TEMPORAL VARIABILITY OF LOTIC MACROINVERTEBRATE COMMUNITIES

Robert Alan Holland

Submitted for the degree of Doctor of Philosophy
Department of Animal and Plant Sciences
University of Sheffield

January 2010

"You cannot step into the same river twice"

Heraclitus of Ephesus (c. 535 – c. 475 BC)

ACKNOWLEDGEMENTS

First and foremost I would like to thank Lorraine Maltby and Phil Warren for all of their help and guidance over the last few years. Their knowledge and enthusiasm have made conducting the research presented in this thesis not only interesting but also immensely enjoyable.

Within the Environment Agency I would like to thank Alice Hiley, Veronique Adrianssen and in particular Bill Brierley for help with the provision of data and useful comments and discussions.

My thanks to Simone Bizzi from the Catchment Science Centre at the University of Sheffield who helped with the use of ArcGIS to derive the River Habitat Survey variables and land use variables utilised in Chapter 4.

I would like to thank Philippe Usseglio-Polatera from the University of Metz for the provision of the trait database that forms the basis for the analysis presented in Chapter 5.

Thanks to Sheffield Museum and particularly Alistair McLean who provided access to the Sheffield District Invertebrate survey data presented in Chapter 6.

I would also like to extend my thanks to my friends and colleagues within both the Department of Animal and Plant Sciences and the Catchment Science Centre for their encouragement and support.

I would like to acknowledge the support of the Environment Agency for England and Wales and the University of Sheffield who jointly funded this research. I would also like to thank David Lerner and Lorraine Maltby at the University of Sheffield for setting up the Catchment Science Centre in which I was based. Facilities were provided by the Department of Animal and Plant Science at the University of Sheffield, for which I thank Professor Malcolm Press and Professor Lorraine Maltby.

Last but by no means least I would like to thank three special people, firstly Jo Harkness for making the times when I wasn't working lots of fun, and secondly my Mum and Dad for their love and support.

SUMMARY

Understanding temporal variability in ecological communities is critical to monitoring and managing biodiversity and ecosystem function. However studies of community variability over large temporal and spatial scales are scarce, partly because suitable data are rare. The principal aim of this study was to use two previously unutilized data sets to examine long-term temporal variability of flowing water macroinvertebrate communities in order to understand better the patterns in this variation and role of environmental factors and community composition in driving it.

The first data source comprised invertebrate data from a national water quality monitoring programme. Four main findings arose from this analysis. Firstly, that temporal variability exhibits a complex pattern across multiple spatial scales. Secondly, that factors relating to climate, land use and local scale habitat stability are important determinants of variability. Thirdly, that these relationship are non-linear with threshold values above which there is a sudden change in temporal variability. Fourthly, that community composition is related to temporal variability with specific taxa contributing disproportionately to community variability due to their biological and ecological traits.

The second data source was a freshwater invertebrate survey carried out within a single catchment in 1979. Re-sampling of selected sites was conducted as part of this study providing a comparison of patterns within and between years. Whilst patterns of community structure were consistent between years, there was considerable variation in the identity of taxa and spatial relationships between communities over time. This seems likely to result from improvements in river quality over the last 30 years.

The combination of long-term and large spatial scale community data has provided unique insight into temporal variability. In the river systems of England and Wales environmental factors exert a strong influence on communities driving variability however, their influence is mediated through the taxa present within the system.

CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

1.1: Introduction

The Millennium Ecosystem Assessment (2005) states that the last few decades have been characterised by a period of disturbance of ecosystems unmatched in human history. Globally, freshwater systems are probably the most threatened ecosystem type (Boon, 2000) with disturbance arising from flow modification, water pollution, overexploitation, destruction and degradation of habitat and invasion by exotic species (Dudgeon et al., 2006). Such disturbance has led to changes in aquatic communities that include the loss of biodiversity (Pimm et al., 1995; Poff et al., 1997; Weiher and Keddy, 1999) and changes in the range/distribution of species, leading to patterns such as the loss of regional distinctiveness (Rahel, 2000; Johnson and Hering, 2009). Whilst there are arguments that such changes should be avoided purely on moral and aesthetic grounds (Gaston and Spicer, 1998) there is a growing understanding that changes in communities can have a profound influence on the function of ecosystems (Loreau, 2000; Millenium Ecosystem Assessment, 2005). Inevitably this may lead to a reduction in the provision of ecosystem services to humankind that are valued at many billions of dollars each year (Costanza et al., 1997; Wilson and Carpenter, 1999).

Central to our ability to monitor and effectively manage biodiversity and associated ecosystem function, is the ability to detect and understand change in communities through time (Odum, 1985; Bunn and Davies, 2000; Fraterrigo and Rusak, 2008). Temporal variability can result from natural (e.g. Minshall, 1988; Behmer and Hawkins, 1986; Beche and Resh, 2007) as well as anthropogenic disturbance (e.g. Snyder et al., 2003; Murphy and Davy-Bowker, 2005; Johnson and Hering, 2009). In order to set scientifically defensible management goals it is vital to resolve the contribution that these differing processes make to variation in communities (Niemi and McDonald, 2004).

Recent years have seen the increasing use of macroinvertebrate communities for the assessment of river system health (Metzeling et al., 2002). A central assumption of many of these assessment methodologies is that in the absence of anthropogenic stress, community composition is persistent through time (Richards et al., 1992; Statzner et al., 1997; Scarsbrook, 2002; Woodward et al., 2002; Milner et al., 2005). However, there is a paucity of data with which to examine this assumption at the spatial and temporal scales at which many potential drivers of temporal variability (e.g. changes in climate, changes in land use) operate (Metzeling et al., 2002; Jackson and Fureder, 2006).

The principal aim of this thesis is to examine temporal variability in lotic macroinvertebrate communities. In this introductory review chapter, I begin by outlining some definitions of measures of community temporal variability (Section 1.2) before considering how communities are assembled based on the selection of traits appropriate for the environmental conditions at a site (Section 1.3). Traits govern the ability of taxa to respond to disturbance and so are central to thinking about temporal variability. Differing types of disturbance are considered (Section 1.4) before the role of environmental predictability is described (Section 1.5). I then consider the role of abiotic (Section 1.6) and biotic (Section 1.7) processes in governing long-term temporal variability. Finally, I consider that long-term spatially extensive data are important for providing insights into temporal variability of communities (Section 1.8) before stating the research aim and objectives (Section 1.9).

1.2: Definitions: Persistence, stability and temporal variability

The terms "persistence" and "stability" describe differing elements of temporal variability, but are often used interchangeably in the literature (Connell and Sousa, 1983). Persistence is used as a term to indicate that a population or species either does not go extinct at a site, or if it does that it re-colonises within a time shorter than is required for the turnover off all individuals at the site (Connell and Sousa, 1983). In essence persistence is about the presence or absence of taxa at a site through time, however it is not concerned with the balance of the community in terms of which taxa are most dominant. The consideration of dominance arises through the definition of stability which is a measure of the community that considers presence of taxa as well as their abundance, and thus reflects changes in the balance of the community through time (Bradt et al., 1999; Magurran, 2004). Since abundances can change without species turnover, measuring stability can record change that would not be evident based on persistence alone. Persistence of taxa within communities is usually high compared to stability (Weatherley and Ormerod, 1990; Robinson et al., 2000) as it requires a major impact, which will be rare, to cause the extinction of a taxon (Russell et al., 1986; Death and Winterbourn, 1994; Beche and Resh, 2007). Changes in the abundance of taxa will be much more frequent, and could arise through processes such as the seasonal variation in numbers associated with their life cycle (Reece et al., 2001).

1.3: Community structure and The Habitat Templet

In order to consider both how and why community structure may change through time, it is useful first to consider the processes that govern the assembly of a community. Poff (1997) and Belyea and Lancaster (1999) present a description of how communities are assembled, based on the idea that taxa must pass through a series of "filters" in order to become established at a site. The central assumption of this model is that environmental conditions remove taxa from the total species pool at increasingly small scales until only those adapted to the local conditions remain, as illustrated in Figure 1.1. The mechanism through which this filtering takes place is based on the selection for traits that maximise a taxon's success within a community. As there is only a limited amount of both time and energy available to an organism throughout its life, resources must be assigned to various attributes to achieve the optimal adaptation to the environment (Korfiatis and Stamou, 1999). Therefore, over evolutionary time, taxa have made a series of evolutionary "choices" about combinations of traits that they possess, and it is through filtering of these traits that the presence of taxa in a community is determined. For example in New Zealand streams where communities are subjected to large, frequent and unpredictable disturbance events relating to periods of high flow, taxa are filtered so that they possess traits that make them less likely to be lost during disturbance (i.e. resistance traits such as streamlined/flattened body, clinger, two or more life stages outside water) and traits that allow them to utilise refuges and quickly re-colonise (i.e. resilience traits such as high adult mobility, small size, habitat generalist) (Townsend et al., 1997). The fundamental relationship between taxa and the environment that this filtering process represents forms the basis for Southwood's (1977) Habitat Templet theory.

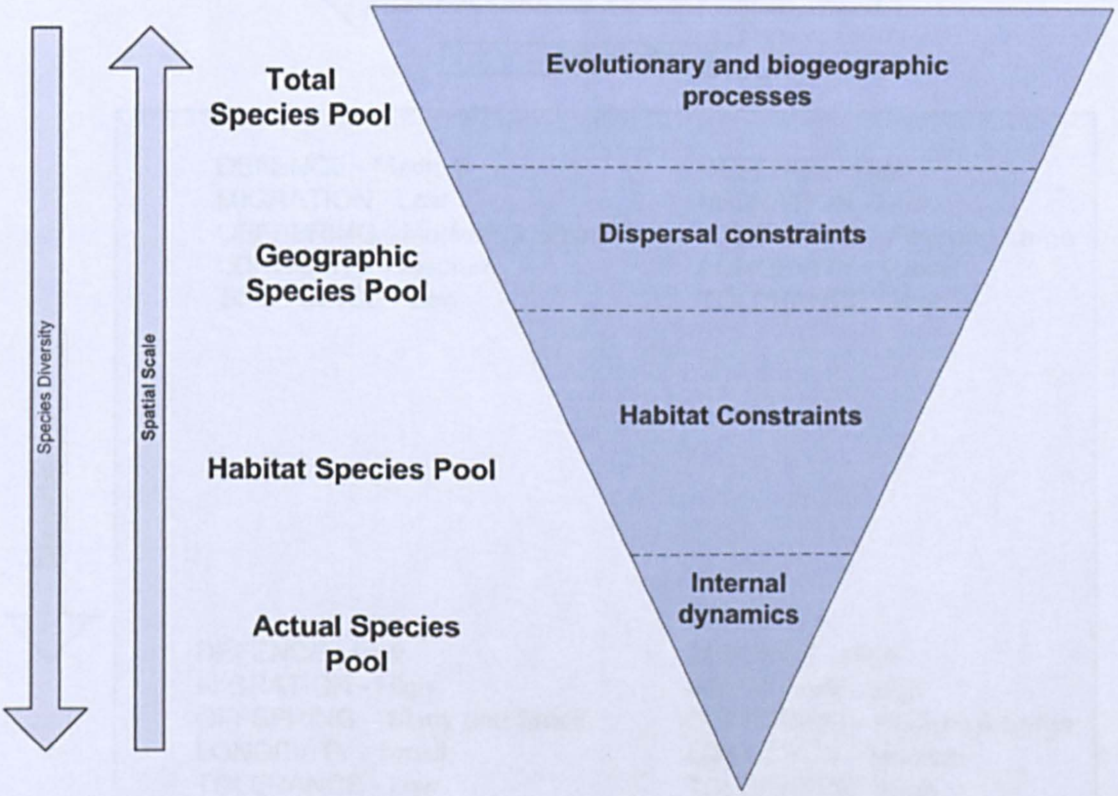


Figure 1.1. Processes defining community structure at decreasing spatial scales based on Poff (1997) and Belyea and Lancaster (1999). On the left the two arrows indicate highest species diversity at the largest spatial scale, representing the total pool of species available within a region. As spatial scale decreases this "pool" of taxa is reduced as taxa pass through a series of filters, shown on the right, that remove taxa which lack the required adaptations.

The Habitat Templet was proposed as the basis for an "ecological periodic table" (Southwood, 1977) as traits represent a fundamental description of the features of taxa. The importance of the Habitat Templet and its success for understanding the relationships between community structure and abiotic and biotic factors is illustrated by Southwood (1988) who shows that Habitat Templets make a consistent set of predictions about the traits of taxa across a broad range of groups from river invertebrates to fungi and flowering plants (Figure 1.2).

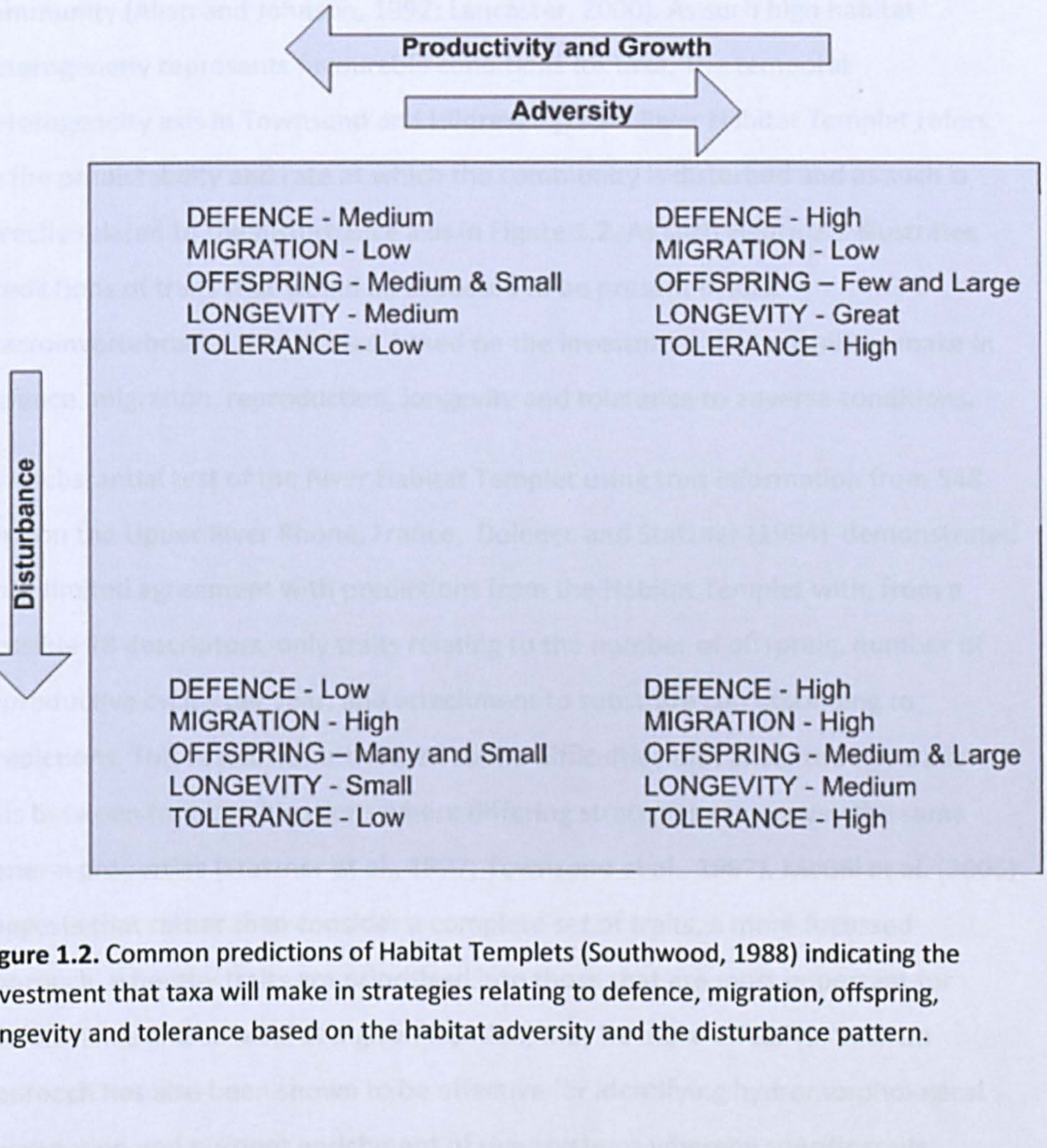


Figure 1.2. Common predictions of Habitat Templets (Southwood, 1988) indicating the investment that taxa will make in strategies relating to defence, migration, offspring, longevity and tolerance based on the habitat adversity and the disturbance pattern.

Taking Southwood's (1977) Habitat Templet, Townsend and Hildrew (1994) established a River Habitat Templet using axes defined by temporal and spatial heterogeneity, as they considered that these descriptors represent the principal forces influencing community structure in lotic systems. Townsend and Hildrew (1994) defined spatial heterogeneity as the provision of refuges in space. This is analogous to the adversity axis in Figure 1.2 that is considered to represent the unfavourableness of the habitat (Southwood, 1977; Greenslade, 1983). In lotic systems spatial heterogeneity plays a key role in reducing the influence of harsh environmental conditions by creating microhabitats, thus ensuring that the impact of adverse conditions or the influence of predators are not uniform across the community (Allan and Johnson, 1997; Lancaster, 2000). As such high habitat heterogeneity represents favourable conditions for taxa. The temporal heterogeneity axis in Townsend and Hildrew's (1994) River Habitat Templet refers to the predictability and rate at which the community is disturbed and as such is directly related to the disturbance axis in Figure 1.2. As such Figure 1.2 illustrates predictions of traits that would be expected to be present in lotic macroinvertebrate communities, based on the investment that organisms make in defence, migration, reproduction, longevity and tolerance to adverse conditions.

In a substantial test of the River Habitat Templet using trait information from 548 taxa on the Upper River Rhone, France, Doledec and Statzner (1994) demonstrated only limited agreement with predictions from the Habitat Templet with, from a possible 18 descriptors, only traits relating to the number of offspring, number of reproductive cycles per year, and attachment to substrate corresponding to predictions. This result was attributed to the difficulties that arise through trade-offs between trait combinations where differing strategies can convey the same general properties (Statzner et al., 1997; Townsend et al., 1997). McGill et al. (2006) suggests that rather than consider a complete set of traits, a more focussed approach, whereby traits are prioritised into those that are most important for performance and survival at a given location, may be more effective. Such an approach has also been shown to be effective for identifying hydromorphological degradation and nutrient enrichment of river systems whereby specific traits present within the community have changed to reflect these stressors (e.g. Bis and Usseglio-Polatera, 2001) and has been suggested as a approach that could differentiate between the impact of multiple stressors on aquatic systems (Statzner and Beche, 2010).

What emerges from this view of the relationship between organisms and their environment is that the conditions to which organisms are typically exposed defines the traits that are present within the community, and this in turn dictates the response of organisms to variability of these habitat conditions (Resh et al., 1988; Poff and Ward, 1990; Allan and Johnson, 1997). This response can take the form of resistance to disturbance whereby taxa are able to remain within the habitat when subjected to a disturbance, or resilience whereby taxa are lost from the habitat but rapidly re-colonise. As such resistance and resilience are central to the definition of persistence. For a community to be persistent following disturbance it must have taxa that are either resistant or resilient. Ecologically, resistance is achieved through the adoption of behavioural, morphological or physiological adaptations to disturbance, such as streamlining or firm attachment, whereas resilience may arise through adaptations that include a short life cycle with high reproduction and rapid dispersal potential (Gasith and Resh, 1999; Usseglio-Polatera et al., 2000). In disturbed conditions, communities typically contain a high proportion of taxa with a small body size, high adult mobility, that are habitat generalists, with flattened bodies and with two or more life stages outside water (Townsend et al., 1997; Ilg and Castella, 2002; Griswold et al., 2008). In contrast, under comparatively stable and predictable conditions, where taxa have not been forced to adapt, taxa will be expected to be longer lived and invest more energy in parental care (Statzner et al., 1997; Diaz et al., 2008).

The implication of this view of lotic communities is that the impact of a large scale disturbance (such as high rainfall events) may be markedly different between communities. Winterbourn (1997) suggested that in harsh and disturbed conditions the Habitat Templet may favour a limited taxa with strong resistance and resilience strategies leading to high persistence through time. Where communities are not subjected to such disturbance, the Habitat Templet suggests that taxa may lack traits for resistance or resilience, and so may be more susceptible to disturbance leading to higher temporal variability. Armitage (2006) demonstrated this relationship when comparing communities from a regulated and unregulated river whereby communities on the unregulated river were more robust to disturbance due to their exposure to highly variable conditions. However, this represents one of the few examples where this relationship has been examined.

1.4: Disturbance and community variability

Disturbance is a natural feature of ecosystems promoting renewal and diversity within the system (White and Jentsch, 2001). Its origins can be either endogenous or exogenous to the community and either natural or anthropogenic (Scheffer et al., 2001; Currie, 2007; Fraterrigo and Rusak, 2008). Disturbance is central to the study of temporal variability in communities as, irrespective of its origin, it represents force acting on the community. How the community is able to respond to this disturbance determines the level of temporal variability.

In a key paper, Resh et al. (1988) define disturbance as "a relatively discrete event characterised by a frequency, intensity or severity outside a predictable range, and that disrupts ecosystems, community, or population structure and changes resources or the physical environment". Whilst agreeing with this definition, Harding et al. (1998) considers that prolonged changes in the physical environment may also represent a form of disturbance to the community. The distinction here is between what Niemi et al. (1990) terms a "pulse" disturbance that represent a discrete short-term event (i.e. a flood or drought), or a "press" disturbance that represents a gradual long-term change in conditions (i.e. eutrophication or acidification). Both types of disturbance have the potential to significantly alter the community through time. Researchers may be interested in studying differing responses dependent on the form that the disturbance takes. For example, if a disturbance was to be categorised as a "pulse" through the system, then research may focus on the speed of recovery. Alternatively, with a "press" disturbance, research may focus on whether there is long-term compositional change.

Long-term processes, these "press" disturbances, can themselves be divided into two distinct categories dependent on whether they have a trajectory of change or whether they occur in a cyclical pattern. Changes in communities in response to increasing water temperature provide an example of such directional change. Daufresne et al. (2003), in a 20 year study on the Upper Rhone River, attributed changes in the fish and invertebrate community to the effect of climate change where they observed warmer water species colonising the area and the loss of cold water species. Similarly, Chessman (2009) established trends in families of macroinvertebrates in response to climate change with differing taxa exhibiting an increase, decline or no change in occurrence over a 13 year period.

Examples of long-term cyclical patterns have been noted through the examination of climatic indices such as the Southern Oscillation Index or the North Atlantic Oscillation. These indices, which may operate over decade long cycles, alternate periods of warmer, wetter climate with colder, drier (e.g. Bradley and Ormerod, 2001; Beche and Resh, 2007). Results from such studies indicate increased variability of macroinvertebrate communities arising through difference in hydrology between years over the length of the cycle (Beche, 2007).

From the perspective of temporal variability, the key part of Resh's (1988) definition of disturbance is "outside a predictable range". If disturbance to the community can be considered as representing a novel stressor, as it is outside the range of disturbance defined by the axis of the Habitat Templet, then taxa may lack mechanisms with which to respond. In the next section, I examine this further by suggesting that predictability is central to understanding the relationship between disturbance and community temporal variability.

1.5: Predictability through time

Table 1.1 summarises a selection of studies where long-term variability of macroinvertebrate communities has been examined. A consistent finding in such studies is that communities exhibit least temporal variability where conditions are constant through time (e.g. Townsend et al., 1987; Richards and Minshall, 1992; Johnson et al., 1994; Scarsbrook, 2002). Constancy ensures a match between the Habitat Templet and traits of taxa present leading to persistence of taxa within the community (Poff and Ward, 1990). Even where taxa may be periodically exposed to harsh conditions, if the timing or frequency of such events is predictable, organisms will possess strategies that allow them to respond. For example seasonal variation in habitat conditions may be quite marked in some environments however may not represent a stressor to communities as taxa will possess mechanisms, such as diapauses, that enable them to cope with such predictable variation in the environment (Reece et al., 2001; Bonada et al., 2007).

Returning to Resh et al.'s (1988) description of disturbance, the other consistent finding of long-term studies is that unusual events lead to variation within the community (Beche, 2006). Viewed within the framework of the Habitat Templet disruption of the community arises as taxa lack mechanisms through which to respond (Allan and Johnson, 1997; Bradt et al., 1999). Such unusual events may include natural disturbance a) to which the community is usually exposed, but where their magnitude crosses a certain threshold (Meffe, 1987; Beche and Resh, 2007), b) that occur unusually close together leading to a cumulative impact (Meffe, 1987; Bradt et al., 1999; Fraterrigo and Rusak, 2008), or c) where the timing is unusual (Boulton et al., 1992). Alternatively, impacts may represent novel stressors to the system, such as those associated with anthropogenic impacts, where taxa may lack mechanisms with which to respond as they have not encountered these stressors during their evolutionary history (Resh, 1988).

A complication to this argument arises through a consideration of the way that organisms perceive their environment (Malmqvist, 2002). When viewed at differing spatial and temporal scales the definition of what represents a disturbance may change. Townsend and Hildrew (1994) consider that this represents one of the most fundamental challenges to our understanding of the relationship between the environment and species traits, and by extension temporal variability. For example shortening of the time required for development to reproductive age may be critical for organisms that live in frequently disturbed habitats (Resh et al., 1988). Boulton et al. (1992) considers that in highly disturbed desert streams, due to fast life cycles, taxa may perceive their environment as being relatively stable through time as they will occupy the stable habitat between disturbance events. Similarly, taxa may select habitats that provide temporally stable conditions within a broadly unstable environment. For example, the deposition of eggs within plant stems may represent a stable microhabitat for a vulnerable portion of a taxon's life history (Richards et al., 1997; Statzner et al., 1997; Usseglio-Polatera et al., 2000).

Table 1.1. Previous long-term studies of macroinvertebrate community variability. For scale "River" refers to studies that have considered sites on the same river, "Catchment" to studies that have considered sites across a number of rivers, "Regional" to studies that have considered a number of catchments and "National" to studies that included a broad geographic spread across a whole country. Superscripts indicate analysis based on data taken from the same study area and using either the same or a portion of the same data. These are; (1) the River Rhone (France); (2) Llyn Brianne Catchment (Wales, UK); (3) Broadstone Stream (England, UK).

Author (s)	Years	Sites	Scale	Region
Meffe and Minckley (1987)	41	1	River	USA (Arizona)
Townsend et al. (1987)	8	27	Catchment	UK (England)
McElravy et al. (1989)	7	1	River	USA (California)
Richards and Minshall (1992)	7	10	Catchment	USA (California)
Johnson et al. (1994)	30	3	River	USA (Kentucky)
Bradt et al. (1999)	20	1	River	USA
Scarsbrook et al. (2000)	17	66	National	New Zealand
Voelz et al. (2000)	15	10	River	USA (Colorado)
Robinson et al. (2000)	6	6	Catchment	USA (Idaho)
Gibbins et al. (2001)	14	4	Catchment	UK (Scotland)
Metzeling et al. (2002)	20	8	Regional	Australia
Scarsbrook (2002)	9	26	Regional	New Zealand
Wagner and Schmidt (2004)	25	1	River	Germany
Aagaard et al. (2004)	15	5	River	Norway
Milner et al. (2005)	8	6	Catchment	USA (Alaska)
Brown et al. 2006	7	3	Catchment	France
Monk et al. (2006)	10	83	Regional	UK (England/Wales)
Armitage (2006)	30	5	2 Rivers	UK (England)
Beche et al. (2006)	19	2	Catchment	USA (California)
Collier (2007)	10	49	Regional	New Zealand
Burgmer et al. (2007)	15	22	17 lakes, 5 rivers	Northern Europe
Chessman (2009)	13	1818	Regional	Australia (N.S.W.)
Webb and King (2009)	11	67	Regional	Australia (Victoria)
Durance and Ormerod (2009)	50	18	Regional	UK (England)
Fruget et al. (2001)[1]	9	5	River	France
Daufresne et al. (2003)[1]	20	7	River	As above
Weatherley and Ormerod (1990)[2]	5	18	Catchment	UK (Wales)
Bradley and Ormerod (2001)[2]	14	8	Catchment	As above
Bradley and Ormerod (2002)[2]	13	11	Catchment	As above
Ormerod and Durance (2009)[2]	25	14	Catchment	As above
Speirs et al. (2000)[3]	25	1	River	UK (England)
Woodward et al. (2002)[3]	30	1	River	As above

1.6: Environmental drivers of temporal variability

Whilst constancy and predictability of environmental conditions are central to temporal variability the specific factors identified as being of central importance are often study or even site specific. Whilst this will in part reflect differences in the range of factors examined between different studies it also indicates that the importance of specific factors may change depending on regional and local context. However, a number of commonalities emerge from the studies presented in Table 1.1 which are now considered in further detail.

1.6.1: Flow regime

Flow regime is considered to be one of the most significant factors affecting river systems (Statzner and Higler, 1986; Poff, 1997; Monk et al., 2006). Periods of unusual flow may have the greatest influence on community structure through time (Statzner and Higler, 1986; Richards et al., 1997; Brown et al., 2007). For example Metzeling et al. (2002) and Beche and Resh (2007) identified periods of low flow, resulting from drought, as being particularly damaging to communities driving ecological instability between years. Similarly, flooding can have a major impact on communities through the shear stress exerted on the substrate (Clausen and Biggs, 2000; Beche and Resh, 2007) or through the removal of food (Collier, 2007) and refuges. However, establishing such links has proved a challenge due to the differing ways that the flow regime can vary and be characterised (Clausen and Biggs, 2000; Konrad et al., 2008) and a lack of data availability (e.g. Bradley and Ormerod, 2001). Clausen and Biggs (2000) reports that over 50 different variables relating to flow regime were used by 6 separate studies examining the influence of flow on macroinvertebrate communities. However, Clausen and Biggs (2000) identified that measures could be grouped into firstly, those that provide a general description of the environment (average flow magnitude and variability), and secondly, those that report the duration, volume and frequency of unusual flow events. In the context of the current study the second group is most relevant as this description corresponds to Connell and Sousa's (1983) definition of disturbance as representing an event outside of the predictable range that taxa experience. For example the frequency with which flows exceeding the three times the median discharge emerge in a number of studies as representing an important determinant of community stability (e.g. Gibbins et al., 2001).

1.6.2: Surrounding Land Use

With the profound influence that human activity has had on river systems (Usseglio-Polatera and Beisel, 2002), an emerging view in river ecology is that watershed scale variables relating to land use provide the most accurate prediction of stream communities with alteration of the broad system overwhelming local factors (Gergel et al., 2002; Urban et al., 2006). Land use, particularly urban and agricultural, affects the river system by altering the geomorphology, thermal regime, water chemistry, nutrient loadings, sediment and organic matter inputs, and flow regime (Paul and Meyer, 2001; Sponseller et al., 2001; Snyder et al., 2003; Townsend et al., 2004). The response of communities to increasing land use pressure is profound and usually results in a decrease in diversity of taxa above certain threshold values (see review in Gergel, 2002). However, the relationship between increasing land use intensity and temporal variability of communities has seldom been examined in lotic systems. In one of the only examples Collier (2007) demonstrated a complex relationship whereby above a certain level of land use stress there is a sudden sharp increase in community variability through time. In Collier's (2007) study metrics relating to Ephemeroptera, Plecoptera and Trichoptera showed a small increase in variability up to moderate levels of land use stress, then a rapidly accelerating increase in variability at highly stressed sites.

1.6.3: Substrate Stability and Habitat Heterogeneity

In a number of studies, substrate stability emerges as a key factor relating to temporal variability (Gibbins et al., 2001; Brown, 2007). Large bed sediment is needed for the provision of refuges and attachment sites (Roy et al., 2003) that will mitigate the influence of impacts, principally relating to flow regime, on communities. However, habitat degradation tends to reduce the substrate size (Heatherly et al., 2007) resulting in an decrease in the stability of the substrate for a given discharge event (Townsend et al., 1997). Communities with the least stable substrate have been demonstrated to be the most influenced by periods of high flow (Gibbins, 2001).

More broadly substrate, together with features such as vegetation, provides habitat heterogeneity, a vital component of the river that will influence community temporal variability as it ensures that disturbance is rarely uniform across the system (Brown, 2007). Heterogeneity provides a principal mechanism through which small scale factors influence the action of larger scale processes by providing refuges for taxa and a source of colonists following disturbance (Townsend et al., 1997; Lancaster, 2000; Brown, 2007; Diaz et al., 2008) and as discussed in section 1.3 is central to the structuring of communities in lotic systems. For example Brown (2003) demonstrated a positive relationship between the number of habitat patches and temporal variability in macroinvertebrate communities concluding that heterogeneity was important in providing refuges from both predation and high flows. Lancaster (2000) demonstrated a high density of taxa in refuges immediately following flooding, with recovery of populations arising through re-colonisation from these areas.

1.6.4: Interactions between factors

Whilst the influence of flow, land use and substrate are considered separately above, in real systems there will be complex interactions between these factors that may enhance or ameliorate the effect of any one factor on the system. For example changes arising through urbanisation of catchments causes both the loss of habitat heterogeneity (Brown, 2003) and an increase in the flashiness of rivers (Walsh et al., 2005) serving to increase the level of disturbance and leading to an increase in temporal variability of communities. Similarly, recent studies have demonstrated that the influence of rising annual temperatures on macroinvertebrate communities can be masked by water quality problems (e.g. Daufresne et al., 2007; Durance and Ormerod, 2007; Durance and Ormerod, 2009) or by variation in discharge (Durance and Ormerod, 2007). Identifying those variables which are likely to exert an overriding influence on communities is therefore central to increasing our understanding of temporal variability in lotic macroinvertebrate communities.

1.7: Biotic processes and Community Assembly

So far environmental factors have been considered to be the principal drivers of temporal variability. However, biotic processes (e.g. predation, competition, disease and dispersal) can also be considered to play a key role in structuring communities over time. Generally, studies indicate that biotic processes may be most important within the community where environmental conditions are most stable through time (Townsend et al., 1987; Richards and Minshall, 1992; Woodward et al., 2002).

Dispersal can be seen to play a central role in determining temporal variability in communities in two ways. Firstly, recovery of communities following disturbance relies on successful re-colonisation of taxa. This re-colonisation has been shown to occur comparatively quickly (e.g. Townsend and Hildrew, 1976; Woodward et al., 2002) meaning that when viewed over long periods of time communities may exhibit little temporal variability. Alternatively, Humphrey et al. (2000) and Beche et al. (2006) demonstrate that taxa in streams with a distinct wet and dry season rely on re-colonisation between seasons, as conditions are too unfavourable for continued persistence over time. This has the effect of leading to high levels of variation in communities between years as the identity of the taxa that re-colonise varies between years (Beche, 2006).

More broadly, work examining the role of stochastic and deterministic (Connell, 1978; Hubbell, 1997; Poff, 1997) processes in defining community composition, suggests that the importance of chance events, such as re-colonisation, may alter with the harshness of the environment (Chase, 2007; Lepori and Malmqvist, 2009). Lepori and Malmqvist (2009) demonstrated that along a gradient of disturbance defined by flow regime, at the ends of the gradient – the least and most disturbed - the community was governed by random processes, leading to sudden shifts in community structure through time. At the most stable sites Lepori and Malmqvist (2009) demonstrated that dynamics of birth, death, colonisation and extinction, particularly in relation to rare taxa, governed community composition through time. Similarly, at the most disturbed sites severe heavy flooding led to random extinction and colonisation. This contrasts with sites subjected to intermediate levels of disturbance where there was a loss of sensitive taxa and a community dominated by more robust taxa that followed a predictable gradient of recovery after disturbance (Lepori and Malmqvist, 2009). These studies (Chase, 2007; Lepori and Malmqvist, 2009) suggest that changes in the physical environment may lead to a shift in the relative importance of deterministic and stochastic processes in shaping the community, which in turn may alter the temporal variability of the community.

Whilst biotic factors such as predation, competition, disease and dispersal can create strong localised patterns within streams (Poff et al., 1997; Belyea and Lancaster, 1999), as demonstrated by Chase (2007) and Lepori and Malmqvist (2009) their importance for determining temporal variability is controlled by the influence of the physicochemical environment (Death and Winterbourn, 1995). For example, Beche and Resh (2007) demonstrated that an increase in the importance of predation and competition within a stream community was driven by the increasing density of taxa in pools following a drought. Similarly, Woodward et al. (2002) showed that acidification of a stream led to the increase in abundance of a dragonfly predator, significantly affecting the structure of the community through time. In both instances, changes in the community through time were driven by biotic processes that resulted from a change in the physicochemical environment. This justifies the focus on environment conditions taken by many of the studies in Table 1.1.

1.8: Biomonitoring in the "Invisible Present"

In Section 1.1 of this chapter I considered that the importance of detecting and understanding temporal variability in communities is due to the implications it has for biodiversity and ecosystem function. One of the most widely used techniques for detecting change in lotic communities is routine monitoring of macroinvertebrates, due to their ubiquity and ability to indicate water quality over extended periods of time (Metcalfe, 1989; Clarke et al., 1996; Bonada et al., 2006). A typical approach for such assessment, and one that is employed in the UK, is to use a predictive model to compare the observed community to that which would be expected to be present in the absence of anthropogenic stress (i.e. RIVPACS:

Wright, 2000; AUSRIVS: Davies, 2000; BEAST: Reynoldson et al., 2000). Underpinning such systems is the assumption that communities that are not subjected to anthropogenic stress will be highly persistent (Bunn and Davies, 2000; Humphrey et al., 2000). However, this assumption has rarely been tested. Irvine (2004) argues if we are unable to quantify the levels of change that arise due to natural processes, we may draw incorrect conclusions about the status of river systems. This may arise through two processes. Firstly, where a "reference" condition is used persistence of the reference community must be assumed otherwise the comparison of impacted communities is being made against a moving baseline. Secondly, for the implementation of legislation such as the Water Framework Directive (European Community, 2000), which classifies status based on a number of categories, failure to understand the level of temporal variability may lead to incorrect classification of the ecological status of sites as the community present may have moved between categories due to temporal variation caused by factors other than anthropogenic stress. Difficulties caused by such variability will be most significant close to class boundaries where smaller variation has an increased probability of leading to misclassification, and most critical at the moderate to good boundary where misclassification may lead to inappropriate management practices (Irvine, 2004).

The difficulty in investing resources over extended time periods means that long-term datasets with which to examine temporal variability of macroinvertebrate communities are comparatively rare (Elliott, 1990; Reid and Ogden, 2005; Holmes, 2006; Fisher et al., 2010). In reviews of the freshwater literature both McElravy et al. (1989) and Jackson and Fureder (2006) demonstrated that few studies are of over five years in length. Without a suitable long-term perspective researchers have been shown to over-emphasise the importance of rare events in shaping communities (Weatherhead, 1986; Boulton et al., 1992). Although events such as droughts (e.g. Beche, 2007; Griswold, 2008) can appear to be catastrophic for communities in the short term, recovery of systems may occur relatively quickly (e.g. Townsend, 1976). Scarsbrook (2002) reports that two studies using the same sites reached markedly differing conclusion about temporal variability of macroinvertebrate communities in New Zealand streams due to difference in the period over which assessment was made. Using data from 1990 – 1996 Scarsbrook et al. (2000) detected directional change for a number of communities however, a subsequent analysis of communities from 1990 – 1998 suggested no trend.

In order to understand the temporal characteristics of a system, it is essential that the length of time that monitoring is undertaken for should reflect dynamics within the system (Strayer et al., 1986; Wiley M.J et al., 1997, Griswold et al., 2008). Many species have particular points in their life history where they are most sensitive to disturbance, for example emergence or oviposition (Townsend et al., 1987; Scarsbrook, 2002). Ensuring that at least one generation has occurred could be considered as the minimum time over which to consider a population in order to assess temporal variability (Connell and Sousa, 1983).

Similarly, many processes (e.g. succession, acid deposition, climate change) are slow, rare, subtle or complex (Jackson and Fureder, 2006) meaning our ability to both detect and understand temporal variability is limited by what Magnuson (1990) refers to as the "invisible present". Historic data are essential to provide us with a context against which we can judge changes in communities through time (Swanson and Sparks, 1990; Janzen, 2009). Such long-term studies have provided us with an understanding of issues of current societal concern such as acidification (e.g. Weatherley and Ormerod, 1987; Woodward et al., 2002), climate change (Burgmer et al., 2007; Chessman, 2009) and biodiversity loss (Rahel, 2000; Fisher et al., 2010) as well a wide range of other biological and ecological processes (see examples in McElravy et al., 1989; Elliott, 1990; Jackson and Fureder, 2006; Reid and Ogden, 2006; Fisher et al., 2010).

Equally important is the availability of spatially explicit data (Jackson and Fureder, 2006), since the examination of spatial patterns can provide important information about those factors that are most significant in determining temporal variability (Li et al., 2000). However, most long-term studies are limited spatially, most typically considering a few sites (e.g. Meffe and Minckley, 1987; Boulton et al., 1992; Woodward et al., 2002; Beche et al., 2006; Daufresne et al., 2003) or more rarely patterns at the catchment (e.g. Brewin et al., 2000; Bradley and Ormerod, 2001) or regional and national scale (e.g. Scarsbrook et al., 1999; Scarsbrook, 2002; Chessman, 2009). As can be seen in Table 1.1 with only 1 exception (Chessman, 2009) all studies are based on less than a hundred sites, predominantly distributed across a single river or catchment. Substituting space for time as a method of understanding ecological processes is problematic due to spatial structuring whereby the assumption that all sites have the same history and environmental characteristics is seldom met (Strayer et al., 1986). However, less attention is paid to the problems of considering temporal variability at small scales. If only a single site is considered then temporal variability of the community may be incorrectly assigned to factors that vary in time whereas the spatial context may be equally important (Collier, 2007). As communities are defined by the interaction of factors acting across multiple spatial scales (Poff, 1997; Belyea and Lancaster, 1999) the ability to understand what drives temporal variability may be compromised by a lack of spatial data.

1.9: Conclusion

In this chapter, it has been argued that understanding the role of natural and anthropogenic factors in determining temporal variability of communities is essential if we are to make informed management decisions to preserve biodiversity and ecosystem function (Section 1.1). Community structure was described as arising within a framework where the typical conditions dictate the traits of taxa present (Section 1.3), and this in turn determines how taxa can respond to changes in the environment (Section 1.4). Predictability of conditions was considered as central to persistence of taxa (Section 1.5) and it was demonstrated that anthropogenic factors often result in an increase in the number and severity of unpredictable events (Section 1.6). Finally, I considered that long-term spatially extensive data are needed to examine temporal variability but due to the difficult in investing time and resources few such data sources exist (Section 1.8).

1.10: Thesis Aim and Objectives

The principal aim of this thesis is to examine long-term temporal variability in lotic macroinvertebrate communities. This aim is addressed through the use of two primary data sources. Firstly, data collected by the Environment Agency (a statutory monitoring body in England and Wales) as part of their environmental assessment programme. These data will be used to address objectives 1-3 below and are described in further detail in Chapter 2. Secondly, historic data deriving from an extensive survey of the river system of Sheffield Metropolitan District conducted in 1979/80. These data will be used to address objective 4 below.

Objective 1: Spatial patterns in temporal variability

Few studies have examined temporal variability at scales that capture large environmental gradients (Section 1.5). Such an examination would provide an important perspective on the scale at which processes that determine temporal variability are operating (Section 1.8) which in turn would provide an indication about the likely identity of these processes (Section 1.6). Therefore, the first objective of this study is to examine spatial patterns in temporal variability of lotic macroinvertebrates across England and Wales.

Objective 2: Environmental drivers of temporal variability

Based on results from objective 1, the second objective of this study is to examine which environmental variables are key determinants of temporal variability in lotic macroinvertebrate communities. The spatial extent of the current dataset provides a gradient of environmental conditions (Section 1.6) with which it is possible to examine the relative contribution of natural and anthropogenic factors to temporal variability.

Objective 3: Community structure and temporal variability

The fundamental relationship between temporal variability and traits of taxa (Section 1.3 and 1.4) is then examined. The objective of this chapter is to examine the relationship between traits of taxa, community composition and temporal variability to determine whether some communities, owing to their constituent taxa and the traits they possess, exhibit more temporal variability than others.

Objective 4: Long-term changes in macroinvertebrate occurrence and distribution in Sheffield Metropolitan District.

Utilising a dataset collected in 1979/80 the changes in lotic macroinvertebrate community structure over a 28 year period across a river system in south Yorkshire are examined. These data provide an unusual combination of time span, and high spatial density of sampling sites, providing an opportunity to examine temporal changes in distribution patterns within one part of a drainage basin. The objective of this section is to describe temporal variability in the system and consider shifts in both taxon occurrence and community structure at these scales.

CHAPTER 2: METHODS

2.1: Introduction

The work in this thesis is based on two sources of freshwater invertebrate community data not directly collected in this study. The first comprises data collected as part of a nationwide water quality monitoring programme conducted by the Environment Agency (Chapters 3-5). The second consists of data on the freshwater invertebrate fauna of the Sheffield area collected during 1979 and 1980. This chapter provides a general description of these data and highlights some of the steps taken in preparing the data for analysis.

Analyses of the first data set (Chapters 3 – 5) use the same measure of community temporal variability. This chapter provides a general introduction to the choice of measure and its particular characteristics and allows extra scope to illustrate the performance of the measure.

2.2: Data Sources

2.2.1: Environment Agency Monitoring Data – the BIOSYS database

The Environment Agency conducts an extensive programme of water quality monitoring every year involving the collection of thousands of samples from running and still waters across England and Wales. Data relating primarily to macroinvertebrate communities are collated and entered into a relational database called BIOSYS.

The BIOSYS database contains information collected over a time period of nearly 40 years. The earliest record within BIOSYS was collected in 1964, which predates the formation of the Environment Agency by 32 years. This reflects the use of the database as a general repository for historic as well as contemporary records from a range of sources (i.e. water companies, private individuals, consultants, angling clubs etc.). The data used for the present study run up to the end of 2005, the date the work commenced. In total there are 297,969 samples from 62,765 sites held within BIOSYS for this period. Figure 2.1 illustrates the number of sites with sampling information available during each year over this period.

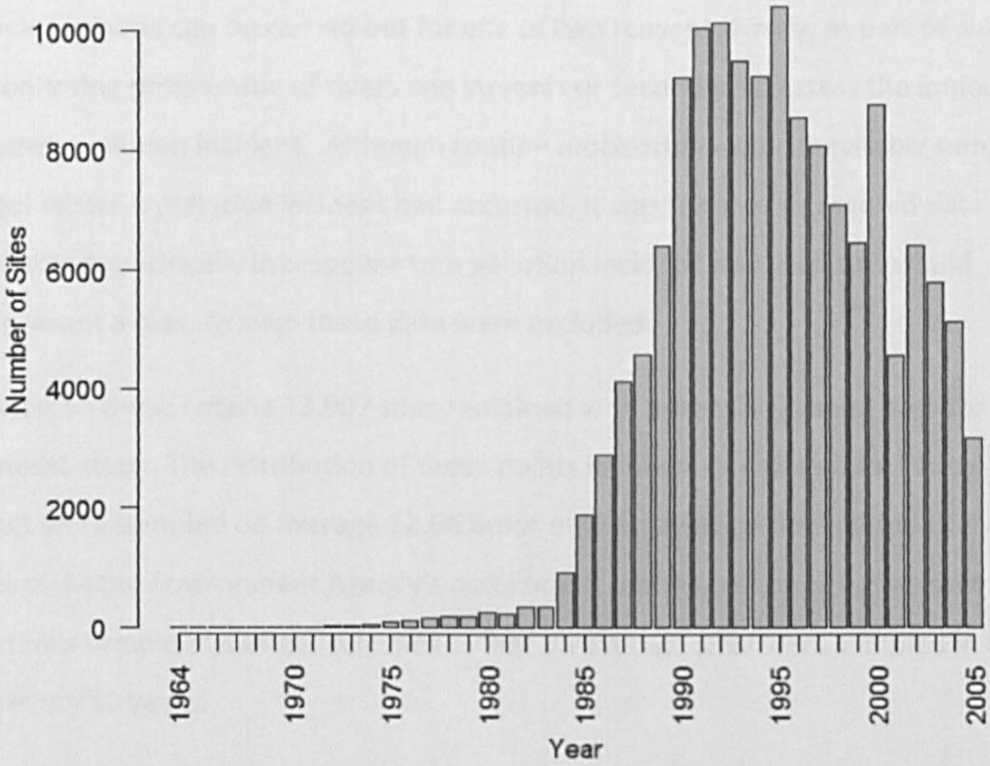


Figure 2.1: Number of Environment Agency sampling sites with macroinvertebrate data by year.

As the principal aim of the current study is to analyse temporal variability, one key requirement is that the data be collected in a consistent manner. In this way variability in community structure through time is driven by either actual change in the community, or by error in the execution of sampling or recording (Clarke, 2000; Dines and Murray-Bligh, 2000), but not by systematic changes in the methods.

Errors cannot be removed, but use of a consistent method means that any methodological errors should be consistent across all samples. Therefore, the first step in selecting data for analysis was to filter the data to select only those samples with a consistent method.

The standardised method used for the collection of macroinvertebrate data is based on a three minute kick sampled and one minute manual search and is described fully in Murray-Bligh et al. (1997) and in further detail in subsequent chapters. The data set was therefore filtered to include only samples collected using this method. Such sampling can be carried out for one of two reasons. Firstly, as part of a routine monitoring programme of rivers and streams or secondly, to assess the impact of a known pollution incident. Although routine monitoring would inevitably sample sites where a pollution incident had occurred, it was felt that if targeted data collected specifically in response to a pollution incident was used this would represent a bias. As such these data were excluded.

Based on these criteria 13,907 sites remained with potentially useful data for the current study. The distribution of these points is shown in Figure 2.2a. These 13,907 sites were sampled on average 12.66 times over an average time period of 9.48 years. As the Environment Agency's assessment method requires both a spring and autumn sample this effectively means that on average sites were sampled in 6 out of every 10 years.

Having selected a dataset collected in a consistent manner the next step was to select a temporal window over which the assessment of community temporal variability was to be made. In selecting sites the aim was to strike a compromise between spatial coverage and temporal length. Differing combinations of time periods were used and it was found that 1990 to 2005 gave the best combination of period length and the spatial coverage. This resulted in the selection of 1574 sites where sampling was conducted over at least 6 years. Of note is the absence of sampling data from the north west of England. The operations team in the north west of England employed a slight deviation from the method used by the rest of the country, by conducting bank side sorting of samples. In all other regions the entirety of the sample was preserved and taken back to the laboratory for sorting. Preliminary analysis of scales of temporal variation (using the approach in Chapter 3), including data from the North West region, indicated a strong bias towards highest levels of temporal variability in this region. To avoid possible confounding of this regional variation of method with other sources of temporal variation, samples from areas known to use bank side rather than laboratory sorting were excluded. The distribution of the remaining points is shown in Figure 2.2b.

Having identified sites the final task required was to standardise the recording of taxonomic information. Two issues arise here. First, examination of the sample data indicated that a mixture of both species and family level data were used, and that this was not consistent through time at any one site. The decision was made to standardise the taxonomic resolution to that of family level to avoid the problem that otherwise variation could arise through the use of differing taxonomic resolution and not through actual changes in the community (Bradley and Ormerod, 2001; Metzeling et al., 2002).

A second potential source of error arose because, during the length of time represented in the dataset, a number of taxonomic revisions had occurred meaning the same species could be recorded under one or more different names over time. The Coded Checklist of Animals Occurring in Fresh Water in the British Isles (Furze, 2007) was used as the standard list of taxa present within England and Wales. All taxa were assigned current names based on this list ensuring consistency across years. Taxa with incorrect spelling were identified by comparing a complete list of all taxa present within the BIOSYS dataset, to the master list of Furze (2007). Where a name in BIOSYS did not correspond to the Furze list in most cases there was an obvious error (i.e. Assellidae [sic] and Asellidae) that could be corrected. In the few instances where it was unclear these records were omitted.

This process resulted in a master list of 160 taxa across all sites with average taxon richness per sample of 21.12, and a total of 1574 sites (Figure 2.2b). It is this data set that forms the basis for the analyses in Chapters 3-5.

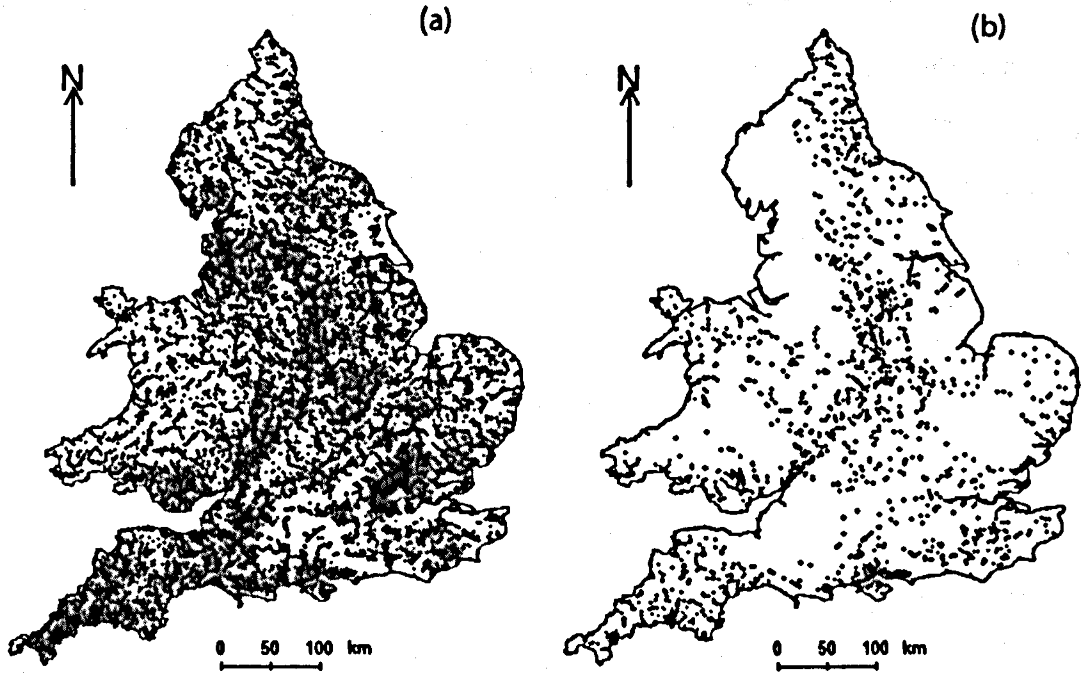


Figure 2.2: Illustrating the distribution of; (a) all 13,907 sample points sampled with a consistent method involving a 3-min kick sample and 1 minute manual search and; (b) those sites with at least years sampling years between 1990 and 2005.

2.2.2: Sheffield Invertebrate Survey

Chapter 6 is based on historic data collected in the Sheffield Metropolitan District over 1979 and 1980 (Zasada and Smith, 1981). The analysis performed in Chapter 6 is based on a subset of the historic data from 25 sites that were re-sampled in 2007. Here a brief description of the full historic dataset is provided.

The entire survey comprised freshwater invertebrate samples taken from 423 sites of which 299 were in streams or rivers. Data detailing the time and location of sampling and the names and abundances of taxa recorded were on record cards archived at the Sheffield City Museum along with preserved samples from the original survey. These data cards were digitised and these digital copies used to extract information and input it into a Microsoft Access database.

The location of each point was plotted in ArcGIS using an 8-figure grid reference. Figure 2.3 shows all 299 stream or river sites sampled within the Sheffield Metropolitan District. As described above for the BIOSYS dataset, taxa names were standardised to family level and consistency in spelling checked using the Furze (2007) list. This resulted in a master list of 67 taxa with average of 7.57 taxa per site. Of these 299 sites 25 were re-sampled in 2007 and form the basis for the historical comparison. These points are indicated in Figure 2.3 by green points. Further details of this dataset are provided in Chapter 6.

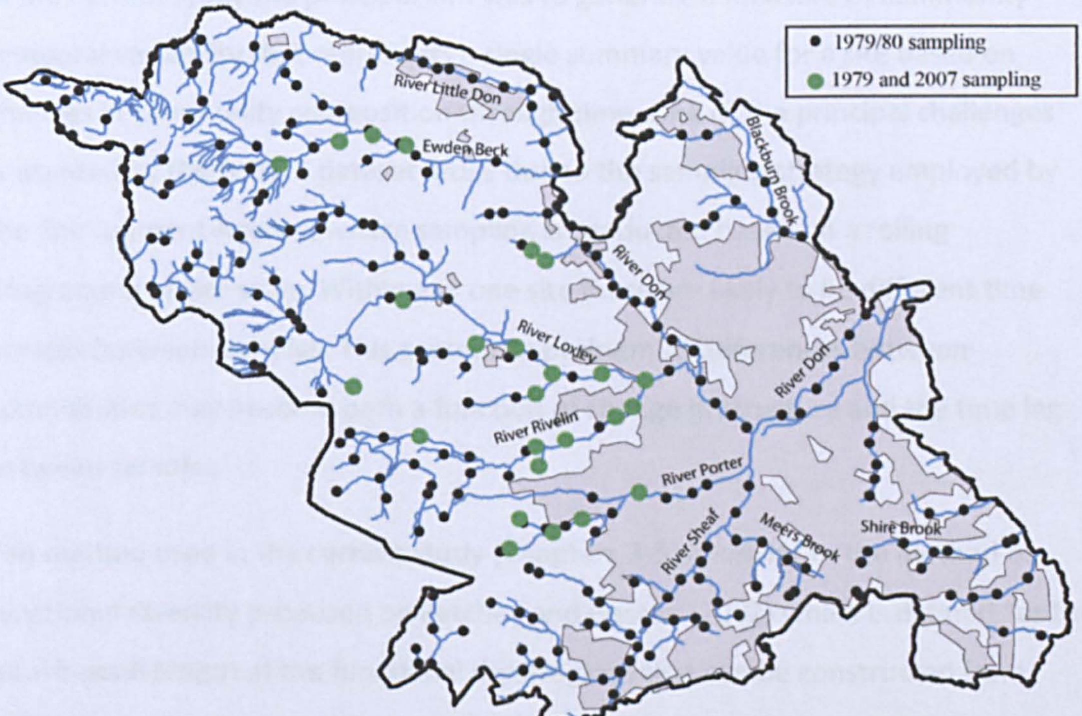


Figure 2.3: All lotic freshwater sites within the Sheffield Metropolitan District sampled as part of the 1979/80 Sheffield Invertebrate Survey. Green points indicate sites sampled in both 1979 and 2007.

2.3: Temporal variability of macroinvertebrate communities

2.3.1: The measure of community change

In the literature temporal variability of macroinvertebrate communities has been assessed using a variety of differing methods depending on the question being examined. However, a common feature of many studies is the availability of inter-annual data. As such the assessment of temporal variability is often based on the calculation of dissimilarity between pairs of years (e.g. McElravy et al., 1989; Bradley and Ormerod, 2001; Milner et al., 2005). Longer term variability can then be assessed by examining features such as the Coefficient of Variation of these dissimilarities (e.g. Beche et al., 2006) or through the use of ordination or regression techniques to examine trends through time (e.g. Gibbins et al., 2001).

In the current study the principal aim was to generate a measure of community temporal variability that represents a single summary value for a site based on changes in community composition through time. One of the principal challenges presented by the BIOSYS dataset arose due to the sampling strategy employed by the Environment Agency, where sampling is conducted based on a rolling programme of site visits. Within any one site there are likely to be different time periods between samples. This presents a problem as differences between communities may become both a function of change in structure and the time lag between samples.

The method used in the current study (Chapters 3-5) is similar to the method of functional diversity proposed by Petchey and Gaston (2002) which is defined "as the total branch length of the functional dendrogram that can be constructed from information about species' functional traits" (Petchey and Gaston, 2007). In the current study Petchey and Gaston's (2002) "functional diversity" is replaced by community composition described by the taxa present and their relative abundance described on a logarithmic scale from 1 to 5 (where Category 1= 1-9; Category 2= 10 – 99; Category 3 = 100 – 999; Category 4 = 1000 – 9999; Category 5 = 10000 - 100000 individuals).

To calculate temporal variability, a time by taxon matrix is constructed for each site and a dissimilarity measure (Kulczynski distance; Section 2.3.2 below) is used to construct a dissimilarity matrix representing all pairwise comparisons between years. This dissimilarity matrix is then used to calculate a dendrogram, where each branch represents the community on one sampling occasion. The measure of temporal variability is then calculated by summing the total branch length of this dendrogram. As total branch length is sensitive to the number of samples used, this was standardized to six samples (i.e. sampling occasions) per site, by random selection from samples taken between 1990 and 2005 whilst maintaining both the first sample (collected in either 1990 or 1991) and the last sample (collected in either 2004 or 2005).

Figure 2.4a illustrates the dendrogram for a hypothetical community that has exhibited little community variability through time. It was composed of 30 taxa in year 1 and lost a single taxon in each subsequent year. As communities share many common taxa between years, the total branch length of the dendrogram is low resulting in a value for community change of 0.14. This contrast with Figure 2.4b, where again a hypothetical community composed of 30 taxa in year 1 is used. In this instance there is a 10 percent change in taxa in each subsequent year resulting in a final community that shares only 50 percent of the original taxa. This results in more dissimilarity between communities through time, a higher total branch length of the dendrogram (note differing scales) and consequently a value of 0.8 for temporal variability.

By examining temporal variability in this way the influence of uneven sampling is reduced. The measure, by grouping communities with the most similar composition closest together, assesses the minimum "spread" of the community through time. Effectively the measure describes the two communities that have the maximum dissimilarity over time, and captures the amount of variation in the remaining communities.

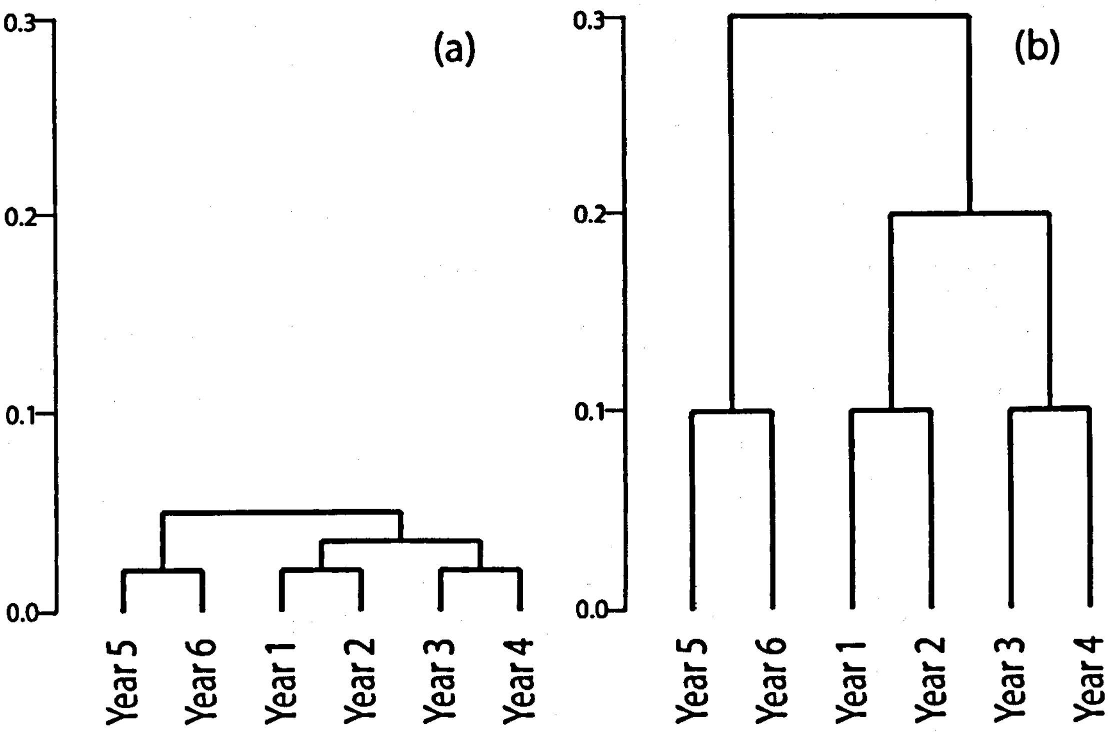


Figure 2.4. Dendrograms representing community similarity for; (a) community composed of 30 taxa losing 1 taxon per year and (b) community composed of 30 taxa losing 10% each year. Note differing scales of y-axis.

2.3.2: Dissimilarity measures

The most fundamental step in comparing communities both in space and time is the calculation of a measure of their dissimilarity. Many multivariate methods, commonly applied to the analysis of ecological communities (e.g. Principal Components Analysis, Correspondence Analysis), are based on the calculation of a matrix of such compositional dissimilarities between samples (Faith et al., 1987). However, there are a wide variety of such measures available all with differing properties (see Legendre and Legendre (1998) for review).

