

# **Verbetering kennisregels ecologische effecten**

Ten behoeve van de KRW Verkenner

STOWA

6 februari 2013  
Definitief rapport  
9X2209





Documenttitel Verbetering kennisregels ecologische  
effecten  
Ten behoeve van de KRW Verkenner  
Verkorte documenttitel EEE3  
Status Definitief rapport  
Datum 6 februari 2013  
Projectnaam EEE3  
Projectnummer 9X2209  
Opdrachtgever STOWA  
Referentie 9X2209/R00003/419090/BW/DenB

Auteur(s) ir. A.H.H.M. Schomaker  
Collegiale toets ir. R.A.E. Knoben  
Datum/paraaf 27-02-2013  
Vrijgegeven door ir. R.A.E. Knoben  
Datum/paraaf 27-02-2013









## INHOUDSOPGAVE

	Blz.
1 INLEIDING	1
2 OPZET	2
2.1 Methodiek	2
2.2 Verbetering datasets	2
2.3 Training neurale netwerken	3
3 RESULTATEN	8

## BIJLAGEN

1. Beschrijving NN langzaam stromende beken
2. Beschrijving NN snel stromende beken
3. Beschrijving NN diepe meren
4. Beschrijving NN ondiepe meren
5. Beschrijving NN kanalen
6. Beschrijving NN sloten
7. Beschrijving NN zwak brake wateren
8. Beschrijving NN brak tot zoute wateren





## 1 INLEIDING

Voor de nieuwe versie van de KRW-verkenner bleek er behoefte te bestaan om de daarin opgenomen kennisregels te verbeteren. Daartoe diende het bestaande Expertsysteem Ecologisch Effecten (EEE2) een update te krijgen naar EEE3. Dit was mogelijk omdat sinds het beschikbaar komen van EEE2 in 2009 de destijds gebruikte datasets steeds completer en betrouwbaarder werden.

Royal HaskoningDHV heeft van STOWA opdracht gekregen om voor EEE3:

- Bestaande datasets aan te vullen en te verbeteren.
- Nieuwe neurale netwerken te trainen voor de acht verschillende watertypen.
- De verkregen netwerken te valideren en tezamen als 'stand alone' executable te compileren voor gebruik in de KRW-verkenner.

In deze korte rapportage wordt de methodiek, opzet en resultaten van de neurale netwerken beschreven. In bijlagen 1 t/m 8 worden de concrete resultaten van deze netwerken voor elk van de acht watertypen gegeven.

## 2 OPZET

### 2.1 Methodiek

Met de aangevulde en verbeterde datasets zijn nieuwe neurale netwerken getraind met de Neural Network Toolbox V7.0 (R2010b) in Matlab V7.11 (R2010b) van Mathworks®. Het daarbij gebruikte type neurale netwerk is een zogenaamd *feed forward back propagation* netwerk dat ook toegepast is voor het trainen van de EEE2 netwerken<sup>1</sup>. Voor elk te trainen netwerk zijn de datasets verdeeld in een set met 80% wateren voor de training en een set van 20% wateren voor validatie.

De aldus verkregen netwerken zijn vervolgens gecompileerd tot de executable EEE3.exe met Matlab Compiler V4.14 (R2010b). Deze executable leest een csv-invoerbestand in en genereert twee uitvoerbestanden, te weten:

- Een csv-bestand waarin eventuele te hoge of te lage waarden van een stuurvariabele als 'outliers' zijn aangegeven.
- Een csv-bestand met voorspelde EKR's voor alle kwaliteitselementen.

Deze executable is aan Deltares geleverd voor koppeling aan de KRW Verkenner.

### 2.2 Verbetering datasets

Voor het trainen van nieuwe neurale netwerken ten behoeve van EEE3 zijn de datasets voor de acht verschillende watertypen zoals gebruikt voor de verdere ontwikkeling van EEE2 in 2009 aangevuld en verbeterd. Tabel 2.1 geeft hiervan een overzicht.

De belangrijkste wijzigingen zijn:

- Vergroten van de datasets en daarmee ook vergroting van de dekking over Nederland.
- Vergroten van het aantal waarnemingen die zijn gebaseerd op KRW-toetsingen van echte biologische monsters. Kennisregels zijn daardoor meer op waarnemingen gebaseerd waardoor minder gebruik is gemaakt van expert kennis.

**Tabel 2.1 Aantal records in de datasets voor ontwikkeling EEE3 en aandeel monsters dat getoetst is aan de KRW-maatlatten (uit Limnodata, Piscaria of uit het KRW-portaal). Resterend aandeel ingevuld op basis van expert judgement. Tussen () het aantal records c.q. percentage getoetst bij EEE2.**

Watertypencluster	Aantal records	Percentage records gebaseerd op echte biologische monsters die aan KRW-maatlatten getoetst zijn.			
		Fytoplankton	Overige waterflora	Macrofauna	Vissen
Langzaam stromende beken	202 (239)		93 (77)	97 (99)	66 (43)
Snel stromende beken	195 (184)		56 (10)	97 (95)	54 (7)
Ondiepe meren	193 (180)	66 (25)	58 (30)	90 (95)	30 (22)
Diepe meren	128 (111)	80 (41)	39 (28)	85 (40)	21 (18)
Kanalen	198 (169)	66 (29)	74 (51)	94 (91)	41 (13)
Sloten	193 (197)		75 (52)	93 (89)	31 (20)
Zwak brakke wateren	203 (202)	93 (95)	59 (39)	96 (95)	15 (9)
Brakke tot zoute wateren	186 (158)	95 (95)	24 (20)	96 (96)	10 (4)

<sup>1</sup> Rapport 'Verdere ontwikkeling Expertsysteem Ecologische Effecten en evaluatie gebruik in de Ex ante evaluatie KRW', voor Planbureau voor de Leefomgeving, Royal Haskoning, 16 april 2009.

Voor zes watertypen zijn de datasets uitgebreid, terwijl twee grotere sets van het EEE2 (langzaam stromende beken en sloten) in omvang iets zijn afgenomen. Bij deze laatste twee datasets zijn de expertoordelen uit de sets geschrapt waar daar ruimte voor was. Criterium daarbij was dat daar waar op die posities van het gehele EKR-bereik in de dataset van EEE3 nu echte data beschikbaar waren de expertoordelen in de set van EEE2 zijn verwijderd.

De andere (meestal voorheen kleinere datasets) zijn in omvang toegenomen. Dit geldt vooral voor de kanalen en de brakke tot zoute wateren. Na deze update is voor alle watertypenclusters op de diepe meren na een dataset van voldoende omvang (circa 190-200) verkregen. Het aantal diepe meren dat in Nederland als waterlichaam is aangewezen is echter zo laag dat een dataset met vergelijkbare omvang niet mogelijk is op dit moment. Ondanks de relatief kleine dataset waren er voor alle kwaliteitselementen van dit type waterlichaam toch goede netwerken te trainen.

Het percentage waarnemingen dat gebaseerd is op getoetste monsters is in de meeste gevallen toegenomen vergeleken met de percentages zoals gebruikt bij EEE2. Bij fytoplankton en macrofauna van brakke tot zoute wateren is dit percentage gelijk gebleven ten opzicht van EEE2 en in geval van fytoplankton en macrofauna bij langzaam stromende beken en ondiepe meren is dat percentage licht afgenomen ten opzichte van EEE2. De reden voor deze afnames is dat extra monsters van andere kwaliteitselementen beschikbaar zijn gekomen waar geen gegevens over fytoplankton en/of macrofauna bij bekend waren. Het aandeel getoetste waarnemingen van macrofyten en vissen van de betreffende watertypen kon daarmee substantieel toenemen ten koste van een kleine afname bij fytoplankton en macrofauna.

## 2.3 Training neurale netwerken

Met behulp van de Matlab scripts die zijn opgesteld om de neurale netwerken voor EEE2 te trainen, zijn nieuwe neurale netwerken getraind met de datasets die ten opzichte van de sets in 2009 zijn aangevuld en verbeterd.

In tabel 2.2 is een overzicht gegeven van de watertypen, de verschillende neurale netwerken en het aantal trainingen dat is uitgevoerd.

Voor elke trainingscyclus is telkens willekeurig 80% van de records in die datasets gebruikt, waarbij validatie heeft plaats gevonden met de overige 20%. Tijdens de training worden de netwerkfout in de trainingsset en de netwerkfout in de validatieset continu gecontroleerd. Beide fouten nemen tijdens de eerste fase van de training gewoonlijk af. Wanneer het netwerk zichzelf overtraint ('overfitting'), zal de fout in de validatieset meestal groter worden. De netwerkparameters worden dan vastgelegd op het minimum van de validatiefout.

De training- en validatiesets van de uiteindelijk vastgestelde neurale netwerken zijn vervolgens ook beschikbaar gesteld voor het trainen van de andere twee typen expertsystemen (PUN's bij Witteveen en Bos en regressieboomanalyses bij Planbureau voor de Leefomgeving).

Voor een specifiek neuraal netwerk zijn telkens twee trainingscycli ingezet, te weten:

1. Een cyclus gericht op maximalisatie van de correlatiecoëfficiënt  $R^2$  tussen voorspelde en gemeten EKR's.
2. Een vervolgcyclus met optimalisatie van de richtingscoëfficiënt RC van de verkregen lineaire regressielijn van voorspelde en gemeten EKR's.

Na elke training in een trainingscyclus met een zekere 80% / 20% verdeling van de dataset wordt de berekende  $R^2$  vergeleken met  $R^2$  van de voorgaande training. Indien dan een hogere  $R^2$  wordt gevonden, worden de daarvoor gebruikte trainings- en validatieset bewaard.

Na deze training op maximale en dan vastgelegde  $R^2$  wordt in de vervolgcyclus doorgetraind op de optimale waarde 1 van de RC. Daarbij wordt ook weer telkens met een nieuwe verdeling van de dataset getraind maar wel telkens met de in de vorige cyclus vastgestelde  $R^2$  als minimum.

Na beide cycli wordt dan een getraind netwerk verkregen dat is gemaximaliseerd op  $R^2$  en is geoptimaliseerd op RC. De uiteindelijke daarbij behorende trainings- en validatiesets zijn daarna verstuurd aan de trainers van de andere twee expertsystemen.

**Tabel 2.2: Watertypen, neurale netwerken en trainingen**

Watertype		Neuraal netwerk voor	Aantal trainingen
Codering	Omschrijving		
R4, R5, R6 en R12	Langzaam stromende beken	<ul style="list-style-type: none"> <li>• Overige waterflora</li> <li>• Macrofauna</li> <li>• Vissen</li> </ul>	Elk netwerk 1.000 op $R^2$ en 10.000 op RC
R13, R14, R15, R17 en R18	Snel stromende beken	<ul style="list-style-type: none"> <li>• Overige waterflora</li> <li>• Macrofauna</li> <li>• Vissen</li> </ul>	Elk netwerk 1.000 op $R^2$ en 10.000 op RC
M20	Diepe meren	<ul style="list-style-type: none"> <li>• Fytoplankton</li> <li>• Overige waterflora</li> <li>• Macrofauna</li> <li>• Vissen</li> </ul>	Elk netwerk 1.000 op $R^2$ en 10.000 op RC
M14, M23 en M27	Ondiepe meren	<ul style="list-style-type: none"> <li>• Fytoplankton</li> <li>• Overige waterflora</li> <li>• Macrofauna</li> <li>• vissen</li> </ul>	Elk netwerk 1.000 op $R^2$ en 50.000 op RC
M3, M4, M6a/b, M7a/b en M10	Kanalen	<ul style="list-style-type: none"> <li>• fytoplankton</li> <li>• overige waterflora</li> <li>• macrofauna</li> <li>• vissen</li> </ul>	Elk netwerk 5.000 op $R^2$ en 50.000 op RC
M1a/b, M2 en M8	Sloten	<ul style="list-style-type: none"> <li>• overige waterflora</li> <li>• macrofauna</li> <li>• vissen</li> </ul>	Elk netwerk 1.000 op $R^2$ en 10.000 op RC
M30	Zwak brakke wateren	<ul style="list-style-type: none"> <li>• fytoplankton</li> <li>• overige waterflora</li> <li>• macrofauna</li> <li>• vissen</li> </ul>	Elk netwerk 5.000 op $R^2$ en 50.000 op RC
M31	Brak tot zoute wateren	<ul style="list-style-type: none"> <li>• fytoplankton</li> <li>• overige waterflora</li> <li>• macrofauna</li> <li>• vissen</li> </ul>	Elk netwerk 5.000 op $R^2$ en 50.000 op RC

Tabel 2.3 geeft een overzicht van de stuurfactoren die bij de training van de verschillende watertypen als invoerparameter zijn gebruikt.

**Tabel 2.3: Watertypen en de gekozen hydromorfologische en chemische stuurvariabelen**

Watertype	Langzaam stromende beken	Snel stromende beken	Diepe meren	Ondiepe meren	Kanalen en vaarten	Sloten	Zwak brakke wateren	Brakke tot zoute wateren
Stuurvariabele								
Oeverinrichting			X	X	X	X	X	X
Peildynamiek			X	X	X	X	X	X
Onderhoud					X	X	X	X
Connectiviteit							X	X
Meandering	X	X						
Verstuwing	X	X						
Beschaduwing	X	X						
Scheepvaart					X			
BZV	X	X						
Chloride jgem							X	X
Totaal P zgem	X	X	X	X	X	X	X	X
Totaal N zgem	X	X	X	X	X	X	X	X

Op de volgende pagina zijn in tabel 2.4 de waarden van de stuurvariabelen per watertypencluster verder uitgewerkt naar klasse-indeling en eenheden.

