This is an author produced version of a paper published in Computers in industry. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:

Woxenius, J., Persson, J. A., Davidsson, P.
Utilising more of the loading space in intermodal line trains – Measures and decision support
Computers in industry, 64 (2) s. 146-154
URL: http://dx.doi.org/10.1016/j.compind.2012.11.007

Access to the published version may require subscription. Published with permission from: Elsevier
Utilising more of the loading space in intermodal line trains – measures and decision support

Johan Woxenius\textsuperscript{1,3,1}, Jan A. Persson\textsuperscript{2,3}, Paul Davidsson\textsuperscript{2,3}

\textsuperscript{1} Department of Business Administration, University of Gothenburg, Box 610, SE-405 30 Gothenburg, Sweden
\textsuperscript{2} Department of Computer Science, Malmö University, SE-205 06 Malmö, Sweden
\textsuperscript{3} School of Computing, Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden

johan.woxenius@handels.gu.se; jan.a.persson@mah.se, paul.davidsson@bth.se

Abstract: The focus of this article is to identify, characterise and qualitatively evaluate the existing measures for increasing the utilisation of loading space in intermodal road-rail freight line trains. Normally, these trains operate like intercity passenger trains, because they make short unit load transhipment stops along rail corridors in order to travel comparatively small and short flows. The measures for loading space utilisation include adapting the train’s capacity, changing the departure times, altering the train routes and sending trucks to different terminals as well as replacing rail transport with trucks. Some of the measures require improved information sharing, and executives can strengthen the effect of the measures by adding decision support systems and price incentives to transport buyers.

Keywords: Booking system, intermodal freight transport, space utilisation, transport planning, yield management

1 Introduction

Intermodal road-rail freight transport systems seem a promising link between outlets and new markets with small and dispersed demand, because these systems have a production profile resembling intercity passenger trains with short stops along a line. Such systems might also help to recapture those markets abandoned by European intermodal operators when they shifted to direct shuttle trains, as the intermodal systems could offer transport services in a comparatively dense terminal network [1, 2]. The direct trains are competitive on the routes they address, but compared to networks offering service between many combinations of terminals, direct trains leave large areas without service.

Intermodal line train systems have been in commercial operation in Japan for decades. They have also been commercially tested in Sweden and are scheduled for implementation in Continental Europe. The Swedish tests showed that the profitability of such systems strongly depends on the average loading space utilisation along the route [3, 4], presuming that all shipments contribute sufficiently to the revenues. The stiff competition that the intermodal
system faces from all-road transport simply does not allow the companies to leave many empty wagons in the train sets.

The present study extends the author’s previous work and applies the findings of a literature review to identify the following measures for increasing loading space utilisation and revenue for line trains. According to these measures, the subjects for further study are as follows:

- adapting the trains’ capacity;
- adapting the departure timing;
- using trucks parallel to rail lines, i.e., replacing badly filled trains with road transport;
- adapting train routes, i.e., optimising the train routes dynamically to changes in demand;
- assigning terminals dynamically, i.e., optimising the assigned departure and arrival terminals for each load unit; and
- applying price incentives, i.e., motivating customers to fill empty slots by offering dynamic pricing.

In order to assess these measures, it is often necessary to improve the information sharing and exchange between the actors involved, including the transport buyers, freight forwarders, road hauliers, terminal operators and rail operators. It is also often necessary to develop and implement software-based decision support systems, such as support systems that rely on operations research methods or agent technology.

The purpose of this article is to identify, characterise and evaluate existing measures for increasing the space utilisation for freight in intermodal line train systems. The article will assess to what degree the measures are likely to fulfil the users’ demands for service quality and how the measures are likely to support the providers’ profitability. Moreover, it will discuss the scientific and commercial maturity of the measures as well as the intermodal system’s need for further development.

This study covers the full door-to-door transport system and focuses on the activities of transport ordering, transport planning, road haulage, transhipment and rail haulage. It uses the term intermodal transport consistently with the definitions presented by the European Conference of Ministers of Transport [5]. In short, an intermodal system must use at least two traffic modes for a single journey while keeping the goods within a unit load. The geographical context of this article is European, and, unless otherwise indicated, intermodal road-rail freight transport (IRRFT) refers to the combination of road and rail using the unit load type containers, swap bodies and potentially semi-trailers.

