Transportation Research Part C 158 (2024) 104413

Contents lists available at ScienceDirect

Transportation Research Part C

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

Data-driven preference-based routing and scheduling for activity-based freight transport modelling

Ali Nadi [a,*], Neil Yorke-Smith [b], Maaike Snelder [a,c], J.W.C. Van Lint [a], Lóránt Tavasszy [a]

[a] Department of Transport and Planning, Delft University of Technology, the Netherlands
[b] STAR Lab, Delft University of Technology, the Netherlands
[c] TNO, The Hague, the Netherlands

ARTICLE INFO

Keywords:
Activity-based tour modelling
Freight transport modelling
Data-driven routing and scheduling
Preference-based vehicle routing
Bayesian optimization

ABSTRACT

Understanding preferences and behaviours in road freight transport is valuable for planning and analysis. This paper proposes a data-driven vehicle routing and scheduling approach for use as a descriptive tool to study road freight transport activities. The model developed seeks to capture planners' or drivers' preferences in order to reproduce observed road freight activities. The model is based on a parametrized time-dependent vehicle routing problem whose parameters can be estimated from a set of observed planned tours. We propose a Bayesian optimization technique for parameter estimation of the model. Empirical results show that the model can fit real-world data accurately and synthesize tour flows close to reality.

1. Introduction

Truck flow patterns can best be understood through the study of freight activities on a transportation network. As opposed to passenger models, freight transport activities have been less researched and hence the literature in this field is relatively limited. There are various reasons for this, among which is that transport modellers often lack observations on firms' activities, whereas disaggregate tour data collection is very expensive. Additionally, modelling freight activities is complex due to the heterogeneity in the transport markets (Khan and Machemehl, 2017; Holguin-Veras and Patil, 2007), the variety of objectives, the dynamic nature of the market, and the ambiguity in the multi-actor decision-making environment (You,Chow,and Ritchie, 2016; Gonzalez-Calderon and Holguín-Veras, 2019). Researchers have long recognized that there should be distinctions between freight and passenger transport modelling due to the complexity associated with logistic decisions (Tavasszy and De Jong, 2013; Gonzalez-Calderon and Holguín-Veras, 2019). Recently, there has been an increasing interest in descriptive agent-based freight simulation models to study these logistic decisions in the context of freight transport. Examples of this trend are TRABAM (Mommens et al., 2018), MASS-GT (de Bok and Tavasszy, 2018), and SimMobility (Sakai et al., 2020).

Tour planning is a tactical operation that firms undertake to transport commodities from the point of production to the point of consumption. Therefore, tour planning has become a crucial component in all the existing microscopic freight simulation models. Most research to date has tended to rely either on statistical/econometric methods (Hunt and Stefan, 2007; de Bok,Tavasszy,and Thoen,

* Corresponding author.
E-mail addresses: A.Nadinajafabadi@tudelft.nl (A. Nadi), N.Yorke-Smith@tudelft.nl (N. Yorke-Smith), Maaike.snelder@tno.nl (M. Snelder), J.W.C.vanLint@tudelft.nl (J.W.C. Van Lint), L.A.Tavasszy@tudelft.nl (L. Tavasszy).

https://doi.org/10.1016/j.trc.2023.104413
Received 27 March 2023; Received in revised form 4 November 2023; Accepted 9 November 2023
Available online 17 November 2023

2022) or on normative operational research tools (Schröder and Liedtke, 2017; Sakai et al., 2020) to generate tours for simulation purposes.

Models based on econometric methods provide statistical insights on tour formation through sequentially constructing tours based on choice models that predict the next trip destination given the current location of the vehicle, until a decision is made to end the tour. The most important drawback with this (trip-based) approach is the ex-ante assumption required about the end result of the tours before being able to use the model to generate tours. Examples are the number of tours generated from zones and the number of stops. Recent research, such as Nuzzolo and Comi (2014) and Thoen et al. (2020), has addressed this issue by introducing a shipment-based tour modelling where the construction of tours is based on assigning shipments to a tour rather than focusing on trips. The choice of shipment assignment is based on a generalized cost that the next shipment adds to the tour. Although the shipment-based model is the most promising tour model so far, such a model is not combinatorial and cannot capture the inter-dependent routing and scheduling characteristics of tours. That is the structure of tours and the way tours spatially and temporally form cannot be completely understood with such models.

Normative combinatorial optimization models, on the other hand, stem from a similar context to which the tours are planned in reality and thus are able to find the optimal routing and scheduling of shipment pickups and deliveries considering real-world constraints. The two main drawbacks of such models are their computation time and lack of generalization. The first issue is more manageable due to the recent advances in both hardware technologies and soft computing algorithms. The latter, however, requires a new methodology development that can deal with heterogeneity in objectives, constraints, and preferences of tour planners in various logistic sectors. The primary contribution of this paper is to bring the advantage of data-driven modelling to normative discrete optimization models, capturing the preferences of planners and even drivers in vehicle routing and scheduling of shipments. This leads to a more realistic tour model for use in freight simulators.

This paper is structured as follows. In the next section, we review related literature on tour modelling. Then we propose a methodology for a data-driven vehicle routing and scheduling model. Afterwards, the model is applied to the case of freight transport in the Netherlands. We conclude the paper by discussing the findings from the tour activities of various industries.

2. Related research

Previous research has developed several tour-based freight transport models. We identify three main classes of tour modelling approaches in the literature. The first class of freight tour modelling is based on the entropy maximization theory in which the most likely tour flow is estimated based on freight trip generation, truck counts on the road network, and total transportation cost or time. In this method, the Lagrangian multiplier associated with zones and transport cost and time can be used to interpret the aggregated tour flow pattern on a network (Sánchez-Díaz,Holguín-Veras,and Ban, 2015). Gonzalez-Calderon and Holguín-Veras (2017) extended this method by adding a heuristic to the model to identify the number and locations of the traffic counts that should be used in the estimation. They also tested the sensitivity of the model to the locations of the traffic counts under different scenarios. Although the entropy maximization method provides valuable insights into general time-dependent tour flow patterns on an aggregate level and bypasses the need for expensive surveys, the method is not able to generate disaggregate tour sequences (Gonzalez-Calderon and Holguín-Veras, 2017).

The second class of tour models relies on the assumption that firms are rational profit maximizers whose behaviour can be predicted based on the theory of utility maximization. These approaches can focus on three units of analysis, either trips, as in (Hunt and Stefan, 2007), shipments as in (Thoen et al., 2020; Nuzzolo and Comi, 2014), or tours (Khan and Machemehl, 2017). Choice modelling is the basis for all these approaches helping to explain behavioural daily tour patterns.

Of this second class of tour models, trip-based models are often estimated based on disaggregate trip data collected through surveys or GPS. In these models, tours are constructed through incremental trip chaining in such a way that the next destination in a tour is estimated based on the conditional probability of the current stop (Hunt and Stefan, 2007). These types of models are relatively easy to implement and provide transport modellers with descriptive statistics and insights into freight transport systems. However, there are several issues with these types of models. Discrete choice methods are not subject to constraints (like time windows) and therefore hardly can capture spatial–temporal characteristics of tours (Heinitz and Liedtke, 2010). Additionally, trip chaining decisions are made once at the tactical level and hence incremental reconstruction of the tour at the operational level is not the context in which carriers plan tours in reality.

Shipment-based tour models, similar to trip-based models, stack a set of choice models to reconstruct tours. The difference is that these choice models consider whether or not assigning a shipment to a tour will maximize the utility of the planner based on a generalized cost. The most comprehensive shipment-based tour modelling is developed by Thoen et al. (2020). This model consists of two binary choice models. The first model is the choice of selecting a shipment to be added to a tour; the second model is the 'end of tour' choice model that makes sure if the tour should be ended due to duration constraints and or capacity limitations. For reconstructing tours, a nearest neighbour search algorithm is used to find the next stop. The shipment-based architecture allows the inclusion of several logistical constraints, such as shipment size, in tour formation. In practice, tours are often the result of an optimization process where optimal time and sequence of trips are treated interdependently. The shipment-based architecture, although promising, does not capture scheduling decisions and also does not take the combinatorial challenge of deciding the visiting order of locations into account.

In order to deal with the issues mentioned regarding the second class of tour-based freight transport modelling, some researchers have proposed using a family of vehicle routing problems (VRP) and simulation in order to directly study the tours. These VRP formalisms model the pickup and delivery behaviour of carriers within a multi-agent micro-simulation framework (Donnelly, 2009; Van

Heerden and Joubert ,2014; Wisetjindawat et al., 2012; Sakai et al., 2020; Siripirote et al., 2020). Although such normative VRP models can perfectly capture space–time constraints, their outcome could deviate from observed tours due to heterogeneity in the tour planning preferences of planners (Canoy and Guns, 2019) or other differences between the stylized model and reality. In some sectors, executed tours deviate from the planned tours because of the tacit knowledge of truck drivers about the receivers' conditions and/or externalities that are not recognizable to the planners or are not easy to put in the objective function of VRP models (Mandi et al., 2021).