**Tabel 2.4a: Stuurvariabelen langzaam en snel stromende beken uitgesplitst naar hydromorfologische en chemische parameters**

Hydromorfologische parameters	Klassen	Waarden en omschrijving
Meandering <sup>2</sup>	5	1=recht+normprofiel, 2=gestrekt+natuurlijker dwarsprofiel, 3=zwak slingerend, 4=slingerend, 5= vrij meanderend
Verstuwing	3	1=sterk gestuwd zonder vistrappen, 2=gestuwd met vistrappen, 3=ongestuwd
Beschaduwing	3	1=onbeschaduwd zonder ruigte op de oevers, 2=gedeeltelijk beschaduwd of ruigte op de oever en 3=grotendeels of geheel beschaduwd (opgaande begroeiing/bos)
Chemische parameters	Eenheid	
BZV*	mg O <sub>2</sub> /l	Zomergemiddelde (april-september), maat voor organische belasting
Totaal P	mg P/l	Zomergemiddelde (april-september), maat voor eutrofiëring
Totaal N	mg N/l	Zomergemiddelde (april-september), maat voor eutrofiëring

\*BZV is Biochemisch Zuurstofverbruik (maat voor de organische belasting van een water)

<sup>2</sup> De mate van meandering wordt uitgedrukt in sinuositeit. Sinuositeit is de lengte van de beek gedeeld door de lengte van het stroomdal. Op basis van sinuositeit kunnen 5 klassen in meandering worden onderscheiden, te weten: 1. Recht (sinuositeit 1) 2. Gestrekt (sinuositeit 1.01-1.05) 3. Zwak slingerend (sinuositeit 1.06-1.25) 4. Slingerend (sinuositeit 1.26-1.5) 5. Meanderend (sinuositeit >1.5)

**Tabel 2.4a: Stuurvariabelen zoete ondiepe en diepe meren uitgesplitst naar hydromorfologische en chemische parameters**

Hydromorfologische parameters	Klassen	Waarden en omschrijving
Oeverinrichting	3	1=beschoeid of steil en onbegroeid, 2=riet/helofyten, 3=moeras+riet/helofyten
Peildynamiek	3	1=tegennatuurlijk, 2=stabiel, 3=natuurlijk
<b>Chemische parameters</b>	<b>Eenheid</b>	
Totaal P	mg P/l	Zomergemiddelde (april-september), maat voor eutrofiëring
Totaal N	mg N/l	Zomergemiddelde (april-september), maat voor eutrofiëring

**Tabel 2.4b: Stuurvariabelen kanalen uitgesplitst naar hydromorfologische en chemische parameters**

Hydromorfologische parameters	Klassen	Waarden en omschrijving
Oeverinrichting	3	1=beschoeid, 2=steil, 3=flauw/moerassig (NVO*)
Peildynamiek	3	1=tegennatuurlijk, 2=stabiel, 3=natuurlijk
Onderhoud <sup>3</sup>	2	1=intensief, 2=extensief
Scheepvaart	2	1=intensief bevaren, 2 niet of nauwelijks bevaren
<b>Chemische parameters</b>	<b>Eenheid</b>	
Totaal P	mg P/l	Zomergemiddelde (april-september), maat voor eutrofiëring
Totaal N	mg N/l	Zomergemiddelde (april-september), maat voor eutrofiëring

\*NVO staat voor Natuurvriendelijke oevers

**Tabel 2.4d: Stuurvariabelen sloten uitgesplitst naar hydromorfologische en chemische parameters**

Hydromorfologische parameters	Klassen	Waarden en omschrijving
Oeverinrichting	3	1=beschoeid, 2=steil, 3=flauw/moerassig (NVO*)
Peildynamiek	3	1=tegennatuurlijk, 2=stabiel, 3=natuurlijk
Onderhoud	2	1=intensief, 2=extensief
<b>Chemische parameters</b>	<b>Eenheid</b>	
Totaal P	mg P/l	Zomergemiddelde (april-september), maat voor eutrofiëring
Totaal N	mg N/l	Zomergemiddelde (april-september), maat voor eutrofiëring

\*NVO staat voor Natuurvriendelijke oevers

<sup>3</sup> De mate van intensiteit van het onderhoud was meestal alleen kwalitatief beschikbaar. Wanneer werd aangegeven dat voor de KRW het onderhoud extensief zou gaan worden uitgevoerd als maatregel, is aangenomen dat het in de huidige situatie intensief is.

**Tabel 2.4e: Stuurvariabelen zwak brakke en brakke tot zoute wateren uitgesplitst in hydromorfologische en chemische parameters**

<b>Hydromorfologische parameters</b>	<b>Klassen</b>	<b>Waarden en omschrijving</b>
Oeverinrichting	3	1=beschoeid, 2=steil, 3=flauw/moerassig (NVO*)
Peildynamiek	3	1=tegennatuurlijk, 2=stabiel, 3=natuurlijk
Onderhoud	2	1=intensief, 2=extensief
Connectiviteit	3	1=geïsoleerd, 2=periodiek geïsoleerd, 3=open verbinding
<b>Chemische parameters</b>	<b>Eenheid</b>	
Chloride gehalte	mg Cl/l	Zomergemiddelde (april-september), maat voor verzoeting
Totaal P	mg P/l	Zomergemiddelde (april-september), maat voor eutrofiëring
Totaal N	mg N/l	Zomergemiddelde (april-september), maat voor eutrofiëring

\*NVO staat voor Natuurvriendelijke oevers

### 3 RESULTATEN

Hierna worden in tabel 3.1 een samenvatting gegeven van de resultaten van de netwerktrainingen voor de kwaliteitselementen van de verschillende watertypen.

Naast correlatiecoëfficiënt  $R^2$  en richtingscoëfficiënt RC zijn in de tabel opgenomen:

- De gemiddelde voorspelfout RMSE (*Rooted Mean Square Error*). Deze is als volgt gedefinieerd:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

In deze formule staat  $y_i$  voor de gemeten EKR en  $\hat{y}_i$  voor de met het expertsysteem voorspelde EKR in water  $i$ . Bij een perfecte voorspelling geldt dat de RMSE-waarde gelijk is aan nul.

- Het percentage voorspelde EKR's dat minder dan 0,1 afwijkt van gemeten EKR.
- De *Coëfficiënt of Determination* (CoD) van de validatieset. De CoD is een maat voor de voorspellingswaarde van een neurale netwerk. De waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen. De formule voor de CoD is:

$$CoD = 1 - \frac{RMSE^2}{var(y_i)}$$

Met 'var( $y_i$ )' de variantie van de metingen in de validatieset.

Voor de gedetailleerde resultaten van de getrainde netwerken voor de verschillende watertypen wordt verwezen naar bijlagen 1 t/m 8.

De verkregen resultaten zijn goed vergelijkbaar met die van EEE2. In algemene zin zijn de EEE3-resultaten enigszins beter, zie RMSE en afwijking EKR in tabel 3.1. De resultaten van beide EEE-versies zijn statistisch niet verder met elkaar vergeleken omdat de CoD's bij EEE2 niet zijn bepaald.

De met deze netwerken gecompileerde executable EEE3.exe dient als invoer een inputdata.csv bestand te hebben zoals weergegeven in tabel 3.2.

De beide uitvoerbestanden van deze EEE3 tool – outputdata.csv en outliers.csv - zijn weergegeven in tabel 3.3 en 3.4.



**Tabel 3.1: Statistische kenmerken van getrainde neurale netwerken voor de verschillende watertypen tussen () waarden van EEE2**

Watertype	Kwaliteitselementen	Trainingset				Validatieset
		R <sup>2</sup>	RC	RMSE	% afwijking EKR < 0.1	CoD
Langzaam stromende beken	Overige waterflora	0.81	0.83	0.066 (0.100)	88 (71)	0.56
	Macrofauna	0.88	0.91	0.060 (0.083)	91 (82)	0.74
	Vissen	0.87	0.91	0.061 (0.079)	90 (85)	0.84
Snel stromende beken	Overige waterflora	0.91	0.95	0.063 (0.070)	87 (85)	0.76
	Macrofauna	0.75	0.75	0.104 (0.100)	68 (66)	0.60
	Vissen	0.82	0.82	0.093 (0.110)	72 (61)	0.84
Diepe meren	Macrofyten	0.66	0.64	0.158 (0.094)	54 (62)	-0.06
	Overige waterflora	0.87	0.87	0.081 (0.120)	75 (73)	0.56
	Macrofauna	0.86	0.88	0.068 (0.130)	89 (59)	0.56
	Vissen	0.85	0.89	0.071 (0.085)	85 (79)	0.69
Ondiepe meren	Macrofyten	0.78	0.77	0.117 (0,086)	61 (82)	0.64
	Overige waterflora	0.84	0.85	0.086 ((0.093)	79 (84)	0.42
	Macrofauna	0.73	0.75	0.089 (0.100)	72 (58)	0.57
	Vissen	0.83	0.87	0.086 (0.071)	75 (84)	0.35
Kanalen	Macrofyten	0.86	0.86	0.076 (0.110)	82 (67)	0.64
	Overige waterflora	0.68	0.75	0.099 (0.088)	67 (78)	-0.21
	Macrofauna	0.78	0.78	0.090 (0.120)	72 (61)	0.27
	Vissen	0.81	0.81	0.070 (0.069)	86 (83)	0.23
Sloten	Overige waterflora	0.79	0.76	0.089 (0.110)	73 (64)	0.49
	Macrofauna	0.77	0.79	0.093 (0.120)	69 (60)	0.49
	Vissen	0.89	0.91	0.057 (0.091)	93 (82)	0.71
Zwak brakke wateren	Macrofyten	0.70	0.79	0.119 (0.130)	72 (60)	-0.18
	Overige waterflora	0.81	0.84	0.084 (0.150)	81 (61)	0.55
	Macrofauna	0.66	0.68	0.104 (0,120)	75 (71)	0.06
	Vissen	0.71	0.79	0.091 (0.089)	77 (74)	0.35
Brak tot zoute wateren	Macrofyten	0.74	0.82	0.100 (0,100)	67 (72)	-0.13
	Overige waterflora	0.84	0.88	0.063 (0.140)	88 (58)	0.45
	Macrofauna	0.68	0.70	0.123 (0.150)	58 (61)	0.40
	Vissen	0.68	0.75	0.086 (0.098)	83 (71)	0.36

**Tabel 3.2: Voorbeeld van invoerbestand EEE3-tool**

ID	Waterbody	Watertype	Meanderi	Shading	Weirs	Banks	LevelControl	Maintenance	Shipping	Connectivity	BOD	Chloride	P-total	N-total
a1	a	R4	10	0.5	0.3	NA	NA	NA	NA	NA	1.7	NA	0.14	4
a2	b	R5	2	1	2	NA	NA	NA	NA	NA	1.5	NA	0.2	2
a3	c	R6	7	7	7	NA	NA	NA	NA	NA	1.5	NA	0.1	3
a4	d	R12	2	3	1	NA	NA	NA	NA	NA	1.8	NA	0.15	2
a5	e	R13	2	2	2	NA	NA	NA	NA	NA	1.2	NA	0.2	1
a6	f	R14	1	3	3	NA	NA	NA	NA	NA	1	NA	0.1	2
a7	g	R15	3	1	2	NA	NA	NA	NA	NA	1.5	NA	0.2	2
a8	h	R17	3	3	1	NA	NA	NA	NA	NA	1.5	NA	0.1	3
a9	f	R18	1	6	4	NA	NA	NA	NA	NA	1.8	NA	0.15	2
a10	g	M1a	NA	NA	NA	1	1	2	NA	NA	NA	NA	0.1	3
a11	h	M1b	NA	NA	NA	2	1	1	NA	NA	NA	NA	0.15	2
a12	i	M2	NA	NA	NA	6	6	6	NA	NA	NA	NA	0.2	5
a13	j	M8	NA	NA	NA	2	2	1	NA	NA	NA	NA	0.1	4
a14	k	M3	NA	NA	NA	1	2	3	2	NA	NA	NA	0.2	4
a15	l	M4	NA	NA	NA	2	1	2	1	NA	NA	NA	0.1	3
a16	m	M6a	NA	NA	NA	1	2	3	2	NA	NA	NA	0.15	2
a17	n	M6b	NA	NA	NA	2	1	1	1	NA	NA	NA	0.2	3
a18	o	M7a	NA	NA	NA	1	2	2	2	NA	NA	NA	0.25	5
a19	p	M7b	NA	NA	NA	2	1	2	1	NA	NA	NA	0.3	5
a20	q	M10	NA	NA	NA	1	2	3	2	NA	NA	NA	0.2	4
a21	r	M14	NA	NA	NA	2	1	NA	NA	NA	NA	NA	0.1	3
a22	s	M23	NA	NA	NA	1	2	NA	NA	NA	NA	NA	0.2	2
a23	t	M27	NA	NA	NA	2	1	NA	NA	NA	NA	NA	0.1	4
a24	u	M20	NA	NA	NA	1	2	NA	NA	NA	NA	NA	0.2	4
a25	v	M30	NA	NA	NA	2	2	2		2		2000	0.5	6
a26	w	M31	NA	NA	NA	1	3	2		3		10000	0.7	3