The article begins with a review of the transport capacity management’s decisions regarding loading space utilisation by examining these decisions on the tactical and operational levels. This review is followed by another review of how passenger transport (mainly intercity bus and rail services) has approached the problem of filling seats along multi-stop routes while generating satisfying revenue. This latter review should be of interest, because some of the passenger transport measures are applicable for filling the slots in intermodal line trains for freight. Yet the different logistics of passenger and freight transport also require this study to benchmark freight transport services before it can identify the best measures for intermodal line train systems. The final sections contain the core of the article, describing and assessing the measures that the article has found feasible for increasing the space utilisation in intermodal line train systems for freight.
2 Transport capacity management

Capacity management in transportation covers different technical resources with significantly different time scales of implementation and duration as well as associated cost structures. These resources are also controlled by a disparate set of actors. Infrastructure is an interesting example of a jointly controlled resource, because it requires decades to plan, finance and build. Yet when the infrastructure is in place, it will be used for decades if not centuries. Infrastructure as such is beyond the scope of this article, but it is important to note because access to this resource is a limiting factor for the execution of transport services. This section focuses on generic measures for increasing the space utilisation of freight transport, including a review of how the issue is dealt with in passenger and freight operations. It thus gives the context of the challenge of increasing the slot utilisation in line trains.

Capacity management is divided into strategic, tactical and operational planning with different magnitude of decisions and different time horizons. The terms are clearly inspired by military terminology. In this article, the sort order follows management practice rather than the sort order strategic, operational and tactical planning applied by the military. The focus is on tactical and operational planning.

2.1 Tactical capacity management

Managers generally address the dimensioning train capacity of transport operators within the medium term, and their plans are thus often referred to as tactical decision-making. Rolling stock capacity is, as it is in most industries, added incrementally. The importance of each decision depends upon the character of the demand, the size of the units, the size of the fleet and the marginal cost of buying and operating an excessive number of units or excessively large units. Adding a few lorries to a large fleet, for example, is a minor decision, but adding a new ship, plane or train departure to the timetable or adding a lorry for the owner-operator to use would involve significant, discrete steps. Some systems, like long-distance trucking, operate in fixed unit sizes. Operators of other systems decide both on number and size of resources, facing trade-offs between frequency and economies of scale according to a certain transport demand. A wide range of parameters, such as flow imbalances, network interdependencies, seasonal and business cycle peaks and long delivery times from vehicle suppliers and vessel suppliers, will complicate the dimensioning of capacity.

Capacity dimensioning has become ever more important to transport operators (see, for example, [6]) as generally meagre profits compels them to feel less willing to keep extra capacity as a backup. Shippers who work hard with lean production also require their transport providers to improve their operations. The lack of slack resources in the system, however, complicates operations, and operating close to full train capacity often implies increasing average costs [7]. In some segments, such as the mail and parcel services, general cargo forwarding [8], ferry operations [9] and public transport, even the users take capacity for granted. The operators are thus forced to over-dimension their systems, but they can, in turn, avoid complicated booking systems [10]. Having excess capacity may also facilitate late bookings at a premium price, but in general, excess capacity weakens the operator’s power when negotiating prices.

2.2 Operational capacity management

Decision making at the tactical level has resulted in a fixed overall train capacity available for operational planning. At this level, managers make their short-term decisions to minimise costs and maximise revenue from operating a system with a fixed set of resources. The time that operators need to allocate resources to a network in order to improve performance differs
across traffic modes. Marin and Salmeron [11], for instance, state that allocating existing resources to a rail network has a medium-term planning horizon; thus, this allocation requires tactical planning. Nevertheless, the planning horizon of rail track access has been decreasing. For example, Indra-Payoong et al. [12] have suggested a scheduling method for container trains with weekly changes in the timetable. In their study, they clearly wish to accelerate and otherwise improve rail transport planning, so they treat the design of a definite train plan as operational planning. Operational planning with a very short or even real-time horizon is another activity [13], which Cheng [14, 15] denotes as rescheduling.

Operational decisions predominantly address scheduling, matching consignments with transport resources and—to some extent—pricing and booking administration. By adopting the perspective of forwarders and shippers, Jarzemskis [16] and Jarzemskiene and Jarzemskis [17] suggest models for middle range short-term allotment planning for individual shipments in intermodal freight trains. The present study’s assumption of a fixed set of resources, however, relates to the full system, and transport systems are often hierarchically organised, so some actors in the hierarchy can treat capacity as a variable also on the operational level. Forwarders that irregularly contract independent, small hauliers constitute an example. These variations in the definition of capacity complicate this analysis of resource utilisation, since these hauliers might take on transport assignments independently or for other, larger transport systems when not occupied by the focal transport system. Therefore, one can generally regard them as not idle, making Advanced Planning and Scheduling systems less pertinent. These systems do allow managers to optimise several actors, but Hvolby and Steger-Jensen [18] explain that managers rarely use them, as each partner belongs to a number of supply chains. Such a network of partners would benefit from a plan for all their operations, not just parts of them.

Again, it is important to remember that there are some major differences between passenger transport and freight transport that affect operational capacity management. Although passenger transport can be very imbalanced in the short run (e.g., commuting peaks in the morning and afternoon), passengers tend to return home, and they are comparatively homogenous when they occupy one seat each. They also make individual decisions and produce parts of the transport service themselves. Freight, on the other hand, is monodirectional by definition. Freight consignments span from a single letter to 500,000 tonnes of crude oil. Usually, firm representatives purchase these consignments, and the transport of every consignment must be managed. Some of the significant differences between passenger and freight transport pertain to operational decision making. The challenges related to traffic mode, distance, time requirements and size of consignments usually differ between the two transport systems. Despite these fundamental differences, passenger and freight transport both face capacity management challenges related to intermodal line trains’ slot utilisation. Accordingly, the following two sections analyse how capacity is managed in passenger and freight transport.

