Synthetically generating traffic scenarios for simulation-based container terminal planning

Marvin Kastner*, Ole Grasse
Institute of Maritime Logistics, Hamburg University of Technology (TUHH),
Hamburg, Germany
*Corresponding author: Marvin Kastner; `marvin.kastner@tuhh.de`

Keywords: Container Terminal; Terminal Planning; Economic Traffic Forecast; Simulation Study; Synthetic Data; Traffic Scenario Generation

Summary

More than 80 % of world trade is delivered via sea, making the maritime supply chain a very important backbone for the economy (UNCTAD, 2020). Containerized trade regularly outperforms other types of transport in terms of growth, coinciding with consistent increases of average container vessel sizes (UNCTAD, 2020). Container terminal operations are heavily affected by this development, since less but larger port calls create unwanted peaks and stress on the terminals and the hinterland. Not all container terminals are affected equally by the described situation. Economic cycles and events such as the global COVID-19 pandemic or the Russian war in Ukraine change the global supply chains, trade characteristics and transport demands between ports in the world.

In 2004, Hartmann proposed an approach to create scenarios for simulation and optimization in the sense of container terminal planning and logistics (Hartmann, 2004). Due to the significant changes in maritime trade over the years, a new approach for generating synthetic container flow data became practical. In 2021, we introduced a rethought and reworked approach on this topic. The proposed tool, named *ConFlowGen*, aims to assist planners, scientists, and other maritime experts with providing comprehensive container flow scenarios based on minimal inputs and assumptions of the user. In this paper, we introduce *ConFlowGen*'s general principle of operation in an exemplary use case in the context of container terminal planning.

1 Context and motivation

Planning a container terminal involves many interrelated activities to ensure that the constructed terminal actually meets the expectations of the shareholders and stakeholders. As one of the first steps, the expected vessel fleet mix and expected cargo flows in an economic region are forecast for the next decades; in the scope of a subsequent analysis, it is then decided how to position the terminal in the market (Port Planning and Construction Committee, 2001). The container terminal operator (or a port planner taking their perspective) needs to decide which cargo flows they try to attract to operate profitably. Based on the cargo flows that are expected to be routed through the terminal, the quay side, land side, and yard area are laid out (Wiese, Suhl, & Kliewer, 2011). At each of the interfaces (typically serving deep sea vessels, feeders, barges, trains, and/or trucks), typically short waiting times and high productivity are required. The yard area serves as a buffer between the discharging of a container from one vehicle and its loading onto another vehicle. The required yard capacity is then estimated based on the expected container volume p.a. and the expected container dwell time. In other words, the expected cargo flows and their characteristics are crucial when planning container terminals.

The first drafts of the quay side, land side, and yard area of the container terminal are often based on rough calculations using expected annual figures and average values (Chu & Huang, 2005). Each draft might differ in terms of its layout or operating system, e.g., for transporting laden containers straddle carriers, rail-mounted gantry cranes and automated guided vehicles, or rubber-tired gantry cranes and tractor/trailer units are commonly used. Simulation comes into play once more specific questions arise and expected terminal key performance indicators are required (Kastner, Lange, & Jahn, 2020). Such questions could be "how are terminal operations affected by supply chain disturbances?" or "can the terminal reach a certain berth productivity when two deep sea vessels are simultaneously served?" By constructing a digital twin of each terminal draft and experimenting with it, the stakeholders and shareholders gain first insights. These then support the decision making processes during the planning phase. In this context, the abstract long-term traffic forecasts need to materialize in concrete traffic scenarios consisting of virtual deep sea vessels, feeders, barges, trucks, and/or trains. These vehicles arrive at the digital twin of the terminal, request some or all containers to be discharged and some to be loaded before they depart again. Such traffic scenarios need to be created automatically and fast so that the digital twins can be thoroughly tested before selecting one of the drafts to be realized.

Whenever a simulation study is executed that covers the container handling processes at a terminal, inevitably at some point a traffic scenario is generated. Somehow, the vehicles and containers must enter and leave the container terminal. The data generation is often fully integrated into the simulation model. Hartmann (2004) has taken a different approach. He generates traffic scenarios by means of a separate software. Based on assumptions and random distributions, complex scenarios are generated. The data is then exported in a tabular format digestible by a simulation model. The same concept of separate scenario generation is further elaborated by Kastner, Grasse, and Jahn (2022) who programmed an adjusted version of the algorithm in Python and shared the implementation online under a free license. The software is called *ConFlowGen* and is invoked via an application programming interface. This allows the user to adjust the assumptions and distributions to their needs. To the best knowledge of the authors, this is currently the only freely accessible tool for traffic scenario generation at container terminals. But is it a suitable tool to support terminal planners?

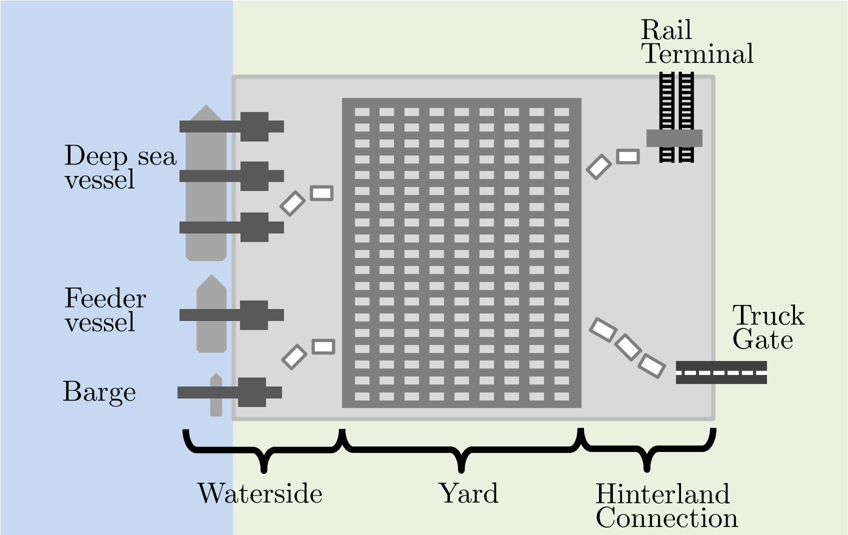

Figure 1: A container terminal and its traffic interfaces considered by *ConFlowGen*

2 On traffic scenario generation with *ConFlowGen*

The tool *ConFlowGen* aims to provide comprehensive, synthetic but yet realistic scenarios in a machine-readable data format for application in simulation studies or other data-driven applications. It's perspective focuses on container terminals including their interfaces to different modes of transport, see Figure 1. *ConFlowGen* allows users to quickly adjust their scenario inputs (e.g., vessel schedules, properties, or distributions) and provides human-readable descriptions and analyses for the resulting scenario.

