Model predictive control for simultaneous planning of container and vehicle routes

Larsen, Rie B.; Atasoy, Bilge; Negenborn, Rudy R.

DOI
10.1016/j.ejcon.2020.06.003

Publication date
2021

Document Version
Final published version

Published in
European Journal of Control

European Journal of Control 57 (2021) 273–283

Contents lists available at ScienceDirect

European Journal of Control

journal homepage: www.elsevier.com/locate/ejcon

Model predictive control for simultaneous planning of container and vehicle routes

Rie B. Larsen*, Bilge Atasoy, Rudy R. Negenborn

Department of Maritime and Transport Technology, Delft University of Technology, Mekelweg 2, Delft 2628 CD, the Netherlands

ARTICLE INFO

Article history:
Received 26 August 2019
Revised 15 May 2020
Accepted 8 June 2020
Available online 12 June 2020

Recommended by Prof. T. Parisini

Keywords:
Synchromodal transport
Model predictive control
A-Modal booking
Container routing
Truck routing
Multimodal transport

ABSTRACT

Transport of containers on a-modal bookings enables transport suppliers to route the containers in accordance with the current state of the synchromodal transport network. At the same time, it enables the transport providers to route their vehicles in real time based on the current need for transportation. The interdependency of the routes of containers and of vehicles has not yet been discussed explicitly in the synchromodal literature. This paper presents a model predictive controller that determines which combination of trucks, trains, and ships to use for transporting the containers and what routes empty and full trucks should use as one integrated problem. The impacts of this integrated problem as opposed to only considering the routes of the containers are shown with experiments on a simulated synchromodal hinterland network performed with both the proposed method and with a method that solely routes the containers. The results indicate an improved vehicle utilization. Furthermore, the integrated problem approach allows for more realistic constraints and costs.

1. Introduction

The longer it takes from the moment a plan is made until it is implemented, the larger is the risk that something unexpected will happen. In container transport this unexpected event could be extreme weather delaying a barge, or an extra control check by customs delaying a container. Traditionally, such events are handled manually, hence making direct truck transport the easiest mode to use. Truck transport is however often the least environmentally friendly and the most man-hour consuming mode of transport. From an environmental, societal, and economical perspective it is therefore desirable to use other modes of transport such as rail and water instead. Multi-modal, intermodal and synchromodal transport, as well as the physical internet, supply chain logistics, etc. are all concepts that enable such a shift away from simplistic solutions and towards overall efficient solutions.

The shift towards an overall efficient approach creates new challenges on both the strategic, network design level, the tactical, flow scheduling level and on the operational, specific movements level. For synchromodal transport it can be argued that the time-horizon of decisions taken on the tactical level becomes closer to the time-horizon of decisions on the operational level [25], when the flows and services can be re-planned based on online information. A key enabler for this change is the concept of a-modal bookings where the service of transport is bought instead of a slot on a specific connection. This lets the transport supplier decide which modes and which vehicles are used to fulfil a specific transport order, and allows the supplier to change this decision during the execution of the transport.

It is however not enough to change decisions in real time, it is also necessary to take good decisions. Smart planning, disruption handling, dynamic switching, and demand aggregation are in [30] identified to be the four categories of necessary actions to obtain synchromodality. Real-time switching and integrated planning are also in the literature review [9] found to be among the 8 most important properties of synchromodality. It is thus agreed upon that the success of synchromodal transport is closely linked to the ability to switch plans when disturbances occur and the ability to plan container moves and equipment use simultaneously.

This paper presents such a framework which chooses modality and routes for containers simultaneously with routes and loading/unloading actions for trucks in real time. The framework uses model predictive control (MPC) to take decisions based on the latest available information with a conscious trade-off between the cost of transport for the containers and the utilization rates of the vehicles.

In the current literature on transport planning under uncertainty, transport suppliers create vehicle routes based on estimations of the demand. In [35], a static plan that accommodates uncertain future events is created by optimizing over

* Corresponding author.
E-mail addresses: r.b.larsen@tudelft.nl (R.B. Larsen), b.atasoy@tudelft.nl (B. Atasoy), r.r.negenborn@tudelft.nl (R.R. Negenborn).

https://doi.org/10.1016/j.ejcon.2020.06.003

different scenarios, while they in [27] are accommodated by using the probabilistic knowledge of the future events in an approximate dynamic programming method. Another approach is to plan truck flows and barge and train schedules ahead of time based on an assumed demand and handle undercapacity during implementation with expensive ad hoc alternatives (e.g., [2,33]).

In the literature there are very few attempts that directly plan container and vehicle routes simultaneously at the operational level. The authors of [32] state that "the flexibility in transportation routes may be used in conjunction with the operational fleet deployment problem. This creates new and more complex optimisation challenges", but the statement is not explored further. A planning model that besides container routes also decides if a specific service is operated or not is presented in [35]. The services are however not routed, which for a scenario with more import than export will lead to overcapacity of empty vehicles on the import side. In other words, the need for vehicles performing round-trips is not considered. In the container route planning model presented in [26], trucks are likewise modelled as links between locations which for a given time can be used or not. It is here taken into account that trucks may not always be available, but the model does not route the trucks. In [24], import containers, trucks, trains and barges are scheduled simultaneously by solving a mixed integer optimization problem. However, all vehicles, including trucks, have pre-determined routes and thus only the departure times are decided. In contrast, the current paper routes the trucks and handles both import and export containers.

Both container and vehicle planning problems have separately been studied extensively in the literature for several different transport systems. In [31], a comprehensive overview of the Operations Research planing models used in multimodal, intermodal, and synchromodal transport can be found. To route containers through a synchromodal network, Di Febbraro et al. [19] finds the k shortest paths through a network where barges and trains depart according to a schedule. This framework does not reconsider decisions on future actions automatically, but the ability to do so when disruptions occur is discussed. In [14], last minute decisions are used to route commodity flows online over a network with scheduled barge and train services, assuming truck capacity is infinite and instantly available. In [28], a similar problem is addressed by learning a preferred policy with Approximate Dynamic Programming. To obtain higher utilization rates of vehicles, the literature on dynamic vehicle routing problems combine pre-defined pick-up and delivery appointments in the most efficient way [23]. Most papers in this category do not relate themselves to intermodal or synchromodal transport. Some accommodate transshipments in their models (e.g., [4,8]) and cover thereby some of the challenges of intermodal transport planning.

The ability to change decisions during transport without confirmations from shippers as well as the increasing volumes to be transported motivate the use of control methods in container transport problems. Model predictive control (MPC) has already been used to address the container routing problem, but has not yet been used to integrate the planning of container and truck routes. In aforementioned [14], receding horizon control is used to plan the container flows in a hinterland network, but in contrast to the current paper, they only consider import and assume trucks are available when needed. In [13], that model is extended to the distributed case, where the geographical network is divided into non-overlapping regions served by different cooperating stakeholders. They consider commodity flows between multiple origins and destinations, but still assume trucks to be instantly available when needed. The container routing problem is furthermore solved distributed in [7] in an MPC-like framework. Trucks are hare considered instantly available and mainly used for last-mile transport.

MPC has also been used for planning and execution of related problems. It has been used to coordinate supply to demand in different supply chains (e.g. [10,17,22,34]). These models generally treat transport as a known input delay, without considering modes and timetables. Reis [1] and Wang and Rivera [21] employed MPC to improve efficiency inside container terminals. The former considers equipment as queues, and is only suitable for small geographical areas, as it does not consider the advantages of handling containers based on their geographical location. The latter considers trucks to be instantaneously available.

The current paper is an extension of the conference paper [12] with improved assumptions and additional simulated experiments that strengthen the conclusions. The MPC formulation has been modified to ensure recursive feasibility when unpredicted events change the truck travel time. In this paper, trucks can drive through nodes without unloading the container they carry and wait for vacant unloading capacity at the node. The results section has furthermore been extended to include several scenarios, two different demand profiles and both nominal cases and cases with uncertain truck travel times. The method's sensitivity to prediction horizon length is furthermore discussed. All simulations are performed on the multi-commodity, synchromodal transport network seen in Fig. 1, considering multi-type trucks as well as scheduled trains and barges.

The paper is organized as follows. In Section 2 the transport network model is introduced. Section 3 presents the control algorithm used for the simultaneous, real-time planning. Section 4 describes the simulation scenarios used to compare the proposed benchmark method to a real-time container routing method, which is presented in the same section. The results of the comparison are presented and discussed in Section 5. Finally in Section 6 the conclusions and directions for future research are discussed.

