

# TAPAS: A multi-agent-based model for simulation of transport chains

Johan Holmgren<sup>1</sup>, Paul Davidsson<sup>1,2</sup>, Jan A. Persson<sup>2</sup>, Linda Ramstedt<sup>3</sup>

<sup>1</sup>*School of Computing, Blekinge Institute of Technology, Box 214, SE-374 24 Karlshamn, Sweden*

<sup>2</sup>*School of Technology, Malmö University, Malmö, Sweden*

<sup>3</sup>*Vectura, Solna, Sweden*

---

## Abstract

We present the Transportation And Production Agent-based Simulator (TAPAS), which is an agent-based model for simulation of transport chains that can be used, e.g., for analysis of transport-related policy and infrastructure measures. TAPAS is more powerful than traditional approaches to freight transport analysis, as it explicitly models production and customer demand, and it captures the interaction between individual transport chain actors, their heterogeneity and decision making processes, as well as time aspects. Whereas traditional approaches rely on assumed statistical correlation, TAPAS relies on causality, i.e., the focus is on the decisions and negotiations that lead to activities. TAPAS is composed of two connected layers, one that simulates the physical activities, e.g., production and transportation, and one that simulates the decision making and interaction between actors. We illustrate TAPAS with a scenario in which the consequences of three transport policy and infrastructure measures are studied.

*Keywords:* Multi-agent-based simulation, Multi-agent systems, Transport chain simulation, Freight transportation, Decision-making

---

## 1 Introduction

Freight transportation causes different types of positive and negative effects on the society. Positive effects typically relate to economy and social welfare, e.g., due to the possibility to consume products that have been produced far away. Negative effects mainly relate to the environment, and typical examples are emissions, congestion and energy use. By applying different types of governmental control policies, infrastructure investments and strategic business strategies, it is sometimes possible to influence how transportation and other activities in transport chains are performed. Control policies include different types of taxes and fees, such as kilometer and fuel taxes, and regulations, such as weight restrictions on vehicles. Infrastructure investments can be made into roads, railway tracks, intermodal freight terminals, industry tracks, etc., and examples of strategic business measures are improvement of timetables and adjustment of vehicle fleets to better meet the demand for transport. Public authorities, in the role of policy makers, often have a wish to reach certain governmental goals, such as obtaining sustainable transport systems and meeting emission targets. A typical ambition of a public authority is to increase the internalization of external costs, e.g., by letting road users pay for the road wear they cause (Button, 1997). However, internalization of external costs may have effects that might be negative on other goals. For instance, it might lead to negative economic development in a region. For enterprises, the goal is typically to maximize profit, e.g., through optimization of their activities (either individually or in collaboration), by reducing lead-times, lowering transport costs, improving delivery accuracy, etc. To avoid undesired consequences and to confirm desired consequences it is important to be able to accurately predict what the consequences will be when applying different types of measures.

Multi-Agent-Based Simulation (MABS), in which one or more of the simulated entities are modeled as agents, can be used for studying freight transport systems. An agent can be viewed as a system that is situated in some environment and capable of autonomous action in that environment in order to meet its design objectives (Wooldridge and Jennings, 1995). We present the Transportation And Production Agent-based Simulator (TAPAS), which is a micro-level model for quantitative impact assessment of, e.g., different types of transport-related policy and infrastructure measures. In TAPAS, which is based on agent technology, the individual actors of a transport chain, e.g., producers, transport operators and customers, are explicitly modeled. TAPAS incorporates the complexity of transport choices, e.g., with respect to consignment size, and route and transport mode. In addition, TAPAS is able to take into account time aspects, such as, the effects of timetables, arrival times, and time-differentiated taxes and fees. This makes TAPAS more powerful than traditional approaches to freight transport simulation as it is able to capture the individual goals of transport chain actors and the interaction between them, as well as their heterogeneity and decision making. Whereas traditional approaches rely on assumed statistical correlation between different parameters, TAPAS relies on causality, i.e., the decisions and negotiations that lead to activities are modeled.

The main application area for TAPAS is to provide decision support to public policy makers by enabling analysis of, e.g., different types of transport-related policy and infrastructure measures. Other possible applications of TAPAS include assisting companies in making tactical and operational decisions, e.g., concerning choice of consignment size and frequency of deliveries, as well as strategic decisions, e.g., concerning choice of producer, adaptation of vehicle fleets, and location of storages and terminals. Moreover, we argue that TAPAS may complement existing approaches (e.g., on the macro-level) in different ways, e.g., by generating transport demand that may serve as input data to other models.

The research concerning TAPAS has been going on for several years, and the work has rendered a number of conference publications (e.g., Bergkvist et al., 2005; Holmgren et al., 2007; Ramstedt et al., 2007; Davidsson et al., 2008). This paper extends the previous work by providing a more detailed technical description and better motivation of the design of the TAPAS simulation model. Also, TAPAS has been extended with new functionality, e.g., load consolidation for demand driven transports has been added, and the model for setting transport prices has been improved.

In the next section we discuss some related work, followed in Section 3 by a detailed description and motivation of the TAPAS simulation model, including an account to the implementation of TAPAS. In Section 4 we illustrate some important functionalities provided by TAPAS by presenting a simulation study that has been conducted using TAPAS. Finally, in Section 5 and Section 6, we conclude the paper with a discussion and some suggestions for future work.

## 2 Related work

In this paper we present TAPAS, which is an agent-based simulation model that can be used, e.g., for impact assessment of transport-related policy and infrastructure measures and evaluation of different types of strategic business strategies. TAPAS is best described as a transport chain simulator, which has been extended with functionality for simulating production. In TAPAS, transportation is modeled in more detail than production, and it only considers a single-tier of the supply chain. Therefore, TAPAS mainly relates to the transport domain, and only partially to the more production-focused supply chain domain. Hence, the work behind TAPAS can be regarded as multi-disciplinary, and we will in this section cover several domains. We focus on research related to freight transport analysis and we also pay some attention to research related to supply chain modeling, focusing on agent-based supply chain models. It should be mentioned that the focus of this section is to give an overview of relevant work, not to conduct a complete literature review since reviews in relevant domains already exist. For instance, Terzi and Cavalieri (2004) contributed with a review focusing on parallel and distributed supply chain simulation. In the transport modeling domain, several reviews exist (e.g., Chow et al., 2010; de Jong et al., 2004; Tavasszy, 2006), and Friedrich and Liedtke (2009) provided a review of current transport modeling literature focusing on two agent-based models.

TAPAS simulates freight transportation, where the focus is on the freight flows and its allocation to vehicles, and models with this focus is often called commodity flow based models. Another related type of model is tour-based models, which instead focus on vehicle movements, i.e., the focus is on tours in which vehicles often start and end in the same location. An example of a tour-based model is the Calgary model presented by Hunt and Stefan (2007). Tour-based models are more often used when the geographical focus is regional or urban, while commodity flow based models often have a national or international focus.

Traditionally, freight transportation has been studied using macro-level (or equation-based) models, such as Samgods (Swahn, 2001), ASTRA (Shade et al., 1999) and SISD (Williams and Raha, 2002). Models of this type take a societal perspective and are based on aggregated coarse-grained data on national or regional levels, and different countries typically use their own models. Moreover, such models are commonly based on the aggregate and sequential four-step approach. Models based on this approach include one or several sub-models that can be described using four steps; production & attraction (trip generation), trip distribution, modal choice, and aggregation (de Jong et al., 2004). Models based on the four-step approach can look rather different, and different steps are typically represented by different types of sub-models, such as system dynamics models (e.g., ASTRA (Shade et al., 1999)), and economic models such as I/O models (e.g., Samgods (Swahn, 2001)) and gravity models (e.g., NEAC (Chen and Tardieu, 2000)). An overview of models based on the four-step approach is provided in (de Jong et al., 2004). Further, sometimes more than one step is represented within the same sub-model, such as in Samgods (Swahn, 2001), in which modal choice and assignment is integrated. An issue with traditional aggregate sequential models is that they do not take the logistical processes into account, e.g., choices of carrier type and ordering strategies, and thus fail to model the level where decisions regarding the actual transports take place. Starting about a decade ago, several aggregate sequential models, which also take logistical aspects into consideration, have been developed. Such models are for example SLAM (Williams and Raha, 2002), SMILE (Tavasszy et al., 1998), GoodTrip (Boerkamps and van Binsbergen, 1999), and the model suggested by de Jong and Ben-Akiva (2007), and they typically make use of disaggregation of aggregate data. However, since these models do not represent individual transport chain actors, the complex interactions (e.g., negotiation) and the decision making processes from which transport plans result are impossible to capture. Moreover, these models can only capture average times, but typically no aspects related to dynamic time (e.g., timetabled transportation), which is crucial when

logistical decisions are coordinated. The decision making in supply chains is subject to both short-term and long-term planning implying that the time dimension of the decisions needs to be considered.

We argue that more precise predictions regarding the effects of transport-related policy and infrastructure measures, and strategic business measures, can be achieved by using micro-level, and in particular agent-based models. Such models enable to also capture the decision making of actors involved in logistical processes. This is important for the ability to capture the causality from which transport activities appear. Agent-based models are often also designed in a way that they explicitly model time aspects.

Since TAPAS is an agent-based model for transport chain simulation, we will in the rest of this section discuss some agent-based models for supply chain and freight transport analysis. As an example, an agent-based supply chain modeling framework with the purpose of assisting managers to analyze and understand the impact (in terms of costs, risks, etc.) of different decisions was presented in (Swaminathan et al., 1998). The framework made use of a library of reusable software components to promote reusability, and it was used for analyzing different inventory policies with the purpose of achieving improved inventory management. The similarity TAPAS has with supply chain models like the one in (Swaminathan et al., 1998) is that it is built around similar structural components, however, the application area of TAPAS is similar to that of traditional freight transport models. In supply chain models, common structural components are activities/processes (such as manufacturing and assembling) and actors/roles (such as manufacturer and customer). Strader et al. (1998) presented a multi-agent-based simulation tool, which was used to show the importance of information technology for supporting the order fulfillment process. This was achieved by studying a 5-tier supply chain network under different demand management and information sharing policies. In a simulation model by Gjerdrum et al. (2001), agent technology and mathematical optimization were combined in order to study different inventory replenishment policies. Mathematical optimization was used to generate locally optimal manufacturing plans, and agents (e.g., customer, warehouse, factory and transport agents) were used to manage the supply chain. In comparison to our approach, the model by Gjerdrum et al. models manufacturing on a more detailed level, but transportation is not modeled on a detail that allows transport-related policies to be studied. Liedtke (2009) presented an agent-based approach to commodity modeling (INTERLOG), which is relevant for our work since its aim is to study the same types of questions as can be studied with TAPAS. The major difference between the models concerns which actors dominate supply chains; in TAPAS the customer is dominating, whereas in INTERLOG the supply chain is dominated by the producers and the transport operators. Moreover, in contrast to TAPAS, INTERLOG assumes a pre-determined, however stochastically generated, transport demand. This makes it impossible to capture changes in transport demand, e.g., caused by changes in logistical structures. Moreover, the literature contains a number of newer agent-based simulation approaches (van der Zee and van der Vorst, 2005; Chatfield et al., 2007; Roorda et al., 2010), which are presented as re-usable non-implemented conceptual frameworks that can be used as guidance when constructing agent-based simulation models. While these models are more on a conceptual level, TAPAS is more concrete as it is an implemented model.

