

## Ontwikkeling en schatting van een gedragsmatig model voor tourformatie

Toen, Sebastiaan; Tavasszy, Lorant; Homem de Almeida Correia, Goncalo; van Duin, Ron; de Bok, Michiel

**Publication date**

2018

**Document Version**

Accepted author manuscript

**Published in**

Bijdragen Vervoerslogistieke Werkdagen 2018

**Citation (APA)**

Toen, S., Tavasszy, L., Homem de Almeida Correia, G., van Duin, R., & de Bok, M. (2018). Ontwikkeling en schatting van een gedragsmatig model voor tourformatie. In M. Kraan, & S. Weijers (Eds.), *Bijdragen Vervoerslogistieke Werkdagen 2018* (2018 ed., pp. 103-117). (Bijdragen Vervoerslogistieke Werkdagen). University Press Zelzate.

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

## **ONTWIKKELING EN SCHATTING VAN EEN GEDRAGSMATIG MODEL VOOR TOURFORMATIE**

S. Thoen, TU Delft/Significance

L.A. Tavasszy, TU Delft

G. Homem de Almeida Correia, TU Delft

J.H.R. van Duin, TU Delft/Hogeschool Rotterdam

M.A. de Bok, TU Delft/Significance

## **Abstract**

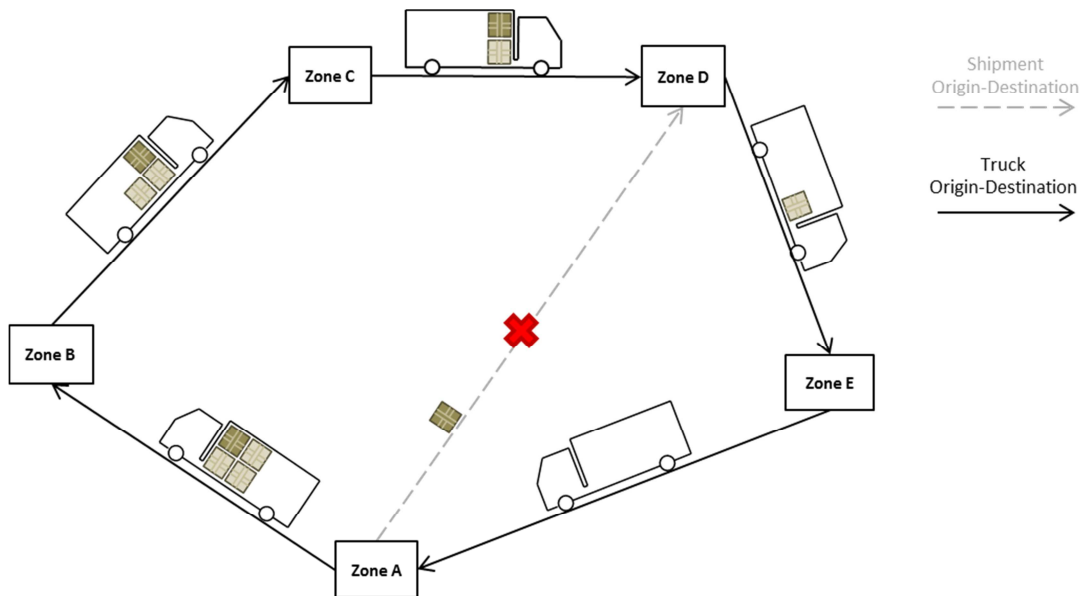
Tourformatie is kenmerkend aan goederenwegvervoer, niet zelden worden meerdere zendingen in een rit vervoerd. Toch wordt dit aspect vaak over het hoofd gezien in goederenvervoermodellen. Wanneer tourformatie wel wordt toegepast, dan worden vaak geen zendingen gemodelleerd, is het modelgedrag niet gebaseerd op empirische data, of wordt slechts een klein segment van goederenvervoer gemodelleerd. In deze paper beschrijven we een tourformatie algoritme voor goederenwegvervoer binnen Nederland. Dit algoritme wijst zendingen toe aan ritten. Twee probabilistische keuzestappen in dit algoritme zijn geschat op empirische data over Nederlandse beroepsvervoerders. Relevante strategieën en randvoorwaarden in tourformatie komen terug in dit algoritme. Daarnaast worden verschillen in tourformatie voor diverse locaties, goederensoorten, en voertuigtypen overwogen. Zo worden op goederenoverslagplaatsen meestal ritten met een enkele zending gevormd, terwijl vanaf distributiecentra vaker ritten met meerdere zendingen voorkomen. Toepassing van het algoritme laat zien dat geobserveerde ritpatronen uitstekend gereproduceerd worden.

### **1. Introductie**

Goederenvervoermodellen worden veelvuldig gebruikt om de hoeveelheid vrachtverkeer te voorspelen in verschillende toekomstscenario's en onder diverse beleidsmaatregelen. Een kenmerkend aspect van goederenwegvervoer is tourformatie, tijdens een rit worden vaak meerdere klanten bezocht om transportkosten te besparen (Hunt & Stefan, 2007; Sánchez-Díaz et al., 2015). Het modelleren van tourformatie leidt tot betere voorspellingen in goederenvervoermodellen, omdat er rekening wordt gehouden met het feit dat zendingen niet noodzakelijkerwijs direct van de laadlocatie naar de loslocatie worden vervoerd (Figuur 1) (Holguín-Veras et al., 2014; Sánchez-Díaz et al., 2015).