2. Model description

The transport network is modelled as a continuous state, discrete time, state-space commodity flow model of a hinterland network. The network is described by an undirected graph, where the nodes represent locations where containers are transferred between modes, locations where containers or trucks are stored or parked for longer periods of time, or scheduled services with high capacity. The arcs represent truck routes between physical locations or (un)loading actions for scheduled services. Vehicles and containers are modelled on separate networks that are coupled by the constraint that containers can only flow on a directed arc if there is at least the same number of trucks flowing on the same arc. If one of the nodes is a train or barge node (a scheduled service), no trucks are required. The main features of the model are:

- Demand is modelled as containers available to the network and needed from the network. Unsatisfied demand is penalized. At all timesteps the demand is fully known over the planning horizon.
- Commodity flows are considered to be continuous variables. This simplifies the model and can capture the desired level of accuracy, see [20].
- Unscheduled vehicles, with trucks as example, are also modelled as continuous variable flows. This again allows for balance between model complexity and accuracy.
- Each scheduled service is modelled separately. Two trains serving the same route are modelled as two nodes.
- A limited number of containers can be (un)loaded to trucks at a given node and a limited number can be loaded to and from the scheduled services at any given time.

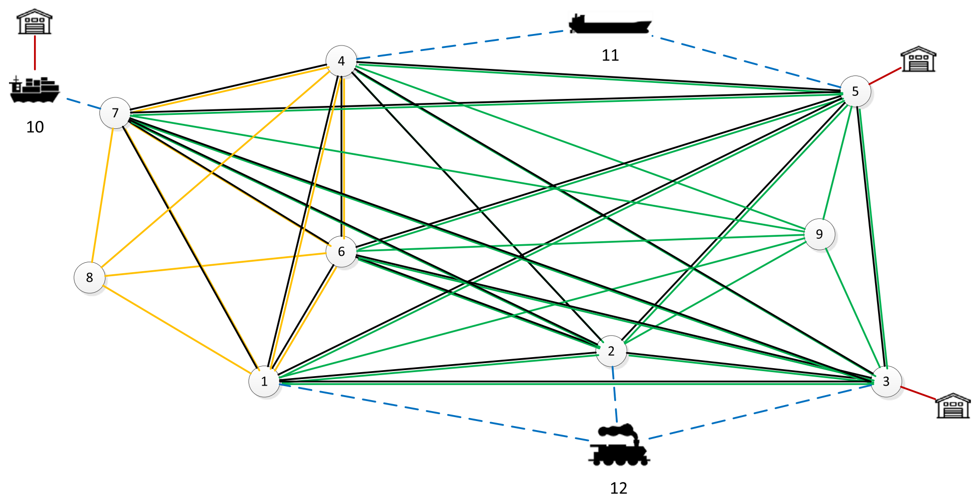

Fig. 1. Example network. Circles 1–9 and icons 10–12 are nodes of the system. Green and yellow lines are long and short distance truck networks, dashed lines indicate time dependent connections (connections to scheduled services) and red lines show the connections between network nodes and their adjacent virtual destination nodes.

- Trucks can wait at a node to be unloaded at a later time or drive through with its load.
- Travel times and capacity limits are known for the planning horizon at all timesteps.
- Terminal operating hours, truck drivers resting hours, and predictable travel time delays due to peak hours are not considered.

The model is an extension of the model presented in [12]. The current model ensures recursive feasibility of the MPC even when truck travel-times are uncertain. Trucks can here wait at nodes or drive through nodes without unloading and loading containers. This ensures that the capacity for unloading and loading trucks is not exceeded if delays cause multiple trucks to arrive at the same time. It furthermore brings the model closer to reality.

The model supports multi-commodity flows for both import and export. The demand profiles at the destinations are created based on time widows for each single container, but as commodity flows are considered, one container of a certain commodity can replace another (similar to the assumption in [14]). In [15,20] it is shown how this classification can be used to keep track on due dates and expiration dates. Trucks are modelled in the same fashion as containers, allowing to distinguish different kinds of vehicles. Each truck network includes a free parking node that represents the trucks that are not being used in the network but are available to the network.

The travel time and capacity limits can vary over time as time dependent parameters. This way, e.g., expected congestions can be modelled as time dependent increased travel times, and lower stacking height on barges due to high water levels can be modelled as time dependent decreased capacity. To simplify the notation they are used without a time indication in the model. It is assumed that when a travel is started, it is also fulfilled. In other words, no decisions can be taken when a truck or container is on an arc. The scheduled services (barge and train) are modelled as nodes with time dependent arc capacities that correspond to the timetable of the respective connection. When the scheduled service is at a terminal, it has a predetermined time slot to unload and hereafter a predetermined time slot to load before it departs according to schedule.

The mathematical description of the transport network is kept general, while the specifications of the network used as example can be found in Section 4 and Fig. 1.

The state x_i of each node $i \in \mathcal{N}$ in the system at every time step k is given by:

$$x_i(k) = \begin{bmatrix} x_i^c(k) \\ x_i^v(k) \\ u^h_{i,m_1}(k) \\ \vdots \\ u^h_{i,m_{n_v}}(k) \\ v_i^h(k) \end{bmatrix}, \tag{1}$$

where $x_i^c(k) \in \mathbb{R}^{n_c}_{\geq 0}$ is the quantity in number of containers of each of the n_c different commodities stacked at node i and $x_i^v(k) \in \mathbb{R}^{n_v}_{\geq 0}$ is the quantity of each of the n_v different truck types parked at node i. In this paper superscripts are used to distinguish variables with similar functions, while subscripts are used for indexing the variables. Notice that most variables are vectors such that different commodities are represented by different elements in the vector. It is for simplicity assumed that all containers are of the same size and that all truck types can transport one container. However, these assumptions can be overcome by introducing additional commodities for containers of different sizes and vehicle capacities different than 1. The vector $u^h_{i,m}(k)$ is the amount in containers of each commodity that are on the way to node i by a truck of type m at time step k. It is necessary to keep a record of the containers that are on the way to node i but have not yet arrived, since each arc in the truck network is associated with a travel time τ_{ji} that acts as a delay. Formally, $u^h_{i,m_1}(k) = [u_{ji,m_1}(k-1)^T \dots u_{ji,m_1}(k-\tau_{ji})^T \dots u_{j'i,m_1}(k-1)^T \dots u_{j'i,m_1}(k-\tau_{j'i})^T]^T$, $\{j \cdots j'\} = \mathcal{T}_i$, $\{m_1 \cdots m_{n_v}\} = [1, n_v]$, where $u_{ji,m}(k) \in \mathbb{R}^{n_c}_{\geq 0}$ is the volume in containers of each commodity that leave node j at time step k on the arc to node i using truck type m. The set $\mathcal{T}_i$ contains all nodes with a truck connection to node i. Likewise, $v_i^h(k) = [v_{ji}(k-1)^T \dots v_{ji}(k-\tau_{ji})^T \dots v_{j'i}(k-1)^T \dots v_{j'i}(k-\tau_{j'i})^T]^T, \{j \cdots j'\} = \mathcal{T}_i$ is the amount of trucks of the different types that are on the way to node i at time step k. Here, $v_{ji}(k) \in \mathbb{R}^{n_v}_{\geq 0}$ is the amount of trucks that leave node j towards i at time step k.

The demand is modelled on virtual destination nodes $d \in \mathcal{D}$ that are adjacent to network nodes. The virtual destination nodes are copies of the network nodes, which instead of modelling the container flows model the satisfaction and accumulation of new demand. It is thus possible for a container to arrive at the network node corresponding to its destination before it is used to satisfy the demand at the virtual destination node. The arc between a virtual destination node and its adjacent network node has unlimited capacity and zero travel time, letting demand being satisfied unrestricted as soon as containers arrive at the network node. The unsatisfied demand (both available and needed containers) at the virtual destination nodes is penalized, while containers stacked at the network node waiting for demand to satisfy are only accumulating storage costs and taking up stack space. We say that node i has outgoing demand when i is the origin of the commodity and that node i has incoming demand when i is the destination. The virtual destination nodes have different dynamics than the nodes in the network, namely

$$x_i^d(k+1) = x_i^d(k) - u_{di}(k) - u_{id}(k) + d_i(k), \tag{2}$$

where $x_i^d(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is the amount of incoming and outgoing demand in containers of each commodity at time step k. Both incoming and outgoing demand are modelled as positive values, since the commodities are defined based on destination. The variable $u_{id}(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is the containers that were available at network node i that are used to satisfy the incoming demand at time step k, and likewise, $u_{di}(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is the containers used to satisfy the outgoing demand. Demand satisfaction can be postponed (hence the integral dynamics), and the new demands $d_i(k) \in \mathbb{R}_{\geq 0}^{n_c}$, that can be satisfied from time step k, act as disturbances to the system and are thus not controllable.

