Open access • Journal Article • DOI:10.1007/S10980-007-9101-Y

Spatial patterns of macroinvertebrate functional feeding groups in streams in relation to physical variables and land-cover in Southwestern France

— Source link

Arthur Compin, Régis Céréghino

Institutions: Paul Sabatier University

Published on: 26 Apr 2007 - Landscape Ecology (Springer Netherlands)

Topics: Landscape ecology and Riparian zone

Spatial patterns of macroinvertebrate functional feeding groups in streams in relation to physical variables and land-cover in southwestern France

Arthur Compin, Régis Céréghino

Spatial patterns of macroinvertebrate functional feeding groups in streams in relation to physical variables and land-cover in southwestern France

Arthur COMPIN* and Régis CEREGHINO

EcoLab, Laboratoire d'Ecologie Fonctionnelle, University Paul Sabatier, Toulouse, France

*Corresponding author : EcoLab, Laboratoire d'Ecologie Fonctionnelle, UMR 5245, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse cedex 9, France.

E-mail : compin@cict.fr, Fax: +33 561 556 096, Phone : +33 561 558 436

Date of the manuscript draft: 26th March 2007

Word count: 6328 words

Abstract

Artificial neural networks were used to quantify the distribution of macroinvertebrate functional feeding groups (FFGs) in relation to physical variables and to land-cover in the Adour-Garonne stream system (SW France; 116,000 km²). The relative abundances of five FFGs were calculated from macroinvertebrate data recorded at 165 sampling sites. Each site was characterized using 5 physical variables (elevation, stream order, stream width, distance from the source, slope) and 3 land-cover variables (% forested, % urban areas, % agricultural areas). The sites were first classified using the Self-Organizing Map algorithm (SOM), according to the physical and land-cover variables. Two major clusters of sites corresponded to anthropogenically-modified and natural areas, respectively. Anthropogenically-modified areas were clearly divided into agricultural and urban landscapes. Each major cluster was divided into 3-4 subsets of sites according to a topographic gradient of physical variables. To examine the variability of the communities, FFG proportions at the 165 sites were examined on the SOM trained with physical and land-cover variables. When the riverine landscape was natural, FFG patterns responded to the upstream-downstream gradient in physical variables. When the landscape was altered by agriculture or urbanization, the effects of land-cover on FFGs overcame the influence of the physical variables. The categorization of the landscape into forested, agricultural, and urban areas was relevant to detect changes in FFG patterns. In light of increasing development along riparian zones, the use of SOMs to detect responses of FFGs to landscape alterations at regional scales exemplifies an effective technique for assessing river health based on ecological indicator groups.

Key words : Community structure, freshwater invertebrates, landscape alterations, river continuum, environmental filters, bioassessment.

Introduction

One of the newer approaches to assessing the ecological health of rivers is the Reference Condition Approach (RCA, reviewed in Bailey et al. 2003), where "ecological health" may be defined in terms of similarity to a pristine undisturbed state. Throughout the world, environmental policies aimed at improving or preserving the biological quality of surface waters increasingly rely on RCAs, and will therefore have an impact on a wide range of people and activities such as water consumers and recreational users, agriculture, industry and business activities (Sachon and Wasson 2000). Reference Condition Approaches were recently developed in Europe, Australia, and Canada (Wright et al. 2000). These techniques use classifications of reference sites from rivers of high biological quality to provide system-specific predictions of the fauna to be expected under undisturbed conditions. By knowing what biota should be present in a given geographic zone, one can estimate the degree to which human activity has altered it (Hawkins et al. 2000; Van Sickle and Hughes 2000) because any site can be assessed by comparing its biota (macroinvertebrates, fish, or diatoms) to the reference sites, and any change in expected assemblages can indicate environmental changes in the area. Reference conditions were previously derived from biogeographic models based on species assemblages (RIVPACS, Wright et al. 2000). However, any model referring only to a region-specific fauna (a species list) is more likely to have local acceptance (Céréghino et al. 2003). Conversely, considering communities through their functional attributes rather than in terms of species assemblages *sensu stricto* can provide an approach more applicable to other systems (Santoul et al. 2005). Macroinvertebrates constitute an important part of animal production within freshwater ecosystems, and are tightly integrated into the structure and function of their habitats (organic matter processing, nutrient retention, food resource for amphibians, fish, or birds; Oertli 1993). During recent decades, the categorisation of stream macroinvertebrates by functional feeding group (FFG) based on morphological and behavioural adaptations to acquire the food resources (Merritt and Cummins 1996) has shown considerable promise as a tool for assessing spatial changes in lotic communities based on environmental conditions (Wallace and Webster 1996; Blasius and Merritt 2002).

The theory that describes the influence of the surrounding landscape on aquatic communities was initiated by Hynes (1975), who, arguing that ‘the valley rules the stream’, recognized that rivers and streams are influenced by the landscapes through which they flow. Later, the River Continuum Concept (Vannote et al. 1980) offered a “reference scheme” describing the continuous gradient of physical characteristics of streams from headwaters to mouth and the resulting functional responses under natural conditions. In upland streams, shredders reduce coarse particulate organic matter to fine particulate organic matter, making it available to collectors that are more abundant downstream. Where the stream widens, algae and moss can grow, photosynthesis is expected to be higher in mid-sized streams (but not in large streams), proportions of shredders fall due to lower inputs of riparian coarse particulate organic matter, and grazer-scrapers become more abundant. Further downstream, collectors predominate and shredders are virtually absent. However, the nature of the riparian and surrounding landscape is definitely modified by human activities (agriculture, forestry, urbanization, etc.), which may subsequently alter the FFG composition of macroinvertebrate communities by modifying food availability (Larranaga et al. 2006).

Detrital inputs in streams are an important source of nutrition for many invertebrates, forming a strong trophic link between plant and animal production. Both riparian- and catchment-scale land-cover are likely to explain significant variation in macroinvertebrate diversity (Strayer et al. 2003; Rios and Bailey 2006), the latter being positively affected by increasing the extent of perennial riparian and upland vegetation (Vondracek et al. 2005). Conversely, the alteration of riparian corridors may alter autochthonous contributions to streams, ultimately changing the food quality available to the aquatic communities (Elliott et al. 2004). Anthropogenic alteration of riparian and watershed vegetation is likely to override geomorphological controls on the distribution of macroinvertebrate FFGs (Maridet et al. 1998). Assuming that the biological assemblages of a river ecosystem integrate the spatial and temporal variability of the riparian and aquatic habitat (Townsend and Hildrew 1994), taxa with certain combinations of adaptations are believed to be selected by the dynamics of local habitat conditions (Southwood 1977). Thus, within a given region, the functional structure of macroinvertebrate communities is expected to vary consistently in relation to land-cover patterns.

