

De distributiestructuur van het goederen wegvervoer in NL

Michiel de Bok, TU Delft/Significance
Lóri Tavasszy, TU Delft
Ali Nadi, TU Delft
Raeed Mohammed, TU Delft

Samenvatting

Goederenvervoer is het transport gedeelte van het logistieke systeem: goederen worden van A naar B vervoerd. Een groot gedeelte is onderdeel van logistieke ketens. Waar data over goederenvervoer in toenemende mate beschikbaar is voor onderzoek en planningsdoeleinden, blijft de beschikbaarheid van logistieke data schaars.

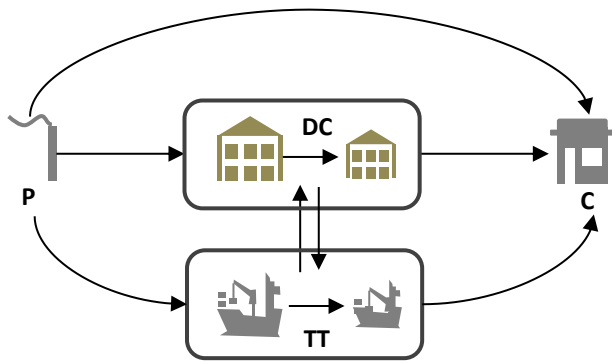
In deze bijdrage presenteren wij de resultaten van een data gedreven onderzoek waarbij 'big' tripdata van transporteurs zijn geanalyseerd op distributiestructuren. Uitdaging daarbij is de transport data te verrijken met logistieke informatie: vond deze rit plaats vanuit een multimodale terminal, een distributiecentra, of kwam deze vanaf een producent? Op de TU Delft hebben we een effectieve methode opgezet om structurele distributiepatronen te ontdekken, ondanks de data-inefficiënties.

De resultaten geven een relevante inkijk in distributiestructuren voor verschillende segmenten in het goederenvervoer: informatie die tot nog toe ontbreekt.

Inleiding

Goederenvervoer is het transport gedeelte van het logistieke systeem: goederen worden van A naar B vervoerd. Een groot gedeelte is onderdeel van logistieke ketens. Waar data over goederenvervoer in toenemende mate beschikbaar is, blijft de beschikbaarheid van logistieke data voor onderzoek en planningsdoeleinden schaars. Voor Nederland als distributieland is de organisatie van distributiekanaal van essentieel belang voor de planning van onze infrastructuurnetwerken. Onstein et al (2018) geeft een inkijk in hoe bedrijven hun distributienetwerken vormgeven: voor de locatiebepaling van nieuwe distributiecentra zijn logistieke kosten, serviceniveau en de klantvraag de belangrijkste (hoofd)factoren. Wat echter nog niet onderzocht is, is de hoeveelheid goederen, die via dergelijke distributiecentra afgehandeld worden: dergelijke data is niet beschikbaar op systeemniveau.

In dit onderzoek willen we kwantitatief inzicht krijgen in de patronen vervoersvolumes per vervoerssegment. Voor distributie structuren maken we onderscheid naar producenten (C) en gebruikers van goederen (P) en in de distributiekanaal maken we onderscheid naar logistieke knopen: distributie centra (DC) en multimodale transshipment terminals (TT).



Figuur 1: Schematisatie van verschillende vervoerssegmenten in distributiekanaal

In deze bijdrage presenteren wij de resultaten van een data gedreven onderzoek waarbij 'big' tripdata van transporteurs is geanalyseerd op distributiestructuren. Uitdaging daarbij is de transport data te verrijken met logistieke informatie: vond deze rit plaats vanuit een multimodale terminal, een distributiecentra, of kwam deze vanaf een producent? Op de TU Delft hebben we een effectieve methode opgezet om structurele distributiepatronen te ontdekken, ondanks de data-inefficiënties.

We benoemen eerst de bronnen die we hebben gebruikt om de analyse op te zetten. Vervolgens beschrijven we de methode waarop we deze bronnen hebben gecombineerd om te komen tot een onderzoeksbestand, verrijkt met de logistieke informatie. Vervolgens bespreken we de resultaten van het onderzoek: welke segmenten maken in welke mate gebruik van distributiekanaal? Ten slotte bespreken we onze conclusies bij de resultaten en bespreken we de verdere toepassingsmogelijkheden en beperkingen van deze data-analyse.

