

Voorspelling van het dagverbruik in het voorzieningsgebied van de Berenplaat

Inleiding

Elk waterleidingbedrijf zal ernaar streven steeds over voldoende drinkwater te beschikken om aan de vraag te kunnen voldoen. Maar bij sommige productieprocessen wordt een directe reactie op snelle schommelingen in de vraag belemmerd door de zorg voor de kwaliteit van het af te leveren water. Die kwaliteit is namelijk gebaat bij een gelijkmatig verloop van het productieproces.



DRS. PAUL K. BAGGELAAR
KIWA NV, Nieuwegein
Afdeling Wiskunde
en Informatica

Het inspelen op het verbruik is beter te verwezenlijken als men de beschikking heeft over korte termijn voorspellingen van dat verbruik. In de watersector is de behoefte daaraan nog niet zo sterk als in de elektriciteits- of gassector, maar de eerste signalen zijn er. Zo heeft de Drinkwaterleiding Rotterdam (DWL) KIWA verzocht onderzoek te doen naar een voorspeller van het uurverbruik in het voorzieningsgebied van de Berenplaat, om een optimale regeling van het productieproces mogelijk te maken. Het onderzoek heeft geleid tot een tweetrapsvoorspeller, waarbij eerst het dagverbruik van de volgende dag wordt voorspeld en dit

Samenvatting

Dit artikel behandelt de ontwikkeling en evaluatie van een methode om het dagverbruik van drinkwater in het voorzieningsgebied de Berenplaat te voorspellen. Deze voorspellingen zijn nodig om het productieproces van drinkwater optimaal te kunnen regelen. Er is gebruik gemaakt van een Box-Jenkins model dat rekening houdt met doordeweekse feestdagen en de verdamping in het gebied. Het model genereert nauwkeurige voorspellingen en heeft zelfs na drie jaar nog een goede voorspelkracht, met een absolute voorspelfout die gemiddeld slechts 2,6% bedraagt.

vervolgens wordt opgesplitst in voorspellingen van het uurverbruik. In dit artikel wordt beschreven hoe de voorspeller van het dagverbruik is ontwikkeld aan de hand van een statistische analyse van het dagverbruik in het verleden.

In het voorgaande artikel [Urbanus en Biemans, deze H_2O] is beschreven hoe hieruit voorspellingen van het uurverbruik worden verkregen en hoe deze worden betrokken bij het productieproces van drinkwater op de Berenplaat. Voor een uitgebreide beschrijving wordt verwezen naar Baggelaar *et al.*, [1991].

Gebruikte methode

Omdat er onvoldoende bekend is over het mechanisme van het dagverbruik om daarvan een theoretisch model op te stellen, is gekozen voor de statistische (of empirische) benadering. De voorspelling wordt daarbij gegenereerd door een model dat is afgeleid aan de hand van een

analyse van gegevens van het dagverbruik en eventuele andere relevante variabelen. Een statistische benadering die in dit opzicht in veel sectoren al vruchten heeft afgeworpen is de Box-Jenkins methode [Box and Jenkins, 1976]. Zo wordt deze in de elektriciteits- en de gassector al geruime tijd met succes toegepast om voorspellingen van het verbruik te genereren.

Voor onze toepassing zijn twee soorten Box-Jenkins model relevant, te weten:

- het univariate model, dat een voorspelling van het verbruik genereert als een gewogen som van het verbruik op voorgaande dagen;
- het transfer model, dat een voorspelling van het verbruik genereert als de combinatie van een gewogen som van het verbruik op voorgaande dagen en gewogen sommen van de waarden van andere relevante variabelen.

Voor theoretische achtergronden van de Box-Jenkins methode wordt verwezen naar het standaardwerk Box and Jenkins [1976]. Voor een praktisch gerichte verhandeling kan worden verwezen naar McLeod [1983].