In summary, the literature contains a number of agent-based models for supply chain and freight transport analysis, however most of them focus on simulating already established chains, while TAPAS partially constructs the transport chain, i.e., product supplier, route and transport operators are dynamically chosen. This is an advantage, since it enabled us to base the TAPAS model on supply and demand for products instead of transport demand, which is the case in most existing freight transport models.

### 3 The simulation model

In this section we present the TAPAS simulation model, which has a two-tier architecture with a *physical simulator* and a *decision making simulator*, as illustrated in Fig. 1. In the physical simulator, all physical entities (e.g., vehicles, production facilities, transportation infrastructure) are modeled, and in the decision making simulator, a set of transport chain decision makers (or roles) are modeled as agents. The main motivation for choosing a two-tier architectural design is that entities in the physical simulator are considered passive, while entities in the decision making simulator act independently and potentially proactively. The two layers are connected in a way that the activities in the physical simulator are initiated by the decisions taken by the decision makers. The agents appear in a hierarchical structure (see Fig. 1) to be able to represent the hierarchical decision making structure that is common in real-world transport chains. In hierarchical decision making structures, information is not automatically made available for all agents, which may result in a certain degree of local optimization. This appears to be rather typical in transport chains today, since often only a limited amount of information is shared (Zhou and Benton Jr., 2007).

TAPAS has mainly been validated by interviews with experts in policy issues and transport modeling, and with practitioners in transportation and logistics. Several simulation studies have been performed with TAPAS (e.g., Davidson et al., 2008), in which the sensitivity of different input parameters has been analyzed, and the results have in some studies been compared to similar studies (Ramstedt, 2008).

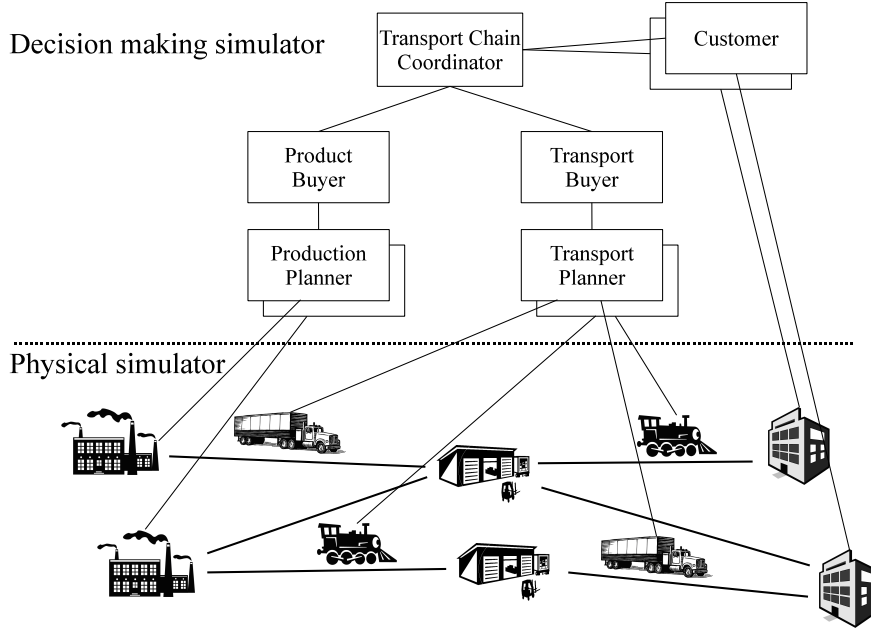


Fig. 1: Overview of the TAPAS simulation model with a physical simulator that is connected to a decision making simulator.

### 3.1 The physical simulator

TAPAS uses a transport network, which can be described as a directed graph  $(N, L)$  with a set of nodes  $N$  and a set of directed links  $L$ . A node can be a customer node, a producer (factory) node, or a connection point (terminal) in the network. For future reference, we let  $N^C \subseteq N$  refer to the customer nodes and  $N^F \subseteq N$  to the producer nodes. A link represents a directed connection between two network nodes and it corresponds to exactly one transport mode (typically road, rail or sea), and two nodes may be connected by multiple links representing different modes.

#### 3.1.1 Production

Commonly used production strategies are pull and push, where production is demand driven in the former and forecast driven in the latter. Pull strategies are often used to decrease inventory levels and to better capture demand variations, while push strategies often use resources better due to the potential for improved planning, e.g., by producing larger batches (Ahn and Kaminsky, 2005). In TAPAS we have chosen a pull strategy, in which products are produced in batches.

TAPAS models a set of product types  $P$ , and  $P_{n^f} \subseteq P$  denotes the set of product types that can be produced in node  $n^f$ . For each product type  $p \in P_{n^f}$ , we let  $c_{n^f p}^{mtrl}$  denote the raw material cost per unit,  $s_{n^f p}^{batch}$  the maximum batch size,  $c_{n^f p}^{prod\ time}$  the production cost per time unit, and  $t_{n^f p}^{batch\ proc}$  the batch processing (production) time. A setup time  $t_{n^f p}^{batch\ setup}$  (e.g., representing changeover) is considered when two batches of different product types are produced in sequence. The expected time

$$t_{n^f p}^{prod}(q) = \lceil q / s_{n^f p}^{batch} \rceil \cdot \left( t_{n^f p}^{batch\ setup} + t_{n^f p}^{batch\ proc} \right) \quad (1)$$

for producing  $q$  units of product type  $p$  in node  $n^f$  is calculated as the number of batches times the batch production time, where the setup time is considered only if the previously produced batch had a different product type. The expected cost

$$c_{n^f p}^{prod\ cost}(q) = t_{n^f p}^{prod}(q) \cdot c_{n^f p}^{prod\ time} + c_{n^f p}^{mtrl} \cdot q, \quad (2)$$

for producing  $q$  units of product type  $p$  is calculated as the production cost plus the cost for raw material.

Hence, TAPAS makes use of a cost model in which the cost for products depends on the time for production. The model has been chosen mainly due to its potential to represent different types of production models, i.e., (1) batch production, (2) continuous production, by setting the batch size to one and the batch setup time to zero, and (3) instant retrieval of products from storage, by setting the time-based production cost to zero, however without considering inventory costs for products.

### 3.1.2 Transportation

In TAPAS, transportation is carried out by a set of vehicles  $V$ , where a vehicle is characterized by its loading capacity (including type of storages), maximum traveling speed, fuel type, and emissions (e.g.,  $\text{NO}_x$ ,  $\text{CO}$  and  $\text{CO}_2$ ) per consumed unit of fuel. It should be noted that we also denote electricity as a fuel type. The amount of fuel that a vehicle consumes depends on the weight of the carried load, and for vehicle  $v \in V$ , we let  $\text{fuel}_v^{\text{empty}}$  denote the fuel consumption (per distance unit) when  $v$  is empty, and  $\text{fuel}_v^{\text{full}}$  the fuel consumption when  $v$  is carrying its maximum load  $w_v^{\text{max}}$ . The fuel consumption  $\text{fuel}_v(w)$  (per distance unit) when  $v$  carries a load of weight  $w$  ( $0 \leq w \leq w_v^{\text{max}}$ ) is approximated linearly according to

$$\text{fuel}_v(w) = \frac{\text{fuel}_v^{\text{full}} - \text{fuel}_v^{\text{empty}}}{w_v^{\text{max}}} \cdot w + \text{fuel}_v^{\text{empty}}. \quad (3)$$

A vehicle has a transport mode (road, rail or sea) and it is only able to travel on links with the same mode. Further, it might be controlled by a timetable (offering so called timetabled transports with scheduled times for departure and arrival) or not (providing transports referred to as demand driven). Transport costs consist of:

- time-based costs (e.g., driver, capital, and administration),
- distance-based costs (e.g., fuel, vehicle wear, and kilometer tax),
- link-based costs (e.g., road tolls), and
- fixed operator-based ordering costs (e.g., administration).

### 3.1.3 Terminals

For terminals, the focus is on modeling times and costs for loading and unloading of vehicles. The times for loading and unloading a vehicle are expressed in terms of fixed and variable times. Fixed times  $t_v^{\text{loading fixed}}$  and  $t_v^{\text{unloading fixed}}$  represent the times it take to prepare a vehicle for loading or unloading. Variable times  $t_{vp}^{\text{loading variable}}$  and  $t_{vp}^{\text{unloading variable}}$  are given for each product type and denote the times needed for loading or unloading one unit of product  $p$ . The costs  $c_v^{\text{loading}}$  and  $c_v^{\text{unloading}}$  for loading and unloading a vehicle are given as costs per time unit. The cost for loading  $q$  units of product type  $p$  onto vehicle  $v$  is calculated as:

$$c_{vp}^{\text{loading}}(q) = c_v^{\text{loading}} \cdot t_{vp}^{\text{loading}}(q), \quad (4)$$

where the loading time is calculated as:

$$t_{vp}^{\text{loading}}(q) = t_v^{\text{loading fixed}} + t_{vp}^{\text{loading variable}} \cdot q. \quad (5)$$

The formulas for unloading cost and time are identical to (4) and (5), except for that the cost and times for loading are replaced with those for unloading. It is also possible to define vehicle specific fixed costs for visiting a terminal, e.g., to represent terminal fees, and terminal processing times in addition to the times for preparing a vehicle for loading and unloading, e.g., to be able to model customs.

## 3.2 The decision making simulator

In TAPAS we consider six transport chain roles (or decision makers); transport chain coordinator, product buyer, transport buyer, transport planner, production planner and customer. These roles can be argued to be present in all transport chains. However, a role can be associated with different logistical operators depending on the organizational settings within the chain. For instance, the transport chain coordinator may be represented by a producer in one transport chain and by a third party company in another.

In TAPAS, each customer node  $n^c \in N^c$  is represented by a customer agent ( $C_{n^c}$ ), each production node  $n^f \in N^f$  by a production planner agent ( $\mathcal{PP}_{n^f}$ ), and transport planner agents ( $\mathcal{TP}$ :s) represent transport providers (operators). There exists one transport chain coordinator agent ( $\mathcal{TCC}$ ), one transport buyer agent ( $\mathcal{TB}$ ), and one product buyer agent ( $\mathcal{PB}$ ). In a transport chain it is often possible to identify additional roles, such as terminal and inventory agents (Ramstedt, 2008), and in TAPAS we have chosen to implicitly capture these roles in the physical simulator instead of modeling them as agents.

### 3.2.1 Interaction protocol for the ordering process

The process of buying resources (products or services) is a commonly occurring, and sometimes difficult, task, especially when the availability of resources are limited, different resources depend on each other, or multiple buyers and suppliers negotiate in parallel. The problem of how to automate the process of buying scarce resources is well covered in the literature (cf. Sandholm and Lesser, 2001; Knabe et al., 2002; Schillo et al., 2002), but according to our knowledge there exists no efficient approach for dealing with resources that depend on each other. When buying resources that depend on each other, it is important that booking is made in a way that there will be no situations where

one or more resources have been booked (and hence become useless) while it has become impossible to book some other resources that are required for accomplishing an overall task. Hence, synchronization and booking should be made in a way that it is guaranteed that either none or all required resources will be booked.

