

Algemene verspreiding

MILIEUMETINGEN UITGEVOERD IN DE WIJDE OMGEVING VAN UMICORE HOBOKEN - RELATIES TUSSEN MILIEUCOMPARTIMENTEN -



VITO

December 2008

Alle rechten, waaronder het auteursrecht, op de informatie vermeld in dit document berusten bij de Vlaamse Instelling voor Technologisch Onderzoek NV ("VITO"), Boeretang 200, BE-2400 Mol, RPR Turnhout BTW BE 0244.195.916.

De informatie zoals verstrekt in dit document is vertrouwelijke informatie van VITO. Zonder de voorafgaande schriftelijke toestemming van VITO mag dit document niet worden gereproduceerd of verspreid worden noch geheel of gedeeltelijk gebruikt worden voor het instellen van claims, voor het voeren van gerechtelijke procedures, voor reclame of antireclame en ten behoeve van werving in meer algemene zin aangewend worden.

SAMENVATTING

Voor het blootstellingsonderzoek van lood bij kinderen in Hoboken in 2008 zijn milieumetingen uitgevoerd bij woningen, scholen en publieke plaatsen verspreid over de gemeenten Hoboken, Wilrijk en Hemiksen.

Hierbij zijn de loodconcentraties gemeten in de bodem, in het straatstof, in PM10 in de binnenlucht, in PM10 in de buitenlucht, in het stof bemonsterd met een veegdoekje en in het stof van de stofzuigerzakken. Bovendien zijn oudere loodconcentraties beschikbaar van de bodem in dat gebied (concentraties van de Universiteit van Gent). Bij de statistische analyse is de relatie onderzocht tussen de loodconcentratie in huisstof (stofzuiger of veegstof) en de loodbelading van straatstof en veegstof enerzijds en het loodgehalte in bodem, straatstof en zwevend stof anderzijds. De methodes hierbij toegepast zijn descriptieve analyse en multiple lineaire regressies.

Er zijn 3 reeksen bodemconcentraties gebruikt voor de statische verwerking: een eerste reeks met de resultaten van VITO, een tweede met de resultaten van de Universiteit van Gent en een derde reeks bestaande uit een combinatie van de resultaten van VITO en van de Universiteit van Gent. De resultaten hiervan zijn samengevat in tabel 19. Het resultaat van de multiple lineaire regressie is afhankelijk van de gekozen reeks met loodconcentraties in de bodem.

In het laatste deel is een brontracering toegepast waarbij de elementen uit de XRF analyse van stof uit de stofzuigerzak (=binnenshuis) gerelateerd zijn aan bodem, straatstof en zwevend stof buiten. De methodes die hierbij toegepast zijn zijn descriptieve analyse, Principal Components Analysis (PCA) en multi-pele lineaire regressie.

Met behulp van de Principal Component Analyse werden de elementen van de stofzuigerzak opgedeeld in vier groepen, waarvan één groep duidelijk afgebakend was als de groep van metalen. Voor de brontracering kon voor deze groep een verband aangetoond worden met het bodemstof. In de regressieanalyse kwam immers duidelijk naar voren dat de metalen binnenshuis (stofzuigerzak) voornamelijk afkomstig zijn uit het bodemstof. Deze verklaarde immers 43% van de variantie van Cd, Pb, As, Zn, Cu, Se en Ni in de stofzuigerzak. Er moet echter wel voorzichtig omgesprongen worden met deze resultaten. Er was immers informatie beschikbaar voor maximaal 29 woningen, wat al zeer weinig is. Bovendien waren er voor deze woningen ontbrekende waarden en werden geëxtrapolerde waarden gebruikt voor zwevend stof buiten.

INHOUD

Samenvatting.....	I
Inhoud	II
Lijst van tabellen	IV
Lijst van figuren	V
Hoofdstuk 1 Probleemstelling.....	1
1.1 <i>Inleiding.....</i>	1
1.2 <i>Doelstelling</i>	1
1.3 <i>Opbouw van de studie</i>	2
1.4 <i>Coördinatie en uitvoering van de studie.....</i>	2
Hoofdstuk 2 Methodologie.....	4
2.1 <i>Metingen.....</i>	4
2.2 <i>Statistiek</i>	5
2.2.1 <i>Descriptieve analyse</i>	5
2.2.2 <i>Identificatie van de covariaten.....</i>	5
2.2.3 <i>Kwantificatie van de covariaten.....</i>	6
Hoofdstuk 3 Resultaten.....	7
3.1 <i>Dataset met bodemconcentraties VITO.....</i>	7
3.1.1 <i>Descriptieve analyse</i>	7
3.1.2 <i>Identificatie van de covariaten.....</i>	9
→ <i>Concentratie Pb in stofzuigerzak (mg/kg ds).....</i>	9
→ <i>Belading Pb in veegmonsters (µg/m²).....</i>	11
→ <i>Concentratie Pb in veegmonsters (mg/kg ds).....</i>	13
3.1.3 <i>Kwantificatie van de covariaten.....</i>	15
3.2 <i>Dataset met bodemconcentraties van de Universiteit van Gent.....</i>	17
3.2.1 <i>Descriptieve analyse</i>	17
3.2.2 <i>Identificatie van de covariaten.....</i>	17
→ <i>Concentratie Pb in stofzuigerzak (mg/kg ds).....</i>	17
→ <i>Belading Pb in veegmonsters (µg/m²).....</i>	19
→ <i>Concentratie Pb in veegmonsters (mg/kg ds).....</i>	21
3.2.3 <i>Kwantificatie van de covariaten.....</i>	23
3.3 <i>Dataset met geïntegreerde bodemconcentraties</i>	25
3.3.1 <i>Descriptieve analyse</i>	25
3.3.2 <i>Identificatie van de covariaten.....</i>	25
→ <i>Concentratie Pb in stofzuigerzak (mg/kg ds).....</i>	25
→ <i>Belading Pb in veegmonsters (µg/m²).....</i>	27
→ <i>Concentratie Pb in veegmonsters (mg/kg ds).....</i>	29
3.3.3 <i>Kwantificatie van de covariaten.....</i>	31

Hoofdstuk 4	Brontracering	34
4.1.1	Inleiding.....	34
4.1.2	Methodologie	34
4.1.3	Resultaten	35
4.1.4	Discussie	39
Hoofdstuk 5	Milieumetingen Pb versus brontracering	40
Hoofdstuk 6	Besluit	41
Hoofdstuk 7	Literatuurlijst	44

LIJST VAN TABELLEN

Tabel 1: Descriptieve statistieken voor de concentratie van Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en concentratie (mg/kg ds) van Pb in de veegmonsters, met en zonder outliers	9
Tabel 2: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers	9
Tabel 3: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers	11
Tabel 4: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers	14
Tabel 5: De (aangepaste) R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters	16
Tabel 6: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat	17
Tabel 7: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers	18
Tabel 8: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers	20
Tabel 9: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers	22
Tabel 10: De (aangepaste) R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters	24
Tabel 11: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat	25
Tabel 12: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers	26
Tabel 13: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers	28
Tabel 14: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers	30
Tabel 15: De aangepaste R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters	32
Tabel 16: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat	33
Tabel 17: Descriptieve statistieken voor de concentratie van elementen in de stofzuiger, het bodemstof, straatstof en zwevend stof buiten (mg/kg ds)	35
Tabel 18: Correlatiematrix met Spearman correlatiecoëfficiënten (%) voor de elementen van de stofzuigerzak. Groen gearceerd: gecorreleerde elementen ($r \geq 50\%$); rood gearceerd: sterk gecorreleerde elementen ($r \geq 80\%$)	36
Tabel 19: Vergelijking resultaten statistische analyses milieumetingen met bodemconcentraties VITO, Universiteit van Gent en geïntegreerde bodemconcentraties.	42

LIJST VAN FIGUREN

Figuur 1: Histogram (links) en Box-and-Whisker plot (rechts) voor de concentratie van Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en concentratie (mg/kg ds) van Pb in de veegmonsters	8
Figuur 2: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)	10
Figuur 3: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers	11
Figuur 4: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)	12
Figuur 5: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers	13
Figuur 6: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)	14
Figuur 7: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers	15
Figuur 8: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)	18
Figuur 9: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers	19
Figuur 10: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)	20
Figuur 11: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers	21
Figuur 12: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)	22
Figuur 13: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers	23
Figuur 14: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)	26
Figuur 15: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers	27
Figuur 16: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)	28
Figuur 17: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers	29
Figuur 18: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)	30
Figuur 19: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers	31

Lijst van figuren

- Figuur 20: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de component twee van de stofzuigerzak bestaande uit de elementen Cd, Pb, As, Zn, Cu, Se en Ni _____ 38
- Figuur 21: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de component drie van de stofzuigerzak bestaande uit de elementen Cr, Ba en TiO₂ _____ 39

HOOFDSTUK 1 PROBLEEMSTELLING

1.1 Inleiding

Sinds de jaren '70 wordt in de wijk Moretusburg te Hoboken, naast de Umicore-vestiging, het loodgehalte in bloed van kinderen tussen 12 maanden en 12 jaar regelmatig opgevolgd. Uit de eerste resultaten bleek het noodzakelijk dat de loodopname bij kinderen beperkt moest worden. Hiervoor werden meerdere acties ondernomen door de overheden en het bedrijf. Zo werd de uitstoot aan lood (Pb) gereduceerd, werden maatregelen getroffen om stof te vermijden, werden de bewoners bewust gemaakt van de problematiek, werden er adviezen rond gezondheid gegeven etc.

De voormalige risicozone Moretusburg is inmiddels gesaneerd.

Volgens het Vlaamse berekeningsmodel voor het inschatten van risico's van bodemverontreiniging is er in de omgeving van Hoboken geen overschrijding meer te verwachten van het huidige gezondheids criterium voor inname van lood. De voormalige risicozone Moretusburg en Hertogvelden zijn reeds gesaneerd, de overige risicozones in Vinkevelden worden in de komende maanden gesaneerd. Lood is een stof waarvoor de gezondheidsrisico's niet alleen kunnen worden ingeschat door modelberekening van de inname (externe belasting), maar ook door het meten van het loodgehalte in bloed (interne belasting).

De resultaten van het biomonitoringsonderzoek dat tweemaal per jaar georganiseerd wordt in de wijk Moretusburg bevestigen de berekende resultaten via het blootstellingsmodel. Buiten de wijk Moretusburg werden de berekende gehalten echter nog niet getoetst aan gemeten waarden (met uitzondering van de school die tweemaal per jaar wordt meegenomen als controle).

1.2 Doelstelling

Om er zeker van te zijn dat er buiten Moretusburg, in de ruimere omgeving van Hoboken, geen overschrijding is van het huidige gezondheids criterium voor lood in bloed van 10 µg/dl, werd het loodbloedgehalte van kinderen onderzocht. De resultaten van dit onderzoek worden vergeleken met de resultaten uit het controlegebied Hemiksem.

Op basis van de huidige milieumetingen in de lucht en in de bodem blijkt dat de afstand tot de fabriek belangrijk is. Bijgevolg wordt ook aandacht geschonken aan deze variabelen.

Eveneens wordt een bestaand blootstellingsmodel verfijnd, specifiek voor de regio Hoboken. Hiervoor worden bij bepaalde kinderen, die deelnamen aan het onderzoek, ook milieumetingen thuis en op school uitgevoerd. De gegevens van lood in bloed en lood in het milieu worden samen verwerkt. Op deze manier wordt meer inzicht gekregen in de belangrijkste blootstellingswegen voor de kinderen.

1.3 Opbouw van de studie

Om deze studie uit te voeren werden volgende deelopdrachten uitgeschreven en uitgevoerd:

1. Rekrutering en staalname bij kinderen in Hoboken door PIH
2. Statistische analyse door Universiteit Hasselt
3. Milieumetingen door VITO
4. Modelleren van de blootstelling van kinderen aan lood in Hoboken door VITO

Dit deelrapport handelt over de milieumetingen en meer bepaald het onderdeel waarin nagegaan werd of er relaties kunnen gelegd worden tussen de metingen uitgevoerd binnen de woningen en de metingen buitenshuis. Het rapport over de milieumeting zelf heeft als referentie 2008/MIM/R/080. Statistisch significante relaties worden gebruikt in de geïntegreerde risico-analyse voor het schatten van binnenshuisconcentraties voor woningen waar geen milieumetingen voorhanden zijn.

1.4 Coördinatie en uitvoering van de studie

Deze studie is gefinancierd door de overeenkomst gesloten tussen de Vlaamse Regering, de OVAM en Umicore NV voor het aanpakken van de verontreiniging in de wijdere omgeving rond de Umicore NV vestigingen.

Het project wordt inhoudelijk volledig gestuurd door de administraties binnen de Technische Werkgroep, bestaande uit:

- team Milieugezondheidszorg Antwerpen van het Vlaams Agentschap Zorg en Gezondheid
- Medisch Milieukundigen bij Logo Stad Antwerpen en bij Logo Mechelen
- Dienst Milieu & Gezondheid van het Departement Leefmilieu, Natuur en Energie
- VMM
- OVAM

De andere leden van de Technische Werkgroep zijn de opdrachtnemers:

- PIH
- VITO
- Universiteit Hasselt

Tevens werd Haskoning toegelaten tot de vergadering als secretaris en voor het aanleveren van de gekende data in de wijdere omgeving van Umicore NV te Hoboken.

In deze Technische Werkgroep, onder voorzitterschap van de OVAM, wordt door de verschillende betrokken organisaties advies gegeven en wordt samengewerkt over de onderzoeksopzet en optimale afstemming van verschillende onderzoeken op hetzelfde domein.

Umicore NV maakte geen deel uit van deze Technische werkgroep. Zij werden wel geïnformeerd over de belangrijkste beslissingen die genomen werden binnen de Technische Werkgroep.

De conclusies van deze studie werden voor publicatie voorgelegd aan een wetenschappelijk team, bestaande uit volgende leden:

- Prof. dr. Em. A. Meheus

- Prof. Lison
- Prof. dr. M. Van Sprundel
- Prof. Dr. E. Smolders
- Agnes Oomen
- Dr. Ward Roekens

HOOFDSTUK 2 METHODOLOGIE

2.1 Metingen

De statistische analyse heeft tot doel de relatie te onderzoeken tussen de loodconcentratie in huisstof (stofzuiger of veegstof) en de loodbelading in huisstof enerzijds en het loodgehalte in bodem, straatstof en zwevend stof anderzijds. Voor de statistische analyse werd gebruik gemaakt van volgende meetgegevens:

- Concentratie in bodem, fractie < 250 μm (mg Pb/kg ds);
- Concentratie in straatstof (mg Pb/kg ds);
- Belading op straatstof ($\mu\text{g Pb/m}^2$);
- Concentratie in PM_{10} in binnenlucht, bepaald met ICP-MS (ng Pb/m^3);
- Concentratie in huisstof:
 - o Resultaat van de veegmonsters (mg Pb/kg ds);
 - o Resultaat van de stofzuigerzak (mg Pb/kg ds);
- Belading op huisstof, resultaat van de veegmonsters ($\mu\text{g Pb/m}^2$)

Aanvullend werd de concentratie op PM_{10} buitenshuis meegenomen. Omdat niet per woning gegevens beschikbaar waren, werd een interpolatie uitgevoerd op basis van de resultaten van de meetposten (Van Deun et al., 2008) en werd aan de hand van de Lambertcoördinaten van de woningen de overeenkomstige loodconcentratie op zwevend stof afgeleid.

Een eerste reeks analyses werd uitgevoerd op de data, zoals bepaald in de milieumetingen (Van Deun et al., 2008). Met het oog op verder gebruik van de regressies en de noodzaak om alternatieve bodemconcentraties te gebruiken, werd eveneens een statistische analyse uitgevoerd, waarbij de bodemconcentraties uit het project milieumetingen vervangen werden door de statistisch geïnterpoleerde bodemconcentraties, afgeleid door de Universiteit van Gent. Doordat deze laatste dataset oudere gegevens bevat, zijn de concentraties afkomstig van het gebied vooraleer sanering werd uitgevoerd. Een deel van de woningen werd in tussentijd gesaneerd. Omdat niet geweten is hoe snel de binnenhuisconcentraties reageren op wijzigingen in de bodemconcentraties, werd daarom een derde analyse uitgevoerd op een geïntegreerde dataset. Hierbij werden concentraties uit de dataset van de Universiteit van Gent gebruikt voor de woningen, waarbij al sanering werd uitgevoerd. Het betreft hier 2 woningen in Moretusburg en 2 woningen in het oostelijk gebied net buiten Moretusburg. Voor de overige woningen werden de milieumetingen uitgevoerd in dit project, gebruikt.

