

MEMORANDUM

VOLGNUMMER : WM 05-02

ONDERWERP : Naar een andere weeralarmsystematiek

SAMENSTELLER : Kees Kok

DATUM : mei 2005

Dit is een persoonlijk memorandum. Slechts de auteur is verantwoordelijk voor de inhoud. Indien niet meer nodig, graag retourneren aan de auteur.

Naar een andere weeralarmsystematiek

Kees Kok

Inhoudsopgave

Abstract	4
I. Inleiding	5
II. Huidige systematiek	6
1. Schematisch overzicht	6
2. Welke events	7
3. Problemen met de huidige systematiek	8
III. Nieuwe situatie	10
1. Voorbeeld: het automatische alertsysteem voor de waterschappen	10
2. Voorstel nieuwe aanpak	12
a. loskoppeling meteo en verwachte impact	
b. eisen aan de definities van de gevaarlijk weer events	
c. alle afzonderlijke stappen zichtbaar maken	
d. alles in kansen	
e. variabele drempels	
f. communicatie naar buiten	
g. de rol van de meteoroloog	
h. semi-continue verwachtingen	
i. presentatie	
j. verificatie	
IV. Hoe kan verificatie bijdragen aan verbetering	20
1. Problemen met de huidige verificatie	20
2. Voorstel nieuwe verificatie	21
V. Slotopmerkingen	25
Dankbetuiging	28
Literatuur	28

Abstract

De huidige weeralarmverwachtingen zijn van slechte kwaliteit. In de afgelopen 6 jaar is meer dan de helft van de weeralarmsituaties "gemist". De gebruikte methode is erg ondoorzichtig, zowel voor afnemers als voor forecasters, en allerlei argumenten van meteorologische aard en van mogelijke (ingeschatte) impact lopen door elkaar. Het is niet duidelijk hoe en waar deze argumenten een rol spelen, en ze kunnen achteraf ook niet meer zichtbaar gemaakt worden. De huidige systematiek leidt dientengevolge tot slechte resultaten, gaat in de besluitvorming gepaard met vele, veelal niet-meteorologische dilemma's, stelt de forecasters voor "onmogelijke" problemen en vormt bovendien een barrière voor potentiële verbeteringen. Tenslotte, terwijl het de bedoeling zou moeten zijn om in een zo vroeg mogelijk stadium zo veel mogelijk informatie te verschaffen over verwacht gevaarlijk weer, neigt de huidige gang van zaken eerder naar het omgekeerde.

De slechte verwachtingen zijn hoofdzakelijk het gevolg van een gebrekkige systematiek en niet noodzakelijkerwijs het gevolg van een slechte inschatting van het verwachte weer door de forecaster, of van een slechte performance van de modellen of van het ontbreken van de nodige tools. Er kan niet blootgelegd worden aan welke onderdelen van het verwachtingsproces iets mankeert. De trend in de kwaliteit van de verwachtingen kan niet bepaald worden, ook niet omdat de eisen aan wat te verwachten steeds aan discussie onderhevig zijn. En dus kan er van de verificatie ook niet een stimulans uitgaan tot verbetering van de verwachtingen.

In dit verhaal wordt een andere systematiek voorgesteld voor het maken van gevaarlijk weer verwachtingen in het algemeen en weeralarmen in het bijzonder. Deze is volledig gebaseerd op de kansen die inherent aan de besluitvorming ten grondslag liggen. Het scheidt de meteorologie van de impact voor de maatschappij zodat het vermengen van argumenten in de totstandkoming van de verwachtingen vermeden wordt. Hierdoor worden grote problemen voor de forecasters opgelost en kan veel meer waarde worden geleverd aan veel meer gebruikers. De voorgestelde systematiek lijkt (niet toevallig) erg op die van het automatische alertsysteem voor de waterschappen.

Door de voorgestelde strikte scheiding tussen meteo en wat de maatschappij er mee doet (al of niet waarschuwen, al of niet andere acties ondernemen) kan het KNMI weer afgerekend worden op puur de meteo (zoals het hoort), i.e. op de skill van de verwachting en de tijdigheid van uitgifte, en dus niet meer op de verwachting van het effect van het verwachte weer op een "gemiddelde" gebruikersgroep.

Willen we als KNMI de maatschappij beter bedienen met gevaarlijk weer verwachtingen (niet alleen weeralarmen), dan moet er maatwerk geboden worden in de vorm van kansen voor verschillende gebruikers. Niet in de zin dat er voor iedere gebruiker een apart verwachtingssysteem ontwikkeld moet worden of aparte verwachtingssystemen t.b.v. weeralarmen. Dit is niet zinnig omdat de wensen en eisen snel veranderen (zie de weeralarm-criteria). Beter is het om veel meer onderzoek te doen aan gevaarlijk weer, de voorspelbaarheid ervan, naar betere modellen, short range ensembles, en meer aandacht te schenken aan opleiding van de forecasters.

Tot slot wordt in dit verhaal ook een andere manier van verificatie voorgesteld die, uitgaande van de voorgestelde systematiek, het mogelijk maakt de kwaliteit van de weeralarmverwachtingen (indirect) te monitoren, en te analyseren waar tekortkomingen zijn en waar bijsturing dient plaats te vinden. Deze methode kan i.t.t. de huidige een bijdrage leveren om de verwachtingen op termijn te verbeteren.

I. INLEIDING

De huidige gang van zaken bij de totstandkoming van bijv. een weeralarm is dat de meteorologische situatie door de meteoroloog ingeschat wordt op grond van uitvoer (van een veelheid) van modellen, eventueel gecombineerd met recente waarnemingen en een stuk eigen ervaring. Deze inschatting is (bijna) noodzakelijkerwijs een *kans*schatting of de voorgeschreven drempelwaarde zal worden overschreden. Vervolgens wordt op grond van een groot aantal zaken (ook subjectieve) een besluit genomen of een weeralarm uitgegeven wordt of niet. Dit laatste is een *categorische* (ja/nee) uitspraak. Deze laatste uitspraak wordt gebaseerd op de inschatting van de meteoroloog (!) gebaseerd niet alleen op de waarschijnlijkheid van optreden van het event maar ook op de verwachte impact op "de" gebruikers. Hierbij worden allerlei overwegingen betrokken zoals de tijd van de dag, de plaats in het land, hoeveel waarschuwingen er onlangs al geweest zijn, etc. en tenslotte ook een inschatting van "de" gebruiker (zonder uitsplitsing naar specifieke gebruikers). Het is geen wonder dat bovenstaande tot allerlei problemen leidt. Uit eerdere evaluaties blijkt dat de kwaliteit van de verwachtingen onvoldoende is en ook is er een groot aantal dilemma's gesignaleerd.

In dit stuk wordt een andere systematiek geschetst voor het maken en uitgeven van gevaarlijk weer verwachtingen en waarschuwingen. De weeralarmen zijn hier een onderdeel van. Het biedt een framework waarbinnen dit soort verwachtingen op een andere manier gemaakt wordt zodat veel meer waarde geleverd kan worden aan een veel grotere groep gebruikers. Het biedt in tegenstelling tot de huidige praktijk een belangrijke mogelijkheid tot verbetering van de skill van de verwachtingen en levert een handvat voor de verificatie van dit soort verwachtingen.

Niet alle aspecten zijn (even) nieuw. Aan een deel van de hier voorgestelde veranderingen wordt op dit moment al gewerkt.

II. Huidige systematiek

II.1. Schematisch overzicht

In de discussie over weeralarm en alles wat daar bij hoort is het goed om de verschillende facetten in de besluitvorming zoveel mogelijk uit elkaar te houden. Deze zaken zijn wel nauw met elkaar verweven maar het verduidelijkt de discussie als je weet over welk deel van het traject je het hebt. Daarom wordt in het volgende eerst een schema gegeven dat het proces om te komen tot een weeralarm in een aantal stappen uiteen haalt. Hoewel niet iedere forecaster precies deze stappen zal doorlopen ziet het schema er ongeveer uit als in de volgende figuur.

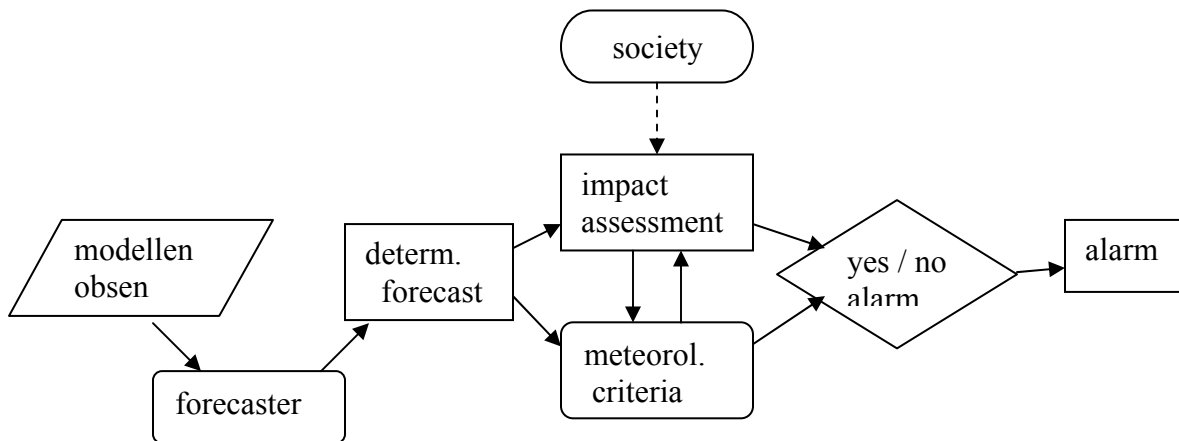


Fig. II.1. Schematisch overzicht van de verschillende stappen in het besluitvormingsproces voor het al of niet uitgeven van een weeralarm. De gestippelde lijn is iets dat on-line geen rol speelt. Zie tekst voor verdere uitleg.