In TAPAS, a plan for fulfilling an order consists of a chain of tasks, starting with a production followed by a sequence of transports for different parts of the distance between the producer and the customer. The tasks in a plan need to be coordinated in a way that delivery is made inside the specified time window (if possible) and the starting and ending times of subsequent tasks are synchronized. If a resource included in a plan is no longer available at the time of booking, the whole plan may need to be reconsidered, and already booked resources may need to be canceled with a potential penalty as a consequence. A possible solution to this problem is to force resource providers to commit to their offers, as in the *contract net protocol* (Smith, 1980). However, such an approach makes it more difficult to find good solutions, as well as obtaining high resource utilization in the system (Knabe et al., 2002). The reason is that resources may be allocated to potential buyers who later on decide not to book them, at the same time as other potential buyers may need to make use of those particular resources in order to accomplish their goals.

In an approach for ordering matching products and transportation, two important challenges need to be considered: (1) products and transportation cannot be booked simultaneously, either products or transportation has to be booked first, and (2) products cannot be booked before finding a matching transport solution, and vice versa. In an earlier version of TAPAS (Bergkvist et al., 2005), the contract net protocol was used for buying products from the supplier offering the best price, and a second contract net protocol was used to buy a matching transportation even though it can be considered hazardous to assume the availability of a cost efficient matching transport solution at the time of buying products.

In the current version of TAPAS we have dealt with this potential problem by applying a sequential ordering process, in which only one customer order is allowed to be processed at the same time. This guarantees that complete plans for obtaining products will be available until all included tasks have been booked. In the ordering process, which uses an interaction protocol originally presented in (Holmgren et al., 2007),  $\mathcal{TCC}$  keeps a queue of incoming order requests that are processed in sequence. A similar approach, however applied to a manufacturing system, is presented in (Komma et al., 2007). The obvious shortcoming of such an approach is that longer order processing times increase the risk that the order queue becomes longer and longer, especially when there is a high ordering frequency in the system, or if order processing times are long. In TAPAS, this is however not a problem, since order processing is assumed to be instantaneous.

In Fig. 2 we illustrate the interaction protocol that is used in TAPAS for fulfilling a customer order, and in Table 1 we describe the message types that are used in the interaction protocol. A new interaction starts when  $\mathcal{TCC}$  receives an order request from a customer, and in two rounds of communication, a plan is first requested and generated (in the first ten messages) and then booked and confirmed (in the next ten messages). Hence it follows that customer agents are able to act proactively, while the other agent types are considered to be reactive. To improve the readability of the diagram, only the main flow of messages is shown. Also, the protocol can be interrupted in most of the steps, e.g., when  $\mathcal{PB}$  fails to find an available supplier, or when  $\mathcal{TB}$  is unable to find a feasible transportation for any product proposal.

### 3.2.2 Customer

Each customer node  $n^c \in N^C$  is operated by a customer agent  $C_{n^c}$  who is responsible for ordering products in quantities that keep inventories at levels that minimize the costs for inventory holding and ordering, while reducing the risk for stockouts. When choosing an ordering quantity, a tradeoff has to be made between two counteracting costs; order cost (including costs for products, transportation, administration, etc.) and inventory holding cost. Typically, the order cost per unit decreases and the inventory holding cost increases with an increasing consignment size, and minimizing these costs gives the *Economic Order Quantity* (EOQ).

In the *Wilson (EOQ) Model* (Wilson, 1934), which is claimed by Alstrom (2001) to be the most commonly used policy for finding the EOQ, the annual costs for ordering ( $DC/Q$ ) and inventory holding ( $HQ/2$ ) is minimized by calculating EOQ according to  $\sqrt{2CD/H}$ , where  $C$  denotes the fixed order cost (independent on order quantity),  $D/Q$  the number of orders per year,  $H$  the annual holding cost, and  $Q/2$  the average quantity in stock. However, the Wilson formula relies on a number of assumptions about the modeled system (Axsäter, 2006). For instance, it assumes constant purchase price per ordered unit and constant order cost, which both are assumed to be independent on order quantity and time of order placement. This implies, e.g., that no volume discounts can be modeled. However, in TAPAS a customer always pays the current cost for products and transportation, and in particular the cost for transportation varies, e.g., due to load consolidation and departures of timetabled transports.

To be able to cope with non-constant order costs, we have used an approach based on an idea for dealing with deterministic time-varying demand (Silver and Meal, 1973), in which time is divided into disjoint time periods, and candidate order quantities are those that cover the demand for a number of future time periods. The quantity that minimizes the costs for ordering and inventory holding is then chosen. In TAPAS, a customer chooses the quantity that

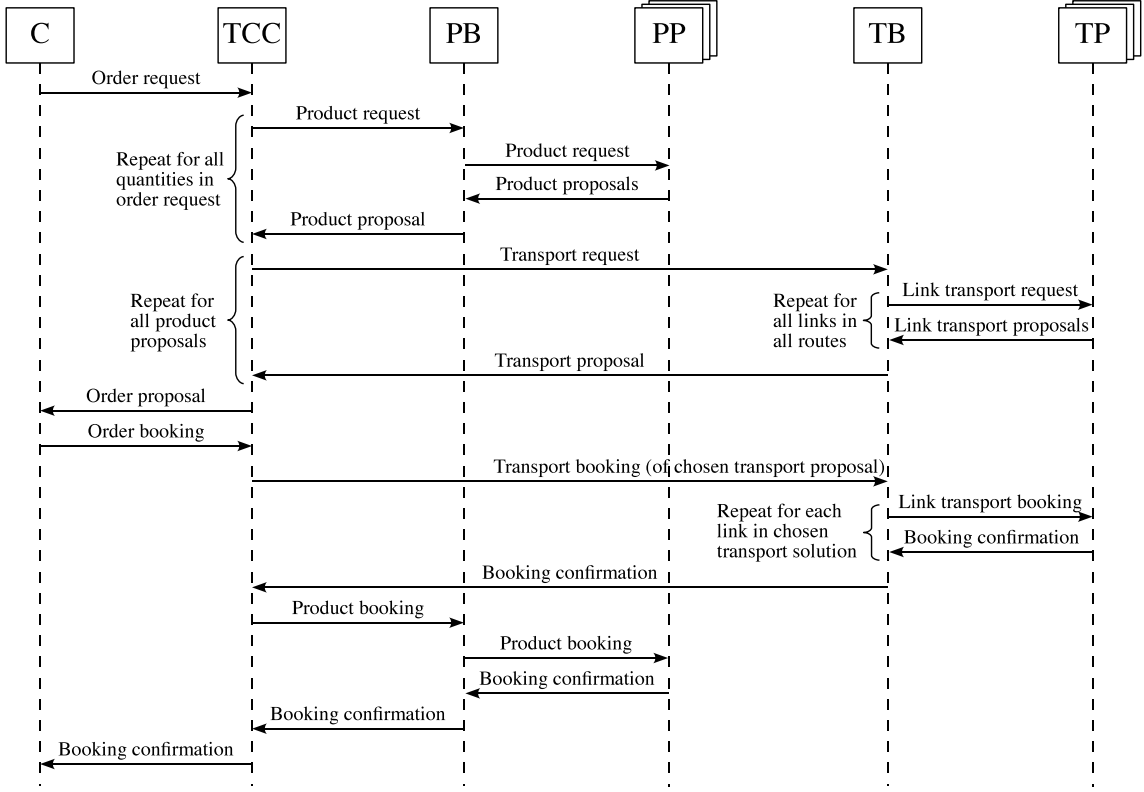


Fig. 2: Overview of the interaction protocol used in TAPAS only showing the main flow of messages.

minimizes the cost for ordering and inventory holding among a set of predetermined candidate quantities. Hence, the choice of candidate order quantities is made differently but the evaluation of quantities is identical in the approaches.

At a particular point in time, the cost  $c(p, q)$  per unit for ordering  $q$  units of product type  $p$  to node  $n^c$  is calculated as

$$c(p, q) = \frac{c^{order}}{q} + \frac{c^{order} \cdot int_p \cdot t^{con}(p, q)}{2 \cdot q},$$

where  $c^{order}$  denotes the current order cost, i.e., costs for products and transportation, including time-based product cost during transportation,  $int_p$  the storage interest for product type  $p$  (in node  $n^c$ ), and  $t^{con}(p, q)$  the expected time for consuming  $q$  units of product type  $p$ .

To be able to determine when to place an order, it is usually assumed that demand is known and constant over time, and that the “order-to-delivery” lead time is fixed. In TAPAS, consumption is stochastic and lead times vary due to several factors, including choice of transport mode and timetabled departures. Hence, the lead time remains unknown until both product supplier and transportation have been chosen. The uncertainty of stochastic consumption is dealt with by modeling safety stock levels, and each customer uses an estimated (approximate) lead time, chosen as an estimated upper bound of the expected lead times for different transport alternatives and suppliers. Since demand is stochastic and lead-times are just estimations, a difficulty with the applied ordering strategy is how to determine when ordering should be considered. The traditional approach is to consider ordering as soon as the inventory level falls below a certain threshold level. This is typically implemented by checking the inventory level in intervals, which in our approach is referred to as *ordering opportunities*. The time between two subsequent ordering opportunities should be small enough to reduce the risk for stock-outs. Note that  $C_{n^c}$  (the customer agent) may have different ordering opportunities for different product types, and for each ordering opportunity a separate decision determines whether or not an order request should be send. A decision whether or not to order at a particular point in time (i.e., an ordering opportunity) is based on the so called *order point*

$$order(p) = saf_p + lead_p \cdot con(p),$$

where  $saf_p$  denotes the safety stock level,  $lead_p$  the estimated “order-to-delivery” lead time, and  $con(p)$  the forecasted consumption per time unit for product type  $p$ . If the current inventory level  $inv^{cur}(p)$  plus the planned deliveries  $del(p, lead_p)$  of product type  $p$  during the lead time period is less than  $order(p)$ , i.e.,

$$inv^{cur}(p, n^c) + del(p, lead_p) < order(p),$$

Table 1: Specification of message types in order of appearance in the interaction protocol. The confirmation messages that are used in the protocol are considered self-explanatory, which is why we have chosen not to specify them.

<b>Order request</b>	Delivery node, product type, a set of candidate order quantities, and for each candidate quantity, a delivery time window, an upper limit for how much delivery is allowed to be delayed, and time-based convex penalty functions for delivering earlier or later than the window.
<b>Product request</b>	Product type and quantity.
<b>Product proposal</b>	Product type, quantity, pickup node, price, earliest time when the products can be picked-up, and supplier identifier.
<b>Transport request</b>	Product type, quantity, pickup and delivery nodes, delivery time window, earliest time when pickup can be made, time-based convex delivery penalty functions, and an upper limit for delivery delay.
<b>Link transport request</b>	Product type, quantity, departure and arrival nodes, and a time window in which transportation is considered.
<b>Link transport proposal (timetabled)</b>	Price (differentiated on transportation, loading and unloading), route identifier, scheduled times for loading, unloading, departure and arrival, and transport operator identifier.
<b>Link transport proposal (demand driven)</b>	Price (differentiated on transportation, loading and unloading), carrier type, availability interval, time needed for transportation, for loading and for unloading, and transport operator identifier.
<b>Transport proposal</b>	Sequence of link transport proposals representing a transport solution and total price for the solution. For demand driven link transport proposals, times for departure, arrival, loading and unloading are specified. Link transport proposals in sequence are if possible re-represented (i.e., merged together) so that each link transport proposal has exactly one loading and exactly one unloading.
<b>Order proposal</b>	Matching product and transport proposals.
<b>Order booking</b>	Order proposal.
<b>Transport booking</b>	Transport proposal.
<b>Link transport booking</b>	Link transport proposal as specified in the transport proposal.
<b>Product booking</b>	Product proposal and specification of pickup time.

