http://efi-plus.boku.ac.at/

Project no.: 0044096

Project acronym: EFI+

Improvement and spatial extension of the European Fish Index

Instrument: STREP

Thematic Priority: Scientific Support to Policies (SSP) - POLICIES-1.5

D3.1 Report on hydromorphological pressures
D3.2 Global pressure index

Due date of deliverable: 31.12.2007

Actual submission date: 07.05.2008

Start date of project: 01.01.2007 Duration: 24 Month

Organisation name of lead contractor for this deliverable:

Institute of HYDROBIOLOGY and AQUATIC ECOSYSTEM MANAGEMENT
Department of WATER, ATMOSPHERE and ENVIRONMENT
BOKU - UNIVERSITY of NATURAL RESOURCES and APPLIED LIFE SCIENCES

Responsible authors: Rafaela Schinegger, Andreas Melcher and Stefan Schmutz

Final version

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission	
RE	Restricted to a group specified by the consortium (including the Commission	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of contents

1. Introduction

The objective of Task 3.1 was to identify human drivers and pressures affecting fish, to analyse interrelations among pressures, to define reference conditions and to produce a global index of all pressure types.

Out of the list of potential pressure variables defined in WP1 and collected in WP2, pressures that show correlations with key fish species were selected. The relationships between drivers (e.g. hydropower, flood protection or agriculture) and resulting pressures (e.g. altered flow, altered habitat conditions or waterquality pressures) were analysed and key drivers in different regions and river types identified. Furthermore, redundant pressure variables were eliminated, synergistic effects were localised and interrelations among pressures were analysed. Finally, pressures were combined into indices that describe particular pressure types (e.g. hydrology, morphology, continuity and waterquality) and further aggregated into global pressure indices (3 different approaches).

The first part of this report describes the pressure variables collected in WP2 and deals with the definition of the dataset used for pressure index calculation. Aside, 5 river types defined as a function of environmental gradients on a European scale are identified and described.

The second part deals with classification and selection of pressure variables for index development by pointing out redundant variables using principle component analysis and correlation tables.

Finally, in the third part, different approaches for calculation of a global pressure index are described and discussed. One of these indices will be used in WP4 to separate reference from degraded conditions and to calibrate the EFI+.

1.1 Selection of pressure analysis dataset

In order to calculate pressure indices in the same manner for all available sites, it was necessary to filter the whole dataset (in total 14221 sites and 29509 fishing occasions), by using as many completed pressure variables as possible (where modality "NoData" is not accepted), or by including as many sites as possible. An alternative method for filtering the data prevented the loss of all sites from any one country. Therefore, 4 variants of the dataset were evaluated:

Tab. 1.1: Showing 4 different "Filter variants" to select the dataset for pressure analysis.

Filter name:	Filter_total_pressure	Filter_max_pressure	Filter_consenus_pressure	Filter_min_pressure
Country_abbreviation	Nr. of sites	Nr. of sites	Nr. of sites	Nr. of sites
AT	847	847	873	879
CH			500	502
DE	593	750	781	781
ES	304	2039	2098	2099
FI	278	278	278	278
FR	89	657	785	989
HU	61	193	193	193
IT	304	322	498	517
LT		55	94	94
NL	182	182	182	182
PL	127	898	907	911
PT	155	923	923	923
RO	254	263	263	263
SE	117	605	605	605
UK			1228	1228
Total	3311	8012	10208	10444
% of total dataset	**23,28**	**56,34**	**71,78**	**73,44**
Nr. of variables complete:	**24**	**20**	**18**	**15**
Description	Excluding all incomplete pressures datasets (NoData for the following variables): Fished_area Wetted_width Run1 H_imp H_hydrop H_waterabstr H_hydromod H_tempimp H_veloincr H_resflush M_sed M_channel M_crossec M_instrhab M_ripveg M_embank M_floodpro W_toxic W_acid W_eutroph W_opoll W_osilt C_B_s_up C_B_s_do	Excluding many incomplete pressures datasets (NoData for the following variables): Fished_area Wetted_width Run1 H_imp H_hydrop H_waterabstr H_hydromod H_resflush M_sed M_channel M_crossec M_instrhab M_embank M_floodpro W_toxic W_acid W_eutroph W_opoll C_B_s_up C_B_s_down	Excluding some incomplete pressures datasets (NoData for the following variables): Fished_area Wetted_width Run1 H_imp H_hydrop H_waterabstr H_hydromod H_resflush M_channel M_crossec M_instrhab M_embank M_floodpro W_acid W_eutroph W_opoll C_B_s_up C_B_s_down	Excluding few incomplete pressures datasets (NoData for the following variables): Fished_area Wetted_width Run1 H_imp H_hydrop H_waterabstr H_hydromod H_resflush M_channel M_crossec M_instrhab M_embank M_floodpro C_B_s_up C_B_s_down

As Tab. 1.1 shows, data were available for 29509 fishing occasions (14221 sites), but complete reporting of all pressure variables was only available for 3311 sites (23,3% of all sites). Using this filter method, 3 whole countries (CH, LT, UK) were lost. Thus, the number of variables that could be used to produce pressure indices was limited. Only 56,3% of data (8012 sites; loosing 2 countries) were available for 20 complete pressure variables.

Therefore, Filter_consensus was chosen for EFI+ pressure analysis, with 18 “complete” variables, and all countries included, representing 10208 sites (71% of all sites). Because pressure information was included in the fishing occasion table, this table contained a lot of replicates for each fishing site (in the total dataset 29509 fishing occasions on 14221 sites). To select only one site per fishing occasion, we therefore selected the youngest date for a site with multiple fishing occasions, after filtering for only the complete variables.
For consideration of standards as the EN 14962 („Water quality – guidance on the scope and selection of fish sampling methods“, CEN 2004), the dataset has been screened according standards like minimum fished area of 100m² or the rule that fished length must at least be 10-times as long the wetted width (for method wading). But these facts have not been considered for pressure analysis because of minor importance for this step. When applying the Filter_consensus as described before, the output of 10208 sites has been taken for all further pressure analyses in this document.

1.2 Nr. of EFI+ sites per country

Fig. 1.2 shows the spatial distribution of sites included in the final dataset. In addition to 13 partners out of the EFI+ consortium, the Netherlands (Institute for Inland Water Management and Waste Water Treatment) and Lithuania (Institute of Ecology) contributed several sites.

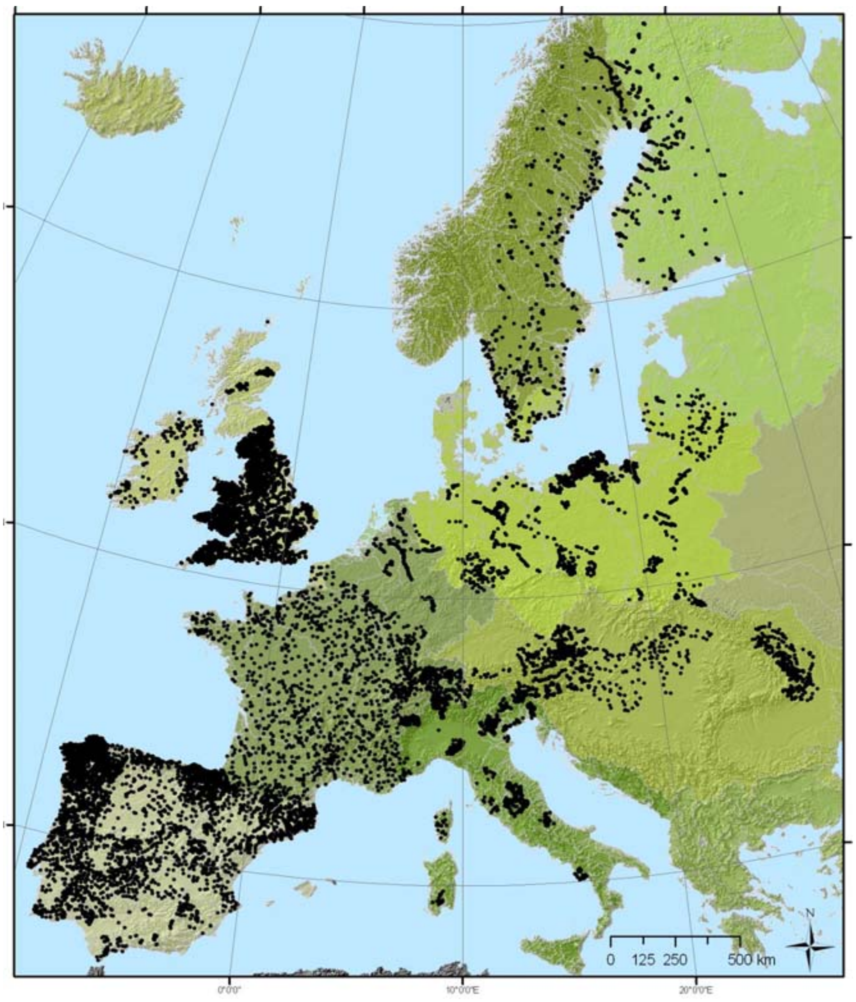


Fig. 1.2.: Spatial distribution of EFI+ sites (Copyright Clemens Trautwein)

Tab. 1.2: A summary and comparison of the number of sites per country in the final datasets.

Country	Nr. of sites selected for pressure analysis	Total nr. of sites in table site/ EFI+ DB	Total nr. of fishing occasion in table fishing occasion/ EFI+ DB
AT	873	938	1172
CH	500	717	969
DE	781	803	1817
ES	2098	4239	5189
FI	278	530	530
FR	785	1145	6570
HU	193	193	193
IT	498	652	1152
LT	94	115	130
NL	182	182	790
PL	907	919	978
PT	923	923	923
RO	263	263	323
SE	605	615	5652
UK	1228	1987	3162
Total	10208	14221	29550

1.3 Pressure groups

As a basis for pressure analyses, the following five pressure groups have been defined in WP1 (see deliverables D. 1.1 – 1.3) and collected in WP 2 (see deliverables D. 2.1 -2.2): hydrological pressures, morphological pressures, connectivity pressures, waterquality pressures and other pressures (that don't fit into these 4 groups).

Each group consists of 2-8 variables giving information on human pressures through direct or indirect descriptors. The variables were defined during preliminary consortium meetings and finalized at the EFI+ Kick-off meeting in January 2007 in Bratislava, to guarantee that each country and geographic region, with its specific pressures, are represented in the EFI+ database. Before doing any further analyses, each variable indicated in the tables has been coded (according a scheme from 1- 5, 1=best, 5=worst), to transform these variables into "numeric values". The coding scheme has been carefully considered and was applied to all variables in the same manner:

Each of the five numbers can be considered as semi-quantitative, with five modalities ranging from

"No" = 1,

"Slight" = 2,

"Weak" or "Intermediate" = 3 and

"High" or "Strong" = 5.

In the following section, descriptive analysis of the pressure variables is described, including affiliation between pressures and groups, pressure intensity, and the total percentage of the dataset completed per variables.

In addition to all pressures collected by partners, information on land-use, population density, irrigated land, and road density was calculated via GIS, on a European scale, for all fishing occasions (N=29550). These variables have already been calculated and will be compared with the pressure information compiled by partners afterwards. *

* In deliverables D. 3.1 and D.3.2, these variables have not been considered for analyses.

Tab. 1.3.1: Overview on all variables representing the group "hydrological pressures". Information on categories of pressure intensity as well as short names for analyses (Varible_code), numeric coding (Coding) and completeness of datasets (% of total sites collected) are described here for each variable.