Weeralarmen zijn ontstaan door de behoefte om de maatschappij te waarschuwen voor grote gevolgen die zwaar weer zou kunnen hebben. Bij "grote" gevolgen moet gedacht worden aan maatschappij-ontwrichtend of veiligheidsrisico vormend voor grote groepen mensen. Een aantal van door het weer veroorzaakte gevaarlijke omstandigheden is een aantal jaren geleden gedefinieerd, plus de daarbij ingeschatte meteorologische criteria. Deze impact en de bijbehorende meteorologische criteria liggen min of meer vast en zijn tot stand gekomen middels besprekingen tussen maatschappelijke instanties (*society* genoemd in de figuur), weerproviders en het KNMI en staan tijdens het verwachtingsproces in principe niet ter discussie (vandaar de stippellijn in het schema). Deze meteorologische criteria zijn op het ogenblik het hele jaar door constant. De vertaling van de impact op de maatschappij naar de meteorologische criteria is zeer onzeker en zeker niet eeneenduidig. Het op grote schaal omwaaien van bomen kan een groot veiligheidsprobleem zijn, maar de vertaling naar een meteorologisch criterium (bijv. een bepaalde windsnelheid) waarboven bomen massaal omwaaien zal in de winter anders zijn dan in de zomer.

Overleg tussen "society" en het KNMI bepaalt uiteindelijk voor welke fenomenen er gewaarschuwd dient te worden. Welke dat op dit moment zijn staat in paragraaf II.2.

In de nu gebruikelijke verwachtingssystematiek (de stroom van links naar rechts in het schema) is het zo dat de meteoroloog de beschikbare uitvoer van verschillende modellen interpreteert, deze combineert met laatste waarnemingen en met eigen inzichten, en uiteindelijk komt tot een min of meer deterministische uitspraak van bijvoorbeeld de maximale windsnelheid. Deze wordt vervolgens getoetst aan de afgesproken weeralarm criteria, gecombineerd, in meer of mindere mate, met de dan te verwachten schade cq risico (als zijn verwachting uitkomt). Als deze schade of risico naar zijn inzicht te groot is dan wordt een weeralarm uitgegeven. Iedere forecaster maakt zijn afwegingen op een andere manier, maar globaal gesproken volgt het een proces zoals in bovenstaand schema. In alle stappen is sprake van impliciete onzekerheid die echter nergens expliciet gemaakt wordt. Omdat er uiteindelijk een ja/nee uitspraak dient te volgen wordt er zo lang mogelijk gewacht tot de onzekerheid zo klein mogelijk is. Het enige zichtbare en verifieerbare in deze systematiek is de meteo informatie die beschikbaar was alsmede de uiteindelijke uitspraak.

In de dagelijkse verwachtingssystematiek voor extreem weer wordt door de forecaster een inschatting gemaakt van de mogelijke gevolgen van het weer. Hij / zij wordt daar ook min of meer op afgerekend. Is het criterium niet gehaald maar is er wel veel schade zal men het niet altijd terecht vinden dat er geen alarm is uitgegeven. Omgekeerd, is het criterium wel gehaald maar bleef schade uit dan kan "geen alarm" een succes zijn. Dit leidt er toe dat de meteoroloog geneigd is niet puur naar de ingeschatte hevigheid van het weerfenomeen te kijken maar ook een inschatting van de impact van het weer te betrekken bij de beoordeling of een weeralarm uitgegeven dient te worden.

II.2. Welke gevaarlijk weer events

De weeralarmcriteria worden bepaald op grond van hun maatschappij-ontwrichtend karakter, hun waarde voor de samenleving, veiligheid van de burger, of iets dergelijks. Deze en andere gevaarlijk weer criteria worden opgesteld in nauw overleg met maatschappelijke instanties, zoals politie, brandweer, waterschappen, etc. (in de figuur samengevat als society).

Randvoorwaarden: volgens de wet zullen ze iets met "gevaarlijk" weer voor de maatschappij te maken moeten hebben. En de grenzen zijn bovendien zodanig gekozen dat het verwachte aantal weeralarmen per jaar niet al te hoog is (i.o.m. de providers).

Op het ogenblik zijn er criteria vastgelegd voor storm / windstoten, sneeuw, ijzel, en onweer. Tot enkele jaren geleden was er ook een criterium voor regen. De lijst en definities zijn af en toe onderwerp van herbezinning. Zie voor meer informatie over de actuele weeralarm grootheden de CWK-site.

II.3. Problemen met de huidige systematiek

De systematiek om te komen tot een weeralarm kent veel problemen en tekortkomingen. Allereerst deugt een aantal van de definities van de te voorspellen grootheden niet. Dit is gemakkelijk (maar na veel overleg) op te lossen door ze te vervangen door correcte, heldere definities. Dit moet uiteraard in overleg gebeuren met gebruikers zoals dat ook voor de huidige definities al bestaat. (Hierover later meer, zie III.2). Daarnaast heerst er nogal wat onbekendheid over de exacte definities (criteria) van de te voorspellen fenomenen niet alleen bij de gebruikers maar ook bij de forecasters. Ook dit is vrij simpel op te lossen.

Een veel fundamenteeler probleem met "het weeralarm" ligt in de hele systematiek om te komen tot het al of niet uitgeven van een weeralarm. In de huidige gang van zaken worden de tussenstappen niet expliciet zichtbaar gemaakt. De meteoroloog vormt zich vanuit de modeloutput, waarnemingen en eigen (meteorologische) expertise een beeld van het verwachte weer. Hieruit wordt samen met een eigen schatting van het mogelijke effect van dit verwachte weer rechtstreeks een categorische uitspraak gedaan over het al of niet optreden van een weeralarmsituatie. Alleen deze eindconclusie is zichtbaar (en verifieerbaar), de tussenstappen die er wel degelijk zijn gemaakt om tot het eindoordeel te komen zijn ondergesneeuwd, worden niet opgeslagen en worden niet meegenomen in de afrekening. Deze gang van zaken heeft grote nadelen. Zo kan niet duidelijk worden waar tekortkomingen aan te wijten zijn en is bijsturing en / of verbetering van deze onderdelen per definitie onmogelijk.

Er worden tijdens het verwachtingsproces allerlei impliciete aannames gemaakt over de gebruikers en over hoe de waarschuwing bij hen over komt. Dit heeft tot gevolg dat er vragen opkomen als "moet er niet eerder gewaarschuwd worden in de Randstad", of "overdag eerder dan 's nachts", etc. Ook wordt de psychologie van de gebruiker meegenomen in de uiteindelijke beslissing. Meteorologen willen anticiperen op wat de gebruiker vindt van hun verwachting. De gebruiker wil voor iedere opgetreden event vooraf gewaarschuwd zijn en tolereert misschien wat makkelijker false alarms. Deze tolerantie is afhankelijk van de recente geschiedenis: is er al een paar keer onterecht niet gewaarschuwd dan zal een volgende keer veel sneller gewaarschuwd worden en omgekeerd. En nog gekker: is er een zwaar extreem voorval gemist dan wordt er voor een kort daarop verwacht minder extreem voorval maar geen alarm uitgegeven om de hoon van het publiek te ontlopen. Dit is al met al een zeer moeilijke positie voor de meteoroloog. Bovendien heeft dit probleem niks met meteorologie te maken. De huidige gang van zaken leidt daarom dan ook niet tot een mechanisme waardoor de meteoroloog (meteorologisch gezien) "betere" verwachtingen gaat maken, zelfs integendeel. Het leidt er in sommige gevallen toe dat achteruitgang in de (meteorologische) skill van de waarschuwingen juist gestimuleerd (beloond) wordt. (Zie ook de tevredenheid van Achmea over een onterecht uitgegeven weeralarm!)

Alle informatie die de meteoroloog heeft over de weersontwikkeling (inclusief zijn onzekerheid hierover) wordt uiteindelijk verdisconteerd in één statement over het gevaar voor "de" gebruiker. Zo'n categorische uitspraak blokkeert de mogelijkheden voor niet alleen "de" gebruiker maar ook alle andere gebruikers (het overgrote deel dus) om meer waarde uit de aanwezige kennis bij de meteoroloog over de verwachte weersontwikkeling te halen. Essentiële informatie wordt dus achtergehouden.

Tenslotte, doordat men gedwongen is een harde (i.e. ja/nee) uitspraak te doen zal zolang mogelijk gewacht worden. Er is geen enkele drijfveer om zo vroeg mogelijk te waarschuwen (er is geen concurrentie en geen harde afrekening, het publiek wordt toch wel gewaarschuwd voor zwaar weer). Het omgekeerde is juist het geval. Men kiest blijkbaar voor het beperken van het aantal false alarms boven het tijdig / eerder waarschuwen van de gebruikers. Dit heeft tot gevolg dat maar liefst 60 procent van de weeralarmsituaties "gemist" wordt! (Zie par. IV.1. Hoe lang van te voren gewaarschuwd wordt wordt niet eens meegenomen in de verificatie!!!)

Terwijl het de bedoeling zou moeten zijn om in een zo vroeg mogelijk stadium zo veel mogelijk informatie te verschaffen over verwacht gevaarlijk weer, neigt de huidige gang van zaken meer naar het omgekeerde. Samengevat kan men stellen dat de huidige systematiek leidt tot slechte resultaten, zowel in skill als in tijdige informatievoorziening, de forecasters voor "onmogelijke" problemen stelt en een barrière vormt voor potentiële verbeteringen.

III. Nieuwe situatie

III.1 Voorbeeld: het automatische alertsysteem voor de waterschappen

Omdat de voorgestelde systematiek veel lijkt op de systematiek die gehanteerd wordt in het automatische waarschuwingingsysteem voor de waterschappen en ook omdat het lekker overzichtelijk is volgt eerst een korte beschrijving van dat systeem.