In order for *ConFlowGen* to generate scenarios, some input data are required. First, assumptions about vehicles as well as their properties and arrival information (e.g., schedules) need to be put in. This can be done for deep sea vessels, feeders, barges, and/or trains. The user can either provide comprehensive sailing lists in a tabular format or define the schedules by setting individual arrival frequencies and the number of inbound containers. As the only vehicle type, trucks are created according to the demand resulting from the other transport modes while approximating the truck arrival distribution at the same time. The number of outbound containers for each vehicle is determined by the number of inbound containers in combination with the assumed modal split. An example for such an origin-destination container flow on vehicle type level is shown in Figure 4. *ConFlowGen* also needs inputs about container properties, including expected distributions of lengths, weights, types, dwell times, and their destinations. Lastly, *ConFlowGen* expects some general scenario details. This includes the first and last day of the schedules to be considered and the vehicle type-dependant modal splits.

During the generation process, vehicle instances are generated according to the provided schedules. For each vehicle, the required amount of container instances is generated that it delivers to the container terminal on its inbound journey. Based on the origin-destination container flow distribution and the container dwell times, each container is assigned a later vehicle to leave the terminal again. In the beginning, *ConFlowGen* starts with an empty yard, and after the very last vehicle has departed, the yard is empty again.

After the generation process, *ConFlowGen* can export the resulting containers and vehicles including all their properties and transportation details in an instance-based, tabular format for further data processing in other software. In addition, *ConFlowGen* also provides comprehensive statistical information, analyses, and visualisations for the generated scenario.

In detail, *ConFlowGen* provides two major types of output. In addition to the resulting data including corresponding visuals after the generation process, *ConFlowGen* also provides detailed previews for the user before the generation process. These assist the user to cross-check their input data and to estimate the impact of those inputs. This is further elaborated in Section 4.2.

3 Method

To vividly demonstrate the purpose, functionality, and features of the tool, in this paper *ConFlowGen* is applied and described along an exemplary use case. In this use case, a traffic scenario is generated and analyzed in depth, which is based on real-life data of the Container Terminal *Tollerort* (CTT) in Hamburg, Germany. The example scenario uses publicly available information such as vessel arrival data and the published port modal split. Step-by-step, our assumptions and input data that are used to generate the traffic scenario are explained. Afterwards, the resulting synthetic scenario is analyzed, discussed and visually validated together with our made assumptions.

In the first step, it is shown how *ConFlowGen* supports to calculate the numbers. Just by providing the assumed traffic distribution and vessel arrivals, the dimensions of the container volume from and to the hinterland are automatically determined. This functionality is called preview in *ConFlowGen* because it is just a rough calculation and not yet a scenario with actual vehicle and container instances. The second step covers the actual generation of the synthetic traffic scenario and the calculation of descriptive statistics. These results should be close to the initially used distribution parameters (e.g., container dwell time, truck arrival distribution, etc.). Such traffic scenario instances are then subject to analysis. The run analyses are based on container instances and consumable by, e.g., simulation models and/or digital twins.

Within this paper's exemplary use case, we demonstrate how *ConFlowGen* can provide valuable assistance when dealing with planning and layout problems, as well as other terminal-related questions.

Typical questions where *ConFlowGen* may be a helpful solution are, e.g., the following:

- Seaside: What is the expected seaside throughput? How large are the peaks created by delayed vessels?
- Hinterland: How many trucks are expected per day or hour to pass the truck gate of the terminal?
- Yard: How many ground slots are needed? What is the required stacking height?

Once we have arrived at a traffic scenario of such detail, the expectations of the container terminal yet to be planned are set. Surely, the amount of traffic must be somehow handled at each interface, i.e., long queues must be avoided as well as traffic jam inside the terminal. Likewise, all the containers must be accommodated in the laden and empty container yard so that sufficient capacity must be available. Based on the traffic scenario, each of the subsystems can be dimensioned accordingly. This is what *ConFlowGen* also supports by creating matching analyses.

Next, the planner needs to examine the system behavior and to estimate the actual productivity of the terminal and its subsystems (e.g., seaside, berth, laden container yard, empty container yard, truck gate, and rail gate). As the interaction of the subsystems is complex, this is further analyzed best within a simulation study based on traditional simulation models or more detailed digital twins. *ConFlowGen* supports the export of traffic scenarios to default tabular formats that are digestible by such software.

4 A synthetic data generation use case

In this chapter, *ConFlowGen* is exemplary applied by creating, visualizing and discussing a synthetic traffic scenario based on the existing container terminal CTT in Hamburg, Germany.

4.1 Scenario description

For our use case, we focus on the month of June 2022 as the observation window. To take into account unavoidable ramp-up and ramp-down processes in the resulting traffic flow data (see Section 4.5; Figure 7), vessel data was also accumulated for the two weeks before and after the observation window, which results in a total data time frame of eight weeks (mid of May to mid of July).

To replicate the modal splits for all possible transport mode relations (vehicle type dependant modal splits), *ConFlowGen* already includes multiple universally valid distributions as default values. To match the individual characteristics of the *Port of Hamburg*'s and CTT's overall modal split (which is regularly published by the Port of Hamburg) we adapted minor parts of *ConFlowGen*'s default values to approximate the scenario output.

