SEVENTH FRAMEWORK PROGRAMME

THEME 6: Environment (including Climate Change)

Adaptive strategies to Mitigate the Impacts of Climate Change on European Freshwater Ecosystems

Collaborative Project (large-scale integrating project)
Grant Agreement 244121
Duration: February 1st, 2010 – January 31st, 2014

Deliverable 2.24: Changes in ecological status at RIvPACS reference condition sites

Lead contractor: **NERC (CEH)**
Other contractors involved: n/a

Due date of deliverable: **Month 48**
Actual submission date: **Month 48**

Work package:2

Contributors: **Helen Vincent[1], François Edwards[1], Roger Baker[1], Mike Dunbar[2,1], Ralph Clarke[3,1]**

1 NERC Centre for Ecology and Hydrology, Wallingford, UK

2 Environment Agency, Reading, UK

3 Bournemouth University, UK

Estimated person months: 4.6

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013) Dissemination Level (add X to PU, PP. RE or CO)		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for Members of the consortium (including the Commission Services)	

Abstract

The WFD requires Member States to determine the ecological status of rivers and streams with respect to deviation from a type-specific reference condition. It is essential that Member States can demonstrate that the biological datasets used to define reference conditions meet the criteria of the WFD. The approach requires that reference sites be at their ecological optima, and are assumed to not change because by definition they are not impacted. We used RIvPACS reference site data and UK Environment Agency monitoring data to identify 81 RIvPACS reference sites that had subsequent monitoring data, and analysed seasonal patterns in ASPT and Ntaxa. Autumn ASPT increased over time in both data sets, but not Ntaxa, indicative of a shift in reference conditions and species replacement. The trend was site dependent, indicating that long term climatic cycles, or shifts in climate, are an unlikely cause. Deviation from the perceived reference condition was common for ASPT and Ntaxa at most sites, as a majority of subsequent samples did not fall within +/- 5% of the RIvPACS reference values. The ASPT and Ntaxa values of the RIvPACS reference samples for a site did not lie within the standard error range of the overall mean ASPT and Ntaxa for 70 and 80% of the sites respectively. ASPT was generally higher in upland areas of the UK and lowest in lowland agricultural areas. Rates of change in ASPT were highest at sites with intermediate ASPT scores. Low and high values of Ntaxa were more dispersed across the UK, though Ntaxa correlated to mean air temperature indicating a north/south gradient. Rates of change in Ntaxa were also highest at sites with intermediate Ntaxa scores, and rates of change were higher for spring samples. These results demonstrate that the fixed reference condition concept may not be realistic and that selection of reference sites should consider long term variability.

Deliverable 2.24: Changes in ecological status at RIvPACS reference condition sites

Helen Vincent, François Edwards, Roger Baker, Mike Dunbar, Ralph Clarke

Centre for Ecology and Hydrology, Wallingford UK

January 2014

EU FP7 REFRESH deliverable 2.24

Table of Contents

Acronyms used in the report

ASPT	Average Score Per Taxon
BIOSYS	The Environment Agency's national database of macroinvertebrate records
CEH	The Centre for ecology and Hydrology, UK
EA	The Environment Agency, UK
NERC	The Natural Environment Research Council, UK
NID	The National Invertebrate Database
NRFA	The National River Flow Archive, UK
Ntaxa	Number of Taxa
RIvPACS	River Invertebrate Prediction and Classification System

List of Tables

List of Figures

1. Introduction

Reference condition refers to the naturalness of a site and its biota in the absence of human disturbance or alteration (Stoddard et al., 2006). The WFD requires Member States to determine the ecological status of surface water bodies, using site level biotic metrics, which are compared to a type-specific reference condition (European Commission, 2000). This forms an ecological quality ratio, calculated on a site by site basis by dividing the observed metric by the reference value for that metric i.e. the ecological status is the distance its observed biota has shifted from an undisturbed state. It is essential that Member States can demonstrate that the biological datasets used to define reference conditions meet the criteria of the WFD.

The reference condition approach consists of four main steps: identification of reference sites, creation of a typology, comparison of test site to reference site, and finally diagnosis of the pressure/impact responsible for deviations from reference condition. This general framework has evolved from a number of bio-assessment programmes of which the most influential with respect to the WFD was the UK RIvPACS project (River Invertebrate Prediction and Classification System, now River Invertebrate Classification Tool) (Davy-Bowker et al., 2008). The objective of this project was to develop a biological classification system for unpolluted rivers based on their macroinvertebrate fauna, and to test whether the fauna could be predicted from environmental factors (physical, chemical, geological, geographical). The resulting model, once fed with the environmental variables of a test site, provides a site specific prediction of the macroinvertebrate fauna expected to occur in the absence of anthropogenic disturbance (Wright et al., 1993). This can then be compared to an actual sample from the test site, through species presence/absence, and through the use of summary metrics and indices (for example the BMWP score and its associated ASPT and Ntaxa).

The first step in the RIvPACS type approach is the selection of unstressed reference sites upon which the prediction model is based. At each reference site, one must then obtain biological data and environmental variables representing the perceived environmental drivers or at least correlates. It is vital to sample an adequate number of reference sites to ensure sufficient coverage of the different stream types in the prediction model. The RIvPACS type approach requires high quality sites for each type of river or at least sites which have been only marginally impacted which can be used to set a realistic target. The WFD describes a reference site as that with no, or only very minor, anthropogenic alterations to the values of the hydrochemistry and hydromorphology, and with biota associated with such undisturbed or minimally disturbed conditions (Pardo et al., 2012). The word "pristine" should be avoided in connection with reference sites. In the UK, like for many Member States, it is unlikely that any pristine sites exist (Nõges et al., 2007). A RIvPACS type approach cannot define a reference condition for a site in the absence of any high quality or marginally impacted reference sites of "similar" physical type. From a practical point of view, if the best available sites are chosen then targets can be set to improve other sites to at least this level. Reference sites are characterised by minimal change to their hydromorphology and physico-chemistry as long as these do not have a significant effect on the ecosystem. This minimally disturbed condition can be found at sites that have escaped all but the broadest-scale human disturbances.

Another constraint is that the RIvPACS reference sites were chosen subjectively by consulting local biologists (from EA precursor organisations such as the National River Authority and regional river Purification Boards) for the location of the best quality sites, and covering the full range of environmental types and river systems. Over time, some reference sites in earlier phases of the project were removed when judged to be of insufficient quality and others were added to improve the representation of some stream types.

RIvPACS/RICT is underpinned by a database that comprises the reference data for all the RIvPACS sites used to develop the RIvPACS statistical models in a readily accessible format (835 sites). The development of this database was commissioned by the Scotland and Northern Ireland Forum for Environmental Research (SNIFFER) at the request of the UK WFD technical advisory group (UKTAG) and was funded by SNIFFER and the Environment Agency of England and Wales. The data collected are now of particular importance for the implementation of the WFD and for the definition of reference conditions. NERC - CEH, under the contract from SNIFFER, has made the RIvPACS dataset freely available in the public domain to assist this process and to comply with current freedom of information legislation.

Although selecting minimally impacted sites underpins the reference condition approach, there is no universally accepted method for the identification and selection of these sites as references (Bowman & Somers, 2005). In selecting reference sites for RIvPACS, many a priori decisions were made based on local knowledge by water managers and stakeholders. Since then, the SNIFFER WD46 (2007) project identified a number of reference sites where anthropogenic pressure were too strong at the time of sampling to provide an adequate reference by our modern standards and these sites were eliminated/replaced from the reference site network. Furthermore, ecological status of each reference site was established from only one round of sampling at each site (albeit over three seasons: spring, summer, autumn). This provided an accurate snapshot of the ecology of each site at the time of sampling, and allowed for the ranking of sites based on ecological quality. Thus, the focus was very much on finding the right sites with little consideration for temporal trends. More recent debates on the reference condition approach suggest that this understanding of temporal variability is crucial in determining the adequacy of a site to act as a reference for other sites (Moorhead, 2013, Friberg et al., 2011). To date, there has been no drive to search for data gathered from the reference sites before or after the baseline RIvPACS sampling, so the particular year when the sampling took place can be placed in a wider timeline. Indeed the reference sampling year may be entirely atypical at a site, for example in terms of stream flow or water temperature. For some sites, temporal variability may have masked the effects of anthropogenic stressors, giving the appearance that the sites were truly of high ecological quality when in fact they could have been degraded sites experiencing a 'good year'. For other sites, that may have been in recovery from a past stressor, or may have been experiencing only weak impacts from a stressor, ecological quality may have been good, but not at the level one would require from a reference site. If ecological quality changes at reference sites, this poses profound questions about the relevance of the reference condition approach as this would in effect be a moving target, and greatly increases the uncertainty in deciding if a site is compliant with the WFD requirements.

In this study, we used Environment Agency data holdings (BIOSYS) to identify RIvPACS reference sites where macroinvertebrate data was available for other sampling occasions. We aimed to identify whether the year of RIvPACS reference sampling fell within the range of variation for each site, using the ASPT and Ntaxa metrics. There are several reasons why ecological quality could have decreased at some sites since the RIvPACS reference sampling, e.g.: either the site has been impacted by anthropogenic stressors since, or the sampling year was atypical. For sites where ecological quality has increased since RIvPACS reference sampling, either a stressor went undetected at the time and has since decreased, or, again, the sampling year was atypical. Such occurrences would imply that the sites are not suitable for use as reference sites, because essentially the denominator in the O/E ratio is not a constant, and thus contribute to the scientific debate over the usefulness of the reference condition concept.

2. Methods

A database was constructed linking the 835 original RIvPACS reference sites with the monitoring sites visited by the Environment Agency (England & Wales only, Scotland and Northern Ireland have their own regulatory bodies) over the years since the reference samples were collected, using the data stored in their BIOSYS database. This matchup was undertaken using National Grid references for EA sites which were sampled a minimum of 10 times over this period. Of the 835 sites, 81 sites were matched to an EA site within 5m (i.e. the same sites), 256 to an EA site within 500m and 292 to an EA site within 1000m using GIS. It was decided to initiate the analysis using the 81 sites that were most closely matched, rather than risk including sites that, while within 1km of one another, might represent significantly different conditions. A full up-to-date extraction for these 81 sites was made from BIOSYS, yielding data from 2735 samples from 81 sites. The merger of the BIOSYS extraction and the RIvPACS reference samples into a fresh database included all the samples from the RIvPACS reference collection (3 samples per site: spring, summer, autumn) and the subsequent (and in some cases prior) samples taken at these 81 matched sites. RIvPACS samples were collected from 1978 – 1993. BIOSYS samples collected between 1979 and 2011, but with increasing samples predominantly collected from the 1990’s onwards. For each sample, the database lists the season and year in which the sample was taken and the BMWP, ASPT, Ntaxa and LIFE(F) scores for each sample. In this study, we focus on ASPT and Ntaxa, the two commonly reported metrics used by the environment agency to assess ecological quality for WFD purposes.

We set a +/- 5% threshold of deviation from seasonal RIvPACS ASPT and Ntaxa scores and calculated the percentage of samples that fell within these thresholds. For samples that deviated by +5% from the RIvPACS reference status, we examined regional biases with correlation analysis of ASPT and Ntaxa to major spatial drivers: altitude, distance from source, mean air temperature and alkalinity. We then calculated the mean of all BIOSYS ASPT and Ntaxa values for a site and compared them seasonally to the RIvPACS sample ASPT and Ntaxa, and also to the mean RIvPACS ASPT and Ntaxa scores for a site (mean of 3 seasons). We calculated the percentage of RIvPACS samples for which ASPT and Ntaxa did not fall within the range of standard error of the mean of the BIOSYS values for a site.

A mixed model approach was used in R to assess temporal trends over the 33 year period from 1978 to 2011. ASPT and NTAXA were used as dependent variables and analysed separately for spring and autumn. Year was centred on the mean year (1996), to allow an interpretable intercept, and coded as “cenYear”. Operator relates to whether the sample was collected by RiVPACS or BIOSYS, B and R respectively. The fixed factor used was coded as R_B. Biosys site identifiers were used, as each Biosys site had a corresponding RiVPACS site, therefore only one code was used (BIOSYS_SITE_ID). The fixed effects part of the model was constructed using two fixed factors, cenYear and R_B as separate predictors and the interaction between them. The random component was constructed using site id as the random intercept term and cenYear as the random slopes term:

Dependent Variable (ASPT/NTAXA) ~ cenYear * R_B + (cenYear|BIOSYS_SITE_ID) + error

The model was run separately for each dependent variable using spring and autumn data to detect seasonal influences. Model residuals were checked for normality. Components of the model were checked using the full model and reduced models, checking different components separately. Model fits were checked using χ^2 tests. Comparison of the AIC and BIC values were also estimated.

3. Results

The 81 selected sites were plotted (Figure 1) demonstrating the spatial coverage for England and Wales. RIvPACS data consisted of 81 samples for each site/season combination, whereas BIOSYS had markedly less data for the summer season than for autumn and spring (Table 1). The vast majority of ASPT and Ntaxa site sample scores in the BIOSYS data were at least 5% higher or lower than the RIvPACS ASPT/Ntaxa for that site (Table 2). There were more samples with higher ASPT than lower ASPT (particularly in autumn), but conversely there were more samples with lower Ntaxa than higher Ntaxa (Figure 2, Figure 3). For Ntaxa, values above+ 5% of the RIvPACS value tended to occur in spring, and those lower than -5% tended to occur in summer and autumn.

The time series of ASPT and Ntaxa were plotted for each site and season (Figures 4 – 6 and 7 – 9 respectively). A number of different scenarios were revealed: sites where later samples fall entirely +/- 5% of the RIvPACS value (e.g. 46195 in Figure 4); sites where variability is much higher but is not consistently above or below the 5% thresholds (e.g. 56037 in Figure 5), sites where there is a clear trend for values consistently lower than the RIvPACS value (e.g. 52133 in Figure 7 or 56016 in Figure 8) and sites where there is a clear trend for values higher than the RIvPACS value (e.g. 1910 in Figure 7 or 9127 in Figure 8).

For samples above the 5% threshold, correlation analyses indicated that only Ntaxa and mean air temperature were correlated to one another (Table 3). Regression analysis indicated that this was a weak trend F = 6.1, p = 0.016, R^2 = 7.2% (Figure 10), however because there was a clear geographical mean air temperature gradient, a similar gradient in autumn Ntaxa values is likely (Figure 11).

The RIvPACS values of ASPT and Ntaxa for a site for each season generally fell outside the range of the standard error of the mean of the BIOSYS samples (**Figures 12 – 14 and Figures 16 – 18** respectively and Figure **15** and Figure **19**) but this varied strongly by site, season and metric. At some sites (e.g. 55556) ASPT always falls outside the standard error range, at others, seasonal fluctuations are evident (e.g. 1910). For Ntaxa many sites also fell outside the standard error range across all seasons, either above (e.g. 55916) or below (e.g. 55556), though some did show seasonal fluctuation (e.g. 36204). Overall, a very high percentage of sites had RIvPACS values of Ntaxa and ASPT that did not fall within the range of error of mean BIOSYS values (Table **4**,Figure **20**). Summer season saw the least number of sites where the RIvPACS Ntaxa and ASPT fall outside of the standard error range of the mean BIOSYS values.

Mixed modeling indicated that both site mean (intercept) and rate of change (slope) for ASPT varied widely between sites, with roughly an equal amount of sites falling above or below the overall mean (all sites), with similar patterns in means and rates in autumn and spring (Figure 21). For Ntaxa, more site means were below the overall mean than for ASPT in both autumn and spring, and rates of change were greater at low and high mean site Ntaxa values (Figure 22). Spatial patterns were apparent, mean

ASPT (Figure 23) was highest in parts of Northern England and Northern Wales, and lowest in the East Midlands and East Anglia area, for both spring and summer. The rates of change were generally the lowest in these areas. Ntaxa (Figure 24) showed a different spatial pattern with sites with high mean Ntaxa more randomly dispersed and fewer than sites with high mean ASPT, and with rates of change highest in Central England. Mean Ntaxa did not differ much with season, but the rate of change was generally lower in autumn that in spring seasons.

