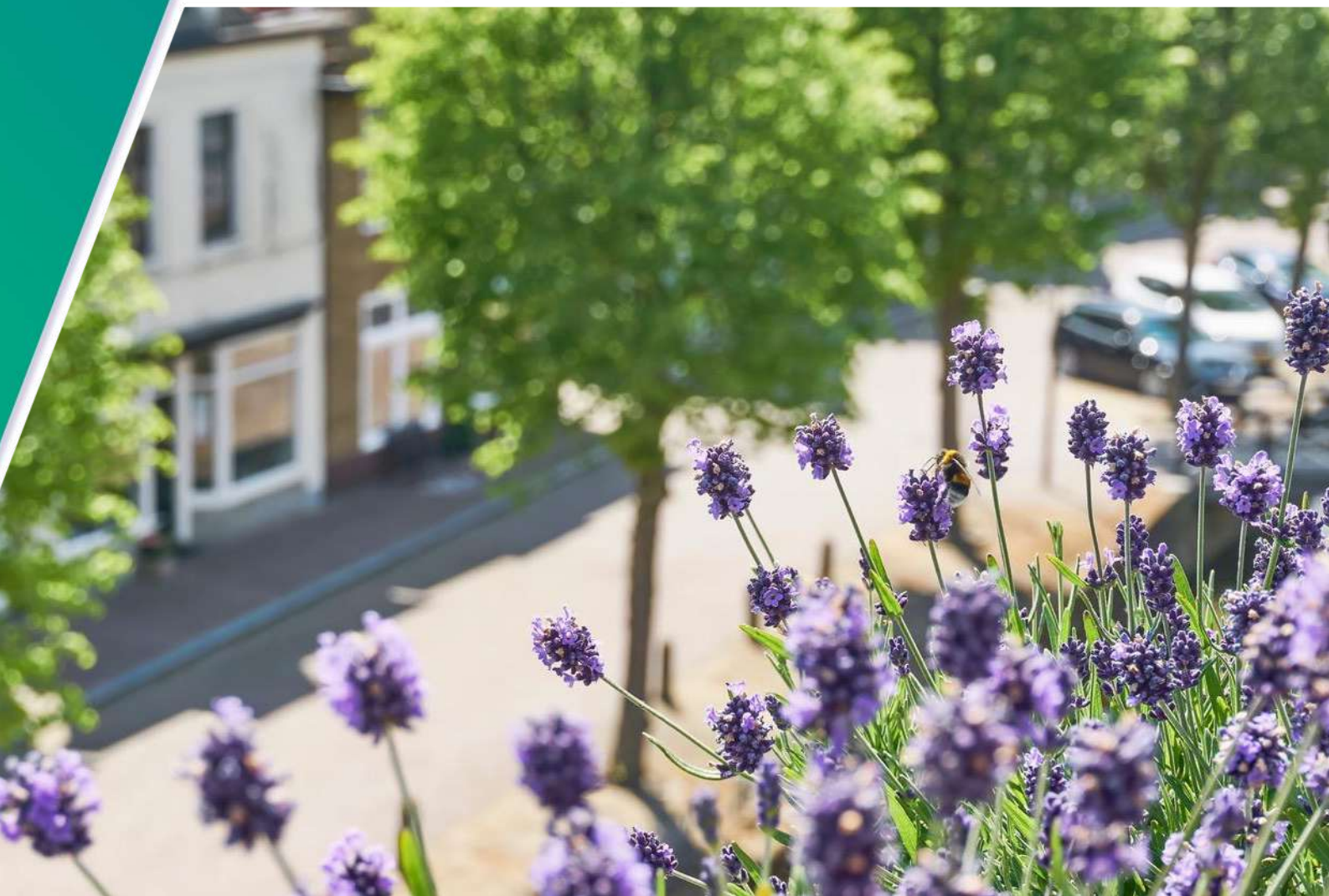




Academische Werkplaats
Gezonde Leefomgeving



Pilot clusteranalyse milieuverkenning Overijssel

Pilot clusteranalyse milieuverkenning Overijssel

2 juni 2025

Auteurs:

Henk Broekhuizen (GGD Twente)

Dit project is financieel mogelijk gemaakt door de Academische Werkplaats Gezonde Leefomgeving met cofinanciering van GGD Twente.

De projectgroep bestond uit:

Henk Broekhuizen (GGD Twente)

Wouter Wolfkamp (GGD Twente)

Het project is uitgevoerd in overleg met de begeleidingscommissie bestaande uit Moniek Zuurbier (GGD Gelderland Midden en coördinator AWGL), Roos van den Broek (GGD Twente, GGD IJsselland), Wicher Wolbers (Provincie Overijssel), Karin Jeurink (Provincie Overijssel) en Justine Blanford (Universiteit Twente)

Met dank aan Floor Leeferink (GGD Twente) en Suzanne Prins (GGD Twente).



SAMENVATTING

Achtergrond: het bovenregionale team Milieu en Gezondheid van de GGD'en in regio oost heeft recent een milieuverkenning uitgevoerd voor de provincie. Uit deze verkenning kwamen twaalf milieugezondheidsthema's naar voren die zich in sommige buurten soms stapelen. Het doel van dit onderzoek was om meer inzicht te krijgen in stapelingsprofielen van buurten middels en ruimtelijke clusteranalyse als methode te verkennen.

Methoden: in dit onderzoek lag de focus op kwetsbaarheid, geluid, groen, temperatuur, NO₂, PM2.5 en bereikbaarheid. Data hierover komt uit openbare data van het CBS, RIVM, de Atlas Leefomgeving en de Klimaat-effectatlas. Waar nodig is data vertaald naar CBS-buurniveau. Voor raster data is dit gedaan met de mediaan waarde in de buurt. Om multicollineariteit te voorkomen, werden sterk ($\text{cor} > 0.8$) correlerende variabelen verwijderd, waarna we ruimtelijke autocorrelatie (Moran's I, LISA) berekenden. De resulterende LISA-waarden vormden de input voor hiërarchische clustering (Ward's methode), waarbij wij de silhouette-score gebruikten om de beste instellingen voor de clustering te bepalen. Met een validatieworkshop werden clusters namen gegeven.

Resultaten: de clusteranalyse onderscheidt negen clusters die verschillen over de thema's. Er is een duidelijke scheiding tussen stedelijke en niet stedelijke clusters, maar ook daarbinnen is verschil. Er zijn twee sterk stedelijke clusters (*bruisende binnenstad* en *kwetsbare buitenwijken*) die vooral te vinden zijn in de grote steden. De meer landelijke clusters verschillen op basis van indicatorwaarden, bv *groene oase* of *rustig buitengebied*. Ten slotte zijn er nog clusters die vooral geconcentreerd zijn rondom mobiliteitsknooppunten of industrie.

Discussie: beschrijvende ruimtelijke clusteranalyse is een effectieve methode om complexe buurniveau-data te 'comprimeren'. Hierdoor kunnen beleidsmakers verdiepend kijken naar de data zonder daarin te verdrinken. Dit kan ze helpen met beleid rondom gericht investeren voor gelijke kansen. Beschrijvende clusteranalyse biedt, in tegenstelling tot normatieve clusteranalyse, geen prioritering van clusters. Het kan beleidsmakers echter wel helpen met het opstellen van denkrichtingen voor oplossing.

Conclusie: wij bevelen aan om beschrijvende clusteranalyse te gebruiken bij gesprekken rondom ruimtelijke milieugezondheidproblematiek. Het is voor daadwerkelijk beleid belangrijk de analyse te gronden in meer data, een gewogen prioritering en een actieve betrokkenheid van beleidsmakers bij de interpretatie en implementatie van de resultaten.

INHOUDSOPGAVE

Samenvatting	3
Hoofdstuk 1 Achtergrond	6
Doelen onderzoek	6
Hoofdstuk 2 Methoden	6
2.1 Dataverzameling	6
2.2 Dataverwerking	8
2.3 Clusteranalyse	8
Hoofdstuk 3 resultaten	9
3.1 Resultaten per indicator	9
3.2 Resultaten clusteranalyse	9
3.3 Beschrijving clusters	10
Hoofdstuk 4 Discussie en conclusie	13
4.1 Discussie	13
4.1.1 Bespreking resultaten	13
4.1.2 Methodologische keuzes	13
4.1.3 Haalbaarheid en toegevoegde waarde bij beleid	14
4.1.4 Advies omtrent gebruik	15
4.1.5 Vervolgstappen	15
4.2 Conclusie	16
Bronnen	17
Bijlage 1: Gebruikte brondata	18
Inkomen en opleidingsniveau	19
Mate van Stedelijkheid	20
Geluid	21
Groen en afstand tot voorzieningen	22
Hitte	23
Luchtkwaliteit (NO ₂ , PM _{2.5})	24
Bijlage 2: Gemeenteprofielen	26
Enschede	26
Losser	27

Almelo	28
Borne	29
Dinkelland	30
Haaksbergen	31
Hengelo	32
Hof van Twente	33
Oldenzaal	34
Tubbergen	35
Twenterand	36
Wierden	37
Hellendoorn	38
Rijssen-Holten	39
Dalfsen	40
Deventer	41
Hardenberg	42
Kampen	43
Olst-Wijhe	44
Ommen	45
Raalte	46
Staphorst	47
Steenwijkerland	48
Zwartewaterland	49
Zwolle	50

HOOFDSTUK 1 ACHTERGROND

Gezondheidsverschillen en een gezonde leefomgeving vormen belangrijke beleidsthema's in de provincie Overijssel. Het bovenregionale team Milieu en Gezondheid van de GGD'en in regio oost heeft recent een milieuverkenning uitgevoerd voor de provincie. Uit deze verkenning kwamen twaalf milieugezondheidsthema's naar voren. Deze thema's stapelen zich in sommige wijken op en kunnen daar specifiek beleid vereisen. Indien er binnen de provincie gebieden zijn met een vergelijkbaar *accumulatieprofiel* (d.w.z. waar risicofactoren opstapelen), kunnen de provincie en gemeenten deze inzichten benutten voor gericht beleid en gezamenlijke kennisontwikkeling.

Doelen onderzoek

In deze studie onderzoeken wij de mogelijkheid om rapportage op het niveau van groepen van gebieden toe te passen. Dit onderzoek komt voort uit de wens van de provincie Overijssel om een diepgaander inzicht te verkrijgen in het omgevingsverkenningsrapport en uit de ambitie van GGD Twente om ruimtelijke clusteranalyse als analysemethode te verkennen.

HOOFDSTUK 2 METHODEN

2.1 Dataverzameling

De te includeren indicatoren zijn gebaseerd op de Overijsselse milieuverkenning en worden gepresenteerd in Tabel 1. Enkele thema's die een minder duidelijke (bewezen) relatie hebben met gezondheid of waar in de milieuverkenning alleen een beschrijving van was opgenomen zonder goede indicator zijn uit de analyse weggelaten. Dit betekent dat er niet gekeken is naar gewasbeschermingsmiddelen, grondwater- en oppervlaktewaterkwaliteit, bodemkwaliteit en licht/donkerte. Geluidshinder, geurhinder en verkoeling zijn geëxcludeerd, omdat er alleen data op gemeenteniveau beschikbaar was. Data over locatie en grootte van veehouderijen zijn geëxcludeerd vanwege commerciële gevoeligheid en de korte tijdsduur van dit project. Alle gegevens waren openbaar beschikbaar via de volgende instellingen: CBS, RIVM, Atlas Leefomgeving en Klimaat-effectatlas.

Naast deze data die gebruikt zal worden voor het clusteren, zijn er ook data verzameld waarmee de gezondheid van inwoners per buurt (en in het verlengde daarvan: per cluster) vergeleken kan worden. Deze komen allen uit de Gezondheidsmonitor Volwassenen en Ouderen 2022, specifiek de SMAP buurt schattingen van RIVM. Er is gekozen om de volgende indicatoren voor ouderen (65+) te gebruiken:

1. Percentage oudere inwoners met goede of zeer goede gezondheid
2. Percentage oudere inwoners met serieus overgewicht
3. Percentage oudere inwoners dat voldoet aan de beweegrichtlijn
4. Percentage oudere inwoners dat (zeer) ernstig eenzaam is

Tabel 1: Overzicht van gebruikte data voor het clusteren. Met 'buurt' wordt bedoeld een buurt volgens de definitie van CBS. * Dit verschilt per bron: wegverkeer (2019-2021), treinverkeer (2019), luchtvaart (2016), industrie (2008), Windturbines (2021). ** Hierbij is de gemiddelde afstand genomen tot deze drie voorzieningen: huisarts, supermarkt, school. Voor categorieën zie <https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/stedelijkheid--van-een-gebied-->

Tabel 1	Meetniveau	Eenheid	Bron en jaar
Huishoudens met een laag inkomen	Buurt	%	CBS-kerncijfers wijken en buurten 2024
Huishoudens met een laag opleidingsniveau	Buurt	%	CBS-kerncijfers wijken en buurten 2024
Mate van stedelijkheid	Buurt	Categorisch***	CBS-kerncijfers wijken en buurten 2024
Geluid	Raster van 10m bij 10m	Gemiddelde Lden avond/nacht	Atlas Leefomgeving 2021*
Groen (excl. agrarisch)	Buurt	% van oppervlakte	Klimaat-effectatlas 2023
Gevoelstemperatuur op een warme zomerdag	Raster van 2m bij 2m	Graden Celsius	Klimaat-effectatlas 2022
NO ₂ concentratie	Raster van 25m bij 25m	Microgram/m ³	Atlas Leefomgeving 2022
PM2.5 concentratie	Raster van 25m bij 25m	Microgram/m ³	Atlas Leefomgeving 2022
Afstand tot voorzieningen**	Buurt	Km	CBS-kerncijfers wijken en buurten 2024

2.2 Dataverwerking

Alle dataverwerkingsstappen zijn uitgevoerd in R. De gebruikte code voor het inladen van data en het uitvoeren van de analyses is publiek beschikbaar op GitHub¹. Voor CBS en RIVM-data is gebruik gemaakt van het *CBSOdataR* package. De andere data zijn gedownload via de website van desbetreffende bron en daarna ingeladen. De analyse wordt uitgevoerd op het niveau CBS-buurtten maar sommige indicatoren zijn alleen als raster beschikbaar (dus een waarde per rechthoekig gebied met bepaalde grootte). Voor dit soort indicatoren is de mediaan van de indicatorwaarde binnen elke buurt berekend. Alle indicatoren zijn genormaliseerd naar een schaal van 0 tot 100. Indicatoren met een hoge onderlinge correlatie (correlatie > 0,8) zijn verwijderd, te beginnen met de indicatoren met de laagste variantie, totdat geen dergelijke hoge correlaties meer aanwezig waren.

2.3 Clusteranalyse

Voor elke indicator is berekend of er ruimtelijke autocorrelatie is m.b.v. Moran's I. Dit is een getal wat meestal tussen de -1 (lokale dispersie: hoge waarden liggen dichtbij lage waarden) of 1 (hoge waarden liggen dicht bij andere hoge waarden) is. Moran's I van 0 betekent dat er geen ruimtelijke autocorrelatie is. Vervolgens zijn ook lokale indicatoren van ruimtelijke autocorrelatie (zogenaamde LISA waarden) bepaald voor elke buurt en elke indicator. LISA-waarden met een p-waarde van meer dan 0,05 zijn op 0 gezet (geen statistisch significante ruimtelijke autocorrelatie). De andere LISA-waarden zijn als volgt gecategoriseerd:

1. lage waarde omringd door lage waarden; een zogenaamde **coldspot**
2. hoge waarde omringd door lage waarden
3. lage waarde omringd door hoge waarden
4. hoge waarde omringd door hoge waarden; een zogenaamde **hotspot**

Vervolgens is een hiërarchische clusteranalyse uitgevoerd op de LISA-waarden om te komen tot een clustertoewijzing voor elke buurt. Voor de clusteranalyse zijn aantal clusters, afstandsmaat en agglomeratiemethode geselecteerd met behulp van een grid search waar geoptimaliseerd is naar de *Silhouette afstand* prestatie maat. Deze maat is hoog naarmate buurten meer lijken op andere buurten in hetzelfde cluster en minder lijken op buurten in andere clusters. Een hoge silhouette afstand betekent dus dat de clusteranalyse een goed onderscheidend vermogen heeft.

