

# Amateurweerstations voor operationele hoge resolutie regenkaarten<sup>1</sup>

LOTTE DE VOS, HIDDE LEIJNSE, AART OVEREEM EN REMKO UIJLENHOET

*Automatische weerstations, verbonden met online platforms, kunnen grote hoeveelheden weerobservaties beschikbaar maken voor operationele toepassingen. Amateurregenmetingen, verkregen van deze platforms, zijn echter gevoelig voor substantiële fouten die typisch zijn voor deze databron. Daarom is een algoritme voor kwaliteitscontrole ontwikkeld, specifiek voor deze regenmetingen. De methode herkent de vier soorten fouten die voorkomen in dergelijke data, is operationeel toepasbaar en maakt daarvoor geen gebruik van aanvullende databronnen. Toepassingen hiervan op een jaar aan data uit Amsterdam en op een maand aan data uit heel Nederland laten zien dat, wanneer slechts 12% van de ruwe data wordt weggelaten, de regenmetingen goed overeenkomen met de referentie: de klimatologische radardataset van het KNMI.*

Artikel

## Inleiding

Voor veel toepassingen zijn hoge resolutie regenmetingen wenselijk. In Nederland meet het KNMI regen met twee radars en twee regenmeternetwerken. De grote reikwijdte van weerradar wordt gecombineerd met de hoge nauwkeurigheid op puntschaal van regenmeters in diverse gecorrigeerde regenproducten. Voor onder meer hydrologische toepassingen is het namelijk noodzakelijk dat radarproducten worden verbeterd door correctie met regenmeterdata. Met de 32 automatische regenmeters van het KNMI kunnen realtimeradarbeelden zinvol worden gecorrigeerd en met dagsommen van het KNMI-handregenmeternetwerk kunnen radarproducten fors worden verbeterd, maar alleen achteraf. Om op operationele basis kwalitatief goede hoge resolutie regenkaarten te kunnen maken zijn grote aantallen in-situsensoren nodig die regen in real time en in kleine tijdstappen meten en rapporteren. Zo'n netwerk zou grote investeringen vragen (in tijd en geld) om te installeren en onderhouden.

Echter, een groot aantal weeramateurs bezit tegenwoordig weerstations met regenmeters waarvan de data automatisch verzameld worden op online weerplatforms. De infrastructuur is dus al aanwezig om grote aantallen regenmetingen te crowdsourcen van deze platforms. De potentie van deze databron is groot, ook ten opzichte van andere regenmeettechnieken zoals radar en commerciële radiostraalverbindingen (De Vos e.a., 2017; 2018). De verwachting is dat de kwaliteit van deze metingen relatief laag is doordat de sensoren veel goedkoper

<sup>1</sup> Dit artikel is een verkorte bewerking van de Engelstalige publicatie van De Vos e.a. (2019)

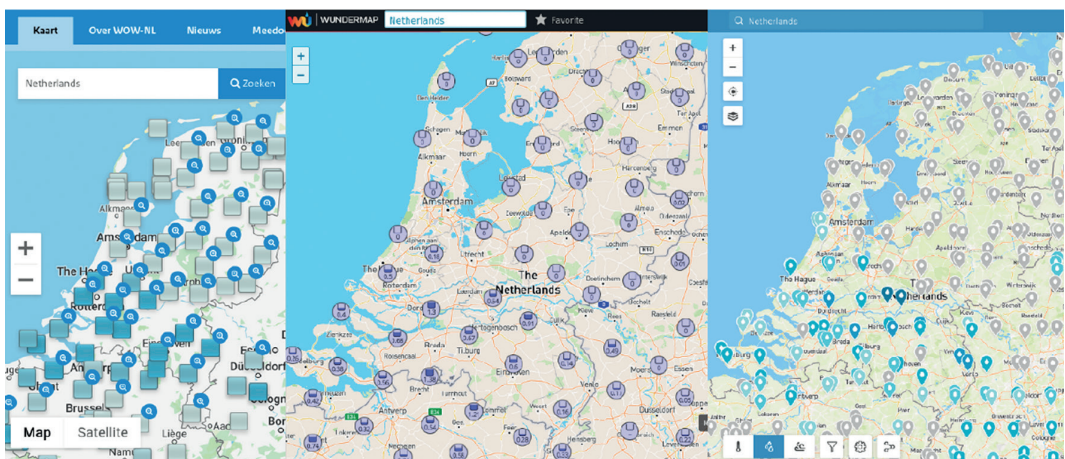
zijn dan de regenmeters die gebruikt worden door het KNMI, en door gebrek aan expertise of geschikte opstellingslocatie geïnstalleerd worden op een manier die de meetnauwkeurigheid negatief beïnvloedt. Vergelijkingen van metingen van diverse weersvariabelen door deze weerstations met een referentie van hoge kwaliteit in Amsterdam lieten, ondanks goede resultaten ten opzichte van andere opportunistische meettechnieken (smartphones-observaties en commerciële radiostraalverbindingen), aanzienlijke meetfouten zien in hun ruwe vorm (De Vos e.a., 2020).