Mixed modeling indicated for spring ASPT (Tables 5 -7) that random slopes and intercepts for sites were important. Variance in ASPT was dependent on site and so was the rate of change. ASPT in autumn (Tables 8 – 10) was the same as ASPT spring except there was a significant time / operator interaction. Ntaxa in spring (Tables 11-13) seems to be dependent on the site, so rates of change and the mean value depends on the site id. The Autumn Ntaxa (Tables 14 – 16) is dependent on site over time in terms of the mean score and the rate of change. The difference between AIC for models fm1 and fm2 is very small, indicating very little difference between the models, and consequently little difference of R_B (Figure 25). This is also the case for the BIC values.

The Mean value and rate of change in ASPT was different for RIVPACS and BIOSYS and showed a clear trend over time for increasing ASPT autumn scores in both data sets, which was site dependent (Figure 26). A similar spring pattern was not significant. Ntaxa did not reflect this pattern in either season.

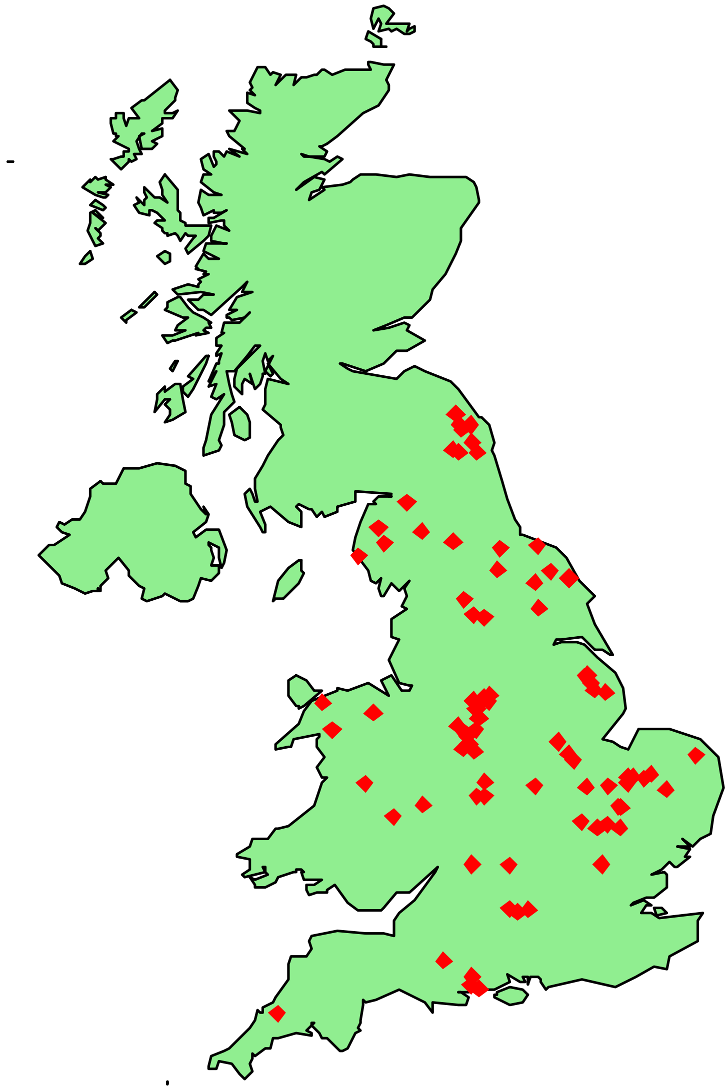

Figure 1: Map of the UK showing the location of the 81 matched sites

Table 1: Number of sites and samples matched between the RIvPACS and BIOSYS databases, matched to within 5meters and split by season. Sites were matched using GIS.

Number of matched sites		81	
Number of samples	**RIvPACS**	**BIOSYS**	**Total**
	243	2492	**2735**
Spring	81	1012	1093
Summer	81	450	531
Autumn	81	1030	1111
			Total 2735

Table 2: Number of BIOSYS samples and sites above, below, and within RIvPACS scores +/-5%.

Number of samples above the 5% RIvPACS level	ASPT	NTAXA
Spring	328	421
Summer	141	138
Autumn	425	383
Number of samples below the 5% RIvPACS level	**ASPT**	**NTAXA**
Spring	245	435
Summer	106	234
Autumn	193	494
Number of BIOSYS sites within RIvPACS +/-5%	**ASPT**	**NTAXA**
Spring	3	0
Summer	7	3
Autumn	2	1

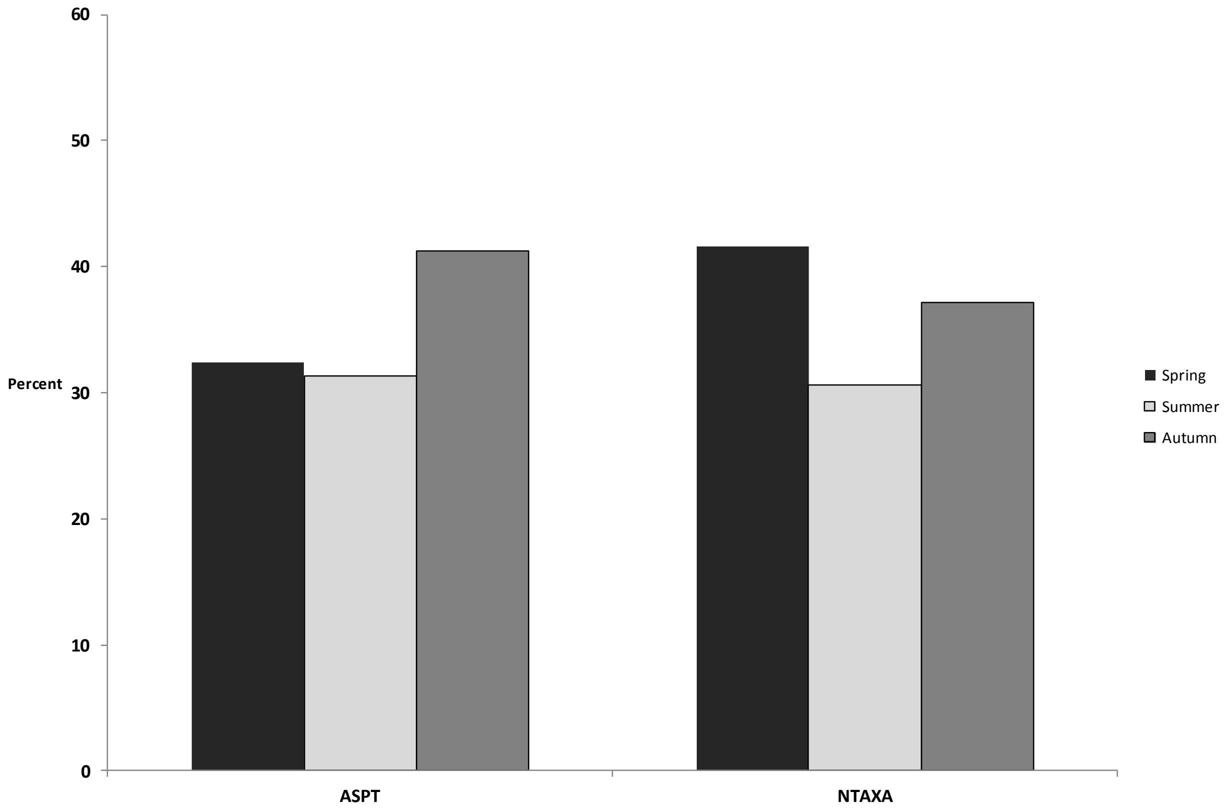

Figure 2: Percent of BIOSYS samples above the RIVPACS value + 5%

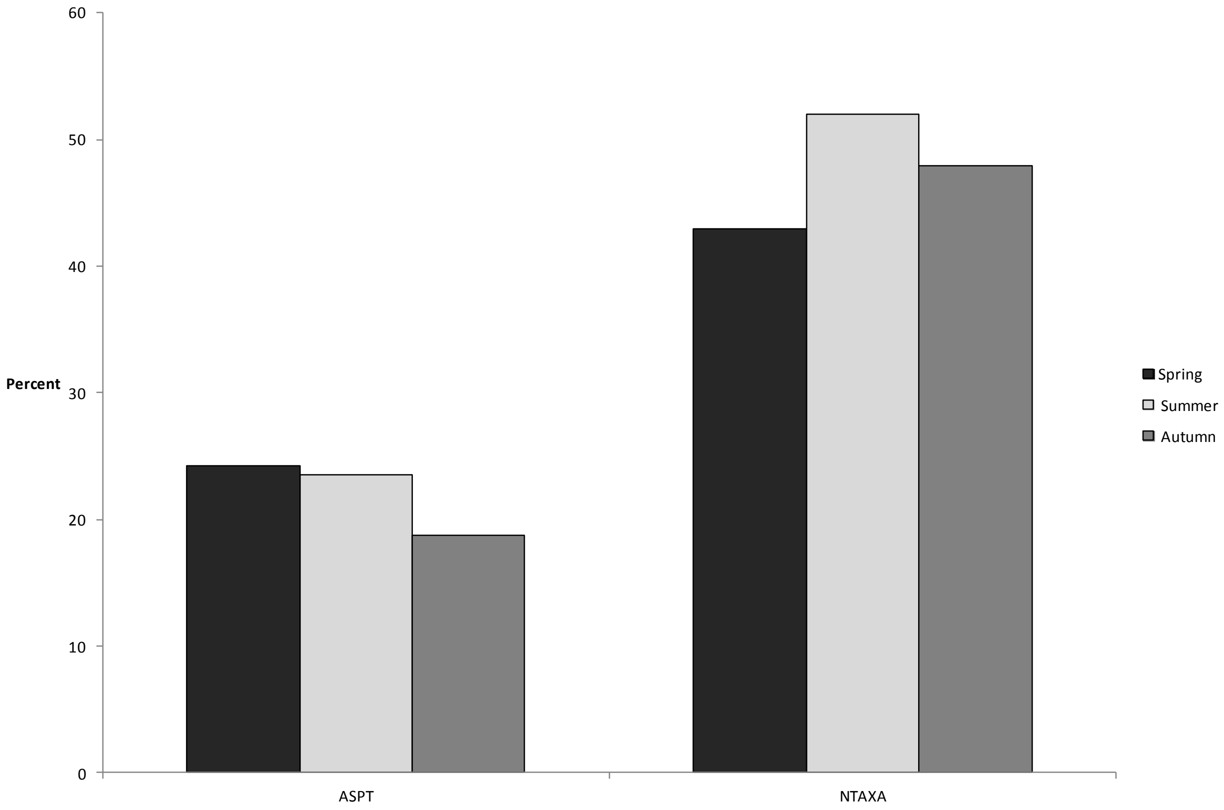

Figure 3: Percent of BIOSYS samples below the RIVPACS value -5%

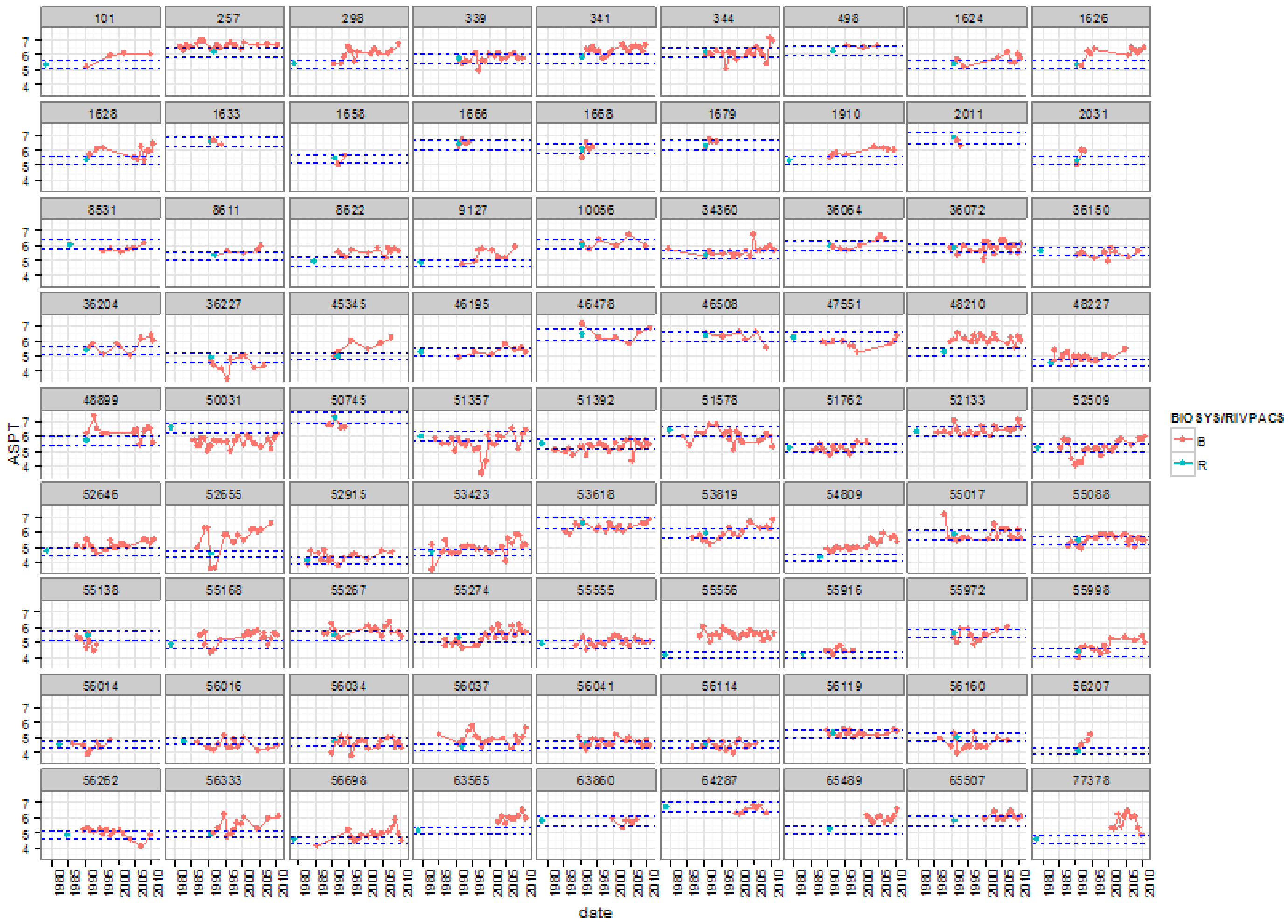

Figure 4: Autumn ASPT across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%.