Beschikbare tijdreeks van het dagverbruik

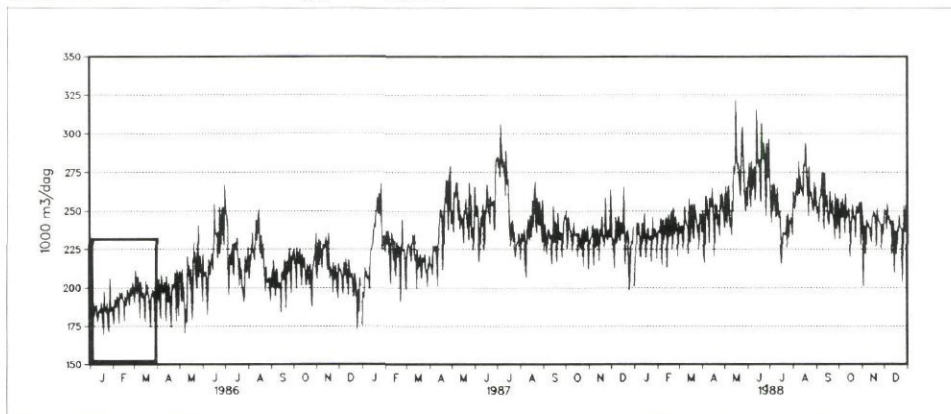
Om tot een voorspeller van het dagverbruik te komen kon worden uitgegaan van de tijdreeks van het dagverbruik van 1986 tot en met 1988 in het voorzieningsgebied van de Berenplaat. Deze is weergegeven in afb. 1. Op deze schaal kunnen in het patroon van het dagverbruik de volgende componenten worden onderscheiden:

- een jaarlijkse cyclus, met een hoger dagverbruik in de zomermaanden;
- een wekelijkse cyclus, met een lager dagverbruik in de weekeinden;
- enkele vrij abrupte en grote wijzigingen van het gemiddelde dagverbruik, vermoedelijk veroorzaakt door zogenaamde 'en-gros' leveringen aan grootverbruikers;
- de geleidelijke opwaartse trend van het dagverbruik.

Modelbouw

Een Box-Jenkins model wordt ontwikkeld volgens de identificatie – schatting – verificatiecyclus, die herhaald wordt doorlopen tot mag worden aangenomen dat het model voldoet aan de vooronderstellingen die eraan ten grondslag liggen. Zo moeten de modelresiduen, de verschillen tussen de werkelijke waarden en de waarden die het model geeft, onafhankelijk van elkaar zijn en afkomstig zijn uit een normale kansverdeling met gemiddelde nul. Dit laatste is vooral van belang met het oog op voorspellingen, omdat de toekomstige modelresiduen dan bij het genereren van voorspellingen op nul kunnen worden gezet. Belangrijke hulpmiddelen bij de identificatie van een univariaat model vormen het auto- en partiële autocorrelogram en bij de identificatie van een transfer model het kruiscorrelogram.

Afb. 1 - Tijdreeks van het dagverbruik in het voorzieningsgebied van de Berenplaat van 1986 tot en met 1988. Het omkaderde deel is vergroot weergegeven in afb. 2.



In afb. 2 is een gedeelte van deze tijdreeks op een grotere tijdschaal weergegeven (januari tot en met maart 1986).

De wekelijkse cyclus is nu nog beter waar te nemen. Het weekminimum valt meestal op zondag en bijna altijd in het weekeinde, wat veroorzaakt wordt door de geringere activiteiten in het weekeinde. Het weekmaximum valt daarentegen geregeld op maandag, wat moet worden toegeschreven aan de 'maandag wasdag' gewoonte.

Voorspelbaarheid van het dagverbruik

Inleiding

Alvorens een operationele voorspeller van het dagverbruik te ontwikkelen, is nagegaan in hoeverre het dagverbruik voorspeld kan worden met een tijdreeksmodel. Hiertoe zijn, met een model voor het dagverbruik in 1986, voorspellingen gegenereerd van het dagverbruik in 1987 en vervolgens geëvalueerd.