Databronnen

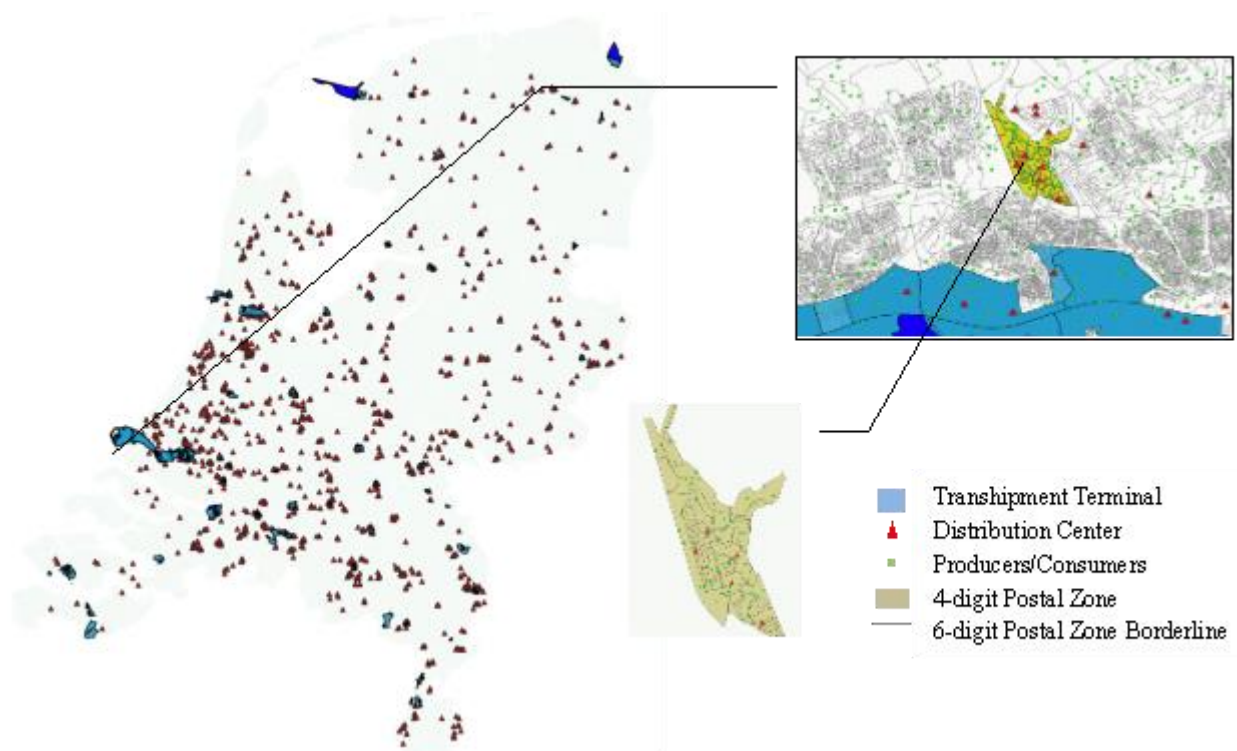
De belangrijkste bron voor dit onderzoek is een rittendagboekje van vrachtvervoer: **de XML-microdata wegvervoer van het CBS**. Dit wordt jaarlijks verzameld door het CBS om de Basisbestanden Goederenwegvervoer te actualiseren. Dit rittendagboekje bevat informatie over het voertuig, de route die het gereden heeft, én de lading die het vervoerde. De combinatie van deze

gegevens in één bron maakt het tot een waardevol onderzoeksbestand die de basis vormt voor de meeste transport statistieken in Nederland. De vragenlijsten bevatten geen vragen meer over de laad- en los locatie van het voertuig, of informatie over de verzender of ontvanger.

Om het invullen van de vragenlijst te vergemakkelijken hebben bedrijven sinds 2013 de mogelijkheid de data rechtstreeks uit hun Transport Management Systemen (TMS) uit te leveren: in het XML-dataformat. Als gevolg van de geautomatiseerde verzameling is de omvang van het onderzoeksbestand aanzienlijk toegenomen. Deze XML data bevat bovendien de locatie gegevens met veel meer detail dan de online vragenlijsten: dit maakt het ook mogelijk de data op lager detail niveau te koppelen aan locatie data.

De eerste soort locatie data die gekoppeld is aan de ritbestanden zijn **distributiecentra** en **multimodale terminals**. De locaties van distributiecentra zijn vastgelegd in een onderzoeksbestand van Rijkswaterstaat. Het onderzoeksbestand met locatie van multimodale terminals is afkomstig van het KiM (KiM, 2019).

Ten slotte is voor de locaties van potentiële verzenders of ontvangers van de goederen is gebruik gemaakt van het **Algemeen BedrijfsRegister (ABR)** van het CBS. Dit bevat de locatie van bedrijven op 6-cijferig postcode niveau, de SBI code van het bedrijf en het aantal werknemers.