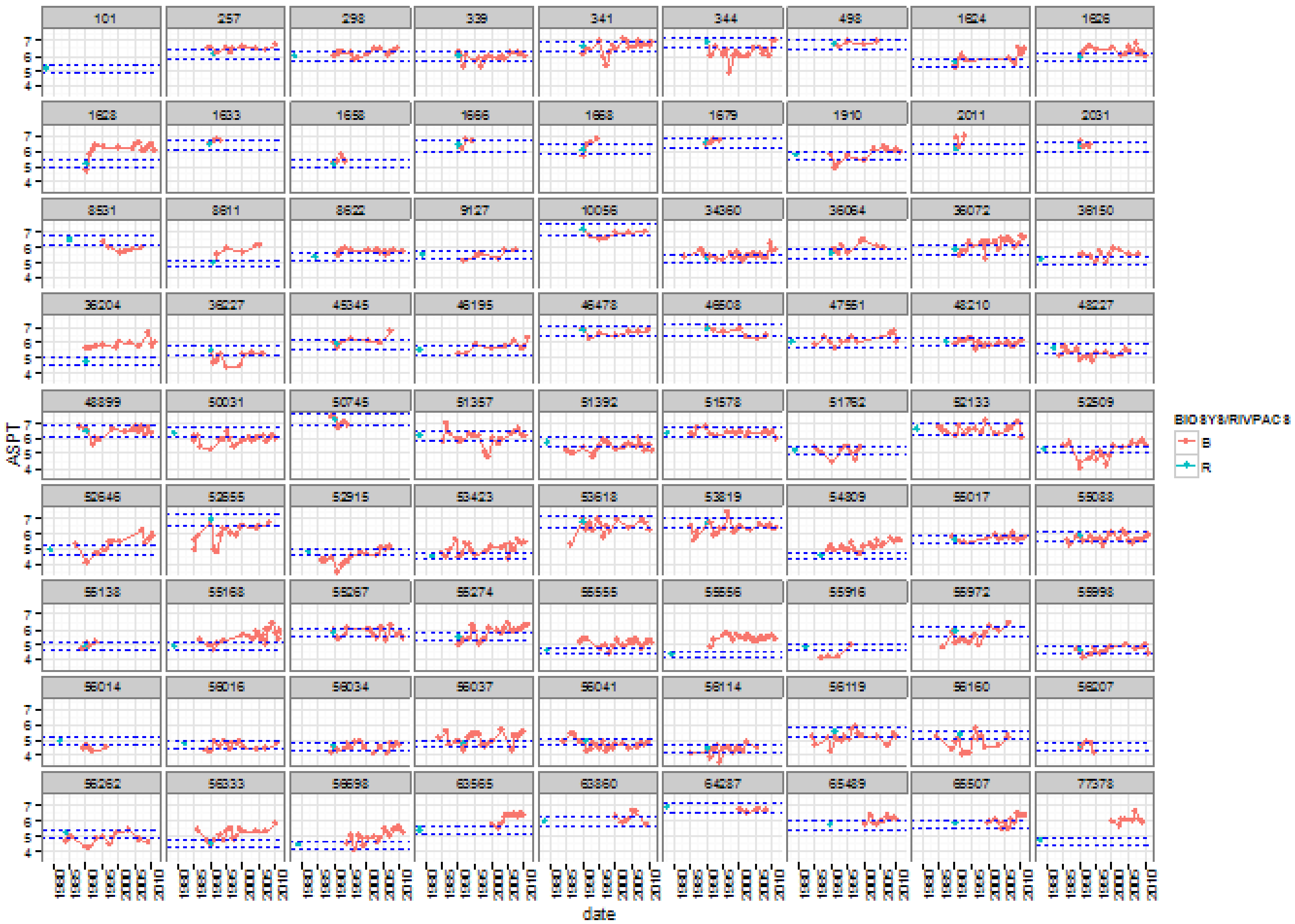

Figure 5: Spring ASPT across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%

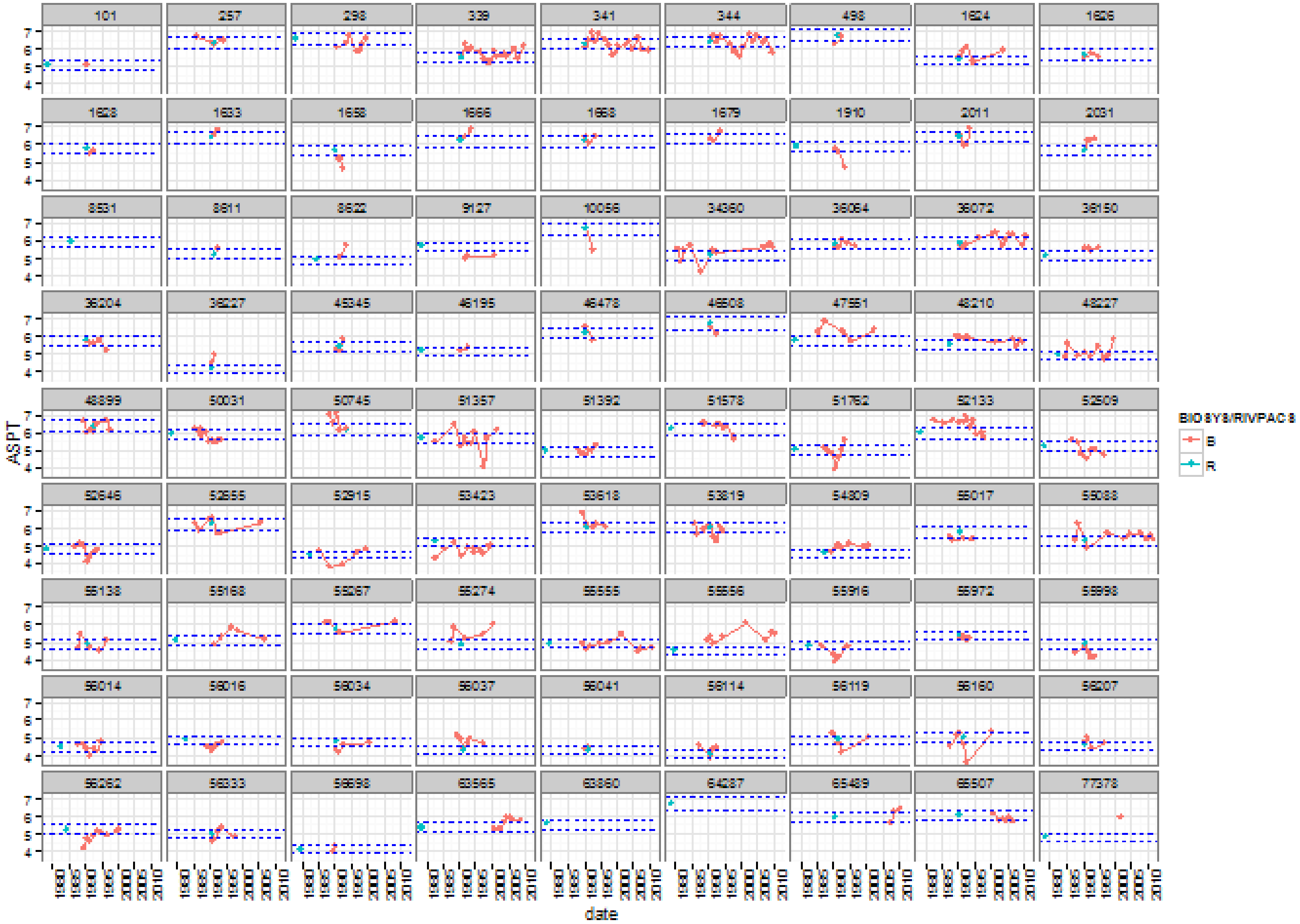

Figure 6: Summer ASPT across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%

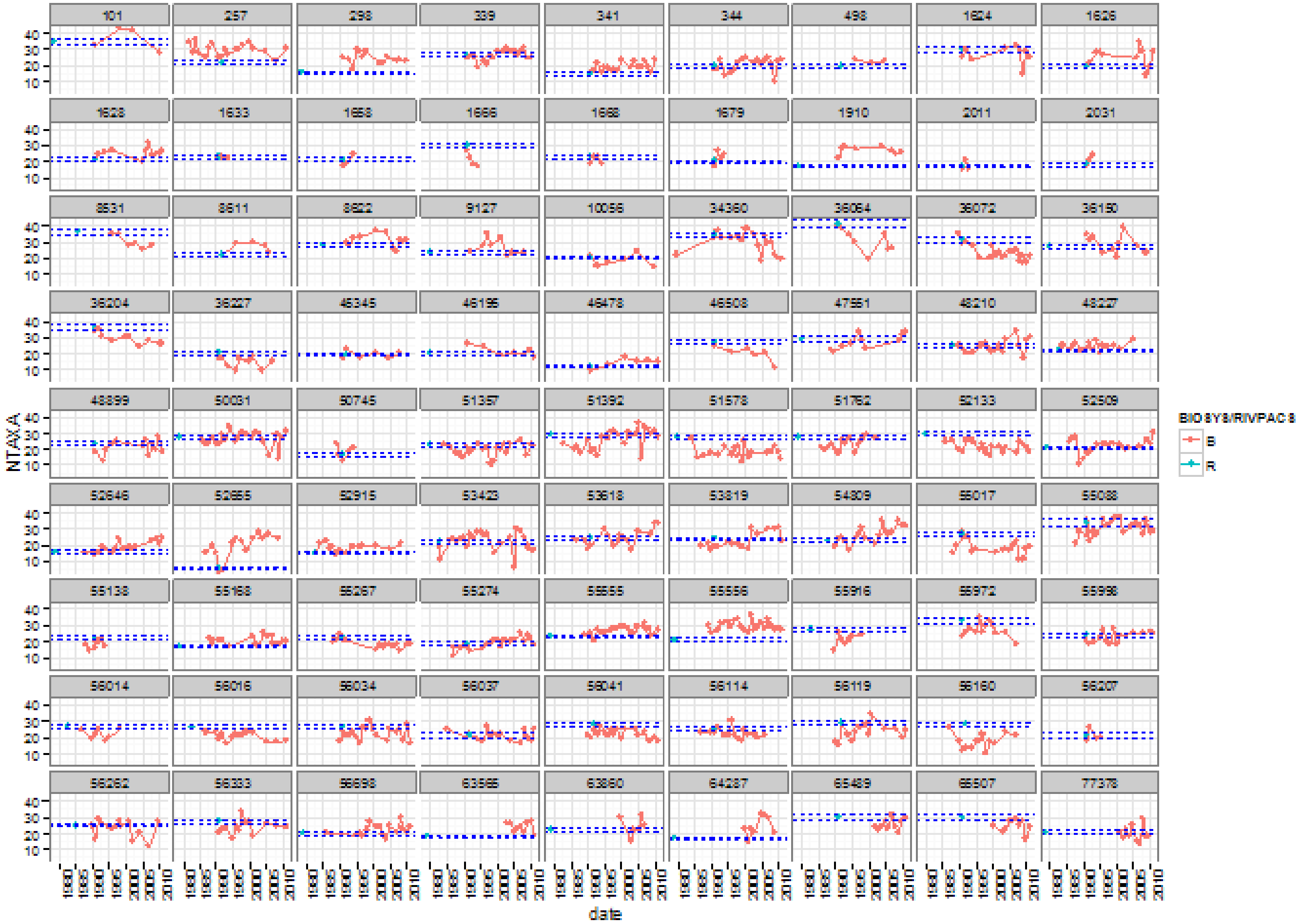

Figure 7: Autumn NTAXA across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%

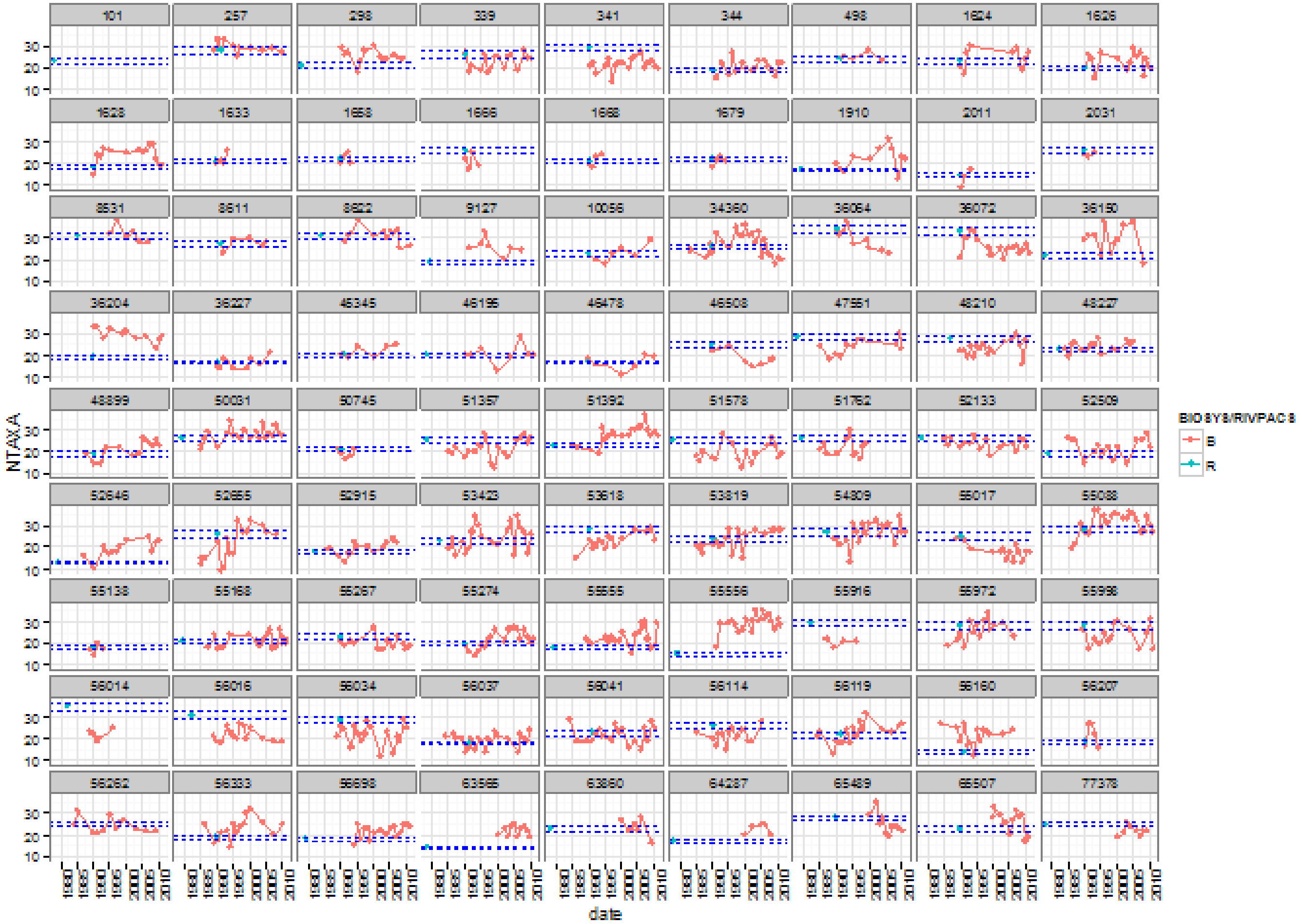

Figure 8: Spring NTAXA across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%

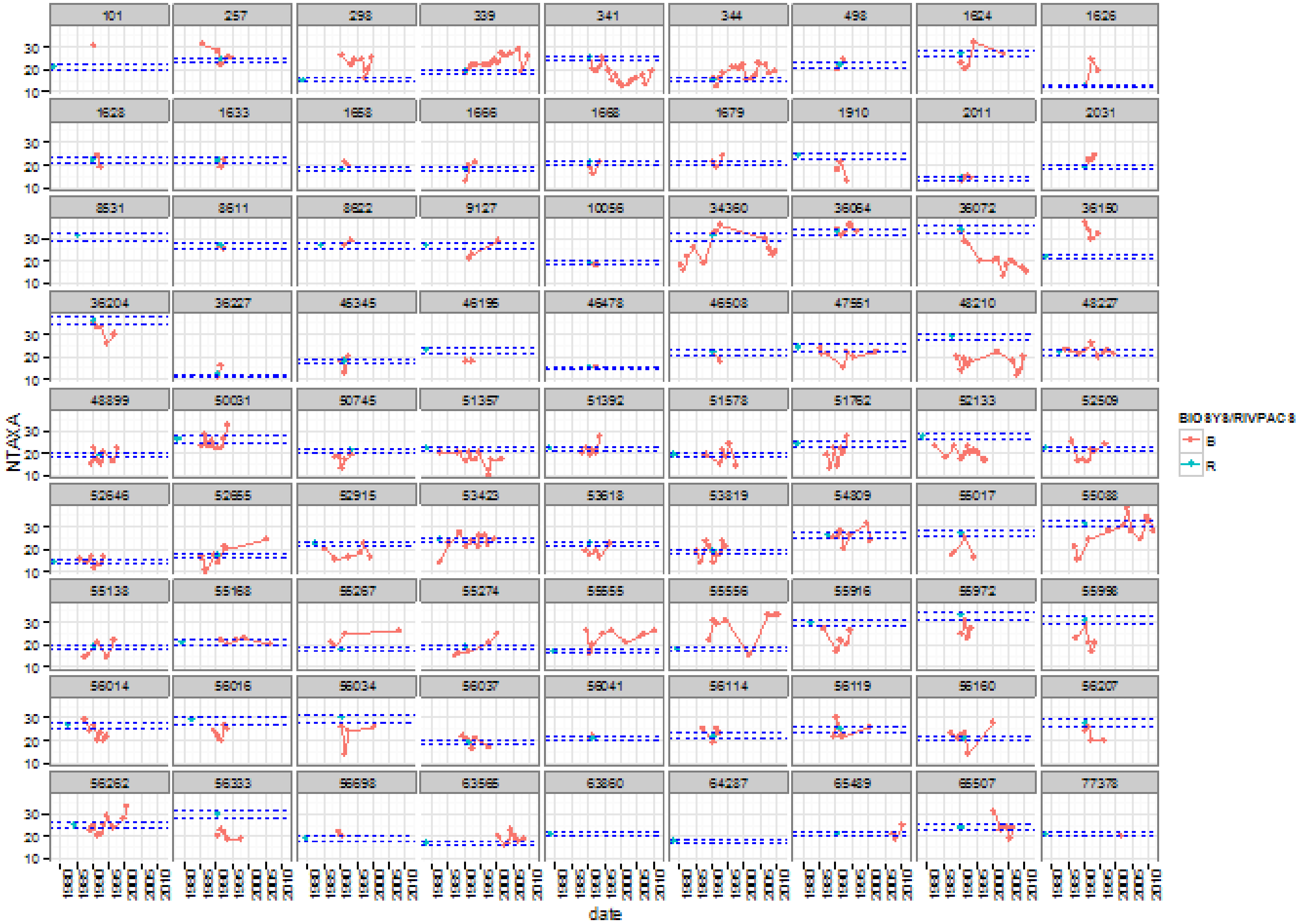

Figure 9: Summer NTAXA across the 81 matched sites. Dashed lines represent RIVPACS values +/- 5%