Tijdreeksmodel dagverbruik 1986

Het dagverbruik in 1986 kan worden beschreven als functie van voorgaande dagverbruiken en het optreden van doordeweekse feestdagen, met het volgende transfer model:

$$\nabla \nabla_7 Z_t = -26,5 \nabla \nabla_7 I_t + (1 - 0,45 B) (1 - 0,95 B^7) a_t \quad (1)$$

met Z het verbruik ($1.000 \text{ m}^3/\text{d}$), I het optreden van een doordeweekse feestdag (met de waarde 0 of 1), a het modelresidu ($1.000 \text{ m}^3/\text{d}$), t de index voor de dag, ∇ de differentie-operator ($\nabla Z_t = Z_t - Z_{t-1}$), ∇_7 de seizoens differentie-operator ($\nabla_7 Z_t = Z_t - Z_{t-7}$), B de verschuivings-operator ($B a_t = a_{t-1}$) en B^7 de seizoens verschuivings-operator ($B^7 a_t = a_{t-7}$).

De gebruikte Box-Jenkins notatie komt op het eerste gezicht afschrikwekkend over, maar heeft als groot voordeel dat

gecompliceerde modellen zeer compact kunnen worden weergegeven. Zo komt bovenstaand model erop neer dat het verbruik op een bepaalde dag beschreven kan worden aan de hand van het dagverbruik, het optreden van doordeweekse feestdagen en het modelresidu, alle tot 8 dagen terug. De standaardafwijking van het modelresidu bedraagt $7,10 \cdot 1.000 \text{ m}^3/\text{d}$, oftewel slechts 3,45% van het gemiddelde dagverbruik in 1986.

Bij de ontwikkeling van dit model

kwamen de volgende zaken aan het licht:

- Als geen rekening wordt gehouden met het optreden van doordeweekse feestdagen, komen er bij de modelresiduen enkele negatieve uitschieters voor op doordeweekse feestdagen, zoals Tweede Paasdag, Koninginnedag, Hemelvaartsdag en Eerste Kerstdag. Dit fenomeen wordt veroorzaakt door het feit dat het vrij regelmatige (en daardoor goed voorspelbare) weekpatroon sterk wordt verstoord door een doordeweekse feestdag. Het verbruik op zo'n feestdag is door de geringere activiteit namelijk minder dan op een gewone doordeweekse dag, zodat het model het verbruik overschat.

- Afgezien van het feit dat dit model wel rekening houdt met het optreden van doordeweekse feestdagen, vertoont zijn vorm grote overeenkomsten met die van een model voor het dagverbruik van drinkwater in Barcelona in 1985 [Quevedo *et al.*, 1988] en die van een model voor het dagverbruik van gas in een aantal regio's van Frankrijk [Borgard *et al.*, 1982]. Blijkbaar vertoont het patroon van het dagelijkse verbruik van dit soort publieke goederen grote overeenkomsten voor bepaalde landen in West-Europa.

Voorspelling van het dagverbruik in 1987

Uitgaande van bovengenoemd transfer

model voor het dagverbruik in 1986 is de voorspelbaarheid van het dagverbruik in 1987 geëvalueerd. Hiertoe is de operationele voorspelsituatie nagebootst, door het dagverbruik van 'morgen' (Z_{t+1}) te voorspellen aan de hand van de beschikbare waarnemingen van het dagverbruik tot en met 'vandaag' ($Z_1 \dots Z_t$). Het functioneren van deze voorspeller kan worden geëvalueerd aan de hand van de relatieve voorspelfout (rvf%), gedefinieerd als:

$$\text{rvf}\% = 100\% \cdot \frac{Z_{t+1} - \hat{Z}_{t+1}}{Z_{t+1}}$$

met Z_{t+1} en $\hat{Z}_{t+1}$ respectievelijk het werkelijke en het voorspelde dagverbruik van morgen. Uit de in afb. 3 weergegeven tijdreeks van de relatieve voorspelfout blijkt dat de voorspelkracht redelijk stationair blijft en dat de voorspelfout overwegend minder dan 5% en zelden meer dan 10% van het betreffende dagverbruik bedraagt (zie ook afb. 4). Concluderend kan worden gesteld dat met een model, afgeleid met de gegevens van een bepaald jaar, nog minstens een jaar goede voorspellingen kunnen worden gegenereerd.