Bij de metingen op zwevend stof lagen 6 van de 20 meetresultaten beneden de detectielimiet. Voor deze punten werd de helft van de detectielimiet gehanteerd.

2.2 Statistiek

2.2.1 Descriptieve analyse

Een descriptieve analyse werd uitgevoerd om een algemeen beeld te krijgen van de data. Een tabel met descriptieve statistieken van de afhankelijke variabelen (concentratie Pb in stofzuiger en belading en concentratie Pb in veegmonsters) werd gemaakt om kenmerken van de verdeling te identificeren zoals het gemiddelde, minimum, maximum, ... (Bland, 2000). Bovendien werden histogrammen en Box-and-whisker plots gemaakt om de verdeling visueel voor te stellen. Van deze Box-and-whisker plots konden ook eendimensionale outliers geïdentificeerd worden (Bland, 2000). Een outlier is een observatie die op een abnormale afstand van andere waardes ligt. De box wordt getekend van het 25^{ste} percentiel tot het 75^{ste} percentiel van de data. Voor het p^{de} percentiel van een distributie is $p\%$ van de observaties gelijk of kleiner dan deze waarde. De box weerspiegelt als dusdanig de middelste 50% van de data. De punt in de box geeft de mediaan weer. De interquartiele range (IR) is het verschil tussen het 75^{ste} en het 25^{ste} percentiel, het is een maat van spreiding. De volgende punten worden berekend: $L1 = 25^{\text{ste}} \text{ percentiel} - 1.5 * IR$, $L2 = 25^{\text{ste}} \text{ percentiel} - 3 * IR$, $U1 = 75^{\text{ste}} \text{ percentiel} + 1.5 * IR$ en $U2 = 75^{\text{ste}} \text{ percentiel} + 3 * IR$. Een lijn wordt getekend van het 25^{ste} percentiel naar de kleinste observatie die groter is dan L1, en van het 75^{ste} percentiel naar de grootste observatie die kleiner is dan U1. Deze lijnen noemen we whiskers, ze geven het gebied van niet-outliers weer. Buiten deze whiskers worden de outliers individueel geplot met een puntje. Outliers die nog buiten L2 en U2 liggen, noemen we extreme outliers, deze worden geplot met een sterretje. De descriptieve plots helpen in het identificeren van eendimensionale outliers, de identificatie van outliers die echter een significant effect hebben op de gefitte modellen zal bestudeerd worden in de volgende paragraaf.

2.2.2 Identificatie van de covariaten

Univariate regressiemodellen werden uitgevoerd om de relatie tussen de respons (concentratie Pb in de stofzuiger en concentratie en belading Pb in veegmonsters) en de covariaten (zwevende lucht buiten en binnen, bodemstof, straatstof, ...) te onderzoeken. Een multipel lineair regressiemodel werd opgebouwd waarbij de p-waarde voor de onafhankelijke variabelen om in het model te worden opgenomen 0.10 was en 0.05 om er in te blijven. Op deze manier wordt er een verklarend model voor ieder van de drie responsen opgebouwd, waarbij de verschillende covariaten gebruikt worden als verklarende factoren (Neter et al., 1996).

Bij de opbouw van het model werd rekening gehouden met het probleem van multicollineariteit. Dit is het aanwezig zijn van een hoge lineaire correlatie tussen twee of meer verklarende variabelen in een regressiemodel (Neter et al., 1996)). Indien er sprake is van multicollineariteit is het moeilijk om het effect van één variabele op de respons te schatten en bovendien zijn de estimates van de regressiecoëfficiënten minder precies. Dit effect wordt nagegaan met behulp van de variantie inflatiefactoren. Als de variantie inflatie factor groter is dan 10, wordt multicollineariteit vastgesteld (Fox, 1991).

Wanneer het multipel lineair regressiemodel wordt gefit, moeten de assumpties van dit parametrische model worden gecheckt. Als de assumpties niet voldaan zijn, moeten er maatregelen genomen worden. Een aantal diagnostische plots kunnen gebruikt worden om deze voorwaarden te checken². Een plot van de residuen ten opzichte van de

voorspelde waarden werd gebruikt om de assumptie van constante varianties, onafhankelijkheid (willekeurigheid) en lineariteit na te gaan. De ideale situatie is dat de error termen random verdeeld zijn rond de waarde nul en dat hun distributie rond nul dezelfde variabiliteit vertoont. Enige afwijking van deze situatie toont aan dat minstens één van deze voorwaarden niet voldaan is. Een plot van de residuen ten opzichte van de verwachte residuen onder normaliteit werd gebruikt om normaliteit te checken, dit noemt men de normal probability plot. Indien de assumptie van normaliteit voldaan is, verwachten we een lineaire relatie tussen de residuen en de verwachte residuen. De diagnostische plots zijn echter informeel en moeten gebruikt worden in combinatie met formele testen om na te gaan wat geobserveerd is niet alleen te wijten is aan kans. De test van constante varianties, the White's General test (White, 1980) is geldig wanneer de error termen normaal en onafhankelijk verdeeld zijn met gemiddelde nul. Het test de nulhypothese van constante varianties (homoscedasticiteit). De Kolmogorov-Smirnov test² werd gebruikt om normaliteit te bepalen. Het test de nulhypothese dat de error termen normaal verdeeld zijn. De lack of fit test (Neter et al., 1996) checkt de geschiktheid van het model, dat is het checken van de lineariteit. De nulhypothese van lineariteit wordt geaccepteerd als de p-waarde groter is dan 0.05.

Zoals eerder vermeld moeten ook de outliers gebaseerd op het opgebouwde lineair regressiemodel geïdentificeerd worden. De outliers met betrekking tot de respons worden bepaald met behulp van de studentized deleted residuen en de outliers met betrekking tot de verklarende factoren met behulp van de leverages of de diagonale elementen van de hat matrix (Neter et al., 1996). Als de absolute waarde van de studentized deleted residue groter is dan 2, dan is deze observatie een outlier met betrekking tot de respons. Als de leverage groter is dan $2 \cdot p/n$ (met p het aantal parameters en n het aantal observaties), dan is deze observatie een outlier met betrekking tot de verklarende factoren. Nadat de outliers geïdentificeerd zijn, wordt verder bepaald of deze ook van invloed zijn op het model. Dit gebeurt met behulp van hun DFFITS (invloed op een enkele gefitte waarde) en Cook's Distance (invloed op alle gefitte waarden). Outliers die ook van invloed bevonden werden, worden verder onderzocht. Er zal opnieuw een multipel lineair regressiemodel gefit worden zonder deze beïnvloedende outliers en gekeken worden of er al dan niet een significant verschil optreedt in de regressieparameters. Als er geen significant verschil voorkomt, worden de outliers behouden in het model, indien niet wordt het model zonder deze beïnvloedende outliers verder gebruikt.

2.2.3 Kwantificatie van de covariaten

De aangepaste R-kwadraat (coëfficiënt van determinatie) (Neter et al., 1996) van het bekomen multipel lineair regressiemodel toont de proportie van variabiliteit in de respons die in rekening gebracht wordt door dit model, aangepast omdat het aantal verklarende factoren in het model in rekening wordt gebracht. De relatieve belangrijkheid van een predictor in een multipel lineair regressiemodel wordt bepaald door de General Dominance Index⁵. Dit is de gemiddelde toename van de coëfficiënt van determinatie geassocieerd met de predictor over alle mogelijke submodellen.

Kwantitatieve relaties tussen de covariaten en de respons worden berekend op basis van de estimates van de regressiecoëfficiënten van het multipel lineair regressiemodel. De regressiecoëfficiënt is de helling van de regressielijn. Dit wil zeggen dat de estimate van de regressiecoëfficiënt van een bepaalde covariate, indien we de overige covariates in het model constant houden, de verandering in het gemiddelde van de respons weergeeft per eenheid dat de desbetreffende covariate stijgt (Neter et al., 1996).

HOOFDSTUK 3 RESULTATEN

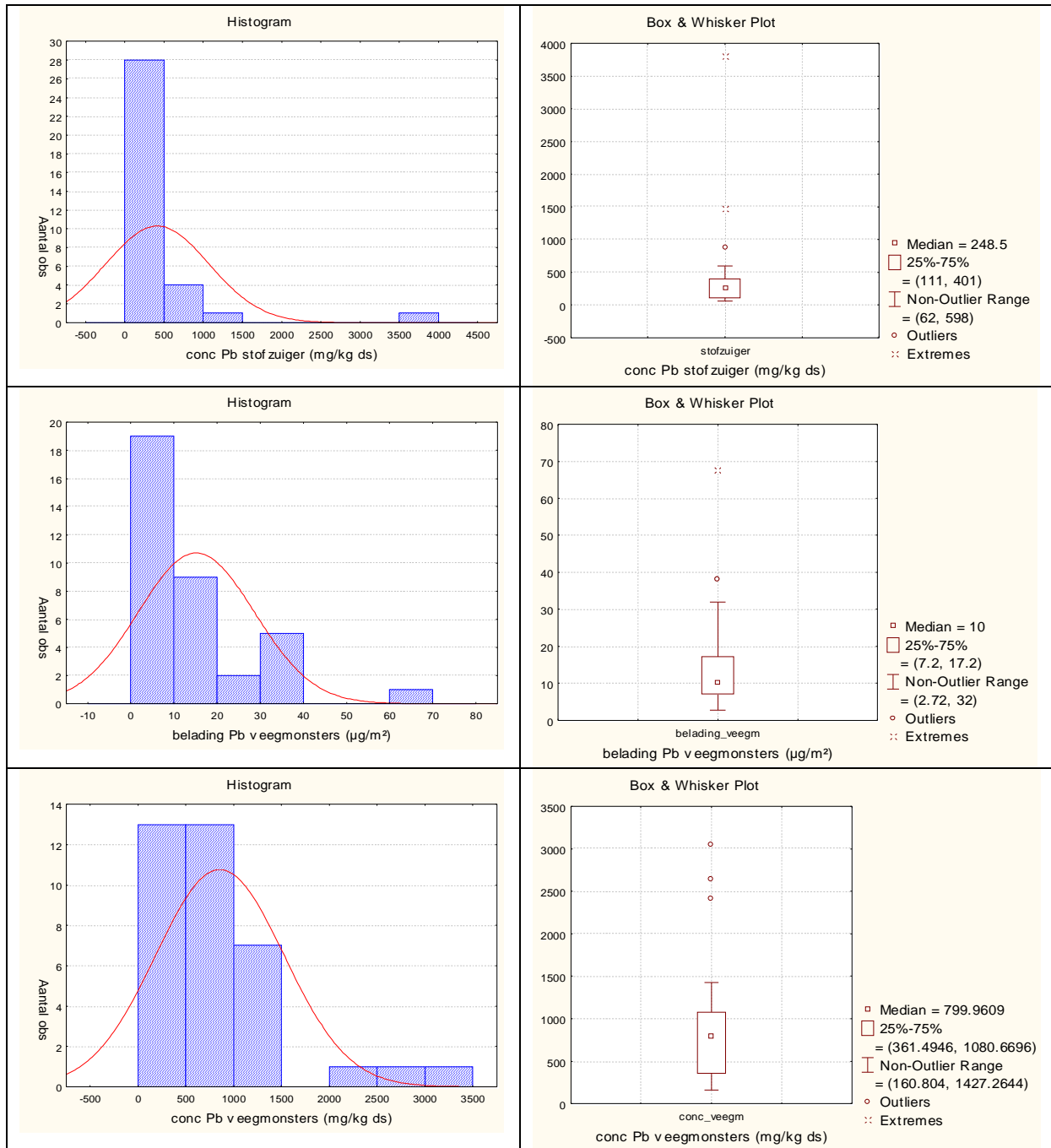
3.1 Dataset met bodemconcentraties VITO

In deze analyse werden als bodemconcentraties de door VITO gemeten concentraties van de bovenste cm van de bodem gebruikt. Het betreft metingen in de tuinen van woningen.

3.1.1 Descriptieve analyse

De histogrammen (Figuur 1 linkerkant) tonen aan dat de verdelingen van zowel de concentratie van Pb in de stofzuiger als de belading en concentratie van Pb in de veegmonsters rechtsscheef verdeeld zijn. Dit kan mede veroorzaakt worden door de outliers. Uit de Box-and-Whisker plots (Figuur 1 rechterkant) werden 3 outliers (woning 3, 50 en 30), waarvan 2 extreme (woning 50 en 30), geïdentificeerd voor de concentratie van lood in de stofzuigerzak. Eveneens 3 outliers (woning 5, 37 en 28), waarvan 1 extreme (woning 28), werden geïdentificeerd voor de belading van lood in de veegmonsters. Voor de concentratie van lood in de veegmonsters werden ook 3 outliers (woning 5, 28 en 30) geïdentificeerd. De descriptieve statistieken met en zonder deze outliers worden gegeven in Tabel 1. Hieruit blijkt dat bij het weglaten van de outliers de gemiddelde en mediaan dichter bij elkaar liggen, en de verdeling dus minder scheef wordt.

Hoofdstuk 3 Resultaten



Figuur 1: Histogram (links) en Box-and-Whisker plot (rechts) voor de concentratie van Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en concentratie (mg/kg ds) van Pb in de veegmonsters

	Concentratie Pb stofzuiger (mg/kg ds)		Belading Pb veegmonsters ($\mu\text{g}/\text{m}^2$)		Concentratie Pb veegmonsters (mg/kg ds)	
	Gehele populatie	Zonder woning 3, 50 en 30	Gehele populatie	Zonder woning 5, 37 en 28	Gehele populatie	Zonder woning 5, 28 en 30
N	28	25	36	33	36	33
Gemiddelde	474,9	286,6	15,0	12,1	848,6	680,5
SD	710,7	152,3	13,4	8,3	666,5	359,5
Mediaan	311,5	293,0	10,0	9,2	800,0	661,0
Min	62,0	62,0	2,7	2,7	160,8	160,8
Max	3800,0	598,0	67,6	32,0	3052,6	1427,3
P10	111,0	111,0	4,0	4,0	252,8	252,8
P90	862,0	558,0	32,0	26,0	1427,3	1154,8

Tabel 1: Descriptieve statistieken voor de concentratie van Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en concentratie (mg/kg ds) van Pb in de veegmonsters, met en zonder outliers

3.1.2 Identificatie van de covariaten

→ Concentratie Pb in stofzuigerzak (mg/kg ds)