Table 3: correlations between the proportion of BIOSYS samples greater than the RIvPACS value + 5% and major regional drivers. * Correlations significant to the 95% level

	Proportion of samples above the RIvPACS Plus 5%					
	Spring		Summer		Autumn	
	NTAXA	ASPT	NTAXA	ASPT	NTAXA	ASPT
Distance from Source	0.134	-0.037	0.064	-0.001	-0.007	0.121
Altitude	0.073	0.142	0.112	-0.007	-0.029	0.035
Alkalinity	0.084	0.013	0.068	0.013	0.147	-0.093
Mean Air Temperature	0.188	0.060	0.012	0.060	**0.268***	0.148

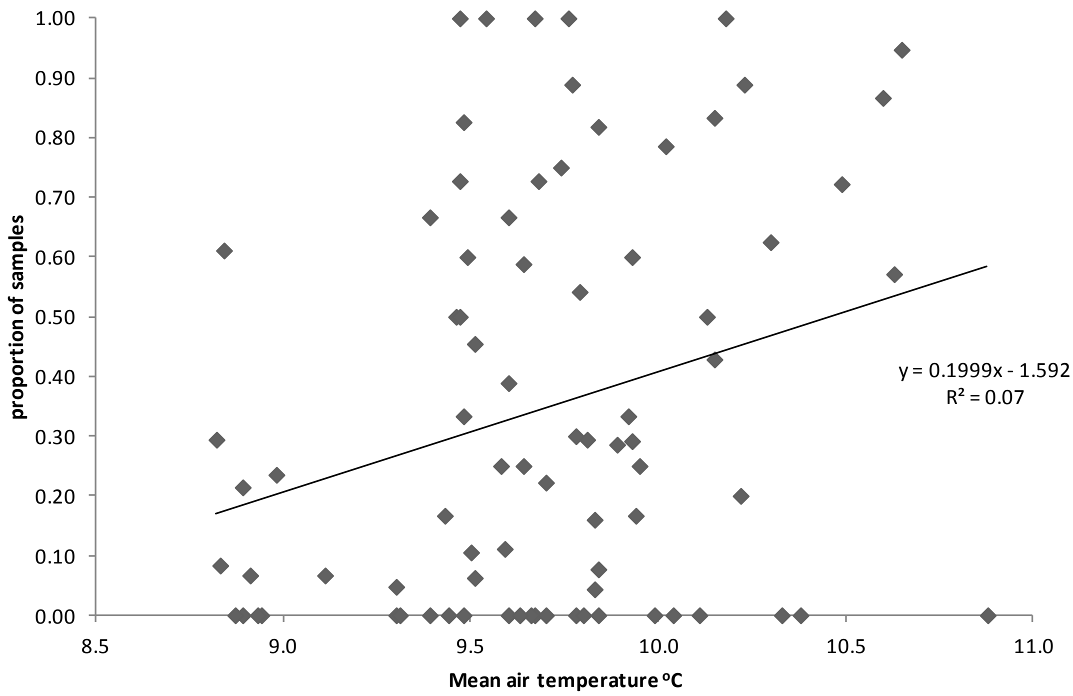

Figure 10: Linear regression analysis between the proportion of BIOSYS samples above the RIVPACS value + 5% for each site and Mean Air Temperature

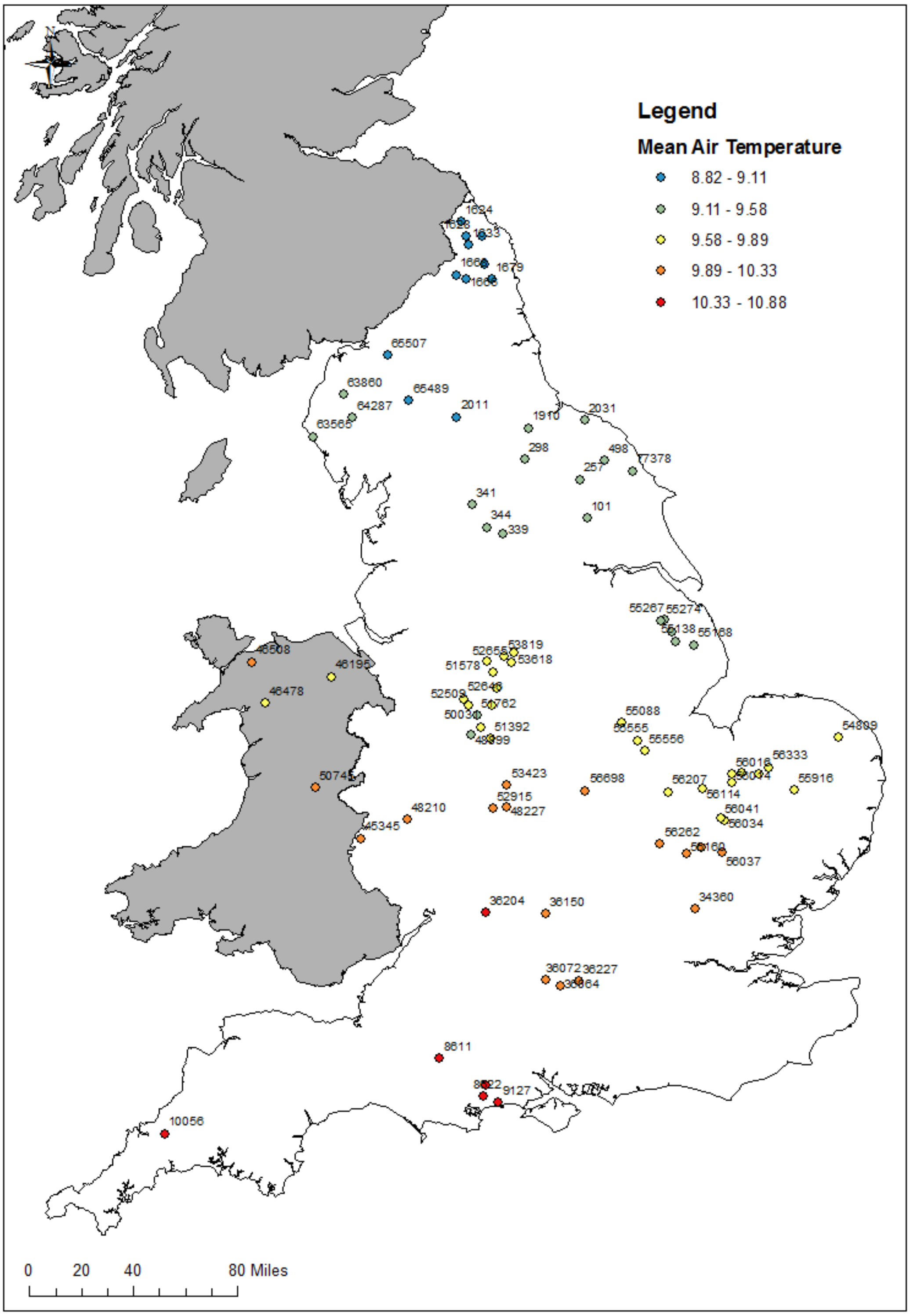

Figure 11: Map showing matched sites and Mean Air Temperature gradient

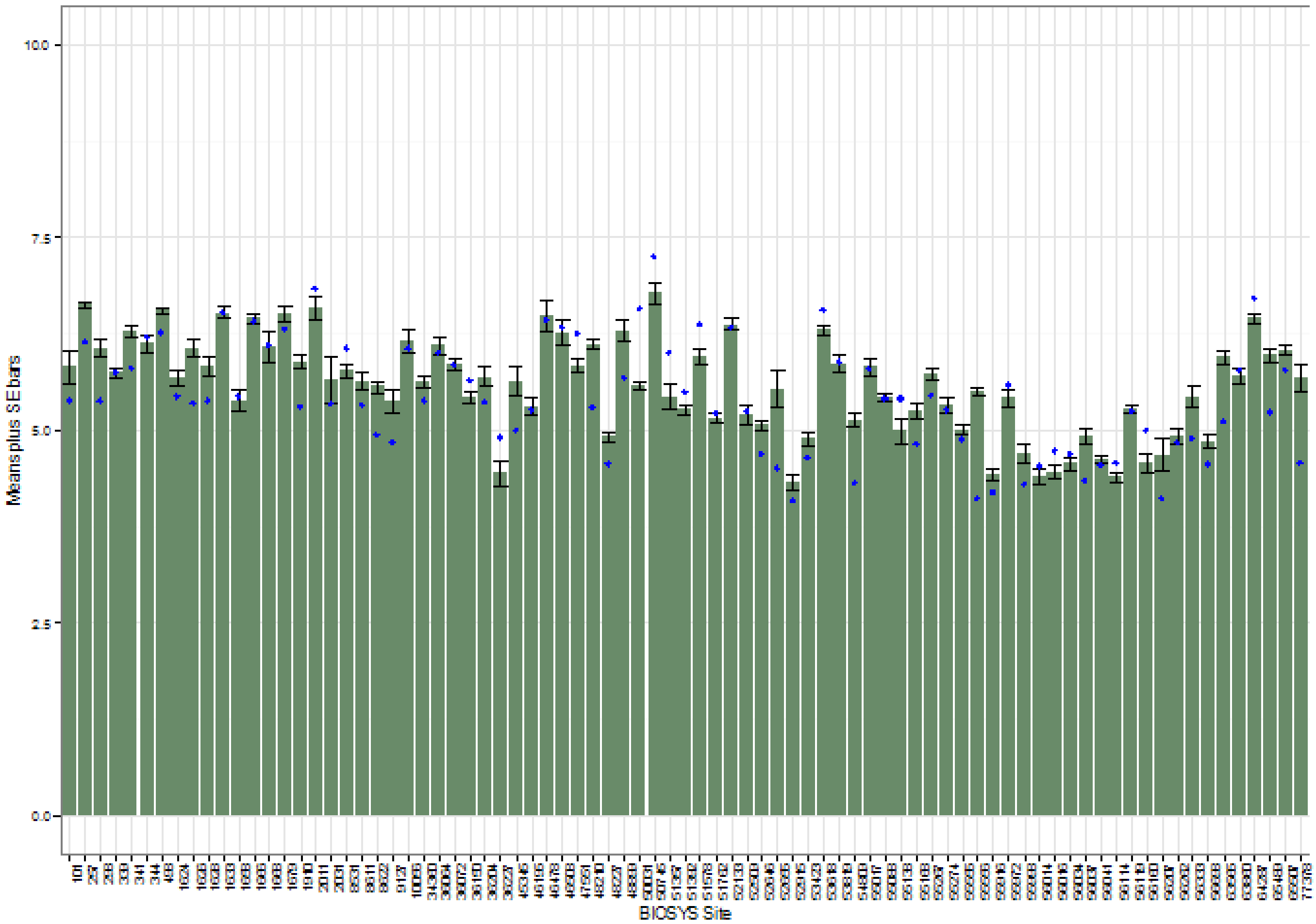

Figure 12: Means +/- SE for BIOSYS Autumn ASPT. Blue points show RIVPACS reference values

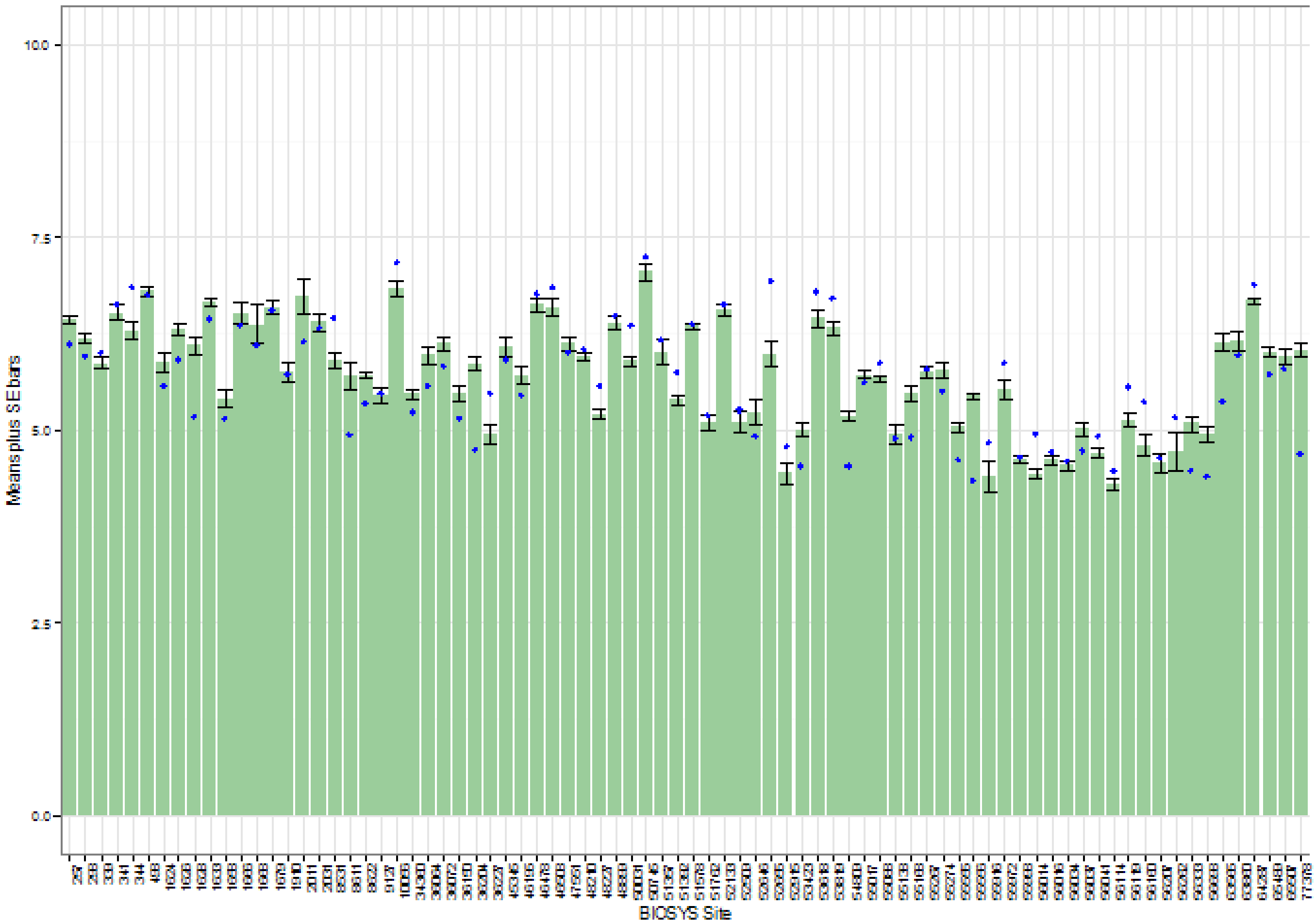

Figure 13: Means +/- SE for BIOSYS Spring ASPT. Blue points show RIVPACS reference values