Only a few studies have touched upon this problem, introducing a new family of VRP which can be called *preference-based routing*. In this class of VRPs, the objective is to minimize the perceived costs or maximize the utility of planners and/or drivers (Mandi et al., 2021). From a methodological perspective, these studies can be divided into two groups. The first group of studies uses a matrix of transition probabilities between the stops instead of a matrix of distance or travel time in conventional VRPs. Canoy and Guns (2019) propose a maximum likelihood routing problem that maximizes the joint transition probabilities while planning tours. They use Markov chain models to estimate the transition probabilities. Mandi et al. (2021) propose a similar approach but adopt neural networks to estimate transition probabilities. Using neural networks allows for adding contextual variables to the prediction of transition probabilities. Both these studies show that this maximum likelihood routing with Markov transition probabilities can learn tours from a set of historical solutions with the reasonable route and arc differences. A problem with these approaches is that they require a large history of complete tour solutions to learn from. Such information may be available in practice for individual firms, where such tools have already been implemented for firms to help planners plan tours. However, generalizable data is hardly accessible for scientific purposes and particularly so for freight simulation. This is mainly because individual firms are not willing to disclose their activities exhaustively due to their customers' privacy.

The second group of studies on preference-based routing in shipment-based tour models employ an inverse optimization approach to calibrate a family of multi-objective VRP models. You, Chow, and Ritchie (2016) use the method of successive averages to estimate the weights of a weighted sum of multiple objectives from a set of observed tour diaries. The calibrated VRP can be used in any micro-simulation framework to simulate the activity of the freight carriers. Similar to the first group, parameter estimation of these methods mostly requires fully observed truck movement patterns from GPS trajectories (Siripirote,Sumalee,and Ho, 2020). However, tour data, if available, is often privately-owned and only partially available to traffic agencies and policy makers, if at all. As opposed to the maximum likelihood routing approach, the parameter estimation of the inverse optimization approach can be adapted in such a way that it can learn from partially observed tour information just for cases where full information is not available. Until now, the literature lacks an efficient parameter estimation method for these cases.

In summary, the mentioned studies underline the importance of data-driven routing in freight transport modelling. However, existing methods are mostly based on either the data collected from a limited and expensive survey, or on fully observed truck movements with trip sequences and schedules obtained from GPS. In either case, existing work has neglected the role of scheduling of tours. Here, a method to estimate disaggregate dynamic tour planning spatially- and temporally based on partially-observed tour data will be a requirement. The regular inverse optimization technique as proposed in You, Chow, and Ritchie (2016) cannot be adopted when the tour information is not fully available. To the best of our knowledge, the development of a method to calibrate a time-dependent VRP model based on shipment flows with partially observed tour data has not yet been explored and is therefore the subject of the current paper. To address this research gap, we propose an efficient parameter estimation method to build a data-driven time-dependent VRP model based on partially observed tour information.

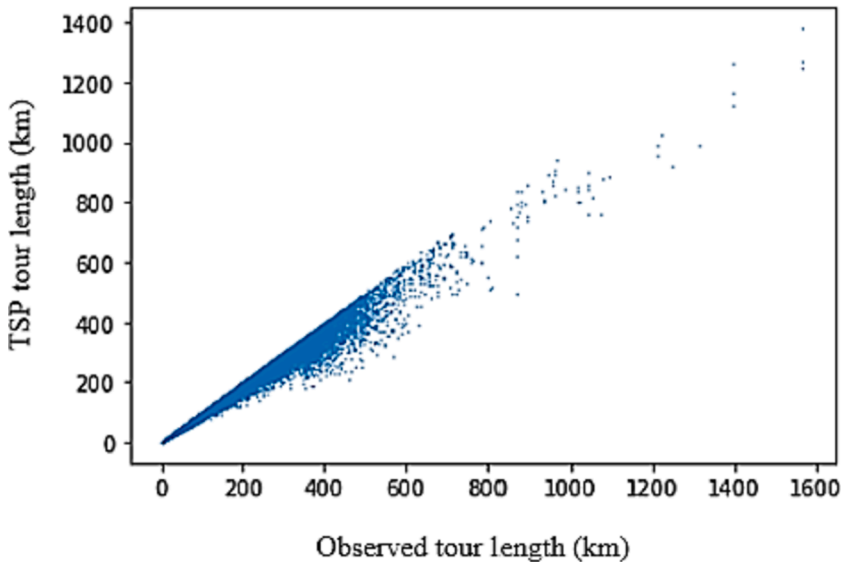

Fig. 1. Deviation of executed tour length from the TSP routes.

This study advances the state-of-the-art the literature in three ways:

1. It designs a new freight routing and scheduling model for transport modelling that can accurately link freight tour activities to the truck flows per time of day in between firms or traffic analysis zones.
2. It utilizes partially available tour information, to make sure that the model is as close as possible to reality, by generating the tour planning patterns of various carriers.
3. It proposes a new and efficient parameter estimation method that combines the strength of both optimization and machine learning approaches, to provide statistical insights with a certain confidence about tour patterns learned from data.

3. Method for data-driven routing and scheduling

Recall that our objective is to infer from historical data the collective pattern of pre-trip routing and scheduling process of freight transport. In the tour planning process, planners often use optimization software that can produce pickup and delivery plans that minimize tour length and travel time. Such optimization problems are known as traveling salesman problems (TSP), for one vehicle, or vehicle routing problems for multiple vehicles.

To explore the potential deviation of planners from TSP tour sequences, we use a large tour database collected by the Central Bureau for Statistics (CBS) in the Netherlands. After data pre-processing, a set of 16,171 tour activities of the 720 vehicles from the largest carriers are selected for this study. We provide more information about this dataset in Section 4.

We use the state-of-the-art open-source Google OR-Tools solver (https://developers.google.com/optimization/) to solve a benchmark TSP for all the tours in the dataset. Fig. 1 compares the length of the observed tours with the TSP routes. From this experiment, we identify that tours are on average 18.5 % longer than TSP solutions. This reveals that planners or drivers, in many cases, do not prefer the tours with the lowest travel times or costs: the data finds that the humans are apparently not rational with respect to the TSP optimization model (Hofstede et al., 2019).

In some cases, planners might use other objectives like fuel consumption and/or the cost of carbon emissions in the objective of the optimization problem. Further, our estimate of times and costs may differ from those used by the firms. However, such objectives and other real-world constraints or their perception of time and costs are not observable to us.

To deal with this problem and capture the preferences of planners and drivers in planning tours, we propose a data-driven routing and scheduling problem where the preferences of the planners can be identified through a set of features associated with zones (or firms). To achieve this, we include a linear weighted sum of these features in the cost of routing and scheduling decisions. We use the Benders decomposition method (Castellucci,Darvish,and Coelho, 2021) with an adaptive large neighbourhood search (ALNS) algorithm to find the optimal sequence of visited zones/firms, and utilize a Bayesian optimization algorithm to iteratively update the parameters associated with the zonal features. We next give details of the mathematical model and the decomposition-based approach.

3.1. Tour modelling formulation with learnable parameters

We formulate each tour in our data set as a time-dependent TSP with pickup and delivery and capacity constraints. In our data, tours are often starting from the base location of carriers or firms to visit consumers of goods, transshipment terminals, manufacturers, retailers, or distribution centres. For each firm in the area of study, this TSP is defined on a graph $G=\{V, A\}$ where V is the set of all nodes and A is the arcs' set. Vertices of this network consist of carriers' depot {0}, pick-up points P and delivery points D. Each link is associated with c_{ij}^m which is a function of features related to the pickup or delivery points and the link between them in time interval m. The higher the c_{ij}^m the lower the inclination of planners to visit j after i in the sequence. The order of visiting locations, however, depends on a combination of the planner's tendency towards all other i and j, and hence can be identified through combinatorial optimization. Each pick-up point i is also associated with demand quantity $q_i > 0$ and its corresponding delivery point $q_{i+m} = -q_i$. For better reference to the indices, parameters and decision variables, we list them as follows:

V	A set of pickup and delivery locations
i,j	Index of pick up or delivery locations $i,j \in$ V
M	A set of time intervals
cap	Maximum capacity of vehicle k
L_i	A load of a vehicle visiting location i
x_{ij}^m	If a vehicle travels from i to j in interval m, then $x_{ij}^m = 1$, otherwise $x_{ij}^m = 0$
β	Vectors of parameters to be estimated for the cost of the links
c_{ij}^m	Cost of the link between i and j in the interval $m \in$ M (with learnable parameters)
t_{ij}^m	Travel time between i and j in the interval m $\in$ M
a_i	Arrival time to location i
s_i	Loading/ unloading duration in location i
q_i	Demand quantity in location i
v_i	Index of the first node in the route that visits node $i \in$ V

