Article:

Rush, Graham, McDarby, Patrick, Edwards, Robin et al. (3 more authors) (2021) Development of an intertidal foraminifera training set for the North Sea and an assessment of its application for Holocene sea-level reconstructions. Marine Micropaleontology. 102055. ISSN: 0377-8398

https://doi.org/10.1016/j.marmicro.2021.102055

Development of an intertidal foraminifera training set for the North Sea and an assessment of its application for Holocene sea-level reconstructions

Graham Rush[a,*], Patrick McDarby[a], Robin Edwards[b], Yvonne Milker[c], Ed Garrett[a], Roland Gehrels[a]

[a]*Department of Environment and Geography, University of York, YO10 5NG, UK*
[b]*School of Natural Sciences, Trinity College Dublin, Ireland*
[c]*Center for Earth System Research and Sustainability, Institute for Geology, University of Hamburg, Germany*

Abstract

Regional datasets of the vertical distribution of intertidal foraminifera are useful to reconstruct Holocene sea-level changes from fossil foraminifera in estuaries and salt marshes. In this paper, we present a new foraminiferal dataset from the Ythan Estuary (Scotland) and combine it with data from eight other coastal sites from England, Denmark and Germany to produce a regional modern training set for the North Sea. We recognise a correlation between foraminifera and tidal elevation which makes the foraminifera suitable as sea-level indicators. We subdivide the data into subregional training sets and develop WA and WAPLS transfer functions. Applying a variety of statistical methods, including detrended canonical analysis, cross-validation by bootstrapping and leave-one-site-out, and the modern analogue technique, we establish the most appropriate transfer function from which to reconstruct early Holocene sea-level changes in a sediment core from the western North Sea coast. Results show that the subregional England/Scotland training set provides the most appropriate sea-level reconstructions, with decimetre-scale uncertainties. The techniques we use in this study, that consider both the modern and fossil assemblages to determine the best training set and transfer function, are suggested as a template for the development of regional transfer functions based on foraminifera and other intertidal microfossils.

Keywords: Sea level, Transfer function, Foraminifera, Salt marsh

---

Highlights

- New foraminifera modern training set for the Ythan Estuary, east Scotland.
- Synthesis of North Sea foraminifera data for sea-level reconstructions.
- Template for determining the most effective regional transfer function.

1. Introduction

Foraminifera have long been recognised as accurate and precise sea-level indicators in Holocene intertidal deposits (Scott and Medioli, 1978). This is because foraminiferal assemblages form in narrow vertical zones across the intertidal zone as a result of variations in tidal submergence. The correlation with elevation relative to the tidal frame provides a tool to reconstruct sea-level changes from fossil intertidal foraminifera preserved in sediment cores (Gehrels, 2000). When the relationship between foraminifera and elevation in the modern environment is quantified, regression methods can be applied to develop predictive transfer functions that are capable of reconstructing sea-level changes from fossil foraminiferal assemblages in subsurface intertidal deposits. Relative sea-level reconstructions based on foraminifera in salt-marsh deposits have been established in many temperate coastal regions with high vertical precision, often quoted as sub-decimetre (e.g. Gehrels *et al.*, 2004; Horton *et al.*, 2009; Hawkes *et al.*, 2010; Engelhart *et al.*, 2011; Wright *et al.*, 2011; Barnett *et al.*, 2016), leading salt marshes to be labelled ‘geological tide gauges’ (Barlow *et al.*, 2013). Consistency is used as an informal indicator and the assumption of uniformitarianism is applied, since it is not possible to assess the accuracy of palaeo reconstructions prior to pre-instrumental records.

---

[*] Corresponding author
Email address: `G.Rush@leeds.ac.uk` (Graham Rush)

Many studies use foraminiferal assemblages in modern samples from a proximal site, collectively known as a training set, to develop local transfer functions that are then applied to reconstruct sea level based on fossil assemblages from a core (e.g. Gehrels *et al.*, 2004, 2005). This method relies on the assumption that the assemblages in the modern environment are an accurate analogue for the past environment and fossil foraminifera preserved in the core. In the absence of appropriate analogues in local training sets, and therefore a likely non-conformity of environmental conditions, a regional training set developed from multiple sites may be more appropriate (Watcham *et al.*, 2013; Barlow *et al.*, 2013). A regional dataset is a collection of training sets from multiple sites along a stretch of coastline, although the size of the region is not defined (Barlow *et al.*, 2013). Regional foraminifera training sets have been developed for regions such as the UK (Horton *et al.*, 1999), southwest Europe (Leorri *et al.*, 2011), Oregon, western USA (Hawkes *et al.*, 2010) and the eastern coast of the USA (Wright *et al.*, 2011; Kemp *et al.*, 2012). The method has been used to produce continuous sea-level records for the late Holocene (e.g. Hawkes *et al.*, 2010; Engelhart *et al.*, 2011) and the early Holocene using foraminifera (Horton *et al.*, 1999) and diatoms (Lawrence *et al.*, 2016). Barlow *et al.* (2013) discuss the relative benefit of local and regional approaches, but in general a regional training set that comprises sites where faunal assemblages are responding to tidal inundation in a similar manner increases natural variability and therefore the potential of providing an analogous environment.

For early and middle Holocene sea-level reconstructions developed from intertidal microfossils, the likelihood that environmental conditions and/or the realised niches of taxa at a single site have remained constant may be small. For example, Edwards and Horton (2000) found that of 26 fossil assemblages from cored intertidal sediments in the south of England, 10 were unlike any other modern surface assemblage. A lack of modern analogues requires the inclusion of additional sites in to a larger training set. Assemblages from multiple sites are more likely to capture a wider range of environmental conditions with a greater chance of providing a suitable modern analogue (Legendre and Fortin, 1989; Juggins and Birks, 2012), with the caveat that the precision of the prediction may be reduced. For example, Lawrence *et al.* (2016) applied a regional diatom transfer function to generate an early Holocene sea-level reconstruction in southwest Scotland based on samples from nine sites on the west coast of Scotland. Yet a key taxon found in the fossil record had become extinct in the area and a further site from the English North Sea coast was included to provide the missing analogue. This reduced the mean precision of the sea-level reconstruction which was calculated at ± 56 cm (Lawrence *et al.*, 2016). These examples highlight that whilst generally reducing precision, regional transfer functions are generally required to ensure reconstructions of early Holocene sea-level changes based on microfossils are more likely to be accurate. We deem a reconstruction to be accurate if it is consistent with others using different training sets and models where foraminifera show a similar inter-site relationship with tidal inundation.

In this paper we compile a new regional training set of intertidal foraminifera for the North Sea basin. The training set includes nine sites from the eastern (England and Scotland) and western (Denmark and Germany) coasts of the North Sea. Data from eight sites have been previously published and we add to this by presenting a new local dataset for a site on the east coast of Scotland, the Ythan Estuary. The aims of this paper are:

1. to study the relationship with elevation relative to sea level of the modern foraminifera in the Ythan Estuary and the regional composite training set, known forthwith as the *North Sea* training set, and hence its suitability for reconstructing sea-level changes.
2. to assess the best choice of transfer function between regional, sub-regional and local training sets, to reconstruct relative sea-level change from an early Holocene sediment core and provide a template for decision making in similar studies.

2. Materials and Methods

2.1. Ythan Estuary training set development

The Ythan Estuary consists of mudflats, salt marshes and freshwater marshes (Fig. 1). The River Ythan is tidal up to 11 km from the mouth of the estuary (Stapleton and Pethick, 1996) and the intertidal area is approximately 1.85 km^2 of which 0.13 km^2 is intact salt marsh. The mean tidal range at the mouth of the estuary is 2.48 m as calculated by taking a weighted average of two nearby tidal gauge stations at Aberdeen (57°8.64' N, 2°4.82' W) and Peterhead (57°29.84' N, -2°13.69' W), 18 km and 15 km away respectively (UK Hydrographic Office, 2016). To the north and west, the estuary is mostly bounded by arable farmland. However, in the 'Sleek of Tarty' area adjacent to the mouth of the Tarty Burn tributary (see Fig. 1) the salt marsh extends into a freshwater marsh and woodland. In this area the salt

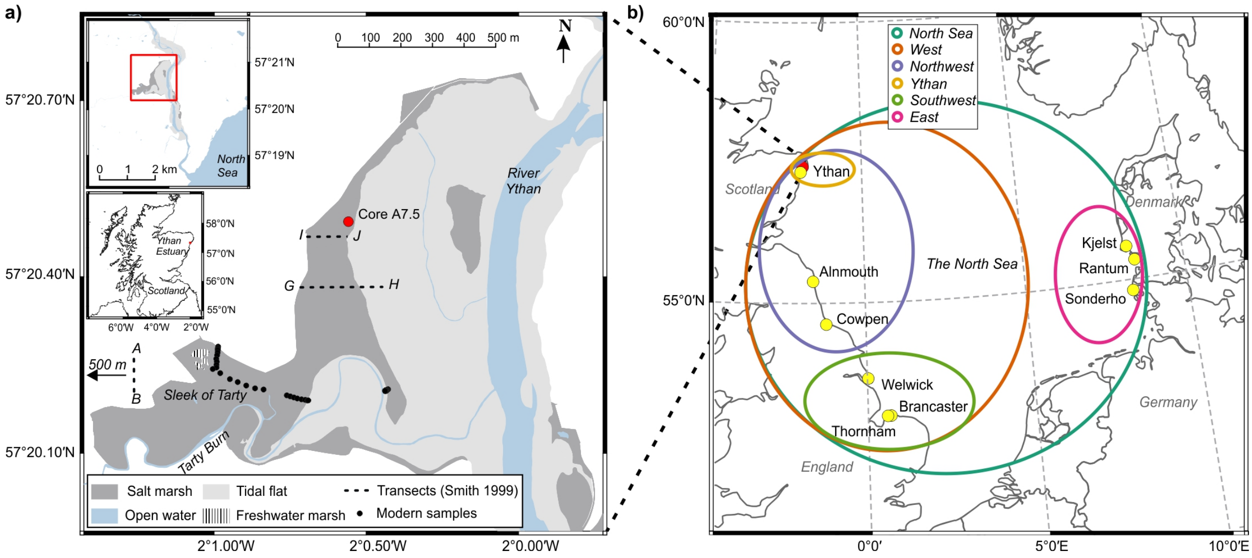


Figure 1: **Map of the Ythan Estuary and North Sea sites.** a) The Ythan Estuary showing the transects of modern samples, the transects from Smith *et al.* (1999) shown in Fig. 6a and the location of core A7.5 described in the text and Fig. 6b. b) Map of modern foraminifera sites (yellow markers) that make up the training sets. Red marker indicates the site of the core used in the reconstruction. Coloured ellipses indicate groupings of sites for regional and subregional datasets. The names for each region or subregion correspond to the main text.

marsh is characteristically undulating with many well defined ponds and creeks. The lower salt marsh generally terminates at a small tidally cut cliff, although occasionally grades into the mudflat.

We established a local benchmark using a Trimble R6 model 3 DGPS with a vertical uncertainty of 0.006 m. The benchmark was tied to two local Ordnance Survey benchmarks to provide elevation measurements relative to the UK national vertical geodetic ordnance datum (OD). Transects were designed to capture the elevation gradient from the highest occurrence of foraminifera (HoF) to mean tide level (MTL) and included a freshwater reed bed, high, middle and low salt marsh zones and mudflat. Sampling was timed to coincide with a large spring tide to maintain standardisation with other studies and to allow sampling at lower elevations (e.g. Avnaim-Katav *et al.*, 2017). Sampling sites were established at c. 5 cm elevation intervals along two transects using a Trimble M1 DR2 total station. Standardised surface samples of 10 cm^2 by 1 cm deep were collected (e.g. Gehrels, 2000). After collection samples were stored at 4 °C before processing.

Preparations of foraminifera samples followed standard laboratory procedures (Gehrels, 2002). The 63 – 500 μm fraction of a 5 cc subsample was wet sieved and stored in a Rose Bengal and ethanol solution within seven days of collection to enable identification of living and dead foraminifera (Murray and Bowser, 2000). Tests containing stained protoplasm within the last few chambers were assumed to be living and only the dead foraminiferal taxa were investigated to avoid seasonal bias (Culver and Horton, 2005) and provide the most representative analogue for material found in fossil records (Horton *et al.*, 1999; Horton and Edwards, 2006; Berkeley *et al.*, 2007; Kemp *et al.*, 2009). Samples were split using a wet splitter (Scott and Hermelin, 1993) and picked and counted until a target of at least 200 individuals was met in even $^1/_8$ splits. If this number was not achieved the full sample was counted. Individuals were identified with reference to the taxonomy of de Rijk (1995); Wright *et al.* (2011); Edwards and Wright (2015); Müller-Navarra *et al.* (2017); Hayward *et al.* (2020). A total of 37 samples were counted and individual taxa expressed as their percentage relative abundance of the sample.