For the seaside of the scenario, a publicly available sailing list, published and maintained by *Hamburger Hafen Logistik AG* (HHLA), was used to deduct the quayside traffic pattern (HHLA, 2022a). The original sailing list data contained 352 entries of arriving vessels of different types and sizes for the given time frame. We filtered the data and distinguished the vessels into two main categories, namely deep sea vessels and feeders, while (for reasons of brevity) excluding non-commercial vessels, tug boats, supply vessels, as well as barges from the data set. The resulting data set includes 236 port calls and is depicted in Figure 2. The vessel sizes cover a range of 124 twenty foot equivalent (TEU) units to 23,992 TEU with accumulations at around 1,000, 5,500, 14,000 and 20,000 TEU. The distinction between feeders and deep sea vessels is defined in the data set and was not conducted by the authors. It can be seen, that the resulting seaside traffic data over time consists of frequent port calls of the rather small feeders together with rarer but regular port calls of the larger deep sea vessels.

For arriving vessels, we assume that 30 % – 60 % of the loaded containers of deep sea vessels are handled at the terminal, using a triangular distribution. For feeders, the handling ratio is set to 30 % – 80 %. Those values are based on Park and Suh (2019). *ConFlowGen* distinguishes different container types via distributions. To keep the scenario simple, we stick to the default container type distribution (see Kastner & Grasse, 2021).

Regarding the amount of trains connecting the terminal, no comprehensive data was publicly available. To generate sufficient rail traffic (train visits) for the scenario, we therefore determined the necessary

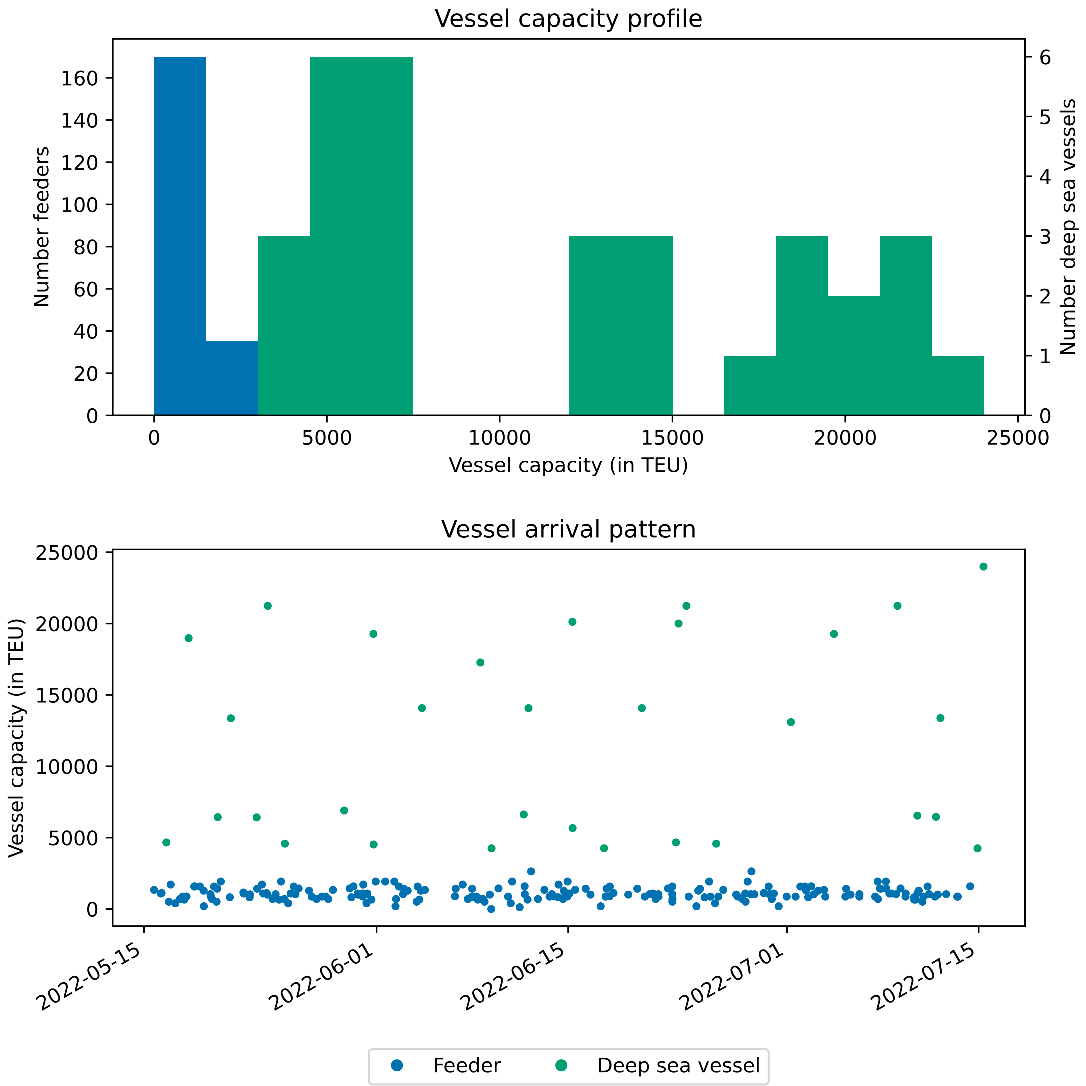

Figure 2: The vessel capacity profile and vessel arrival pattern

number of trains for each weekday to fulfill the scenarios overall transport requirements induced by the other modes of transport. We assume, that an average train can load up to 96 TEU and use *ConFlowGen*'s default distribution for the train utilization.

To keep the exemplary scenario brief and easy to understand we used some further simplifications. Barges and inland vessels are fully ignored in this scenario, which is acceptable since they are only responsible for a rather small traffic volume share. Also, we do not account for deep sea to deep sea transshipment (interlining), since Hamburg is not a major interlining-transshipment hub but rather has a strong hinterland connectivity. For simplification, also feeder-to-feeder transshipment is ignored. Furthermore, we only account for 20' and 40' container sizes while ignoring other dimensions such as high cubes, 45' containers, open boxes, etc. 20' and 40' containers are considered in a 1/3 to 2/3 ratio, (see Statistisches Bundesamt, 2022).

