Standardisation of river classifications:

Framework method for calibrating different biological survey results against ecological quality classifications to be developed for the Water Framework Directive

Contract No: EVK1-CT 2001-00089
8th deliverable, due 31/12/04, entitled:

Inter-calibration and harmonisation of "invertebrate methods"

(Paper version)

Compiled by Leonard Sandin[5], Nikolai Friberg[10], Mike T Furse[1], Ralph Clarke[1], Sören Larsen[10]

Partners no 1 (CEH, United Kingdom), 5 (SLU, Sweden) and 10 (NERI, Denmark)

A project under the 5th Framework Programme
Energy, Environment and Sustainable Development
Key Action 1: Sustainable Management and Quality of Water

Contents

1 INTRODUCTION: Objectives and format of the deliverable

1.1 OBJECTIVES (TAKEN FROM THE DESCRIPTION OF WORK)

The aims of this report are:

- To compare and intercalibrate the results of proven assessment methods that already comply with CEN standards, e.g. RIVPACS in Great Britain, IBGN in France, Saprobic Systems in Austria and Germany, EBEOSWA in The Netherlands and IBE in Italy, since these methods are not likely to be discarded by their respective Member States.

- To compare and inter-calibrate the results of different invertebrate methods, particularly in terms of errors, precision, relation to reference conditions.

- To investigate the cost-effectiveness of different approaches to the collection and processing of macroinvertebrate samples involving varying levels of resource intensity.

The general objectives of STAR include:

- Inter-calibration of the assessments derived in different river types, eco-regions and Member States
- Quantification of the errors associated with the field and laboratory protocols used to obtain the data

The collection and processing of data required to meet these objectives form Workpackage 7 (WP7) and Workpackage 8 (WP8) of the project.

In combination the source workpackages provide:

- high quality, consistently identified macro-invertebrate data from sites sampled and sorted in eleven Member States and three NAS using standardised field and laboratory protocols for the most commonly used field techniques (two in each country)

- databases of reliable quality on ecological river quality from other national and international projects

- the metrics used by EU Member States and NAS for assessing the Ecological Status of their lotic waterbodies.

1.2 FORMAT OF THE DELIVERABLE

The deliverable is comprised of one component:

- This written document

1.3 PARTICIPATING PARTNERS

18 of 22 partners participated in one or several of WP 7, 8, and 11. These were as follows:

- Centre for Ecology and Hydrology ……. United Kingdom
- University of Duisburg-Essen……..... Germany
- BOKU – University of Agricultural Sciences…….. Austria
- Alterra Green World research……. The Netherlands
- Swedish University of Agricultural Sciences……Sweden
- Masaryk University Brno……. Czech Republic
- Hellenic Centre for Marine Research, IIW……... Greece
- Consiglio Nazionale delle Ricerche……..Italy
- University of Évora……. Portugal
- National Environmental Research Institute…….. Denmark
- Vuzkumny ustav vodohospodarsky T.G. Masaryka…… Czech Republic
- Autonomous Province of Bolzano……Italy
- University of Metz……France
- Research Institute Senckenberg……Germany
- University of Łódź ……Poland
- University of Latvia…….. Latvia
- Slovak Academy of Science……. Slovak Republic
- Comenius University of Bratislava…….. Slovak Republic

2 METHODS

2.1 MACROINVERTEBRATES ASSESSMENT METHODS

2.1.2 Sampling methods of macroinvertebrates

At all sites sampled in the STAR project the STAR-AQEM method (section 2.2) was used. In total, 15 partners in 13 countries took STAR-AQEM samples (Table 2.1). For each STAR-AQEM site, samples were also taken using national macroinvertebrate assessment protocols to enable a comparison between methods. Eight different national assessment protocols were compared to the STAR-AQEM method. As some countries do not have a national protocol they used a slightly modified version of the UK RIVPACS sample protocol (section 2.3) hereafter denoted the RIVPACS method. These countries were Austria, Germany and Greece (Table 2.1).

Table 2.1 Methods applied by the STAR partners

Institution (STAR Partner No.)	Methods applied
CEH (1) UK	STAR-AQEM and RIVPACS
Univ. D-Essen (2) Germany	STAR-AQEM and RIVPACS
BOKU (3) Austria	STAR-AQEM and RIVPACS
SLU (5) Sweden	STAR-AQEM and Swedish national method
Masaryk University (6) Czech Rep.	STAR-AQEM and PERLA
HCMR-IIW (7) Greece	STAR-AQEM and RIVPACS
CNR-IRSA (8) Italy	STAR-AQEM and IBE
University of Evora (9) Portugal	STAR-AQEM and PMP
NERI (10) Denmark	STAR-AQEM and DSFI
Water Inst. Brno (12) Czech Rep.	STAR-AQEM and PERLA
University of Metz (14) France	STAR-AQEM and IBGN
Senckenberg (15) Germany	STAR-AQEM and RIVPACS
University of Lodz (17) Poland	STAR-AQEM and Polish national method
University of Latvia (20) Latvia	STAR-AQEM and LVS 240:1999
Comenius University (22) Slovak Rep.	STAR-AQEM and PERLA

In the next sections each of the methods employed during the course of the STAR project are briefly described. For more detailed information see under “protocols” on the STAR project homepage (www.eu-star.at).

2.1.2.1 The STAR-AQEM sampling method

Aim

The methods are based on the Rapid Bioassessment Protocols (Barbour et al. 1999), the procedures of the Environment Agency (Environment Agency 1999a), the Austrian Guidelines “Saprobiology” (Moog et al. 1999) and ISO 7828. These guidelines have been tested and adapted by the AQEM partners (see www.aqem.de) to provide standardised procedures for collecting and analysing macroinvertebrate samples within the STAR-AQEM stream assessment procedures and be further adapted to meet the requirements of the STAR project.

The description does not aim at, nor is it able of competing with or replacing the references cited above. The information given here focuses on the application of the STAR-AQEM approach to guarantee a standardised procedure.

Sampling

A sample consists of 20 “sampling units” taken from all microhabitat types at the sampling site with a share of at least 5% coverage. A “sampling unit” is a stationary sampling accomplished by positioning the net and disturbing the substrate for a distance that equals the square of the frame width upstream of the handnet (0.25 x 0.25 m; 500 µm mesh size) or using a Surber sampler (0.25 m x 0.25 m = 0.0625 m^2 and a 500 µm mesh size). The 20 “sampling units” must be distributed according to the share of microhabitats. For example, if the habitat in the sampling reach is 50% psammal (sand), then 10 “sampling units” must be taken there. The categories of microhabitat composition are to be taken from the site protocol (parameters 23 and 24). This procedure results in sampling of approximately 1.25 m^2 stream bottom area.

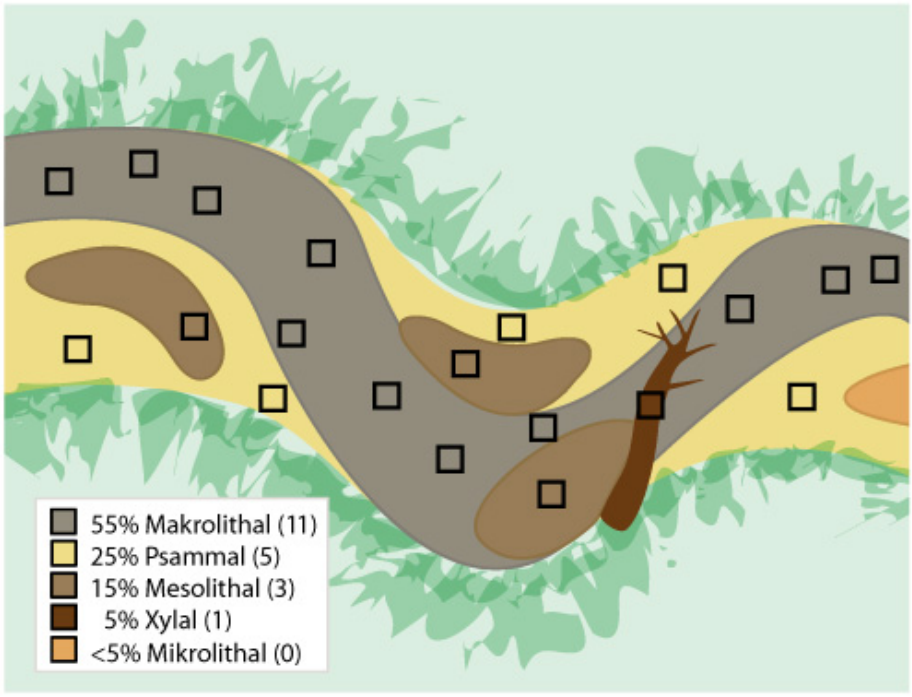


Figure 2.1: Example of sampling unit positions in a theoretical sampling siteaccording to the ‘multi habitat sampling’ method applied in STAR-AQEM.

Sampling starts at the downstream end of the reach and proceeds upstream. When sampling the ‘sampling units’ the hand-net is used either as a kick net, or for ‘jabbing’, ‘dipping’ or ‘sweeping’. When kick-sampling, hold the net vertically with the frame at right angles to the current, downstream from your feet, and disturb the stream bed vigorously by kicking or rotating the heel of your boot to dislodge the substratum and the fauna within a depth of at least 10-15 cm. Disturb the substrate in the 0.25 x 0.25 m area upstream of the net. Hold the net close enough for the invertebrates to flow into the net with the current, but far enough away for most of the sand and gravel to drop before entering the net. The surface of soft sediments and fine or organic microhabitats should be sampled by pushing the hand-net gently through the uppermost 2-5 cm of the substratum. In shallow waters with a strong current an open Surber sampler can be used instead of a hand-net. To sample with an open Surber sampler in slow-flowing areas the sediment within the Surber frame can be disturbed using the hands, in the normal fashion, and then a current created by pushing water through the net with the hands to

trap the animals. It is possible to use different devices for different microhabitats, as long as the same area is sampled. For a detailed description see the AQEM Manual.

Sample procedure in the field

Large wood and stones can be removed after being rinsed and inspected for clinging or sessile organisms. Any organisms found have to be placed into the sample container. Large and fragile organisms (e.g. Ephemeroptera) or species that cannot be preserved (e.g. Tricladida, Oligochaeta) should be picked out of the sample in the field (a maximum of 30 representative organisms in total). These organisms should be stored in a small separate container containing only organisms but no substrate. They must be kept separately for auditing purposes. Large and rare organisms, which can easily be determined in the field (such as large mussels), should be removed from the sample and be placed back in the stream after they have been added to the sampling protocol. Transfer the sample from the net to sample container(s) and preserve with formalin (4% final concentration) or in enough 95% ethanol to cover the sample immediately after collection.

Laboratory procedure

Samples are sorted in trays without magnification. Subsampling are employed as depicted in the following flowchart:

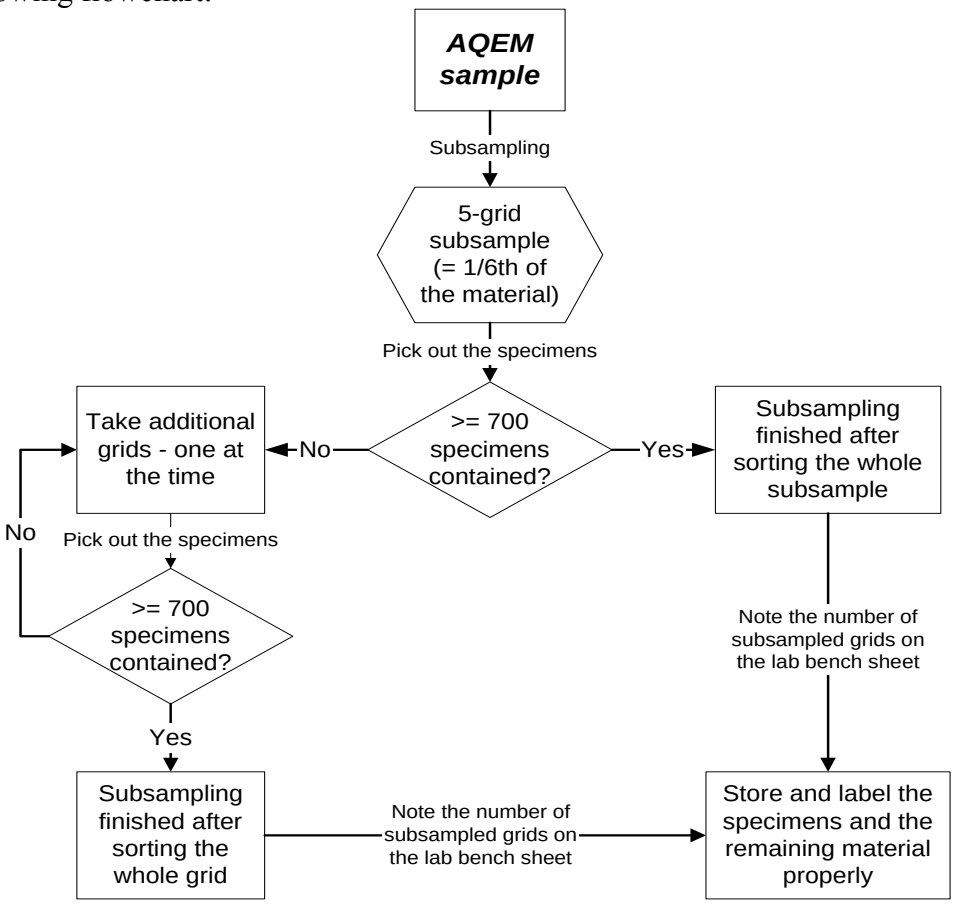


Figure 2.2 Flowchart of the subsampling procedure. The example is based on the original subsampler after (Caton 1991) with 30 grids, each 6 x 6 cm. A minimum amount of 1/6th of the material has to be subsampled, containing a minimum number of 700 specimens.

Taxonomic resolution will depend on the geographically region sampled but always aiming at identifying macroinvertebrates to the best attainable level. Oligochaeta and Chironomidae below the

subfamily level are the only taxa not identified to species in all countries. For a detailed description of the laboratory method, see the AQEM manual.

2.1.2.2 The RIVPACS sampling method

Aim

This RIVPACS sampling method employed in STAR is based on the Environment Agency's standard sampling and analysis manual (BT001) and describes the methods used by regulatory authorities in England, Wales, Scotland, Northern Ireland and Isle of Man for collecting and analysing samples of invertebrates to assess the quality of rivers. It is based on methods developed and required for RIVPACS. Other systems and methods used by the Environment Agency also rely on these methods.

Sampling

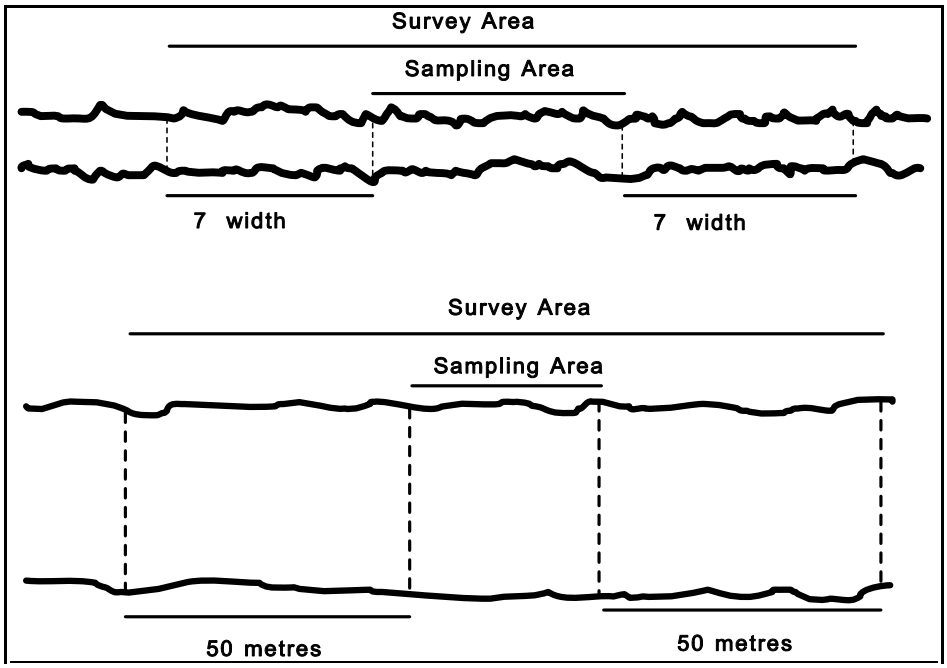


Figure 2.3 The RIVPACS survey area and the sampling area

At each site, you should define both the sampling area from which the samples are collected and a more extensive survey area. The boundaries of the sampling area must be recorded in a site manual. The boundaries of the survey area only need to be recorded if they are not obvious, for instance if the sampling area is not central in the survey area.

The sampling area covers the whole width of the stream wherever possible, but its length will depend on the width of the stream and the variability of its habitats. The sampling area must be a single area of river bed whose major habitat types can be sampled in the recommended sampling period. It must not be a collection of separate places along an extended length of river, for instance to include both riffles and pools in an attempt to increase the variety of animals captured. This would cause over-sampling and result in an apparent under-prediction by RIVPACS. The sampling area will usually be between five and twenty metres long. It will be longer in narrow streams than in wide rivers. Sample environmental data (i.e. water width, mean water depth and substratum composition) collected for RIVPACS relates to the average conditions in the whole sampling area. Parts of the sampling area that are inaccessible for sampling are still considered to be a part of it. The sampling area is therefore more than simply the precise location from which the invertebrate sample is collected.

The survey area extends either seven channel widths or 50 m either side of the sampling area, (see Figure 2.3), whichever is the shorter (this depends on the width). The survey area must be similar to the sampling area that it encompasses. This will ensure that differences between samples from the same site caused by slight differences in the area that is sampled are minimised, and will enable the sampling area to be extended for conservation assessments or to allow replicate samples to be taken. Placing the sampling area within a survey area also ensures that the sampling area is homogenous, and not an isolated area distinct from its immediate surroundings, which may be vulnerable to damage.

The sampling comprises of two elements: manual searching and pond-net sampling.

The manual search is in two parts that, together, last one minute although the time spent on each part may vary. The first part is to seek and collect animals living on the water surface, such as whirligig beetles, water crickets and pond skaters. This must be done before any other sampling, because these animals are easily disturbed and will either leave the Sampling Area or be much more difficult to find later. They are best caught with a pond-net. Most surface dwellers are very active and they should be secured in a tied bag or vial immediately after capture. Whilst searching for these animals, note the area covered by different habitats within the Sampling Area, so that you can apportion the sampling effort amongst them in the main sample.

The second part of the search is for animals from habitats that are not sampled effectively by the methods use to collect the main sample. Pick-off animals attached or clinging to the submerged stems of emergent plants, rocks, logs, or other solid objects, with forceps or a stiff paint brush. Examine rocks at several places across the river to cover the different biotopes and areas covered by different sized substrata. Always search for animals attached to floating-leafed plants. Inspect the under-surfaces of floating leaves as well as the upper surface and stems.

The whole search must last one minute. It is standardised by time alone, and not by searching a certain number of rocks or locations. This period only covers the time spent actually searching, and excludes the time spent moving around the site. A stopwatch or watch with second hand must be used to ensure that the cumulative time spent actively searching is one minute.

The pond-net sampling.

The pond-net can be undertaken in different ways depending on the nature of the Survey Area.
Different biotopes at the same site may be sampled by a combination of the methods described below.
The total sampling time must be three minutes.

If a site comprises discrete habitats, apportion the sampling effort according to their cover in the Sampling Area. If a site appears to be homogeneous in character, continuous diagonal transects will suffice for most of the sample.

Always move upstream and diagonally across the stream a number of times whilst sampling, rather than straight upstream. This will ensure that a greater number of habitats are sampled, even if they are not apparent, and therefore a higher proportion of the taxa present at the site are collected (see Woodiwiss, 1980).

The three minutes covers only the time spent actively sampling, and excludes the time spent emptying the net, or moving around the site. It is recommended that sampling is done in short bursts of 15-20 seconds. There will be 9 to 12 bursts in a three minute sample, which is worth remembering when apportioning the sampling effort to the different habitats. A stopwatch or watch with second hand must be used to ensure that the cumulative time spent actively sampling is precisely three minutes. If

two people are on-site, it may be easier for one to time the sampling with the stopwatch so that the other can concentrate on collecting the sample. The sampler should call-out to the timekeeper when to start and stop the watch, and the timekeeper can remind the sampler when each sampling burst should end. When kick sampling, hold the net vertically with the frame at right-angles to the current, downstream from your feet, and resting firmly on the river bed; disturb the stream bed vigorously by kicking and rotating the heel of your boot to dislodge the substratum and the fauna within it to a depth of about 10 cm. Lifting and disturbing the substratum with your heel and toe by rotating your foot is particularly effective. There is no need to kick-up froth. Hold the net close enough for the invertebrates to flow into the net with the current, but far enough away for most of the sand and gravel to drop before entering the net. Hold the net further away where the substratum is finer or the current swifter, to prevent it clogging. Move large stones by hand if they cannot be shifted by foot, and sample the finer sediment that collects beneath them. Where the stream bed is soft silt or clay, kick sampling is ineffective because the net will become blocked rapidly. Instead, skim the bottom edge of the net gently through the top few centimetres of the substratum, which is where most of the animals will be found. Alternatively, stir-up the surface of the sediment by foot or with the back of the net, and pass the open net through the clouded water. Rinse the silt away through the net frequently, by agitating the net in the current or at the water surface. Specified sampling approaches also exists for other habitat types such as boulders, vegetation etc.

Sample procedure in the field

The sample should be rinsed to remove silt and clay, and to discard stones, wood, and large fragments of vegetation before removing the sample from the net. Drain the sample before putting it into a collecting jar or polythene bag. Do not add water to the sample. Fill the sample containers to no more than about half-full with collected material. This will leave sufficient room for fixative or preservative, and an air space. Never cram material into a sample container, and never fill it completely: use an additional container instead. Every container must be labelled.

Laboratory procedure

All sorting and identification for RIVPACS analysis must be undertaken in the laboratory and not in the field. The whole sample must be sorted. When washing the sample the 500 μm sieve is mandatory. Everything retained on it, or on a larger aperture sieves, is considered to be part of the sample. Sorting is undertaken without magnification but using a bench lamp shining from the side and slightly forwards to minimise reflections and shadows. There is no time limit for sorting a sample. Sorting time will depend on the sample and the experience of the sorter. Samples identified to family will normally take about 2 hours to sort. Identification are normally taken to one of two levels: the family level used in calculation of BMWP and the species level used in RIVPACS species level analysis. Family level identification takes normally about 2-3 hours whereas species level identification can take up to 2 days. Aabundance of each taxon should be recorded using a logarithmic scale of abundance (Table 2.2).

Table 2.2. Abundance categories used in RIVPACS

Category	Abundance
A	1 – 9
B	10 – 99
C	100 – 999
D	1000 – 9999
E	10000+

.

2.1.2.3 The Nordic sampling method

The Nordic methods (the Swedish and the Danish sampling method) are considered together as a previous intercalibration exercise among the Nordic countries showed no significant differences in species composition from samples using the two methods (Skriver 2000).

2.1.2.3.1 The Swedish sampling method

Aim
The monitoring of benthic fauna in running waters aims at describing the status and detecting changes in the benthic fauna communities. Species composition generally reflects environmental perturbation and the method can therefore be used to assess the effects of air pollution, land use and other encroachments or measures within the catchment area. The analysis of benthic fauna in running waters is especially suited for assessing the acidification status of a site.

The monitoring type – Benthic fauna timeseries is primarily aimed at detecting temporal changes in benthic fauna and secondarily for comparing different localities or sites. The benthic fauna samples should always be taken from a well defined substratum type to minimise variation and attain the goal of the time series monitoring.

Sampling
The samples should be taken from well defined sampling sites. A sampling site is defined as the whole wetted width of the stream, along a ten-meter stretch along that is as homogeneous as possible regarding substratum composition, vegetation, water depth, and water velocity. The water depth should not exceed one meter and the water velocity should preferably be greater than 10 cm/s. Sampling should not be done is areas that dry out for parts of the year and should be placed at a distance of at least 100 m from a lake-outlet. Sampling should primarily be restricted to riffle areas and hard bottom substratum, since it is the preferred habitat for kick-sampling. Sampling should take place downstream of a 50 m long homogeneous sampling area that do not differ considerably from the sampling site regarding substratum, vegetation and water velocity, to minimise the effects of habitats not represented in the sampling site.

A minimum of five replicate samples should be taken from each sampling site with a hand-net according to the kick-sampling method as described in the European standard SS-EN 27 828 (see below). Each sample is stored and analysed separately. The samples should be spatially distributed over the entire sampling site, but sampling near the shore should be avoided to minimise the effects of different habitats. Sampling contains the whole area from 0-1 m depth and assessment of substratum composition and vegetation should be done from this whole area.

The sampling methodology and necessary equipment for sampling benthic fauna using a hand-net is described in the European standard (SS-EN 27 828). The net is held against the bottom substratum, perpendicular to the stream, and the sampler disturbs the loose substratum upstream of the opening of the net using the foot at an area as wide as the net. The net is placed close enough to the foot, for the animals to be moved into the net by the current, but far enough away, that most gravel and sand particles settles before entering the net. The net is then moved upstream and the procedure is repeated at a stretch of 1 m during 1 minute. The net is removed from the stream after the sample has been taken and the material is collected at the bottom of the net, before being transferred to a sieve or plastic container. Organic material and phytobenthos is scrubbed off stones, twigs and the like and this large material is then discarded.

The bottom substratum is disturbed and the loose material is collected by moving the net through the water column in those cases where the stream velocity is so low that the disturbed animals do not enter

the net with the help of the current. Disturbance of the substratum of movement of the net is done at a total stretch of 1 m during 1 minute and then treated as above.

Sample procedure in the field

A total of five samples should be taken at each sampling site and transferred to separate containers where they are conserved with 96% EtOH to a final concentration of ca 70%. Samples should not be sorted in the field, but sorting and identification should be done in the laboratory. The samples should be marked both within (on a piece of paper with a pencil) and outside of the container

Laboratory procedure

2.1.2.3.2 The Danish (DSFI) sampling method

Aim

The Danish Stream Fauna Index is a standardised method, which replaces the old subjective method from 1970. The DSFI was introduced as the official method for biological assessment of running waters in Denmark from 1998 (danish environmental protection agency 1998). DSFI is currently used yearly at 1051 stations in the National Monitoring Programme for the Aquatic Environment NOVA 2003 (Bøgestrand 1999). In addition, DSFI is widely used by regional water authorities.

Although DSFI has proven to be sensitive towards several stressors (e.g. hydromorphological changes, low pH etc.), it has primarily been developed to detect the impact of organic pollution. All macroinvertebrate taxa used in DSFI are indicators of organic pollution either by being tolerant or sensitive towards low oxygen levels. The sensitivity of DSFI is therefore highest with respect to organic pollution. Organic pollution from urban areas is a major problem in Lithuanian streams and rivers. This is especially evident in densely populated areas and in stream and rivers with low summer discharges.

Sampling
The sampling procedure is standardised, and includes, in principle, sampling of all microhabitats at the site. Sampling is undertaken using a standard handnet with a 25 x 25 cm opening and a tapering netbag with a mesh size of 0.5 mm (European Standard EN 27 828). Sampling is done at three transects across the stream lying about 10 m apart, four kick samples are taken at each transect 25%, 50%, 75% and 100% from one of the stream banks (Figure 2.3). If stream width is less than one meter, i.e. the width of four handnet heads, the transects should be placed diagonally in an upstream direction. Sampling is started at the downstream transect and progresses upstream. The 12 kick samples are pooled for further analysis. The kick samples are collected by placing the handnet on the stream bed, and then placing a foot on the stream bed in front of the handnet, with the toes pointing downstream. The foot is then moved backwards about 40 cm against the current, and animals and sediment are swept by the current into the net. Once the sediment has settled, the procedure is repeated at the same spot, without having moved the net. At low current velocities, however, a slightly different sampling approach has to be used. After kicking into the bottom substrate with the foot it is necessary to move the handnet actively in the upstream direction to compensate for the low current velocity.

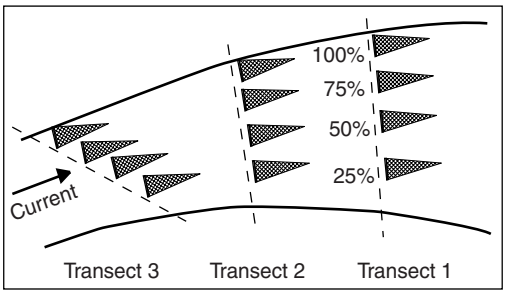


Figure 2.4. DSFI sampling methodology. Diagonally sampling (Transect 3) is undertaken if the stream width is less than 1m.

In deep rivers, the standard method of sampling may be impossible because sampling is performed along transects. In this case it is recommended to sample all available substrate types present along the bank.

Since many animals such as flatworms, leeches, snails and caddis larvae with stone cases adhere firmly to the substrate and tend to be under-represented, kick sampling is supplemented by 5 minutes of hand-picking from submerged stones and large wooden debris. The animals collected by hand-picking are kept separately from the kick sample.

Sample procedure in the field

The pooled kick sample and the hand-picked sample, which together constitute the fauna sample, are preserved separately in the field and are subsequently analysed in the laboratory. If necessary, the kick sample is sieved in the field or in the laboratory in a sieve with mesh size 0.5 mm. The pooled kick sample should not in volume exceed 0.5 – 1 l. of material.

Laboratory procedure

The macroinvertebrates are sorted and identified in the laboratory. No sub-sampling procedure is allowed. However, sorting and identification are generally not necessary when two specimens of a taxon have been identified in the kick sample or one specimen in the hand-picked sample. Some taxa have to be found in higher numbers, i.e. Gammarus, “other Trichoptera”, Simuliidae, Oligochaeta, Asellus and Chironomus (see below for further explanation). The macroinvertebrates have to be identified at least to the taxonomic level indicated in Table 2.3.

Table 2.3. Minimum level of identification in Danish Stream Fauna Index (DSFI)

Taxonomic group	Taxa used in Danish Stream Fauna Index (DSFI)
Turbellaria (flatworms)	Tricladida
Oligochaeta (true worms)	Tubificidae, Oligochaeta
Hirudinea (leeches)	Helobdella, Erpobdella
Malacostraca (crustaceans)	Asellus, Gammarus
Plecoptera (stoneflies)	Amphinemura, Brachyptera, Capnia, Isogenus, Isoperla, Isoptena, Leuctra, Nemoura, Nemurella, Perlodes, Protonemura, Siphonoperla, Taeniopteryx
Ephemeroptera (mayflies)	Ametropodidae, Baetidae, Caenidae, Ephemeridae, Ephemerellidae, Heptageniidae, Leptophlebiidae, Siphlonuridae
Megaloptera (alder-fly)	Sialis
Coleoptera (beetles)	Elmis, Limnius, Elodes
Trichoptera (caddis larvae) with cases	Beraeidae, Brachycentridae, Hydroptilidae, Goeridae, Glossosomatidae, Leptoceridae, Lepidostomatidae, Limnephilidae, Molannidae, Odontoceridae, Phryganeidae, Sericostomatidae
Trichoptera (caddis larvae) without cases	Ecnomidae, Hydropsychidae, Philopotamidae, Polycentropodidae, Psychomyiidae, Rhyacophilidae
Diptera (flies and midges)	Psychodidae, Chironomus, Chironomidae, Eristalinae, Simuliidae
Gastropoda (snails)	Ancylus, Lymnaea
Lamellibranchia (mussels)	Sphaerium

2.1.2.4 The French (IBGN) sampling method

Aim
The "Standardised Global Biological Index" (IBGN) is widely used in monitoring programmes in France since 1992.

Sampling

Substrate types and flow are assessed using a survey protocol. Eight different samples are taken from major (representative) habitats in channel. Substrate composition has the main priority in defining mesohabitats and consequently were the samples are going to be taken if less than 8 substrate types are present at the sampling site, the same substrate types are sampled in areas differing in water velocity. Both substrate and velocity classes are predefined in the protocol. Sampling should be undertaken using a Surber sampler (area: 0.25x0.20 m = 0.05 m^2, 500 µm mesh size) where appropriate i.e. when water depth and velocity allows using the Surber sampler. Deeper and more stagnant water bodies/areas are sampled using a hand net (opening: 0.25x0.20 m, 500 µm mesh size): The net is pulled a 0.5 m distance or a similar distance is covered using sweeping movements depending on conditions at the site.

Sample procedure in the field
Samples are preserved in the field. Each of the 8 samples is kept separate. No sorting or removal of excess material occurs in the field.

Laboratory procedure
Each of the 8 samples are sorted and identified separately to produce 8 individual and one combined taxalist. Sorting is undertaken using a stereomicroscope. Samples are identified to fixed taxonomic level consisting of 152 taxa. Insecta, Crusteacea, Mollusca, Achaeta and Turbellaria are identified to the family level whereas Porifera, Cnidaria, Bryozoa, Oligochaeta, Nematomorpha, Nemertina and Hydracarina are identified either to order, class or phylum. The methods allow three different options when recording abundance. It is sufficient to find 3 (or 10) individuals of taxa belonging to "indicator groups" (given in the IBGN assessment system) and one individual of other taxa. Alternatively, abundance can be recorded using abundance classes or by counting all individuals in the sample.

2.1.2.5 The Italian (IBE) sampling method

Aim:

Sampling:
4 sampling seasons per year are requested by national legislation (allowed 3 seasons). Flood periods must be avoided.

No estimation a priori assessment of habitat composition is undertaken. The presence and the scale of dominance of the inorganic habitats are recorded (Table 2.4)

Table 2.4. Example of the recording of inorganic habitats

Inorganic Habitats	Order of occurrence
Rock	4
Boulders	2
Stones	1
Gravel	3
Sand and silt	Not present

The presence of the following organic habitats is also recorded

- Aquatic vegetation (% of cover recorded)
- Periphyton (recorded as absent/few/abundant)
- Filamentous Bacteria (recorded as absent/few/abundant)

A number of kick-samples is undertaken on a transect that should be completed from bank to bank. Where the transect can not be completed (deep water, fast flowing etc.), the sampling is performed from one bank to the middle of the river and return to the same bank. The number of kick-samples is not fixed and depends on the substrate type (fewer kicks if homogeneous).
The sampling device used is a hand net. The mesh size reported in the manual is 21mesh/cm (0.475 mm). Usually the mouth is 20x25. Since the sampling is not quantitative and the number of kick-samples is variable, could be difficult to estimate a sampling area. E.g. in our experience (CNR-IRSA) for STAR samples a good estimation could be 0.9 m^2.

Sample procedure in the field
Usually live sorting in the field is performed. Sorting in lab can be undertaken in case of adverse climate condition. A first identification of the specimens is performed directly in the field and a preliminary value of the index can be assessed.

Laboratory procedure
Sorted sample are preserved in alcohol and further identification is performed in the lab to confirm or not the first assessment.
Identification level is Genus (for Plecoptera, Ephemeroptera, Odonata, Hirudinae and Triclada) and Family (for the others orders). A minimum number or specimens (different for each taxon) must be collected to consider valid the taxa.

2.1.2.6 The Polish sampling method

Aim

The BMWP method provides a score for each macroinvertebrate family that is primarily dependent on its sensitivity to organic pollution. This method was intended to be applied in Poland, operating with a modified BMWP discrimination table.

Sampling

At each sampling section (reach: transect 100m) four quantitative samples are taken, using core sampler - from different substrate patches and morphodynamic units (runs, riffles, pools). The investigated bottom surface per quantitative sample covers an area of 95 cm². Each quantitative sample is kept and analysed separately. At each sampling section (reach: transect 100m) one qualitative sample is collected from all dominated types of river channel habitats. The kick-net sampling method is used (mesh size 300 µm). Studied reaches are mapped for a variety of physical and morphological variables (e.g. organic debris, LWD, erosional and depositional areas, habitat modifications, etc. see below: a list of environmental variables). The precise locations of sampling microhabitats for quantitative samples are marked.

Sample procedure in the field

Samples are preserved with 4% formalin in the field and transported to the laboratory.

Laboratory procedure

In the laboratory, the biological material is sieved, by using hand-net or sieve (300 µm mesh-size). Each sample should be completely sorted. In the case of high abundance of Oligochaeta and/or Chironomidae, sub-sampling of the whole sample is applied. The whole sample is portioning by grid system (4x4 squares), and sub-samples are selected randomly. The organisms from each sub-sample are put into vial separately with detailed description (e.g. 1/16 of the sample number X, vial no. n). All sorted animals are counted and transferred to 70% ethanol. Macroinvertebrates are counted and identified to the family level (except Oligochaeta and some Diptera families).

2.1.2.7 The Czech (PERLA) sampling method

Aim

The PERLA prediction system is a biological method of ecological status assessment of running waters in the Czech Republic, within the STAR project it was also used in the Slovak Republic. The method is based on the comparison of an observed site with a reference site. It takes the natural variability of the environment and within biological communities into consideration and corresponds with present trends in the EU.

The PERLA prediction system is based on the prediction of macroinvertebrate community composition at a specific site using several environmental variables and on the subsequent comparison of the predicted (target) community with the macroinvertebrate community actually found on the site assessed. The application of PERLA requires the compilation of a reference data set for the given geographical region.

The PERLA prediction system is based on the RIVPACS approach

Sampling

Characteristic stretch of the stream: Its length is equal to the 7-fold stream width or to 50 m (depending on which distance is shorter) upstream and downstream of the sampling stretch of the stream. Based on the characteristic stretch the values of some environmental variables are assessed (slope, character of substrate, water plant vegetation, degree of shading, riparian vegetation).

Sampling stretch of the stream: usually, it is not possible to sample the entire characteristic stretch, therefore, a shorter one – the sampling stretch – is defined in its centre. This has to include all habitats present within the characteristic stretch. In smaller streams (river channel width under 5 m), the entire sampling stretch and almost the entire characteristic stretch are sampled. In bigger streams (river channel width of 5m or above), individual sampling sites are selected within the sampling stretch of the stream.

Sampling sites: sites where macrozoobenthos samples are taken. Selection of sampling sites: all habitats present within the sampling stretch are recorded:
- sites of various stream velocity
- various distances from the banks
- grass tufts on the banks, with parts floating in the water
- branches or trunks lying in the water
- roots
- sites with various substrates (stones, sand, fine sediments, etc.)
- tufts of water plants floating in the stream
- calmer water with water plants near to the banks etc.

For each habitat, its percentage area of the total bottom area of the sampling stretch of the stream is estimated. The corresponding sampling time is allocated to each individual habitat (multihabitat sampling). Total net sampling time is 3 minutes. The standard method Three Minute Semiquantitative Kick Sampling using a hand net of 0,5 mm mesh size is employed. Sampling is conducted moving in up-stream direction, thus avoiding disturbance of the area not sampled yet. The substrate infront of the net is disturbed by the foot or hand of the person doing the sampling. Complementary to that, objects from the bottom are left up and investigated so that species firmly attached to them are not overlooked or underestimated. Submerged water plants, leaves of terrestrial plants, tree roots, branches and trunks fallen into the stream, etc. are also washed of infront of the net. A detailed specification of hand sampling in flowing shallow water in of foot sampling in deeper water is given in the standard ČSN EN 27828 (ISO 7828).

Sample procedure in the field

In the field, coarse anorganic sediments are removed from the sample by decanting, twigs, leaves are taken away and the sample is well washed to remove fine sediments.

On site, the sample should be presorted on white photo-pan into glas test-tubes according to taxonomic groups to reduce potential mechanical damage of brittle insect larvae. Surplus water is removed from the rest of the sample, which is placed into a PVC container with a volume of 1 – 2 litres. The sorting of the remaining organisms is done in the laboratory. It is essential that all sampling equipment (particularly the net) are extremely thoroughly washed after each sampling.

The samples are preserved by adding a 40% formaldehyde solution to a resulting 4% concentration. Some organisms are better preserved in ethanol (molluscs, crustaceans). The use of any different preservative should be indicated directly on the sample (ČSN EN ISO 5667-3).

The samples are labled by putting lables inscribed by plain pencil (graphite) into the container or test-tubes. From the outside, the container is marked by a water-resistant felt pen. The label or description on the outside of the container should always include the code identifying the sample, and further information on the sampling date, the stream and the site. A sampling protocol is filled in for each sample.

Laboratory procedure

Biological samples are identified on the lowest taxonomic level possible, i.e. usually on the species level.

The results of identification are recorded in the identification protocol.

The identification protocol has to include:

- code identification of the sample,
- name of stream,
- name of cross section,
- sampling date,
- date of identification,
- name of the person who conducted the identification,
- list of recorded taxa with indication of development stages and their abundances,
- the percentage of sample processed.

2.1.2.8 The Portuguese sampling method (PMP)

Sampling:
Samples should be taken in spring avoiding the influence of floods.
Before sampling the number of habitats at the reached should be assessed (Table 2.5). The percentage of cover of each habitat will be determined.

Table 2.5. Habitat quantification. 6 main habitats are identified

	Habitats	Granulometry	Empiric scale
Inorganic habitats	Bolders and rock	> 256 mm	Higher than an A4 sheet
	Stones	64 – 256 mm	Egg < stones < A4 sheet
	Gravel	2 – 64 mm	Lower than an egg
	Sand and silt	< 2 mm	
Organic habitats	Macrophytes and algae		
	Coarse particulate organic matter		

The stretch to be sampled will dependent on stream width (10 times the average width). A fixed stretch of 50 metres will be sampled when stream average width is higher than 5 metres. The stretch must represent the existent habitat diversity.

The sampling device employed is a hand net with a mouth of 25 cm and mesh size of 0.5 mm. 10 kick samples of one metre length covering the different habitats according to its percentage of cover (now in Portugal only 6 kick samples are being performed).

Sample procedure in the field
Samples will be preserved inside plastic containers using as fixative formalin (10%) or alcohol (90%).
Live sorting is allowed if it occurs during a period no longer than 48 hours after the sampling

Laboratory procedure
Samples must be washed gently under current water using sieves of 1mm (coarse fraction) and 0.5mm (fine fraction). The washed samples will placed inside white plastic trays and the existent organisms removed by eye neck with forceps. The organisms will be preserved with alcohol (70%).
Only fine fraction (>0.5mm and < 1mm) can be subsampled. The fine fraction must be homogenized. A portion of approximately 10% of the total weight of the fine fraction will be taken. If in this portion no more than 200 organisms are present, another similar portion must be taken. This procedure is repeated until in the sum of the all portion more than 200 organisms are sorted. The total amount of organisms of the fine fraction is estimated according to the total percentage of portions sorted.

2.1.2.9 The Latvian sampling method

Aim

The Latvian Standard determines the method and procedure for the assessment of long-term impact of pollution in small streams; the method is based on cenosis of benthic macroinvertebrates. This method is applied for the assessment of biological quality of small rivers and streams at full length or at stretches, as well as for the determination of local impact of pollution. The method is used for assessment of long-term impact of organic pollution.

The method is used for the control of biological quality of small rivers and streams of rithral and potamal type, with current velocity above 0.1 m/s. The method can be applied for the investigation of the whole river or it's single stretches, as well as for the establishing of a local anthropogenic impact, for example, in the intake area of wastewaters.

Sampling

A typical river stretch of 20 – 50 m is selected for sampling, where all the biotopes are investigated (by type of river-bed, composition of bottom, aquatic vegetation and current velocity) and their relative occurrence is determined. Occurrence of various biotopes in river stretches is given in Figure 2.5 The macroinvertebrates are taken with a bottom scraper or picked with forceps from stones or branches or other underwater objects. At the selected reaches of rivers 20 individual samples of benthos are taken and tested like one median sample. The individual samples are taken according to the occurrence of all biotopes. For example, if 50 % of bottom consists of sand, 50 % of samples are taken from sandy biotopes. Organisms, picked from stones and branches, are considered as individual samples.The samples should be taken conversely to the current direction, in order to prevent disturbance of confused bottom to biotopes downstream the sampling site.

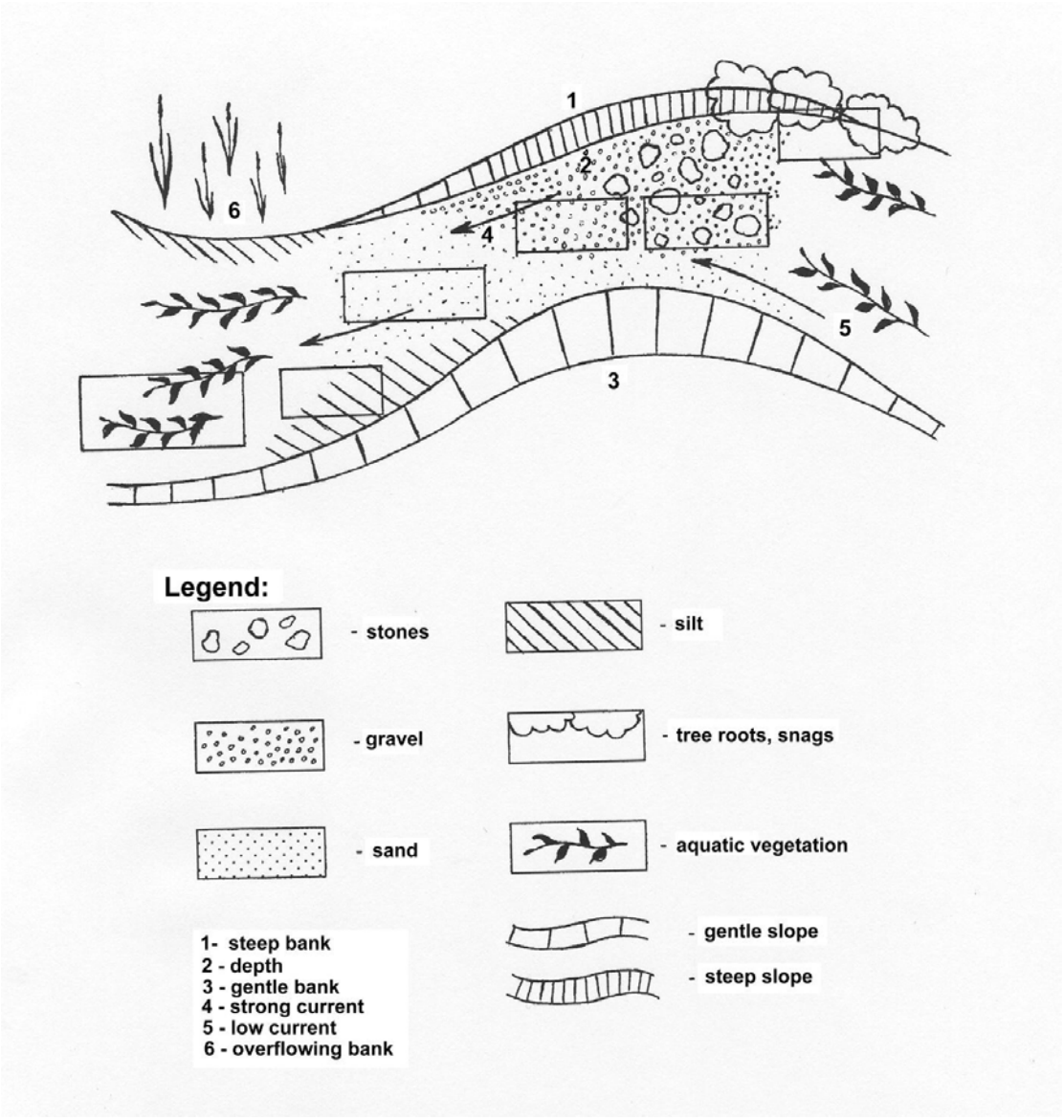


Figure 2.5. Occurrence of various biotopes in the reaches of river

Sample procedure in the field
The samples are put in a sorting tray and investigated at the stream to the relevant taxonomical level, the number of individuals is counted and results are put in the protocol of results. A magnifying glass and keys of identification are used [6 – 10 or other]. If it is impossible to identify the organisms at the field, they must be put in vials and fixed in ethyl alcohol (70%) or formalin (4%). In case of necessity, non-fixed samples can be analysed at the laboratory.

Laboratory procedure
The fixed organisms should be kept in a dark place. Time of storage is unlimited.

At least 12 indicator organisms should be taken to obtain statistically significant results, the sum of relative occurrence of organisms should be at least 30. The saprobity index is calculated.

2.2 COMPARISON OF METHODS USED

The majority of sampling methods employed by the different countries have many features in common (Table 2.6). The majority of methods involve an *a priori* assessment of habitats at the sampling site, exceptions being the RIVPACS method and the DSFI method. In RIVPACS habitats are sampled in proportion to their occurrence, which is subjectively assessed by the surveyor while sampling. DSFI uses a fixed sampling grid that should cover most habitats without introducing a sampling bias due to variability in how surveyors assess number of habitats present.

All methods except the Swedish method use a multi-habitat sampling approach. In contrast, it is the only method, which take replicate samples to assess inter-sample variability. Most methods use standard hand nets with a width of 25 cm and mesh bag with a 500 µm mesh size in accordance to the CEN standard EN 27 828. The samples are therefore semi-quantitative A Surber sampler can be used when employing the STAR-AQEM method, while it is obligatory when using the French IBGN protocol with the exception of sampling in lentic areas. The Polish method uses both a quantitative core sampler and a hand net. Mesh sizes used varies between 300 and 1000 µm. Three of the methods (RIVPACS, DSFI and PERLA) include a pick sample of attached macroinvertebrates.

When samples are obtained using a hand net the area sampled cannot be completely fixed. However, as the sampling effort should be similar as long the sampling protocol is followed, the number of individuals obtained in range samples should be directly comparable. In addition, the area sampled can be roughly estimated from the area disturbed in front of net multiplied with net width (Table 2.6). Using this approximation the area sampled can be compared among methods. The smallest area sampled is 0.4 m^2 (IBGN) and the largest is 2.25 m^2 (STAR-RIVPACS and PERLA) using the assumptions given in the footnote to Table 2.6 To further enable an inter-method comparison, catch per unit effort (CPUE) was calculated for each method using sampled area and mesh size (calculation are given in footnote, Table 2.6). CPUEs ranged from 0.32 (IBGN) to 1.8 (PERLA).

Handling in the field and processing of samples in the laboratory will affect the quality of the assessment result. Field sorting, collection of some species from the sample in the field and removal of excess material can all potentially reduce sample quality by the loss of species. Field handling is extremely dependent on the surveyors' abilities and is affected by weather conditions, time pressure etc. Three methods used field sorting of the whole sample (IBE, PERLA and the Latvian method), four collected some species to further identification in the field (STAR-AQEM, IBE, PERLA and the Latvian method) and excess material was removed using most methods. Only when using DSFI, IBGN, the Polish method and PMP are removal of excess material in the field is not allowed.

Live sorting is only standard when applying the Italian IBE protocol (Table 2.6). When using RIVPACS and Portuguese PMP live sorting is optional but dead sorting is recommended. All other methods rely on the sorting of dead material. Obligatory live sorting is likely to affect quality negatively as it introduces a time constraint on the sorting procedure.

Only the STAR-AQEM method allows sub-sampling of the entire sample. Even though this is not obvious from the CPUE value of the STAR-AQEM method, it collects large amounts of inorganic material, organic debris and plants, which makes sub-sampling necessary. Sub-sampling can potentially reduce the number of species found and hence affect sample quality negatively and increase sampling variance (assesed in Section 5.4).

With regard to sorting under magnification, enumeration of all individuals collected and identification to the best attainable taxonomic level, the methods investigated are highly variable. Sorting under magnification increases the likelihood of finding all species present in sample, even the smaller specimens. Enumeration of all individuals and identification to the best attainable level possible increase the biological information in the sample and hence potentially the quality of the assessment.

To allow an inter-comparison of methods used, a handling-processing score is given in Table 2.6. The score is based on giving the value 1 to each of the handling-processing steps which are considered to be positive for overall assessment quality (0 if negative, see footnote to Table 2.6 for further details) i.e. a high score indicates a high quality method (8 is maximum). The handling-processing score ranges between 1 (IBE) to 7 (Swedish, Polish and Portuguese (PMP) methods) with most methods obtaining scores of either 4 or 5. It should be noted the handling-processing score is subjective and that sample treatment of the indivual methods reflects local conditions, cost-effectiveness and other considerations which is not assessed.

Table 2.6. The score is based on giving the value 1 to each of the handling-processing steps which are considered to be positive for the overall assessment quality (0 if negative, see footnote to Table 2.6 for further details) i.e. a high score indicates a high quality method (8 is maximum).

		STAR-AQEM	RIVPACS	Nordic methods		The French (IBGN) sampling method	The Italian (IBE) sampling method	The Polish sampling method	The Czech (PERLA) sampling method	The Portugeuse (PMP) sampling method	The Latvian sampling method
				The Swedish method	The Danish (DSFI) method						
Strategy	*A priori* habitat assessment	Y	N	Y	N	Y	Y?	Y	Y	Y	Y
	Multi-habitat/number of habitats	Y/V	Y/V	N/1	Y/V	Y/8	Y/V	Y/V	Y/V	Y/V	Y/V
	No. of samples/replicates	20/none	1/none	5/5	12/none	8/none	1/none	4(6)/none	1/none	10/none	20/none
Effort	Sampling device	Hand net or Surber	Hand net	Hand net	Hand net	Surber	Hand net	Core sampler and hand net	Hand net	Hand net	Hand net
	Width of sampling device (m)	0.25	0.25	0.25	0.25	0.25	0.25	0.0095 m^2 0.25	0.25	0.25	0.205
	Mesh size (µm)	500	1000	500	500	500	475	300	500	500	1000
	Kicking technique (area $sample^{-1}$, time used, distance sampled)	Fixed area (0.0625 m^2)	Time (3 min)	Time/distance (1 m 1 min^{-1})	Fixed area (0.1 m^2)	Fixed area (0.05 m^2)	Variable	Fixed area (0.0225 m^2 and 0.0625 m^2)	Time (3min)	Fixed area 1 m sampled (0.205 m^2)	Fixed area 1 m sampled (0.205 m^2)
	Pick sampling (effective sampling time)	N	Y (1 min)	N	Y (5 min)	N	N	N	N	N	Y
	Area covered (m^2)*	1.25	2.25	1.25	1.20	0.4	≈ 0.9[4]	0.09 + 0.375	2.25	2.05	4.1
	CPUE#	1	0.9	1	0.96	0.32	0.77	0.62	1.8	1.64	1.64
Handling	Field sorting	N	N	N	N	N	Y	N	Y	N	Y
	Some species collected from sample in the field	Y	N[1]	N	N	N	Y	N	Y	N	Y
	Excess material removed	Y	Y	Y	N	N	Y	N	Y	N	Y
Processing	Live sorting	N	N[2]	N	N	N	Y	N	N	N[2]	N
	Sub-sampling	Y	N	N	N	N	N	N	N	N[3]	N
	Use of magnification sorting	N	N	Y	N	Y	N	N	N	N	Y
	Enumeration of all** individuals	Y	Y	Y	N	N	N	Y	Y	Y	Y
	Identification to species level##	Y	N	Y	N	N	N	Y	Y	N	N
	Handling/processing score***	4	5	7	5	6	1	7	4	7	4

Notes to table 2.6

* : As not all methods sample a fixed area (most of the methods use hand nets and the area will not be complete constant among samples in the same way a Surber or a core sampler is), the area covered is estimated in the following way: width of sampling device x distance travelled/or distance disturbed in front of the device. With respect to the RIVPACS and the PERLA method it is assumed that sampling distance is 1 m per 20 seconds.

: CPUE is calculated as the area covered and mesh size using the following formula: CPUE = (sample area/1.25 m^2)/(mesh size/0.5 mm). The pick sample is not included in the estimation of CPUE.

** : enumeration of all species could be a total count of all individuals in sample or putting them into abundance classes. In latter case the actual number will often be based on estimation.

: identification to the species level means that all taxa are identified best attainable level and that the subsequent index calculation, to which the sampling method was developed, is at least partly based on species information. It should be noted, however, that all sampling methods will provide data that can be used on various taxonomic levels.

*** : Handling/processing score is calculated by assigning a score of either 0 or 1 to each step in the handling and processing procedure of a sample given in the table. The score one is given if the step is assumed to be beneficial to the overall quality of the taxon list produced: if field sorting is not undertaken the score is 1, if no species are removed (1), if no excess material removed (1), if no live sorting is undertaken (1), if no sub-sampling is undertaken (1), if sorting is done using magnification (1), if all individuals are enumerated (1) and if identification is done to the species level (1). The underlaying assumptions behind this quality scoring is that any sample treatment in the field could potentially affect sample quality negatively by the loss of individuals or by introducing elements of subjectivity (surveyor skills, weather conditions). In the laboratory, the most thorough sorting and identification procedure scores the highest values.

[1] : except rare species which are released into the stream or river again, in the case where it is possible to identify them properly in the field and no biomass data is necessary
[2] : dead sorting is recommended but live sorting is optional
[3] : only the fine fraction (> 0.5 mm and < 1 mm) can be subsampled
[4] : as the number of kick samples varies, the area is set using expert judgement (CNR-IRSA)

3 DATA HANDLING AND ANALYTICAL APPROACH

3.1 SPATIAL SCALES COVERED (COUNTRY, BIOREGIONAL TYPE, PAN-EUROPEAN)

The data collected for this work package cover 13 countries (Austria, Czech Republic, Denmark, France, Germany, Greece, Italy, Latvia, Poland, Portugal, Slovakia, Sweden, and UK). The sampling included 22 stream types, where five were defined as being of the STAR project type "Core stream type 1" (mid altitude, 200-500 m.a.s.l., and with a "small" catchment area 10-100 km^2), seven were of the STAR project type "Core stream type 2" (lowland, <200 m.a.s.l., and "medium" catchment areas 100-1000 km^2), whereas ten other stream types were defined as STAR project type "Additional stream type" (having a different characterisation). These stream types are situated in 11 Ecoregions according to Illies definition (Illies, 1978; as used in the Water Framework Directive), these were regions 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, and 18.

3.2 TAXONOMIC ADJUSTMENT OF THE DATA

Each country has adjusted all of their own taxonomical data, so that there are no biases within each country's dataset caused by differences in taxonomic resolution used (e.g. between sampling seasons, where at some seasons it might be more difficult to identify certain taxa because they are in early instars [e.g. being small]). The taxonomic adjustments were made using common rules within the project. There are three main ways of adjusting the taxonomic data:

• aggregating species to a higher taxonomic level
• omitting a higher taxonomic level
• distributing individuals which are "only" determined to genus level according to the relative share of individuals determined to species level (e.g. 200 individuals determined as *Baetis sp.* could be divided among *Baetis fuscatus* (60 individuals determined) and *Baetis rhodani* (140 individuals determined) according to their relative occurrence 30:70).

All methods can be used within one data set. The choice of the best suited method should be made depending on the taxonomic group at question, based on a combination of individuals occurring and their abundance and the ecological relevance of the species/taxon within the respective taxonomic group. If species either occur in many samples, the abundance of specimens is significant or species differ in their ecological demands , they should be kept separated as individual taxa in the data set and not added at a more coarse taxonomic level.

When applying any of the methods described above, the following criteria should be applied for taxonomic adjustment:
• Taxonomic adjustment always takes place at the best attainable taxonomic level, preferable at species level.
• When a genus is generally identified to species level, with the exception of only a few specimens, the genus level is omitted and specimens determined as Genus sp. are distributed among the species kept.
• When the frequency of occurrence of a genus is more than 20% of the frequencies of occurrence of the underlying species together, all species are aggregated to genus level.
• The 20%-criteria is not a strict rule. In borderline situations a decision can be made based on the ecological indicative value of the genus or of the species in combination with its/their abundance.
• When species and groups/aggregates or genus and family are present, the same criteria are applied for taxonomic adjustment as at the genus and respective species level.
For more details see the AQEM manual; www.aqem.de.

3.3 ENVIRONMENTAL DATA AND RELATIONSHIPS

Within each country one or several STAR project stream types were sampled (see above). Within each country macroinvertebrate samples were taken in two different sampling seasons (all partners sampled in spring and one additional season) (see Table 3.1). A single anthropogenic stressor was also defined for each stream type sampled within the project, with the three main stressor types being:

- Organic pollution (including eutrophication)
- Toxic pollution
- Habitat degradation

For each stream type in each country a pre-defined number of sites were sampled for each level of Ecological status (as defined by the Water Framework Directive), typically ca 3 'High' sites, ca 3 'Good' sites, ca 2 'Moderate' sites, and' ca 2 'Bad sites. For streams affected by hydromorphological degradation, sites with a 'Bad' ecological status were considered unlikely and were not taken into consideration when defining number of streams sampled for each status class. The a priori status classification used to decide where to sample were based on previously collected data and expert judgement using mainly physico-chemical data.

Table 3.1 The stream types, seasons and macroinvertebrate sampling methods used within each country. Only types and seasons where both sampling methods have been used at the same sites are listed.

Country	Types sampled	Seasons sampled	Methods used
Austria	A05 and A06	Spring and summer	RIVPACS and S-A
Czech Republic	C04 and C05	Spring and summer	PERLA and S-A
Denmark	K02	Spring and summer	DSFI and S-A
France	F08	Spring and autumn	IBGN and S-A
Germany	D03, D04 and D06	Spring and summer	RIVPACS and S-A
Greece	H04, H05, H06, and H07	Spring and summer	RIVPACS and S-A
Italy - CNR	I06	Spring and winter	IBE and S-A
Italy - Bolzano	I05	Spring and summer	IBE and S-A
Latvia	L02	Spring and autumn	LVS and S-A
Poland	O02	Spring and autumn	Polish and S-A
Portugal	P04	Spring and autumn	Portuguese and S-A
Slovakia	V01	Spring and autumn	PERLA and S-A
Sweden	S05 and S06	Spring and autumn	Swedish and S-A
United Kingdom	U15 and U23	Spring and autumn	RIVPACS and S-A

It was not always possible to take samples using the two sampling methods used within each country, and for the comparison of sampling methods, data were only included in the analysis where both sampling methods for macroinvertebrates were sampled at the same site in the same stream in the same season. The number of samples used for these comparisons therefore differed between types, seasons, and methods used (Table 3.2).

Table 3.2 The number of samples used in the comparison of sampling methods for each type, season and macroinvertebrate sampling methods within each country.

Country	Types sampled	Seasons sampled	Methods used
Austria	54 and 26	40 and 40	40 and 40
Czech Republic	56 and 40	48 and 48	48 and 48
Denmark	44	22 and 22	22 and 22
France	48	24 and 24	24 and 24
Germany	48, 48, and 24	60 and 60	60 and 60
Greece	10, 40, 20, and 20	20 and 70	45 and 45
Italy - CNR	38	16 and 22	19 and 19
Italy - Bolzano	40	40 and 40	20 and 20
Latvia	96	48 and 48	48 and 48
Poland	102	50 and 52	51 and 51
Portugal	40	20 and 20	20 and 20
Slovakia	24	12 and 12	12 and 12
Sweden	62 and 46	52 and 56	54 and 54
United Kingdom	52 and 48	50 and 50	50 and 50

3.4 COMMON METRICS USED

All national metrics should be used (i.e. BMWP/ASPT, DSFI, IBE etc). when comparing classifications at the national level. However, it is not relevant to test national methods and specific metrics (e.g. for a certain stream type) on the general data set. Therefore a group of metrics was selected which is generally applicable and covers various types of stress. The metrics vary in intrinsic properties as to which features of the macroinvertebrate community they respond, i.e. structural (incl. sensitivity), functional or life cycle properties. The general metrics selected are shown in Table 3.3. These metrics are calculated from species data using the various national methods and the STAR-AQEM method. This allows for a direct comparison of the performance of the national method compared with the STAR-AQEM method for each country individually.

Table 3.3. Common metrics used for the comparison of national methods and the STAR-AQEM method.

Metric	Type
M1 maximal size	Trait
M2 number of reproductive cycles per year	Trait
M7 Locomotion and substrate relation	Trait
M12 Current velocity (Preferendum)	Trait
Saprobic Index (Zelinka & Marvan)	Structural (sensitivity)
Abundance	Structural
ASPT	Structural (sensitivity)
Shannon-Wiener index	Structural (diversity)
EPT-taxa	Structural (sensitivity)
No. of taxa	Structural (diversity)
No. of families	Structural (diversity)
Oligochaeta [%]	Structural (insensitivity)
RETI	Functional
%Grazers	Functional
%Gatherers	Functional
%Shredders	Functional

3.5 ANALYSIS OF DATA AT THE NATIONAL LEVEL

The purpose is to compare each of the national methods to the STAR-AQEM method. This was accomplished by performing a Students *t*-test, or a non-parametric Sign test (Sokal & Rohlf, 1995) if the differences in metric values between the STAR-AQEM and national method for a given site and season were not normally distributed. These tests were performed on the 16 metrics given in Table 3.3. Furthermore the correlation between the STAR and national method was investigated by Spearman's rank correlation.

For metrics with high correlation, the functional relationship between the STAR-AQEM and national method was investigated and estimated. For a number of selected metrics we plotted their dependence on CPUE, latitude, stressor type/strength/gradient, habitat complexity etc. and possible regressions were investigated.

Box and whisker plots were used to summarise and compare the data for the STAR and national methods for given country and season. This type of plots was also used for plotting CPUE versus selected metrics.

3.6 ANALYSIS SPECIES DATA

The species data (as opposed to the metrics or index data) were used to compare the sampling methods (STAR-AQEM and one other method) in terms of community composition and assess if there are biases in taxa collected using one versus the other method. The analysis was performed using Mantel tests (Mantel, 1967). This test compares two dissimilarity matrices (in this case one with data collected using the STAR-AQEM method and one dataset collected using the other [e.g. national] sampling method. The null hypothesis of the test is of no relationships between the two dissimilarity matrices. The inter-sample community similarities was based on the Bray-Curtis similarity index using the raw taxonomic abundances. The analysis were performed in PCOrd version 4.25 (McCune & Mefford, 1999) using Mantel´s asymptotic approximation, based on the algorithm by Douglas & Endler (1982).

Further investigations of differences in taxonomic composition collected using the two compared methods involved the Indicator value approach (Dufrêne& Legendre, 1997) was also performed using the PC-Ord program. The method tries to find indicator taxa for groups of samples (in this case the groups were defined as the two different sampling methods). The method takes into account both the specificity of a certain taxon and also the fidelity of that taxon to a certain group. Taxa or indicators are defined using an Indicator Value (IV) which goes from 0 to 100 %, where taxa or indicators who are only found in one group, and within all sites of that group receives an IV of 100%. The statistical significance of the IV values is tested using a randomisation procedure (in our case using 999 permutations). Differences in number of taxa collected using the two methods were also assessed using Students *t*-test with all data from all stream types and seasons together, within each country.

3.7 STATISTICAL ANALYSIS METHODS

3.7.1 Univariate methods

The comparison of ecological classifications was done using chi-square analyses of samples taken using the two methods in the same stream in the same season. Not all countries have Water Framework Directive compliant ecological status classification methods, but most countries had some method that could be used for this analysis (Table 3.4).

Table 3.4 Ecological status classification systems based on macroinvertebrates used in the comparisons of sampling methods.

Country	Ecological status classification based on macroinvertebrates
Austria	National Austrian multimteric index
Czech Republic	Multimetric index developed in AQEM project
Denmark	Danish Stream Fauna Index
France	Not possible to compare IBGN with STAR-AQEM samples
Germany	German Saprobic system and multimetric module "Generel degradation"
Greece	BMWP, ASPT, BMWP (Spanish), IBE quality class, Shannon-Wiener index
Italy Brughero	Indice Biotico Esteco
Italy Bolzano	Indice Biotico Esteco
Latvia	Saprobity Index (SI)
Poland	Modified BMWP score for Poland
Portugal	IM9 index developed during the AQEM project
Slovakia	Slovak Saprobic index
Sweden	Swedish Ecological Quality Criteria (ASPT, DSFI and an acid index)
United Kingdom	RIVPACS predictive system and GQA classes

Direct relationships (correlations) between community composition of macroinvertebrates and a number of environmental factors (land use in the catchment, hydromorphological factors, physical and chemical variables (pollution), and substratum at the sampling site) were analysed using linear regression. The analysis were performed by correlating the first and second axes of a Non Metric Multidimensional Scaling (NMDS) ordination of the sample scores with the first four axes (gradients) of a Principal Component Analysis (PCA) of the environmental data for the same sample and sites. The inter-sample community similarities was based on the Bray-Curtis similarity index using the raw taxonomic abundances.

3.7.2 Multivariate methods

Comparisons of the community structure among sampling methods, stream types, sampling seasons, and ecological status (pre-defined) were done in two different ways. Firstly by comparing ordination scores of the first and second axes of a Non Metric Multidimensional Scaling (NMDS) using the Bray-Curtis similarity index (based on raw abundances), all calculated the computer program PCOrd. The analysis was run specifying two axes, with 0.0005 as stability criterion, standard deviations in stress over the last 30 iterations using 100 iterations in the analyses. Statistical comparisons of differences were simply made by Students *t*-test of the ordination scores for each axis separately and defining two groups based on either method of sampling, stream type, sampling season or the pre-defined ecological status gradient. The ordination diagrams in the report are based on these analyses.

Secondly comparisons were made using Canonical Correspondence Analysis (CCA) (ter Braak, 1987) using the computer program CANOCO version 4.5 (ter Braak & Smilauer, 1998). Here the sampling method, stream type, season, and pre-defined ecological classification were entered in the analyses as dummy variables (coded 0 and 1) and in the analysis the importance of each variable in explaining the community composition data were calculated and compared. In a second step the macroinvertebrate sampling methods dummy variable was automatically selected and the amount of variation explained simultaneously (i.e. interaction terms) by sampling method and the three other parameters were evaluated.

4 COMPARABILITY OF INVERTEBRATE SAMPLING METHODS

4.1 COMPARISON OF THE STAR METHODOLOGY AND NATIONAL ASSESSMENT METHODS

4.1.1 Correlation between STAR-AQEM and National methods

The majority of the 16 metrics analysed using values derived from the STAR-AQEM method and the various national methods correlated significantly, and positively to each other (Table 4.1). Only a few correlations were negative. However, despite being significant a substantial number of correlations had coefficients below 0.7. Overall, number of EPT-taxa was the metric that was most highly correlated when compared among countries. Also the RETI index was highly correlated in most countries. The metric with the overall weakest correlation in an inter-country comparison was abundance. Especially 4 countries exhibited strong correlations between their national method and the STAR-AQEM method. These were Czech Republic, Germany, Sweden and the UK. In contrast, especially Italy but also Denmark and Portugal had many weak correlations although some lack of significance can be explained from the low number of sites in these countries. Strong correlations do not necessarily mean that methods will provide identical results. However, they show that results from the different method can be compared.

4.1.2 Comparison of AQEM-STAR and national methods

The 16 metrics were calculated from samples obtained using the various national methodologies and the STAR-AQEM method. Only main samples were used so that each site was represented by one sample per season. Performance of the national method and the STAR-AQEM method was tested using pair-wise comparisons for each country individually. For both seasons combined, no overall clear pattern emerged with respect to the differences between metric results obtained using STAR-AQEM and national methods (Table 4.2). Some national methods performed better (i.e. scored significant higher values) than the STAR-AQEM method in some countries and vice versa in other countries. Within countries there was in most cases not a consistent pattern when comparing metrics: compared to the STAR-AQEM method, some metrics would score higher when calculated using data obtained by the national method while other would score lower.

In most cases (64% of the countries) the various national methods yielded significantly higher EPT-taxa values than the STAR-AQEM method. A similar pattern was evident with respect to number of families in 73% of the countries significantly more families were found using the national method. In contrast, the STAR-AQEM method yielded significantly more EPT-taxa and families in 9% and 27% of the countries, respectively.

The STAR-AQEM method appeared to perform better than the national methods in Italy and Latvia. With respect to Italy, this could reflect the very low handling-processing score compared with the STAR-AQEM method as well as the other national methods (Table 2.6). With respect to Latvia, the differences are likely to reflect that a number of taxa are not considered in the national method and they will consequently not appear in the taxa list (section 2.10). In Sweden and Portugal, the national method performed consistently better than the STAR-AQEM method. This could relate to the use of subsampling in the STAR-AQEM methodology, which might reduce the number of taxa found. In the case of Sweden, the higher number of taxa (all and EPT) and families might reflect that the sampling effort is concentrated in riffles which are the most species rich in stream ecosystems (e.g. Brown & Brussock, 2001). In Denmark and Germany, significantly more individuals were found when employing the STAR-AQEM method whereas the opposite was true with respect to number of EPT-

taxa and families. Again, this might reflect that taxa are lost when subsampling the large STAR-AQEM sample.

Several countries used the RIVPACS method as their national method (Austria, Germany, Greece and UK; Table 2.6). In addition, the Czech PERLA system is very closely related to the RIVPACS method (Table 2.6). Overall, there were no clearly consistent results among these countries. This could indicate that differences in how sampling is undertaken among countries are as important as the intrinsic differences in the methods employed.

Table 4.1. Correlation matrix between the STAR-AQEM method and the respective national methods. In the top panel for each country are correlations calculated on spring samples and in the lower panel correlations are calculated using summer/autumn samples. Significant correlations are denoted: * p<0.05; ** p<0.005: ***p<0.0005. Note that the number of samples varies among countries and between seasons. Therefore, similar correlation coefficients might not have the same p-value.

	Abundance	Number of taxa	Saprobic	ASPT	Shannon - Wiener	Grazers (%)	Shredders (%)	Gatherers (%)	RETI	Oligochaeta (%)	EPT – taxa	Number of families	M1	M2	M7	M12
Austria	0.34	0.80***	0.62**	0.63**	0.84***	0.87***	0.80***	0.71**	0.85***	0.69**	0.74***	0.73***	No data	No data	No data	No data
	0.66**	0.71***	0.84***	0.51*	0.70**	0.71***	0.79***	0.72***	0.76***	0.29	0.77***	0.65**	No data	No data	No data	No data
Czech Rebublic	0.63**	0.86***	0.96***	0.89***	0.73***	0.71***	0.93***	0.85***	0.82***	0.80***	0.94***	0.87***	No data	No data	No data	No data
	0.39	0.90***	0.83***	0.90***	0.60**	0.65**	0.94***	0.72***	0.89***	0.87***	0.94***	0.87***	No data	No data	No data	No data
Denmark	0.67*	0.71*	0.82**	0.91***	0.14	0.55	0.28	0.72*	0.70*	0.49	0.72*	0.85**	0.43	0.58*	0.63*	0.76***
	0.17	0.85**	0.68*	0.95***	0.76**	0.73*	0.61*	0.87***	0.87***	0.54	0.95***	0.90***	0.73**	0.95***	0.89***	0.84***
France	0.76*	0.70*	Not possible	0.83**	0.83**	0.87**	0.85**	0.86**	0.83**	0.59	0.85**	0.44	No data	No data	No data	No data
	0.13	0.51	Not possible	0.78**	0.82**	0.81**	0.80**	0.55	0.84**	0.73*	0.83**	0.57	No data	No data	No data	No data
Germany	0.66***	0.51*	0.93***	0.78***	0.68***	0.92***	0.83***	0.76***	0.70***	0.45*	0.86***	0.43*	No data	No data	No data	No data
	0.77***	0.66**	0.89***	0.74***	0.82***	0.87***	0.82***	0.60**	0.72***	0.19	0.65**	0.53*	No data	No data	No data	No data
Greece	0.05	0.93***	Not possible	0.61	0.78*	0.61	0.59	0.77*	0.66*	0.76*	0.79*	0.92***	No data	No data	No data	No data
	0.13	0.60***	Not possible	0.70***	0.79***	0.84***	0.34*	0.67***	0.47*	-0.18	0.64***	0.58***	No data	No data	No data	No data
Italy	0.10	0.49	0.52	0.39	0.56	0.18	0.50	0.35	0.38	No data	0.58	0.35	No data	No data	No data	No data
	-0.26	0.37	0.55	0.73*	0.22	0.63*	0.52	0.67*	0.37	No data	0.65*	0.33	No data	No data	No data	No data
Latvia	0.30	0.62**	0.64**	0.20	0.40	0.49*	0.65**	0.55*	0.73***	0.28	0.46*	0.63**	0.54**	0.17	0.33	0.60***
	0.49*	0.66**	0.79***	0.73***	0.38	0.37	0.85***	0.59**	0.63**	0.58**	0.25	0.69***	0.12	0.76***	0.37	0.73***
Poland	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	0.67***	0.39	0.46*	0.73***
	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	0.56**	0.84***	0.67***	0.70***
Portugal	0.37	-0.18	0.88**	0.75*	0.59	0.30	0.88**	0.88**	0.09	0.67*	0.45	-0.09	No data	No data	No data	No data
	0.54	0.82*	0.84**	0.63	0.72*	0.55	0.92***	0.66	0.76*	0.55	0.56	0.86**	No data	No data	No data	No data
Slovakia	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	0.99***	0.99***	0.55	0.84*
	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	No data	0.49	0.60	1.00***	0.89*
Sweden	0.50*	0.80***	0.85***	0.84***	0.85***	0.82***	0.72***	0.70***	0.82***	0.59**	0.83***	0.74***	No data	No data	No data	No data
	0.67***	0.64***	0.93***	0.91***	0.51*	0.85***	0.68***	0.78***	0.81***	0.53**	0.83***	0.69***	No data	No data	No data	No data
UK	0.82***	0.77***	0.80***	0.88***	0.77***	0.89***	0.83***	0.77***	0.77***	0.62**	0.95***	0.88***	0.86***	0.93***	0.72***	0.91***
	0.60**	0.90***	0.93***	0.92***	0.66***	0.58**	0.85***	0.64**	0.64**	0.39	0.95***	0.89***	0.78***	0.89***	0.90***	0.86***

Table 4.2 Significant differences between the AQEM-STAR method (S) and the respective national methods (N). Analyses are based on all sites and seasons within each country.

	Abun-dance	Number of taxa	Saprobic index	ASPT	Shannon - Wiener	%Grazers	%Shred-ders	%Gat-herers	RETI	%Oligo-chaeta	EPT – taxa	Number of families	M1	M2	M7	M12
Austria					S > N*	S > N*	S < N**				S < N*	S < N*	N/A	N/A	N/A	N/A
Czech	S < N*	S < N*	S > N*		S > N**			S > N*		S > N***	S < N***	S < N*	N/A	N/A	N/A	N/A
Denmark	S > N***	S < N***									S < N***	S < N***				
France	S < N***		S < N***		S > N***						S < N**	S < N***	N/A	N/A	N/A	N/A
Germany	S > N*										S < N**	S < N*	N/A	N/A	N/A	N/A
Greece		S > N***	N/A									S > N***	N/A	N/A	N/A	N/A
Italy	S > N***	S > N***					S < N*			N/A	S > N**	S > N***	N/A	N/A	N/A	N/A
Latvia	S > N***	S > N*			S < N***			S > N***		S > N***		S > N***	S > N***		S > N*	S > N***
Poland	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		S < N*		S < N*
Portugal		S < N**									S < N*	S < N*	N/A	N/A	N/A	N/A
Slovakia	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	S > N*			
Sweden		S < N***				S < N*	S < N*				S < N*	S < N***	N/A	N/A	N/A	N/A
UK	S > N***		S > N*		S > N**	S < N*	S < N***	S > N***	S < N***	S > N***		S < N*	S < N*	S > N***	S < N***	S > N**

p < 0.05; ** p < 0.005; ***p < 0.0005

In conclusion, the STAR-AQEM method was not superior to the majority of national methods. In contrast, the STAR-AQEM method appears to collect fewer taxa (all and EPT) and families than the majority of the national methods. The most likely explanation for this finding is that species are lost during the sub-sampling procedure employed by the STAR-AQEM method. Two methods, the Italian IBE method and the Latvian method, appear to loose information about the macroinvertebrate community to a degree that might affect the assessment of ecological stream quality. They yielded almost in all cases lower metric values than the STAR-AQEM method. Laboratory processing (IBE) and identification of more species (Latvian method) would properly improve their performance.

4.1.3 Comparison of seasonal variability between the AQEM-STAR method and the national methods

Austria

For the majority of metrics tested there were no significant differences between the STAR-AQEM and the RIVPACS samples either in spring or summer (Figs App. 1.1a-k). There was, however, a change in which method performed best (i.e. gave significant higher metric values) between spring and summer. In spring, the RIVPACS method yielded significantly higher values for number of taxa, number of EPT taxa, number of families ($p<0.05$) and %shredders ($p<0.005$). In contrast, when analysing summer samples the STAR-AQEM method yielded significantly higher values with respect to number of taxa, ASPT ($p<0.05$), abundance ($p<0.005$) and Shannon Wiener index ($p<0.0005$).

Czech Republic

The % Oligochaeta were significantly higher in STAR-AQEM samples both spring ($p<0.05$) and summer ($p<0.005$) compared with PERLA samples (Fig. App.1.2g). In contrast, number of EPT taxa was significantly higher when using the PERLA method ($p<0.05$ in spring; $p<0.0005$ summer)(Fig. App.1.2c). In spring, abundance ($p<0.05$) and number of families ($p<0.05$) were higher in PERLA samples whereas the opposite were true for Shannon-Wiener index ($p<0.005$) and %gatherers ($p<0.005$). In summer, significantly more taxa ($p<0.05$) and a higher RETI score ($p<0.05$) where found in PERLA samples. Overall, the differences between the STAR-AQEM and PERLA method appeared to be small and independent of season (Figs. App.1.2a-k). However, the higher %oligochaeta and lower number of EPT-taxa in STAR-AQEM samples both seasons might reflect the habitats sampled as PERLA is very similar to the RIVPACS method (see UK section below).

Denmark

The significant differences between the STAR-AQEM method and the DSFI method were consistent between spring and summer (Figs. App.1.3a-k). Abundance in the STAR-AQEM sample was significantly higher both in spring ($p<0.005$) and summer ($p<0.0005$) compared to the DSFI sample (Fig. App.1.3a). In contrast, DSFI yielded significant higher values both spring and summer with respect to number of taxa ($p<0.005$ (spring), $p<0.0005$ (summer)), number of EPT-taxa ($p<0.005$ (spring), $p<0.005$ (summer)) and number of families ($p<0.05$ (spring), $p<0.0005$ (summer)). The trait metrics M1 was significantly higher in the DSFI sample in spring ($p<0.05$) whereas M7 was significantly higher in the AQEM-STAR sample ($p<0.05$).

France

The most significant differences between the STAR-AQEM and the IBGN were found in the spring samples (Figs. App. 1.4a-k). Abundance ($p<0.005$), number of taxa ($p<0.05$), Saprobic index ($p<0.005$), number of EPT-taxa ($p<0.05$) and number of families ($p<0.0005$) were all significantly higher in the IBGN sample in spring. The Shannon-Wiener Index and % shredders were significantly lower ($p<0.005$ and 0.0005, respectively) in IBGN samples than in AQEM-STAR samples in spring. In autumn samples, only abundance ($p<0.005$), Saprobic index ($p<0.0005$) and number of families ($p<0.0005$) were significant different between the two methods. In all cases yielded IBGN higher values. That the Saprobic index were significantly higher both spring and summer indicates that the additional taxa found using the IBGN method could be species tolerant towards organic pollution, as higher Saprobic values indicate a higher degree of degradation.

Germany

There was almost no significant differences between the STAR-AQEM method and RIVPACS method when analysing the German spring and summer samples separately (Figs. App. 1.5a-k). In spring, RIVPACS yielded significantly more %grazers ($p<0.05$) and number of EPT-taxa ($p<0.05$) and significantly less individuals (abundance, $p<0.05$) in summer. The high degree of similarity between the two sampling methods might reflect that they have been taken by the same group of surveyors, which have estimated the number and composition of habitats in the same manner.

Greece

In both seasons, the STAR-AQEM method collected a significantly higher number taxa ($p<0.0005$ in spring; $p<0.005$ in summer) and families ($p<0.0005$ in spring; $p<0.005$ in summer) than the RIVPACS method (Figs App. 1.6d&k). All other differences between the two methods were not significant ($p>0.05$; Figs App. 1.6a-c;e-j) and as for Germany this might reflect that the surveyors did the sampling and habitat assessment very similarly.

Italy

The Italian data exhibited strong seasonal differences in the number of significant differences between the STAR-AQEM method and the IBE method (Figs. App. 1.7a-k). In spring, only the number of individuals caught (abundance) was significantly higher ($p<0.005$) using the STAR-AQEM compared to the IBE method. In summer, however, 8 of the 11 metrics tested were significantly different between the two methods. The following were significantly higher in STAR-AQEM samples: abundance ($p<0.005$); number of taxa ($p<0.0005$); number of EPT-taxa ($p<0.005$) and number of families ($p<0.0005$). In contrast, the Saprobic index value was significantly higher ($p<0.05$) using the IBE method as were Shannon-Wiener ($p<0.05$) and %shredders ($p<0.05$). The results clearly indicate that there is a loss of individuals and species when employing the IBE method during summer. This could relate the sampling procedure in which the effort put into sampling reflects the availability of suitable habitat. In summer with low flow and narrow wetted perimeter the number of habitats could be assessed to be low. Another explanation could be handling and processing of the samples, which are highly dependent on field identifications.

Latvia

The Latvian method collected consistently significant ($p<0.0005$) fewer individuals both spring and summer compared to the STAR-AQEM method (Fig App.1.8a). The same was true for trait M1 metric which was significantly higher in both seasons using the STAR-AQEM method (Fig. App. 1.8g; $p<0.005$ in spring; $p<0.0005$ in summer). All other significant differences between the two methods varied with season: number of taxa ($p<0.0005$); %gatherers ($p<0.005$); number of EPT-taxa ($p<0.005$) and number of families ($p<0.005$) were significantly higher in the STAR-AQEM samples in spring but not in summer ($p>0.05$). RETI was significantly higher ($p<0.05$) using the Latvian method in spring. In summer, %gatherers ($p<0.0005$); %oligochaeta ($p<0.005$) and M7 ($p<0.05$) were significantly higher using the STAR-AQEM method, whereas the opposite was true for the Shannon-Wiener index which was significantly higher ($p<0.005$) using the Latvian method (Figs. App. 1.8b-f;h-l). The STAR-AQEM method clearly appears to collect more species during spring than the Latvian method whereas this is not the case when samples are taken during summer.

Poland

Almost no data. The trait metrics M2 and M12 were significantly higher ($p<0.05$) when employing the Polish method in spring compared with the STAR-AQEM method (Fig App.1.9b&d). There were no significant differences in summer with respect to the trait metrics (Figs. App. 1.9a-d).

Portugal

There was very few significant differences between the Portuguese PMP method and the STAR-AQEM method (Figs. App. 1.10a-l). There was no significant differences in spring whereas in autumn the number of taxa, EPT-taxa and families were significantly higher using the PMP method ($p<0.05$). The high degree of similarity between the two methods is likely to reflect that the PMP method is developed from the original AQEM method.

Slovakia

Almost no data. The trait metrics M1 and M2 were significantly higher ($p<0.05$) when employing the STAR-AQEM method in autumn compared with the RIVPACS method (Figs. App. 1.11a&b). There were no significant differences in spring with respect to the trait metrics (Figs. App. 1.11a-d).

Sweden

Number of taxa were significantly higher in both spring ($p<0.005$) and autumn ($p<0.0005$) samples using the Swedish method compared to the STAR-AQEM method (Fig. App. 1.12l). In addition, %grazers was significantly higher ($p<0.05$) in springs samples, and number EPT-taxa ($p<0.05$) and families ($p<0.0005$) in autumn samples, when employing the Swedish method (Fig. App.1.12c;d;f). The only metric that was significantly higher in STAR-AQEM samples was %shredders in spring ($p<0.05$; Fig. App. 1.12j). These results might reflect thathigher sampling effort in riffles by the Swedish method is especially important in autumn. This could reflect life cycle patterns or flow conditions, where low flow are likely to increase the number of species in riffle habitats.

United Kingdom

Out of the 32 metrics tested (spring and autumn combined), 21 were significantly different when comparing the RIVPACS method with STAR-AQEM method (Figs. App. 1.13a-p). There was no overall clear consistent pattern with respect to which method yielded the highest metric values or with season. RIVPACS samples yielded significant higher values in both spring and autumn for the following metrics: %shredders ($p<0.0005$ in spring; $p<0.05$ autumn); RETI ($p<0.005$ in spring; $p<0.005$ autumn) and M7 ($p<0.005$ in spring; $p<0.005$ autumn). The opposite (i.e. STAR-AQEM yielded higher values both seasons) was true for the following metrics: abundance ($p<0.005$ in spring; $p<0.05$ autumn); Saprobic index ($p<0.05$ in spring; $p<0.05$ autumn); %oligochaeta ($p<0.0005$ in spring; $p<0.0005$ autumn) and M2 ($p<0.005$ in spring; $p<0.05$ autumn). ASPT ($p<0.05$); number of EPT-taxa ($p<0.05$) and families ($p<0.005$) were significantly higher in spring when employing the RIVPACS method whereas the opposite was true with respect to the Shannon-Wiener index ($p<0.05$). In autumn, %grazers ($p<0.005$) were significantly higher in RIVPACS samples whereas the opposite was true for %gatherers ($p<0.0005$) and M12 ($p<0.05$). The metric values probably reflect that not the same habitats were sampled using the two methods. Especially the spring samples differ in this respect. The higher number of e.g. oligochaeta and lower number of EPT-taxa in STAR-AQEM samples compared with the RIVPACS samples in spring are likely to reflect that more samples STAR-AQEM are taken in slow flowing, depositional habitats than the RIVPACS samples. A reason for this difference could be that habitats with a coverage less than 5% are sampled using the RIVPACS method whereas this is not the case for the STAR-AQEM method which do not assess rare habitats.

4.1.4 Inter-country comparison of metric performance

Abundance

There was no relationship between abundance of macroinvertebrates in samples and sampling area or CPUE (Figs. 4.1 & 4.2). The French methods IBGN had a significant higher number of individuals than all methods and at the same time the lowest sampling area and CPUE. If the IBGN method is omitted from the data set there is a tendency for an increase in number of individuals caught with increasing CPUE. Abundance was significantly lower in samples with a handling/processing score of 1 (the IBE method) whereas abundance varied independently of the score in the range 4 to 7 (Fig. 4.3).

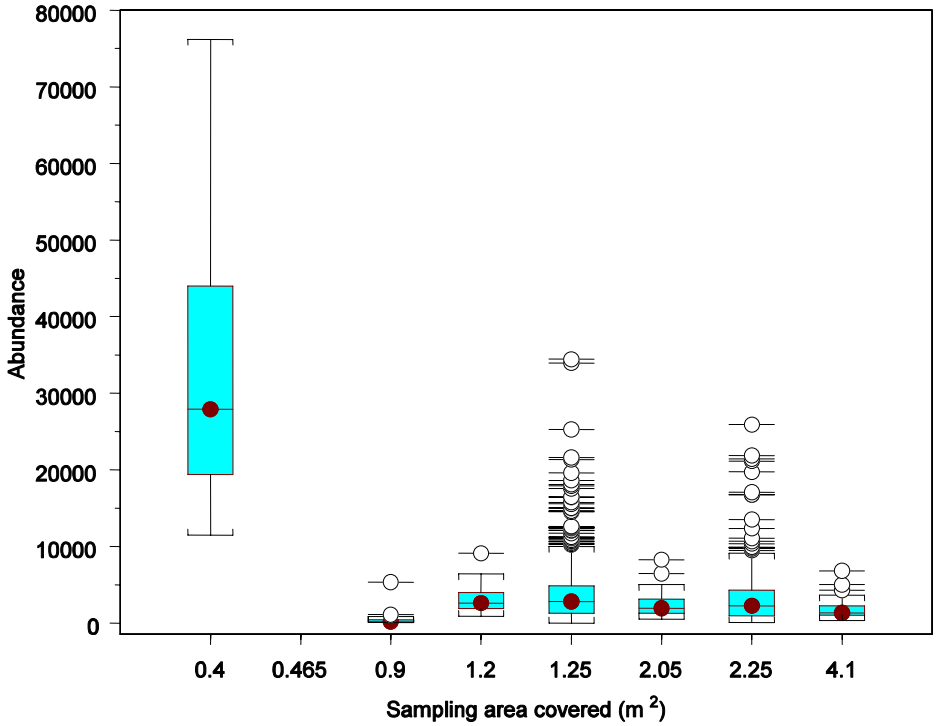


Figure 4.1 Abundance in relation to sampling area covered.

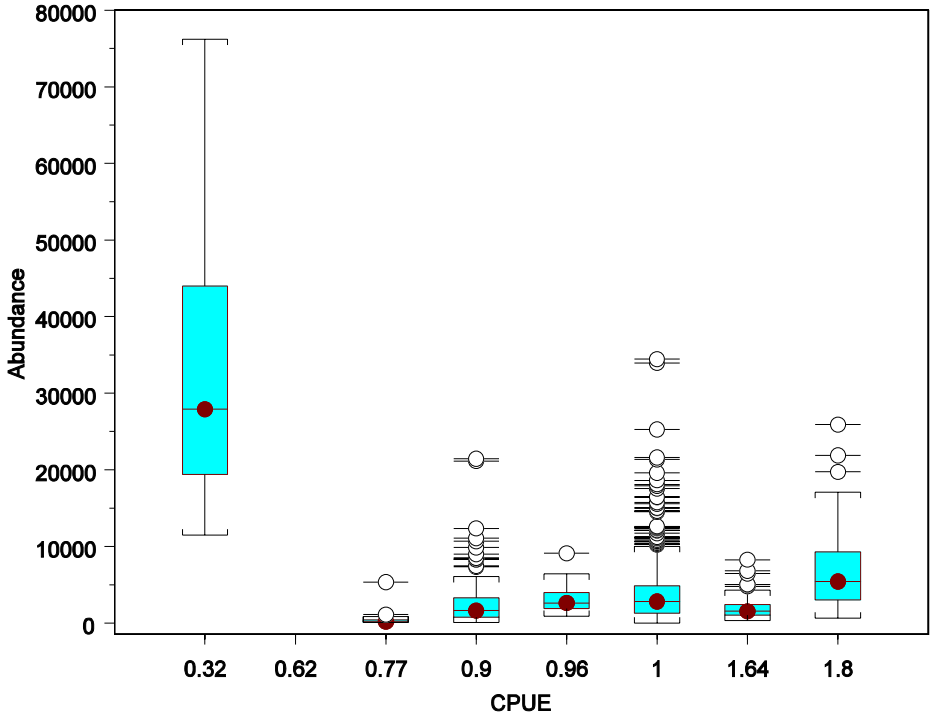


Figure 4.2 CPUE in relation to sampling area covered.

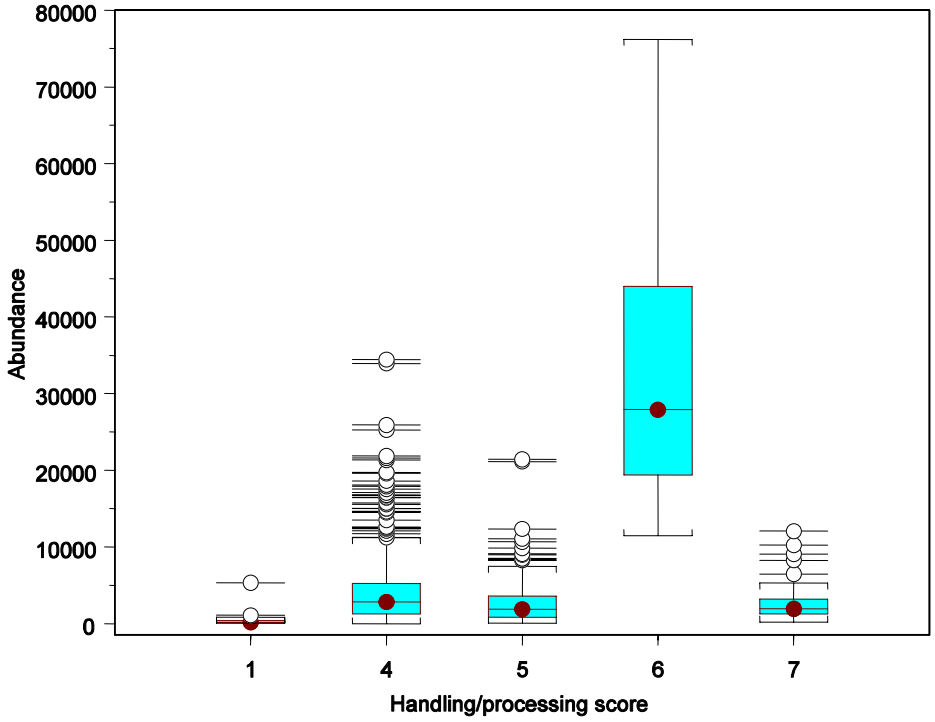


Figure 4.3 Abundance in relation to handling/processing score.

There was no clear relationship between the number of EPT-taxa and sampling area and CPUE (Fig. 4.4 & 4.5). The IBGN method caught a similar number of EPT-taxa as the other methods despite the small area covered and low CPUE. There was a tendency that the number of EPT-taxa found increased with increasing handling/processing score indicating that these taxa are lost during sample treatment (Fig. 4.6).

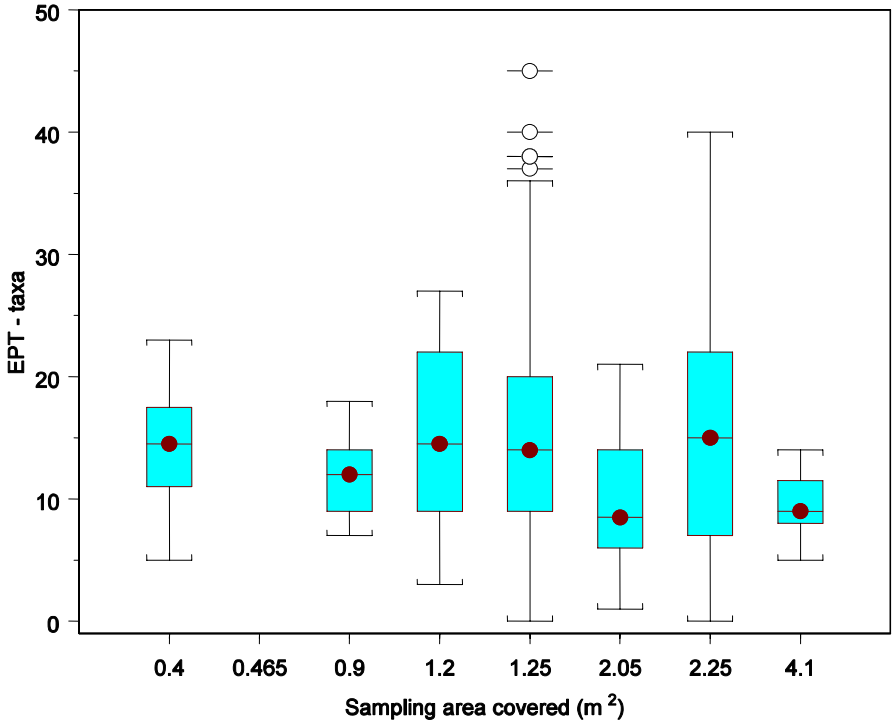


Figure 4.4 EPT-taxa in relation to sampling area.

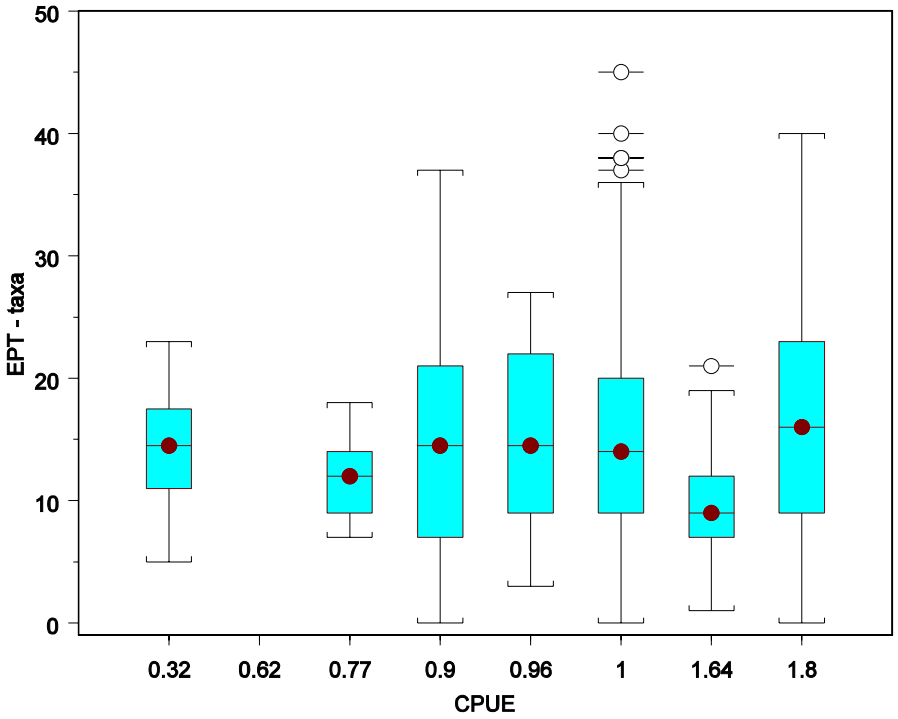


Figure 4.5 EPT-taxa in relation to CPUE.

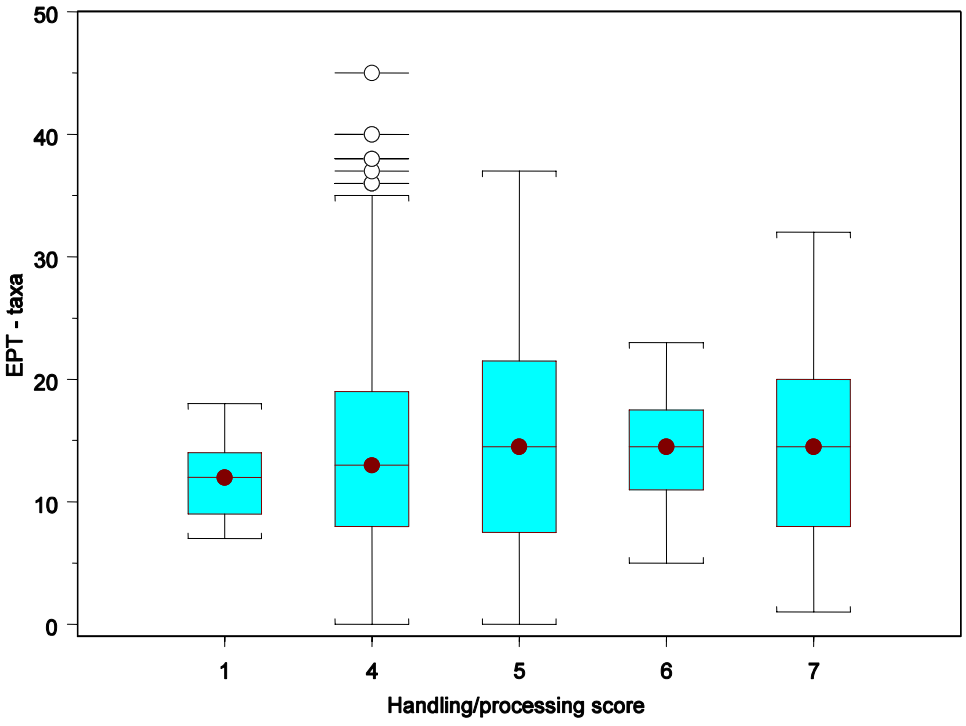


Figure 4.6 EPT-taxa in relation to handling/processing score.

The number of families found was not related to sampling area and CPUE (Fig. 4.7 & 4.8). As with abundance, the method with the smallest sampling area and CPUE caught the largest number of families (the IBGN method). There was no effect of the handling/processing score on the number of families (Fig. 4.9)

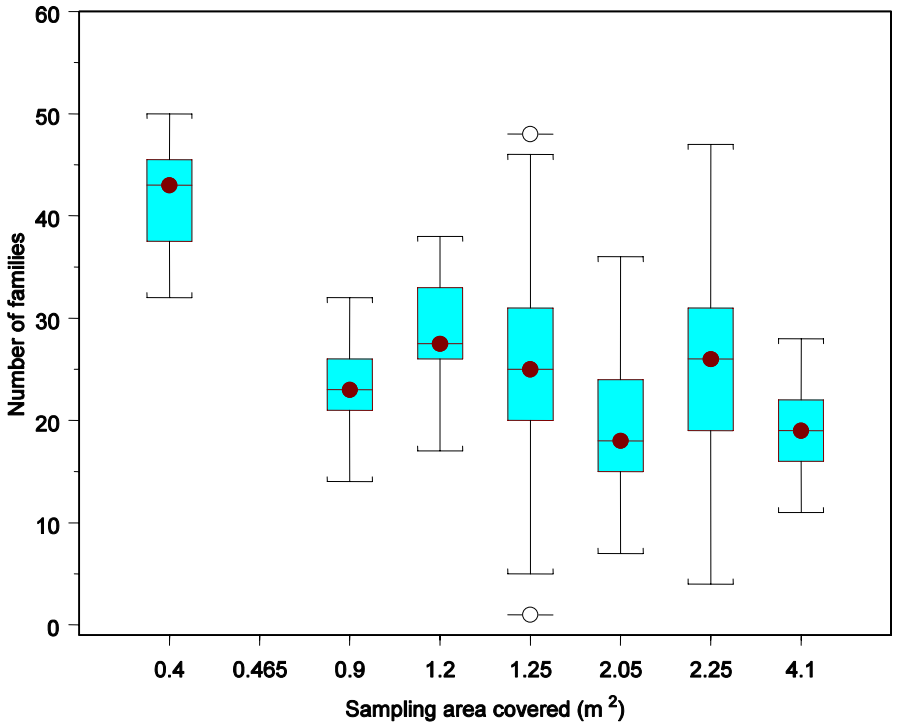


Figure 4.7 Number of families in relation to sampling area.

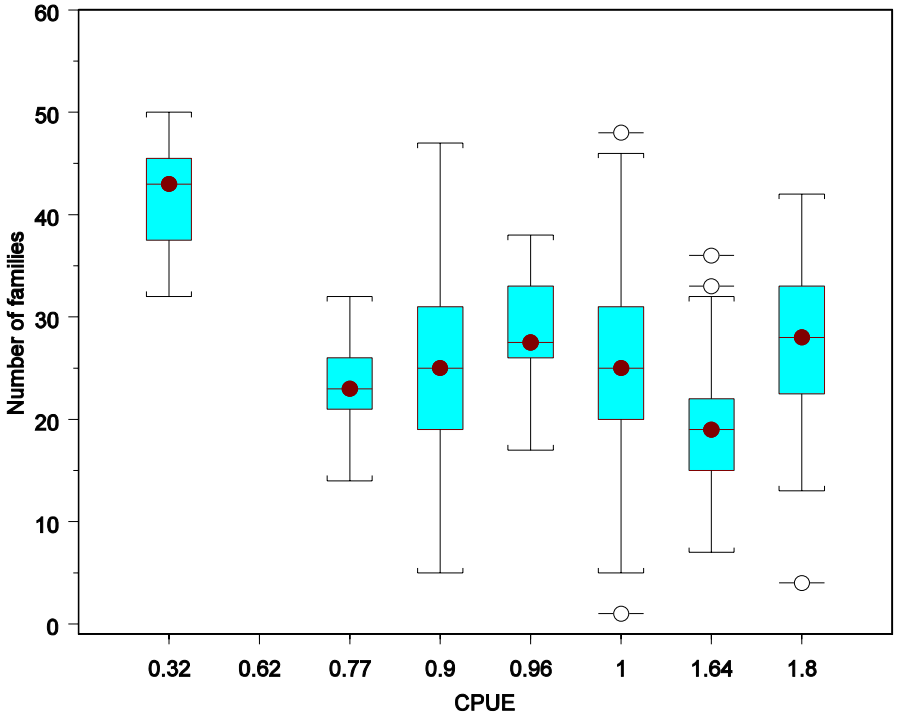


Figure 4.8 Number of families in relation to CPUE

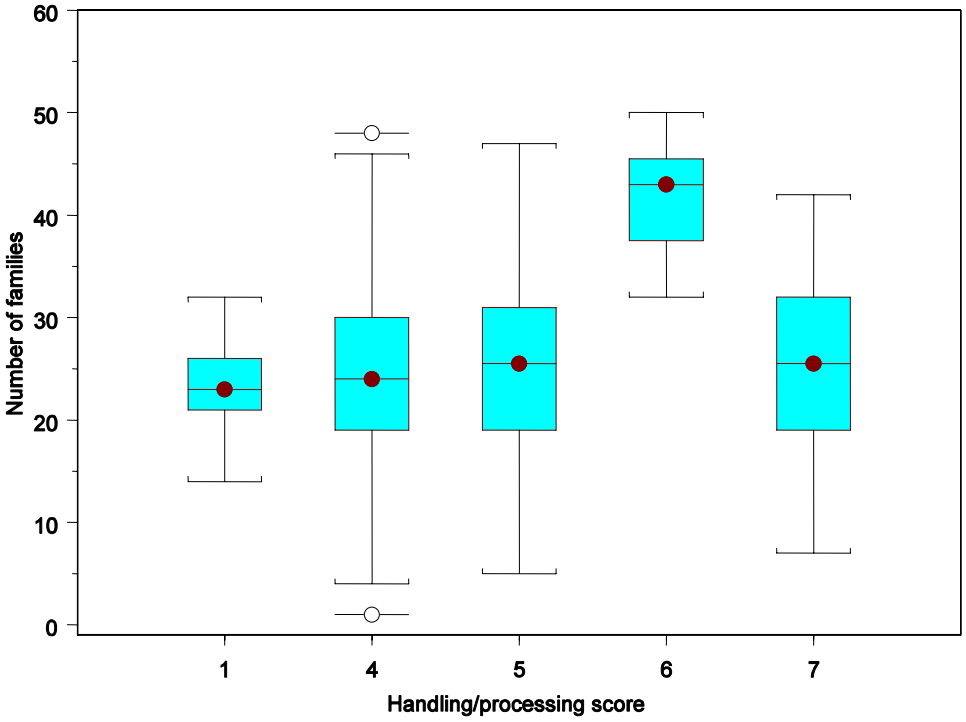


Figure 4.9 Number of families in relation to handling/processing score.