In the present study the requirement was to characterise differences between ecological communities over time. In this instance a key property of the dissimilarity measure is that it will reach its maximum value where two communities have no taxa in common (Beals, 1984; Quinn and Keough, 2002). Some dissimilarity measures (e.g. Euclidian) do not possess this property as they consider that double negatives, where a taxon is absent in both samples, actually constitutes a level of similarity as the communities have a shared lack of taxa.

Measures including the Bray-Curtis, Kulczynski and Canberra distance conform to this requirement. Of these measures Faith et al. (1987), using a series of model communities, demonstrate that that Kulczynski distance performed best based on its ability to describe dissimilarities between a range of simulated communities. In the current study the Kulczynski distance was therefore used to calculate dissimilarities between communities. The measure is defined as;

$$\beta kul = 1 - \left(\frac{1}{2}\right)\left\{\frac{\sum_i min\ (X_{ij}, X_{ik})}{\Sigma(X_{ij})} + \frac{\sum_i min\ (X_{ij}, X_{ik})}{\Sigma(X_{ik})}\right\}$$

Where the formula calculates the dissimilarity between two objects j and k, based on attributes $i = 1$ to N (Faith et al., 1987). Within the literature the Bray-Curtis distance is perhaps the most commonly employed dissimilarity measure, and it shares many of the properties of the Kulczynski distance (Quinn and Keough, 2002). Whilst Faith et al. (1987) considered the Bray-Curtis measure to be robust, the author demonstrated that in some instances (where beta diversity was low) the Kulczynski distance performed marginally better. However, to demonstrate the comparability of the two dissimilarity measures, the measure of temporal variability used in the current study was calculated based on dissimilarities between communities using both the Kulczynski and Bray-Curtis measures. As illustrated in Figure 2.5 there is a strong positive correlation between temporal variability calculated using the two measures (Pearson's correlation, r = 0.98, d.f. = 1572, p <0.001). Although the distribution of the points suggests that the use of the Bray-

Curtis would lead to higher estimates of temporal variability (as can be seen by the movement of the relationship above the 1:1 line) results using either measure would be broadly comparable.

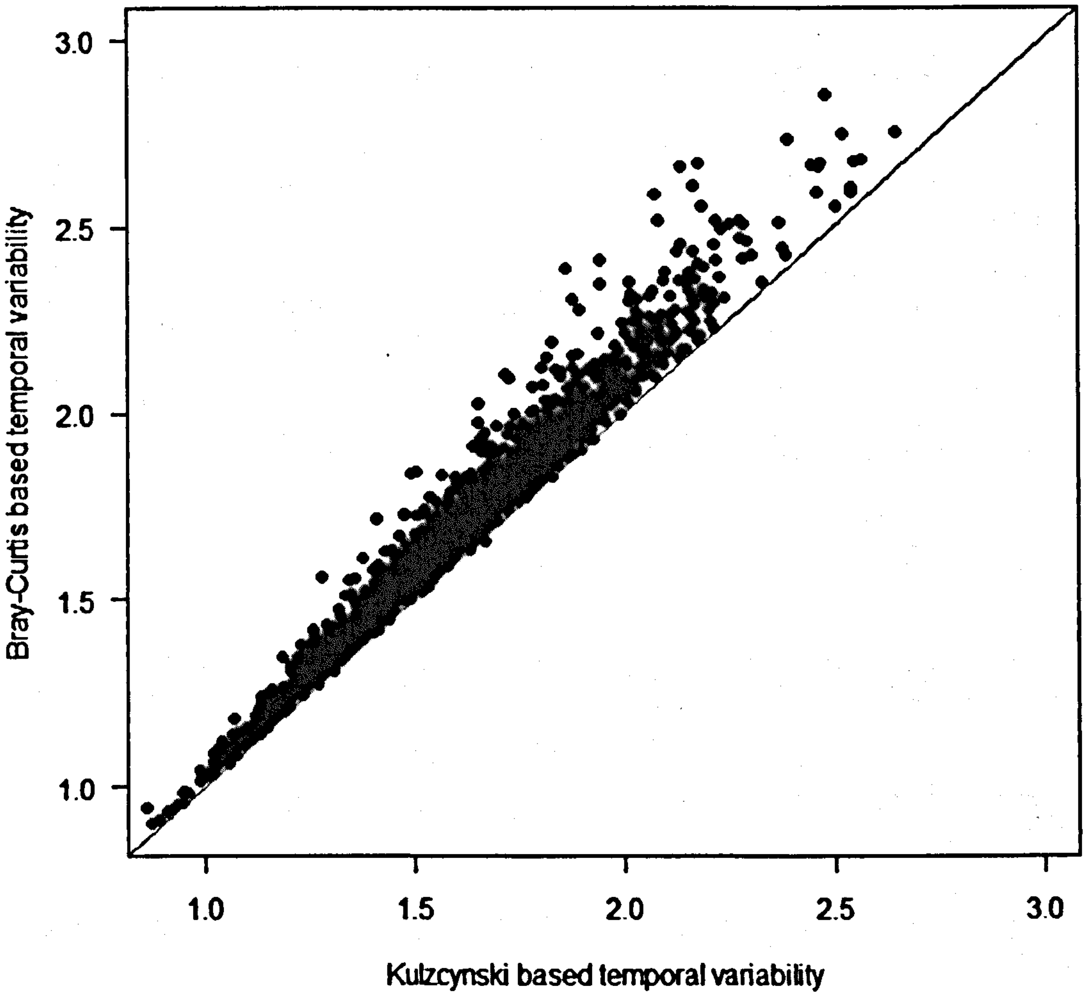


Figure 2.5: Relationship between temporal variability calculated using the Kulzcynski and Bray-Curtis measures of community dissimilarity.

2.3.3: Performance of the measure of community change

In order to interpret results from temporal variability it is important to understand how factors such as community size influence comparison between communities.

As such I present a number of examples of how the metric performs. In all cases model communities were constructed with a taxon richness of between 6 and 40, as this represent the typical size of real communities that will subsequently be considered in Chapters 3 to 5. For each community, differences are based on 6 descriptions of the communities analogous to the 6 years that will subsequently be used in Chapters 3-5. Although the dissimilarity measures is sensitive to the abundance of taxa, giving more weight to the loss or gain of the most dominant taxa, for clarity examples below use a constant abundance across all taxa.

In the first example hypothetical communities ranging in size from 6 to 40 taxa were considered to lose 1 taxon each year. This is illustrated in Table 2.1 for a community that in year 1 contains 10 taxa. Figure 2.6 illustrates the performance of the metric. As the size of the community increases from 6 to 40 there is a negative relationship between number of taxa and community change. This illustrates that as the percentage of the community that changes decreases so too does community change.

Table 2.1: Example of community used to illustrate properties of the community change measure. In this example 1 taxon was lost each year from an initial taxon richness of 10 resulting in a halving the size of the community through time.

	Taxon 1	Taxon 2	Taxon 3	Taxon 4	Taxon 5	Taxon 6	Taxon 7	Taxon 8	Taxon 9	Taxon 10
Year 1	3	3	3	3	3	3	3	3	3	3
Year 2	3	3	3	3	0	3	3	3	3	3
Year 3	3	3	3	0	0	3	3	3	3	3
Year 4	3	3	0	0	0	3	3	3	3	3
Year 5	3	0	0	0	0	3	3	3	3	3
Year 6	0	0	0	0	0	3	3	3	3	3

In the second example the community was considered to have two possible states and to switch between these states over time. This is illustrated in Table 2.2 for a community where a total of ten taxa are present through time. In odd numbered years taxa 1 - 5 were present, and in even number years taxa 6 – 10. For each of the hypothetical communities used in the analysis there was complete dissimilarity between the two states irrespective of total taxon richness. As can be seen in Figure 2.7 this results in an identical level for the value of community change. Also note that comparing the y-axis between Figure 2.6 and 2.7 that community change is higher in these latter communities. This is driven by the complete dissimilarity between communities in alternate years.

Table 2.2: Example of community used to illustrate properties of the community change measure. In this example communities fluctuate between two stable states through time.

	Taxon 1	Taxon 2	Taxon 3	Taxon 4	Taxon 5	Taxon 6	Taxon 7	Taxon 8	Taxon 9	Taxon 10
Year 1	3	3	3	3	3	0	0	0	0	0
Year 2	0	0	0	0	0	3	3	3	3	3
Year 3	3	3	3	3	3	0	0	0	0	0
Year 4	0	0	0	0	0	3	3	3	3	3
Year 5	3	3	3	3	3	0	0	0	0	0
Year 6	0	0	0	0	0	3	3	3	3	3

In the third example, four communities of 10, 20, 30 and 40 taxa are used. Each community was changed by 10 percent of the total taxa in year 1 with each subsequent year (i.e. the community with 10 taxa lost 1 taxa each year, the community with 20 taxa lost 2 taxa each year, etc), in a step wise fashion, so that after six years the first and last community shared 50% of the same taxa. As can be seen in Figure 2.8 communities with the smallest number of taxa are recorded as exhibiting the least temporal variability. Over the 6 years the community that started with 10 taxa had lost 5 taxa, whereas the community that started with 40 taxa had lost 20. This indicates that the measure is sensitive to changes in absolute numbers of taxa rather than relative numbers.

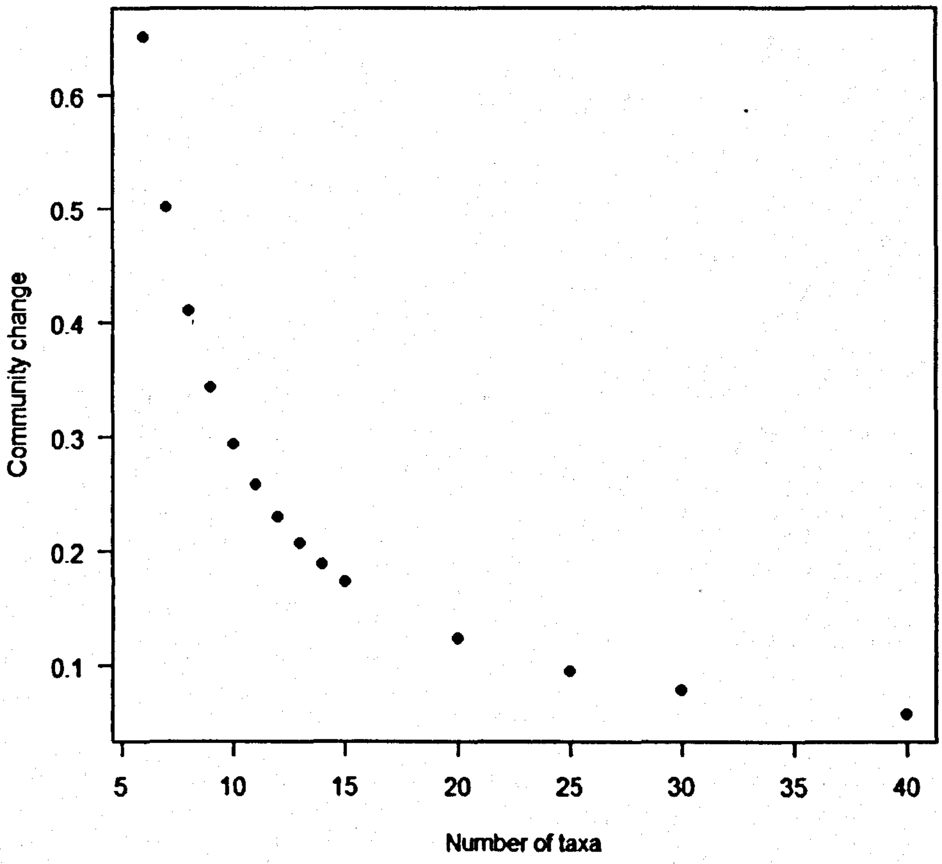


Figure 2.6: The relationship between number of taxa and community change based on communities where all taxa have the same abundance. Differences in communities arise through the loss of 1 taxon in each year of sampling.

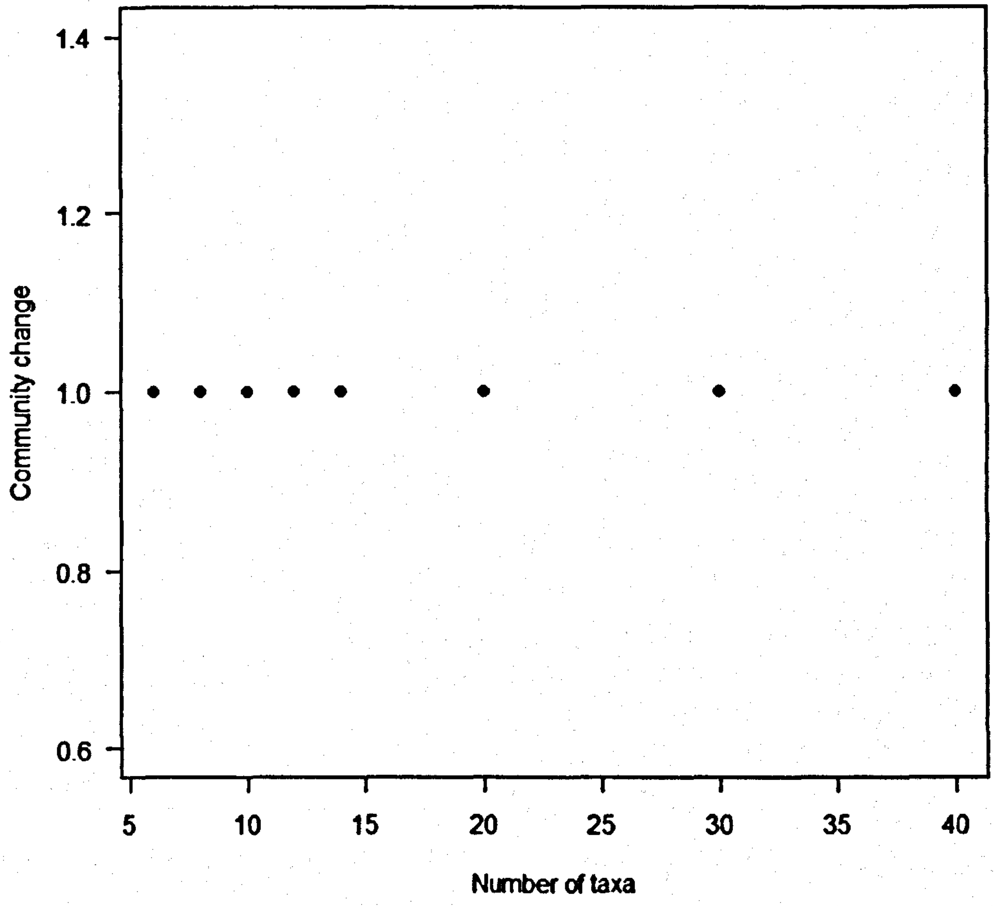


Figure 2.7: Relationship between number of taxa and community change based on communities. Differences in communities arise through the switching of taxa between two states.

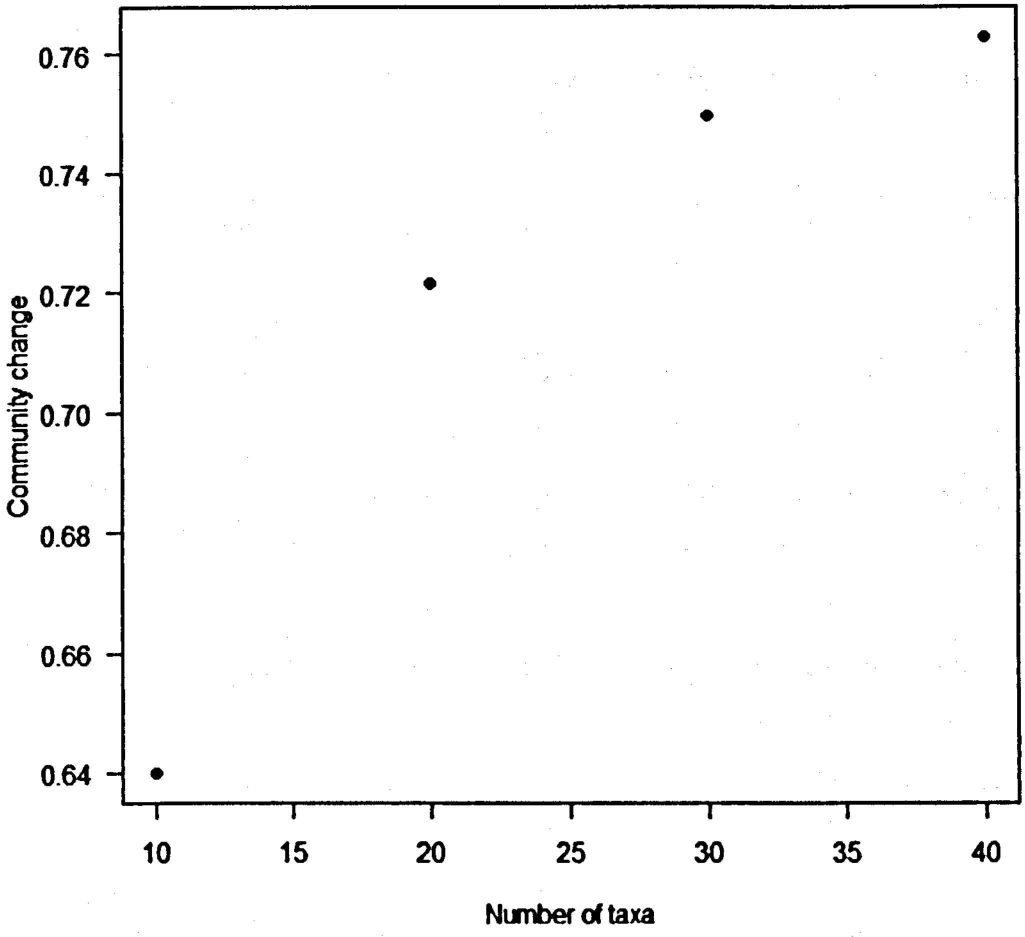


Figure 2.8: The relationship between number of taxa and community change based on the loss of 10% of the original number of taxa in the community with each subsequent year.

2.3.4: Characterising community change

A number of differing patterns of temporal variability are commonly reported in the literature. These include where communities exhibit stochastic variation around a mean composition (e.g. Voelz et al., 2000; Gibbins et al., 2001), where communities are impacted by a sudden event and subsequently recover (e.g. Meffe and Minckley, 1987; Bradt et al., 1999) or where communities exhibit directional change over time with wide difference in the community composition between the first and last sample (e.g. Daufresne et al., 2004; Webb and King, 2009).

Distinguishing these differing types of change may be informative when considering the reasons for temporal variability within a community. Collins et al. (2000) proposed a method to separate these forms based on regression analysis of temporal variability over increasing time lags between communities. If communities fluctuate around a mean composition over long periods of time there will be zero slope. This contrasts with communities that are impacted and then recover in subsequent years where there will be a negative slope. This is illustrated in Figure 2.9a based on a hypothetical community where composition in year 3 was markedly different from all other communities, simulating the impact of a pulse disturbance. Finally, if communities undergo directional change there will be an increase in the amount of dissimilarity between communities with increasing time and a positive slope. This is illustrated in Figure 2.9b where the community is undergoing directional change away from one composition towards another. In this hypothetical community there is the loss of 10% of taxa from the previous year and the addition of 10% of previously unrecorded taxa. As the time lag between samples increases so does the dissimilarity between the communities. This method is discussed further in Chapter 3 section 3.3.

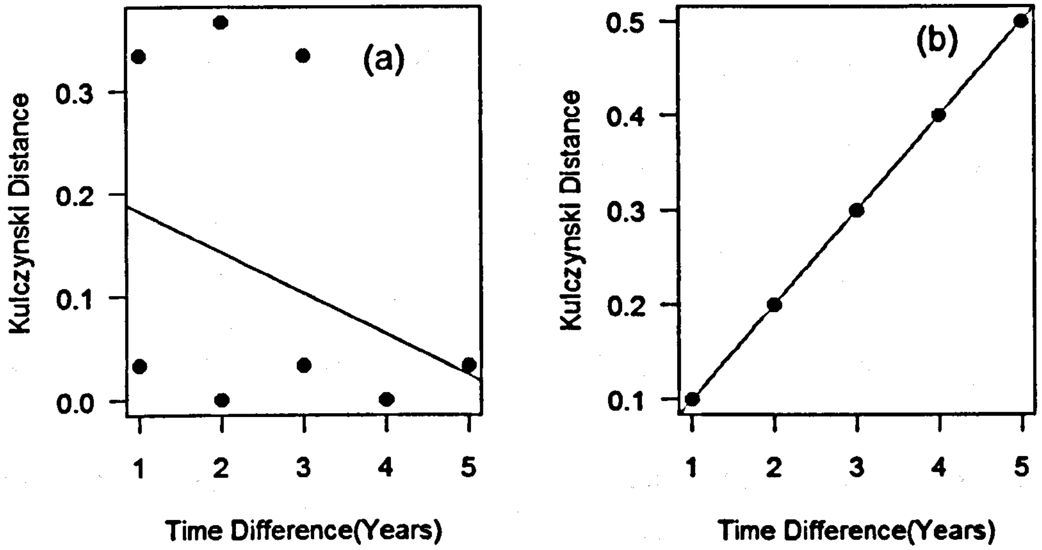


Figure 2.9: The dissimilarity between communities with increasing time lags based on; (a) a hypothetical community with a markedly different structure in year 3 representing the impact of a pulsed disturbance; (b) a community exhibiting directional change in structure through time.

2.3.5: Conclusion

In the preceding sections a measure of temporal variability has been outlined that allow the characterisation of temporal variability in communities over long periods of time. Although the choice of dissimilarity metric used has the potential to influence findings the Kulczynski distance is considered to be robust. Based on the measure of temporal variability outlined above, the impact of the loss of a single taxon from a community will be lowest in the communities with high richness. Conversely, the effect of the loss of similar proportions of taxa will be highest in the most taxon rich communities.

CHAPTER 3: SPATIAL PATTERNS IN THE TEMPORAL VARIABILITY OF MACROINVERTEBRATE COMMUNITIES IN RIVERS

3.1: Introduction

Long-term studies have made important contributions to our understanding of aquatic and terrestrial systems, and are essential for understanding the role of gradual and cyclical processes, or infrequent events, in shaping communities (Franklin et al. 1990; Jackson and Fureder, 2006). In lotic systems long-term studies have suggested that factors such as flow regime (Meffe and Minckley, 1989; Scarsbrook, 2002), rainfall and drought (McElravy et al. 1989), climatic cycles (Bradley and Ormerod, 2001), and constancy of habitat (Johnson et al., 1994; Robinson et al., 2000; Brewin et al., 2000; Scarsbrook, 2002) can have important effects on how community structure changes over time. However, our scope for examining these drivers in greater depth is limited by the scarcity of long-term data.

Jackson and Fureder (2006) show that of 5645 studies examining stream macroinvertebrates published between 1987 and 2004, only 46 were 5 years or more in length. Long-term studies increase the probability of observing events or processes that are determinants of community structure. In short-term studies greater importance may be assigned to unusual events leading to erroneous conclusions about their overall importance (Weatherhead, 1986). Similarly, communities that appear to be unstable over short time periods may be cyclical in nature leading to a changing understanding of their dynamics with increasing study length (Rahel, 1990).

Coupled with this lack of temporal data most long-term studies of lotic communities have a limited spatial extent. Studies most commonly range from a single site to a few sites within a river (e.g. Boulton et al., 1992; Daufresne et al., 2003; Beche et al., 2006) with only a limited number considering greater spatial extents, for example the catchment (e.g. Bradley and Ormerod, 2001; Brewin et al., 2000) or national (Scarsbrook et al., 2000) scale. In many systems important environmental factors are often spatially structured (Dormann et al., 2007). The lack of temporal data at large spatial scales limits our ability to examine the role of different environmental factors in controlling variation through time. As with macroinvertebrate community structure, it would be expected that spatial patterns in temporal variability will reflect important physical and chemical gradients (Li et al., 2000). Examining these gradients is most readily achieved at small scales (Irvine, 2004) however results from such studies may not be transferable to other systems where the importance of local drivers may be affected by processes acting at a larger scale (Weatherley and Ormerod, 1990).

The lack of large scale temporal data may have profound implications for management practices in freshwater systems as biological assessment methods based on macroinvertebrate communities often assume persistence in the absence of anthropogenic stress (Richards and Minshall, 1992; Bunn and Davies, 2000; Robinson et al., 2000; Metzeling et al., 2002; Milner et al., 2005). To have confidence in the conclusions that are drawn from biological monitoring it is essential to understand and quantify the level of expected variability (Irvine, 2004). Difficulties in making the commitments of resources required to instigate and maintain large-scale, long-term sampling programmes have until recently limited our ability to examine such patterns (Elliott, 1990). However, the development of national biomonitoring programmes for water quality have themselves provided resources that allow research to be conducted at large spatial scales (for example Moss et al., 1987; Scarsbrook et al., 2000; Johnson et al., 2007; Chessman, 2009). With their continued use such programmes have started to yield long-term data that increasingly form the basis for studies of long-term change (Jackson and Fureder, 2006).

Here the spatial and temporal extent of a biomonitoring dataset are utilised to examine spatial structure in the patterns of temporal variation in macroinvertebrate communities in English and Welsh rivers. The principal aim is to identify spatial patterns and scales of community change in order to evaluate the relative influence of regional and local factors in controlling temporal variability of macroinvertebrate communities. This is the first step towards identifying the types of environmental, or other, processes that might be important in controlling community dynamics.

3.2: Methods

3.2.1: Monitoring Data

In England and Wales the Environment Agency (EA) currently conducts a programme of General Quality Assessment of rivers based on macroinvertebrate communities. The technique was developed from 1977 onwards and compares the observed fauna at a site with the fauna predicted to occur based on a range of environmental conditions, and in the absence of anthropogenic stress (Wright, 2000). The data resulting from the national monitoring programme comprise over a quarter of a million samples from 62765 sites and sampling information is available covering a period from the late 1970s to the present day. Sample sites are located in most rivers in England and Wales and information for each site details the date of the sample, the identity and log abundance category of each taxon (Murray-Bligh et al., 1997), along with environmental data used to predict the expected community (Wright, 2000).

Sampling takes place in spring (March – May), summer (June – August) or autumn (September – November) mainly on a rolling three year programme. Samples are collected using a standard protocol: a three minute kick sample and one minute manual search in which all the major habitats in the reach are sampled proportionally (Wright, 2000). Taxa are generally identified to family or species and log abundance scores assigned. Following identification, sample data are submitted to a central repository where they are stored in a database called BIOSYS.

3.2.2: Data selection and the BIOSYS dataset

The BIOSYS database has been designed to store information for all the EA's monitoring programmes. It contains records for sampling programmes carried out using a number of different techniques. To ensure comparability of samples a subset of the BIOSYS database was compiled that contained information collected using the method described above. Because the EA uses a rolling programme of site visits very few sites had contiguous yearly data available. Instead site selection was based on the availability of at least 6 samples taken between 1990 and 2005. As there is known seasonal variation in macroinvertebrate community composition due to life histories (Rosillon, 1985) each season was examined separately. Spring samples provided the best spatial coverage.

Due to the time span of the data collection, various taxonomic changes have occurred as classifications have been revised. Names were updated to the most recent taxonomy using the Coded Checklist of Animals Occurring in Fresh Water in the British Isles (Furze, 2007). An additional complication is that there are sometimes differences in the taxonomic level at which specimens are recorded both through time at a single site, and between sites. To remove the influence of this variation on measurement of community difference, identification was standardised to the family level (Bradley and Ormerod, 2001; Metzeling et al., 2002).

3.2.3: Measuring community change

There are many techniques for comparing compositional changes in communities based on presence/absence data (Koleff et al., 2003) or abundance (Legendre and Legendre, 1998). Each measure has strengths and limitations and the most appropriate measure will be determined by the nature of the data and the aim of the study. In this instance the aim was to produce a single summary value for a site that describes the total change in the community through time based on the identities and relative abundances of taxa present.

The measure of temporal variability used in this study is similar to the measure of Functional Diversity proposed by Petchey and Gaston (2002). The measure is based on the total branch length of a dendrogram constructed using information about community composition in samples taken over time at a site. First, the similarity between communities is calculated to produce a diagonal matrix of pairwise comparisons. This matrix is used to construct a dendrogram with each of the branches representing a sample occasion (year). As dissimilarity of communities increases so does the length of the branch between any two points. Total branch length therefore represents a measure of dissimilarity across all the samples at a site, through time, giving an aggregate measure of community change. In this method the choice of both the distance measure and linkage method are important as they influence the shape of the dendrogram and so the total branch length. To inform the choice of distance measure and linkage method artificial samples representing constancy, directional trends and stochastic variation over time, were simulated. The combination of Kulczynski's distance (Faith et al., 1987) and average linkage clustering was shown to produce results that most clearly distinguished the pattern of change in the simulated communities.

As total branch length is sensitive to the number of samples used, this was standardized to six samples per site, by random selection from samples taken between 1990 and 2005. To gain an understanding of the impact that using such a random approach would have on measurements of temporal variability, the performance of this approach was compared with contiguous yearly data from 1990 to 1999. These data represent all sites identified in the dataset that have a continuous run of yearly sample information. The two approaches yielded results that were strongly correlated (Pearson's correlation, $r = 0.89$, $n = 43$, $p < 0.001$). The advantage of the random approach is that it increases the number of sites available for this study from 43 (for contiguous years) to 1574 (random years) dramatically improving the spatial extent of data available.

The type of community change was characterised as being either stochastic or directional using a method adapted from Collins et al. (2000). The method is based on regression analysis of the relationship between the time interval between pairs of samples and the difference in their community composition. Using this technique it is possible to characterise temporal variability into three categories based on the regression line. Firstly, if communities are increasingly dissimilar with increasing time lag between them this will result in a positive slope. Secondly, a zero slope indicates no directional change in the community. The dissimilarity in community composition is relatively constant irrespective of time lag between samples. Thirdly, a negative slope will characterise communities as having changed and then returned to a state similar to the original composition. In this instance points at the largest and smallest time lags will be most similar. This slope may be indicative of communities exhibiting cyclical dynamics or recovery from a perturbation.

To separate significant from non-significant relationships Mantel tests (1000 permutations) were used to test the relationship between the distance and dissimilarity matrices (Legendre and Legendre, 1998). In taking this approach the power to distinguish between stochastic variation (no slope) and cyclical dynamics or recovery from a perturbation (negative slope) is lost as the latter may be best described by non-linear regression making them unsuitable for examination with a Mantel test. In the context of this study it was considered more important to distinguish between communities exhibiting a trajectory of change (positive slope) and those exhibiting longer term stability due to either stochastic variation (no slope) or cyclical dynamics (negative slope). Subsequently the term stochastic is used to imply a community that does not exhibit a directional change in community composition.

The distribution of sites exhibiting either directional or stochastic change was examined for regional bias. ArcGIS was used to divide the country into a ten by ten grid. Squares with no sample points in were deleted. This grid size was chosen as it kept a statistically valid number of sample sites in each of the squares. For both directional and stochastic change a bootstrapped chi-squared statistic (10000 permutations) was calculated for the difference between the observed number of points in each grid square and the number that would be expected by chance. The contribution of each square to the overall chi-squared statistic was then examined to gauge bias towards a particular type of change.

3.2.4: Identifying spatial structure

Spatial structure was examined using Principal Coordinate Analysis of Neighbour Matrices (PCNM) (Borcard and Legendre, 2002). PCNM analysis is a new approach for examining spatial structures based on the idea that the spatial arrangement of data points can be translated into explanatory variables for use in regression or multivariate analysis (Dormann et al., 2007). The technique has been used to examine spatial patterns in chlorophyll *a* concentrations (Borcard et al., 2004), orabitid mites (Borcard et al., 2004; Dray et al., 2006), littoral fish communities (Brind'Amour et al., 2005), bird species richness (Diniz-Filho and Bini, 2005), defoliation by the spruce budworm (Bellier et al., 2007) and lake macrophyte communities (Briers, 2006). It offers the advantage over other techniques, such as trend surface analysis, in that it is able to identify all spatial scales that can be represented within a given dataset (Borcard et al., 2004). In this way PCNM analysis provides information about the dominant scales at which processes are acting on communities. Detailed descriptions with examples of its application are provided by Borcard and Legendre (2002), Borcard et al. (2004), Diniz-Filho and Bini (2005), Dray et al.(2006), Griffith and Peres-Neto (2006), and Dormann et al. (2007).

PCNM analysis is based on an ordination technique called Principal Coordinates Analysis (PCoA). A more in-depth discussion detailing the mathematical basis for the technique is provided by Legendre and Legendre (1998). Here I provide a simplified version designed to illustrate how spatial descriptors are derived.

Spatial coordinates of the 1574 sites are used to construct a matrix containing Euclidean distances between all pairwise combinations of sites. This matrix is truncated (Borcard and Legendre, 2002) and the resultant neighbour matrix submitted to PCoA. As with other ordination techniques the purpose of PCoA is to produce linear combinations of variables that summarise variation in the original dataset. PCoA produces a series of eigenvalues with associated eigenvectors that describe the data in ordination space. To understand PCNM it is useful to examine what these eigenvalues and eigenvectors represent.

The eigenvalues, also termed latent roots (Quinn and Keough, 2002) represent the amount of variance explained by each of the axis in the ordination. As the PCoA is based on Euclidean distances the largest eigenvalues represent the greatest variance in the data and therefore the largest spatial scale. Decreasing eigenvalues represent progressively smaller scales. Associated with each eigenvalue is a list of eigenvectors that indicates the contribution of each of the original variables (sample sites) to that axis of the ordination.

Quinn and Keough (2002) states that one way of understanding PCoA is as a translation of dissimilarities between objects into Euclidean distances representing their arrangement in multidimensional space. The author considers that when the original dissimilarities are based on Euclidean distances the resultant ordination will be a representation of the complete spatial structure of the original dataset. In essence this is what PCNM analysis produces. Although there are complications due to the truncation process, which results in negative eigenvalues (Borcard and Legendre, 2002), the overall result is the production of a description of the complete spatial structure in the data, with each axis describing a different spatial pattern at varying scales.

To illustrate this further Figure 3.1 shows a spatial filter defined by the largest eigenvalue. This represents the most variance in the data and therefore the broadest scale. The values for the eigenvectors of each sample are represented as squares. The size of squares is proportional to the value of their eigenvector with white and black squares indicating negative and positive values respectively. The largest squares therefore represent the sites that contribute most to that axis of the ordination.

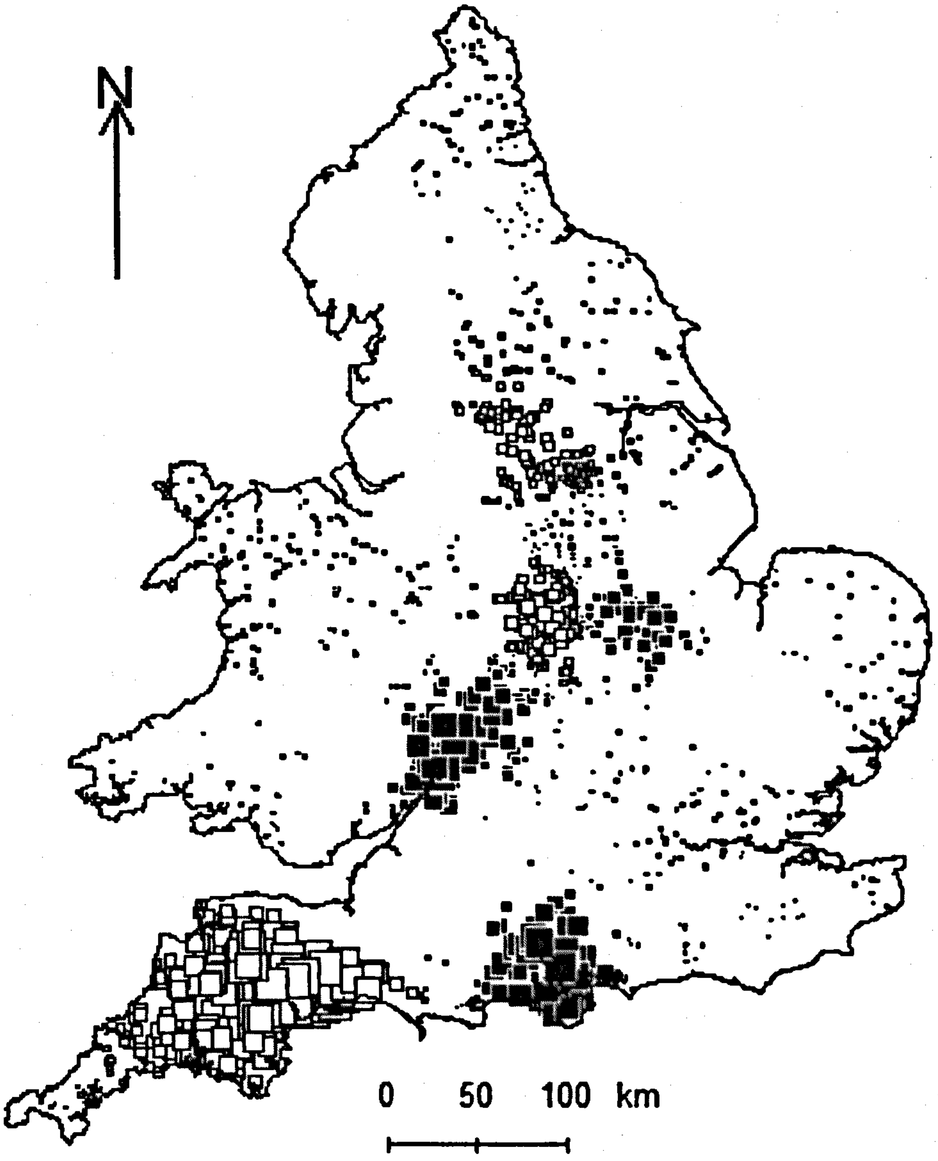


Figure 3.1: Moran's eigenvector map representing the largest spatial filter produced from PCNM analysis. Black squares indicate positive and white squares negative eigenvectors with the size of the squares represent the contribution of each specific sample point to the axis of the ordination.

Taking each eigenvalue in turn and plotting the eigenvectors of each sample point reveals all of the spatial patterns described in the data. These spatial filters can now be used as explanatory variables to examine patterns of spatial autocorrelation. Retaining all the spatial filters associated with positive eigenvalues may overcorrect for spatial autocorrelation (Griffith and Peres-Neto, 2003). Following Whittingham et al. (2006), Aikaike's Information Criterion (AIC) was used to select a subset of spatial filters in a multiple regression that represents patterns of temporal variability at differing spatial scales.

All calculations were carried out using R 2.8 (R Development Core Team, 2008).

3.3: Results

A total of 1574 sites met the criteria for measuring temporal variability. Long-term data were unavailable for the north west of England due to differences in sampling method on some occasions. The distribution of sites used in the study is shown in Figure 3.2.

Large scale patterns in temporal variability are shown in Figure 3.3. This figure represents an average value for community change for catchments across England and Wales. Catchment boundaries were derived by the Centre for Ecology and Hydrology using Digital Elevation Models and an automated GIS tool. There are gaps in the coverage (white areas on Figure 3.3) that represent catchments with no study sites within their boundaries. Catchment characteristics exert a considerable influence on communities so provide a useful unit with which to visualise large patterns. Figure 3.3 is characterised by a region of high change corresponding to catchments in central England. Ignoring catchments in the north east there is a trend towards increasing temporal variability from Cornwall and Wales in the west of the country to the Anglia region in the east. Within the north east of England there is a more complex pattern of temporal variability with a patchwork of high and low levels of change.

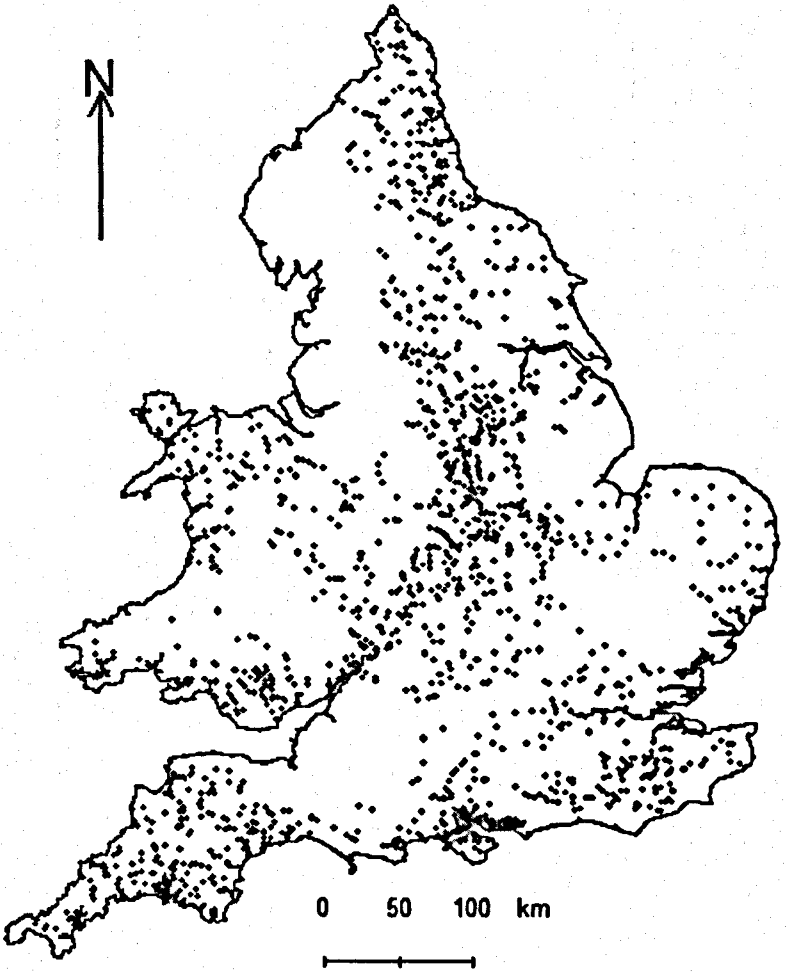


Fig 3.2: Distribution of 1574 sample sites used in the current study.

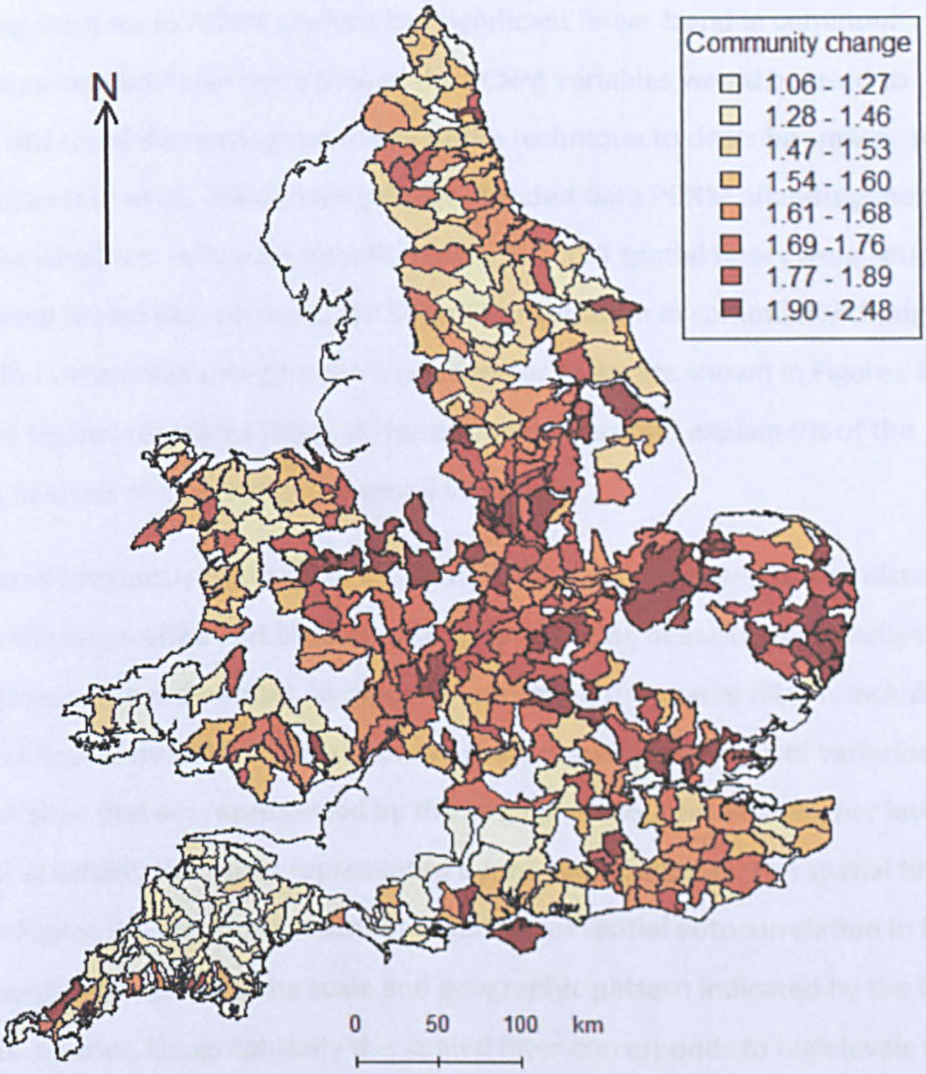


Figure 3.3: Levels of macroinvertebrate community change over 15 years by catchment. White areas represent areas with no data available.

Representation of the data in this way is useful in providing an initial visual analysis. Whilst patterns may be complicated by the irregular shape of catchments and uneven distribution of sampling sites this description is supported by a significant linear trend detected based on the latitude and longtitude of individual points (r^2 = 0.147, F-value = 136, d.f. = 2, 1571, p-value <0.001).

However, a considerable amount of the variation in community change remains unaccounted for by this broad scale trend. PCNM analysis provides an analysis of the underlying point pattern that is able to consider smaller scale processes. Before submitting the data to PCNM analysis the significant linear trend in community change was removed otherwise 50% of the PCNM variables would be used to describe this trend decreasing the ability of the technique to describe smaller scale patterns (Borcard et al., 2004). Using the de-trended data PCNM analysis generated 830 PCNM variables. Following selection using AIC, 193 spatial filters were retained in the spatial model that accounts for 54% of the variation in community change. Spatial filters visualised using Moran's eigenvector maps are shown in Figures 3.4 to 3.6. These figures represent three of these spatial filters and explain 9% of the variation in levels of community temporal variability.

As discussed previously the scale of the spatial filter is represented by the distance between the large white and black squares (representing negative and positive eigenvectors in the ordination). In multiple regression the spatial filter is included in the overall spatial model if it contributes to understanding patterns of variation. In Figure 3.4 sites that are represented by the large black squares have higher levels of temporal variability than sites represented by the white squares. The spatial filter shown in Figure 3.4 therefore describes a pattern of spatial autocorrelation in levels of community change, with the scale and geographic pattern indicated by the black and white squares. Geographically this spatial filter corresponds to high levels in South Yorkshire and Nottinghamshire.

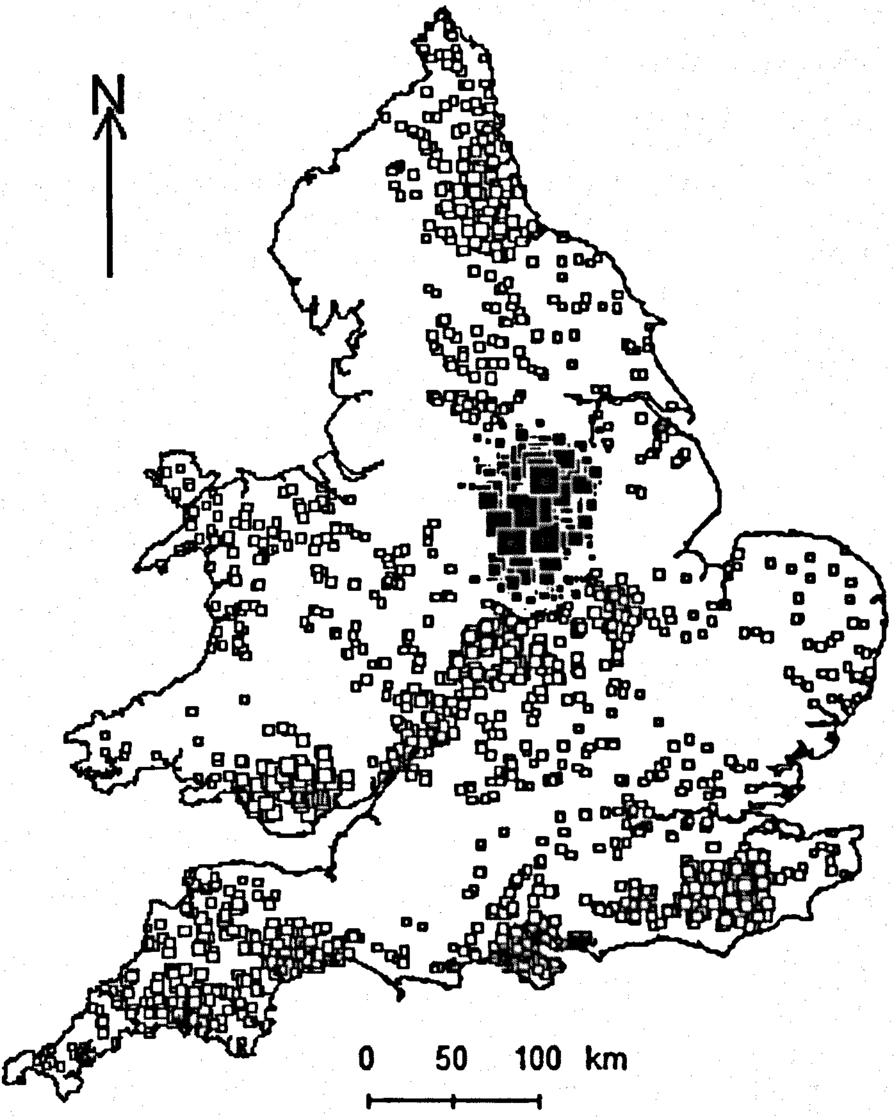


Figure 3.4: Moran's eigenvector map illustrating a spatial pattern in temporal variability. The filter describes spatial autocorrelation between sites exhibiting high (black squares) and low (white squares) levels of community change.

Figure 3.5 represents the largest spatial filter describing spatial autocorrelation in the extreme south east of the country. This filter captures low levels of community change in this region compared with sites at a distance of circa 150 km away. Geographically this corresponds to low levels of change in Cornwall and Devon compared with higher levels around the River Severn and in Hampshire. Note that whilst this figure represents that same PCNM filter as shown in Figure 3.1 colours have been reversed to indicate sites with higher (black) and lower (white) temporal variability. This scheme is consistent across all figures where patterns of community change are described in the current chapter in order to aid interpretation. However, in other studies spatial filters are often represented using black squares to indicate positive eigenvectors and white squares to indicate negative eigenvectors.

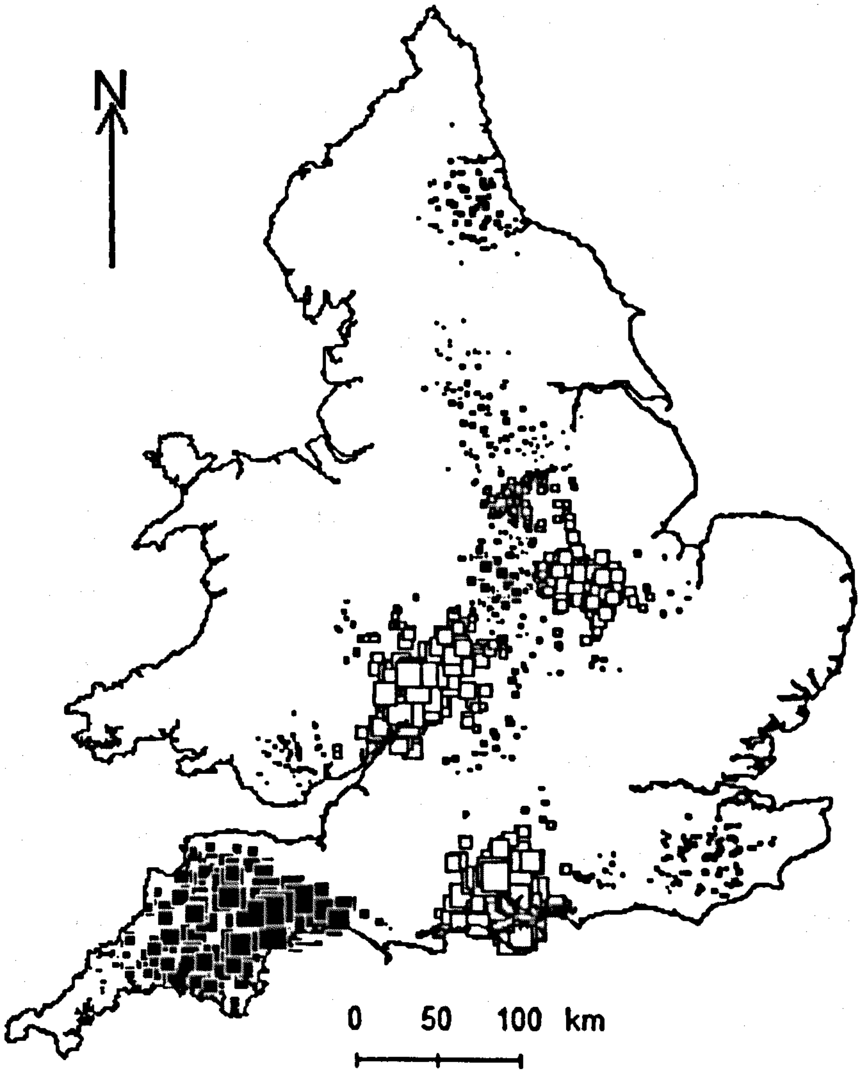


Figure 3.5: Moran's eigenvector map illustrating a spatial pattern in temporal variability. This figure represents the largest spatial filter. The filter describes spatial autocorrelation between sites exhibiting high (black squares) and low (white squares) levels of change.

Figure 3.6 describes spatial patterns, within the area of high community change at the centre of the country, at a scale of circa 40 km. Once again the larger white and black squares represent negative and positive ends of the ordination axis respectively, with lower levels of community change associated with the white squares.

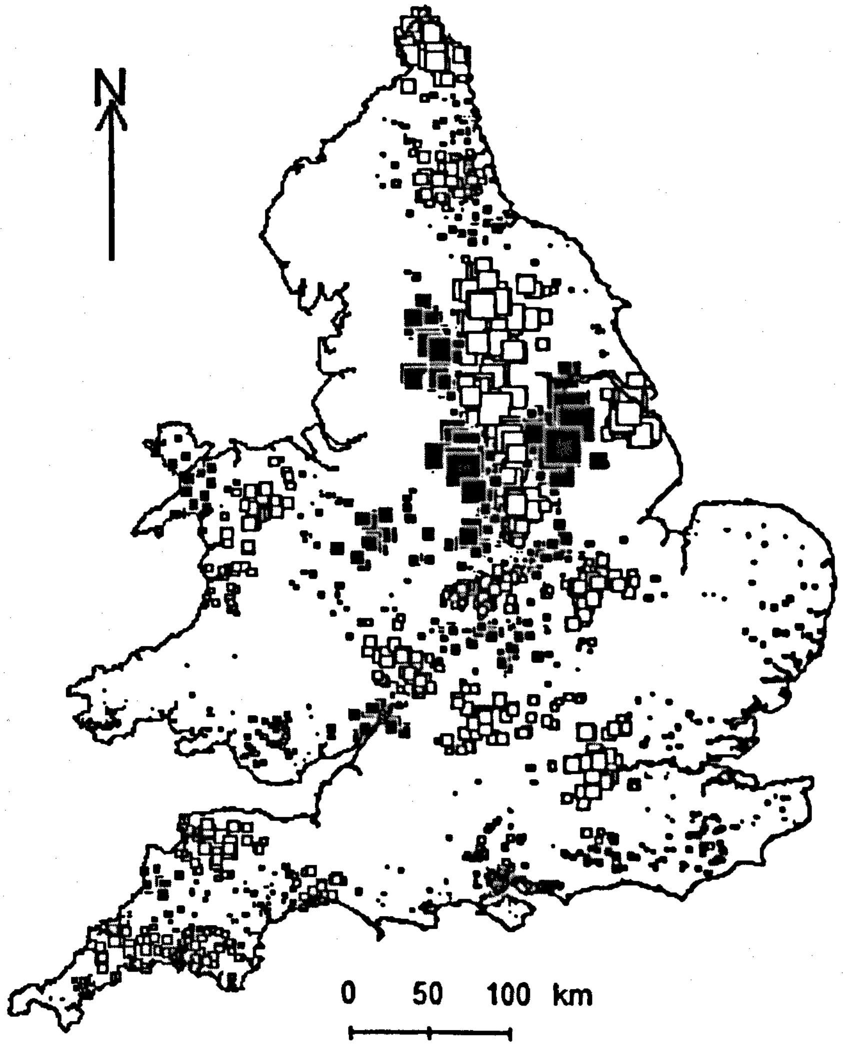


Figure 3.6: Moran's eigenvector map illustrating a spatial pattern in temporal variability. This figure illustrates the complexity of spatial relationship that can be described. The filter describes spatial autocorrelation between sites exhibiting high (black squares) and low (white squares) levels of change.

A comparison of Figure 3.3 (community change by catchment) and Figure 3.4 to 3.6 indicates some agreement in patterns of community change at these large scales. For example high levels of change in the centre of England are shown in both figures. At the scales represented in Figure 3.4 to 3.6 visualising patterns in this way provides a useful insight into variance within the data. However, PCNM analysis produced 193 spatial filters. The interpretation of smaller scale filters is more problematic as such representations results in extremely complex maps. It is important to examine the contribution of these small scales as they represent forces acting on the community.

Figure 3.7 describes the relationship between the scale of the filter and its contribution towards the total explanatory power of the spatial model. This figure reveals spatial structuring at three distinct scales differentiated by steps in the cumulative r-squared value at 35 and 83 kilometres. Forty two percent of the total explanatory power of the model is accounted for by small scale filters indicating relationships between sites less than 35 km apart. The next contributions towards the model are from spatial filters from 35 - 83 km that together account for a further 42% of the explanatory power of the model. Finally, 16% of the explanatory power of the model is composed of a five spatial filter with scales of between 83 and 150 km. It can therefore be concluded that community change is been driven by factors operating regionally (Figure 3.4 and Figure 3.4 to 3.6) and locally (as indicated in Figure 3.7).

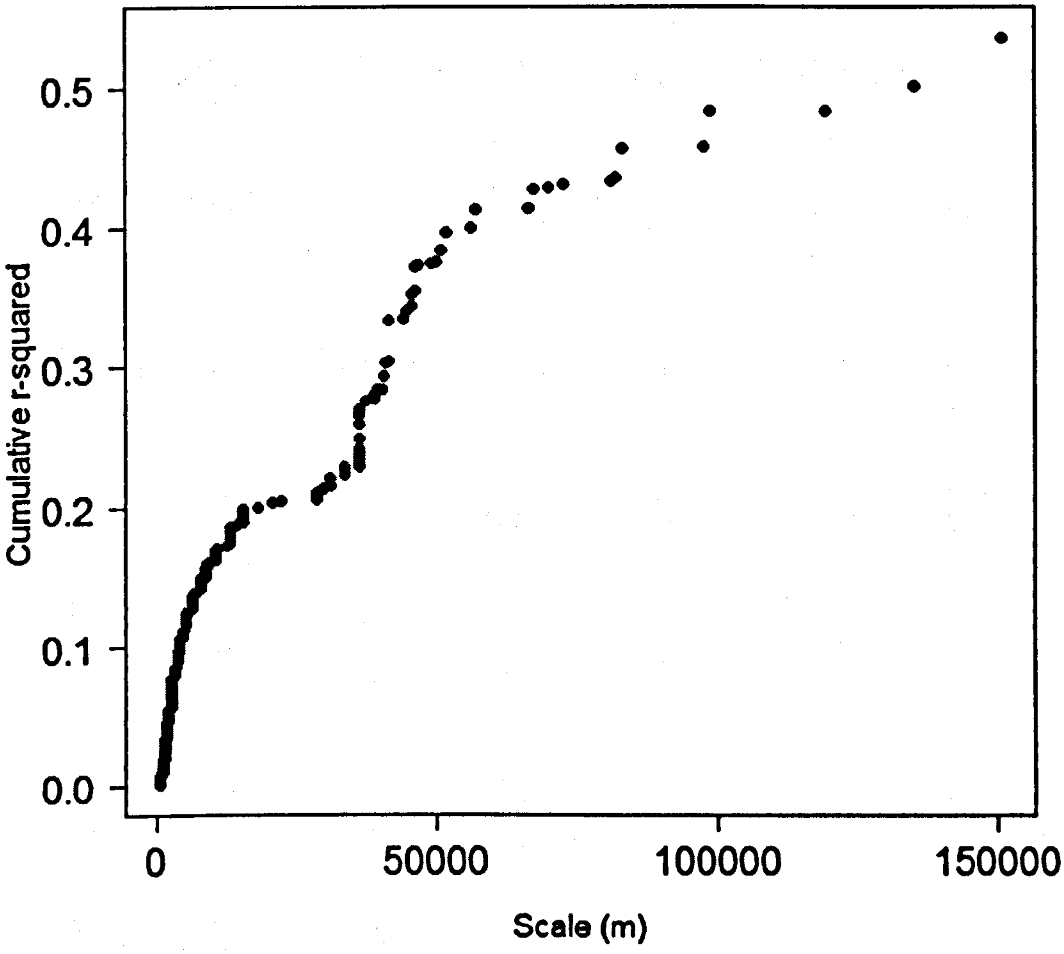


Figure 3.7: Contribution of different scale spatial filters to r-squared value of the spatial model.

Analysis of the type of community change showed 75% of sites had communities exhibiting stochastic variation in composition between 1990 and 2005 as indicated by non-significant Mantel tests for the community/time distance matrices. The distribution of these points is shown in Figure 3.8a. The distribution of sites with a significant Mantel test and therefore exhibiting directional change in community structure is shown in Figure 3.8b. To examine whether there was a bias in the distribution of sites exhibiting the two different forms of change was examined using chi-squared tests based on the number of points within a grid with a cell size of 50 km x 50 km. As a number of grid squares had either no points or low numbers a randomisation procedure (1000 permutations) was used to generate bootstrapped confidence intervals to test whether the distribution of points was other than would be expected by chance. For communities that exhibited stochastic variation through time there was no evidence of a bias in the distribution of points (Chi-squared = 51.54; bootstrapped 95% confidence intervals 41.41 to 75.85). However, for communities that exhibited directional change there was evidence that the points were not distributed randomly (Chi-squared = 145.8; bootstrapped 95% confidence intervals 42.23 to 77.97). As can be seen in figure 3.8b there is a clear bias towards directional change for sites in the centre of the country.

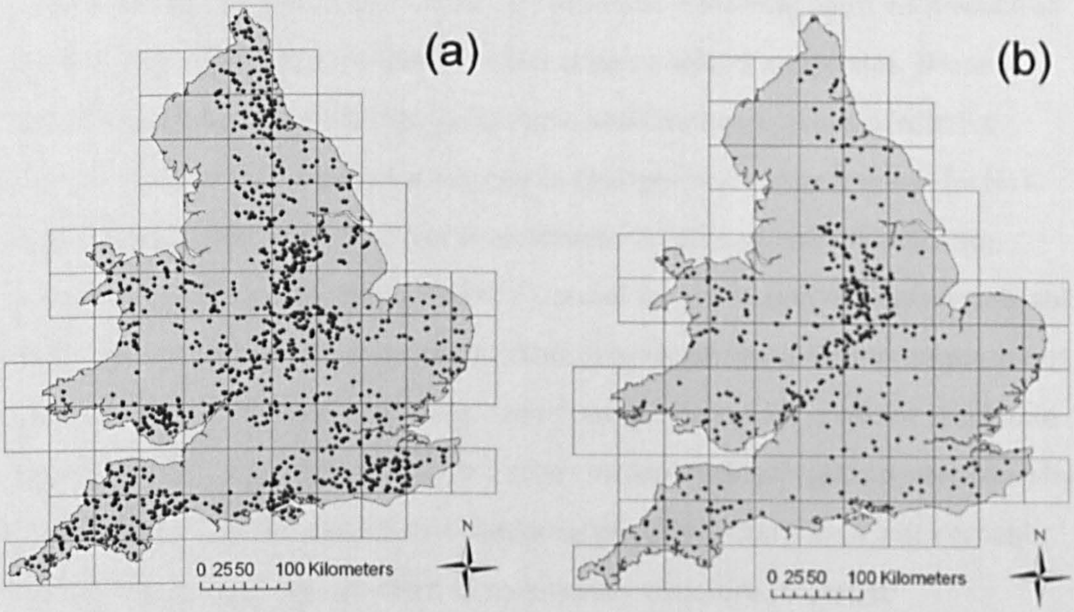


Figure 3.8: Analysis of the types of change indicating (a) stochastic variation or (b) directional change in community structure.