4.2 Reviewing the input data

As previously mentioned, *ConFlowGen* can not only create synthetic data and its analyses after an elaborated generation process, it also provides the user with a fast and guiding feedback, called previews. Previews provide a first impression on what kind of data will be generated based on the set input distributions and schedules. At this stage, some internal simplifications are made to save calculation time. Among others, no container instances are generated and operational constraints are neglected. The preview function is designed to answer typical rough-planning user questions like „will my modified schedule input still match my modal split between trucks and trains?“

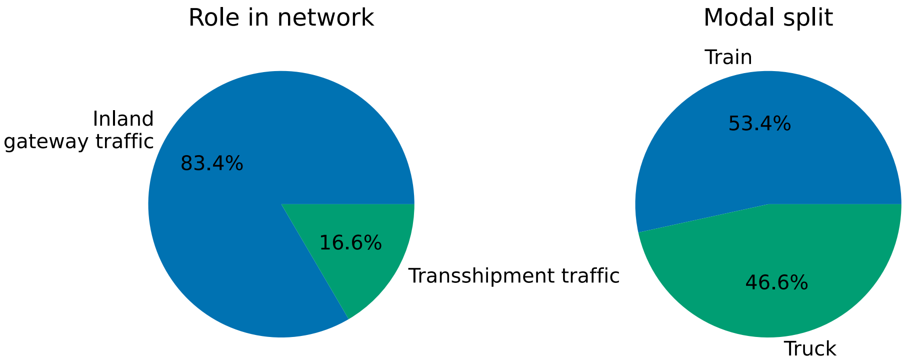

Figure 3: The role of the terminal in its network and the modal split in hinterland as estimated based on the distributions alone.

In Figure 3, two typical previews derived from the use case are shown. The left pie-chart depicts the role in network via the estimated relation of inland gateway traffic and transshipment traffic. *ConFlowGen* determines those without generating each of the container or vehicle instances. It is just an estimation based on the distributions which contain seaside to seaside (transshipment) and seaside to landside (inland gateway) transport relations. If the scenario, e.g., aims for a 20 % transshipment, planning the inputs can be adapted instantly without running the generation. On the right side, the resulting modal split between the two modes train and truck is estimated. In our example use case, we aimed for a slight surplus on the train side, which is observable in the later analyses figures as well.

After having an appropriate input scenario set and performed a visual inspection of our previews for plausibility, we can start the actual generation. This will give us the container flow data needed together with comprehensive analyses and figures.

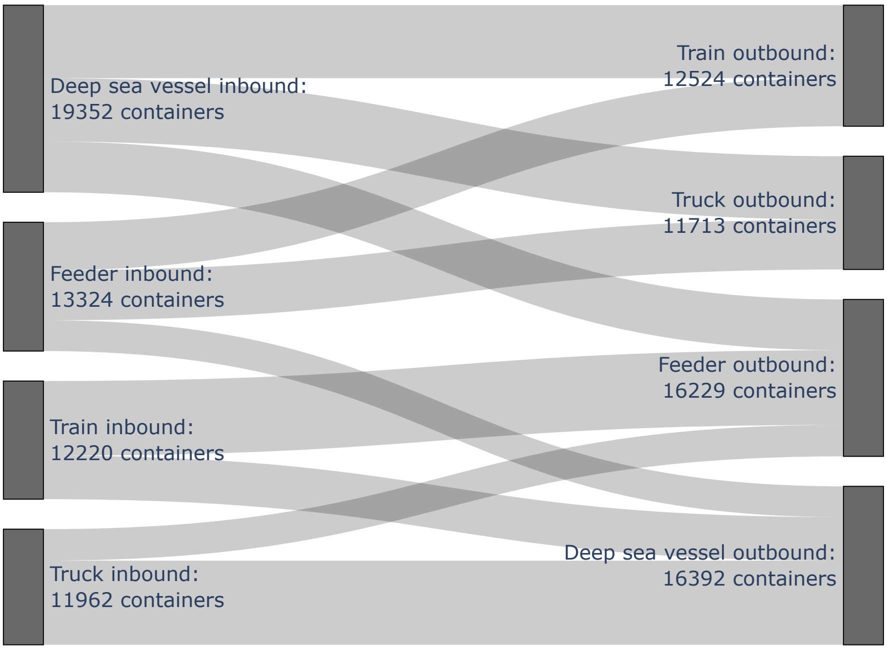

Figure 4: The container flow from its inbound vehicle type to its outbound vehicle type for June 2022 measured in containers.

4.3 Summarizing the generated container flow

After execution of the generation process, we can now take a look into the data outputs and the corresponding analyses. Each container is delivered to the terminal by one vehicle on its inbound journey and is picked up again to leave the terminal by a second vehicle on its outbound journey. This relations are depicted in Figure 4 as a Sankey diagram. For the created scenario, there are around 19,000 containers delivered via deep sea vessels, which are then transported further on trains, trucks and feeders. On the output side, the terminal exports around 16,000 containers via deep sea vessels. Therefore, regarding the deep sea traffic, the terminal can be characterized with a slight import surplus. It can be seen, that there is no direct deep sea to deep sea transshipment in the scenario. The same can be seen for feeders as there is no connection between the feeders on the inbound side (shown left) and the outbound side (shown right). Moreover, the modelled terminal implies an overall strong rail connectivity, since more than 50 % of the hinterland's inbound and outbound traffic is transported via trains. Each of those characteristics match our given scenario inputs and previews.

For each type of container (i.e., standard, reefer, empty, and dangerous goods) and inbound-outbound relationship (i.e., each pairwise combination of deep sea vessel, feeder, train, and truck), one container dwell time distribution is assumed. Exemplary, the container dwell time distribution for standard containers from deep sea vessels to trucks are depicted in Figure 5. The figure shows *ConFlowGen*'s underlying input distribution for this exact transport relation (grey line), as well as a histogram representation of the dwell times of the generated container instances (blue beams). It can be seen, that most of the containers for this typical import relation stay within the terminal for around 40 hours until they are picked up by a truck and are transported to the hinterland. A minimum processing time can also be identified in the figure, since even the fastest containers dwell for 10 to 20 hours. On the right side of the plot, it is visible that a small amount of containers need up to 210 hours until they are transported further. A distribution like this matches typical container terminals which charge a storage fee after a certain period of time.

