

# **Stedelijkheidsgraden in transportmodellen: Een analyse over de impact van de ruimtelijke omgeving op reisgedrag**

Janine Grauss – ProRail/TU Delft – janinetimmerman@outlook.com

Justin Hogenberg – ProRail – Justin.Hogenberg@prorail.nl

Niels van Oort – TU Delft – N.vanOort@tudelft.nl

## **Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 21 en 22 november 2024, Utrecht**

### **Samenvatting**

Nederland is een klein land, maar kent veel gebieden met uiteenlopende ruimtelijke kenmerken. Verschillen in de ruimtelijke omgeving hebben een effect op reisgedrag, wat het een uitdaging maakt om reisgedrag in Nederland te vangen met één landelijk transportmodel.

Het Landelijk Model Systeem (LMS) is een transportmodel voor heel Nederland. Om onderscheid te maken in vervoersgedrag tussen verschillende gebieden, speelt de stedelijkheidsgraad een belangrijke rol. Deze variabele deelt Nederland op in zes clusters op basis van de bevolkingsdichtheid. Onderzoek Verplaatsingen in Nederland (OViN) is enquêtedata over reisgedrag. Zowel OViN als LMS data laten een correlatie zien tussen de stedelijkheidsgraad en de gemiddelde modaliteitskeuze. Zo neemt autogebruik af naarmate de stedelijkheidsgraad toeneemt en OV gebruik toe.

Echter, in tegenstelling tot het LMS, kunnen volgens OViN gebieden met dezelfde stedelijkheidsgraad toch een erg verschillende modal split hebben. Zo hebben Zoetermeer, Delft en Leiden dezelfde stedelijkheidsgraad, maar wordt er meer autogereden en minder gefietst in Zoetermeer. Om dit verschil in reisgedrag goed te vangen, zijn aanvullende ruimtelijke kenmerken nodig. Dit kan door middel van de 'D-variabelen': dichtheid (Density); diversiteit van landgebruik (Diversity); ontwerp van het wegennetwerk (Design); bereikbaarheid van bestemmingen (Destination accessibility); kwaliteit van het OV (Distance to transit) en vraagsturing (Demand management).

We hebben de LMS-zones geclusterd op basis van deze D-variabelen. Deze clusters onthullen nieuwe gebieden die niet de standaard 'stedelijkheidsgraadtrends' volgen. Bijvoorbeeld, de centra van middelgrote steden (bijv. Enschede) hebben een lagere stedelijkheidsgraad dan de buitenwijken van hoog-stedelijke gebieden, maar hebben toch een lager autogebruik volgens OViN. Voornamelijk autobestuurder- en fietsgebruik in deze gebieden wordt momenteel niet goed gemodelleerd door het LMS.

Om het werkelijke effect van de ruimtelijke omgeving te kwantificeren, is er met 'propensity score matching' gecorrigeerd voor demografische kenmerken. Hieruit volgt dat de ruimtelijke omgeving gemiddeld verantwoordelijk is voor meer dan 50% van de verschillen in modal split tussen de clusters en dat verschillen in de demografie een kleinere impact hebben. Dit betekent dat het essentieel is om ruimtelijke kenmerken te gebruiken in een transportmodel. Niet alle verschillen in modaliteitskeuze tussen gebieden kunnen worden gevangen door demografische kenmerken.

Uit ons onderzoek blijkt dat alleen de stedelijkheidsgraad onvoldoende in staat is om de ruimtelijke omgeving te differentiëren en dat het LMS daardoor de modaliteitskeuze in relevante gebieden onvoldoende goed modelleert. Door meer aspecten van de ruimtelijke omgeving toe te voegen, wordt het LMS beter geschikt om beleid te testen dat invloed heeft op de ruimtelijke omgeving.

## 1. Inleiding

Het Landelijk Model Systeem (LMS) is een transportmodel dat wordt beheerd door Rijkswaterstaat en dat wordt gebruikt om voorspellingen te doen voor het spoor- en wegennetwerk in Nederland (Hofman, 2017). Het LMS is een complex model dat in staat is om veel verschillende effecten die invloed hebben op reisgedrag mee te nemen. Geen enkel model is echter in staat om de realiteit compleet na te bootsen en transportmodellen worden steeds verder verbeterd. Nederland is een klein land, maar bestaat uit veel verschillende gebieden. Deze gebieden variëren van hoog stedelijke gebieden met een uitgebreid (openbaar) vervoersnetwerk tot landelijke gebieden waar mensen erg afhankelijk zijn van hun auto. Deze verschillen in omgeving hebben een effect op het reisgedrag van mensen (e.g. Kent et al., 2023; Cao et al., 2009), wat het een uitdaging maakt om reisgedrag in een heel land te vatten met slechts één model.

In het LMS speelt de stedelijkheidsgraad een belangrijke rol bij het onderscheid maken in verschil in reisgedrag in verschillende gebieden. Deze variabele deelt Nederland op in 6 clusters op basis van de bevolkingsdichtheid, maar het is onduidelijk of alleen de bevolkingsdichtheid genoeg is om onderscheid te maken tussen gebieden. De hypothese van dit onderzoek is dat door alleen of voornamelijk gebruik te maken van variabelen in transportmodellen die gebaseerd zijn op de bevolkingsdichtheid, er onvoldoende onderscheid wordt gemaakt tussen gebieden die verschillend reisgedrag vertonen. Dit geeft de volgende onderzoeksvraag:

*In hoeverre is de stedelijkheidsgraad in staat om verschil in reisgedrag in verschillende gebieden mee te nemen in transportmodellen en hoe kunnen deze verschillen realistischer worden meegenomen in dezelfde transportmodellen?*

Om de scope van het onderzoek te limiteren, is er alleen gefocust op modal split en niet op andere aspecten van reisgedrag.

Dit paper is als volgt opgebouwd. Hoofdstuk 2 geeft de methodiek die is gebruikt. Vervolgens geeft hoofdstuk 3 een korte samenvatting van de aanwezige literatuur en een overzicht van relevante factoren in de LMS-documentatie. Hoofdstuk 4 geeft de resultaten en hoofdstuk 5 de conclusies.

Dit paper is gebaseerd op een Master thesis van de studie 'Traffic & Transport Engineering' aan de TU Delft en is uitgevoerd als afstudeerstage bij ProRail. Deze thesis is tot stand gekomen onder begeleiding van de twee coauteurs en Adam Pel en Bart van Arem van de TU Delft. Voor het volledige onderzoek, zie Timmerman (2024).

## 2. Methodiek

Dit onderzoek bestaat uit twee delen: een theoriegedeelte en een data-analyse. Het eerste gedeelte bestaat uit een literatuuronderzoek en een analyse van de LMS-documentatie. In het literatuuronderzoek is uitgezocht welke ruimtelijke factoren een effect hebben op reisgedrag. Vervolgens zijn deze factoren uit de literatuur vergeleken met de ruimtelijke factoren die worden meegenomen in het LMS. Dit geeft een beeld van in hoeverre het LMS de ruimtelijke omgeving meeneemt bij het modelleren van de modal split.

