning period into weeks and try to assign trips to weeks in such a way that each of the weeks has similar characteristics. Let the set of weeks be denoted by W. We use an integer programming model to assign trips to weeks. Let the binary variables y_{kw} indicate whether trip $k \in \hat{K}$ is assigned to week $w \in W$. Feasibility requires $\sum_{j \in S} y_{kw} = 1 \; \forall k \in \hat{K}$.

Next, consider a characteristic a that needs to be smoothed. Let a_k denote the value of the characteristic in trip k and let $A = \sum_{k \in \hat{K}} a_k$ be the total value of the characteristic over the planning period. We let $[\frac{A}{|W|} - \tau_a, \frac{A}{|W|} + \tau_a]$ be the desirable range for the weekly value of characteristic a, where τ_a is a pre-specified parameter. The following integer program distributes the trips over the weeks so as to best match the desirable range:

$$\min \sum_{w \in W} \delta_w^2$$

subject to

$$\sum_{k \in \hat{K}} a_k y_{kw} - (\frac{A}{|W|} + \tau_a) \leq \delta_w, \quad w \in W$$
$$(\frac{A}{|W|} - \tau_a) - \sum_{k \in \hat{K}} a_k y_{kw} \leq \delta_w, \quad w \in W,$$

where δ_w represents the violation of the desirable range in week w, either above or below.

The above optimization has one short-coming: it does not explicitly take into account the fact that weeks form an ordered sequence, and it may thus happen that in a solution the value is at the lower end of the desirable range for the first weeks in the planning period and at the upper end of the desirable range for last weeks of the planning period. That is, the optimization only indirectly tries to ensure a spread of values over the planning period.

An alternative integer program, which explicitly incorporates the sequential nature of time, seeks to match the perfect cumulative value of the characteristic at the end of each week w in the planning period, i.e., it seeks to match $w\frac{A}{|W|}$. To be more precise

$$\min \sum_{w \in W} \delta_w^2$$

subject to

$$\sum_{\substack{w' \in W: \\ w' \leq w}} \sum_{k \in \hat{K}} a_k y_{kw'} - w\frac{A}{|W|} \leq \delta_w, \quad w \in W$$

$$w\frac{A}{|W|} - \sum_{\substack{w' \in W: \\ w' \leq w}} \sum_{k \in \hat{K}} a_k y_{kw'} \leq \delta_w, \quad w \in W.$$

We found that in practice the most effect smoothing strategy, in terms of efficiency as well as solution quality, is to embed both integer programs in a hierarchical scheme. That is, first distribute the trips over the weeks so as to best match the desirable range for the weekly value, which gives a maximum violation δ_a, if any, of the desirable range in any week of the planning period, followed by distributing trips over the weeks so as to minimize the violation of the cumulative end-of-week targets subject to the constraint that each weekly value has to be within the range $[\frac{A}{|W|} - \tau_a - \delta_a, \frac{A}{|W|} + \tau_a + \delta_a]$.

Trip balancing is thus accomplished using a hierarchical integer program in which the following objectives are optimized in order: deviation from average weekly mine tonnage, from perfect end-of-week mine tonnage, from average weekly brand-operator tonnage, from end-of-week brand-operator tonnage, from average weekly terminal tonnage, from end-of-week terminal tonnage, from average weekly terminal count, and from end-of-week terminal count.

Note that the smoothing simultaneously attempts to ensure small weekly variances of high volume tonnages and regular spacing over the planning horizon of low volume tonnages.

In practice, since the generated shipping stems will be used in simulation studies, the planning period is usually expanded (and thus the mine production forecasts scaled up) so that the planning period includes warmup weeks as well as post warm-up weeks. As a consequence, separate tonnage targets and constraints have to be introduced for just the post warm up period. The main effect of this is that the integer programs get larger and therefore, rather than solving for all $|W|$ weeks at once, a two-pass approach is employed. In the first pass, W is partitioned into buckets, each bucket containing 6 or 7 weeks, and trips are assigned to a bucket. In the second pass, each trip is then assigned to a week from the bucket to which that trip was assigned in the first pass.

The final step in the trip balancing and timing module is to assign an arrival time (in hours) to the trips in each week. The trips in a week are ordered randomly and an exponential distribution with rate $\frac{n}{h}$ is used to obtain inter-arrival times between consecutive trips, where n is the number of trips in the time bucket and $h = 168$, the number of hours in a week. The strategic planning team at HVCCC performed an in-depth analysis that showed that an exponential distribution was a good fit to ship inter-arrival times.

3.6 Accommodating a changing environment

As mentioned in the introduction, generating shipping stems is further complicated by changes to the environment. How to account for a new terminal? How to account for the decommissioning of existing mines and the introduction of new mines?

A simple and effective way to handle these complexities, and the approach that we have adopted, is to have an expert with an understanding of the HVCC and of the producers operating in the HVCC "manipulate" the set of historical trips by creating new "historical" trips based on his or her beliefs about which mines will be used to source product in place of decommissioned mines, about what new brands and associated brand recipes are likely to be seen in the future, about what type of trips can be expected to be loaded at a new terminal, etc.