3.4: Discussion

Temporal variation in macroinvertebrate community structure has not been examined at the spatial extent of the present study before in UK rivers and rarely in other lotic systems (e.g. Scarsbrook et al., 2000; Chessman, 2009). The present study has demonstrated spatial structure in temporal variability at regional and local scales suggesting a number of possible mechanisms governing community structure through time. In addition to the amount of variation, communities were characterised as exhibiting either directional or stochastic change in their structure. Communities in the south and east of England exhibited stochastic variation through time with communities varying around an average composition. Sites in the centre of the country exhibited an increased bias towards directional change with increasing dissimilarity in community structure over a 15 year period. These contrasting forms of variation imply different mechanisms are driving temporal variability in the community.

Although not considering temporal variation, studies by Wright (2000) and Murphy and Davy-Bowker (2005) were conducted at a similar spatial scale using data that forms part of the current study. As such they provide a starting point with which to consider factors that influence communities across England and Wales. These previous studies demonstrate that geographic position has a strong predictive power for community composition related to changes in physicochemical factors. During the development of the River Invertebrate Prediction and Classification System (commonly referred to as RIVPACS), a tool for predicting macroinvertebrate community structure in the absence of anthropogenic stress, a distinct sequence of macroinvertebrate assemblages along a north west, south east gradient at pristine sites were identified (Wright, 2000). In a study incorporating impacted sites Murphy and Davy-Bowker (2005) described a similar large scale pattern with a discernable north/south and east/west gradient in community structure related to physicochemical gradients and environmental stress.

Patterns of community change in the current study at least in part reflect spatial patterns from this previous work. Figure 3.3 provides an indication of average temporal variability summarised by catchment and demonstrates a number of spatial patterns. Regression analysis of the underlying point data revealed a significant linear trend in community change across England and Wales accounting for 14.7% in the total variation of communities through time. Higher levels of change in the centre and towards the east of the country, and lower levels of change to the west and the north, correspond to physicochemical and anthropogenic stress gradients described by Wright (2000) and Murphy and Davy-Bowker (2005). Murphy and Davy-Bowker (2005) identified urban run-off and organic pollution together with inputs associated with agriculture as the most influential factors disrupting lotic macroinvertebrate communities across England and Wales. The broad scale trend in community change might plausibly be linked to these factors. For example increasing variation in community change towards the east of the country corresponds to areas of high agricultural production. Conversely areas of low community change to the south west correspond with areas of comparatively low agricultural production and urbanisation. These relationships are explored in further detail in chapter 5.

The linear relationship is complicated by considerable landscape heterogeneity which may account for the relatively low amount of variation explained. For example in the north east of England and Wales figure 3.3 indicates a complex patchwork of catchments exhibiting high and low levels of community change. In these instances catchments exhibiting high levels of variation often correspond with major metropolitan areas. Aquatic invertebrate community structure is known to respond strongly to urbanisation (Roy et al., 2003; Snyder et al., 2003) with watershed scale variables overriding the importance of local physical and chemical properties (Urban et al., 2006). This justifies the use of the PCNM technique which is able to examine smaller scale spatial relationships such as may arise within a heterogenous landscape.

The more detailed analysis of the underlying point pattern provided by PCNM indicates broad regional differences in levels of temporal variability. For the three spatial filters presented in Figure 3.4 to 3.6 it is possible to identify areas of high variability in the centre of the country (Figure 3.4), finer spatial structure within this region (Figure 3.6) and, at the broadest scale, an area of low variability in the south west of the country (Figure 3.5). In total these and other large scale filters accounted for 58% of the total explanatory power of the spatial model describing levels of temporal variability across England and Wales. Temporal variability is clearly being influenced by factors acting at a regional scale. Previous studies indicate that constancy of habitat may be a central driver of persistence (Richards and Minshall, 1992; Johnson et al., 1994; Scarsbrook, 2002). The observed large scale patterns may result from regional differences in geology (Richards and

Minshall, 1992; Gibbins et al., 2001), climate (Fruget et al., 2001; Collier, 2007) and land use (Weatherley and Ormerod, 1990; Fruget et al., 2001; Roy et al., 2003) that influence the amount of disturbance to which the community is subjected (Poff and Ward, 1990).

As well as regional patterns PCNM analysis identified a considerable influence of smaller spatial filters (Figure 3.7) in describing temporal variability. Forty two percent of the total explanatory power of the spatial model is composed of spatial filters with a scale less than 35 km indicating considerable heterogeneity in patterns of community change across the study region. Lotic systems may be viewed as representing a longitudinal gradient of environmental conditions (Vannote et al., 1980) with localised characteristics having a strong influence on the community (Heino et al., 2007). In such systems it would be expected that differences in factors such as substratum (Davies et al., 2000; Malmqvist, 2002; Bonada et al., 2007), flow regime (Bonada et al., 2007; Meffe and Minckley, 1989; Burgherr and Ward, 2001; Gibbins et al., 2001; Scarsbrook, 2002), vegetation (Weatherley and Ormerod, 1990), and stream size (Beche and Resh, 2007) would result in differing levels of temporal variability over relatively small scales. Davies et al. (2000) considers that modified systems will be characterised by increased local variability in conditions that will exert an influence on communities shaping both structure and temporal variability. River systems in England and Wales are significantly impacted by human activities with nearly 90% of sites surveyed by the EA considered to be influenced to some degree (Davy-Bowker et al., 1999). Statzner and Higler (1986) state that the response of communities to an event may be different when they are living in sub-optimal compared to ideal conditions. The importance of small scale spatial patterns and the history and current status of English and Welsh rivers points towards local differences in physicochemical conditions as being potentially important determinants of temporal variability.

By demonstrating the importance of multiple scales in describing temporal variability the results of the present study also suggests that interaction between factors acting across these scales are important in controlling temporal variation. Collier (2007) considers that the influence of land use may be secondary to that of climate in controlling variability in community composition. However, the author reports that slight modification of the environment may lead to increased interannual variability in community composition suggesting an interaction between large and small scale processes. Townsend et al. (2004) demonstrates that such 'cross-scale' interactions may cause considerable variation in community composition. The difficulty in making generalisations about drivers of temporal variability based on previous studies may arise from complex interactions across scales that affect the influence of any one specific driver on the community. This finding has implications for monitoring procedures based on macroinvertebrate community structure where understanding and quantifying the level of expected variability is a stated goal (Irvine, 2001). Observed patterns will be dependent on the spatial scale and location over which temporal variation is examined (Downes et al., 1993). Studies of variability arising through sampling methodology (e.g. Clarke, 2000; Clarke et al., 2002) are often based on a detailed examination of samples from a few sites. Quantifying the level of expected variation in communities in such a way in the current context may produce results that over- or underestimate the level of variation dependent on local and regional conditions.

The type of variation in community structure at each study sites was characterised as indicating a bias towards communities exhibiting stochastic variation in composition through time. Such change has been considered as representing long-term stability in communities in previous studies (Gibbins et al., 1994; Scarsbrook, 2002) and may be indicative of the role of random processes in driving change where community composition alters in response to seasonal and interannual variation in conditions. For example a difference in hydrological regime between years has been demonstrated to drive variation in community structure (McElravy et al., 1989; Boulton et al., 1992; Wagner et al., 2000; Gibbins et al., 2001; Jackson and Fureder, 2006) with certain taxa being more sensitive to the impacts than others (Bradt et al., 1999).

In contrast to this stochastic variation, communities in the centre of the country were demonstrated to exhibit high levels of directional change in composition through time. Alteration of the community in this way suggests a fundamental and long-term influence on habitat conditions and an associated adaptation of the community. Davy-Bowker et al. (1999) describe a net improvement in water quality in English and Welsh rivers since systematic analysis began in 1990. Directional change in this instance may be a result of a shift in communities towards taxa characteristic of a clean water fauna (Scarsbrook et al., 2000). The regional bias towards directional change in central England corresponds to an area historically influenced by industrial and urban pollution (Murphy and Davy-Bowker, 2005). This supports changes in water quality as a potential driving force however, further analysis focusing on community composition is required to separate this from other potential drivers such as global climate change (Daufresne et al., 2003; Burgmer et al., 2007).

For both stochastic and directional change it is important to acknowledge that sampling error will also contribute to observed levels of variation (Clarke and Hering, 2006). Although all samples were collected using the same method differences between operators and site characteristics (Clarke et al., 2002) have been shown to influence sample accuracy along with the error that arises due to sample processing and taxonomic identification (Clarke and Hering, 2006). This source of error will make a more important contribution to our understanding of processes where communities are shown to be exhibiting stochastic variation as it will be difficult to separate the relative contribution of random variation caused by natural processes from that caused by sampling error. In communities characterised as exhibiting directional change, variation is due to a more fundamental shift in community composition through time and is unlikely to arise purely for methodological reasons.

In conclusion, little is known about the processes that control long-term variability in macroinvertebrate communities. Understanding such variability is of both fundamental interest and has practical implications for monitoring the health of ecosystems. The study identified a broad spatial pattern in temporal change across England and Wales that can plausibly be linked to known difference in environmental conditions. The use of a novel method of spatial analysis revealed that a large proportion of the total temporal variation on communities is explained by processes acting across multiple spatial scales beneath the broad scale trend emphasising the importance of regional and local scales in explaining patterns in variation. Within the context of previous studies of long-term change these patterns suggest that regional and local differences in the physicochemical environment exert control on the community. However, by demonstrating complex spatial structure results from the current study indicate that it may be difficult to make generalisations about the expected levels of variability in a community at a specific site where long-term data are unavailable.

CHAPTER 4: THE INFLUENCE OF ENVIRONMENTAL FACTORS ON THE TEMPORAL VARIABILITY OF LOTIC MACROINVERTEBRATE COMMUNITIES

4.1: Introduction

Studies in both pristine (e.g. Moss et al., 1987) and stressed (e.g. Murphy and Davy-Bowker, 2005) habitats have demonstrated that in lotic systems the physicochemical environment is the principal factor shaping macroinvertebrate communities (Death and Winterbourn, 1995). Factors such as flow regime (Voelz et al., 2000), temperature (Vannote and Sweeney, 1980; Ward and Stanford, 1982), substrate (Gurtz and Wallace, 1984; Brown, 2007) and land use (e.g. Allan, 2004) acting across multiple temporal and spatial scales (Minshall, 1988; Malmqvist, 2002) produce a Habitat Templet (Southwood, 1977; Southwood, 1988) that, depending on both historical and contemporary context, defines community structure (Robinson et al., 2002). To become established at a site a taxon must possess traits that allow reproductive success under a given set of conditions (Poff, 1997; Belyea and Lancaster 1999) defined by the spatial and temporal characteristics of the environment (Southwood, 1977; Southwood, 1988; Townsend and Hildrew, 1994). The significance of this process is that at any local site there will be selection for a specific suite of traits to be present within the community, and this will limit the response of the taxa to disturbance (Meffe and Minckley, 1987; Boulton et al., 1992; Wagner and Schmidt, 2004). Communities exposed to wide ranging conditions will contain taxa with traits conveying resistance or resilience allowing them to remain *in situ* or quickly recover following disturbance (Wallace, 1990). In systems that typically exhibit little environment variation communities may be more controlled by biotic interactions (Townsend et al., 1987; Richards and Minshall, 1992; Woodward, 2002), and as such may lack appropriate response mechanisms to disturbance leading to temporal variability in community structure (e.g. Armitage, 2006).

Whilst conceptually there is a clear link between the physicochemical environment and community variability, identifying which environmental factors are most important in governing this relationship presents a considerable challenge. Hierarchy theory (Simon, 1962) suggests that large scale factors should exert an overriding influence on communities as there effects are transferred down to progressively smaller spatial scales. This contrasts with multiscale theory (Wu and Loucks, 1995) that emphasises the importance of factors acting across spatial scales. In freshwater systems whilst some studies have highlighted the importance of large scale factors such as climate (Bradley and Ormerod, 2001; Beche and Resh, 2007), others consider that catchment (Sponseller et al., 2001; Roy et al., 2003; Urban et al., 2006), or reach scale (Richards and Minshall, 1992; Brown, 2007) properties may be most important in controlling community temporal variability. Although certain drivers of change emerge as a common theme in many studies, the hierarchical nature of river systems and the interactions between different environmental factors across multiple scales (Frissell et al., 1986; Allan, 2004) make it difficult to separate out specific influences on variability. Large scale factors (e.g. catchment area, underlying geology, catchment relief) may often subsume the importance of smaller scales factors in controlling communities (Davies et al., 2000). For example Walsh et al. (2007) demonstrated that the beneficial influence of riparian buffers on stream communities can be rapidly lost due to the large scale influence of urbanisation. However, it has been demonstrated that small scale factors may in turn reduce the influence of larger scale processes, for example by the provision of refuges for taxa during hydrological events (Gurtz and Wallace, 1984; Richards and Minshall, 1992) suggesting that rather than considering the influence of environmental factors operating in a top down way, a multiscale approach with interactions across spatial scales may be most appropriate.

As stream ecosystems represent one of the most heavily degraded habitats in the world with both regional and local environmental conditions being heavily modified (Boon, 2000; Johnson et al., 2007) it is important to understand the influence that this altered Habitat Templet might have on macroinvertebrate communities. Anthropogenic stress will often result in a change in catchment or reach scale properties and as such may interrupt the interactions between factors at different scales (Allan, 2004) potentially altering temporal properties of communities and affecting ecosystem integrity. Opportunities to explore the link between temporal variability and environmental factors are limited by the relative scarcity of long-term spatially explicit data (McElravy et al., 1988; Jackson and Fureder 2006). With the majority of studies being under five years in length (Jackson and Fureder, 2006) there is a problem of perception in defining which environmental drivers are affecting community variability. Weatherhead (1986) suggests that short term studies may tend to overestimate the importance of unusual events (e.g. floods, droughts) for the community. As aquatic communities show rapid recovery from disturbance (Townsend and Hildrew, 1976; Voelz et al., 2000) a long-term perspective is important for understanding the relationship between environmental factors and community variability. By understanding the long-term level of temporal variability and how this relates to environmental factors it is possible to establish a 'base-line' with which to judge the impact of specific stressors thus putting the present into context (Elliott, 1990; Magnuson, 1990). Of equal importance is the availability of large scale data as a lack of such data reduces our ability to examine temporal variability of macroinvertebrate communities under contrasting sets of conditions (Collier, 2007) limiting our understanding of what drives change in a heterogeneous landscape.

In the present study data collected as part of an environmental monitoring programme across England and Wales was used to explore the relationship between temporal variability of macroinvertebrate communities and environmental factors. Long-term temporal variability of macroinvertebrate communities has only rarely been examined in UK river systems (e.g. Townsend et al., 1987; Gibbins et al., 2001; Woodward et al., 2002; Armitage, 2006; Durance and Ormerod, 2009; Ormerod and Durance, 2009) and never at the spatial extent represented by the current study. Environmental data were derived from a range of source including GIS layers (e.g. climate, land use) and data collected specifically to assess the condition of river systems (River Habitat Survey: Raven et al., 1997; RIVPACS: Wright, 2000). There is a considerable problem in analysing and interpreting such data where relationships may be non-linear and unbalanced making traditional statistical techniques difficult to use (Urban et al., 2006). To overcome this problem a novel machine learning technique, Random Forests (Breiman, 2001), was used to address the primary objective, which was to examine the relationship between environmental factors and long-term temporal variability of lotic macroinvertebrate communities across a broad range of environmental conditions.

4.2: Methods

4.2.1: Macroinvertebrate data

Macroinvertebrate community composition was derived from data collected as part of the Environment Agency's (EA's) General Quality Assessment (GQA) of rivers. Sample are collected from most river in England and Wales using a standard method (Murray-Bligh et al., 1997) based on a three minute kick sample and one minute manual search. Sorting and identification of preserved samples takes place in the laboratory where taxa are identified primarily to family level. Following identification a proportion of sample are subjected to Analytical Quality Control (AQC) to measure the final quality of the results of the survey. Following AQC, survey results are entered into BIOSYS, a relational database that contains information from all biological monitoring conducted by the EA.

Data from 3305 sites were extracted from BIOSYS representing sites where macroinvertebrate samples were collected between 1990 and 2005. To maximise the number of sites available for analysis, sites were selected where the first sample was taken in 1990 or 1991 and the last sample in 2004 or 2005, with at least four separate sampling occasions between. For each sample names of taxa were updated to the most recent taxonomy using Furze (2007). Within the data there were inconsistencies in the level of identification through time with a mixture of species and family level information. As this would have a considerable impact on any calculations of community change (Bradley and Ormerod, 2001; Metzeling et al., 2002) identification was standardised to the family level.

4.2.2: Environmental Data

The first group of environmental variables represent those associated with the GQA method (Wright, 2000). These variables describe site specific conditions including substrate composition, channel width, channel depth, slope and discharge category, as well as broader scale variables including stream order, altitude and distance from source (Murray-Bligh et al., 1997). Environmental attributes are characterised as being either time invariant (e.g. altitude) or time variant (e.g. channel width). For each site the first sampling occasion (i.e. 1990 or 1991) was taken as the baseline year and environmental attributes extracted from BIOSYS. For time variant measures all subsequent values were extracted from the database and measures of variability through time calculated.

A second set of environmental attributes were derived from the River Habitat Survey, a method developed to assess the naturalness of rivers based on their physical attributes and focusing particularly on those features considered to be important for wildlife (Raven et al., 1997). Fox et al. (1998) provides a comprehensive review of the development of this system together with details of the variables recorded. The EA has undertaken a comprehensive assessment of much of the river network within the UK using this method. Using ArcGIS, spatial congruence between RHS and the 3305 GQA Biology sites was examined. Using a 1000 metre buffer, 415 paired GQA and RHS sites were identified, the distribution of which are shown in Figure 4.1. From this dataset a series of variables were extracted that describe both local conditions within the 500 meter survey length, and upstream characteristics based on average scores of RHS sites upstream of the survey length. Two sets of environmental data were extracted relating to the naturalness of the river system. First, four differing features that act as proxies for geomorphologic processes were selected namely the number of pools and riffles, channel bars, the vegetation structure and the amount of woody debris. Second, a set of anthropogenic factors relating to the presence of culverts, bridges, weirs, sluices and other modification of the river habitat.

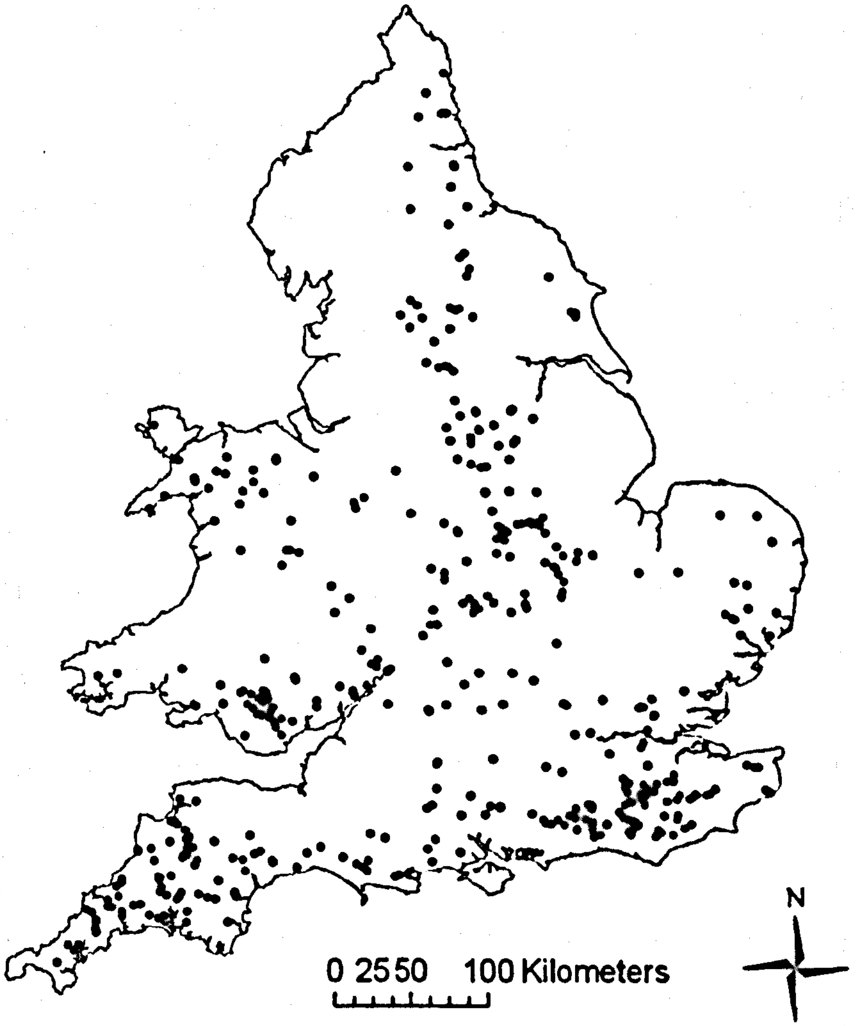


Figure 4.1: The distribution of the 415 sites identified as having matching GQA and RHS data based on a 1000 meter buffer.

For each site ArcGIS was used to calculate the upstream catchment based on the Ordnance Survey 50 metre panorama digital elevation model. Percentage land cover for these newly derived catchments was calculated based on a simplification of land cover classes from the Land Cover Map 2000 (LCM2000) into six categories representing; (i) grassland, moor and heath, (ii) urban and suburban areas, (iii) arable land, (iv) deciduous forest, (v) coniferous forest and (vi) unclassified, this latter category including inland and coastal water.

Gridded observation datasets of long-term climate data from 1990 to 2005 were obtained from the UK Met Office UK Climate Projections (UKCP09) programme. Using the method of Moss et al. (1987) mean annual air temperature was calculated as the mean of January, April, July and October means, with mean annual air temperature range calculated as the difference between the January and July means. Mean annual rainfall was calculated based on annual values from 1990 to 2005 with these figures being used to calculate the coefficient of variation (CV) of mean annual rainfall providing an indication of the predictability of this value. For each measure new 5 km x 5 km gridded data were produced and ArcGIS used to extract values based on the 415 GQA points.

For each of the 415 sites a total of 38 environmental variables were available (see Table 4.1). These were characterised into four spatial scales based on a similar classification by Feld and Hering (2007). Regional variables capture information about the sites geographic position in England and Wales and describe large scale characteristics such as underlying geology and climate. Catchment variables primarily provide a description of land use within the catchment. Reach scale variables characterise the general river habitat and indicate the extent that the river system has been influenced by human activity. Finally, site scale variables provide site specific descriptors of channel stability and substrate composition.

Table 4.1: Environmental variables used in the current analysis. In each case a brief description is provided. For variables derived from BIOSYS Murray-Bligh (1997) provides detailed method, and for those derived from the River Habitat Survey consult Fox (1998).

Factor	Derivation	Description
Regional		
Easting	BIOSYS	Geographic position based on British National Grid System.
Northing	BIOSYS	Geographic position based on British National Grid System.
Geological Drift	BGS Geo. Drift Map	Geological drift categories based on extraction to points in ArcGIS
Mean annual temperature	Met Office UKCP09	5 km x 5 km gridded data from 1990 to 2005 calculated from Met Office long-term monitoring programme.
Temperature range	Met Office UKCP09	5 km x 5 km gridded data from 1990 to 2005 calculated from Met Office long-term monitoring programme.
Mean annual rainfall	Met Office UKCP09	5 km x 5 km gridded data from 1990 to 2005 calculated from Met Office long-term monitoring programme.
CV Annual Temperature	Met Office UKCP09	5 km x 5 km gridded data from 1990 to 2005 calculated from Met Office long-term monitoring programme.
Catchment		
Upstream Catchment	OS 50m panorama digital elevation model	Calculated with ArcGIS based on OS 50m panorama digital elevation model.
2000 Arable L.C.	LCM2000	Includes classes (i) tilled land and (ii) scrub and orchard
2000 Broad Leaved L.C.	LCM2000	Includes class deciduous woodland.
2000 Coniferous L.C.	LCM2000	Includes class coniferous woodland.
2000Grassland L.C.	LCM2000	Includes (i) grass heath, (ii) mown turf, (iii) semi natural swards, (iv) rough/marsh grass, (v) moorland grass, (vi) open shrub moor, (vii) dense shrub moor, (viii) Bracken, (ix) Dense shrub heath, (x) open shrub heath, (xi) Upland bog, (xii) lowland bog, and (xiii) Ruderal weeds.
2000 Urban L.C.	LCM2000	Includes classes (i) urban, (ii) suburban and (iii) inland bare ground.
Stream Order	CEH river network	Strahler stream order.
Altitude	BIOSYS	Obtained from Ordnance Survey 1:50 000 scale map.
Distance from source	BIOSYS	Obtained from Ordnance Survey 1:50 000 scale map using an curvimeter or planimeter.
Reach		
Pools and Riffles	RHS	Number of pools and riffles in 500 meter stretch (RHS sweep up).
Bars	RHS	Number of bars and riffles in 500 meter stretch (RHS sweep up)
Channel Bars	RHS	Number of channel bar and riffles in 500 meter stretch (present in the spot checks)
Veg. total	RHS	A score index for vegetation structure joining left and right bank. Different score has been given at each category present in the RHS (bare, uniform, simple, complex)
Woody	RHS	Presence of woody debris (3=Extensive, 1=Present, 0=None)
Hard modification (U.S.)	RHS	UPSTREAM (average of RHS sites present in the upstream catchment) presence of re-sectioned banks, bank protections, embankments along the ten spot check sites
Number of culverts (U.S.)	RHS	UPSTREAM (average of RHS sites present in the upstream catchment) presence of culverted along the ten spot check sites
Soft modification (U.S.)	RHS	UPSTREAM (average of RHS sites present in the upstream catchment) presence of weirs, sluices, ford and bridge along the ten spot check sites
Hard modification (500 m)	RHS	LOCAL (the RHS sites analysed) presence of re-sectioned banks, bank protections, embankments along the ten spot check sites
Number of culverts (500 m)	RHS	LOCAL (the RHS sites analysed) presence of culverted along the ten spot check sites
Soft modification (500 m)	RHS	LOCAL (the RHS sites analysed) presence of weirs, sluices, ford and bridge along the ten spot check sites
Slope	BIOSYS	Measures as m per km from Ordnance Survey 1:50 000 scale maps.

Table 4.1 continued.

Factor	Derivation	Description
Code	RHS	Twelve categories of different drift geology are present in the RHS database. An ordinal indicator, based on sediment size, from fine to coarse, has been defined, looking at the composition of sediments in each geological category. This variable does not take into account the catchment geology.
Site		
% Boulders/Cobbles	BIOSYS	% cover of particles less than 0.06 mm on their longest axis. Estimated visually over whole sampling Area.
% Pebbles/Gravels	BIOSYS	% cover of particles between 0.06 mm and 2 mm on their longest axis. Estimated visually over whole sampling Area.
% Sand	BIOSYS	% cover of particles between 2 mm and 64 mm on their longest axis. Estimated visually over whole sampling Area.
% Silica/clay	BIOSYS	% cover of particles over 64 mm on their longest axis. Estimated visually over whole sampling Area.
CV Depth	BIOSYS	Stream depth is measured as predominant conditions at site. The coefficient of variation was calculated based on all available data from 1990 to 2005.
CV Width	BIOSYS	Stream width is measured at the water surface. The coefficient of variation was calculated based on all available data from 1990 to 2005.
Variance Mean Phi Units	BIOSYS	Standard deviation of mean Phi units for all samples between 1990 and 2005.
Discharge category	BIOSYS	Estimate of the mean annual discharge at a site based on existing data.
Change Chemical Grade	EA GQA programme	Extracted from Environment Agency database and calculated as the change in chemical GQA score from 1990 to 2005.

4.2.3: Community change

The measure of temporal variability used in the current study is similar to the measure of functional diversity proposed by Petchey and Gaston (2002). It represents a single summary value for a site that indicates total change in the identity and relative abundance of taxa present. For each of the 415 sites Kulczynski's distance (Faith et al., 1987) was used to calculate the similarity between communities through time and create a diagonal matrix of pairwise comparisons. This matrix is then used to construct a dendrogram using average linkage clustering. Each of the branches of the resulting dendrogram represents a sampling occasion at the site. As dissimilarity between samples increases so too does branch length. Total branch length therefore represents a measure of dissimilarity across all the samples at a site giving an aggregate measure of community change.

As branch length is sensitive to the number of samples used the first and last sample taken at each site were selected, then four other samples for the intervening years were selected at random. This approach is necessary as due to the design of the EA's monitoring programme few sites are sampled on an annual basis. Of the 3305 sites initially identified only 43 had contiguous yearly data available. To examine the effect of the randomisation approach community variability was calculated for these 43 sites using both the contiguous and randomly selected data. Pearson's correlation indicated a strong relationship between the two measures (Pearson's correlation, $r = 0.89$, $n = 43$, $p < 0.001$). This suggests that the random approach provides a strong indication of temporal variability in the community. The usefulness of the approach is that it increases the number of sites available from 43 (contiguous) to 415 (random) greatly increasing the spatial scale at which drivers of temporal variability can be examined.

4.2.4: Community change and Environmental Drivers

The relationship between temporal variability and environmental variables was explored using Random Forests (Breiman, 2001). This technique belongs to the family of machine learning techniques that use an algorithm to learn the relationship between predictors and response variables (Breiman, 2001). The approach differs from commonly used statistical techniques where predictor variables and their relationship to the response variable are described within the statistical model (Elith et al., 2008; Peters et al., 2007). Machine learning assumes that there is a complex and unknown relationship between the response and its predictors and, by observing the relationship between the two, the algorithm attempts to find patterns within the data (Prasad et al., 2006, Hochachka et al., 2007). The focus of such machine learning techniques is on predictive accuracy and including any predictor variable that is potentially informative (Hochachka et al., 2007).

Random Forests are based on regression trees, an approach to the analysis of ecological data that has become increasingly popular due to its robustness and ease of interpretation (De'ath and Fabricius, 2000). In regression trees, the data are repeatedly partitioned into two mutually exclusive groups that are as homogenous as possible. Where multiple predictors are used, at each node the homogeneity of groups based on splits produced by all predictors is examined and the explanatory variable that maximises homogeneity is chosen (De'ath and Fabricius, 2000; Prasad et al., 2006). Initially, the tree is grown to its maximum and then pruned back to an optimal size using a range of techniques such as cross-validation (Prasad et al., 2006; Crawley, 2007).

Random forests are, as the name suggests, composed of a number of regression trees, typically 500 to 2000. Each regression tree is grown using a bootstrap sample of two thirds of the entire dataset, with a random subset of predictor variables used at each node to generate the best split (Liaw and Wiener, 2002; Prasad et al., 2006; Peters et al., 2007). The use of a random subset of predictors in the construction of the trees has been demonstrated to increase the predictive performance compared with a single tree (Peters et al., 2007). The predictions of each tree in the Random Forest are aggregated using averaging to construct the final model (Liaw and Wiener, 2002). Accuracy of model predictions is assessed by using each regression tree model to predict the data not in the bootstrap sample (the remaining one third). This out-of-bag (OOB) data allows an estimation of the error rate for each regression tree, and once it is aggregated across all trees for the Random Forest (Liaw and Wiener, 2002). This estimate of the error is used to predict the percentage of variance explained by the random forest model using the equation

$$1 - \frac{MSE_{OOB}}{\hat{\sigma}_y^2}$$

Where MSE_{OOB} represents the mean of squared residuals from the OOB estimates (see Liaw and Wiener, 2002 for a detailed explanation).

Major criticisms of machine learning techniques such as Random Forests is that they are a "black box" and as such do not readily lend themselves to interpretation and scrutiny. The most widely used technique to aid interpretation is that of partial dependence plots that describe the effect of an individual predictor after accounting for the average effect of all other predictors (Hochachka et al., 2007).

For example to understand the influence of the % of urban catchment on temporal variability (i) replace all values for % urban catchment with the value of 1 whilst leaving all other values the same to create a synthetic data set; (ii) calculate the amount of temporal variability by passing this dataset through the data mining model constructed from the real data; (iii) the predicted value for temporal variability in the synthetic dataset is then the partial dependence value for 1% urban catchment; (iv) repeat this procedure for other values of % urbanisation to build up the partial dependence plot.

The 415 sites selected for this study were analysed using the "randomForest package" (Liaw and Wiener, 2002) in R 2.8.1 (R Development Core Team, 2008). Twelve randomly chosen variables were used in each tree and a total of 10,000 trees were used to construct the random forest. Variable importance was assessed based on the Mean Decrease Accuracy (%IncMSE) (Kuhn et al., 2008) where higher values of %IncMSE indicate higher variable importance.

4.2.5: Spatial Scale

For the environmental variables identified as being most important, correlograms were computed to examine spatial structure. Correlograms were calculated using Sturge's rule to determine the optimal number of distance classes (Legendre and Legendre, 1998). Moran I statistic with a Bonferroni correction was calculated for each distance class to determine whether there was significant spatial autocorrelation.

4.3: Results

Environmental variables explain 36% of the variance in temporal variability. Figure 4.2 illustrates the importance of each variable within the Random Forest model based on mean decrease in accuracy when the variable is omitted. A mixture of geographic position (northing, easting), large scale descriptors (temperature range, mean annual rainfall), land use characteristics and site scale factors (variance in Phi units and variance in channel width) emerge as important predictors. With the exception of variance in Phi units and channel width most explanatory power is attributed to catchment and regional scale variables.

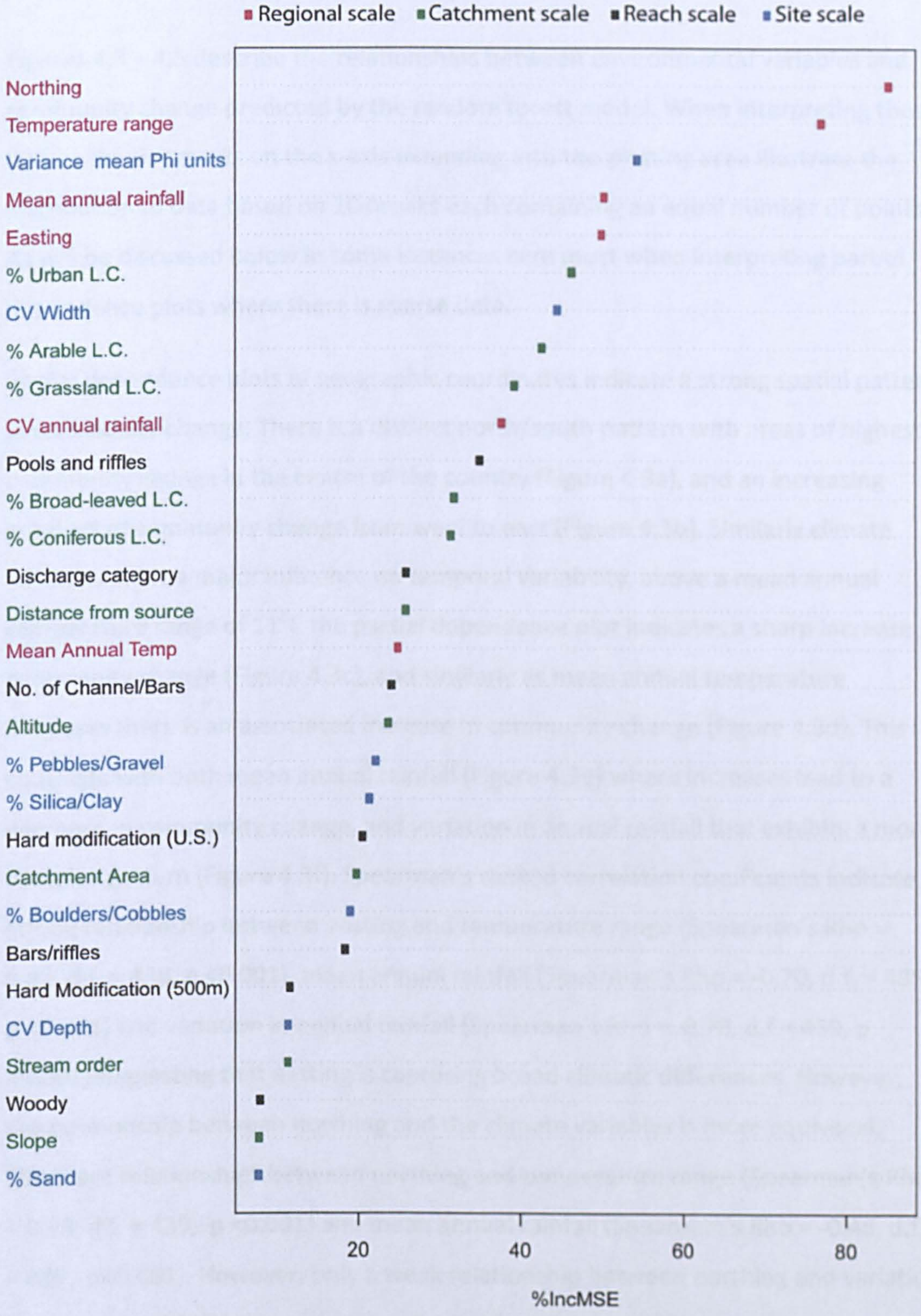


Figure 4.2: Relative importance of each predictor in the Random Forest model. Colours indicate characterisation of processes as regional, catchment, reach or site scale based on the classification of Feld and Hering (2007).

Figures 4.3 – 4.5 describe the relationships between environmental variables and community change predicted by the random forest model. When interpreting these figures the tick marks on the x-axis extending into the plotting area illustrate the distribution of data based on 10 breaks each containing an equal number of points. As will be discussed below in some instances care must when interpreting partial dependence plots where there is sparse data.

Partial dependence plots of geographic coordinates indicate a strong spatial pattern in community change. There is a distinct north/south pattern with areas of highest community change in the centre of the country (Figure 4.3a), and an increasing gradient of community change from west to east (Figure 4.3b). Similarly climate variables exert a major influence on temporal variability, above a mean annual temperature range of 11°C the partial dependence plot indicates a sharp increase in community change (Figure 4.3c), and similarly as mean annual temperature increases there is an associated increase in community change (Figure 4.3d). This contrasts with both mean annual rainfall (Figure 4.3e) where increases lead to a decrease in community change, and variation in annual rainfall that exhibits a more complex pattern (Figure 4.3f). Spearman's ranked correlation coefficients indicate strong relationship between easting and temperature range (Spearman's Rho = 0.81, d.f. = 439, $p < 0.001$), mean annual rainfall (Spearman's Rho = -0.70, d.f. = 439, $p < 0.001$) and variation in annual rainfall (Spearman's Rho = 0.78, d.f.= 439, $p < 0.001$) suggesting that easting is capturing broad climatic differences. However, the relationship between northing and the climate variables is more equivocal. There are relationships between northing and temperature range (Spearman's Rho = 0.38, d.f. = 439, $p < 0.001$) and mean annual rainfall (Spearman's Rho = -0.46, d.f. = 439, $p < 0.001$). However, only a weak relationship between northing and variation in annual rainfall (Spearman's Rho = 0.10, d.f. = 439, $p < 0.05$). These results suggest that whilst there is strong west/east climate gradient this relationship is less strong from north to south.

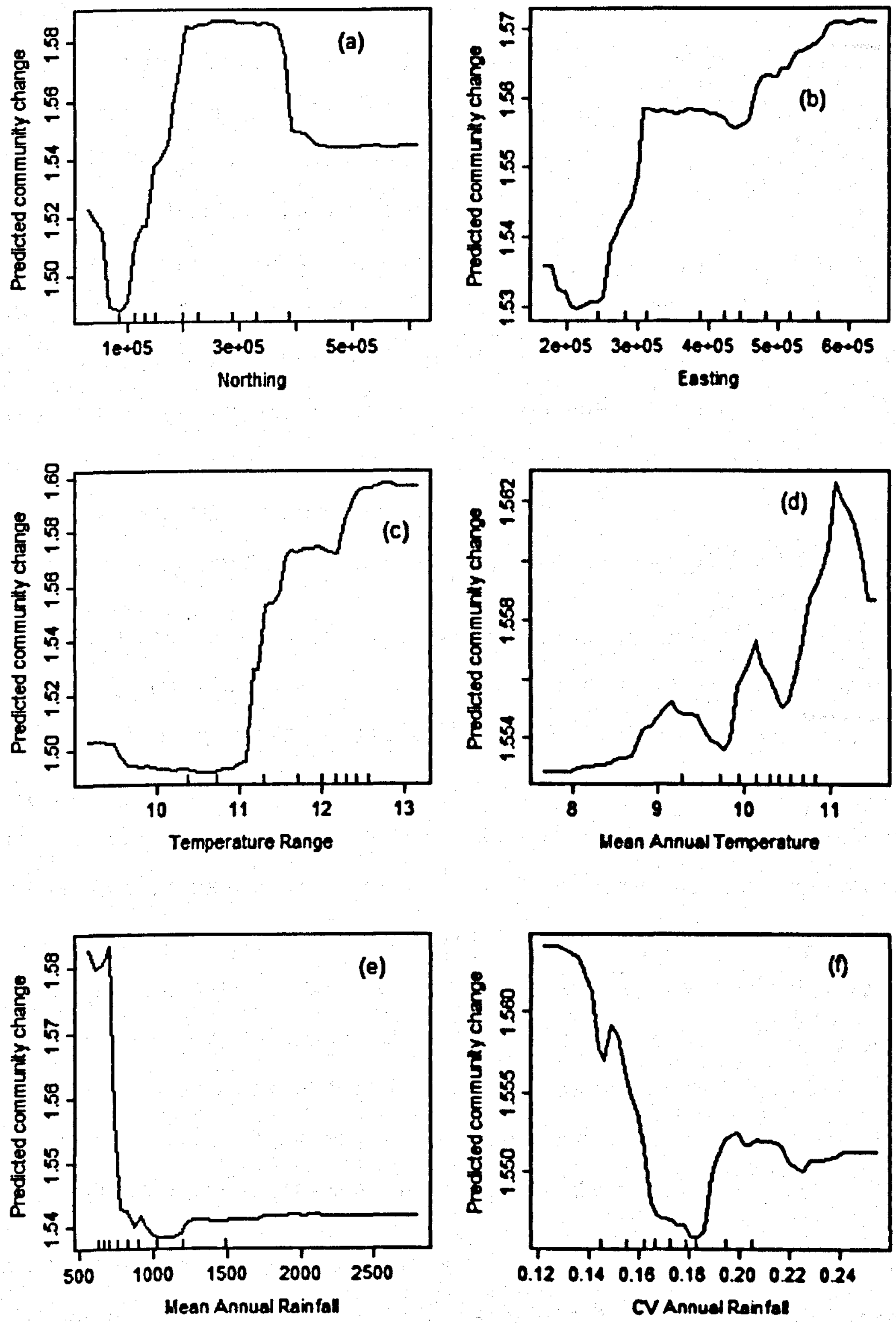


Figure 4.3: Partial dependence plot of regional scale environmental descriptors (a) Northing, (b) Easting, (c) annual temperature range °C, (d) mean annual temperature °C, (e) mean annual rainfall (mm) and (f) coefficient of variation of mean annual rainfall.

At the catchment scale (Figure 4.4a-e), land use has a major influence on temporal variability. Increases in land use associated with human activity (urbanisation/ arable crops/coniferous woodland) are associated with an increase in the temporal variability of communities. This is most marked for urbanisation and coniferous woodland (Figure 4.4a and 4.3e) where a relatively small percentage of the upstream catchment under these land uses (between 10% and 20%) results in sharp increase in community change. For arable land (Figure 4.4c) a greater proportion of the catchment may be utilised (*circa* 50%) before there is an increase in community change. These findings contrast with grassland land use where increasing cover in the catchment results in decreasing community change (Figure 4.4b). Results for percentage of mixed woodland present a more complex picture where initially there is an increase in community change then, as *circa* 30% of the catchment becomes mixed woodland, a decrease in community change (Figure 4.4d).

Correlations between land use categories reflect clear regional differences in land use across England and Wales, most strongly for arable land use and both mixed woodland (Spearman's Rho = -0.51, d.f. = 439, p <0.001) and coniferous woodland (Spearman's Rho = -0.63, d.f. = 439, p <0.001) reflecting the dominance of arable land use to the east of the country. Land use will be related to distance to source in a number of complex ways where, for example, coniferous woodland is likely to be found closer to the source as forestry is concentrated in many upland areas. The direct measure of this relationship (Figure 4.4f) indicates that there are decreasing levels of community change with increasing distance from source. Similarly, land use is strongly correlated with environmental factors reflecting congruence between these regional factors. Table 4.1 summarises correlations between the most significant predictors from the Random Forest model.

For both % Urban and % Coniferous land use it should be noted that there were comparatively few samples that included extremely high values. As such care must be taken in the interpretation of community change where land cover values exceed *circa* 30%. However, there can be a high degree of confidence in the relationship below these values as most data points are concentrated here.

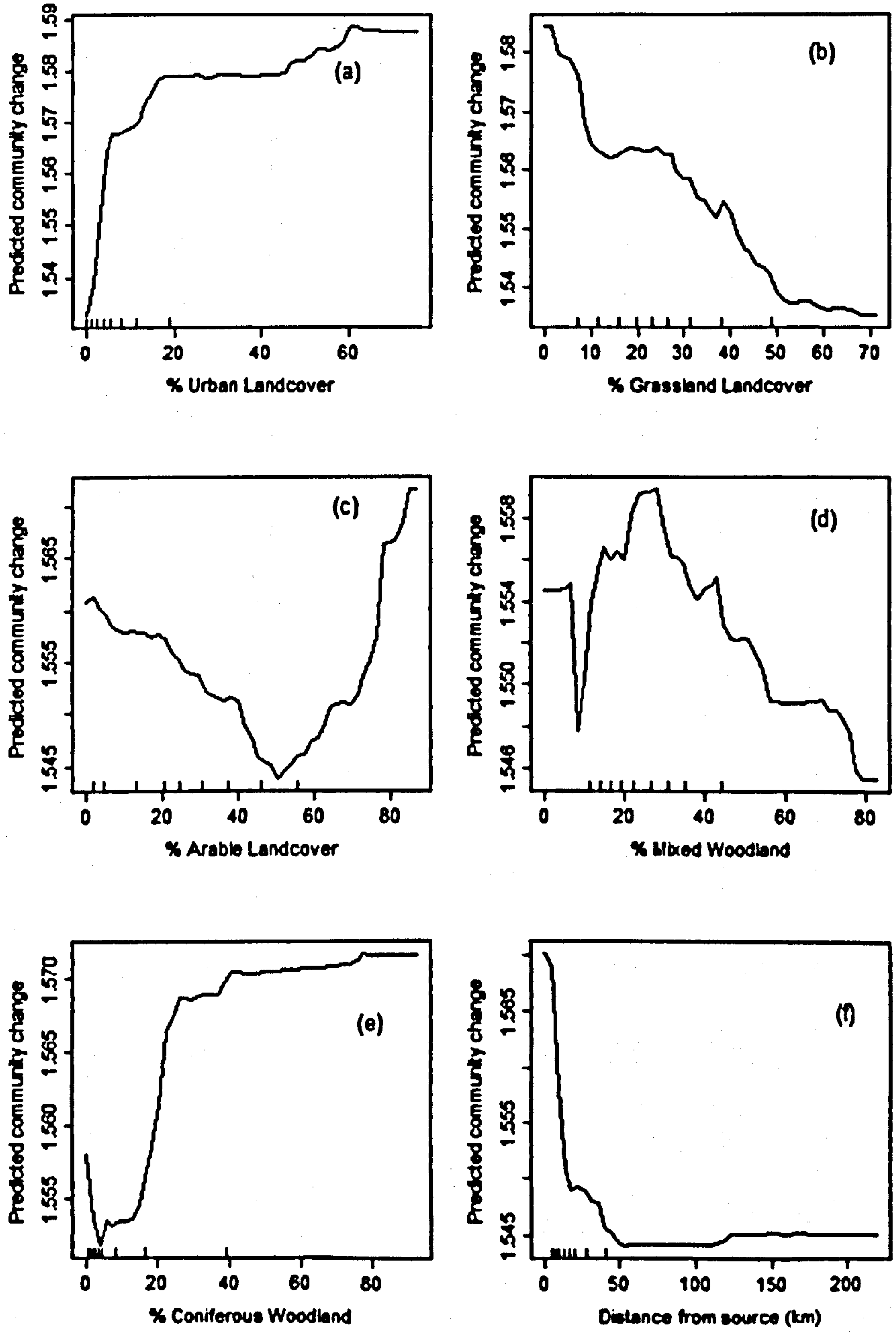


Figure 4.4: Partial dependence plots of catchment scale variables indicating the influence of (a) % urban land cover, (b) % grassland, (c) % arable, (d) % mixed woodland, (e) % coniferous woodland and (f) distance from source (km).

Table 4.2: Spearman's Rho correlations between environmental descriptors identified as being most important in the Random Forest model. Significance is indicated by coloured cells at ■ p < 0.05; ■ p <0.01 and ■ p <0.001.

	Easting	Northing	% Arable	% Broadleaved	% Coniferous	% Grassland	% Urban	Distance from source	Altitude	CV Channel Width	CV Mean Phi Units	Hard modification	Pools and Riffles	Temp. Range	Mean Annual Temp.	Mean Annual Rainfall
Northing	0.24															
% Arable	0.48	0.04														
% Broad Leaved	0.12	-0.09	-0.51													
% Coniferous	-0.28	0.19	-0.63	0.27												
% Grassland	-0.41	-0.41	-0.02	-0.11	-0.18											
% Urban	0.29	0.37	0.1	-0.03	-0.06	-0.39										
Distance from source	-0.14	0.21	-0.01	-0.07	0.18	-0.01	0.08									
Altitude	-0.26	0.14	-0.28	0.02	0.15	-0.10	0.04	-0.17								
CV Channel Width	0.19	0.26	0.12	-0.06	-0.02	-0.03	0.02	-0.11	0.06							
CV Mean Phi Units	0.49	0.12	0.29	0.03	-0.22	-0.03	0.18	-0.19	-0.22	0.17						
Hard modification	0.18	0.32	0.26	-0.17	-0.12	-0.3	0.51	0.21	-0.01	0	0.11					
Pools and Riffles	-0.29	-0.16	-0.27	0.13	0.07	0.14	-0.24	-0.2	0.24	-0.06	-0.25	-0.33				
Temp. Range	0.81	0.38	0.56	-0.05	-0.32	-0.44	0.54	0.01	-0.09	0.19	0.37	0.4	-0.31			
Mean Annual Temp.	0.16	-0.62	0.26	0.05	-0.33	0.27	-0.09	-0.05	-0.64	-0.16	0.14	-0.09	-0.09	0.10		
Mean Annual Rainfall	-0.70	-0.46	-0.68	0.19	0.43	0.22	-0.44	-0.08	0.27	-0.24	-0.41	-0.36	0.31	-0.83	-0.12	
CV Annual Rainfall	0.78	0.1	0.35	0.16	-0.25	-0.26	0.18	-0.17	-0.12	0.14	0.36	0.10	-0.13	0.67	0.17	-0.54

Finally, at the site scale variation in substrate (Figure 4.5a) and channel width (Figure 4.5b) lead to increasing community change. The reach scale provides the only direct measure of flow regime (discharge category) with Figure 4.5c illustrating that low discharge results in the highest levels of community variability. As the number of pools and riffles within the 500 m section of the river increases there is a decrease in community change (Figure 4.5d). As with % urban and coniferous land cover, care must be taken in interpreting the relationship at higher values for variance in substrate, channel width and the number of pools and riffles, as higher values are rare.

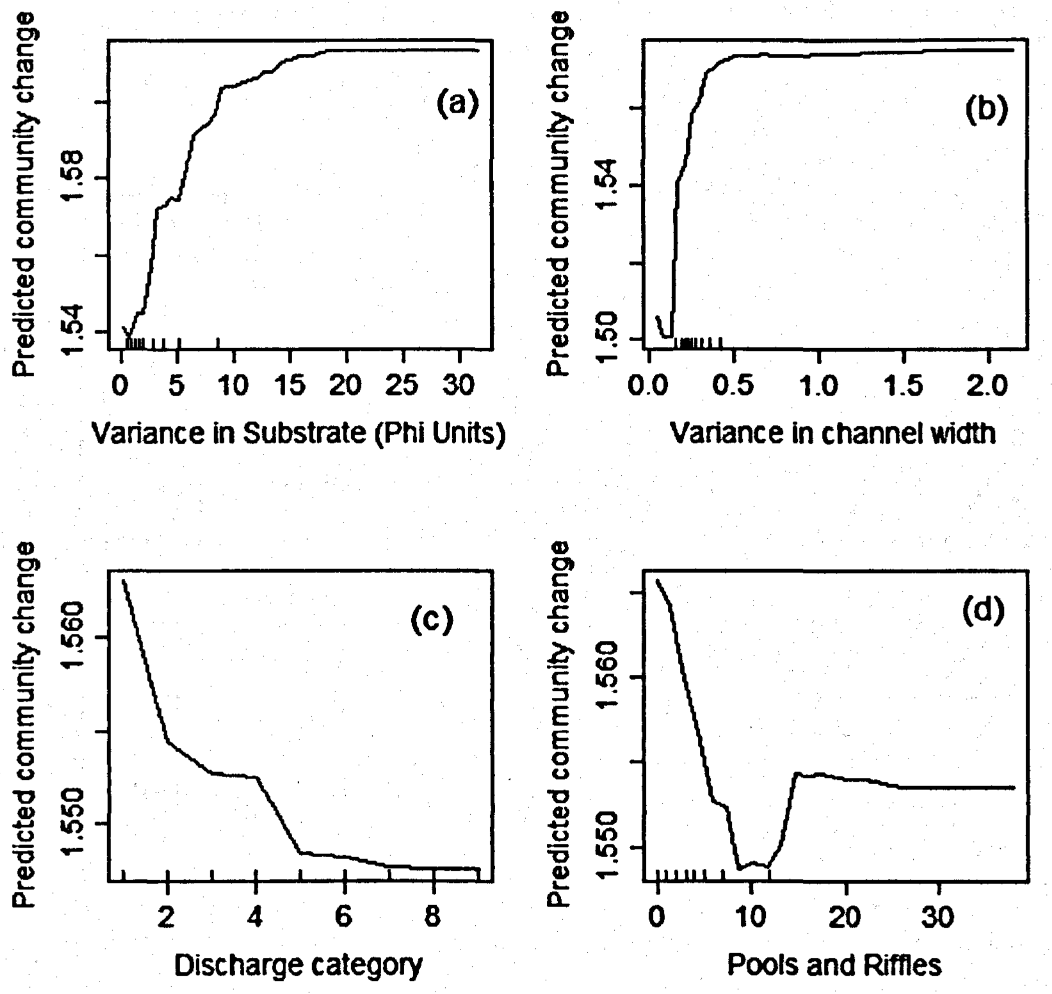


Figure 4.5: Frequency of occurrence of in site scale variables indicating (a) variation in substrate composition 1990-2005, (b) variation in channel width 1990-2005, (c) discharge category and (d) number of pools and riffles.

Moran's I statistic was used to examine the extent to which sites that are close together have similar environmental conditions. Figure 4.6 illustrates correlograms for predictor variables identified as being most important for the prediction of community temporal variability. In most instances variables exhibited spatial autocorrelation at scales of up to 150 km. Mixed woodland, pools and riffles, and variation in channel width exhibited similarities at smaller scales (circa 50km; Figure 4.6h-j). Community change itself shows spatial autocorrelation at scales of up to 200 km (Figure 4.6i) indicating similar levels of community temporal variability at spatial scales corresponding to environmental drivers.

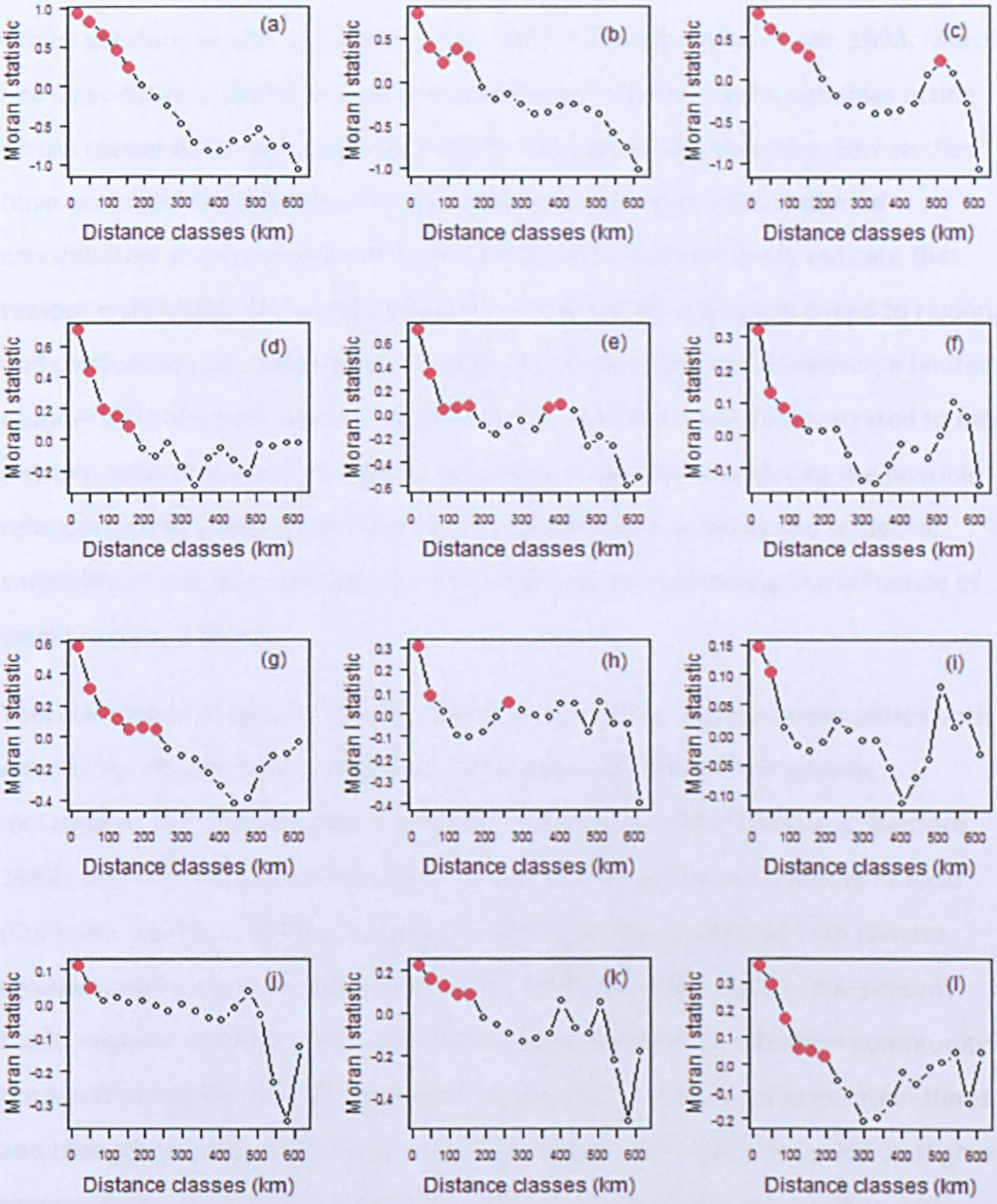


Figure 4.6: Correlograms of (a) temperature range, (b) annual rainfall, (c) variation in rainfall, (d) % arable, (e) % conifer, (f) % urban, (g) % grassland, (h) % mixed woodland, (i) pools and riffles, (j) variation in width and (k) variation in Phi units. Finally (l) represents correlogram of community temporal variability. Red points indicate significant spatial autocorrelation based on Moran I statistic after Bonferroni correction.

4.4: Discussion

Whilst previous studies (e.g. Moss et al., 1987; Richards and Minshall, 1993; Murphy and Davy-Bowker, 2005) have demonstrated that environmental variables acting across spatial scales are accurate predictors of community structure, few studies have examined the link between such variables and temporal variation of communities at comparable scales. Results from the current study indicate that temporal variability of macroinvertebrate communities is strongly linked to regional and catchment scale variables relating to climate and land use. However, a limited number of site specific factors, specifically bed substrate, are demonstrated to have a strong relationship with temporal variability. I begin by considering the possible relationships between large scale climate and land use patterns and temporal variability of macroinvertebrate communities before considering the influence of smaller scale variables.

Thermal regime exerts a profound effect on macroinvertebrate communities both directly by influencing the physiology of organism affecting their growth, metabolism and reproduction (Vannote and Sweeney, 1980; Ward and Stanford, 1982), and indirectly by altering factors such as the quality and quantity of food (Cummins and Klug, 1979). Changes in thermal regime associated with climate change (Daufresne et al., 2004; Chessman, 2009) or, at the scale of the present study, regional climatic gradients has been demonstrated to influence community temporal variability through processes such as the shortening of generation times and changes in interspecific interactions (Burgmer et al., 2007; Berg, 2010). Higher temperatures can have a de-stabilising action on community composition with increased turnover of taxa and replacement by other members of the local species pool (Hillebrand et al., 2009). In the current study air temperature range was demonstrated to hold the most predictive power of all the environmental variables considered. Air temperature range is based on the difference between average temperatures in January and July, and as such represents seasonal variation in the habitat. Such predictable change is not usually considered to represent a stressor to macroinvertebrate communities as taxa adjust their life cycles in anticipation of the changing conditions (Wolda, 1988; Reece et al., 2001; Bonada et al., 2007). Seasonal variation has even been demonstrated to enhance diversity through the temporal separation of taxa within the annual cycle (Ward and Stanford, 1992). However, Townsend et al. (1987) found a similar relationship between annual water temperature range and community temporal variability, that the authors attributed to temperature range being indicative of instability in the environment. Wide variation may result in conditions exceeding the tolerance limits of taxa leading to local extinction. Similarly, taxa may rely on a dormant phase or dispersal mechanisms that removes them from unfavourable conditions but leads to increasing community temporal variability due to the stochastic nature of re-colonisation (Wolda, 1988; Beche and Resh, 2007).

Partial dependence plots indicated a positive relationship between mean annual temperature and community change. However, of the regional and catchment scale environmental variables, mean annual temperature held the least predictive power after catchment area. The strong correlation between altitude and mean annual temperature suggests that the negative relationship with community change may be driven by higher annual temperatures in lowland areas where rivers are more prone to anthropogenic disturbance (Murphy and Davy-Bowker, 2005; although see Ormerod and Durance, 2009). In this case land use, which also has a distinct upland to lowland structure, clearly represents a more powerful predictor.

As well as temperature, climate variables relating to rainfall were also considered in the current study. These rainfall variables were employed primarily as indicators of broad regional climatic patterns. However, rainfall exerts a considerable influence on macroinvertebrate communities primarily through its association with flow regime (Poff et al. 1997b; Konrad et al., 2008). Ecologically meaningful flow elements may be magnitude, frequency above magnitude, duration, timing or predictability and flashiness or rate of change (Clausen and Biggs, 2000; Poff et al., 1997b). Although there is a clear link between flow regime and precipitation (e.g. Jones and Lister, 1998; Trigo et al., 2004) defining the specific relationship is complex with the timing, duration and intensity of rainfall together with catchment characteristics influencing flow at a specific site (Poff et al., 1997). Gibbins et al. (2001) suggests that there might be the need for extremely detailed hydrological data to pick up subtle changes in conditions that will affect macroinvertebrate communities. Such gauged data is unavailable for many of the sites considered in the current study. Therefore the two measures of rainfall employed here provide an indication of areas with water stress (mean annual rainfall) and areas where there is likely to be the most variation in flow regime between years (CV annual rainfall).

Mean annual rainfall emerged as the fourth most important predictor for temporal variability of macroinvertebrate communities in the current study. Figure 4.3e indicates a sharp decrease in temporal variability of communities above a mean annual rainfall of *circa* 750 mm per year. Areas with low mean annual rainfall may be subjected to periods of drought that have been shown to significantly impact communities through a number of processes that include alteration of the habitat characteristics (Beche and Resh, 2007; Bonada et al., 2007; Bêche et al., 2009), concentration of taxa in pools leading to increased biotic interaction (Boulton et al., 1992; Brown, 2007) and an increase in the importance of other stochastic community processes (Beche and Resh, 2007; Chase, 2007).

Although of less importance, the other measure of rainfall used in the current study, namely variation in annual rainfall, exhibited a more complex sinusoidal relationship with high levels of temporal variability associated with both low and high values of variation in rainfall. Beche and Resh (2007) demonstrated that high levels of variation in precipitation influenced temporal variation of communities by affecting stream discharge and habitat quality and quantity between years. Conversely, low levels of variation in rainfall may be indicative of relatively benign flow regime where the component taxa within communities may lack adaptations that provide resistance or resilience to unusual events such as floods (e.g. Armitage, 2005). Lepori and Malmqvist (2009) in a study examining the relationship between community structure and flow regime, considered that communities that were the least and most stressed where governed primarily by stochastic processes whereas those at intermediate levels of disturbance were governed by deterministic processes. The implications of Lepori and Malmqvist's (2009) study are that the sinusoidal response noted in the relationship between variation in annual rainfall and community temporal variability could arise due to the importance of stochastic processes under high and low levels of environmental variation. Collier (2007) noted such a similar sinusoidal response when examining the impact of increasing land use stress on aquatic communities. Collier (2007) considered that at intermediate levels of disturbance there is a loss of fragile species and replacement by tolerant taxa resulting in increases persistence, whereas at the most disturbed sites species are predominantly vagile relying on fast colonisation and resulting in considerable community variation through time.