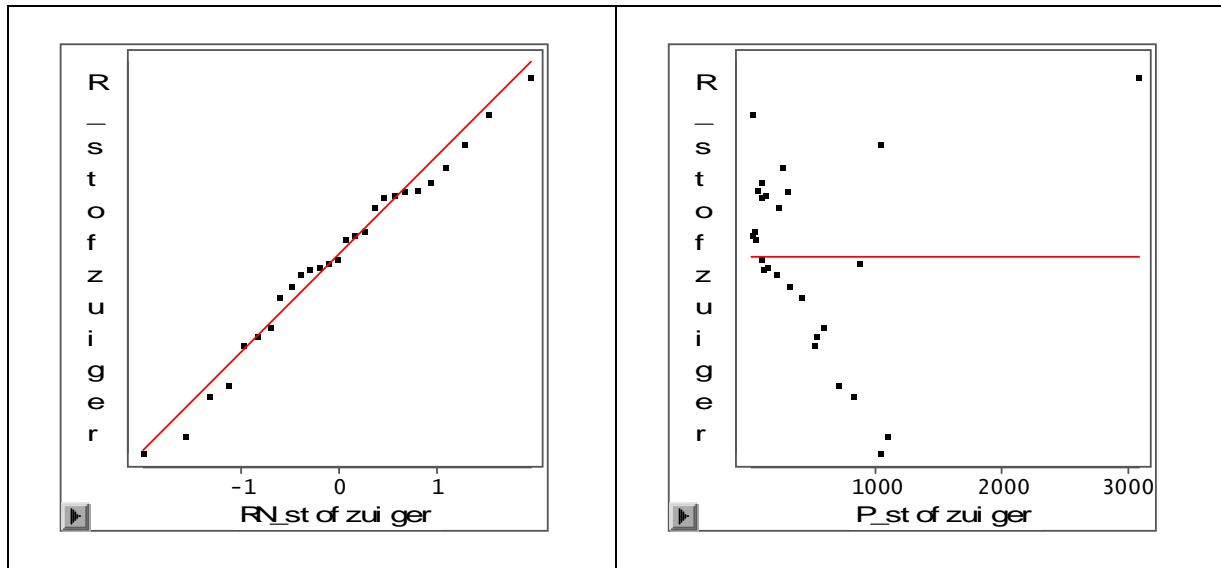
Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de stofzuigerzak bleven de volgende significante variabelen over in het finale model (N=27): concentratie Pb in bodem in fractie < 250 μm (mg/kg ds) en de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 2). Dit model heeft een aangepaste R-kwadraat van 0.7382, dit wil zeggen dat dit model 74% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart. De variantie inflatiefactoren (Tabel 2) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5, 30 en 44		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		-81,68 (-297,89; 134,53)	0,4432	0	103,19 (-69,38; 275,76)	0,2274	0
Concentratie Pb bodem	mg/kg ds	1,23 (0,62; 1,85)	0,0004	1,27	1,00 (0,46; 1,55)	0,0010	1,00
Belading Pb straatstof	$\mu\text{g}/\text{m}^2$	0,04 (0,02; 0,05)	< 0,0001	1,27	ns	ns	ns

ns: not significant

Tabel 2: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers

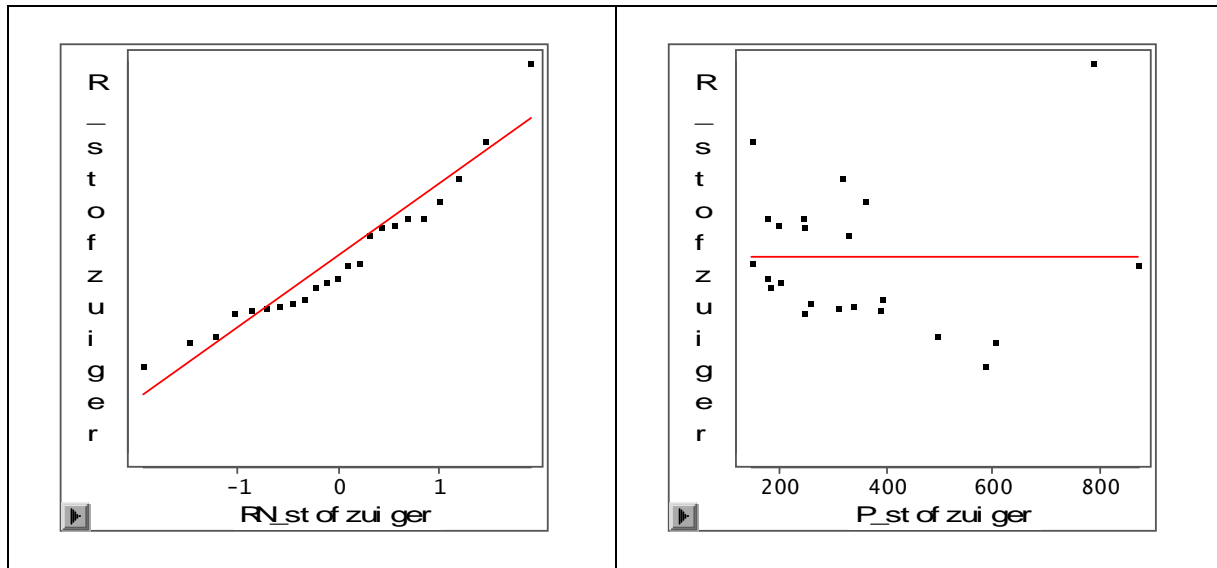
De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 2 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 2 toont dat de variabiliteit van de error termen verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0038) werden constante varianties verworpen.



Figuur 2: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waarden (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden drie outliers gevonden met betrekking tot de concentratie van Pb in de stofzuiger, en met behulp van de leverages werden vijf outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk drie en vier van hen waren van invloed. Wanneer we het model zonder deze 4 beïnvloedende outliers (namelijk woning 2, 5, 30 en 44) fitten, wordt de concentratie van Pb in straatstof insignificant (p -waarde=0.6854). Het finaal gefitte model ($N=23$), zonder de beïnvloedende woningen, heeft als verklarende factor enkel de concentratie van Pb in de bodem in fractie < 250 μ m (mg/kg ds) (Tabel 2). Dit model heeft een R-kwadraat van 0.4079, dit wil zeggen dat dit model 41% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p -waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 3 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 3 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p -waarde=0.0506) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 3, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test uitgevoerd die de geschiktheid bevestigde aangezien de p -waarde (lack of fit)=0.7646.



Figuur 3: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers

→ **Belading Pb in veegmonsters ($\mu\text{g}/\text{m}^2$)**

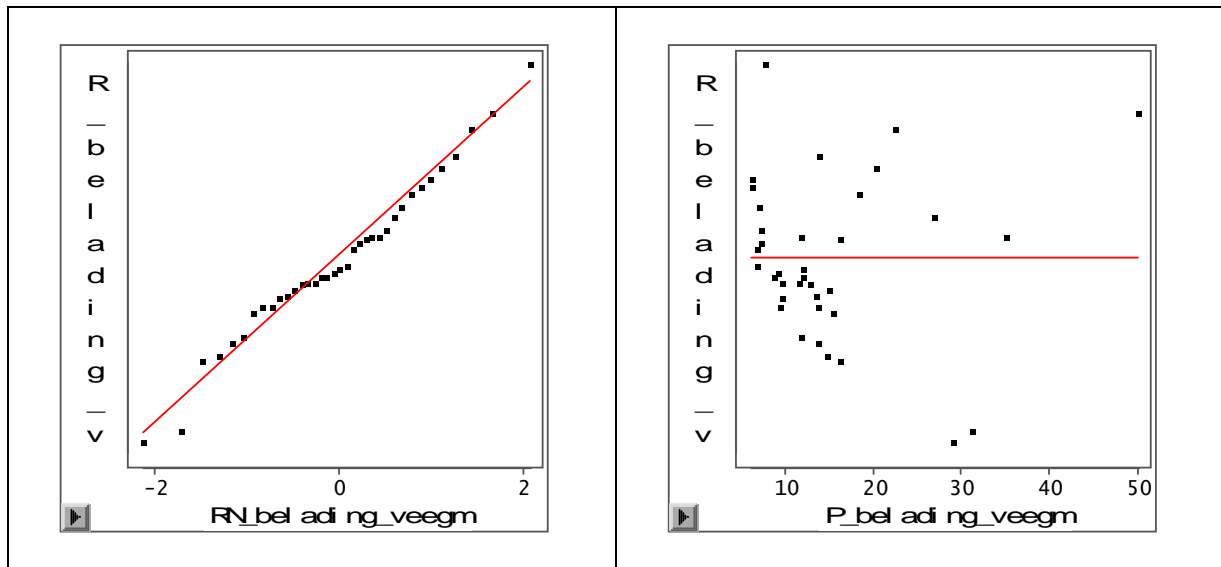
Bij het opbouwen van een multipel lineair regressiemodel voor de belading van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model (N=36): concentratie Pb in zwevend stof buitenshuis (ng/m^3) en of er al dan niet gerookt wordt in de woning (0-1) (Tabel 3). Dit model heeft een aangepaste R-kwadraat van 0.4670, dit wil zeggen dat dit model 47% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) verklaart. De variantie inflatie factoren (Tabel 3) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid of score	Met beïnvloedende woningen			Zonder beïnvloedende woningen 18, 26, 28, 32 en 62		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		-1.46 (-8.65; 5.74)	0.6830	0	3.60 (-3.94; 11.15)	0.3357	0
Conc Pb zwevend stof buiten	ng/m^3	0.18 (0.10; 0.26)	< 0.0001	1.03	0.18 (0.09; 0.26)	0.0002	1.01
Roken	0 = niet roken in woning 1 = roken in woning	11.38 (1.61; 21.15)	0.0237	1.03	18.74 (3.60; 33.87)	0.0171	1.04
Huis gesaneerd	0 = huis niet gesaneerd 1 = huis gesaneerd	ns	ns	ns	-6.71 (-12.10; -1.32)	0.0165	1.02

ns: niet significant

Tabel 3: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 4 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 4 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0984) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 4, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.1220.

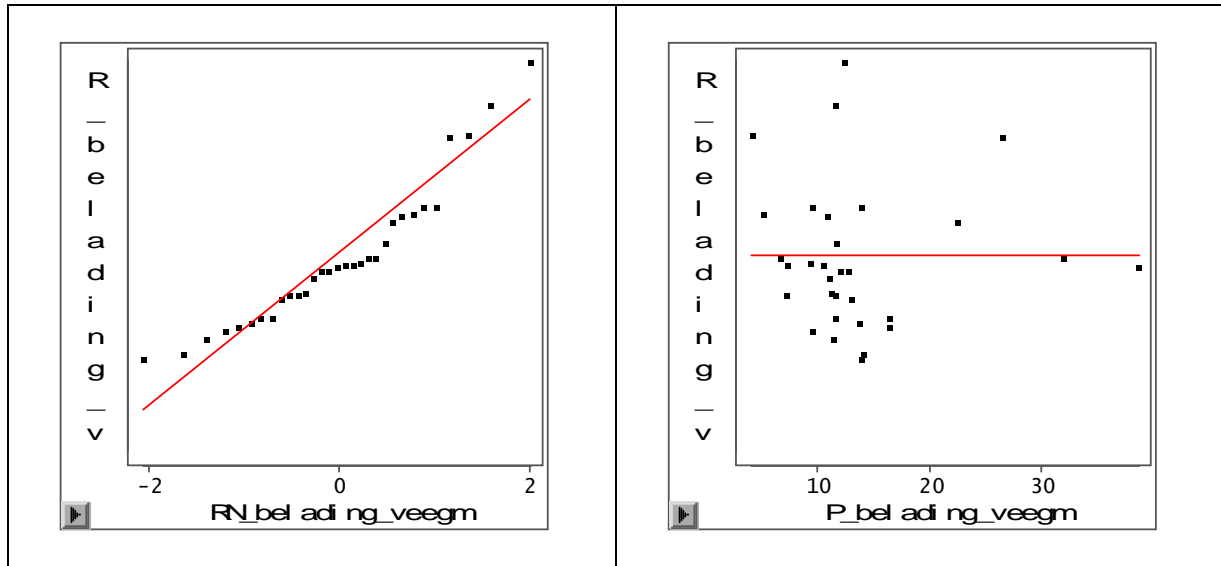


Figuur 4: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden vier outliers gevonden met betrekking tot de belading van Pb in de veegmonsters, en met behulp van de leverages werden zeven outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk drie en vijf van hen waren van invloed. Wanneer we het model zonder deze vijf beïnvloedende outliers (namelijk woning 18, 26, 28, 32 en 62) fitten, wordt de variabele of het huis al dan niet gesaneerd werd significant (p-waarde = 0.0165). Het finaal gefitte model (N=31), zonder de beïnvloedende woningen, heeft als verklarende factoren: concentratie Pb in zwevend stof buitenshuis (ng/m^3), of er al dan niet gerookt wordt in de woning (0-1) en of het huis al dan niet gesaneerd werd (0-1) (Tabel 3). Dit model heeft een aangepaste R-kwadraat van 0.4923, dit wil zeggen dat dit model 49% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) verklaart. De variantie inflatie factoren (Tabel 3) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.017. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 5 links) waar de punten niet op een rechte lijn liggen. De rechterkant van Figuur 5 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's

General test (p -waarde=0.9336) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 5, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p -waarde (lack of fit)=0.4844.



Figuur 5: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers

→ **Concentratie Pb in veegmonsters (mg/kg ds)**

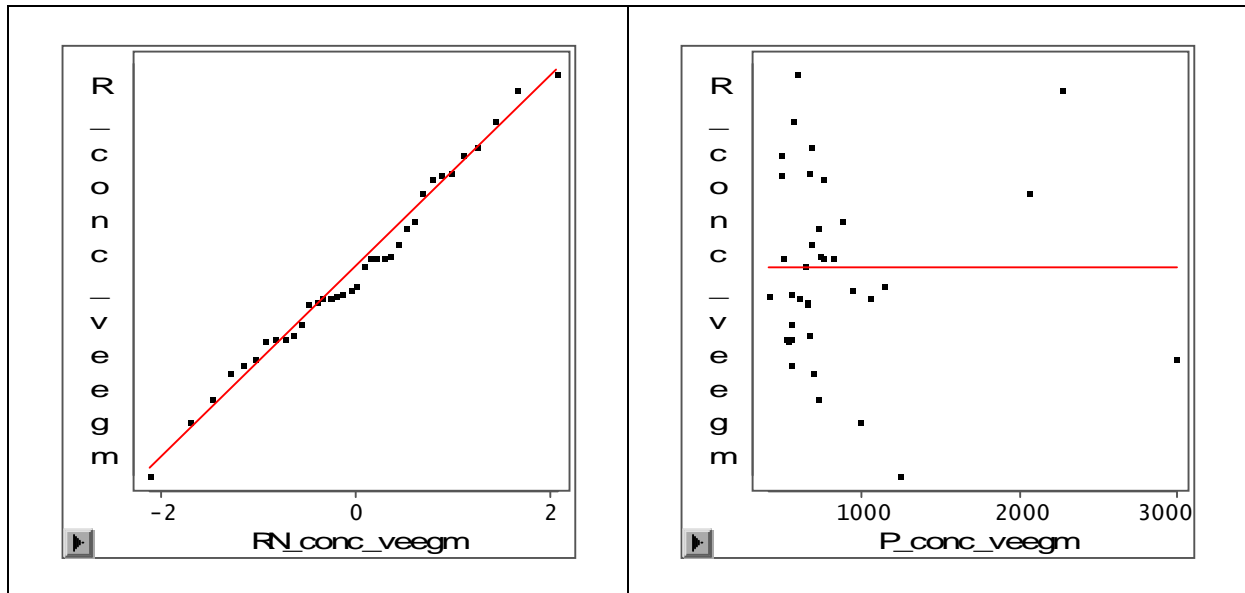
Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model ($N=34$): concentratie Pb in zwevend stof buitenshuis (ng/m^3), de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) en de concentratie Pb in de bodem in fractie $< 250\mu\text{m}$ ($\text{mg}/\text{kg ds}$) (Tabel 4). Dit model heeft een aangepaste R-kwadraat van 0.6668, dit wil zeggen dat dit model 67% van de variabiliteit in de concentratie van Pb in de veegmonsters ($\text{mg}/\text{kg ds}$) verklaart. De variantie inflatie factoren (Tabel 4) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid of score	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5 en 30		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		108.42 (-242.56; 459.40)	0.5329	0	310.64 (-19.29; 640.56)	0.0640	0
Conc Pb zwevend stof buiten	ng/m ³	5.00 (1.20; 8.80)	0.0117	1.43	4.57 (0.15; 9.00)	0.0433	1.88
Belading straatstof	Pb µg/m ²	0.03 (0.02; 0.04)	0.0001	1.50	0.02 (0.00; 0.04)	0.0188	1.88
Conc bodem	Pb mg/kg ds	0.56 (-0.00; 1.13)	0.0506	1.06	ns	ns	ns

ns: not significant

Tabel 4: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 6 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 6 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.4409) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 6, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.