¹ https://github.com/HenkBroekhuizen/awgl_clusteranalyse_overijssel

HOOFDSTUK 3 RESULTATEN

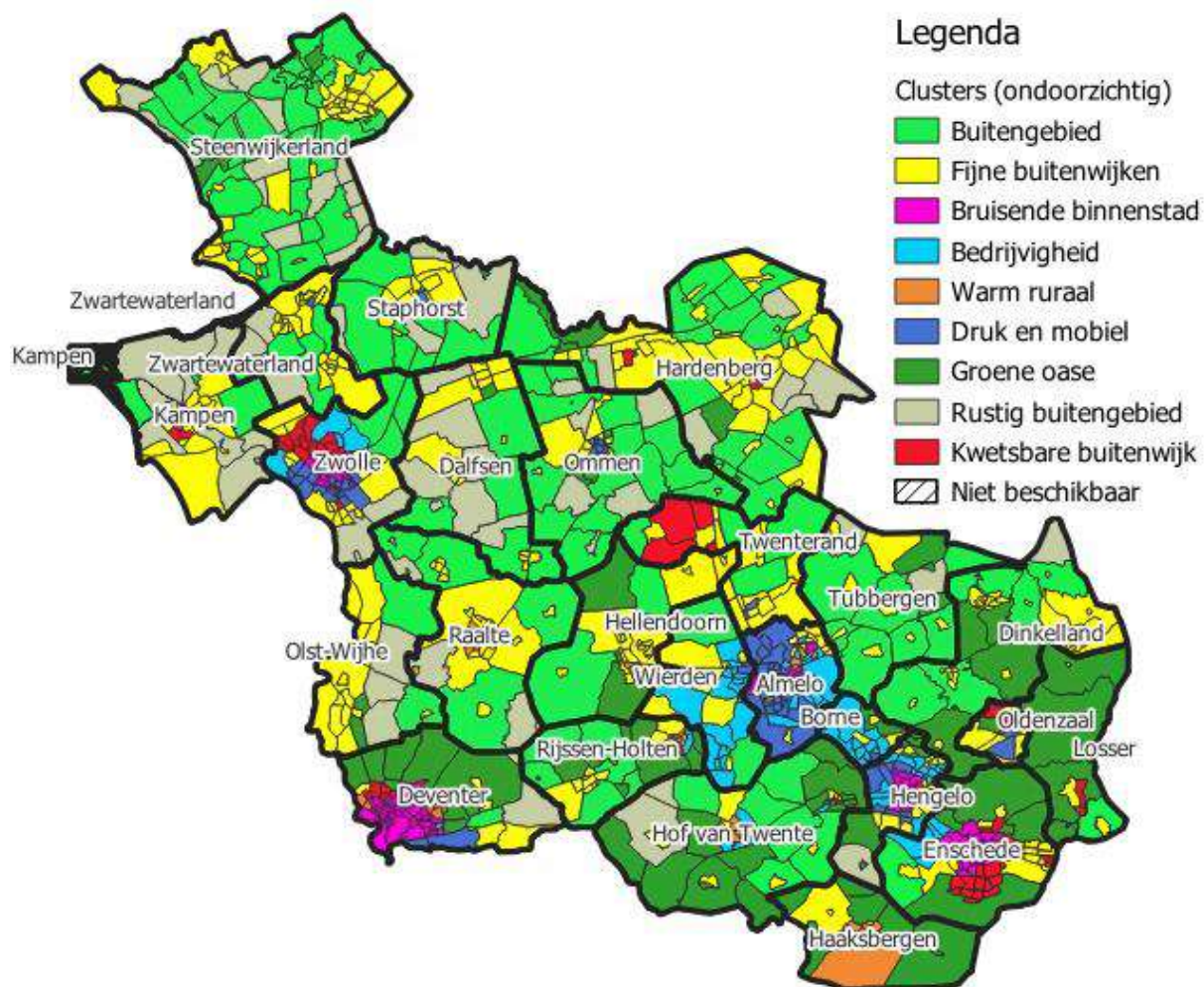
3.1 Resultaten per indicator

Appendix A presenteert de data, zowel de ruwe data als de data die per buurt is geaggregeerd. De indicatoren over *inkomen*, *opleiding* en *stedelijkheid* laten een duidelijke clustering zien rondom bewoond gebied. *Geluidoverlast* is groter in stedelijke gebieden dan in de meeste landelijke gebieden, maar concentreert zich ook in de periferie van stedelijke gebieden, waar buurten zich dicht bij een of meerdere hoofdwegen bevinden. Grote wegen zoals de A50, A1, A37 en de route tussen Almelo en Coevorden eveneens duidelijk zichtbaar. Interessant is dat, hoewel de lokale ruimtelijke clustering rond wegen zichtbaar is in zowel de ruwe als de mediaanplots, de ruimtelijke autocorrelatie in gebieden rondom de hoofdwegen statistisch niet significant was. Het is niet verrassend dat de *boombedekking* buiten stedelijke gebieden groter is. Deze concentreert zich vooral in het zuidelijke deel van de provincie en in de omgeving van Ommen. De *afstand tot voorzieningen* is groter in landelijke gebieden en vertoont ook een clustering in het noordelijke deel van de provincie. In de *hitte* data is het stedelijk hitte-eiland duidelijk zichtbaar, samen met een klein noord-zuid temperatuurgradiënt dat wordt versterkt door de normalisatiestap. Het is belangrijk op te merken dat het effect van het "stedelijk" hitte-eiland zichtbaar is in alle steden en dorpen, niet alleen in de grotere steden². Desondanks bevinden de belangrijkste hotspots zich in de grote Twentse steden, enkele kleinere Twentse plaatsen, en in Raalte en Deventer in IJsselland. Aangezien NO_2 voornamelijk afkomstig is van (intensief) verkeer, is het niet verrassend dat er twee (reeksen van) hotspots zijn: één langs de A1 tussen Deventer en Duitsland en één langs de A50 tussen Zwolle en Drenthe in het noorden. De concentratie van $PM_{2.5}$ is hoger in Twente, met name in de steden en het industriegebied rondom Universiteit Twente.

3.2 Resultaten clusteranalyse

Hiërarchische clustering met *Ward's D-agglomeratie* leverde de hoogste silhouette afstanden op. Ward's methode minimaliseert de totale variantie binnen clusters door telkens de twee clusters samen te voegen die de kleinste toename in variantie geven. Die toename is een gewogen zogenaamde *Gower-afstand* tussen de clustercentra. Deze afstandsmaat normaliseert de verschillen tussen variabelen en berekent vervolgens een gewogen gemiddelde van deze verschillen. De silhouette afstand is hoog bij twee clusters, aangezien het algoritme daar een eenvoudige scheiding maakt tussen stedelijke en landelijke gebieden. Er is een tweede lokaal optimum bij negen clusters. Na een korte daling begint de silhouettescore daarna weer te stijgen. Echter, een groot aantal clusters is moeilijk te interpreteren. Het is dus nodig om een evenwicht te zoeken tussen interpretatiegemak en clusteringkwaliteit. Hierbij is gekozen voor negen clusters in de eindresultaten. Een kaart met de clustertoewijzing per buurt is te vinden in Figuur 1.

² Zie ook <https://klimaatadaptatienederland.nl/@216530/hitte-en-droogte-de-kleine-kernen-en-het-landelijk/>



Figuur 1: Clustertoewijzingen per CBS-buurt. Elke kleur is een ander cluster. Een kaart per gemeente is te vinden in bijlage 2 van dit rapport.

3.3 Beschrijving clusters

Elk van de clusters heeft een eigen accumulatieprofiel van de meegenomen indicatoren, zie Tabel 2. Om discussies over de analyse makkelijker te maken is het handig om clusters namen te geven (i.p.v. cluster 1, cluster 2, etc.). Hiervoor is een korte interpretatiesessie georganiseerd met twee adviseurs gezondheidsbevordering, één adviseur milieu en gezondheid en één epidemioloog. Deze collega's hebben namen aan elk van de clusters gegeven. Bijvoorbeeld op basis van indicatoren die in een bepaald cluster opvallend hoog of laag waren. De naam van een cluster heeft als doel om het makkelijker te maken om (het profiel van) dat cluster te onthouden. De naam is niet alomvattend of precies, het kan bv. zijn dat een stedelijke buurt een soortgelijk profiel heeft als buitenwijken en daardoor onder cluster 'fijne buitenwijken' valt, terwijl het geen buitenwijk is. Naast de indicatoren die gebruikt zijn om buurten te verdelen in clusters, is er ook data verzameld over vier gezondheidsindicatoren per buurt (Tabel 3).

Tabel 2: Indicatorwaarden per cluster. AV=afstand tot basisvoorzieningen.

	Stedelijkheid	Laag inkomen	Laag opleiding	AV	Groen	NO ₂	PM _{2.5}	Geluid	Hitte
Buitengebied	Niet	2%	15 %	4	46 %	10	8	46	34
Fijne buitenwijken	Niet	2%	23 %	1	34 %	10	9	49	36
Bruisende binnenstad	Sterk	6%	22 %	1	20 %	13	9	55	39
Bedrijvigheid	Weinig	2%	25 %	1	33 %	12	9	52	37
Warm ruraal	Matig	3%	22 %	1	28 %	11	9	52	39
Druk en mobiel	Weinig	3%	25 %	1	25 %	12	9	58	37
Groene oase	Niet	2%	17 %	2	64 %	10	9	42	35
Rustig buitengebied	Niet	3%	18 %	5	52 %	9	8	38	34
Kwetsbare buitenwijk	Sterk	6%	22 %	1	31 %	11	9	50	37

Stedelijke clusters

De resultaten laten een duidelijk spectrum zien van landelijke profielen naar stedelijke profielen. Exemplarisch van de stedelijke kant van dit spectrum is het cluster *bruisende binnenstad*. Buurten in dit cluster bevinden zich voornamelijk in het centrum van Overijsselse grote steden. Er woont en gebeurt veel in dit cluster, wat zich vertaalt naar weinig ruimte voor groen, veel hitte en veel luchtverontreiniging. De ervaren gezondheid van inwoners van dit cluster is het op één-na laagst, alhoewel er relatief weinig overgewicht is en gemiddelde eenzaamheid. Het tweede stedelijke cluster heet *kwetsbare buitenwijk*. Buurten in dit cluster bevinden zich vaak aan de buitenkant van de (grotere) steden. Ten opzichte van de bruisende binnenstad wonen er meer mensen met een laag opleidingsniveau. Buurten in dit cluster hebben een iets beter accumulatieprofiel dan de bruisende binnenstad: er is iets meer groen, en iets minder luchtverontreiniging, etc. Desondanks zijn de gezondheidsuitkomsten in dit

cluster het meest nadelig van alle clusters: de ervaren gezondheid is het laagst en er is veel overgewicht en eenzaamheid.

Tabel 3: Gezondheidsindicatoren per cluster voor inwoners >65 jaar.

	Ervaren gezondheid (zeer) goed	Ernstig overgewicht	Voldoet aan beweegrichtlijn	(Ze)er ernstig eenzaam
Buitengebied	65%	21%	36%	8%
Fijne buitenwijken	66%	17%	43%	9%
Bruisende binnenstad	60%	20%	39%	12%
Bedrijvigheid	66%	17%	45%	9%
Warm ruraal	62%	18%	41%	10%
Druk en mobiel	66%	16%	45%	9%
Groene oase	67%	15%	45%	9%
Rustig buitengebied	69%	16%	40%	9%
Kwetsbare buitenwijk	57%	20%	38%	13%

Landelijke clusters

Aan de andere kant van het spectrum liggen de sterk landelijke clusters *buitengebied*, *fijne buitenwijken*, *groene oase*, en *rustig buitengebied*. Hiertussen wordt het onderscheid gemaakt op basis van indicatorwaarden die tussen deze clusters verschillen. Buurten in het cluster *buitengebied* kennen een grote afstand tot voorzieningen, er zijn ook veel coldspots wat betreft geluid, hitte, en luchtverontreiniging. Er wonen weinig mensen met een laag inkomen. Hetzelfde geldt voor buurten in het cluster *groene oase*. Dit is het cluster met het meeste groen. Buitengebied buurten komen vooral voor in IJsselland en het noorden van Twente, groene oase buurten vooral ten noorden van Deventer en in het zuiden/zuidoosten van Twente. Het cluster *fijne buitenwijk* omvat (het gebied rondom) een aantal grotere plaatsen zoals Steenwijk, Denekamp en Raalte. Dit impliceert dat deze dorpen een soortgelijk profiel hebben als het omliggende gebied. Buurten in dit cluster hebben een profiel van veel groen en relatief weinig hitte, maar wel meer geluid en luchtverontreiniging dan de eerste twee genoemde 'groene' landelijke clusters. Het laatste landelijke cluster, *rustig buitengebied*, bevindt zich vooral in IJsselland en heeft het op één-na-hoogste percentage groen. Het wordt gekarakteriseerd door de laagste waarden voor geluid, luchtverontreiniging, hitte. De afstand tot voorzieningen is het hoogst.

Clusters tussen het stedelijke en landelijke in

De laatste drie clusters worden, meer dan de andere clusters, bepaald door specifieke hoge waarden van indicatoren. Het *warm ruraal* cluster bevat buurten in dorpen die gemiddeld erg warm zijn, haast even warm als de bruisende binnenstad. Voorbeelden hiervan zijn Haaksbergen en Wierden. Buurten in het cluster *druk en mobiel* liggen veelal rondom auto- en treinverkeersknooppunten, zoals de zuidkant van Deventer en Zwolle en de westkant van Almelo. Er is veel luchtverontreiniging en geluid. Buurten in het cluster *bedrijvigheid* hebben een soortgelijk, maar iets groener profiel, en bevinden zich relatief vaak in de gebieden tussen de drie Twentse steden.

HOOFDSTUK 4 DISCUSSIE EN CONCLUSIE

4.1 Discussie

4.1.1 Bespreking resultaten

Dit project onderzocht ruimtelijke clustering als methode om accumulatieprofielen van milieugezondheidsindicatoren in provincie Overijssel in kaart te brengen op CBS-buurniveau. De belangrijkste kracht van het project was de integratie van meerdere gegevensbronnen in één analyse. Dit biedt een diepgaander inzicht in ruimtelijke patronen van milieugezondheid binnen de provincie. Alhoewel een grove clustering van accumulatieprofiel puur en alleen op basis van stedelijkheid wel kloppend is, gaat dat voorbij aan specifieke (en wellicht beleidsrelevante) verschillen tussen buurten. Want ook buurten die hetzelfde zijn qua stedelijkheid kunnen verschillen op tal van andere indicatoren. Er zijn bijvoorbeeld enkele buurten in kleinere steden waarvan het accumulatieprofiel vergelijkbaar is met dat van een dichtbebouwde stedelijke wijk. De resultaten tonen ook aan dat de groepering van buurten over grenzen gaat van gemeenten en GGD-regio's.

4.1.2 Methodologische keuzes

Tijdens het onderzoek moesten verschillende methodologische keuzes worden gemaakt, deels ingegeven door gegevensbeschikbaarheid en deels door tijdsbeperkingen. Twee kwesties verdienen nadere toelichting: gegevensbeschikbaarheid en ruimtelijke afbakeningen. In eerdere versies van de analyse werden gemeentelijke data uit GGD-enquêtes en het Mullier Instituut gebruikt voor indicatoren rondom geurhinder en de beschikbaarheid van gezonde voorzieningen in de buurt. Hierbij werd aangenomen dat alle buurten binnen een gemeente dezelfde waarde hadden. Dit vertekende echter de clusteranalyse, doordat deze dan een zwaar gewicht gaf aan deze indicatoren en daardoor voornamelijk gemeentelijke grenzen weerspiegelde. Uiteindelijk werd geurhinder geëxcludeerd uit de analyse. Voor de beschikbaarheid van gezonde voorzieningen werd het gemiddelde van meerdere andere voorzieningen als proxy gebruikt, met de aanname dat gezonde voorzieningen zich op vergelijkbare afstanden bevinden. Dit leidde waarschijnlijk tot een overschatting van de afstand tot gezonde voorzieningen, aangezien sommige voorzieningen, zoals lokale parken, wijdverspreider zijn dan de voorzieningen uit de CBS-database die werden gebruikt.

Wat betreft de ruimtelijke afbakeningen bestaat er een aanzienlijke heterogeniteit in buurtgroottes binnen de provincie: stedelijke buurten zijn doorgaans veel kleiner dan landelijke buurten. Aangezien clustering plaatsvindt op basis van één mediaanwaarde per indicator per buurt, kunnen waarden in grotere gebieden vertekend worden wanneer de onderliggende verdeling scheef is. Dit is met name het geval wanneer slechts een deel van een groot gebied bewoond is en de rest bestaat uit landbouwgrond of bos. De variatie in buurtgrootte leidt ook tot verschillen in connectiviteit: sommige gebieden zijn slechts verbonden met één ander gebied. Zo bestaat een klein dorp uit één buurt die volledig wordt omringd door een grotere landelijke buurt. Dit resulteert in een 'eiland-effect' dat vooral in landelijke gebieden optreedt. Dit maakt het minder waarschijnlijk dat grote buurten deel zijn van cold- of hotspots en verlaagt daardoor het onderscheidend vermogen van de clusteranalyse. Een verbeterrichting is het gebruik van slechts rasterdata, bv door gebruik te maken van data uit CBS-vierkanten³.

4.1.3 Haalbaarheid en toegevoegde waarde bij beleid

Een clusteranalyse kan waarde hebben voor beleidsmakers omdat het meer inzicht kan geven in aan milieugezondheid gerelateerde problematiek waar veel data (veel indicatoren over veel gebieden) tegelijk van belang zijn bij het opstellen van beleid. Een voorbeeld hiervan is *gericht investeren voor gelijke kansen*. Dit beleid vormgeven kan niet als er alleen heel algemene data is, bv op provincie of gemeenteniveau. Het interpreteren van de resultaten van een clusteranalyse geeft beleidsmakers een kans in de onderliggende data te duiken zonder te verdrinken in alle data die gepaard gaat met een analyse op buurtniveau. Om een idee te krijgen: in de provincie zijn er 1089 CBS-buurten, met twaalf thema's betekent dat $1089 \times 12 = 13.068$ gegevenspunten die in een tabel verwerkt en geïnterpreteerd moeten worden. Dit is veel meer dan de $9 \times 12 = 108$ in de huidige clusteranalyse. De clusteranalyse biedt dus een geschikt startpunt voor het verder verdiepen van specifieke delen van de dataset, zonder al voor de analyse start in de data te "verdrinken". GGD-adviseurs, die kennis hebben van zowel milieugezondheid als van onderzoek, kunnen beleidsmakers hierbij helpen.

Het zou ook kunnen lonen om te kijken naar clusters die op elkaar lijken op sommige eigenschappen (bv stedelijkheid) maar verschillen in andere eigenschappen (bv NO₂). In onze resultaten zijn er bijvoorbeeld meerdere clusters met een lage stedelijkheid, maar met een verschillend profiel w.b. luchtverontreiniging. De beschrijvende clusteranalyse kan het begin zijn van een zoektocht naar het antwoord op de vraag welke beleidsmaatregelen er in de 'schonere' gebieden zijn uitgevoerd om te komen tot dit profiel.

Doordat gebruik gemaakt wordt van openbare data en het R script beschikbaar is, is het uitvoeren van de clusteranalyse goedkoop en niet tijdrovend. De randvoorwaarde is wel dat er in de organisatie enige technische kennis beschikbaar is om de analyses te draaien en de resultaten zodanig te verwerken dat het bruikbaar is voor beleidsmaker, bv door een kaart te maken met R of een GIS-tool zoals ArcGIS of QGIS. Organisaties kunnen met data de analyse verrijken door indicatoren toe te voegen waarop geclusterd wordt of door indicatoren toe te voegen waarmee clusters kunnen worden vergeleken (zoals in dit project is gedaan met gezondheidsuitkomsten). Het is bij het toevoegen van data wel zaak te waken voor data die

³ <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/kaart-van-100-meter-bij-100-meter-met-statistieken>

hoog correleert met al meegenomen data en hiervoor te corrigeren door hoog gecorreleerde kolommen te verwijderen.

In dit project zijn beschrijvende clusters gevonden. Het doel daarvan is om te weten welke buurten er op elkaar lijken, gegeven de indicatoren die we meenemen. Het clusteranalyse algoritme houdt geen rekening met het relatieve belang (bv. impact op gezondheid) van indicatoren over de meegenomen range van waarden. Er zijn andere projecten geweest die wel meer normatieve clusters hebben gezocht. Het doel daarvan is te weten welke buurten de hoogste prioriteit hebben om aan te pakken in beleid, bv door alle indicatoren te normaliseren naar een getal tussen 0 en 1 en deze dan te aggregeren. Voorbeelden hiervan zijn de Twentse buurtbarometer⁴ (breed, niet specifiek gezondheid of leefomgeving) of een analyse van het RIVM over risicofactoren in de leefomgeving⁵. Beide van deze projecten geven indicatoren gelijke gewichten, maar het is mogelijk om indicatoren gewichten te geven. Een beschrijvende clusteranalyse als uitgevoerd in dit project kan helpen bij het duiden van de prioriteitsscores en het ontwerpen van beleid. Dit kan door bijvoorbeeld te kijken naar soortgelijke buurten met een overeenkomstig profiel, ook over gemeentegrenzen heen. Eenzelfde soort systematiek op gemeenteniveau is al gemaakt door de VNG⁶.

4.1.4 Advies omtrent gebruik

Beschrijvende ruimtelijke clusteranalyse is een goede toevoeging aan de gereedschapskist van onderzoekers die met data-analyses adviseurs en beleidsmakers ondersteunen op het gebied van milieugezondheid.