We adopt a compact form of a standard two-index formulation of the time-dependent pickup and delivery problem from Castellucci, Darvish, and Coelho (2021) and Furtado et al. (2017). In this paper we are not proposing a new formulation; rather, we propose a

method that can be used to make any class of TSP learnable from a set of real-world historical tours. We choose the following formulation because it is aligned with the structure of tours that is observable to us. One can adapt this formulation based on the need and observability of real-world objectives and constraints.

$$min\sum_{m\in M}\sum_{i\in V}\sum_{j\in V}c_{ij}^{m}x_{ij}^{m} \tag{1}$$

subject to:

$$\sum_{m\in M}\sum_{i\in V}x_{ij}^{m}=1;\forall j\in P\cup D \tag{2}$$

$$\sum_{m\in M}\sum_{j\in V}x_{ij}^{m}=1;\forall i\in P\cup D \tag{3}$$

$$\sum_{m\in M}\sum_{i\in V}x_{ij}^{m}=\sum_{m\in M}\sum_{i\in V}x_{ji}^{m};\forall j\in V,m\in M \tag{4}$$

$$\sum_{m\in M}\sum_{j\in V}x_{0j}^{m}=1;\forall j\in V,m\in M \tag{5}$$

$$a_j\geq a_i+s_i+t_{ij}^{m}-M\left(1-x_{ij}^{m}\right);\forall i\in V,j\in V,m\in M \tag{6}$$

$$L_j\geq L_i+q_i-M\left(1-x_{ij}^{m}\right);\forall i\in V,j\in V,m\in M \tag{7}$$

$$Max\{0,q_i\}\leq L_i\leq min\{cap,cap+q_i\} \tag{8}$$

$$a_{i+n}\geq a_i+s_i+t_{i,i+n}^{m};\forall i\in P,m\in M \tag{9}$$

$$v_{i+n}=v_i;\forall i\in P \tag{10}$$

$$v_j\geq j.x_{0j}^{m};\forall j\in P\cup D,m\in M \tag{11}$$

$$v_j\geq j.x_{0j}^{m}-n\left(x_{0j}^{m}-1\right);\forall j\in P\cup D,m\in M \tag{12}$$

$$v_j\geq v_i-n\left(x_{ij}^{m}-1\right);\forall i,j\in P\cup D,m\in M \tag{13}$$

$$v_j\leq v_i-n\left(1-x_{ij}^{m}\right);\forall i,j\in P\cup D,m\in M \tag{14}$$

$$a_i+s_i-t(m-1)X_{ij}^{m}\geq 0;\forall i\in V,j\in V0\},m\in M \tag{15}$$

$$a_i+s_i\leq t.m+\left(1-X_{ij}^{m}\right)(T-t.m);\forall i,j\in V,m\in M \tag{16}$$

$$x_{ij}^{m}=\{0,1\} \tag{17}$$

The objective function presented in Equation (1) minimizes the perceived cost of routing and scheduling for planners/drivers. Equations (2) and (3) make sure that each node is visited exactly once. Equation (4) ensures the degree of balance for all nodes. Equation (5) represents the constraint for the truck being departed from the depot. The load and time consistency is guaranteed with constraints presented by Equations (6) and (7). The inequality constraint proposed in Equation (8) ensures the capacity limitation of vehicles. Equation (9) guarantees the precedence of pickup and delivery locations, as pick-up points always must be before delivery points. Equations (10) to (14) reflect the pairing relation which makes sure that all related pick-up and delivery locations are on the same tour. We defined the lower and upper bound on departure times in equations (15) and (16) ensuring that departure time from each location is linked with its associated interval *m*. Finally, Equation (17) defines the domain of the decision variable. The above formulas represent in general how tours can be planned by firms based on a dataset. We did not include the time windows constraint. As mentioned previously, not all the tour information is available for transport modellers. Real-world constraints like time windows, if available, can add to the accuracy of the model.

To capture the spatial and temporal preferences of tour planners from a set of partially observed tours information, we define c_{ij} (presented in Equation (1)) to represent the cost of visiting *j* after *i* from the planner's perspective. We link this cost to a set of features that relate to the characteristics of visiting zones or firms and also the link between them:

$$c_{ij}^{m} = \sum_{f}^{F} \beta_{f}^{m} \chi_{fij}^{m} + b^{m} \tag{18}$$

In Equation (18), β_f^m represents the importance of feature f and χ_{fij}^m represents the value of each feature in time interval m and b^m is a bias term that captures the aggregate costs that planners may consider but is not observable to us. We introduce the features that can contribute to the planner's preferences. This implies that we estimate different parameters for similar features at different tour departure times. These features are as follows:

- χ_{1ij}^m: travel time cost per hour between firm/zone i and j when the trip starts at time interval m. This can be derived by multiplying a fixed cost per hour by the travel time between i and j.
- χ_{2ij}^m: transport cost per kilometre between firm/zone i and j when the trip starts at time interval m. This can be derived by multiplying a fixed cost per kilometre by the distance between i and j.
- χ_{3ij}^m: vehicle load while traversing from i to j and the trip starts at time m.
- χ_{4ij}^m: ratio of the number of commodities for being picked up or delivered in j over i and the trip starts at time m.
- χ_{5ij}^m: ratio of the weight of commodities for being picked up or delivered in j over i when the trip starts at time m.
- χ_{6ij}^m: transport cost per kilometre between depot and j over transport cost per kilometre between the depot and i when the trip starts at time m.
- χ_{7ij}^m: travel time cost per hour between depot and j over travel time cost per hour between the depot and i when the trip starts at time m.
- χ_{8ij}^m: transport cost per kilometre between j and depot over transport cost per kilometre between the depot and i when the trip starts at time m.
- χ_{9ij}^m: travel time cost per hour between j and depot over travel time cost per hour between i and depot when the trip starts at time m.
- χ_{10ij}^m: 1 if j is a producer/consumer of goods and i is the distribution centre and 0 otherwise.
- χ_{11ij}^m: 1 if both i and j are distribution centres and 0 otherwise.
- χ_{12ij}^m: 1 if the commodity between any i and j is of type NSTR 0 (agricultural products) and 0 otherwise.
- χ_{13ij}^m: 1 if the commodity between any i and j is of type NSTR 1 (food products) and 0 otherwise.
- χ_{14ij}^m: 1 if the commodity between any i and j is of type NSTR 6 (construction materials) and 0 otherwise.
- χ_{15ij}^m: 1 if the commodity between any i and j is of type NSTR 7 (fertilizers) and 0 otherwise.
- χ_{16ij}^m: 1 if the commodity between any i and j is of type NSTR 8 (chemical) and 0 otherwise.
- χ_{17ij}^m: 1 if the commodity between any i and j is of type NSTR 9 (machinery and others) and 0 otherwise.

Together, these features help to identify the preferences of planners in the routing and scheduling of different commodities between different logistics firms. Besides feature parameters, there might be some other necessary but unknown parameters in the model that can be estimated from the observed data. In this study, for example, the loading and unloading duration of commodities, s_i, is not reported. We therefore assume that the loading and unloading of commodities is a random process that its service time has a cumulative exponential probability distribution with an unknown average service time μ. Hence the probability that the service time in location i, will be less than or equal to t is:

$$P(s_i \leq t) = 1 - e^{-\mu t} \tag{19}$$

The unknown average service time μ has to be estimated along with the other parameters.

3.2. Parameter estimation of the tour model

Having described the mathematical model, we now describe how to estimate its parameters. In general, we begin with initializing the parameter of the models with some initial values. Then we use the ALNS search algorithm and Benders decomposition to generate route sequences and departure time schedules for the tours. We iteratively update the parameters of the model minimizing the deviation between generated and observed tour characteristics. This process requires repeatedly solving the TSP model after each update in parameters. Since TSP models are computationally expensive, we cannot easily adopt a complete optimization algorithm to search for a better set of parameters. Alternatives are local search algorithms, perturbation stochastic approximation methods, or gradient descent algorithms – all of which may stop in a local optimum and hence report inefficient results.

To deal with this problem, we adopt from artificial intelligence the model-based Bayesian optimization technique that is successfully applied in hyper-parameter optimization of black-box functions like deep learning neural network models. Model-based algorithms are a class of global optimization methods that can be used to find the minimum of functions that are expensive to evaluate, do not have derivatives available, and can only be measured under noisy environments or simulations. Examples of such algorithms are LineBO (Kirschner et al., 2019), DONE (Bliek,Verwer,and de Weerdt, 2021) and SMAC (Hutter et al., 2011). These algorithms are mainly based on two principal goals: to explore the search space sufficiently for a global minimum, and to obtain a good solution in as few function evaluations as possible. These methods use a surrogate surface fitted to the parameter space in order to reduce the number of function evaluations while searching for the global optimum. In this paper, we propose the same approach to estimate the parameters of the tour model.

In Fig. 2 we illustrate the parameter estimation algorithm which can be explained in five steps as follows:

Step 1: Prepare instances of observed tours with all the matrices of features related to visiting locations in different time intervals.

Step 2: Generate n initial parameters $\Theta=\{\theta_1,\theta_2,\ldots,\theta_n\}$ for the tour model using Latin hypercube sampling. Each θ_n includes all β parameters, bias terms b, and the μ parameter. We started with 15 initial samples of all parameters.