Daarom stellen wij een methode voor kwaliteitscontrole (in het Engels Quality Control of QC) voor die om kan gaan met de typische fouten in regenmetingen van amateurweerstations, zodat uit deze databron van hoge *kwantiteit* ook een regenproduct van hoge *kwaliteit* gemaakt kan worden. Deze QC-methode is in real time toepasbaar, waardoor dit soort regenmetingen operationeel gebruikt kan worden. Omdat in Nederland het aantal regenmetingen uit amateurweerstations veel hoger is (enkele duizenden) dan het aantal automatische regenmeters van het KNMI (32), biedt dit bijvoorbeeld perspectief voor een verdere verbetering van de reallimeradarbeelden.

## Online weerstationplatforms

Er zijn verschillende online platforms waar weeramateurs hun weerstations mee kunnen verbinden (afbeelding 1). Populaire voorbeelden zijn WOW-NL ([www.wow.knmi.nl](http://www.wow.knmi.nl)), Wundermap ([www.wunderground.com/wundermap](http://www.wunderground.com/wundermap)) en Netatmo Weathermap ([www.weathermap.netatmo.com](http://www.weathermap.netatmo.com)). Op kaarten worden de meest recente observaties van verschillende weersvariabelen weergegeven als symbolen. Elk van deze platforms heeft zijn voor- en nadelen, maar we richten ons in deze studie op de Netatmo Weathermap, omdat alle stations die daarmee verbonden zijn van hetzelfde type zijn, en omdat de dichtheid van dit netwerk het hoogst is in Nederland.

De Netatmo regenmeter bestaat uit een kantelbakje onder een opvangtrechter met een diameter van 13 cm die omslaat wanneer er 0.101 mm regen opgevan-



Afbeelding 1 Schermafdrucken van de weerkaart van respectievelijk de WOW-NL, Wundermap en Netatmo Weathermap platforms. Door de kaarten op de webpagina's in te zoomen worden meer meetlocaties zichtbaar.

gen is. Het aantal kantelingen wordt elke circa 5 minuten draadloos doorgegeven aan de binnenmodule, vanwaar de informatie via het wifinetwerk van de eigenaar naar het Netatmo platform wordt gestuurd. Doordat de momenten waarop de weerstations hun data naar het platform sturen niet gesynchroniseerd zijn en kunnen afwijken (meestal enkele seconden, soms langer) heeft de meetreeks onregelmatige tijdsintervallen. Metingen kunnen door de eigenaar van het weerstation bijgehouden worden op een smartphone-app of online dashboard. Iemand kan zijn/haar regensensor desgewenst kalibreren door er een vaste hoeveelheid water doorheen te gieten. Het volume van de kantelbak kan daarom een andere waarde krijgen dan de nominale waarde van 0.101 mm.

