Fish in Lakes Classification Tool (FIL2) for Eco Region 17

Summary of development and outputs

Inland Fisheries Ireland – Research and Development Division

Fiona Kelly, Andrew Harrison and Lynda Connor

Agri-Food and Biosciences Institute (AFBI) - Fisheries and Aquatic Systems Branch

Michelle Allen, Trudyann Kelly and Robert Rosell




Overview

- FIL1 (NS Share 2005-2006)
- FIL2 – 2010
- Typology suitable for fish
- Discriminant model-classification rules
- EQR model and classification of boundaries
- Ecological status
- Testing FIL2 on data from Scotland, England & Wales
- Preliminary comparison of FIL2 and EQR4
- Questions

NS Share Project – FIL1 (Fish in lakes tool) (2005 to 2006)

- 3 Fish types (salmonid, perch and roach)
- Multimetric approach
- The individual metrics were not scored and combined in the manner employed by other researchers (e.g. Karr).
- Combines metrics using a classification rule (Discriminant analysis)
- Not fully WFD compliant as no EQRs

Only 80 lakes

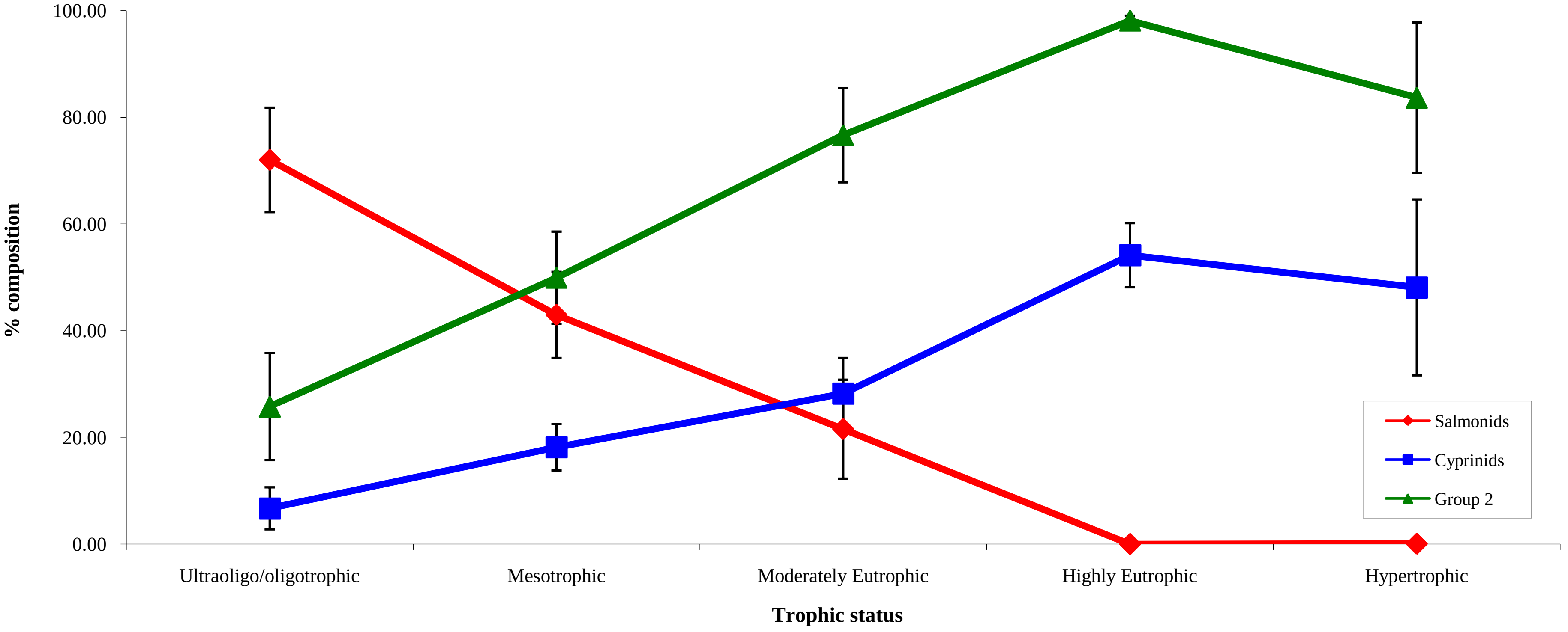
FIL2 - February to December 2010

Objectives

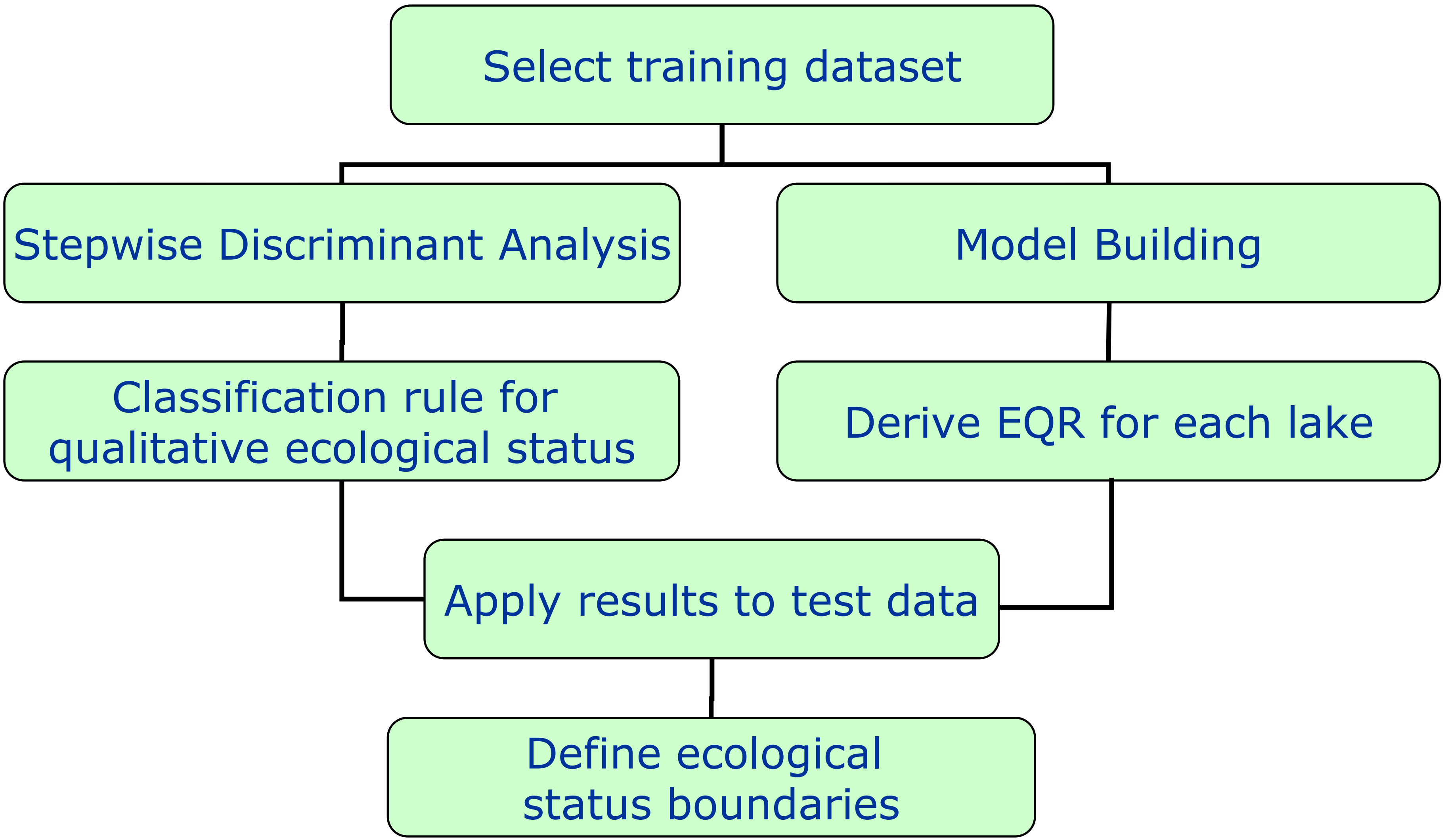
- Update FIL1 using new additional data (2007 to 2009)
- Test typology
- Derive an Ecological Quality Ratio (EQR) for each lake [0,1]
- Assign quantitative EQR boundaries to the 5 qualitative ecological status classes – high, good, moderate, poor and bad
- Assess if the end product is transferable between member states - intercalibration to be finalised by June 2011

First questions to be addressed

1. Can fish be used to classify ecological status?
2. Is there a change in the structure of the fish community in relation to a decrease in trophic status/water quality?



Overview of the Solution



Fish in lakes Data (2005 to 2009)

- 137 lakes surveyed = training dataset
- 14 lakes surveyed on two occasions = test dataset (1st survey)
- Abiotic metrics (x 8)
- Pressure metrics (x 9)
- Fish metrics (x 145)
  - species composition, reproduction, abundance, biomass, condition, age, length

Typology for lakes – Ecoregion 17

Typology	Impact Class					Total
	High	Good	Moderate	Poor	Bad	
Low ALK <4m <50Ha	5	11	2	0	1	19
Low ALK <4m ≥50Ha	0	6	0	0	0	6
Low ALK ≥4m <50Ha	1	4	0	0	0	5
Low ALK ≥4m ≥50Ha	5	14	0	0	0	19
Med ALK <4m <50Ha	0	5	3	3	3	14
Med ALK <4m ≥50Ha	0	8	3	2	0	13
Med ALK ≥4m <50Ha	0	5	3	0	0	8
Med ALK ≥4m ≥50Ha	1	11	1	0	0	13
High ALK <4m <50Ha	0	4	2	2	1	9
High ALK <4m ≥50Ha	2	9	1	0	2	14
High ALK ≥4m <50Ha	2	4	3	0	0	9
High ALK ≥4m ≥50Ha	2	6	0	0	0	8
Total	18	87	18	7	7	137

Question: Is it appropriate for fish?