Tabel 3.3: Voorbeeld van uitvoerbestand outputdata.csv

ID	Waterbody	Watertype	Meandering	Shading	Weirs	Banks	LevelControl	Maintenance	Shipping	Connectivity	BOD	Chloride	P-total	N-total	EQRfytoplankton	EQRaquaflora	EQRmacroinv	EQRfish
a1	a	R4	10	0.5	0.3	NA	NA	NA	NA	NA	1.7	NA	0.14	4	NA	0.51284	0.66277	0.38281
a2	b	R5	2	1	2	NA	NA	NA	NA	NA	1.5	NA	0.2	2	NA	0.49136	0.30431	0.31885
a3	c	R6	7	7	7	NA	NA	NA	NA	NA	1.5	NA	0.1	3	NA	0.93785	0.86214	0.91209
a4	d	R12	2	3	1	NA	NA	NA	NA	NA	1.8	NA	0.15	2	NA	0.52964	0.53334	0.52623
a5	e	R13	2	2	2	NA	NA	NA	NA	NA	1.2	NA	0.2	1	NA	0.50315	0.47833	0.43354
a6	f	R14	1	3	3	NA	NA	NA	NA	NA	1	NA	0.1	2	NA	0.50772	0.38774	0.44606
a7	g	R15	3	1	2	NA	NA	NA	NA	NA	1.5	NA	0.2	2	NA	0.5907	0.48234	0.47787
a8	h	R17	3	3	1	NA	NA	NA	NA	NA	1.5	NA	0.1	3	NA	0.61543	0.61575	0.44075
a9	f	R18	1	6	4	NA	NA	NA	NA	NA	1.8	NA	0.15	2	NA	0.43841	0.37512	0.42047
a10	g	M1a	NA	NA	NA	1	1	2	NA	NA	NA	NA	0.1	3	NA	0.3321	0.37495	0.47325
a11	h	M1b	NA	NA	NA	2	1	1	NA	NA	NA	NA	0.15	2	NA	0.29899	0.37095	0.47713
a12	i	M2	NA	NA	NA	6	6	6	NA	NA	NA	NA	0.2	5	NA	0.85665	0.64703	0.95021
a13	j	M8	NA	NA	NA	2	2	1	NA	NA	NA	NA	0.1	4	NA	0.33927	0.49802	0.5735
a14	k	M3	NA	NA	NA	1	2	3	2	NA	NA	NA	0.2	4	0.47211	0.37762	0.38936	0.43983
a15	l	M4	NA	NA	NA	2	1	2	1	NA	NA	NA	0.1	3	0.93493	0.1279	0.29451	0.97472
a16	m	M6a	NA	NA	NA	1	2	3	2	NA	NA	NA	0.15	2	0.52034	0.37925	0.57012	0.5285
a17	n	M6b	NA	NA	NA	2	1	1	1	NA	NA	NA	0.2	3	0.92999	0.077455	0.1194	0.97314
a18	o	M7a	NA	NA	NA	1	2	2	2	NA	NA	NA	0.25	5	0.4608	0.37596	0.16707	0.34282
a19	p	M7b	NA	NA	NA	2	1	2	1	NA	NA	NA	0.3	5	0.92203	0.31511	0.27583	0.95844
a20	q	M10	NA	NA	NA	1	2	3	2	NA	NA	NA	0.2	4	0.47211	0.37762	0.38936	0.43983
a21	r	M14	NA	NA	NA	2	1	NA	NA	NA	NA	NA	0.1	3	0.39174	0.33884	0.39231	0.21369
a22	s	M23	NA	NA	NA	1	2	NA	NA	NA	NA	NA	0.2	2	0.33377	0.13267	0.39523	0.17629
a23	t	M27	NA	NA	NA	2	1	NA	NA	NA	NA	NA	0.1	4	0.37261	0.32579	0.40229	0.18618
a24	u	M20	NA	NA	NA	1	2	NA	NA	NA	NA	NA	0.2	4	0.25174	0.18503	0.29217	0.27509
a25	v	M30	NA	NA	NA	2	2	2		2		2000	0.5	6	0.3445	0.21557	0.57727	0.50995
a26	w	M31	NA	NA	NA	1	3	2		3		10000	0.7	3	0.87426	0.44087	0.74205	0.85524

**Tabel 3.4: Voorbeeld van uitvoerbestand outliers.csv**

If environmental factor has value 1, its value is outside the data limits of the trainingset of the corresponding neural network											
Outliers slow flowing brook											
ID	Meandering	Shading	Weirs	BOD	P-total	N-total					
a1	1	1	1	0	0	0					
Outliers fast flowing brook											
ID	Meandering	Shading	Weirs	BOD	P-total	N-total					
a7	0	0	0	0	0	0					
Outliers fast flowing brook											
ID	Meandering	Shading	Weirs	BOD	P-total	N-total					
a8	0	0	0	0	0	0					
Outliers fast flowing brook											
ID	Meandering	Shading	Weirs	BOD	P-total	N-total					
a9	0	1	1	0	0	0					
Outliers Ditch											
ID	Banks	Level control	Maintena	P-total	N-total						
a10	0	0	0	0	0						
Outliers Ditch											
ID	Banks	Level control	Maintena	P-total	N-total						
a11	0	0	0	0	0						
Outliers Ditch											
ID	Banks	Level control	Maintena	P-total	N-total						
a12	1	1	1	0	0						

## **Bijlage 1**

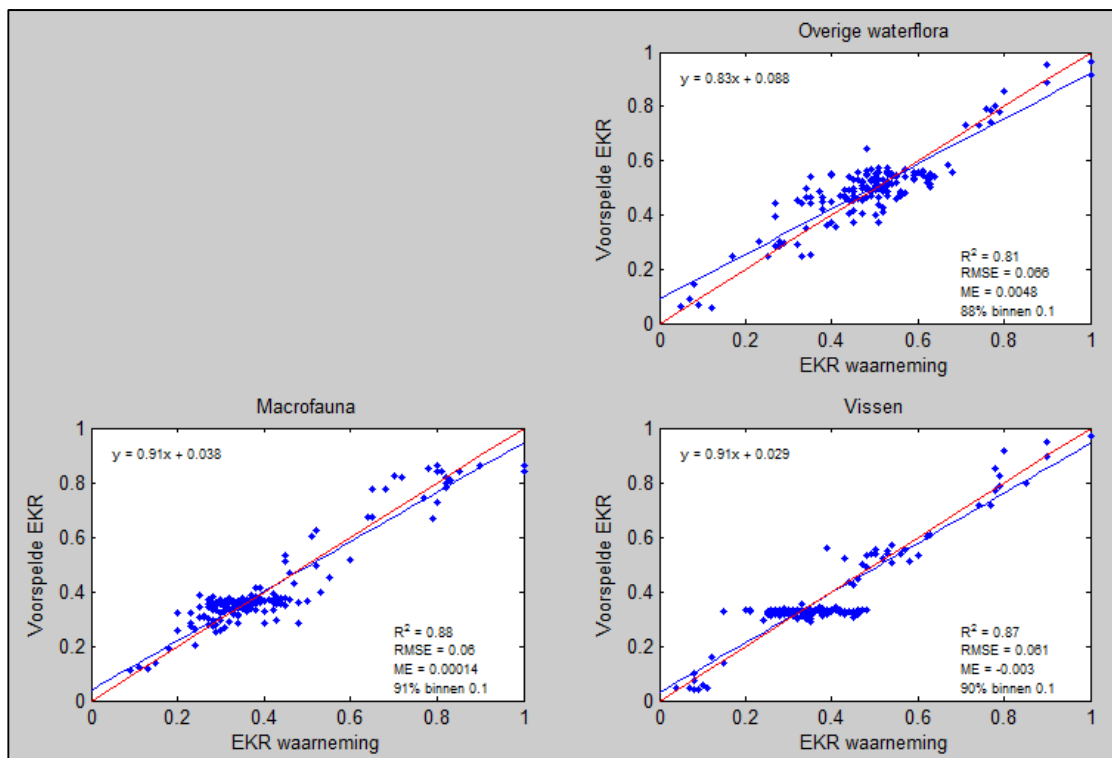
### **Beschrijving NN langzaam stromende beken**



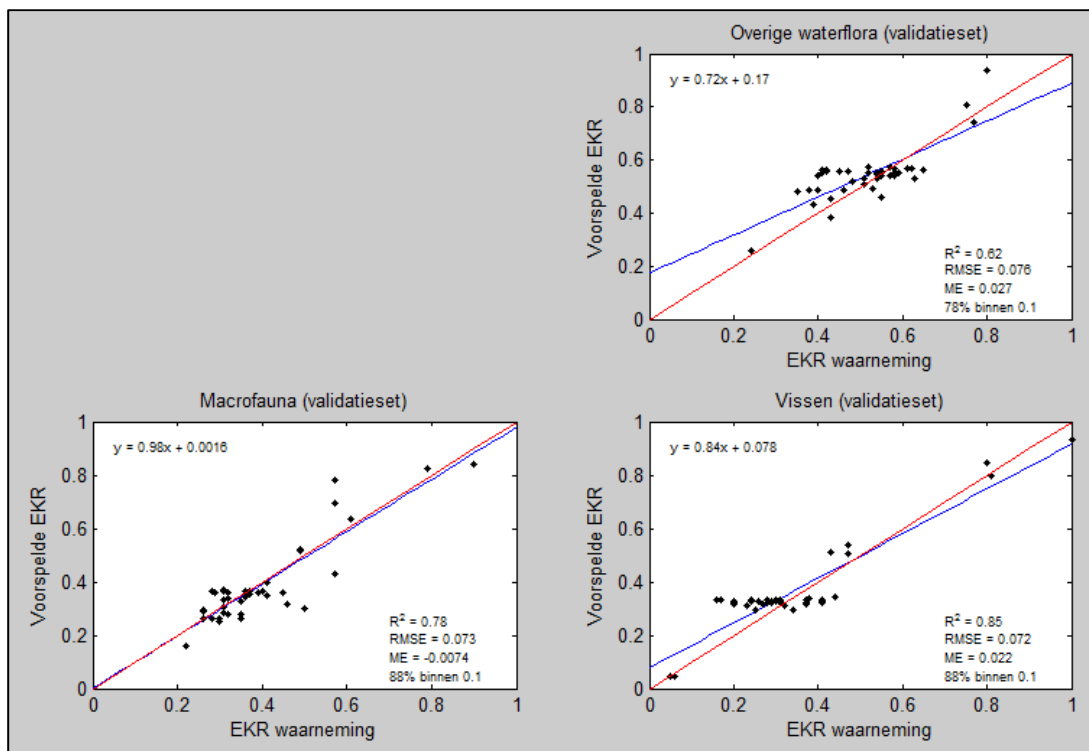
### Neuraal netwerken (EEE3) voor watertypen R4 R5 R6 R12 Langzaam stromende beken

- Kwaliteitselementen (EKR): overige waterflora, macrofauna en vissen
- Stuurfactoren: meandering, beschaduwing, verstuwing, BZV, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 202
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 1.000 x op  $R^2$  en 10.000 x op richtingscoëfficiënt

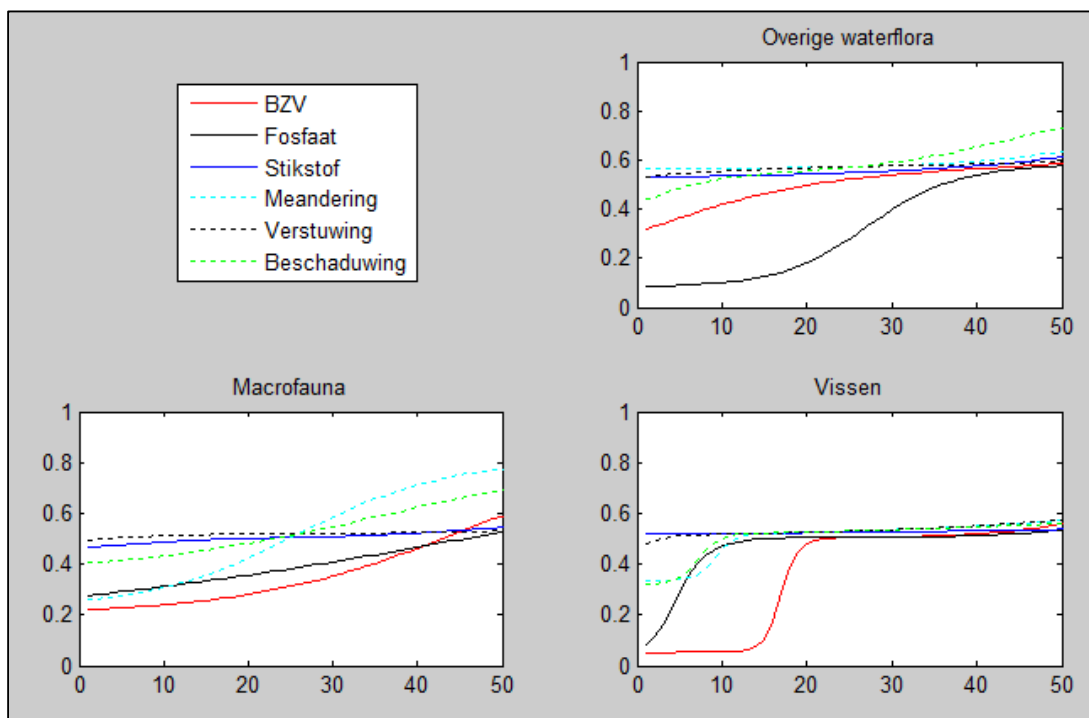
#### Resultaat neuraal netwerk op trainingset



