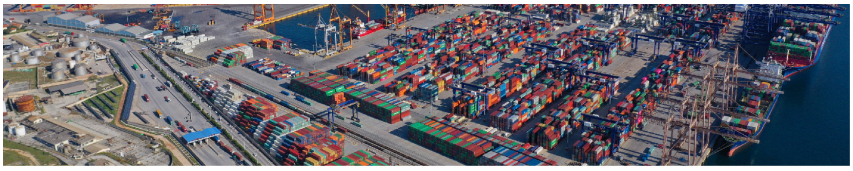


IPIC 2021 | 8th International Physical Internet Conference | Athens

PhysICAL

On Modelling and Solving Green Collaborative Tactical Transportation Planning

Lukas Gosch[1,2,*], Matthias Prandtstetter[1], Karl F. Doerner[2]
[1]AIT Austrian Institute of Technology, Center for Energy, Vienna, Austria
[2]University of Vienna, Austria
[*]Corresponding author: gosch.lukas@gmail.com

Abstract: *This paper presents a new mathematical model for tactical transportation planning in a horizontal collaboration defined by warehouse sharing and the joint organization of transport. The model features intermodal transport, handling and storage capacities, diverse products and realistic tariff structures with volume discounts. Furthermore, it allows for sustainable planning by associating estimated CO_2 equivalent (CO_2e) emissions to each logistic operation and then either optimize for transportation costs, emissions or both objectives. Subsequently, we derive a mixed-integer formulation for exact solution approaches and a hybrid heuristic to solve large-scale instances. The hybrid heuristic is composed of a slope scaling matheuristic we generalized to non-negative integer variables and a second local search based refinement step, which reroutes flow of multiple products at once along lowest-cost paths in the network. Results obtained from simulating collaboration in the Danube Region using the regional available transportation infrastructure including railway and shipping networks reveal significant saving potentials in both costs and CO_2e emissions. Cost minimizing solutions always lead to reductions of the carbon footprint. However, minimizing for emissions can significantly further this reduction but requires a minimum size of the collaboration to operate cost-efficiently.*

Conference Topic: *Interconnected freight transport, logistics and supply chain*

Keywords: *Freight Transportation Planning; Horizontal Collaboration; Intermodal; Emissions Reduction; Sustainability; Capacitated Network Design; Space Time Network; Hybrid Heuristic; Mixed-Integer Linear Programming; Local Search*

1 Introduction

The transport sector is responsible for over 14% of the total anthropogenic greenhouse gas emissions. Approximately 43% of these emissions can be attributed to freight transportation (Sims et al., 2014). Furthermore, the OECD predicts a tripling of the global goods traffic until 2050 (OECD, 2003; International Transport Forum, 2019) and when continuing with the status-quo, transport emissions are predicted to increase at the fastest rate compared to any other energy-related end-use sector (Sims et al., 2014). As a result, urgent measures must be taken to reduce the rising emissions in freight logistics. Central solution proposals for a more sustainable transport of goods are: modal shifts, increasing freight load factors for example through consolidation, and better optimized and integrated transport networks (Sims et al., 2014). A recent paradigm in logistics addressing all these points at once is collaboration between competitors taking the form of sharing warehouses and jointly organizing freight transportation. This so called horizontal collaboration (European Union, 2001) is a major stepping stone of the EU's strategy to achieve climate-neutrality by 2050 (ALICE, 2014, 2019) and an integral part of the Physical Internet vision (Montreuil, 2011).

Therefore, this paper introduces and solves a new mathematical model to optimally and sustainably plan and use logistic resources in a horizontal collaboration. Planning and decision making in logistics and supply chain management is structured into three different levels, strategic (long-term), tactical (medium-term) and operational (short-term) (Crainic, 2000). The model introduced here assumes that the relevant transportation infrastructure is already in place and is concerned with tactical planning which sets the general conditions for operational decisions.

1.1 Problem Description

Competing enterprises with diverse sets of products want to collaborate to reduce logistic costs and emissions. This collaboration is characterized by sharing transportation and opening up warehouses for mutual usage, both enabled by packing goods in modular and standardized load units. The companies together have a network of warehouses spread over different geographical regions with differing product demands, but not every enterprise on its own has a warehouse in every demand region. Furthermore, different companies can have product demands in similar geographical regions. Consequently, they could share certain transport routes. Now, to deliver goods into a demand region on time, the companies want to devise a shared transportation strategy which makes effective use of new consolidation and storage potentials and existing intermodal infrastructure whose usage is unlocked through higher product volumes. As these planning issues require some lead time, they are interested in developing a tactical freight plan. Hence, the model's main focus is in finding optimal paths freight should take through the existing network. This includes tariff, transport mode and storage choices making effective use of spatial and temporal consolidation potentials and of opportunities for economics of scale. The model is not concerned with concrete vehicle routing, packing problems or similar as these decisions are part of operational planning.

Key Model Aspects

Sustainable planning is achieved by associating estimated CO_2 equivalent (CO_2e) emissions to each logistic operation and optionally pricing them allowing for *internalization*. Internalization refers to estimating costs of wider effects of business activities on the community and ecosystem and integrating them into the companies budgets (McKinnon et al., 2015). As a result, the developed model allows to optimize for transportation costs, CO_2e emissions or both. The ***diversity of products*** is met with a holistic commodity-modeling approach. This includes if necessary, special transportation conditions such as cooling. Then, through adding a time dimension, the model allows for expiry-aware shared routing of perishable and non-perishable goods.

The model considers a network of facilities which can be warehouses, factories or transshipment points, connected by (capacitated) transportation relations of arbitrary types in space and time. As a result, ***intermodal transportation*** possibilities are integrated capturing their full implications on emission, cost and delivery time. Additionally, each node in the network can be endowed with ***handling capacity limits***. ***Storage*** possibilities are represented by transport relations in time between the same facility. Finally, the model incorporates ***realistic tariff structures*** with all-unit volume discounts often found in practice (Munson and Jackson, 2015). We design graph-structures to linearly model these tariffs. These allow us to formulate the problem as a ***capacitated network design problem*** (Magnanti and Wong, 1984).

1.2 Overview of the Paper

In Section 2 our model is formally introduced. Due to the linear mixed-integer programming formulation being NP-hard, we develop and combine two different types of heuristics to successfully

solve large instances. In Section 3.1 we generalize the slope scaling heuristic (Kim and Pardalos, 1999; Crainic et al., 2004) to non-negative integer variables making it applicable to our model. Slope scaling is a matheuristic originally developed for network design problems with binary variables and constitutes an integral part of state-of-the art hybrid heuristics for this problem type (Gendron et al., 2018; Akhavan Kazemzadeh et al., 2021). In Section 3.2 a local-search based approach - using the slope scaling heuristic as a fast and effective construction heuristic - is developed. Based on ideas from Harks et al. (2016), it jointly reroutes flow of multiple commodities using lowest-cost paths in the network. The aforementioned heuristics as well as a commercial MIP-solver are applied to generated problem instances based on the real intermodal transportation infrastructure in the Danube Region. The results of which can be found in Section 4. They show significant savings in costs and emissions. A conclusion can be found in Section 5.

2 Mathematical Model

This section presents the formal model developed and used in this work. Fundamentally, two stages of the underlying network are differentiated. First, a time-expanded network is given (see Section 2.1). This forms the basis on which the heuristics (see Section 3) operate. However, non-linear transportation tariffs between nodes result in a non-linear objective. Therefore, in a second stage simple arcs representing these transport relations can be replaced with more complex graph-gadgets to obtain a linear mixed-integer programming (MIP) formulation called the tariff-expanded network (see Appendix B). This allows to apply exact MIP-solvers to the problem[1]. However, only the heuristic solution approaches manage to effectively solve large instances. We call the developed network design problem the generalized tactical transportation problem (GTTP), as it can be shown to include the tactical transportation problem introduced by Harks et al. (2016) as a special case.

