

# WORDT DE KWALITEIT VAN LOW-COST FIJNSTOFSENSOREN ONDERSCHAT?

## EEN VERGELIJKEND ONDERZOEK EN ZOEKTOCHT NAAR MOGELIJKHEDEN OM HIERVAN DE MEETRESULTATEN TE VERBETEREN

*Met de komst van low-cost fijnstofsensoren is de interesse om zelf fijnstof te gaan meten toegenomen. Wie is er nu niet nieuwsgierig naar de luchtkwaliteit in de eigen achtertuin, de lucht die je dagelijks inademt? Individueel of in kleine groepen worden meetinitiatieven gestart waarbij kennis en data worden gedeeld. Maar hoe betrouwbaar zijn die goedkope sensoren nu eigenlijk? Dit artikel beschrijft een burgerinitiatief wat op die vraag een antwoord probeert te vinden.*

ANDRÉ VAN DER WIEL, DIETER PIENKA

### | Achtergrond

In het kader van informatie en meerwaarde is het 'Visibilis project' ontstaan. Visibilis staat voor zichtbaarheid. Het is de kernboodschap van een *citizen science* project dat gestart is in mei 2018<sup>1</sup>. Het gaat om het zichtbaar maken van de onzichtbare deeltjes in de atmosfeer, die in belangrijke mate de kwaliteit van de

lucht bepalen. Data omzetten naar bruikbare informatie is voor *citizen science* een belangrijk doel, namelijk bewustwording van de leefomgeving. Het samenstellen van de benodigde apparatuur, de ontwikkeling van de software en het analyseren van de data is zelfstandig en binnen het project uitgevoerd. Met hulp van het RIVM en de DCMR is meetlocatie 'Berghaven' te Hoek van Holland beschikbaar gekomen (zie figuur 1), waardoor het mogelijk werd de sensormetingen recht-

streeks te vergelijken met de daar opgestelde professionele apparatuur.

### | De opstelling

Een weerbestendige koffer vol met sensoren en randapparatuur was tijdens het project geïnstalleerd op het dak van de meetlocatie, vlak naast de officiële fijnstofanalysers van DCMR: de 'BAM1020' en de 'Leckel'. De opstelling op deze locatie biedt de mogelijkheid om de prestaties van de sensoren onderling te vergelijken →



Figuur 1: Locatie Visibilis project te Berghaven (Hoek van Holland)

met de meetresultaten van de officiële apparatuur. Als wij weten wat de prestaties van de sensoren zijn en hoe ze zich onderling verhouden, dan wordt het interessant om te kijken naar de mogelijkheid van het opstellen van kalibratie- en validatieprocedures tegen de officiële meetapparatuur.

### *'Doel is om met een algoritme de betrouwbaarheid van de fijnstofsensoren te verbeteren'*

Een fijnstofsensoren gebruikt voor de meting een laser welke gevoelig is voor luchtvochtigheid en in mindere mate ook voor temperatuur. Het is een fenomeen waar veel burgerwetenschappers mee worstelen. Vanwege deze gevoeligheid zijn deze parameters meegenomen in het kalibratieproces. Het uiteindelijke doel is door middel van een algoritme de betrouwbaarheid van de fijnstofsensoren te verbeteren. De resultaten van het Visibilis project zijn in een rapport beschreven en zijn te downloaden via <https://www.sca->

[peler.com/index.php/2019/12/12/project-visibilis-resultaten-vastgelegd-in-rapport/](https://www.sca-peler.com/index.php/2019/12/12/project-visibilis-resultaten-vastgelegd-in-rapport/).

#### | **Sensorkoffer**

De sensoren en randapparatuur zijn ondergebracht in een gemodificeerde 'Dylos' koffer. De koffer is voorzien van beluchtingsgaten aan de onderkant en zij-kanten en deze zijn voorzien van horren-gaas, zodat insecten niet naar binnen kunnen kruipen. De koffer staat op rubberen pootjes, waardoor er ook een luchtstroom onder de koffer kan ontstaan. Een extra ventilator zorgt voor het stimuleren van de interne doorstroming. De samenstelling van de koffer is als volgt:

- Dylos DC1100 Pro serieel gekoppeld (USB) aan Raspberry Pi
- Nova SDS011 serieel gekoppeld (USB) aan Raspberry Pi
- Nova SDS011 als Lufdaten variant met NodeMCU/Wifi gekoppeld aan Raspberry Pi
- 3x Plantower PMSA003 aan Wemos D1 Mini Pro met Wifi gekoppeld aan Raspberry Pi
- 3x Bosch meteo BME280 aan Wemos D1 Mini Pro met Wifi gekoppeld aan Raspberry Pi (meteo in de sensor koffer)

- Dallas DS18B20 One Wire GPIO gekoppeld aan Raspberry Pi (buitentemperatuur)