Resultaat neurale netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren





*Aantal goede netwerken*

- Overige waterflora: 280
- Macrofauna: 203
- Vissen: 296

*Statistische gegevens*

Parameter	Betekenis	Overige waterflora	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination	0.56	0.74	0.84
R	Correlatie coëfficiënt	0.90	0.94	0.93
R2	Kwadraat corr coëff	0.81	0.88	0.87
RC	Richtingscoëfficiënt regressielijn	0.83	0.91	0.91
Std RC	Standaard afwijking van RC	0.010	0.011	0.011
Int	Intercept van regressielijn	0.088	0.038	0.029
ME	Gemiddelde fout	0.005	0.0001	-0.003
MSE	Gemiddelde van gekwadrateerde fouten	0.004	0.004	0.004
RMSE	Standaard deviatie (wortel van MSE)	0.066	0.060	0.061
% binnen 0,1 EKR		88	91	90



## **Bijlage 2**

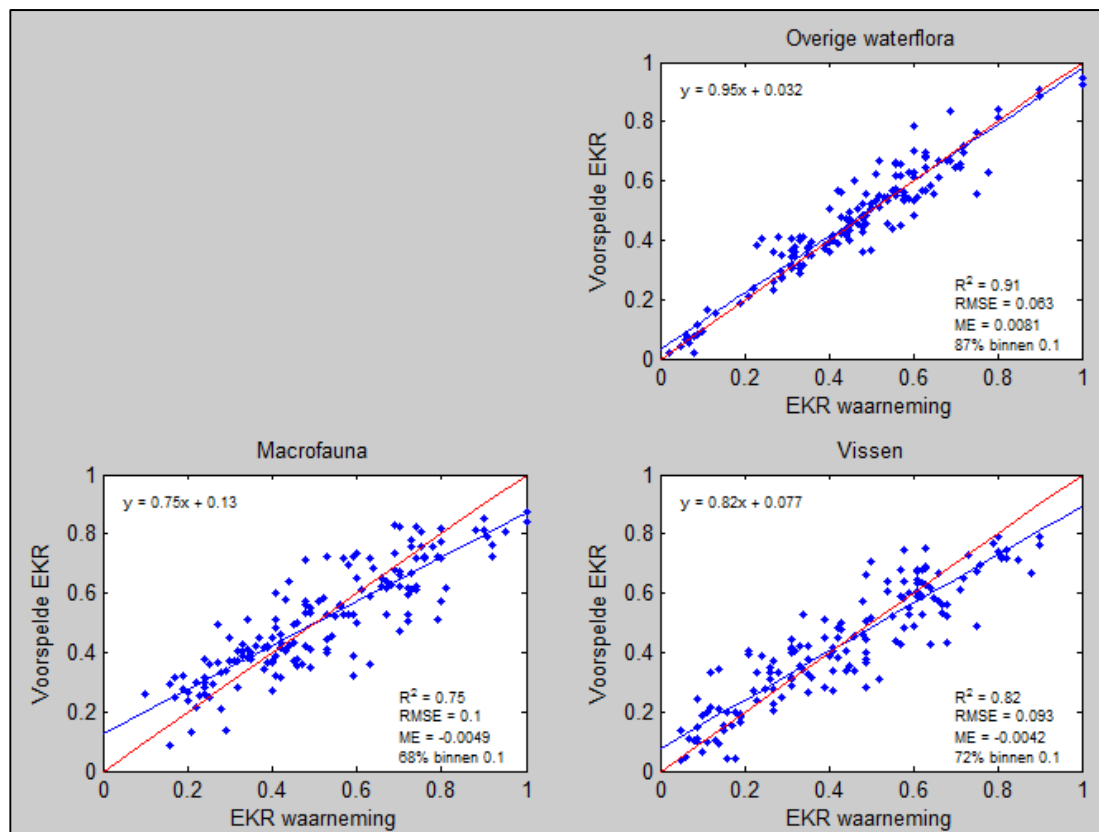
### **Beschrijving NN snel stromende beken**



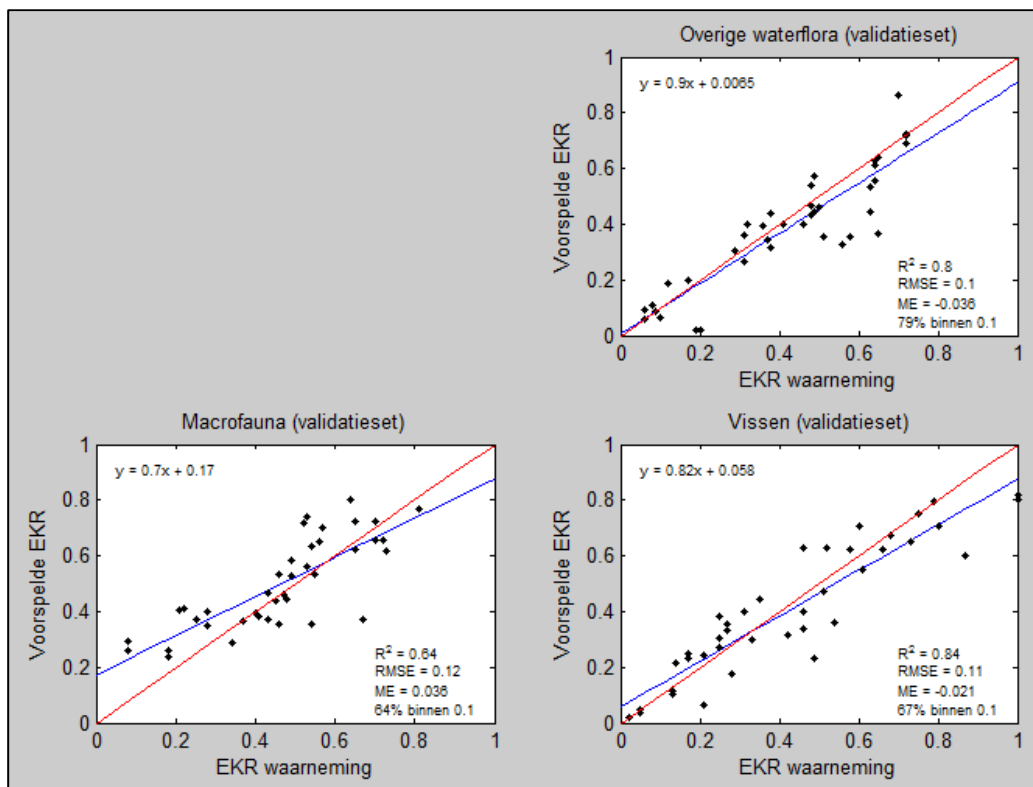
### Neuraal netwerken (EEE3) voor watertypen R13 R14 R15 R17 R18 Snelstromende beken

- Kwaliteitselementen (EKR): overige waterflora, macrofauna en vissen
- Stuurfactoren: meandering, beschaduwing, verstuwing, BZV, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 195
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 1.000 x op  $R^2$  en 10.000 x op richtingscoëfficiënt

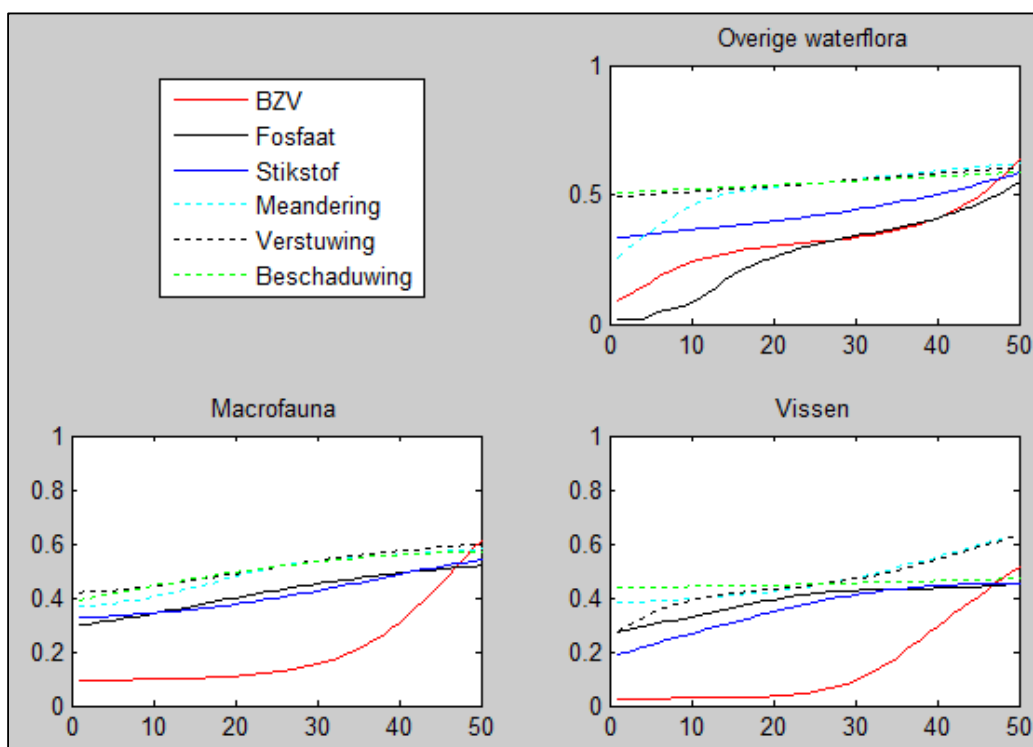
#### Resultaat neuraal netwerk op trainingset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Overige waterflora: 290
- Macrofauna: 83
- Vissen: 127

*Statistische gegevens*

Parameter	Betekenis	Overige waterflora	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	0.76	0.60	0.84
R	Correlatie coëfficiënt	0.95	0.87	0.91
R2	Kwadraat corr coëff	0.91	0.75	0.82
RC	Richtingscoëfficiënt regressielijn	0.95	0.75	0.82
Std RC	Standaard afwijking van RC	0.010	0.016	0.016
Int	Intercept van regressielijn	0.032	0.125	0.077
ME	Gemiddelde fout	0.008	-0.005	-0.004
MSE	Gemiddelde van gekwadrateerde fouten	0.004	0.011	0.009
RMSE	Standaard deviatie (wortel van MSE)	0.063	0.104	0.093
% binnen 0,1 EKR		87	68	72

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.





## **Bijlage 3**

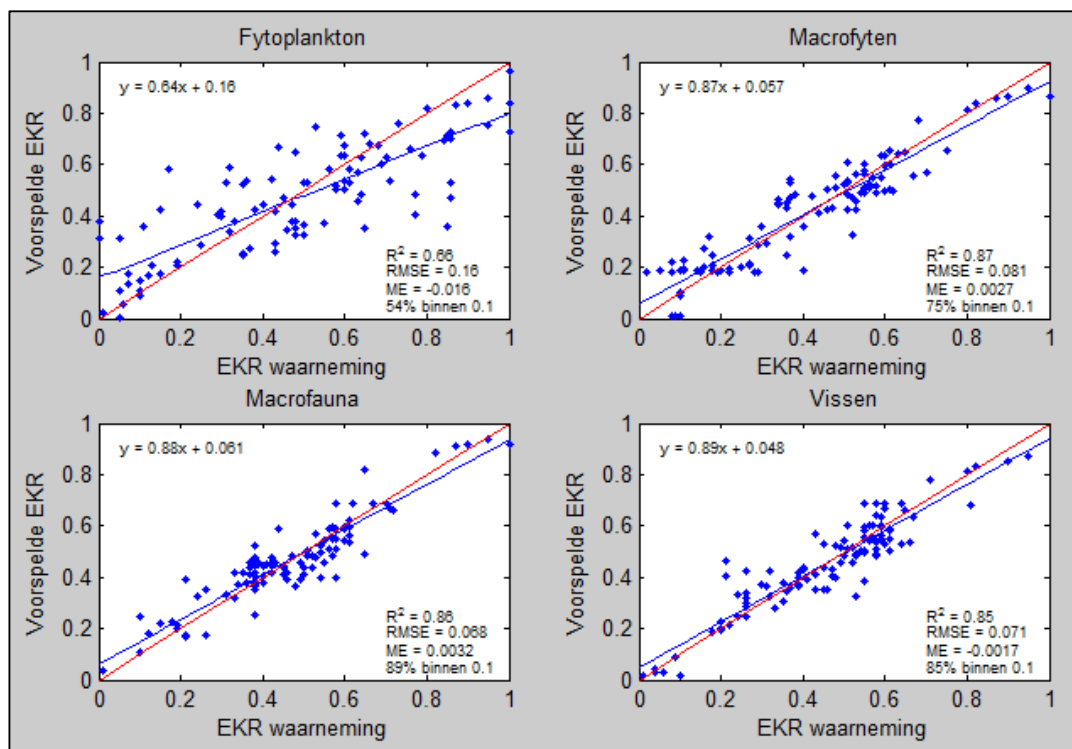
### **Beschrijving NN diepe meren**



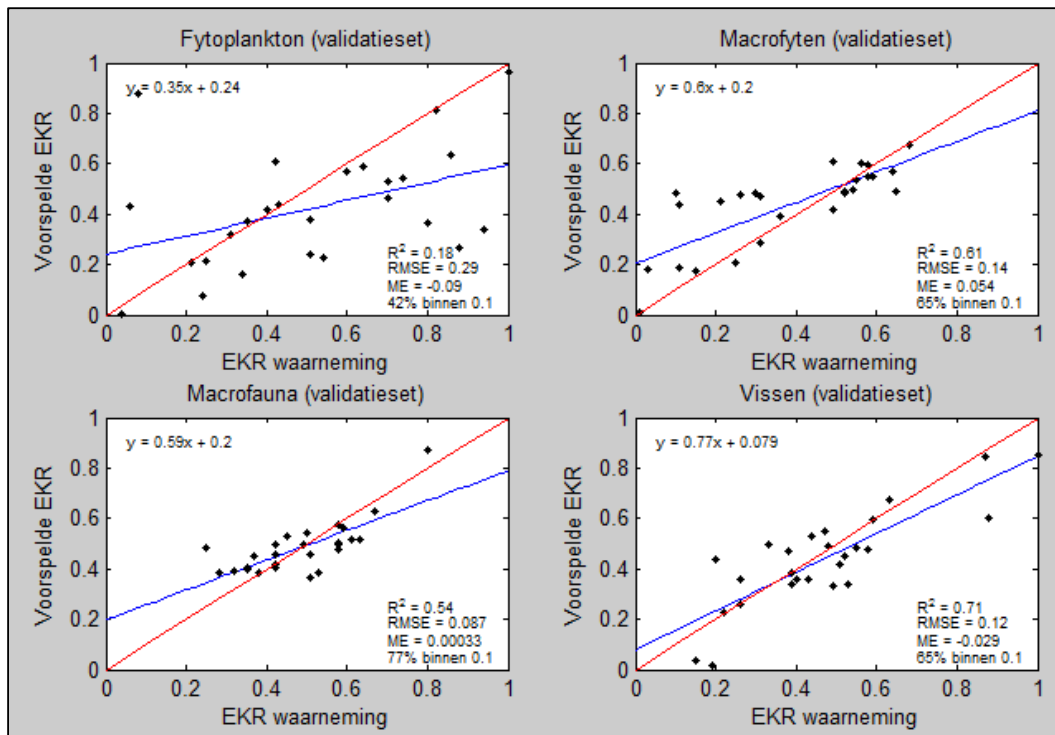
### Neuraal netwerken (EEE3) voor watertypen M16 en M20 Diepe meren

