

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

Graham Rush<sup>a,\*</sup>, Patrick McDarby<sup>a</sup>, Robin Edwards<sup>b</sup>, Yvonne Milker<sup>c</sup>, Ed Garrett<sup>a</sup>, Roland Gehrels<sup>a</sup>

<sup>a</sup>Department of Environment and Geography, University of York, YO10 5NG, UK

<sup>b</sup>School of Natural Sciences, Trinity College Dublin, Ireland

<sup>c</sup>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

---

## 1 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)

35 Many studies use foraminiferal assemblages in mod- 87  
36 ern samples from a proximal site, collectively known as 88  
37 a training set, to develop local transfer functions that are 89  
38 then applied to reconstruct sea level based on fossil as- 90  
39 semblages from a core (e.g. Gehrels *et al.*, 2004, 2005). 91  
40 This method relies on the assumption that the assem- 92  
41 blages in the modern environment are an accurate ana- 93  
42 logue for the past environment and fossil foraminifera 94  
43 preserved in the core. In the absence of appropriate 95  
44 analogues in local training sets, and therefore a likely 96  
45 non-conformity of environmental conditions, a regional 97  
46 training set developed from multiple sites may be more 98  
47 appropriate (Watcham *et al.*, 2013; Barlow *et al.*, 2013). 99  
48 A regional dataset is a collection of training sets from 100  
49 multiple sites along a stretch of coastline, although the 101  
50 size of the region is not defined (Barlow *et al.*, 2013). 102  
51 Regional foraminifera training sets have been devel- 103  
52 oped for regions such as the UK (Horton *et al.*, 1999), 104  
53 southwest Europe (Leorri *et al.*, 2011), Oregon, west- 105  
54 ern USA (Hawkes *et al.*, 2010) and the eastern coast 106  
55 of the USA (Wright *et al.*, 2011; Kemp *et al.*, 2012). 107  
56 The method has been used to produce continuous sea- 108  
57 level records for the late Holocene (e.g. Hawkes *et al.*, 109  
58 2010; Engelhart *et al.*, 2011) and the early Holocene 110  
59 using foraminifera (Horton *et al.*, 1999) and diatoms 111  
60 (Lawrence *et al.*, 2016). Barlow *et al.* (2013) discuss 112  
61 the relative benefit of local and regional approaches, but 113  
62 in general a regional training set that comprises sites 114  
63 where faunal assemblages are responding to tidal inun- 115  
64 dation in a similar manner increases natural variability 116  
65 and therefore the potential of providing an analogous 117  
66 environment. 118

67 For early and middle Holocene sea-level reconstruc-  
68 tions developed from intertidal microfossils, the likeli-  
69 hood that environmental conditions and/or the realised  
70 niches of taxa at a single site have remained constant  
71 may be small. For example, Edwards and Horton (2000)  
72 found that of 26 fossil assemblages from cored inter-  
73 tidal sediments in the south of England, 10 were unlike  
74 any other modern surface assemblage. A lack of mod-  
75 ern analogues requires the inclusion of additional sites  
76 in to a larger training set. Assemblages from multiple  
77 sites are more likely to capture a wider range of envi-  
78 ronmental conditions with a greater chance of provid-  
79 ing a suitable modern analogue (Legendre and Fortin,  
80 1989; Juggins and Birks, 2012), with the caveat that the  
81 precision of the prediction may be reduced. For exam-  
82 ple, Lawrence *et al.* (2016) applied a regional diatom  
83 transfer function to generate an early Holocene sea-level  
84 reconstruction in southwest Scotland based on samples  
85 from nine sites on the west coast of Scotland. Yet a  
86 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 recon-  
struction which was calculated at  $\pm 56$  cm (Lawrence  
*et al.*, 2016). These examples highlight that whilst gen-  
erally 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 differ-  
ent 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 (Eng-  
land and Scotland) and western (Denmark and Ger-  
many) 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 be-  
tween 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 (Sta-  
pleton and Pethick, 1996) and the intertidal area is ap-  
proximately 1.85 km<sup>2</sup> of which 0.13 km<sup>2</sup> 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 Hydro-  
graphic 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

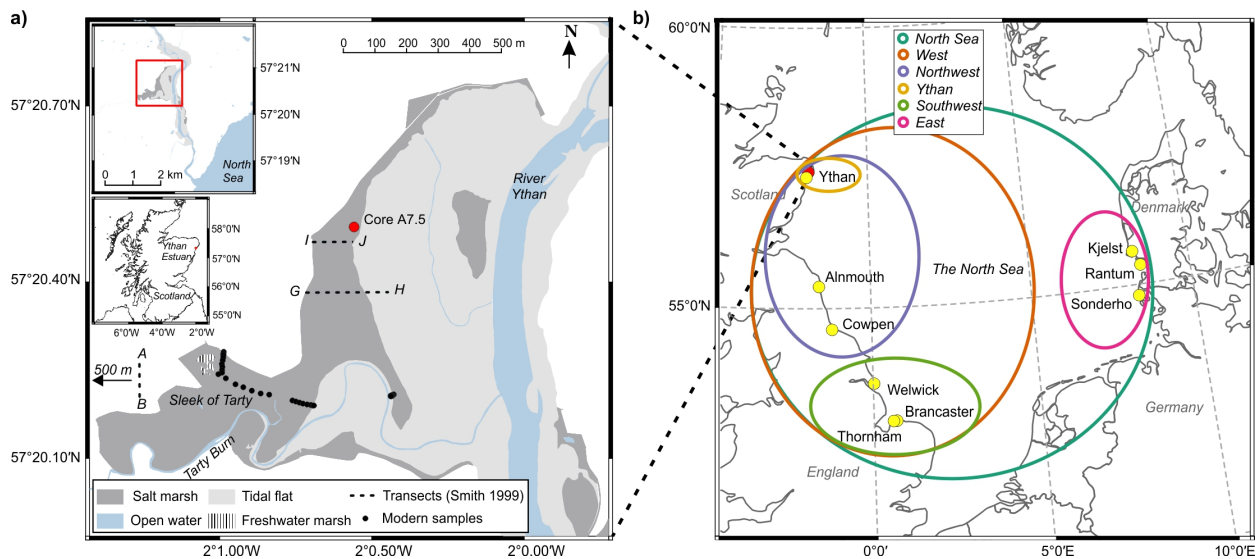


Figure 1: **Map of the Ythan Estuary and North Sea sites.** a) The Ythan Estuary showing the transects of modern samples, the transects from Smith *et al.* (1999) shown in Fig. 6a and the location of core A7.5 described in the text and Fig. 6b. b) Map of modern foraminifera sites (yellow markers) that 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.

136 marsh is characteristically undulating with many well  
 137 defined ponds and creeks. The lower salt marsh gener-  
 138 ally terminates at a small tidally cut cliff, although oc-  
 139 casionally grades into the mudflat.

140 We established a local benchmark using a Trimble R6  
 141 model 3 DGPS with a vertical uncertainty of 0.006 m.  
 142 The benchmark was tied to two local Ordnance Survey  
 143 benchmarks to provide elevation measurements relative  
 144 to the UK national vertical geodetic ordnance datum  
 145 (OD). Transects were designed to capture the elevation  
 146 gradient from the highest occurrence of foraminifera  
 147 (HoF) to mean tide level (MTL) and included a fresh-  
 148 water reed bed, high, middle and low salt marsh zones  
 149 and mudflat. Sampling was timed to coincide with a  
 150 large spring tide to maintain standardisation with other  
 151 studies and to allow sampling at lower elevations (e.g.  
 152 Avnaim-Katav *et al.*, 2017). Sampling sites were es-  
 153 tablished at c. 5 cm elevation intervals along two tran-  
 154 sects using a Trimble M1 DR2 total station. Standard-  
 155 ised surface samples of 10 cm<sup>2</sup> by 1 cm deep were col-  
 156 lected (e.g. Gehrels, 2000). After collection samples  
 157 were stored at 4 °C before processing.

158 Preparations of foraminifera samples followed stan-  
 159 dard laboratory procedures (Gehrels, 2002). The 63 –  
 160 500 µm fraction of a 5 cc subsample was wet sieved  
 161 and stored in a Rose Bengal and ethanol solution within  
 162 seven days of collection to enable identification of liv-  
 163 ing and dead foraminifera (Murray and Bowser, 2000).  
 164 Tests containing stained protoplasm within the last few

165 chambers were assumed to be living and only the  
 166 dead foraminiferal taxa were investigated to avoid sea-  
 167 sonal bias (Culver and Horton, 2005) and provide the  
 168 most representative analogue for material found in fos-  
 169 sil records (Horton *et al.*, 1999; Horton and Edwards,  
 170 2006; Berkeley *et al.*, 2007; Kemp *et al.*, 2009). Sam-  
 171 ples were split using a wet splitter (Scott and Hermelin,  
 172 1993) and picked and counted until a target of at least  
 173 200 individuals was met in even 1/8 splits. If this num-  
 174 ber was not achieved the full sample was counted. In-  
 175 dividuals were identified with reference to the taxon-  
 176 omy of de Rijk (1995); Wright *et al.* (2011); Edwards  
 177 and Wright (2015); Müller-Navarra *et al.* (2017); Hay-  
 178 ward *et al.* (2020). A total of 37 samples were counted  
 179 and individual taxa expressed as their percentage rela-  
 180 tive abundance of the sample.