To facilitate manipulating a set of historical trips, we have developed a simple tool that automates some of the tasks that frequently occur in this process. The tool takes as input a list of mines to be decommissioned and a list of new mines with for each new mine a "most similar" existing mine. The tool performs the following actions:

- For a new mine $m \in M$ with a most similar mine $\overline{m} \in M$, it takes every trip $\overline{k} \in K$ that sourced product from $\overline{m}$, i.e., $\overline{m} \in m(\overline{k})$, and, for every $b \in b(\overline{k})$ such that $\overline{m} \in m(b, \overline{k})$, creates a copy k of trip $\overline{k}$ in which $\overline{m}$ is replaced by m (i.e., $m(b, k) = (m(b, \overline{k}) \setminus \{\overline{m}\}) \cup \{m\}$ and $t_{mbk} = t_{\overline{m}b\overline{k}}$). In fact, a new trip is not just created for each singleton component $\{m\} \subseteq M$ with most similar mine $\{\overline{m}\} \subseteq m(b, \overline{k})$, but for all subsets of new mines $M' \subseteq M$ with most similar mines $\overline{M}' \subseteq m(b, \overline{k})$ that contribute to brand $b \in b(\overline{k})$ in the cargo-profile of trip $\overline{k}$. Trip likelihoods of the original and the generated trips are set to be equal and to sum up to one.
- It removes all the trips that sourced coal from a mine that is decommissioned.

4 Computational results

In this section, we report the results of two sets of computational experiments. The first set of experiments focuses on validating the proposed approach. The second set of experiments explores its ability to generate useful future shipping stems. All computational experiments were conducted on a Dell PowerEdge 2950 with dual quad core 3.16GHz Intel Xeon X546C processors and 64Gb of RAM running Red Hat Enterprise Linux 5 using CPLEX v12.3 running in deterministic mode using a single thread. Our approach has been implemented using the AMPL modeling language and system for formulating, solving and analyzing large-scale optimization problems (Fourer et al. [9]). The use of a powerful and flexible modeling language has greatly facilitated the development and maintenance of the approach.

4.1 Historic mine production

To validate the proposed approach, we generate shipping stems with the forecast mine production for each mine set to the mine's actual production in 2008. The generated shipping stem should thus exhibit characteristics that are similar (not necessarily identical) to the shipping stem that occurred in 2008. Appropriate permissable and desirable ranges were determined in conjunction with the strategic planning team at HVCCC. (Note that permissable and desirable range information was not available for the historical shipping stem.)

Computational parameters that influence the results of the proposed approach are the over-sampling factor and the time limit imposed for each integer programming solve. The default values for these parameters are 1.2 and 1 hour. These default parameters were determined through initial computational experiments. They provide a good balance between computational efficiency and quality of results. The proposed approach is not very sensitive to the over-sampling factor and the time limit parameters. For example, increasing the over-sampling factor to 1.5 and increasing the time limit to 1.5 hours does not lead to substantially different results.

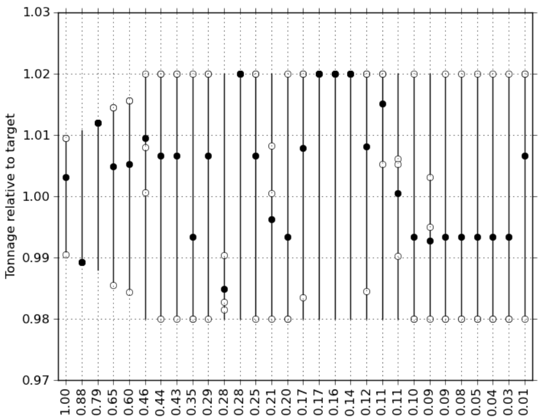

Figure 4: Mine tonnages relative to targets

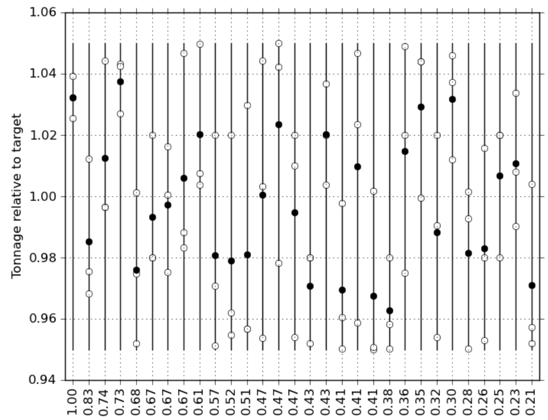

Figure 5: Mine average component tonnages relative to targets

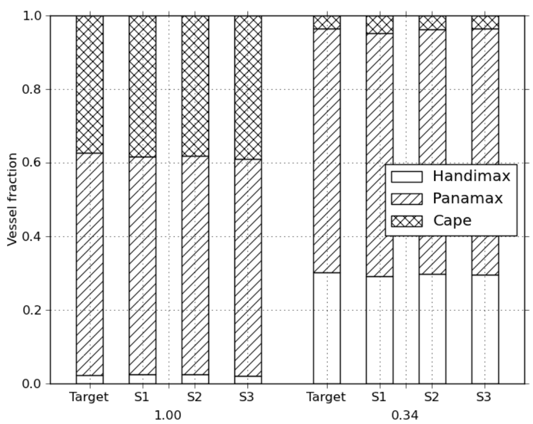

Figure 6: Terminal-vessel fractions

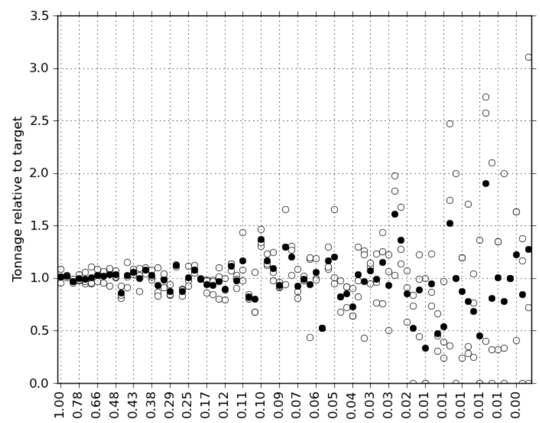

Figure 7: Brand tonnages relative to history

Due to the sampling that is part of our proposed approach, it is possible to generated multiple shipping stems for a single data and parameter set. The HVCCC strategic planning team exploits this possibility by typically generating three shipping stems, performing an examination of each of them, and then selecting one to be used as input for a detailed simulation study of the system performance. Therefore, the results that we present for our validation experiment also involve three generated shipping stems.