At regional scales, ecological data such as FFG and environmental variables often vary and covary in a nonlinear fashion (Lek and Guegan 2000). Thus, nonlinear modelling methods such as artificial neural networks (ANNs) should theoretically be preferred for dealing with such data (Blayo and Demartines 1991). Combining clustering and ordering abilities, the Self-Organizing Map algorithm (SOM, unsupervised neural network, Kohonen 1982, 2001) is a powerful tool to visualize high-dimensional data. This technique has shown particular relevance to pattern detection in biological communities in relation to environmental data because the gradient distribution of some biological variables can be visualized in a SOM previously trained with environmental variables (Park et al. 2003). We used a SOM algorithm to interpret the functional variability of the communities with respect to physical and land-cover variables. Self-organizing maps and other ANN techniques have been successfully implemented in various aspects of ecological modelling such as classifying groups (Levine et al. 1996), patterning complex relationships (Tuma et al. 1996), predicting population and community development (Recknagel et al. 1997), modelling habitat suitability (Özesmi and Özesmi 1999), and assessment of water quality (Walley et al. 2000). In this study we used SOM to address two research objectives: (1) to quantify spatial pattern in the

Adour-Garonne stream system (SW France) based on riparian land-cover characteristics and physical variables related to the location along a stream continuum; and (2) to examine how FFG communities respond to riparian land-cover and topographic stream position.

Methods

Study area

The Adour-Garonne stream system (SW France) has a drainage basin area of 116,000 km². Land-cover in this area was historically temperate forest, but is now deeply modified by human activities and development. One-hundred sixty-five sites ranging from high mountain (2500 m a.s.l.) to plain (10 m a.s.l.) areas were sampled (Fig. 1). The substrate of the river beds was dominated by boulders (mean particulate diameter > 200 mm) and pebbles (100 – 200 mm) in mountain areas (> 600 m a.s.l., mean annual discharge ~ 1 to 4 $m^3 s^{-1}$), to stones (20 – 100 mm) and gravel (2 – 20 mm) in piedmont areas (600 – 300 m a.s.l., mean annual discharge ~ 20 to 40 $m^3 s^{-1}$), and sand (0.2 – 2 mm) and silt (< 0.2 mm) in plain areas (< 300 m a.s.l., mean annual discharge ~ 400 $m^3 s^{-1}$).

Field data

For each site, a Geographic Information System (GIS, Mapinfo professional 7.8) was used to arbitrarily delineate a geographical buffer zone including the sampling site, and a [1000 m-long x 100 m-large] riparian corridor located immediately upstream from the site. This size falls within that of the “Reach Buffer” *sensu* Allan (2004), and is well suited to assign a land-cover influence to each site. Sampling sites were then characterized using 5 physical and 3 land-cover variables. The physical variables were elevation a.s.l (m), stream order, stream width (m), distance from the source (km), and slope (‰). The 3 land-cover variables were percent area within a buffer zone covered by forest (areas occupied by forest and woodlands with native or exotic coniferous and/or deciduous trees; scrub and herbaceous vegetation associations), urban development (industrial, commercial and transport units; artificial and non-agricultural vegetated areas), and agricultural (arable lands, permanent crops and pasture). Digital land-cover information was obtained from the CORINE land-cover database for Europe (CLC 2000, European Environment Agency, http://www.eea.europa.eu/; see also Cruiskshank and Tomlison 1996). This database was generated from orthorectified satellite images and provides thematic GIS map layers including up to 44 land-cover classes with a mapping scale of 1:100,000.

These 8 variables were chosen because they characterise the location of sampling sites within the stream system and within the regional landscape mosaic, and they are easy to describe using a GIS. The use of simple variables in a successful final model could reduce the effort and cost of data collection for water management applications.

Each site was sampled twice for macroinvertebrates, in summer (June – August) and winter (December – February), to take into account species seasonality. At each period, eight sample-units were taken from all substratum types using a standard Surber sampler (sampling area 0.1 m², mesh size 0.3 mm). Sample-units were distributed in proportion to the relative abundance of the substrata. This scheme was adapted from the IBGN protocol ("Indice Biologique Global Normalisé", AFNOR 1992), which is the standardised biological index used by French administrations in charge of environmental surveys. It was adapted by Verneaux et al. (1982) from the "Trent Biotic Index" (Woodiwiss 1964), which was originally designed for Great Britain.

In the laboratory, invertebrates were keyed to genus or family and counted. They were then partitioned into "functional feeding groups" (Cummins 1974; Cummins and Klug 1979), based on invertebrate morphological and behaviour adaptations to acquire their food resources (Merritt and Cummins 1996). These groups were: (1) shredders (SH, feed on coarse particulate organic matter > 1000 µm in size); (2) collector-filterers (FC, sift fine particulates of 1000 to 0.45 µm from the flowing column of water); (3) collector-gatherers (GC, gather fine particulates of organic matter from the debris and sediments on the bed of the stream); (4) grazers-scrapers (GS, scrape off and consume the organic layer of algae, microorganisms and dead organic matter attached on stones and other substrates); and (5) predators (PR, feed on other animals). The relative abundances (%) of the various FFG were calculated from their mean density estimates (ind.m^{-2}) obtained by pooling the summer and winter samplings carried out at each site.

Modelling procedure

The SOM Toolbox (version 2) for Matlab® developed by the Laboratory of Information and Computer Science at the Helsinki University of Technology was used (http://www.cis.hut.fi/projects/somtoolbox/, see Vesanto et al. 1999 for practical instructions). The strengths and weaknesses of the SOM and other ANN techniques in comparison with conventional multivariate analyses were discussed in Giraudel and Lek (2001) and Gevrey et al. (2003). The SOM algorithm is an unsupervised learning procedure, which transforms the multi-dimensional input data into a two-dimensional map subject to a topological (neighbourhood preserving) constraint (detailed in Kohonen 2001). Unsupervised (or self-organized) learning means that the ANN receives a number of different input patterns, discovers significant features in these patterns and learns how to classify input data into appropriate categories. The SOM thus plots the similarities of the data by grouping similar data items together, in a way that can be described as follows:

- The virtual samples (visualized as hexagonal cells) are initialized with random samples drawn from the input data set.
- The virtual samples are updated in an iterative way: (1) a sample unit is randomly chosen as an input unit; (2) the Euclidean distance between this sample unit and every virtual sample is computed; (3) the virtual sample closest to the input is selected and called 'best matching unit' (BMU); and (4) the BMU and its neighbours are moved a bit towards the input unit.