2.1 Time-Expanded Network

The time-expanded network $G_{\mathcal{T}}$ is constructed for a certain time horizon T (e.g. 7, 14 or 30 days) with the individual time periods summarized in the set $\mathcal{T} = \{1, ..., T\}$. $G_{\mathcal{T}}$ is constructed from a set of base nodes copied T-times resulting into a time-expanded node set $\mathcal{V}_{\mathcal{T}}$. Base nodes are either: physical facilities (e.g. warehouses, factories or transshipment points) with each facility being part of a demand region, demand nodes used to represent demand regions, or bin nodes to remove unused or expired goods form the system. There is only one bin node in the set of base nodes. Then, the arc set $\mathcal{A}_{\mathcal{T}}$ consists of transport relations between these nodes in different time periods. Transport relations can be of arbitrary types and we distinguish transportation modes $\mathcal{M} = \{L, R, S\}$ consisting of lorry (L), rail (R) and ship (S), storage arcs connecting the same facility at two consecutive time periods denoted by type C, and artificial arcs connecting each facility to its regional demand node as well as to the bin node. These use the type-symbol Ω.

Network Structure

The facilities are connected among themselves based on physical realities. If for example two facilities $i, j \in \mathcal{V}_{\mathcal{T}}$ are connected by a road network, then there exists an arc $a = (i, j, L) \in \mathcal{A}_{\mathcal{T}}$ with corresponding non-zero distance d_a and non-zero travel time τ_a. Storage arcs have a distance of zero but a travel-time of one. For each time-period, there is one bin node and each facility is connected to it with distance and travel time of zero. Facilities in a demand region are connected to the associated demand node with distance and travel times of zero using mode Ω. Hence, demand

[1] If one only optimizes for emissions, the time and tariff expanded networks are equivalent. This is due to the (very unfortunate) physical impossibility of volume discounts on emissions.

in the region the warehouse is located in can be satisfied directly out of stock. This is due to the fact that the model is only concerned with successful transportation into a region to fulfill its demand and the inner-regional distribution to customers should be addressed by an operational model. Furthermore, each facilities is connected to every other demand region except its own using mode L with non-zero distances and travel times. These connections represent direct deliveries by lorries into a demand region without first delivering into a regional facility or using any form of intermodal transport.

2.2 Commodities

All facilities are supplied with a certain stock of products, also called commodities, each time period acting as *source* nodes. Demand nodes have associated commodity demands for each time period acting as *sinks*. The set of commodities is denoted by $\mathcal{K}$. In general, products can be highly diverse in their volume and weight henceforth called the products *properties* (see also Harks et al. (2016)). As an example, it does make a non-negligible impact on the transportation costs and available resources if one transports a certain number of styrofoam sheets compared to the same number of steel beams. Therefore, each commodity $k \in \mathcal{K}$ has an associated extent p_{kl} for each property $l \in \mathcal{P}$. Additionally, products have different transportation and storage characteristics. As an example, some goods may require cooling to specific temperatures or can be perishable. To model different transportation requirements, we introduce a *product type* concept and associate to each commodity a specific type from a set of types Σ. Only goods of the same type can be transported together. In our instances, we separate two types of goods, those whose storage-containers require electricity and those whose storage-containers don't. Additionally, to address perishability commodities $k \in \mathcal{K}$ can have varying *lifetimes* Δt_k. As a result, a commodity is defined as perishable if its lifetime is smaller than the number of time-periods in the network. Consequently, commodities can be grouped into a set of perishable ones $\mathcal{K}_\Delta$ and a set of non-perishable ones $\mathcal{K}_\mathcal{T}$. In order to correctly handle perishability in the model, each perishable commodity has an associated *production timestamp* $t \in \mathcal{T}$ and is counted as expired in time-periods greater than $t + \Delta t_k$. With $x_{k_t} \in \mathbb{R}_0^+$ we denote the *flow* of perishable commodity k produced at time t and with $x_k \in \mathbb{R}_0^+$ the flow of non-perishable commodity k. These can be pooled together into a *flow vector* $x \in \mathbb{R}_0^{|\mathcal{K}_\mathcal{T}|+|\mathcal{K}_\Delta||\mathcal{T}|}$. The summed up extent of property $l \in \mathcal{P}$ over all commodities of type $\sigma \in \Sigma$ in x is given by the flow-sum function $P_l^\sigma(x) \coloneqq \sum_{k \in \mathcal{K}_\mathcal{T}^\sigma} p_{kl} x_k + \sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{K}_\Delta^\sigma} p_{kl} x_{k_t}$. The sets $\mathcal{K}_\mathcal{T}^\sigma$ and $\mathcal{K}_\Delta^\sigma$ comprise all non-perishable or perishable goods of type σ.

2.3 Tariffs

Each transport relation has one[2] associated tariff for each product type. Harks et al. (2016) give a good overview of possible tariffs and their resulting cost functions. In this work, we are especially concerned with *all-unit quantity discounts* as they are the most prevalent (Munson and Jackson, 2015). Corresponding tariffs usually have a set of cost levels $\mathcal{N} = \{1, ..., N\}$ defining different cost rates depending on the shipped volume or weight. Therefore, Harks et al. (2016) develop a cost function with different linear cost rates depending on the sent flow extent in some property $l \in \mathcal{P}$ as follows. Each cost level $n \in \mathcal{N}$ is applied starting from transport volume β_n and has the linear cost rate c_n. As a result, the cost function by Harks et al. (2016) for flow vector x and a specific

[2] Multiple available tariffs can be modeled by parallel arcs.

product type[3] σ is given as

$$C(x) = \min_{n \in \mathcal{N}} \{c_n \cdot \max\{P_l^\sigma(x), \beta_n\}\} \tag{1}$$

However, only linearly pricing transport volumes does not adequately capture the diversity of all-unit volume discount tariffs employed by freight transportation companies. Especially, it is not applicable to constant cost levels for different transport weight or quantity intervals or a combination of a fixed-cost rate with volume-dependent linear costs resulting in piecewise-linear cost levels (see Figure 1 in Appendix B). An overview of such tariffs with more complex cost structures is given in Kempkes and Koberstein (2010). Below, we formalize these tariffs used in our work.

2.3.1 Constant Cost Levels with All-Unit Discounts

Constant cost levels for different weight or quantity intervals are, among others, often encountered in rail cargo shipments. In them for each cost level $n \in \mathcal{N}$ a different fixed-cost rate f_n is assumed. In theory, these cost levels can have arbitrary interval lengths. We assume a fixed property- and type-dependent interval length of B_l^σ and interpret it as the capacity in direction $l \in \mathcal{P}$ of one installed *transport unit (TU)* on this transport relation. Exemplary, for rail or ship cargo shipments one transport unit could be one ISO container (see Figure 1 in Appendix B). Then, the number of installed TUs for product type σ is given by $\tilde{y}^\sigma = \max_{l \in \mathcal{P}} \{\lceil \frac{P_l^\sigma(x)}{B_l^\sigma} \rceil\}$. In the mixed-integer problem installed TUs are additional integer decision variables (see Appendix B.2). Now, β_n denotes the minimum number of TUs at which cost level n starts to be applicable. The starting cost of the n-th level is b_n and represents the base cost of shipping β_n containers. This yields the cost function (2).

$$C(x) = \min_{n \in \mathcal{N}} \{f_n \cdot \max\{\tilde{y}^\sigma - \beta_n, 0\} + b_n\} \tag{2}$$

Exemplary, b_n can represent the price of β_n TUs with the new cost rate f_n, i.e. $b_n = f_n \beta_n$. If $f_n \leq f_{n-1}$, this represents an all-unit discount structure. However, if b_n represents the costs of $\beta_n - 1$ TUs to the previous price rate f_{n-1} plus the costs of one TU to the current rate f_n, it would represent incremental discounts. Note that to formulate the cost function not in terms of installed TUs, but in term of single-property levels (as for example weight levels), replace $\tilde{y}^\sigma$ with $\tilde{y}_l^\sigma = \lceil \frac{P_l^\sigma(x)}{B_l^\sigma} \rceil$.