Recall that there are a number of characteristics of a shipping stem that are of interest. We will present and discuss a few of the key characteristics.

One of the most important characteristics of a shipping stem are the mine tonnages, as forecast mine productions are the starting point for shipping stem generation. In Figure 4, we show for each of the mines the lower and upper limit of the desirable range about the mine tonnage target (where the mine tonnage target is set at 1.0) and the mine tonnage of the three shipping stems generated (as a fraction of the mine tonnage target), each represented by an empty circle, and the average mine tonnage over the three shipping stems generated, represented by a solid circle. The mines are listed in order of non-increasing target tonnage, where the tonnage of each mine is represented along the axis as a fraction of the largest mine tonnage. We see that for the mines with the highest target tonnages, the desirable ranges are carefully chosen, whereas for the mines with smaller tonnages the desirable range is simply set to plus and minus 2% about the target tonnage value. Furthermore, we see that for all mines the mine tonnage of the three generated shipping stems is within the desirable range.

Not only is the annual tonnage of a mine important, but also the average amount of coal that a mine contributes to a trip in a shipping stem (as that relates to operational efficiencies). In Figure 5, we show for each of the mines the lower and upper limit of the desirable range about the mine's average component tonnage per trip target (where the mine's average component tonnage per trip target is set at 1.0) and the mine's average component tonnage in the three shipping stems generated (as a fraction of the mine's average component tonnage per trip target), each represented by an empty circle, and the average of the mine's average component tonnage over the three generated shipping stems, represented by a solid circle. The mines are listed in order of non-increasing target average component tonnage per trip, where this tonnage is represented along the axis as a fraction of the largest target average component tonnage per trip. We see that for all mines the mine's average component tonnage per trip in the three generated shipping stems is within the desirable range.

In Figure 6, we present for each of the two terminals the mix of vessels visiting the terminal, i.e., for each of the terminals we show, in the form of stacked bars, for each of the vessel classes what fraction it constitutes of the vessels visiting the terminal. More specifically, we show for each of the terminals and for each vessel class, the target fraction (the fraction observed in the historical shipping stem) and the fraction in each of the three shipping stems generated. We see that for both terminals and for all vessel classes the vessel fractions in the three generated stems are close to the target vessel fractions.

In Figure 7, we show for each of the brands loaded during the year the tonnage of that brand loaded in each of the three generated shipping stems (as a fraction of the tonnage of the brand loaded in the historical stem), represented by empty circles, and the average tonnage loaded over the three shipping stems generated, represented by a solid circle. The brands are listed in order of non-increasing historical tonnage, where the tonnage of each brand is represented along the axis as a fraction of the largest brand tonnage. We see that for the brands with the highest historical tonnages, the generated shipping stems match the historical tonnages quite well, but that for brands with low historical tonnages the deviations can be substantial. This is not surprising as some of the brands with low historical tonnages have occurred only on a few historic trips and if they occur in one fewer or one more trip in the generated shipping stems, the deviation can be substantial. The changes in brand proportions in the cargo-profiles and the changes in component proportions in brand recipes for the trips in the generated shipping stems are all within the specified limits.

Finally, we investigate how well the balancing module smooths out the trips. For the balancing experiments, we have used the information from the historical shipping stem to determine the desirable range for the weekly value of a characteristic. More specifically, for characteristic a, we set τ_a equal to the maximum deviation from the weekly average, either up or down, observed in any of the weeks in the historical shipping stem. In Figure 8, we show the range of these deviations from the weekly average for each mine in the historical stem in the form of a line, i.e., from the largest deviation below the weekly average observed in any of the 52 weeks to the largest deviation above the weekly average observed in any of the 52 weeks. We show the range of the deviations from the weekly average for each mine in the generated stems in the form of a rectangle (i.e., the average of the maximum deviations over the three generated stems). The mines are listed in non-increasing order of total annual tonnage sourced from the mine. Similarly, in Figure 9 we show the range of the deviations from the cumulative end-of-week target values for each mine in the historical stem in

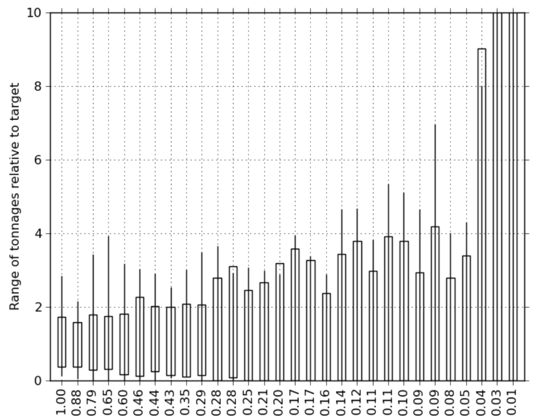

Figure 8: Balance mine weekly tonnage.