### 2.3 Operational capacity management in passenger transport

Except for taxis, corporate jets and some full charter services (which are all disregarded here), passenger transport service plans typically aim to consolidate passengers who would otherwise make individual decisions about how and when to travel. These services therefore need a certain degree of compromise between particular demands. Most travellers are aware of the major features of the transport service they wish to use and have access to a relatively fixed and predictable timetable. The travellers then choose between traffic modes, operators and departures.
With a fixed set of resources and a published timetable, passenger transport operators must first allocate resources to different routes and departures in the network and then attempt to maximise revenue. In some cases, operators use a booking system for allocating resources, but many operators in the ferry, coach and public transport segments fully avoid booking systems. When demand exceeds capacity, the operators who do not use booking will add to the vehicle’s capacity, or the passengers must simply wait for the next departure.

Dynamic pricing is a major tool for increasing total revenues. The art of sales is basically to provide products and then persuade as many customers as possible to pay as much as possible for each item. It is difficult to mark up the price of physical products, since customers can easily compare them and they can be stored and resold. A common trick that retailers play is to sell the product at a high price in the beginning of the product’s life cycle and successively lower the price to entice more buyers, referred to as the price skimming strategy [19]. For services, the trick is the inverse: sellers increase the price for the same service over time. For example, the cost of a transport service may depend on when the customer desires to travel with that service. This art of understanding customer behaviour to maximise revenue is known as transport revenue management or yield management (for an overview, see [20]). One common denominator across any passenger transport service that chooses to apply yield management is that tickets are cheapest well in advance and very shortly before each departure. However, waiting to purchase a ticket shortly before departure does not guarantee a seat, and tickets with such guarantees are increasingly expensive.

Yield management can be successful only when customers can be segmented or at least cannot resell the service. It also requires that customers have fair but not full knowledge of the firm’s pricing strategy and its available capacity. Yield management does not work if customers cannot negotiate rational choices between offers. The number of tickets sold in each price category depends on the likelihood that the capacity will be fully used. Due to the comparatively low variable cost of adding a passenger to these systems, some operators offer monthly and seasonal tickets for frequent users, and such tickets do not add a marginal price whenever a customer uses the service.

2.4 Operational capacity management in freight transport

The mono-directional character of freight transport implies that transport providers should pay significant attention to evening out imbalances [21] and re-positioning their resources, which is why prices can differ significantly between the two directions of a single route. This difference is obvious in container shipping. Sometimes, firms pay no fee at all for sending empty containers to China.

Freight operators also have the option of routing resources dynamically after the demand. This option is rarely available to operators of collective passenger transport, perhaps with the exception of airport hotel shuttle services. A classic example of dynamic freight routing is tramp shipping, where ships are routed according to current demand. Another example of dynamic routing is the large road transport segment of part loads (i.e., loads that are too small to receive a full transport resource but too large for efficient terminal handling). Typically, one truck picks up five to ten loads from consignors in one area and delivers them to consignees in another area. The routing is planned differently each day depending on demand, but the driver is still subject to time window restrictions.

Some freight transport systems like mail, express parcel, less than truckload, wagonload rail and liner shipping are very complicated to design, but once they are in operation, they follow strict rules of execution. For large market segments, however, dispatchers maintain a fixer mentality and use experience, rules of thumb and the occasional decision support system for
operational capacity management. Full load transport, on the other hand, is rather simple to plan and execute.

Freight transport services are sold in a different way than passenger transport services are. The former generally trades through business-to-business relationships and includes substantial discounts for frequent users. The character of business relations between transport providers and transport buyers is diverse and, according to Wedel [7], they range from deep, multi-year relationships where a single transport provider is designated via frequent transport buyers having basic agreements based on tariffs with several transport providers that they use selectively, to occasional TBtransport buyers. According to Henriksson and Persson [22], only 26% of all Swedish transport buyers had contracted only one transport provider, while 53% had three or more contracts. Andersson and Normar [23] foresee further differentiation between types of transport provider-transport buyer relationships. Finally, it is important to remember that ordering transport services is still often characterised by routine [24]. The conditions for using yield management in freight transport thus differ from passenger transport.