The remaining nodes in the network are described as in (1) and have the same dynamics. For describing the dynamics three sets are defined for each node i: $\mathcal{T}_i$ as introduced earlier, $\mathcal{S}_i$ and $\mathcal{D}_i$. The set $\mathcal{S}_i$ contains all nodes to which i is linked via a time-dependent arc connection. If node i is a scheduled service, $\mathcal{S}_i$ contains the terminals it serves, and if node i is a terminal, $\mathcal{S}_i$ contains the scheduled services that depart from here. Notice that if i is a scheduled service $\mathcal{T}_i = \emptyset$. Likewise $\mathcal{D}_i$ contains the adjacent destination node for node i. This set contains maximum one element. The dynamics of $x_i^c(k)$ is

$$\begin{aligned} x_i^c(k+1) =& x_i^c(k) + \sum_{m\in[1,n_v]} \sum_{j\in\mathcal{T}_i} \left(u_{ji,m}(k-\tau_{ji}) - u_{ij,m}(k)\right) \\ &+ \sum_{s\in\mathcal{S}_i} (u_{si}(k) - u_{is}(k)) + \sum_{d\in\mathcal{D}_i} (u_{di}(k) - u_{id}(k)), \end{aligned} \tag{3}$$

where the control action $u_{is}(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is the containers moved from node i over a time-dependent connection to node s. If node i is a barge, $u_{is}(k)$ is unloading containers at terminal s. $u_{si}(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is the reverse movement.

As there are no scheduled services nor demand in the truck network the dynamics hereof is given by:

$$x_i^v(k+1) = x_i^v(k) + \sum_{j\in\mathcal{T}_i} \left(v_{ji}(k-\tau_{ji}) - v_{ij}(k)\right). \tag{4}$$

The two networks are connected by the constraint that containers cannot be moved without a truck if they are transported on a truck-arc.

$$\sum_{m\in[1,n_v]} \mathbf{1}_{n_c} u_{ij,m}(k) \leq \mathbf{1}_{n_c} v_{ij}(k) \;\forall\; j \in \mathcal{T}_i. \tag{5}$$

The bold $\mathbf{1}_a = \{1\}^a$ is a row vector of size a with all ones.

The network is furthermore constrained by capacities:

$$\mathbf{1}_{n_c} x_i^c(k) \leq c_i^c \tag{6}$$

$$x_i^v(k) \leq c_i^v \tag{7}$$

$$-c_i^m \leq \sum_{m\in[1,n_v]} 1_{n_c} \sum_{j\in\mathcal{T}_i} \text{abs}\left(u_{ji,m}(k-\tau_{ji}) - u_{ij}(k)\right) \leq c_i^m \tag{8}$$

$$\mathbf{1}_{n_c} u_{si}(k) \leq c_{si}(k)\;,\; s \in \mathcal{S}_i \tag{9}$$

$$\mathbf{1}_{n_c} u_{is}(k) \leq c_{is}(k)\;,\; s \in \mathcal{S}_i, \tag{10}$$

where, at location i, the scalar c_i^c is the maximum number of containers that can be stored, $c_i^v \in \mathbb{R}_{\geq 0}^{n_v}$ is the maximum number of vehicles of each kind that can be parked ((7) is to be satisfied element wise). The notation abs$(\cdot)$ is the element-wise absolute value of a vector. The scalar c_i^m is the maximum number of containers that can be moved to and from trucks within one time step at location i. Notice that this constraint is the crane capacity and thus does not effect containers that remain on the same truck. Trucks can thus drive through nodes without limitations. This is different from the model presented in [12]. The schedules of the barge and train connections are implemented by the time varying crane speeds $c_{si}(k)$ and $c_{is}(k)$. To illustrate, assume i is a barge and s is a terminal. When the barge is at the terminal and can be unloaded $c_{si}(k) = 0$ and $c_{is}(k) \neq 0$, and when the barge can be loaded $c_{si}(k) \neq 0$ and $c_{is}(k) = 0$, otherwise $c_{si}(k) = 0$ and $c_{is}(k) = 0$.

3. Proposed control method

To achieve an efficient execution of container transport and truck routing that can adapt to delays online, a convex MPC is proposed. The control variables are, for all $i \in \mathcal{N}$, the amount of departing trucks and the containers they bring, $v_{ij}(k)$, $\forall\, j \in \mathcal{T}_i$ and $u_{ij,m}(k)$, $\forall\, j \in \mathcal{T}_i$, the quantity to load and unload for scheduled services, $u_{ij}(k)$, $\forall\, j \in \mathcal{S}_i$, and the amount of demand to satisfy $u_{di}(k)$ and $u_{id}(k)$, $d \in \mathcal{D}_i$.

The proposed control model is based on Ref. [12] and extended, such that trucks can arrive to a node and continue driving without unloading the container it carries. Trucks with containers are furthermore able to wait at a node until the crane is available to unload them, this is modelled as a road leading back to the same node the next timestep, hence $\tau_{ii} = 1\;\forall\, i \in \mathcal{N}$. This is a more realistic assumption than what was used in [12].

It is assumed that the controller has an accurate model for the dynamics of the transport system, and access to accurate information of the state of the global system every ΔT minute. Furthermore, a prediction of the future demand is assumed available to the controller. At each time $t = i\Delta T$, $i \in \mathbb{N}$ the controller gets up to date information and uses it to find the sequence of decisions that will minimize a cost function over a prediction horizon T_p. Only the decisions that require an action at this timestep $t = i\Delta T$ are implemented, and when $t = (i+1)\Delta T$, the process starts over.

The dynamics presented in Section 2 is known by the controller, but since only trucks that load or unload containers require crane movements, the decision variable $z_{i,m}(k) \in \mathbb{R}_{\geq 0}^{n_c}$ is introduced to represent the containers departing at timestep k from node i on the same vehicle which they arrived with and have not been unloaded from. This way (8) can be formulated as a convex constraint. Only crane movements are restricted and bare a cost. Containers arriving and leaving on the same trucks do not. The subscript $m \in [1, n_v]$ denotes the different vehicle types.

The cost to be optimized by the MPC is the total cost of transporting the containers. It is assumed that the transport provider has pre-approved all incoming orders, which means that the deadline and payment from the shipper for each container is fixed. The planning tool should thus minimize the cost the transport provider needs to pay to fulfil the accepted orders, namely storing of containers, (un)loading of vehicles, slots on scheduled services, movement of trucks and parking of trucks. It is assumed that there is a central planner that can decide which plan will be followed. To evaluate what the best sequence of decisions is, the

MPC controller solves the optimization problem (11)–(17), where the measured state (1) for node i at time t is denoted by $\tilde{x}_i(t)$. The decision vector U contains all inputs $u_{ij,m}(k)$, $v_{ij}(k)$, $u_{id}(k)$ and $u_{di}(k)$ for all $i \in \mathcal{N}$ and $k \in [0, T_p - 1]$. The time-invariant weight M_i^c is the cost of storing a container at node i, while M_i^v is the cost of parking a truck. M_{ij}^t is the cost of a truck journey from i to j and M_i^l is the cost associated with moving a container from a stack to a truck or vice versa. Moving a container to or from a scheduled service has the cost M_i^s, which is only paid at the terminals. Transport by scheduled service is paid per container per time step as the container storage cost M_i^c. The cost of unsatisfied demand is a quadratic term scaled by M_d^i, which lets less delays be significantly cheaper than more delays.