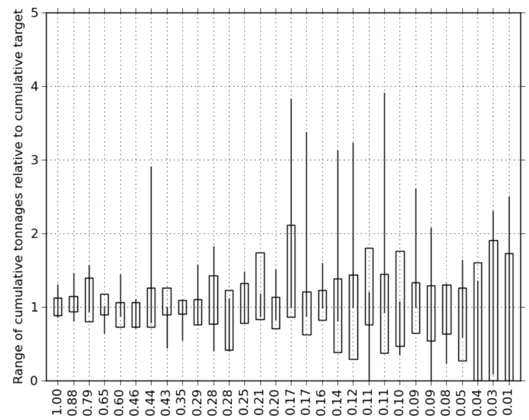

Figure 9: Balance mine end-of-week tonnage.

the form of a line and the range of the deviations from the cumulative end-of-week target values for each mine in the generated stems in the form of a rectangle. We see that the proposed methodology results in more smoothness than observed in the historic stem, but not excessively so; for all the mines there is noticeable and sufficient variation in the tonnage sourced from week to week, which is what the HVCCC is looking for. In fact, they are very interested in analyzing shipping stems with more smoothness as they expect to have more control over the ship arrivals in the future. We note too that the larger deviations are occurring for the mines from which only a small tonnage is sourced. In many cases, coal components from these mines are only sourced a few times a year which leads to somewhat meaningless weekly average values. For example, we see that for the two mines from which the smallest tonnage is sourced, the largest weekly value is more than ten times the weekly average, which is simply a result of the fact that no coal is sourced from these mines in most of the weeks in the planning period. The cumulative end-of-week targets for these mines are much more important, and we see that the deviations from these cumulative end-of-week targets is relatively small also for these mines.

Thus, the methodology does a good job of managing the average and the spread. Managing the average is essential for mines with a high forecast production whereas managing the spread is essential for mines with a low forecast production.

4.2 Forecast mine production

The next set of experiments explores whether the proposed approach can generate useful future shipping stems. The strategic planning team at HVCCC created two data sets, one in which the system throughput is 140 Mtpa and one in which the system throughput is 160 Mtpa. These data sets are a good indication of the strategic planning challenges the HVCCC faces, namely a significant increase in annual throughput: from a little over 90 Mtpa in 2008 to an anticipated 225 Mtpa in 2013. Both data sets contain a new terminal (which came into operation early 2010), have four new mines and four decommissioned mines, and have six new brands and five existing brands that are no longer used. The latter information, i.e., the information regarding new brands and existing brands that are no longer used represents an assessment of the planning team and is based on knowledge of the locations of the new and decommissioned mines and of the producers that own the new and decommissioned mines. To accommodate the changing environment, the planners created about 1500 new trips to the new terminal. About 300 of these trips had at least one of the new brands in the cargo profile; the vast majority had one new brand in the cargo-profile, but a few had two new brands and one even had three new brands in the cargo-profile. The new mines contribute only to the new brands and therefore only send coal to the new terminal. About 200 of the newly created trips involve coal from a new mine.

Because the results for 160Mtpa data set are similar to the results for the 140Mtpa data set, we only present the results for the 140Mtpa data set.

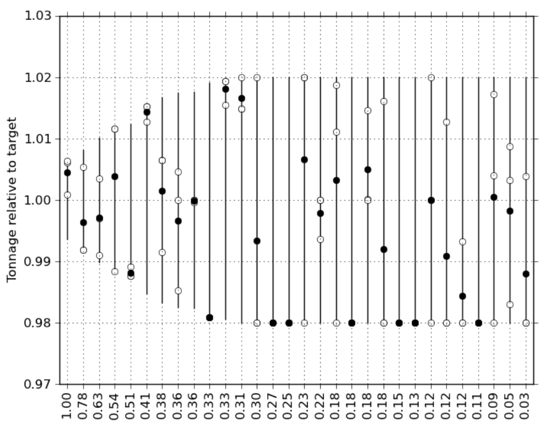

Figure 10: Mine tonnages relative to targets (140 Mtpa scenario)

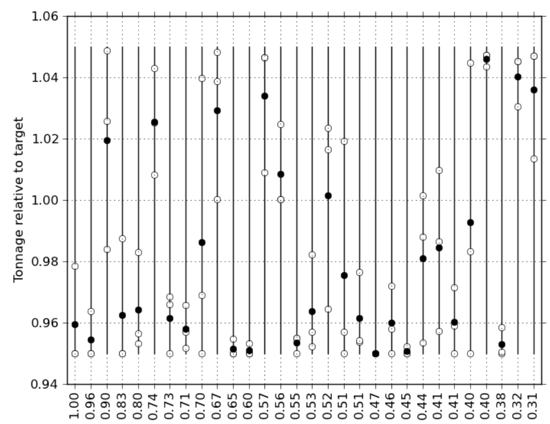

Figure 11: Mine average component tonnages relative to targets (140 Mtpa scenario)

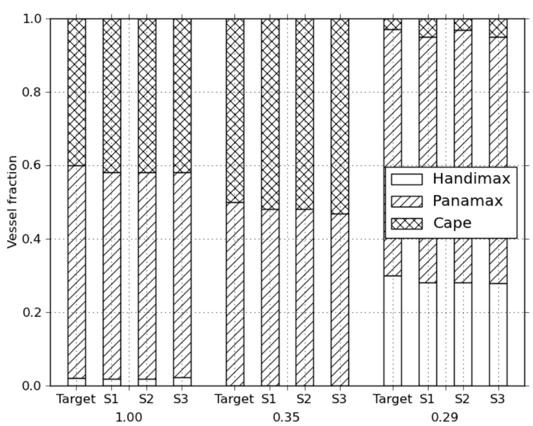

Figure 12: Terminal-vessel fractions (140 Mtpa scenario)

In Figure 10, we show for each of the mines the lower and upper limit of the desirable range about the mine tonnage target and the mine tonnages in the three shipping stems generated (as well as the average mine tonnage over the three shipping stems). We see a pattern similar to the one we observed in the validation experiment described above. The mine tonnages are within the desirable range for all mines for each of the three generated shipping stems.