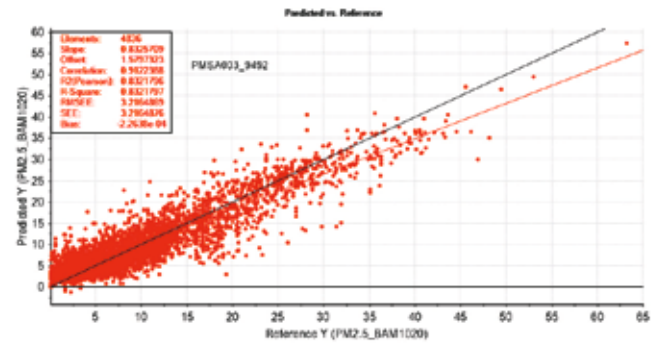
#### | **Database**

Een database met 6131 uurwaarden is samengesteld en dient als basis voor de kalibratie en validatie van de kalibratie algoritmes, kortweg modellen genoemd. Voor alle sensoren geldt dat een uurwaarde het gemiddelde is van 3600 secondewaarden. In geval van missende data zijn deze niet meegenomen in de uurwaarde. Met deze hoge resolutie is het mogelijk tijdelijke piekwaarden goed te meten. Per type fijnstofsensoren is een model gekalibreerd en gevalideerd tegen de BAM1020. Er is gekozen voor de BAM1020 omdat deze monitor in uurwaarden kan meten.

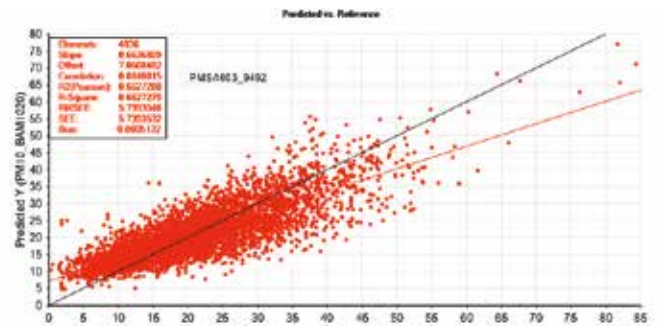


Figuur 2: De koffer gevuld met sensoren en randapparatuur

Figuur 3: PM<sub>2,5</sub>-model  
Plantower PMSA003 sensor



Figuur 4: PM<sub>10</sub>-model  
Plantower PMSA003 sensor



Deze monitor meet echter niet volgens de Europese gravimetrische referentiemethode. De BAM1020 wordt op de locatie gebruikt voor het meten van PM<sub>2,5</sub> en PM<sub>10</sub> en is door DCMR gekalibreerd tegen de Sven Leckel SEQ47/50 die wel een officiële referentiemonitor is. Op deze wijze kan de BAM1020 toch als referentie worden gebruikt voor de kalibratie en validatie van de modellen.

Een korte studie heeft aangetoond dat de onderlinge verschillen in de data van de drie Bosch BME280 meteosensoren binnen de nauwkeurigheid van de sensor zijn gebleven. Voor de kalibratie en validatie van de modellen is de input voor temperatuur en relatieve luchtvochtigheid van de 'gemiddelde' Bosch meteosensor gebruikt. De Dylos sensor meet de deeltjesconcentraties per 0,01 kubieke voet voor deeltjes > 0,5µm en deeltjes > 2,5µm. De Nova sensor meet de fijnstofconcentraties in µg/m<sup>3</sup> voor PM<sub>2,5</sub> en PM<sub>10</sub>, de Plantower sensor voor PM<sub>1</sub>, PM<sub>2,5</sub> en PM<sub>10</sub>. De Plantower sensor kan tevens de deeltjesconcentratie per 0,1 liter meten voor zes deeltjesklassen, namelijk > 0,3 µm, > 0,5 µm, > 1 µm, > 2,5 µm, > 5 µm en > 10 µm.

De database is volgens een alternerend patroon gesplitst in twee delen: 83% is gebruikt als kalibratie set en 17% als validatie set. Uurwaarden uit de validatie set zijn niet gebruikt voor de kalibratie, dit maakt de validatie set onafhankelijk en is tevens een belangrijk aspect in het toetsen van de betrouwbaarheid van modellen.

### Modellen

Voor het Visibilis project zijn voor elk type sensor modellen ontwikkeld voor PM<sub>2,5</sub> en PM<sub>10</sub>. De robuustheid van de modellen is getoetst met de onafhankelijke validatieset. In het Visibilis rapport wordt de ontwikkeling van de modellen zeer uitvoerig beschreven, maar voor dit artikel hebben we er voor gekozen de focus te leggen op één type sensor en kiezen we voor de Plantower PMSA003. Deze sensor kan

naast PM<sub>2,5</sub> en PM<sub>10</sub> ook de deeltjesconcentraties meten en dat maakt deze sensor interessant om te onderzoeken wat de beste modellen gaat opleveren. Het model kan gezien worden als een algoritme dat toegepast wordt op de sensorwaarde, met als doel de betrouwbaarheid van de sensor te verbeteren.