$C_{nc}$  asks for order proposals (in an order request) for a set of quantities  $Q$ .

A delivery time window should be chosen in a way that there will be enough available inventory capacity at the time of delivery, and that the risk for stock-out before delivery is kept at a minimum. For an order quantity of  $q$  units, the delivery time window is calculated as

$$\left[ t^{cur} + a, t^{cur} + \max \left\{ a, \frac{inv^{cur}(p) + del(p, lead_p) - saf_p}{con(p)} \right\} \right],$$

where  $t^{cur}$  denotes the current time,  $inv_p^{max}$  the maximum allowed inventory level for product  $p$ , and parameter

$$a = \max \left\{ 0, \frac{inv^{cur}(p) + del(p, lead_p) + q - inv_p^{max}}{con(p)} \right\},$$

makes sure that there will be enough available inventory capacity at the time of delivery.

### 3.2.3 Transport chain coordinator

$\mathcal{TCC}$  is responsible for managing all customer orders, and received order requests are put in a queue for sequential processing. For each quantity  $q \in Q$  in an order request,  $\mathcal{TCC}$  tries to find the least cost feasible combination of products and transportation, which are represented in an order proposal. The order proposals (one for each  $q \in Q$ ) are returned to the customer who chooses the “best” proposal (i.e., the EOQ). To be able to generate order proposals,  $\mathcal{TCC}$  first asks  $\mathcal{PB}$  for relevant product proposals for each quantity  $q \in Q$ , by sending product requests. Then for each received product proposal, it asks  $\mathcal{TB}$  for a set of matching transport proposals, by sending transport requests.



### 3.2.4 Product buyer

$\mathcal{PB}$  operates in between  $\mathcal{TCC}$  and the  $\mathcal{PP}$ :s, and it is mainly responsible for finding products that satisfy product requests that are received from  $\mathcal{TCC}$ . When receiving a product request  $\mathcal{PB}$  forwards it to all  $\mathcal{PP}$ :s, and the product proposals that are returned are forwarded to  $\mathcal{TCC}$  for further processing. When the customer has chosen an order proposal, which represents EOQ,  $\mathcal{PB}$  sends a product booking message to the chosen product supplier.

The behavior of  $\mathcal{PB}$  is not particularly advanced, and it can be argued that  $\mathcal{TCC}$  could communicate directly with the  $\mathcal{PP}$ :s instead of communicating through  $\mathcal{PB}$ . We include  $\mathcal{PB}$  in the interaction protocol mainly for structural purposes (we argue that all transport chains should have a product buyer), and due to its potential for future improvement. An example of a possible refinement is to let  $\mathcal{PB}$  operate more proactively in the ordering process, e.g., by filtering out all product proposals except the best one in a particular geographic area.

### 3.2.5 Production planner

In TAPAS, each production node  $n^f \in N^F$  is operated by a production planner agent  $\mathcal{PP}_{n^f}$  whose main responsibility is to plan the production in  $n^f$ . When  $\mathcal{PP}_{n^f}$  receives a product request asking for a product type  $p \in P_{n^f}$  and quantity  $q$ , it generates a product proposal containing specification of the earliest time  $t_{n^f}^{prod}$  when the requested products can be made available for pickup at  $n^f$ , and a price  $c_{n^f}^{prod\ price}$ , which is generated according to Equation (2), however with the difference that the expected production time  $t_{n^f p}^{prod}(q)$  that is used as input to Equation (2) is calculated in Equation (1) assuming that a batch setup time is considered for all batches.

A production cost that is calculated using the model described in Section 3.1.1 depends on whether or not there need to be batch setups involved in the production, and in particular it depends on the batch sequencing, which typically varies according to the changing demand for products. It is therefore not possible to use Equation (1) to generate the production time when using the production cost model as a product price model without assuming that, e.g., a batch setup time is included in each produced batch. Otherwise, there would potentially be unrealistic dependencies between the production cost, which depends on factors such as batch sequencing, and the price for products, which normally is rather static. In summary, the production cost model used in TAPAS assumes that setup is considered only when there is a change in product type, and the product price model assumes that setup time is considered in all batches. Hence, the product price is (1) an upper bound of the production cost, and (2) it does not depend on batch sequencing, which is important in a model like TAPAS.

To enable an incoming order to be processed as early as possible, it is assumed that the existing production plan can be rescheduled so that already scheduled products are produced earlier than originally planned, assuming there is available production capacity for rescheduling. We have chosen not to consider the additional inventory costs that might be a consequence of rescheduled production, since that would considerably increase the modeling complexity. In Fig. 3 we illustrate how a production plan can be updated to enable early processing of an incoming order.

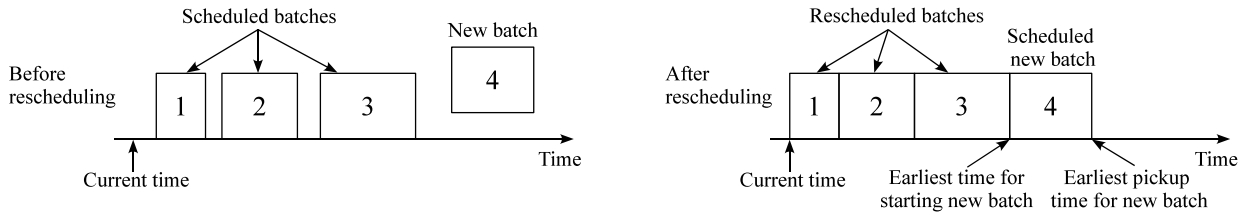


Fig. 3: Illustration of how three scheduled batches (1-3) can be rescheduled in order to enable early processing of an incoming order (4).

The rescheduling principle illustrated in Fig. 3 is used by  $\mathcal{PP}$ :s both when generating product proposals and when managing product bookings, i.e., when receiving product booking messages. When  $\mathcal{PP}_{n^f}$  receives a product request, it calculates the earliest possible pickup time  $t_{n^f}^{prod}$  by checking the earliest time that all scheduled productions can be finished. This is achieved by imagining a sequence of all scheduled productions (starting from the time of the request) in a way that there are no gaps in between any productions.

When receiving a product booking message,  $\mathcal{PP}_{n^f}$  first checks if there is enough available capacity for production immediately before the pickup time that is specified in the product booking message. In other words, it checks whether or not the production can be planned without rescheduling. If that is the case it schedules the production so that the booked products is expected to be available immediately before the requested pickup time. It should be noted that such a procedure typically creates gaps in the production plan, as illustrated to the left in Fig. 3. If there is insufficient capacity for production immediately before the specified pickup time,  $\mathcal{PP}_{n^f}$  has to reschedule the existing production plan so that the requested products can be produced immediately before the pickup time. The reason that productions

are always scheduled as late as possible it that it is believed that such a policy in most cases creates a minimum need for rescheduling. The rescheduling procedure can be seen as recursive, since moving a scheduled production earlier in time might create a need to reschedule the production immediately before that production, which might cause a need to reschedule the production even before that, etc. It should be noted that is assumed that the pickup time specified in a product booking message never violates the earliest pickup time given in the corresponding product proposal.

### 3.2.6 Transport buyer

In TAPAS, the transport buyer ( $\mathcal{TB}$ ) is responsible for creating transport solutions in order to satisfy transport requests received from  $\mathcal{TCC}$ . A transport solution is obtained by solving a shortest (cheapest) path problem with additional constraints for representing the following complicating factors:

1. Vehicle capacities and previous bookings must be considered, e.g., the remaining capacity on a booked vehicle should be made available for future bookings.
2. There is a mix of timetabled and demand driven transports.
3. Terminal activities such as unloading, loading, and reloading should be explicitly considered.
4. Delivery has to occur inside the delivery time window, otherwise a penalty cost should be applied.
5. Time-based, link-based, distance-based and fixed transport costs (as discussed in Section 3.1.2), and time-based costs, e.g., for tied up capital and deterioration should be considered.

The literature contains a number of approaches to intermodal freight planning (e.g., Caramia and Guerriero, 2009; Ziliaskopoulos and Wardell, 2000; Chang, 2008; Jansen et al., 2004). However, none of them applies to our problem since either (1) they do not handle all complicating factors that we need to consider, or (2) the combination of timetabled and demand driven transports are handled in an unsatisfactory way, e.g., by representing a demand driven transport as multiple timetabled departures, i.e., one for each time period. Such a modeling approach may have severe affects on the performance of the algorithm if the number of time-periods is high. Therefore, we have developed a customized algorithm for finding transport solutions. A compressed pseudo code description of the proposed algorithm is given in Algorithm 1. In the algorithm, a set of precompiled (manually specified) routes  $R_{n_f n_c}$ , expressed as sequences of network nodes is defined for each potential pair of pickup and delivery nodes  $n_f \in N^F$  and  $n_c \in N^C$ . Since routes are defined as node sequences, two sequenced nodes in a route may be connected by multiple links, e.g., corresponding to different transport modes. The motivation for using precompiled routes is that useful routes generally are known or can be found in advance.

In the algorithm,  $\mathcal{TB}$  processes the routes  $r \in R_{n_f n_c}$  in sequence to be able to find the least cost transport solution, possibly including penalty cost, for which a transport proposal is created and returned to  $\mathcal{TCC}$ . The algorithm is applied when  $\mathcal{TB}$  has received all link transport proposals for each pair of sequenced nodes  $(n_i, n_j)$  included in any route  $r \in R_{n_f n_c}$  (see the interaction protocol in Fig. 2). The link transport proposals that are returned from the  $\mathcal{TP}$ :s should have departure and arrival times, or intervals of availability for demand driven transports, inside an interval  $[t_{n_f}^{prod}, t_{n_c}^{late} + k^{late}]$ , where  $t_{n_f}^{prod}$  is the earliest time the products can be available for pickup at  $n_f$ ,  $t_{n_c}^{late}$  is the end of the delivery time window  $[t_{n_c}^{early}, t_{n_c}^{late}]$ , and  $k^{late} \geq 0$  is a constant that is used to allow delivery to be made later than the delivery time window. However, penalty costs as specified by the customer are applied to deliveries earlier or later than the delivery time window.

In the search for the optimal transport solution the algorithm uses a search tree  $T$ , in which each tree node  $\eta$  (except for the root of the tree) represents a link transport proposal with departure (network) node  $dep(\eta)$  and arrival node  $arr(\eta)$ . Instead of representing a transport proposal, the root  $\rho$  represents the delivery time window and serves as the parent of all tree nodes representing transports arriving at  $n_c$ . The search starts in  $n_c$  and proceeds backwards (along the routes  $r \in R_{n_f n_c}$ ) towards  $n_f$  by gradually expanding the search tree. For a search tree node  $\eta$ , one child node will be added for each link transport proposal representing a transport arriving at  $dep(\eta)$ . Hence, it follows that a link transport proposal will be represented by multiple tree nodes in order to enable evaluation of each possible combination of link proposals along the routes. Each search tree node  $\eta$  has a cumulative cost  $c(\eta)$ , which corresponds to the cost for moving the goods from  $dep(\eta)$  to  $n_c$  (including unloading at  $n_c$ ). In the algorithm the following search strategies are applied:

1. Tree nodes for which the transports they represent are infeasible in combination with the transports represented by their ancestor nodes, as well as nodes that necessarily lead to transport solution that are more expensive than the so far best found solution will be pruned (i.e., no child nodes will be added).
2. A depth-first based traversal strategy is adopted in order to potentially find valid solutions early to enable early pruning of more expensive tree branches.
3. Tree nodes representing timetabled link transport proposals are expanded before nodes representing demand driven link transport proposals.