Ontwikkeling van een operationeel tijdreeksmodel

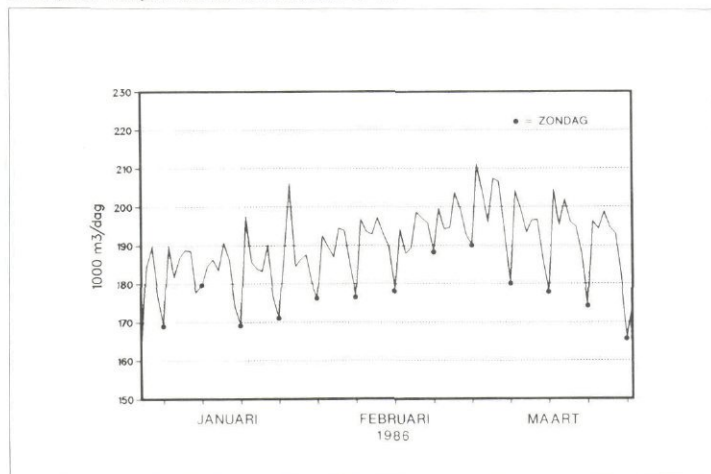
Inleiding

Vervolgens is een operationeel model voor de voorspelling van het dagverbruik vanaf 1989 ontwikkeld aan de hand van het dagverbruik in 1988. Daarbij is ook nagegaan of het model verbeterd kon worden door rekening te houden met meteorologische omstandigheden en de bouwvakvakantie.

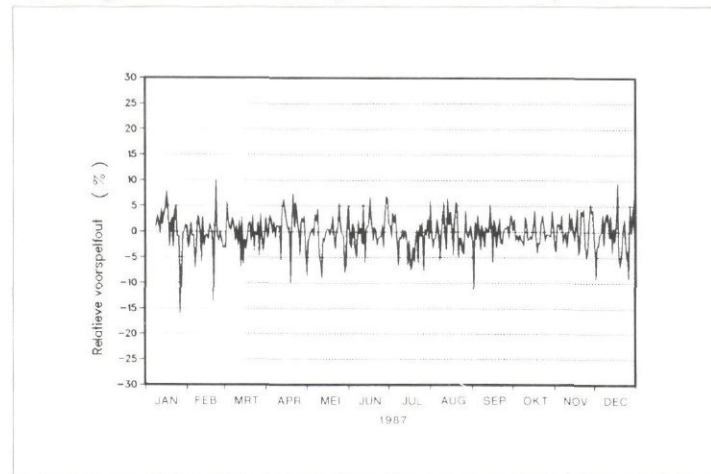
Invloed meteorologische omstandigheden

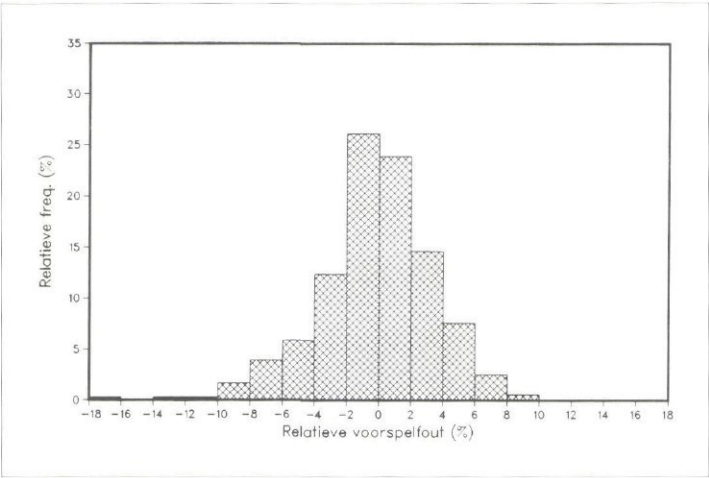
Er mag worden aangenomen dat het dagverbruik naast het optreden van

Afb. 2 - Tijdreeks van het dagverbruik in het voorzieningsgebied van de Berenplaat van januari tot en met maart 1986.



Afb. 3 - Tijdreeks van de relatieve voorspelfout bij gebruik van het transfer model voor het dagverbruik in 1986 om het dagverbruik in 1987 te voorspellen.