## 181 2.2. Regional training set compilation

182 Existing modern salt-marsh foraminifera training  
 183 sets, collected from around the North Sea, include Al-  
 184 nmouth, Brancaster, Cowpen, Thornham and Welwick  
 185 (Horton and Edwards, 2006), Brancaster (Gehrels *et al.*,  
 186 2001), Kjelst (Gehrels and Newman, 2004), Rantum,  
 187 Sønderho (Müller-Navarra *et al.*, 2017) and Tümlau  
 188 (Müller-Navarra *et al.*, 2016) (see Fig. 1). All of the  
 189 data was included except Tümlau, because of human in-  
 190 terference that makes it unsuitable for sea-level recon-  
 191 structions (Müller-Navarra *et al.*, 2016), and two Al-  
 192 nmouth transects because of our doubts over the accuracy

193 of the elevation measurements. Tidal datums were taken 242  
 194 from the original publications and the tidal range was 243  
 195 modelled for each site using the TPX08-ATLAS global 244  
 196 model of ocean tides (Egbert and Erofeeva, 2010). The 245  
 197 data are summarised in Table 1 and throughout the text. 246

198 In order for the local training sets to be incorpo- 247  
 199 rated into a regional model, taxonomy was standard- 248  
 200 ised using the World Register of Marine Species (Hay- 249  
 201 ward *et al.*, 2020) (Supplementary info Table 1) and 250  
 202 the sample elevation was standardised using a standard- 251  
 203 ised water level index (SWLI) which accounts for dif-  
 204 ferences in tidal range between the sites (Zong and Hor-  
 205 ton, 1999; Gehrels, 2000). The lowest common phylo-  
 206 genic level was applied which resulted in some taxa be-  
 207 ing grouped at genus level. The only common environ-  
 208 mental variable across all sites was elevation relative to  
 209 a local datum. While inundation frequency is a more di-  
 210 rect ecological parameter (Gehrels, 2000; Gehrels *et al.*,  
 211 2001; Müller-Navarra *et al.*, 2017), elevation is a linear  
 212 approximation that is widely used and is applied 252  
 213 here in the absence of inundation information across 253  
 214 all sites. Generally MTL and mean high water spring 254  
 215 tide (MHWS) or mean higher high water (MHHW) have 255  
 216 been used for standardisation between sites with dif- 256  
 217 ferent tidal ranges in sea-level studies; however, trans- 257  
 218 fer functions are sensitive to the choice of tidal da- 258  
 219 tum (Woodroffe and Long, 2010; Wright *et al.*, 2011). 259  
 220 Woodroffe and Long (2010) demonstrated that highest 260  
 221 astronomical tide (HAT) causes less distortion in the 261  
 222 upper tidal range, a key locale in foraminifera recon- 262  
 223 structions (Gehrels *et al.*, 2004). Wright *et al.* (2011) 263  
 224 investigated this cross site standardisation further and 264  
 225 concluded that the well-established ecological relation- 265  
 226 ship between HoF and the upper limit of marine influ- 266  
 227 ence is more effective yet. However, sampling up to the 267  
 228 elevation of HoF is only definitively reached at two of 268  
 229 the sites (Brancaster and Ythan) and we therefore use 269  
 230 the highest common datum, HAT, in the standard SWLI 270  
 231 equation (Eq. 1) (Horton *et al.*, 1999) 271

$$272 \quad SWLI_n = \frac{100(h_n - MTL_s)}{HAT_s - MTL_s} + 100 \quad (1) \quad 273$$

232 such that where  $SWLI_n$  is the standardised water level 274  
 233 index of the sample ( $n$ ),  $h_n$  is the height of the sample in 275  
 234 the local datum,  $MTL_s$  and  $HAT_s$  are the mean tide level 276  
 235 and highest astronomical tide of the site ( $s$ ) in the local 277  
 236 datum. 278

237 The full regional training set, known forthwith as the 279  
 238 *North Sea training set*, was sub-divided in order to test 280  
 239 the effect on model performance of training set size and 281  
 240 tidal range, that naturally coincides with geographic re- 282  
 241 gions (Fig. 1b) owing to the North Sea tidal set up. 283

The data was screened such that samples were re-  
 242 tained if total counts were  $> 75$  in order to be able to re-  
 243 construct sea-level change at a decimetre- and decadal-  
 244 scale (Kemp *et al.*, 2020). The nine sites provided a  
 245 total of 265 samples of foraminifera assemblages from  
 246 52 different taxa. The screening of the raw data of each  
 247 training set resulted in between 17 and 35 % (*Northwest*  
 248 and *Ythan* respectively) of samples being removed. The  
 249 large variation in samples being removed is largely due  
 250 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		No. of taxa	Tidal range (m)	References
		Raw	Screened			
Alnmouth	Estuarine back-barrier	20	20	29	3.13	Horton and Edwards (2006)
Brancaster	Open coast back-barrier	59	49	24	3.89	Gehrels <i>et al.</i> (2001); Horton and Edwards (2006)
Cowpen	Estuarine back-barrier	31	30	24	3.21	Horton and Edwards (2006)
Kjelsf	Open embayment	27	12	12	1.04	Gehrels and Newman (2004)
Rantann	Open coast back-barrier	33	28	14	1.75	Müller-Navarra <i>et al.</i> (2017)
Sænderho	Open embayment	14	12	12	1.66	Müller-Navarra <i>et al.</i> (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

284 WA by adjusting taxa optima (Juggins and Birks, 2012; 336  
 285 Kemp and Telford, 2015). Each extra component can be 337  
 286 seen as adding extra, hypothetical, environmental vari- 338  
 287 ables. WAPLS generally outperforms WA because; 1) 339  
 288 'edge effects' that affect WA where optima are overes- 340  
 289 timated at the low end and underestimated at the high 341  
 290 end (Mohler, 1983) can be reduced or eliminated by 342  
 291 WAPLS; and/or 2) because in reality the composition  
 292 of taxa assemblages is influenced by additional factors  
 293 other than elevation and WAPLS can exploit the resul-  
 294 tant structured pattern in the residuals after WA; and/or  
 295 3) the fit of outliers may be improved (Birks, 1995;  
 296 Juggins and Birks, 2012). WAPLS is thus favoured in  
 297 many sea-level studies (Barlow et al., 2013). We chose  
 298 the best performing model with the fewest numbers of  
 299 components, limited to three, following the rule of par-  
 300 simony (Birks, 2012), only selecting successive compo-  
 301 nents if a significant performance increase, based on the  
 302 root mean squared error of prediction (RMSEP) and  $R^2$   
 303 values (Birks, 1998), of  $> 5\%$  is observed. These are  
 304 from here on known as the WAPLS transfer functions  
 305 with the number of components given as  $c$  (e.g. *WAPLS-*  
 306 *cI*).

307 Locally weighted transfer functions (LW) can take  
 308 advantage of both local and regional training sets by  
 309 using a set number of closest analogues for each fos-  
 310 sil sample as defined by Modern analogue technique  
 311 (MAT) and predict the sea level using these (Kemp  
 312 and Telford, 2015). We chose the 50 closest analogues  
 313 (Birks, 2012) in the final training set and then dynam-  
 314 ically developed a WAPLS transfer function for each fos-  
 315 sil sample, given as LW-WAPLS.

316 To assess the best choice of transfer function for  
 317 a sea-level reconstruction, we collected a core (A7.5)  
 318 from the Ythan Estuary using a 50 mm Russian corer  
 319 at  $57^{\circ}20.49' N$ ,  $2^{\circ}0.54' W$ . The site was selected based  
 320 on published stratigraphy by Smith *et al.* (1999) who  
 321 dated the peat at the base of a core at 5.19 m core depth  
 322 to c. 8.5 ka cal BP. This basal peat is overlain by salt-  
 323 marsh and mudflat deposits, and topped by a sand  
 324 deposit that is attributed to the Storegga tsunami of c. 8.15  
 325 ka cal BP (Dawson *et al.*, 2011; Bondevik *et al.*, 2012).  
 326 We took 1-cm-thick subsamples from the core at appro-  
 327 priate depths with respect to lithological and biological  
 328 changes in an iterative manner between the basal peat  
 329 at 5.17 m and the base of the sand deposit at 2.28 m  
 330 core depth. Thus the core spans a period in the early  
 331 Holocene that is of interest as it potentially includes  
 332 the sea-level rise associated with the drainage of glacial  
 333 lakes Agassiz and Ojibway at ca. 8.4 ka cal BP (Bar-  
 334 ber *et al.*, 1999; Li *et al.*, 2012; Hijma and Cohen, 2010,  
 335 2019; Lawrence *et al.*, 2016). A total of 24 samples

336 were prepared and foraminifera counted as described  
 337 above.