The main challenge for the algorithm is to deal with the differences between timetabled and demand driven transports, where a previously booked demand driven transport is regarded as timetabled. The difficulty is that timetabled

---

**Algorithm 1** Find the least cost transport solution for a transport request, given a pickup node  $n_f$ , a delivery node  $n_c$ , and link transport proposals for all connections in the routes in  $R_{n_f n_c}$ . The cost for a tree node corresponds to the cost for traveling from the departure node of the link transport proposal it represents until unloading at  $n_c$ .

---

**procedure** FINDBESTTRANSPORTSOLUTION

create rootNode ▷ rootNode represents the delivery time window

**initialization:** rootNode.cost := 0, rootNode.bestChild := null, bestCost :=  $\infty$

**for** each route  $r \in R_{n_f n_c}$  **do**

**for** each link transport proposal  $t$  in  $r$  which has arrival node  $n_c$  **do**

        newChild = createChild(rootNode,  $r$ ,  $t$ , bestCost)

**if** newChild.cost < rootNode.bestCost **then**

            rootNode.bestChild := newChild, bestCost := newChild.cost

**return** best transport solution with cost bestCost

▷ the best solution is obtained from the best child

**procedure** CREATECHILD(parentNode,  $r$ ,  $t$ , bestCost)

create new node currentNode

**initialization:** currentNode.bestChild := null, currentNode.bestChildCost :=  $\infty$

**if**  $t$  is infeasible in combination with the transport represented by parentNode **then**

    currentNode.cost :=  $\infty$ , **return** currentNode

**if**  $t$  is demand driven **then**

    set preliminary times for departure and arrival, and potentially for loading and unloading \*

calculate currentNode.cost

▷ includes cost for transportation, loading, unloading, penalty, etc.

currentNode.cost = currentNode.cost + parentNode.cost

**if**  $t$ .departureNode =  $n_f$  **then return** currentNode.cost

▷ a complete transport solution is found

**if** currentNode.cost  $\geq$  bestCost **then** currentNode.cost :=  $\infty$  **return** currentNode

**for** each link transport proposal  $t'$  in  $r$  that is connected to the arrival node of  $t$  **do**

    newChild = createChild(currentNode,  $r$ ,  $t'$ , bestCost)

**if** newChild.cost < currentNode.bestChildCost **then**

        currentNode.bestChild := newChild, currentNode.bestChildCost := newChild.cost

**return** currentNode

\* Potentially involving a backtracking towards rootNode in order to represent updated preliminary times for a sequence of demand driven ancestor nodes.

---

transports have fixed departure times while demand driven transports are flexible concerning departure and arrival times. For simplicity of presentation, we here refer to tree nodes representing timetabled and demand driven transports as timetabled and demand driven nodes.

When expanding a search tree node, the feasibility is checked for all added child nodes, however in different ways for timetabled and demand driven nodes, as explained below. Based on the departure and arrival times of its ancestors, for a demand driven node the algorithm specifies preliminary times for departure, arrival and potentially for loading and unloading. For a demand driven node it also specifies a possible *interval of operation*, defining when the activities for the represented transport are allowed to take place. For a demand driven transport, the interval of operation is based on the interval of availability that is specified in the link transport proposal, as well as on the operation times of the ancestors, and it may need to be used further down in the tree, e.g., when a timetabled successor requires that the preliminary specified times for one or more demand driven ancestors should be updated. For a parent node  $\eta$  and child node  $\zeta$ , we let  $z(\zeta, \eta) \in \{0, 1\}$  denote whether or not the products will be reloaded ( $z(\zeta, \eta) = 1$ ) or not ( $z(\zeta, \eta) = 0$ ) between the transports represented by  $\zeta$  and  $\eta$ , and  $c^{\text{reloading}}(\zeta, \eta) = c^{\text{unloading}}(\zeta) + c^{\text{loading}}(\eta)$  the cost for reloading. Since it is only known by the child whether or not the goods need to be reloaded, the cost for loading has to be represented in the child node. Moreover, since a parent typically has multiple child nodes, a reconsidering of the preliminary times for a demand driven node needs to be represented in the particular successor node that requires that the times should be updated.

**The root.** As mentioned above, the root node  $\rho$  represents the delivery time window, and a timetabled transport arriving at  $n_c$  (i.e., a timetabled child of  $\rho$ ) will always be valid since a penalty cost can be applied if delivery is made outside  $[t_{n_c}^{\text{early}}, t_{n_c}^{\text{late}}]$ . Similarly, a demand driven transport is valid whenever the time interval for when it is available, as specified in the link transport proposal, is at least as long as the time it takes for the vehicle to perform its activities. For a demand driven transport, preliminary times for unloading, arrival and departure are set as late as possible in the delivery time window, and if delivery inside the delivery time window is impossible due to the interval of availability, the preliminary

times are heuristically chosen either before or after the delivery time window in a way that the penalty is minimized. The general form for describing the cost for a node that connects to  $\rho$  is:

$$c(\zeta) = c^{transp}(\zeta) + c^{unloading}(\zeta) + c^{time}(\zeta) + pen(\zeta),$$

where  $c^{transp}(\zeta)$  is the transport cost,  $c^{unloading}(\zeta)$  the unloading cost,  $pen(\zeta)$  the potential penalty cost that need to be considered when adding  $\zeta$ , and  $c^{time}(\zeta)$  a time-based product cost, i.e., the mathematical product of the product value, the time between departure and unloading, and a time-based factor (or interest rate). The customer chooses this time-based factor, and it may represent a cost for tied-up capital, deterioration, time to market, etc., for the particular type of product.

**Intermediary nodes.** When adding a child node  $\zeta$  to an intermediary search tree node  $\eta$  (i.e.,  $\eta$  is not the root), four cases can occur depending on what types of transports  $\eta$  and  $\zeta$  represent. It should be noted that the transport represented by  $\zeta$  occurs before the transport represented by  $\eta$ . With the notation introduced above, the cost  $c(\zeta)$  for reaching  $\zeta$  is described as:

$$c(\zeta) = c(\eta) + c^{transp}(\zeta) + z(\zeta, \eta) \cdot c^{reloading}(\eta, \zeta) + c^{time}(\zeta) + pen(\zeta),$$

where in particular  $pen(\zeta)$  and  $c^{time}(\zeta)$  depends on several factors. The base case is that a time-based cost  $c^{time}(\zeta)$  should be generated for the time between the departure times of  $\zeta$  and  $\eta$ .

If both  $\eta$  and  $\zeta$  are timetabled, it is only necessary to conduct a feasibility check to ensure that  $\zeta$  will be further expanded only if the arrival time of  $\zeta$  is earlier than the departure time of  $\eta$ , and if  $\eta$  and  $\zeta$  represent transports by different vehicles there also need to be time for reloading.

When  $\eta$  is timetabled and  $\zeta$  is demand driven, the time for departure and loading in node  $\eta$  are strictly specified in the corresponding link transport proposal, and the times for the activities represented by  $\zeta$  should be synchronized with the activities in node  $\eta$ . Since  $\eta$  represents a timetabled vehicle while  $\zeta$  represents a demand driven one (a vehicle is assumed to offer either timetabled or demand driven transports), it is assumed that reloading will occur between  $\zeta$  and  $\eta$ . The unloading of  $\zeta$  will be preliminary scheduled immediately before  $\eta$  starts loading, and the actual transport (departure and arrival) immediately before that. If  $\zeta$  is unavailable at the preferred time of operation, instead the activities in  $\zeta$  should be preliminary scheduled earlier in the period of availability in a way that (as late as possible) the penalty is minimized.

The case when both  $\eta$  and  $\zeta$  are demand driven is similar to the previous case. However, if reloading needs to occur between  $\eta$  and  $\zeta$  (i.e.,  $\eta$  and  $\zeta$  represent different vehicle types), the loading of  $\eta$  needs to be preliminary scheduled, as this was impossible to do when  $\eta$  was created. However, the preferred time for loading in  $\eta$  might be infeasible due to the interval of operation for  $\eta$ , which means that the preliminary scheduled activities for  $\eta$  will have to be rescheduled (actually delayed) and additional penalties, as well as an additional time-based cost ( $c^{time}(\zeta)$ ) should be considered. Delaying  $\eta$  means that additional demand driven nodes, i.e., ancestors of  $\eta$ , also may have to be delayed, depending on how their activities are preliminary planned. Actually, the preliminary times of all sequential demand driven nodes that can be found in a back-track search (towards the root  $\rho$ ) until either a timetabled node or the root is reached, may need to be rescheduled. If (in the back-track search) a timetabled node  $\psi$  is reached before reaching  $\rho$ , the preliminary times for the activities represented by the demand driven ancestors between  $\zeta$  and  $\psi$  should be synchronized also considering the loading time of  $\psi$  (note that reloading will occur since  $\psi$  is timetabled and its child in the current branch is demand driven). If  $\rho$  is reached before any timetabled node could be found, i.e.,  $\rho$  and  $\zeta$  is connected by a sequence of demand driven transports, a penalty for delivering outside the specified delivery time window may have to be applied as a possible consequence of rescheduling the node connected to  $\rho$ . Another reason for rescheduling a sequence of demand driven ancestors is a limited availability interval of  $\zeta$ . Of course, limited intervals of operation for one or more of the demand driven transports that need to be rescheduled may limit how rescheduling may be conducted, and in worst case  $\zeta$  will be infeasible, and hence pruned. Since each of the, potentially rescheduled, ancestors may have multiple successors, it is important to emphasize that extra penalty costs and time-based costs that are a consequence of the rescheduling of ancestors have to be represented in  $\zeta$ . Also, since a rescheduled node may have other branches, which require different or no rescheduling, a specification of how activities for the ancestors are rescheduled should be represented in  $\zeta$ .

The case when  $\eta$  is demand driven and  $\zeta$  is timetabled is similar to the case when both  $\eta$  and  $\zeta$  are demand driven. The difference is that for  $\zeta$ , the times for loading, unloading, and traveling are represented in the corresponding link transport proposal. As in the previous case, it might be necessary to reschedule  $\eta$  and a sequence of demand driven ancestors of  $\eta$  whenever the times of  $\zeta$  causes infeasibility with the preliminary times of  $\eta$ . In the same way as discussed above, penalty cost and an additional time-based cost may have to be considered.

**Leaf nodes.** Leaf nodes, which represent transports departing from  $n_f$ , are processed exactly as intermediary nodes, with the difference that loading also need to be scheduled for  $\zeta$ . This might lead to infeasibility and extra penalty cost

and time-based cost, in the same way as discussed above. In general, the cost for a leaf node can be calculated as:

$$c(\zeta) = c(\eta) + c^{transp}(\zeta) + z(\zeta, \eta) \cdot c^{reloading}(\zeta, \eta) + c^{loading}(\zeta) + c^{time}(\zeta) + pen(\zeta).$$

When the best sequence of tree nodes (a transport solution) for a transport request has been determined, the preliminary times for all demand driven transports in the solution need to be permanently set. This may require that demand driven transports included in the solution needs to be rescheduled according to the preferences of nodes further down in the tree (i.e., closer to the leaf). If more than one node specifies a rescheduling of the same demand driven ancestor, it is the preferences of the node closest to the leaf that should be applied. The reason is that the preferences of a node that is added later in the tree also takes into account the preferences of nodes that have been added earlier.