3 Intermodal line train systems

In this section, the article narrows its scope to intermodal road-rail freight transport (IRRFT). This section characterises the line train production system in technical and administrative terms. Cardebring et al. [25] conducted a survey of European intermodal operators and found a wide range of production arrangements in use, but there is clear evidence for claiming that IRRFT is conventionally produced. Currently, the dominant production paradigm in IRRFT is night-leaps directly between large-scale transhipment terminals using gantry cranes and reach stackers [4]. According to Woxenius and Bärlak [26], IRRFT is still following a trend of abandoning true networks for even more direct trains between major conurbations and ports. In addition, Gouvernal and Daydou [27] found that the United Kingdom’s use of dedicated trains has increased dramatically over the last few years. Finally, North America is providing service to a decreasing number of terminals [28].

The advantages of direct trains are evident when flows are large enough for a satisfying frequency and long enough for the speed and low running costs of rail to compensate for the costs and time consumed by large-scale vertical handling. Yet operating with other network and terminal designs allows IRRFT to compete also at routes characterised by small volumes or short distances. This study regards short distance as shorter than the 500 kilometres, which is often scheduled within Europe [1, 29-31] and within Japan [32]. Short distance also refers to 500 miles (approx. 800 kilometres) in the USA [28]. A small volume refers to a volume that is less than economically viable for direct trains [33].

In order to compete with all-road transport in markets with small volumes or short distances, firms regularly advocate for line trains, also referred to as corridor trains [1, 2, 4, 31, 34-38]. Combining the offer of transport between pairs of relatively closely located terminals with the one between end points, means that geographical markets not accessible for conventional IRRFT requiring a comparatively large transport demand between two terminals [1]. The corridor traffic design is not a new invention; it is well established in passenger intercity services and, as mentioned, it is also used for IRRFT in Japan. Nevertheless, European and US rail freight services rarely use this design.

Designers apply the corridor concept to IRRFT by using line-trains that stop 15 to 30 minutes for transhipment approximately every 100 kilometres. This study’s interpretation of a corridor follows what is outlined by Woxenius et al. [38], but it differs from Rizzoli et al.’s [39, p. 61]
design, as they define it as a “privileged point-to-point railway connection between two terminals.”

Operators of line train services explore the borderland of competition between road and rail, since they aim for the overnight transport market of semi-finished and manufactured goods at the relatively short distances of 200-500 kilometres. This is definitely the home ground of road transport, and some argue that politicians and environmentalists are unrealistic when asking for significantly more intermodal freight transport. During the day, operators move containers to ports, transport less time-sensitive commodities and empty unit loads since the otherwise common demand for night-time transport does not apply. Serving a larger number of terminals than direct trains implies less economies of scale in rail and terminal operations, as elaborated by Ballis and Golias [40] but on a system level, it facilitates economies of scope, as elaborated by Jara-Diaz and Basso [41].

The intermediate terminals must be efficient. When transhipment consumes too much time and money, the total lead time and price is not attractive to transport buyers, as Bontekoning and Kreutzberger [42] explained. Improving inefficient terminals requires substantial improvement to their cost-quality ratio; incremental improvement of conventional terminals and related shunting operations would not suffice. Side track terminals are more promising, because unit loads can be transhipped under the overhead contact wire.

The three-year commercial test of the Swedish Light-combi concept showed that the maximum train speed of 110 km/h sufficed for covering a distance of 650 km overnight, including four intermediate transhipment stops. It also showed that the load plan is crucial for decreasing the time spent at terminals, because load plans shorten the handling equipment’s driving distances. Overall, the technology and logistics of the Swedish Light-combi concept worked smoothly, but since the system was used only for distribution from a central warehouse to shops spread around Sweden, the trains suffered from route imbalances [3]. Hence, trains were decreasingly filled with loaded swap bodies during the routes.

The distance and time consumption [43] of pre and post-haulage (PPH) by road is crucial when the rail haulage distance is smaller. Both Bukold [34] and Rutten [29] show that the break-even distance of IRRFT is most sensitive to PPH costs when it competes with all-road, since PPH costs often account for 50% of the costs incurred in an IRRFT chain [44]. Costs particularly depend on the rate of empty hauls and the daily number of hauls performed by the hauliers, as specifically shown by Nierat [45]. Arnold et al. [46], however, found that the location of terminals and thus the PPH distance had little or even no impact on the IRRFT’s market share, but their results come from a study of long border-crossing distances through the Iberian Peninsula.

Transport corridor systems require much tighter co-operation with hauliers, especially in cases when lorries are part of the transhipment technology or when unmanned terminals are used. In addition, PPHs are fewer and shorter in transport corridor systems than they are in conventional IRRFT, compelling the corridor system hauliers to engage in other businesses besides PPH. The maximum local road haulage distance is considered 50 km, but when the flows fluctuate, lorries can be used for longer haulage in order to avoid stops where only a few unit loads are shifted.