Whilst regional processes relating to temperature and rainfall were demonstrated to hold the most predictive power, descriptors relating to land use also emerged as important determinants of temporal variability. Catchment land use reflects the degree of human activity (Feld, 2004) with results from the current study indicating increased levels of temporal variability as urban, arable and coniferous land increases, and trends towards decreasing community temporal variability under grassland and mixed woodland. The influence of arable and urban land use is well documented within the literature where they have been demonstrated to exert a considerable influence on the catchment through alterations to water chemistry (Paul and Meyer, 2001; Chadwick et al., 2006; Giller and O'Halloran, 2004), riparian vegetation (Richards and Minshall, 1992; Chadwick et al., 2006; Urban et al., 2006), thermal regimes (Paul and Meyer, 2001), the magnitude, timing and duration of flow events (Konrad et al., 2008), rates of channel erosion (Richards and Minshall, 1992; Roy et al., 2003; Townsend et al., 2004) and bed substrate (Davies et al., 2000; Paul and Meyer, 2001; Roy et al., 2003; Heatherly et al., 2007). As such urbanisation and agriculture are considered to be two of the leading factors impacting freshwater aquatic systems worldwide (Paul and Meyer, 2001). Similarly coniferous forests have been demonstrated to exert considerable influence on catchments through alteration of the hydrology, sedimentation, habitat and energy inputs and in particular the exacerbation of acidification processes under certain conditions (Clenaghan et al., 1998).

Urban et al. (2006) demonstrated that the effect of land use may be to alter both the severity of the disturbance and to restrict dispersal. Many long-term studies (e.g. McElravy et al., 1989; Boulton et al., 1992; Bradt et al., 1999) have noted that unusual or intense disturbance is often needed to alter community composition. As such the pattern of increasing community variation under increasing land use pressure in the present study may in part be explained by an increase in the frequency of such intense events. This is likely to be coupled with increasing homogenisation of habitats associated with anthropogenic activities such as urbanisation (Brown, 2003). Such homogenisation removes refuges for existing species that act as sources of taxa for recolonisation following the disturbance leading to a uniform impact on the community (Lancaster, 2000; Brown, 2007; Fraterrigo and Rusak, 2008). The loss of lateral connectivity in urban environments is considered to lead to increased temporal variability in this way as there are no longer sources of colonists following disturbance (Usseglio-Polatera and Beisel, 2002). Similarly, channel modification has been demonstrated to lead to fragmentation of instream communities by barriers that may also lead to increased variability through the isolation of populations (Urban et al., 2006). Metapopulation and landscape ecology highlights the importance of interlinked but separate populations for maintaining biodiversity (Lancaster, 2000) so where communities rely on such processes for maintenance of the community through time isolation through the loss of longitudinal or lateral connectivity will result in increased temporal variability of the community.

Results from the current study indicate community change exhibits a non-linear response to the percentage of arable, urban and coniferous land within the catchment. Above 10 to 20 percent urban and coniferous and 50 percent arable there is a sharp increase in community change. These results suggest a "tipping point" above which there is a sudden change in community properties and has been noted in a number of systems worldwide (Scheffer et al., 2001). For aquatic macroinvertebrates a number of studies have detailed such dramatic shifts in community structure (e.g. Paul and Meyer, 2001; Roy et al., 2003; Collier, 2007) with differences in this tipping point suggested to be related to the intensity of the impact. For example Paul and Meyer (2001) demonstrated that impervious surfaces covering more than 10 percent of the catchment resulted in floods that peaked more rapidly and were shorter in duration leading to a greater impact on communities. Doak et al. (1999) considered that the intensity of agricultural production, and not the percentage of the catchment under this land use, may exert the most influence on the community. In some instance the authors report that up to 80% of land use can be low intensity agricultural with little affect on the stream community.

The majority of reach and site scale variables in the current study emerge as being of less importance than broader scale factors such as climate and land use. This finding agrees with previous studies that have suggested that broad scale variables can exert an overriding influence on macroinvertebrate communities (Davies et al., 2000). Exceptions to this are variation in channel width through time and variation in substrate composition, with increasing variation leading to increasing community change in both cases. Both variables are clearly indicative of habitat stability that, as with temperature range, may lead to community variability. Similar to the tipping points in land use there appears to be a sharp increase in community change with increasing variation of both these measures. This again suggests that there are critical thresholds for the communities however, care needs to be taking in interpreting this result as there are fewer sites that exhibit particularly high levels of variation allowing less confidence in the relationship. However, the most important site scale variable, variation in substrate, has been demonstrated to be an important determinant of community variability in a number of other studies (e.g. Gibbins et al., 2001; Roy et al., 2003) with habitat degradation often leading to a reduction in substrate size and stability (Heatherly et al., 2007). Substrate conditions influence community structure through processes such as the provision of attachment sites, refuges against predators or flooding and the provision of periphyton food (Poff and Ward, 1990; Roy et al., 2003; Brown, 2007). As such where conditions exhibit the most variation it would be expected that there will be a similar high level of variation in community composition in response. The increased levels of variability predicted in the current study may therefore result from increased mortality as conditions become unsuitable, or as demonstrated by Paul and Meyer (2001) can results from an increased propensity for invertebrates to enter the drift as sedimentation increases.

The importance of reach and site scale variables that incorporate a measure of temporal variability highlights a potential weakness in the current analysis where for certain variables, most notably land use, our data lacks a temporal element. Land use change has been shown to be one of the principal driving forces behind ecosystem variability throughout the world (Sponseller et al., 2001; Collier, 2007) where much variation is characterised by loss of agricultural or forested land and replacement with urban (Roy et al., 2003). The process of conversion itself causes considerable disruption to the aquatic system and leads to marked changes in the community (Booth, 1997). Land use change over the course of the study could therefore represent an important source of community temporal variability. In river systems in England and Wales it could be suggested that there will not have been a significant change in land use between 1990 and 2005 at all but a few sites. However, land use change in the longer term may be exerting a more subtle effect on community variability. The ability of a community to resist and recovery from disturbance is a function of the biota's experience of the environment (Poff and Ward, 1990) with the affects of stressors likely to be incorporated into the community over relatively long time scales (Richards and Minshall, 1992).   Previous studies have shown a considerable time lag between changes in the physicochemical environment and the assemblage. For example land use representing conditions up to fifty years ago have been demonstrate to be an accurate predictor of community composition in the present day (Harding et al.,1998). As such there is the potential that historic land use patterns might provide a stronger predictor of community temporal variability however, this relationship needs further examination (Gergel, 2002).

Geographic position, most importantly northing, was demonstrated to be a strong indicator of community temporal variation with partial dependence plots indicating an increasing gradient from west to east, and a humped relationship from south to north. Geographic position captures large scale patterns in temporal variability of macroinvertebrate communities. In the current analysis there was evidence of significant spatial autocorrelation between factors at scales up to 150 km. This captures strong regional characteristics between areas of England and Wales that have been shown to be important predictors of macroinvertebrate community structure (Wright, 2000; Murphy and Davy-Bowker, 2005). As such the inclusion of geographic coordinates in the current study might have been expected to increase the predictive accuracy of the model as such spatial descriptors can act as surrogates for unmeasured variables (Murphy and Davy-Bowker, 2005). Murphy and Davy-Bowker (2005) demonstrated that a predictive model of macroinvertebrate community composition across a similar area was 76% as accurate using just spatial position as opposed to a range of environmental factors. Results from the current analysis indicated that the Random Forest model was only able to explain 36% of the variation even with spatial descriptors included. This suggests that other complex processes that are not geographically structured are important for controlling community change. Often the assumption is that in river systems a single point is representative of points upstream, this may not be the case (Townsend et al., 2004). Local environmental factors exert considerable influence on communities and have been demonstrated to vary not only at the site and reach scale but also within riffles and microhabitats (Boyero 2003).

In conclusion this study has demonstrated that a range of environmental factors acting across multiple scales influence temporal variability of macroinvertebrate communities. This suggests that large scales variables, climate and land use, offer a practical route through which levels of expected temporal variability within macroinvertebrate communities might be assessed. Results suggest that the inclusion of a limited set of reach and site scale variables, most notably variation in substrate, would increase the predictive power of any models. Such characterisation is an important requirement of any programme that uses biological elements to assess environmental stress (Irvine, 2005). However, the study has also demonstrated that much variation in community structure through time remains unaccounted for by the variables typically considered in the assessment of river health.

CHAPTER 5: RELATIONSHIPS BETWEEN TAXON TRAITS, COMMUNITY STRUCTURE AND TEMPORAL VARIABILITY IN LOTIC MACROINVERTEBRATE COMMUNITIES

5.1: Introduction

How and why communities change through time is a fundamental question in ecology that has received considerable attention in recent years due to increasing concerns about the loss of biodiversity and associated ecosystem function (Yachi and Loreau, 1999; Vitousek et al., 1997; Loreau, 2000). Increasing our understanding of why communities change has broad implications for monitoring, management and conservation of ecosystems in landscapes that are being increasingly dominated by human activity (Vitousek et al., 1997; Palmer et al., 2004). As temporal variation in communities can arise through processes that are both natural and anthropogenic it is essential to understand the relative importance of each if we are to make informed management decisions (Niemi and McDonald, 2004).

Community structure arises through the influence of abiotic and biotic processes acting across multiple scales (Minshall, 1988; Poff, 1997; Belyea and Lancaster, 1999). These processes represent a Habitat Templet (Southwood, 1977; Southwood, 1988) that for a given set of conditions select for traits of taxa that represent the optimum investment in strategies for defence, migration, reproduction, longevity and tolerance to adverse conditions (Townsend and Hildrew, 1994; Korfiatis and Stamou, 1999). In lotic systems, Townsend and Hildrew (1994) proposed a River Habitat Templet where selection for these traits is based on the temporal and spatial heterogeneity of the environment. In Townsend and Hildrew's (1994) definition, temporal heterogeneity refers to the frequency and magnitude of variation in conditions from their long-term average, whereas spatial heterogeneity refers to the provision of refuges as these ameliorate or modify the influence of such disturbance. For highly disturbed sites, selection based on these criteria would result in a community where taxa possess resistance mechanisms (such as streamlining or firm attachment) with which to remain at a site (Gasith and Resh, 1999) or resilience mechanisms (such as flight) allowing rapid recolonisation following disturbance (Usseglio-Polatera et al., 2000; Vieira et al., 2004). This contrasts with more stable environmental conditions where selection may favour taxa with stronger competitive abilities (Townsend et al., 1987; Bonada et al., 2007).

The question that arises from this understanding of how communities are shaped is whether some communities, owing to their constituent taxa and the traits that they possess, exhibit less temporal variability than others shaped by a differing combination of factors. Commonly temporal variability of communities is considered to be most strongly influenced by unpredictable events as they represent stresses to which the community is not adapted (Meffe and Minckley, 1987; Boulton et al., 1992; Bradt et al., 1999; Wagner and Schmidt, 2004; Beche et al., 2006). If typical conditions are relatively benign taxa may lack mechanisms to cope with even moderate disturbance, whereas if communities are subjected to harsh conditions taxa would be expected to possess resistance or resilience mechanisms meaning that severe disturbance, which would be unusual, would be needed to cause temporal change. For example Armitage (2006) demonstrated that communities downstream of a reservoir were highly fragile and susceptible to disturbance compared with a parallel stream not influenced by the reservoir. Armitage (2006) attributes this to the community downstream of the reservoir being exposed to a narrower range of environmental conditions and therefore lacking mechanisms to respond to large scale disturbance such as rainfall events. Communities exposed to harsh conditions have been demonstrated to exhibit little variation through time (e.g. Scarsbrook, 2002; Milner et al., 2005) as the specific traits required to ensure success at such sites result in a limited community of highly specialised taxa (Beche et al., 2006; Ilg and Castella, 2006). However, in extreme conditions the response mechanism required by taxa may in itself lead to increased temporal variability within the community. Humphrey et al. (2000) and Beche et al. (2006) demonstrate that in seasonal Mediterranean type streams there is a distinct wet and dry season fauna that rely on recolonisation as their principal strategy. The reliance on an inherently stochastic process leads to low persistence through time in terms of taxonomic identity (Townsend et al., 1987; Humphrey et al., 2000) indicating a clear link between the strategies of taxa, community structure and temporal variability. Such findings suggest a mechanism through which community structure, by reflecting the range of response mechanisms available for the taxa, may be related to temporal variability.

Differences in temporal variability between communities have important implications for biomonitoring programmes where change in community structure is assumed to arise due to anthropogenic stress (Richards et al., 1992; Statzner et al., 1997; Scarsbrook, 2002; Woodward et al., 2002; Milner et al., 2005). To have confidence in conclusions from such programmes it is important to understand whether some communities, due to their constituent taxa and the range of traits that they possess, are likely to exhibit higher levels of temporal variation than others. The ability to address this question at temporal and spatial scales sufficient to capture long-term change in a variety of communities has been limited until fairly recently by a lack of such data (McElravy et al., 1989; Jackson and Fureder, 2006). Here data from a national water quality monitoring programme are used to; (a) examine the contribution that individual taxa make to temporal variability; (b) determine the role that differing traits of taxa play in the persistence of taxa within the community; (c) examine the relationship between community composition and temporal variability; (d) examine whether temporal variability is related to changes in the Habitat Templet or arises due to the strategies of taxa within communities.

5.2: Methods

5.2.1: Data

The Environment Agency conducts a rolling programme of monitoring designed to assess water quality based on macroinvertebrate community structure (Wright, 2000). This programme is based on a standard method (Murray-Bligh et al., 1997) designed to allow comparisons of water quality across England and Wales. Taxa are sampled using a three minute kick sample and one minute manual search. Samples are preserved and macroinvertebrates identified in the laboratory, usually to the family level, and their identity and log abundance recorded. Following a quality assurance procedure data are stored on a central database called BIOSYS that contains over a quarter of a million samples from 62765 sites collected over the last decades.

As the monitoring strategy employed by the EA is based on three year rolling programme few sites were available with contiguous yearly sampling data. Instead site selection was based on the availability of at least 6 samples taken between 1990 and 2005, where the first sample was taken in either 1990 or 1991 and the final sample in 2004 or 2005. As macroinvertebrate community composition changes seasonally (Rosillon, 1985) each sampling season was examined separately. The choice of spring sampling provided the widest spatial coverage and resulted in 1574 sites being selected for further analysis.

Taxonomic resolution was standardised to the family level for each sample as information at the species level was not consistently available and mismatches in the level of taxonomic identification between years would influence measurement of community change (Bradley and Ormerod, 2001; Metzeling et al., 2002). Due to taxonomic changes over time names of families were standardised using the Coded Checklist of Animals Occurring in Fresh Water in the British Isles (Furze, 2007).

5.2.2: Characterising community temporal variability

The measure of community change used in the current study is based on the total branch length of a dendrogram that represents the similarity between samples taken at a site over time. It is similar to the measure of functional diversity proposed by Petchey and Gaston (2002). For each site, based on the taxa present and their log abundance, a dissimilarity matrix was calculated using Kulczynski distance (Faith et al., 1987). Agglomerative hierarchical clustering was then used on the resultant dissimilarity matrix to construct a dendrogram where each branch represents a sample taken at the site. As the dissimilarity between community compositions in samples increases, so does the total branch length of the dendrogram providing an aggregate measure of community change.

As branch length will clearly be sensitive to the number of samples this was standardised to six samples for each site. Where more samples were available the first and last sample were retained (1990 or 1991 and 2004 or 2005 respectively) and four other years were randomly selected. Of the 1574 sites used in the analysis 43 have contiguous data from 1990 to 1999. To examine the affect that using random as opposed to contiguous data has on the assessment of temporal variability for each of these sites community change was calculated using both sets of data. Pearson's correlation indicated a strong positive relationship between calculated levels of community change based on random and contiguous data ($r = 0.89$, $n = 43$, $p < 0.001$).

5.2.3: Persistence of individual taxa along a gradient of temporal variability

Pooled data from the 1574 sites used in the study represents a gradient of temporal variability. For each taxon, logistic regression was used to analyse the relationship between the proportion of samples through time (at each site) in which the taxon occurred and the magnitude of temporal variability at the site. Depending on the form of the relationship and its statistical significance, taxa were classified into one of three groups as shown in Figure 5.1 (page 108). Taxa were classified as being "High Frequency" taxa if they occurred in a high proportion of samples along the entire gradient of temporal variation (Figure 5.1a) where a high proportion was characterised as being in excess of 50% of samples at the site through time. This contrasts with a second group of taxa that were present in a consistently low proportion of samples through time (Figure 5.1b) that were classified as "Low Frequency" taxa. These Low Frequency taxa occurred in less than 50% of sample through time at sites along the gradient of temporal variability. Finally, "Variable Frequency" taxa were considered to be those where the proportion of samples in which species occurred changed significantly along the gradient of community temporal variability (Figure 5.1c).

5.2.4: Trait affinities and temporal variability

To examine the role that traits of taxa have in influencing persistence each taxon was assigned a series of trait descriptors based on a subset of data from a European database of autecological information (Usseglio-Polatera et al., 2000). As detailed by Usseglio-Polatera et al. (2000) trait descriptors relate to the life cycle of taxa (size, aquatic stages, life cycle duration, potential number of generations per year), traits that convey resilience or resistance (dispersal, resistance stages) and general physiological and behavioural mechanisms (respiration, reproduction). In addition saprobity was included as an indication of the likely sensitivity of the taxa to pollution. Within the database traits are described as "modalities" based on a fuzzy coding procedure (Chevenet et al., 1994) that avoids the need to assign traits to a single category. Modalities represent different possible traits for a particular feature of the taxa, for example respiration may be through a gill, plastron, spiracle etc.. Scores are presented as affinities for a particular modality with 0 indicating no affinity and 5 indicating high affinity.

Information within the database is at the species or genus levels whereas the data used in the current study has been standardised to the family level. For each family an average of the affinities was taken as providing an indication of the typical trait affinities for each taxa. Resh et al. (1994) suggests that the use of family level data in this way is still likely to allow differentiation between trait affinities as grouping taxa into higher taxonomic orders is based on similarities in characteristics.

Having assigned trait affinities to each taxon Linear Discriminant Analsysis was used to examine whether trait affinities provide an accurate predictor of the classification of taxa into High, Variable or Low Frequency classes. LDA uses *a priori* knowledge of groupings to identify which descriptors best assign individuals to groups, and through cross validation using a subset of the original data provides a measure of this classification accuracy (Quinn and Keough, 2002). Strong classification accuracy would suggest distinct differences between the three groups. Traits that are most important for the classification are assessed based on the loading of the LDA axis. Following LDA Kruskal-Wallis tests (followed by multiple comparison tests) were used to examine each trait separately to identify significant difference and similarities.

5.2.5: Community composition and temporal variability

To examine the relationship between community structure and temporal variability cluster analysis was used to group communities based on similarity in composition. As both the method of clustering and the number of end groups has the potential to greatly influence the grouping, three differing clustering techniques with five differing sized end groups were calculated and differences assessed.

For each of the 1574 sites the Kulczynski distance was used to calculate a dissimilarity matrix based on the community structure in the first sample. Divisive hierarchical clustering and agglomerative clustering using Ward and Complete linkage methods (Quinn and Keough, 2002) were used to group sites based on the dissimilarity matrix. In the resulting dendrogram branches were "cut" at different similarity levels to create either 8, 10, 12, 15 or 20 end groups. To simplify interpretation these communities were then assigned numbers reflecting rank order of average community change. Formal testing of significant differences in average community change between groups was conducted using a non-parametric test (Kruskal-Wallis, followed by multiple comparison test) as the data violated assumptions for the parametric test (ANOVA).

5.2.6: Environment or trait driven temporal variability

The link between traits of taxa and environmental variables described by the Habitat Templet suggests that examining changes in community trait composition through time could be informative about factors influencing community temporal variability. Communities may exhibit high levels of taxonomic change through time, with little variation in traits of taxa, indicating relatively consistent environmental conditions where change is driven by biological processes such as extinction or colonisation. Alternatively, high levels of change in trait composition may indicate alteration of the Habitat Templet, with community change driven by changes in the constituent taxa in response to the environmental conditions selecting against existing strategies.

To examine this relationship a measure of community variability based on traits of taxa within communities was calculated using a similar method as was employed for taxonomic variability. At each site an average trait profile was calculated for each sample. This was achieved by taking an average value for each of the trait descriptors. For example the trait "Maximum Potential Size" is composed of 7 descriptors (≤ 0.25 cm; > 0.25-0.5 cm; > 0.5-1 cm; > 1-2 cm; > 2-4 cm; > 4-8 cm and >8cm), with each taxon assigned a score of between 0 and 5 based on its affinity for each descriptor as discussed above. For each descriptor an average affinity value was calculated using information for all taxa in the sample. Therefore if there is a change in the Habitat Templet it would be expected that this would be reflected in a change in the average affinities for descriptors through time. For example if a change in the physicochemical environment favour smaller taxa this would result in an increase in the average affinity of the community for smaller size descriptors.

Having calculated an average trait profile for each sample at a site, trait temporal variability was calculated using an identical method to that used to calculate temporal variability based on taxonomic identity. A dissimilarity matrix was calculated using Kulczynski distance and then agglomerative hierarchical clustering used to construct a dendrogram where branch length represents the difference between trait profiles of the samples, and thus provides an aggregate measure of temporal variability of traits.

The relationship between trait and taxonomic based temporal variability was then examined using robust regression, including 97.5% of the data, to remove the influence of outliers (Reimann et al., 2008).

5.3: Results

5.3.1: Contribution of individual taxa

Logistic regression identified three groups of taxa based on the relationship with increasing community change. Table 5.1 summarises membership of these three categories by broad taxonomic groups.

Seven families were identified as representing a core of High Frequency taxa that occur in a high proportion of samples through time irrespective of the level of community variability. In each case logistic regression indicated a significant linear trend as shown in Figure 5.1a. This category is composed of the Sphaeriidae (Bivalvia), Hydrobiidae (Gastropoda), Gammaridae and Asellidae (Crustacea), Erpobdellidae and Glossiphonidae (Hirudinea) and Chironomidae (Diptera).

Table 5.1: Membership of broad groupings of taxa to each of the three groupings.

Taxa Group	High	Low	Driver
Bivalvia	1	1	0
Coleoptera	0	7	3
Crustacea	2	0	0
Diptera	1	2	2
Ephemeroptera	0	1	6
Gastropoda	1	3	1
Heteroptera	0	5	1
Hirudinea	2	1	0
Megaloptera	0	1	0
Odonata	0	5	0
Oligochaeta	0	2	0
Plecoptera	0	1	5
Trichoptera	0	6	10
Turbellaria	0	1	2

A second group containing 37 families represents taxa found in a small proportion of samples through time irrespective of the level of temporal change. These taxa are termed Low Frequency taxa and consist of 7 families of Coleoptera, 6 families of

Trichoptera, 5 families of both Heteroptera and Odonata, 3 families of Gastropoda, 2 families of Diptera and single families from the Bivalvia, Crustacea, Ephemeroptera, Hirudinea, Megaloptera, Plecoptera and Turbellaria. The typical relationship between community change and persistence is illustrated in Figure 5.1b. In a number of instances no significant linear relationship between community change and persistence was detected. However, on examination in all instances taxa were found to occur consistently in less than 50% of samples through time along the gradient of temporal change and as such were included in the group.

Finally, logistic regression identified 30 families that exhibited a significant negative relationship with increasing community change (Figure 5.1c). These represent taxa that exhibit varying persistence along the gradient of temporal change. Where communities vary little through time these taxa exhibit long-term persistence, however as community temporal variation increases this group of taxa are predicted to occur in a decreasing proportion of samples. These Variable Frequency taxa include 10 families of Trichoptera, 6 families of Ephemeroptera, 5 families of Plecoptera, 3 families of Coleoptera, 2 families of both Turbellaria and Diptera and 1 family of both Gastropoda and Heteroptera.

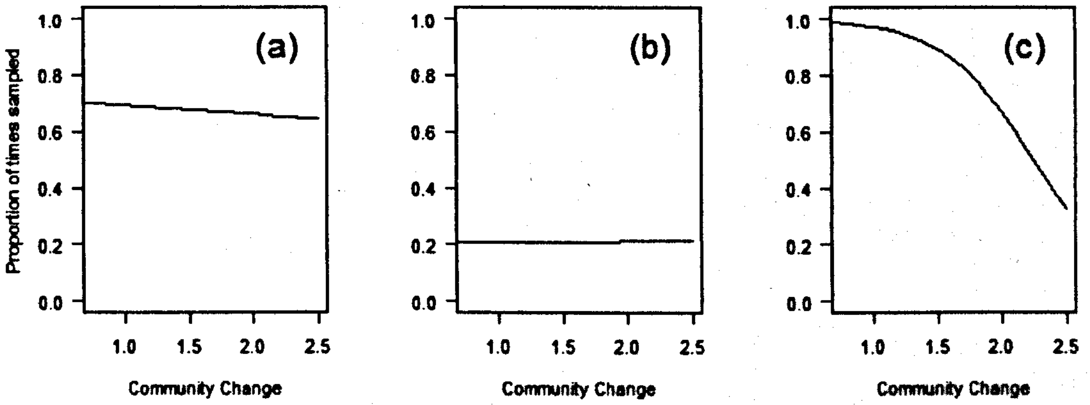


Figure 5.1: Typical relationship between the proportion of times that a taxon is sampled along the gradient of community temporal variation. Three distinct relationships are shown representing (a) taxa present in a high proportion of samples termed High Frequency taxa, (b) taxa present in small proportion of samples termed Low Frequency taxa, and (c) taxa that are present in decreasing proportion of samples as community variation increases termed Variable Frequency taxa.

5.3.2: Relationship between temporal change and traits of taxa

Linear Discriminant Analysis (LDA) was able to derive a clear separation of the three categories based on traits (Figure 5.2). Annotations on the axis of Figure 5.2 indicate the most important traits for separating groups based on LDA axis loadings. The strongest separation on axis 1 is between High Frequency taxa and the other groups and is driven by a higher affinity for these taxa to produce free clutches of eggs or reproduce through ovoviparity, the presence of an aquatic adult and tolerance of polysaprobic conditions. Cross validation using a subset of the original data showed that the classification was 77% accurate, indicating that traits provide a meaningful way of discriminating between the three categories. Errors most commonly arose where Variable Frequency taxa are categorised as being Low Frequency taxa and vice versa. Both Variable and Low Frequency taxa were rarely misclassified as High Frequency taxa.

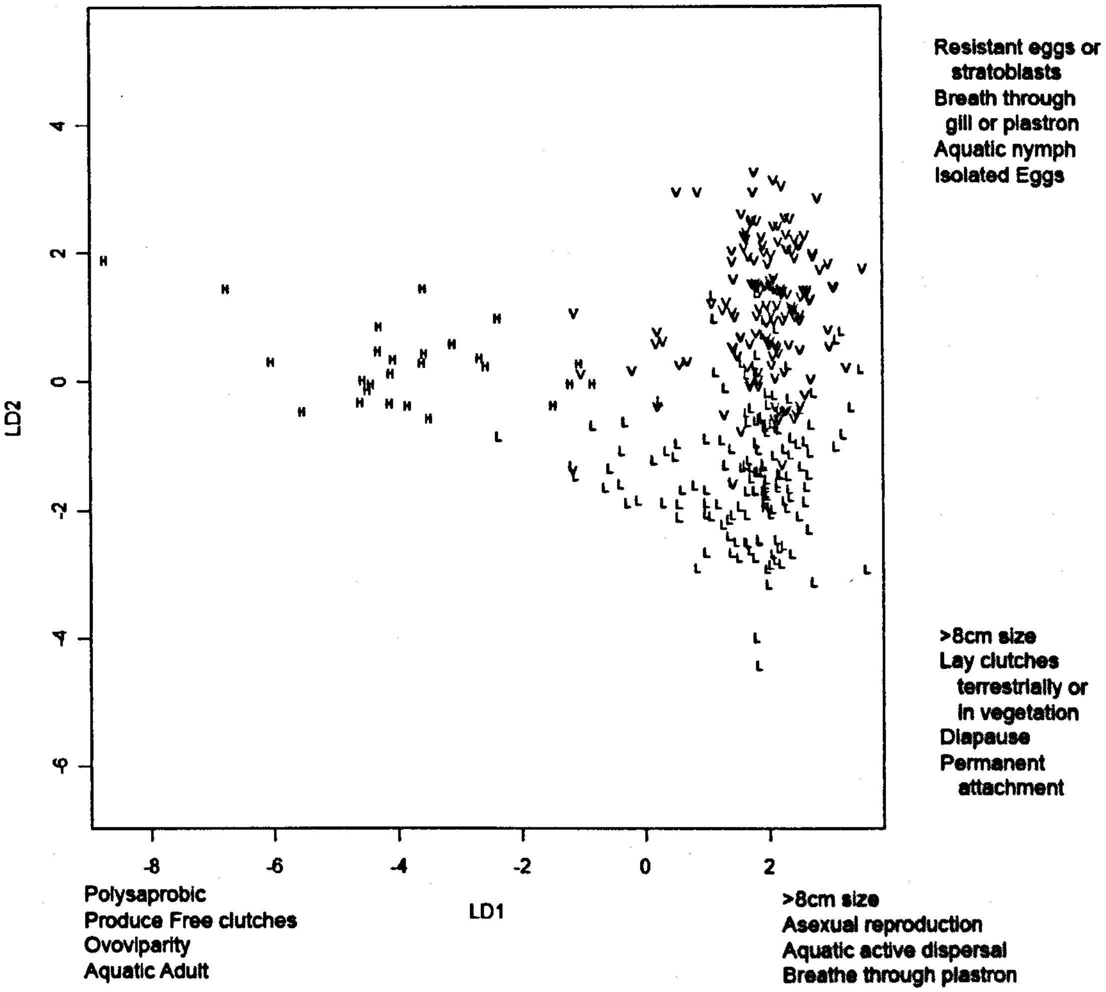


Figure 5.2: Linear Discriminant Analysis separating (H) High Frequency taxa, (L) Low Frequency taxa, and (V) Variable Frequency taxa. Accuracy of classification through cross validation is 77%.

The first set of traits considered relate to the life cycle of taxa. The trait maximum potential size (Figure 5.3a-g) was similar across the three groups with the only significant difference being for High Frequency taxa to reach 2-4cm. Whilst all three groups had a strong affinity for a life cycle duration of ≤ 1 year there was a stronger affinity for this in Variable Frequency taxa (Figure 5.3h) whereas both High and Low frequency taxa had an increased affinity to a life cycle duration of > 1 year (Figure 5.3i). All three categories had a strong affinity towards 1 life cycle per year suggesting this is a common strategy across taxa (Figure 5.3k). However, there was a stronger affinity for Low Frequency taxa to have multiple life cycles per year

(Figure 5.3l). There is a high affinity for eggs and larvae (Figure 5.3m-n) to be found within the aquatic environment for all three categories. Major difference are exhibited as an increased affinity for Variable Frequency taxa to be present as nymphs (Figure 5.3o), and for High and Low Frequency taxa to be present as aquatic adults (Figure 5.3p).

Almost uniquely amongst the three categories High Frequency taxa employ ovoviparity as a reproductive strategy (Figure 5.4a). However, the dominant reproductive strategy across all groups is based on the production of fixed clutches of eggs (Figure 5.4d). Alternative strategies include the production of free clutches (Figure 5.4e) in High Frequency taxa. An alternate strategy for Low Frequency taxa to produce clutches in vegetation (Figure 5.4f).

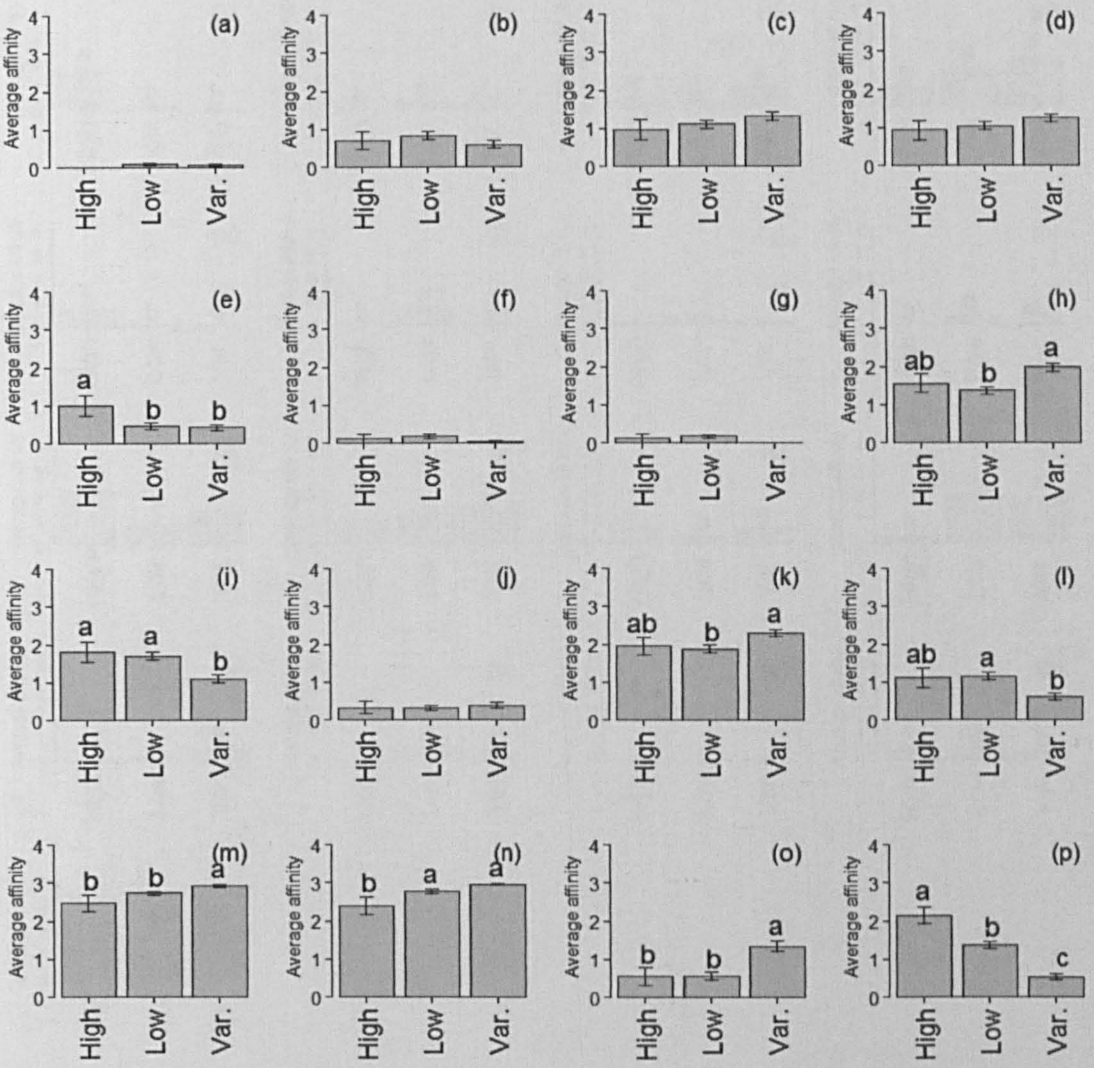


Figure 5.3: Modalities for life cycle traits:- **(a-g) Maximum potential size** (a) ≤ 0.25 cm, (b) > 0.25-0.5 cm, (c) > 0.5-1 cm, (d) > 1-2 cm, (e) > 2-4 cm, (f) > 4-8 cm and (g) >8cm; **(h-i) Life cycle duration of** (h) ≤ 1 year or (i) > 1 year; **(j-l) Potential number of life cycles per year** (j) <1, (k) 1, or (l) >1; (m-p) Aquatic stages (m) egg, (n) larva, (o) nymph or (p) adult. Letters indicate significant differences between trait affinities based on multiple comparison tests following Kruskal-Wallis. Where letters are absent there was no significant difference between categories.

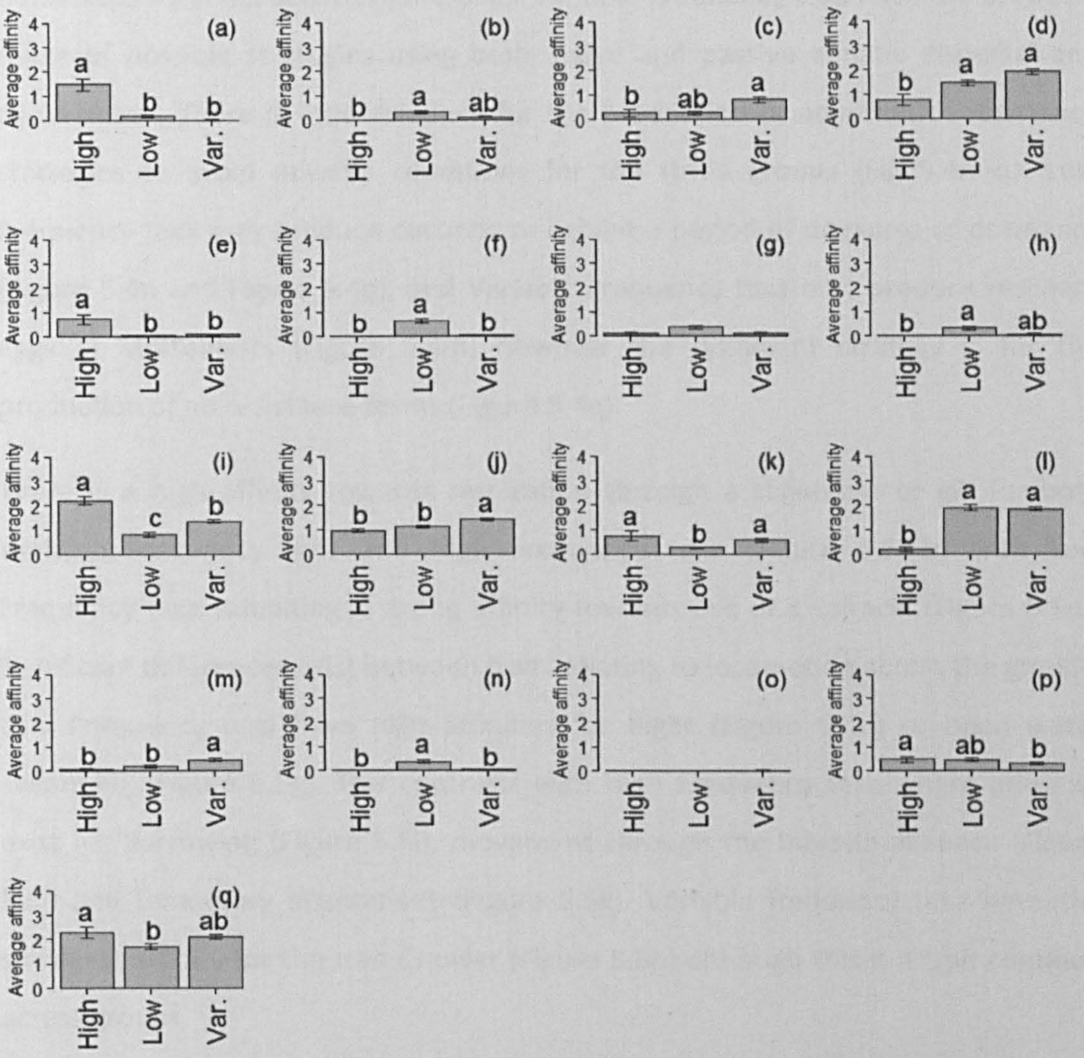


Figure 5.4: Modalitites for resistence and resilience strategies. **(a-h) Reproductive strategy** (a) ovoviviparity, (b) isolated eggs, free, (c) isolated eggs, cemented, (d) clutches, cemented or fixed, (e) clutches, free, (f) clutches, in vegetation, (g) clutches, terrestrial and (h) asexual reproduction; **(i-l) Dispersal** (i) aquatic passive, (j) aquatic active, (k) aerial passive, and (l) aerial active**; (m-q) Resistance forms** (m) eggs, statoblasts, (n) cocoons, (o) housings against desiccation, (p) diapause or dormancy, (q) none. Letters indicate significant differences between trait affinities based on multiple comparison tests following Kruskal-Wallis. Where letters are absent there was no significant difference between categories.

Dispersal and resistance stages are key features that may influence persistence of taxa by conveying resistance or resilience. There were marked differences in dispersal strategy between the three groups. High Frequency taxa had a strong affinity for aquatic passive dispersal with Low Frequency taxa exhibiting the least affinity for this form (Figure 5.4i). Instead dispersal by Low Frequency taxa was dominated by aerial active (Figure 5.4l). Variable Frequency taxa have the broadest range of possible strategies using both active and passive aquatic dispersal and aerial active. There is little evidence for the production or adoption of resistance strategies to avoid adverse conditions for the three groups (Fig 5.4m-q). Low Frequency taxa may produce cocoons or exhibit a period of diapause or dormancy (Figure 5.4n and Figure 5.4p), and Variable Frequency taxa may produce resistant eggs or stratoblasts (Figure 5.4m) however the dominant strategy is for the production of no resistance forms (Figure 5.4q).

There is a high affinity towards respiration through a tegument or gill for both Variable Frequency taxa and High Frequency taxa (Figure 5.5a-b) with Low Frequency taxa exhibiting a strong affinity towards use of a spiracle (Figure 5.5d). Significant differences exist between traits relating to locomotion across the groups. Low Frequency taxa have high affinities for flight (Figure 5.5e) or open water swimming (Figure 5.5g). This contrasts with High Frequency taxa where affinities exist for burrowing (Figure 5.5i), movement through the interstitial space (Figure 5.5j) and temporary attachment (Figure 5.5k). Variable frequency taxa have the strongest affinity for the trait Crawler (Figure 5.5h) although this is a trait common across groups.

Finally, Variable Frequency taxa seem to exhibit a stronger preference for nutrient poor conditions ranging from xenosaprobic to oligosaprobic (Figure 5.5m-n). This contrasts with High Frequency taxa that exhibit affinities for a range of conditions from the least to the most nutrient rich (Figure 5.5m-q).

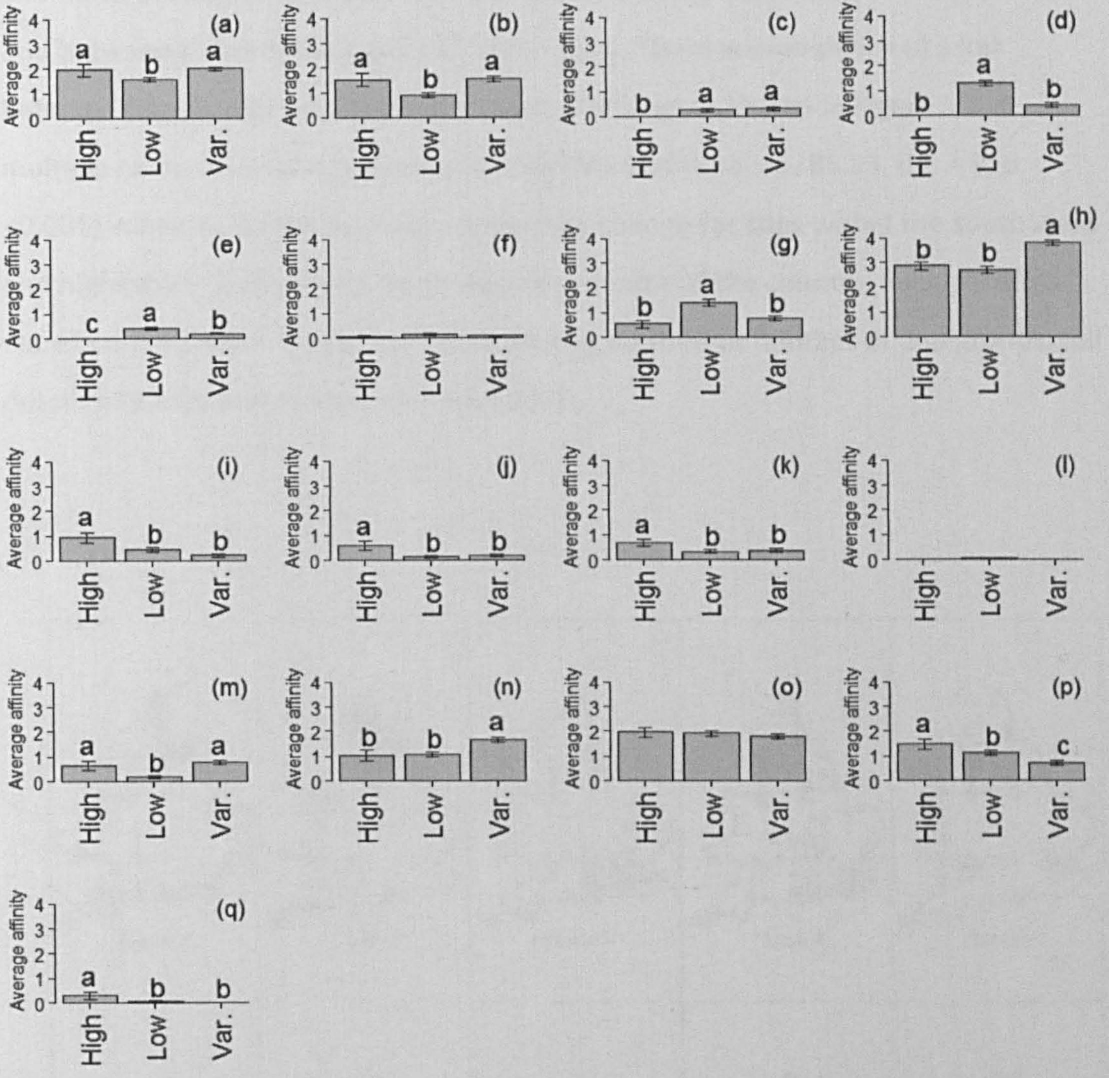


Figure 5.5: Modalities for respiration, locomotion and saprobity. (a-d**) Type of respiration** (a) tegument, (b) gill, (c) plastron, (d) spiracle ; (e-l) **Locomotion** (e) flier, (f) surface swimmer, (g) full water swimmer, (h) crawler , (i) burrower, (j) interstitial, (k) temporarily attached, or (l) permanently attached; (m-q) **Saprobity;** (m) xenosaprobic, (n) oligosaprobic, (o) b-mesosaprobic, (p) a-mesosaprobic, and (q) polysaprobic. Letters indicate significant differences between trait affinities based on multiple comparison tests following Kruskal-Wallis. Where letters are absent there was no significant difference between categories.

5.3.3: Community composition and temporal change

Irrespective of the clustering method and number of end groups used a clear spatial pattern within communities emerged. This is illustrated in Figure 5.6 based on Ward linkage with 10 end groups. There are similarities between communities to the west and north of the country, between communities in the south east of the country, and between those in the centre of the country. There was evidence of a link between community structure and temporal change as shown in Figure 5.7. A multiple comparison test following Kruskal-Wallis (Chi-sq = 1185.23, d.f. = 9, p <0.001) indicates lowest levels of community change for sites within the south west and highest levels of change for those in the centre of the country. Such patterns were broadly similar irrespective of clustering method or number of end groups, full details of which are provided in Appendix 1.

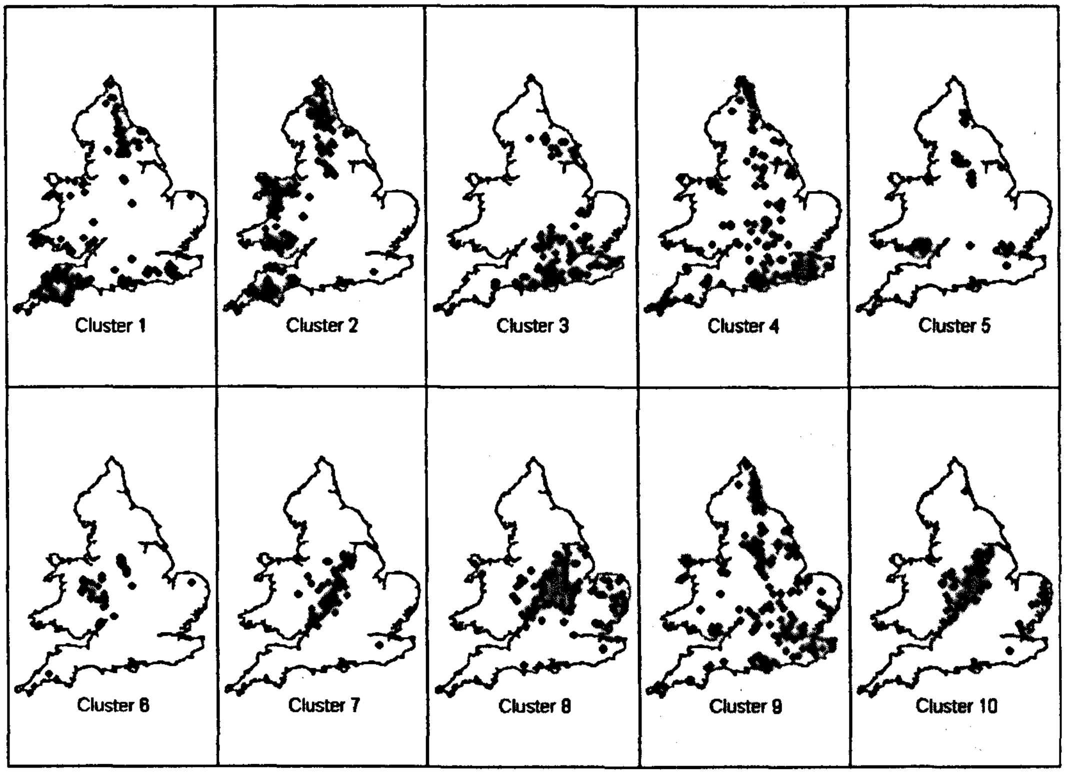


Figure 5.6: Clustering of communities based on compositional similarity using Ward linkage and 10 end groups.

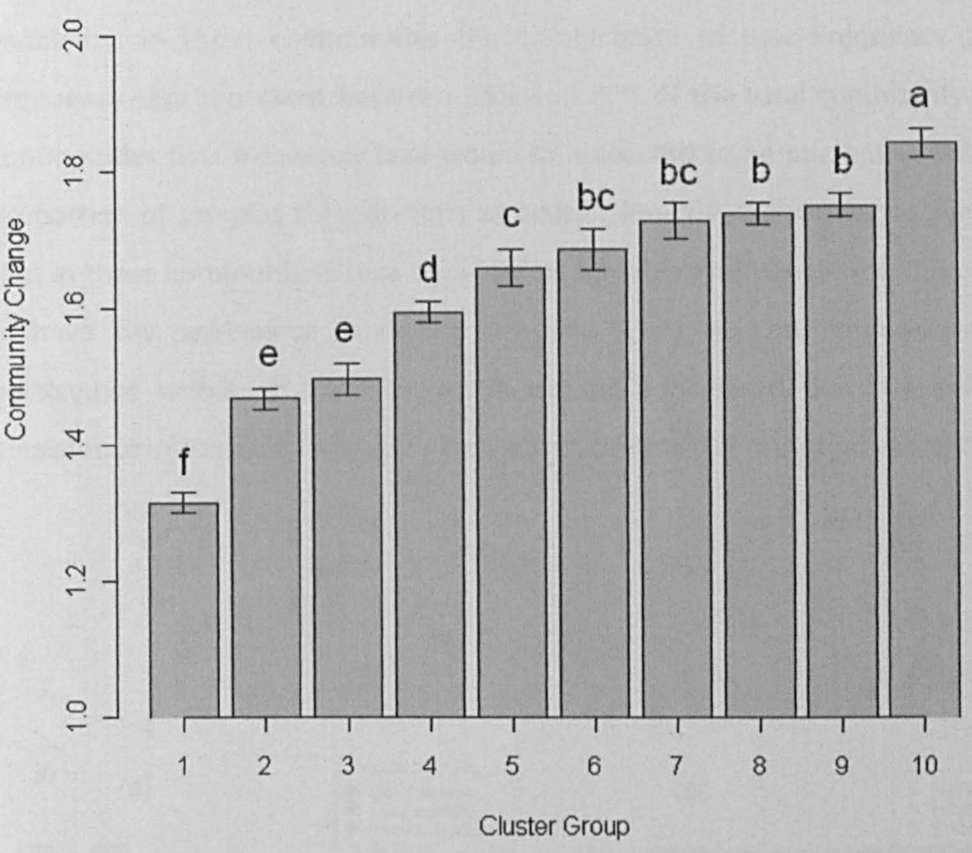


Figure 5.7: Average values for community change 1990-2005. Cluster group refers to those derived from Ward linkage and 10 end groups as shown in Figure 5.6. Letters indicate significant differences in community change based on multiple comparison tests following Kruskal-Wallis.

Community structure was described in terms of the typical percentage composition of each of the three groups. Communities in Figure 5.8 correspond to those shown in Figure 5.7. This represents a gradient of increasing temporal variation from low (Community 1) to high (Community 10). Communities that exhibit the lowest rates of change are composed of a large proportion of taxa classified as Variable Frequency taxa (communities 1, 2 and 3). Logistic regression indicates these taxa will be present in a high proportion of samples through time as community change is low. Combined with the High Frequency taxa the majority of the community (between 70% and 86%) is composed of taxa predicted to be persistent through time.

Communities 6 to 10 exhibit statistically the highest community temporal variability. In these communities the combination of Low Frequency and High Frequency taxa represent between 35% and 79% of the total community. In these communities Low frequency taxa would be expected to be present in only a small proportion of samples through time at a site. Similarly, logistic regression predicts that in these communities taxa classified as Variable Frequency would be expected to have low persistence as shown in Figure 5.1(c). As Low Frequency taxa lack persistence within all communities it can be considered that the shift in the persistence of Variable Frequency taxa drives community temporal variability.

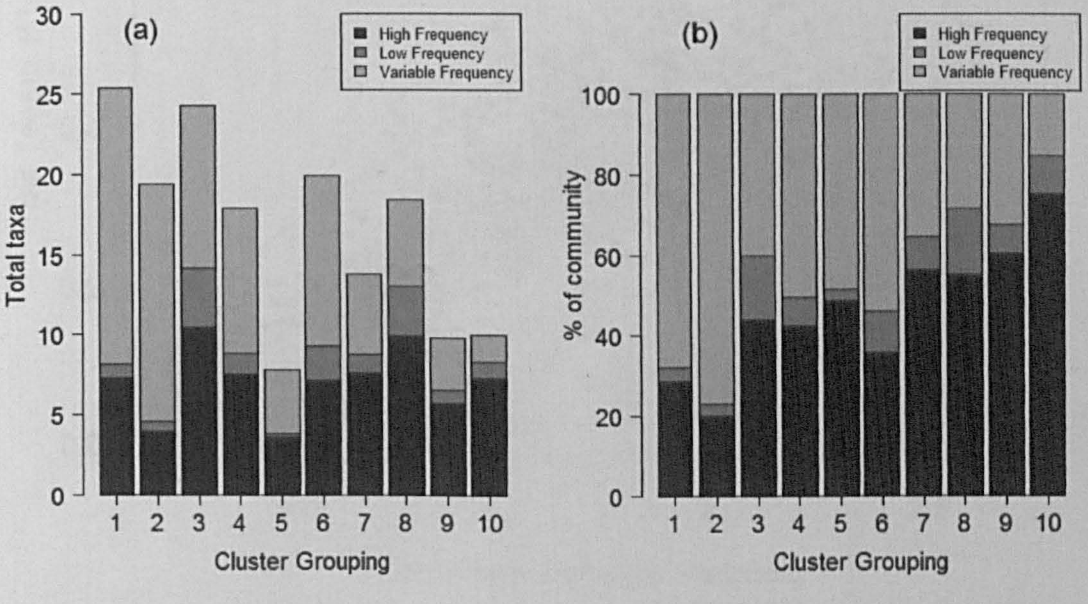


Figure 5.8: Composition of communities based on (a) average number of taxa within each category, (b) % of community composed of taxa from each category. In both cases cluster grouping refers to the communities as shown in Figure 5.7. As such implicit in the figure is an increase in community variability from group 1 to 10. (Letters indicate significant differences between trait affinities based on multiple comparison tests following Kruskal-Wallis. Where letters are absent there was no significant difference between categories).

5.3.4: Taxonomic or trait variability

Robust regression indicates a strong positive correlation (r = 0.83, d.f. = 1534, p <0.001) between temporal variability based on trait and taxonomic composition (Figure 5.9). Where communities vary little through time there was a strong relationship between the two measures. However, as variability increases there are an increased number of outliers. This would suggest that in communities that vary most through time there may be more marked changes in the trait structure.

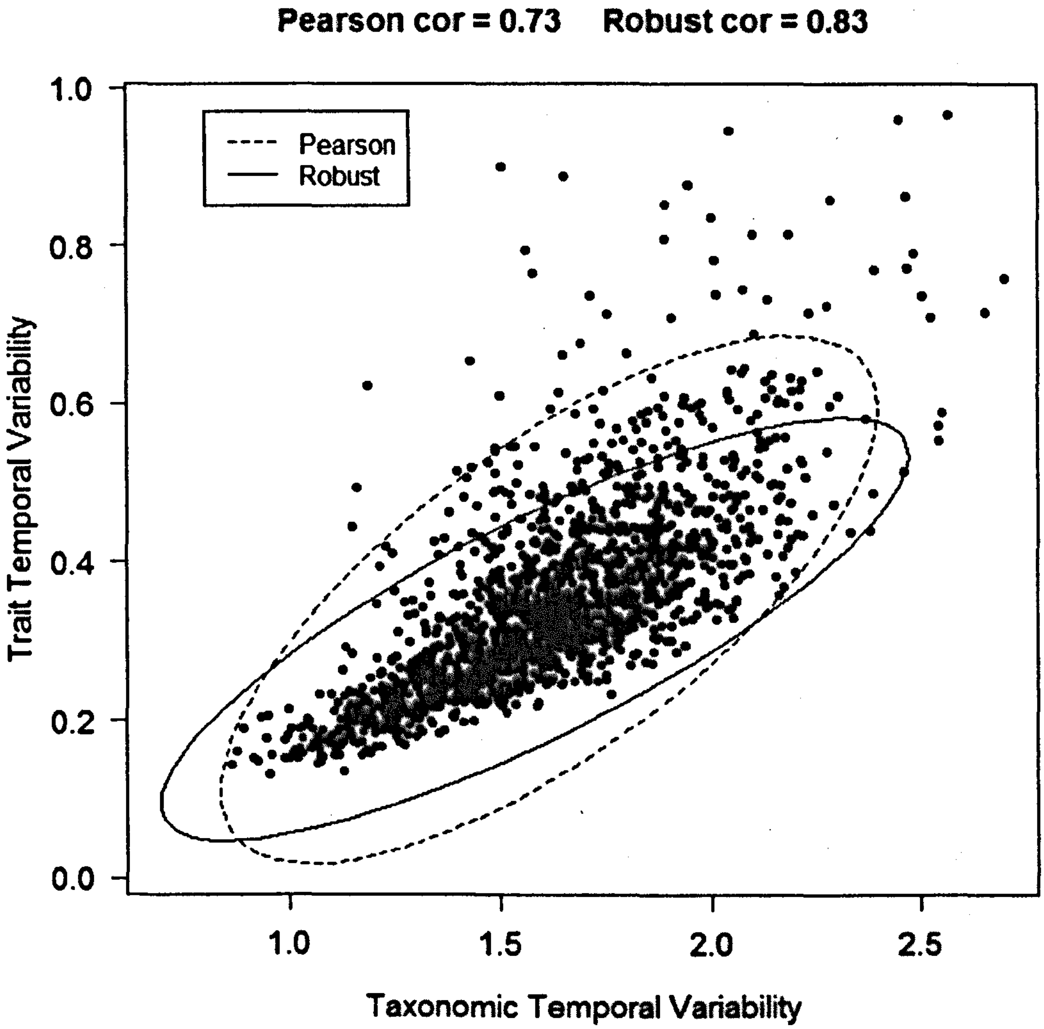


Figure 5.9. Pearson and Robust correlation between temporal variability of communities based on trait and taxonomic descriptors.

5.4: Discussion

To be persistent at a site organisms can adopt strategies such as a short life cycle and high reproductive potential or can be long lived and maximise persistence through behaviour, morphology or physiological mechanisms such as streamlining as both strategies provide responses to disturbance (Gasith and Resh, 1999). Results from the current study suggest that High Frequency taxa achieve persistence within communities through these two contrasting mechanisms. Resistance mechanisms are present within the Gastropoda and Bivalvia, long lived taxa with protective shells enabling them to escape disturbance, which exhibit both temporary and permanent attachment reducing the impact of event such as spates, and have investment in young through ovoviparity. Similarly, the Gammaridae and Asellidae are long lived, present in the aquatic environment as adults and invest energy in brood care. This contrasts with the other representatives of the group, most notably the Diptera, that are characterised by high reproductive potential and fast colonisation, traits providing the population with considerably resilience (Resh et al., 1988).

Results indicate that there was not a significant difference between many of the traits of Variable Frequency taxa and High Frequency taxa. This result suggests that there are only a limited number of features that make Variable Frequency taxa either more susceptible to disturbance or limited their recovery potential or both. Variable Frequency taxa display good dispersal potential, suggesting that recolonisation following disturbance is unlikely to limit their recovery potential. However, many of these taxa are present in the aquatic environment primarily as a nymph, which represents a form susceptible to disturbance (Diaz et al., 2008). Compared with High Frequency taxa, Variable Frequency taxa have a low affinity for attachment, burrowing or use of the interstitial space potentially making them more prone to disturbance during high flow events. As such in frequently disturbed sites, although they may rapidly colonise, they lack traits that confer resistance properties allowing the establishment of a persistent population. Variable Frequency taxa also have a low affinity for more than one life cycle per year potentially restricting their recovery following disturbance. Taxonomically this group is dominated by the Ephemeroptera, Plecoptera and Trichoptera, taxa that can be considered to be sensitive to disturbance (Feld and Hering, 2007).

Results from logistic regression indicate a third group of taxa that exhibit consistently low persistence irrespective of the level of community temporal change. These Low Frequency taxa exhibit high dispersal potential as well as a high affinity for traits such as presence of an aquatic adult, and possess other adaptations such as breathing through a spiracle that would be expected to provide resistance against disturbance. They have a low affinity for a short life cycle and the potential for more than one life cycle per year, adaptation that would be predicted to lead to temporal persistence within communities. Taxonomically this group was dominated by the Coleoptera, Heteroptera and Odonata, taxa that are often associated with anthropogenically stressed sites, as their traits mean they are able to move away from and avoid pressures (Diaz et al., 2008). For example, the Coleoptera are best described as mobile generalists with the majority respiring at the water surface making them resistant to low oxygen concentrations. They have a well developed adult structure that protects against variation in the physicochemical environment and have the option to move within or leave the water should conditions become unfavourable (Richoux, 1994). Typical members of the Heteroptera and Odonata favour areas with low hydrological connectivity (Paillex et al., 2009) however, because of good dispersal abilities they are found in a wide range of habitats. The reliance of these taxa on dispersal probably accounts for the low proportion of samples in which these taxa were found.

The examination of the relationship between persistence and traits of taxa suggests that certain taxa may be more closely constrained by environment conditions than others owing to the strategies that they employ, and that this leads to differences in the persistence of taxa demonstrated in this study. Major groups such as the Ephemeroptera, Plecoptera (Usseglio-Polatera and Tachet, 1994) and Trichoptera (Tachet et al., 1994) that are considered to show strong relationship with trait affinities and habitat characteristics and have traits that may be considered to make them more susceptible to disturbance and as such show varying persistence within communities. Other groups, for example the Gastropoda and Diptera, employ markedly differing strategies that focus on either resistance or resilience to disturbance, resulting in persistence within the community through differing mechanisms. Finally, a range of taxa typified by the Coleoptera show little congruence between the expected level of persistence based on their traits and their observed persistence within the community, suggesting that the adaptive mechanisms that they employ free them from rigid environmental constraints.