### 3.2.7 Transport planner

A transport planner agent ( $\mathcal{TP}$ ) represents an owner of a fleet of vehicles  $V_T \subseteq V$ , which are assumed to operate in some particular geographic area, and the main responsibility of  $\mathcal{TP}$  is to plan the vehicles in  $V_T$  by assigning them transport tasks. When  $\mathcal{TP}$  receives a link transport request from  $\mathcal{TB}$ , specifying start node  $n_s$ , end node  $n_e$ , and a window  $[t^s, t^e]$ , it creates a set of link transport proposals that are returned to  $\mathcal{TB}$ . Since timetabled vehicles depart and arrive according to timetables, while demand driven vehicles are flexible regarding when they are able to operate, it follows that link transport proposals for the two different types of transports need to contain somewhat different information. For a timetabled transport, a link transport proposal specifies the scheduled times for loading, unloading and transportation, and for a demand driven transport, instead the time needed for traveling, loading and unloading, as well as a time interval in which the transport is available for operation is specified. For a timetabled vehicle, a separate link transport proposal will be generated for each scheduled departure between  $n_s$  and  $n_e$ , which has departure and arrival inside  $[t^s, t^e]$ . For demand driven transports, a separate link transport proposal can be generated for each available demand driven vehicle, but since this may result in a large number of similar (or identical) link transport proposals, instead one proposal, specifying an interval of availability is generated for each different vehicle type in  $V_T$ . Some advantages with this approach are that the interval of availability for a vehicle type may be considerably larger than those for the individual vehicles of the particular type, and the number of link transport proposals can be reduced. Moreover, link transport proposals for previously booked demand driven transports are generated as for timetabled transports, and each scheduled demand driven departure (from  $n_s$  to  $n_e$  within the requested time interval) with remaining capacity should be represented in a separate link transport proposal, which specifies the planned times for departure, arrival, loading, and unloading.

Typically, a vehicle carries products for more than one customer (i.e., load consolidation), and a buyer of transport capacity should pay a reasonable share of the total transport cost. However, when giving a link transport proposal,  $\mathcal{TP}$  only knows about already booked consignments. Future bookings can only be guessed, or estimated from forecasts. Also, how much goods will be carried on the return transport is something that potentially could have an effect on the price for transport capacity. We do not model return transports explicitly, but return transports can implicitly be represented by adjusting different cost components and the average load utilization factor, and it is the responsibility of the user to decide if and how return transports should be considered. In summary, the major difficulty for  $\mathcal{TP}$  is to determine prices for those transport bookings that do not utilize the full capacity of a vehicle.

An obvious approach for setting the transport price is to charge a price that is based on the weight and size of the goods to be transported, as well as on an average load utilization factor. In this approach, no volume discounts are modeled, and the cost for buying transport capacity is calculated as:

$$c^{order} = \frac{c^T \cdot f^L}{f^{avg}} + c^{fuel}(w), \quad (6)$$

where  $c^T$  is the total transport cost (including fuel cost for an empty transport),  $f^L$  the load utilization factor for the requested load,  $f^{avg}$  the assumed average vehicle load utilization factor (potentially different for different vehicle types, geographical areas and transport operators), and  $c^{fuel}(w)$  the fuel addition cost based on the weight  $w$  of the requested load. The load factor ( $f^L$ ) is calculated as the requested quantity  $q$  divided by the maximum loading capacity of the requested product type, restricted either by the volume or weight capacity of the vehicle. As mentioned above, the fuel cost for a transport is included in  $c^T$ , and the fuel addition cost  $c^{fuel}(w)$  is calculated as:

$$c^{fuel}(w) = l \cdot (fuel_v(w) - fuel_v^{empty}), \quad (7)$$

where  $l$  is the traveled distance and  $fuel_v(w)$ , which is the fuel consumption per distance unit when carrying a load of weight  $w$ , is calculated according to Equation (3).

In addition to the linear pricing model described above, we provide a pricing model in which a risk cost is added to bookings that leave transport capacity unbooked. The risk cost is based on an estimated probability  $d$ , at the time of

giving a link transport proposal, that there will be at least one more booking before the departure. Typically, the risk costs for bookings decrease as the amount of booked transport capacity increases, and it is the responsibility of the user of TAPAS to specify how the probabilities should be calculated.

In the pricing model with a risk cost, the order cost is calculated as:

$$c^{order} = c^T \cdot f^L + c^{fuel(w)} + risk, \quad (8)$$

where the risk cost is calculated as:

$$risk = (1 - d) \cdot c^T \cdot f^R, \quad (9)$$

and  $f^R$  is the remaining transport capacity, at the time of booking, in percentage of the total capacity.

In summary, we propose two different approaches for setting the price for transport capacity; one where the price depends linearly on how much capacity is ordered, and one where early bookings gets a larger share of the transport cost to cover for the uncertainty regarding future bookings.

### 3.3 Implementation

TAPAS is implemented as a discrete event simulator, since we found it appropriate to model the studied domain as a chronological sequences of events, representing activities such as order processing, loading, unloading, departures, and arrivals. It makes use of a next-event time advance mechanism, instead of fixed-increment time advance approach, due to its ability to simulate time in an arbitrary level of detail. TAPAS is implemented in the Java language, and the decision making simulator is implemented using the Java Agent DEvelopment Framework (JADE) platform (Bellifemine et al., 2007).

Agents operate autonomously and typically asynchronously, and in agent-based (parallel) simulation it is important to carefully consider how to synchronize activities with the simulation clock. There are two approaches for synchronizing parallel simulation (cf. Xuehui et al., 2009); in *conservative protocols* activities are always simulated in chronological order, and in *optimistic protocols* activities are allowed to be simulated in non-chronological order, however with the consequence that the simulation needs to be recovered when synchronization errors occur. TAPAS makes use of a conservative approach, in which a *synchronization agent* is used to ensure that activities always are simulated in chronological order. The synchronization agent manages the simulation clock and it maintains the *event queue*, in which future events are sorted in chronological order. It should be noted that, at a particular point of time only some future events are typically known; events are continuously generated as the simulation proceeds over time. The behavior of the synchronization agent can be described as a loop in which it (1) removes the first event (i.e., next in time) from the event queue and (2) advanced the simulation clock to the time of the event that was removed from the queue. It also sorts new events into the event queue whenever they are created.

Activities are in TAPAS represented by start and end events, which makes it straightforward to simulate that activities may take longer time than expected, i.e., delays in activities. In TAPAS it is possible to model production and transportation using user-defined probability distributions. Start events are created as consequences of other activities, e.g., since order processing occurs at fixed points in time the start event for next order processing is created when processing the end event for the previous order processing, and events for starting production and transportation are generated at the time of booking those activities. The end event for an activity is created at the start of the activity (i.e., when the start event is processed). The delay for an activity is known already when the activity starts, and it is therefore captured in the time for the end event of the activity.

Those physical entities that perform activities (not entities such as links and nodes) are in TAPAS connected to agents. For example, vehicles are connected to transport planners and production facilities to production planners. When the synchronization agent processes an event, it sends a message to the agent connected to that event, providing informing about the activity that should start or end (depending on whether the processed event is a start or end event). In case the activity cannot start (or end) at the particular point of time, the agent returns a message informing about that. The synchronization agent is equipped with a mechanism for handling delays and dependencies between activities. In addition to managing the event queue, the synchronization agent also keeps a list of events that for different reasons are unable to start, e.g., since they are waiting for other, potentially delayed, activities to end. For example, the loading of a vehicle in a producer depot may have to wait for a delayed production to finish. Whenever an activity has ended (i.e., an end event has been processed) the synchronization checks if any activity in the waiting list could start as a consequence of the ended activity. If that is the case, the synchronization agent creates a start event for the activity that might start and schedules it in the event queue as early as possible.

## 4 Scenario and simulation study

In this section we illustrate the functionality included in TAPAS by briefly describing a simulation study that has been conducted in a scenario around the Southern Baltic Sea. The scenario is an extension of a scenario that has been

studied in collaboration with partners in an EU financed project called EastWest Transport Corridor (<http://www.eastwesttc.org>). It contains one logistical terminal, in Kaunas (Lithuania), which provides two types of products, and three customers; one in Sweden (Älmhult) and two in Denmark (Copenhagen and Esbjerg). In the scenario, which is illustrated in Fig. 4, transportation of 20 ft ISO-containers (TEUs) from Kaunas to the three customers is offered by five transport providers. Transportation by sea and rail is assumed to follow timetables while transportation by road is demand driven. A detailed description of the scenario, including input parameters for all entities, timetables, etc., is given in (Holmgren, 2011).

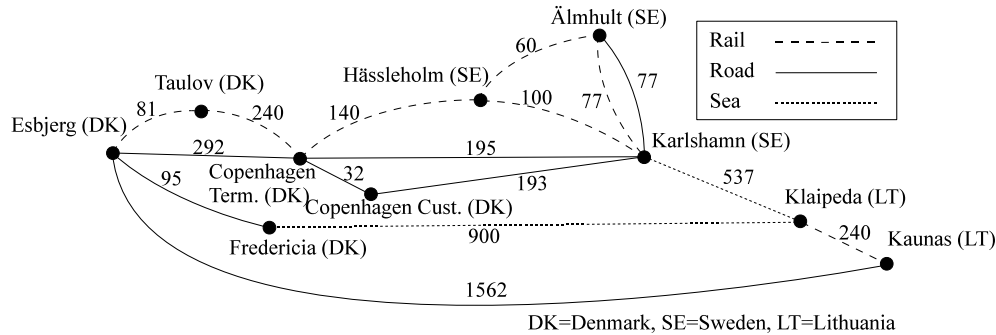


Fig. 4: Illustration of the transport network used in the scenario, where the numbers on the links represent distances in kilometers.

We studied the consequences of three transport policy and infrastructure measures aimed at achieving a modal shift from road to rail and sea transportation, which is a goal within the European Union (European Commission, 2001); an introduction of a kilometer tax for heavy trucks in Sweden and parts of Denmark, a CO<sub>2</sub> tax for all types of transports in the studied area, and a new direct railway link between Karlshamn and Älmhult (the so-called SouthEast Link). The kilometer tax level is suggested in (Friberg et al., 2007) and it is differentiated based on the euro class, as well as on the total weight of the trucks, the CO<sub>2</sub> tax levels are similar to the levels discussed in (SIKA, 2008), and the SouthEast Link is an infrastructure project that is currently discussed in Sweden. In the study, we investigated two different timetables for the SouthEast link.

The scenario and its results have been validated through interviews with domain experts, and a visualization of the scenario helped us discover unrealistic assumptions and to facilitate the communication of assumptions and simulation results. We have performed sensitivity analyses of different input parameters in order to understand what influence different parameters have on the results. Moreover, the estimated transport cost structures, i.e., the relations between time-based and distance-based costs have been compared to the cost structures used in the Samgods model (Swahn, 2001).

In the simulation study we considered the following experimental settings:

- S0.** The base case refers to the current situation in which none of the studied measures are applied.
- S1.** S0 + a kilometer tax of 0.15 euro/km for trucks operating in Sweden, and between Copenhagen Terminal and Copenhagen Customer.
- S2.** S0 + a CO<sub>2</sub> tax for all vehicles operating in the modeled region. Tax levels from 0.10 euro/kg up to 0.30 euro/kg in steps of 0.05 were considered. We let S2.x refer to setting S2 with a CO<sub>2</sub> tax level of 0.x euro/kg.
- S3.** S0 + a new railway link between Karlshamn and Älmhult (i.e., the SouthEast Link). Two timetables, which are synchronized in different ways with ferry arrivals in Karlshamn (from Klaipeda), were considered: (a) worse synchronization and (b) better synchronization (see (Holmgren, 2011) for specification of timetables).