2.2. *Regional training set compilation*

Existing modern salt-marsh foraminifera training sets, collected from around the North Sea, include Alnmouth, Brancaster, Cowpen, Thornham and Welwick (Horton and Edwards, 2006), Brancaster (Gehrels *et al.*, 2001), Kjelst (Gehrels and Newman, 2004), Rantum, Soønderho (Müller-Navarra *et al.*, 2017) and Tümlau (Müller-Navarra *et al.*, 2016) (see Fig. 1). All of the data was included except Tümlau, because of human interference that makes it unsuitable for sea-level reconstructions (Müller-Navarra *et al.*, 2016), and two Alnmouth transects because of our doubts over the accuracy of the elevation measurements. Tidal datums were taken from the original publications and the tidal range was modelled for each site using the TPXO8-ATLAS global model of ocean tides (Egbert and Erofeeva, 2010). The data are summarised in Table 1 and throughout the text.

In order for the local training sets to be incorporated into a regional model, taxonomy was standardised using the World Register of Marine Species (Hayward *et al.*, 2020) (Supplementary info Table 1) and the sample elevation was standardised using a standardised water level index (SWLI) which accounts for differences in tidal range between the sites (Zong and Horton, 1999; Gehrels, 2000). The lowest common phylogenic level was applied which resulted in some taxa being grouped at genus level. The only common environmental variable across all sites was elevation relative to a local datum. While inundation frequency is a more direct ecological parameter (Gehrels, 2000; Gehrels *et al.*, 2001; Müller-Navarra *et al.*, 2017), elevation is a linear approximation that is widely used and is applied here in the absence of inundation information across all sites. Generally MTL and mean high water spring tide (MHWS) or mean higher high water (MHHW) have been used for standardisation between sites with different tidal ranges in sea-level studies; however, transfer functions are sensitive to the choice of tidal datum (Woodroffe and Long, 2010; Wright *et al.*, 2011). Woodroffe and Long (2010) demonstrated that highest astronomical tide (HAT) causes less distortion in the upper tidal range, a key locale in foraminifera reconstructions (Gehrels *et al.*, 2004). Wright *et al.* (2011) investigated this cross site standardisation further and concluded that the well-established ecological relationship between HoF and the upper limit of marine influence is more effective yet. However, sampling up to the elevation of HoF is only definitively reached at two of the sites (Brancaster and Ythan) and we therefore use the highest common datum, HAT, in the standard SWLI equation (Eq. 1) (Horton *et al.*, 1999)

$$SWLI_n = \frac{100(h_n - MTL_s)}{HAT_s - MTL_s} + 100 \tag{1}$$

such that where $SWLI_n$ is the standardised water level index of the sample (n), h_n is the height of the sample in the local datum, MTL_s and HAT_s are the mean tide level and highest astronomical tide of the site (s) in the local datum.

The full regional training set, known forthwith as the *North Sea training set,* was sub-divided in order to test the effect on model performance of training set size and tidal range, that naturally coincides with geographic regions (Fig. 1b) owing to the North Sea tidal set up.

The data was screened such that samples were retained if total counts were > 75 in order to be able to reconstruct sea-level change at a decimetre- and decadal-scale (Kemp *et al.*, 2020). The nine sites provided a total of 265 samples of foraminifera assemblages from 52 different taxa. The screening of the raw data of each training set resulted in between 17 and 35 % (*Northwest* and *Ythan* respectively) of samples being removed. The large variation in samples being removed is largely due to zero counts being included in some sites.

Table 1: Summary of intertidal sites included in the regional training sets. Geomorphic classifications are based on Allen (2000). The number of taxa and samples are given for the raw data and after screening.

Site	**Classification**	**No. of samples** Raw / Screened	**No. of taxa**	**Tidal range (m)**	**References**
Alnmouth	Estuarine back-barrier	20 / 20	29	3.13	Horton and Edwards (2006)
Brancaster	Open coast back-barrier	59 / 49	24	3.89	Gehrels *et al.* (2001); Horton and Edwards (2006)
Cowpen	Estuarine back-barrier	31 / 30	24	3.21	Horton and Edwards (2006)
Kjelst	Open embayment	27 / 12	12	1.04	Gehrels and Newman (2004)
Rantum	Open coast back-barrier	33 / 28	14	1.75	Müller-Navarra *et al.* (2017)
Sønderho	Open embayment	14 / 12	12	1.66	Müller-Navarra *et al.* (2017)
Thornham	Open coast back-barrier	24 / 23	23	3.62	Horton and Edwards (2006)
Welwick	Estuarine back-barrier	20 / 20	20	3.92	Horton and Edwards (2006)
Ythan	Estuarine back-barrier	37 / 24	15	2.48	This study

2.3. *Transfer functions and data analysis*

Detrended canonical correspondence analysis (DCCA) (Hill and Gauch, 1980) was applied using the software CANOCO version 5.1 (ter Braak and Smilauer, 2012) to test whether the taxa response is linear or unimodal along the elevation gradient. All datasets had a DCCA axis gradient length greater than two standard deviations, generally accepted as being the threshold, such that unimodal statistical models are appropriate for exploration of the training sets (Birks, 1995). We therefore developed unimodal transfer functions for the six training sets using two different techniques: weighted averaging with classical deshrinking (WA) (ter Braak and Barendregt, 1986) and weighted averaging with partial least square regression (WAPLS) (ter Braak and Juggins, 1993).

WA based transfer functions take the average of all the optima of all the taxa weighted such that those with greater abundances are given more prominence. This causes the vertical range to be compressed and thus a 'deshrinking' correction is necessary. Inverse deshrinking, equivalent to WAPLS component 1, is more suitable for reconstructions close to the middle of the environmental gradient, whilst classical deshrinking is more suitable for reconstructions close to the limits of the gradient (Juggins and Birks, 2012; Kemp and Telford, 2015). We aim to capture the maximum environmental gradient and we therefore favour the use of the classical method for the WA transfer functions.

WAPLS based transfer functions (components 2 and higher) exploit correlations that remain in the residuals after fitting SWLIs that are not taken into account in

WA by adjusting taxa optima (Juggins and Birks, 2012; Kemp and Telford, 2015). Each extra component can be seen as adding extra, hypothetical, environmental variables. WAPLS generally outperforms WA because; 1) 'edge effects' that affect WA where optima are overestimated at the low end and underestimated at the high end (Mohler, 1983) can be reduced or eliminated by WAPLS; and/or 2) because in reality the composition of taxa assemblages is influenced by additional factors other than elevation and WAPLS can exploit the resultant structured pattern in the residuals after WA; and/or 3) the fit of outliers may be improved (Birks, 1995; Juggins and Birks, 2012). WAPLS is thus favoured in many sea-level studies (Barlow et al., 2013). We chose the best performing model with the fewest numbers of components, limited to three, following the rule of parsimony (Birks, 2012), only selecting successive components if a significant performance increase, based on the root mean squared error of prediction (RMSEP) and R^2 values (Birks, 1998), of > 5 % is observed. These are from here on known as the WAPLS transfer functions with the number of components given as c (e.g *WAPLS-c1*).

Locally weighted transfer functions (LW) can take advantage of both local and regional training sets by using a set number of closest analogues for each fossil sample as defined by Modern analogue technique (MAT) and predict the sea level using these (Kemp and Telford, 2015). We chose the 50 closest analogues (Birks, 2012) in the final training set and then dynamically developed a WAPLS transfer function for each fossil sample, given as LW-WAPLS.

To assess the best choice of transfer function for a sea-level reconstruction, we collected a core (A7.5) from the Ythan Estuary using a 50 mm Russian corer at 57°20.49' N, 2°0.54' W. The site was selected based on published stratigraphy by Smith *et al.* (1999) who dated the peat at the base of a core at 5.19 m core depth to c. 8.5 ka cal BP. This basal peat is overlain by saltmarsh and mudflat deposits, and topped by a sand deposit that is attributed to the Storegga tsunami of c. 8.15 ka cal BP (Dawson *et al.*, 2011; Bondevik *et al.*, 2012). We took 1-cm-thick subsamples from the core at appropriate depths with respect to lithological and biological changes in an iterative manner between the basal peat at 5.17 m and the base of the sand deposit at 2.28 m core depth. Thus the core spans a period in the early Holocene that is of interest as it potentially includes the sea-level rise associated with the drainage of glacial lakes Agassiz and Ojibway at ca. 8.4 ka cal BP (Barber *et al.*, 1999; Li *et al.*, 2012; Hijma and Cohen, 2010, 2019; Lawrence *et al.*, 2016). A total of 24 samples were prepared and foraminifera counted as described above.

The resultant WA, WAPLS and LW-WAPLS transfer functions were used to predict SWLIs from the core fossil foraminifera. This was converted to an indicative meaning in metres by reversing Eq. 2 using the following equation:

$$I_f = \frac{(SWLI_f - 100)(HAT_b - MTL_b)}{100} + MTL_b \quad (2)$$

where I_f is the indicative meaning of the fossil sample ($_f$) in the local datum, $SWLI_f$ is the predicted SWLI of the fossil sample ($_f$), HAT_b and MTL_b are the highest astronomical tide and mean tide level of the site (b) in the local datum. The transfer functions also produce prediction sample specific errors (*SSE*) for the fossil samples that is converted in the same manner and used as the range of the reconstructions (range = $I_f \pm SSE$).

To assess training set and transfer function performance and decide on the most suitable to use for our core reconstruction, we conducted qualitative and quantitative analyses in a step-wise manner. All analyses were carried out in R version 3.6.1 using the packages rioja (Juggins, 2017), vegan (Oksanen *et al.*, 2019) and fpc (Hennig, 2019). We used the following methods:

1. Detrended Correspondence Analysis (DCA) with fossil samples passively projected to assess the samples and taxa groupings and relationships (Edwards and Wright, 2015).
2. Constrained correspondence analysis (CCA) to quantify the proportion of variance explained.
3. Cluster analysis using partitioning around medoids (PAM) with Euclidean distances (Kaufmann and Rousseeuw, 1990; Rousseeuw, 1987) to recognise clusters of modern and fossil foraminiferal assemblages. The highest average silhouette width was used to determine the appropriate number of partitions (Kemp *et al.*, 2013) .
4. Modern analogue technique (MAT) using the minimum dissimilarity coefficient (MinDC) to measure the dissimilarity between fossil samples and the closest modern analogue using the commonly applied Chord-squared distance metric (Kemp and Telford, 2015). We applied the conservative method of defining 'good' and 'close' modern analogues for fossil samples as having a MinDC lower than the 5th and 20th percentiles respectively and samples with a larger MinDC defined as 'poor' (Watcham *et al.*, 2013; Barlow *et al.*, 2013).

5. Transfer function performance statistics using bootstrapping (boot) and leave-one-site-out (LOSO) cross-validation to assess the quality of the transfer functions and the independence of samples (Telford and Birks, 2005; Payne *et al.*, 2012; Kemp *et al.*, 2013).
6. If using WAPLS with > 2 components the pattern of taxa optima updates were analysed to assess the effect of introducing more statistical complexity that may distort reconstructions (Wright *et al.*, 2011).
7. Ranges of the different core reconstructions were compared to evaluate the likely accuracy of each regional, subregional and local transfer function.
8. The statistical significance of the reconstructions was measured by comparing the reconstructions from the training sets against a set of 999 randomly generated data using redundancy analysis ordination (Telford and Birks, 2011).

3. Results and Discussion

3.1. Training set development

3.1.1. Ythan Estuary data

Foraminifera occur in the Ythan Estuary from the lowest sampled point of just above MTL across the elevation gradient of the tidal zone to just above HAT. Both dead and living individuals were found up to an elevation of 2.45 m OD (SWLI = 202). A total of 15 taxa were found although four had a maximum abundance < 10 %. The majority of taxa show some relationship with elevation (see Fig. 2). Taxa elevation optima (the abundance weighted average of SWLI in all samples in which they occur (ter Braak and Barendregt, 1986)) and tolerances (the abundance weighted standard deviation (Birks *et al.*, 1990)) following WA show that the different taxa occupy a gradient of different elevations (Fig. 2c). The highest marsh samples, above MHWS at 1.96 m OD where tidal inundation is infrequent, are low in total foraminifera counts and hence only one was retained following screening. In these samples *Balticammina pseudomacrescens* appears in relatively high abundance much like in the eastern North Sea marshes (Gehrels and Newman, 2004; Müller-Navarra *et al.*, 2017). The salt-marsh samples are almost entirely composed of agglutinated taxa. *Entzia macrescens*, also commonly referred to as *Jadammina macrescens,* is the most common salt-marsh taxon and is found in similar abundances across the marsh. *Miliammina fusca* is found in greatest abundances in the lower marsh like in many marshes globally (see Berkeley *et al.*, 2007). Cluster analysis (Fig. 2a) shows that the salt-marsh samples can be subdivided into three clusters that are also somewhat evident in the DCA plot (Fig. 2b). A low marsh cluster is formed with high abundance of *Miliammina fusca*, a mid-high marsh cluster with a mixture of agglutinated taxa and a final cluster with *Haplophragmoides* spp. found in unusually high abundances (> 75 %) compared to other UK salt marshes (Horton *et al.*, 1999). There is a clear shift to assemblages dominated by calcareous species at 1.22 m OD just below the transition from salt marsh to mudflat, along with occasional *Trochammina ochracea* and *Haplophragmoides* spp. PAM clustering displays two distinct clusters of these mudflat samples, one with *Haynesina germanica* most abundant and featuring high abundances of *Buliminella elegantissima* and *Brizalina variabilis*, and a second with few *Buliminella elegantissima* and *Brizalina variabilis* and dominated by *Elphidium williamsoni*, *Trochammina ochracea* and *Haplophragmoides* spp. Though usually found in the vegetated zone (Berkeley *et al.*, 2007) *Haplophragmoides* spp. are also found on mudflats in the east of the North Sea (Müller-Navarra *et al.*, 2017). CCA shows that 24.59 % of the variance can be explained by elevation. The modern foraminifera surface distributions of the sampled tidal zone in the Ythan Estuary display a well-established relationship with elevation and are therefore suitable sea-level indicators.

