

Onderzoek Kwaliteitscontrole Neerslagmetingen

Samenvatting onderzoek kwaliteitscontrole voor regenmeters door Lotte de Vos

In aanvulling op “Crowdsourced personal weather stations show great potential for operational rainfall monitoring” door L.W. de Vos, H. Leijnse, A. Overeem & R. Uijlenhoet, ingediend bij Geophysical Research Letters (GRL).

Achtergrond

Om regenmeterdata te kunnen gebruiken in neerslagproducten is het zeer belangrijk dat grove fouten er in real-time uitgehaald kunnen worden. Binnenkort zullen regenmetergegevens van waterschappen en gemeenten in real-time voor het KNMI beschikbaar komen, en de potentie van het inmengen van deze regenmeters in radarneerslagproducten is groot, mits deze data voldoen aan bovengenoemde eisen. Daarom is er een kwaliteitscontrole (QC)-methodiek ontwikkeld. Omdat het KNMI bij het begin van dit onderzoek nog niet klaar was om regenmetergegevens van waterschappen en gemeentes in te winnen, is ervoor gekozen om het onderzoek uit te voeren met andere niet-officiële WMO regenmeters: de regenmeters van het amateurweerstation Netatmo. Deze regenmeters hebben de overeenkomst met regenmeters van waterschappen en gemeentes dat de data in real-time met hoge temporele resolutie ingewonnen kunnen worden, en dat de afstand tussen regenmeters beperkt is. De fouten die voorkomen in deze amateur-regenmeter zullen waarschijnlijk groter zijn en vaker voorkomen dan bij andere regenmeters. We nemen aan dat het gebruik van deze Netatmo dataset ook bruikbare methodieken oplevert voor regenmeters van waterschappen en gemeenten. Dit zal uiteraard getest worden zodra er een langere dataset is opgebouwd van deze regenmeters.

Automatische persoonlijke weerstations (PWS) zijn populaire meetinstrumenten die particulieren kunnen aanschaffen om het weer in hun omgeving te meten. In veel gevallen worden de metingen in real-time gedeeld en gevisualiseerd op online platforms. Netatmo is een populair platform waaraan alle Netatmo weerstations gelinkt zijn. Deze weerstations meten druk, temperatuur, relatieve luchtvochtigheid, en in het geval dat extra modules zijn aangesloten, ook regen en/of wind in intervallen van ~5 min. Het Netatmo meetnetwerk is veel dichter dan de regenmeetnetwerken van het KNMI, met gemiddeld 1 station per 11 km² over heel Nederland, waarbij de netwerk dichtheid in steden veel hoger is dan in ruraal gebied.

Deze databron levert in theorie een grote hoeveelheid observaties, maar zijn gevoelig voor meetfouten. Voor crowdsourced PWS regenmetingen zijn er verschillende oorzaken voor fouten, e.g. wind effecten, foute kalibratie, afgeschermd locatie, tipping mechanisme blokkade door scheefwaaier of obstakels in de regenmeter, etc. Deze uiten zich op 4 verschillende manieren:

- Lange periodes van nulmetingen: Faulty zero (FZ)
- Plotselinge hoge metingen: High influx (HI)
- Afwijkende metingen ten opzichten van naburige stations: Station outlier (SO)
- Systematische onder- of overschatting: Bias

Om operationeel gebruik te maken van dit type regendata, namelijk gecrowdsourced van weer platforms met variabele update intervallen en zonder beschikbare meta-data, is een real-time toepasbaar kwaliteitscontrole (QC) nodig die in staat is om deze fouten te herkennen.

Onderzoek

Dit onderzoek maakt gebruik van twee PWS datasets van het Netatmo platform.

- Regenmetingen van alle operationele PWSen in het Amsterdam stedelijke gebied (~575 km²) over een periode van 2 jaar.
- Regenmetingen van alle operationele PWSen in Nederland in mei 2018.