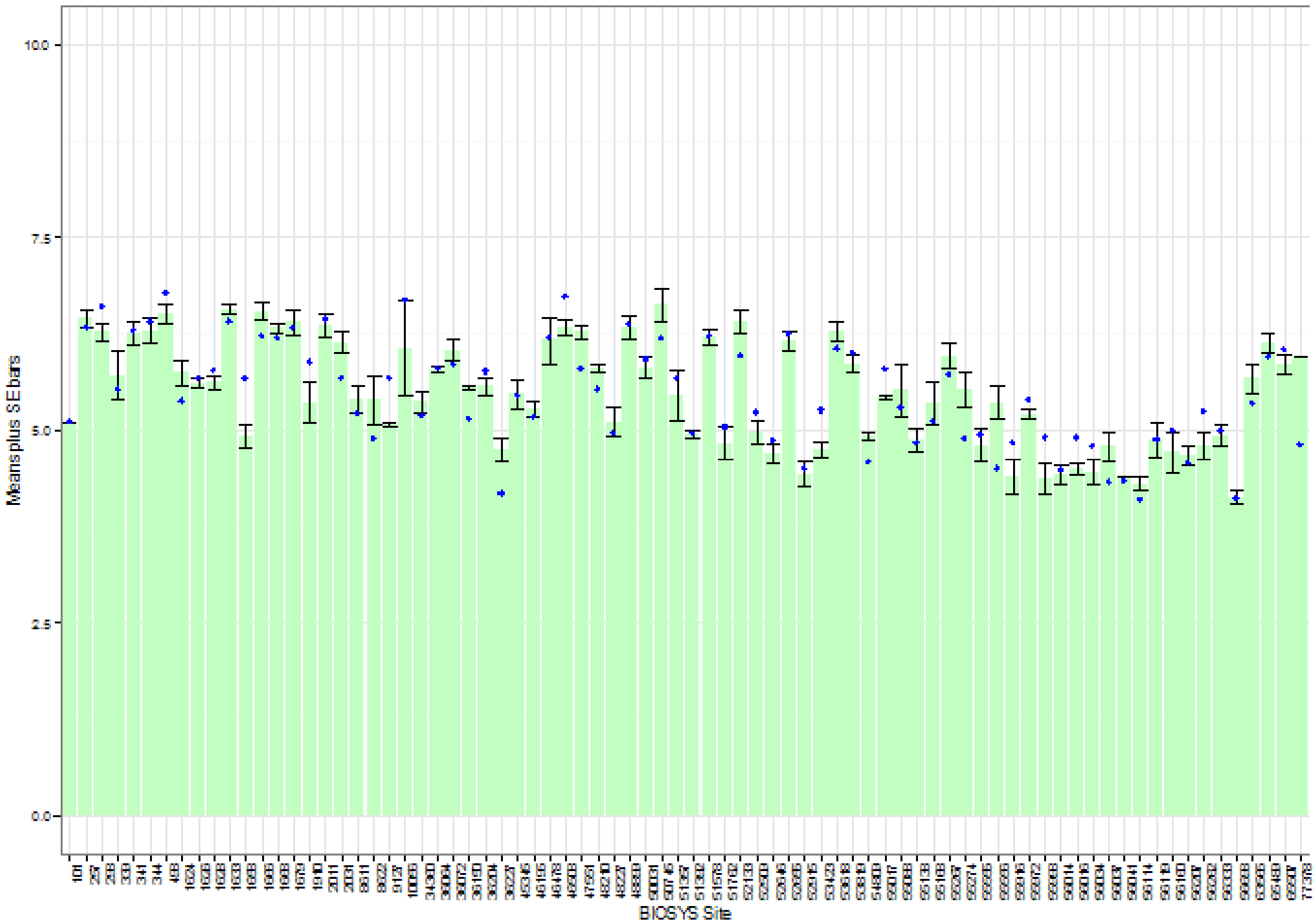

Figure 14: Means +/- SE for BIOSYS Summer ASPT. Blue points show RIVPACS reference values

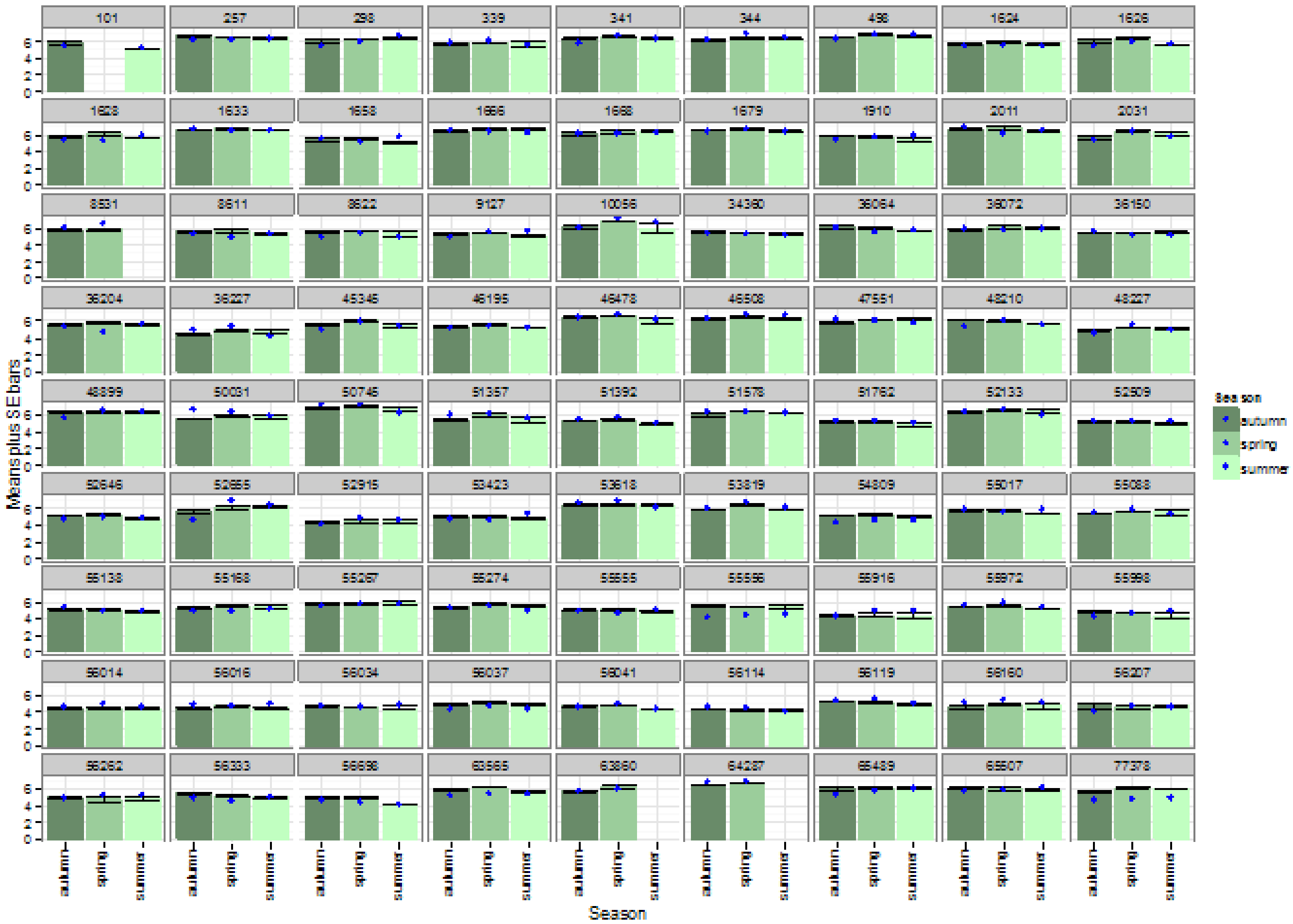

Figure 15: Means +/- SE across seasons for ASPT

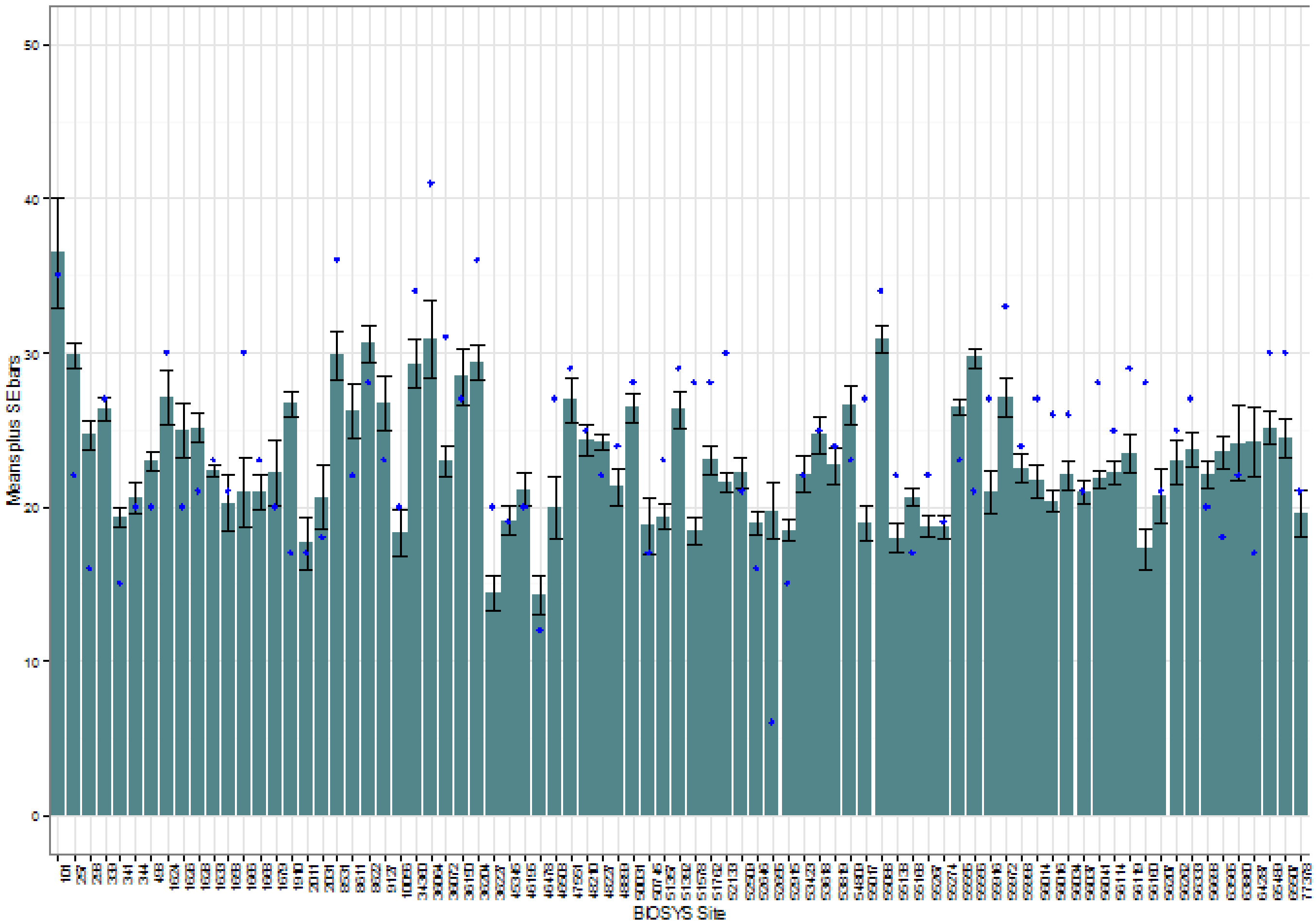

Figure 16: Means +/- SE for BIOSYS Autumn NTAXA. Blue points show RIVPACS reference values

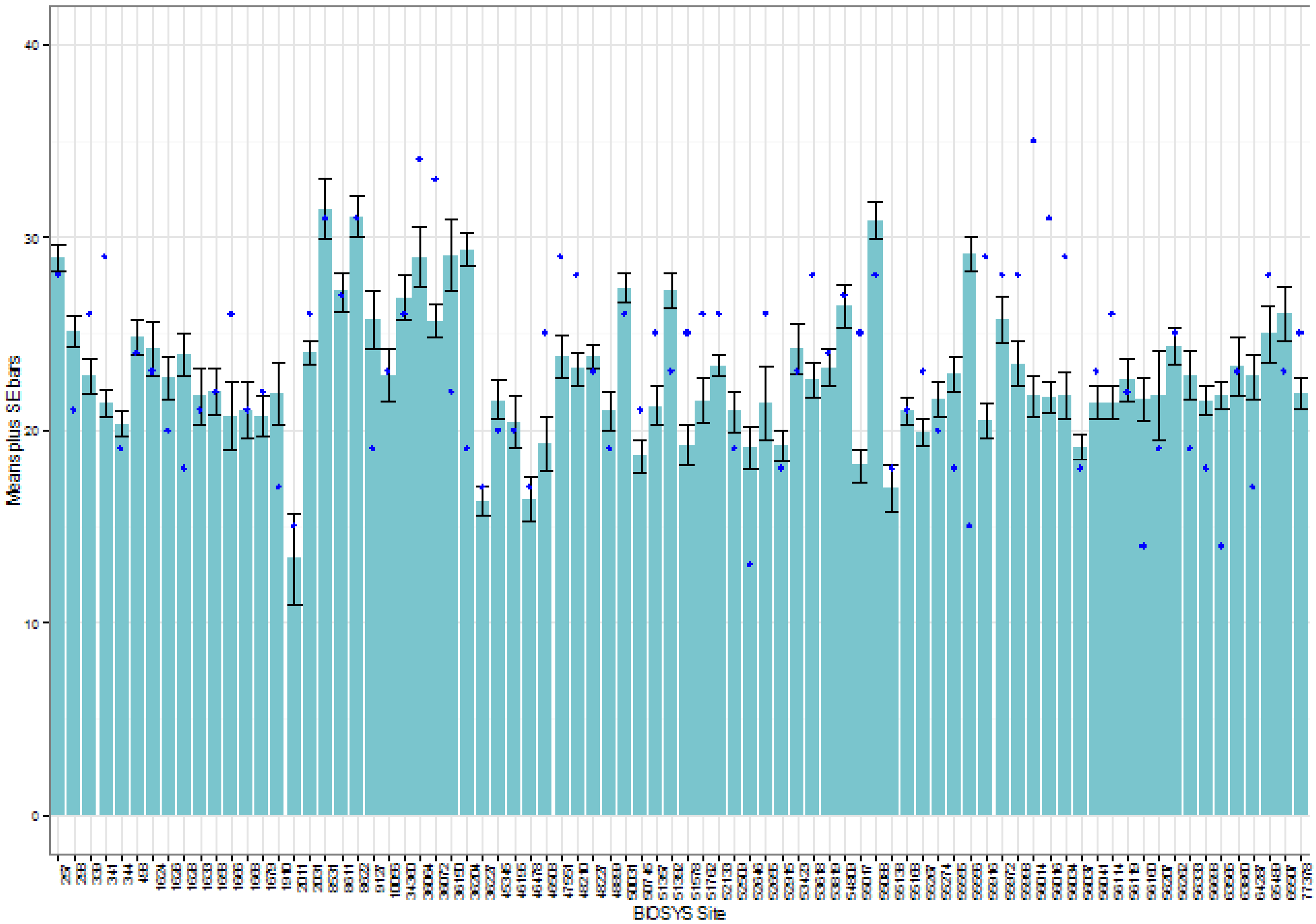

Figure 17: Mean +/- SE for BIOSYS Spring NTAXA.