Category	Variable_code	Pressure type	Criterion (unit)	No pressure	Pressure intensity				Total %
Hydrological pressures	H_imp	Impoundment	Natural flow velocity reduction on site due to impoundment	no (no impoundment)	weak	strong			**95.1**
		Coding		**1**	**3**	**5**			
	H_hydrop	Hydropeaking	Site affected by hydropeaking	no (no hydropeaking)	partial	yes			**98.8**
		Coding	**Frequencies**	**1**	**3**	**3**			
	H_waterabstr	Water abstraction	Is the site affected by water flow alteration/minimum flow (water abstraction)	no (no water abstraction)	weak to medium (less than half of the mean annual flow)	strong (more than half of mean annual flow)			**89,3**
		Coding		**1**	**3**	**5**			
	H_resflush	Reservoir flushing	Is the fish fauna affected by flushing of reservoirs upstream of the site?	no	yes				**98.4**
		Coding		**1**	**3**				
	H_hydromod	Hydrograph modification	Seasonal hydrograph modification due to hydrological alteration (e.g. water storage for irrigation, hydropower,…)	no	yes				**97.8**
		Coding		**1**	**3**				
	H_tempimp	Temperature	Is there an impact on water temperature	no (no temperature modification)	permanent increase (due to cooling water input e.g. for nuclear or thermal power plant, or others)	permanent decrease (e.g. permanent temperature reduction due to coldwater input from impoundment)	summer increase (temperature increase due to epilimnion release)	summer decrease (temperature decrease and/or winter increase due to hypolimnion water release)	**86.8**
		Coding		**1**	**3**	**3**	**3**	**3**	
	H_veloincr	Flow velocity increase	Is there an impact on flow conditions (mean velocity) due to channelisation, floodprotection, etc.	no	yes				**84.2**
		Coding		**1**	**3**				

Tab. 1.3.2: Overview on all variables representing the group “morphological pressures”. Information on categories of pressure intensity as well as short names for analyses (Variable_code), numeric coding (Coding) and completeness of datasets (% of total sites collected) are described here for each variable.

Category	Variable_code	Pressure type	Criterion (unit)	No pressure	Pressure intensity				Total %
Morphological pressures	M_channel	Channelisation	Alteration of natural morphological channel plan form; intensity of straightening	no	intermediate	straightened			**93.8**
		Coding	**Frequencies**	**1**	**3**	**5**			
	M_crosssec	Channelisation	Alteration of cross section	no	intermediate	technical crossec./U-profil			**92.9**
		Coding		**1**	**3**	**5**			
	H_veloincr	Flow velocity increase	Is there an impact on flow conditions (mean velocity) due to channelisation, floodprotection, etc.	no	yes				**84.2**
		Coding		**1**	**3**				
	M_instrhab	Channelisation	Alteration of instream habitat conditions	no	intermediate	high			**92.6**
		Coding	**Frequencies**	**1**	**3**	**5**			
	M_embankm	Channelisation	Artificial embankment	no (natural shoreline)	slight (local presence of artifical material for embankment)	intermediate (continuous embankment but permeable (e.g. rip rap))	high (continuous, no permeability (e.g. concrete walls))		**91.6**
		Coding		**1**	**2**	**3**	**5**		
	M_ripveg	Channelisation	Alteration of riparian vegetation close to shoreline	no	slight	intermediate	high (no riparian vegetation)		**84.4**
		Coding		**1**	**2**	**3**	**5**		
	M_floodpr	Floodprotection	presence of dykes for flood protection	no	yes				**92.1**
		Coding		**1**	**3**				
	M_sediment	Sedimentation	Input of fine sediment (mainly mineral input; bank erosion, erosion from agricultural land)	no	weak (slight reduction of sediment porosity)	medium	high (coarse sediment clogged)		**89.6**
		Coding		**1**	**2**	**3**	**5**		
	M_remfloodpl	Floodprotection	If the river has a former floodplain - Proportion of connected floodplain still remaining. Floodplain = area connected during the flood.	> 50%	10-50%	less than 10 %	some waterbodies remaining	no	**75.5**
		Coding		**2**	**3**	**4**	**4**	**4**	

Tab. 1.3.3: Overview on all variables representing the group “connectivity pressures”. Information on categories of pressure intensity as well as short names for analyses (Varible_code), numeric coding (Coding) and completeness of datasets (% of total sites collected) are described here for each variable.

Category	Variable_code	Pressure type	Criterion (unit)	No pressure	Pressure intensity		Total %
	C_B_s_up	Migration barriers - river segment scale	Barriers on segment level upstream	no	partial	yes	**98.3**
		Coding		**1**	**3**	**3**	
	C_Bn_s_up	Migration barriers - river segment scale	Number of barriers upstream	no = 0	Seg 1 (less than 1 or 1), Seg 5 (less than 2 or 2), Seg 10 (less than 3 or 3),	Seg 1 (lmore than 1), Seg 5 (lmore than 2), Seg 10 (lmore than 3),	**82.7**
		Coding		**1**	**3**	**4**	
	C_Bd_s_up	Migration barriers - river segment scale	Distance to next barrier in the segment - upstream	0 = no barrier	Seg 1 (more than 250m), Seg 5 (more than 1250m), Seg 10 (more than 2500m),	Seg 1 (less than 250m), Seg 5 (less than 1250m), Seg 10 (less than 2500m),	**34.0**
		Coding		**1**	**3**	**4**	
Connectivity pressures	C_B_s_do	Migration barriers - river segment scale	Barriers on segment level downstream	no	partial	yes	**98.3**
		Coding		**1**	**4**	**4**	
	C_Bn_s_do	Migration barriers - river segment scale	Number of barriers downstream	no = 0	Seg 1 (less than 1 or 1), Seg 5 (less than 2 or 2), Seg 10 (less than 3 or 3),	Seg 1 (lmore than 1), Seg 5 (lmore than 2), Seg 10 (lmore than 3),	**81.3**
		Coding		**1**	**3**	**4**	
	C_Bd_s_do	Migration barriers - river segment scale	Distance to next barrier in the segment - downstream	0 = no barrier	Seg 1 (more than 250m), Seg 5 (more than 1250m), Seg 10 (more than 2500m),	Seg 1 (less than 250m), Seg 5 (less than 1250m), Seg 10 (less than 2500m),	**31.0**
		Coding		**1**	**3**	**4**	
	C_B_c_do	Migration barriers - catchment scale	Presence of downstream barriers on the catchment scale	no	partial: migration possible for good swimmers (e.g. salmon) or for particular situations/years	yes: definite barriers for most species most of the time	**98.5**

Tab. 1.3.4: Overview on all variables representing the groups "waterquality pressures" and "others". Information on categories of pressure intensity as well as short names for analyses (Variable_code), numeric coding (Coding) and completeness of datasets (% of total sites collected) are described here for each variable.

Category	Variable_code	Pressure type	Criterion (unit)	No pressure	Pressure intensity				Total %
Waterquality pressures	W_osilt	Water quality	Is organic siltation observed?	no	yes				**84.3**
		Coding		**1**	**3**				
	W_toxic	Water quality	Toxic prioriy substances (organic and nutrient appearance)	no or very minor (e.g. atmospheric input far away, no contamination in the segment itself)	weak (important risk, link to particular substance)	high concentration (a clearly kown input)			**88.8**
		Coding	**Frequencies**	**1**	**3**	**5**			
	W_acid	Water quality	Acifidication	no	yes				**97.4**
		Coding		**1**	**3**				
	W_eutroph	Water quality	Is there an artificial eutrophication?	no	low	intermediate (occurance of green algae)	extreme (oxygen depletion, increase of primary production)		**93.7**
		Coding		**1**	**3**	**4**	**5**		**88.8**
	W_opoll	Water quality	Is organic pollution observed?	no	weak	strong			
		Coding		**1**	**3**	**5**			
	W_index	Water quality	National water quality index - in 5 classes	indication of classes (1 to 5; 5 = worst)	2	3	4	5	**83.0**
		Coding	**Frequencies**	**1**	**2**	**3**	**4**	**5**	
Others	O_nav	Navigation	Navigation intensity	no	low	medium	strong (e.g. large ships)		**88.2**
		Coding		**1**	**2**	**3**	**4**		
	O_cores	Colinear connected reservoir	Presence, absence of fish farms, fish ponds and agricultural use, must be connected to the river:	no	yes				**91.2**

Tab. 1.3.5: Overview on completeness of datasets per variable and country “OK” indicates the number of sites with valid values for each variable; N/A gives information on missing data for each variable. Fields marked in yellow represent pressures with critical number of sites without valid information on specific pressures. These pressures will not be considered for the pressure index (see chapter 3).

		Country name																													
		AT		CH		DE		ES		FI		FR		HU		IT		LT		NL		PL		PT		RO		SE		UK	
		OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A	OK	N/A
	H_imp	937	1	717	0	803	0	4239	0	304	226	1107	38	193	0	652	0	115	0	182	0	919	0	923	0	263	0	615	0	1262	725
	H_hydrop	934	4	717	0	803	0	4239	0	304	226	1115	30	193	0	652	0	114	1	182	0	916	3	923	0	263	0	615	0	1987	0
	H_waterabstr	931	7	717	0	803	0	2417	1822	304	226	1081	64	193	0	564	88	115	0	182	0	919	0	923	0	263	0	615	0	1987	0
	H_hydromod	938	0	717	0	803	0	4239	0	304	226	1093	52	193	0	563	89	95	20	182	0	915	4	923	0	263	0	615	0	1980	7
	H_tempimp	910	28	717	0	803	0	4239	0	304	226	1079	66	193	0	583	69	97	18	182	0	913	6	923	0	263	0	615	0	0	1987
	H_veloincr	936	2	717	0	803	0	3724	515	304	226	1102	43	193	0	570	82	91	24	182	0	916	3	923	0	263	0	615	0	0	1987
	H_resflush	905	33	717	0	803	0	4239	0	304	226	1118	27	193	0	652	0	115	0	182	0	918	1	923	0	263	0	615	0	1987	0
	M_sed	938	0	82	635	795	8	2943	1296	304	226	1100	45	193	0	652	0	106	9	182	0	915	4	923	0	263	0	615	0	1980	7
	M_channel	938	0	717	0	803	0	3820	419	281	249	1116	29	193	0	546	106	115	0	182	0	919	0	923	0	263	0	615	0	1645	342
	M_crossec	938	0	557	160	803	0	3719	520	281	249	1112	33	193	0	551	101	115	0	182	0	919	0	923	0	263	0	615	0	1684	303
	M_instrhab	938	0	534	183	803	0	3716	523	281	249	1110	35	193	0	547	105	115	0	182	0	919	0	923	0	263	0	615	0	1684	303
	M_ripveg	938	0	717	0	803	0	3715	524	310	220	1113	32	193	0	547	105	115	0	182	0	918	1	923	0	263	0	615	0	0	1987
Pressure	M_embank	938	0	552	165	803	0	3719	520	310	220	1115	30	193	0	548	104	115	0	182	0	918	1	923	0	263	0	615	0	1393	594
	M_floodpro	938	0	717	0	803	0	3699	540	310	220	1117	28	193	0	544	108	115	0	182	0	916	3	923	0	263	0	615	0	1382	605
	M_remfloodpl	938	0	316	401	751	52	2775	1464	310	220	1080	65	193	0	652	0	6	109	182	0	284	635	923	0	254	9	615	0	0	1987
	W_toxic	938	0	640	77	780	23	4187	52	310	220	875	270	193	0	384	268	115	0	182	0	914	5	923	0	263	0	615	0	1980	7
	W_acid	938	0	717	0	803	0	4238	1	310	220	1057	88	193	0	652	0	115	0	182	0	915	4	923	0	263	0	615	0	1980	7
	W_index	938	0	713	4	803	0	3573	666	310	220	681	464	193	0	492	160	115	0	182	0	912	7	923	0	263	0	557	58	1972	15
	W_eutroph	938	0	713	4	803	0	4058	181	310	220	1094	51	193	0	628	24	115	0	182	0	917	2	923	0	263	0	615	0	1980	7
	W_opoll	932	6	244	473	803	0	3899	340	310	220	901	244	193	0	578	74	115	0	182	0	917	2	923	0	263	0	615	0	1980	7
	W_osilt	938	0	244	473	803	0	4198	41	310	220	1086	59	193	0	522	130	114	1	182	0	917	2	923	0	263	0	615	0	0	1987
	O_nav	938	0	717	0	803	0	4239	0	310	220	1127	18	193	0	652	0	115	0	182	0	919	0	923	0	263	0	615	0	0	1987
	O_collconn	938	0	0	717	803	0	3401	838	310	220	1108	37	193	0	652	0	115	0	182	0	918	1	923	0	263	0	615	0	1987	0
	C_B_c_do	938	0	717	0	799	4	4239	0	310	220	1105	40	193	0	652	0	115	0	182	0	919	0	923	0	263	0	615	0	1987	0
	C_B_s_up	938	0	694	23	799	4	4235	4	310	220	1076	69	193	0	652	0	115	0	182	0	919	0	923	0	263	0	615	0	1987	0
	C_B_s_do	938	0	694	23	799	4	4224	15	310	220	1076	69	193	0	652	0	115	0	182	0	919	0	923	0	263	0	615	0	1987	0
	C_Bn_sup	938	0	373	344	799	4	4235	4	310	220	344	801	193	0	652	0	18	97	182	0	919	0	923	0	263	0	615	0	1987	0
	C_Bn_sdo	938	0	341	376	799	4	4215	24	310	220	312	833	193	0	652	0	12	103	182	0	919	0	923	0	263	0	615	0	1987	0