3.1.2. Compiled North Sea data

The foraminifera of the eight previously published sites and the Ythan (see Table 1) combined in the regional *North Sea* data occur at elevations from 273 SWLI, well above HAT, to 70, below MTL (Fig. 3). The highest 14 samples (> 210 SWLI), well above HAT, are all from the sites in Denmark and Germany. This could be due to the wind-induced setup that occurs in these areas (Bartholdy *et al.*, 2004; Pedersen *et al.*, 2009), where tides > 3 m above HAT in Esbjerg have been recorded (Bartholdy *et al.*, 2010). These samples are generally restricted to *Balticammina pseudomacrescens* and *Entzia macrescens* which are species able to survive despite infrequent flooding (Berkeley *et al.*, 2007). Although micro tidal sites are generally desired because they provide higher precision reconstructions (Callard *et al.*, 2011; Barlow *et al.*, 2013; Edwards and Wright, 2015) they will be more influenced by this effect and hence the micro tidal range of the eastern sites (all < 1 m) may be causing the extremely high SWLI values seen in these sites a phenomenon also seen in some sites in the US (Wright *et al.*, 2011). The assemblages of the remaining salt-marsh samples are dominated by ag-

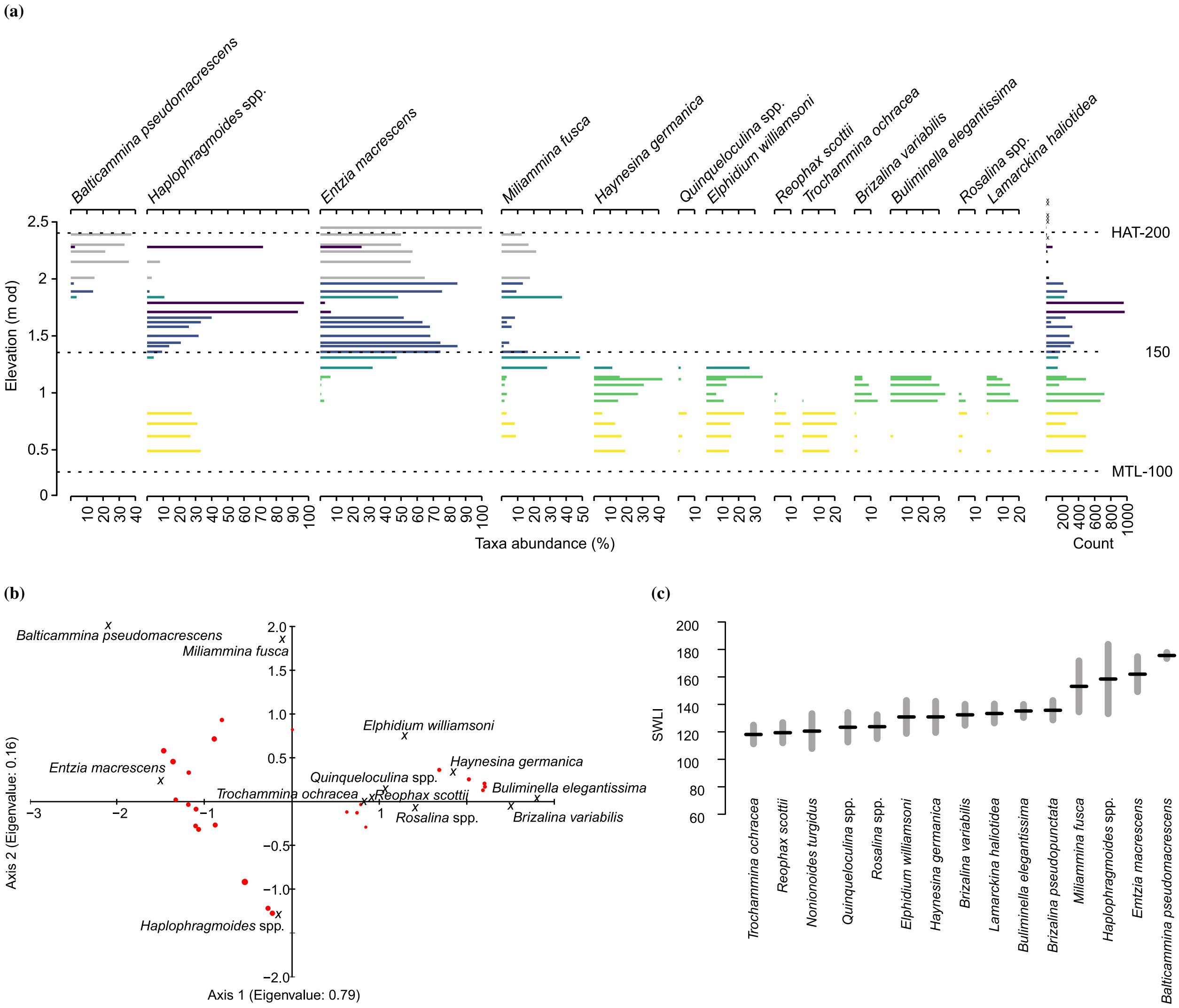


Figure 2: **Summary diagrams of the modern foraminifera assemblages for the Ythan Estuary**: a) Foraminifera abundances as percentages of the count of total dead specimens (given in the final column) against elevation. Only taxa with > 5 % maximum abundance are shown. Grey bars denote samples that were excluded following screening and Xs mark where no foraminifera were found. The bars are coloured according to PAM clustering of the dataset (average silhoutte width = 0.53). Dashed lines indicate tidal datums and SWLI values. b) Detrended correspondence analysis (DCA) of surface foraminifera assemblages showing samples in circles, sized according to SWLI, and taxa. c) Taxa tolerances (grey bar) and optima (black lines) of the Ythan Estuary foraminifera taxa following WA with classical deshrinking.

glutinated foraminifera. *Entzia macrescens* is abundant across the marsh, as is *Miliammina fusca* although in lower abundances. *Trochammina inflata* appears mainly restricted to the higher marsh, while *Haplophragmoides* spp. become more prevalent in the middle-lower marsh mainly at Rantum, Sønderho and Ythan. Calcareous taxa become more abundant in the lower marsh, although are found much higher particularly in eastern marshes which may be due to the wind-induced setup and/or high pH promoting better preservation (Müller-Navarra *et al.*, 2016). Relatively high abundances of *Ammonia* spp. are found at Brancaster and Thornham compared to the other sites, while *Haynesina germanica* is notably abundant at Welwick. Foraminifera were not sampled below 140 SWLI, approximately at the transition between the salt marsh and the mudflat, at Brancaster, Kjelst and Thornham. However, the assemblages from the sites that were sampled show a clear dominance of calcareous taxa. The taxa abundances appear to show a relationship with elevation across the training set, although it may be reduced at the eastern sites due to local conditions.

The relationship between taxa and elevation is confirmed by the cluster analysis (Figs. 3 and 4) and DCA (Fig. 5). PAM clustering produces the highest average silhouette width when samples are split into three clusters, shown in Figs. 3 and 4. Cluster 1 mainly comprises the mudflat samples (defined as SWLI approximately $<$ 140) dominated by calcareous taxa from a mixture of sites. Cluster 2 mainly comprise samples where *Haplophragmoides* spp. are abundant as described above. While *Haplophragmoides* spp. are not rare taxa and are found at varying elevations relative to tidal inundation in other marshes (e.g. Berkeley *et al.*, 2007, for a review), they are marked out in these samples by its high abundance that cause high intra-cluster similarity and low dissimilarity with other samples. Finally, samples in cluster 3 contain the majority of salt-marsh samples. The majority of samples are dominated by *Entzia macrescens* and encompass the full elevation range of the salt marsh. A cluster of samples that are almost entirely found above 170 SWLI and comprise *Entzia macrescens*, *Miliammina fusca* and *Trochammina inflata* is identified from within cluster 3 if we add extra clusters.

A feature of the data appears to be a small degree of clustering by sites. This could be due to a lack of independence of samples because of spatial auto-correlation that may cause transfer function performance statistics to be overly optimistic (Telford and Birks, 2005; Payne *et al.*, 2012), or that each site has unique controls that could make a regional transfer function either problematic or arguably more robust (Legendre and Fortin, 1989). To investigate this further LOSO cross-validation was applied and is discussed below.

The DCA plots summarise the unconstrained relationship between samples, taxa and SWLI (Fig. 5). The plot shows that the samples tend to align towards axis 1 and appear somewhat correlated with SWLI with general clustering of higher and lower SWLI samples. The relationship is confirmed by CCA such that 7.13 % of the variance can be explained by elevation, and that the ratio of the axis constrained by SWLI and the first unconstrained axis is 1.04, indicating it is an important variable in explaining taxa distributions (Kemp *et al.*, 2013). A number of samples diverge from axis 1 in the DCA plot and may therefore be influenced by secondary variables. These samples are those described above, where *Haplophragmoides* spp. are found at higher elevations. Although other environmentally variables are generally correlated with elevation (i.e. salinity, pH), it is evident that other environmental variables are affecting the taxa niches which is not unexpected considering the inclusion of multiple sites from differing environments and tidal ranges. The taxa-sample relationships shown in Fig. 5b confirm the effect of *Haplophragmoides* spp. and also notably how *Balticammina pseudomacrescens* drives the samples with the highest elevations. Collections of taxa are evident along axis 1, where agglutinated and calcareous taxa show a general partition. Some modern Ythan samples appear to be consistently distal from other sites and/or axis 1 suggesting they are providing assemblages with somewhat differing response to elevation. However, similarities with other samples from across the North Sea suggest comparable conditions may be occurring at a wide range of sites.

PAM and DCA analysis suggests that the foraminifera assemblages that make up the *North Sea* training set are appropriate as sea-level indicators. The agglutinated taxa appear particularly well suited, as observed by others, although may lack more definitive zonation (e.g. Horton *et al.*, 1999; Gehrels, 2000; Gehrels *et al.*, 2005; Horton and Edwards, 2006; Kemp *et al.*, 2013; Barnett *et al.*, 2016). Calcareous taxa are useful indicators of tidal flats in the data and hence useful for transfer functions where a wide environmental range may be sought. However, the lowest occurrence of some taxa is not sampled meaning the full range is not captured, which could be problematic for predicting the lowest range of some fossil samples and so these should be treated with an element of caution (Woodroffe, 2009). Calcareous taxa also occur across a wide elevation range in the region, which may impact the predictive ability of the transfer functions. This wide range may be due to in-wash resulting in allochthonous foraminifera (Murray, 2003; Horton and Murray, 2006). Some studies attempt to avoid this effect by either not sampling the tidal flat (e.g. Gehrels, 2000; Kemp *et al.*, 2013)) and/or removing these so termed 'exotic' taxa (e.g. Horton and Edwards, 2006; Kemp *et al.*, 2009; Leorri *et al.*, 2011; Mills *et al.*, 2013) resulting in better predictability in some cases. However, Mills *et al.* (2013) observe that performance does not improve in their Mersey training set and that this method does not preclude the mixing of lower marsh agglutinated foraminifera. We include all samples as our core includes a fossil mudflat environment and thus the lower samples are critical as well as providing a lower limit to some of the agglutinated marsh taxa (Wright *et al.*, 2011). However, other cores that are limited to salt marsh environments may warrant investigation of the effect of removing lower elevation samples or taxa.

