



PBL-Notitie

## **De ecologische kwaliteit van Nederlands oppervlaktewater: zijn de voorspellingen beter geworden sinds 2009?**

P. Cleij, H. Visser  
contact: [hans.visser@pbl.nl](mailto:hans.visser@pbl.nl)

Datum: 26-2-2013  
Publicatienummer: 1081

## 1 Aanleiding

In 2009 heeft Royal Haskoning een rapport opgesteld over het verbeteren van datasets en de afleiding van ecologische rekenregels voor de KRW-verkenner. Voorspelling-modellen werden afgeleid op basis van neurale netwerken (Royal Haskoning) en regressiebomen (PBL). Deze modellen werden getest op 29 EKR-deelmaatlatten, verspreid over 8 watertypen. Zie rapport nr. 9T6271, 9 februari 2009: "Verbeteren datasets en afleiding ecologische rekenregels voor de KRW-verkenner". Het rapport is te downloaden van:

[http://publicwiki.deltares.nl/download/attachments/32079883/Evers+et+al+\(2009\)+Verbeteren+datasets+en+afleiding+ecologische+rekenregels+voor+de+KRW-verkenner.pdf?version=1&modificationDate=1271684015000](http://publicwiki.deltares.nl/download/attachments/32079883/Evers+et+al+(2009)+Verbeteren+datasets+en+afleiding+ecologische+rekenregels+voor+de+KRW-verkenner.pdf?version=1&modificationDate=1271684015000)

Over de benadering van regressiebomen heeft het PBL een rapport gepubliceerd in 2008 ("Stuurfactoren voor de ecologische kwaliteit van regionaal oppervlaktewater"). Voor details van de methode wordt verwezen naar dit rapport. Het is te downloaden van:

<http://www.pbl.nl/sites/default/files/cms/publicaties/500140002.pdf>

In 2012 heeft een update plaatsgevonden van dezelfde 29 EKR-deelmaatlatten. Datasets zijn uitgebreid en opgeschoond door Royal Haskoning. Elk van de 29 datasets is nu bij benadering 200 wateren groot. Stuurvariabelen bij elk van de 29 maatlatten zijn identiek gebleven aan die in 2009.

In deze notitie wordt gekeken of de regressieboom-benadering uit 2009 verbeterd wordt door de keuze van PUNN's op de vernieuwde dataset uit 2012. Hierbij wordt gekeken naar vooruitgang in voorspelkracht van de gehanteerde wiskundige technieken.

## 2 Oud en Nieuw in de KRW-verkenner

In deze notitie geven we een inschatting van het verschil in voorspelkracht van het oorspronkelijke ecologische model van de nieuwe KRW-Verkenner in de vorm van een serie regressieboom-modellen en het nieuwe default ecologische model in de vorm van een serie PUNN-modellen. Onderscheiden wordt de verbetering door gebruik van de nieuwe dataset voor training en validatie en de additionele verbetering door de overgang van regressiebomen op PUNN's als model-type.

De vergelijking is primair gemaakt o.b.v. van de Coefficient of Determination voor training op de hele dataset (definitie in Visser (2013)). Reden hiervoor is dat deze CoD-training voor de hele dataset de enig beschikbare performance indicator is voor de oude regressieboom-modellen. Voor de nieuwe regressieboom-modellen is deze indicator alsnog bepaald. Voor de PUNN-modellen is de indicator niet beschikbaar (er heeft geen training op de hele dataset plaatsgevonden). In plaats daarvan is gebruik gemaakt van een (ruwe) schatting van de gemiddelde CoD-training voor de hele dataset.

Hierbij is ervan uitgegaan dat er een systematisch verschil zit tussen de CoD-training voor de trainingset en de CoD-training voor de hele dataset. Als gevolg van de wijze van verdeling van de (nieuwe) dataset in training- en validatieset is de inschatting dat de CoD-training voor de trainingset gemiddeld hoger uitkomt dan voor de hele dataset. Het verschil tussen de gemiddelde CoD-training voor de hele dataset en de trainingset is hierbij geschat aan de hand van het overeenkomstige verschil voor de (nieuwe) regressiebomen.

In aanvulling hierop wordt de verbetering van de voorspelkracht als gevolg van de overgang van regressiebomen naar PUNN's ook geïllustreerd aan de hand van de gemiddelde CoD-waarde voor de validatieset (CoD-validatie). Zie onderstaande tabel voor de relevante vergelijkingsgegevens.

Uit de tabel is af te lezen dat de gemiddelde CoD-training voor de hele dataset van de oorspronkelijke regressieboom-modellen **0.65** is. Voor de nieuwe regressiebomen is dit **0.73**. Deze verbetering van **0.08** kan dus geheel op conto geschreven worden van de nieuwe dataset. Voor de PUNN's is de geschatte waarde voor de gemiddelde CoD-training voor de hele dataset 0.75. Deze additionele verbetering van 0.02 zal dus het gevolg zijn van de overgang van regressiebomen op PUNN's. Dat de overgang van regressiebomen naar PUNN's een verbetering oplevert wordt (meer dan) bevestigd door een vergelijking van de gemiddelde CoD-waarden voor de validatieset: 0.49 voor de (nieuwe) regressiebomen tegen 0.60 voor de PUNN's.

Conclusie is dus dat zowel het gebruik van de nieuwe dataset als de overgang van regressiebomen op PUNN's waarschijnlijk zal leiden tot een verbetering van de voorspelkracht van het ecologische model van de KRW-Verkenner.

Hierbij moet worden aangetekend dat de gemaakte vergelijking niet ideaal is en idealiter gebaseerd zou moeten zijn op waarden van performance indicatoren (CoD, RMSE, percentage binnen 0.1, ...) voor validatiesets, die het resultaat zijn van een willekeurige trekking uit de volledige dataset. De (verschillende) wijze waarop de oude en nieuwe modellen getraind zijn, laten een dergelijke vergelijking echter niet toe.