2. River types based on environmental gradients

In order to get an overview about the spatial distribution of EFI+ sites and to find sites with similar environmental conditions, we created a new variable (river type). This variable characterizes different types of environmental gradients across Europe, by only considering 6 of the most important environmental variables that are available for all sites. Therefore, the following analyses have been conducted:

2.1 Factor analysis

First, variables which are available for the majority of sites and which are important descriptors of environmental gradients were selected: Altitude, actual river slope (after checking correlation with variable valley slope), wetted width, size of catchment, latitude and longitude. We conducted a factor analysis using these input variables. The factor analysis identified 3 factors and explained 79,5% of the variance. The factors can be summed as follows:

FAC1_catch (contains variables size_of_catchment and wetted_width),

FAC2_geo (contains variables latitude and longitude) and

FAC3_alt (contains variables altitude and actual river slope).

Table 2.1.1: Total variance of 79.464, explained for 6 environmental input variables

Comp	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,139	35,651	35,651	2,139	35,651	35,651	1,704	28,404	28,404
2	1,444	24,064	59,715	1,444	24,064	59,715	1,586	26,439	54,843
3	1,185	19,748	79,464	1,185	19,748	79,464	1,477	24,620	79,464
4	,620	10,334	89,798						
5	,317	5,291	95,089						
6	,295	4,911	100,000						

Table 2.1.2:Rotated component matrix, describing he composition of factors FAC1_catch, Fac2_geo and FAC3_alt.

Component	1 FAC1_catch	2 Fac2_geo	3 Fac3_alt
Size_of_catchment	,919		
Wetted_width	,917		
Longitude		,898	
Latitude		,858	
Actual_river_slope			,837
Altitude			,829

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 4 iterations.

2.2 Cluster analysis

The next step was to use the 3 factors as input data for cluster analysis, done by using the K-means method processed by SPSS© 15 (this software was used for all further analyses). The output of the cluster analysis was 6 clusters representing different types of environmental gradients.

Tab 2.2.1: 6 final cluster centers, describing the composition out of the 3 different factors, representing types of environmental gradients.

	Cluster					
	1	2	3	4	5	6
FAC1_catch	-,10910	3,33127	22,48782	-,01040	6,82281	-,21245
FAC2_geo	-,99866	,43188	-,88885	,30408	,08138	,62153
FAC3_alt	,10376	-,40471	,79682	3,34955	-,26195	-,31118

Due to the fact that cluster number 3 contained only 3 sites (located in the Netherlands at River Meuse with a very large catchment), this cluster was combined with cluster 6. Additionally, the cluster-order was re-structured in an ordinal way, and it now represents 5 river types as a function of environmental gradients.

Tab 2.2.2: : Types of env. variablilty after reduction from 6 to 5 clusters.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	325	3,2	3,2	3,2
	2	3493	34,2	34,6	37,8
	3	5854	57,3	58,0	95,8
	4	297	2,9	2,9	98,7
	5	132	1,3	1,3	100,0
	Total	10101	99,0	100,0	
Missing	System	107	1,0		
Total		10208	100,0	100	

2.3 Discriminant analysis

Next, discriminant analysis was used to test the output of the cluster analysis. The explained variance of discriminant analysis was around 83,3%.

Tab 2.3.1: Comparison of cluster 5 (counts after cluster analyses and optimizing) and discrim 5 (counts after discriminant analysis).

		cluster5					Total
		1	2	3	4	5	
discrim5	1	320	861	376	0	0	1557
	2	4	2493	2	0	0	2499
	3	1	78	5342	0	0	5421
	4	0	61	134	261	3	459
	5	0	0	0	36	129	165
Total		325	3493	5854	297	132	10101

2.4 Description of the five river types

Tab. 2.4.1: Distribution of sites per country and type of environmental variablility.

		River type					Total
		1	2	3	4	5	
Country	AT	95	0	767	11	0	873
	CH	311	7	182	0	0	500
	DE	0	0	452	211	118	781
	ES	492	1587	4	15	0	2098
	FI	0	0	227	51	0	278
	FR	270	91	342	53	0	756
	HU	1	0	157	25	5	188
	IT	208	14	215	2	0	439
	LT	0	0	88	5	0	93
	NL	0	0	147	7	23	177
	PL	25	0	843	19	18	905
	PT	77	800	2	44	0	923
	RO	26	0	229	8	0	263
	SE	4	0	593	8	0	605
	UK	48	0	1173	0	1	1222
Total		1557	2499	5421	459	165	10101

Typology :

In following, the five river types based on environmental gradients are described shortely, Fig. 2.4.1 A – F on the next page describe the distribution of types as a function ot the 6 selected environmental variables.

Type	Description	Country	Name
1	Small sized rivers, high altitude and slope, mountainous areas of Europe	Switzerland, Spain, France,Austria, Italy	**Alpine type**
2	Small sized rivers with medium altitude, low slope, south-western part of Europe	Portugal, Spain, France, Italy	**Iberian type**
3	Medium sized rivers, medium altitude, low slope, central Europe	UK, Romania, Poland, Lithuania, Hungary, Scandinavia, Austria	**Central type**
4	Large sized rivers, large catchment but low altitude, low slope, central/western Europe	Netherlands, Germany, Poland	**Large type**
5	Extreme large sized rivers, very large catchment size, extreme low slope and altitude, central/western Europe	Netherlands, Germany	**Very large type**

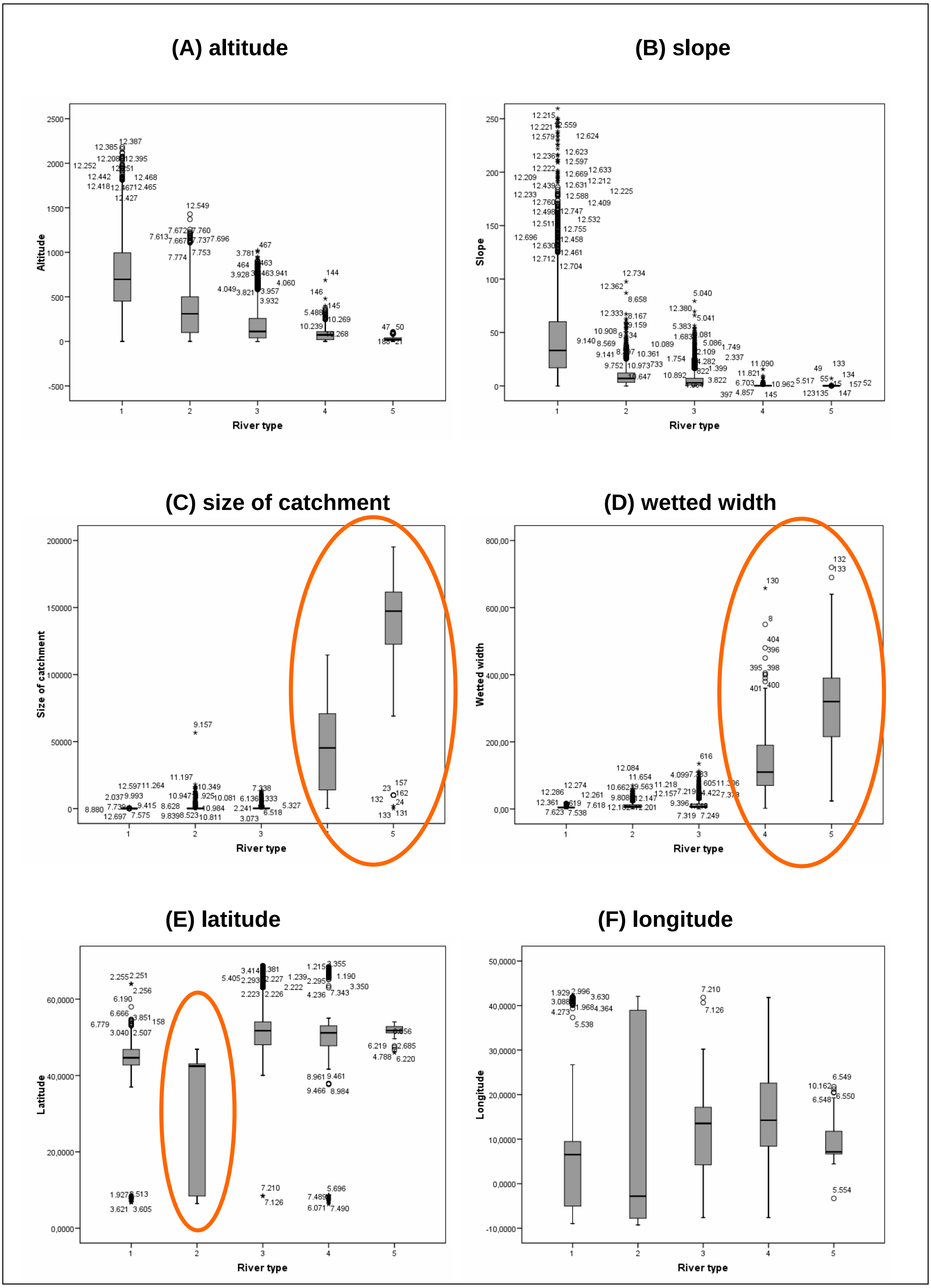


Figure 2.4.1: Boxplots, showing the distribution of environmental gradients for the 5 river types, described by (A) altitude [m], (B) slope [%o], (C) size of catchment [km²], (D) wetted width [m]], (E) latitude, (F) longitude.