In de wetenschappelijke literatuur wordt in toenemende mate aandacht besteed aan het gedragsmatig modelleren van tourformatie. Twee stromingen zijn te onderscheiden: (1) wiskundige optimalisatiemodellen en (2) algoritmes met discrete keuze-elementen.

Vaak worden Vehicle Routing Problems opgelost om ritten te vormen die een set zendingen vervoert van laad- naar loslocatie (bijv. Boerkamps & van Binsbergen, 1999; Wisetjindawat et al., 2006; Anand et al., 2014). Deze wiskundige optimalisatiemodellen worden normaliter gebruikt om de optimale oplossing aan een transportplanner voor te schrijven, maar kunnen ook worden toegepast om het gedrag van een transportplanner te beschrijven. Er kleven twee nadelen vast aan deze manier van tourformatie modelering: (1) de rekentijd is erg hoog en neemt sterk toe naarmate meer zendingen moeten worden toegewezen, (2) kalibratie op basis van empirische data is uitermate complex en kan niet op een statistische manier plaatsvinden (bijv. You et al., 2016).



*Figuur 1. Wanneer meerdere zendingen in een rit vervoerd worden, dan hoeft een trip van de laadlocatie van een zending (Zone A) naar de loslocatie (Zone D) niet noodzakelijkerwijs gemaakt te worden.*

Het alternatief is het ontwikkelen van een algoritme met discrete keuze-elementen. Hunt & Stefan (2007) zijn de eersten die een dergelijke aanpak voor tourformatie voorstelden, welke zij toepasten in een simulatiemodel voor commercieel verkeer in Calgary, Canada. Eerst wordt voorspeld hoeveel ritten starten in iedere zone, waarna een voertuigtype wordt gekozen per rit. Deze rit wordt vervolgens incrementeel uitgebreid door telkens de keuze te maken tussen een trip naar een volgende zone of een trip terug naar de standplaats. Het sterke punt van deze manier van modelleren is dat gedrag in het model op statistische wijze kan worden gekalibreerd op empirische data. Een zwaktepunt is dat ritten direct worden gegenereerd, er worden dus geen zendingen gemodelleerd.

Om verschillende redenen is het belangrijk om zendingen te modelleren alvorens tourformatie plaatsvindt. Allereerst bepalen de zendingen van een vervoerder de mogelijkheden en beperkingen voor tourformatie. Tourformatie is in feite het toewijzen van zendingen aan voertuigen, waarbij beperkingen gerelateerd aan deze zendingen, zoals de voertuigcapaciteit, een grote rol kunnen spelen. Daarnaast kunnen we in een model met zendingen op expliciete manier meenemen hoe ritpatronen verschillen tussen goederengroepen (Holguín-Veras et al., 2014). Tot slot kunnen relevante en meer gedetailleerde scenario's en beleidsmaatregelen worden onderzocht wanneer zendingen gemodelleerd worden. Hierbij kan men denken aan de effecten van toenemende orderfrequenties en afnemende ordergroottes (Boerkamps & van Binsbergen, 1999).

In dit onderzoek is een tourformatie model ontwikkeld dat zowel statistisch gekalibreerd is op empirische data als expliciet gebaseerd is op zendingen. Verder vraagt dit model relatief weinig rekenkracht en is het model breed toepasbaar. Waar andere modellen slechts focussen op een klein

segment binnen goederenvervoer, zoals vervoer van voedsel en boeken (Boerkamps & van Binsbergen, 1999), maken onze data en methode geen dergelijke begrenzingsen.

In Paragraaf 2 wordt dit tourformatie algoritme uitgelegd. Paragraaf 3 beschrijft de data die we gebruiken voor schatting en validatie. Paragraaf 4 en 5 rapporteren respectievelijk de schattings- en validatieresultaten. Conclusies en aanbevelingen zijn geformuleerd in Paragraaf 6.

