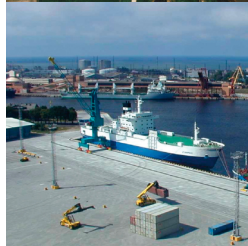



Voorspellen ecologische kwaliteits ratio op basis van product unit neural networks



Voorspellen ecologische kwaliteitsratio op basis van product unit neural networks

| referentie | projectcode | status |
|--------------------|-----------------------------|-----------------|
| STO179-1/smid3/001 | STO179-1 | definitief |
| projectleider | projectdirecteur | datum |
| dr.ir.A.C. de Niet | prof.dr.ir.F.H.L.R. Clemens | 15 oktober 2012 |

| autorisatie | naam | paraaf |
|-------------|----------------------|---|
| goedgekeurd | R.J.Brederveld, MSc. |  |

| INHOUDSOPGAVE | blz. |
|---|--------------------|
| 1. INLEIDING | 1 |
| 1.2. Modelleerwedstrijd | 1 |
| 1.3. Voorspellingstool | 2 |
| 1.4. Leeswijzer | 2 |
| 2. DATASET | 3 |
| 2.1. Opbouw | 3 |
| 2.2. Samenstelling | 3 |
| 2.3. Opdeling | 4 |
| 2.4. Schaling | 4 |
| 3. PRODUCT UNIT NEURAAAL NETWERK | 5 |
| 3.1. Principes achter neurale netwerken | 5 |
| 3.2. Beperkingen van gewone neurale netwerken | 6 |
| 3.3. Product Unit Neuraal Netwerk | 6 |
| 3.3.1. Grote voorspellende kracht | 6 |
| 3.3.2. Netwerk kan worden vereenvoudigd | 6 |
| 3.3.3. Inzicht in relatie tussen input en output | 6 |
| 3.3.4. Transporteerbaar | 7 |
| 3.4. PUNN-toolbox | 7 |
| 3.5. Het ontwikkelen van PUNN's voor de EKR-score | 7 |
| 4. RESULTAAT | 8 |
| 4.1. Beoordelingscriteria | 8 |
| 4.2. Prestaties | 8 |
| 4.3. Voorbeelden van PUNN's | 9 |
| 4.3.1. EKR Fytoplankton zwak brakke wateren | 10 |
| 4.3.2. EKR Vissen zwak brakke wateren | 10 |
| 4.3.3. EKR Fytoplankton diepe meren | 10 |
| 4.3.4. EKR Waterflora diepe meren | 10 |
| 4.3.5. EKR Macrofauna langzaam stromende beken | 11 |
| 4.3.6. EKR Vissen ondiepe meren | 11 |
| 4.4. Evaluatie | 11 |
| 5. CONCLUSIE | 13 |
| laatste bladzijde | 13 |
| BIJLAGEN | aantal blz. |
| I BIJLAGE I MATLABCODE EKR-PUNN'S | 16 |

1. INLEIDING

De Kader Richtlijn Water verplicht veel waterschappen maatregelen te nemen de ecologische kwaliteit in waterlichamen te verbeteren. Die maatregelen zijn vaak kostbaar, dus is het van belang vooraf een goed beeld te hebben van het effect van verschillende maatregelen op de ecologische kwaliteit. De ecologische kwaliteit van een waterlichaam wordt uitgedrukt in de Ecologisch Kwaliteitsratio (EKR). Verschillende hydromorfologische en chemische stuurvariabelen beïnvloeden de EKR. Om een betrouwbare schatting te geven van de effectiviteit van maatregelen heeft Deltares in opdracht van STOWA in de KRW-verkenner een tool opgenomen waarmee bij verschillende maatregelen een EKR-score wordt voorspeld.

In de vigerende versie van de KRW-verkenner is de voorspelling van de EKR-scores gebaseerd op regressiebomen. Deze regressiebomen zijn door het Planbureau voor de Leefomgeving ontwikkeld op basis van een dataset met de ecologische toestand van een groot aantal waterlichamen in Nederland.

In de afgelopen jaren zijn veel nieuwe gegevens verzameld over de ecologische toestand van verschillende waterlichamen. Er zijn nieuwe waterlichamen bijgekomen en de betrouwbaarheid van de gegevens van waterlichamen die al in de dataset zaten is verder verbeterd. Het STOWA wil dat deze nieuwe kennis in de KRW-verkenner wordt verwerkt. Daarvoor moet een nieuw model worden ontwikkeld op basis van de nieuwe dataset. Daarbij is meteen de vraag gesteld of regressiebomen wel het beste model opleveren. Zijn andere modellerings technieken zoals neurale netwerken niet beter?