Spatial distribution of the five river types

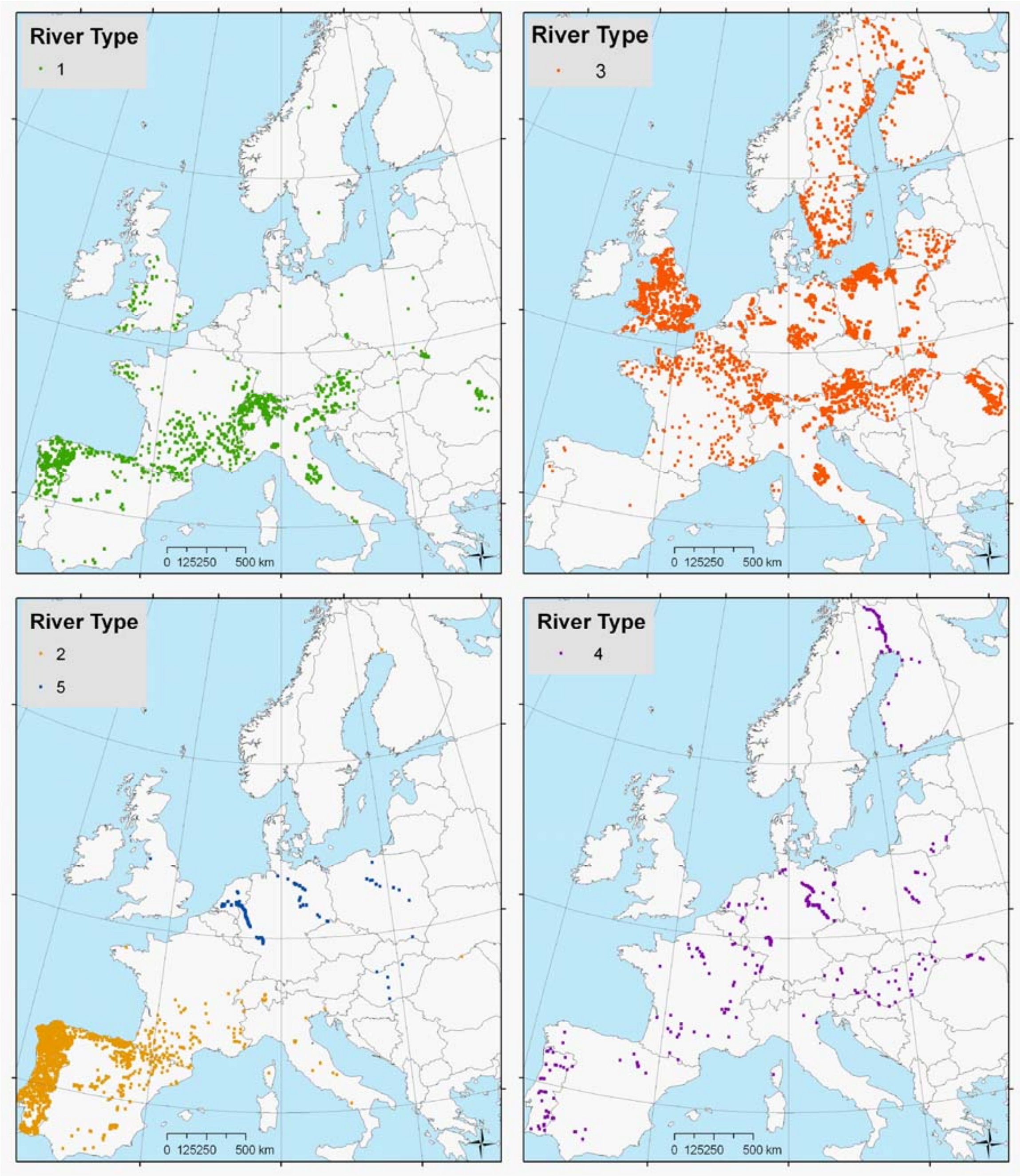


Fig. 2.4.2: Spatial distribution of the five river types across Europe (Copyright Clemens Trautwein).

Finally, the 5 river types were used for all further analyses of the pressure dataset. In general, the focus for index development will be set on river types 1, 2 and 3 (including 9477 sites), sites located in types 4 and 5 (including 624 sites) will be more important and interesting for assessment of large floodplain rivers. But nevertheless, the occurrence of types 4 & 5 must be kept in mind for all further steps of EFI+ development, because they represent a very diverse dataset that should not be combined up with data for smaller rivers.

Note that the 5 river types calculated here represent a combination of environmental attributes and were identified only for testing output of analyses with pressure indices and GIS variables. This new variables is not intended to represent a comprehensive classification of European river types.

3. Variable selection for pressure index development

To point out redundant variables and to reduce the number of variables used for index calculation, we used PCA (principle component analysis). This procedure quantifies categorical variables while reducing the dimensionality of the data (SPSS Categories® 11.0 manual). Categorical principal components analysis is also known by the acronym CATPCA for categorical principal components analysis.
The goal of principal component analysis is to reduce an original set of variables into a smaller set of uncorrelated components that represent most of the information found in the original variables. The technique is most useful when a large number of variables prohibit effective interpretation of the relationships between objects (subjects and units).
First, PCA was done for each pressure group with the total dataset and afterwards with a split dataset (stratified by all five river types and by country). To test and specify the outputs of PCA, correlation analysis and cross tables (for variables that potentially can be summed up) have been done in the next step.

3.1 Hydrological pressures

Status: 6 variables have been collected

Variable	Index development	Comment
Impoundment (H_imp)	yes	
Hydropeaking (H_hydrop)	yes	
Water abstraction (H_waterabstr)	yes	
Hydrograph modification (H_hydromod)	yes	
Reservoir flushing (H_resflush)	yes	
Temperature impact (H_tempimp)	no	Completeness 86,6%, but variable is completely missing in one country

(A) **(B)**

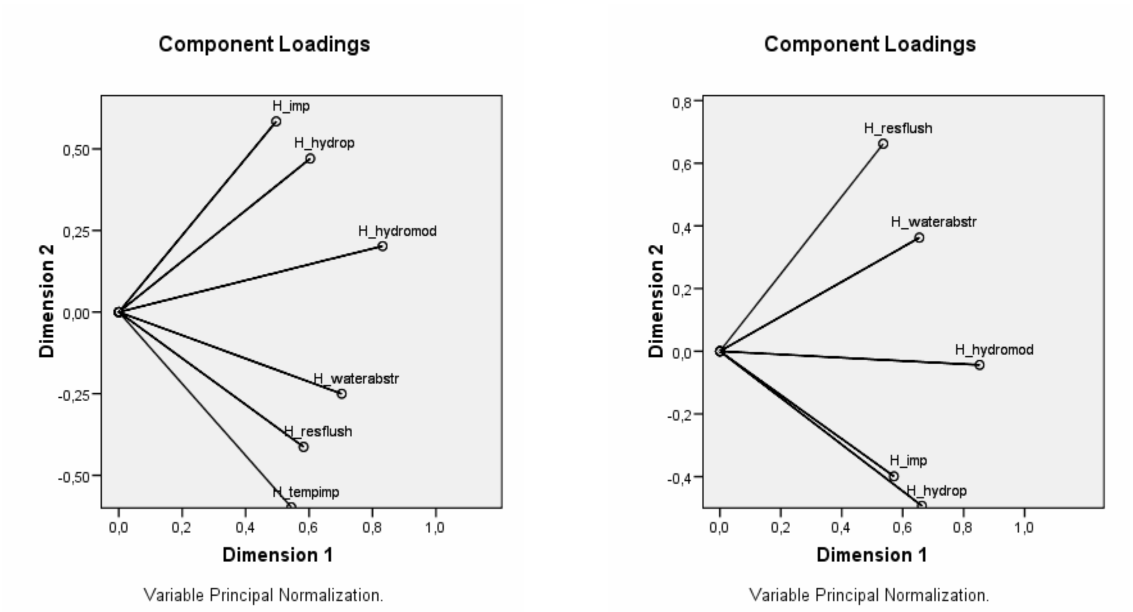


Figure 3.1.1: PCA with all 6 variables in group "hydromorphological pressures" (A) and the 5 variables selected for pressure index calculation (B).

PCA has been done for the whole group "hydrological pressures" (see figure 3.1.1), and for the five river types (described in section 2 of this report). The output suggested that all hydrological variables act differently (also in the 5 river types) and should be integrated into the pressure type specific index and into the overall index separately. The correlation table below (Tab. 3.1.1) strengthens this hypothesis.

Tab. 3.1.1: Spearman's rank correlation for all variables in group "hydrological pressures", showing that none of them are highly correlated.

Correlations

			H_imp	H_hydrop	H_waterabstr	H_hydromod	H_tempimp	H_resflush
Spearman's rho	H_imp	Correlation Coefficient	1,000	,281**	,178**	,362**	,020	,161**
		Sig. (2-tailed)	.	,000	,000	,000	,063	,000
		N	10208	10208	10208	10208	8910	10208
	H_hydrop	Correlation Coefficient	,281**	1,000	,233**	,514**	,135**	,097**
		Sig. (2-tailed)	,000	.	,000	,000	,000	,000
		N	10208	10208	10208	10208	8910	10208
	H_waterabstr	Correlation Coefficient	,178**	,233**	1,000	,445**	,409**	,292**
		Sig. (2-tailed)	,000	,000	.	,000	,000	,000
		N	10208	10208	10208	10208	8910	10208
	H_hydromod	Correlation Coefficient	,362**	,514**	,445**	1,000	,288**	,358**
		Sig. (2-tailed)	,000	,000	,000	.	,000	,000
		N	10208	10208	10208	10208	8910	10208
	H_tempimp	Correlation Coefficient	,020	,135**	,409**	,288**	1,000	,357**
		Sig. (2-tailed)	,063	,000	,000	,000	.	,000
		N	8910	8910	8910	8910	8910	8910
	H_resflush	Correlation Coefficient	,161**	,097**	,292**	,358**	,357**	1,000
		Sig. (2-tailed)	,000	,000	,000	,000	,000	.
		N	10208	10208	10208	10208	8910	10208

**. Correlation is significant at the 0.01 level (2-tailed).

Aside from these analyses, the relationship between hydrological and morphological variables was screened. The output demonstrated that the variable H_veloincr (velocity increase due to channelisation and floodprotection), was primarily part of the group hydrological pressures, but it belongs with the instream channel variables in the morphology section. Therefore, the variable H_veloincr was recoded into M_H_veloincr and henceforth is considered as a morphological variable.

Variables finally used for Hydrological index calculation:

H_imp, H_hydrop, H_waterabstr, H_hydromod, H_resflush, H_tempimp.

3.2 Morphological pressures

Status: 9 variables have been collected:

Variable	Index developme	Comment
Channelisation (M_channel)	yes	
Cross section (M_crosssec)	yes	
Instream habitat (M_instrhab)	yes	
Embankment (M_embankm)	yes	
Floodprotection (M_floodpr)	yes	
Riparian vegetation (M_ripveg)	no	Completeness 84,6%, but variable is total missing for one country
Velocity increase (M_H_veloincr)	no	Completeness 84,2%, but one country is totally missing, redundancy
Sedimentation (M_sediment)	no	Completeness 89,9%, but one country is totally missing, redundancy
Remaining floodplain (M_remfloodpl)	no	Inconsistent information, low completeness (75,5%)

PCA for the morphology section was completed for the whole group and again, separately, for the five river types. As described before, the variable velocity increase was added to the morphological pressure group and included into the following analyses.