2.3.2 Piecewise-Linear Cost Levels with All-Unit Discounts

Tariffs with different piecewise linear cost levels are very similar to the constant cost levels case except that for each discount level n, we additionally assume variable costs c_n for the actual volume or weight of flow. The resulting cost function writes

$$C(x) = \min_{n \in \mathcal{N}} \{f_n \cdot \max\{\tilde{y}^\sigma - \beta_n, 0\} + c_n \cdot \max\{P_l^\sigma(x) - B_l^\sigma(\beta_n - 1), 0\} + b_n\} \tag{3}$$

Again, setting $b_n = f_n \beta_n + c_n B_l^\sigma(\beta_n - 1)$ with $f_n \leq f_{n-1}$ and $c_n \leq c_{n-1}$ results in all-unit discounts. Note that the linear costs only depend on one property. This cost function exemplary arises when modelling transportation costs with lorries. A fixed costs part arises for each commissioned truck due to various singular factors such as driver wages. Linear costs arise due to increasing fuel consumption based on its actual weight.

[3] Harks et al. (2016) do not have a product type concept and aggregate the property-extent over all products.

3 Heuristics

In Section 3.1 we generalize the slope scaling heuristic (SSC) to non-negative integer variables and develop an approximate linear program for our model. SSC is used as a fast construction heuristic for the local search presented in Section 3.2. We also experimented with iterative linear programming (Gendron et al., 2018) as theoretical properties such as finite convergence can be generalized to non-negative integer variables. However, we found it does not scale to larger instances.

3.1 Slope Scaling

Slope Scaling is an iterative solution process based on solving an approximate linear program $\mathcal{LP}$ exactly. The resulting solution $\tilde{x}$ is used to construction a feasible solution $(\tilde{x},\tilde{y})$ to the original problem $\mathcal{MIP}$. Then, $(\tilde{x},\tilde{y})$ is used to update the coefficients in the objective of $\mathcal{LP}$ yielding $\mathcal{LP}'$ in such a way that the objective function value $v(\mathcal{MIP}(\tilde{x},\tilde{y}))$ equals $v(\mathcal{LP}'(\tilde{x}))$. We denote the set of linearization coefficients in the linear program at iteration t as $\rho(t)$ and the corresponding linear program as $\mathcal{LP}(\rho(t))$. $\mathcal{LP}(\rho(t))$ takes the form of a network flow problem and the formal definition is presented in Appendix C. After efficiently solving $\mathcal{LP}(\rho(t))$, a solution $(\tilde{x},\tilde{y})$ to $\mathcal{MIP}$ is obtained by setting $\tilde{y}^{\sigma}_a = \max_{l\in\mathcal{P}}\{\lceil \frac{P^{\sigma}_l(\tilde{x}_a)}{B^{\sigma}_l}\rceil\}, \forall a \in \mathcal{A}_{\mathcal{T}}$. Denote with $\lambda_1 f^{\sigma}_a$ the weighted costs and with $\lambda_2\Delta^{\sigma}_a$ the weighted emissions of one TU (see $\mathcal{MIP}$-objective in Appendix B.3) and with $l(a)$ the property used for linear pricing on arc a. Now, using $(\tilde{x},\tilde{y})$, $\rho(t)$ can be updated as follows

$$\rho^{\sigma}_a(t+1) = \begin{cases} (\lambda_1 f^{\sigma}_a + \lambda_2\Delta^{\sigma}_a)\tilde{y}^{\sigma}_a \,/\, P^{\sigma}_{l(a)}(x_a) & \text{if } P^{\sigma}_{l(a)}(x_a) > 0 \\ \rho^{\sigma}_a(t) & \text{otherwise} \end{cases} \tag{4}$$

resulting in $v(\mathcal{MIP}(\tilde{x},\tilde{y})) = v(\mathcal{LP}(\rho(t+1)))$. Then, $\mathcal{LP}(\rho(t+1)))$ is solved and the above solution construction and update scheme repeated for $t+1$. Initial cost estimates are set to the linearized costs of one full TU on the respective connection, i.e. $\rho^{\sigma}_a(t=0) = (\lambda_1 f^{\sigma}_a + \lambda_2\Delta^{\sigma}_a)/B^{\sigma}_{al(a)}$. The slope scaling heuristic starts at $t=0$ and runs for as many iterations until an already seen solution is produced (resulting in a loop) or a specific time limit is reached. Additionally, it is possible to omit $\tilde{y}^{\sigma}_a$ from the update equation (4). Then, $v(\mathcal{MIP}(\tilde{x},\tilde{y})) \neq v(\mathcal{LP}(\rho(t+1)))$ but for $t \geq 1$ the update scheme becomes *monotonic* on each tariff level with respect to increasing flow extent. This can be interpreted as favoring higher transport volumes on an arc no matter the filling rate of the last TU, playing into the economics of scale. The results in Section 4 indicate that the monotonic update scheme outperforms the update scheme (4) matching objective-costs. This could indicate that the prime paradigm behind designing a slope scaling mechanism - matching objective-costs (Kim and Pardalos, 1999) - is not necessarily key to its success. Other properties such as monotonicity can be equally important. This insight can be used to devise novel problem specific update rules diverting from pure objective-costs matching.

3.2 Local Search

The presented local search is based upon the *flow decomposition theorem* (Ahuja et al., 1993) which states that every non-negative arc flow can be represented as a path and cycle flow. Thus, the basic idea of the local search is to construct a path-decomposition of flow on the acyclic time-expanded graph and then randomly dissolve paths and reroute the flow. The employed moves and path decomposition calculations are inspired by Harks et al. (2016). A path-decomposition $\mathcal{P}$ consists of a set of tuples (P, f_P). Each tuple consists of a path P in $\mathcal{G}_{\mathcal{T}}$ and the transported commodities $f_P \in \mathbb{R}_0^{|\mathcal{K}_{\mathcal{T}}|+|\mathcal{K}_{\Delta}||\mathcal{T}|}$. Paths carry only commodities of a specific type and due to the

topology of the network, either end in a demand node or a bin node. Therefore, we call a path either a demand-path or a bin-path. Note that for a given flow $\mathcal{P}$ is not necessarily unique.

Now, we distinguish two types of neighbourhoods. The first consists of solutions constructed by removing one demand-path and all bin-paths and then, repairing the solution by rerouting the flow. The second neighbourhood differs by instead removing a group of demand-paths sharing the same transport relation. Therefore, the neighbourhoods are characterized by a given path-decomposition and the employed rerouting scheme R and are denoted $\mathcal{N}_1(\mathcal{P},R)$ and $\mathcal{N}_2(\mathcal{P},R)$, respectively. A move is defined by randomly choosing a solution from $\mathcal{N}_1(\mathcal{P},R)$ or $\mathcal{N}_2(\mathcal{P},R)$ and accepting it, if it improves upon the current solution. Furthermore, we distinguish two rerouting schemes: a *heaviest first* rerouting and a *cheapest relative cost* rerouting. The heaviest first rerouting searches a (source s, sink t)-pair allowing for the highest delivery weight and calculates a cheapest path from s to t and adds the found path to the path decomposition. It iteratively repeats these two steps until all demand has been satisfied or infeasibility is detected. The cheapest relative cost rerouting differs by choosing the (s,t)-pair which results in the cheapest path relative to the transported weight. The rerouting schemes are handling-capacity aware and if necessary, the flow of goods f_P is adapted to exactly fulfill handling-capacity limits. Arcs not allowing for additional flow are not anymore considered in finding cheapest paths. Lastly, two move-dependent ways to calculate the path decomposition are employed. For moves in $\mathcal{N}_1$ the path decomposition is calculated in a depth-first search manner. Starting at a source, for each incident arc, the maximum-weight flow vector that could be assigned to a path traversing this arc is calculated. The arc with the maximal maximum-weight-flow vector is chosen. For moves in $\mathcal{N}_2$ a bidirectional depth-first search is employed. Starting from the most heavily used arc with respect to unassigned flow, it chooses arcs in both directions that maximize the savings if one would reduce their flow by the assigned one.