The number of taxa was - as for the other metrics tested - not related to sampling area and CPUE (Fig. 4.10 & 4.11). There was a high degree of variability which appears to be method specific and that cannot be explained from single variables as sampling area and CPUE. With respect to handling/processing score there was a tendency that the number of taxa found increases with increasing score (Fig. 4.12).

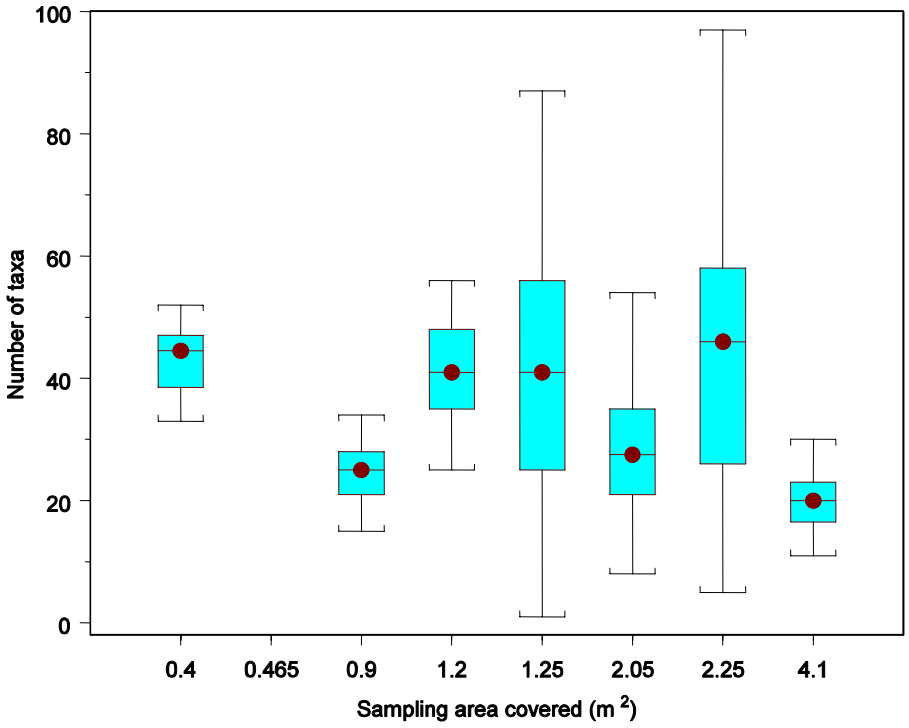


Figure 4.10 Number of taxa in relation to sampling area

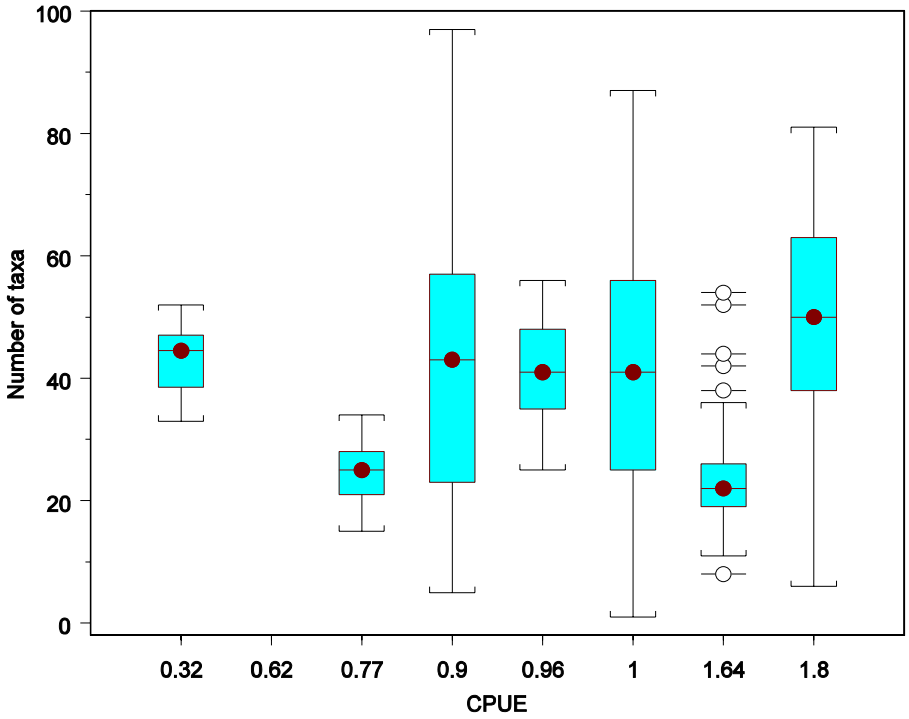


Figure 4.11 Number of taxa in relation to CPUE

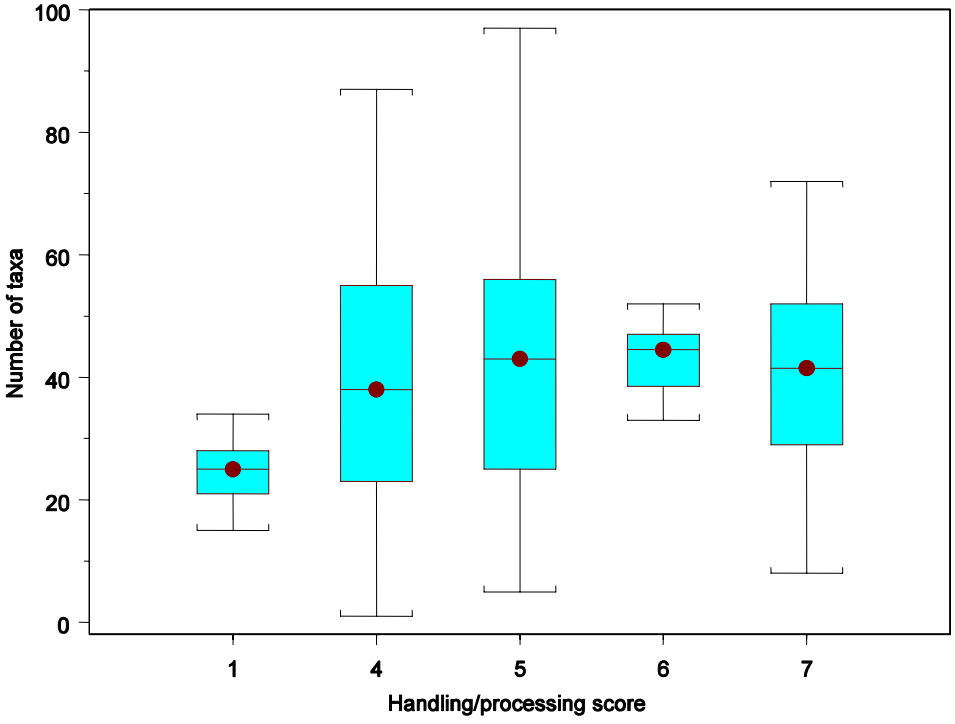


Figure 4.12 Number of taxa in relation to handling/processing score

4.2 COMPARISONS OF ECOLOGICAL CLASSIFICATIONS

The ecological classifications based on macroinvertebrates were compared among samples taken using the STAR-AQEM sampling method and the other (national etc) method. For the Austrian data 19 of the 40 pairs (one STAR-AQEM and one RIVPACS sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 47.5% (Table 4.3). Only one RIVPACS sample was classified outside one class above or below the STAR-AQEM classification, i.e. 97.5% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the RIVPACS method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total six sites out of 40 (15.0%) were classified "across" this boundary. In all cases the samples were classified as "moderate" according to one method and 'high' or 'good' by the other method.

Table 4.3 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the RIVPACS sampling method (top-bottom) for Austrian samples.

	High	Good	Moderate	Poor	Bad
High	5	6			
Good	6	4	3		
Moderate	1	2	4	2	
Poor				1	1
Bad					5

For the Czech Republic data 32 of the 48 pairs (one STAR-AQEM and one PERLA sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 66.7% (Table 4.4). Only two PERLA sample were classified outside one class above or below the STAR-AQEM classification, i.e. 95.8% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the PERLA method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total seven sites out of 48 (15%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'high' or 'good' by the other method.

Table 4.4 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the PERLA sampling method (top-bottom) for Czech Republic samples.

	High	Good	Moderate	Poor	Bad
High	12	4	1		
Good	1	8	6		
Moderate			4	1	1
Poor				5	1
Bad				1	3

For the Danish data 15 of the 22 pairs (one STAR-AQEM and one DSFI sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 68.2% (Table 4.5). No DSFI samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the DSFI method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total five sites out of 22 (23%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'good' by the other method.

Table 4.5 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the DSFI sampling method (top-bottom) for Danish samples.

	High	Good	Moderate	Poor	Bad
High	10	1			
Good		2	3		
Moderate		2	2		
Poor				1	1
Bad					

For the German data 39 of the 60 pairs (one STAR-AQEM and one RIVPACS sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 65.0% (Table 4.6). No RIVPACS samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the RIVPACS method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total eleven sites out of 60 (18.3%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'good' by the other method.

Table 4.6 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the RIVPACS sampling method (top-bottom) for German samples.

	High	Good	Moderate	Poor	Bad
High		2			
Good		23	4		
Moderate		7	7	1	
Poor			4	7	1
Bad				2	2

For the Greek data 27 of the 45 pairs (one STAR-AQEM and one RIVPACS sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 60.0% (Table 4.7). No RIVPACS samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the RIVPACS method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total eight sites out of 45 (17.8%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'good' by the other method.

Table 4.7 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the RIVPACS sampling method (top-bottom) for Greek samples.

	High	Good	Moderate	Poor	Bad
High	4	2			
Good	3	13	2		
Moderate		6	8	4	
Poor			1	1	
Bad					1

For the Italian CNR data 6 of the 19 pairs (one STAR-AQEM and one IBE sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 31.6% (Table 4.9). No IBE sample were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the IBE method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total two sites out of 19 (10.5%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'good' by the other method. The IBE values were calculated using the AQEMrap assessment software, at the time of production of this report, the index was to some degree miscalculated using this software. Comparisons can therefore only be made within this dataset, but no conclusions should be made outside this dataset.

Table 4.9 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the IBE sampling method (top-bottom) for Italian CNR samples.

	High	Good	Moderate	Poor	Bad
High	5				
Good	11	1			
Moderate		2			
Poor					
Bad					

For the Italian Bolzano data 15 of the 20 pairs (one STAR-AQEM and one IBE sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 75.0% (Table 4.10). No IBE sample were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the IBE method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. None of the sites out of 20 (0%) were classified "across" the boundary. None of the Bolzano sites were at all classified outside the high-good boundary, with 14 out of 20 sites classified as having a high ecological status according to the IBE method. The IBE values were calculated using the AQEMrap assessment software, at the time of production of this report, the index was to some degree miscalculated using this software. Comparisons can therefore only be made within this dataset, but no conclusions should be made outside this dataset.

Table 4.10 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the IBE sampling method (top-bottom) for Italian Bolzano samples.

	High	Good	Moderate	Poor	Bad
High	14				
Good	5	1			
Moderate					
Poor					
Bad					

For the Latvian data, the comparison looks a little bit different, since the Latvian classification system consists of eight classes from "very clean" to "strongly polluted". The sites sampled in the STAR project were all classified as "clean to slightly polluted", "slightly polluted" or "slightly polluted to polluted". A total of 34 of the 42 pairs (one STAR-AQEM and one LV sample taken in the same stream in the same season) were classified into the same ecological class in all cases "slightly polluted" i.e. 81.0% (Table 4.11). No LV samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits, a reason for this was that all STAR samples except two were classified into one quality class (slightly polluted).

Table 4.11 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the LV sampling method (top-bottom) for Latvian samples. The Latvian classification system consists of eight classes: very clean, clean, clean to slightly polluted, slightly polluted to polluted, polluted, polluted to strongly polluted, strongly polluted.

	Clean to slightly polluted	Slightly polluted	Slightly polluted to polluted
Clean to slightly polluted		6	
Slightly polluted		34	2
Slightly polluted to polluted			

For the Polish data data 27 of the 51 pairs (one STAR-AQEM and one Polish sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 82.4% (Table 4.12). One Polish samples was classified outside one class above or below the STAR-AQEM classification, i.e. 98% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the Polish method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total two sites out of 51 (3.9%) were classified "across" the boundary, in both cases the sites were classified as High-Good by the STAR-AQEM method and Moderate by the Polish method.

Table 4.12 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the Polish sampling method (top-bottom) for Polish samples.

	High	Good	Moderate	Poor	Bad
High	31	4	1		
Good	2	5	1		
Moderate			1	1	
Poor				2	
Bad					3

For the Portuguese data 11 of the 20 pairs (one STAR-AQEM and one Portuguese sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 55.0% (Table 4.13). No Portuguese samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the Portuguese method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total two sites out of 20 (10.0%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'good' by the other method.

Table 4.13 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the Portuguese sampling method (top-bottom) for Portuguese samples.

	High	Good	Moderate	Poor	Bad
High	2	3			
Good	4	5	2		
Moderate			3		
Poor				1	
Bad					

For the Slovak Republic data 6 of the 11 pairs (one STAR-AQEM and one PERLA sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 54.5% (Table 4.14). No PERLA samples were classified outside one class above or below the STAR-AQEM classification, i.e. 100% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the RIVPACS method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total one site out of eleven (9.1%) was classified "across" the boundary, this sample was classified as "moderate" according to one method and 'good' by the other method.

Table 4.14 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the PERLA sampling method (top-bottom) for Slovak republic samples.

	High	Good	Moderate	Poor	Bad
High	1				
Good	1	1			
Moderate		1	2	3	
Poor				1	
Bad					1

For the Swedish data 38 of the 56 pairs (one STAR-AQEM and one kick sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 67.9% (Table 4.15). Only one kick sample was classified outside one class above or below the STAR-AQEM classification, i.e. 98% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the kick sampling method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total ten sites out of 56 (18%) were classified "across" the boundary.

Table 4.15 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the kick sampling method (top-bottom) for Swedish samples.

	High	Good	Moderate	Poor	Bad
High	1	2	1		
Good	2	17	3	1	
Moderate		5	11	3	
Poor			1	8	
Bad					1

For the UK data 31 of the 50 pairs (one STAR-AQEM and one RIVPACS sample taken in the same stream in the same season) were classified into the same ecological class (from 'high' to 'bad'), i.e. 62.0% (Table 4.16). Only one RIVPACS sample was classified outside one class above or below the STAR-AQEM classification, i.e. 98% of the samples using the two methods were classified within these limits (e.g. a sample classified as 'moderate' by STAR-AQEM were also classified as 'good', 'moderate' or 'poor' according to the RIVPACS method). Another way to compare the two sampling methods and differences in ecological classification is relationships between sites classified as 'high' or 'good' by one method and 'moderate' to 'bad' with the other method or vice versa, since the 'good' 'moderate' boundary is the most important according to the Water Framework Directive. In total six sites out of 50 (12.0%) were classified "across" the boundary, the samples were in all cases classified as "moderate" according to one method and 'high' or 'good' by the other method.

Table 4.16 Comparison of classification of macroinvertebrate samples taken either using the STAR-AQEM sampling method (left-right) and the RIVPACS sampling method (top-bottom) for UK samples.

	High	Good	Moderate	Poor	Bad
High	16	7			
Good		4	6	1	
Moderate			2	2	
Poor				4	1
Bad				2	5

4.3 SUMMARY OF COMPARISONS OF ECOLOGICAL CLASSIFICATIONS

The number of ecological quality classifications giving the same quality class using the two sampling methods (STAR-AQEM versus either a national method or the RIVPACS method) gave quite consistent results. Between 31.6% and 82.4% of the samples were classified into the same class using the two macroinvertebrate methods.

The results from the countries sampling using the RIVPACS method (Austria, Germany, Greece, and UK) varied between 47.5% and 65% with most countries consistently a little bit above 60%.The IBE method differed most compared to the STAR-AQEM method when comparing the ecological classification (only 31.6% were the same in the CNR dataset). In 13 out of 19 cases did the IBE method classify a site one step lower than the STAR-AQEM method did. In the Bolzano dataset, no sites were classified across the good-moderate border, but that was because all sites seemed to have a high or good ecological status and no sites were near the good-moderate boundary.

Very few classifications differed more than one class in any direction using the two methods and six of the countries had all of their compared samples within one quality class in each direction (e.g. if the national/RIVPACS methods classifies a site as “moderate” then the STAR-AQEM method shouldn’t assess the same site as worse than “poor” or better than “good”). Finally, according to the Water Framework Directive, the good-moderate boundary is the most important, since sites with a quality below this boundary has to be restored. Generally less than 20% of the sites were classified “across” this border using the two methods, where almost in all cases one method classified the site as having a “good” and the other a “moderate” status.

Generally the PERLA and Polish sampling methods seemed to be most similar to the STAR-AQEM method. 14.6% of the sites in the Czech Republic, 9.1% in the Slovak Republic and only 3.9% of the sites in Poland were classified across the border. One reason for the Polish results was also that many of the Polish sites were classified as “high” by both sampling methods, and the ecological quality was thus far from the good-moderate class boundary where the misclassification rate is the highest.

Country	**Same class**	**Not more than one away**	**Outside good-moderate**
Austria	47.5 %	97.5 %	15.0 %
Czech Republic	66.7 %	95.8 %	14.6 %
Denmark	68.2%	100.0 %	22.7%
France	--	--	--
Germany	65.0%	100.0%	18.3%
Greece	60.0%	100.0%	17.8%
Italy CNR	31.6%	100%	10.5%
Italy Bolzano	75.0%	100.0%	0.0%
Latvia	81.0%*	100.0%*	--
Poland	82.4%	98.0%	3.9%
Portugal	55.0%	100.0%	10.0%
Slovak Republic	54.5%	100.0%	9.1%
Sweden	67.9%	98.2%	17.9%
United Kingdom	62.0%	98.0%	12.0%

** The Latvian classification system consists of eight classes and can therefore not easily be compared with the other assessment systems.*

4.4 COMPARISON OF DIFFERENCES IN TAXA COMPOSITION

The taxonomic composition for all sites sampled using STAR-AQEM versus the national (or RIVPACS) sampling method was compared using Mantel tests (see above) (Table 4.17). For all comparisons, there were significant similarities between the STAR-AQEM versus the national or RIVPACS method used in each country. The STAR-AQEM and the RIVPACS method gave very similar results (method used in Austria, Germany, Greece, and the UK). The results were also very similar for the two Nordic methods (DSFI in Denmark and the Swedish standards method). The PERLA method on the other hand gave quite different results; it came out very similar to the STAR-AQEM method in the Czech Republic, but not in the Slovak Republic. The least similar results were obtained when comparing the Italian IBE method and the STAR-AQEM method and for the Slovak PERLA samples.

Table 4.17 Mantel tests comparing the STAR-AQEM samples with the national (or RIVPACS) sampling method using a similarity measure.

Country	**Mantel statistic (r)**	**t**	**p-value**
Austria	0.667	8.3572	< 0.00000001
Czech Republic	0.777	19.1458	< 0.00000001
Denmark	0.676	5.9804	< 0.00000001
France	0.393	5.2286	0.00000021
Germany	0.740	21.1542	< 0.00000001
Greece	0.518	9.2608	< 0.00000001
Italy (CNR)	0.370	3.3702	0.00078
Italy (Bolzano)	0.263	2.8499	< 0.005
Latvia	0.371	5.6389	0.00000002
Poland	0.299	5.3873	0.00000009
Portugal	0.325	3.2816	0.00106795
Slovakia	0.423	2.4175	0.01580459
Sweden	0.580	12.4218	< 0.00000001
United Kingdom	0.672	13.1037	< 0.00000001

Table 4.18 Differences in taxonomic groups (number of taxa) caught using the two sampling methods, compared using two sample *t*-tests. Significant differences denoted with N/R = higher number of taxa caught using the national or RIVPACS method, S-A = significant higher number of taxa caught using the STAR-AQEM sampling method

No of taxa per group	AT	CZ	DK	F	D	GR	IT CNR	IT Bolz.	LV	PL	PO	SV	SE	UK
Porifera	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Coelenterata	n.s.	n.s.	n.s.	<0.001 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Cestoda	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Trematoda	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Turbellaria	n.s.	n.s.	n.s.	<0.005 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Nematoda	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	<0.05 N/R	n.s
Nematomorpha	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Gastropoda	n.s.	<0.05 S-A	n.s.	<0.005 N/R	n.s.	n.s.	<0.05 S-A	n.s.	<0.05 S-A	n.s.	n.s	n.s	n.s	n.s
Bivalvia	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Polychatea	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Oligochaeta	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	<0.005 N/R	n.s	<0.0005 N/R	<0.005 S-A
Hirudinea	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	<0.05 S-A	n.s.	n.s	n.s	n.s	n.s
Crustacea	<0.001 N/R	n.s.	n.s.	<0.05 N/R	n.s.	n.s.	<0.05 S-A	n.s.	n.s.	n.s.	n.s	<0.001 N/R	n.s	n.s
Araneae	n.s.	n.s.	n.s.	n.s	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s	n.s
Ephemeroptera	n.s.	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	<0.01 S-A	n.s.	<0.005 N/R	n.s	n.s	n.s
Odonata	n.s.	n.s.	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s
Plecoptera	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	<0.01 N/R	n.s	n.s
Heteroptera	n.s.	n.s.	<0.05 N/R	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s
Planipennia	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s
Megaloptera	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s
Trichoptera	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	<0.005 S-A	n.s.	n.s.	n.s.	n.s.	n.s	n.s	n.s
Lepidoptera	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	<0.01 N/R	n.s
Coleoptera	n.s.	n.s.	<0.05 N/R	<0.005 S-A	n.s.	n.s.	<0.01 S-A	n.s.	n.s.	<0.001 N/R	n.s.	n.s	n.s	n.s
Diptera	n.s.	n.s.	n.s.	n.s.	<0.001 S-A	n.s.	<0.05 S-A	<0.05 S-A	n.s.	<0.001 S-A	n.s.	n.s	<0.0001 N/R	n.s
Bryozoa	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	Not calc.	n.s.	n.s
Hydrachnidia	n.s.	n.s.	n.s.	n.s.	n.s.	<0.001 S-A	n.s.	n.s.	n.s.	n.s.	n.s.	Not calc.	n.s.	n.s
Others	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	Not calc.	n.s.	n.s
EPT-Taxa	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	<0.05 S-A	n.s.	n.s.	n.s.	n.s.	Not calc.	n.s.	n.s

4.5 INDICATOR SPECIES ANALYSIS

When comparing how many taxa were indicative of the STAR-AQEM versus national/RIVPACS samples in terms of number of taxa captured, more taxa were over-represented in national/RIVPACS samples. In France five taxa were over-represented when sampling was taken using the IBGN method compared to the STAR-AQEM method and in Sweden four taxa were over-represented when samples was taken using the Swedish standard method as opposed to the STAR-AQEM method. Only one partner (the Italian CNR partner) had a high number of taxonomic groups over-represented when sampling using the STAR-AQEM method versus the national (IBE) method (Table 4.18). All countries who used the RIVPACS sampling method had only one or two taxonomic groups over-represented using any method, whereas France (using IBGN) and Italy (CNR using the STAR-AQEM) method had the highest number of over-represented taxonomic groups.

Table 4.19 Number of taxonomic groups (in terms of number of taxa) over-represented in samples taken using the national/RIVPACS method versus the STAR-AQEM method for each partner.

Country	National/RIVPACS	STAR-AQEM
AT	1	0
CZ	2	1
DK	3	0
F	5	1
D	0	1
GR	0	1
IT CNR	0	6
IT Bolz.	1	1
LV	0	3
PL	1	1
PO	2	0
SV	2	0
SE	4	0
UK	0	1
Total	21	16

The taxonomic groups most commonly over-represented in the study were Diptera (over-represented five times), Gastropoda, Crustacea, and Colepotera (over-represented four times), and Oligochaeta and Ephemeroptera (over-represented three times) (Table 4.20).

Table 4.20 Number of countries where the individual taxonomic groups is over-represented using the national/RIVPACS (N/R) sampling method or the STAR-AQEM (S-A) sampling method.

No of taxa per group	N/R	S-A
Porifera	0	0
Coelenterata	1	0
Cestoda	0	0
Trematoda	0	0
Turbellaria	1	0
Nematoda	2	0
Nematomorpha	1	0
Gastropoda	1	3
Bivalvia	0	0
Polychatea	0	0
Oligochaeta	2	1
Hirudinea	0	1
Crustacea	3	1
Araneae	0	0
Ephemeroptera	2	1
Odonata	1	0
Plecoptera	2	0
Heteroptera	1	0
Planipennia	0	0
Megaloptera	0	0
Trichoptera	0	1
Lepidoptera	1	0
Coleoptera	2	2
Diptera	1	4
Bryozoa	0	0
Hydrachnidia	0	1
Others	0	0
EPT-Taxa	0	1

The taxa indicative (over-represented) in RIVPACS samples in Austrian streams all belonged to the groups of Oligochaeta, Diptera, Coleoptera, Trichoptera, and Crustacea (Table 4.21) whereas the samples over-represented in STAR-AQEM samples belonged to the groups Diptera, Ephemeroptera, and Trichoptera. A total of 17 statistically significant indicator taxa were found for RIVPACS samples, whereas much fewer (five) indicator taxa were found for STAR-AQEM samples.

Table 4.21 Indicator (over-represented) taxa in comparing the two sampling methods in Austrian streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
RIVPACS	Oligochaeta	TUBIFICIDAE	Aulodrilus japonicus	66.2	0.0005
RIVPACS	Diptera	CHIRONOMIDAE	Brillia bifida	65.3	0.0012
RIVPACS	Coleoptera	GYRINIDAE	Orectochilus villosus Lv.	64.2	0.0174
RIVPACS	Diptera	LIMONIIDAE	Hexatoma sp.	63.3	0.0122
RIVPACS	Coleoptera	ELMIDAE	Esolus sp. Lv.	53.3	0.005
RIVPACS	Diptera	CHIRONOMIDAE	Prodiamesa olivacea	50.2	0.0281
RIVPACS	Diptera	CHIRONOMIDAE	Brillia flavifrons	46.9	0.0011
RIVPACS	Oligochaeta	LUMBRICIDAE	Eiseniella tetraedra	44.2	0.0016
RIVPACS	Trichoptera	LIMNEPHILIDAE	Potamophylax sp.	43.9	0.0276
RIVPACS	Diptera	CHIRONOMIDAE	Heleniella sp.	40.4	0.0104
RIVPACS	Coleoptera	DYTISCIDAE	Oreodytes sanmarkii Lv.	32.6	0.0227
RIVPACS	Diptera	CHIRONOMIDAE	Synorthocladius semivirens	30.4	0.037
RIVPACS	Coleoptera	SCIRTIDAE	Scirtidae Gen. sp. Lv.	24.3	0.032
RIVPACS	Trichoptera	LIMNEPHILIDAE	Halesus sp.	24.1	0.0094
RIVPACS	Crustacea	GAMMARIDAE	Gammarus fossarum	19	0.0083
RIVPACS	Diptera	CHIRONOMIDAE	Nilotanypus dubius	18.7	0.0168
RIVPACS	Diptera	CHIRONOMIDAE	Thienemanniella sp.	17.5	0.0389
STAR-AQEM	Diptera	CHIRONOMIDAE	Macropelopia sp.	48.5	0.0145
STAR-AQEM	Diptera	CHIRONOMIDAE	Apsectrotanypus trifascipennis	45.7	0.0003
STAR-AQEM	Ephemeroptera	BAETIDAE	Baetis rhodani	39.7	0.0009
STAR-AQEM	Ephemeroptera	LEPTOPHLEBIIDAE	Paraleptophlebia submarginata	29.7	0.0228
STAR-AQEM	Trichoptera	SERICOSTOMATIDAE	Sericostoma sp.	25.5	0.0329

The taxa indicative (over-represented) in PERLA samples in Czech Republic streams belonged to the groups of Bivalvia, Diptera, Ephemeroptera, and Oligochaeta, (Table 4.22) whereas the samples over-represented in STAR-AQEM samples belonged to the groups Oligochaeta, Diptera, and Ephemeroptera. A total of five statistically significant indicator taxa were found for PERLA samples, whereas seven indicator taxa were found for STAR-AQEM samples.

Table 4.22 Indicator (over-represented) taxa in comparing the two sampling methods in Czech Republic streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
PERLA	Bivalvia	SPHAERIIDAE	Pisidium subtruncatum	43.3	0.04
PERLA	Diptera	EMPIDIDAE	Chelifera stigmatica	36.8	0.0101
PERLA	Ephemeroptera	LEPTOPHLEBIIDAE	Habroleptoides confusa	22.9	0.0007
PERLA	Ephemeroptera	LEPTOPHLEBIIDAE	Habrophlebia lauta	20.8	0.0405
PERLA	Oligochaeta	LUMBRICIDAE	Lumbricidae Gen. Sp.	16.7	0.0054
STAR-AQEM	Oligochaeta	NAIDIDAE	Nais alpina	48.7	0.0036
STAR-AQEM	Diptera	LIMONIIDAE	Paradelphomyia sp.	43.1	0.0243
STAR-AQEM	Ephemeroptera	LEPTOPHLEBIIDAE	Paraleptophlebia sp.	38.4	0.0101
STAR-AQEM	Diptera	PEDICIIDAE	Pedicia straminea	35.4	0.0492
STAR-AQEM	Diptera	PSYCHODIDAE	Pericoma sp.	24.7	0.0463
STAR-AQEM	Diptera	PSYCHODIDAE	Pneumia sp.	24.3	0.0132
STAR-AQEM	Diptera	SIMULIIDAE	Simulium latipes	18.3	0.0072

The taxa indicative (over-represented) in DSFI samples in Danish streams belonged to the groups of Coleoptera, Crustacea, Diptera, Ephemeroptera, Oligochaeta, Plecoptera, and Trichoptera (Table 4.23) whereas the samples over-represented in STAR-AQEM samples belonged to the groups Bivalvia, Coleoptera, Diptera, Ephemeroptera, Plecoptera, Trichoptera, and Turbellaria. A total of nine statistically significant indicator taxa were found for DSFI samples, whereas sixteen indicator taxa were found for STAR-AQEM samples.

Table 4.23 Indicator (over-represented) taxa in comparing the two sampling methods in Danish streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
DSFI	Coleoptera	ELMIDAE	Elmis aenea	72.7	0.0069
DSFI	Diptera	CHIRONOMIDAE	Prodiamesinae Gen.	63	0.0352
DSFI	Diptera	LIMONIIDAE	Hexatoma sp.	58.1	0.0008
DSFI	Crustacea	ASELLIDAE	Asellus aquaticus	57.8	0.0374
DSFI	Plecoptera	PERLODIDAE	Isoperla difformis	40.9	0.001
DSFI	Plecoptera	CHLOROPERLIDAE	Isoptena serricornis	36.4	0.0054
DSFI	Ephemeroptera	BAETIDAE	Nigrobaetis niger	36	0.046
DSFI	Trichoptera	LEPIDOSTOMATIDAE	Lepidostoma hirtum	29.3	0.0423
DSFI	Oligochaeta	PROPAPPIDAE	Propappus volki	22.7	0.0488
STAR-AQEM	Ephemeroptera	BAETIDAE	Baetis rhodani	79.9	0.0006
STAR-AQEM	Trichoptera	GOERIDAE	Silo pallipes	70.6	0.0001
STAR-AQEM	Diptera	CHIRONOMIDAE	Diamesinae Gen.	66.6	0.0001
STAR-AQEM	Diptera	LIMONIIDAE	Eloeophila sp.	63.9	0.002
STAR-AQEM	Coleoptera	ELMIDAE	Limnius volckmari	60	0.0021
STAR-AQEM	Turbellaria	DUGESIIDAE	Dugesia gonocephala	53.8	0.0285
STAR-AQEM	Ephemeroptera	EPHEMERIDAE	Ephemera danica	48.7	0.0212
STAR-AQEM	Trichoptera	GLOSSOSOMATIDAE	Agapetus ochripes	45.5	0.0007
STAR-AQEM	Bivalvia	SPHAERIIDAE	Sphaerium sp.	39.1	0.0029
STAR-AQEM	Coleoptera	HYDRAENIDAE	Hydraena gracilis	35.4	0.0093
STAR-AQEM	Coleoptera	GYRINIDAE	Orectochilus villosus	32.4	0.0418
STAR-AQEM	Plecoptera	TAENIOPTERYGIDAE	Brachyptera risi	31.8	0.0089
STAR-AQEM	Plecoptera	NEMOURIDAE	Nemoura flexuosa	27.3	0.0218
STAR-AQEM	Trichoptera	ODONTOCERIDAE	Odontocerum albicorne	27.3	0.0212
STAR-AQEM	Trichoptera	POLYCENTROPODIDAE	Polycentropus flavomaculatus	21.6	0.0473
STAR-AQEM	Plecoptera	NEMOURIDAE	Amphinemura standfussi	20.7	0.047

The taxa indicative (over-represented) in IBGN samples in French streams belonged to the groups of Bivalvia, Coelenterata, Coleoptera, Crustacea, Diptera, Ephemeroptera, Gastropoda, Hirudinea, Hydrachnidia, Oligochaeta, Trichoptera, and Turbellaria (Table 4.24). No taxa were over-represented in the STAR-AQEM samples compared to in the IBGN samples. A total of 29 significantly significant indicator taxa were found for the IBGN samples.

Table 4.24 Indicator (over-represented) taxa in comparing the two sampling methods in French streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
IBGN	Diptera	EMPIDIDAE	Empididae Gen. Sp.	85.1	0.0004
IBGN	Hydrachnidia	[Ph:Hydrachnidia]	Hydrachnidia Gen. sp.	82.5	0.0001
IBGN	Diptera	CERATOPOGONIDAE	Ceratopogonidae Gen. sp.	81.1	0.0001
IBGN	Diptera	LIMONIIDAE	Limoniidae Gen. sp.	81.0	0.0010
IBGN	Diptera	CHIRONOMIDAE	Chironomidae Gen. sp.	80.3	0.0001
IBGN	Oligochaeta	[Kl:Oligochaeta]	Oligochaeta Gen. sp.	79.3	0.0001
IBGN	Gastropoda	""ANCYLIDAE""	Ancylidae"" Gen. sp.	79.2	0.0001
IBGN	Diptera	PSYCHODIDAE	Psychodidae Gen. sp.	76.9	0.0001
IBGN	Trichoptera	HYDROPTILIDAE	Hydroptilidae Gen. sp.	76.4	0.0041
IBGN	Bivalvia	SPHAERIIDAE	Sphaeriidae Gen. sp.	74.6	0.0168
IBGN	Trichoptera	HYDROPSYCHIDAE	Hydropsychidae Gen. sp.	74.3	0.0310
IBGN	Ephemeroptera	BAETIDAE	Baetidae Gen. sp.	74.2	0.0128
IBGN	Gastropoda	LYMNAEIDAE	Lymnaeidae Gen. sp.	72.4	0.0156
IBGN	Coleoptera	ELMIDAE	Elmidae Gen. sp. Ad.	71.1	0.0250
IBGN	Coelenterata	HYDRIDAE	Hydridae Gen. sp.	68.9	0.0003
IBGN	Diptera	SIMULIIDAE	Simuliidae Gen. sp.	66.8	0.0089
IBGN	Coleoptera	ELMIDAE	Elmidae Gen. sp. Lv.	65.4	0.0207
IBGN	Trichoptera	RHYACOPHILIDAE	Rhyacophilidae Gen. sp.	62.5	0.0024
IBGN	Hirudinea	ERPOBDELLIDAE	Erpobdellidae Gen. sp.	61.1	0.0223
IBGN	Trichoptera	PSYCHOMYIIDAE	Psychomyiidae Gen. sp.	59.9	0.0192
IBGN	Ephemeroptera	LEPTOPHLEBIIDAE	Leptophlebiidae Gen. sp.	53.8	0.0493
IBGN	Gastropoda	BITHYNIIDAE	Bithyniidae Gen. sp.	52.6	0.0100
IBGN	Trichoptera	LIMNEPHILIDAE	Limnephilidae Gen. sp.	49.6	0.0094
IBGN	Turbellaria	PLANARIIDAE	Planariidae Gen. sp.	48.7	0.0012
IBGN	Coleoptera	HALIPLIDAE	Haliplidae Gen. sp. Lv.	46.7	0.0448
IBGN	Crustacea	ASELLIDAE	Asellidae Gen. sp.	46.5	0.0228
IBGN	Hirudinea	PISCICOLIDAE	Piscicolidae Gen. sp.	38.0	0.0494
IBGN	Diptera	ANTHOMYIIDAE	Anthomyiidae Gen. sp.	35.9	0.0057
IBGN	Coleoptera	SCIRTIDAE	Scirtidae Gen. sp. Lv.	27.6	0.0207

The taxa indicative (over-represented) in STAR-AQEM samples in German streams belonged to the groups of Coleoptera, Crustacea, Diptera, Ephemeroptera, Heteroptera, Oligochaeta, Plecoptera, and Trichoptera (Table 4.25). Only one taxa were over-represented in the RIVPACS samples compared to in the STAR-AQEM samples, i.e., *Gammarus pulex*. A total of 14 significantly significant indicator taxa were found for the STAR-AQEM samples.

Table 4.25 Indicator (over-represented) taxa in comparing the two sampling methods in German streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
RIVPACS	Crustacea	GAMMARIDAE	Gammarus pulex	33.4	0.0374
STAR	Diptera	CHIRONOMIDAE	Tanypodinae Gen. sp.	50.0	0.0128
STAR	Diptera	CHIRONOMIDAE	Polypedilum sp.	49.6	0.0039
STAR	Diptera	CHIRONOMIDAE	Tvetenia sp.	49.0	0.0028
STAR	Coleoptera	HYDRAENIDAE	Hydraena gracilis Ad.	43.3	0.0020
STAR	Coleoptera	ELMIDAE	Oulimnius tuberculatus Ad.	35.0	0.0001
STAR	Diptera	CHIRONOMIDAE	Prodiamesa olivacea	33.2	0.0100
STAR	Plecoptera	CHLOROPERLIDAE	Siphonoperla sp.	20.2	0.0344
STAR	Oligochaeta	LUMBRICULIDAE	Lumbriculus variegatus	17.3	0.0288
STAR	Heteroptera	CORIXIDAE	Corixidae Gen. sp.	17.1	0.0241
STAR	Coleoptera	DYTISCIDAE	Hydroporinae Gen. sp. Lv.	14.5	0.0123
STAR	Diptera	PSYCHODIDAE	Psychodidae Gen. sp.	14.4	0.0438
STAR	Ephemeroptera	CAENIDAE	Caenis horaria/robusta	13.8	0.0455
STAR	Trichoptera	HYDROPSYCHIDAE	Hydropsyche incognita	10.8	0.0281
STAR	Diptera	EMPIDIDAE	Hemerodromia sp.	10.0	0.0302

There were only one taxa indicative (over-represented) in STAR-AQEM samples in Greek streams and one taxa in RIVPACS samles. The STAR-AQEM samples were over-represented by Hydrachnidae, whereas RIVPACS samples were over-represented by Capniidae (Table 4.26).

Table 4.27 Indicator (over-represented) taxa in comparing the two sampling methods in Greek streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
RIVPACS	Plecoptera	CAPNIIDAE	Capniidae Gen. sp.	32.7	0.0467
STAR	Hydrachnidia	HYDRACHNIDAE	Hydrachnidae Gen. sp.	40.9	0.0050

All taxa indicative (over-represented) in Italian-CNR samples were taken using the STAR-AQEM sampling method. These taxa belonged to the groups Coleoptera, Crustacea, Diptera, Ephemeroptera, Gastropoda, Odonata, Oligochaeta, Plecoptera, and Trichoptera (Table 4.28).

Table 4.28 Indicator (over-represented) taxa in comparing the two sampling methods in Italian-CNR streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
STAR	Ephemeroptera	CAENIDAE	Caenis sp.	97.6	0.0001
STAR	Diptera	CHIRONOMIDAE	Chironomidae Gen. sp.	96.8	0.0001
STAR	Coleoptera	ELMIDAE	Elmidae Gen. sp. Lv.	92.5	0.0002
STAR	Plecoptera	LEUCTRIDAE	Leuctra sp.	90.4	0.0001
STAR	Trichoptera	HYDROPSYCHIDAE	Hydropsychidae Gen. sp.	80.8	0.0101
STAR	Odonata	GOMPHIDAE	Onychogomphus sp.	72.7	0.0025
STAR	Trichoptera	POLYCENTROPODIDAE	Polycentropodidae Gen. sp.	71.2	0.0018
STAR	Diptera	SIMULIIDAE	Simuliidae Gen. sp.	70.8	0.0273
STAR	Trichoptera	LEPTOCERIDAE	Leptoceridae Gen. sp.	67.3	0.0022
STAR	Gastropoda	HYDROBIIDAE	Bythinella sp.	61.6	0.0227
STAR	Trichoptera	BERAEIDAE	Beraeidae Gen. sp.	60.9	0.0034
STAR	Trichoptera	PHILOPOTAMIDAE	Philopotamidae Gen. sp.	60.7	0.0338
STAR	Diptera	ATHERICIDAE	Athericidae Gen. sp.	60.6	0.0053
STAR	Trichoptera	HYDROPTILIDAE	Hydroptilidae Gen. sp.	56.9	0.0027
STAR	Coleoptera	HYDROPHILIDAE	Hydrophilidae Gen. sp. Lv.	56.8	0.0009
STAR	Ephemeroptera	BAETIDAE	Centroptilum luteolum	48.9	0.0222
STAR	Gastropoda	PHYSIDAE	Physa sp.	47.0	0.0353
STAR	Diptera	LIMONIIDAE	Limoniidae Gen. sp.	46.4	0.0257
STAR	Crustacea	[Kl:Crustacea]	Ostracoda Gen. sp.	42.1	0.0030
STAR	Diptera	PSYCHODIDAE	Psychodidae Gen. sp.	42.1	0.0032
STAR	Diptera	EMPIDIDAE	Empididae Gen. sp.	41.0	0.0088
STAR	Coleoptera	DYTISCIDAE	Dytiscidae Gen. sp. Lv.	38.2	0.0291
STAR	Coleoptera	DRYOPIDAE	Dryopidae Gen. sp. Lv.	38.1	0.0205
STAR	Trichoptera	PSYCHOMYIIDAE	Psychomyiidae Gen. sp.	31.6	0.0186
STAR	Oligochaeta	LUMBRICULIDAE	Lumbriculidae Gen. sp.	31.4	0.0298
STAR	Ephemeroptera	BAETIDAE	Procloeon sp.	25.0	0.0441

All taxa indicative (over-represented) in Italian-Bolzano samples were taken using the STAR-AQEM sampling method. These taxa belonged to the groups Diptera, Ephemeroptera, Hydrachnidia, Oligochaeta, Plecoptera, Trichoptera, and Turbellaria (Table 4.29).

Table 4.29 Indicator (over-represented) taxa in comparing the two sampling methods in Italian-Bolzano streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
STAR	Ephemeroptera	BAETIDAE	Baetis sp.	95.2	0.0001
STAR	Diptera	CHIRONOMIDAE	Chironomidae Gen. sp.	80.0	0.0001
STAR	Turbellaria	PLANARIIDAE	Crenobia sp.	80.6	0.0001
STAR	Diptera	EMPIDIDAE	Empididae Gen. sp.	92.6	0.0001
STAR	Oligochaeta	ENCHYTRAEIDAE	Enchytraeidae Gen. sp.	93.6	0.0001
STAR	Hydrachnidia	[Ph:Hydrachnidia]	Hydrachnidia Gen. sp.	80.0	0.0001
STAR	Plecoptera	LEUCTRIDAE	Leuctra sp.	88.2	0.0002
STAR	Trichoptera	LIMNEPHILIDAE	Limnephilidae Gen. sp.	86.2	0.0003
STAR	Oligochaeta	LUMBRICULIDAE	Lumbriculidae Gen. sp.	83.1	0.0011
STAR	Plecoptera	NEMOURIDAE	Nemoura sp.	45.0	0.0020
STAR	Plecoptera	NEMOURIDAE	Protonemura sp.	80.8	0.0116
STAR	Diptera	PSYCHODIDAE	Psychodidae Gen. sp.	30.0	0.0188
STAR	Ephemeroptera	HEPTAGENIIDAE	Rhithrogena sp.	39.5	0.0188
STAR	Trichoptera	RHYACOPHILIDAE	Rhyacophilidae Gen. sp.	66.6	0.0415
STAR	Diptera	STRATIOMYIIDAE	Stratiomyiidae Gen. sp.	73.2	0.0491

Taxa indicative (over-represented) in Latvian samples belonged to the groups Diptera, Trichoptera, Hirudinea, and Oligochaeta. Taxa indicative (over-represented) in STAR-AQEM samples belonged to the groups Diptera, Ephemeroptera, Hirudinea, Odonata, Trichoptera, Oligochaeta, Bivalvia, and Porifera (Table 4.30). Comparisons in Latvian streams differ from the other comparisons, since the Latvian sampling method only includes some 60 indicator taxa, the STAR-AQEM samples has therefore been adjusted to these taxa in the comparisons.

Table 4.30 Indicator (over-represented) taxa in comparing the two sampling methods in Latvian streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
Latvian	Diptera	CHIRONOMIDAE	Chironomidae Gen. sp.	74.0	0.0001
Latvian	Trichoptera	LIMNEPHILIDAE	Limnephilus sp.	15.7	0.0119
Latvian	Hirudinea	PISCICOLIDAE	Piscicola geometra	43.6	0.0137
Latvian	Oligochaeta	TUBIFICIDAE	Tubificidae Gen. sp.	12.5	0.0254
STAR	Ephemeroptera	BAETIDAE	Baetidae Gen. sp.	32.1	0.0001
STAR	Diptera	CERATOPOGONIDAE	Ceratopogonidae Gen. sp.	77.5	0.0001
STAR	Ephemeroptera	EPHEMERIDAE	Ephemera sp.	77.9	0.0001
STAR	Hirudinea	ERPOBDELLIDAE	Erpobdella sp.	72.1	0.0005
STAR	Odonata	GOMPHIDAE	Gomphus sp.	63.7	0.0010
STAR	Ephemeroptera	LEPTOPHLEBIIDAE	Habrophlebia sp.	62.0	0.0091
STAR	Trichoptera	LEPTOCERIDAE	Mystacides sp.	21.8	0.0127
STAR	Trichoptera	RHYACOPHILIDAE	Rhyacophila sp.	57.1	0.0212
STAR	Oligochaeta	LUMBRICULIDAE	Lumbriculidae Gen. sp.	18.6	0.0305
STAR	Bivalvia	UNIONIDAE	Unionidae Gen. sp.	33.1	0.0360
STAR	Trichoptera	[Ord:Trichoptera]	Trichoptera Gen. sp.	34.2	0.0457
STAR	Porifera	SPONGILLIDAE	Spongillidae Gen. sp.	30.2	0.0492

Taxa indicative (over-represented) in Polish samples belonged to the groups Diptera, Megaloptera, Oligochaeta, Coleoptera, Trichoptera, and Heteroptera. Taxa indicative (over-represented) in STAR-AQEM samples belonged to the groups Diptera, Trichoptera, and Oligochaeta (Table 4.31). A large number of taxa were indicative of the STAR-AQEM sampling method, most of these were Diptera (Chironomidae) taxa.

Table 4.31 Indicator (over-represented) taxa in comparing the two sampling methods in Polish streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
Polish	Diptera	CHIRONOMIDAE	Chironomidae Gen. sp.	96.3	0.0001
Polish	Megalopter	SIALIDAE	Sialis lutaria	42.5	0.0175
Polish	Oligochaet	TUBIFICIDAE	Potamothrix hammoniensis	33.0	0.0122
Polish	Coleoptera	ELMIDAE	Elmis maugetii Lv.	32.4	0.0136
Polish	Oligochaet	TUBIFICIDAE	Limnodrilus udekemianus	28.9	0.0293
Polish	Diptera	CERATOPOGONIDAE	Probezzia seminigra	27.4	0.0139
Polish	Diptera	LIMONIIDAE	Pilaria sp.	24.6	0.0024
Polish	Coleoptera	CHRYSOMELIDAE	Donacia sp. Lv.	23.5	0.0003
Polish	Trichoptera	LIMNEPHILIDAE	Anabolia furcata	23.5	0.0003
Polish	Diptera	LIMONIIDAE	Eloeophila sp.	21.9	0.0263
Polish	Heteroptera	NOTONECTIDAE	Notonecta glauca ssp.	19.7	0.0481
Polish	Diptera	PEDICIIDAE	Dicranota sp.	15.9	0.0309
Polish	Oligochaet	TUBIFICIDAE	Rhyacodrilus coccineus	15.7	0.0047
Polish	Diptera	CERATOPOGONIDAE	Alluaudomyia sp.	15.7	0.0047
Polish	Oligochaet	TUBIFICIDAE	Psammoryctides barbatus	15.6	0.0052
Polish	Diptera	LIMONIIDAE	Phylidorea sp.	15.3	0.0134
Polish	Oligochaet	TUBIFICIDAE	Peloscolex sp.	11.8	0.0248
STAR	Diptera	CHIRONOMIDAE	Tanytarsini Gen. sp.	49.0	0.0001
STAR	Diptera	CHIRONOMIDAE	Procladius (Holotanypus) sp.	49.0	0.0001
STAR	Diptera	CHIRONOMIDAE	Orthocladiinae Gen. sp.	45.1	0.0001
STAR	Diptera	CHIRONOMIDAE	Polypedilum (Tripodura) sp.	43.1	0.0001
STAR	Diptera	CHIRONOMIDAE	Micropsectra sp.	39.2	0.0001
STAR	Diptera	CHIRONOMIDAE	Chironomus sp.	37.3	0.0001
STAR	Diptera	CHIRONOMIDAE	Microtendipes pedellus-Gr.	31.4	0.0001
STAR	Diptera	CHIRONOMIDAE	Thienemannimyia-Gr. Gen. sp.	31.4	0.0001
STAR	Diptera	CHIRONOMIDAE	Epoicocladius flavens	29.4	0.0001
STAR	Diptera	CHIRONOMIDAE	Paratendipes sp.	29.4	0.0001
STAR	Diptera	CHIRONOMIDAE	Chironomini Gen. sp.	27.5	0.0001
STAR	Diptera	CHIRONOMIDAE	Prodiamesa olivacea	27.5	0.0001
STAR	Diptera	CHIRONOMIDAE	Thienemanniella sp.	27.5	0.0001
STAR	Diptera	CHIRONOMIDAE	Cladotanytarsus sp.	25.5	0.0001
STAR	Diptera	CHIRONOMIDAE	Tvetenia sp.	25.5	0.0003
STAR	Diptera	CHIRONOMIDAE	Apsectrotanypus trifascipennis	23.5	0.0001
STAR	Diptera	CHIRONOMIDAE	Tanypodinae Gen. sp.	23.5	0.0004
STAR	Diptera	CHIRONOMIDAE	Rheotanytarsus sp.	21.6	0.0005
STAR	Diptera	CHIRONOMIDAE	Tanytarsus usmaensis	21.6	0.0005
STAR	Diptera	CHIRONOMIDAE	Cryptochironomus sp.	21.6	0.0008
STAR	Trichoptera	LIMNEPHILIDAE	Anabolia laevis	21.6	0.0009
STAR	Diptera	CHIRONOMIDAE	Clinotanypus nervosus	19.6	0.0016
STAR	Diptera	CHIRONOMIDAE	Tanytarsus brundini	19.6	0.0020
STAR	Diptera	CHIRONOMIDAE	Polypedilum pedestre-Agg.	19.6	0.0010
STAR	Diptera	CHIRONOMIDAE	Polypedilum (Pentapedilum) sp.	19.6	0.0012
STAR	Diptera	CHIRONOMIDAE	Parametriocnemus stylatus	17.6	0.0025
STAR	Diptera	CHIRONOMIDAE	Microtendipes sp.	17.6	0.0035

Table 4.31 (continued)

Method	Group	Family	Taxa	IV	p-value
STAR	Diptera	CHIRONOMIDAE	Cricotopus (Isocladius) sp.	17.6	0.0027
STAR	Diptera	CHIRONOMIDAE	Polypedilum (Polypedilum) sp.	17.6	0.0032
STAR	Diptera	CERATOPOGONIDAE	Bezzia sp.	16.9	0.0177
STAR	Diptera	CHIRONOMIDAE	Nanocladius sp.	15.7	0.0043
STAR	Diptera	CHIRONOMIDAE	Corynoneura sp.	15.7	0.0058
STAR	Diptera	CHIRONOMIDAE	Polypedilum nubeculosum-Gr.	15.7	0.0064
STAR	Diptera	CHIRONOMIDAE	Ablabesmyia monilis	15.7	0.0068
STAR	Diptera	CHIRONOMIDAE	Polypedilum exsectum	15.7	0.0053
STAR	Oligochaet	NAIDIDAE	Dero sp.	15.4	0.0174
STAR	Diptera	CHIRONOMIDAE	Potthastia longimana	13.7	0.0120
STAR	Diptera	CHIRONOMIDAE	Dicrotendipes sp.	13.7	0.0129
STAR	Trichoptera	SERICOSTOMATIDAE	Notidobia ciliaris	13.2	0.0189
STAR	Diptera	CHIRONOMIDAE	Synorthocladius semivirens	11.8	0.0247

Only three taxa were indicative (over-represented) in Portuguese samples. They were all related to the Portuguese sampling method, thus there were no taxa indicative of the STAR-AQEM sampling method. The indicator taxa belonged to the groups Diptera, Ephemeroptera, and Coleoptera (Table 4.32).

Table 4.32 Indicator (over-represented) taxa in comparing the two sampling methods in Portuguese streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
Portuguese	Diptera	PEDICIIDAE	Dicranota sp.	58.9	0.0260
Portuguese	Ephemeroptera	BAETIDAE	Baetis scambus	30.0	0.0205
Portuguese	Coleoptera	HYDROPHILIDAE	Berosus sp. Lv.	25.0	0.0468

Six taxa were indicative (over-represented) in Slovak republic samples, four of these were over-represented in PERLA samples and two in STAR-AQEM samples. The indicator taxa indicative of PERLA samples belonged to the groups Coleoptera, Diptera, and Bivalvia, whereas both the taxa indicative of STAR-AQEM samples belonged to the group Diptera.(Table 4.33).

Table 4.33 Indicator (over-represented) taxa in comparing the two sampling methods in Slovak republic streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
PERLA	Coleoptera	ELMIDAE	Elmis maugetii	76.2	0.0291
PERLA	Diptera	CHIRONOMIDAE	Corynoneura lobata	63.5	0.0184
PERLA	Bivalvia	SPHAERIIDAE	Pisidium sp.	57.3	0.0127
PERLA	Diptera	CHIRONOMIDAE	Rheocricotopus fuscipes	51.0	0.0446
STAR	Diptera	CERATOPOGONID	Bezzia sp.	78.2	0.0040
STAR	Diptera	CHIRONOMIDAE	Parametriocnemus stylatus	41.7	0.0376

Seven taxa were indicative (over-represented) in Swedish samples, all of these were over-represented in STAR-AQEM samples. The indicator taxa indicative of the STAR-AQEM samples belonged to the group Diptera, Oligochaeta, Hirudinea, and nematoda.(Table 4.34).

Table 4.34 Indicator (over-represented) taxa in comparing the two sampling methods in Swedish streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
STAR	Oligochaeta	NAIDIDAE	Nais sp.	46.3	0.0003
STAR	Diptera	CHIRONOMIDAE	Eukiefferiella sp.	37.8	0.0116
STAR	Diptera	PSYCHODIDAE	Psychodidae Gen. sp.	20.3	0.0414
STAR	Diptera	CHIRONOMIDAE	Potthastia gaedii-Gr.	14.8	0.0071
STAR	Oligochaeta	LUMBRICIDAE	Lumbricidae Gen. sp.	14.5	0.0316
STAR	Hirudinea	GLOSSIPHONIIDA	Glossiphonia complanata	13.0	0.0120
STAR	Nematoda	[Kl:Nematoda]	Nematoda Gen. sp.	11.6	0.0477

Twelve taxa were indicative (over-represented) in UK samples, all of these were over-represented in STAR-AQEM samples and none in RIVPACS samples. The indicator taxa indicative of STAR-AQEM samples belonged to the groups Diptera and Oligochaeta (Table 4.35).

Table 4.35 Indicator (over-represented) taxa in comparing the two sampling methods in UK streams. IV = Indicator Value (higher values indicate a stronger affinity to a group), p-value = statistical significance of the relationship between method and taxa.

Method	Group	Family	Taxa	IV	p-value
STAR	Diptera	CHIRONOMIDAE	Eukiefferiella claripennis	56.9	0.0012
STAR	Diptera	CHIRONOMIDAE	Polypedilum sp.	51.8	0.0254
STAR	Oligochaeta	TUBIFICIDAE	Rhyacodrilus coccineus	49.2	0.0309
STAR	Diptera	CHIRONOMIDAE	Thienemanniella sp.	49.2	0.0404
STAR	Oligochaeta	TUBIFICIDAE	Tubifex ignotus	49.1	0.0424
STAR	Diptera	CHIRONOMIDAE	Rheocricotopus sp.	44.2	0.0267
STAR	Diptera	CERATOPOGONIDAE	Ceratopogonidae Gen. sp.	34.5	0.0010
STAR	Oligochaeta	TUBIFICIDAE	Limnodrilus hoffmeisteri	23.8	0.0123
STAR	Oligochaeta	LUMBRICIDAE	Lumbricidae Gen. sp.	20.5	0.0197
STAR	Oligochaeta	LUMBRICULIDAE	Lumbriculidae Gen. sp.	19.9	0.0216
STAR	Oligochaeta	NAIDIDAE	Stylaria lacustris	16.7	0.0403
STAR	Diptera	CHIRONOMIDAE	Nanocladius rectinervis	11.5	0.0238

4.6 VARIATION EXPLAINED BY STREAM TYPE, SEASON, AMOUNT OF STRESS AND MACROINVERTEBRATE SAMPLING METHOD

Six countries sampled in at least two stream types and the amount of variation in macroinvertebrate community composition explained by type differed between 16.0% in the Czech Republic and 67.9% of the total explained variation in Greece (Table 4.36). Two different seasons were sampled in all countries and season explained between 11.6% of the total explained variation (in Greece) and 56.0% of the total explained variation in Latvia. The pre-defined stress gradient (here divided into sites pre-defined as having a high or good ecological status versus those pre-defined as having a moderate, poor or bad ecological status) explained between 15.3% (in Greece) and 55.3% of the total explained variation in France. Stream type, differences between seasons, and the pre-defined stress gradient were always statistically significant explanatory variables. The difference in sampling method on the other hand, generally only explained a smaller part of the total explained variation (except in a few cases). Sampling methods was a statistically significant explanatory variable and explained a relatively high amount of the total explained variation in Poland, Latvia, and Italy (CNR) (Table 4.36).

Table 4.36 Variation explained by stream type, season, amount of stress, and sampling method (out of total explained variation). Values in italics are statistically significant with p <0.05.

Country	% Type	% Season	% Stress	% Method	% TEV
Austria	*33.3%*	*27.6%*	*29.1%*	*11.6%*	15.1%
Czech Republic	*16.0%*	*40.0%*	*38.8%*	*8.2%*	19.7%
Denmark	--	*45.8%*	*48.6%*	8.6%	14.9%
France	--	*38.6%*	*55.3%*	5.9%	25.8%
Germany	*60.6%*	*17.5%*	*16.0%*	5.9%	15.8%
Greece	*67.9%*	*11.6%*	*15.3%*	9.7%	17.7%
Italy-CNR	--	*53.6%*	*31.5%*	*19.2%*	24.5%
Italy-Bolzano	--	*51.2%*	*38.5%*	9.0%	20.5%
Latvia	--	*56.0%*	*19.9%*	*25.2%*	16.0%
Poland	--	19.7%	*30.0%*	*58.2%*	6.8%
Portugal	--	*46,0%*	*41,7%*	15,5%	18.4%
Slovak Republik	--	*38.3%*	*39.6%*	20.6%	26.0%
Sweden	*30.3%*	*33.1%*	*32.9%*	6.6%	10.3%
United Kingdom	*25.4%*	*27.7%*	*40.4%*	8.8%	12.6%

There were generally no joint explanatory effects of sampling method with either stream type, season or amount of stress (Table 4.37). Only in the Polish samples were seasons related to sampling method and in Portugal was there a relation between stress and sampling method. When looking at these comparisons one must of course take into account the fact that e.g. differences in how much variation is explained by type in relation to the other factors in dependent on how large differences there are in types analysed. In Germany for example, there are a large difference in taxonomic composition among the three stream types, and therefore season and stress comes out as less important (when analysing season as an explanatory variable e.g. within stream type D03 in Germany, this variable explains 24.7% whereas stress within the stream type explains 63.8% and sampling method 12.8%).

Table 4.37 Variation explained jointly by choice of sampling method and stream type, season and amount of stress (out of total explained variation)

Country	% Type	% Season	% Stress
Austria	0.2%	0.4%	0.9%
Czech Republic	0.0%	1.1%	1.1%
Denmark	--	0.0%	0.0%
France	--	0.5%	0.0%
Germany	0.1%	0.3%	0.1%
Greece	1.6%	0.1%	0.1%
Italy-CNR	--	2.1%	0.4%
Italy-Bolzano	--	0.0%	0.0%
Latvia	--	0.7%	0.4%
Poland	--	6.5%	1.6%
Portugal	--	0.3%	5.6%
Slovakia	--	0.0%	0.0%
Sweden	0.0%	0.5%	0.0%
United Kingdom	0.0%	0.0%	0.3%

4.7 ORDINATIONS SHOWING COMPARISONS OF SAMPLING METHODS

Another way to evaluate the importance of stream types, season, amount of stress and sampling method is to run an unconstrained ordination (in this case non-metric multidimensional scaling) and then statistically test differences in ordination scores among samples taken e.g. in different stream types, seasons etc.Here differences in stream types in macroinvertrebrate community composition was found in Austria, Germany, Greece, Sweden, and UK (Table 4.38). Statistical differences among seasons were found in Austria, Czech Republic, Italy (CNR), Latvia, Slovak Republic, Sweden, and UK. Differences among sites pre-classified as having different amount of human perturbation (stress) were found in Austria, Czech Republic, France, Portugal, Slovak Republic, and UK. Finally differences among sampling methods were found in France, Italy (CNR and Bolzano), Latvia, and Poland.

Table 4.38 Comparisons of NMS ordination axes in relation to stream type, season, amount of stress, and macroinvertebrate sampling method using two sample *t*-tets (p-values).

Country	Type	Season	Stress	Method
Austria Axis 1	< 0.05	< 0.0001	ns	ns
Austria Axis 2	ns	ns	< 0.05	ns
Czech Republic Axis 1	ns	< 0.0001	< 0.05	ns
Czech Republic Axis 2	ns	< 0.005	< 0.0001	ns
Denmark Axis 1	--	ns	ns	ns
Denmark Axis 2	--	ns	ns	ns
France Axis 1	--	ns	ns	< 0.0001
France Axis 2	--	ns	<0.001	ns
Germany Axis 1	< 0.0001	ns	ns	ns
Germany Axis 2	< 0.0001	ns	ns	ns
Greece Axis 1	< 0.005	ns	ns	ns
Greece Axis 2	< 0.005	ns	ns	ns
Italy-CNR Axis 1	--	ns	ns	< 0.0001
Italy-CNR Axis 2	--	< 0.05	ns	ns
Italy-Bolzano Axis 1	--	ns	ns	< 0.0001
Italy-Bolzano Axis 2	--	ns	ns	< 0.0001
Latvia Axis1	--	ns	ns	< 0.005
Latvia Axis 2	--	< 0.001	ns	< 0.001
Poland Axis 1	--	ns	ns	< 0.001
Poland Axis 2	--	ns	ns	< 0.001
Portugal Axis 1	--	ns	< 0.0005	ns
Portugal Axis 2	--	ns	ns	ns
Slovak Republic Axis 1	--	ns	< 0.05	ns
Slovak Republic Axis 2	--	< 0.001	ns	ns
Sweden Axis 1	< 0.01	< 0.05	ns	ns
Sweden Axis 2	ns	< 0.0001	ns	ns
United Kingdom Axis 1	< 0.005	< 0.001	< 0.001	ns
United Kingdom Axis 2	ns	ns	< 0.001	ns

Austria

There was no difference in macroinvertebrate community structure among samples taken using the STAR-AQEM versus the RIVPACS approach in Austrian streams (Figure 4.13). None of the two sample *t*-test of ordination scores along any of the axes showed a significant difference in ordination scores. Axis one (T = -0.01, p > 0.05, df = 76), axis two (T = 1.76, p > 0.05, df = 75).

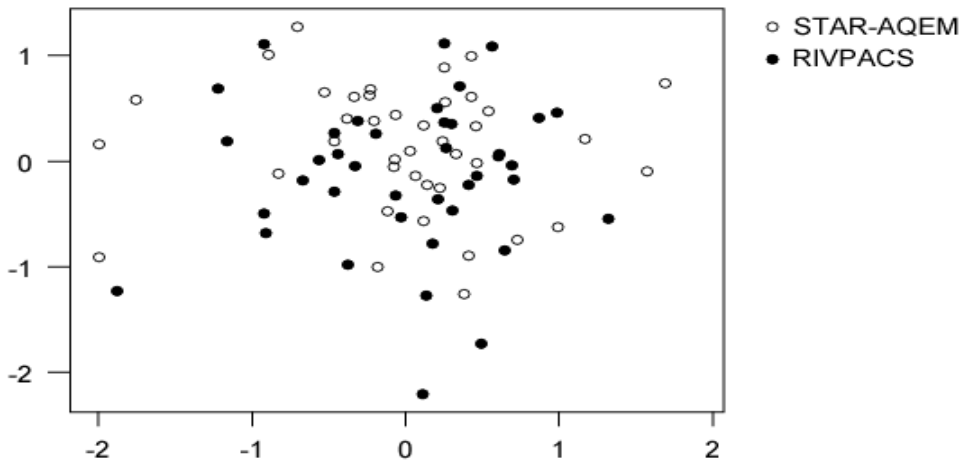


Figure 4.13 Austrian macroinvertebrate data divided into samples taken using the STAR-AQEM approach = open circles and the RIVPACS approach = closed circles.

Czech Republic

There was no difference in macroinvertebrate community structure among samples taken using the STAR-AQEM versus the PERLA approach in Czech Republic streams. None of the two sample *t*-test of ordination scores along any of the axes showed a significant difference in ordination scores (Figure 4.14). Axis one (T = 0.20, p > 0.05, df = 93), axis two (T = -0.92, p > 0.05, df = 93).

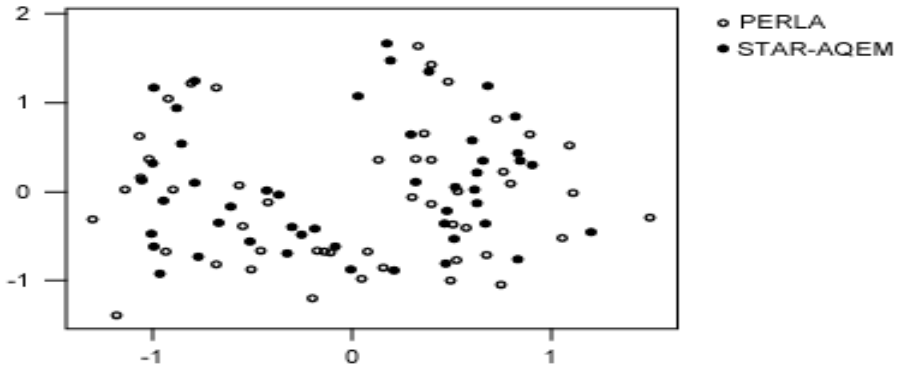


Figure 4.14 Czech Republic macroinvertebrate data divided into samples taken using the STAR-AQEM approach = closed circles and the PERLA approach = open circles. The final stress of the NMS was 58.97 for 2 axes.

Denmark

There was no difference in macroinvertebrate community structure among samples taken using the STAR-AQEM versus the DSFI approach in Danish streams. None of the two sample *t*-test of ordination scores along any of the axes showed a significant difference in ordination scores (Figure 4.15). Axis one (T = 0.05, p > 0.05, df = 37), axis two (T = 1.28, p > 0.05, df = 41).

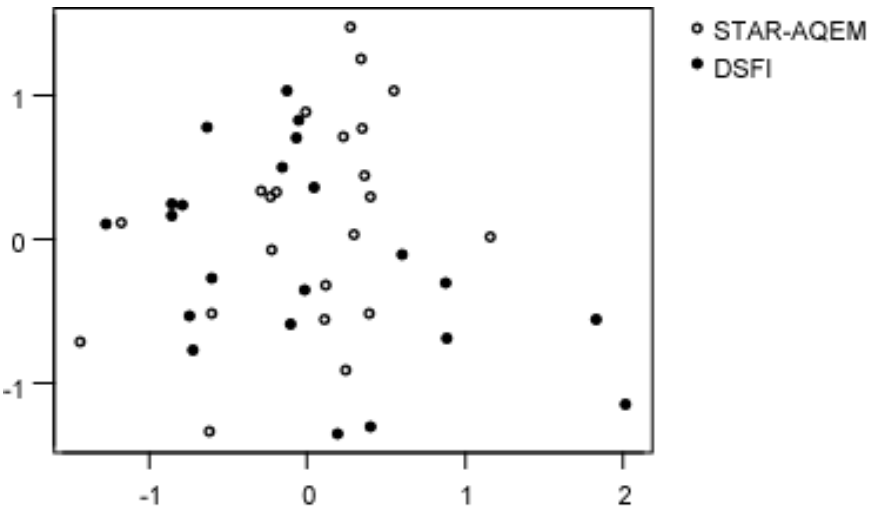


Figure 4.15 Danish macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the DSFI method = closed circles.

France

There were clear differences in community structure among samples taken using the STAR-AQEM versus the IBGN approach in French streams. The two sample *t*-test of ordination scores along the first axes showed a significant difference in ordination scores (Figure 4.16). Axis one (T = -6.63, p < 0.0001, df = 37), axis two (T = 1.28, p > 0.05, df = 45).

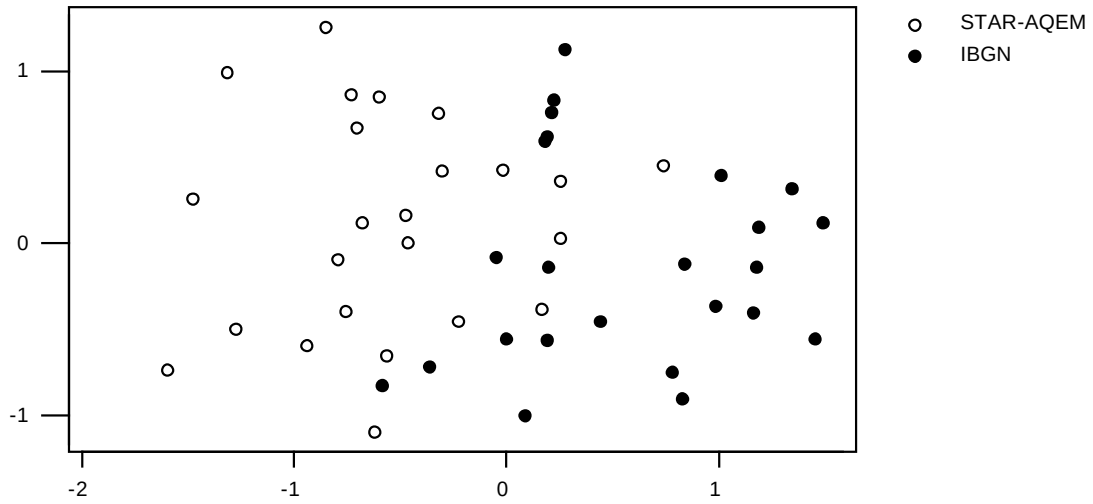


Figure 4.16 French macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the IBGN method = closed circles.

Germany

There were no differences in community structure among samples taken using the STAR-AQEM versus the RIVPACS approach in German streams. The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores (Figure 4.17). Axis one (T = -0.37, p > 0.05, df = 117), axis two (T = -1.52, p > 0.05, df = 117).

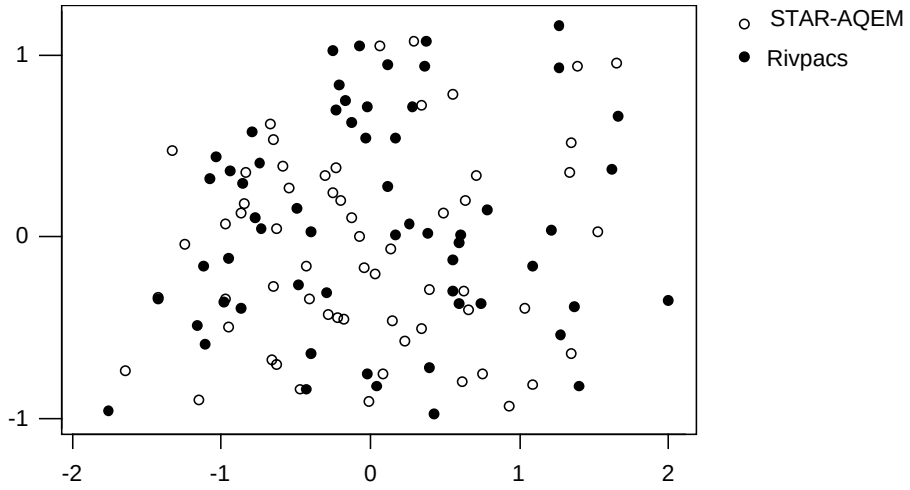


Figure 4.17 German macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the RIVPACS method = closed circles.

Greece

There were no differences in community structure among samples taken using the STAR-AQEM versus the RIVPACS approach in Greek streams (figure 4.18). The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores. Axis one ($T = -0.55$, $p > 0.05$, $df = 87$), axis two ($T = 1.12$, $p > 0.05$, $df = 87$).

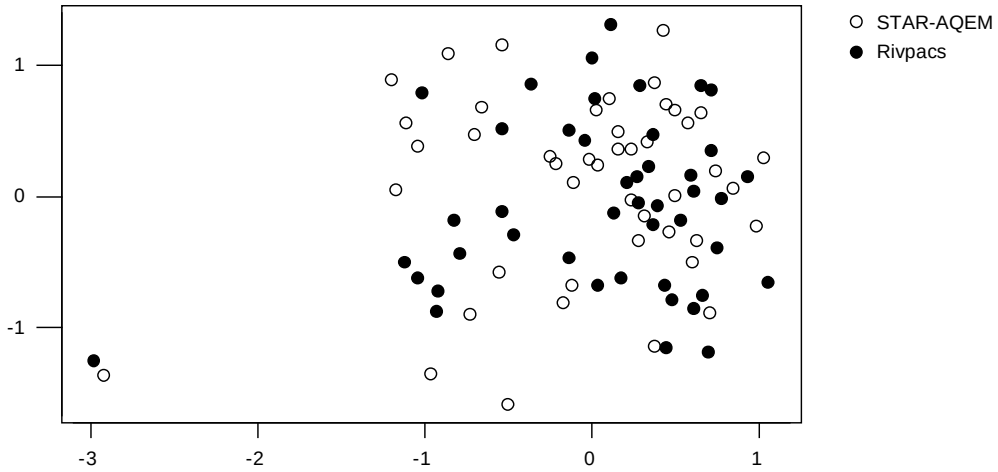


Figure 4.18 Greece macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the RIVPACS method = closed circles.

Italy-CNR

There was a clear difference in community structure among samples taken using the STAR-AQEM versus the IBE approach in Italian-CNR streams. The two sample *t*-test of ordination scores along the first axes showed a significant difference in ordination scores, whereas no such difference was seen along axis two (the two sampling methods still divided the all samples into two clear groups) (Figure 4.19). Axis one ($T = -11.01$, $p < 0.001$, $df = 35$), axis two ($T = -0.19$, $p > 0.05$, $df = 35$).

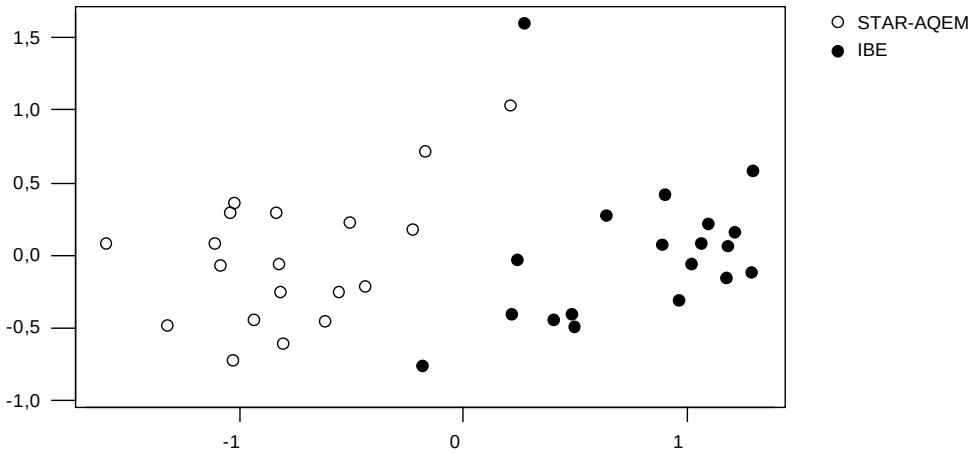


Figure 4.19 Italian-CNR macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the IBE method = closed circles.

Italy-Bolzano

There was a clear difference in community structure among samples taken using the STAR-AQEM versus the IBE approach in Italian-Bolzano streams. The two sample *t*-test of ordination scores along the first axes showed a significant difference in ordination scores, which was also the case along the second axis of the NMS (Figure 4.20). Axis one ($T = -4.63$, $p < 0.001$, $df = 33$), axis two ($T = 7.70$, $p < 0.001$, $df = 33$).

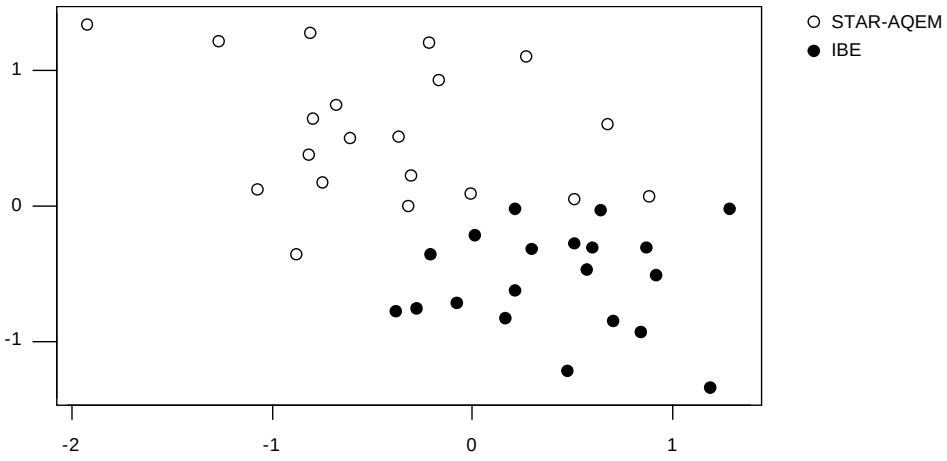


Figure 4.20 Italian-Bolzano macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the IBE method = closed circles.

Latvia

There was a clear difference in community structure among samples taken using the STAR-AQEM versus the Latvian sampling approach. The two sample *t*-test of ordination scores along the first axes showed a significant difference in ordination scores, which was also the case along the second axis of the NMS (Figure 4.21). Axis one ($T = 3.32$, $p < 0.005$, $df = 86$), axis two ($T = -6.85$, $p < 0.0001$, $df = 86$). No such differences were found for the sites pre-classfied as having a high or good ecological status versus a moderate, poor or bad ecological status. Whereas for Latvian samples, the second ordination axis were related to the sampled season.

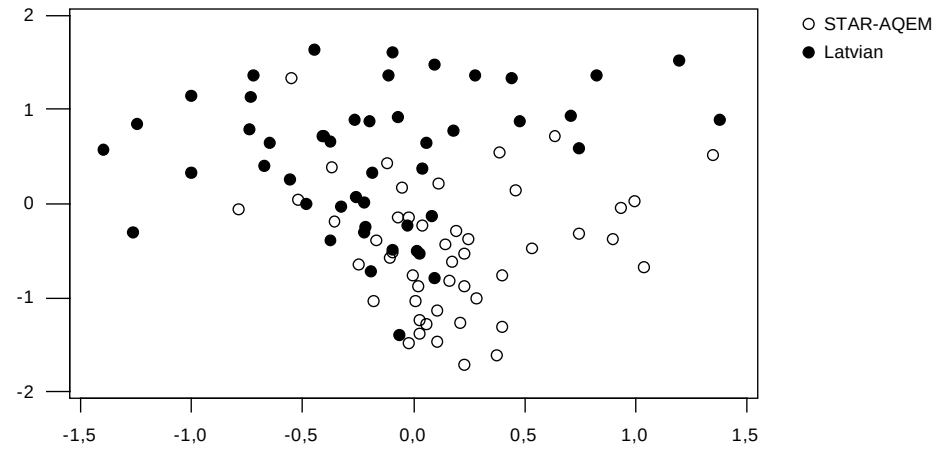


Figure 4.21 Latvian macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the Latvian method = closed circles.

Poland

There was a clear difference in community structure among samples taken using the STAR-AQEM versus the Polish sampling approach. The two sample *t*-test of ordination scores along the first axes showed a significant difference in ordination scores, which was also the case along the second axis of the NMS (Figure 4.22). Axis one ($T = 6.58$, $p < 0.0001$, $df = 98$), axis two ($T = 5.51$, $p < 0.0001$, $df = 84$). No such differences were found for the different seasons or sites pre-classfied as having a high or good ecological status versus a moderate, poor or bad ecological status.

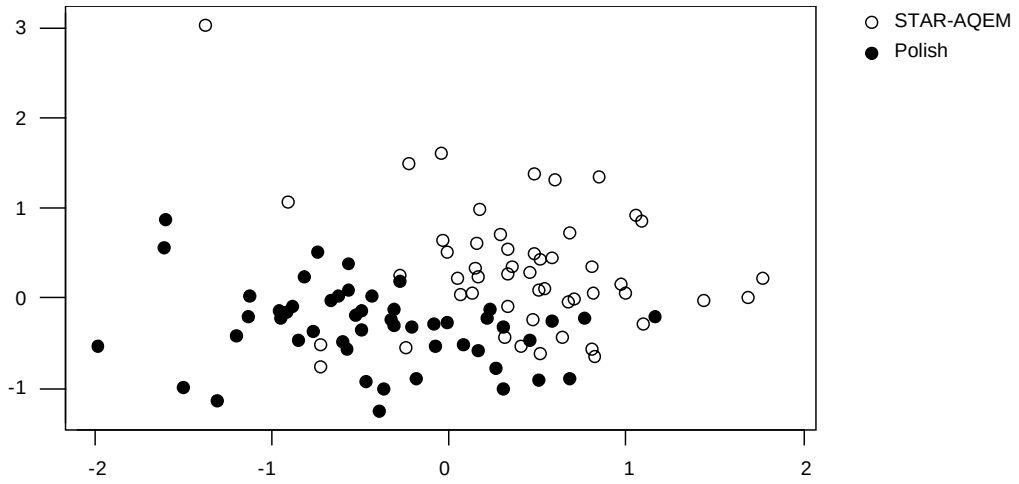


Figure 4.22 Polish macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the Polish method = closed circles.

Portugal

There was a no difference in community structure among samples taken using the STAR-AQEM versus the Portuguese sampling approach. The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores, which was also the case along the second axis of the NMS. Axis one ($T = -1.03$, $p > 0.05$, $df = 36$), axis two ($T = 1.00$, $p > 0.05$, $df = 19$). There were a difference along the first NMS ordination axis for sites pre-classfied as having a high or good ecological status versus a moderate, poor or bad ecological status with ($T = 3.85$, $p = 0.0005$, $df = 37$).

Slovak Republic

There was a no difference in community structure among samples taken using the STAR-AQEM versus the PERLA sampling approach (Figure 4.23). The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores, which was also the case along the second axis of the NMS. Axis one ($T = -0.62$, $p > 0.05$, $df = 18$), axis two ($T = -0.10$, $p > 0.05$, $df = 3$). There were a difference along the first NMS ordination axis for sites pre-classfied as having a high or good ecological status versus a moderate, poor or bad ecological status with ($T = 4.42$, $p < 0.05$, $df = 4$) and there were also a significant difference along the second NMS axis in ordination scores when comparing sites sampled in spring versus autumn ($T = -6.86$, $p < 0.0001$, $df = 21$).

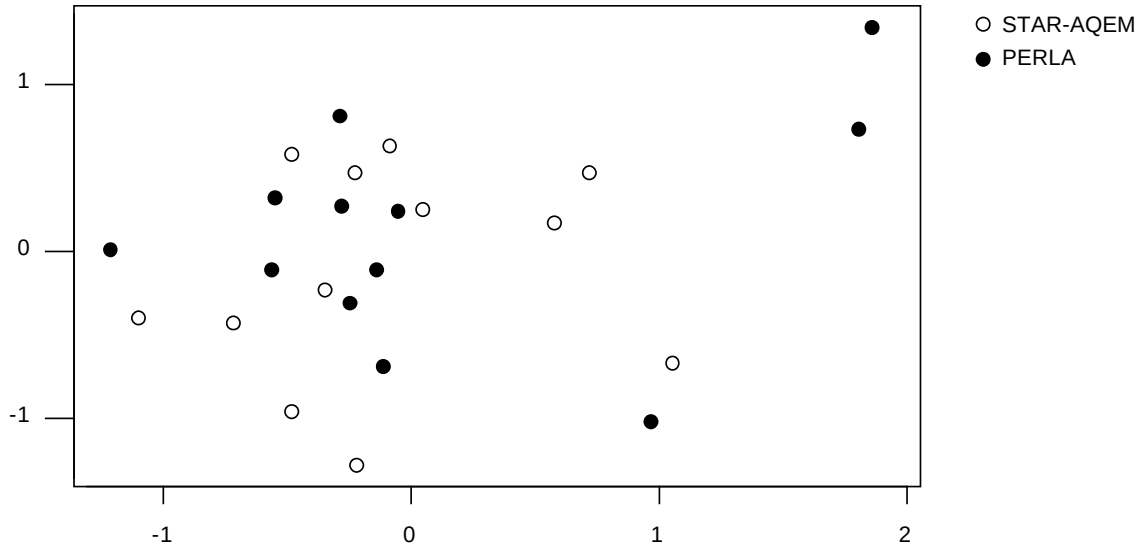


Figure 4.23 Slovak Republic macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the PERLA method = closed circles.

Sweden

There was a no difference in community structure among samples taken using the STAR-AQEM versus the Swedish standards sampling approach (Figure 4.24). The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores, which was also the case along the second axis of the NMS. Axis one ($T = 1.04$, $p > 0.05$, $df = 105$), axis two ($T = -0.83$, $p > 0.05$, $df = 105$). There were a difference along the first NMS ordination axis for samples taken in the two types sampled in Sweden ($T = 2.64$, $p < 0.05$, $df = 97$) and along both the first and second axis for samples taken in the two seasons (spring and autumn) sampled in Sweden axis one ($T = 2.25$, $p < 0.05$, $df = 105$), axis two ($T = -9.38$, $p < 0.0001$, $df = 105$).

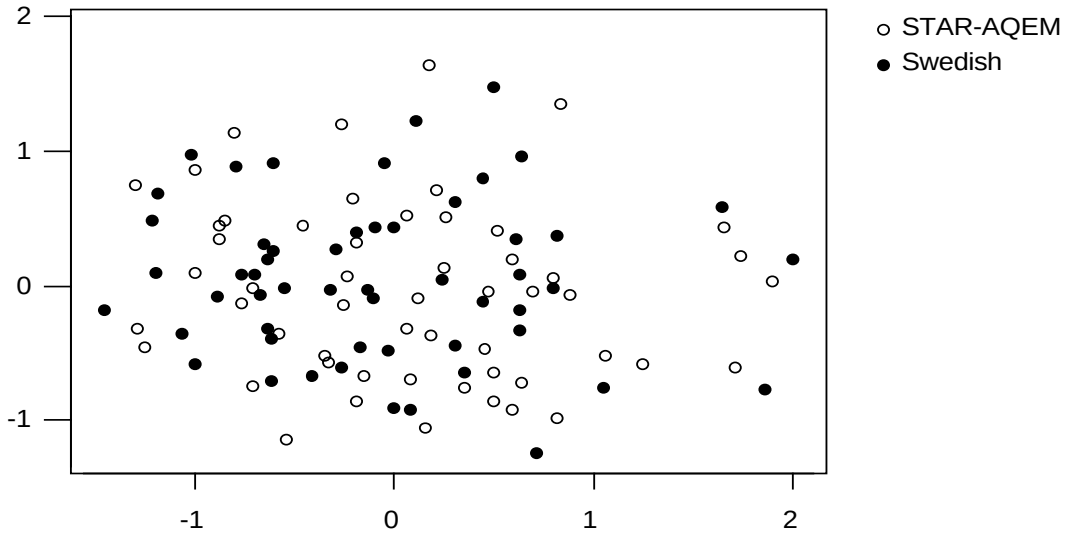


Figure 4.24 SwedishSwedish macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the PERLA method = closed circles.

United Kingdom

There was a no difference in community structure among samples taken using the STAR-AQEM versus the RIVPACS sampling approach (Figure 4.25). The two sample *t*-test of ordination scores along the first axes showed no significant difference in ordination scores, which was also the case along the second axis of the NMS. Axis one ($T = -1.87$, $p > 0.05$, $df = 97$), axis two ($T = -0.42$, $p > 0.05$, $df = 97$). There was a difference along the first and second NMS ordination axis for sites pre-classfied as having a high or good ecological status versus a moderate, poor or bad ecological status with axis one ($T = 5.01$, $p < 0.0001$, $df = 83$) and axis 2 ($T = -7.91$, $p < 0.0001$, $df = 84$). There were also a significant difference along the first NMS axis in ordination scores when comparing sites sampled in spring versus autumn and sites sampled in the two stream types; seasons ($T = -4.34$, $p < 0.0001$, $df = 94$), types ($T = 3.19$, $p < 0.005$, $df = 96$),

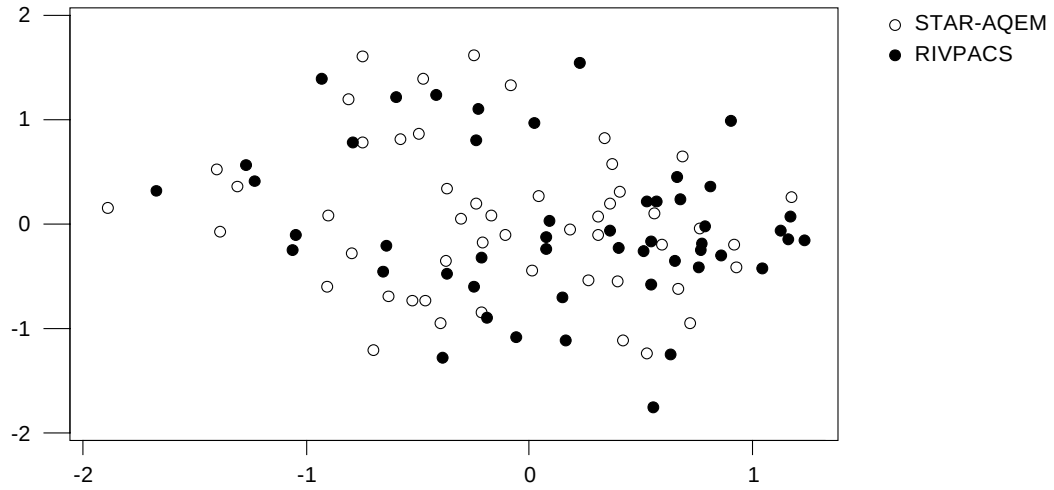


Figure 4.25 UK macroinvertebrate data divided into samples taken using the STAR-AQEM sampling method = open circles versus the RIVPACS method = closed circles.

4.8 TYPOLOGICAL/ENVIRONMENTAL DIFFERENCES

To compare if the different macroinvertebrate sampling methods used in the STAR project covered different habitats, a standardised Prinicipal Components Analysis (PCA) of arcsine transformed percentage data from both mineral and biological substratum types was performed. The two first axes of the PCA explained 18.3% of the total variance in the substratum data.

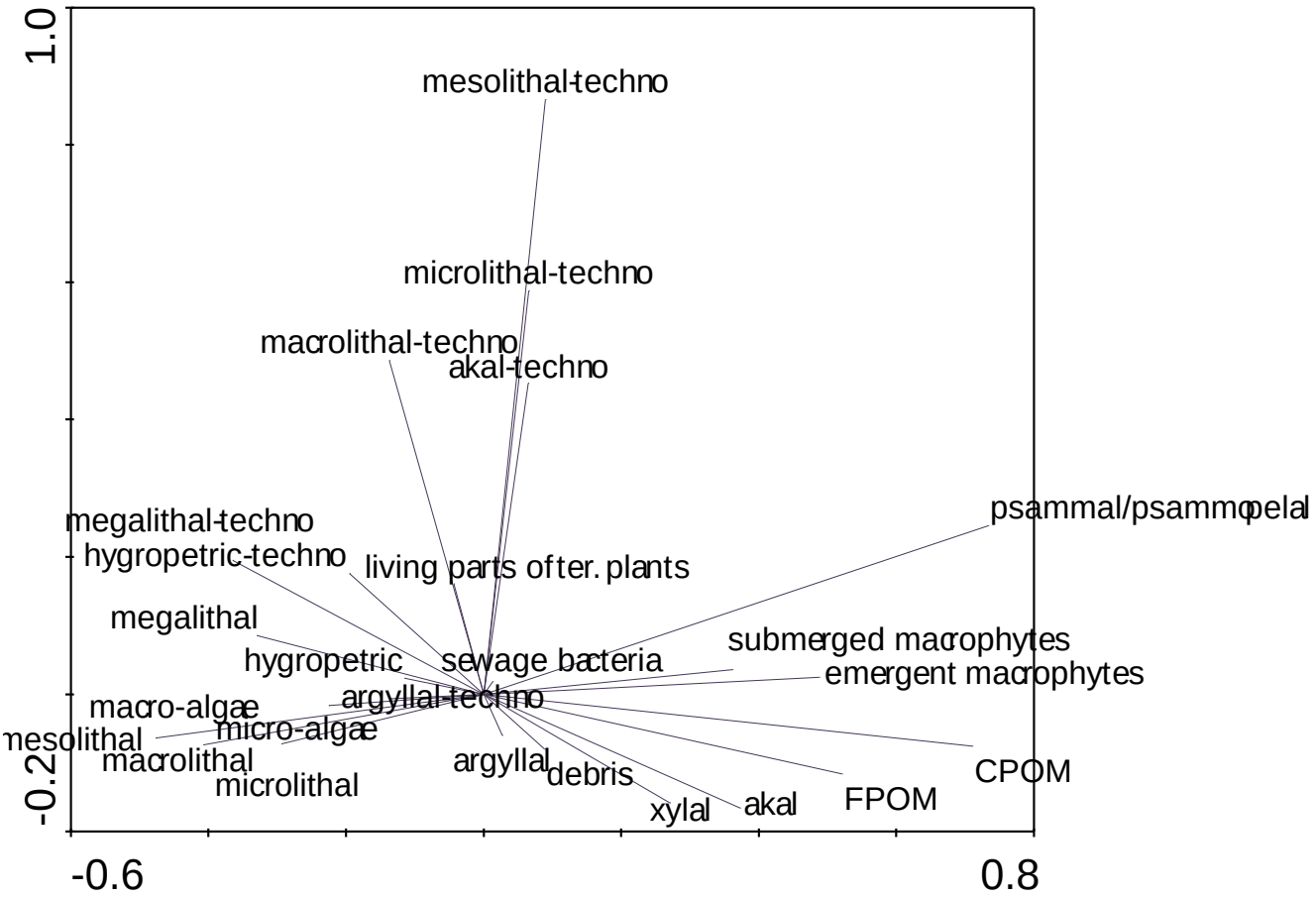


Figure 4.26 Substratum data from 1547 macroinvertebrate samples (both STAR-AQEM and national/RIVPACS samples including replicates) from 12 STAR countries (all except Portugal). The standardised Principal Components Analysis shows the 23 mineral and biological components for all sites.

There was a clear first axis going from coarse substratum types and micro and macro algae to the left in the ordination (e.g. mesolithal with a substratum size of 6-20 cm, macrolithal 20-40 cm, and microlithal 2-6 cm), whereas to the right along the first axis substratum types such as psammal/psammopelal (i.e. very fine material), Fine Particulate Organic Matter (FPOM), Course Particluate Organic Matter (CPOM), and akal (with a particle size of 0.2-2 cm), (Figure 4.26). The second axis divided a few German sites with a non-natural substratum (technolithal) of different kinds from the rest of the sites sampled within the STAR project (Figure 4.26 and Figure 4.27).

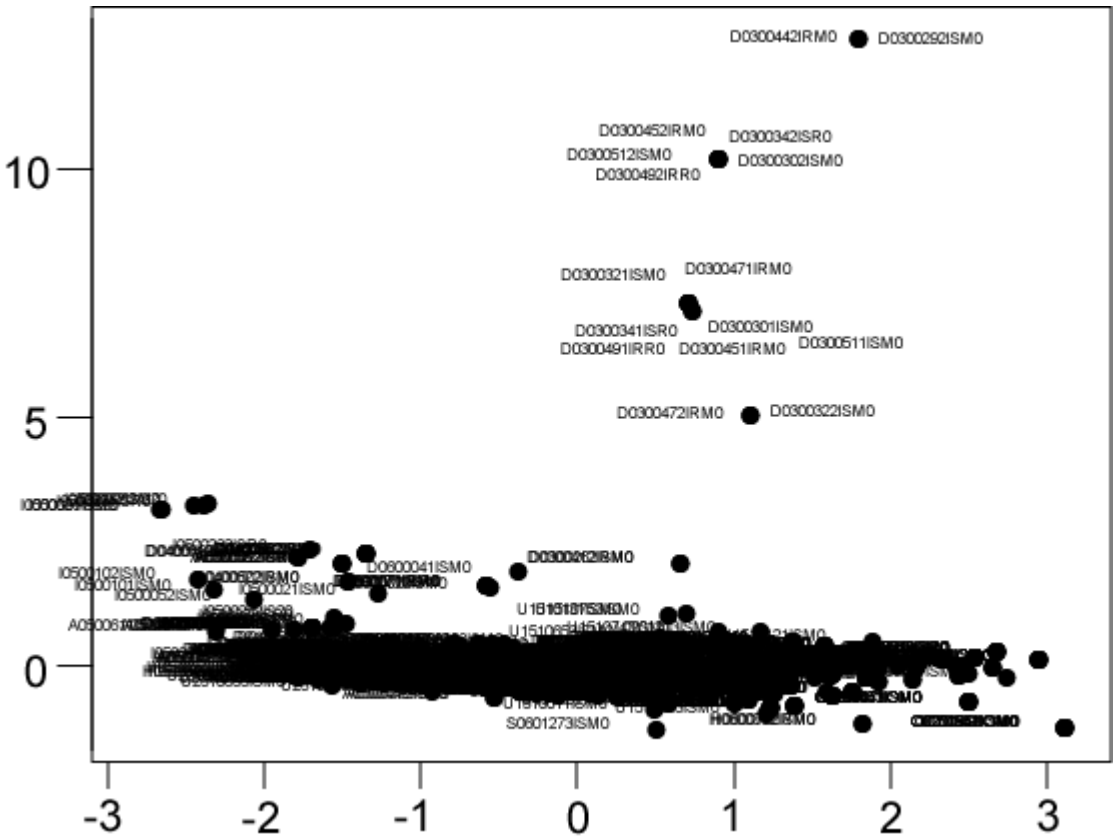


Figure 4.27 Substratum data from 1547 macroinvertebrate samples (both STAR-AQEM and national/RIVPACS samples including replicates) from 12 STAR countries (all except Portugal). The outlier sites along axis 2 are all from German samples (coding starts with a D).

Since the first axis of the PCA clearly captured the main difference in substratum type (from course material to the left to fine material to the right in the ordination), this gradient was used to compare if the different macroinvertebrate sampling methods used in the STAR project captured different substratum types.

Table 4.39 Comparison of substratum composition (both mineral and biological) for each country in the STAR project involved in sampling of core and additional stream types. Comparisons of the different stream types, seasons, sampling methods and levels of ecological stress (pre-defined) within each country using *t*-tests, except for comparisons of substratum composition in German streams, where test of differences of the 3 types was done using a one-way ANOVA.

Country	Types	Seasons	Methods	Stress
Austria	ns	ns	ns	ns
Czech Republic	$p < 0.0001$	ns	ns	< 0.05
Denmark	--	$p < 0.05$	ns	ns
France	--	ns	ns	< 0.005
Germany	$p < 0.001$	ns	ns	< 0.001
Greece	$p < 0.001$	ns	ns	ns
Italy-CNR	--	ns	< 0.001	ns
Italy-Bolzano	--	ns	< 0.001	ns
Latvia	--	ns	< 0.001	ns
Poland	--	ns	ns	< 0.001
Portugal*	--	ns	ns	< 0.05
Slovakia	--	ns	ns	ns
Sweden	$p < 0.01$	< 0.05	ns	ns
United Kingdom	< 0.001	ns	< 0.05	ns

* the Portuguese data were analysed separately from the remainder of the substratum dataset

There were no differences in substratum composition for either types, sampling seasons, sampling methods or ecological quality in the Austrian streams. In the Czech Republic streams, such differences were found both among the two stream types sampled (C04 and C05) (Figure 4.28) and among samples taken at sites with different ecological stress (Figure 4.29, Table 4.39).

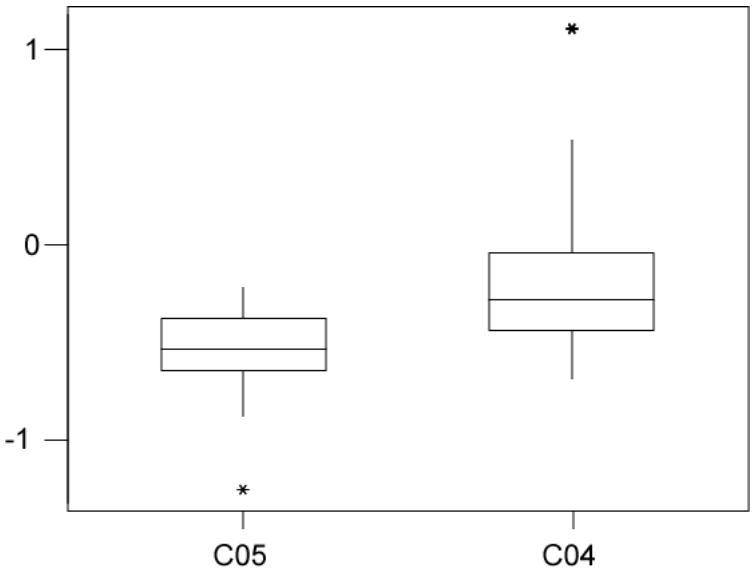


Figure 4.28 Substratum composition expressed as Principal Component Scores, where negative scores indicate coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at the two stream types sampled in the Czech Republic, C04 and C05 (see above).

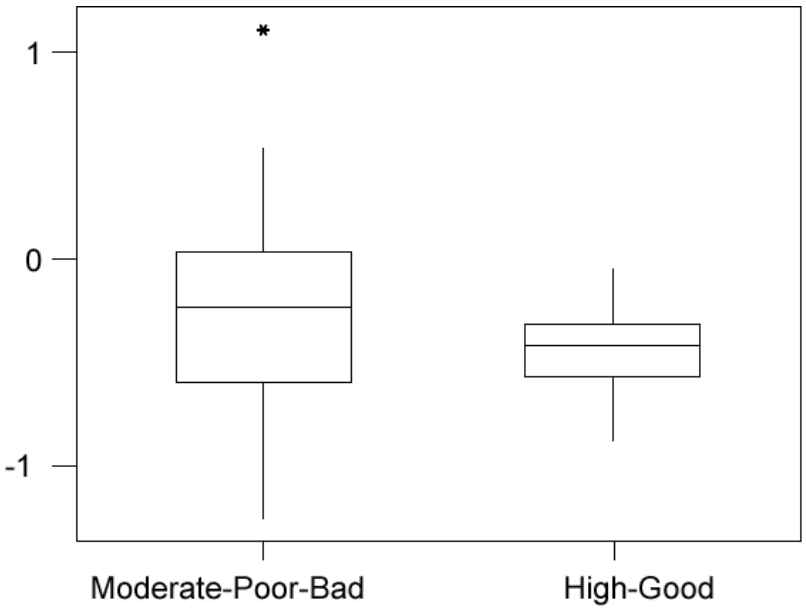


Figure 4.29 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at sites with different ecological status (pre-classified) in the Czech Republic, High-Good vs Moderate-Poor-Bad ecological status (see above).

In the Danish streams, there were statistical differences in substratum composition among the sampled seasons, where the substratum for samples taken in the summer were classified as finer than the substratum sampled in spring Figure 4.30, Table 4.39). No such difference was found for either sampling method used (DSFI versus STAR-AQEM or for the pre-defined stressor gradient.

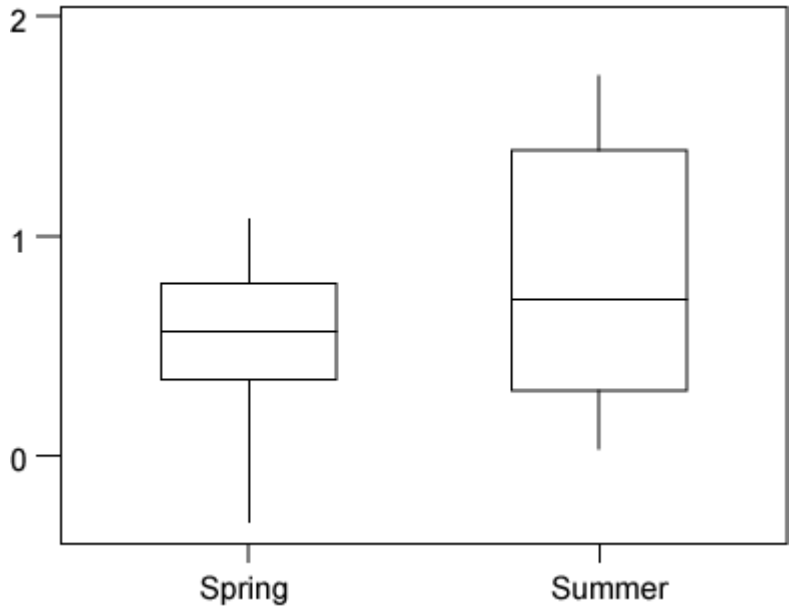


Figure 4.30 Substratum composition expressed as Principal Component Scores, where negative scores indicate coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken in different seasons in Denmark, spring versus summer (see above).

In the French streams, there was no difference in substratum composition for samples taken using the STAR-AQEM versus the IBGN method. There was no difference in substratum composition for samples taken in different seasons either. There was, however, a clear difference in substratum composition among samples taken at sites with a pre-defined high or good ecological status versus a moderate, poor, or bad status (Figure 4.31). Sites pre-classified as having a high or good ecological status had a more course substratum composition than sites classified as having a moderate or worse ecological status.

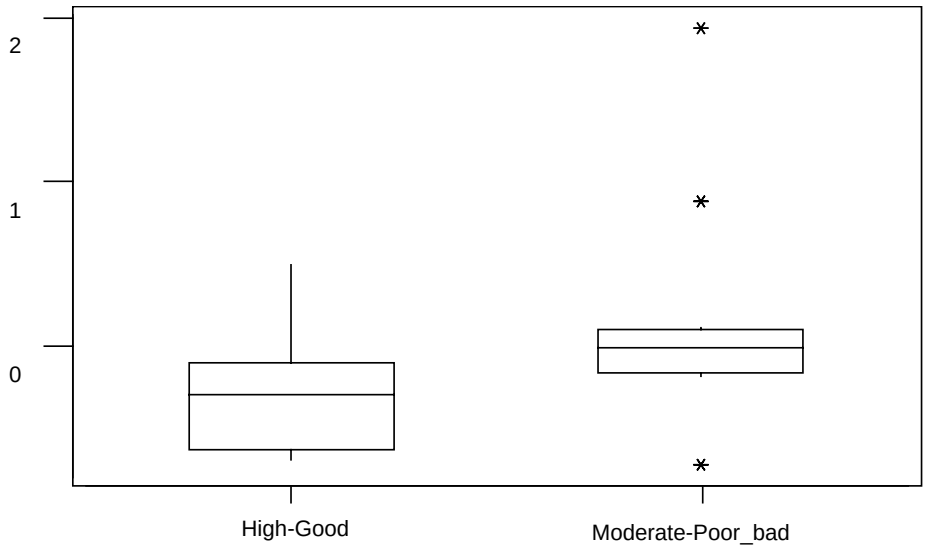


Figure 4.31 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at sites with different pre-classified ecological status in France, pre-classified as having a high or good versus a moderate, poor or bad ecological status (see above).

In the German streams, there was no difference in substratum composition for samples taken using the STAR-AQEM versus the RIVPACS method. There was no difference in substratum composition for samples taken in different seasons either. There were, however, a clear difference in substratum composition among the sampled stream types, where samples taken in type D03 had higher PCA axis scores (i.e. fine substratum types) as opposed to stream types D04 and D06 (Figure 4.33). This is because in the stream type D03 (lowland sandy streams) course substratum types indicates a detoriated condition, whereas in stream types D04 and D06 course substratum types are normal. When comparing substratum composition among samples taken at sites with a pre-defined high or good ecological status versus a moderate, poor, or bad status in stream types D04 and D06 (Figure 4.32a), the sites having a high or good ecological status had a finer substratum composition than sites classified as having a moderate or worse ecological status. In stream type D03, finer substratum composition was related to higher ecological status, which is natural since the stream type in its unstressed state has a sandy substratum (Fig. 4.32b).