For each setting we simulated 420 days with a precision of 1 min. We made simulation runs with 10 sets of random generator seeds for variation of consumption (different seeds were used for different customers). Each set of seeds was used for all settings, which enabled us to make pair-wise comparisons of results for different settings.

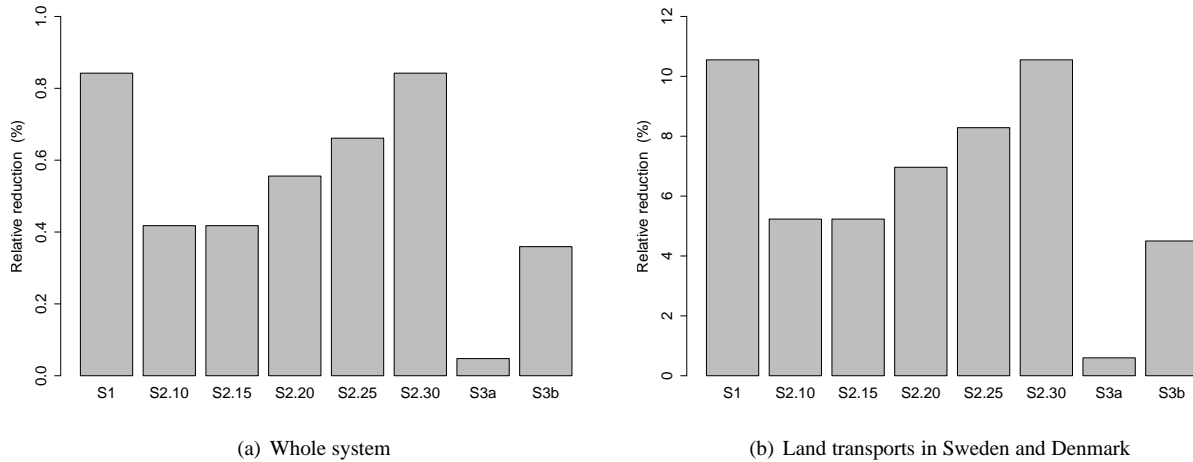
From a larger set of available routes (see (Holmgren, 2011) for specification of all routes available in the scenario), we observed that the following five routes were used, however in different proportions for different settings:

- Route 1.** Kaunas (Rail) Klaipeda (Sea) Karlshamn (Road) Älmhult
- Route 2.** Kaunas (Rail) Klaipeda (Sea) Karlshamn (Rail) Älmhult
- Route 3.** Kaunas (Rail) Klaipeda (Sea) Karlshamn (Rail) Copenhagen Terminal (Road) Copenhagen Customer
- Route 4.** Kaunas (Rail) Klaipeda (Sea) Karlshamn (Road) Copenhagen Customer
- Route 5.** Kaunas (Rail) Klaipeda (Sea) Fredericia (Road) Esbjerg

An important indicator of the impact of the studied measures is the route choice, which we illustrate with the per-

centage of TEUs transported using different routes. In Table 2 it can be seen that all of the studied measures caused a shift towards routes involving more rail transports and less road transports. However, for different measures the shift was observed in different parts of the network. For transportation to Älmhult, the only measure that showed an effect on the route choice is the SouthEast Link. In the settings without the SouthEast Link, all TEUs were transported on road between Karlshamn and Älmhult (i.e., Route 1). In the settings with the SouthEast Link (S3a and S3b) we observed shifts toward Route 2 using the SouthEast Link (i.e., railway). In S3a in average 5.8% and in S3b in average 43.6% of the TEUs were transported using Route 2. Not surprisingly, in the setting with better timetable synchronization (S3b) we observed a higher shift than in the setting with slightly worse synchronization (S3a). In all settings, transportation of all TEUs to Esbjerg were made using Route 5, with sea transportation from Klaipeda to Fredericia followed by road transportation from Fredericia to Esbjerg. This is reasonable due to the fact that the long distance makes it economically tractable to use sea transportation instead of land transportation through Sweden and Denmark. For Copenhagen Customer, which can be reached only by truck, the results vary for the different settings. For settings S0 and S3, in which no measures were applied for transportation on the routes to Copenhagen Customer, all transports were made using road transportation from Karlshamn directly to Copenhagen Customer (i.e., Route 4). In setting S1 (kilometer tax), a 100% shift towards Route 3 using rail between Karlshamn and Copenhagen Terminal followed by road transportation between Copenhagen Terminal and Copenhagen Customer was observed. For settings S2 (CO<sub>2</sub> tax), a gradually increasing shift towards Route 3 was observed as the tax level was increased.

Fig. 5: For each studied measure, the relative reduction (in percentage) of CO<sub>2</sub> emissions for (a) the whole system, and (b) land transports in Sweden and Denmark.



A positive consequence of achieving a shift from road to rail transportation is a reduced amount of CO<sub>2</sub> emissions. All observed reductions of CO<sub>2</sub> emissions in the studied system was a consequence of a modal shift from road to rail in Sweden and Denmark. In Table 2 we show the average CO<sub>2</sub> emissions (in tons), as well as the average transport work (in tonne-kilometer), which is another important indicator of a changed transport pattern, for transportation of one TEU from Kaunas to each customer. Moreover, in Fig. 5 we present the relative CO<sub>2</sub> reduction for (a) the whole system, and (b) land transports in Sweden and Denmark. The main reason for observing a small reduction of CO<sub>2</sub> emissions when considering all simulated transports is that a significant share of the transports in all settings were made using sea and rail transportation between Kaunas and Karlshamn and between Klaipeda and Esbjerg. Further, the transport costs are affected in different ways by different measures and it may be important to analyze the positive effects (e.g., reduced CO<sub>2</sub> emissions) in relation to the economic impact caused by applying measures. For the studied measures, we show in Table 2 the average costs for transporting one TEU to the different customers.

## 5 Discussion

A general assumption in TAPAS is that decision makers are local cost minimizers. It is possible to model decision makers in other ways, e.g., by letting them behave according to simple rules that are coded into the agents. However, it is often reasonable to assume that real-world decision makers strive towards behaving optimally. Moreover, it may sometimes be relevant to let other factors than cost, e.g., environmental impact, reliability and punctuality, influence the decision making. However, most types of factors can be represented in the decision making process by different



Table 2: For each setting, the average taken over 10 replications for a number of relevant types of output data.

		S0	S1	S2.10	S2.15	S2.20	S2.25	S2.30	S3a	S3b
<b>Share of TEUs (in percentage) transported using different routes (differentiated on customer)</b>										
Älmhult	Route 1	100	100	100	100	100	100	100	94.2	56.4
	Route 2	0	0	0	0	0	0	0	5.8	43.6
Copenhagen	Route 3	0	100	49.6	49.6	66.0	78.5	100	0	0
	Route 4	100	0	50.4	50.4	34.0	21.5	0	100	100
Esbjerg	Route 5	100	100	100	100	100	100	100	100	100
<b>Average transport work (in tonkm) for transportation of one TEU to each customer (differentiated on transport mode)</b>										
Älmhult	Rail	2640	2640	2640	2640	2640	2640	2640	2689	3009
	Road	847	847	847	847	847	847	847	798	478
	Sea	5907	5907	5907	5907	5907	5907	5907	5907	5907
Copenhagen	Rail	2640	5280	3949	3949	4382	4713	5280	2640	2640
	Road	2123	352	1245	1245	954	733	352	2123	2123
	Sea	5907	5907	5907	5907	5907	5907	5907	5907	5907
Esbjerg	Rail	2640	2640	2640	2640	2640	2640	2640	2640	2640
	Road	1045	1045	1045	1045	1045	1045	1045	1045	1045
	Sea	9900	9900	9900	9900	9900	9900	9900	9900	9900
<b>Average CO<sub>2</sub> emissions (in tons) for transportation of one TEU from Kaunas to each customer</b>										
Älmhult		0.623	0.623	0.623	0.623	0.623	0.623	0.623	0.622	0.616
Copenhagen		0.679	0.660	0.670	0.670	0.667	0.664	0.660	0.679	0.679
Esbjerg		0.885	0.885	0.885	0.885	0.885	0.885	0.885	0.885	0.885
<b>Average cost (in euro) for transporting one TEU to each customer.</b>										
Älmhult		364.2	368.5	426.0	456.9	487.8	518.7	549.6	364.7	364.3
Copenhagen		554.8	566.3	627.3	660.9	695.4	729.4	764.5	554.9	554.8
Esbjerg		519.0	519.0	606.9	650.8	694.7	738.6	782.6	519.0	519.0

types of costs. For instance, uncertainty regarding reliability and punctuality may be represented as internal costs in organizations, and environmental impact may be represented as governmental taxes and fees.

The cost models that are used in TAPAS (e.g., for production, transportation and ordering) have mainly been chosen to enable real-world resemblance to be obtained in modeled scenarios, and we argue that this is the best way to define cost models in a simulation model like TAPAS. For instance, the model for estimating production costs allows the user to define three types of production strategies; batch production, continuous production, and instant retrieval of products from inventory. Moreover, two different models for setting transport prices are provided, and the user of TAPAS decides on case basis which model is best to use. However, there is room for improvement of the cost and pricing models. For instance, it could be relevant to provide the possibility to model product-dependent fixed costs and preparation times in terminals to be able to represent special requirements (time and other resources) that are needed for loading and unloading of particular product types. Also, it could be relevant to introduce fixed costs in a production line, e.g., to represent wastage when changing product types. However, this would require a more advanced model for determining production costs. Further, it would be relevant to extend and improve the models for calculating transport costs, e.g., since different types of time aspects often have an important influence on the decision making regarding which transport alternatives are most appropriate (Sommar and Woxenius, 2007). A possible extension would be to model punctuality as a cost component to let it have an influence on the decision making. Moreover, train transport costs include many different cost components, which interact in a rather complex way, and a possible extension would be to make use of some of the ideas in the train transport cost model proposed by Troche (2009). Similarly, the ship transport cost model could be further refined, e.g., by also taking into account fuel consumption at berth.

An important advantage of agent models is the modular structure, which enables adjustments of local models. To develop TAPAS to function even better as a general transport chain simulation model, a library of different types of models, strategies, and behaviors would facilitate the simulation of different scenarios. Moreover, the possibility to achieve benefits by advanced cooperation, such as joint planning within the transport chain, is an aspect that would be relevant to investigate further. As an example, the transport providers, as well as the customer could benefit from letting  $\mathcal{TB}$  negotiate directly with the customer regarding the delivery time window when there are cost efficient transport alternatives that have been disregarded due to penalties. To enable cooperation between decision makers to be captured in TAPAS, it would be necessary to extend the proposed interaction protocol. A possible approach would be to model agents as finite state machines, in which different types of behavior are allowed only when an agent is in certain states. This would make it easier to customize the communication between agents for particular scenarios, which would further

increase the generalizability of TAPAS.

Another aspect that would be interesting to consider is dynamic agent behavior, which would further improve the possibility to accurately mimic real-world situations. It would be interesting to consider agents that learn from experience how to improve their behavior. For instance, it would be possible for customer agents to improve their ordering strategies by dynamically updating their safety stock levels when it is realized that some product types frequently run out of stock. Moreover, it would be interesting to allow agents to dynamically enter or leave the system, which would make it possible to study how situations evolve over longer periods of time.

