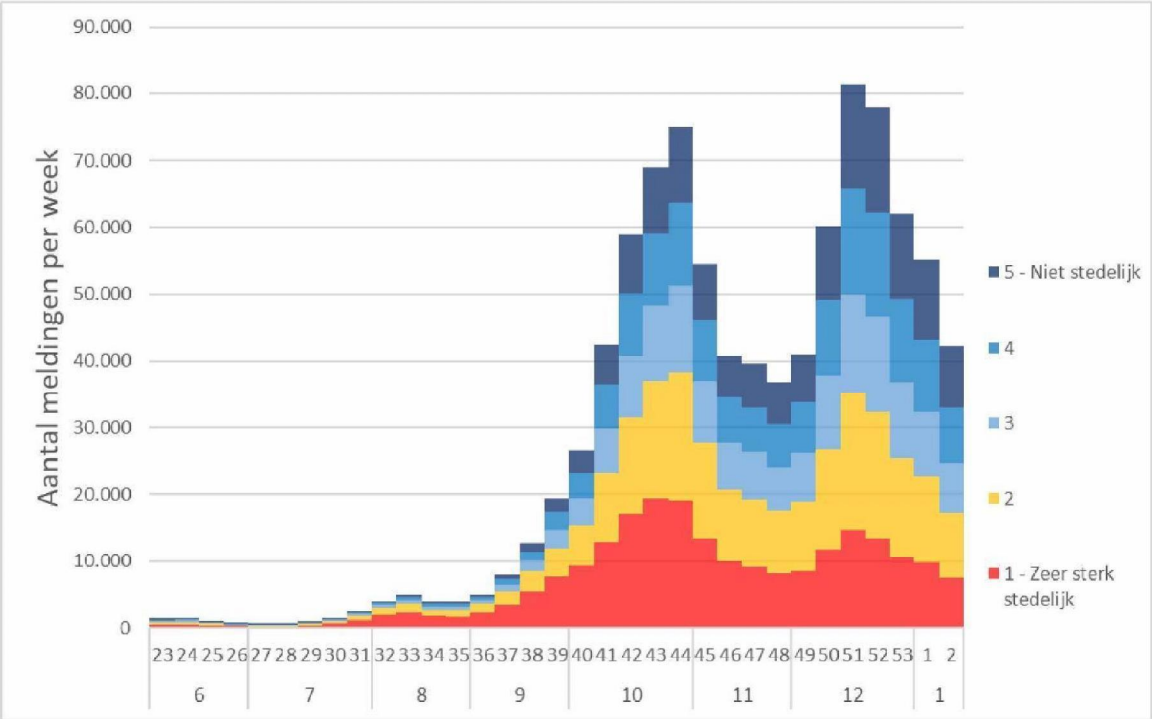
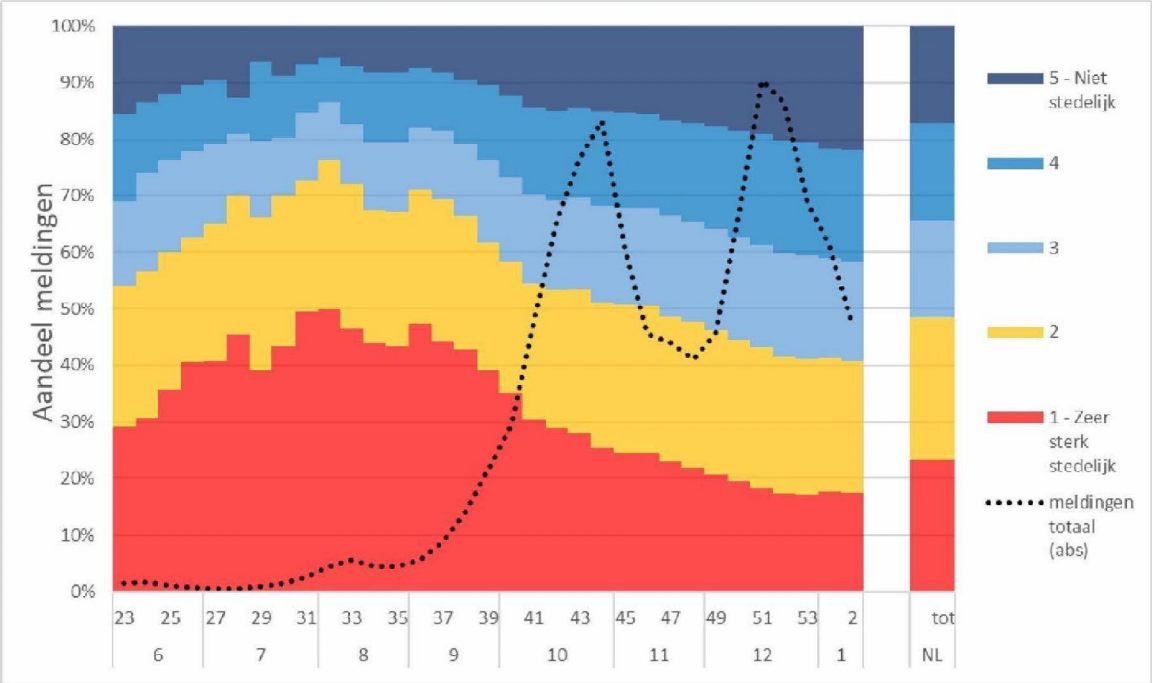
1. Meldingen naar Stedelijkheid