Step 3: For each parameter setting, we solve the tour model for all instances prepared in step 1. We employ ALNS using simulated annealing to provide approximate solutions for carriers having a large number of customers in the planning horizon. Since our model is time-dependent and the tour solving algorithm has to find actual departure time from each node considering traffic conditions, exact solutions for such problems require longer computation time. To help the algorithm search both time and space dimensions faster, we adopt a logic-based Benders decomposition method as presented in (Castellucci,Darvish,and Coelho, 2021). The main idea of this technique relies on heuristically decomposing and solving smaller problems that lead to faster and more efficient ways of solving the main optimisation problem. We thus decompose the time-dependent tour model presented in Equations 1–17 into master and sub-problems.

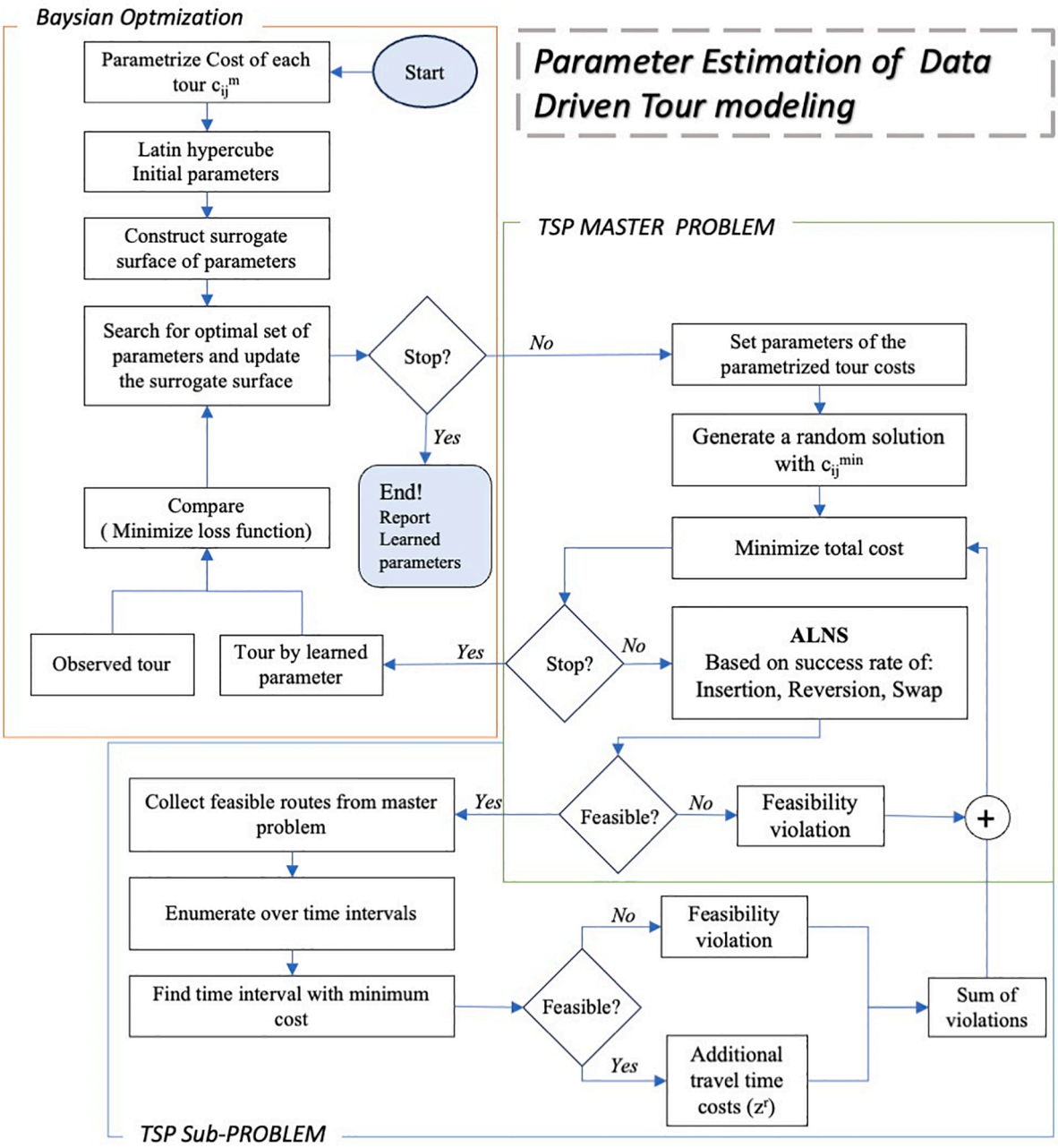

Fig. 2. Procedure of model construction and parameter generation.

Master problem:

In the master problem, we calculate the routes with constant travel time throughout the day. Essentially, the master problem serves as a relaxation of the time-dependent TSP in which the problem is reformulated by enumerating all possible routes for a selected tour (vehicle). To this end, we define an auxiliary variable c_{ij}^{min} as the minimum travel time along arc i and j.

$$c_{ij}^{min} = \min_{m \in M} c_{ij}^{m} \tag{20}$$

Equation (20) only concerns the time-dependent attributes in c_{ij}^{m} that is χ_{1ij}^{m}, χ_{7ij}^{m}, and χ_{9ij}^{m}. Let R be a set of feasible routes for a tour. These routes are defined as a sequence of nodes departing from a depot. For any feasible route, we assume z^r is the time-dependent tour cost of the route r. We define variable $x_{ij} \in \{0, 1\}$ to indicate if a link between i and j is traversed. Additionally, we introduce $c_r \geq 0$ which is the additional time-dependent tour costs. This term serves to penalize the master problem's objective function in response to the feedback received from the sub-problem. The master problem is the reformulation of the proposed TSP model in Equations (1)–(14). To avoid formula repetition, we will only present the objective function of the master problem. The constraints 2–14 remain unchanged, except for the omission of index m.

$$Min \sum_{i \in V} \sum_{j \in V} c_{ij}^{min} x_{ij} + c_r \tag{21}$$

Note that the cost of a tour is separated into minimum cost c_{ij}^{min} and additional cost c_r. The feasible solution of the master problem always remains feasible after applying the time-dependent travel time due to time-independent constraints in the master problem. To properly consider the additional time-dependent costs of traversing each link for a feasible solution of the master problem, we generate the violation term using equation (22).

$$c_r \geq z^r - \sum_{(i,j) \in R} c_{ij}^{min} \tag{22}$$

Sub problem:

Upon discovering a feasible solution to the master problem, we use a sub-problem to compute the complete set of arrival and departure times using time-dependent travel time matrices. Subsequently, we integrate the feasibility violations determined in the sub-problem back into the master problem's objective function. To formulate the sub-problem, consider having a feasible solution to the master problem that consists of a set of nodes V_r. The objective function of the sub-problem (Equation (23)) is to determine the optimal interval m in which the total time-dependent cost of the master problem's feasible tour reaches its minimum value. For any $x_{ij} = 1$ in this tour, we ensure the same arc is used in the sub-problem using constraint (24). The timing relationship between the consecutive stops and the boundaries of the departure time from each visit is given by Equations (25) and (26). Finally, constraints (27) and (28) enforce the domain of variables.

$$Min z_r \tag{23}$$

Subject to:

$$\sum_{m \in M} x_{ij}^{m} = 1; i, j \in V_r \tag{24}$$

$$a_i + s_i + t(m-1)x_{ij}^{m} \geq 0; \forall i, j \in V_r, m \in M \tag{25}$$

$$a_i + s_i \leq t.m + \left(1 - x_{ij}^{m}\right)(T - t.m); \forall i, j \in V_r, m \in M \tag{26}$$

$$x_{ij}^{m} \in \{0, 1\} \tag{27}$$

$$z_r \geq 0 \tag{28}$$

To solve the sub-problem, we enumerate over all time intervals seeking the optimal solution that designates the departure time of the vehicle that minimizes its time-dependent tour cost after visiting all nodes in V_r following the links in the feasible solution provided by the master problem. For more details on how this method performs in solving time-dependent TSP as compared to other methods, we refer our readers to Castellucci, Darvish, and Coelho (2021).

Adaptive large neighbourhood search.