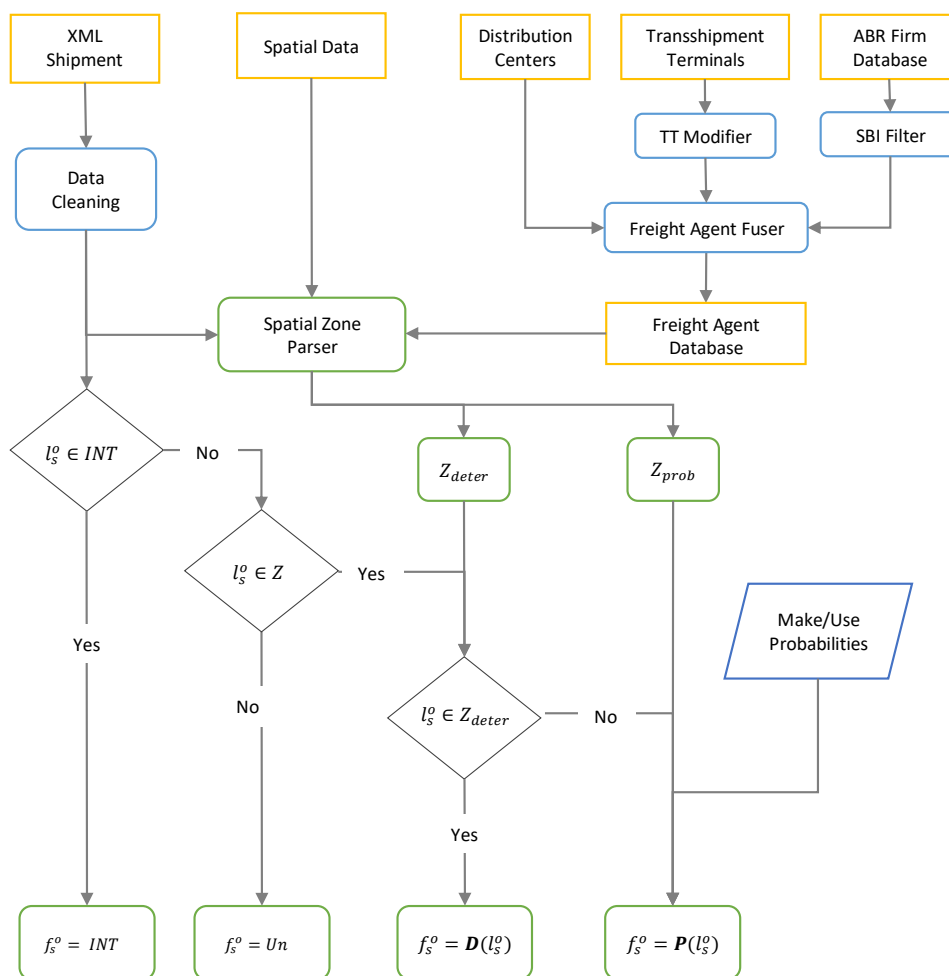
Figuur 2: Visualisatie van gebruikte data (n.b. de XML- en bedrijfsregister data zijn niet weergegeven)

Methode

De basisgedachte bij de data fusie is de data te koppelen op basis van hun geografische locatie. Dit kan voor het gedeelte van de data waarvoor de adresinformatie volledig is direct gebeuren. Op postcode 6 niveau kan een unieke link gelegd worden of een laad- of loslocatie bij een distributiecentrum plaatsvond, een multimodale terminal, of bij een producent of consument. De

externa databases zijn samengevoegd tot een “Freight Firms” database: een bestand met hierin de relevante actoren voor goederenvervoer. Hierin worden 4 verschillende logistieke functies onderscheiden: multimodale terminals (55), distributie centra (ca. 1600), en producent en consumenten (ruim 300.000 bedrijven).

De toedeling is echter meer complex dan deze basisgedachte, om twee redenen. Allereerst is de adres informatie van de laad- en loslocatie niet altijd volledig. Daarnaast kunnen op een laad of loslocatie meerdere bedrijven gevestigd. Daarom is een hiërarchisch algoritme uitgewerkt wat stapsgewijs de laad- en loslocaties koppelt aan de externe databases. Het algoritme wordt beschreven in het volgende stroomschema. Voor detail beschrijving van het algoritme zie Mohammed et al. (2023).