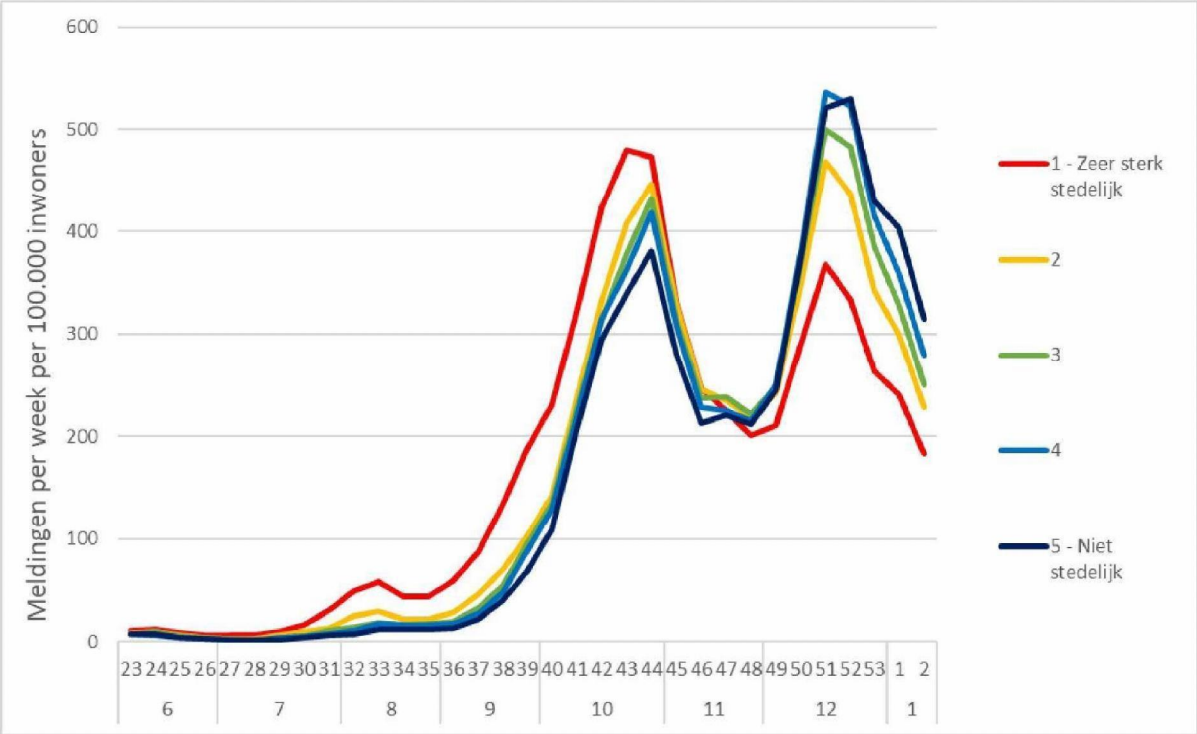
Omgevingsadressendichtheid, CBS-vierkanten (500 x 500m) 2019

Meldingen per pc6, gewogen voor afstand tot testlocatie.

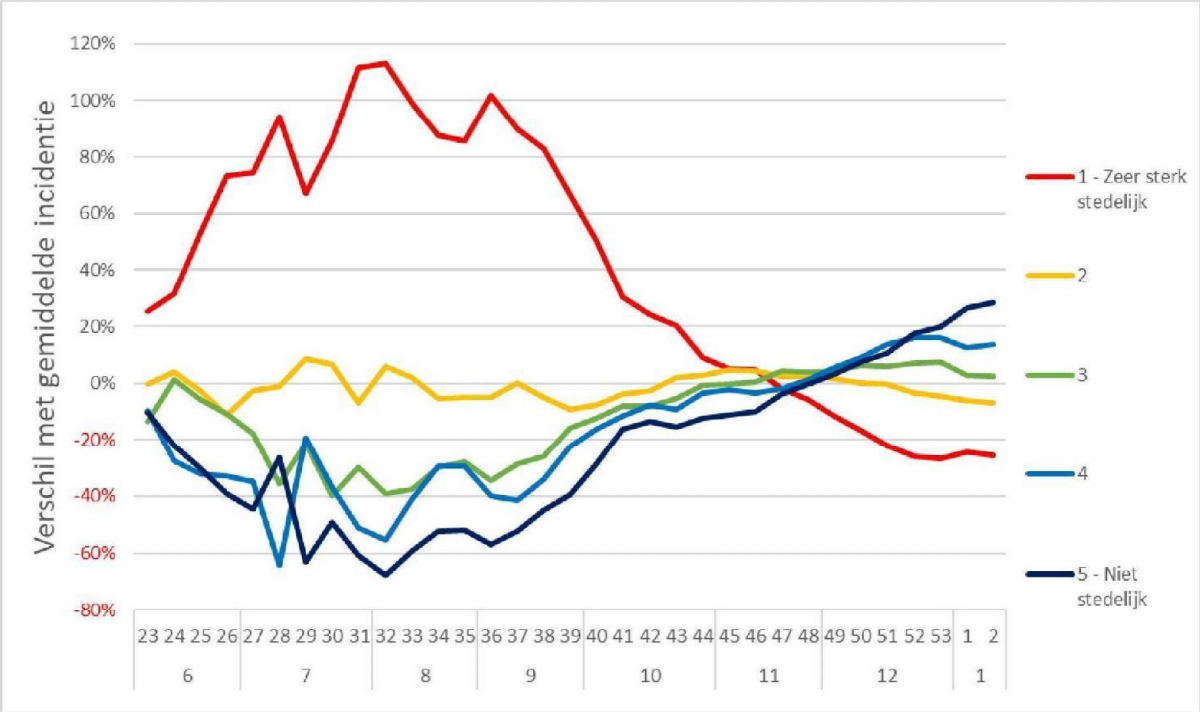




NB: De rechter stapelbalk geeft ter referentie de verdeling van Nederlandse bevolking aan.