We utilize an adaptive large neighbourhood search (ALNS) with simulated annealing (SA) as the solver for solving the above time-dependent pickup and delivery TSP. The idea of this algorithm was first proposed by (Ropke and Pisinger, 2006). We use the same principles to search through solution space for an approximation of the optimum solution. Given that SA is inherently a discrete algorithm, the solutions of a TSP can be coded as a permutation of the nodes' indices. SA involves an iterative procedure where, in each iteration, a new solution is generated and evaluated based on the preceding solution. Many heuristics have been devised to generate solutions for SA through permutation operations. ALNS enables the solver to select multiple heuristics, drawing from their respective success rates in prior iterations. This can take place by assigning a weight to each heuristic, reflecting its historical efficacy. These weights are updated periodically during the runtime of the SA algorithm. The selection of each heuristic from the set heuristics $H =$

$\{h_i \vee i = 1, 2, \cdots k\}$ is based on the probabilities calculated as follows:

$$P(h_i) = \frac{w_{h_i}}{\sum_{j=1}^{k} w_{h_j}} \tag{29}$$

In ALNS, we update the weights of the heuristics after P_u iterations (see Equation (30).

$$w_{h_i} = \begin{cases} (1-\varnothing)w_{h_i} + \varnothing \frac{s_{h_i}}{n_{h_i}}, if n_{h_i} > 0 \\ (1-\varnothing)w_{h_i}, if n_{h_i} = 0 \end{cases} \tag{30}$$

where s_{h_i} is the number of time that the heuristic h has been successful in improving the solution over n_{h_i} that is the number of time that this heuristic has been used during the P_u iterations.

A factor $0 \leq \varnothing \leq 1$ is defined in equation (30) to control the impact of the recent success of a heuristic on its weight. It is worth mentioning that the success of a heuristic depends on three conditions, namely (1) if the new solution ($x^{'}$) found by the heuristic is the best one so far; or (2) if the new solution ($x^{'}$) improves the current solution (x); or (3) if the solution does not improve the current solution but is accepted. The third condition is due to the SA that accepts such solutions based on Boltzmann probability $e^{-(f(x^{'})-f(x))/T}$ that depends on the cost difference of the new and current solutions controlled by a so-called temperature factor $T \geq 0$. For more information on the ALNS with SA please refer to Ropke and Pisinger (2006). For the implementation of the tour model, we use open-source Python packages, in particular the state-of-the-art Google OR-Tools.

Step 4: Given the observed tour sequences and generating tours based on initial zone feature weights Θ, we can evaluate the tour model by comparing observed and generated tours. Previous research introduced arc and route deviation metrics to measure similarity or dissimilarity between two sequences. However, we cannot use these metrics as our observed tours do not reveal such information. In our case where the tour information is partially available, we compare available tour information using the following loss function:

$$l(\theta) = \frac{1}{N}\sum_{n=1}^{N}\left(r_n - \hat{r}_n\right)^2 \tag{31}$$

where N denotes the number of observed tours, r_n is the observed tour information (i.e., tour length, tour duration, tour start time and end time), and $\hat{r}_n$ is the tour information generated by the tour model.

Step 5: In this step, we iteratively update the initial parameters θ minimizing the loss function defined in equation (31). We adopt the following steps of the Bayesian optimization algorithm to find parameters θ.

<u>1- Surrogate construction:</u> Fit a surrogate surface f to the evaluated Θ samples using a Multivariate Gaussian process with mean μ and covariance Σ (see Equations 32–34).

$$f(\Theta)N(\mu, \Sigma) \tag{32}$$

$$\Sigma = K(\theta, \theta^{'}) = \begin{bmatrix} K(\theta_1, \theta_1) & \cdots & K(\theta_1, \theta_n) \\ \vdots & \ddots & \vdots \\ K(\theta_n, \theta_1) & \cdots & K(\theta_n, \theta_n) \end{bmatrix} \tag{33}$$

$$K(\theta, \theta^{'}) = exp\left(\frac{(\theta - \theta^{'})^2}{2d^2}\right) \tag{34}$$

The kernel function k(θ,θ$^{'}$) determines how smooth the surrogate surface could be and the length parameter d scales the correlation between parameters θ and $\theta^{'}$.

<u>2- Search for minimum:</u> Find the minimum of the interpolated surface for the sample of parameters. This gives us the alternative point with the lowest mean sample values among the already evaluated points.

$$\theta_{min} = \underset{\theta \in \Theta}{argmin} f(\theta) \tag{35}$$

<u>3- Next evaluation point:</u> the next evaluation point θ* in parameter space is where the expected improvement (EI) measure is maximum. This EI measurement is introduced in Equations (36) and (37) to balance exploration and exploitation.

$$EI(\theta^*) = E[\max(\bar{f}(\theta_{min}) - f^*(\theta), 0) \tag{36}$$

$$f^*(\theta)N(\bar{f}(x), \sigma(\theta)) \tag{37}$$

Where $\bar{f}(\theta)$ is the mean stochastic prediction at point θ and $\sigma^2(\theta)$ is the estimate of prediction error.

<u>4- Update the surrogate surface:</u> With the new parameter θ*, we update the covariance matrix and consequently the surrogate surface (see Equations 38–42). The process iteratively continues until the stopping criteria apply.

$$J(\theta^*)|J(\theta_1), \cdots, J(\theta_n)N(\mu(\theta^*), \overline{K}(\theta^*)) \tag{38}$$

$$\mu(\theta^*) = k_{\theta^*}(\theta^*)^T k_{\theta^*\theta^*}^{-1} g \tag{39}$$

$$g = (J(\theta_1), \cdots, J(\theta_n))^T \tag{40}$$

$$K_{\theta^*}(\theta^*)^T = (k(\theta^*, \theta_1), \cdots, k(\theta^*, \theta_n)) \tag{41}$$

$$\overline{K}(\theta^*) = k(\theta^*, \theta_1) + k_{\theta^*}(\theta^*)^T k_{\theta^*\theta^*}^{-1} k_{\theta^*}(\theta^*) \tag{42}$$

4. Empirical validation

Having described the proposed tour model, we now apply it to a real-world set of shipment data and imperfect tour information in the Netherlands. We first describe the databases used and the data preprocessing that was applied, before estimating the parameters of the tour model.

4.1. Firms and carriers' tour databases

For the development of this tour model, we use carriers' tour data collected by the Statistics Netherlands (CBS). This data is available only with the permission of CBS. The data includes over 2.7 million records of shipments across the country for the years 2013–2015. In this data collection, the largest third-party carriers are legally obliged to fill in a form that collects one week of their activities. In most cases, the data has been exported directly from the firm's transportation management system automatically through an XML data structure.

The data includes the geographic locations of loading and unloading of shipments, their commodity type, and shipment size. It also provides information about the capacity of the vehicles and is enriched with other data sources (e.g., firm establishment data) to express locations of important activities including distribution centres, transshipment terminals, and producer/consumer of goods. The data provides aggregate information about each tour as well. This information is about the total tour costs and times, shipments that belong to one tour, the first and last visited locations, and the start and end time and location of the tour. Although the data includes very detailed information about the shipments, it does not provide us with the order of trips in a tour. The tour database only includes the total tour duration. However, the intermediate travel time between each visited customer is not provided. To impute these

Table 1
Descriptive statistics of tour data.

Tours characteristics	Number of tours	Percentage of tours
Number of stops		
3–5	6734	44 %
6–10	7378	48 %
>10	1183	8 %
Tour length (km)		
0–50	649	4 %
50–100	1576	10 %
100–200	4020	26 %
200–400	7170	47 %
>400	1880	12 %
Tour Duration (hour)		
0–3	1651	11 %
4–6	2925	19 %
7–8	3171	21 %
9–10	3095	20 %
>10	4453	29 %
Commodity types		
0	1530	10 %
1	4457	29 %
6	183	1 %
7	41	0 %
8	98	1 %
9	8986	59 %
Logistics activities	Number of shipments	Percentage of shipments
Intermediate Loading locations		
DC	17,083	16 %
TT	1861	2 %
P/C	88,454	82 %
Intermediate Loading locations		
DC	5128	5 %
TT	2277	2 %
P/C	99,993	93 %

intermediate travel times and distances, we use the calibrated national Dutch regional traffic model. This model helped us to create skim matrixes of the morning, midday and afternoon travel times between each traffic analysis zones in the Netherlands. For more information about the available data see de Bok and Tavasszy (2018) and Thoen et al. (2020).

After preprocessing the data, we selected a subset of 16,171 tour activities (107,398 shipments) of the 720 vehicles from the largest carriers. This subset does not include direct tours. We excluded direct tours from our sample because such tours do not reveal the routing preferences of the planners. The proposed method after training can still generate direct tours if the capacity of the vehicle is reached or the VRP cannot bundle a shipment with others due to the costs of other tours. Table 1 reports the descriptive statistics of the selected tours.

4.2. Results of the model estimation

We hold out 20 % of the tour data for validation, and train the tour model based on the rest of the tour data using the Bayesian optimization algorithm of Section 3.2. We draw the index of the training set from a random uniform distribution to avoid any bias in the selected data. We set the number of initial parameter samples in Bayesian optimization to 15. Parameters associated with features are bounded between −10 and 10. The parameter of the service time generator is assumed to be an integer between 20 and 120 min. We specify 0.05 as the stopping criteria meaning that the model does not improve the solutions by at least 5 % in 5 consecutive iterations, the algorithm stops updating the parameters. We activate this criterion after 50 iterations. We also set the maximum number of iterations to 100 based on initial trials. All the features are normalized between 0 and 1 using a min–max scaler. For tours with 10 or greater stops, the TSP solver automatically switches between exact and heuristic algorithms for solving the TSP models.