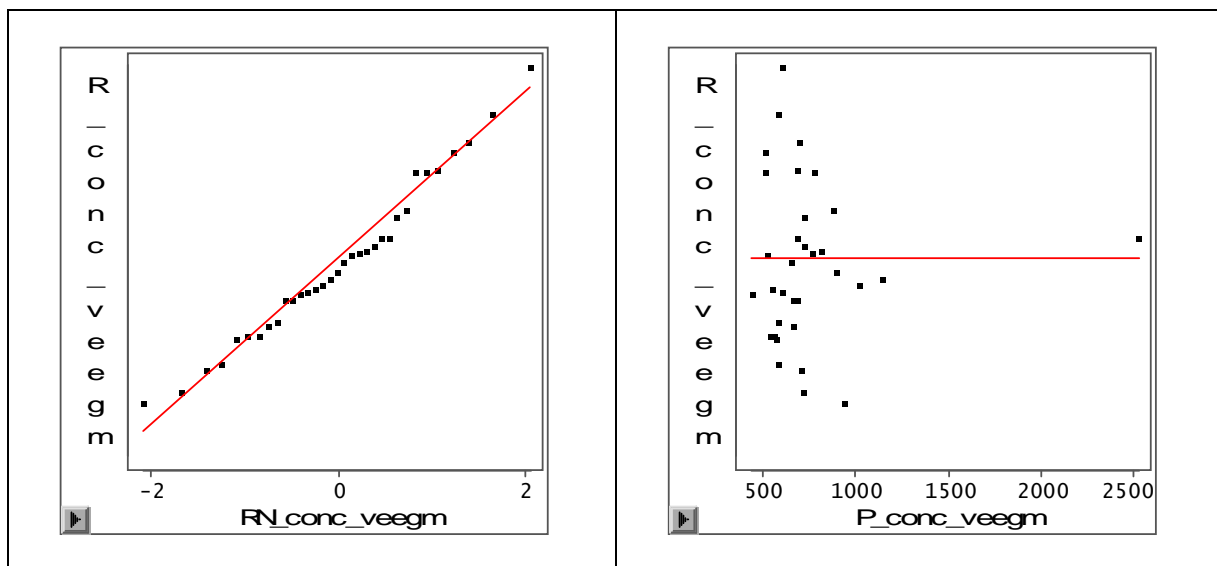


Figuur 6: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden twee outliers gevonden met betrekking tot de concentratie van Pb in de veegmonsters, en met behulp van de leverages werden vijf outliers geïdentificeerd met betrekking tot

de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk één en drie van hen waren van invloed. Wanneer we het model zonder deze drie beïnvloedende outliers (namelijk woning 2, 5 en 30) fitten, wordt de variabele concentratie Pb in bodem in fractie < 250 μm insignificant (p-waarde=0.4898). Het finaal gefitte model (N=33), zonder de beïnvloedende woningen, heeft als verklarende factoren: concentratie Pb in zwevend stof buitenshuis (ng/m^3) en de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 4). Dit model heeft een aangepaste R-kwadraat van 0.4956, dit wil zeggen dat dit model 50% van de variabiliteit in de concentratie van Pb in de veegmonsters (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 4) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 7 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 7 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.8028) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 7, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 7: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers

3.1.3 Kwantificatie van de covariaten

De modellen zonder beïnvloedende outliers die geselecteerd werden in de vorige paragraaf, werden gebruikt om de relatie tussen iedere verklarende factor en de afhankelijke variabele, zowel voor concentratie Pb in de stofzuigerzak, als de hoeveelheid Pb in de veegmonsters als de concentratie Pb in de veegmonsters, te kwantificeren.

De aangepaste R-kwadraat (coëfficiënt van determinatie) van de bekomen multipel lineaire regressiemodellen toont de proportie van variabiliteit in de respons die in

rekening gebracht wordt door dit model, aangepast omdat het aantal verklarende factoren in het model in rekening wordt gebracht. De relatieve belangrijkheid van de verklarende factor in een multipel lineair regressiemodel werd bepaald door de General Dominance Index. Dit is de gemiddelde toename van de coëfficiënt van determinatie geassocieerd met de predictor over alle mogelijke submodellen.

Voor de concentratie van lood in de stofzuiger werd 41% van de variabiliteit verklaard door het model, die in dit geval geheel bepaald werd door de concentratie van lood in de bodem, en dus niet aangepast moest worden voor het aantal verklarende factoren in het model. Voor de belading van lood in de veegmonsters werd 49% van de variabiliteit verklaard door het model, waarbij de concentratie van lood in zwevend stof buitenshuis de grootste bijdrage leverde (33%), met een kleiner percentage voor het al dan niet roken in huis (11%) en het feit of het huis al dan niet gesaneerd is (10%). Voor de concentratie van lood in de veegmonsters werd 50% van de variabiliteit verklaard door het model, met een ongeveer gelijke bijdrage van de concentratie van lood in zwevend stof buitenshuis (25%) en de hoeveelheid lood in straatstof (28%) (Tabel 5).

	<i>Concentratie Pb stofzuiger (mg/kg ds)</i>	<i>Belading veegmonsters (µg/m²)</i>	<i>Pb Concentratie Pb veegmonsters (mg/kg ds)</i>
Covariaten	<i>General Dominance Index (%)</i>		
Concentratie Pb bodem	40.8	ns	ns
Concentratie Pb zwevend stof buiten	ns	33.2	25.0
Roken	ns	11.4	ns
Huis gesaneerd	ns	9.7	ns
Belading Pb straatstof	ns	ns	27.7
	<i>Aangepaste* R-kwadraat (%)</i>		
Model	40.8	49.2*	49.6*

ns: niet significant

Tabel 5: De (aangepaste) R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading (µg/m²) en de concentratie (mg/kg ds) Pb in de veegmonsters

Vertrekkende vanuit het multipel lineair regressiemodel werden de estimates van de regressiecoëfficiënten gebruikt om de relatie tussen de respons en de significante determinanten te kwantificeren. Een positieve relatie werd vastgesteld tussen de concentratie van lood in de bodem en lood in de stofzuiger, waarbij een stijging van lood in de bodem met één eenheid tot een ongeveer recht evenredige stijging van lood in de stofzuiger met één eenheid leidde. Voor de belading van Pb in de veegmonsters werd een positieve relatie bevonden met de concentratie van Pb in zwevend stof buitenshuis. Roken in de woning leidde eveneens tot een grotere hoeveelheid van Pb in de veegmonsters. Sanering van de woning zorgde daarentegen voor een daling van deze hoeveelheid. De concentratie van Pb in de veegmonsters werd positief beïnvloed door de concentratie van Pb in zwevend stof buitenshuis en de belading van Pb in straatstof (Tabel 6).

		Concentratie stofzuiger (mg/kg ds)	Pb	Belading veegmonsters ($\mu\text{g}/\text{m}^2$)	Pb	Concentratie veegmonsters (mg/kg ds)	Pb
Covariaten	Waarde	Verandering					
Concentratie Pb bodem	Stijging met 1 mg/kg ds	Stijging met 1 mg/kg ds	1	ns		ns	
Concentratie Pb zwevend stof buiten	Stijging met 1 ng/m ³	ns		Stijging met 0.18 $\mu\text{g}/\text{m}^2$	0.18	Stijging met 4.57 mg/kg ds	4.57
Roken	Ja	ns		18.74 $\mu\text{g}/\text{m}^2$ hoger dan indien niet		ns	
Huis gesaneerd	Ja	ns		6.71 $\mu\text{g}/\text{m}^2$ lager dan indien niet		ns	
Belading straatstof Pb	Stijging met 1 $\mu\text{g}/\text{m}^2$	ns		ns		Stijging met 0.02 mg/kg ds	0.02

ns: niet significant

Tabel 6: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat

3.2 Dataset met bodemconcentraties van de Universiteit van Gent

In deze analyse werden als bodemconcentraties de door de Universiteit van Gent geïnterpoleerde concentraties gebruikt. Deze bodemconcentraties werden door Haskoning opgezocht aan de hand van de Lambertcoördinaten van de woningen. De resultaten zijn bepaald op de bovenste 30 cm van bodems op openbaar terrein.

3.2.1 Descriptieve analyse

De descriptieve analyse is opgenomen onder 3.1.1.

3.2.2 Identificatie van de covariaten

→ **Concentratie Pb in stofzuigerzak (mg/kg ds)**

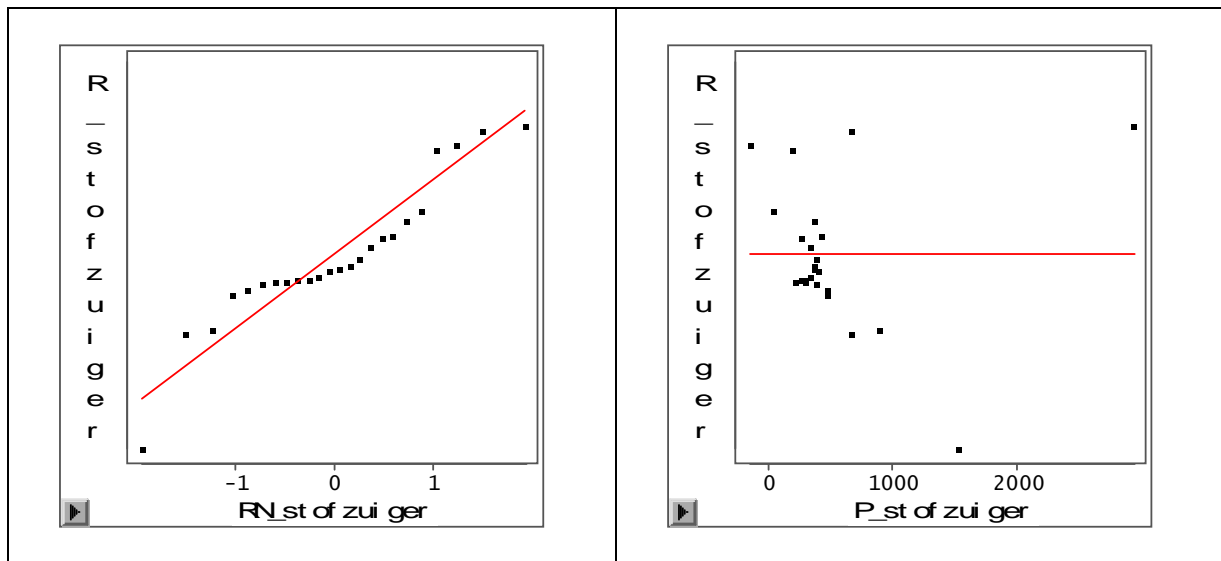
Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de stofzuigerzak bleven de volgende significante variabelen over in het finale model (N=24): concentratie Pb in bodem in fractie < 250 μm (mg/kg ds) en de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 7). Dit model heeft een aangepaste R-kwadraat van 0.6119, dit wil zeggen dat dit model 61% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 7) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5, 30 en 44		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		444.27 (114.59; 773.95)	0.0107	0	2.50 (-267.63; 272.64)	0.9849	0
Concentratie Pb bodem	mg/kg ds	-0.76 (-1.55; 0.03)	0.0592	1.05	ns	ns	ns
Belading straatstof Pb	$\mu\text{g}/\text{m}^2$	0.05 (0.04; 0.07)	< 0.0001	1.05	ns	ns	ns
Ouderdom woning	jaren	ns	ns	ns	6.98 (2.02; 11.95)	0.0079	1.00

ns: niet significant

Tabel 7: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.081. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 8 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 8 toont dat de variabiliteit van de error termen verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0027) werden constante varianties verworpen.

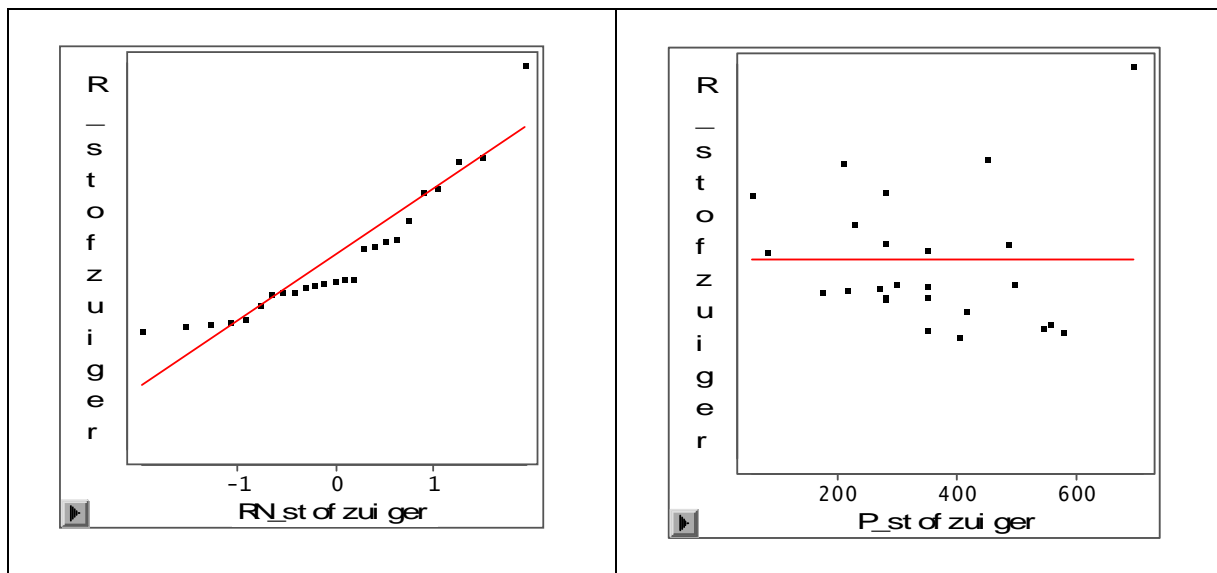


Figuur 8: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden twee outliers gevonden met betrekking tot de concentratie van Pb in de stofzuiger, en met behulp van de leverages werden twee outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk twee en twee van hen waren van invloed. Wanneer we het model zonder deze drie beïnvloedende outliers (namelijk woning 2, 5 en 30) fitten, worden de concentratie van Pb in straatstof insignificant (p-

waarde=0.2143) en concentratie Pb in bodem in fractie < 250 μ m (p-waarde=0.9088). Het finaal gefitte model (N=25), zonder de beïnvloedende woningen, heeft als verklarende factor enkel de ouderdom van de woning (Tabel 7). Dit model heeft een R-kwadraat van 0.2691, dit wil zeggen dat dit model 27% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart.

De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test kleiner is dan 0.01. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 9 links) waar de punten niet op een rechte lijn liggen. De rechterkant van Figuur 9 toont dat de variabiliteit van de error termen ongeveer verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0009) werden constante varianties verworpen. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 9, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test uitgevoerd die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.0940.