1.2. Modelleerwedstrijd

Om tot het beste nieuwe model te komen heeft het STOWA drie partijen uitgenodigd om aan een 'wedstrijd' mee te doen. Elke partij is gevraagd op basis van specifieke expertise een zo goed mogelijk model te bouwen. De uitgenodigde partijen zijn:

- planbureau voor de Leefomgeving (PBL): regressiebomen;
- royal Haskoning DHV (RHDHV): neurale netwerken;
- witteveen+Bos (W+B): product unit netwerken.

Voor een eerlijke beoordeling en evaluatie van de kwaliteit van verschillende modellerings technieken is de dataset opgesplitst in twee delen. Het trainingsdeel (80 %) wordt gebruikt voor de ontwikkeling van de modellen. Het validatiedeel (20 %) wordt alleen gebruikt om vast te stellen hoe de verschillende modellen presteren op nieuwe data. Het validatiedeel mag op geen enkele wijze worden betrokken in de tot stand koming van de modellen.

In dit rapport zijn aanpak en resultaten beschreven van de modellering van Witteveen+Bos op basis van product unit netwerken. Een Product Unit Neuraal Netwerk (PUNN) is een speciaal type neuraal netwerk. Een PUNN - ook wel white-box netwerk genoemd - is gebaseerd op producten in plaats van sommen. Het voordeel van een PUNN is dat het mogelijk is het getrainde netwerk te vereenvoudigen zodat alleen de bepalende relaties overblijven. Een PUN kan daardoor meer inzicht geven in de onderliggende structuur dan een traditioneel neuraal netwerk.

1.3. Voorspellingstool

Het resultaat van de modellering met PUNN's is ontsloten door een tool (executable) op te leveren die kan worden geïntegreerd in de KRW-verkenner. De tool voorspelt de EKR-scores op basis van stuurvariabelen. Over de eisen aan de in- en output is overleg geweest met Deltares.

1.4. Leeswijzer

In hoofdstuk 2 is de dataset beschreven waarop de modellering is gebaseerd. In hoofdstuk 3 wordt beschreven wat een PUNN is en hoe deze op basis van de datasets tot stand zijn gekomen. In hoofdstuk 4 wordt ingegaan op de prestaties van de PUNN's voor het voorspellen van de EKR scores.

2. DATASET

In dit hoofdstuk is beschreven op basis van welke dataset de modellen zijn gebouwd.

2.1. Opbouw

De dataset bestaat uit gegevens van een groot aantal waterlichamen in Nederland. Per waterlichaam is een Ecologische Kwaliteitsratio (EKR) score bepaald voor de volgende kwaliteitselementen:

- fytoplankton;
- overige waterflora;
- macrofauna;
- vissen.

De EKR-scores hebben een waarde tussen 0 en 1 en wordt bepaald met KRW-maatlatten. De maatlat wordt voor natuurlijke watertypen opgedeeld in de klassen Slecht (0-0.2), Ontoereikend (0.2-0.4), matig (0.4-0.6), goed (0.6-0.8) en zeer goed (0.8-1.0).

De ecologische kwaliteit wordt bepaald door inrichting, belasting en onderhoud van een waterlichaam. De invloedsfactoren zijn gevat in de volgende 12 stuurvariabelen:

- oeverinrichting;
- peildynamiek;
- onderhoud;
- connectiviteit;
- meandering;
- verstuwning;
- beschaduwning;
- scheepvaart;
- BZV;
- chloride concentratie;
- totaal P concentratie;
- totaal N concentratie.

Welke stuurvariabelen er precies toe doen hangt af van het watertype. De watertypen zijn samengevoegd in de volgende watertypenclusters:

- langzaam stromende beken (R4, R5, R6 en R12);
- snel stromende beken (R13, R14, R15, R17 en R18);
- diepe meren (M20);
- ondiepe meren (M14, M23 en M27);
- kanalen (M3, M4, M6a/b, M7a/b en M10);
- sloten (M1a/b, M2 en M8);
- zwak brakke wateren (M30);
- brakke tot zoute wateren (M31).

2.2. Samenstelling

De dataset is in opdracht van STOWA samengesteld, gecontroleerd en geleverd door Royal Haskoning DHV. Per watertypencluster is één Excel-file geleverd met waarin per locatie is aangegeven wat de waarde is van de stuurvariabelen en de EKR-scores. De dataset is deels gebaseerd op metingen, deels op schattingen en deels op een expert-oordeel. In een aantal gevallen zijn kunstmatige wateren geïntroduceerd zodat ook extreme situaties in de database zijn opgenomen.

2.3. Opdeling

Per watertypencluster en per EKR-score zijn de waterlichamen gesplitst in twee categorieën. Een deel voor training en een deel voor validatie. Het trainingsdeel mag worden gebruikt voor het ontwikkelen en trainen van de voorspellingsmodellen. Het validatie-deel mag daar op geen enkele wijze bij betrokken worden en dient alleen om achteraf de kwaliteit te beoordelen van het model. De indeling is door Royal Haskoning DHV bepaald en meegeleverd met de data.