## **2. Methodologie**

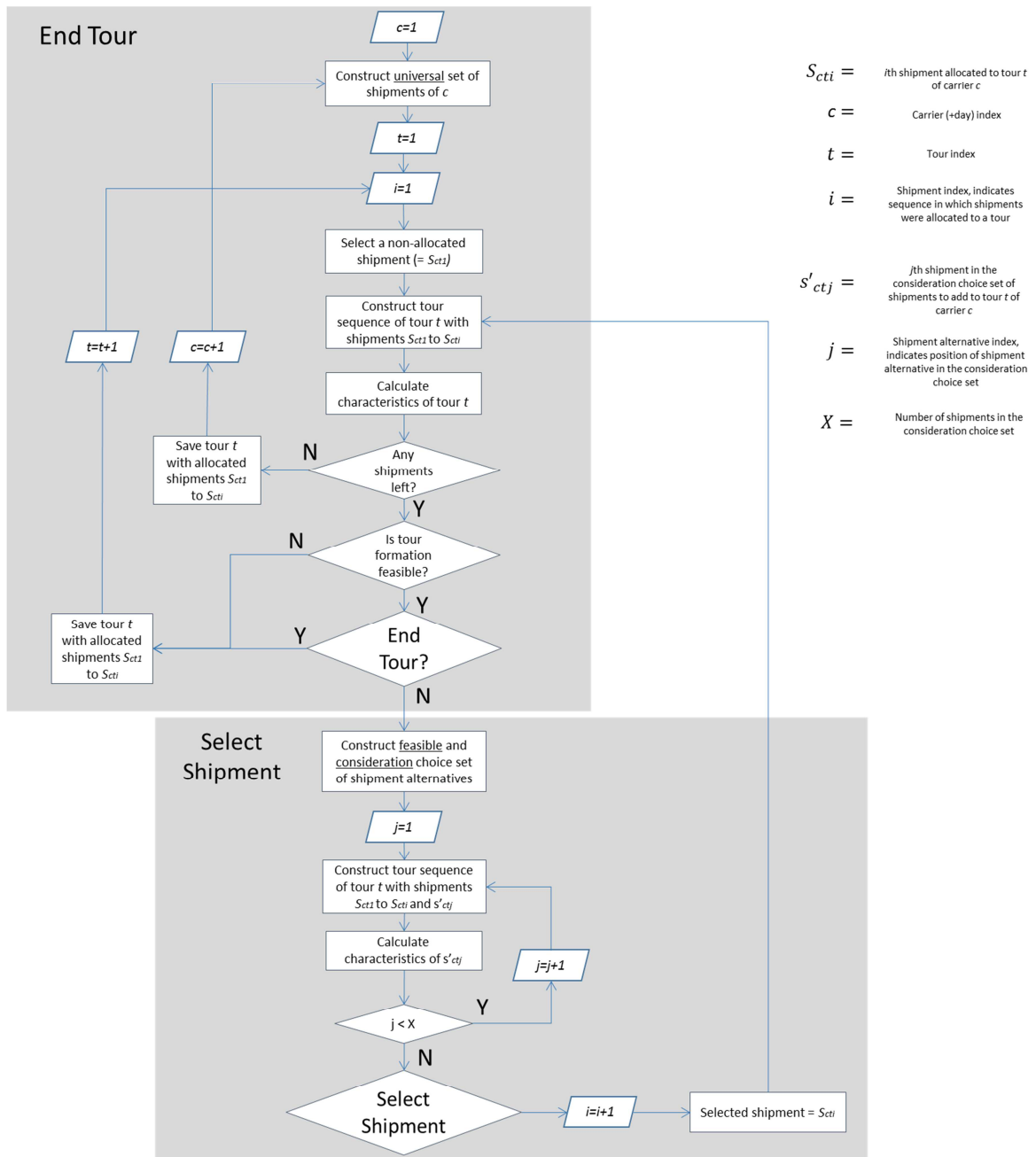
Het tourformatie model wijst zendingen toe aan ritten in een iteratief proces. Figuur 2.1 beschrijft het algoritme in een stroomschema. In dit algoritme zijn twee keuzestappen te onderscheiden: de zogeheten End Tour en de Select Shipment stap. De End Tour stap is een binair logistisch regressiemodel; de afhankelijke variabele is binair, waarbij de waarde 0 betekent dat een extra zending wordt toegevoegd aan de rit en de waarde 1 betekent dat de rit wordt beëindigd. De Select Shipment stap is een multinomiaal logistisch regressiemodel; als de rit niet wordt beëindigd, dan wordt in dit model een specifieke zending gekozen om aan de rit toe te voegen. Naast de End Tour en Select Shipment stap, is ook een toursequentie algoritme gespecificeerd om de volgorde van locaties in de rit te bepalen.

In Paragraaf 2.1 en 2.2 worden de respectievelijk het End Tour en het Select Shipment proces binnen het tourformatie algoritme beschreven.

### ***2.1 End Tour stap***

Het vormen van ritten start met het definiëren van de universele keuze set, de set zendingen die samen in een rit terecht kan komen. In ons algoritme zijn dit alle zendingen van dezelfde vervoerder die afgeleverd moeten worden op dezelfde dag. Vervolgens wordt willekeurig een zending gepakt om de constructie van een nieuwe rit in te luiden.

De toursequentie wordt daarna bepaald. Dit is de volgorde waarin de laad- en loslocaties van de zendingen in de rit worden bezocht. Op basis hiervan berekenen we de reistijd van de rit tot dan toe. De toursequentie wordt bepaald aan de hand van twee variaties op een nearest-neighbor-search. In een nearest-neighbor-search is de volgende locatie telkens de dichtstbijzijnde onbezochte locatie. In de eerste variatie bezoeken we eerst alle laadlocaties en daarna de loslocaties, in de tweede variatie worden laad- en loslocaties om de beurt bezocht. De variatie die leidt tot de kortste rit wordt gekozen.



Figuur 2. Een stroomschema van het tourformatie algoritme.

Vervolgens wordt gecheckt of de rit (zoals geconstrueerd tot dusver) nog voldoet aan de volgende vier randvoorwaarden:

- Het totaal vervoerde gewicht is niet hoger dan het laadvermogen
- De rit duurt niet langer dan negen uur, i.v.m. de Rust- en Rijttijdenwet
- Er wordt geen cement/beton vervoerd, voor deze zendingen zien we zelden tot nooit ritten met meerdere zendingen omdat deze zendingen vaak groot en tijdsgevoelig zijn (Khan & Machemehl, 2017)

- Er zijn nog zendingen over binnen een straal van  $a$  km

Als de rit niet aan deze randvoorwaarden voldoet, dan wordt de rit beëindigd en starten we een nieuwe rit. Als wel aan de randvoorwaarden is voldaan, dan wordt met het End Tour keuzemodel de kans bepaald dat de rit alsnog wordt beëindigd. De verklarende variabelen van dit End Tour keuzemodel worden gerapporteerd en toegelicht in Paragraaf 4.