- Kwaliteitselementen (EKR): fytoplankton, macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peilbeheer, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 128
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 1.000 x op  $R^2$  en 10.000 x op richtingscoëfficiënt

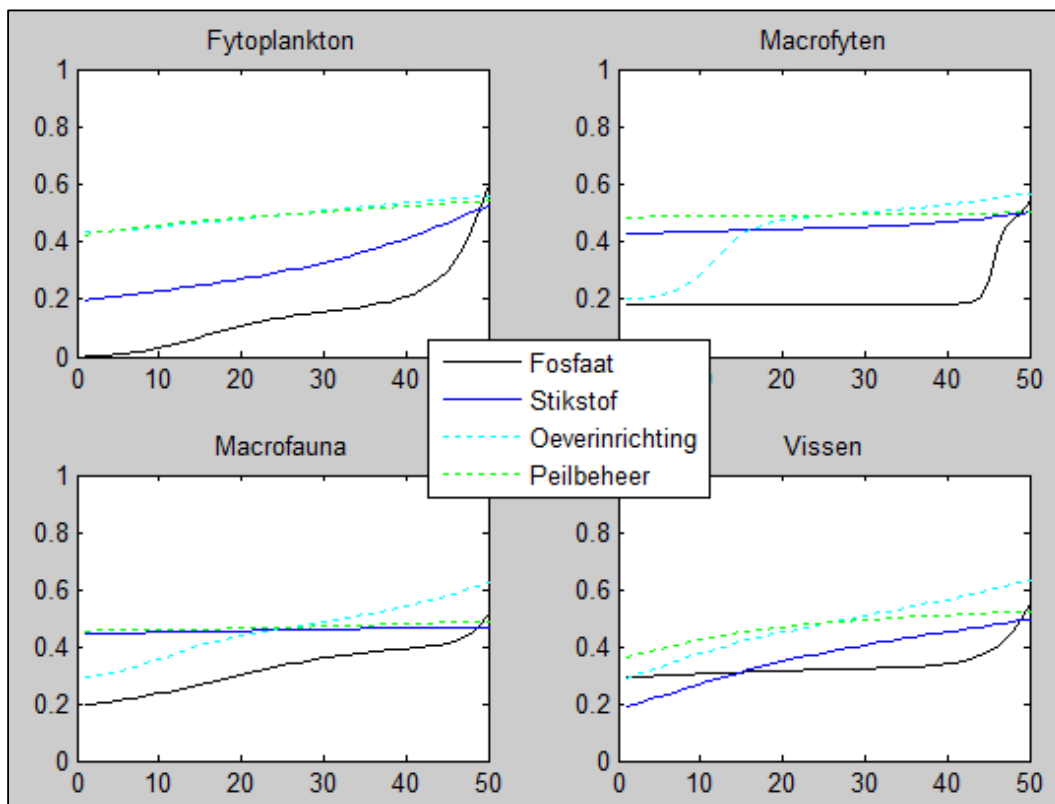
#### Resultaat neuraal netwerk op trainingset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Fytoplankton: 104
- Macrofyten: 480
- Macrofauna: 718
- Vissen: 732

*Statistische gegevens*

Parameter	Betekenis	Fytoplankton	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determinatie <sup>1</sup>	-0.06	0.56	0.56	0.69
R	Correlatie coëfficiënt	0.82	0.93	0.92	0.92
R2	Kwadraat corr coëff	0.66	0.87	0.86	0.85
RC	Richtingscoëfficiënt regressielijn	0.64	0.87	0.88	0.89
Std RC	Standaard afwijking van RC	0.030	0.017	0.014	0.014
Int	Intercept van regressielijn	0.158	0.057	0.061	0.048
ME	Gemiddelde fout	-0.016	0.003	0.003	-0.002
MSE	Gemiddelde van gekwadeerde fouten	0.025	0.005	0.005	0.005
RMSE	Standaard deviatie (wortel van MSE)	0.158	0.081	0.068	0.071
% binnen 0,1 EKR		54	75	89	85

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.



## **Bijlage 4**

### **Beschrijving NN ondiepe meren**

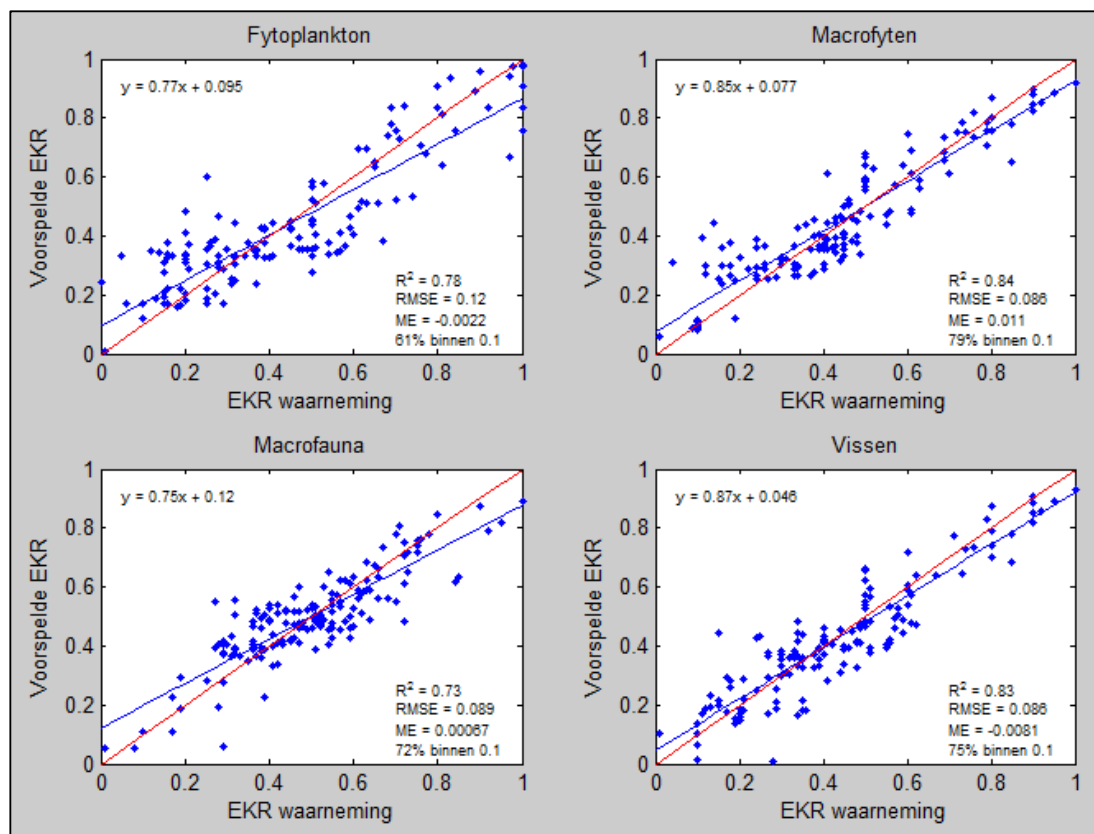




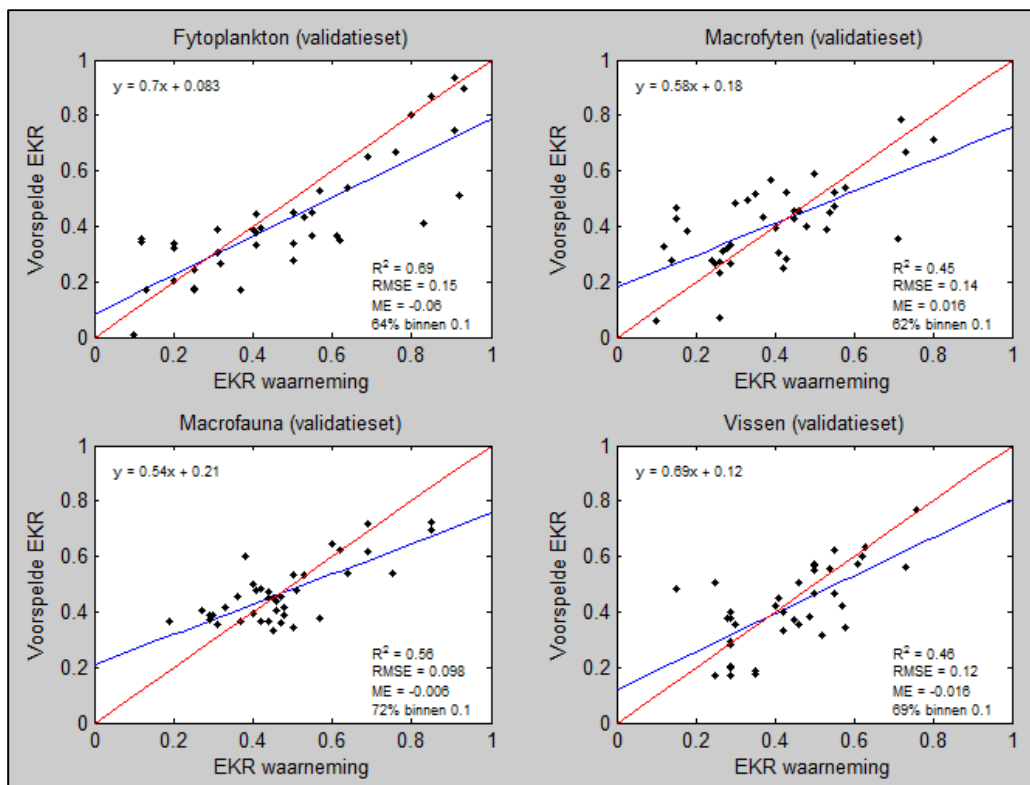
### Neuraal netwerken (EEE3) voor watertypen M14 M23 en M27 Ondiepe meren

- Kwaliteitselementen (EKR): fytoplankton, macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peilbeheer, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 193
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 1.000 x op  $R^2$  en 50.000 x op richtingscoëfficiënt

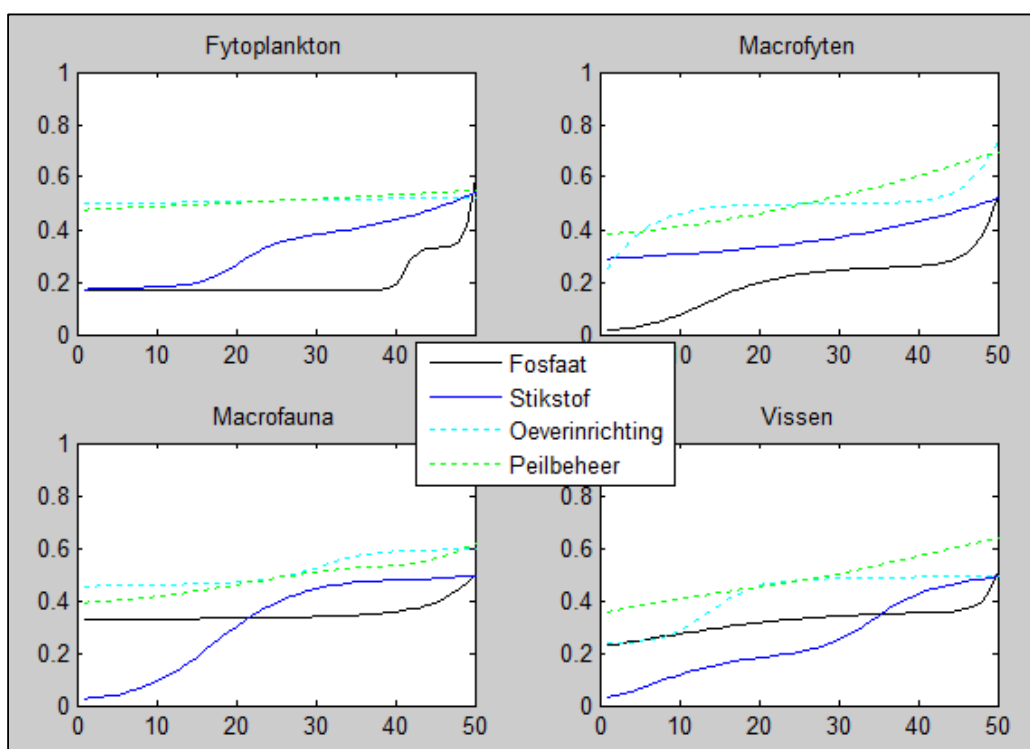
#### Resultaat neuraal netwerk op trainingset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Fytoplankton: 92
- Macrofyten: 1556
- Macrofauna: 389
- Vissen: 953

*Statistische gegevens*

Parameter	Betekenis	Fytoplankton	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	0.64	0.42	0.57	0.35
R	Correlatie coëfficiënt	0.88	0.92	0.86	0.91
R2	Kwadraat corr coëff	0.78	0.84	0.73	0.83
RC	Richtingscoëfficiënt regressielijn	0.77	0.85	0.75	0.87
Std RC	Standaard afwijking van RC	0.020	0.014	0.014	0.015
Int	Intercept van regressielijn	0.095	0.077	0.122	0.046
ME	Gemiddelde fout	-0.002	0.011	0.001	-0.008
MSE	Gemiddelde van gekwadrateerde fouten	0.014	0.007	0.008	0.007
RMSE	Standaard deviatie (wortel van MSE)	0.117	0.086	0.089	0.086
% binnen 0,1 EKR		61	79	72	75

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.