4.1.5 Vervolgstappen

Om nuttig te zijn voor beleidsmakers en de onderzoekers die hen ondersteunen, hebben beide groepen stappen te zetten.

Beleidsmakers kunnen faciliteren dat betere data beschikbaar is. Veel data over hoe mensen hun leefomgeving ervaren, zoals geluid- en geurhinder, komen uit de gezondheidsmonitors die periodiek wordt uitgevoerd door GGD'en. De steekproefgroottes hiervan zijn vaak niet groot genoeg om schattingen op wijk (laat staan buurt) niveau te kunnen berekenen en RIVM SMAP schattingen zijn niet haalbaar voor alle indicatoren. Beleidsmakers kunnen steekproefophogingen, waarbij gemeenten de GGD betalen om meer vragenlijsten te versturen en te analyseren, faciliteren of aanmoedigen. Hierdoor is de kans groter dat er genoeg data is om wijk- of buurtschattingen te berekenen. Data over de locatie en grootte van veehouderijen waren niet beschikbaar vanwege commerciële gevoeligheid. Hierdoor missen we een deel van het plaatje en hebben de analyses minder waarde voor beleidsmakers.

Voor wat betreft de rol van onderzoekers is een logische vervolgstap om de analyse rondom gezondheidsuitkomsten uit te breiden. De berekende gezondheidsuitkomsten per cluster in dit onderzoek waren informatief, maar boden mogelijk een vertekend beeld, omdat bijvoorbeeld

⁴ <https://kennispunttwente.nl/twentsebuurtbarometer>

⁵ <https://gisportal.rivm.nl/arcgis/apps/storymaps/stories/85d006bc9c124413b8b08238d3f2c742>

⁶ <https://amigo.waarstaatjegemeente.nl/>

de leeftijdsopbouw van de bevolking niet werd meegenomen. Daarnaast verdient de invloed van lineaire bronnen versus puntbronnen op de clusterresultaten verdere aandacht. Grote wegen waren bijvoorbeeld duidelijk zichtbaar in de ruwe NO₂-data en gemiddelde buurtwaarden, maar werden niet als hotspots herkend in de LISA-analyse, omdat het effect van wegen op de mediaanwaarde van grote gebieden klein is. Een tweede vervolgstap voor onderzoekers is het uitbreiden van de geografische en temporele grenzen van de analyse. De eerste uitbreiding is (relatief) eenvoudig, aangezien de gebruikte opendata veelal landelijk beschikbaar is. Echter, de interpretatie wordt hierdoor aanzienlijk complexer, gezien de omvang van de resultaten en het grote aantal betrokken belanghebbenden. Het uitbreiden van de tijdshorizon kan relevante inzichten opleveren over trends in milieugezondheid, maar vereist ook een uitgebreidere dataset en mogelijk interpolatie voor jaren met missende indicatorwaarden. Toch kunnen hiermee belangrijke vragen worden beantwoord, zoals: zijn clusters stabiel over tijd, en zo niet, wat veroorzaakt de veranderingen? Welke factoren hebben ertoe geleid dat sommige buurten bijvoorbeeld zijn verschoven van een ongunstige clusters naar gunstigere cluster? Dit zou de basis kunnen vormen voor kennisuitwisseling tussen gemeenten.

4.2 Conclusie

Dit project toonde aan dat een clusteranalyse een waardevol instrument kan zijn om meer inzicht te verkrijgen in de complexiteit van milieugezondheid binnen de provincie Overijssel. De werkelijke waarde van de methoden en resultaten vereist echter meer data en kan alleen worden vastgesteld in samenwerking met beleidsmakers bij de GGD, gemeenten en de provincie.

BRONNEN

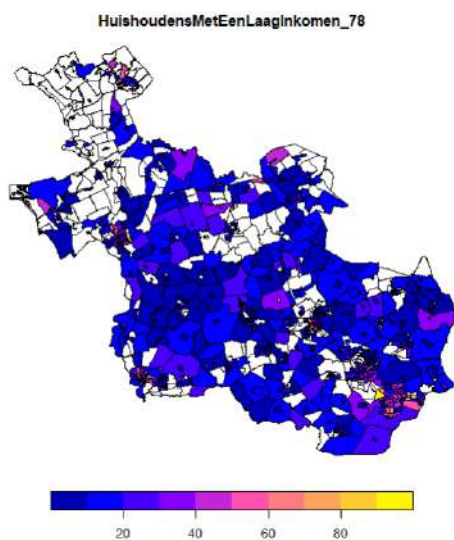
- 1) Klimaatadaptatie Nederland. (2019). *Hitte en droogte in de kleine kernen en het landelijk gebied van Overijssel*. Kennisportaal Klimaatadaptatie. Geraadpleegd op 2 juni 2025. Beschikbaar op: <https://klimaatadaptatienederland.nl/@216530/hitte-en-droogte-de-kleine-kernen-en-het-landelijk/>
- 2) Centraal Bureau voor de Statistiek (CBS). (2025). *Kaart van 100 meter bij 100 meter met statistieken*. Nederland Regionaal – Geografische data. Geraadpleegd op 2 juni 2025. Beschikbaar op: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/kaart-van-100-meter-bij-100-meter-met-statistieken>
- 3) Kennispunt Twente. (z.d.). *Twentse BuurtBarometer*. Geraadpleegd op 2 juni 2025. Beschikbaar op: <https://kennispunttwente.nl/twentsebuurtbarometer>
- 4) Rijksinstituut voor Volksgezondheid en Milieu (RIVM). (z.d.). *Gezondheid en Ruimte - Kwetsbaarheid en Kansen in de Stapeling van Omgevingsfactoren*. Geraadpleegd op 2 juni 2025. Beschikbaar op: <https://gisportal.rivm.nl/arcgis/apps/storymaps/stories/85d006bc9c124413b8b08238d3f2c742>
- 5) Vereniging van Nederlandse Gemeenten (VNG). (2025). *AMIGO – Algoritmische Matchmaker voor Ideale Gemeentelijke Oriëntatie*. Geraadpleegd op 2 juni 2025. Beschikbaar op: <https://amigo.waarstaatjegemeente.nl/>

BIJLAGE 1: GEBRUIKTE BRONDATA

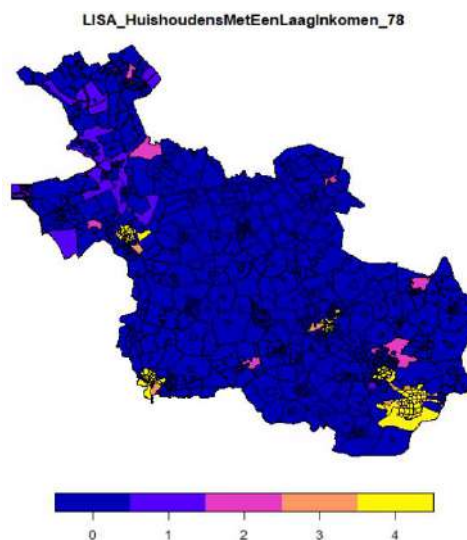
Deze appendix laat per indicator zien:

1. De ruwe data (alleen voor indicatoren met raster data als bron). Voor definities en bron zie tabel 1.
2. Data per buurt (genormaliseerd tussen 0=minimum en 100=maximum)
3. LISA waarden per buurt

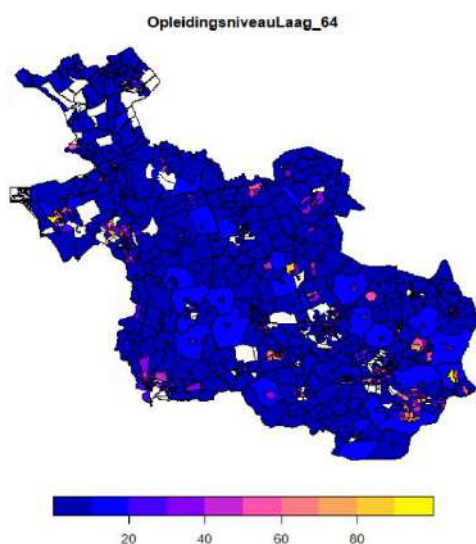
Inkomen en opleidingsniveau



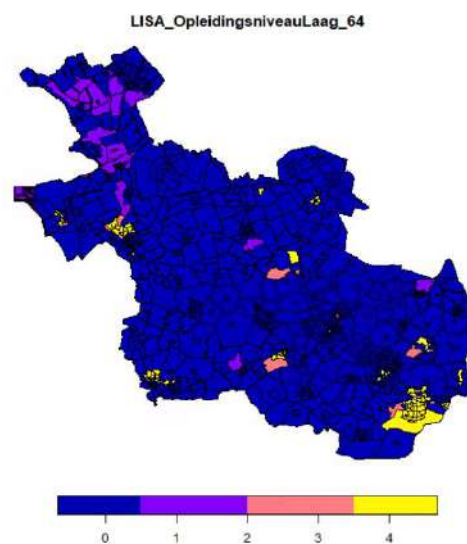
Huishoudens met een laag inkomen per CBS buurt. Bron: CBS kerncijfers wijken en buurten 2024.



Local indicators of spatial autocorrelation per CBS buurt voor de indicator 'huishoudens met een laag inkomen'.

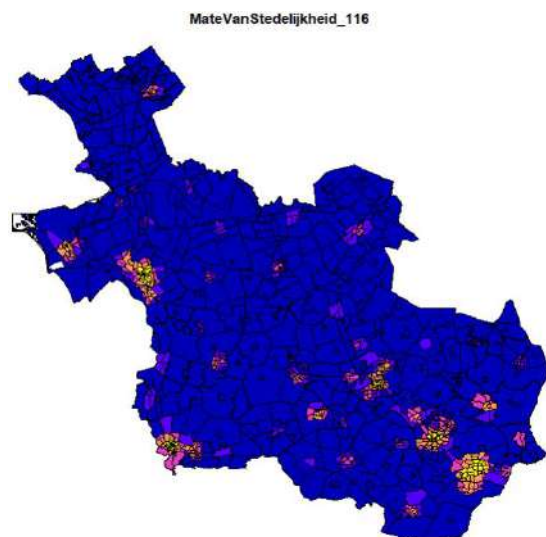


Huishoudens met een laag opleidingsniveau per CBS buurt. Bron: CBS kerncijfers wijken en buurten 2024.

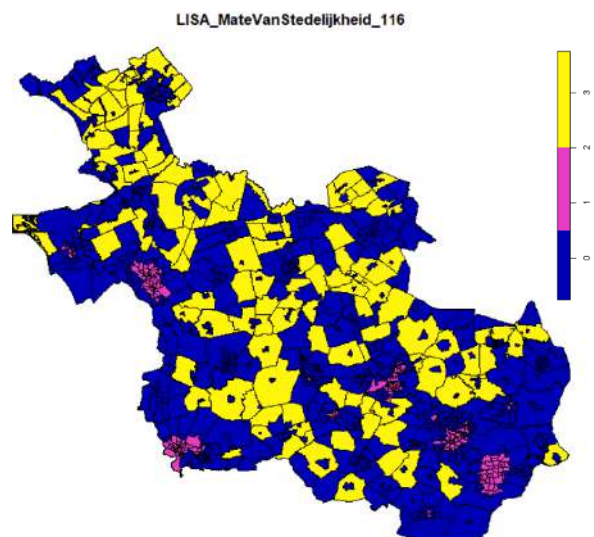


Local indicators of spatial autocorrelation per CBS buurt voor de indicator 'huishoudens met een lage opleiding'.

Mate van Stedelijkheid

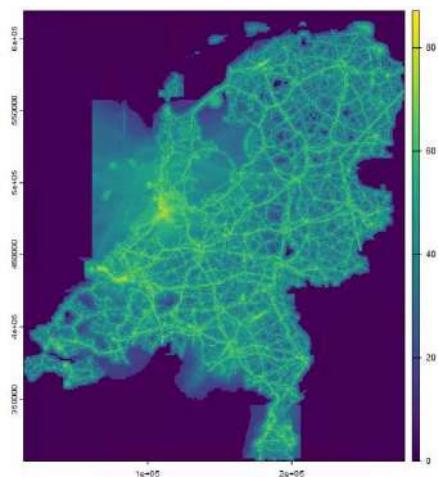


Mate van stedelijkheid per CBS buurt. Bron: CBS kerncijfers wijken en buurten 2024.

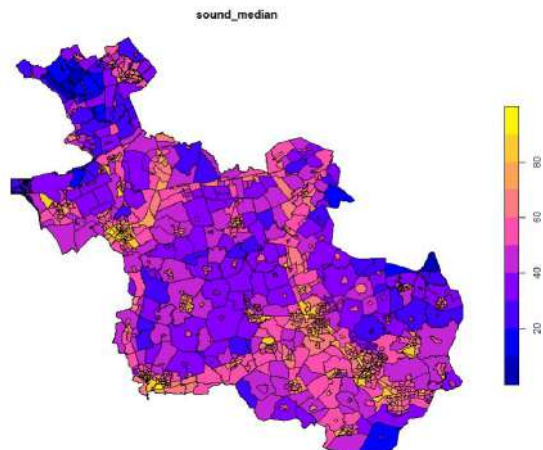


Local indicators of spatial autocorrelation per CBS buurt voor de indicator 'stedelijkheid'.

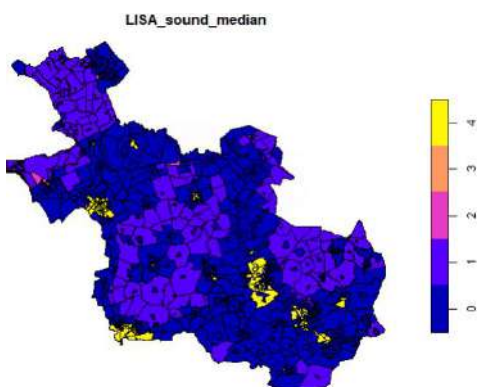
Geluid



Gemiddelde Lden avond/nacht per vierkant van 10 bij 10 meter. Bron: Atlas leefomgeving 2021.

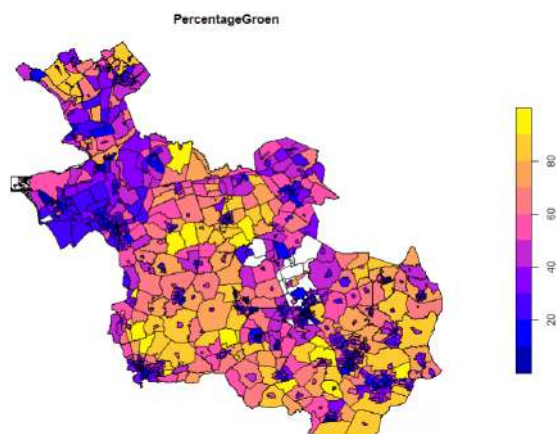


Mediaan Lden avond/nacht per CBS buurt.

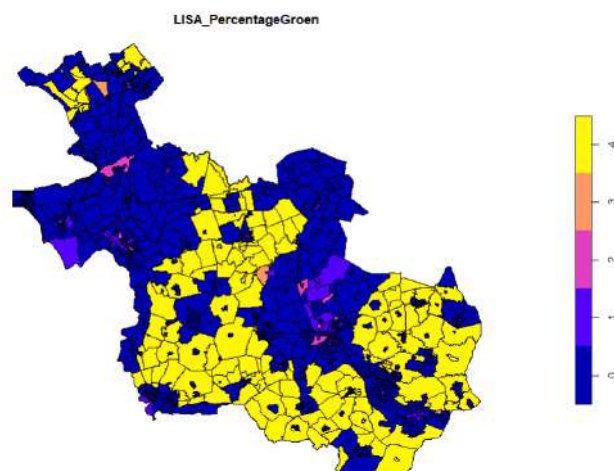


Local indicators of spatial autocorrelation
per buurt voor de indicator 'Lden
avond/nacht'.

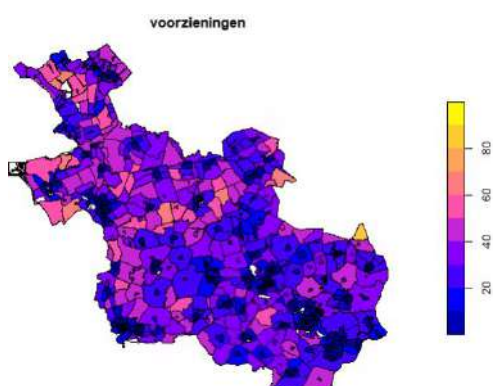
Groen en afstand tot voorzieningen



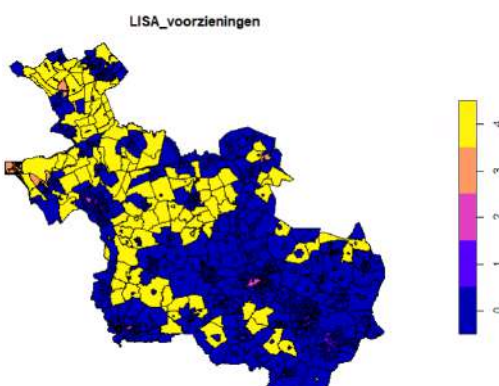
Percentage oppervlakte groen per CBS buurt. Bron: klimaateffectatlas 2023.



Local indicators of spatial autocorrelation per CBS buurt voor de indicator 'percentage oppervlakte groen'.

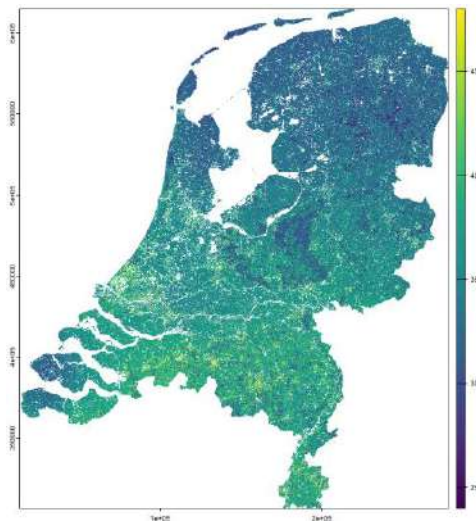


Afstand tot voorzieningen per CBS buurt. Bron: CBS kerncijfers wijken en buurten 2024.

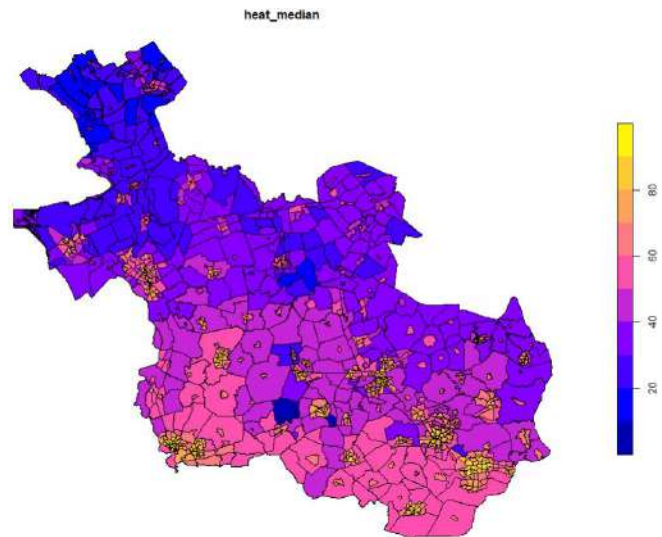


Local indicators of spatial autocorrelation voor de indicator 'afstand tot voorzieningen'.