The role of the former national, in terms of ownership as well as geographical coverage, railway administrations is diminishing in European IRRFT. Although they remain significant operators and owners of intermodal operators, they are now challenged by ports, shipping lines, logistics service providers and specialised new entrants [26, 47]. They are now far from being the only potential operators of line train systems. Moreover, it is difficult to predict the types of organisations that a commercial implementation would include and the scope of the
services produced by the system. One extreme is a transport provider integrating PPH, terminal handling and rail haulage for transporting consignments between the doors of consignors and consignees. Another extreme is a specialised rail haulage company that offers a trunk line train service with intermediary stops according to a timetable during which hauliers or transport buyers themselves can transship unit loads if there are empty slots in the train.

4 Measures for increasing loading space utilisation and revenue in line train systems

The focus of this section is to scrutinise the measures for capacity management of line train systems for unit loads that are available at the operational level. Hence, sections 2 and 3 are combined here. Regarding decision-making, the system under study is seen as a discrete event logistics system. Although the flow of unit loads is continuous at some stages, the managers make decisions about discrete events, and their actions fulfil Mönch et al.’s [48, p. 558] definition of such a system.

Admittedly, reducing average costs and maximising revenue by a high load factor on the trains is only a secondary goal; the primary goal from the operators’ perspective is to maximise the business economic profit. Nevertheless, with comparatively low marginal costs in rail operations, producing more transport services does normally improve the economic result of the operations. With a proper understanding of the marginal costs and the character of the demand, managers can make wise operational decisions that will improve profitability.

The organisational structure and the scope of offered services obviously affect the nature of operational decision-making regarding slot utilisation. Yet to avoid the problems with system boundaries and structure for the remainder of the article, this study will treat the line train system as a single entity that encompasses a set of roles or functions. The system producing the service can be regarded as having medium complexity that is rather far from either ends of the complexity scale designed by Mönch et al. [48, p. 558]. That is, the system in question sits right in the middle of a spectrum that ranges from a single warehouse to a global supply network. This medium complexity also avoids problems associated with centralised vs. decentralised decision-making, which were investigated by Newman and Yano [49]. The assumption that all roles aim at improving the system performance is admittedly a simplification, since there is often an element of opportunistic behaviour in a serially coupled production system.

The main system functions are defined as PPH, rail haulage and transhipment, and the roles are transport buyer, consignor, consignee, truck driver, train dispatcher and train engine driver. Note that the terminals are assumed to be unmanned, since dedicated personnel at terminals with few transhipments would be too costly. Thus, either the truck driver or the rail engine driver has to execute the transhipment function. Regardless, the transhipments need to be fast enough not to prolong the overall transport time. There is often a role overlap between transport buyer, consignor and consignee, but a third party can also act as transport buyer towards the line train system. Of particular interest are the roles that are authorised to decide which service level to offer and which price to accept.

4.1 Adapting train capacity

The first measure that addresses loading space utilisation is obviously to adapt the rail haulage capacity to the actual demand or at least the estimated demand. This measure regards the frequency of departures and the number of wagons in each train, as modelled by Nozick and Morlok [50], and it does not regard the track and terminal capacity, as elaborated by Kozan [51], since the latter capacity is regarded as fixed at the operational decision level.

8
The first option for increasing loading space utilisation is to wait until the train is fully loaded before departing, which is a similar strategy to the operation of dolmus taxis in Istanbul or small car ferries crossing rivers and narrow straits. This option is, however, rarely accepted by transport buyers using IRRFT, and it also implies a risk of sub-optimal utilisation of the operator’s resources and the rail infrastructure.

The second option is to depart according to a timetable while using a number of trains below demand to create a shortage. This is not a good option, either, since queues at terminal gates tend to disappear when transport buyers are disappointed with the service [12, 52]. Operating below demand implies lost revenues, and it presents the immediate risk of losing the transport buyers with the highest willingness to pay. Given their budget, these buyers are likely to be the first to abandon the service for all-road transport.

The third option is to operate according to a fixed frequency while varying the number of wagons (i.e., unit load slots) that are available in each train. It is important to remember that the marginal hauling costs relating to one empty rail wagon are very low compared to the operational costs of forming trains with shunting and break-tests as well as administrative planning costs. A better motive for splitting trains between departures is to make more efficient use of a fixed number of available wagons in a network. This procedure was previously applied by the Scandinavian intermodal operator CargoNet. Since they demanded booking information from the transport buyers as late as on the day of departure, they had to reposition wagons in their network depending on prognoses rather than on bookings. After a sincere effort to improve the prognosis tool [53], they decided in 2006 to operate only shuttle trains (i.e., trains with a fixed number of wagons operating between two terminals) with a multi-purpose wagon type [54].

There is, nevertheless, a clear trend in European IRRFT to operate shuttles rather than dynamically composed trains and networks [33]. As mentioned above, decisions about capacity serve as input to the operational decision levels investigated below. Therefore, the discussion of capacity sizing will be revisited, mainly in section 4.9.