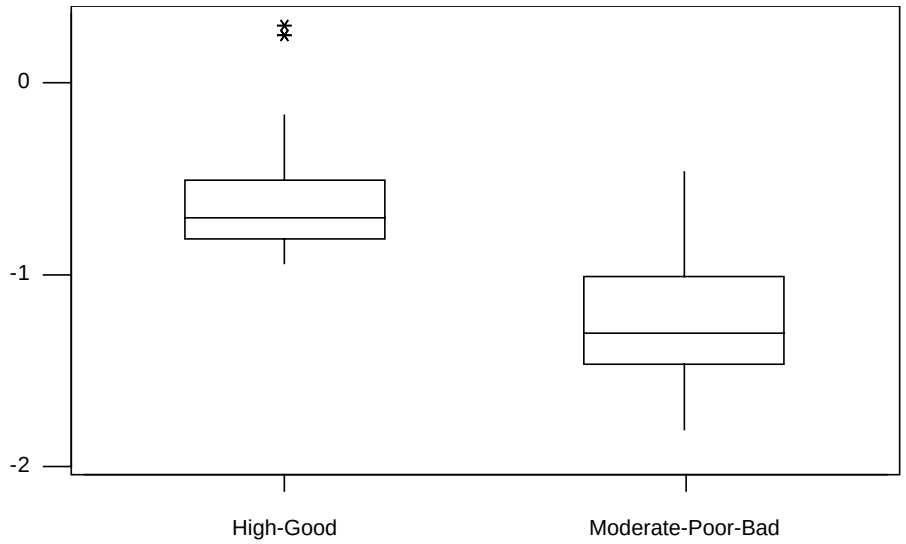


Figure 4.32a Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at sites with different pre-classified ecological status in German stream types D04 and D06, pre-classified as having a high or good versus a moderate, poor or bad ecological status (see above).

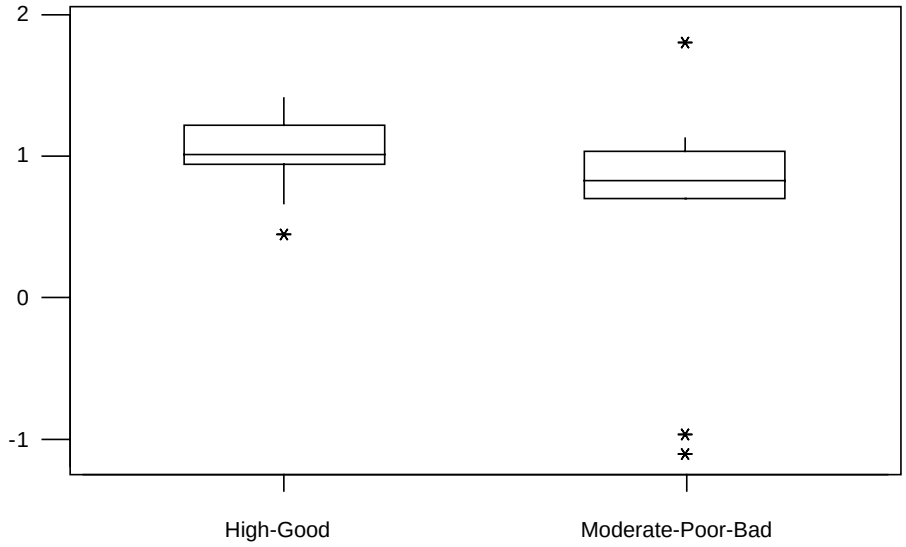


Figure 4.32b Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at sites with different pre-classified ecological status in German stream steam D03, pre-classified as having a high or good versus a moderate, poor or bad ecological status (see above).

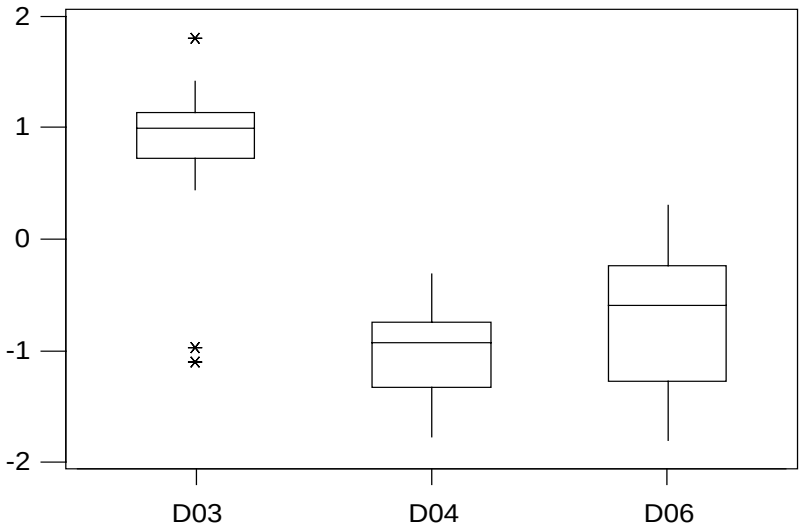


Figure 4.33 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at the three stream types sampled in the Germany, D03, D04, and D05 (see above).

In the Greek streams, there was no difference in substratum composition for samples taken using the STAR-AQEM versus the RIVPACS method. There was no difference in substratum composition for samples taken in different seasons or difference sin pre-defined ecological status either. There were, however, a clear difference among the four types sampled in Greek streams (Figure 4.34), where samples taken in type H07 had higher PCA axis scores (i.e. fine substratum types) as opposed to stream types H04, H05, and H06.

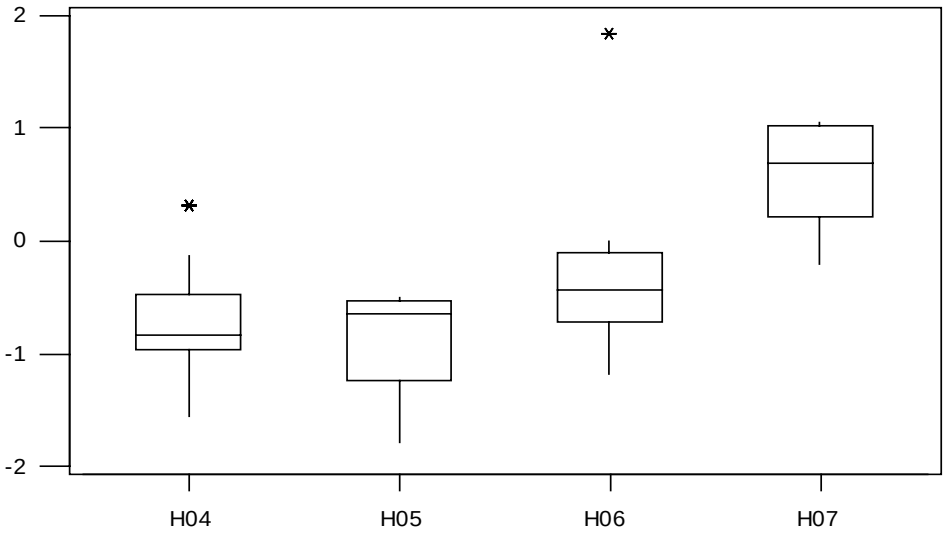


Figure 4.34 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken at the four stream types sampled in the Greece, H04, H05, H06, and H07 (see above).

There were no differences in substratum composition for either sampling season or pre-defined ecological quality in the Italian streams sampled by CNR. There was, however, a clear difference in substratum composition among the two sampling methods, where the STAR-AQEM samples generally were taken on more course substratum (lower PCA scores) as opposed to the IBE method, where most samples were taken on identical (finer) substratum (Figure 4.35, Table 4.39).

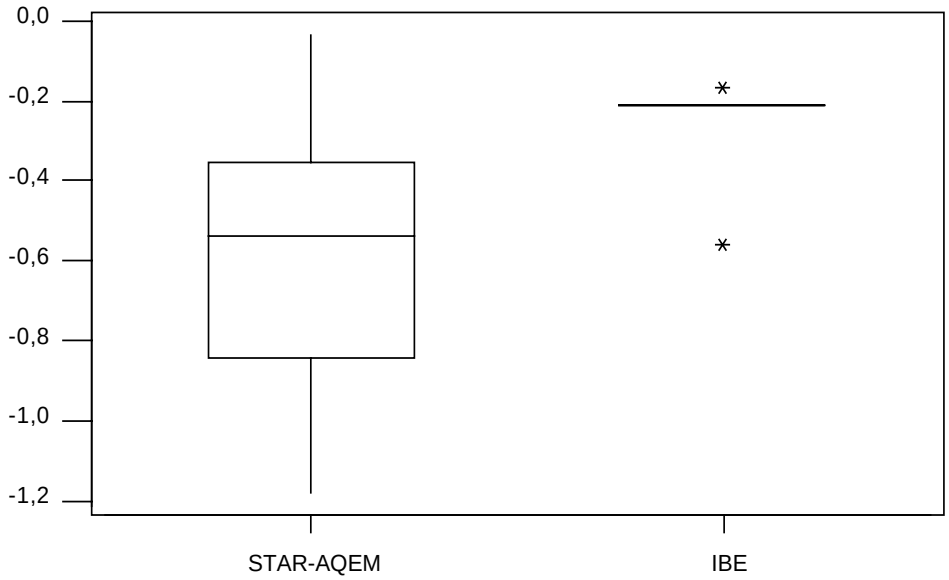


Figure 4.35 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken using the two sampling methods in Italy by CNR, STAR-AQEM versus the IBE method (see above).

There were no differences in substratum composition for either sampling season or pre-defined ecological quality in the Italian streams sampled by Bolzano. There was, however, a clear difference in substratum composition among the two sampling methods, where the STAR-AQEM samples generally were taken on more course substratum (lower PCA scores) as opposed to the IBE method, where all samples were taken on identical (finer) substratum (Figure 4.36, Table 4.39).

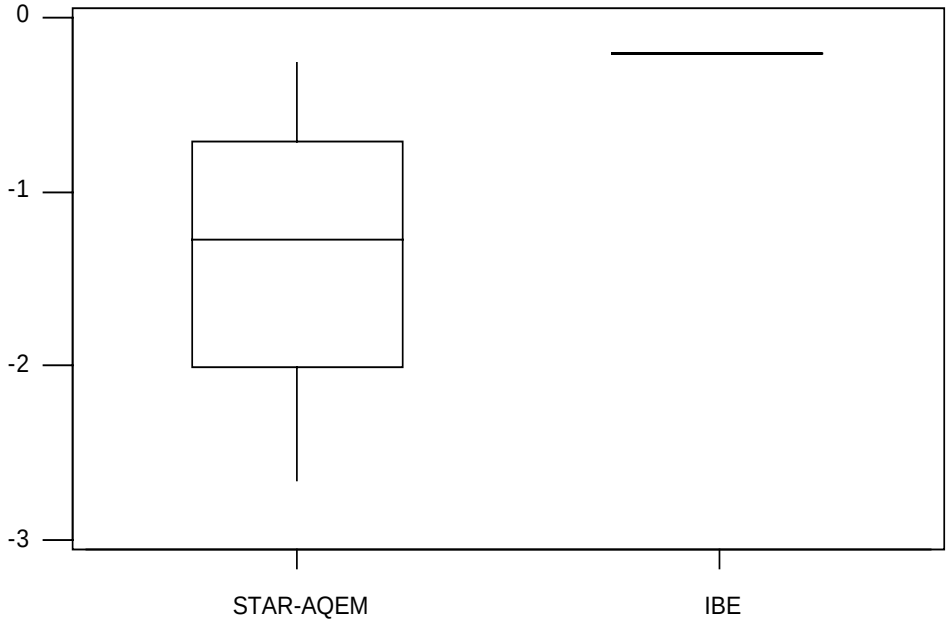


Figure 4.36 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken using the two sampling methods in Italy by Bolzano, STAR-AQEM versus the IBE method (see above).

There were no differences in substratum composition for either sampling season or pre-defined ecological quality in the Latvian streams. There was, however, a clear difference in substratum composition among the two sampling methods, where the STAR-AQEM samples generally were taken on a finer substratum (higher PCA scores) as opposed to the Latvian method, where all samples were taken on identical (courser) substratum (Figure 4.37, Table 4.39).

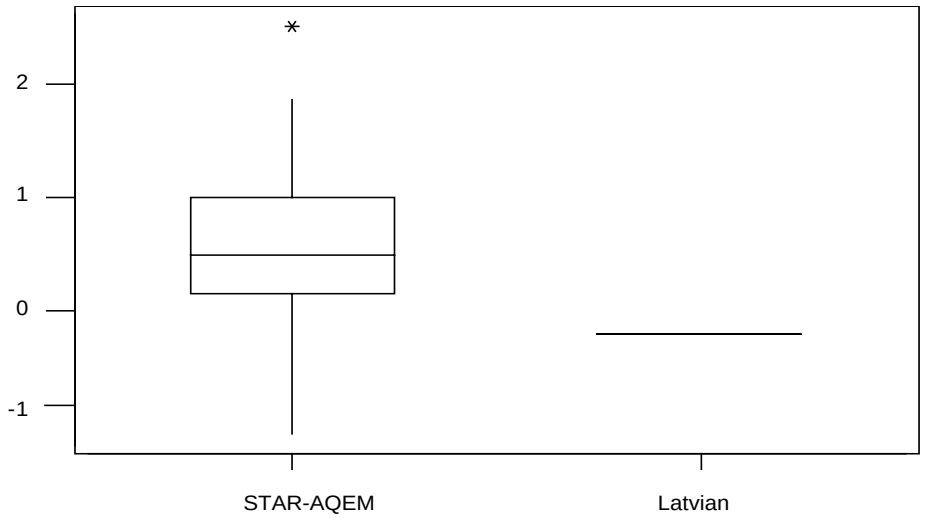


Figure 4.37 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken using the two sampling methods in Latvia with the STAR-AQEM versus the Latvian sampling method (see above).

There were no differences in substratum composition for either sampling season or the two sampling methods in Latvian streams. There was, however, a clear difference in substratum composition among samples with different pre-defined ecological quality, where sites pre-classifed as having a High-Good status generally were taken on a courser substratum (lower PCA scores) than those pre-classified as Moderate-Poor-Bad (Figure 4.38, Table 4.39).

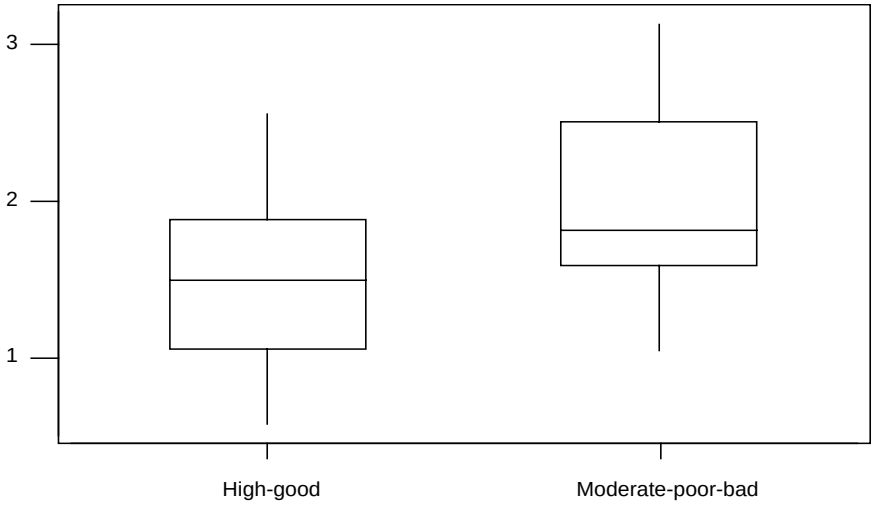


Figure 4.38 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken from Latvian sites with different pre-classfications of ecological status.

The Portuguese data for substratum type were analysed separately from the rest of the datset, here positive PCA scores on axis 1 were related to course substratum. A statistical difference was found among samples taken at sites pre-classified as having ahigh or good ecological status versus the other quality classes (Fig. 4.39). Sites with a high or good status generally had a courser substratum type than the other quality classes.

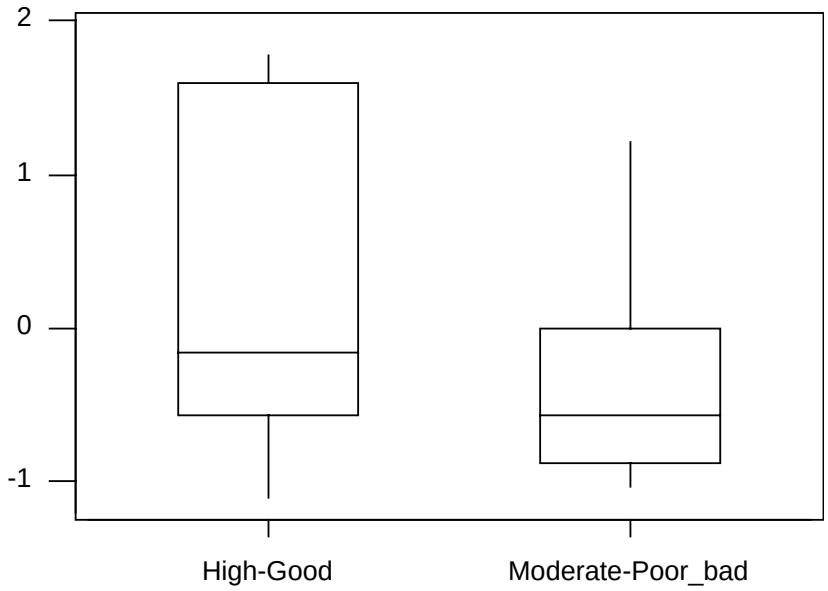


Figure 4.39 Substratum composition expressed as Principal Component Scores, where positive scores indicates coarse substratum types, whereas negative scores indicate fine substratum types (as opposed to all other PCA substartum analysis in this report). Comparisons of samples taken Portuguese sites with different pre-classfication of ecological status.

There were no differences in substratum composition for sampling season, pre-defined ecological quality or sampling method used in the Slovak republic streams (Table 4.39).

There were no differences in substratum composition for either samples taken at sites with different pre-defined ecological quality or samples taken using the two sampling methods in Swedish streams (Figure 4.40, Table 4.39). There was, however, a clear difference in substratum composition among samples taken in the two stream types, where the type S05 generally had a finer substratum than the type S06. There was also a difference in substratum composition in the two sampled seasons, where samples in spring generally were taken on a courser substratum than samples taken in the autumn.

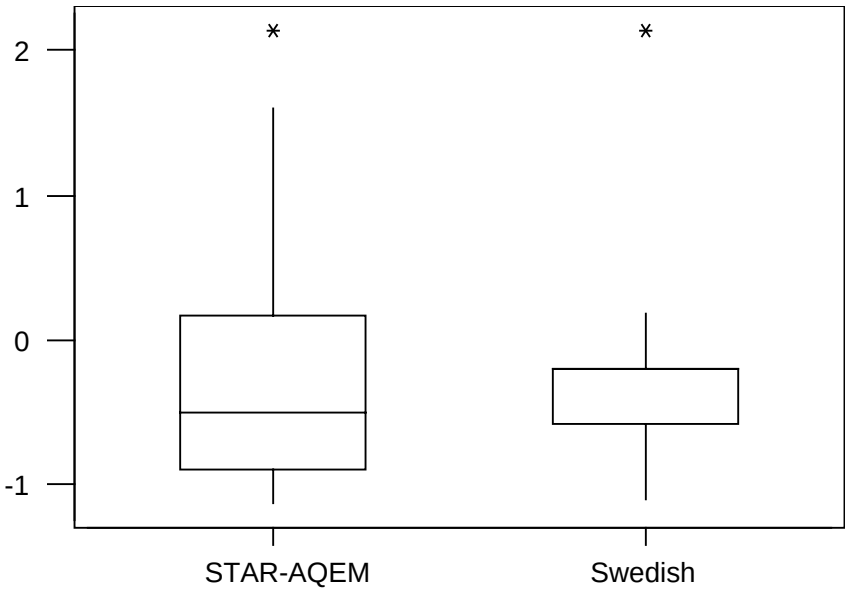


Figure 4.40 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken using the two sampling methods in Sweden with the STAR-AQEM versus the Swedish sampling method (see above).

There were no differences in substratum composition for either sampling season or samples with different pre-defined ecological quality in the UK streams. There was, however, a clear difference in substratum composition among samples taken in the two stream types, where the type U15 generally had a courser substratum than the type U23. There were also differences in substratum composition for the RIVPACS versus the STAR-AQEM samples, where the RIVPACS substratum types all were classified the same and where the STAR-AQEM samples generally were sampled on more fine substratum types (Figure 4.41, Table 4.39).

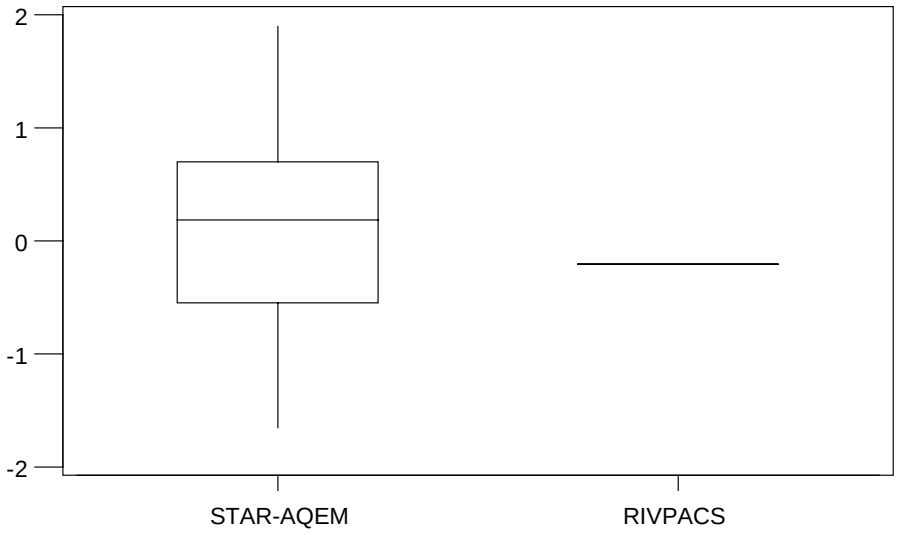


Figure 4.41 Substratum composition expressed as Principal Component Scores, where negative scores indicates coarse substratum types, whereas positive scores indicate fine substratum types (see above). Comparisons of samples taken using the two sampling methods in UK with the STAR-AQEM versus the RIVPACS sampling method (see above).

5. REPLICATE SAMPLING AND SUB-SAMPLING VARIABILITY

5.1 REPLICATE SAMPLING PROGRAMME WITHIN STAR

All assessments of the ecological status of a river site using macroinvertebrate sampling are subject to uncertainty and errors. Most quantitative assessments of the biological status of water bodies are based on the values of biological indices or metrics derived from the taxonomic composition of the sample, where the metric is intended to measure some specific aspect of general feature of the biota. An index of ecological quality or status is of little value without some knowledge of its levels of uncertainty (Clarke 2000, REFCOND 2003). In particular, it is important to have quantitative estimates of the effects of sampling variation on the value of any biotic index or metric used to assess the ecological status of a river site. Replicate sample values will vary because of inherent natural small-scale spatial heterogeneity in the fauna at a site. Sampling methods and derived indices which are very prone to high levels of variation between replicate samples will tend to provide less reliable estimates of ecological quality ratios and ecological status for a site and have less power and confidence to detect changes in ecological quality (Clarke *et al* 2002).

The STAR project was therefore designed to include a extensive replicated sampling programme within the main field sampling programme.

As part of the STAR field sampling programme, STAR-AQEM samples were taken at all sites by each participating partner. At each site in nearly all of the main stream types, each partner also collected samples using a notional "national" method. This was normally a widely used protocol within the individual partner's Member State, but in Germany, Austria and Greece where there was no existing common national sampling protocol, the UK RIVPACS protocol was used. The sampling methods are described in detail in Section 2. Both STAR-AQEM and national samples were collected in two seasons - spring and either summer or autumn - the precise months involved varied because of climatic differences across Europe.

To assess sampling variability, each partner took a second replicate field sample using each sampling method in each sampling season at a subset (usually six) of their sites. These sites were carefully selected within each sampled stream type to cover a range of perceived (i.e. pre-classified) qualities of sites from 'high' and 'good' to 'moderate' or 'poor'/'bad'. This was important because the sampling variability of one or more metrics may depend on the quality of a site; poorer quality sites with fewer taxa present might be less variable in some taxonomic richness/diversity metrics, but more variable in metrics based on some form of average stress-tolerance score of the taxa present (e.g. ASPT or a Saprobic index). The two samples taken from a site in any one season were referred to as the 'main' and 'replicate' samples.

The STAR-AQEM method protocol involves a standardised method of laboratory sub-sampling of the macroinvertebrate field sample. The sample material is spread out as evenly as possible on a tray marked out with a 6 by 5 grid of cells. The STAR-AQEM protocol requires the biologist to randomly select five of the 30 grid cells and identify and count all of the macroinvertebrate specimens in these five cells. If necessary additional cells are randomly selected until at least 700 individuals have been identified.

This sub-sampling procedure will introduce an additional source of variation in the recorded taxonomic composition for the site and hence in the values of metrics for the site at that time. This source of variation was not assessed in the previous AQEM project which devised the AQEM (now STAR-AQEM) sampling method protocol. Variation in taxonomic composition and metric values between replicate field samples taken from the same site at the same time will be due to both sampling spatial variation in the field and laboratory sub-sampling effects.

To quantify the size of the sub-sampling source of variation, especially in relation to field sampling variability, STAR project partners took a second replicate sub-sample from one of the replicate STAR-AQEM samples for all or most of the sites at which two replicate samples were taken (Table 5.1). The two sub-samples taken from a field sample were referred to as the 'main' and 'replicate' sub-samples.

Table 5.1 Sites in each stream type and country for which two STAR-AQEM field samples ('main' and 'replicate') were taken and for which two sub-samples ('main' and 'replicate') were taken from one field sample in at least one season (1=spring, 2=summer, 3=autumn). Site code 'xxx.y' indicates replicate sub-samples only taken at site 'xxx' in season 'y'. Brackets indicate sites with replicate samples but no replicate sub-samples; Bold indicates no replicate sample, but two-samples from a single field sample.

Country	Stream Type	Description	Seasons sampled	n Sites	STAR site codes
Austria	A05	small-sized, shallow mountain streams	1 + 2	5	600 603 607 609.2 952.1
	A06	small-sized crystalline streams of the ridges of the Central Alps	1 + 2	5	701 702.2 706 708.1 (704.2 708.2)
Czech Republic	C04	small-sized, shallow mountain streams	1 + 2	3	614 620 625
	C05	small-sized streams in the Central sub-alpine Mountains	1 + 2	3	713 717 722
Germany	D03	medium-sized lowland streams	1 + 2	2	649 659
	D04	small-sized, shallow mountain streams	1 + 2	2	627 634
	D06	small-sized Buntsandstein-streams	1 + 2	2	816 821
France	F08	small-sized, shallow headwater streams in Eastern France	1 + 3	6	724 725 726 728 729 733.3 (733.1)
Greece	H04	small-sized calcareous mountain streams in Western, Central and Southern Greece	1 + 2	6	(735 737 738 739 753 756)
Italy	I05	small-sized streams in the southern calcareous Alps	1 + 2	3	849 855 856
	I06	small-sized calcareous streams in the Central Apennines	1 + 2	6	(836 837 840 842 843.2 845)
Denmark	K02	medium-sized lowland streams	1 + 2	6	662 663 665 667 671 673
Latvia	L02	medium-sized lowland streams	1 + 3	13	(996.1 997 1002.1 1005.1 1006 1007 1010 1013.1 1016 1017 1027.1 1030.1 1034.1)
Poland	O02	medium-sized lowland streams	1 + 3	5	895 897.3 903 913 916.1 (897.1 915.3 916.3 1036)
Portugal	P04	medium-sized streams in lower mountainous areas of S. Portugal	1 + 3	3	**859 860** 867 (863 864 865 866.3 868)
Sweden	S05	medium-sized lowland streams	1 + 3	5	685.1 689.1 691 **695.3 697.3** (685.3 689.3)
	S06	medium-sized streams on calcareous soils	1 + 3	3	875 876 878
UK	U15	small-sized, shallow lowland streams	1 + 3	3	639 642 648
	U23	medium-sized lowland streams	1 + 3	3	674 678 681

5.2 CALCULATION OF METRIC VALUES

The values of almost all of the metrics were calculated using AQEM Assessment Software version 2.3 (AQEMrap). This was available to all partners (as well as the public) as a downloadable package from the AQEM web site http://www.aqem.de/products.htm. Instructions for using this software are included with the software (AQEM Consortium, 2004). The taxonomic data were first exported as Excel spreadsheets from the STAR macroinvertebrate database held in AQEMdip on the STAR web site and then taxonomically adjusted to a consistent national level. These files were then imported into the metric calculation software and the metrics values for each sample and sub-sample for each sampling method exported to Excel files. These metric files for each STAR partner were then combined with the relevant site characteristics meta data into a single dataset for statistical analysis within the Minitab Release 14 statistics package (http://www.minitab.com).

The analyses reported here are for 27 ecological quality metrics intended to represent a wide range of aspects and responses of the macroinvertebrate fauna (Table 5.2). These metrics include all of the 16 metrics (Table 3.4) used in Sections 3 and 4 to compare sampling methods, together with other potentially important or common inter-calibration metrics. The STAR database permits similar analyses to be made for other available metrics.

All macroinvertebrate samples taken from stream types in Italy, Greece and Portugal whether using the STAR-AQEM, RIVPACS or national method, were only identified to family level. The three Saprobic indices, which require data identified to species or genus level, were therefore not calculated for sites from these stream types. The metrics measuring percentage or proportional abundance of specific guilds (%Rheophilic, %Littoral, %Grazers/scrapers, %Shredders, %Gatherer/collectors and RETI) were calculated using family level data for Italy, Greece and Portugal, but using species level data for all other countries. Similarly, the metrics measuring the total 'Number of taxa' and the Shannon-Wiener diversity index will depend on the taxonomic resolution of the data; higher taxonomic resolution will obviously lead to more individual taxa being recorded and probability more variability in results. The sampling standard deviation (SD) of these metrics for the stream types based on family data may not be comparable those based on species and genus level data.

The value of the new metrics Log(Sel_EPTD+1) and 1-GOLD, which are two of the six proposed Inter-calibration Common Metrics (ICMs) (Buffagni *et al.* 2004) were calculated separately from AQEMrap. Replicate STAR-AQEM sub-samples values of the metrics Log(Sel_EPTD+1) and (1-GOLD) were only available for two stream types C04 and C05 in the Czech Republic; however values were available for both a 'main' and 'replicate' sample for some sites in many stream types. The Italian national IBE metric is included for provisional information, although there are still some unresolved problems in its calculation within AQEMrap and it may be inappropriate for many non-Italian stream types.

5.3 STATISTICAL METHODS USED TO QUANTIFY VARIABILITY IN METRIC VALUES

The statistical analysis concentrates on assessing the sampling and sub-sampling variability in many of the most commonly used macroinvertebrate-based metrics.

Statistical analysis of variance (ANOVA) techniques were used to estimate the variances in the observed metric values due to each source. The analyses used the hierarchical nested ANOVA procedure in the Minitab Release 14 statistics package (http://www.minitab.com),

which is based on equating ANOVA observed mean squares to their expected vales; negative variance estimates were set to zero.

In particular, for the STAR-AQEM sample data, hierarchical nested ANOVA (Minitab Release 14, 2004) was used to estimate the following variance components:

σ_U^2 = variance due to differences between replicate sub-samples within a sample

σ_R^2 = variance due to differences between replicate samples within a site and season

σ_I^2 = variance between individual site means within a season and stream type

σ_J^2 = variance between season means within a stream type

σ_K^2 = variance between stream type means

This approach correctly identifies that part of the overall variance between replicate samples which is merely the consequence of sub-sampling (namely σ_U^2) from that due to real differences between the two samples in the fauna obtained (namely σ_R^2). The overall variance (denoted σ_E^2) between replicate samples taken using the STAR-AQEM method is the sum of the two components, namely: $\sigma_E^2 = \sigma_U^2 + \sigma_R^2$. The relative importance of sub-sampling effects to sampling effects was assessed and measured by the statistic: $P_{sub} = 100\sigma_U^2 / \sigma_E^2$.

For the sample data collected using either the RIVPACS protocol or the 'national' protocol, replicated sub-sampling was not involved and hierarchical nested ANOVA was used to estimate the overall variance (σ_E^2) in metric values between replicate samples, together with the between-season and between-site variances.

If a particular metric and sampling method are to be effective in discriminating the ecological status classes of river sites within a stream type, then the overall replicate sampling variance (σ_E^2) should be small relative to the total variability in metric values within the stream type.

The average total variance in metric values over all sampled sites within any one stream type is measured by: $\sigma_T^2 = \sigma_E^2 + \sigma_I^2 + \sigma_J^2$.

The size of the overall replicate sampling variance (σ_E^2) relative to the total variance (σ_T^2) within a stream type is then given by the statistic: $P_{samp} = 100\sigma_E^2 / \sigma_T^2$. This is a better practical measure of the precision of each metric than using the usual coefficient of variation (CV) determined as the ration of the replicate SD to the replicate mean. This is because the typical actual range of values a metric takes with real samples rarely includes values near zero, so a low CV may not indicate high precision in practice. As an example, a metric may have a sampling SD of say 0.5 on replicate means ranging from around 5.0 to 6.0; giving a CV of 10% or less. However, because of the limited range of values of the metric (roughly 4 -7), the percentage P_{samp} of total variance due to sampling is much higher at around 40%.

This method of estimating P_{samp} is the best approach if the observed metric values are subsequently to be compared against a single reference condition value for a stream type or site, regardless of season, because in such cases the relevant average total variance in metric values within any one stream type should include the between season variance σ_J^2. This is the approach used throughout this report to estimate and compare the relative sizes (P_{samp}) of the sampling variances for each metric and sampling method.

However, if the observed metric values are subsequently to be compared against a season-specific reference condition value for a stream type or site (as done in the UK RIVPACS bioassessment system and software), then the relevant average total variance σ_T^2 in metric values within any one stream type should exclude the variance σ_J^2 due to differences between seasons in average metric value. Analyses reported below show that σ_J^2 is usually much less than inter-site variance σ_I^2 such that the estimates of P_{samp} provided in this study will also give a reasonable guide to the relative precision of metrics in the case where inter-season variance is excluded.

Because it was only possible to take replicate samples (and STAR-AQEM sub-samples) at a few (2-7) sites in each stream type of each STAR partner, estimates of the above variance components for individual stream types may be imprecise. Therefore, to obtain more robust estimates for a particular sampling method, the variance components (and their relative size) for a particular metric are also derived using all of the sites for which the method was used in a particular country, and also for all sites regardless of country. The variance components are usually quoted in the tables in their standard deviations (SD) form where SD is the square root of the variance. This is done because the SD are in the same units as the metric values and hence it is much easier to understand their practical size. For, example $SD_U = \sqrt{\sigma_U^2}$ denotes the SD due to STAR-AQEM sub-sampling and $SD_E = \sqrt{\sigma_E^2}$ denotes the overall SD due to variability between replicate sample values. When a SD is based on only two values (x_1 and x_2) then the SD is equal to the absolute value of their difference divided by the square root of two (i.e. $|x_1 - x_2|/\sqrt{2} = 0.71|x_1 - x_2|$).

Frequently in ecology, the replicate sampling variability in a biotic index of taxonomic abundance, richness or composition often increases with the value of the index. For example, Clarke *et al.* (2002) found that the variance in the number of macroinvertebrate taxa found in replicate RIVPACS samples increased roughly in proportion to the average number of taxa found in samples from the same site, but that by transforming the data, the replicate variability in the square root of the number of taxa was roughly constant and did not depend on the physical type or ecological quality of the sites.

Using a similar approach for the STAR dataset, Taylor's Power Law regressions of (natural) log replicate variance against (natural) log replicate mean were used to estimate the best data transformation to reduce the systematic variability in the replicate standard deviation of metric values (Taylor 1961; Elliott 1997, Clarke *et al.* 2002). The log-log regression slope *b* indicates the mathematical power (m^b) of the replicate mean (m) with which the replicate variance increases. A slope approximately equal to zero indicates that the replicate sampling variance does not vary systematically and no transformation is necessary. When the slope is approximately 1.0, 1.5 or 2.0, then a transformation of metric values using, respectively, a square root ($\sqrt{x}$), double square root ($\sqrt{\sqrt{x}}$) or logarithmic ($\log(x)$) transformation will lead to more equitable variances. The double square root transformation is similar in effect to taking logarithms, but is preferred because it has the advantage of not needing to add an arbitrary constant (e.g. $\log(x+1)$) when a metric has some values of zero. Spearman rank correlations between replicate SD and replicate mean were also calculated.

These analyses were used to determine whether, for a particular sampling method and metric, the sampling and other variances should be analysed and estimated using the metric's untransformed or transformed values. For reasons of consistency and robustness, only one transformation was used for any single metric regardless of sampling method or stream type. The decision of which single transformation, if any, to use for each metric, was based on

assessments of patterns in the overall replicate sampling variability for the STAR-AQEM and RIVPACS methods, as the other methods were only used on relatively few sites in a single region (Table 5.2)

Many of the selected metrics are percentages (range 0-100) or proportions (0-1) which are based on the fraction of all individuals or of all taxa which are in a particular group or have particular characteristics. The replicate sample values of such metrics tend to be less variable when their values for a site are very low (near zero) or very high (near 100%) and most variable at intermediate values (20-80%). In such cases, the arcsine transformation of the square root of the proportions x (i.e. arcsine($\sqrt{x}$)) is the standard transformation used in statistical analyses to make the sampling variance more equitable (e.g. Sokal & Rohlf 1995). Visual inspection of plots of replicate SD versus replicate mean values of such metrics for the STAR data suggested that this transformation was appropriate and it was applied and recommended for all such metrics which are percentages or proportions. When the metric x is a percentage rather than a proportion, then the exact transformation is arcsine($\sqrt{(x/100)}$). If the values of such metrics in a dataset are all less than 50%, then the replicate SD may appear to increase with the replicate mean, but it is safer to retain a sensible "model" for the variation about the whole potential range of 0-100%.

Table 5.2 Spearman rank correlations between replicate sampling SD and replicate mean value for the STAR-AQEM and RIVPACS methods on untransformed (x) and appropriately transformed (f(x)) metric values. Based on all available sites from all available countries and stream types.

Metric	Transform f(x)	STAR-AQEM		RIVPACS	
		x	f(x)	x	f(x)
Abundance [ind/m²]	√√x	0.77	0.33	0.64	0.03
Number of taxa	√x	0.45	0.18	0.58	0.40
Number of Families	√x	0.23	0.00	0.31	0.13
Number of EPT taxa	√x	0.36	0.01	0.38	0.10
Saprobic Index (Zelinka & Marvan)	x	-0.07		0.19	
German Saprobic Index (new version)	x	0.45		0.25	
Czech Saprobic Index	x	0.07		0.24	
Average score per Taxon (ASPT)	x	-0.10		0.16	
IBE	x	0.16		0.16	
Diversity (Shannon_Wiener_Index)	x	-0.19		-0.25	
% Rheophilic preference (Type RP)	asin(√(x/100))	0.12	-0.04	-0.24	-0.36
% Type RP (abundance classes)	asin(√(x/100))	-0.09	-0.18	-0.14	-0.19
% Littoral preference	asin(√(x/100))	0.15	-0.02	-0.23	-0.38
% Grazers/Scrapers	asin(√(x/100))	0.26	0.00	-0.01	-0.19
% Shredders	asin(√(x/100))	0.74	0.27	0.47	-0.03
% Gatherers/Collectors	asin(√(x/100))	0.20	0.17	0.27	0.13
% Oligochaeta	asin(√(x/100))	0.82	0.45	0.82	0.30
% EPT individuals	asin(√(x/100))	0.33	0.13	0.28	0.19
% EPT (abundance classes)	asin(√(x/100))	0.06	-0.08	0.22	0.15
% EPT Taxa	asin(√(x/100))	0.01	-0.01	0.19	0.13
RETI	asin(√x)	0.02	-0.13	-0.44	-0.44
Log(Sel_EPTD+1)	x	-0.18		-0.55	
1 –GOLD	asin(√x)	-0.08	-0.07	-0.35	-0.17
Trait m1 : max body size (≤ 1cm)	asin(√x)	0.02	0.05	-0.01	0.05
Trait m2 : reproductive cycles per year (>1)	asin(√x)	0.28	0.23	0.49	0.49
Trait m7 : locomotion+substrate relation (crawler)	asin(√x)	-0.15	-0.15	-0.54	-0.54
Trait m12 : current velocity preferred (<25cm/s)	asin(√x)	-0.07	0.01	0.08	0.08

Table 5.2 shows the single transformation which was judged to be the most appropriate for each of the individual metrics. The table also shows the Spearman rank correlation between the overall replicate sampling standard deviation (SD_E) and the replicate sampling mean across all sites and seasons. Ideally there would be little or no correlation between the SD and the mean so that a single estimate of SD could be used for all sites. The high correlations for the metrics 'Abundance', 'Number of taxa', '%Shredders' and '%Oligochaeta' are all greatly reduced on the transformed scales for the metric values, as intended.

Subsequent analyses to estimate sampling and other variances were conducted on the transformed values of each metric, using the transformation specified in Table 5.2.

Where appropriate Kruskal-Wallis one-way analysis of variance of the ranks of the replicate sampling SD (SD_E) was used to test for statistically significant difference in SD_E between stream types.

5.4 ESTIMATES OF SAMPLING AND SUB-SAMPLING VARIABILITY IN METRIC VALUES FOR THE STAR-AQEM METHOD

5.4.1 Estimates of sub-sampling variability

The effect of only identifying the individuals within a sub-sample fraction of a STAR-AQEM sample was assessed by assessing the size and pattern of differences in metric values obtained from sites where two replicate sub-samples were taken from the same sample (Table 5.1).

Figures 5.1-5.6 plot the difference between two replicate sub-samples against the average of the two values for untransformed values of a range of metric types. The differences rather than the SD of the two replicate sub-samples are shown in these initial plots to aid visual understanding. Figure 5.1 shows difference in the 'Number of taxa' found in two replicate sub-samples from the same STAR-AQEM sample. In the majority of cases the difference is less than five and often less than two. However there are occasional large differences, the most extreme of which was one sample from Denmark (stream type K02) where 25 taxa were found in one sub-sample and 44 in the other. In terms of 'Number of Families', the difference between sub-samples was no more than three in the majority of cases, but there was a difference of nine families for one Swedish sample (stream type S05) and of 10 families for the same Danish sample as mentioned above (Figure 5.2). The sub-sampling variability in 'Number of taxa' and 'Number of Families' tends to be less when fewer taxa/families are present, which is why the sampling and sub-sampling SD of these metrics were estimated on the square root transformed values, for which the SD was less systematically variable.

Table 5.3 gives the estimates of the standard deviation in untransformed metric values due to STAR-AQEM sub-sampling, with separate estimates for each STAR stream type for which replicate sub-sample values were obtained. Table 5.4 gives the same information but, where indicated, using the transformed values of particular metrics. Because the metric 'Log(Sel_EPTD+1)' was only available for replicate STAR-AQEM sub-samples for two stream types C04 and C05 in the Czech Republic, the overall estimate of its sub-sample SD may not be reliable for other stream types.

These latter estimates could be used in the STARBUGS software package (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sub-sampling variability in individual metric values on the uncertainty of multi-metric assessments of the ecological status of sites. These estimates can also be used to provide information on the expected uncertainty in metrics values due to sub-sampling effects for sites where no replicate sub-samples have been taken.

Some stream types might be expected to give more 'nuisance' material of small-scale debris than others, which might influence the ability to distribute the macroinvertebrates evenly between the grid cells. However, there are no obvious systematic consistent differences between stream types in the pattern and extent of sub-sampling variation in metric values (Figure 5.1-5.6).

For any new site, it may be most reliable to use the estimates of sub-sampling (or sampling) SD in Table 5.4 for the same stream type, if available. This may be most appropriate when the metric values and variability are highly dependent on either the stream type or the precise taxonomic resolution used by each STAR partner (e.g. as for the metric 'Number of taxa'). However, there are only a few STAR sites in each stream type for which replicate samples were taken, so the individual estimates of SD may themselves be imprecise. Therefore, if there are no obvious major differences in the SD between stream types, it may be more robust

to use estimates based on the information from a combination of stream types, or even the median SD value across all stream types, as given in Table 5.4.

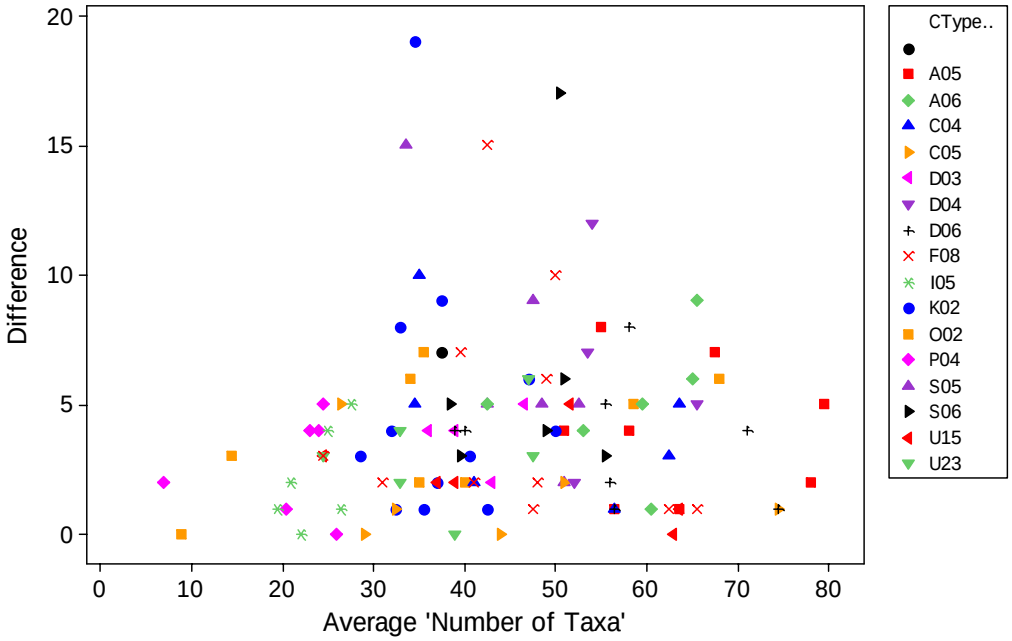


Figure 5.1 STAR-AQEM method: Difference between two replicate sub-samples plotted against the average of the two values for untransformed values of the metric 'Number of taxa', for all available STAR sites and seasons with replicated sub-samples. Symbols denote STAR stream type of each site.

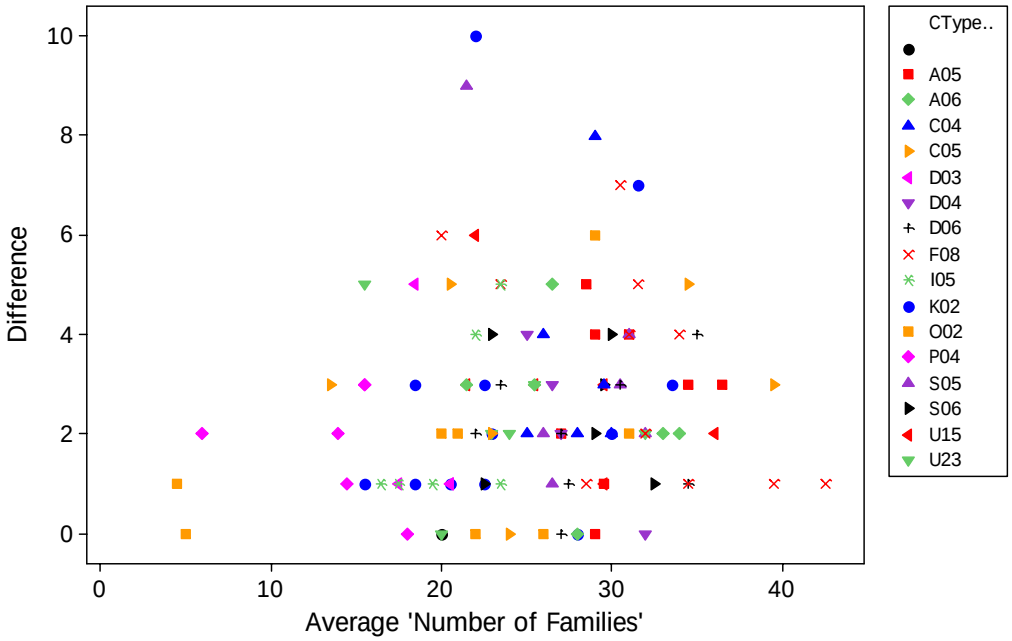


Figure 5.2 STAR-AQEM method: Difference between two replicate sub-samples plotted against the average of the two values for untransformed values of the metric 'Number of Families'.

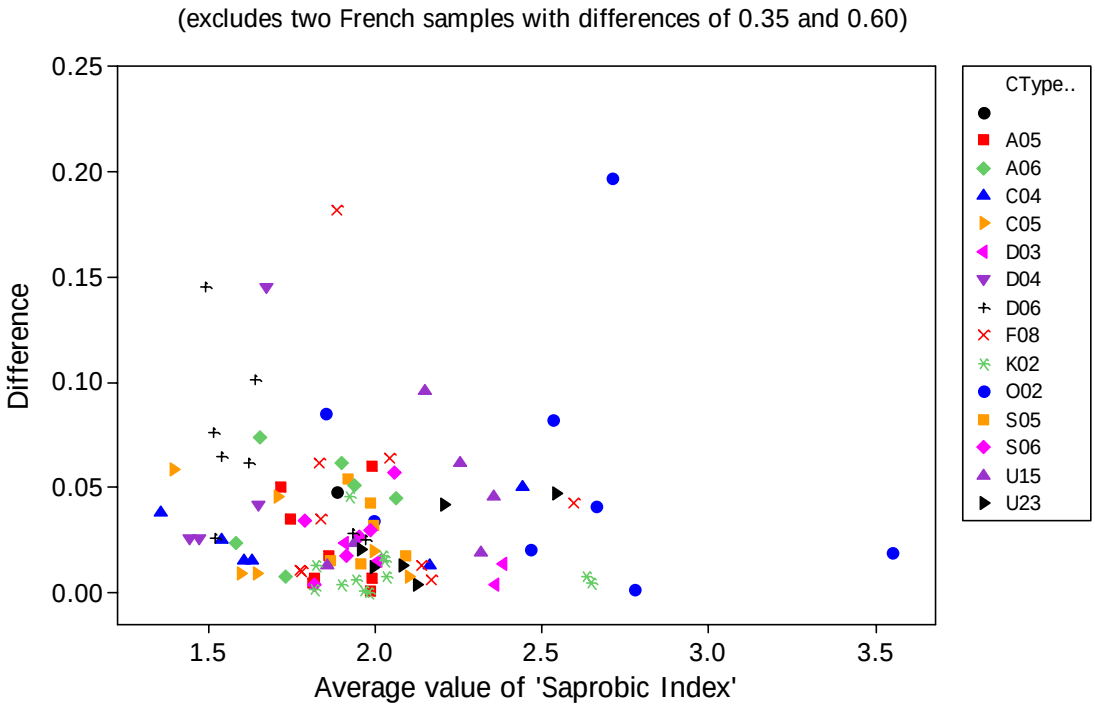


Figure 5.3 STAR-AQEM method: Difference between the two replicate sub-samples plotted against the average of the two values for untransformed values of the metric 'Saprobic Index (Zelinka & Marvan)', for all available STAR sites and seasons with replicated sub-samples. Symbols denote STAR stream type of each site.
'.

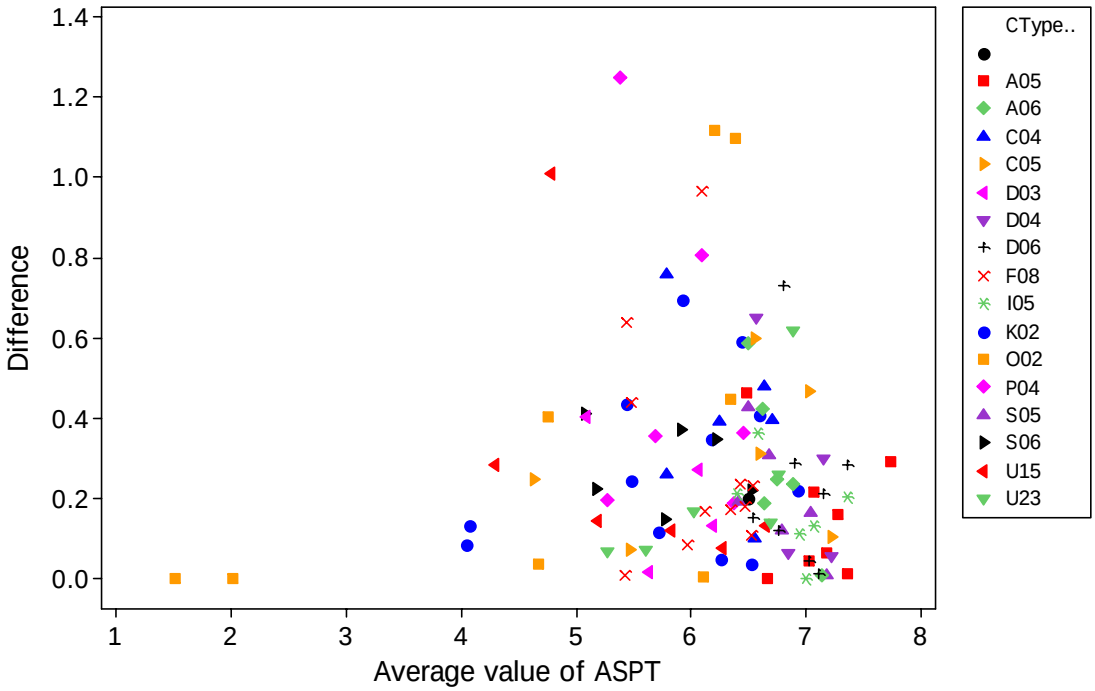


Figure 5.4 STAR-AQEM method: Difference between two replicate sub-samples plotted against the average of the two values for untransformed values of the metric 'Average Score per Taxon (ASPT)

.

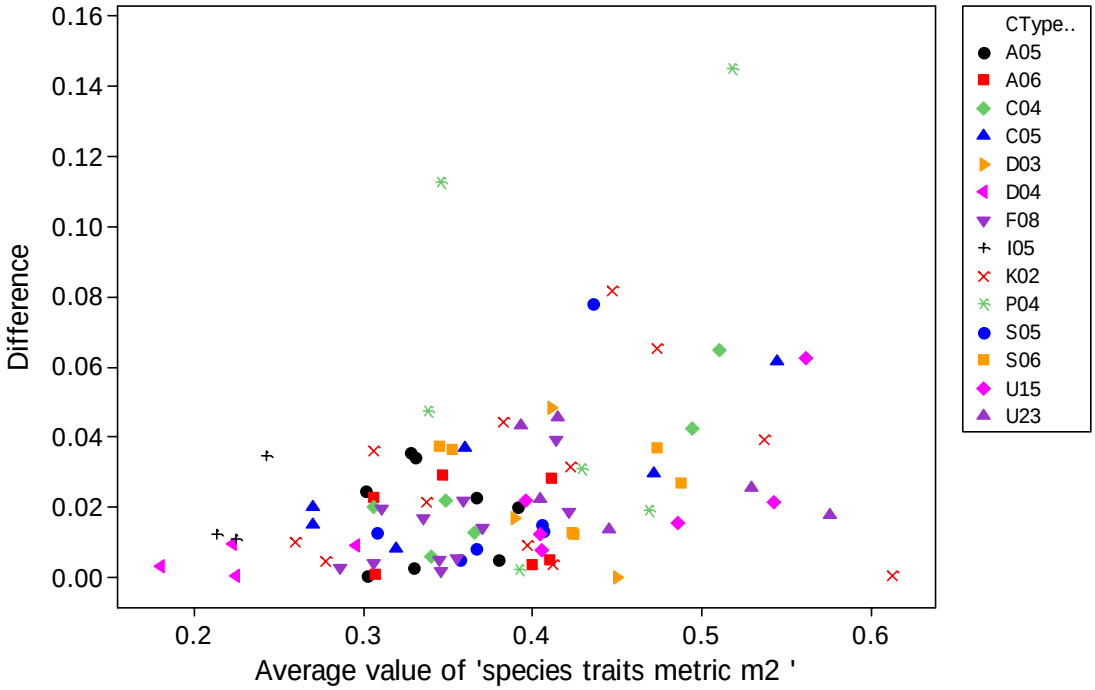


Figure 5.5 STAR-AQEM method: Difference between the two replicate sub-samples plotted against the average of the two values for untransformed values of traits metric 'Trait m2: % individuals with >1 reproductive cycle per year' for all available STAR sites and seasons with replicated sub-samples. Symbols denote STAR stream type of each site

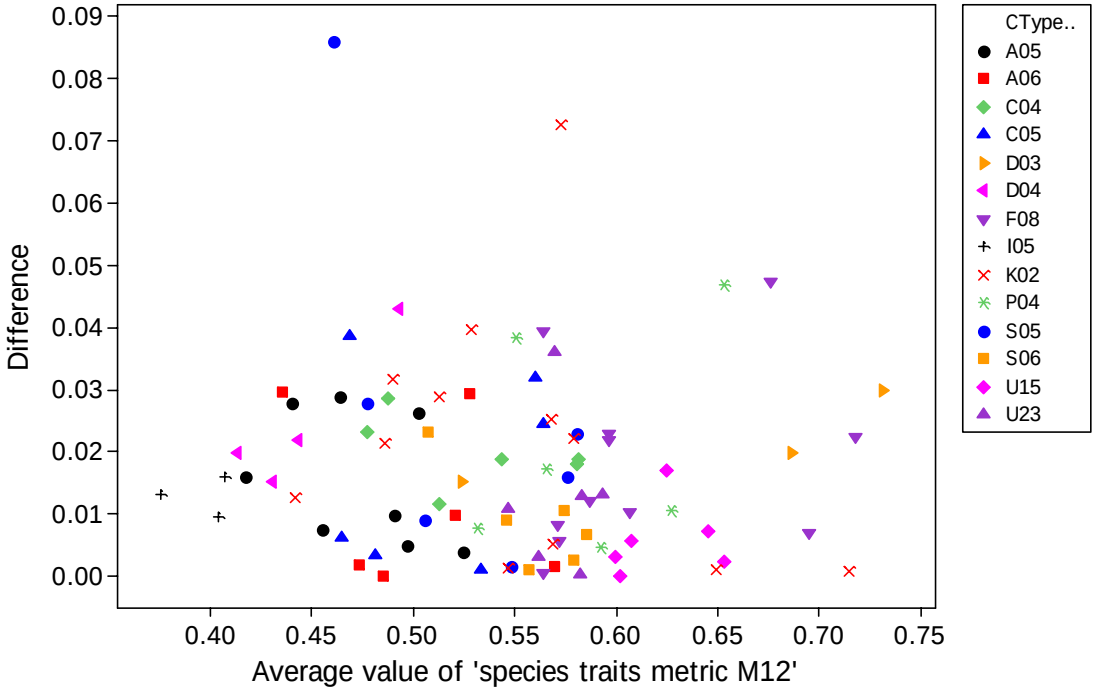


Figure 5.6 STAR-AQEM method: Difference between two replicate sub-samples plotted against the average of the two values for untransformed values of traits metric 'Trait m2: % individuals preferring current velocity < 25cm/s)'.

Table 5.3 STAR-AQEM method: Estimate of the standard deviation (SD_U) in (untransformed) metric values due to sub-sampling, separately for each STAR stream type. Missing values indicate where metric values were not available or appropriate.

	Stream Type															
Metric	A05	A06	C04	C05	D03	D04	D06	F08	I05	K02	O02	P04	S05	S06	U15	U23
Abundance [ind/m²]	538	464	862	625	319	611	609	4160	398	1128	2930	141	893	866	1068	602
Number of taxa	3.32	3.92	3.70	1.61	2.76	5.27	2.99	4.44	1.98	5.00	3.19	2.27	5.66	5.66	1.89	2.48
Number of Families	2.24	2.06	2.90	2.45	1.87	1.90	1.66	2.82	1.94	2.80	1.75	1.50	3.10	1.98	2.38	1.96
Number of EPT taxa	2.21	1.71	2.43	2.48	1.77	3.34	2.52	2.01	0.96	2.50	2.11	1.12	2.14	1.80	1.04	0.96
Saprobic Index	0.022	0.035	0.021	0.023	0.011	0.055	0.054	0.155		0.011	0.059		0.023	0.023	0.037	0.020
German Saprobic new	0.035	0.024	0.076	0.087	0.041	0.051	0.013	0.025		0.046	0.073		0.056	0.040	0.079	0.034
Czech Saprobic	0.047	0.049	0.045	0.034	0.022	0.044	0.049	0.041		0.110	0.088		0.065	0.035	0.084	0.058
ASPT	0.154	0.237	0.315	0.251	0.178	0.255	0.221	0.282	0.145	0.247	0.420	0.461	0.172	0.214	0.311	0.206
IBE	0.304	0.700	0.770	0.548	0.458	0.854	0.357	0.742	0.356	0.735	0.760	0.796	1.098	0.600	0.735	0.404
Diversity SW	0.112	0.066	0.108	0.058	0.162	0.145	0.071	0.107	0.036	0.168	0.101	0.107	0.225	0.102	0.129	0.086
% Rheophilic	2.53	2.79	2.97	1.03	2.45	3.22	5.50	2.97	3.47	2.10	2.14	6.22	9.41	2.89	3.89	5.97
% Rheophilic (ab-class)	2.19	2.58	3.57	2.81	3.24	3.03	2.14	3.90		3.89	3.33	7.41	3.12	2.95	3.00	4.00
% Littoral	1.71	1.46	1.47	1.38	0.62	1.69	1.63	3.28		1.42	1.10	2.96	5.77	1.41	0.86	2.77
% Grazers/Scrapers	2.20	1.42	2.02	0.35	0.85	3.68	1.53	2.07	1.27	2.21	1.10	4.96	7.12	1.55	0.74	1.96
% Shredders	0.91	1.00	1.64	0.70	2.32	0.67	1.46	1.40		1.56	0.59	0.09	1.40	1.15	1.92	0.99
% Gatherers/Collectors	1.33	1.08	1.26	0.95	0.45	1.47	0.79	3.59		3.09	3.28	6.74	5.38	0.90	2.32	1.29
% Oligochaeta	1.26	1.57	1.57	0.44	0.39	0.20	0.22	3.02	0.06	2.06	4.76	2.30	0.74	0.83	2.20	3.12
% EPT individuals	2.94	1.97	3.86	1.32	3.40	4.06	2.10	2.57	3.91	2.19	1.06	3.57	16.36	3.22	1.04	2.79
% EPT (ab-class)	2.33	0.92	2.54	5.78	4.73	1.63	1.40	3.44		2.37	2.98	4.15	3.47	2.59	2.53	3.65
% EPT Taxa	3.23	1.98	4.60	7.97	5.22	1.48	2.51	6.29	3.42	3.44	4.05	6.00	2.19	2.67	3.26	4.27
RETI	0.020	0.016	0.015	0.015	0.022	0.019	0.016	0.030	0.013	0.026	0.017	0.068	0.086	0.019	0.021	0.012
Log(Sel_EPTD+1)			0.081	0.086												
1 –GOLD	0.034	0.018	0.034	0.010	0.030	0.026	0.018	0.038	0.030	0.036	0.027	0.036	0.176	0.043	0.032	0.034
Trait m1:max size ≤1cm	0.018	0.019	0.023	0.022	0.028	0.012		0.017	0.011	0.023		0.026	0.028	0.010	0.008	0.018
Trait m2 : >1 cycle	0.016	0.014	0.024	0.024	0.021	0.005		0.012	0.016	0.027		0.056	0.024	0.021	0.021	0.022
Trait m7 : crawler loco.	0.011	0.016	0.029	0.019	0.026	0.010		0.010	0.008	0.020		0.036	0.012	0.013	0.014	0.015
Trait m12:current<25cm	0.013	0.012	0.015	0.016	0.016	0.019		0.016	0.009	0.021		0.019	0.027	0.008	0.006	0.012

Table 5.4 STAR-AQEM method: Estimate of the standard deviation (SD_U) in transformed (f(x) metric values due to sub-sampling, separately for each STAR stream type and the median (Med) values across stream types.

	Stream Type																
Metric	f(x)	A05	A06	C04	C05	D03	D04	D06	F08	I05	K02	O02	P04	S05	S06	U15	U23	Med
Abundance [ind/m²]	√√x	0.320	0.243	0.256	0.229	0.281	0.425	0.288	0.804	0.134	0.390	0.544	0.129	0.928	0.347	0.486	0.282	0.304
Number of taxa	√x	0.212	0.255	0.298	0.149	0.216	0.356	0.205	0.338	0.195	0.422	0.257	0.250	0.460	0.404	0.150	0.201	0.253
Number of Families	√x	0.203	0.202	0.273	0.251	0.217	0.188	0.155	0.272	0.204	0.286	0.176	0.214	0.321	0.191	0.244	0.226	0.216
Number of EPT taxa	√x	0.261	0.212	0.335	0.283	0.245	0.320	0.230	0.288	0.138	0.348	0.280	0.356	0.260	0.213	0.232	0.134	0.261
Saprobic Index	x	0.022	0.035	0.021	0.023	0.011	0.055	0.054	0.155		0.110	0.059		0.023	0.023	0.037	0.020	0.023
German Saprobic new	x	0.035	0.024	0.076	0.087	0.041	0.051	0.013	0.025		0.046	0.073		0.056	0.040	0.079	0.034	0.041
Czech Saprobic	x	0.047	0.049	0.045	0.034	0.022	0.044	0.049	0.041		0.110	0.088		0.065	0.034	0.084	0.058	0.046
ASPT	x	0.154	0.237	0.315	0.251	0.178	0.255	0.221	0.282	0.145	0.247	0.420	0.461	0.172	0.214	0.311	0.206	0.242
IBE	x	0.304	0.700	0.770	0.548	0.458	0.854	0.357	0.742	0.356	0.735	0.760	0.796	1.098	0.600	0.735	0.404	0.718
Diversity SW	x	0.112	0.066	0.108	0.058	0.162	0.145	0.071	0.107	0.036	0.168	0.101	0.107	0.225	0.102	0.129	0.086	0.107
% Rheophilic	asin	0.033	0.029	0.032	0.013	0.026	0.036	0.058	0.034	0.036	0.036	0.031	0.090	0.100	0.030	0.042	0.070	0.035
% Rheophilic (ab-class)	asin	0.022	0.026	0.037	0.028	0.038	0.030	0.022	0.046		0.049	0.035	0.077	0.031	0.031	0.031	0.040	0.031
% Littoral	asin	0.018	0.015	0.017	0.015	0.014	0.017	0.016	0.037		0.019	0.015	0.034	0.069	0.015	0.010	0.031	0.017
% Grazers/Scrapers	asin	0.023	0.016	0.022	0.004	0.019	0.040	0.017	0.028	0.013	0.025	0.022	0.067	0.084	0.016	0.009	0.023	0.022
% Shredders	asin	0.022	0.014	0.019	0.011	0.023	0.013	0.018	0.019		0.025	0.011	0.007	0.044	0.017	0.022	0.015	0.018
% Gatherers/Collectors	asin	0.014	0.012	0.013	0.010	0.005	0.017	0.010	0.044		0.036	0.043	0.073	0.059	0.011	0.026	0.013	0.014
% Oligochaeta	asin	0.025	0.029	0.019	0.016	0.025	0.027	0.026	0.042	0.011	0.030	0.069	0.069	0.022	0.025	0.039	0.037	0.027
% EPT individuals	asin	0.035	0.022	0.046	0.014	0.048	0.042	0.022	0.032	0.042	0.026	0.014	0.044	0.178	0.036	0.017	0.033	0.034
% EPT (ab-class)	asin	0.028	0.010	0.029	0.059	0.052	0.017	0.014	0.039		0.026	0.033	0.088	0.035	0.027	0.032	0.038	0.032
% EPT Taxa	asin	0.037	0.022	0.049	0.081	0.057	0.015	0.025	0.070	0.034	0.037	0.043	0.117	0.022	0.028	0.042	0.044	0.040
RETI	asin	0.020	0.017	0.016	0.015	0.022	0.020	0.018	0.035	0.014	0.029	0.025	0.081	0.094	0.020	0.023	0.013	0.020
Log(Sel_EPTD+1)	x			0.081	0.086													0.084
1 –GOLD	asin			0.038	0.011													0.025
Trait m1:max size ≤1cm	asin	0.018	0.019	0.023	0.022	0.028	0.012		0.017	0.011	0.023		0.031	0.030	0.010	0.008	0.018	0.019
Trait m2 : >1 cycle	asin	0.017	0.014	0.025	0.024	0.021	0.006		0.013	0.018	0.027		0.057	0.024	0.021	0.021	0.022	0.021
Trait m7 : crawler loco.	asin	0.011	0.016	0.029	0.019	0.026	0.010		0.010	0.008	0.020		0.038	0.012	0.013	0.014	0.015	0.015
Trait m12:current<25cm	asin	0.013	0.012	0.015	0.016	0.017	0.019		0.017	0.009	0.021		0.019	0.028	0.008	0.006	0.012	0.016

5.4.2 Relative importance of sub-sampling variation to field sampling variation with the STAR-AQEM method

A STAR-AQEM sample is based on 20 sampling units taken in proportion to the estimated percentage cover of each major habitat type at the site. However, the taxonomic composition of replicate field samples will still vary because of small-scale spatial heterogeneity in habitat and patchiness in macroinvertebrate distribution and density within a site. Different samples will involve taking sampling units from different locations within the site. These differences are likely to be even greater when the replicate samples are taken completely independently by different personnel. Thus there will be real differences in taxonomic composition and derived metric values between replicate field samples. With STAR-AQEM field samples, there is a separately identifiable source of variation in metric values due to sub-sampling and only identifying and counting the macroinvertebrate individuals in a fraction (minimum one-sixth) of the whole sample. Because of the replicate sub-sampling and replicated field sampling design in the STAR filed sampling programme, it is possible to separate and estimate the variances in metric values due to field sampling (σ_R^2) and sub-sampling (σ_U^2). It is important to be able to quantify the relative importance of STAR-AQEM sub-sampling on the overall variance in metric values obtained for replicate samples. The relative influence of sub-sampling effects were expressed by the percentage ($P_{sub} = 100\sigma_U^2 / \sigma_E^2$) of the overall variance ($\sigma_E^2 = \sigma_U^2 + \sigma_R^2$) in metric values between replicate field samples which is due specifically to sub-sampling variation. Estimates were calculated completely independntly for each stream type, but because replicate STAR-AQEM sub-samples were usually taken at six or less sites within any one stream type, individual estimates are likely to be imprecise and volatile. Therefore, the more statistically robust median value of P_{sub} across all stream types is also given for each metric (Table 5.5).

STAR- AQEM sub-sampling variation causes a relatively large part of the overall variance between replicate sample values for many metrics, and is estimated on average (see median values in Table 5.5) to contribute more than 50% of the overall variance between replicate samples for 12 of the 27 metrics analysed. In general, sub-sampling variance has a large effect on those metrics which are based on the number of taxa present, such as number of families and number of EPT taxa.

Sub-sampling variation is estimated to be responsible for more than half of the overall replicate sampling variance in ASPT (Average Score Per Taxon) for the vast majority of stream types (Table 5.5). ASPT only depends on the presence, rather than abundance, of BMWP (Biological Monitoring Working Party) families. Several taxa in the sample at very low abundances may be found in one sub-sample but not another. The metric '% EPT taxa' measuring the percentage of all taxa which belong to the EPT group (Emphemeroptera + Plecoptera + Trichoptera), which is also only dependent on the presence of each taxon, is also relatively variable between sub-samples. Given that ASPT is an important component of the proposed Inter-calibration Common Metrics (ICMs) (Buffagni *et al.* 2004), this may merit further investigation for consistency across individual stream types.

The metrics based on relative abundance (i.e. percentage composition) of one of more taxonomic groups seem to be less prone to the effects of sub-sampling with less than one third of overall sampling variability in metric values usually due to sub-sampling.

Table 5.5 STAR-AQEM method: Estimates of the percentage (P_{sub}) of the overall variance between replicate samples which is due to sub-sampling for transformed (f(x) metric values, separately for each STAR stream type and the median (Med) values across stream types.

		Stream Type																
Metric	f(x)	A05	A06	C04	C05	D03	D04	D06	F08	I05	K02	O02	P04	S05	S06	U15	U23	Med
Abundance [ind/m²]	√√x	15	18	14	9	100	100	51	49	6	40	38	3	100	5	77	20	29
Number of Taxa	√x	12	50	72	13	30	100	100	87	95	60	54	46	80	100	5	29	57
Number of Families	√x	44	40	100	65	100	74	83	61	100	52	94	41	71	72	38	93	72
Number of EPT Taxa	√x	29	32	100	100	30	93	100	75	53	50	100	100	58	100	42	100	84
Saprobic Index	x	50	46	14	15	15	100	100	50		94	98		30	37	13	39	42
German Saprobic new	x	100	46	66	100	100	100	37	6		100	65		100	74	75	28	74
Czech Saprobic	x	66	25	8	5	10	90	41	18		100	16		100	13	32	5	21
ASPT	x	20	86	100	100	26	100	75	100	25	75	100	76	28	100	45	52	76
IBE	x	41	82	100	50	49	100	85	54	57	63	97	100	76	25	63	25	63
Diversity SW	x	50	15	16	12	43	52	20	44	4	16	20	9	78	21	25	12	20
% Rheophilic	asin	39	17	12	2	6	51	100	9	35	6	1	27	100	12	13	9	12
% Rheophilic (ab-class)	asin	60	37	60	43	63	80	40	61		66	2	100	23	75	79	53	60
% Littoral	asin	12	4	15	10	7	16	16	20		4	4	15	100	24	3	8	12
% Grazers/Scrapers	asin	36	16	20	1	13	57	24	60	10	10	12	95	100	9	4	4	14
% Shredders	asin	52	14	26	2	4	91	40	20		18	7	4	83	18	9	5	18
% Gatherers/Collectors	asin	8	16	3	5	0	47	34	37		10	20	48	100	5	9	1	10
% Oligochaeta	asin	15	77	2	16	66	99	65	14	38	3	9	32	50	26	35	13	29
% EPT individuals	asin	29	11	28	3	59	45	12	9	25	4	9	24	100	13	26	4	18
% EPT (ab-class)	asin	52	6	57	100	59	42	35	68		21	100	100	79	56	46	53	56
% EPT Taxa	asin	56	19	82	100	91	17	50	100	70	28	100	100	65	45	65	100	68
RETI	asin	12	14	6	1	3	15	100	61	10	7	10	100	100	24	6	1	11
Log(Sel_EPTD+1)	x			34	12													23
1 –GOLD	asin			45	1													23
Trait m1:max size ≤1cm	asin	90	56	39	48	34	23		14	27	36		73	53	45	27	69	42
Trait m2 : >1 cycle	asin	35	47	76	100	92	16		50	53	45		100	83	100	29	100	64
Trait m7 : crawler loco.	asin	23	80	100	85	100	37		24	31	39		100	64	82	27	100	72
Trait m12:current<25cm	asin	63	75	78	100	100	91		40	90	87		39	100	27	6	64	76

In summary, sub-sampling variation is a major or non-negligible part of the overall replicate sample variability in many commonly used metrics. Sorting and identifying a larger fraction of the sample would reduce this source of variation; in the extreme, sorting the whole sample would eliminate it. However, all extra identification increases costs. It is only possible to determine the cost effectiveness of extra sub-sampling effort by sorting all 30 tray cells of a STAR-AQEM sample and doing repeated computerised random combination of increasing numbers of cells macroinvertebrates to assess the rate of reduction in sub-sorting variance.

These results also highlight the importance of always trying to spread and distribute the sample material as evenly as possible amongst the 30 grid cells on a sorting tray for any STAR-AQEM macroinvertebrate sample.

Although sub-sampling contributes a major part of the overall inter-replicate variance in numerous metrics, overall inter-replicate variance may still be small compared to the range in metric values amongst sites of varying quality and thus such metrics may still have high precision to detect differences between sites. The practical size and importance of overall replicate variability is assessed below.

5.4.3 Overall replicate sampling variability for the STAR-AQEM method

The combined effect of field sampling spatial variability and subsequent laboratory sub-sampling of STAR-AQEM samples determines their overall variability in metrics values for a site at any point in time. It is the overall replicate sampling variability in a metric's values which determine it uncertainty and precision in site bioassessments.

Figures 5.7 – 5.12 show the pattern of estimates of the overall replicate sampling SD (SD_E) in selected metrics for individual sites (and seasons) in relation to the average of the replicate values at that site and season. The individual estimates of sampling SD are based on just two sample values (i.e. excluding the second replicate sub-sample for the samples with two sub-samples). Both 'Number of taxa' and 'Number of EPT taxa' show a tendency for sampling SD to be higher on sites with more taxa (Figures 5.7-5.8), supporting the justification for transforming the metrics to a square root scale before estimating a single average sampling SD for all sites in a stream type or across stream types.

Sampling variability in the percentage of the total abundance in a sample comprised by key taxonomic or feeding groups also tended to increase with the relative abundance of that group at the site (Table 5.2). For example, SD_E for the metric '% Shredders' (percentage of all individuals which feed by shredding plant matter and detritus) increases with the relative abundance of 'Shredders' (Figure 5.9(a), Spearman rank correlation $r_S = 0.74$); an arcsine transformation of the metric values reduces the systematic pattern to the uncertainty and the correlation between replicate SD and replicate mean (Figure 5.9(b), $r_S = 0.27$).

Amongst the four species trait metrics analysed, only 'Trait m2' (related to the proportion of individuals from species with more than one reproductive cycle per year) showed any tendency to have higher sampling variability with higher site values – although this was more noticeable for metric values based on the RIVPACS sampling method (Table 5.2). Species trait metric 'Trait m1' based on body size appears to be more susceptible to sampling variation than species trait metric 'Trait m12' based on water velocity preference – this is assessed further below.

.

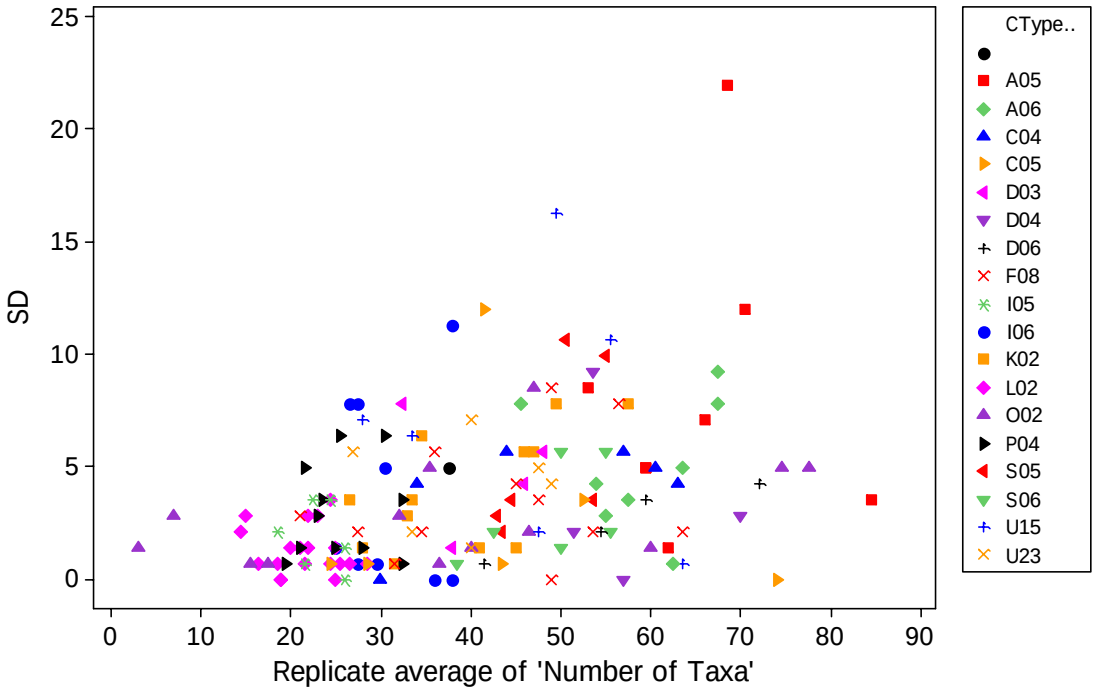


Figure 5.7 STAR-AQEM method: Replicate sample SD in relation to replicate mean for untransformed values of the metric 'Number of taxa' for all available STAR sites and seasons with replicated sub-samples. Symbols denote STAR stream type of each site

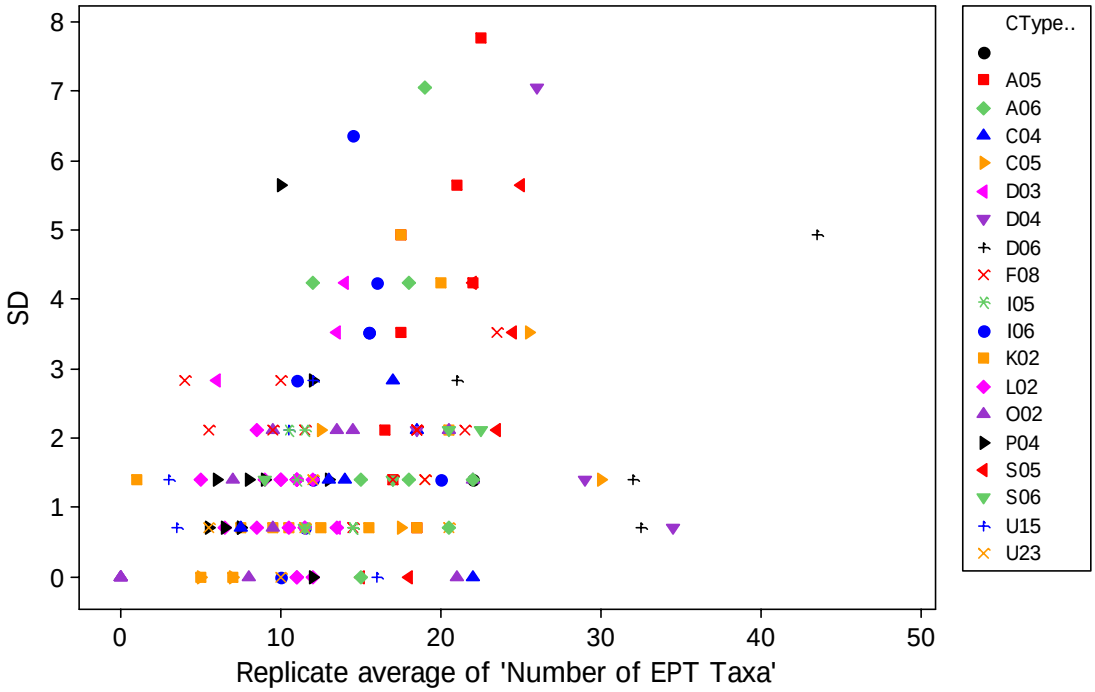


Figure 5.8 STAR-AQEM method: Replicate sample SD in relation to replicate mean for the metric 'Number of EPT taxa'.

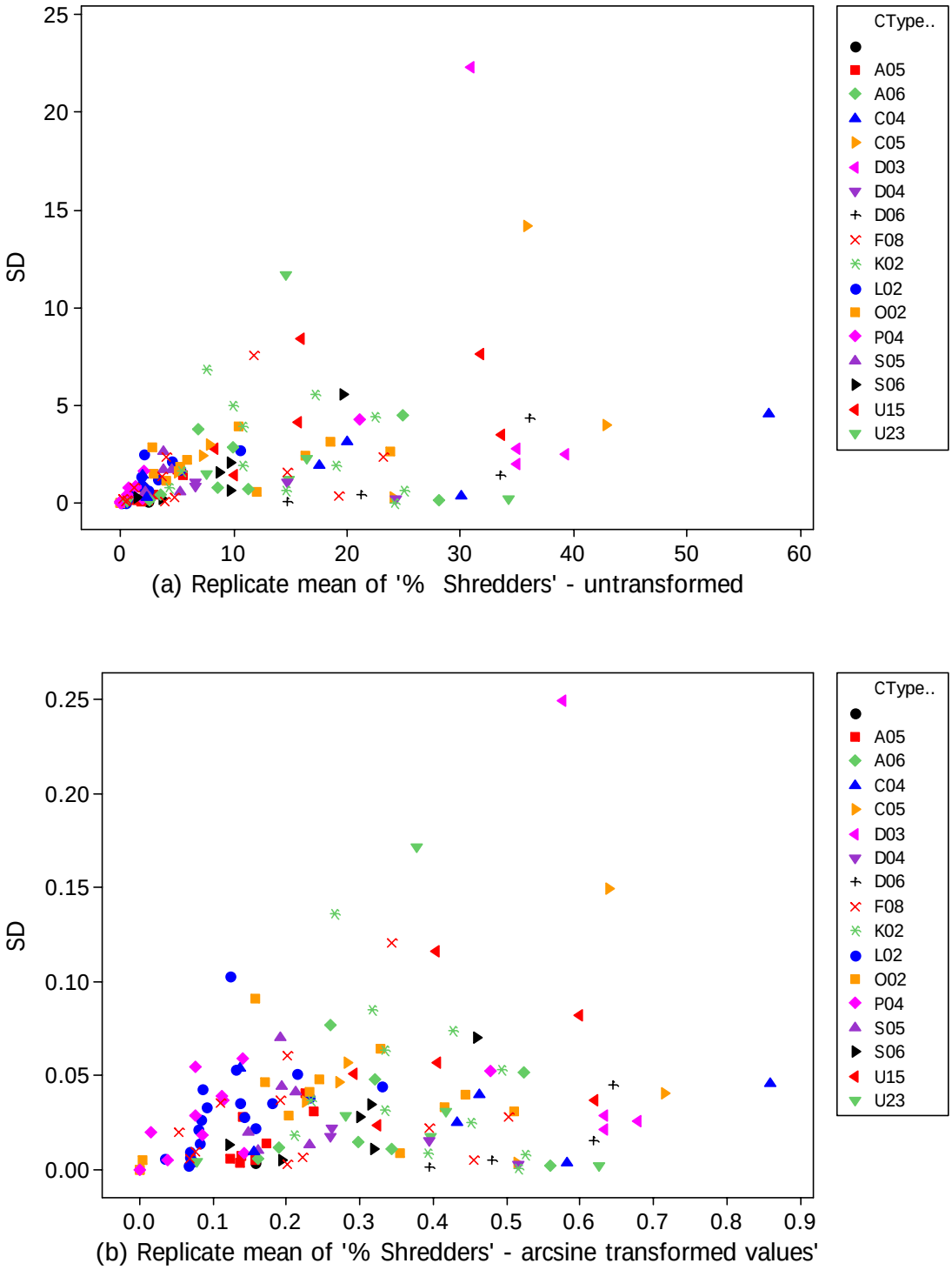


Figure 5.9 STAR-AQEM method: Replicate sample SD in relation to replicate mean for the metric '% Shredders': using (a) untransformed values and (b) arcsine transformed values of the metric. Symbols denote STAR stream type of each site.

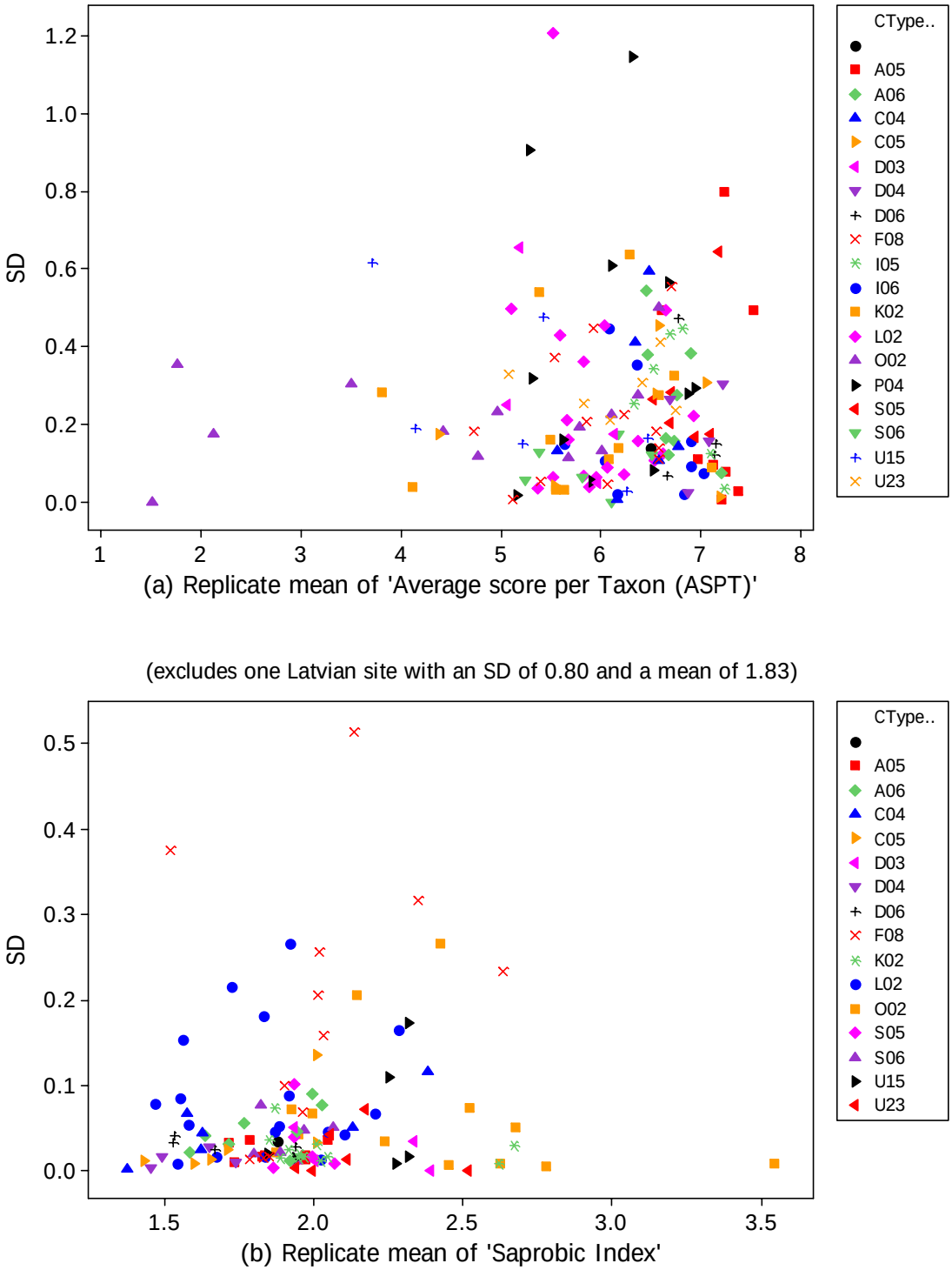


Figure 5.10 STAR-AQEM method: Replicate sample SD in relation to replicate mean for untransformed values of the metrics (a) ASPT, and (b) 'Saprobic Index (Zelinka & Marvan)'. Symbols denote STAR stream type of each site.

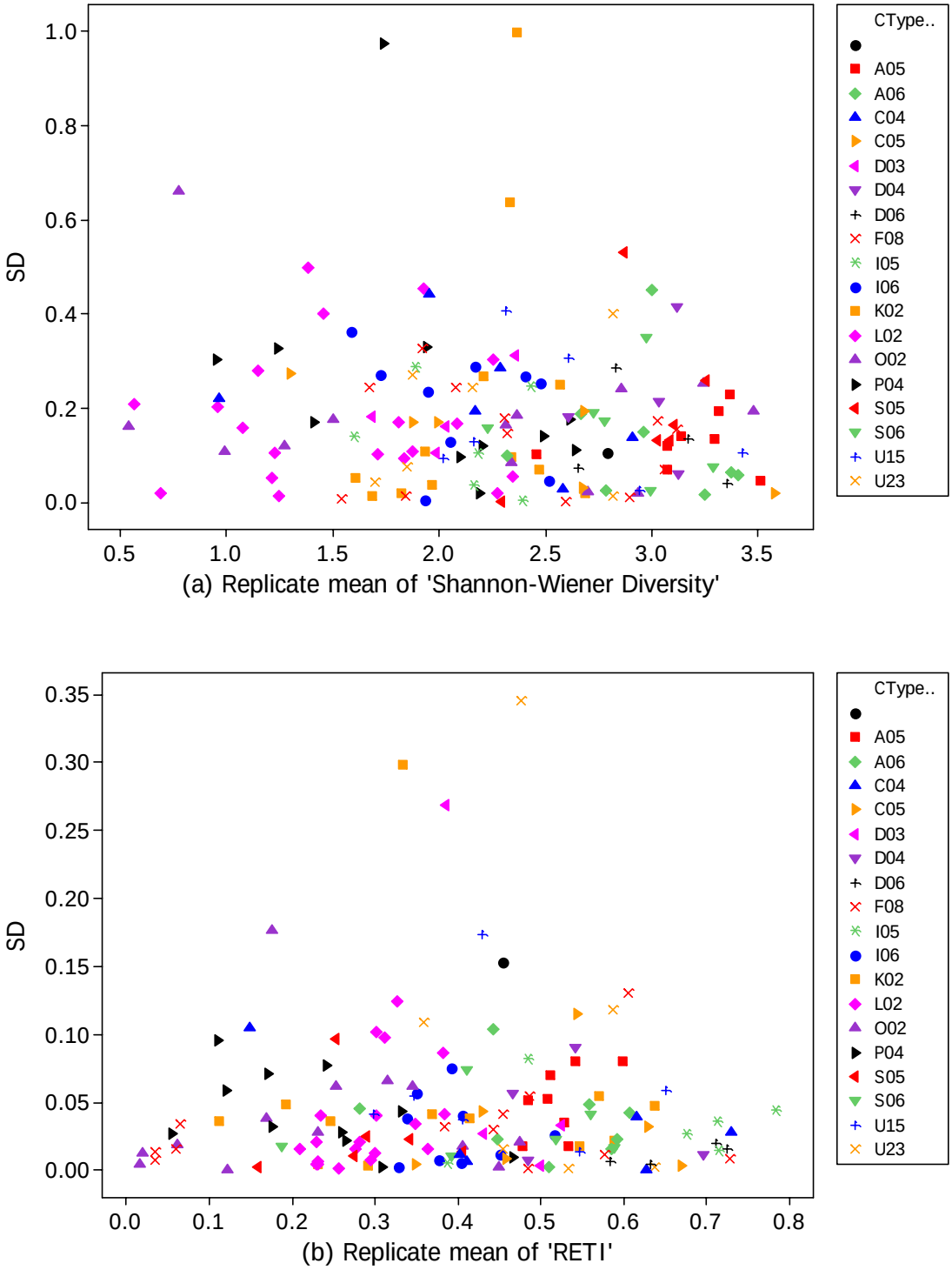


Figure 5.11 STAR-AQEM method: Replicate sample SD in relation to replicate mean for untransformed values of the metrics (a) 'Shannon-Weiner Diversity Index' and (b) 'RETI' (Rhithron Feeding Type Index). Symbols denote STAR stream type of each site.

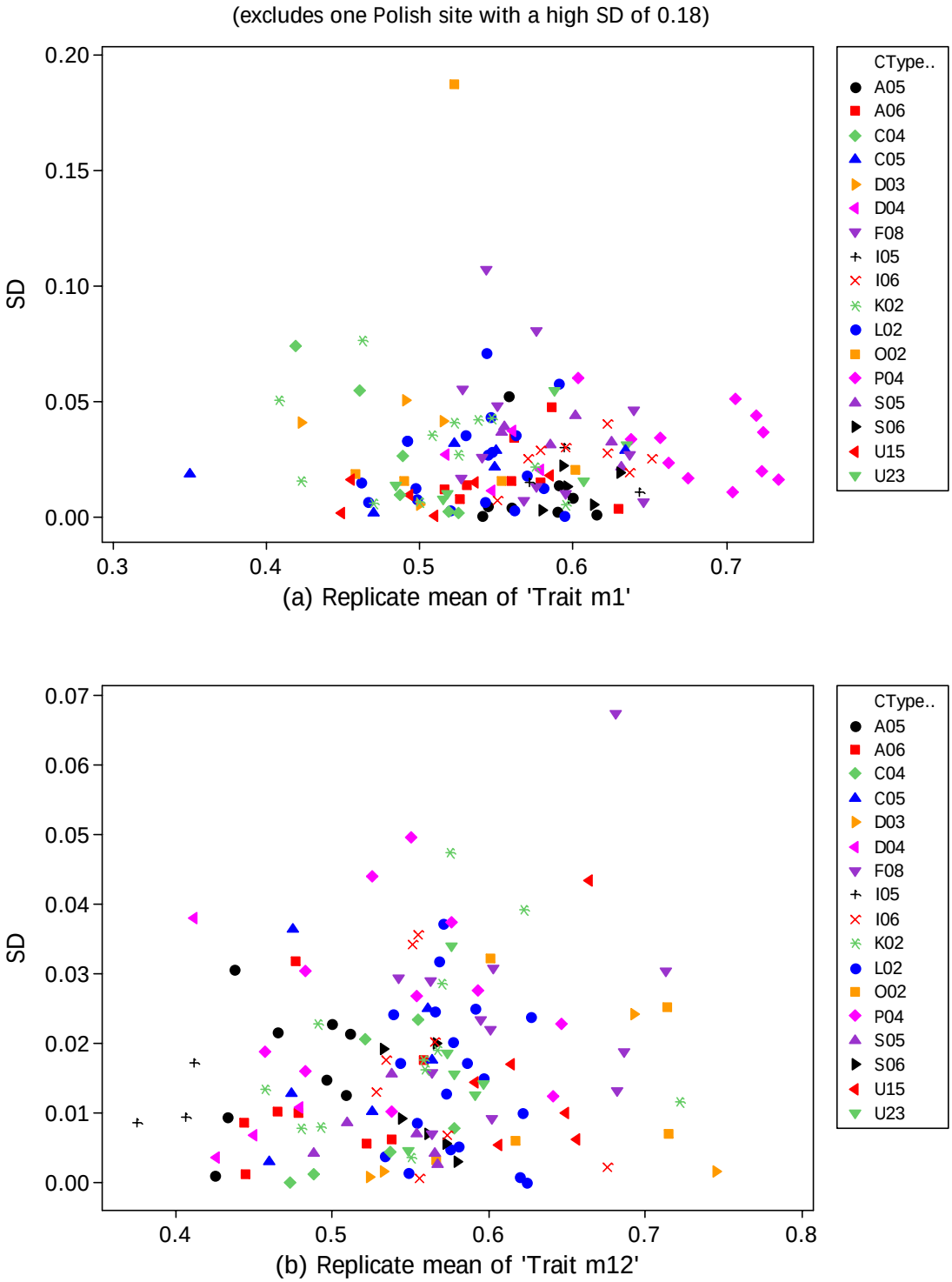


Figure 5.12 STAR-AQEM method: Replicate sample SD in relation to replicate mean for untransformed values of two species traits metrics (a) 'Trait m1 : max body size ≤1cm', and (b) 'Trait m12 : current <25cm/s'. Symbols denote STAR stream type of each site.

Table 5.6 STAR-AQEM method: Estimate of the overall standard deviation (SD_E) in untransformed metric values of replicate samples from sites in each STAR stream type

Metric	Stream Type A05	A06	C04	C05	D03	D04	D06	F08	H04	I06	K02	L02	O02	P04	S05	S06	U15	U23
Abundance [ind/m²]	1360	1113	2043	1615	319	328	1132	3700	79	2572	1704	1722	3429	2134	204	2248	1260	1930
Number of taxa	10.07	5.80	4.56	5.14	5.30	4.92	2.98	4.30	2.55	5.54	4.70	1.66	3.65	3.64	6.43	3.55	8.87	4.67
Number of Families	3.97	2.77	1.71	2.36	1.27	2.60	2.06	3.33	2.70	5.09	2.87	1.55	1.14	2.86	3.45	1.89	4.28	2.57
Number of EPT taxa	4.39	3.40	1.68	1.80	3.12	3.77	2.96	2.24	2.01	3.27	2.08	1.16	1.32	2.13	3.34	1.61	1.96	0.96
Saprobic Index	0.025	0.053	0.062	0.058	0.032	0.017	0.032	0.243			0.030	0.215	0.102		0.045	0.044	0.085	0.035
German Saprobic new	0.019	0.036	0.081	0.063	0.040	0.055	0.021	0.097			0.041	0.159	0.094		0.028	0.044	0.079	0.068
Czech Saprobic	0.046	0.110	0.163	0.159	0.077	0.032	0.054	0.105			0.102	0.185	0.206		0.039	0.095	0.121	0.265
ASPT	0.381	0.303	0.309	0.262	0.363	0.217	0.258	0.265	0.437	0.211	0.293	0.375	0.247	0.534	0.333	0.108	0.339	0.299
IBE	0.485	0.634	0.424	0.858	0.660	0.543	0.255	1.009	0.545	0.897	0.727	0.525	0.869	0.618	1.057	1.314	0.580	0.785
Diversity SW	0.141	0.186	0.254	0.169	0.205	0.253	0.163	0.168	0.272	0.236	0.362	0.230	0.240	0.354	0.262	0.192	0.223	0.224
% Rheophilic	4.65	6.61	8.06	9.66	7.39	5.20	3.57	9.18	11.44	10.09	12.64	7.54	20.09	15.45	5.58	7.86	10.80	22.20
% Rheophilic (ab-class)	2.84	4.28	3.52	3.74	4.64	3.37	2.41	5.21	8.50		5.31	3.97	19.85	6.40	4.61	3.01	2.75	6.71
% Littoral	5.12	7.97	3.80	3.72	3.46	3.29	4.31	7.78	1.97		7.24	6.28	3.56	8.56	2.67	3.01	5.31	8.93
% Grazers/Scrapers	4.61	3.54	3.14	2.95	2.60	4.06	3.35	3.09	4.38	4.93	5.93	3.52	4.56	3.81	2.15	4.54	3.29	9.99
% Shredders	0.86	2.35	2.47	6.27	11.37	0.89	2.30	2.48	1.47		3.53	1.21	2.12	1.46	1.50	2.52	5.29	4.92
% Gatherers/Collectors	4.67	2.66	6.72	3.95	6.74	2.26	0.97	7.04	7.02		10.09	3.46	7.84	9.32	3.27	4.28	8.54	10.78
% Oligochaeta	3.69	1.60	11.23	1.05	0.42	0.11	0.17	6.97	0.22	1.92	13.64	4.24	17.31	6.13	0.94	1.88	4.50	9.31
% EPT individuals	5.53	5.62	5.84	6.50	2.24	6.48	5.98	9.38	8.16	9.11	11.64	7.64	3.81	7.35	12.35	6.81	2.72	15.13
% EPT (ab-class)	3.24	3.70	3.05	5.24	4.99	2.98	2.24	3.77	8.04		3.38	3.41	2.88	5.75	4.29	2.83	2.94	5.21
% EPT Taxa	4.30	5.02	5.51	6.29	4.68	3.97	4.31	5.40	9.83	4.85	3.80	3.28	2.87	6.44	2.98	3.66	3.61	3.21
RETI	0.056	0.048	0.047	0.052	0.136	0.054	0.013	0.046	0.044	0.038	0.093	0.052	0.060	0.051	0.043	0.038	0.081	0.156
Log(Sel_EPTD+1)	0.182		0.114	0.268			0.127	0.277	0.354	0.333	0.215			0.439	0.186			0.087
1 –GOLD	0.078	0.056	0.075	0.093	0.032	0.083	0.059	0.108	0.078	0.073	0.125	0.089	0.057	0.075	0.123	0.100	0.152	0.216
Trait m1:max size ≤1cm	0.019	0.023	0.039	0.024	0.039	0.026		0.048	0.040	0.027	0.037	0.030	0.085	0.035	0.035	0.014	0.012	0.028
Trait m2 : >1 cycle	0.027	0.018	0.022	0.016	0.017	0.013		0.018	0.034	0.015	0.036	0.026	0.095	0.040	0.017	0.018	0.035	0.024
Trait m7 : crawler loco.	0.026	0.013	0.023	0.018	0.018	0.013		0.022	0.034	0.019	0.034	0.016	0.074	0.029	0.013	0.010	0.021	0.016
Trait m12:current<25cm	0.019	0.014	0.013	0.021	0.012	0.020		0.029	0.039	0.021	0.023	0.018	0.019	0.030	0.008	0.013	0.021	0.019

Table 5.7 STAR-AQEM method: Estimate of the overall standard deviation (SD_E) in transformed (f(x)) metric values due to sampling, separately for sites in each STAR stream type and the median (Med) values across stream types.

Metric	Stream Type f(x)	A05	A06	C04	C05	D03	D04	D06	F08	H04	I06	K02	L02	O02	P04	S05	S06	U15	U23	Med
Abundance [ind/m²]	√√x	0.746	0.511	0.659	0.656	0.258	0.222	0.529	1.185	0.257	0.995	0.673	0.676	0.984	0.711	0.311	1.498	0.430	0.622	0.659
Number of taxa	√x	0.620	0.378	0.324	0.398	0.432	0.334	0.186	0.321	0.282	0.496	0.357	0.185	0.317	0.358	0.449	0.247	0.664	0.383	0.358
Number of Families	√x	0.352	0.261	0.166	0.250	0.144	0.258	0.191	0.314	0.299	0.478	0.300	0.175	0.172	0.331	0.317	0.177	0.406	0.274	0.261
Number of EPT taxa	√x	0.489	0.417	0.212	0.199	0.477	0.377	0.253	0.358	0.343	0.430	0.377	0.191	0.185	0.346	0.342	0.196	0.333	0.134	0.342
Saprobic Index	x	0.025	0.053	0.062	0.058	0.032	0.017	0.032	0.243			0.030	0.215	0.102		0.045	0.044	0.085	0.035	0.049
German Saprobic new	x	0.019	0.036	0.081	0.063	0.040	0.055	0.021	0.097			0.041	0.159	0.094		0.028	0.044	0.079	0.068	0.063
Czech Saprobic	x	0.046	0.110	0.163	0.159	0.077	0.032	0.054	0.105			0.102	0.185	0.206		0.039	0.095	0.121	0.265	0.105
ASPT	x	0.381	0.303	0.309	0.262	0.363	0.217	0.258	0.265	0.437	0.211	0.293	0.375	0.247	0.534	0.333	0.108	0.339	0.299	0.299
IBE	x	0.485	0.634	0.424	0.858	0.660	0.543	0.255	1.009	0.545	0.897	0.727	0.525	0.869	0.618	1.057	1.314	0.580	0.785	0.660
Diversity SW	x	0.141	0.186	0.254	0.169	0.205	0.253	0.163	0.168	0.272	0.236	0.362	0.230	0.240	0.354	0.262	0.192	0.223	0.224	0.224
% Rheophilic	asin	0.053	0.079	0.096	0.100	0.105	0.056	0.038	0.109	0.146	0.130	0.143	0.092	0.313	0.176	0.062	0.080	0.112	0.244	0.100
% Rheophilic (ab-class)	asin	0.029	0.044	0.036	0.038	0.054	0.034	0.025	0.065	0.096		0.057	0.042	0.311	0.068	0.051	0.032	0.030	0.069	0.043
% Littoral	asin	0.052	0.081	0.048	0.048	0.054	0.034	0.044	0.090	0.022		0.085	0.079	0.084	0.089	0.033	0.033	0.061	0.116	0.058
% Grazers/Scrapers	asin	0.047	0.040	0.050	0.033	0.051	0.044	0.038	0.041	0.052	0.063	0.077	0.042	0.070	0.067	0.029	0.051	0.041	0.112	0.047
% Shredders	asin	0.021	0.038	0.035	0.072	0.127	0.016	0.024	0.043	0.053		0.059	0.038	0.042	0.035	0.040	0.035	0.069	0.073	0.039
% Gatherers/Collectors	asin	0.049	0.030	0.075	0.046	0.090	0.025	0.012	0.079	0.075		0.109	0.036	0.102	0.104	0.034	0.050	0.089	0.112	0.063
% Oligochaeta	asin	0.067	0.028	0.120	0.036	0.037	0.020	0.029	0.112	0.028	0.057	0.158	0.063	0.248	0.121	0.030	0.046	0.059	0.108	0.059
% EPT individuals	asin	0.066	0.066	0.080	0.079	0.047	0.068	0.062	0.098	0.094	0.095	0.126	0.086	0.049	0.094	0.130	0.097	0.039	0.174	0.080
% EPT (ab-class)	asin	0.036	0.040	0.033	0.054	0.061	0.031	0.022	0.047	0.083		0.055	0.036	0.031	0.060	0.043	0.033	0.038	0.054	0.039
% EPT Taxa	asin	0.048	0.054	0.058	0.065	0.054	0.041	0.043	0.064	0.100	0.049	0.067	0.034	0.032	0.068	0.030	0.040	0.045	0.033	0.048
RETI	asin	0.057	0.049	0.066	0.053	0.145	0.054	0.015	0.053	0.046	0.039	0.104	0.056	0.081	0.075	0.049	0.039	0.084	0.163	0.056
Log(Sel_EPTD+1)	x	0.182		0.114	0.268			0.127	0.277	0.354	0.333	0.215			0.439	0.186			0.087	0.182
1 –GOLD	asin	0.084	0.061	0.091	0.100	0.040	0.088	0.077	0.117	0.094	0.076	0.135	0.099	0.097	0.093	0.133	0.103	0.165	0.241	0.097
Trait m1:max size ≤1cm	asin	0.020	0.023	0.040	0.024	0.039	0.026		0.048	0.042	0.028	0.037	0.031	0.086	0.037	0.036	0.014	0.012	0.028	0.030
Trait m2 : >1 cycle	asin	0.028	0.020	0.023	0.017	0.018	0.016		0.019	0.037	0.017	0.038	0.028	0.102	0.040	0.018	0.019	0.035	0.024	0.022
Trait m7 : crawler loco.	asin	0.026	0.013	0.023	0.019	0.018	0.013		0.022	0.035	0.020	0.035	0.016	0.083	0.029	0.013	0.010	0.021	0.016	0.020
Trait m12:current<25cm	asin	0.019	0.015	0.013	0.021	0.013	0.020		0.031	0.039	0.021	0.024	0.019	0.020	0.030	0.008	0.013	0.022	0.019	0.020

The replicate sampling SD for both ASPT and the Saprobic index (Zelinka & Marvan) do not vary systematically with the replicate mean value for a site (Figure 5.10). Although the Spearman rank correlation between SD and replicate mean is very weak, there is some suggestion that the very low quality sites with very low ASPT values (< 3) may be less variable between samples, but there are too few such sites to make inferences.