In Figure 11, we show for each of the mines the lower and upper limit of the desirable range about the mine's average component tonnage per trip target and the mine's average component tonnage per trip in the three shipping stems generated (as well as the average of the mine's average

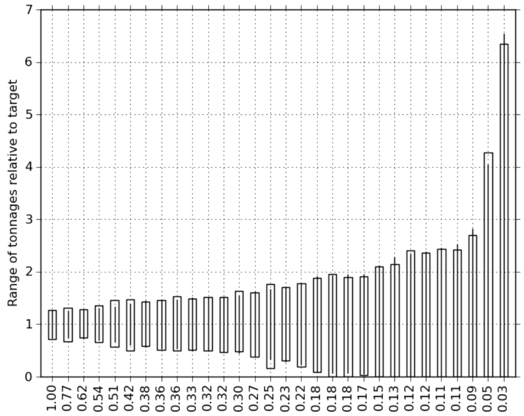

Figure 13: Balance mine weekly tonnage (140 Mtpa scenario)

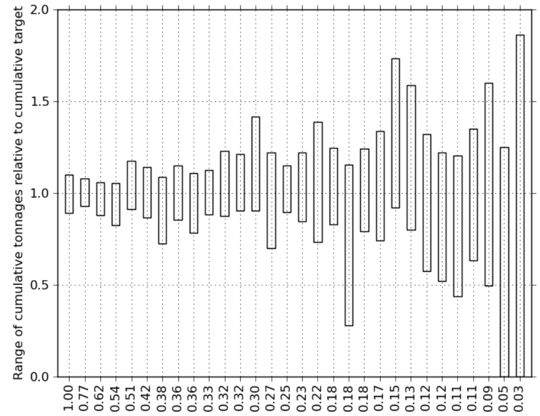

Figure 14: Balance mine end-of-week tonnage (140 Mtpa scenario)

component tonnage per trip over the three shipping stems). The mine's average component tonnage is within the desirable range for all mines for each of the three generated shipping stems.

In Figure 12, we present for each of the three terminals the mix of vessels visiting the terminal, i.e., for each of the terminals we show, in the form of stacked bars, for each of the vessel classes what fraction they constitute of the vessels visiting the terminal. More specifically, we show for each of the terminals and for each vessel class, the target fraction and the fraction in each of the three shipping stems generated. We observe that the new terminal (with label 0.35) does not accept visits by vessels in the Handimax class, which is the class with the smallest vessels. Furthermore, we see that the terminal-vessel fractions are all close to their targets in each of the three generated shipping stems.

Next, we examine the trip balancing (Figures 13 and 14). As mentioned before, the strategic planning team of the HVCCC expects that they will have more control over vessel arrivals in the future and therefore wants the generated shipping stems to exhibit more smoothness than the historic shipping stem. Consequently, they have provided relatively narrow desirable ranges for the weekly values. We observe, first and foremost, that the larger forecast system tonnage (and the associated larger forecast mine tonnages) allow for better smoothing than what was observed in the historical shipping stem. In the historical data, deviations from the weekly average could be in the order of 60 to 80 percent even for the mines with the largest sourced tonnages, whereas in the 140 Mtpa scenario this does not happen and the deviations are in the order of 10 to 30 percent. We see that it was not possible to be within the desirable range for all of the mines. We see too that the cumulative end-of-week tonnages are close to the targets even for mines with small forecast tonnage.

What the future will bring in terms of demand for coal from the Hunter Valley is and will remain a moving target. There is only one certainty: for the next few years the demand will continue to go up. At the writing of this paper, the HVCCC is generating shipping stems for data sets including a fourth terminal and with mine production forecasts totalling more than 250 Mtpa.

5 Implementation and Impact

The development of the stem generation tool was initiated by the strategic planning team at HVCCC. The strategic planning team at HVCCC conducts numerous studies throughout the year to assess the impact of potential capacity expansions and operational changes on the throughput of the coal chain. Increasing the throughput of the coal chain is necessary to be able to accommodate the growth in demand for coal. Most of these studies are carried out with the assistance of one of four different simulation models or an optimization model, each of which requires a shipping stem as input. It is important to realize that shipping stems do not simply represent a detailed view of forecast demand. A shipping stem also needs to reflect any changes in demand management policies or terminal design and stockpile management strategies, because such changes result in different ship arrival patterns, cargo sizes, and vessel profiles at each of the terminals. This is one of the reasons why the stem generation tool provides a variety of ways to control the characteristics of the shipping stem being generated, e.g., importance ranking of characteristics and target values, desirable ranges, and permissable ranges for characteristic values. Shipping stems should also be able to reflect planned or anticipated changes to the coal chain itself, e.g., new terminals, new mines, and new brands.