## 6 Conclusion and future work

We have presented a micro-level model for transport chain simulation. By using agent technology, we were able to simulate the decision making activities as well as the interaction between actors, which is difficult, if possible at all, using traditional techniques. By conducting a simulation study in a scenario around the Southern Baltic Sea, we have shown that it is possible to use TAPAS for studying various types of transport policy and infrastructure measures. Moreover, we have shown (1) that it is possible to deal with the complexities of mixing timetabled and demand driven transports in the algorithmic approach (2) that it is possible to simulate the principles of EOQ by letting the customer select the best order quantity among a set of possible quantities, and (3) the interaction framework is appropriate for capturing the ordering process in a multi-agent-based transport chain simulation model for analysis, e.g., of transport-related policy and infrastructure measures.

The studied scenario contains 12 agents; 3 customers, 1 production planner, 5 transport planners, 1 transport chain coordinator, 1 product buyer, and 1 transport buyer. We believe it is possible to scale up the size of scenarios to a few hundred producers and customers without reimplementing TAPAS, however, this need to be confirmed in simulation experiments. The goal is to be able to use TAPAS for studying larger regions containing thousands of producers and customers. A possible approach to improve the scalability would be to enable TAPAS to be executed on multiple processors.

Below we give a few pointers to further development of TAPAS. Extensions of the modeling on the agent level include: (1) integration of sophisticated optimization algorithms in the agents to improve the quality of their decisions, (2) allowing agents to communicate outside the proposed interaction protocol to enable enhanced cooperation between transport chain actors, (3) allowing agents to learn from experience, and (4) allowing agents to dynamically enter and leave the agent system. Examples of interesting extensions to the physical simulator are: (1) including external aspects that affect transportation, e.g., by simulating other traffic on the links, which would enable us to study the effects of link congestion, (2) simulating that real world ordering processing is time consuming (currently order processing is assumed to be instantaneous), and (3) allowing customer orders to be processed in parallel, which would require that aspects related to resource allocation in the ordering process need to be carefully considered.

## References

- Ahn, H., Kaminsky, P., 2005. Production and distribution policy in a two-stage stochastic push-pull supply chain. *IIE Transactions* 37 (7), 609–621.
- Alstrom, P., 2001. Numerical computation of inventory policies, based on the  $EOQ/\sigma_x$  value for order-point systems. *International Journal of Production Economics* 71 (1-3), 235–245.
- Axsäter, S., 2006. *Inventory Control*, 2nd Edition. Springer, New York, USA.
- Bellifemine, F. L., Caire, G., Greenwood, D., 2007. *Developing Multi-Agent Systems with JADE* (Wiley Series in Agent Technology). John Wiley & Sons, Chichester, England.
- Bergkvist, M., Davidsson, P., Persson, J. A., Ramstedt, L., 2005. A hybrid micro-simulator for determining the effects of governmental control policies on transport chains. In: *Lecture Notes in Computer Science: Vol. 3415. Multi-Agent and Multi-Agent-Based Simulation*. Springer, pp. 236–247.
- Boerkamps, J., van Binsbergen, A., 1999. Goodtrip - a new approach for modelling and evaluation of urban goods distribution. In: *Proceedings of City Logistics I. 12-14 July, Cairns, Australia*, pp. 175–186.
- Button, K., 1997. *Transport Economics*, 2nd Edition. Edward Elgar Publishing Limited, Glos, UK.
- Caramia, M., Guerriero, F., 2009. A heuristic approach to long-haul freight transportation with multiple objective functions. *Omega* 37 (3), 600–614.
- Chang, T. S., 2008. Best routes selection in international intermodal networks. *Computers & Operations Research* 35 (9), 2877–2891.
- Chatfield, D. C., Hayya, J. C., Harrison, T. P., 2007. A multi-formalism architecture for agent-based, order-centric supply chain simulation. *Simulation Modelling Practice & Theory* 15 (2), 153–175.

- Chen, M., Tardieu, P., 2000. The NEAC model, answering questions in a European context. Paper presented at THINK-UP seminar 1, Paris.
- Chow, J. Y. J., Yang, C. H., Regan, A. C., 2010. State-of-the art of freight forecast modeling: lessons learned and the road ahead. *Transportation* 37 (6), 1011–1030.
- Davidsson, P., Holmgren, J., Persson, J. A., Ramstedt, L., 2008. Multi agent based simulation of transport chains. In: *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems (AAMAS 2008)*. International Foundation for Autonomous Agents and Multiagent Systems, May 12-16, Estoril, Portugal, pp. 1153–1160.
- de Jong, G., Ben-Akiva, M., 2007. A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B* 41 (9), 950–965.
- de Jong, G., Gunn, H. F., Walker, W., 2004. National and international freight transport models: an overview and ideas for further development. *Transport Reviews* 24 (1), 103–124.
- European Commission, 2001. White paper - European transport policy for 2010: time to decide. Tech. rep., Luxemburg.
- Friberg, G., Flack, M., Hill, P., Johansson, M., Vierth, I., McDaniel, J., Lundgren, T., Hesselborn, P., Bångman, G., 2007. Kilometerskatt för lastbilar - effekter på näringar och regioner. Redovisning av ett regeringsuppdrag i samverkan med itps, rapport 2007:2, Swedish Institute for Transport and Communications Analysis (SIKA).
- Friedrich, H., Liedtke, G., 2009. Consideration of logistics for policy analysis with freight transport models. In: *8th Conference on Applied Infrastructure Research*. October 9-10, Berlin, Germany.
- Gjerdrum, J., Shah, N., Papageorgiou, L. G., 2001. A combined optimization and agent-based approach to supply chain modelling and performance assessment. *Production Planning & Control* 12, 81–88.
- Holmgren, J., 2011. An extended eastwest transport corridor (ewtc) scenario. Available at: [http://www.bth.se/tek/jhm.nsf/attachments/ewtc\\_pdf/\\$file/ewtc.pdf](http://www.bth.se/tek/jhm.nsf/attachments/ewtc_pdf/$file/ewtc.pdf).
- Holmgren, J., Davidsson, P., Persson, J. A., Ramstedt, L., 2007. An agent based simulator for production and transportation of products. In: *The 11th World Conference on Transport Research*. June 8-12, Berkeley, USA.
- Hunt, J., Stefan, K., 2007. Tour-based microsimulation of urban commercial movements. *Transportation Research Part B* 41 (9), 981–1013.
- Jansen, B., Swinkels, P., G.J.A., Teeuwen, B. d., Fleuren, H., 2004. Operational planning of a large-scale multi-modal transportation system. *European Journal of Operational Research* 156 (1), 41–53.
- Knabe, T., Schillo, M., Fischer, K., 2002. Improvements to the FIPA contract net protocol for performance increase and cascading applications. In: *International Workshop for Multi-Agent Interoperability at the German Conference on AI (KI-2002)*.
- Komma, V., Jain, P., Mehta, N., 2007. Agent-based simulation of a shop floor controller using hybrid communication protocols. *International Journal of Simulation Modelling* 6 (4), 206–217.
- Liedtke, G., 2009. Principles of micro-behavior commodity transport modeling. *Transportation Research Part E* 45 (5), 795–809.
- Ramstedt, L., 2008. Transport policy analysis using multi-agent-based simulation. Doctoral Dissertation Series No. 2008:09, Blekinge Institute of Technology, Sweden.
- Ramstedt, L., Davidsson, P., Holmgren, J., Persson, J., 2007. On the use of micro-level simulation for estimation of the effects of governmental control policies. In: *The 11th World Conference on Transport Research*. June 8-12, Berkeley, USA.
- Roorda, M. J., Cavalcante, R., McCabe, S., Kwan, H., 2010. A conceptual framework for agent-based modelling of logistics services. *Transportation Research Part E: Logistics and Transportation Review* 46 (1), 18–31.
- Sandholm, T. W., Lesser, V. R., 2001. Leveled commitments contracts and strategic breach. *Games and Economic Behaviour* (35), 212–270.
- Schillo, M., Kray, C., Fischer, K., 2002. The eager bidder problem: a fundamental problem of DAI and selected solutions. In: *Proceedings of the first international joint conference on Autonomous agents and multiagent systems (AAMAS 2002)*. ACM, July 15-19, Bologna, Italy, pp. 599–606.
- Shade, W., Martino, A., Roda, M., 1999. Assessment of transport strategies. In: *Proceedings of the 17th International Conference of the System Dynamics Society and the 5th Australian & New Zealand Systems Conference*. July 20-23, Wellington, New Zealand.
- SIKA, 2008. Vilken koldioxidkatt krävs för att nå framtida utsläppsmål? Pm 2008:4, Swedish Institute for Transport and Communications Analysis (SIKA).
- Silver, E., Meal, H., 1973. A heuristic for selecting lot size requirements for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment. *Production and Inventory Management* 14, 64–74.
- Smith, R. G., 1980. The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers* C-29 (12), 1104–1113.

- Sommar, R., Woxenius, J., 2007. Time perspectives on intermodal transport of consolidated cargo. *European Journal of Transport and Infrastructure Research* 7 (2), 163–182.
- Strader, T., Lin, F., Shaw, M., 1998. Simulation of order fulfillment in divergent assembly supply chains. *Journal of Artificial Societies and Social Simulation* 1 (2).
- Swahn, H., 2001. The Swedish national model systems for goods transport SAMGODS - a brief introductory overview. SAMPLAN-report 2001:1, Swedish Institute for Transport and Communications Analysis (SIKA).
- Swaminathan, J. M., Smith, S. F., Sadeh, N. M., 1998. Modeling supply chain dynamics: A multiagent approach. *Decision Sciences* 29 (3), 607–632.
- Tavasszy, L., 2006. Freight modelling - An overview of international experiences. In: *TRB Conference on Freight Demand Modelling: Tools for Public Sector Decision Making*. September 25-27, Washington DC, USA.
- Tavasszy, L., Smeenk, B., Ruijgrok, C., Nov. 1998. A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research* 5 (6), 447–459.
- Terzi, S., Cavalieri, S., 2004. Simulation in the supply chain context: a survey. *Computers in Industry* 53 (1), 3–16.
- Troche, G., 2009. Activity-based rail freight costing - a model for calculating transport costs in different production systems. Ph.D. thesis, KTH, Royal Institute of Technology, Sweden.
- van der Zee, D., van der Vorst, J., 2005. A modeling framework for supply chain simulation: Opportunities for improved decision making. *Decision Sciences* 36 (1), 65.
- Williams, I., Raha, N., 2002. Review of freight modelling. Tech. rep., Cambridge, UK, final Report DfT Integrated Transport and Economic Appraisal.
- Wilson, R., 1934. A scientific routine for stock control. *Harvard Business Review* 13 (1), 116–129.
- Wooldridge, M. J., Jennings, N. R., 1995. Intelligent agents: theory and practice. *The Knowledge Engineering Review* 10 (2), 115–152.
- Xuehui, W., Lei, Z., Nong, X., Yuhua, T., 2009. Time management in parallel discrete event simulation. In: *International Forum on Information Technology and Applications (IFITA 2009)*. IEEE, May 15-17, Chengdu, China, pp. 209–212.
- Zhou, H., Benton Jr., W., 2007. Supply chain practice and information sharing. *Journal of Operations Management* 25 (6), 1348–1365.
- Ziliaskopoulos, A., Wardell, W., 2000. An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays. *European Journal of Operational Research* 125 (3), 486–502.