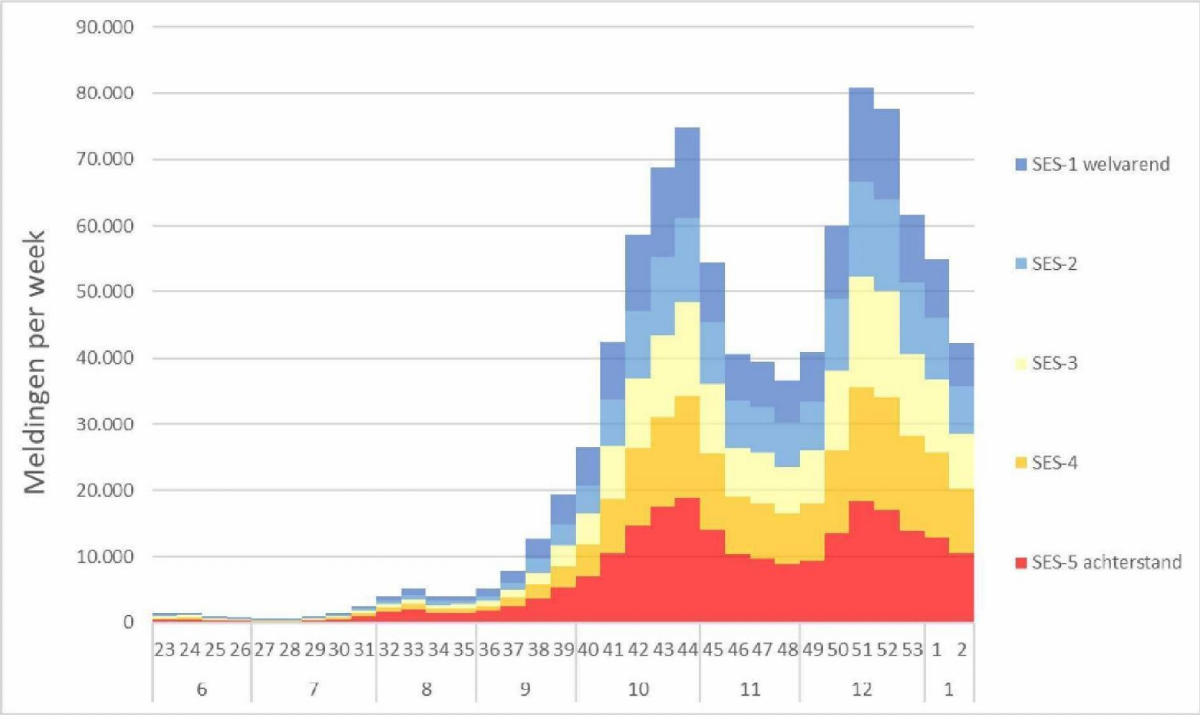
Stedelijkheidsklasse 1 (zeer sterk stedelijk) is ruim 4 maanden lang oververtegenwoordigd, t/m de top van de 2e golf. Maar de laatste 2 maanden juist ondervertegenwoordigd.

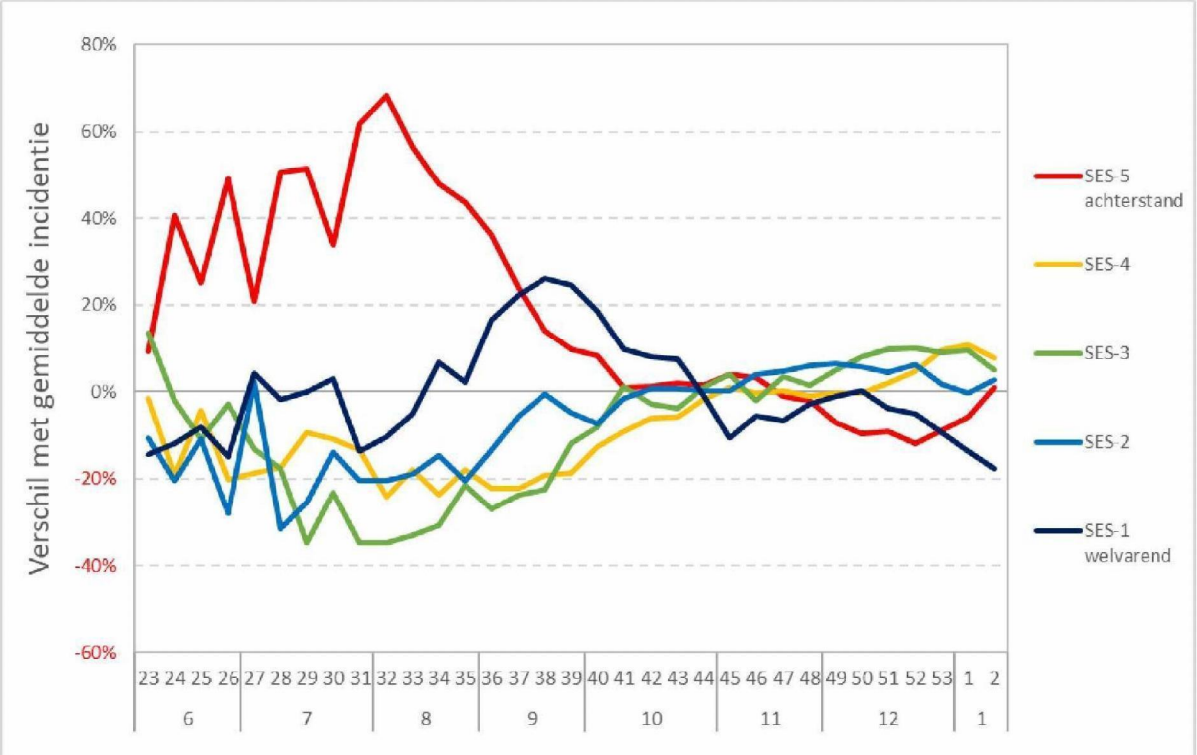


2. Meldingen naar SES

Socoiaal Economische Status (SCP, 2017), per pc4-gebied, ingedeeld in 5 kwintielen.