To examine a shipping stem, either manually or automatically created, the HVCCC strategic planning team uses a formidable Excel spreadsheet with a large number of worksheets, each containing various pivot tables and graphs. To provide a flavor of the information that is used by the HVCCC strategic planning team to examine a shipping stem, we reproduce a few of the graphs in Figures 15-18. The two graphs at the top (Figures 15 and 16) show the historical weekly production in tonnes at two of the mines; one mine with a medium-size production and one mine with a low production. The dashed lines are regression lines that ideally should be horizontal. If the regression lines are not horizontal, then the production is not smoothed out over the year, which is a reason for concern. The graphs clearly show the variation through the year in terms of production, even for a mine with a medium-size production. The planners will have to decide whether this is acceptable or whether the shipping stem has to be modified in order to produce a smoother production pattern at these mines, which better reflects what is likely to happen in reality. The third graph (Figure 17) shows the weekly throughput in tonnes at one of the terminals and the fourth graph (Figure 18) shows the trip counts for the same terminal, i.e., the number of vessels arriving at each terminal in each week. Again, the regression lines should be horizontal, which they are in this case. This is an indication for the planners that the shipping stem reflects the pattern that is likely to be observed at the terminal. Even though they are related, the terminal tonnage and terminal trip counts provide different information. The overheads involved in handling each vessel make high trip counts undesirable; they put severe strain on the vessel handling components of the port operation.

The graphs illustrate that there is substantial variation through the year at both the mines (in terms of production) and at the terminals (in terms of tonnage and trip counts). This variation in combination with the complex interactions among the different entities in the Hunter Valley coal export chain, e.g., mines, terminals, and vessels, makes it very difficult to create a shipping stem with desirable characteristics. For example, manually trying to adjust the shipping stem in such a way that the regression line at the low production mine becomes more horizontal may lead to undesirable trip counts at one or more terminals. Trying to correct for these undesirable trip counts at terminals may lead to undesirable brand-operator tonnages, and so on and so forth. Simultaneously considering all the complex interactions when manipulating a shipping stem has become too hard for a human planner. As the graphs show, there is significant variation in activity during the year at the various entities in the coal chain. As any logistics expert knows, reducing

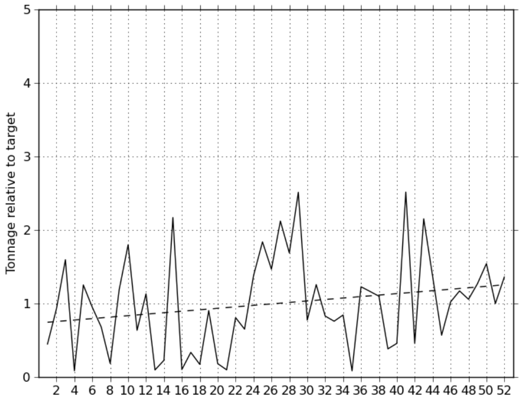

Figure 15: Medium-size mine.

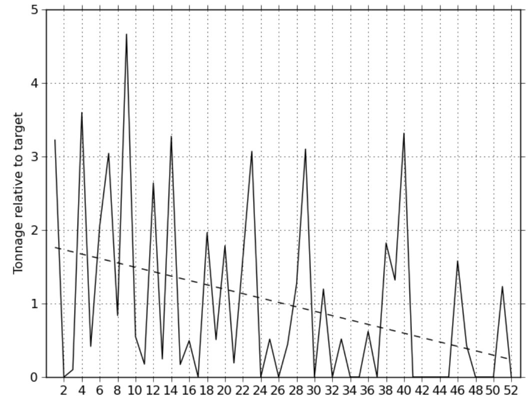

Figure 16: Small mine.

Figure 17: Terminal tonnage.

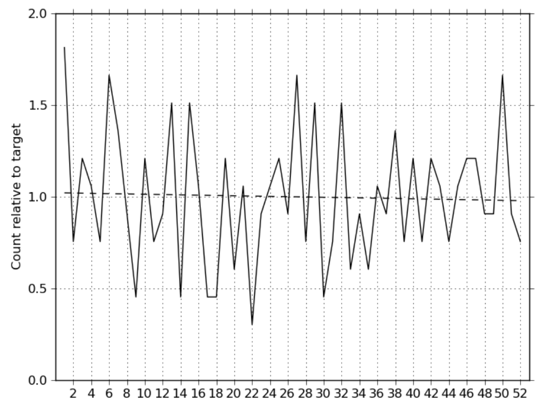

Figure 18: Terminal trip count.

variation is a crucial element in any strategy aimed at improving operational efficiencies, or, in the case of the Hunter Valley coal export chain, aimed at increasing throughput. To better understand the impact of reduced variation on throughput, the HVCCC needs to be able to generate shipping stems that resemble history, but also reflect possible demand management strategies. Without our optimization-based stem generation tool this is virtually impossible.

The major challenge during the development of the stem generation tool was the lack of insight into brand information. Brand information is highly confidential and not released to the strategic planning team by the mining companies. Furthermore, an in-depth analysis of historic shipping stems revealed both clear variations in brand recipes and clear correlations between the brands in cargo-profiles. Reaching consensus on what it means for a shipping stem to exhibit an appropriate amount of variation and an appropriate amount of correlation was difficult and developing methodology that produces shipping stems that exhibit the desired behavior was not easy. Another challenge was to balance the planning team's desire to introduce an element of randomness in the shipping stems generated and the desire to control the characteristics of the shipping stems generated. Developing technology that can produce even a single shipping stem satisfying the various permissable ranges and meeting the different targets was a complex undertaking. Complementing that technology with an appropriate random sampling mechanism was non-trivial. A practical challenge encountered during the development of the stem generation tool was the substantial effort involved in validating shipping stems produced by the tool. The strategic planning team has a small number of staff and their services are in high demand. Finding the time to carefully analyze a shipping stem, which involves examining a year's worth of ship arrivals and cargo profiles, was not always easy. This challenge was exacerbated by the fact that the insights gained during the development of the tool inevitably lead to changes in the requirements and thus, in turn, to even more evaluations.