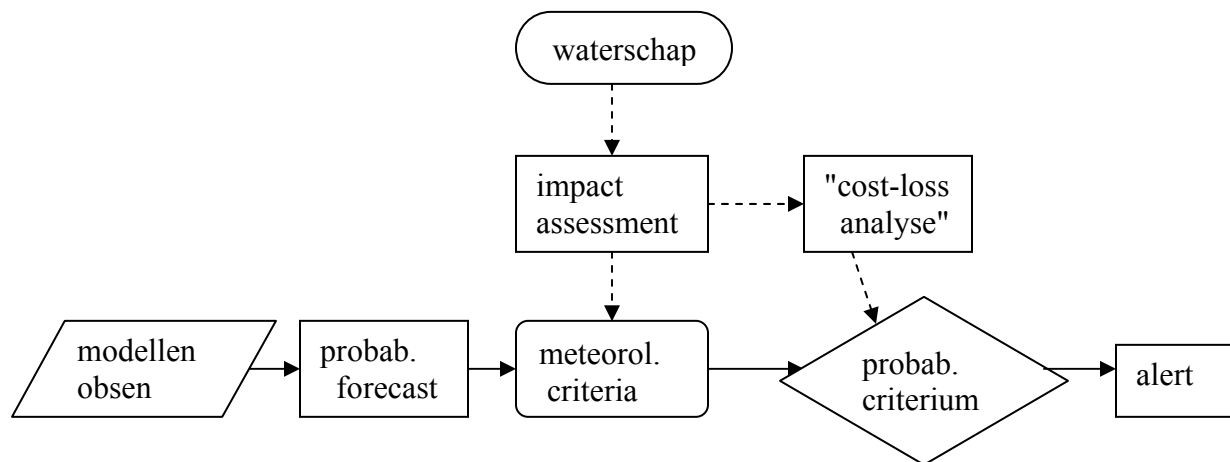


Fig. III.1. Schematisch overzicht van het alertsysteem voor de waterschappen. Hier is alleen de probabilistische tak geschetst. De gestippelde lijnen spelen on-line geen rol. Zie de tekst voor verdere uitleg.

Na uitvoerig overleg met ieder deelnemend waterschap is allereerst bepaald wat de kritische neerslaggrenzen zijn waarboven er grote schade op zou treden. Deze grenzen worden hoofdzakelijk bepaald door de maximale pomp- en lozingscapaciteit van het betreffende waterschap in combinatie met de bergingscapaciteit. Dit leidt tot een aantal kritische drempels, en in het geval van de waterschappen zijn deze dit direct eeneenduidig gekoppeld aan meteorologische criteria. Bijv. 20mm / dag en 30 mm / 2dagen. In deze criteria kan ook eventueel de reeds gevallen hoeveelheid neerslag betrokken worden (i.e. de afname van de opslagcapaciteit). Een tweede stuk informatie dat afkomstig is van de waterschappen is de hoeveelheid schade (losses) die geleden wordt bij overschrijding van de kritische drempel in combinatie met de kosten die gemaakt zouden moeten worden om die schade te vermijden (costs). Met behulp van zo'n cost-loss-achtige analyse is er een schatting gemaakt bij welk kanspercentage waarmee de overschrijding van een drempel verwacht wordt, het betreffende waterschap zal moeten beginnen voorzorgsmaatregelen te treffen. Is bijvoorbeeld de geleden schade erg groot bij teveel neerslag en zijn de kosten ter voorkoming relatief laag dan zal men eerder geneigd zijn voorzorgs-maatregelen te treffen (i.e. bij een lage kans). Is daarentegen de cost-loss verhouding erg groot dan zal men pas bij een hoge kans actie ondernemen.

Per waterschap is soms een tiental kritische drempels (per seizoen) gedefinieerd met bijbehorende probabilistische cost-loss drempels. Er is geen limiet aan het aantal per waterschap. Het systeem is zodanig ingericht dat desnoods per uur van de dag verschillende criteria en kansdrempels gekozen kunnen worden. Deze zijn zoals gezegd rechtstreeks afkomstig van de waterschappen. Ze zijn vooraf afgesproken en vastgelegd en in de operationele praktijk wordt alleen nog rekening gehouden met deze criteria; er wordt dus niet meer gekeken naar de impact. Hiermee is de meteorologie (inclusief de reeds gevallen neerslag) losgekoppeld van de rest. Ieder uur wordt er voor alle drempelwaarden van alle waterschappen een overschrijdingskans berekend uitgaande van de modelverwachting en andere relevante informatie. Is de kans groter dan de opgegeven cost-loss kans dan krijgt het betreffende waterschap automatisch een alertmail met alle relevante informatie. Wat er vervolgens mee gedaan wordt (peilbeheerders waarschuwen, pompen, etc.) is aan het waterschap. In dit hele proces is geen meteoroloog betrokken. Deze is slechts beschikbaar voor vragen na een gegeven alert en bij toelichting off-line hoe het beste met de verwachtingen omgegaan kan worden.

Hoewel de waterschappen een groot aantal alertdrempels hebben kunnen opgeven zijn er situaties denkbaar dat een drempel tijdelijk een andere waarde zou moeten hebben. Een voorbeeld hiervan is het uitvallen van een gemaal. In zo'n situatie zou men eigenlijk bij een lagere neerslagdrempel gewaarschuwd willen worden. Daarom is er voor de waterschappen altijd de mogelijkheid om de kansverwachtingen voor andere drempels te bekijken. De kansverwachtingen worden namelijk uurlijks automatisch gemaakt voor een hele serie drempelwaarden. Analoog zijn er situaties denkbaar dat men incidenteel bij een andere kansdrempel gewaarschuwd zou willen worden. Ook voor dit doel kunnen staan alle verwachtingen ter inzage. Op deze manier hoeven niet dagelijks de afgesproken drempelwaarden aangepast te worden. Het alertsysteem voorziet dus in een set kansverwachtingen voor iedere overschrijdingsdrempel die altijd "klaar staan", dat automatisch een alert genereert en verzendt als de opgegeven kansdrempels overschreden worden, en dat geraadpleegd kan worden bij allerlei onvoorziene omstandigheden.

Het waterschapsysteem staat model voor de voorgestelde gevaarlijk weer (en weeralarm) systematiek. Voor meer informatie over het alertsysteem voor de waterschappen wordt verwezen naar Kok et al. (2005) en naar <http://info.knmi.nl/wm-am/stat/wmstat.html>.

III.2 Voorstel nieuwe aanpak

In dit hoofdstuk wordt een voorstel besproken voor een nieuwe aanpak van de weeralarmproblematiek. Deze aanpak verschilt in principe niet veel van de aanpak voor het verwachten van andere gevaarlijk weerfenomenen. Niet alle aspecten zijn nieuw t.o.v. de huidige situatie.

Voorstel:

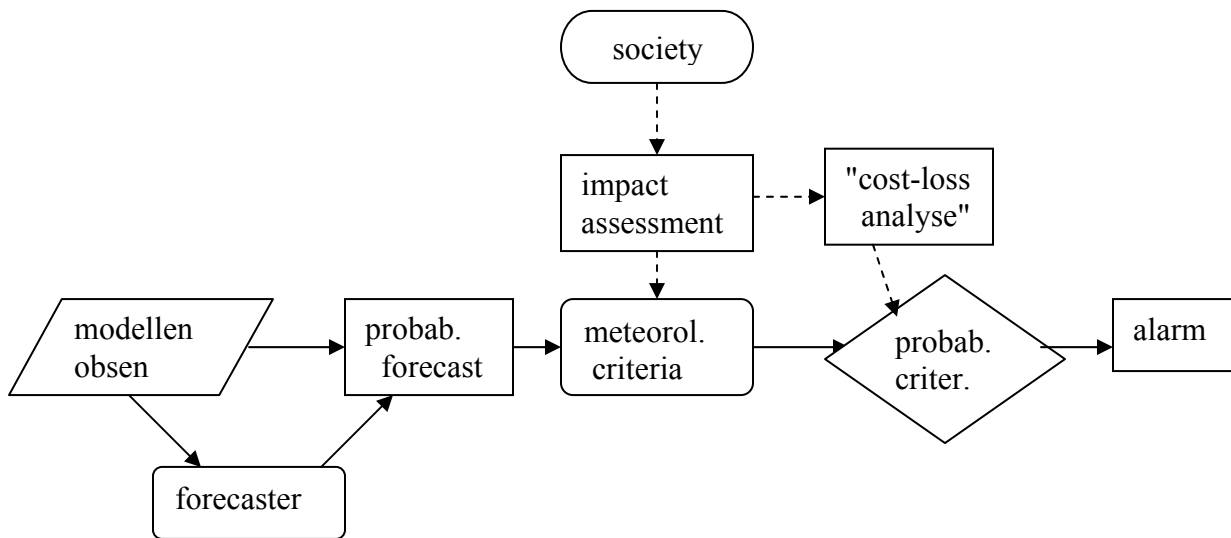


Fig. III.2. Schematisch overzicht van de voorgestelde stappen in het besluitvormingsproces voor het al of niet uitgeven van een weeralarm. De gestippelde lijn is iets dat on-line geen rol speelt. Zie tekst voor verdere uitleg.

In het bovenstaande schema staan de off-line processen aangeduid door middel van stippellijnen, het on-line verwachtingsproces staat van links naar rechts aangegeven met getrokken pijlen. De on-line inzet is uiteraard weersafhankelijk. Nadere uitleg over de verschillende onderdelen van het schema komt in de volgende paragrafen aan de orde.

In de volgende subparagrafen wordt een aantal van de meest in het oog springende veranderingen t.o.v. de huidige systematiek besproken. Terwille van de leesbaarheid wordt een aantal argumenten meerdere keren genoemd.