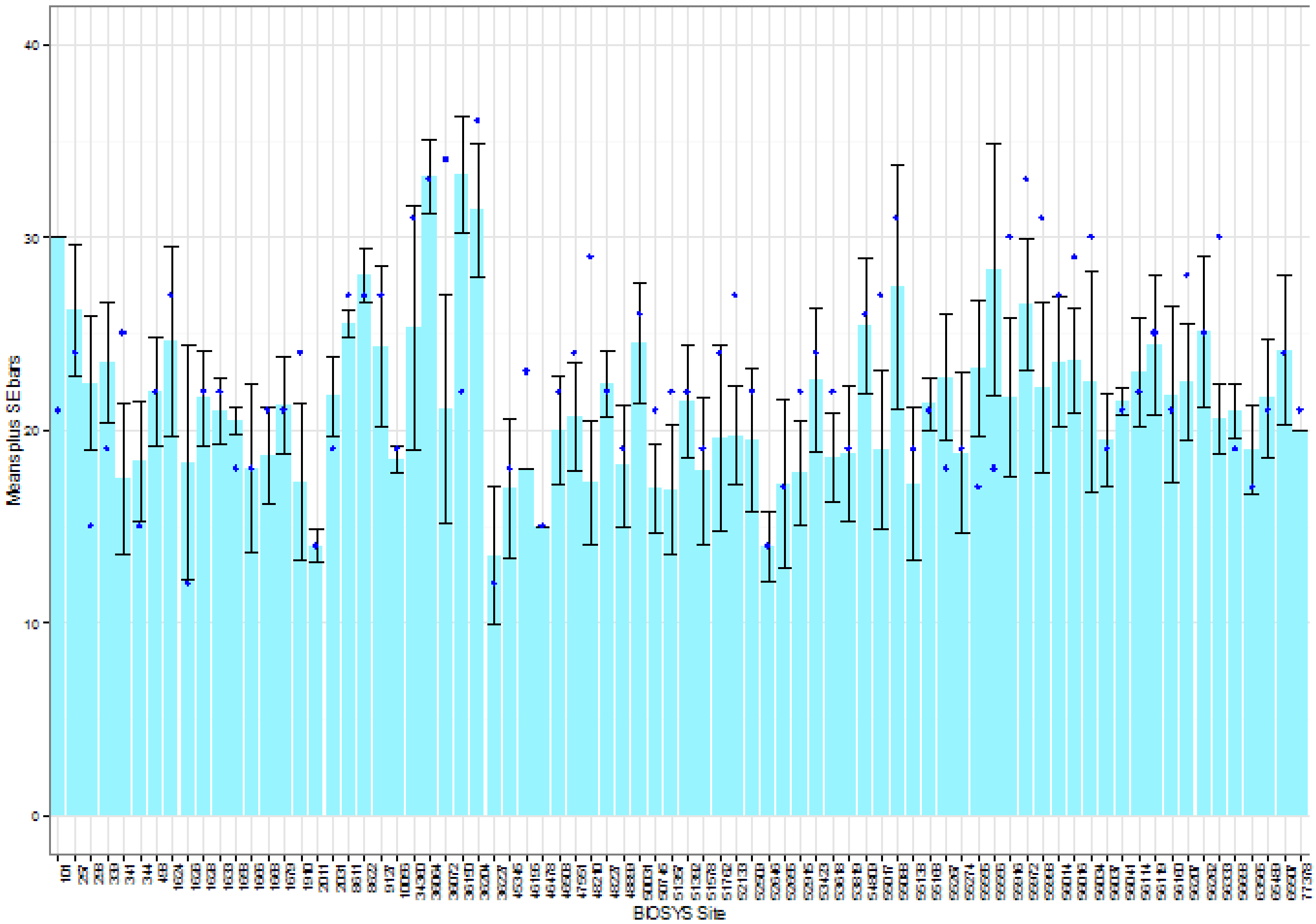

Figure 18: Mean +/- SE fro BIOSYS Summer NTAXA.

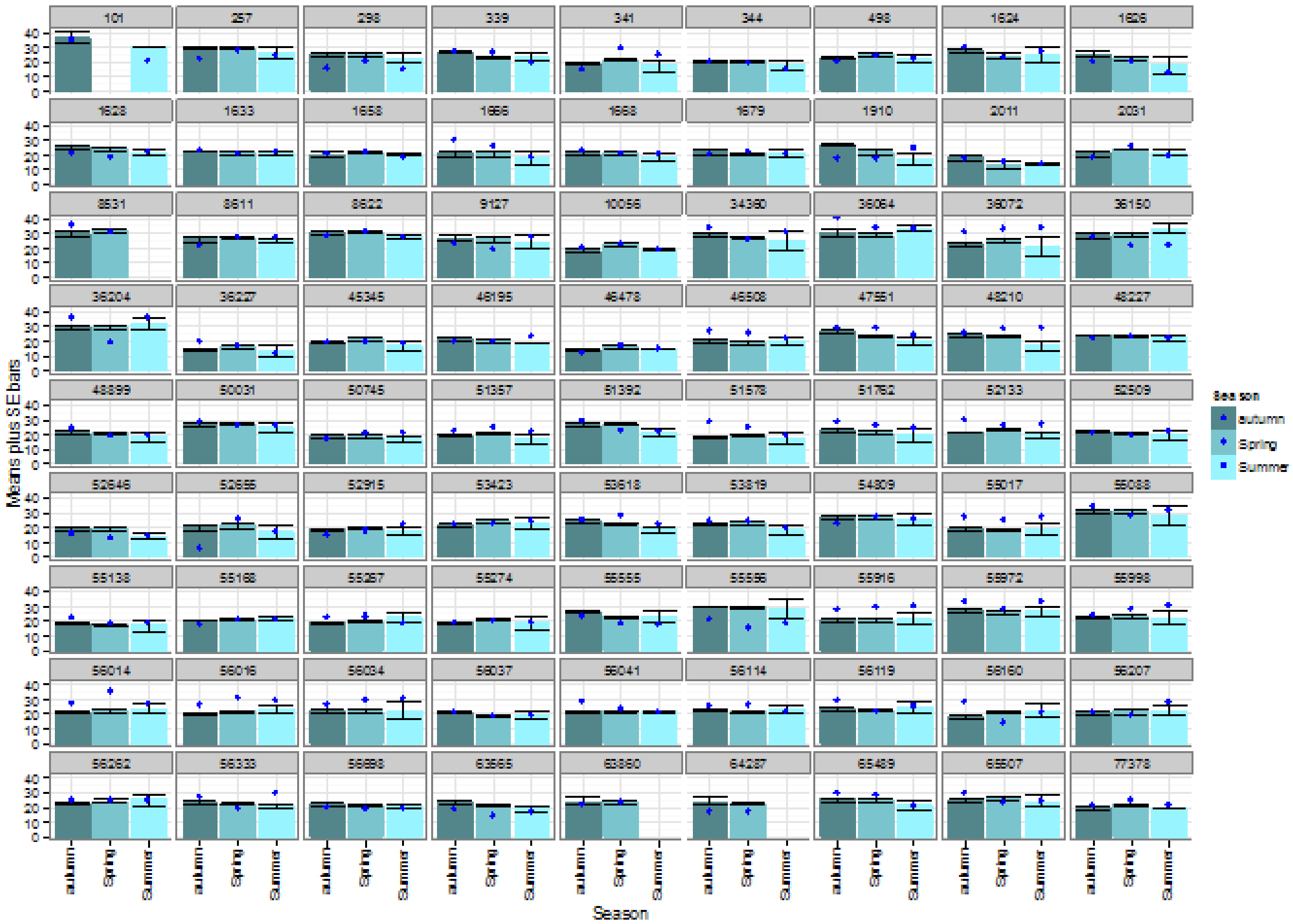

Figure 19: Means +/- SE across seasons for Ntaxa

Table 4: Sites where the RIvPACS metric lie outside of the BIOSYS mean +/- the standard error bars

		Number of	Sites where the RIVPACS value is outside of the SE bars for the BIOSYS data			
		> BIOSYS SE	< BIOSYS SE	total samples outside of BIOSYS SE	Total no. of sites sampled	Percent
NTAXA	Spring	29	29	58	80	73
	Summer	23	12	35	78	45
	Autumn	39	26	65	81	80
ASPT	Spring	28	36	64	80	80
	Summer	21	27	48	78	62
	Autumn	19	40	59	81	73

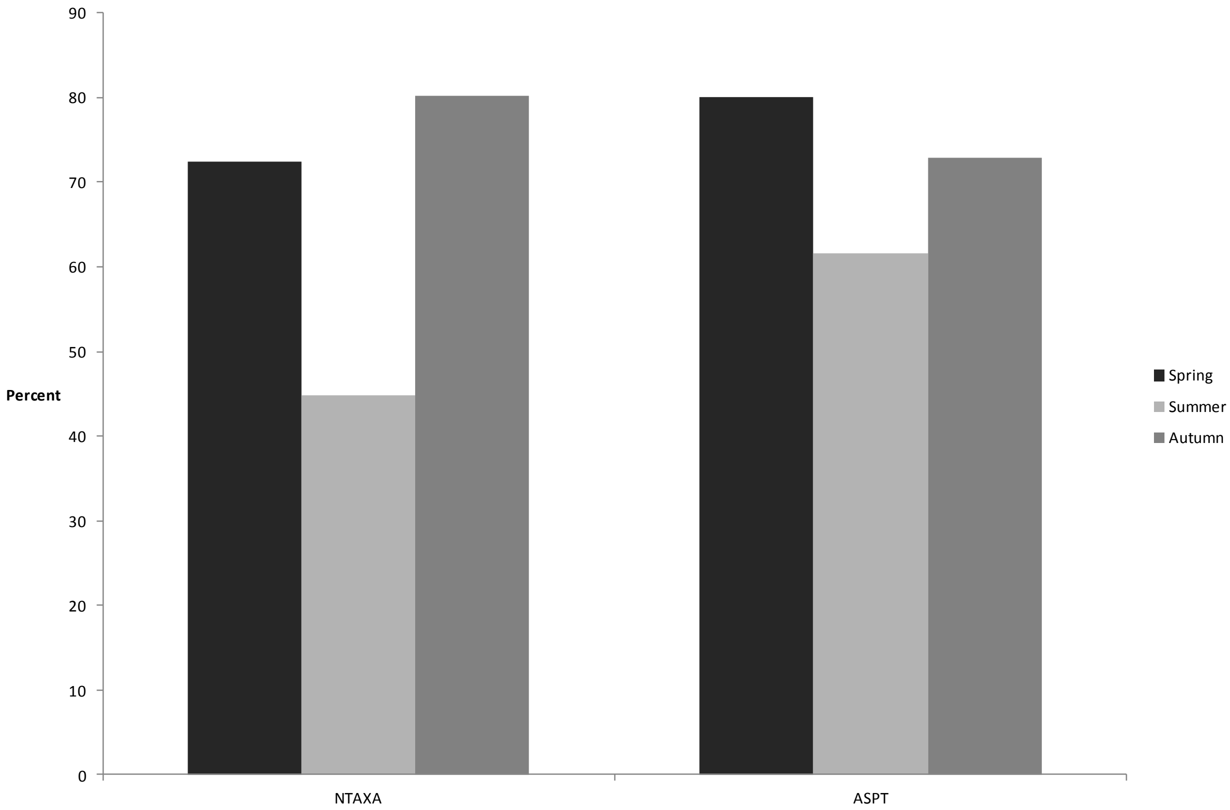

Figure 20: Percentage of sites where the RIvPACS reference value does not fall within the BIOSYS mean +/- SE