## **Bijlage 5**

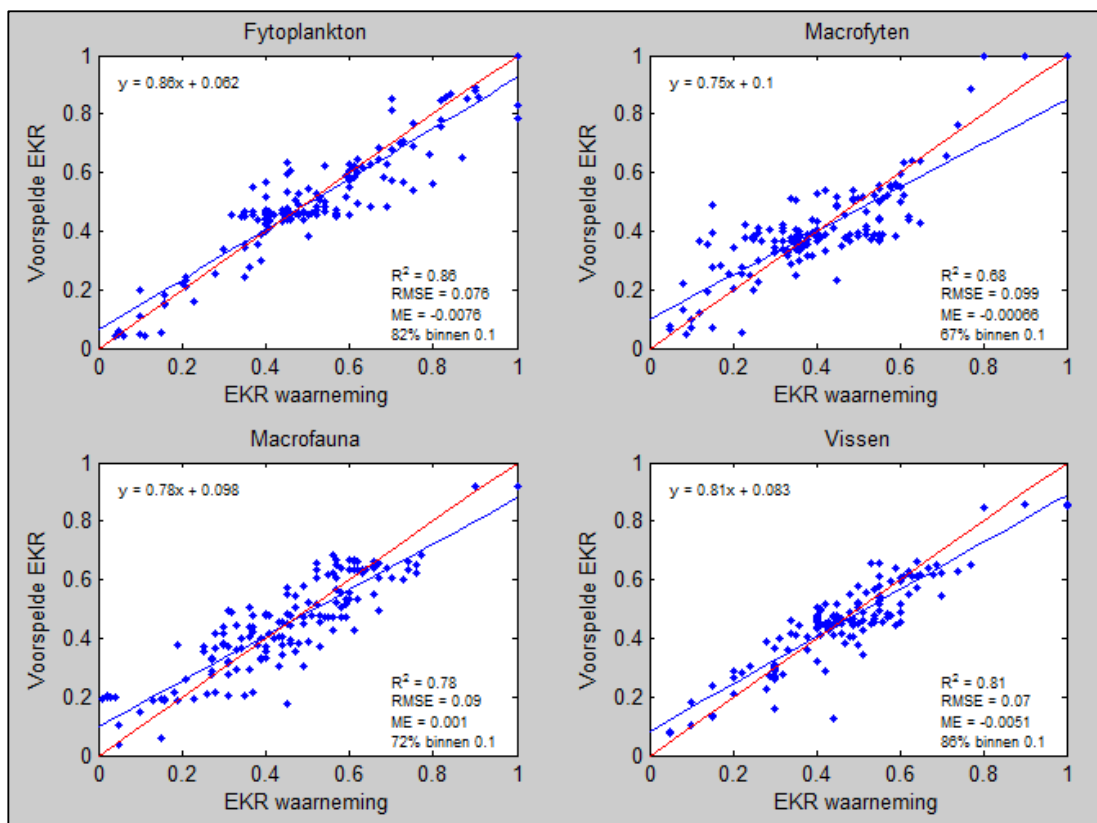
### **Beschrijving NN kanalen**



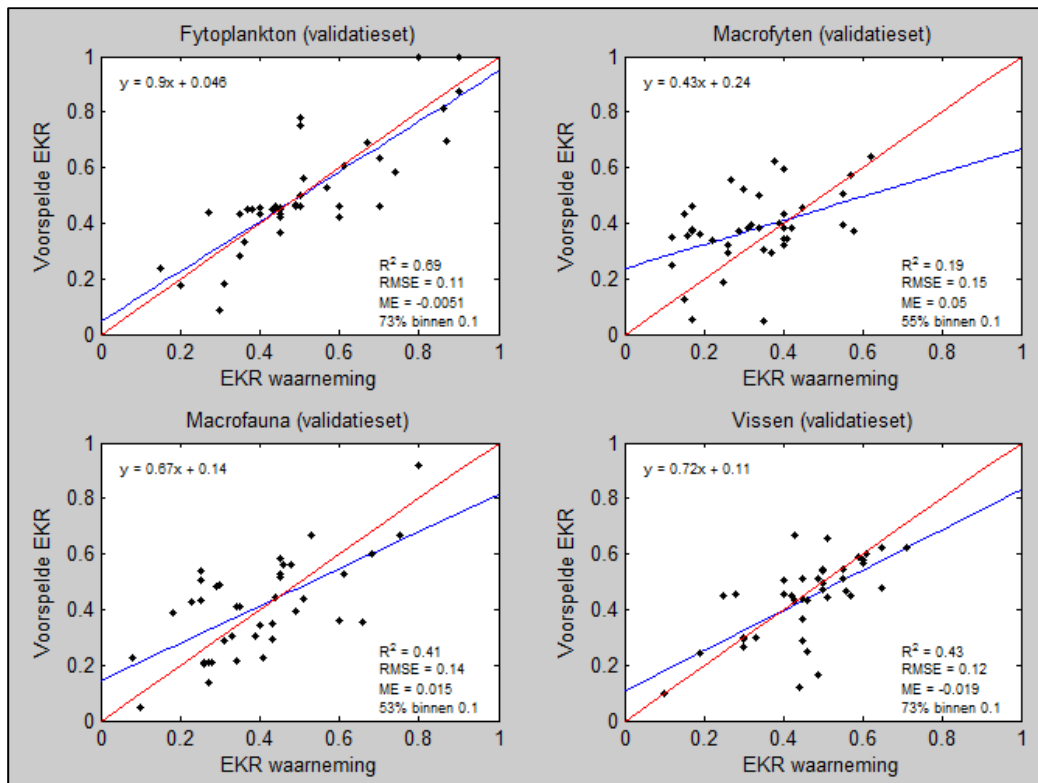
### Neuraal netwerken (EEE3) voor watertypen M3 M6ab M7ab en M10 Kanalen

- Kwaliteitselementen (EKR): fytoplankton, macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peilbeheer, onderhoud, scheepvaart, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 198
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 5.000 x op  $R^2$  en 50.000 x op richtingscoëfficiënt

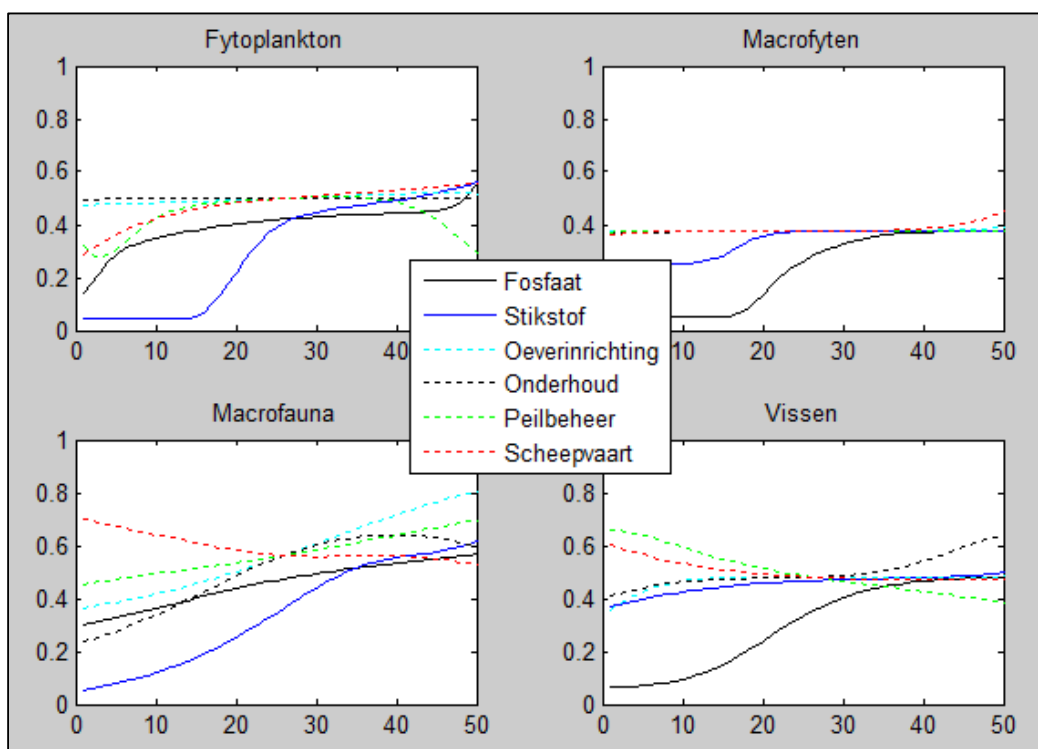
#### Resultaat neuraal netwerk op trainingset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren





*Aantal goede netwerken*

- Fytoplankton: 478
- Macrofyten: 1546
- Macrofauna: 821
- Vissen: 1189

*Statistische gegevens*

Parameter	Betekenis	Fytoplankton	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	0.64	-0.21	0.27	0.23
R	Correlatie coëfficiënt	0.93	0.82	0.88	0.90
R2	Kwadraat corr coëff	0.86	0.68	0.78	0.81
RC	Richtingscoëfficiënt regressielijn	0.86	0.75	0.78	0.81
Std RC	Standaard afwijking van RC	0.012	0.018	0.015	0.012
Int	Intercept van regressielijn	0.062	0.100	0.098	0.083
ME	Gemiddelde fout	-0.008	-0.001	0.001	-0.005
MSE	Gemiddelde van gekwadrateerde fouten	0.006	0.008	0.008	0.005
RMSE	Standaard deviatie (wortel van MSE)	0.076	0.099	0.090	0.070
% binnen 0,1 EKR		82	67	72	86

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.