In de data-analyse zijn de synthetische matrices van het LMS voor 2018 (Rijkswaterstaat, Water, Verkeer en Leefomgeving [RWS WV], 2018) vergeleken met Onderzoek Verplaatsingen in Nederland (OVIN). OVIN data is enquête data waarin mensen uit heel Nederland zijn gevraagd om alle reizen die ze hebben gemaakt op één dag bij te houden. Deze dataset is vervolgens verwerkt en aangevuld met extra informatie uit overheidsregisters (Centraal Bureau van Statistiek & Rijkswaterstaat [CBS & RWS], 2017c). De dataset bevat niet alleen data over de reis zelf (e.g. modaliteit, afstand), maar ook persoonlijke gegevens (e.g. leeftijd, huishoudsamenstelling). OVIN data van 2015-2017 is gebruikt om de huidige versie van het LMS te kalibreren. Door het vergelijken van de LMS-matrices met OVIN, kan een beeld worden gevormd van de accuraatheid van het LMS. De data-analyse bestaat uit drie delen. Voor dit onderzoek zijn de OVIN datasets van 2013-2017 gebruikt, die voorafgaand aan de analyse zijn gestapeld (CBS & RWS, 2014; CBS & RWS, 2015, CBS & RWS, 2017a, CBS & RWS, 2017b, CBS & RWS, 2017c). Naast OVIN- en LMS-data, is er ook data verzameld van de ruimtelijke kenmerken van Nederland die volgde uit de literatuur. Zie Timmerman (2024) voor meer informatie over het verzamelen en verwerken van alle data.

De eerste stap is een onderzoekende analyse. Hierbij zijn eerst de gemiddelde modaliteitskeuzes van elke stedelijkheidsgraad met elkaar vergeleken. Vervolgens is er ingezoomd op verschillende gebieden en is de modaliteitskeuze op zoneniveau vergeleken van gebieden met (grotendeels) dezelfde stedelijkheidsgraad.

De tweede stap is een clusteranalyse. Hierbij zijn de zones in het LMS geclusterd op basis van ruimtelijke kenmerken door middel van een hiërarchische cluster methode. De ruimtelijke kenmerken die gebruikt zijn voor het clusteren zijn gebaseerd op de resultaten uit het literatuuronderzoek. Het doel van deze clustermethode is om clusters te vinden die grotere verschillen laten zien in (gemiddeld) reisgedrag dan de stedelijkheidsgraden (e.g. de stedelijkheidsgraden 1 en 2 hebben een erg vergelijkbare modal split,  $p > 0.05$ ) en om clusters te vinden met patronen in reisgedrag die niet terug te zien zijn wanneer er alleen wordt gekeken naar de stedelijkheidsgraad of naar andere variabelen gebaseerd op de bevolkingsdichtheid. Omdat er geen geschikte indicatoren gevonden zijn om dit proces te optimaliseren, is door middel van een handmatig proces bepaald welke variabelen deel uitmaken van de uiteindelijke clustersets en uit hoeveel clusters deze clustersets bestaan. Zie Timmerman (2024) voor meer details over dit proces en de clustermethode.

De verschillende clusters die gemaakt zijn in de clusteranalyse hebben vermoedelijk andere demografische karakteristieken. Om een schatting te maken van wat het 'werkelijke' effect is van de ruimtelijke omgeving wordt 'propensity score matching' (PSM) gebruikt (e.g. Liu et al., 2024; Pot et al., 2023). Bij PSM worden observaties (in dit geval reizen) uit twee clusters met vergelijkbare demografische karakteristieken aan elkaar gekoppeld. Observaties die niet gekoppeld kunnen worden, worden verwijderd. Dit resulteert in twee nieuwe clusters met een vergelijkbare populatie, maar met andere ruimtelijke kenmerken, waarvan het reisgedrag vergeleken kan worden (Cao et al., 2009). De observaties worden gekoppeld aan de hand van een 'propensity score'.

Stel, clusters 0 en 1 worden met elkaar vergeleken. De propensity score is de kans dat een observatie behoort tot cluster 1, gebaseerd op de demografische karakteristieken. Deze score wordt berekend voor alle observaties in beide clusters (Liu et al., 2024) en is voor dit onderzoek berekend met logistische regressie (Pot et al., 2023). Om te controleren of PSM gelukt is en de populaties van beide clusters voldoende op elkaar lijken, wordt de 'standard mean difference' (SMD) berekend voor elke demografische eigenschap en voor elk clusterpaar, zie vergelijking 1.

$$SMD = \frac{100(\bar{X}_1 - \bar{X}_0)}{\sqrt{\frac{S_1^2 + S_0^2}{2}}} \quad (1)$$

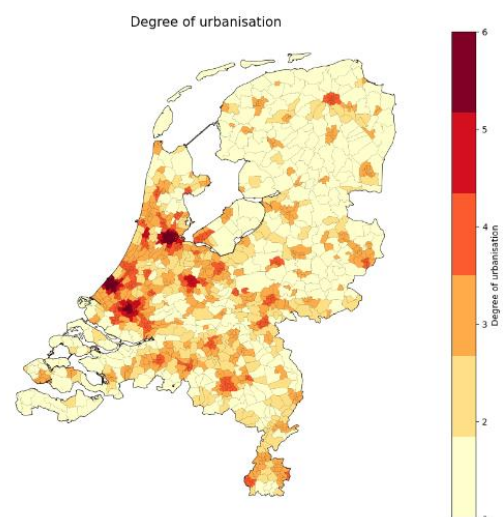
Hierbij zijn  $\bar{X}_0$  en  $\bar{X}_1$  de gemiddelde waarden voor een demografische eigenschap van cluster 0 en 1 en zijn  $S_0$  en  $S_1$  de standaarddeviaties (Liu et al., 2024). Een waarde van  $|SMD| < 10\%$  betekent dat de clusters genoeg op elkaar lijken (Oakes & Johnson, 2006).

Nadat alle observaties gekoppeld zijn, kan het effect van de ruimtelijke omgeving op reisgedrag worden geschat. Dit is het verschil in gemiddeld reisgedrag tussen twee clusters na PSM. Dit wordt ook wel het 'average treatment effect' (ATE) genoemd. Het verschil in gemiddeld reisgedrag voorafgaand aan het matchen wordt het 'observed effect' (OBE) genoemd. Het werkelijke effect van de ruimtelijke omgeving kan ook worden uitgedrukt door te kijken naar de verhouding tussen het verschil in reisgedrag voor en na PSM (de ATE OBE ratio) (Liu et al., 2024). Dit is met de aanname dat mensen met vergelijkbare externe karakteristieken gemiddeld genomen vergelijkbare reispreferenties hebben. Om te testen of de resultaten statistisch gezien significant zijn, is een t-test gebruikt. Een nadeel van deze methode is dat slechts twee clusters tegelijk met elkaar vergeleken kunnen worden.

### 3. Literatuur & LMS documentatie

#### 3.1 Stedelijkheidsgraad

De stedelijkheidsgraad is een maat van hoe stedelijk een gebied is. Het LMS deelt Nederland op in 1406 zones en voor elke zone wordt de stedelijkheidsgraad berekend, zie figuur 1. Er zijn zes stedelijkheidsgraden. De stedelijkheidsgraad wordt bepaald aan de hand van de bevolkingsdichtheid van de zone zelf en die van de omringende zones (RWS WVL, 2021). Dit is de eigen definitie van het LMS. Er is geen universele definitie van de stedelijkheidsgraad.



Figuur 1: Stedelijkheidsgraden volgens het LMS.

#### 3.2 D-variabelen

In de literatuur worden vaak de zogenaamde D-variabelen gebruikt om de ruimtelijke omgeving uit te drukken (Ewing & Cervero, 2010; Kent et al., 2023):

- *Density* (dichtheid): Een variabele per oppervlakte eenheid, e.g. bevolkingsdichtheid, banendichtheid.
- *Diversity* (diversiteit): Variabelen voor de diversiteit van landgebruik, e.g. percentage land dat wordt gebruikt voor de landbouw.
- *Design* (ontwerp): De karakteristieken van het wegennetwerk, e.g. dichtheid van het wegennetwerk.
- *Destination accessibility* (bereikbaarheid van bestemmingen): een maat voor hoe makkelijk het is om verschillende bestemmingen te bereiken, e.g. aantal supermarkten in de nabije omgeving.
- *Distance to transit* (afstand tot openbaar vervoer (OV)): De bereikbaarheid en kwaliteit van het OV-netwerk, e.g. afstand tot de bushalte, aantal treinen per uur.
- *Demand management* (sturing van vraag): Maatregelen om bepaald reisgedrag te stimuleren of te ontmoedigen, e.g. parkeertarief.