The training was broken down into two parts:

- Ordering phase (the 3 000 first steps): when this first phase takes place, the samples are highly modified in a wide neighbourhood of the BMU.
- Tuning phase (7 000 steps): during this phase, only the virtual samples adjacent to the BMU are lightly modified.

At the end of training, the BMU is determined for each sample, and each sample is set in the corresponding hexagon of the SOM map. Neighbouring samples on the grid are expected to represent adjacent clusters of samples. Consequently, sites appearing distant in modelling space (according to physical and land-cover variables) represent expected differences among sites in real environmental characteristics.

In order to bring out relationships between environmental and FFG variables, we introduced the FFG variables (%SH, %GC, %FC, %GS, %PR) into a SOM previously trained with the 5 physical and 3 land-cover variables. The structure of the SOM for our study consists of two layers of neurons connected by weights (or connection intensities): the input layer composed of 13 neurons (one per variable) connected to the 165 sites, and the output layer composed of 66 neurons (see below) visualized as hexagonal cells organized on an array with 11 rows and 6 columns (Fig. 2). However, during the training, we used a mask to give a null weight to the 5 FFG variables, whereas physical and land-cover variables were given a weight of 1 so that the search for the BMU was based on the 8 physical and land-cover variables only. Setting mask value to zero for a given component (here for each of the five FFGs) removes the effect of that component on organization (Vesanto et al. 2000; Vesanto and Hollmen 2003; Sirola et al. 2004; Raivio 2006). The values for FFGs were thus visualized on the SOM previously trained with physical and land-cover variables only. The number of output neurons (map size) is important to detect the deviation of the data. If the map size is too small, it might not explain some important differences that should be detected. Conversely, if the map size is too big, the differences are too small. To select a map size, we followed the procedure described in Park et al. (2003): the network was trained with different map sizes and we chose the optimum size based on local minimum values for quantization and topographic errors. Quantization error is the average distance between each data vector and its BMU and, thus, measures map resolution. Topographic error represents the proportion of all data vectors for which 1st and 2nd BMUs are not adjacent, and is used for the measurement of topology preservation. The number of 66 output neurons retained for this study fitted well the heuristic rule suggested by Vesanto et al. (2000) who reported that the optimal number of map units is close to $5\sqrt{n}$, where n is the number of samples.

Finally, the unified-matrix (U-matrix, Ultsch and Siemon 1990) was used as a distance matrix to identify the cluster boundaries on the SOM map (Vesanto and Hollmen 2003). Significant differences between SOM clusters were tested on the value of FFG variables in the output neurons of the SOM (the 66 hexagons), using t-tests and one-way ANOVAs followed by post-hoc tests (Tukey HSD tests).

Results

Classification of sampling sites

After training the SOM with physical and land-cover variables, the U-matrix helped to derive 7 clear clusters of sampling sites (Fig. 2a). Sites in clusters A1-A3 (top areas of the map) and N0-N3 (bottom areas) corresponded to 2 major land-cover patterns: anthropogenically-modified (A) and natural (N) landscapes, respectively. Anthropogenically-modified areas were clearly divided into agricultural (A1 and A3) and urban (A2) landscapes. Then, each major section of the SOM was divided into 3 or 4 sub-clusters of sites according to physical variables, and, specifically, according to their location within the upstream – downstream continuum. Independent of land-cover, A1 and N1, A2 and N2, A3 and N3 corresponded to mountain, piedmont and plain areas, respectively.

Distribution of macroinvertebrate FFGs

When the distribution of FFGs was visualized on the SOM previously trained with physical and land-cover variables (using a shading scale, see Fig. 3), all functional groups were present in the 7 clusters. Gathering-Collectors was the dominant FFG (55 – 78 %), FC represented 22.7%, while other groups always represented less than 17 % of the invertebrate community. Shredders (SH) had their highest percentages in upland stream sections while collectors (FC, GC) predominated further downstream, a pattern that corresponds well with the River Continuum Concept model. Predators (PR) and scrapers (SC) did not show clear patterns along the river system, as they predominated in clusters A1, A3 and N0.

The range of FFG proportions differed among clusters (Fig. 3, Table 1). In anthropogenically-modified areas, there were significant differences in the relative abundance of all FFGs except FC between agricultural (A1 and A3) and urban (A2) landscapes. Gathering-collectors had their highest percentages in urban landscapes, while SC and PR (and to a lesser extent SH) dominated in agricultural landscapes. In natural landscapes, there was no significant difference in %GC among clusters. Other FFGs showed a longitudinal distribution pattern, with PR, SC and SH dominating in upland areas (clusters N0 and N1) and FC dominating in downstream areas (N3).

Finally, there were significant differences in the relative abundance of all five FFGs between pairs of Anthropogenically-modified *vs* Natural clusters (Fig. 4). In anthropogenically-modified landscapes, higher %GC values exceeded those observed in natural environments. Conversely, SH and FC showed higher percentage values in natural landscapes, compared to anthropogenically-modified landscapes at similar locations within the stream system.

Discussion

Land cover influences the chemical and biological characteristics of river ecosystems (Moore and Palmer 2005) and the structure of lotic macroinvertebrate communities may be subsequently influenced by land-use practices within catchments (Sponseller et al. 2001). For example, cleared riparian vegetations are detrimental to shredders as leaf litter input is reduced (Davies et al. 2005; Lecerf et al. 2005). Fertilizer runoff entering a stream enhances the development of periphytic algae thus favouring grazer-scrapers (Delong and Brusven 1998). On the other hand, phosphorus fertilization can also increase abundance of moss, and moss abundance may cause displacement of scrapers (Slavik et al. 2004). Gathering-collectors seem to be the only group able to find sufficient food in urban streams (Suren and McMurtrie 2005). When the riverine landscape was natural, we found that FFG patterns mainly responded to the upstream-downstream gradient in physical conditions, as predicted by the River Continuum Concept (Figs 2 and 3). However, when the riverine landscape was altered by agriculture or urbanization, there were significant differences between urban and agricultural clusters of sites in terms of FFG distributions. These results suggest that the effects of anthropogenic activities on the functional structure of macroinvertebrate communities may overcome the influence of the upstream-downstream gradient of the physical variables in streams (Fig. 4, Table 1).