Results presented in Figure 5.6, Figure 5.7 and in Appendix 1 suggest a link between community composition and temporal variability. In demonstrating that communities that exhibit most change are least diverse, and that temporal variability is driven by changing persistence of sensitive taxa (e.g. Ephemeroptera, Trichoptera and Plecoptera) results from the current study suggests that anthropogenic disturbance may play a key role in controlling temporal variability. Spatial patterns in community structure correspond to gradients of anthropogenic disturbance (Murphy and Davy-Bowker, 2005) across England and Wales that are known to influence community structure. Charvet et al. (2000) demonstrates that communities in natural or semi-natural conditions are functionally diverse but develop towards more specialised (i.e. less diverse structure) when disturbed by anthropogenic impacts. Similarly increasing temporal variability, as was used as the gradient with which to examine community structure, is considered to be indicative of stressed communities (Odum, 1985; Fraterrigo and Rusak 2008) with results from the current study demonstrating that this response arises through the affect that disturbance has on a specific group of taxa (i.e. Variable Frequency taxa). Such findings provide a potential mechanism for the relationship between temporal variability and community structure as described. There is known to be a continuous redistribution of taxa through drift in the water column (Townsend and Hildrew, 1976) with recolonisation of disturbed habitats occurring rapidly (McCabe and Gotelli, 2000). Whilst there is a stable core of highly persistent taxa in all communities, increasing temporal variability may arise due to continued colonisation of sensitive taxa into sites where conditions do not permit long-term persistence. Of the taxa considered in the current study Brittain and Eikeland (1988) demonstrate that most drift is by the Ephemeroptera, Plecoptera, Trichoptera and Coleoptera suggesting that these will colonise disturbed sites rapidly. However, this will mean they occur in sub-optimal conditions where they will not persist (Poff and Ward, 1990; Resh et al., 1994) due to either a direct mortality effect or through the increased rates of drift noted in disturbed sites where they may colonise but then rapidly move on as conditions are unsuitable (Brittain and Eikeland, 1988). This idea is supported by other studies that have demonstrated a close relationship between traits and environmental features following disturbance, and in headwater streams where the input of taxa through drift is smaller (Townsend et al., 1997). These examples suggest that taxa often occur in sub-optimal conditions where mis-matches between their traits and the environmental conditions mean that their ability to persist through time may be limited.

In one of the few studies to examine long-term variability of trait based measures Beche et al. (2006) found little variation in traits over time compared with large variations in community composition. In Beche's (2006) study taxa relied on constant recolonisation (a resilience strategy) due to large seasonal difference in environmental conditions. The authors considered that the low variation in trait based temporal variability arose due to trait "underdispersion" (Weiher and Keddy, 1999) where the harshness of the abiotic environment meant that the geographic species pool was composed of taxa with a limited set of traits. So whilst taxonomically the community varied markedly over time, there were only a limited number of possible traits resulting in low trait variability even with high species turnover.

In the current study there was a strong positive relationship between temporal variability based on taxonomic and trait based measures (Figure 5.9). This suggests that rather than the adaptive strategies of taxa driving temporal variability as in Beche et al.'s (2006) study, environmental factors play a central role in determining temporal variability. River systems in England and Wales have the potential to support a diverse fauna both taxonomically and in terms of the traits present. The relationship between trait and taxonomic based variability described in the current study can therefore arise in two ways. Firstly, as has already been discussed colonisation of sites is a constant process through the invertebrate drift. Mismatches between the Habitat Templet and the traits of colonising taxa will result in high trait and taxonomic based variability as these colonists are lost. Secondly, the relationship described in Figure 5.9 could be driven by variability in environmental conditions themselves. In response to such environmental variability, which represents a change in the Habitat Templet, taxa that previously possessed traits allowing long-term persistence will be lost and replaced by taxa which possess traits suited to the new Habitat Templet.

Findings from the current study have important implication for biomonitoring. Results indicate that communities with a high proportion of Variable Frequency taxa exhibit less temporal variability through time than those dominated by High Frequency taxa. As Variable Frequency taxa are those that are most sensitive to disturbance these findings suggest that any observed change in these communities would be cause for concern as taxa would be expected to be persistent through time. However, communities with a high proportion of High Frequency taxa would be expected to exhibit high levels of temporal variation possibly due to the continual colonisation and loss of sensitive taxa. In these communities, where alteration in community structure may be a desired management goal, as it is likely to be indicative of less stressful conditions, any observed changes in composition is more likely to result from random changes in the community through time. To have confidence that any changes results from a change in conditions it may be necessary to invest more resources in monitoring the community through time.

CHAPTER 6: SHEFFIELD FRESHWATER INVERTEBRATE SURVEY:

28 YEARS ON

6.1: Introduction

Long-term studies have made important contributions to not only ecological science but also to the identification and understanding of processes of broader societal concern (Strayer et al., 1986). However, whilst there is agreement among ecologists about the important contribution that long-term studies of systems can make, variation in communities over long periods of time has received little attention compared with seasonal or interannual variation (Johnson et al., 1994). In fresh water systems long-term studies have led to increased understanding of phenomena such as demographic balance (e.g. Speirs et al., 2000), persistence of taxa (e.g. Townsend et al., 1987), competitive coexistence (e.g. Elliott, 2006) and seasonality (e.g. Wolda, 1988). They have identified and increased understanding of issues such as acidification (e.g. Weatherley and Ormerod, 1987; Woodward et al., 2002) and climate change (Burgmer et al., 2007; Chessman, 2009), and have helped to shape management practices (e.g. Bradley and Ormerod, 2002; Ormerod and Durance, 2009) and inform the design of monitoring programmes (e.g. Scarsbrook et al., 2000; Milner et al., 2005). However, despite their importance McElravy et al. (1989) and Jackson and Fureder (2006) demonstrated that globally very few studies exist of more than five years in length. In the UK, with only a few exceptions (e.g. Townsend et al., 1987; Weatherley and Ormerod, 1990; examples cited in Elliott, 1990), temporal variability in macroinvertebrate community structure over long time periods has seldom been examined. This can be attributed to both a lack of historic datasets (Johnson et al., 1994) and to the difficulty in maintaining historic records in an accessible form (Strayer et al., 1986). Here one such dataset (Zasada and Smith, 1981) collected in Sheffield Metropolitan District, South Yorkshire, is used to examine changes in macroinvertebrate community over a 28 year period from 1979 to 2007.

The Sheffield Metropolitan District contains approximately 150 km of running water that lie predominantly within the Don catchment system (Zasada and Smith, 1981). Until comparatively recently some of the waterways in the Don system were considered to be among the most heavily polluted in Europe, as a result of the extensive development of heavy industry in the area since the 18th Century (Firth, 1997). In common with many other river systems in England and Wales, in the last 20 years the Don catchment has seen a marked improvement in water quality driven by investment from industry and other organisations (Durance and Ormerod, 2009).

In 1979-1980 a study was carried out, under the direction of the Sheffield City Museums service that set out to document the occurrence of freshwater invertebrates in both still and flowing waters across the entirety of Sheffield Metropolitan District. The survey area was defined by the administrative boundary of the city. Zasada and Smith (1981) consider that within the area there are three distinct regions. To the west are the upland areas, mainly above 305 meters, that are characterised by acid water flowing over Millstone Grits and peat deposits. To the east of these, and on the outskirt to the west of the city of Sheffield, are the lower reaches of the rivers Loxley, Rivelin, Porter and Sheaf which, after crossing rocks of the Coal Measures, become less acidic and less turbulent. Finally, the remainder of the rivers, flowing through the city itself, are characterised as slower moving and were heavily polluted with industrial and domestic waste. The aim of the 1979 study was to sample one flowing water and one still water habitat within each of the 400 1 km x 1 km grid squares covering the survey area, though in practice not all grid squares had both, or in some cases either, habitat. In total 423 sites were visited, of which 299 were lotic. The history of Sheffield and its industrial past makes such a spatially explicit dataset of particular interest in terms of examining recovery of systems from severe degradation. Such spatial coverage is rare in most long-term datasets (Richards and Minshall, 1992; Collier, 2007) adding to the value of the historic records as it allows the examination of communities across a range of environmental conditions.

The present study focussed on a series of streams flowing from the uplands on the western outskirts of the city, relatively undisturbed by industry in their upper reaches, which were considered in the 1979 survey to represent sites with the richest and most varied invertebrate fauna (Figure 6.1). Although historically extensively utilized for small scale water powered industries, by the early 20th Century, most of this activity had ceased on these rivers and by the middle of the century the systems could be considered rural (South Yorkshire Historic Environment Characterisation). The principal factors influencing these sites in the present day are considered to be the presence of upstream reservoirs and the associated influence that this will have on hydrology and geomorphology (Konrad et al., 2008). As these sites were considered to be the most diverse in the historic study they represent good candidate sites with which to examine persistence of taxa over long time periods. As industry was considered to have halted by the middle of the 20th century in these systems, they present a rare opportunity to examine the extent that communities recover following the removal of pressures (Harding et al., 1998) and their structure may provide insight into the historic pressure on the system (Metcalfe, 1989).

The principal aim of the current study is to examine changes in the occurrence and distributions of invertebrates at the same locations for sample dates 28 years apart. This is done by first comparing the overall distributions of taxa, taxon and trait-based measures of community structure, and spatial structure of communities, between sample dates, and second, by directly calculating the extent of community change at each site across the sample period. Differences in methodology between years are also examined to assess the limitations of the data, and the usefulness of the dataset for further studies into the macroinvertebrate fauna of the system.

6.2: Methods

6.2.1: Historic data

The Sheffield Invertebrate survey was sponsored by the Sheffield City Museums and was conducted during 1979 and 1980 by a two person team whose remit was to *"collect, identify and record water-beetles, shrimps, caddis flies and the many other denizens of local ponds and streams"* (Zasada and Smith, 1981). Collection of macroinvertebrates was conducted using a naturalists pond net with a 1-mm mesh size *"on a 'maximum return' basis, whereby sweeps were made through the various microhabitats at each site such as the water surface, under stones, the substratum (gravel, mud or plant debris), amongst submerged aquatic plants and so on, in an attempt to obtain as many different species as possible. ... No attempt was made to standardise the collecting operation, to a given number of sweeps, duration or area."* (Zasada and Smith, 1981). Collected specimens were preserved in Steedman's B, a mixture of propylene phenoxetol and propylene glycol. Identification was carried out by members of the museum and sampling team although Zasada and Smith (1981) also provides a list of specialist taxonomists who helped with identification or confirmation of specimens, suggesting a high degree of accuracy.

Zasada and Smith (1981) provide a detailed synopsis of the findings of the study illustrating the distribution of all taxa. For the current study the original record cards, which were archived in the city museum, were digitised and the information entered into a Microsoft Access database that now holds records from all 423 sites sampled in the original survey. Using ArcGIS the location of all sample points within the system was plotted. The distribution of sites was examined and 25 sites identified on western tributaries of the Don for re-sampling in 2007. The basis for choosing these sites was that they represented a number of parallel streams with a similar history allowing useful comparisons to be made.

6.2.2: Sampling in 2007

Sampling took place over the autumn of 2007 with sampling dates within the season matched as close as practicable to that of the 1979 survey (in all cases sampling took place within 14 days of the original sample date within the season). The distribution of sampling sites is shown in Figure 6.1. The location of 1979 survey points was described using an 8 digit grid reference. Assuming the sites of the original survey were recorded accurately the use of a GPS system allowed the original sampling point to be located within a 10 metre square.

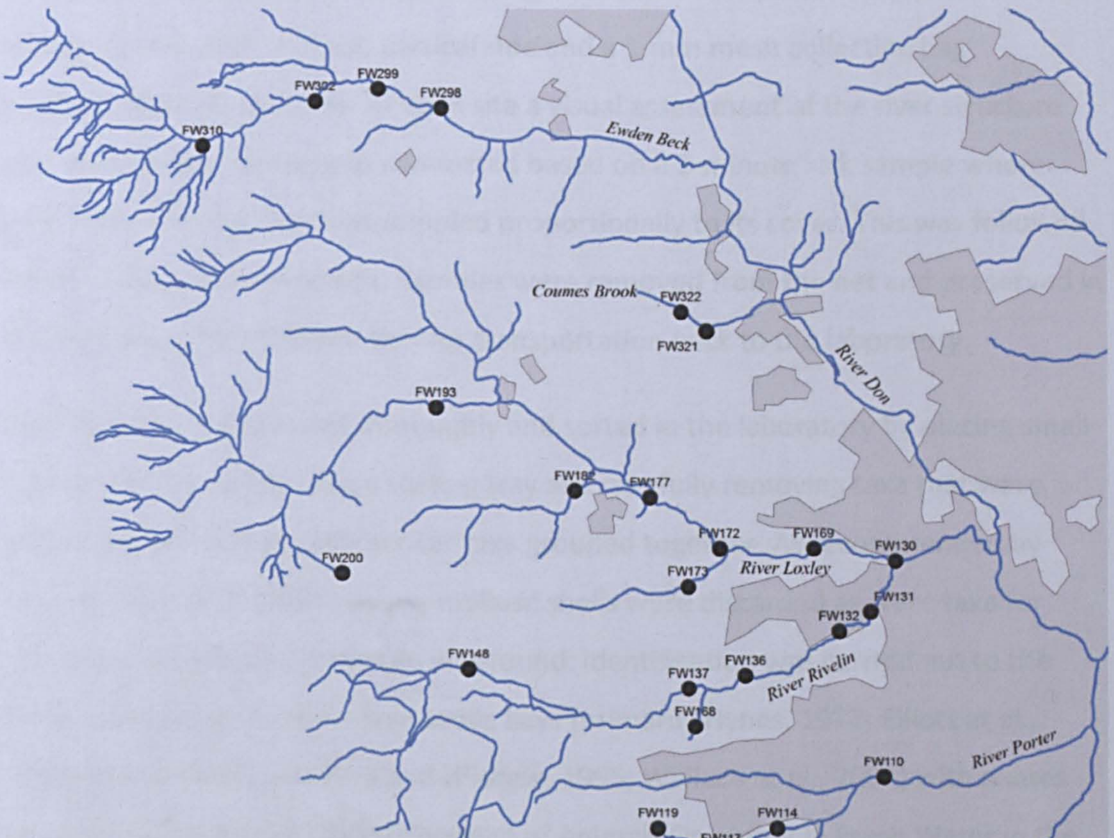


Figure 6.1: Distribution of joint 1979 and 2007 sampling sites on the Sheffield river network.

As described above, the original 1979 survey was based on a "maximum return" method of sampling with no effort to standardise. Samples could be argued to be comparable as there was consistency in the people conducting the survey, and therefore a consistent understanding of the intensity of sampling that constituted the maximum return. For the 2007 survey a decision was made to not try to replicate this method since there was insufficient information to know whether we were using the same criteria of completeness as the previous survey, but rather to use an established sampling protocol. The benefits of this were, firstly it would ensure consistency between samples in 2007, and secondly it would allow other researchers to replicate the sampling in the future at the same sites or in the wider river system and would provide samples of potential comparability to those used elsewhere in the study. Sampling was conducted using the method of Murray-Bligh et al. (1997). Each site was sampled using a standard FBA-pattern pond net that has a 25 cm lower edge, a 22 cm vertical side and a 1-mm mesh collecting bag approximately 50 cm deep. At each site a visual assessment of the river structure was made and sampling was conducted based on a 3-minute kick sample where each habitat in the river was sampled proportionally to its cover. This was followed by a 1-minute manual search. Samples were removed from the net and preserved in the field using 70% IMS solution for transportation back to the laboratory.

Each sampled was washed thoroughly and sorted in the laboratory by placing small portions of the sample into a sorting tray and carefully removing taxa that were placed in Petri dishes, with similar taxa grouped together. As recommended by Murray-Bligh et al. (1997) empty mollusc shells were discarded as were taxa for which only part of the specimen was found. Identification was carried out to the family level using standard taxonomic keys (primarily Hynes, 1977; Elliott et al., 1988; Friday, 1988; Edington and Hildrew, 1995; Wallace et al., 2003) with names standardised using the Coded Checklist of Animals Occurring in Fresh Water in the British Isles (Furze, 2007).

6.2.3: Methodological comparison

A noticeable feature of the 1979 data was the low recorded abundance of individuals in each sample compared to 2007 (Table 6.1). In 1979 an average of 20.8 individuals were sampled at each site compared with 389.2 in 2007. In Zasada and

Smith (1981) there is no explicit statement of whether in 1979 all sampled individuals were retained at each site or whether only a proportion of the total taxa sampled, thought to be representative of all taxa present, were retained for subsequent identification – though given the low numbers the latter seems the most likely. In order to examine the influence that the markedly different abundances may have on comparisons between years the program "Analytic Rarefaction 1.3" (http://www.uga.edu/strata/software/index.html), which uses the methods of Raup (1975) and Tipper (1977), was used to produce rarefaction curves based on random sampling of the 2007 data for each site. This enabled an examination of the number of individuals that would need to be sampled in 2007 to achieve the same estimate of taxon richness as 1979.

Table 6.1 summarises the result of this analysis for taxon richness and abundance for the 1979 and 2007 surveys. Results from the rarefaction analysis demonstrate that if the 2007 method was employed in 1979, it would have been necessary to sample on average 40 individuals, representing a significant increase from the actual number of individuals sampled in 1979 (paired-sample t-test, t = 4.32, d.f.= 24, p <0.001). This suggests that the 1979 method, where on average only 20.8 individuals were collected, was either a very efficient sampling method or that not all individuals were retained. Whilst it is difficult to separate these affects this result does suggest that the 2007 method, which sampled 389.2 individuals per site on average, may inflate the estimate of taxon richness making comparisons of absolute values between years difficult. This could either arise through the more intensive sampling or through the more detailed picking procedure that can be undertaken in the laboratory compared with the bank-side sorting used in 1979. As the aim of the current study was to compare communities both within and between years, a number of strategies were used to address this issue. Firstly, unadjusted data were used for comparisons within years and for descriptions of the most common taxa present. Secondly, comparisons between years were based both on the unadjusted data as well as 2007 data with the rarest taxa (defined as those with abundance of 5 or less) excluded. As demonstrated by Clarke and Hering (2006) the 2007 method has been demonstrated to sample an increasing proportion of rare taxa as sampling effort increases.

Table 6.1: Taxon richness and abundance for sites sampled in 1979 and 2007. Also presented is the results from rarefaction indicating; (a) the number of individuals that would have needed to be collected in 2007 to equal 1979 taxon richness; (b) 2007 standardised taxon richness based on the abundance of individuals sampled in 1979.

SITE_ID	**Taxon Richness**		**Abundance**		**Rarefaction**	
	1979	2007	1979	2007	(a) Required sampling intensity 2007	(b) Standardised 2007 Taxon Richness based on 1979 abundance
FW110	9	14	17	203	46	6
FW114	1	16	12	97	1	7
FW117	15	25	37	259	40	14.5
FW118	7	22	24	541	12	10.1
FW119	13	26	66	760	133	10.9
FW130	4	14	11	151	9	4.6
FW131	11	24	26	483	36	11.9
FW132	15	21	44	342	120	11
FW136	17	19	52	303	133	13.4
FW137	11	23	29	250	32	11.4
FW148	9	22	22	243	20	9.5
FW168	6	16	18	362	16	6.4
FW169	8	19	18	340	48	7.3
FW172	9	20	18	386	44	6.9
FW173	1	25	1	556	3	1
FW177	8	11	29	162	33	7.7
FW182	3	8	3	64	6	2.2
FW193	9	19	43	542	48	9.5
FW200	2	10	6	132	3	3.2
FW298	5	6	8	26	17	3.7
FW299	5	18	13	229	20	4.9
FW302	9	13	13	373	110	3.2
FW310	5	9	5	394	62	2
FW321	1	26	1	1151	1	1
FW322	4	21	5	1381	11	2.6

The second methodological question that arises is whether some taxa were systematically excluded from the 1979 sample due to the sampling method. As will be discussed in further detail in the main results section, there were 15 taxa that were uniquely sampled in 2007. Utilising the wider extent of the 1979 data (representing 299 sites) 6 of these taxa were demonstrated to be recorded at other sites within Sheffield Metropolitan District during the 1979 survey. This suggests that their absence could be attributed to biological and not methodological issues. For the remaining 9 taxa absent in 1979 (composed of two families of the Diptera, two of the Ephemeroptera, one of Plecoptera and four of the Trichoptera) whilst the specific families were not recorded in the wider sampling area, the presence of other representatives of these groups, with similar morphological and behavioural features, suggests once again that methodological reasons alone may not explain their absence in 1979. Therefore it can be concluded that the sampling method used in 1979 did not appear to systematically exclude these taxa.

A documentable difference between the 1979 and 2007 methodology arises through the treatment of Gastropoda and Bivalvia taxa. In the 1979 survey it is explicitly stated that where empty shells were found, these were collected and the presence of taxa recorded. For the 2007 protocol, as the method was designed for biomonitoring, empty shells are specifically excluded from the sample.

6.2.4: Community structure and composition

For each site taxon richness was calculated as the number of families present. To examine the influence of the differing methods, the 2007 data were calculated both as "unadjusted" counts of taxon richness and "standardised" taxon richness which is the estimate of richness (using rarefaction) based on the number of individuals sampled in 1979.

Site by taxon matrices were produced for each of the sampling years and dissimilarity measures calculated to examine variation between communities both within and across years. As the methodologies differed between years comparisons of communities was based primarily on taxa presence or absence. Differences in community structure were calculated based on Jaccard's dissimilarity index

(Legendre and Legendre, 1998). This index is a similarity measure with values ranging from 0 (representing identical communities) to 1 (no shared taxa) and takes the form;

$$\beta_{\mathrm{jac}} = 1 - \left(\frac{a}{a+b+c}\right)$$

Where **a** is the number of shared taxa between the two samples, **b** is the number of taxa unique to the first sample and **c** is the number of taxa unique to the second sample (Koleff et al., 2003). As other measures used in the study were all based on dissimilarity between communities, values for Jaccards index are reported as β_{Jac} calculated as;

Community structure was also described using a trait-based approach where taxa were assigned to broad groups based on either their biological or ecological characteristics as described by Usseglio-Polatera et al. (2001). Biological traits are described by 8 categories that separate primitive forms of aquatic life (e.g. Porifera, Oligochaeta) from more complex forms (e.g. Insecta) and large, long lived, slow reproducers from smaller, faster reproducing taxa. Table 6.2 summarises the principal biological features of these groups. Similarly grouping by ecological traits describes well known gradients in aquatic ecology that correspond to a shift in taxa from those that prefer high flows, coarse bed substrate , nutrient poor upland habitats to those with an affinity for slow flowing, nutrient rich lowland rivers, summarised in Table 6.3. This classification was based on a database of 472 taxa (Usseglio-Polatera et al., 2000; Usseglio-Polatera et al., 2001) with relationships between traits analysed using the fuzzy coding procedure of Chevenet et al. (1994). Classification based on biological and ecological traits is primarily at the species or genus level compared with the current study where identification was standardised to the family level. In most instances taxa could be assigned to a single biological or trait group. However, where the Usseglio-Polatera et al. (2001) classification indicated that species within a family could belong to more than one trait group taxa in the current study were assigned to each possible group.

For both 1979 and 2007 data communities at each site were described based on the number of taxa that belong to each of the biological and ecological trait groups. Dissimilarities between communities both within and across years were then calculated using the Kulczynski distance defined as;

$$\beta kul = 1 - \left(\frac{1}{2}\right)\left\{\frac{\sum_i min\ (X_{ij}, X_{ik})}{\sum(X_{ij)}} + \frac{\sum_i min\ (X_{ij}, X_{ik})}{\sum(X_{ik)}}\right\}$$

Where the formula calculates the dissimilarity between two objects *j* and *k*, based on attributes *i* = 1 to N (Faith et al., 1987). Values of 0 represent complete similarity and values of 1 complete dissimilarity between communities.

Table 6.2: Biological traits of taxa corresponding to the classification scheme of Usseglio-Polatera et al. (2001).

Biological trait category	Description	Typical taxa
A	Large and long lived taxa. Reproduction through various techniques including ovovipartiy and asexual reproduction. Tegumental respiration and permanent attachment. Filter feeders of microorganisms or detritus.	Porifera and Bryozoa
B	Medium to large sized crawlers or burrowers. Ovoviparity as the main reproductive technique. Respiration and dispersal are aquatic. Various life cycles numbers and duration per year.	Crustacea, Hirudinea and Bivalvia
C	Medium sized taxa. Monovoltine. Crawlers that reproduce through either cemented eggs or clutches. Mainly predators or shredders.	Turbellaria, Plecoptera, Trichoptera and Diptera
D	Homogenous group of medium to large taxa. Semivoltine with a long life cycle duration. Adults are strong fliers with excellent dispersal capabilities.	Odonata
E	Small and medium sized. Aquatic respiration. Short-lived and uni- or plurivoltine.	Crustacea, Diptera, Trichoptera, Gastropoda.
F	Medium sized. Monovoltine organisms that produce egg masses or clutches. Aquatic respiration. Eggs or larvae have phase of quiescence to avoid adverse conditions.	Trichoptera, Ephemeroptera and Plecoptera
G	Small to medium size. Aerial respiration. Shredders or piercers. A range of life cycle durations and number of reproductive cycles per year.	Coleoptera and Heteroptera
H	Multivoltine. Burrowers or occupying the interstitial space. Feed on fine detritus and/or microorganisms	Oligochaeta

Table 6.3: Ecological traits of taxa corresponding to the classification scheme of Usseglio-Polatera et al. (2001).

Ecological trait category	Description	Typical taxa
A	Taxa typical of cold waters. Use coarse mineral substrate or vegetation. Oligotrophic.	Trichoptera, Plecoptera, Gasteropoda, Ephemeroptera and Turbellaria.
B	Lowland streams. Taxa considered to be oligosaprobic and eurythermic. Same substrate preferences as A	Trichoptera, Plecoptera and Coleoptera.
C	Colonise banks in lowland rivers. Mineral substrate, from cobbles to sand or in vegetation. Slow to medium current velocities. Oligo – to Beta-mesosaprobic that prefer oligotrophic or mesotrophic habitats.	Trichoptera, Coleoptera and Ephemeroptera.
D	Banks or side arms of lowland rivers. Wide substrate preferences. Oligotrophic or mesotrophic habitats. Mainly eurythermic and oligo- to beta-mesasaprobic.	Diptera and Trichoptera.
E	Eurythermic or thermophilous. Beta to gamma mesosaprobic. Wide substrate preferences. Lentic riverine habitat and lakes and ponds.	Oligochaeta, Bivalvia, Ephemeroptera and Trichoptera.
F	Found in all habitat types. Traits intermediate between E and G. Live in lentic or stagnant habitats, especially macrophytes.	Trichoptera, Diptera, Odonata, Coleoptera, Heteroptera, Gastropoda and Trichoptera
G	Taxa typical of stagnant waters outside river systems. Eurythermis and beta to gamma mesosaprobic. Macrophytes, organic detritus or mud as a habitat.	Coleoptera, Odonata, Heteroptera, Diptera and Trichoptera.

6.2.5: Comparison of Matrices

Relationships between community structure based on both taxonomic and trait-based descriptions were examined using a series of Mantel tests (Quinn and Keough, 2002). In each case the null hypothesis was that there was no correlation between the distance matrices of taxon or trait-based community structure between years – i.e. sites that were most similar to each other in 1979 were not those that were most similar in 2007.

To examine spatial structure in communities, easting and northing's of site locations were used to calculate Euclidean distances between sites across the river system. This measure represents the straight line distance between points. Using a Digital Rivers Network the Reticular distance between sample points was also calculated using the "Matrix Distance" Visual Basic script in ArcGIS developed by Dussault (2009). Reticular distance is a measure of distance between points along the river network and as such may provide a different insight into the relationships between points than that provided by Euclidean distance.

For each year, spatial relationships between sites were examined using Mantel tests on the correlation between the Euclidean or Reticular distance matrix and taxon or trait-based community structure. Significant relationships indicate spatial structure within the system.

6.2.6: Biological Indices

Finally, biological indices that indicate the sensitivity of taxa to organic pollution were used to examine whether this measure could plausibly be linked to differences in communities between 1979 and 2007. Firstly, the taxa present in both 1979 and 2007 were assigned their Biological Monitoring Working Party (BMWP) Score, which is indicative of their sensitivity to organic pollution, and the two years compared to examine whether there has been a net loss or gain of sensitive taxa between the years. Secondly, for each site an Average Score Per Taxon (ASPT) was calculated by summing the total of the BMWP scores for each taxon present at a site and then dividing this figure by the total number of taxa used. ASPT values below four are generally regarded as indicating poor water quality.

6.3: Results

Taxon richness was highest in 2007, with an average of 17.9 taxa per site, compared with 7.48 in 1979. Lowest taxon richness in 2007 was at a site in Ewden Beck (Figure 6.1; site FW298) where 6 taxa were sampled, this contrasts with 1979 where at three sites only a single taxon was found (Figure 6.1: sites FW114, FW173, FW321). Maximum taxon richness in 2007 was within the Porter Brook (Figure 6.1; site FW131) where 26 taxa were sampled contrasting with 1979 where a maximum of 17 taxa where found at a site on the Rivelin (Figure 6.1; site FW136).

A paired-sample t-test indicated a significant difference in taxon richness between the two sampling years (paired-sample t-test, $t = 8.50$, d.f.= 24, $p < 0.001$). In order to address the potential effect of sampling intensity, as discussed above, taxon richness in 2007 was standardised using the number of individuals sampled in 1979 (Table 6.1), and taxon richness between years was compared. For this 'standardized richness' there was no significant difference in taxon richness between years (paired-sample t-test, $t = 1.28$, d.f. = 24, $p > 0.05$).

Spatial patterns in taxon richness were examined using a series of Mantel tests in order to identify whether there was consistency in taxon richness at sites between years. Figure 6.2a and 6.2b illustrate the sites ranked using unadjusted taxon richness for both 1979 and 2007. There was no evidence of similarity in the spatial pattern of taxon richness when comparing the two years based on these data (Mantel test $r = 0.08$, $p > 0.05$). However, using the 2007 standardised values for taxon richness, there was evidence of a strong relationship in the spatial pattern of taxon richness between years (Mantel test, $r = 0.67$, $p < 0.001$). Spatially, there was evidence of structuring of taxon richness in 2007 based on both Euclidean (Mantel test, $r = 0.21$, $p < 0.05$) and Reticular (Mantel test, $r = 0.21$, $p < 0.05$) distance between sites, indicating sites that are closer together have similar taxonomic richness. This spatial structuring was also present using standardised taxon richness for both Euclidean (Mantel test, $r = 0.41$, $p < 0.001$) and Reticular (Mantel test, $r = 0.30$, $p < 0.01$) distance. There was no evidence for this relationship based on either Euclidian (Mantel test, $r = -0.06$, $p > 0.05$) or Reticular (Mantel test, $r = 0.06$, $p > 0.05$) distance in 1979.

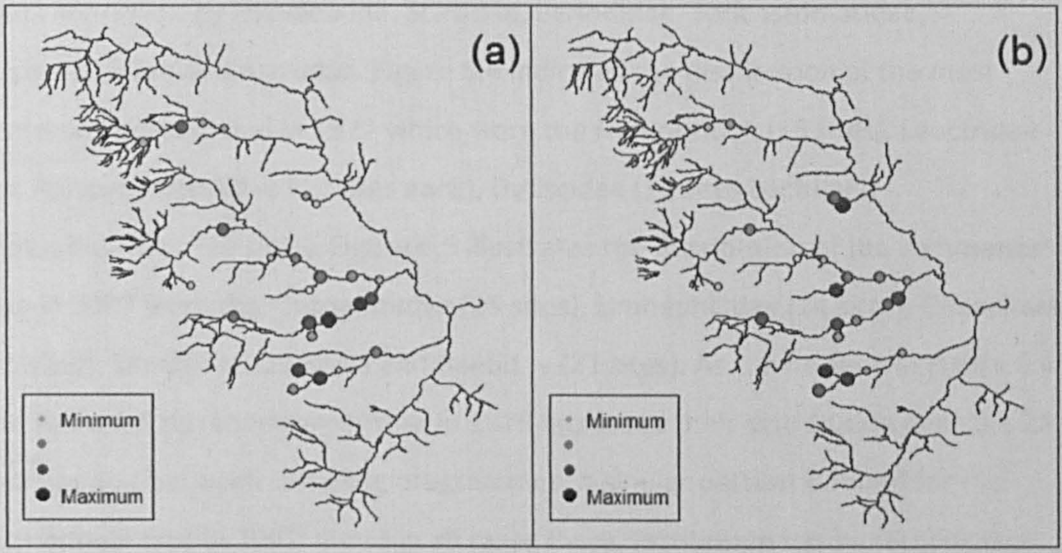


Figure 6.2: Taxon richness in (a) 1979 and (b) 2007 with size and colour of circles proportional to the rank of taxon richness within years.

Sites sampled in 2007 were on average more similar in composition (average β_{Jac} 0.56) than those in 1979 (average β_{Jac} 0.87). Using standardised 2007 data there was an increase in dissimilarity between sites (β_{Jac} 0.67) although this remained less than the dissimilarity between 1979 sites. There was no evidence that the pattern of dissimilarities between 1979 and 2007 communities were consistent between the years (Mantel test, $r = 0.24$, $p > 0.05$). Community composition was found to be spatially structured in 2007 with Mantel statistics indicating a positive relationship based on both Euclidean (Mantel test, $r = 0.34$, $p < 0.05$) and Reticular (Mantel test, $r = 0.30$, $p < 0.05$) distances. When rare taxa (present in an abundance of 5 or less in the 2007 samples) are removed the strength of this relationship increases for Euclidean (Mantel test, $r = 0.41$, $p < 0.001$) and remains the same for Reticular distance (Mantel test, r = 0.30, p <0.01). This result indicates that sites closer together have more taxa in common. There was no evidence for this relationship in 1979 (Euclidean - Mantel test, r = 0.13, p >0.05; Reticular - Mantel test, r = 0.08, p >0.05).

A comparison of taxa by rank order of occurrence across each of the 25 sites between 1979 and 2007, indicated a shift in the most commonly sampled taxa between years (Figure 6.3; Spearman's Rho = 0.26, p <0.05). Similarities between years are driven by Hydracarina, Scirtidae, Perlodidae, Sericostomatidae, Rhyacophilidae and Asellidae. Figure 6.4 indicates the distribution of the most common taxa sampled in 1979 which were the Nemouridae (15 sites), Leuctridae and Polycentropodidae (12 sites each), Dytiscidae (11 sites each) and Leptophlebiidae (10 sites). Figure 6.5 illustrates the distribution of the commonest taxa in 2007 were the Chironomidae (25 sites), Limnephilidae (24 sites), Oligochaeta (23 sites), Simuliidae (21 sites) and Baetidae (21 sites). As can be seen in Figure 6.4 and Table 6.4 the commonest taxa in 1979 increased their distribution over the 28 years period between sampling programmes. A similar pattern is found for the commonest taxa in 2007 where in all cases there distribution has increased since 1979. The Simuliidae represent a previously unrecorded taxon.

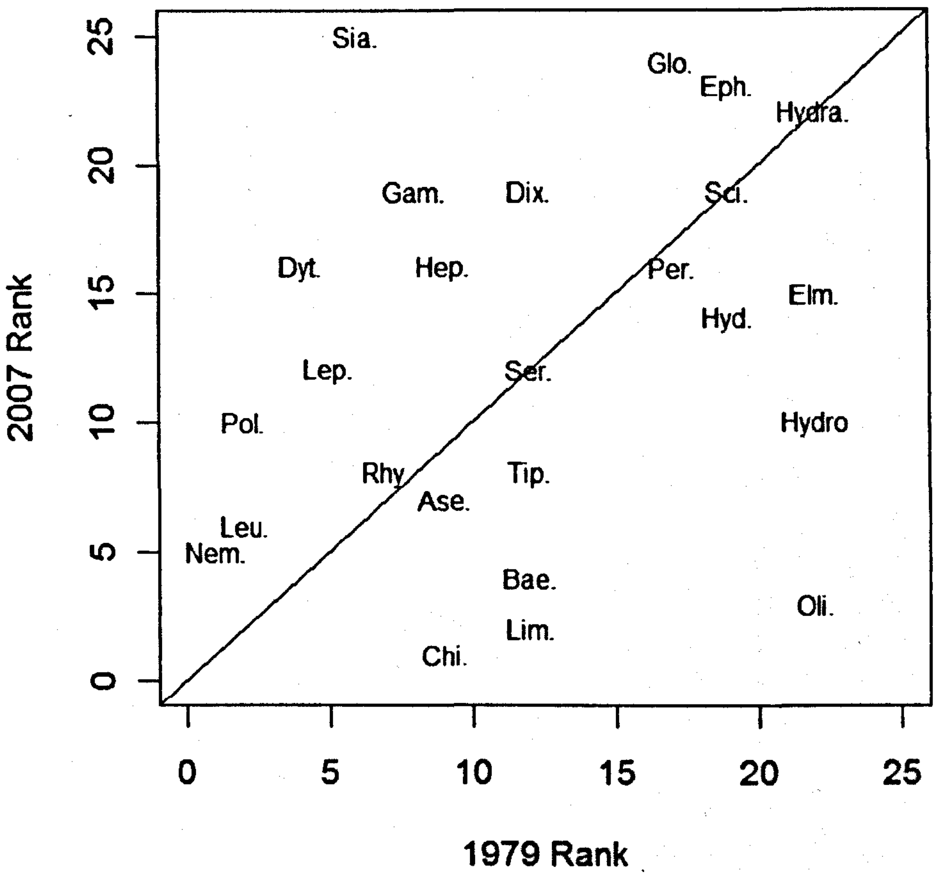


Figure 6.3: Differences between ranks of taxa present in both 1979 and 2007 (1 = taxa present at highest number of sites). Taxonomic identity is indicated by the first three letters of the family names from Table 6.4 for most taxa. However, "Hyd." indicates Hydrophilidae, "Hydra." indicates Hydracarina and "Hydro." indicates Hydropsychidae.

Table 6.4: Taxa sampled during both time periods indicating the number of sites at which they were sampled and the rank of this value (1 = occurrence at most sites).

Family	No. of sites		Rank no. of sites	
	1979	2007	1979	2007
Asellidae	5	18	10	7
Dytiscidae	11	11	4	13
Elmidae	1	12	14	12
Hydrophilidae	2	13	13	11
Scirtidae	2	7	13	15
Chironomidae	5	25	10	1
Dixidae	4	7	11	15
Tipulidae	4	17	11	8
Baetidae	4	21	11	4
Ephemeridae	2	4	13	18
Heptageniidae	5	11	10	13
Leptophlebiidae	10	14	5	10
Gammaridae	7	7	8	15
Hydracarina	1	6	14	16
Sialidae	9	1	6	21
Leuctridae	12	19	2	6
Nemouridae	15	20	1	5
Perlodidae	3	11	12	13
Glossosomatidae	3	2	12	20
Hydropsychidae	1	15	14	9
Limnephilidae	4	24	11	2
Polycentropodidae	12	15	3	9
Rhyacophilidae	8	17	7	8
Sericostomatidae	4	14	11	10
Oligochaeta	1	23	23	3
Sphaeriidae	3	-	12	-
Hydraenidae	1	-	14	-
Limoniidae	7	-	8	-
Crangonyctidae	4	-	11	-
Hydrobiidae	6	-	9	-
Lymnaeidae	8	-	7	-
Planorbidae	8	-	7	-
Veliidae	6	-	9	-
Erpobdellidae	3	-	12	-
Glossiphoniidae	3	-	12	-
Goeridae	2	-	13	-
Dendrocoelidae	1	-	14	-
Ceratopogonidae	-	18	-	7
Empididae	-	19	-	6

Table 6.4 continued.

Family	No. of sites	Rank no. of sites	Family	No. of sites
	1979	2007		1979
Psychodidae	-	7	-	15
Simuliidae	-	21	-	4
Caenidae	-	2	-	20
Ephemerellidae	-	2	-	20
Capniidae	-	8	-	14
Chloroperlidae	-	4	-	18
Brachycentridae	-	3	-	19
Ecnomidae	-	2	-	20
Hydroptilidae	-	1	-	21
Lepidostomatidae	-	4	-	18
Odontoceridae	-	5	-	17
Philopotamidae	-	2	-	20
Psychomyiidae	-	12	-	12

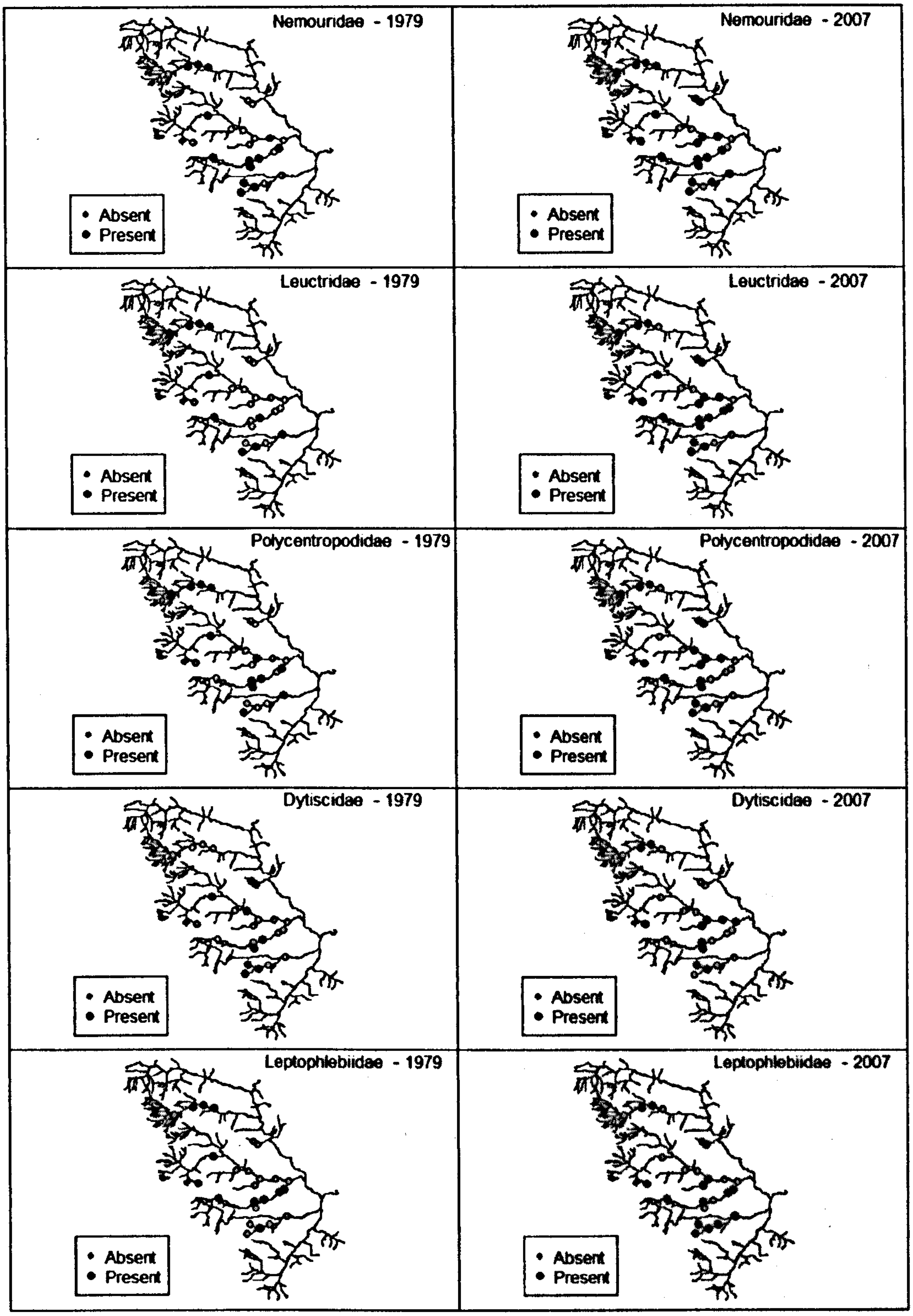


Figure 6.4: The distribution of the 5 most common taxa in 1979 and their corresponding distribtions in 2007.

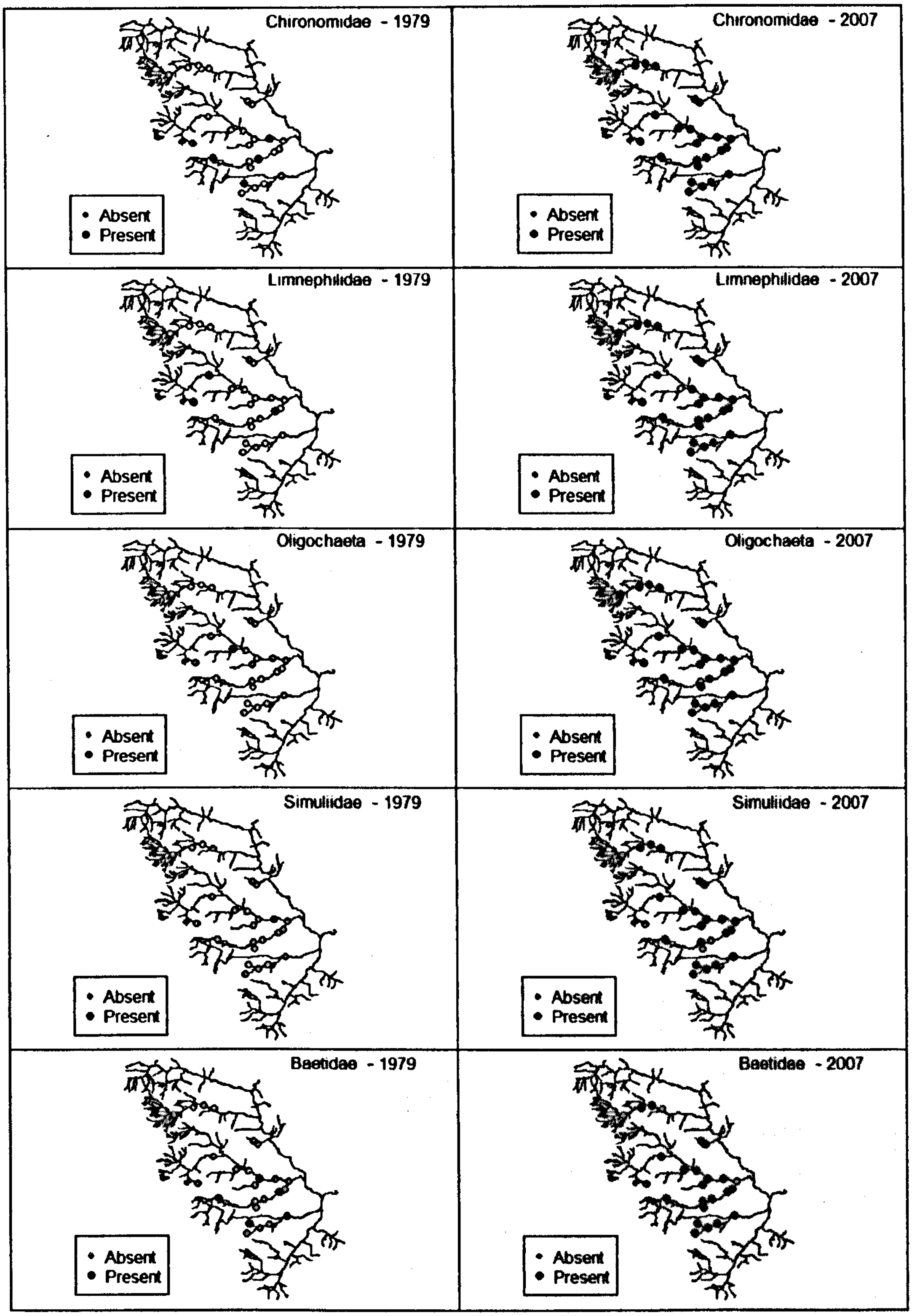


Figure 6.5: The distribution of the 5 most common taxa in 2007 and their corresponding distribtion in 1979.

Table 6.4 also indicates taxa unique to either 1979 or 2007 and indicates the loss of 12 taxa since 1979 and the introduction of 15 taxa previously not sampled. Figure 6.6 illustrates the difference between major groupings of taxa indicating the addition of 7 Trichoptera, 2 Ephemeroptera and 2 Plecoptera, and the loss of 3 Gastropoda, 2 Hirudinea and Bivalvia. Loss of the Gastropoda and Bivalvia can be attributed in part to the method of sampling which in 1979 recorded presence based on empty shells, as indicated in the comparison of methods.

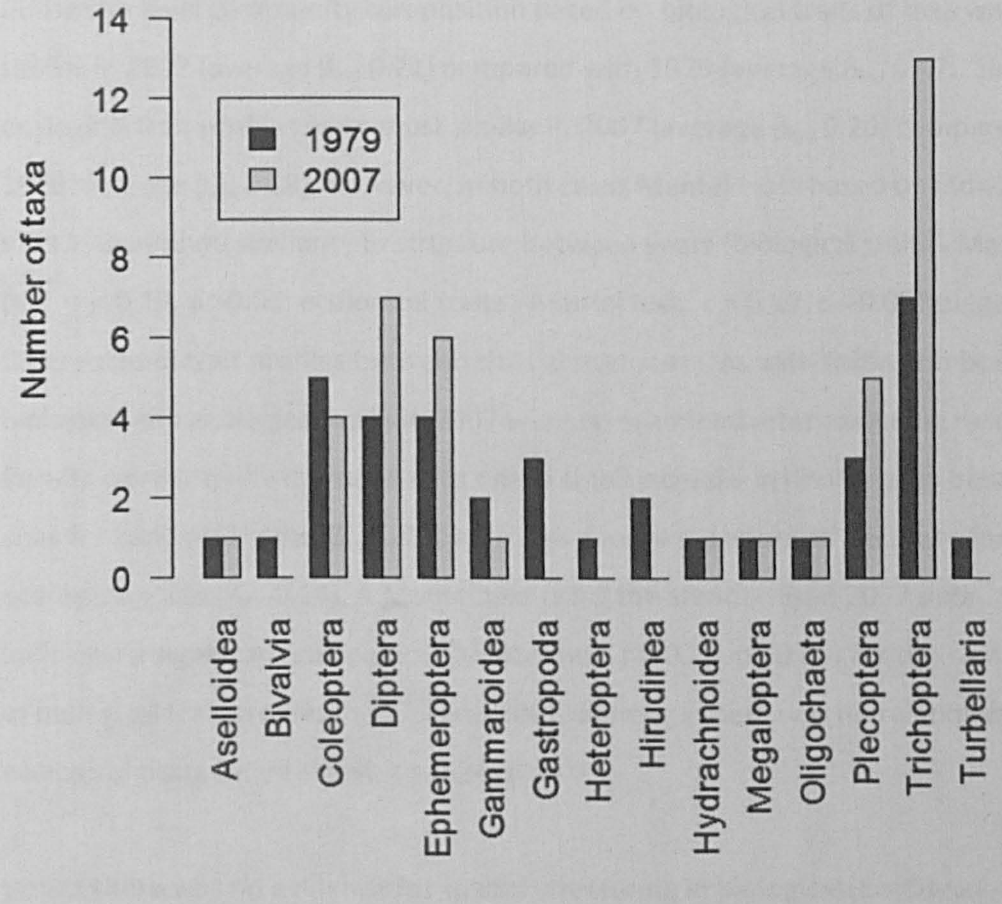


Figure 6.6: Taxa present in 1979 and 2007 summarised by major taxonomic groups.

Community composition was also described using an amalgamation of taxa into trait groups that describe similarity based on biological and ecological features.

Across all sites, the occurrence of biological and ecological trait groups were examined by counting the number of sites where each trait group was recorded and Chi-square statistics used to examine differences between years. This analysis indicated that for both biological (Chi-Sq = 2.423, d.f. = 5, $p > 0.05$) and ecological traits (Chi-Sq = 2.572, d.f. = 6, $p > 0.05$) there was no significant difference between years. Commonest biological traits were C, E, and F that are indicative of medium sized taxa that employ a broad range of biological strategies, as described in Table 6.2. In the case of ecological traits in both 1979 and 2007 categories C and D were commonest. These are indicative of taxa of lowland rivers with meso- to oligotrophic conditions on a range of substrate types, as described in Table 6.3.

At the site level community composition based on biological traits of taxa was most similar in 2007 (average β_{kul} 0.21) compared with 1979 (average β_{kul} 0.47). Similarly, ecological trait profiles were most similar in 2007 (average β_{kul} 0.20) compared with 1979 (average β_{kul} 0.38). However, in both cases Mantel tests based on individual sites indicated no similarity in structure between years (biological traits - Mantel test, $r = 0.19$, $p > 0.05$; ecological traits - Mantel test, $r = 0.17$, $p > 0.05$) suggesting differences in trait profiles between sites across years. As with taxon composition, biological and ecological traits in 2007 were re-examined after excluding rare taxa. Results were broadly the same with only a small increase in dissimilarity between sites for biological traits (β_{kul} 0.23) although a more pronounced decrease for ecological traits (β_{kul} 0.24). A Mantel test using the standardised 2007 data indicated a significant correlation (Mantel test, $r = 0.23$, $p < 0.05$) for the comparison of biological trait profiles in 1979 and 2007, although there was no relationship for ecological traits (Mantel test, $r = 0.24$, $p > 0.05$).

Whilst there was no evidence for spatial structuring in biological (Euclidean - Mantel test, $r = 0.06$, $p > 0.05$; Reticular - Mantel test, $r = 0.04$, $p > 0.05$) and ecological traits (Euclidean - Mantel test, $r = 0.10$, $p > 0.05$; Reticular - Mantel test, $r = 0.10$, $p > 0.05$) for 1979 data, there was evidence of a relationship between ecological traits and Euclidean distance (Mantel test, $r = 0.23$, $p < 0.05$) and biological traits and Reticular distance (Mantel test, $r = 0.19$, $p < 0.05$) in 2007. However, there was only marginal evidence of a relationship between biological traits and Euclidean distance (Mantel test, $r = 0.19$, $p = 0.05$) and ecological traits and Reticular distance (Mantel test, $r = 0.20$, $p = 0.05$) in 2007. When rare taxa were removed from the 2007 analysis there was a significant relationship based on all distance measures for both biological (Euclidean - Mantel test, $r = 0.18$, $p < 0.05$; Reticular - Mantel test, $r = 0.20$, $p < 0.05$) and ecological (Euclidean - Mantel test, $r = 0.26$, $p < 0.05$; Reticular - Mantel test, $r = 0.25$, $p < 0.05$) traits.

On average, dissimilarity between communities at each site between years was high (β_{Jac} 0.85) indicating significant changes in composition between years (Figure 6.7a). Variability was lower based on dissimilarity in biological (β_{kul} 0.35; Figure 6.7b) and ecological (β_{kul} 0.33; Figure 6.7c) traits suggesting some stability in this measure of the community. Mantel tests where used to compare whether there was evidence of spatial structure in the dissimilarity matrices. For dissimilarities between taxa (Euclidean - Mantel test, $r = 0.17$, $p=0.05$; Reticular - Mantel test, $r = 0.18$, $p = 0.05$; Fig 7a), biological traits (Euclidean - Mantel test, $r = -0.09$, $p > 0.05$; Reticular - Mantel test, $r = -0.01$, $p > 0.05$; Figure 7c) and ecological traits (Euclidean - Mantel test, $r = 0.17$, $p = 0.05$; Reticular - Mantel test, $r = 0.18$, $p = 0.05$; Figure 7c) it was concluded that there was little evidence of spatial structure.

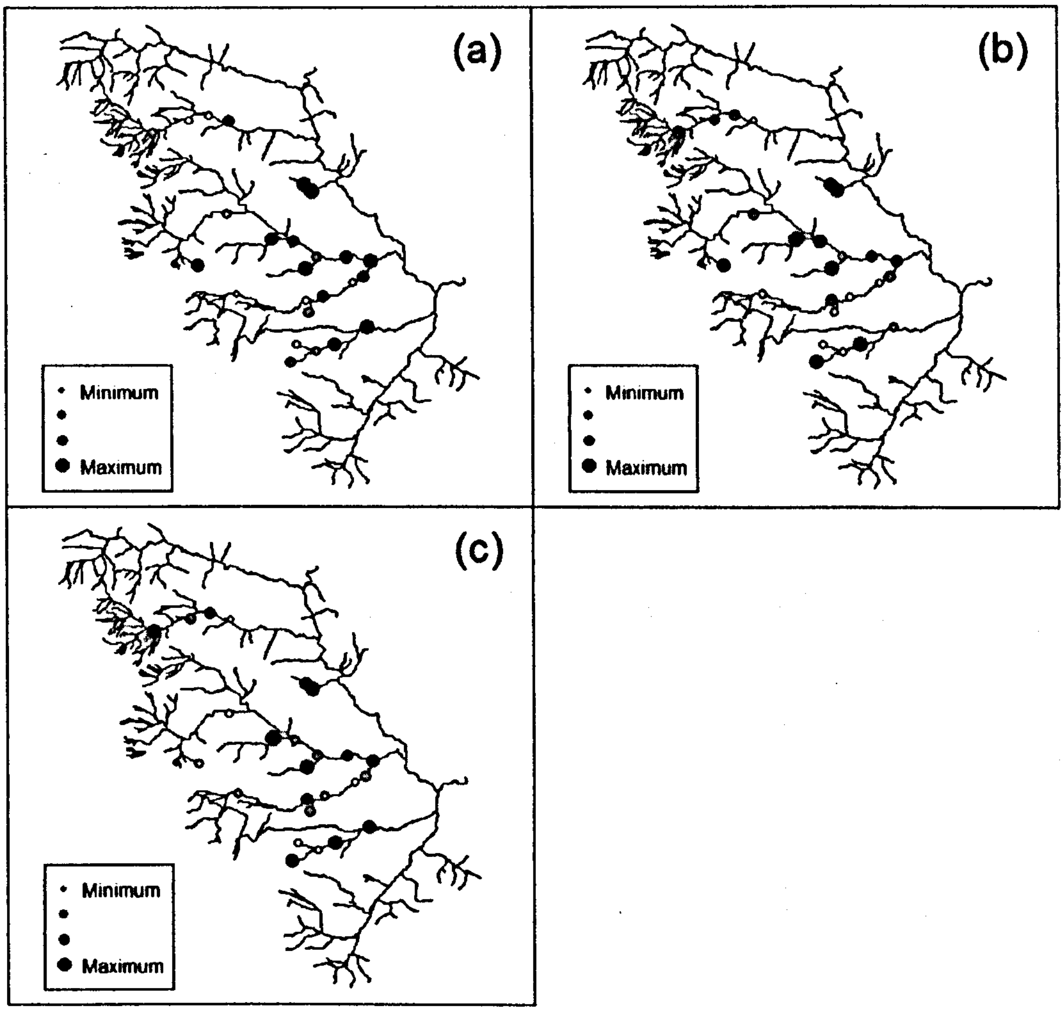


Figure 6.7: Dissimilarities in communities based on (a) community structure, (b) biological traits and (c) ecological traits. Colour and size of points illustrates rank dissimilarity for each measure.

Finally, to compare the difference in ecological quality of the river systems between years ASPT was calculated at each site for both 1979 and 2007 (excluding those sites at which only a single taxon was sampled). Average ASPT in 1979 across all sites was 5.04 compared with 5.25 in 2007. A paired sample t-test indicated no significant difference between ASPT in 1979 compared with 2007 (paired-sample t-test, $t = 1.845$, d.f. = 24, $p > 0.05$). However, in order to examine changes to the broader pool of species present within the river system a comparison based on the BMWP values of taxa unique to each year was made. This indicated that there was a significant difference between the average score of those found solely in 1979 (average 4.72) compared with those unique to 2007 (average 8.5) (paired-sample t-test, t = -3.89, d.f. = 16, p <0.001) suggesting an increase in the number of sensitive taxa.

6.4: Discussion

The data on freshwater invertebrate distributions and diversity in 1979/80 provide a valuable resource against which more recent sample data can be compared, providing the rare opportunity to contrast site-specific data, at multiple sites, across a time period of almost 30 years. The results are complicated by potential issues of sample comparability, but nonetheless suggest some significant and ecologically interesting changes over this period. The implications of the methodological issues are firstly discussed to highlight those results that warrant caution in their interpretation.

The comparison of the methods presented above (section 6.2.3) suggests the analysis of patterns within years is likely to be robust. In 1979, consistency of staff conducting the survey suggests that similar collection processes would be used at sites, allowing valid comparisons to be made between sites across the river system. As the stated goal was to assess the freshwater biodiversity of the Sheffield Metropolitan District it can be assumed that the method they employed was designed to fulfil this aim. Therefore, in the current study comparisons of patterns within years (highest/lowest taxon richness), and comparisons between years (spatial patterns in measures of taxon richness, similarity of composition based on taxonomic and trait descriptions) are valid as both sampling techniques will have captured trends in these measures and conclusions are reliant on relative rather than absolute values. Similarly, differences in area wide assessments (most common taxa, unique and shared taxa between years) can be considered to be robust. In all cases whilst absolute values may vary, both methods should capture a description of the biological communities present. However, caution must be applied to results examining differences over time at the site level. Clearly, the 2007 sampling regime captured many more individuals, and whilst standardising samples based on rarefaction or the exclusion of rare taxa provides one way of exploring what adjustment of the data to match on abundance might look like, there is no good evidence to indicate that the comparison it generates is any more realistic than the unadjusted data. With these caveats in mind I now consider the findings from the study.

Taxon richness for the entire set of sites is broadly similar between both 1979 (37 families) and 2007 (40 families). At the site level, taxon richness was higher in 2007 than 1979 based on unadjusted data, though not when standardised based on 1979 sampling abundances. As such it was concluded that a difference in taxon richness between years arises primarily due to sampling method. Similarly, whilst there is little correlation in the spatial pattern of taxon richness between years for sites based on unadjusted taxon richness (Mantel test on dissimilarity matrices), the strong positive correlation using adjusted 2007 data suggests that sites with the highest/lowest taxon richness in 1979 continued to be amongst the sites with the highest/lowest taxon richness in 2007. This relationship is shown in Figure 6.2 to be primarily driven by high taxon richness in the Porter Brook and Rivelin and low richness in Ewden Beck. Based on Euclidian and Reticular distances between sites, taxon richness in 2007 was found to be spatially structured with sites closest together exhibiting similar values, a pattern absent in 1979. It can be concluded that whilst there is little evidence to suggest a systematic change in taxon richness over time, there are stronger spatial relationships between communities within the river system in 2007 than in 1979. This spatial relationship between taxon richness in 2007 extends to community structure where there was also evidence that communities are more similar to each other in 2007 than 1979. Taken together these results may be indicative of a strengthening of linkages between communities. The loss of such linkages often results from anthropogenic stress on systems that restrict dispersal (Malmqvist, 2002; Urban et al., 2006) limiting immigration and restricting the ability of the community to recover from disturbance (Palmer et al., 1996; Fagan, 2002).

When rare taxa were excluded from the 2007 communities there was an increase in dissimilarity between sites suggesting that these taxa may play an important role in driving similarity between communities. As the measure of community dissimilarity used in the present study was based on presence or absence, similarity may arise due to this large scale dispersal of taxa across the whole system. This is consistent with known species abundance patterns that are frequently described by a log normal model (Cao et al., 1998) where there is dominance by a few taxa present in high abundances and then a long "tail" of rare taxa. Whilst, in the current study, the identity of the most dominant taxa at sites may change spatially, dispersal within the water column leads to a wide distribution of taxa throughout the system leading to their presence in low numbers across many sites (Brittain and Eikeland, 1988; Petersen et al., 2004) and driving the similarity between communities observed in 2007. Whilst this provides a statistical mechanism for the increased similarity between sites in 2007, features of the taxa and the known history of the river systems suggest a number of possible ecological reasons for these patterns.

When rare taxa are excluded from the 2007 data, sites are still more similar than those in 1979 and exhibit spatial structure absent in 1979, suggesting structuring across the catchment. Additionally, between 1979 and 2007 there was a marked change in the identity of the taxa most commonly sampled (Figure 6.4 and 6.5), together with the addition of 15 previously unrecorded taxa. Previous studies suggest communities might be expected to persist through time unless there is a significant alteration in environmental conditions (e.g. Johnson et al., 1994; Robinson et al., 2000; Brewin et al., 2000; Humphrey et al., 2000). An examination of the previously unrecorded taxa in terms of their sensitivity to organic pollution provides one possible mechanisms for these observed changes. Whilst results indicated no significant difference between ASPT values across years (1979 average = 5.04, 2007 average = 5.25), a comparison of the taxa found across all sites suggests that there has been an increase in the number of taxa sensitive to pollution. The average BMWP score of those taxa unique to 2007 was 8.5, compared with an average value of 4.72 for those taxa unique to 1979. Whilst sensitive taxa were found in 1979, it is possible that ecological characteristics such as the ability to utilise a wide range of habitats (e.g. the Heptageniidae and Leptophlebidae) could account for their presence (Usseglio-Polatera and Tachet, 1994). Taxa unique to 2007 (e.g. Ephemerellidae, Chloroperlidae) often have more specific habitat requirements (Tachet et al., 1994) and as such may be more sensitive to system wide disturbance as they may lack flexibility to occupy refuges.