In our experiment, the number of time intervals that the departure time of tours may fall in is $m = 3$. This indicate morning peak, afternoon peak, and off-peak periods. For feature selection, we started with features 1 to 6 from the list of features that are introduced in Section 3.1. These features can explain planners' preferences intuitively. Then we consecutively add/remove other features to the model only if they improve the final loss function value. Table 2 shows the estimated parameters of the tour model.

In Table 2, each estimate is associated with the standard deviation of estimates around the surrogate surface (see Equation (26)). The columns 'std' shows how reliable estimates are in terms of the standard deviation. In general, we can see that the estimated parameters are less reliable for the afternoon as compared to morning and off-peak periods. This is probably because there are relatively lower trip legs in tours that fall in the afternoon period and therefore the algorithm has fewer observations that it can learn from.

The result shows that both transport costs and travel time costs between i and j (χ^m_{1ij} and χ^m_{1ij}), increase the cost c_{ij} and therefore lower the preference of planners to keep j after i is the sequence of trips. The transport cost has a higher impact on the planner preferences as compared to travel time cost. The higher weight for travel time in the afternoon indicates that carriers value afternoon travel times more than morning or off-peak periods. This means that carriers are more likely to avoid including a link in a tour that has a larger travel time during the afternoon.

Similarly, transport costs in the afternoon are more important than the morning and off-peak periods. This indicates that trips that are scheduled in the afternoon within a tour are more likely to be shorter than trips in the morning or noon. However, the standard deviations of estimates for a time in the afternoon, although acceptable (we accept estimates with std < 5e-2), are slightly higher than the morning and afternoon estimates. This is because of the lower number of tour samples in this category for parameter estimation.

The load of vehicles χ^m_{3ij} has significant negative impacts on the planners' cost c_{ij} and therefore, if the vehicle has a higher load on

Table 2
Estimated parameters of the model.

No	features	Morning peak		Afternoon peak		Off-peak	
		Estimates	std	Estimates	std	Estimates	std
1	χ^m_{1ij}	1.12	1.8e-4	2.1	1.52e-05	1.064	2.18e-4
2	χ^m_{2ij}	3.52	1.5e-3	4.8	0.014	4.15	2.87e-5
3	χ^m_{3ij}	−2.39	1.42e-5	−3.89	1.7e-4	−1.66	1.3e-5
4	χ^m_{4ij}	−0.251	1.4e-8	−0.312	3.29e-4	−0.107	1.3e-2
5	χ^m_{5ij}	−0.097	1.2e-3	–	–	−0.042	4.38e-2
6	χ^m_{6ij}	−2.379	1.4e-8	−3.289	1.3e-2	−1.263	3.29e-9
7	χ^m_{7ij}	−1.246	1.5e-6	−1.956	1.7e-4	−0.954	1.3e-5
8	χ^m_{8ij}	−2.074	2.1e-2	3.173	3.7e-3	1.063	1.73e-5
9	χ^m_{9ij}	−1.331	1.87e-3	1.212	2.7e-3	1.023	3.8e-5
10	χ^m_{10ij}	−0.471	6.3e-7	−0.423	3.9e-3	−0.529	2.2e-8
11	χ^m_{11ij}	0.516	2.5e-3	–	–	0.294	4,1e-3
12	χ^m_{12ij}	−0.170	1.6e-4	−0.289	1.7e-4	0.363	2.4e-6
13	χ^m_{13ij}	0.055	2.3e-4	−0.181	1.2e-3	−0.541	3.1e-5
14	χ^m_{14ij}	−0.134	1.7e-2	–	–	–	–
15	χ^m_{15ij}	−0.041	1.3e-2	–	–	–	–
16	χ^m_{16ij}	−0.314	1.83e-5	−0.155	3.96e-4	0.251	3.1e-3
17	χ^m_{17ij}	0.268	6.5e-4	0.164	2.8e-3	−0.512	3.8e-6
Service time estimate: 52 min							

the link i-*j* then the cost of this link is less from the planners' perspective. The ratio of the number of commodities in *j* over *i* (χ^m_{4ij}) also has a significant negative impact on the planners' cost which implies that if there is a larger number of commodities in *j* as compared to *i* for pickup or delivery, the perceived cost of planners traversing the link i-j in a tour will become less. The same explanation emerges for the ratio of weights χ^m_{4ij}, however, the magnitude of the impact is relatively lower.

Parameters associated with features χ^m_{6ij} and χ^m_{7ij} indicate how travel time and cost of the depot to *i* and *j*, can influence the planners' inclination to keep link i-*j* in a tour. As the figures suggest, these features have negative impacts on the total cost (positive impact on planners' preferences). This intuitively shows that if *j* is further away from depot as compared to *i* or the travel time between depot and *j* is higher than the travel time between the depot and *i*, the planners would like to keep *j* after *i* in the tour.

Parameters associated with features χ^m_{8ij} and χ^m_{9ij} indicates how travel time and cost of *i* and *j* to the depot, can influence the planners' preferences to keep *j* after *i* in a tour. The sign of the parameters shows that, during off-peak and afternoon peak periods, these features have positive impacts on the total cost of traversing from *i* to *j*, from the planners' perspective. This means that if the depot is further away from *j* as compared to *i* or the travel time between *j* and depot is higher than the travel time between *i* and depot, the planners do not prefer to keep j after i in the tour. For the morning peak period, however, the sign is counter-intuitively negative. One possible reason for this could be that planners do not value trips back to the depot in the morning as during this period trips in tours are more outgoing flows. The lower std for these parameters also shows relatively low confidence, which means probably there are fewer observations to support this parameter.

The negative sign of features χ^m_{10ij} indicates that the trips between distribution centres and producer–consumer have lower perceived costs from the planners' point of view. On contrary, if both i and j are distribution centres χ^m_{11ij}, then planners prefer not to put *j* after *i* in the trip sequence of a tour. Given our dataset, the model cannot estimate a significant parameter for the afternoon due to a lack of observations.

Features χ^m_{12ij} to χ^m_{17ij} relates the planners' preferences for routing and scheduling of five different commodity types. Unlike other features that interdependently explain the preferences of planners towards routing and scheduling of trips inside a tour, these features mostly influence only the scheduling decisions. This is because the commodity type category does not change within a tour in 95 % of our sample dataset. The positive sign of these commodity types within a particular time interval means that planners have less interest to schedule trips carrying the commodity type within that interval. For example, the weights of χ^m_{12ij} is positive for the off-peak period and negative for morning and afternoon periods. This implies that planners prefer to schedule trips with agricultural products during off-peak periods.

The model also suggests generating on average 52 min of service time using exponential distribution for (un)loading locations. The service time generator has a minimum bound of 20 min. Drivers' break times are estimated collectively with the service time. Our dataset does not report on when and where they stop to rest but reports on the actual tour duration; hence the model can only capture these extra times along with the service time of each zone.

4.3. Model performance

To study the performance of the model, we report on the mean absolute percentage error (MAPE) and the coefficient of determination (R-squared) regarding estimated and observed tours' length and duration. To scope our model's merits, we also compare the fit of the proposed data-driven TSP model and the benchmark TSP model individually against real-world tours. With this, we aimed to demonstrate the real-world applicability and advantages of our proposed framework. we believe that this comparison serves as a baseline for highlighting the innovative nature of our model.

The high R^2 measurement confirms the overall good performance of the data-driven TSP model to generate real-world tour data. Table 3 shows that the purposed model can predict tour length with 5.1 % error on the train and 9.3 % error on the test datasets. The benchmark TSP model with 18.6 % and 17.9 % errors on train and test datasets respectively shows higher deviation from real-world tour data. As opposed to the benchmark TSP, the data-driven model also shows relatively high similarity to reality regarding tour duration. However, the lower performance of the model in predicting tour duration as compared to the tours' length is because of the stochasticity in the service time of the visiting locations.

By contrasting the two approaches, we emphasize the benefits of leveraging real-world data to better understand and predict routes, which is a significant departure from using the traditional TSP in freight transport modelling and simulation. Please note that by comparing these approaches against real-world data, we are not generally claiming any superiority of our model over the traditional

Table 3
Performance of the model.

KPI	Model	Train data		Test data	
		MAPE(%)	R^2	MAPE(%)	R^2
Tour length (km)	Data-driven TSP	5.1	0.92	9.3	0.9
	Benchmark TSP	18.6	0.81	17.9	0.82
Tour duration (min)	Data-driven TSP	12.3	0.88	12.7	0.86
	Benchmark TSP	23.2	0.68	21.8	0.71

TSP method rather we challenge the usability of the traditional non-data-driven TSP models in the simulation of tour planning in the freight transport system.

We acknowledge that a direct quantitative comparison between the two may not provide a comprehensive assessment of our model's merits. However, we firmly believe that demonstrating the improvements achieved by our data-driven approach in terms of route understanding, prediction accuracy, and the ability to simulate real-world freight transport scenarios contributes to advancing the field. This comparison is intended to underscore the practical implications and potential applications of our model within the context of freight transport logistics.