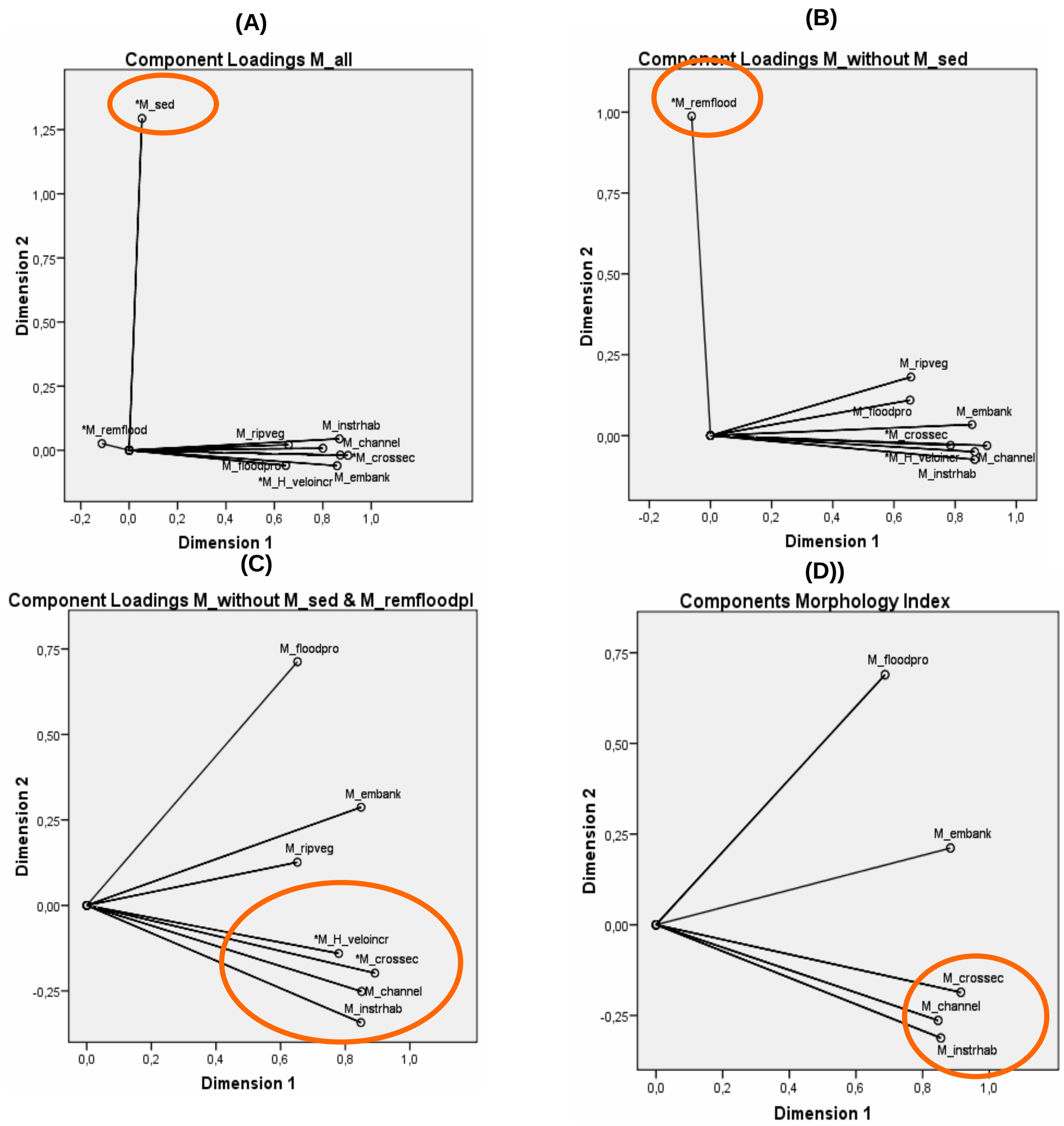


Fig. 3.2.1, A-D: PCA for all variables out of “morphological pressures”, including M_H_veloincr. Fig. A and Fig. B show that M_sed and M_remfloodpl correspond to different components than all others; Fig. C shows that some variables might be redundant; Fig. D shows the loadings of variables finally considered for Morphological pressure index calculation.

The first output of PCA was that variables M_channel, M_instrhab, M_H_veloincr and M_crossec act in a similar manner and correspond to the same component (axis). To point out clear correlations between these variables, M_sed and M_remfloodpl were stepwise excluded from PCA, because they create different components (axes).

Tab. 3.2.1 confirms the hypothesis that M_channel, M_instrhab, M_H_veloincr and M_crossec are highly correlated, but variable velocity increase had to be excluded from further analyses, because of incomplete information (as well as riparian vegetation and sedimentation).

Tab. 3.2.1: Spearman's rank correlation for the four variables M_channel, M_instrhab, M_H_veloincr and M_crossec, showing that they are highly correlated.

Correlations

		M_sed	M_channel	M_crossec	M_instrhab	M_ripveg	M_embank	M_floodpro	M_remfloodpl
M_sed	Corr. Coeff.	1,000	,043**	,046**	,103**	,125**	,093**	,085**	-,032*
	Sig. (2-tailed)	.	,000	,000	,000	,000	,000	,000	,006
	N	9705	9705	9705	9705	8473	9705	9705	7492
M_channel	Corr. Coeff.	,043**	1,000	,738**	,659**	,477**	,630**	,438**	-,068*
	Sig. (2-tailed)	,000	.	,000	,000	,000	,000	,000	,000
	N	9705	10208	10208	10208	8976	10208	10208	7749
M_crossec	Corr. Coeff.	,046**	,738**	1,000	,756**	,521**	,693**	,492**	-,064*
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000
	N	9705	10208	10208	10208	8976	10208	10208	7749
M_instrhab	Corr. Coeff.	,103**	,659**	,756**	1,000	,534**	,590**	,402**	-,107*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000
	N	9705	10208	10208	10208	8976	10208	10208	7749
M_ripveg	Corr. Coeff.	,125**	,477**	,521**	,534**	1,000	,477**	,400**	-,005
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,663
	N	8473	8976	8976	8976	8976	8976	8976	7748
M_embank	Corr. Coeff.	,093**	,630**	,693**	,590**	,477**	1,000	,656**	-,084*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000
	N	9705	10208	10208	10208	8976	10208	10208	7749
M_floodpro	Corr. Coeff.	,085**	,438**	,492**	,402**	,400**	,656**	1,000	-,038*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,001
	N	9705	10208	10208	10208	8976	10208	10208	7749
M_remfloodpl	Corr. Coeff.	-,032**	-,068**	-,064**	-,107**	-,005	-,084**	-,038**	1,000
	Sig. (2-tailed)	,006	,000	,000	,000	,663	,000	,001	.
	N	7492	7749	7749	7749	7748	7749	7749	7749

According to Tab. 3.2.1, M_channel, M_instrhab and M_crossec show a high correlation and can be pooled into one variable (M_morph_instream, see next page). This decision is supported by the following crosstables. The number of cases in crosstables (given in %) for min/min (no/no) pressures and max/max pressures (extreme/high) always should be at least 70% of total values.

Tab. 3.2.2: Crossable, showing the relationship between variable instream habitat and alteration of cross-section. Relation between No/No and Technical crosssec/Strong in both cases more than 80%.

			M_instrhab			Total
			1 No	3 Weak	5 Strong	
M_crossec	1 No	Count	5589	1097	136	6822
		%	81,9%	16,1%	2,0%	100,0%
	3 Intermediate	Count	286	1196	334	1816
		%	15,7%	65,9%	18,4%	100,0%
	5 Technical crosssec	Count	46	207	1317	1570
		%	2,9%	13,2%	83,9%	100,0%
Total		Count	5921	2500	1787	10208
		%	58,0%	24,5%	17,5%	100,0%

Tab. 3.2.3: Cross table, showing the relationship between variable instream habitat and channelisation. Relation between No/No and Straightened/Yes represents in both cases more than 70%.

			M_instrhab			Total
			1No	3 Weak	5 Strong	
M_channel	1 No	Count	5502	1370	291	7163
		%	76,8%	19,1%	4,1%	100,0%
	3 Intermediate	Count	363	746	385	1494
		%	24,3%	49,9%	25,8%	100,0%
	5 Straightened	Count	56	384	1111	1551
		%	3,6%	24,8%	71,6%	100,0%
Total		Count	5921	2500	1787	10208
		%	58,0%	24,5%	17,5%	100,0%

Variables describing floodplains

Tab. 3.2.4: Crosstable, showing the relation between M_remfloodpl and Floodplain_site.

			Floodplain						Total
			NoData	Large	Medium	No	Small	Some	
Floodplain_site		Count	128	19	3	65	1	4	220
		%	58,2%	8,6%	1,4%	29,5%	,5%	1,8%	100,0%
	No	Count	1779	5	11	5282	16	9	7102
		%	25,0%	,1%	,2%	74,4%	,2%	,1%	100,0%
	Yes	Count	552	326	353	1086	516	53	2886
		%	19,1%	11,3%	12,2%	37,6%	17,9%	1,8%	100,0%
Total		Count	2459	350	367	6433	533	66	10208
		%	24,1%	3,4%	3,6%	63,0%	5,2%	,6%	100,0%

Variable M_remfloodpl always corresponds to a different component than in PCA, and in contrast to all other variables of group morphological pressures it gives opposite information (high = best case). This variable was compared with two other available "floodplain" variables, M_floodpr (floodprotection) and M_floodplain_site (presence of former floodplain, table site) to determine whether their outputs could be pooled. Tab. 3.2.4 shows that especially M_remfloodpl and M_floodplain_site do not fit together, e.g in table sites it's indicated that 68,8 percent of sites had no former floodplain, but in remaining floodplain, the information is given that there are no more floodplains remaining. Due to these facts and the "low" completeness of variable M_remfloodpl (available only for 75% of the pressure analysis dataset, see Tab. 1.4.2), we propose to take this variable out of the index development. But nevertheless, information on former floodplain or remaining floodplain percentage will be useful for further analyses, especially for large floodplain rivers assessment.

Variables finally used for the Morphology index calculation:

A new variable (M_floodpl_total) for further floodplain analyses will be created as follows: In cases where the variable flooplain_site is equal to "no", this information will be adopted, in all other cases information out of variable M_remfloodplain will be taken.

Redundant variables M_channel, M_instrhab and M_crossec will be summed up in variable M_morph_instr (by taking the arithmetic mean value of these 3 variables).

Final selection: Morph_instream, M_embankm, M_ripveg, M_floodpr, M_sediment,

3.3 Connectivity pressures

Status: 7 variables have been collected

Variable	Index development	Comment
Barriers on catchment down (C_catch_down)	no	90% impacted
Barriers on segment down (C_B_s_do)	yes	
Barriers on segment up (C_B_s_up)	yes	
Number of barriers on segment down (C_Bn_s_do)	no	completeness (81,3%) & redundancy
Number of barriers on segment up (C_Bn_s_up)	no	completeness (82,7) & redundancy
Distance to next barrier on segment down (C_Bd_s_do)	no	completeness (31%)
Distance to next barrier on segment up (C_Bd_seg_up)	no	completeness (34%)

Due to the low completeness of variables describing distance to barrier on the segment level (only 31% as well as 34% complete for variables C_Bd_s_up and C_Bd_s_do, see completeness of variables per country in Tab. 1.4.5) and the fact that different units (meters as well as kilometers) have been indicated by some partners in the updated version of EFI+ database, these variables are not used for index calculation.

But later on, they may be useful for assessing connectivity and diadromous species.

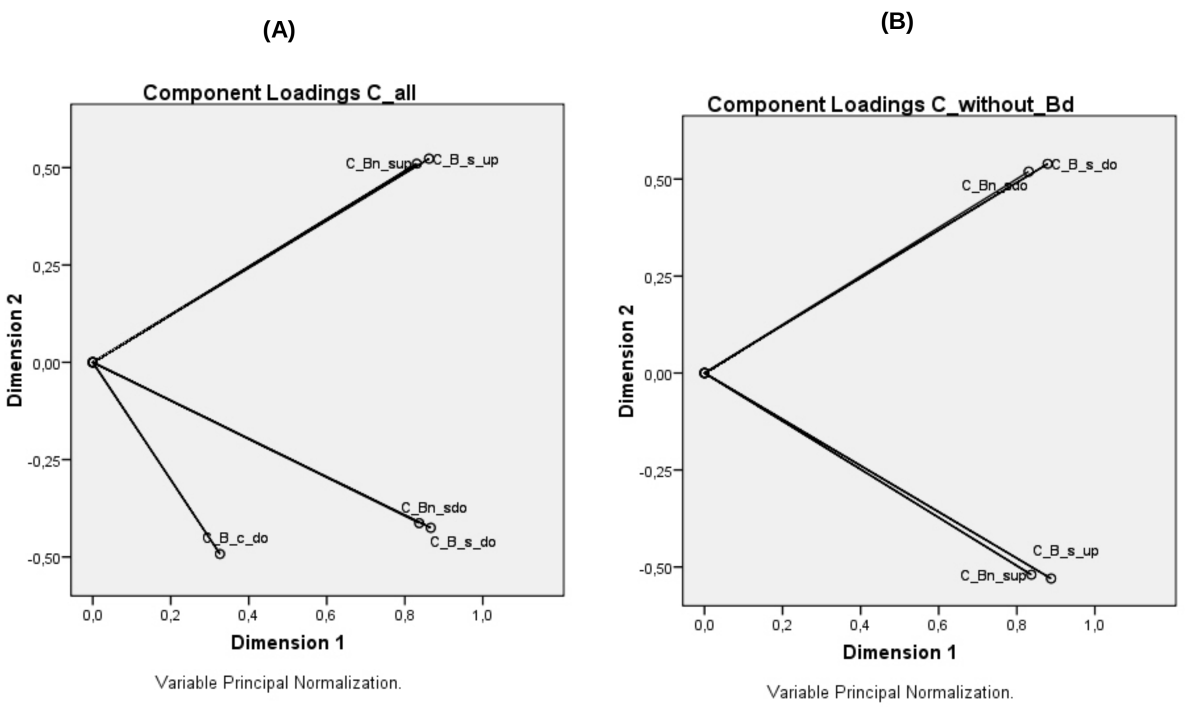


Fig. 3.3.1, A & B: PCA, showing that variables on segment up and on segment down correspond to the same component (axis), variable on catchment not.