Figuur 9: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers

→ **Belading Pb in veegmonsters (μ g/m²)**

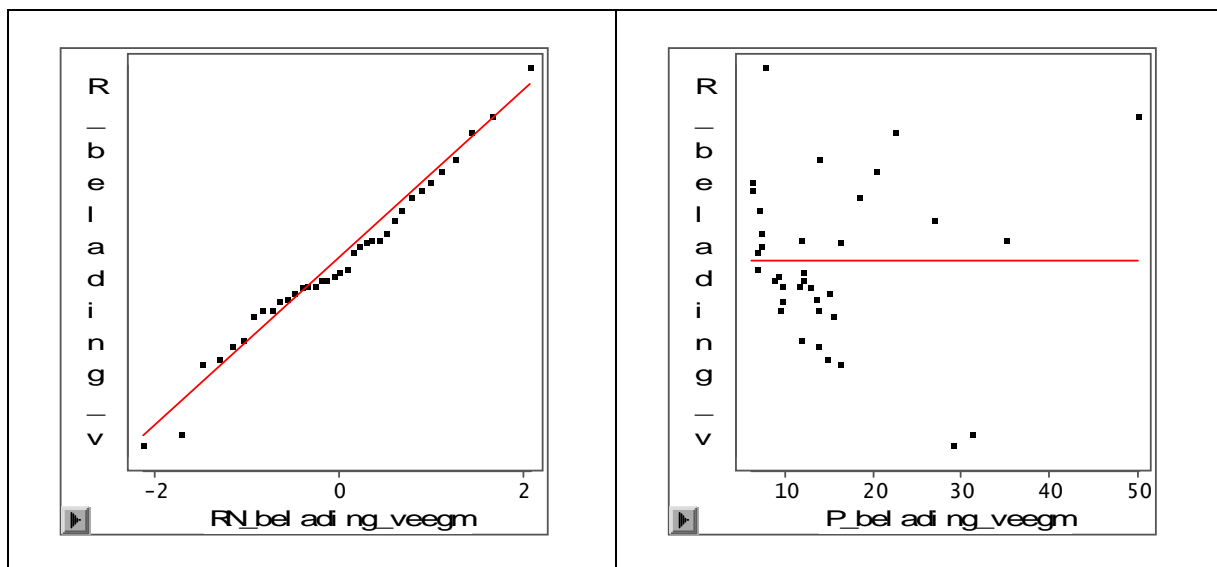
Bij het opbouwen van een multipel lineair regressiemodel voor de belading van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model (N=36): concentratie Pb in zwevend stof buitenshuis (ng/m³) en of er al dan niet gerookt wordt in de woning (0-1) (Tabel 8). Dit model heeft een aangepaste R-kwadraat van 0.4670, dit wil zeggen dat dit model 47% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters (μ g/m²) verklaart. De variantie inflatie factoren (Tabel 8) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Met beïnvloedende woningen				Zonder beïnvloedende woningen 18, 26, 28, 32 en 62		
	Eenheid of score	Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		-1.46 (-8.65; 5.74)	0.6830	0	3.60 (-3.94; 11.15)	0.3357	0
Conc Pb zwevend stof buiten	ng/m ³	0.18 (0.10; 0.26)	< 0.0001	1.03	0.18 (0.09; 0.26)	0.0002	1.01
Roken	0 = niet roken in woning 1 = roken in woning	11.38 (1.61; 21.15)	0.0237	1.03	18.74 (3.60; 33.87)	0.0171	1.04
Huis gesaneerd	0 = huis niet gesaneerd 1 = huis gesaneerd	ns	ns	ns	-6.71 (-12.10; -1.32)	0.0165	1.02

ns: niet significant

Tabel 8: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers

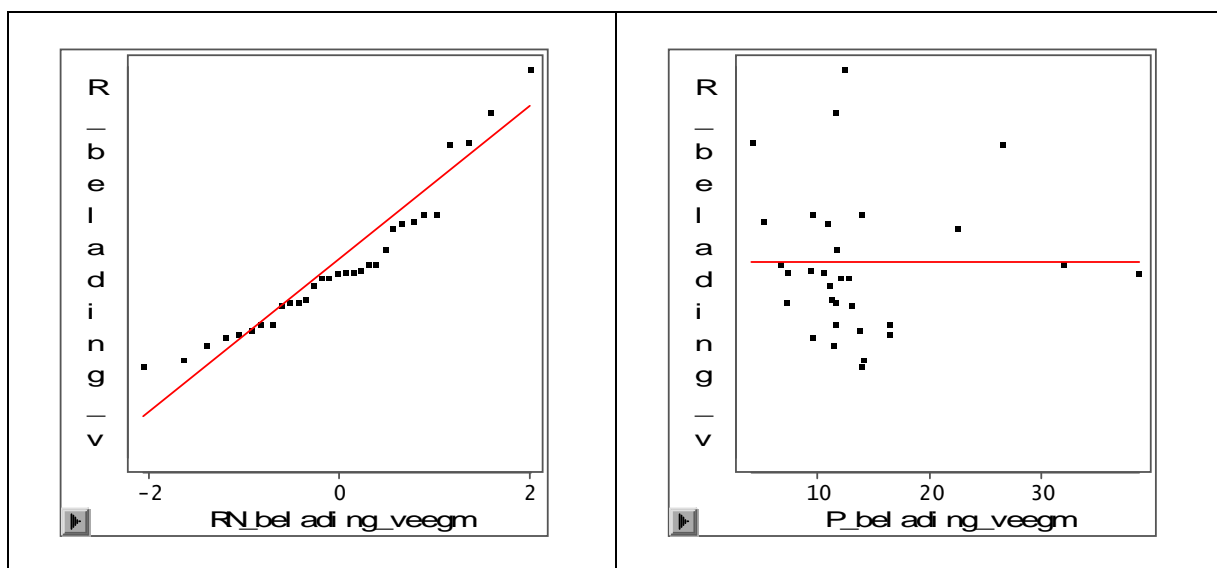
De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 10 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 10 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0984) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 10, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.1220.



Figuur 10: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden vier outliers gevonden met betrekking tot de belading van Pb in de veegmonsters, en met behulp van de leverages werden zeven outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk drie en vijf van hen waren van invloed. Wanneer we het model zonder deze vijf beïnvloedende outliers (namelijk woning 18, 26, 28, 32 en 62) fitten, wordt de variabele of het huis al dan niet gesaneerd werd significant (p-waarde = 0.0165). Het finaal gefitte model (N=31), zonder de beïnvloedende woningen, heeft als verklarende factoren: concentratie Pb in zwevend stof buitenshuis (ng/m^3), of er al dan niet gerookt wordt in de woning (0-1) en of het huis al dan niet gesaneerd werd (0-1) (Tabel 8). Dit model heeft een aangepaste R-kwadraat van 0.4923, dit wil zeggen dat dit model 49% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) verklaart. De variantie inflatie factoren (Tabel 8) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.017. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 11 links) waar de punten niet op een rechte lijn liggen. De rechterkant van Figuur 11 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.9336) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 11, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.4844.



Figuur 11: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers

→ **Concentratie Pb in veegmonsters ($\text{mg}/\text{kg ds}$)**

Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model

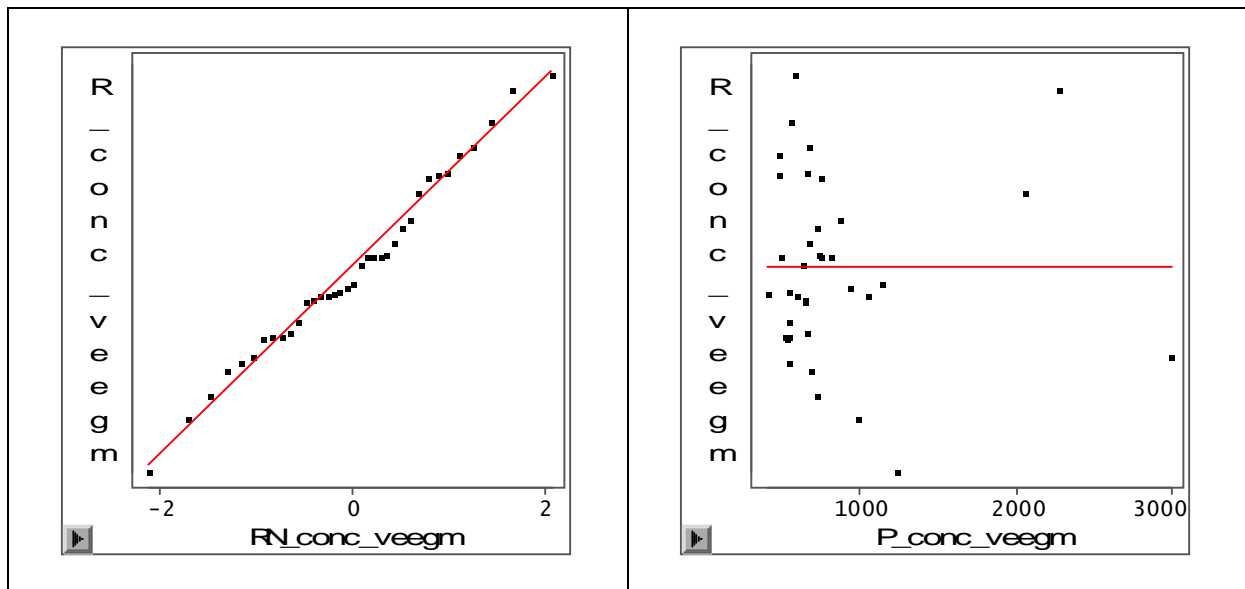
(N=36): concentratie Pb in zwevend stof buitenshuis (ng/m³) en de hoeveelheid Pb in straatstof (µg/m²) (Tabel 9). Dit model heeft een aangepaste R-kwadraat van 0.6353, dit wil zeggen dat dit model 64% van de variabiliteit in de concentratie van Pb in de veegmonsters (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 9) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid of score	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5 en 30		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		275.02 (-34.36; 584.40)	0.0796	0	261.85 (-66.83; 590.52)	0.1142	0
Conc Pb zwevend stof buiten	ng/m ³	4.71 (0.86; 8.56)	0.0180	1.42	5.06 (0.50; 9.62)	0.0307	1.69
Belading Pb straatstof	µg/m ²	0.03 (0.02; 0.04)	< 0.0001	1.42	0.03 (-0.00; 0.07)	0.0693	1.69

ns: not significant

Tabel 9: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers

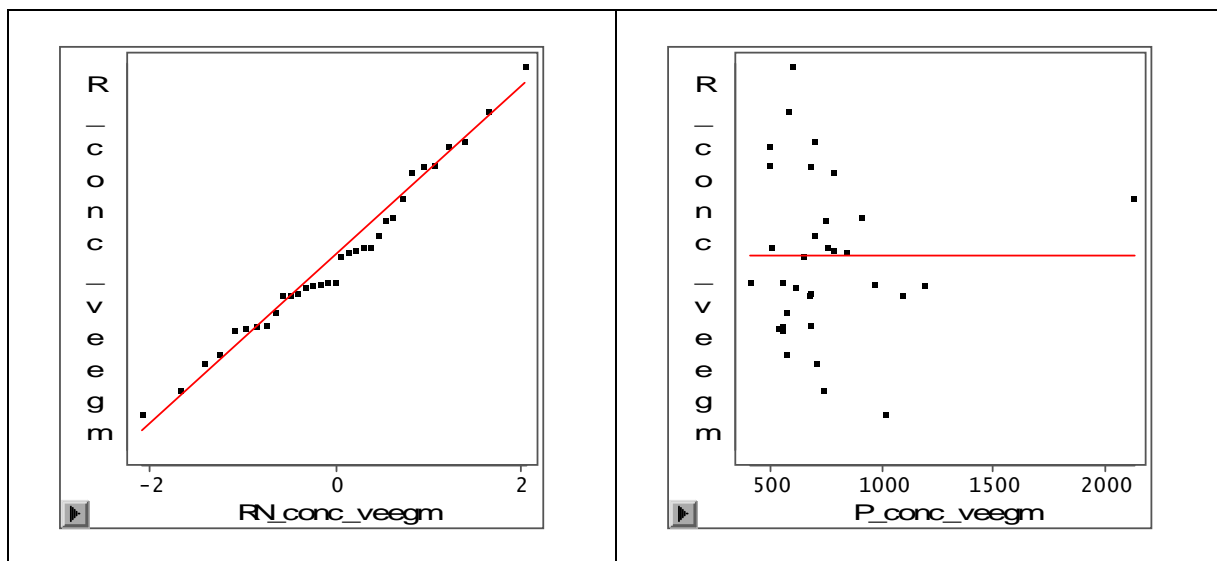
De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 12 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 12 toont dat de variabiliteit van de error termen verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0095) werden constante varianties verworpen. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 12, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 12: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden drie outliers gevonden met betrekking tot de concentratie van Pb in de veegmonsters, en met behulp van de leverages werden vijf outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk twee en drie van hen waren van invloed. Wanneer we het model zonder deze drie beïnvloedende outliers (namelijk woning 2, 28 en 30) fitten, blijven dezelfde variabelen behouden in het model. Het finaal gefitte model (N=33), zonder de beïnvloedende woningen, heeft dus als verklarende factoren: concentratie Pb in zwevend stof buitenshuis (ng/m^3) en de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 9). Dit model heeft een aangepaste R-kwadraat van 0.4056, dit wil zeggen dat dit model 41% van de variabiliteit in de concentratie van Pb in de veegmonsters (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 9) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 13 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 13 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.7946) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 13, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 13: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers

3.2.3 Kwantificatie van de covariaten

De modellen zonder beïnvloedende outliers die geselecteerd werden in de vorige paragraaf, werden gebruikt om de relatie tussen iedere verklarende factor en de afhankelijke variabele, zowel voor concentratie Pb in de stofzuigerzak, als de hoeveelheid Pb in de veegmonsters als de concentratie Pb in de veegmonsters, te kwantificeren.

De aangepaste R-kwadraat (coëfficiënt van determinatie) van de bekomen multipel lineaire regressiemodellen toont de proportie van variabiliteit in de respons die in rekening gebracht wordt door dit model, aangepast omdat het aantal verklarende factoren in het model in rekening wordt gebracht. De relatieve belangrijkheid van de verklarende factor in een multipel lineair regressiemodel werd bepaald door de General Dominance Index. Dit is de gemiddelde toename van de coëfficiënt van determinatie geassocieerd met de predictor over alle mogelijke submodellen.

Voor de concentratie van lood in de stofzuiger werd 27% van de variabiliteit verklaard door het model, die in dit geval geheel bepaald werd door de ouderdom van de woning, en dus niet aangepast moest worden voor het aantal verklarende factoren in het model. Voor de belading van lood in de veegmonsters werd 49% van de variabiliteit verklaard door het model, waarbij de concentratie van lood in zwevend stof buitenshuis de grootste bijdrage leverde (33%), met een kleiner percentage voor het al dan niet roken in huis (11%) en het feit of het huis al dan niet gesaneerd is (10%). Voor de concentratie van lood in de veegmonsters werd 41% van de variabiliteit verklaard door het model, met een ongeveer gelijke bijdrage van de concentratie van lood in zwevend stof buitenshuis (24%) en de hoeveelheid lood in straatstof (21%) (Tabel 10).

	<i>Concentratie Pb stofzuiger (mg/kg ds)</i>	<i>Belading veegmonsters (µg/m²)</i>	<i>Pb Concentratie Pb veegmonsters (mg/kg ds)</i>
Covariaten	<i>General Dominance Index (%)</i>		
Concentratie Pb zwevend stof buiten	<i>ns</i>	33.2	23.6
Belading Pb straatstof	<i>ns</i>	<i>ns</i>	20.7
Roken	<i>ns</i>	11.4	<i>ns</i>
Huis gesaneerd	<i>ns</i>	9.7	<i>ns</i>
Ouderdom woning	26.9	<i>ns</i>	<i>ns</i>
	<i>Aangepaste* R-kwadraat (%)</i>		
Model	26.9	49.2*	49.6*

ns: niet significant

Tabel 10: De (aangepaste) R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading (µg/m²) en de concentratie (mg/kg ds) Pb in de veegmonsters

Vertrekkende vanuit het multipel lineair regressiemodel werden de estimates van de regressiecoëfficiënten gebruikt om de relatie tussen de respons en de significante determinanten te kwantificeren. Een positieve relatie werd vastgesteld tussen de ouderdom van de woning en lood in de stofzuiger. Voor de belading van Pb in de veegmonsters werd een positieve relatie bevonden met de concentratie van Pb in zwevend stof buitenshuis. Roken in de woning leidde eveneens tot een grotere hoeveelheid van Pb in de veegmonsters. Sanering van de woning zorgde daarentegen voor een daling van deze hoeveelheid. De concentratie van Pb in de veegmonsters werd positief beïnvloed door de concentratie van Pb in zwevend stof buitenshuis en de belading van Pb in straatstof (Tabel 11).

		Concentratie stofzuiger (mg/kg ds)		Pb	Belading veegmonsters ($\mu\text{g}/\text{m}^2$)	Pb	Concentratie veegmonsters (mg/kg ds)	Pb
Covariaten	Waarde	Verandering						
Concentratie Pb zwevend buiten	Stijging met 1 ng/m^3	met	ns		Stijging met 0.18 $\mu\text{g}/\text{m}^2$		Stijging met mg/kg ds	5.06
Belading straatstof	Pb Stijging met 1 $\mu\text{g}/\text{m}^2$	met	ns		ns		Stijging met mg/kg ds	0.03
Roken	Ja		ns		18.74 $\mu\text{g}/\text{m}^2$ hoger dan indien niet		ns	
Huis gesaneerd	Ja		ns		6.71 $\mu\text{g}/\text{m}^2$ lager dan indien niet		ns	
Ouderdom woning	Stijging met 1 jaar	met	Stijging met mg/kg ds	6.98	ns		ns	

ns: niet significant

Tabel 11: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat

3.3 Dataset met geïntegreerde bodemconcentraties

In deze analyse werden voor de gesaneerde woningen als bodemconcentraties de door de Universiteit van Gent geïnterpoleerde concentraties gebruikt. Voor de niet-gesaneerde woningen werden de door VITO gemeten concentraties gebruikt.

3.3.1 Descriptieve analyse

De descriptieve analyse is opgenomen onder 3.1.1.