Een zevende D-variabele is de Demografie (Ewing & Cervero, 2010). Hoewel dit geen aspect is van de ruimtelijke omgeving, is het wel een belangrijke variabele om rekening mee te houden. Het effect van de ruimtelijke omgeving wordt meestal minder wanneer de demografie wordt meegenomen in de analyse (Cao et al., 2009).

### 3.3 LMS-documentatie

Het LMS bestaat uit verschillende samenhangende modules. De bevolkingsmodule (D4) en de frequentiebepaling en nutbepaling van de groeifactor module (D7) zijn geanalyseerd om uit te zoeken hoe de ruimtelijke omgeving wordt meegenomen in het LMS. De stedelijkheidsgraad en andere ruimtelijke factoren komen voornamelijk voor in deze modules. Bij het Nutbepalingsmodel wordt het nut van elke vervoersmiddel-bestemming-dagdeel alternatief bepaald. Dit deel van het LMS is dus het meest relevant voor de modal split (RWS WVL, 2021).

De D-variabelen zijn tot op zekere hoogte meegenomen in het LMS. Uit deze analyse bleek dat het LMS (naast uiteraard demografische karakteristieken) vooral *Density* en *Distance to transit* variabelen meeneemt in de modellen. Met name de stedelijkheidsgraad (die ook een *Density* variabele is) komt vaak terug in het LMS. De *Distance to transit* variabelen zijn hoofdzakelijk gerelateerd aan het modelleren van trein gebruik (inclusief voor- en natransport), en niet aan het gebruik van bus, tram en metro (BTM) als hoofdmodaliteit. *Demand management* komt een aantal keer voor door bijvoorbeeld variabelen voor parkeertarief en -vergunningen. Er zijn enkele variabelen die gerekend kunnen worden onder *Diversity* of *Destination accessibility*. Deze variabelen zijn voornamelijk gerelateerd aan banen (e.g. aandeel of aantal banen in verschillende sectoren) en niet aan andere aspecten van de ruimtelijke omgeving (e.g. de diversiteit van (niet baan gerelateerd) landgebruik). Er zijn geen expliciete variabelen gerelateerd aan *Design*, maar dit is vermoedelijk impliciet meegenomen bij het bepalen van de bereikbaarheid van herkomst-bestemming (HB) paren.

De meeste D-variabelen in het LMS zijn gerelateerd aan het modelleren van trein, autobestuurder en autopassagier. De D-variabelen voor BTM, fiets en lopen zijn zeer beperkt. Op basis van deze bevindingen wordt verwacht dat het LMS relatief goed is in het modelleren van treingebruik in verschillende gebieden en dat modaliteiten die sterk

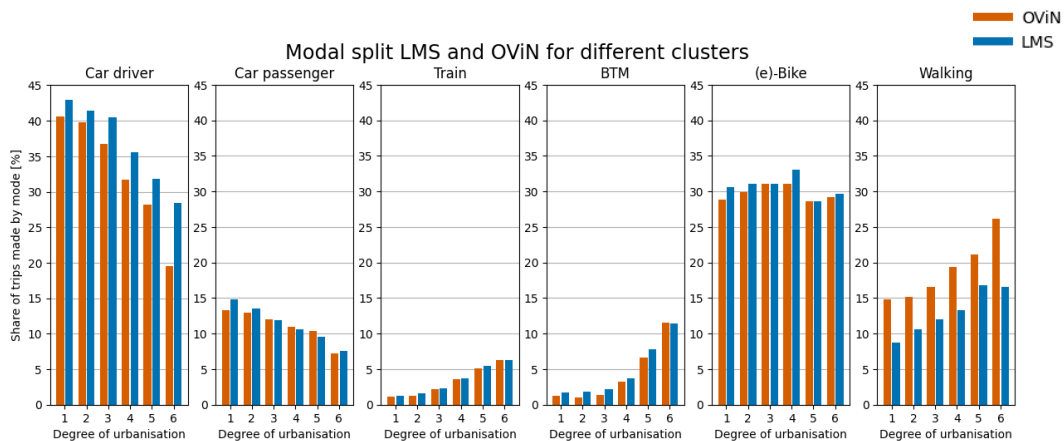
afhankelijk zijn van de bevolkingsdichtheid beter zullen worden gemodelleerd dan modaliteiten die afhankelijker zijn van andere type variabelen. Modaliteiten waarvan het gebruik minder sterk gecorreleerd is met de bevolkingsdichtheid zullen vermoedelijk minder goed worden gemodelleerd.

## 4. Modal split analyse

### 4.1 Onderzoekende analyse

#### Modal split per stedelijkheidsgraad

Eerst is er gekeken naar de gemiddelde modal split op basis van de stedelijkheidsgraad, zie figuur 2. Zowel OViN en LMS volgen grotendeels dezelfde trend: autobestuurder neemt af bij een hogere stedelijkheidsgraad (d.w.z. een hogere bevolkingsdichtheid) en trein, BTM en lopen neemt toe. Fietsgebruik lijkt redelijk gelijk te blijven met een kleine dip voor de hoogste twee stedelijkheidsgraden. Vergelijkbare resultaten voor OViN en LMS zijn logisch, omdat het OViN is gebruikt voor het kalibreren van het LMS en de stedelijkheidsgraad een belangrijke rol speelt in het LMS.



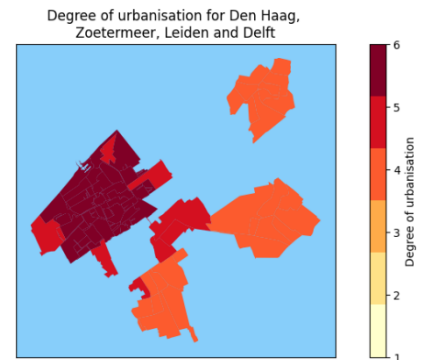
Figuur 2: Modal split per stedelijkheidsgraad volgens OViN en LMS. De modal split is bepaald aan de hand van het aantal vertrekkende trips in een zone.

Er zijn echter ook verschillen op te merken tussen OViN en LMS. Het aandeel autobestuurders wordt overschat in het LMS, met name in de hoog stedelijke gebieden. Zo schat het LMS dat in zones met een stedelijkheidsgraad van 6 gemiddeld 28% van de reizen wordt gemaakt met de auto. Volgens OViN zijn dat minder dan 20% van de reizen. Het omgekeerde is te zien wanneer er wordt gekeken naar het aantal reizen dat lopend worden gemaakt. Voor alle stedelijkheidsgraden wordt lopen onderschat door het LMS.