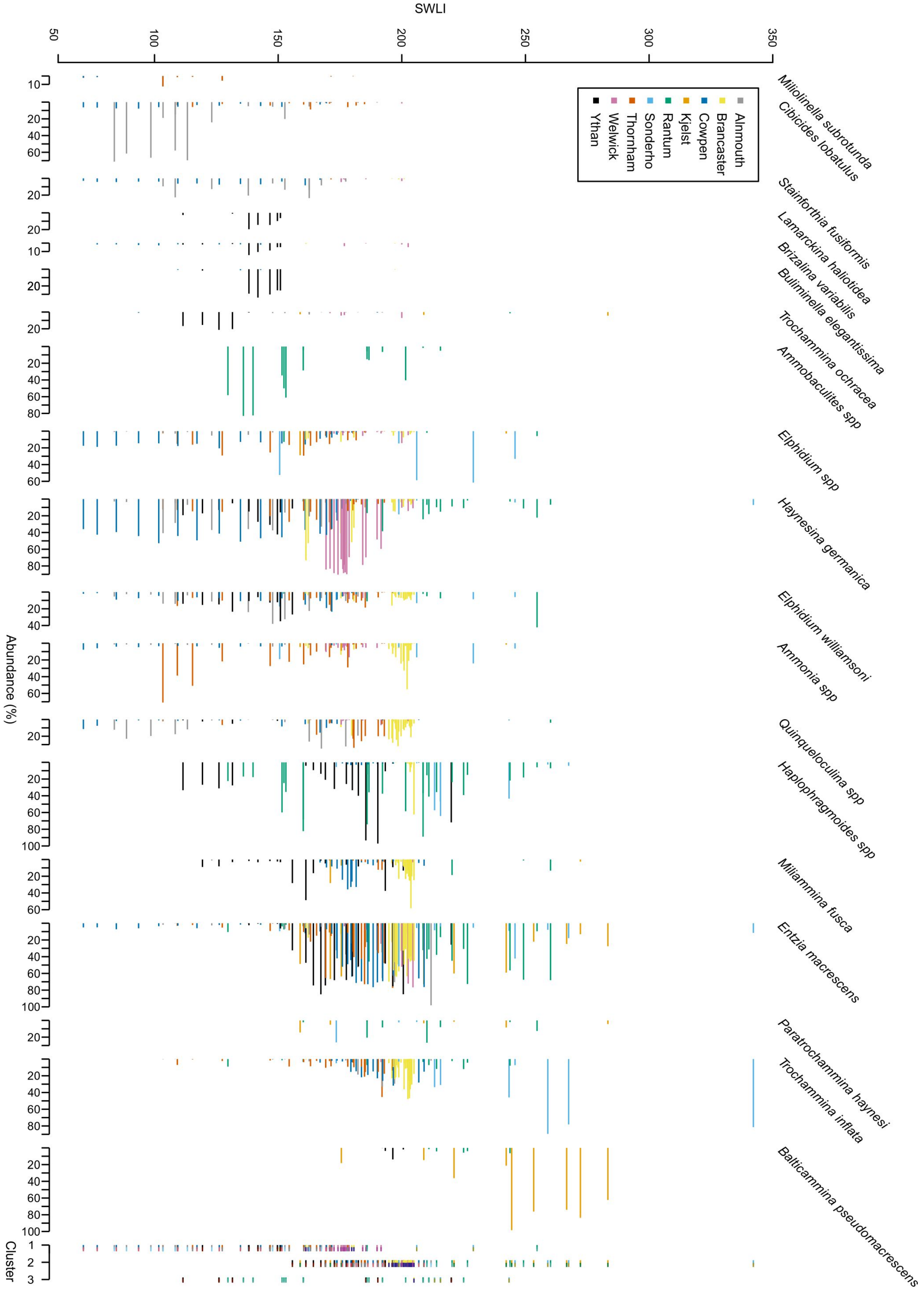

SWLI
50
100
150
200
250
300
350
Alnmouth
Brancaster
Cowpen
Kjelst
Rantum
Sonderho
Thornham
Welwick
Ythan
Miliolinella subrotunda
Cibicides lobatulus
Stainforthia fusiformis
Lamarckina haliotidea
Brizalina variabilis
Buliminella elegantissima
Trochammina ochracea
Ammobaculites spp
Elphidium spp
Haynesina germanica
Elphidium williamsoni
Ammonia spp
Quinqueloculina spp
Haplophragmoides spp
Miliammina fusca
Entzia macrescens
Paratrochammina haynesi
Trochammina inflata
Balticammina pseudomacrescens
Abundance (%)
Cluster
1
2
3

Figure 3: **The combined North Sea foraminifera data of 218 samples from 9 sites ordered by SWLI value as shown in the first column**. Taxa are included that have a maximum abundance > 10 % and occur in more than 10 % of the samples. The bar colour denotes the site corresponding to the legend. The final column shows the groups according to PAM clustering and marked by clusters 1-3.

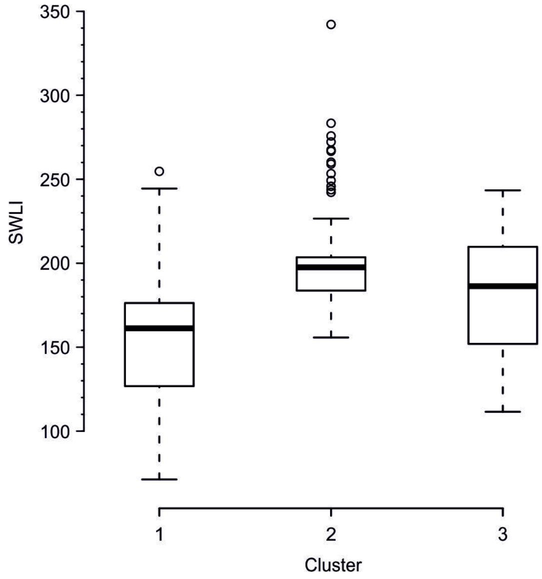


Figure 4: **Boxplot of the clusters following PAM clustering of the *North Sea* training set.** The boxplots represent the 2.5, 16, 50, 84 and 97.5 percentiles for each cluster. Outliers are represented by the circles.

3.2. *Comparison of training sets and transfer functions*

3.2.1. *Correspondence and cluster analysis*

In order to assess which of the regional, subregional and local training sets is most appropriate for reconstructing sea-level from the core samples shown in Fig. 6 we began by plotting DCA, with the core samples passively projected, and PAM, with modern and fossil samples combined. The core lithology and fossil foraminifera assemblages are summarised in Fig. 6 and show salt marsh clays that are dominated by *Entzia macrescens* and *Milliammina fusca* abruptly transitioning to estuarine silts comprising mainly calcareous taxa. DCA results for all of the training sets shown in Fig. 7 demonstrate an alignment of modern samples with axis 1 and a correlation with elevation. The samples are broadly clustered according to elevation, with a separation between salt-marsh and mudflat samples. There is also a clear distinction between core samples from within the apparent salt-marsh and mudflat zones in all of the plots. The fossil samples appear to show a wider dispersal and closer similarity to modern samples with increasing training set size, whereby the *North Sea* and *West* appear to perform well. Although DCA shows that many Ythan samples are occasionally distal from other sites and/or axis 1 in the regional training sets, it is important to observe that they appear to provide closer matches to many fossil samples. This is perhaps not surprising as they come from the same site and that conditions and taxa response show similarities between modern and 8000 years ago. There are clearly still exceptions that may be due to different taxa niches and emphasise the importance of developing regional training sets. The DCA analysis is supported by plotting of the PAM results (Supplementary info Fig. 2) where clusters of modern samples become more clearly grouped by elevation as the training sets reduce in size, although there are overlaps in SWLI values in all. The fossil samples are dispersed between all clusters of each training set. However, the fossil samples become more closely bunched to each other, a measure of the similarity between samples, in the *Northwest* and *Ythan* versions and appear most widespread in the *West*. The correspondence and cluster analysis suggest that the *North Sea* and *West* training sets are best suited for reconstructing sea level based on the fossil foraminifera in the Ythan core.

3.2.2. *Modern Analogue Technique*

An assessment of the five closest modern analogues identified by MAT shows that almost all of the closest analogues for every core sample are from sites on the west of the North Sea (Supplementary info Table 2). Only Kjelst from the east provides any of the five closest analogues. The Ythan provides 51 % of the analogues showing that inclusion of the local data is imperative, particularly for the fossil mudflat samples. All sites from the west, with the exception of Thornham, contribute analogues within the closest two, suggesting that in agreement with DCA and cluster analysis, inclusion of the sites in the *West* training set is required as a minimum to provide acceptable modern analogues to reconstruct sea level from core A7.5.

3.2.3. *Transfer Function performance*

We applied transfer functions using WA with classical deshrinking, WAPLS using the different training sets and LW. Generally performance statistics, shown in Table 2, improve when transfer functions use more

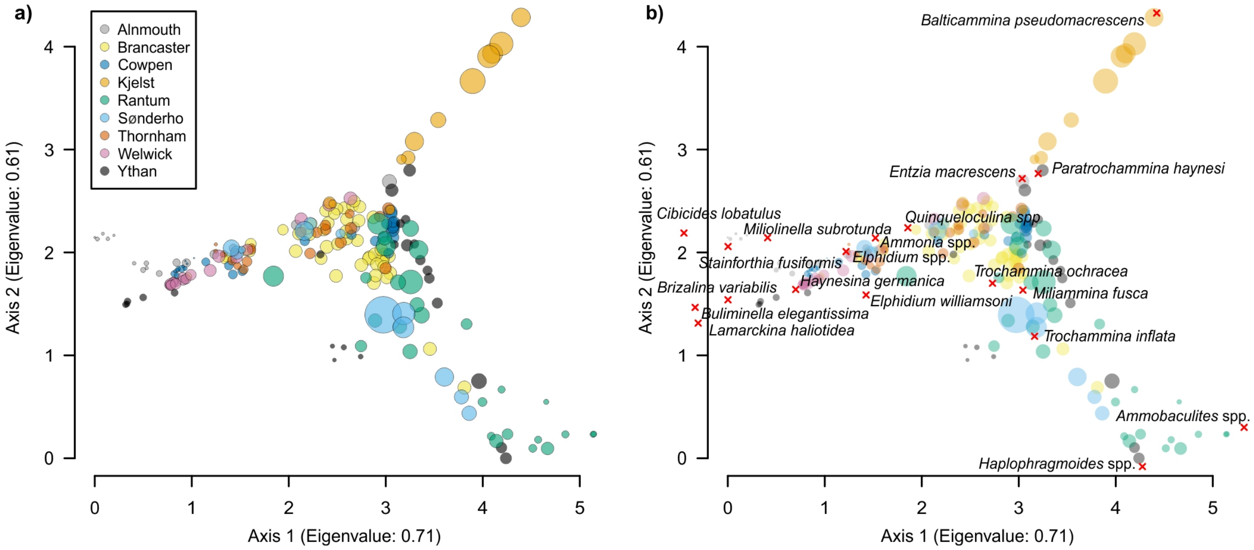


Figure 5: **Detrended canonical analysis (DCA) of the North Sea training set**. (a) Samples are plotted against elevation and coloured according to site and sized according to SWLI value. (b) Species are added and labelled if the maximum abundance > 10 %.

localised training sets in common with other sea-level studies (Horton and Edwards, 2006; Woodroffe and Long, 2010; Barlow *et al.*, 2013). The exceptions are the *East* transfer functions that perform particularly poorly, and the *Northwest* that performs worse than the *West*. The taxa optima and tolerance plots (Fig. 8) demonstrate that the taxa optima within the eastern sites are highly variable and hence the poor performance of the *East*, which in turn negatively impacts the performance of the *North Sea* transfer functions. Whilst this provides a wider range of environmental response it comes at the expense of less precision. All of the WA transfer functions have a higher RMSEP and average bias than the corresponding WAPLS models, although they do tend to have a lower maximum bias, and variable R^2 values. This suggests that classical deshrinking performs better at the extreme ends of the gradient, while inverse deshrinking (used in WAPLS models) provides higher accuracy for those in the middle as indicated by Birks (1995) and Juggins and Birks (2012). Comparison of the predicted SWLI against observed SWLI (Fig. 9) shows that although better fitting for the most extreme sample may be better in WA transfer functions, WAPLS appears to be perform better across the full elevation gradient.

Table 2: Performance statistics for the final transfer functions from each training set. The values given are the cross-validated statistics using leave-one-site-out (LOSO) or bootstrapping (boot). The bootstrapped statistics are converted to metres based on the Ythan tidal regime for core A7.5. Significance is based on 999 trials and the * denotes significant results ($p < 0.05$).

	$\text{SWLI}_{\text{loso}}$	$\text{SWLI}_{\text{boot}}$				Ythan A7.5 $(\text{m})_{\text{boot}}$		Significance
Model	**RMSEP**	**RMSEP**	R^2	**Av. Bias**	**Max. Bias**	**RMSEP**	**Max. bias**	p
North Sea-WA	38.59	33.88	0.55	0.09	108.93	0.56	1.80	0.12
North Sea-WAPLS-c1	32.57	26.53	0.58	0.41	112.18	0.44	1.85	0.07
North Sea-LW-WAPLS	34.27	24.86	0.60	1.81	118.10	0.41	1.95	0.10
West-WA	33.51	25.63	0.61	-0.04	33.33	0.42	0.55	0.10
West-WAPLS-c2	34.02	19.86	0.66	0.38	45.87	0.33	0.76	0.02*
West-LW-WAPLS-c2	NA	16.98	0.73	0.51	37.23	0.28	0.61	0.01*
Northwest-WA	33.51	23.34	0.70	-0.16	30.07	0.39	0.50	0.09
Northwest-WAPLS-c2	26.8	26.77	0.52	1.52	41.31	0.44	0.68	0.06
Ythan-WA	NA	22.55	0.58	1.64	41.83	0.37	0.69	0.17
Ythan-WAPLS-c2	NA	17.4	0.70	1.12	14.25	0.29	0.24	0.05
Southwest-WA	28.32	16.99	0.56	-0.25	39.04	0.28	0.64	NA
Southwest-WAPLS-c2	25.57	12.46	0.67	-0.18	36.87	0.21	0.61	NA
East-WA	45.94	37.62	0.54	1.63	71.75	0.62	1.18	NA
East-WAPLS-c1	44.54	40.21	0.29	7.32	66.23	0.66	1.09	NA

To assess the independence of samples, either because of training set samples being clustered by sites or the possibility of spatial auto-correlation, we compared cross-validation using LOSO with bootstrapping. Spatial auto-correlation, whereby nearby samples tend to resemble one another more than randomly selected locations, can occur when samples are collected along transects as is common procedure in sea-level research, as opposed to the more ecologically sound methods such as random sampling (Telford and Birks, 2005). If spatial auto-correlation is present it can result in overly-optimistic RMSEP values and hence misguided model choice (Payne *et al.*, 2012; Kemp and Telford, 2015). LOSO cross-validation removes all samples from one site and predicts the SWLI for them using the remaining sites and repeats the process for each site (Payne *et al.*, 2012). In all regions $\text{RMSEP}_{\text{loso}}$ was greater than