3.3.2 Identificatie van de covariaten

→ Concentratie Pb in stofzuigerzak (mg/kg ds)

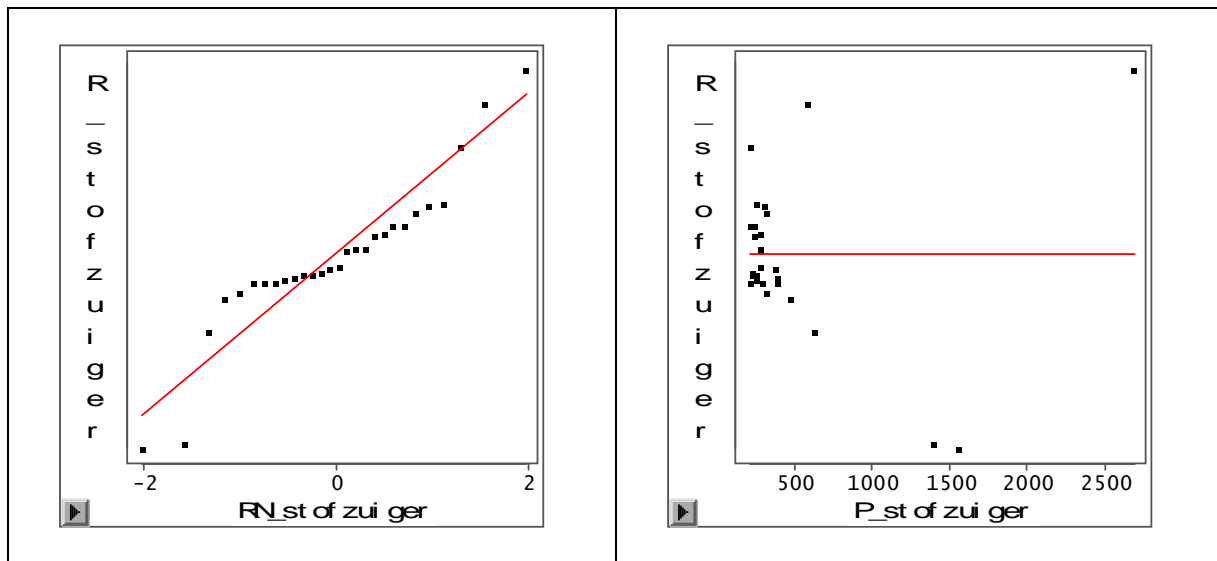
Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de stofzuigerzak bleven de volgende significante variabelen over in het finale model (N=28): de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 12). Dit model heeft een R-kwadraat van 0.5849, dit wil zeggen dat dit model 58% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 12) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5, 30 en 44		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		202.60 (-0.90; 406.09)	0.0509	0	-78.08 (-335.62; 179.46)	0.5352	0
Concentratie Pb bodem	mg/kg ds	ns	ns	ns	0.53 (0.09; 0.96)	0.0193	1.11
Belading straatstof Pb	$\mu\text{g}/\text{m}^2$	0.05 (0.03; 0.07)	< 0.0001	1	ns	ns	ns
Ouderdom woning	jaren	ns	ns	ns	5.10 (0.28; 9.92)	0.0391	1.11

ns: not significant

Tabel 12: Multipel lineair regressiemodel voor de concentratie van Pb in de stofzuiger (mg/kg ds), met en zonder de beïnvloedende outliers

De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test kleiner is dan 0.01. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 14 links) waar de punten zo niet op een rechte lijn liggen. De rechterkant van Figuur 14 toont dat de variabiliteit van de error termen verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde < 0.0001) werden constante varianties verworpen.

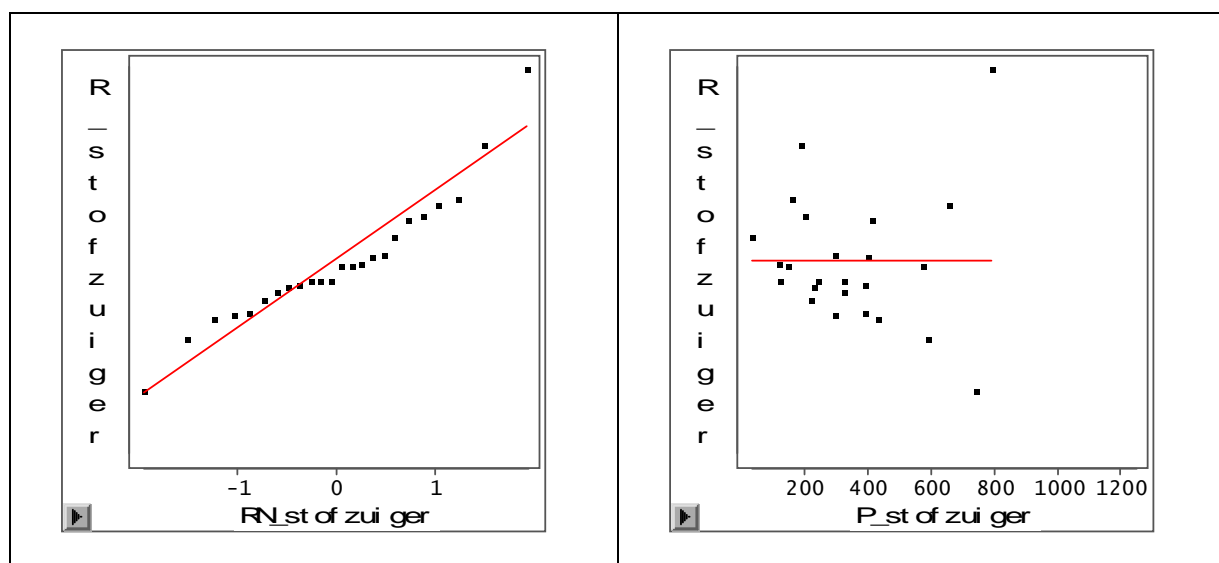


Figuur 14: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden vier outliers gevonden met betrekking tot de concentratie van Pb in de stofzuiger, en met behulp van de leverages werden twee outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk drie en twee van hen waren van invloed. Wanneer we het model zonder deze drie beïnvloedende outliers (namelijk woning 2, 5 en 30) fitten, wordt de belading van Pb in straatstof insignificant (p-

waarde=0.1256). Het finaal gefitte model (N=24), zonder de beïnvloedende woningen, heeft als verklarende factor de concentratie van Pb in de bodem in fractie < 250 μ m (mg/kg ds) en de ouderdom van de woning (jaren) (Tabel 12). Dit model heeft een R-kwadraat van 0.3871, dit wil zeggen dat dit model 39% van de variabiliteit in de concentratie van Pb in de stofzuigerzak (mg/kg ds) verklaart.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.088. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 15 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 15 toont dat de variabiliteit van de error termen verschillend is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0027) werden constante varianties verworpen. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 15, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 15: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de stofzuiger (mg/kg ds) zonder de beïnvloedende outliers

→ **Belading Pb in veegmonsters (μ g/m²)**

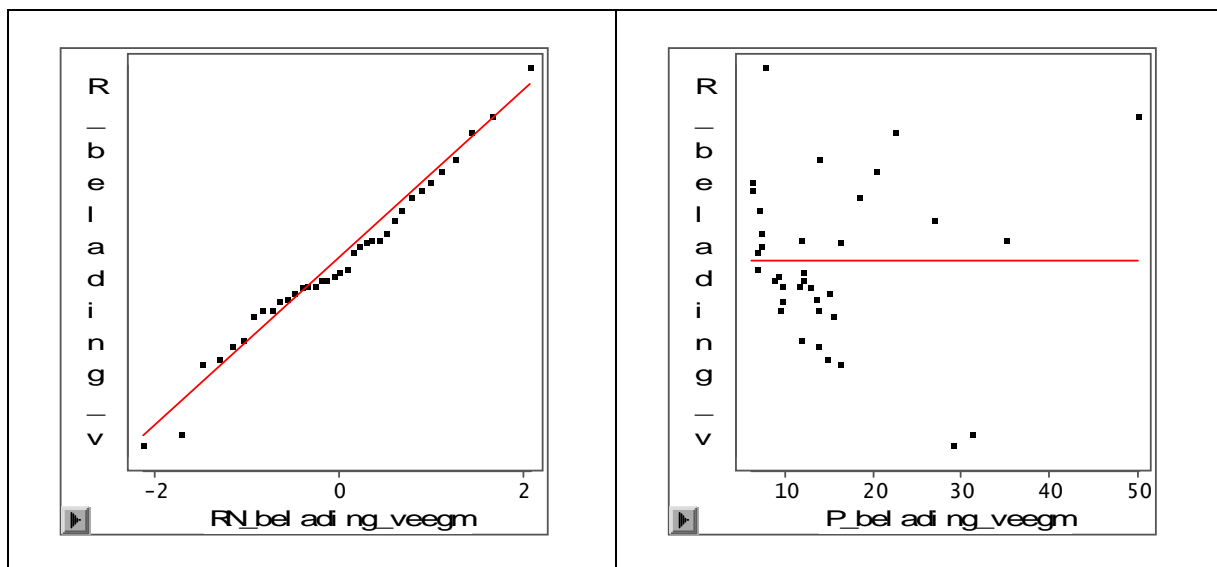
Bij het opbouwen van een multipel lineair regressiemodel voor de belading van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model (N=36): concentratie Pb in zwevend stof buitenshuis (ng/m³) en of er al dan niet gerookt wordt in de woning (0-1) (Tabel 13). Dit model heeft een aangepaste R-kwadraat van 0.4670, dit wil zeggen dat dit model 47% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters (μ g/m²) verklaart. De variantie inflatie factoren (Tabel 13) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Met beïnvloedende woningen				Zonder beïnvloedende woningen 18, 26, 28, 32 en 62		
	Eenheid of score	Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		-1.46 (-8.65; 5.74)	0.6830	0	3.60 (-3.94; 11.15)	0.3357	0
Conc Pb zwevend stof buiten	ng/m ³	0.18 (0.10; 0.26)	< 0.0001	1.03	0.18 (0.09; 0.26)	0.0002	1.01
Roken	0 = niet roken in woning 1 = roken in woning	11.38 (1.61; 21.15)	0.0237	1.03	18.74 (3.60; 33.87)	0.0171	1.04
Huis gesaneerd	0 = huis niet gesaneerd 1 = huis gesaneerd	ns	ns	ns	-6.71 (-12.10; -1.32)	0.0165	1.02

ns: niet significant

Tabel 13: Multipel lineair regressiemodel voor de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$), met en zonder de beïnvloedende outliers

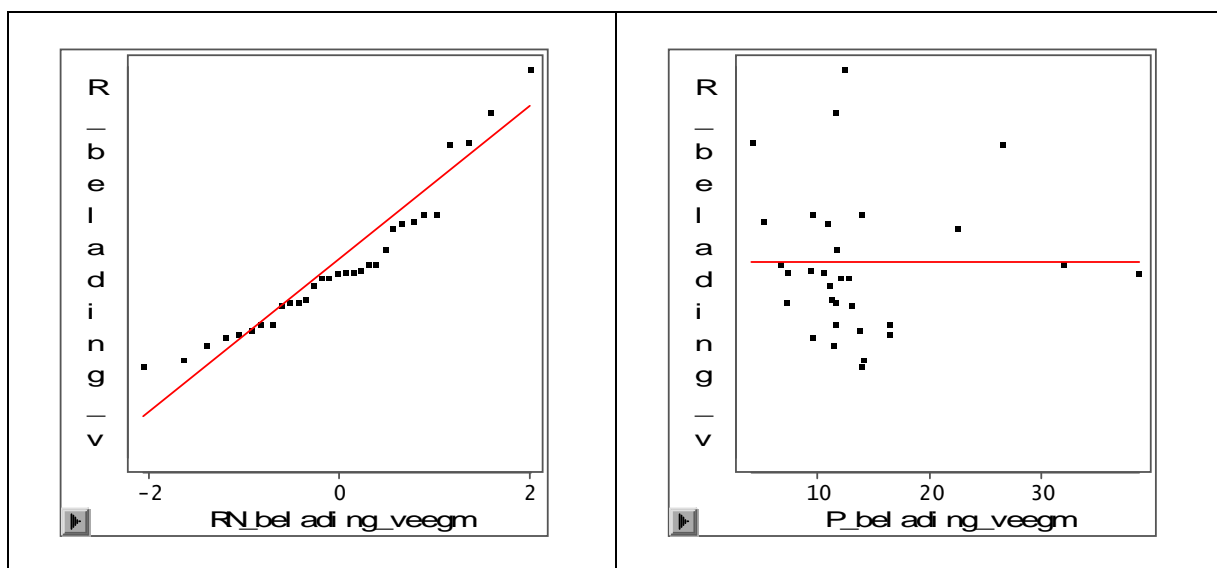
De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 16 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 16 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.0984) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechterkant van Figuur 16, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.1220.



Figuur 16: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden vier outliers gevonden met betrekking tot de belading van Pb in de veegmonsters, en met behulp van de leverages werden zeven outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk drie en vijf van hen waren van invloed. Wanneer we het model zonder deze vijf beïnvloedende outliers (namelijk woning 18, 26, 28, 32 en 62) fitten, wordt de variabele of het huis al dan niet gesaneerd werd significant (p-waarde = 0.0165). Het finaal gefitte model (N=31), zonder de beïnvloedende woningen, heeft als verklarende factoren: concentratie Pb in zwevend stof buitenshuis (ng/m^3), of er al dan niet gerookt wordt in de woning (0-1) en of het huis al dan niet gesaneerd werd (0-1) (Tabel 13). Dit model heeft een aangepaste R-kwadraat van 0.4923, dit wil zeggen dat dit model 49% van de variabiliteit in de hoeveelheid van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) verklaart. De variantie inflatie factoren (Tabel 13) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.017. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 17 links) waar de punten niet op een rechte lijn liggen. De rechterkant van Figuur 17 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.9336) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 17, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt. Om het lineair model te checken, werd de lack of fit test gebruikt die de geschiktheid bevestigde aangezien de p-waarde (lack of fit)=0.4844.



Figuur 17: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de belading van Pb in de veegmonsters ($\mu\text{g}/\text{m}^2$) zonder de beïnvloedende outliers

→ **Concentratie Pb in veegmonsters (mg/kg ds)**

Bij het opbouwen van een multipel lineair regressiemodel voor de concentratie van lood in de veegmonsters bleven de volgende significante variabelen over in het finale model

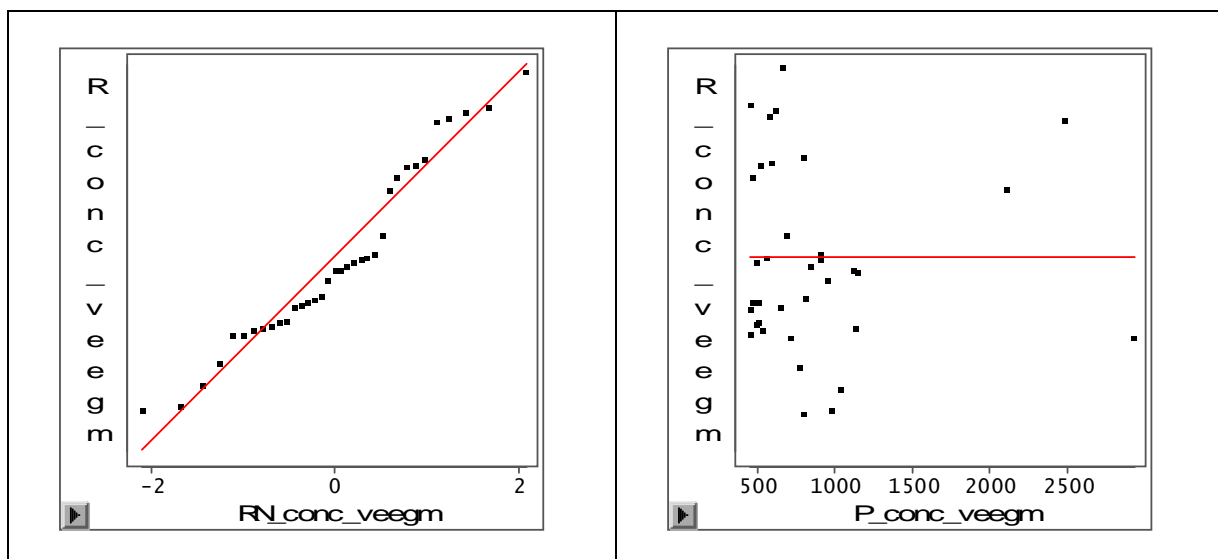
(N=35): de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) en de concentratie Pb in de bodem in fractie < 250 μm (mg/kg ds) (Tabel 14). Dit model heeft een aangepaste R-kwadraat van 0.6872, dit wil zeggen dat dit model 69% van de variabiliteit in de concentratie van Pb in de veegmonsters (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 14) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