BMWP scores represent one method of linking the features of taxa with the prevalent environmental conditions. Such trait-based approaches have received considerable attention in recent years as they provide a way to examine changes in the aquatic environment that can be linked directly to habitat conditions (Davies et al., 2000) and may represent a differing perspective to processes acting on the community (Usseglio-Polatera et al., 2001). The idea stems from the Habitat Templet (Southwood, 1977; Townsend and Hildrew, 1994) which states that significant difference in trait profiles, both spatially and temporally, can be attributed to differences in the physical environment. In the current study community structure was examined based on groupings of taxa according to their wider biological and ecological traits (as described in Usseglio-Polatera et al., 2001). Previous studies (Charvet et al., 2000; Archaimbault et al., 2005; Beche et al., 2006; Bonada et al., 2007) have demonstrated that through time such trait based descriptors of structure are often more stable than taxon based measures. Results from the current study are consistent with these findings with lower values for dissimilarity both between and within years when compared to taxonomic composition. Although the dissimilarity measures are not directly comparable, as they are based on differing statistical techniques, the community based metric was extremely high (average β_{Jac} 0.85) suggesting almost complete dissimilarity between communities. This contrasts with the trait based measures which had values that were indicative of lower total variation (Biological traits β_{kul} 0.35; Ecological traits β_{kul} 0.33). However, the differing dissimilarity measures used, and differences in the descriptions of the communities (i.e. a consistent number of traits vs. a changing number of taxon) make this comparison difficult. More compelling evidence for a difference between years arises through the presence of spatial structure in trait based measures of the community in the 2007 samples when rare taxa are removed, as this suggests structuring of environmental conditions across the river system absent from the 1979 samples. This would also be consistent with a general improvement in water quality since 1979.

Whilst an improvement in water quality represents a plausible explanation for difference in communities between years it is interesting to question when this improvement may have taken place. There is known to be a time lag between change in the physicochemical environment and recovery of the community (Wiens, 2002) with in some cases community structure being more accurately predicted by conditions up to 40 years previous. For example Harding et al. (1998) demonstrated that across 24 streams in the United States, whole watershed land use in the 1950s was a more accurate predictor of community structure than contemporary (1990) conditions. In the current study, differences in community structure between 1979 and 2007 may not represent a large change in habitat or water quality over the 28 year period between studies, but rather may be representative of points on a trajectory of change reflecting longer term historic conditions in the system, perhaps dating back to the earlier part of the twentieth century. The speed of recovery of a system is known to be significantly influenced by the availability of local colonists, with the regional species pool playing a central role in re-colonisation following disturbance (Naeslund and Norberg, 2006). The large scale historic degradation of river systems across Sheffield Metropolitan District suggests that this regional species pool may have been significantly impacted in the past reducing the number of such colonists. Changes in community structure between 1979 and 2007 may reflect the slow arrival of taxa from other river systems with similarity between communities in the 2007 study reflecting the increased length of time over which colonisation has had to occur.

In conclusion whilst methodological problems undoubtedly exist with the data, findings from the current study indicate that useful comparisons can still be made between years, and that differences may be attributable to biological and ecological mechanisms. The use of both community and trait based approaches for the examination of temporal variability suggest that there has been a fundamental shift in the community that can be attributed to changes in the physicochemical environment. These differences could be accounted for by either changes in conditions over the last 27 years, or lag in the recovery of communities as recolonisation takes place following improvement in habitat conditions in the years prior to the 1979 study.

CHAPTER 7: GENERAL DISCUSSION

7.1: Aim of thesis and principal findings

Understanding change in communities is one of the fundamental challenges in ecology and has wide implications for the preservation of biodiversity and the protection of ecosystem function. The central aim of this study was to examine temporal variability of lotic macroinvertebrate communities in order to; (a) identify spatial patterns; (b) examine the relationship with environmental factors; and (c) examine the role of community composition. The two datasets utilised in the current study are amongst the most spatially and temporally diverse ever used to examine long-term temporal variability of lotic macroinvertebrate communities. The first dataset represents the extent of a national water quality monitoring programme and was utilised in Chapters 3 -5. The second dataset represents historic data collected almost 30 years ago across a river system in central England and was utilised in Chapter 6.

The first objective of the current study was to identify spatial structure in community temporal variability as such structure will reflect important processes acting on communities (Li et al., 2000). For example similar levels of temporal variability between communities across large areas would indicate that large scale factors (i.e. climate, geology) influence patterns of variability. Alternatively, more complex spatial patterns would highlight the importance of local variation in the environment. The analysis described in Chapter 3 demonstrated structure in macroinvertebrate temporal variability across a range of scales from the regional to local, suggesting that temporal variability is not controlled by solely large or small scale processes, but rather is influenced by a range of factors that are operating at the site, catchment and regional level. Although this finding conforms in part to the view of the river system as a hierarchy of interacting processes (Frissell et al., 1986) where large scale factors exert "top-down" pressure on communities however, it indicates that local variation in conditions has the potential to alter the influence of these large scale factors, either by ameliorating the pressure or enhancing it, resulting in differing levels of temporal variability between sites even in close proximity.

Using a range of environmental descriptors derived from GIS and monitoring data, the second objective was to examine the relationship between these environmental factors and temporal variability of macroinvertebrate communities (Chapter 4). Regional scale factors relating to climate (low rainfall; high annual temperature range) and human land use (coniferous woodland; urban; arable farming) were demonstrated to lead to increased temporal variability. These findings agreed with results from previous studies (e.g. Weatherley and Ormerod, 1990; Townsend et al., 1997; Robinson et al., 2000; Gibbins et al., 2001; Diaz et al., 2008) suggesting that climate and land use are a universal influence on macroinvertebrate communities in lotic systems. At the site scale, which represents the local environment of the macroinvertebrate communities, both increasing substrate and channel variability lead to increased temporal variability within communities. Both variables are indicative of habitat stability with previous studies emphasising the importance of substrate stability, in particular, in relation to community temporal variability (e.g. Townsend et al., 1997; Gibbins et al., 2001; Roy et al., 2003).

In contrast to the majority of previous studies, the spatial extent of the monitoring data used in the current study allowed the examination of levels of temporal variability across geographic areas that encompass large environmental gradients. Results suggest that the relationship between temporal variability and environmental drivers was in many cases non-linear with a clear threshold beyond which variability changed sharply. Such a relationship has rarely been demonstrated before in freshwater lotic systems in relation to temporal variability of communities (Collier, 2007) although a number of studies have demonstrated changes in community structure above certain threshold values (e.g. Doak et al., 1999; Paul and Meyer, 2001; Roy et al., 2003).

The next objective was to examine whether some communities owing to their constituent taxa were more variable than others (Chapter 5). Community composition was shown to have a spatial structure across England and Wales, as demonstrated in previous studies (e.g. Wright, 2000). This structure was a good indicator of the level of temporal variability, with communities to the south and west exhibiting less change through time than those in the centre of the country. This finding is plausibly linked to known gradients of environmental stressors across England and Wales (Chapter 2 - 3; Murphy and Davy-Bowker, 2005). Theory arising through the Habitat Templet indicates that such environment gradients will "filter" taxa based on the traits, selecting for those that ensure survival and reproductive success under a given set of conditions. Previous studies indicate that this may have two implications for temporal variability. Firstly, the traits selected for might themselves lead to temporal variability as they confer resilience properties to the taxa based on rapid dispersal and recolonisation (e.g. Beche et al., 2006; Lepori and Malmqvist, 2009). Secondly, communities may exhibit more temporal variability as they lack both resistance and resilience mechanisms with which to respond to disturbance.

Communities were clustered based on similarity of composition and it was demonstrated that some groupings of communities were more variable than others through time. Using the gradient of temporal change across all sites, the persistence of individual taxa within the community was examined to identify whether some taxa were more persistent than others. This analysis identified three groupings of taxa based on how persistent they were and whether their persistence changed along the gradient of temporal change of the communities in which they occurred. Persistent taxa (termed "High Frequency" in Chapter 5) were considered to possess contrasting sets of traits that conveyed either resistance or resilience to disturbance, ensuring long-term persistence within communities. Taxa that were not persistent (termed "Low Frequency" in Chapter 5) where characterised as possessing traits typical of mobile generalists and exhibited little persistence as a result of continual dispersal throughout the system. Finally, a group of taxa were identified possessing traits that made them potentially more susceptible to disturbance. These (termed "Variable Frequency" in Chapter 5) are present within lotic systems in life stages that are particularly susceptible to disturbance, and exhibit a reduced reproductive potential compared with other taxa. Communities dominated by these "Variable Frequency" taxa were highly persistent through time as they are indicative of minimally disturbed sites. This contrasts with communities where such "Variable Frequency" taxa were present in a low proportion of the community where temporal variability was driven by colonisation and local extinction of these taxa as they lacked strategies to ensure continual persistence. Few studies have examined the direct relationship between traits and long-term temporal variability in freshwater systems with Beche et al. (2006) representing a notable exception. However, although traits may not be considered explicitly, studies commonly describe differing levels of persistence amongst members of a community at a site and as such are implicitly reporting the same relationship described in this study. For example Bradt et al. (1999) demonstrated a significant decrease amongst the Trichoptera, *Psychomyia* and *Leucotrichia* over a 20 year period. Chessman (2009) demonstrated that from a total of 124 families, 33 exhibited increasing trends, 37 decreasing trends and 54 no significant trends over 13 years associated with climatic shifts. Similarly, other studies report a persistent "core" of taxa. For example Brown et al. (2006) demonstrated little variation in a core of 15 taxa within a community that the authors attribute to a common pool of well adapted taxa. In this way results from the current study are consistent with demonstrated patterns in communities over long temporal periods. However, results from the current study provide a new perspective by demonstrating the specific roles of key traits in governing persistence.

The final objective was to compare historic and contemporary patterns in macroinvertebrate communities, and to examine temporal changes over a 28 year period within parallel river systems within a single catchment. This was addressed in Chapter 6 using a dataset that incorporated a re-sampling of a 1979/1980 survey of freshwater sites in the Sheffield Metropolitan District. Although measures such as taxon richness and ASPT remained consistent between years there was evidence of a change in taxon identity. Communities were considered to be more similar in 2007 than 1979 and exhibited spatial structure absent from the 1979 data. Although previous studies of temporal variability have reported little variation between years (e.g. Richards and Minshall, 1992; Johnson et al., 1994; Bradt et al., 1999) such persistence or stability is associated with constancy of environmental conditions. Based on historic knowledge of the river systems (Firth, 1997) changes in community structure were suggested to be driven by historic improvements in water quality. As such, long-term variability in community structure is driven by changing environmental conditions, a finding consistent with other studies (e.g. Bradley and Ormerod, 2002; Daufresne et al., 2003).

In conclusion the current study has provided insights into the spatial patterns, environmental drivers and the role of community composition in controlling temporal variability in macroinvertebrate communities. I now consider the implication of these findings in a broader ecological context and from a management perspective. Finally, I consider the lessons from the two datasets used here for the implementation of effective long-term monitoring capable of yielding insights into the long-term temporal dynamics of communities.

7.2: Synthesis

One of the key features of this study is the spatial extent over which temporal change could be examined. In a review of previous studies relating to long-term temporal variability of lotic macroinvertebrate communities (Chapter 1; Table 1) it was demonstrated that with the exception of one study (Chessman, 2009) the scale and extent of the data used in Chapters 3-5 is unique. This is important as the spatial scale at which a phenomenon is examined has the potential to greatly influence the conclusion that is drawn about relationships (Farnsworth, 1998). As such, the present study provides a unique opportunity to examine the applicability of findings from other studies of similar communities over broader areas.

Based on the analysis of spatial patterns of temporal change, results from the current study confirm that a multiscale approach (Wu and Loucks, 1995), whereby the interaction of factors acting across spatial scales influence communities, is most appropriate in lotic systems. The complex spatial patterns (Chapter 3) present across the spatial extent of the data used in the current analysis clearly demonstrate that, although large scale regional processes do exert some influence on temporal variability, much of the variation in communities is driven by processes acting at the catchment, reach or site scales. Although such complex spatial patterns have not been demonstrated for temporal variability of lotic macroinvertebrate communities before, due to limits in the availability of either spatial or temporal data, it conforms to the view of the river system whereby differences at the regional, catchment, river and site scale lead to complex patterns in macroinvertebrate assemblage structure (Parsons et al., 2003). Results from the current study provide an interesting perspective on previous studies of temporal variability in lotic systems. Whilst it can be argued that a number of processes emerge as key drivers of temporal variability in previous studies (Chapter 1; Section 1.6), there remains considerable disagreement about which factors are most important. For example, many studies consider flow regime to be central (e.g. Statzner et al., 1988) whereas other find no relation with discharge patterns (e.g. Bradley and Ormerod, 2001). The analysis presented here suggests that rather than a single factor or group of factors exerting control over temporal variability, the relative importance of factors may differ on a site and study basis. For example whilst flow regime may be important for some communities, its influence may be overridden by land use pressures. This has considerable implications for biomonitoring and is discussed further in section 6.3.

Previous studies have also emphasised the importance of constancy of habitat for temporal persistence or stability within communities (e.g. Townsend et al., 1987; Richards and Minshall, 1992; Johnson et al., 1994; Scarsbrook, 2002). In the current study (Chapter 4) direct measures of environmental variability were demonstrated to have a positive relationship with increasing levels of temporal variability (i.e. temperature range, bed substrate, channel width). Furthermore, measures that are known to be associated with environmental instability (i.e. urban/arable/coniferous land use) were also demonstrated to be related to an increase in community temporal variability. As such, results from the current study support the statement that constancy of conditions is important.

Results indicating a strong positive correlation between trait and taxonomic based community variability (Chapter 5; Section 5.3.4) represent one of the few instances where temporal variability in trait-based measures have been considered over long temporal scales in freshwater systems. The strong relationship between temporal variability of traits and taxonomic based structure adds support to the view that the physicochemical environment can be considered as an overriding influence controlling the temporal variability of macroinvertebrate communities in river systems (Winterbourn, 1997). However, this result also provides further support for the ideas presented in the Habitat Templet, which whilst having a sound theoretical basis has proved difficult to test practically (see Doledec and Statzner, 1994; Townsend et al., 1997; Heino, 2005). Many of the difficulties in linking environmental conditions to the Habitat Templet have been attributed to trade-offs (McElravy et al., 1989; Chevenet et al., 1994; Townsend et al., 1997) where, for example, resilience traits (e.g. small body size, fast reproduction) confer a similar advantage as resistance traits (e.g. large body, streamlining), as taxa can be persistent through rapid recolonisation following disturbance, or by remaining *in situ*. McElravy et al. (1989) suggested that difficulties associated with creating the link between habitat and traits could be overcome by focusing on individual traits. Similarly, as has been argued in relation to the use of traits for biomonitoring, to effectively interpret the loss of taxa from a community it is essential to identify those traits that relate directly to disturbance (Statzner and Beche, 2010). The results presented in the current study suggest that stream macroinvertebrates possess a limited set of traits (i.e. high affinity for one life cycle per year; low affinity for attachment, burrowing or the use of the interstitial space; see Chapter 5) that make taxa susceptible to disturbance and lead to temporal variability. These traits might be suggested to represent "time sensitive" features of the taxa that influence long-term persistence and reproductive success. This contrasts with the remaining traits considered in the present analysis that might be considered to be more closely linked to short term survival at a site. Taxa that possess these time sensitive features are therefore good candidates for the assessment of environmental conditions through time.

Another interesting implication of the findings in Chapter 5 is that it suggests that regional scale selection for traits may play an important role in controlling temporal variability of communities. To arrive at a specific site taxa must pass through a series of filters at increasingly small spatial scales (Poff, 1997; Belyea and Lancaster, 1999; Chapter 1, Figure 1.1). As discussed above one of the principal drivers of change is colonisation by taxa that lack traits to ensure long-term persistence. A number of other studies (e.g. Winterbourn, 1997; Beche et al., 2006) have demonstrated that where there is a small regional species pool, composed of taxa with little trait diversity due to harsh regional conditions, communities are persistent and stable through time. In such instances, due to the limited regional fauna, stability arises as re-colonisation will inevitably be from the small pool of available taxa. By contrast lotic systems in England and Wales could be considered to have a relatively large regional species pool from which taxa can be drawn following disturbance. As was demonstrated in the current study in areas with low site diversity there may be high temporal variability due to immigration of new species from the large regional pool. If site diversity is high the chance of colonisation by previously unrecorded taxa decreases resulting in low temporal variability.

Temporal variability is an important consideration for the assessment of ecosystem function and provision of ecosystem services. Of central concern is reliability in the provision of such services (Naeem, 1998) where "reliability" is the probability that a system will provide a consistent level of performance over time. In the current study, although it was shown that some communities were more variable than others (Chapter 5) it would be interesting to examine the relationship between this community-based variability, and variability of aggregate measures indicating ecosystem function. Specifically are the core taxa (the "High Frequency" taxa) ensuring continued ecosystem function, or are such communities ecologically impaired? This question could be addressed through the use of composite indices relating to functional groups (e.g. Heino, 2005). The results of any such analysis could have important implications for the maintenance of goods and services that humans rely on (Millennium Ecosystem Assessment, 2005).

One of the most significant findings of the current study is that the response of communities to increasing anthropogenic disturbance (anthropogenic land use i.e.

arable/urban/conifer) does not appear to be linear (Chapter 4), but rather to be associated with certain thresholds above which there is a sharp increase in community variability. This finding is consistent with predictions of increased temporal variability of stressed communities (Odum, 1985) and is of concern as increasing variability is thought to be an indicator in potential regime shifts (Scheffer et al., 2001; Carpenter and Brock, 2006; Koch et al., 2009) whereby the properties of the system change to a new state. Future work could examine whether communities that exhibit the most variation are also those with impaired function.

7.3: Implications for Biomonitoring

Current assessment of ecological quality of lotic systems in England and Wales is based on a reference approach where sites are classified based on the difference between the observed community and the community that would be expected in the absence of anthropogenic stress (Wright, 2000). The assumption of this approach is that communities are persistent in the absence of anthropogenic disturbance (Richards and Minshall, 1992; Statzner et al., 1997; Bunn and Davies, 2000; Scarsbrook, 2002; Woodward et al., 2002; Milner et al., 2005).

Although results from the current study lend broad support to this statement by indicating that measures associated with anthropogenic disturbance of river systems (i.e. land use; see Chapter 4) lead to increases in the temporal variability of communities, large scale drivers not associated with anthropogenic activity (i.e. temperature range and mean annual rainfall) were also found to be amongst the most important predictors of temporal variability in lotic macroinvertebrate communities. This finding is consistent with a number of other studies that have demonstrated considerable temporal variability in communities even in pristine sites (e.g. Scarsbrook, 2002; Milner et al., 2006). This has two important implications for biomonitoring.

Firstly, temporal variability represents a difficulty in the establishment of reference conditions, the "expected" communities, against which communities sampled at other sites are compared. Such reference conditions are often based on the collection of 1 year's samples from a specific site, thought to be representative of the least anthropogenically disturbed community (Wright, 2000; Milner et al., 2005). The structure of this community (or more specifically an aggregated description of this community across a number of sites) is then used as a baseline "expected" community. When biomonitoring is carried out the "observed" community at a site that might be subjected to stress is compared with this "expected" community to assess ecological status. However, as the "expected" community was established using a limited temporal window it may not be representative of the community that would be present at some point in the future. For example in the current study I have demonstrated that climatic processes, which have a strong regional pattern, are related to temporal variability. Therefore reference conditions established in areas with a high annual temperature range, and low annual rainfall (demonstrated in Chapter 4 to lead to increasing temporal variability) might not be representative of the community structure over time. In essence, temporal variability driven by natural processes leads to a moving baseline of conditions on which communities are assessed.

The second implication of temporal variability is for the community being assessed, the "observed" community. Changes in the status of the community through time are assumed to be the result of anthropogenic disturbance. As was argued above large scale natural processes may influence temporal variability suggesting that variation may be natural in some cases and will vary regionally. Therefore it may be necessary to attach differing levels of significance to any observed variability in communities depending on this climate context. However, results from the current study indicating a complex spatial pattern of temporal variability (Chapter 3) suggest that temporal variability may present an even more complex problem. The implication of this multiscale control is that even within sites close together there may be considerable differences in the temporal variation of the communities due to local differences in environmental factors (Chapter 4) or community composition (Chapter 5). This is supported by recent studies (Brown, 2003; Durance and Ormerod, 2009) showing that a combination of factors acting across scales has the potential to ameliorate or enhance the impact of a specific drivers of interest (i.e. changes in water quality; climate change). But just how much of a problem is this?

In a final piece of analysis the EQI biology score, which represent the difference between the observed and expected community at a site, was calculated for each of the 1574 communities used in Chapters 3-4 based on their community composition in the first year of sampling. As can be seen in Figure 7.1, there is a negative correlation (Pearson's correlation, $r = -0.39$, d.f. =- 67.03, $p<0.001$) between EQI and temporal variability indicating that the most degraded communities are also those that exhibit the highest temporal variability. A similar relationship exists for other commonly used indices such as Average Score Per Taxon (Pearson's correlation, $r = -0.48$, d.f. = 35.61, $p<0.01$) and Biological Monitoring Working Party Score (Pearson's correlation, $r = -0.59$, d.f. = 43.01, $p<0.001$). This suggests that at the majority of reference sites (that by definition have the highest EQI values) temporal variability would be low compared with those communities that are ecologically impaired. However, communities do change so a multiyear assessment may be required to strengthen the predictive model. Alternatively, sites with low EQI score will typically exhibit higher levels of temporal variability therefore a more intensive monitoring programme may be required to establish whether any shift in EQI score can be attributed to a change in water quality. However, understanding this relationship presents a complex challenge as two possible scenarios exist.

Firstly. as was suggested in chapter 5, higher levels of temporal variability could be driven by colonisation and extinction processes with taxa arriving in the community but being unable to persist over long periods due to impacts on specific parts of their life cycle. In this first instance communities could be expected to exhibit small shifts in their EQI score as a result of the transient taxa however the mean EQI score would be expected to remain relatively constant. However, in the second scenario communities may possess multiple states with large variation in their EQI score. The argument for this scenario is that communities with high EQI scores possess taxa that are sensitive to disturbance and may therefore shift between states of high and low ecological quality. This second scenario may begin to explain the scatter of points within figure 7.1 where there can be large difference between the levels of temporal variation for communities for any given EQI score. For example for an EQI score of above 1 (indicating communities expected in pristine systems) the level of community change ranges between 0.87 and 2.20. Towards the upper limit of this range communities may exhibit large variation in their EQI score through time. The scatter in Figure 7.1 is therefore driven by communities shifting between multiple states reinforcing the argument for long term monitoring to establish the true ecological status of a community.

For the majority of predictive models the calculation of a score representing the difference between observed and expected communities does not represent the end point of the assessment. Rather these scores are used to assign sites to categories based on their ecological status. There are legislative requirements for sites to achieve a certain standard, for example in the Water Framework Directive (European Community, 2000) all sites must achieve good ecological status. Irvine (2004) identified lack of understanding of temporal variability as a major problem for the assessment of ecological status whereby temporal variation in communities could have the effect of either moving the status of a site up, or more importantly down a grade and thereby leading to incorrect management decisions about the site being taken. Much recent work has focussed on the impact of methodological problems whereby it has been demonstrated that close to class boundaries errors in data collection leads to an increases in the probability that a sample will be assigned to the wrong class (Irvine, 2004). Results from the current study suggest a similar problem may arise through temporal variability, however its importance may vary between communities.

In Figure 7.1 vertical lines indicate class boundaries relating to ecological quality targets. Based on the correlation between EQI and community change (moving horizontally across figure 7.1) where ecological quality is high (Class A and B) there is low temporal variability and therefore a low probability of assigning a site to an incorrect class. However, as ecological quality decreases, there is increasing temporal variability and therefore an associated increase in the probability of incorrect classification due to temporal variability. However, as discussed above, for any particular EQI score there is also considerable variationin the amount of community change exhibited at sites (moving verticall through figure 7.1). Less confidence may be placed in the classification of certain communities due to the long term variability that occurs within them based on the physicochemical and ecological characteristics that they possess. The implication of these results is that less monitoring effort can be put into assessing the status of communities that occur in the bottom right of figure 7.1. For all other communities more detailed long terms surveys are required to accurately understand shifts in the ecological status of communities through time.From a legislative point of view the greatest problem arises at the boundary that represents pass or fail. In Figure 7.1 this boundary arises between categories B and C where failure may result in fines. However, it represent a problem across all boundaries as an important part of biomonitoring, and a key reason for the adoption of a category based system, is to aid in the communication of the effectiveness of management measures.

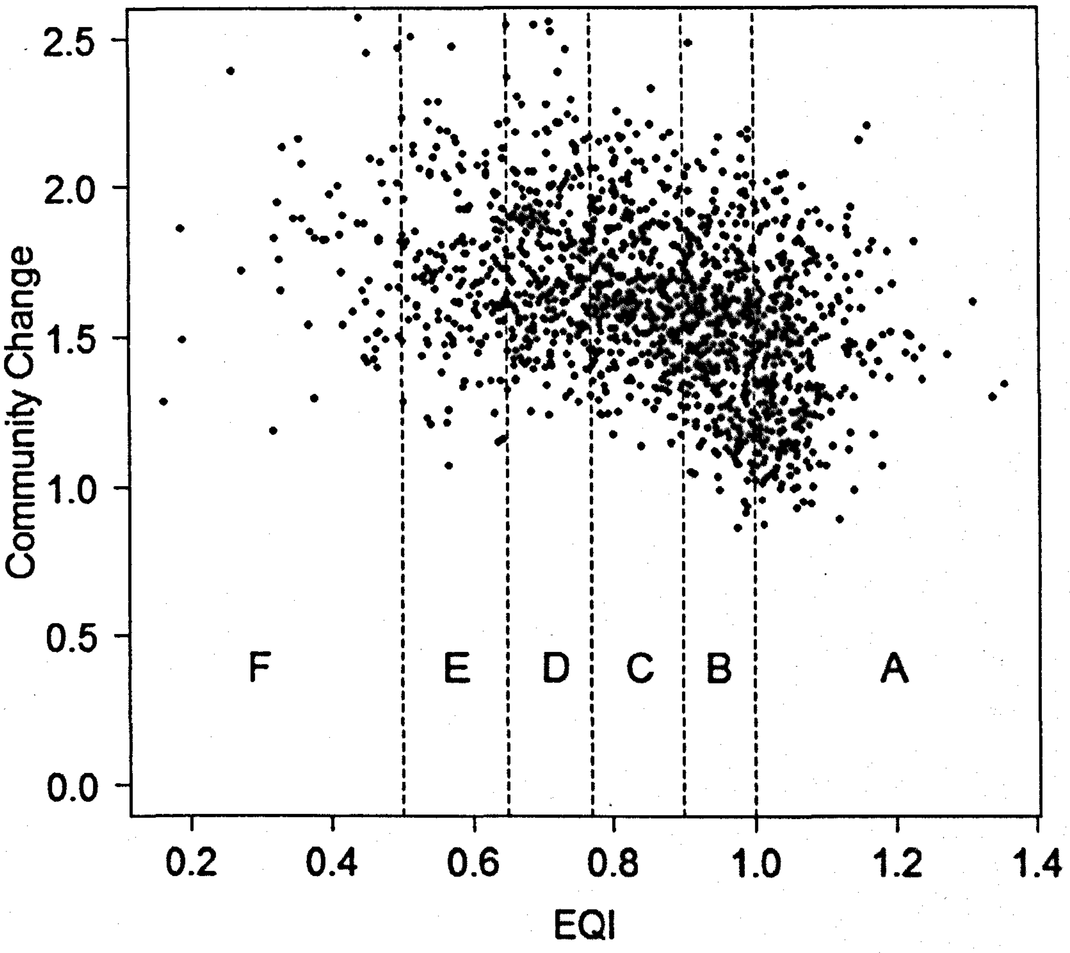


Figure 7.1. Relationship between Community Change from 1990 to 2005 and values for EQI (Pearson's correlation, r = -0.39. d.f. = 67.03, p <0.001) which represents the difference between observed and expected macroinvertebrate communities and as such represents proportional similarity of the community to its unstressed state (Helmsley-Flint, 2000). Letters indicate Environmental Quality classes from A – very good, to F – bad.

7.4: Problems in the use of long-term data and areas for further work

Strayer et al. (1986) and other authors (e.g. Elliott,1990) provide a detailed account of many of the general challenges faced in collecting long-term data, which include problems in funding, justification, site selection, collection and management of data over periods that may exceed the working life of many practitioners. A common feature uniting the two datasets used in this thesis, and a feature of many such dataset identified by Strayer et al. (1986), is that in the case of both BIOSYS and Sheffield Invertebrate Survey data were not collected with the specific aim of generating long-term data. In this final section, I consider a number of problems that have arisen in utilising the datasets in this way, and consider how the collection of long-term data could be altered to enhance the value of such data for future research.

7.4.1: Multiple vs. single points in time

A major difference between the two datasets used in the current study is that whereas BIOSYS sampled data over multiple points in time, Sheffield Invertebrate Survey simply represents two points in time. Having multiple points in time has a number of benefits as it allows an examination of trends in the community, for example whether communities vary around a mean composition or exhibit a trajectory of change (e.g. Scarsbrook, 2002). Similarly, depending on the spacing of samples, multiple points in time may allow us to examine the influence of cyclical processes (e.g. the North Atlantic Oscillation; Bradley and Ormerod, 2001) or identify the influence of gradual changes in conditions (e.g. climate change; Daufresne et al., 2004; Chessman, 2009).

In the context of the current study the availability of multiple points in time is important as it increases the confidence that can be placed in the degree of community temporal variability recorded at a site. For example, if just data from two points in time had been used it is possible that major impacts of climate or drought before one of the sampling periods could have led to the conclusion of high temporal variability in all communities. In the approach taken in the current study (Chapter 2; Section 2.2) the impact of such unusual event on the community will be averaged out over all samples.

More broadly continuous monitoring increases the probability that you will observe an unusual event that may significantly influence the community (Lawton, 1988). Such unusual events often provide fresh insight into the factors that shape communities by allowing an examination of the response and recovery of the system (e.g. Meffe and Minckley, 1987). However, in order to interpret the influence that such unusual events have on communities it is important to have a baseline with which to judge change. Such a baseline is only available when long-term continuous monitoring data are available.

Data from two points in time can still make a valuable contribution to our understanding of communities through time where it is possible to use a weight of evidence approach to explain any recorded variability. If historic environmental data are available, by examining changes in both abiotic and biotic measures it is possible to draw conclusion about drivers of change in communities. For example in the analysis of Sheffield Invertebrate survey (Chapter 6) the known history of the river system can plausibly be linked to changes noted in the community. Similarly where data exist for multiple sites it can be informative to examine differences and similarities in temporal variability between sites. If one community changes more than others then it is interesting to explore the possible reasons for this difference.

One way of adding value to historic datasets is to establish a baseline measure of variability. Contemporary sampling could be conducted over a few years to establish typical values of interannual variability that would provide a baseline level of change. By comparing the amount of change over 28 years (difference between historic and contemporary sampling) with the level of interannual variability in the present day it would be possible to assess the significance of any variation. Such an approach was used by Johnson et al. (1994) who demonstrated that temporal variability over 30 years was comparable with variability of the community over 1 year, leading the author to conclude that there was little long-term change in the community. The major problem with this approach for the Sheffield data is the use of a non-standardised method. However, for other data sets where there is a repeatable method such an approach could add value.

7.4.2: Repeatable, well documented method to ensure comparability across long time periods.

In Chapter 6 differences in sampling regime between years for the Sheffield Invertebrate survey were considered to make comparisons at the level of individual sites problematic. This contrasts with the BIOSYS dataset that uses a standardised method (Murray-Bligh et al., 1997) allowing comparisons across space and time. In the only other study of a comparable size to the data used in Chapters 3-5 (Chessman, 2009) the ability to examine patterns at a site were impacted to some degree by a change in method that resulted in increasing sub sampling through time. Chessman (2009) considered that this made it difficult to separate the influence of increasing climate change from changes in water quality. As such, a primary consideration when conducting sampling should be to use a method that can be well documented and is repeatable allowing researchers in future years to replicate the assessment.

7.4.3: Integration of data collection and the availability of long-term environmental data

In Chapter 4 the relationship between environmental factors and community temporal variability was examined using a range of data sources. One of the features of this analysis was that although data from the River Habitat Survey (RHS, Chapter 4: Section 4.2) and BIOSYS are collected by the same agency, there was little congruence between sites, with only 415 of a possible 1574 within 1000 meters. Increasing the congruence between sites would allow a more detailed examination of the relationship between river habitat features and aquatic communities, which have been demonstrated in other studies to be important (e.g. Buckton and Ormerod, 1997; Manel et al., 2000)

Similarly, both flow data and chemical data are collected across river systems in England and Wales. However, congruent monitoring data was unavailable for the current study. Flow regime has, in a number of studies, been considered to represent a master variable determining community structure (e.g. Statzner and Higler, 1986; Monk et al., 2006) and so is potentially important for the study of temporal variability. Similarly, changes in biological indices are related to the general improvement in chemical quality of rivers (Durance and Ormerod, 2009), however a more explicit link would provide further information with which to investigate temporal variability. Monitoring technology that use sensors and remote mechanisms to automatically and continuously monitor conditions provide one route through which this could be achieved (see review Glasgow et al., 2004).

The key point is not to do with increasing the amount of available data, but rather that integrating existing collection schemes may improve our ability to examine relationships between the physicochemical environment and communities.

7.4.4: Taxonomic Resolution

The analysis of both datasets used in the current study was based on family level data to ensure consistency between years (see Chapter 2). The use of such data are consistent with the identification level used in biomonitoring however, the use of family level data has been demonstrated to lead to a decrease in the level of observed community variability (Metzeling et al., 2002; Webb and King, 2009) as there can be different responses of species within a family to the same stress (Extence et al., 1999; Burgmer et al., 2007). This raises a number of interesting questions relevant to the present study. Firstly, will the effect be consistent across the entire country? Areas of low family level diversity such as upland regions of Wales, may have high species level diversity due to the presence of a few species rich families. How would this influence observed patterns? Secondly, the trait analysis presented in Chapter 5 was based on an average trait profile for families. Examining species level data may enable us to establish a stronger link between individual species traits and persistence of taxa within communities.

7.4.5: Large vs. small scale studies

The final point to raise relates to a number of the issues identified above concerning taxonomic resolution, integration of data collection, multiple sample points in time, etc.. Specifically what are the benefits of using a large dataset as opposed to a smaller number of sites that could be more intensively examined? The current study has provided valuable insight into the processes that drive temporal variability specifically by being able to describe spatial patterns, compare variability over environmental gradients and examine differing responses amongst communities. However, the insights that it has provided raise a number of questions about, for example, the relationship between structure and function of systems, importance of taxonomic resolution, influence of chemical and flow variables, that may now be better addressed through the use of smaller scale intensive studies. The current study could inform the design of these more intensive studies by suggesting that it would be appropriate to examine such issues using a number of contrasting communities and under a number of differing land use types. Differing mechanisms may influence the communities under this scenario and a more focussed approach might further enable us to investigate the causal mechanisms of temporal variability in lotic macroinvertebrate communities.

7.5: Conclusion

In conclusion, the current study has demonstrated that temporal variability in macroinvertebrate communities across England and Wales is driven by processes acting at multiple spatial scales. I have demonstrated a clear relationship between environmental factors and temporal variability of lotic macroinvertebrate communities. This relationship appears to be related to the traits of taxa within the community that govern their response to disturbance. The combination of both spatial and temporal scale makes the current study an important contribution to the understanding of the temporal dynamics of communities. However, there still exists much uncertainty that might only be addressed through a joined up approach to the study of river systems.

REFERENCES

Aagaard, K., J. O. Solem, T. Bongard and O. Hanssen (2004). "Studies of aquatic insects in the Atna River 1987–2002." Hydrobiologia **521**: 87-105.

Allan, D. J. and L. B. Johnson (1997). "Catchment-scale analysis of aquatic ecosystems." Freshwater Biology **37**: 107-111.

Allan, J. D. (2004). "Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems." Annual Review of Ecology, Evolution, and Systematics **35**(1): 257-284.

Archaimbault, V., P. Usseglio-Polatera and J.-P. Bossche (2005). "Functional Differences Among Benthic Macroinvertebrate Communities in Reference Streams of Same Order in a Given Biogeographic Area." Hydrobiologia **551**(1): 171-182.

Armitage, P. D. (2006). "Long-term faunal changes in a regulated and an unregulated stream - Cow Green thirty years on." River Research and Applications **22**(9): 947-966.

Beals, E. W. (1984). "Bray Curtis ordination: an effective strategy for analysis of multivariate ecological data." Advances in Ecological Research **14**: 1-55.

Bêche, L. A., P. G. Connors, V. H. Resh and A. M. Merenlender (2009). "Resilience of fishes and invertebrates to prolonged drought in two California streams." Ecography **32**(5): 778-788.

Beche, L. A., E. P. McElravy and V. H. Resh (2006). "Long-term seasonal variation in the biological traits of benthic-macroinvertebrates in two Mediterranean-climate streams in California, U.S.A." Freshwater Biology **51**(1): 56-75.

Beche, L. A. and V. H. Resh (2007). "Short-term climatic trends affect the temporal variability of macroinvertebrates in California 'Mediterranean' streams." Freshwater Biology **52**(12): 2317-2339.

Behmer, D. J. and C. P. Hawkins (1986). "Effects of overhead canopy on macroinvertebrate production in a Utah stream." Freshwater Biology **16**(3): 287-300.

Bellier, E., P. Monestiez, J.-P. Durbec and J.-N. Candau (2007). "Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches." Ecography **30**(3): 385-399.

Belyea, L. and J. Lancaster (1999). "Assembly rules within a contingent ecology." Oikos **86**(3): 402-416.

Bis, B. and P. Usseglio-Polatera (2004). Species Trait Analysis Standardisation of River Classification: Framework method for calibrating different biological survey results against ecological quality classifications to be developed for the Water Framework directive.

Bonada, N., S. Doledec and B. Statzner (2007). "Taxonomic and biological trait differences of stream macroinvertebrate communities between Mediterranean and temperate regions: implications for future climatic scenarios." Global Change Biology **13**(8): 1658-1671.

Bonada, N., N. Prat, V. H. Resh and B. Statzner (2006). "Developments in Aquatic Insect Biomonitoring: A Comparative Analysis of Recent Approaches." Annual Review of Entomology **51**(1): 495-523.

Bonada, N., M. Rieradevall and N. Prat (2007). "Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network." Hydrobiologia **589**(1): 91-106.

Boon, P. J. (2000). "The development of integrated methods for assessing river conservation value." Hydrobiologia **422/423**: 413–428.

Borcard, D. and P. Legendre (2002). "All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices " Ecological Modelling **153**(1-2): 51-68.

Borcard, D., P. Legendre, C. Avois-Jacquet and H. Tuomistoc (2004). "Dissecting the spatial structure of ecological data at multiple scales." Ecology **85**(7): 1826-1832.

Boulton, A. J., C. G. Peterson, N. B. Grimm and S. G. Fisher (1992). "Stability of an Aquatic Macroinvertebrate Community in a Multiyear Hydrologic Disturbance Regime." Ecology **73**(6): 2192-2207.

Boyero, L. (2003). "Multiscale patterns of spatial variation in stream macroinvertebrate communities." Ecological Research **18**(4): 365-379.

Bradley, D. C. and S. J. Ormerod (2001). "Community persistence among stream invertebrates tracks the North Atlantic Oscillation." Journal of Animal Ecology **70**(6): 987-996.

Bradley, D. C. and S. J. Ormerod (2002). "Long-term effects of catchment liming on invertebrates in upland streams." Freshwater Biology **46**(4): 161-171.

Bradt, P., M. Urban, N. Goodman, S. Bissell and I. Spiegel (1999). "Stability and resilience in benthic macroinvertebrate assemblages." Hydrobiologia **403**(0): 123-133.

Breiman, L. (2001). "Random forests." Machine Learning **45**(1): 5 - 32.

Brewin, P. A., S. T. Buckton and S. J. Ormerod (2000). "The seasonal dynamics and persistence of stream macroinvertebrates in Nepal: do monsoon floods represent disturbance?" Freshwater Biology **44**(4): 581-594.

Briers, R. A. (2006). Dissecting patterns in lake macrophyte communities at multiple spatial scales. Fourteenth Annual IALE(UK) conference. Oxford Brookes University, IALE(UK).

Brind'Amour, A., D. Boisclair, P. Legendre and D. Borcard (2005). "Multiscale spatial distribution of a littoral fish community in relation to environmental variables." Limnology and Oceanography **50(2)**: 465-479.

Brittain, J. E. and T. J. Eikeland (1988). "Invertebrate drift: A review." Hydrobiologia **166**(1): 77-93.

Brown, B. (2007). "Habitat heterogeneity and disturbance influence patterns of community temporal variability in a small temperate stream." Hydrobiologia **586**(1): 93-106.

Brown, B. L. (2003). "Spatial heterogeneity reduces temporal variability in stream insect communities." Ecology Letters **6**: 316-325.

Brown, L., A. Milner and D. Hannah (2006). "Stability and Persistence of Alpine Stream Macroinvertebrate Communities and the Role of Physicochemical Habitat Variables." Hydrobiologia **560**(1): 159-173.

Brown, L. E., D. M. Hannah and A. M. Milner (2007). "Vulnerability of alpine stream biodiversity to shrinking glaciers and snowpacks." Global Change Biology **13**(5): 958-966.

Buckton, S. T. and S. J. Ormerod (1997). "Use of a new standardized habitat survey for assessing the habitat preferences and distribution of upland river birds." Bird Study **44**: 327-337.

Bunn, S. E. and P. M. Davies (2000). "Biological processes in running waters and their implications for the assessment of ecological integrity." Hydrobiologia **422-423**(0): 61-70.

Burgherr, P. and J. V. Ward (2001). "Longitudinal and seasonal distribution patterns of the benthic fauna of an alpine glacial stream (Val Roseg, Swiss Alps)." Freshwater Biology **46**(12): 1705-1721.

Burgmer, T., H. Hillebrand and M. Pfenninger (2007). "Effects of climate-driven temperature changes on the diversity of freshwater macroinvertebrates." Oecologia **151**(1): 93-103.

Cao, Y., D. D. Williams and N. E. Williams (1998). "How important are rare species in aquatic community ecology and bioassessment?" Limnology and Oceanography **43**(7): 1403-1409.

Carpenter, S. R. and W. A. Brock (2006). "Rising variance: a leading indicator of ecological transition." Ecology Letters **9**(3): 311-318.

Chadwick, M. A., D. R. Dobberfuhl, A. C. Benke, A. D. Huryn, K. Suberkropp and J. E. Thiele (2006). "Urbanization affects stream ecosystem function by altering hydrology, chemistry, and biotic richness." Ecological Applications **16**(5): 1796-1807.

Charvet, S., B. Statzner, P. Usseglio-Polatera and B. Dumont (2000). "Traits of benthic macroinvertebrates in semi-natural French streams: an initial application to biomonitoring in Europe." Freshwater Biology **43**(2): 277-296.

Chase, J. M. (2007). "Drought mediates the importance of stochastic community assembly." Proceedings of the National Academy of Sciences **104**(44): 17430-17434.

Chessman, B. C. (2009). "Climatic changes and 13-year trends in stream macroinvertebrate assemblages in New South Wales, Australia." Global Change Biology **15**(11): 2791-2802.

Chevenet, F., S. Dolodec and D. Chessel (1994). "A fuzzy coding approach for the analysis of long-term ecological data." Freshwater Biology **31**(3): 295-309.

Clarke, R. (2000). Uncertainty in estimate of biological quality based on RIVPACS`. Assessing the biological quality of fresh waters: RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association: 39-54.

Clarke, R. and D. Hering (2006). "Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS." Hydrobiologia **566**(1): 433-439.

Clarke, R. T., M. T. Furse, J. F. Wright and D. Moss (1996). "Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna." Journal of Applied Statistics **23**(2-3): 311 - 332.

Clarke, R. T., M. T. Furse, J. M. Gunn, J. M. Winder and J. F. Wright (2002). "Sampling variation in macroinvertebrate data and implications for river quality indices." Freshwater Biology **47**(9): 1735-1751.

Clausen, B. and B. J. F. Biggs (2000). "Flow variables for ecological studies in temperate streams: groupings based on covariance." Journal of Hydrology **237**(3-4): 184-197.

Clenaghan, C., P. S. Giller, J. O'Halloran and R. Hernan (1998). "Stream macroinvertebrate communities in a conifer afforested catchment in Ireland: relationships to physico-chemical and biotic factors." Freshwater Biology **40**(1): 175-193.

Collier, K. J. (2007). "Temporal patterns in the stability, persistence and condition of stream macroinvertebrate communities: relationships with catchment land-use and regional climate." Freshwater Biology **53**(3): 603-616.

Collins, S. L., F. Micheli and L. Hartt (2000). "A method to determine rates and patterns of variability in ecological communities." Oikos **91**(2): 285-293.

Connell, J. H. (1978). "Diversity in Tropical Rain Forests and Coral Reefs." Science **199**(4335): 1302-1310.

Connell, J. H. and W. P. Sousa (1983). "On the Evidence Needed to Judge Ecological Stability or Persistence." The American Naturalist **121**(6): 789-824.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton and M. van den Belt

(1997). "The value of the world's ecosystem services and natural capital." Nature **387**(6630): 253-260.

Crawley, M. J. (2007). The R Book, Wiley.

Cummins, K. W. and M. J. Klug (1979). "Feeding ecology of stream invertebrates." Annual Review of Ecology and Systematics **10**: 147-172.

Currie, D. J. (2007). "Disentangling the roles of environment and space in ecology." Journal of Biogeography **34**(12): 2009-2011.

Daufresne, M., M. C. Roger, H. Capra and N. Lamouroux (2004). "Long-term changes within the invertebrate and fish communities of the Upper Rhone River: effects of climatic factors." Global Change Biology **10**(1): 124-140.

Davies, N. M., R. H. Norris and M. C. Thoms (2000). "Prediction and assessment of local stream habitat features using large-scale catchment characteristics." Freshwater Biology **45**(3): 343-369.

Davies, P. E. (2000). Development of a national river bioassessment system (AUSRIVAS) in Australia. Assessing the biological quality of fresh waters: RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association: 113-124.

Davy-Bowker, J., M. T. Furse, R. T. Clarke and M. J. Gravelle (1999). Analysis of 1995 survey data. Phase 2 post-survey appraisal. Unit 1: Taxon Distribution studies.

De'ath, G. and K. E. Fabricius (2000). "Classification and regression trees: a powerful yet simple technique for ecological data analysis." Ecology **81** (11): 3178-3192.

Death, R. G. and M. J. Winterbourn (1994). "Environmental Stability and Community Persistence: A Multivariate Perspective." Journal of the North American Benthological Society **13**(2): 125-139.

Death, R. G. and M. J. Winterbourn (1995). "Diversity Patterns in Stream Benthic Invertebrate Communities: The Influence of Habitat Stability." Ecology **76**(5): 1446-1460.

Diaz, A. M., M. L. S. Alonso and M. R. V.-A. Gutierrez (2008). "Biological traits of stream macroinvertebrates from a semi-arid catchment: patterns along complex environmental gradients." Freshwater Biology **53**(1): 1-21.

Dines, A. D. and J. A. D. Murray-Bligh (2000). Quality assurance and RIVPACS. Assessing the biological quality of fresh waters: RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association.

Diniz-Filho, J. A. F. and L. M. Bini (2005). "Modelling geographical patterns in species richness using eigenvector-based spatial filters." Global Ecology and Biogeography **14**(2): 177-185.

Doak, D. F., D. Bigger, E. K. Harding, M. A. Marvier, R. E. O'Malley and D. Thomson (1998). "The Statistical Inevitability of Stability-Diversity Relationships in Community Ecology." The American Naturalist **151**(3): 264-276.

Doledec, S. and B. Statzner (1994). "Theoretical habitat templets, species traits, and species richness: 548 plant and animal species in the Upper Rhone River and its floodplain." Freshwater biology **31**(3): 523-538.

Dormann, C. F., J. M. McPherson, M. B. Araujo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Daniel Kissling, I. Kuhn, R. Ohlemuller, P. R. Peres-Neto, B. Reineking, B. Schroder, F. M. Schurr and R. Wilson (2007). "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review." Ecography **30**(5): 609-628.

Downes, B. J., P. S. Lake and E. S. G. Schreiber (1993). "Spatial variation in the distribution of stream invertebrates: implications of patchiness for models of community organization." Freshwater Biology **30**(1): 119-132.

Dray, S., P. Legendre and P. R. Peres-Neto (2006). " Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM) " Ecological Modelling **196**(3-4): 483-493.

Dudgeon, D., A. H. Arthington, M. O. Gessner, Z.-I. Kawabata, D. J. Knowler, C. Leveque , R. J. Naiman, A.-H. Prieur-Richard, D. Soto, M. L. J. Stiassny and C. A. Sullivan (2006). "Freshwater biodiversity: importance, threats, status and conservation challenges." Biological Reviews **81**(02): 163-182.

Durance, I. and S. J. Ormerod (2007). "Climate-change effects on upland stream macroinvertebrates over a 25 year period." Global Change Biology **13**(5): 942-957.

Durance, I. and S. J. Ormerod (2009). "Trends in water quality and discharge confound long-term warming effects on river macroinvertebrates." Freshwater Biology **54**(2): 388-405.

Dussault, G. (2009). Matrix Distance.

Edington, J. M. and A. G. Hildrew (1981). A key to the caseless caddis larvae of the British Isles. Ambleside, Cumbria, UK, Freshwater Biological Association.

Elith, J., J. R. Leathwick and T. Hastie (2008). "A working guide to boosted regression trees." Journal of Animal Ecology **77**(4): 802-813.

Elliott, J. M. (1990). "The need for long-term investigations in ecology and the contribution of the Freshwater Biological Association. " Freshwater Biology **23**: 1-5.

Elliott, J. M. (1992). "Variation in the population density of adult sea trout, *Salmo trutta*, in 67 rivers in England and Wales." Ecology of Freshwater Fishes **1**: 5-11.

Elliott, J. M. (2006). "Periodic habitat loss alters the competitive coexistence between brown trout and bullheads in a small stream over 34 years." Journal of Animal Ecology **75**: 54-63.

Elliott, J. M., U. H. Humpesch and T. T. Macan (1988). Larvae of the British Ephemeroptera: a Key with Ecological Notes

European Community (2000). Directive 2000/60/EC of October 23 2000 of the European Parliament and of the Council establishing a framework for community action in the field of water policy. Official Journal of the European Community. **L327**: 1–72.

Extence, C. A., D. M. Balbi and R. P. Chadd (1999). "River flow indexing using British benthic macroinvertebrates: a framework for setting hydroecological objectives." Regulated Rivers: Research & Management **15**(6): 545-574.

Fagan, W. F. (2002). "Connectivity, fragmentation, and extinction risk in dendritic metapopulations." Ecology **12**: 3243–3249.

Faith, D. P., P. R. Minchin and L. Belbin (1987). "Compositional dissimilarity as a robust measure of ecological distance." Vegetatio **69**: 57-68.

Farnsworth, E. J. (1998). "Issues of Spatial, Taxonomic and Temporal Scale in Delineating Links between Mangrove Diversity and Ecosystem Function." Global Ecology and Biogeography Letters **7**(1): 15-25.

Feld, C. (2004). "Identification and measure of hydromorphological degradation in Central European lowland streams." Hydrobiologia **516**(1): 69-90.

Feld, C. K. and D. Hering (2007). "Community structure or function: effects of environmental stress on benthic macroinvertebrates at different spatial scales." Freshwater Biology **52**: 1380-1399.

Firth, C. J. (1997). 900 years of the Don Fishery - Domesday to the dawn of the new millenium, Environment Agency.

Fisher, J. A. D., K. T. Frank and W. C. Leggett (2010). "Dynamic macroecology on ecological time-scales." Global Ecology and Biogeography **19**(1): 1-15.

Fox, P. J. A., M. Naura and P. Scarlett (1998). "An account of the derivation and testing of a standard field method, River Habitat Survey." Aquatic Conservation: Marine and Freshwater Ecosystems **8**(4): 455-475.

Franklin, J. F., C. S. Bledsoe and J. T. Callahan (1990). "Contributions of the Long-Term Ecological Research Program." BioScience **40**(7): 509-523.

Fraterrigo, J. M. and J. A. Rusak (2008). "Disturbance-driven changes in the variability of ecological patterns and processes." Ecology Letters **11**(7): 756-770.

Friday, L. E. (1988). A key to the adults of British water beetles. Taunton, Somerset, UK, Field Studies Council.

Frissell, C. A., W. J. Liss, C. E. Warren and M. D. Hurley (1986). "A hierarchical framework for stream habitat classification: Viewing streams in a watershed context." Environmental Management **10**(2): 199-214.

Fruget, J.-F., M. Centofanti, J. Dessaix, J.-M. Olivier, J.-C. Druart and P.-J. Martinez (2001). "Temporal and spatial dynamics in large rivers: example of a long-term monitoring of the middle Rhone River " Annales de Limnologie **37**(3): 237-251.

Furze, M. (2007). Coded Checklist of Animals Occurring in Fresh Water in the British Isles, CEH.

Gasith, A. and V. H. Resh (1999). "Streams in mediterranean climate regions: Abiotic Influences and Biotic Responses to Predictable Seasonal Events." Annual Review of Ecology and Systematics **30**(1): 51-81.

Gaston, K. J. and J. I. Spicer (1998). Biodiversity. An Introduction, Blackwell Science Ltd.

Gergel, S. E., M. G. Turner, J. R. Miller, J. M. Melack and E. H. Stanley (2002). "Landscape indicators of human impacts to riverine systems." Aquatic Sciences - Research Across Boundaries **64**(2): 118-128.

Gibbins, C. N., C. F. Dilks, R. Malcolm, C. Soulsby and S. Juggins (2001). "Invertebrate communities and hydrological variation in Cairngorm mountain streams." Hydrobiologia **462**(1 - 3): 205-219.

Giller, P. S. and J. O'Halloran (2004). "Forestry and the aquatic environment: studies in an Irish context." Hydrology and Earth System Sciences **8(3)**: 314-326.

Glasgow, H. B., J. M. Burkholder, R. E. Reed, A. J. Lewitus and J. E. Kleinman (2004). "Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies." Journal of Experimental Marine Biology and Ecology **300**(1-2): 409-448.

Griffith, D. A. and P. R. Peres-Neto (2006). "Spatial modeling in ecology: the flexibility of eigenfunction spatial analyse. " Ecology **87**(10): 2603–2613.

Griswold, M. W., R. W. Berzinis, T. L. Crisman and S. W. Golladay (2008). "Impacts of climatic stability on the structural and functional aspects of macroinvertebrate communities after severe drought." Freshwater Biology **53**(12): 2465-2483.

Gurtz, M. E. and J. B. Wallace (1984). "Substrate-Mediated Response of Stream Invertebrates to Disturbance." Ecology **65**(5): 1556-1569.

Harding, J. S., E. F. Benfield, P. V. Bolstad, G. S. Helfman and E. B. D. Jones (1998). "Stream biodiversity: The ghost of land use past." Proceedings of the National Academy of Sciences of the United States of America **95**(25): 14843-14847.

Heatherly, T., M. Whiles, D. Gibson, S. Collins, A. Huryn, J. Jackson and M. Palmer (2007). "Stream insect occupancy-frequency patterns and metapopulation structure." Oecologia **151**(2): 313-321.

Heino, J. (2005). "Functional biodiversity of macroinvertebrate assemblages along major ecological gradients of boreal headwater streams." Freshwater Biology **50**(9): 1578-1587.

Heino, J., H. Mykra, H. Hamalainen, J. Aroviita and T. Muotka (2007). "Responses of taxonomic distinctness and species diversity indices to anthropogenic impacts and natural environmental gradients in stream macroinvertebrates." Freshwater Biology **52**(9): 1846-1861.

Hillebrand, H., J. Soininen and P. Snoeijs (2009). "Warming leads to higher species turnover in a coastal ecosystem." Global Change Biology **16**(4): 1181-1193.

Hochachka, W. M., R. Caruana, D. Fink, A. R. T. Munson, M. Riedewald, D. Sorokina and S. Kelling (2007). "Data-Mining Discovery of Pattern and Process in Ecological Systems." Journal of Wildlife Management: 2427-2437.

Holmes, N. T. H. (2006). "The importance of long-term data sets in science and river management." Aquatic Conservation: Marine and Freshwater Ecosystems **16**(4): 329-333.

Hubbell, S. P. (1997). "A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs." Coral Reefs **16**(0): 9-21.

Humphrey, C. L., A. W. Storey and L. Thurtell (2000). AUSRIVAS: operator sample processing errors and temporal variability - implcations for model sensitivity. Assessing the Biological Quality of Fresh Waters: RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association: 143-164.

Hynes, H. B. N. (1977). A Key to the Adults and Nymphs of the British Stoneflies (Plecoptera), with Notes on their Ecology and Distribution Ambleside, Cumbria, UK, Freshwater Biological Association.

Ilg, C. and E. Castella (2006). "Patterns of macroinvertebrate traits along three glacial stream continuums." Freshwater Biology **51**(5): 840-853.

Irvine, K. (2004). "Classifying ecological status under the European Water Framework Directive: the need for monitoring to account for natural variability." Aquatic Conservation: Marine and Freshwater Ecosystems **14**: 107-112.

Jackson, J. K. and L. Fureder (2006). "Long-term studies of freshwater macroinvertebrates: a review of the frequency, duration and ecological significance." Freshwater Biology **51**(3): 591-603.

Janzen, H. H. (2009). "Long-term ecological sites: musings on the future, as seen (dimly) from the past." Global Change Biology **15**(11): 2770-2778.

Johnson, P. D., K. M. Brown and C. V. Covell Jr. (1994). "A comparison of the macroinvertebrate assemblage in Doe Run Creeek, Kentucky: 1960 and 1990." Journal of the North American Benthological Society **13**(4): 496-510.

Johnson, R. K., M. T. Furse, D. Hering and L. Sandin (2007). "Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programmes." Freshwater Biology **52**(5): 939-958.

Johnson, R. K. and D. Hering (2009). "Response of taxonomic groups in streams to gradients in resource and habitat characteristics." Journal of Applied Ecology **46**: 175-186.

Jones, P. D. and D. H. Lister (1998). "Riverflow reconstructions for 15 catchments over England and Wales and an assessment of hydrologic drought since 1865." International Journal of Climatology **19**: 999-1013.

Koch, E. W., E. B. Barbier, B. R. Silliman, D. J. Reed, G. M. E. Perillo, S. D. Hacker, E. F. Granek, J. H. Primavera, N. Muthiga, S. Polasky, B. S. Halpern, C. J. Kennedy, C. V. Kappel and E. Wolanski (2009). "Non-linearity in ecosystem services: temporal and spatial variability in coastal protection." Frontiers in Ecology and the Environment **7**(1): 29-37.

Koleff, P., K. J. Gaston and J. J. Lennon (2003). "Measuring beta diversity for presence-absence data." Journal of Animal Ecology **72**(3): 367-382.

Konrad, C. P., A. M. D. Brasher and J. T. May (2008). "Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States." Freshwater Biology **53**(10): 1983-1998.

Korfiatis, K. J. and G. P. Stamou (1999). "Habitat Templets and the Changing Worldview of Ecology." Biology and Philosophy **14**(3): 375-393.

Kuhn, S., B. Egert, S. Neumann and C. Steinbeck (2008). "Building blocks for automated elucidation of metabolites: Machine learning methods for NMR prediction." BMC Bioinformatics **9**(1): 400.

Lancaster, J. (2000). "Geometric Scaling of Microhabitat Patches and Their Efficacy as Refugia during Disturbance." Journal of Animal Ecology **69**(3): 442-457.

Legendre, P. and L. Legendre (1998). Numerical Ecology, Elsevier.

Lepori, F. and B. Malmqvist (2009). "Deterministic control on community assembly peaks at intermediate levels of disturbance." Oikos **118**(3): 471-479.

Li, J., A. Herlihy, W. Gerth, P. Kaufmann, S. Gregory, S. Urquhart and D. P. Larsen (2001). "Variability in stream macroinvertebrates at multiple spatial scales." Freshwater Biology **46**(1): 87-97.

Liaw, A. and M. Wiener (2002). Classification and Regression by randomForest. R News. **2(3)**: 18--22.

Loreau, M. (2000). "Biodiversity and Ecosystem Functioning: Recent Theoretical Advances." Oikos **91**(1): 3-17.

Magnuson, J. J. (1990). "Long-Term Ecological Research and the Invisible Present." BioScience **40**(7): 495-501.

Magurran, A. E. (2004). Measuring Biological Diversity, Blackwell Science Ltd.

Malmqvist, B. (2002). "Aquatic invertebrates in riverine landscapes." Freshwater Biology **47**(4): 679-694.

Manel, S., S. T. Buckton and S. J. Ormerod (2000). "Testing Large-Scale Hypotheses Using Surveys: The Effects of Land Use on the Habitats, Invertebrates and Birds of Himalayan Rivers." Journal of Applied Ecology **37**(5): 756-770.

McCabe, D. J. and N. J. Gotelli (2000). "Effects of disturbance frequency, intensity, and area on assemblages of stream macroinvertebrates." Oecologia **124**: 270-279.

McElravy, E. P., G. A. Lamberti and V. H. Resh (1989). "Year-to-year variation in the aquatic macroinvertebrate fauna of a northern California stream." Journal of the North American Benthological Society **8**(1): 51-63.

McGill, B. J., B. J. Enquist, E. Weiher and M. Westoby (2006). "Rebuilding community ecology from functional traits." Trends in Ecology and Evolution **21**(4): 178-185.

Meffe, G. K. and W. L. Minckley (1987). "Persistence and Stability of Fish and Invertebrate Assemblages in a Repeatedly Disturbed Sonoran Desert Stream." American Midland Naturalist **117**(1): 177-191.

Metcalfe, J. L. (1989). "Biological water quality assessment of running waters based on macroinvertebrate communities: History and present status in Europe." Environmental Pollution **60**(1-2): 101-139.

Metzeling, L., D. Robinson, S. Perriss and R. Marchant (2002). "Temporal persistence of benthic invertebrate communities in south-eastern Australian streams: taxonomic resolution and implications for the use of predictive models." Marine and Freshwater Research **53**: 1223-1234.

Millennium Ecosystem Assessment (2005). Ecosystem and Human Wellbeing: Synthesis. Washington, DC. Island Press.

Milner, A. M., S. C. Conn and L. E. Brown (2005). "Persistence and stability of macroinvertebrate communities in streams of Denali National Park, Alaska: implications for biological monitoring." Freshwater Biology **51**(2): 373-387.

Minshall, G. W. (1988). "Stream ecosystem theory: a global perspective." Journal of the North American Benthological Society **7**(4): 163-288.

Monk, W. A., P. J. Wood, D. M. Hannah, D. A. Wilson, C. A. Extence and R. P. Chadd (2006). "Flow variability and macroinvertebrate community response within riverine systems." River Research and Applications **22**: 595-615.

Moss, D., M. T. Furse, J. F. Wright and P. D. Armitage (1987). "The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data." Freshwater Biology **17**(1): 41-52.

Murphy, J. F. and J. Davy-Bowker (2005). "Spatial structure in lotic macroinvertebrate communities in England and Wales: relationship with physical, chemical and anthropogenic stress variables." Hydrobiologia **534**(1): 151-164.

Murphy, J. F. and P. S. Giller (2000). "Seasonal dynamics of macroinvertebrate assemblages in the benthos and associated with detritus packs in two low-order streams with different riparian vegetation." Freshwater Biology **43**: 617-631.

Murray-Bligh, J. A. D., M. T. Furze, F. H. Jones, R. J. M. Gunn, R. A. Dines and J. F. Wright (1997). Procedure for collecting and analysing macro-invertebrate samples for RIVPACS, Institute of Freshwater Ecology and Environment Agency.

Naeem, S. (1998). "Species Redundancy and Ecosystem Reliability." Conservation Biology **12**(1): 39-45.

Naeslund, B. and J. Norberg (2006). "Ecosystem consequences of the regional species pool." Oikos **115**(3): 504-512.

Niemi, G. J., P. Devore, N. Detenbeck, D. Taylor, A. Lima and J. Pastor (1990). "Overview of Case Studies on Recovery of Aquatic Systems from Disturbance." Environmental Management **14**(5): 571-587.

Niemi, G. J. and M. E. McDonald (2004). "Application of Ecological Indicators." Annual Review of Ecology, Evolution, and Systematics **35**(1): 89-111.

Odum, E. P. (1985). "Trends Expected in Stressed Ecosystems." BioScience **35**(7): 419-422.

Ormerod, S. J. and I. Durance (2009). "Restoration and recovery from acidification in upland Welsh streams over 25 years." Journal of Applied Ecology **46**(1): 164-174.

Paillex, A., S. Dolédec, E. Castella and S. Mérigoux (2009). "Large river floodplain restoration: predicting species richness and trait responses to the restoration of hydrological connectivity." Journal of Applied Ecology **46**(1): 250-258.