2.4. Schaling

De stuurvariabelen verschillen in grootte en spreiding. Omdat het voor het ontwikkelen van de modellen beter is de stuurvariabelen dezelfde orde van grootte hebben, zijn de variabelen geschaald. Voor product unit neurale netwerken moeten de inputvariabelen strikt positief zijn. Aan die eis is voor alle variabelen voldaan, zodat er geen verplaatsing (shift) nodig is. Een overzicht van de schalingsfactoren is weergegeven in onderstaande tabel.

Tabel 2.1. Grenzen en schalingsfactoren voor de stuurvariabelen

| stuurvariabele | min | max | schalingsfactor |
|-----------------------|------|-------|-----------------|
| oeverinrichting | 1 | 3 | 2 |
| peildynamiek | 1 | 3 | 2 |
| onderhoud | 1 | 2 | 1,5 |
| connectiviteit | 1 | 3 | 2 |
| meandering | 1 | 5 | 3 |
| verstuwung | 1 | 3 | 2 |
| beschaduwung | 1 | 3 | 2 |
| scheepvaart | 1 | 2 | 1,5 |
| BZV | 0,5 | 20 | 2 |
| chloride concentratie | 100 | 15000 | 1000 |
| totaal P concentratie | 0,01 | 10 | 0,1 |
| totaal N concentratie | 0,67 | 100 | 3 |

De minima en maxima komen uit het rapport 'Verdere ontwikkeling Expertstelsysteem Ecologische Effecten en evaluatie gebruik in de Ex ante evaluatie KRW'(april 2009) dat Royal Haskoning heeft opgesteld in opdracht van PBL.

De data is bij het ontwikkelen van de modellen niet getoetst op deze minima en maxima. Dat wil zeggen dat er geen aanpassingen zijn gedaan om te zorgen dat de waarden tussen de aangegeven minima en maxima liggen. Bij een aantal waarden in de dataset is gebleken dat deze buiten de grenzen van de stuurvariabelen liggen. In de training van de modellen zijn deze waarden gewoon meegenomen.

Bij het toepassen van de modellen wordt wel gecontroleerd of de waarden van de stuurvariabelen tussen de minima en maxima liggen. Als dat niet het geval wordt afgerond naar het minimum of het maximum.

De schalingsfactoren worden zowel bij de training als bij het toepassen van de modellen gebruikt.

3. PRODUCT UNIT NEURAAAL NETWERK

In dit hoofdstuk wordt kort uitgelegd wat de principes van een neuraal netwerk zijn. Vervolgens wordt het product unit neuraal netwerk (PUNN) geïntroduceerd en wordt aangegeven wat de belangrijkste verschillen zijn ten opzichte van gewone neurale netwerken.

3.1. Principes achter neurale netwerken

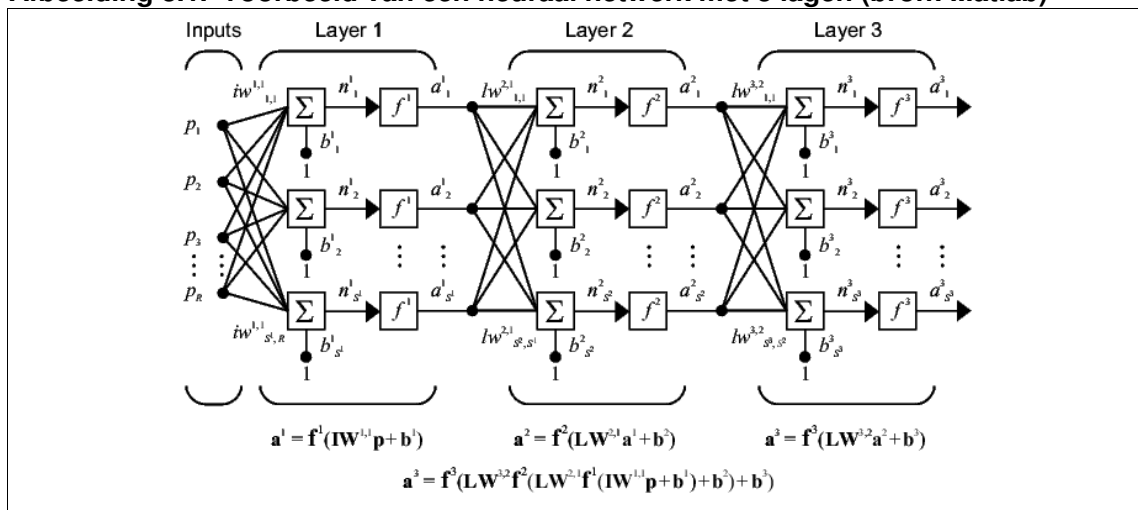
Neurale netwerken zijn geïnspireerd door het menselijk brein. Het menselijk brein bestaat uit neuronen die aan elkaar gekoppeld zijn en leert door training en herhaling nieuwe begrippen en situaties. Hersens werken met elektrische signalen, wanneer er bij een neuron een signaal binnenkomt, kan het neuron dit signaal doorgeven. Afhankelijk van de signaalsterkte geeft het neuron een sterker of zwakker signaal door. Het neuron manipuleert dus de input en geeft een aangepast signaal door. Dat principe is vertaald naar een wiskundige techniek waarbij een kunstmatig neuron de inputvariabelen weegt en sommeert met de volgende formule:

$$\sum_{u=1}^{n_x} w_u x_u$$

Het resultaat wordt gemanipuleerd via een transferfunctie en dat wordt als output doorgegeven.