Variabelen	Eenheid of score	Met beïnvloedende woningen			Zonder beïnvloedende woningen 2, 5 en 30		
		Estimate (95% BI)	p-waarde	VIF	Estimate (95% BI)	p-waarde	VIF
Intercept		377.03 (180.32; 573.75)	0.0005	0	426.30 (221.35; 631.25)	0.0002	0
Belading straatstof	Pb $\mu\text{g}/\text{m}^2$	0.02 (0.01; 0.04)	0.0022	1.92	0.03 (-0.00; 0.07)	0.0677	1.60
Conc bodem	Pb mg/kg ds	0.85 (0.36; 1.33)	0.0012	1.92	0.67 (0.11; 1.23)	0.0212	1.60

ns: not significant

Tabel 14: Multipel lineair regressiemodel voor de concentratie van Pb in de veegmonsters (mg/kg ds), met en zonder de beïnvloedende outliers

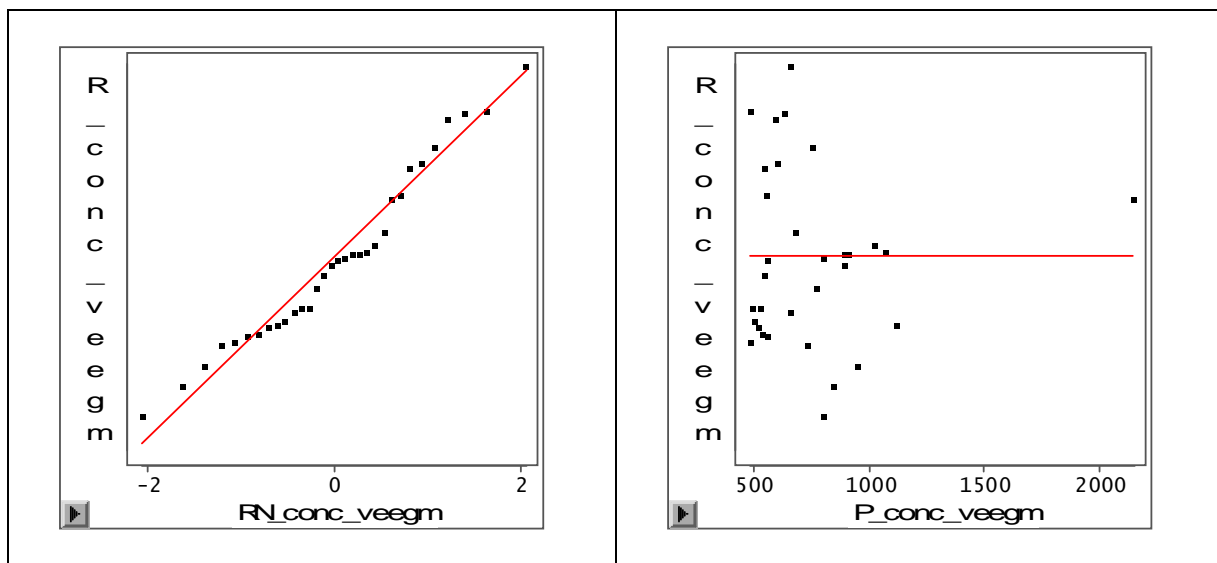
De assumptie van normaliteit van de error termen was niet voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test gelijk is aan 0.036. Dit kan ook geconcludeerd worden van de normal probability plot (links) waar de punten niet op een rechte lijn liggen. De rechterkant van Figuur 18 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.4143) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 18, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 18: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds)

Bij het testen voor outliers met behulp van de studentized deleted residuen werden twee outliers gevonden met betrekking tot de concentratie van Pb in de veegmonsters, en met behulp van de leverages werden vier outliers geïdentificeerd met betrekking tot de verklarende factoren. Op basis van hun DFFITS en Cook's Distance werd nagegaan of deze outliers van invloed waren, respectievelijk twee en drie van hen waren van invloed. Wanneer we het model zonder deze drie beïnvloedende outliers (namelijk woning 2, 28 en 30) fitten, wordt de variabele hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) op het randje insignificant (p -waarde=0.0677). We behouden deze variabele echter in het finale model vermits het echt wel op het randje is, en om toch met een multipel model te kunnen verder werken. Het finaal gefitte model ($N=32$), zonder de beïnvloedende woningen, heeft als verklarende factoren: de concentratie Pb in de bodem in fractie < $250\mu\text{m}$ (mg/kg ds) en de hoeveelheid Pb in straatstof ($\mu\text{g}/\text{m}^2$) (Tabel 14). Dit model heeft een aangepaste R-kwadraat van 0.4193, dit wil zeggen dat dit model 42% van de variabiliteit in de concentratie van Pb in de veegmonsters (mg/kg ds) verklaart. De variantie inflatie factoren (Tabel 14) zijn niet groter dan 10, dit geeft aan dat er geen serieus multicollineariteit probleem aanwezig is.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p -waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 19 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 19 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p -waarde=0.6001) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 19, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 19: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de concentratie van Pb in de veegmonsters (mg/kg ds) zonder de beïnvloedende outliers

3.3.3 Kwantificatie van de covariaten

De modellen zonder beïnvloedende outliers die geselecteerd werden in de vorige paragraaf, werden gebruikt om de relatie tussen iedere verklarende factor en de

afhankelijke variabele, zowel voor concentratie Pb in de stofzuigerzak, als de hoeveelheid Pb in de veegmonsters als de concentratie Pb in de veegmonsters, te kwantificeren.

De aangepaste R-kwadraat (coëfficiënt van determinatie) van de bekomen multipel lineaire regressiemodellen toont de proportie van variabiliteit in de respons die in rekening gebracht wordt door dit model, aangepast omdat het aantal verklarende factoren in het model in rekening wordt gebracht. De relatieve belangrijkheid van de verklarende factor in een multipel lineair regressiemodel werd bepaald door de General Dominance Index. Dit is de gemiddelde toename van de coëfficiënt van determinatie geassocieerd met de predictor over alle mogelijke submodellen.

Voor de concentratie van lood in de stofzuiger werd 39% van de variabiliteit verklaard door het model, met een ongeveer gelijke bijdrage van lood in de bodem (24%) en ouderdom van de woning (20%). Voor de belading van lood in de veegmonsters werd 49% van de variabiliteit verklaard door het model, waarbij de concentratie van lood in zwevend stof buitenshuis de grootste bijdrage leverde (33%), met een kleiner percentage voor het al dan niet roken in huis (11%) en het feit of het huis al dan niet gesaneerd is (10%). Voor de concentratie van lood in de veegmonsters werd 42% van de variabiliteit verklaard door het model, met een ongeveer gelijke bijdrage van de concentratie van lood in de bodem (26%) en de hoeveelheid lood in straatstof (27%) (Tabel 15).

	Concentratie Pb stofzuiger (mg/kg ds)	Belading Pb veegmonsters ($\mu\text{g}/\text{m}^2$)	Concentratie Pb veegmonsters (mg/kg ds)
Covariaten			
General Dominance Index (%)			
Concentratie Pb bodem	24.1	ns	26.0
Concentratie Pb zwevend stof buiten	ns	33.2	ns
Belading Pb straatstof	ns	ns	27.2
Roken	ns	11.4	ns
Huis gesaneerd	ns	9.7	ns
Ouderdom woning	19.9	ns	ns
Aangepaste R-kwadraat (%)			
Model	38.7	49.2	41.9

ns: niet significant

Tabel 15: De aangepaste R-kwadraat van het model en de General Dominance Index van de verklarende factoren voor de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters

Vertrekkende vanuit het multipel lineair regressiemodel werden de estimates van de regressiecoëfficiënten gebruikt om de relatie tussen de respons en de significante determinanten te kwantificeren. Een positieve relatie werd vastgesteld tussen de concentratie van lood in de bodem en lood in de stofzuiger. Eveneens werd er een positieve relatie vastgesteld tussen de ouderdom van de woning en de concentratie van lood in de stofzuiger. Hoe ouder de woning, hoe hoger de concentratie van lood in de stofzuiger. Voor de belading van Pb in de veegmonsters werd een positieve relatie bevonden met de concentratie van Pb in zwevend stof buitenshuis. Roken in de woning leidde eveneens tot een grotere hoeveelheid van Pb in de veegmonsters. Sanering van de woning zorgde daarentegen voor een daling van deze hoeveelheid. De concentratie van Pb in de veegmonsters werd positief beïnvloed door de concentratie van Pb in de bodem en de belading van Pb in straatstof (Tabel 16).

		Concentratie stofzuiger (mg/kg ds)	Pb	Belading veegmonsters ($\mu\text{g}/\text{m}^2$)	Pb	Concentratie veegmonsters (mg/kg ds)	Pb
Covariaten	Waarde	Verandering					
Concentratie Pb bodem	Stijging met 1 mg/kg ds	Stijging met 1 mg/kg ds	0.53	ns		Stijging met 1 mg/kg ds	0.67
Concentratie Pb zwevend stof buiten	Stijging met 1 ng/m^3	ns		Stijging met 1 $\mu\text{g}/\text{m}^2$	0.18	ns	
Belading straatstof	Pb Stijging met 1 $\mu\text{g}/\text{m}^2$	ns		ns		Stijging met 1 mg/kg ds	0.03
Roken	Ja	ns		18.74 $\mu\text{g}/\text{m}^2$ hoger dan indien niet		ns	
Huis gesaneerd	Ja	ns		6.71 $\mu\text{g}/\text{m}^2$ lager dan indien niet		ns	
Ouderdom woning	Stijging met 1 jaar	Stijging met 1 mg/kg ds	5.10	ns		ns	

ns: niet significant

Tabel 16: Verandering in de concentratie Pb in de stofzuiger (mg/kg ds) en de belading ($\mu\text{g}/\text{m}^2$) en de concentratie (mg/kg ds) Pb in de veegmonsters voor een gegeven variatie van een significante covariaat

HOOFDSTUK 4 BRONTRACERING

4.1.1 Inleiding

In deze studie is het de bedoeling om de verschillende elementen die voorkomen in de stofzuigerzak te relateren met de elementen afkomstig van zwevend stof buiten, straatstof en bodemstof. Het is met andere woorden de bedoeling om de herkomst van de elementen in de stofzuigerzak (=binnenshuis) te achterhalen.

4.1.2 Methodologie

Een **descriptieve analyse** werd uitgevoerd om een algemeen beeld te krijgen van de data. Een tabel met descriptieve statistieken van de concentratie van de elementen in de stofzuiger, het bodemstof, het straatstof en het zwevend stof buiten werd gemaakt om kenmerken van de verdeling te identificeren namelijk het aantal observaties, het gemiddelde en de standaard deviatie (Bland, 2000). Een correlatiematrix geeft de paarsgewijze Spearman correlatiecoëfficiënten weer tussen de verschillende elementen van de stofzuigerzak. De correlatiecoëfficiënt r geeft de mate van samenhang weer tussen twee variabelen en ligt tussen -1 en 1 . De correlatiewaarde 1 is de perfecte (positieve) samenhang, de correlatiewaarde 0 geeft weer dat er geen enkele samenhang is en de correlatiewaarde -1 geeft weer dat er perfecte (negatieve) samenhang is. We zullen deze coëfficiënten uitdrukken in percentages. Als indicatie wordt genomen dat een correlatie van 50% of meer erop wijst dat de variabelen gecorreleerd zijn. Indien deze hoger is dan 70% zijn de variabelen sterk gecorreleerd.

Om de 27 elementen van de stofzuigerzak die waarschijnlijk onderling samenhangen te analyseren, zullen we deze elementen groeperen om als zodanig verder te analyseren. **Principal Components Analysis (PCA)** is een procedure om uw aantal variabelen te reduceren. Deze techniek is nuttig wanneer je informatie verkregen hebt over een aantal variabelen en je denkt dat deze informatie overtollig is. Met overtolligheid bedoelen we hier dat de variabelen onderling gecorreleerd zijn met elkaar. Met behulp van PCA kunnen de geobserveerde variabelen gereduceerd worden in een kleiner aantal "principal components" (artificiële variabelen) dat het grootste deel van de variantie in de geobserveerde variabelen bepaalt. Deze principal components kunnen dan verder gebruikt worden in daaropvolgende analyses. Een principal component of hoofdcomponent is een lineaire combinatie van optimaal gewogen geobserveerde variabelen. Het aantal componenten bepaald door PCA is gelijk aan het aantal geobserveerde variabelen. De eerste paar componenten zijn echter voldoende om genoeg variantie te verklaren. Dus enkel deze worden behouden voor verdere analyse.

De resulterende componenten van de PCA zullen verder geanalyseerd worden door **multiële lineaire regressie**. De methodologie hiervan is opgenomen onder 2.2.2 en 2.2.3. Hiervoor werden dezelfde componenten samengesteld voor bodemstof, straatstof

en zwevend stof buiten. De bedoeling is om iedere component van de stofzuigerzak te relateren met dezelfde component van bodemstof, straatstof en zwevend stof buiten.

4.1.3 Resultaten

Uit Tabel 17 met de **descriptieve statistieken** kunnen we zien dat de metingen van de concentraties van de elementen voor de stofzuigerzak, het bodemstof en straatstof in dezelfde grootteorde liggen. In zwevend stof buiten liggen de concentraties van de verschillende elementen beduidend lager in vergelijking met de andere compartimenten. Voor de stofzuigerzak beschikken we maar voor 28 woningen over gegevens. Voor bodemstof zijn er hier en daar enkele woningen die ontbreken. Voor straatstof zijn er echter voor het grootste deel van de elementen maar metingen beschikbaar voor 11 woningen en voor zwevend stof buiten zijn een aantal elementen zelfs niet bepaald (Tabel 17). Deze kleine aantallen verklaren mede de grote standaard deviaties die weergegeven worden in Tabel 17.

Uit de **correlatiematrix** (Tabel 18) blijkt dat er veel samenhang is tussen de verschillende elementen in de stofzuigerzak. In het groen zijn de Spearman correlaties weergegeven hoger dan 50%, in het rood deze die hoger zijn dan 70% en waarvan bijgevolg de elementen sterk gecorreleerd zijn.