The resultant WA, WAPLS and LW-WAPLS trans-  
 338 fer functions were used to predict SWLIs from the core  
 339 fossil foraminifera. This was converted to an indicative  
 340 meaning in metres by reversing Eq. 2 using the follow-  
 341 ing equation:  
 342

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

343 where  $I_f$  is the indicative meaning of the fossil sample  
 344 ( $f$ ) in the local datum,  $SWLI_f$  is the predicted SWLI  
 345 of the fossil sample ( $f$ ),  $HAT_b$  and  $MTL_b$  are the highest as-  
 346 tronomical tide and mean tide level of the site ( $b$ ) in the  
 347 local datum. The transfer functions also produce predic-  
 348 tion sample specific errors ( $SSE$ ) for the fossil samples  
 349 that is converted in the same manner and used as the  
 350 range of the reconstructions ( $range = I_f \pm SSE$ ).

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

- 358 1. Detrended Correspondence Analysis (DCA) with  
 359 fossil samples passively projected to assess the  
 360 samples and taxa groupings and relationships (Ed-  
 361 wards and Wright, 2015).
- 362 2. Constrained correspondence analysis (CCA) to  
 363 quantify the proportion of variance explained.
- 364 3. Cluster analysis using partitioning around medoids  
 365 (PAM) with Euclidean distances (Kaufmann and  
 366 Rousseeuw, 1990; Rousseeuw, 1987) to recognise  
 367 clusters of modern and fossil foraminiferal assem-  
 368 blages. The highest average silhouette width was  
 369 used to determine the appropriate number of parti-  
 370 tions (Kemp *et al.*, 2013).
- 371 4. Modern analogue technique (MAT) using the min-  
 372 imum dissimilarity coefficient (MinDC) to mea-  
 373 sure the dissimilarity between fossil samples and  
 374 the closest modern analogue using the commonly  
 375 applied Chord-squared distance metric (Kemp and  
 376 Telford, 2015). We applied the conservative  
 377 method of defining 'good' and 'close' modern ana-  
 378 logues for fossil samples as having a MinDC lower  
 379 than the 5<sup>th</sup> and 20<sup>th</sup> percentiles respectively and  
 380 samples with a larger MinDC defined as 'poor'  
 381 (Watcham *et al.*, 2013; Barlow *et al.*, 2013).

- 382 5. Transfer function performance statistics using  
383 bootstrapping (boot) and leave-one-site-out  
384 (LOSO) cross-validation to assess the quality of  
385 the transfer functions and the independence of  
386 samples (Telford and Birks, 2005; Payne *et al.*,  
387 2012; Kemp *et al.*, 2013).  
388 6. If using WAPLS with > 2 components the pat-  
389 tern of taxa optima updates were analysed to assess  
390 the effect of introducing more statistical complex-  
391 ity that may distort reconstructions (Wright *et al.*,  
392 2011).  
393 7. Ranges of the different core reconstructions were  
394 compared to evaluate the likely accuracy of each  
395 regional, subregional and local transfer function.  
396 8. The statistical significance of the reconstructions  
397 was measured by comparing the reconstructions  
398 from the training sets against a set of 999 randomly  
399 generated data using redundancy analysis ordina-  
400 tion (Telford and Birks, 2011).

### 401 3. Results and Discussion

#### 402 3.1. Training set development

##### 403 3.1.1. Ythan Estuary data

404 Foraminifera occur in the Ythan Estuary from the  
405 lowest sampled point of just above MTL across the ele-  
406 vation gradient of the tidal zone to just above HAT. Both  
407 dead and living individuals were found up to an eleva-  
408 tion of 2.45 m OD (SWLI = 202). A total of 15 taxa  
409 were found although four had a maximum abundance  
410 < 10 %. The majority of taxa show some relationship  
411 with elevation (see Fig. 2). Taxa elevation optima (the  
412 abundance weighted average of SWLI in all samples  
413 in which they occur (ter Braak and Barendregt, 1986))  
414 and tolerances (the abundance weighted standard deviation  
415 (Birks *et al.*, 1990)) following WA show that the  
416 different taxa occupy a gradient of different elevations  
417 (Fig. 2c). The highest marsh samples, above MHS  
418 at 1.96 m OD where tidal inundation is infrequent, are  
419 low in total foraminifera counts and hence only one was  
420 retained following screening. In these samples *Baltic-*  
421 *cammina pseudomacrescens* appears in relatively high  
422 abundance much like in the eastern North Sea marshes  
423 (Gehrels and Newman, 2004; Müller-Navarra *et al.*,  
424 2017). The salt-marsh samples are almost entirely com-  
425 posed of agglutinated taxa. *Entzia macrescens*, also  
426 commonly referred to as *Jadammina macrescens*, is the  
427 most common salt-marsh taxon and is found in simi-  
428 lar abundances across the marsh. *Miliammina fusca* is  
429 found in greatest abundances in the lower marsh like  
430 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 abun-  
dances (> 75 %) compared to other UK salt marshes  
(Horton *et al.*, 1999). There is a clear shift to assem-  
blages dominated by calcareous species at 1.22 m OD  
just below the transition from salt marsh to mudflat,  
along with occasional *Trochammina ochracea* and *Hap-*  
*lophragmoides* spp. PAM clustering displays two dis-  
tinct clusters of these mudflat samples, one with *Hay-*  
*nesina germanica* most abundant and featuring high  
abundances of *Buliminella elegantissima* and *Brizalina*  
*variabilis*, and a second with few *Buliminella elegan-*  
*tissima* and *Brizalina variabilis* and dominated by *El-*  
*phidium williamsoni*, *Trochammina ochracea* and *Hap-*  
*lophragmoides* spp. Though usually found in the veg-  
etated 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.

##### 405 3.1.2. Compiled North Sea data

406 The foraminifera of the eight previously published  
407 sites and the Ythan (see Table 1) combined in the re-  
408 gional *North Sea* data occur at elevations from 273  
409 SWLI, well above HAT, to 70, below MTL (Fig. 3). The  
410 highest 14 samples (> 210 SWLI), well above HAT, are  
411 all from the sites in Denmark and Germany. This could  
412 be due to the wind-induced setup that occurs in these  
413 areas (Bartholdy *et al.*, 2004; Pedersen *et al.*, 2009),  
414 where tides > 3 m above HAT in Esbjerg have been  
415 recorded (Bartholdy *et al.*, 2010). These samples are  
416 generally restricted to *Balticammina pseudomacrescens*  
417 and *Entzia macrescens* which are species able to survive  
418 despite infrequent flooding (Berkeley *et al.*, 2007). Al-  
419 though micro tidal sites are generally desired because  
420 they provide higher precision reconstructions (Callard  
421 *et al.*, 2011; Barlow *et al.*, 2013; Edwards and Wright,  
422 2015) they will be more influenced by this effect and  
423 hence the micro tidal range of the eastern sites (all <  
424 1 m) may be causing the extremely high SWLI values  
425 seen in these sites a phenomenon also seen in some sites  
426 in the US (Wright *et al.*, 2011). The assemblages of  
427 the remaining salt-marsh samples are dominated by ag-