Door verschillende neuronen naast elkaar te zetten en de output van de ene laag als input voor de volgende te gebruiken ontstaat een neuraal netwerk. Een voorbeeld daarvan is weergegeven in afbeelding 3.1 Er zijn erg veel vrijheden in de keuze van het aantal neuronen, de manier waarop deze zijn gekoppeld en welke transferfuncties worden toegepast.

Afbeelding 3.1. Voorbeeld van een neuraal netwerk met 3 lagen (bron: Matlab)



Voor een effectief gebruik van neurale netwerken is het van belang de structuur van het netwerk af te stemmen op het onderliggende problemen. Ook moet het netwerk op de juiste manier worden getraind om te voorkomen dat het netwerk uitstekend voorspelt op de trainingset, maar daarbuiten de plank volledig mislaat. Als aan die voorwaarden is voldaan kan het een krachtige tool zijn om modellen te ontwikkelen op basis van grote datasets.

3.2. Beperkingen van gewone neurale netwerken

De standaard neurale netwerken hebben beperkingen. Één van de belangrijkste nadelen is het feit dat het model een black-box is. Als het neurale netwerk een goed model oplevert, is het volkomen onduidelijk waarom. Het levert geen inzicht in de aard van de relatie tussen de input en de output van het netwerk. Daarnaast is het lastig een getraind neurale netwerk buiten de omgeving van Matlab toe te passen. Het kan wel, maar vraagt een inspanning die groter wordt naarmate het neurale netwerk complexer wordt.

3.3. Product Unit Neuraal Netwerk

Er is een alternatieve formulering voor neurale netwerken die niet gebaseerd is op een som maar op een product: een product unit neuraal netwerk (PUNN). Een PUNN bestaat uit een lineaire combinatie van een constante en termen van de volgende vorm:

$$\prod_{u=1}^{n_x} x_u^{p_u}$$

In plaats van een sommatie van gewogen inputs, wordt gebruik gemaakt van een product waarbij de gewichten de rol van een machtsverheffing hebben gekregen. Bij gewone neurale netwerken worden de gewichten bepaald door training. Bij een PUNN geldt dat voor de machten. Een PUNN heeft in vergelijking met een gewoon neurale netwerk een paar voordelen die kort zullen worden beschreven.

3.3.1. Grote voorspellende kracht

Het eerste voordeel van PUNN's is dat producten een grotere ruimte beschrijven dan sommen. Een PUNN kan daardoor met minder knopen dan een standaard neurale netwerk hetzelfde resultaat bereiken.

3.3.2. Netwerk kan worden vereenvoudigd

Als een PUNN is getraind, valt direct uit het getrainde netwerk af te lezen welke inputs een grote rol spelen en welke er juist minder toe doen. Een input waarvan de machtsverheffing dicht bij 0 ligt, is vrijwel constant en kan mogelijk worden verwijderd. Via een iteratief proces van verwijdering van niet-relevante nodes en hertraining kan een PUNN worden vereenvoudigd. Dit proces heet snoeien (Engels: pruning). Het resultaat van snoeien is een sterk vereenvoudigd netwerk waarin alleen belangrijke structuren overblijven.

3.3.3. Inzicht in relatie tussen input en output

Uit wat overblijft van een PUNN na trainen en snoeien valt af te lezen wat de relatie is tussen de in- en output. Uit de machten is direct te zien wat de aard van het verband is. Juist daarom zijn PUNN's goed bruikbaar voor het ontdekken van verbanden uit grote datasets van nog grotendeels onbegrepen problemen.

3.3.4. Transporteerbaar

Een PUNN is een formule die zonder veel problemen expliciet kan worden uitgeschreven. Dat maakt het heel eenvoudig een getrainde PUNN in een andere omgeving (Excel, externe code) te gebruiken.

3.4. PUNN-toolbox

Voor het toepassen van PUNN's heeft Witteveen+Bos een toolbox ontwikkeld. De toolbox is een uitbreiding op de Neural Network toolbox van Matlab. De toolbox maakt het mogelijk PUNN's op te zetten, te trainen, te snoeien en te exporteren in formule-vorm naar een tekstfile. De toolbox is door Witteveen+Bos meerdere keren succesvol toegepast. Bijvoorbeeld bij het ontwikkelen van een metamodel voor PCLake. De PUNN's zijn opgenomen in een website, waardoor een deel van de resultaten van PCLake (de waarden voor kritische belasting) voor een groter publiek is ontsloten.