Typology suitable for fish

Reference condition defined as:

– % Natural land	Ref > 80%
– % Agricultural	Ref < 20%
– % Urban land	Ref <20%
– % Forestry	Ref <20%
– TP mean	Ref <12ug/l P (rejection threshold <20ug/l P)
– Chlor a mean	Ref < 8mg/l (rejection threshold <15mg/l)
– pH	Ref = 6-9
– Connectivity	Ref = no barriers or no natural impassable barriers

↓

43 lakes

Developing a typology appropriate for fish

Abiotic variables

Lake area (Ha)

Max D (m)

Altitude (m)

U/S trib area

Alkalinity (mg/l CaCO3)

Easting (Irish national grid in m)

Northing (Irish national grid in m)

Distance from tidal limit (m)

Correlation with

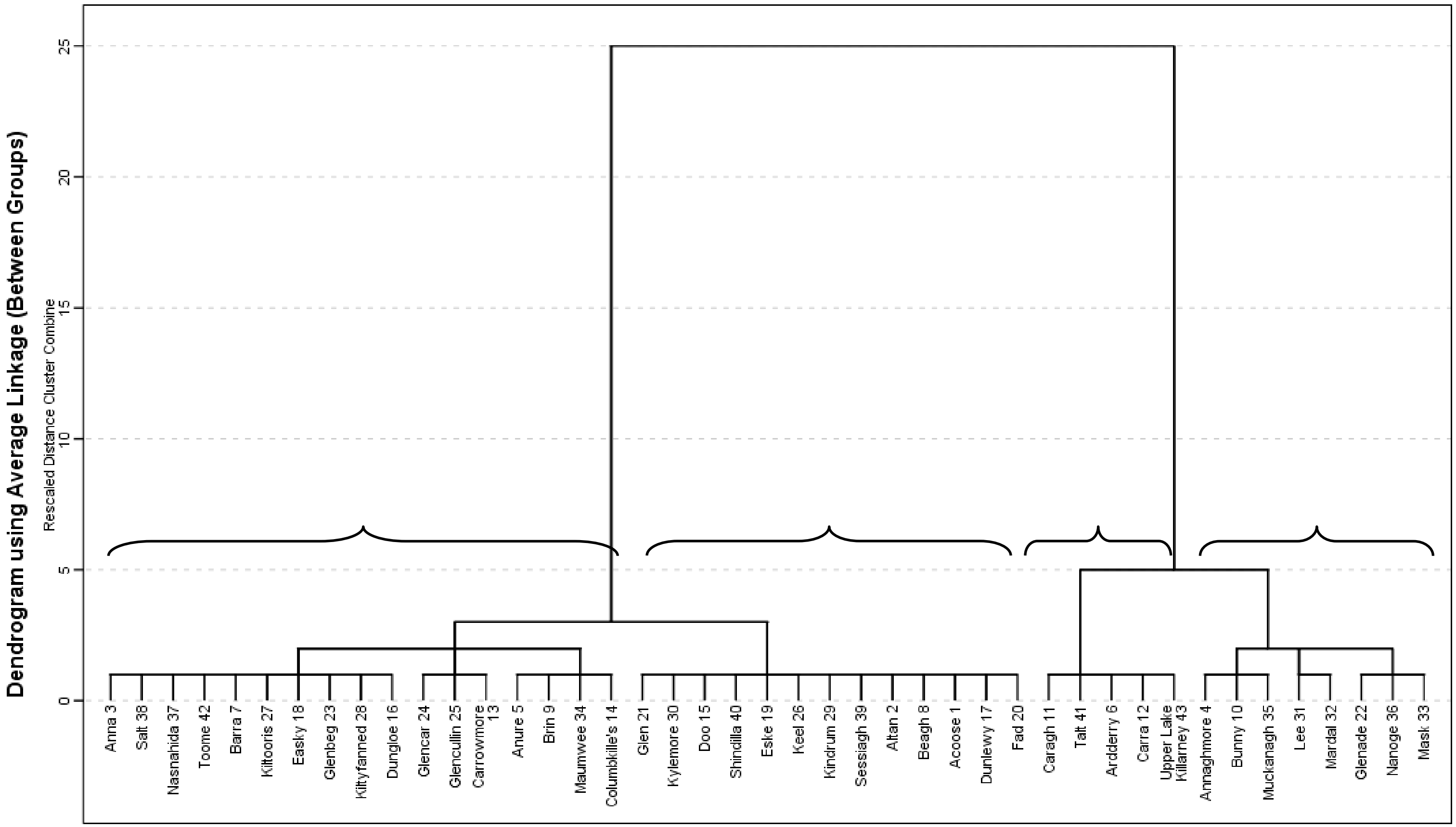
Fish metrics
(raw and square root transformed)

=

alkalinity and depth most significant variables influencing the fish community in reference lakes

PCA, Cluster and Stepwise Discriminant analysis

Defining typology: Cluster analysis - reference lakes

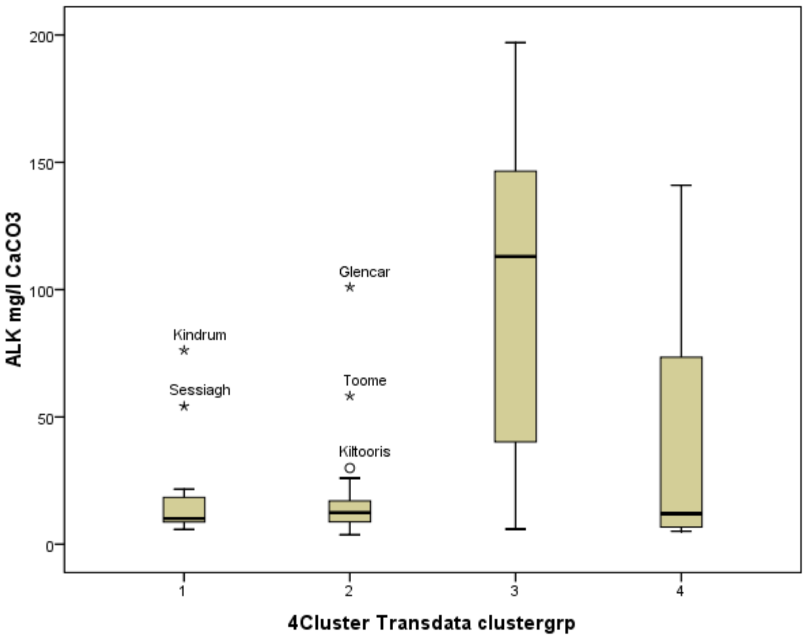


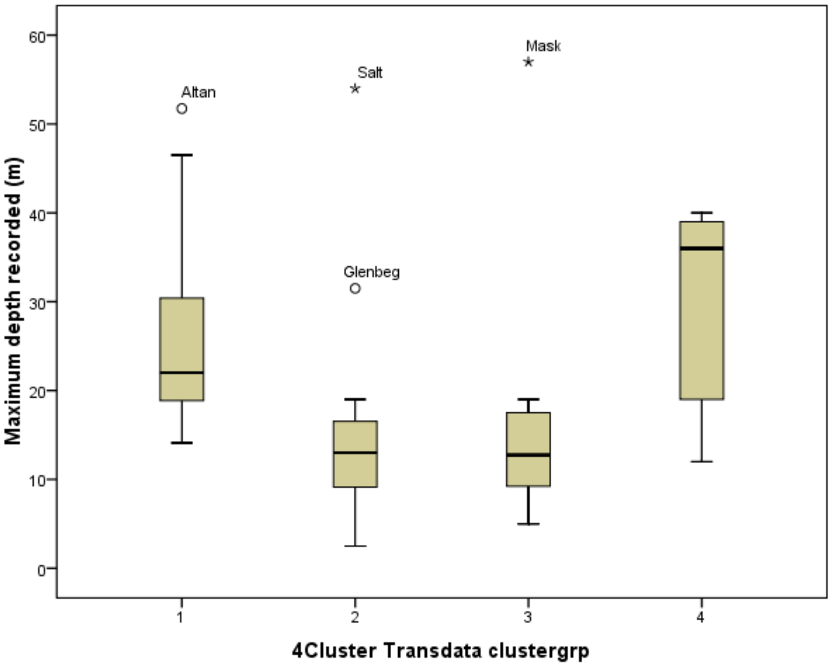
Cluster analysis of fish metrics - 4 groups of lakes

(Ward linkage method on square root transformed data)

Results: Typology

no clear cut-points, cross-over in range of values





Boxplots of alkalinity and Max Depth grouped by cluster membership when analysing the square root transformed fish metrics using the Ward linkage method

Typology for fish in lakes in Ireland

Alkalinity	mg/l $CaCO_3$	Max Depth	(m)
Low	<67	Shallow	<17
Low	<67	Deep	>17
High	>67	Shallow	<17
High	>67	Deep	>17

1. **Optimum cut-point for ALK was 67 with 36 of the 43 lakes being correctly classified into Low and High ALK**

2. **The optimum cut-point for depth was 17 with 35 of the 43 lakes being correctly classified into Shallow and Deep depth**

Next steps

1. Define a pressure gradient
2. Stepwise discriminant analysis and classification rules for assigning qualitative ecological status
3. EQR model
4. Define Boundaries

1. Define a Pressure gradient

Combine the TP_MEAN and CHLOR_MAX using a 'lowest common denominator' approach

E.g. TP_MEAN = high, CHLOR_MAX = good…..overall = good

TP_MEAN = good, CHLOR_MAX = poor…..overall = poor

	Total P (ug/l P)	Max Chl a (mg/m3)
Ultraoligo/Oligotrophic (HIGH)	TP_MEAN <12	<8.0
Mesotrophic (GOOD)	12<= TP_MEAN <35	>8- <=25
Moderately Eutrophic (MODERATE)	35<= TP_MEAN <60	>25 - <=50
Highly Eutrophic (POOR)	60<= TP_MEAN <100	>50 - <=75
Hypertrophic (BAD)	TP_MEAN >100	>75

Number of lakes by typology and impact class for the training dataset

Combined poor and bad

Typology	Impact Class				Total
	High	Good	Mod	Poor/bad	
Low Alkalinity/Shallow Depth	11	25	4	8	48
Low Alkalinity/Deep Depth	5	18	4	2	29
High Alkalinity/Shallow Depth	3	17	11	12	43
High Alkalinity/Deep Depth	3	12	1	1	17
Total	22	72	20	23	137

2. Classification rule for qualitative ecological status

- Discriminant analysis classification rules were developed for each of the 4 typologies using a stepwise procedure.
- The discriminant analysis process will objectively identify which of the metrics will maximise the between to within impact class variability.
- Sparse metrics, metrics with missing values and species-specific metrics for
  EEL, PIKE, PRED, ROACHBREAM, ROACHRUDD, RUDDBREAM and TENCH were excluded from the analysis
- Sparse was defined to be those metrics with less than 10% non-zero values.