To summarize, the local search works by in the beginning constructing a solution using slope scaling, followed by choosing an initial neighbourhood $\mathcal{N}_i(\mathcal{P},R)$ and then, repeating the following steps until convergence or a time limit:

1. Construct a path-decomposition for $\mathcal{N}_i(\mathcal{P},R)$.
2. Follow randomly chosen improving neighbouring solutions until the improvement over a certain number of iterations falls below a threshold.
3. Change R and go to step 1. If R has been changed in the last iteration, change i instead.

For the results in this paper, we chose $\mathcal{N}_1(\mathcal{P},\textit{cheapest relative cost})$ as an initial neighbourhood.

4 Results

In this section, we present the results of applying a state-of-the-art MIP solver and our heuristics to instances modelling horizontal cooperation in the Danube region.

4.1 Data

The instances are based on a dataset introduced by Wolfinger et al. (2019) including locations of major cities, train stations and Danube ports in Austria, Slovakia, Hungary, Romania, Serbia, and Bulgaria. In each city, we have added additional random locations and recalculated all road-distances using Ariadne (Prandtstetter et al., 2013). Distance and time-matrices of other transport modes were taken from Wolfinger et al. (2019). Additionally, each city is interpreted as a demand region and we randomly distribute commodity demands and stocks. Generated instances range in size from two collaborating warehouses in two different regions up to the exhaustion of our memory

limits during the solution process. More concretely, input instances can be grouped into having two ($R2$), fife ($R5$) or ten regions ($R10$). $R2$-instances are generated with either one or four warehouses per region. $R5$-instances are generated with two or four warehouses per region. For $R10$-instances the five more populated regions have four warehouses and five lower populated ones have two. Warehouses are again grouped representing association to different companies. Some regions are equipped with a cross-dock, train station or port. All instances have a time period of one day and are generated in three different time horizon (T) versions of 7, 14 or 30 days. Lastly, each instance has a version with active handling constraints (tight t) and without (loose l). Therefore, instance groups are identified using $R<\#>\_T<\#>\_t/l$ and using a $*$ instead of $<\#>$ to indicate averages over all respective instance groups. Tariffs with three levels of volume discounts are employed. Piecewise-linear cost levels are used for lorry connections and constant cost levels for rail and ship connections. On these connections, TUs always correspond to 40-feet ISO containers. Handling costs are integrated into the arc costs. Prices are set based on information from an industry partner.

In total 60 instances have been generated leading to 180 optimization problems considering optimizing for costs, emissions or both. The input instances and the code to generate them together with an in-depth description of the data (including price and emission tables) will be made available on https://github.com/saper0/gttp-data.

4.2 Experiments

Solutions calculated on our test instances are evaluated with respect to a *direct delivery solution*. In a direct delivery solution demand for a product is satisfied by a lorry shipment without the possibility of cooperation. This means consolidation is only possible across warehouses associated to the same company. Direct delivery instances are generated by removing transshipment nodes and arcs connecting warehouses from different companies from the original input instances. Next, a MIP-solver is applied on the adapted instances for one hour. The thereby calculated direct delivery solutions are either optimal or proven to be within a 1% optimality gap. All experiments were written in C++ using CPLEX 12.10, had a time limit of one hour and were performed on an Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60GHz CPU with 32 GB RAM.

4.2.1 Algorithmic Insights

Table 1 shows that the local search based approach, using one iteration[4] of slope scaling (SSC) to generate a starting solution, outcompetes all other solution approaches by large margins. Additionally, SSC using the more monotonic linearization procedure slightly outperforms the objective cost-matching one. This highlights the sensibleness of problem specific slope scaling mechanisms parting from the paradigm of matching objective costs between the approximate and original problem. Lastly, the MIP solver - initialized with the direct delivers solution to aid its solution process - struggles with emission optimization and completely fails if transportation costs should be optimized. This is due to the fact that when transportation costs are included, the MIP operates on the tariff expanded network leading to a model blow up quickly exceeding the memory limits.

4.2.2 Managerial Insights

Table 1 shows the average percentual improvements of the objective functions in different instance groups obtained by applying the aforementioned solution approaches. The following discussion concerns the best obtained result on each instance (consistently by local search, see Section 4.2.1). Optimizing for emissions universally leads to significant CO_2e reduction possibilities averaging

[4] Applying SSC until convergence and then applying local search leads to slightly inferior solutions.

Table 1: Average percentual improvements (reductions) of the objective functions in different instance groups compared to the direct delivery solution. Hence, for emissions the improvement on CO_2e, for costs the improvement on the transportation costs and in both the improvement on the sum of emissions and costs by pricing one tonne of CO_2e with 100€ (Delft, 2019) (using the avoidance cost approach (McKinnon et al., 2015)) is shown. Due to its random components, the local search has been run with ten different seeds. Its reported results are averages over these runs including their sample standard deviation. For the MIP approach, the number of solvable instances versus all instances in the instance group are shown. The other solution approaches can solve all instances except two in $r10\_T*\_*$ due to exceeding memory limits.

Optimize for		Local Search [%]	SSC (monotonic) [%]	SSC [%]	MIP [%]
Emissions	$r2\_T*\_*$	30.7 ± 0.1	29.9	29.7	27.3 (24/24)
	$r5\_T*\_*$	31.9± 0.2	29.5	29.2	15.3 (22/24)
	$r10\_T*\_*$	30.9 ± 0.4	28.9	28.5	9.2 (10/12)
	$r*\_T*\_*$	31.2 ± 0.2	29.6	29.3	19.4 (56/60)
Costs	$r2\_T*\_*$	6.0 ± 0.2	0.9	-1.6	2.0 (22/24)
	$r5\_T*\_*$	20.1 ± 0.4	12.5	10.3	0.0 (7/24)
	$r10\_T*\_*$	23.6 ± 0.6	16.9	14.5	0.0 (1/12)
	$r*\_T*\_*$	14.9 ± 0.4	8.5	6.8	1.5 (30/60)
Both	$r2\_T*\_*$	8.0 ± 0.3	2.2	1.5	2.1 (22/24)
	$r5\_T*\_*$	21.5 ± 0.4	14.9	12.8	0.0 (9/24)
	$r10\_T*\_*$	24.3 ± 0.5	18.5	16.3	0.0 (1/12)
	$r*\_T*\_*$	16.4 ± 0.3	10.2	8.7	1.4 (32/60)

31.2%±10.2%[5] (13.8% - 53.0%) *independent* of the size of the collaboration. However, we find emission reduction potentials are strongly correlated with the planning horizon. A planning horizon of 7 days lead to on average 20.5%±3.7% CO_2e reduction, 14 days to 33.0%±5.2% and 30 days to 41.1%±7.6%. This is due to the fact that intermodal transport has longer lead times. Therefore, some changeover time from lorry transportation is needed to make effective use of its possibilities. Concerning transportation cost optimization, contrary to emission minimization we find a clear correlation with the size of the collaboration increasingly enabling cost-efficient consolidation (see Appendix A for more results on consolidation). Significant cost reductions are possible on all looked upon instances averaging 14.9%±9.2% (0.4% - 30.5%). Similar to optimizing for emissions, we additionally find a dependence on the planning horizon with less savings potential for a horizon of 7 days (12.0%±8.9%) than 14 days (15.9%±9.2%) than 30 days (17.0%±9.1%). This can be explained by increased consolidation possibilities (see Appendix A).