3.5. Het ontwikkelen van PUNN's voor de EKR-score

Per Watertypencluster en per EKR is één PUNN ontwikkeld. Daarbij zijn de volgende uitgangspunten gehanteerd:

- de PUNN bestaat uit 4 termen plus een constante;
- elke PUNN wordt 2000 keer getraind na random initialisatie;
- de beste 30 PUNN's worden gesnoeid;
- de beste gesnoeide PUNN wordt opgeleverd.

Bij het trainen van de PUNN's is het doel het minimaliseren van de RMSE (root mean squared error). De RMSE is ook gebruikt om te bepalen welke (gesnoeide) PUNN het beste is. De RMSE is de standaard keuze, maar ligt ook voor de hand omdat grote fouten door kwadratering relatief zwaar wegen. De optimalisatie zal er dus voor zorgen dat er minder grote afwijkingen zijn dan wanneer de absolute fout zou zijn geminimaliseerd. Daar komt bij dat de RSME één van de beoordelingcriteria voor de prestatie van de verschillende modellen is.

De PUNN's zijn getraind en gesnoeid op basis van geschaalde data. De schaling kan achteraf expliciet in de PUNN's worden verwerkt. Dat verandert niets aan de machten zelf. Alleen de factor voor elk deelproduct verandert.

Wat het trainen van PUNN's voor de EKR-score voor resultaat heeft opgeleverd is beschreven in het volgende hoofdstuk.

4. RESULTAAT

Per Watertypencluster en per EKR score is een PUNN getraind en gesnoeid. In totaal gaat het om 29 netwerken.

4.1. Beoordelingscriteria

In overleg met STOWA, Deltares, PBL en Royal Haskoning DHV zijn 4 beoordelingscriteria vastgesteld waarop de kwaliteit van de modellen wordt beoordeeld. Het gaat om de kengetallen R-kwadraat, het percentage binnen 0.10, de RMSE en de 'coefficient of determination' (gecorrigeerde R-kwadraat). De kengetallen worden berekend over de validatieset, die op geen enkele wijze is betrokken bij de training van de netwerken.

4.2. Prestaties

In de tabellen is aangegeven hoe de netwerken van de verschillende EKR's presteren op de bovengenoemde beoordelingscriteria. De kengetallen zijn ook voor de trainingsset weergegeven.

Tabel 4.1. Kwaliteit EKR-PUNN's voor Langzaam stromende beken

| EKR | training | | | | validatie | | | |
|------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| waterflora | 0.81 | 87 % | 0.066 | 0.81 | 0.54 | 78 % | 0.080 | 0.49 |
| macrofauna | 0.84 | 85 % | 0.070 | 0.84 | 0.81 | 85 % | 0.064 | 0.79 |
| vissen | 0.86 | 88 % | 0.062 | 0.86 | 0.84 | 80 % | 0.081 | 0.80 |

Tabel 4.2. Kwaliteit EKR-PUNN's voor Snelstromende beken

| EKR | training | | | | validatie | | | |
|------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| waterflora | 0.92 | 94 % | 0.056 | 0.92 | 0.88 | 85 % | 0.072 | 0.88 |
| macrofauna | 0.77 | 72 % | 0.10 | 0.76 | 0.52 | 46 % | 0.13 | 0.50 |
| vissen | 0.82 | 72 % | 0.093 | 0.82 | 0.86 | 74 % | 0.097 | 0.86 |

Tabel 4.3. Kwaliteit EKR-PUNN's voor Sloten

| EKR | training | | | | validatie | | | |
|------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| waterflora | 0.78 | 66 % | 0.096 | 0.78 | 0.47 | 69 % | 0.097 | 0.45 |
| macrofauna | 0.78 | 71 % | 0.088 | 0.78 | 0.55 | 56 % | 0.13 | 0.51 |
| vissen | 0.84 | 88 % | 0.068 | 0.84 | 0.71 | 79 % | 0.089 | 0.67 |

Tabel 4.4. Kwaliteit EKR-PUNN's voor Kanalen

| EKR | training | | | | validatie | | | |
|--------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| fytoplankton | 0.81 | 78 % | 0.089 | 0.81 | 0.77 | 73 % | 0.095 | 0.73 |
| waterflora | 0.72 | 74 % | 0.091 | 0.72 | 0.32 | 58 % | 0.12 | 0.13 |
| macrofauna | 0.77 | 71 % | 0.091 | 0.77 | 0.63 | 65 % | 0.11 | 0.58 |
| vissen | 0.81 | 85 % | 0.071 | 0.81 | 0.71 | 80 % | 0.072 | 0.70 |