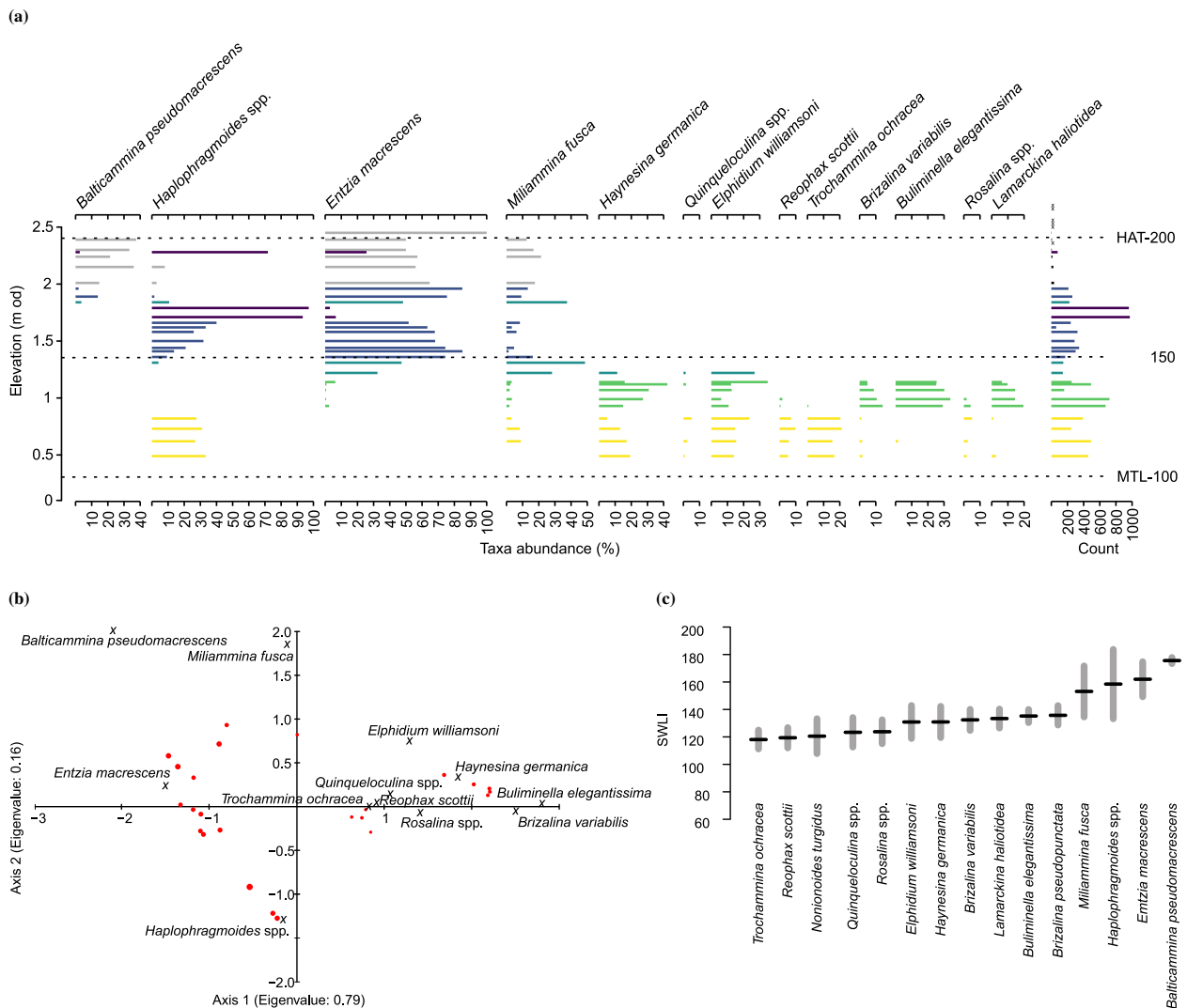


Figure 2: **Summary diagrams of the modern foraminifera assemblages for the Ythan Estuary:** a) Foraminifera abundances as percentages of the count of total dead specimens (given in the final column) against elevation. Only taxa with > 5 % maximum abundance are shown. Grey bars denote samples that were excluded following screening and Xs mark where no foraminifera were found. The bars are coloured according to PAM clustering of the dataset (average silhouette 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.

482 glutinated foraminifera. *Entzia macrescens* is abundant 482  
 483 across the marsh, as is *Miliammina fusca* although in 483  
 484 lower abundances. *Trochammina inflata* appears mainly 484  
 485 restricted to the higher marsh, while *Haplophragmoides* 485  
 486 spp. become more prevalent in the middle-lower marsh 486  
 487 mainly at Rantum, Sønderho and Ythan. Calcareous 487  
 488 taxa become more abundant in the lower marsh, 488  
 489 although are found much higher particularly in eastern 489  
 490 marshes which may be due to the wind-induced setup 490  
 491 and/or high pH promoting better preservation (Müller- 491

Navarra *et al.*, 2016). Relatively high abundances of 492  
 493 *Ammonia* spp. are found at Brancaster and Thornham 493  
 494 compared to the other sites, while *Haynesina germanica* 494  
 495 is notably abundant at Welwick. Foraminifera were not 495  
 496 sampled below 140 SWLI, approximately at the transition 496  
 497 between the salt marsh and the mudflat, at Bran- 497  
 498 caster, Kjelst and Thornham. However, the assemblages 498  
 499 from the sites that were sampled show a clear domi- 499  
 500 nance of calcareous taxa. The taxa abundances appear 500  
 501 to show a relationship with elevation across the training

502 set, although it may be reduced at the eastern sites due 554  
503 to local conditions. 555

504 The relationship between taxa and elevation is confirmed 556  
505 by the cluster analysis (Figs. 3 and 4) and DCA 557  
506 (Fig. 5). PAM clustering produces the highest average 558  
507 silhouette width when samples are split into three clusters, 559  
508 shown in Figs. 3 and 4. Cluster 1 mainly comprises 560  
509 the mudflat samples (defined as SWLI approximately < 561  
510 140) dominated by calcareous taxa from a mixture of 562  
511 sites. Cluster 2 mainly comprise samples where *Haplophragmoides* 563  
512 spp. are abundant as described above. 564  
513 While *Haplophragmoides* spp. are not rare taxa and 565  
514 are found at varying elevations relative to tidal inundation 566  
515 in other marshes (e.g. Berkeley *et al.*, 2007, for 567  
516 a review), they are marked out in these samples by its 568  
517 high abundance that cause high intra-cluster similarity 569  
518 and low dissimilarity with other samples. Finally, sam- 570  
519 ples in cluster 3 contain the majority of salt-marsh sam- 571  
520 ples. The majority of samples are dominated by *Entzia* 572  
521 *macrescens* and encompass the full elevation range of 573  
522 the salt marsh. A cluster of samples that are almost 574  
523 entirely found above 170 SWLI and comprise *Entzia* 575  
524 *macrescens*, *Miliammina fusca* and *Trochammina in-* 576  
525 *flata* is identified from within cluster 3 if we add extra 577  
526 clusters. 578

527 A feature of the data appears to be a small degree of 579  
528 clustering by sites. This could be due to a lack of inde- 580  
529 pendence of samples because of spatial auto-correlation 581  
530 that may cause transfer function performance statistics 582  
531 to be overly optimistic (Telford and Birks, 2005; 583  
532 Payne *et al.*, 2012), or that each site has unique con- 584  
533 trols that could make a regional transfer function either 585  
534 problematic or arguably more robust (Legendre and 586  
535 Fortin, 1989). To investigate this further LOSO cross- 587  
536 validation was applied and is discussed below. 588

537 The DCA plots summarise the unconstrained relation- 589  
538 ship between samples, taxa and SWLI (Fig. 5). The 590  
539 plot shows that the samples tend to align towards axis 1 591  
540 and appear somewhat correlated with SWLI with general 592  
541 clustering of higher and lower SWLI samples. The 593  
542 relationship is confirmed by CCA such that 7.13 % of 594  
543 the variance can be explained by elevation, and that the 595  
544 ratio of the axis constrained by SWLI and the first uncon- 596  
545 strained axis is 1.04, indicating it is an important 597  
546 variable in explaining taxa distributions (Kemp *et al.*, 598  
547 2013). A number of samples diverge from axis 1 in the 599  
548 DCA plot and may therefore be influenced by secondary 600  
549 variables. These samples are those described above, 601  
550 where *Haplophragmoides* spp. are found at higher el- 602  
551 evations. Although other environmentally variables are 603  
552 generally correlated with elevation (i.e. salinity, pH), it 604  
553 is evident that other environmental variables are affect-

554 ing the taxa niches which is not unexpected considering 555  
556 the inclusion of multiple sites from differing environ- 557  
558 ments and tidal ranges. The taxa-sample relationships 559  
560 shown in Fig. 5b confirm the effect of *Haplophrag-*  
561 *moides* spp. and also notably how *Balticammina pseu-*  
562 *domacrescens* drives the samples with the highest ele-  
563 vations. Collections of taxa are evident along axis 1,  
564 where agglutinated and calcareous taxa show a general  
565 partition. Some modern Ythan samples appear to be  
566 consistently distal from other sites and/or axis 1 sug-  
567 gesting they are providing assemblages with somewhat  
568 differing response to elevation. However, similarities  
569 with other samples from across the North Sea suggest  
570 comparable conditions may be occurring at a wide range  
571 of sites.