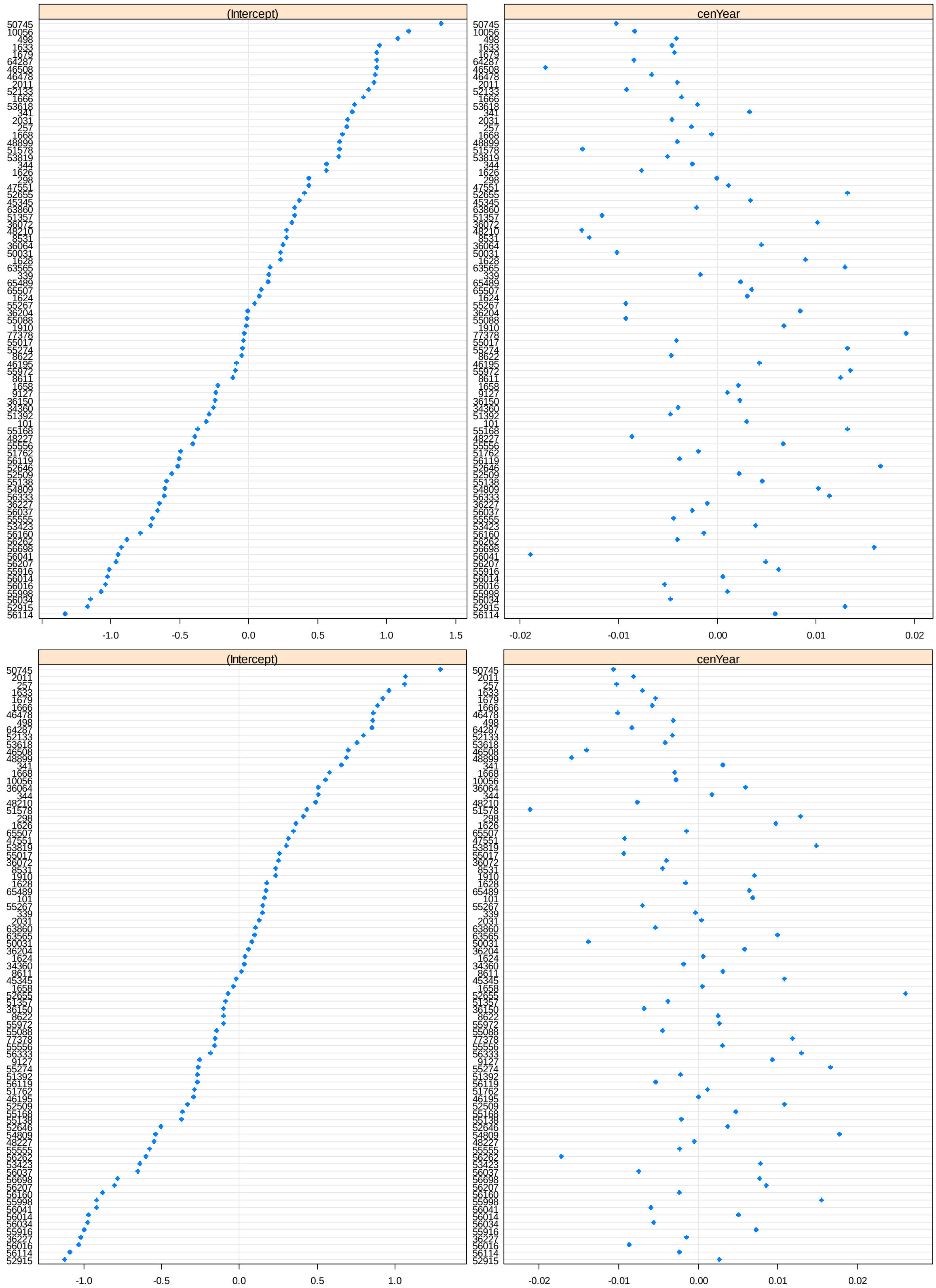

Figure 21: Random effects plot for ASPT, top = spring; bottom = autumn

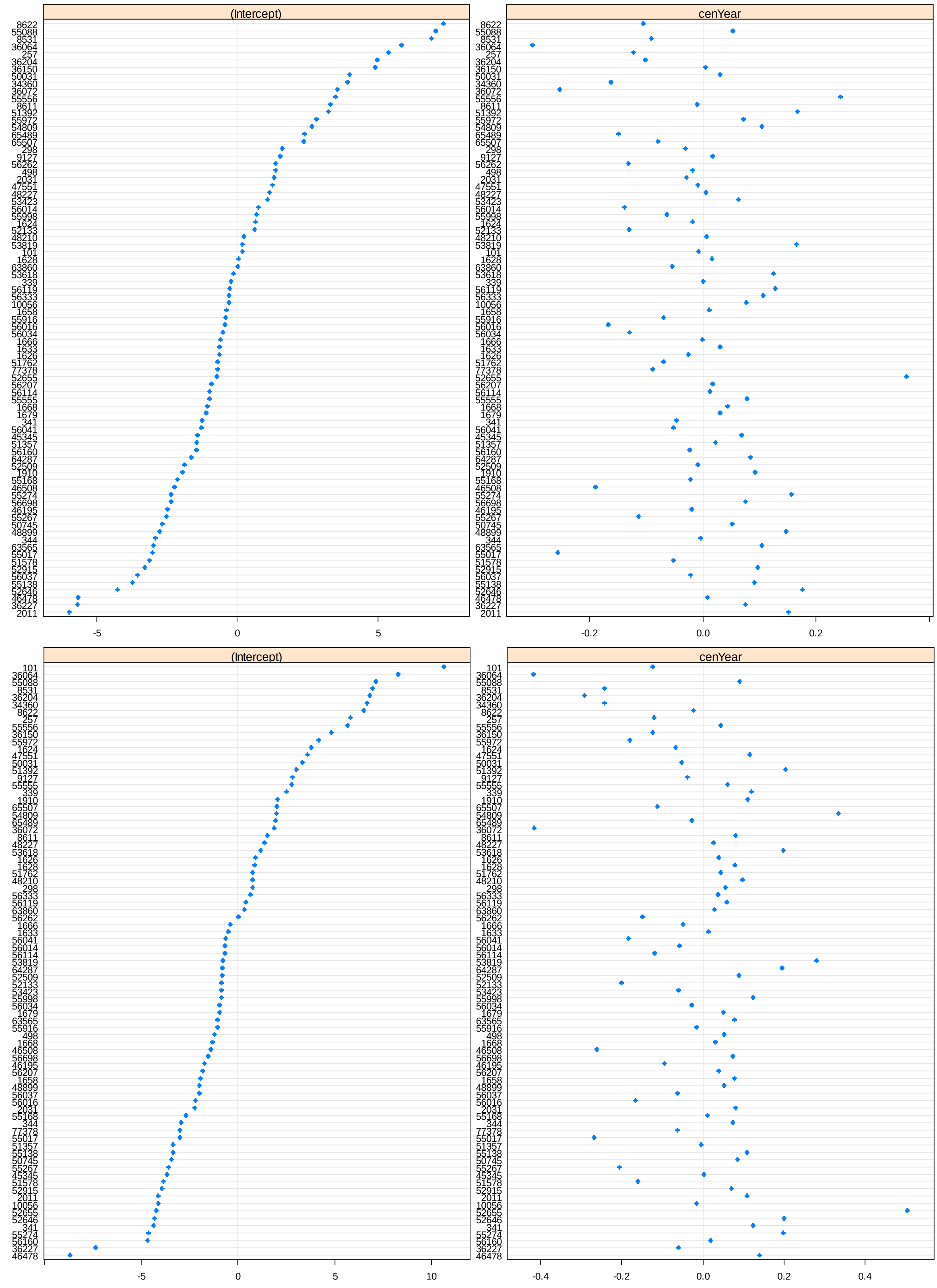

Figure 22: Random effects plot for Ntaxa, top = spring: bottom = autumn

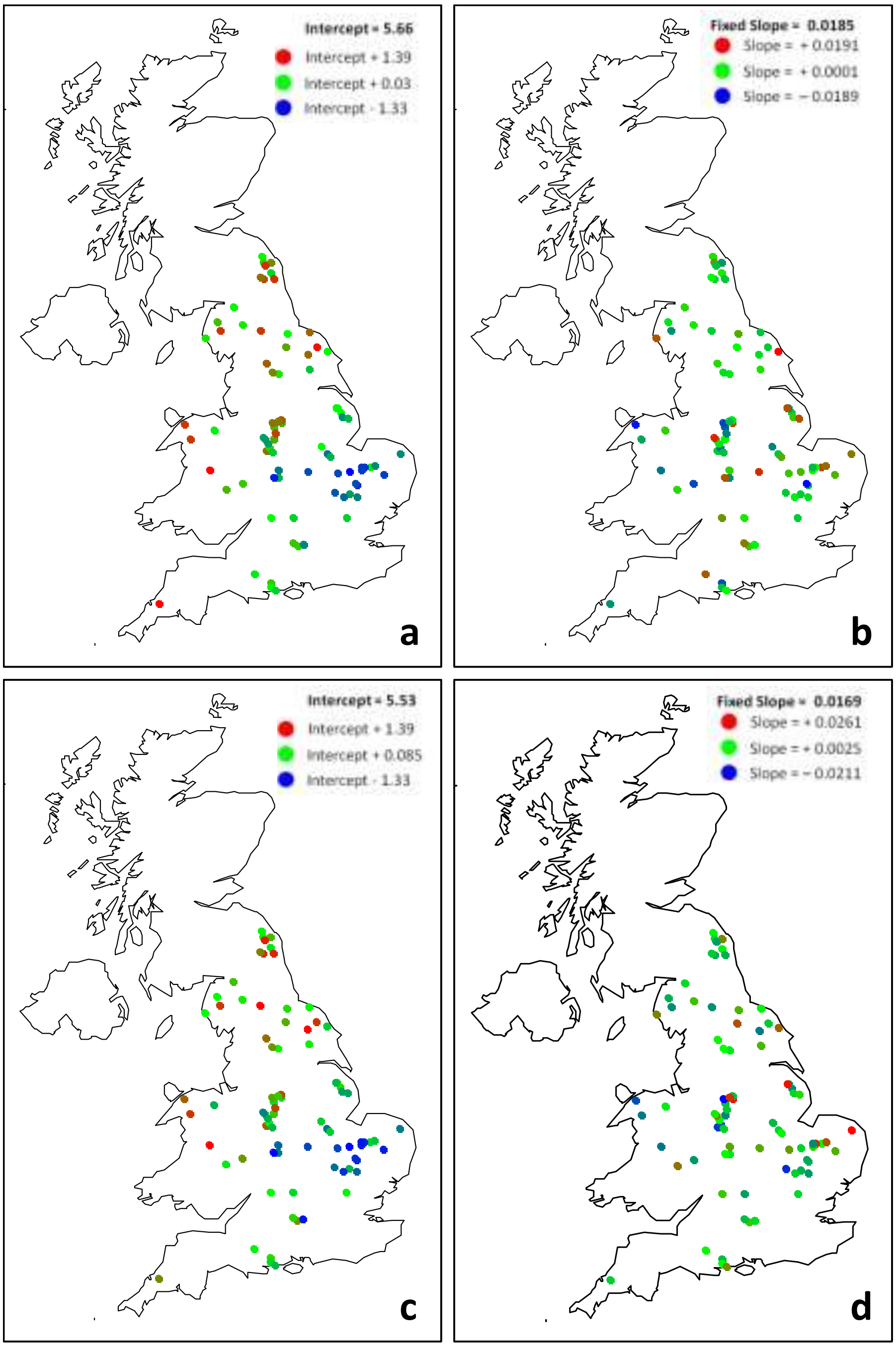

Figure 23: Matched sites showing the results of the mixed effects model for ASPT. a) and b) show the random intercepts and slopes, respectively for spring. c) and d) show the random intercepts and slopes, respectively for autumn

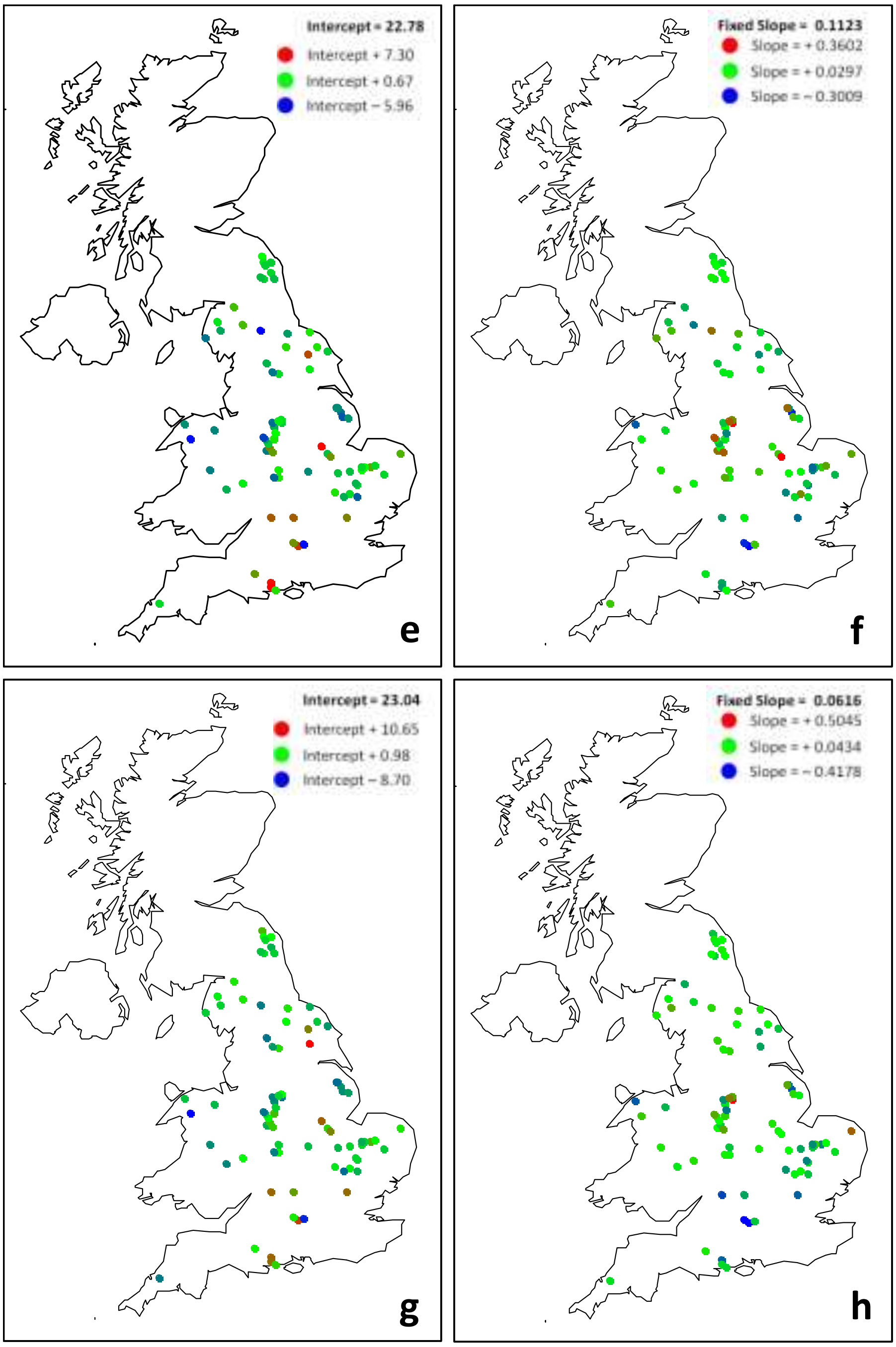

Figure 24: Matched sites showing the results of the mixed effects model for Ntaxa. e) and f) show the random intercepts and slopes, respectively for spring. g) and h) show the random intercepts and slopes, respectively for autumn

Table 5: Fixed effects of the ASPT Spring mixed effects model, showing the estimate and Standard Error terms l. Bold figures show significant t-values

	Sample No.	Estimate	S.E.	t	Deleted model component
Fm1 Intercept cenYear R_BR cenYear:R_BR	 1093	 5.662561 0.018561 0.020647 -0.014075	 0.074918 0.002155 0.086689 0.008177	 **75.58** **8.61** 0.24 -1.72	n/a Full Model
Fm2 Intercept cenYear cenYear:R_BR	 1093	 5.662894 0.018555 -0.015771	 0.074940 0.002157 0.004061	 **75.57** **8.60** **-3.88**	Operator (R_BR)
Fm3 Intercept R_BR R_BB:cenYear R_BR:cenYear	 1093	 5.662561 0.020647 0.018561 0.004486	 0.074918 0.086689 0.002155 0.008082	 **75.58** 0.24 **8.61** 0.56	Time (cenYear)
Fm4 Intercept cenYear R_BR	 1093	 5.664748 0.017914 0.150012	 0.074700 0.002102 .0150012	 **75.83** **8.52** **3.48**	Time operator interaction (cenYear: R_B)
Fm5 Intercept cenYear R_BR cenYear:R_BR	 1093	 5.668713 0.017771 0.046604 -0.009613	 0.074756 0.001576 0.085669 0.007792	 **75.83** **11.27** 0.54 -1.23	Random slopes - Fixed slopes only
Fm6 Intercept	 1093	 5.79800	 0.07291	 **79.52**	Intercept only

Table 6: Random effects for Spring ASPT, showing the variance, standard deviation (s.d.) and AIC and BIC model criterion.

	N	Groups	AIC	BIC	Variance	s.d.
Fm1	1093	81	1038	1078		
BIOSYS_SITE_ID (Intercept)					0.4402783	0.66353
cenYear					0.0001423	0.01193
Residual					0.1067346	0.32670
Fm2	1093	81	1036	1071		
BIOSYS_SITE_ID (Intercept)					0.440713	0.66386
cenYear					0.000143	0.01196
Residual					0.106717	0.32668
Fm3	1093	81	1038	1078		
BIOSYS_SITE_ID (Intercept)					0.4402783	0.66353
cenYear					0.0001423	0.01193
Residual					0.1067346	0.32670
Fm4	1093	81	1039	1074		
BIOSYS_SITE_ID (Intercept)					0.4377271	0.66161
cenYear					0.0001359	0.01166
Residual					0.1072536	0.32750
Fm5	1093	81	1058	1088		
BIOSYS_SITE_ID (Intercept)					0.4386	0.6623
Residual					0.1148	0.3388
Fm6	1093	81	1089	1114		
BIOSYS_SITE_ID (Intercept)					0.45385	0.67368
cenYear					0.00039	0.01975
Residual					0.10854	0.32946

Table 7: Model validation by comparison of model components, using χ^2 distributions. Significance codes: 0 * = 0.001, ** = 0.01, * = 0.05**

Anova comparisons	DF	χ^2	P
Fm1 vs Fm2	1	0.0564	0.8123
Fm1 vs Fm3	0	0	1
Fm1 vs Fm4	1	2.9325	0.08682
Fm1 vs Fm5	2	24.35	**5.158e-06 *** **
Fm1 vs Fm6	3	57.665	**1.853e-12 *** **

Table 8: Fixed effects of the ASPT Autumn mixed effects model, showing the estimate and Standard Error terms. Bold figures show at or near significant t-values.

	No. Observations	Estimate	S.E.	T	Deleted model component
Fm11 Intercept cenYear R_BR cenYear:R_BR	1111	 5.538296 0.016975 -0.154508 -0.019164	 0.068505 0.002340 0.093934 0.008662	 **80.85** **7.25** -1.64 **-2.21**	n/a Full Model
Fm21 Intercept cenYear cenYear:R_BR	1111	 5.534356 0.017087 -0.006820	 0.068191 0.002337 0.004334	 **81.16** **7.31** -1.57	Operator (R_BR)
Fm31 Intercept R_BR R_BB:CenYear R_BR:cenYear	1111	 5.538296 -0.154508 0.016975 -0.0021189	 0.068505 0.093934 0.002340 0.008597	 **80.85** -1.64 **7.25** -0.25	Time (cenYear)
Fm41 Intercept cenYear R_B	1111	 5.538794 0.016154 0.025643	 0.068263 0.002304 0.047063	 **81.14** **7.01** 0.54	Time operator interaction (cenYear: R_B)
Fm51 Intercept cenYear R_B cenYear:R_BR	1111	 5.541459 0.016731 -0.42901 -0.017168	 0.068638 0.001718 0.092772 0.008282	 **80.74** **9.74** -1.54 -2.07	Random slopes - Fixed slopes only
Fm61 Intercept	1111	 5.62655	 0.06644	 **84.68**	Intercept only

Table 9: Random effects for Autumn ASPT, showing the variance, standard deviation (s.d.) and AIC and BIC model criterion.

	No. of observations	Groups	AIC	BIC	Variance	s.d.
Fm1	1111	81	1220	1260		
BIOSYS_SITE_ID(Intercept)					0.3646017	0.60382
cenYear					0.0001631	0.01277
Residual					0.1278577	0.35757
Fm2	1111	81	1221	1256		
BIOSYS_SITE_ID(Intercept)					0.3615325	0.6013
cenYear					0.0001614	0.0127
Residual					0.1283039	0.3582
Fm3	1111	81	1220	1260		
BIOSYS_SITE_ID(Intercept)					0.3646017	0.60382
cenYear					0.0001631	0.01277
Residual					0.1278577	0.35757
Fm4	1111	81	1223	1258		
BIOSYS_SITE_ID(Intercept)					0.3618486	0.60154
cenYear					0.0001604	0.01266
Residual					0.1285952	0.35860
Fm5	1111	81	1238	1268		
BIOSYS_SITE_ID(Intercept)					0.3664	0.6053
Residual					0.1367	0.3697
Fm5	1111	81	1257	1282		
BIOSYS_SITE_ID(Intercept)					0.3656875	0.60472
cenYear					0.0004365	0.02089
Residual					0.1283996	0.35833

Table 10: Model validation by comparison of model components, using χ^2 distributions. Significance codes: 0 * = 0.001, ** = 0.01, * = 0.05**

	DF	χ^2	P
Fm1 / Fm2	1	2.6908	0.1009
Fm1 / Fm3	0	0	1
Fm1 / Fm4	1	4.8667	0.02738*
Fm1 / Fm5	2	21.9	1.755e-05***
Fm1 / Fm6	3	42.767	2.758e-09 ***

Table 11: Fixed effects of the Spring NTAXA fixed effects model, showing the estimate and Standard Error terms. Bold figures show at or near significant t-values.