$$\min_U \sum_{k=0}^{T_p} \left(\sum_{i \in \mathcal{N}} \left(M_i^c x_i^c(k) + M_i^v x_i^v(k) + \sum_{j \in \mathcal{T}_i} M_{ij}^t v_{ij}(k) \right.\right.$$
$$+ \sum_{m \in [1,n_v]} M_i^l \left(\sum_{j \in \mathcal{T}_i} \left(u_{ij,m}(k) + u_{ji,m}(k - \tau_{ji}) \right) - 2 z_{i,m}(k) \right)$$
$$\left.\left. + \sum_{s \in \mathcal{S}_i \cap \mathcal{N}_i} \left(M_i^s (u_{si}(k) + u_{is}(k)) \right) \right) + \sum_{i \in \mathcal{D}} (x_i^d(k))^T M_i^d x_i^d(k) \right) \tag{11}$$

$$\text{s.t} \quad (2)-(7), (9), (10) \quad \forall\, i \in \mathcal{N}, \quad \forall\, k \in [0, Tp - 1] \tag{12}$$

$$z_{i,m}(k) \le \sum_{j \in \mathcal{T}_i} u_{ij,m}(k) \quad \forall\, i \in \mathcal{N},\ \forall\, m \in [1, n_v],\ \forall\, k \in [0, Tp - 1] \tag{13}$$

$$z_{i,m}(k) \le \sum_{j \in \mathcal{T}_i} u_{ji,m}(k - \tau_{ji}) \forall\, i \in \mathcal{N}, \quad \forall\, m \in [1, n_v],\ \forall\, k \in [0, Tp - 1] \tag{14}$$

$$\sum_{m \in [1,n_v]} \sum_{j \in \mathcal{T}_i} \left(u_{ij,m}(k) + u_{ji,m}(k - \tau_{ji}) \right) - 2 z_{i,m}(k) \le c_i^m \quad \forall\, i \in \mathcal{N} \tag{15}$$

$$v_{ij}(k) = 0 \quad \forall\, i \in \mathcal{N},\ \forall\, j \in \mathcal{T}_i, \quad \forall\, k > T_p - \tau_{ij} \tag{16}$$

$$x_i(k = 0) = \tilde{x}_i(t) \quad \forall\, i \in \mathcal{N} \tag{17}$$

Typically, MPC ensures recursive feasibility of the optimization problem and stability of the controlled system by special constraints and costs at the end of the prediction horizon [18]. The synchromodal transport system described in this paper is inherently marginally stable and recursively feasible, but as the actions taken within the prediction horizon will effect the state of the system in the future and thus the long-term (infinity) cost, considerations regarding the two concepts are important. The methods to address these challenges often impose conservatism that will cause underutilization of the scheduled services in the synchromodal transport problem, see,e.g., [5]. A way to address the long-term cost of the MPC problems, when no formulation of the expected infinity costs and constraints exist, is to use a long prediction horizon, see, e.g., Guo et al. [3] or Van Riessen et al. [6]. The current literature on this assumes different symmetric cost functions around a reference point (here the global zeros-state) that lies in the interior of the feasible set. If the transport cost is formulated based on absolute numbers and the reference point is set to be a vector of very small positive numbers instead the origin, then the assumptions hold and only the time-varying constraints prevent a calculation of the necessary length of the prediction horizon. To ensure the controller sees the consequences of its decisions, only trucks that will arrive within the prediction horizon are allowed to depart (16).

Table 1
Costs parameters.

$M_i^v = 1 \cdot \mathbf{1}_{n_v}\ \forall i \in [1, 7]$	$M_i^v = 0 \cdot \mathbf{1}_{n_v}\ \forall i \in [8, 9]$
$M_i^c = 1.2 \cdot \mathbf{1}_{n_c}\ \forall i \in [1, 7] \backslash \{6\}$	$M_6^c = 0.12 \cdot \mathbf{1}_{n_c}$
$M_{11}^c = 1.2 \cdot \mathbf{1}_{n_c}$	$M_{12}^c = 1.6 \cdot \mathbf{1}_{n_c}$
$M_i^l = 3 \cdot \mathbf{1}_{n_c}\ \forall i \in [1, 7]$	$M_i^s = 3 \cdot \mathbf{1}_{n_c}\ \forall i \in [1, 5]$
$M_{ij}^t = \tau_{ij} \cdot 3 \cdot \mathbf{1}_{n_c}\ \forall i, j \in [1, 9]$	$M_3^d = 30$
$M_5^d = 30$	$M_{10}^d = 30$

4. Simulation experiments

To evaluate the potential benefits of simultaneous routing of containers and trucks, simulation experiments of hinterland transport scenarios have been carried out. The experiments are performed both with the planning method presented in Section 3 that determines container and truck routes simultaneously and with a benchmark method that considers truck capacity to be infinite and instantly available. To focus on the added value of simultaneous routing, the benchmark method is an MPC-based method that has the same parameters and constraints as the proposed method except for the cost structure and assumptions related to the movement of empty trucks. In each experiment, the applied control method decides the routing over 600 timesteps. The simulations were performed in Matlab with Yalmip [16] and Gurobi.

In this section, first the parameters of the MPC are discussed, then the benchmark method is introduced followed by descriptions of the hinterland transport scenarios.

4.1. MPC parameters

The choice of costs and prediction horizon has significant impact on the MPC's resulting control since the MPC's prediction horizon is finite and without estimations of the infinity cost. For the presented results, the proportional costs shown in Table 1 are used. They are chosen to reflect the expenses from a system-wide perspective. To encourage movement and capture the cost of unnecessary crane-moves at small stacks, the costs of stacking containers and parking trucks are fairly high except at the central stack (node 6) and the parking lots (node 8 and 9), respectively. In the literature this cost is often either disregarded, e.g. in [24] or very low, e.g. [14] where it is less than 0.1% of the hourly transport cost by barge. The results presented in Section 5 are simulated based on an update rate of $\Delta T = 15$ minutes and a prediction horizon $T_p = 80$ timesteps. In the remainder of this paper, time is measured in timesteps, not minutes, as the model does not consider opening hours.

Different prediction horizon lengths allows the MPC to take different events into consideration. The longer the horizon is, the better overview over available connections the MPC will have. However, as increasing the prediction horizon length leads to increased computation time, a trade-off has to be established. It is generally advisable to choose prediction horizons long enough that the MPC can foresee both departure and arrival of the scheduled services at all times and truck roundtrips from parking node over container origin to container destination. For one of the scenarios that will be introduced in Section 4.3 (P1U with deterministic travel time) the realized costs when using simultaneous planning is shown for different prediction horizon lengths in Fig. 2. The bar diagram shows that for smaller prediction horizons ($T_p = 40$ and $T_p = 50$) the MPC cannot foresee the benefit of sending an empty

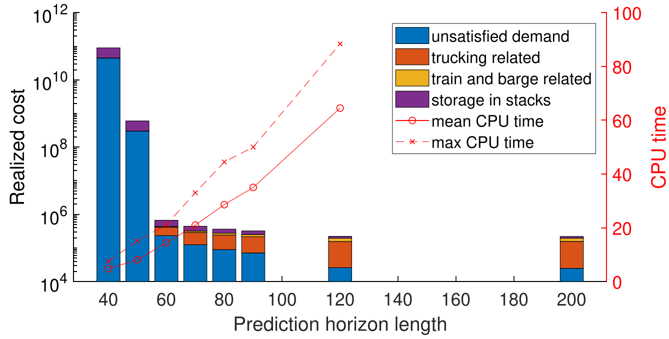

Fig. 2. Comparison of realized cost and computation time in seconds per timestep for different prediction horizon lengths for scenario P1 with unbalanced demand.

Table 2
Travel times on truck networks in time steps.

		End node								
		1	2	3	4	5	6	7	8	9
Start node	1	1	15	27	1	25	1	1	1	30
	2	15	1	12	25	20	14	16	–	20
	3	27	12	1	23	10	28	28	–	5
	4	1	15	25	1	23	1	1	1	30
	5	25	20	10	23	1	23	23	–	5
	6	1	14	28	1	23	1	1	1	30
	7	1	16	28	1	23	1	1	1	30
	8	1	–	–	1	–	1	1	1	–
	9	30	20	5	30	5	30	30	–	1

truck from its initial parking spot at node 9 to pick up an import container in node 7, as the delivery of that container in the inland terminals will lie outside the predicted future. When $T_p = 50$ the MPC can however predict the delivery of an import container in virtual demand node 5, if an empty truck is send from node 5 to pick up said container. For longer prediction horizons, where the MPC can foresee truck roundtrips to all destinations, the cost still decreases when the prediction horizon gets longer. The portion of the cost that is used on transport compared to the cost of unsatisfied demand also increases with increased prediction horizons. However, as the computation time increases with the prediction horizon, upcoming simulation experiments are limited to $T_p = 80$.