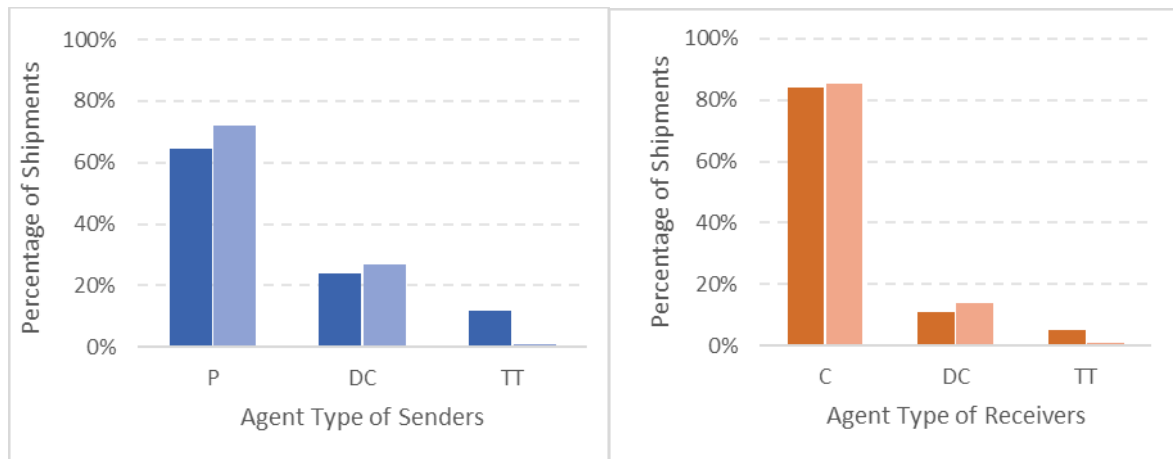


Figuur 3: Datafusie algoritme

Als de laad- en loslocatie exact bekend zijn, en op de locatie is een distributiecentrum, multimodale terminal of uniek bedrijf gevestigd, dan wordt de zending uniek gekoppeld. Daarbij wordt verondersteld dat als eerste gekoppeld wordt aan de terminal, anders een distributiecentrum, of ten slotte aan een individueel bedrijf. De probabilistische toedeling wordt gebruikt als er meerdere bedrijven zijn waaraan de zending gekoppeld kan worden: de toedeling wordt dan gebaseerd op een kansfunctie die rekening houdt met de goederensoort, de industriesector van de betreffende

bedrijven, en de aanbod/gebruik aandeel van deze bedrijfssector in de productie en consumptie van deze goederensoort. Deze laatste aandelen zijn beschikbaar in algemene aanbod/gebruik statistiek van het CBS. De probabilistische toedeling wordt ook gebruikt indien de locatie van de zending niet compleet is en alleen de postcode 4 beschikbaar is. Er wordt dan rekening gehouden met een groter gebied waarbinnen de zending is afgeleverd en dus een groter aantal bedrijven wat mogelijk de verzender of ontvanger is.

De volgende figuur vat de uitkomsten van de procedure samen: het aandeel zendingen wat geladen of gelost wordt op logistieke knopen, zowel uit de deterministische als de probabilistische toedeling.



Figuur 4: Aandeel herkomst- en bestemmingstypen: Producent (P), Distributiecentrum (DC), Terminal (TT) of Consument (C), met onderscheid deterministische matching (donker) of probabilistische (licht).

Resultaten

De samengevoegde trip- en locatiedata is vervolgens geanalyseerd in samenstelling naar distributiestructuur. Daarbij kunnen we kijken naar het aandeel per distributie segment (direct of via distributiecentrum/transshipment terminal) en de kenmerken van de lading: gemiddeld gewicht, aantal zendingen per trip (mate van consolidatie), en ritafstanden. De volgende tabel vat de resultaten samen.

Tabel 1: Globale statistieken uit het onderzoeksbestand

Distributie Segment	Aandeel zendingen (%)	Zending grootte gem. (ton)	Aantal zendingen in rit gem.	Aandeel ritten (%)	Ritafstand gem. (km)
P-C	60.1	14.4	1.9	79.6%	54.9
P-DC	7.5	2.6	4.2	4.5%	73.2
DC-C	21	1.1	8.8	6.0%	100.9
DC-DC	7.6	1.4	7.6	2.5%	88.5
P-TT	1.2	8.9	1.2	2.5%	86.9
TT-C	2.1	11.9	1.3	4.1%	104.3
DC-TT	0.2	3.3	2.2	0.2%	104.5
TT-DC	0.2	14.3	1.3	0.4%	110.7
TT-TT	0.1	3.7	1.1	0.23%	99.5

