

Neerslag-kansverwachting voor het waterbeheer

JANTINE BOKHORST, WAGENINGEN UNIVERSITEIT / HYDROLOGIC

ARNOLD LOBBRECHT, HYDROLOGIC / UNESCO-IHE

In de meteorologie wordt in toenemende mate gebruik gemaakt van kansverwachtingen. Vroeger was de weersverwachting gebaseerd op één modeluitkomst. Deze deterministische verwachting zegt echter niets over de zekerheid van die verwachting. Tegenwoordig wordt de deterministische verwachting dan ook vaak gecombineerd met een kansverwachting, die de mogelijke variatie in het verwachte weerbeeld wel tot uiting brengt. Ook voor het operationele waterbeheer biedt deze ontwikkeling perspectief. Op sommige gebeurtenissen kan men beter bedacht zijn, al is de kans dat ze optreden nog zo klein. In dit artikel worden enkele aspecten van het gebruik en de betrouwbaarheid van neerslag-kansverwachtingen in het waterbeheer verkend.

Atmosferische modellen worden voortdurend verbeterd. Er komt steeds meer inzicht in atmosferische processen en de rekencapaciteit van supercomputers neemt toe, waardoor de atmosfeer steeds beter numeriek beschreven kan worden. Toch blijft de voorspelbaarheid van het weer beperkt. Ten dele wordt dit veroorzaakt door het feit dat numerieke modellen altijd een benadering zullen blijven van de exacte fysische processen in de atmosfeer. Een andere belangrijke oorzaak ligt in de beginvoorwaarden van het model. De initiële condities worden vastgesteld op basis van waarnemingen. Deze waarnemingen zijn echter beperkt in aantal en nauwkeurigheid. Aangezien de atmosfeer een chaotisch systeem is,

kunnen kleine foutjes in de beginvoorwaarden uitgroeien en van grote invloed zijn op de betrouwbaarheid van een verwachting. Deze betrouwbaarheid is echter niet af te lezen aan de uitkomsten van een deterministisch model.

Het 'European Centre for Medium-Range Weather Forecasts' (ECMWF) in Reading, Engeland, heeft in 1992 een hulpmiddel ontwikkeld om een schatting te kunnen maken van de invloed van de genoemde onzekerheden op de uitkomsten. Naast een wereldomvattend deterministisch model met een gridafstand van 40 x 40 km wordt ook een model gebruikt met een gridafstand van 80 x 80 km. Dit grovere model wordt 51 keer doorgerekend. De 51

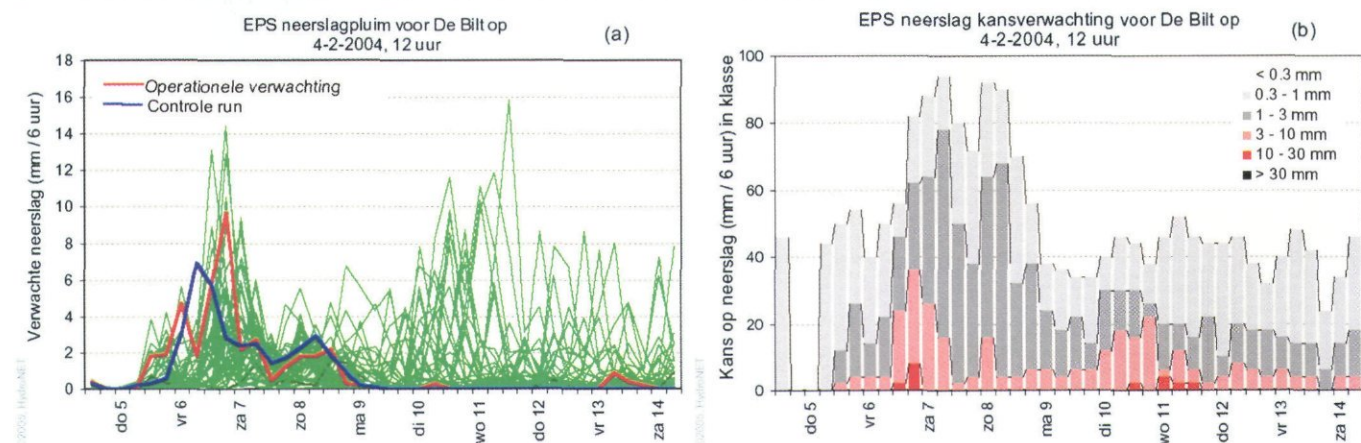
runs (of ensembleleden) van dit 'Ensemble Prediction System' (EPS) hebben onderling licht afwijkende beginvoorwaarden. Eén run heeft dezelfde voorwaarden als het deterministische model: de controlerun. De initiële condities voor de overige 50 runs worden opgesteld door aan de onverstoorte beginvoorwaarden kleine verstoringen toe te voegen die binnen de onzekerheid liggen. De verstoringen worden zo gekozen dat de 50 nieuwe voorwaarden zo veel mogelijk een uniforme ruimtelijke spreiding en een gelijke waarschijnlijkheid hebben. Elke twaalf uur wordt het EPS-model doorgerekend tot tien dagen vooruit.