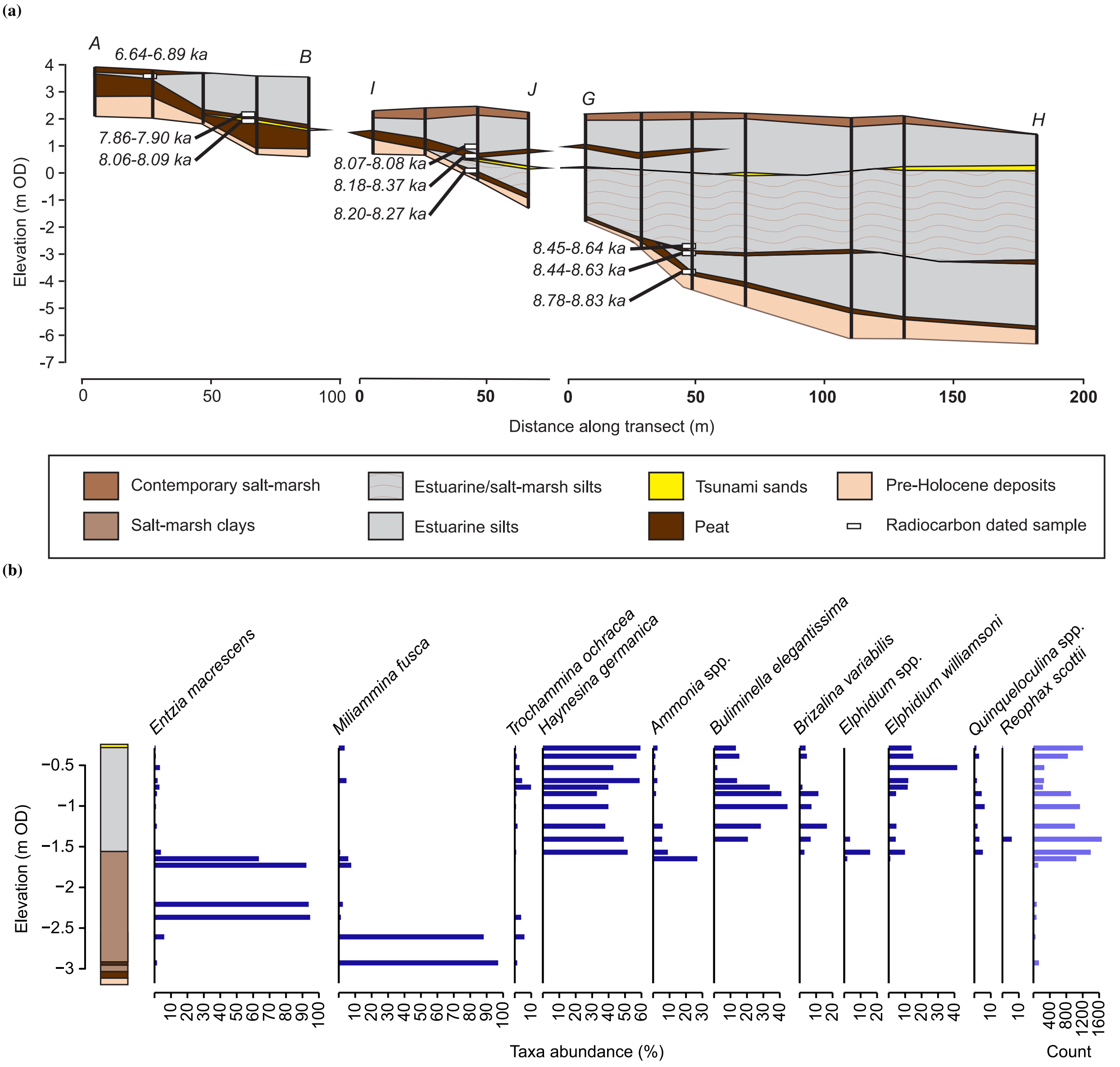


Figure 6: **Summary of Ythan palaeo data**: a) The lithology and calibrated radiocarbon ages of the transects sampled in the work by Smith *et al.* (1999) that correspond to Fig. 1. b) Lithology and fossil foraminifera assemblages of core A7.5 from the Ythan Estuary. Foraminifera abundances are given for taxa that have a maximum abundance > 5 %.

$\text{RMSEP}_{\text{boot}}$ (see Table 2). However, the differences are relatively small suggesting only a limited degree of possible spatial auto-correlation may be present, or alternatively the clustering of samples between sites reflects variable local environmental conditions. The inclusion of many sites also minimises the effect of spatial auto-correlation (Legendre and Fortin, 1989; Telford and Birks, 2005) whilst simultaneously making the larger regional transfer functions more robust against environmental changes occurring at a particular site (Barnett *et al.*, 2016). Fig. 10 shows that the sites perform relatively consistently suggesting that one particular site is not being unduly effected and that all sites likely display some environmental variability.

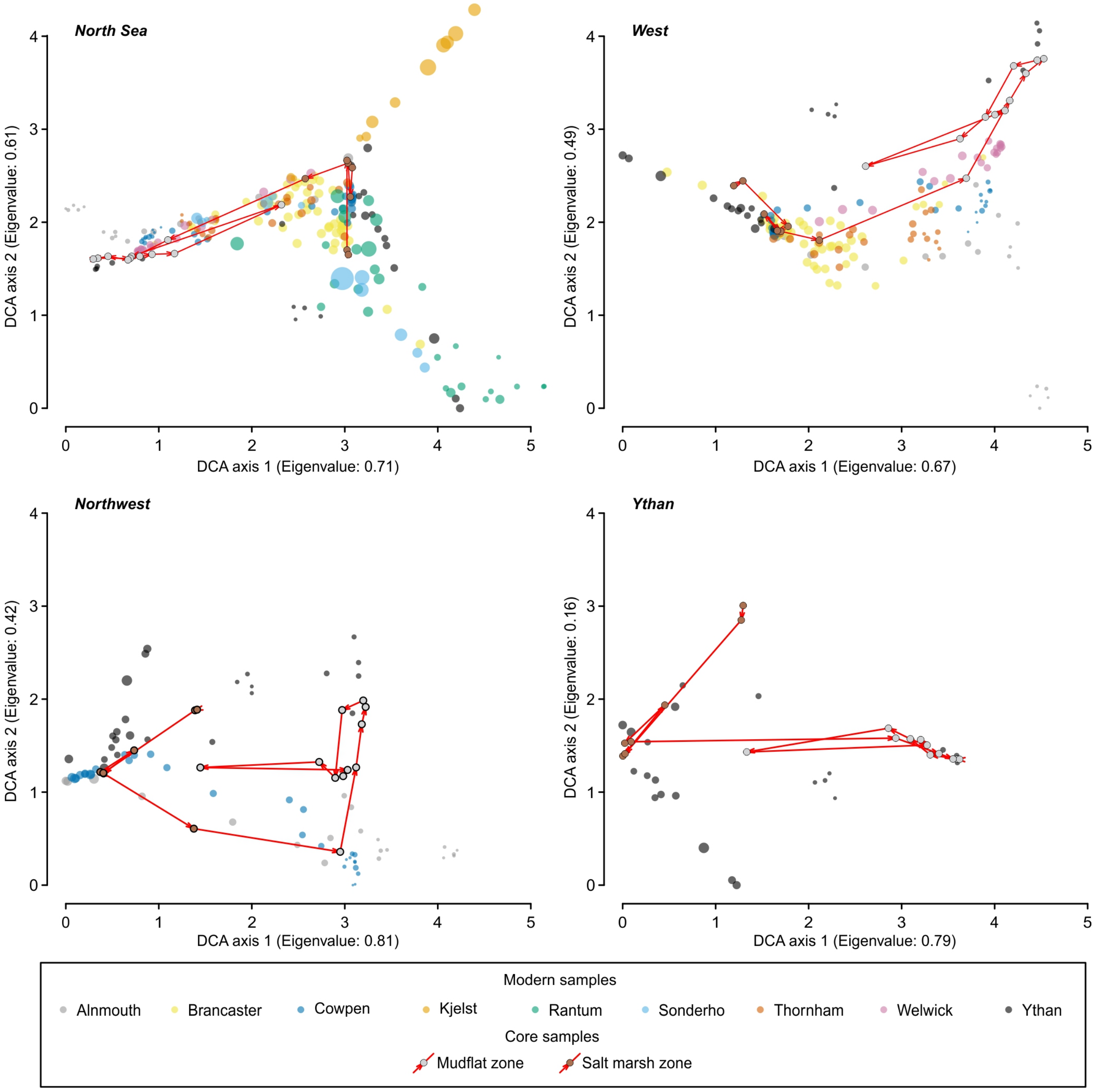
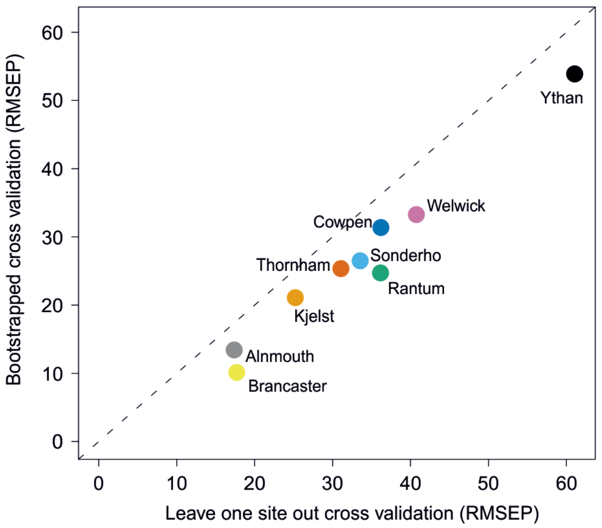


Figure 7: **Correspondence analysis of training sets and fossil data** showing the four sub-regional training sets and Ythan core samples using detrended correspondence analysis (DCA). The modern samples are coloured by site and the size is scaled by SWLI. The core samples are coloured by the lithology of the core and reflect the stratigraphic succession shown by the red lines with arrows indicating the stratigraphic order.



To further assess transfer function performance and understand the effect of deshrinking methods and adding extra complexity in WAPLS with components > 2, we compared scatterplots of observed versus predicted SWLI and the updated taxa optima (Wright *et al.*, 2011). Fig. 8 shows that when inverse deshrinking (WAPLS component 1) as opposed to classical is applied low elevation taxa are under predicted and high

Figure 10: **Comparison of root mean square error prediction**

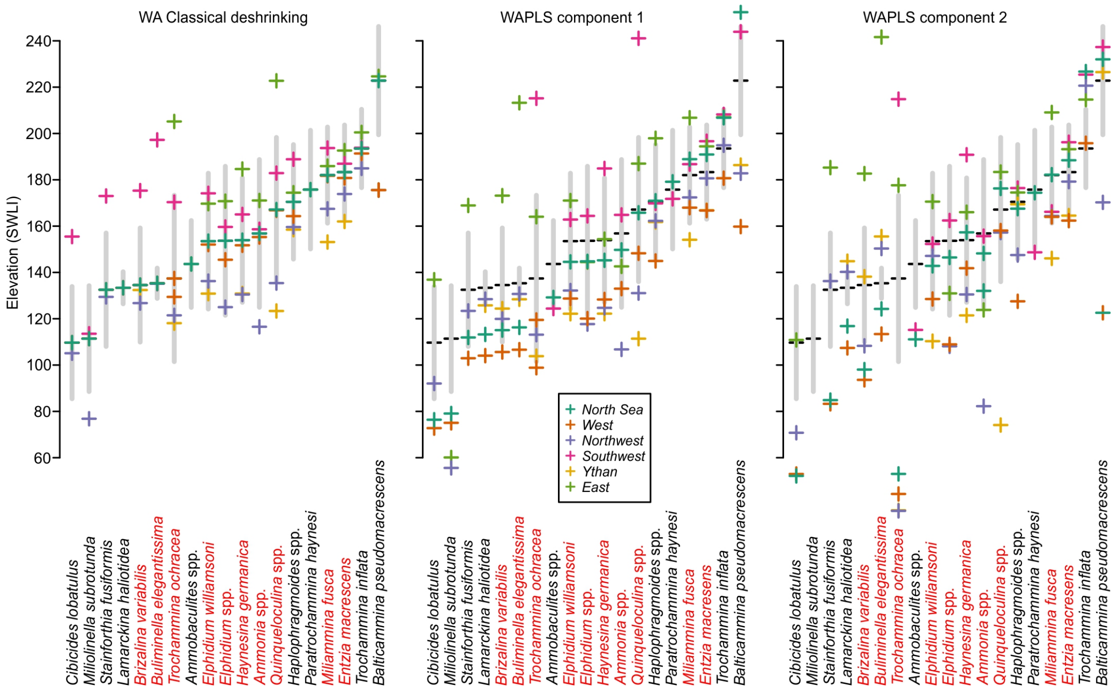


Figure 8: **Comparison of the progressive taxa optima updates** for the different training sets from WA with classical and inverse (equivalent to WAPLS component 1) to more complex WAPLS transfer functions with multiple components. Taxa tolerances are shown for WA by the grey bar. Taxa are included that have a maximum abundance > 10 % and occur in more than 10 % of the samples. Taxa labels coloured red are taxa that are also found in the Ythan Estuary core (A7.5)