4.4. Model evaluation

The tour distance and number of stop distributions are two important aggregate features that should be captured well by tour models. To evaluate the performance of our model to reproduce aggregate tour characteristics for the purpose of freight modelling, we simulate the model on the available shipment data. This time, however, we do not specify the link between shipments and tours. In other words, only shipment origins and destinations (OD) is available to the firms. In this experiment, firms do the routing and scheduling for multiple vehicles to meet all the shipments on the planning day. Note that shipment OD is available at 6-digit postcode level. Multiple firms may be active in a 6-digit postcode. We assume, however, that all the shipments that belong to a postcode belong to one single hypothetical firm. This assumption can be refined using advanced population and shipment synthesis models, if available. We also add 11,265 direct tours to our sample for model evaluation.

For each firm, we use benchmark TSP and the data-driven TSP model to route and schedule tours. We group tour length into different clusters and compared the estimated tour distance distribution with the observations in Fig. 3. The results show that the

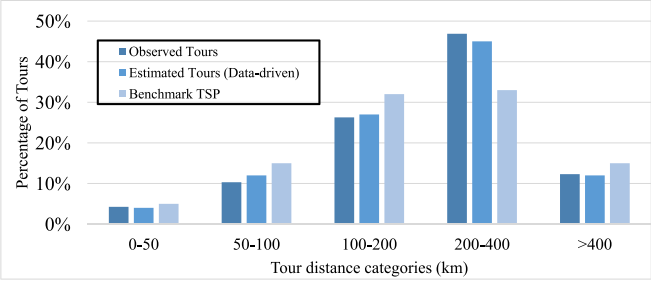

Fig. 3. Estimated vs. observed tour distance distribution.

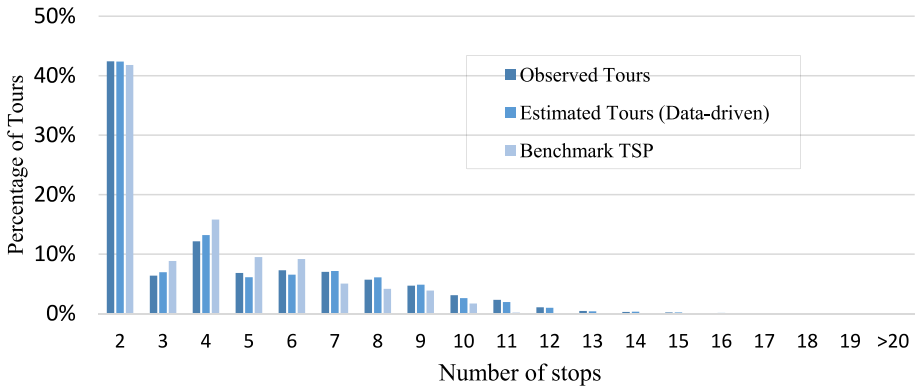

Fig. 4. Estimated vs. observed tour number of stops distribution.

a) Confusion Matrix for departure time

		Estimated		
		Off-peak	Morning	Afternoon
Observed	Off-peak	87%	11%	3%
	Morning	15%	82%	4%
	Afternoon	4%	0%	96%

b) Confusion Matrix for end time

		Estimated		
		Off-peak	Morning	Afternoon
Observed	Off-peak	82%	11%	9%
	Morning	7%	81%	12%
	Afternoon	2%	3%	95%

Fig. 5. Observed and estimated departure time and end time intervals.

model is able to approximate this feature reasonably well. Unlike the benchmark TSP model, the proposed data-driven TSP model can re-produce tour length distribution.

Similarly, in Fig. 4 we can see that our data-driven model captures the general pattern in the frequency of tours with different numbers of stops. In general, it can be concluded that tours with a smaller number of stops (3 and 4) are slightly overestimated.

We also compare the coincidence rate of estimated and observed tours' departure and end time in Fig. 5. After estimating the departure time and the end time of tours using the proposed model, we count the number of tours that their departures and end times fall into the same time intervals as the observe tours. The figure shows that our model can correctly predict departure (up to 96 %) and end (up to 95 %) times of the tours.

4.5. Sensitivity analysis

In addition to the model validation, we test the sensitivity of our model towards the increase or decrease of travel time. From the data, we identify the most frequent link (i-j) in all tours. We increase and then decrease travel time on this link for both morning and afternoon by 10 %, 20 % and 50 %. The results in Fig. 6 show that the model is sensitive to changes in travel time and assigns fewer trips to this link if we increase the travel time.

We also experiment with the sensitivity of the model to changes in travel costs per kilometre. Changes in transport costs can happen due to, for instance, tolling systems, increases in labour costs, fuel costs, or carriers' fixed costs. Fig. 7 indicates that the proposed model responds logically to changes in travel costs.

A comparison between the sensitivity of the tours to cost and time shows that carriers are more sensitive to transport costs than travel time. The sensitivity of this model towards time and cost – along with its ability to model tours spatially and temporally – makes this modelling approach a strong and more realistic tool, as compared to conventional models, for policy implications and impact

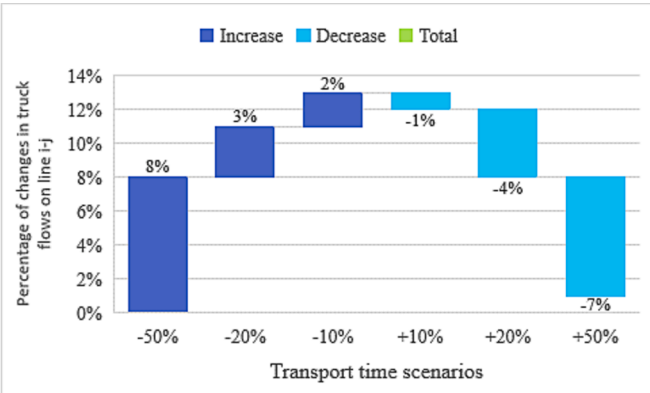

Fig. 6. Sensitivity analysis for travel time increase/decrease on the most frequent link i-j.

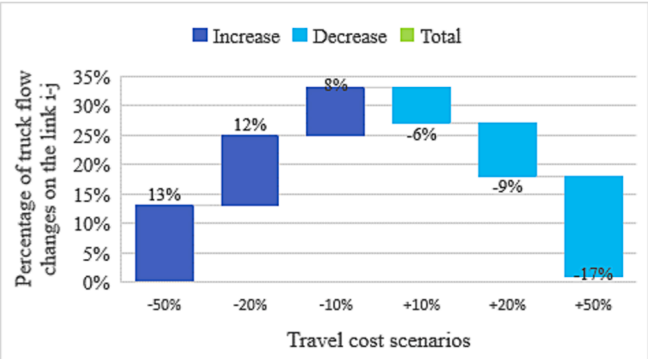

Fig. 7. Sensitivity analysis of the model on transport cost variations.

assessment.

5. Discussion and conclusion

This research investigated the possibility of developing a data-driven tour model to capture the activity of freight carriers from disaggregate and partially available tour data. We parametrized a time-dependent TSP model to capture the aggregate spatial and temporal correlations between trip legs in a tour. We used Bayesian optimization to estimate the parameters of this model. The results from the model show that the estimated model not only fits the individual tour flow data with high accuracy but also it captures aggregate tour features like tour distance distributions and the number of stops. The model is sensitive to time and cost changes and can be used as a tool in traffic simulations for integrated traffic and freight transport modelling.

The proposed model has practical capabilities, and also limitations. Estimating a general tour model which can represent the collective tour activities of large carriers can become an undetermined problem, especially when a limited imperfect amount of disaggregated data is available. Researchers have been dealing with these problems in transportation science for decades (for example in OD estimation problem (Krishnakumari et al., 2020). The performance and accuracy of models in these cases depend largely on the amount of prior information used to help the model converge to the real behaviour of road users. Our proposed method shows that additional, although imperfect, tour information like tour duration, cost, start time, and end time can provide a robust and accurate estimation of a single tour model with meaningful parameters that can reproduce the preferences of carriers in planning tours.

It is important to note that while certain features may resemble human planners' behaviour, the underlying reasons for these decisions remain largely unknown. For instance, a common observation might be that tours generated by our model tend to avoid selecting links in the afternoon with high travel times. Although this pattern suggests the preference of human planners for efficient routes, it does not necessarily imply that this is the exact reason behind the observed behaviour. Understanding the exact motivations behind the decisions captured by our tour generation model remains a challenge. Consequently, efforts to interpret and explain the decision-making processes of these systems remain an ongoing research area.

However, the lack of a causal relationship between the parameters and the tour outcomes does not hinder our model capability for its domain application. This is because our model has a primary objective of generating tours like human planners and the meaningful parameters incorporated within it offer valuable insights for conducting what-if analyses. These parameters effectively capture the likelihood and preferences of human planners under specific conditions, allowing us to understand the potential outcomes of varying factors within the tour generation process.