## ***2.2 Select Shipment stap***

Als de rit niet wordt beëindigd, dan kiezen we welke zending wordt toegevoegd aan de rit. Eerst genereren we een keuzeset met zendingen die mogelijk kunnen worden toegevoegd. Dit zijn de zendingen van dezelfde vervoerder en dezelfde datum, die nog niet zijn toegewezen aan een rit. Daarnaast worden zendingen die bestaan uit cement, zendingen die tot een overschrijding van het laadvermogen leiden, en zendingen buiten een straal van  $a$  km verwijderd uit de keuzeset. Als er geen zendingen overblijven na deze selectieprocedure dan wordt de rit alsnog beëindigd. Een subset van  $y$  zendingen wordt vervolgens gesampled om het model efficiënt te houden qua rekentijd.

De gekozen zending wordt toegevoegd aan de rit, waarna opnieuw wordt overwogen om de rit te beëindigen. Dit proces herhaalt zich tot de rit daadwerkelijk wordt beëindigd, vervolgens tot alle zendingen van de vervoerder op deze dag zijn toegewezen aan een rit, en tot slot tot alle vervoerders en dagen zijn afgegaan. De tweede zending die we toevoegen aan een rit hoeft niet noodzakelijkerwijs de tweede zending te zijn die wordt geladen en gelost in de rit. Met het nearest-neighbor-search toursequentie algoritme bepalen we telkens opnieuw de volgorde waarin de laad- en loslocaties worden bezocht.

## **3. Data**

### ***3.1 Big data van het CBS***

Om de coëfficiënten van het End Tour en het Select Shipment keuzemodel te schatten, maken we gebruik van de XML microdata van het Centraal Bureau voor de Statistiek (CBS). Deze gigantische dataset bevat informatie over ongeveer 2.6 miljoen zendingen. Doordat het CBS vervoerders toestaat een XML-interface te installeren, kan data automatisch worden afgelezen uit de transportsystemen van deze vervoerders. Een gevolg is dat er een zelfselectie van grote beroepsvervoerders is in de data, omdat deze beroepsvervoerders vaker de geavanceerde transportsystemen hebben die nodig zijn om een XML-interface te gebruiken.

Veel relevante informatie over deze zendingen is gerapporteerd, zoals het bruto gewicht en het NSTR goederentype. Daarnaast is bekend welke zendingen in dezelfde rit zijn vervoerd met welk voertuigtype. Ook zijn de laad- en loslocaties ingevuld, soms met coördinaten, maar meestal op PC6 of PC4 niveau.

Verrijking van de XML data met andere databronnen biedt veel nieuwe inzichten en mogelijkheden. Data over buurten kan gekoppeld worden aan de laad- en loslocaties. Op basis van het Algemeen Bedrijven Register en Kerncijfers wijken en buurten 2015 kunnen we bijvoorbeeld bepalen of een rit een goederenoverslagterminal, een logistieke knoop (distributiecentrum), of een verstedelijkt gebied heeft bezocht. Matrices met reistijden en afstanden tussen buurten kunnen worden gebruikt om een logische volgorde van laad- en loslocaties af te leiden voor de zendingen in een rit. De reistijden en buurtkenmerken zijn alleen verkregen voor Nederlandse buurten. De scope van dit onderzoek is dan ook goederenwegvervoer binnen Nederland.

Een minpunt voor dit onderzoek is dat lege ritten niet gerapporteerd worden. Volgens de definitie van een rit in de data, start een rit op de locatie waar goederen in een leeg voertuig worden geladen, en eindigt een rit op de standsplaats of op de locatie waar het voertuig leeggeraakt. Als gevolg zijn er opvallend veel directe ritten (één laadlocatie, één loslocatie) en ritten met een korte afstand waar te nemen in de data (Tabel 1).

### ***3.2 Descriptieve analyses***

In Tabel 2 zien we dat het percentage directe ritten aanzienlijk verschilt tussen goederengroepen, voertuigtypen, en locatietypen. Ritten die cement vervoeren bestaan vrijwel nooit uit meerdere stops omdat de zendingen groot en erg tijdsgevoelig zijn (Khan & Machemehl, 2017). Wanneer agrarische en voedingsproducten worden vervoerd, wordt vaker een rit met meerdere stops waargenomen dan wanneer olie, benzine, metalen, of bouwmaterialen worden vervoerd. Verschillende kenmerken van deze goederen kunnen verklaren waarom zendingen vaker gecombineerd kunnen worden in een rit, zoals: de ordergroottes, het gemak van laden en lossen, en de ruimtelijke spreiding van verladings- en ontvangers.

Vrachtwagens met aanhangers worden voornamelijk gebruikt voor directe ritten. Om goederen uit de vrachtwagen te lossen, moet de aanhanger ontkoppeld worden. Dit vermoeilijkt een rit waarin op efficiënte wijze meerdere klanten worden bediend. De 'overig/speciaal' categorie bestaat voor een groot deel uit bestelbusjes. Omdat bestelbusjes een kleiner laadvermogen hebben, kunnen minder vaak meerdere zendingen worden vervoerd in een rit.