Tabel 4.5. Kwaliteit EKR-PUNN's voor Ondiepe meren

| EKR | training | | | | validatie | | | |
|--------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| fytoplankton | 0.85 | 73 % | 0.095 | 0.85 | 0.79 | 67 % | 0.13 | 0.72 |
| waterflora | 0.86 | 75 % | 0.079 | 0.86 | 0.42 | 59 % | 0.14 | 0.37 |
| macrofauna | 0.74 | 77 % | 0.087 | 0.74 | 0.58 | 64 % | 0.097 | 0.57 |
| vissen | 0.85 | 81 % | 0.080 | 0.85 | 0.44 | 69 % | 0.11 | 0.40 |

Tabel 4.6. Kwaliteit EKR-PUNN's voor Diepe meren

| EKR | training | | | | validatie | | | |
|--------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| fytoplankton | 0.69 | 50 % | 0.15 | 0.69 | 0.64 | 46 % | 0.17 | 0.63 |
| waterflora | 0.85 | 76 % | 0.086 | 0.85 | 0.82 | 73 % | 0.091 | 0.81 |
| macrofauna | 0.86 | 87 % | 0.066 | 0.86 | 0.61 | 85 % | 0.080 | 0.60 |
| vissen | 0.86 | 90 % | 0.069 | 0.86 | 0.70 | 69 % | 0.12 | 0.69 |

Tabel 4.7. Kwaliteit EKR-PUNN's voor Zwak brakke wateren

| EKR | training | | | | validatie | | | |
|--------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| fytoplankton | 0.59 | 54 % | 0.14 | 0.59 | 0.55 | 51 % | 0.13 | 0.54 |
| waterflora | 0.81 | 75 % | 0.082 | 0.81 | 0.76 | 76 % | 0.10 | 0.73 |
| macrofauna | 0.68 | 77 % | 0.10 | 0.68 | 0.37 | 66 % | 0.095 | 0.35 |
| vissen | 0.77 | 79 % | 0.078 | 0.77 | 0.71 | 78 % | 0.099 | 0.71 |

Tabel 4.8. Kwaliteit EKR-PUNN's voor Brak tot zoute wateren

| EKR | training | | | | validatie | | | |
|--------------|----------------|------------------------|-------|-------------------|----------------|------------------------|-------|-------------------|
| | R ² | perc. bin- nen 0.10 | RMSE | coef of determ | R ² | perc. bin- nen 0.10 | RMSE | coef of determ |
| fytoplankton | 0.57 | 58 % | 0.12 | 0.57 | 0.56 | 59 % | 0.11 | 0.55 |
| waterflora | 0.82 | 84 % | 0.067 | 0.82 | 0.62 | 73 % | 0.090 | 0.60 |
| macrofauna | 0.78 | 68 % | 0.10 | 0.78 | 0.50 | 46 % | 0.15 | 0.47 |
| vissen | 0.76 | 86 % | 0.073 | 0.76 | 0.50 | 57 % | 0.13 | 0.43 |

4.3. Voorbeelden van PUNN's

Om een idee te geven wat de PUNN's nu opleveren worden een paar voorbeelden gegeven van PUNN's. Het merendeel voor relaties waar het vereenvoudigen (snoeien) een interessant resultaat oplevert. Er zijn ook modellen waar helemaal niet vereenvoudigd kan worden. Daarvan is ook één voorbeeld gegeven.

De formules zijn op basis van de geschaalde stuurvariabelen.

De getallen in de formules zijn afgerond tot drie significante cijfers. In de opgeleverde code is met een aanmerkelijk hogere precisie gewerkt. Dat heeft als reden dat er cijferverlies kan optreden, waardoor het berekende resultaat onnauwkeurig wordt.

4.3.1. EKR Fytoplankton zwak brakke wateren

Stuurvariabelen: Oeverinrichting (Oe), Peildynamiek (Pe), Onderhoud (O), Connectiviteit (C), Chloride (Cl), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Fytoplankton is als volgt:

$$EKR = -0.690 - 0.0565 \frac{O^{3.26}}{P} - 0.378 \frac{C^{3.48} P^{0.435} N^{0.692}}{Pe^{4.71} Cl^{1.47}} \\ + 1.50 \frac{1}{Pe^{0.0574} P^{0.0887} N^{0.154}} + 0.490 \frac{C^{3.50} P^{0.451} N^{0.697}}{Pe^{4.24} Cl^{1.46}}$$

In deze formule komt oeverinrichting (Oe) niet voor. Voor de EKR-score fytoplankton is dat blijkbaar geen echte stuurvariabele.