Watertype	Kwal. elem.	CoD - training					CoD - validatie		
		RB		PUNN			RB		PUNN
		Oud	Nieuw	Nieuw		Nieuw	Nieuw	Nieuw	
		Hele set <sup>1)</sup>	Trainingset <sup>2)</sup>	Hele set <sup>3)</sup>	Trainingset <sup>2)</sup>	Hele set <sup>4)</sup>	Validatieset <sup>2)</sup>	Validatieset <sup>2)</sup>	
1 Langzaam stromende beken	Waterflora	0,58	0,76	0,73	0,81	-	0,48	0,49	
	Macrofauna	0,6	0,87	0,83	0,84	-	0,71	0,79	
	Vissen	0,74	0,85	0,83	0,86	-	0,8	0,8	
2 Snelstromende beken	Waterflora	0,86	0,9	0,89	0,92	-	0,72	0,88	
	Macrofauna	0,73	0,72	0,7	0,76	-	0,41	0,5	
	Vissen	0,72	0,75	0,77	0,82	-	0,72	0,86	
3 Sloten	Waterflora	0,7	0,76	0,45	0,78	-	0,3	0,45	
	Macrofauna	0,64	0,74	0,75	0,78	-	0,45	0,51	
	Vissen	0,69	0,86	0,81	0,84	-	0,75	0,67	
4 Kanalen	Fytoplankton	0,63	0,71	0,72	0,81	-	0,64	0,73	
	Waterflora	0,57	0,72	0,61	0,72	-	0,04	0,13	
	Macrofauna	0,51	0,79	0,73	0,77	-	0,54	0,58	
	Vissen	0,48	0,84	0,81	0,81	-	0,53	0,7	
5 Ondiepe meren	Fytoplankton	-	0,8	0,77	0,85	-	0,58	0,72	
	Waterflora	-	0,86	0,81	0,86	-	0,45	0,37	
	Macrofauna	-	0,78	0,73	0,74	-	0,56	0,57	
	Vissen	-	0,84	0,8	0,85	-	0,51	0,4	
6 Diepe meren	Fytoplankton	0,83	0,8	0,72	0,69	-	0,3	0,63	
	Waterflora	0,79	0,87	0,86	0,85	-	0,54	0,81	
	Macrofauna	0,45	0,86	0,83	0,86	-	0,41	0,6	
	Vissen	0,75	0,85	0,84	0,86	-	0,72	0,69	
7 Zwak brakke wateren	Fytoplankton	0,43	0,45	0,52	0,59	-	0,25	0,54	
	Waterflora	0,69	0,6	0,71	0,81	-	0,7	0,73	
	Macrofauna	0,67	0,72	0,64	0,68	-	-0,04	0,35	
	Vissen	0,7	0,65	0,64	0,77	-	0,64	0,71	
8 Brak tot zoute wateren	Fytoplankton	0,43	0,61	0,5	0,57	-	0,43	0,55	
	Waterflora	0,72	0,74	0,73	0,82	-	0,54	0,6	
	Macrofauna	0,56	0,74	0,69	0,78	-	0,18	0,47	
	Vissen	0,66	0,58	0,63	0,76	-	0,37	0,43	
<b>Gemiddeld (zonder Ondiepe meren):</b>		<b>0,65</b>	<b>0,76</b>	<b>0,73</b>	<b>0,79</b>	<b>0,75</b>	<b>0,49</b>	<b>0,60</b>	
		<i>Oude model (regressie- bomen obv oude dataset)</i>		<i>Verbete- ring door gebruik nieuwe dataset</i>		<i>Verbetering door gebruik nieuwe dataset en PUNN's</i>	<i>Regressie- bomen obv nieuwe dataset</i>	<i>Verbetering door gebruik PUNN's</i>	

1) Evers *et al.* (2009- p. 40 - tabel 5.1)

2) Visser, H. (2013)

3) Persoonlijke mededeling Hans Visser (PBL), d.d. 2012-11-15.

4) Schatting gemiddelde CoD PUNN hele dataset volgens:

[Gem. CoD PUNN trainingset] + ([Gem. CoD RB hele dataset] - [Gem. CoD RB trainingset]).

## Literatuur

Evers, C.H.M., Keukelaar, F., Schomaker, A.H.H.M., 2009. Verbeteren datasets en afleiding ecologische rekenregels voor de KRW-verkenner. Rapport Deltares/PBL, 9 februari 2009.

Schomaker, A.H.H.M., 2012. Verbetering kennisregels ecologische effecten ten behoeve van de KRW-verkenner. Rapport Royal Haskoning DHV, 3 december 2012.

Visser, H., Puijtenbroek, P.J.T.M. van, Janssen, P.H.M., 2008. Stuurfactoren voor de ecologische kwaliteit van regionaal oppervlaktewater. Een statistische analyse met regressiebomen voor de Ex-ante evaluatie KRW. PBL-rapport 500140002/2008.

Visser, H., 2013. De ecologische kwaliteit van Nederlands oppervlaktewater: een vergelijking van drie modelleringstechnieken. PBL-notitie 1045, 31 januari 2013.

Visser, H., Wortelboer, R., 2013. De ecologische kwaliteit van Nederlands oppervlaktewater: een analyse met regressiebomen. PBL-notitie 1075, 25 februari 2013.

**Informatie uit deze notitie mag worden overgenomen op voorwaarde van bronvermelding: "PBL-notitie De ecologische kwaliteit van Nederlands oppervlaktewater: zijn de voorspellingen beter geworden sinds 2009?, rapportnummer 1081, Den Haag: Planbureau voor de Leefomgeving."**