4.2. Benchmark method

To illustrate the impact of performing simultaneous planning, the proposed method is compared to a benchmark method that assumes trucks are instantaneously available and only optimizes the container routes in the hinterland network. The benchmark method is an MPC controller with the same update rate and prediction horizon as the proposed method. This ensures that the differences in the results obtained by the two methods only show the impact of considering container and truck routes simultaneously compared to assuming trucks instantaneously available. For the same reason, the constraints of the proposed method are used in the MPC problem of the benchmark method. Instantly available trucks are implemented in the benchmark method by ensuring sufficient capacity is available at all times in all nodes and and by assigning the travel cost M^t_{ij} to the transported container instead of the truck. In the literature, it is common to assign travel costs this way (e.g. [24]). Furthermore, the handling cost M^l_i in the benchmark model is charged per departing truck to discourage the movement of empty trucks. The benchmark MPC solves thus the optimization problem (18) and (19).

$$\min_{U} \sum_{k=0}^{T_p} \left(\sum_{i\in\mathcal{N}} \left(M^c_i x^c_i(k) + \sum_{j\in\mathcal{T}_i} \left(\sum_{m\in[1,n_c]} M^t_{ij} u_{ij,m}(k) + M^l_i v_{ij}(k) \right) + \sum_{s\in\mathcal{S}_i\cap\mathcal{N}_i} \left(M^s_i (u_{si}(k) + u_{is}(k)) \right) \right) + \sum_{i\in\mathcal{D}} (x^d_i(k))^T M^d_i x^d_i(k) \right) \tag{18}$$

$$\text{s.t} \qquad (12), (13), (14), (15), (16), (17) \tag{19}$$

4.3. Simulation scenarios

The hinterland transport network used for the simulation experiments can be seen in Fig. 1. It consists of three virtual destinations: one adjacent to ship connections, and two adjacent to inland terminals, from where last-mile delivery and pick-up are assumed to be arranged. The ships arrive and depart according to a predetermined schedule. The network has a barge and a train connection with fixed schedules. In the port area (between node 1,4,6,7,8) port vehicles transport the containers (yellow network), while long distance trucks are responsible for the remaining routes (green network). In this example the two truck networks are not overlapping, but the proposed planning method is able to address overlaps as well. The travel times τ_{ij} for both networks can be seen in Table 2.

The initial container state and the initial states of arriving containers and vehicles are zero in all scenarios, $\tilde{x}^c_i(t=0) = \mathbf{0}_{n_c}$, $\tilde{u}^h_i(t=0) = \mathbf{0}_{n_c}$ and $\tilde{v}^h_i(t=0) = \mathbf{0}_{n_v}$ $\forall i \in \mathcal{N}$.

The demand profiles for the virtual demand nodes 3, 5, and 10 were generated based on individual transport orders with an allowable lead time of minimum 40 time steps. Two different demand profiles were used, as seen in Fig. 3. One where significantly more containers are imported (destination 3 and 5) than exported (destination 10/ship), and one where the import and export are proportional, which are referred to as unbalanced and balanced demand, respectively. When empty and full containers are considered as disconnected problems (e.g. [11]) the unbalanced demand is more common in Europe. Balanced demand is on the other hand more representative if movements of all containers are considered in one problem (e.g. [29]). It is assumed that the controller at all timesteps has access to an accurate demand prediction for the prediction horizon.

Three scenarios with different levels of available resources have been used in the experiments. In the first scenario the constraints on container storage, truck parking, truck (un)loading and ship (un)loading are sufficiently loose that they do not become active during the simulation experiments. These constraints are all tightened to restrictive levels for the second scenario. The third scenario combines the constraints from the second scenario with limitations on the number of trucks available. Since the third scenario requires control of the total number of trucks in the system, the benchmark method is, like most methods in synchromodal transport literature, not applicable and only simulation results for the proposed method are presented.

A summary of the differences between the available resources in the three scenarios can be found in Table 3 together with the non-zero initial states. The train and barge schedules, capacity and maximum (un)loading rates are the same for all three scenarios with the following values: $c^c_{11} = 80$, $c_{11i} = c_{i11} \in \{0, 50\}\ \forall i \in \{4, 5\}$ and $c^c_{12} = 45$, $c_{12i} = c_{i12} \in \{0, 30\}\ \forall i \in \{1, 2, 3\}$ for the barge and train respectively.

If a truck is delayed, then not only the container it currently transports is affected, but also the containers it was scheduled to transport in the future. The MPC can react to delays and reschedule, such that other trucks transport the most urgent containers. This, however, require that other trucks are available. The simulations from this paper are therefore performed for both the

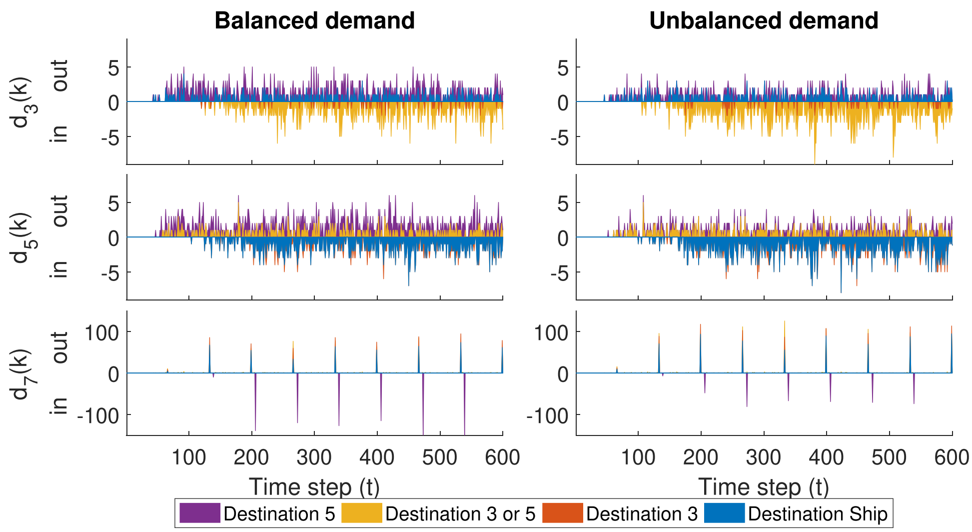

Fig. 3. Demand profile at the three virtual destination nodes for both balanced and unbalanced demand. The quantity of new demand $d_i(k)$ is shown over time steps. Outgoing demand is shown as positive and the incoming demand as negative.

Table 3
The difference in parameters and initial conditions between the performed simulations. The five scenarios are evaluated using both the balanced and the unbalanced demand profiles.

		Resources		
		Unrestricting $c_i^c = 1000\ \forall i \in [1, 7]$ $c_i^m = 300\ \forall i \in [1, 7]$ $c_{12,7}, c_{7,12} \in \{0,\ 200\}$	**Tight** $c_i^c = 75\ \forall i \in [1, 5]$ $c_i^c = 250\ \forall i \in \{6, 7\}$ $c_i^m = 20\ \forall i \in [1, 5]$ $c_i^m = 50\ \forall i \in \{6, 7\}$ $c_{12,7}, c_{7,12} \in \{0,\ 40\}$	**Limited trucks** $c_i^c = 75\ \forall i \in [1, 5]$ $c_i^c = 250\ \forall i \in \{6, 7\}$ $c_i^m = 20\ \forall i \in [1, 5]$ $c_i^m = 50\ \forall i \in \{6, 7\}$ $c_{12,7}, c_{7,12} \in \{0,\ 40\}$
Method	**Proposed** $c_8^v = [1000 \quad 0]^T$ $c_9^v = [0 \quad 1000]^T$ $\tilde{x}_i^v(0) = [0 \quad 0]^T\ \forall i \in \{1, 7\}$	Scenario **P1** $c_i^v = [100 \quad 100]^T\ \forall i \in \{1, 4, 6, 7\}$ $e_i^v = [0 \quad 100]^T \forall \in \{2, 3, 5\}$ $\tilde{x}_8^v(0) = [1000 \quad 0]^T$ $\tilde{x}_9^v(0) = [0 \quad 1000]^T$	Scenario **P2** $c_i^v = [10 \quad 10]^T \forall i \in \{1, 4, 6, 7\}$ $c_i^v = [10 \quad 10]^T\ \forall b \in \{2, 3, 5\}$ $\tilde{x}_8^v(0) = [1000 \quad 0]^T$ $\tilde{x}_9^v(0) = [0 \quad 1000]^T$	Scenario **P3** $c_i^v = [10 \quad 10]^T \forall i \in \{1, 4, 6, 7\}$ $\tilde{x}_i^v(0) = [0 \quad 0]^T\ \forall b \in \{2, 3, 5\}$ $\tilde{x}_8^v(0) = [8 \quad 0]^T$ $\tilde{x}_9^v(0) = [0 \quad 75]^T$
	Benchmark $c_8^v = [0 \quad 0]^T$ $c_9^v = [0 \quad 0]^T$ $\tilde{x}_8^v(0) = [0 \quad 0]^T$ $\tilde{x}_9^v(0) = [0 \quad 0]^T$	Scenario **B1** $c_i^v = [0 \quad 1000]^T\ \forall i \in \{1, 4, 6, 7\}$ $c_i^v = [0 \quad 1000]^T\ \forall i \in \{2, 3, 5\}$ $\tilde{x}_i^v(0) = [500 \quad 500]^T\ \forall i \in \{1, 4, 6, 7\}$ $\tilde{x}_i^v(0) = [500 \quad 500]^T\ \forall i \in \{2, 3, 5\}$	Scenario **B2** $c_i^v = [1000 \quad 1000]^T\ \forall i \in \{1, 4, 6, 7\}$ $c_i^v = [0 \quad 1000]^T\ \forall i \in \{2, 3, 5\}$ $\tilde{x}_j^v(0) = [500 \quad 500]^T\ \forall i \in \{1, 4, 6, 7\}$ $\tilde{x}_j^v(0) = [500 \quad 500]^T\ \forall i \in \{2, 3, 5\}$	Not applicable