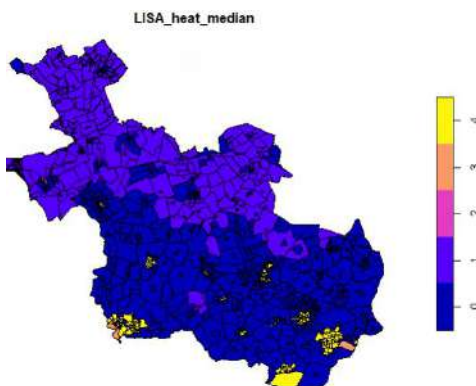
Hitte



Hitte (gevoelstemperatuur op een warme zomerdag) per vierkant van 2 bij 2 meter. Bron Klimateffectatlas 2022.

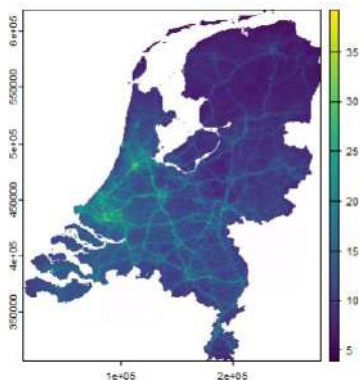


Mediane hitte (gevoelstemperatuur op een warme zomerdag) per CBS buurt.

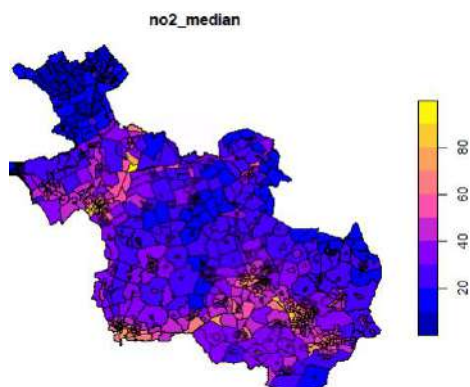


Local indicators of spatial autocorrelation voor de indicator 'hitte'.

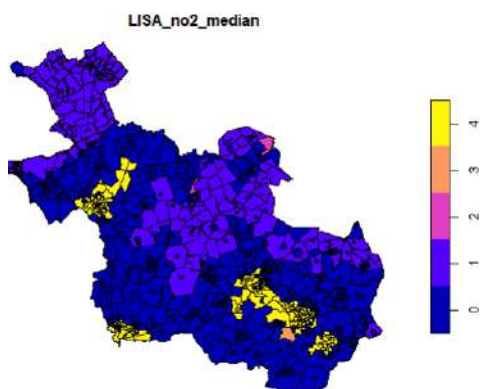
Luchtkwaliteit (NO₂, PM_{2.5})



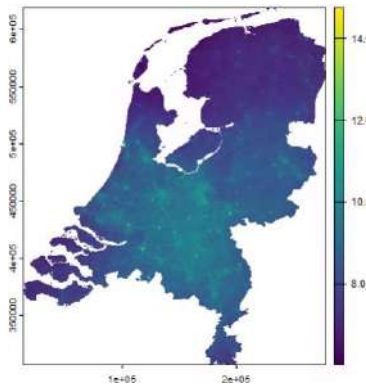
NO₂ concentratie per 25 bij 25 meter. Bron: atlas leefomgeving 2022.



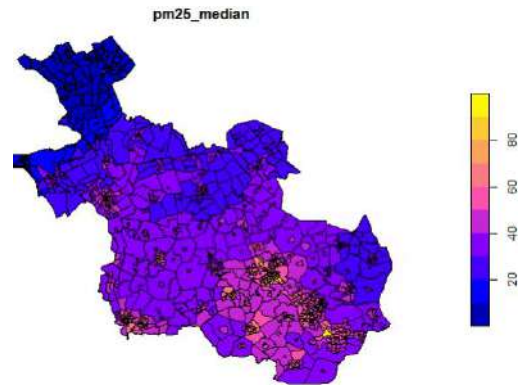
Mediane NO₂ concentratie per CBS buurt.



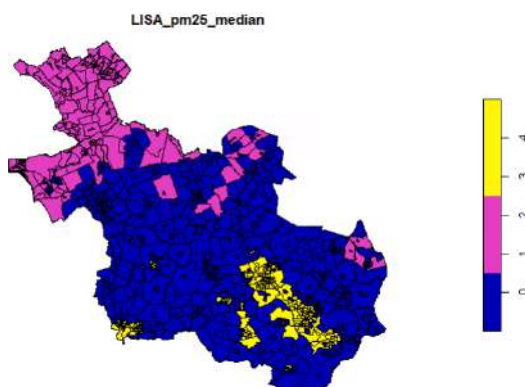
Local indicators of spatial autocorrelation voor de indicator 'NO₂'.



PM2.5 concentratie per vierkant van 25 bij 25 meter. Bron: Atlas leefomgeving 2022.



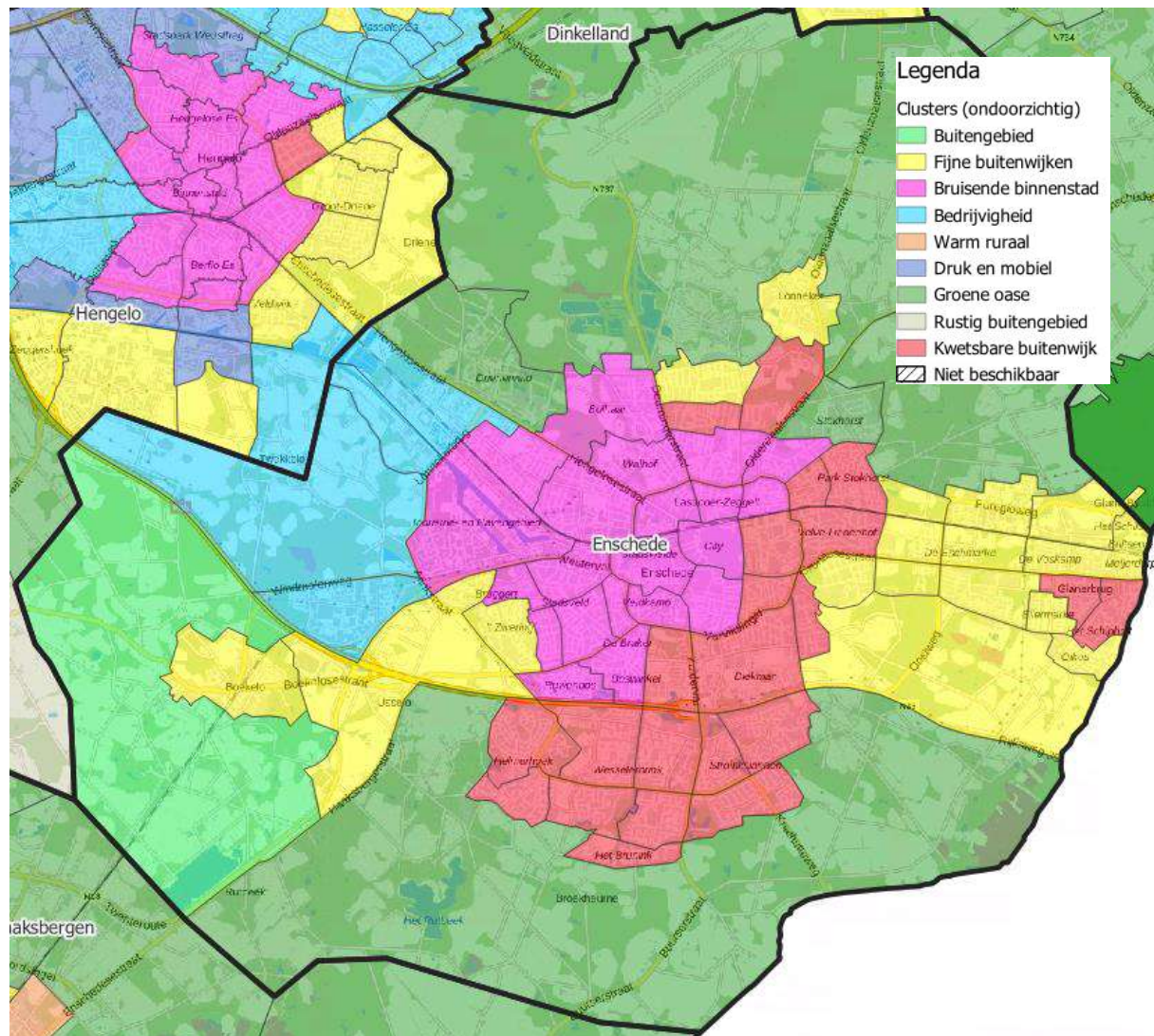
Mediane PM2.5 concentratie per CBS buurt.



Local indicators of spatial autocorrelation voor de indicator 'PM2.5.

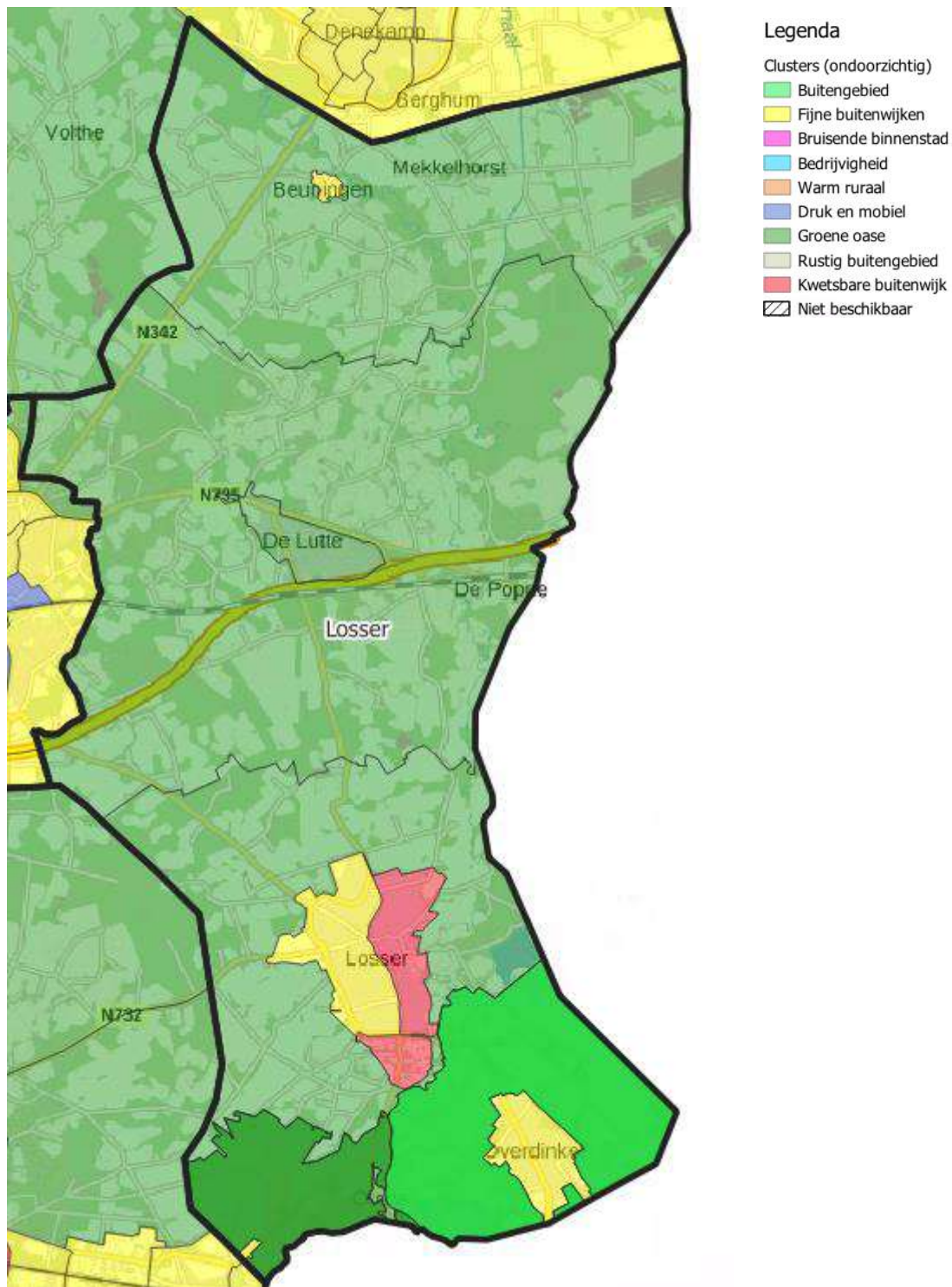
BIJLAGE 2: GEMEENTEPROFIELEN

Enschede



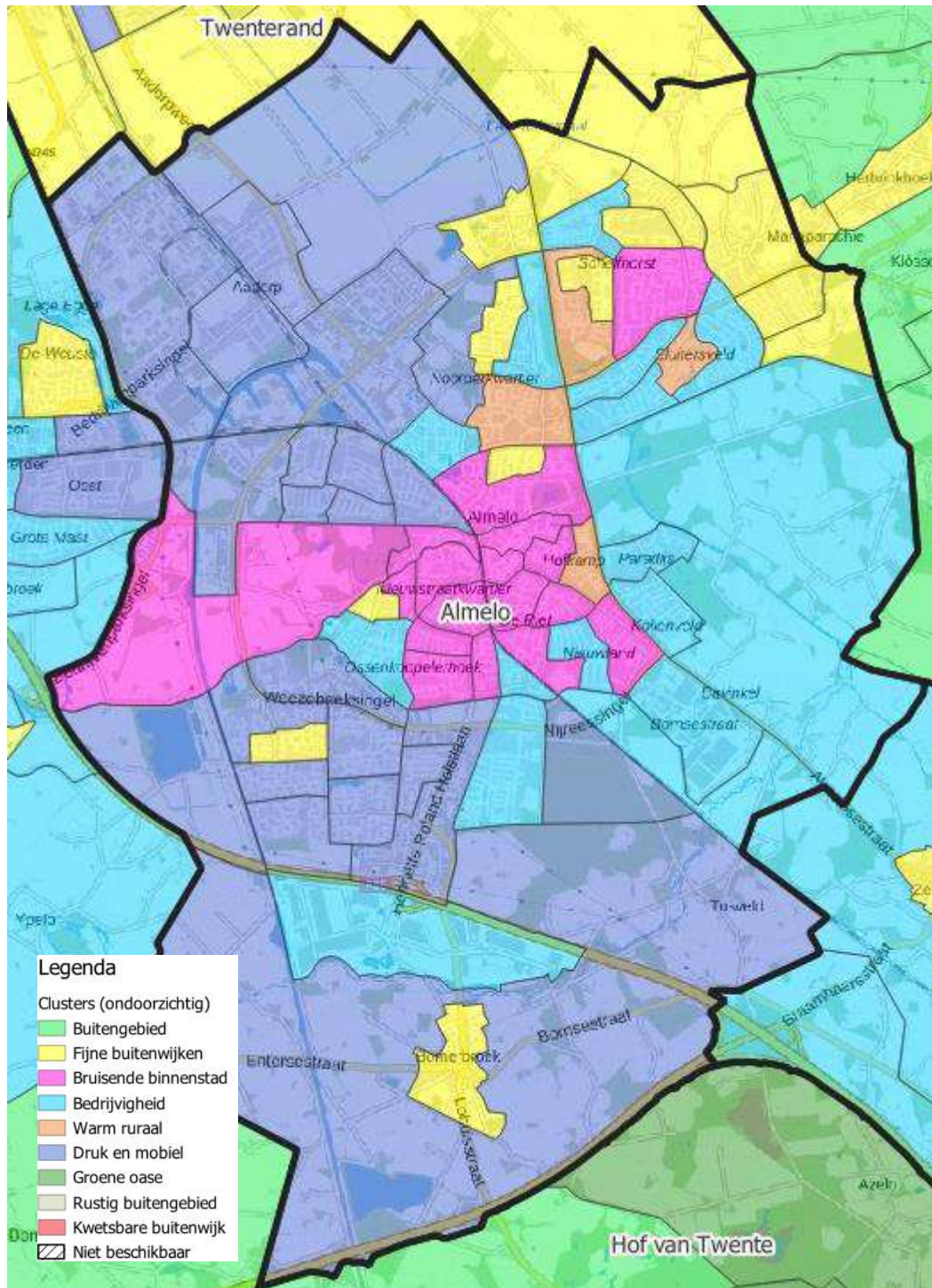
Clusterindelingen per CBS buurt in gemeente Enschede. Buurten in de binnenstad vallen vooral onder cluster 'bruisende binnenstad'. Buurten in het zuiden en oosten van de stad vallen vaker onder cluster 'kwetsbare buitenwijk'. Er zijn enkele 'fijne buitenwijk' buurten ten zuidwesten en ten oosten van Enschede. Het gebied rondom station Enschede kennispark valt onder cluster 'bedrijvigheid'. De rest van de landelijke delen van gemeente Enschede valt vooral onder 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Losser