Meldingen per pc6, gewogen voor afstand tot testlocatie.



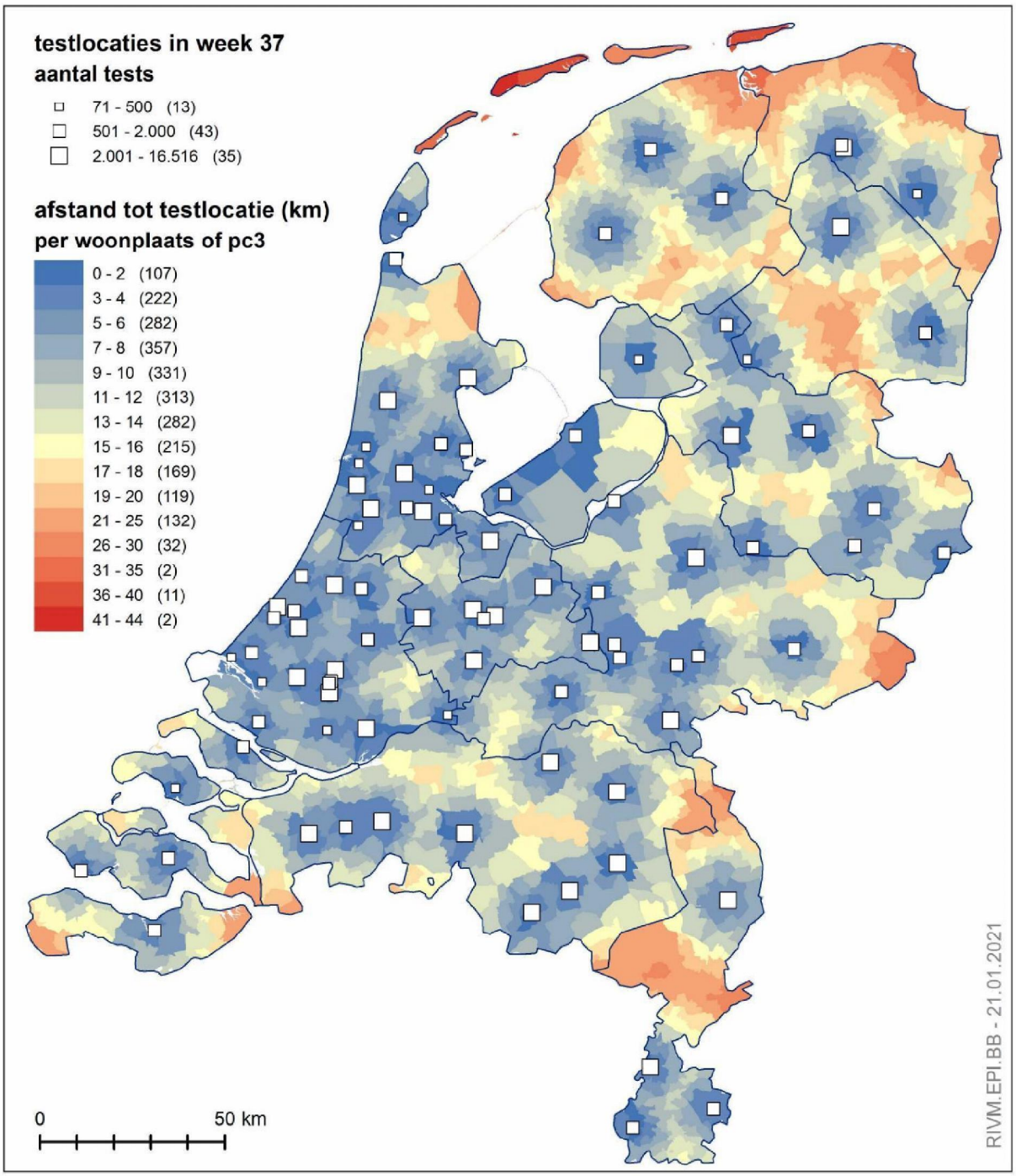


3. Invloed van afstand tot testlocatie

Beschrijving van de gevolgde methode om gewichten mbt afstand tot testlocatie te bepalen.