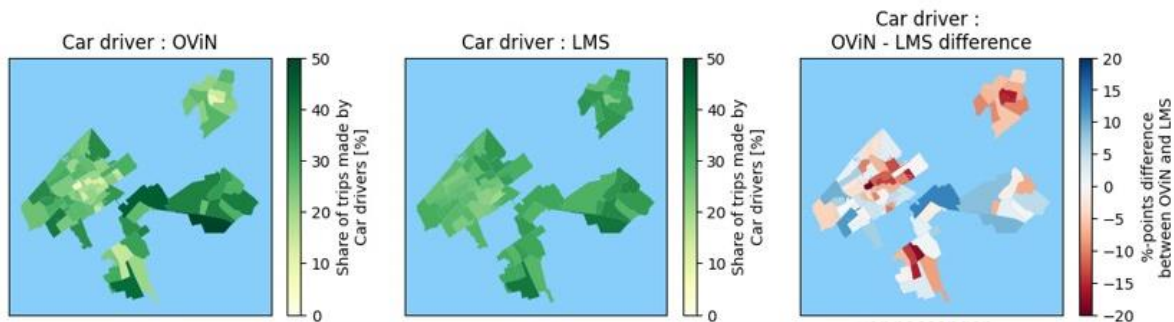
#### Modal split Den Haag, Leiden, Zoetermeer en Delft

Vervolgens is er ingezoomd op verschillende gebieden om te kijken hoe de modal split verschilt op zoneniveau binnen een stad. Hiervoor is uitgebreid gekeken naar Amsterdam, Zeeland en Den Haag, Leiden, Zoetermeer en Delft. Dit laatste gebied zal verder worden toegelicht voor de modaliteiten autobestuurder en fietser.

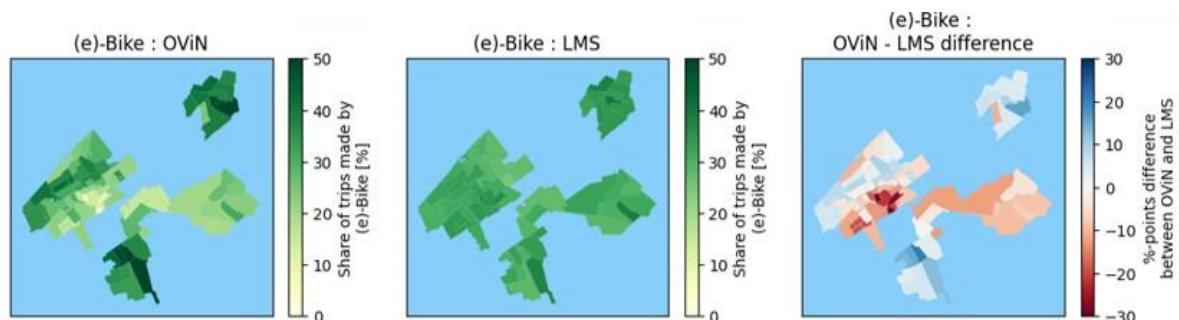
Figuur 3 laat de stedelijkheidsgraden zien voor Den Haag, Leiden, Zoetermeer en Delft. Den Haag heeft grotendeels een stedelijkheidsgraad van 6 (en 5 in de buitenste wijken) en Leiden, Zoetermeer en Delft hebben grotendeels een stedelijkheidsgraad van 4. De figuren 4 en 5 laten het aandeel autobestuurder- en fietsgebruik zien in de steden. Het linker figuur geeft dit weer voor OViN, het middelste figuur voor LMS en het rechter figuur geeft het verschil weer voor OViN en LMS. Hierbij betekent rood dat het LMS het gebruik van die modaliteit overschat en blauw dat het LMS het gebruik onderschat.



Figuur 3: Stedelijkheidsgraad Den Haag, Leiden, Zoetermeer, Delft



Figuur 4: Aandeel autobestuurder in Den Haag, Leiden, Zoetermeer en Delft volgens OViN en LMS en het verschil.



Figuur 5: Fietsgebruik in Den Haag, Leiden, Zoetermeer en Delft voor OViN en LMS en het verschil.

Deze figuren laten zien dat volgens OViN het aandeel reizen door een autobestuurder in Zoetermeer gemiddeld genomen hoger is dan in de andere drie steden. Vooral in de centra van Den Haag, Leiden en Delft wordt er duidelijk minder autogereden dan in Zoetermeer. Voor fietsgebruik is het tegenovergestelde te zien. In de steden Delft en Leiden wordt er veel gefietst, terwijl dit in Zoetermeer en delen van Den Haag soms wel de helft minder is.

Het LMS laat echter hele andere patronen zien. Het valt allereerst op dat het LMS het fietsgebruik meer lijkt 'uit te smeren'. Waar in OViN duidelijk zones met hoog of laag modaliteitsgebruik te zien zijn, zit er volgens het LMS veel minder variatie in het modaliteitsgebruik. Er is wel een beetje variatie te zien (e.g. lager autogebruik in het centrum van den haag en een hoger fietspercentage in Leiden), maar de verschillen zijn veel kleiner dan te zien is in OViN.

De rechter figuren laten duidelijk zien dat autogebruik in Zoetermeer en buitenwijken van Den Haag voornamelijk onderschat wordt door het LMS, terwijl autogebruik in

Leiden, Delft en het centrum van Den Haag wordt overschat. Voor fiets is deels het tegenovergestelde te zien. Fietsgebruik in Zoetermeer en het centrum van den Haag wordt overschat, terwijl het in de overige gebieden wordt onderschat.

Dit voorbeeld laat duidelijk zien dat gebieden met dezelfde stedelijkheidsgraad een hele verschillende modal split kunnen hebben en dat er dus meer variabelen dan alleen de bevolkingsdichtheid nodig zijn om goed onderscheid te kunnen maken tussen steden met verschillend reisgedrag. Daarnaast laat dit voorbeeld zien dat het LMS op het moment niet in staat is om de verschillen in reisgedrag tussen steden met dezelfde stedelijkheidsgraad op een juiste manier te modelleren voor alle modaliteiten. Deze patronen zijn ook te zien op andere plekken in Nederland en zijn dus niet alleen een probleem voor de omgeving van Den Haag. Op dit punt is dus verbetering mogelijk in het LMS.

De belangrijkste bevindingen na het analyseren van alle modaliteiten voor verschillende gebieden staan hieronder samengevat:

- Zowel op nationaal als regionaal niveau, wordt het aandeel autobestuurder overschat en het aandeel lopen onderschat. In de centra van steden wordt het autogebruik het sterkst overschat door het LMS.
- Wanneer er alleen gekeken wordt naar de stedelijkheidsgraden op nationaal niveau, lijkt fietsgebruik redelijk constant te zijn door het hele land en goed gemodelleerd te worden door het LMS. OViN laat echter zien dat fietsgebruik erg varieert tussen en binnen steden. Dit wordt op het moment niet goed gemodelleerd door het LMS. Dit is te zien zowel in grote steden (e.g. Amsterdam) als in meer landelijke gebieden (e.g. Zeeland).
- Treingebruik wordt relatief goed gemodelleerd. Het absolute aandeel treinreizen is niet altijd accuraat, maar het LMS is over het algemeen goed in staat om zones met veel of weinig trein gebruik te identificeren. BTM bevat meer verschillen tussen OViN en LMS, maar vaak worden grotere gebieden met veel of weinig BTM gebruik wel goed geïdentificeerd. Hierbij is het belangrijk om te beseffen dat OViN ook maar slechts een steekproef is, en hierdoor minder accuraat kan zijn bij lager gebruik van een modaliteit. Het aandeel trein en BTM ligt in veel gebieden in Nederland onder de 5%.

#### *4.2 Clusteranalyse*

Voor de clusteranalyse zijn er 48 variabelen verzameld aan de hand van de D-variabelen die volgden uit de literatuur. Hieruit blijkt dat alle Density variabelen sterk gecorreleerd zijn met alle modaliteiten, behalve fietsgebruik. Op basis van deze informatie is het dus logisch om Density variabelen te gebruiken in een transportmodel als representatie van de ruimtelijke omgeving.