The estimated sampling SD in the 'Saprobic index' generally higher for sites from France (stream type F08) and Latvia (L02). This may be because the Saprobic index is not valid for the taxonomic level of identification used by these STAR partners, and highlights the more general problem of only using metrics in situations for which they are appropriate.

Estimates of the average overall replicate sampling SD (SD_E) for each metric are therefore given separately for sites in each stream type in Tables 5.6 and 5.7 for untransformed and transformed metric values respectively. All sites within a stream type will have sampled in the same way and individuals identified to a consistent taxonomic level.

The estimates of overall replicate sampling SD_E in Table 5.7 (based where appropriate on transformed metric values) can be used in the STARBUGS software package (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sampling variability in individual metric values on the uncertainty of multi-metric assessments of the ecological status of sites. These estimates can also be used to provide information on the expected uncertainty in metrics values due to sampling variation for sites in the same stream type where only one sample has been taken at a point in time.

However, there are only a few STAR sites in each stream type for which replicate samples were taken, so the individual estimates of SD_E may themselves be imprecise. Therefore, if there are no obvious major differences in the SD_E between stream types, it may be more robust to use estimates based on the information from a combination of stream types, or even the median sampling SD value across all stream types, as given in Table 5.7, or the overall estimate of SD_E in Table 5.8 based on the hierarchal ANOVA of all stream types simultaneously.

5.4.4 Relative precision of different metrics derived from STAR-AQEM samples

If a particular metric is to be effective in discriminating the ecological status classes of river sites within a stream type, then the overall replicate sampling variance (σ_E^2) should be small relative to the total variance in metric values amongst all sites across the range of ecological qualities within the stream type. The total variance (σ_T^2) is the sum of the variances due to replicate sampling (σ_E^2), seasonal differences within sites (σ_I^2) and differences between sites (σ_J^2).The size of the overall replicate sampling variance (σ_E^2) relative to the total variance (σ_T^2) within a stream type is given by the statistic: $P_{samp} = 100\sigma_E^2 / \sigma_T^2$.

Values of P_{samp} were first calculated using the hierarchical ANOVA approach for sites from all stream types combined which effectively uses the values of the individual variance components averaged across stream types for which replicate samples were available (Table 5.8). The estimates for the three Saprobic indices exclude data from Greece, Italy and Portugal because their macroinvertebrate identification was only to family level. The values of P_{samp} vary between 9% for the German new Saprobic Index up to 35% for the metric '%Rheophilic based on abundance classes', with an average value of 18%. The new German

and the Czech Saprobic indices appear to have the lowest overall average relative sampling variance when based on ANOVA across all stream types with metric values available. The German new Saprobic index weights taxa by their abundance class rather than their raw abundance as in the original Saprobic index, which may explain why it is relatively more stable and less prone to sampling variance. Similarly, the metric '%EPT individuals' had a lower percentage sampling variance when based on abundance classes (P_{samp}=12%) compared to abundances (18%).

Table 5.8 STAR-AQEM method: Estimates of the average standard deviations (SD) in metric values due to each of the hierarchical effects of sub-sampling (SD_U), field sampling (SD_R), sites (SD_I), seasons (SD_J) and stream types (SD_K). $P_{samp} = 100\sigma_E^2 / \sigma_T^2$ = percentage of overall variance $(\sigma_T^2 = SD_E^2 + SD_I^2 + SD_J^2)$ within a stream type due to the overall replicate sampling variance $(SD_E^2 = SD_U^2 + SD_R^2)$. Estimates are based on, and applicable to, transformed (f(x)) values of metrics as indicated. Based on ANOVA of all available sites averaging across stream types.

Metric	f(x)	SD_U	SD_R	SD_E	SD_I	SD_J	SD_K	P_{samp}	P_{samp} rank
Abundance [ind/m²]	√√x	0.458	0.610	0.763	1.176	0.347	1.312	28	26
Number of taxa	√x	0.297	0.271	0.402	0.884	0.357	1.095	15	7
Number of Families	√x	0.234	0.190	0.301	0.638	0.138	0.398	18	13
Number of EPT taxa	√x	0.271	0.191	0.332	0.869	0.000	0.757	13	6
Saprobic Index	x	0.061	0.110	0.126	0.259	0.039	0.197	19	20
German Saprobic new	x	0.052	0.065	0.083	0.256	0.020	0.266	9	1
Czech Saprobic	x	0.064	0.126	0.141	0.382	0.088	0.343	11	2
ASPT	x	0.271	0.183	0.327	0.748	0.000	0.583	16	8
IBE	x	0.671	0.415	0.789	1.634	0.135	1.699	19	19
Diversity SW	x	0.120	0.221	0.251	0.504	0.246	0.378	17	11
% Rheophilic	asin	0.048	0.131	0.140	0.242	0.081	0.179	23	23
% Rheophilic (abund classes)	asin	0.039	0.092	0.100	0.137	0.002	0.114	35	27
% Littoral	asin	0.027	0.064	0.069	0.125	0.058	0.120	20	21
% Grazers/Scrapers	asin	0.032	0.047	0.057	0.121	0.033	0.098	17	12
% Shredders	asin	0.021	0.047	0.051	0.138	0.041	0.110	11	3
% Gatherers/Collectors	asin	0.034	0.068	0.076	0.157	0.035	0.071	18	15
% Oligochaeta	asin	0.037	0.098	0.105	0.167	0.040	0.132	27	25
% EPT individuals	asin	0.052	0.078	0.094	0.196	0.042	0.151	18	14
% EPT (abund. classes)	asin	0.038	0.030	0.048	0.132	0.000	0.111	12	4
% EPT Taxa	asin	0.052	0.022	0.056	0.117	0.000	0.109	19	17
RETI	asin	0.036	0.064	0.073	0.152	0.000	0.111	19	18
Log(Sel_EPTD+1)	x	0.084	0.260	0.273	0.584	0.220	0.521	16	9
1 –GOLD	asin	0.057	0.097	0.113	0.236	0.030	0.150	18	16
Trait m1 : max size ≤ 1cm	asin	0.021	0.029	0.036	0.062	0.015	0.052	24	24
Trait m2 : >1 cycle	asin	0.025	0.023	0.034	0.077	0.008	0.055	16	10
Trait m7 : crawler locomotion	asin	0.019	0.020	0.028	0.054	0.013	0.032	20	22
Trait m12 : current <25cm/s	asin	0.017	0.015	0.023	0.058	0.021	0.062	12	5
Average (range)								18 (9-35)	

Amongst the four species trait metrics analysed, the metric 'Trait m12' based on water current preference of taxa had the lowest overall percentage sampling variance (12%). The six proposed Inter-calibration Common Metrics (ICMs) (Buffagni *et al.* 2004) all had percentage sampling variances less than 20% which is encouraging (Table 5.8).The percentage sampling variance was higher for 'Number of Families' (18%) than for 'Number

of EPT taxa' (13%), and marginally more than that for 'Shannon-Wiener diversity index' (17%). The other three proposed ICM metrics, ASPT (16%), 'Log(Sel_EPTD+1)' (16%) and '1-GOLD' (18%) all had similar overall percentage sampling variances when averaged across all available sites and stream types. This is encouraging for the use of the latter two metrics which are recently devised metrics.

It may be that some metrics such as the Saprobic indices are not appropriate for some of the stream types or taxonomic levels used by some partner countries. Values of P_{samp} were therefore also calculated for each individual stream type for which sufficient data existed (Table 5.9). In Table 5.9, estimates for a stream type only depend on the sample data for that stream type and so are not influenced by potentially inappropriate sample data from other stream types. To compare the relative susceptibility of metrics to sampling variation, the median value of P_{samp} across stream types was calculated for each metric (Table 5.9).

The STAR field sampling programme included sites from 'high' or 'good' quality to 'poor'/'bad' quality. For a fixed size of replicate sampling variance, the percentage variance P_{samp} will be less in stream types for which a wider range of qualities of sites were sampled. This should be remembered when comparing values of P_{samp} across stream types for any particular metric. However, comparisons of the values of P_{samp} between metrics within stream types are completely valid because they are all based on the same set of sites. The values of P_{samp} for the 27 metrics were therefore ranked within each stream type and then the ranks averaged across stream types to give the column 'Mean rank' in Table 5.9. (For stream types where estimates were not calculated for some metrics, the ranks were re-scaled to give a range of 1-27 to ensure comparability across ranks).

The percentage sampling variance for any particular metric varies considerably between stream types. This is partly due to the small number of sites in each type, but also because the percentage also depends on the overall range of values of the metric across the sites within the type. For example, P_{samp} is high for Austrian stream types A05 and A06 because those sites were chosen to assess stress from degradation in stream morphology rather than eutrophication and so do not have a great range of values of ASPT (Figure 5.10(a)), making sampling variability a greater proportion of total variability (Table 5.9).

However, overall patterns in P_{samp} are detectable. The original Saprobic index, the German new Saprobic index and the Czech Saprobic index appear to have the lowest percentage sampling variance with median values of only 3%, 5% and 6% respectively (Table 5.9). This suggests that these Saprobic indices have amongst the lowest susceptibilities to sampling variation and can be estimated with the greatest relative precison within a stream type. Sampling variance tends to be less than 10% of the total variance in Saprobic metric values within any one stream type. In contrast, ASPT, another indicator of organic pollution stress (but based on only the presence-absence of families), has highly levels of percentage sampling variance within most stream types.

Although values were not available for many stream types, the new metric 'Log(Sel_EPTD+1) also appears to have relatively low replicate sampling variance with a median value of P_{samp} of only 7%.

Amongst the four species trait metrics analysed, 'Trait m1' is subject to the greatest sampling variation relative to its total variability within a stream type.

Table 5.9 STAR-AQEM method: ANOVA estimates of the percentage of overall variance (P_{samp}) within each stream type due to the overall replicate sampling variance (SD_E^2). Estimates are based on, and applicable to, transformed (f(x)) values of metrics as indicated and only given for stream types with at least 3 site/season combinations with replicate samples. Med = median P_{samp} ; Mean rank = average of the ranks of P_{samp} values within each stream type.

Metric	f(x)	A05	A06	C04	C05	D03	D04	D06	F08	H04	I05	I06	K02	L02	O02	P04	S05	S06	U15	U23	Med	Mean Rank
Abundance [ind/m²]	√√x	21	11	16	38	4	6	34	55	11	6	48	60	30	23	22	3	70	7	15	21.5	18
Number of taxa	√x	15	16	6	7	29	18	3	9	11	54	66	20	13	3	31	32	13	32	9	15.5	16
Number of Families	√x	11	17	10	9	8	24	7	18	13	46	63	22	12	2	35	28	14	26	9	15.5	17
Number of EPT taxa	√x	22	28	4	2	56	16	7	15	27	60	77	13	16	1	28	20	7	6	1	15.5	15
Saprobic Index	x	2	10	3	4	3	1	2	60				1	46	5		6	3	10	1	3	8
German Saprobic new	x	2	7	6	3	6	14	2	29				2	16	5		3	4	10	5	5	8
Czech Saprobic	x	2	6	4	8	6	1	2	29				6	29	18		4	13	14	24	6	9
ASPT	x	42	40	15	4	31	16	19	15	36	57	18	9	31	3	43	29	3	11	7	17	17
IBE	x	6	9	4	18	18	12	3	23	16	92	47	15	27	7	18	40	74	8	13	16.5	16
Diversity SW	x	3	14	9	5	15	19	7	9	18	20	36	41	13	7	35	37	14	14	17	14	15
% Rheophilic	asin	2	8	8	13	7	11	6	13	38	6	32	13	11	74	21	7	15	17	60	12	12
% Rheophilic (ab-class)	asin	3	14	5	9	5	39	13	9	29			7	14	94	15	15	11	5	22	12	12
% Littoral	asin	9	35	6	6	9	7	26	40	3			25	14	44	36	4	5	17	51	15.5	15
% Grazers/Scrapers	asin	5	13	16	8	16	10	6	3	10	8	24	23	34	41	21	5	17	18	80	16	14
% Shredders	asin	6	5	2	13	64	1	2	10	9			18	16	7	7	46	10	18	19	10	11
% Gatherers/Collectors	asin	8	4	12	7	37	2	1	8	30			27	21	16	19	7	19	43	61	14	14
% Oligochaeta	asin	25	1	12	3	16	9	21	40	50	25	36	48	16	56	38	8	13	5	10	16	16
% EPT individuals	asin	7	14	17	18	4	11	13	25	11	16	31	19	13	4	9	36	29	6	44	15	14
% EPT (ab-class)	asin	7	24	5	9	33	7	8	25	36			13	9	2	15	17	9	6	10	9	11
% EPT Taxa	asin	15	44	21	16	45	13	28	44	52	64	20	21	10	2	26	11	13	9	5	18	18
RETI	asin	8	17	8	8	99	12	1	3	12	7	10	24	39	20	20	12	6	29	97	12	14
Log(Sel_EPTD+1)	x	6		3	16			3	14	27		35	6			29	8			1	7	11
1 –GOLD	asin	7		7	20			6	16	17		20	45			9	41			17	16.5	17
Trait m1:max size ≤1cm	asin	2	17	60	10	25	45	55	64	21	26	45	41	28	99	27	32	6	4	11	27.5	20
Trait m2 : >1 cycle	asin	6	5	5	3	9	5	9	10	15	20	10	12	33	73	18	12	10	18	5	10	11
Trait m7 : crawler loco.	asin	27	4	7	6	18	6	18	31	24	11	28	21	18	72	17	10	8	18	4	17.5	15
Trait m12:current<25cm	asin	12	10	6	11	2	12	15	19	16	22	14	11	15	14	20	2	6	17	10	12	12

The proposed ICM metrics of Number of EPT taxa, ASPT, Shannon-Wiener diversity and (1-GOLD) also have highly variable estimates of P_{samp} but with similar intermediate size median values of 15.5%, 17%, 14% and 16.5% respectively.

In summary, the mean rank of the P_{samp} values for each metrics suggests that most of these selected 27 metrics have average replicate sampling variances of 10-20% of the total variance in metrics values within a stream type (Table 5.8). This suggests that the precision of such metrics is sufficient to indicate gross changes in the ecological status of sites, but there will be considerable uncertainty in the assignment of sites to particular status classes.

The estimates of SD_E derived here can be used in the software program STARBUGS (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sampling variation on the uncertainty in assignment of sites to ecological status classes.

5.5 REPLICATE SAMPLING VARIABILITY IN METRIC VALUES FOR THE RIVPACS METHOD

5.5.1 Assessment of sampling variability using the RIVPACS method

The RIVPACS method is described in section 2.1.2.2. The RIVPACS method of field sampling and subsequent sample processing was used in by the STAR partners in Austria (stream types A05 and A06), Germany (stream types D03, D04 and D06), Greece (stream types H04, H05, H06 and H07) and the UK (stream types U15 and U23). Within the careful design of the STAR field sampling programmes, two RIVPACS samples ('main' and 'replicate') were taken at the same time at all, or at the vast majority, of the sites where replicate STAR-AQEM samples were taken. This enabled direct comparison between the RIVPACS and STAR-AQEM methods of their replicate sampling SD in metric values, both in numerical terms (SD_E) and as a percentage (P_{samp}) of the total variance with a stream type. In Austria, Germany and the UK, replicate RIVPACS samples were taken from sites in each of the sampled stream types, but in Greece, replicate sampling was confined to six sites in one stream type H04 (Table 5.10).

Table 5.10 Number of sites (and their STAR site codes) in each stream type and country for which two replicate RIVPACS samples ('main' and 'replicate') were taken in at least one season (1=spring, 2=summer, 3=autumn). Site code 'xxx.y' indicates replicate samples only taken at site 'xxx' in season 'y'.

Country	Stream Type	Description	Seasons sampled	n Sites	STAR site codes
Austria	A05	small-sized, shallow mountain streams	1 + 2	4	600 603 607.2 609.2
	A06	small-sized crystalline streams of the ridges of the Central Alps	1 + 2	4	701 702.2 706 708
Germany	D03	medium-sized lowland streams	1 + 2	2	649 659
	D04	small-sized, shallow mountain streams	1 + 2	2	627 634
	D06	small-sized Buntsandstein-streams	1 + 2	2	816 821
Greece	H04	small-sized calcareous mountain streams in Western, Central and Southern Greece	1 + 2	6	735 737 738 739 753 756
UK	U15	small-sized, shallow lowland streams	1 + 3	3	639 642 648
	U23	medium-sized lowland streams	1 + 3	3	674 678 681

Figures 5.13-5.15 plot the difference in (untransformed) metric values between two replicate RIVPACS samples taken from the same site against the average of the two values for a range of metrics. The differences rather than the SD of the two replicate samples are shown in these initial plots to aid visual understanding.

Table 5.11 gives estimates of the replicate sampling standard deviation (SD_E) in untransformed values of each of the metrics, with separate estimates for each STAR stream type for which replicate sample values were available. The estimate of sampling SD for individual metrics do vary between stream types. At the time of analysis, values of the new metrics 'Log(Sel_EPTD+1)' and '(1-GOLD)' were only available for RIVPACS samples from Greek stream type H04; values for the new species trait metrics were only available for the UK RIVPACS samples (stream types U15 and U23).

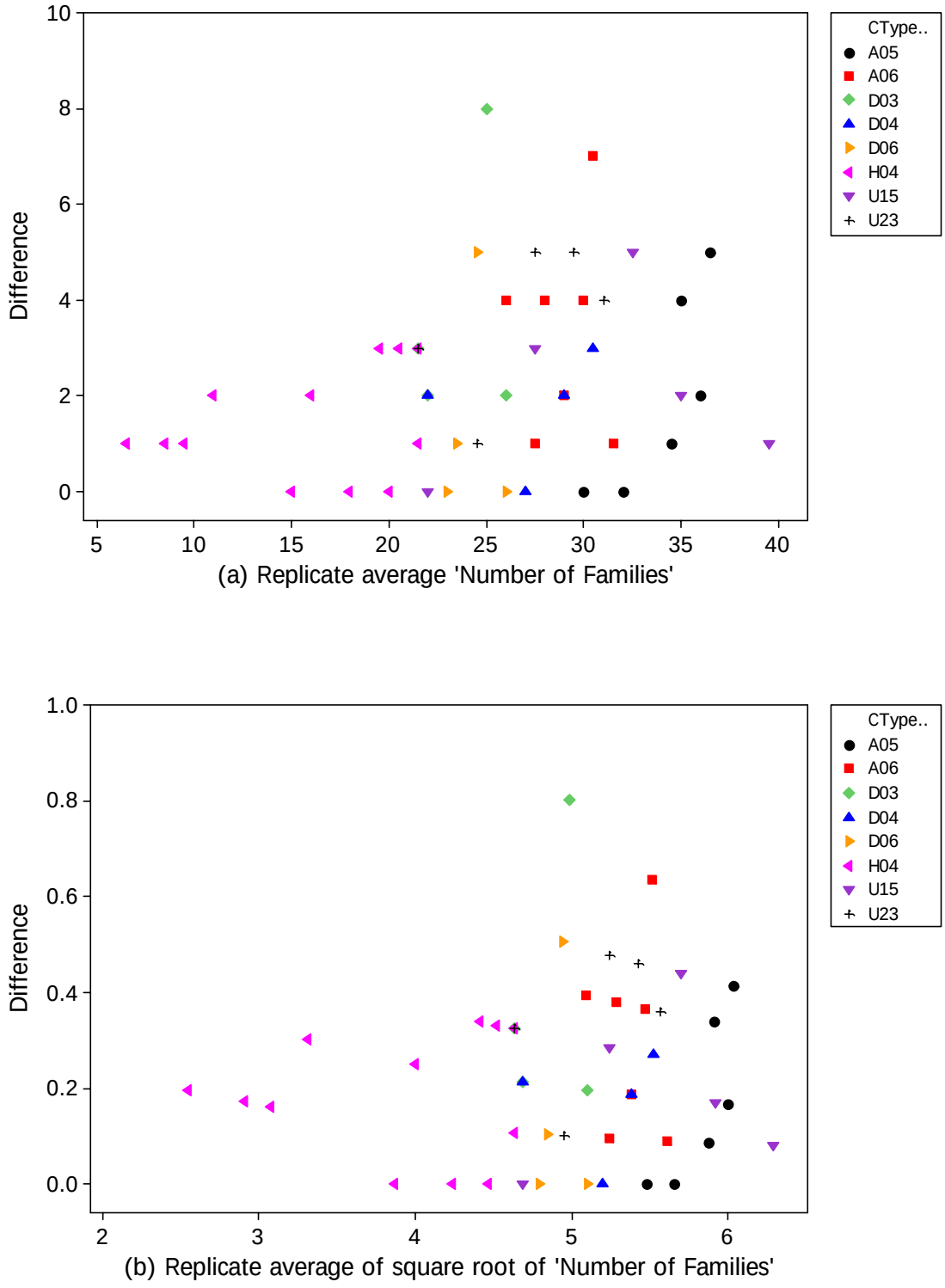
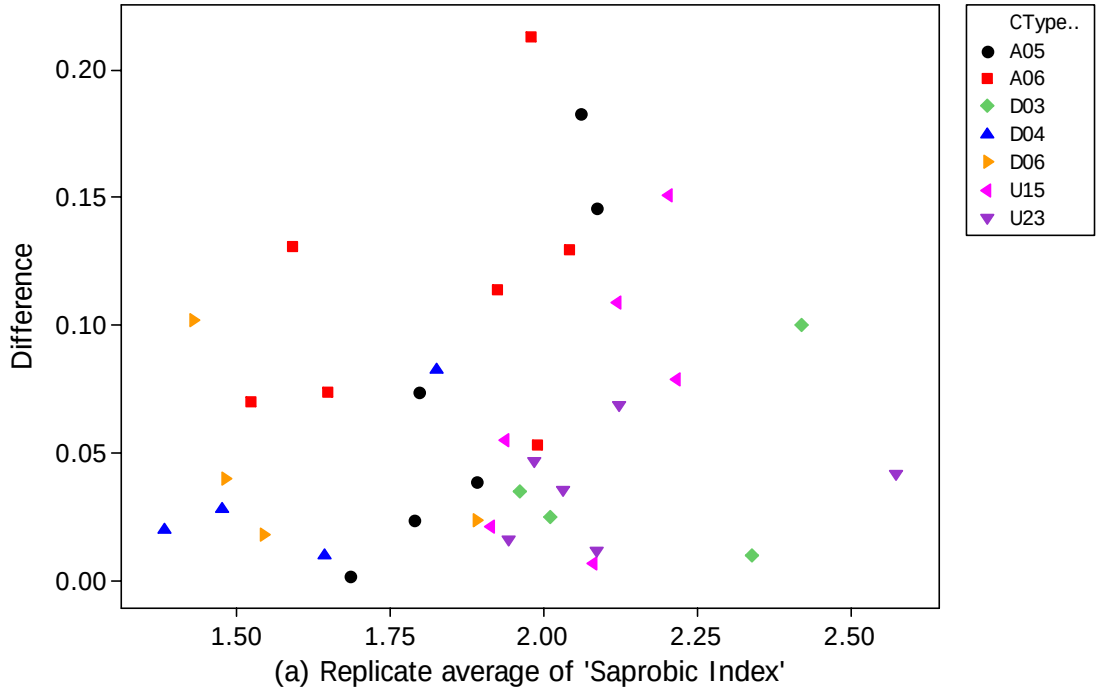


Figure 5.13 RIVPACS method: Difference in 'Number of Families' recorded between two replicate samples plotted against the average of the two values for (a) untransformed and (b) square root transformed values. Plot (but not analyses) excludes one UK site with 18 and 35 families in two replicate samples)



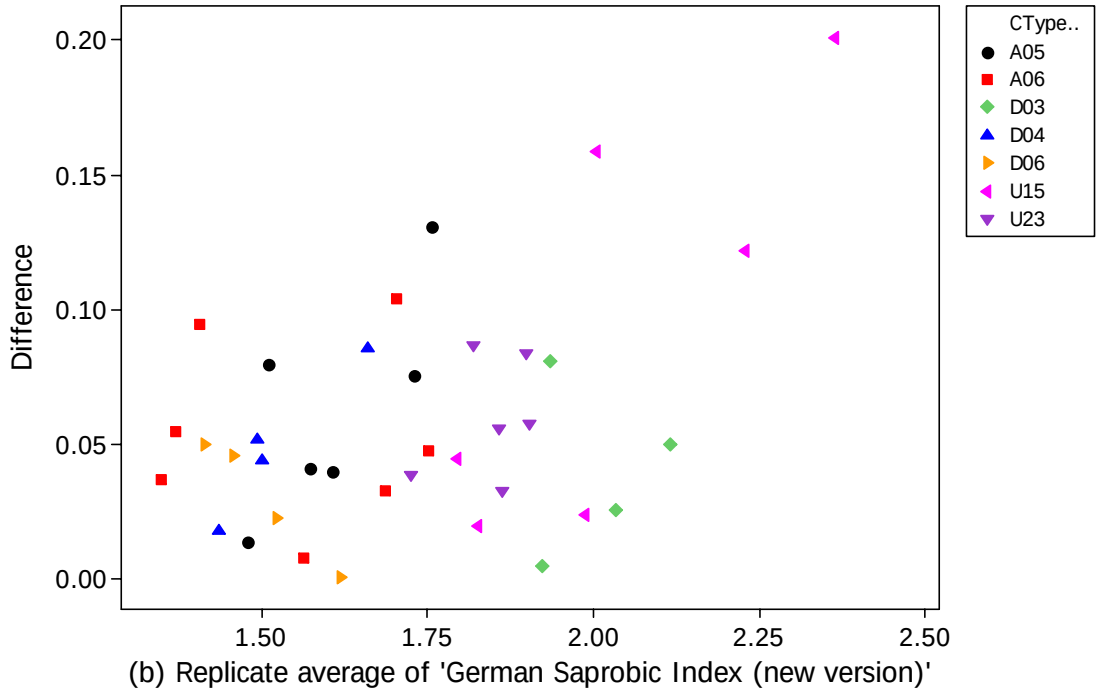


Figure 5.14 RIVPACS method: Difference in metric values between two replicate samples plotted against the average of the two values for untransformed values of metrics (a) 'Saprobic Index (Zelinka & Marvan)' and (b) German Saprobic Index (new version).

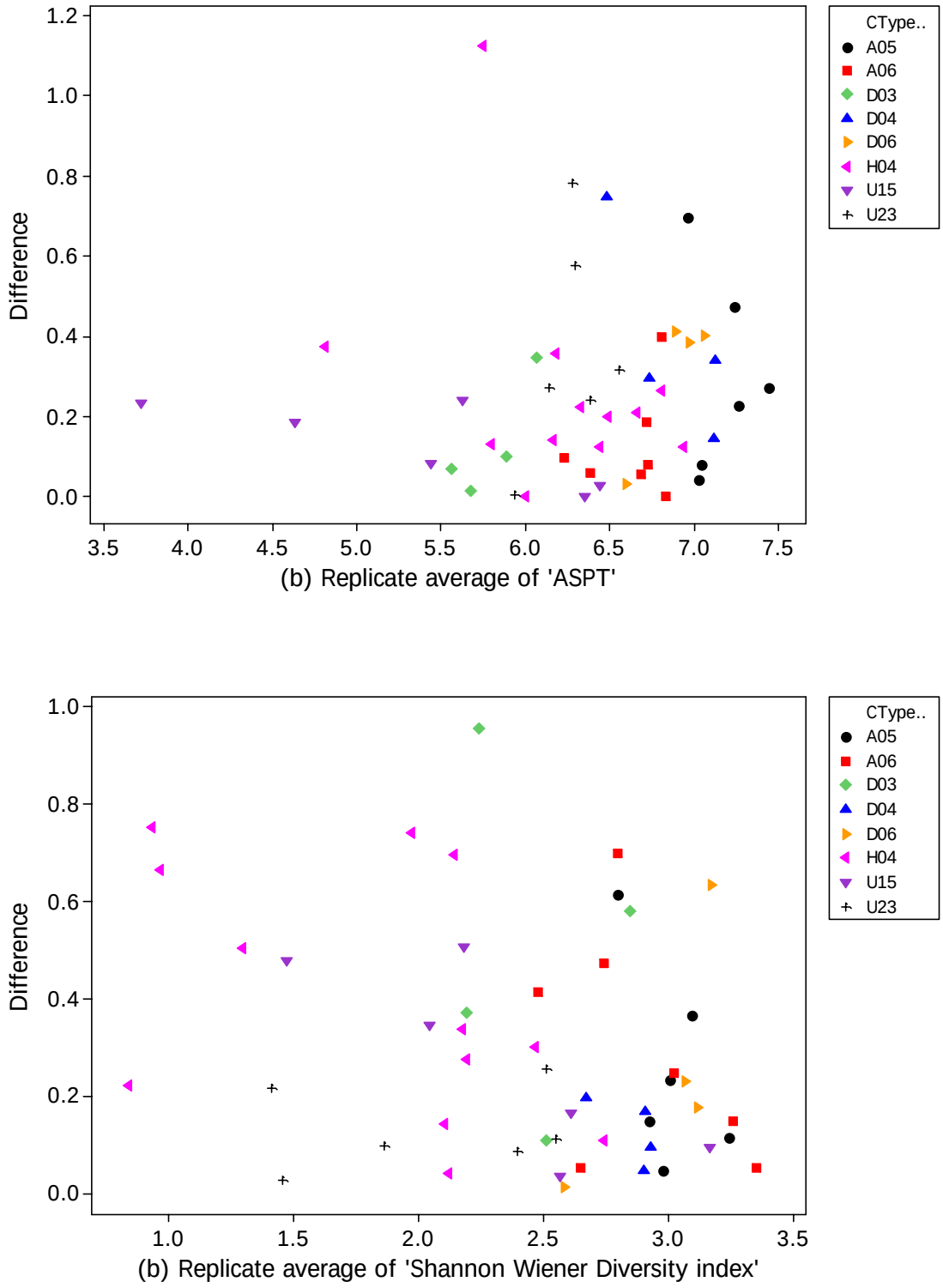


Figure 5.15 RIVPACS method: Difference in metric values between two replicate samples plotted against the average of the two values for untransformed values of metrics (a) ASPT and (b) Shannon-Wiener Diversity index.

The difference between replicate RIVPACS samples in the 'Number of Families' recorded tends be less when there are fewer families present at a site (i.e. as measured by the average of the two replicate values (Figure 13(a)). This is a similar pattern to that found for STAR-AQEM samples, supporting the overall decision to estimate variance components based on the square root transformed values for 'Number of Families' and other taxonomic richness metrics. After transformation, the replicate differences did not vary consistently with the average Family richness of the site, even though site richness varied considerably with especially fewer families recorded in the Greek sites from stream type H04 (Figure 5.13(b)).

Table 5.11 RIVPACS method: Estimates of the standard deviation (SD_E) in untransformed metric values of replicate samples from sites in each STAR stream type. Missing values denote stream types where particular metric values were not available/appropriate.

Metric	SD_E in Untransformed metric values							
	Stream Type							
	A05	A06	D03	D04	D06	H04	U15	U23
Abundance [ind/m²]	1587	730	443	354	818	217	597	3152
Number of taxa	5.76	6.21	7.16	5.69	4.53	1.47	8.96	8.07
Number of Families	1.96	2.71	3.18	1.46	1.80	1.27	5.23	2.90
Number of EPT taxa	2.68	1.39	2.67	2.92	2.62	1.21	1.71	2.50
Saprobic Index	0.072	0.087	0.039	0.032	0.040		0.061	0.029
German Saprobic new	0.052	0.044	0.035	0.039	0.025		0.083	0.044
Czech Saprobic	0.123	0.099	0.090	0.087	0.031		0.106	0.074
ASPT	0.266	0.124	0.131	0.314	0.246	0.274	0.114	0.314
IBE	0.936	0.626	0.686	0.255	0.381	0.742	1.115	0.854
Diversity SW	0.224	0.263	0.418	0.098	0.246	0.332	0.231	0.108
% Rheophilic	4.95	9.58	8.61	4.84	8.92	14.20	4.21	5.82
% Rheophilic (ab-class)	3.06	4.22	2.66	5.34	1.98	8.87	4.60	3.97
% Littoral	3.16	2.01	2.49	3.94	1.07	5.49	1.49	2.65
% Grazers/Scrapers	4.78	3.68	3.42	2.99	3.76	8.02	2.05	3.38
% Shredders	1.26	4.86	12.80	2.85	4.15	1.36	2.78	1.84
% Gatherers/Collectors	3.63	2.51	8.72	4.27	1.56	9.41	8.99	2.84
% Oligochaeta	4.40	2.30	0.83	1.24	0.34	0.73	8.55	5.92
% EPT individuals	9.35	5.07	6.38	5.53	7.31	9.91	2.56	6.30
% EPT (ab-class)	3.63	2.05	5.05	1.86	3.98	8.01	3.23	4.33
% EPT Taxa	3.53	2.76	4.34	4.48	4.77	7.98	3.40	4.66
RETI	0.056	0.039	0.116	0.050	0.009	0.054	0.049	0.030
Log(Sel_EPTD+1)						0.216		
1 –GOLD						0.076		
Trait m1:max size ≤1cm							0.021	0.025
Trait m2 : >1 cycle							0.039	0.020
Trait m7 : crawler loco.							0.031	0.026
Trait m12:current<25cm							0.014	0.013

Equally interesting is that the sampling SD of the square root of 'Number of Families' does not vary systematically between stream types – a Kruskal-Wallis one-way ANOVA of the ranks of the individual site sampling SD shows no statistically significant differences between stream types (test $p = 0.274$). This suggests that the overall ANOVA estimate of sampling SD of 0.268 based on all of the sites sampling using the RIVPACS method. This estimate is similar but slightly higher than the equivalent estimate of 0.228 derived by Clarke *et al*. (2002) based on single season replicate sampling of a set of 16 UK sites covering a wide range of stream types and qualities.

Differences (and SD) between replicate RIVPACS sample values of the original Saprobic Index of Zelinka and Marvan (Figure 5.14(a)) and either the new version of the German Saprobic index (Figure 5.14(b)) or the Czech Saprobic index showed no major trends or differences between stream types. Although the estimates of the average sampling SD for each Saprobic index do vary between stream types (Table 5.11), Kruskal-Walllis ANOVA tests of the ranked sampling SD did not detect any statistically significant differences in SD between stream types (test $p = 0.157$, 0.699 and 0.468 respectively). However, this may be partly because replicate RIVPACS samples were only taken from a few (2-6) sites in each stream type (Table 5.10).

Differences between replicate sample values of ASPT show no systematic patterns – the replicate variability does not change with the average value of ASPT for a site (Figure 5.15(a)). Equally interesting is that the sampling SD does not vary systematically between stream types – a Kruskal-Wallis one-way ANOVA of the ranks of the individual site sampling SD shows no statistically significant differences between stream types (test $p = 0.128$). The overall ANOVA estimate of the sampling SD of ASPT, averaged over all stream types where RIVPACS samples were taken, was 0.239. This is very close to the value of 0.249 obtained for equivalent RIVPACS single season samples by Clarke *et al*. (2002), which is encouraging about the repeatability and robustness of these uncertainty analyses.

Differences in the Shannon-Wiener diversity index between replicate samples are often large relative to the actual values of the index, especially for some UK and German sites (Figure 5.15(b)), suggesting that this diversity index is highly susceptible to sampling variation and hence may be of low precision. The general precision of metrics derived from samples based on the RIVPACS method is discussed below.

Tables 5.12-5.15 provide estimates of the RIVPACS sampling SD (SD_E) for the transformed (where appropriate) values of each metric, separately for each stream type, and country, for which RIVPACS samples were taken.

This information is useful for people working within any particular stream type or country, and in particular, the estimates sampling SD (SD_E) obtained here can be used as provisional estimates of the expected sampling uncertainty in metric values obtained using the RIVPACS method for other new sites in the same stream types or country. However, the individual estimates of relative precision are highly variable because of the low number of sites involved.

Hierarchical ANOVA analyses were therefore also used for each country to estimate the average variance in metrics values within stream types due to differences between sites within a season (SD_I^2) and due to differences between seasons in average metric values (SD_J^2). From these variance component estimates, the average replicate sampling variance (SD_E^2) for each metric could be expressed as a percentage (P_{samp}) of average overall variance for the metric within a stream type, separately for each country (Tables 5.12-5.15).

Low values of P_{samp} indicate that a metric has high statistical precision in the sense that the variation in replicate samples from the same site are small compared to the total variability amongst all sites (of varying ecological quality) within a stream type.

Tables 5.12-5.15 also give, for each country, the corresponding estimates of SD_E and P_{samp} for replicate samples taken using the STAR-AQEM method from (mostly) the

same sites as used for the replicate RIVPACS sampling. This provides a valid direct comparison of the relative precision of the two methods in terms of susceptibility to sampling variability, separately for sites in Austria, Germany, Greece and the UK.

5.5.2 RIVPACS method in Austria and comparison with STAR-AQEM method

Using the RIVPACS method to sample the streams in Austria gave high precision to estimates of most metrics with nearly two-thirds (62%) of metrics having replicate sampling variances of less than 10% of the total variance in metric values within any one stream type (Table 5.12). The average value of P_{samp} was only 9 % and no metric had a value of P_{samp} greater than 24% - this suggests high sampling repeatability of all aspects of the macroinvertebrate community structure. For these Austrian stream types, the RIVPACS and STAR-AQEM methods were, on average, about equally precise (Table 5.12).

Table 5.12 Comparisons of RIVPACS (R) and STAR-AQEM (S-A) methods used in Austria for overall standard deviations (SD_E) and percentage variance (P_{samp}) due to replicate sampling for transformed (f(x)) values of metrics. Based on hierarchical ANOVA within each stream type (A05 and A06) and averaging across stream types. Values in bold denote RIVPACS < STAR-AQEM.

		SD_E RIVPACS		Average SD_E		P_{samp}	
Metric	f(x)	A05	A06	R	S-A	R	S-A
Abundance [ind/m²]	√√x	0.891	0.518	0.715	0.640	24	17
Number of taxa	√x	0.353	0.404	**0.038**	0.513	**8**	14
Number of Families	√x	0.164	0.251	**0.215**	0.310	**6**	12
Number of EPT taxa	√x	0.272	0.159	**0.219**	0.454	**5**	23
Saprobic Index	x	0.072	0.087	0.080	0.042	13	6
German Saprobic new	x	0.052	0.044	0.048	0.029	**4**	5
Czech Saprobic	x	0.123	0.099	0.111	0.084	6	5
ASPT	x	0.266	0.124	**0.202**	0.344	**9**	40
IBE	x	0.936	0.626	0.784	0.565	13	7
Diversity SW	x	0.224	0.263	0.246	0.165	19	5
% Rheophilic	asin	0.060	0.104	0.087	0.068	8	5
% Rheophilic (abundance classes)	asin	0.031	0.043	0.038	0.037	7	6
% Littoral	asin	0.032	0.022	**0.027**	0.068	**3**	19
% Grazers/Scrapers	asin	0.049	0.044	0.046	0.043	10	6
% Shredders	asin	0.029	0.056	0.046	0.031	10	6
% Gatherers/Collectors	asin	0.038	0.026	**0.032**	0.041	**4**	6
% Oligochaeta	asin	0.067	0.045	0.056	0.051	7	6
% EPT individuals	asin	0.099	0.064	0.082	0.066	13	9
% EPT (abundance classes)	asin	0.038	0.022	**0.031**	0.038	**8**	11
% EPT Taxa	asin	0.037	0.029	**0.033**	0.051	**11**	22
RETI	asin	0.056	0.042	**0.049**	0.053	**7**	10
Average (range)						9 (3-24)	11 (5-40)
Metrics (%) with P_{samp} < 10%						13 (62%)	12 (57%)
Number of metrics with smaller value of SD_E or P_{samp}				9	12	10	11

5.5.3 RIVPACS method in Germany and comparison with STAR-AQEM method

Using the RIVPACS method to sample the three STAR stream types (D03, D04 and D06) in Germany gave variable levels of sampling precision for metric values. Six metrics (29%), including the three Saprobic metrics, had replicate sampling variances of less than 10% of the total variance in metric values within any one stream type – indicating very high sampling repeatability and precision (Table 5.13).

Table 5.13 Comparisons of RIVPACS (R) and STAR-AQEM (S-A) methods used in Germany for overall standard deviations (SD_E) and percentage variance (P_{samp}) due to replicate sampling for transformed (f(x)) values of metrics. Based on hierarchical ANOVA within each stream type (D03, D04, D06) and averaging across stream types. Values in bold denote RIVPACS < STAR-AQEM.

		SD_E RIVPACS			SD_E		P_{samp}	
Metric	f(x)	D03	D04	D06	R	S-A	R	S-A
Abundance [ind/m²]	√√x	0.486	0.502	0.334	0.447	0.363	17	12
Number of Taxa	√x	0.519	0.382	0.324	0.416	0.333	24	14
Number of Families	√x	0.323	0.139	0.182	0.228	0.203	18	12
Number of EPT Taxa	√x	0.341	0.276	0.248	**0.291**	0.380	**12**	20
Saprobic Index	x	0.039	0.032	0.040	0.037	0.028	2	2
German Saprobic new	x	0.035	0.039	0.025	**0.034**	0.041	**2**	7
Czech Saprobic	x	0.090	0.087	0.031	0.075	0.057	4	2
ASPT	x	0.131	0.314	0.246	**0.242**	0.286	27	23
IBE	x	0.686	0.255	0.381	**0.476**	0.515	**10**	11
Diversity SW	x	0.418	0.098	0.246	0.286	0.210	25	13
% Rheophilic	asin	0.094	0.049	0.093	0.081	0.072	12	7
% Rheophilic (abundance classes)	asin	0.027	0.054	0.020	**0.037**	0.040	7	7
% Littoral	asin	0.047	0.040	0.011	**0.036**	0.045	**4**	11
% Grazers/Scrapers	asin	0.055	0.032	0.039	**0.043**	0.045	**6**	10
% Shredders	asin	0.162	0.041	0.049	0.100	0.075	38	21
% Gatherers/Collectors	asin	0.093	0.045	0.018	0.061	0.054	14	11
% Oligochaeta	asin	0.045	0.079	0.040	0.057	0.030	55	15
% EPT individuals	asin	0.074	0.058	0.074	0.069	0.060	11	8
% EPT (abundance classes)	asin	0.054	0.019	0.041	**0.040**	0.042	**14**	16
% EPT Taxa	asin	0.046	0.045	0.049	0.047	0.046	**21**	24
RETI	asin	0.122	0.050	0.009	**0.077**	0.090	**27**	32
Average (range)							17 (2-55)	13 (2-32)
Metrics (%) with P_{samp} < 10%							6 (29%)	6 (29%)
Number of metrics with smaller value of SD_E or P_{samp}					9	12	8	11

The average value of Psamp was still only 17%. In Germany, the overall sampling variability in metric values based the RIVPACS method was similar, but slightly higher than that based on the STAR-AQEM method. Interestingly, the RIVPACS method gave even higher precision (in terms of Psamp) for the German Saprobic Index (new version) than the STAR-AQEM method (Table 5.13). The percentage sampling variance of ASPT for the German stream types was much higher (23-27%) than that any of the Saprobic metrics for both the RIVPACS and STAR-AQEM methods. However, the main stress operating within the stream types sampled in Germany was considered to degration of stream morphology rather than organic pollution. Therefore both ASPT and the Saprobic indices, designed primarily to indicate biological impacts of organic pollution, do not cover a large range of values within these stream types (Figure 5.15). Within that constraint, the Saprobic species-based metrics appear to be relatively less susceptible to sampling varation.

5.5.4 RIVPACS method in Greece and comparison with STAR-AQEM method

Replicate RIVPACS samples were taken in each of two seasons at six sites in one stream type (H04) in Greece. It is important to remember, than macroinvertebrates were only identified to Family level, rather than genus or species level – this is why the Saprobic metrics are excluded (Table 5.14). The RIVPACS method led to higher percentage sampling variance (P_{samp}) for several metrics than was found for the other countries where the RIVPACS method was used. The average value of P_{samp} for Greek sites was 27% and only three metrics 'Number of taxa', 'Number of families' and the recently proposed ICM (Inter-calibration Common Metric) 'Log(Sel_EPTD+1)' (Table 5.14). In comparison with the STAR-AQEM method applied to the same sites in Greece, only 6 of the 20 metrics assessed had smaller sampling standard deviations (SD_E) for the RIVPACS method. However, the STAR-AQEM method had similar average levels of percentage sampling variance (25%) and only two (different) metrics with sampling variances less than 10% of total variance within the stream type. Thus the RIVPACS and STAR-AQEM give roughly equal moderately high sampling percentage sampling variances as applied within the STAR project to this stream type in Greece. Values for the new proposed ICM 'Log(Sel_EPTD+1)' were only available for sites in Greece for samples obtained using the RIVPACS method. It is encouraging that its estimated percentage sampling variance was only 8%, indicating high precision and repeatability, albeit only tested on these six sites in one stream type. It is then disappointing the equivalent figure based on the STAR-AQEM method for the same sites in 27% (Table 5.14).

Table 5.14 Comparisons of RIVPACS (R) and STAR-AQEM (S-A) methods used in Greece (stream type H04) for overall standard deviations (SD_E) and percentage variance (P_{samp}) due to replicate sampling for transformed (f(x)) values of metrics. Based on hierarchical ANOVA and averaging across all sampled sites. Values in bold denote RIVPACS < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	R	S-A	R	S-A
Abundance [ind/m²]	√√x	0.513	0.257	21	11
Number of taxa	√x	**0.182**	0.282	**5**	11
Number of Families	√x	**0.157**	0.299	**4**	13
Number of EPT taxa	√x	**0.291**	0.343	**20**	27
ASPT	x	**0.274**	0.437	**19**	36
IBE	x	0.742	0.545	32	16
Diversity SW	x	0.332	0.272	26	18
% Rheophilic	asin	0.160	0.146	**27**	38
% Rheophilic (abund classes)	asin	0.118	0.096	**25**	29
% Littoral	asin	0.065	0.022	15	3
% Grazers/Scrapers	asin	0.108	0.052	34	10
% Shredders	asin	0.056	0.053	42	9
% Gatherers/Collectors	asin	0.105	0.075	44	30
% Oligochaeta	asin	0.047	0.028	75	50
% EPT individuals	asin	0.116	0.094	17	11
% EPT (abund. classes)	asin	0.083	0.083	40	36
% EPT Taxa	asin	**0.089**	0.100	**51**	52
RETI	asin	0.061	0.046	18	12
Log(Sel_EPTD+1)	x	**0.216**	0.354	**8**	27
1 –GOLD	asin	0.107	0.094	**16**	17
Average (range)				27 (4-75)	23 (3-52)
Metrics (%) with P_{samp} < 10%				3 (14%)	2 (10%)
Number of metrics with smaller value of SD_E or P_{samp}		6	14	9	11

5.5.5 RIVPACS method in the UK and comparison with STAR-AQEM method

The RIVPACS method has been used by the UK environment agencies to assess the biological quality of rivers for 15 years. However, most national surveys and monitoring programmes have concentrated on the use of just RIVPACS observed to expected ratios of the metrics 'Number of BMWP taxa/families' and ASPT. The STAR project is the first time such a wide range of metrics have been calculated for sampling obtained using the RIVPACS method. It is therefore encouraging that, of the 25 metrics assessed, 15 (60%) had sampling variances which formed less than 10% of the total variance in each metric's values within any one stream type (Table 5.15). ASPT, the three Saprobic metrics and most of the percentage abundance of selected taxa all had such low sampling variability and hence high precision. However, the two taxonomic richness metrics 'Number of taxa' and 'Number of Families' both had much higher sampling variances with estimated values of P_{samp} of 32% and 36% respectively. This low precision and repeatability of these two richness metrics was not found for the RIVPACS method for either the Austrian or Greek streams and merits further investigation.

Table 5.15 Comparisons of RIVPACS (R) and STAR-AQEM (S-A) methods used in the UK for overall standard deviations (SD_E) and percentage variance (P_{samp}) due to replicate sampling for transformed (f(x)) values of metrics. Based on hierarchical ANOVA within stream types (U15 and U23) and averaging across stream types. Values in bold denote RIVPACS < STAR-AQEM.

		SD_E RIVPACS		Average SD_E		P_{samp}	
Metric	f(x)	U15	U23	R	S-A	R	S-A
Abundance [ind/m²]	√√x	0.353	0.812	0.626	0.535	15	11
Number of taxa	√x	0.702	0.564	0.637	0.542	32	20
Number of Families	√x	0.509	0.267	0.406	0.346	36	16
Number of EPT taxa	√x	0.288	0.326	0.308	0.254	6	4
Saprobic Index	x	0.061	0.029	**0.048**	0.065	**3**	5
German Saprobic new	x	0.083	0.044	**0.067**	0.074	**6**	7
Czech Saprobic	x	0.106	0.074	**0.091**	0.206	**3**	21
ASPT	x	0.114	0.314	**0.236**	0.320	**5**	9
IBE	x	1.115	0.854	0.993	0.699	19	11
Diversity SW	x	0.231	0.108	**0.180**	0.223	**10**	15
% Rheophilic	asin	0.046	0.064	**0.056**	0.190	**2**	41
% Rheophilic (abund classes)	asin	0.049	0.040	**0.045**	0.053	**7**	14
% Littoral	asin	0.016	0.039	**0.030**	0.093	**3**	38
% Grazers/Scrapers	asin	0.026	0.040	**0.034**	0.085	**9**	57
% Shredders	asin	0.035	0.027	**0.031**	0.071	**3**	18
% Gatherers/Collectors	asin	0.094	0.030	**0.070**	0.101	**23**	53
% Oligochaeta	asin	0.134	0.084	0.112	0.087	14	8
% EPT individuals	asin	0.030	0.078	**0.059**	0.126	**5**	33
% EPT (abund. classes)	asin	0.036	0.049	**0.043**	0.046	**7**	8
% EPT Taxa	asin	0.041	0.051	0.046	0.039	9	6
RETI	asin	0.050	0.031	**0.041**	0.130	**6**	65
Trait m1 : max size ≤ 1cm	asin	0.021	0.025	0.023	0.022	14	9
Trait m2 : >1 cycle	asin	0.040	0.020	0.031	0.030	12	10
Trait m7 : crawler locomotion	asin	0.032	0.026	0.029	0.019	26	8
Trait m12 : current <25cm/s	asin	0.015	0.013	**0.014**	0.020	**4**	12
Average (range)						11 (2-36)	20 (4-65)
Metrics (%) with P_{samp} < 10%						15 (60%)	9 (36%)
Number of metrics with smaller value of SD_E or P_{samp}				15	10	15	10

In the UK, when compared with the STAR-AQEM method used at the same sites, the RIVPACS method tended to give both lower sampling SD (SD_E) and lower percentage sampling variance the for the majority of metrics. This suggests that the STAR-AQEM method would offer improvement in precision for stream assessment in the UK over the current RIVPACS method.

5.5.6 Summary across stream types

As replicate RIVPACS samples could only be taken at a few sites in each country, estimates of precision within each country are likely to be imprecise and volatile. There hierarchical ANOVA was used to combine the data across all four countries where RIVPACS samples were taken to derive overall estimates of sampling SD and sampling precision for the RIVPACS method and how it compares with the STAR-AQEM method used at the same sites (Table 5.16).

Table 5.16 Comparisons of RIVPACS (R) and STAR-AQEM (S-A) methods for overall standard deviations (SD_E) and percentage variance (P_{samp}) due to replicate sampling for transformed (f(x)) values of metrics. Based on hierarchical ANOVA and averaging across all sites and stream types with replicate RIVPACS sampling. [1] denotes excludes Greek sites, [2] Greek sites only, [3] UK sites only. Values in bold denote RIVPACS < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	R	S-A	R	S-A
Abundance [ind/m²]	√√x	0.587	0.538	20	7
Number of taxa	√x	**0.434**	0.476	18	9
Number of Families	√x	**0.268**	0.295	14	12
Number of EPT taxa	√x	**0.278**	0.381	8	8
Saprobic Index [1]	x	0.059	0.047	7	2
German Saprobic new [1]	x	0.051	0.050	5	2
Czech Saprobic [1]	x	**0.094**	0.129	**4**	6
ASPT	x	**0.239**	0.320	10	10
IBE	x	0.773	0.592	17	5
Diversity SW	x	0.266	0.198	22	8
% Rheophilic	asin	**0.103**	0.119	**12**	14
% Rheophilic (abund classes)	asin	0.068	0.043	16	5
% Littoral	asin	**0.042**	0.071	**7**	13
% Grazers/Scrapers	asin	0.065	0.059	20	10
% Shredders	asin	0.063	0.060	17	10
% Gatherers/Collectors	asin	0.071	0.068	21	17
% Oligochaeta	asin	0.073	0.060	15	5
% EPT individuals	asin	**0.084**	0.087	12	9
% EPT (abund. classes)	asin	0.053	0.042	15	5
% EPT Taxa	asin	0.057	0.046	20	7
RETI	asin	**0.058**	0.093	**13**	30
Log(Sel_EPTD+1) [2]	x	0.216	0.144	10	3
1 –GOLD	asin	0.107	0.096	16	9
Trait m1 : max size ≤ 1cm [3]	asin	**0.023**	0.025	14	7
Trait m2 : >1 cycle [3]	asin	0.031	0.025	12	4
Trait m7 : crawler locomotion [3]	asin	0.029	0.019	26	7
Trait m12 : current <25cm/s [3]	asin	0.014	0.018	4	3
Average (range)				14 (4-26)	8 (2-30)
Metrics (%) with P_{samp} < 10%				6 (22%)	19 (70%)
Number of metrics with smaller value of SD_E or P_{samp}		11	16	4	21

The metrics 'Number of EPT taxa', 'Saprobic index (Zelinka & Marvan)', 'German Saprobic index new', 'Czech Saprobic index', 'ASPT', '%Littsral' and 'Log(Sel_EPTD+1) all have average percentage sampling variances (P_{samp}) of 10% or less for the RIVPACS method (Table 5.16). The three Saprobic metrics have values of P_{samp} less than 10% for each of eight stream types sampled using the RIVPACS method. This suggests that these metrics are highly robust to sampling effects using the RIVPACS method. This high sampling precision for these saprobic metrics is generally about the same as that obtained when the STAR-AQEM method is used at the same sites and stream types (Table 5.12-5.16).

The general level of percentage sampling variance for metric values obtaining using the RIVPACS method tends to be highest for the Greek sites in stream type H04, for which macroinvertebrates were only identified to family rather than species or genus level.

Eleven of the 27 metrics analysed had lower overall sampling SD for the RIVPACS method than for the STAR-AQEM method. These include all three taxonomic richness metrics, namely 'Number of taxa', 'Number of Families' and 'Number of EPT taxa'. However, because the total variance within a stream type is slightly higher (not shown) for several metrics for the STAR-AQEM method, once converted to relative variance (P_{samp}), the RIVPACS method had lower percentage sampling variance for only four metrics (Table 5.16).

Overall, the RIVPACS method seems to give slightly lower sampling precision (i.e. higher P_{samp}) for more than half of the metrics. However, the differences between the two methods in estimated values of both SD_E and P_{samp} is small for many metrics and probably within the estimation error of this dataset.

Either the stream type- or country-specific estimates of sampling SD (SD_E) in Tables 5.12-5.15 or the overall estimates averaged across stream types (Table 5.16) can be used in the software program STARBUGS to assess the effect of sampling variation on uncertainty in assessments of ecological status based on the RIVPACS method.

Within the careful design of the STAR field sampling programmer, two STAR-AQEM samples were taken at the same time at the vast majority of the sites where replicate RIVPACS samples were taken. This has enabled direct comparison between the RIVPACS and STAR-AQEM methods of their replicate sampling SD in metric values, both in numerical terms (SD_E) and as a percentage (P_{samp}) of the total variance with a stream type.

Obviously, this comparison of sampling methods excludes information on their relative costs of taking and processing a sample, which is highly relevant to their cost-effectiveness, but beyond the scope of this report. However the component costs of obtaining and processing RIVPACS and STAR-AQEM method samples were assessed and compared for some sites in a separate study within the STAR project and reported in Vlek (2004) which forms Deliverable N1 of the STAR project. Vlek (2004: section 5.3.3 found that, on average across the sampled sites, STAR-AQEM samples took 18 hours to process (including sorting and identification), whilst RIVPACS samples took only 9 hours – half the amount of time. Therefore the RIVPACS are likely to be more cost-effective than STAR-AQEM samples, at least when the aim is to base site assessments on one or more of the metrics assessed here.

5.6 REPLICATE SAMPLING VARIABILITY IN METRIC VALUES FOR OTHER 'NATIONAL' METHODS

5.6.1 Preliminary overall assessment

Many countries involved in the STAR project already had a sampling method for macroinvertebrates which was either already used on a national scale or was proposed for national use. These methods are hereafter referred to as the 'National' methods. Most of STAR partners participating in the field sampling programme who did not take samples using the RIVPACS methods, took and analysed samples using their 'National' method at all or most of the sites where they also took STAR-AQEM samples.

For most of these stream types, replicate 'National' method samples (in addition to the 'main' sample) were also taken at the same time at all, or nearly all, of the same selected subset of sites at which replicate STAR-AQEM samples were taken (Table 5.17). Taking replicate samples at the same set of sites using both methods provided valid direct comparisons of the sampling SD for individual metrics between the 'National' and STAR-AQEM method, because both sampling methods were then based on the same range of site qualities and within-site habitat heterogeneities – both of which could influence sampling variability in macroinvertebrate composition and derived metric values.

Table 5.17 Number of sites (and their site codes) in each stream type and country for which two 'National' method samples ('main' and 'replicate') were taken in at least one season (1=spring, 2=summer, 3=autumn). Site code 'xxx.y' indicates replicate samples only taken at site 'xxx' in season 'y'.

Country (National method)	Stream Type	Seasons sampled	n Sites	STAR site codes
Czech Republic (PERLA)	C04	1 + 2	3	614 620 625
	C05	1 + 2	3	713 717 722
France (IBGN)	F08	1 + 3	6	724 725 726 728 729 733
Italy (IBE)	I06	1 + 2	6	836 837.2 840 842 843 845
Denmark (DFSI)	K02	1 + 2	6	662 663 665 667 671 673
Latvia	L02	1 + 2	6	997 1006 1007 1010 1016 1017
Poland	O02	1 + 3	7	895 897 903 913 915.3 916 1036
Portugal (PMP)	P04	1 + 3	6	863 864 865 866 867 868
Sweden	S05	1 + 3	3	685 689 691
	S06	1 + 3	3	875 876 878

Figures 5.18-5.20 plot the difference in (untransformed) metric values between two replicate samples taken by the various 'National' methods against the average of the two values for a range of metrics. Each point denotes an individual site in one season. The differences in metric values of the two replicate samples rather than their SD are shown in these initial plots to aid visual understanding. All of the 'National' methods are shown together on the same plot for compactness, but differences in replicate sampling variability and sampling SD (SD_E) should be interpreted with caution as the different methods may collect and identify varying amounts of macroinvertebrates. For example, methods recording larger numbers of taxa (or families) at a site are likely to record greater inter-sample variation in number of taxa (or families). The sampling variance expressed as a percentage (P_{samp}) of the total variance in the metric's values within a stream type is more appropriate for comparing the sampling precision of methods.

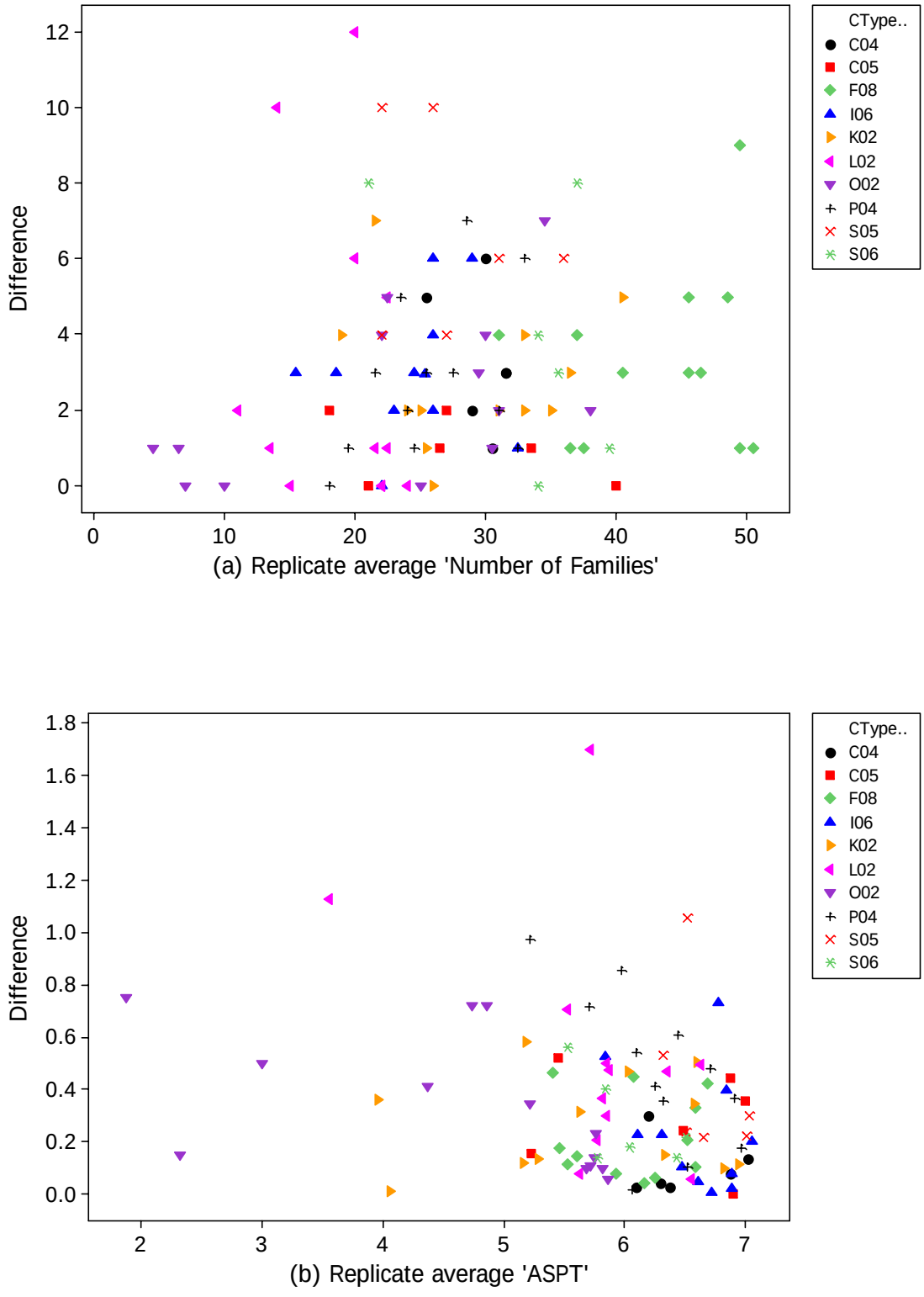


Figure 5.16 'National' sampling methods Difference in metric values between two replicate samples plotted against the average of the two values for untransformed values of metrics (a) 'Number of Families' and (b) 'Average Score per Taxon' (ASPT).

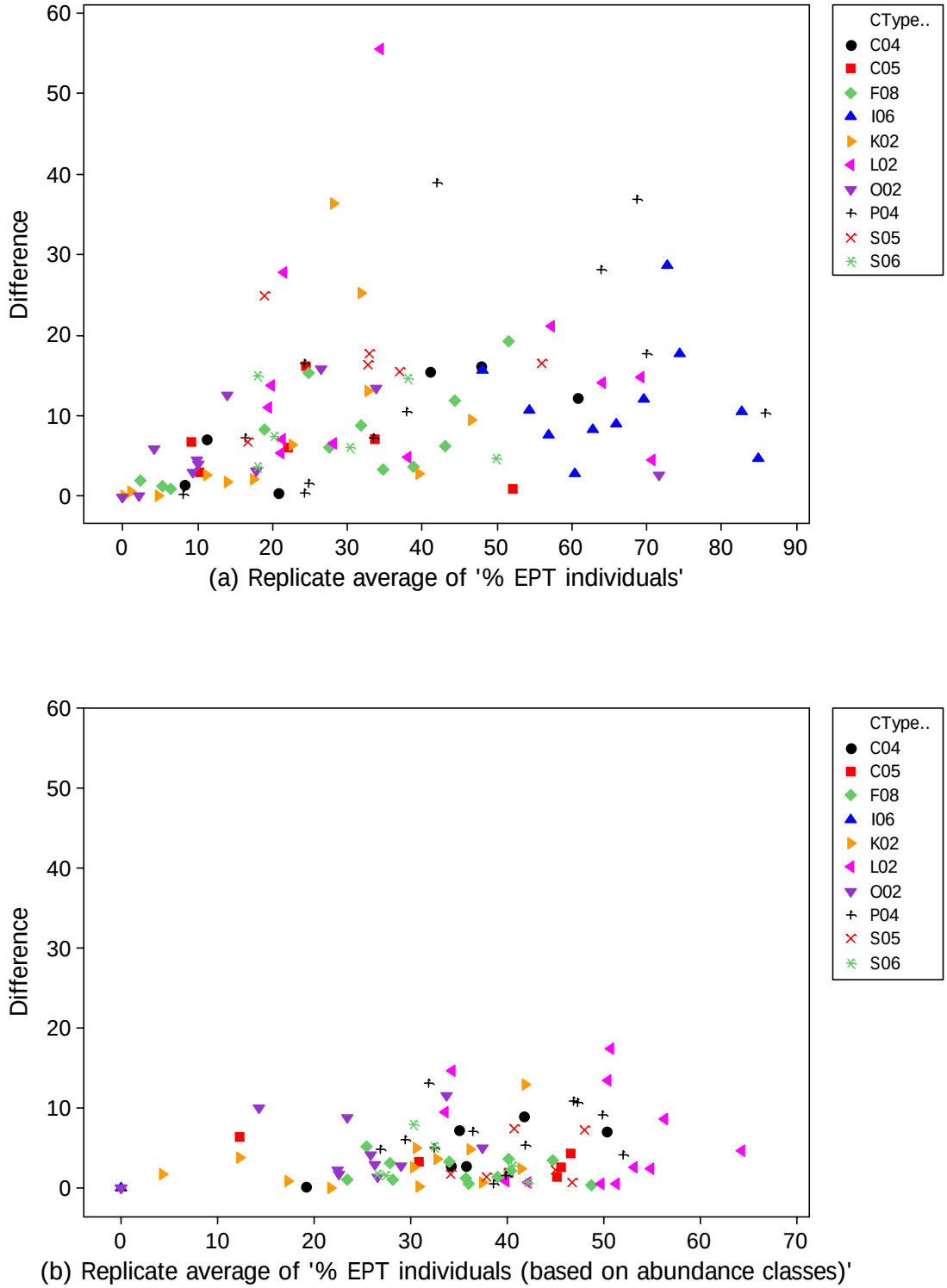


Figure 5.17 'National' sampling methods Difference in metric values between two replicate samples plotted against the average of the two values for untransformed values of metrics (a) '% EPT individuals' and (b) '% EPT individuals (based on abundance classes)' – not available for type I06.

.

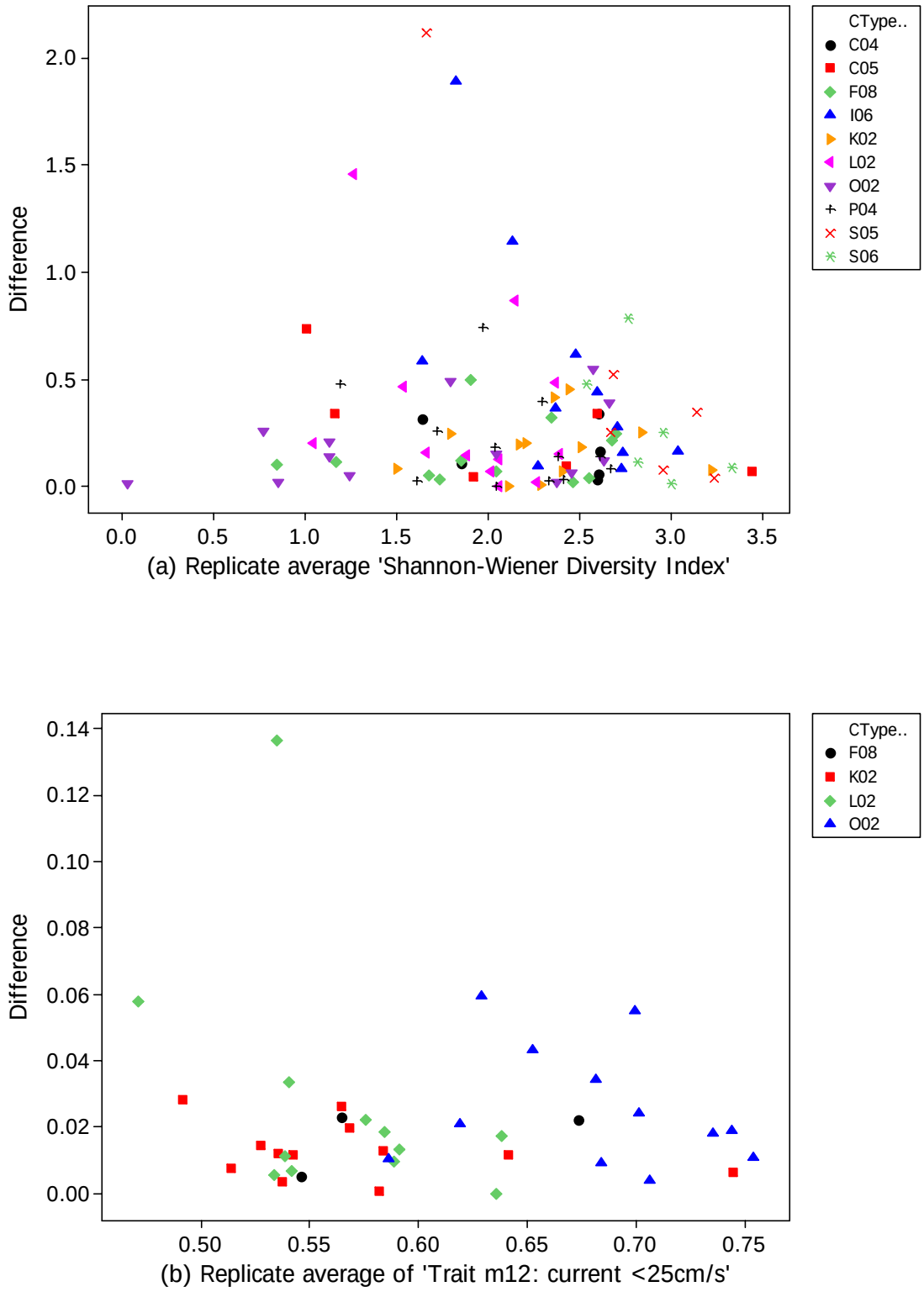


Figure 5.18 'National' sampling methods Difference in metric values between two replicate samples plotted against the average of the two values for untransformed values of metrics (a) 'Shannon-Wiener Diversity index' and (b) 'Trait m12: current <25cm/s'

Estimates of the replicate sampling standard deviation (SD_E) for both untransformed and transformed values of each of the metrics are given in Tables 5.18 and 5.19 respectively, with separate estimates for each STAR stream type. The use of some metrics given in these tables may be inappropriate for some stream types, sampling methods or the taxonomic resolution used, but they are given here for completeness.

Table 5.18 'National' methods: Estimates of the overall replicate sampling standard deviation (SD_E) in untransformed metric values for each STAR stream type.

Method	Czech		French	Italian	Danish	Latvian	Polish	Portu guese	Swedish	
Stream Type	C04	C05	F08	I06	K02	L02	O02	P04	S05	S06
Abundance [ind/m²]	2646	1802	10100	702	548	709	3715	679	452	2119
Number of taxa	5.00	2.52	2.75	2.75	3.08	4.06	4.02	3.79	11.94	6.47
Number of Families	2.65	0.91	2.84	2.46	2.38	3.61	2.20	2.48	5.03	3.58
Number of EPT taxa	1.12	2.22	1.26	1.58	1.38	2.35	2.01	2.30	5.35	1.78
Saprobic Index	0.061	0.050	0.000	0.409	0.043	0.134	0.103	0.078	0.034	0.094
German Saprobic new	0.044	0.082		0.211	0.059	0.245	0.071	0.043	0.059	0.045
Czech Saprobic	0.152	0.090		0.136	0.116	0.171	0.151	0.092	0.049	0.055
ASPT	0.098	0.237	0.186	0.227	0.226	0.494	0.296	0.384	0.369	0.223
IBE	0.635	0.387	0.490	0.736	0.329	1.131	1.011	1.185	0.893	0.500
Diversity SW	0.147	0.258	0.146	0.527	0.164	0.380	0.184	0.213	0.642	0.280
% Rheophilic	5.04	8.02	5.74	11.61	6.87	11.04	4.85	16.91	18.86	9.60
% Rheophilic (ab-class)	5.56	4.75	2.62		4.07	6.45	2.98	6.09	7.42	5.49
% Littoral	2.63	4.46	2.53		4.54	6.55	2.05	8.28	7.55	4.95
% Grazers/Scrapers	2.20	1.41	2.56	7.44	4.54	3.91	1.69	3.74	6.31	4.48
% Shredders	3.07	5.08	1.80		4.29	3.11	2.74	1.27	1.49	1.92
% Gatherers/Collectors	3.36	2.13	5.30		9.17	6.10	5.31	8.17	10.92	4.41
% Oligochaeta	2.76	0.57	4.96	1.07	9.90	8.19	7.44	4.83	2.46	3.03
% EPT individuals	7.69	5.81	6.46	9.51	9.77	14.75	5.14	13.84	12.12	6.90
% EPT (ab-class)	4.07	2.65	1.88		3.31	6.13	3.83	5.29	3.14	2.95
% EPT Taxa	3.70	3.45	2.30	3.72	2.17	6.04	4.74	5.01	1.49	3.10
RETI	0.041	0.045	0.045	0.095	0.070	0.054	0.060	0.067	0.083	0.067
1 –GOLD	0.048	0.091	0.077	0.065	0.110	0.149	0.082	0.128	0.138	0.072
Trait m1:max size ≤1cm			0.011		0.018	0.049	0.074			
Trait m2 : >1 cycle			0.019		0.029	0.047	0.051			
Trait m7 : crawler loco.			0.010		0.024	0.046	0.075			
Trait m12:current<25cm			0.013		0.011	0.032	0.022			

The difference between replicates in the 'Number of families' recorded is less than four in the majority of sites for all 'National' methods, except for the Swedish method (Figure 5.16). For the Swedish 'National' method, the difference in recorded 'Number of Families' is between 4 and 10 for many sites, leading to the highest replicate sampling variance with SD_E for untransformed 'Number of Families' of 11.94 and 6.47 for stream types S05 and S06 respectively (Table 5.18). Even after transformation to the square root scale, the sampling SD for these two stream types is still higher than that for any other stream type and 'National' method (Table 5.19), including nearly all of those based on the 'RIVPACS' method (Tables 5.12-5.15). Only stream type (U15) in the UK when sampled by the RIVPACS method had such a high level of sampling variability in transformed 'Number of Families'. Sampling variability in 'Number of Families' was also very high for two sites sampled using the Latvian 'National' method, with replicate values of 9 and 19 for one site and 14 and 26 for the other (Figure 5.16).

Table 5.19 'National' methods: Estimates of the overall replicate sampling standard deviation (SD_E) in transformed metric (f(x)) values for each STAR stream type.

Method	Czech		French	Italian	Danish	Latvian	Polish	Portu guese	Swedish	
Stream Type	C04	C05	F08	I06	K02	L02	O02	P04	S05	S06
Abundance [ind/m²]	0.587	0.601	0.983	1.148	0.411	0.657	1.202	0.566	0.628	0.751
Number of taxa	0.357	0.185	0.201	0.262	0.252	0.475	0.312	0.295	0.876	0.507
Number of Families	0.247	0.096	0.213	0.248	0.233	0.436	0.214	0.237	0.500	0.340
Number of EPT taxa	0.122	0.251	0.164	0.204	0.214	0.401	0.332	0.305	0.589	0.223
Saprobic Index	0.061	0.050			0.043	0.134	0.103	0.078	0.034	0.094
German Saprobic new	0.044	0.082	0.211		0.059	0.245	0.071	0.043	0.059	0.045
Czech Saprobic	0.152	0.090	0.136		0.116	0.171	0.151	0.092	0.049	0.055
ASPT	0.098	0.237	0.186	0.227	0.226	0.494	0.296	0.384	0.369	0.223
IBE	0.635	0.387	0.490	0.736	0.329	1.131	1.011	1.185	0.893	0.500
Diversity SW	0.147	0.258	0.146	0.527	0.164	0.380	0.184	0.213	0.642	0.280
% Rheophilic	0.052	0.110	0.060	0.137	0.080	0.141	0.069	0.192	0.220	0.100
% Rheophilic (ab-class)	0.057	0.049	0.029		0.045	0.070	0.073	0.067	0.075	0.057
% Littoral	0.028	0.053	0.030		0.049	0.098	0.038	0.086	0.081	0.054
% Grazers/Scrapers	0.026	0.017	0.029	0.103	0.054	0.050	0.026	0.051	0.090	0.048
% Shredders	0.050	0.054	0.032		0.053	0.066	0.047	0.031	0.053	0.035
% Gatherers/Collectors	0.042	0.023	0.058		0.096	0.064	0.060	0.086	0.150	0.052
% Oligochaeta	0.038	0.025	0.078	0.052	0.132	0.096	0.119	0.085	0.065	0.068
% EPT individuals	0.082	0.073	0.071	0.107	0.112	0.173	0.070	0.152	0.142	0.082
% EPT (ab-class)	0.042	0.034	0.020		0.036	0.063	0.046	0.055	0.032	0.032
% EPT Taxa	0.038	0.042	0.025	0.038	0.025	0.062	0.056	0.053	0.015	0.033
RETI	0.043	0.048	0.049	0.120	0.072	0.057	0.068	0.078	0.115	0.069
1 –GOLD	0.056	0.118	0.083	0.094	0.118	0.168	0.094	0.144	0.159	0.075
Trait m1:max size ≤1cm			0.011		0.018	0.050	0.075			
Trait m2 : >1 cycle			0.020		0.031	0.049	0.054			
Trait m7 : crawler loco.			0.010		0.024	0.046	0.083			
Trait m12:current<25cm			0.014		0.011	0.032	0.024			

Figure 5.17 shows the distribution of differences between replicate samples taken using the 'National' methods in each of two metrics measuring the percentage of all individuals at the site which are EPT taxa (Emphemeroptera + Plecoptera + Trichoptera). It is clear that sampling variability is usually less when the metric is based on abundance classes ('%EPT abundance-classes' in Figure 5.17(b)) than when based on raw abundance ('%EPT individuals' in Figure 5.17(a)).

The next sub-sections describe the estimates of sampling SD of metric values for each 'National' method and country in turn. In particular, Tables 5.20-5.27 give estimates of the replicate sampling SD (SD_E) and sampling variance expressed as a percentage (P_{samp}) of the total variance in metric values within a stream type; with a direct comparison of these estimates for the 'National' method with those obtained with samples taken and processed using the STAR-AQEM method from all (or almost all) of the same sites at the same time.

5.6.2 Czech 'National' (PERLA) method

The Czech PERLA method is described in section 2.1.2.7. Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the Czech 'National' (PERLA) method are given for selected metrics in Table 5.20. The estimates are provided averaged across the two sampled stream types C05 and C06; but estimates of SD_E for the individual stream types can be obtained from Table 5.18 (untransformed metrics) and Table 5.19 (transformed, where appropriate). Comparisons are made with samples taken using the STAR-AQEM method at the same sites at the same time.

Table 5.20 Czech 'national' method (PERLA): Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method, averaged across stream types C04 and C05. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.594**	0.600	**8**	18
Number of taxa	√x	**0.284**	0.342	**5**	6
Number of Families	√x	**0.187**	0.274	**5**	16
Number of EPT taxa	√x	**0.198**	0.281	**3**	6
Saprobic Index	x	**0.056**	0.048	3	2
German Saprobic new	x	**0.066**	0.079	**3**	5
Czech Saprobic	x	**0.125**	0.132	3	3
ASPT	x	**0.181**	0.271	**3**	8
Diversity SW	x	0.210	0.191	8	6
% Rheophilic	asin	0.086	0.080	**6**	7
% Rheophilic (abund classes)	asin	0.053	0.041	8	8
% Littoral	asin	0.042	0.039	5	4
% Grazers/Scrapers	asin	**0.022**	0.035	**2**	8
% Shredders	asin	0.052	0.048	4	4
% Gatherers/Collectors	asin	**0.034**	0.054	**3**	7
% Oligochaeta	asin	**0.032**	0.074	**1**	5
% EPT individuals	asin	0.078	0.069	**10**	12
% EPT (abund. classes)	asin	**0.038**	0.043	**5**	7
% EPT Taxa	asin	**0.040**	0.059	**7**	17
RETI	asin	**0.046**	0.049	**4**	5
1 –GOLD	asin	0.092	0.078	7	7
Average (range)				5 (1-10)	8 (2-18)
Metrics (%) with P_{samp} < 10%				20 (95%)	17 (81%)
Number of metrics with smaller value of SD_E or P_{samp}		14	7	14	3

The PERLA method has a very high sampling precision within the two sampled stream types with values of P_{samp} of 10% or less for all of the metrics analysed. The precision is usually roughly the same as for the STAR-AQEM method; where they differ the PERLA method appears to be more precise. For example, the value of P_{samp} for 'Number of Families' is estimated to be 16% for the STAR-AQEM method, but only 5% for the PERLA method.

Overall, the Czech 'National' method gives very low variability in metric values between replicate samples and hence high sampling precision, mostly as good or better than that for the STAR-AQEM method (Table 5.20).

5.6.3 French 'National' (BGN) method

The French IBGN method of sampling and sample processing is described in section 2.1.2.4. Within STAR, samples taken using the IBGN method were identified to family level for most groups, but samples taken from the same French sites using the STAR-AQEM method were identified to a more detailed or level. This will explain some differences between the two sampling methods.

Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the French 'National' method are given for selected metrics in Table 5.21. Comparisons are made with samples taken using the STAR-AQEM method at the same sites at the same time.

Table 5.21 French 'National' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method for stream type F08. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.983**	1.185	**34**	55
Number of taxa	√x	**0.201**	0.321	18	9
Number of Families	√x	**0.213**	0.314	19	18
Number of EPT taxa	√x	**0.164**	0.358	**7**	15
ASPT	x	**0.186**	0.265	**14**	15
Diversity SW	x	**0.146**	0.168	**6**	9
% Rheophilic	asin	**0.060**	0.109	**5**	13
% Rheophilic (abund classes)	asin	**0.029**	0.065	**8**	9
% Littoral	asin	**0.030**	0.090	**5**	40
% Grazers/Scrapers	asin	**0.029**	0.041	5	3
% Shredders	asin	**0.032**	0.043	**6**	10
% Gatherers/Collectors	asin	**0.058**	0.079	10	8
% Oligochaeta	asin	**0.078**	0.112	**17**	40
% EPT individuals	asin	**0.011**	0.098	**10**	25
% EPT (abund. classes)	asin	**0.020**	0.047	**5**	25
% EPT Taxa	asin	**0.025**	0.064	**8**	44
RETI	asin	**0.049**	0.053	5	3
Trait m1 : max size ≤ 1cm	asin	**0.011**	0.048	**11**	64
Trait m2 : >1 cycle	asin	0.020	0.019	**8**	10
Trait m7 : crawler locomotion	asin	**0.010**	0.022	**25**	31
Trait m12 : current <25cm/s	asin	**0.014**	0.031	**5**	19
Average (range)				11 (5-34)	22 (3-64)
Metrics (%) with P_{samp} < 10%				12 (57%)	6 (29%)
Number of metrics with smaller value of SD_E or P_{samp}		20	1	16	5

The French 'National' method gives high sampling precision (i.e. P_{samp} <10%) for just over half (57%) of the metrics assessed and especially for those metrics based on percentage abundance of taxa with selected characteristics. Sampling SD were less variable between replicate samples taken using the 'National' method than for those taken using the STAR-AQEM method 20 of the 21 metrics assessed; and this was also true to a lesser extent when expressed in terms of percentage (P_{samp}) of total variance within stream type (F08). The two methods give more similar lower precisions for the metrics based on taxonomic richness, namely for 'Number of taxa' and 'Number of Families' (Table 5.21).

5.6.4 Italian 'National' (IBE) method

The Italian IBE method of sampling and sample processing is described in section 2.1.2.5.

Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the Italian 'National' (IBE) method are given for selected metrics in Table 5.22. Comparisons are made with samples taken using the STAR-AQEM method at the same time at the same sites in stream type I06. Fewer metrics were assessed because all Italian macroinvertebrate samples (both IBE and STAR-AQEM) were only identified to family level, and some metrics required lower level identification.

Table 5.22 Italian 'National' method (IBE): Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	1.311	0.995	73	48
Number of taxa	√x	**0.258**	0.496	**21**	66
Number of Families	√x	**0.245**	0.478	**25**	63
Number of EPT taxa	√x	**0.211**	0.430	**12**	77
ASPT	x	0.263	0.211	38	18
IBE	x	**0.831**	0.897	**37**	47
Diversity SW	x	0.520	0.236	56	36
% Grazers/Scrapers	asin	0.100	0.063	49	24
% Oligochaeta	asin	**0.052**	0.057	**14**	36
% EPT individuals	asin	0.110	0.095	**25**	31
% EPT Taxa	asin	**0.038**	0.049	**18**	20
RETI	asin	0.115	0.039	67	10
1 –GOLD	asin	0.098	0.076	22	20
Average (range)				35 (12-73)	38 (10-77)
Metrics (%) with P_{samp} < 10%				0	0
Number of metrics with smaller value of SD_E or P_{samp}		6	7	7	6

The Italian (IBE) 'National' method of sampling and subsequent sorting and processing of samples appears to lead to relatively high variability in metric values between replicate samples, at least relative to the total variance in metric values amongst the sampled sites within the stream type I06 for which data were available. The replicate sampling variance was not less than 10% of the total variance within the stream type for any of the 13 metrics analysed.

However, it is very interesting to note that on average, the Italian IBE method appeared to give as repeatable results and the same replicate sampling precision as the STAR-AQEM sampling method in terms of both the sampling SD (SD_E) and the percentage sampling variance (P_{samp}) of metrics (Table 5.22). Both methods, at least as carried out in Italy for the Italian sites within the STAR project, give amongst the highest overall percentage sampling variances and lowest precisions of all countries, stream types and methods.

5.6.5 Danish 'National' (DFSI) method

The Danish DSFI method of sampling and sample processing is described in section 2.1.2.3.2.

Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the Danish 'National' (DFSI) method are given for selected metrics in Table 5.23. Comparisons are made with samples taken using the STAR-AQEM method at the same time at the same sites..

Table 5.23 Danish 'National' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method for stream type K02. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.411**	0.673	**18**	60
Number of taxa	√x	**0.252**	0.357	**12**	20
Number of Families	√x	**0.233**	0.300	**17**	22
Number of EPT taxa	√x	**0.214**	0.377	**5**	13
Saprobic Index	x	0.043	0.030	3	1
German Saprobic new	x	0.059	0.041	4	2
Czech Saprobic	x	0.116	0.102	9	6
ASPT	x	**0.226**	0.293	**5**	9
Diversity SW	x	**0.164**	0.362	**13**	41
% Rheophilic	asin	**0.080**	0.143	**4**	13
% Rheophilic (abund classes)	asin	**0.045**	0.057	**6**	7
% Littoral	asin	**0.049**	0.085	**10**	25
% Grazers/Scrapers	asin	**0.054**	0.077	**16**	23
% Shredders	asin	**0.053**	0.059	18	18
% Gatherers/Collectors	asin	**0.096**	0.109	30	27
% Oligochaeta	asin	**0.132**	0.158	**33**	48
% EPT individuals	asin	**0.112**	0.126	22	19
% EPT (abund. classes)	asin	**0.036**	0.055	**7**	13
% EPT Taxa	asin	**0.025**	0.067	**4**	21
RETI	asin	**0.072**	0.104	**18**	24
Trait m1 : max size ≤ 1cm	asin	**0.018**	0.037	**9**	41
Trait m2 : >1 cycle	asin	**0.031**	0.038	**9**	12
Trait m7 : crawler locomotion	asin	**0.024**	0.035	**10**	21
Trait m12 : current <25cm/s	asin	**0.011**	0.024	**6**	11
Average (range)				13 (3-39)	22 (1-60)
Metrics (%) with P_{samp} < 10%				12 (50%)	5 (21%)
Number of metrics with smaller value of SD_E or P_{samp}		21	3	18	5

The Danish Stream Fauna Index (DSFI) was not included amongst the analysed metrics partly because it only takes the seven possible integer values (1-7), so its sampling variation needs to be summarised in a different way, but also because, it is so close to the five class system of the WFD, it might be revised to immediate conform to five values (Friberg *pers comm.*).

The sampling SD (SD_E) for the Danish 'National' method was less than that based on STAR-AQEM method at the same sites and time for 21 of the 24 metrics analysed (Table 5.23).

More impressively, 75% (18) of the metrics had lower percentage sampling variances within the stream type for the Danish 'National' method than for the STAR-AQEM method.

Half of the 24 metrics assessed had replicate sampling variances which formed less than 10% of the total variance in the metric values amongst all sites within the stream type (Table 5.23). Moreover, only three metrics measuring percentage abundance of selected taxa (%Gatherers/Collectors, %Oligochaeta and %EPT individuals) had values of P_{samp} greater than 20%.

Thus the Danish (DFSI) method seems to lead, in most cases, to metric values with low sampling variances and thus high sampling precision and repeatability.

5.6.6 Latvian 'National' method

The Latvian method is described in section 2.1.2.9. Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the Latvian 'National' method are given for selected metrics in Table 5.24. Comparisons are made with samples taken using the STAR-AQEM method at the same time at the same sites in stream type L02.

Table 5.24 Latvian 'national' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method for stream type L02. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.657**	0.676	31	30
Number of taxa	√x	0.475	0.185	57	13
Number of Families	√x	0.436	0.175	55	12
Number of EPT taxa	√x	0.401	0.191	70	16
Saprobic Index	x	**0.134**	0.215	**14**	46
ASPT	x	0.494	0.375	43	31
Diversity SW	x	0.380	0.230	56	13
% Rheophilic	asin	0.141	0.092	32	11
% Rheophilic (abund classes)	asin	0.070	0.042	43	14
% Littoral	asin	0.098	0.079	20	14
% Grazers/Scrapers	asin	0.050	0.042	**19**	34
% Shredders	asin	0.066	0.038	46	16
% Gatherers/Collectors	asin	0.064	0.036	39	21
% Oligochaeta	asin	0.096	0.063	30	16
% EPT individuals	asin	0.173	0.086	47	13
% EPT (abund. classes)	asin	0.063	0.036	37	9
% EPT Taxa	asin	0.062	0.034	54	10
RETI	asin	0.057	0.056	**19**	39
Trait m1 : max size ≤ 1cm	asin	0.050	0.031	61	28
Trait m2 : >1 cycle	asin	0.049	0.028	48	33
Trait m7 : crawler locomotion	asin	0.046	0.016	98	18
Trait m12 : current <25cm/s	asin	0.032	0.019	47	15
Average (range)				43 (14-98)	21 (9-96)
Metrics (%) with P_{samp} < 10%				0 (0%)	1 (5%)
Number of metrics with smaller value of SD_E or P_{samp}		1	21	3	19

When using the Latvian 'National' method, all of the 22 metrics assessed had replicate sampling variances which contributed more than 10% of the total variance in metrics values amongst all sites within the sampled stream type L02. The average percentage sampling variance (P_{samp}) was 43% - this was the highest average amongst all countries, stream types and methods. Replicate samples were taken using the STAR-AQEM standard method at all of the Latvian sites at which replicate samples were taken by the 'National' method. Appendix 1 Figure 8d shows that the number of families recorded in 'National' method samples was on average only slightly less than that recorded in STAR_AQEM samples. However, the 'National' method gave higher values of sampling SD (SD_E) for all except one metric and higher values of P_{samp} for all except three of the 22 metrics analysed (Table 5.24).

In summary, the current Latvian 'National' method gives metrics values which are highly susceptible to sampling variation and hence of low precision – so it may benefit from refinement. These comparisons with the STAR-AQEM method exclude the relative costs and hence cost-effectiveness of the two methods.

5.6.7 Polish 'National' method

The Polish method is described in section 2.1.2.6. Estimates of the average sampling standard deviation (SD_E) between replicate samples taken using the Polish 'National' sampling method are given for selected metrics (transformed where appropriate) in Table 5.25. This Table also gives the sampling variance expressed as a percentage (P_{samp})of the total variance in metric values within the sampled stream type O02, together with the values of SD_E and P_{samp} values for the same sites for comparison.

Table 5.25 Polish 'national' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method for stream type O02. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	1.202	0.984	20	23
Number of taxa	√x	**0.312**	0.317	4	3
Number of Families	√x	0.214	0.172	3	2
Number of EPT taxa	√x	0.332	0.185	4	1
Saprobic Index	x	0.103	0.102	**3**	5
ASPT	x	0.296	0.247	6	3
Diversity SW	x	**0.184**	0.240	**6**	7
% Rheophilic	asin	**0.069**	0.313	**11**	74
% Rheophilic (abund classes)	asin	**0.073**	0.311	**8**	94
% Littoral	asin	**0.038**	0.084	**7**	44
% Grazers/Scrapers	asin	**0.026**	0.070	**4**	41
% Shredders	asin	0.047	0.042	9	7
% Gatherers/Collectors	asin	**0.060**	0.102	**4**	16
% Oligochaeta	asin	**0.119**	0.248	**7**	56
% EPT individuals	asin	0.070	0.049	10	4
% EPT (abund. classes)	asin	0.046	0.031	5	2
% EPT Taxa	asin	0.056	0.032	6	2
RETI	asin	**0.068**	0.081	**12**	20
Trait m1 : max size ≤ 1cm	asin	**0.075**	0.086	**63**	99
Trait m2 : >1 cycle	asin	**0.054**	0.102	**21**	73
Trait m7 : crawler locomotion	asin	0.083	0.083	**38**	72
Trait m12 : current <25cm/s	asin	0.024	0.020	21	14
Average (range)				12 (3-63)	30 (1-99)
Metrics (%) with P_{samp} < 10%				15	10
Number of metrics with smaller value of SD_E or P_{samp}		11	9	12	8

The Polish 'National' method for macroinvertebrate sampling appears to give replicate sampling SD for metric values of about the same absolute size as for many other 'national' methods (Tables 5.16, 5.18 and 5.19). However, once expressed as a percentage of the total variances in metrics values with across all sites with the sampled stream type (O02), the Polish 'National' method appears to have high relative precision with values of P_{samp} of less than 10% for 15 of the 20 metrics assessed (Table 5.25). The four traits metrics had the lowest relative precision with P_{samp} ranging from 21-63%.

The Polish 'National' method was noticeably more precise for some of the metrics based on selective percentage abundances (e.g. %Rheophilic, %Littoral, %Grazers/Scrapers).

5.6.8 Portuguese 'National' (PMP) method

The Portuguese PMP method is described in section 2.1.2.8. Estimates of the average sampling standard deviation (SD_E) between replicate samples taken using the Portuguese 'National' sampling method are given for selected metrics (transformed where appropriate) in Table 5.26. This Table also gives the sampling variance expressed as a percentage (P_{samp})of the total variance in metric values within the sampled stream type P04, together with the values of SD_E and P_{samp} values for the same sites for comparison.

Table 5.26 Portuguese 'national' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method for stream type P04. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.566**	0.711	31	22
Number of taxa	√x	**0.295**	0.358	**15**	31
Number of Families	√x	**0.237**	0.331	**15**	35
Number of EPT taxa	√x	**0.305**	0.346	**11**	28
Saprobic Index	x	**0.078**	0.091	18	17
ASPT	x	**0.384**	0.534	**29**	43
IBE	x	1.185	0.618	49	18
Diversity SW	x	**0.213**	0.354	**13**	35
% Rheophilic	asin	0.192	0.176	30	21
% Rheophilic (abund classes)	asin	**0.067**	0.068	37	15
% Littoral	asin	**0.086**	0.089	50	36
% Grazers/Scrapers	asin	**0.051**	0.067	**11**	21
% Shredders	asin	**0.031**	0.035	**6**	7
% Gatherers/Collectors	asin	**0.086**	0.104	**15**	19
% Oligochaeta	asin	**0.085**	0.121	**22**	38
% EPT individuals	asin	0.152	0.094	22	9
% EPT (abund. classes)	asin	**0.055**	0.060	**11**	15
% EPT Taxa	asin	**0.053**	0.068	**12**	26
RETI	asin	0.078	0.075	23	20
1 –GOLD	asin	0.144	0.093	19	9
Average (range)				22 (6-50)	23 (7-43)
Metrics (%) with P_{samp} < 10%				1 (5%)	3 (15%)
Number of metrics with smaller value of SD_E or P_{samp}		15	5	11	9

Using the Portuguese 'National' method gave replicate sampling variance in metric values which on average, contributed 22% (range 6-50%) of the total variance in metric values amongst all sampled sites within stream type P04.

When compared with the STAR-AQEM method used at the same time at the same site, the 'National' method gave smaller estimates of sampling SD (SD_E) for 75% of the metric assessed. However, once expressed as percentage sampling variance (P_{samp}), the Portuguese 'National' method appears to give about the same average level of sampling variance and precision as the STAR-AQEM method (Table 5.26).

The new metric '1-GOLD' (equal to the proportion of identified individuals which are not selected Gastopoda, Oligochaeta or Diptera) and proposed by the Portuguese partners, had a estimated percentage sampling variance of 19% when based on samples obtained using the Portuguese method compared to 9% using STAR-AQEM method.

5.6.9 Swedish 'National' method

The Swedish method of sampling and sample processing is described in section 2.1.2.3.1.

Estimates of the average sampling standard deviation (SD_E) and percentage sampling variance (P_{samp}) due to variation between replicate samples taken using the Swedish 'National' method are given for selected metrics in Table 5.27. The estimates are provided averaged across the two sampled stream types S05 and S06; but estimates of SD_E for the individual stream types can be obtained from Table 5.18 (untransformed metrics) and Table 5.19 (transformed, where appropriate). Comparisons are made with samples taken using the STAR-AQEM method at the same sites at the same time.

Table 5.27 Swedish 'national' method: Estimates of the average replicate sampling SD (SD_E) and percentage sampling variance (P_{samp}) in transformed (f(x)) metric values and comparisons with the STAR-AQEM method, averaged across stream types S05 and S06. Values in bold denote National < STAR-AQEM.

		SD_E		P_{samp}	
Metric	f(x)	National	STAR-AQEM	National	STAR-AQEM
Abundance [ind/m²]	√√x	**0.692**	1.028	**21**	30
Number of taxa	√x	0.716	0.415	60	32
Number of Families	√x	0.427	0.296	39	30
Number of EPT taxa	√x	0.445	0.253	26	10
Saprobic Index	x	0.071	0.035	10	3
ASPT	x	0.304	0.232	19	8
Diversity SW	x	0.495	0.218	93	22
% Rheophilic	asin	0.171	0.086	49	14
% Rheophilic (abund classes)	asin	0.067	0.046	24	14
% Littoral	asin	0.069	0.047	21	9
% Grazers/Scrapers	asin	0.072	0.054	28	18
% Shredders	asin	0.045	0.041	26	16
% Gatherers/Collectors	asin	0.112	0.044	64	12
% Oligochaeta	asin	0.067	0.036	21	10
% EPT individuals	asin	**0.116**	0.124	**30**	32
% EPT (abund. classes)	asin	**0.032**	0.036	**6**	7
% EPT Taxa	asin	**0.026**	0.032	**5**	6
RETI	asin	0.095	0.055	**7**	13
Average (range)				30 (5-93)	16 (3-32)
Metrics (%) with P_{samp} < 10%				3 (17%)	5 (28%)
Number of metrics with smaller value of SD_E or P_{samp}		4	14	5	13

The replicate sampling SD of all except four of the 18 metrics assessed was higher using the Swedish 'National' method than using the STAR-AQEM method. However, especially for some richness-related metrics, this could be partly because the Swedish 'National' method tends to record more taxa (average = 48 and 53 taxa per sample in stream types S05 and S06) than the STAR-AQEM method (average = 43 and 44 respectively).

Once sampling variance is standardised as a percentage of total variance for that method within each stream type, the Swedish 'National' method is still less precise than the STAR-AQEM method for the majority of metrics. The 'National' method

was as precise or slightly better for %EPT individuals (based on abundance classes), %EPT Taxa and the RETI metrics – all of which were the only metrics for which P_{samp} was less than 10%.

Only five of the 18 metrics had percentage sampling variances less than 20% and the average value was 30% when based on the 'National' method. This compares with an average value of P_{samp} of 16% when metric values are based on samples obtained using the STAR-AQEM method at the same time at the same sites.

These comparisons with the STAR-AQEM method ignore the relative sampling and processing times, costs and hence cost-effectiveness of the two methods. They merely estimate relative sampling precision.

5.7 IMPLICATIONS FOR UNCERTAINTY IN ASSESSMENTS OF THE ECOLOGICAL STATUS OF SITES

The STAR replicated field sampling programme and the subsequent statistical analysis of variation in metric values reported here, have in many cases provided the first systematic quantitative estimates and comparisons of the sampling variances of a wide range of commonly used metrics for both the STAR-AQEM and a wide range of national sampling methods and procedures for macroinvertebrates.

The estimates of overall replicate sampling SD_E reported in the Tables above can be used in the STARBUGS simulation software package (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sampling variability in individual metric values on the uncertainty of multi-metric assessments of the ecological status of sites. STARBUGS was produced as Deliverable 9 within the STAR project and can be downloaded from the STAR web-site (www.eu-star.au).

On providing estimates of SD_E derived here as input to the software program STARBUGS, the program will use stochastic simulations of potential metric values for site to assign sites probabilistically to ecological status classes based on pre-defined class boundaries.

The analyses reported here concludes that for many metrics their sampling variation was less variable when the metric values (x) were transformed to either the square root ($\sqrt{x}$), double square root ($\sqrt{\sqrt{x}}$) or the arcsine transformation of the square root for proportions (arcsine($\sqrt{x}$)), or the arcsine($\sqrt{(x/100)}$) transformation for metrics which are percentages (Table 5.2).

The STARBUGS program requires the user to specify the transformation scale, if any, on which the supplied estimates of the sampling SD of each metric are based. This determines how potential values of each metric for the site are simulated, as explained in Table 5.28

Table 5.28 Mathematical procedure used to simulate random sampling values of metrics with sampling SD (σ) which are constant on a particular transformation scale. *X* denotes the user-supplied untransformed observed value for a site. *Z* denotes a random standard normal deviate with a mean of zero and SD of σ.

Transformation	Mathematical notation	Simulated value of metric in untransformed units
none	x	$X + Z$
square root	$\sqrt{x}$	$(\sqrt{X} + Z)^2$
double square root	$\sqrt{\sqrt{x}}$	$(\sqrt{\sqrt{X} + Z})^4$
arcsine square root for proportions	$\arcsin(\sqrt{x})$	$\sin(\arcsin(\sqrt{X}) + Z))^2$
arcsine square root for percentages	$\arcsin(\sqrt{(x/100)})$	$\sin(\arcsin(\sqrt{(X/100)}) + Z))^2$

Perhaps most importantly, the estimates of sampling SD derived here can also be used to provide information on the expected uncertainty in metrics values due to sampling variation for other sites in the same stream types sampling using the same method, but when only one sample has been taken at a point in time and thus there is no replication. Of course, the validity of their use depends on the assumption that exactly the same field sampling and laboratory sorting and identification procedures have been used by staff with a similar level of training.

As a guide to the likely levels of uncertainty in assignment of sites to ecological status classes, Figure 5.19 and Table 5.29 show the probability of misclassifying the a site of any particular true quality according to the size of the errors or uncertainty in the index (or metric) values expressed as a percentage of the width of the status classes for the index (see Clarke *et al.* 1996 for the mathematical derivation).

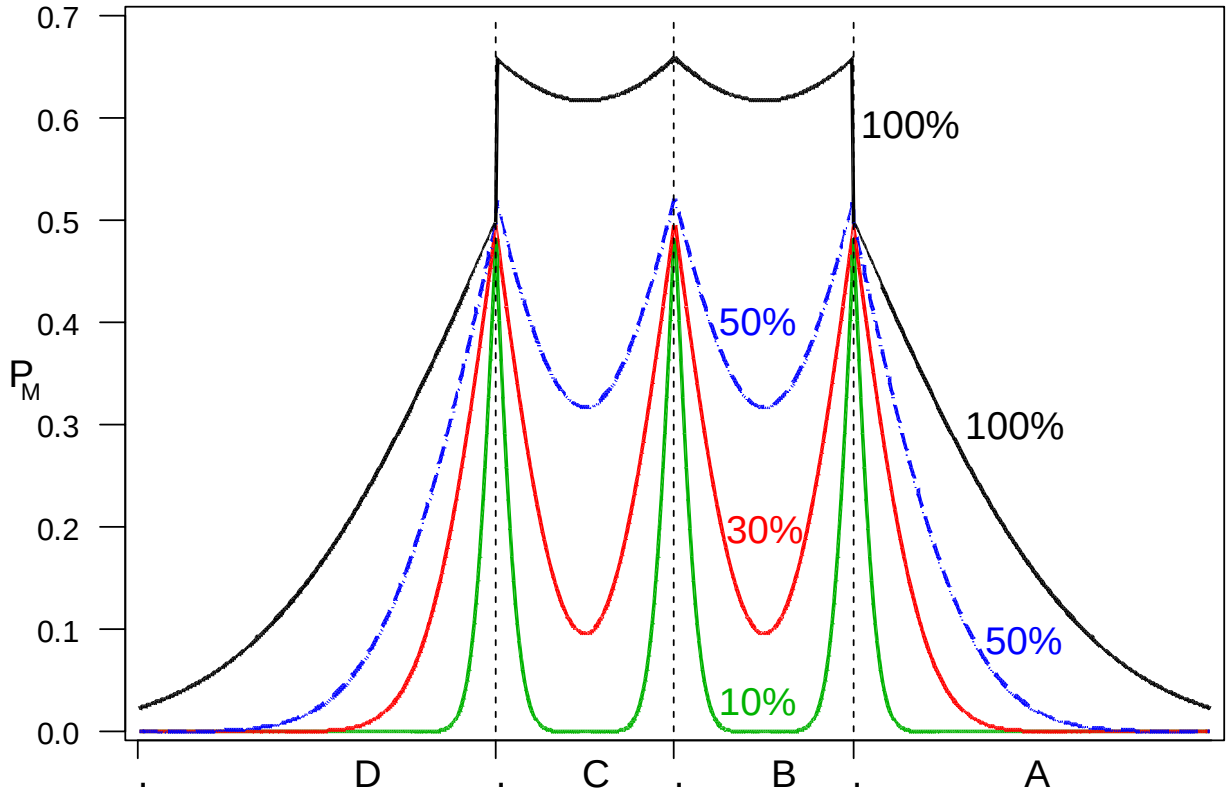


Figure 5.21 Plot of the probability (P_M) of classifying a site into a different status class versus its true index value (I_E) for a range of sampling standard deviations (σ) in the observed index value (I_O). For illustration, the index has been divided into four classes (A, B, C, D) with the middle two classes each of width W. Plots are shown for σ = 10%, 30%, 50% and 100% of W, where the broken line indicates the 50% plot (adapted from Figure2 of Clarke (2000)).

Table 5.29 Mean and range of misclassification rates (P_M) for sites with true qualities in a middle (i.e. not top or bottom) ecological status class for each of a range of error standard deviations (σ) in their observed quality index values, where σ is expressed as a percentage of the width of each middle class.

σ	Mean % misclassification	Range
10%	8%	0 - 50%
30%	24%	10 - 50%
50%	39%	32 - 52%
100%	63%	62 - 66%

When the sampling standard deviation of the index values is only 10% of the width of a band, then sites whose true quality lies in the middle of the status class would never be misbanded. Sites whose true quality lies on the border of two classes will always have at least a 50% chance of being placed in the wrong class. With error standard deviations of 10% of class width, the overall misclassification rate for sites in a middle class, such as bands B or C in Figure 5.19, (assuming an even spread of true qualities) is only 8% (Table 5.29). If however, the error standard deviation is 50% of the class width, then even the sites in the centre of a middle class have a roughly one in three chance of being placed in the wrong class and roughly 40% of all sites in the

class will be mis-placed into either a higher or lower class. If the error standard deviation is equal to the class width (i.e. 100%), as if possible for metrics with high sampling variability, then all sites whose true quality lies within a middle class will more likely than not be placed in the wrong class (Table 5.29). The probability of misbanding a site from the top or bottom classes (bands A or D in Figure 5.19) is only half that for middle classes, or only one-quarter if the (top or bottom) class width is twice that of the middle classes.