Uit de algemene statistiek blijkt dat ongeveer 60% van alle zendingen, en 80% van de ritten direct zijn tussen een producent en consument (P-C). Daarnaast blijkt dat ook een groot gedeelte van de zendingen van en naar distributiecentra vervoerd wordt: ruim 20% van de zendingen gaat van een distributiecentrum naar consument (6 % van het aantal ritten), ruim 7% van de zendingen gaat van producent naar distributiecentrum (4,5% van het aantal ritten) en ruim 7% van de zendingen betreft vervoer tussen distributiecentra (2,5% van het aantal ritten). Dit kan vergeleken worden met de beperkte schaarse bronnen die beschikbaar zijn. Davydenko (2015) schatte in het verleden dat distributiecentra en logistieke knopen verantwoordelijk zijn voor zo'n 40% van het totaal aantal trips. Qua orde grootte komt dit wel overeen met het aandeel zendingen, maar gemeten in aantallen ritten is het aandeel in onze analyse lager: 13% van het aantal ritten gaat van/naar een distributiecentrum en 34% van het aantal zendingen.

Internationale voorbeelden van statistieken op deze schaal zijn schaars. Recent hebben Yang et al. (2023) de ritten van 2,6 miljoen vrachtwagens in China geanalyseerd: hier bleek 10% van de ritten een 'logistic warehouse' te zijn (Yang et al, 2023). Dit aandeel is vergelijkbaar met de 13% die wij in onze data voor Nederland vinden. Op basis van deze vergelijking gaan wij ervanuit dat de structuren die gevonden zijn representatief zijn.

Directe zendingen blijken gemiddeld ook een groter gemiddeld gewicht (14.4 T) te hebben dan indirecte zendingen via distributiekanaalen. Dit lijkt plausibel omdat afnemers van grotere ladingseenheden sneller gebruik maken van een full truck load, wat efficiënt is vanuit transportkosten. Grote lading vervoeren via een distributiekanaal kan leiden tot onnodige hogere transport en overslag kosten.

Indirect Vervoer daarentegen, is vaker geconsolideerd: het gemiddeld aantal zendingen per rit is veel hoger bij de ritten vanuit een distributie centrum of daartussen. Dit laat zien dat distributiekanaalen gebruikt worden om zendingen te consolideren, wat leidt tot efficiëntere transportkosten. Dit is een natuurlijk een belangrijke reden voor het bestaan van distributiekanaalen. We kunnen de ritpatronen van goederenvervoer hierdoor ook beter begrijpen en beschrijven: het type rondrit wordt voor een deel verklaard of het een rondrit is die vertrekt vanuit een DC voor een bundel afleveringen. Vanuit de producent naar dc zien we minder consolidatie: dit kan ook verklaard worden dat de producent een product in grote batch aanlevert aan de groothandel, waarna het in kleinere hoeveelheden wordt doorverkocht.

Bij elkaar zijn dit algemene patronen die goed aansluiten bij de logica van distributie structuren: consolideren van individuele zendingen (voor kosten efficiency). Dit geeft vertrouwen in de bruikbaarheid van dergelijke statistiek. Nieuw is het op grote schaal kunnen kwantificeren van het bestaan van deze structuren.

De resultaten laten per goederengroepen grote variaties zien. Daarom zijn de belangrijkste indicatoren ook per goederengroep uitgesplitst: zie Tabel 2 en 3 voor het aandeel per segment, de zendinggrootte en het aantal gecombineerde zendingen in een rit (consolidatie). De eerste groep betreft de niet bulk goederensoorten: stukgoed, (half) fabricaten, landbouwproducten (niet gekoeld) en versproducten. Daarnaast worden vier bulk groepen onderscheiden: mineralen, bedrijfsafval, chemische producten en bouwmaterialen.

Bij de niet bulkgoederen, zien we vooral een hoog aandeel van gebruik distributiecentra (de gekleurde segmenten in de taartdiagrammen). We zien dit bij alle goederengroepen: bij landbouwproducten zijn de meeste zending gerelateerd aan distributiecentra, bij stukgoederen en (half)fabricaten is dit ongeveer de helft, en bij versproducten is dit ongeveer één derde.

Bij landbouwproducten is ook duidelijk te zien dat de grote zendingen met direct transport van producent naar afnemers gaan: dit is te zien aan de grijze lijn in de cumulatieve distributiefunctie voor zendinggrootte in de middelste kolom. Bij de overige niet bulk groepen laat de CDF van zendinggrootte zien dat er meer groetere zendinggroottes voorkomen in het direct transport (de P-C lijn is de onderste lijn in de figuur).