We demonstrated that FFG proportions responded to broad land-cover categories: "forested area"; "agricultural area"; and "urban area" (Figs 2, 3). To assign a land-cover influence to a given site, the "local reach" is assumed to be the appropriate size for the buffer zone. It is described by Allan (2004) as a buffer of 100 m to several hundred meters in width on each bank, and some hundreds of meters to a kilometer in length. In practice, the width of the buffer zone is often adapted to the landscape characteristics of the studied areas, and ranges from 30 m (Basnyat et al. 2999; Rios and Bailey 2003) to 100 m on each side of the river (Sliva and Williams 2001). Our results suggest that the 1000 m-long x 100 m-large buffer zone was appropriate to detect changes in FFG compositions in a large stream system. Therefore, our approach should be relevant to identify and to delineate areas of concern in integrated management at watershed levels. At the site scale, these categories are almost certainly too large to allow accurate differentiations of the effects of land-cover on stream communities. Existing tools such as functional indices using FFG distributions would be relevant at such a spatial scale, however, we suggest that the local response of FFGs to more specific sub-categories of land-cover should be further investigated. For instance, Dolédec et al. (2006) recently analysed the effects of agricultural development (ungrazed to extensively grazed pasture, cattle farming) on the species traits of invertebrates from a grassland stream, and observed that species traits helped to differentiate the consequences of land-cover intensification in stream communities.

In undisturbed rivers, all FFGs are present irrespective of the river section studied and its geographical region (Bij de Vaate and Pavluk 2004). This scheme concurs with our own results. Therefore, on a local scale (a stream section), using upstream reference sites in a river system and then comparing spatial and/or temporal patterns of FFG abundances and/or proportions under natural and disturbed conditions remains of value to assess ecosystem degradation (Merritt et al. 2002). Relevant tools based on functional classifications of macroinvertebrates were already designed to this end – for instance, the Index of Trophic Completeness (Pavluk et al. 2000), a quality score based on the presence of twelve trophic guilds (defined by the diet, the feeding behaviour, and the food size) in benthic invertebrate communities. However, at the river to stream system scale, most European freshwaters are impacted by human activities, which lead to losses of taxa and/or discontinuities in the distribution of the fauna. Specifically, the Serial Discontinuity Concept (Ward and Stanford 1983) described well the modifications in abiotic and/or biotic parameters due to disturbance in an affected river section. The discontinuity can be "negative" (modifications towards upstream conditions) or "positive" (modifications towards downstream conditions). Landscape alterations influence successional schemes (this study), and therefore functional processes in river ecosystems (Ward 1998). Since the challenge of recent applied research is to assess models having the broadest capability of predicting spatial patterns of community organization (see the European Water Framework Directive, detailed in Sachon and Wasson 2002), this situation raises concerns about the possibility to develop reference schemes based on the functional structure of macroinvertebrate communities on a broad scale, because there is little chance to find a river which fits the River Continuum Concept along its whole course.

If stream management is needed to maintain or restore freshwater biodiversity, our study supports the idea that action plans should be designed at a landscape scale (Ward 1998). Therefore, to be effective, management efforts should be based on explicit spatial distribution schemes. In light of development along riparian zones, our ability to detect responses of FFGs to landscape alterations at regional scales exemplifies a cost-effective technique for assessing river health based on ecological indicator groups.

Acknowledgements

We wish to thank Dr. D. Anderson and three anonymous Reviewers for their constructive comments on an earlier version of this paper.

References

A.F.N.O.R. (1992) Essais des eaux. Détermination de l'Indice Biologique Global Normalisé (I.B.G.N.). Norme française T 90–350:1–8

Allan JD (2004) Landscapes and riverscapes: the influence of land–use on stream ecosystems. Annu Rev Ecol Evol Syst 35:257–284

Bailey RC, Norris RH, Reynoldson TB (2003) Bioassessment of freshwater ecosystems using the reference condition approach. Kluwer Academic Publishers, New York, USA

Basnyat P, Teeter LD, Flynn KM, Lockaby BG (1999) Relationships between landscape characteristics and nonpoint source pollution inputs to coastal estuaries. Environe Manage 23:539–549

Bij de Vaate A, Pavluk TI (2004) Practicability of the Index of Trophic Completeness for running waters. Hydrobiologia 519:49–60

Blasius BJ, Merritt RW (2002) Field and laboratory investigations on the effects of road salt (NaCl) on stream macroinvertebrate communities. Environ Pollut 120:219–231

Blayo F, Demartines P (1991) Data analysis: how to compare Kohonen neural networks to other techniques ? In: Prieto A (ed.) Artificial Neural Networks. International Workshop IWANN '91. Springer-Verlag, Berlin, Germany, pp 469–476

Céréghino R, Park YS, Compin A, Lek S (2003) Predicting the species richness of aquatic insects in streams using a limited number of environmental variables. J N Am Benthol Soc 22:442–456

Cruickshank MM, Tomlison RW (1996) Application of CORINE land cover methodology to the UK. Some issues raised from Northern Ireland. Global Ecol Biogeogr 4/5:235–248

Cummins KW (1974) Structure and function of stream ecosystems. Bioscience 24:631–641

Cummins KW, Klug MJ (1979) Feeding Ecology of Stream Invertebrates. Ann Rev Ecol Syst 10:147–172

Davies PE, Cook LSJ, McIntosh PD, Munks SA (2005) Changes in stream biota along a gradient of logging disturbance, 15 years after logging at Ben Nevis, Tasmania Forest. Ecol Manage 219:132–148

Delong MD, Brusven MA (1998) Macroinvertebrate community structure along the longitudinal gradient of an agriculturally impacted stream. Environ Manage 22:445–457

Dolédec S, Phillips N, Scarsbrook M, Riley RH, Townsend CR (2006) Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. J N Am Benthol Soc 25:44–60

Elliott SR, Naiman RJ, Bisson PA (2004) Riparian influences on the biophysical characteristics of seston in headwater streams. Northwest Sci 78:150–157

Gevrey M, Dimopoulos L, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 160:249–264

Giraudel JL, Lek S (2001) A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. Ecol Model 146(1–3):329–339

Hawkins CP, Norris RH, Gerritsen J, Hughes RM, Jackson SK, Johnson RK, Stevenson RJ (2000) Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. J N Am Benthol Soc 19:541–556