case without delays (nominal case) and the case where the trucks may be delayed (uncertain case). The delays are not predicted by the MPC, but are added to the system by changing the distribution of the incoming container and truck flows, $u_i^h(k)$ and $v_i^h(k)$. At each timestep, a percentage of the trucks that are almost arriving is delayed exactly one timestep. One truck can be delayed several times, this corresponds to the trucks not informing the central planner in advance if they foresee a longer delay. It is assumed that the probability that an empty truck is delayed is the same as the probability that a loaded truck is delayed, hence the differences between the predicted state at time k computed at $k-1$ and the measured state are $\tilde{u}_{ji,m}(t-\tau_{ji}) = \beta u_{ji,m}(-\tau_{ji}|t-1)$, $\tilde{u}_{ji,m}(1-\tau_{ji}) = (1-\beta)u_{ji,m}(-\tau_{ji}|t-1) + u_{ji,m}(1-\tau_{ji}|t-1)$, $\tilde{v}_{ji}(t-\tau_{ji}) = \beta v_{ji}(-\tau_{ji}|t-1)$, and $\tilde{v}_{ji}(1-\tau_{ji}) = (1-\beta)v_{ji}(-\tau_{ji}|t-1) + u_{ji}(1-\tau_{ji}|t-1)$ where β is the share of the trucks that are delayed and the notation $a(k|t)$ indicates the prediction of $a(k)$ computed at time t. The parameter β is drawn per arc in the network per timestep k from the truncated normal distribution seen in Fig. 4.

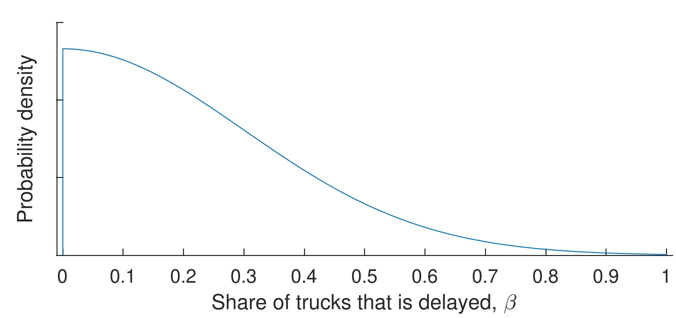

Fig. 4. The probability that a share of the trucks on an arc is delayed by 15 minutes. If no trucks are delayed $\beta = 0$ and if all trucks are delayed $\beta = 1$.

5. Results and discussion

This section presents the results of the comparison between the proposed method that considers the movements of trucks and

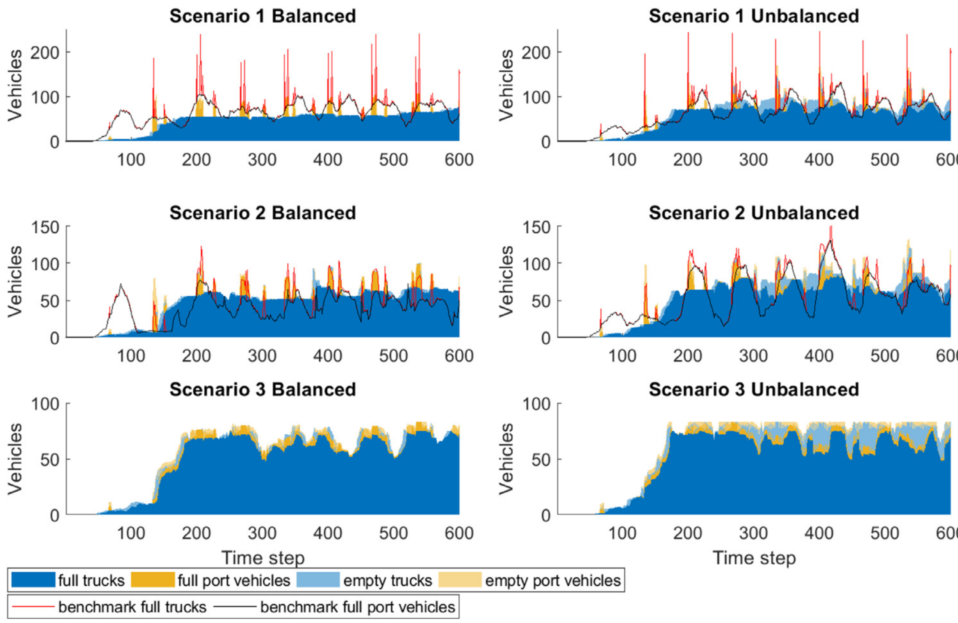

Fig. 5. Number of driving vehicles $\sum_{i\in\mathcal{N},j\in\mathcal{N}}\sum_{l=0}^{\tau_{ij}-1} v_{ij}(k-l)$, i.e., for each time step the y-axis shows the number of vehicles that has departed a node but has not yet reached the next node. The shaded area is the number of vehicles for the proposed method with the light part being the portion of empty vehicles. The lines are the full vehicles when using the benchmark method. The information is stacked for each method.

Table 4
The average and maximum CPU time for the MPC to compute outputs at one time step t. No highlight: nominal case. Grey highlight: case with uncertain travel time.

	Balanced demand					Unbalanced demand				
	P1	P2	P3	B1	B2	P1	P2	P3	B1	B2
Average CPU time (s)	31	29	26	23	25	29	27	26	22	22
	27	30	29	22	25	26	26	29	22	23
Maximum CPU time (s)	42	41	34	30	38	44	36	39	31	31
	44	53	37	29	42	43	35	47	30	30

containers simultaneously and the benchmark method that assumes infinite and instant truck capacity. Both methods compute what actions to take fast due to the optimization problems' convex nature. The average and maximum time the MPCs needed to compute the inputs at any timestep t are shown in Table 4 for all scenarios. The benchmark method is faster than the proposed method, but both are significantly faster than the chosen update rate of $\Delta T = 15$ min, i.e., the real-time performance is guaranteed. In a real-world implementation, one could thus choose to increase the prediction horizon or decrease the update rate. The trends from the results presented in this section are expected to hold in such cases too.

The results show very significant differences in the utilization of the transport modes for the proposed and the benchmark methods. The utilization do however not differ much between the nominal scenarios and their counterparts with uncertain travel time. Hence, in the following only the vehicle utilization in the nominal scenarios will be discussed.

In Fig. 5 the results for the nominal scenarios with unbalanced and balanced demand are presented. The solid color blocks show the results for the proposed method, while the lines show the results produced by the benchmark method. The dark blue area indicates how many hinterland trucks were transporting containers at a given time, while the translucent blue indicates how many empty hinterland trucks were driving in the network. The yellow and translucent yellow show the same for the port vehicles. In scenario 3 with unbalanced demand profile all 75 hinterland trucks and 8 port vehicles are driving either empty or full at nearly all timesteps $k > 200$. For the benchmark method, only information about the vehicles that transport containers exists due to the assumption of instant and infinite truck capacity. Thus for the benchmark method no information on empty vehicles is available to be shown. The full hinterland trucks are shown in black and the full port vehicles are shown in red. The information is stacked in the same manner as the information for the proposed method, i.e. for scenario 1 with unbalanced demand the benchmark method used 72 hinterland vehicles and 173 port vehicles to transport containers at timestep $k = 400$. Notice that all experiments start with an empty system and a slowly increasing demand profile.