Variables finally used for Connectivity index calculation:

Our analyses indicated that three variables, C_bn_seg_up, C_bn_seg_down and C_catch_down, should be removed from the pressure analysis. C_catch_down, in particular, corresponds to a different component than the other connectivity variables in PCA and cannot be combined with the other connectivity variables. Logically, the large spatial extent of this variable (presence or absence of a barrier anywhere downstream of the survey site) also limits its usefulness. The data on total number of barriers does show clear trends using both PCA and correlation analysis and therefore will be used for further continuum analyses, but not for the continuum-pressure index.

Tab. 3.3.1: Spearman's rank correlation between connectivity variables. Variables on segment up and on segment down are highly correlated.

			C_B_c_do	C_B_s_up	C_B_s_do	C_Bn_sup	C_Bn_sdo
Spearman's rho	C_B_c_do	Correlation Coefficient	1,000	,133**	,192**	,176**	,233**
		Sig. (2-tailed)	.	,000	,000	,000	,000
		N	10203	10203	10203	9446	9379
	C_B_s_up	Correlation Coefficient	,133**	1,000	,432**	,979**	,453**
		Sig. (2-tailed)	,000	.	,000	,000	,000
		N	10203	10208	10208	9448	9381
	C_B_s_do	Correlation Coefficient	,192**	,432**	1,000	,452**	,973**
		Sig. (2-tailed)	,000	,000	.	,000	,000
		N	10203	10208	10208	9448	9381
	C_Bn_sup	Correlation Coefficient	,176**	,979**	,452**	1,000	,522**
		Sig. (2-tailed)	,000	,000	,000	.	,000
		N	9446	9448	9448	9448	9225
	C_Bn_sdo	Correlation Coefficient	,233**	,453**	,973**	,522**	1,000
		Sig. (2-tailed)	,000	,000	,000	,000	.
		N	9379	9381	9381	9225	9381

**. Correlation is significant at the 0.01 level (2-tailed).

Both, PCA and correlation analyses showed that both variables describing the segment down and both variables describing the segment up are redundant (see also Tab. 3.3.1).

Variables qualified for pressure index calculation: C_seg_down, C_seg_up.

3.4 Waterquality pressures

Status: 6 variables have been collected

Variable	Index development	Comment
Acidification (W_acid)	yes	
Water quality index (W_index)	no	inconsistencies
Toxic substances (W_toxic)	no	completeness is 88,8%, but totally missing for one country
Organic siltation (W_osilt)	no	completeness is 83,4%, but totally missing for one country
Organic pollution (W_opoll)	yes	
Eutrophication (W_eutroph)	yes	

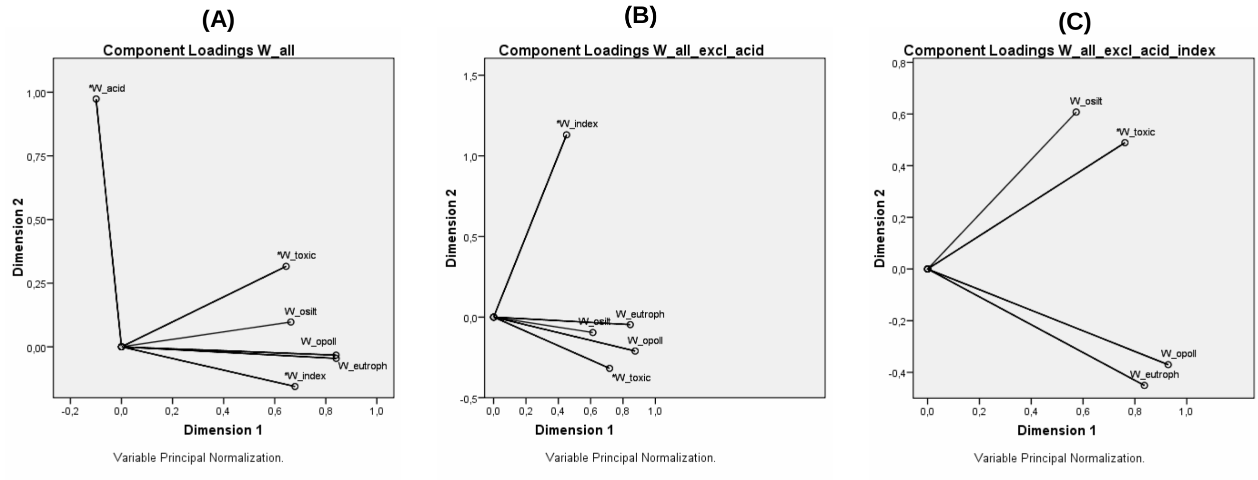


Fig. 3.4.1, A-C: PCA for all variables out of "waterquality pressures". Fig. A to C show that W_index and W_acid correspond to different components (axes), and that W_osilt and W_toxic as well as W_eutrph and W_poll seem to be redundant.

PCA-analyses for the total dataset as well as for the split dataset (into 5 river types) showed, that variables W_acid and W_index correspond to different components, but variables W_toxic and W_osilt as well as W_eutroph and W_opoll might be redundant. Tab 3.4.1 tests this hypothesis with Spearman's rank correlation.

Tab. 3.4.1: Spearman's rank correlations between all variables of group "waterquality pressures". Although some variables show correlations, all waterquality variables are integrated separately into the pressure index, because they describe totally different effects.

Correlations

			W_toxic	W_acid	W_index	W_eutroph	W_opoll	W_osilt
Spearman's rho	W_toxic	Correlation Coefficient	1,000	-,014	,202**	,374**	,446**	,523**
		Sig. (2-tailed)	.	,180	,000	,000	,000	,000
		N	9451	9451	9227	9451	9451	8203
	W_acid	Correlation Coefficient	-,014	1,000	-,008	-,025*	-,109**	-,071**
		Sig. (2-tailed)	,180	.	,421	,011	,000	,000
		N	9451	10208	9903	10208	10208	8545
	W_index	Correlation Coefficient	,202**	-,008	1,000	,554**	,422**	,397**
		Sig. (2-tailed)	,000	,421	.	,000	,000	,000
		N	9227	9903	9903	9903	9903	8264
	W_eutroph	Correlation Coefficient	,374**	-,025*	,554**	1,000	,596**	,461**
		Sig. (2-tailed)	,000	,011	,000	.	,000	,000
		N	9451	10208	9903	10208	10208	8545
	W_opoll	Correlation Coefficient	,446**	-,109**	,422**	,596**	1,000	,398**
		Sig. (2-tailed)	,000	,000	,000	,000	.	,000
		N	9451	10208	9903	10208	10208	8545
	W_osilt	Correlation Coefficient	,523**	-,071**	,397**	,461**	,398**	1,000
		Sig. (2-tailed)	,000	,000	,000	,000	,000	.
		N	8203	8545	8264	8545	8545	8545

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Special case water quality index:

As already noticed during data collection, variable W_index (Waterquality index in 5 classes) is a problematic variable, because different indices were used by different partner-countries. Either one or both chemical and biotic data were included in this variable. The following cross tables compare this variable with some other important water quality-variables, to estimate whether the water quality index is correlated with other water quality problems.

Tab. 3.4.2: Relation between waterquality index and toxic substances. The table shows that W_index is not able to detect problems with toxics substances correctly, 20% of waterquality class 1 and class 2 are intermediate or highly impacted by toxic substances.

			W_toxic			Total
			1 No	3 Intermediate	5 High	1
W_index	1	Count	1868	283	133	2284
		%	81,8%	12,4%	5,8%	100,0%
	2	Count	2432	550	214	3196
		%	76,1%	17,2%	6,7%	100,0%
	3	Count	1224	745	403	2372
		%	51,6%	31,4%	17,0%	100,0%
	4	Count	644	193	59	896
		%	71,9%	21,5%	6,6%	100,0%
	5	Count	268	136	75	479
		%	55,9%	28,4%	15,7%	100,0%
Total		Count	6436	1907	884	9227
		%	69,8%	20,7%	9,6%	100,0%

Tab. 3.4.3: Relationship between waterquality index and a combined variable of eutrophication and organic pollution (always taking into account the worst value of both). The table shows that W_index is not able to detect eutroph_opll problems correctly, 9% of waterquality class 1 and 54,2% of waterquality class 2 are impacted by eutrophication and/or organic pollution (at least 3 = weak).

			W_eutoph_opoll_max				Total
			1	3	4	5	1
W_index	1	Count	2161	118	24	72	2375
		%	91,0%	5,0%	1,0%	3,0%	100,0%
	2	Count	1547	1423	140	263	3373
		%	45,9%	42,2%	4,2%	7,8%	100,0%
	3	Count	606	1321	437	233	2597
		%	23,3%	50,9%	16,8%	9,0%	100,0%
	4	Count	159	366	232	249	1006
		%	15,8%	36,4%	23,1%	24,8%	100,0%
	5	Count	63	166	138	185	552
		%	11,4%	30,1%	25,0%	33,5%	100,0%
Total		Count	4536	3394	971	1002	9903
		%	45,8%	34,3%	9,8%	10,1%	100,0%

Conclusion for waterquality pressures:

The water quality index will not be used for pressure index calculations because it isn't able to detect other important water quality pressures. After testing the outputs of PCA with correlation analysis (Fig. 3.4.1), it was decided to integrate all remaining water quality-variables separately into the pressures index. Although some variables show correlations, they describe totally different effects. Due to the fact that W_toxic and W_osilt are not complete for all sites selected for pressure analysis, these variables were rejected for pressure analysis.

Variables selected for Water quality index calculation: W_acid, W_opoll, W_eutroph.

3.5 Variables retained for index development - Frequencies

Tab. 3.5.1: The table describes which variables will be summed up or kept separately for index calculation. Additional, Frequencies of occurrence n % are indicated for each variable and category (fields below each category)..

Cat.	Code_grouped	Variable_code	No pressure	Pressure intensity			
Hydrology	H_impound	H_impound	no (1)	weak (3)	strong (5)		
		Frequency in %	88,5	6,2	5,3		
	H_hydrop	H_hydrop	no (1)	partial (3)	yes (3)		
		Frequency in %	92,3	7,7			
	H_waterabstr	H_waterabstr	no (1)	weak to medium (3)	strong (5)		
		Frequency in %	71	17,2	11,8		
	H_resflush	H_resflush	no (1)	yes (3)			
		Frequency in %	96,1	3,9			
	H_hydromod	H_hydromod	no (1)	yes (3)			
		Frequency in %	81,8	18,2			
Morphology	Morph_instr	M_channel	no (1)	intermediate (3)	straightened (5)		
		Frequency in %	71,1	14,8	14,2		
		M_crosssec	no (1)	intermediate (3)	technical crossec (5)		
		Frequency in %	67,9	18	14,1		
		M_instrhab	no (1)	intermediate (3)	high (5)		
		Frequency in %	57,8	25,5	16,6		
	M_embankm	M_embankm	no (1)	slight (2)	intermediate (3)	high (5)	
		Frequency in %	65,8	13,1	10,9	10,1	
	M_floodpr	M_floodpr	no (1)	yes (3)			
		Frequency in %	77,0	23,0			
Connectivity	C_seg_up	C_bsegup	no (1)	partial (3)	yes (3)		
		Frequency in %	76,8	23,2			
	C_seg_down	C_bsegdown	no (1)	partial (4)	yes (4)		
		Frequency in %	78,0	22,0			
	W_acid	W_acid	no (1)	yes (3)			
		Frequency in %	97,5	2,5			
	W_eutroph	W_eutroph	no (1)	low (3)	intermediate (4)	extreme (5)	
		Frequency in %	59,0	22,4	16,9	1,7	
	W_orgpoll	W_orgpoll	no (1)	weak (3)	strong (5)		
		Frequency in %	62,8	28,6	8,6		

4. Pressure index calculation

After retaining variables for index calculation, the remaining pressures were combined to pressure type indices (e.g. hydrology, morphology, continuity) and further aggregated into a global pressure index. In the following, we describe 4 different approaches for the EFI+ pressure index:

4.1 Different index approaches

4.1.1 MEAN Index - “FAME” approach

As already analysed in the FAME project, one approach to estimate the total impacts of pressures is to compute the mean value of all variables selected for index calculation.