III.2a. loskoppeling meteo en verwachte impact

Er dient een absolute loskoppeling te zijn tussen het verwachtingsproces en de inschatting van de impact die het verwachte weer op de maatschappij zou kunnen hebben. De tweede moet als het ware vooraf gebeurd zijn ("off-line") en in afspraken vastgelegd waarna de forecaster zich on-line alleen met het weer bezighoudt. In deze sectie wordt het off-line gedeelte besproken. Dit deel is samengevat in de onderstaande figuur.

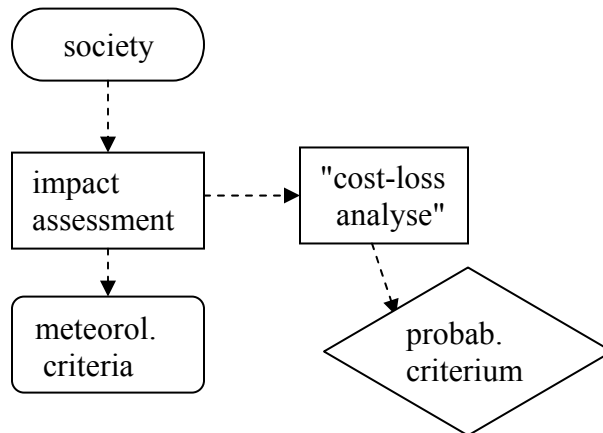


Fig. III.3. Schematisch overzicht van de stappen die leiden tot de vaststelling van de meteorologische gevaarlijk weer criteria en de bijbehorende kansdrempels. Zie tekst voor verdere uitleg.

De gevaarlijk weer fenomenen waarvoor gewaarschuwd moet worden, worden in overleg met allerlei instanties bepaald. Dit gebeurt op min of meer dezelfde manier als nu al het geval is. De criteria daarbij zijn het maatschappij-ontwrichtend karakter, hun waarde voor de samenleving, veiligheid van de burger, of iets dergelijks. De betrokken maatschappelijke instanties, zoals politie, brandweer, waterschappen, etc. worden hier verder samengevat als **society**. Deze criteria, die puur gebaseerd zijn op impact assessment (maatschappij-ontwrichtend, veiligheid, etc.) variëren in de tijd en de plaats. Denk aan spits – geen spits, zomer – winter, Randstad – Veluwe. Alle gevaarlijk weer items hebben per definitie betrekking op gebieden en niet op punten / locaties. Het aantal criteria kent in principe geen beperking; er is op dit moment echter één randvoorwaarde: volgens de wet zullen ze iets met "gevaarlijk" weer voor de maatschappij te maken moeten hebben. En de impact-grenzen zijn bovendien zodanig gekozen dat het verwachte aantal weeralarmen per jaar niet al te hoog is (i.o.m. de providers). Voor andere gevaarlijk weer verwachtingen zullen er minder restricties zijn.

De maatschappij-ontwrichtende, etc. fenomenen moeten vertaald worden naar meteorologische criteria. Bij welke windsnelheid, bij welke sneeuwintensiteit, etc. treedt de ontwrichting op. Ook dit is weer tijds- en plaatsafhankelijk. Bomen vangen 's zomers meer wind dan 's winters, etc.

Tenslotte moet ook bepaald worden bij welk kanspercentage de individuele gebruiker(sgroep) gewaarschuwd wil worden. Zie ook III.2.d en e. Absolute zekerheid of een gevaarlijk weer situatie op zal treden kan namelijk niet gegeven worden, dus er moet gewerkt worden met kansen. Dit is niet een nadeel maar juist een groot voordeel (zie later). Bij welke kans men gewaarschuwd wil worden hangt o.a. af van de mogelijkheid cq de wens om voorzorgsmaatregelen te treffen of in gang te zetten, het publiek (of andere betrokkenen) in te lichten, etc. Bij "lage" kansen zal men dit niet willen doen en bij "hoge" juist wel. Er vindt dus een soort "cost-loss analyse" plaats om uit te maken welke kansdrempel gehanteerd moet worden. Maar ook deze drempel is uiteraard weer tijds- en plaatsgebonden.

Grofweg samengevat: de maatschappij (society) bepaalt (in belangrijke mate) **wat** maatschappij-ontwrichtend, etc. is, samen met de meteorologen en andere deskundigen wordt bepaald wat de bijbehorende **meteorologische drempels** zijn, en samen wordt bepaald **bij welke kans** van optreden er gewaarschuwd moet worden. Daarnaast moeten in dit overleg afspraken gemaakt worden vanaf hoe lang van te voren men ingelicht wil worden. De kansdrempels, etc. kunnen uiteraard ook weer een functie van de voorspeltermijn zijn. Hiermee zijn "voorwaarschuwingen" dus onderdeel van dezelfde systematiek.

Het vaststellen van deze criteria zal een iteratief proces zijn. Het ontwikkelen van objectieve voorspelmethodes die precies "passen" op afgesproken criteria is dus niet erg zinvol; na ieder event of bijna-event zouden de regels daarover aangepast kunnen worden. Effort dient zich veeleer te richten op onderzoek naar de verbetering van het inzicht over het ontstaan van "gevaarlijk weer", naar betere modellen en naar de betere toepasbaarheid van de modellen (in brede zin) voor gevaarlijk weer situaties.

III.2b. *eisen aan de definities van de gevaarlijk weer events*

De huidige weeralarm criteria zijn niet altijd correct gedefinieerd. Er moeten de volgende eisen aan gesteld worden.

* De criteriadefinities moeten *mece* zijn: mutually exclusive and collectively exhaustive, dwz. een opgetreden situatie behoort eeneenduidig wel of niet tot het event.

Dus niet zoals het oorspronkelijke criterium voor zwaar onweer: "Op grote schaal (ten minste ter grootte van een provincie) met minstens 15 ontladingen per minuut binnen een straal van 15 kilometer". Gebiedsgroottes moeten exact gedefinieerd zijn (in aantal vierkante km's en in de vorm). Uitdrukkingen als "eventueel gepaard gaand met" zijn uit den boze, voegen althans niets toe. Hetzelfde geldt voor "op grote schaal" en "grote kans op".

* De definitie moet altijd tenminste 3 componenten bevatten: drempelwaarde (hoeveelheid, prikwaarde), tijd (interval, duur, intensiteit) en gebiedsgrootte.

NB. Of we achteraf eenduidig a.d.h.v. de waarnemingen kunnen bepalen of een event is opgetreden of niet doet er hier niet toe. Dit mag niet de definitie van het event beïnvloeden. Zie ook hoofdstuk IV.2.

Bij onweer of regen speelt dit probleem wat minder door de relatief goede dekking van de waarnemingen, maar voor ijzel, hagel of windstoten kan vaak niet eenduidig worden vastgesteld of aan de gegeven criteria is voldaan of niet. Criteria waarin de meetbaarheid of verifieerbaarheid verdisconteerd wordt moeten vermeden worden. Het gaat er namelijk niet om (voor de gebruikers) of op waarnemstations het te voorspellen fenomeen optreedt, maar of het ergens in het betreffende gebied optreedt. Zie ook IV.2.D.

NB. Deze definities staan in principe volkomen los van de vraag of er objectieve verwachtingstechnieken voor zijn. Ze zijn puur gebaseerd op wat de maatschappij belangrijk vindt. Een ander ding is dat er wel een indicatie van de skill gegeven moet kunnen worden. Deze bepaalt namelijk mede of het wel zinnig is om verwachtingen te gaan maken / uitgeven voor een bepaald event. Als de skill minimaal zou zijn, dan zal zo'n event (voorlopig) niet opgenomen worden in het rijtje events die in aanmerking komen voor een weeralarm.

NB. De meteorologische criteria bepalen de te voorspellen grootheid. Dit mag niet verward worden met de verwachte kans dat de te voorspellen grootheid optreedt ofwel met de kans dat er aan de criteria voldaan wordt. Deze staat los van de definities.

III.2c. *alle afzonderlijke stappen zichtbaar maken*

Alle afzonderlijke stappen en processen die bij de totstandkoming van gevaarlijk weer verwachtingen een rol spelen, zoals aangeduid in de figuur aan het begin van dit hoofdstuk, moeten expliciet zichtbaar gemaakt en gemonitord worden. Dit geldt zowel voor de off-line zaken (zie III.2.a) als voor alle on-line zaken. Alleen een analyse van al deze deelprocessen kan inzicht verschaffen waar problemen vandaan komen en waar eventueel bijsturing of extra inzet noodzakelijk is. Onderdeel van zo'n analyse is een grondige verificatie. Voor verdere details hierover zie hoofdstuk IV.

III.2d. *alles in kansen*

Het weer heeft een duidelijk stochastisch karakter. Dit geldt met name voor extreem weer (en dus voor gevaarlijk weer). Informatie over gevaarlijk weer kan dus per definitie het best in de vorm van kansen gebeuren. Alle verwachtingen dienen dan ook gegeven te worden in probabilistische vorm. Dit geldt dus niet alleen voor alle eventuele tussenstappen maar ook voor de verwachting die naar de gebruikers gaat. Deze verwachting moet niet boven een bepaalde kansdrempel alsnog als categorische statement uitgegeven worden maar moet in kansen gecommuniceerd worden.