Op goederenoverslaglocaties worden veel directe ritten geobserveerd. Dit zijn waarschijnlijk grotere zendingen afkomstig van een producent (Friedrich et al., 2014), bijvoorbeeld containers in de haven van Rotterdam. Vanaf distributiecentra zien we vaker ritten met meerdere stops. Dit zijn kleinere zendingen die vervoerd worden naar plaatsen van consumptie (Friedrich et al., 2014), waardoor meerdere zendingen in een voertuig passen. Daarnaast hebben distributiecentra het voordeel een grote set zendingen beschikbaar te hebben met dezelfde laadlocatie, wat consolidatie faciliteert. Tot slot worden in stedelijke gebieden relatief vaak ritten met meerdere stops gemaakt. De concentratie



van laad- en loslocaties maakt het makkelijker om op efficiënte wijze meerdere klanten te bezoeken in een rit.

Ritkenmerk	Aantal ritten
<i>Aantal stops</i>	
1-2 (direct)	365,905 (92.44%)
3-5	18,538 (4.68%)
6-10	10,008 (2.53%)
>10	1,361 (0.34%)
<i>Ritafstand [km]</i>	
0-20	172,341 (43.54%)
20-40	82,995 (20.97%)
40-60	19,508 (4.93%)
60-80	14,976 (3.78%)
80-100	13,747 (3.47%)
100-120	14,383 (3.63%)
120-140	12,578 (3.18%)
140-160	11,749 (2.97%)
160-180	6,944 (1.75%)
180-200	6,748 (1.70%)
≥200	39,843 (10.07%)
<i>Cement</i>	179,468 (45.35%)
<i>NSTR goederensoort</i>	
0: vee en agrarische producten	9,541 (2.41%)
1: voeding, veevoer	20,617 (5.21%)
2-5: benzine, olie, metalen	746 (0.19%)
6: bouwmaterialen	45,279 (11.44%)
7: mest	457 (0.12%)
8: chemische producten	210,151 (53.09%)
9: machines & overig	109,021 (27.54%)
<i>Voertuigtype</i>	
Vrachtwagen	194,875 (49.37%)
Vrachtwagen + aanhanger	37,660 (9.54%)
Trekker-oplegger	160,049 (40.55%)
Overig/speciaal	2,134 (0.54%)
<i>Type locatie bezocht in rit</i>	
Overslag	102,679 (25.94%)
Distributiecentrum	176,249 (44.53%)
Verstedelijk gebied	146,098 (36.91%)

Ritkenmerk	Percentage directe ritten
<i>Gemiddelde</i>	92.44%
<i>Gemiddelde (zonder cement)</i>	86.18%
<i>Cement</i>	
<i>NSTR goederensoort</i>	100.00%
0: vee en agrarische producten	73.05%
1: voeding, veevoer	64.11%
2-5: benzine, olie, metalen	96.92%
6: bouwmaterialen	97.67%
7: mest	77.02%
8: chemische producten	95.25%
9: machines & overig	84.14%
<i>Voertuigtype</i>	
Vrachtwagen	72.92%
Vrachtwagen + aanhanger	96.44%
Trekker-oplegger	85.26%
Overig/speciaal	97.29%
<i>Type locatie bezocht in de rit</i>	
Overslag (als laadlocatie)	96.35%
Overslag (als loslocatie)	96.01%
Distributiecentrum (als laadlocatie)	68.45%
Distributiecentrum (als loslocatie)	70.35%
Verstedelijk gebied	60.97%

#### 4. Schattingsresultaten

Vijftig procent van de data is gebruikt voor het schatten van het End Tour en het Select Shipment keuzemodel. Er is een hoog percentage directe ritten in de data, veel van de ritten wordt na de eerste zending dus al beëindigd (Tabel 1). Daarom is het End Tour keuzemodel apart geschat voor de eerste zending en voor latere zendingen. Tabel 2 geeft de schattingsresultaten van het End Tour (1<sup>e</sup> zending) keuzemodel weer, een positieve coëfficiënt vergroot de kans dat een rit wordt beëindigd op

het moment dat deze rit slechts één zending heeft toegewezen gekregen. In Tabel 3 zijn de schattingsresultaten van het End Tour (latere zendingen) keuzemodel gerapporteerd, een positieve coëfficiënt vergroot de kans dat de rit wordt beëindigd nadat deze reeds twee of meer zendingen heeft toegewezen gekregen. De cellen in Tabel 3 geven de Beta parameter en de standaardfout weer.

#### ***4.1 Schattingsresultaten End Tour keuzemodel***

In het eerste End Tour model is de kans dat de rit wordt beëindigd groter voor zendingen met een kortere reistijd. Voor zulke korte zendingen hebben vervoerders een voorkeur voor directe ritten; directe ritten zijn makkelijker te plannen en te volgen, en voor korte zendingen zijn kostenbesparingen d.m.v. consolidatie lager dan voor lange afstand zendingen. De kans dat de rit wordt beëindigd neemt ook toe naarmate een groter deel van het laadvermogen wordt gebruikt; als strategie voor kostenbesparing maken vervoerders liever ritten met een hoog laadvermogen. Tot slot vinden we de effecten voor locatietypen, voertuigtypen, en goederensoorten uit Tabel 2 terug in deze schattingen.

Het tweede End Tour model heeft overwegend dezelfde effecten als het eerste End Tour model. Echter, de coëfficiënt voor reistijd is positief in plaats van negatief. Als de rit meerdere zendingen heeft, dan neemt de kans toe dat de rit wordt beëindigd naarmate de reistijd langer is omdat de kans toeneemt dat de werkdag van de vrachtwagenchauffeur volgepland is met deze rit. Daarnaast zijn er twee additionele variabelen in het tweede End Tour model. De rit wordt eerder beëindigd als er geen zendingen meer te vinden zijn met een laad- en/of loslocatie nabij de locaties die reeds worden bezocht in de rit. En als het aantal stops in de rit hoger is, dan neemt de kans dat de rit wordt beëindigd af; de rit is toch al complex met meerdere stops, het toevoegen van een extra zending is dan minder onaantrekkelijk.