### Typische fouten in amateurregenmetingen

We onderscheiden vier soorten fouten in deze amateurregenmetingen: een veel voorkomende fout is een systematische onder- of overschatting in de metingen. Een aannemelijke oorzaak hiervoor kan bijvoorbeeld zijn dat de regenmeter beschut staat opgesteld en daardoor minder regen meet dan daadwerkelijk gevallen is, of dat het volume van het kantelbakje dat is ingesteld niet correspondeert met het daadwerkelijke volume. Dit kan komen door fabricagefouten, of doordat bij kalibratie het water sneller door de sensor stroomde dan het kantelmechanisme kon verwerken, waardoor een deel langs het mechanisme stroomde, en de nieuw gekalibreerde waarde te hoog werd.

Een andere veel voorkomende fout is een lange periode van nulmetingen doordat het kantelmechanisme geblokkeerd is. Dit kan worden veroorzaakt door vaste neerslag of verstoppingen in de sensor (bijvoorbeeld een spinnenweb in de regenmeter dat het kantelbakje tegenhoudt of bladeren in de trechter die het water niet doorlaten naar het kantelbakje). De sensor blijft dan actief observaties doorgeven van nul kantelingen, ook als het regent.

Plotselinge hoge waarden die niet aan regen zijn gerelateerd zijn zeldzamer, maar kunnen tot grote fouten leiden. Ze kunnen worden veroorzaakt door het bewegen van de regenmeter, bijvoorbeeld tijdens verplaatsen of schoonmaken, waarbij het kantelbakje kan kantelen en er ten onrechte regen wordt geregistreerd. Ook kan het zijn dat een hoeveelheid sneeuw of hagel door de sensor is opgevangen, die bij dooi geregistreerd wordt als een intense regenpiek.

In de laatste foutencategorie die we onderscheiden zitten situaties waarin de regendynamiek die door de sensor gemeten wordt niet overeenkomt met de dynamiek die je op die locatie verwacht. Het kan zijn dat de gerapporteerde stationslocatie niet (meer) klopt, bijvoorbeeld na een verhuizing. Ook zijn er zeldzame tijdreeksen gevonden waar, door een fout op het platform, regen werd geregistreerd in de vorm van dagaccumulaties in plaats van accumulaties over vijf minuten.

### Kwaliteitscontrole

Doordat al deze fouten geregeld voorkomen in regenobservaties van weerstations van weeramateurs kan de ruwe dataset niet direct gebruikt worden voor hydrologische toepassingen. We hebben een automatische methode voor kwa-

liteitscontrole ontwikkeld die gebruik maakt van het gegeven dat de ruimtelijke dichtheid van deze weerstations hoog is. De aanname is dat nabijgelegen stations een vergelijkbare regendynamiek meten, en dat afwijkende metingen herkend kunnen worden door te vergelijken met hun burens.

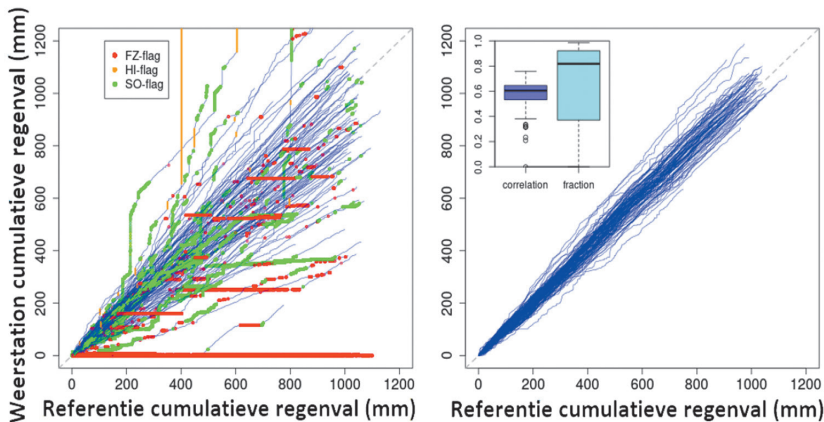
Metingen van een station worden vergeleken met de beschikbare metingen in een straal van 10 km. Als een station langer dan een bepaalde duur actief nulmetingen registreert, terwijl de mediaan van de buurmetingen niet nul is, worden deze metingen geclassificeerd als foutieve nulmetingen. Metingen worden genegeerd totdat het station weer een niet-nulmeting doorgeeft. Als een station een veel hogere hoeveelheid registreert dan de mediaan van de burens wordt dit als een foutieve hoge meting herkend. De bovengrens is dynamisch, zodat een hoge meting meer verdacht is bij droog weer dan wanneer de burens aanzienlijke regen meten. Hiermee worden niet alle regenpieken zonder meer als "fout" betiteld. Ook moet er minimaal een bepaald aantal naburige metingen zijn voordat een meting als zodanig wordt geclassificeerd. Een uitzonderings-situatie waarin de burenvergelijking niet volstaat is als opgevangen sneeuw in de opvangbakjes smelt en in meerdere stations gelijktijdig leidt tot regenobservaties veroorzaakt door het smeltwater. Hoewel dit zelden voorkomt is dit niet onmogelijk in het Nederlandse klimaat.