## **Bijlage 6**

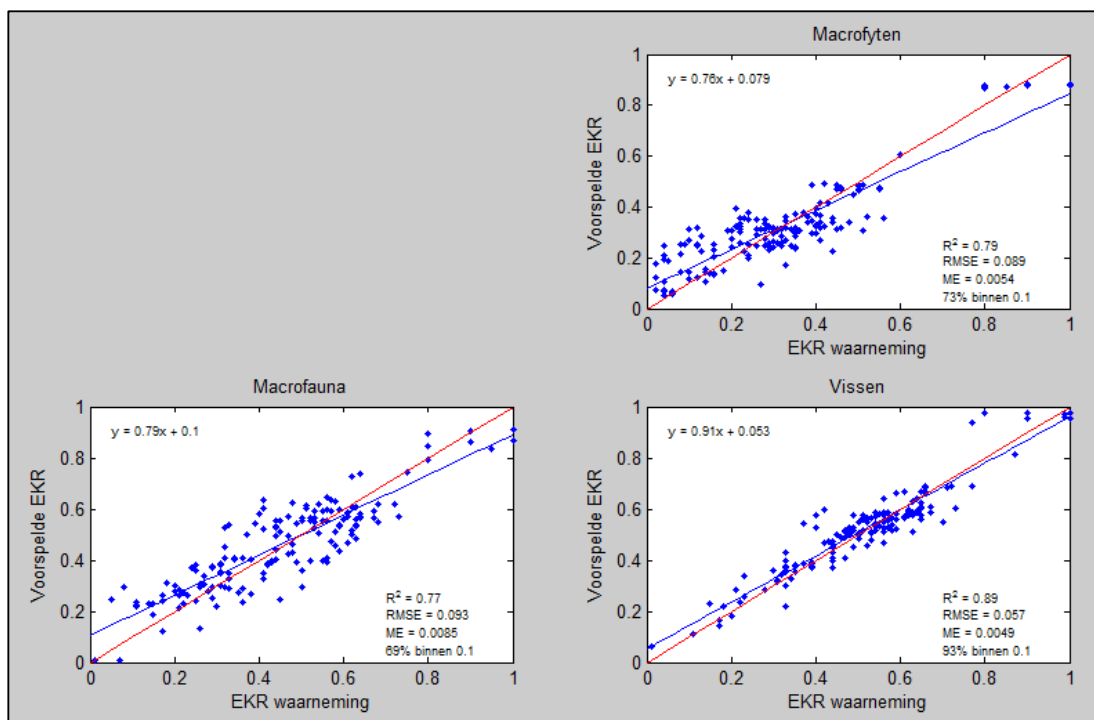
### **Beschrijving NN sloten**



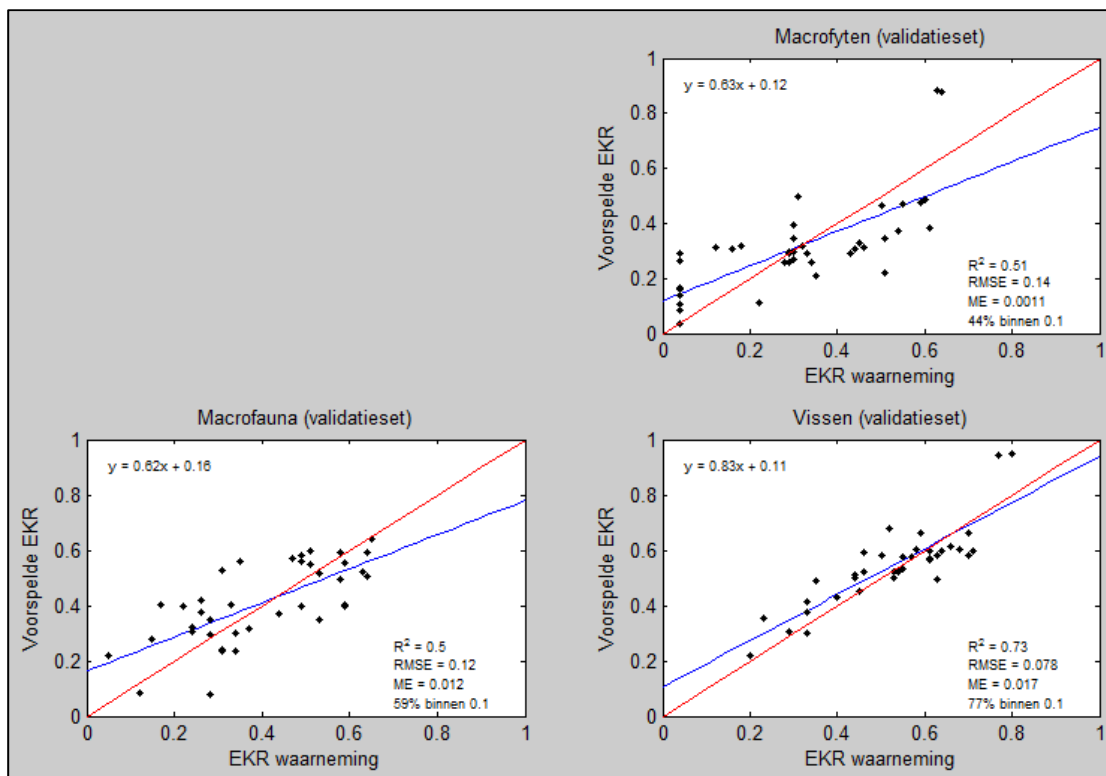
### Neuraal netwerken (EEE3) voor watertypen M01 M02 M08 Sloten

- Kwaliteitselementen (EKR): macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peildynamiek, onderhoud, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 193
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 1.000 x op  $R^2$  en 10.000 x op richtingscoëfficiënt

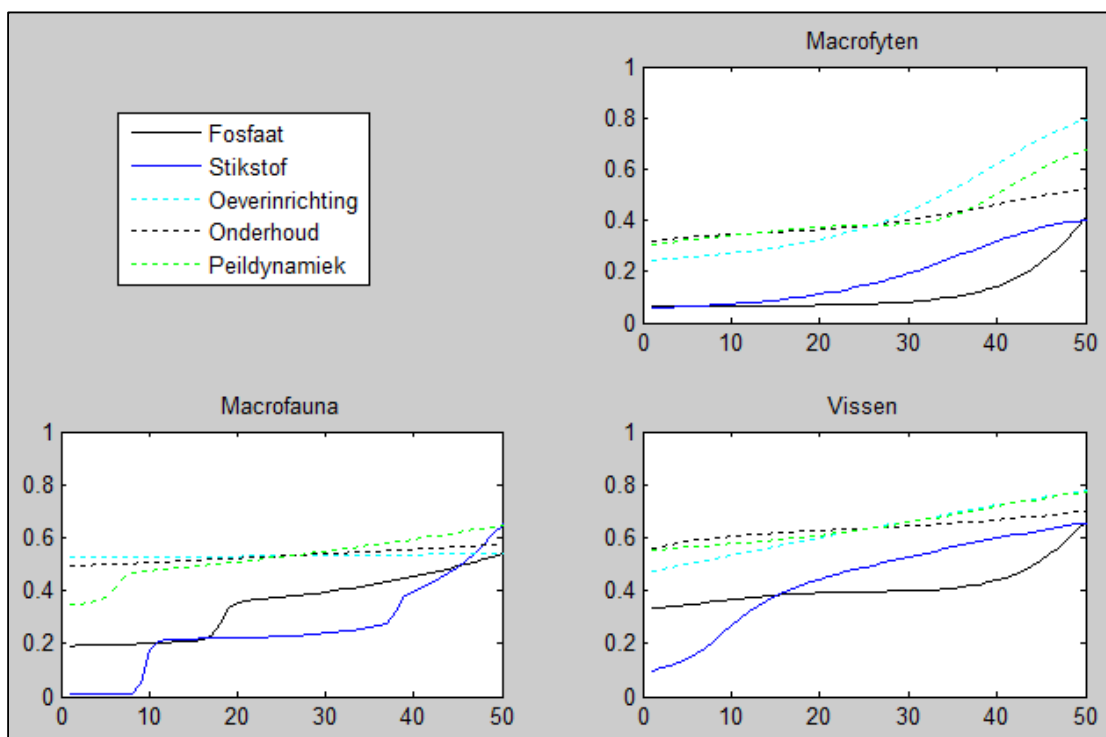
#### Resultaat neuraal netwerk op trainingset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Macrofyten: 894
- Macrofauna: 134
- Vissen: 658

*Statistische gegevens*

Parameter	Betekenis	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	0.49	0.49	0.71
R	Correlatie coëfficiënt	0.89	0.88	0.94
R2	Kwadraat corr coëff	0.79	0.77	0.89
RC	Richtingscoëfficiënt Regressielijn	0.76	0.79	0.91
Std RC	Standaard afwijking van RC	0.020	0.015	0.008
Int	Intercept van regressielijn	0.079	0.105	0.053
ME	Gemiddelde fout	0.005	0.008	0.005
MSE	Gemiddelde van gekwadrateerde fouten	0.009	0.009	0.003
RMSE	Standaard deviatie (wortel van MSE)	0.089	0.093	0.057
% binnen 0,1 EKR		73	69	93

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.





## **Bijlage 7**

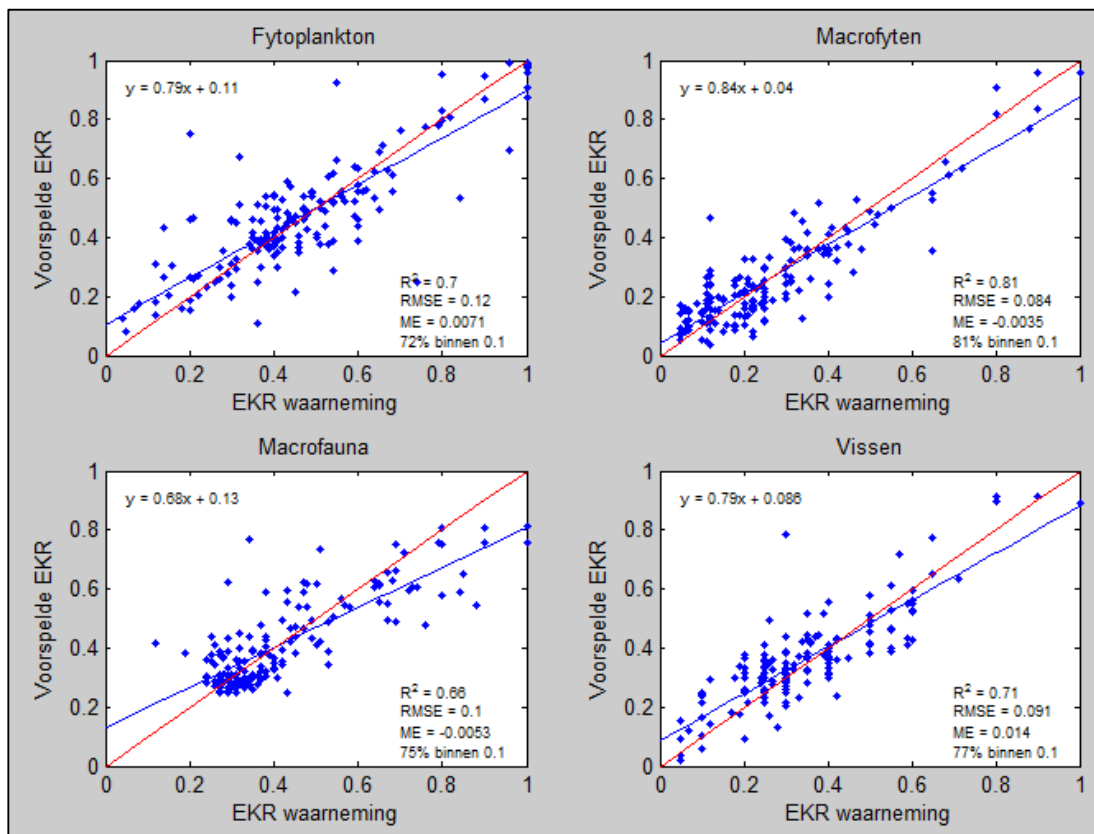
### **Beschrijving NN zwak brakke wateren**



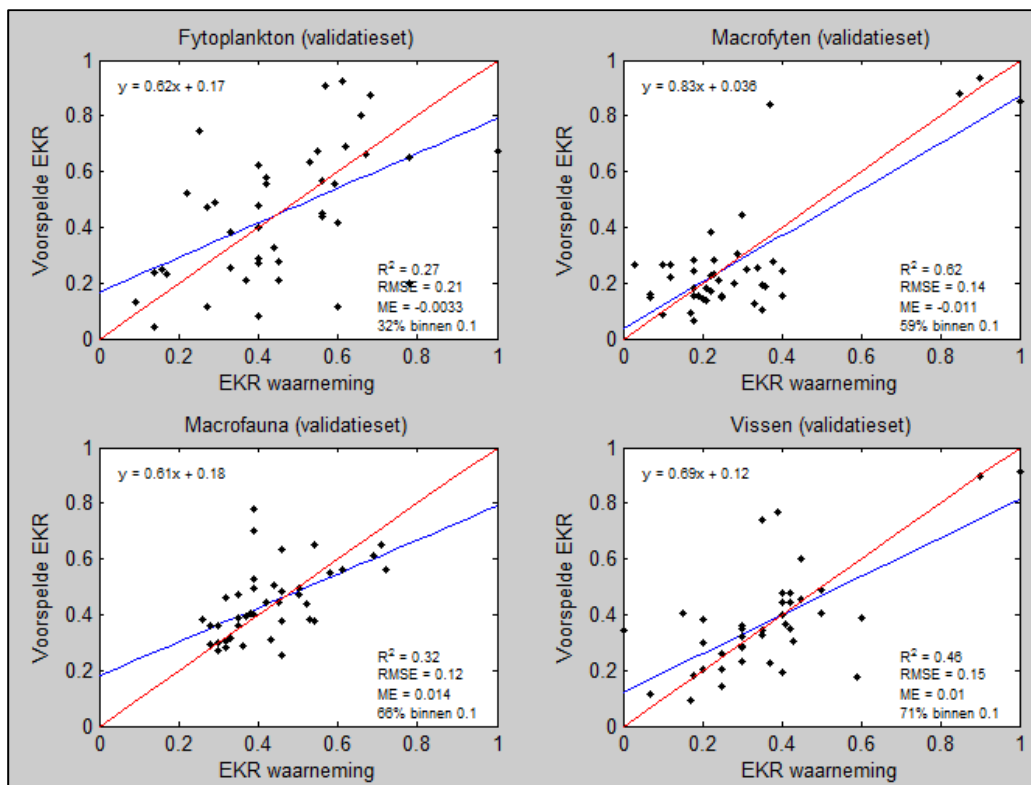
## Neuraal netwerken voor watertype M30 Zwak brakke wateren

- Kwaliteitselementen (EKR): fytoplankton, macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peildynamiek, onderhoud, connectiviteit, chloride, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 203
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 5000 voor maximaliseren van correlatie ( $R^2$ ) tussen waargenomen en voorspelde EKR plus 50.000 voor maximaliseren van de richtingscoëfficiënt van de regressielijn van beide EKR-sets