13 fish metrics used for the DA classification

CORE METRICS

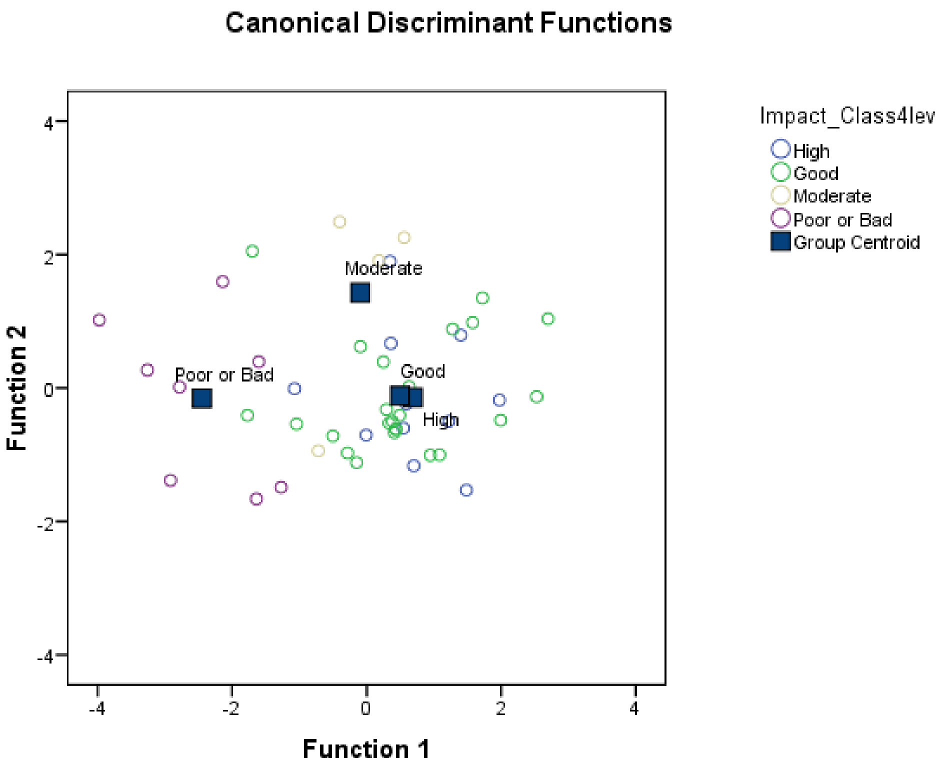
- TOT_BPUE: Sum of mean biomass per unit effort (excl. eels and adult salmon)
- NAT_BPUE: Sum of mean biomass per unit effort of native (group 1) fish species
- PERCH_BIO: Mean perch biomass per unit effort

OPTIMAL METRICS (stepwise procedure)

- RHEO_BIO: % individuals (based on BPUE excl. eels & adult salmon) that are rheophilic
- SPE_EVEN: Species evenness/dominance (1/D=1/(Nmax/Ntot) (Nmax= no. inds represented by the most abundant species, Ntot=total number of individuals in the sample (eels captured in fyke nets excluded) (Based on total number of fish captured)
- ROACH_BPUE: Mean biomass per unit effort ((g) of fish per linear metre of net - gill nets and fyke nets)
- BREAM_%_IND: % composition of bream based on CPUE (BREAM_CPUE/TOTAL_CPUE*100)
- PHYT_%_BIO: % individuals (based on BPUE excl. eels and adult salmon) that are phytophilic
- 2_%_BIO: % biomass of Group 2 species (based on BPUE excl. eels and adult salmon), inc hybrids
- CYP_BIO: % biomass (based on BPUE excl. eels and adult salmon) of cyprinid species, inc hybrids
- RUDD_%_IND: % composition of rudd based on CPUE (RUDD_CPUE/TOTAL_CPUE*100)
- MAX_L_DOM_BIO: Maximun length of dominant species (based on BPUE excl. eels and adult salmon)
- LITH_IND: % individuals (based on CPUE excl. eels and adult salmon) that are lithophilic

(per linear metre of net used – gill nets and fykes)

Low ALK/Shallow lakes: 6 fish metrics



Plot of the first two discriminant functions based on the 6 metrics

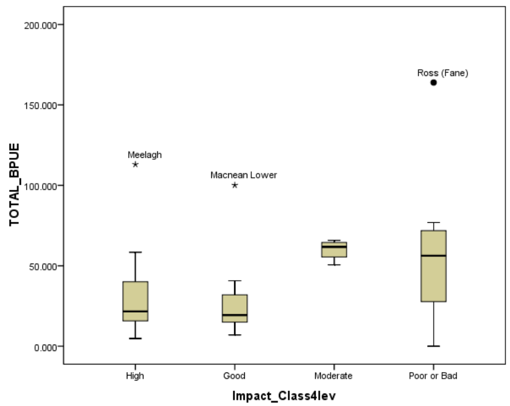
92.1% of the variation within the data was explained by the first two discriminant functions .

Classification of the discriminant analysis rule

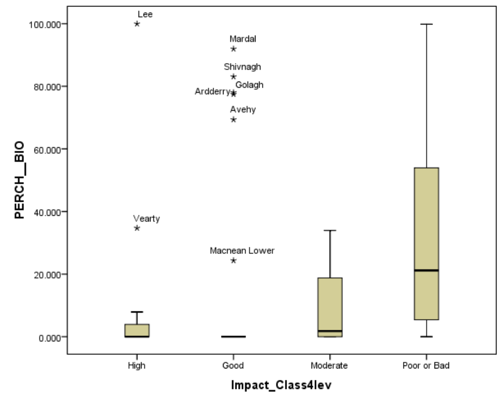
	Predicted Impact Class				
Actual Impact Class	**High**	**Good**	**Mod**	**Poor/Bad**	**Total**
High	7	3	1	0	11
Good	4	**17**	2	2	25
Moderate	0	1	**3**	0	4
Poor or Bad	0	0	0	**8**	8
Total	11	21	6	10	48

72.9% of the lakes were correctly classified and when jack-knifing 52.1% were correctly classified.

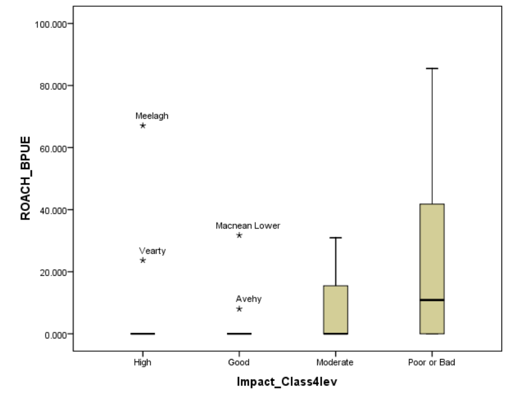
Low alkalinity shallow lakes: 6 fish metrics significant