4.2 Adapting departure timing

Although time usually scores high as a parameter for the choice between road and IRRFT [52, 55], timing is often more important than speed for transporting goods with relatively high value over shorter distances. The basic principle is still production and sales in the daytime and transport overnight. Examples of transport services executed during the daytime include full loads in just-in-time schedules, distribution from local warehouses and documents with couriers. Neither the processing industry nor the other systems operating continuously adhere to the night-leap production cycle, but most other assignments are overnight.

The definition of the line train market mentioned above implies that a large share of the transported goods belongs to strictly time-coordinated supply chains. Sommar and Woxenius [55] demonstrated that if the departure times of trains do not fit the supply chain demands, then the transport buyers simply use trucking instead. Developing timetables based upon efficient time utilisation of rail rolling stock, such as those suggested by Woxenius et al. [38], would face difficulties if implemented. For example, CargoNet had to discontinue a day-train between Gothenburg and Stockholm due to lack of demand for the mid-day departure, despite the fact that it transported maritime containers that generally do not apply to the overnight rhythm. For the same IRRFT segment, Indra-Payoong et al. [12] focused on schedule designs that minimise operational costs by matching departure times with a given demand. Consequently, timing departures according to demand is a major measure for attracting loads to the trains.
4.3 Using trucks parallel to rail lines

The former measure assumes that all long haulage will be performed by rail. However, there is no end in itself for a transport provider to maximise the use of trains if certain unit loads are more rationally transported by road. Situations in which this concern is relevant include a very low general filling grade of a train, a severe imbalance or a sudden technical problem. One example of a specific situation would be a few surplus unit loads remaining at a terminal: trucks could handle these loads and be the substitute for a second train departure. Another conceivable operational decision is to truck a unit load to avoid a stop and thus save total transport time. Furthermore, a firm may use trucks more regularly when it wishes to build up flows to the scale where trains would be economically feasible or when diminishing flows no longer justify train operations. However, these lattermost scenarios involve the tactical level of dimensioning the system. One example is that trucking was seen as an integrated tool in the advanced implementation plan of the Light-Combi concept in Sweden [3, 4].

4.4 Adapting train routes

By using good information exchange and a booking situation for a planning period, the train dispatcher can apply the traffic principle of dynamic routes [33] when this dispatcher has alternative paths to use through the network. The dispatcher could route trains differently for each departure to maximise the utilisation of train resources or to maximise revenue. This scenario is, however, different from the routing problems modelled by Barnhart and Ratliff [56], who compared road and rail haulage of semi-trailers. It also differs from the study by Newman and Yano [28], who allocated a set of containers to a set of available intermodal train services. Moreover, Jarzemski [16] and Jarzemskiene and Jarzemskis [17] studied allotment booking in intermodal shuttle trains.

This measure copies the production profile of part load trucking or buses routed from an airport to hotels depending on the actual demand for travel. It can also imply that the dispatcher avoids stops with no demand for transhipment and adopts the timetable accordingly in order to shorten the total transport time.

It is crucial to note that dynamic routing requires easy access to track infrastructure via either general slack capacity or dummy slots booked in advance. Track access plans are currently updated about twice a year, and the dynamic allocation of slots is complicated by the increasing number of operators and the fact that passenger trains would still require fixed timetables and would still be sensitive to disturbances. In the long run, however, daily track access schemes seem realistic.

4.5 Assigning terminals dynamically

When dispatchers fine-tune the demand and supply of slots in a line train system, they may encounter situations with a short demand overlap, such as a capacity bottleneck between only two stops. Such a slight over-booking on intercity trains is possible when a surplus passenger might stand for a few stops until a seat is free. Over-booking is obviously not possible in a train with a fixed number of slots. The situation can be solved either by denying transport of any of the overlapping unit loads or by re-routing a truck to a farther terminal. The latter measure presumes a complete and two-way information flow, and it might also require price-incentives for the transport buyer whose unit load has to be routed to the second-best located terminal. In an integrated system, though, the cost can be absorbed as long as the marginal cost, including the change in terminal, is lower than the marginal revenue.
4.6 Applying price incentives

From the transport provider’s perspective, the purpose of applying price incentives is to maximise revenue, that is, to apply yield management. Yield management has been explored extensively in the airline industry [20] and to a lesser extent in train operations [57]. The literature on yield management freight rail transport appears meagre. The line-based IRRFT service discussed here satisfies the following characteristics of yield management: buyers cannot easily resell the service, and the opportunities for transport buyers to buy service on a speculative basis are limited. Furthermore, in order for price incentives to have an effect, some transport buyers must be price sensitive. Compared to passengers, transport buyers are generally less price-sensitive and are more frequently bounded by long-term contracts, but at least some transport buyers might consider changing their behaviour in line with price incentives.