However, the substantial benefits of having an efficient and effective stem generation tool meant that there was a strong desire to see the development through until the end. The benefits of the stem generation tool are threefold. First, it will free up valuable staff time as the time to prepare a complete shipping stem with certain desired characteristics will be reduced from several weeks to several hours. Second, the ability to configure shipping stems allows for demand profiles to be tailored to each terminal's operating design requirements. Third, with the capacity to easily and quickly tailor shipping stems comes the possibility, not presently available to the strategic planning team, to "fine tune" terminal operating and inbound rail and load point infrastructure designs based on slight demand profile differences as reflected in the generation of alternative shipping stems. The latter is important from a whole-of-coal-chain perspective as the ability to analyze more and different expansion options based on demand driven variations leads to a higher level of confidence in the final recommendations related to inbound infrastructure requirements.

To truly appreciate the above points, it is helpful to provide some additional insight into the rapidly changing environment in the HVCC. In 2008, there were two coal terminals operated by Port Waratah Coal Services (PWCS), which together handled a little over 92Mt of coal. In 2009, PWCS completed expansions that increased the capacity to about 110Mtpa. In 2010, the Newcastle Coal Infrastructure Group (NCIG) opened a new coal terminal able to handle 30 Mtpa. By 2012, PWCS expects to have completed further expansions that will increase its capacity to 130 Mtpa, and NCIG expects that by 2013, when their next expansion phase is fully operational, it will provide capacity of up to 53 Mtpa. Beyond that there are already plans to further expand the NCIG terminal to a capacity of 66 Mtpa, and plans for a fourth terminal (T4) with a capacity of up to 120 Mtpa are on the drawing board. This increase in terminal capacity is necessary to accommodate new mines in the expanding Ulan and Gunnedah regions. Of course, the rail infrastructure needs to be upgraded and expanded as well to be able to achieve higher throughput. Several track projects are either in progress or planned, involving the addition of extra passing loops and new track. But the challenge for the strategic planning team goes well beyond guiding and analyzing the impact of infrastructure expansions. The two PWCS terminals are *cargo assembly* terminals, whereas the NCIG is a *dedicated stockpile* terminal, and, most likely, so will be the new T4 terminal. That is, there is a trend away from a pull system towards a push system. However, effectively operating dedicated stockpile terminals requires much more careful demand management as only a limited number of brands will be available at any time of year and larger cargo sizes and smooth ship arrivals are necessary to achieve the maximum throughput of such a terminal. To perform optimization and simulation studies to analyze and evaluate the impact of the changing environment and to investigate other future options, such as inland terminals and new rail corridors, it is essential that shipping stems with pre-specified characteristics can be generated efficiently; preferably multiple shipping stems with the same pre-specified characteristics. The ability to generate multiple shipping stems with similar characteristics is also critical for the HVCCC's whole-of-coal-chain simulation model. Because many decisions of organizations across the coal chain are based on the predictions of this model, it is subject to an independent audit in order to provide confidence in the validity its predictions. Auditors have raised particularly the requirement that multiple stems with similar characteristics are provided for the audit of this simulation model.

The stem generation tool is completely built in the AMPL modeling environment. This ensures that the technology is relatively easy to extend in case new features are desirable. All input and output files are delimiter-separated ASCII files that can be written and read from most spreadsheet and database management systems. No sophisticated graphical user interface had to be developed as the members of the strategic planning team make extensive use of Microsoft Excel and are all expert Excel users and comfortable with importing and exporting ASCII files.

As mentioned earlier, the development of the stem generation tool took much longer than expected, but doing it right was more important than doing it quickly. As Rob Oyston, leader of the strategic planning team says: "Without the ability to generate trusted shipping stems with characteristics representing different future scenarios in terms of cargo sizes, arrival rates, and smoothness, we would not be able to demonstrate convincingly to the producers and other stakeholders that the planned or proposed infrastructure expansions and demand management policies are necessary to achieve the increased throughput required to accommodate the soaring demand for coal".

However, the greatest and most enduring impact of the stem generation tool is the fact that it allows the planning team to continue to do what it is charged to do. The scenarios that the planning team is looking at right now, with a throughput of more than 250 Mtpa, can no longer be generated manually. The complexity of preparing a shipping stem exhibiting all the desired characteristics for that level of demand has become too much for the human brain to accomplish, even when given weeks to do so. It has become too difficult to keep track of the intricate relationships between the various characteristics and to anticipate the consequences of a change to one part of the shipping stem on other parts of the shipping stem. The manual construction process simply did not converge anymore. The only way forward was the use of optimization-based stem generation technology.

6 Conclusions

Even though everyone is familiar with the expression "Garbage in equals garbage out", relatively little research is dedicated to the issues surrounding input data generation, which is especially relevant for strategic planning projects and exercises.

Here we have described a mathematical programming based approach for generating shipping stems, i.e., for generating input data to simulation and optimization models. Importantly, key parameters of the data generated can be adjusted by the user to obtain data with different properties, and hence different impacts on the supply chain. For example, the proportions of different vessel types, or their breakdown by terminal can be changed, or the smoothness of the demand arrival over time can be adjusted. Adjustable data, able to reflect future anticipated demand, is essential to organizations such as the HVCCC for their forward planning. The HVCCC uses such data to answer what-if questions about future infrastructure or operational plans, and in particular to understand what the limits of operational improvements are. These limits dictate how far delay to infrastructure expansion is possible, and hence what capital expenditure savings can be gained from investment in operational improvements. Exercising control over key features of demand is one important area in which operational improvements can be realized.

The HVCCC now maintains four different simulation models focussed on different aspects of the Port of Newcastle coal operation (one an overarching whole-of-coal-chain model [21]) to better understand system-wide effects, and to assess alternative congestion risk mitigation strategies [20]. They also use an optimization model to determine the best combinations of infrastructure expansion options to investigate. Without these tools, the HVCCC would not be able to achieve its target capacity. All these tools require shipping stems. The approach we describe in this paper enables such data to be generated efficiently and effectively.