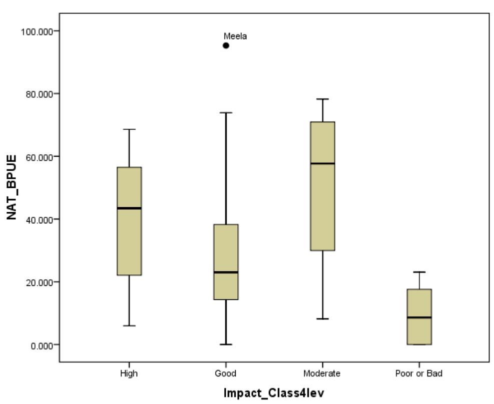
TOT_BPUE



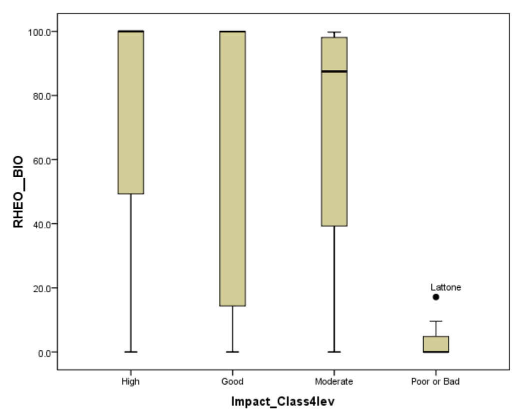
PERCH_BIO



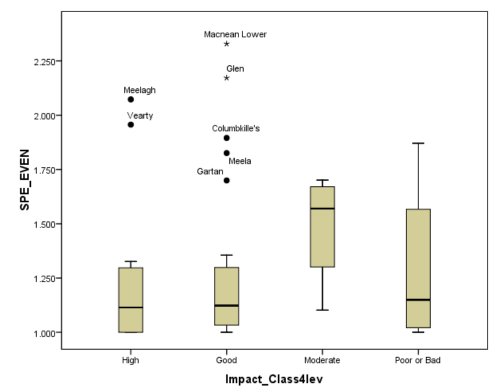
ROACH_BPUE



NAT_BPUE



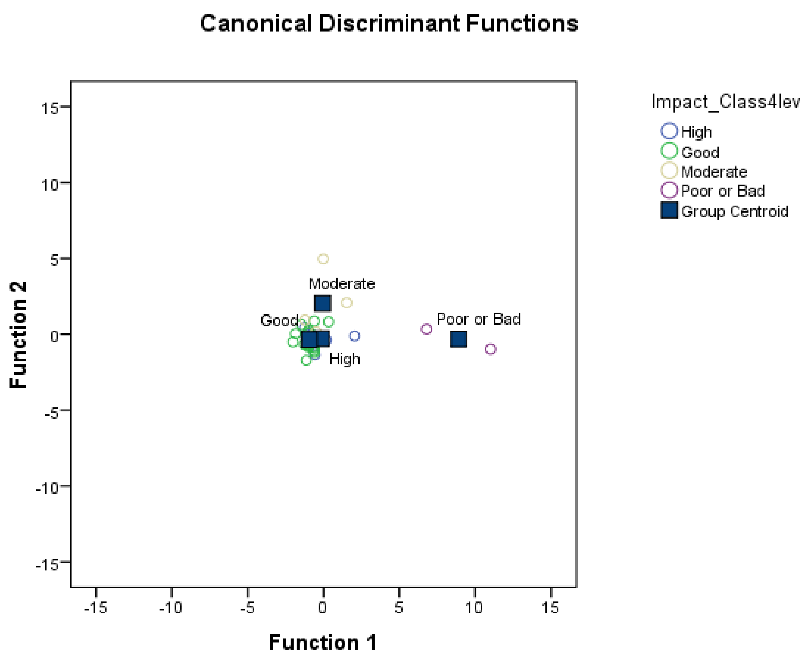
RHEO_BIO



SPE_EVEN

Typology-specific boxplots for the metrics identified by the discriminant analysis

Low ALK/Deep Lakes: 5 fish metrics significant



Plot of the first two discriminant functions based on the 5 metrics

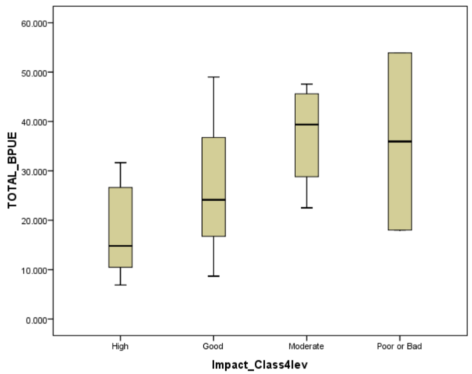
98.7% of the variability within the data was explained by the first two discriminant functions

Classification of the discriminant analysis rule

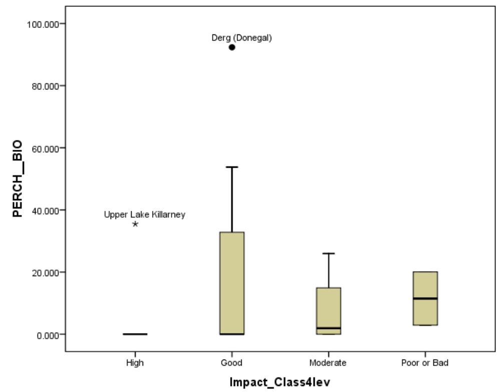
	Predicted Impact Class				
Actual Impact Class	**High**	**Good**	**Mod**	**Poor/Bad**	**Total**
High	**3**	2	0	0	5
Good	5	**12**	1	0	18
Moderate	0	2	**2**	0	4
Poor/ Bad	0	0	0	**2**	2
Total	8	16	3	2	29

65.7% of the lakes were correctly classified and when jack-knifing 51.7% were correctly classified.

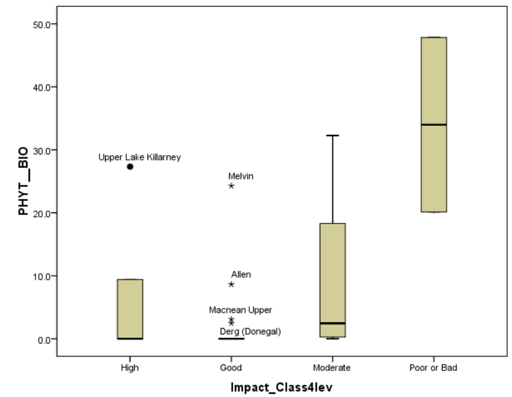
Low alkalinity deep lakes: 5 fish metrics significant



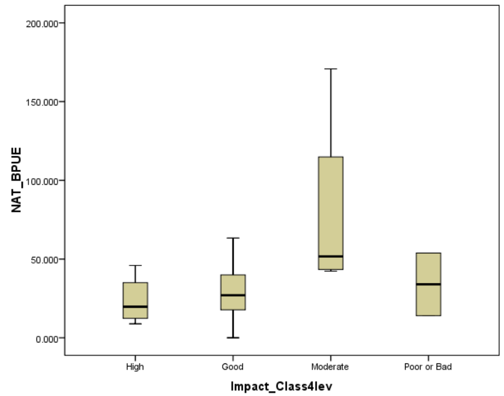
TOT_BPUE



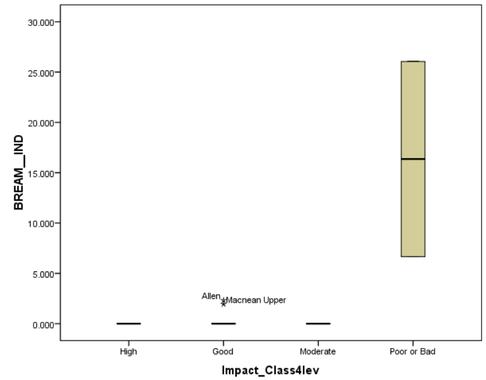
PERCH_BIO



PHYT_BIO



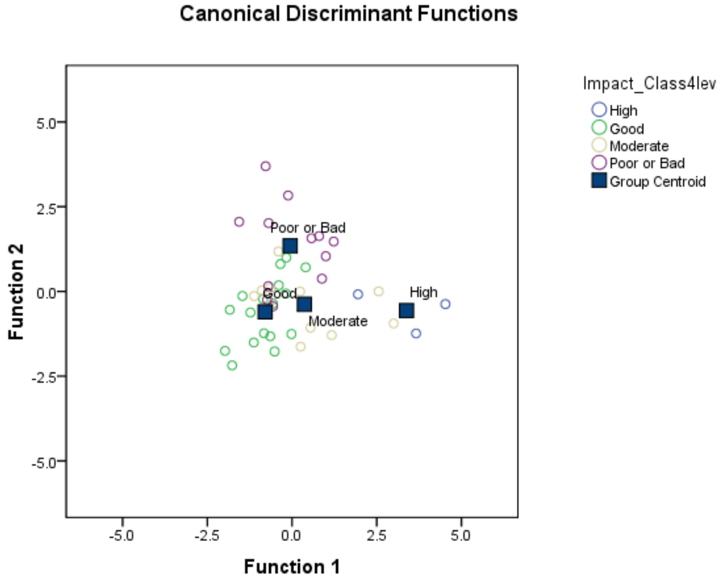
NAT_BPUE



BREAM_IND

Typology-specific boxplots for the metrics identified by the discriminant analysis

High alkalinity shallow lakes: 6 fish metrics significant



Plot of the first two discriminant functions based on the 6 metrics

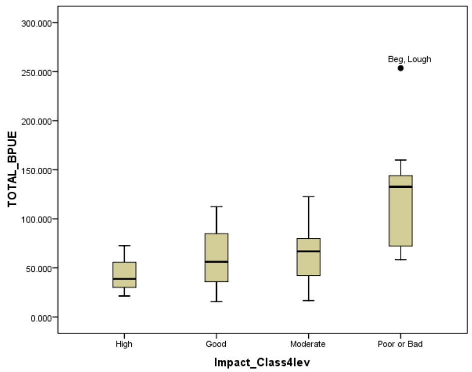
Classification of the discriminant analysis rule

	Predicted Impact Class				
Actual Impact Class	**High**	**Good**	**Mod**	**Poor/Bad**	**Total**
High	**3**	0	0	0	3
Good	0	**13**	1	3	17
Moderate	2	4	**4**	1	11
Poor or Bad	0	3	1	**8**	12
Total	5	20	6	12	43

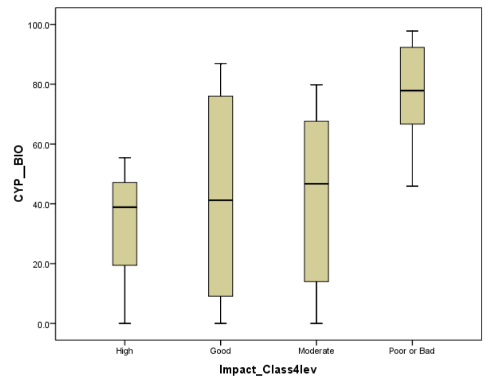
97.6% of the variation in the data was explained by the first two discriminant functions of the discriminant rule

65.1% of the lakes were correctly classified and when jack-knifing 53.3% of the lakes were correctly classified.