4.3.2. EKR Vissen zwak brakke wateren

Stuurvariabelen: Oeverinrichting (Oe), Peildynamiek (Pe), Onderhoud (O), Connectiviteit (C), Chloride (Cl), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Vissen is als volgt:

$$EKR = 0.395 + 0.24 \frac{Pe^{1.5}}{N^{0.388}} - 0.0575 \frac{Cl^{0.293} P^{0.401} N^{0.224}}{O^{1.28}} \\ - 0.173 \frac{Pe^{0.857}}{Cl^{0.545} P^{0.214}} + 0.0166 \frac{Oe^{1.52} C^{1.27} Cl^{0.64} P^{0.47}}{Pe^{1.58} O^{1.85}}$$

Alle stuurvariabelen komen in deze formule voor.

4.3.3. EKR Fytoplankton diepe meren

Stuurvariabelen: Oeverinrichting (Oe), Peildynamiek (Pe), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Fytoplankton voor diepe meren is als volgt:

$$EKR = -0.735 + 0.726 \frac{1}{N^{0.288}} + 0.0156 N^{1.54} + 0.358 \frac{N^{0.311}}{P^{0.409}}$$

De oeverinrichting en de peildynamiek komen niet in de formule voor en zijn dus geen echte stuurvariabele. Ecologisch gezien is dat geen verrassing.

4.3.4. EKR Waterflora diepe meren

Stuurvariabelen: Oeverinrichting (Oe), Peildynamiek (Pe), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Overige waterflora voor diepe meren is als volgt:

$$EKR = 7.05 - 0.578 \frac{Pe^{3.13}}{Oe^{2.46}} + 0.319 \frac{Pe^{3.55}}{Oe^{3.15}} + 0.00399 \frac{1}{Pe^{7.17}} - 6.38 \frac{P^{0.0156}}{Pe^{0.178}}$$

In deze formule komt N-totaal niet voor en speelt P-totaal een beperkte rol. De EKR-score Waterflora wordt voornamelijk bepaald door de peildynamiek en de verhouding tussen peildynamiek en oeverinrichting. Vanuit ecologisch perspectief is conform de verwachting. In diepe plassen zijn de N- en P-concentraties over het algemeen erg laag. Het voorkomen van waterplanten wordt daarom vooral door het doorzicht (lichtklimaat) en begroeibaar areaal (gerelateerd aan oeverinrichting en peildynamiek) bepaald.

4.3.5. EKR Macrofauna langzaam stromende beken

Stuurvariabelen: Meandering (M), Verstuwing (St), Beschaduwning (Sch), BZV (B), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Macrofauna voor langzaam stromende beken is als volgt:

$$EKR = -0.0358 + 0.31M^{0.745}St^{1.31} + 0.627\frac{M^{0.875}Sch^{0.637}B^{0.172}}{St^{3.08}} - 0.545\frac{M^{0.793}Sch^{0.562}B^{0.171}P^{0.0143}}{St^{3.29}} + 0.144\frac{1}{St^{1.86}}$$

Totaal fosfaat en totaal stikstof komen niet of nauwelijks voor in deze formule. Vooral de (complexe) relatie met verstuwing springt er uit. In beken is vooral substraat, inrichting en stromingssnelheid van belang. De verblijftijd is meestal dusdanig kort dat nutriënten zoals P en N niet tot uitdrukking komen.

4.3.6. EKR Vissen ondiepe meren

Stuurvariabelen: Oeverinrichting (Oe), Peildynamiek (Pe), P-totaal (P), N-totaal (N). De relatie tussen de stuurvariabelen en de EKR Vissen voor ondiepe meren is als volgt:

$$EKR = 0.201 + 1.33\frac{Oe^{0.139}}{Pe^{1.06}P^{0.771}N^{0.153}} - 0.405\frac{Oe^{0.529}}{Pe^{1.43}P^{1.01}N^{0.33}} - \frac{1}{Oe^{0.452}Pe^{2.28}P^{0.624}N^{0.107}} + 0.266\frac{1}{Oe^{0.262}Pe^{3.72}P^{0.394}N^{0.471}}$$

Bij deze PUNN bleken het niet mogelijk te snoeien zonder kwaliteit te verliezen in de voorspelling. Daardoor valt uit deze relatie minder duidelijk iets af te lezen.

4.4. Evaluatie

Als gekeken wordt naar de prestaties van de PUNN's valt op dat er grote verschillen zijn tussen de modellen. Er zijn watertypenclusters en EKR's waarvoor de PUNN uitstekend presteert. Bijvoorbeeld voor Waterflora bij Snel stromende beken ligt de COD (coefficient of determination) voor training- en validatieset erg hoog (circa 0,90). Maar er zijn ook modellen waarvan de kwaliteit matig is. Voor Fytoplankton bij Brakke tot zoute wateren bijvoorbeeld met een COD van 0,55. Dat heeft waarschijnlijk voor een groot deel te maken met de samenstelling van de dataset. Over het algemeen ligt de COD van de trainingset dicht bij de COD van de validatieset. Dat geeft vertrouwen in de modellen. Het laat zien dat de netwerken niet overtraind zijn.