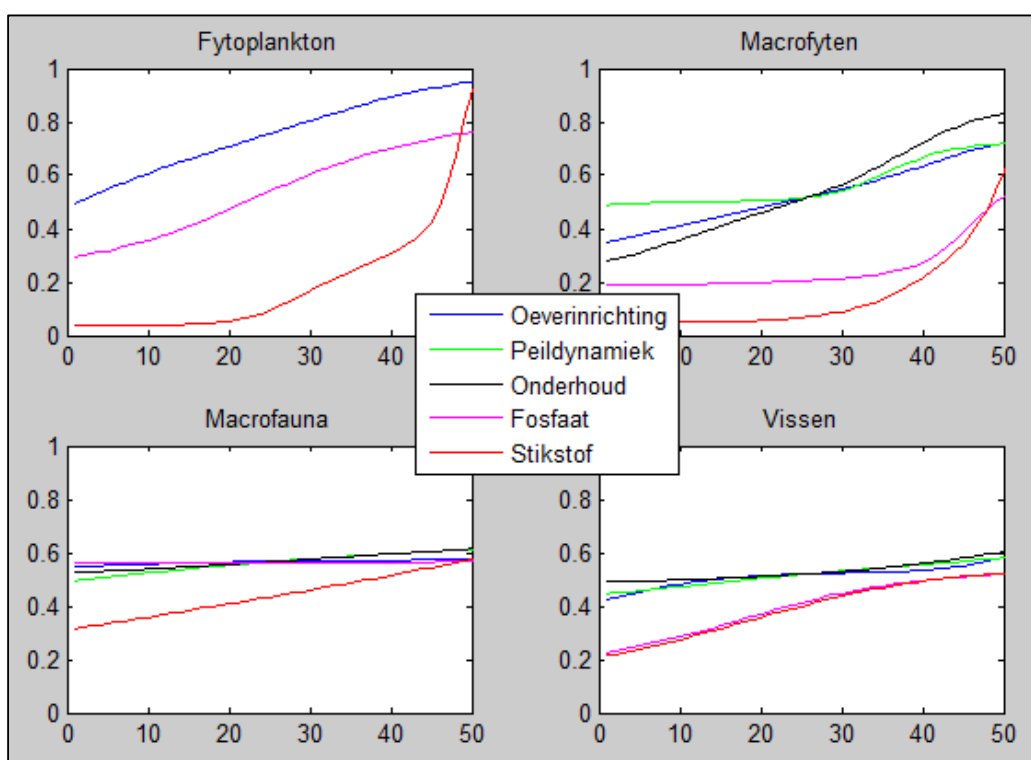
### Resultaat neuraal netwerk op gehele dataset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Fytoplankton: 2569
- Macrofyten: 1993
- Macrofauna: 960
- Vissen: 1072

*Statistische gegevens*

Parameter	Betekenis	Fytoplankton	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	-0.18	0.55	0.06	0.35
R	Correlatie coëfficiënt	0.84	0.90	0.81	0.84
R2	Kwadraat corr coeff	0.70	0.81	0.66	0.71
RC	Richtingscoëfficiënt regressielijn	0.79	0.84	0.68	0.79
Std RC	Standaard afwijking van RC	0.018	0.021	0.019	0.018
Int	Intercept van regressielijn	0.106	0.040	0.131	0.086
ME	Gemiddelde fout	0.007	-0.004	-0.005	0.014
MSE	Gemiddelde van gekwadeerde fouten	0.014	0.007	0.011	0.008
RMSE	Standaard deviatie (wortel van MSE)	0.119	0.084	0.104	0.091
% binnen 0,1 EKR		72	81	75	77

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.



## **Bijlage 8**

### **Beschrijving NN brak tot zoute wateren**

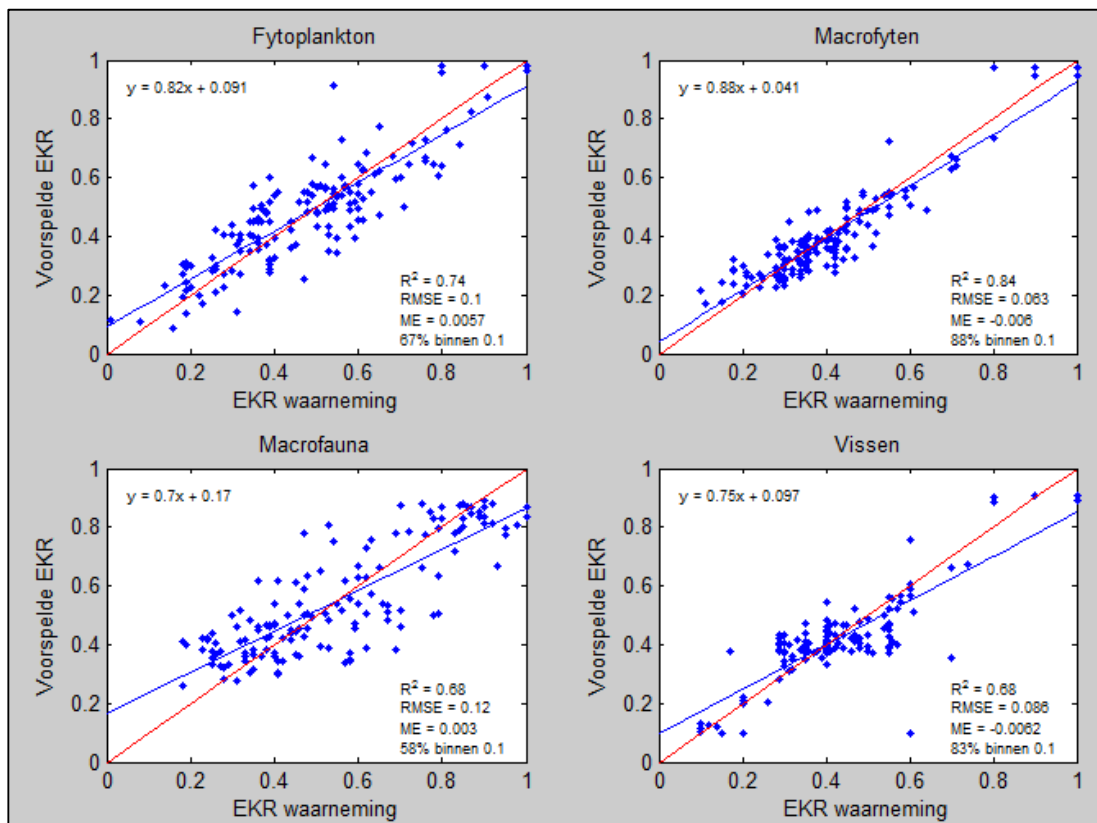




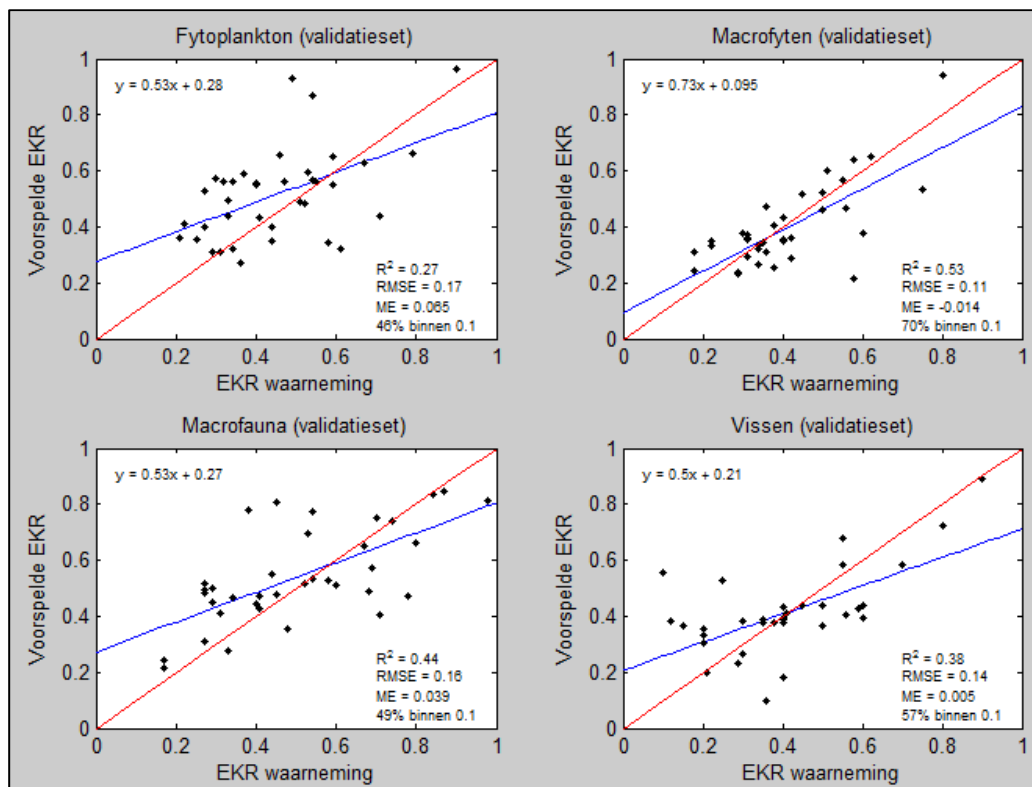
## Neuraal netwerken voor watertype M31 Brak zoute wateren

- Kwaliteitselementen (EKR): fytoplankton, macrofyten, macrofauna en vissen
- Stuurfactoren: oeverinrichting, peildynamiek, onderhoud, connectiviteit, chloride, totaal fosfaat, totaal stikstof
- Aantal waarnemingen in dataset: 186
- Toegepast software pakket: Neural Network Toolbox 7.0 in Matlab 7.11.0.584 (R2010b)
- Aantal data gebruikt voor training: 80%
- Aantal data gebruikt voor validatie: 20%
- Aantal trainingen: 5000 voor maximaliseren van correlatie ( $R^2$ ) tussen waargenomen en voorspelde EKR plus 50.000 voor maximaliseren van de richtingscoëfficiënt van de regressielijn van beide EKR-sets

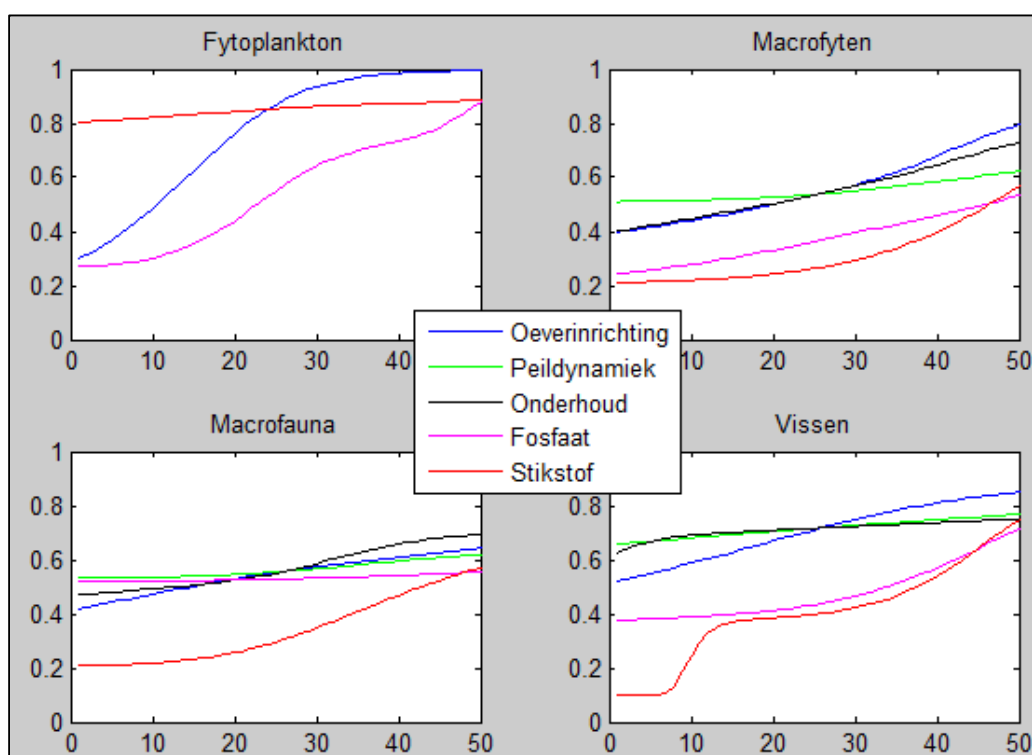
### Resultaat neuraal netwerk op gehele dataset



Resultaat neuraal netwerk op validatieset



Gevoeligheid voor verandering stuurfactoren



*Aantal goede netwerken*

- Fytoplankton: 8723
- Macrofyten: 1839
- Macrofauna: 353
- Vissen: 931

*Statistische gegevens*

Parameter	Betekenis	Fytoplankton	Macrofyten	Macrofauna	Vissen
COD validatie	Coëfficiënt of determination <sup>1</sup>	-0.13	0.45	0.40	0.36
R	Correlatie coëfficiënt	0.86	0.92	0.82	0.82
R2	Kwadraat corr coeff	0.74	0.84	0.68	0.68
RC	Richtingscoëfficiënt regressielijn	0.82	0.88	0.70	0.75
Std RC	Standaard afwijking van RC	0.016	0.012	0.018	0.016
Int	Intercept van regressielijn	0.091	0.041	0.167	0.097
ME	Gemiddelde fout	0.006	-0.006	0.003	-0.006
MSE	Gemiddelde van gekwadrateerde fouten	0.010	0.004	0.015	0.007
RMSE	Standaard deviatie (wortel van MSE)	0.100	0.063	0.123	0.086
% binnen 0,1 EKR		67	88	58	83

<sup>1</sup> COD is een maat voor de voorspellingswaarde van een ANN. Waarde dient groter te zijn dan 0 en zo dicht mogelijk bij 1 te liggen.