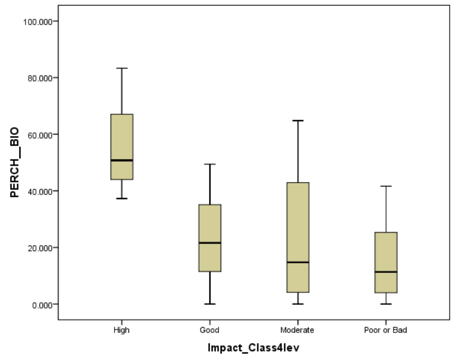
High alkalinity shallow lakes: 6 fish metrics significant



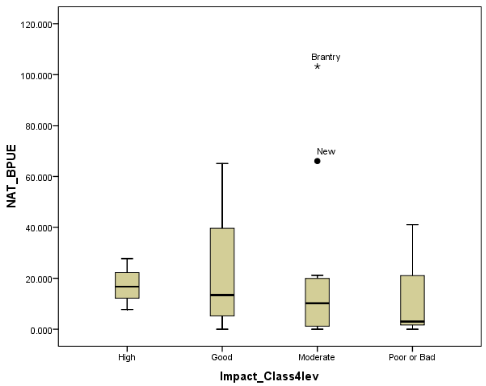
TOT_BPUE



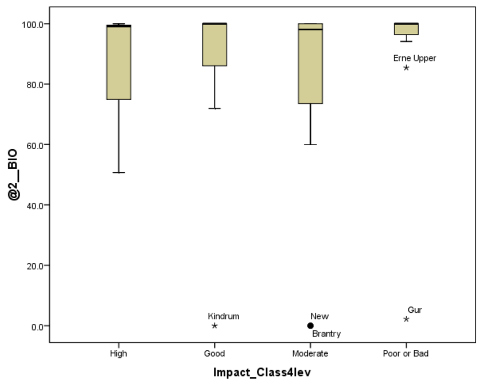
CYP_BIO



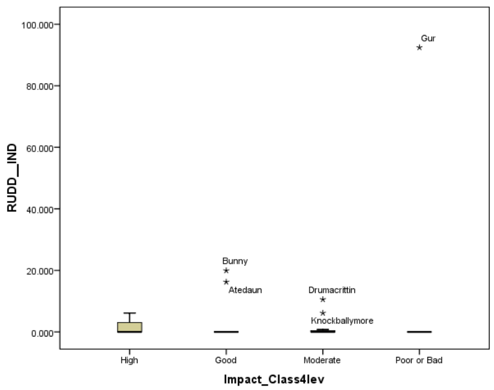
PERCH_BIO



NAT_BPUE



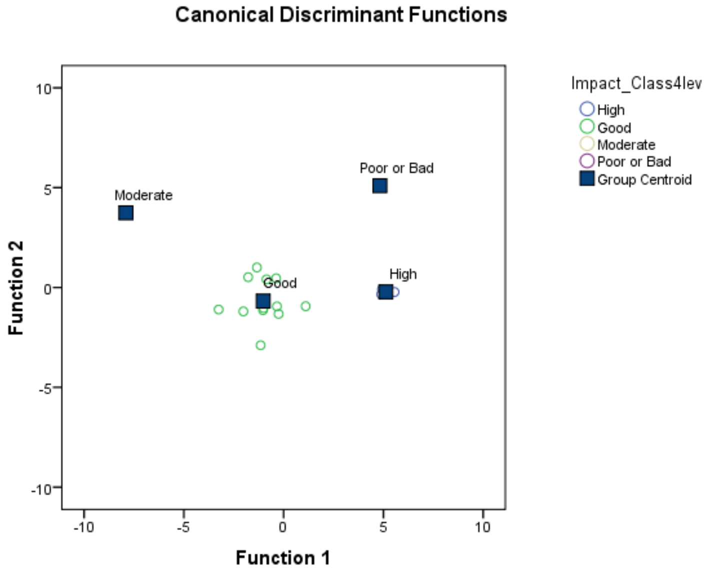
%at2_BIO



RUDD_IND

Typology-specific boxplots for the metrics identified by the discriminant analysis

High alkalinity deep lakes: 5 fish metrics significant



Plot of the first two discriminant functions based on the 5 metrics

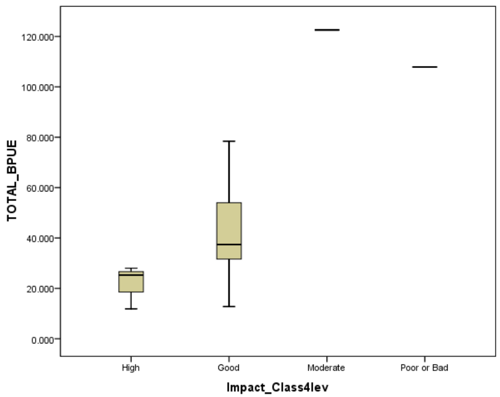
Classification of the discriminant analysis rule

	Predicted Impact Class				
Actual Impact Class	**High**	**Good**	**Mod**	**Poor/ Bad**	**Total**
High	**3**	0	0	0	3
Good	0	**12**	0	0	12
Moderate	0	0	**1**	0	1
Poor or Bad	0	0	0	**1**	1
Total	**3**	**12**	**1**	**1**	**17**

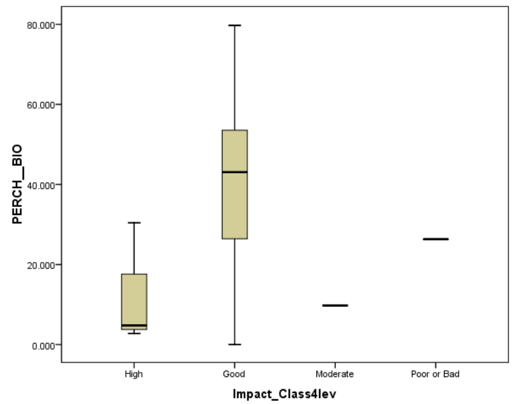
98.3% of the variability was explained by the first two discriminant functions of the discriminant rule

100% correct classification of the lakes (Table 18) and when jack-knifing 82.4% of the lakes were correctly classified

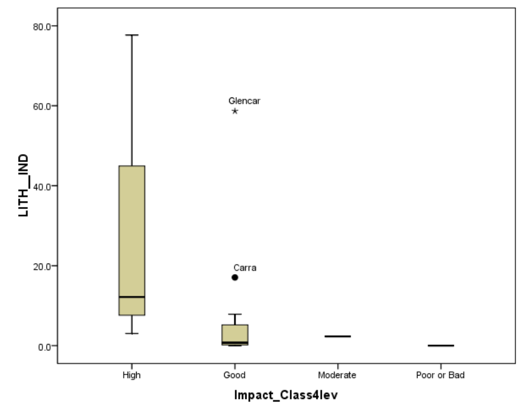
High alkalinity deep lakes: 5 fish metrics



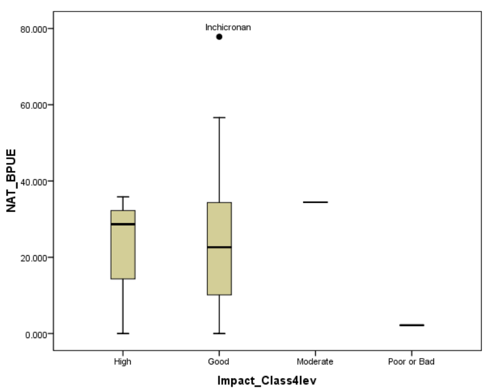
TOT_BPUE



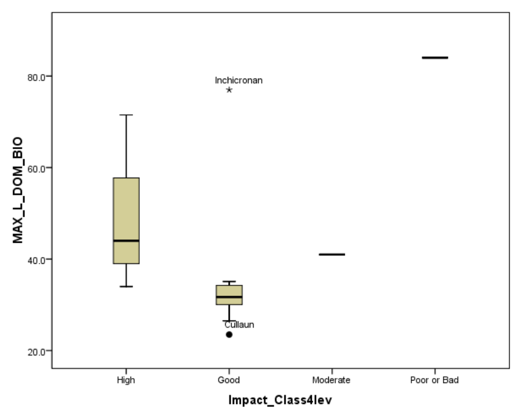
PERCH_BIO



LITH_IND



NAT_BPUE



MAX_L_DOM

Typology-specific boxplots for the metrics identified by the discriminant analysis

3. Developing the EQR model (& setting boundaries)

1. Create a pseudo EQR as a response to develop the model (starting point)
2. EQR model building
3. Set boundaries

Creating a Pseudo EQR

Two weighted scores were created by:

1. Multiplying mid-points of the impact class using TP mean or Chlor max with the posterior probabilities for each impact class from the DA rules and summing - this gives a weighted value
2. Scores for the test dataset were assumed to be normally distributed
3. A z score was calculated and associated probability/EQR was derived
4. The two EQRs were derived and averaged to create an overall EQR
5. Lakes were ranked according to the pseudo EQR and EO used to assess if in correct relative order

Posterior probability for each lake is the probability that the lake belongs to a particular impact class based on the typology specific fish metrics

Typology	**Lake**	**P(1) High**	**P (2) Good**	**P(3) Moderate**	**P(4) Poor/bad**
Low alk/shallow	Akibbon	0.1327	0.0975	**0.7675**	0.0023
Low alk Deep	Glencullin	**0.7277**	0.2383	0.0339	0.0001
High alk/shallow	Egish	0.0198	0.0623	0.2829	**0.6349**
High alk/deep	Glencar	0.000	**1.000**	0.000	0.000

EQR Model Building

- The pseudo EQR was assumed to be logit-normally distributed
- The logit normal pseudo EQR was regressed against the fish metrics (sparse excluded)
- Only fish metrics that were significant at 25% were brought forward for the typology specific multivariate regression analysis
- Most of the fish metrics used in the DA classification were significant
- Output is an EQR and confidence limits for each lake