It should be noted that the transport provider offers transport services between a number of combinations of origin and destinations along a rail line, which makes the problem different from yield management on a direct route, as is often the case in air transport. Since the load space utilisation along the line typically varies significantly, the pricing between specific pairs of terminals is the most obvious basis for price incentives. It is suitable to increase prices for services between particular terminals where there is a significant risk of shortage of capacity and, correspondingly, decrease prices to attract transport demand between terminals with surplus capacity. This may drive some customers away from the service, but hopefully enough customers would consider another terminal for loading and unloading. See section 4.4 for details about adapting train routes and section 4.5 for details about the dynamic assignment of terminals, which both can be interconnected to price incentives.

Another base for price incentives is the conditions of the service regarding the precisions on the arrival time, i.e. a wider time window, or to offer a special category for stand-by shipments. This allows the rail operator to store some unit loads, which can be moved by alternative departures and thereby increase the total load utilisation. It can also be used for allocating less total capacity (e.g. fewer wagons) and still have the same service level (here defined as the probability of being able to offer the customer a desired service).

A price incentive based on the point of time of sale (i.e., an agreement) on the transport is suitable, implying that the incentive is dynamic. Typical factors to consider when deciding the price are the competition, the time remaining to departure, current load utilisation and planned or forecasted load utilisation. The purpose of considering current load utilisation in relation to forecasted utilisation levels is to be able to capitalise on the effects of price incentives correctly and in a timely manner. The dynamic pricing could be further advanced by including modelling tools for computing the opportunity costs such as input for pricing, which Yan et al. [58] had calculated.

4.7 Improving information sharing

Although rules for order time and general distribution of obligations between the system roles are established above the operational level, they obviously influence the improvement potential at the operative level. To a large extent, improvements to information sharing are a prerequisite for several other measures, but it is also an operative measure in its own right.

The availability of demand information is a typical issue here. Newman and Yano [49] demonstrated how knowledge of the demand facilitates centralized decision-making, better use of resources and, ultimately, reduced costs. This outcome is not surprising, as many studies suggest sharing information transparently as a more or less universal solution to supply chain problems, such as the investigation by Hvolby and Trienekens [59]. Information
transparency is particularly effective for smaller firms [60]. Insufficient information sharing will always present a major bottleneck in streamlining transport services.

Yet there is hope, since information about the expected movements is often available in the time-coordinated supply chains well before the line train departures. The problem is that the information is rarely forwarded to the transport provider if doing so is not stipulated in the contract, which significantly reduces the provider’s ability to make improvements [61]. Managers at the operational level thus need to tell higher level managers that improving operations require that that conditions need to be renegotiated.

4.8 Applying decision support systems

In order to realise the measures identified in sections 4.1-4.6, information must be available, especially for the transport provider, in order to develop efficient plans. It is also important to make the information available to the transport buyers. For instance, providers can communicate price incentives to buyers, which was elaborated in the previous section. The important role of the decision support system in the communication context is to make information available to decision-makers in a structured manner and thus facilitate the optimisation of a plan across several partners in the supply chain. Hvolby and Steger-Jensen [18] investigated the benefits of a decision support system in an Advanced Planning and Scheduling context.

Normative techniques, such as optimisation encapsulated in decision support systems, are potentially very supportive. These techniques can support a number of measures, including decisions on train capacity, departure times, allocation of parallel trucks, train routes and terminal assignments. Many of these decisions are made under uncertainty, but they are based on some forecasts of the uncertainty in the demand. A common practice is to consider the decision problems as deterministic while including some implicit considerations of the uncertainty, such as considering the estimates of additional transport buyers.

Many of the measures above interact with each other in a rather complex way, such as price incentives with capacity allocation or time tables with demand adaptation [12]. Therefore, it is important to study carefully the combinations of policies related to these measures and not just the policies themselves. The complexity of this problem may not seem overwhelming with computerised support. To develop suitable decision support systems, the present study suggests using mathematical optimisation (for applications in the rail sector, see [62]) and agent technology (for applications in the field of transport, see [63]) in order to study these measures and their interactions.

Agent technology could be used both for strategic and operational decisions. For the latter, it could be used to prioritise goods, such as choosing which goods should be left waiting for the next train when a train is full. Each goods unit could then be represented by a software agent that knows the utility function of the goods, that is, how important is it that the goods catch the current train. The agents then negotiate to determine which buyer should suffer the least by waiting, perhaps by using some market-based approach like a computational auction [64]. The agent can even be integrated in the goods itself, as described by Jevinger et al. [65]. The provider then gets something similar to yield management, but the customers are software agents rather than human beings in this case. For strategic decisions, agent-based simulations can be used where the different decisions in the transport chain are modelled as software agents (see [66]). By using this type of micro-level simulation, a transport provider can study potential policies, including timetable or pricing policies, through what-if scenarios. The provider would play out these scenarios by inputting different system assumptions, such as transport buyer preferences. Whelan and Johnson [67] conducted a study that was similar to
these what-if scenarios, focusing on passenger traffic and overcrowding. These simulations could help providers to decide the parameter values of the policy rules and estimates that they tested in the optimisation models.