In summary:

The STAR project's extensive replicated sampling programme and the subsequent analysis of results has provided the first ever quantitative comparative study of the susceptibility of each of a wide range of established and 'National' macroinvertebrate sampling methods and a wide range of metrics to uncertainty resulting from the effects of field sampling variability subsequent sub-sampling and laboratory (or bank-side) procedures and protocols.

Further analyses of the results within this report, beyond the time limitations of the STAR project, would be useful to draw conclusions across sampling methods about the general relative precision of individual metrics. It is intended that such further interpretations will be included in an agreed paper on sampling variation effects within the STAR project special issue of the journal Hydrobiologia.

The preliminary estimates of sampling standard deviation reported here will provide valuable provisional estimates for use in any assessment of uncertainty in any single metric or multi-metric assessments of river quality based on any of the sampling methods tested within the STAR project.

6. SUMMARY

- In total, 15 partners in 13 countries took STAR-AQEM samples. For each STAR-AQEM site, samples were also taken using national macroinvertebrate assessment protocols to enable a comparison between methods. Eight different national assessment protocols were compared to the STAR-AQEM method. As some countries do not have a national protocol they used a slightly modified version of the UK RIVPACS sample protocol hereafter denoted the RIVPACS method. These countries were Austria, Germany and Greece.

- The data collected for this work package cover 13 countries (Austria, Czech Republic, Denmark, France, Germany, Greece, Italy, Latvia, Poland, Portugal, Slovakia, Sweden, and UK). The sampling included 22 stream types, where five were defined as being of the STAR project type "Core stream type 1" (mid altitude, 200-500 m.a.s.l., and with a "small" catchment area 10-100 km^2), seven were of the STAR project type "Core stream type 2" (lowland, <200 m.a.s.l., and "medium" catchment areas 100-1000 km^2), whereas ten other stream types were defined as STAR project type "Additional stream type" (having a different characterisation). These stream types are situated in 11 Ecoregions according to Illies definition (Illies, 1978; as used in the Water Framework Directive), these were regions 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, and 18.

- For each stream type in each country a pre-defined number of sites were sampled for each level of Ecological status (as defined by the Water Framework Directive), typically ca 3 'High' sites, ca 3 'Good' sites, ca 2 'Moderate' sites, and' ca 2 'Bad sites. Within each stream type one of three main types of stress were assessed, i.e., organic pollution/eutrophication, hydromorphological degradation, and toxic pollution.

- The majority of sampling methods employed by the different countries have many features in common. The majority of methods involve an *a priori* assessment of habitats at the sampling site, exceptions being the RIVPACS method and the DSFI method. In RIVPACS habitats are sampled in proportion to their occurrence, which is subjectively assessed by the surveyor while sampling. DSFI uses a fixed sampling grid that should cover most habitats without introducing a sampling bias due to variability in how surveyors assess number of habitats present. All methods except the Swedish method use a multi-habitat sampling approach. In contrast, it is the only method, which take replicate samples to assess inter-sample variability. Most methods use standard hand nets with a width of 25 cm and mesh bag with a 500 µm mesh size in accordance to the CEN standard EN 27 828. The samples are therefore semi-quantitative A Surber sampler can be used when employing the STAR-AQEM method, while it is obligatory when using the French IBGN protocol with the exception of sampling in lentic areas. The Polish method uses both a quantitative core sampler and a hand net. Mesh sizes used varies between 300 and 1000 µm. Three of the methods (RIVPACS, DSFI and PERLA) include a pick sample of attached macroinvertebrates.

- To allow an inter-comparison of methods used, a handling-processing score was given to each method. The score is based on giving the value 1 to each of

the handling-processing steps which are considered to be positive for overall assessment quality (0 if negative, i.e. a high score indicates a high quality method (8 is maximum). The handling-processing score ranges between 1 (IBE) to 7 (Swedish, Polish and Portuguese (PMP) methods) with most methods obtaining scores of either 4 or 5.

- All national metrics should be used (i.e. BMWP/ASPT, DSFI, IBE etc). when comparing classifications at the national level. However, it is not relevant to test national methods and specific metrics (e.g. for a certain stream type) on the general data set. Therefore a group of metrics was selected which is generally applicable and covers various types of stress. The metrics vary in intrinsic properties as to which features of the macroinvertebrate community they respond, i.e. structural (incl. sensitivity), functional or life cycle properties. The common metrics are calculated from species data using the various national methods and the STAR-AQEM method. This allows for a direct comparison of the performance of the national method compared with the STAR-AQEM method for each country individually.

- The 16 metrics were calculated from samples obtained using the various national methodologies and the STAR-AQEM method. Only main samples were used so that each site was represented by one sample per season. Performance of the national method and the STAR-AQEM method was tested using pair-wise comparisons for each country individually. For both seasons combined, no overall clear pattern emerged with respect to the differences between metric results obtained using STAR-AQEM and national methods. Some national methods performed better (i.e. scored significant higher values) than the STAR-AQEM method in some countries and vice versa in other countries. Within countries there was in most cases not a consistent pattern when comparing metrics: compared to the STAR-AQEM method, some metrics would score higher when calculated using data obtained by the national method while other would score lower.

- In most cases (64% of the countries) the various national methods yielded significantly higher EPT-taxa values than the STAR-AQEM method. A similar pattern was evident with respect to number of families in 73% of the countries significantly more families were found using the national method. In contrast, the STAR-AQEM method yielded significantly more EPT-taxa and families in 9% and 27% of the countries, respectively.

- The STAR-AQEM method was not superior to the majority of national methods. In contrast, the STAR-AQEM method appears to collect fewer taxa (all and EPT) and families than the majority of the national methods. The most likely explanation for this finding is that species are lost during the sub-sampling procedure employed by the STAR-AQEM method. Two methods, the Italian IBE method and the Latvian method, appear to loose information about the macroinvertebrate community to a degree that might affect the assessment of ecological stream quality. They yielded almost in all cases lower metric values than the STAR-AQEM method. Laboratory processing (IBE) and identification of more species (Latvian method) would properly improve their performance.

- When comparing the ecological classification of sites/samples using the two sampling methods (STAR-AQEM versus either a national method or the

RIVPACS method) gave quite consistent results. Between 31.6% and 82.4% of the samples were classified into the same class using the two macroinvertebrate methods. The results from the countries sampling using the RIVPACS method (Austria, Germany, Greece, and UK) varied between 47.5% and 65% with most countries consistently a little bit above 60%.The IBE method differed most compared to the STAR-AQEM method when comparing the ecological classification (only 31.6% were the same in the CNR dataset).

- Very few classifications differed more than one class in any direction using the two methods and six of the countries had all of their compared samples within one quality class in each direction (e.g. if the national/RIVPACS methods classifies a site as "moderate" then the STAR-AQEM method shouldn't assess the same site as worse than "poor" or better than "good"). Finally, according to the Water Framework Directive, the good-moderate boundary is the most important, since sites with a quality below this boundary has to be restored. Generally less than 20% of the sites were classified "across" this border using the two methods, where almost in all cases one method classified the site as having a "good" and the other a "moderate" status. Generally the PERLA and Polish sampling methods seemed to be most similar to the STAR-AQEM method. 14.6% of the sites in the Czech Republic, 9.1% in the Slovak Republic and only 3.9% of the sites in Poland were classified across the border. One reason for the Polish results was also that many of the Polish sites were classified as "high" by both sampling methods, and the ecological quality was thus far from the good-moderate class boundary where the misclassification rate is the highest.

- The taxonomic composition for all sites sampled using STAR-AQEM versus the national (or RIVPACS) sampling method was compared using Mantel tests (see above) (Table 4.17). For all comparisons, there were significant similarities between the STAR-AQEM versus the national or RIVPACS method used in each country. The STAR-AQEM and the RIVPACS method gave very similar results (method used in Austria, Germany, Greece, and the UK). The results were also very similar for the two Nordic methods (DSFI in Denmark and the Swedish standards method). The PERLA method on the other hand gave quite different results; it came out very similar to the STAR-AQEM method in the Czech Republic, but not in the Slovak Republic. The least similar results were obtained when comparing the Italian IBE method and the STAR-AQEM method and for the Slovak PERLA samples.

- When comparing how many taxa were indicative of the STAR-AQEM versus national/RIVPACS samples in terms of number of taxa captured, more taxa were over-represented in national/RIVPACS samples. In France five taxa were over-represented when sampling was taken using the IBGN method compared to the STAR-AQEM method and in Sweden four taxa were over-represented when samples was taken using the Swedish standard method as opposed to the STAR-AQEM method. Only one partner (the Italian CNR partner) had a high number of taxonomic groups over-represented when sampling using the STAR-AQEM method versus the national (IBE) method. All countries who used the RIVPACS sampling method had only one or two taxonomic groups over-represented using any method, whereas France (using IBGN) and Italy (CNR using the STAR-AQEM) method had the highest number of over-represented taxonomic groups.

- Six countries sampled in at least two stream types and the amount of variation in macroinvertebrate community composition explained by type differed between 16.0% in the Czech Republic and 67.9% of the total explained variation in Greece. Two different seasons were sampled in all countries and season explained between 11.6% of the total explained variation (in Greece) and 56.0% of the total explained variation in Latvia. The pre-defined stress gradient (here divided into sites pre-defined as having a high or good ecological status versus those pre-defined as having a moderate, poor or bad ecological status) explained between 15.3% (in Greece) and 55.3% of the total explained variation in France. Stream type, differences between seasons, and the pre-defined stress gradient were always statistically significant explanatory variables. The difference in sampling method on the other hand, generally only explained a smaller part of the total explained variation (except in a few cases). Sampling methods was a statistically significant explanatory variable and explained a relatively high amount of the total explained variation in Poland, Latvia, and Italy (CNR).

- Within the STAR field sampling programme, a second 'replicate' sample was taken at between two and six sites in most stream types, usually in both sampling seasons, using both the STAR-AQEM and the 'National' or RIVPACS methods. These replicate sampling sites were chosen to cover a range of perceived qualities. This enabled the estimation of the variance and standard deviation (SD) in values of individual metrics due to replicate field sampling variation In addition, at most of these replicate sampling sites, a second 'replicate' sub-sample was taken from one of the two STAR-AQEM samples. This enabled the estimation and separation of sub-sampling effects from field sampling effects for STAR-AQEM samples.

- Sampling variances for a metric often showed a tendency to increase with the site mean value for the metric. For metrics expressed as proportions or percentages, replicate sample values were often more variable for sites with intermediate mean values (around 0.5 or 50%). To overcome this and derive a more representative single estimate of sampling variance for all sites within a stream type for a particular metric and method, the metric values were often transformed (either square roots or double square roots, or arcsine transformed for proportions or percentages (see Table 5.2)). Variance components and standard deviations were estimated and compared on these transformed scales.

- The practical size and importance of overall replicate variance was estimated by expressing the variance as a percentage (P_{samp}) of the total variace in metric values amongst all sites within a stream type. A low percentage indicates that the combination of sampling method and metric has high statistical precision compared to variability amongst sites of differing quality. High percentages indicate low sampling precision and low repeatability and hence that such a combination of sampling method and metric is unlikely to have much power to detect differences in ecological status class. Obviously, metrics with high sampling precision and repeatability may still not be good ecological metrics or relibaility indicators of ecological status class.

- STAR-AQEM sub-sampling variation causes a relatively large part of the overall variance between replicate sample values for many metrics, and is estimated on average (see Table 5.5) to contribute more than 50% of the overall variance between replicate samples for 12 of the 27 metrics analysed.

In general, sub-sampling variance is relatively large for those metrics which are based on the numbers of taxa present, such as number of families and number of EPT taxa. Sorting and identifying a larger fraction of the sample would reduce this source of variation; in the extreme, sorting the whole sample would eliminate it. However, all extra identification increases costs, which were not assessed in this study, but see Vlek *et al*. (2004). Although sub-sampling contributes a major part of the overall inter-replicate variance in numerous metrics, overall inter-replicate variance may still be small compared to the range in metric values amongst sites of varying quality and thus such metrics may still have high precision to detect differences between sites.

- The original Saprobic index, the German new Saprobic index and the Czech Saprobic index appear to have the lowest percentage sampling variance with median values of only 3%, 5% and 6% respectively. The proposed ICM metrics of Number of EPT taxa, ASPT, Shannon-Wiener diversity and (1-GOLD) also have highly variable estimates of P_{samp} across stream types, but with similar intermediate size median values of 15.5%, 17%, 14% and 16.5% respectively. Most of these 27 metrics analysed have average replicate sampling variances of 10-20% of the total variance in metrics values within a stream type (Table 5.8). This suggests that the precision of such metrics based on the STAR-AQEM method is sufficient to indicate gross changes in the ecological status of sites, but there will be considerable uncertainty in the assignment of sites to particular status classes. The estimates of SD_E derived here can be used in the software program STARBUGS (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sampling variation on the uncertainty in assignment of sites to ecological status classes based on the STAR-AQEM sampling method.

- Replicate RIVPACS samples were taken from sites in each of the sampled stream types in Austria, Germany and the UK, but in Greece, replicate sampling was confined to six sites in one stream type. Using the RIVPACS method to sample the streams in Austria gave high precision to estimates of most metrics with nearly two-thirds (62%) of metrics having percentage sampling variances (P_{samp}) of less than 10% with an average of only 9 % and a maximum of only 24% - this suggests high sampling repeatability of all aspects of the macroinvertebrate community structure. For these Austrian stream types, the RIVPACS and STAR-AQEM methods were, on average, about equally precise.

- Within the three stream types sampled in Germany, the average percentage sampling variance (P_{samp}) for the RIVPACS method was 17%, but ranged from only 2-4% for the three saprobic indices up to 55% for the metric '%Oligochaeta'. The estimated values of P_{samp} for ASPT for the sampled German sites and stream types were relatively high for both the RIVPACS and STAR-AQEM methods at 27% and 23% respectively. Overall, the RIVPACS and STAR-AQEM methods gave similar sampling SD and percentage sampling variances, although the STAR-AQEM was estimated to be more precise for slightly more metrics. For the sampled Greek stream type (H04), the RIVPACS method gave higher percetnage sampling variances (P_{samp}) with an average value of 27%, ranging from 4% for 'Number of Families' to 75% for '%Oligochaeta'. Overall in Greece, the RIVPACS method was slightly less precise than the STAR-AQEM method.

- Within the two stream types sampled in the UK, the RIVPACS method led to lower sampling SD and lower percentage sampling variances for the majority of metrics compared to the STAR-AQEM method. The average value of P_{samp} for the RIVPACS and STAR-AQEM methods in the UK were 11% (range 2-36%) and 20 (range 4-65%) respectively. For the RIVPACS method, sampling variance was less than 10% of total variance within a stream type for nearly two-thirds of the 27 metrics analysed, compared to only one-third for the STAR-AQEM method.

- Vlek (2004) found that, on average across the sampled sites, STAR-AQEM samples took 18 hours to process (including sorting and identification), whilst RIVPACS samples took only 9 hours – half the amount of time. As the RIVPACS method led to no more than marginally higher average percentage sampling variances within the four countries where both methods were used, the RIVPACS method may be more cost-effective than the STAR-AQEM method, at least when the aim is to base site assessments on one or more of the metrics assessed here.

- Most of the other partners involved in the STAR sampling programme took replicate samples using their 'National' method at the same sites as replicate STAR-AQEM samples were taken. The two methods were compared in terms of estimated sampling SD and especially the percentage sample variance (P_{samp}) for each metric.

- Within the two stream types sampling in the Czech Republic, sampling variance was low compared to total variance in metric values within a stream type for both the 'National' PERLA mathod and the STAR-AQEM method, with average values of of P_{samp} of only 5% and 8% respectively. Moreover, P_{samp} was no more than 10% for all of the metrics analysed. The 'PERLA' method gave lower sampling SD and lower P_{samp} than the STAR-AQEM method for the majority of metrics, indicating its higher repeatability and precision – although relative costs, unassessed, may also vary.

- Within the stream type sampled in France, the 'National' (BGN) method gave lower sampling SD and lower P_{samp} for the nearly all of the metrics analysed, with an average value of P_{samp} of 11% compared to 22% for the STAR-AQEM method.

- As used in Italy in STAR stream type I06, both the 'National' (IBE) method and the STAR-AQEM method gave amongst the highest percentage sampling variances of any stream type and method, although the analyses were based on only those metrics considered to be valid for community data based on identification to family level. For the IBE method, the average P_{samp} was 35% with a range from 12% for the metric 'Number of EPT taxa' to 73% for the total 'Abundance' metric. The STAR-AQEM method appeared to be equally prone to sampling variation with an average P_{samp} of 38% (range 10-77%). Part of this explanation may be that the sites sampled in Italy may have covered a narrower range of ecological qualities so that total sampled variability within the stream type is relatively low. In Italy, samples are sorted in the field which may cause extra variability between replicate sample derived metric values.

- The Latvian 'National' method was the most prone to sampling variability and led to higher values of P_{samp} than the STAR-AQEM method for almost all metrics based on the same set of sampled sites.

- Both the Danish (DSFI) and Polish 'National' methods lower sampling variances than the STAR-AQEM method for the majority of metrics measured within for the same set of sampled sites in their respective countries. In Denmark, the average value of P_{samp} was 13% based on the DSFI method compared to 22% for the STAR-AQEM method for the same sites. Half of the metrics had P_{samp} values less than 10% for the DSFI – suggesting the method provides high sampling precision and repeatability of results.
- In Portugal, the 'National' method and STAR-AQEM method gave similar intermediate levels of sampling precision with average values of P_{samp} of 22% and 23% respectively. In Sweden, the 'National' method led to higher percentage sampling variances than the STAR-AQEM method for the majority of metrics with average values of P_{samp} of 30% and 16% respectively.

- The STAR replicate sampling study has provided the first comparative estimates of the susceptibility to sampling variability of a range of macroinvertebrate sampling methods and metrics.

- The estimates of replicate sampling standard deviation (SD_E) reported here can be used in the STARBUGS simulation software package (STAR Bioassessment Uncertainty Guidance Software, Clarke 2004) to assess the effect of sampling variability in individual metric values on the uncertainty of multi-metric assessments of the ecological status of sites. STARBUGS was produced as Deliverable 9 within the STAR project and can be downloaded from the STAR web-site (www.eu-star.au).

7 REFERENCES

Brown, A. V. & P. P. Brussock, 2001. Comparisons of benthic invertebrates between riffles and pools. Hydrobiologia 220:99-108.

Clarke, R.T. (2000). Uncertainty in estimates of river quality based on RIVPACS. In: Assessing the biological quality of freshwaters: RIVPACS and similar techniques. Wright, J.F., D.W. Sutcliffe and Furse, M..T. (eds). *pp 39-54. Freshwater Biological Association*, Ambleside.

Clarke, R.T., Furse, M.T., Gunn, R.J.M., Winder, J.M. & Wright, J.F. (2002). Sampling variation in macroinvertebrate data and implications for river quality indices. *Freshwater Biology*, **47**, 1735-1751.

Clarke, R.T. (2004). 9th STAR deliverable. Error/uncertainty module software STARBUGS (STAR Bio Assessment Uncertainty Guidance Software) User Manual.

Douglas, M. E. & J. A. Endler. 1982. Journal of Theoretical Ecology 99:777-795.

Dufrêne, M. and P. Legendre, 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs, 67 : 345-366.

Illies, J., (ed.), 1978: Limnofauna Europaea. Gustav Fischer Verlag, Stuttgart.

Mantel, N. 1967. Cancer Research 27:209-220.

McCune, B. & Mefford, M.J. (1999): PC-ORD. Multivariate Analysis of Ecological Data, Version 4.33. MjM Software Design, Gleneden Beach, Oregon, USA. 237 pp.

Sokal, R.R. & Rohlf, J.R. (1995). Biometry, 3rd edition. Freeman and Company, New York.

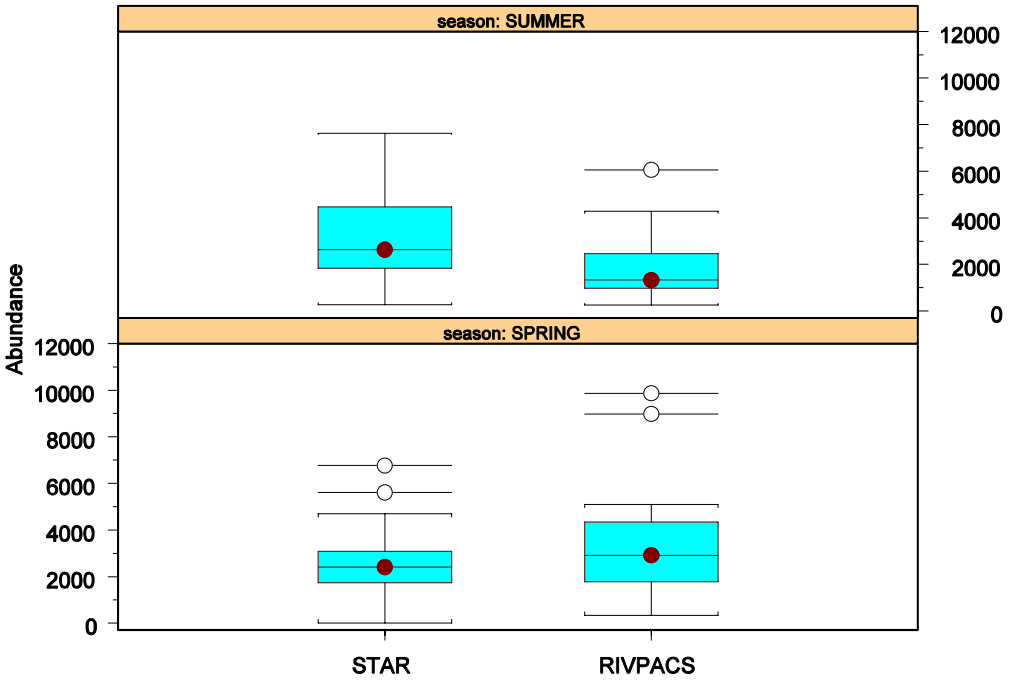
ter Braak, C. J. F. 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. – Vegetatio 69: 69–77.

ter Braak, C. J. F. and Smilauer, P. 1998. CANOCO reference manual and user's guide to CANOCO for Windows: software for canonical community ordination (ver. 4). – Microcomputer Power, Ithaca, NY.

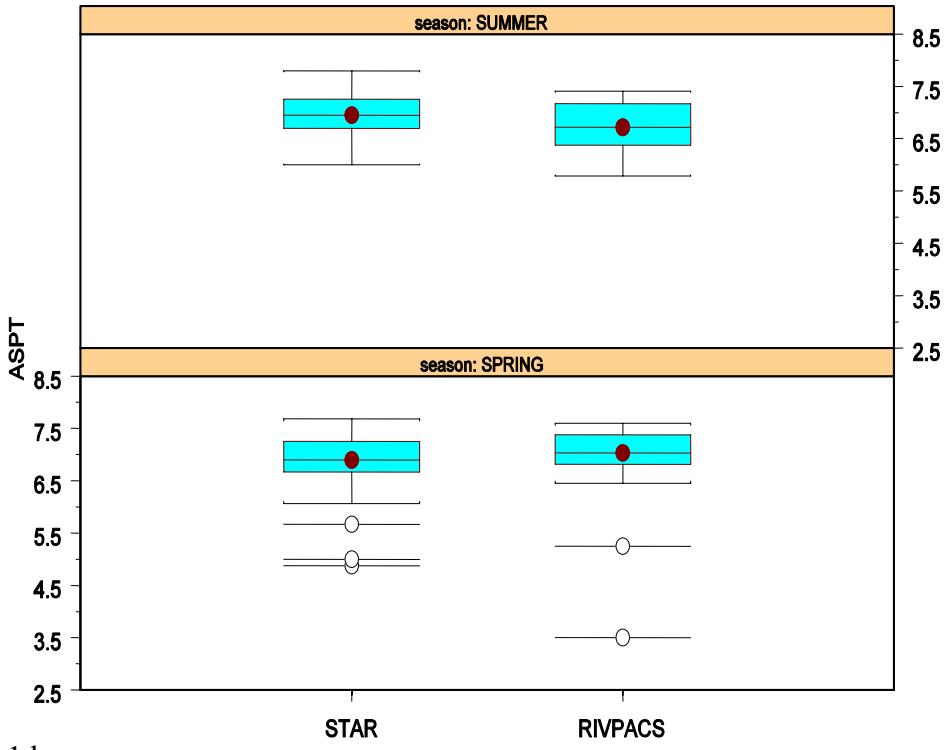
Vlek, H. (2004) STAR Deliverable N1: Comparison of (cost) effectiveness between various macroinvertebrate field and laboratory protocols.

Appendix 1. Comparison of seasonal variability between the AQEM-STAR method and the national methods

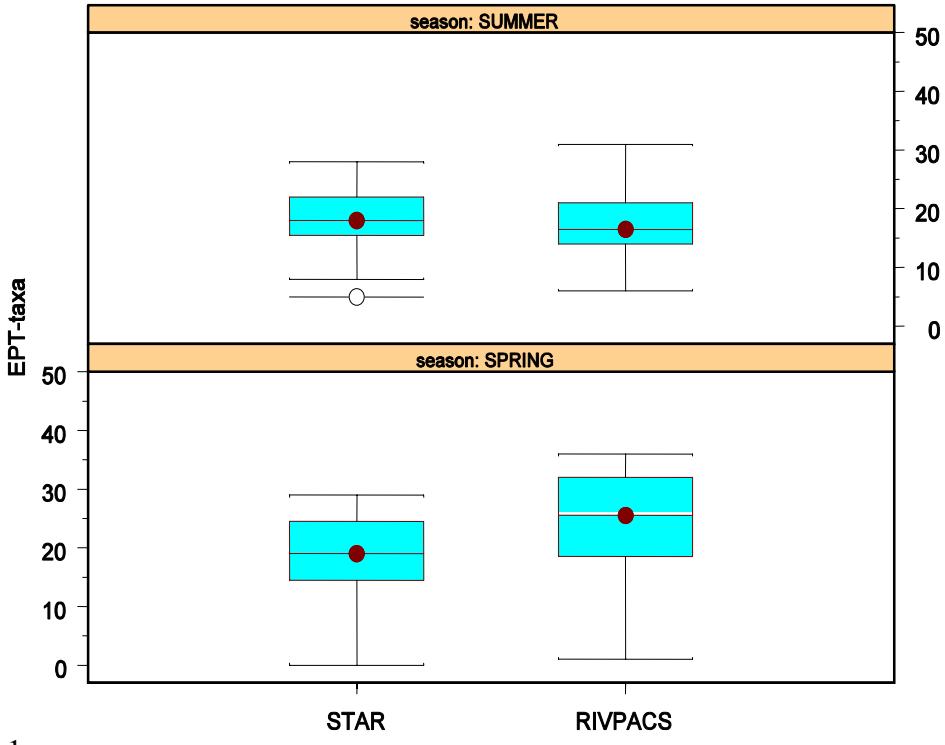
1 - Austria



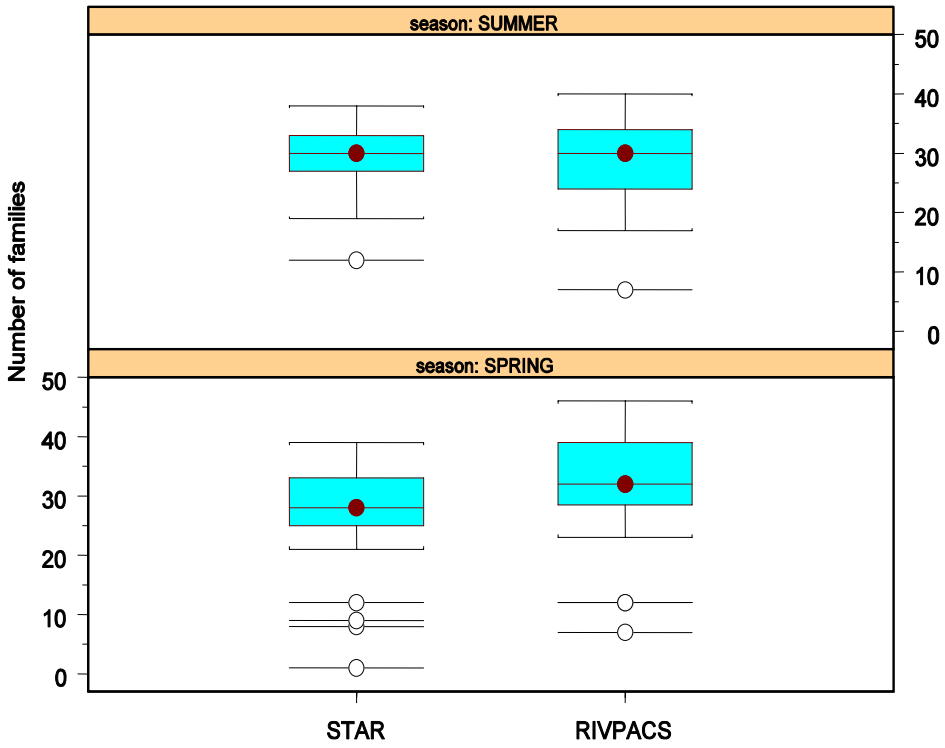
1 a


season: SUMMER
season: SPRING
ASPT
8.5
7.5
6.5
5.5
4.5
3.5
2.5
STAR
RIVPACS

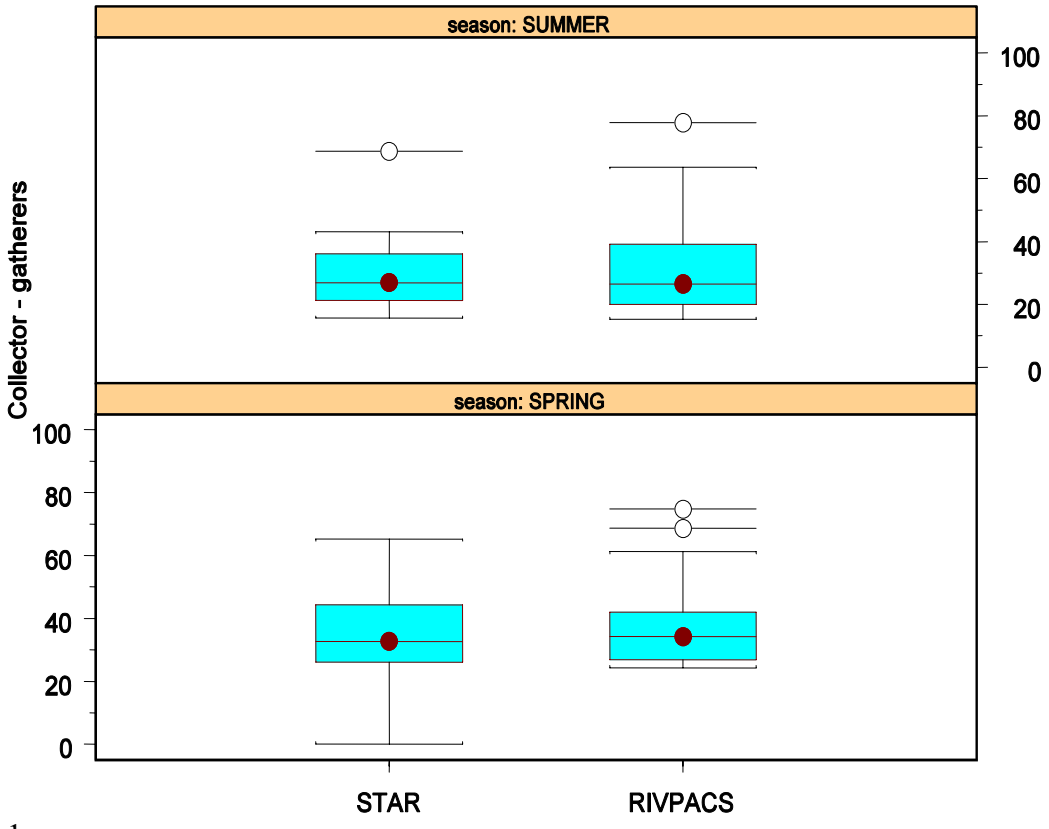
1 b


season: SUMMER
season: SPRING
EPT-taxa
50
40
30
20
10
0
STAR
RIVPACS

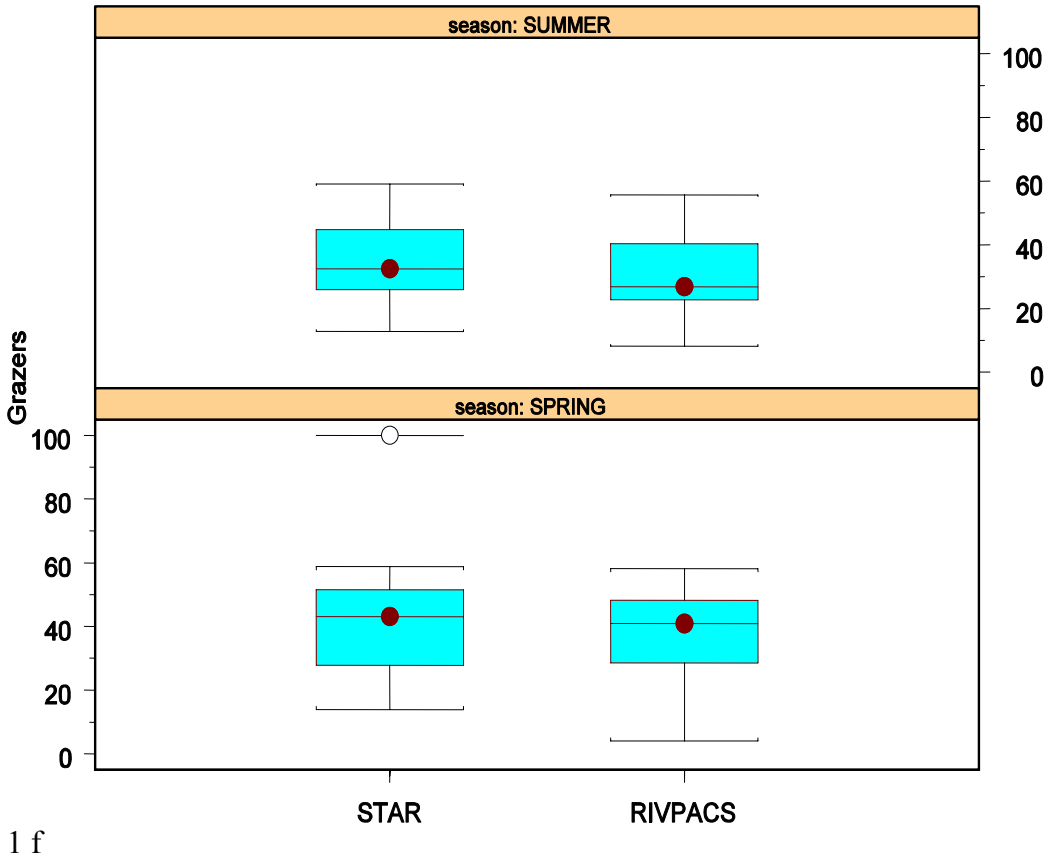
1 c


season: SUMMER
season: SPRING
Number of families
0
10
20
30
40
50
STAR
RIVPACS

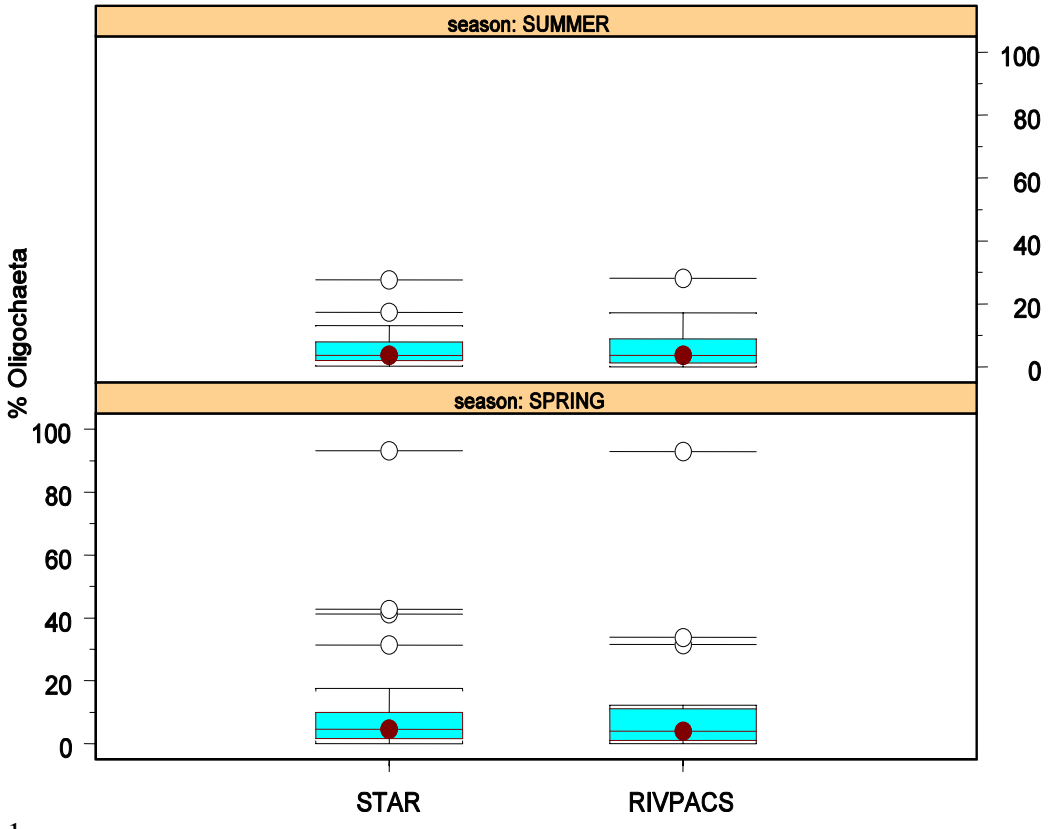
1 d


season: SUMMER
season: SPRING
Collector - gatherers
0
20
40
60
80
100
STAR
RIVPACS

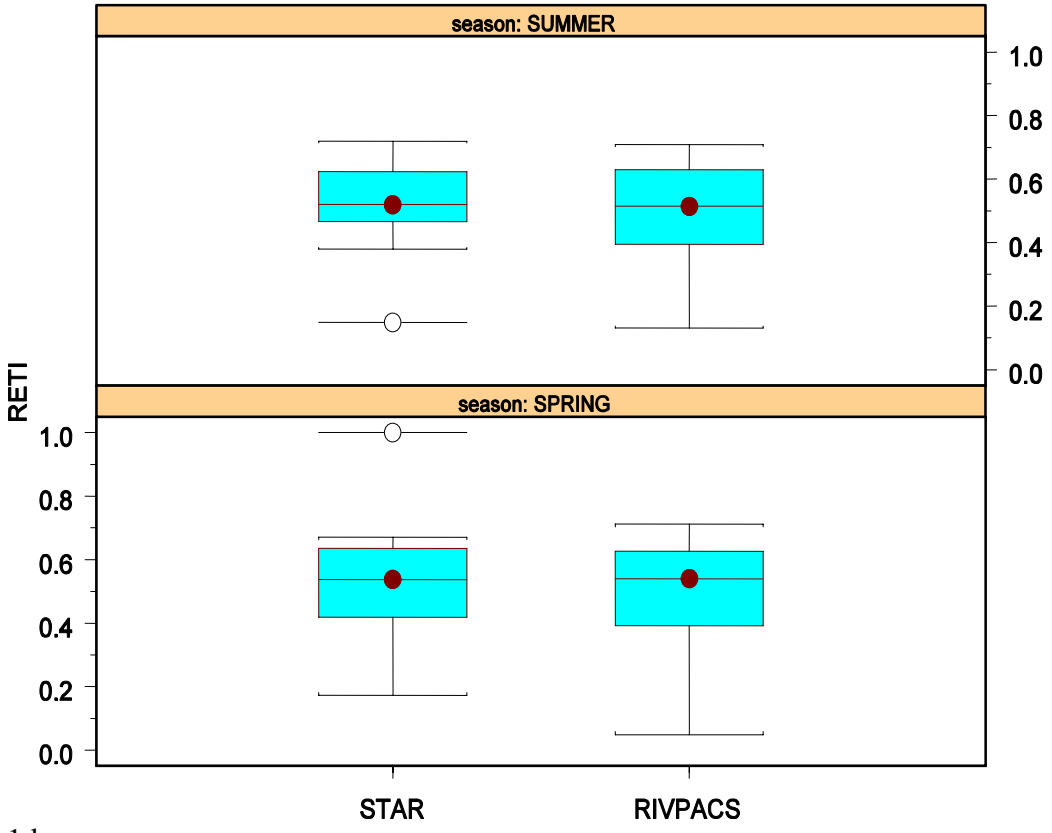
1 e


season: SUMMER
season: SPRING
Grazers
100
80
60
40
20
0
STAR
RIVPACS

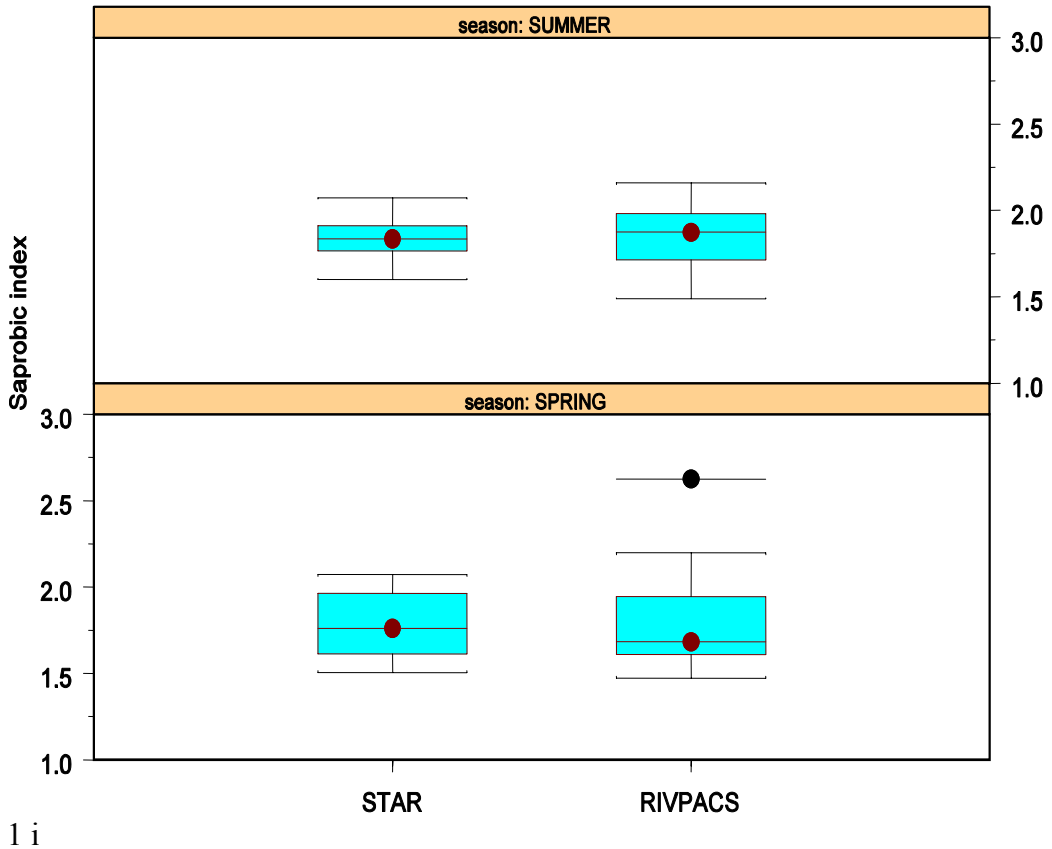
1 f


season: SUMMER
season: SPRING
% Oligochaeta
100
80
60
40
20
0
STAR
RIVPACS

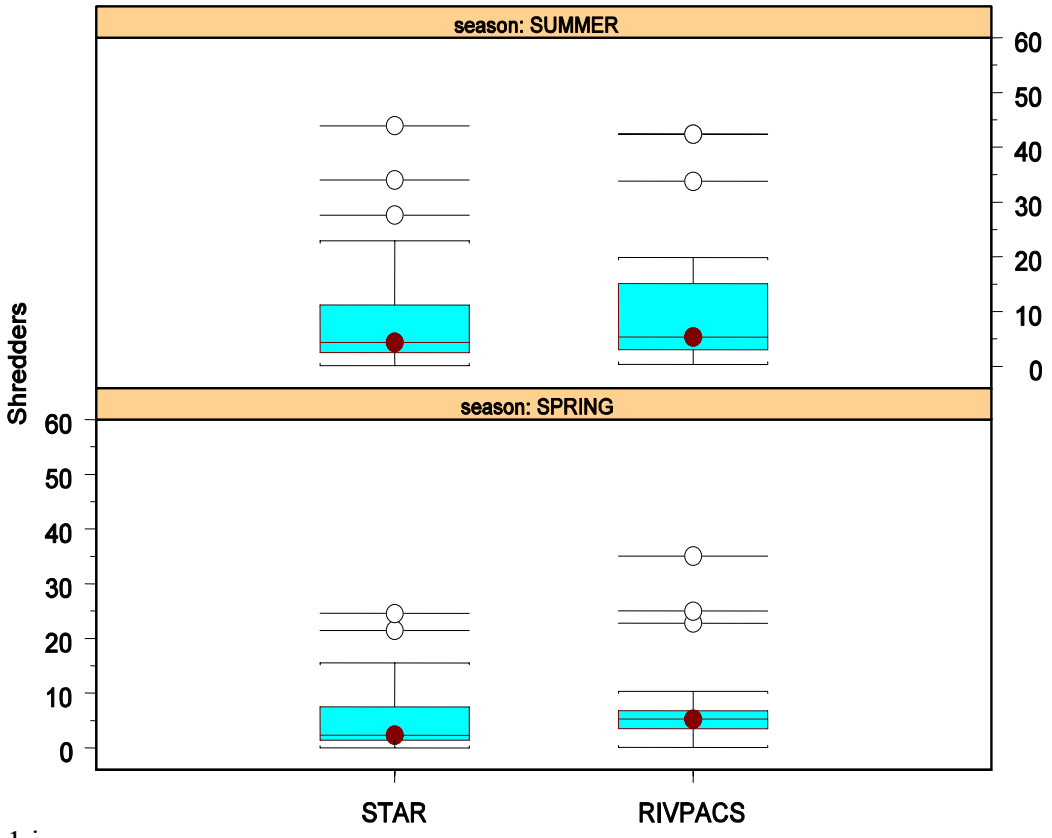
1 g


season: SUMMER
season: SPRING
RETI
1.0
0.8
0.6
0.4
0.2
0.0
STAR
RIVPACS

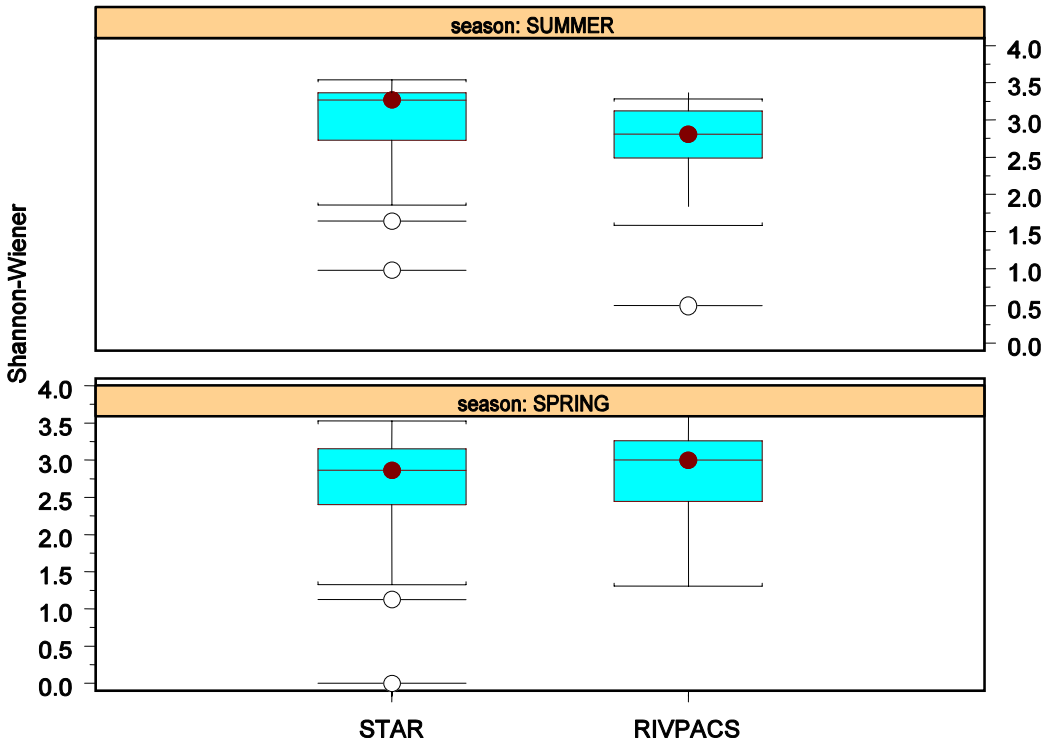
1 h


season: SUMMER
season: SPRING
Saprobic index
3.0
2.5
2.0
1.5
1.0
STAR
RIVPACS

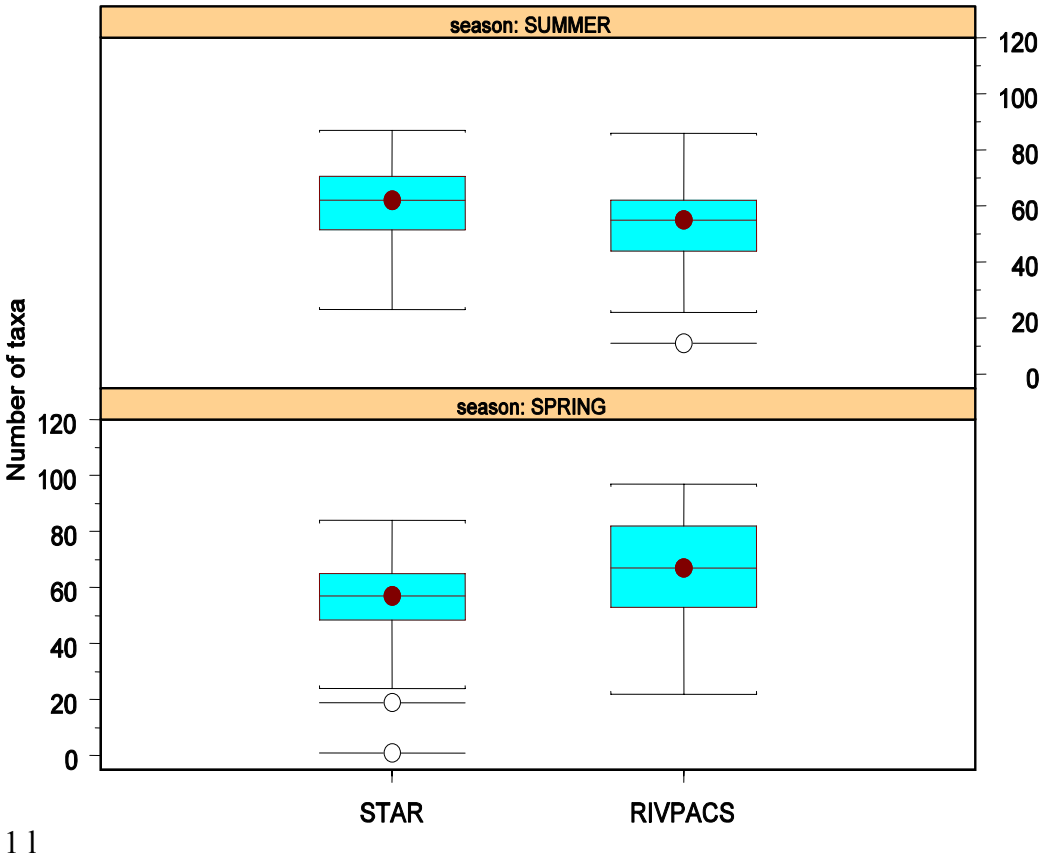
1 i


season: SUMMER
season: SPRING
Shredders
60
50
40
30
20
10
0
STAR
RIVPACS

1 j

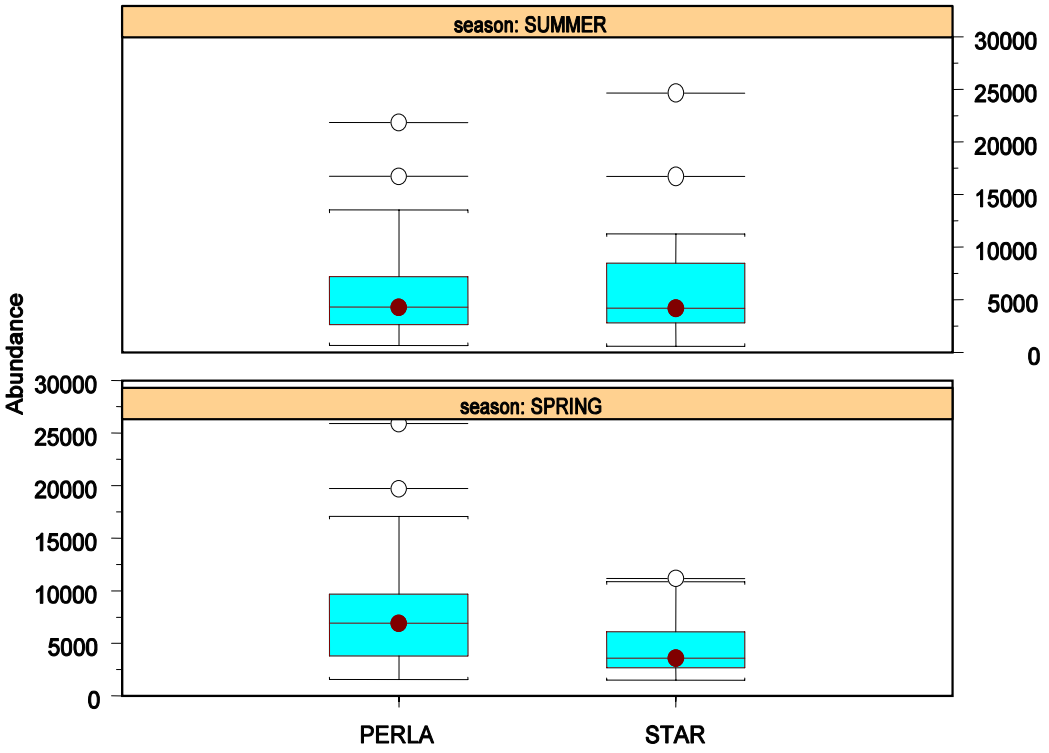

season: SUMMER
season: SPRING
Shannon-Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
STAR
RIVPACS

1 k

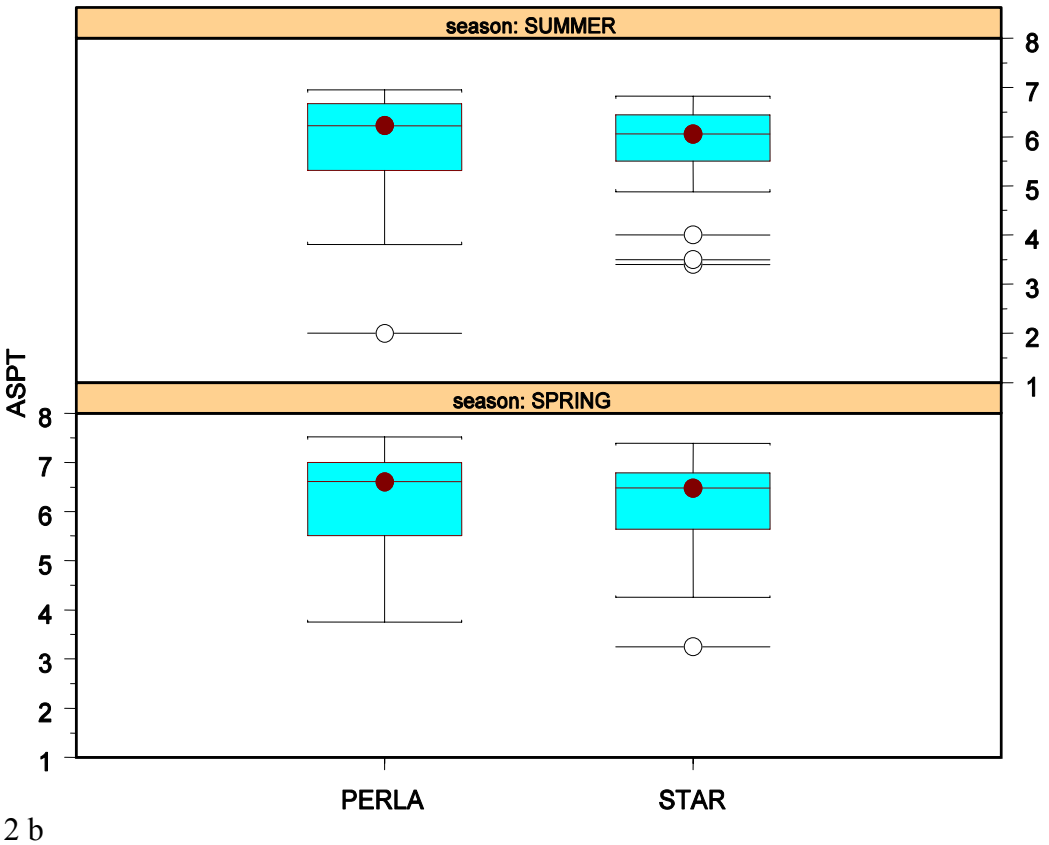

season: SUMMER
season: SPRING
Number of taxa
120
100
80
60
40
20
0
STAR
RIVPACS

1 1

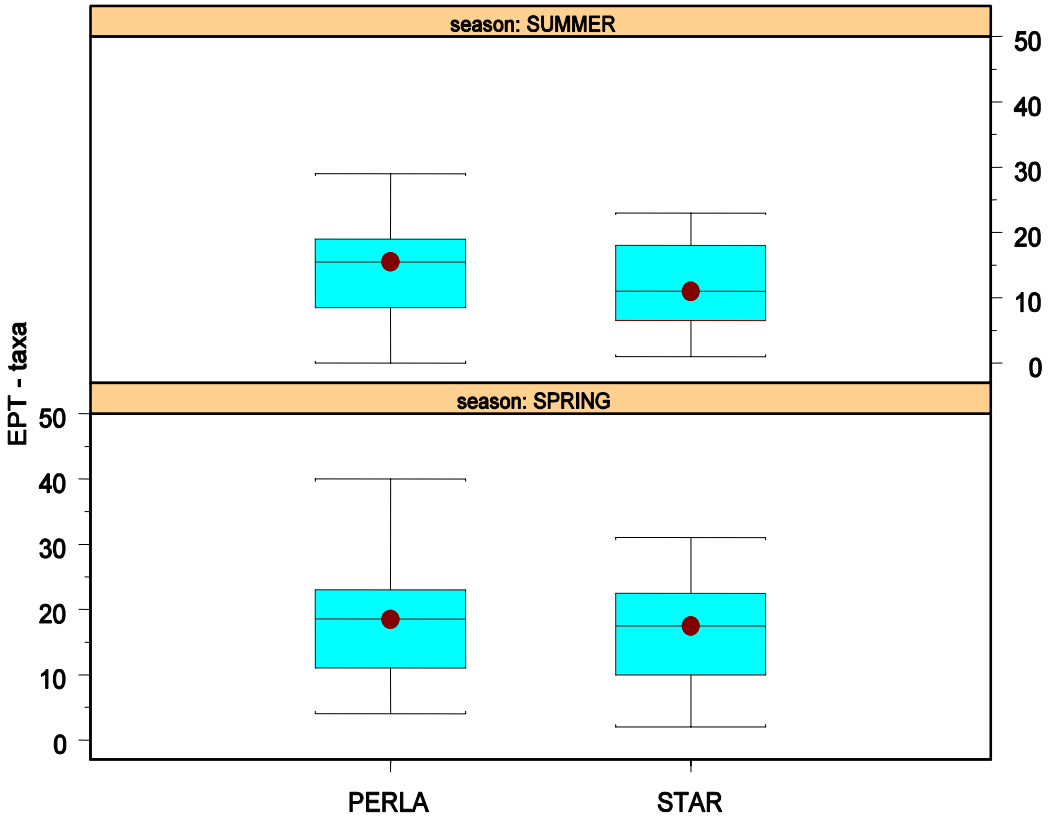
2 – Czech Republic



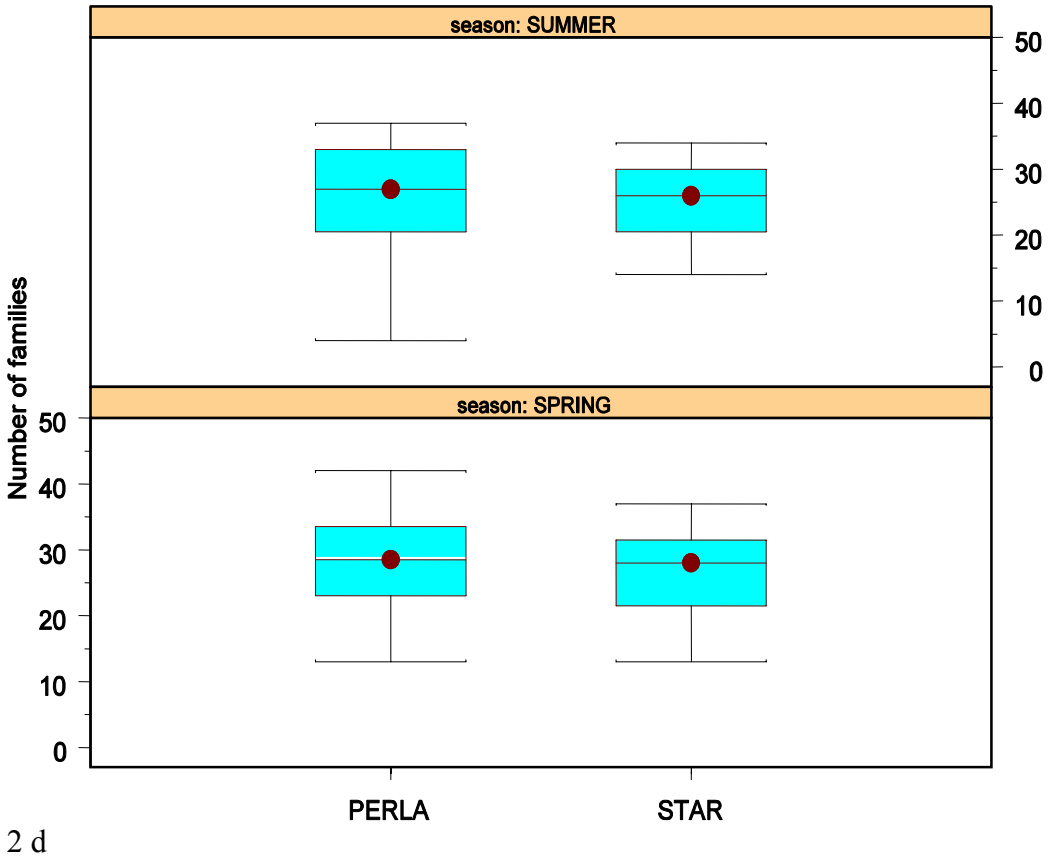
2 a



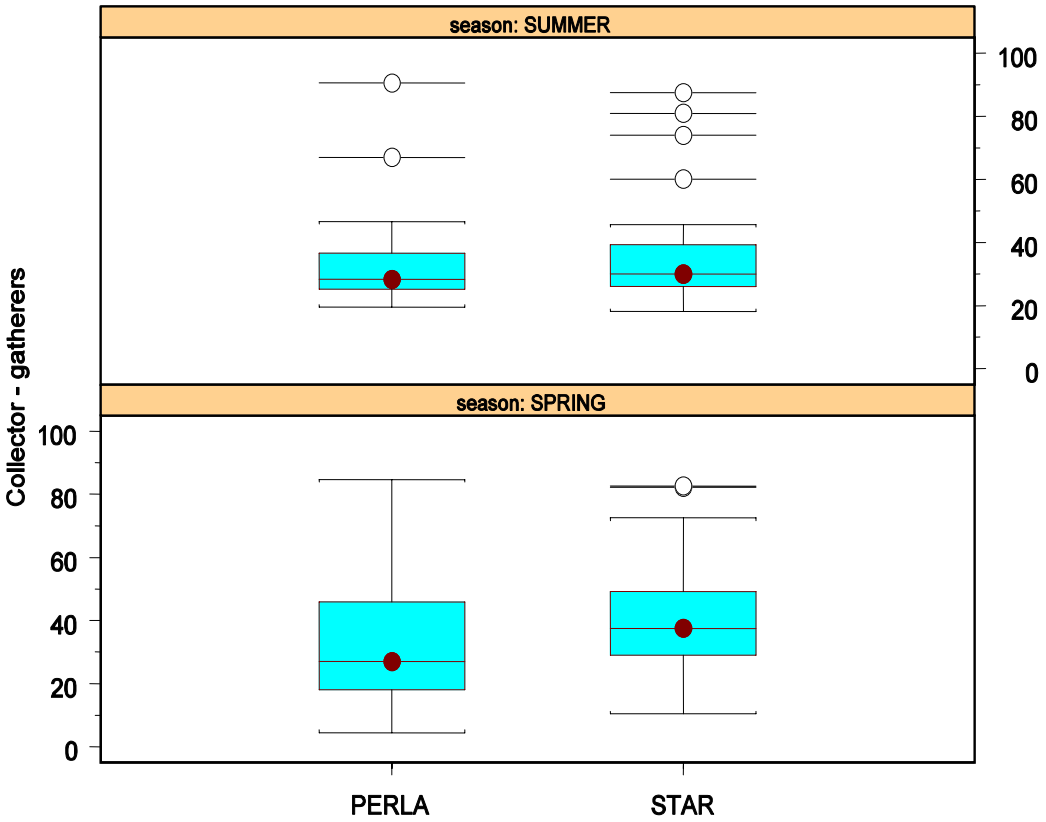
2 b


season: SUMMER
season: SPRING
EPT - taxa
0
10
20
30
40
50
PERLA
STAR

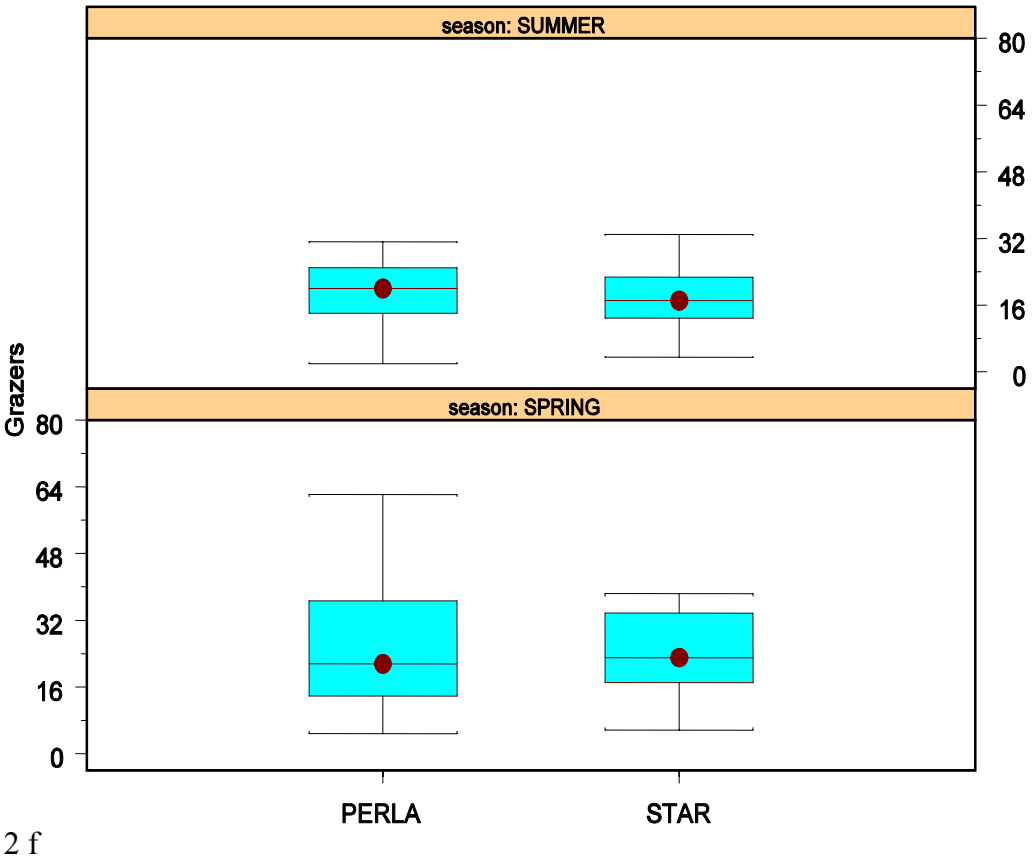
2 c


season: SUMMER
season: SPRING
Number of families
0
10
20
30
40
50
PERLA
STAR

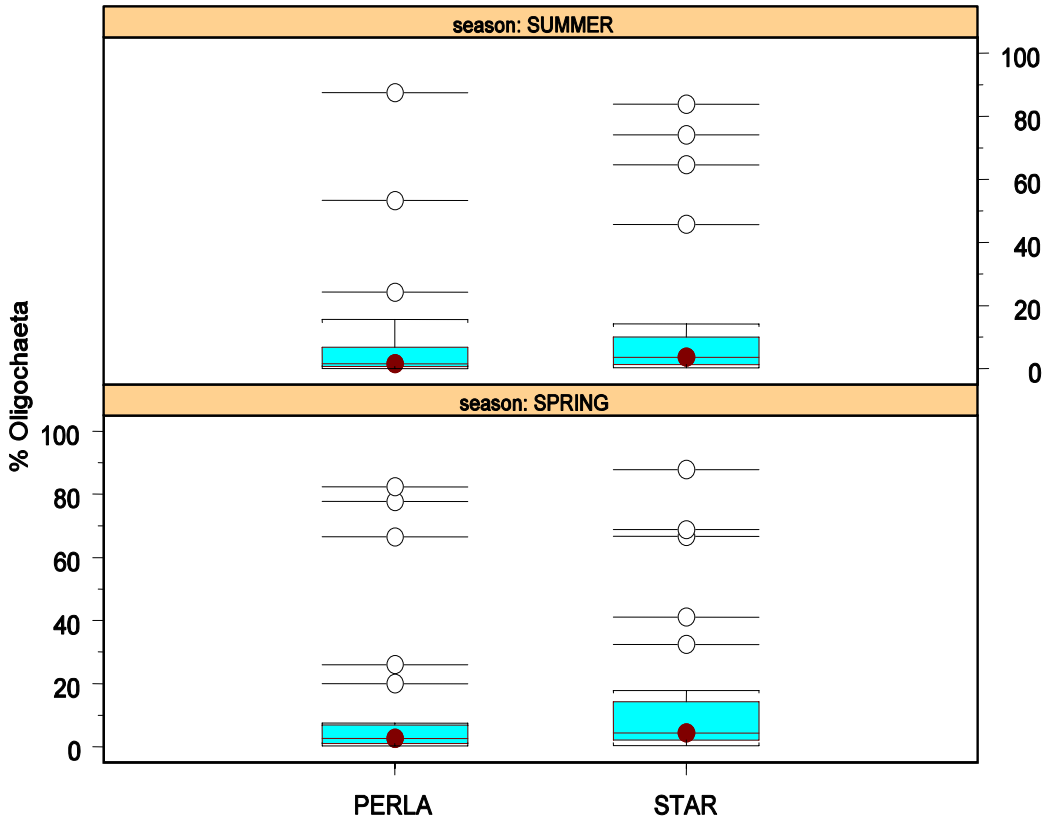
2 d


season: SUMMER
season: SPRING
Collector - gatherers
0
20
40
60
80
100
PERLA
STAR

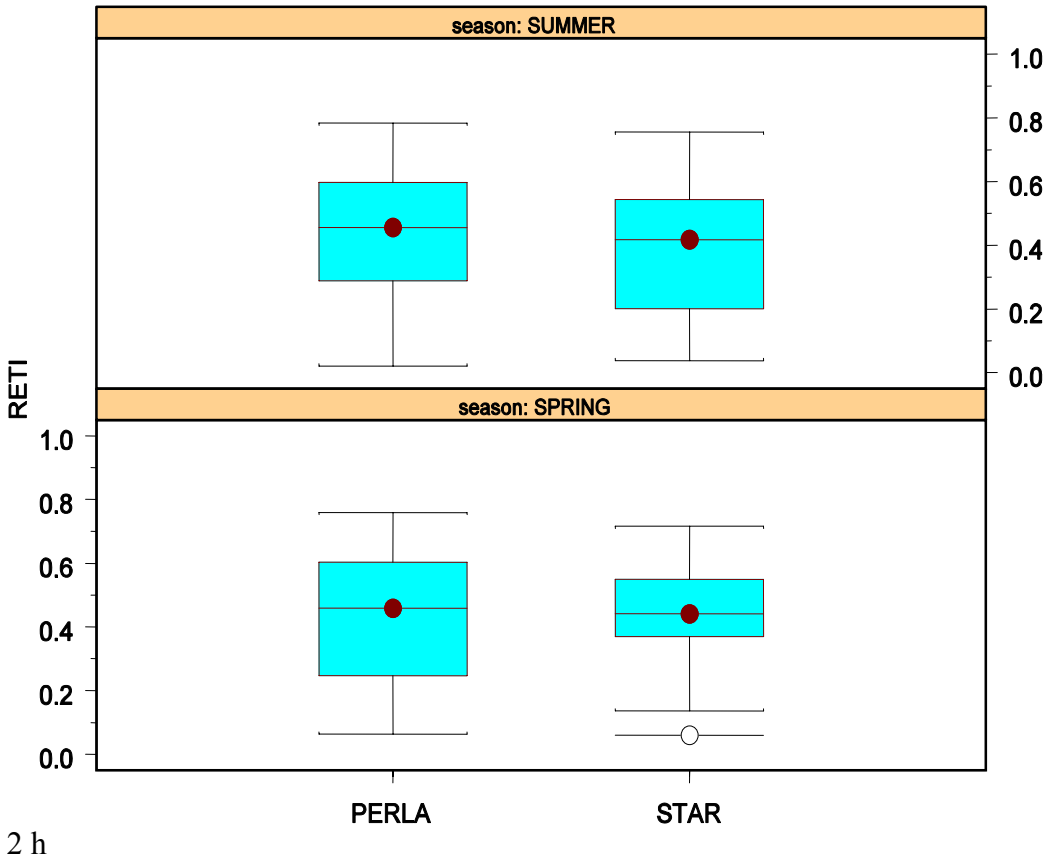
2 e


season: SUMMER
season: SPRING
Grazers
0
16
32
48
64
80
PERLA
STAR

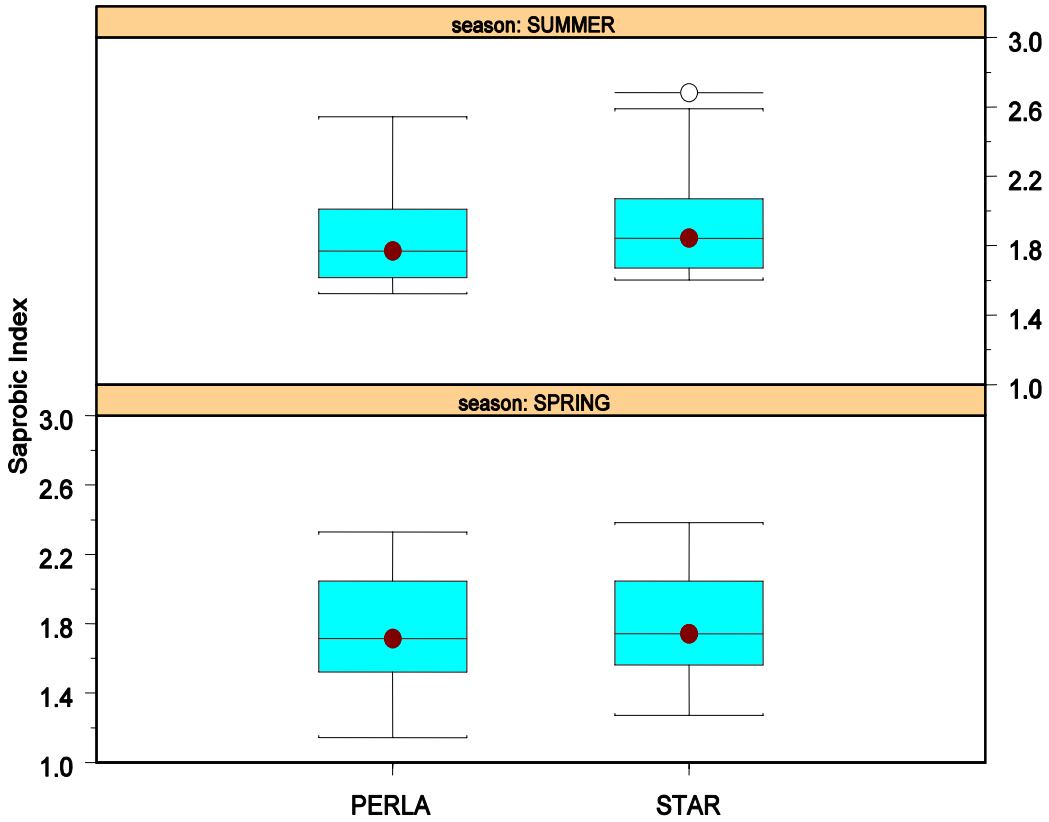
2 f


season: SUMMER
season: SPRING
% Oligochaeta
0
20
40
60
80
100
PERLA
STAR

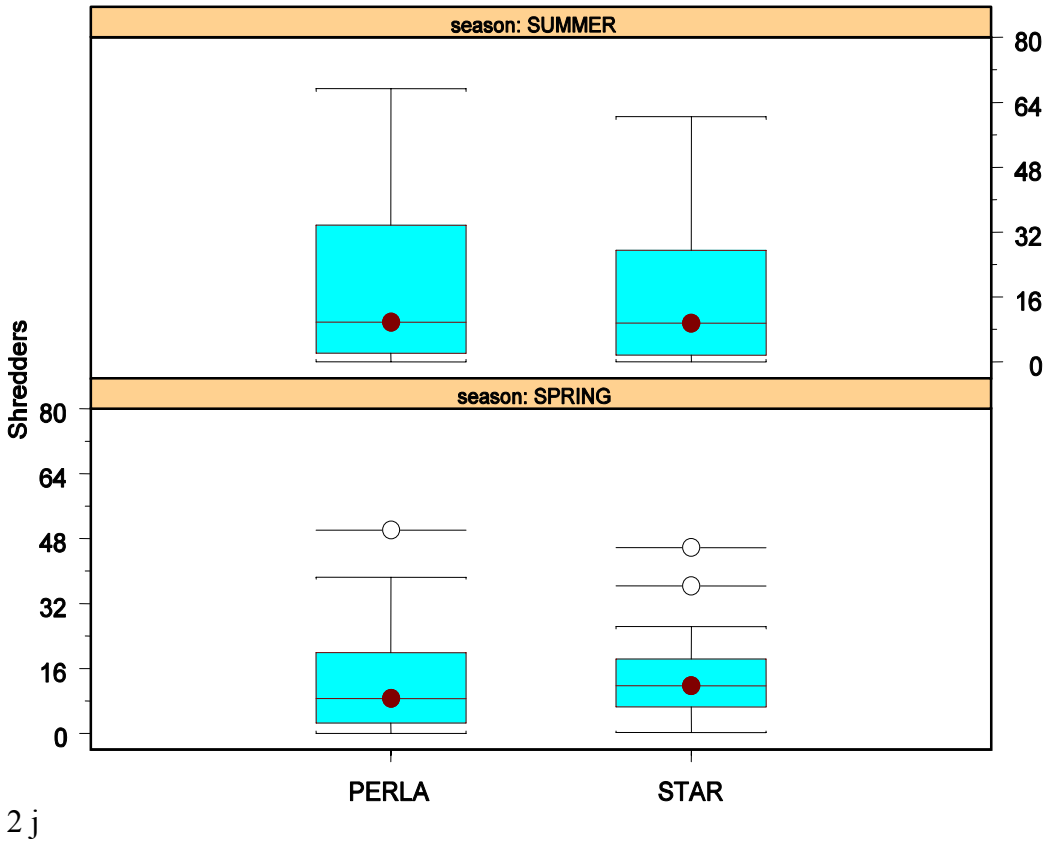
2 g


season: SUMMER
season: SPRING
RETI
0.0
0.2
0.4
0.6
0.8
1.0
PERLA
STAR

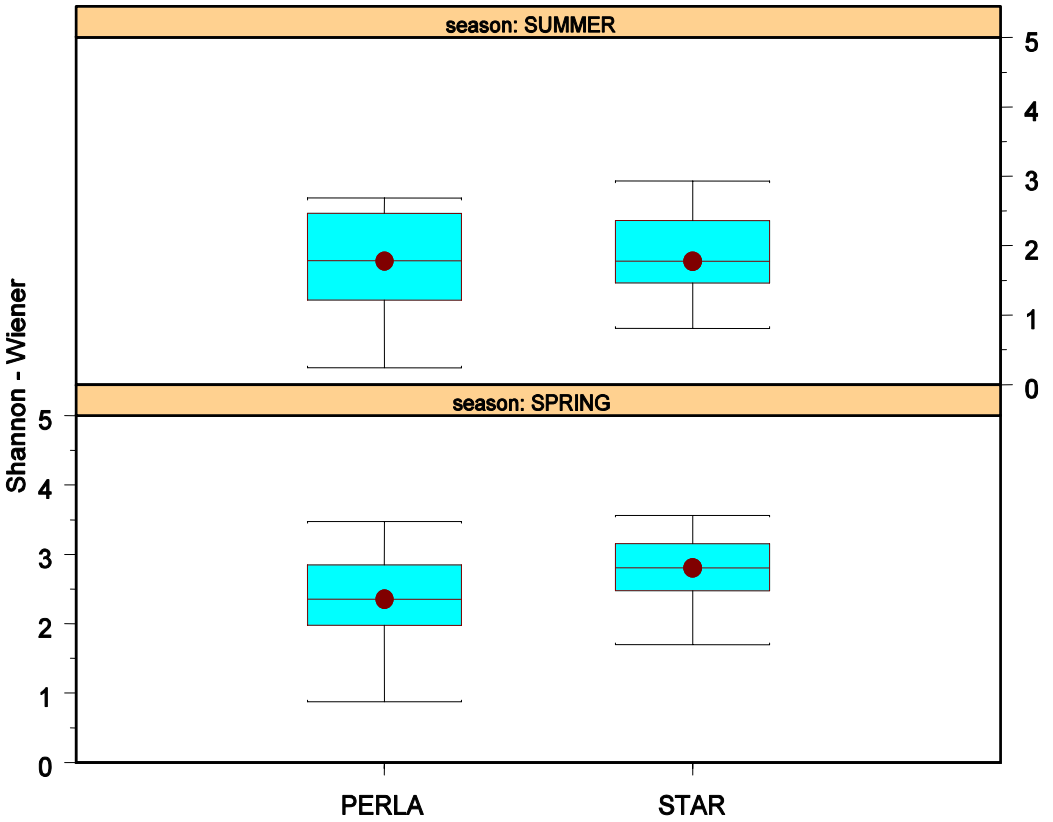
2 h


season: SUMMER
season: SPRING
Saprobic Index
3.0
2.6
2.2
1.8
1.4
1.0
PERLA
STAR

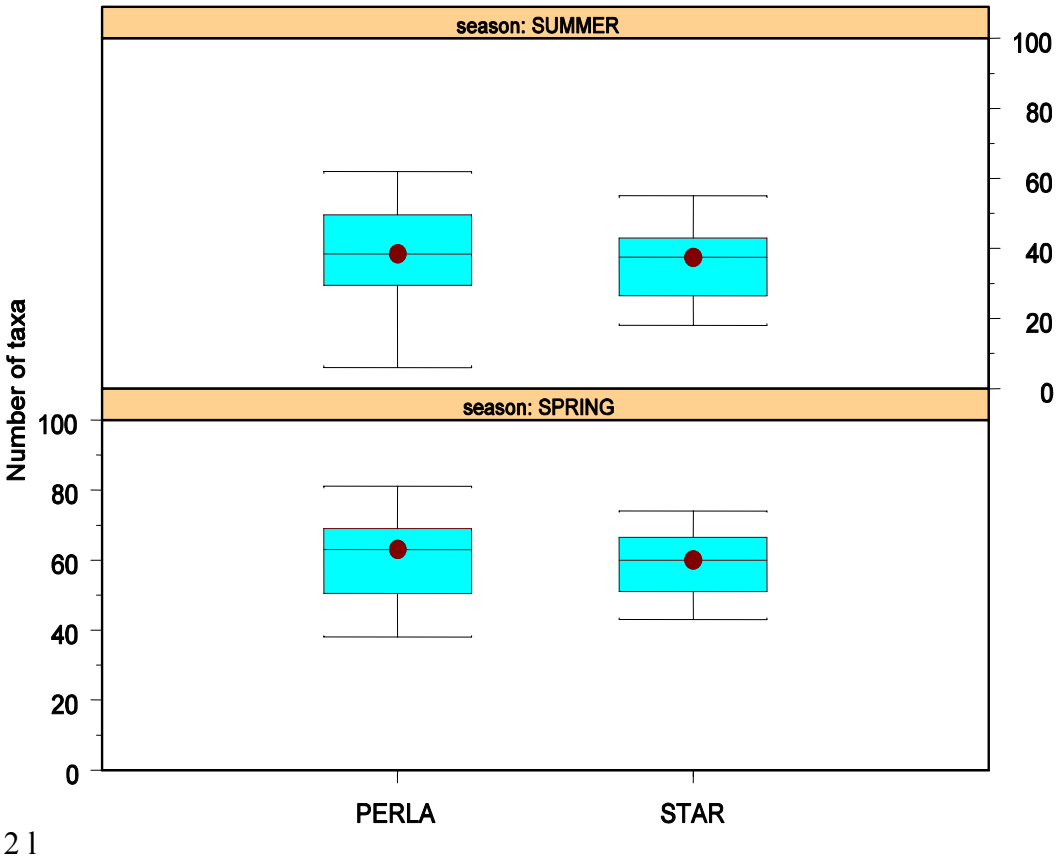
2 i


season: SUMMER
season: SPRING
Shredders
80
64
48
32
16
0
PERLA
STAR

2 j

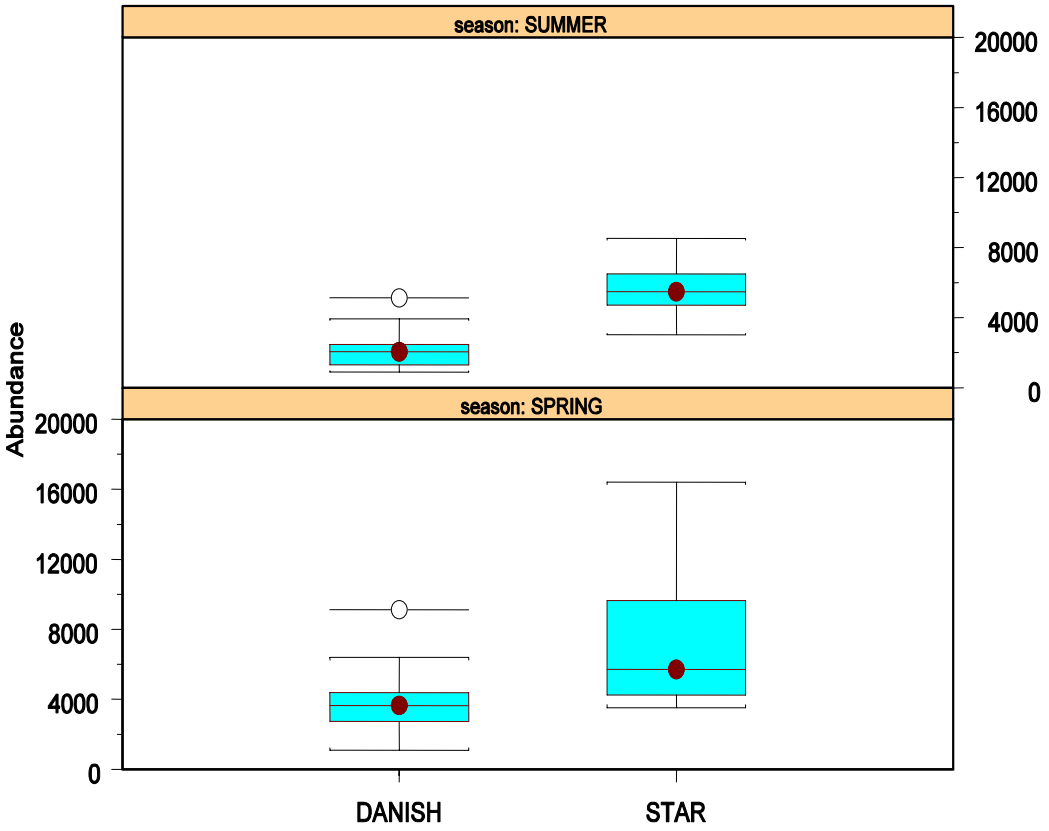

season: SUMMER
season: SPRING
Shannon - Wiener
0
1
2
3
4
5
PERLA
STAR

2 k

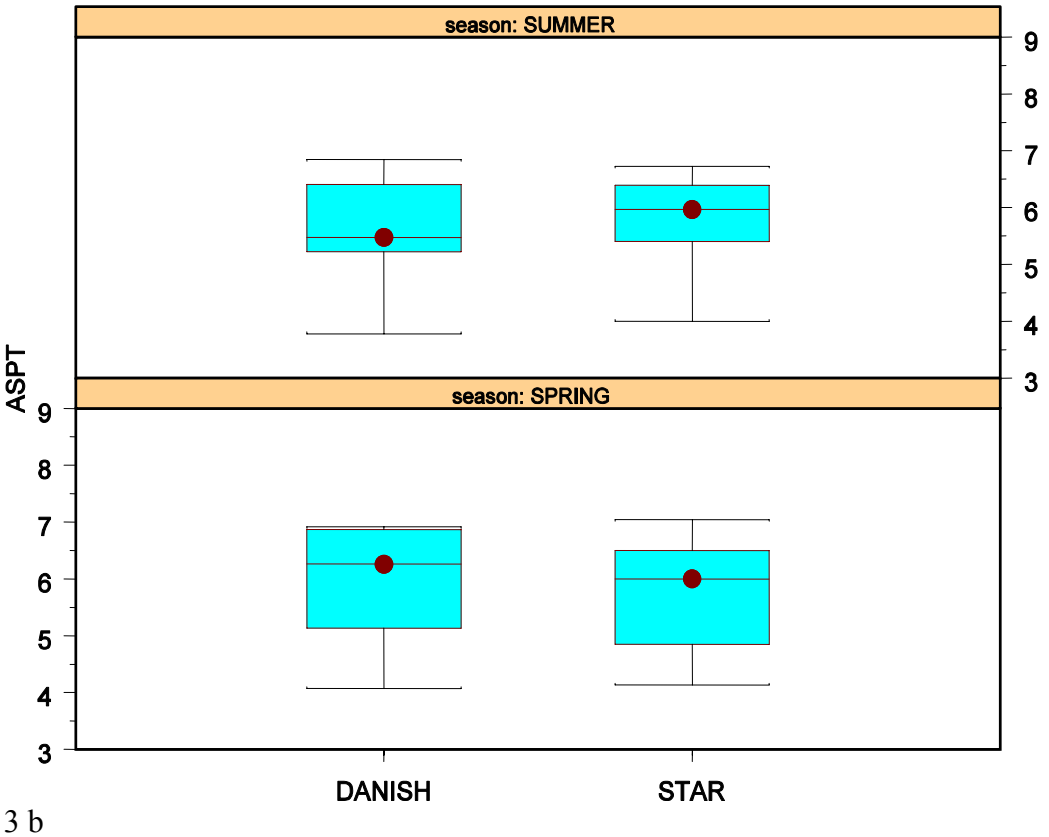

season: SUMMER
season: SPRING
Number of taxa
0
20
40
60
80
100
PERLA
STAR

2 l

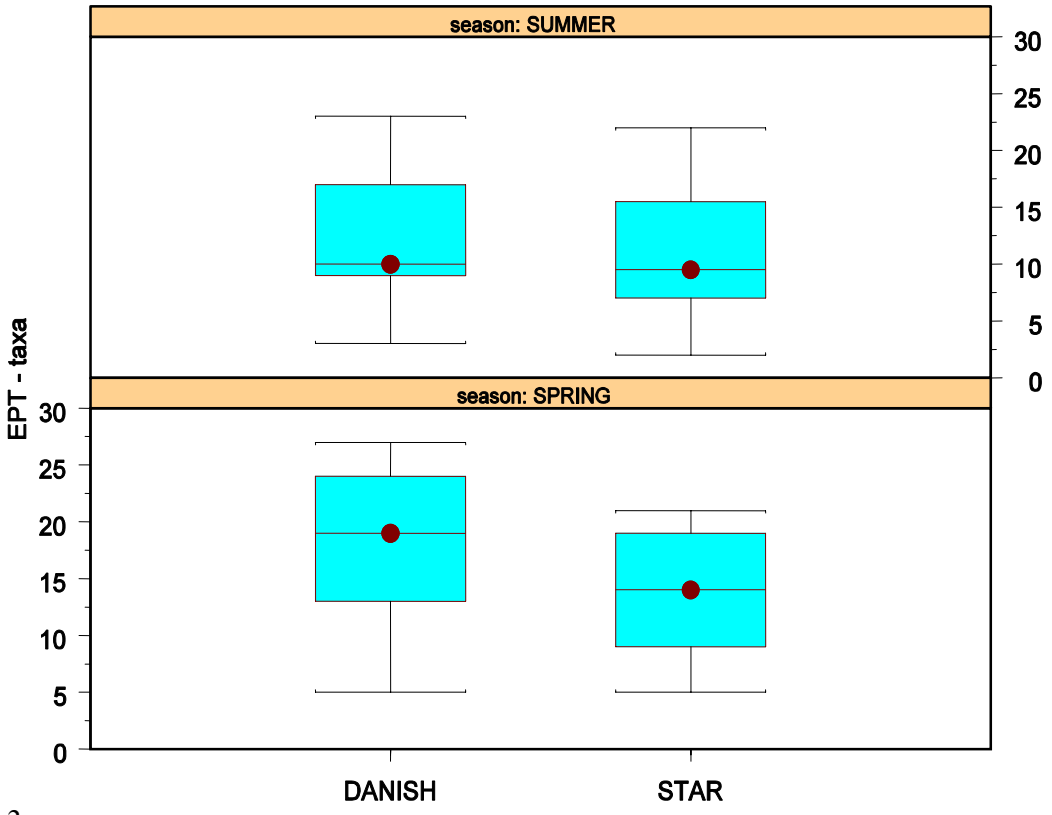
3 - Denmark



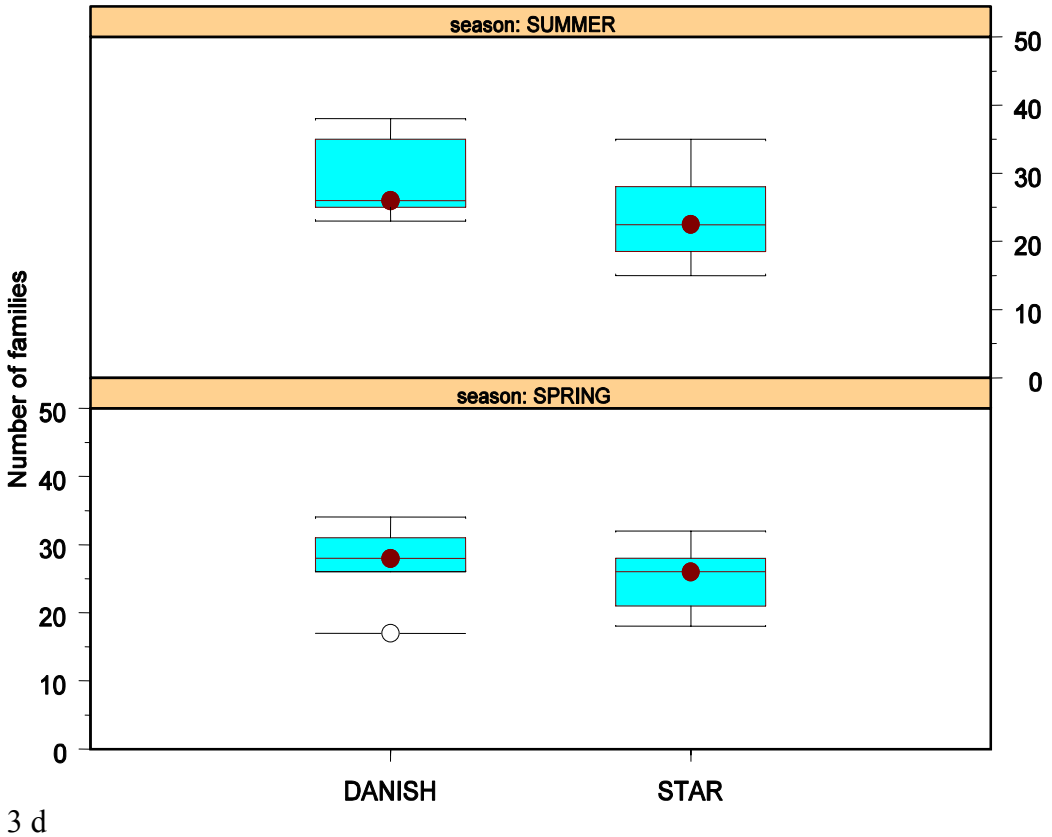
3a



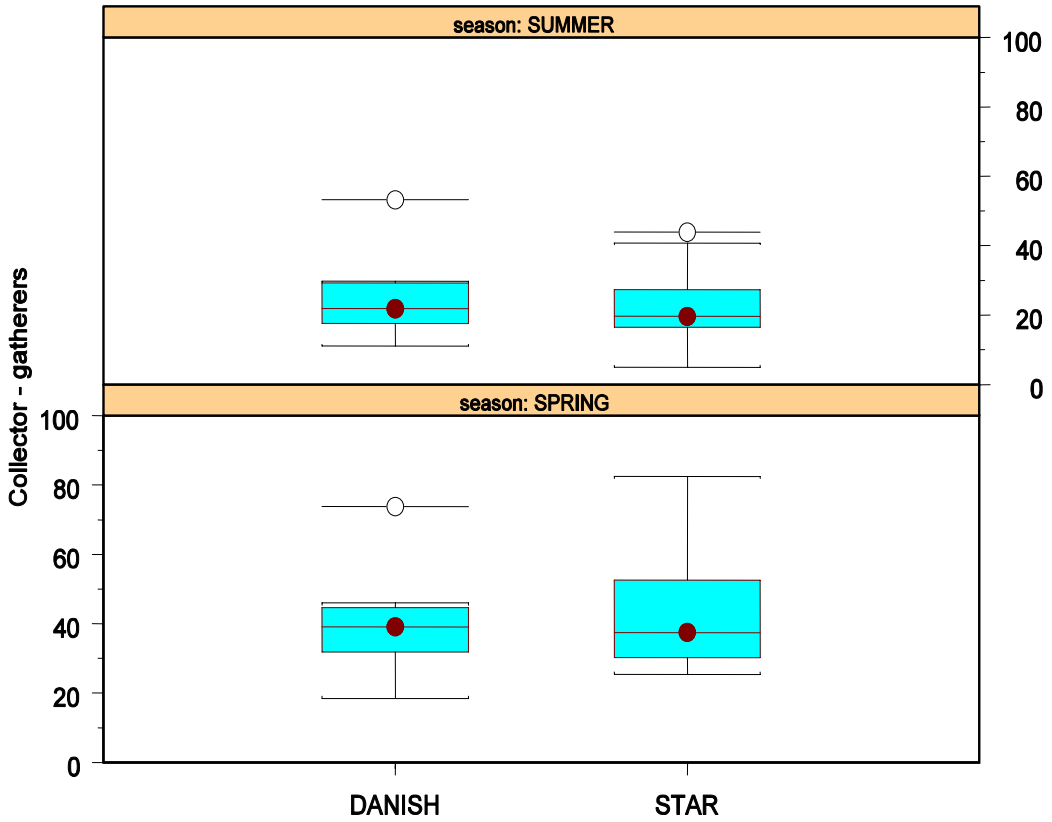
3 b


season: SUMMER
season: SPRING
EPT - taxa
0
5
10
15
20
25
30
DANISH
STAR

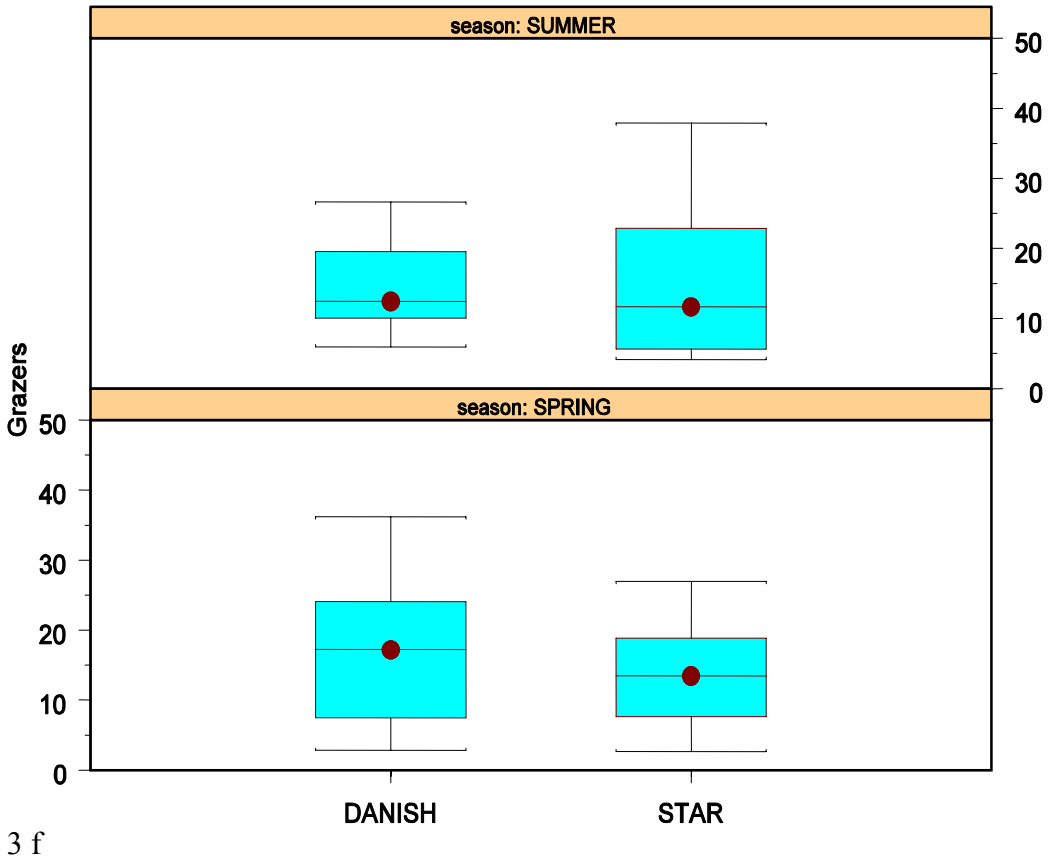
3 c


season: SUMMER
season: SPRING
Number of families
0
10
20
30
40
50
DANISH
STAR

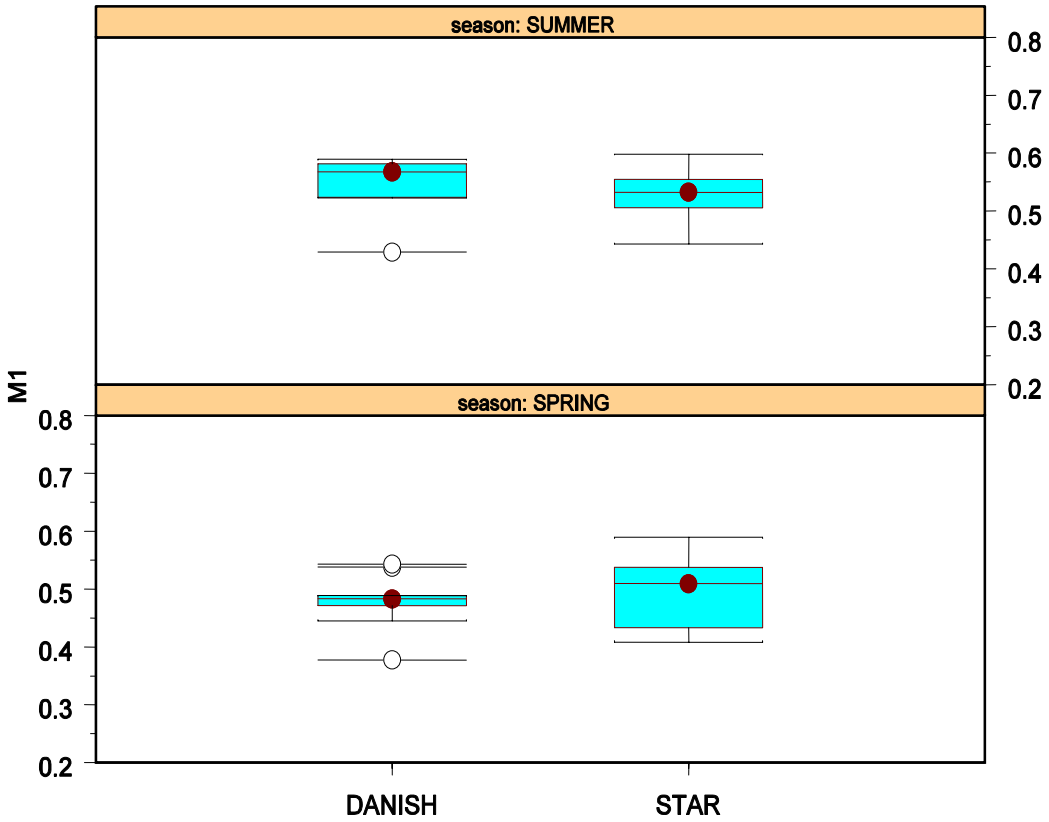
3 d


season: SUMMER
season: SPRING
Collector - gatherers
0
20
40
60
80
100
DANISH
STAR

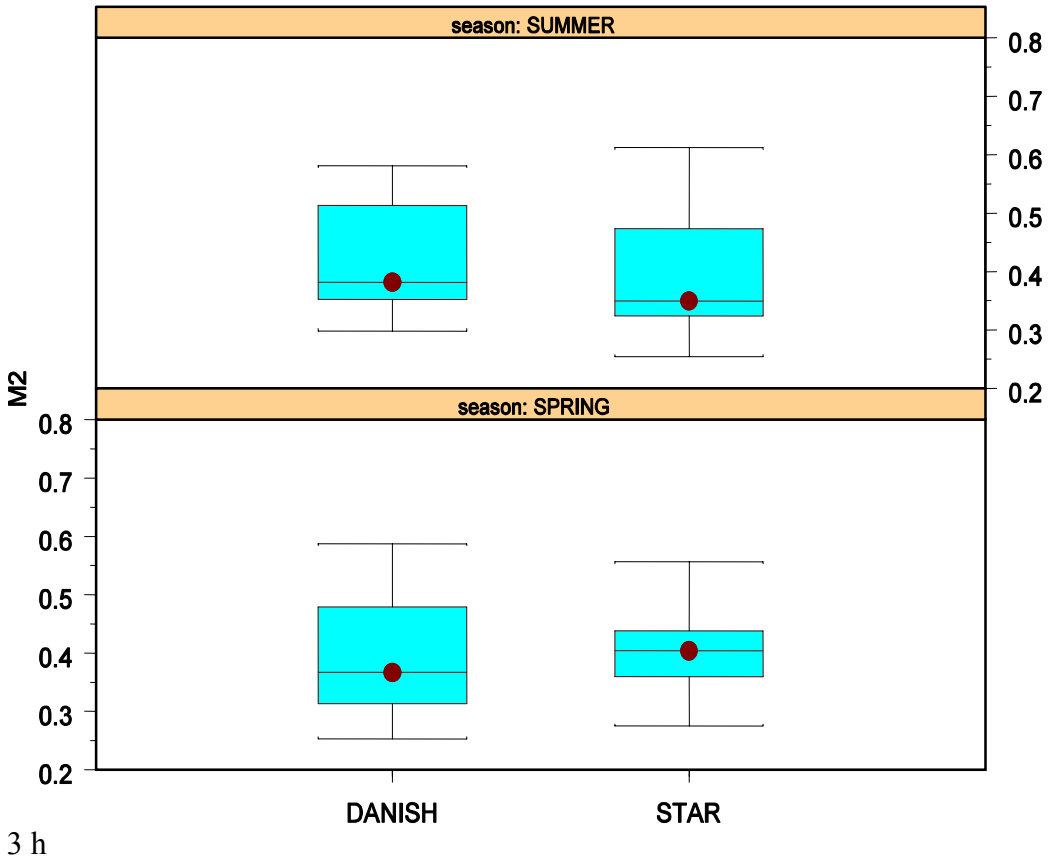
3 e


season: SUMMER
season: SPRING
Grazers
0
10
20
30
40
50
DANISH
STAR

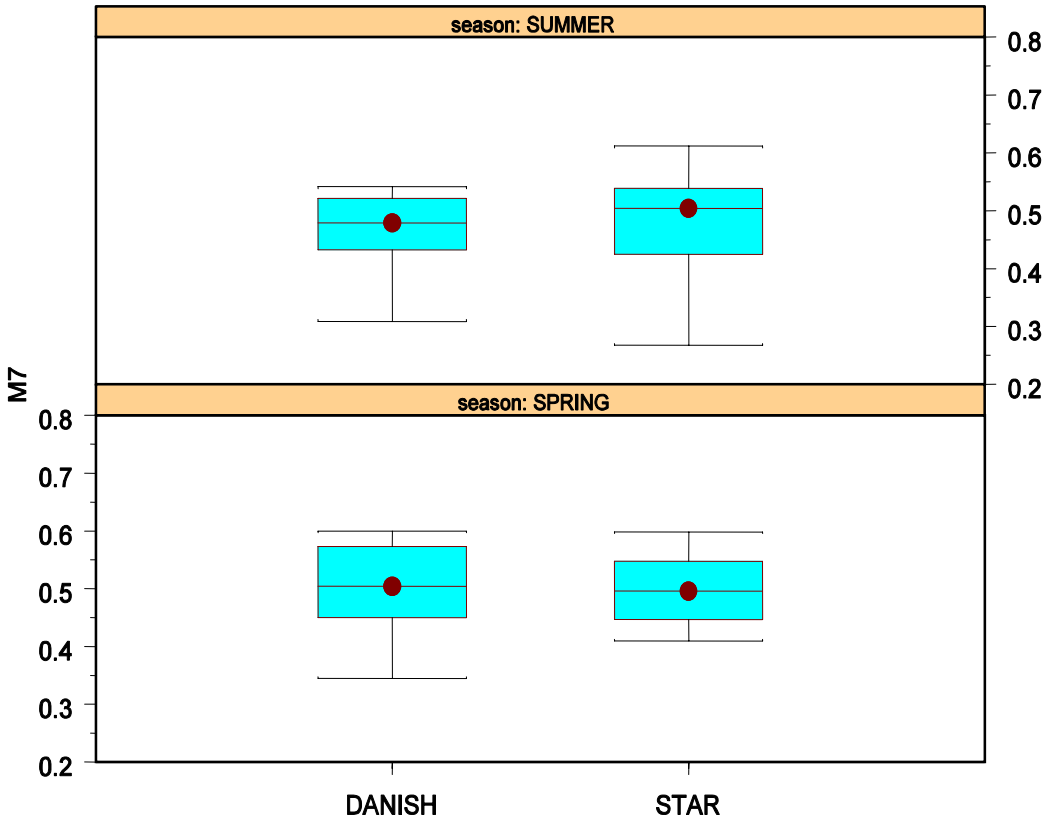
3 f


season: SUMMER
season: SPRING
M1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
DANISH
STAR