EQR Model: Low alk/shallow lakes

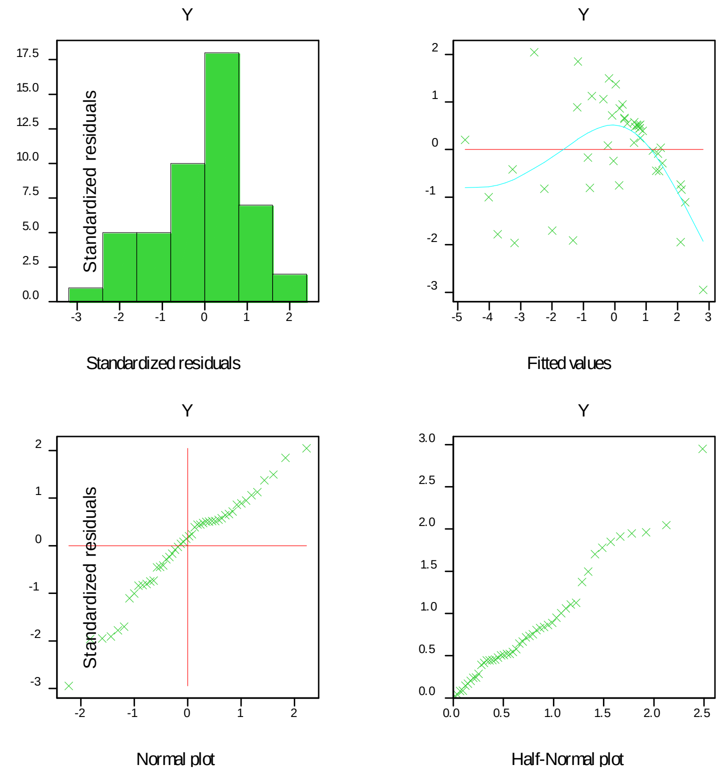
TOTAL_BPUE
NAT_BPUE
SPE_EVEN
RHEO_BIO
ROACH_BIO

significant ($p<0.05$)

PERCH_BIO

not significant ($P=0.958$)

Model explained 86% of the variation within the data



Residual diagnostics of the EQR model for Low ALK/Shallow lakes

End model=

$$\operatorname{logit}(EQR_{LowALK/ShallowDepth}) = -6.340 - 0.065 \times TOTAL\_BPUE + 0.027 \times NAT\_BPUE + 0.057 \times PERCH\_BIO + 2.325 \times SPE\_EVEN + 0.054 \times RHEO\_BIO + 0.090 \times ROACH\_BPUE$$

EQR Model: Low alk/deep lakes

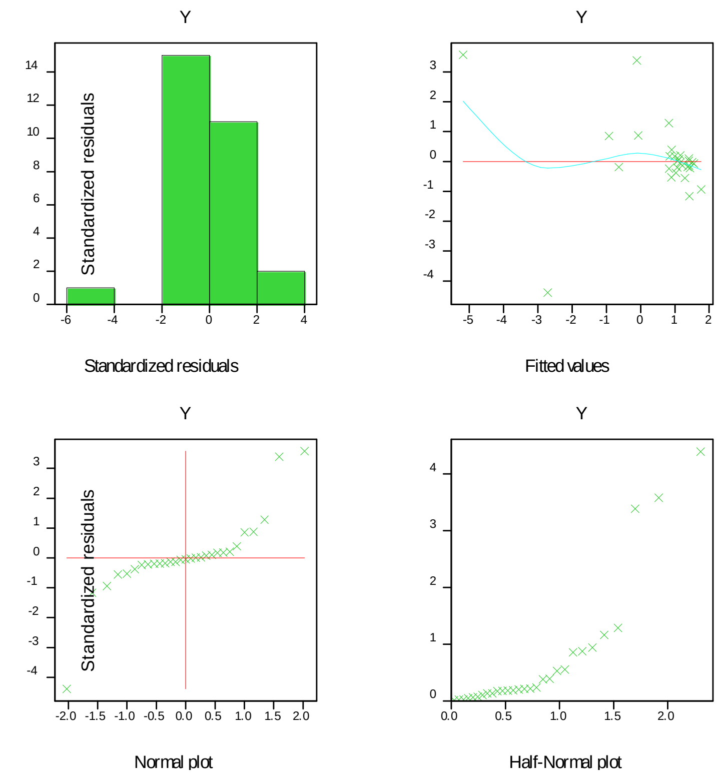
TOTAL_BPUE
PHYT_BIO
BREAM_IND

significant ($p<0.05$)

NAT_BPUE
PERCH_BIO

not significant ($p>0.05$)

Model explained 85.5% of the variation within the data



Residual diagnostics of the EQR model for Low ALK/Deep lakes

End model =

$$\operatorname{logit}\left(EQR_{LowALK/DeepDepth}\right) = 1.666 - 0.001 \times TOTAL\_BPUE - 0.014 \times NAT\_BPUE + 0.003 \times PERCH\_BIO - 0.063 \times PHYT\_BIO - 0.188 \times BREAM\_IND$$

EQR Model: High alk/shallow lakes

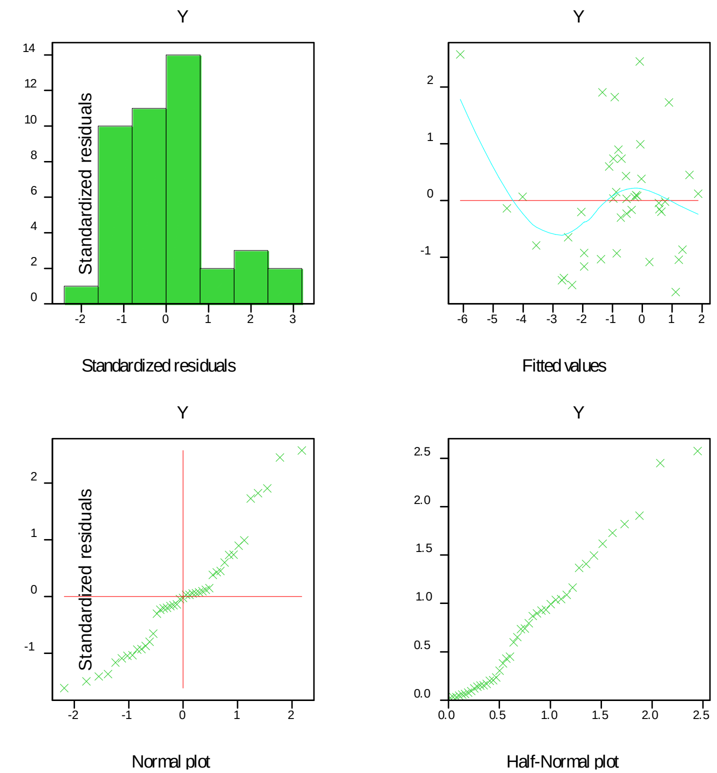
TOTAL_BPUE
PERCH_BIO

significant ($p<0.05$)

NAT_BPUE
%_2_BIO
CYP_BIO
RUDD_IND

not significant ($P>0.05$)

Model explained 85.6% of the variation within the data



Residual diagnostics of the EQR model for High ALK/shallow lakes

End model

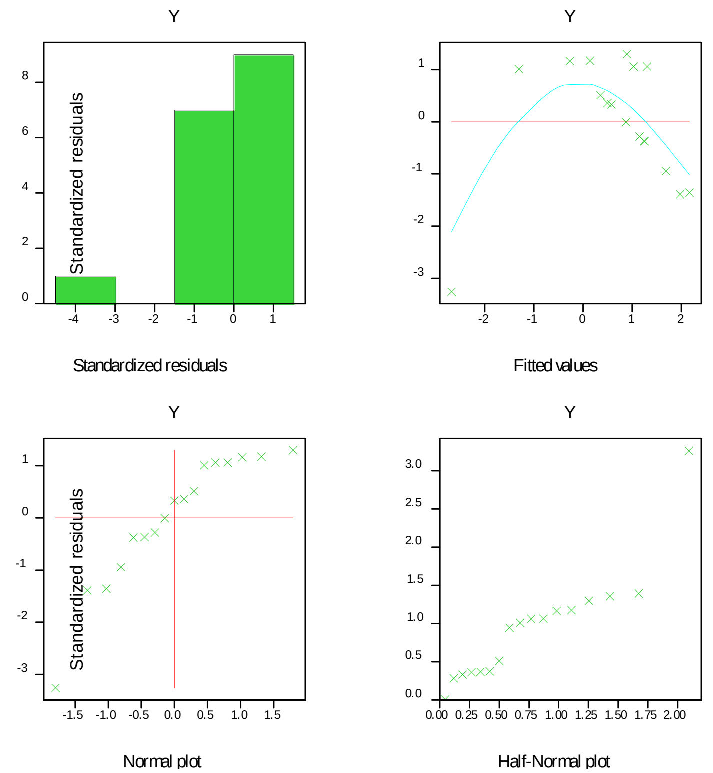
$$
\begin{aligned}
logit(EQR_{HighALK/ShallowDepth}) = {} & 0.059 - 0.032 \times TOTAL\_BPUE + 0.021 \times NAT\_BPUE + \\
& 0.027 \times PERCH\_BIO + 0.001 \times \%at2\_BIO + \\
& 0.010 \times CYP\_BIO - 0.018 \times RUDD\_IND
\end{aligned}
$$

EQR Model: High alk/deep lakes

TOTAL_BPUE } significant (P<0.05)

NAT_BPUE
PERCH_BIO
MAX_L_DOM_BIO
LITH_IND
} **not significant (P>0.05)**

Model explained 63.6% of the variation within the data



Residual diagnostics of the EQR model for High ALK/shallow lakes

End model

$$logit\left(EQR_{HighALK/DeepDepth}\right) = 3.390 - 0.029 \times TOTAL\_BPUE + 0.008 \times NAT\_BPUE - 0.005 \times PERCH\_BIO - 0.034 \times MAX\_L\_DOM\_BIO - 0.004 \times LITH\_IND$$