Hynes HBN (1975) The stream and its valley. Verh Int Ver Theor Ang Limnol 19:1–15

Kohonen T (1982) Self-organized formation of topologically correct feature maps. Biol Cyber 43:59–69

Kohonen T (2001) Self-Organizing Maps, 3rd edn. Springer-Verlag, Berlin, Germany

Larranaga A, Larranaga S, Basaguren A, Elosegi A, Pozo J (2006) Assessing impact of eucalyptus plantations on benthic macroinvertebrate communities by a litter exclusion experiment. Ann Limnol – Int J Lim 42:1–8

Lecerf A, Dobson M, Dang CK, Chauvet E (2005) Riparian plant species loss alters trophic dynamics in detritus-based stream ecosystems. Oecologia 146:432–442

Lek S, Guégan JF (2000). Artificial Neuronal Networks: application to ecology and evolution. Springer-Verlag, Berlin, Germany

Levine ER, Kimes DS, Sigillito VG (1996) Classifying soil structure using neural networks. Ecol Model 92:101–108

Maridet L, Wasson JG, Philippe M, Amoros C, Naiman RJ (1998) Trophic structure of three streams with contrasting riparian vegetation and geomorphology. Archiv Hydrobiol 144:61–85

Merritt RW, Cummins KW (1996) An introduction to the aquatic insects of North America, 3rd edn. Kendall/Hunt, Dubuque, USA

Merritt RW, Cummins KW, Berg MB, Novak JA, Higgins MJ, Wessell KJ, Lessard JL (2002) Development and application of a macroinvertebrate functional-group approach in the bioassessment of remanant river oxbows in southwest Florida. J N Am Benthol Soc 21:290–310

Moore AA, Palmer MA (2005) Invertebrate biodiversity in agricultural and urban headwater streams: Implications for conservation and management. Ecol Appl 15:1169–1177

Oertli B (1993) Leaf litter processing and energy flow through macroinvertebrates in a woodland pond (Switzerland). Oecologia 96:466–477

Özesmi S, Özesmi U (1999) An artificial neural network approach to spatial habitat modelling with interspecific interaction. Ecol Model 116:15–31

Park YS, Céréghino R, Compin A, Lek S (2003) Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. Ecol Model 160:265–280

Pavluk TI, Bij de Vaate A, Leslie HA (2000) Development of an Index of Trophic Completeness for benthic macroinvertebrate communities in flowing waters. Hydrobiologia 427:135–141

Raivio K (2006) Analysis of soft handover measurements in 3G network. In: Proceedings of the 9th ACM international Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems (Terromolinos, Spain, October 02 - 06, 2006), MSWiM '06. ACM Press, New York, pp 330–337

Recknagel F, French M, Harkonen P, Yabunaka KI (1997) Artificial neural network approach for modelling and prediction of algal blooms. Ecol Model 96:11–28

Rios SL, Bailey RC (2006) Relationship between riparian vegetation and stream benthic communities at three spatial scales. Hydrobiologia 553:153–160

Sachon G, Wasson JG (2002) La directive Eau de l'Union européenne. Conséquences pour la recherche. Nat Sci Soc 10:93–95

Santoul F, Cayrou J, Mastrorillo S, Céréghino R (2005) Spatial patterns of the biological traits of freshwater fish communities in S.W. France. J Fish Biol 66:301–314

Sirola M, Lampi G, Parviainen J (2004) Using self-organizing map in a computerized decision support system. In: Pal NR, Kasabov N, Mudi RK, Pal S, Parui SK (eds) Neural Information Processing, 11th International Conference, ICONIP 2004, Calcutta, India, November 22-25, 2004, Proceedings. Springer-Verlag, Berlin, Germany, pp 136–141

Slavik K, Peterson BJ, Deegan LA, Bowden WB, Hershey AE, Hobbie JE (2004) Long-term responses of the Kuparuk River ecosystem to phosphorus fertilization. Ecology 85:938–954

Sliva L, Williams DD (2001) Buffer zone versus whole catchment approaches to studying land use impact on river water quality. Wat Res 35:3462–3472

Southwood TRE (1977) Habitat, the templet for ecological strategies? J Anim Ecol 46:337–365

Sponseller RA, Benfield EF, Valett HM (2001) Relationships between land use, spatial scale and stream macroinvertebrate communities. Freshwater Biol 46:1409–1424

Strayer DL, Beighley RE, Thompson LC, Brooks S, Nilsson C, Pinay G, Naiman RJ (2003) Effects of land cover on stream ecosystems: roles of empirical models and scaling issues. Ecosystems 6:407–23

Suren AM, McMurtrie S (2005) Assessing the effectiveness of enhancement activities in urban streams: II. Responses of invertebrate communities. Riv Res Appl 21:439–453

Townsend CR, Hildrew AG (1994) Species traits in relation to a habitat templet for river systems. Freshwater Biol 31:265–275

Tuma A, Haasis HD, Rentz O (1996) A comparison of fuzzy expert systems, neural networks and neuro-fuzzy approaches controlling energy and material flows. Ecol Model 85:93–98

Ultsch A, Siemon HP (1990) Kohonen's self organizing feature maps for exploratory data analysis. In: Proceeding of the INNC'90 International Neural Network Conference. Kluwer Academic Publishers, Dordrecht, The Nederlands, pp 305–308

Van Sickle J, Hughes RM (2000) Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. J N Am Benthol Soc 19:370–384

Vannote RL, Minshall GW, Cummins KW, Sedell JR, Cushing CE (1980) The River Continuum Concept. Can J Fish Aquat Sci 37:130–137

Verneaux J, Galmiche P, Janier F, Monot A (1982) Une nouvelle méthode pratique d'évaluation de la qualité des eaux courantes. Un indice biologique de qualité générale. Ann Sci Univ Franche Comté 4:11–21

Vesanto J, Hollmen J (2003) An automated report generation tool for the data understanding phase. In: Abraham A, Jain L (eds) Innovations in Intelligent Systems: Design, Management and Applications, Studies in Fuzziness and Soft Computing, chapter 5. Springer-Verlag, Berlin, Germany

Vesanto J, Himberg J, Alhoniemi E, Parhankangas J (1999) Self-Organising Map in Matlab: the SOM Toolbox. In: Proceedings of the Matlab Digital Signal Processing Conference. Espoo, Finland, pp 35–40

Vesanto J, Himberg J, Alhoniemi E, Parhankangas J (2000) SOM Toolbox for Matlab 5. Technical Report A57, Neural Networks Research Centre,Helsinki University of Technology, Helsinki, Finland