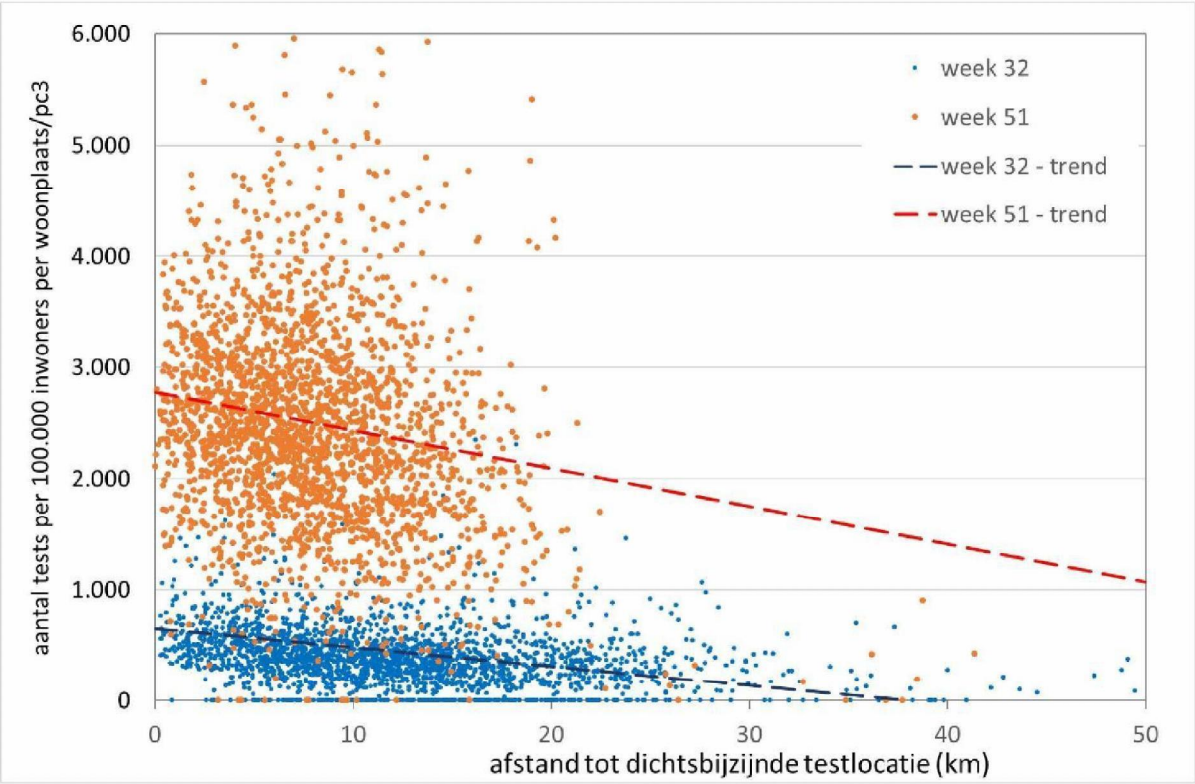
a) Allereerst is het lastig dat er steeds meer testlocaties bijkomen, maar soms ook weer verdwijnen of worden verplaatst. Ik kon dus niet alle tests en testlocaties van het hele jaar opgeteld gebruiken. Ik heb nu afstanden en testincidentie bepaald op weekbasis.

Zie bijgevoegd voorbeeldkaartje van de situatie in week 37.



b) Wanneer je de testincidentie uitzet tegen de afstand-tot-dichtstbijzijnde-testlocatie zie je altijd wel een verband, en kan er een regressielijn worden bepaald. Zie onderstaande grafiek met de scatterplot afstand x testincidentie voor week 32 en 51. De berekende trendlijn is gewogen obv het aantal inwoners.

Iedere stip is een woonplaats of een pc3-wijk van een grote stad.



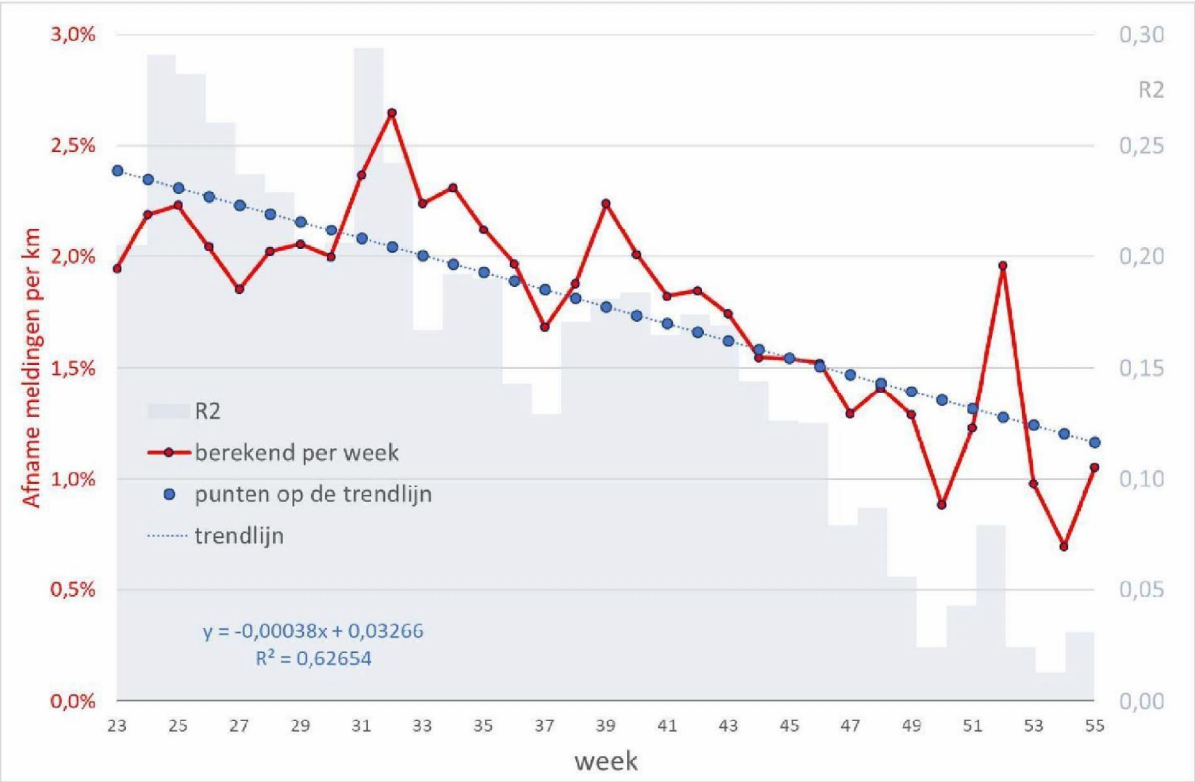
Het verband wordt in de loop van het jaar wel steeds zwakker. Waarschijnlijk omdat er steeds meer testlocaties bijkomen, en dus veel mensen meer op kortere afstand wonen en de reistijd geen groot probleem meer is. Maar misschien wordt de testbereidheid sowieso hoger door toenemend aantal besmettingen in alle gebieden

Zie bijv. het verschil tussen de weken in de scatterplot. De trendlijn van week 32 daalt relatief sterker, en de testincidentie zal na 38 km op het 0-punt zijn. In week 51 is dat pas na 81 km.

c) Om deze relatie te kunnen gebruiken als correctiefactor bij het aantal meldingen ben ik op zoek naar een correcte formule hiervoor.