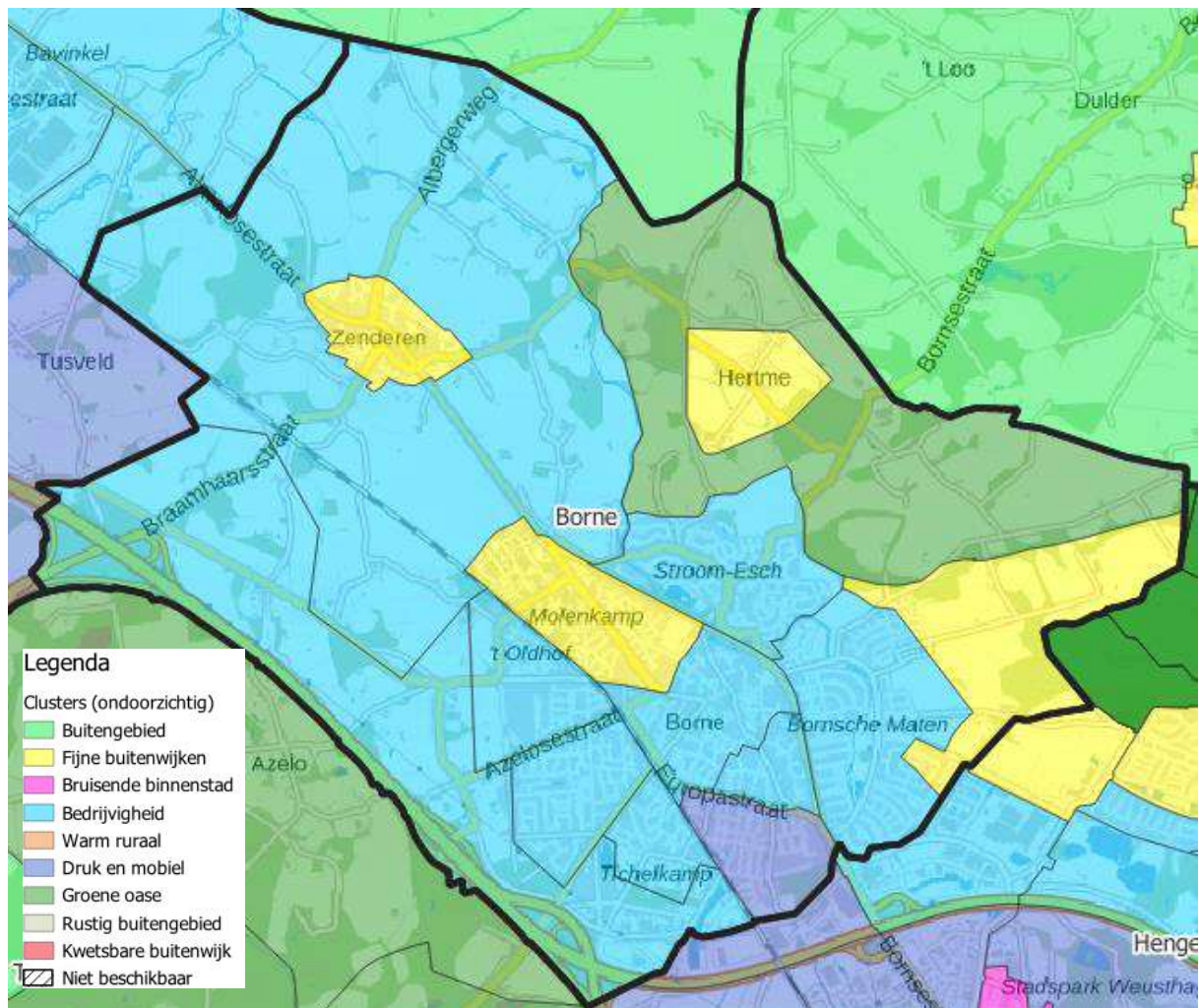
Clusterindelingen per CBS buurt in gemeente Losser. Twee van de drie CBS buurten in Losser zelf vallen onder 'kwetsbare buitenwijken'. Overdinkel en Losser west vallen onder 'fijne buitenwijk'. De rest van de gemeente valt vooral onder 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Almelo



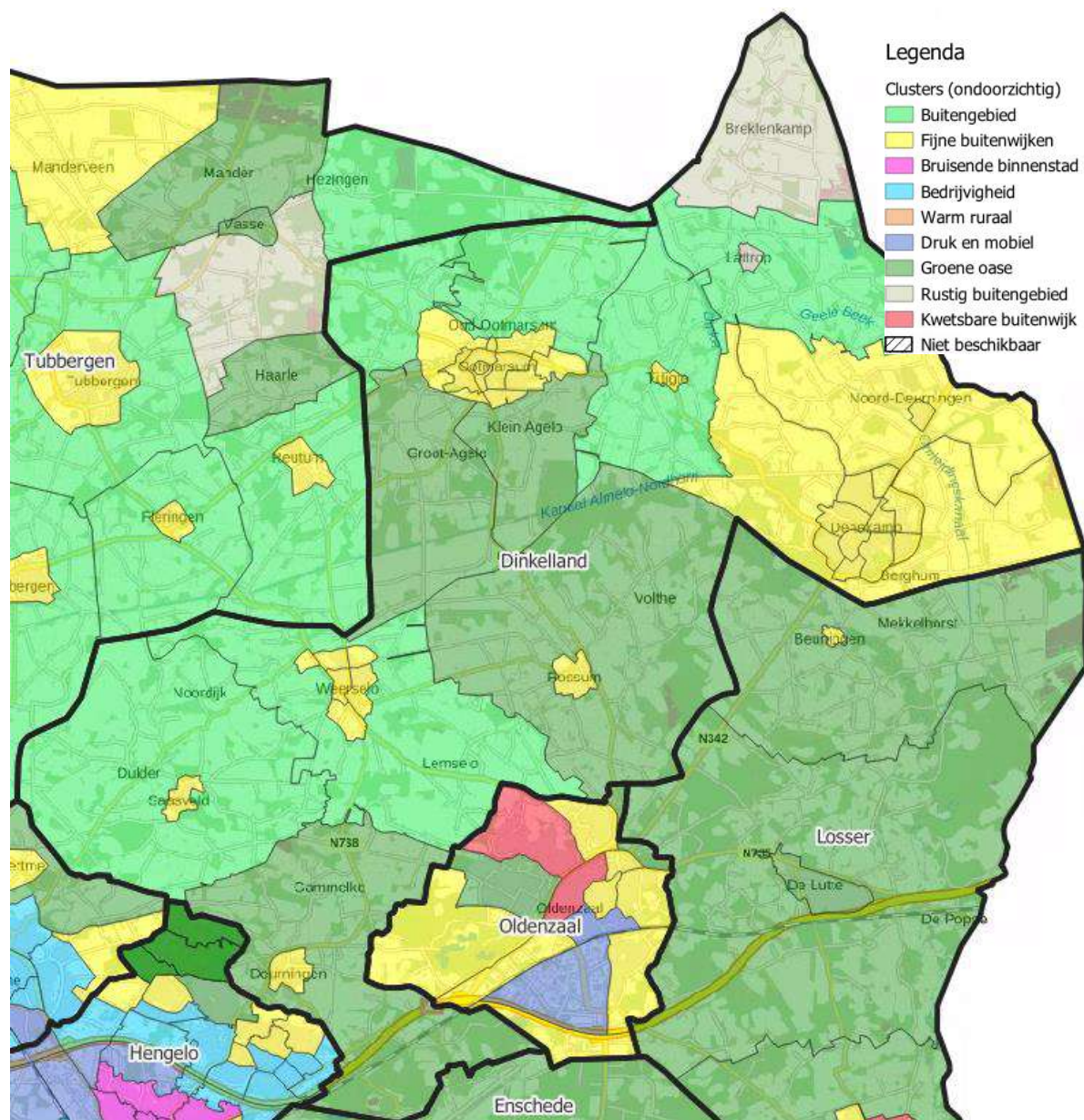
Clusterindelingen per CBS buurt in gemeente Almelo. Rondom de 'bruisende binnenstad' buurten is er een lapwerk van 'druk en mobiel' en 'bedrijvigheid' buurten'. Enkele losse buurten en buurten rondom Schelfhorst vallen onder 'fijne buitenwijken'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Borne



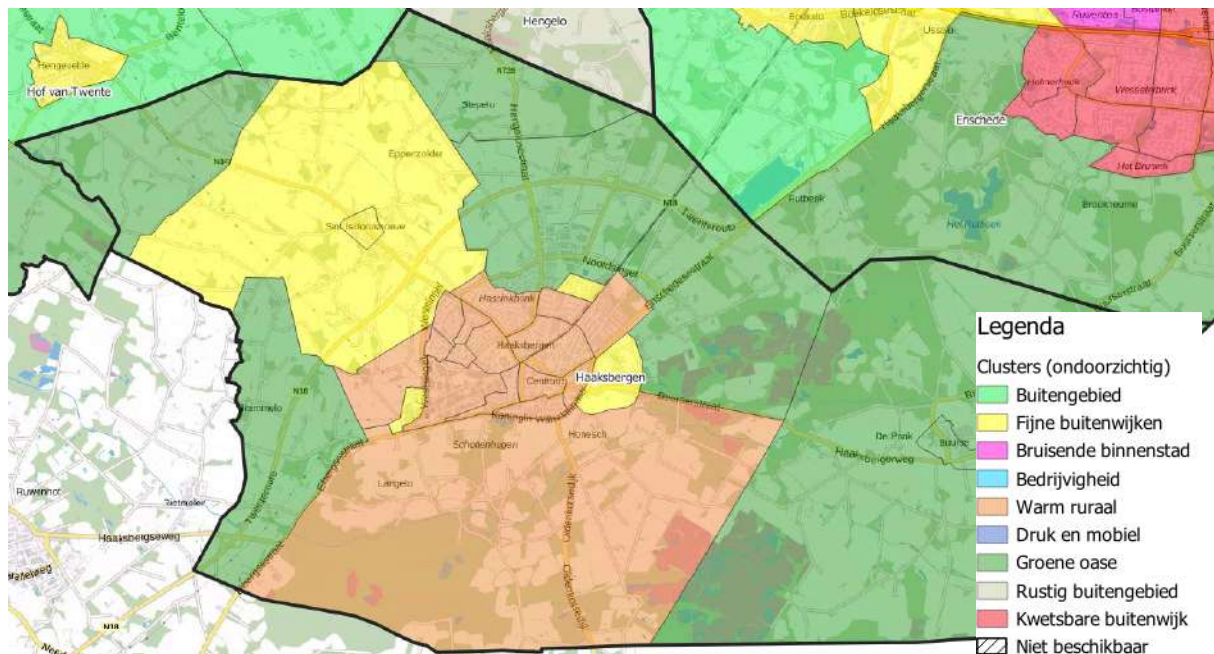
Clusterindelingen per CBS buurt in gemeente Borne. Buurt de Molenkamp en de dorpen Zenderen en Hertme vallen onder cluster 'fijne buitenwijken'. De rest van de gemeente heeft vooral buurten die onder cluster 'bedrijvigheid' vallen. 'Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Dinkelland



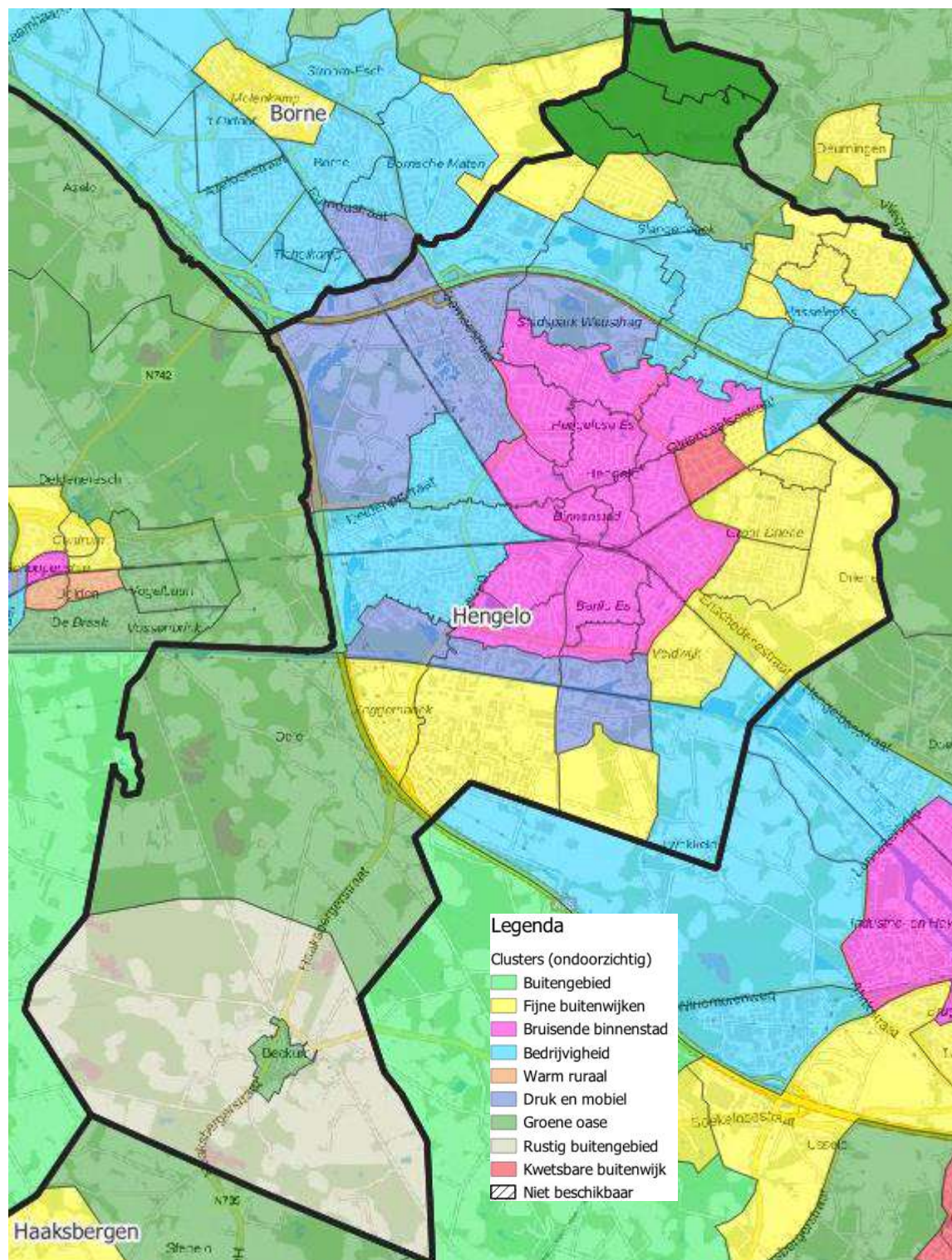
Clusterindelingen per CBS buurt in gemeente Dinkelland. De kernen van deze gemeente vallen onder cluster 'fijne buitenwijk'. Landelijke gebieden vallen onder clusters 'buitengebied' en 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Haaksbergen



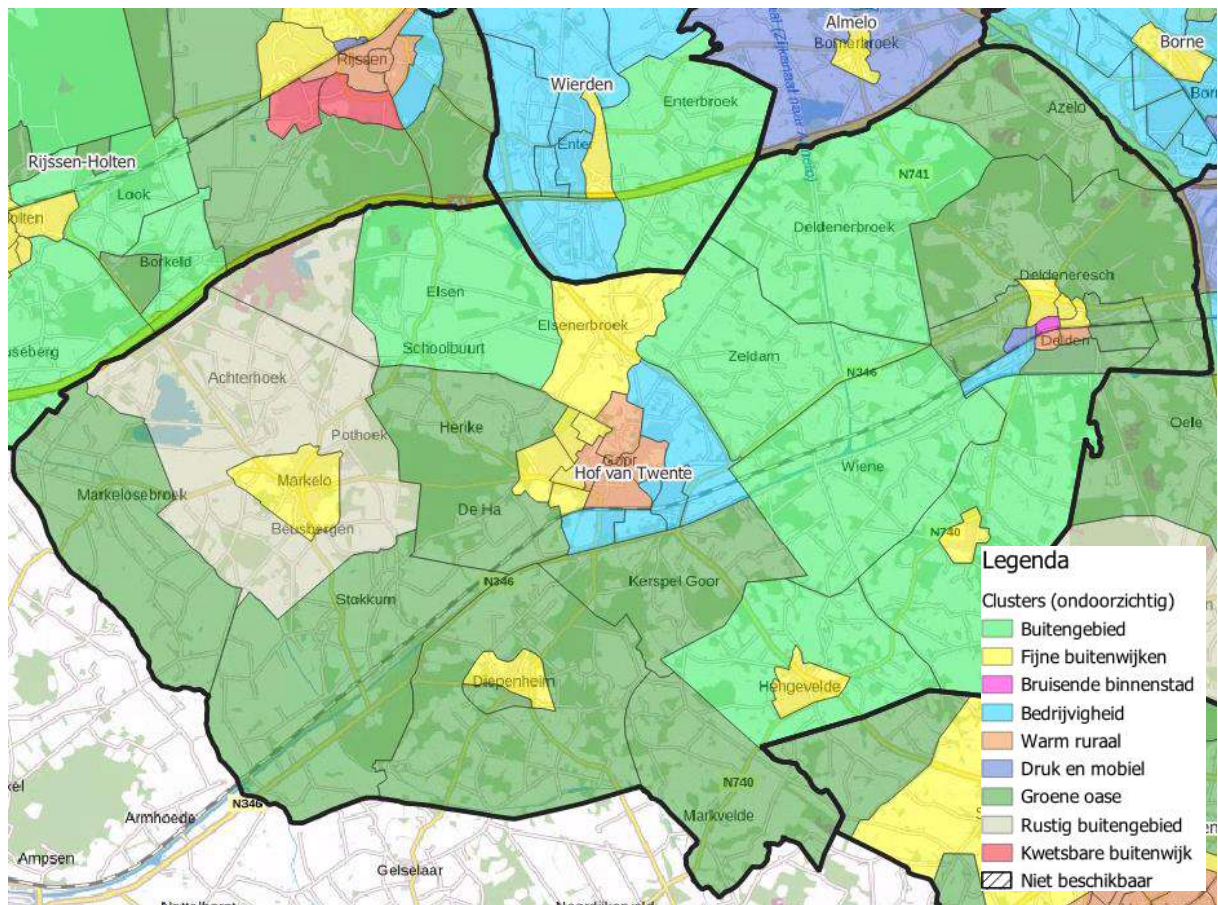
Clusterindelingen per CBS buurt in gemeente Haaksbergen. Buurten in Haaksbergen vallen vooral onder cluster 'warm ruraal'. In de rand van Haaksbergens zijn er enkele 'fijne buitenwijken'. Het buitengebied valt onder 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Hengelo