Om te herkennen of een station een afwijkende neerslagdynamiek registreert, wordt er vergeleken met metingen van elk naburig station over een periode van minimaal twee weken met minimaal honderd intervallen van vijf minuten van regenachtig weer. Als de correlatie met de meeste buurstations lager is dan 0.15 dan wordt het station als afwijkend geclassificeerd. Omdat de vergelijkingsperiode lang is, duurt het even voor een station met deze fout uit de selectie wordt gehaald. Als de correlatie wel hoog is, maar er een systematisch verschil is in de regenhoeveelheden van een station ten opzichte van de meeste burens, worden de metingen wel opgenomen in de dataset, maar vermenigvuldigd met een correctiefactor. Omdat amateurweerstations de neiging hebben om de werkelijk gevallen hoeveelheid regen te onderschatten, kan ervoor gekozen worden om al te beginnen met een algehele correctiefactor op het netwerk. De waarde van die factor kan bepaald worden in een offline vergelijkingsstudie van het netwerk met een betrouwbare regenreferentie (zoals weerradar gecorrigeerd met officiële regenmeters).

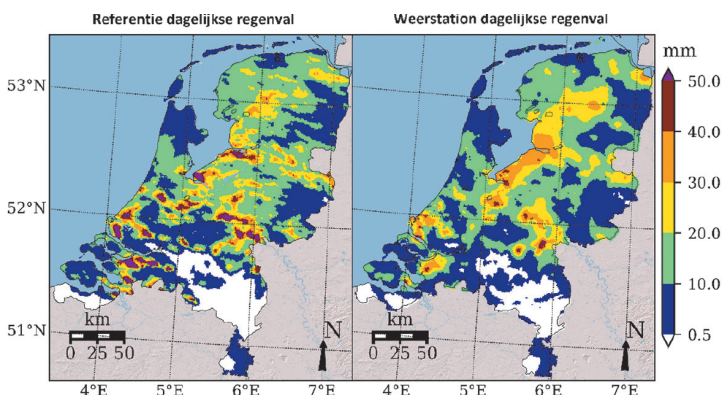
De ontwikkelde methode bestaat uit vier modules die de vier soorten fouten in regenmetingen van amateurweerstations met onregelmatige meetintervallen automatisch herkennen, losse intervallen in real time kunnen classificeren en daarbij geen referentiemetingen nodig hebben. In totaal zijn er in dit algoritme voor automatische kwaliteitscontrole 11 instelbare parameters die het gedrag van de methode beïnvloeden. Door deze parameters te variëren kan de nadruk gelegd worden op een strengere of minder strenge filtering van de ruwe dataset. Door geclassificeerde intervallen uit de dataset te halen blijft er een dataset van betrouwbare metingen over. Een gedetailleerde uitleg van de modules en de parameters is gegeven in De Vos e.a. (2019).