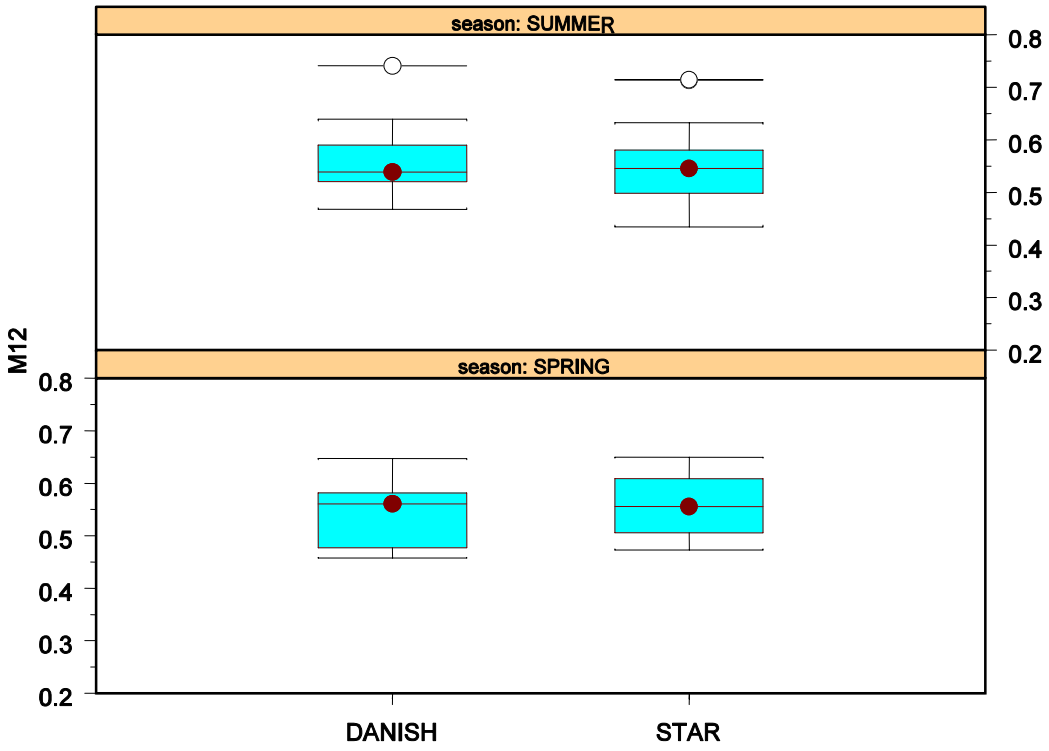
3 g


season: SUMMER
season: SPRING
M2
0.8
0.7
0.6
0.5
0.4
0.3
0.2
DANISH
STAR

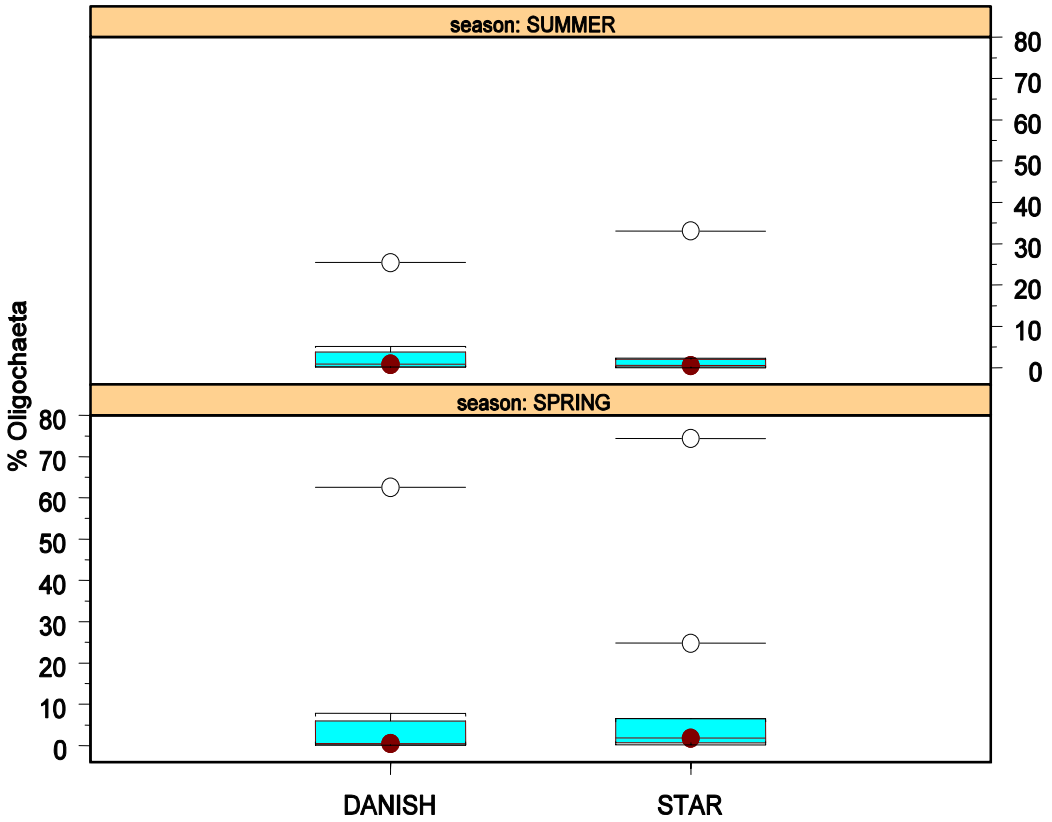
3 h


season: SUMMER
season: SPRING
M7
0.8
0.7
0.6
0.5
0.4
0.3
0.2
DANISH
STAR

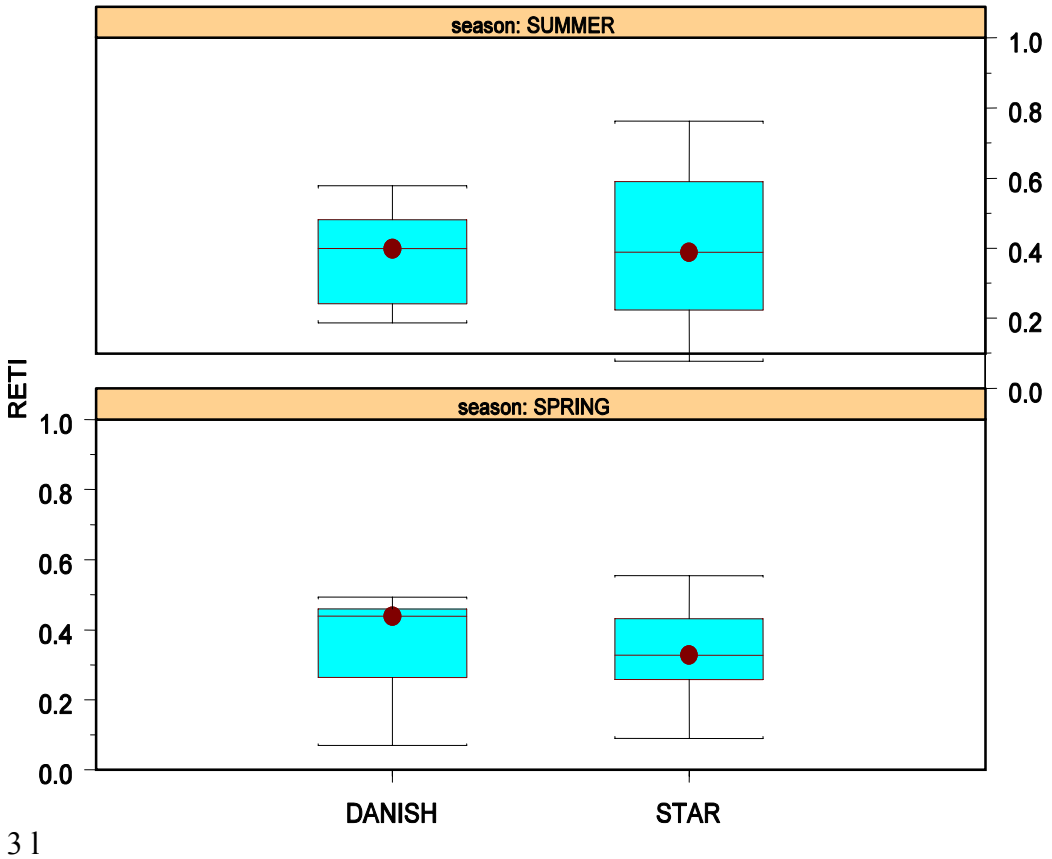
3 i


season: SUMMER
season: SPRING
M12
0.8
0.7
0.6
0.5
0.4
0.3
0.2
DANISH
STAR

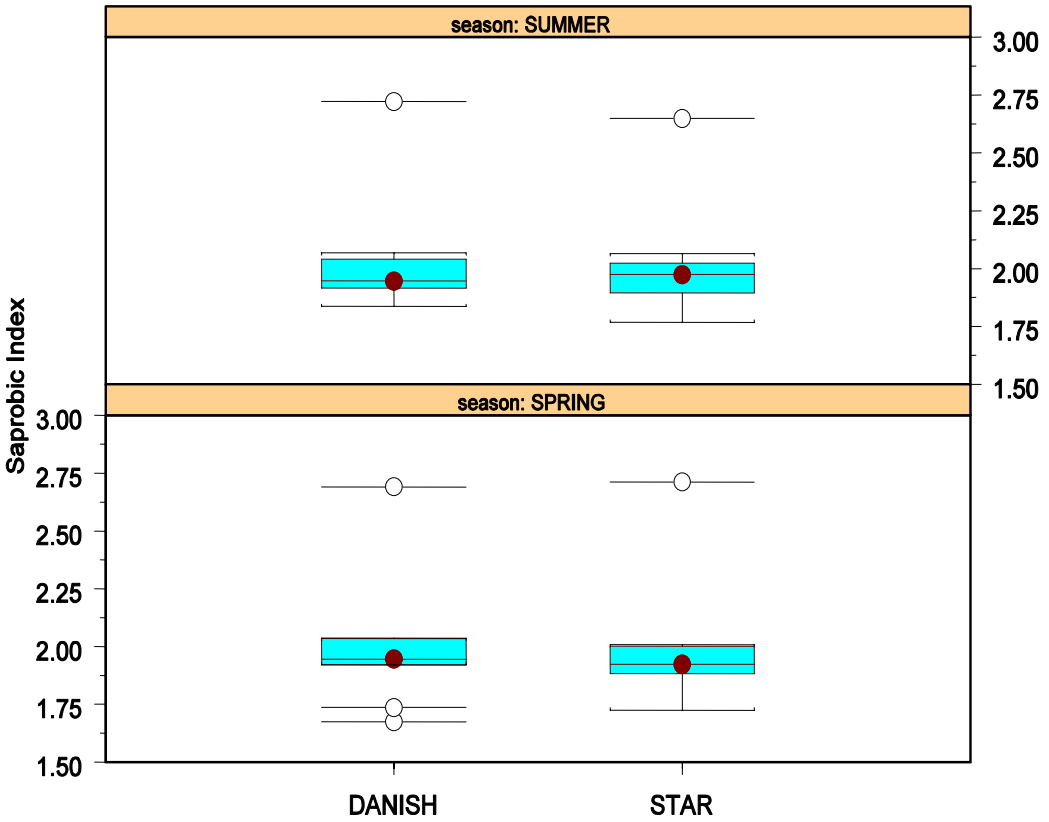
3 j


season: SUMMER
season: SPRING
% Oligochaeta
0
10
20
30
40
50
60
70
80
DANISH
STAR

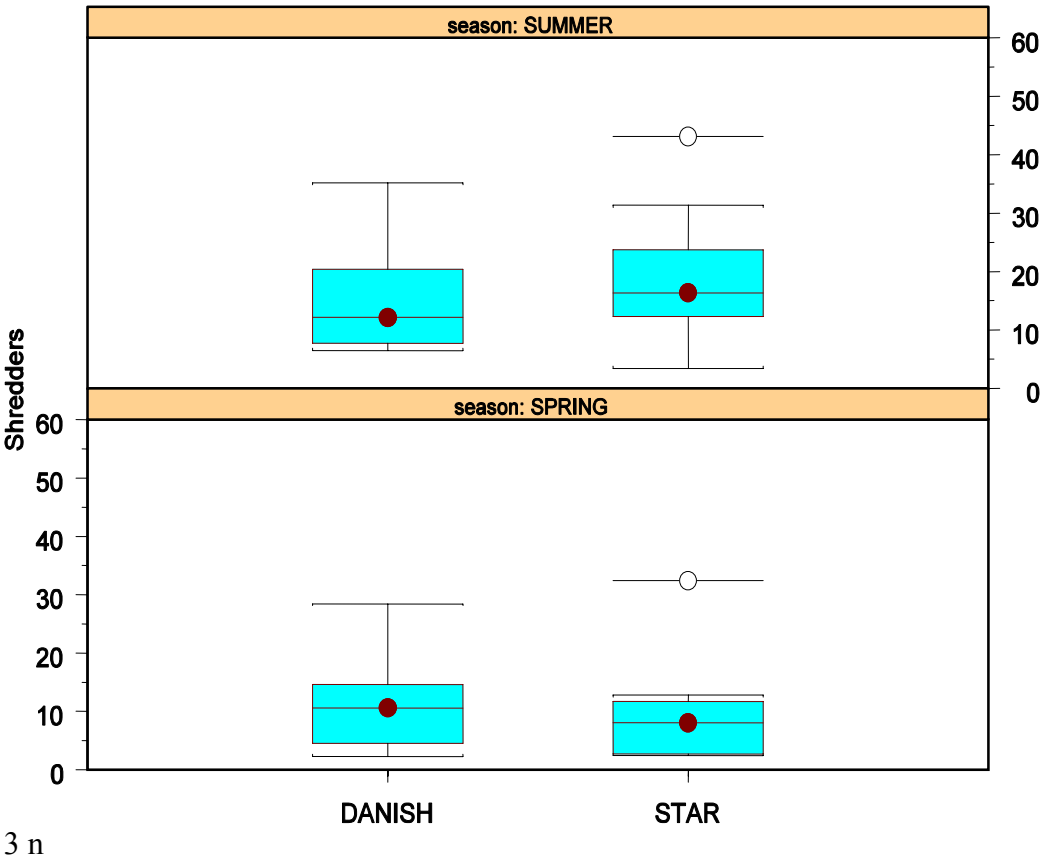
3 k


season: SUMMER
season: SPRING
RETI
0.0
0.2
0.4
0.6
0.8
1.0
DANISH
STAR

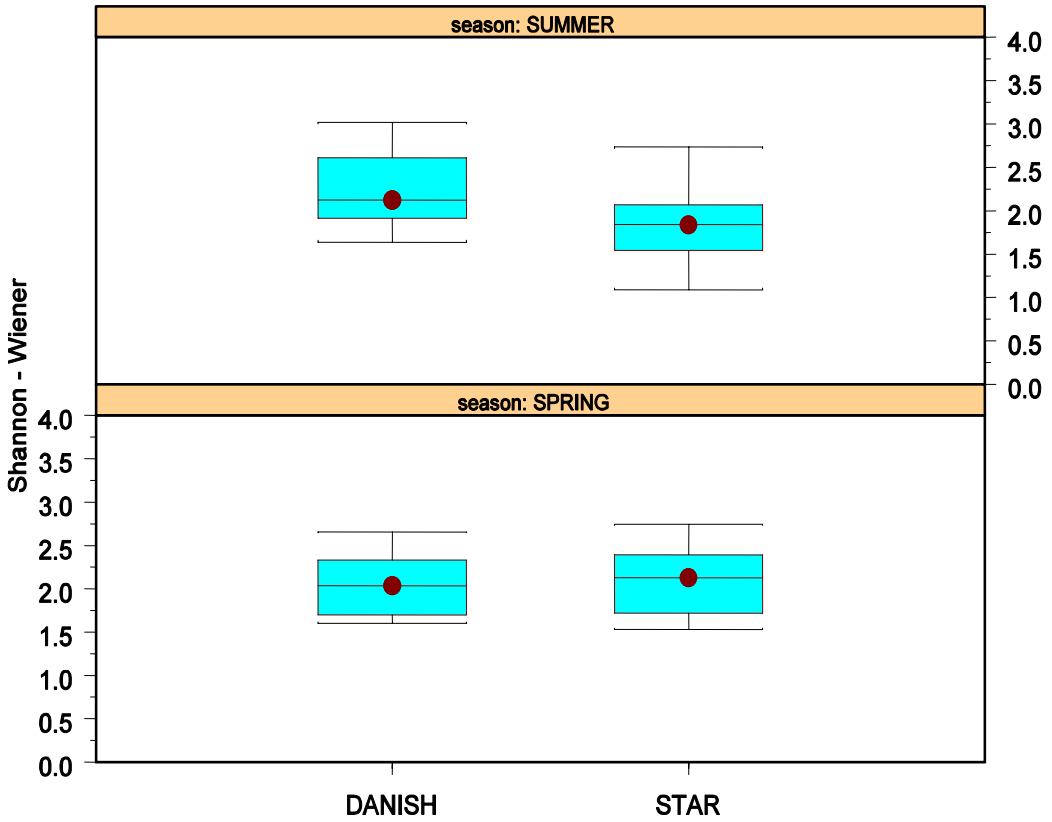
3 l


season: SUMMER
season: SPRING
Saprobic Index
3.00
2.75
2.50
2.25
2.00
1.75
1.50
DANISH
STAR

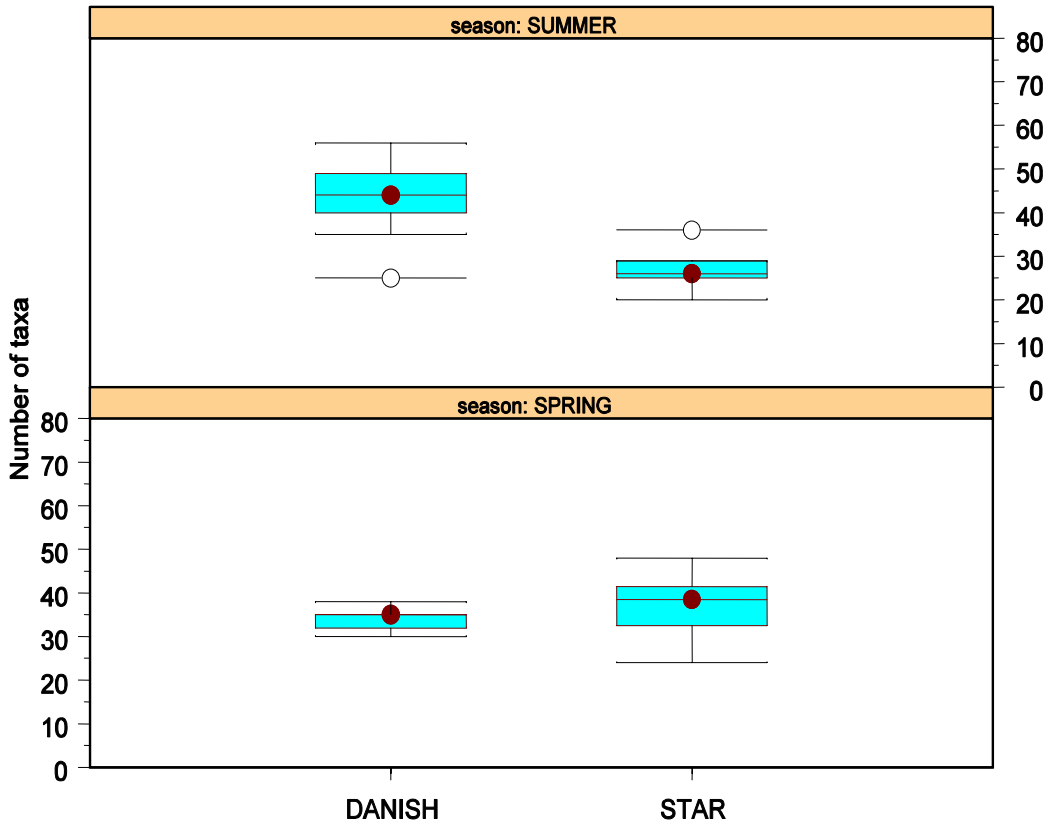
3 m


season: SUMMER
season: SPRING
Shredders
60
50
40
30
20
10
0
DANISH
STAR

3 n

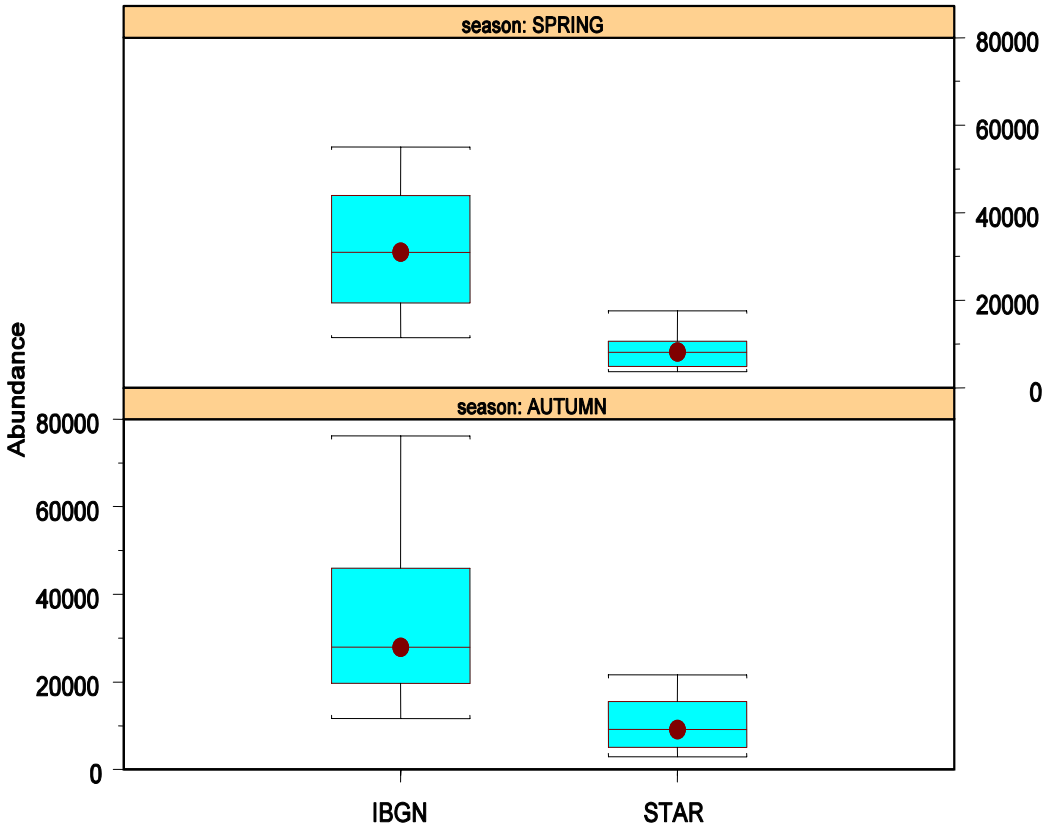

season: SUMMER
season: SPRING
Shannon - Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
DANISH
STAR

3 o

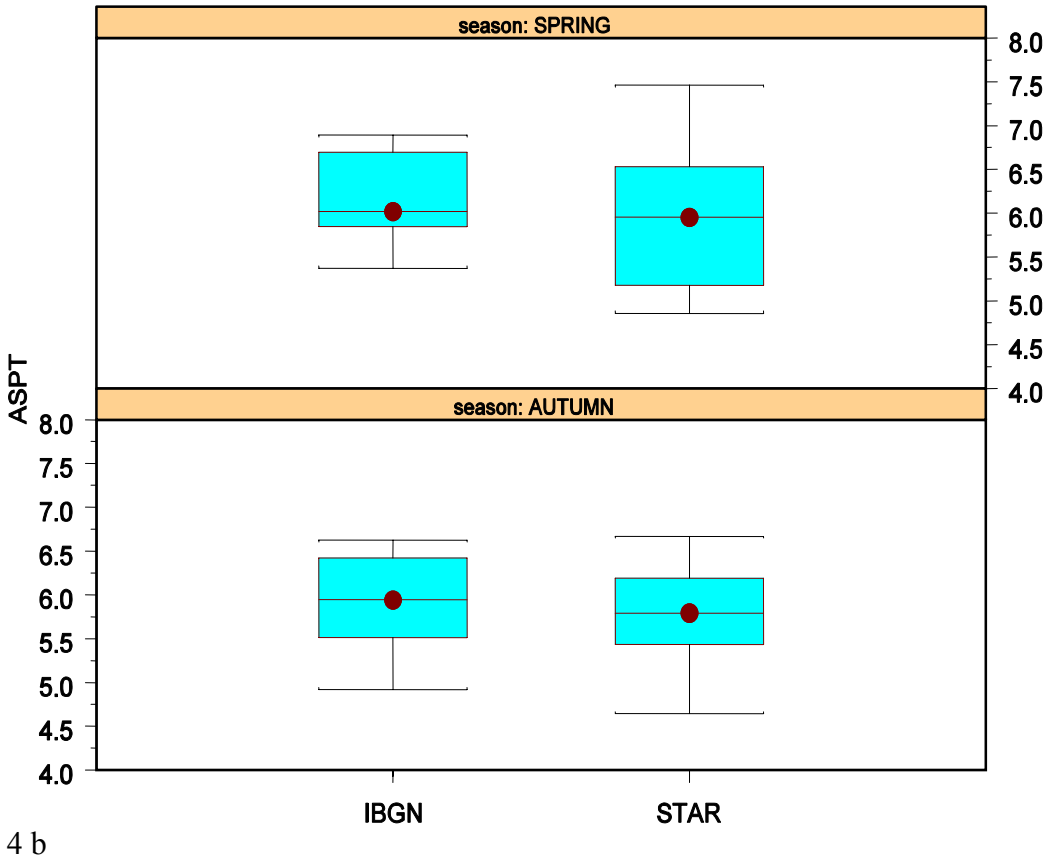

season: SUMMER
season: SPRING
Number of taxa
80
70
60
50
40
30
20
10
0
DANISH
STAR

3 p

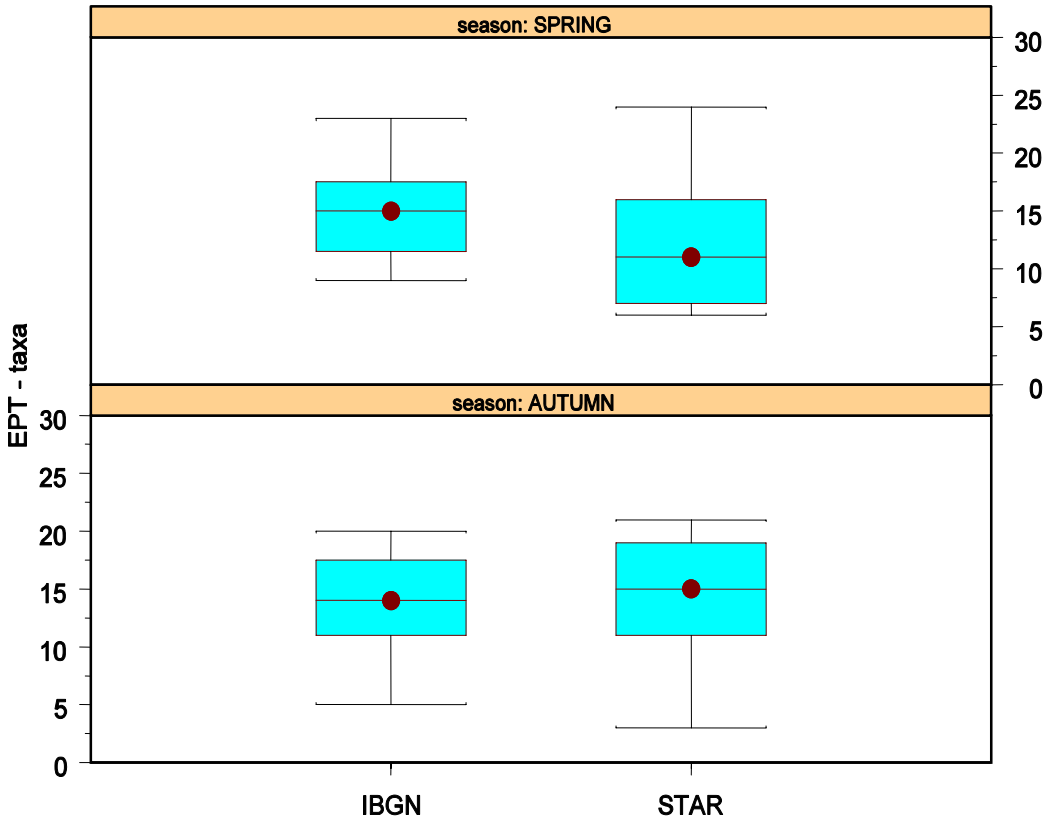
4 – France



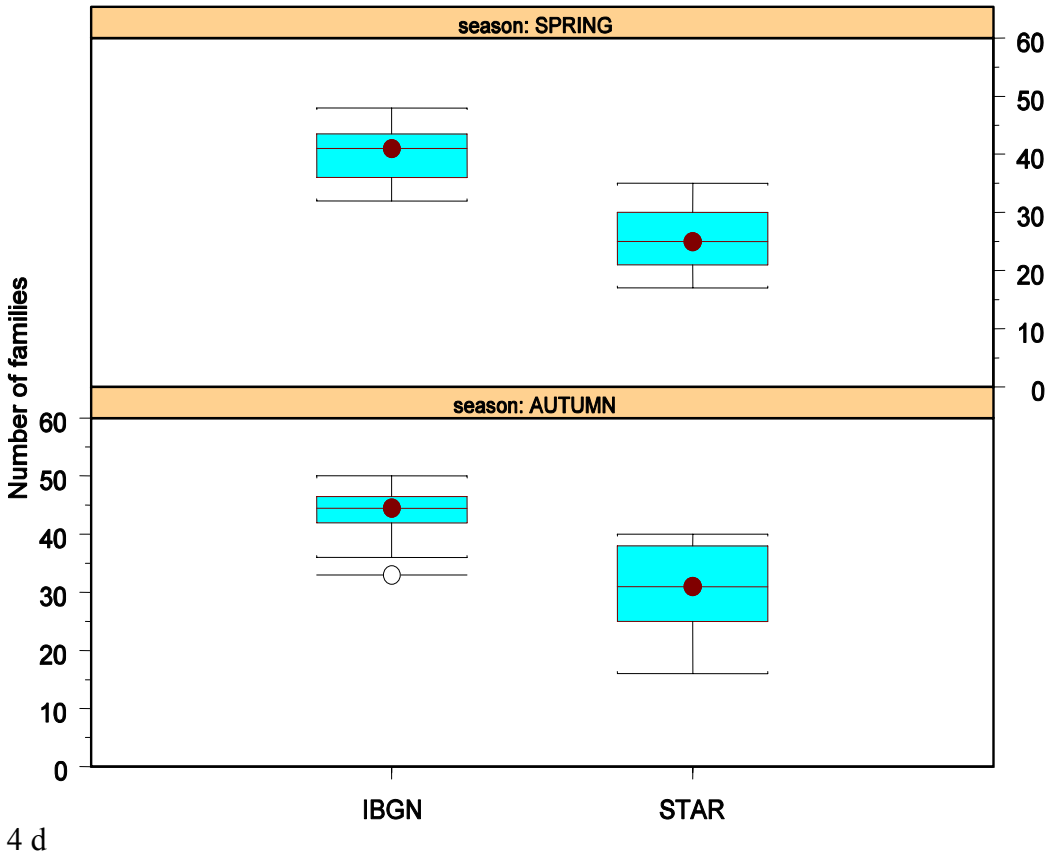
4 a



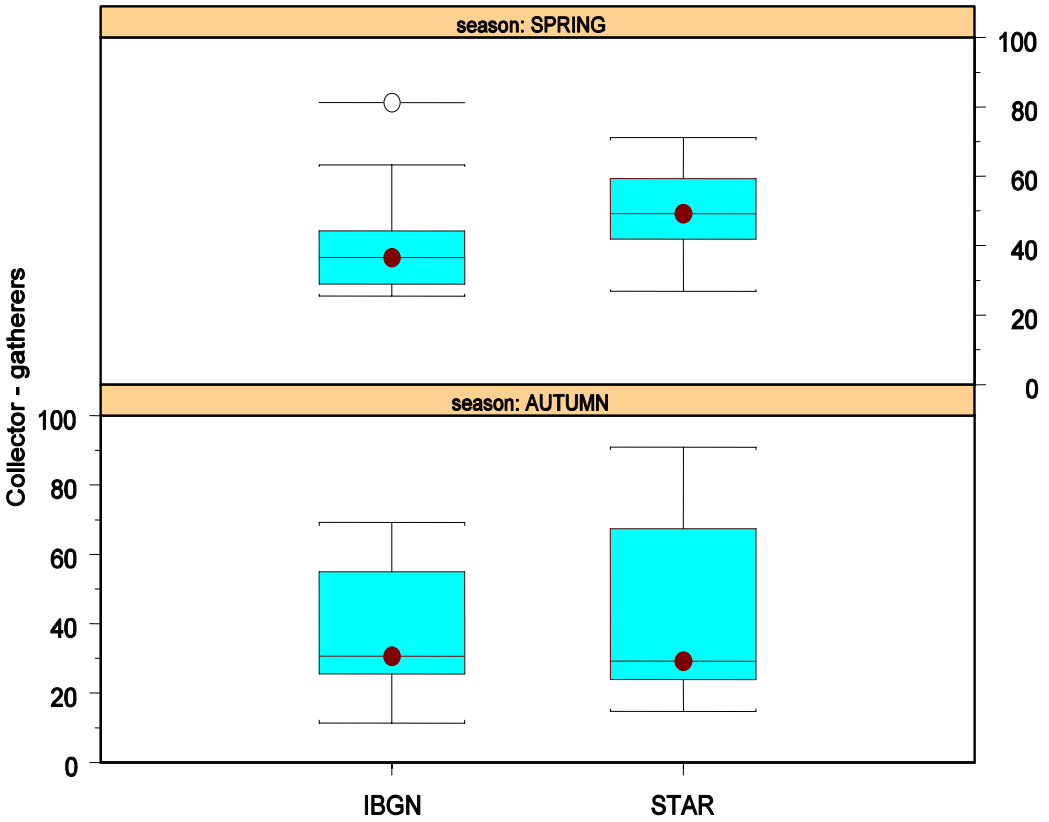
4 b


season: SPRING
season: AUTUMN
EPT - taxa
0
5
10
15
20
25
30
IBGN
STAR

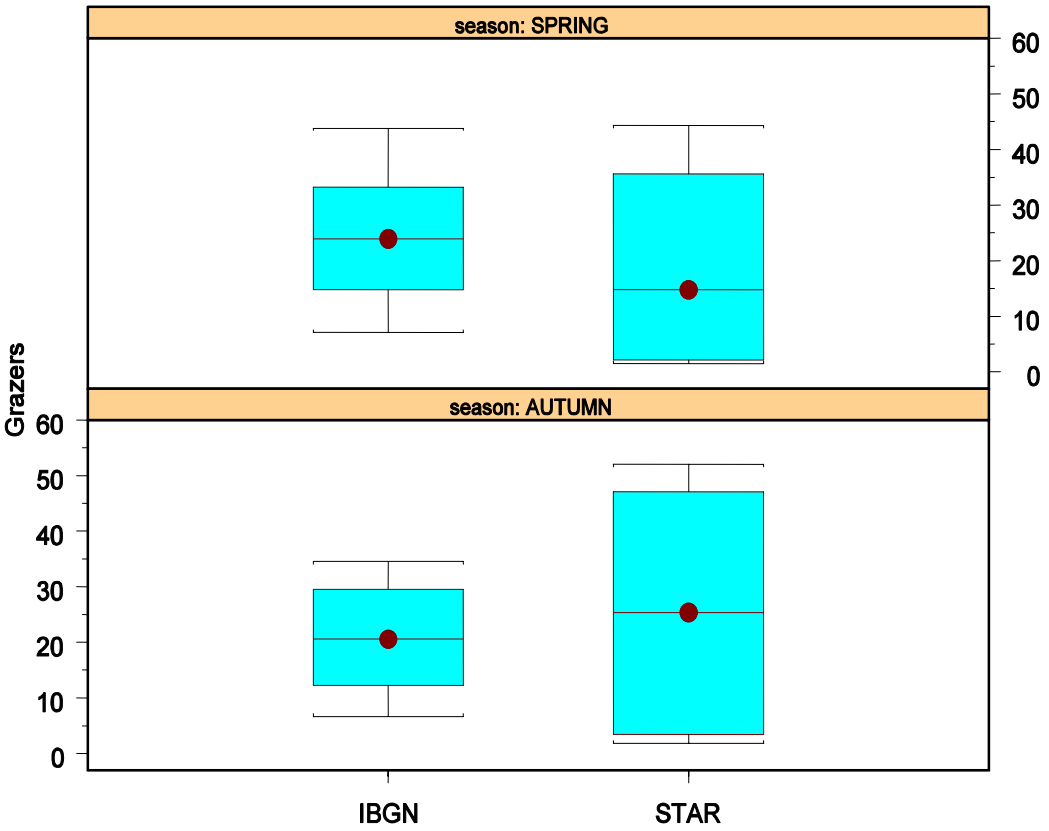
4 c


season: SPRING
season: AUTUMN
Number of families
0
10
20
30
40
50
60
IBGN
STAR

4 d

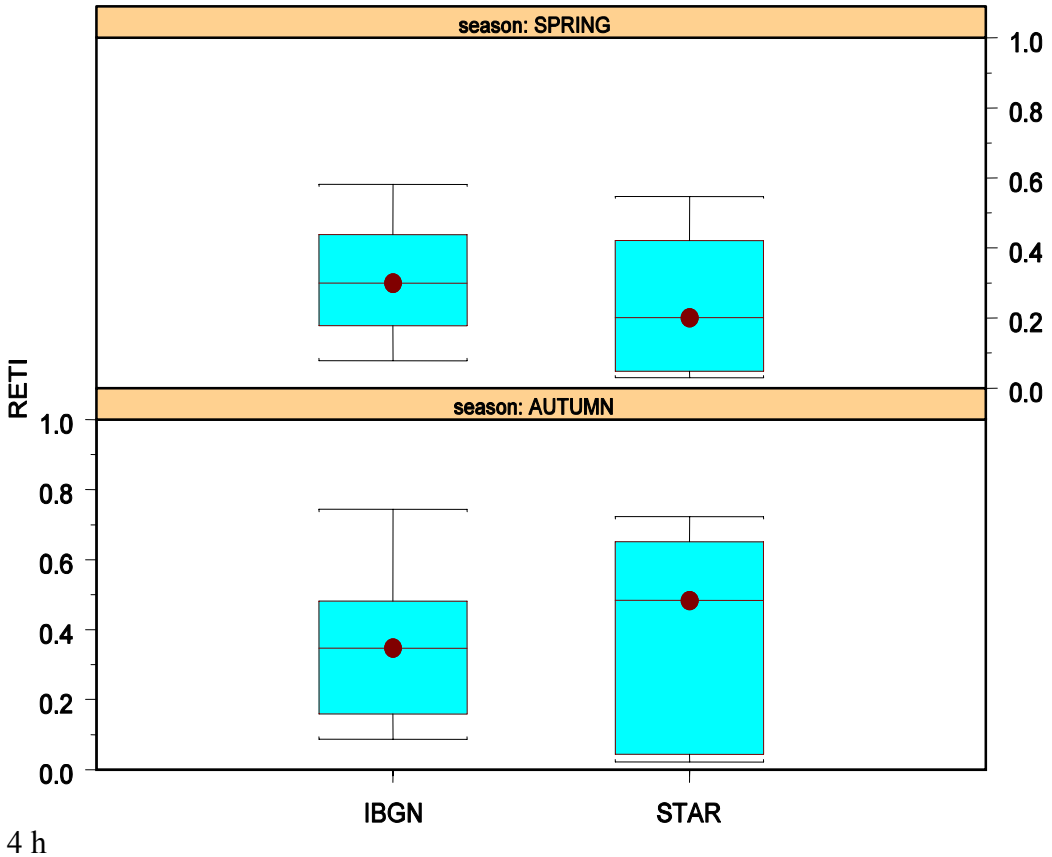

season: SPRING
season: AUTUMN
Collector - gatherers
0
20
40
60
80
100
IBGN
STAR

4 e

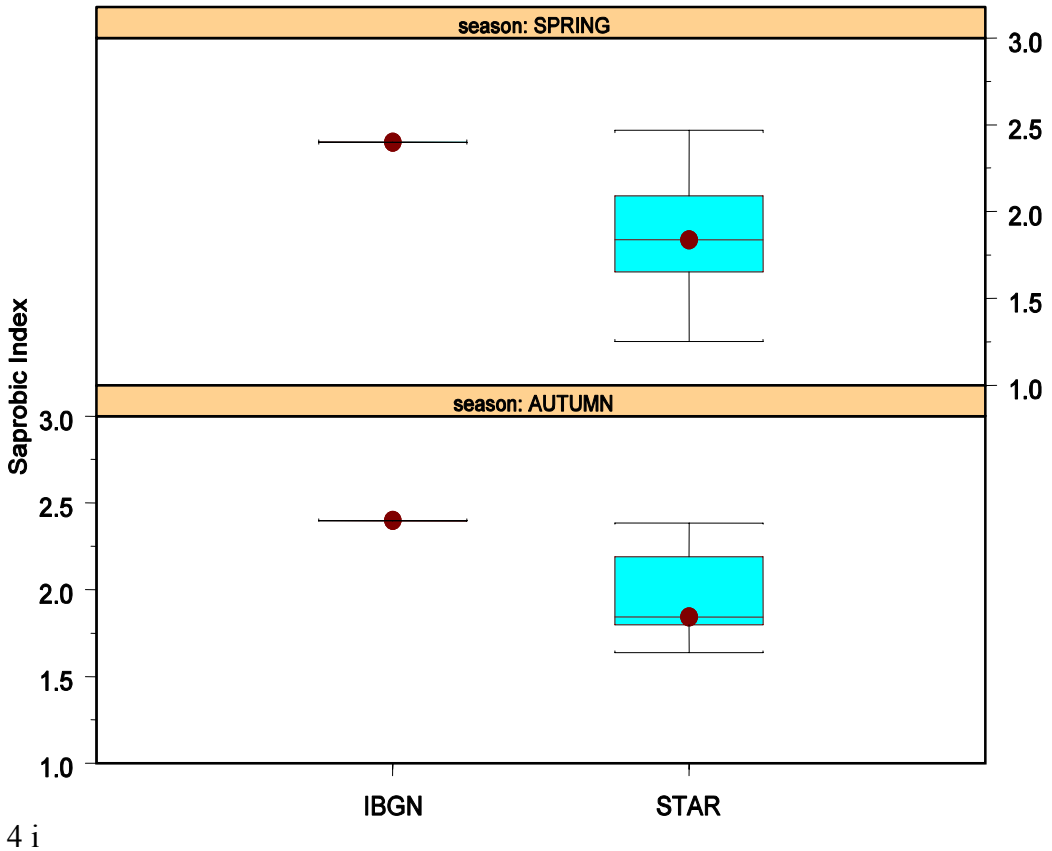

season: SPRING
season: AUTUMN
Grazers
0
10
20
30
40
50
60
IBGN
STAR

4 f

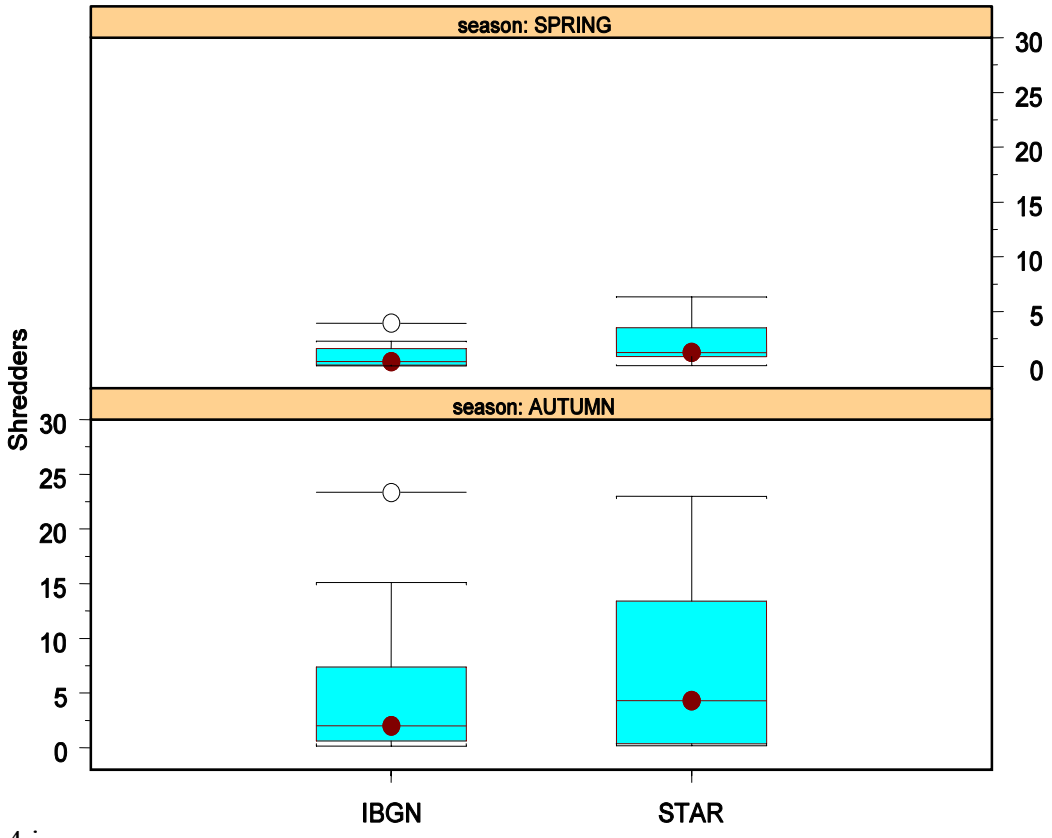
r
4 g


season: SPRING
season: AUTUMN
RETI
1.0
0.8
0.6
0.4
0.2
0.0
IBGN
STAR

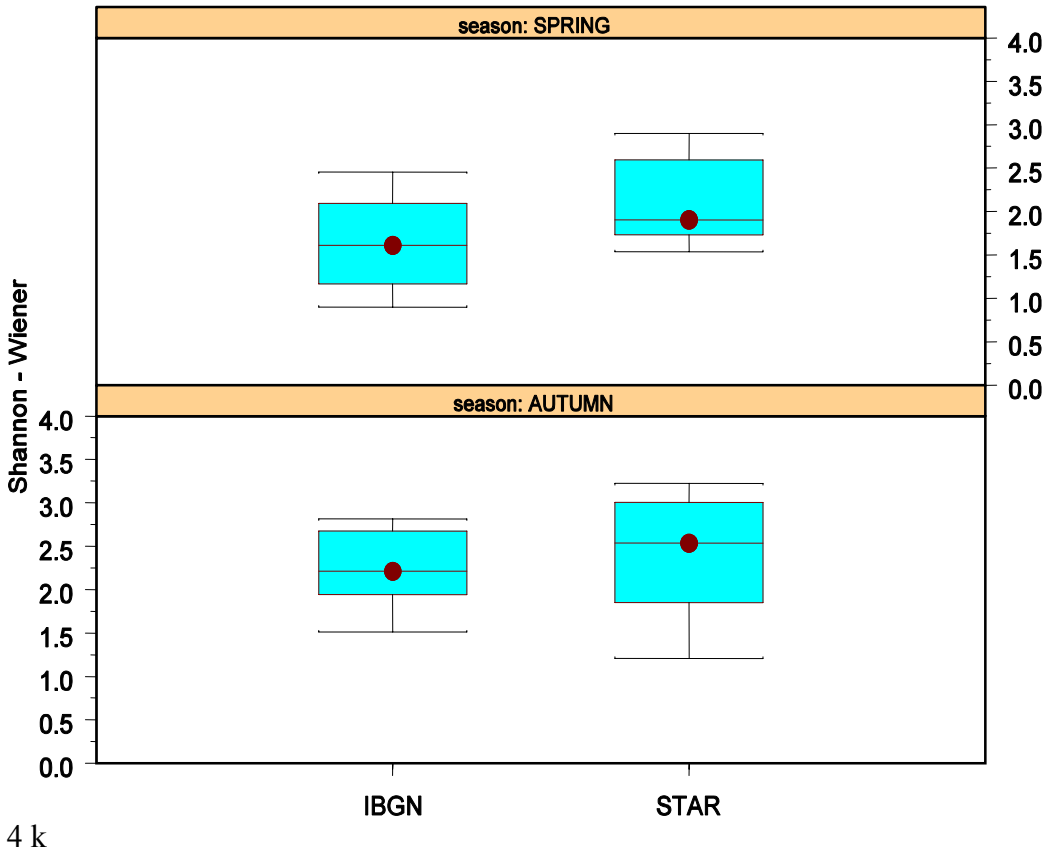
4 h


season: SPRING
season: AUTUMN
Saprobic Index
3.0
2.5
2.0
1.5
1.0
IBGN
STAR

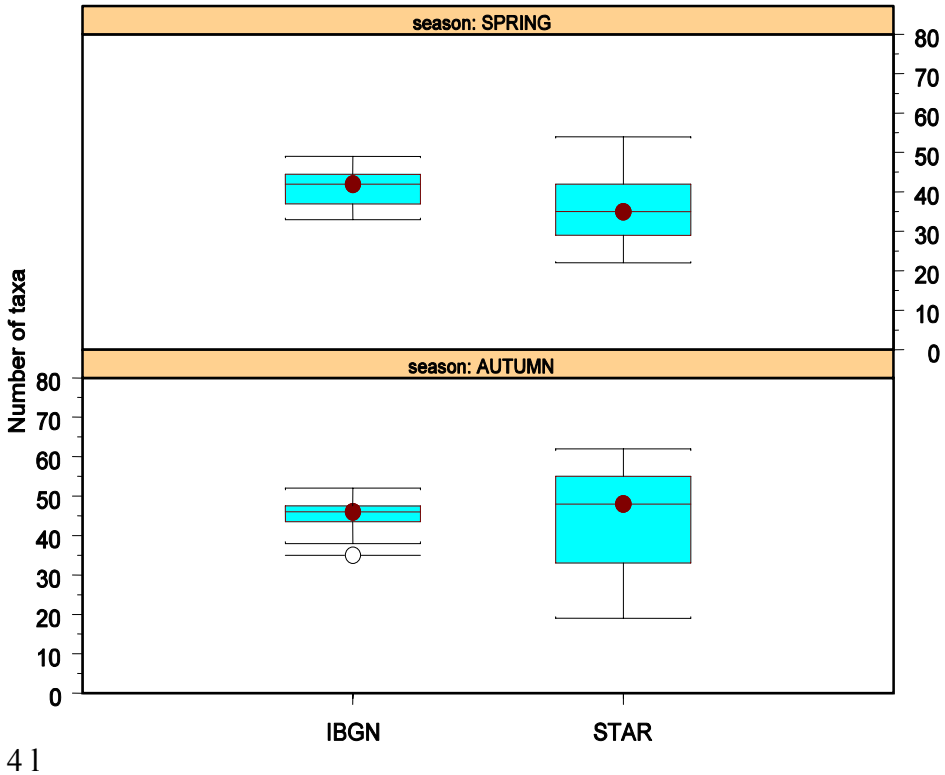
4 i


season: SPRING
season: AUTUMN
Shredders
30
25
20
15
10
5
0
IBGN
STAR

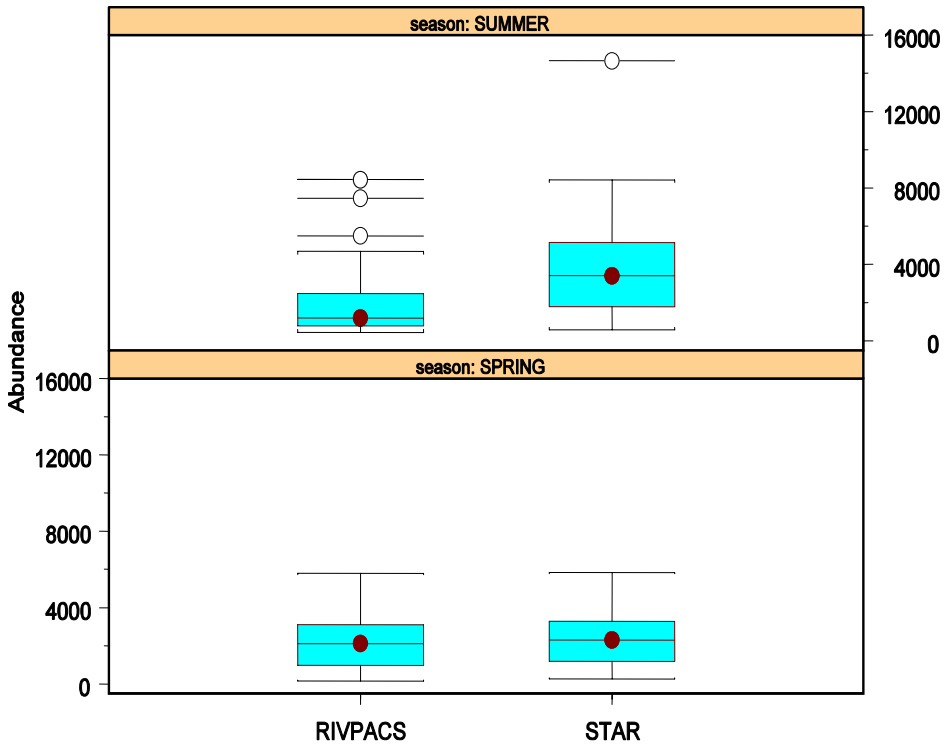
4 j


season: SPRING
season: AUTUMN
Shannon - Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
IBGN
STAR

4 k

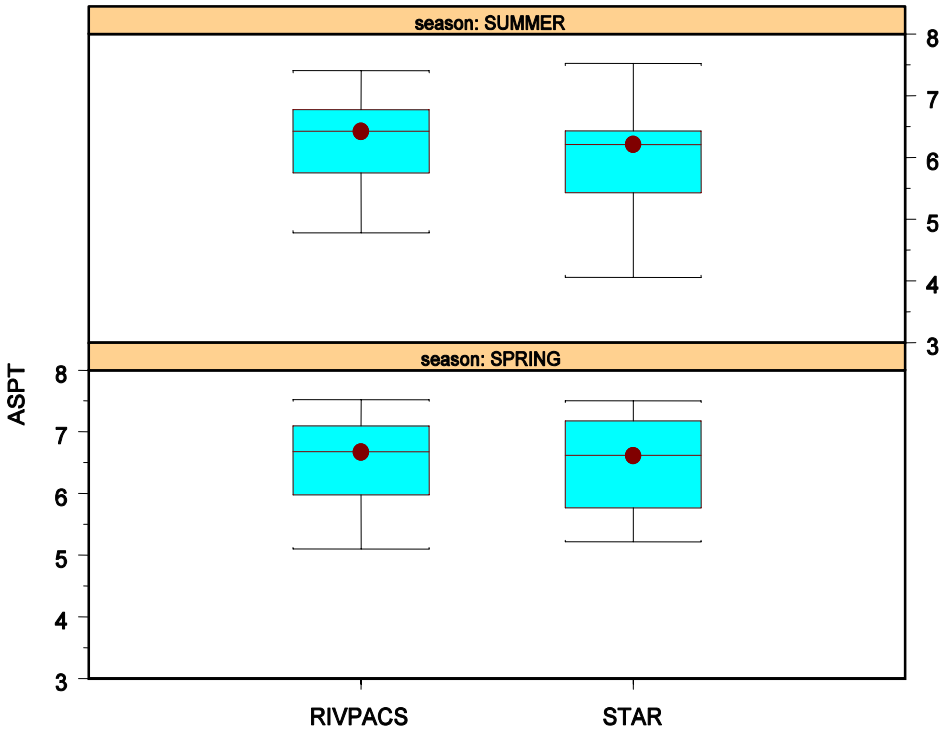

season: SPRING
season: AUTUMN
Number of taxa
0
10
20
30
40
50
60
70
80
IBGN
STAR

4 1

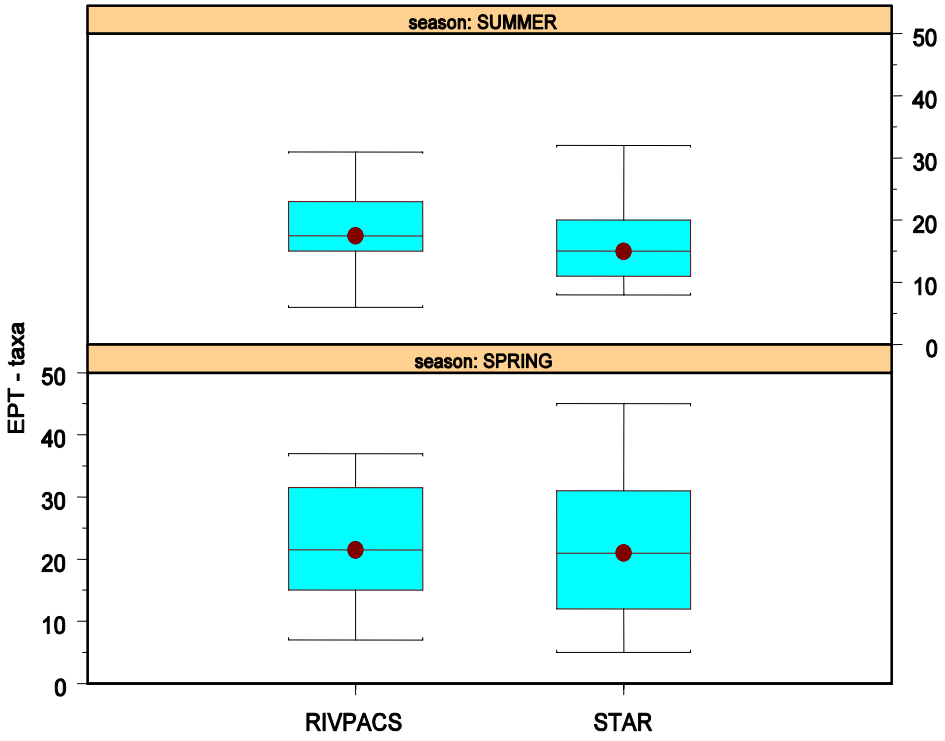

season: SUMMER
season: SPRING
Abundance
0
4000
8000
12000
16000
RIVPACS
STAR

5 - Germany

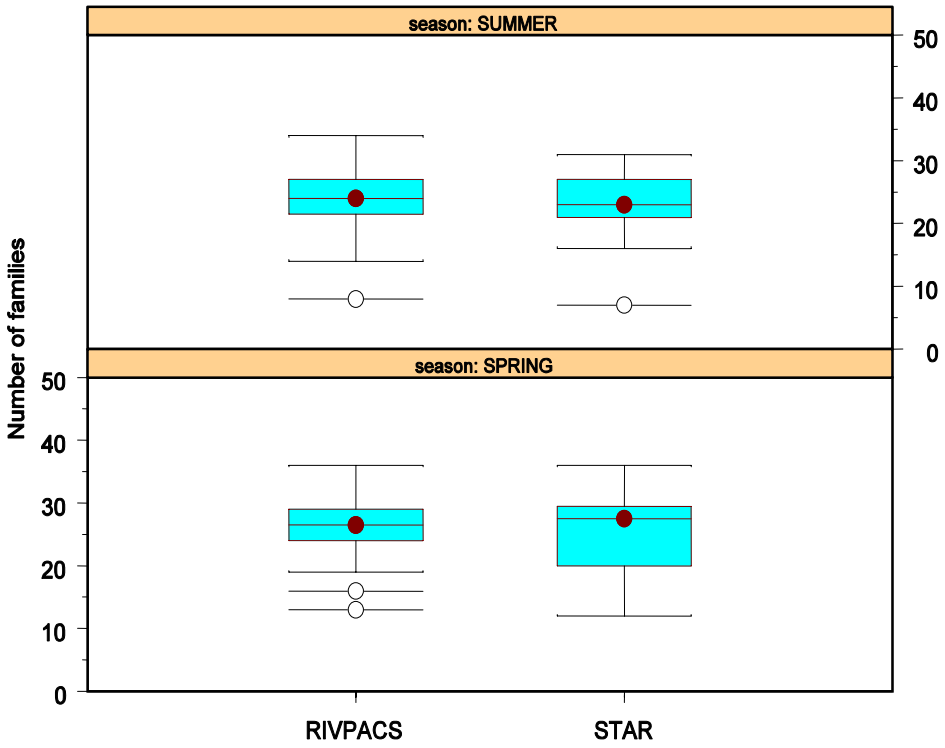
5 a



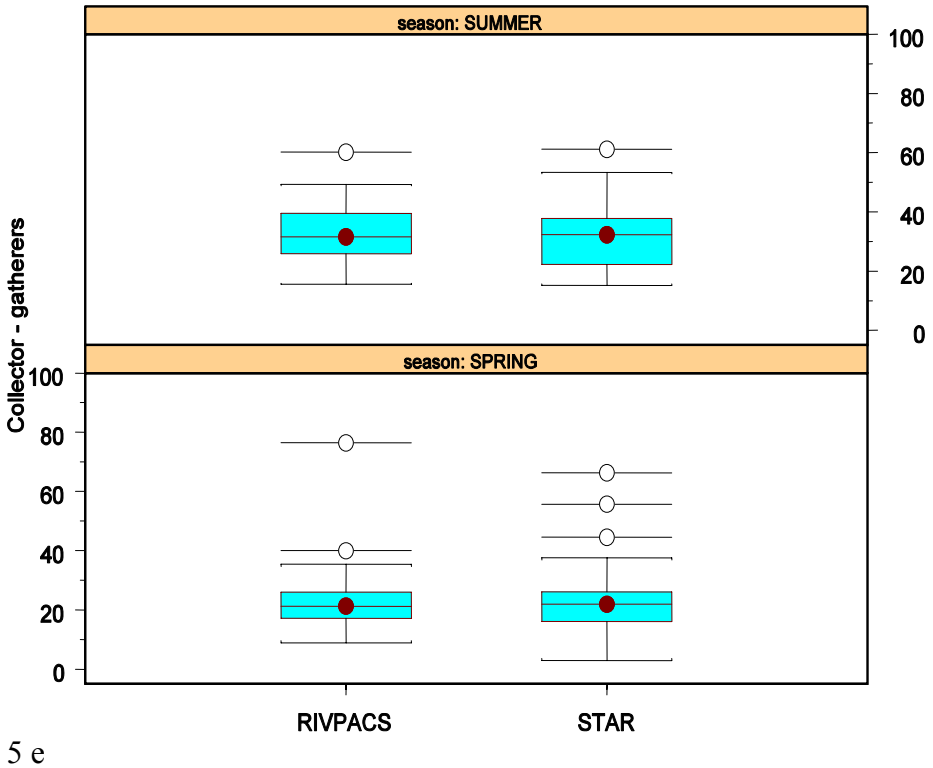
5 b



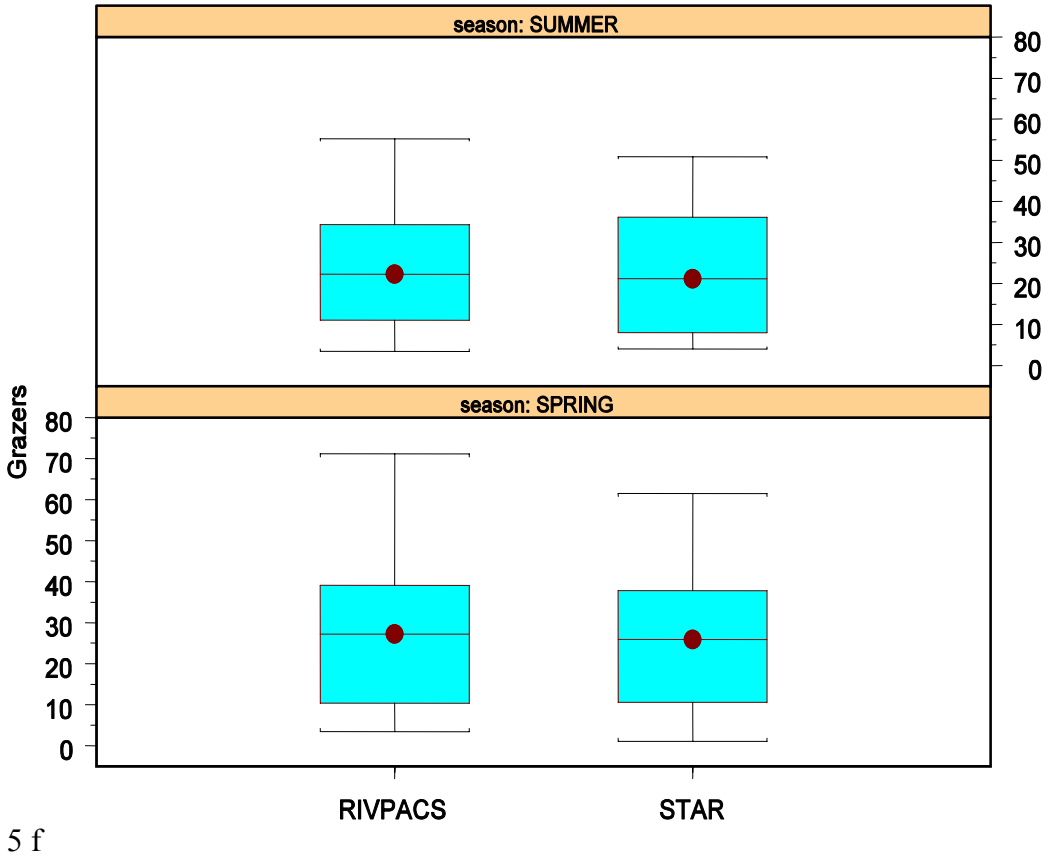
5 c


season: SUMMER
season: SPRING
Number of families
0
10
20
30
40
50
RIVPACS
STAR

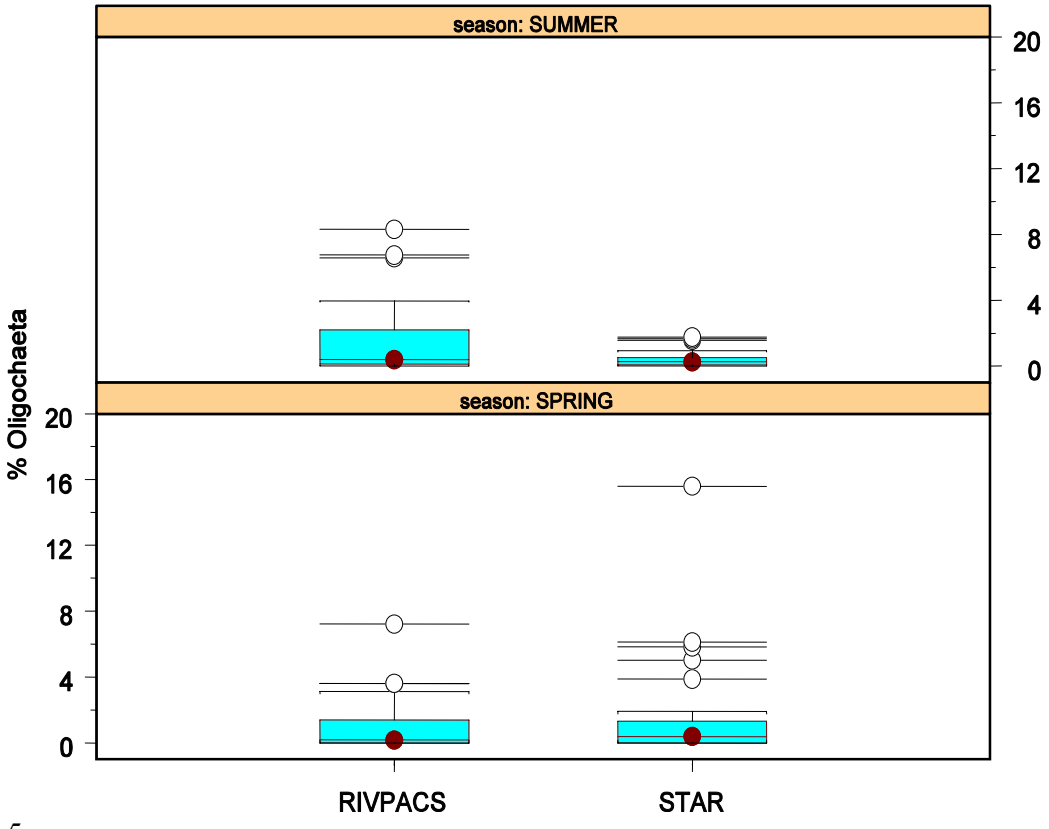
5 d


season: SUMMER
season: SPRING
Collector - gatherers
0
20
40
60
80
100
RIVPACS
STAR

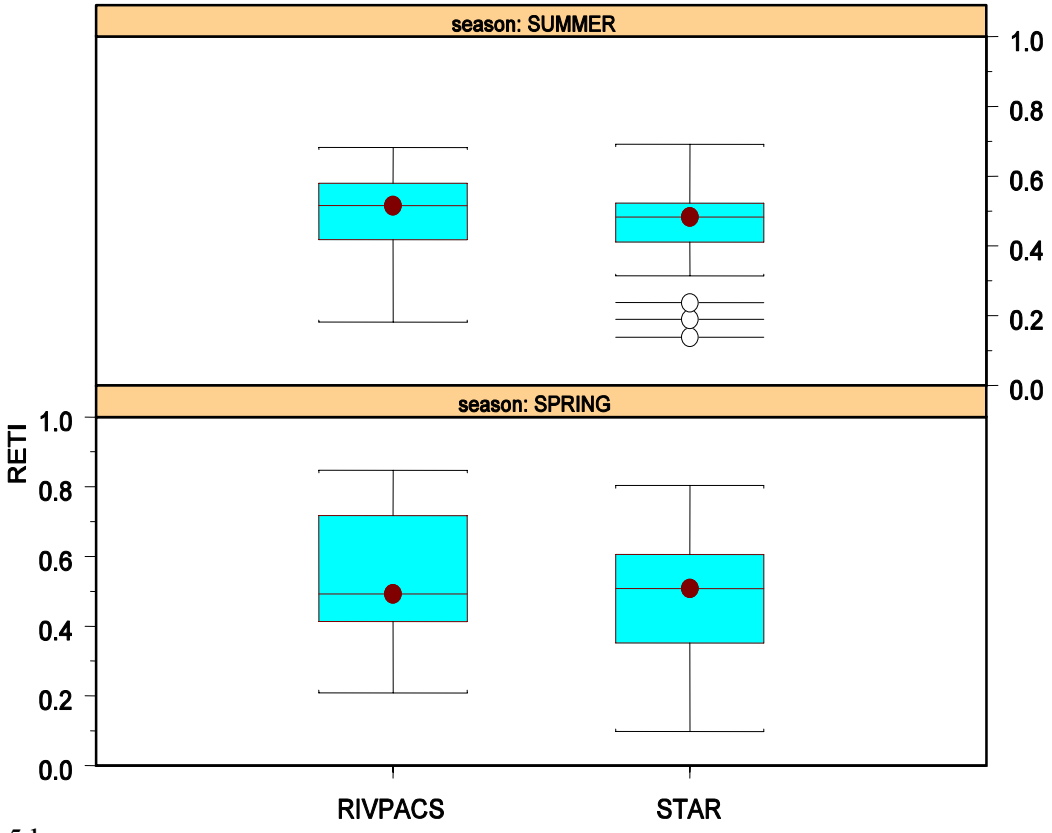
5 e


season: SUMMER
season: SPRING
Grazers
0
10
20
30
40
50
60
70
80
RIVPACS
STAR

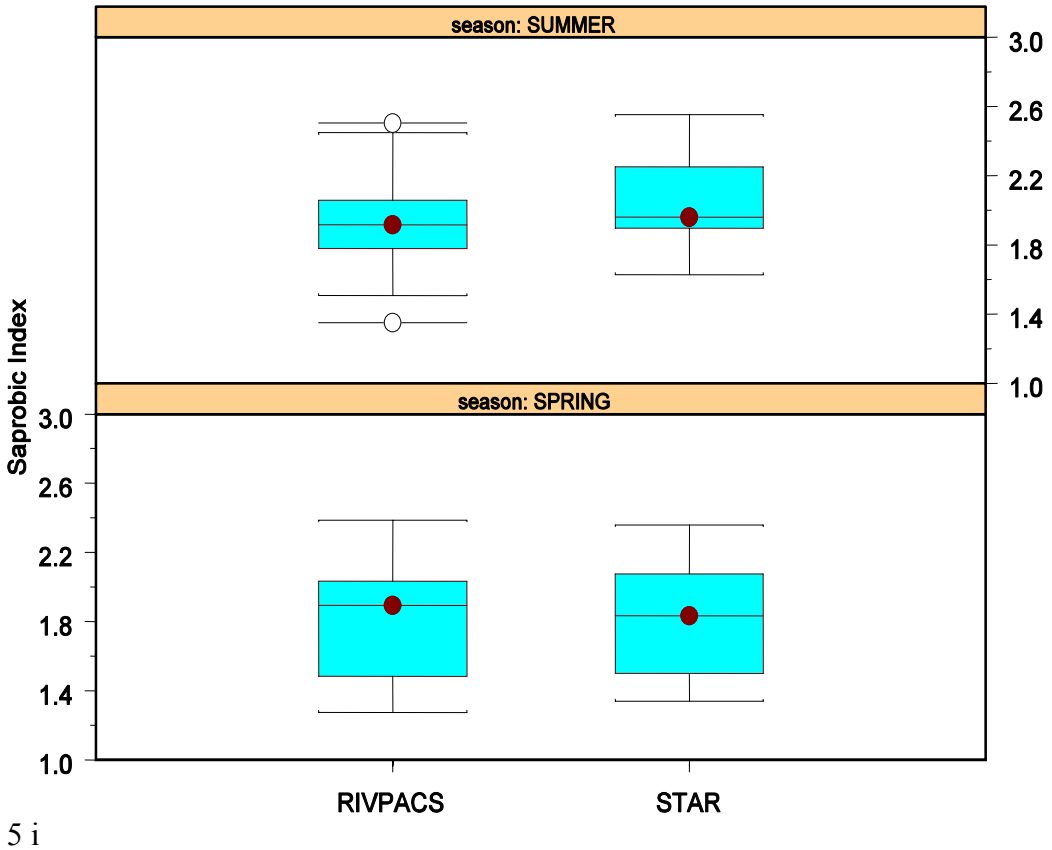
5 f


season: SUMMER
season: SPRING
% Oligochaeta
0
4
8
12
16
20
RIVPACS
STAR

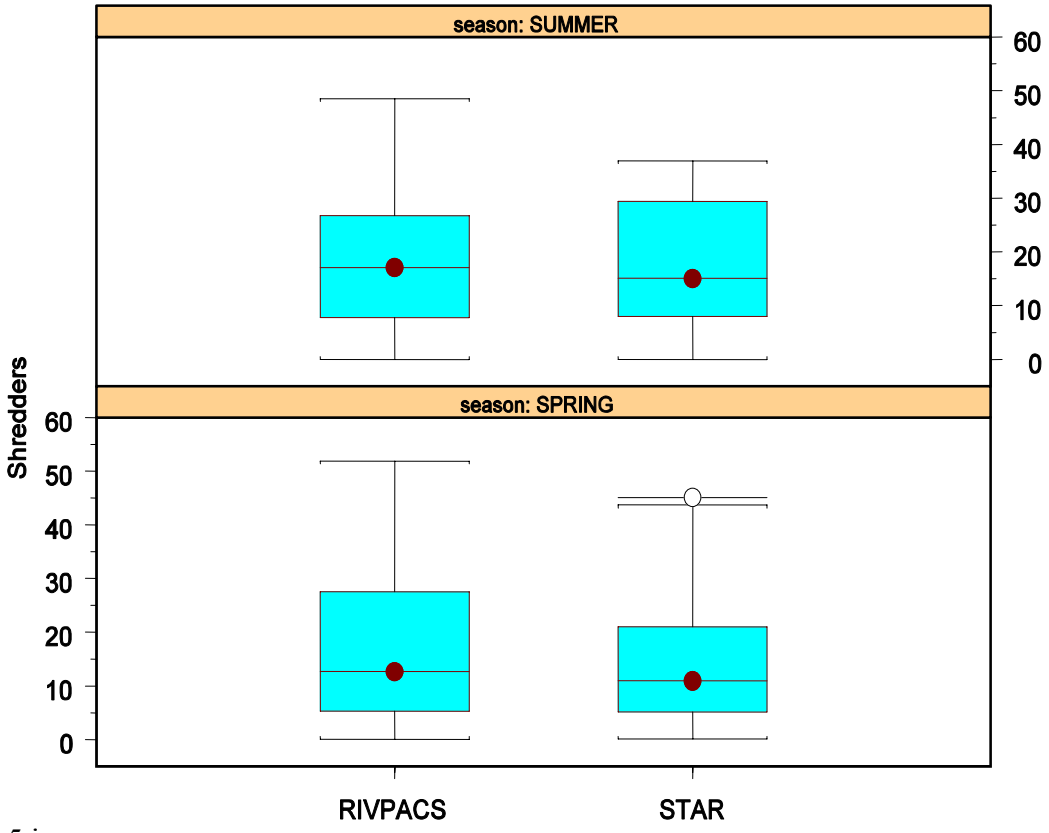
5 g


season: SUMMER
season: SPRING
RETI
1.0
0.8
0.6
0.4
0.2
0.0
RIVPACS
STAR

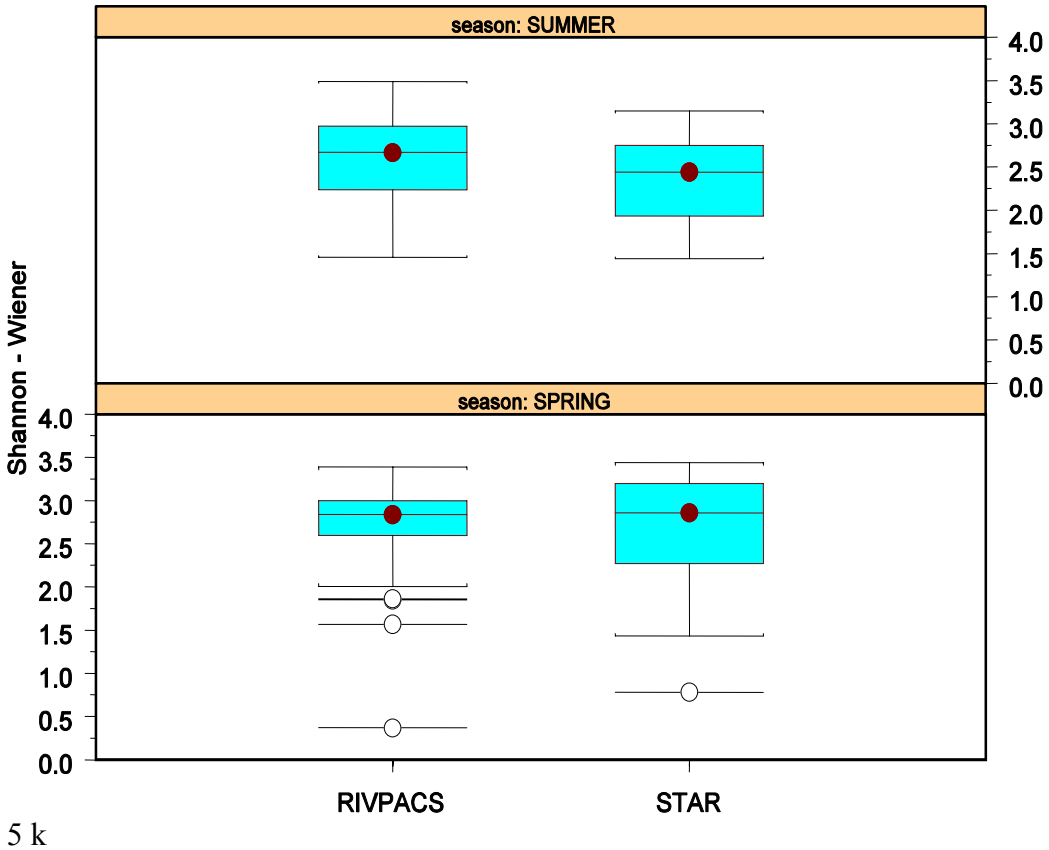
5 h


season: SUMMER
season: SPRING
Saprobic Index
3.0
2.6
2.2
1.8
1.4
1.0
RIVPACS
STAR

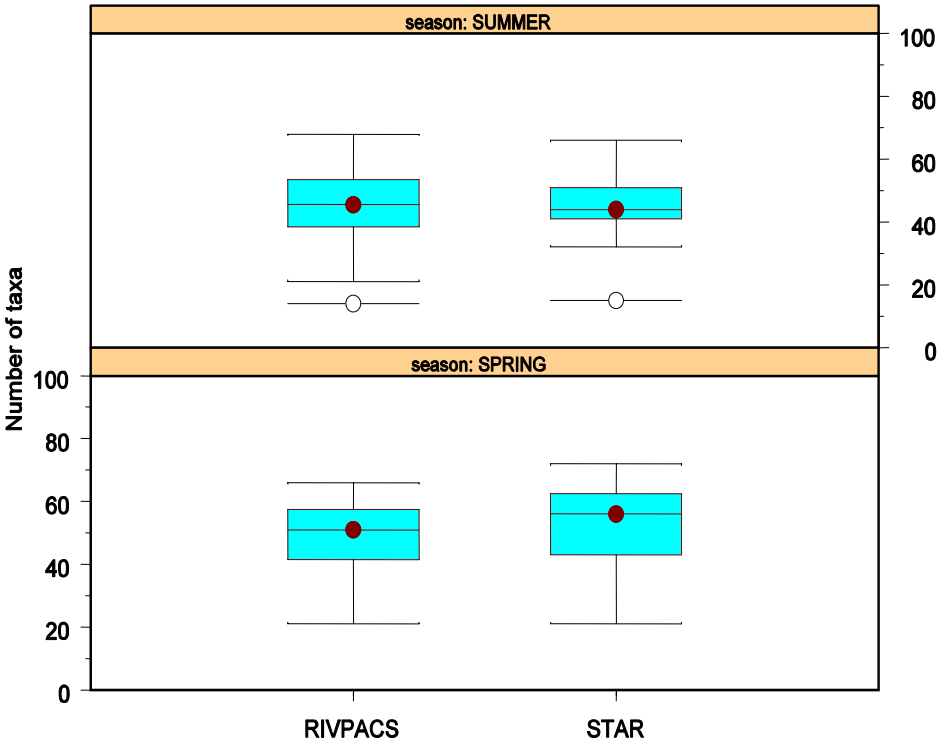
5 i


season: SUMMER
season: SPRING
Shredders
0
10
20
30
40
50
60
RIVPACS
STAR

5 j

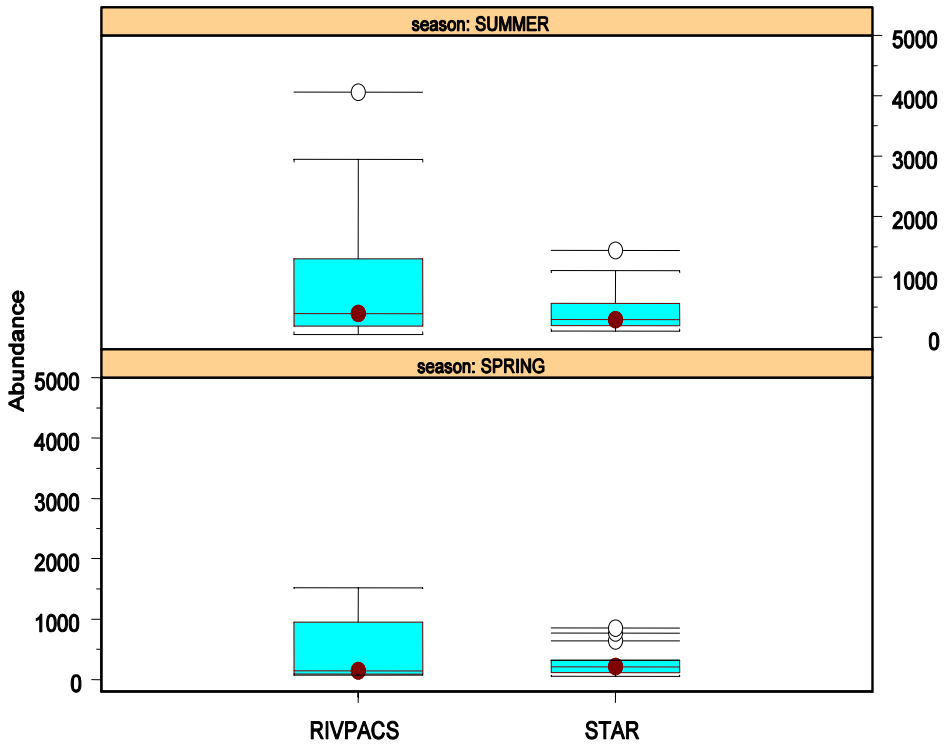

season: SUMMER
season: SPRING
Shannon - Wiener
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
RIVPACS
STAR

5 k

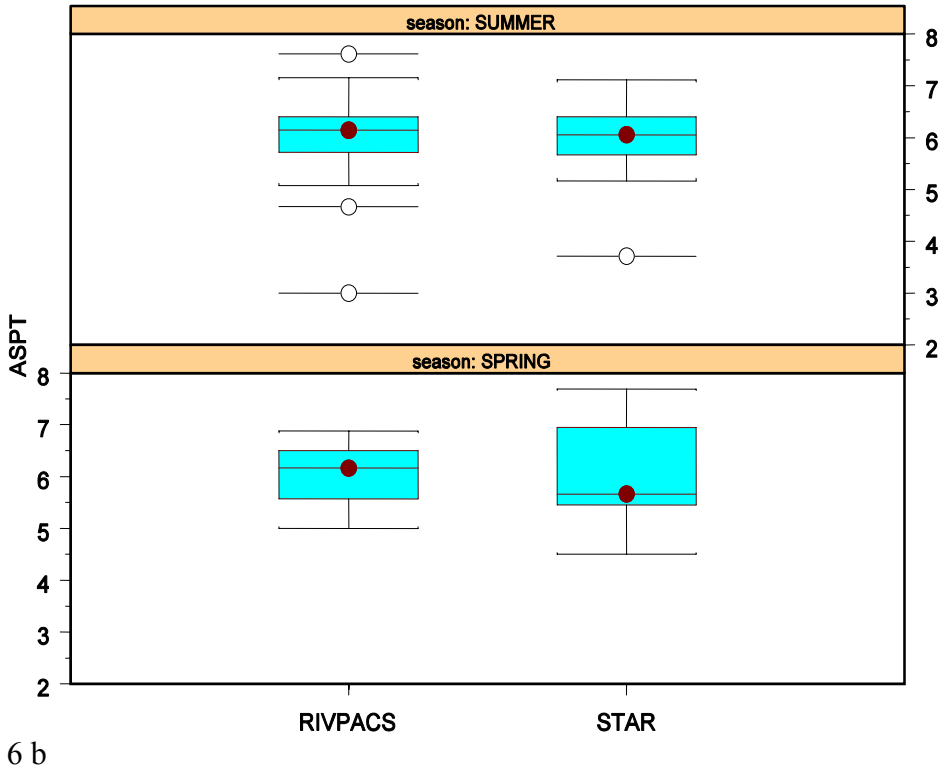

season: SUMMER
season: SPRING
Number of taxa
0
20
40
60
80
100
RIVPACS
STAR

5 1

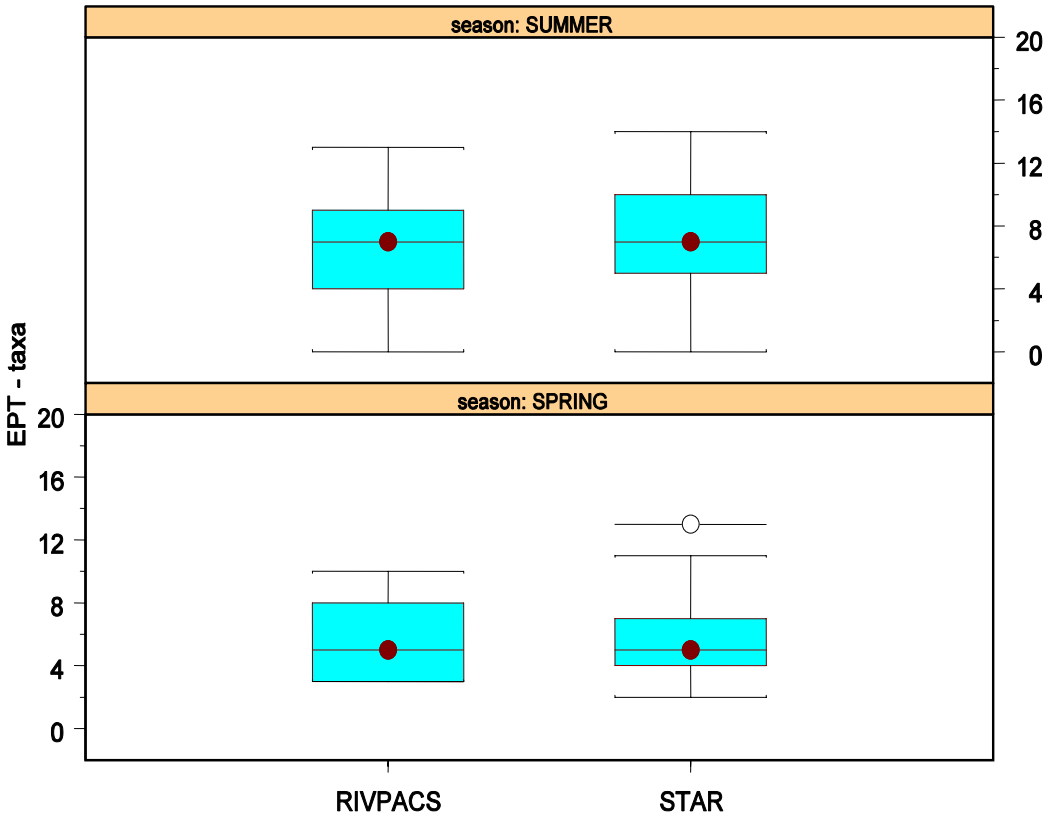
6 – Greece



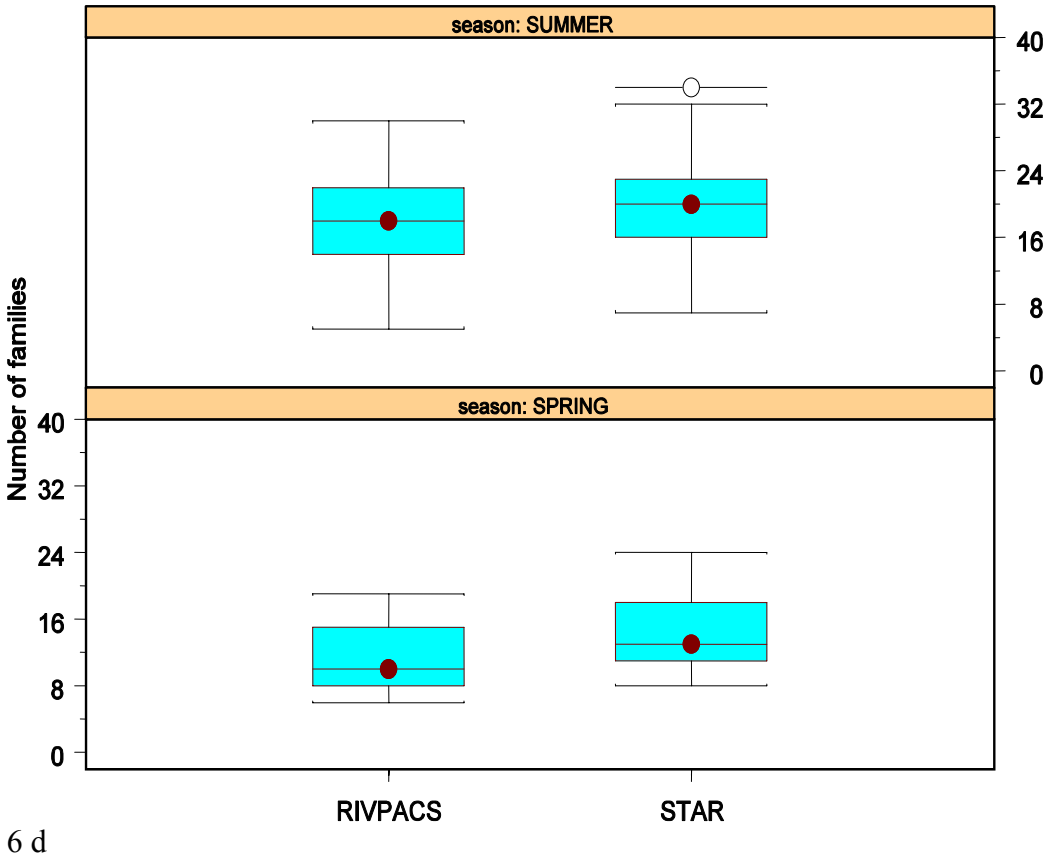
6 a



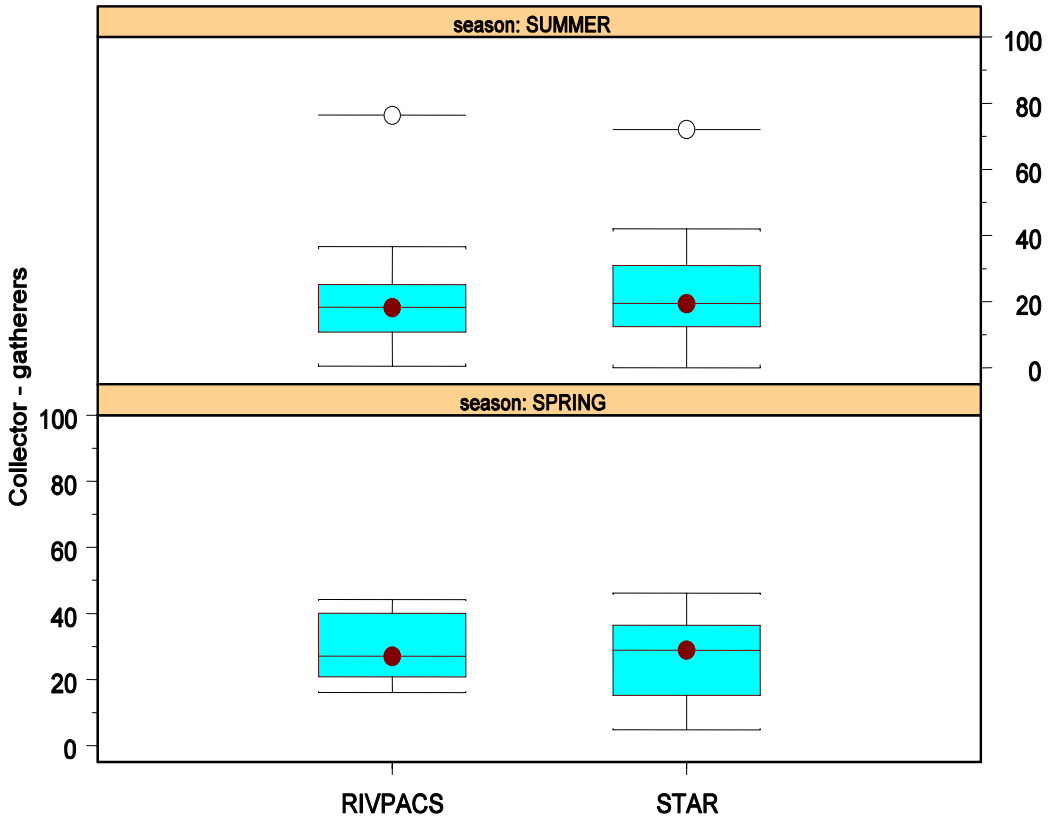
6 b


season: SUMMER
season: SPRING
EPT - taxa
0
4
8
12
16
20
RIVPACS
STAR

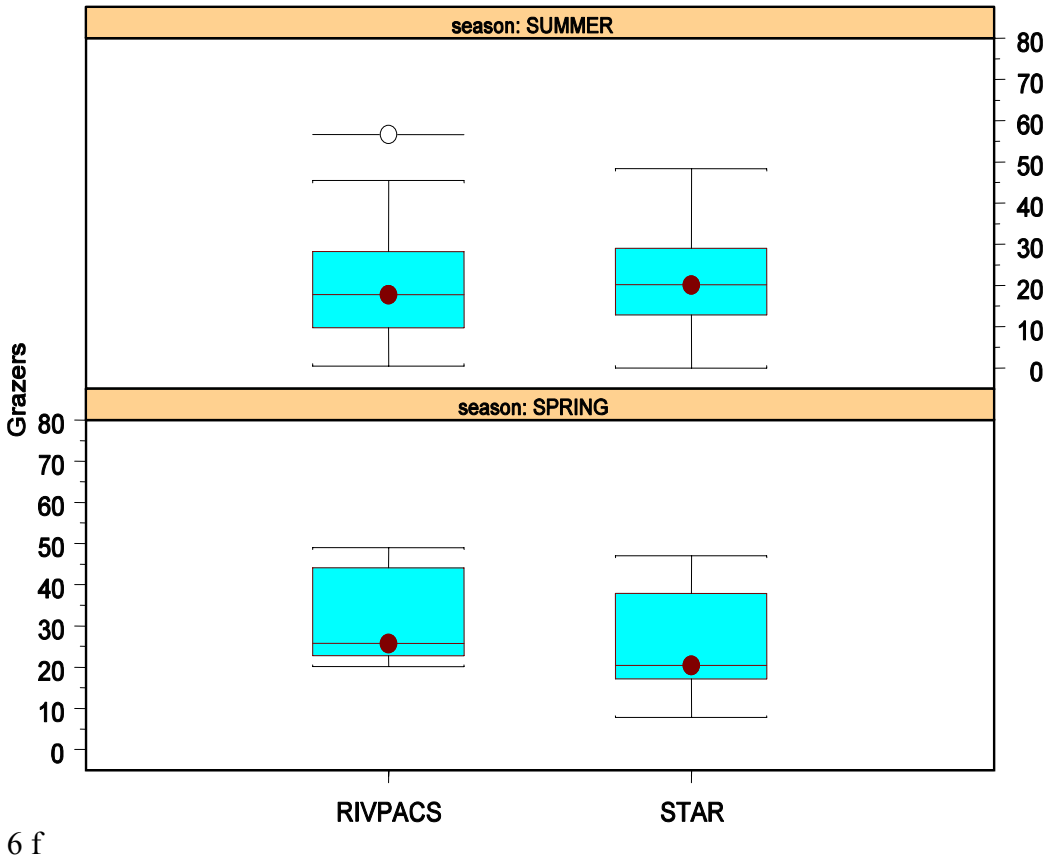
6 c


season: SUMMER
season: SPRING
Number of families
0
8
16
24
32
40
RIVPACS
STAR

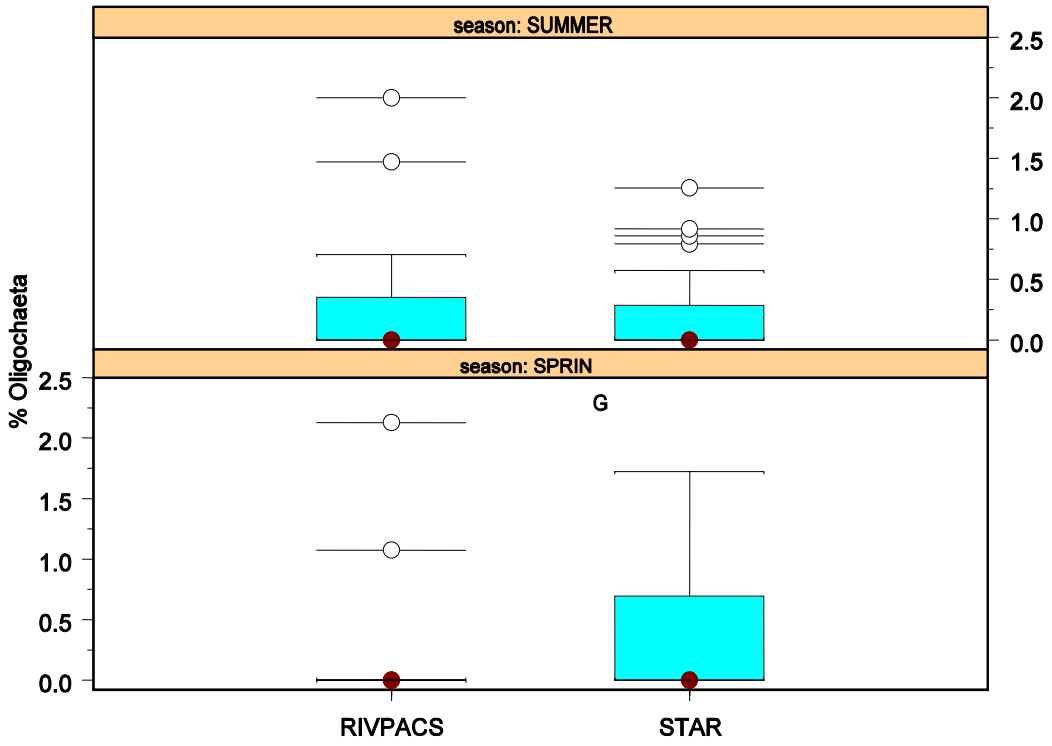
6 d


season: SUMMER
season: SPRING
Collector - gatherers
0
20
40
60
80
100
RIVPACS
STAR

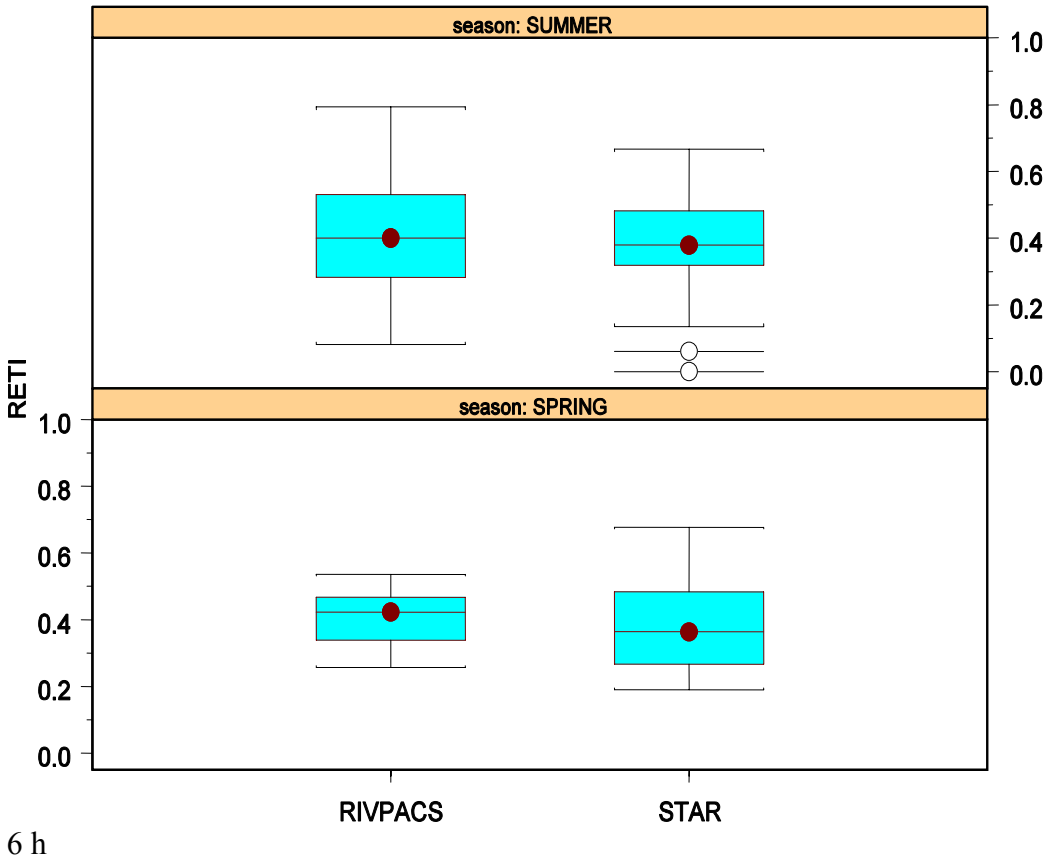
6 e


season: SUMMER
season: SPRING
Grazers
0
10
20
30
40
50
60
70
80
RIVPACS
STAR

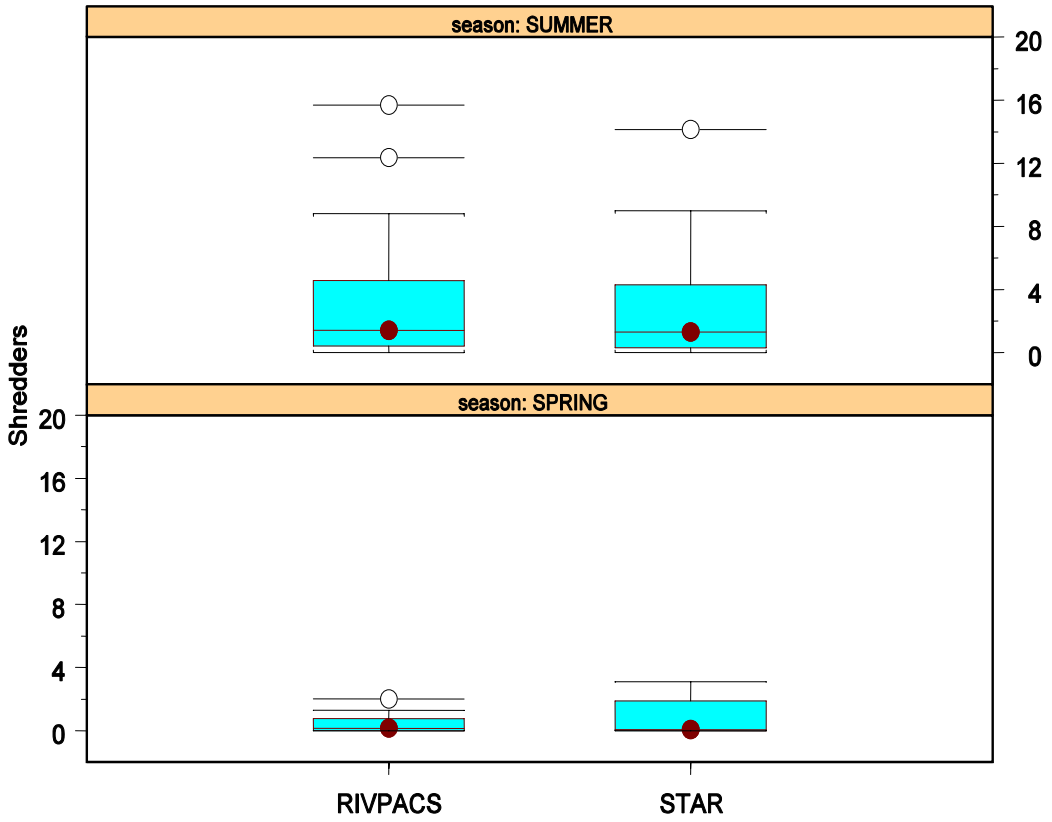
6 f


season: SUMMER
season: SPRIN
% Oligochaeta
G
2.5
2.0
1.5
1.0
0.5
0.0
RIVPACS
STAR