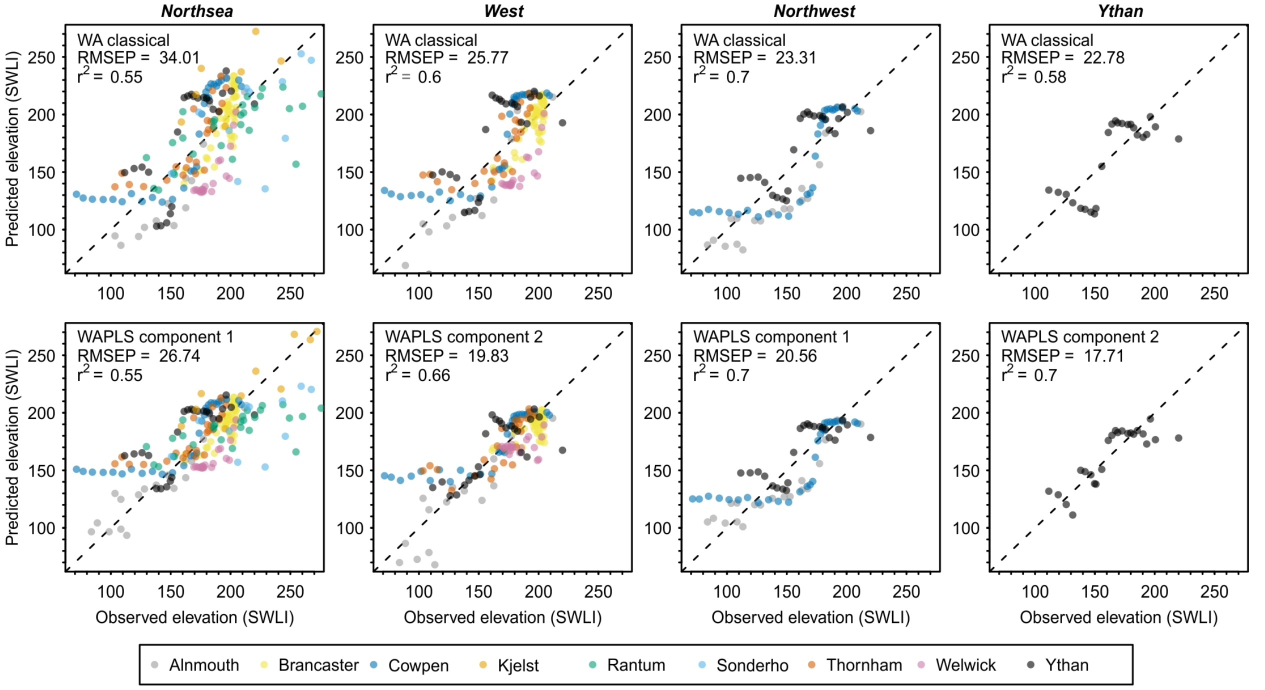


Figure 9: **Scatterplots of the observed SWLI against the predicted SWLI** for WA classical deshrinking and WAPLS components 1 or 2 depending on which was selected (see text). The samples are coloured according to site as shown in the legend.

salt-marsh taxa are over predicted, although this is less apparent in Fig. 9 where WAPLS consistently provides more accurate predictions. This is likely because the most abundant taxa tend to have optima towards the centre of the elevation gradient in the training set and are therefore more suitable for inverse deshrinking. Although there is noticeable variability between training sets in Fig. 8, the *North Sea* and *West* taxa optima are very similar and both appear relatively stable to the effect of adding extra components in WAPLS transfer functions. Adding a second component to the *West* WAPLS transfer function can be seen to produce accurate sample predictions shown in Fig. 9 by updating many of the taxa optima (Fig. 8). However, of the taxa that receive major updates only *Trochammina ochracea* is found in the core and in low abundances and therefore the reconstructions will not be distorted.

Despite the good performance, structure remains in the dataset (see Fig. 9) such that some of the samples found across the elevation gradient of the mudflat (particularly those of Cowpen) have very similar predictions because the samples have very similar taxa abundances. This demonstrates the difficulty in the application of mudflat samples. One option would be to remove these samples but they represent the full environmental gradient of the samples and so we favour retaining all samples but treating predictions with SWLI values < 140 with extra caution.

Based on transfer function performance statistics alone, the inclination would be to choose the more localised training sets for sea-level reconstructions for core A7.5. However, incorporating more sites provides more robustness against the possible effects of clustered samples and add additional variability in a taxa-elevation relationship that is likely important for providing analogues for early Holocene foraminifera. Taxa optima are also more stable with a greater number of sites and suggest these may also be valid.

3.2.4. Relative sea-level reconstructions

We applied the preferred WA and WAPLS transfer functions to reconstruct the palaeo marsh elevation relative to MTL at the Ythan Estuary using the fossil assemblages found in core A7.5 (see Fig. 6). All of the reconstructions show a similar pattern of indicative meaning change, with periods of rapid decreases (i.e., relative sea-level rise) at core depths of around -3 m and -1.6 m set against a trend of more gradual decrease (Fig. 11). The reconstructions produced notable differences depending on the training set and whether WA or WAPLS transfer functions were used. The reconstructions using the *North Sea* and *West* training sets display very similar patterns to each other, although the precision, as quantified by the sample specific error, increases from *North Sea* to *West*. The *West* model occasionally predicts slightly lower SWLI which is due to the exclusion of the eastern sites that generally have higher taxa optima. The ranges of the *Northwest* and *Ythan* reconstructions show divergence from the *North Sea* and *West* at differing periods and magnitudes across WA and WAPLS transfer functions. The number of good and close modern analogues as assessed by the MinDC become greater with increased number of sites in the training sets, for example increasing from nine for the *Ythan* to 24 for the *West* and *North Sea* (Fig. 11). The extra samples in the *North Sea* and *West* training sets evidently provide better analogues that also drive differing SWLI predictions from the others and hence suggest the *Northwest* and *Ythan* produce inaccurate reconstructions.

The ranges of the WA and WAPLS reconstructions show overlap in all cases, although there is a probable difference in the magnitude of both the first and second decrease in indicative meaning. The WA based reconstructions tend to predict lower surface elevations for mudflat samples and higher surface elevations for salt-marsh samples than the WAPLS models. This results in the first period of change being apparently larger in the WAPLS reconstructions compared to the WA, whilst the second is greater in the WA reconstructions. The fossil samples that bound the periods of rapid change appear indicative of mudflats and middle-upper salt marsh with taxa, from towards the gradient ends; predictions of assemblages in these zones are therefore important. Fig. 8 shows that WAPLS appears better at removing the edge-effects than WA even when one component is used (Mohler, 1983) and should thus provide more accurate reconstructions of the fossil mudflat samples. Small changes in species abundances towards the ends of the gradient can substantially alter predictions (Gehrels, 2000). At the upper end this is generally beneficial because of the high precision of high salt-marsh species (Scott and Medioli, 1978; Gehrels, 2000), but can be problematic at lower ends where taxa tolerance can be less precise due to wide apparent niches and taxa sometimes being found at sub-tidal elevations (Berkeley *et al.*, 2007). Lower elevation samples are sometimes removed because of the uncertainty in taxa response to elevation on the mudflats (Edwards and Wright, 2015) especially those that show nonlinearity with elevation (Hamilton and Shennan, 2005). However, despite the difficulties of trying to reconstruct mudflat environments, including the lower elevation samples is required to provide appropriate modern analogues and are therefore justified here. WAPLS transfer functions are consequently likely to provide better predictions of both the higher and lower elevation samples.

3.2.5. Statistical significance of reconstructions

We tested whether the reconstructions trained on the different training sets explained more variance in the core samples than the majority of reconstructions trained on randomly derived sets of data (n = 999) as proposed by Telford and Birks (2011). The results show that only the *West-WAPLS-c2* produced a statistically significant reconstruction ($p < 0.05$) (Table 2). The *West-WA* and *North Sea-WA* and *North Sea-WAPLS-c1* reconstructions perform relatively well and the similarity of reconstructions suggests they these are still valid. Telford and Birks (2011) argue that palaeoclimatology should not be exempt from interpreting common constraints of insignificant results which can be equally applied to sea-level studies, although Kemp *et al.* (2013) show that some salt-marsh cores do not provide sufficient downcore variability to outperform random data. In this case the fossil assemblages do provide notable variety and thus the significant result is somewhat relevant and provides support for the *West-WAPLS-c2* transfer function.

3.3. Training set and transfer function selection

The challenge of which combination of training set and transfer function model to apply is not a straightforward decision and may lead to different outcomes as described above in the core reconstruction (seen in Fig. 11) and by Barlow *et al.* (2013). Including all samples from a wider geographic region and a longer elevation gradient captures wider taxa tolerances and provides more modern analogues. One could therefore argue for including as wide a range as possible. However, the challenge is minimising the range of the reconstructed indicative meanings as much as possible by keeping variability of taxa tolerances to a minimum, whilst still providing an analogue for the past environment that can produce suitably precise and accurate quantitative reconstructions. We have built on suggested methods for developing transfer functions (Wright *et al.*, 2011; Watcham *et al.*, 2011; Barlow *et al.*, 2013; Kemp and Telford, 2015) with particular consideration for early Holocene sea-level reconstructions to assess this.

Assessment of transfer function performance statistics alone could lead to the more localised models being preferred due to the lower RMSEP and maximum bias. However, MAT, DCA and PAM silhouette plots show that fossil samples often lack appropriate analogues and similarity with the modern environment in the *Ythan* and *Northwest* training sets. The lack of analogues and poor clustering of the Ythan and fossil samples suggest that the likelihood of the foraminifera-elevation relationship at a single site remaining unchanged over a period of thousands of years is unlikely. Even relatively local regional microfossil training sets such as the *Northwest* here and the West coast of Scotland diatom set (Lawrence *et al.*, 2016) do not necessarily provide suitable modern microfossil analogues and therefore including sites that encompass different environmental conditions in a larger training set appears necessary. A shift in the reconstructed ranges of the reconstructions when moving to the more localised training sets from the *West* and *North Sea* training sets is evident further suggesting the reconstructions from the more localised training may lose accuracy. On the basis of missing analogues and lack of similarity we rule out using the *Ythan* and *Northwest* training sets for reconstructing the Ythan core.

DCA and cluster analysis suggests that the *North Sea* and *West* models both perform well and indeed both have no 'poor' analogues according to MAT. The *North Sea* training set has more spread in the modern foraminifera tolerances due to the eastern sites. This may simply be adding unnecessary noise, or on the other hand be providing additional analogues and thus better accuracy. However, the foraminifera in the eastern sites appear to be overly affected by the wind-driven water levels that reduces their utility for reconstructing sea level reflected in their higher optima. Using tide loggers as opposed to relying on modelled tidal data may help overcome this issue that will be particularly important at micro-tidal sites. Both *West* and *North Sea* training sets do still produce very similar reconstructions for either WA and WAPLS transfer functions suggesting that the small increase in precision in the reconstructions and better predictive ability from the *West* training set does not come with a loss of consistency. It therefore appears that the modern environment captured by the sites in the *West* training set provides a suitable analogue for the conditions found at the Ythan Estuary around 8 ka without the noise introduced by eastern sites. We therefore select the *West* training set for our transfer functions.

The WA and WAPLS transfer functions result in similar reconstructions, with overlapping ranges. The main difference is the prediction of mudflat samples which cause periods of smaller magnitude changes in the WAPLS-derived indicative meanings. WAPLS arguably provides better accuracy for this zone by removing edge effects (Mohler, 1983; Juggins and Birks, 2012) and indeed towards the centre. Furthermore, the

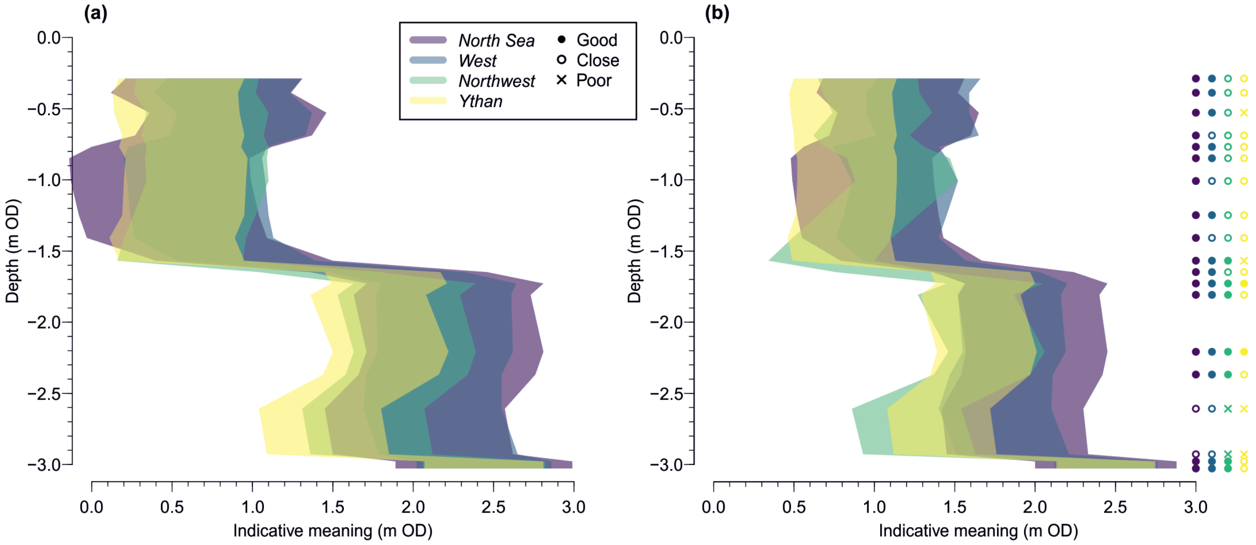


Figure 11: **Reconstructed indicative meaning of core A7.5 samples** from the different training sets using a) WA and b) WAPLS based transfer functions. 'Good', 'Close' and 'Poor' are measures of the closeness of modern analogues and are taken from the modern analogue technique (MAT) with MinDC values of < 5, 5 – 20 and > 20$^{\text{th}}$ percentiles respectively.