5 Characteristics of measures

In the table below, the measures are characterised with respect to their potential effect on the transport buyer and the transport provider. The table also proposes combinations for these measures and in what time frame each measure would contribute to increased system performance. As pointed out above, the interactions are complex and require further studies in their specific context. To support these future studies, this table presents hypotheses of which measures would make efficient combinations.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Involved</th>
<th>Proposed combinatio</th>
<th>Time frame</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TB</td>
<td>TP</td>
<td>Short</td>
</tr>
<tr>
<td>1. Adapt train capacity</td>
<td>(X) X</td>
<td>2,7,8</td>
<td>X X X</td>
</tr>
<tr>
<td>2. Adapt departure timing</td>
<td>X X</td>
<td>1,7,8</td>
<td>X</td>
</tr>
<tr>
<td>3. Use trucks parallel to rail lines</td>
<td>X X</td>
<td>5,7,8</td>
<td>X X</td>
</tr>
<tr>
<td>4. Adapt train routes</td>
<td>X X</td>
<td>3,7,8</td>
<td>X</td>
</tr>
<tr>
<td>5. Assign terminals dynamically</td>
<td>X X</td>
<td>3,6,7,8</td>
<td>X</td>
</tr>
<tr>
<td>6. Apply price incentives</td>
<td>X X</td>
<td>5,7,8</td>
<td>X</td>
</tr>
<tr>
<td>7. Improve information sharing</td>
<td>X X</td>
<td>All</td>
<td>X X X</td>
</tr>
<tr>
<td>8. Apply decision support systems</td>
<td>X X</td>
<td>All</td>
<td>X X X</td>
</tr>
</tbody>
</table>

TB=Transport buyer, TP=Transport provider

The measures that the transport provider controls (i.e., measures 3 and 4 in Table 1) are obviously the easiest to apply. The other measures affect the transport buyer to varying extents. Regarding the latter measures, the actors may mutually decide upon some changes, whereas other changes might have to be regulated in the contract or as a general agreement so that the transport provider keeps the authority to make changes within certain limits.

Each measure is believed to contribute to the efficiency of the line train system, but Table 1 demonstrates that they are preferably applied in combinations. The complexity of a change depends on how many measures the provider applies, and relatively complex changes are likely to necessitate a decision support system before the provider can hope to control the system or evaluate the efficiency of different solutions.

Table 1 also shows that all measures defined and analysed here apply to the short-term management of line train systems. Measures 1, 7 and 8, however, are not decisions made exclusively by operational management, since they also affect performance in the medium and long terms. Capacity sizing in particular is a matter of continuous attention at all management levels.

6 Implications and conclusions

There is a complex relationship between the decisions affecting the load space utilisation in IRRFT along corridors. These decisions include strategic business positioning, tactical resource dimensioning and operative short-term resource allocation such as routing, timing and engaging additional truck capacity. Some of the operational measures presented here only
relate to the core of the system, which is the operation of the line trains. Other measures relate to the execution of PPH, and still others require the transport buyers’ involvement.

For the lattermost type of measure, which promises the best resource utilisation, the price incentive schemes affecting the transport buyer’s behaviour are of most interest. Often, the providers use price incentives in a one-way manner with price offers based on fixed departures, each with a fixed number of slots. More advanced combinations are, however, feasible for line train systems with a limited number of transport buyers, along with possibilities for changing the departure time and route of a train with relatively short notice. With a good two-way information exchange, an iterative process can be established to improve resource utilisation even further. The providers can test transport buyers’ attitudes towards the different measures in the toolbox and predict a satisfying load utilisation before enacting any measures. The general priorities of a limited set of regular transport buyers can also be captured in a multi-agent-based simulation application. The transport provider can then either avoid some iteration of manual transport buyer feedback or at least test alternatives in the simulation tool before suggesting a change to the transport buyers.

The implications of the work presented in this article mainly pertain to prospective operators of IRRFT systems based on line trains. Such systems address a very difficult segment of IRRFT, but the rewards might be comparatively large for the successful operator. The main reason the rewards can be so large is that European transport work in the 150-500-kilometre transport segment is about twice the size of transport work in the segment above 500 kilometres. It would be difficult to implement a successful system given the economies of scale and scope and the commercial risks involved. Therefore, the first to implement a successful system is not likely to be easily challenged. The cherry pickers would rather keep focusing on the direct shuttles for many years to come.

Still, the competition with all-road transport is fierce and direct trains will dominate intermodal transport on many geographic markets. Some of the measures also apply to conventional IRRFT with direct trains between large-scale terminals, while other measures might attract attention from line-based passenger transport, although that application seems more commercially developed than freight transport.

The implementation of line-based IRRFT requires substantial effort and a number of careful decisions. Line-based IRRFT has faced a certain lack of success so far in Europe. Success is believed to depend on a development of the system in discrete steps in order to attract demand, which stresses the need for continuous implementation and re-evaluation of measures and their associated analyses.

References