In een aantal gevallen (bijvoorbeeld EKR Waterflora van Kanalen) zijn er wel frappante verschillen tussen de kwaliteit van het model bij training en validatie. Dat zou er op kunnen duiden dat de validatieset niet voldoende representatief is voor de trainingset. Oftewel er zitten te veel bijzondere gevallen in. Dat kan zijn veroorzaakt door de wijze waarop beide sets zijn samengesteld door Royal Haskoning DHV. Achteraf is gebleken dat de opdeling van validatie- en training afhangt van de keuze van het beste gewone neurale netwerk. Dat heeft tot gevolg dat afwijkende combinaties van stuurvariabelen en EKR een voorkeur hebben om in de validatieset terecht komen. De validatieset is daardoor niet volledig random, maar verschilt significant van de trainingset. Een correcte voorspelling van de validatieset is daardoor moeilijker. Desalniettemin blijken de PUNN's in veel gevallen in staat een model op te leveren dat op beide sets voldoende goed presteert.

In de Tabel 4.9 en 4.10 is een oordeel gegeven over de kwaliteit van de PUNN's op basis van de validatieset en de testset. Het oordeel is gebaseerd op de drie evaluatiecriteria:

- meer dan 75 % van de dataset heeft een absolute fout kleiner dan 0,1;
- de RMSE is kleiner dan 0,1;
- de COD is groter dan 0,75.

Als een model op deze drie criteria positief test krijgt het de score uitstekend (++). Wordt aan 2 van de 3 criteria voldaan, dan is de score goed (+). Bij één positieve score is het model voldoende (o). Als aan geen enkel criterium is voldaan, scoort het model matig (-). Bij de tabel moet worden opgemerkt dat de criteria niet onafhankelijk zijn. Het is bijvoorbeeld vrij onwaarschijnlijk dat een model een grote fout (RMSE) heeft in combinatie met een hoge COD.

Tabel 4.9. Oordeel kwaliteit PUNN's op basis van validatieset

| | fytoplankton | waterflora | macrofauna | vissen |
|-------------------------|--------------|------------|------------|--------|
| lanzaam stromende beken | nvt | + | ++ | ++ |
| snel stromende beken | nvt | ++ | - | + |
| sloten | nvt | o | - | + |
| kanalen | o | - | - | + |
| ondiepe meren | - | - | o | - |
| diepe meren | - | + | + | - |
| zwak brakke wateren | - | + | o | + |
| brak tot zoute wateren | - | o | - | - |

++ = uitstekend, + = goed, o = voldoende, - = matig

Tabel 4.10. Oordeel kwaliteit PUNN's op basis van trainingset

| | fytoplankton | waterflora | macrofauna | vissen |
|-------------------------|--------------|------------|------------|--------|
| lanzaam stromende beken | nvt | ++ | ++ | ++ |
| snel stromende beken | nvt | ++ | + | + |
| sloten | nvt | + | + | ++ |
| kanalen | ++ | o | + | ++ |
| ondiepe meren | + | ++ | + | ++ |
| diepe meren | - | ++ | ++ | ++ |
| zwak brakke wateren | - | ++ | + | ++ |
| brak tot zoute wateren | - | ++ | + | ++ |

++ = uitstekend, + = goed, o = voldoende, - = matig

Uit de laatste tabel blijkt dat de PUNN's op de trainingset bijna altijd goed of uitstekend presteren. Alleen Fytoplankton blijkt moeilijker te voorspellen voor diepe meren, zwak brakke wateren en brak tot zoute wateren. De prestaties verslechteren als wordt gekeken naar de validatieset. Dat heeft voor een deel te maken met de opbouw van de validatieset. Toch is in veel gevallen de kwaliteit nog voldoende tot goed. Waar de kwaliteit als matig wordt bestempeld, geldt in de meeste gevallen dat de RMSE tussen 0,10 en 0,15 ligt.

Bij veel EKR-modellen is het mogelijk de PUNN's te vereenvoudigen, wat inzicht geeft in de relatie tussen de stuurvariabelen en de EKR-score. In het ene geval is dat wat duidelijker dan in het andere.

5. CONCLUSIE

Uit dit rapport blijkt dat het goed mogelijk is om PUNN's af te leiden uit de geleverde dataset. De PUNN's leveren over het algemeen voldoende betrouwbare modellen om te worden opgenomen in de KRW-verkenner. Daarnaast kunnen ze gebruikt worden om na te denken over aard van de relaties tussen stuurvariabelen en EKR-scores. De gesnoeide PUNN's geven gekwantificeerd inzicht in de bijdragen van de verschillende sturingsfactoren. Het is de moeite waard daar een uitgebreidere ecologische interpretatie van te maken dan in dit rapport is gedaan.