Vondracek B, Blann KL, Cox CB, Nerbonne JF, Mumford KF, Nerbonne BA, Sovell LA, Zimmerman JKH (2005) Land use, spatial scale, and stream systems: Lessons from an agricultural region. Environe Manage 36:775–791

Wallace JB, Webster JR (1996) The Role of Macroinvertebrates in Stream Ecosystem Function. Ann Rev Entomol 41:115–139

Walley WJ, Martin RW, O'Connor MA (2000) Self-organising maps for classification of river quality from biological and environmental data. In: Denzer R, Swayne DA, Purvis M, Schimak G (eds) Environmental software systems: environmental information and decision support. IFIP Conference Series, Kluwer Academic Publishers, Boston, Massachusset, USA, pp 27–41

Ward JV (1998) Riverine landscapes: Biodiversity patterns, disturbance regimes, and aquatic conservation. Biol Conserv 83:269–278

Ward JV, Stanford JA (1983) The intermediate disturbance hypothesis: an explanation for biotic diversity patterns in lotic systems. In: Fontaine TD, Bartell SM (eds) Dynamics of lotic ecosystems. Ann Arbor Sciences, Ann Arbor, Michigan, USA, pp 347–356

Woodiwiss FS (1964) The biological system of stream classification used by the Trent river Board. Chemistr and Industry 11:443–447

Wright JF, Sutcliffe DW, Furse MT (2000) Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK

Figure and tables

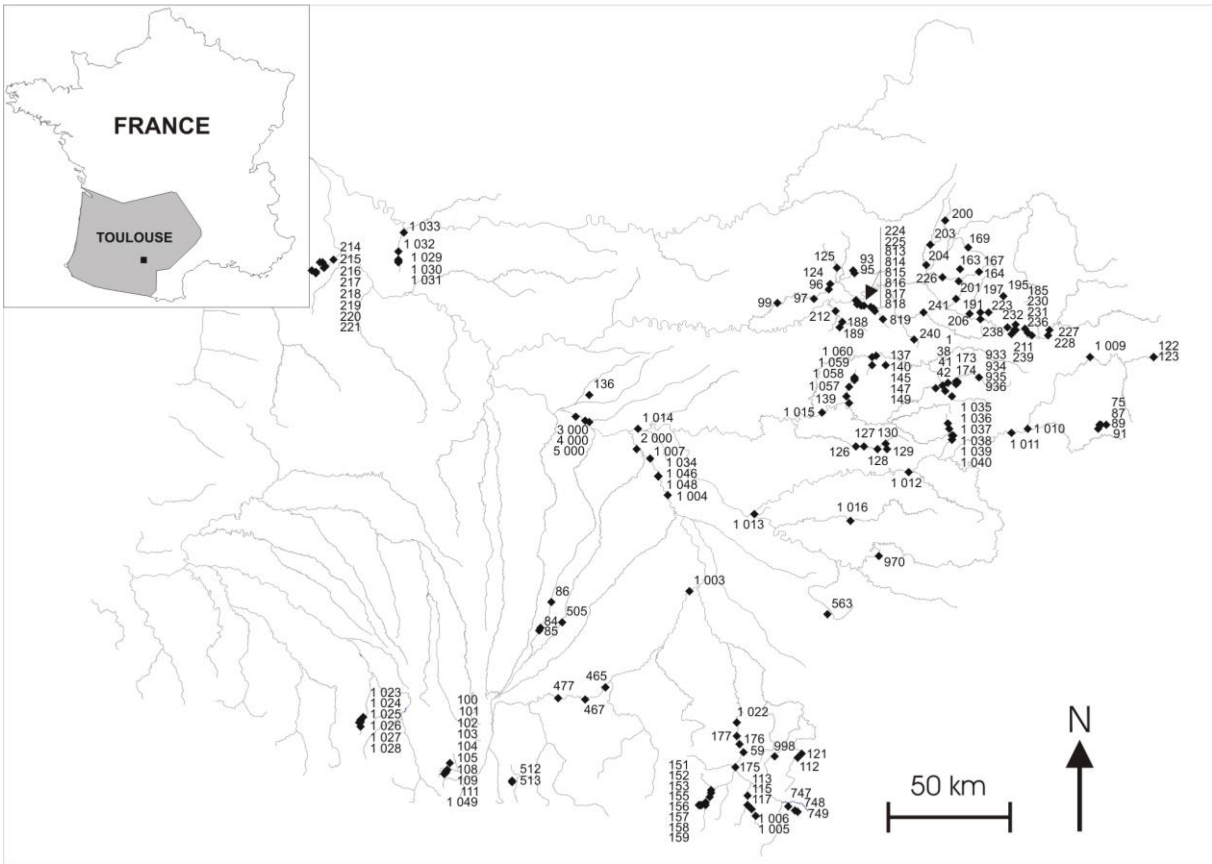

Figure 1. Map of the Adour Garonne stream system, and location of the 165 sampling sites, shown by squares and adjacent numbers.