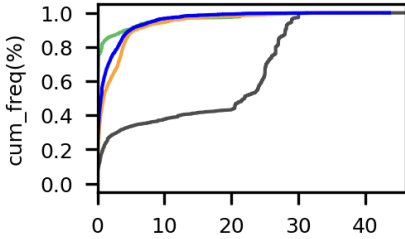
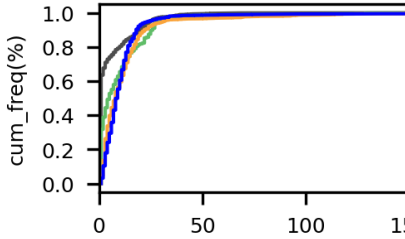

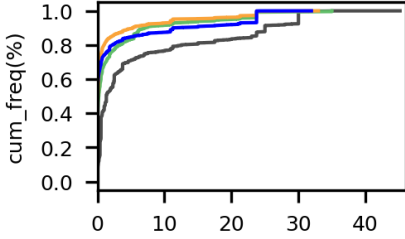
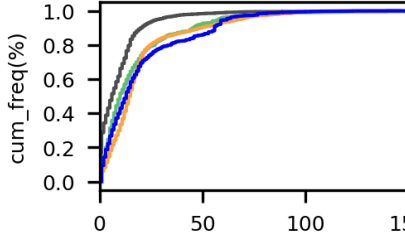

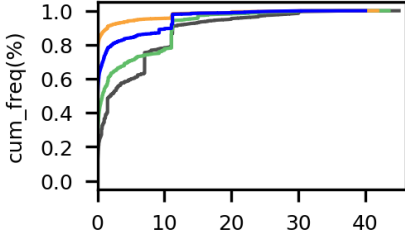
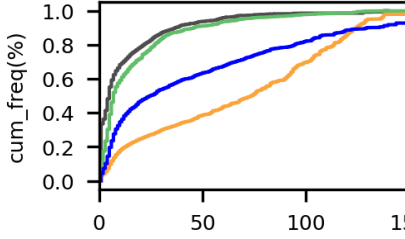

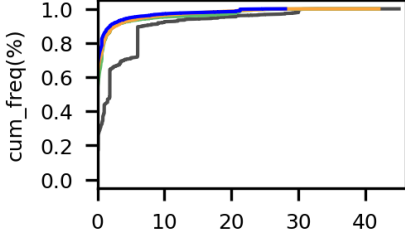
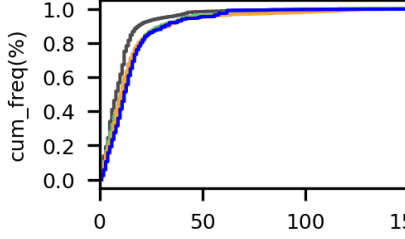
De cumulatieve distributiefunctie voor consolidatie laat zien dat bij indirect transport ook grotere aantallen zendingen per rondrit voorkomen. Vooral bij (half)fabricaten is dit duidelijk te zien. Vooral de stromen van distributiecentra naar consumenten hebben veel grote aantallen zendingen per rondrit. De stromen van producent naar distributiecentra blijken minder gefragmenteerd en hebben minder zendingen per rondrit. Ook dit kan logisch verklaard worden door de distributiefunctie: de voorraad van ene producent wordt via het distributiecentrum geleverd aan meerdere afnemers.

Versproducten en stukgoederen hebben ook veel kleinere zendingen maar de mate van consolidatie is niet duidelijk groter. We vermoeden dat om praktische redenen de consolidatie niet groter kan zijn: voor deze producten geldt dat tijdige levering belangrijker is in verband met houdbaarheid van de producten, en ook voor stukgoederen kan just-in-time management voor een deel van de producten ervoor zorgen dat de consolidatie minder is. Een tweede reden kan zijn dat de volume capaciteit in deze productgroepen vaker de bottle necks is. Nog meer zendingen in één rondrit is niet mogelijk omdat de volume capaciteit al bereikt is; dit is eerder al geconcludeerd door Davydenko en Tavasszy (2015).