Tabel 3. Schattingsresultaten End Tour (1<sup>e</sup> zending)

keuzemodel.	
End Tour (1 <sup>e</sup> zending)	
R <sup>2</sup> Nagelkerke	0.442
-2 LL	47315
Percentage correct	84.8
N	75255
<i>constante</i>	1.684 (0.029)
$\sqrt{\text{reistijd rit [h]}}$	-1.698 (0.037)
<i>(gewicht/laadvermogen)<sup>2</sup></i>	5.471 (0.102)
<i>overslagterminal in rit</i>	1.588 (0.037)
<i>distributiecentrum in rit (laad)</i>	-0.578 (0.026)
<i>distributiecentrum in rit (los)</i>	-0.475 (0.026)
<i>verstedelijkt gebied in rit</i>	-0.461 (0.038)
<i>voertuigtype [vrachtwagen]</i>	-1.295 (0.039)
<i>[vrachtwagen + aanhanger]</i>	1.850 (0.049)
<i>[trekker-oplegger]</i>	-
<i>[anders/speciaal]</i>	-
<i>NSTR goederensoort [0: vee en agrarische producten]</i>	-0.736 (0.047)
<i>[1: voeding, veevoer]</i>	-0.659 (0.032)
<i>[2-5: benzine, olie, metalen]</i>	1.495 (0.337)
<i>[6: bouwmaterialen]</i>	1.452 (0.058)
<i>[7: mest]</i>	0.713 (0.253)
<i>[8: chemische producten]</i>	0.583 (0.045)
<i>[9: machines &amp; overig]</i>	-

Tabel 4. Schattingsresultaten End Tour (latere zendingen)

keuzemodel.	
End Tour (latere zendingen)	
R <sup>2</sup> Nagelkerke	0.292
-2 LL	37894
Percentage correct	81.8
N	44618
<i>constante</i>	-2.526 (0.062)
<i>reistijd rit [h]</i>	0.386 (0.014)
<i>gewicht/laadvermogen</i>	3.286 (0.057)
<i>afstand dichtstbijzijnde zending [km]</i>	0.009 (0.001)
<i>ln(aantal stops)</i>	-0.911 (0.042)
<i>overslagterminal in rit</i>	0.526 (0.047)
<i>distributiecentrum in rit (laad)</i>	-0.191 (0.036)
<i>distributiecentrum in rit (los)</i>	0.094 (0.036)
<i>verstedelijkt gebied in rit</i>	-0.145 (0.032)
<i>voertuigtype [vrachtwagen]</i>	-1.968 (0.061)
<i>[vrachtwagen + aanhanger]</i>	-0.954 (0.088)
<i>[trekker-oplegger]</i>	-
<i>[anders/speciaal]</i>	-
<i>NSTR goederensoort [0: vee en agrarische producten]</i>	2.226 (0.059)
<i>[1: voeding, veevoer]</i>	0.871 (0.035)
<i>[2-5: benzine, olie, metalen]</i>	-
<i>[6: bouwmaterialen]</i>	0.556 (0.081)
<i>[7: mest]</i>	-1.105 (0.327)
<i>[8: chemische producten]</i>	1.517 (0.063)
<i>[9: machines &amp; overig]</i>	-

#### 4.2 Schattingsresultaten Select Shipment keuzemodel

Het Select Shipment keuzemodel heeft drie verklarende variabelen. De kans dat een zending de volgende additionele zending is neemt allereerst af als de zending meer kosten toevoegt aan de rit, simpelweg omdat vervoerders transportkosten willen minimaliseren. De additionele gegeneraliseerde kosten zijn een gewogen som van de additionele tijd (€45,12/uur) en afstand (€0,45/km). Deze

kostenkengetallen omvatten bijvoorbeeld brandstof- en loonkosten, en zijn gebruikt in het BasGoed vrachtmodel (Significance, 2018).

Daarnaast wordt een zending eerder gekozen als deze minder stops toevoegt aan de rit, ofwel wanneer de zending geladen of gelost moet worden op locaties die reeds worden aangedaan in de rit. Dit leidt tot simpelere ritten met minder stops, wat dientengevolge ook kan leiden tot minder benodigde tijd voor parkeren en laden/lossen.

Tot slot neemt de kans dat een zending wordt gekozen toe als de zending van hetzelfde goederensoort is als de overige zendingen in de rit. Sommige goederen kunnen niet samen in een vrachtwagen vervoerd worden (bijv. vee en chemische producten), waardoor het merendeel van de gerapporteerde ritten in de data enkel zendingen van dezelfde goederensoort vervoert.

*Tabel 5. Schattingsresultaten Select Shipment keuzemodel.*

Select Shipment	
$R^2_{McFadden}$	0.187
LL	-63256
N	43409
<i>additionele gegeneraliseerde kosten [€]</i>	-0.005 (0.000)
<i>additionele aantal stops</i>	-1.039 (0.010)
<i>zelfde goederengroep als rit</i>	2.313 (0.038)

## 5. Validatie

Om het tourformatie algoritme met geschatte coëfficiënten te valideren, construeren we ritten met de zendingen in de resterende vijftig procent van de data waarop niet geschat is. Vervolgens vergelijken we geaggregeerde statistieken van de gerapporteerde en de voorspelde ritten in deze validatiedata. De voorspelde statistieken zijn gemiddeld genomen over drie runs om willekeur te verminderen.