## Resultaten in Nederland

De methode voor kwaliteitscontrole is ontworpen en gekalibreerd (i/e. de 11 parameterwaarden zijn bepaald) op basis van maand 2 t/m 13 van een dataset van 25 maanden aan regenmetingen in Amsterdam. Ter validatie is de methode vervolgens toegepast op de laatste 12 maanden van de dataset, waarbij zowel voor kalibratie als voor validatie de voorgaande maand als opstarttijd werd gebruikt. De linkerfiguur in afbeelding 2 laat zien welke intervallen door de kwaliteitscontrole als fout geclassificeerd worden. De rode, oranje en groene kleuren bij respectievelijk de horizontale, verticale en niet-diagonale lijnstukken laten zien dat de kwaliteitscontrole in staat is om zonder referentiedata de foute observaties te herkennen. De gefilterde dataset komt zeer goed overeen met de



Afbeelding 2 Cumulatieve regenval volgens de amateurweerstations in Amsterdam tijdens het validatiejaar, ten opzichte van de KNMI referentie. Links zijn alle metingen weergegeven, waarbij als fout geclassificeerde intervallen gekleurd zijn (rood voor foutieve nulmetingen (FZ), oranje voor foutieve hoge waarden (HI) en groen voor een afwijkende dynamiek (SO)). Rechts zijn deze gekleurde intervallen uit de dataset weggehaald en is de dynamische correctiefactor toegepast. De boxplots laten de spreiding van de correlaties en van de fracties metingen die overblijven zien voor alle weerstations.



Afbeelding 3 Regenaccumulaties tussen 29 mei 2018 08:00 UTC en 30 mei 2018 08:00 UTC volgens de referentie en uit amateurweerstations na QC die minimaal 95% databeschikbaarheid gedurende de dag hadden.

klimatologische radardataset van het KNMI (die met één tot twee maanden vertraging beschikbaar komt). Dit blijkt uit de goede overeenkomst tussen de amateurmetingen en de referentie van de cumulatieve regenmetingen in de rechterfiguur in afbeelding 2. Hierbij is slechts 12% van de ruwe data weggehaald.

De ontwikkelde methode voor automatische kwaliteitscontrole is ook toegepast op een dataset van een maand (mei 2018) voor heel Nederland. Afbeelding 3 laat zien dat de resulterende regenkaarten goed overeenkomen met de regenkaart gebaseerd op de KNMI-referentie. Dit suggereert dat deze databron potentie heeft om operationeel gebruikt te worden, bijvoorbeeld voor de real-timecorrectie van radargegevens. Hierbij moet wel worden opgemerkt dat de dichtheid van het netwerk in de randstad (veel) hoger is dan daarbuiten. Dat zal, vanwege de vergelijking met naburige stations in de methode, tot een grotere onzekerheid in gebieden met een lagere stationsdichtheid leiden.

## Conclusies

De resultaten laten zien dat de ontwikkelde methode voor automatische kwaliteitscontrole een geschikte manier is om de nuttige data uit de regenmetingen van amateurweerstations te extraheren. Hierbij blijft een substantieel deel van de data beschikbaar (88%) van een toch al aanzienlijke dataset. De methode is in theorie operationeel toepasbaar en er zijn geen aanvullende metingen nodig doordat er gebruik wordt gemaakt van het principe dat er genoeg naburige metingen zijn die de regendynamiek correct kunnen waarnemen. Deze aanname is alleen geldig in het geval van een relatief dicht netwerk. Door parameters van de kwaliteitscontrole aan te passen kan bijvoorbeeld de afstand waarop stations nog steeds als burens worden beschouwd worden vergroot, waarmee de methode ook toepasbaar wordt op minder dichte netwerken, bijvoorbeeld het Netatmo netwerk in andere gebieden of netwerken van andere typen (amateur-) weerstations. In minder dichte sensornetwerken moet mogelijk met grotere tijdstappen gewerkt worden, zodat dezelfde regendynamiek meetbaar is over het gebied waarvoor de vergelijking plaatsvindt.

Andere aanpassingen zijn mogelijk nodig in gebieden waar de regendynamiek niet consistent is in alle richtingen, bijvoorbeeld in bergachtige gebieden. De burensselectie op basis van afstand zou in die situatie mogelijk aangevuld kunnen worden met een selectie op basis van hoogte, zodat stations alleen vergeleken worden met burens die zich ongeveer op dezelfde hoogte bevinden. Een andere mogelijke uitdaging is het toegang krijgen tot de data, die weliswaar gegenereerd worden door de bezitters van de weerstations, maar daarna eigendom worden van de bedrijven achter de platforms.