De modelresultaten kunnen op verschillende manieren worden gepresenteerd. Afbeelding 1a geeft een neerslagpluim weer: per zes uur is van alle runs de verwachte neerslag uitgezet tegen de tijd. Ook de uitkomst van het deterministische model (operationele verwachting) is weergegeven. Aan het verschil tussen de controle run en de operationele verwachting is af te lezen wat de gevolgen zijn van het grovere grid voor de modeluitkomsten. De modeluitkomsten kunnen ook gegroepeerd worden in klassen. Het percentage van de leden dat binnen een bepaalde klasse valt, geeft de kans dat de werkelijke neerslag in die klasse valt (afbeelding 1b). De pluim kan ook nog verder verwerkt worden tot een droogte- of wateroverlastverwachting, door bijvoorbeeld de ensembleleden te gebruiken als invoer voor een waterbalansmodel (afbeelding 2a) of een gedetailleerd model van een watersysteem (afbeelding 2b)³.

Statistische verificatie

In theorie zou elk ensemblelid even waarschijnlijk moeten zijn en zou de spreiding van het ensemble een goede maat moeten zijn voor de betrouwbaarheid van de deterministische verwachting. Maar is dat ook zo in de praktijk? Een algemene, veelgebruikte methode voor de

Afb. 1: Presentaties van EPS-neerslagverwachting voor De Bilt (4 februari 2004): (a) modeluitkomsten per zes uur, (b) kansverdeling per zes uur (KNMI, 2004; HydroNET, 2004).



verificatie van EPS-ensembles is het Talagrand-diagram²⁾. De 51 ensembleleden delen de ruimte van mogelijke neerslaghoeveelheden op in 52 intervallen. Het onderste interval loopt van 0 tot het onderste ensemblelid, het bovenste interval loopt van het hoogste ensemblelid tot oneindig. In theorie heeft elk interval circa twee procent kans uiteindelijk de werkelijke neerslag te bevatten. Door de intervallen van onder naar boven te nummeren van 1 tot 52 en te tellen hoe vaak in elk interval de werkelijke neerslag voorkwam, kan gecontroleerd worden of deze veronderstelling correct was. In afbeelding 3 zijn de Talagrand-diagrammen weergegeven van de neerslagverwachting voor De Bilt voor zes uur, drie dagen en tien dagen vooruit. Drie dagen vooruit wil zeggen dat de verwachte zes-uursom voor 72 uur vooruit is vergeleken met de gemeten zes-uursom voor dat tijdstip. Gekeken is naar de EPS-ensembles van 12:00 uur 's middags over de periode september 2002 t/m december 2004.

Het Talagrand-diagram voor de eerste zes uur toont een grote overbevolking van de twee buitenste intervallen: de pluim is veel te smal. Maar liefst 43,5 procent van de waarnemingen ligt buiten de pluim, terwijl dat in theorie maar circa vier procent had mogen zijn. Met een toenemende verwachtingstermijn vlak de grafiek af. Voor alle verwachtingstermijnen

zijn de onderste intervallen te vol; dan volgt een te leeg middengedeelte en het bovenste interval is weer iets te vol. De overbevolking van de onderste intervallen ontstaat doordat het vaak voorkomt dat veel ensembleleden kleine hoeveelheden neerslag verwachten, terwijl er in werkelijkheid niets valt³⁾. Dit verschijnsel zou ten dele veroorzaakt kunnen worden door het model zelf. Het is echter ook te wijten aan de gebruikte verificatiemethode.