Table 2 in Appendix A shows that as expected optimizing for transportation costs simultaneously leads to emission reductions which are more pronounced the larger the size of the collaboration and the more time periods considered. However, emission reductions are not as large as when optimizing for them only. The picture is different when optimizing for emissions only. Here, solutions for small instances with collaborating companies between two regions results in slightly increased costs (7.0%±5.9%). Interestingly, the calculated small cost increases together with the average emission reductions of roughly 30% are consistent with a real life case study conducted in the EU in which four companies (shippers) from two regions collaborated (Jacobs et al., 2013). Hence, our results showcase the existence of a minimum necessary size of the collaboration such

[5]Different to Table 1, in the next two paragraphs the standard deviations are calculated based on the average results of the different instances in the referred to instance-group and not on the instance-group averages over multiple runs.

that decisions aiming at emission reduction also result in cost savings. Note that internalizing CO_2e emissions results in a good trade-off between optimizing for costs or emissions only.

5 Conclusion

In this paper we introduced a new mathematical model for tactical transport planning in a horizontal collaboration. The model incorporates realistic tariff structures, intermodal transport, handling capacities and storage possibilities. Furthermore, it can be applied to plan transportation with a diverse set of products among others supporting perishability or special transport needs such as cooling. The model allows for sustainable planning through internalizing CO_2e costs. Hence, it can be used to minimize transportation costs, emissions or both together. Subsequently, we developed a mixed-integer formulation of the model for exact solution approaches and a memory efficient hybrid heuristic. The hybrid heuristic is composed of two parts, a matheuristic which we generalized to non-negative integer variables and a local search. These solution techniques were successfully applied to generated problem instances based on the real transportation infrastructure - including railway and shipping - in the Danube Region. This revealed significant saving potentials in both costs and emissions. As expected, optimizing for transportation costs automatically leads to a reduced carbon footprint. However, optimizing for the carbon footprint only necessitates a minimum size of the collaboration for transportation costs to simultaneously decrease.

Our results support the assumption that horizontal collaboration in warehouse-sharing and unlocking intermodal freight transport opportunities positively impact emissions and climate-neutral actions. Therefore, further efforts need to be made to promote collaboration in real-world settings.

Acknowledgements

We want to thank Seragiotto Clovis for performing the computations with Ariadne and Georg Brandtstätter for his excellent compute cluster support. This work received funding by the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) in the research program “Mobilität der Zukunft” under grant number 877710 (PhysICAL).

References

Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice hall, 1993.

Mohammad Rahim Akhavan Kazemzadeh, Tolga Bektaş, Teodor Gabriel Crainic, Antonio Frangioni, Bernard Gendron, and Enrico Gorgone. Node-based lagrangian relaxations for multicommodity capacitated fixed-charge network design. *Discrete Applied Mathematics*, 2021.

ALICE. *Global supply network coordination and collaboration: research and innovation roadmap*. 2014. URL `https://www.etp-logistics.eu/wp-content/uploads/2015/08/W46mayo-kopie.pdf`.

ALICE. *A framework and process for the development of a ROADMAP TOWARDS ZERO EMISSIONS LOGISTICS 2050*. 2019. URL `http://www.etp-logistics.eu/wp-content/uploads/2019/12/Alice-Zero-Emissions-Logistics-2050-Roadmap-WEB.pdf`.

Tolga Bektaş, Jan Fabian Ehmke, Harilaos N. Psaraftis, and Jakob Puchinger. The role of operational research in green freight transportation. *European Journal of Operational Research*, 274(3):807–823, 2019.

Dimitris Bertsimas and John Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, 1st edition, 1997. ISBN 1886529191.

Teodor Gabriel Crainic. Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272 – 288, 2000.

Teodor Gabriel Crainic, Bernard Gendron, and Geneviève Hernu. A slope scaling/lagrangean perturbation heuristic with long-term memory for multicommodity capacitated fixed-charge network design. *J. Heuristics*, 10:525–545, 09 2004.

CE Delft. *Handbook on the external costs of transport*. Publications Office of the European Union, Luxembourg, 2019.

European Union. Commission notice: guidelines on the applicability of article 81 of the ec treaty to horizontal cooperation agreements. *2001/C3/02*, 2001.

Bernard Gendron, Saïd Hanafi, and Raca Todosijević. Matheuristics based on iterative linear programming and slope scaling for multicommodity capacitated fixed charge network design. *European Journal of Operational Research*, 268(1):70 – 81, 2018.

Tobias Harks, Felix König, Jannik Matuschke, Alexander Richter, and Jens Schulz. An integrated approach to tactical transportation planning in logistics networks. *Transportation Science*, 50(2):439–460, 2016.

International Transport Forum. *ITF Transport Outlook 2019*. 2019. URL `https://www.oecd-ilibrary.org/content/publication/transp_outlook-en-2019-en`.

Kurt Jacobs, Steven Vercammen, and Sven Verstrepen. *Creation of an Orchestrated Intermodal Partnership Between Multiple Shippers*. 2013. URL `http://www.co3-project.eu/wo3/wp-content/uploads/2011/12/CO3-D4-2-exec-summary.pdf`.

Jens Kempkes and Achim Koberstein. A resource based mixed integer modelling approach for integrated operational logistics planning. *Lecture Notes in Business Information Processing*, 46:281–294, 01 2010.

Dukwon Kim and Panos M. Pardalos. A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure. *Operations Research Letters*, 24(4):195–203, 1999.

T. L. Magnanti and R. T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18(1):1–55, 1984.

Alan McKinnon, Michael Brown, Maja Piecyk, and Anthony Whiteing. *Green Logistics : improving the environmental sustainability of logistics*. Kogan Page Limited, 3th edition, 2015.

Benoit Montreuil. Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, 3(2-3):71–87, May 2011.

Charles L. Munson and Jonathan Jackson. Quantity Discounts: An Overview and Practical Guide for Buyers and Sellers. *Foundations and Trends® in Technology, Information and Operations Management*, 8(1–2):1–130, June 2015. Publisher: Now Publishers, Inc.

OECD. *Delivering the Goods*. 2003. URL `https://www.oecd-ilibrary.org/content/publication/9789264102828-en`.

Matthias Prandtstetter, Markus Straub, and Jakob Puchinger. On the way to a multi-modal energy-efficient route. In *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, pages 4779–4784, 2013.

Ralph Sims, Roberto Schaeffer, Felix Creutzig, Xochitl Cruz-Núñez, Márcio D'Agosto, Delia Dimitriu, M.J. Meza, Lewis Fulton, S. Kobayashi, Oliver Lah, Alan Mckinnon, P. Newman, Minggao Ouyang, J.J. Schauer, Daniel Sperling, and Geetam Tiwari. *Transport. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 09 2014.

David Wolfinger, Fabien Tricoire, and Karl F. Doerner. A matheuristic for a multimodal long haul routing problem. *EURO Journal on Transportation and Logistics*, 8(4):397–433, 2019.