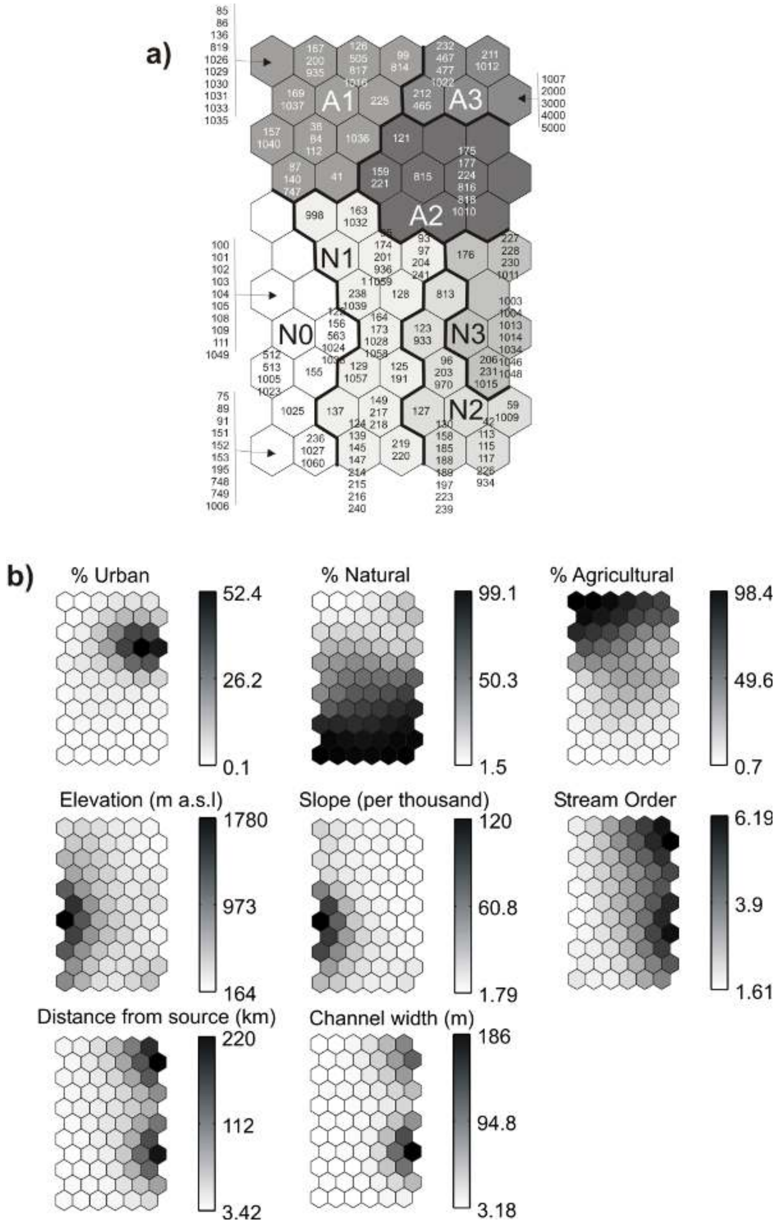

Figure 2. a) Distribution of sampling sites on the self-organising map (SOM) according to the 3 land-cover and 5 physical variables, and clustering of the trained SOM. Codes correspond to sampling sites (Fig. 1). Clusters A1 – N3 were derived from the Unified-matrix clustering. A = Anthropogenically-modified areas, N = Natural areas. Sites that are neighbours within clusters are expected to have similar features. b) Gradient analysis of each variable on the trained SOM, with visualization in shading scale (dark = high values, light = low values). The ordinate of the SOM represents the gradient of anthropogenic modification (from low [bottom] to high [top]), whereas the abscissa of the map represents the gradients of stream order, distance from the source, and stream width (from low [left] to high [right]), and the slope and elevation a.s.l. (from low [right] to high [left]). With the exception of the cluster N0 (high mountain sites, no persistent human presence), pairs of A and N clusters (A1 and N1, A2 and N2, A3 and N3) represent river sections with similar position in the upstream-dowstream gradient of physical conditions, but correspond to different land-use characteristics (agricultural or urban *vs* natural).

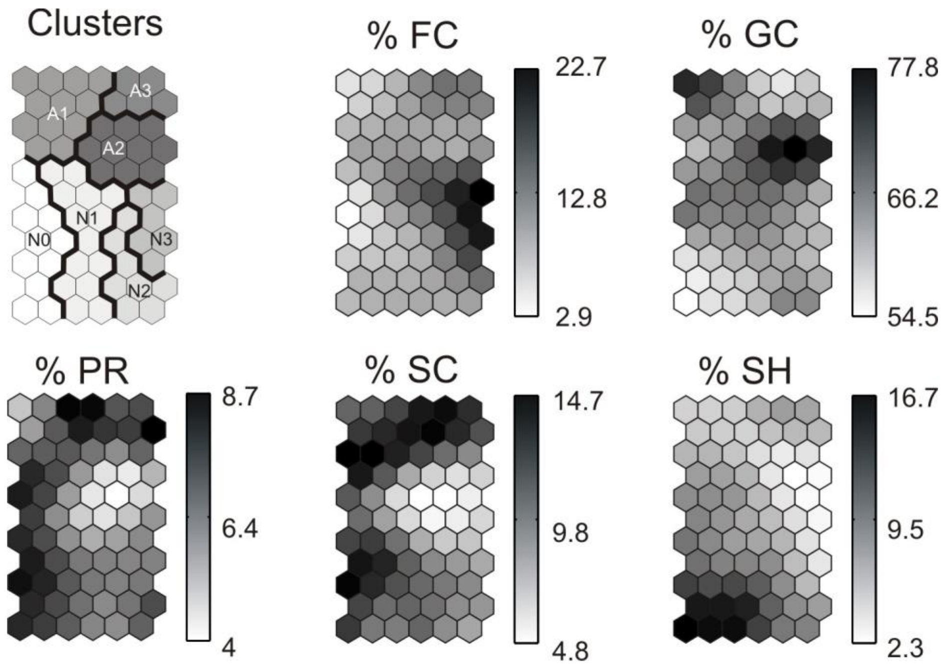

Figure 3. Visualization of Functional Feeding Group percentage compositions on the Self-Organizing Map trained trained with land-cover and physical variables. The mean value of each variable was calculated in each output neuron of the trained SOM. Dark represents a high value, while light is low. SH = Shredders, FC = collector-filterers, GC = collector-gatherers, GS = grazers-scrapers, PR = predators.