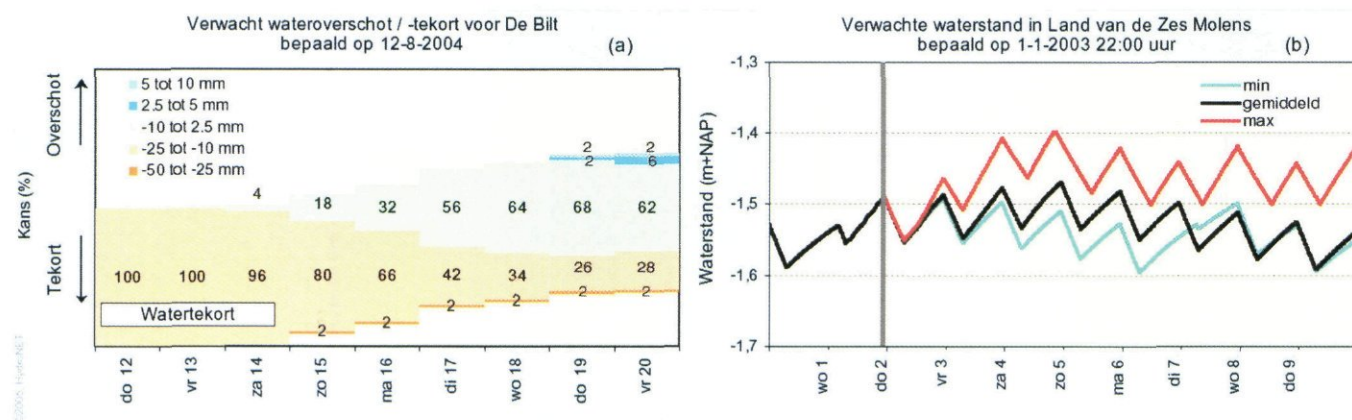
De gridbox-gemiddelde EPS-verwachtingen zijn namelijk vergeleken met puntwaarnemingen. Eigenlijk zou deze verificatie moeten plaatsvinden met gebiedsgemiddelde waarden. Het is te verwachten dat de verdeling over de intervallen dan beter is. De kans dat ergens in het gebied iets valt, is groter dan de kans dat het precies boven de regenmeter in De Bilt regent. De onderste intervallen zullen daardoor minder overbevolkt raken. Ook de plaatselijke uitschieters die nu het bovenste interval te vol maken, zullen worden uitgemiddeld over het hele gebied. Helaas is het erg moeilijk gebiedsgemiddelde zes-uurssommen te genereren die voldoende nauwkeurig zijn voor deze statistische verificatie. Het aantal grondstations met betrouwbare uurmetingen is beperkt en als gevolg hiervan kunnen drie-uursradargegevens nog niet optimaal gecorrigeerd worden⁴⁾.

Verificatie EPS voor het waterbeheer

Wat zegt nu het gegeven dat EPS-ensembles statistisch al dan niet correct zijn? Het zegt alleen iets over de (gemiddelde) juistheid van de kansklassen. De betrouwbaarheid van de kansklassen is een belangrijk uitgangspunt voor een kansverwachting, maar zegt niet alles. Het zegt bijvoorbeeld niets over de presaties van EPS wat de weersverwachting voor morgen betreft.