Afb. 4 - Histogram van de relatieve voorspelfout bij gebruik van het transfer model voor het dagverbruik in 1986 om het dagverbruik in 1987 te voorspellen.

doordeweekse feestdagen ook wordt beïnvloed door de meteorologische omstandigheden. Dit is nagegaan aan de hand van tijdreeksen van enkele meteorologische variabelen, representatief voor het gebied. Het betrof het aantal zonuren, de gemiddelde temperatuur, de minimumtemperatuur, de maximumtemperatuur, de neerslag en de Makkink-verdamping, alle op dagbasis. Uit de laatste twee is verder nog de tijdreeks van het potentieel neerslagoverschot berekend. Uit een beschouwing van de dynamische relaties bleek het dagverbruik het sterkst gerelateerd aan de Makkink-verdamping, gevolgd door het aantal zonuren en het potentieel neerslagoverschot. Daarop is een evaluatie uitgevoerd van transfer modellen met verschillende combinaties van deze drie meteorologische variabelen, aangevuld met de doordeweekse feestdag, als verklarende variabelen. Het model met als verklarende variabelen de Makkink-verdamping en de doordeweekse feestdag bleek daarbij veruit het meest geschikt. Het toevoegen van andere meteorologische variabelen aan het model gaf geen verbetering, omdat deze sterk gecorreleerd zijn aan de Makkink-verdamping. Het resulterende transfer model was het volgende:

$$\nabla \nabla_7 Z_t = \frac{6,53}{(1 - 0,44 B)} \nabla \nabla_7 X_t - 29,44 \nabla \nabla_7 I_t + \frac{(1 - 0,51 B - 0,20 B^2 + 0,13 B^3)}{(1 - 0,95 B^7)} a_t \tag{2}$$

met Z het verbruik (1.000 m³/d), t de index voor de dag, X de Makkink verdamping (mm/d), I het optreden van een doordeweekse feestdag (0 of 1) en a het modelresidu (1.000 m³/d). De standaardafwijking van het modelresidu bedraagt 7,61·1.000 m³/d, slechts 3,06% van het gemiddelde dagverbruik in 1988. Volgens dit model kan het verbruik

op een bepaalde dag worden beschreven aan de hand van het verbruik tot 9 dagen terug, de Makkink-verdamping tot 8 dagen terug, het optreden van een doordeweekse feestdag tot 9 dagen terug en het modelresidu tot 11 dagen terug.

De vorm van dit transfer model vertoont overeenkomsten met die van een transfer model voor het dagverbruik van gas in Engeland [Lyness, 1984], al fungeert in dat model alleen de temperatuur als verklarende variabele.

In tabel I is voor een aantal verschillende modellen van het dagverbruik in 1988 de standaardafwijking van het modelresidu vermeld, waaruit blijkt dat vooral het toevoegen van de Makkink-verdamping als verklarende variabele een grote verbetering van het model oplevert.

Beschikbaarheid gegevens Makkink-verdamping

Er is een belangrijke praktische beperking bij het genereren van voorspellingen met dit ontwikkelde transfer model. Voor de voorspelling van het verbruik van morgen, moet namelijk ook worden uitgegaan van de Makkink-verdamping van morgen, een

waarde die op het moment van de voorspelling uiteraard nog niet beschikbaar is. Wel kan steeds rond 2 uur 's nachts een prognose van de Makkink-verdamping worden geleverd door Meteo Consult, gebaseerd op kennis van het weersverloop en ervaring. In eerste instantie is daarvan gebruik gemaakt. Een evaluatie van deze methode is uitgevoerd op basis van de gegevens van de afgelopen 8 maanden (12 januari tot en met 1 augustus 1991, 202 dagen). Daarbij is ook nagegaan in hoeverre de voorspelbaarheid van het verbruik van morgen wordt beïnvloed door de Makkink-waarde. Totaal zijn daarbij 4 methoden beschouwd, waarbij in bovengenoemd transfer model (2) voor de voorspelling van het verbruik van morgen, de volgende Makkink-waarden zijn gebruikt:

- 1. de werkelijke Makkink-waarde van morgen
- 2. de morgenochtend rond 2 uur geleverde prognose van de Makkink-waarde van morgen
- 3. de werkelijke Makkink-waarde van vandaag
- 4. geen Makkink-waarde

Bij elke methode, uitgezonderd 4, worden voor de historische Makkink-waarden steeds de werkelijke waarden ingevuld. Relevante karakteristieken van de relatieve voorspelfout zijn vermeld in tabel II.