Het is al sinds tenminste Thompson (1952) bekend dat kansverwachtingen veel meer waarde hebben voor gebruikers dan deterministische of categorische verwachtingen. Dit omdat kansverwachtingen inherent veel meer nuttige informatie bevatten. De inherente onzekerheid van de te voorspellen grootheden is skilful te kwantificeren en zeer nuttig te gebruiken. Er bestaat inmiddels een overweldigende hoeveelheid literatuur over de meerwaarde van kansverwachtingen en hoe die door alle

potentiële gebruikers kunnen worden benut zodanig dat er maximaal profijt van gemaakt kan worden. Categorische (ja/nee) verwachtingen hebben daarentegen zeer grote nadelen (e.g. Murphy, 1991). Ze leiden in het algemeen tot overforecasting (hier niet!! zie hoofdstuk IV.1), hebben weinig of geen nut voor gebruikers, en leiden tot *hedging* bij de forecasters, i.e. een verwachting uitgeven die niet overeenkomt met wat werkelijk gedacht wordt. Dit laatste helpt het leerproces van forecasters om zeep en hun eventuele toegevoegde waarde t.o.v. modeloutput.

Bij eventuele misverstanden over of misinterpretatie van kansverwachtingen moet goed gekeken worden waar deze verwarring vandaan komt. Het blijkt dat in de meeste gevallen de event definitie de meeste ellende veroorzaakt en niet het niet kunnen omgaan met kansen. Het publiek blijkt juist het uitdrukken van de onzekerheid te prefereren boven de suggestie van absolute zekerheid die bij een deterministische (of categorische) verwachting verondersteld wordt, en prefereert zelfs een kans uitgedrukt in getalvorm boven die in kwalitatieve termen (Murphy, 1977 en Murphy et al., 1980).

In het algemeen kan gesteld worden dat overal waar onzekerheid bestaat over de gebruikte informatie of over modelverwachtingen of conclusies deze onzekerheid gekwantificeerd meegenomen moet worden. (zelfs de onzekerheid in de waarnemingen moet "in kansen" meegenomen worden, zie hoofdstuk IV).

III.2e. *variabele drempels*

Zoals gezegd hangt de vertaling van weer naar impact sterk af van de tijd en plaats. Dit is deels te voorzien en dan is het handig om dit vast te leggen in de set drempels (zie II.2.a). Deels kan het incidenteel zijn (open-lucht feesten, Koninginnedag, etc.) en ook onverwachts. Op zulke momenten zouden de drempelwaarden (de meteorologische en / of de kansdrempels) bijvoorbeeld lager moeten zijn dan anders. Ook daarvoor zou het KNMI (basis) gevaarlijk weer verwachtingen moeten hebben die daarbij kunnen helpen. Het initiatief tot het inwinnen van de meteo info en wat er mee gedaan wordt ligt bij de gebruikers. (zie ook f en h). Door verwachtingen te maken voor verschillende drempels kunnen meer gebruikers flexibel bediend worden en bovendien niet alleen in de afgesproken standaard situaties.

Ook kan gedacht worden aan de mogelijkheid voor de gebruiker om zelf interactief de grenzen tijdelijk aan te passen. Incidenteel kan dan bijvoorbeeld een drempel lager gelegd worden, waarna bij overschrijding automatisch gewaarschuwd wordt. Later wordt de drempel door de gebruiker weer teruggezet.

III.2f. *communicatie naar buiten*

Het KNMI levert de kansverwachtingen voor de afgesproken fenomenen. Bij kansen groter dan de afgesproken kansdrempel wordt de "geabonneerde" gebruiker(sgroep), liefst **automatisch**, ingelicht. Hierbij wordt uiteraard gespecificeerd welke grootte, kracht, intensiteit, duur, gebied, forecasttijd, etc. het betreft. Maar ook de verwachte kans van optreden wordt altijd vermeld. Het is essentieel dat ook voor weeralarmen de kansen gecommuniceerd worden naar "het publiek". Zoals eerder gezegd, kan het voorkomen dat tegelijkertijd meerdere drempels overschreden worden.

Het besluit om het publiek te waarschuwen of om andere instanties in te lichten op grond van de uitgegeven verwachting ligt bij voorkeur bij de gebruiker(sgroep) zelf. Deze ligt in ieder geval niet bij de forecaster. Deze is op zijn hoogst achter de hand om inlichting te verschaffen over de meteorologie.

Alle uitstaande gevaarlijk weer verwachtingen zijn in principe allemaal zichtbaar voor de gebruikers (lieftst bij het hele publiek). Op deze manier kunnen ze ook snel informatie inwinnen in het geval dat door een of andere incidentele reden de drempelwaarden anders liggen dan vooraf gedacht. Een voorbeeld hiervan kan zijn dat de impact op de samenleving op een bepaalde dag veel groter kan zijn dan normaal (bijv. op Koninginnedag). In zo'n situatie kan de gebruiker dan weer eventueel zorg dragen voor waarschuwing in de media. Voorwaarde hiervoor is wel dat er ook voor een aantal "minder gevaarlijke criteria" verwachtingen beschikbaar zijn. Dit is ook om andere redenen noodzakelijk (zie hoofdstuk IV.2).

III.2g. *de rol van de meteoroloog*

Hierbij moet onderscheid gemaakt worden tussen de rol van de meteoroloog als forecaster (on-line) en die als voorlichter (vooral off-line). De on-line taken spelen zich uiteraard uitsluitend af na weersafhankelijke inzet.

De rol van **de meteoroloog als forecaster** is beperkt tot de puur meteorologische kant. En hij / zij speelt dus alleen (eventueel) een rol bij de eerste stap in de figuur (Fig. III.2), i.e. bij de vertaling van de uitvoer van de modellen en eventueel gebruik makend van recente waarnemingen naar kansuitspraken over de betreffende fenomenen. De forecaster is dus niet direct betrokken bij het naar buiten brengen van de waarschuwing. Hij / zij bemoeit zich uitsluitend met de meteorologische kant, en niet met de "problemen" van de afnemer ("het is al de derde keer deze week", "er is nu wel erg veel (of juist erg weinig) volk op straat", etc). Deze afweging is in principe vooraf in allerlei afspraken vastgelegd (zie III.2a). Ook hier is er een analogie met het alertsysteem voor de waterschappen (zie III.1).

Vooralsnog zullen veel modelverwachtingen via de forecaster lopen (zie figuur III.2) om getransformeerd te worden tot een definitieve (kans)verwachting. De forecaster kan zijn vertrouwen in de modelverwachting gebruiken om, óf de deterministische verwachting te voorzien van een kans, óf de door het model geleverde kansen aan te passen. Hierin zit dus een groot subjectief element. Juist hierin ligt de (waarschijnlijk) toegevoegde waarde van de forecaster. Door de focus van het KNMI op gevaarlijk weer is dit een van zijn belangrijkste taken. Om de uitgegeven kansverwachtingen van waarde te kunnen laten zijn is het echter noodzakelijk dat deze reliable zijn. Dit vergt veel kennis over kansen.

Herhaaldelijk is vastgesteld dat forecasters met enige skill "subjectieve" kansverwachtingen kunnen maken voor niet-gevaarlijk weer elementen (e.g. Daan and Murphy, 1982). Voor gevaarlijk weer is dit vooralsnog geen onomstreden zaak. Maar zolang de modellen nog geen "goede" kansverwachtingen kunnen maken voor gevaarlijk weer ligt hier dus een essentiële rol voor de forecaster. Verwacht mag worden dat dit voor de forecasttermijn tot ca 12 uur vooruit daarom voorlopig de belangrijkste taak van de KNMI-forecaster zal zijn. De kansinformatie verkregen uit korte termijn ensemble systems zal hierbij steeds belangrijker worden.

De rol van de **meteoroloog als voorlichter** speelt zich voornamelijk af off-line bij de totstandkoming van de afspraken over de gevaarlijk weer elementen, de exacte definities hiervan en eventueel bij de ondersteuning hoe optimaal om te gaan met kansinformatie (i.e. de "cost-loss analyse"). Ook dit vergt nogal wat kennis op het gebied van kansverwachtingen. Verwacht mag worden dat gebruikers voor de eventuele ondersteuning voor het optimaal gebruik maken van kansinformatie steeds vaker buiten het KNMI zullen zoeken.

III.2h. *semi-continue verwachtingen*

Zodra de gevaarlijk weer forecaster opgeroepen is (weersafhankelijke inzet) wordt het weer permanent (extra) bewaakt en worden zodra nieuwe informatie daartoe aanleiding geeft, nieuwe verwachtingen uitgegeven en vastgelegd. Dit kan betekenen dat 12 uur van te voren een hoge kans op optreden van iets "gevaarlijks" gegeven wordt die dus automatisch naar de gebruiker gaat, en enkele uren later de kans veel lager geschat wordt. Bijv. door nieuwe informatie of door een andere forecaster op dienst. De gebruiker is dus semi-continue voorzien van alle verwachtingen. In de voorwaarschuwingsfase kan iets soortgelijks van kracht zijn.

Het kan dus best voorkomen dat uitgegeven waarschuwingen (ter grootte van een weeralarm bijvoorbeeld) later weer ingetrokken worden. Dit is inherent aan de beperkte voorspelbaarheid van het (extreme) weer. Ook is het zo dat niet iedere forecaster op grond van dezelfde informatie tot dezelfde kansinschatting komt. Ook dit is niet erg. Het maakt onderdeel uit van het leerproces. En het blijft onderdeel van iedere verwachtingssystematiek zolang de verwachtingen nog niet rechtstreeks geleverd worden door volledig objectieve systemen. Deze onzekerheid van de inschatting van forecasters mag niet misbruikt worden om de uiteindelijke kansuitspraak van de forecaster te beïnvloeden. Dit gevaar dreigt als de ingeschatte kans juist gelijk of vlak in de buurt ligt van een kansdrempel waarboven gewaarschuwd zou moeten worden. Juist dan is er het gevaar om allerlei gebruikersoverwegingen in het oordeel te betrekken. Dit is nou juist een van de redenen waarom de verwachtingen niet zullen verbeteren in de huidige systematiek.