Principle: Arithmetic mean values

Waterquality	Morphology	Continuum	Hydrology
1	1	3	5

Index (exemplary): 2,5 *(1+1+3+5 = 10; 10/4 = 2,5)*

Characteristics: Only the average is considered, single pressures are underrepresented. Multiple impacted sites are well represented.

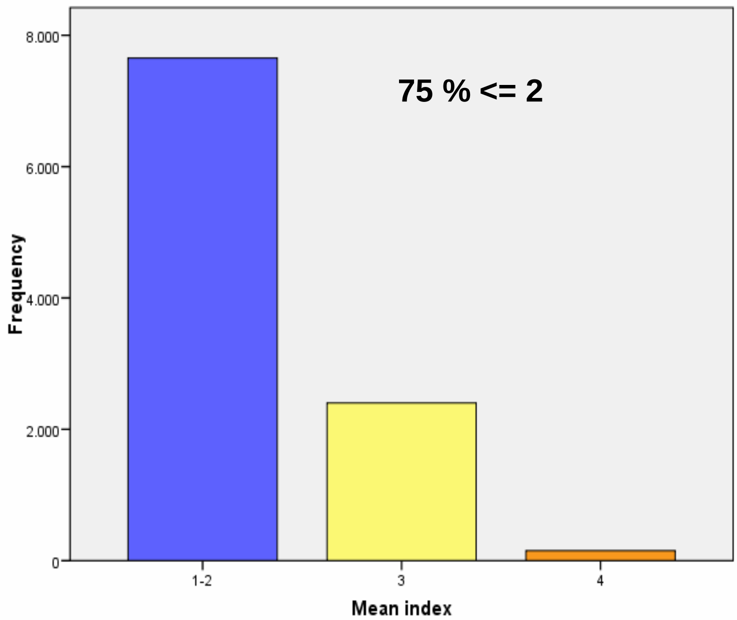


Fig.4.1.1: Distribution of Mean Index in 3 classes: 1-2, 3, 4; more than 75% of sites are classified better than 3.

4.1.2 MAX Index - “Worst case” approach

Another possible index classified each site by taking the worst impact class of all variables chosen for the index.

Principle: One out – all out

Waterquality	Morphology	Continuum	Hydrology
1	1	3	5

Index (exemplary): 5

Characteristics: The worst variable always rates the total index value.

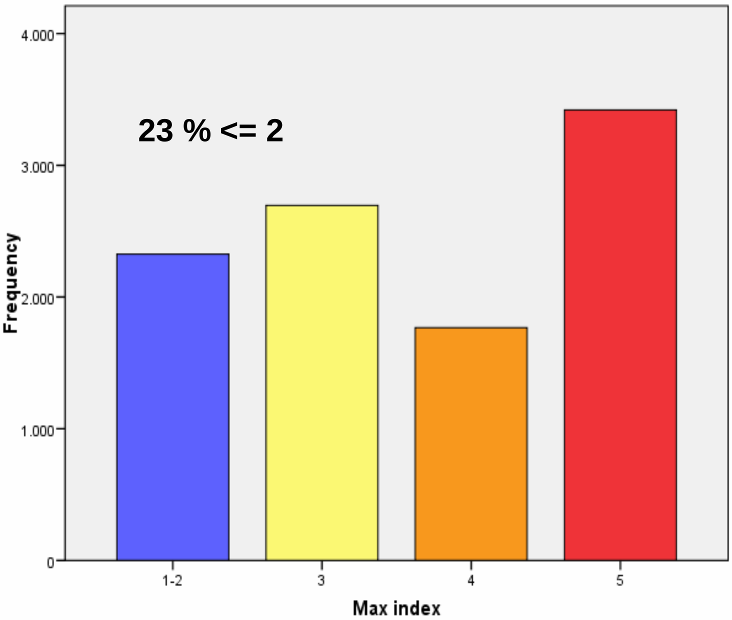


Fig. 4.1.2: Distribution of Max Index in 3 classes: 1-2, 3, 4, 5; more than 23% of sites are classified better than 3, but around 35% are classified as 5.

4.1.3 Degraded Index - “Average worse case” approach

Principle: Mean value of all values worse than 2

Waterquality	Morphology	Continuum	Hydrology
1	1	3	5

Index (exemplary): 4 *(3+5 = 8; 8/2 = 4)*

Characteristics: At least 2 values for classification needed.

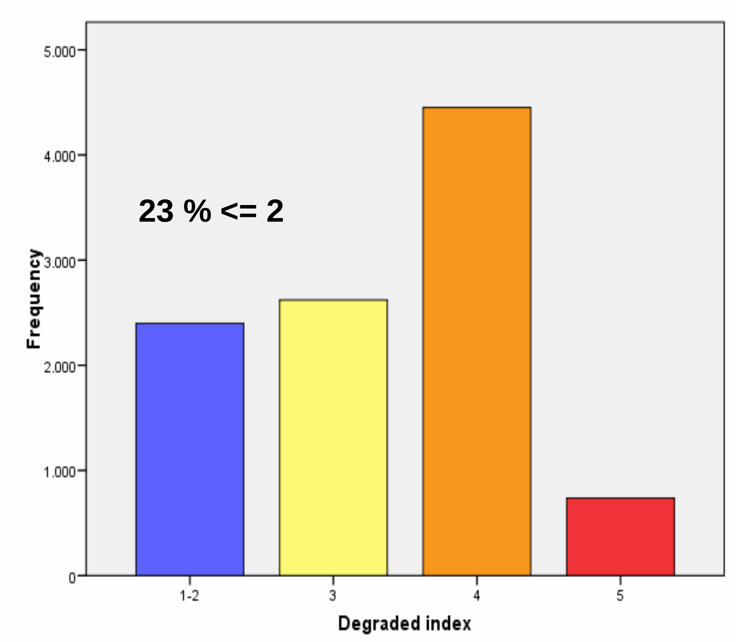


Fig. 4.1.3: Distribution of Degraded Index in 4 classes: 1-2, 3, 4, 5; more than 23% of sites are classified better than 3, but around 40% are classified as 4.

As Fig. 4.1.4 shows, the 3 different approaches show different behaviour especially in classes 1-2. E.g. the MEAN index classifies around 80% of sites in class 1-2, DEG Index and MAX index only 23%. The different approaches have been calculated for all sites, Fig. 4.1.5 A & B show the distribution and frequency of different index approaches per country.

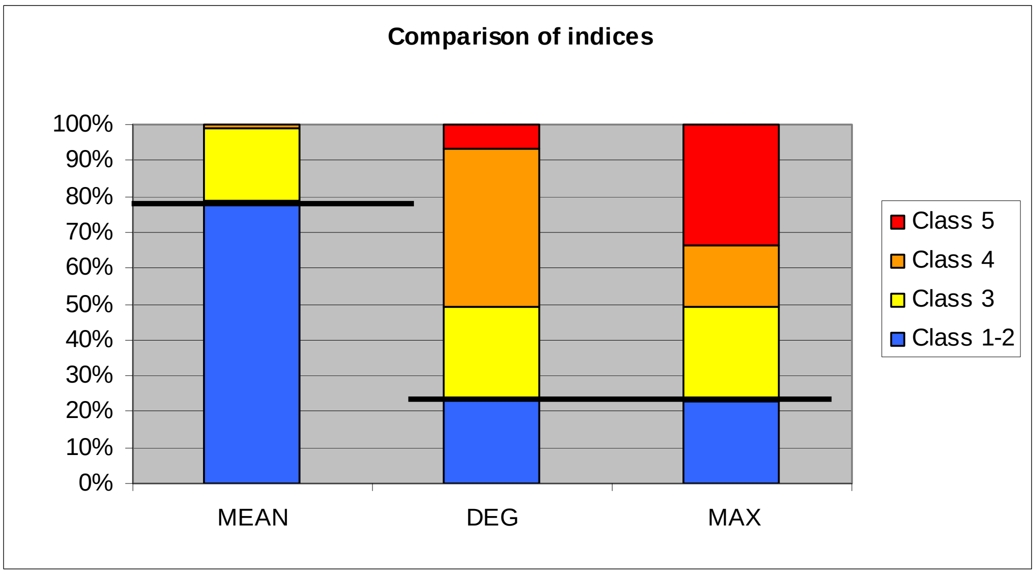
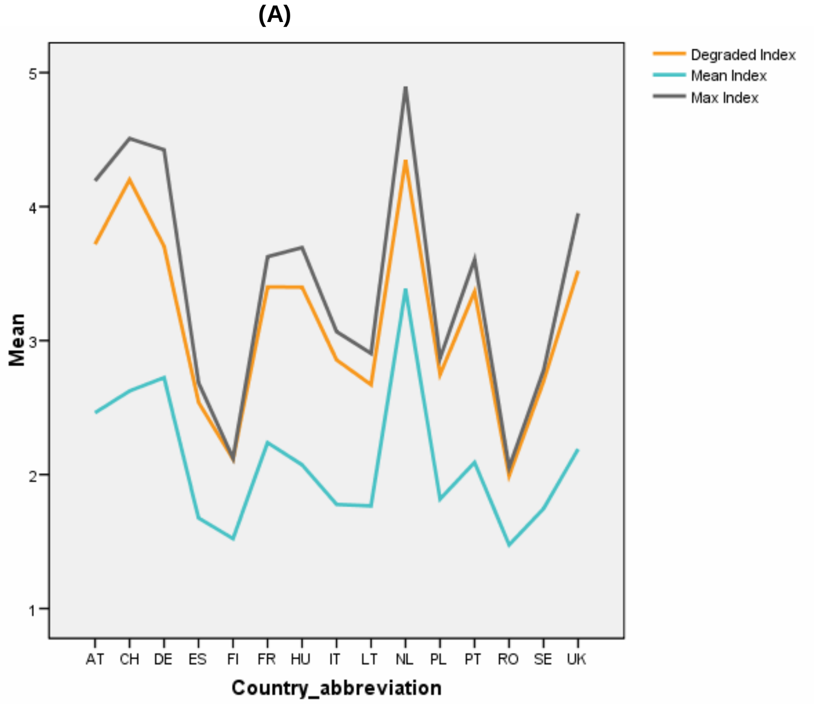


Fig. 4.1.4: Comparison of 3 different index approaches (MEAN, DEG, MAX) in classes.

(A)



(B)

Country		Mean index	Degraded index	Max index
AT	Mean	2,46	3,72	4,19
	N/Sites	873	873	873
CH	Mean	2,62	4,20	4,51
	N/Sites	500	500	500
DE	Mean	2,72	3,70	4,42
	N/Sites	781	781	781
ES	Mean	1,68	2,54	2,69
	N/Sites	2098	2098	2098
FI	Mean	1,52	2,12	2,12
	N/Sites	278	278	278
FR	Mean	2,24	3,40	3,63
	N/Sites	785	785	785
HU	Mean	2,07	3,40	3,69
	N/Sites	193	193	193
IT	Mean	1,78	2,86	3,07
	N/Sites	498	498	498
LT	Mean	1,77	2,67	2,90
	N/Sites	94	94	94
NL	Mean	3,38	4,35	4,90
	N/Sites	182	182	182
PL	Mean	1,82	2,75	2,87
	N/Sites	907	907	907
PT	Mean	2,09	3,36	3,60
	N/Sites	923	923	923
RO	Mean	1,48	1,99	2,05
	N/Sites	263	263	263
SE	Mean	1,75	2,70	2,78
	N/Sites	605	605	605
UK	Mean	2,19	3,52	3,95
	N/Sites	1228	1228	1228
Total	Mean	2,06	3,14	3,41
	N/Sites	10208	10208	10208

Fig. 4.1.5.: Distribution of the three different indices per country (A), mean values of indices per country (B).