*Table 2: This tables gives more in-depth analytics on the solutions obtained by local search. The **first column** showcases percentual improvements of the respective quantity (costs or emissions) not optimized for compared to the direct delivery solution. Hence, for emission minimization the effects on the costs are shown. For cost optimization the effects on CO_2e emissions. When optimizing for both, Table 1 shows a weighted sum of costs and emissions. Hence, here both quantities are given separate. The **second column** gives a measure on the relative intermodal consolidation. It measures in percent how much less outgoing rail, ship and cross-dock to cross-dock transport units are ordered relative to the incoming truckloads (lorry-TUs). The **third column** shows the number of ordered ship and rail TUs relative to all ordered TUs. Again, reported results are averages over all runs including their sample standard deviation over the runs. The **last column** shows the percentual increase of the average filling rate calculated over all ordered TUs compared to the direct delivery solution.*

Optimize for	Results in [%]	Costs or Emissions		Consolidation	Rail & Ship	Filling Rate
Emissions	$r2\_T*\_*$	-7.0 $\pm$ 0.2		25.4$\pm$ 0.2	6.3 $\pm$ 0.2	-5.6 $\pm$ 0.5
	$r5\_T*\_*$	13.4 $\pm$ 0.4		13.1$\pm$ 1.2	7.2 $\pm$ 0.1	7.4 $\pm$ 0.8
	$r10\_T*\_*$	17.1 $\pm$ 0.6		20.2$\pm$ 1.8	7.9 $\pm$ 0.1	14.1 $\pm$ 0.8
	$r*\_T*\_*$	5.6 $\pm$ 0.3		19.4$\pm$ 0.9	6.9 $\pm$ 0.1	3.2 $\pm$ 0.7
Costs	$r2\_T*\_*$	8.5 $\pm$ 0.4		10.9$\pm$0.7	3.2 $\pm$ 0.2	2.5 $\pm$ 1.2
	$r5\_T*\_*$	23.5 $\pm$ 0.4		15.6$\pm$0.9	2.8 $\pm$ 0.1	14.6 $\pm$ 0.7
	$r10\_T*\_*$	24.3 $\pm$ 0.4		19.5$\pm$1.0	2.9 $\pm$0.1	23.5 $\pm$ 0.9
	$r*\_T*\_*$	17.4 $\pm$ 0.4		14.3 $\pm$0.8	3.0 $\pm$ 0.1	11.1 $\pm$ 1.0
Both (Costs | Emissions)	$r2\_T*\_*$	4.8 $\pm$ 0.3	20.0 $\pm$ 0.8	10.2$\pm$0.5	4.3 $\pm$ 0.2	1.5 $\pm$ 1.0
	$r5\_T*\_*$	19.8 $\pm$ 0.4	28.3 $\pm$ 0.4	13.9$\pm$0.9	4.0 $\pm$ 0.1	14.5 $\pm$ 0.8
	$r10\_T*\_*$	23.6 $\pm$ 0.6	27.8 $\pm$ 0.4	18.2$\pm$1.1	4.1 $\pm$ 0.1	22.7 $\pm$ 0.8
	$r*\_T*\_*$	14.3 $\pm$ 0.4	24.8 $\pm$ 0.5	13.1$\pm$0.7	4.1 $\pm$ 0.1	10.5 $\pm$ 0.9

A Further Results

Table 2 highlights intermodal consolidation and the use of greener transport modes such as ship and rail. It shows that when optimizing for costs or jointly for costs and emissions, consolidation effects are more pronounced when the size of the collaboration increases. Furthermore, internalizing emissions leads to higher usage of rail and ship connections. The small decrease in the measured consolidation effect when optimizing for both objectives compared to costs only can be explained due to the fact that greener intermodal opportunities start to be viable with less transport volume. This is due to the fact that emission reductions compensate high intermodal consolidation requirements which are otherwise necessary so that the chosen intermodal route is cost-efficient. When optimizing for emission the results show that there are a lot of consolidation and intermodal routing opportunities which could drastically cut emissions but are not yet viable from a cost perspective. An emission minimizing solution results in more than double the use of intermodal containers than optimizing for costs only.

Note that Table 2 only shows intermodal consolidation. Another types of consolidation involves shipments between warehouses. These can happen either to send out consolidated shipments or to bring in a large volume of products into a region to satisfy its demands over multiple time periods using one consolidated shipment. Such a large shipment of products then requires storage in a regional warehouse over one or multiple time periods. We find when optimizing for emissions, TUs on warehouse to warehouse connections make up on average 3.7%$\pm$4.2%[6](0% - 14.9%) of

[6]In this paragraph in-group standard deviations are reported.

the total number of ordered TUs. When optimizing for costs they make up 2.6%±2.5%(0% - 8.4%) and when optimizing for both they make up 2.4%±2.5%(0% - 8.5%). This small decrease when optimizing for both objectives can be explained by the increased intermodal usage. Without horizontal collaboration (in the direct delivery solutions) this quantity is effectively zero. Lastly, the average delivery time of a shipment on the last arc to a demand node drops on average to 0.3±0.2 days when optimizing for emissions compared to on average 1.2±0.2 days in the direct delivery solutions. Optimizing for costs results in 0.7±0.1 days and optimizing for both objectives results in 0.6±0.2 days. This indicates that products are very often shipped into a demand region by means different to direct delivery by lorry.

B Tariff-Expanded Network

Due to the complex tariff structures employed, if one would define a mixed-integer model on the time-expanded graph $G_{\mathcal{T}}$, the arcs representing transport relations would have highly non-linear cost functions. This would make it difficult to apply established and powerful solution techniques for integer linear programs. To address this problem, two specific concepts are used. First, we introduce a second set of decision variables y representing the number of installed *transport units (TUs)* on each arc. A TU represents a certain amount of capacity bought for product flow on this arc and is a non-negative integer variable. These TUs can be containers, trucks or similar. This concept naturally leads to a capacitated network design formulation but alone is not enough to result into a linear formulation. A similar concept is used by Harks et al. (2016) and more general *units of facility installed* (taken from Crainic (2000)) is common in the network design literature.

Secondly, each simple arc in $G_{\mathcal{T}}$ representing a transport relation is replaced by a more complex graph structure known as *graph gadgets*. These gadgets are again made up of arcs and nodes and - together with the above decision variables - allow to linearly model the non-linear cost structures induced by the employed tariffs. The resulting network $G = (\mathcal{V}, \mathcal{A})$ constructed from $G_{\mathcal{T}}$ is called *tariff-expanded*. Employing tariff-expansion allows to derive a linear mixed-integer model defined on G which is a variant of the fixed-charge network flow problem.

B.1 Graph Gadgets

Graph gadgets were first introduced by Harks et al. (2016) for their employed transportation tariffs. In this section, we extend their approach by deriving graph gadgets for tariffs with constant cost levels t_{const} and for tariffs with piecewise-linear cost levels t_{lin}. We assume the cost levels are discount structures, i.e. having monotonic cost levels $C_n(x,y) \leq C_{n-1}(x,y)$. We start by describing the graph gadget construction for the piecewise-linear case. First, the number of transport containers depending on the flow $\tilde{y}$ is replaced by the decision variable y leading to a cost function $C(x,y)$ depending on both x and y. Therefore, it is possible and done in practice to book more TUs than necessary to get a higher discount.