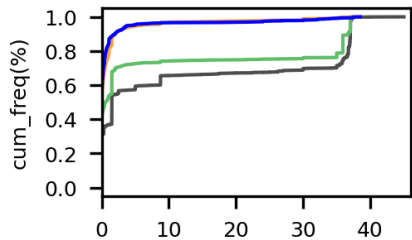
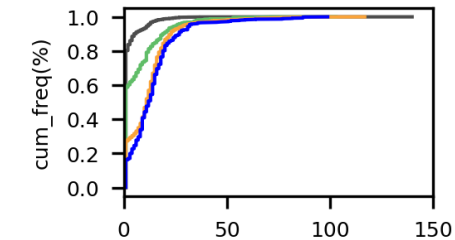
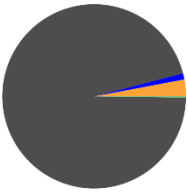
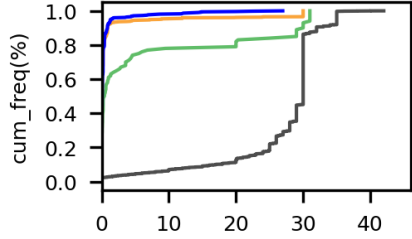
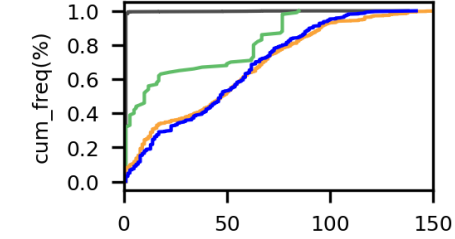

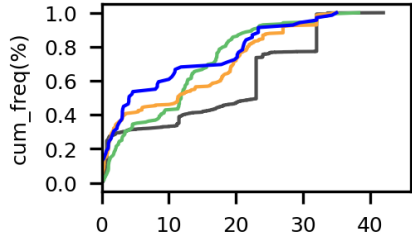
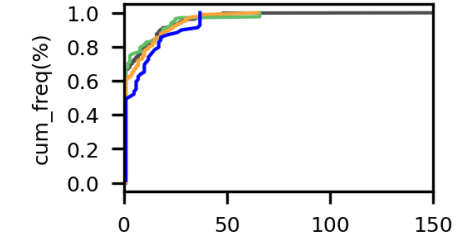
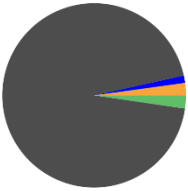
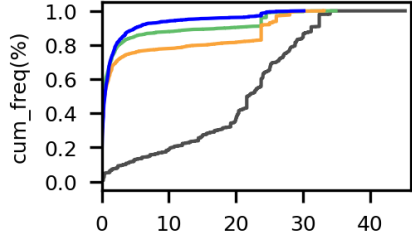
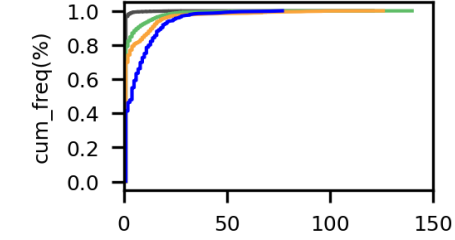
Bij de bulk goederengroepen mineralen, chemicaliën en bouw materiaal zien we veel direct transport van producent naar consument. De zendingen zijn ook significant zwaarder dan de andere goederengroepen. Bouwmaterialen laten nog het meeste zendingen via distributiecentra zien. Veel bevoorrading aan de bouwplaats loopt ook via de bouw groothandel wat deze structuur verklaart.

Een groot gedeelte van de natte bulk (chemische producten) en droge bulk (mineralen en bedrijfsafval) worden vervoerd in grote zendingen rond de 20 à 30 ton, en het meeste zijn directe zendingen (grijze lijnen). Vooral mineralen lijken standaard in ladingen van 30 Ton vervoerd te worden naar de afnemers. Er zijn ook een beperkt aantal lading stromen van minerale producten die juist in grote bundelingen van zendingen via distributiekanaal vervoerd worden: hierbij gaat het waarschijnlijk om het vervoer en distributie van zeldzame mineralen voor de industrie.

Bij bedrijfsafval hebben ook een groot gedeelte van de ladingen die via distributiekanaal gaan een grote omvang: het gaat hier om verzamel depots voor dit afval.

Goederensoort	Aandeel per logistiek segment	ECDF van zendinggrootte	ECDF mate van consolidatie (zendingen/rit)
Landbouwproducten (niet-gekoeld)			
Stukgoed			
(Half)fabricaten			
Versproducten			

Tabel 2: Beschrijvende statistieken niet-bulk groepen. Legenda: [— P-C — P-DC — DC-C — DC-DC]

Goederensoort	Aandeel per logistiek segment	ECDF van zendinggrootte	ECDF mate van consolidatie (zendingen/rit)
Bouwmaterialen			
Mineralen			
Bedrijfsafval			
Chemische producten			

Tabel 3: Beschrijvende statistieken bulk groepen.