Door de verschillen in de totale testincidentie per week (hoogte op de y-as) verschilt helaas ook de afgeleide regressieformule (y =a*km + b). Om te kijken hoe steil de regressielijn daalt heb ik daarom het "afnamepercentage" (=b/a) genomen. *(Wist even de juiste term hiervoor niet.)*

En deze waarde is wel goed vergelijkbaar over de verschillende weken. Gemiddeld ligt dit afnamepercentage op 1,75%. Dit komt neer op een daling van de testincidentie met 17,5 ppt voor elke 10 km; en dat bij afstanden boven de 57 km niemand meer de moeite neemt om zich te laten testen. Dit afnamepercentage wordt per week wel steeds lager. Zie figuur.



Ik wil rekening mee houden met deze daling ipv één uniforme waarde te gebruiken. Om de getoonde grilligheid van de lijn te vermijden gebruik daarbij ik niet de per week berekende waarde, maar de waarde op de ingetekende trendlijn.

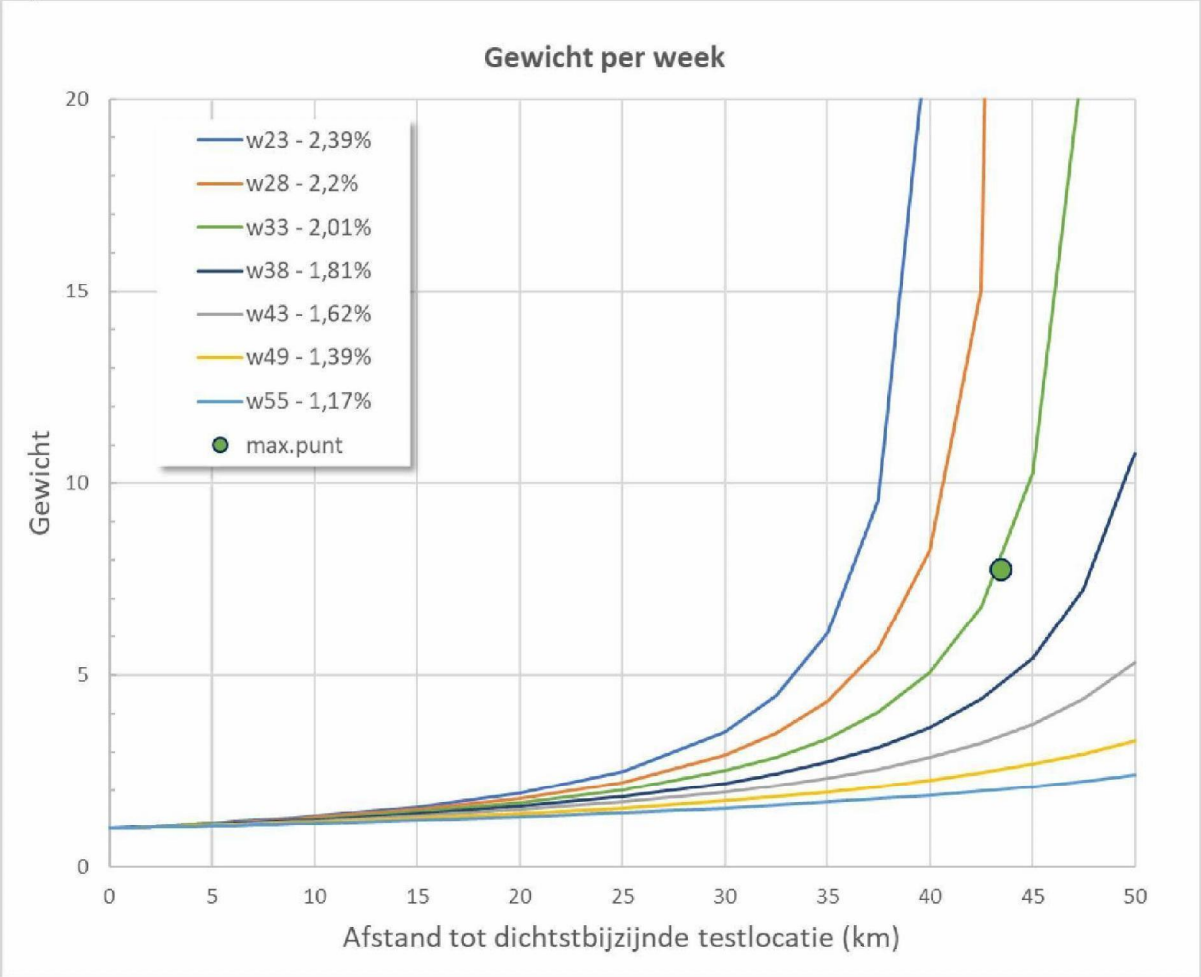
Ook de betrouwbaarheid wordt steeds lager: de R2 van de regressielijn is gezakt van 0,291 in week 24 naar 0,024 in week 53. Zie de grijze vlakken in bijgevoegd figuur. Op de scatterplots zie je in de latere weken eerder een diffuse puntenwolk dan punten mooi nabij een lijn. Het is dus nog maar de vraag of in de laatste weken van het jaar dit eigenlijk wel betrouwbaar gebruikt kan worden. Daarentegen zijn dan de afnamepercentages lager, én door de vele testlocaties wonen meer mensen op kortere afstand. De berekende gewichten zullen daardoor in de meeste gevallen erg laag zijn (dichtbij 1), waardoor de invloed van afstand uiteindelijk niet zo groot meer zal zijn.

d) De gedachte is nu om dit verschijnsel te gebruiken om de meldingen te wegen: hoe verder weg van een testlocatie, hoe hoger het gewicht van de aanwezige cases.