	No. Observations	Estimate	S.E.	T	Deleted model component
<u>Fm1</u> Intercept cenYear R_BR cenYear:R_BR	1093	 22.78739 0.11249 1.33472 -0.02077	 0.36953 0.02598 0.99979 0.09423	 **61.67** **4.33** 1.33 -0.22	n/a Full Model
<u>Fm2</u> Intercept cenYear cenYear:R_BR	1093	 22.81249 0.11184 -0.12999	 0.36924 0.02620 0.04689	 **61.78** **4.27** -2.77	Operator (R_BR)
<u>Fm3</u> Intercept R_BR R_BB:CenYear R_BR:cenYear	1093	 22.78739 1.33472 0.11249 0.09173	 0.36953 0.99979 0.02598 0.09365	 **61.67** 1.33 **4.33** 0.98	Time (cenYear)
<u>Fm4</u> Intercept cenYear R_B	1093	 22.78981 0.11159 1.52621	 0.36935 0.02562 0.49693	 **61.70** **4.36** **3.07**	Time operator interaction (cenYear: R_B)
<u>Fm5</u> Intercept cenYear R_B cenYear:R_BR	1093	 22.72555 0.12019 1.83661 0.01928	 0.36574 0.01815 0.98909 0.08943	 **62.14** **6.62** 1.86 0.22	Random slopes - Fixed slopes only
<u>Fm6</u> Intercept	1093	 23.1526	 0.3657	 **63.32**	Intercept only

Table 12: Random effects for Spring NTAXA, showing the variance, standard deviation (s.d.) and AIC and BIC model criterion.

	No. of observations	Groups	AIC	BIC	Variance	s.d.	R^2
Fm1 BIOSYS_SITE_ID(Intercept) cenYear Residual	1093	81	6256	6296	 9.28218 0.02374 14.24591	 3.0467 0.1541 3.7744	
Fm2 BIOSYS_SITE_ID(Intercept) cenYear Residual	1093	81	6256	6291	 9.28709 0.02455 14.25348	 3.0475 0.1567 3.7754	
Fm3 BIOSYS_SITE_ID(Intercept) cenYear Residual	1093	81	6256	6296	 9.28218 0.02374 14.24591	 3.0467 0.1541 3.7744	
Fm4 BIOSYS_SITE_ID(Intercept) cenYear Residual	1093	81	6254	6289	 9.28282 0.02364 14.24859	 3.0468 0.1538 3.7747	
Fm5 BIOSYS_SITE_ID(Intercept) Residual	1093	81	6282	6312	 9.067 15.554	 3.011 3.944	
Fm5 BIOSYS_SITE_ID(Intercept) cenYear Residual	1093	81	6271	6296	 9.54501 0.03355 14.30149	 3.0895 0.1832 3.7817	

Table 13: Model validation by comparison of model components, using χ^2 distributions. Significance codes: 0 * = 0.001, ** = 0.01, * = 0.05**

	DF	χ^2	P
Fm1 / Fm2	1	1.7688	0.1835
Fm1 / Fm3	0	0	1
Fm1 / Fm4	1	0.0484	0.8259
Fm1 / Fm5	2	29.675	3.599e-07 ***
Fm1 / Fm6	3	20.423	0.0001387 ***

Table 14: Fixed effects of the Autumn NTAXA fixed effects model, showing the estimate and Standard Error terms. Bold figures show at or near significant t-values.

	No. Observations	Estimate	S.E.	T	Deleted model component
Fm1	1111				n/a
Intercept		23.04529	0.45024	**51.18**	Full Model
cenYear		0.06180	0.03070	**2.01**	
R_BR		2.50768	1.05400	**2.38**	
cenYear:R_BR		0.09763	0.09822	0.99	
Fm2	1111				Operator
Intercept		23.10508	0.44747	**51.63**	(R_BR)
cenYear		0.05950	0.03105	1.92	
cenYear:R_BR		-0.10584	0.04872	**-2.17**	
Fm3	1111				Time
Intercept		23.04529	0.45024	**51.18**	(cenYear)
R_BR		2.50768	1.05400	**2.38**	
R_BB:CenYear		0.06180	0.03070	**2.01**	
R_BR:cenYear		0.15944	0.09906	1.61	
Fm4	1111				Time operator interaction (cenYear: R_B)
Intercept		23.04149	0.44896	**51.32**	
cenYear		0.06570	0.03057	**2.15**	
R_B		1.59716	0.52144	**3.06**	
Fm5	1111				Random slopes - Fixed slopes only
Intercept		22.98877	0.44286	**51.91**	
cenYear		0.06457	0.01946	**3.32**	
R_B		2.93327	1.05352	**2.78**	
cenYear:R_BR		0.13553	0.09397	1.44	
Fm6	1111				Intercept only
Intercept		23.3200	0.4413	**52.85**	

Table 15: Random effects for Spring NTAXA, showing the variance, standard deviation (s.d.) and AIC and BIC model criterion.

	No. of observations	Groups	AIC	BIC	Variance	s.d.
Fm1 BIOSYS_SITE_ID(Intercept) cenYear Residual	1111	81	6500	6540	 14.48935 0.03944 15.61402	 3.8065 0.1986 3.9515
Fm2 BIOSYS_SITE_ID(Intercept) cenYear Residual	1111	81	6503	6538	 14.32518 0.04087 15.68745	 3.7849 0.2022 3.9607
Fm3 BIOSYS_SITE_ID(Intercept) cenYear Residual	1111	81	6500	6540	 14.48935 0.03944 15.61402	 3.8065 0.1986 3.9515
Fm4 BIOSYS_SITE_ID(Intercept) cenYear Residual	1111	81	6499	6534	 14.39423 0.03996 15.62571	 3.7940 0.1999 3.9529
Fm5 BIOSYS_SITE_ID(Intercept) Residual	1111	81	6553	6583	 13.95 17.78	 3.735 4.216
Fm5 BIOSYS_SITE_ID(Intercept) cenYear Residual	1111	81	6506	6531	 14.59486 0.04223 15.74348	 3.8203 0.2055 3.9678

Table 16: Model validation by comparison of model components, using χ^2 distributions. Significance codes: 0 * = 0.001, ** = 0.01, * = 0.05**

	DF	χ^2	P
Fm1 / Fm2	1	5.6195	0.01776 *
Fm1 / Fm3	0	0	1
Fm1 / Fm4	1	0.9818	0.3218
Fm1 / Fm5	2	57.466	3.323e-13 ***
Fm1 / Fm6	3	11.996	0.007398 **

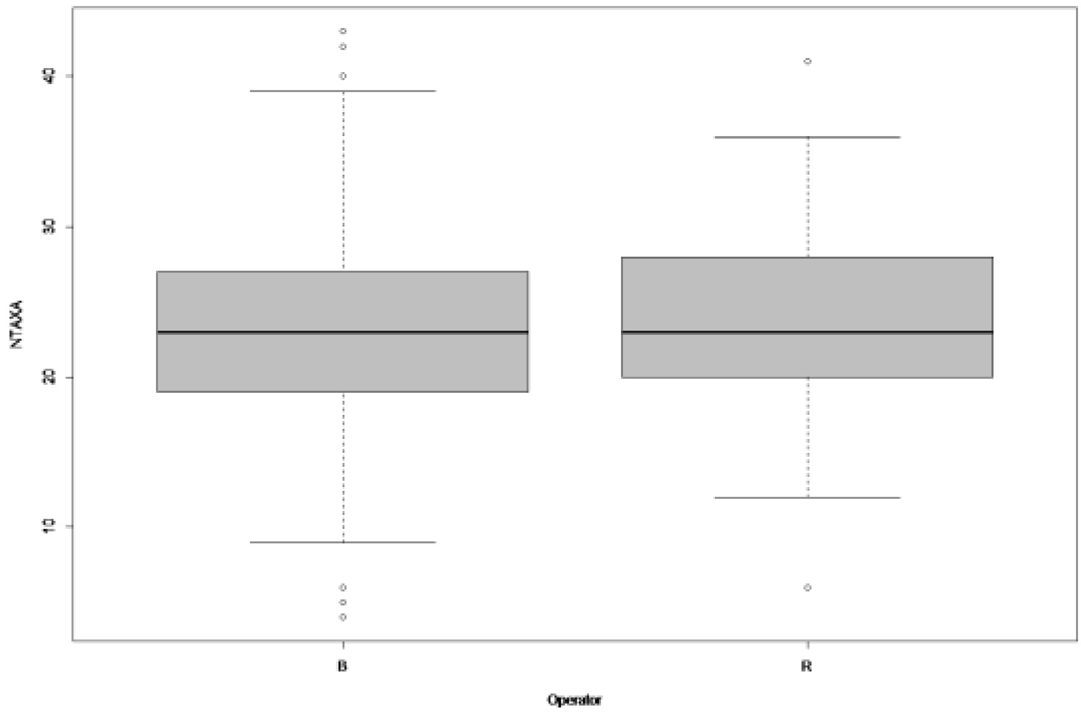

Figure 25 Boxplot of Ntaxa values for B Biosys data and R RIvPACS data showing means and standard errors.

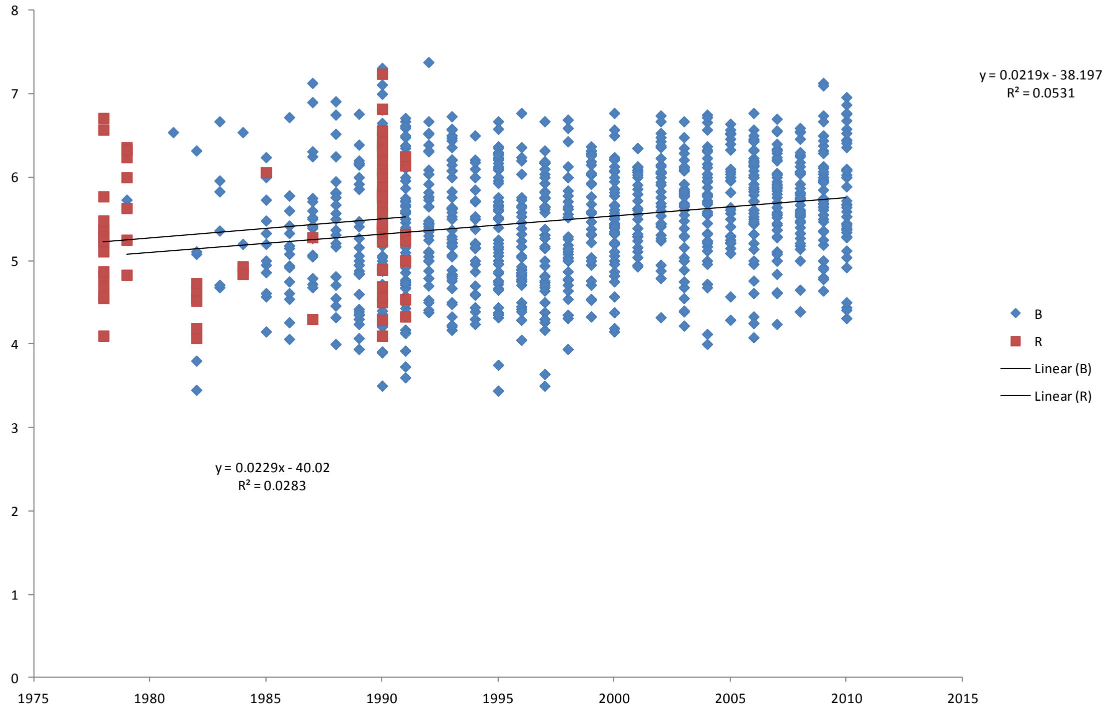

Figure 26: Linear regression analysis of ASPT over time for autumn RIvPACS and BIOSYS samples

4. Conclusions

- This study identified 81 sites used as RIvPACS reference sites that had been monitored at least ten times since by the Environment Agency.
- Deviation from the perceived reference condition was the norm at most sites when assessed with ASPT and Ntaxa, to the point that virtually no samples fell within +/- 5% of the RIvPACS reference values for any of the sites.
- The RIvPACS samples scores did not fall within the standard error range of the mean ASPT and Ntaxa for respectively 70 to 80% of the sites in both Spring and Autumn. Summer samples showed less deviation but the sample size for this season was much smaller than the other two seasons.
- Upward changes from reference condition in ASPT and Ntaxa did not correlate well with map variables, though mean air temperature showed a weak correlation to Ntaxa. Potential increases in temperature with climate change may thus have spatially variable consequences for the ecological quality of rivers, e.g. with more rapid changes in the south of the UK.
- ASPT of the sites exhibited spatial patterns with higher values in upland areas of the UK and lowest in lowland agricultural and urban areas. Rates of change in ASPT were however highest at sites with intermediate ASPT scores.
- Spatial patterns were less evident when using Ntaxa as low and high values were more dispersed across the UK. Rates of change in Ntaxa were also highest at sites with intermediate Ntaxa scores, and rates of change were much higher for spring samples than for autumn samples.

- There is a clear increasing trend in autumn ASPT over time in both data sets, indicative of a shift in reference conditions. This trend has been observed nationally before in impacted rivers, and is attributed to chemical improvements in water quality (Vaughan & Ormerod, 2012), but this is the first time that it has been detected in the reference site network, where improvements in water quality should not be applicable. The fact that Ntaxa does not have the same trend could indicate species replacement at some sites, and is consistent with chemical improvements in water or habitat quality at reference sites. Because the increasing ASPT trend was site dependent, it is less likely that long term climatic cycles, or shifts in climate, are the cause unless they act at differently at small regional scales. These sites could have been in recovery from a past stressor, including natural events such as floods and droughts, or impacted by an undetected stressor at the time of RIvPACS sampling.
- The results raise important questions about the use of the reference condition concept in the WFD. In this study, we have observed that a reference based on year's sampling is insufficient to characterize a reference site because variability is high even at these minimally impacted sites. If the reference is essentially a moving target, it becomes harder to apply EQR's and confidence in assessments will be low.

5. Literature cited

Bowman, M. F. & Somers, K. M. (2005) Considerations when using the reference condition approach for bioassessment of freshwater ecosystems. *Water Quality Research Journal of Canada,* **40,** 347-360.

Davy-Bowker, J., Clarke, R. T., Corbin, T. A., Vincent, H., Pretty, J. L., Hawczak, A., Murphy, J. & Jones, J. I. (2008) *River invertebrate classification tool: Final Report.* . Scottish and Northern Ireland Forum for Environmental Research (SNIFFER Project WFD72c). Edinburgh, UK.

European Commission. (2000) Directive 2000/60/EC of the European Parliament and of the Council 23 October 2000: Establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities*.

Friberg, N., Bonada, N., Bradley, D. C., Dunbar, M. J., Edwards, F. K., Grey, J., Hayes, R. B., Hildrew, A. G., Lamouroux, N., Trimmer, M. & Woodward, G. (2011) Biomonitoring of Human Impacts in Freshwater Ecosystems: The Good, the Bad and the Ugly. *Advances in Ecological Research* (ed W. Guy), pp. 1-68. Academic Press.

Moorhead, K. K. (2013) A Realistic Role for Reference in Wetland Restoration. *Ecological Restoration,* **31,** 347-352.

Nõges, P., Van de Bund, W., Cardoso, A. C. & Heiskanen, A. S. (2007) Impact of climatic variability on parameters used in typology and ecological quality assessment of surface waters - implications on the Water Framework Directive. *Hydrobiologia,* **584,** 373-379.

Pardo, I., Gómez-Rodríguez, C., Wasson, J.-G., Owen, R., van de Bund, W., Kelly, M., Bennett, C., Birk, S., Buffagni, A., Erba, S., Mengin, N., Murray-Bligh, J. & Ofenböeck, G. (2012) The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. *Science of the Total Environment,* **420,** 33-42.

Stoddard, J. L., Larsen, D. P., Hawkins, C. P., Johnson, R. K. & Norris, R. H. (2006) Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecological Applications,* **16,** 1267-1276.

Vaughan, I. P. & Ormerod, S. J. (2012) Large-scale, long-term trends in British river macroinvertebrates. *Global Change Biology,* **18,** 2184-2194.

Wright, J. F., Furse, M. & Armitage, P. (1993) RIVPACS - a technique for evaluating the biological quality of rivers in the UK. *European Water Poolution Control,* **3,** 15-25.