572 PAM and DCA analysis suggests that the 573  
574 foraminifera assemblages that make up the *North* 575  
576 *Sea* training set are appropriate as sea-level indicators. 577  
578 The agglutinated taxa appear particularly well suited, as 579  
579 observed by others, although may lack more definitive 580  
580 zonation (e.g. Horton *et al.*, 1999; Gehrels, 2000; 581  
582 Gehrels *et al.*, 2005; Horton and Edwards, 2006; Kemp 583  
584 *et al.*, 2013; Barnett *et al.*, 2016). Calcareous taxa are 584  
585 useful indicators of tidal flats in the data and hence 585  
586 useful for transfer functions where a wide environ- 586  
587 mental range may be sought. However, the lowest 587  
588 occurrence of some taxa is not sampled meaning the 588  
589 full range is not captured, which could be problematic 589  
590 for predicting the lowest range of some fossil samples 590  
591 and so these should be treated with an element of 591  
592 caution (Woodroffe, 2009). Calcareous taxa also occur 592  
593 across a wide elevation range in the region, which may 593  
594 impact the predictive ability of the transfer functions. 594  
595 This wide range may be due to in-wash resulting in 595  
596 allochthonous foraminifera (Murray, 2003; Horton and 596  
597 Murray, 2006). Some studies attempt to avoid this 597  
598 effect by either not sampling the tidal flat (e.g. Gehrels, 598  
599 2000; Kemp *et al.*, 2013)) and/or removing these so 599  
600 termed 'exotic' taxa (e.g. Horton and Edwards, 2006; 600  
601 Kemp *et al.*, 2009; Leorri *et al.*, 2011; Mills *et al.*, 601  
602 2013) resulting in better predictability in some cases. 602  
603 However, Mills *et al.* (2013) observe that performance 603  
604 does not improve in their Mersey training set and that 604  
605 this method does not preclude the mixing of lower 605  
606 marsh agglutinated foraminifera. We include all sam- 606  
607 ples as our core includes a fossil mudflat environment 607  
608 and thus the lower samples are critical as well as 608  
609 providing a lower limit to some of the agglutinated 609  
610 marsh taxa (Wright *et al.*, 2011). However, other cores 610  
611 that are limited to salt marsh environments may warrant 611  
612 investigation of the effect of removing lower elevation 612  
613 samples or taxa. 613



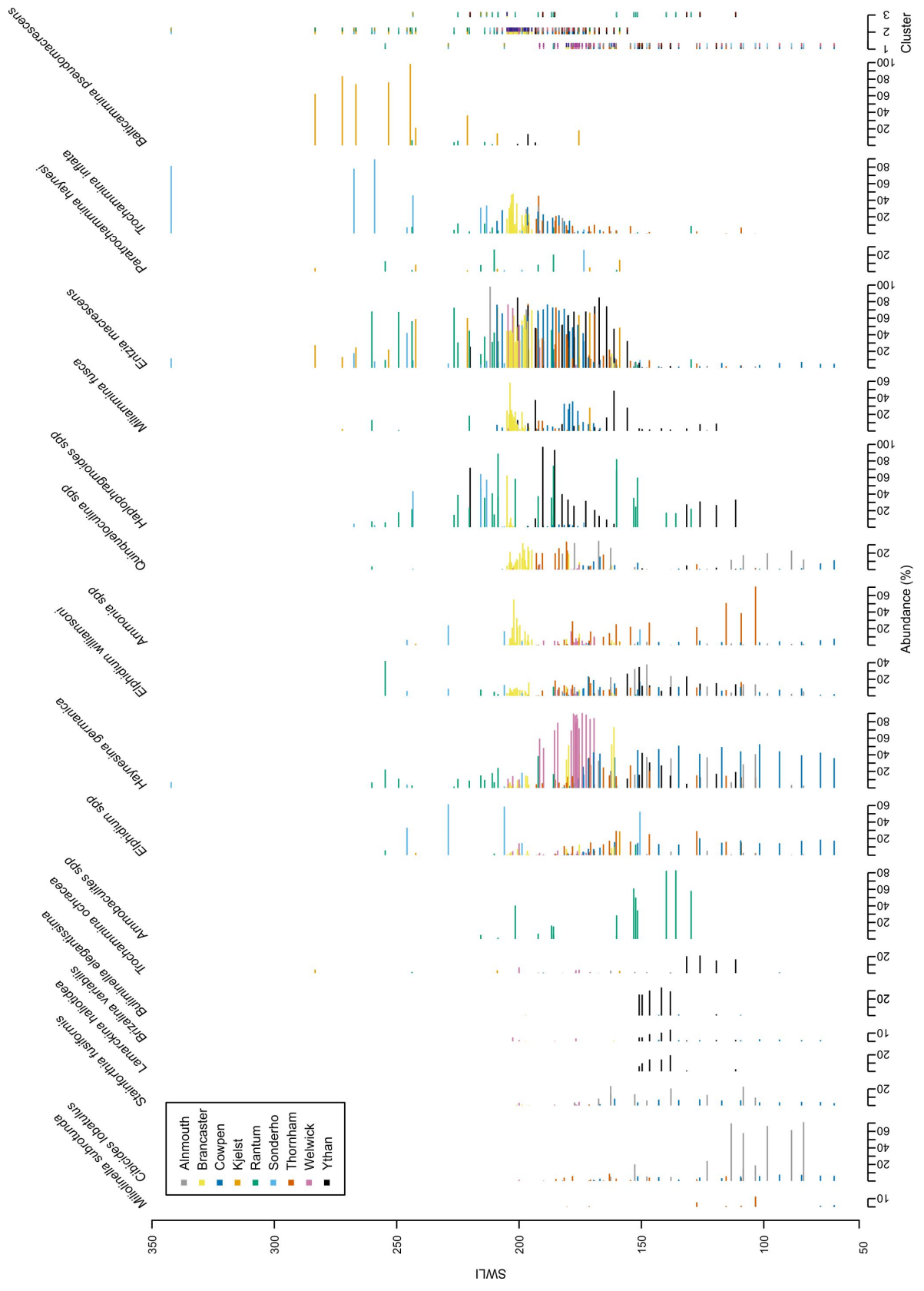


Figure 3: The combined North Sea foraminifera data of 218 samples from 9 sites ordered by SWLI value as shown in the first column. Taxa are included that have a maximum abundance > 10 % and occur in more than 10 % of the samples. The bar colour denotes the site corresponding to the legend. The final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

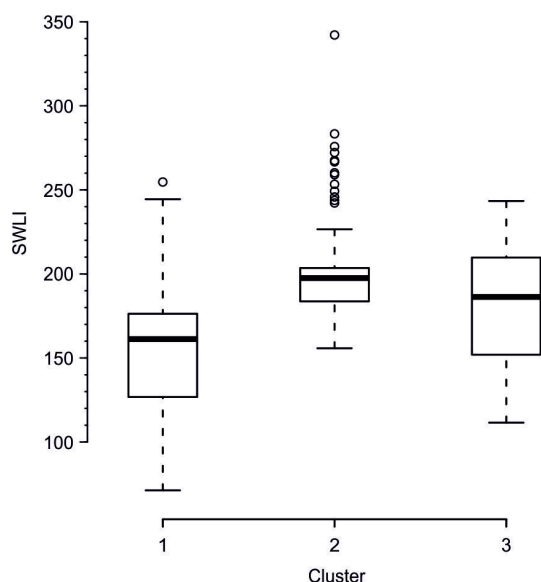


Figure 4: Boxplot of the clusters following PAM clustering of the North Sea training set. The boxplots represent the 2.5, 16, 50, 84 and 97.5 percentiles for each cluster. Outliers are represented by the circles.

### 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

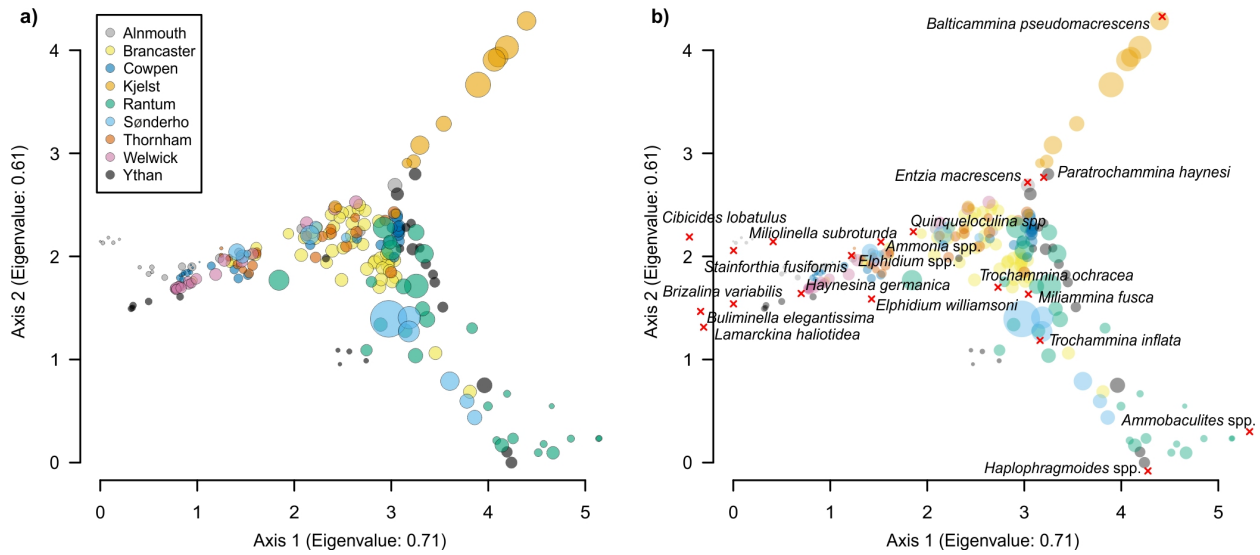


Figure 5: **Detrended canonical analysis (DCA) of the North Sea training set.** (a) Samples are plotted against elevation and coloured according to site and sized according to SWLI value. (b) Species are added and labelled if the maximum abundance > 10 %.