In Tabel 6 en Figuur 3 zien we dat het tourformatie algoritme de werkelijke verdeling van het aantal stops op uitstekende wijze weet te reproduceren. Ook voor de ritafstand is dit het geval. Het aantal stops binnen verschillende goederengroepen reproduceert het algoritme ook vrij goed (Tabel 7). Daarnaast hebben veranderingen in modelspecificaties slechts zeer geringe invloed ( $\alpha$  = straal in km waarbuiten zendingen niet overwogen worden,  $\gamma$  = aantal zendingen gesampled voor de Select Shipment keuzeset). We kunnen dus concluderen dat het model erg robuust is voor modelspecificaties en geobserveerde ritpatronen sterk weet na te bootsen.

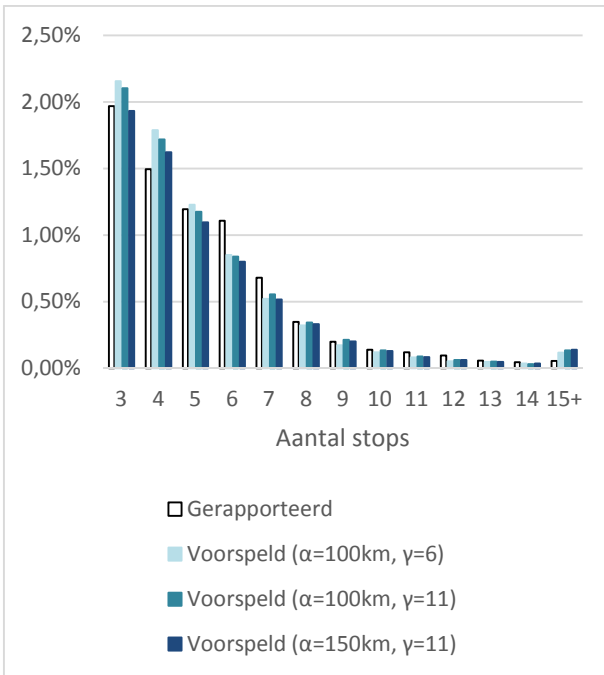
*Tabel 6. Percentage gerapporteerde en voorspelde directe ritten.*

Percentage direct ritten
--------------------------

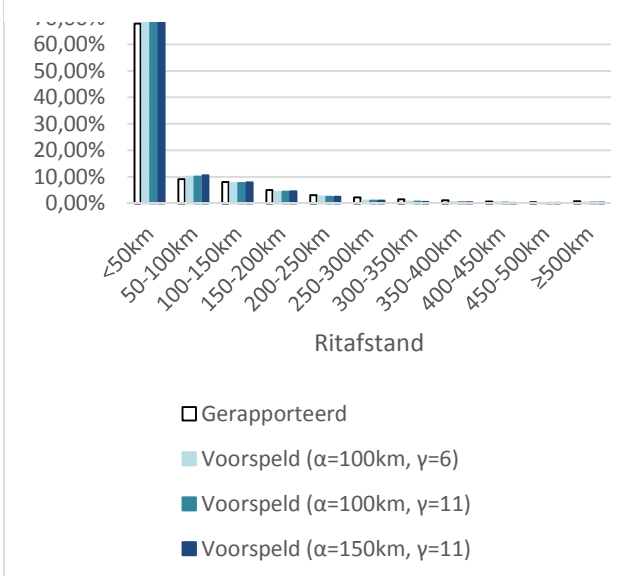
Gerapporteerd	Voorspeld ( $\alpha=100\text{km}$ , $\gamma=6$ )	Voorspeld ( $\alpha=100\text{km}$ , $\gamma=11$ )	Voorspeld ( $\alpha=150\text{km}$ , $\gamma=11$ )
92.50%	92.51%	92.56%	93.01%

Tabel 7. Percentage gerapporteerde en voorspelde directe ritten per goederengroep.

Goederensoort	Percentage direct ritten	
	Gerapporteerd	Voorspeld ( $\alpha=100\text{km}$ , $\gamma=6$ )
<i>0: vee en agrarische producten</i>	72.51%	75.24%
<i>1: voeding, veevoer</i>	64.80%	82.30%
<i>2-5: benzine, olie, metalen</i>	97.89%	96.16%
<i>6: bouwmaterialen</i>	97.49%	95.81%
<i>7: mest</i>	77.88%	80.29%
<i>8: chemische producten</i>	99.31%	99.45%
<i>9: machines &amp; overig</i>	84.14%	82.02%



Figuur 3. Percentage gerapporteerde en voorspelde ritten per aantal stops.



Figuur 4. Percentage gerapporteerde en voorspelde ritten per ritafstand.

## 6. Conclusies en aanbevelingen

Deze paper beschrijft een tourformatie algoritme dat ritten vormt uit zendingen. Schattingsresultaten van de twee probabilistische keuzestappen in dit algoritme zijn gerapporteerd. Daarna is het tourformatie algoritme toegepast op validatiedata. In deze laatste paragraaf bespreken we de sterke

punten en tekortkomingen van het tourformatie algoritme, waarna aanbevelingen voor vervolgonderzoek worden gedaan.