Palmer, M., E. Bernhardt, E. Chornesky, S. Collins, A. Dobson, C. Duke, B. Gold, R. Jacobson, S. Kingsland, R. Kranz, M. Mappin, M. L. Martinez, F. Micheli, J. Morse, M. Pace, M. Pascual, S. Palumbi, O. J. Reichman, A. Simons, A.

Townsend and M. Turner (2004). "Ecology for a Crowded Planet." Science **304**: 1251-1252.

Palmer, M. A., J. D. Allan and C. A. Butman (1996). "Dispersal as a regional process affecting the local dynamics of marine and stream benthic invertebrates." Trends in Ecology and Evolution **11**: 322-326.

Parsons, M., M. C. Thoms and R. H. Norris (2003). "Scales of Macroinvertebrate Distribution in Relation to the Hierarchical Organization of River Systems." Journal of the North American Benthological Society **22**(1): 105-122.

Paul, M. J. and J. L. Meyer (2001). "Streams in the urban landscape." Annual Review of Ecology and Systematics **32**(1): 333-365.

Petchey, O. L. and K. J. Gaston (2002). "Functional diversity (FD), species richness and community composition." Ecology Letters **5**: 402-411.

Petchey, O. L. and K. J. Gaston (2007). "Dendrograms and measuring functional diversity." Oikos **116**(8): 1422-1426.

Peters, J., B. D. Baets, N. E. C. Verhoest, R. Samson, S. Degroeve, P. D. Becker and W. Huybrechts (2007). "Random forests as a tool for ecohydrological distribution modelling." Ecological Modelling **207**(2-4): 304-318.

Petersen, I., Z. Masters, A. G. Hildrew and S. J. Ormerod (2004). "Dispersal of adult aquatic insects in catchments of differing land use." Journal of Applied Ecology **41**: 934-950.

Pimm, S. L., G. J. Russell, J. L. Gittleman and T. M. Brooks (1995). "The Future of Biodiversity." Science **269**(5222): 347-350.

Poff, N. and J. Ward (1990). "Physical habitat template of lotic systems: Recovery in the context of historical pattern of spatiotemporal heterogeneity." Environmental Management **14**(5): 629-645.

Poff, N. L. (1997). "Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology." Journal of the North American Benthological Society **16**(2): 391-409.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks and J. C. Stromberg (1997). "The Natural Flow Regime." BioScience **47**(11): 769-784.

Prasad, A., L. Iverson and A. Liaw (2006). "Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction." Ecosystems **9**(2): 181-199.

Quinn, G. P. and M. J. Keough (2002). Experimental design and Data Analysis for Biologists, Cambridge University Press.

R Development Core Team (2008). R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Rahel, F. J. (1990). "The Hierarchical Nature of Community Persistence: A Problem of Scale." The American Naturalist **136**(3): 328-344.

Rahel, F. J. (2000). "Homogenization of Fish Faunas Across the United States." Science **288**(5467): 854-856.

Raup, D. M. (1975). "Taxonomic diversity estimation using rarefaction." Paleobiology **1**: 333-342.

Raven, P. J., N. T. H. Holmes, F. H. Dawson and M. Everard (1998). "Quality assessment using River Habitat Survey data." Aquatic Conservation: Marine and Freshwater Ecosystems **8**: 477-499.

Reece, P. F., T. B. Reynoldson, J. S. Richardson and D. M. Rosenberg (2001). "Implications of seasonal variation for biomonitoring with predictive models in the Fraser River catchment, British Columbia." Canadian Journal of Fisheries and Aquatic Sciences **58**(7): 1411-1417.

Reid, M. A. and R. W. Ogden (2006). "Trend, variability or extreme event? The importance of long-term perspectives in river ecology." River Research and Applications **22**(2): 167-177.

Reimann, C., P. Filzmoser, R. Garrett and R. Dutter (2008). Statistical Data Analysis Explained. West Sussex, England, John Wiley & Sons, Ltd.

Resh, V. H., A. V. Brown, A. P. Covich, M. E. Gurtz, H. W. Li, G. W. Minshall, S. R. Reice, A. L. Sheldon, J. B. Wallace and R. C. Wissmar (1988). "The Role of Disturbance in Stream Ecology." Journal of the North American Benthological Society **7**(4): 433-455.

Resh, V. H., A. G. Hildrew, B. Statzner and C. R. Townsend (1994). "Theoretical habitat templets, species traits, and species richness: a synthesis of long-term ecological research on the Upper Rhone River in the context of concurrently developed ecological theory." Freshwater Biology **31**(3): 539-554.

Reynoldson, T. B., K. E. Day and T. Pascoe (2000). The development of the BEAST: a predictive approach for assessing sediment quality in the North American Great Lakes. Assessing the biological quality of freshwaters. RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association.

Richards, C., R. Haro, L. Johnson and G. Host (1997). "Catchment and reach-scale properties as indicators of macroinvertebrate species traits." Freshwater Biology **37**(1): 219-230.

Richards, C. and G. W. Minshall (1992). "Spatial and temporal trends in stream macroinvertebrate communities: the influence of catchment disturbance." Hydrobiologia **241**: 173-184.

Richoux, P. (1994). "Theoretical habitat templets, species traits, and species richness: aquatic Coleoptera in the Upper Rhone River and its floodplain." Freshwater Biology **31**(3): 377-395.

Robinson, C. T., G. W. Minshall and T. V. Royer (2000). "Inter-annual patterns in macroinvertebrate communities of wilderness streams in Idaho, U.S.A." Hydrobiologia **421**(1): 187-198.

Robinson, C. T., K. Tockner and J. V. Ward (2002). "The fauna of dynamic riverine landscapes." Freshwater Biology **47**(4): 661-677.

Rosillon, D. (1985). "Seasonal variations in the benthos of a chalk trout stream, the River Samson, Belgium." Hydrobiologia **126**(3): 253-262.

Roy, A. H., A. D. Rosemond, M. J. Paul, D. S. Leigh and J. B. Wallace (2003). "Stream macroinvertebrate response to catchment urbanisation (Georgia, U.S.A.)." Freshwater Biology **48**(2): 329-346.

Russell, G. J., J. M. Diamond, S. L. Pimm and T. M. Reed (1995). "A Century of Turnover: Community Dynamics at Three Timescales." Journal of Animal Ecology **64**(5): 628-641.

Scarsbrook, M. R. (2002). "Persistence and stability of lotic invertebrate communities in New Zealand." Freshwater Biology **47**(3): 417-431.

Scarsbrook, M. R., I. K. G. Boothroyd and J. M. Quinn (2000). "New Zealand's National River Water Quality Network: long-term trends in macroinvertebrate communities." New Zealand Journal of Marine and Freshwater Research **34**: 289-302.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke and B. Walker (2001). "Catastrophic shifts in ecosystems." Nature **413**(6856): 591-596.

Simon, H. A. (1962). "The Architecture of Complexity." Proceedings of the American Philosophical Society **106**(6): 467-482.

Snyder, C. D., J. A. Young, R. Villella and D. P. Lemarié (2003). "Influences of upland and riparian land use patterns on stream biotic integrity." Landscape Ecology **18**(7): 647-664.

Southwood, T. R. E. (1977). "Habitat, the Templet for Ecological Strategies?" Journal of Animal Ecology **46**(2): 336-365.

Southwood, T. R. E. (1988). "Tactics, strategies and templets." Oikos **52**(1): 3-18.

Speirs, D. C., W. S. C. Gurney, A. G. Hildrew and J. H. Winterbottom (2000). "Long-term demographic balance in the Broadstone stream insect community." Journal of Animal Ecology **69**(1): 45-58.

Sponseller, R. A., E. F. Benfield and H. M. Valett (2001). "Relationships between land use, spatial scale and stream macroinvertebrate communities." Freshwater Biology **46**(10): 1409-1424.

Statzner, B. and L. A. Beche (2010). "Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems?" Freshwater Biology **55**(s1): 80-119.

Statzner, B., J. A. Gore and V. H. Resh (1988). "Hydraulic Stream Ecology: Observed Patterns and Potential Applications." Journal of the North American Benthological Society **7**(4): 307-360.

Statzner, B. and B. Higler (1986). "Stream hydraulics as a major determinant of benthic invertebrate zonation patterns." Freshwater Biology **16**(1): 127-139.

Statzner, B., K. Hoppenhaus, M.-F. Arens and P. Richoux (1997). "Reproductive traits, habitat use and templet theory: a synthesis of world-wide data on aquatic insects." Freshwater Biology **38**(1): 109-135.

Strayer, D., J. S. Glitzenstein, C. G. Jones, J. Kolasa, G. E. Likens, M. J. McDonnell, G. G. Parker and S. T. A. Prickett (1986). "Long-term ecological studies: an illustrated account of their design, operation, and importance to ecology." Occasional Publication of the Institute of Ecosystem Studies.(2): 1-38.

Swanson, F. J. and R. E. Sparks (1990). "Long-Term Ecological Research and the Invisible Place." BioScience **40**(7): 502-508.

Tachet, H., P. Usseglio-Polatera and C. Roux (1994). "Theoretical habitat templets, species traits, and species richness: Trichoptera in the Upper Rhone River and its floodplain." Freshwater Biology **31**(3): 397-415.

Tipper, J. C. (1979). "Rarefaction and rarefiction - the use and abuse of a method in paleontology." Paleobiology **5**: 423-434.

Townsend, C. R., S. Doledec and M. R. Scarsbrook (1997). "Species traits in relation to temporal and spatial heterogeneity in streams: a test of habitat templet theory." Freshwater Biology **37**(2): 367-387.

Townsend, C. R., B. J. Downes, K. Peacock and C. J. Arbuckle (2004). "Scale and the detection of land-use effects on morphology, vegetation and macroinvertebrate communities of grassland streams." Freshwater Biology **49**(4): 448-462.

Townsend, C. R. and A. G. Hildrew (1976). "Field Experiments on the Drifting, Colonization and Continuous Redistribution of Stream Benthos." Journal of Animal Ecology **45**(3): 759-772.

Townsend, C. R. and A. G. Hildrew (1994). "Species Traits in Relation to a Habitat Templet for River Systems." Freshwater Biology **31**(3): 265-275.

Townsend, C. R., A. G. Hildrew and K. Schofield (1987). "Persistence of Stream Invertebrate Communities in Relation to Environmental Variability." Journal of Animal Ecology **56** (2): 597-613.

Townsend, C. R., M. R. Scarsbrook and S. Doledec (1997). "The Intermediate Disturbance Hypothesis, Refugia and Biodiversity in Streams." Limnology and Oceanography **42**(5): 938-949.

Trigo, R. M., D. P. Vazquez, T. J. Osborn, Y. Castro-Diez, S. Gamiz-Fortis and M. J. Esteban-Parra (2004). "North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian peninsula." International Journal of Climatology **24**: 925-944.

Urban, M. C., D. K. Skelly, D. Burchsted, W. Price and S. Lowry (2006). "Stream communities across a rural-urban landscape gradient." Diversity and Distributions **12**(4): 337-350.

Usseglio-Polatera, P. and J.-N. Beisel (2002). "Longitudinal changes in macroinvertebrate assemblages in the meuse river: anthropogenic effects versus natural change." River Research and Applications **18**: 197-211.

Usseglio-Polatera, P., M. Bournaud, P. Richoux and H. Tachet (2000). "Biological and ecological traits of benthic freshwater macroinvertebrates: relationships and definition of groups with similar traits." Freshwater Biology **43**(2): 175-205.

Usseglio-Polatera, P., P. Richoux, M. Bournaud and H. Tachet (2001). "A functional classification of benthic macroinvertebrates based on biological and ecological traits: application to river condition assessment and stream management." Arch. Hydrobiol. **139**(1): 53-83.

Usseglio-Polatera, P. and H. Tachet (1994). "Theoretical habitat templets, species traits, and species richness: Plecoptera and Ephemeroptera in the Upper Rhone River and its floodplain." Freshwater Biology **31**(3): 357-375.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell and C. E. Cushing (1980). "The River Continuum Concept." Canadian Journal of Fisheries and Aquatic Sciences **37**: 131-136.

Vannote, R. L. and B. W. Sweeney (1980). "Geographic Analysis of Thermal Equilibria: A Conceptual Model for Evaluating the Effect of Natural and Modified Thermal Regimes on Aquatic Insect Communities." The American Naturalist **115**(5): 667-695.

Vieira, N. K. M., W. H. Clements, L. S. Guevara and B. F. Jacobs (2004). "Resistance and resilience of stream insect communities to repeated hydrologic disturbances after a wildfire." Freshwater Biology **49**(10): 1243-1259.

Vitousek, P. M., H. A. Mooney, J. Lubchenco and J. M. Melillo (1997). "Human Domination of Earth's Ecosystems." Science **277**: 494-499.

Voelz, N. J., S.-H. Shieh and J. V. Ward (2000). "Long-term monitoring of benthic macroinvertebrate community structure: a perspective from a Colorado river." Aquatic Ecology **34**(3): 261-278.

Wagner, R., T. Dapper and H. H. Schmidt (2000). "The influence of environmental variables on the abundance of aquatic insects: a comparison of ordination and artificial neural networks." Hydrobiologia **422-423**(1): 143-152.

Wagner, R. and H.-H. Schmidt (2004). "Yearly discharge patterns determine species abundance and community diversity: Analysis of a 25 year record from the Breitenbach." Arch. Hydrobiol. **161**(4): 511-540.

Wallace, I. D., B. Wallace and G. N. Philipson (2003). Keys to the Case-bearing Caddis Larvae of Britain and Ireland. Ambleside, Cumbria, UK, Freshwater Biological Associtation.

Wallace, J. B. (1990). "Recovery of Lotic Macroinvertebrate Communities from Disturbance." Environmental Management **14**(5): 605-620.

Walsh, C. J., A. H. Roy, J. W. Feminella, P. D. Cottingham, P. M. Groffman and R. P. M. II (2005). "The urban stream syndrome: current knowledge and the search for a cure." Journal of the North American Benthological Society **24**(3): 706-723.

Walsh, C. J., K. A. Waller, J. Gehling and R. M. Nally (2007). "Riverine invertebrate assemblages are degraded more by catchment urbanisation than by riparian deforestation." Freshwater Biology **52**(3): 574-587.

Ward, J. V. and J. A. Stanford (1982). "Thermal Responses in the Evolutionary Ecology of Aquatic Insects." Annual Review of Entomology **27**(1): 97-117.

Weatherhead, P. J. (1986). "How Unusual are Unusual Events?" The American Naturalist **128**(1): 150-154.

Weatherley, N. S. and S. J. Ormerod (1990). "The Constancy of Invertebrate Assemblages in Soft-Water Streams: Implications for the Prediction and Detection of Environmental Change." Journal of Applied Ecology **27**(3): 952-964.

Webb, J. A. and E. L. King (2009). "A Bayesian hierarchical trend analysis finds strong evidence for large-scale temporal declines in stream ecological condition around Melbourne, Australia." Ecography **32**(2): 215-225.

Weiher, E. and P. Keddy (1999). Assembly rules as general constraints on community composition. Ecological Assembly Rules. Perspectives, advances, retreats. P. Keddy and E. Weiher, Cambridge University Press: 251-271.

White, P. S. and A. Jentsch (2001). "The search for generality in studies of disturbance and ecosystem dynamics." Progress in Botany **62**: 399-450.

Whittingham, M. J., P. A. Stephens, R. B. Bradbury and R. P. Freckleton (2006). "Why do we still use stepwise modelling in ecology and behaviour?" Journal of Animal Ecology **75**(5): 1182-1189.

Wiens, J. A. (2002). "Riverine landscapes: taking landscape ecology into the water." Freshwater Biology **47**(4): 501-515.

Wiley M.J, S. L. Kohler and P. W. Seelbach (1997). "Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams." Freshwater Biology **37**(1): 133-148.

Wilson, M. A. and S. R. Carpenter (1999). " Economic valuation of freshwater ecosystem services in the United States: 1971 to 1997." Ecological Applications **9**(3): 772-783.

Winterbourn, M. J. (1997). New Zealand mountain stream communities: Stable yet disturbed? Evolutionary Ecology of Freshwater Animals: concepts and case studies. B. Streit, T. Stadler and C. M. Lively. Basel, Switzerland.

Wolda, H. (1988). "Insect Seasonality: Why?" Annual Review of Ecology and Systematics **19**: 1-18.

Woodward, G., J. I. Jones and A. G. Hildrew (2002). " Community persistence in Broadstone Stream (U.K.) over three decades." Freshwater Biology **47**(8): 1419-1435.

Wright, J. F. (2000). An introduction to RIVPACS. Assessing the biological quality of fresh waters: RIVPACS and other techniques. J. F. Wright, D. W. Sutcliffe and M. T. Furse. Ambleside, The Freshwater Biological Association: 1-24.

Wu, J. and O. L. Loucks (1995). "From Balance of Nature to Hierarchical Patch Dynamics: A Paradigm Shift in Ecology." The Quarterly Review of Biology **70**(4): 439.

Yachi, S. and M. Loreau (1999). "Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis." Proceedings of the

National Academy of Sciences of the United States of America **96**(4): 1463-1468.

Zasada, K. A. and E. H. Smith (1981). Freshwater Invertebrates of the Sheffield District. Sorby Record Special Series. **4**.

APPENDIX 1

Appendix 1 relates to chapter 5 and provides examples of community groups formed using differing clustering methods and numbers of end groups. A1.1 – A1.5 uses agglomerative clustering with Ward linkage and end groups ranging in size from 8 to 20 communities. A1.6 – A1.10 uses agglomerative clustering with Complete linkage and end groups ranging in size from 8 to 20 communities. Finally, A1.11 – A1.15 uses divisive clustering and end groups ranging in size from 8 to 20 communities.

In all instances similar spatial patterns emerge representing communities clustering to the west and north east, to the south and north east and finally in the centre of the study area. This spatial structure in communities corresponds to increasing levels of temporal variability. At low numbers of end groups Divisive clustering fails to separate out communities into this spatial pattern however when the number of end groups increases to 15-20 a clear separation into areas of high temporal variation in the centre of the country and low variation to the west and north East emerges.

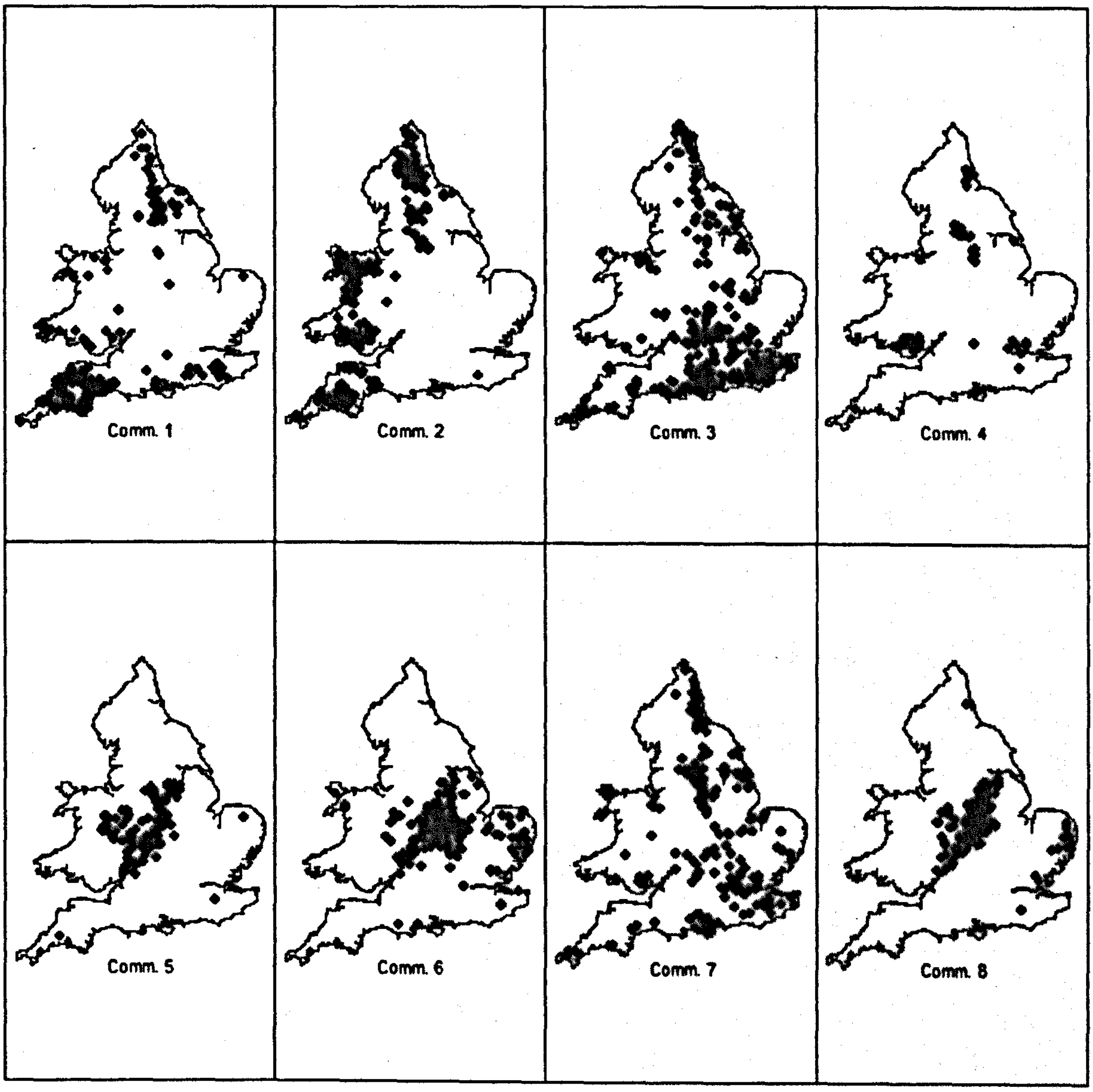


Figure A1.1(a): Spatial structure of communities clustered into 8 end groups using Ward Linkage.

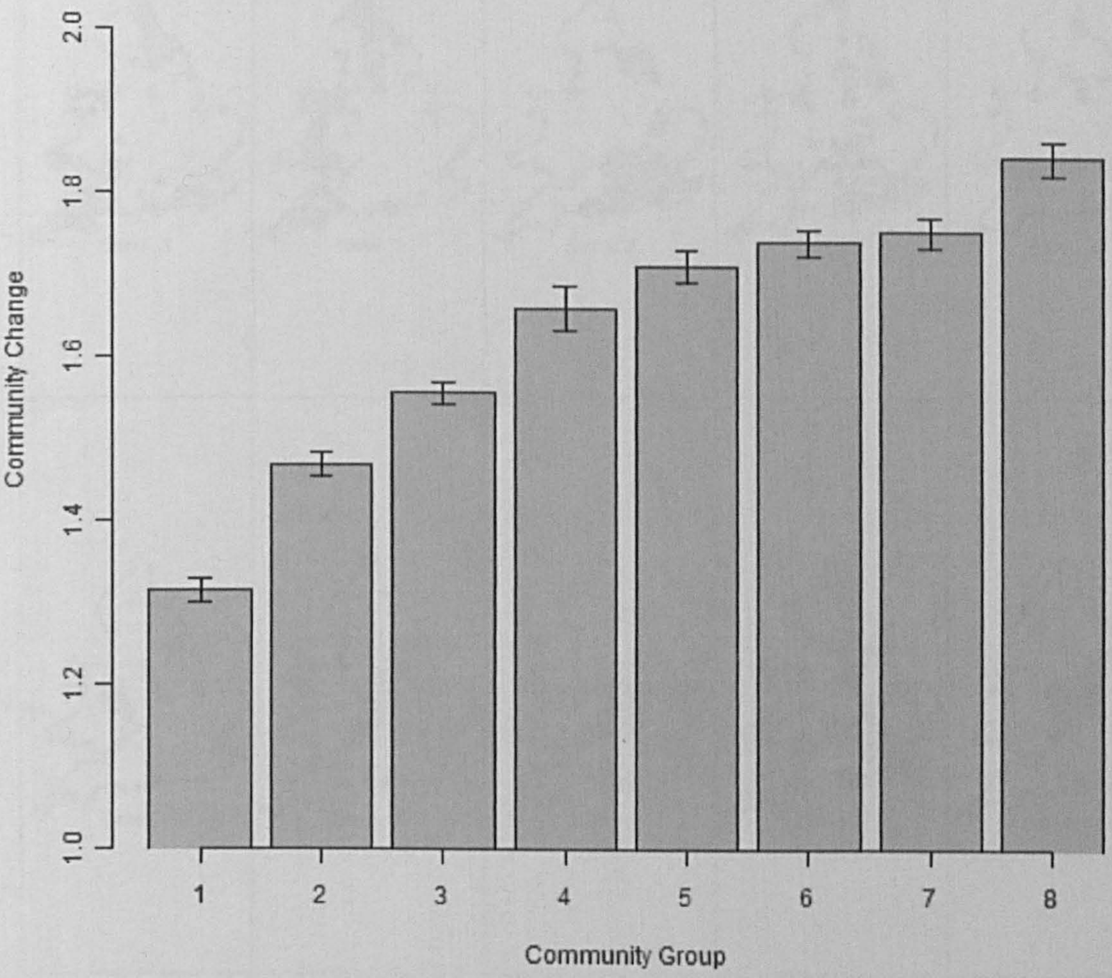


Figure A1.1(b): Average community change for communities clustered into 8 end groups using Ward Linkage.

Figure A1.2(a): Spatial structure of communities clustered into 10 end groups using Ward Linkage.

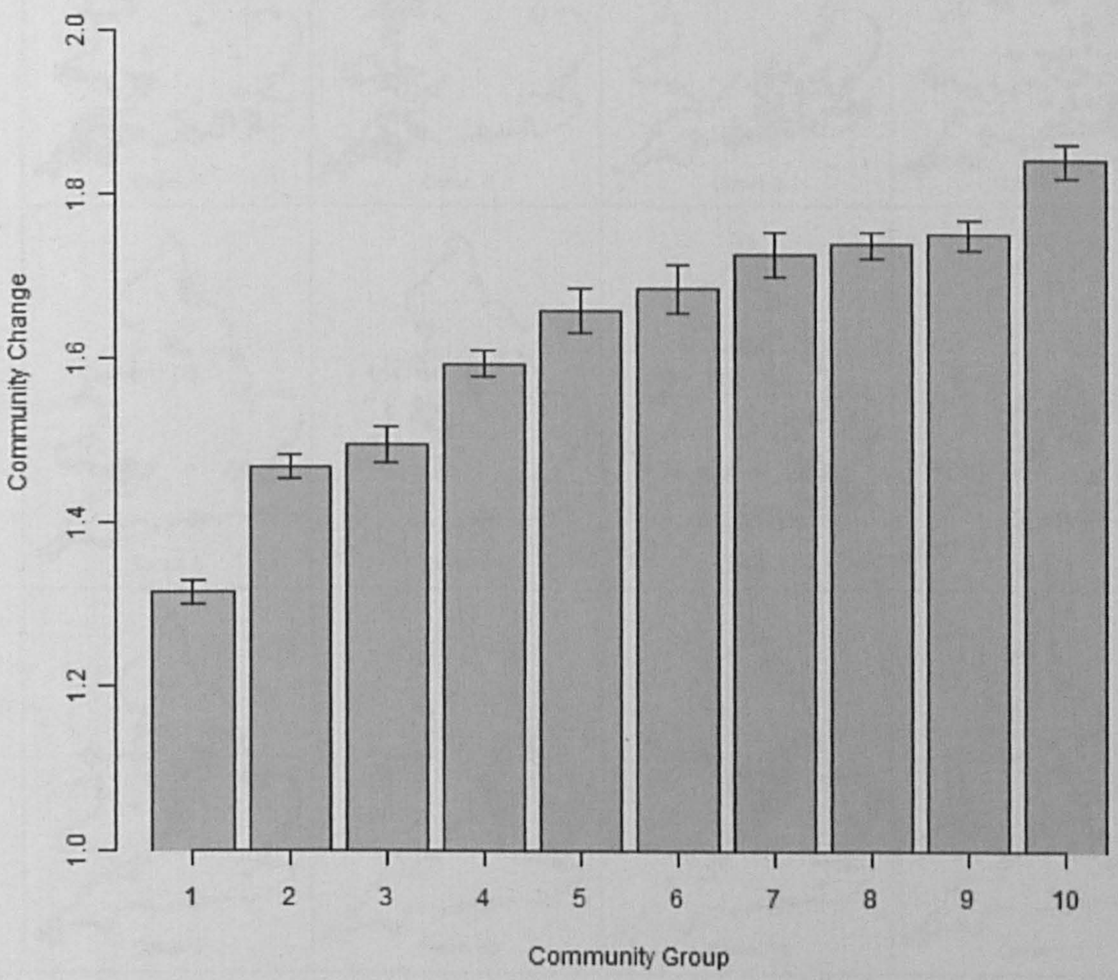
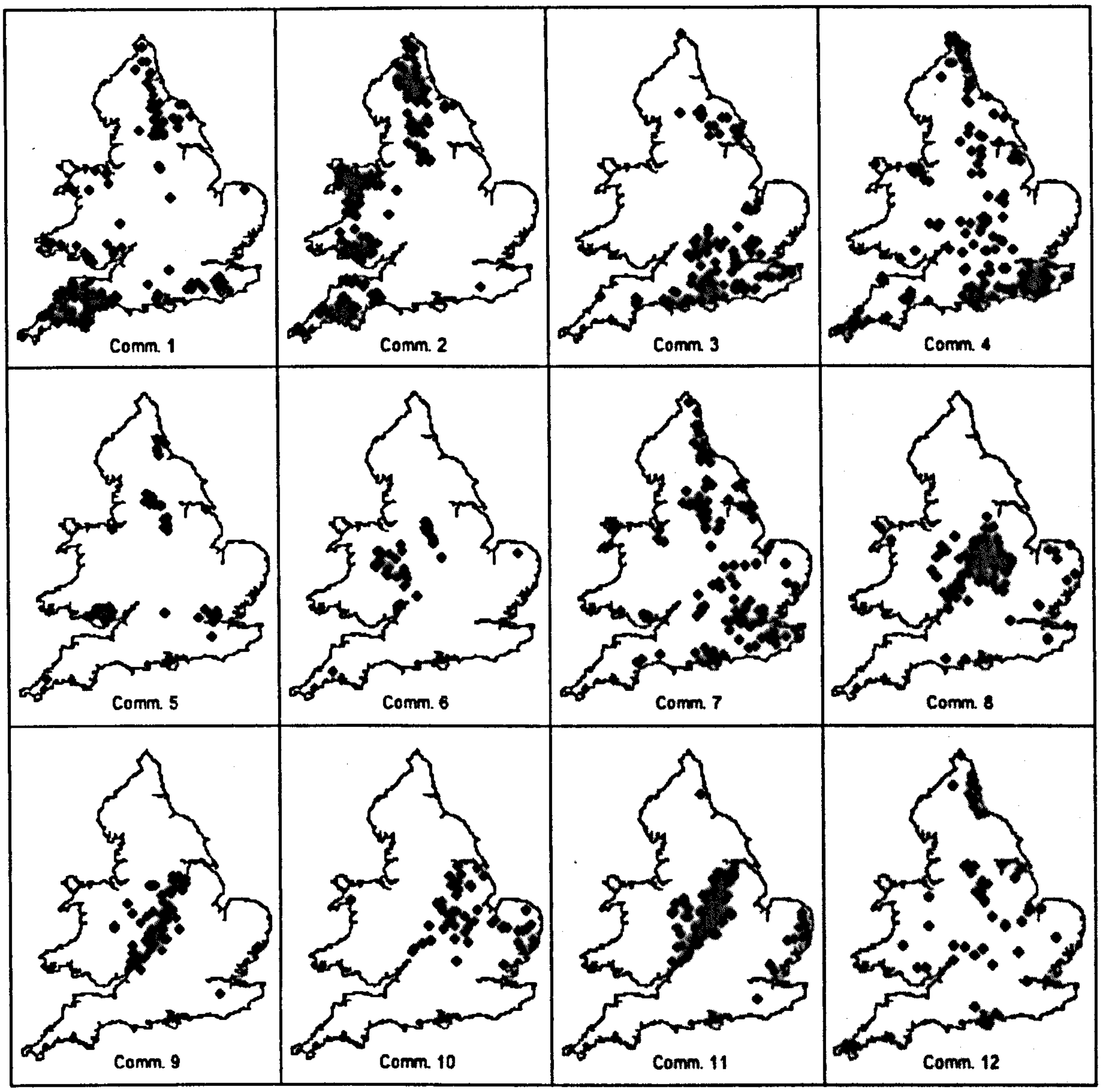


Figure A1.2(b): Average community change for communities clustered into 10 end groups using Ward Linkage.



FigureA1.3(a): Spatial structure of communities clustered into 12 end groups using Ward Linkage.

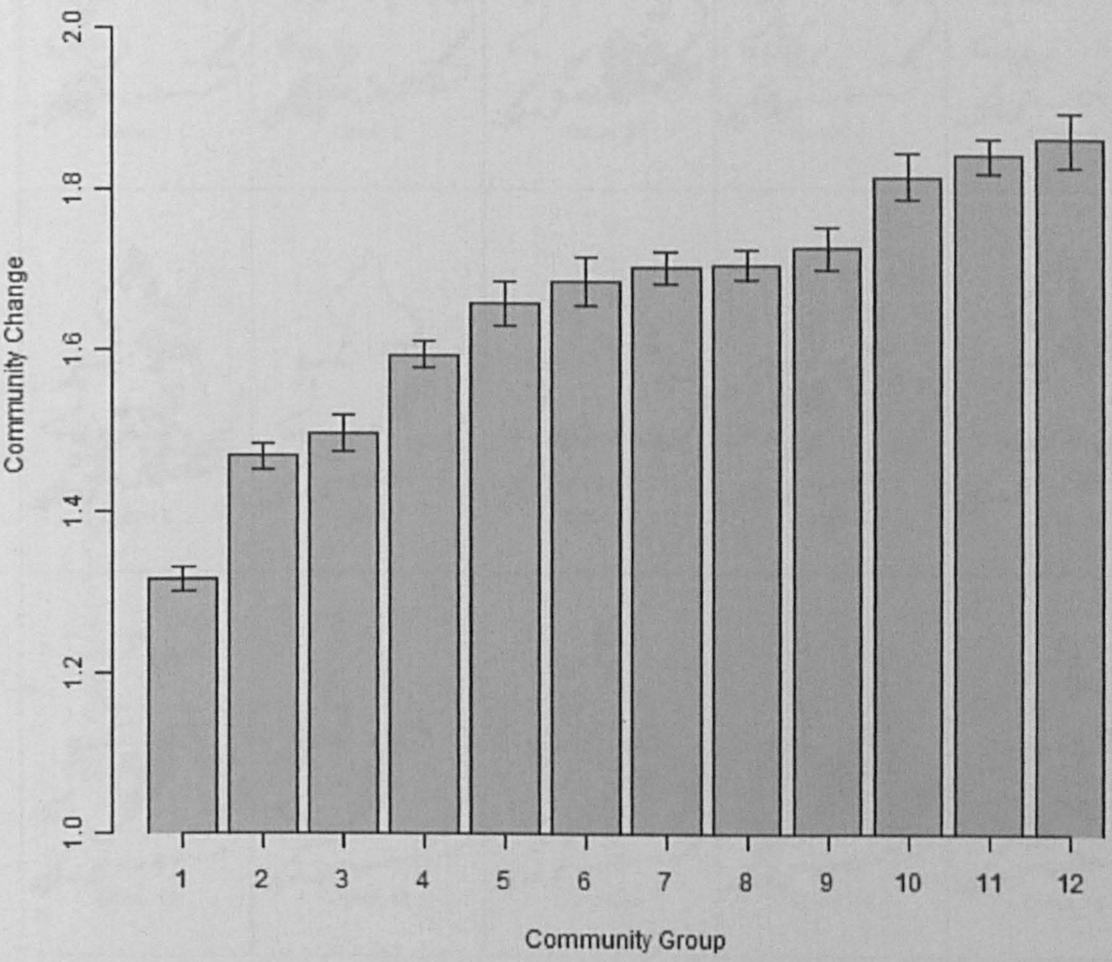


Figure A1.3(b): Average community change for communities clustered into 12 end groups using Ward Linkage.

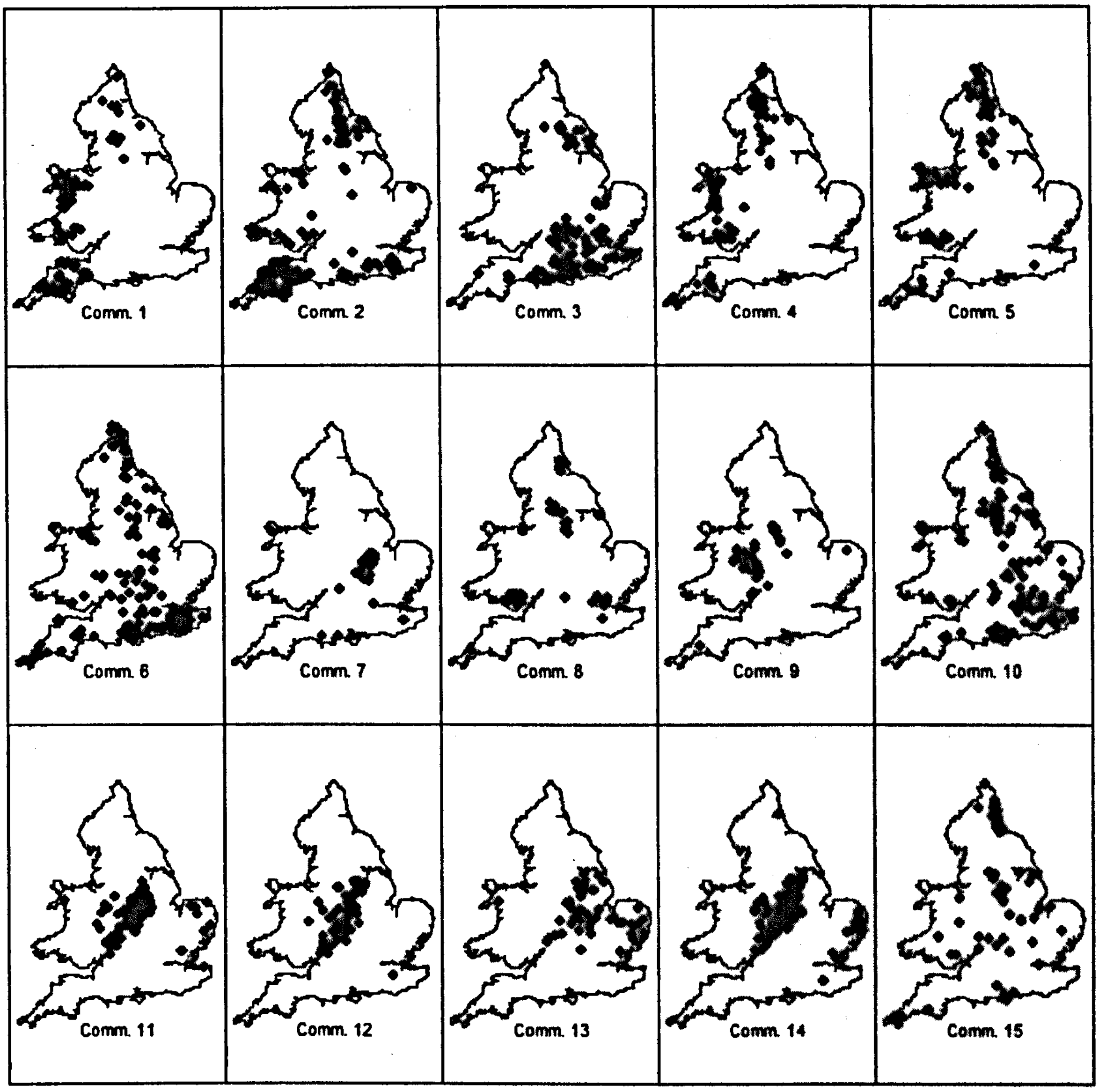


Figure A1.4(a): Spatial structure of communities clustered into 15 end groups using Ward Linkage.

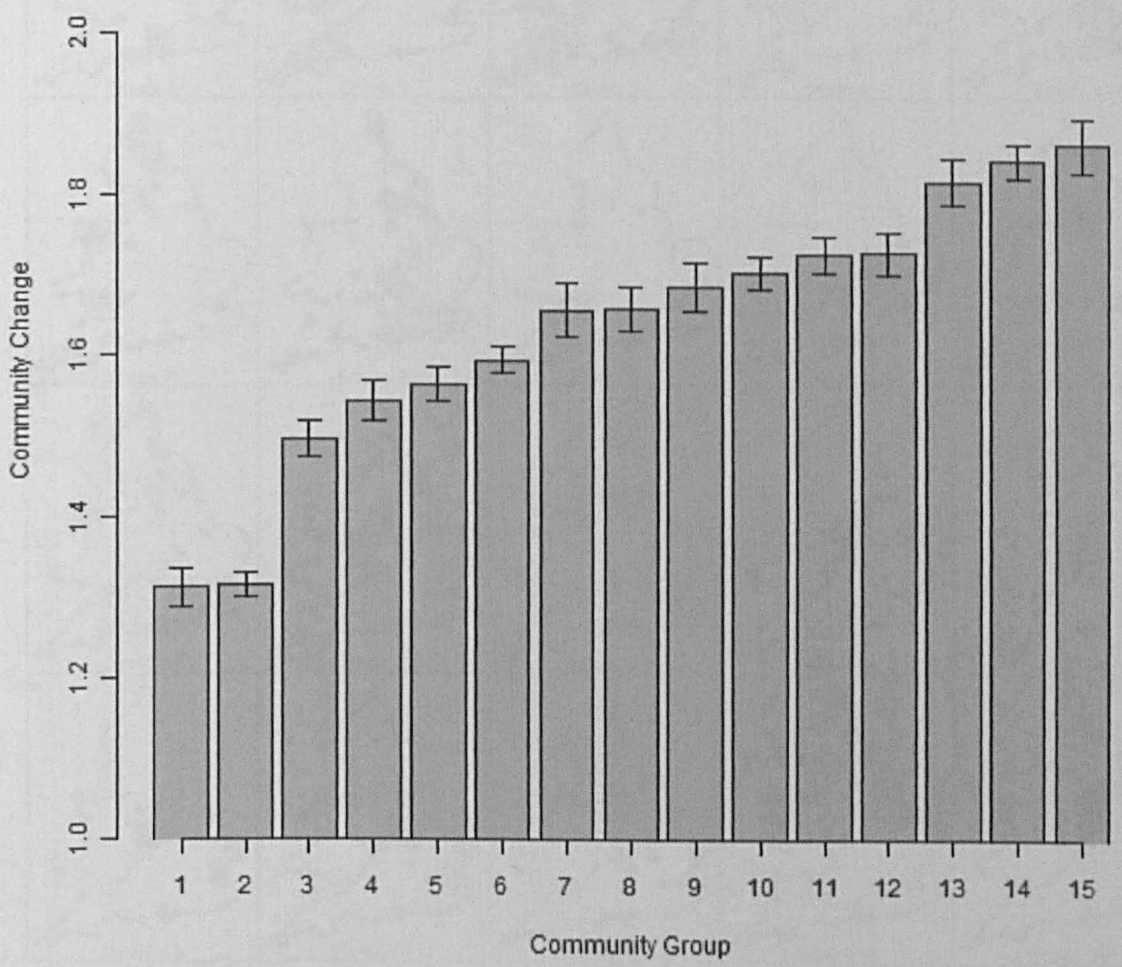


Figure A1.4(b): Average community change for communities clustered into 15 end groups using Ward Linkage.

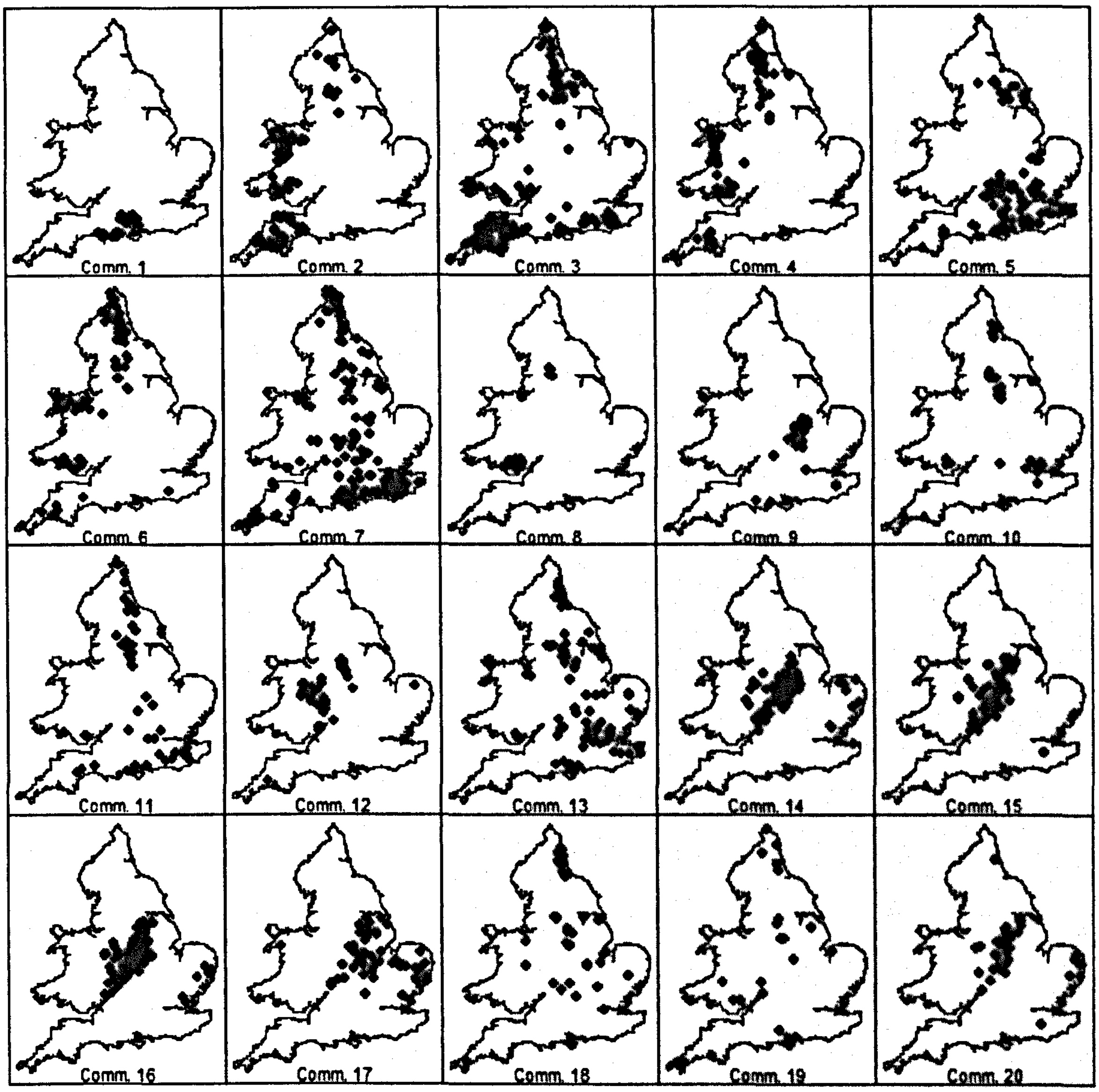


Figure A1.5(a): Spatial structure of communities clustered into 20 end groups using Ward Linkage.

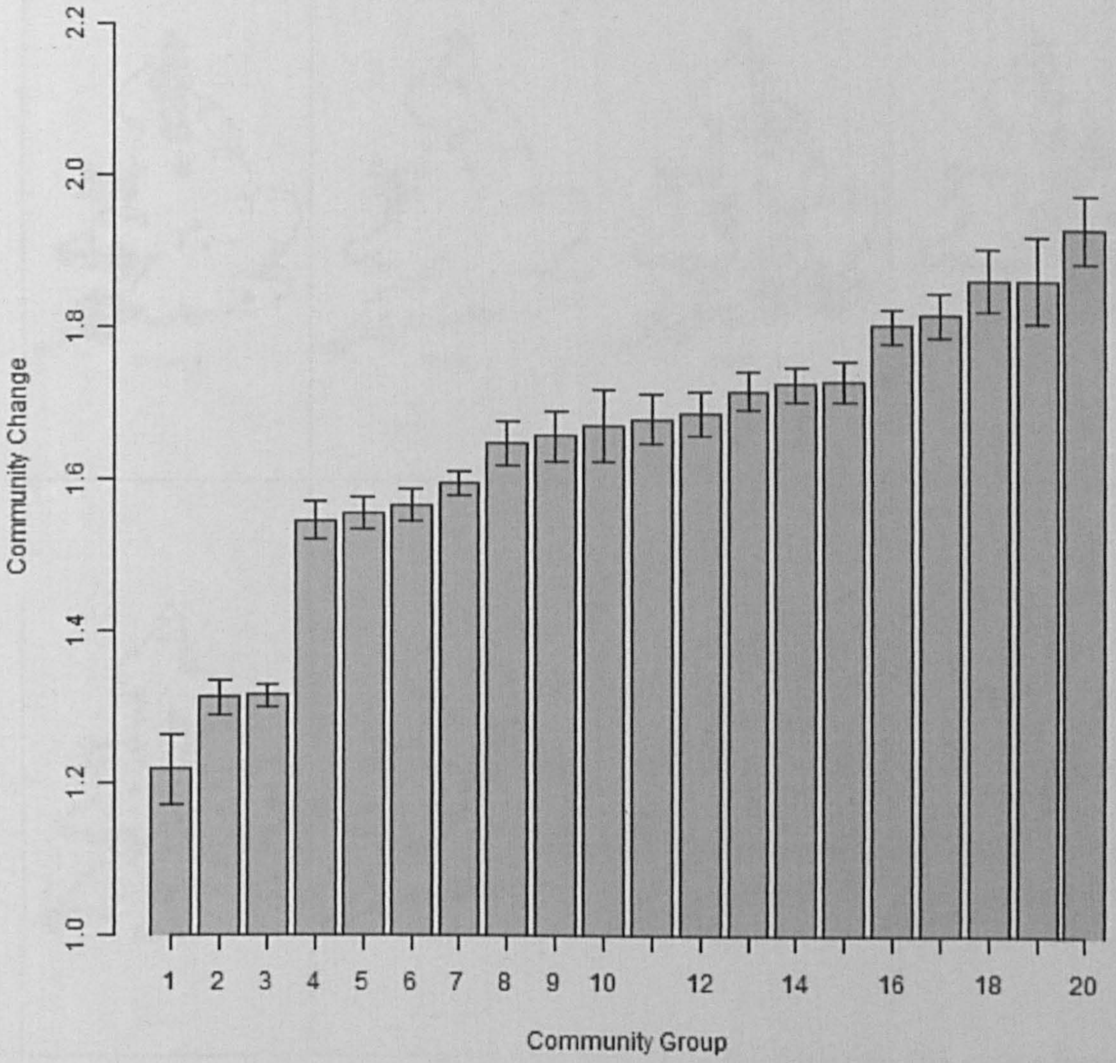


Figure A1.5(b): Average community change for communities clustered into 20 end groups using Ward Linkage.

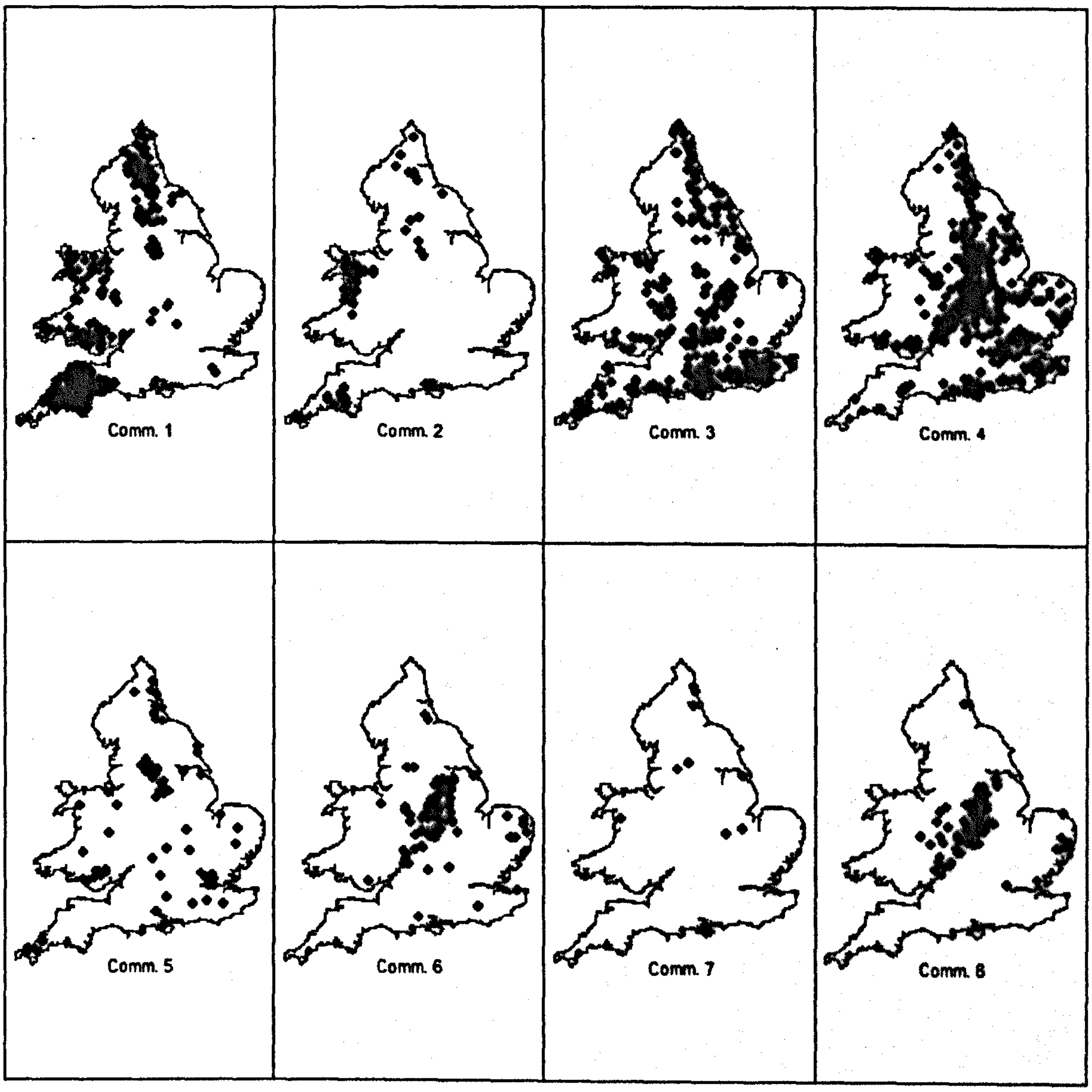


Figure A1.6(a): Spatial structure of communities clustered into 8 end groups using Complete Linkage.

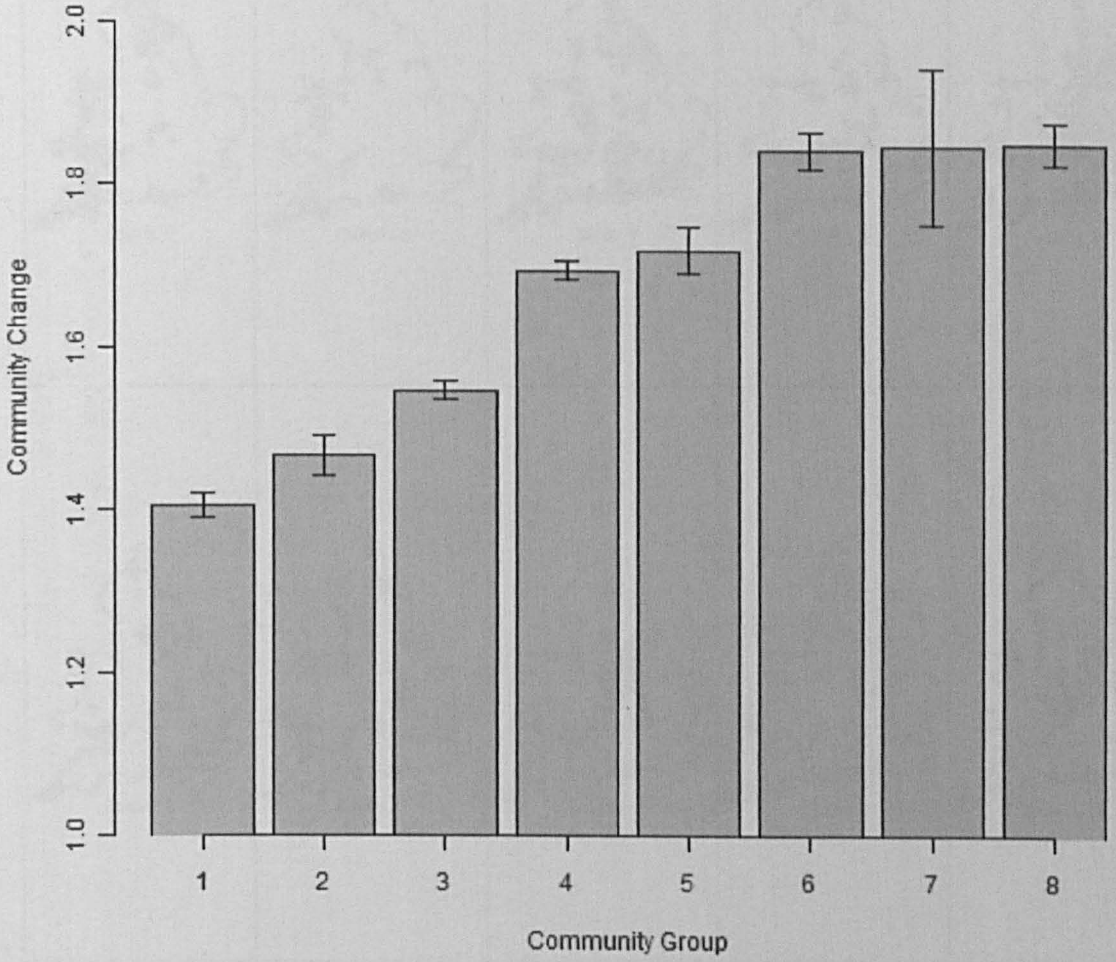


Figure A1.6(b): Average community change for communities clustered into 8 end groups using Complete Linkage.

Figure A1.7(a): Spatial structure of communities clustered into 10 end groups using Complete Linkage.

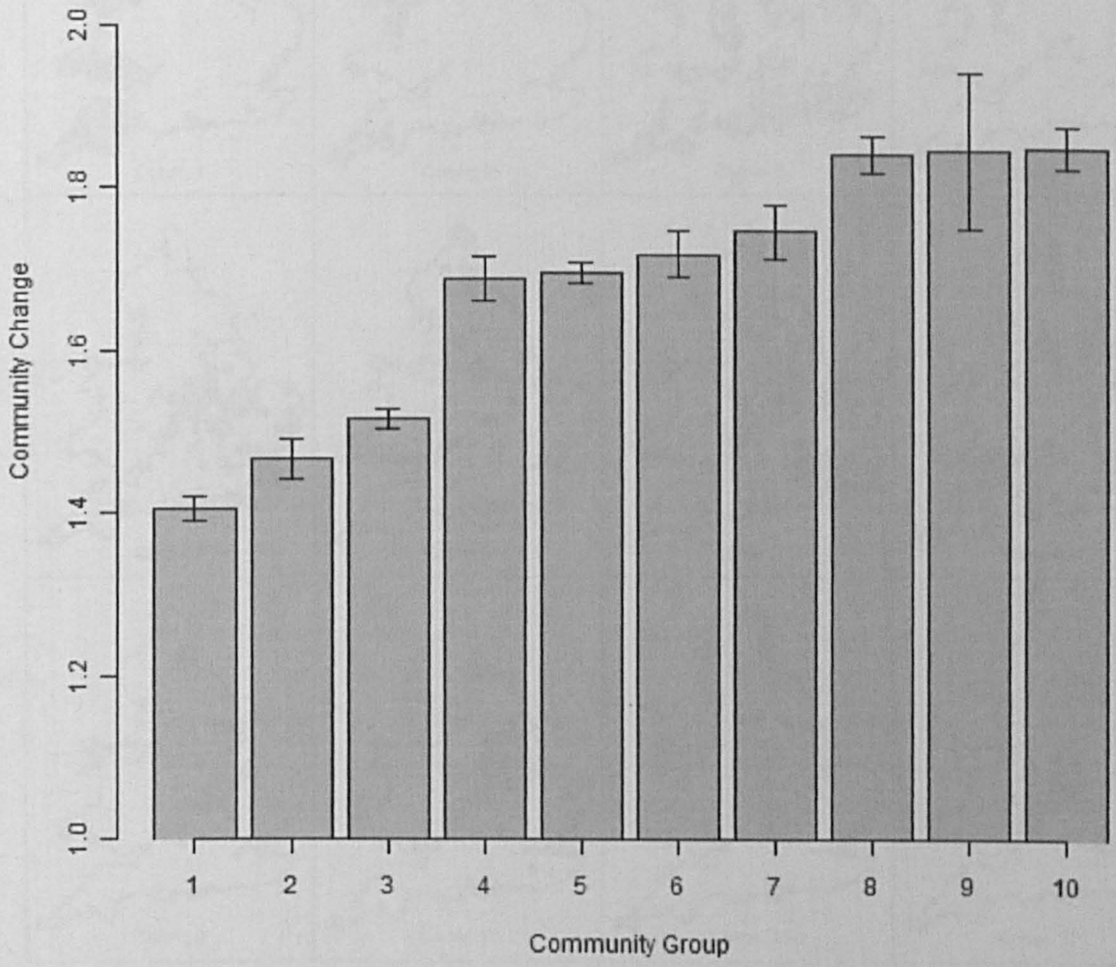


Figure A1.7(b): Average community change for communities clustered into 10 end groups using Complete Linkage.

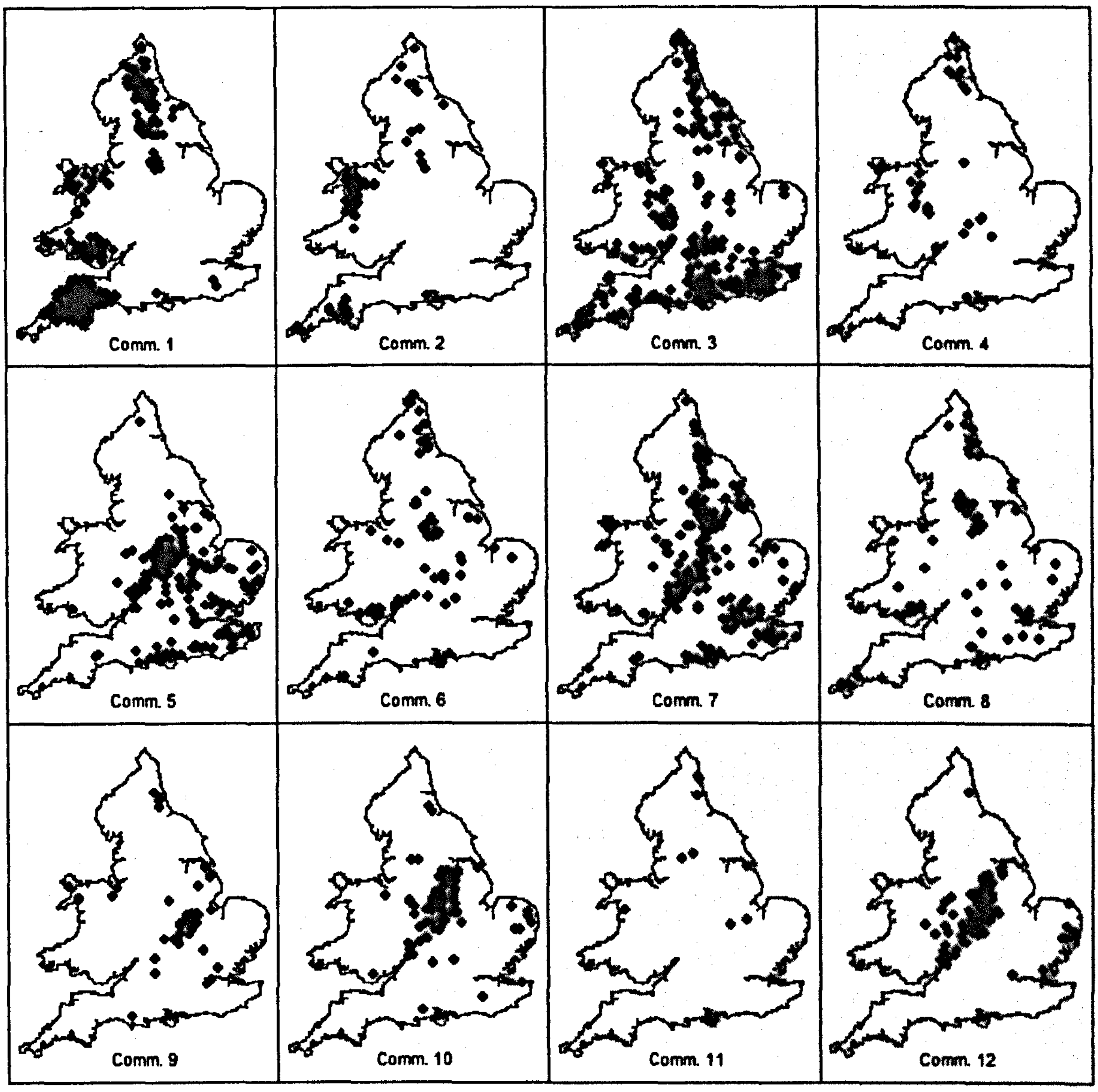


Figure A1.8(a): Spatial structure of communities clustered into 12 end groups using Complete Linkage.