Om de "verdwenen" personen te compenseren kom ik uit op de volgende formule:

Gewicht = $1/(1\text{-afstand}_{km}*\text{afnamepercentage}_{week})$.

Dit gewicht verschilt dus per week en met afstand. Zie onderstaande figuur voor de op deze manier berekende gewichten voor 7 verschillende weken. In de eerste weken van het jaar liggen de gewichten hoger dan in de latere weken.



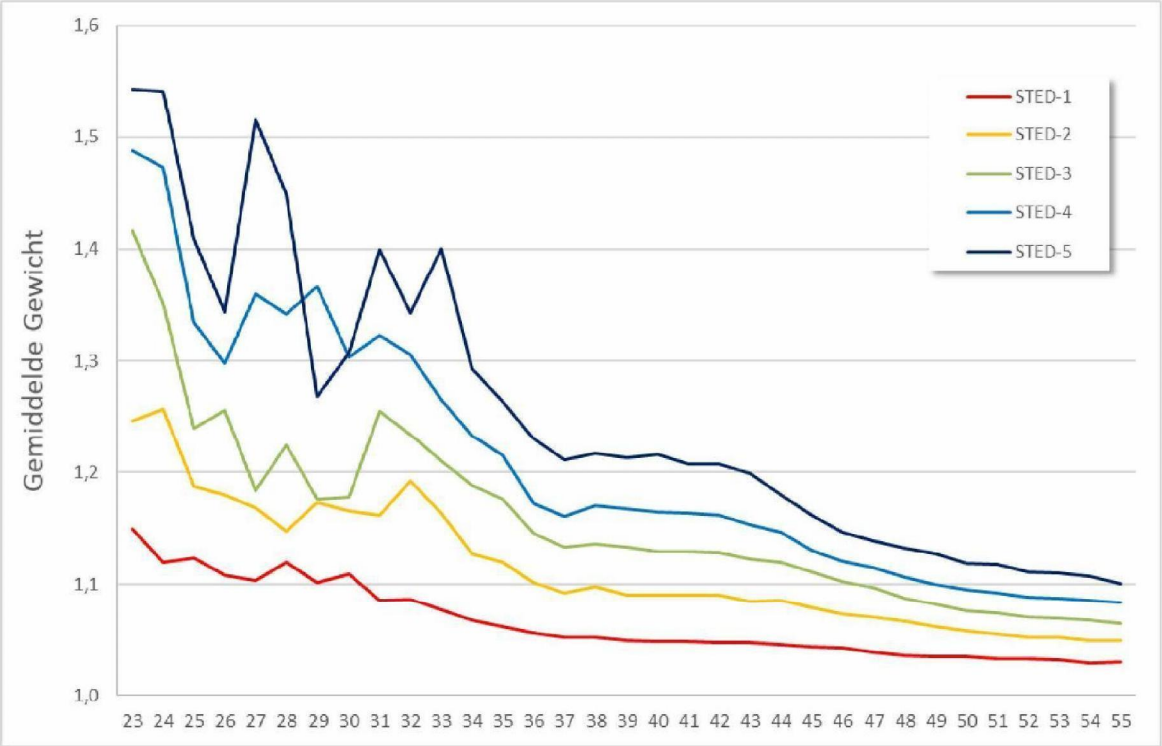
Door het exponentiele karakter van deze formule gaat op grotere afstanden het gewicht wel flink omhoog. Dat lijkt me misschien niet zo realistisch. Wellicht verstandig om een max waarde in te stellen, bijv. gewicht=10

In de dataset met meldingen bleek het maximale voorkomende gewicht op 7,5 te liggen, bij een melding in week 33 op een afstand van 43 km tot een testlocatie. Zie ingetekend bolletje.

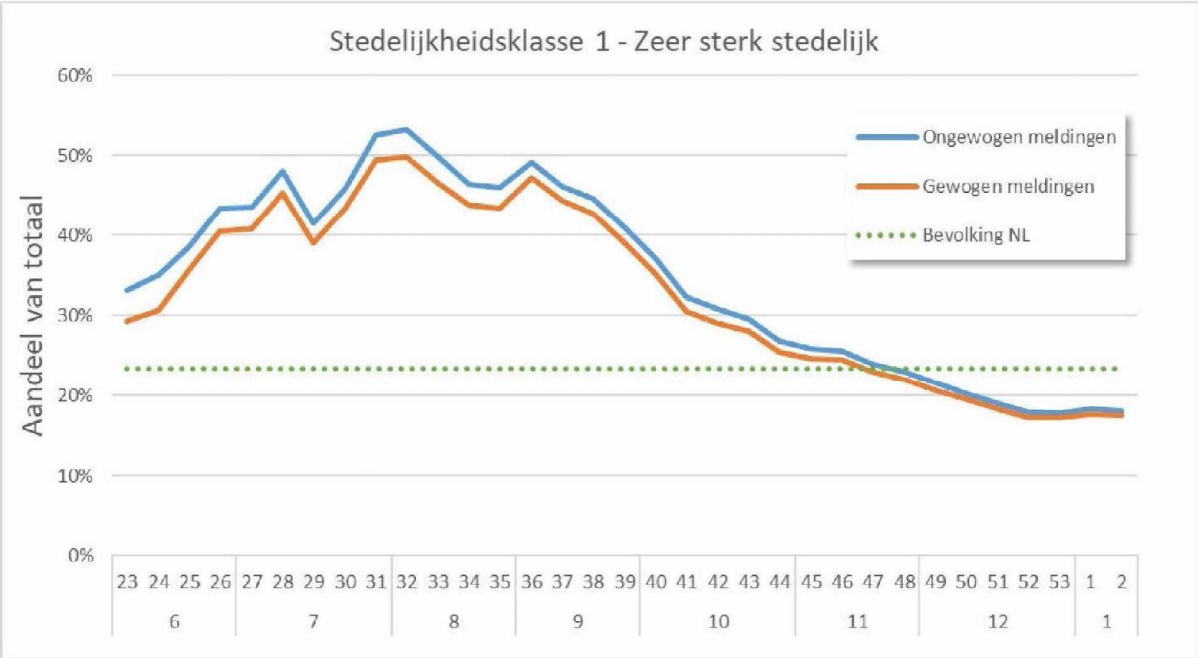
Verreweg de meeste meldingen bevinden zich binnen 20 km van een testlocatie, en krijgen daarmee een gewicht <2.