III.2i. *presentatie*

De uitstaande cq. uitgegeven kansverwachtingen voor alle gevaarlijk weer elementen voor alle opgegeven drempelwaarden, alsmede voor enkele minder extreme drempelwaarden (zie paragrafen III.2.f en IV.2) moeten te allen tijde beschikbaar zijn. Niet alleen in bulletinvorm maar in ieder geval ook in grafische vorm (kanskaartjes). Deze moet zichtbaar zijn voor alle directe afnemers (abonnees), bijvoorbeeld via internet. Dit is noodzakelijk omdat zowel het meteorologisch criterium als de cost-loss verhouding van speciale omstandigheden kan afhangen (zie III.2.f), waardoor tijdelijk andere drempels van belang zijn. De grafische presentatie moet liefst voor iedereen beschikbaar zijn. Op deze manier groeit de bekendheid met en de acceptatie van kansverwachtingen in het algemeen en die voor extremen in het bijzonder.

Deze kanskaarten bevatten duidelijke info over begin- en eindtijd waarop de kans betrekking heeft, de grootte van het gebied, etc. en moeten zodra de inschatting van het fenomeen verandert aangepast worden. In de presentatie staan ook de gebieden aangegeven waar al een eventsituatie heerst.

Er kan zelfs aan gedacht worden dat op grond van de kansen in de kanskaartjes en de individuele drempelwaarden van de "aangesloten" gebruikers(groepen) binnen het betreffende gebied automatisch een waarschuwing naar hen gestuurd wordt. Deze waarschuwing zou dus getriggered worden door software losgelaten op de kaartjes. Dergelijke systemen bestaan al. Verdere speculatie over de presentatie valt buiten de scope van het verhaal.

III.2j. *verificatie*

Gevaarlijk weer komt niet al te vaak voor. Die zeldzaamheid zorgt ervoor dat we per definitie niet veel "soortgelijke" gevallen tegenkomen binnen een bepaalde verificatieperiode, en verificatie zal dus niet snel tot conclusies kunnen leiden. Om toch het een en ander te kunnen leren en om de kwaliteit van de gevaarlijk weer verwachtingen in te kunnen schatten dient de manier waarop de verwachtingen geverifieerd worden grondig te veranderen. Hier zal in het volgende hoofdstuk nader op ingegaan worden.

IV. Hoe kan verificatie bijdragen aan verbetering

IV.1 Problemen met de huidige verificatie

Sinds 1999 worden er weeralarmen uitgegeven. Weeralarmsituaties, verdeeld over 5 fenomenen, zijn in de afgelopen 6 jaar ca 10 keer per jaar voorgekomen. In deze paragraaf wordt een zeer summiere analyse van de verwachtingen voor deze gevallen gegeven. Bij de interpretatie van de onderstaande "verificatie" moet, naast het geringe aantal gevallen, noodgedwongen rekening gehouden worden met de volgende problemen:

- * niet alle predictanddefinities kloppen en / of ze zijn gewijzigd gedurende de verificatieperiode
- * er is niet altijd zeker of een weeralarmsituatie is opgetreden of niet
- * de forecasttijd is niet meegenomen

Desondanks legt een voorzichtige "verificatie" een aantal grote tekortkomingen van de weeralarm-verwachtingen bloot.

Hieronder staan voor de 5 meteorologische grootheden waar weeralarmen voor gegeven kunnen / konden worden de "verificatie" tabellen. Deze zijn ontleend aan een recente notitie van Robert Mureau. Horizontaal staan de verwachtingen (FC) met y = yes en n = no, verticaal de waarnemingen (OBS). Doordat niet altijd is vast te stellen of een weeralarmsituatie is opgetreden of niet, bijv. omdat het waarneemnetwerk niet dicht genoeg is voor bepaalde fenomenen, zou het werkelijke aantal opgetreden gevallen nog iets hoger kunnen liggen dan hier vermeld. Dit zou de resultaten nog een graadje slechter maken.

grootheid:	zware sneeuwval		zware windstoten/storm		zwaar onweer		overvloedige neerslag		ijzel		
	aantal:	10	13	19	9	10					
\	FC	y	n	y	n	y	n	y	n	y	n
O	y	2	6	9	2	3	16	1	8	6	2
B	n	2		2		0		0		2	
S											

In totaal (alle grootheden bij elkaar):

	y	n	
y	21	34	55
n	6		
	<hr/>		
	27		

Wat opvalt is de geringe probability of detection (POD): nog geen 40% (21 van de 55) van de weeralarmsituaties wordt als zodanig onderkend. Ook de underforecasting met een factor 2 valt op (55 opgetreden weeralarmsituaties, 27 keer er een voorspeld). Men wacht blijkbaar tot het moment dat men vrijwel zeker is. De false alarm ratio is dan ook laag. Deze slechte resultaten hebben vooral te maken met de in hoofdstuk II.3 gesignaleerde problemen en waarschijnlijk niet of veel minder met gebrekkige skill van de forecaster of van de modellen. Waar de oorzaken liggen is in de huidige verwachtingen systematiek niet te achterhalen (zie hoofdstuk II.3). Of er over de jaren een verbetering opgetreden is of niet, valt op deze manier uiteraard ook moeilijk vast te stellen. Hetzelfde geldt voor de toekomst; op de huidige manier duurt het waarschijnlijk ettelijke jaren voordat een eventuele trend significant zou kunnen worden aangetoond.

Tot nog toe is er alleen wat betreft de weeralarmen een voorzichtige poging ondernomen om te verifiëren. Maar verder dan het turven van de (veronderstelde) hits, missers en false alarms is het niet gekomen. Het aantal uren vooruit dat gewaarschuwd is, dat toch in belangrijke mate de waarde van de verwachting bepaalt, wordt niet meegenomen. Andere gevaarlijk weer verwachtingen worden in het geheel niet systematisch geverifieerd.

IV.2 Voorstel nieuwe verificatie

Een van de belangrijkste doelen van verificatie is om er dingen uit te leren die aangewend kunnen worden om de verwachtingen op termijn te kunnen verbeteren dan wel om tekortkomingen te signaleren. De huidige systematiek van (m.n) de weeralarmen is volstrekt onvoldoende om dit doel te verwezenlijken. Bijsturing van de (meteorologische) verwachtingsmethode van de forecaster is niet of nauwelijks mogelijk op grond van de huidige verificatie. In deze paragraaf wordt een viertal oorzaken hiervan aangegeven tesamen met een aantal suggesties ter verbetering.

A. Alles in kansverwachtingen

Afgezien van het feit dat het geven van categorische verwachtingen veel minder waarde heeft voor gebruikers (zie III.2d) belemmert het ook de mogelijkheden om iets te leren uit de verwachtingen. Alle gevaarlijk weerverwachtingen en ook die voor weeralarmen dienen te worden gegeven in de vorm van kansverwachtingen in plaats van categorische (ja/nee) uitspraken. Zelfs als besloten wordt om alleen boven een bepaalde kansdrempel een weeralarm in categorische vorm uit te brengen dan moet toch met name de onderliggende kansverwachting geverifieerd worden. Onderdeel hiervan kan dan zijn om specifiek naar de resultaten te kijken in de buurt van (en boven) de drempelwaarden.

B. Alle deelprocessen apart monitoren en verifiëren

Alle in de voorgestelde aanpak geïdentificeerde stappen (zie Fig. III.2) moeten apart vastgelegd en geverifieerd worden. Dit geldt voor alle gebruikte informatie uit alle modellen, de eventuele aanpassingen door de meteoroloog, alle waarnemingen die betrokken zijn bij het onderzoek of een criterium gehaald is of niet (zie D. hieronder), en de uiteindelijke, al of niet uitgegeven, kansverwachtingen. Focus van de verificatie kan liggen in de buurt of boven de opgegeven kansdrempels, maar niet alleen (zie C. hieronder). Tenslotte moet uiteraard de forecasttijd expliciet meegenomen worden in de verwachtingen en de verificatie.

De verificatie van alle deterministische en probabilistische verwachtingen van alle modellen en de forecaster gebeurt op de gebruikelijke manier. Bestudering van de verificatiecijfers kan eventueel aanleiding geven tot aanpassing van de methodieken, verbetering in de postprocessingmodules, verbetering van de skill van de forecaster, etc.

Daarnaast kan apart geëvalueerd worden hoe een al of niet gegeven "alarm" (als de verwachte kans hoger was dan de opgegeven drempel) of (voor)waarschuwing overgekomen is, i.e. hoe het ervaren is door de gebruiker. Dit kan aanleiding geven tot een aanpassing van de definitie van de predictand of van de kritieke kansdrempel (volgens de procedure beschreven in III.2.a), of tot andere communicatie afspraken.

Bovenstaande verificatie en de evaluatie van de (voor)waarschuwingen (alarmen) mogen niet door elkaar heen lopen. De laatste zal niet leiden tot verbetering van de verwachtingen. En mag de forecaster dus ook niet beïnvloeden. Deze moet zich beperken (on-line) tot de puur meteorologische zaken.

Zaken als "wel false alarm maar toch iedereen gelukkig" hebben niks met de skill van de forecast te maken, en zou als aanmoediging opgevat kunnen worden om bij lagere (meteorologische) drempelwaarden een alarm uit te doen gaan. Maar het zou hoogstens aanleiding mogen zijn om (na overleg) de afgesproken grenzen bij te stellen of om de kansdrempel anders te kiezen.

C. Uitbreiding van de verificatie set

Bij verificatie moet altijd onderscheid gemaakt worden tussen de feitelijke *verificatie* in getalvorm en de *interpretatie* van de verificatiecijfers. Bij dit laatste moet ook de *significantie* van de verificatie scores betrokken worden. Dit nou is juist een groot probleem bij verificatie van gevaarlijk weer verwachtingen en weeralarms.