We hebben een methode ontworpen die met behulp van 'neighbor checks' de 4 typische fouten voor deze databron herkent. Hierbij wordt gebruik gemaakt van de hoge dichtheid aan metingen en de aanname dat de meerderheid van naburige stations in staat zijn om een regenbui correct waar te nemen. Wanneer een station afwijkende metingen heeft ten opzichte van naburige stations voor een langere periode dan vooraf ingestelde drempelwaarden, dan worden deze metingen op een gegeven moment een FZ, HI of SO vlag = 1 toegekend. Als er te weinig informatie beschikbaar is om te bepalen of een meting afwijkt wordt een vlag van -1 toegekend. Afhankelijk van de striktheid van de QC kan gekozen worden om deze -1 vlag intervallen mee te nemen of uit te sluiten.

Om te compenseren dat de PWS regen typisch onderschatten is er een algehele bias correctie factor voor de hele dataset bepaald. Dit is de enige stap waarvoor externe regenmetingen nodig zijn. De default bias correctie factor wordt op alle stations toegepast, en indien nodig (op basis van neighbor checks en ingestelde drempelwaarden) worden deze per station in de tijd aangepast. Na de inloop tijd heeft elk station een individuele bias correctie die kan veranderen wanneer de station metingen daartoe aanleiding geven. Het is aan te raden om deze algehele bias correctie factor periodiek te herhalen aangezien het netwerk verandert, door systematische veranderingen in PWS metingen over levensduur, introductie van nieuwe stations en stations die niet meer bijdragen aan de dataset.

Het QC design en de keuzes voor de 11 filter parameters zijn bepaald op basis van het eerste jaar van de Amsterdam dataset. Vervolgens is het filter getest op het tweede jaar en de nationale dataset. De metingen worden voor en na toepassen van QC gevalideerd met het radar-regenmeter product van het KNMI, waarbij PWSen vergeleken worden met de waarde in de referentie in het desbetreffende pixel.

Resultaten

De QC is succesvol in het filteren van de typische fouten in de PWS metingen (FZ, HI, SO en bias). Dit blijkt uit de betere match met de radar-regenmeter referentie als alle gevlagde intervallen zijn weggelaten uit de dataset. Alle individuele filters laten verbetering zien bij validatie, en de gehele QC leidt tot een significante verbetering ten opzichte van de ruwe dataset voor uurlijkse, en 5-min tijdreeksen. In het tweede jaar van de Amsterdam 5-min dataset zorgt de QC voor een verbetering in Pearson correlatie coëfficiënt van 0.07 naar 0.58, de coëfficiënt van variatie van 53.24 naar 7.19, en een

relatieve bias van -11.1% naar +2.6%, door alleen de on-gevlagde intervallen mee te nemen (dit beslaat 88% van de totale dataset).

Het filter is toegepast op de dataset van alle stations in Nederland. Hiermee kunnen geïnterpoleerde kaarten gemaakt kunnen worden van dagelijkse regenval volgens PWS na QC, die qua patroon en hoeveelheid goed overeenkomen met de kaarten gemaakt met het radar-regenmeter product.

Vooruitzicht

Het onderzoek laat zien dat PWS metingen een veelbelovende bron van regenmetingen kunnen zijn, omdat ze een grote hoeveelheid observaties genereren, die na QC accuraat regen kunnen beschrijven. Dit is met name veelbelovend in steden, waar het aantal PWS het hoogst ligt. Het is aangetoond dat de methode van het QC filter goed in staat is om foutieve observaties te herkennen zonder dat daar externe metingen of meta-data voor nodig is. De principes zijn ook toepasbaar op andere regenmeternetwerken dan die van Netatmo. De belangrijkste volgende stap wordt het ontwerpen van operationele software die de QC in real-time kan toepassen op een datastroom van een regenmeter netwerk (zowel regenmeters van waterschappen en gemeentes als amateur regenmeters, zoals WOW of Netatmo komen hiervoor in aanmerking). Deze metingen kunnen dan gebruikt worden om een de radar-regenmeter producten te verbeteren.