e) Het gemiddelde gewicht van alle cases over de hele periode is 1,088

De gewichten zijn zoals verwacht het hoogst voor de minst stedelijke gebieden en aan het begin van de periode, aangezien die gemiddeld verder weg van de testlocaties bevinden.

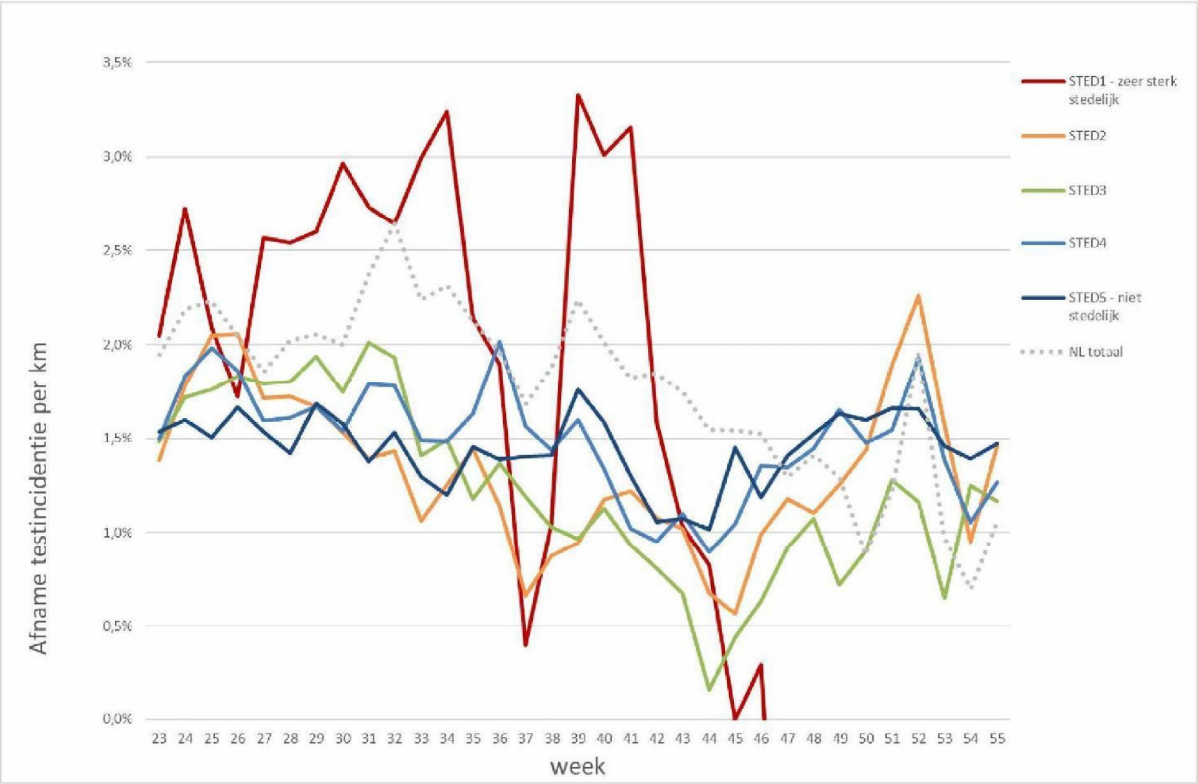


De met deze gewichten gemaakte grafieken verschillen slechts licht met die obv ongewogen meldingen. Bijv. het aandeel van stedelijkheidsklasse-1 daalt met 4–0,5 ppt (zie figuur). Het grote patroon blijft gelijk.



f) Tot slot nog de vraag of dat verschijnsel van afnemende testincidentie onafhankelijk is van de meldingenincidentie. Oftewel: heb je op het platteland – ver weg van de teststraat - minder meldingen door lagere testbereidheid, of zijn er sowieso minder besmettingen en tests?

Dat is lastig te bepalen. Wanneer je die afnamepercentages uitsplitst naar de 5 stedelijkheidsklassen zie ik hier iig weinig verschil tussen. In elke klasse zie je het fenomeen van dalende testincidentie-met-afstand-tot-testlocatie. Zie figuur.

Alleen de meest stedelijke klasse (STED=1) heeft soms afwijkende waarden, en daalt op een gegeven moment onder nul (omgekeerdere relatie?: hoe verder weg van de testlocatie, hoe hoger de testbereidheid?). Maar dit heeft waarschijnlijk te maken dat deze stedelijksheidsklasse verreweg de minste plaatsen heeft (66 vd 2576), dus maar weinig punten in een scatterplot. Bovendien liggen deze punten vaak op korte afstand van een testlocatie, waardoor de spreiding op de x-as maar laag is, en een trendlijn niet echt uit de verf komt.

Het onderscheid naar stedelijkheid is verder niet gebruikt bij de bepaling van de gewichten.