In the case of deep sea vessels, feeders, and trains, containers are assigned their outbound vehicle according to the container dwell time and capacity availability of the vehicles. This ensures that all vehicles are working to their capacity. Trucks operate in a different way – they are created according to the truck arrival distribution in combination with the container dwell time distribution. In Figure 6, the distribution (scaled by the expected number of trucks per week for the time period) and the truck

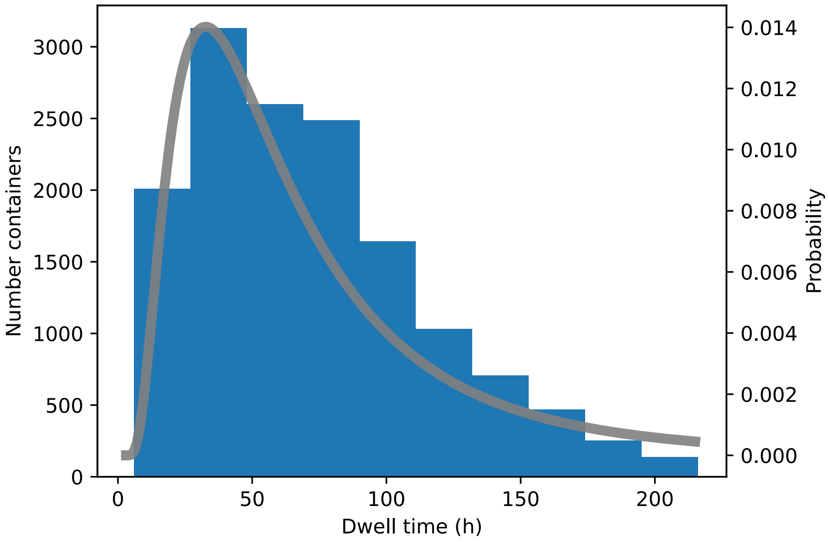

Figure 5: The container dwell time distribution for standard containers in June 2022 for the flow relation from deep sea vessel to truck.

arrivals per hour at the truck gate for each calendar week (CW) are shown. The first line graph is plotted for the ISO CW 19 in 2022 which started at 9th May and lasted until 15th May. The last ISO CW 28 started at 11th July and lasted until 17th July. In other terms, the line graphs cover the whole range when vessels have been arriving at the terminal. The initially assumed truck arrival distribution is the default distribution as provided by *ConFlowGen*. In Germany, it is forbidden to transport goods on the road on Sundays. This even affects the Saturday afternoons when close to no-one starts or ends their journey at container terminals. The largest variation of the truck arrivals is seen on Thursdays and Fridays. Since at Sundays the truck gate is closed but the seaside operations continue, export containers must be delivered by truck before Sunday. Due to the low utilization on Saturdays, this leads to peaks on Fridays. Similarly, if a container is discharged on a Sunday, this affects the truck gate throughput on the following days.

4.4 Filling and emptying the yard

When planning simulation studies, one might be tempted to start operations with a certain amount of containers inside the yard right from the beginning. This allows to approach busier operations faster. At the same time, however, each container might be moved by some equipment and thus the control systems must be able to access all relevant information regarding the arrival and future departure of a container. For each container, the information must be available and this is the reason why *ConFlowGen* generates this information independent from the later usage of the data inside the simulation. Even though the simulation might not start with an empty yard, from an information point of view there is a container that has arrived in the yard first. Likewise, there is always a container that leaves the yard last. But does the yard fill level reach a steady state where for some period of time the inbound and outbound flows are evened out?

The yard fill level, the throughput at the truck gate and the vessel arrivals at the seaside are depicted in Figure 7. As *ConFlowGen* does not have any concept of container handling times, the vessels arriving at the quay are summarized by day (24 h window) and in terms of their TEU capacity. Especially deep sea vessels might take longer than a fraction of a day to be discharged and loaded before they depart again. In other terms, if one was to set up a simulation study and actually report the berthing times and actual seaside throughput, the graph would be much smoother. While the same is theoretically true for trucks at the truck gate, the impact of time-consuming activities is much lower, especially as they are

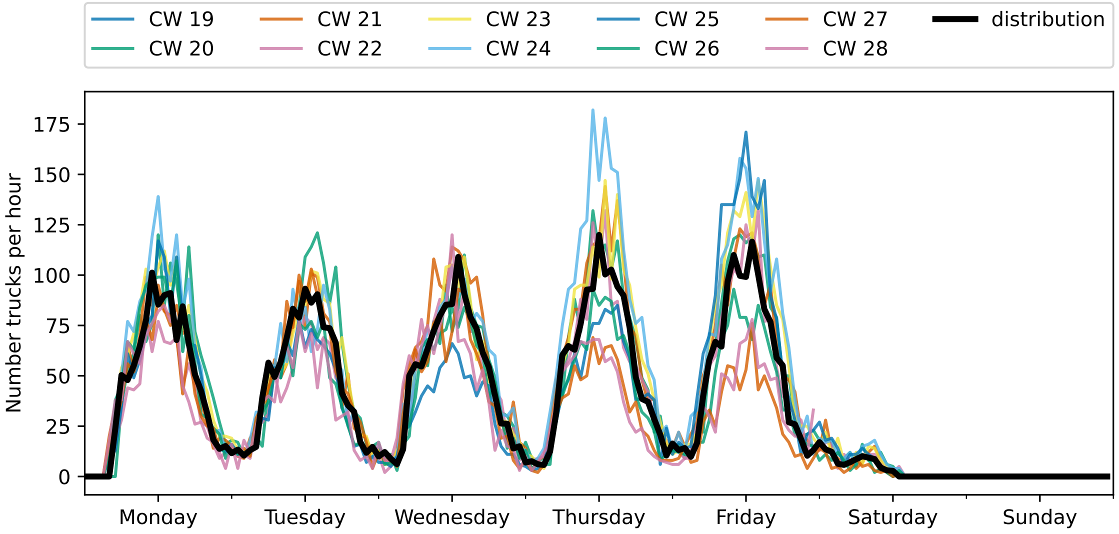

Figure 6: The initially set truck arrival distribution (in black) and the randomly drawn truck arrivals at the truck gate for each calendar week (CW).

summarized by 4h windows.