672 localised training sets in common with other sea-level  
 673 studies (Horton and Edwards, 2006; Woodroffe and  
 674 Long, 2010; Barlow *et al.*, 2013). The exceptions are  
 675 the *East* transfer functions that perform particularly  
 676 poorly, and the *Northwest* that performs worse than the  
 677 *West*. The taxa optima and tolerance plots (Fig. 8)  
 678 demonstrate that the taxa optima within the eastern sites  
 679 are highly variable and hence the poor performance of  
 680 the *East*, which in turn negatively impacts the perform-  
 681 ance of the *North Sea* transfer functions. Whilst this  
 682 provides a wider range of environmental response it  
 683 comes at the expense of less precision. All of the WA  
 684 transfer functions have a higher RMSEP and average  
 685 bias than the corresponding WAPLS models, although  
 686 they do tend to have a lower maximum bias, and vari-  
 687 able  $R^2$  values. This suggests that classical deshrink-  
 688 ing performs better at the extreme ends of the gradi-  
 689 ent, while inverse deshrinking (used in WAPLS models)  
 690 provides higher accuracy for those in the middle as in-  
 691 dicated by Birks (1995) and Juggins and Birks (2012).  
 692 Comparison of the predicted SWLI against observed  
 693 SWLI (Fig. 9) shows that although better fitting for the  
 694 most extreme sample may be better in WA transfer func-  
 695 tions, WAPLS appears to perform better across the  
 696 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 ( $p < 0.05$ ).

Model	SWLI <sub>train</sub>		SWLI <sub>test</sub>			Ythan A7.5 (m) <sub>test</sub>		Significance $p$
	RMSEP	RMSEP	$R^2$	Av. Bias	Max. Bias	RMSEP	Max. bias	
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 RMSEP<sub>loso</sub> was greater than

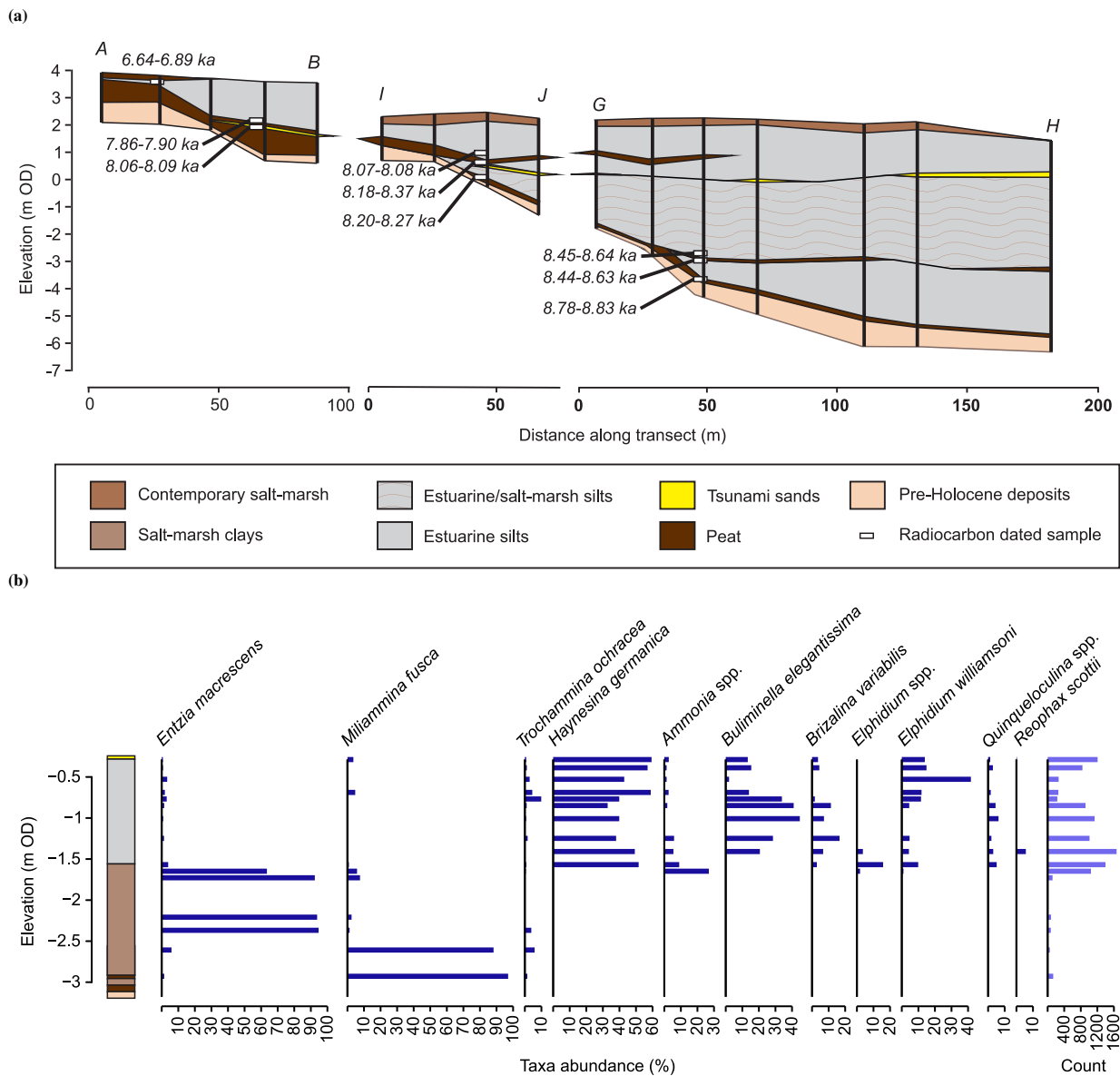


Figure 6: **Summary of Ythan palaeo data:** a) The lithology and calibrated radiocarbon ages of the transects sampled in the work by Smith *et al.* (1999) that correspond to Fig. 1. b) Lithology and fossil foraminifera assemblages of core A7.5 from the Ythan Estuary. Foraminifera abundances are given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

714 RMSEP<sub>boot</sub> (see Table 2). However, the differences are  
 715 relatively small suggesting only a limited degree of pos-  
 716 sible spatial auto-correlation may be present, or alterna-  
 717 tively the clustering of samples between sites reflects  
 718 variable local environmental conditions. The inclusion  
 719 of many sites also minimises the effect of spatial auto-  
 720 correlation (Legendre and Fortin, 1989; Telford and  
 721 Birks, 2005) whilst simultaneously making the larger  
 722 regional transfer functions more robust against environ-  
 723 mental changes occurring at a particular site (Barnett

724 *et al.*, 2016). Fig. 10 shows that the sites perform rela-  
 725 tively consistently suggesting that one particular site is  
 726 not being unduly effected and that all sites likely display  
 727 some environmental variability.