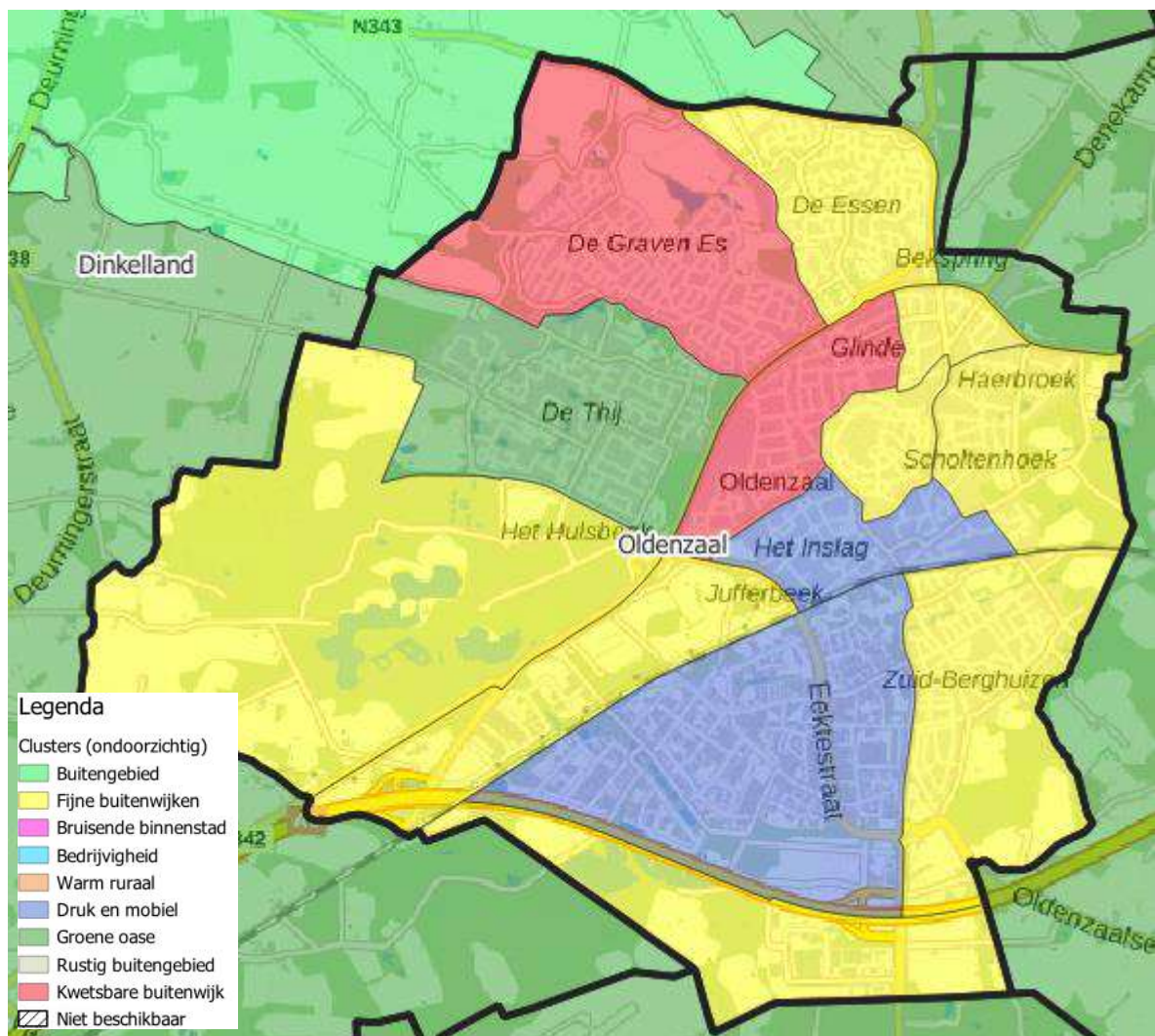
Clusterindelingen per CBS buurt in gemeente Hengelo. Rondom (afgezien van oost) van de 'bruisende binnenstad' buurten zijn er 'druk en mobiel' en 'bedrijvigheid' buurten te vinden. De rand van de stad in het zuiden en oosten bevat 'fijne buitenwijk' buurten. Buurt 'de Noork' valt onder cluster 'kwetsbare buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Hof van Twente



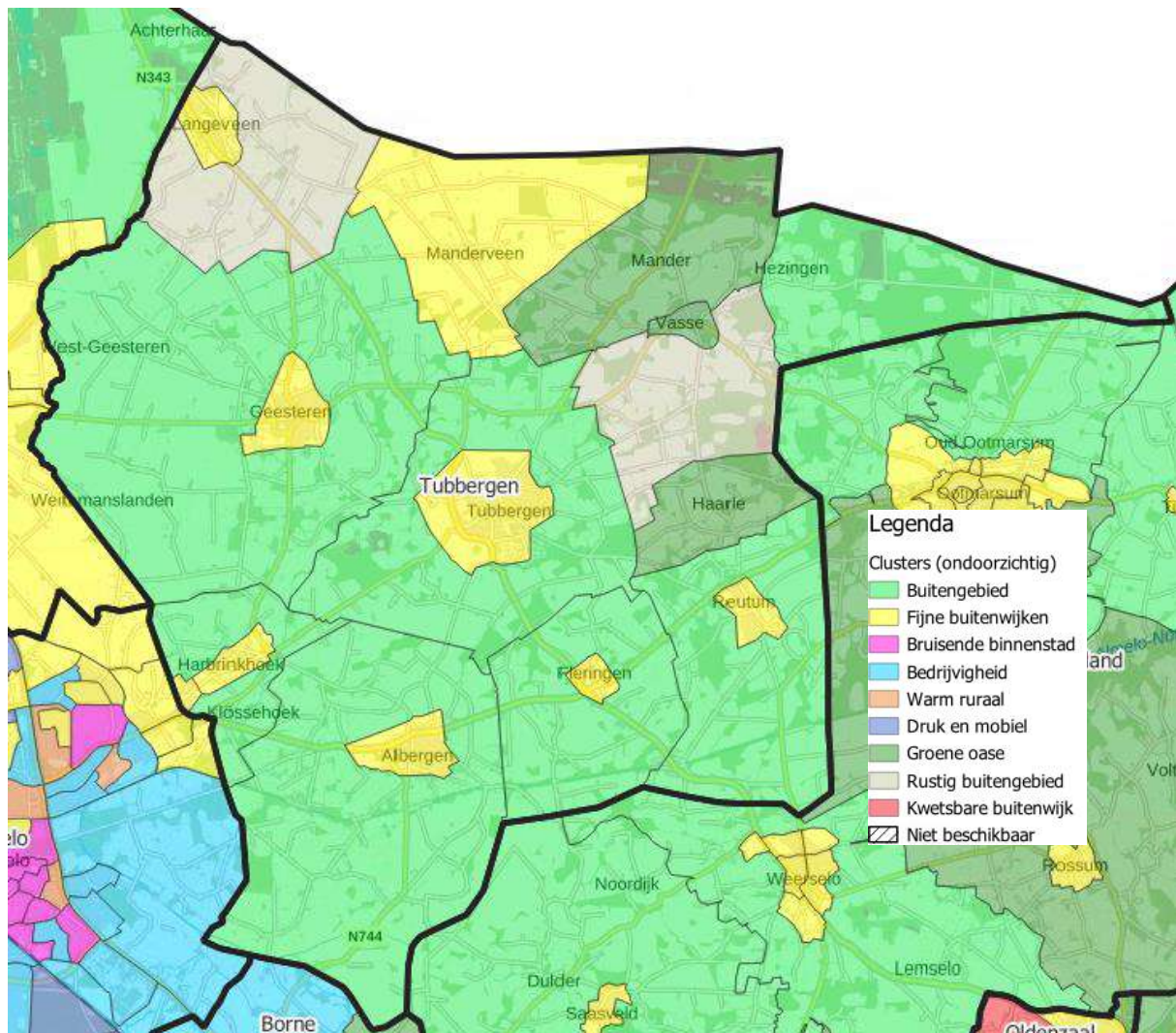
Clusterindelingen per CBS buurt in gemeente Hof van Twente. Goor en Delden hebben een combinatie van 'warm ruraal' en 'fijne buitenwijk' buurten. De kernen in het buitengebied vallen vooral onder 'fijne buitenwijken' waaromheen het landelijk gebied valt onder 'buitengebied' of 'groene oase'. Enkele (kleine) buurten dichtbij het spoor vallen onder 'bedrijvigheid'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Oldenzaal



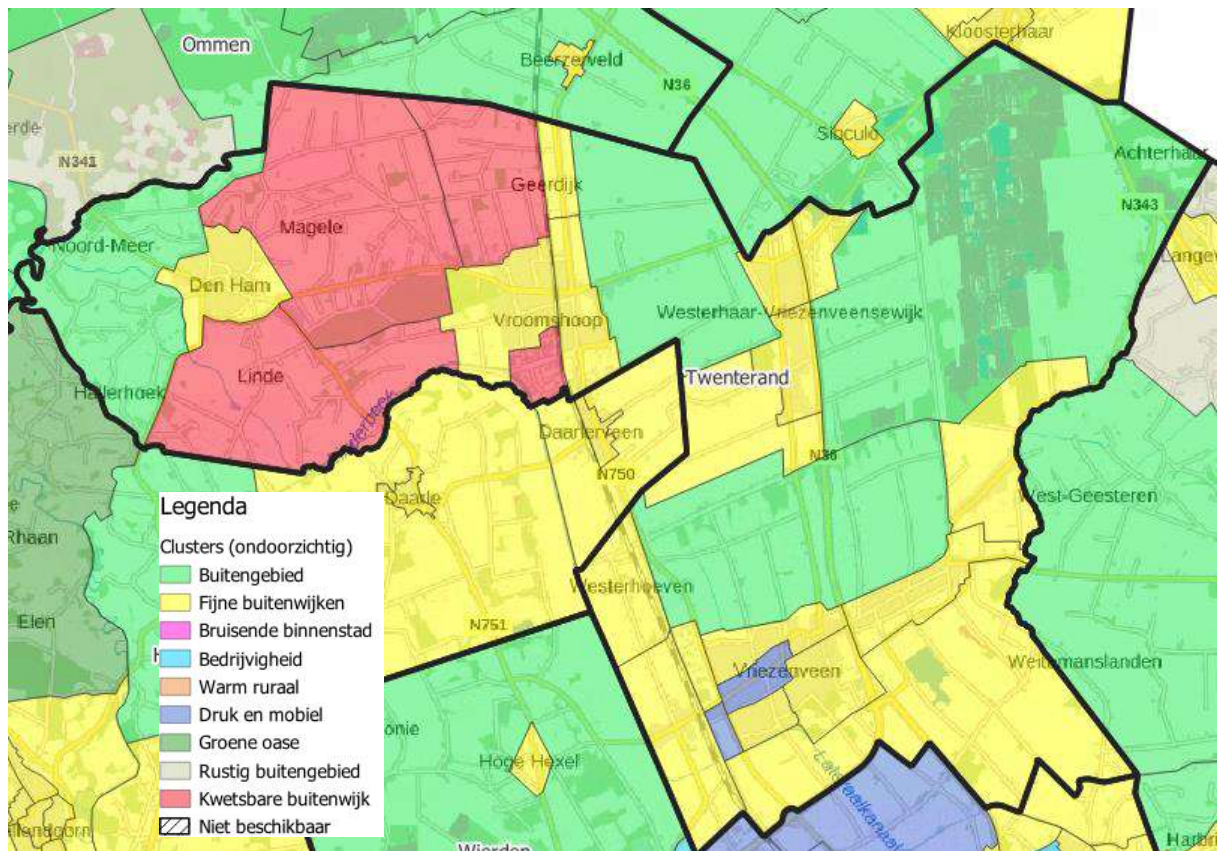
Clusterindelingen per CBS buurt in gemeente Oldenzaal. De kern van Oldenzaal bestaat uit twee 'kwetsbare buitenwijken' en twee 'druk en mobiel' buurten, met daaromheen een ring van 'fijne buitenwijken'. Buurt 'de Thij' wordt gekarakteriseerd als een 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Tubbergen



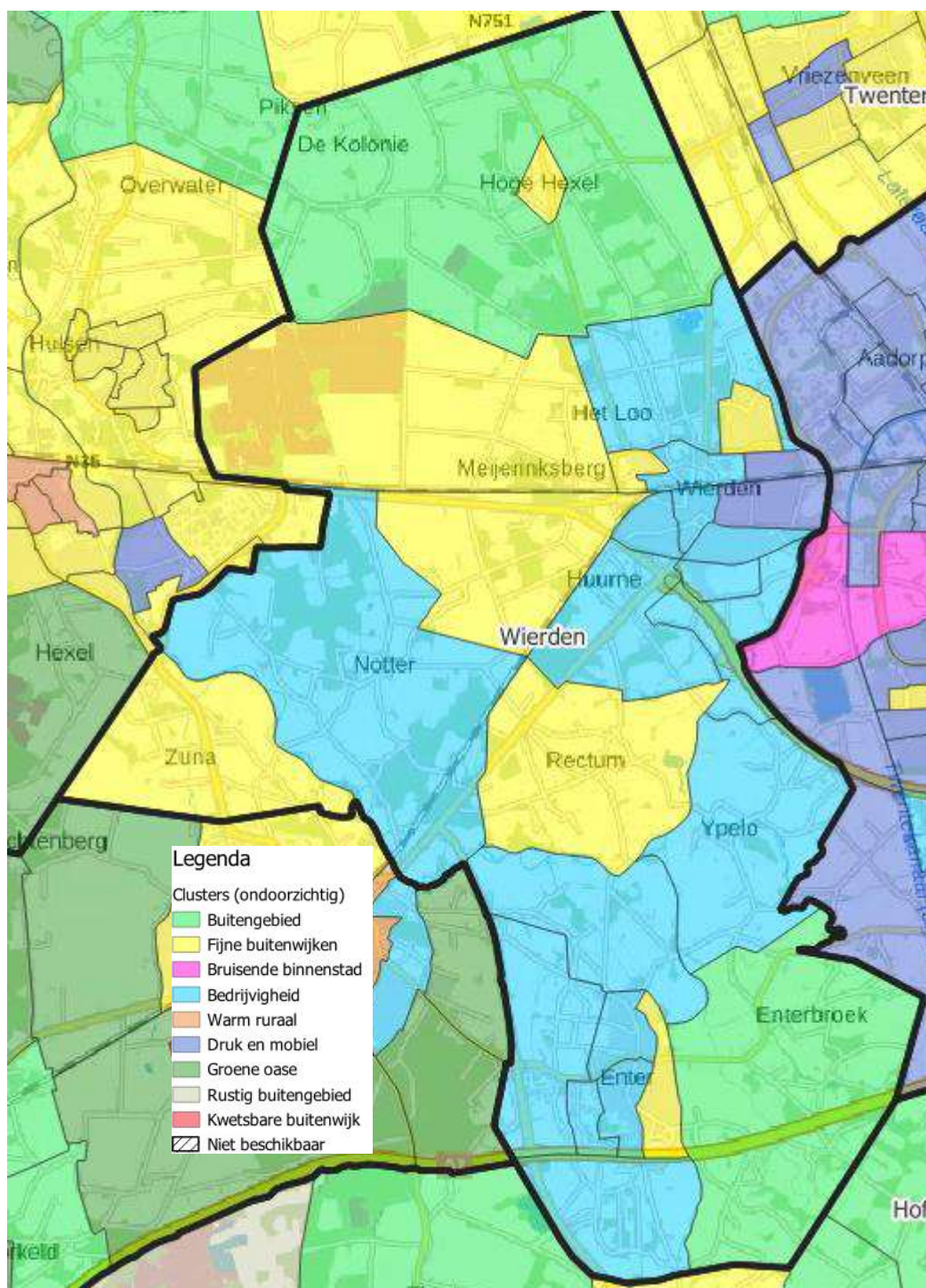
Clusterindelingen per CBS buurt in gemeente Tubbergen. Deze gemeente bestaat uit enkele 'fijne buitenwijken' omringt door (vooral) 'buitengebied' buurten. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Twenterand



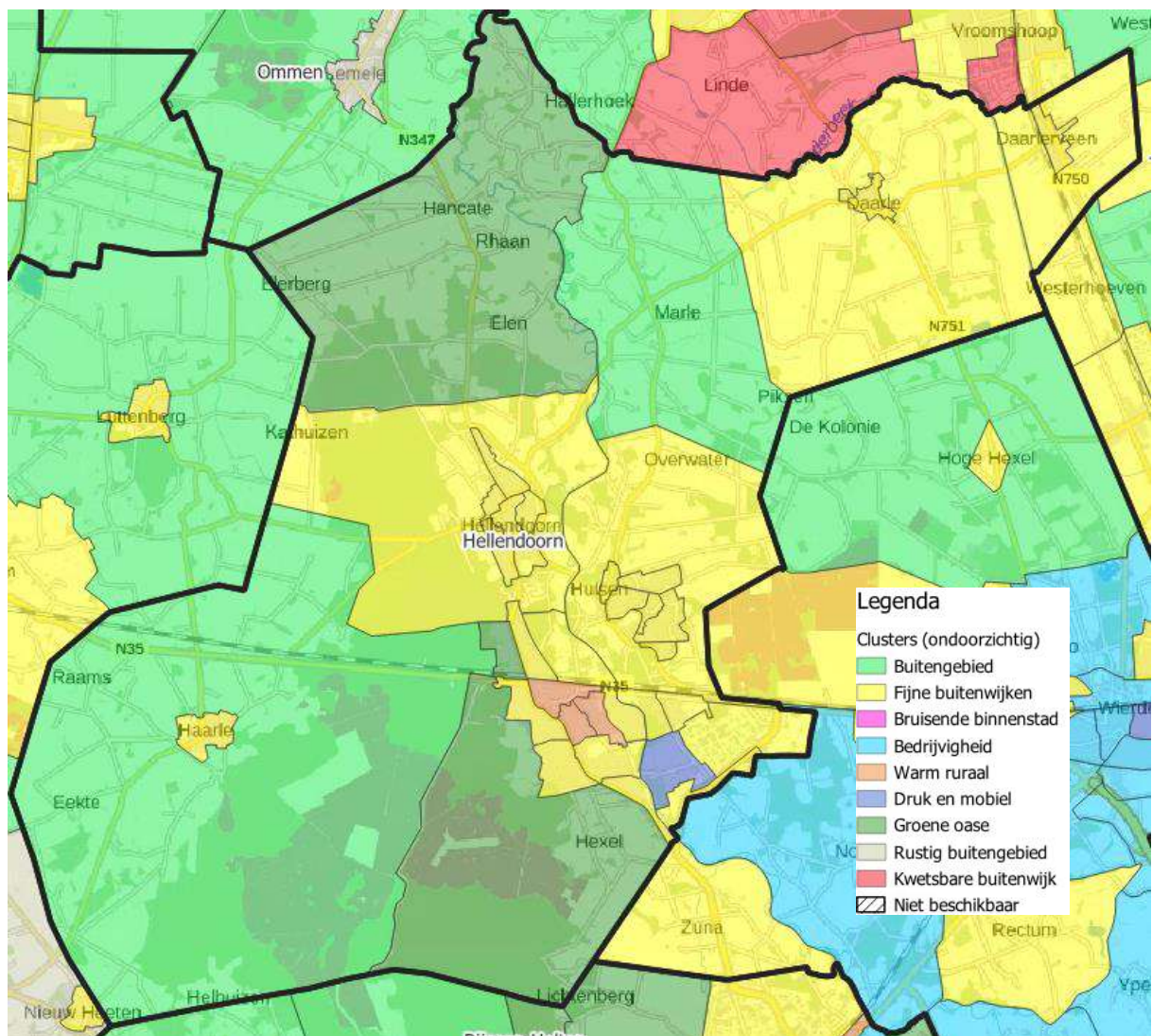
Clusterindelingen per CBS buurt in gemeente Twenterand. Er zijn vooral 'fijne buitenwijk' buurten en wat 'buitengebied'. De buurten tussen Den Ham en Vroomshoop vallen onder 'kwetsbare buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Wierden