West-WAPLS-c2 is the only transfer function to produce a statistically significant reconstruction. Analysis of the taxa optima updates (Fig. 8) when an extra component is added in the *West-WAPLS-c2* transfer function show a number of calcareous taxa from some sites are shifted to well above HAT and so these more complex transfer functions should be treated with caution (Wright *et al.*, 2011). However, the taxa found in the fossil assemblages are not shifted so dramatically and therefore, in this instance, we accept the updates as appropriate for reconstructing the core. Whilst WA and WAPLS both produce similar reconstructions and the decision between the two does not therefore produce significantly different results we prefer the *West-WAPLS-c2* transfer function owing to the better fitting of predicted SWLI, the reduction in edge-effects compared to WA, the reconstruction being statistically significant and without producing dramatic taxa optima updates of the important foraminifera. The *West-WAPLS-c2* transfer function produces a mean sample specific error of 38 cm for the Ythan fossil samples which can be considered good, particularly in light of the approximately 50 cm precision obtained in the Lawrence *et al.* (2016) sea-level reconstruction.

A LW transfer function may circumvent the problem of including a wide range of sites and the associated noise. However, simply running a LW transfer function using the *North Sea* training set (*North Sea-LW-WAPLS*) produces inconsistent and unrealistic reconstructions despite the precision being improved. This appears to be because many samples from the eastern sites with SWLI's > 250 are included and have a greater influence than in the full training set. On the other hand, having established the most appropriate sites to include in the modern training set (i.e the *West* training set) and then running a LW transfer function produces consistent predictions alongside an overall improvement in the precision. Notably, the precision of the tidal-flat samples in the core is unaltered while the predictions of the salt-marsh samples is improved by around ± 7 cm. This is consistent with the better predictive ability of salt-marsh foraminifera and demonstrates the benefit of using a LW transfer function, notwithstanding the importance of establishing a suitable training set by way of a thorough assessment in advance.

4. Conclusions

We have produced a North Sea training set of modern foraminifera based on eight previously published sites and one new site (Ythan Estuary, Scotland). The foraminifera cover an elevation gradient from the highest elevation at which foraminifera occur to mean tide level. The foraminifera display a relationship with elevation relative to sea level. Foraminifera in marshes in the east (Denmark and Germany) were shown to be very variable, often displaying exceptionally broad elevation ranges that may be due to additional environmental factors such as wind build up that is enhanced by the relatively small tidal range.

We assessed the effectiveness of a modern regional training set for reconstructing early Holocene sea level at a coastal site in the western North Sea (Ythan Estuary, Scotland) by dividing the data into different sub-regional training sets and by comparing the results of parallel analyses. We applied a step-wise approach that considered understanding the core lithology and the fossil samples and understanding of the core alongside the modern samples in each of the training sets. In summary we used the following approach, which we also recommend for choosing the most effective transfer function and training set in similar studies:

1. We qualitatively assessed how appropriate for reconstructing sea level each training set was by applying detrended correspondence analysis (DCA) and partitioning around medoids (PAM) clustering. The results highlighted that clustering of modern samples is apparent and is occasionally driven by certain taxa, however these are generally rare or absent in the fossil record and therefore of lesser importance in this example. The modern samples from smaller, more localised, training sets show a more clearly defined relationship with elevation but often lack similarity with fossil samples. Other methods of correspondence and cluster analysis are available but some form should be included alongside fossil samples to enable a truer understanding of the suitability of different training sets in each context.
2. The modern analogue technique (MAT) was used as a statistical measure of the similarity between fossil and modern samples. More localised training sets produced fewer 'good' or 'close' modern analogues, while larger training sets produce no 'poor' analogues. This step is almost certainly necessary to validate that the training set is providing a modern environment analogous to that found at the reconstructed site for the period of interest and should be assessed in conjunction with more qualitative approaches such as in steps 1 and 4.
3. We ran transfer functions using WA with classical deshrinking and WAPLS components 1 or 2 with cross-validation by bootstrapping and by leave-one-site-out (LOSO) to reconstruct palaeo marsh surface elevation changes from the Ythan core. Precision improved in more localised regional models although at a loss of predictive ability. LOSO cross-validation showed that the precision of reconstructions are not unduly over-optimistic and that inter-site variability is present and likely capture different conditions.
4. We analysed the taxa updates when extra components were added in the WAPLS transfer functions. The taxa are occasionally altered beyond the extent of the sampled elevation range and caution is therefore necessary in choosing these more complex transfer functions. However, comparison with fossil samples showed that the relevant taxa are not fundamentally altered in our example. The updates complement MAT in showing the suitability of analogues. This test is a necessary step and should be carried out along with knowledge of fossil samples.
5. We compared all of the reconstructions to assess the accuracy. All reconstructions showed a similar pattern of surface elevation change, although differences were evident in the uncertainty between WA and WAPLS and when trained on different training sites. As opposed to reconstructions from the more localised training sets, the *West* models consistently plotted inside the range of the *North Sea* predictions suggesting that the *West* models retain the accuracy alongside an improvement in model performance. This procedure provides a good understanding of the consistency and hence the likelihood of producing accurate reconstructions.
6. We ran locally weighted transfer functions using the 50 closest analogues. The *West-LW-WAPLS* showed an improvement in performance because of improved precision of the salt-marsh sample predictions whilst still enabling the retention of the tidal-flat samples in the training set. This step is evidently worth exploring, particularly when a training set covering such a wide elevation gradient is used, but with the caveat that it should not be taken without an assessment in advance to fully understand the training set.
7. We tested the significance of the reconstructions. Our results showed that only the *West-WAPLS-c2* and *West-LW-WAPLS-c2* model significantly outperformed transfer functions run on randomly generated data. This test should be ideally included in studies at this stage, although bearing in mind that many cores used in sea-level studies may not have enough variability to outperform random data it may not always be relevant.

Combining data from multiple sites will almost always be necessary to produce early Holocene sea-level reconstructions that are most likely to be accurate. Thus an approach that utilises qualitative and quantitative techniques to assess which training set and transfer function is most suitable is also necessary. We have shown how our approach can provide an evidence-based decision that should help ensure a model is chosen that has good performance and produces predictions that are plausibly accurate. We chose the *West* WAPLS transfer function as the best performing across the full suite of analysis for reconstructing relative sea levels in the Ythan core. However, the decision may well be different for different cores and regions so we recommend that a similar procedure to ours should be followed even when using the same training set. Although we focus on foraminifera here, the techniques are equally applicable to other microfossils such as diatoms and testate amoebae. We advocate that similar step-wise approaches to ours are adopted when assessing model choice and that accuracy be prioritised over precision.

Acknowledgements

We are grateful to the editor Richard Jordan and two anonymous reviewers for their comments that greatly improved the early version of the paper. G.R was funded by a NERC studentship through the ACCE (Adapting to the Challenges of a Changing Environment) Doctoral Training Partnership (Grant No. 2210800). The modern Ythan samples were collected, prepared and counted as part of P.M's University of York BSc thesis: 'Surface distributions of salt-marsh foraminifera for the Ythan estuary, NE Scoltand: an unprecedented modern analogue for Holocene sea-level studies', 2018. We thank W McKay and the Scottish Natural Heritage for providing support and access to the field site. RE, YM and RG worked on and provided the previously published data. We thank the technicians at the Department of Environment and Geography, University of York, particularly Maria Gehrels for her assistance with laboratory work on the Ythan samples, and the countless other people who contributed to the collection and collation of the foraminifera data from the existing sites. We also acknowledge PALSEA, a working group of the International Union for Quaternary Sciences (INQUA) and Past Global Changes (PAGES), which in turn received support from the Swiss Academy of Sciences and the Chinese Academy of Sciences.

Data Availability

Supplementary data to this article can be found online at `https://doi.org/10.1016/j.marmicro.2021.102055`. The full North Sea modern foraminifera is also available on figshare at `https://figshare.com/authors/Graham_Rush/11546401` along with the R code to carry out the analysis. A vignette for running the code is provided at `https://rpubs.com/GrahamRush#`

References

Allen, J. R. L. (2000), 'Morphodynamics of Holocene salt marshes: A review sketch from the Atlantic and Southern North Sea coasts of Europe', *Quaternary Science Reviews* **19**(12), 1155–1231.

Avnaim-Katav, S., Gehrels, W. R., Brown, L. N., Fard, E. and MacDonald, G. M. (2017), 'Distributions of salt-marsh foraminifera along the coast of SW California, USA: Implications for sea-level reconstructions', *Marine Micropaleontology* **131**, 25–43.

Barber, D. C., Dyke, A., Hillaire-Marcel, C., Jennings, A. E., Andrews, J. T., Kerwin, M. W., Bilodeau, G., McNeely, R., Southon, J., Morehead, M. D. and Gagnon, J. M. (1999), 'Forcing of the cold event of 8,200 years ago by catastrophic drainage of Laurentide lakes', *Nature* **400**(6742), 344–348.

Barlow, N. L., Shennan, I., Long, A. J., Gehrels, W. R., Saher, M. H., Woodroffe, S. A. and Hillier, C. (2013), 'Salt marshes as late Holocene tide gauges', *Global and Planetary Change* **106**, 90–110.

Barnett, R. L., Garneau, M. and Bernatchez, P. (2016), 'Salt-marsh sea-level indicators and transfer function development for the Magdalen Islands in the Gulf of St. Lawrence, Canada', *Marine Micropaleontology* **122**, 13–26.

Bartholdy, A. T., Bartholdy, J. and Kroon, A. (2010), 'Salt marsh stability and patterns of sedimentation across a backbarrier platform', *Marine Geology* **278**(1-4), 31–42.

Bartholdy, J., Christiansen, C. and Kunzendorf, H. (2004), 'Long term variations in backbarrier salt marsh deposition on the Skallingen peninsula - The Danish Wadden Sea', *Marine Geology* **203**(1-2), 1–21.

Berkeley, A., Perry, C. T., Smithers, S. G., Horton, B. P. and Taylor, K. G. (2007), 'A review of the ecological and taphonomic controls on foraminiferal assemblage development in intertidal environments', *Earth-Science Reviews* **83**(3-4), 205–230.

Birks, H. (1995), Chapter 6: Quantitative palaeoenvironmental reconstructions, *in* 'Statistical Modelling of Quaternary Science Data'.

Birks, H. J. (1998), 'Numerical tools in palaeolimnology - Progress, potentialities, and problems', *Journal of Paleolimnology* **20**(4), 307–332.

Birks, H. J. B. (2012), Overview of numerical methods in palaeolimnology, *in* H. B. Birks, A. F. Lotter, S. Juggins and P. Smol, J, eds, 'Tracking environmental change using lake sediments', Springer, pp. 19–92.

Birks, H. J., Line, J. M., Juggins, S., Stevenson, A. C. and Ter Braak, C. J. (1990), 'Diatoms and pH reconstruction', *Philosophical Transactions - Royal Society of London, B* **327**(1240), 263–278.

Bondevik, S., Stormo, S. K. and Skjerdal, G. (2012), 'Green mosses date the Storegga tsunami to the chilliest decades of the 8.2 ka cold event', *Quaternary Science Reviews* **45**, 1–6.

Callard, S. L., Gehrels, W. R., Morrison, B. V. and Grenfell, H. R. (2011), 'Suitability of salt-marsh foraminifera as proxy indicators of sea level in Tasmania', *Marine Micropaleontology* **79**, 121–131.

Culver, S. J. and Horton, B. P. (2005), 'Infaunal marsh foraminifera from the outer banks, North Carolina, U.S.A.', *The Journal of Foraminiferal Research* **35**(2), 148–170.

Dawson, A., Bondevik, S. and Teller, J. T. (2011), 'Relative timing of the Storegga submarine slide, methane release, and climate change during the 8.2 ka cold event', *Holocene* **21**(7), 1167–1171.

de Rijk, S. (1995), 'Salinity control on the distribution of salt marsh foraminifera (Great Marshes, Massachusetts)', *The Journal of Foraminiferal Research* **25**(2), 156–166.