The first containers are delivered to the terminal by truck even before May (see Figure 7). These activities are very unlikely and rare in number. First clearly visible activities start in the first week of May. During this time, the containers which will be placed on vessels on their outbound journey are delivered. Starting from 15th May, the truck gate throughput seems to be in a steady state until approximately 15th July. This is the same time period during which vessels arrive at the seaside. In the last week of July and in August, the last containers that are still in the yard are picked up by trucks until the yard is empty again.

The used yard capacity in Figure 7 starts to grow slowly over May but has reached a steady state at 20,000 – 25,000 TEU even before the beginning of June. The peaks at the seaside directly impact the yard fill level. At the beginning of July, the yard fill level starts to drop again and it is far below 5,000 TEU even before August begins. This yard fill level seems reasonable when compared to other terminal figures such as the *Container Terminal Altenwerder* with 39,000 TEU (Brinkmann, 2005, p. 312). Moreover, the terminal capacity can be estimated based on the terminal area and the estimated TEU per hectare. When assuming that straddle carriers can stack 500 TEU per hectare (Brinkmann, 2005, p. 244) and the terminal area is 60 hectare (HHLA, 2022b), then the total capacity is approximately 30,000 TEU. Given the challenging shape of the terminal (HHLA, 2022b), even a yard capacity lower than 30,000 TEU is reasonable.

4.5 A closer look at the ramp up and ramp down period

In the previous section, the long period of a steady state in the yard is shown. During June 2022, the yard fill level fluctuates slightly and this reflects the change of yard utilization in real-life data. The phase from May until the beginning of June is the ramp-up period. The yard fill level in Figure 7 starts at zero and quickly approaches 20,000 TEU. The phase from the end of June until August is the ramp-down period during which the yard fill level decreases back to zero. The ramp-up and ramp-down period are necessarily part of the data generation because logically one vehicle must be the first to deliver a container to the terminal and one vehicle must be a last to pick up the last container from the terminal. The ramp-up and ramp-down periods do not reflect the main phase of operation and these time periods might thus be excluded from performance evaluation. But how is the steady state of the system achieved? This is explored in the remainder of this subsection.

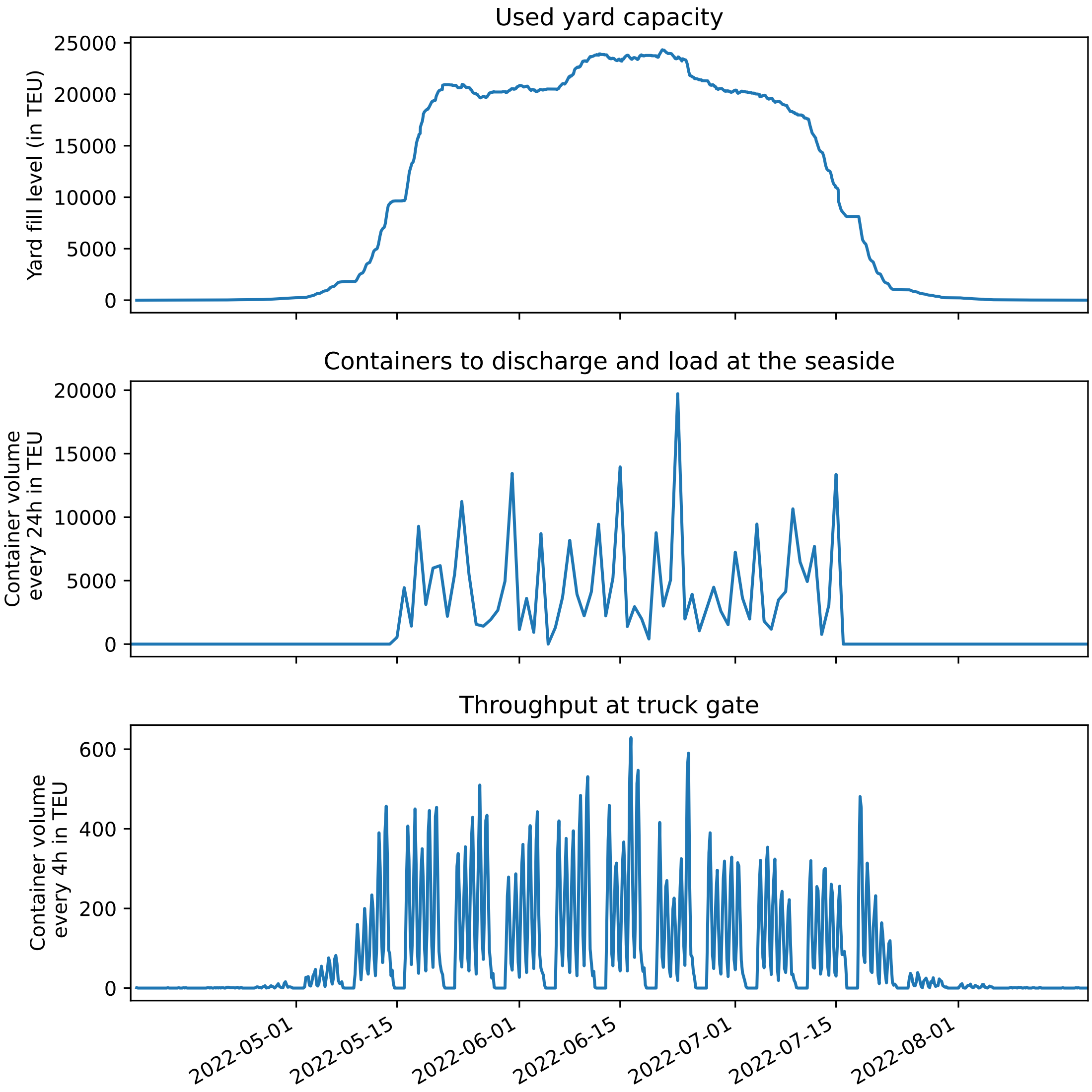

Figure 7: The relationship of yard fill level, vessel arrivals, and truck arrivals.