EQR class boundaries

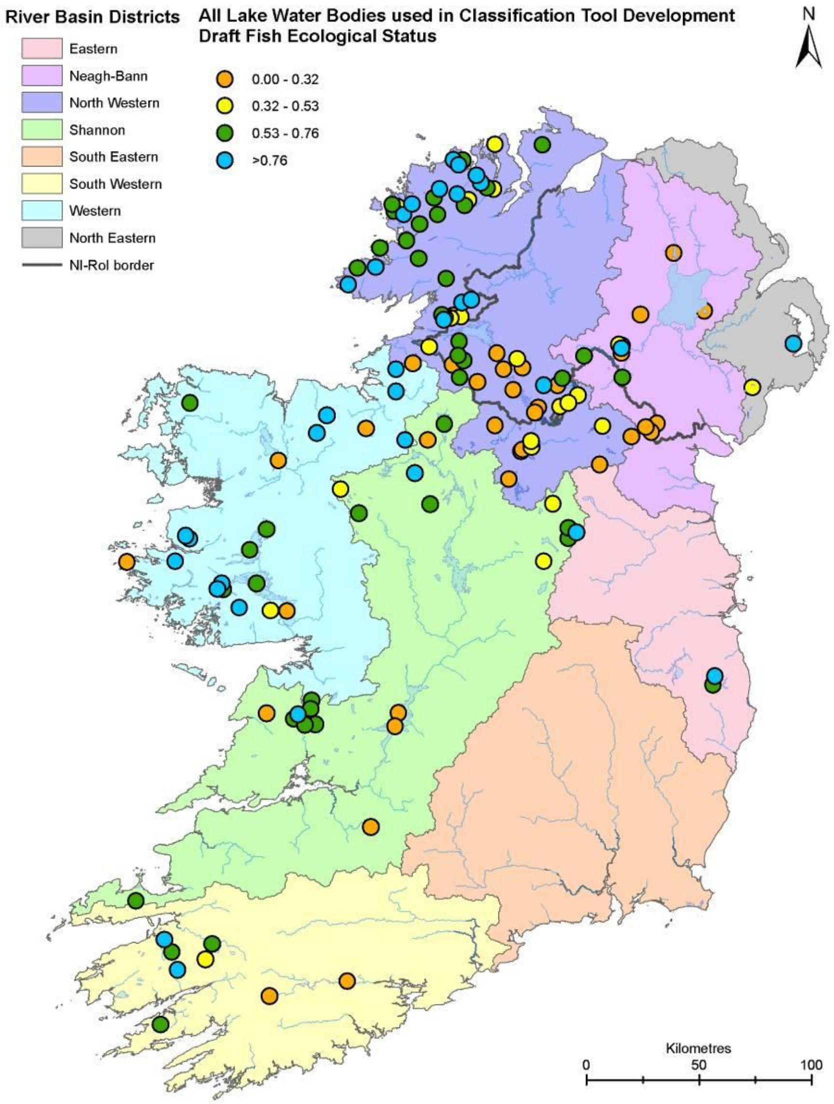
- Boundaries to be set (High/Good, Good/moderate, Moderate/Poor and Poor/Bad
- A range of cut points were investigated
- EQR impact classes were cross-tabulated with the DA impact classes
- Cut-point was determined when the maximum correct classification from the cross-tabulation of EQR impact class was achieved for that impact class
- **Resulted in an overall agreement between the EQR model and the DA rules of 56.9%**

Boundary	**EQR**
High/Good	0.76
Good/moderate	0.53
Moderate/poor-bad	0.32

Classification method		**DA**			
		High	**Good**	**Mod**	**Poor/Bad**
EQR	**High***	**14**	19	0	0
	Good*	9	**31**	3	0
	Moderate*	3	12	**8**	0
	Poor/Bad*	1	7	5	**25**

Do we need to set a boundary for poor/bad (EO?)

Classification of lakes in Ecoregion 17 with FIL2



High = 33 lakes

Good = 43 lakes

Moderate = 23 lakes

Poor/bad = 38 lakes

Example of FIL2 output

Redefined Typology	Lake name	Pseudo EQR	EQR	EQR Lower CI	EQR Upper CI	EQR CP	DA Classification	Chl and TP Classification
Low ALK/Shallow Depth	Glencullin	0.8295	**0.8911**	**0.8397**	**0.9274**	High	High	High
Low ALK/Deep Depth	Beagh	0.8049	**0.8027**	**0.7395**	**0.8536**	High	High	Good
Low ALK/Deep Depth	Altan	0.8194	**0.8252**	**0.7508**	**0.8809**	High	High	High
High ALK/Shallow Depth	New	0.3918	**0.5591**	**0.365**	**0.7367**	Good	Moderate	Moderate
Low ALK/Shallow Depth	Meela	0.67	**0.6489**	**0.4722**	**0.7925**	Good	High	Good
High ALK/Deep Depth	Carra	0.7046	**0.7599**	**0.5859**	**0.8762**	Good	Good	Good
High ALK/Shallow Depth	Sheelin	0.8313	**0.4779**	**0.3456**	**0.6133**	Moderate	High	Moderate
Low ALK/Deep Depth	Melvin	0.6047	**0.4822**	**0.3337**	**0.6339**	Moderate	Moderate	Good
High ALK/Deep Depth	Owel	0.7046	**0.4327**	**0.2296**	**0.6613**	Moderate	Good	Good
High ALK/Deep Depth	Muckno	0.0091	**0.0646**	**0.0117**	**0.2865**	Poor/Bad	Poor/Bad	Poor/Bad
High ALK/Shallow Depth	Gur	0.0102	**0.0105**	**0.0026**	**0.0414**	Poor/Bad	Poor/Bad	Poor/Bad
High ALK/Shallow Depth	Egish	0.0589	**0.1242**	**0.0814**	**0.185**	Poor/Bad	Poor/Bad	Poor/Bad

Next Steps

- Test data from Scotland and England and Wales to investigate if model appropriate to use
- Complete intercalibration and harmonise boundaries

Test data from Scotland and England and Wales to investigate if model appropriate to use

- 6 Scottish lakes
  - Loch Builg, Loch Doon, Loch Eck, Loch Insh, Loch Girlsta and Loch Lomond
  - Good agreement with previous classifications (Ian Winfield, CEH)
  - Scotland to adopt FIL2 (provisional) but will test further lake data
  - 4/4/2011 – received data from an additional 20 lakes
    - 3 SEPA surveys (Irish method)
    - 17 Marine Scotland surveys (reduced CEN)
  - FIL2 EQRs and status sent (4/4/2011) to SEPA, awaiting feedback

- 9 lakes England and Wales
  - Bassenthwaite Loch, Wastwater, Loweswater, Windermere, Llyn Tegid, Rostherne Mere, etc.
  - Classifications sent to EA and CEH
  - Initial feedback good (Ian Winfield, CEH)
  - Will circulate further (EA)

3 additional species

Intercalibration

- Complete intercalibration and harmonise boundaries – Northern GIG
  - Option 3 (same sites)
    - FIL2 Vs EQR4
    - FIL2 Vs EQR8
    - FIL2 Vs Norway tool
  - Option 2 – (no common sites) - common metric
    - FIL2, EQR4, EQR8 and NORWAY TOOL

Irish fish species Vs Nordic fish species

Group 1 (Natives)	Group 2 (Non-natives influencing ecology)	Group 3 (Non-natives generally not influencing ecology)
Brown trout	Roach	Tench
Sea trout	Perch	Rudd
Salmon	Pike	Stoneloach
Char	Bream	Gudgeon
Pollan	Dace	
Eel	Carp	
Shad	Rainbow trout	
3-spined stickleback	Chub	
9-spined stickleback	Minnow	
Brook lamprey		
River lamprey		
Sea lamprey		
Flounder		

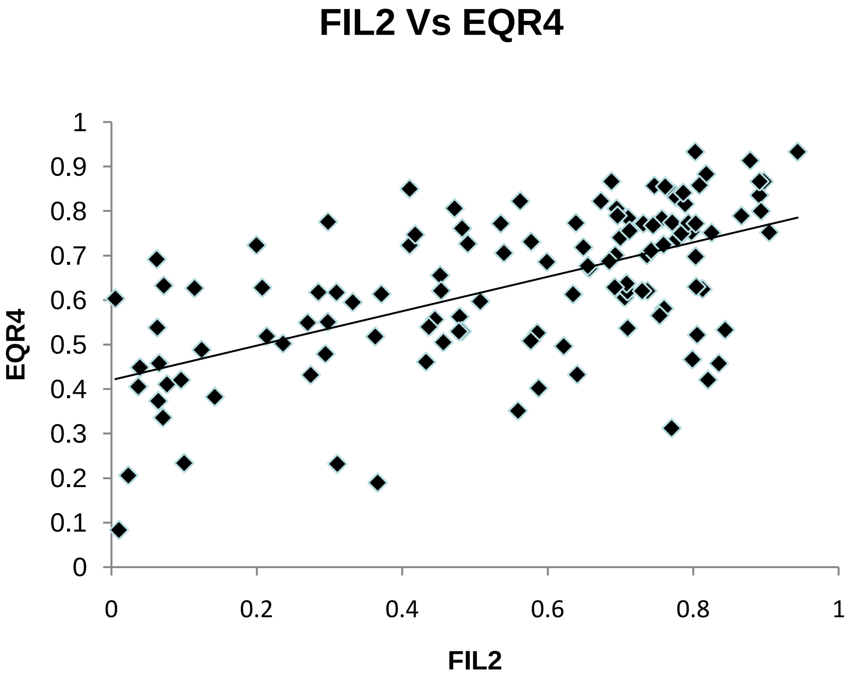
23 additional fish species, what categories to group them into?