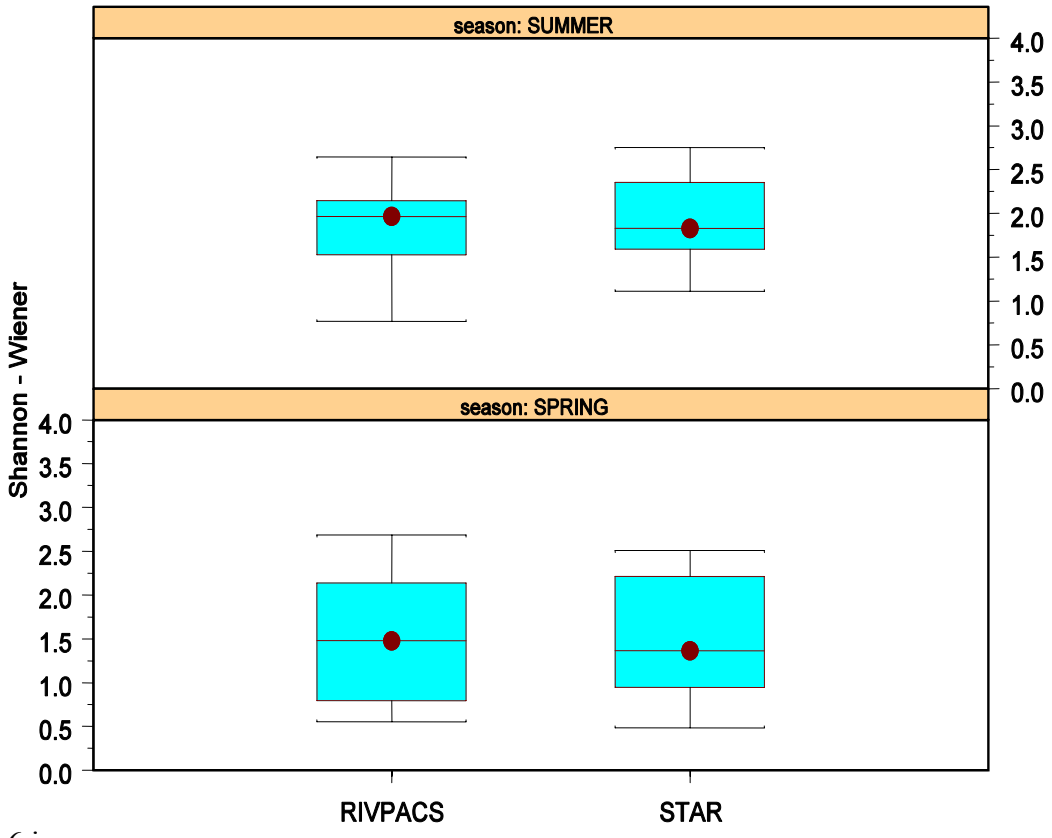
6 g


season: SUMMER
season: SPRING
RETI
1.0
0.8
0.6
0.4
0.2
0.0
RIVPACS
STAR

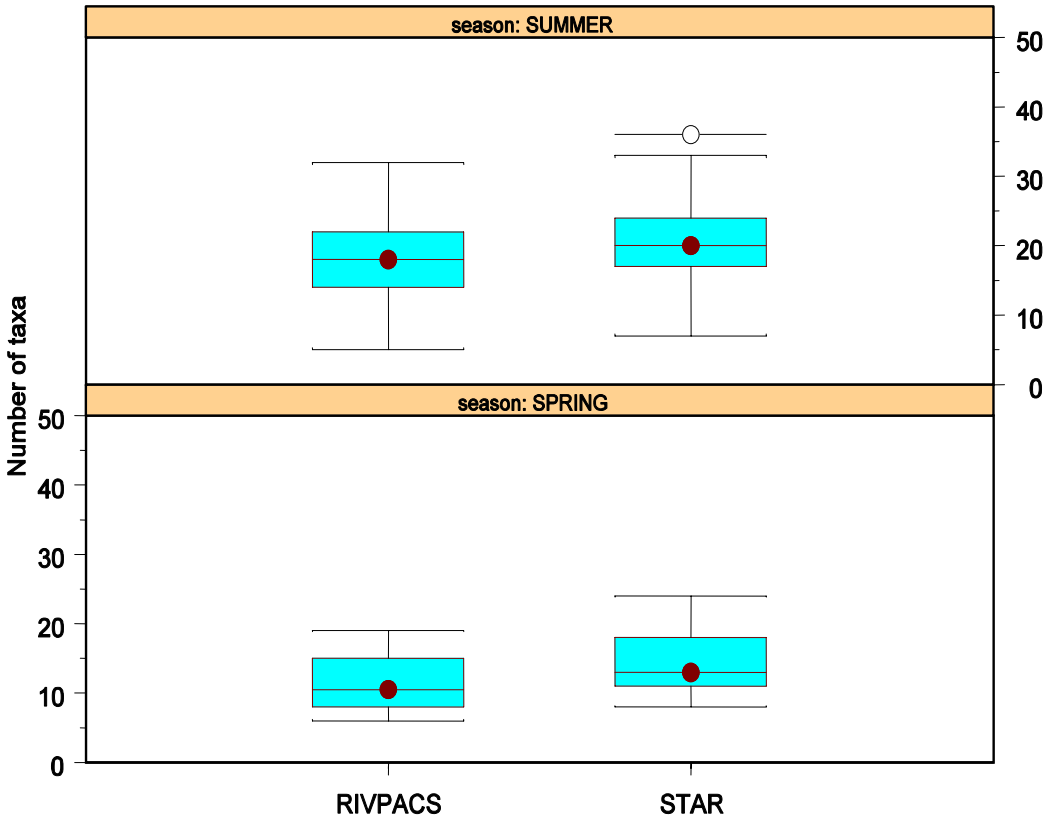
6 h


season: SUMMER
season: SPRING
Shredders
20
16
12
8
4
0
RIVPACS
STAR

6 i

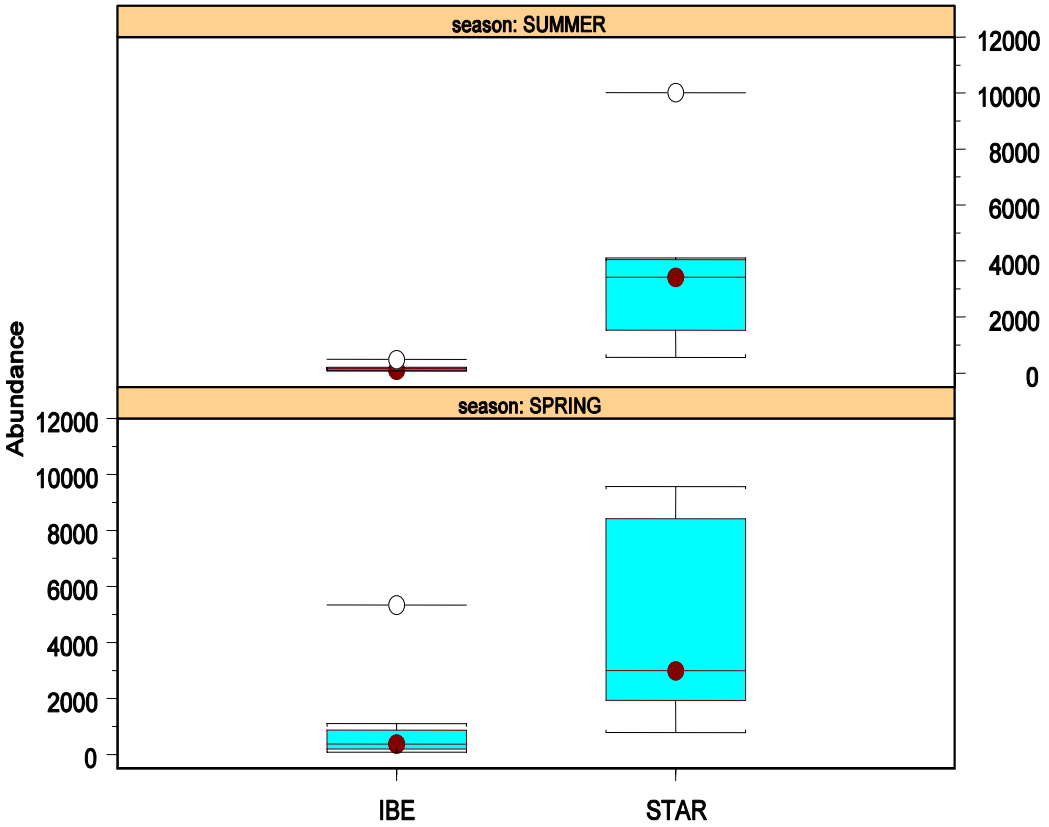

season: SUMMER
season: SPRING
Shannon - Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
RIVPACS
STAR

6 j

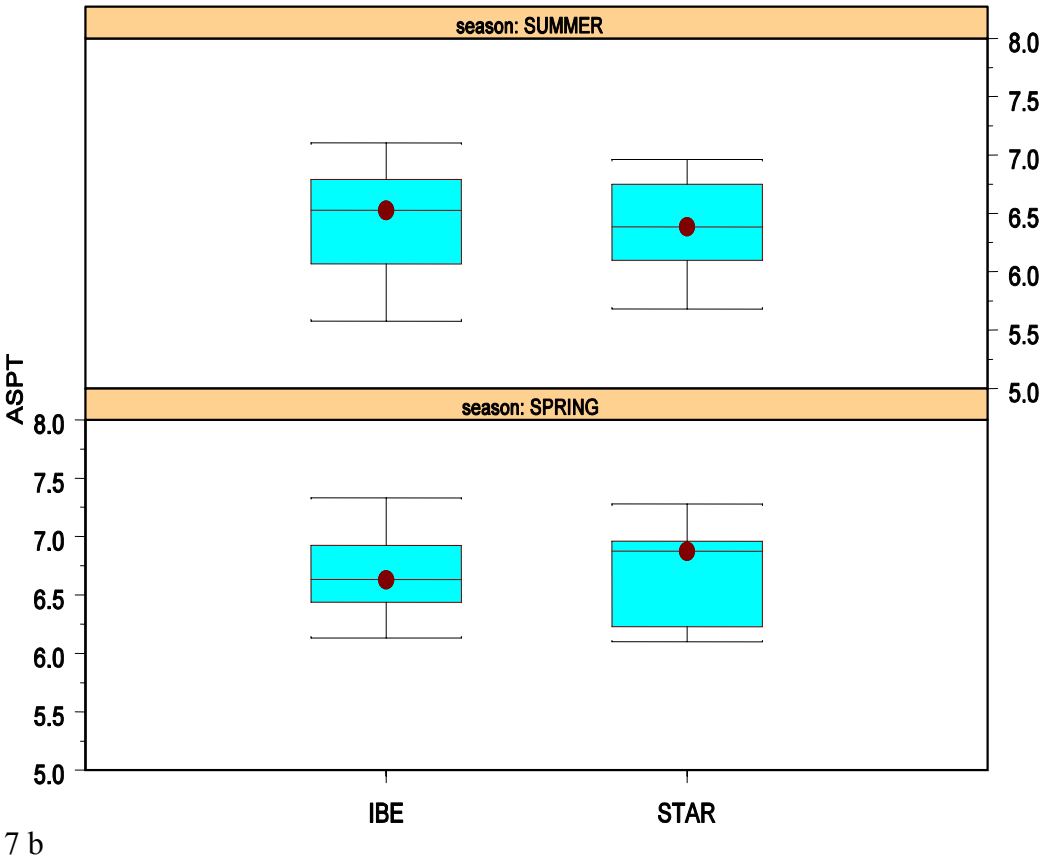

season: SUMMER
season: SPRING
Number of taxa
0
10
20
30
40
50
RIVPACS
STAR

6 k

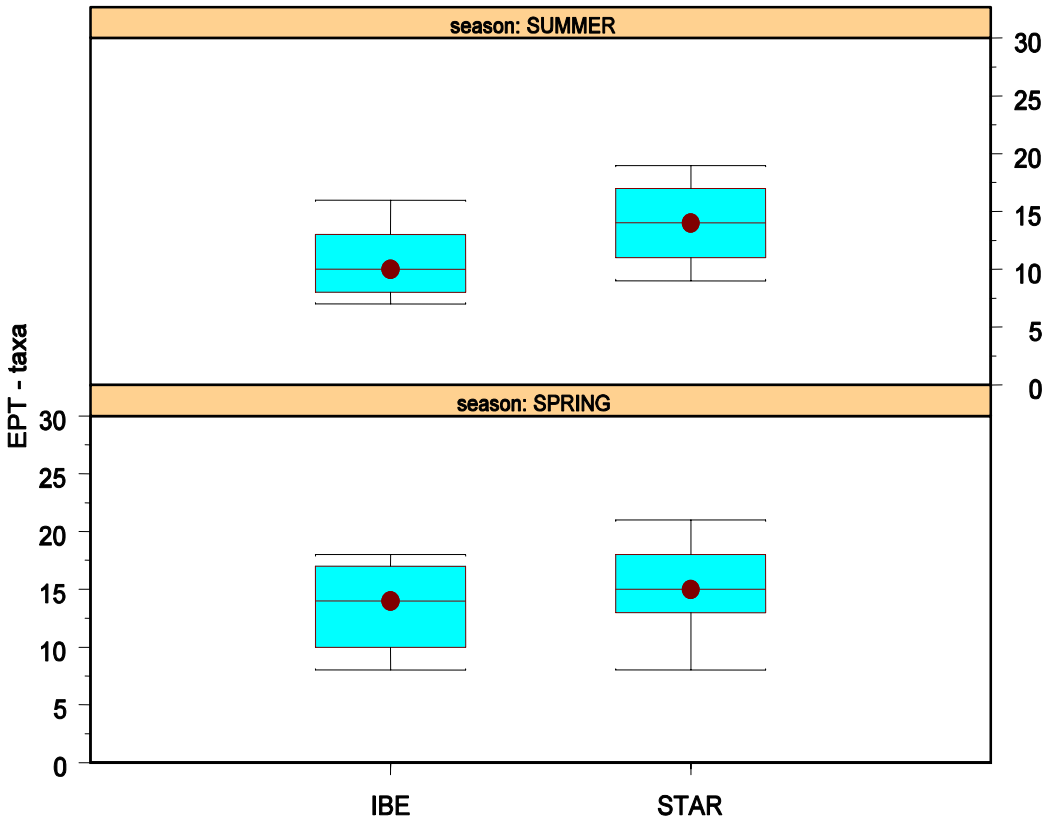
7 - Italy


season: SUMMER
season: SPRING
Abundance
12000
10000
8000
6000
4000
2000
0
IBE
STAR

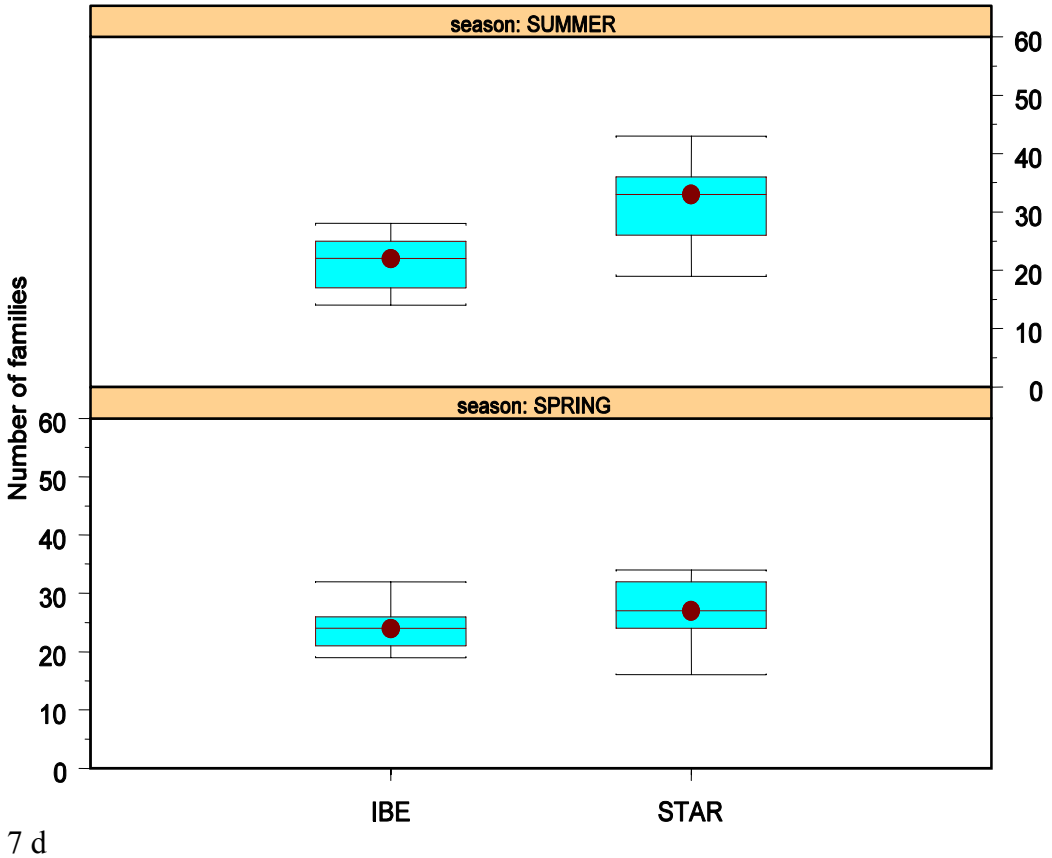
7 a


season: SUMMER
season: SPRING
ASPT
8.0
7.5
7.0
6.5
6.0
5.5
5.0
IBE
STAR

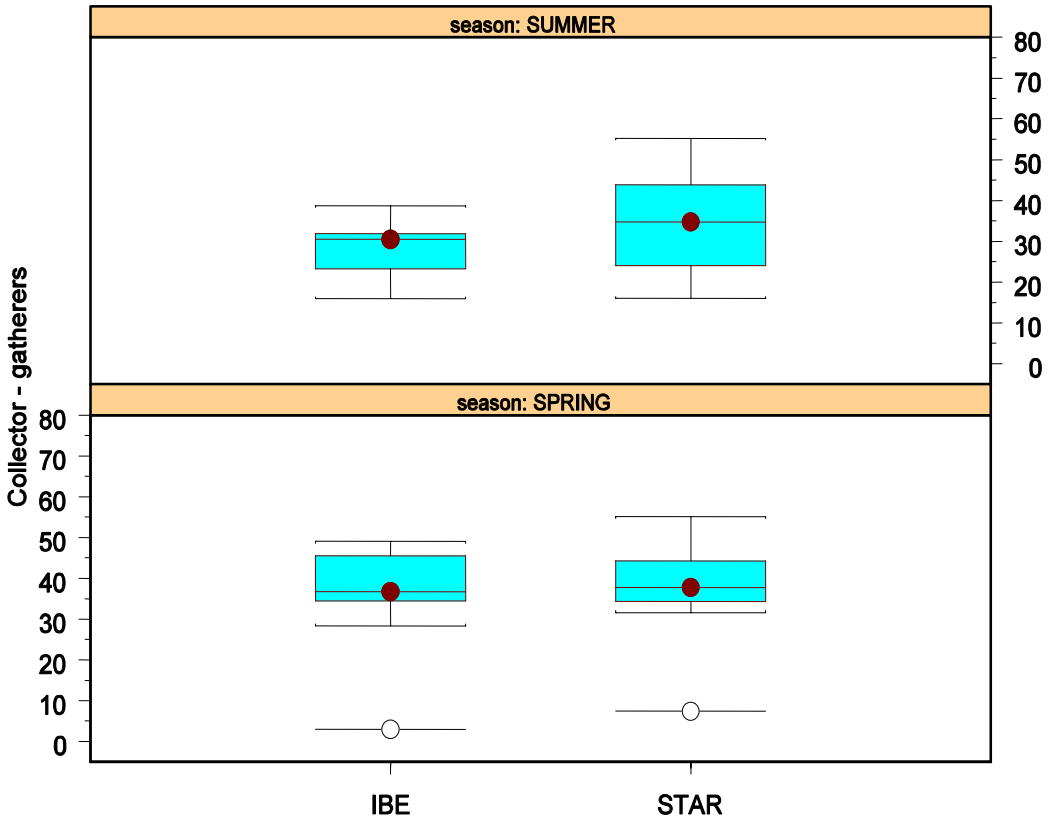
7 b


season: SUMMER
season: SPRING
EPT - taxa
0
5
10
15
20
25
30
IBE
STAR

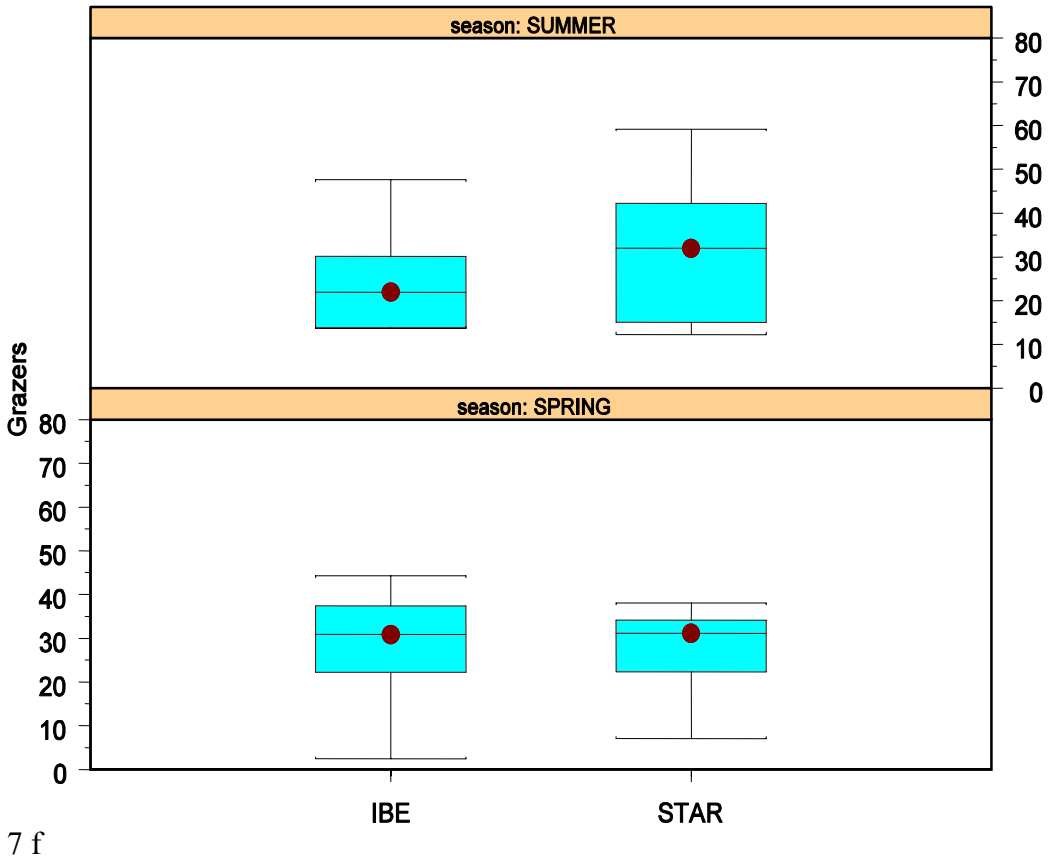
7 c


season: SUMMER
season: SPRING
Number of families
0
10
20
30
40
50
60
IBE
STAR

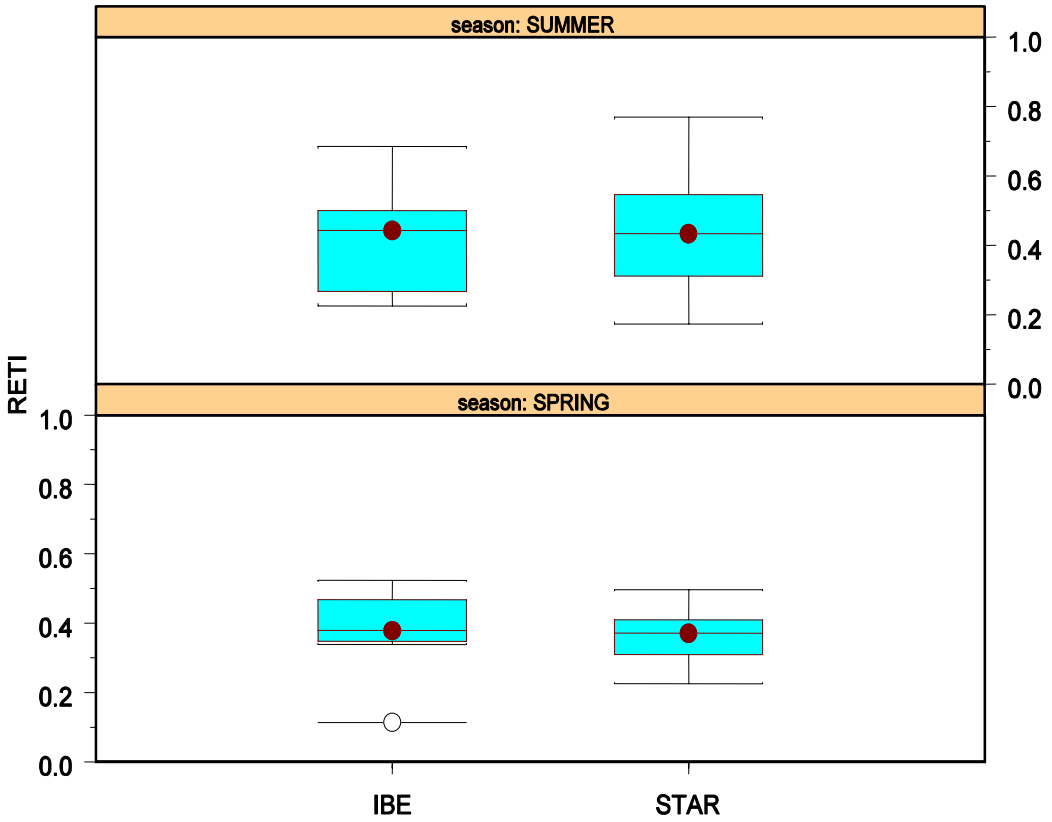
7 d


season: SUMMER
season: SPRING
Collector - gatherers
0
10
20
30
40
50
60
70
80
IBE
STAR

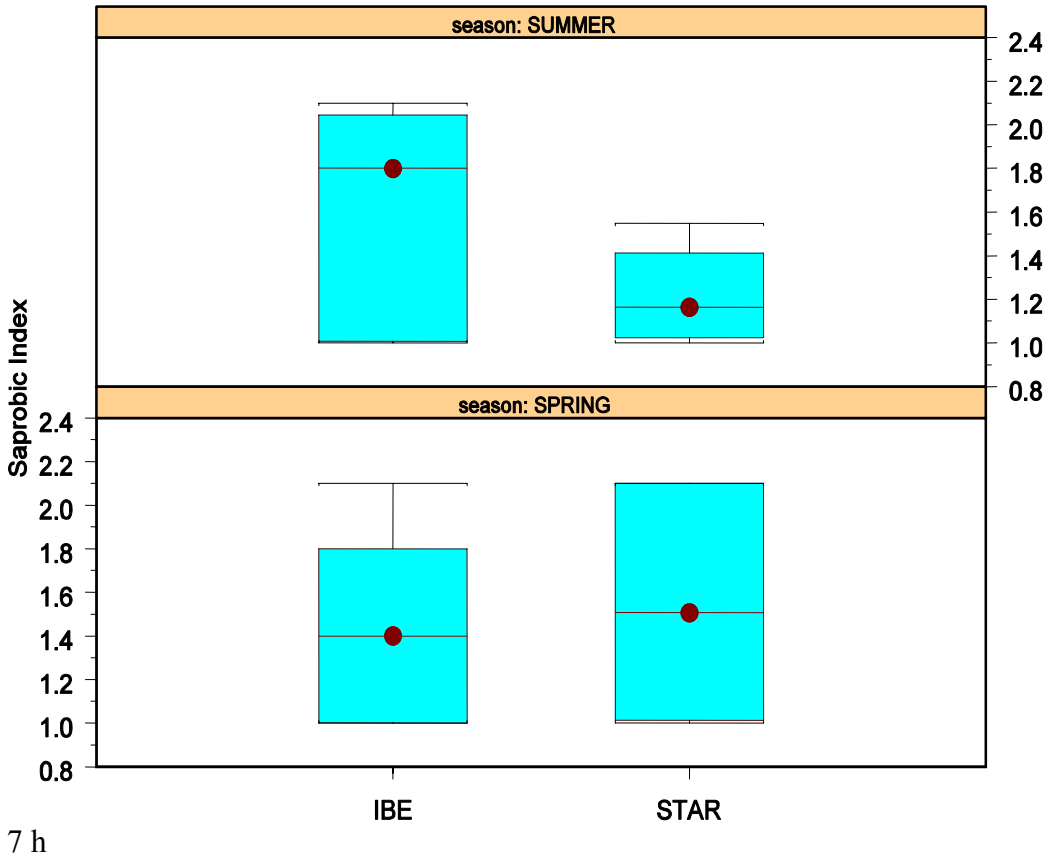
7 e


season: SUMMER
season: SPRING
Grazers
0
10
20
30
40
50
60
70
80
IBE
STAR

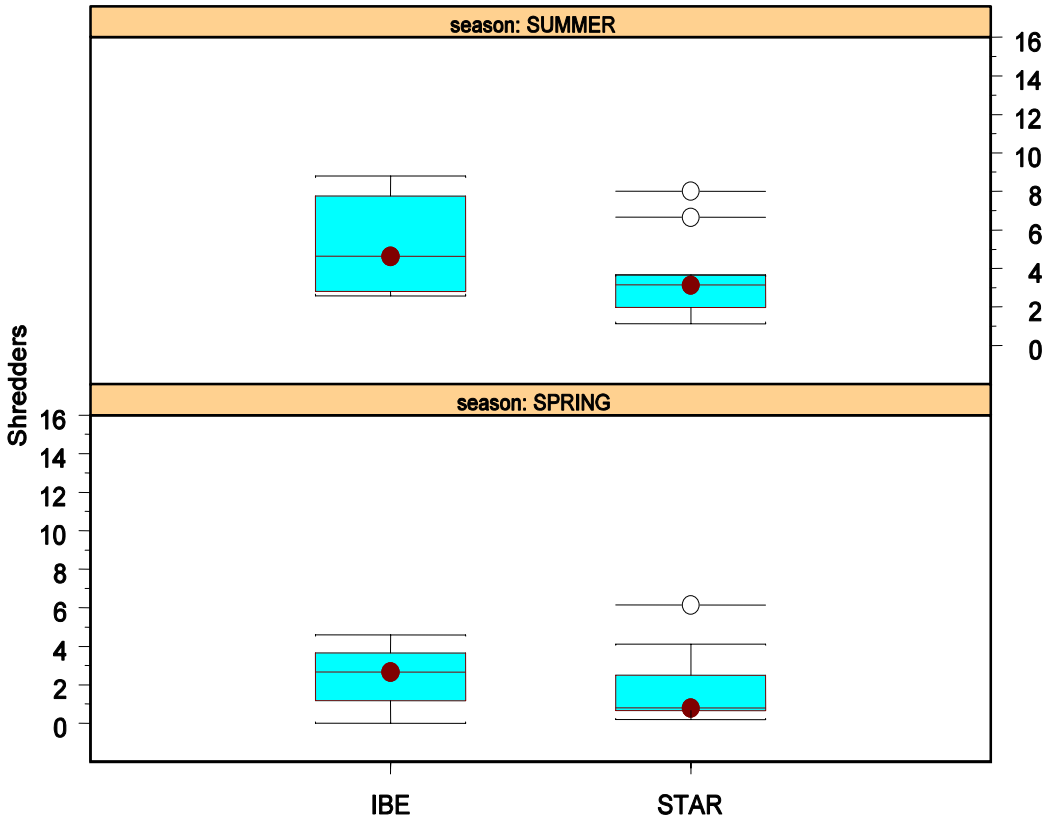
7 f


season: SUMMER
season: SPRING
RETI
1.0
0.8
0.6
0.4
0.2
0.0
IBE
STAR

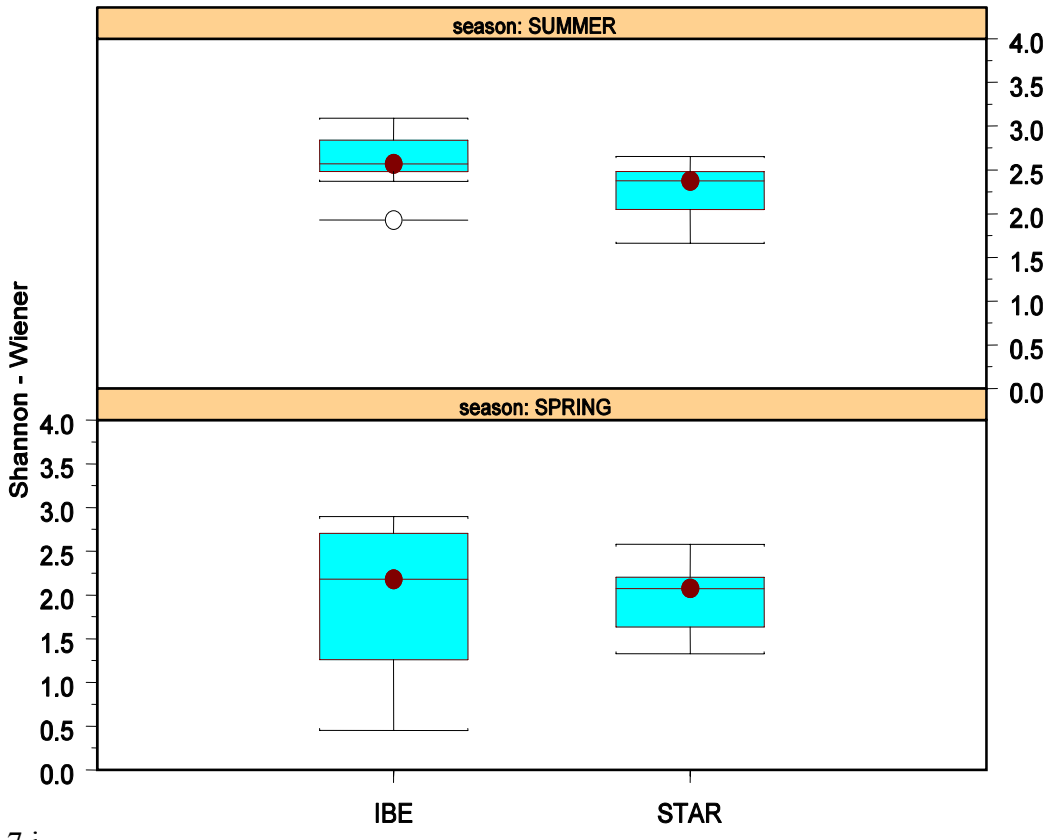
7 g


season: SUMMER
season: SPRING
Saprobic Index
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
IBE
STAR

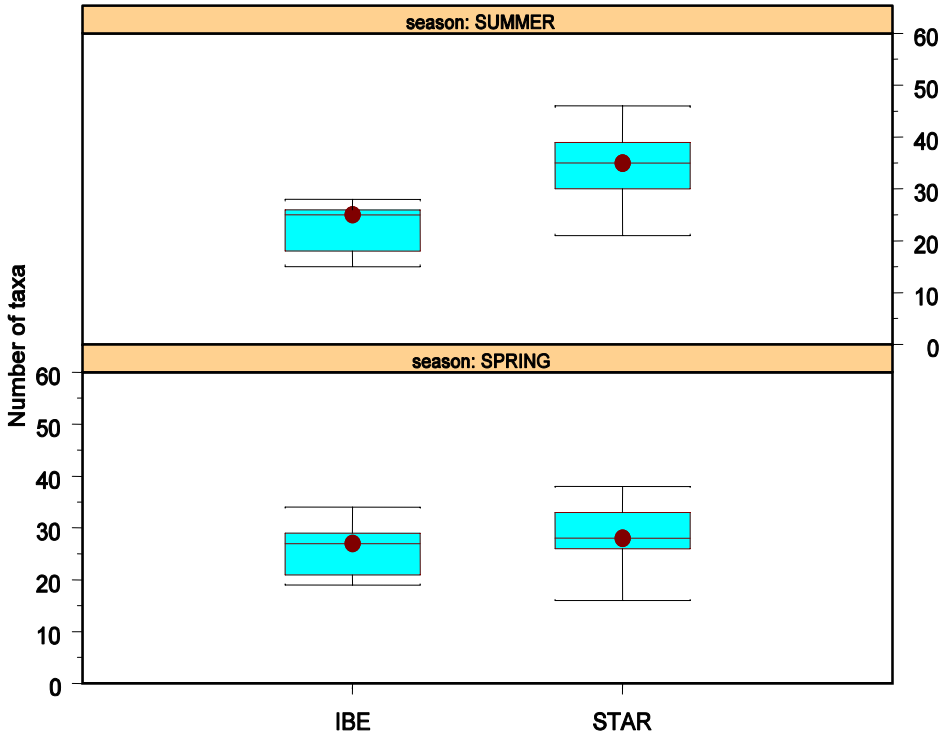
7 h


season: SUMMER
season: SPRING
Shredders
0
2
4
6
8
10
12
14
16
IBE
STAR

7 i

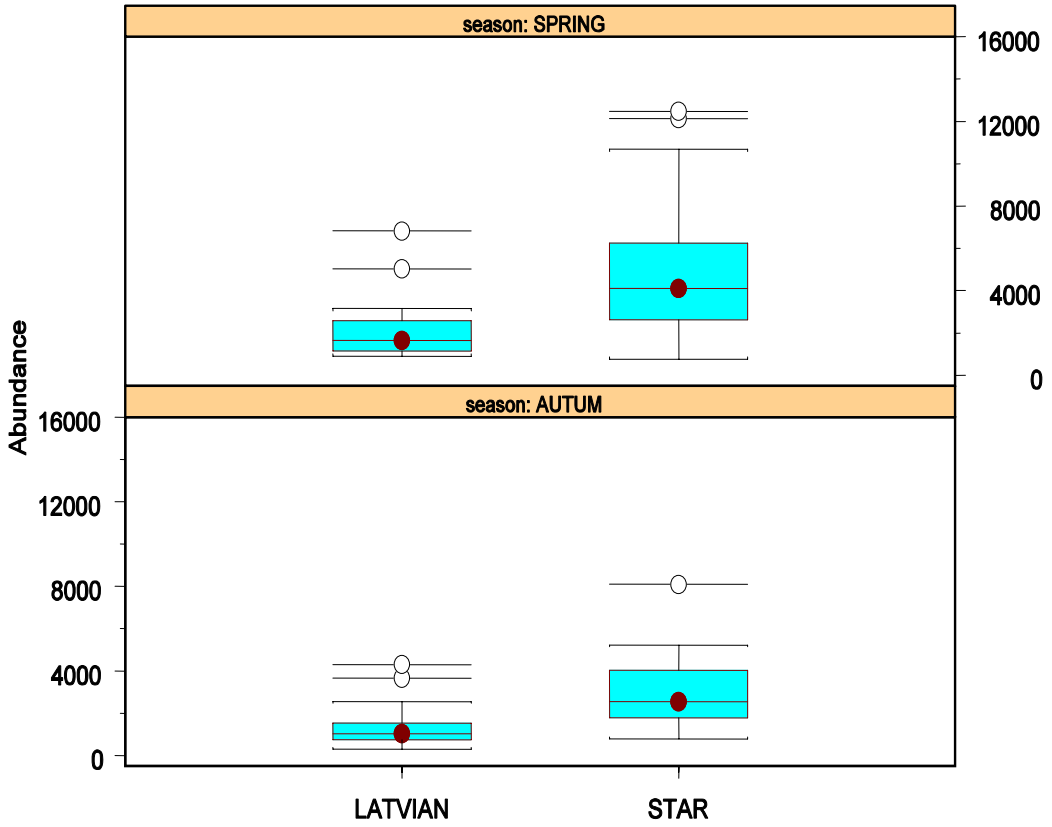

season: SUMMER
season: SPRING
Shannon - Wiener
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
IBE
STAR

7 j

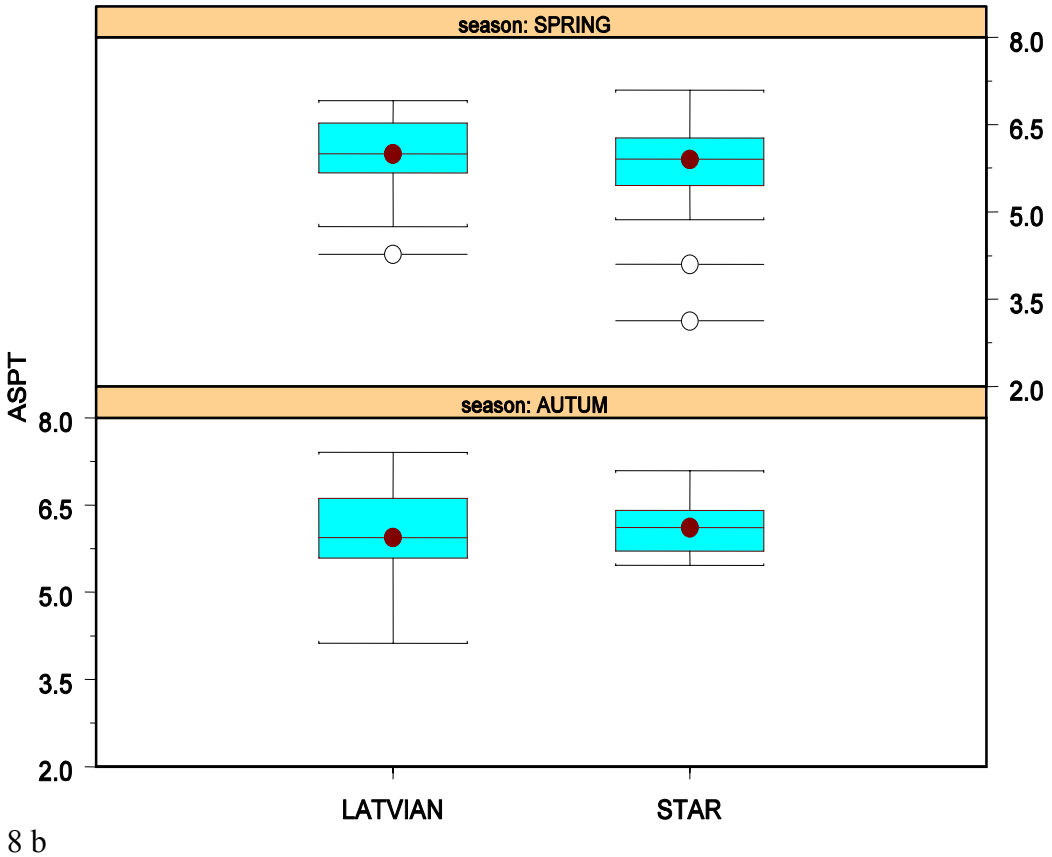

season: SUMMER
season: SPRING
Number of taxa
0
10
20
30
40
50
60
IBE
STAR

7 k

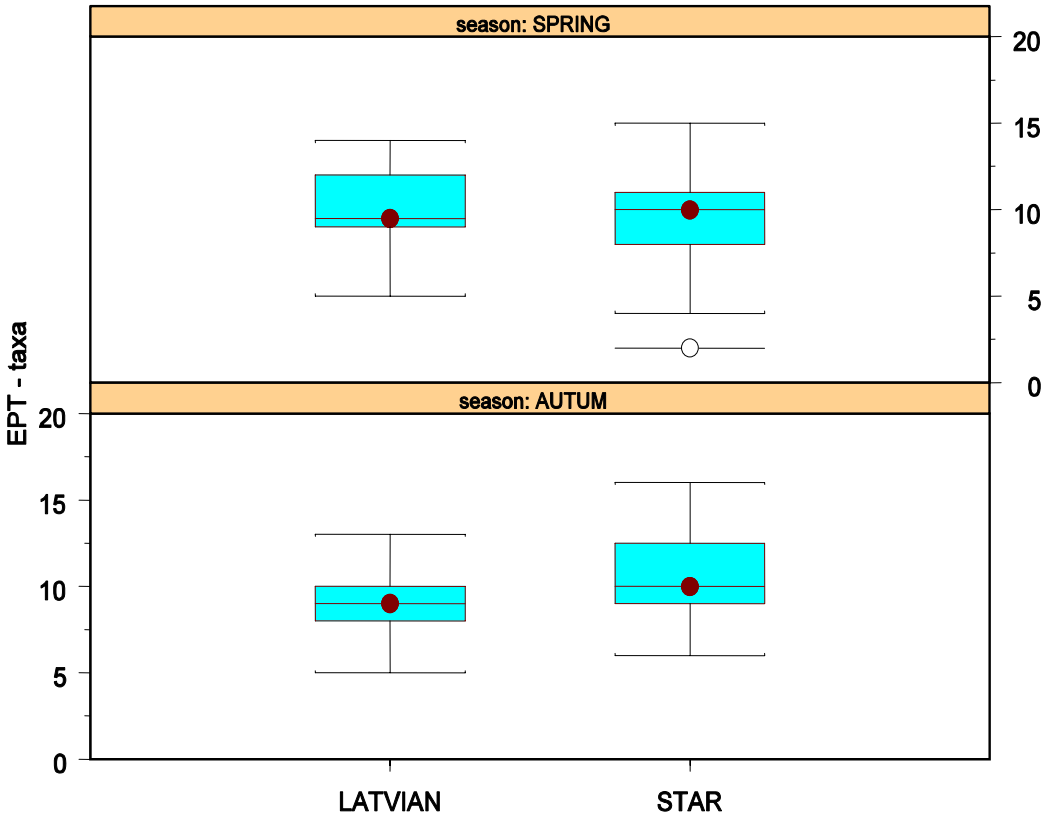
8 - Latvia



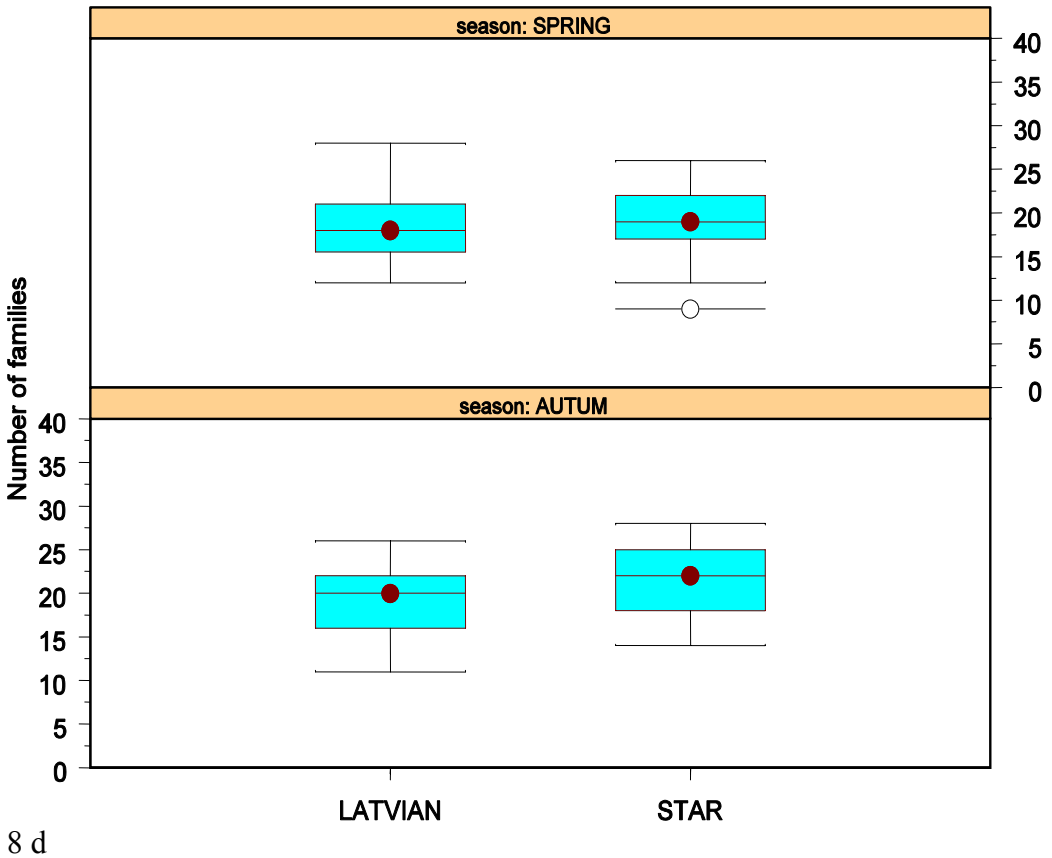
8 a



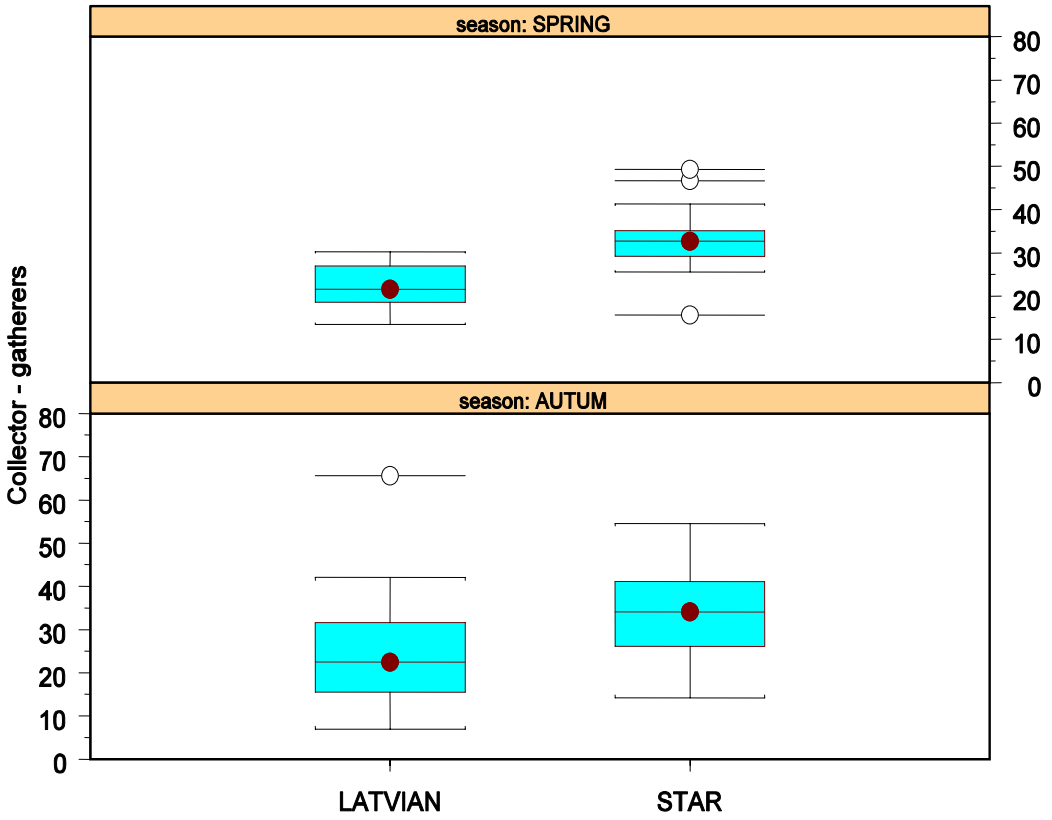
8 b


season: SPRING
season: AUTUM
EPT - taxa
0
5
10
15
20
LATVIAN
STAR

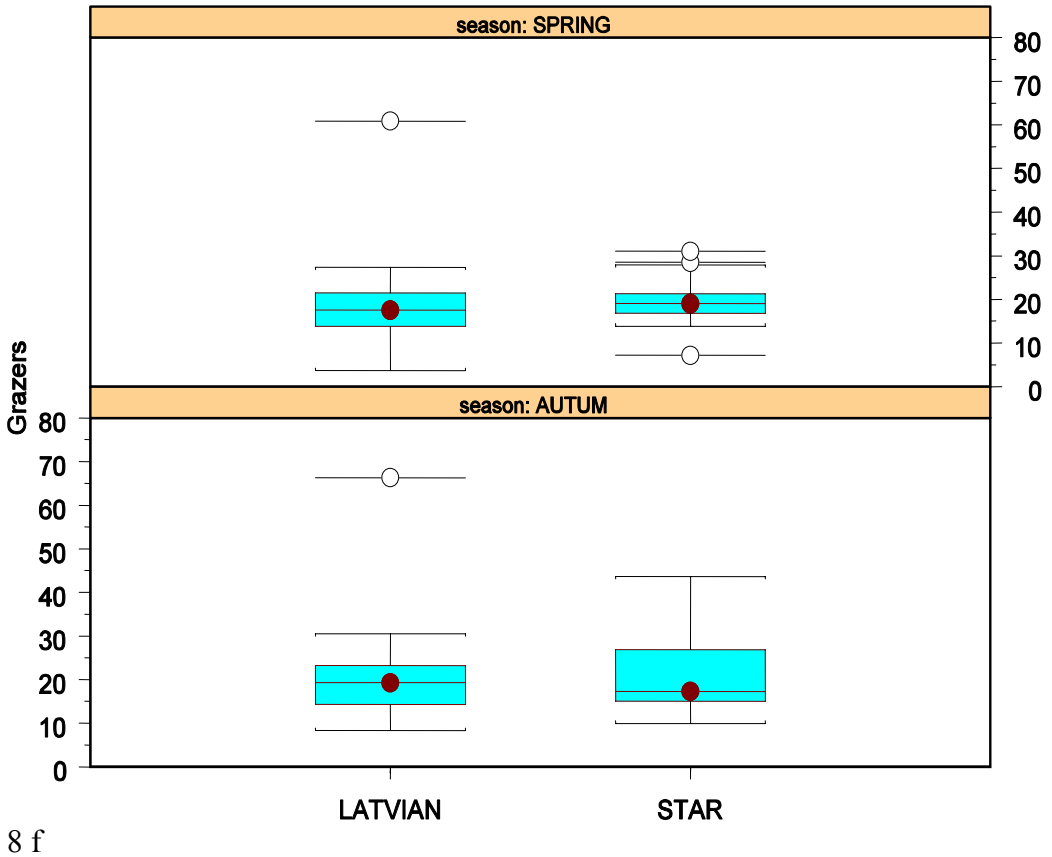
18c


season: SPRING
season: AUTUM
Number of families
0
5
10
15
20
25
30
35
40
LATVIAN
STAR

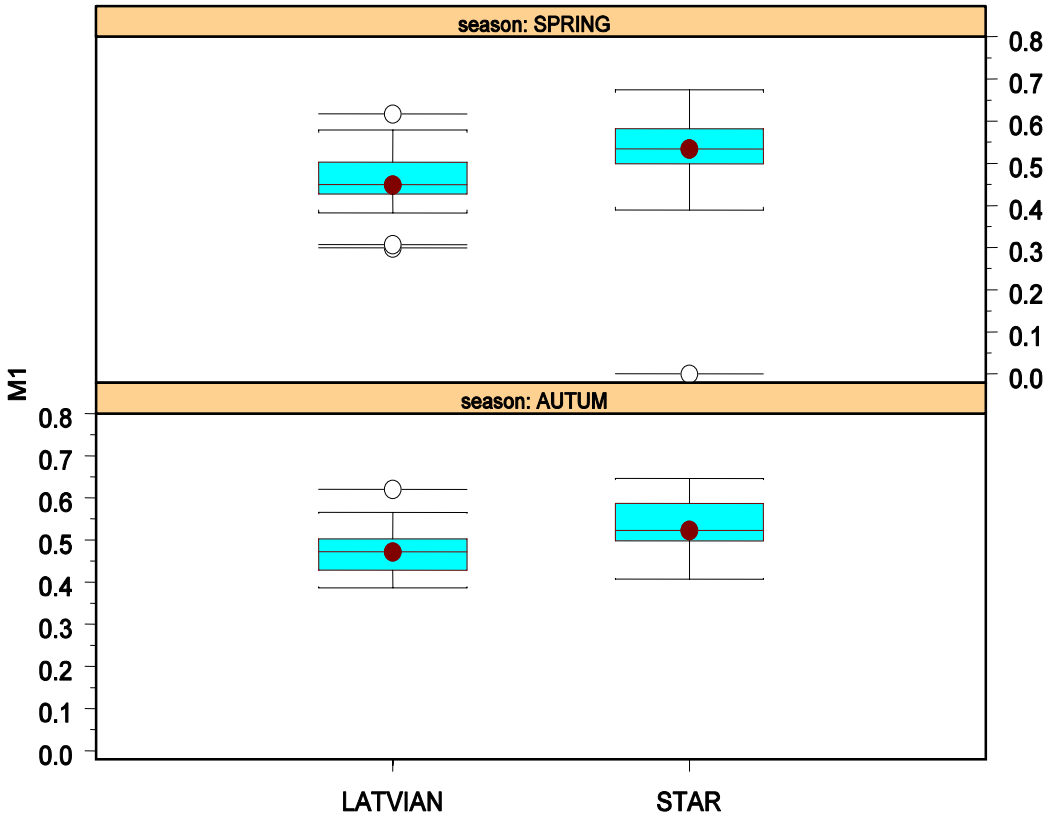
8 d


season: SPRING
season: AUTUM
Collector - gatherers
0
10
20
30
40
50
60
70
80
LATVIAN
STAR

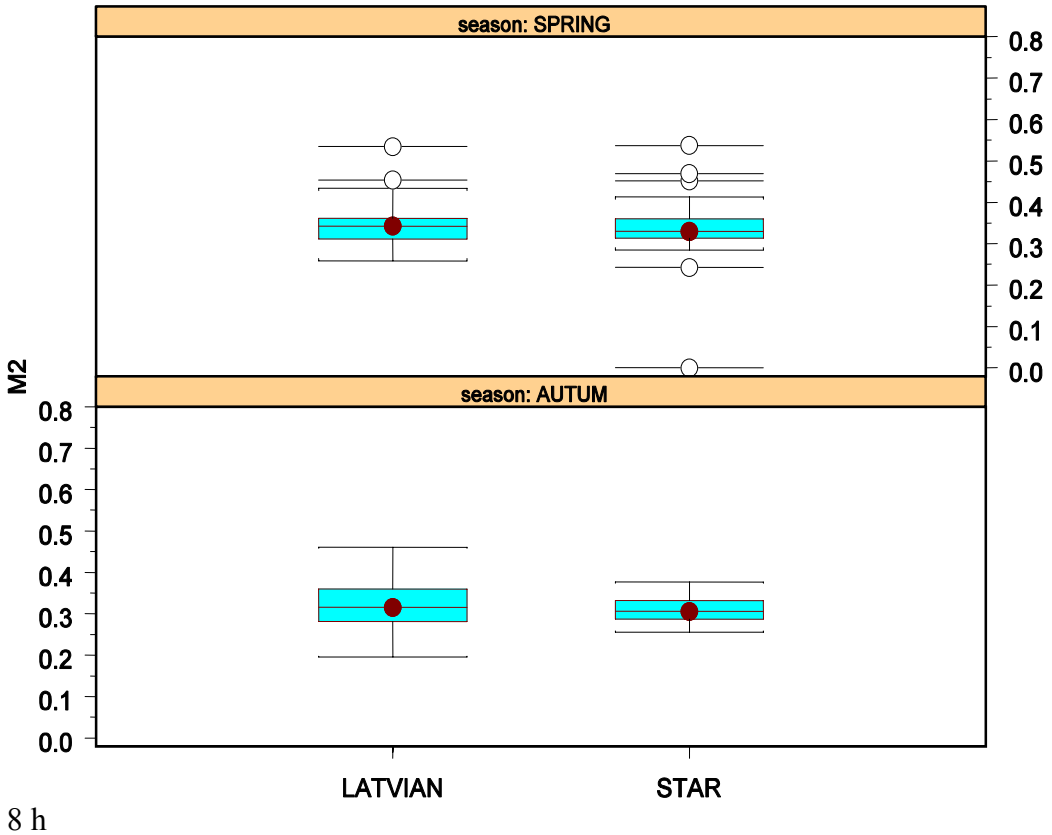
8 e


season: SPRING
season: AUTUM
Grazers
0
10
20
30
40
50
60
70
80
LATVIAN
STAR

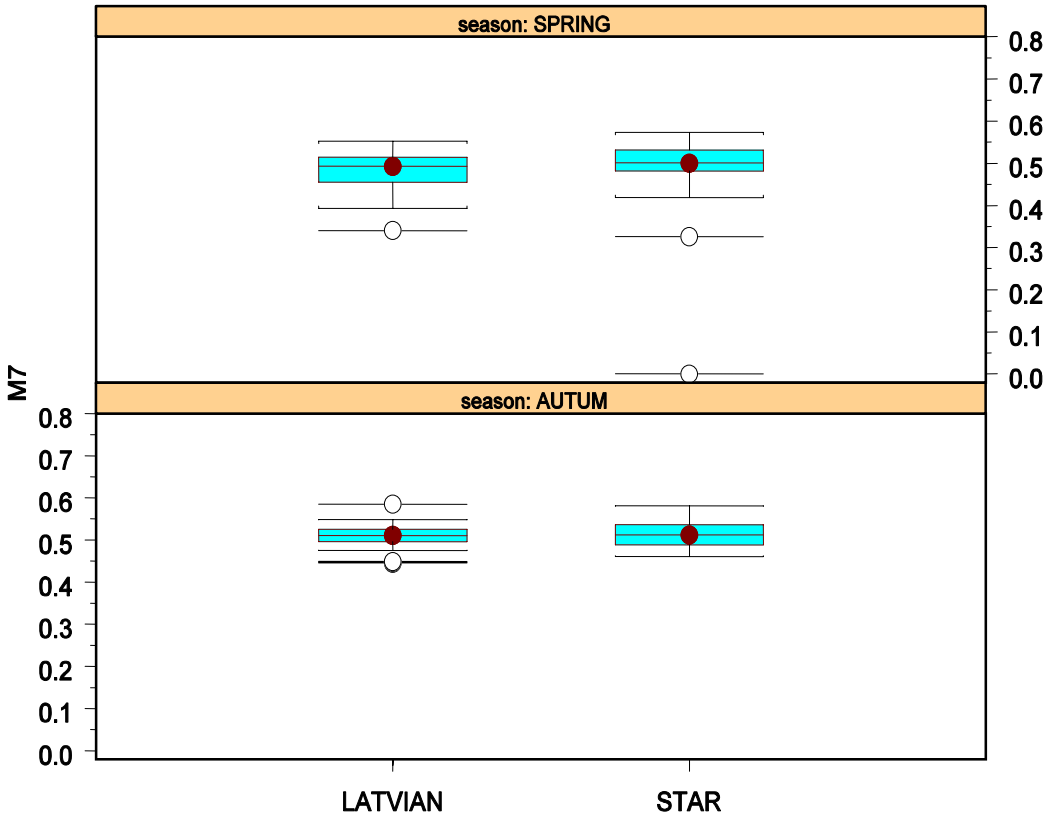
8 f


season: SPRING
season: AUTUM
M1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
LATVIAN
STAR

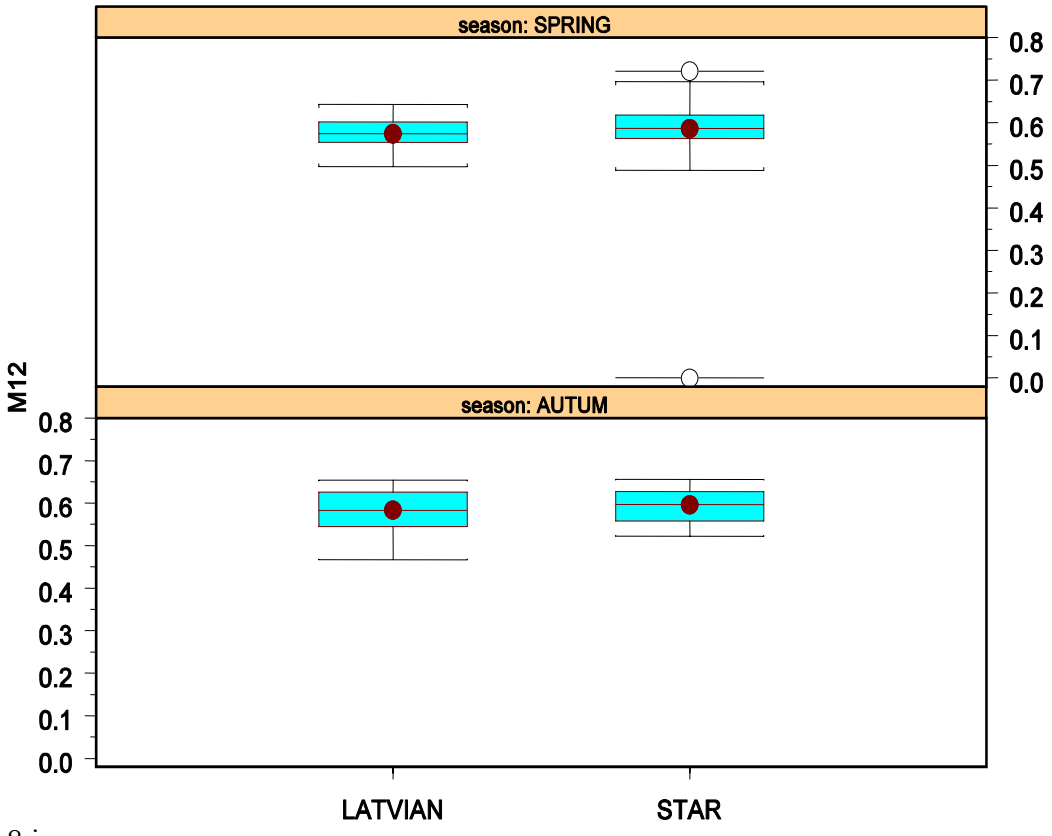
8 g


season: SPRING
season: AUTUM
M2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
LATVIAN
STAR

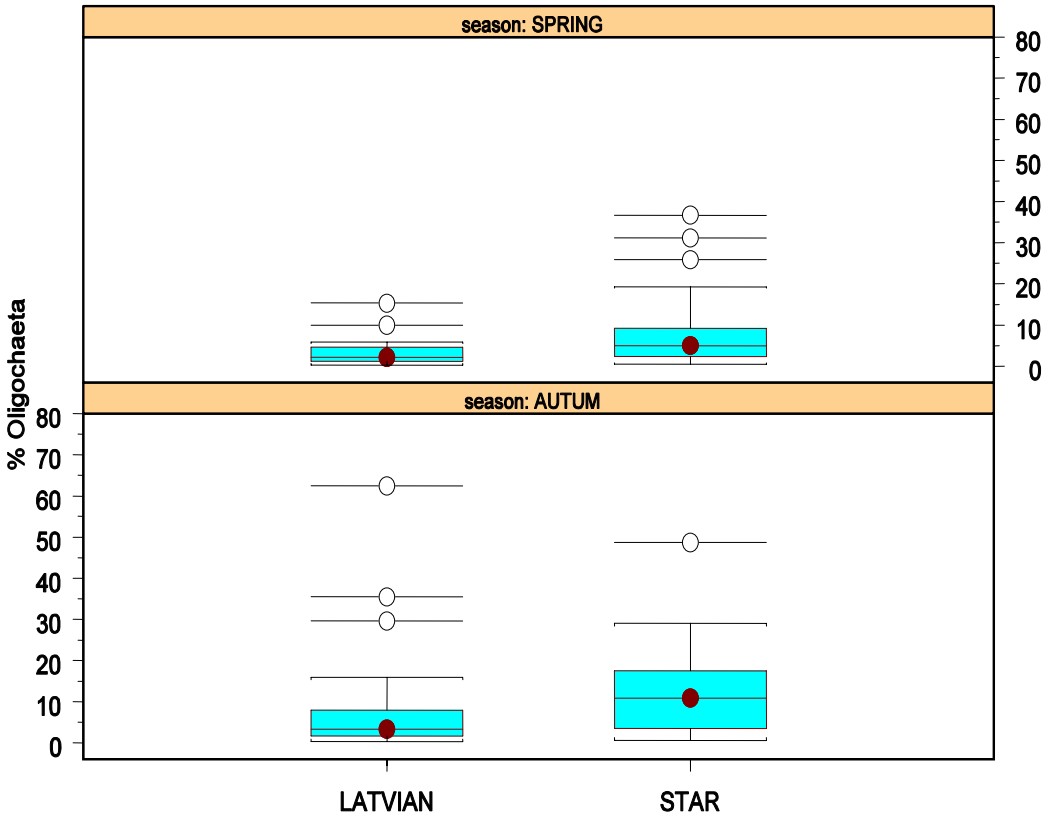
8 h


season: SPRING
season: AUTUM
M7
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
LATVIAN
STAR

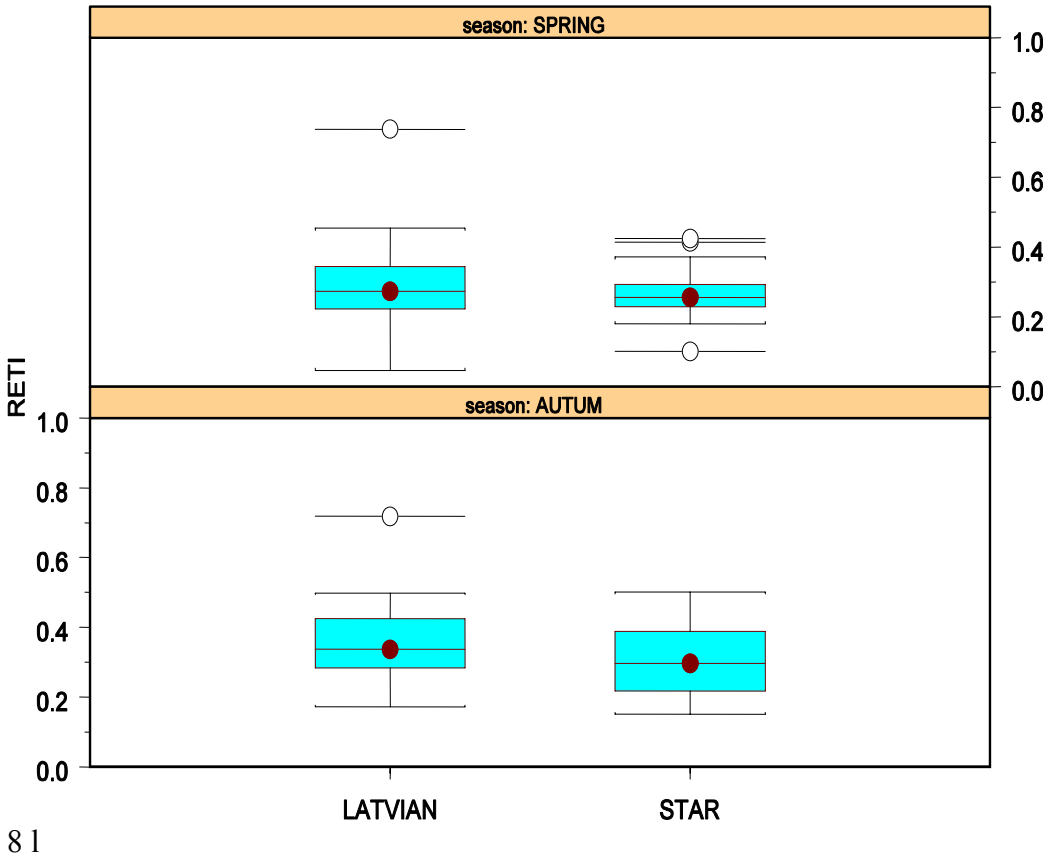
8 i


season: SPRING
season: AUTUM
M12
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
LATVIAN
STAR

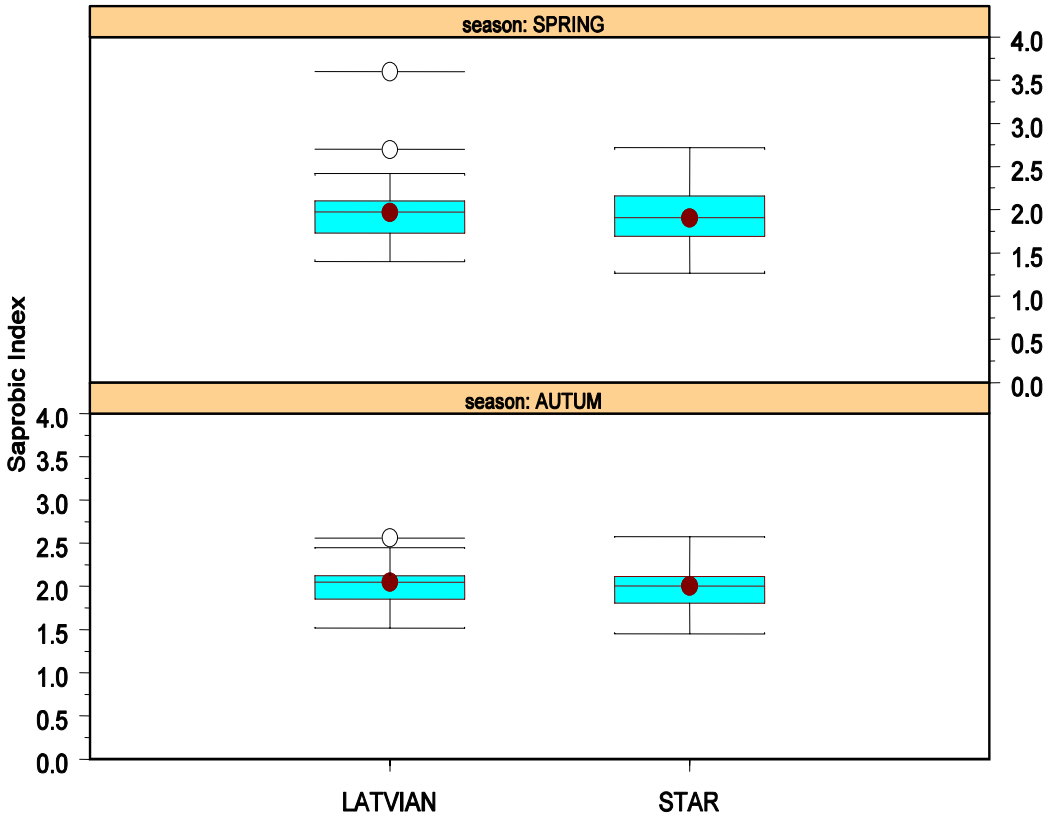
8 j


season: SPRING
season: AUTUM
% Oligochaeta
0
10
20
30
40
50
60
70
80
LATVIAN
STAR

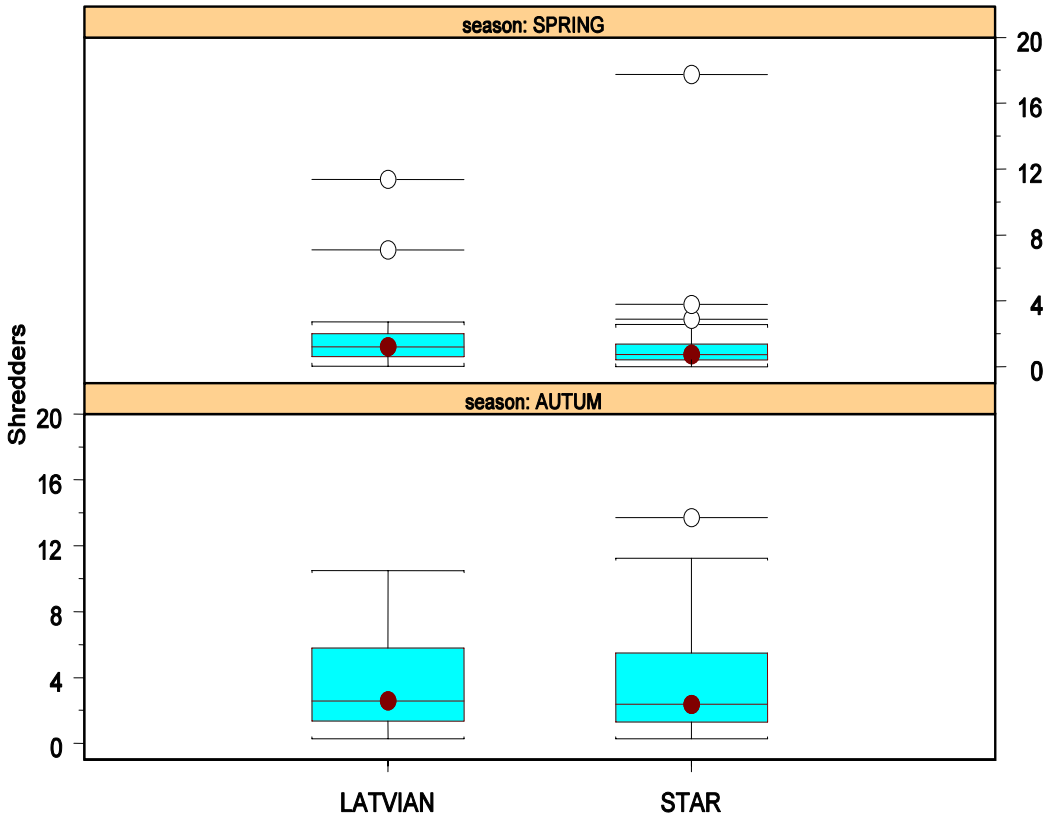
8 k


season: SPRING
season: AUTUM
RETI
0.0
0.2
0.4
0.6
0.8
1.0
LATVIAN
STAR

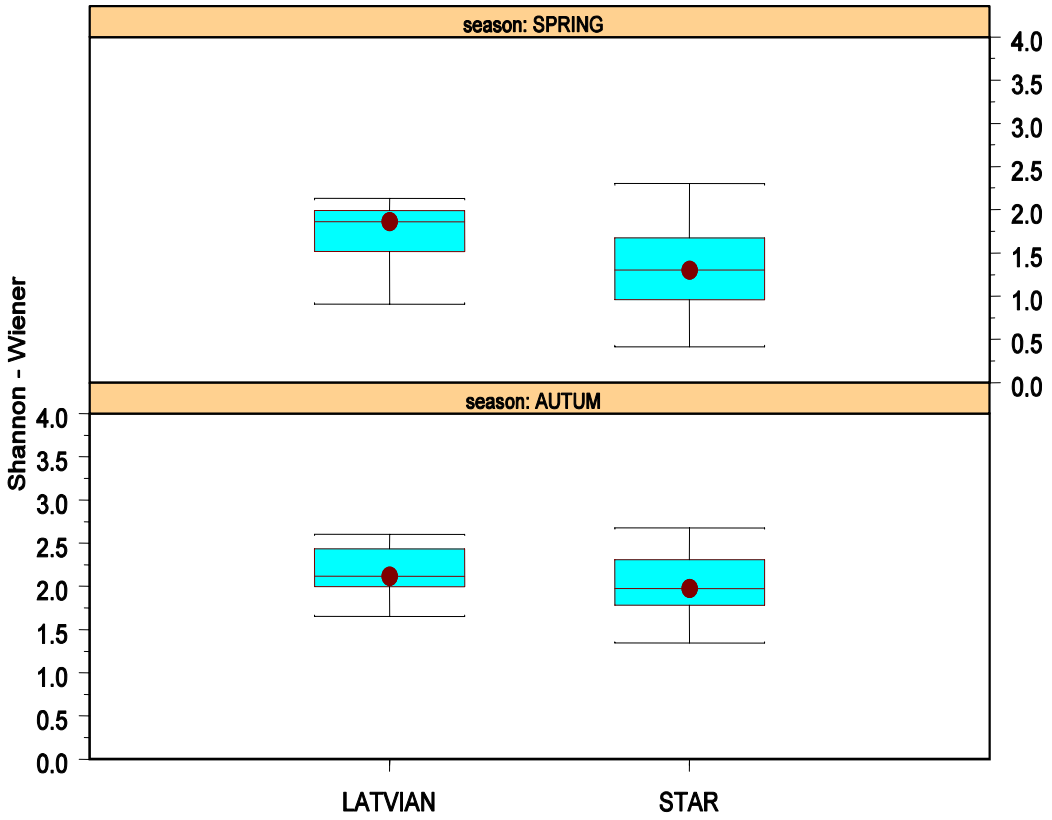
8 l


season: SPRING
season: AUTUM
Saprobic Index
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
LATVIAN
STAR

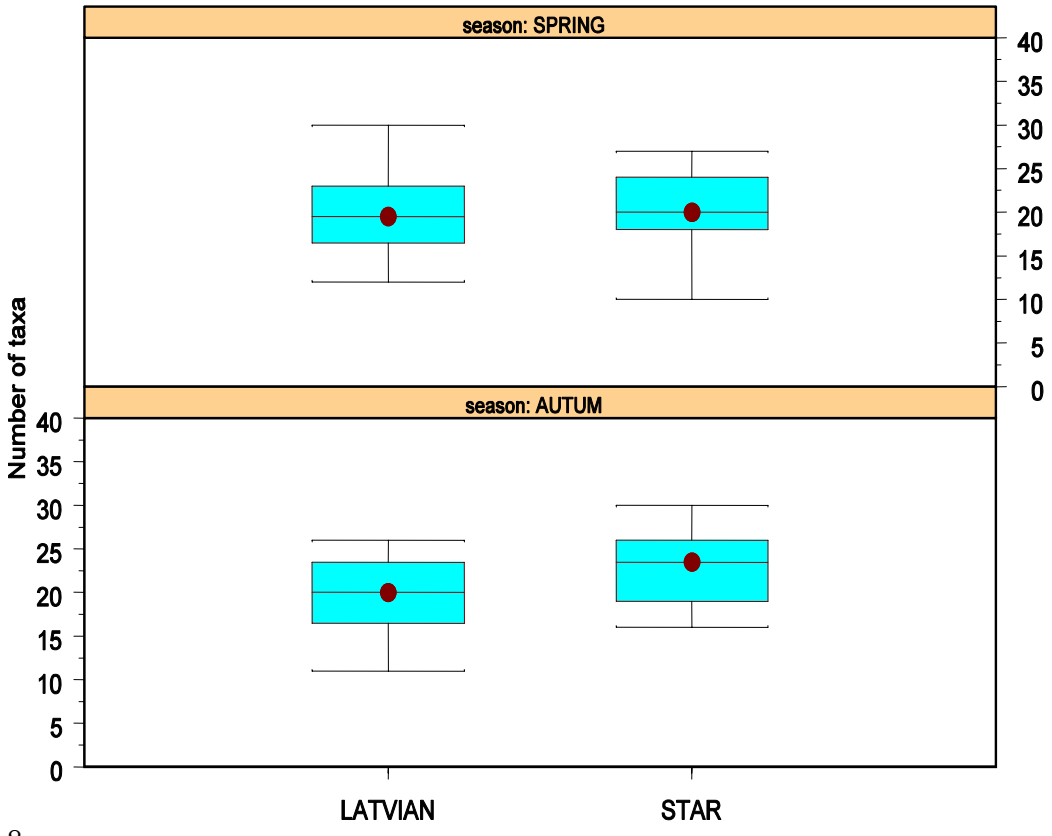
8 m


season: SPRING
season: AUTUM
Shredders
20
16
12
8
4
0
LATVIAN
STAR

8 n

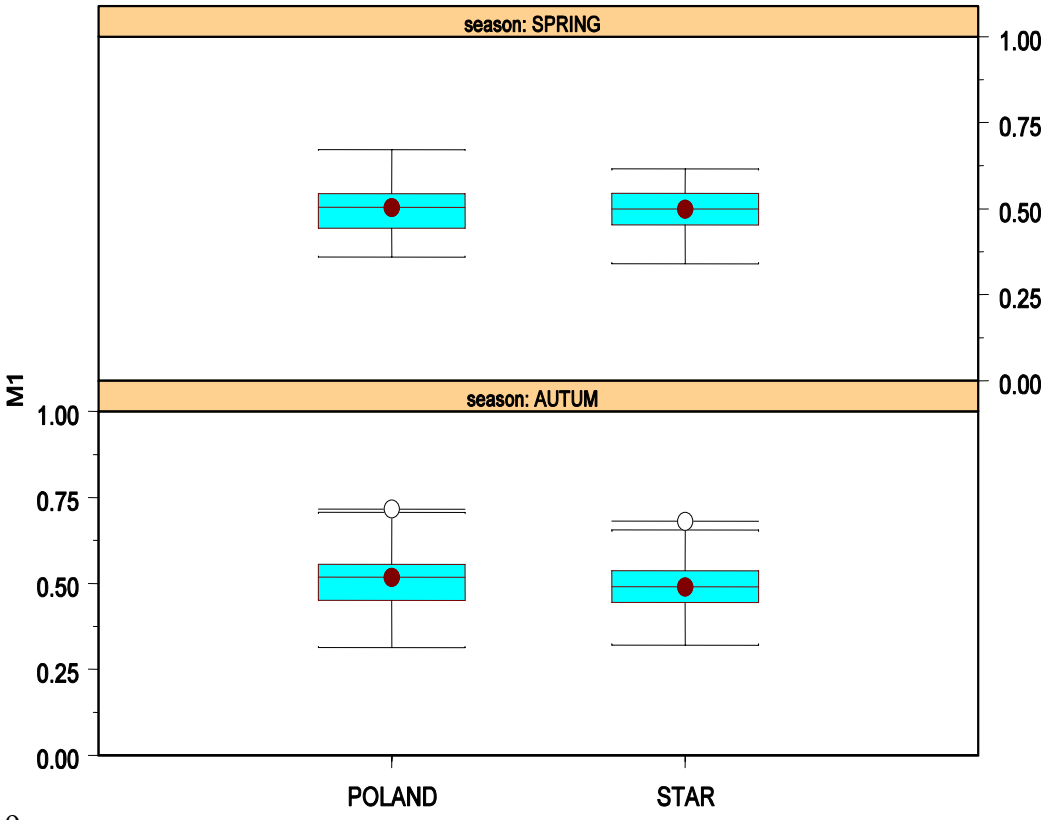

season: SPRING
season: AUTUM
Shannon - Wiener
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
LATVIAN
STAR

8 o

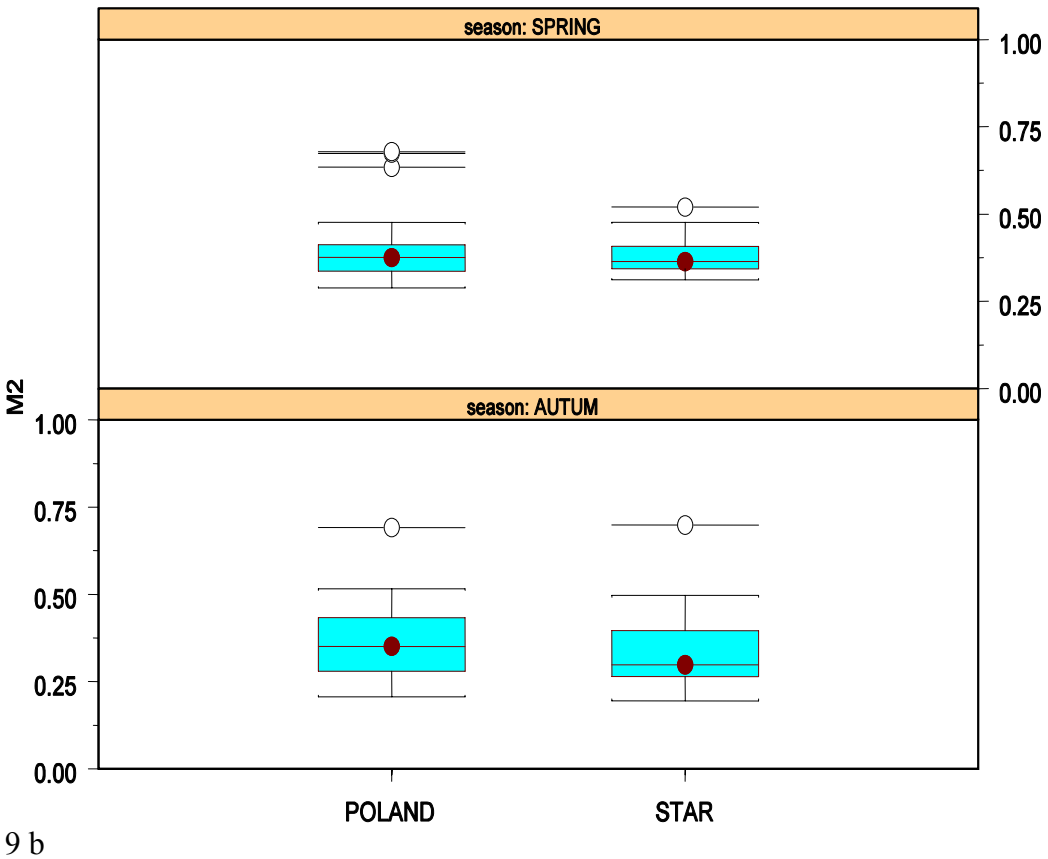

season: SPRING
season: AUTUM
Number of taxa
0
5
10
15
20
25
30
35
40
LATVIAN
STAR

8 p

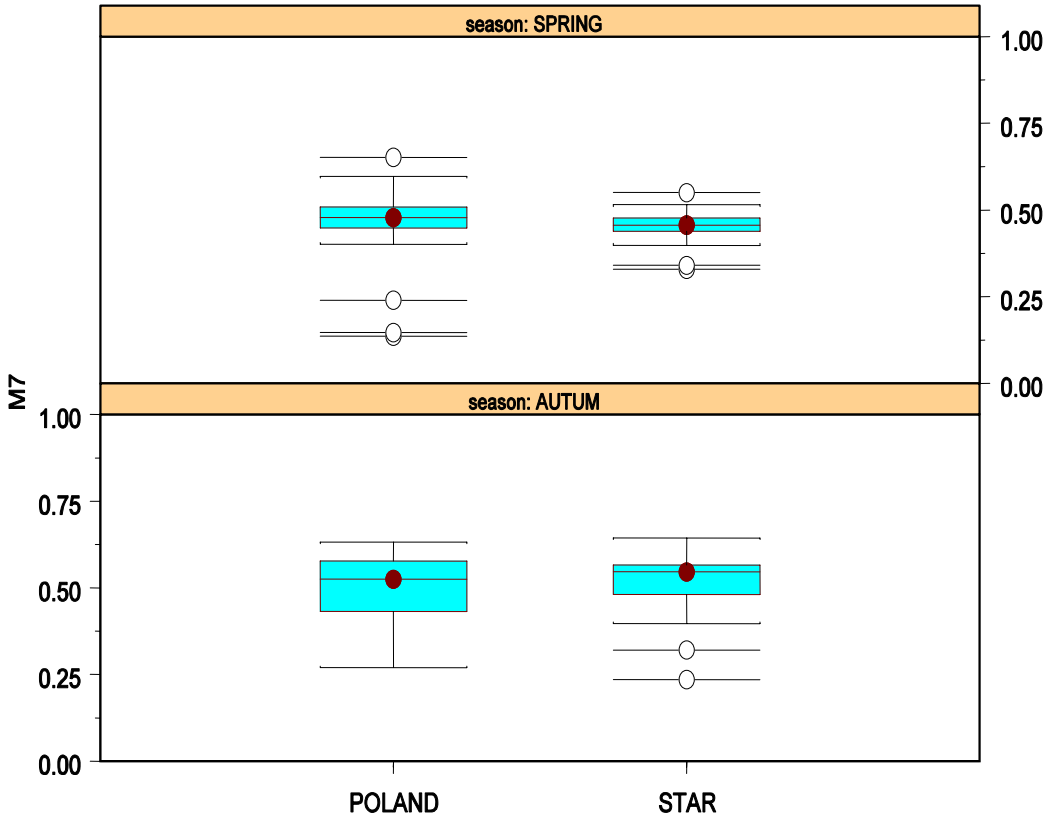
9 – Poland



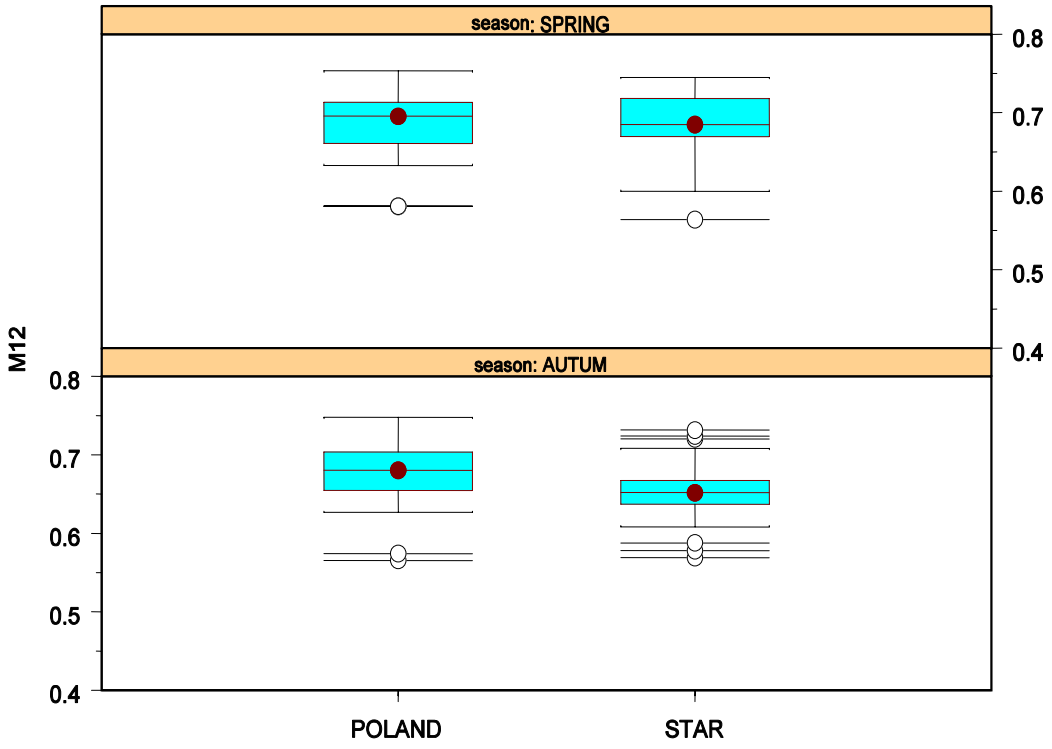
9 a



9 b

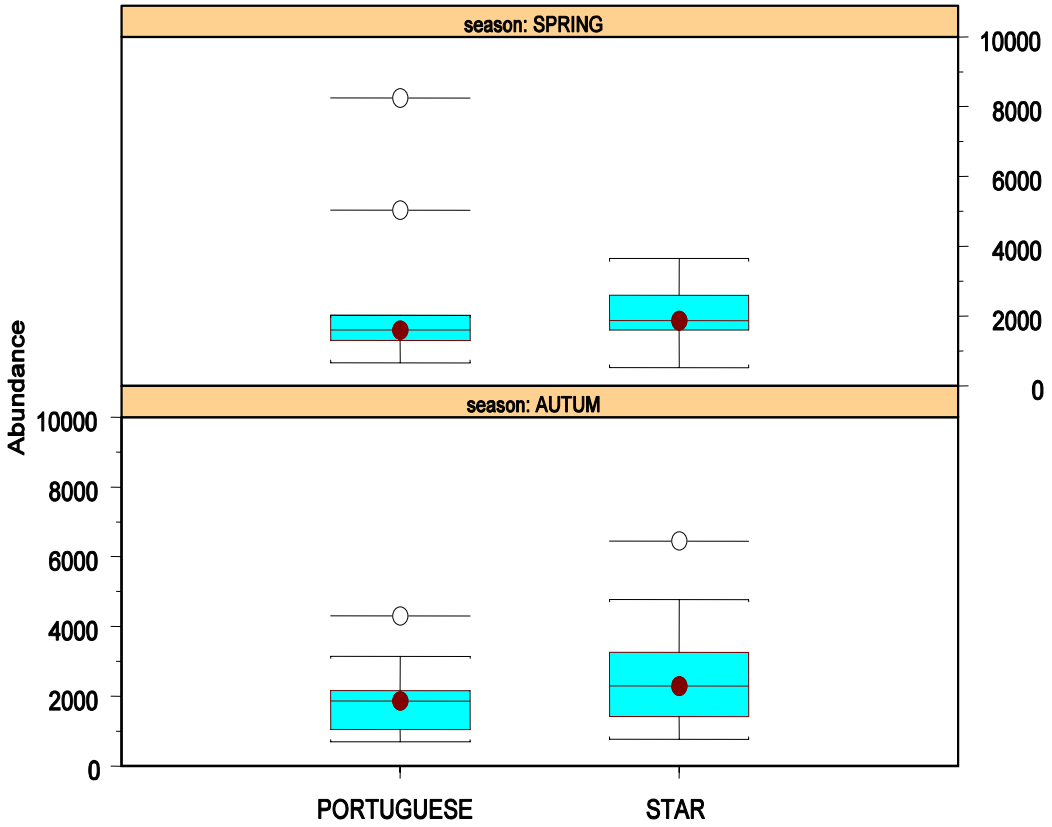

season: SPRING
season: AUTUM
M7
1.00
0.75
0.50
0.25
0.00
POLAND
STAR

9 c

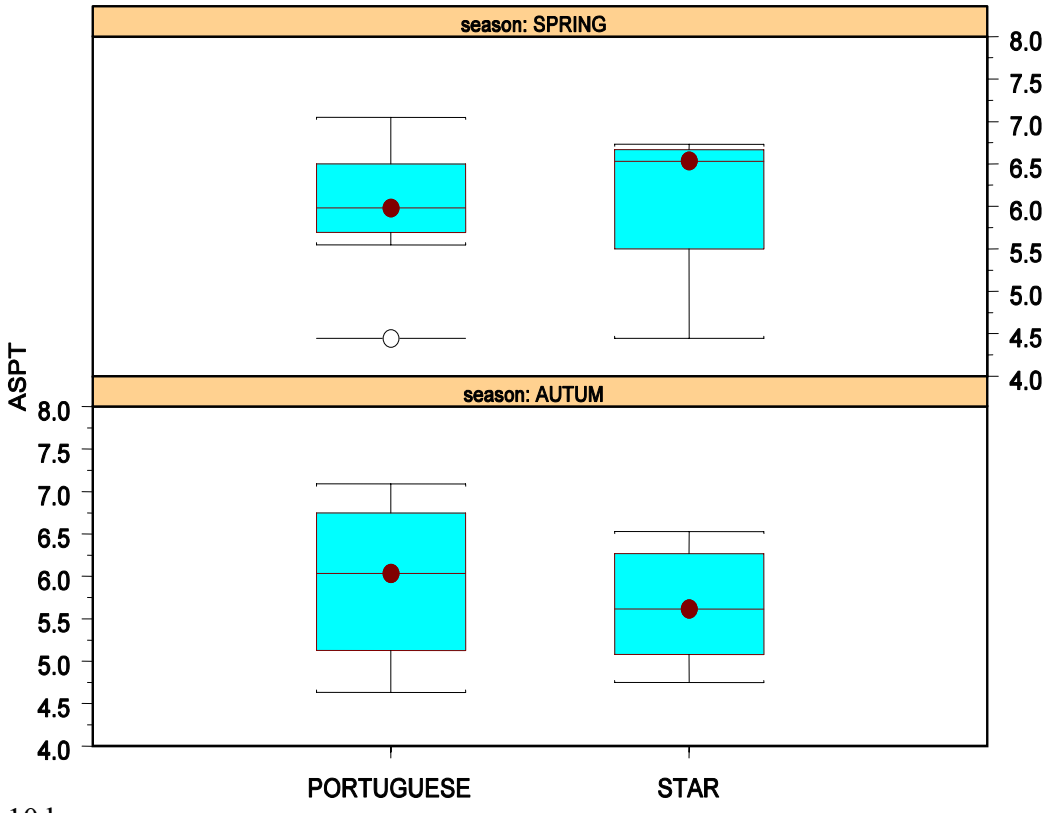

season: SPRING
season: AUTUM
M12
0.8
0.7
0.6
0.5
0.4
POLAND
STAR

9 d

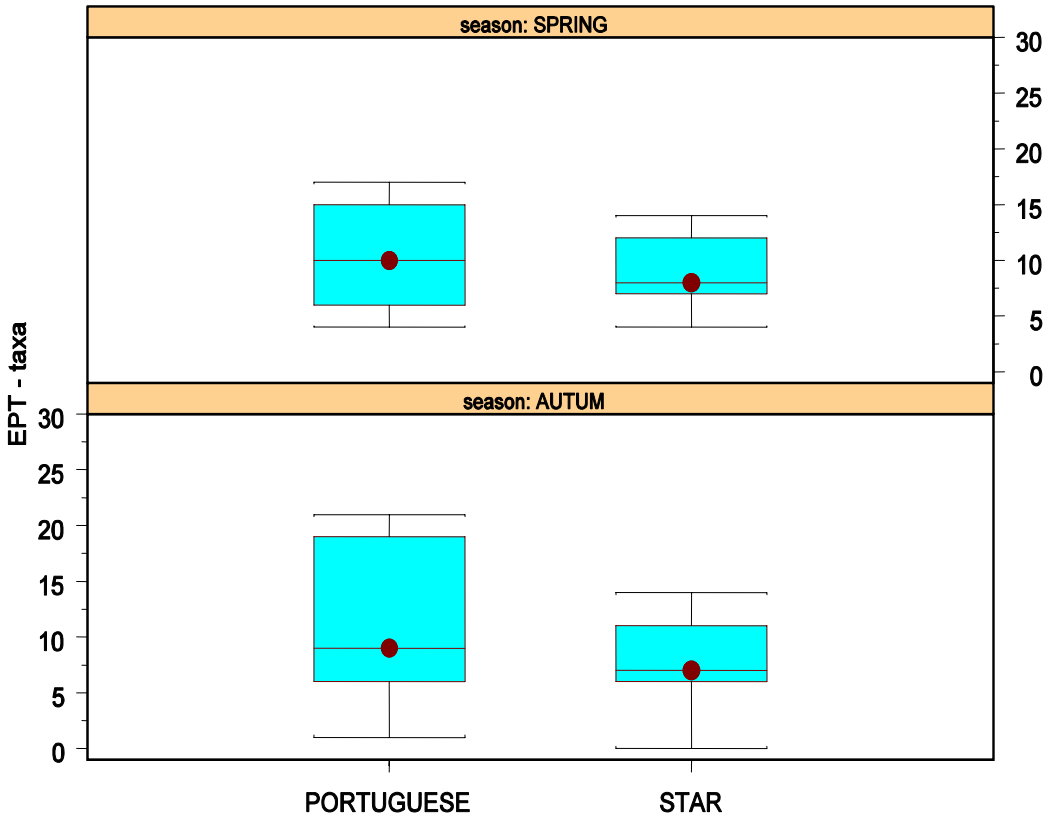
10 – Portugal



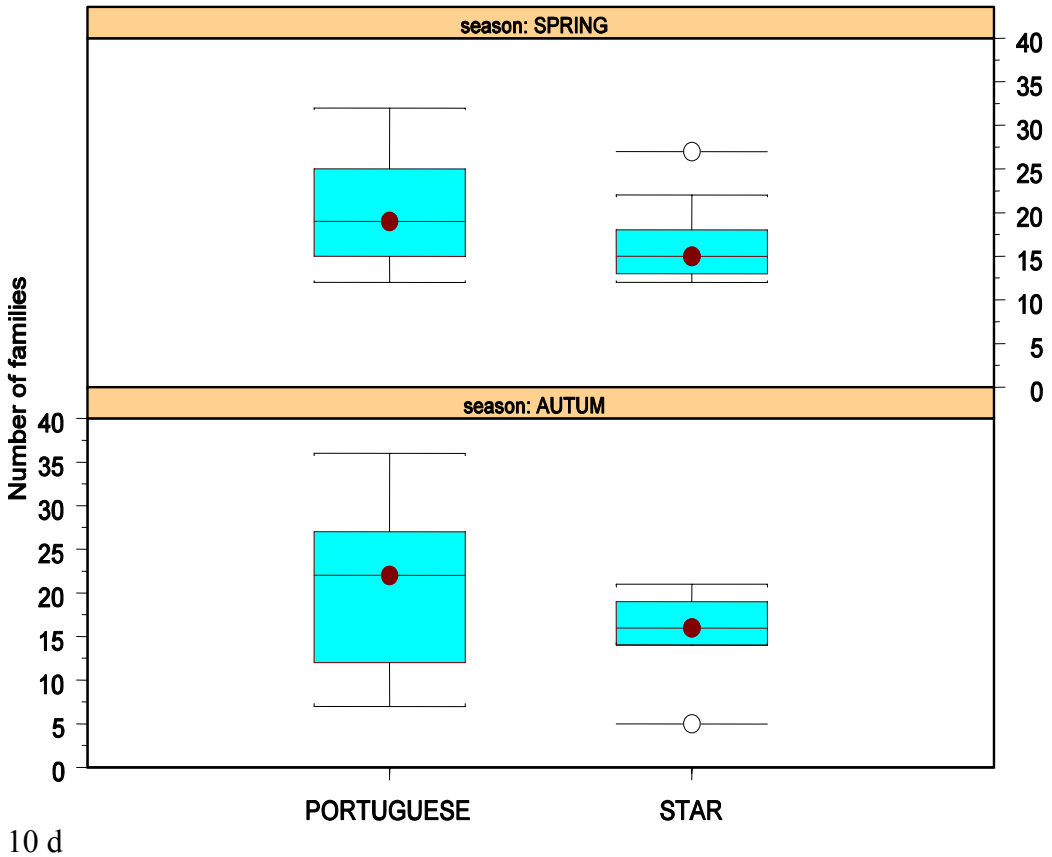
10 a



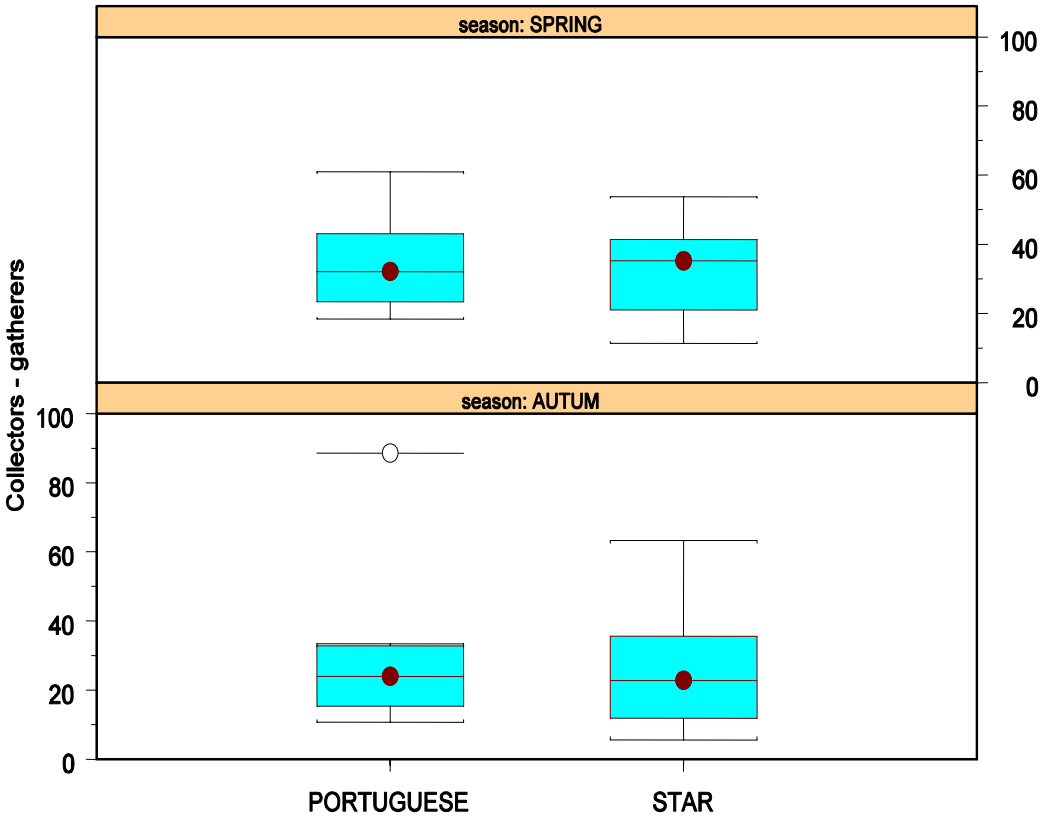
10 b


season: SPRING
season: AUTUM
EPT - taxa
0
5
10
15
20
25
30
PORTUGUESE
STAR

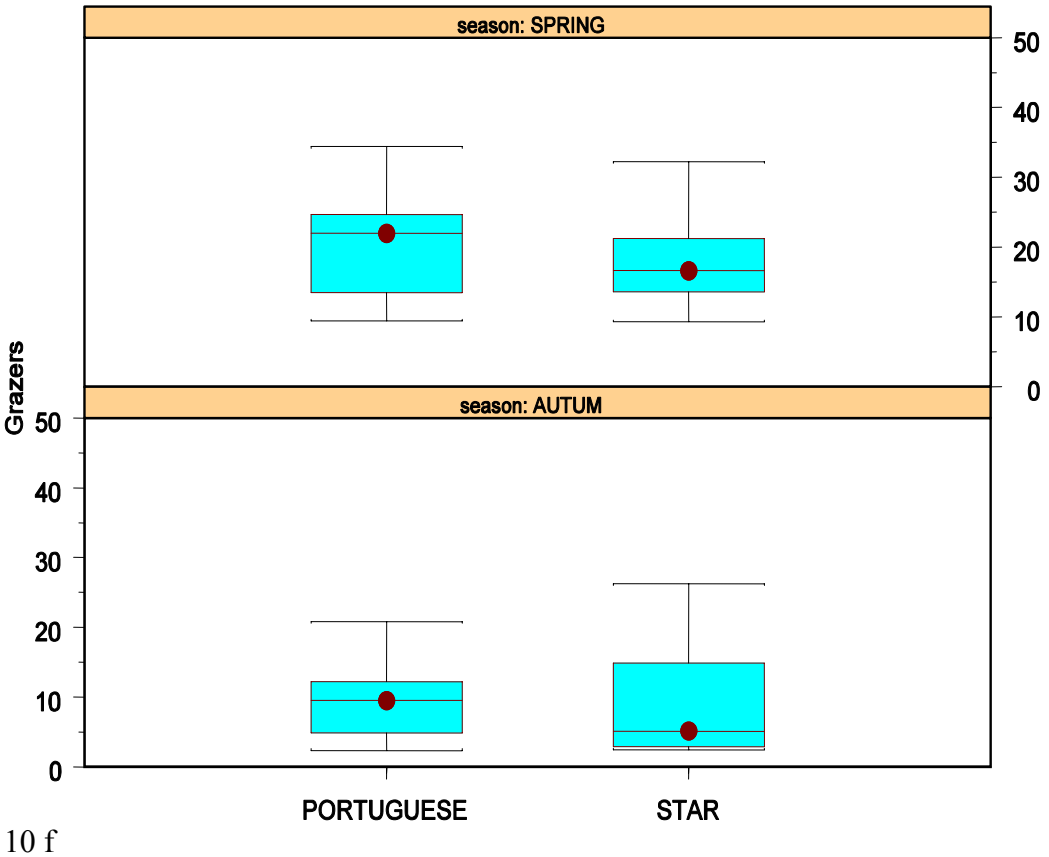
10 c


season: SPRING
season: AUTUM
Number of families
0
5
10
15
20
25
30
35
40
PORTUGUESE
STAR