Verificatie van verwachtingen voor extreme weer (weeralarmen) is per definitie iets van lange adem. Het komt veel te weinig voor om significante uitspraken te kunnen doen over de skill. Weeralarm situaties treden (per definitie) maximaal ca 10 maal per jaar op, verdeeld op dit moment over een vijftal verschillende meteorologische fenomenen (storm / windstoten, onweer, sneeuw, ijzel en (tot voor kort) neerslag).

Tel daarbij de steeds veranderende inzichten (en dus criteria) over welke items nu precies verwacht moeten worden, en de vele veranderingen in modellen en waarneemtechnieken de afgelopen jaren en in de toekomst, dan is het duidelijk dat (standaard)verificatie van extreem weerverwachtingen niet snel (if ever) iets significant oplevert.

Er zijn 2 manieren om iets verder te komen met de verificatie. Allereerst levert de verificatie van kansverwachtingen al iets sneller bruikbare resultaten dan die van categorische. Bovendien kan de set verwachtingen uitgebreid worden door vanaf het moment dat 'de gevaarlijk weer meteoroloog' opgeroepen is (weersafhankelijke inzet) semi-continue (zie III.2h) verwachtingen te maken voor het betreffende fenomeen. Het betreft dan weliswaar nog steeds maar "1 event" maar er kan dan meer ervaring opgedaan worden. In de verificatie moet ook de voorspeltermijn meegenomen worden.

Bovenstaande levert nog steeds niet genoeg gevallen op om profijt te kunnen hebben van verificatie. Een tweede, veel belangrijkere, methode is om in de verificatie niet (alleen) naar de kansverwachtingen te kijken van het gevaarlijk weer of naar die bij weeralarmen, maar ook "soortgelijke" verwachtingen te betrekken voor iets minder extreem weer. Bij "soortgelijk" kan gedacht worden aan gebiedskansen voor het zelfde fenomeen maar voor lagere drempelwaarden of aan kansen berekend op grotere gebieden. Bijvoorbeeld, als "minstens windkracht 11" een gevaarlijk weer limiet zou zijn, dan zou niet alleen de kans daarop berekend moeten worden maar ook de kans op Bf 10 en Bf 9. (Dit is ook om andere redenen nodig, zie III.2e). Dit heeft het voordeel dat het aantal gevallen veel groter is, en dus ook de significantie van de verificatie scores. Verbetering van de scores op drempelwaarden (net) onder de gevaarlijk weer drempels zegt hoogstwaarschijnlijk ook iets over de skill bij hogere drempels. Dit geldt zowel voor de forecasters als voor de modellen. Bovendien zouden resultaten van modelverificatie bij lagere drempels kunnen leiden tot (statistische) methoden om iets te zeggen over hogere drempels. Een ander voordeel is dat de forecaster meer ervaring op kan doen in het verwachten van (gebieds)kansen en (o.a.) eventuele systematische afwijkingen t.o.v. reliability kan corrigeren. Ervaring die toegepast kan worden bij extremere situaties.

D. Onzekerheid in de waarnemingen meenemen

Omdat de gevaarlijk weer verwachtingen altijd gelden voor gebieden (bijv. "op de schaal van een provincie") in plaats van voor gedefinieerde stations, is het niet altijd duidelijk of een criterium gehaald is of niet. Dit is uiteraard afhankelijk van de predictand. Voor de meeste predictands zijn we afhankelijk van stationswaarnemingen. Dit geldt niet voor onweer en (in mindere mate) voor neerslag. Voor deze grootheden is het redelijk zeker vast te stellen of aan een (gebieds)criterium is voldaan. Bij onweer wordt iedere ontlading precies gelocaliseerd; waargenomen neerslag met behulp van radar heeft een goede ruimtelijke dekking maar heeft nog fikse calibratieproblemen. Voor andere predictands (bijv. windstoten), zijn we voor de "detectie" van het fenomeen afhankelijk van een klein aantal stationswaarnemingen. Hier heb je problemen met de representativiteit van de waarnemingen maar vooral met de geringe ruimtelijke dekking. Als het niet op een waarneemstation is waargenomen wil dat niet

zeggen dat het event niet is opgetreden binnen het gegeven gebied. De kans hierop is te kwantificeren, tot troost voor de forecasters¹.

Voor de verificatie is het van cruciaal belang om ook de onzekerheid in de detectie mee te nemen. Dit leidt tot veel "eerlijker" scores en waar eerder iets uit te leren valt. Stel je voor dat een weeralarm event door te weinig waarnemingen ten onrechte geklassificeerd wordt als "niet opgetreden" (of andersom) dan heeft dit grote gevolgen voor de verificatie cijfers, te meer omdat er al zo weinig gevallen zijn. Voor de *interpretatie* van deze scores is het verder van belang zich te realiseren dat de inherente onzekerheid in de waarnemingen sterke invloed heeft op de maximaal haalbare scores.

Per geval moet er dus een schatting gemaakt worden of het event opgetreden is of niet. Deze schatting kan weer uitgedrukt worden in een kans. De schatting hangt vooral af van de schaal van het fenomeen in relatie tot het waarneemnet. Hier zijn theoretische bespiegelingen voor te bedenken (maar vallen buiten het doel van dit rapport). Een schematische voorstelling is gegeven in onderstaande figuur. Vooral bij het niet waargenomen zijn is het vaak onzeker of het niet ergens anders is opgetreden. Maar ook eventuele twijfel over de waarnemingen zelf kan tot uitdrukking gebracht worden. De nu gebruikelijke 2 x 2 contingentietabel waarin ja/nee verwachtingen uitgezet worden tegen ja/nee opgetreden (zoals in IV.1), wordt vervangen door een scatterplot waarin de voorspelde kans uitgezet wordt tegen de kans dat het fenomeen inderdaad is opgetreden. Een voorbeeld hoe dat er uit zou kunnen zien staat hieronder.

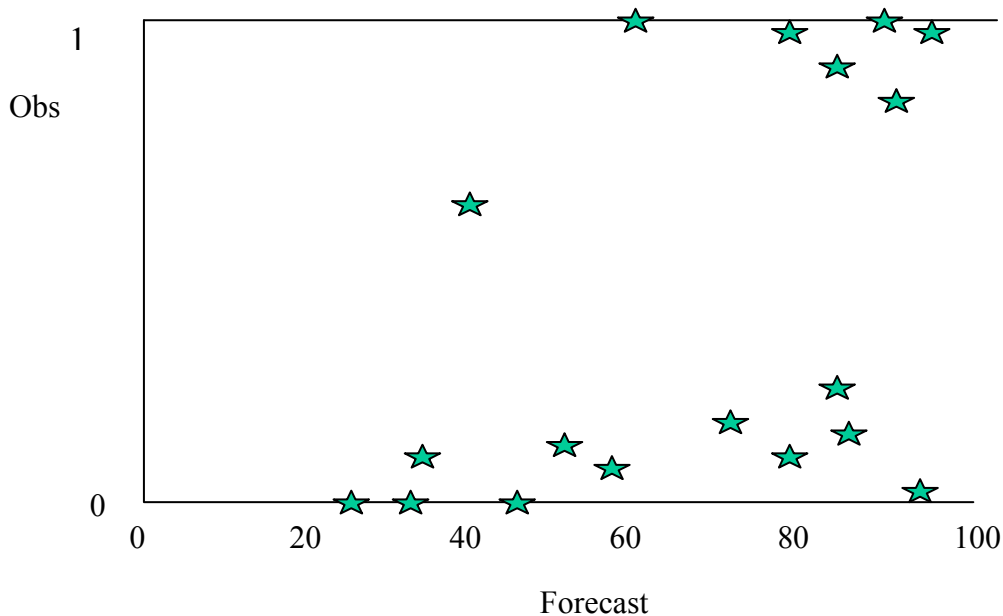


Fig. IV.1. Verificatie diagram. De forecasts zijn gegeven in procenten. Verticaal staat de geschatte kans (tussen 0 en 1) dat een event daadwerkelijk is opgetreden. Zie tekst voor verdere uitleg.

¹ "Bij het uitgeven van een alarm voor zeer zware windstoten geldt een criterium van 56 kts. Wanneer vervolgens de gemeten windstoten niet boven de 55 kts uitkomen wordt dat gevoeld als misser. Wanneer echter de 57 kts op enkele plaatsen gehaald worden, wordt dit beleefd als een prachtig succes." (Nolet, 2005)

V. Slotopmerkingen

De huidige systematiek die gehanteerd wordt om te komen tot het al of niet uitgeven van een weeralarm schiet zwaar tekort. De skill van de weeralarmverwachtingen is zeer slecht. Ruim 60% van alle weeralarmsituaties wordt in het geheel gemist. Bovendien worden de gebruikers zwaar tekort gedaan, niet alleen door de slechte skill maar ook door het niet tijdig waarschuwen en door ze (al of niet afgesproken) veel informatie te onthouden. De enige reden dat het KNMI deze verwachtingen desondanks nog mag maken lijkt te zijn dat het als "gevaarlijk weer instituut" (vooral nog) een monopoliepositie op dit soort verwachtingen heeft. Lukt het niet om de skill drastisch te verbeteren, dan zullen we de monopoliepositie verliezen, zoals dat ook gebeurd is met "de Noordzee verwachtingen" n.a.v. de K13-proef in 1991.

De redenen van de slechte weeralarmverwachtingen zijn velerlei. Zo is er ondanks 6 jaar praktijk nog steeds grote onduidelijkheid (bij publiek én forecasters) over de exacte definitie van de situaties waarvoor een weeralarm uitgegeven kan worden, en welke factoren hierbij nou precies een rol mogen spelen. Maar belangrijker is dat in het verwachtingsproces een vermenging plaats vindt van de meteorologie en de impact die het verwachte weer op "het" publiek zal hebben. Het meenemen van de overwegingen van de gebruiker hoe die de verwachting zal gebruiken of er op zal reageren schaadt de kwaliteit van de verwachtingen (en dus de uiteindelijke belangen van de gebruiker) en staat toekomstige verbetering van de verwachtingen in de weg. Dus ook hier gaat de "klantgerichtheid" veel te ver. Tenslotte leidt de huidige systematiek er toe dat er pas gewaarschuwd gaat worden als er (bijna) absolute zekerheid is (dus vrijwel altijd te laat), en dat er geen enkele incentive is om verwachtingen kwalitatief beter te maken, eerder het tegendeel.