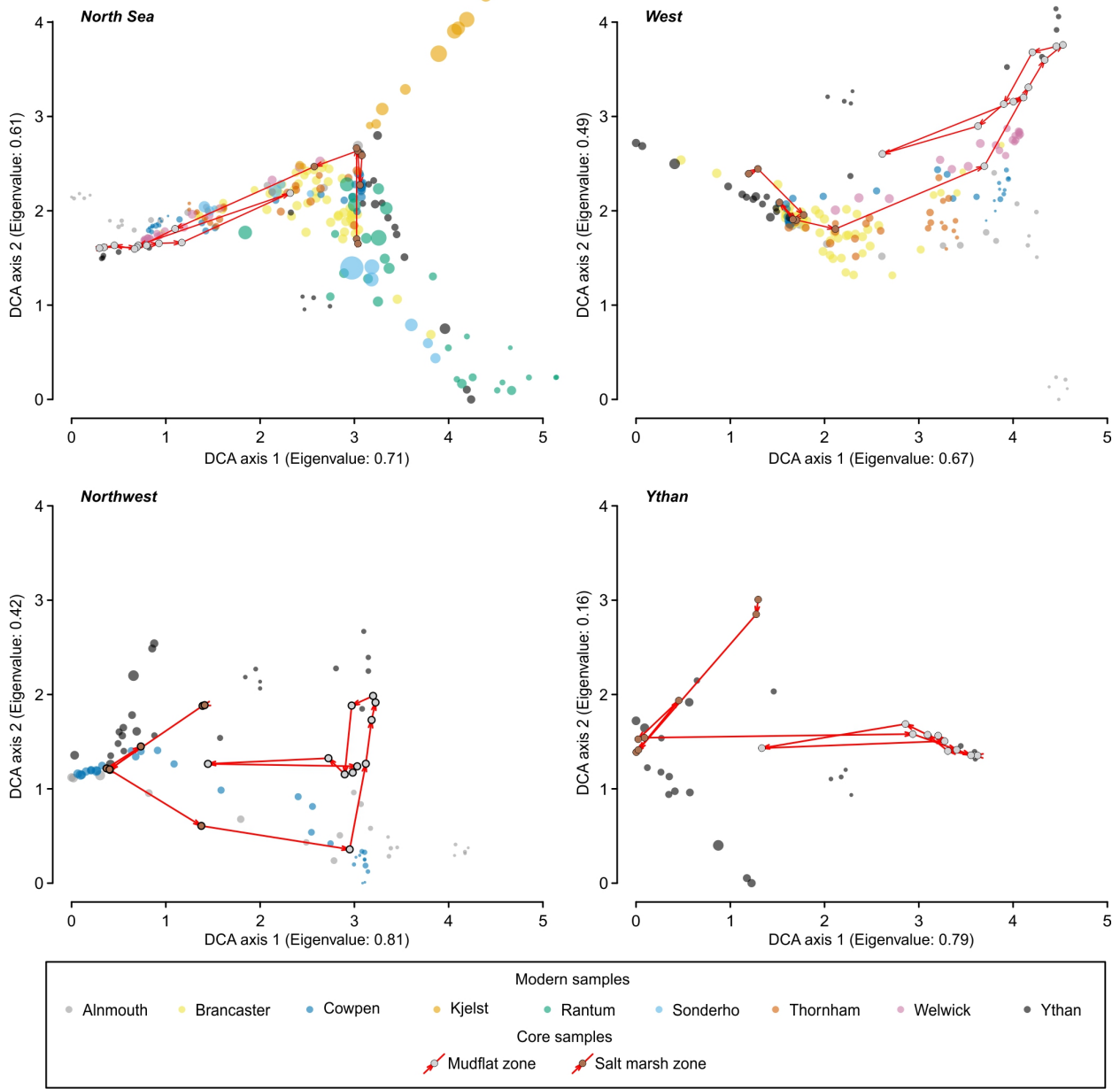


Figure 7: **Correspondence analysis of training sets and fossil data** showing the four sub-regional training sets and Ythan core samples using detrended correspondence analysis (DCA). The modern samples are coloured by site and the size is scaled by SWLI. The core samples are coloured by the lithology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

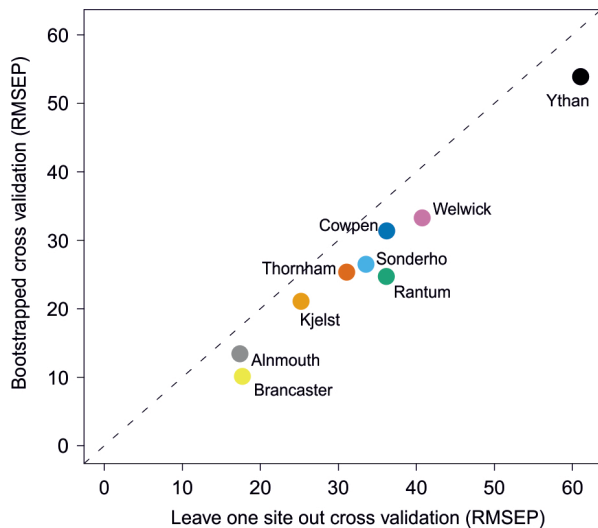


Figure 10: Comparison of root-mean-square error prediction

728 To further assess transfer function performance and  
 729 understand the effect of deshrinking methods and  
 730 adding extra complexity in WAPLS with components  
 731 > 2, we compared scatterplots of observed versus pre-  
 732 dicted SWLI and the updated taxa optima (Wright *et al.*,  
 733 2011). Fig. 8 shows that when inverse deshrinking  
 734 (WAPLS component 1) as opposed to classical is ap-  
 735 plied low elevation taxa are under predicted and high

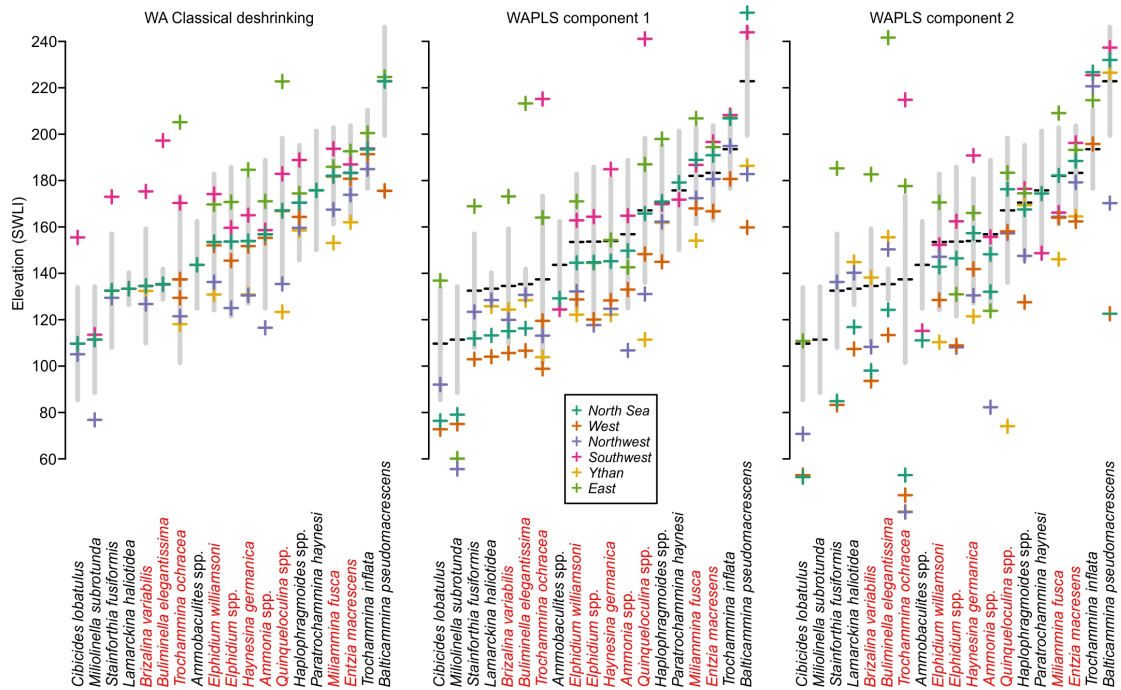


Figure 8: **Comparison of the progressive taxa optima updates** for the different training sets from WA with classical and inverse (equivalent to WAPLS component 1) to more complex WAPLS transfer functions with multiple components. Taxa tolerances are shown for WA by the grey bar. Taxa are included that 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)