	Concentratie stofzuiger (mg/kg ds)			Concentratie bodemstof (mg/kg ds)			Concentratie straatstof (mg/kg ds)			Concentratie zwevend stof buiten (mg/kg ds)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
NaO	28	6707	2433	26	3531	1069	20	5770	1776	/	/	/
MgO	28	5200	1770	26	2804	1142	20	7260	1714	/	/	/
Al ₂ O ₃	28	22417	4448	26	28446	8551	20	35605	9068	29	0.38	0.16
SiO ₂	28	535239	124817	26	802900	83735	20	677685	94606	29	9.24	0.70
P ₂ O ₅	28	1738	540	26	2027	733	20	1992	1008	29	0.38	0.23
K ₂ O	28	7129	1450	26	8008	2767	20	9240	1577	29	3.28	0.44
CaO	28	44014	22038	26	5531	2668	20	28610	13813	29	18.70	2.02
TiO ₂	28	3007	2343	26	2327	642	20	2550	566	29	0.53	0.04
MnO	28	174	50	26	205	126	20	378	127	29	0.36	0.02
Fe ₂ O ₃	28	14343	6231	26	14396	8406	20	24195	6648	29	20.32	0.58
Cr	28	69	49	26	45	16	20	83	30	29	0.23	0.01
S	28	13358	11209	25	1588	866	11	4937	1209	29	2.71	0.95
Cl	28	5324	2765	25	68	60	11	328	219	/	/	/
V	28	18	11	25	34	18	11	32	13	29	0.07	0.00
Co	28	13	7	25	8	5	11	8	5	/	/	/
Ni	28	46	29	25	16	9	11	49	23	29	0.14	0.00
Cu	28	288	336	25	75	52	11	340	215	29	0.71	0.11
Zn	28	717	491	25	176	128	11	657	312	29	1.72	0.12
As	28	22	23	26	29	24	11	39	33	29	0.59	0.24
Se	28	3	3	25	2	2	11	13	15	/	/	/
Br	28	52	191	25	5	2	11	4	3	/	/	/
Rb	28	25	6	25	30	9	11	33	6	/	/	/
Sr	28	246	376	25	46	13	11	78	12	/	/	/
Zr	28	167	47	25	251	73	11	155	44	/	/	/
Cd	28	5	7	26	6	6	11	15	13	29	0.82	0.04
Ba	28	663	834	25	299	91	11	394	124	/	/	/
Pb	28	475	711	26	299	275	11	819	640	29	1.61	0.74

Tabel 17: Descriptieve statistieken voor de concentratie van elementen in de stofzuiger, het bodemstof, straatstof en zwevend stof buiten (mg/kg ds)

Hoofdstuk 4 Brontracering

	NaO	MgO	Al ₂ O ₃	SiO ₂	P ₂ O ₅	K ₂ O	CaO	TiO ₂	MnO	Fe ₂ O ₃	Cr	S	Cl	V	Co	Ni	Cu	Zn	As	Se	Br	Rb	Sr	Zr	Cd	Ba	Pb	
NaO	100																											
MgO		100																										
Al ₂ O ₃			100																									
SiO ₂				100																								
P ₂ O ₅					100																							
K ₂ O						100																						
CaO							100																					
TiO ₂								100																				
MnO									100																			
Fe ₂ O ₃										100																		
Cr											100																	
S												100																
Cl													100															
V														100														
Co															100													
Ni																100												
Cu																	100											
Zn																		100										
As																			100									
Se																				100								
Br																					100							
Rb																						100						
Sr																							100					
Zr																								100				
Cd																									100			
Ba																										100		
Pb																											100	

Tabel 18: Correlatiematrix met Spearman correlatiecoëfficiënten (%) voor de elementen van de stofzuigerzak. Groen gearceerd: gecorreleerde elementen (r≥50%); rood gearceerd: sterk gecorreleerde elementen (r≥80%)

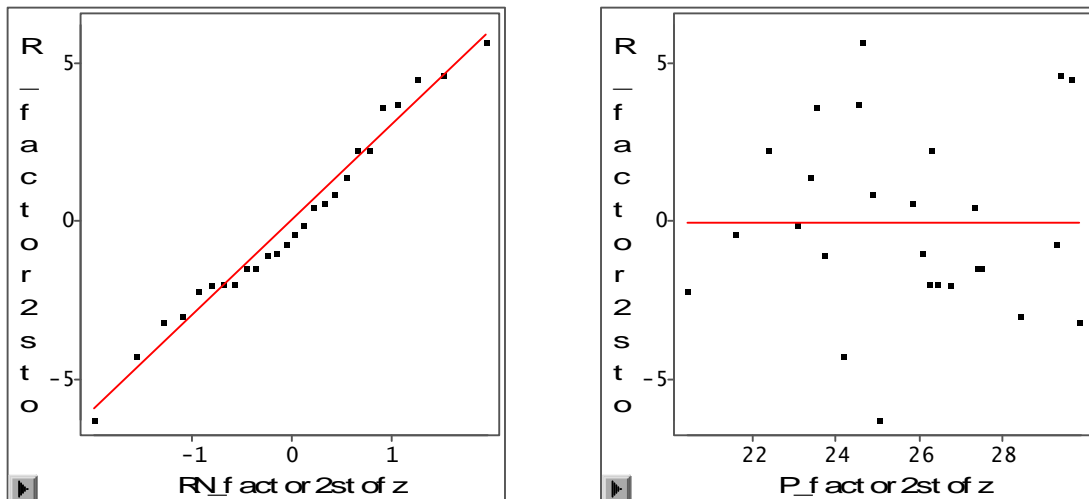
Met behulp van **Principal Component Analysis (PCA)** kunnen we de 27 elementen van de stofzuigerzak reduceren tot vier componenten die samen 70% van de variantie van de elementen in de stofzuigerzak bepalen. De belangrijkste component draagt 34.50% bij en bestaat uit Rb, K₂O, Fe₂O₃, Al₂O₃, MnO en Zr. De tweede component draagt 15% bij en bestaat uit Cd, Pb, As, Zn, Cu, Se en Ni. Deze groep kunnen we duidelijk onderscheiden als de groep van metalen. De derde component draagt 13% bij en bestaat uit Cr, Ba en TiO₂. De vierde component draagt 8% bij en bestaat uit Sr, S en CaO.

Deze vier resulterende componenten van de PCA werden verder geanalyseerd door **multiële lineaire regressie**. Hiervoor werden dezelfde vier componenten samengesteld voor zwevend stof buiten, bodemstof en straatstof. De bedoeling is dan om iedere component van de stofzuigerzak te relateren met dezelfde component van zwevend stof buiten, bodemstof en straatstof.

Bij het opbouwen van een multipel lineair regressiemodel voor component één bestaande uit Rb, K₂O, Fe₂O₃, Al₂O₃, MnO en Zr, bleef de component van bodemstof over als significante variabele in het finale model (N=26). Dit model heeft een R-kwadraat van 0.2321, dit wil zeggen dat dit model 23% van de variabiliteit van component één (Rb, K₂O, Fe₂O₃, Al₂O₃, MnO en Zr) in de stofzuiger verklaart. Wanneer we checken voor beïnvloedende outliers, werd woning 42 geïdentificeerd als een beïnvloedende observatie op het model. Wanneer we deze woning weglaten uit de analyses werd er geen significant model meer bekomen voor component 1.

Bij het opbouwen van een multipel lineair regressiemodel voor component twee bestaande uit Cd, Pb, As, Zn, Cu, Se en Ni, bleef de component van bodemstof over als significante variabele in het finale model (N=26). Dit model heeft een R-kwadraat van 0.4293, dit wil zeggen dat dit model 43% van de variabiliteit van component twee (Cd, Pb, As, Zn, Cu, Se en Ni) in de stofzuiger verklaart. Wanneer we checken voor beïnvloedende outliers, werden geen woningen geïdentificeerd die van invloed waren op het model.

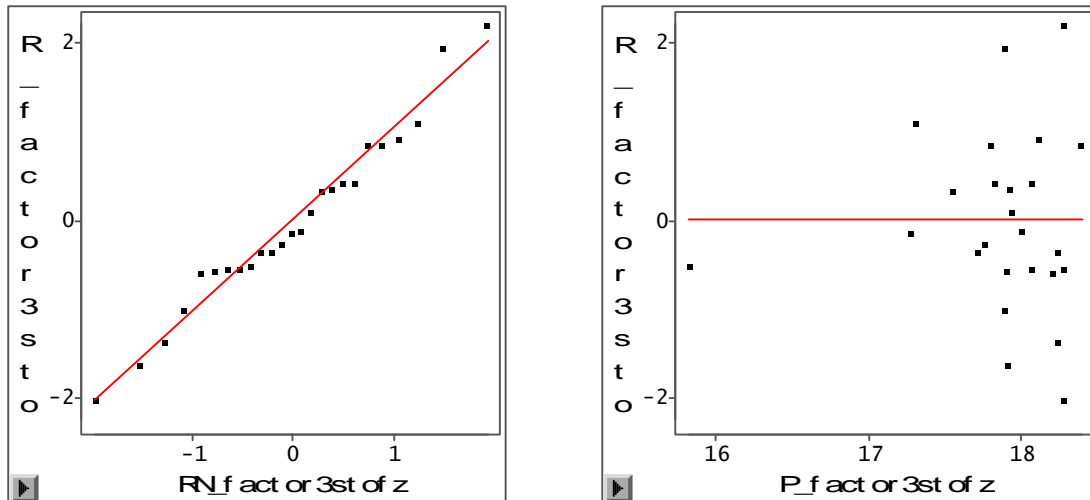
De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 20 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 20 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.8706) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 20, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 20: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de component twee van de stofzuigerzak bestaande uit de elementen Cd, Pb, As, Zn, Cu, Se en Ni

Bij het opbouwen van een multipel lineair regressiemodel voor component drie bestaande uit Cr, Ba en TiO_2 , bleef de component van bodemstof over als significante variabele in het finale model ($N=26$). Dit model heeft een R-kwadraat van 0.2314, dit wil zeggen dat dit model 23% van de variabiliteit van component drie (Cr, Ba en TiO_2) in de stofzuiger verklaart. Wanneer we checken voor beïnvloedende outliers, werd woning 17 geïdentificeerd als een beïnvloedende observatie op het model. Wanneer we deze woning weglaten uit de analyses bleef component drie van bodemstof significant bijdragen aan het finale model ($N=25$). Dit model heeft een aangepaste R-kwadraat van 0.2064, dit wil zeggen dat dit model 21% van de variabiliteit van component drie (Cr, Ba en TiO_2) in de stofzuiger verklaart.

De assumptie van normaliteit van de error termen was voldaan voor dit model gezien de p-waarde van de Kolmogorov-Smirnov test groter is dan 0.15. Dit kan ook geconcludeerd worden van de normal probability plot (Figuur 21 links) waar de punten zo goed als op een rechte lijn liggen. De rechterkant van Figuur 21 toont dat de variabiliteit van de error termen ongeveer gelijk is voor de verschillende niveaus, en ook door White's General test (p-waarde=0.2725) werden constante varianties vastgesteld. De voorwaarde van onafhankelijkheid kan eveneens worden vastgesteld van de rechter Figuur 21, die toont dat de error termen random verdeeld zijn rond het gemiddelde nul en er geen bepaald patroon gevolgd wordt.



Figuur 21: Normal probability plot (links) en plot van de residuen ten opzichte van de voorspelde waardes (rechts) voor het multipel lineair regressiemodel van de component drie van de stofzuigerzak bestaande uit de elementen Cr, Ba en TiO_2

Voor component vier bestaande uit Sr, S en CaO kon er geen significant model bekomen worden. De andere milieucompartimenten dragen hier niet bij tot de variantie in de stofzuigerzak.

4.1.4 Discussie

Met behulp van de Principal Component Analyse werden de elementen van de stofzuigerzak opgedeeld in vier groepen, waarvan één groep duidelijk afgebakend was als de groep van metalen. Voor de brontracering kon voor deze groep een verband aangetoond worden met het bodemstof. In de regressieanalyse kwam immers duidelijk naar voren dat de metalen binnenshuis (stofzuigerzak) voornamelijk afkomstig zijn uit het bodemstof. Deze verklaarde immers 43% van de variantie van Cd, Pb, As, Zn, Cu, Se en Ni in de stofzuigerzak.

Er moet echter wel voorzichtig omgesprongen worden met deze resultaten. Er was immers informatie beschikbaar voor maximaal 29 woningen (zie Tabel 17), wat al zeer weinig is. Bovendien was er voor deze woningen vaak het probleem van ontbrekende waardes, voornamelijk voor straatstof, waar er maar voor 11 woningen alle data beschikbaar waren. Dit kan mede een verklaring zijn voor het feit dat straatstof nooit als significante verklarende variabele werd bevonden. Voor zwevend stof buiten waren een aantal elementen niet gemeten. Daarbovenop kwamen deze metingen van zwevend stof buiten niet van de woningen zelf, maar waren het geëxtrapoleerde metingen van de scholen in de dichtstbijzijnde omgeving (rekening houdende met de windrichting). Deze waren gemeten voor zes scholen in de buurt.

HOOFDSTUK 5 MILIEUMETINGEN PB VERSUS BRONTRACERING

In dit hoofdstuk vergelijken we de resultaten voor Pb verkregen uit de milieumetingen van Hoofdstuk 3 en de brontracering van Hoofdstuk 4.

In hoofdstuk 3 hebben we ons toegespitst op de analyse van Pb als element op zich. Pb wordt echter waarschijnlijk niet als element op zich binnengebracht in de woning. Het is daarom dat we in hoofdstuk 4 hebben gekeken met welke andere elementen Pb samenhangt. Uit de Principal Component Analyse kwam duidelijk naar voren dat Pb als het ware een cluster vormt samen met andere metalen namelijk Cd, As, Zn, Cu, Se en Ni. Als we dan verder gingen onderzoeken uit welk(e) compartiment(en) deze metalen afkomstig zijn kwam bodemstof als enige verklarende factor uit de lineaire regressieanalyse. De factor bodem bepaalde hier 43% van de concentratie van deze groep metalen in de stofzuiger. Dit komt alsnog overeen met de regressieanalyse van hoofdstuk 3, waar de concentratie van Pb in de stofzuiger voor 41% bepaald werd door de concentratie van Pb in de bodem (zie Tabel 5). Het milieucompartiment bodem komt dus in beide analyses naar voren als belangrijkste verklarende factor voor de concentratie van Pb binnenshuis.

Opnieuw moet hier wel de opmerking bij gemaakt worden dat er weinig observaties beschikbaar waren voor de analyses van de brontracering van hoofdstuk 4, mits er maar in 28 woningen metingen gedaan werden. Daar bovenop kwam dan nog dat er voor straatstof voor een aantal elementen maar aan 11 woningen metingen gedaan werden, en voor zwevend stof buiten een aantal elementen zelfs helemaal niet gemeten werden. Als extra onzekerheid waren de metingen van de elementen van zwevend stof buiten aan de woningen geëxtrapoleerde data van metingen aan de scholen en de VMM meetposten in de omgeving. Hier kunnen we dan als opmerking maken hoe exacter de metingen, hoe betrouwbaarder de resultaten. Misschien is het daarom niet al te verbazend dat de milieucompartimenten straatstof en zwevend stof buiten niet als significant verklarende factoren geïdentificeerd konden worden in deze analyses.

Voor de analyses van de milieumetingen van hoofdstuk 3 werden de metingen voor zwevend stof buiten eveneens geëxtrapoleerd. Voor deze analyses waren er wel minder ontbrekende waarden, en was er voor ieder milieucompartiment voldoende informatie beschikbaar. Er waren echter weer maar maximum 28 woningen beschikbaar om een regressiemodel voor te ontwikkelen.

HOOFDSTUK 6 BESLUIT

Op basis van milieumetingen, met drie verschillende reeksen bodemconcentratie-metingen, werd nagegaan of er een relatie is tussen de concentratie of belading van lood in huisstof en de concentraties of belading in bodem, straatstof en zwevend stof binnen. Een overzicht van de bekomen significante relaties is opgenomen in Tabel 19. De weergegeven significante variabelen zijn deze bekomen na verwijdering van outliers uit de dataset.

		<i>Bodemconcentraties VITO</i>	<i>Bodemconcentraties Universiteit van Gent</i>	<i>Geïntegreerde bodemconcentraties</i>
Concentratie stofzuigerzak Variabelen model:	Pb finale	<i>Concentratie Pb bodem</i>	<i>Ouderdom woning</i>	<i>Concentratie Pb bodem Ouderdom woning</i>
Belading Pb veegmonsters Variabelen model:	Pb finale	<i>Concentratie Pb zwevend stof buiten Roken Huis gesaneerd</i>	<i>Concentratie Pb zwevend stof buiten Roken Huis gesaneerd</i>	<i>Concentratie Pb zwevend stof buiten Roken Huis gesaneerd</i>
Concentratie veegmonsters Variabelen model:	Pb finale	<i>Concentratie Pb zwevend stof buiten Belading Pb straatstof</i>	<i>Concentratie Pb zwevend stof buiten Belading Pb straatstof</i>	<i>Concentratie Pb bodem Belading Pb straatstof</i>

Tabel 19: Vergelijking resultaten statistische analyses milieumetingen met bodemconcentraties VITO, Universiteit van Gent en geïntegreerde bodemconcentraties.

HOOFDSTUK 7 LITERATUURLIJST

1. Bland M. An introduction to medical statistics. 3 ed. New York: Oxford University Press, 2000.
2. Neter J, Kutner M, Nachtsheim C, Wasserman W. Applied linear statistical models. New York: McGraw-Hill, 1996.
3. Fox J. Regression diagnostics: an introduction. CA: Sage, 1991.
4. White H. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 1980;**48**:817-838.
5. Chao YC, Zhao Y, Kupper LL, Nylander-French LA. Quantifying the relative importance of predictors in multiple linear regression analyses for public health studies. *J Occup Environ Hyg* 2008;**5**(8):519-29.
6. Van Deun, J., Berghmans, P., Brabers, R., Daems, J., Tirez, K., Govaerts, E., Van Holderbeke, M., Standaert, A., Cornelis, C. (2008). Uitvoeren van milieumetingen voor het bepalen van de externe belasting aan lood in de wijde omgeving van Hoboken. Rapport 2008/IMS/F/080, VITO, Mol.