By utilizing the meaningful parameters of our model, we can simulate and assess the potential impacts of different variables on tour outcomes. For example, we can analyze the effect of adjusting pickup times and durations, modifying constraints, or introducing new destinations. This capability empowers us to explore various "what-if" scenarios, gaining valuable insights into how changes in the model's parameters might influence the generated tours and consequently freight and logistics system.

Although the model validation shows that the proposed model can accurately predict the routing and scheduling of carriers, the high accuracy of the model on the test data is not a surprise as the test data also come from the same sample distribution (but unseen data) on which we build the model. The ideal way of evaluating the model performance would be to use secondary data such as truck counts on road networks to validate and calibrate our model. This, however, requires integrating this model into a more comprehensive freight transport model.

To the best of our knowledge, there is no comparable study that we can compare our results with across all dimensions. de Jong et al. (2016) estimated models to explain the time-period preferences of receivers in road freight transport. They have applied this model to assess the impact of these preferences on the time-period choices of carriers. De Jong et al. show that road freight carriers are relatively insensitive to travel time changes. In contrast, carriers are more sensitive to peak hours if they realize an increase in transport costs. de Jong et al. (2016) only consider the time dimension in freight transport without considering the impact of routing. Interestingly, our model, which treats both routing and scheduling of carriers, shows the same behaviour in tour planning. The spatial–temporal connection that trips have in our model makes this method a good candidate for spatial and temporal analysis. One can use this model, for instance, to study the impact of off-peak policies or to investigate the impact of an increase or decrease in transport cost (e.g., fuel cost, distance/shipment size-related costs) or cost of travel time (e.g., toll system), on the activities of carriers. These abilities make this an interesting tool for policy analysis with capabilities that go beyond current models.

Seen as a tool for analysts, our work can also be used to understand the impact of freight activities on the traffic system. The activities resulting from this model can be easily translated into a time-dependent truck OD table, which can be coupled with a traffic simulation model to investigate the inter-connection between freight and traffic. In this study, the model used the travel time matrices from a traffic simulation model to estimate tours. In return, the time-dependent truck OD that resulted from this model can be also used as input to traffic simulation models. This could provide interesting insights into the inter-relation between logistics and traffic systems. Additional future research direction is the use of auxiliary data such as truck counts on the road networks that can be used to enhance the accuracy of the VRP calibration and time-dependent truck OD matrix estimation.

CRediT authorship contribution statement

Ali Nadi: Conceptualization, Investigation, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Neil Yorke-Smith:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Maaike Snelder:** Conceptualization, Project administration, Writing – review & editing, Supervision. **J.W.C. Van Lint:** Supervision. **Lóránt Tavasszy:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the anonymous reviewers for helping us improving the quality of the paper. This work was supported by the Dutch Research Council (NWO), TKI Dinalog, Commit2Data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat, and TNO under project 'ToGRIP-Grip on Freight Trips' with grant number 628.009.001, and also partially supported by TAILOR, a project funded by EU Horizon 2020 research and innovation program under grant number 952215. The authors thank CBS and Rijkswaterstaat for providing shipment data and travel time matrices, respectively.

References

Bliek, L., Verwer, S., de Weerdt, M., 2021. Black-box combinatorial optimization using models with integer-valued minima. Ann. Math. Artif. Intell. 89, 639–653.

Canoy, Rocsildes, and Tias Guns. 2019. "Vehicle routing by learning from historical solutions." In *Principles and Practice of Constraint Programming: 25th International Conference, CP 2019, Stamford, CT, USA, September 30–October 4, 2019, Proceedings 25*, 54-70. Springer.

Castellucci, P.B., Darvish, M., Coelho, L.C., 2021. A Benders Decomposition Algorithm for the Time-dependent Vehicle Routing Problem. Université de Montreal, Bureau de Montreal.

de Bok, M., Tavasszy, L., 2018. An empirical agent-based simulation system for urban goods transport (MASS-GT). Procedia Comput. Sci. 130, 126–133.

de Bok, M., Tavasszy, L., Thoen, S., 2022. Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in The Netherlands. Transp. Policy 124, 119–127.

de Jong, G., M.K., Ruijs, K., van Houwe, P., Borremans, D., 2016. A time-period choice model for road freight transport in Flanders based on stated preference data. Transportation Research Part E: Logistics and Transportation Review 86, 20–31.

Donnelly, R., 2009. A hybrid microsimulation model of urban freight transport demand. University of Melbourne. Unpublished PhD thesis,.

Furtado, Maria Gabriela S., Pedro Munari, Reinaldo Morabito, 2017. Pickup and delivery problem with time windows: a new compact two-index formulation. Oper. Res. Lett., 45: 334-41.

Gonzalez-Calderon, C.A., Holguín-Veras, J., 2019. Entropy-based freight tour synthesis and the role of traffic count sampling. Transportation Research Part e: Logistics and Transportation Review 121, 63–83.

Heerden, V., Quintin, and Johan W Joubert., 2014. Generating intra and inter-provincial commercial vehicle activity chains. Procedia Soc. Behav. Sci. 125, 136–146.

Heinitz, F.M., Liedtke, G.T., 2010. Principles of Constraint-Consistent Activity-Based Transport Modeling. Comput. Aided Civ. Inf. Eng. 25, 101–115.

Hofstede, G.J., Jonker, C.M., Verwaart, T., Yorke-Smith, N., 2019. The lemon car game across cultures: evidence of relational rationality. Group Decis. Negot. 28, 849–877.

Holguin-Veras, J., Patil, G.R., 2007. Integrated origin–destination synthesis model for freight with commodity-based and empty trip models. Transp. Res. Rec. 2008, 60–66.

Hunt, J.D., Stefan, K.J., 2007. Tour-based microsimulation of urban commercial movements. Transp. Res. B Methodol. 41, 981–1013.

Hutter, Frank, Holger H Hoos, and Kevin Leyton-Brown. 2011. "Sequential model-based optimization for general algorithm configuration." In *Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy, January 17-21, 2011. Selected Papers 5*, 507-23. Springer.

Khan, M., Machemehl, R., 2017. Analyzing tour chaining patterns of urban commercial vehicles. Transp. Res. A Policy Pract. 102, 84–97.

Kirschner, J., Mutny, M., Hiller, N., Ischebeck, R., Krause, A., 2019. Adaptive and safe Bayesian optimization in high dimensions via one-dimensional subspaces. In: International Conference on Machine Learning. PMLR, pp. 3429–3438.

Krishnakumari, P., Van Lint, H., Djukic, T., Cats, O., 2020. A data driven method for OD matrix estimation. Transportation Research Part c: Emerging Technologies 113, 38–56.

Mandi, Jayanta, Rocsildes Canoy, Víctor Bucarey, and Tias Guns. 2021. "Data driven vrp: A neural network model to learn hidden preferences for vrp." In *CP*, 1-42. arXiv:2108.04578.

Mommens, K., Lebeau, P., Verlinde, S., Van Lier, T., Macharis, C., 2018. Evaluating the impact of off-hour deliveries: An application of the TRansport Agent-BAsed model. Transp. Res. Part D: Transp. Environ. 62, 102–111.

Nuzzolo, A., Comi, A., 2014. Urban freight demand forecasting: a mixed quantity/delivery/vehicle-based model. Transportation Research Part e: Logistics and Transportation Review 65, 84–98.

Ropke, S., Pisinger, D., 2006. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Transp. Sci. 40, 455–472.

Sakai, T., Alho, A.R., Bhavathrathan, B.K., Chiara, G.D., Gopalakrishnan, R., Jing, P., Hyodo, T., Cheah, L., Ben-Akiva, M., 2020. SimMobility Freight: An agent-based urban freight simulator for evaluating logistics solutions. Transportation Research Part e: Logistics and Transportation Review 141, 102017.

Schröder, S., Liedtke, G.T., 2017. Towards an integrated multi-agent urban transport model of passenger and freight. Res. Transp. Econ. 64, 3–12.

Siripirote, T., Sumalee, A., Ho, H.W., 2020. Statistical estimation of freight activity analytics from Global Positioning System data of trucks. Transportation Research Part e: Logistics and Transportation Review 140, 101986.

Sánchez-Díaz, I., Holguín-Veras, J., Ban, X.J., 2015. A time-dependent freight tour synthesis model. Transp. Res. B Methodol. 78, 144–168.

Tavasszy, Lorant, and Gerard De Jong. 2013. *Modelling freight transport* (Elsevier).

Thoen, S., Tavasszy, L., de Bok, M., Correia, G., van Duin, R., 2020. Descriptive modeling of freight tour formation: A shipment-based approach. Transportation Research Part e: Logistics and Transportation Review 140, 101989.

Wisetjindawat, W., Yamamoto, K., Marchal, F., 2012. A commodity distribution model for a multi-agent freight system. Procedia Soc. Behav. Sci. 39, 534–542.

You, S.I., Chow, J.YJ., Ritchie, S.G., 2016. Inverse vehicle routing for activity-based urban freight forecast modeling and city logistics. Transportmetrica a: Transport Science 12, 650–673.