In the previous Figure 7, the earlier truck arrivals have been depicted. Generally, trucks are the earliest means of transport that arrive at a container terminal. They start to fill the yard for specific other vehicle types, similar to the distribution indicated by the Sankey diagram in Figure 4. Thus, there are some containers to be loaded onto the first deep sea vessels and feeders to arrive. Starting from 15th May, deep sea vessels, feeders, and trains start to arrive. As visible in Figure 8, more containers are discharged from the first vehicles than there are containers being loaded onto them. Trains start at an outbound-to-inbound ratio of close to 0 %, i.e., they leave the terminal (nearly) empty. Feeders start with an outbound-to-inbound ratio of 80 %, i.e., much closer to the balanced 100 % ratio. Over the next days, the ratio increases until a steady state of the inbound-to-outbound ratio is reached. The steady state ratio can be deduced from Figure 4. If there are more outbound containers for a vehicle type than inbound containers, as it is the case for feeders, then the expected outbound-to-inbound ratio is expected to be larger than 100 %. This means that in this case more containers are loaded onto feeders than discharged. Such unpaired container flows exist in real life due to trade imbalances and complex empty container repositioning approaches. The trains never reach an outbound-to-inbound ratio of over 100 % because they are assumed to be block trains that arrive at the container terminal fully loaded and also depart again fully loaded. Feeders and deep sea vessels are capped at 120 % outbound-to-inbound ratio while also satisfying the TEU capacity constraint of each vehicle. The cap value is a flexible assumption that could be changed by a user at any time.

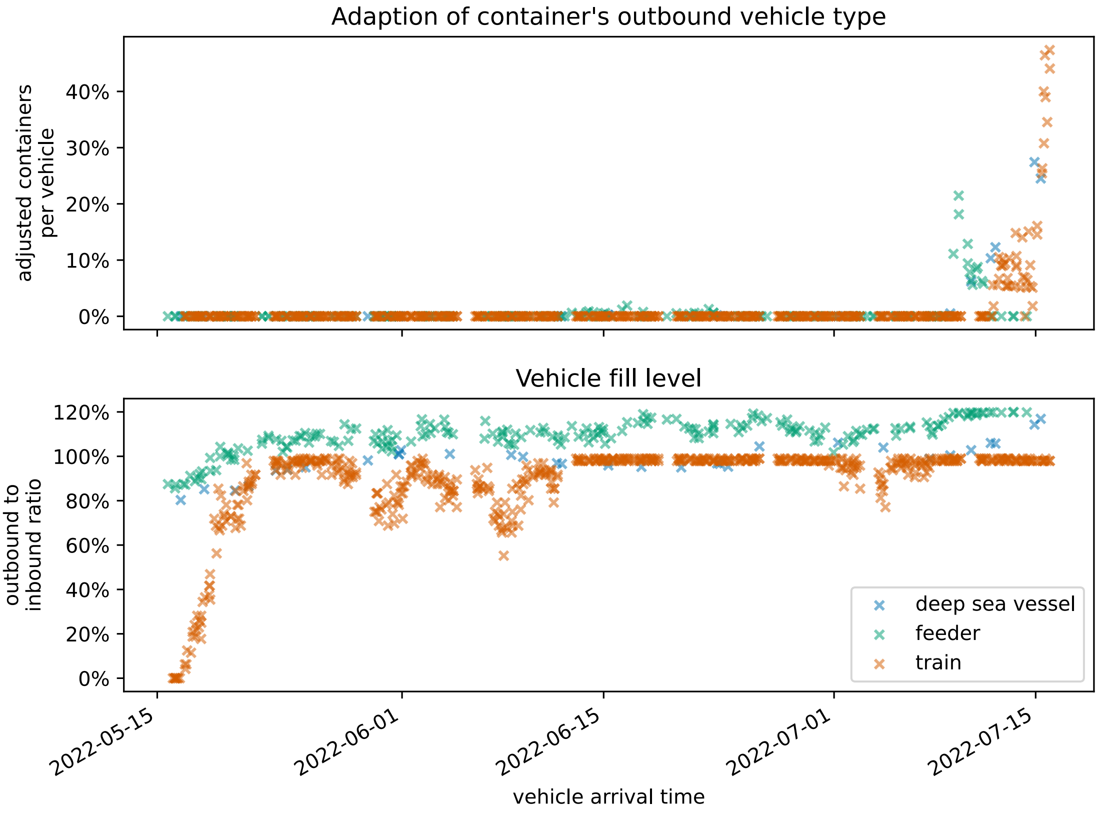

Figure 8: Outbound to inbound ratio per vehicle and the share of containers per vehicle that had to be assigned a new vehicle type

In the steady-state phase, there are a few disturbances depicted in Figure 7. Trains are sometimes not completely filled with containers. This is due to the assumed container dwell times in combination with the vessel arrival pattern. At 5th June, Pentecost took place in Germany that slowed down business activities all over the country and are thus also part of the vessel arrival patterns.

In the ramp-down phase starting in July, the outbound-to-inbound ratio slightly increases. Trains continue to be fully booked on their outbound journeys and feeders that used to already show a high outbound-to-inbound ratio now repeatedly reach the cap of 120 %. The deep sea vessels that are less in number but larger in impact on the yard now also show an outbound-to-inbound ratio of more than 100 %. The last deep sea vessel even approaches 120 %. Due to the non-existence of further vehicles after 15th June, all containers destined to leave the yard by a certain vehicle type are packed on a vehicle of the corresponding type. In some instances, this would lead to constraint violations (either the 120 % cap or the TEU capacity of the vehicle would be exceeded). In those instances, the vehicle type for the outbound journey of the container is adjusted and an alternative vehicle is looked up. In Figure 8, the share of containers that have undergone such adjustment is shown. Generally speaking, these adjustments can happen even in the steady state phase when minor supply chain hick-ups occur. The ramp-down phase of the terminal corresponds to a major change in the supply chain and thus a large amount of adjusted containers is reasonable. In the last week of July, the share of adjusted containers per vehicle (here: trains, feeders, and deep sea vessels) goes up to over 40 % for trains and 20 % – 30 % for the last two deep sea vessels. Also for feeders, values beyond 10 % are recorded. This re-assignment of containers to new vehicle types might skew the overall inbound-to-outbound flow (compare Figure 4) because vehicles are chosen according to their availability. This is another argument why the last phase might show more artifacts of the synthetic generation. Thus, it is recommended to stop recording performance metrics before the ramp-down period starts and exclude it from further evaluation.