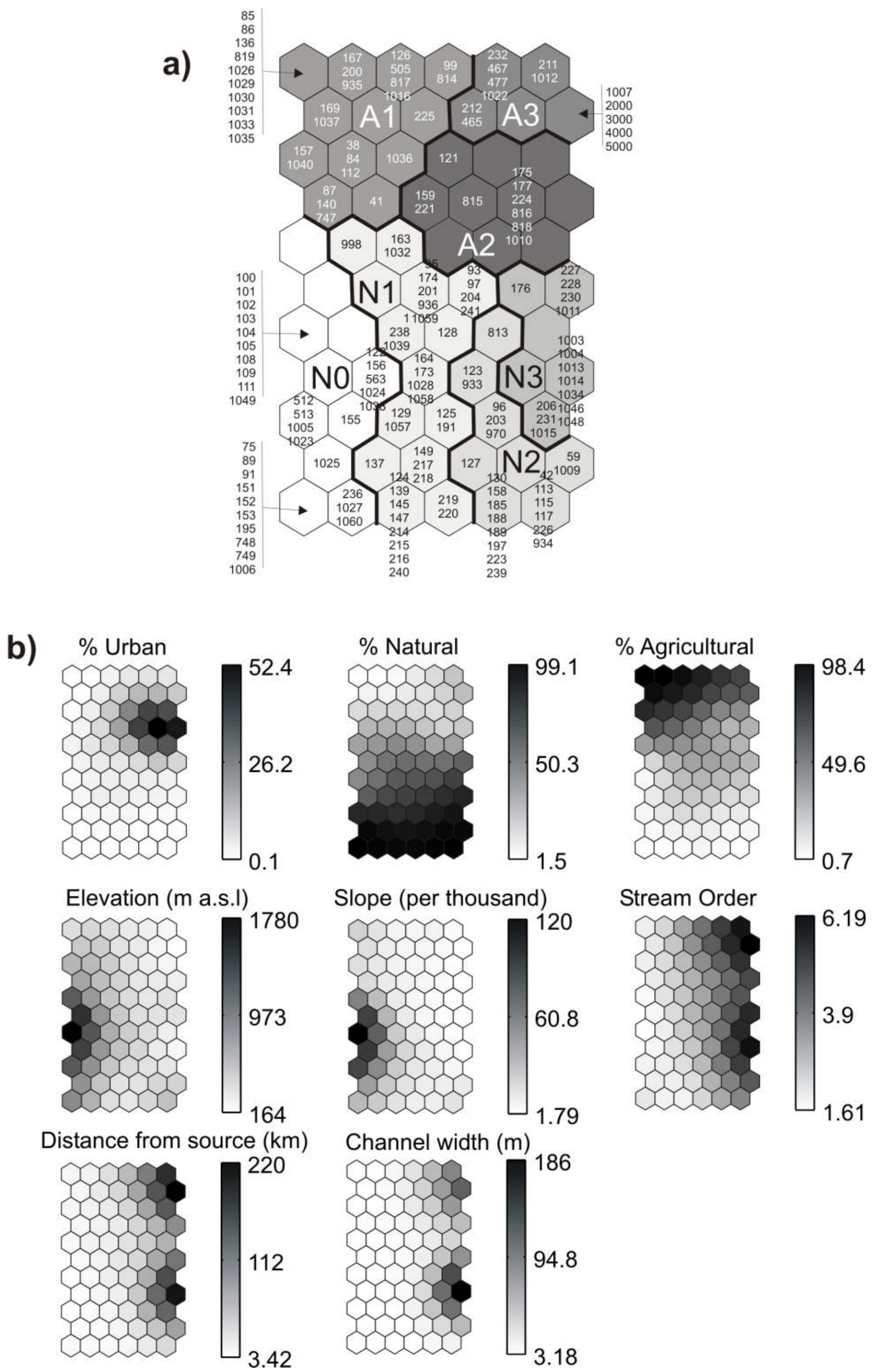

Figure 4. Boxplots of FFG percentage distributions in the seven clusters derived from the SOM analysis, with comparison of pairs of A and N clusters (A1 and N1, A2 and N2, A3 and N3, see text and Fig. 3). n= number of SOM units (hexagons) per cluster. White boxes correspond to forested areas, grey boxes are for agricultural sites, and black boxes are for urban areas. The top, mid- and bottom-line of each box-plot represent the 75th, 50th and 25th percentiles, respectively; the horizontal lines represent the 10th and 90th percentiles. NS: non-significant, *p < 0.05, **p < 0.001, ***p< 0.0001. SH = shredders, FC = collector-filterers, GC = collector-gatherers, GS = grazers-scrapers, PR = predators.

Table 1. Results of one-way ANOVAs testing for differences in FFG percentage distributions among SOM clusters for anthropogenically-modified (clusters A1 – A3) and natural (clusters N0- N3) areas. Post-hoc tests (Tukey Honest Significant Difference) were applied for means comparison when differences were significant ($p \leq 0.05$). For all selected post-hoc procedures, homogeneous subsets are defined. The means for each level of the independent variable are listed in their corresponding homogenous subset. SH = shredders, FC = collector-filterers, GC = collector-gatherers, GS = grazers-scrapers, PR = predators.

Anthropogenically-modified (A1 – A3)

One-way ANOVAs

FFG		Sum of Squares	df	Mean Square	F	Sig.
FC	Between Groups	0.008	2	0.004	9.050	0.001
	Within Groups	0.010	24	0.000		
	Total	0.018	26			
GC	Between Groups	0.054	2	0.027	19.653	0.000
	Within Groups	0.033	24	0.001		
	Total	0.088	26			
PR	Between Groups	0.003	2	0.001	13.873	0.000
	Within Groups	0.002	24	0.000		
	Total	0.005	26			
SC	Between Groups	0.019	2	0.010	28.377	0.000
	Within Groups	0.008	24	0.000		
	Total	0.027	26			
SH	Between Groups	0.003	2	0.002	11.381	0.000
	Within Groups	0.003	24	0.000		
	Total	0.007	26			

Tukey tests

FFG	subset*	Cluster A1	A2	A3	Sig.
FC	2		0.115	0.128	0.385
	1	0.087			1
GC	3		0.719		1
	2	0.654			1
	1			0.595	1
PR	2	0.071		0.077	0.505
	1		0.053		1
SC	2	0.129		0.132	0.922
	1		0.075		1
SH	2	0.059		0.067	0.417
	1		0.039		1

Natural (N0 – N3)

One-way ANOVAs

FFG		Sum of Squares	df	Mean Square	F	Sig.
FC	Between Groups	0.066	3	0.022	41.522	0.000
	Within Groups	0.019	35	0.001		
	Total	0.085	38			
GC	Between Groups	0.004	3	0.001	1.266	0.301
	Within Groups	0.041	35	0.001		
	Total	0.045	38			
PR	Between Groups	0.002	3	0.001	16.737	0.000
	Within Groups	0.001	35	0.000		
	Total	0.003	38			
SC	Between Groups	0.011	3	0.004	14.025	0.000
	Within Groups	0.009	35	0.000		
	Total	0.020	38			
SH	Between Groups	0.022	3	0.007	10.176	0.000
	Within Groups	0.026	35	0.001		
	Total	0.048	38			

Tukey tests

FFG	subset*	Cluster N0	N1	N2	N3	Sig.
FC	3				0.197	1
	2		0.106	0.119		0.637
	1	0.067				1
GC	1	0.612	0.635	0.638	0.621	0.414
PR	2	0.078				1
	1		0.062	0.064	0.060	0.604
SC	2	0.123				1
	1		0.092	0.091	0.076	0.158
SH	2	0.118	0.104	0.088		0.111
	1				0.045	1

*(for alpha=0.05)