Doordat het algoritme op zendingen is gebaseerd, overwegen we dat zendingen de mogelijkheden en beperkingen bepalen voor tourformatie. Belangrijke beperkingen, zoals het laadvermogen en de beschikbaarheid van zendingen, worden expliciet gemodelleerd. Door kalibratie op inclusieve en empirische data komen ook verschillende voorkeuren en strategieën terug in het modelgedrag; reistijd en afstand worden geminimaliseerd door nabije zendingen te groeperen en een voorkeur voor een simpele planning komt naar voren in de constructie van directe ritten voor zendingen over korte afstand. Daarnaast houdt het algoritme rekening met de verschillen in tourformatie op verschillende locatietypen en voor diverse goederengroepen en voertuigtypen. Zo worden directe ritten vaker gemaakt op goederenoverslaglocaties, terwijl ritten vanaf een distributiecentrum vaak meerdere stops hebben.

De XML data rapporteert geen aankomst- en vertrektijden op tussenstops en geen lege trips, daarom houdt het tourformatie algoritme nog geen rekening met verblijftijden op stops, venstertijden, en lege ritten. Daarnaast versimpelt het algoritme het complexe tourformatie proces door telkens één zending toe te voegen en door de keuze voor deze zending slechts af te laten hangen van de zendingen die ervoor aan de rit zijn toegewezen. In werkelijkheid is het een meer geïntegreerd keuzeprocess, een transportplanner kiest niet herhaaldelijk een enkele zending. Toch laat toepassing van het tourformatie algoritme op validatiedata zien dat geobserveerde ritpatronen uitstekend gereproduceerd worden. Veel belangrijke gedragsmatige aspecten van tourformatie komen terug in het modelgedrag.

Het tourformatie algoritme kan worden toegepast in een vrachtmodel met zendingen tussen buurten in Nederland. Vanwege de zelfselectie in de XML data is het model alleen toepasbaar op beroepsvervoer. Daarnaast is het model enkel geschikt voor de Nederlandse context; in andere landen en regio's verschillen onder andere de wet- en regelgeving en de ruimtelijke verdeling van vraag en aanbod van goederen, waardoor modelcoëfficiënten zullen verschillen.

De belangrijkste vervolgstappen zijn gerelateerd aan de integratie van dit tourformatie algoritme in een vrachtmodel. Allereerst moet een realistische set zendingen gegenereerd worden om deze vervolgens toe te wijzen aan vervoerders. Daarnaast is integratie van tourformatie met een voertuigtypekeuzemodel nodig, het is een uitdaging om op correct wijze te modelleren dat het voertuigtype niet noodzakelijkerwijs vaststaat alvorens de tourformatie, het voertuigtype kan namelijk ook tijdens de tourformatie heroverwogen worden. Zowel het genereren van zendingen als de voertuigtypekeuze is onder ontwikkeling voor het MASS-GT voor Rotterdam en omgeving (de Bok et al., 2018). Met een set zendingen en een voertuigtypekeuzemodel kunnen we op basis van een netwerktoedeling en wegtellingen beter bepalen in hoeverre dit tourformatie algoritme leidt tot betere voorspellingen. Tot slot is een methode voor het bepalen van lege ritten een relevante vervolgstap.

## VERWIJZINGEN

- Anand, N., van Duin, J., & Tavasszy, L. (2014). Ontology based multi-agent system for urban freight transportation. *International Journal of Urban Sciences* 18(2), 133-153.
- Boerkamps, J., & van Binsbergen, A. (1999). GoodTrip - A new approach for modelling and evaluation of urban goods distribution. *International Conference on City Logistics 1st*, (pp. 175-186). Cairns, Australia.
- de Bok, M., Tavasszy, L., Bal, I., & Thoen, S. (2018). The incremental development path of an empirical agent-based simulation system for urban goods transport (MASS-GT). *World Conference on Transport Research*. Mumbai, India.
- Friedrich, H., Tavasszy, L., & Davydenko, I. (2014). Distribution Structures. In L. Tavasszy, & G. De Jong, *Modeling Freight Transport* (pp. 65-87). London, UK: Elsevier.
- Holguín-Veras, J., González-Calderón, C., Sánchez-Díaz, I., Jaller, M., & Campbell, S. (2014). Vehicle-Trip Estimation Models. In L. Tavasszy, & G. de Jong, *Modeling Freight Transport* (pp. 143-162). London, UK: Elsevier.
- Hunt, J., & Stefan, K. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research Part B* 41, 981–1013.
- Khan, M., & Machemehl, R. (2017). Analyzing tour chaining patterns of urban commercial vehicles. *Transportation Research Part A* 102, 84-97.
- Sánchez-Díaz, I., Holguín-Veras, J., & Ban, X. (2015). A time-dependent freight tour synthesis model. *Transportation Research Part B* 78, 144–168.
- Significance. (2018). *Schattingsrapport Basgoed 2018 (Werkversie 23 april 2018)*. Den Haag, the Netherlands: Significance.
- Wisetjindawat, W., Sano, K., Matsumoto, S., & Raathanachonkun, P. (2006). Micro-simulation model for modeling freight agents interactions in urban freight movement. *Transportation Research Board 86th Annual Meeting*. Washington, DC, USA.
- You, S. I., Chow, J. Y., & Ritchie, S. G. (2016). Inverse vehicle routing for activity-based urban freight forecast modeling and city logistics. *Transportmetrica A: Transport Science* 12(7), 650-673.