For the second step, assume one t_{lin} is active on a transport relation. Then, the graph gadget is constructed using the steps outlined in figure 2. At first, the transport relation is replaced by one arc for each cost level $n \in \mathcal{N}$ with cost function $C_n(x,y)$ (see Figures 2a and 2b). Each of these arcs is replaced by the graph structure shown in Figure 2c. Variable y_1 associated to e_1 is set to be binary, i.e. $y_1 \leq 1$. The TU-capacity of e_1 is set to infinity. The costs of e_1 are set to the starting cost of the tariff level this graph structure represents, i.e. $c(e_1) = b_n$. This means any flow over this graph structure has to pay at least the starting costs of this tariff. As these already include the costs of $\beta_n - 1$ fully filled TUs, $c(e_2) = 0$ with $y_2 \leq \beta_n - 1$. As b_n also includes the price of the

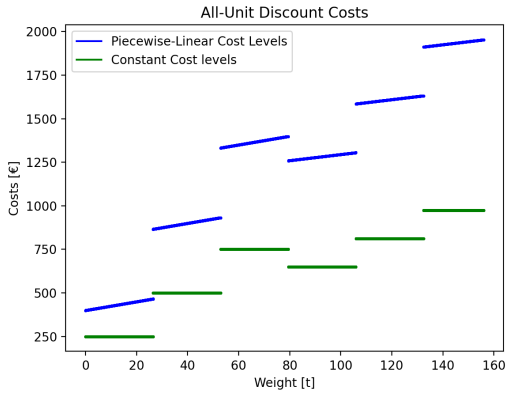
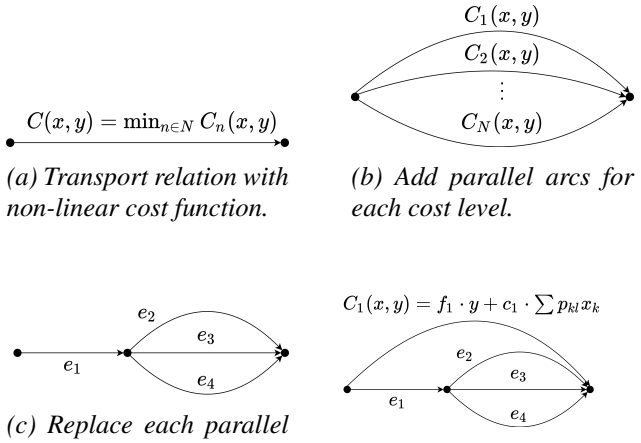


Figure 1: In this work, we use constant cost levels for rail and ship connections and piecewise-linear cost levels for lorry tariffs. The displayed tariffs have one cost level for the first three transport units (TUs) and a second discounted cost level starting with the fourth ordered TU. In this work, one TU always corresponds to one 40-feet ISO container.



(a) Transport relation with non-linear cost function.

(b) Add parallel arcs for each cost level.

(c) Replace each parallel arc with four arcs and one gadget node.

(d) Simplify graph structure of first tariff level.

Figure 2: Graph gadget construction.

empty β_n-th TU, costs on e_3 are set to the linear costs only $c(e_3) = c_n \cdot P_l^\sigma(x_3)$ and e_3 is restricted to be chosen only once $y_3 \leq 1$. The capacity in property $l \in \mathcal{P}$ of each TU on e_2 to e_4 is set to the physical capacity B_l^σ of one real TU employed on this transport relation. Lastly, the costs of e_4 are set to the cost rate of the new tariff level $c(e_4) = f_n \cdot y_4 + c_n \cdot P_l^\sigma(x_4)$. An optimal solution will automatically first fill arc e_2 then e_3 and only then use e_4. If each tariff level is replaced by such a graph structure, again by an argument of optimality, only this graph structure will be chosen resulting into the minimum cost tariff level for a given flow x. However, the graph structure for the first tariff level can be simplified to only one arc with $C_1(x,y) = f_1 \cdot y + c_1 \cdot P_l^\sigma(x)$ as $b_1 = \beta_1 = 0$. If only a small number of tariff-levels exists, this significantly reduces model size. The graph gadget developed includes a graph gadget for constant cost tariff levels as special case when $c_n = 0$.

From a pure mathematical perspective, it is possible to model the employed tariffs using only two outer edges. However, this would not preserve the number of employed TUs for counting purposes necessary for handling capacity restrictions (see Section B.4).

B.2 Decision Variables

To formulate the objective and constraints, we formally define the decision variables on the network $G = (\mathcal{V}, \mathcal{A})$. Note that the node set $\mathcal{V}$ decomposes into a set of facilities $\mathcal{F}$, demand nodes $\mathcal{D}$, bin nodes $\mathcal{B}$ and gadget nodes $\mathcal{G}$ and that each nodes is uniquely identified by in index-time pair (i, τ). As a result, an arc $a \in \mathcal{A}$ connecting (i, τ) with (j, τ') is identified by a five tuple (i, j, τ, τ', m) with m representing the type of the transport relation. Arcs connecting to and from a gadget node inherit the transport mode m of the original transport relation.

The *flow of non-perishable goods* $f_a^k \in \mathbb{R}_0^+$ is defined on all arcs $a \in \mathcal{A}$ and for each non-perishable commodity $k \in \mathcal{K}_\mathcal{T}$. The *flow of perishable goods* $f_a^{k_t} \in \mathbb{R}_0^+$ is defined on all arcs $a = (i, j, \tau, \tau', m) \in \mathcal{A}$ with $\tau' \leq t + \Delta t_k$ and for each perishable commodity $k \in \mathcal{K}_.$. The second decision variable $y_a^\sigma \in \mathbb{N}_0$ describing the *number of transport units* of type $\sigma \in \Sigma$ installed is defined for all $a \in \mathcal{A}$

and $\sigma \in \Sigma$. All flow-variables on an arc a can be collected into a flow vector x_a. These can be again stacked to one big flow vector x. Analogously a TU vector y is constructed.

B.3 Objective

The objective consists of a weighted sum of costs and emissions (for a review on green network design see McKinnon et al. (2015), for green planning techniques in general see Bektaş et al. (2019)). The (linear) cost term includes system-wide transportation, handling and storage costs as given by equation (5). Handling costs are incorporated into the transportation arcs and dependent on the connected facilities and mode of transportation.

$$C(x,y) = \sum_{a \in \mathcal{A}} \sum_{\sigma \in \Sigma} \left[f_a^\sigma y_a^\sigma + c_a^\sigma P_{l(a)}^\sigma(x_a) \right] \tag{5}$$

Fixed-costs per transport unit of type σ on arc a are f_a^σ. Flow extent in property $l \in \mathcal{P}$ of type σ on arc a is linearly priced by c_a^σ. The priced flow-extent property $l \in \mathcal{P}$ depends on the employed tariff. Therefore, it is dependent on the looked upon arc a and written as a function thereof $l(a)$. The emissions term (6) looks very similar, except it replaces fixed-costs with fixed CO_2e emissions Δ_a^σ per TU and variable CO_2e emissions δ_a^σ per flow extent.

$$\Delta(x,y) = \sum_{a \in \mathcal{A}} \sum_{\sigma \in \Sigma} \left[\Delta_a^\sigma y_a^\sigma + \delta_a^\sigma P_{l(a)}^\sigma(x_a) \right] \tag{6}$$

The combined objective reads

$$\min_{x,y} \lambda_1 C(x,y) + \lambda_2 \Delta(x,y) \tag{7}$$

Different choices of λ_1 and λ_2 allow different CO_2e pricing schemes. CO_2e pricing can be ignored by setting $\lambda_2 = 0$. Minimizing for emissions only is enabled by setting $\lambda_1 = 0$.

B.4 Constraints

The model is presented in a cut-set formulation and $\delta^-(v)$ and $\delta^+(v)$ have their usual meaning of all incoming or outgoing arcs, respectively, from node $v \in \mathcal{V}$. Additionally, mode-specific cut-sets are defined for each facility $v \in \mathcal{F}$ with $\delta_m^-(v)$ representing the set of all incoming arcs of type m. One could define $\delta_m^+(v)$ analogously. But to formulate handling capacity constraints in the model, we need to capture the correct number of outgoing transport units of a specific type using the arcs in $\delta_m^+(v)$. In its analogous definition, it will be distorted due to the graph-gadgets. To correct this distortion, in $\delta_m^+(v)$ we don't include the arcs connecting node v to graph-gadget nodes and in return add the outgoing arcs of graph-gadget nodes of type m.