4.2 Number of pressure types affected – multiple pressures

To get an overview about the intensity and number of impacts per site, and to point out countries or river types that are impacted by multiple pressures, new variables indicating the number of pressures categorised higher than 2 (at least weak or strong impact) have been calculated for each pressure group and overall. In addition to the different indices approaches, the number of pressures per site was calculated as follows:

Nr_pressure_groups_affected: Counting number of values worse than 2 in all pressure groups (highest possible value = 4; means each pressure group is affected).

The following figures describe the output of this variable, to show the intensity and number of pressures per site, and to point out countries or river types that are impacted by multiple pressures.

(A) (B)

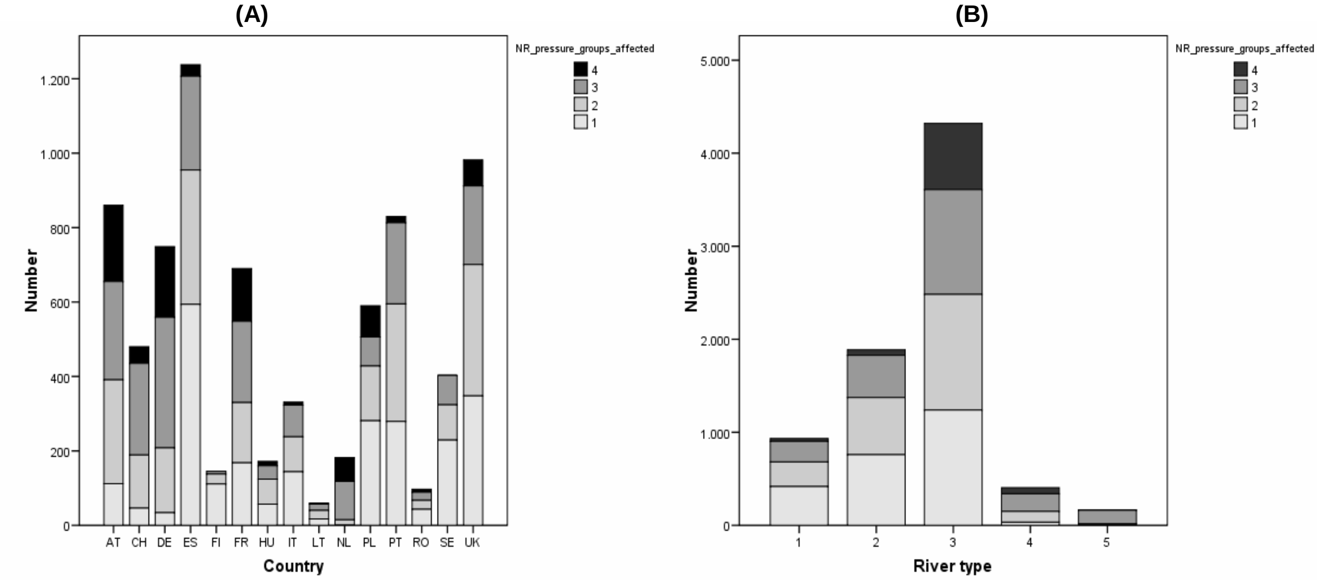


Fig. 4.2.1.: Number of affected pressure groups per country (A) and river type (B).

4.3 Hydromorphological pressures

Finally, first analyses focusing on hydromorphological pressures have been done, to get in idea about the number of sites impacted by hydromorphological pressures only. Therefore, the index HMC has been calculated (Degraded index of Groupindex hydrology, morphology and connectivity) and has been confronted with W_groupindex (average worse index of variables indicated on page 25).

Tab. 4.3.1: Number of sites with HMC pressures (hydromorphologcial pressures, pink), WQ pressures (waterquality pressures, blue), overall pressues (HMC + WQ, yellow) and no pressures (grey).

			W_index				Total
			1+2	3	4	5	
HMC_index	1+2	Count	2399	880	191	95	3555
		%	67,48	24,75	5,37	2,67	100
	3	Count	942	799	321	215	2277
		%	41,37	35,09	14,10	9,44	100
	4	Count	952	1684	527	495	3658
		%	26,03	46,04	14,41	13,53	100
	5	Count	206	271	81	150	708
		%	29,10	38,28	11,44	21,19	100
Total		Count	4499	3634	1120	955	10208
		%	44,07	35,60	10,97	9,36	100

According to the Tab. 4.3.1, 2100 sites (20,6 %) of the total dataset (10208 sites) are affected by hydromorphological pressures without interaction of waterquality pressures and therefore are qualified for analyses focusing on hydromorphological pressures only. On the other hand, 5709 sites (55,9 %) of data are impacted by hydromorphological and waterquality pressures. The remaining 2399 sites (23,5 %) represent sites in Class 1 and 2.

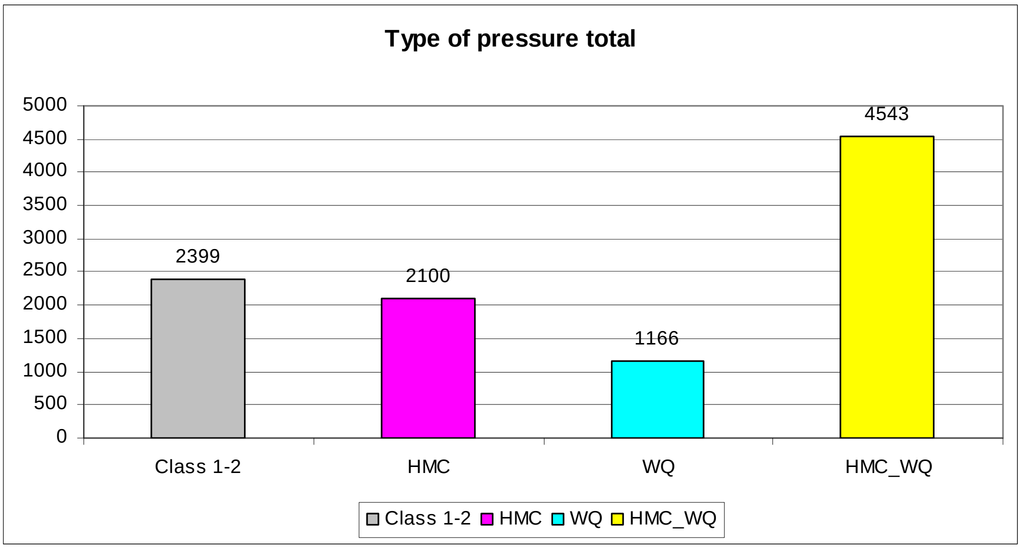


Fig. 4.3.1: Sites affected by HMC, WQ, HMC and WQ and sites with low pressures.

4.4 Weighted degraded Index - "Combined" approach

Principle: Multiplication of "Degraded Index" and "Number of affected pressure groups"

After the EFI+ consortium meeting in Paris in April 2008, the "Degraded Index" which has finally been proposed as global EFI+ index has been reconsidered.

As described in section 4.3, the "Degraded Index" does not detect single pressures and therefore can't meet one of the challenges of EFI+, to detect multiple impacts. Therefore, another approach has been considered, integrating "Degraded Index" and "Number of pressure groups affected" by a multiplication of these factors.

Waterquality	Morphology	Continuum	Hydrology
1	1	3	5

Degraded Index (exemplary): 4

Number of pressure groups affected (value higher than 2): 2

Weighted degraded Index: *(4 * 2) = 8*

Characteristics: Considering sites affected by single and multiple pressures, multiple affected sites are weighted and therefore are classified as more severe than sites with single pressures (e.g. 2x3 =6, 1x5 = 5).

Count

		Affected pressure groups				
		1,00	2,00	3,00	4,00	Total
Weighted_DEG_ Index_class	3	1944	0	0	0	1944
	4	411	0	0	0	411
	5	108	0	0	0	108
	6	0	1162	0	0	1162
	7	0	745	0	0	745
	8	0	303	0	0	303
	9	0	63	557	0	620
	10	0	7	878	0	885
	11	0	0	458	0	458
	12	0	0	203	108	311
	13	0	0	86	257	343
	14	0	0	4	258	262
	15	0	0	1	146	147
	16	0	0	0	92	92
	17	0	0	0	14	14
	18	0	0	0	4	4
Total		2463	2280	2187	879	7809

Tab. 4.4.1: Number of pressures (affected groups) compared with weighted degraded index.

As Tab. 4.4.1 shows, single pressures as well as combined pressures are detected by the "Weighted degraded approach". Fig. 4.2.1 below shows the index distribution for EFI+ sites. It's important to notice that 1944 sites are impacted by one "Weak" pressure only, and it has to be discussed if these sites will be considered for analyses dealing with "Strong pressures".

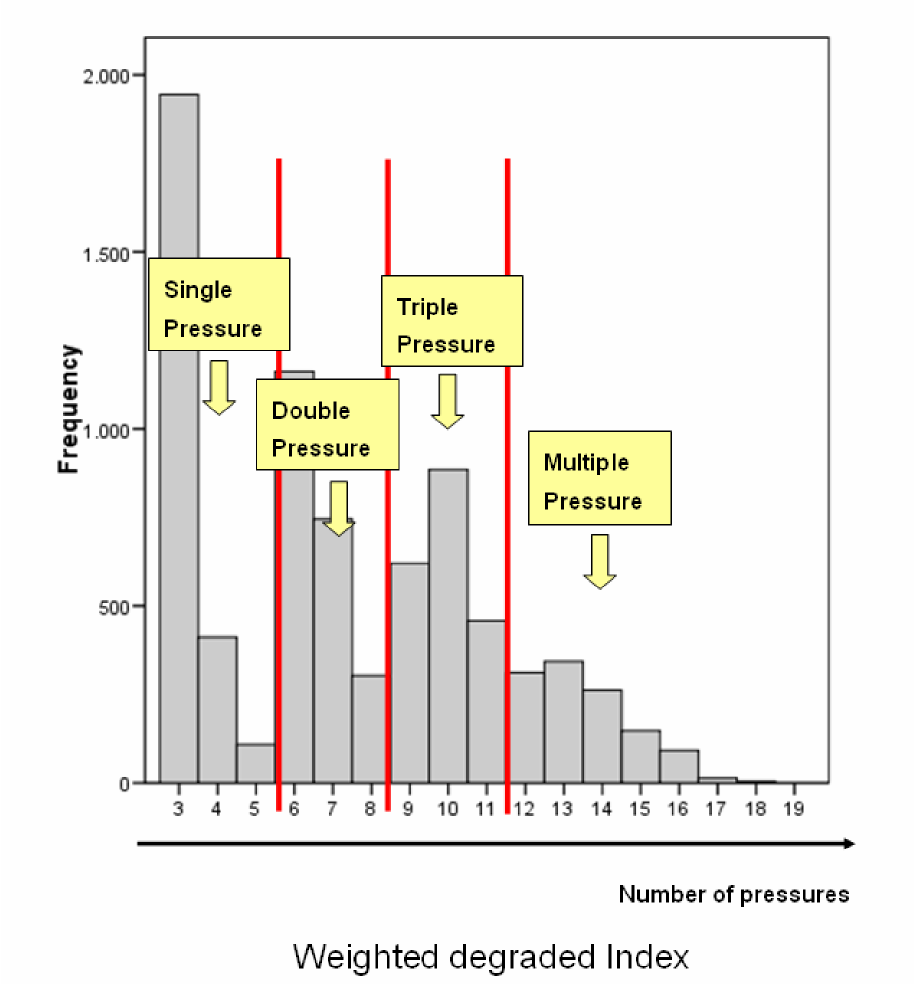


Fig. 4.4.1: Distribution of Weighted DEG Index for EFI+ sites.

5. Conclusion

After testing different approaches for pressure index calculation, a lot of discussions within the EFI+ consortium about pro's and con's of these variants followed. Especially at the EFI+ working group meeting in Paris in April 2008, all partners were confronted with the procedure and the problems of the "degraded approach", which can't separate single pressures from multiple ones. Due to the fact that the new approach, the "Weighted degraded index" detects both, single pressures and multiple pressures, it's now suggested to use it as the final approach for global pressure index calculation instead of the "Degraded index.