De gemiddelde relatieve voorspelfout geeft een indicatie van een eventuele systematische afwijking, terwijl de gemiddelde absolute relatieve voorspelfout een geschikte maat is om de voorspelkracht te beoordelen. Uit de tabel blijkt dat, zoals te verwachten, het verbruik van morgen het best voorspeld kan worden met de werkelijke Makkink-waarde van morgen (methode 1), maar die is uiteraard niet op tijd beschikbaar. Van de overige methoden geeft de huidige

TABEL I - Standaardafwijking van het modelresidu van een aantal modellen voor het dagverbruik in 1988.

	Modelresidu	
	st.afw. 1.000 m ³ /d	%st.afw.
univariaat model	9,78	3,93
transfer model met doordeweekse feestdag	9,03	3,63
transfer model met Makkink-verdamping én doordeweekse feestdag	7,61	3,06

st.afw.: standaardafwijking
%st.afw.: relatieve standaardafwijking (%), volgens: %st.afw. = 100% · st.afw./Z
Z = gemiddeld dagverbruik in 1988

TABEL II - Gemiddelde van de relatieve voorspelfout en van de absolute relatieve voorspelfout bij gebruik van verschillende Makkink-waarden.

Methode	1	2	3	4
gemidd. rvf (%)	- 0,09	- 0,85	- 0,07	- 0,08
gemidd. rvf (%)	2,16	2,61	2,99	3,11

(2), met de 's ochtends vroeg geleverde prognose van de Makkink-waarde, de beste voorspellingen. Wèl geeft deze gemiddeld een overschatting van het dagverbruik, wat wordt veroorzaakt door het feit dat de prognose van de Makkink-waarde gemiddeld te hoog ligt. Het mag overigens opmerkelijk worden genoemd dat met het transfer model dat is afgeleid aan de hand van de gegevens van 1988, in 1991 nog steeds zulke goede voorspellingen kunnen worden gegenereerd.

Invloed bouwvakvakantie

Tenslotte is aan de hand van de verbruikscijfers voor 1988 nagegaan of bij voorspellingen ook rekening moet worden gehouden met de bouwvakvakantie. Hiertoe is aan het transfer model met de Makkink-verdamping en de doorde-weekse feestdag, de bouwvakvakantie als verklarende variabele toegevoegd. Deze is geschematiseerd tot een [0,1] variabele, die alleen in de periode 18 juli tot en met 5 augustus 1988 de waarde 1 en de rest van dat jaar de waarde 0 heeft. Omdat uit de geschatte parameters van dit transfer model bleek dat de bouwvakvakantie geen significant effect had op het dagverbruik, is deze niet opgenomen in het operationele model.

Tenslotte

In het hier beschreven geval kon het dagverbruik van drinkwater goed beschreven en voorspeld worden met een Box-Jenkins tijdreeksmodel, dat rekening houdt met doordeweekse feestdagen en de verdamping in het gebied. Omdat de Box-Jenkins methode in vergelijkbare sectoren al zeer vaak geschikt is gebleken om voorspellingen te genereren, mag vooralsnog worden aangenomen dat hij ook in de watersector ruim toepasbaar zal zijn.

De boven beschreven voorspeller functioneert inmiddels al geruime tijd binnen het bedrijfsvoeringsmodel van de Berenplaat, dat ontwikkeld is door DWL Rotterdam [Urbanus en Biemans, deze H₂O]. KIWA zal in overleg met DWL Rotterdam het bedrijfsvoeringsmodel verder ontwikkelen en algemeen gebruik door de drinkwaterbedrijven mogelijk maken. Maar het voorspeldeel zal uiteraard voor elke situatie apart moeten worden ontwikkeld.

Literatuur

Baggelaar, P. K., Urbanus, J. F. X. en Biemans, R. A. G. (1991). 'Optimalisatie bedrijfsvoeringsmodel Berenplaat met behulp van voorspellingen van het verbruik'. KIWA SWE 91.013. KIWA NV, 1991, Nieuwegein
Borgard, C., George, G., Jenkins, G. M. and

In memoriam ir. H. Haverkate

Aan het begin van dit nieuwe jaar werden wij opgeschrikt door het droeve bericht dat onze vriend en oud-collega Hans Haverkate was overleden. Uit de grote belangstelling bij de crematieplechtigheid op 4 januari van dit jaar, bleek welk een vooraanstaande positie hij in onze bedrijfstak heeft ingenomen.