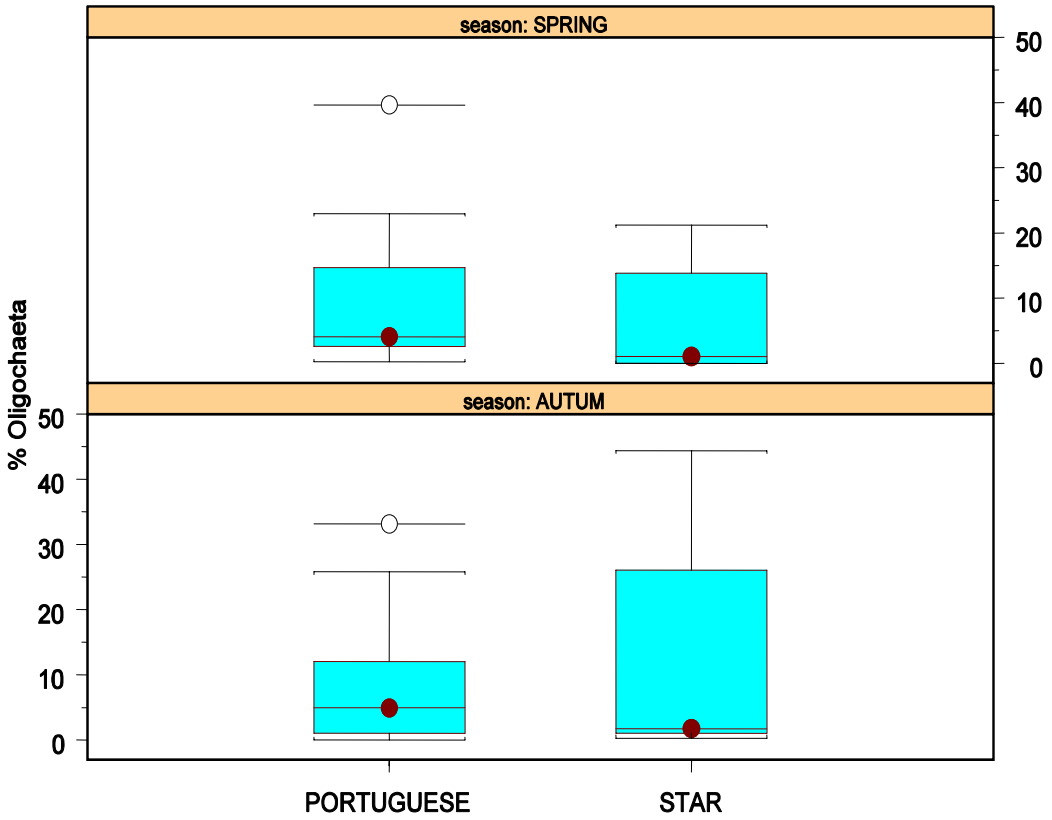
10 d


season: SPRING
season: AUTUM
Collectors - gatherers
0
20
40
60
80
100
PORTUGUESE
STAR

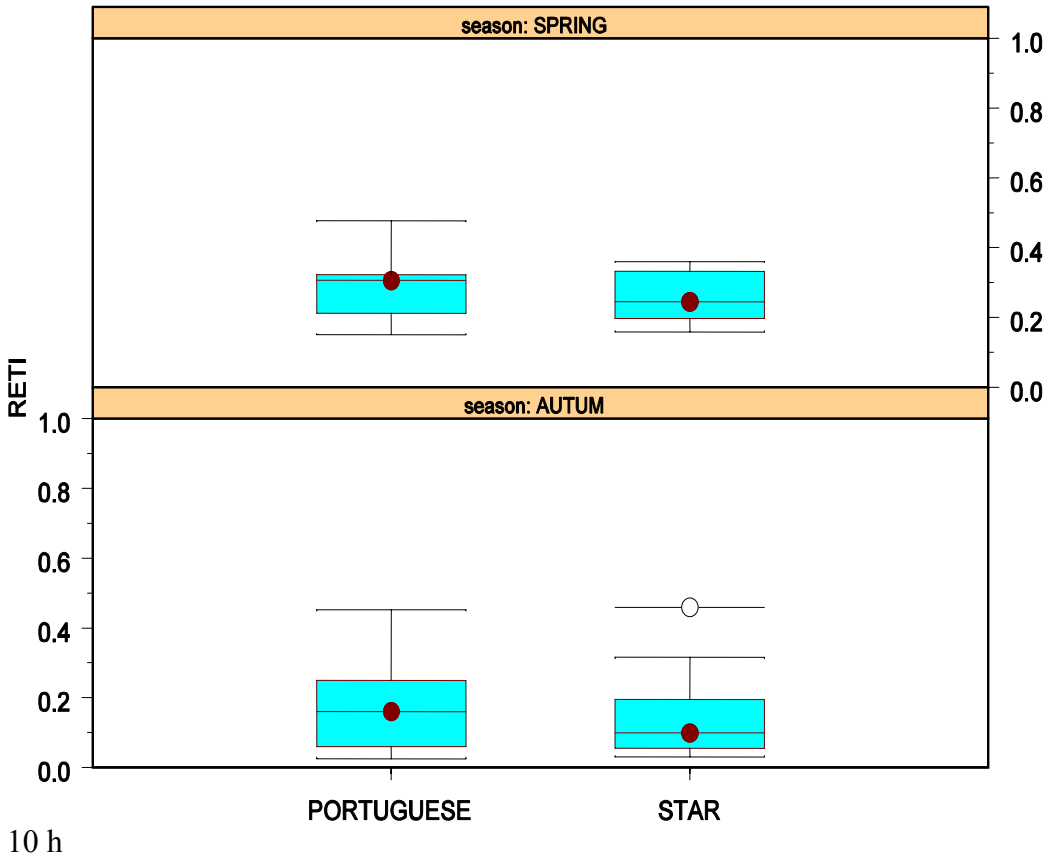
10 e


season: SPRING
season: AUTUM
Grazers
0
10
20
30
40
50
PORTUGUESE
STAR

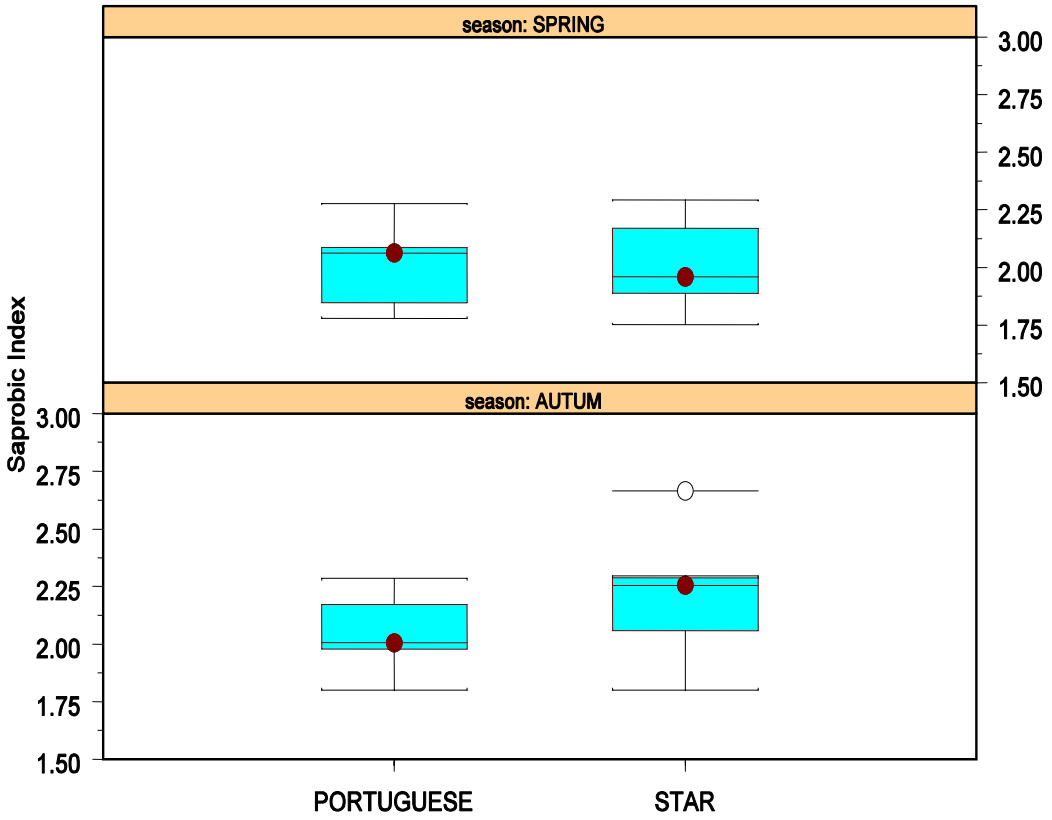
10 f


season: SPRING
season: AUTUM
% Oligochaeta
0
10
20
30
40
50
PORTUGUESE
STAR

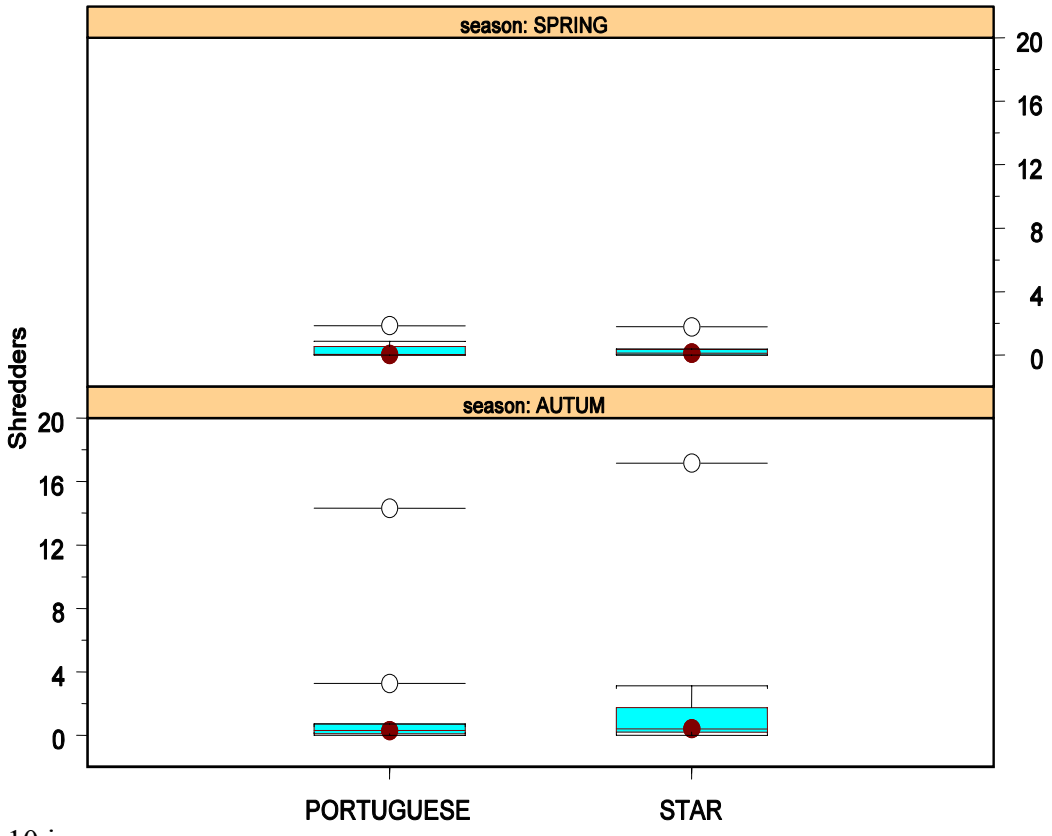
10 g


season: SPRING
season: AUTUM
RETI
0.0
0.2
0.4
0.6
0.8
1.0
PORTUGUESE
STAR

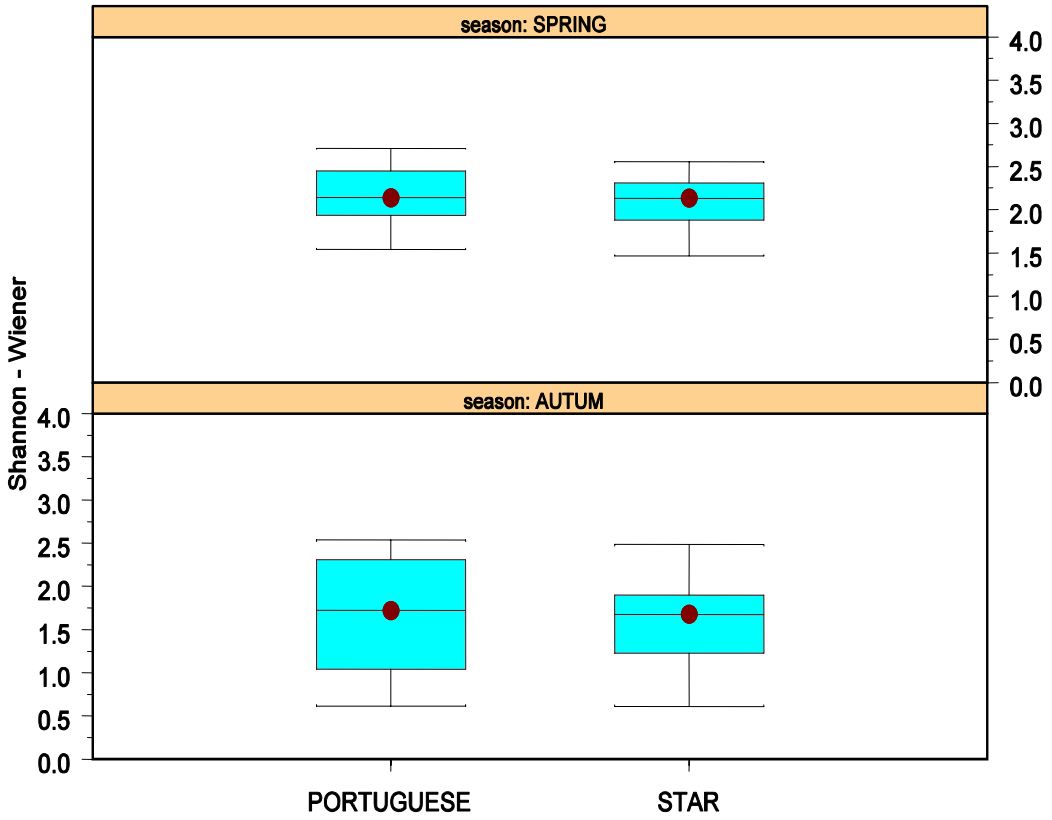
10 h


season: SPRING
season: AUTUM
Saprobic Index
3.00
2.75
2.50
2.25
2.00
1.75
1.50
PORTUGUESE
STAR

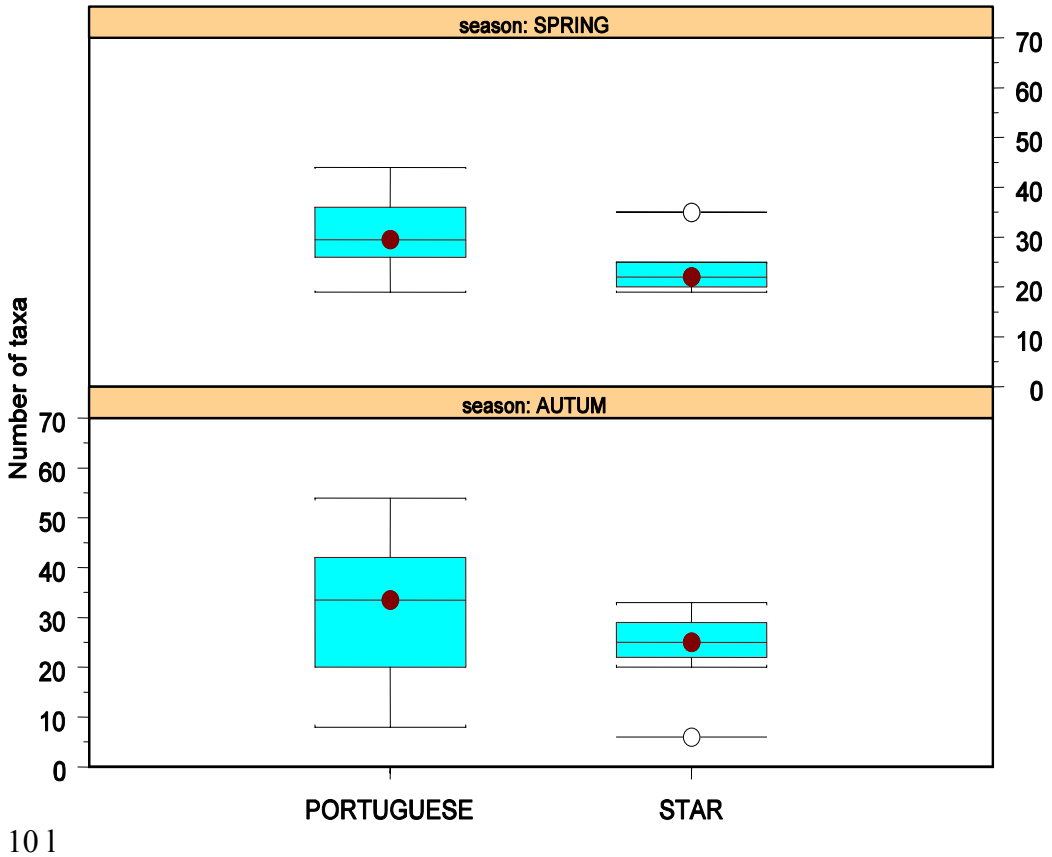
10 i


season: SPRING
season: AUTUM
Shredders
20
16
12
8
4
0
PORTUGUESE
STAR

10 j

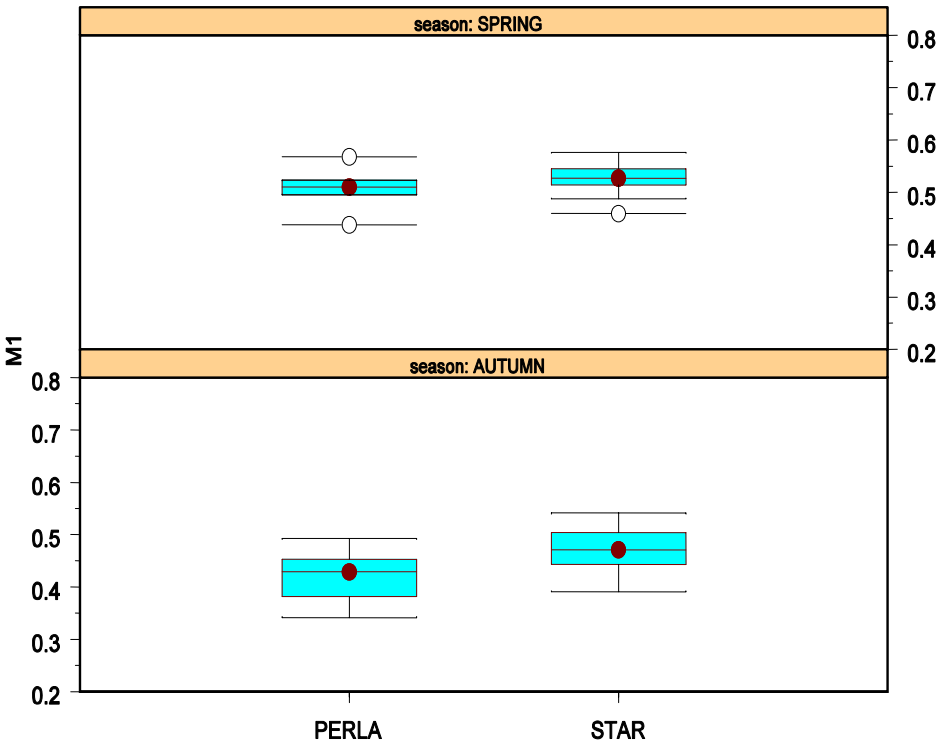

season: SPRING
season: AUTUM
Shannon - Wiener
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
PORTUGUESE
STAR

10 k

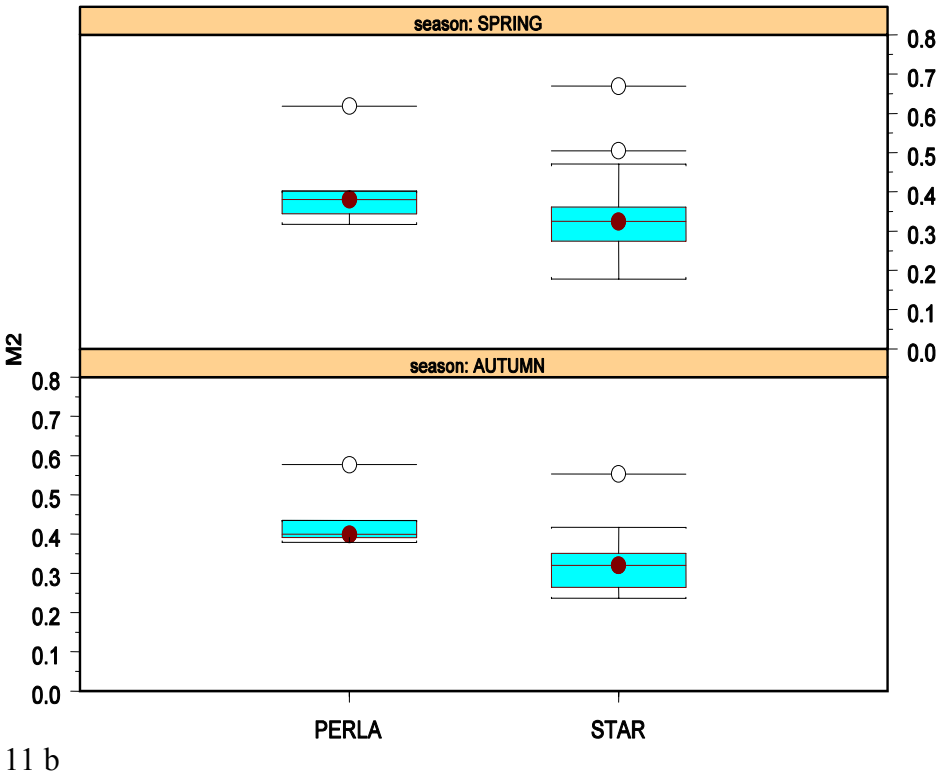

season: SPRING
season: AUTUM
Number of taxa
0
10
20
30
40
50
60
70
PORTUGUESE
STAR

10 l

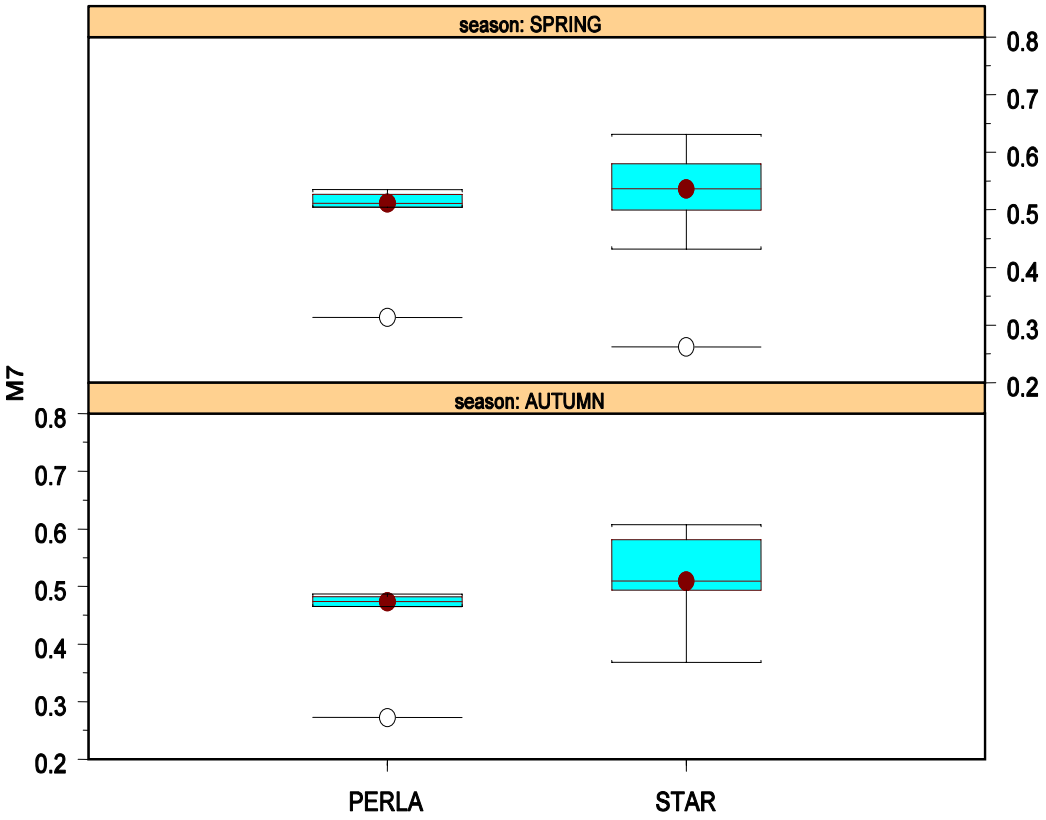
11 - Slovakia



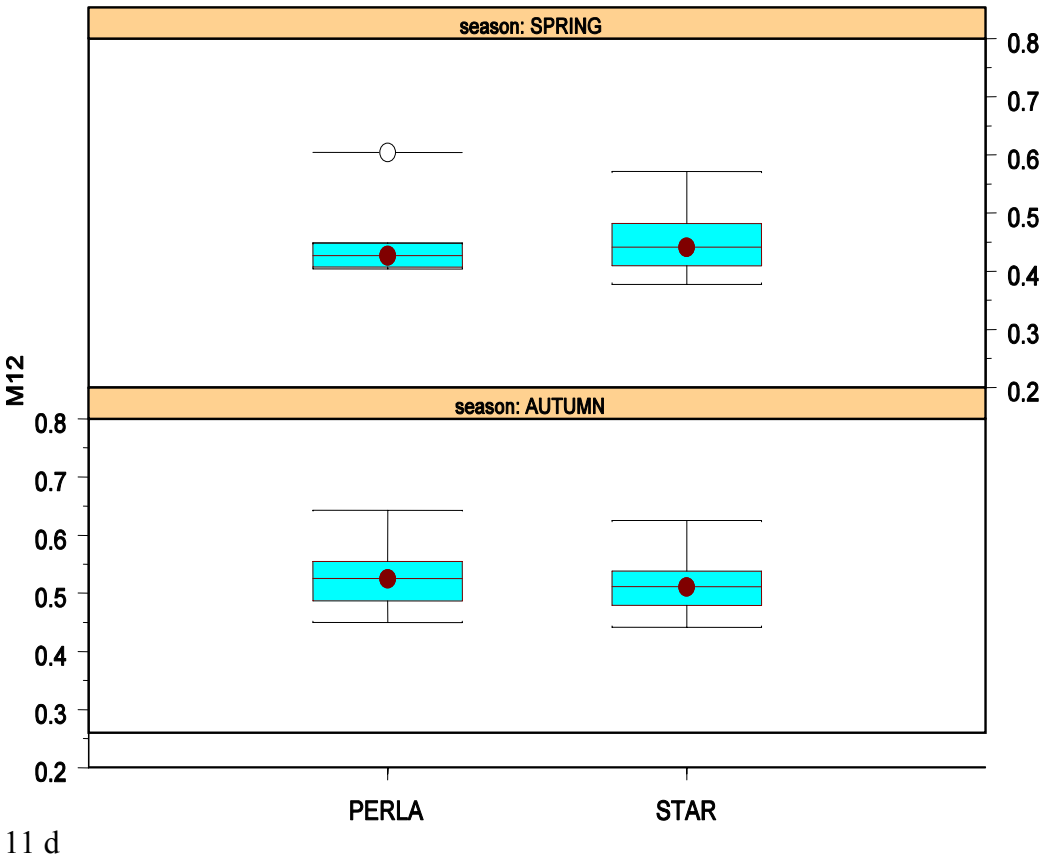
11a



11 b

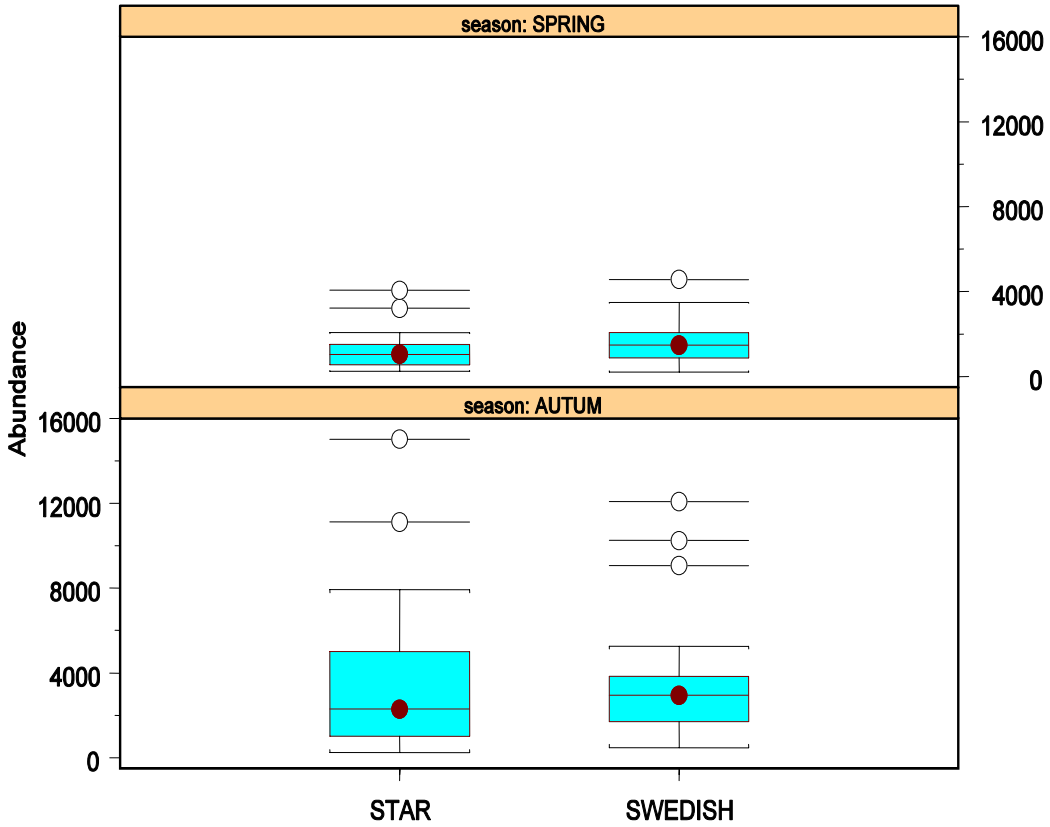

season: SPRING
season: AUTUMN
M7
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PERLA
STAR

11 c

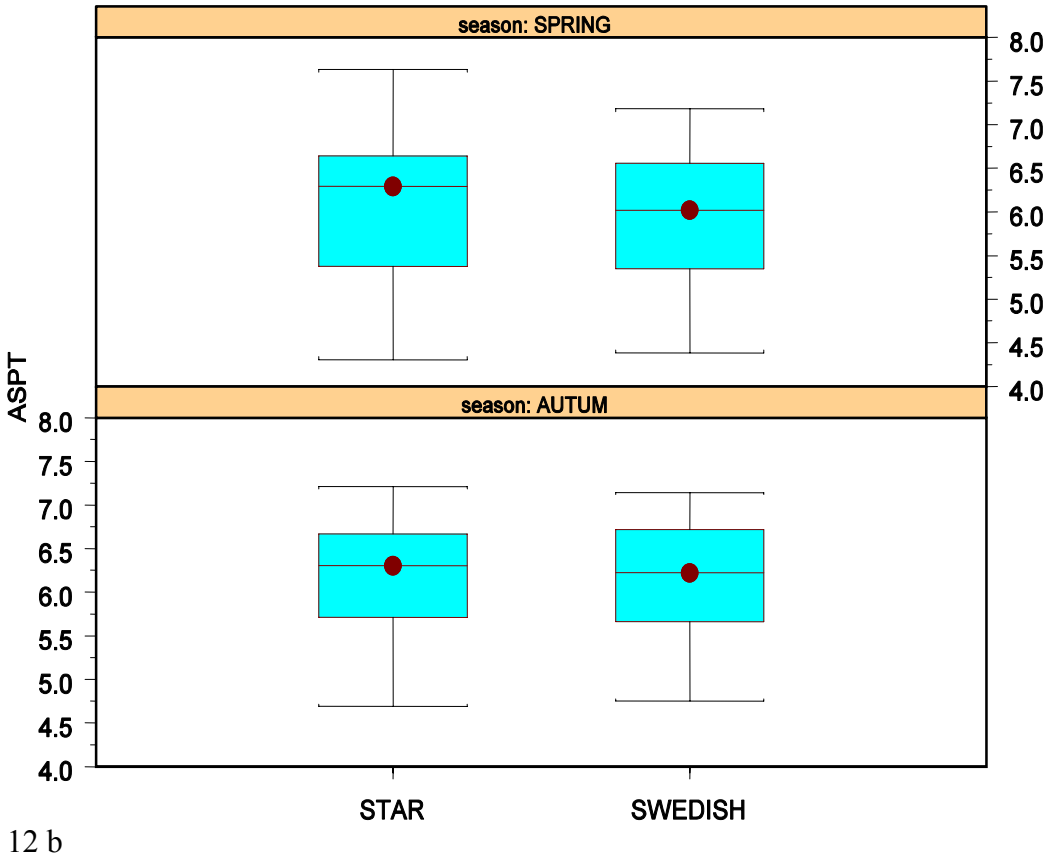

season: SPRING
season: AUTUMN
M12
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PERLA
STAR

11 d

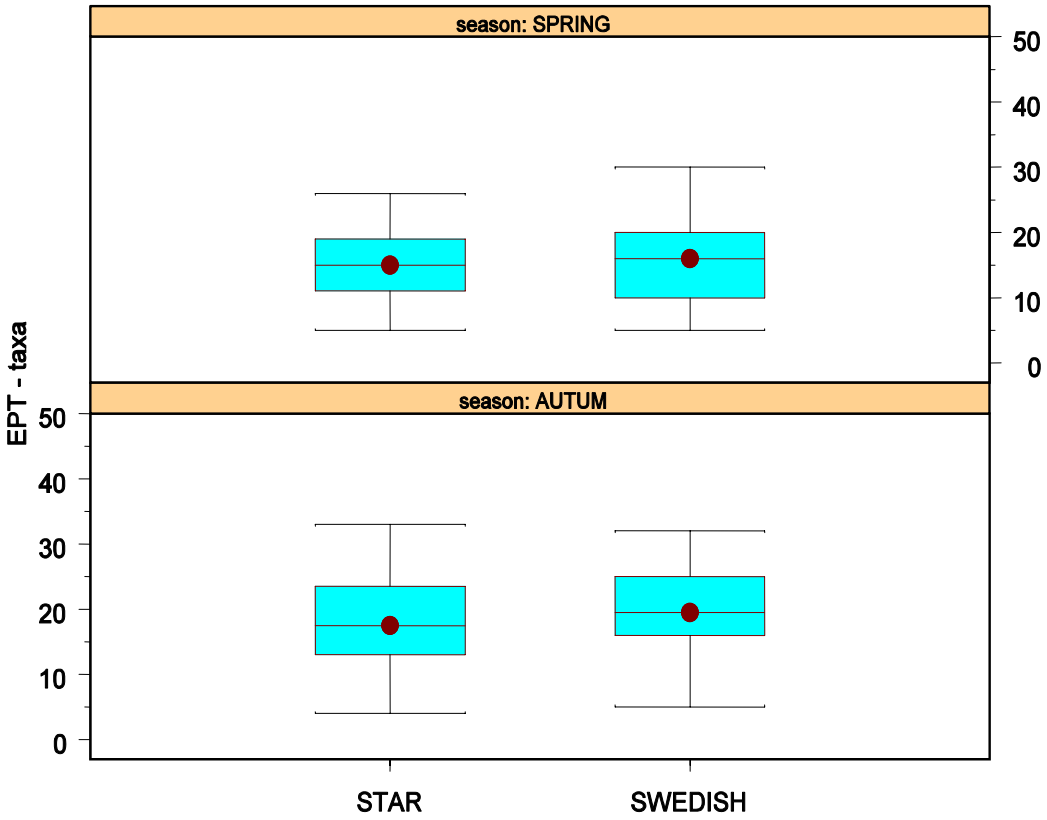
12 - Sweden



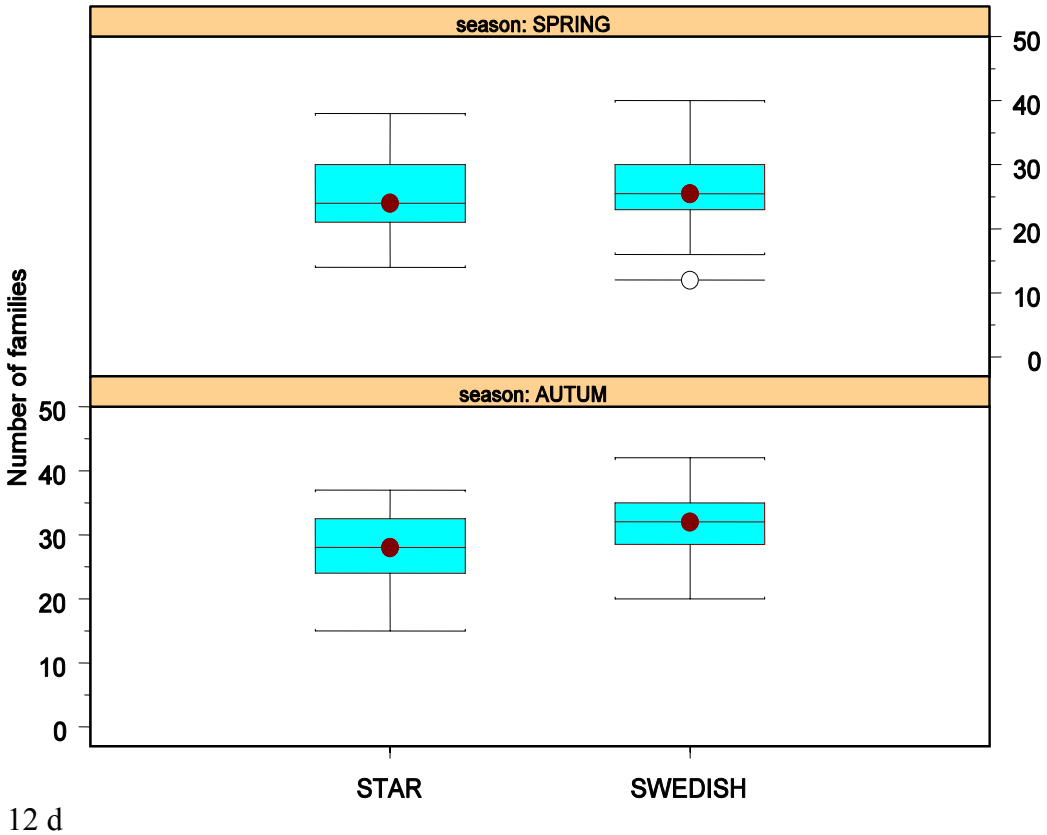
12 a



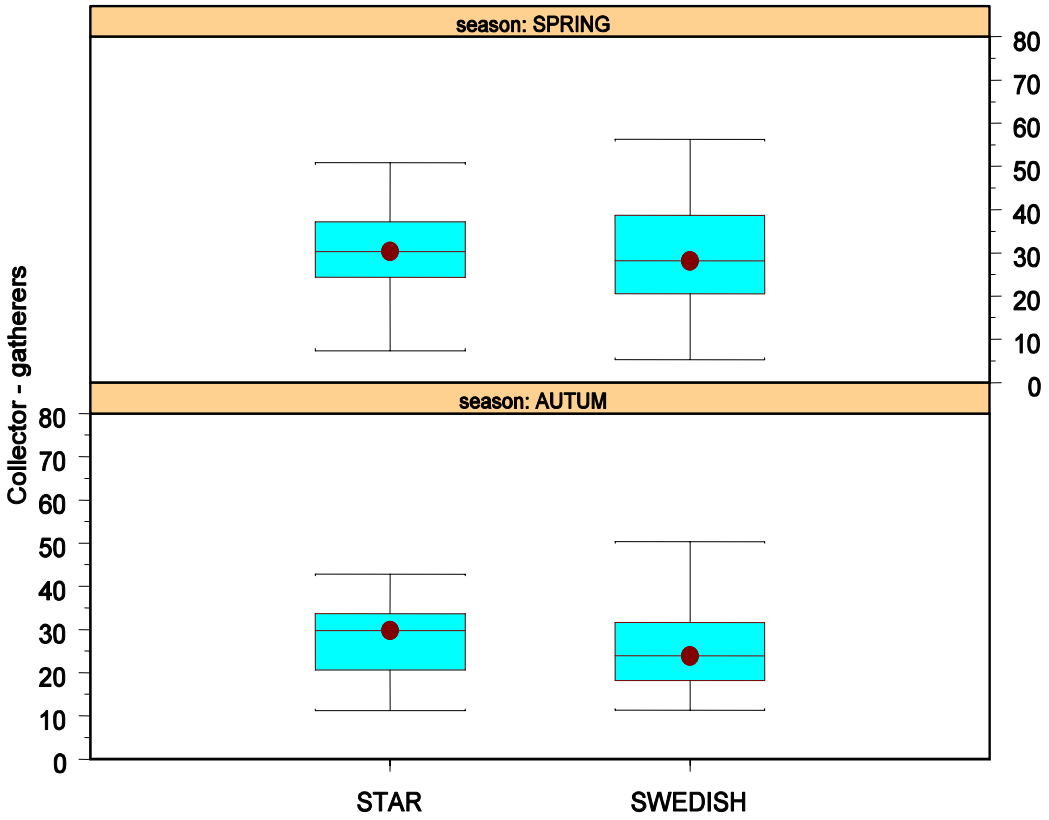
12 b


season: SPRING
season: AUTUM
EPT - taxa
0
10
20
30
40
50
STAR
SWEDISH

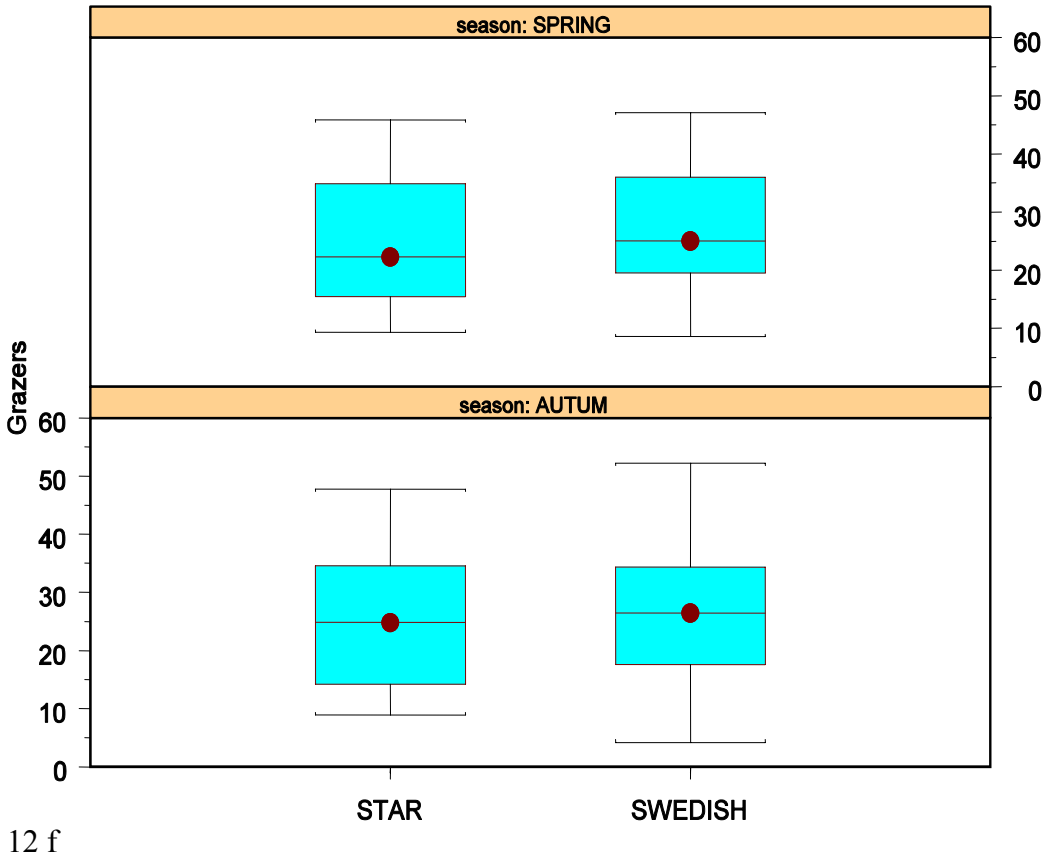
12 c


season: SPRING
season: AUTUM
Number of families
0
10
20
30
40
50
STAR
SWEDISH

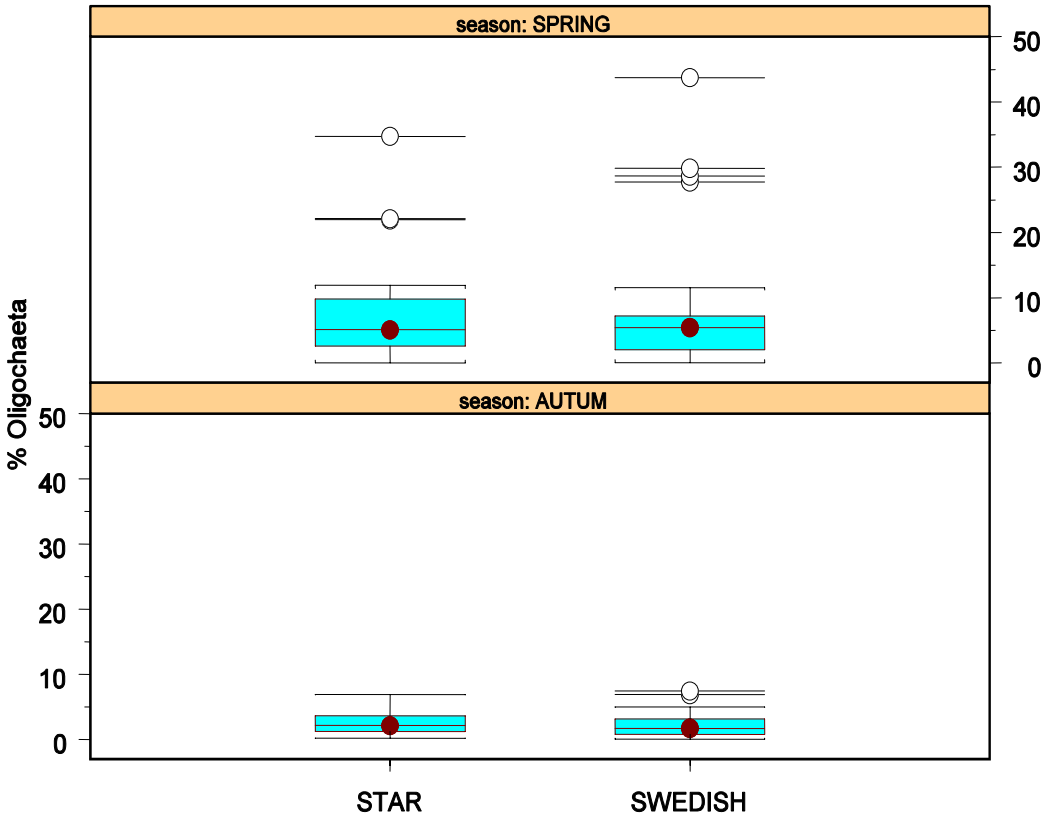
12 d


season: SPRING
season: AUTUM
Collector - gatherers
0
10
20
30
40
50
60
70
80
STAR
SWEDISH

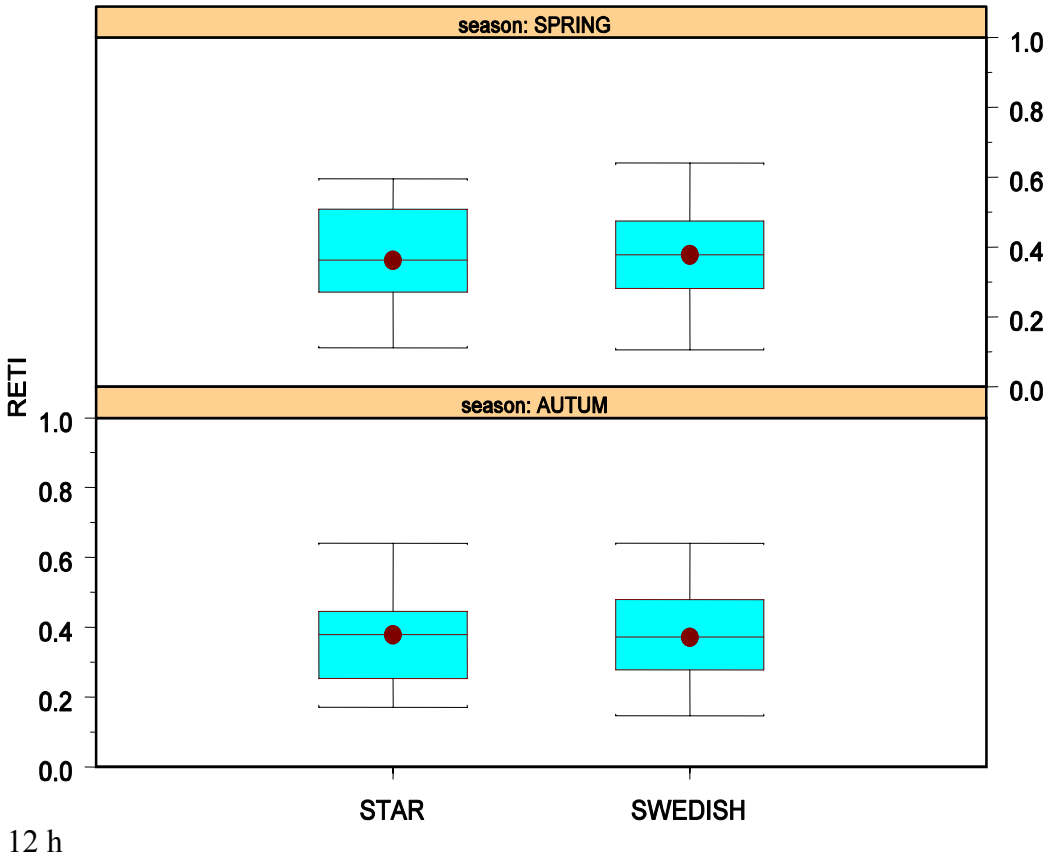
12 e


season: SPRING
season: AUTUM
Grazers
0
10
20
30
40
50
60
STAR
SWEDISH

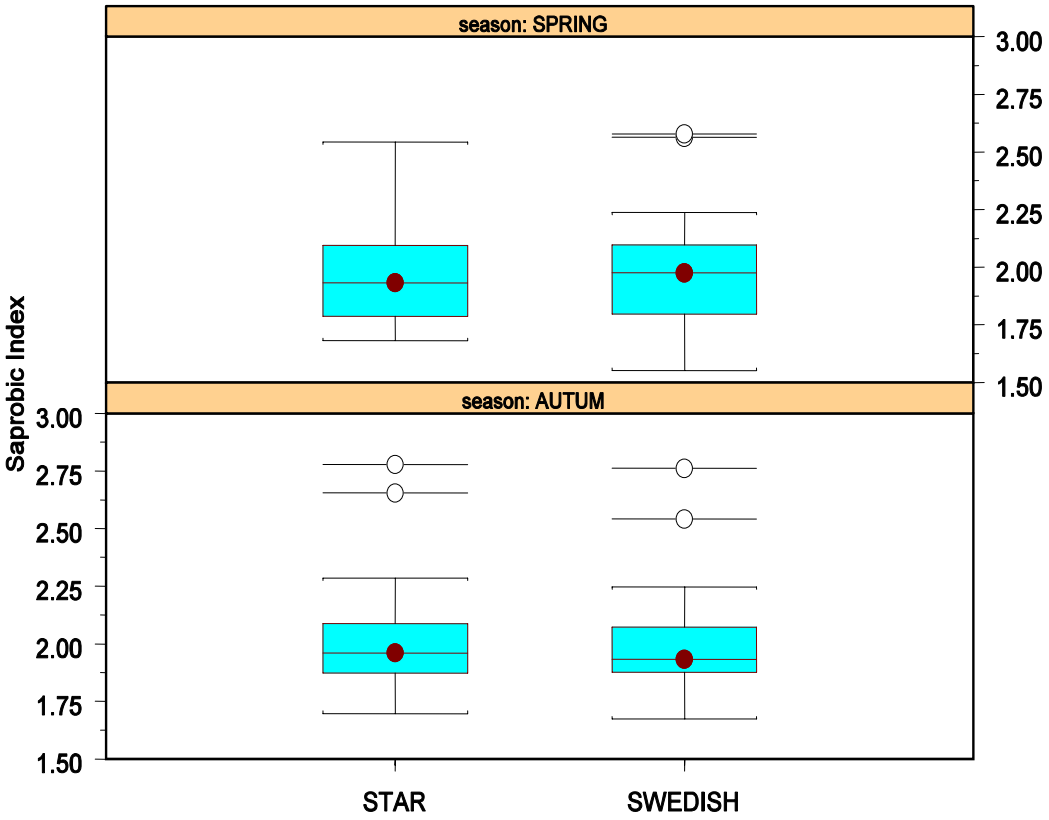
12 f


season: SPRING
season: AUTUM
% Oligochaeta
0
10
20
30
40
50
STAR
SWEDISH

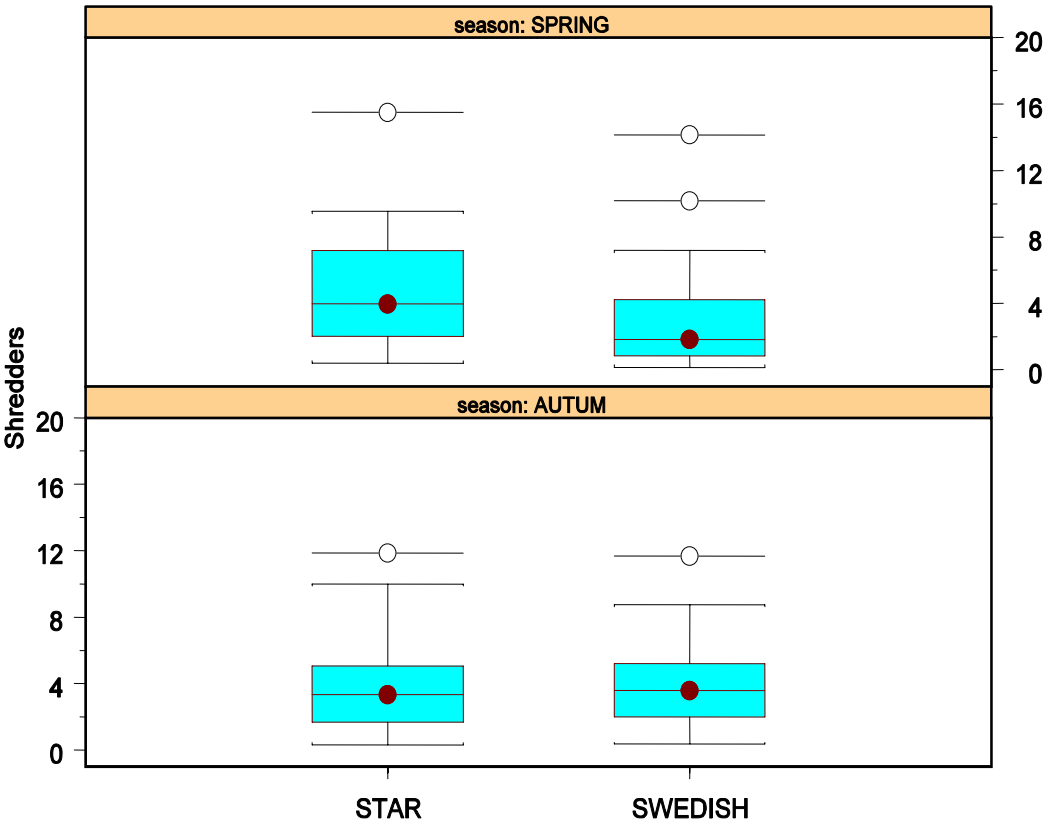
12 g


season: SPRING
season: AUTUM
RETI
0.0
0.2
0.4
0.6
0.8
1.0
STAR
SWEDISH

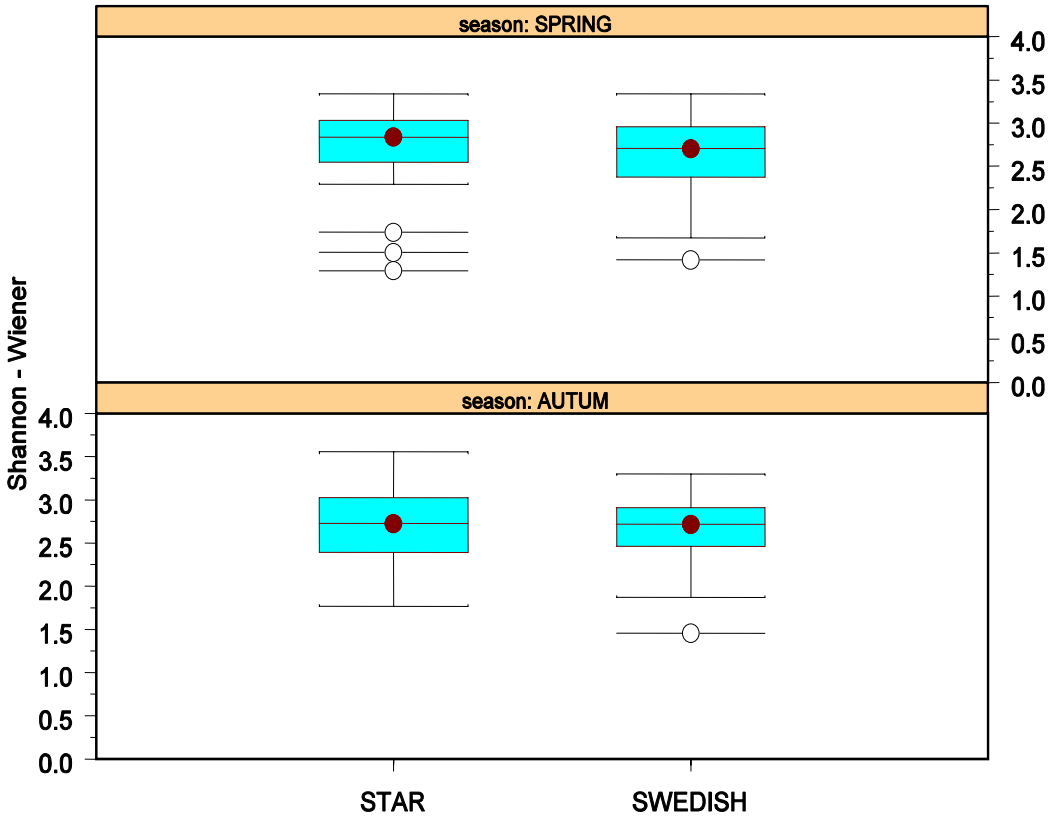
12 h


season: SPRING
season: AUTUM
Saprobic Index
3.00
2.75
2.50
2.25
2.00
1.75
1.50
STAR
SWEDISH

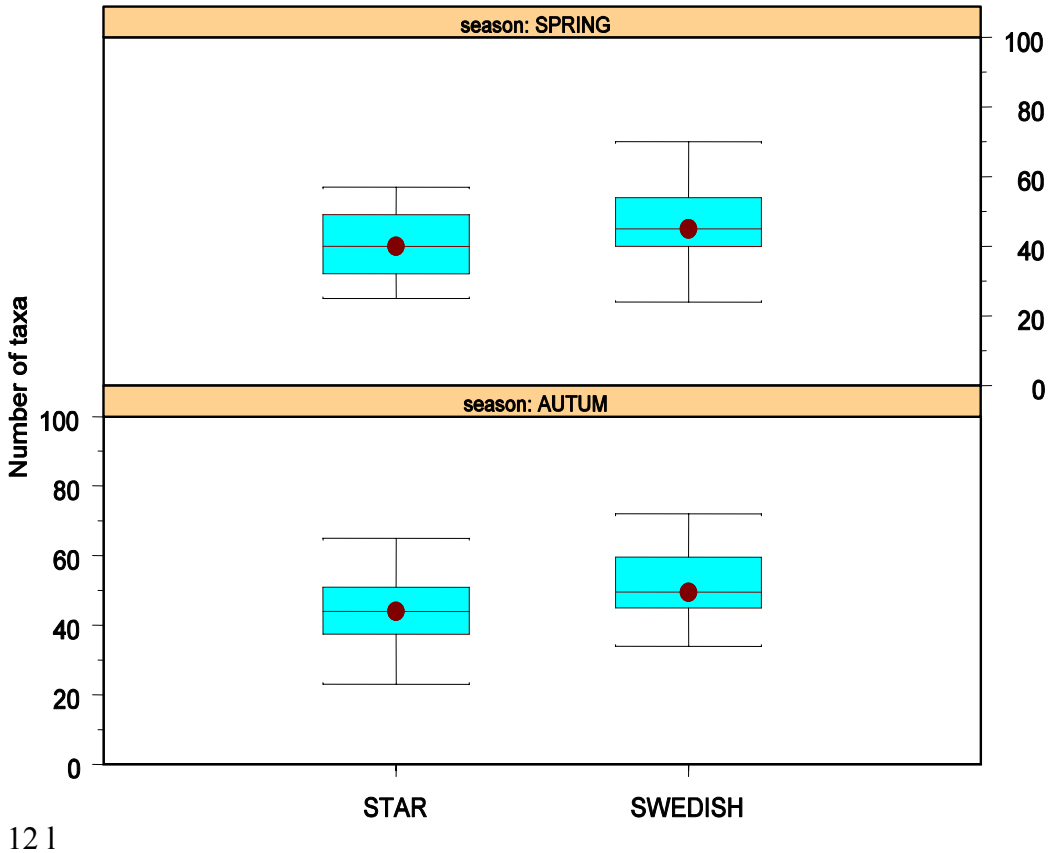
12 i


season: SPRING
season: AUTUM
Shredders
20
16
12
8
4
0
STAR
SWEDISH

12 j

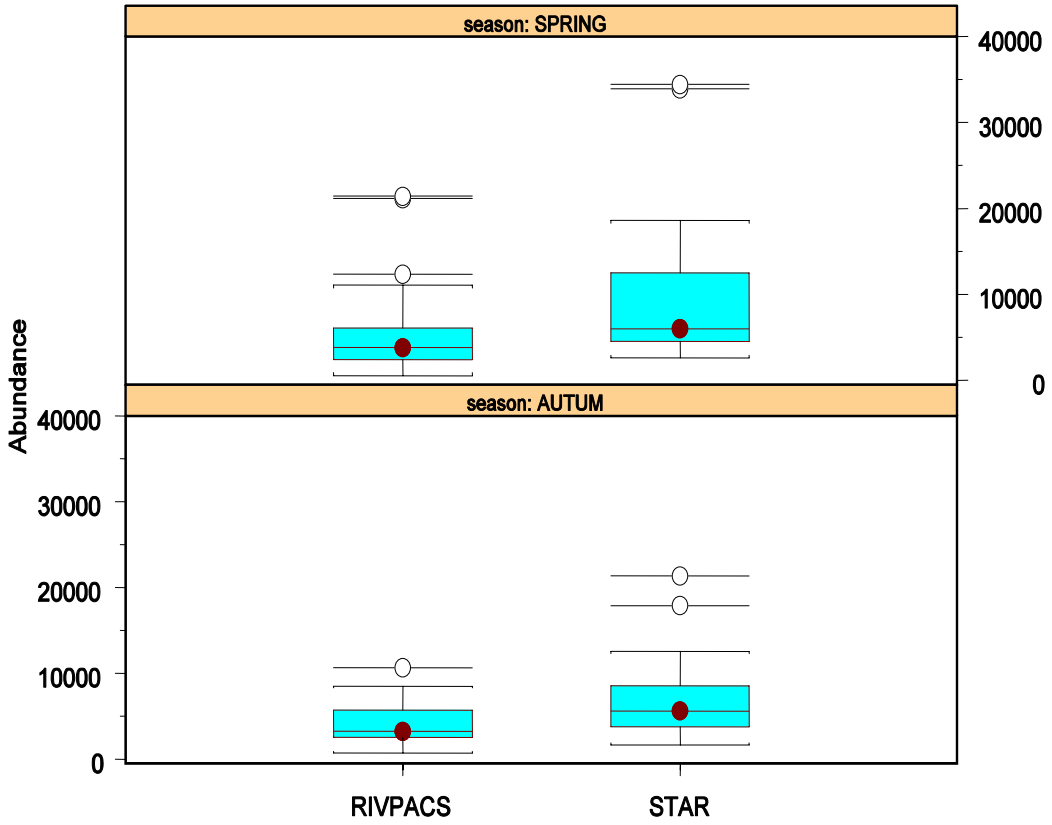

season: SPRING
season: AUTUM
Shannon - Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
STAR
SWEDISH

12 k

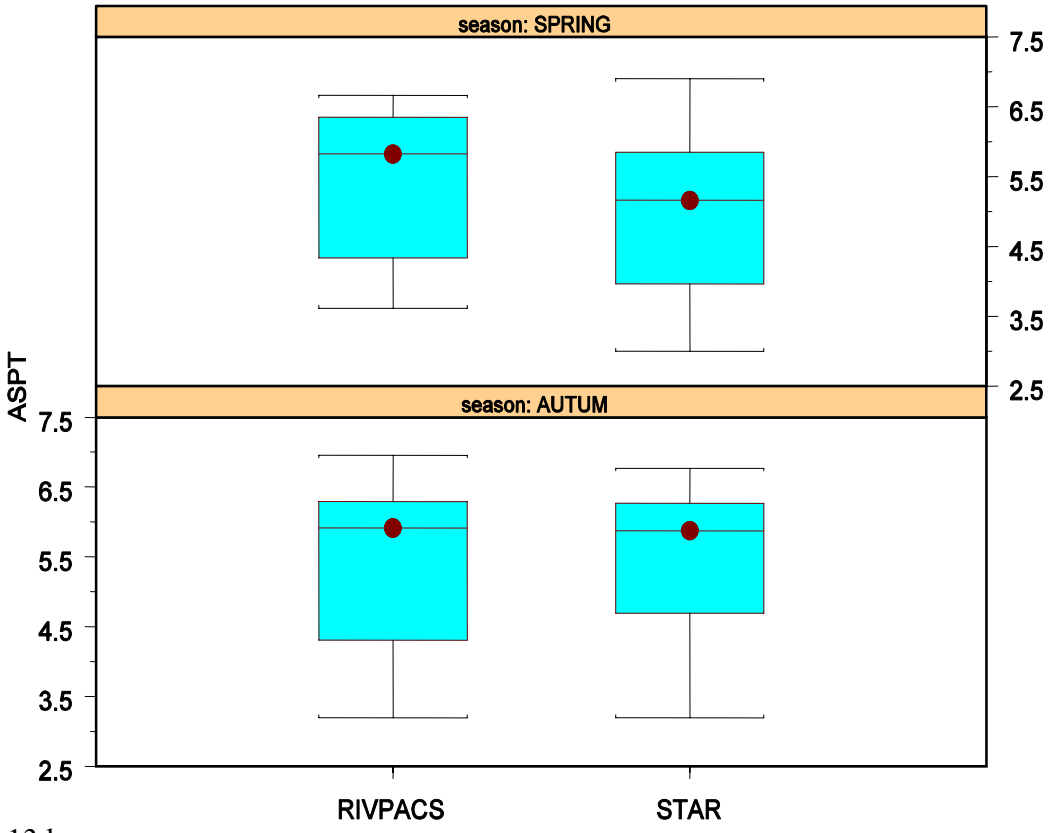

season: SPRING
season: AUTUM
Number of taxa
100
80
60
40
20
0
STAR
SWEDISH

12 l

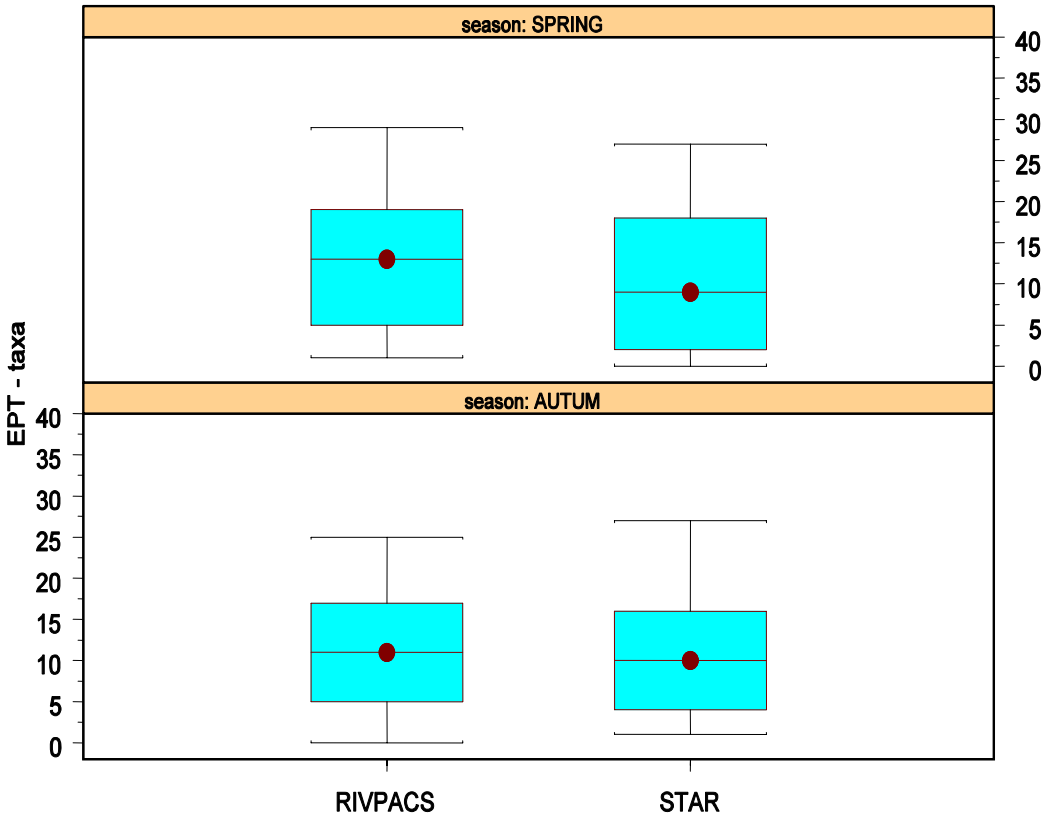
13 - U.K.


season: SPRING
season: AUTUM
Abundance
40000
30000
20000
10000
0
RIVPACS
STAR

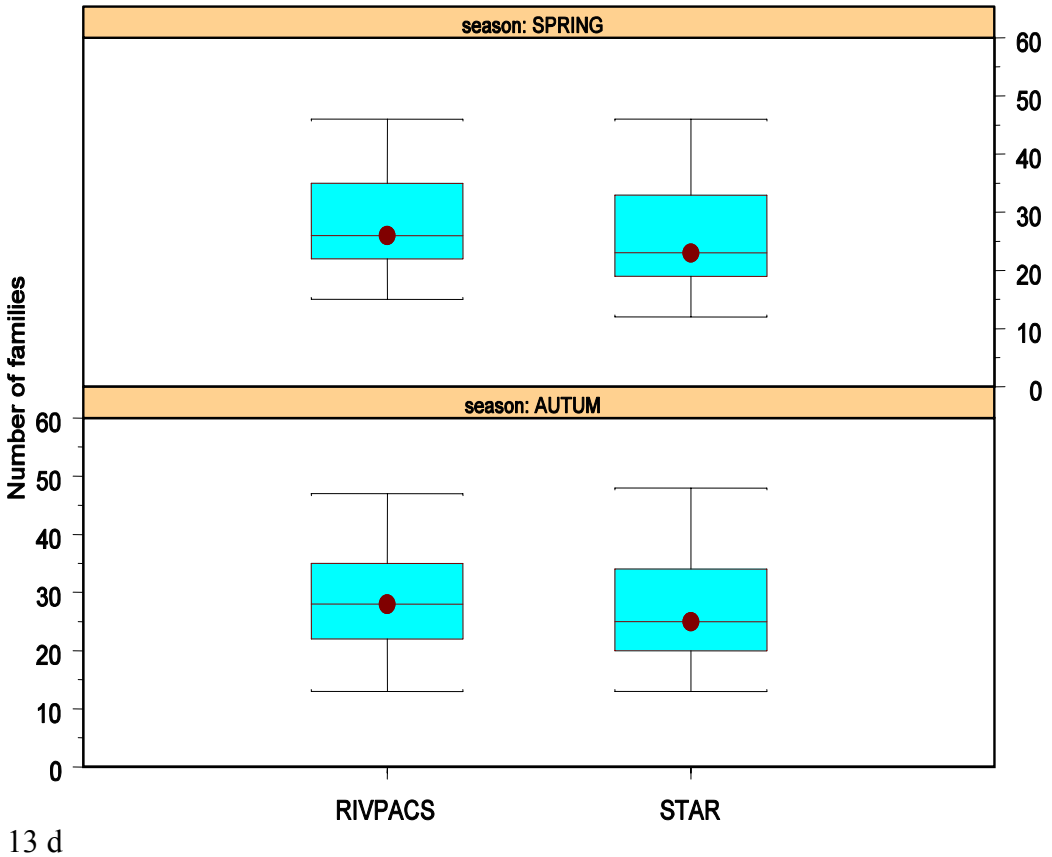
13 a


season: SPRING
season: AUTUM
ASPT
7.5
6.5
5.5
4.5
3.5
2.5
RIVPACS
STAR

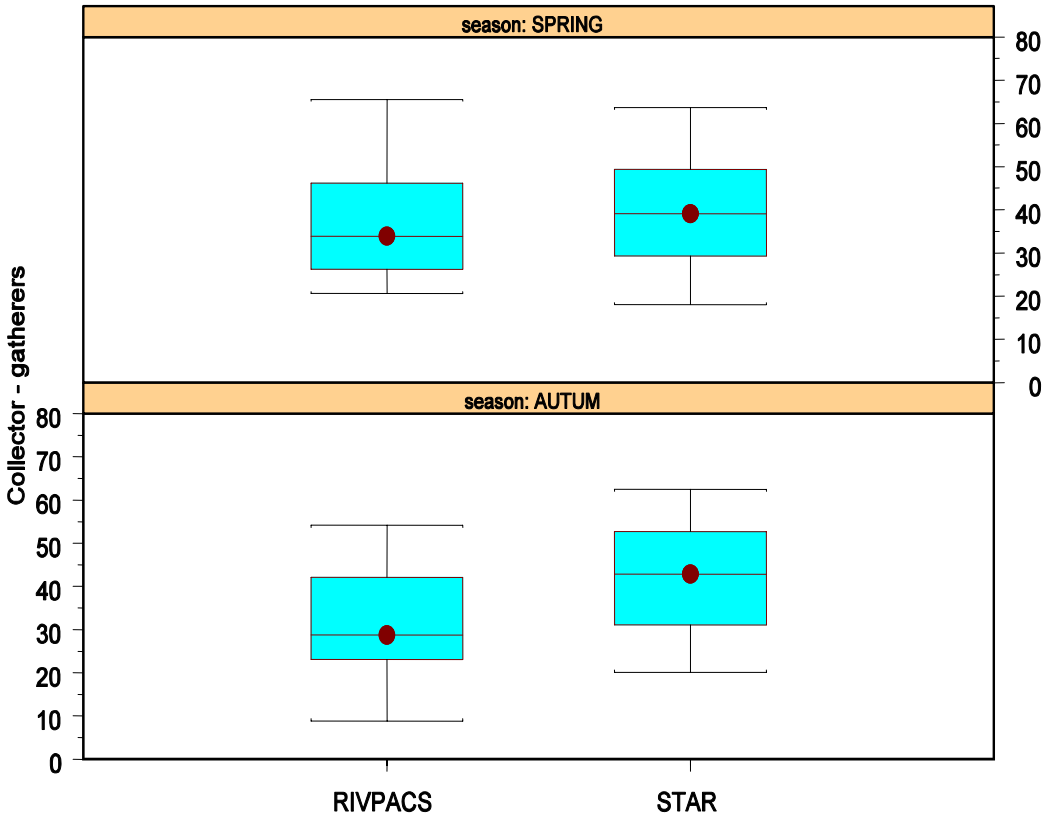
13 b


season: SPRING
season: AUTUM
EPT - taxa
0
5
10
15
20
25
30
35
40
RIVPACS
STAR

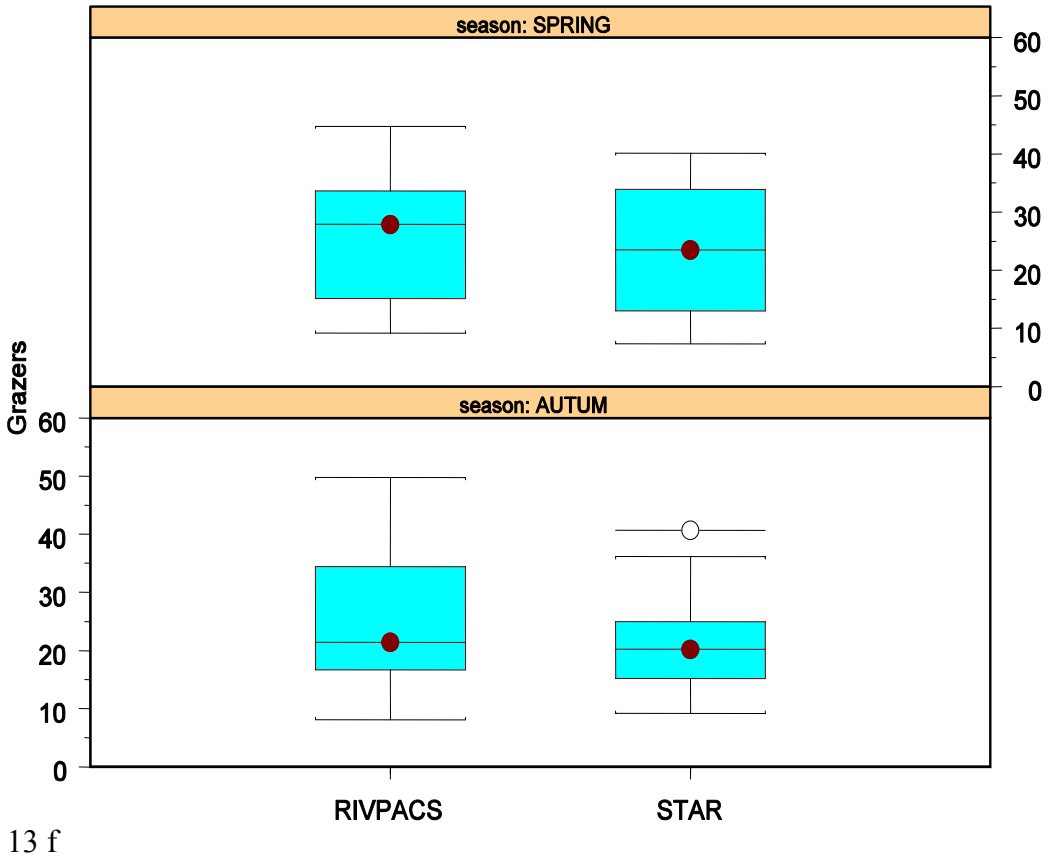
13 c


season: SPRING
season: AUTUM
Number of families
0
10
20
30
40
50
60
RIVPACS
STAR

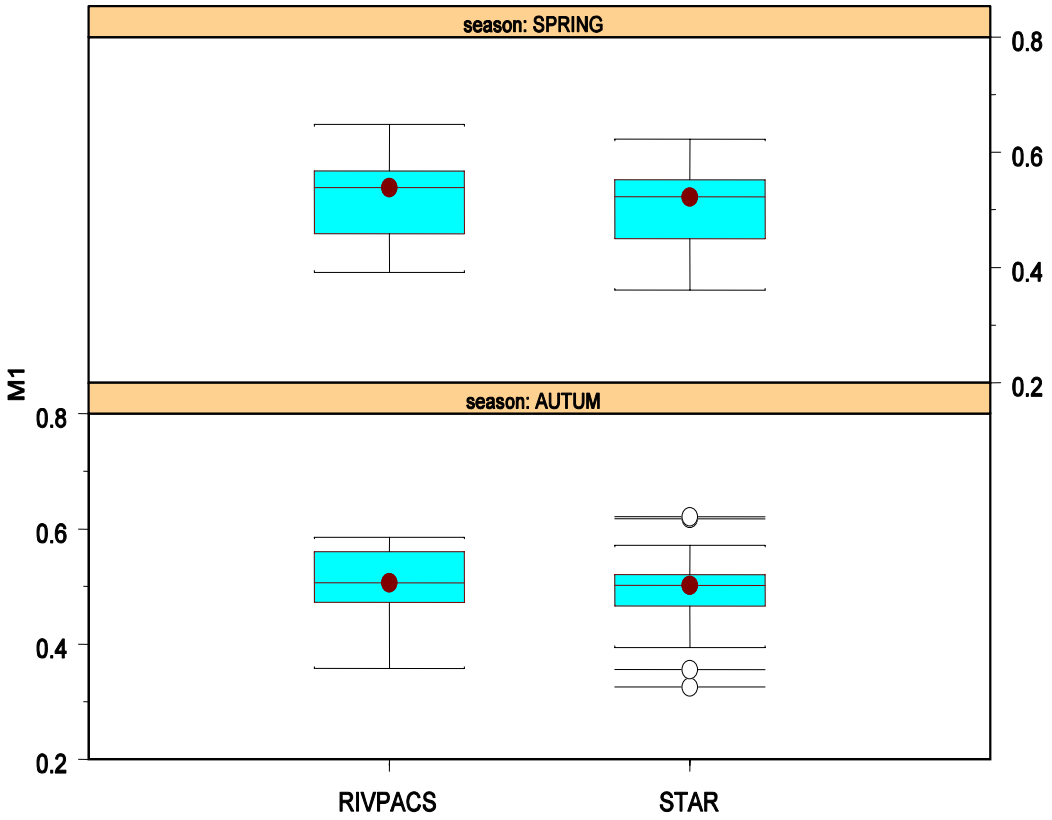
13 d


season: SPRING
season: AUTUM
Collector - gatherers
0
10
20
30
40
50
60
70
80
RIVPACS
STAR

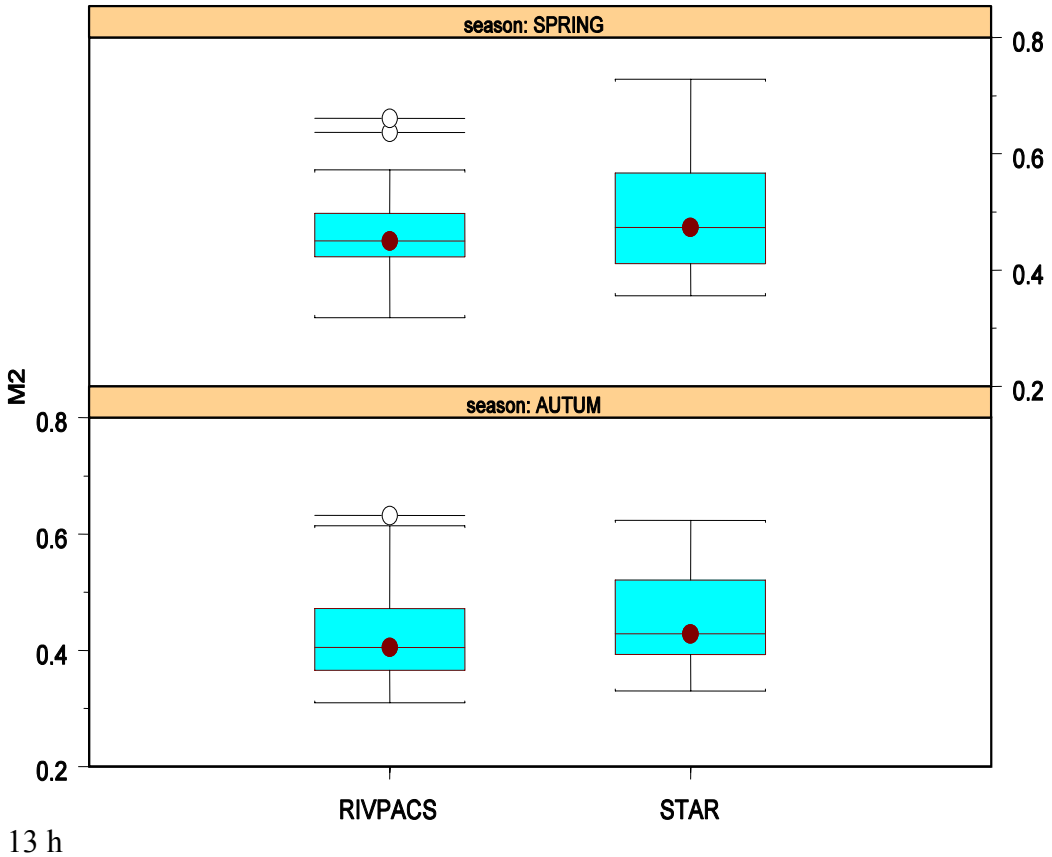
13 e


season: SPRING
season: AUTUM
Grazers
0
10
20
30
40
50
60
RIVPACS
STAR

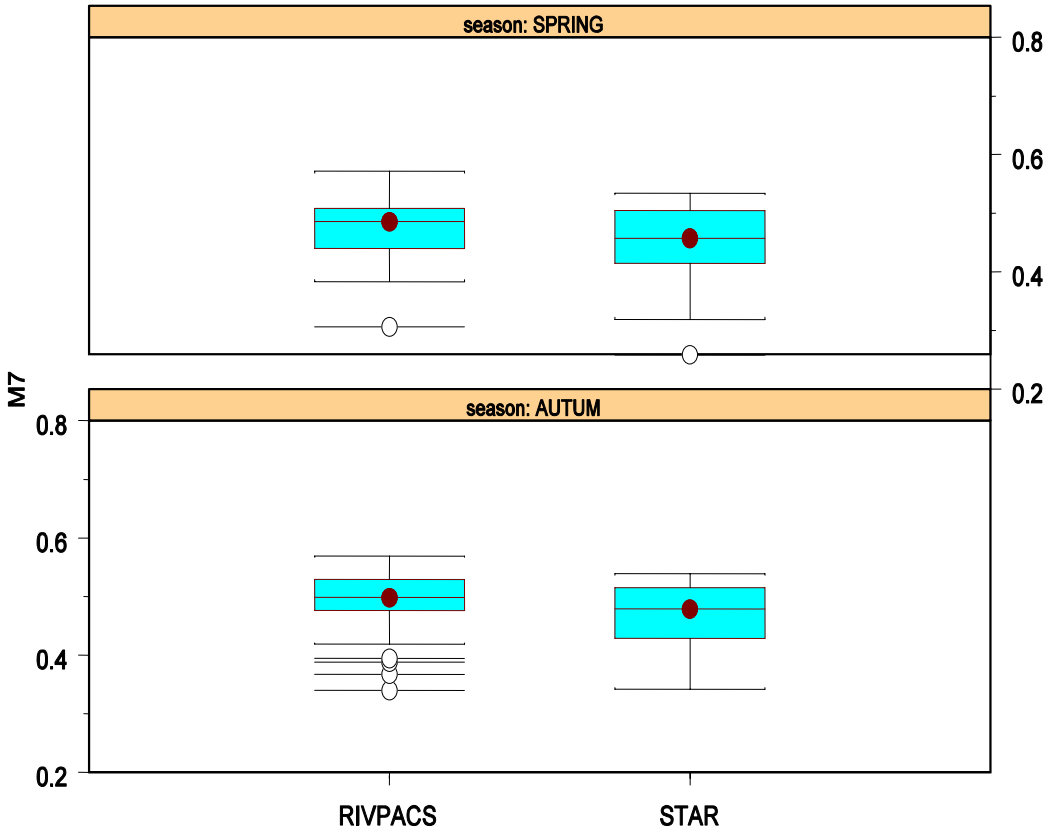
13 f


season: SPRING
season: AUTUM
M1
0.8
0.6
0.4
0.2
RIVPACS
STAR

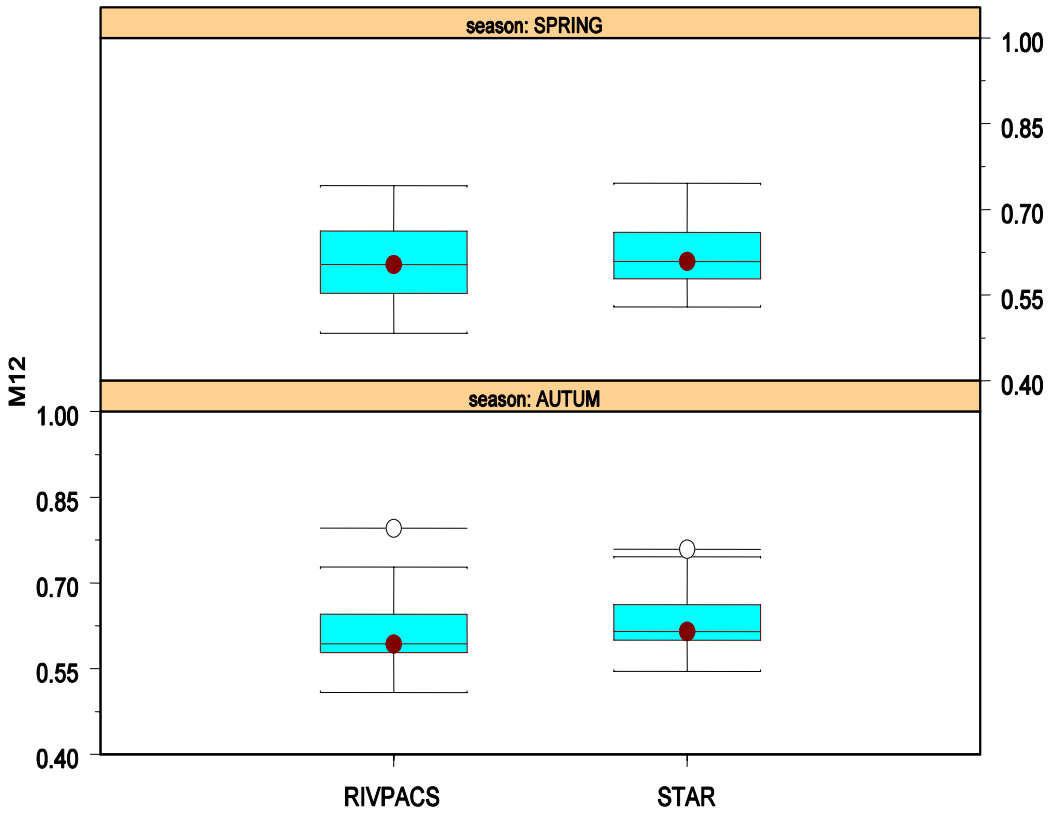
13 g


season: SPRING
season: AUTUM
M2
0.8
0.6
0.4
0.2
RIVPACS
STAR

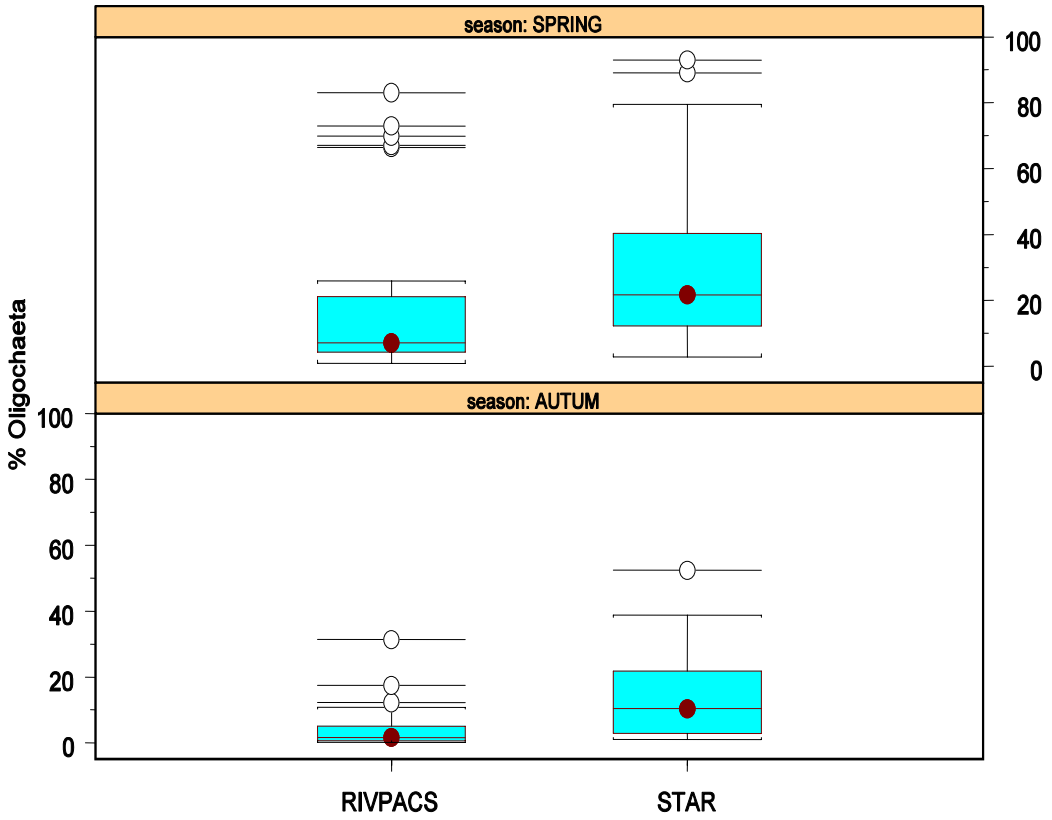
13 h


season: SPRING
season: AUTUM
M7
0.8
0.6
0.4
0.2
RIVPACS
STAR

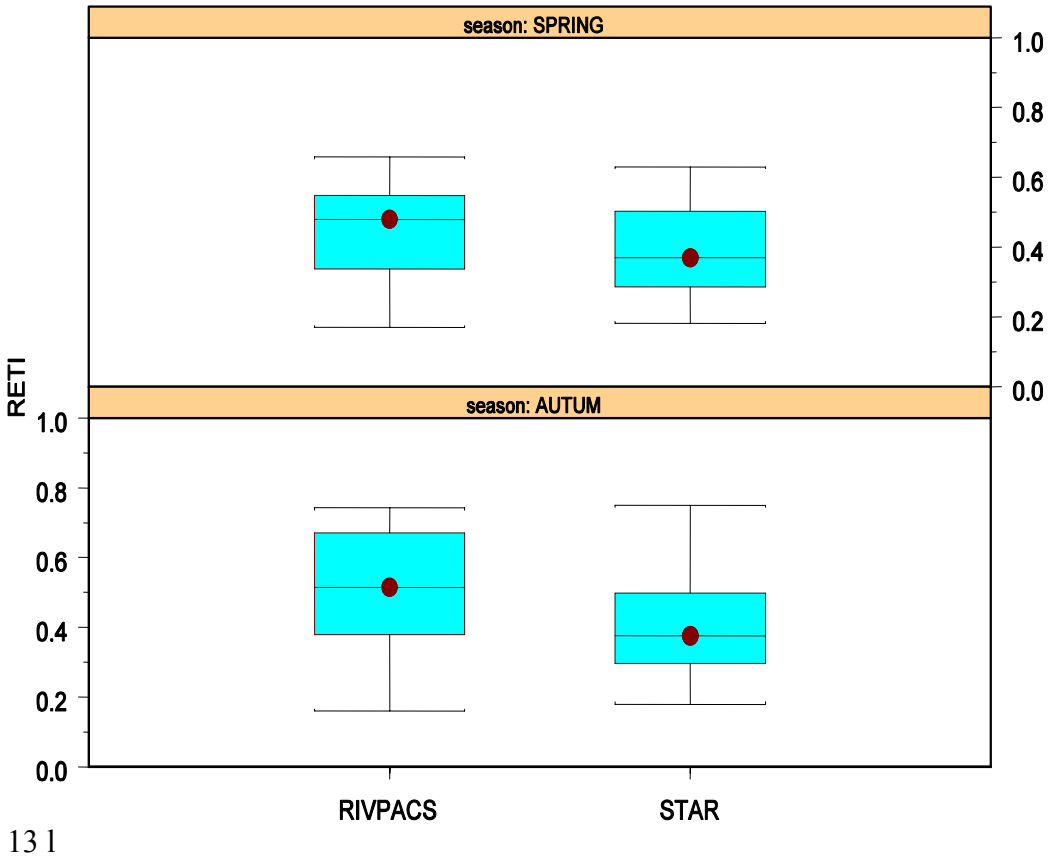
13 i


season: SPRING
season: AUTUM
M12
1.00
0.85
0.70
0.55
0.40
RIVPACS
STAR

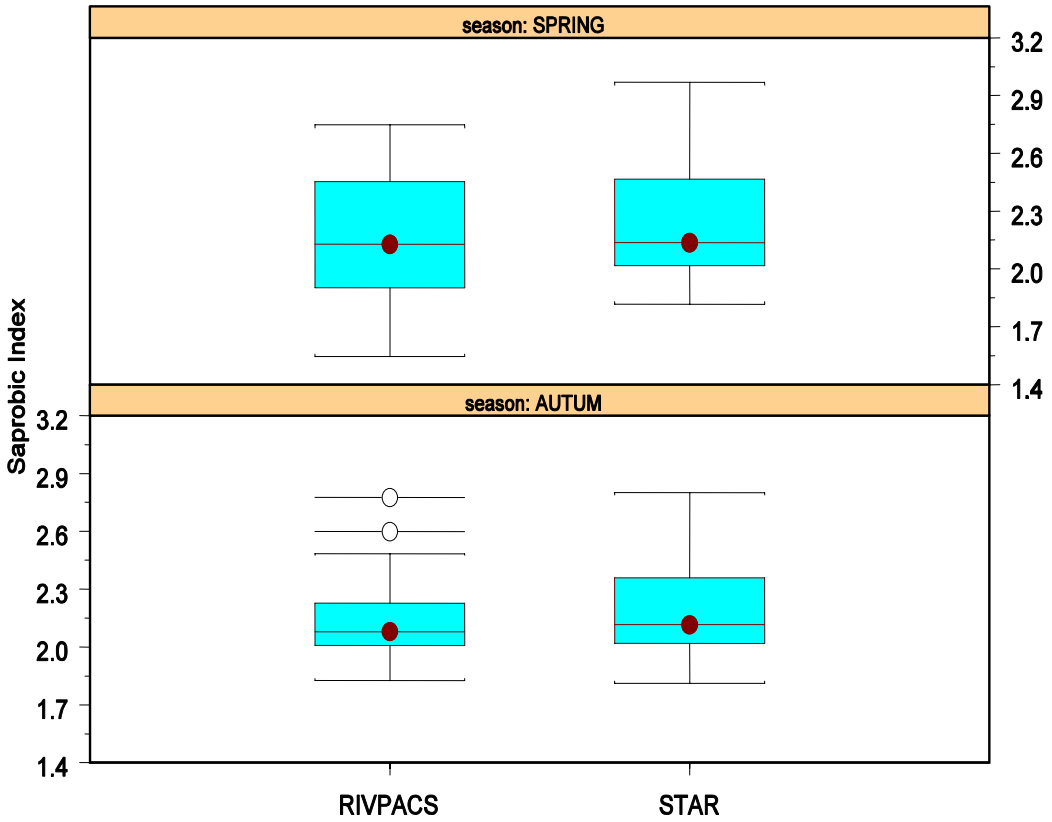
13 j


season: SPRING
season: AUTUM
% Oligochaeta
100
80
60
40
20
0
RIVPACS
STAR

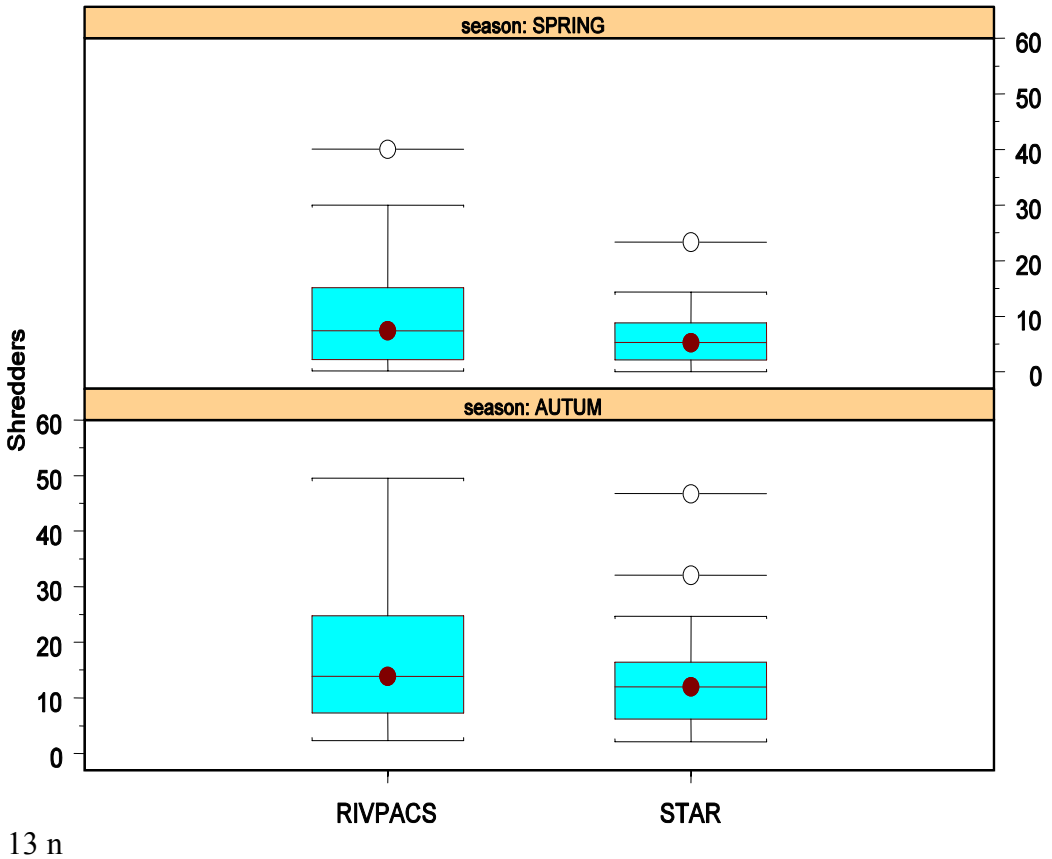
13 k


season: SPRING
season: AUTUM
RETI
1.0
0.8
0.6
0.4
0.2
0.0
RIVPACS
STAR

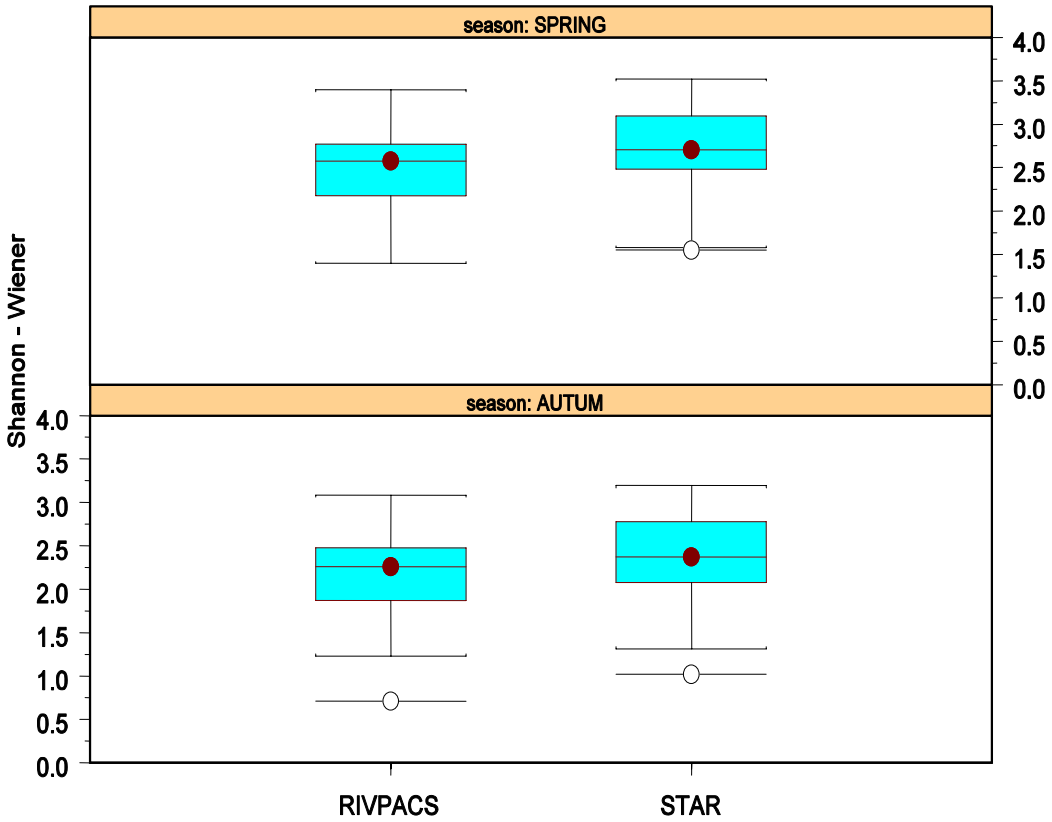
13 l


season: SPRING
season: AUTUM
Saprobic Index
3.2
2.9
2.6
2.3
2.0
1.7
1.4
RIVPACS
STAR

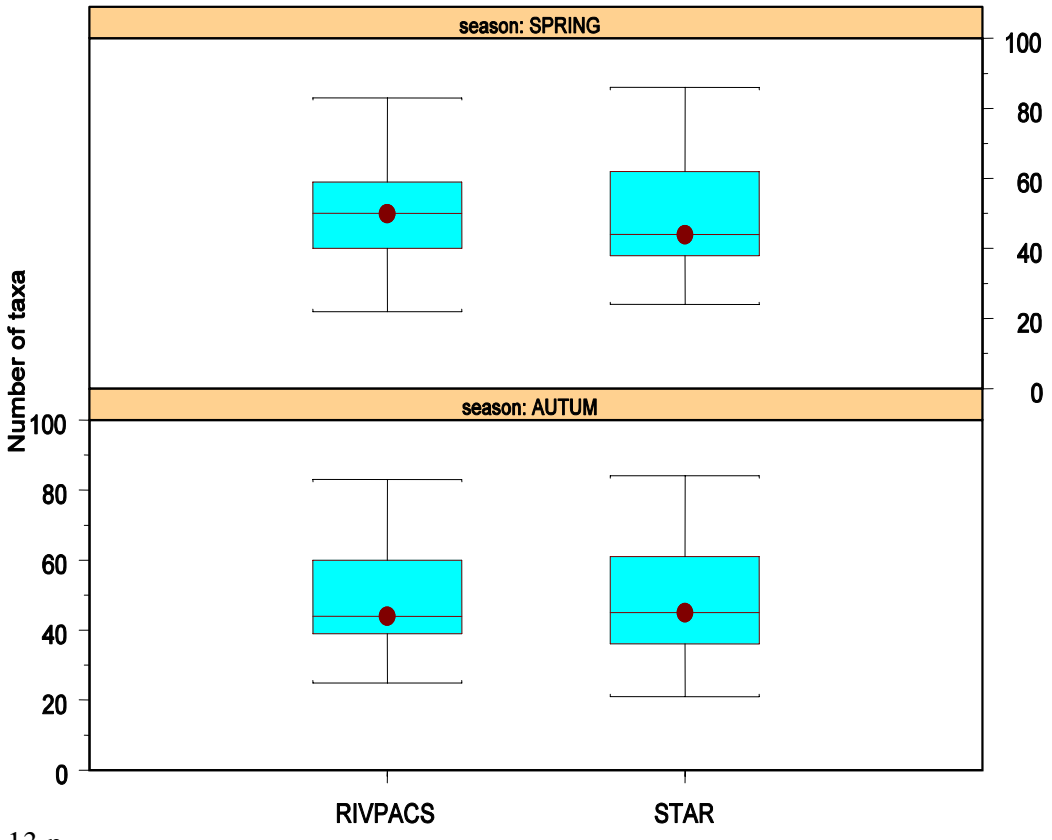
13 m


season: SPRING
season: AUTUM
Shredders
60
50
40
30
20
10
0
RIVPACS
STAR

13 n


season: SPRING
season: AUTUM
Shannon - Wiener
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
RIVPACS
STAR

13 o


season: SPRING
season: AUTUM
Number of taxa
100
80
60
40
20
0
RIVPACS
STAR

13 p