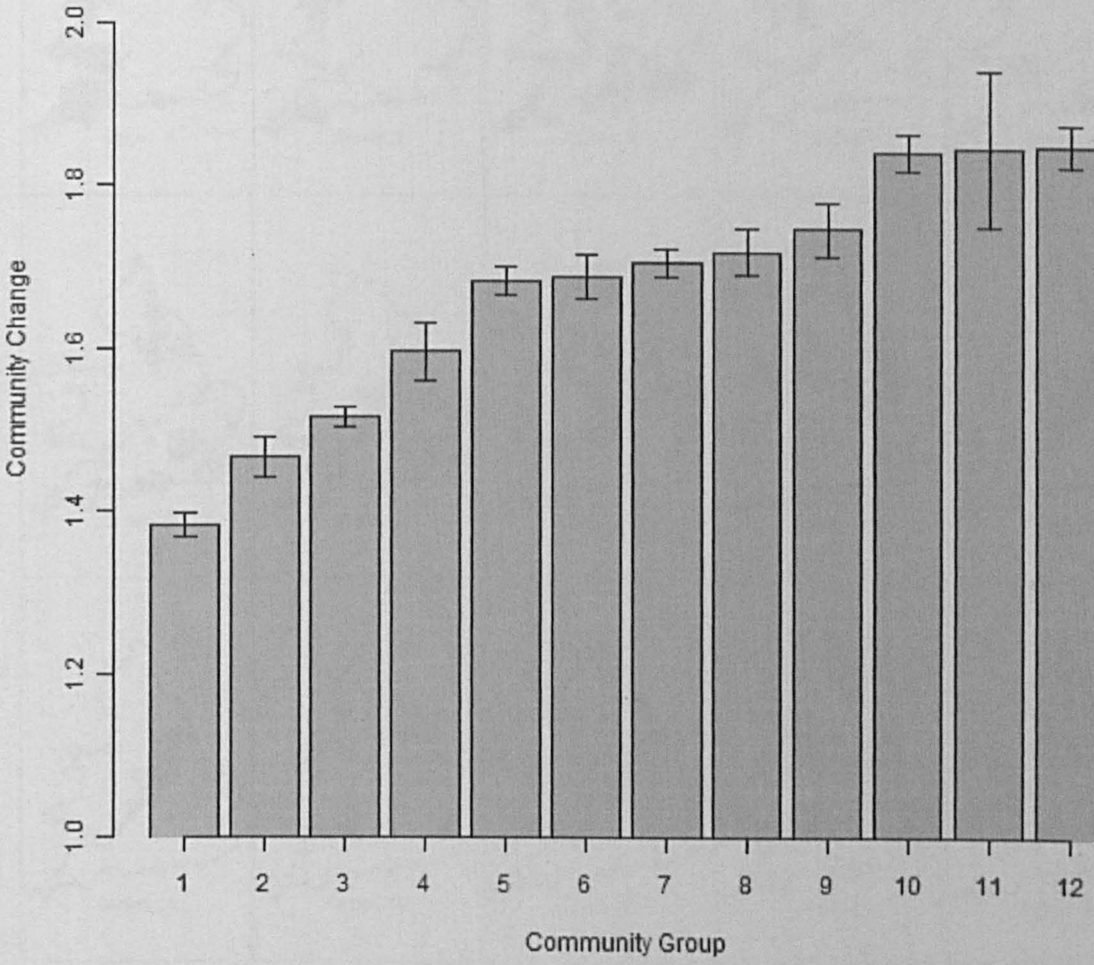


Figure A1.8(b): Average community change for communities clustered into 12 end groups using Complete Linkage.

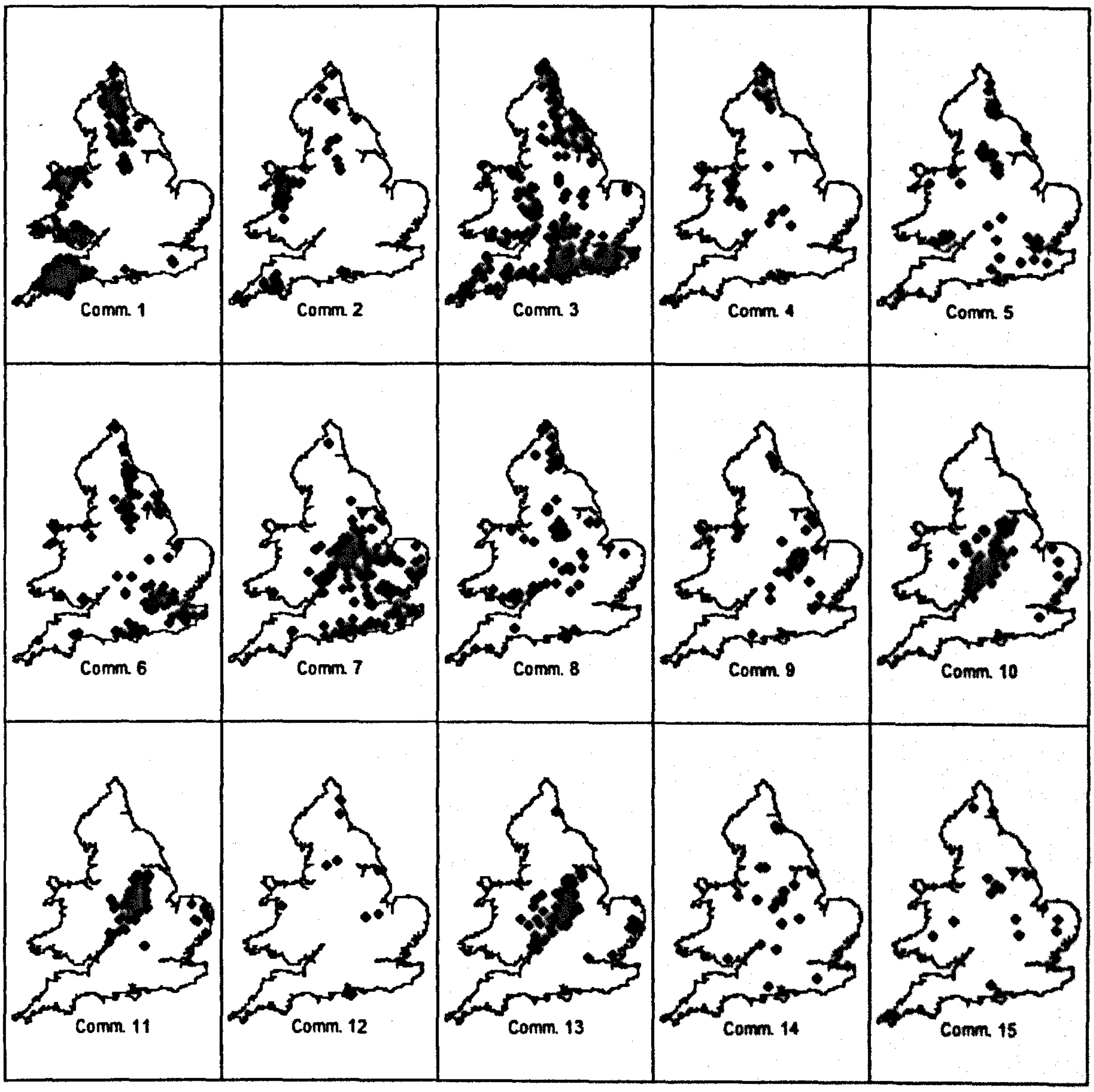


Figure A1.9(a): Spatial structure of communities clustered into 15 end groups using Complete Linkage.

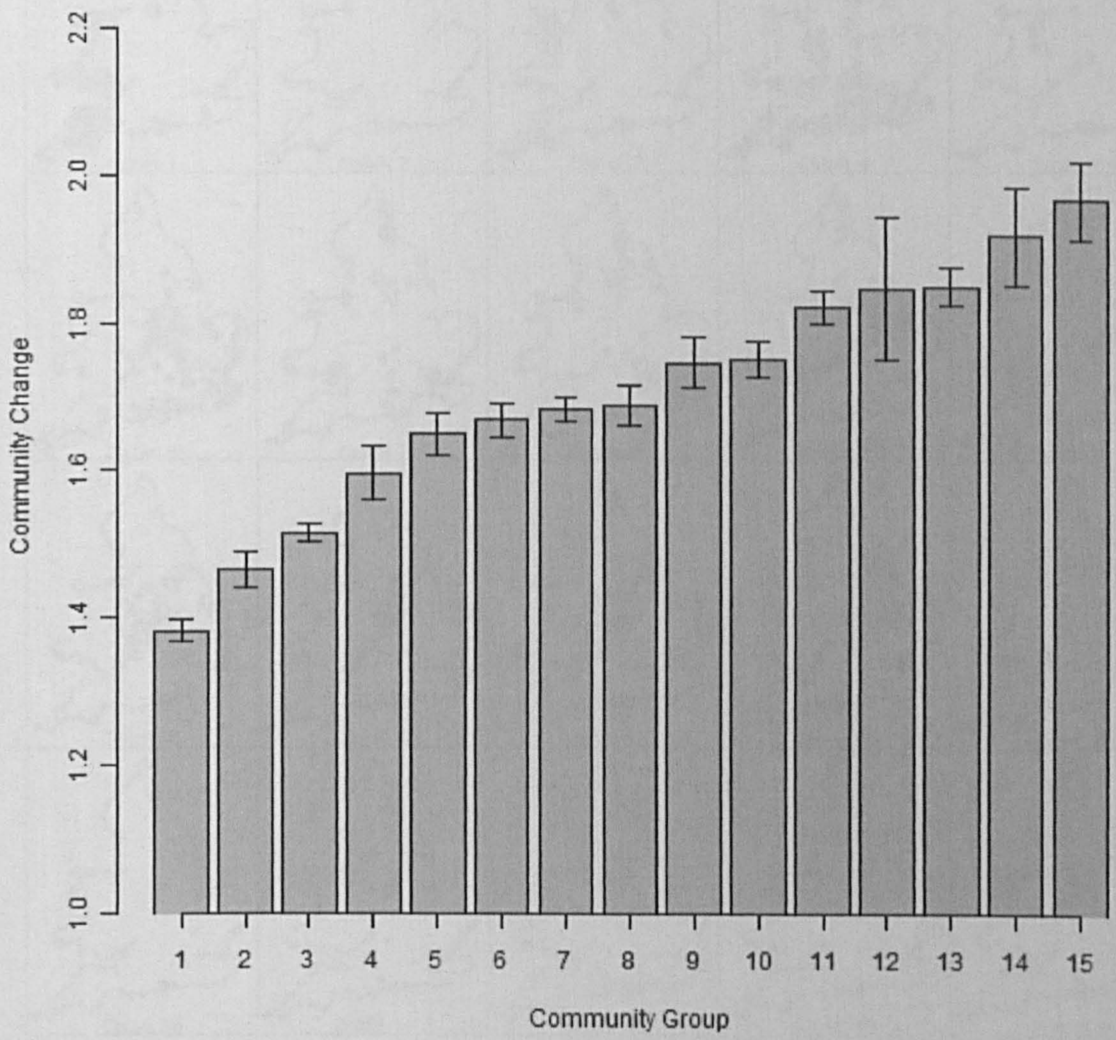


Figure A1.9(b): Average community change for communities clustered into 15 end groups using Complete Linkage.

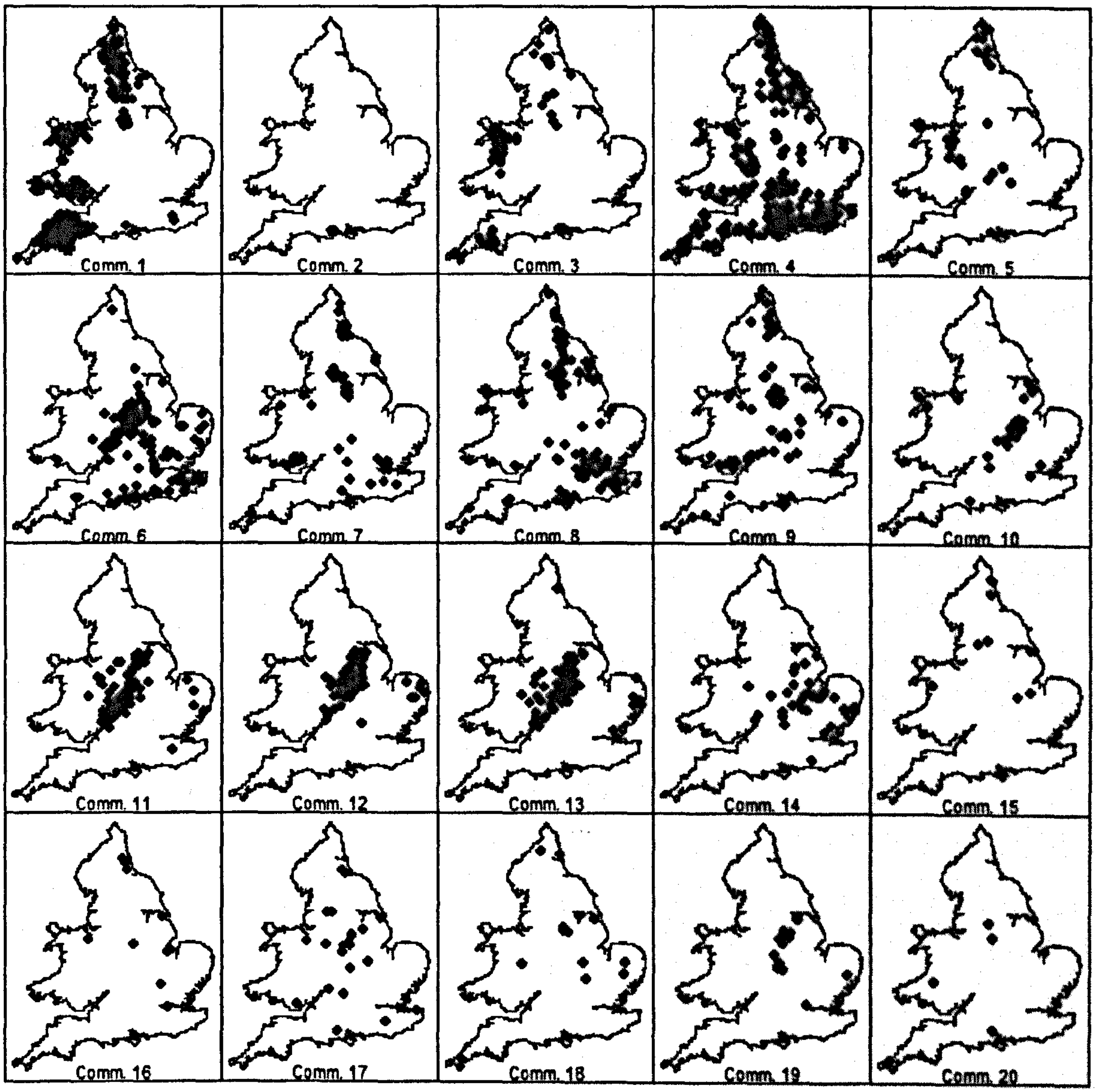


Figure A1.10(a): Spatial structure of communities clustered into 20 end groups using Complete Linkage.

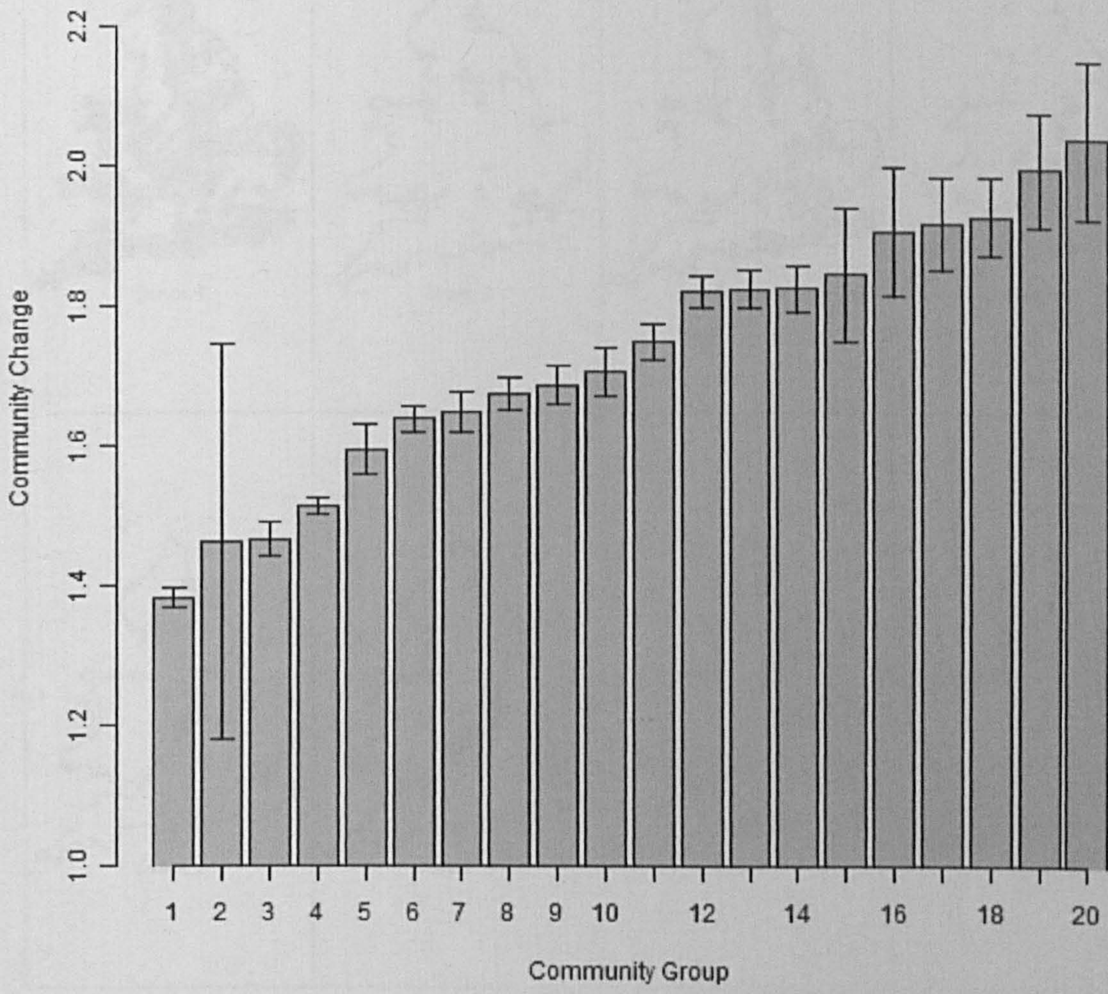


Figure A1.10(b): Average community change for communities clustered into 20 end groups using Complete Linkage.

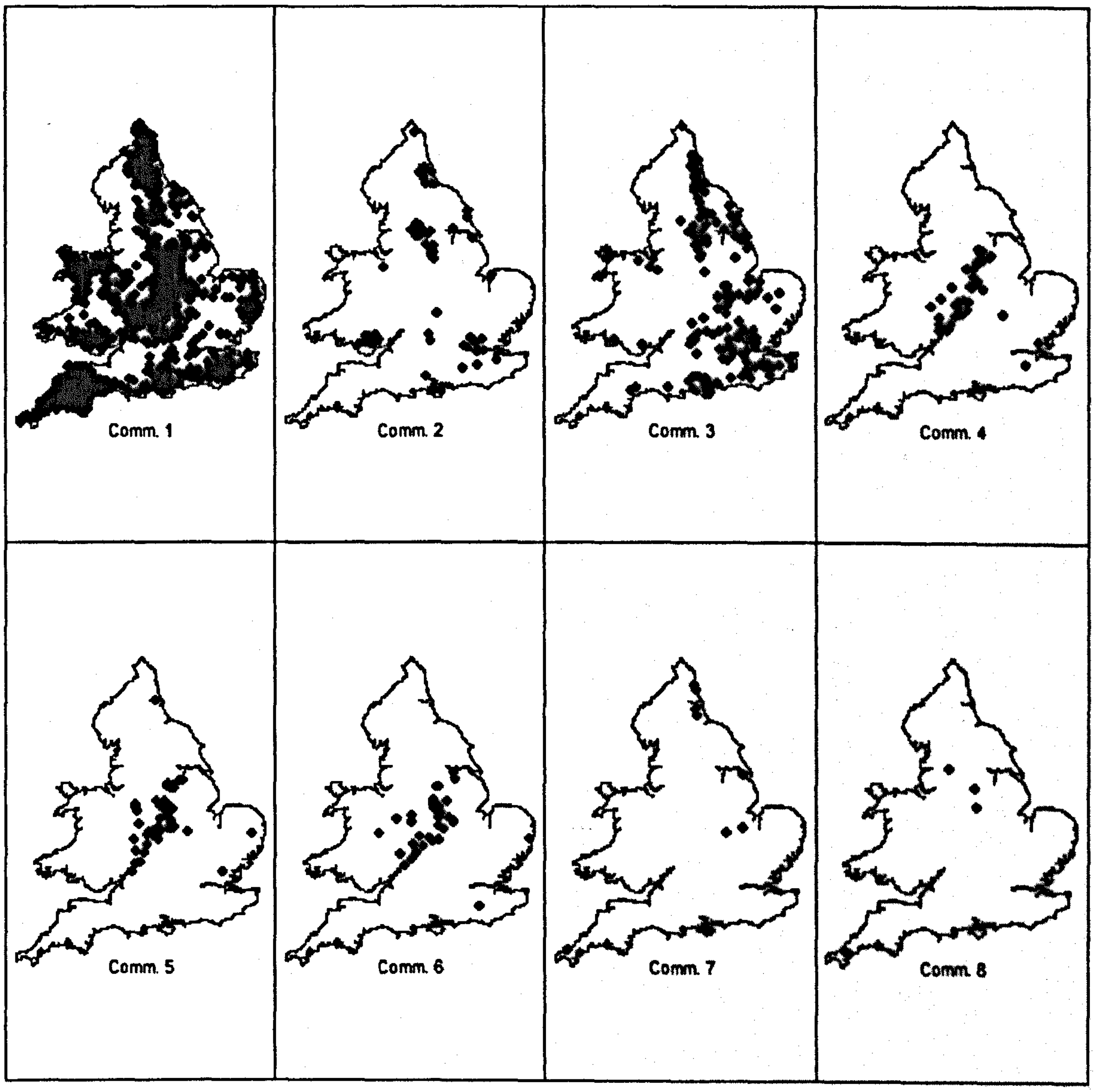


Figure A1.11(a): Spatial structure of communities clustered into 8 end groups using Divisive Clustering.

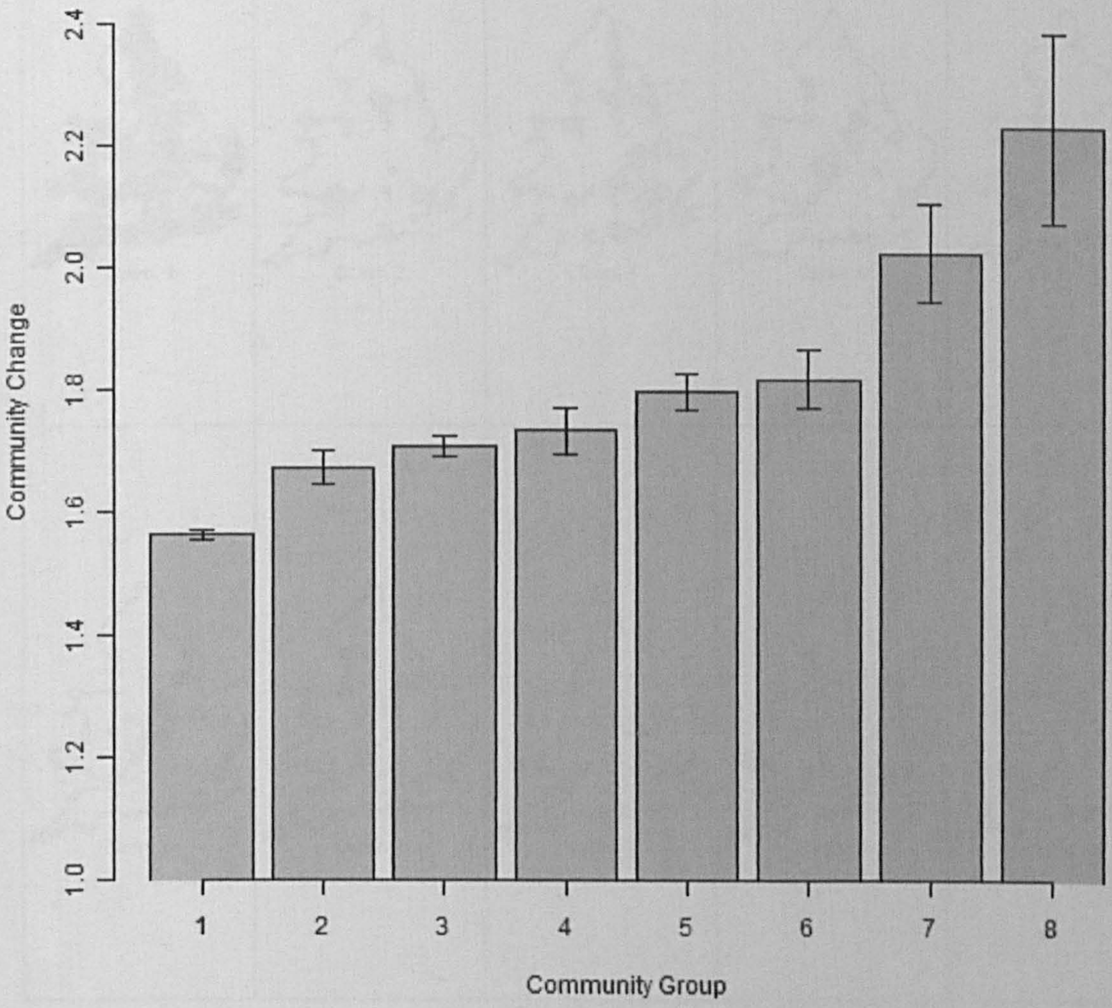


Figure A1.11(b): Average community change for communities clustered into 8 end groups using Divisive Clustering.

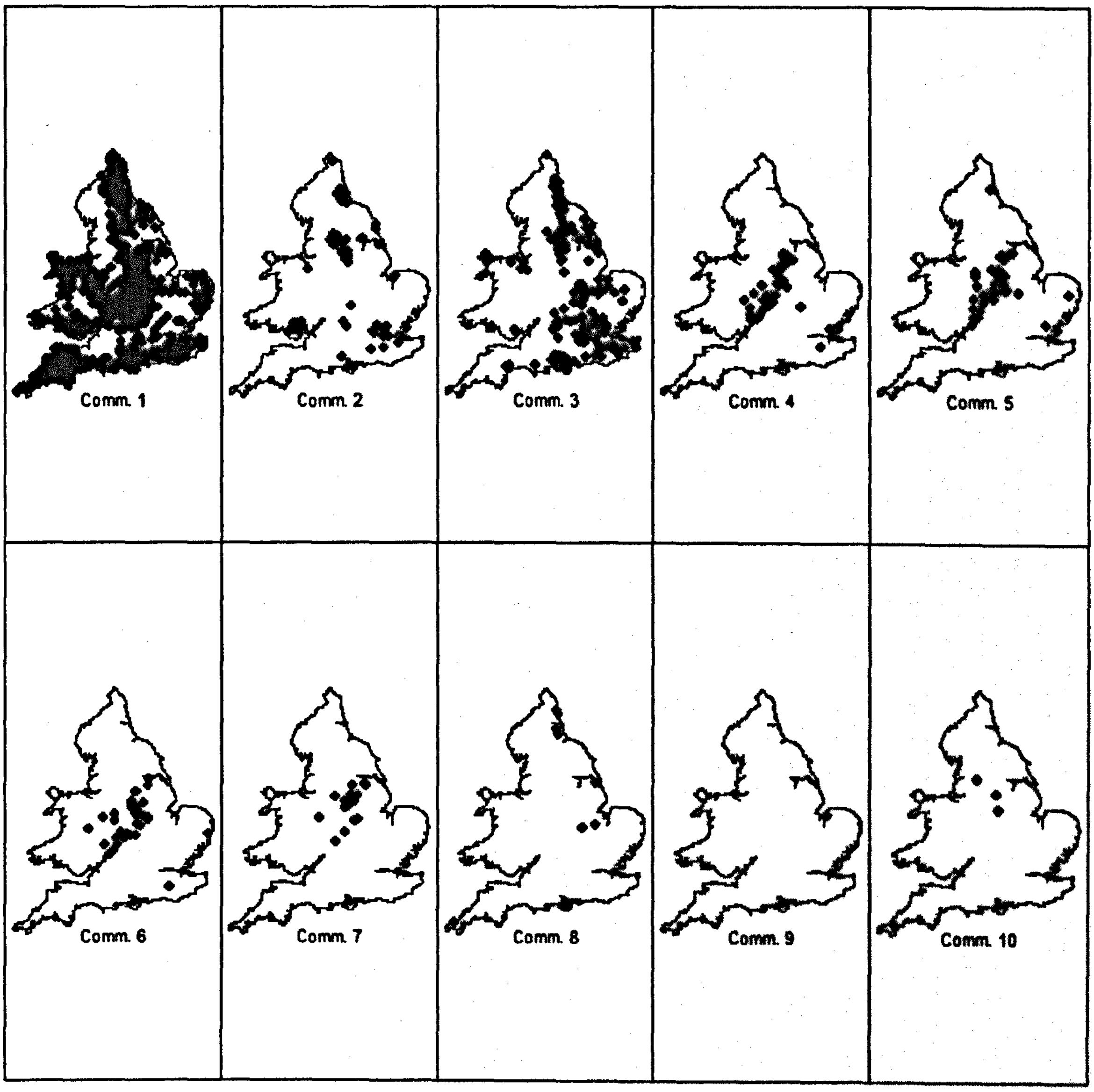


Figure A1.12(a): Spatial structure of communities clustered into 10 end groups using Divisive Clustering.

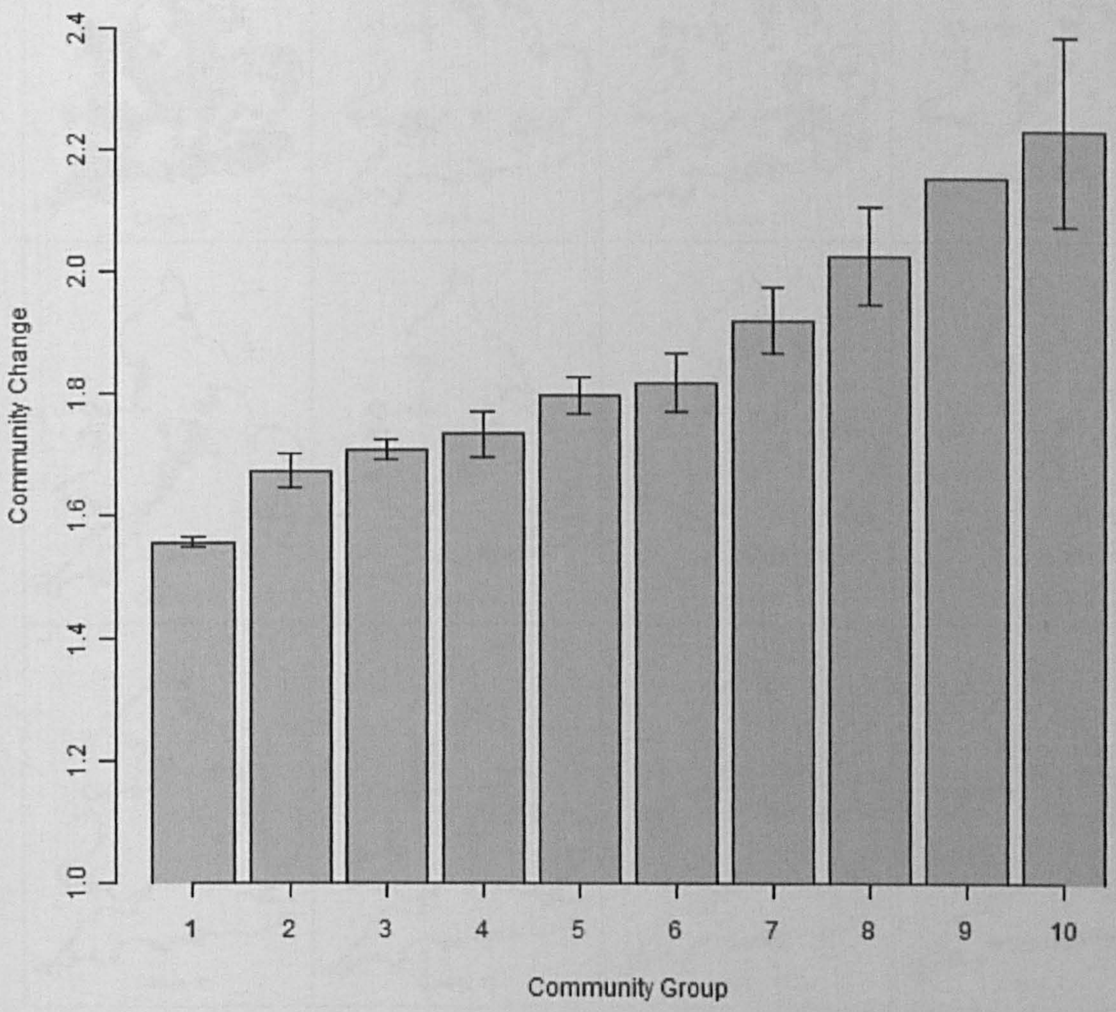


Figure A1.12(b): Average community change for communities clustered into 10 end groups using Divisive Clustering.

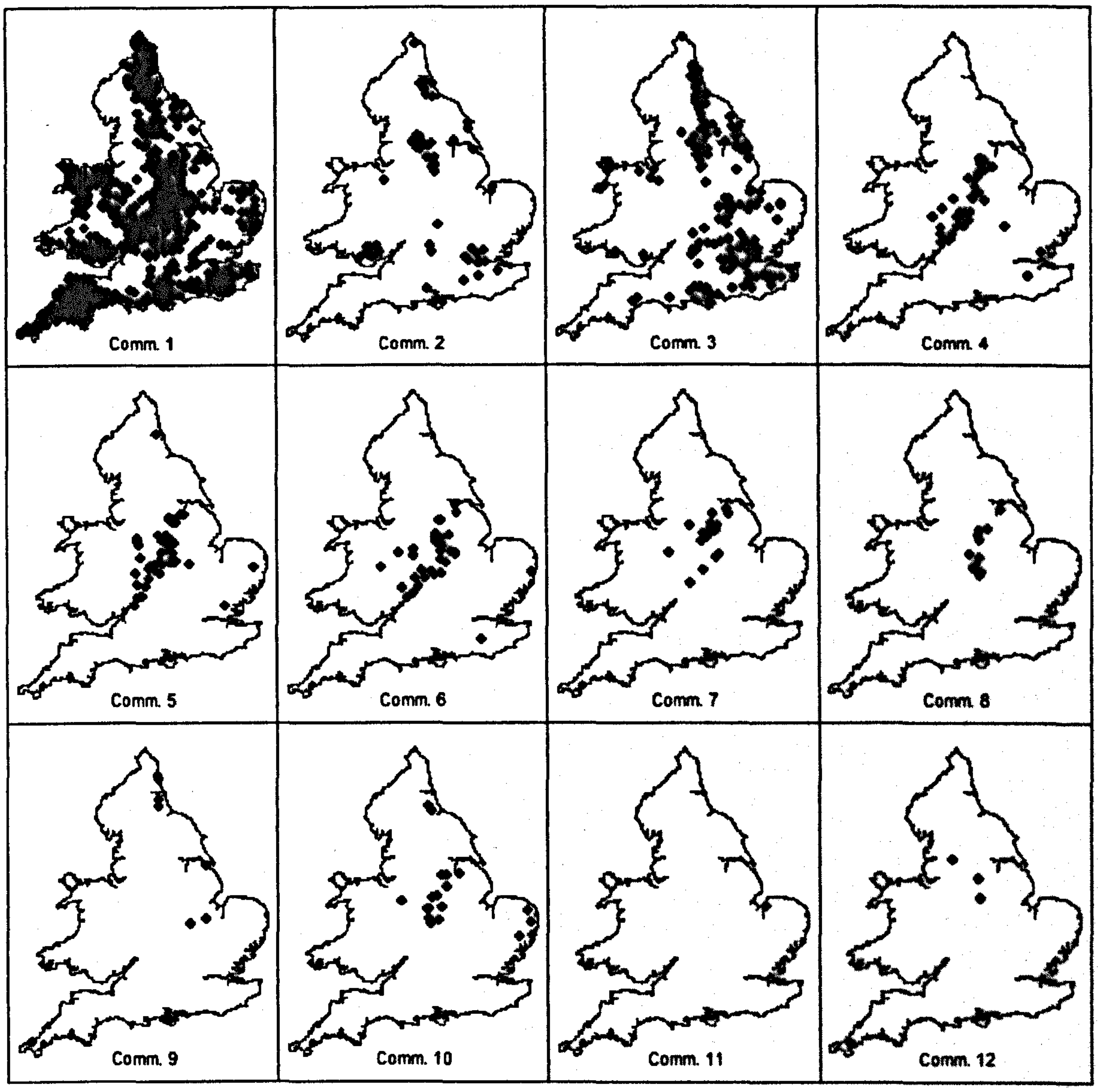


Figure A1.13(a): Spatial structure of communities clustered into 12 end groups using Divisive Clustering.

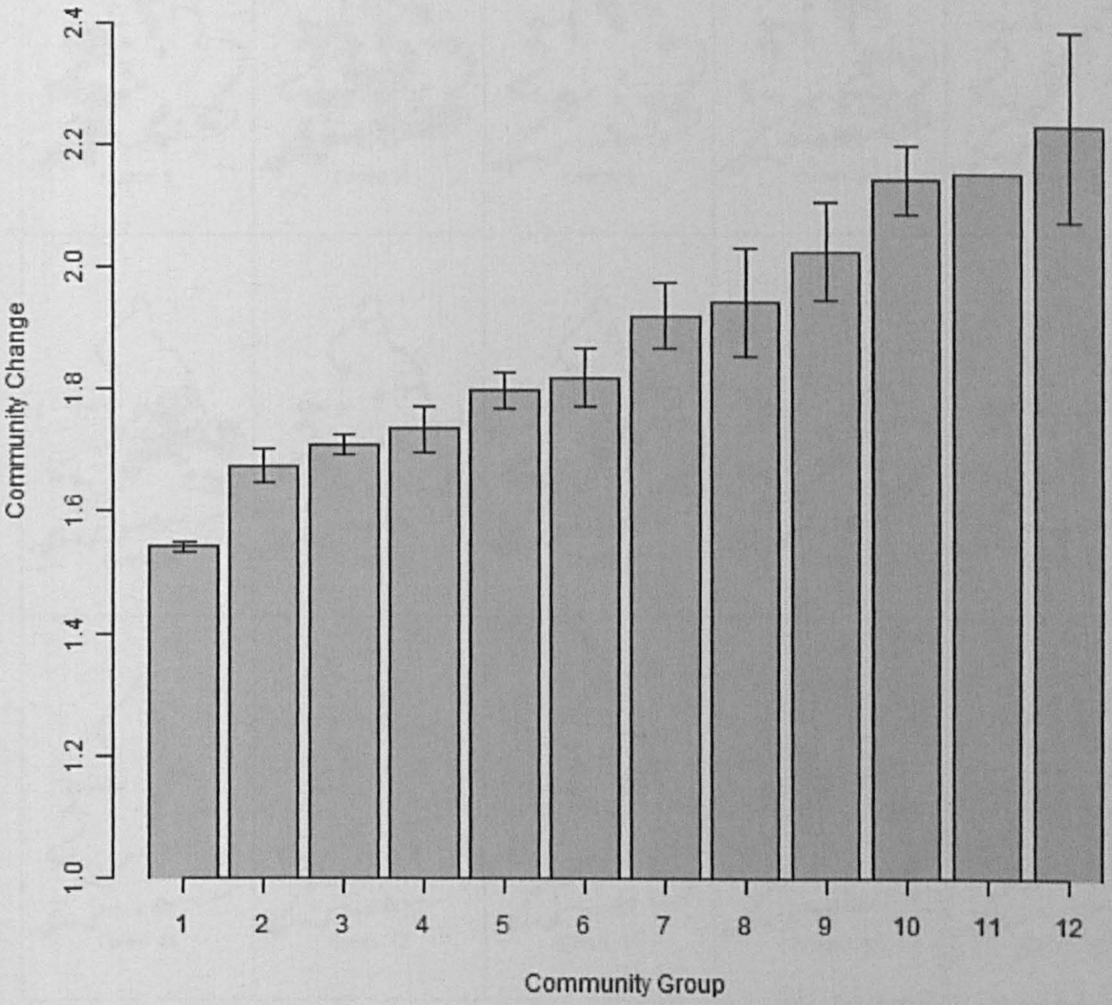


Figure A1.13(b): Average community change for communities clustered into 12 end groups using Divisive Clustering.

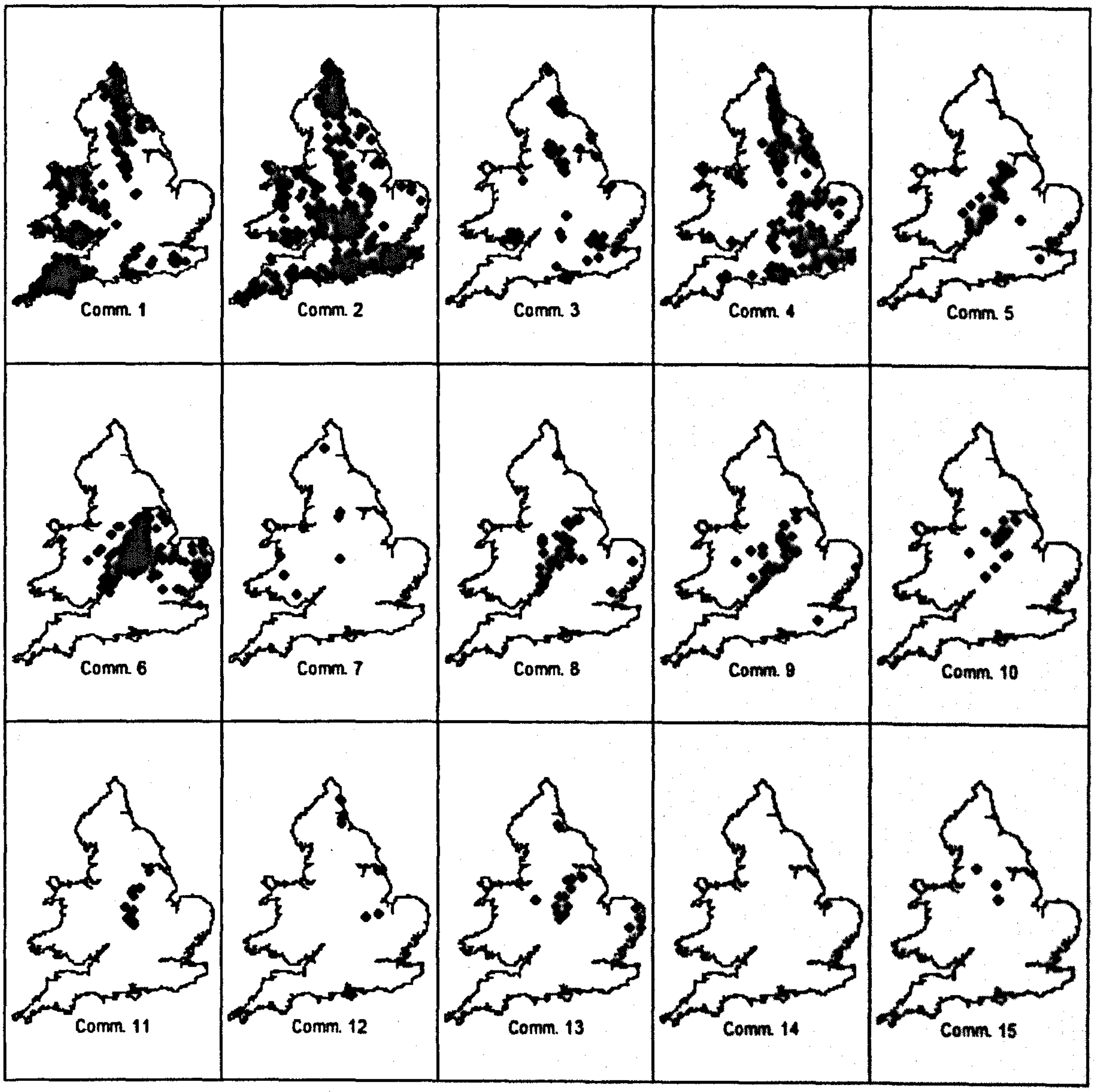


Figure A1.14(a): Spatial structure of communities clustered into 15 end groups using Divisive Clustering.

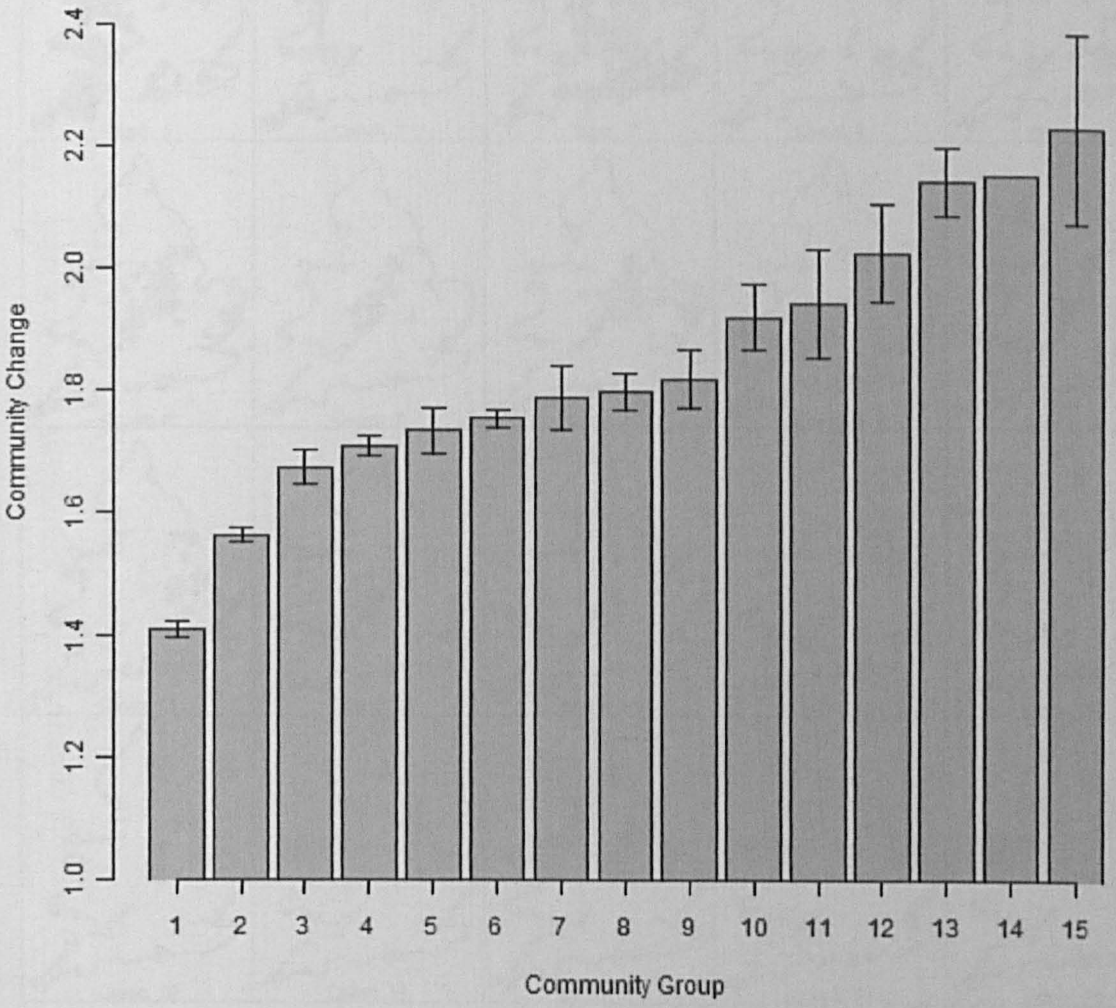


Figure A1.14(b): Average community change for communities clustered into 15 end groups using Divisive Clustering.

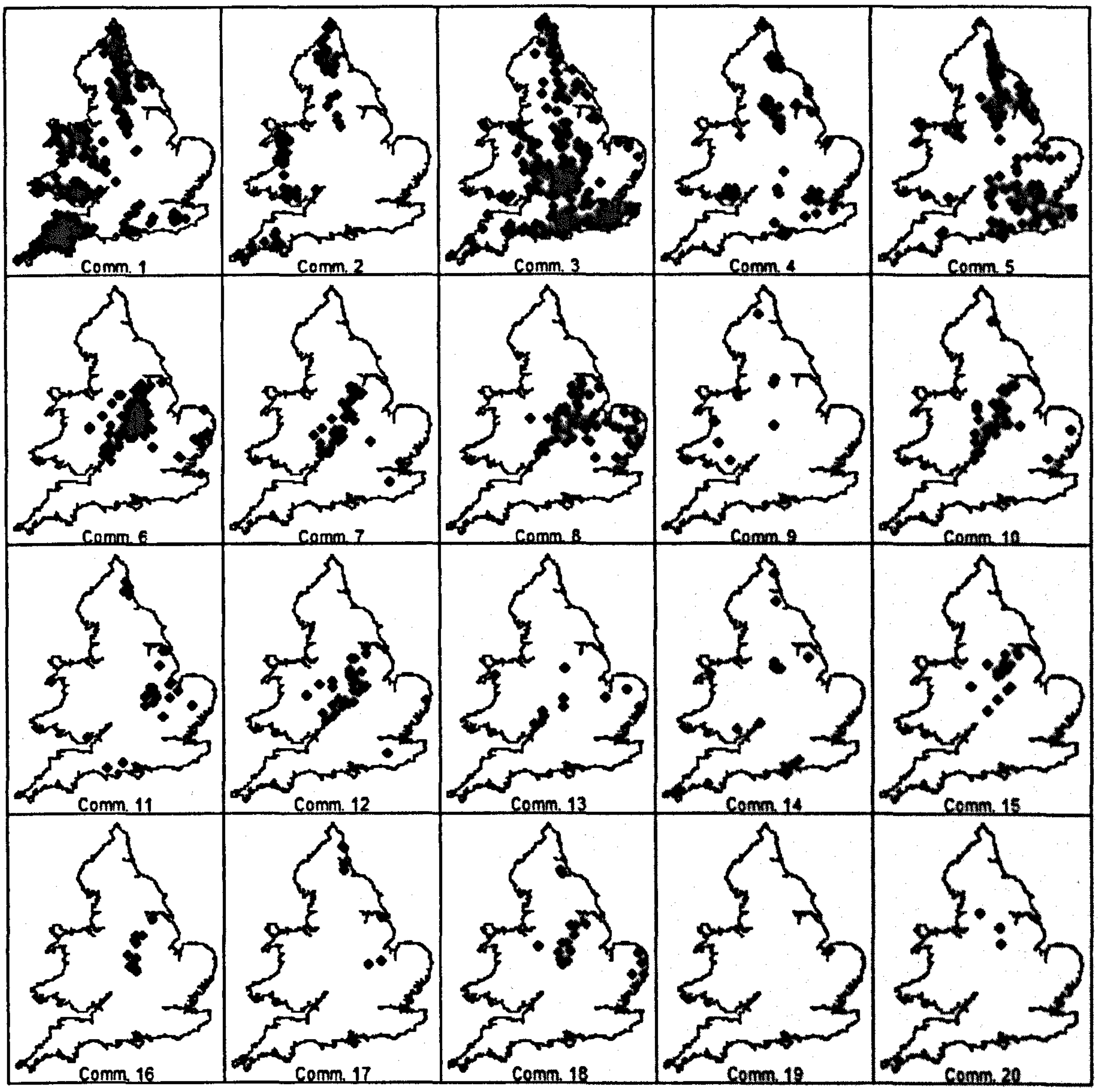


Figure A1.15(a): Spatial structure of communities clustered into 20 end groups using Divisive Clustering.

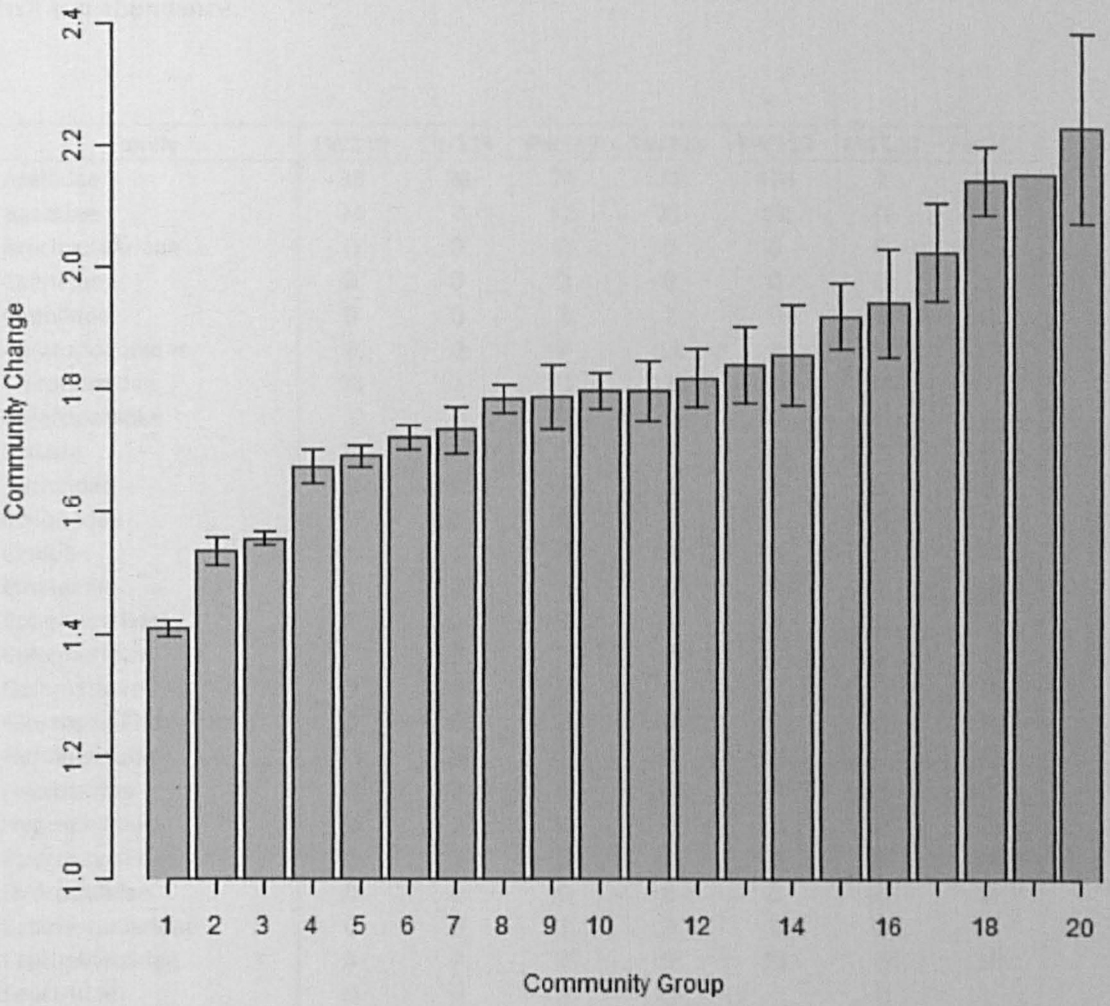


Figure A1.15(b): Average community change for communities clustered into 20 end groups using Divisive Clustering.

APPENDIX 2

Table A2.1: Results from the 2007 sampling of Sheffield Metropolitan District indicating taxa and abundance.

Family	FW110	FW114	FW117	FW118	FW119	FW130	FW131
Asellidae	33	20	74	31	478	7	11
Baetidae	24	7	12	21	61	0	74
Brachycentridae	0	0	0	0	0	0	1
Caenidae	0	0	0	0	0	0	1
Capniidae	0	0	1	2	0	0	0
Ceratopogonidae	2	0	7	11	8	3	6
Chironomidae	98	7	41	111	23	44	115
Chloroperlidae	1	0	0	2	0	0	0
Dixidae	0	0	5	8	16	1	0
Dytiscidae	0	0	1	0	1	1	0
Ecnomidae	0	0	0	1	0	0	0
Elmidae	1	1	8	0	8	2	14
Empididae	1	2	3	37	3	1	10
Ephemerellidae	0	0	0	0	0	0	1
Ephemeridae	0	0	1	0	0	0	7
Gammaridae	0	0	0	0	0	0	0
Glossosomatidae	0	0	0	0	0	0	0
Heptageniidae	19	26	15	0	84	0	92
Hydracarina	0	0	0	0	0	1	2
Hydrophilidae	0	1	12	0	2	0	2
Hydropsychidae	0	1	0	1	0	0	30
Hydroptilidae	0	0	0	0	0	0	0
Lepidostomatidae	0	0	1	0	0	0	1
Leptophlebiidae	8	7	2	15	23	0	38
Leuctridae	0	0	9	87	7	0	1
Limnephilidae	2	3	10	34	9	6	2
Nemouridae	1	2	0	97	3	0	0
Odontoceridae	0	0	2	1	1	0	0
Oligochaeta	6	10	9	24	2	67	40
Perlodidae	0	0	3	17	5	0	0
Philopotamidae	0	0	0	0	0	0	0
Polycentropodidae	0	0	1	2	1	0	0
Psychodidae	0	0	0	1	1	0	0
Psychomyiidae	0	0	4	0	4	1	3
Rhyacophilidae	2	6	10	0	4	0	5
Scirtidae	0	2	12	23	3	0	0
Sericostomatidae	0	0	0	0	1	2	24
Sialidae	0	0	0	0	0	0	0
Simuliidae	5	1	8	11	1	12	1
Tipulidae	0	1	8	4	10	3	2

Table A2.1 continued

Family	FW132	FW136	FW137	FW148	FW168	FW169	FW172
Asellidae	1	62	6	4	42	5	3
Baetidae	38	34	15	31	79	69	134
Brachycentridae	0	0	0	0	0	0	0
Caenidae	0	0	0	0	0	0	0
Capniidae	0	0	0	6	0	0	1
Ceratopogonidae	1	3	8	1	1	4	5
Chironomidae	55	38	52	20	138	92	96
Chloroperlidae	0	0	0	0	1	0	0
Dixidae	0	0	0	0	0	2	0
Dytiscidae	0	0	2	0	1	1	1
Ecnomidae	0	0	1	0	0	0	0
Elmidae	23	8	5	0	0	2	1
Empididae	2	3	1	2	2	7	3
Ephemerellidae	0	0	0	0	0	0	0
Ephemeridae	0	0	1	1	0	0	0
Gammaridae	0	0	0	0	24	2	9
Glossosomatidae	0	0	0	0	0	0	0
Heptageniidae	87	76	15	2	0	0	0
Hydracarina	0	1	0	0	0	0	0
Hydrophilidae	3	7	1	1	0	0	0
Hydropsychidae	4	28	62	0	1	62	15
Hydroptilidae	11	0	0	0	0	0	0
Lepidostomatidae	0	0	0	0	0	0	0
Leptophlebiidae	1	0	23	1	0	0	1
Leuctridae	4	6	11	74	1	3	3
Limnephilidae	1	11	7	25	19	15	24
Nemouridae	1	3	4	25	8	2	2
Odontoceridae	0	0	0	0	0	0	0
Oligochaeta	53	8	0	2	34	10	38
Perlodidae	1	1	5	5	0	0	0
Philopotamidae	0	0	0	0	0	0	0
Polycentropodidae	0	0	1	2	9	8	2
Psychodidae	0	0	0	1	1	0	0
Psychomyiidae	1	2	2	0	0	3	2
Rhyacophilidae	12	6	18	16	1	0	2
Scirtidae	0	0	0	3	0	0	0
Sericostomatidae	17	5	8	2	0	14	22
Sialidae	0	0	0	0	0	0	0
Simuliidae	24	0	1	7	0	38	22
Tipulidae	2	1	1	12	0	1	0

Table A2.1 continued

Family	FW173	FW177	FW182	FW193	FW200	FW298	FW299
Asellidae	0	1	0	6	0	0	0
Baetidae	207	28	25	22	0	0	5
Brachycentridae	0	0	0	1	0	0	0
Caenidae	17	0	0	0	0	0	0
Capniidae	0	0	1	0	4	0	1
Ceratopogonidae	3	0	0	6	0	0	3
Chironomidae	224	19	8	316	74	5	24
Chloroperlidae	1	0	0	0	0	0	0
Dixidae	4	0	0	0	0	0	0
Dytiscidae	1	0	0	0	9	0	3
Ecnomidae	0	0	0	0	0	0	0
Elmidae	0	0	0	0	0	0	1
Empididae	1	0	0	4	0	0	2
Ephemerellidae	0	0	0	11	0	0	0
Ephemeridae	0	0	0	0	0	0	0
Gammaridae	0	7	0	1	0	0	0
Glossosomatidae	1	0	0	1	0	0	0
Heptageniidae	4	0	2	0	0	0	0
Hydracarina	1	0	0	0	0	1	0
Hydrophilidae	2	0	0	4	10	0	0
Hydropsychidae	7	66	2	50	0	0	0
Hydroptilidae	0	0	0	0	0	0	0
Lepidostomatidae	0	0	0	0	0	0	0
Leptophlebiidae	1	0	0	0	0	0	1
Leuctridae	19	0	0	40	15	0	71
Limnephilidae	14	7	0	13	2	1	3
Nemouridae	16	0	0	7	9	7	93
Odontoceridae	1	0	0	0	0	0	0
Oligochaeta	5	11	1	1	5	1	2
Perlodidae	7	0	0	10	3	0	0
Philopotamidae	0	0	0	0	0	0	0
Polycentropodidae	2	2	0	0	0	0	2
Psychodidae	2	0	1	0	0	0	0
Psychomyiidae	0	0	0	0	0	0	0
Rhyacophilidae	0	1	0	1	0	0	8
Scirtidae	0	0	0	0	0	0	0
Sericostomatidae	1	5	0	15	0	0	0
Sialidae	0	0	0	0	0	0	1
Simuliidae	9	15	24	33	0	11	4
Tipulidae	6	0	0	0	1	0	4

Table A2.1 continued

Family	FW302	FW310	FW321	FW322
Asellidae	0	0	153	102
Baetidae	11	0	113	19
Brachycentridae	0	0	2	0
Caenidae	0	0	0	0
Capniidae	0	0	0	3
Ceratopogonidae	0	0	30	1
Chironomidae	17	14	320	35
Chloroperlidae	0	0	0	0
Dixidae	0	0	7	0
Dytiscidae	3	0	0	0
Ecnomidae	0	0	0	0
Elmidae	0	0	0	0
Empididae	0	0	1	3
Ephemerellidae	0	0	0	0
Ephemeridae	0	0	0	0
Gammaridae	0	0	110	849
Glossosomatidae	0	0	0	0
Heptageniidae	0	0	0	0
Hydracarina	0	0	0	3
Hydrophilidae	0	0	18	2
Hydropsychidae	0	0	19	2
Hydroptilidae	0	0	0	0
Lepidostomatidae	0	0	4	1
Leptophlebiidae	1	0	1	0
Leuctridae	27	65	14	204
Limnephilidae	6	7	243	90
Nemouridae	294	298	1	11
Odontoceridae	0	0	0	1
Oligochaeta	3	5	35	0
Perlodidae	0	0	20	0
Philopotamidae	0	0	8	1
Polycentropodidae	2	2	4	1
Psychodidae	0	0	2	0
Psychomyiidae	0	1	10	1
Rhyacophilidae	3	1	5	0
Scirtidae	0	0	1	46
Sericostomatidae	1	0	1	0
Sialidae	0	0	0	0
Simuliidae	1	0	14	2
Tipulidae	4	0	15	4