Legenda: [— P-C — P-DC — DC-C — DC-DC]

Discussie en verder onderzoek

In deze bijdrage willen we de toepassingsmogelijkheden laten zien van het gebruik van big data voor goederen (weg)vervoer in Nederland. Om echt meerwaarde te creëren met dergelijke geautomatiseerde data is het toch noodzakelijk om ontbrekende informatie op een slimme manier aan te vullen. Het verzamelen van trip data uit planningssystemen van vervoerders leidt tot onderzoeksbestanden met grote dichtheid, maar er blijven onvolkomenheden. In dit geval is het ontbreken van gedetailleerde adresinformatie ondervangen door het soort laad- en loslocatie in dit gedeelte van de warnemingen met een probabilistische toedeel algoritme toch in te vullen. Dit geeft ons de mogelijkheid om structuren in de volledige data te ontdekken.

Wij hebben gebruik gemaakt van een heuristische data-gebaseerde methode geschikt voor afleiden distributiepatronen in het goederenvervoer: wat is het gebruik van distributiekanaalen en wat is de aard en de omvang van deze goederenstromen? Op basis van de resultaten is gekeken naar de omvang en samenstelling naar transport karakteristieken van de verschillende vervoersbewegingen in verschillende segmenten van een distributiekanaal. De resultaten laten patronen zien die passen bij de algemene verwachting bij hoe goederenvervoer georganiseerd wordt. Dit geeft ook vertrouwen in de toepassingsmogelijkheden om de data te gebruiken voor beschrijvende modellen van logistieke keuzes. Aan de TU Delft wordt dit bestand gebruikt voor de ontwikkeling van nieuwe agent-gebaseerde simulatiemodellen voor logistiek (De Bok en Tavasszy, 2018).

Verder onderzoek in dit domein is nodig: zo is het met deze data nog niet mogelijk om de onderlinge verbanden van opeenvolgende transportbewegingen in een distributieketen te onderzoeken. En eerste verkenning is uitgevoerd door CBS en TNO (Jacobs et al, 2018; de Blois et al, 2018). Daaropvolgend wordt in een vervolgproject door IenW en CBS ingezet op grotere dataverzameling door het betrekken van meer logistieke spelers en het gebruik maken van OTM standaarden (Hazelhorst en Jacobs, 2020). Ook dergelijke dataverzameling zal samengaan met data onvolkomenheden. De ervaring opgedaan in de hier gepresenteerde analyse kan waardevol zijn om dergelijke verzamelingen te verrijken.

Referenties

Davydenko IY. (2015) Logistics Chains in Freight Transport Modelling. Delft University of Technology; 2015.

Davydenko, IY, L Tavasszy (2015) Vrachtauto's zijn toch beter benut? Een eerste kijk naar gewicht, volume en oppervlakte benutting in Nederlands wegvervoer, Paper gepresenteerd op de Vervoerslogistieke Werkdagen 2015.

de Blois CJ, R Fransen, M Jacobs, J van Meijeren, P Tilanus (2018) Constructie van multimodale transportketendata, rapport TNO 2018 R11516, 14 december 2018.

De Bok, M, L Tavasszy (2018) De incrementele ontwikkeling van een data-driven agent-gebaseerd simulatiemodel voor goederenvervoer. Paper gepresenteerd op de Vervoerslogistieke Werkdagen, 15 en 16 november 2018, Vaals.

Hazelhorst, WO, M Jacobs (2020) Containervervoer in de keten beter in beeld.
Voorlichtingsbijeenkomst Rijkswaterstaat 15 september 2020.

Jacobs, M, C de Blois, D Dahlmans (2018) Data voor multimodale transportketens: uitbreiding in-, uit- en doorvoerstatistiek en koppeling containerdata, Paper gepresenteerd op de Vervoerslogistieke Werkdagen, 15 en 16 november 2018, Vaals.

Jacobs, M, R Meijers (2022) Rapport Pilot Containervervoer, CBS: Heerlen.

Kennisinstituut voor Mobiliteitsbeleid (2019) Benutting Multimodale Achterlandknooppunten, KIM/Ministerie van Infrastructuur en Waterstaat, december 2019

Mohammed, R, A Nadi, L Tavasszy, M de Bok (2023) A Data Fusion Approach To Identify Distribution Chain Segments In Freight Shipment Databases, Transportation Research Records.

Onstein, S., Ektesaby, M., Rezaei, J., van Damme, D., & Tavasszy, L. (2018). Welke factoren bepalen de ruimtelijke inrichting van distributienetwerken? *Logistiek +: Tijdschrift Voor Toegepaste Logistiek*, 2018(6), 72-87

Yang, Y., Jia, B., Yan, X.Y., Li, J., Yang, Z., Gao, Z. (2022). "Identifying intercity freight trip ends of heavy trucks from GPS data." *Transportation Research Part E: Logistics and Transportation Review*. Vol. 157 DOI: 10.1016/j.tre.2021.102590.