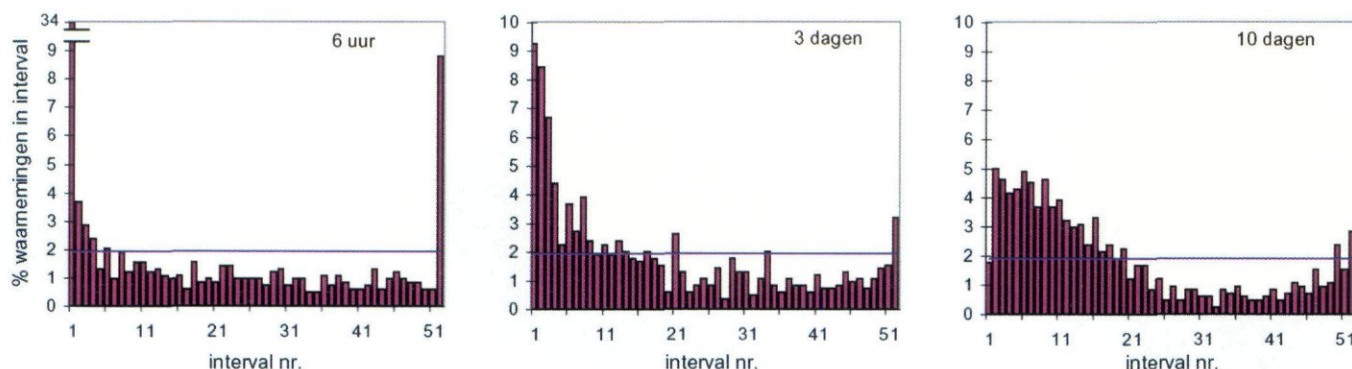
In afbeelding 4a is de gemeten driedaagse neerslagsom weergegeven voor De Bilt, samen met het 50%-percentiel van de verwachte som voor de eerste drie dagen. De vorm van beide curven is globaal gelijk, wat er op wijst dat EPS in ieder geval informatie levert over het weer van morgen. Tussen verwachting en werkelijkheid komen echter afwijkingen voor. Dat is niet meer dan logisch: het 50%-percentiel zegt alleen dat er eventueel kans is dat er meer of minder neerslag valt dan de aangegeven lijn.

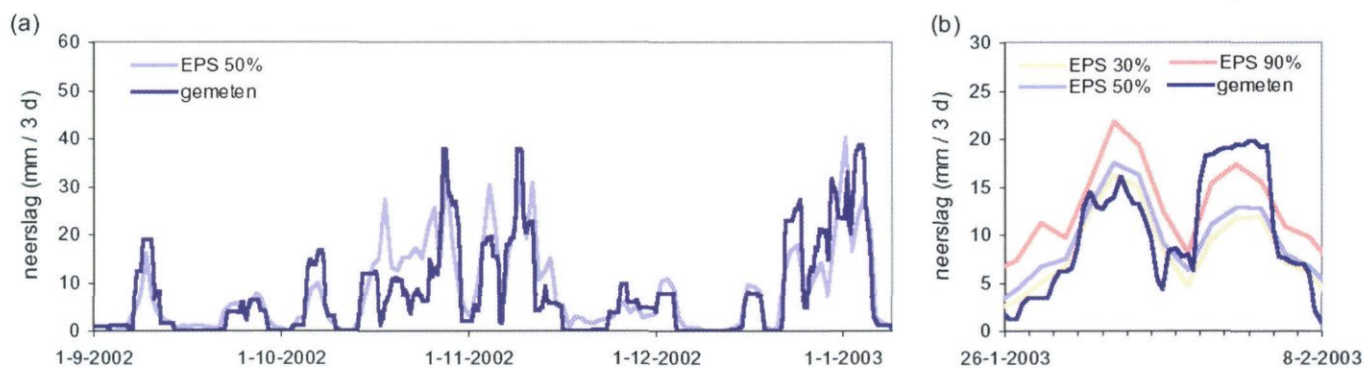
Zoals geïllustreerd is in afbeelding 4b, is het inherent aan een kansverwachting dat de ene keer de neerslagsom onder het 30%-percentiel van het EPS-ensemble kan vallen en de andere keer ruim boven het 90%-percentiel. Dat maakt een goede beoordeling van de situatie lastig. Daarbij komt nog dat een meteorolo-



Afb. 2: Verwerkte EPS-neerslagverwachtingen: (a) verwerking tot wateroverschot en -tekortverwachting, (b) verwerking tot waterstandverwachting met behulp van een minimaal (30%-percentiel), gemiddeld (50%-percentiel) en maximaal neerslagscenario (90%-percentiel) (HydroNET, 2004).

Afb. 3: Talagrand-diagrammen voor De Bilt (september 2002 t/m december 2004): respectievelijk zes uur, drie dagen en tien dagen vooruit.