In all scenarios, less containers are moved by port vehicles when containers and trucks are routed simultaneously. The number of port vehicles needed in peaks is also significantly lower. In scenario P1 with unbalanced demand, port vehicles were driving 3925 timesteps, while only transporting containers 2041 of those timesteps. The maximum amount of containers that were transported by port vehicles at any timestep was 74 in this scenario. In the corresponding scenario B1, 174 containers were transported at one timestep. This is likely to incur even higher costs for driving empty. The only scenario where the port vehicles are almost continuously in use is P3, where only a very limited number is available. The ability to take decision based on the actual number of trucks available thus have a large impact on the realistic viability of the operational decisions. The methods presented in this paper are intended for the operational level, but if they are used on the tactical level to dimension a truck fleet, the results show that it is important to consider container and truck routes simultaneously, since the peaks here give a realistic indication of the necessary fleet size.

The number of vehicles driving empty is as expected higher in the scenarios with unbalanced demand since the trucks have to be replaced to the port before they can transport new containers. The benchmark method always has trucks available, and do thus not need to wait. The fluctuation in the truck usage is therefore much higher for the benchmark method. The same trends can be seen for

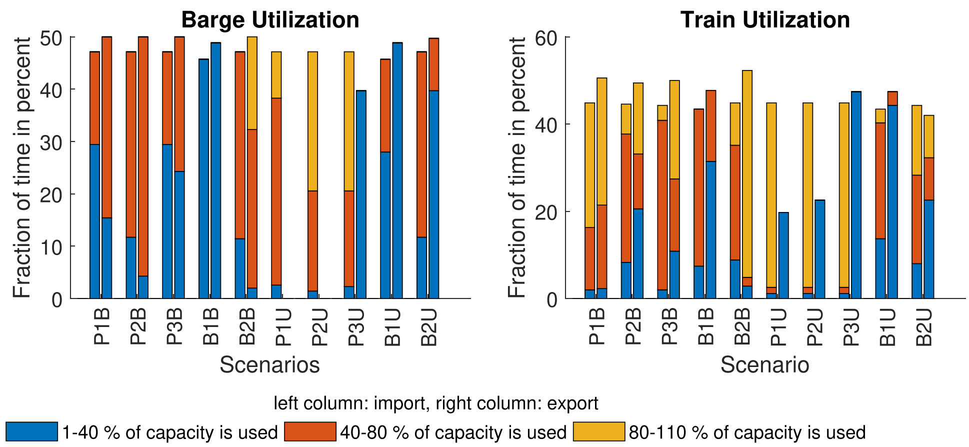

Fig. 6. The occupation rate at the barge and the train for each simulation. Only results for $k \in [250, 600]$ are used as the initial zero-states impacts the earlier results. Example of how to read figure: In scenario P3 with balanced demand profile (P3B) the import takes up (1–40]% of the barge capacity 29% of the time. In the same scenario the export takes up (40–80]% of the capacity 50-24=26% of the time.

the scenarios with balanced demand, however with smaller peaks. The number of vehicles driving empty with the proposed method is very low with empty to full ratios of less than 3% in all scenarios with balanced demand. This indicates that considering containers and trucks simultaneously provides benefits such that trucks may wait for a new transport demand before departing from a node.

The utilization of the barge and train is shown in Fig. 6. It shows the time the barge and train operate at different utilization rates. The results are shown in separate bars for import and export. Hence, a barge-import-utilization rate of (40–80%] is achieved when (40–80%] of the barge capacity is filled with import containers. Only results for timsteps $t \geq 250$ are used to generate the data, as the network starts empty and only around this timestep is fully developed according to Figs. 3 and 5. The results are furthermore shown in percentage of this time-interval rather than timesteps. Hence in scenario 1 with balanced demand and the controller considering simultaneous routing (P1B), import containers take up (1–40%] of the barge capacity 29% of the time, (40–80%] of the barge capacity 18% of the time and never takes up more than 80% of the barge capacity. In the formulation of the model, it is not specified when a container should be unloaded from a scheduled service at the time it is loaded to the service, it is thus possible for both import and export containers to stay on the barge or train while they travel in both directions. This do however only occur in the beginning of the simulations and never while $t > 250$.

The number of containers that are not delivered in time varies largely between the different scenarios, and from the nominal cases to the cases with uncertain travel time. Since the model considers commodity flows, the deadline of each container is not considered, instead the satisfaction of demand for a given commodity at a given destination is discussed. In Table 5 the unsatisfied demand at the three destinations are shown over all scenarios for $t > 200$. The numbers count each timestep a demand was not satisfied, so if a container of a given commodity was needed at time $t = 300$ but is only satisfied at $t = 305$, it adds 5 to the count. The demand at the ship (node 1) follows the capacity at which the ship can be (un)loaded. In scenario 1 all demand is thus to be satisfied at the first timestep possible, while the demand is spread over more timesteps in scenario 2 and 3.

When both container and truck routes are considered, more demand is left unsatisfied in most cases. This is especially pronounced in the case of unbalanced demand, where there is more import than export. In order to reduce the cost of empty trucks driving from the hinterland to the port, the proposed method

Table 5
Unsatisfied demand where each delayed container is counted each timestep. Grey highlighted rows correspond to the scenarios with travel time disturbances while the others are from nominal scenarios.

	Balanced demand					Unbalanced demand				
node	P1	P2	P3	B1	B2	P1	P2	P3	B1	B2
3	26	26	29	19	25	38	37	1435	17	24
	126	110	96	105	65	180	144	1924	176	107
5	25	28	28	19	26	39	37	1163	18	23
	97	80	65	120	50	165	120	1546	152	86
Ship	3	2	5	2	2	3	1	22	3	1
	8	15	44	6	5	6	8	33	4	2

prefers the use of scheduled services as seen in Fig. 6. This results in longer travel times and thus more unsatisfied demand. The scheduled services are priced per slot, so the MPC does not consider the cost of sailing/driving back empty for those services. All scenarios prioritize satisfying the more costly ship demand over the inland demand and therefore the unsatisfied demand is much lower at the ship destination.

When the travel times of trucks become uncertain, all scenarios with sufficient numbers of trucks available adjust the plans and obtain results with expected increases in the unsatisfied demand. Scenario P3 with unbalanced demand does not have enough vehicles available to transport the containers quick enough. This is also visible in Fig. 5 where all vehicles are driving either empty or full at all times. The results from this scenario does thus not represent an implementation we recommend, but serves to ensure that even in cases with under capacity the MPC does not violate the constraints or render infeasible. Eventually the MPC would accumulate very high costs which may give numerical issues, but in a real-world implementation high rates of unsatisfied demand will raise awareness and foster changes in either fleet size or accepted demand.

When the constraints are unrestricting more demand is left unsatisfied with the proposed method under deterministic travel times. However when the truck travel time becomes uncertain, the performance is very similar to the performance of the benchmark method. This is due to the benchmark method's constant availability of trucks in combination with the cheap stacking of containers in node 6. For the benchmark method in scenario 1 the cheapest option is often to truck the containers just in time to satisfy the demand and otherwise keep them at node 6. When trucks and containers are considered simultaneously, it becomes cheaper for the MPC to truck containers ahead of demand if there is otherwise an empty truck driving in that direction. The stacks at node 3 and 5 do thus contain containers with destination at their adjacent virtual destinations more often for the proposed method or when the capacity in node 6 is a limiting factor. More local storage increases the chance that a delay in the truck travel time will not cause a delay in demand satisfaction.

6. Conclusions and future research

The often used assumption that trucks are instantly available at any location in the synchromodal network significantly changes what the optimal actions are. A plan under this assumption is thus likely to perform worse in reality where only a finite number of trucks are available. The proposed method routes trucks and containers simultaneously and successfully smooths out peaks in the needed number of trucks, even when it has a large number of trucks available. This creates better plans for companies that hire third-party trucks for excess capacity, as the company's basic fleet will be utilized better between peaks and less third-party trucks will be needed to serve the peaks.

The proposed method can limit the total number of trucks available in the system. When the plan is optimized for the actually available number of trucks, infeasible plans are avoided and the utilization rates of the available trucks are improved, i.e. the number of empty trucks driving in the network is reduced to the benefit of the environment, society and economy. The proposed method furthermore gives information on how many trucks are to be relocated between specified locations. This information can be used in hindsight or already at the planning stage to indicate beneficial volume-changes to the transport company's sales department.

The proposed method is based on MPC, which previously has been used for container transport planning in some instances, but not for planning containers and trucks simultaneously. It is shown that the MPC can adjust the plan when the travel times are subject to disturbances. When containers and trucks are planned simultaneously, the MPC is encouraged to transport containers when an empty truck is available at the right location and not only based on deadlines, which makes the system less sensitive to travel time delays.

A core assumption in the presented method is that commodity flows accurately represent container transport. This may be the case for large quantity flows of non-perishable goods, but may not hold when the deadline of each container is important. An investigation of the impact of this assumption on performance and computation time will comprise an important line of our future research.