The constraints read as follows:

$$\sum_{a\in\delta^+(v)} x_a^k - \sum_{a\in\delta^-(v)} x_a^k = b_v^k \qquad \forall k \in \mathcal{K}_\mathcal{T}, v \in \mathcal{V}\setminus\mathcal{B} \tag{8}$$

$$\sum_{a\in\delta^+(v)} x_a^{k_t} - \sum_{a\in\delta^-(v)} x_a^{k_t} = s_v^{k_t} \qquad \forall k \in \mathcal{K}_\Delta, t \in \mathcal{T}, v \in \mathcal{F}\cup\mathcal{G} \tag{9}$$

$$\sum_{t=0}^{\tau}\sum_{a\in\delta^-(v)} x_a^{k_t} = \omega_v^k \qquad \forall k \in \mathcal{K}_\Delta, v = (i,\tau) \in \mathcal{D} \tag{10}$$

$$\sum_{k\in\mathcal{K}_\mathcal{T}^\sigma} p_{kl}x_a^k + \sum_{k\in\mathcal{K}_\Delta^\sigma}\sum_{t=0}^{\tau} p_{kl}x_a^{k_t} \le B_{al}^\sigma y_a^\sigma \qquad \forall l \in \mathcal{P}, \sigma \in \Sigma, a \in \mathcal{A} \tag{11}$$

$$y_a^\sigma \le u_a^\sigma \qquad \forall \sigma \in \Sigma, a \in \mathcal{A} \tag{12}$$

$$\sum_{\sigma\in\Sigma}\sum_{a\in\delta_m^\circ(v)} y_a^\sigma \le h_v^{m\circ} \qquad \forall \circ \in \{+,-\}, m \in \mathcal{M}, v \in \mathcal{F} \tag{13}$$

$$\sum_{\sigma\in\Sigma}\sum_{a\in\delta_m^-(v)\cup\delta_m^+(v)} y_a^\sigma \le h_v^m \qquad \forall m \in \mathcal{M}, v \in \mathcal{F} \tag{14}$$

The constraints can be grouped into *flow-conservation constraint* (8)-(10) and *capacity constraints* (11)-(14). Constraint (8) defines flow-conservation for non-perishable goods with a positive b_v^k to represent a source for commodity k and negative to represent demand for commodity k. The following two constraints handle the more complex case of perishable goods. Equation (9) concerns the conservation of flow through facility and gadget nodes. For gadget nodes $s_v^{k_t} = 0$ as they have no supply of goods. If a facility acts as a source for commodity k in time period t, $s_v^{k_t} > 0$. The third flow-conservation constraint (10) regards demand satisfaction of perishable goods. Each region has a time-dependent demand ω_v^k for perishable commodity k which must be exactly fulfilled by its incoming flow. Demand for a commodity k can be satisfied by any commodity k_t independent of its production timestamp t.

Constraint (11) links the flow of goods with the necessary transport units. B_{al}^σ refers to the maximal extend of property l a TU on arc a for commodities of type σ can transport. For this study on arcs of mode $m \in \mathcal{M}$, B_{al}^σ always corresponds to one 40-ft ISO container except in the case when a graph-gadget requires adjusting the capacity. Constraint (12) establishes a maximum number of TUs for type σ installable on arc a. Together with (11) it capacitates the flow on a given arc. Storage is capacitated by setting $B_{al}^\sigma = 1$ and u_a^σ to the real world warehouse capacity for this property. Equation (13) capacitates incoming $(-)$ and outgoing $(+)$ handling operations separately. This captures the operational truth in cross-docks with separated incoming and outgoing docks or conceptually similar architectures like transshipment-points with mode-switches. Lastly, (14) capacitates the sum of containers handled both for incoming and outgoing operation. This is relevant for facilities like a classic warehouse with a specific number of docks for lorries not strictly split into incoming and outgoing docks. During handling commodity types are not distinguished (exemplary, a container crane is oblivious to the fact that a specific container has a cooling module or not).

Note that we have devised gadget-specific strengthening constraints for the mixed-integer formulation presented. However, they only result in small improvements on the solution performance not enough to solve most of the large instances. Therefore, we omit their presentation.

C Slope Scaling - Formal Model

The linear programming objective takes the form

$$v(\mathcal{LP}(\rho(t))) = \min_{x} \sum_{a\in\mathcal{A}} \sum_{\sigma\in\Sigma} \left[\left(\lambda_1 c_a^\sigma + \lambda_2 \delta_a^\sigma + \rho_a^\sigma(t)\right) \cdot P_{l(a)}^\sigma(x_a) \right] \tag{15}$$

with x defined as in the original problem (compare to the $\mathcal{MIP}$-objective in Appendix B.3). The linear program uses the same flow-conservation constraints (8) - (10). The other constraints are adapted as follows

$$\sum_{k\in\mathcal{K}_\mathcal{T}^\sigma} p_{kl} x_a^k + \sum_{k\in\mathcal{K}_\Delta^\sigma} \sum_{t=0}^{\tau} p_{kl} x_a^{k_t} \leq B_{al}^\sigma u_a^\sigma \qquad \forall l \in \mathcal{P}, \sigma \in \Sigma, a \in \mathcal{A} \tag{16}$$

$$\sum_{\sigma\in\Sigma} \sum_{a\in\delta_m^\circ(v)} P_l^\sigma(x_a) \leq B_{ml}^\sigma h_v^{m\circ} \qquad \forall l \in \mathcal{P}, \circ \in \{+,-\}, m \in \mathcal{M}, v \in \mathcal{F} \tag{17}$$

$$\sum_{\sigma\in\Sigma} \sum_{a\in\delta_m^-(v)\cup\delta_m^+(v)} P_l^\sigma(x_a) \leq B_{ml}^\sigma h_v^m \qquad \forall l \in \mathcal{P}, m \in \mathcal{M}, v \in \mathcal{F} \tag{18}$$

Constraint (16) merges (11) and (12). The adapted handling constraints (17) and (18) use the mild assumption of same TU property-extents B_{ml}^σ on one mode m, valid for all our input instances. An exact solution $\tilde{x}$ of $\mathcal{LP}(\rho(t))$ can be calculated efficiently by linear programming algorithms. If an instance has no capacitating handling constraints, $\mathcal{LP}(\rho(t))$ takes the form of a multicommodity minimum-cost flow problem and we find for this case the *network simplex* algorithm (Bertsimas and Tsitsiklis, 1997) outperforms more general simplex algorithms.

As described in the main text, after solving $\mathcal{LP}(\rho(t))$, a solution $(\tilde{x},\tilde{y})$ to $\mathcal{MIP}$ is obtained by setting $\tilde{y}_a^\sigma = \max_{l\in\mathcal{P}}\{\lceil \frac{P_l^\sigma(\tilde{x}_a)}{B_{m(a)l}^\sigma} \rceil\}$ ($m(a)$ denotes the mode used by arc a). Implicitly estimating the number of TUs by summing up flow-extent over multiple arcs as done in (17) and (18) could lead to minor differences with the actual number of used TUs. If in instances with handling constraints some are slightly violated, we repair the solution by rerouting flow exceeding handling capacities using a variant of our local search moves (see Section 3.2).

Furthermore, we apply slope scaling on the much more memory-efficient time-expanded network by replacing $a \in \mathcal{A}$ with $a \in \mathcal{A}_\mathcal{T}$ and in each iteration adapt the cost factors c_a^σ and f_a^σ to the best available tariff level for the chosen $\tilde{y}_a^\sigma$. To the best of our knowledge, this results into the first application of a slope scaling procedure to a changing cost function.