Afb. 4: (a) Vergelijking van de gemeten driedaagse neerslagsom en het 50%-percentiel van de neerslagsom voor de eerste drie dagen voor De Bilt, (b) EPS-percentielen en gemeten driedaagse neerslagsommen voor De Bilt.

loog niet alleen EPS gebruikt, maar ook andere informatiebronnen en ervaring. Een waterbeheerder heeft niet de beschikking over deze aanvullende informatie en bovendien niet de tijd om elke dag de beschikbare informatie te analyseren. Het zou dus nuttig zijn als de analyse automatisch verloopt en een vast percentiel als criterium kan worden gebruikt voor al dan niet anticiperend ingrijpen. Welke percentielwaarde gehanteerd moet worden, hangt niet alleen af van de kwaliteit van EPS voor het verwachten van grote neerslagsommen. Ook is van belang hoe vaak men bereid is onterecht te reageren. Altijd reageren als minimaal één run hoge sommen verwacht, geeft de meeste zekerheid, maar heeft wel tot gevolg dat naar verhouding erg vaak ten onrechte wordt ingegrepen. In afbeelding 5a wordt dit geïllustreerd aan de hand van een voorbeeld. Gekeken is naar de EPS-ensembles voor De Bilt van 12:00 uur 's middags over de periode september 2002 t/m december 2004. Gesteld is dat gereageerd moet worden als er meer dan 25 mm neerslag voor de komende drie dagen wordt verwacht. Een verwachting was goed als het begin van een extreme periode exact op het juiste moment werd verwacht, maar ook als de ver-

wachting binnen marges van twaalf uur te vroeg of te laat was. 'Goed, maar te laat' wil zeggen dat de periode met grote neerslagsommen wel voorzien was, maar dat het begin van deze periode meer dan twaalf uur te laat werd verwacht.

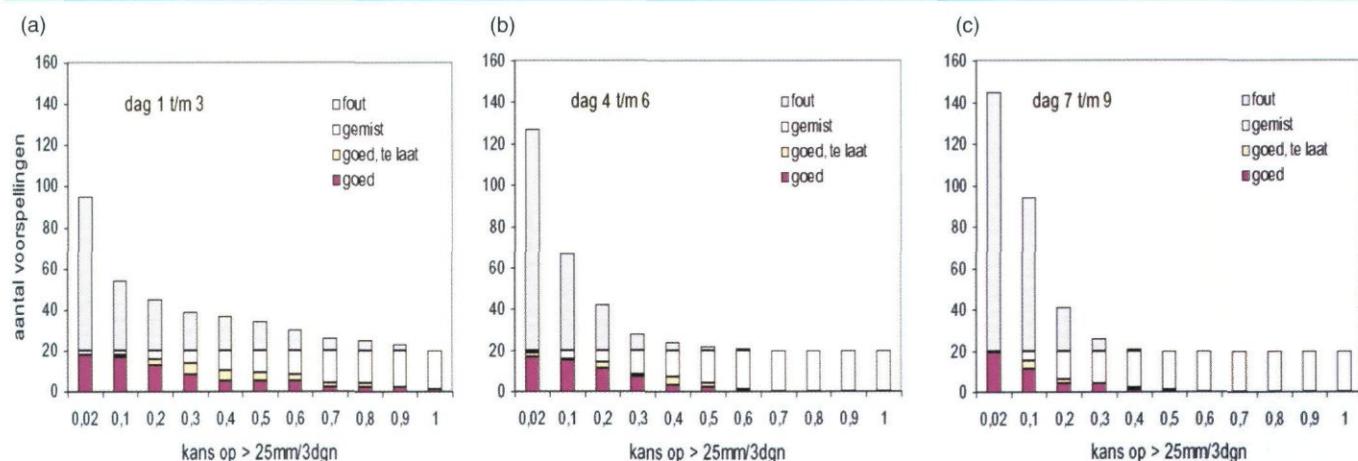
Zoals eerder gesteld is altijd reageren op minimaal één run boven de 25 mm het meest veilig in verwachte situaties van wateroverlast. In dat geval werd 18 van de 20 keer dat er daadwerkelijk meer dan 25 mm viel, goed gereageerd en werd er twee keer gemist. Dit goede resultaat ging echter gepaard met 77 keer een onterechte ingreep. Werd pas gereageerd als tien procent van de runs meer dan 25 mm in drie dagen aangaf, dan werd twee keer een extreme gebeurtenis gemist en één keer te laat gereageerd. Het aantal onterechte reacties was teruggebracht naar 34.