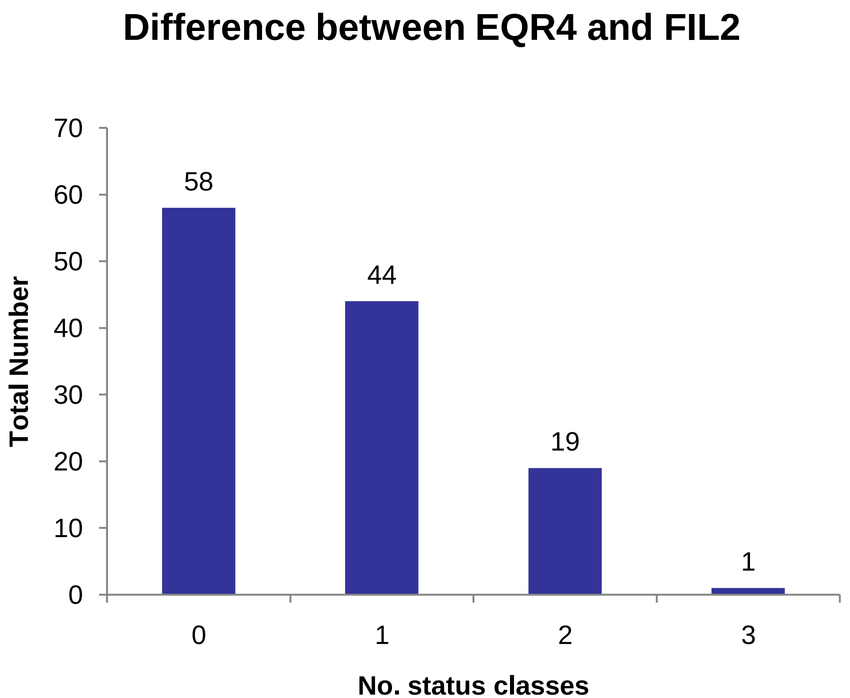
Classifying Nordic fish species into groups for input to FIL2

Group 1 (Natives)	Group 2 (Non-natives influencing ecology)	Group 3 (Non-natives generally not influencing ecology)
Brown trout	Roach	Tench
Sea trout	Perch	Rudd
Salmon	Pike	Stoneloach
Char	Bream	Gudgeon
Pollan	Dace	
Eel	Carp	
Shad	Rainbow trout	
3-spined stickleback	Chub	
9-spined stickleback	Minnow	
Brook lamprey		
River lamprey		
Sea lamprey		
Flounder	**Ruffe, pikeperch, whitebream**	**Bullhead, alpine bullhead**
Vendace, whitefish, peled	**Ide, dace, spirlin,**	**burbot**
Smelt, grayling	**zope, bleak, asp, etc.**	

FIL2 Vs EQR4 – Preliminary analysis

Irish sites



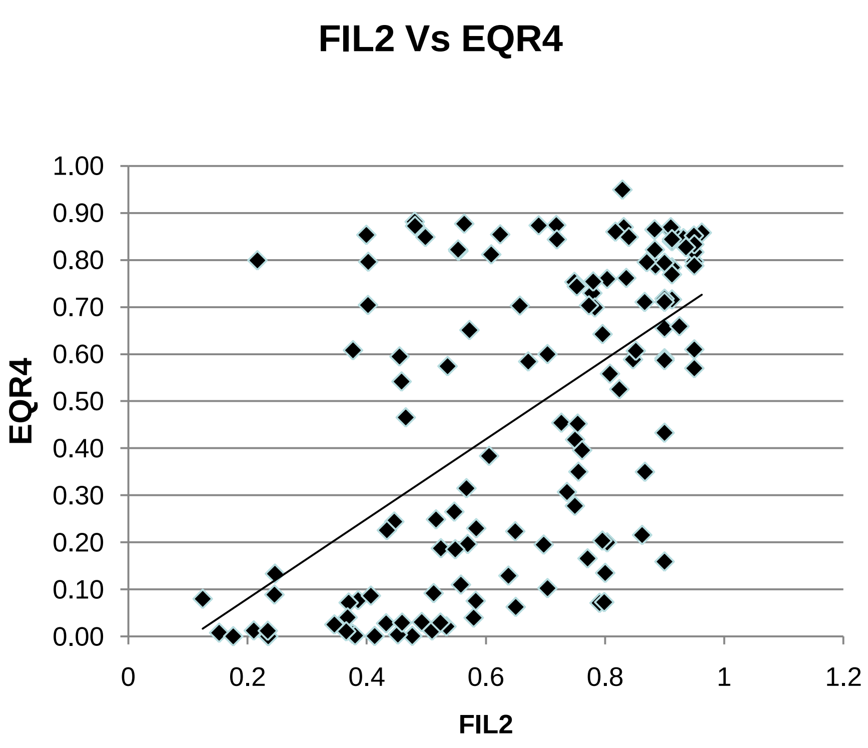


$r^2 = 0.348$
$r = 0.59$

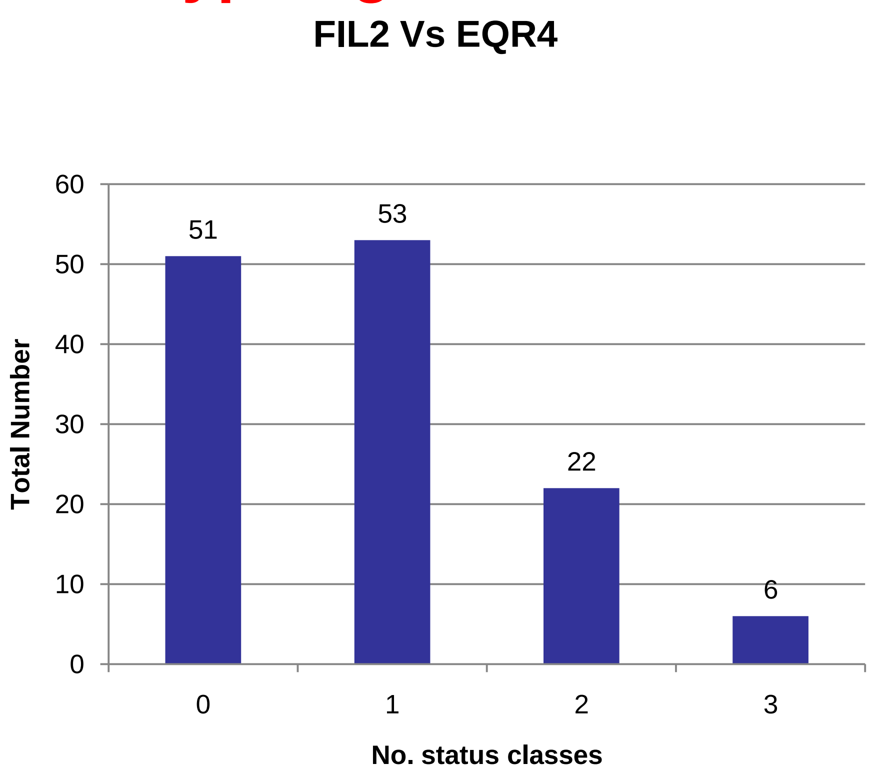
44% same status
33% one class diff
14% 2 classes diff
1% 3 classes diff

FIL2 Vs EQR4 – Preliminary analysis

Finnish Sites – 2 typologies



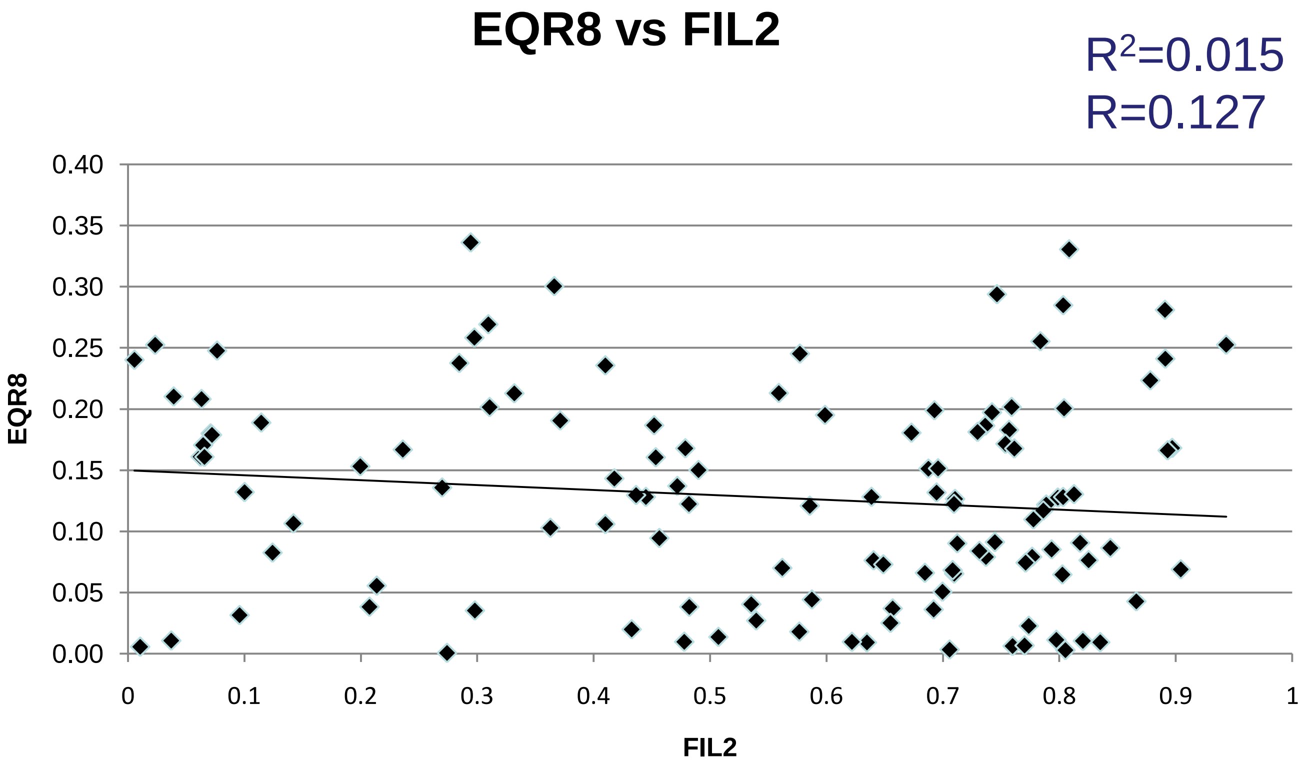
$R^2 = 0.328$
R=0.575



39% - same status
40% - one class diff
17% - 2 classes diff
5% - 3 classes diff

FIL2 Vs EQR8 – Preliminary analysis

Irish sites



EQR8 classified Irish lakes as Moderate, poor and bad

Why is correlation so poor? Non-native species?

Work to be completed

1. Complete calculation of metrics for sites from Sweden and Norway
2. Examine relationship between EQR8 and FI and FIL2
3. Examine relationship between FIL2 and common metric
4. Continue intercalibration and harmonise boundaries

Thank You

Website: www.wfdfish.ie