Naast het directeurschap van de Waterleiding Maatschappij Oostelijk Gelderland heeft hij vele functies van gewicht vervuld. Zo was hij voorzitter van de VWN, voorzitter van de Speurwerkadviescommissie en president-commissaris van KIWA NV. Als voorzitter van de VWN heeft hij nationaal en internationaal veel vrienden gemaakt. Daarbij heeft hij de contacten met de verschillende waterleidingcollega's uit de ons omringende landen sterk gestimuleerd. Zijn verdiensten voor de VWN werden nog eens duidelijk toen hij in 1977 tot erelid van de vereniging werd benoemd.

Als voorzitter van de Speurwerkadviescommissie speelde hij een grote rol in de vaststelling van het onderzoekprogramma van de bedrijfstak, juist in een periode dat het speurwerk van de grondwaterbedrijven in omvang en betekenis sterk toenam. Zijn betrokkenheid bij de bescherming van het grondwater bleek onder meer ook uit het voorzitterschap van de Technische Commissie Bodembescherming. Hij werd in 1975 – hij was toen bijna vijftig jaar – op voordracht van het bestuur van de VEWIN tot lid van de Raad van Commissarissen van KIWA NV benoemd. Al spoedig werd hij in dat college geconfronteerd met het afwijzen van een door directie en commissarissen voorbereid voorstel om het gehele instituut in Nieuwegein te vestigen. Naderhand werd alsnog door aandeelhouders ingestemd met een voorstel om alleen de Hoofdafdeling Speurwerk in Nieuwegein te huisvesten.

Bij zijn herbenoeming in 1979 als lid van de Raad van Commissarissen – hij was

toen al een jaar eerder als opvolger van de heer Zandveld tot president-commissaris gekozen – was hij getuige van het slaan van de eerste paal voor de nieuwe speurwerkvestiging in Nieuwegein.



Toen Hans Haverkate in 1983 als president-commissaris aftrad, kon hij met voldoening terugkijken op een periode waarin KIWA ondanks de economische recessie een sterke groei had door-gemaakt.

Nadat hij in januari 1987 afscheid had genomen van zijn bedrijf, na een dienstverband van ruim vijftientig jaar, ontving hij van de VWN de hoogste onderscheiding van onze bedrijfstak, de 'Van Marle Prijs'.

Wij zullen ons Hans Haverkate blijven herinneren als een wijze en beminnelijke collega die bij moeilijke problemen goed kon relativeren en die bij het nemen van besluiten het menselijke aspect op de voorgrond plaatste.

Jos van Winkelen

McLeod, G. (1982). 'Analysing a Large Number of Energy Time Series for a Utility Company'. Pp. 113-149 in: Case Studies in Time Series Analysis, Volume 1, Edited by Jenkins, G. M. and McLeod, G. Time Series Library, Gwilym Jenkins and Partners Ltd., Lancaster.
Box, G. E. P. and Jenkins, G. M. (1976). 'Time Series Analysis, Forecasting and Control'. Holden-Day, San Francisco.
Lyness, F. K. (1984). 'Gas Demand Forecasting'. The Statistician, 33 (1984), pp. 9-21.
McLeod, G. (1983). 'Box Jenkins in Practice, Volume I: Univariate Stochastic and Transfer Function/ Intervention Analysis'. Time Series Library, Gwilym

Jenkins and Partners Ltd., Lancaster.
Quevedo, J., Cembrano, G., Valls, A. and Serra, J. (1988). 'Time Series Modelling of Water Demand - A study on short-term and long-term predictions'. Pp. 268-288 in: Computer Applications in Water Supply, Volume 1, Systems Analysis and Simulation, Edited by Coulbeck, B. and Orr, C. H., Research Studies Press Ltd. Wiley and Sons Inc., New York.
Urbanus, J. F. X. en Biemans, R. A. G. (1991). 'DWL Rotterdam optimaliseert bedrijfsvoering productie- bedrijven'. H₂O, dit nummer.

• • •