Door middel van een handmatig proces zijn deze 48 variabelen geclusterd. Hierbij moesten de clusters voldoen aan verschillende criteria (bijv. het aantal clusters mag niet te groot zijn om overfitten te voorkomen en niet te weinig om genoeg differentiatie te laten zien; Leiden en Delft moeten deel zijn van een ander cluster dan Zoetermeer). Er is geprobeerd om met zo weinig mogelijk variabelen, clustersets te maken die interessante patronen in reisgedrag blootleggen. Uiteindelijk zijn er twee verschillende clustersets gemaakt waarvan er één verder zal worden toegelicht. Zie Timmerman (2024) voor de andere clusterset.

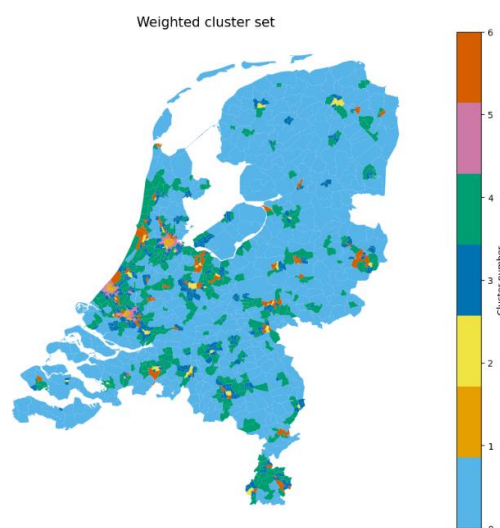


Figuur 6 laat de clusters zien op de kaart van Nederland en tabel 1 geeft een overzicht van de variabelen die gebruikt zijn om deze clusterset te maken.

Tabel 1: Lijst met alle variabelen die gebruikt zijn voor het maken van de gewogen clusterset.

D-variabele	Variabele	Gewicht
Density	Bevolkingsdichtheid	1
	Bevolkingsdichtheid, inclusief omringende zones	1
	Banendichtheid	1
Diversity	Banendichtheid, inclusief omringende zones	1
	Aandeel land dat wordt gebruikt voor diensten (e.g. winkels, restaurants, gemeentehuis)	2
Design	Aandeel huizen gebouwd voor 1945	2
	Wegendichtheid	4
Destination accessibility	Gemiddelde minimale afstand naar nuttige bestemmingen (e.g. supermarkt, ziekenhuis, scholen, recreatie)	4
Distance to transit	Aantal bushaltes binnen 2.5 km	2
	Aantal tram-/ metrohaltes binnen 2.5 km	2
Demand management	Parkeertarief	4

De verschillende variabelen zijn gebaseerd op verschillende databronnen (RWS WVL, 2020; CBS, 2022; CBS & Kadaster, 2019; CBS & ESRI Nederland, 2019; RWS, 2022; University of Groningen Geodienst, 2021) en zijn waar nodig geaggregeerd van postcode 4 (PC4) niveau naar de LMS-zones. Elke variabele in deze clusterset heeft een gewicht, waardoor de verschillende D-variabelen elk even zwaar meewegen. Figuur 7 laat de verdeling van elke variabele zien over de verschillende clusters en figuur 8 laat de modal split zien voor de clusters.

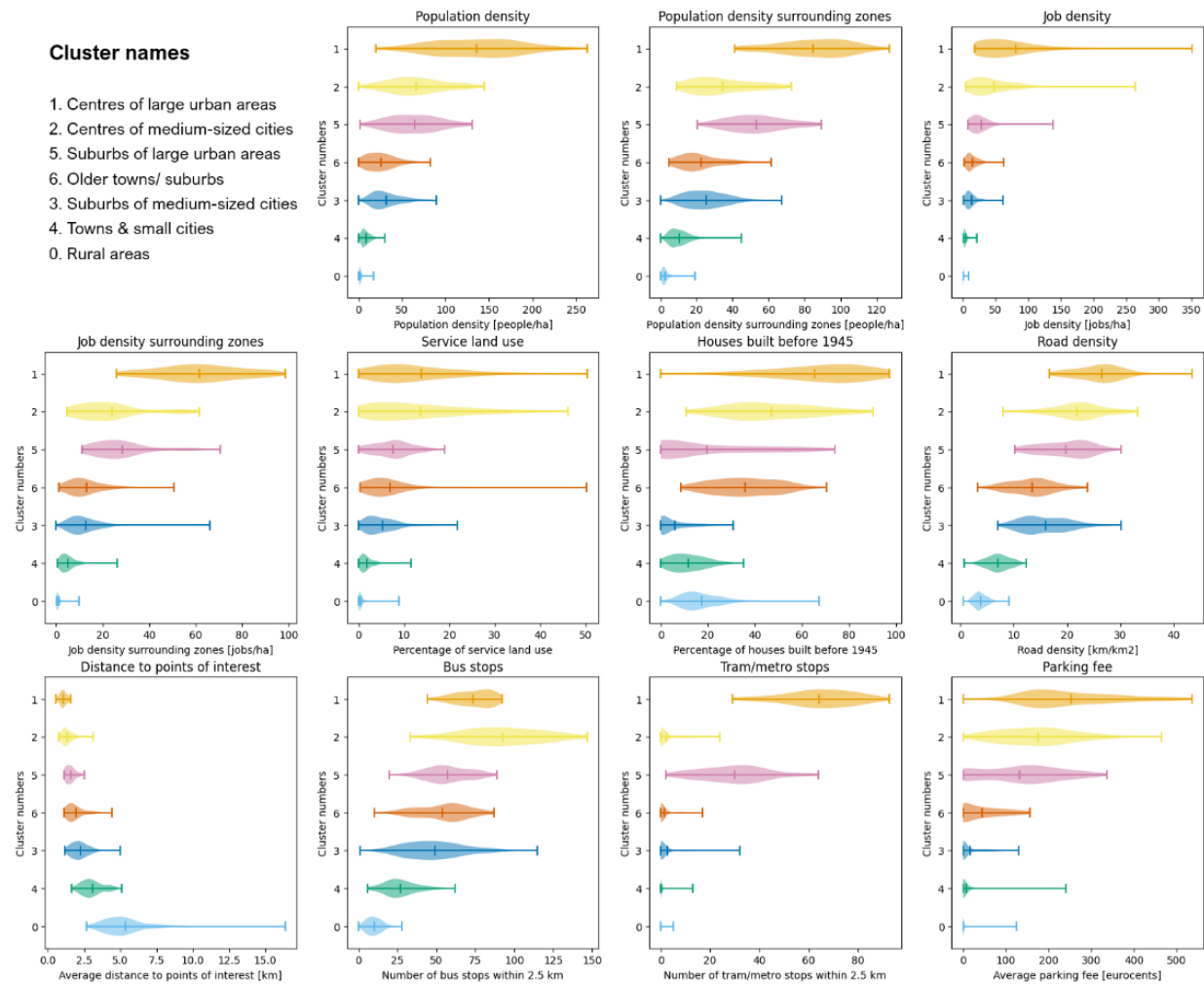


Figuur 6: Kaart van Nederland met de verdeling van de verschillende clusters voor de gewogen clusterset.