### Meervoudige Lineaire Regressie

De modellen zijn gebaseerd op het principe van Meervoudige Lineaire Regressie (MLR), waarbij de respons Y een lineaire correlatie beschrijft met de verschillende variabelen (X). De algemene formule voor een MLR-model is als volgt:

$$Y = B_0 + B_1 * X_1 + B_2 * X_2 + B_3 * X_3 + \dots + B_n * X_n$$

B<sub>0</sub> is de constante en B<sub>1</sub> tot en met B<sub>n</sub> zijn de regressie coëfficiënten (rc). De waarden voor B kunnen negatief of positief zijn en de grootte geeft de bijdrage aan voor het model. Voor de PMSA003 sensor zijn twee typen MLR-modellen ontwikkeld; voor beide typen geldt dat de temperatuur (T) en relatieve luchtvochtigheid (RV) behoren tot de X-variabelen. Het verschil zit in het feit dat voor het eerste type model de PM-waarde als X-variabele wordt gebruikt. De PM-waarde staat voor *Particulate Matter* en de eenheid is µg/m<sup>3</sup>.

Voor het tweede type model worden de zes deeltjesklassen (PN: *Particle Numbers*) als X-variabelen gebruikt. De eenheid van PN is deeltjes per 0,1 liter. Dat resulteert in drie X-variabelen voor het eerste type

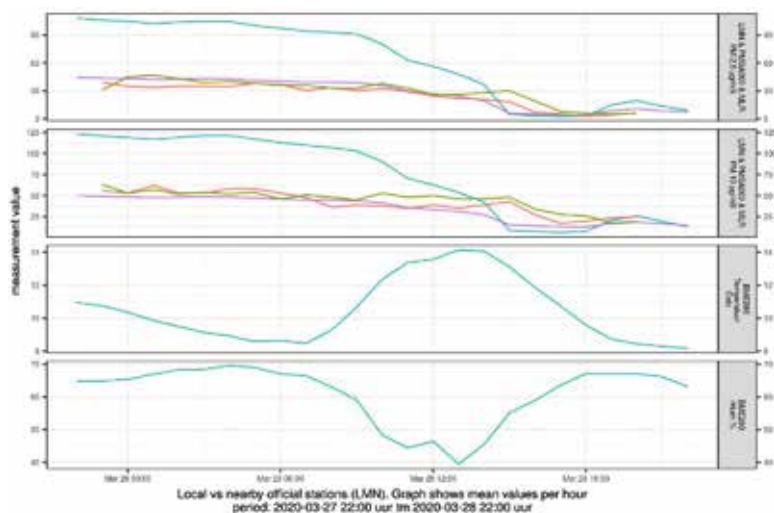
model en acht X-variabelen voor het tweede type model.

De resultaten van de kalibratie en validatie zijn met elkaar vergeleken en op basis van deze evaluatie is de sensor met de beste prestaties gekozen als referentie PMSA003 sensor. Vervolgens is onderzocht in welke mate de modellen van de referentie PMSA003 sensor overdraagbaar zijn naar de andere twee PMSA003 sensoren. Als laatste stap is het beste model voor PM<sub>2,5</sub> en PM<sub>10</sub> toegepast op een onafhankelijke PMSA003 sensor in het burger meetnetwerk van 'Scapeler'. Het is een real time validatie waarbij de modelresultaten direct vergeleken worden met referentiestationen uit het Landelijk Meetnet Luchtkwaliteit (LML).

### De resultaten

De resultaten van de kalibratie en validatie van de verschillende typen modellen zijn in het Visibilis rapport uitvoerig beschreven. De modellen gebaseerd op T, RV en PN laten de beste prestaties zien. De nauwkeurigheid van het model is weergegeven als RMSE, wat staat voor *Root Mean Standard Error of Estimation* (één sigma). De nauwkeurigheid van het PM<sub>2,5</sub>-model en het PM<sub>10</sub>-model is respectievelijk 3 µg/m<sup>3</sup> en 6 µg/m<sup>3</sup>. De nauwkeurigheid van de modellen gebaseerd op T, RV en PM is respectievelijk 4 µg/m<sup>3</sup> en 8 µg/m<sup>3</sup>.