Edwards, R. J. and Horton, B. P. (2000), 'Reconstructing relative sea-level change using UK salt-marsh foraminifera', *Marine Geology* **169**(1–2), 41–56.

Edwards, R. J. and Wright, A. (2015), Foraminifera, *in* I. Shennan, A. J. Long and B. P. Horton, eds, 'Handbook of Sea-level Research', John Wiley & Sons, Ltd, Chichester, chapter 13, pp. 191–217.

Egbert, G. and Erofeeva, L. (2010), 'The OSU TOPEX/Poseidon Global Inverse Solution TPXO 8 Atlas v. 1 Africa'.

Engelhart, S. E., Peltier, W. R. and Horton, B. P. (2011), 'Holocene relative sea-level changes and glacial isostatic adjustment of the US Atlantic coast', *Geology* **39**(8), 751–754.

Gehrels, W. R. (2000), 'Using foraminiferal transfer functions to produce high-resolution sea-level records from salt-marsh deposits, Maine, USA', *Holocene* **10**(3), 367–376.

Gehrels, W. R. (2002), Intertidal foraminifera as palaeoenvironmental indicators, *in* 'Quaternary environmental micropalaeontology', Arnold Publishers, pp. 91–114.

Gehrels, W. R., Kirby, J. R., Prokoph, A., Newnham, R. M., Achterberg, E. P., Evans, H., Black, S. and Scott, D. B. (2005), 'Onset of recent rapid sea-level rise in the western Atlantic Ocean', *Quaternary Science Reviews* **24**(18-19), 2083–2100.

Gehrels, W. R., Milne, G. A., Kirby, J. R., Patterson, R. T. and Belknap, D. F. (2004), 'Late Holocene sea-level changes and isostatic crustal movements in Atlantic Canada', *Quaternary International* **120**(1), 79–89.

Gehrels, W. R. and Newman, S. W. G. (2004), 'Salt-marsh foraminifera in Ho Bugt, western Denmark, and their use as sea-level indicators', *Geografisk Tidsskrift-Danish Journal of Geography* **104**(1), 97–106.

Gehrels, W. R., Roe, H. M. and Charman, D. J. (2001), 'Foraminifera, testate amoebae and diatoms as sea-level indicators in UK saltmarshes: A quantitative multiproxy approach', *Journal of Quaternary Science* **16**(3), 201–220.

Hamilton, S. and Shennan, I. (2005), 'Late Holocene relative sea-level changes and the earthquake deformation cycle around upper Cook Inlet, Alaska', *Quaternary Science Reviews* **24**(12-13), 1479–1498.

Hawkes, A. D., Horton, B. P., Nelson, A. R. and Hill, D. F. (2010), 'The application of intertidal foraminifera to reconstruct coastal subsidence during the giant Cascadia earthquake of AD 1700 in Oregon, USA', *Quaternary International* **221**(1-2), 116–140.

Hayward, B. W., Le Coze, F., Varchard, D., Gross, O. and Gross, O. (2020), 'World modern foraminifera database'.
URL: *https://doi.org/10.14284/305*

Hennig, C. (2019), *fpc: Flexible Procedures for Clustering*. R package version 2.2-3.

Hijma, M. P. and Cohen, K. M. (2010), 'Timing and magnitude of the sea-level jump preluding the 8200 yr event', *Geology* **38**(3), 275–278.

Hijma, M. P. and Cohen, K. M. (2019), 'Holocene sea-level database for the Rhine-Meuse Delta, The Netherlands: Implications for the pre-8.2 ka sea-level jump', *Quaternary Science Reviews* **214**, 68–86.

Hill, M. O. and Gauch, H. G. (1980), 'Detrended correspondence analysis: An improved ordination technique', *Vegetatio* **42**(1-3), 47–58.

Horton, B. P. and Edwards, R. J. (2006), 'Quantifying Holocene sea-level change using intertidal foraminifera: lessons from the British Isles', *Cushman Foundation Special Publication* **40**, 1–97.

Horton, B. P., Edwards, R. J. and Lloyd, J. M. (1999), 'A foraminiferal-based transfer function: Implications for sea-level studies', *Journal of Foraminiferal Research* **29**(2), 117–129.

Horton, B. P. and Murray, J. W. (2006), 'Patterns in cumulative increase in live and dead species from foraminiferal time series of Cowpen Marsh, Tees Estuary, UK: Implications for sea-level studies', *Marine Micropaleontology* **58**(4), 287–315.

Horton, B. P., Peltier, W. R., Culver, S. J., Drummond, R., Engelhart, S. E., Kemp, A. C., Mallinson, D., Thieler, E. R., Riggs, S. R., Ames, D. V. and Thomson, K. H. (2009), 'Holocene sea-level changes along the North Carolina Coastline and their implications for glacial isostatic adjustment models', *Quaternary Science Reviews* **28**(17-18), 1725–1736.

Juggins, S. (2017), *rioja: Analysis of Quaternary Science Data*. R package version 0.9-21.

Juggins, S. and Birks, H. J. B. (2012), Quantitative environmental reconstructions from biological data, *in* H. B. Birks, A. F. Lotter, S. Juggins and P. Smol, J, eds, 'Tracking Environmental Change Using Lake Sediments. Volume 5: Handling and Numerical Techniques', Springer, pp. 431–494.

Kaufmann, L. and Rousseeuw, P. J. (1990), *Finding groups in data: an introduction to cluster analysis*, John Wiley, New York.

Kemp, A. C., Horton, B. P. and Culver, S. J. (2009), 'Distribution of modern salt-marsh foraminifera in the Albemarle-Pamlico estuarine system of North Carolina, USA: Implications for sea-level research', *Marine Micropaleontology* **72**(3-4), 222–238.

Kemp, A. C., Horton, B. P., Vann, D. R., Engelhart, S. E., Grand Pre, C. A., Vane, C. H., Nikitina, D. and Anisfeld, S. C. (2012), 'Quantitative vertical zonation of salt-marsh foraminifera for reconstructing former sea level; an example from New Jersey, USA', *Quaternary Science Reviews* **54**, 26–39.

Kemp, A. C. and Telford, R. J. (2015), Transfer functions, *in* I. Shennan, A. Long and B. Horton, eds, 'Handbook of Sea-level Research', 1 edn, John Wiley & Sons, Ltd, pp. 470–499.

Kemp, A. C., Telford, R. J., Horton, B. P., Anisfeld, S. C. and Sommerfield, C. K. (2013), 'Reconstructing Holocene sea level using salt-marsh foraminifera and transfer functions: Lessons from New Jersey, USA', *Journal of Quaternary Science* **28**(6), 617–629.

Kemp, A. C., Wright, A. J. and Cahill, N. (2020), 'Enough is enough, or more is more? testing the influence of foraminiferal count size on reconstructions of paleo-marsh elevation', *Journal of Foraminiferal Research* **50**, 266–278.

Lawrence, T., Long, A. J., Gehrels, W. R., Jackson, L. P. and Smith, D. E. (2016), 'Relative sea-level data from southwest Scotland constrain meltwater-driven sea-level jumps prior to the 8.2 kyr BP event', *Quaternary Science Reviews* **151**, 292–308.

Legendre, P. and Fortin, M. J. (1989), 'Spatial pattern and ecological analysis', *Vegetatio* **80**(2), 107–138.

Leorri, E., Fatela, F., Cearreta, A., Moreno, J., Antunes, C. and Drago, T. (2011), 'Assessing the performance of a foraminifera-based transfer function to estimate sea-level changes in northern Portugal', *Quaternary Research* **75**(1), 278–287.

Li, Y. X., Törnqvist, T. E., Nevitt, J. M. and Kohl, B. (2012), 'Synchronizing a sea-level jump, final Lake Agassiz drainage, and abrupt cooling 8200 years ago', *Earth and Planetary Science Letters* **315**, 41–50.

Mills, H., Kirby, J., Holgate, S. and Plater, A. (2013), 'The distribution of contemporary saltmarsh foraminifera in a macrotidal estuary: an assessment of their viability for sea-level studies', *Journal of Ecosystem & Ecography* **3**, 1–16.

Mohler, C. L. (1983), 'Effect of sampling pattern on estimation of species distributions along gradients', *Vegetatio* (54), 97–102.

Müller-Navarra, K., Milker, Y. and Schmiedl, G. (2016), 'Natural and anthropogenic influence on the distribution of salt marsh foraminifera in the Bay of Tümlau, German North Sea', *The Journal of Foraminiferal Research* **46**(1), 61–74.

Müller-Navarra, K., Milker, Y. and Schmiedl, G. (2017), 'Applicability of transfer functions for relative sea-level reconstructions in the southern North Sea coastal region based on salt-marsh foraminifera', *Marine Micropaleontology* **135**, 15–31.

Murray, J. W. (2003), 'Patterns in the cumulative increase in species from foraminiferal time-series', *Marine Micropaleontology* **48**(1-2), 1–21.

Murray, J. W. and Bowser, S. S. (2000), 'Mortality, protoplasm decay rate, and reliability of staining techniques to recognize 'living' foraminifera: a review', *The Journal of Foraminiferal Research* **30**(1), 66–70.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E. and Wagner, H. (2019), *vegan: Community Ecology Package*. R package version 2.5-5.

Payne, R. J., Telford, R. J., Blackford, J. J., Blundell, A., Booth, R. K., Charman, D. J., Lamentowicz, , Lamentowicz, M., Mitchell, E. A., Potts, G., Swindles, G. T., Warner, B. G. and Woodland, W. (2012), 'Testing peatland testate amoeba transfer functions: Appropriate methods for clustered training-sets', *Holocene* **22**(7), 819–825.

Pedersen, J. B. T., Svinth, S. and Bartholdy, J. (2009), 'Holocene evolution of a drowned melt-water valley in the Danish Wadden Sea', *Quaternary Research* **72**(1), 68–79.

Rousseeuw, P. J. (1987), 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis', *Journal of computational and applied mathematics* **20**, 53–65.

Scott, D. B. and Hermelin, J. O. R. (1993), 'A device for precision splitting of micropaleontological samples in liquid suspension', *Journal of Paleontology* **67**(1), 151–154.

Scott, D. S. and Medioli, F. S. (1978), 'Vertical zonations of marsh foraminifera as accurate indicators of former sea-levels', *Nature* **272**(5653), 528–531.

Smith, D. E., Firth, C. R., Brooks, C. L., Robinson, M. and Collins, P. E. F. (1999), 'Relative sea-level rise during the Main Postglacial Transgression in NE Scotland, UK', *Transactions of the Royal Society of Edinburgh-Earth Sciences* **90**, 1–27.

Stapleton, C. and Pethick, J. (1996), *Coastal processes and management of Scottish estuaries-III: The Dee, Don and Ythan Estuaries*, Scottish Natural Heritage.

Telford, R. J. and Birks, H. (2011), 'A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages', *Quaternary Science Reviews* **30**(9-10), 1272–1278.

Telford, R. J. and Birks, H. J. (2005), 'The secret assumption of transfer functions: Problems with spatial autocorrelation in evaluating model performance', *Quaternary Science Reviews* **24**(20-21), 2173–2179.

ter Braak, C. J. F. and Barendregt, L. G. (1986), 'Weighted averaging of species indicator values: Its efficiency in environmental calibration', *Mathematical Biosciences* **78**(1), 57–72.

ter Braak, C. J. F. and Juggins, S. (1993), Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages, *in* 'Twelfth international diatom symposium', Springer, pp. 485–502.

ter Braak, C. J. F. and Smilauer, P. (2012), 'Canoco reference manual and user's guide: software for ordination, version 5.0'.

UK Hydrographic Office (2016), *Admiralty Tide Tables, 2016, Volume 1, United Kingdom and Ireland (Including European channel ports).*, United Kingdom Hydrographic Office.

Watcham, E. P., Bentley, M. J., Hodgson, D. A., Roberts, S. J., Fretwell, P. T., Lloyd, J. M., Larter, R. D., Whitehouse, P. L., Leng, M. J., Monien, P. and Moreton, S. G. (2011), 'A new Holocene relative sea level curve for the South Shetland Islands, Antarctica', *Quaternary Science Reviews* **30**(21-22), 3152–3170.

Watcham, E. P., Shennan, I. and Barlow, N. L. M. (2013), 'Scale considerations in using diatoms as indicators of sea-level change: lessons from Alaska', *Journal of Quaternary Science* **28**(2), 165–179.

Woodroffe, S. A. (2009), 'Recognising subtidal foraminiferal assemblages: implications for quantitative sea-level reconstructions using a foraminifera-based transfer function', *Journal of Quaternary Science* **24**(3), 215–223.

Woodroffe, S. A. and Long, A. J. (2010), 'Reconstructing recent relative sea-level changes in West Greenland: Local diatom-based transfer functions are superior to regional models', *Quaternary International* **221**(1-2), 91–103.

Wright, A. J., Edwards, R. J. and van de Plassche, O. (2011), 'Reassessing transfer-function performance in sea-level reconstruction based on benthic salt-marsh foraminifera from the Atlantic coast of NE North America', *Marine Micropaleontology* **81**(1-2), 43–62.

Zong, Y. and Horton, B. P. (1999), 'Diatom-based tidal-level transfer functions as an aid in reconstructing Quaternary history of sea-level movements in the UK', *Journal of Quaternary Science* **14**(2), 153–167.