In summary, the steady state of the yard can be observed from two perspectives. First, the yard fill level shows a clear ramp-up and ramp-down period as well as a long-lasting steady state inbetween. Second, the ramp-up and ramp-down period are identifiable by vehicle properties. The ramp-up period is distinguishable from the steady state period by a lower outbound-to-inbound ratio per vehicle. This ratio later increases until it reaches a steady state. During the ramp-down period, more containers are re-assigned to outbound vehicle types according to the remaining transport capacities. The mechanisms that lead to this specific ramp-up and ramp-down behavior are part of the general algorithm of *ConFlowGen*. Currently, it is up to the user to identify the steady state visually and decide which time period is representative enough for the simulation study or terminal planning task at hand.

5 Conclusions

In this paper, we introduce our synthetic data generation software and we show how to create valid, comprehensive container traffic scenarios with *ConFlowGen* for container terminals. The data is generated based on few assumptions and minimal information; it might further serve as an input for a simulation study time-efficiently. Even if just the traffic scenario itself is thoroughly examined (without running a simulation study), several insights can be generated.

When planning the terminal layout based on average values instead of concrete traffic scenarios, the outcome is supposed to be very similar. But especially when supply chain hick-ups or complex planning decisions are to be considered, it might be expedient and even easier to work with container instances in a detailed traffic scenario than with abstract expected values. When using concrete container traffic examples, aspects like the variation in yard fill level or the required throughput at the various terminal interfaces are already clearly visible. The variation of vessel arrivals based on example data are a more hands-on approach. *ConFlowGen* allows the users to estimate appropriate buffer capacities for the yard and the terminal interfaces based on likely scenarios.

With *ConFlowGen*, we have developed and introduced a convenient way to synthetically generate container flow data for detailed planning tasks and comprehensive scientific purposes with a sufficiently high quality. Nevertheless, the presented tool can also be understood and described as an intermediate or workaround solution to tackle a persistently predominant issue in research and applied science: Real data are often scarce, confidential and/or hard to obtain, especially, e.g., from companies in highly competitive environments. Still, those kinds of data are essential to obtain innovative research results, develop new technologies and/or to generate new insights. Synthetic data generators like *ConFlowGen* can not solve this fundamental issue but may alleviate it at least a little bit.

References

Brinkmann, B. (2005). *Seehäfen: Planung und Entwurf.* Springer Berlin Heidelberg.

Chu, C.-Y., & Huang, W.-C. (2005). Determining container terminal capacity on the basis of an adopted yard handling system. *Transport Reviews*, *25*(2), 181–199. DOI: 10.1080/0144164042000244608

Hartmann, S. (2004). Generating scenarios for simulation and optimization of container terminal logistics. *OR Spectrum*, *26*(2), 171–192. DOI: 10.1007/s00291-003-0150-6

HHLA. (2022a). *HHLA coast sailing list.* Retrieved from `https://coast.hhla.de/report?id=Standard-Report-Segelliste`

HHLA. (2022b). *Technische Daten Tollerort (CTT).* Retrieved from `https://hhla.de/unternehmen/tochterunternehmen/container-terminal-tollerort-ctt/technische-daten`

Kastner, M., & Grasse, O. (2021). *Conflowgen documentation.* Retrieved from `https://conflowgen.readthedocs.io/en/latest/index.html`

Kastner, M., Grasse, O., & Jahn, C. (2022). Container flow generation for maritime container terminals. In *Dynamics in logistics. Proceedings of the 8th international conference LDIC 2022, Bremen, Germany* (pp. 133–143). Bremen, DE: Springer Cham. DOI: 10.1007/978-3-031-05359-7_11

Kastner, M., Lange, A.-K., & Jahn, C. (2020). Expansion planning at container terminals. In M. Freitag, H.-D. Haasis, H. Kotzab, & J. Pannek (Eds.), *International conference on dynamics in logistics (LDIC 2020)* (pp. 114–123). Springer International Publishing. DOI: 10.1007/978-3-030-44783-0_11

Park, N. K., & Suh, S. C. (2019). Tendency toward mega containerships and the constraints of container terminals. *Journal of Marine Science and Engineering*, *7*(5). Retrieved from `https://www.mdpi.com/2077-1312/7/5/131` DOI: 10.3390/jmse7050131

Port Planning and Construction Committee. (2001). *IAPH guidelines for port planning and design.* International Association of Ports and Harbors (IAPH).

Statistisches Bundesamt. (2022). *Statistisches Bundesamt.* Retrieved from `https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Transport-Verkehr/Gueterverkehr/Publikationen/Downloads-Schifffahrt/seeschifffahrt-monat-2080500211084.pdf`

UNCTAD. (2020). *UNCTAD 2020.* Geneva. Retrieved from `https://unctad.org/system/files/official-document/rmt2020_en.pdf`

Wiese, J., Suhl, L., & Kliewer, N. (2011). Planning container terminal layouts considering equipment types and storage block design. In J. W. Böse (Ed.), *Handbook of terminal planning* (pp. 219–245). New York, NY: Springer New York. DOI: 10.1007/978-1-4419-8408-1_12

On the main author

Marvin Kastner received a M.Sc. degree in Informatics from the Hamburg University (UHH), Hamburg, Germany, in 2017. Shortly after graduation, he began working as a research associate and PhD candidate under the supervision of Prof. Dr.-Ing. Carlos Jahn at the Institute of Maritime Logistics at the Hamburg University of Technology (TUHH). Marvin Kastner works on simulation-based optimization in the maritime supply chain with a focus on tactical and strategic terminal planning activities at seaport container terminals. He is the main developer and maintainer of *ConFlowGen*. He has elaborated most of the concepts that are also relevant for his ongoing PhD thesis on supporting port planners to set up insightful simulation studies.