In figuur 3 en 4 is een grafische weergave te zien van twee MLR-modellen voor de →



Figuur 5: Real time validatie PMSA003 sensor

PMSA003 sensor voor respectievelijk  $PM_{2,5}$  en  $PM_{10}$ . Deze modellen zijn gebaseerd op T, RV en PN als X-variabelen. De X-as geeft de referentiewaarde aan volgens de BAM1020 en de Y-as geeft de voorspelde waarde aan volgens het model. De zwarte lijn is de ideale kalibratielijne ( $Y=X$ ) en de rode lijn is de regressielijne door de modelpunten. De modellen bevatten 4836 uurwaarden uit de kalibratieset; dit aantal is exclusief uitbijters.

De reproduceerbaarheid van de BAM1020 monitor voor uurwaarden is  $4 \mu\text{g}/\text{m}^3$ . De nauwkeurigheid van de  $PM_{2,5}$ -modellen is vergelijkbaar met de reproduceerbaarheid van de BAM1020. Een verklaring voor de slechtere modelprestaties voor  $PM_{10}$  kan zijn dat de sensoren minder gevoelig zijn voor deeltjes  $> 2,5 \mu\text{m}$ . Dit wordt ook in andere publicaties beschreven. De sensor is in principe een deeltjesteller en een intern algoritme zet de deeltjesconcentratie om in een massaconcentratie. Hierbij lijkt informatie verloren te gaan die belangrijk is voor het model gebaseerd op T, RV en PN.

Een onderzoek naar het succes van overdraagbaarheid van de modellen van de referentie PMSA003 sensor naar de andere twee PMSA003 sensoren, heeft aangetoond dat de beste resultaten worden bereikt met het model gebaseerd op T, RV en PM. Dit zou men niet verwachten, omdat de modellen gebaseerd op T, RV en PN de beste prestaties laten zien. Een oorzaak zou kunnen zijn dat de scheiding en gevoeligheid van de zes deeltjesklassen niet consistent verloopt voor de verschil-

lende sensoren. Een verbetering van de overdraagbaarheid van de modellen gebaseerd op T, RV en PN kan worden bereikt door normalisatie van de deeltjesklassen. Validatie van de modellen na normalisatie heeft aangetoond dat overeenkomstige prestaties worden bereikt met de oorspronkelijke validatie van de drie PMSA003 sensoren.

#### Real time validatie

De modellen van de referentie PMSA003 sensor, gebaseerd op T, RV en PM als X-variabelen, is geïnstalleerd op een willekeurige PMSA003 sensor die actief is in het 'Scapeler' burgermeetnetwerk. Gekozen is voor een sensor op de meetlocatie in 's-Gravenpolder in Zeeland. Op deze locatie zijn de PMSA003 sensor en de Bosch BME280 sensor geïnstalleerd in dezelfde sensorhut. De resultaten van deze real time validatie zijn te bekijken via [https://openiod.org/SCAPE604/images/R/apri-sensor/aprisensor\\_sgravenpolder-mlr.png](https://openiod.org/SCAPE604/images/R/apri-sensor/aprisensor_sgravenpolder-mlr.png).

De modellen zijn actief op de dataserver en lezen automatisch de data in van de PMSA003 sensor ( $PM_{2,5}$  en  $PM_{10}$ ) en de BME280 sensor (T en RV). De modellen voorspellen de waarden voor  $PM_{2,5}$  en  $PM_{10}$  als gemeten met de BAM1020. Een voorbeeld van de real time validatie is weergegeven in figuur 5. Zowel de originele sensorwaarden (blauwe curves) als de voorspelde waarden met de modellen (lila curves) kunnen vergele-

'Low-cost sensoren  
weer iets meer in  
de richting van het  
serieuze speelveld'

ken worden met de echte waarden gemeten op twee naburige LML-stations. De gekozen LML-stations zijn Ossendrecht en Breda, respectievelijk 32 km en 63 km hemelsbreed ten oosten van 's-Gravenpolder.

De real time validatie is drie maanden na installatie geëvalueerd. De conclusie is dat de uitkomst van de MLR-modellen dichter bij de waarden van de LML-stations liggen.

#### Tot slot

Al met al een technisch verhaal zult u denken. Dat is ook best wel zo, maar het brengt de low-cost sensoren weer een stukje verder in de richting van de kwaliteit die nodig is in het serieuze speelveld. De resultaten van dit project zijn dan ook de aanleiding voor een vervolg onderzoek aan de universiteit van Leiden. Hier is het laatste woord dus nog niet over gesproken.

*Dieter Pientka en André van der Wiel bundelen naast hun werkend bestaan hun krachten onder de naam 'Scapeler'. Dieter Pientka is gespecialiseerd in statistiek en chemometrie. André van der Wiel is gespecialiseerd in informatiesystemen en sensoren.*

[https://www.scapeler.com](https://www.scapeler.com;);  
[info@scapeler.com](mailto:info@scapeler.com) ■

#### Referenties

- <https://www.samenmetenaanluchtkwaliteit.nl/visibilis>