Los van de aanpassingen die nodig kunnen zijn voor specifieke toepassingen laat deze studie duidelijk de mogelijkheid van regenmetingen uit amateurweerstations voor operationele regenmonitoring zien. Ons vervolgonderzoek richt zich nu op het gebruik van deze databron, na kwaliteitscontrole, voor correctie van operationele weerradardata, zodat duidelijk wordt tot welke kwaliteitsverbetering van real-time radarbeelden dit zou kunnen leiden. Daarmee zouden ook de hydrologische wetenschap en het operationele waterbeheer in Nederland gediend zijn.



## Softwarebeschikbaarheid en overige informatie:

De ontwikkelde methode voor automatische kwaliteitscontrole van amateurregenmetingen is voor iedereen beschikbaar gemaakt. Zie voor de computercode en gedetailleerde documentatie: [www.github.com/LottedeVos/PWSQC](http://www.github.com/LottedeVos/PWSQC). De gebruikte datasets van amateurweerstations zijn ook online beschikbaar (<https://data.4tu.nl/repository/uuid:6e6a9788-49fc-4635-a43d-a2fa164d37ec>), net zoals het referentie radarproduct van het KNMI ("Radar precipitation climatology" op <http://climate4impact.eu>).

## Dankwoord

De auteurs bedanken de eigenaars van amateurweerstations voor het maken en delen van regenmetingen, en Netatmo voor het beschikbaar maken van de regenobservatie-datasets via de Netatmo weathermap.

## Literatuur

- De Vos, L. W., Leijnse, H., Overeem, A., & Uijlenhoet, R.** (2017) The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam; in: *Hydrology and Earth System Sciences*, vol 21(2), pag 765-777.
- De Vos, L. W., Raupach, T. H., Leijnse, H., Overeem, A., Berne, A., & Uijlenhoet, R.** (2018) High-resolution simulation study exploring the potential of radars, crowdsourced personal weather stations, and commercial microwave links to monitor small-scale urban rainfall; in: *Water Resources Research*, vol 54(12), pag 10.293-10.312.
- De Vos, L. W., Leijnse, H., Overeem, A., & Uijlenhoet, R.** (2019) Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring; in: *Geophysical Research Letters*, vol 46(15), pag 8820-8829.
- De Vos, L. W., Droste, A. M., Zander, M. J., Overeem, A., Leijnse, H., Heusinkveld, B. G., Steeneveld, G. J. & Uijlenhoet, R.** (2020) Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area; in: *Bulletin of the American Meteorological Society* (in press).

## Summary Amateur weather stations for operational high-resolution rainfall maps

*A large number of weather stations are linked to online platforms, which enables crowdsourcing of a large number of rainfall observations in an operational way. These measurements are sensitive to four types of errors. A methodology is proposed to identify these errors based on the crowdsourced rainfall observations alone, with the use of neighbor checks and predetermined parameters. Applying this quality control methodology on a one-year urban dataset in Amsterdam shows that by excluding 12% of the raw data and applying a dynamic bias correction, a good match is found with the reference rainfall product from KNMI, which is not available in real time. This shows that crowdsourced rainfall observations could be useful for operational purposes after this quality control. The same method showed good results when applied on a national dataset. Future work is focused on merging this data source with operational weather radar data for a high quality rainfall product.*

## Auteur

LOTTE DE VOS

Wageningen University & Research  
lotte.devos@wur.nl

HIDDE LEIJNSE

Koninklijk Nederlands Meteorologisch Instituut  
hidde.leijnse@knmi.nl

AART OVEREEM

Koninklijk Nederlands Meteorologisch Instituut  
Wageningen University & Research  
aart.overeem@knmi.nl

REMKO UIJLENHOET

Wageningen University & Research  
remko.uijlenhoet@wur.nl