We hope that the presentation of our efforts in this area, which we thought would be a matter of a few weeks of work, but which ended up taking more than a year, indicate not only the importance of research in this area, but that such research can also be interesting and rewarding.

Acknowledgment

We like to acknowledge the valuable contributions of Bhaswar Choudhury, Tracey Giles, Rob Oyston, and Dan Tengku from HVCCC. Without their patience, support, and feedback, it would have been impossible to develop the approach for shipping stem generation presented. We also thank the HVCCC who solely funded the initial stages of this work, and jointly, together with the Australian Research Council, funded the later stages under the ARC Linkage Grant no. LP0990739.

References

[1] E. van Asperen, R. Dekker, M. Polman, and H.D. Arons. "Allocation of ships in a port simulation". In *Proceedings of the 15th European Simulation Symposium 2003 - Simulation in Industry*, A. Verbraeck and V. Hlupic (Eds.), SCS European Publishing House, 551-557, 2003.

[2] V. Barichard, X. Gandibleux, and V. T'Kindt (Eds.) "Multiobjective Programming and Goal Programming. Theoretical Results and Practical Applications". *Lecture Notes in Economics and Mathematical Systems 618*, Springer, 2009.

[3] I.C. Bilegan, T.G. Crainic, and M. Gendreau. "Forecasting freight demand at intermodal terminals using neural networks - an integrated framework. *Sixth Triennial Symposium on Transportation Analysis (TRISTAN)*, Phuket, Thailand, 2007.

[4] N.L. Boland and M.W.P. Savelsbergh. "Optimizing the Hunter Valley Coal Chain". In *Managing Supply Disruptions*. H. Gurnani, A. Mehrotra, and S. Ray (eds.) Springer-Verlag London Ltd., 275-302, 2012.

[5] J.Y.J. Chow, C.H. Yang, and A.C. Regan. "State-of-the art of freight forecast modeling: lessons learned and the road ahead", *Transportation 37*, 1011-1030, 2010.

[6] M. Ehrgott. *Multicriteria Optimization*. Springer, 2005.

[7] R. Epstein, M. Goic, A. Weintraub, J. Catalan, P. Santibanez, R. Urrutia, R. Cancino, S. Gaete, A. Aguayo, and F. Caro. "Opimizing long-term production plans in underground and open pit copper mines." *Operations Research 60*, 4-17, 2012.

[8] M.M. Fioroni, L.A.G. Franzese, C.E. Zanin, J.A.S. Quintans, L.D. Pereira, I.R. de Santana, P. Savastano, S.S. Cordeiro, L.F. da Silva, and V.L.D. Benevides. "Matching production planning and ship arrival scheduling by simulation". In *Proceedings of the 2010 Winter Simulation Conference*, B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Ycesan (Eds.), IEEE, 1990-1997, 2010.

[9] R. Fourer, D.M. Gay, and B.W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press, Cole Publishing Company, 2002.

[10] N.G. Hall, and M.E. Posner. "The Generation of Experimental Data for Computational Testing in Optimization". In *Empirical Methods for the Analysis of Optimization Algorithms*, T. Bartz-Beielstein, M. Chiarandini, L. Paquete and M. Preuss (Eds.), Springer, 2010.

[11] D. Jagerman and T. Altiok. "Vessel arrival process and queueing in marine ports handling bulk materials", *Queueing Systems 45*, 223-243, 2003.

[12] D.F. Jones and M. Tamiz. *Practical Goal Programming*. Springer, 2010.

[13] N. Lang and A. Veenstra. "A quantitative analysis of container vessel arrival planning strategies", *OR SPECTRUM 32*, 477-499, 2010.

[14] C.H. Laih, B. Lin, and K.Y. Chen. "Effects of the optimal port queuing pricing on arrival decisions for container ships", *Applied Economics 39*, 1855-1865, 2007.

[15] F. Margot. "Symmetry in Integer Linear Programming". In *50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art*, M. Jnger, T.M. Liebling, D. Naddef, G.L. Nemhauser, W.R. Pulleyblank, G. Reinelt, G. Rinaldi, L.A. Wolsey (Eds.), Springer, 2010.

[16] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. John Wiley & Sons, Chichester, 1988.

[17] A. Newman, E. Rubio, R. Caro, A. Weintraub, K. Eurek. "A review of operations research in mine planning." *Interfaces 40*, 222-245, 2010.

[18] C. H. Reilly. "Synthetic optimization problem generation: Show us the correlations!", *INFORMS Journal on Computing 21*, 458467, 2009.

[19] G. Singh, D. Sier, A.T. Ernst, O. Gavriliouk, R. Oyston, T. Giles, and P. Welgama. "A mixed integer programming model for long term capacity expansion planning: A case study from the Hunter Valley Coal Chain", *European Journal of Operational Research 220*, 210-224, 2012.

[20] J. Vandervoort. "Hunter Valley Coal Chain Coordinator", *CEDA Annual Conference*, Newcastle, Australia, September 9th, 2010, available at `http://www.hvccc.com.au/Communications/MiscellaneousPresentations.aspx`.

[21] P.S. Welgama and R.Oyston. "Study of options to increase the throughput of the Hunter Valley coal chain". In *Proceedings of MODSIM 2003*, Townsville, July 2003, pp. 1841–1846.

[22] L.A. Wolsey. *Integer Programming*. John Wiley & Sons, Chichester, 1998.