Afbeeldingen 5b en 5c tonen vergelijkingen van de verwachte driedaagse sommen voor dag 4 t/m 6 en dag 7 t/m 9. Nog steeds werden de meeste extreme perioden voorzien als altijd gereageerd werd als er minimaal één run meer dan 25 mm in drie dagen verwachtte. Het aan-

tal onterechte reacties dat daarmee gepaard ging, nam echter enorm toe. Waren dat er in afbeelding 5a nog 77, in afbeelding 5b en 5c gaat het om respectievelijk 107 en 125. Hieruit kan een belangrijke conclusie worden getrokken: als alleen wordt gelet op overschrijdingskansen van meerdaagse sommen, presteren de EPS-ensembles van de eerste drie dagen het best. Uit de statistische verificatie kwamen deze echter als slechtste naar voren. Een andere constatering is dat het aantal treffers en foute voorspellingen voor langere verwachtingstermijnen sneller afneemt bij grotere overschrijdingskansen. Dit wijst erop dat de pluim steeds verder uitwaaert richting de klimatologische kansverdeling. De statistische juistheid van de pluim is niet veranderd, de prestatie echter wel.

In de voorbeeldanalyse van afbeelding 5 is 25 mm in drie dagen gebruikt als grenswaarde voor de neerslagsom. Een waterbeheerder zal uit ervaring weten hoeveel neerslag zijn beheersgebied ongeveer kan verwerken. Met behulp van de analyse kan dan een indruk worden verkregen van de waarde van verschillende overschrijdingskansen en kan een keuze worden gemaakt voor een alarmprempel.

Afb. 5: Verificatie van overschrijdingskansen voor de eerste drie dagen, dag 4 t/m 6 en dag 7 t/m 9. De Bilt, september 2002 t/m december 2004.



Het optreden van wateroverlast en de daarmee gepaard gaande kosten zijn echter van meer factoren afhankelijk dan alleen van de neerslag. Het gedrag en de huidige toestand van het watersysteem zijn van belang, maar ook de plaats van de wateroverlast en de manier van ingrijpen. Gaat het om een woonwijk of een weidegebied bestaat ingrijpen uit het aanzetten van een pomp of het preventief oplopen van kaden?

Een volgende stap in het voorkómen van hoogwater is dan ook de EPS-ensembles als invoer te gebruiken voor een model van het watersysteem (afbeelding 2b)^{3),4)}. Dan wordt een goed beeld verkregen van de invloed van de neerslaghistorie, kan goed gezien worden waar problemen op kunnen treden en kan gekeken worden naar de effecten van verschillende maatregelen. De analyse van afbeelding 5 kan herhaald worden, maar dan met gemeten en verwachte waterstanden. De uitkomsten kunnen ook goed aan een kosten-batenanalyse worden onderworpen, waardoor een betere keuze gemaakt kan worden voor een alarmdrempel dan alleen op basis van neerslag.

Conclusies

Als EPS-verwachtingen vergeleken worden met puntwaarnemingen, blijkt dat de theoretische kansverdeling van de ensembles niet helemaal juist is. Het gegeven dat een kansverwachting statistisch correct is, zegt echter niet veel over de prestaties van de verwachting voor het weer van morgen. Die zijn beter te beoordelen aan de hand van overschrijdingskansen van neerslagsommen. Uit een dergelijke analyse voor driedaagse neerslagsommen van De Bilt komt naar voren dat de eerste drie dagen van het EPS-ensemble het best presteren, ondanks dat ze statistisch het minst correct zijn.

Op basis van de analyse van overschrijdingskansen kan een eerste keuze worden gemaakt voor een verantwoorde alarmdrempel. Deze keuze kan verder worden onderbouwd met een modelstudie en een kosten-batenanalyse.

De statistische verificatie, de analyse van overschrijdingskansen en een eventuele modelstudie kunnen verbeterd worden door nauwkeurige gebiedsdekkende (radar)neer-

slagsommen te gebruiken. Het is te verwachten dat de aanname dan beter klopt dat alle ensemblesleden even waarschijnlijk zijn. 

LITERATUUR

- 1) Holleman I. (2004). Weerradar en de neerslag van augustus 2004. Meteorologica nr. 4, pag. 18-22.
- 2) Kok C. (2001). Calibration of EPS derived probabilities. Scientific report WR-2001-04. KNMI.
- 3) Lobbrecht A., G. Hiemstra, M. Talsma en Z. Vonk (2003). Neerslaginformatie voor het waterbeheer. H₂O nr. 23, pag. 22-25.
- 4) Lobbrecht A., S. Loos, G. Versluis en M. Talsma (2004). Neerslaginformatie voor hoogwaterverwachting in De Overwaard. H₂O nr. 18, pag. 29-31.