De reden van de slechte kwaliteit ligt (hoogstwaarschijnlijk) niet in het feit dat de forecasters zulke slechte gevaarlijk weer verwachtingen zouden maken. Maar het ligt vooral in het feit dat ze de verwachte impact van het verwachte weer (moeten) meewegen in hun beoordeling of er een alarm uitgegeven moet worden of niet. Deze impact assessment is geen taak van het KNMI of zou het niet moeten zijn. Het KNMI zou zich in zijn berichtgeving moeten beperken tot puur meteorologische zaken; gebruikers zijn zeer wel in staat om op grond van een verwachting een inschatting te maken van de impact. *"The public-service function must be separated from the weather-forecasting service"*. Voor zover de berichtgeving over de mogelijke impact gedelegeerd is aan het KNMI dient die volkomen los te staan van de meteo. Wel kan off-line eventueel advies gegeven worden aan gebruikers hoe om te gaan met de verwachtingen.

In dit rapport wordt een totaal andere systematiek voorgesteld voor het maken van gevaarlijk weer verwachtingen in het algemeen en weeralarmen in het bijzonder. Een groot aantal aanbevelingen voor veranderingen t.o.v. de huidige situatie is beschreven. Naast het veel strakker definiëren van de te voorspellen grootheden wordt er een strikte scheiding voorgesteld van meteo (on-line) en impact (off-line). Dit leidt tot een totaal andere rol van de meteoroloog en misschien wel van het KNMI. Het leidt er ook toe dat er veel zinniger gebruik gemaakt wordt van de kennis van de meteorologen, dat er veel meer waarde is voor veel meer gebruikers (e.g. Katz and Murphy, 1997), en dat er een mogelijkheid geboden wordt om uiteindelijk de skill van

de verwachtingen op te schroeven. Een noodzakelijkheid voor een gevaarlijk weer instituut.

Wil er meer geleerd kunnen worden van de weeralarmverwachtingen dan kan er niet alleen gekeken worden naar de verwachte en opgetreden weeralarmsituaties. Deze zijn per definitie niet erg frequent en ook nog eens zeer divers. Daardoor valt uit verificatie hiervan per definitie niks of zeer weinig te leren. Dit rapport biedt een uitweg voor dit probleem middels een voorstel om de verificatie anders aan te pakken. Het vastleggen en monitoren van alle afzonderlijke stappen in het verwachtingsproces maakt hier deel van uit. Alleen dan kan geanalyseerd worden welke onderdelen bijsturing behoeven.

Onderdeel van de voorgestelde systematiek is ook om alle verwachtingen te geven in kansen. Het weer is "stochastisch" en zeker het gevaarlijke weer. Dit is een gegeven en daar kan dus maar beter gebruik van gemaakt worden. De maatschappij is zeer wel in staat om om te gaan met kansen zoals blijkt uit talloze studies. "Nu de forecaster nog", zoals veel valt te horen. Deze heeft in het verleden aangetoond dat zijn / haar kansinschatting skill heeft voor het niet-gevaarlijke weer. Om dit ook voor het "gevaarlijke" weer te bereiken is dringend meer ervaring en kennis nodig bij de forecasters. Zolang modellen (inclusief de statistische) nog niet rechtstreeks kansen geven voor gevaarlijk weerfenomenen is de kansinschatting van de forecaster cruciaal. Maar de toegevoegde waarde hiervan is nog lang niet bewezen.

"De gebruiker wil geen kansen", is een ongefundeerd statement. Natuurlijk wil de gebruiker het liefst zekerheid maar gegeven de onontkoombare onzekerheid van het weer wil ie zo goed mogelijk geïnformeerd worden over die onzekerheid. Die biedt juist, mits skilful, de toegevoegde waarde en mag hem dus niet onthouden worden. De mate van onzekerheid kan door iedere afzonderlijke gebruiker gebruikt worden in zijn of haar beslissingsondersteuning. Het KNMI kan eventueel toelichting geven op de verwachting, maar zal in steeds mindere mate betrokken worden bij verdere beslissingsondersteuning. Anderen hebben veel meer expertise op dat gebied. Wij moeten ons beperken tot levering van meteorologische informatie (in brede zin) van zo hoog mogelijke kwaliteit.

Het stochastische karakter van het weer houdt ook in dat in de verwachtingen altijd "missers" (i.e. lage kans voorspeld en toch gevaarlijk weer opgetreden) en "false alarms" (hoge kans en toch geen gevaarlijk weer) zullen voorkomen. Dit is inherent aan de beperkte voorspelbaarheid. Door ook de kansen te communiceren naar de gebruikers (het publiek) zijn missers en false alarms prima te begrijpen en zullen makkelijker geaccepteerd worden.

Willen we de maatschappij beter bedienen met gevaarlijk weer verwachtingen, dan moet er maatwerk geboden worden in de vorm van kansen voor verschillende gebruikers. Niet in de zin dat er voor iedere gebruiker een apart verwachtingssysteem ontwikkeld moet worden. Dit is niet zinnig omdat de wensen snel zullen veranderen (zie de weeralarm-criteria). Beter is het om veel meer onderzoek te doen naar de "basis-ingrediënten" die nodig zijn voor informatie over gevaarlijk weer en voor de uiteindelijke verwachtingen: de kennis van gevaarlijk weer, de voorspelbaarheid ervan, betere modellen, short range ensembles, etc.

Door de voorgestelde strikte scheiding tussen meteo en wat de maatschappij er mee doet (al of niet waarschuwen, al of niet andere acties ondernemen) kan het KNMI weer afgerekend worden op puur de meteo (zoals het hoort), i.e. op de skill van de verwachting en de tijdigheid van uitgifte, en dus niet meer op de verwachting van het effect van het verwachte weer op een gemiddelde gebruikersgroep.

Dankbetuiging

Veel dank is verschuldigd aan Ben Wichers Schreur, Jeanette Onvlee, Daan Vogelezang en Robert Mureau voor de vele discussies over dit onderwerp in de afgelopen jaren. Tevens worden Maurice Schmeits, Sander Tijm, Seijo Kruizinga, Rinske Krabbe en Gertie Geertsema bedankt voor het kritisch doorlezen van een eerdere versie van dit rapport en voor hun suggesties.

Literatuur

- Daan, H. and A. H. Murphy, 1982. Subjective probability forecasting in The Netherlands: some operational and experimental results. *Meteorol. Rdsch.*, **35**, 99-112.
- Katz, R. W. and A. H. Murphy, 1997. *Economic value of weather and climate forecasts*. Cambridge University Press, 222pp.
- Kees Kok, Daan Vogelezang, Ben Wichers Schreur en Iwan Holleman, 2005. Beschrijving van het automatisch waarschuwingssysteem voor extreme neerslag-hoeveelheden t.b.v. de waterschappen. KNMI Technisch Rapport. In voor-bereiding.
- Murphy, A. H., 1977. On the misinterpretation of precipitation probability forecasts. *Bull. Am. Meteorol. Soc.*, **58**, 1297-1299.
- Murphy, A. H., S. Lichtenstein, B. Fischhoff and R. L. Winkler, 1980. Misinterpretations of precipitation probability forecasts. *Bull. Amer. Meteorol. Soc.*, **61**, 695-701.
- Murphy, A. H., 1991. Probabilities, odds and forecasts of rare events. *Wea. Forecasting*, **6**, 302-307.
- Marco Nolet, 2005. Weeralarm 8 januari 2005. Bijlage Meteorologische evaluatie.
- Thompson, J. C., 1952. On the operational deficiencies in categorical weather forecasts. *Bull. Amer. Meteor. Soc.*, **33**, 223-226.

Uitgegeven WM-memoranda

2003

- WM 03-01 Vertical profiles in the stable boundary layer
Wim de Rooy
- WM 03-02 Conjoint MET divisional meeting, Montreal September 2002
Wil van Dijk
- WM 03-03 25e bijeenkomst COST TC on Meteorology
Lissabon, 7-8 november 2002
Dick Blaauboer
- WM 03-04 ?
Radjesh Oemraw
- WM 03-05 26e bijeenkomst COST TC on Meteorology
Brussel, 26-27 mei 2003
Dick Blaauboer

2004

- WM 04-01 27e bijeenkomst COST TC on Meteorology
Exeter, 13-14 november 2003
Dick Blaauboer
- WM 04-02 Geslaagde signaalfunctie Hirlam bij extreme neerslag
van 12 – 13 augustus 2004 in Noord Nederland
Wim de Rooy, Kees Kok, Sander Tijm en Daan Vogelezang
- WM 04-03 28ste bijeenkomst van de COST Technical Committee
on Meteorology, Langen, 22-23 april 2004
Dick Blaauboer
- WM 04-04 29ste bijeenkomst van de COST Technical Committee
on Meteorology, Nice , 30 september – 1 oktober 2004
Dick Blaauboer

2005

- WM 05-01 Verificatieresultaten van het kansverwachtingssysteem
voor (zwaar) onweer over het warme halfjaar van 2004
Maurice Schmeits, Daan Vogelezang, Kees Kok
- WM 05-02 Naar een andere weeralarmsystematiek
Kees Kok
- WM 05-03 30ste Bijeenkomst van de COST Technical Committee
on Meteorology, Larnaca, 18 maart 2005
Dick Blaauboer