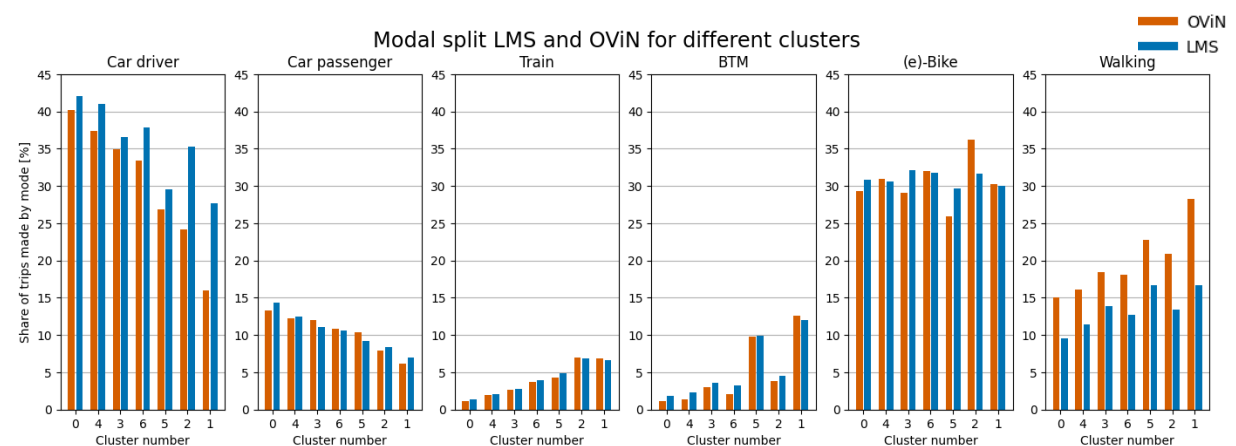
De meest opvallende clusters in deze clusterset zijn de centra van de middelgrote steden (cluster 2) en de buitenwijken van grote hoog-stedelijke gebieden (cluster 5). Vooral de middelgrote steden worden niet goed gemodelleerd door het LMS. Hier zal verder op worden ingegaan.

Beide clusters hebben een vergelijkbare bevolkingsdichtheid, maar wanneer de bevolkingsdichtheid van de omringende zones wordt meegenomen, is de dichtheid hoger in de buitenwijken. Dit komt omdat de buitenwijken de gebieden zijn rond de grootste steden (Amsterdam, Rotterdam en Den Haag), terwijl de centra van de middelgrote steden voornamelijk worden omringd door gebieden met een lagere dichtheid. De stedelijkheidsgraad is gebaseerd op de bevolkingsdichtheid, inclusief de omringende steden. Hierdoor hebben de buitenwijken voornamelijk een stedelijkheidsgraad van 5 of 6, terwijl de middelgrote steden meestal een stedelijkheidsgraad hebben van 4 of 5. Andere opvallende verschillen tussen deze clusters zijn het aandeel huizen dat is gebouwd voor 1945, wat veel hoger is in de middelgrote steden; het aandeel land dat

wordt gebruikt voor diensten; en het OV-netwerk. De buitenwijken hebben veel tram- en/of metrohaltes, terwijl de middelgrote steden dit nauwelijks hebben. Wel hebben zij meer bushaltes.



Figuur 7: Verdeling van de verschillende ruimtelijke variabelen over de clusters voor de gewogen cluster-set.



Figuur 8: Modal split voor de cluster-set volgens OViN en LMS. De modal split is bepaald aan de hand van het aantal vertrekkende trips in een zone.

Op basis van de stedelijkheidsgraad zou verwacht worden dat autogebruik in de buitenwijken lager is dan in de middelgrote steden, OV gebruik en lopen hoger, en fietsen vergelijkbaar. Dit is echter niet waar voor elke modaliteit.

Volgens OViN is het gebruik van de auto hoger in de buitenwijken en het treingebruik lager. BTM is daarentegen wel significant hoger in de buitenwijken en ook wordt er meer gelopen. Ten slotte is er een groot verschil te zien in het aandeel fietsgebruik. In de middelgrote steden ligt het fietsgebruik met wel 10 %-punten hoger dan in de buitenwijken.

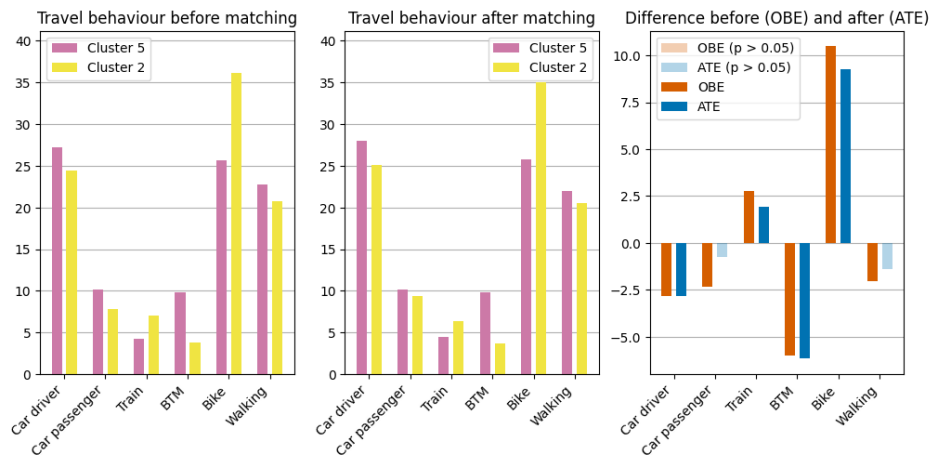
Deze verschillen worden vaak niet goed gemodelleerd door het LMS. Zo 'denkt' het LMS dat het aandeel autobestuurder significant hoger is in de middelgrote steden dan in de buitenwijken. Ook verwacht het LMS dat fietsgebruik vergelijkbaar is. De andere modaliteiten worden wel redelijk gemodelleerd (al wordt lopen nog steeds onderschat).

De belangrijkste bevindingen na het analyseren van alle clusters en de twee verschillende clustersets (zie Timmerman, 2024), zijn dat over het algemeen het LMS de grootste fouten maakt in de modal split bij clusters die niet de aan standaard 'stedelijkheidsgraadtrend' voldoen. Met deze trend wordt bedoeld dat bij een toenemende stedelijkheidsgraad autogebruik afneemt, OV en lopen toeneemt en fietsgebruik redelijk gelijk blijft. De grootste verschillen zijn te zien bij autobestuurder en fietsgebruik. Het LMS is in staat om de trends uit OViN voor de andere modaliteiten redelijk te volgen. Daarnaast laten deze clusters zien dat door meer aspecten van de ruimtelijke omgeving toe te voegen, nieuwe gebieden zichtbaar worden met een eigen reispatroon, die niet zichtbaar zijn wanneer er alleen wordt gekeken naar de stedelijkheidsgraad. Alle gebruikte data voor dit onderzoek is openbaar toegankelijk. Het is dus niet nodig om nieuwe data te verzamelen om deze reispatronen bloot te leggen.

#### *4.3 Propensity score matching*

Verschillende mensen hebben verschillende reisvoorkeuren en wonen op verschillende plekken. Hierdoor zou het kunnen dat verschil in reisgedrag tussen clusters veroorzaakt wordt door verschillen in de demografie en niet door verschil in de ruimtelijke omgeving. In de vorige analyses is er nog geen rekening gehouden met de demografie, maar met PSM is het mogelijk om hiervoor te corrigeren. PSM is uitgevoerd voor elk clusterpaar van beide clustersets en voor alle stedelijkheidsgraadparen. De gebruikte observaties zijn alle reizen volgens OViN en de bijbehorende demografische karakteristieken. Het was niet mogelijk om dezelfde analyse te doen voor LMS-data, omdat bij de beschikbare data geen demografische karakteristieken bekend waren voor elke reis. Er zijn 10 verschillende type karakteristieken meegenomen: leeftijd, geslacht, huishoudinkomen, huishoud grootte, huishoud type (e.g. 1 persoon, 2 oudergezin), maatschappelijke participatie (e.g. fulltime werker, student), type educatie, aantal auto's per huishouden, rijbewijsbezit en bezit van studenten OV. Voor alle clusterparen en demografische karakteristieken geldt:  $|SMD| < 10\%$ , wat erop wijst dat PSM gelukt is. Zie Timmerman (2024) voor meer details.