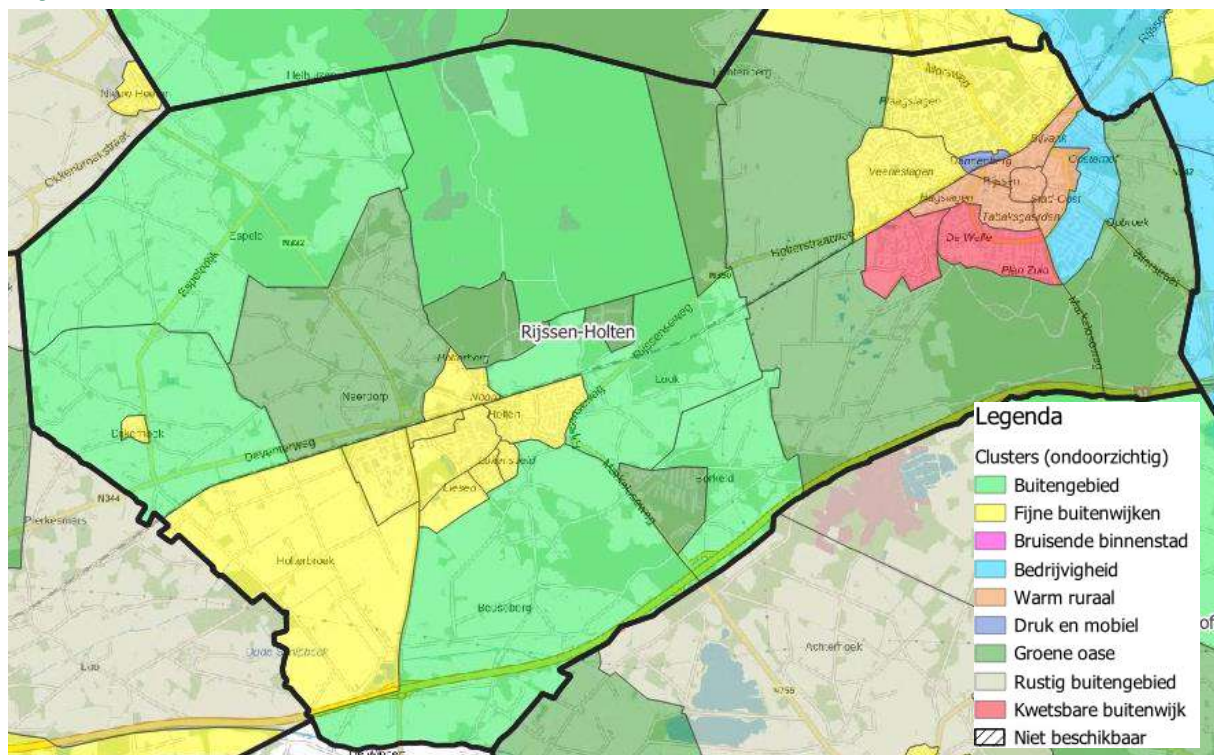
Clusterindelingen per CBS buurt in gemeente Wierden. Veel buurten vallen onder de clusters 'fijne buitenwijk' of 'bedrijvigheid'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Hellendoorn



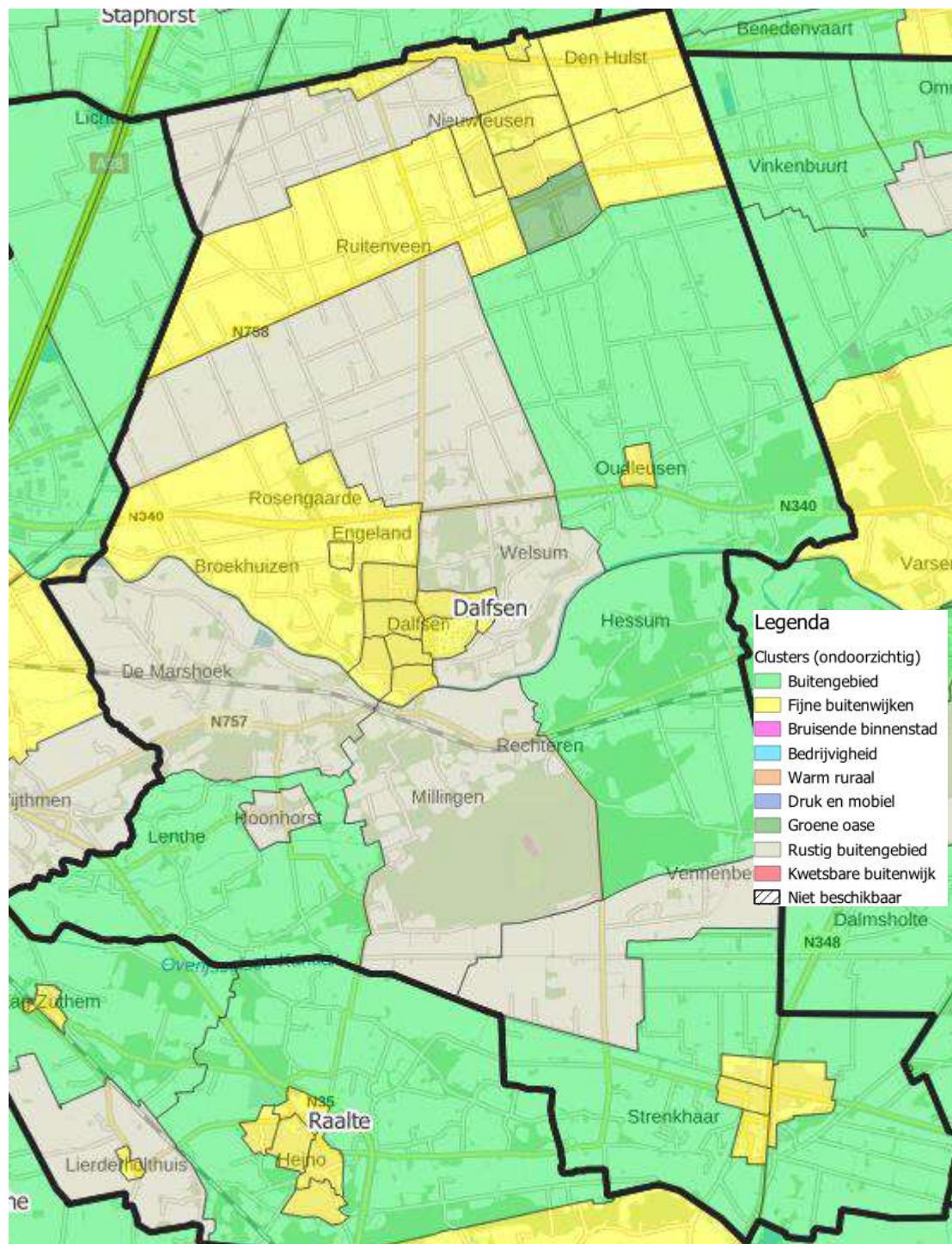
Clusterindelingen per CBS buurt in gemeente Hellendoorn. Buurten in Hellendoorn zelf en het gebied eromheen vallen vooral onder cluster 'fijne buitenwijk'. Buurten in Nijverdal vallen onder cluster 'warm ruraal'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Rijssen-Holten



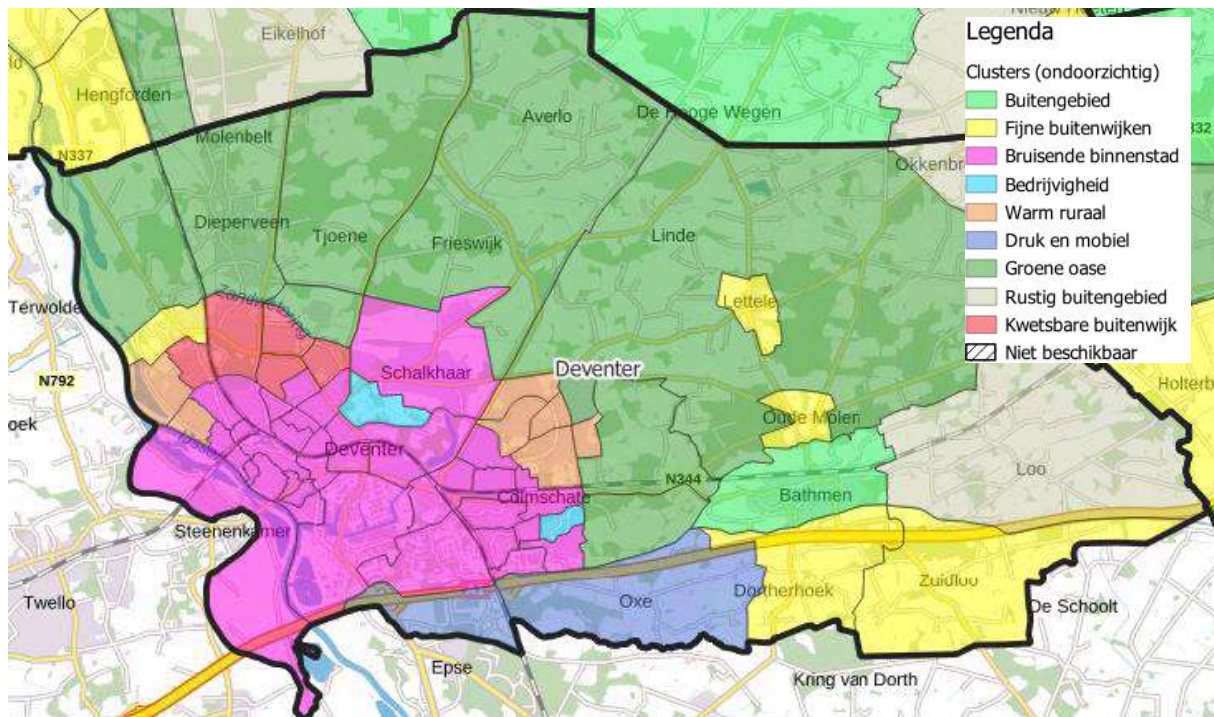
Clusterindelingen per CBS buurt in gemeente Rijssen-Holten. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden. Buurten in Holten vallen allen onder cluster 'fijne buitenwijk'. Het gebied rondom Holten en tussen Rijssen en Holten is in 'buitengebied' of 'groene oase'. De binnenstad van Rijssen valt onder 'warm ruraal', het noorden onder 'fijne buitenwijk' en het zuidwesten onder 'kwetsbare buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Dalfsen



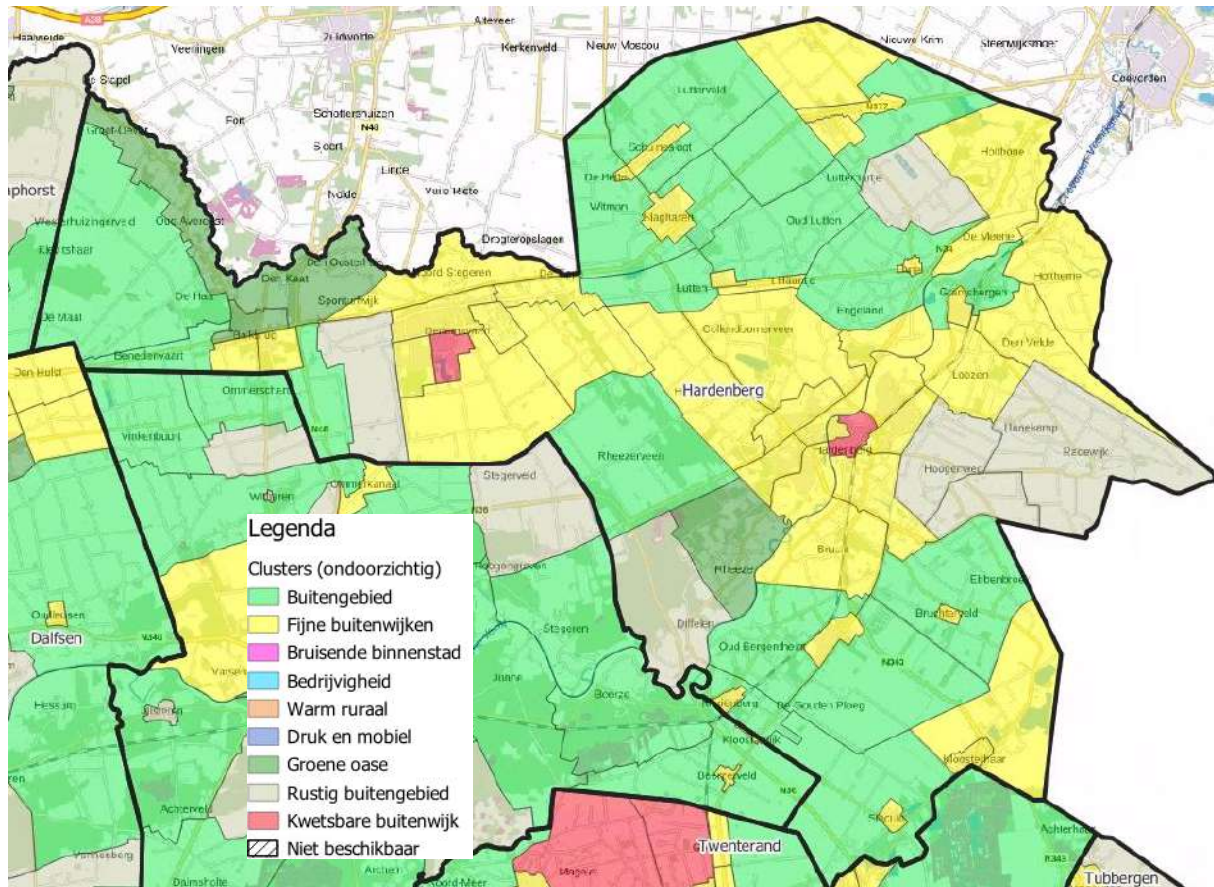
Clusterindelingen per CBS buurt in gemeente Dalfsen. Dalfsen en de kernen in het buitengebied vallen onder cluster 'fijne buitenwijk'. Deze worden omringd door buurten in clusters 'rustig buitengebied' en 'buitengebied'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Deventer



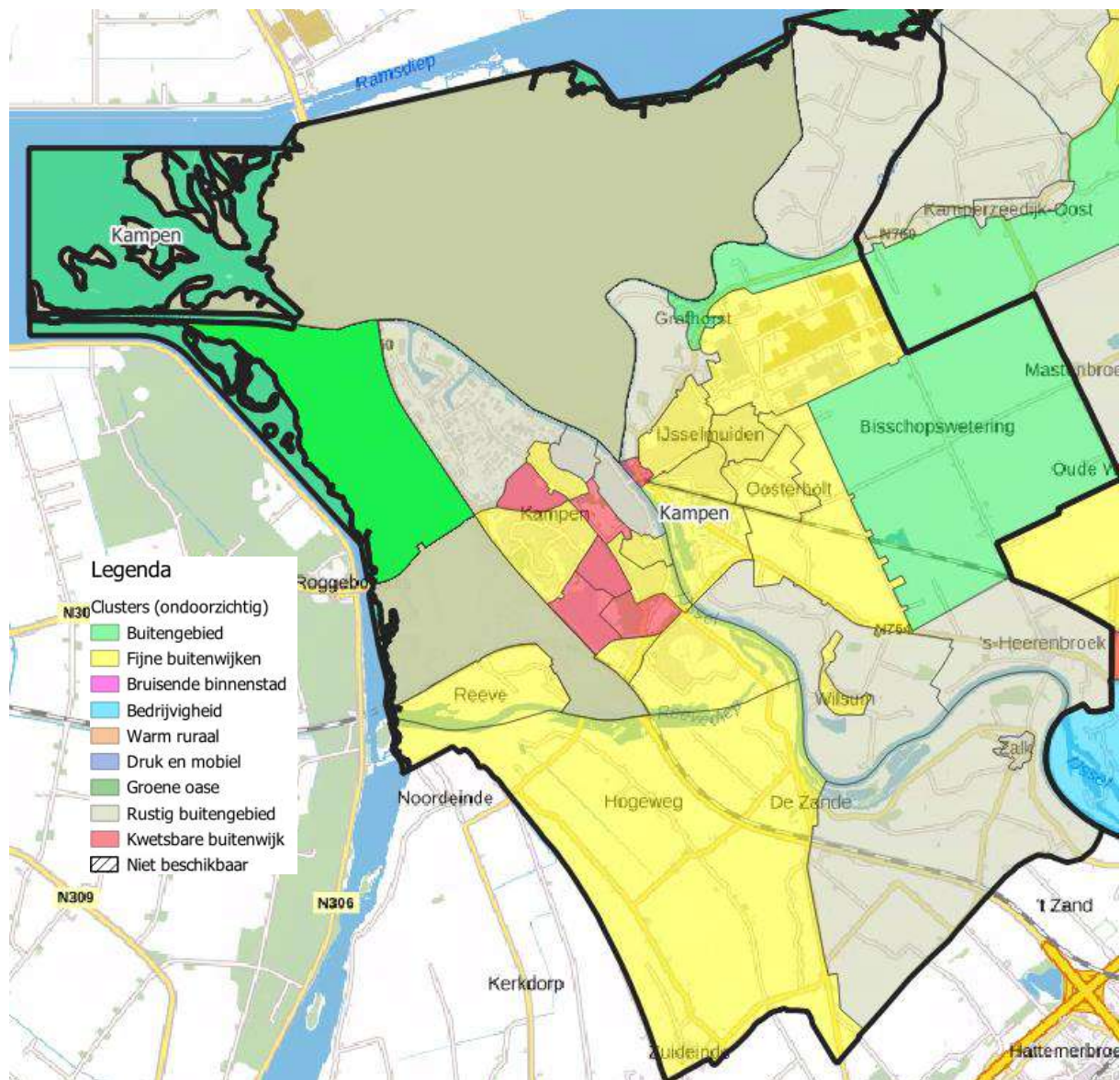
Clusterindelingen per CBS buurt in gemeente Deventer. Stad Deventer valt grotendeels onder 'bruisende binnenstad'. In het noorden zijn enkele buurten die onder cluster 'kwetsbare buitenwijk' vallen. Het buitengebied is vooral 'groene oase'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Hardenberg