The presented method assumes a global controller with perfect predictions and unconditional authority to take and implement all decisions. This description only fits very large transport companies with well-integrated departments and enough orders to make qualified forecasts of demand. Therefore, it is in future research relevant to study the sensitivity of the proposed method to the quality of the demand and parameter predictions and how robust MPC techniques can help alleviate the decreased performance due to uncertainty. It is furthermore of both theoretical and practical relevance to study the problem assuming multiple decision makers. Questions arising, when the network is distributed, involve information sharing, profit distribution and exception handling. Future research on this topic should consider both networks of homogeneous and heterogeneous agents. An example of the latter is adding barge-agents that decide departure times instead of using scheduled departures as assumed in this paper. Another line of research that would prepare the proposed model to real world implementations is the consideration of operation hours of terminals.

Acknowledgment

This research is supported by the project “Complexity Methods for Predictive Synchromodality” (project 439.16.120) of the Netherlands Organisation for Scientific Research (NWO).

References

[1] A. Alessandri, C. Cervellera, M. Cuneo, M. Gaggero, G. Soncin, Management of logistics operations in intermodal terminals by using dynamic modelling and nonlinear programming, Marit. Econ. Logist. 11 (1) (2009) 58–76.
[2] B. Behdani, Y. Fan, B. Wiegmans, R. Zuidwijk, Multimodal schedule design for synchromodal freight transport systems, Eur. J. Transp. Infrastruct. Res. 16 (3) (2016) 424–444.
[3] A. Boccia, L. Grüne, K. Worthmann, Stability and feasibility of state constrained mpc without stabilizing terminal constraints, Syst. Control Lett. 72 (2014) 14–21.
[4] P. Buijs, J.A.L. Alvarez, M. Veenstra, K.J. Roodbergen, Improved collaborative transport planning at dutch logistics service provider fritom, Interfaces 46 (2) (2016) 119–132.
[5] L. Chisci, J.A. Rossiter, G. Zappa, Systems with persistent disturbances: predictive control with restricted constraints, Automatica 37 (7) (2001) 1019–1028.
[6] M.S. Darup, M. Cannon, Some observations on the activity of terminal constraints in linear mpc, in: Proceedings of the European Control Conference, 2016, pp. 770–775.
[7] A. Di Febbraro, N. Sacco, M. Saeednia, An agent-based framework for cooperative planning of intermodal freight transport chains, Transp. Res. C: Emerg. Technol. 64 (2016) 72–85.
[8] M. Drexl, Applications of the vehicle routing problem with trailers and transshipments, Eur. J. Oper. Res. 227 (2) (2013) 275–283.
[9] W. Guo, W.B. van Blokland, G. Lodewijks, Survey on characteristics and challenges of synchromodal transportation in global cold chains, in: Proceedings of the International Conference on Computational Logistics, 2017, pp. 420–434.
[10] T. Hipólito, J. Nabais, R. Carmona Benítez, M. Ayala Botto, R.R. Negenborn, Mpc approach for synchronized supply chains of perishable goods, in: Proceedings of International Conference on Industrial Engineering and Systems Management, 2017.
[11] A. Hosseini, T. Sahlin, An optimization model for management of empty containers in distribution network of a logistics company under uncertainty, J. Ind. Eng. Int. (2018).
[12] R.B. Larsen, B. Atasoy, R.R. Negenborn, Simultaneous planning of container and vehicle-routes using model predictive control, in: Proceeding of the European Control Conference, 2019, pp. 2177–2182.
[13] L. Li, R.R. Negenborn, B. De Schutter, Distributed model predictive control for cooperative synchromodal freight transport, Transp. Res. E: Logist. Transp. Rev. 105 (2017) 240–260.
[14] L. Li, R.R. Negenborn, B.D. Schutter, Intermodal freight transport planning a receding horizon control approach, Transp. Res. C: Emerg. Technol. 60 (2015) 77–95.
[15] X. Lin, R.R. Negenborn, M.B. Duinkerken, G. Lodewijks, Quality-aware modeling and optimal scheduling for perishable good distribution networks: the case of banana logistics, in: T. Bektaş, S. Coniglio, A. Martinez-Sykora, S. Voß (Eds.), Computational Logistics, Springer International Publishing, 2017, pp. 483–497.
[16] J. Löfberg, YALMIP: a toolbox for modeling and optimization in Matlab, in: Proceedings of International Symposium on Computer Aided Control System Design, 2004.
[17] J.M. Maestre, D.M. de la Pena, E.F. Camacho, Distributed mpc: a supply chain case study, in: Proceedings of the Conference on Decision and Control (CDC) held jointly with Chinese Control Conference, IEEE, 2009, pp. 7099–7104.
[18] D. Mayne, An apologia for stabilising terminal conditions in model predictive control, Int. J. Control 86 (11) (2013) 2090–2095.
[19] M.R. Mes, M.E. Iacob, Synchromodal transport planning at a logistics service provider, in: Logistics and Supply Chain Innovation, Springer, 2015, pp. 23–36.
[20] J.L. Nabais, R.R. Negenborn, M.A. Botto, Model predictive control for a sustainable transport modal split at intermodal container hubs, in: Proceedings of the International Conference on Networking, Sensing and Control, 2013, pp. 591–596.
[21] J.L. Nabais, R.R. Negenborn, R. Carmona-Benítez, M.A. Botto, Cooperative relations among intermodal hubs and transport providers at freight networks using an mpc approach, in: Proceedings of International Conference on Computational Logistics, Springer, 2015, pp. 478–494.
[22] E. Perea-Lopez, B.E. Ydstie, I.E. Grossmann, A model predictive control strategy for supply chain optimization, Comput. Chem. Eng. 27 (8–9) (2003) 1201–1218.
[23] H.N. Psaraftis, M. Wen, C.A. Kontovas, Dynamic vehicle routing problems: Three decades and counting, Networks 67 (1) (2016) 3–31.
[24] W. Qu, J. Rezaei, Y. Maknoon, L. Tavasszy, Hinterland freight transportation replanning model under the framework of synchromodality, Transp. Res. E: Logist. Transp. Rev. 131 (2019) 308–328.
[25] V. Reis, Should we keep on renaming a +35-year-old baby? J. Transp. Geogr. 46 (2015) 173–179.

[26] H.G. Resat, M. Turkay, A discrete-continuous optimization approach for the design and operation of synchromodal transportation networks, Compute. Ind. Eng. 130 (2019) 512–525.
[27] A.E.P. Rivera, M.R. Mes, Anticipatory freight selection in intermodal long-haul round-trips, Transp. Res. E: Logist. Transp. Rev. 105 (2017) 176–194.
[28] A.P. Rivera, M. Mes, Service and transfer selection for freights in a synchromodal network, in: Proceedings of the International Conference on Computational Logistics, Springer, 2016, pp. 227–242.
[29] F. Schulte, R.G. González, S. Voß, Reducing port-related truck emissions: coordinated truck appointments to reduce empty truck trips, in: F. Corman, S. Voß, R.R. Negenborn (Eds.), Computational Logistics, Springer International Publishing, 2015, pp. 495–509.
[30] P. Singh, M. van Sinderen, R. Wieringa, Synchromodal transport: pre-requisites, activities and effects, in: Proceedings of the International Conference on Information Systems Logistics and Supply Chain, 2016, pp. 1–4.
[31] M. SteadieSeifi, N.P. Dellaert, W. Nuijten, T. Van Woensel, R. Raoufi, Multimodal freight transportation planning: a literature review, Eur. J. Oper. Res. 233 (1) (2014) 1–15 .
[32] B. Van Riessen, R.R. Negenborn, R. Dekker, Synchromodal container transportation: an overview of current topics and research opportunities, in: Proceedings of International Conference on Computational Logistics, Delft, The Netherlands, 2015a, pp. 386–397.
[33] B. Van Riessen, R.R. Negenborn, R. Dekker, G. Lodewijks, Service network design for an intermodal container network with flexible transit times and the possibility of using subcontracted transport, Int. J. Shipp. Transp. Logist. 7 (4) (2015b) 457–478.
[34] W. Wang, D.E. Rivera, Model predictive control for tactical decision-making in semiconductor manufacturing supply chain management, IEEE Trans. Control Syst. Technol. 16 (5) (2008) 841–855.
[35] Y. Xu, C. Cao, B. Jia, G. Zang, Model and algorithm for container allocation problem with random freight demands in synchromodal transportation, Math. Probl. Eng. 2015 (2015).