Figuur 9 laat een voorbeeld zien voor de clusters 2 en 5 (de middelgrote steden en de buitenwijken). Het linker figuur laat de modal split zien volgens OViN voor PSM, het middelste figuur laat de modal split zien na PSM en het rechter figuur laat verschil in (gemiddelde) modal split zien voor en na PSM. Een lichtblauw balkje in het rechter figuur betekent dat de verschillen in modal split niet meer significant zijn na PSM (i.e. de t-test resulteerde in  $p > 0.05$ ).



Figuur 9: Modal split voor PSM, na PSM en de verschillen in modal split voor de clusters 2 (middelgrote steden) en 5 (buitenwijken van hoogstedelijke gebieden).

Hier is te zien dat het verschil tussen het aandeel autobestuurders voor PSM (OBE) en na PSM (ATE) ongeveer even groot is. Aan de hand van deze resultaten is de ruimtelijke omgeving dus verantwoordelijk te zijn voor ongeveer 100% van de verschillen in het aandeel autobestuurder tussen deze clusters. Verschillen in demografische karakteristieken tussen deze clusters spelen dus een verwaarloosbare rol. Aan de andere kant is te zien dat het verschil in het aandeel autopassagiers en looptaandeel na PSM is weggevallen ( $p < 0.05$ ). In deze gevallen worden de verschillen tussen de clusters dus voornamelijk veroorzaakt wordt door verschil in demografische karakteristieken van beide populaties. De ruimtelijke omgeving speelt hier een beperkte rol. Bij trein, BTM en fietsen is de impact van de ruimtelijke omgeving weer heel hoog.

Deze analyse is gedaan voor alle clusterparen van beide clustersets en voor de stedelijkheidsgraden. De resultaten van PSM laten zien dat de ruimtelijke omgeving een grote rol speelt in het verklaren van de verschillen in modal split tussen de clusters. In de meeste gevallen is de ruimtelijke omgeving verantwoordelijk voor meer dan 50% van de verschillen (ATE OBE ratio  $> 0.5$ ) en spelen dus de verschillen tussen demografische karakteristieken een kleinere rol. Dit laat zien dat het belangrijk is om variabelen van de ruimtelijke omgeving mee te nemen in een transportmodel. Verschillen in reisgedrag tussen gebieden zijn dus niet alleen te verklaren door verschil in de demografie. Gemiddeld genomen heeft de ruimtelijke omgeving de grootste invloed op BTM, fiets en lopen (ATE OBE ratio  $> 0.8$ ). Voor die modaliteiten kan het extra relevant zijn om meer D-variabelen toe te voegen aan een transportmodel.

Daarnaast blijkt dat het effect van de ruimtelijke omgeving (i.e. de ATE OBE ratio) gemiddeld groter bij het vergelijken van de modal split van de cluster paren, dan bij het vergelijken van de stedelijkheidsgraden met elkaar. Dit kan betekenen dat de clustersets die gemaakt zijn voor dit onderzoek beter in staat zijn om gebieden aan te wijzen die verschillen tonen in reisgedrag door verschillen in de ruimtelijke omgeving, dan de stedelijkheidsgraden.

Ten slotte is er niet één duidelijk antwoord te geven op de vraag wat het effect van de ruimtelijke omgeving is op de modal split. Verschillende clusterparen hebben andere ATE OBE ratios en deze ratios zijn ook anders voor de verschillende modaliteiten. Zo is het effect van de ruimtelijke omgeving gemiddeld genomen meer dan 50%, maar zijn er ook clusterparen waarbij er geen significante verschillen waren ( $p < 0.05$ ) voor of na het matchen voor een of meerdere modaliteiten. Het daadwerkelijke effect van de ruimtelijke

omgeving is dus erg afhankelijk van de gebieden die vergeleken worden. Bijvoorbeeld, het effect van de ruimtelijke omgeving op BTM is gemiddeld groter wanneer een van de clusters een uitgebreid tram/metro netwerk heeft (i.e. clusters 1 en 5).

## **5. Conclusies en aanbevelingen**

### *5.1 Conclusies*

Dit onderzoek heeft meer duidelijkheid gegeven over het effect van de stedelijkheidsgraad op reisgedrag. Ook hebben we onderzocht hoe verschil in reisgedrag tussen verschillende gebieden beter kan worden meegenomen in transportmodellen. De stedelijkheidsgraad is een variabele op basis van de bevolkingsdichtheid. Density variabelen zijn sterk gecorreleerd met reisgedrag, wat de stedelijkheidsgraad een logische keuze maakt om de ruimtelijke omgeving te representeren. Echter, alleen de stedelijkheidsgraad is niet genoeg. Door alleen te kijken naar de bevolkingsdichtheid, gaan belangrijke nuances tussen gebieden verloren. Zo vallen bijvoorbeeld de centra van middelgrote steden weg. Ondanks een lagere bevolkingsdichtheid dan bijvoorbeeld de buitenwijken, kenmerken deze zones zich door een lager auto- en BTM-aandeel en een hoger fiets- en treinaandeel dan in de buitenwijken. Dit is niet te zien wanneer er alleen gekeken wordt naar de stedelijkheidsgraad. Daarnaast kunnen zones met dezelfde stedelijkheidsgraad een heel andere modal split hebben en is er veel verschil in reisgedrag te zien tussen verschillende steden. Zo hebben Delft, Zoetermeer en Leiden dezelfde stedelijkheidsgraad, maar wordt er in Zoetermeer meer autogereden en minder gefietst. De stedelijkheidsgraad doet dus onvoldoende recht aan het diverse mobiliteitsgedrag in Nederland.

Het LMS is over het algemeen in staat om de verschillende trends die in OViN te zien zijn voor autopassagiers, trein en BTM goed mee te nemen voor verschillende clusters. Het aandeel lopen wordt significant onderschat, maar het LMS is in staat om clusters met hoger en lager gebruik correct te identificeren.

Voor autobestuurder en fiets worden de volgende belangrijke trends echter niet goed gemodelleerd door het LMS. Het aantal gemodelleerde autobestuurder trips wordt in het LMS sterk beïnvloed door de stedelijkheidsgraad. Dit is grotendeels een evenredig verband. In dit paper is echter aangetoond dat er gebieden in Nederland zijn waar dit evenredige verband niet opgaat. In die gebieden modelleert het LMS autogebruik dus niet goed. Verder modelleert het LMS een vergelijkbaar aandeel voor fietsgebruik door heel het land. Volgens OViN zijn er echter grote verschillen te zien tussen gebieden.

De verschillen in de modal split tussen verschillende clusters en tussen de stedelijkheidsgraden worden veroorzaakt door zowel verschil in demografische karakteristieken als door verschil in de ruimtelijke omgeving. Wanneer er wordt gecorrigeerd voor verschillen in de bevolking door middel van propensity score matching, blijkt dat de ruimtelijke omgeving gemiddeld genomen de grootste rol speelt. Het daadwerkelijke effect van de ruimtelijke omgeving is echter sterk afhankelijk van welke clusters met elkaar vergeleken worden. Dit betekent dat het essentieel is om ruimtelijke kenmerken te gebruiken in een transportmodel. Niet alle verschillen in modaliteitskeuze tussen gebieden kunnen worden gevangen door demografische kenmerken.