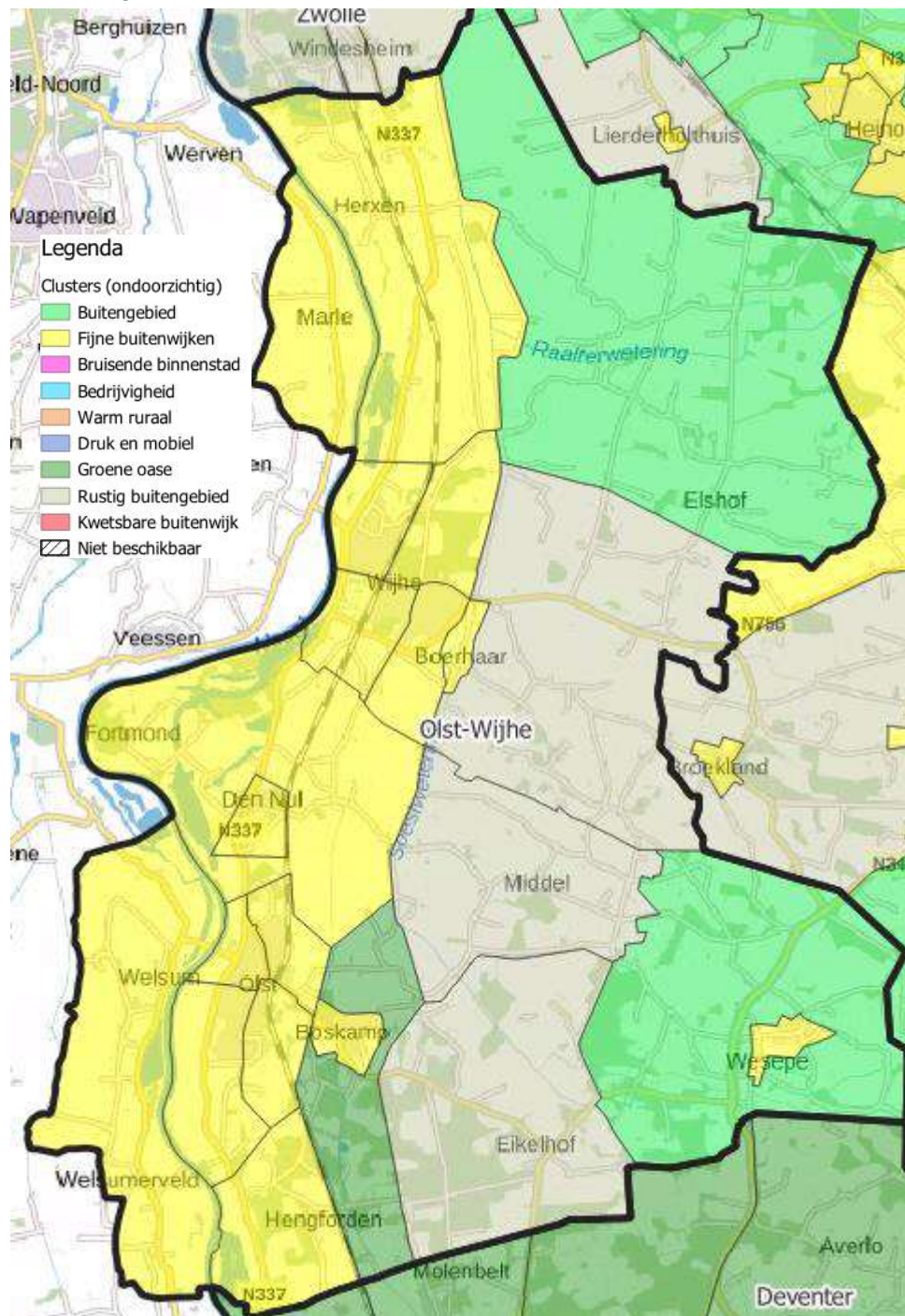
Clusterindelingen per CBS buurt in gemeente Hardenberg. Het centrum van Hardenberg valt onder cluster 'kwetsbare buitenwijk'. De rest van de gemeente valt onder 'fijne buitenwijk' of buitengebied. Dedemsvaart zuid-oost valt onder cluster 'kwetsbare buitenwijken'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Kampen



Clusterindelingen per CBS buurt in gemeente Kampen. De buurten in Kampen vallen vooral onder 'fijne buitenwijk' of 'kwetsbare buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

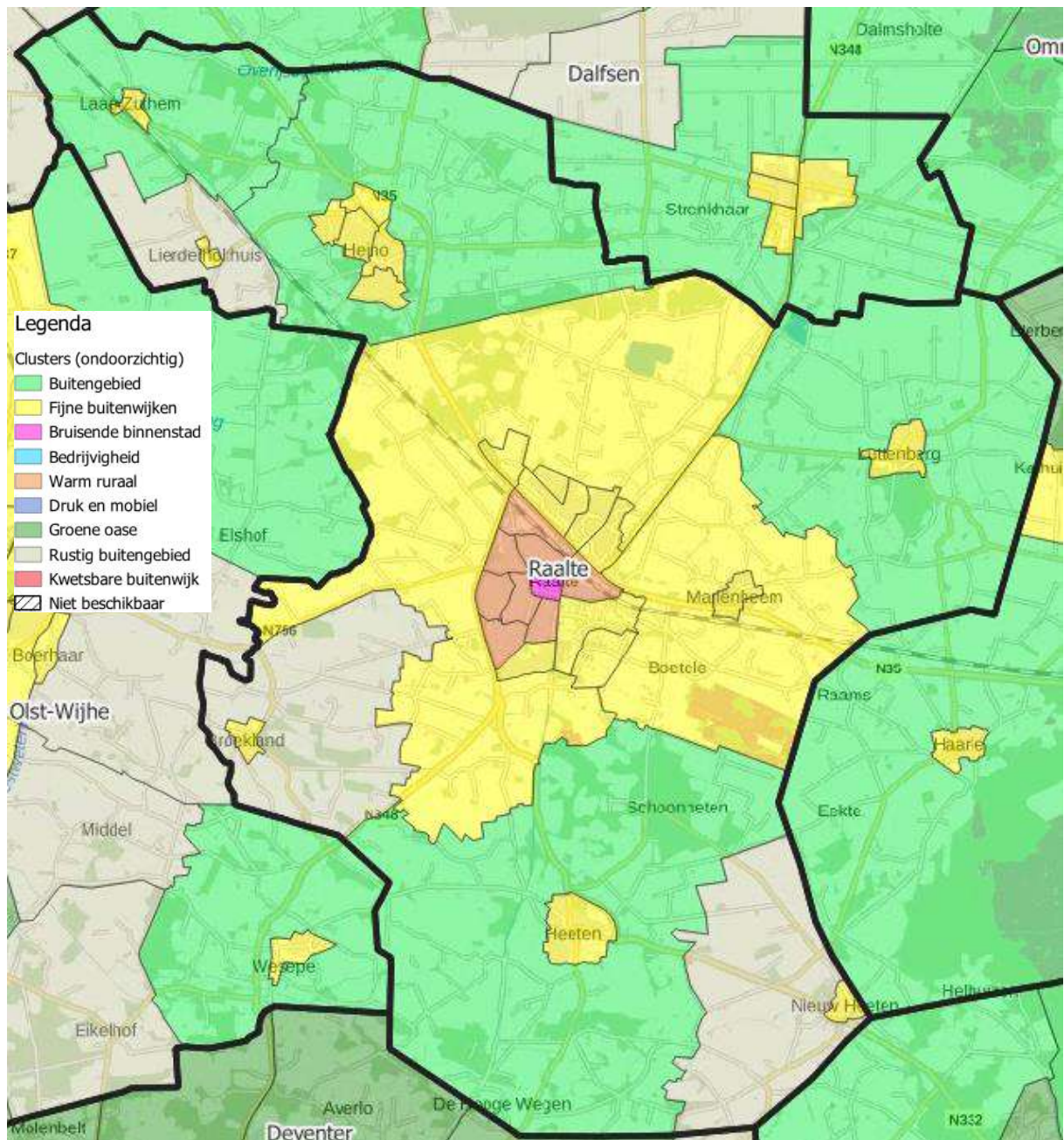
Olst-Wijhe



Clusterindelingen per CBS buurt in gemeente Olst-Wijhe. Buurten in het westen vallen allen onder het cluster 'fijne buitenwijk', in het oosten onder (rustig) 'buitengebied'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

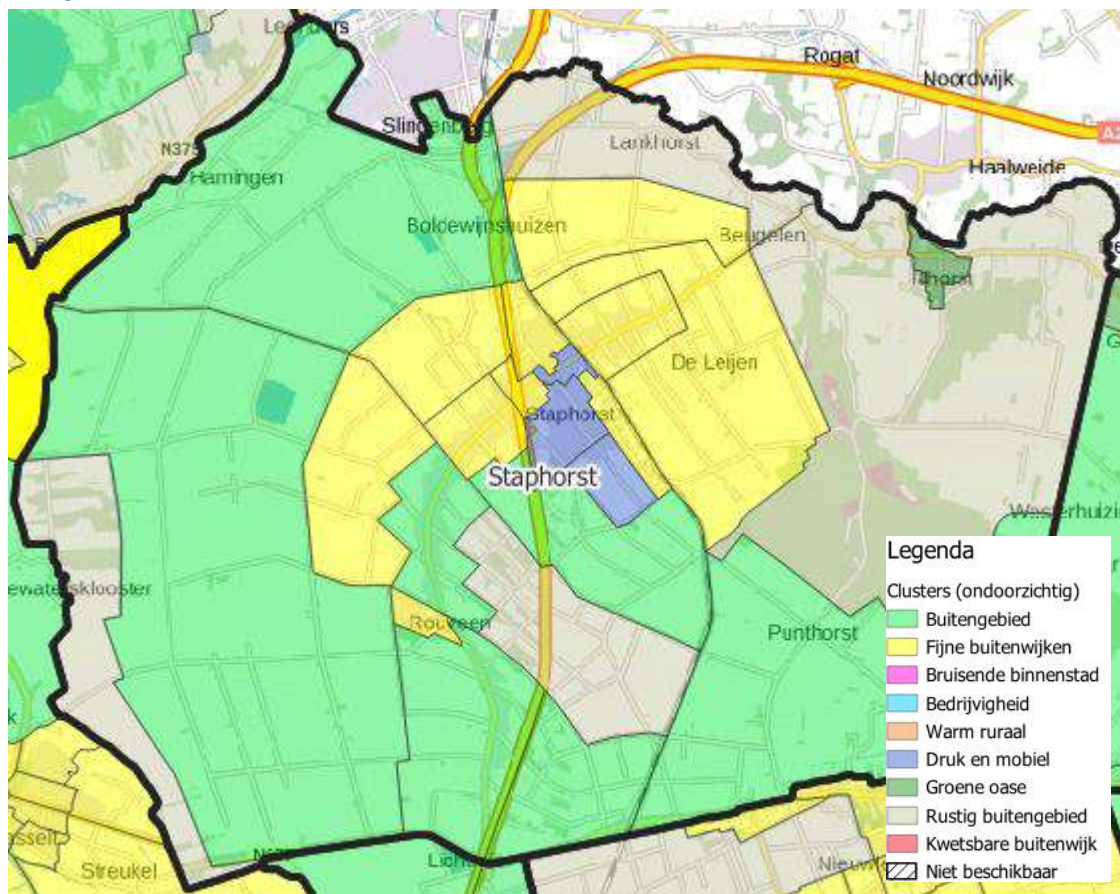


Raalte



Clusterindelingen per CBS buurt in gemeente Raalte. Het centrum van Raalte valt onder het cluster 'bruisende binnenstad', buurten daaromheen onder het cluster 'warm ruraal'. De rest van de gemeente valt vooral onder 'fijne buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Staphorst



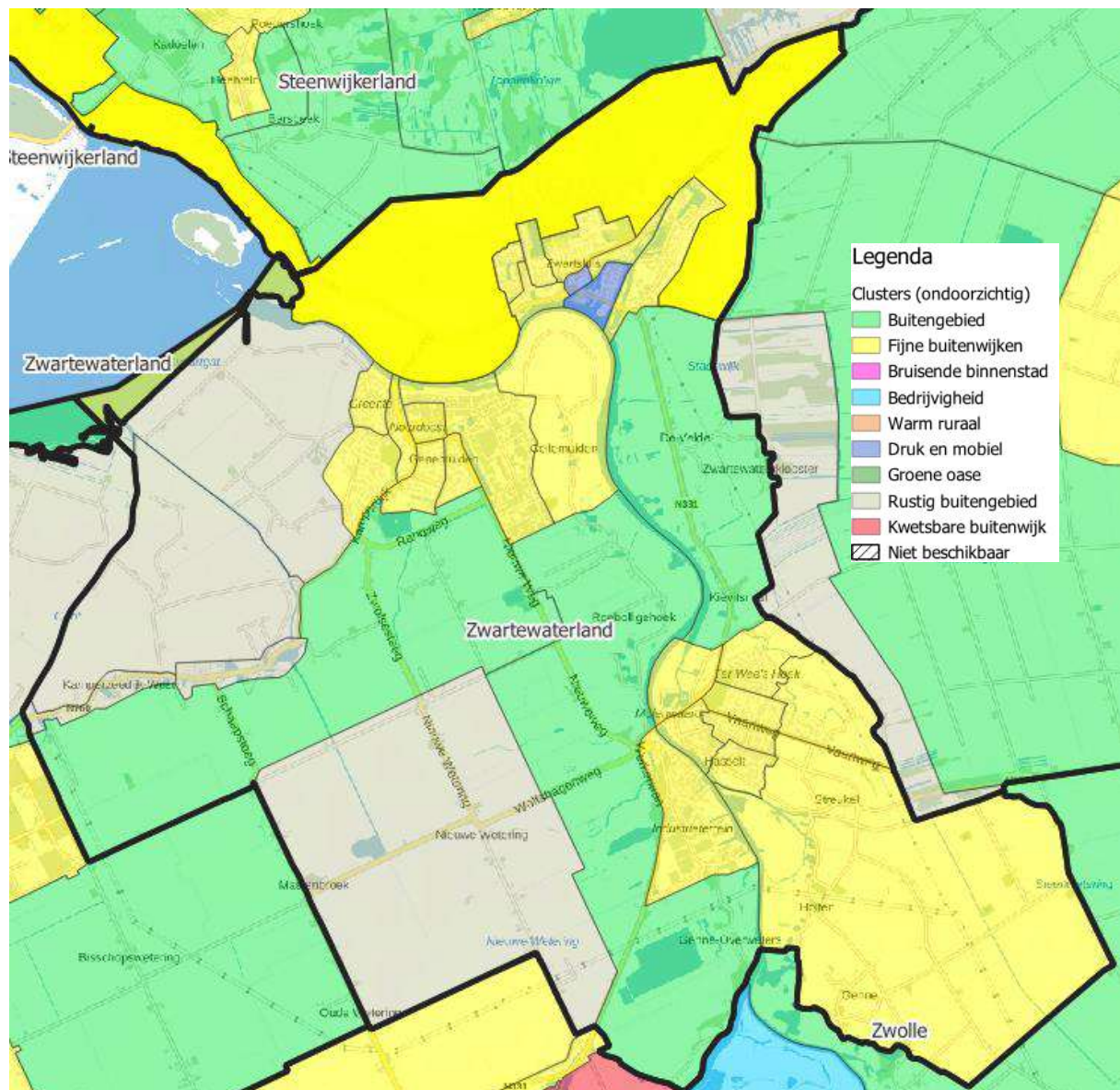
Clusterindelingen per CBS buurt in gemeente Staphorst. Buurten in Staphorst zelf vallen vooral onder de clusters 'fijne buitenwijken' en 'druk en mobiel'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Legenda

[Green]	Buitengebied
[Yellow]	Fijne buitenwijken
[Pink]	Bruisende binnenstad
[Blue]	Bedrijvigheid
[Orange]	Warm ruuraal
[Dark Blue]	Druk en mobiel
[Light Green]	Groene oase
[Grey]	Rustig buitengebied
[Red]	Kwetsbare buitenwijk
[Hatched Box]	Niet beschikbaar

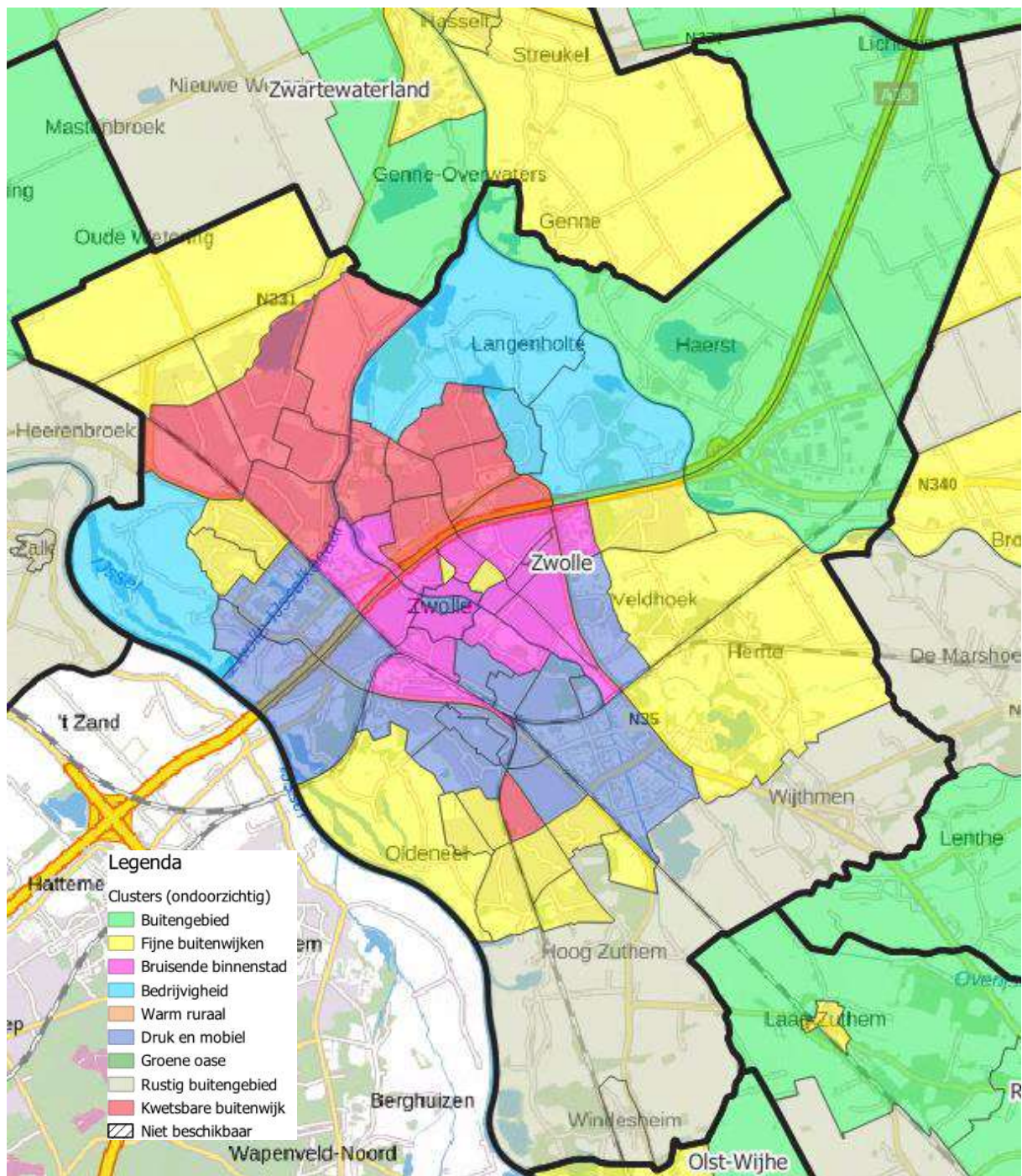


Zwartewaterland



Clusterindelingen per CBS buurt in gemeente Zwartewaterland. In Zwartsluis vallen twee buurten in het cluster 'druk en mobiel'. De rest van de gemeente valt vooral onder clusters 'fijne buitenwijken' en (rustig) 'buitengebied'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.

Zwolle



Clusterindelingen per CBS buurt in gemeente Zwolle. Ten noorden van de 'bruisende binnenstad' centrumbuurtten zijn er buurten die in het cluster 'kwetsbare buitenwijken' vallen. Buurten in het zuiden oosten van de stad vallen onder het cluster 'druk en mobiel' of 'fijne buitenwijk'. Meer uitleg over de clustereigenschappen is in hoofdstuk 3.3 te vinden.