Samengevat, de karakteristieken van de ruimtelijke omgeving spelen een grote rol in de verschillen in modal split tussen verschillende gebieden. In het LMS is er voornamelijk voor autobestuurder en fietsgebruik verbetering mogelijk. Een mogelijke optie om het LMS te verbeteren is door het vervangen van de stedelijkheidsgraad met een nieuwe variabele, zoals (een verbeterde versie van) de clusters die gemaakt zijn voor dit onderzoek. Een andere, makkelijkere, optie die vermoedelijk tot een verbetering zal leiden van het LMS, is het toevoegen van extra D-variabelen aan het model. Doordat meer aspecten van de ruimtelijke omgeving worden meegenomen, wordt er beter onderscheid gemaakt tussen gebieden met verschillen in reisgedrag.

## 5.2 Aanbevelingen

Het LMS wordt momenteel gebruikt door organisaties zoals het Ministerie van Infrastructuur en Waterstaat en ProRail voor het maken van beleid. Het is mogelijk om in het LMS verschillende scenario's te runnen om te kijken wat het effect van beleid is op reisgedrag (Hofman, 2017). Dit onderzoek toont aan dat het LMS beperkt geschikt is voor het testen van beleid met betrekking tot de ruimtelijke omgeving. Het LMS overschat, bijvoorbeeld, sterk het gebruik van de auto in de centra van steden. Wanneer deze steden autoluw beleid willen ontwikkelen, zal het LMS minder geschikt zijn om de effecten van dit beleid te testen. Op het moment is het LMS al niet goed in staat om het effect van de ruimtelijke omgeving op autogebruik mee te nemen. Hierdoor zullen de voorspelde effecten van het beleid extra grote onzekerheden bevatten. Uit de PSM volgt dat de ruimtelijke omgeving gemiddeld verantwoordelijk is voor meer dan 50% van de verschillen in modal split tussen clusters. Dat betekent echter dat ook een significant deel van de verschillen veroorzaakt wordt door verschillen in de demografie. In dit onderzoek is niet gekeken naar de modellering van de populatie. Mochten in deze modelering grote verschillen aanwezig zijn met de werkelijkheid, zou dit ook een deel van de verschillen tussen OVIN en LMS kunnen verklaren. Ten slotte, in dit onderzoek is er alleen gekeken naar het effect van de ruimtelijke omgeving op de modal split van de herkomstzone. Om een completer beeld te krijgen van het effect van de ruimtelijke omgeving, zal ook moeten worden gekeken naar andere aspecten van reisgedrag, zoals bestemmingskeuze, reisafstand, reistijd, reisfrequentie en dagdeelkeuze en naar de karakteristieken van de bestemmingszone.

## Literatuur

- Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009). Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews*, 29(3), 359–395. <https://doi.org/10.1080/01441640802539195>
- Centraal Bureau voor de Statistiek. (2022). *Bodemgebruik, wijk- en buurtcijfers 2017*. [https://opendata.cbs.nl/portal.html?\\_la=nl&\\_catalog=CBS&tableId=85217NED&theme=30](https://opendata.cbs.nl/portal.html?_la=nl&_catalog=CBS&tableId=85217NED&theme=30)
- Centraal Bureau voor de Statistiek & ESRI Nederland. (2019). *Statistische gegevens per vierkant en postcode 2019*. <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode>

- Centraal Bureau voor de Statistiek & Kadaster. (2019). *Wijk- en Buurtkaart 2017*.  
<https://www.cbs.nl/nlnl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2017>
- Centraal Bureau voor de Statistiek & Rijkswaterstaat. (2014). *Onderzoek Verplaatsingen in Nederland 2013 - OViN 2013*. <https://doi.org/10.17026/dans-x9h-dsdg>
- Centraal Bureau voor de Statistiek & Rijkswaterstaat. (2015). *Onderzoek Verplaatsingen in Nederland 2014 - OViN 2014*. <https://doi.org/10.17026/dans-x95-5p7y>
- Centraal Bureau voor de Statistiek & Rijkswaterstaat. (2017a). *Onderzoek Verplaatsingen in Nederland 2015 - OViN 2015 versie 2.0*. <https://doi.org/10.17026/dans-z2v-c39p>
- Centraal Bureau voor de Statistiek & Rijkswaterstaat. (2017b). *Onderzoek Verplaatsingen in Nederland 2016 - OViN 2016*. <https://doi.org/10.17026/dans-293-wvf7>
- Centraal Bureau voor de Statistiek & Rijkswaterstaat. (2017c). *Onderzoek Verplaatsingen in Nederland 2017 - OViN 2017*. <https://doi.org/10.17026/dans-xxt-9d28>
- Ewing, R., & Certero, R. (2010). Travel and the Built Environment. *Journal Of The American Planning Association*, 76(3), 265–294.  
<https://doi.org/10.1080/01944361003766766>
- Hofman, F. (2017). *Het Landelijk Model Systeem*. Ministerie van Infrastructuur en Waterstaat, Rijkswaterstaat Water, Verkeer en Leefomgeving (RWS, WVL).  
[https://open.rijkswaterstaat.nl/publish/pages/34566/verkeer\\_en\\_vervoer\\_het\\_landelijk\\_model\\_systeem.pdf](https://open.rijkswaterstaat.nl/publish/pages/34566/verkeer_en_vervoer_het_landelijk_model_systeem.pdf)
- Kent, J. L., Crane, M., Waidyatillake, N., Stevenson, M., & Pearson, L. (2023). Urban form and physical activity through transport: a review based on the d-variable framework. *Transport Reviews*, 43(4), 726–754.  
<https://doi.org/10.1080/01441647.2023.2165575>
- Liu, J., Xiao, L., & Wang, B. (2024). The varying effects of residential built environment on travel behavior of internal migrants and locals. *Travel Behaviour And Society*, 34, 100692. <https://doi.org/10.1016/j.tbs.2023.100692>
- Oakes, J. M., & Johnson, P. J. (2006). Propensity score matching for social epidemiology. *Methods in social epidemiology*, 1, 370–393.
- Pot, F. J., Koster, S., & Tillema, T. (2023). Perceived accessibility and residential self-selection in the Netherlands. *Journal Of Transport Geography*, 108, 103555.  
<https://doi.org/10.1016/j.jtrangeo.2023.103555>
- Rijkswaterstaat. (2022). *Nationaal Wegenbestand - Wegvakken 2018-10-1*.  
[https://downloads.rijkswaterstaatdata.nl/nwb-wegen/geogegevens/shapefile/Nederland\\_totaal/01-10-2018/Wegvakken/](https://downloads.rijkswaterstaatdata.nl/nwb-wegen/geogegevens/shapefile/Nederland_totaal/01-10-2018/Wegvakken/)
- Rijkswaterstaat, Water, Verkeer en Leefomgeving. (2018). *SES HB-matrices LMS*.
- Rijkswaterstaat, Water, Verkeer en Leefomgeving. (2020). *Sociaaleconomische gegevens en shapefile LMS 2018*.
- Rijkswaterstaat, Water, Verkeer en Leefomgeving. (2021). *Technische Documentatie Groeimodel 4*.
- Timmerman, J.D. (2024). Degrees of urbanisation in transport modelling: Analysing the impact of the spatial environment on travel behaviour using cluster analysis and propensity score matching. *TU Delft MSc Thesis*.  
<https://repository.tudelft.nl/record/uuid:1345133b-69eb-4166-8ea2-12e09fcb9226>
- University of Groningen Geodienst. (2021). *Openbaar vervoer Nederland - Haltes*.  
<https://hub.arcgis.com/datasets/RUG::openbaar-vervoer-nederland/about?layer=0>