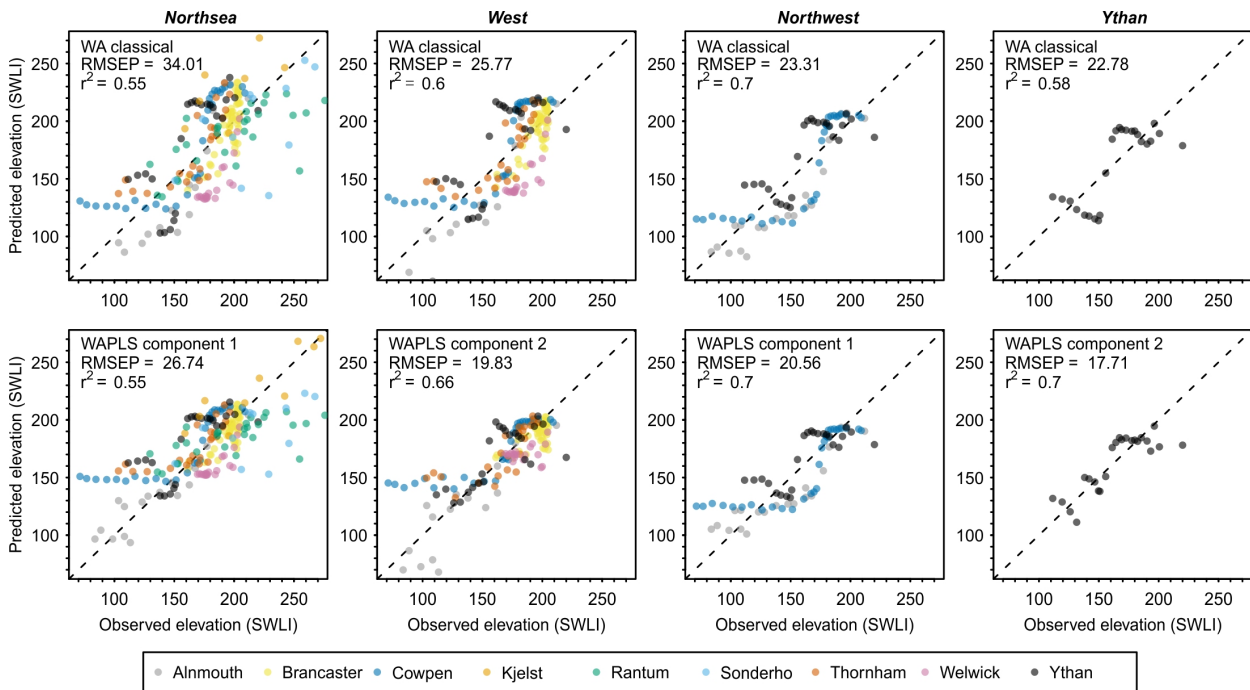


Figure 9: **Scatterplots of the observed SWLI against the predicted SWLI** for WA classical deshrinking and WAPLS components 1 or 2 depending on which was selected (see text). The samples are coloured according to site as shown in the legend.

736 salt-marsh taxa are over predicted, although this is less 787  
737 apparent in Fig. 9 where WAPLS consistently provides 788  
738 more accurate predictions. This is likely because the 789  
739 most abundant taxa tend to have optima towards the 790  
740 centre of the elevation gradient in the training set and 791  
741 are therefore more suitable for inverse deshrinking. Al- 792  
742 though there is noticeable variability between training 793  
743 sets in Fig. 8, the *North Sea* and *West* taxa optima 794  
744 are very similar and both appear relatively stable to 795  
745 the effect of adding extra components in WAPLS trans- 796  
746 fer functions. Adding a second component to the *West* 797  
747 WAPLS transfer function can be seen to produce accu- 798  
748 rate sample predictions shown in Fig. 9 by updating 799  
749 many of the taxa optima (Fig. 8). However, of the taxa 800  
750 that receive major updates only *Trochammina ochracea* 801  
751 is found in the core and in low abundances and therefore 802  
752 the reconstructions will not be distorted. 803

753 Despite the good performance, structure remains in 804  
754 the dataset (see Fig. 9) such that some of the samples 805  
755 found across the elevation gradient of the mudflat (par- 806  
756 ticularly those of Cowpen) have very similar predictions 807  
757 because the samples have very similar taxa abundances. 808  
758 This demonstrates the difficulty in the application of 809  
759 mudflat samples. One option would be to remove these 810  
760 samples but they represent the full environmental gradi- 811  
761 ent of the samples and so we favour retaining all sam- 812  
762 ples but treating predictions with SWLI values < 140 813  
763 with extra caution. 814

764 Based on transfer function performance statistics 815  
765 alone, the inclination would be to choose the more lo- 816  
766 calised training sets for sea-level reconstructions for 817  
767 core A7.5. However, incorporating more sites provides 818  
768 more robustness against the possible effects of clus- 819  
769 tered samples and add additional variability in a taxa- 820  
770 elevation relationship that is likely important for pro- 821  
771 viding analogues for early Holocene foraminifera. Taxa 822  
772 optima are also more stable with a greater number of 823  
773 sites and suggest these may also be valid. 824

#### 774 3.2.4. Relative sea-level reconstructions 825

775 We applied the preferred WA and WAPLS transfer 826  
776 functions to reconstruct the palaeo marsh elevation re- 827  
777 lative to MTL at the Ythan Estuary using the fossil as- 828  
778 semblages found in core A7.5 (see Fig. 6). All of the re- 829  
779 constructions show a similar pattern of indicative mean- 830  
780 ing change, with periods of rapid decreases (i.e., rela- 831  
781 tive sea-level rise) at core depths of around -3 m and 832  
782 -1.6 m set against a trend of more gradual decrease ( 833  
783 Fig. 11). The reconstructions produced notable differ- 834  
784 ences depending on the training set and whether WA 835  
785 or WAPLS transfer functions were used. The recon- 836  
786 structions using the *North Sea* and *West* training sets 837  
838

display very similar patterns to each other, although the 839  
precision, as quantified by the sample specific error, in- 840  
creases from *North Sea* to *West*. The *West* model oc- 841  
casionaly predicts slightly lower SWLI which is due 842  
to the exclusion of the eastern sites that generally have 843  
higher taxa optima. The ranges of the *Northwest* and 844  
*Ythan* reconstructions show divergence from the *North 845*  
*Sea* and *West* at differing periods and magnitudes across 846  
WA and WAPLS transfer functions. The number of 847  
good and close modern analogues as assessed by the 848  
MinDC become greater with increased number of sites 849  
in the training sets, for example increasing from nine 850  
for the *Ythan* to 24 for the *West* and *North Sea* (Fig. 11). 851  
The extra samples in the *North Sea* and *West* training 852  
sets evidently provide better analogues that also drive 853  
differing SWLI predictions from the others and hence 854  
suggest the *Northwest* and *Ythan* produce inaccurate re- 855  
constructions. 856

857 The ranges of the WA and WAPLS reconstructions 858  
show overlap in all cases, although there is a probable 859  
difference in the magnitude of both the first and second 860  
decrease in indicative meaning. The WA based recon- 861  
structions tend to predict lower surface elevations for 862  
mudflat samples and higher surface elevations for salt- 863  
marsh samples than the WAPLS models. This results in 864  
the first period of change being apparently larger in the 865  
WAPLS reconstructions compared to the WA, whilst the 866  
second is greater in the WA reconstructions. The fos- 867  
sil samples that bound the periods of rapid change ap- 868  
pear indicative of mudflats and middle-upper salt marsh 869  
with taxa, from towards the gradient ends; predictions 870  
of assemblages in these zones are therefore important. 871  
Fig. 8 shows that WAPLS appears better at remov- 872  
ing the edge-effects than WA even when one compo- 873  
nent is used (Mohler, 1983) and should thus provide 874  
more accurate reconstructions of the fossil mudflat sam- 875  
ples. Small changes in species abundances towards the 876  
ends of the gradient can substantially alter predictions 877  
(Gehrels, 2000). At the upper end this is generally ben- 878  
eficial because of the high precision of high salt-marsh 879  
species (Scott and Medioli, 1978; Gehrels, 2000), but 880  
can be problematic at lower ends where taxa tolerance 881  
can be less precise due to wide apparent niches and taxa 882  
sometimes being found at sub-tidal elevations (Berkeley 883  
*et al.*, 2007). Lower elevation samples are sometimes 884  
removed because of the uncertainty in taxa response to 885  
elevation on the mudflats (Edwards and Wright, 2015) 886  
especially those that show nonlinearity with elevation 887  
(Hamilton and Shennan, 2005). However, despite the 888  
difficulties of trying to reconstruct mudflat environ- 889  
ments, including the lower elevation samples is required 890  
to provide appropriate modern analogues and are there-

839 fore justified here. WAPLS transfer functions are con- 888  
840 sequently likely to provide better predictions of both the 889  
841 higher and lower elevation samples. 890

### 842 3.2.5. Statistical significance of reconstructions 891

843 We tested whether the reconstructions trained on 892  
844 the different training sets explained more variance in 893  
845 the core samples than the majority of reconstructions 894  
846 trained on randomly derived sets of data ( $n = 999$ ) as 895  
847 proposed by Telford and Birks (2011). The results show 896  
848 that only the *West-WAPLS-c2* produced a statistically 897  
849 significant reconstruction ( $p < 0.05$ ) (Table 2). The 898  
850 *West-WA* and *North Sea-WA* and *North Sea-WAPLS-c1* 899  
851 reconstructions perform relatively well and the similar- 900  
852 ity of reconstructions suggests they these are still valid. 901  
853 Telford and Birks (2011) argue that palaeoclimatology 902  
854 should not be exempt from interpreting common con- 903  
855 straints of insignificant results which can be equally ap- 904  
856 plied to sea-level studies, although Kemp *et al.* (2013) 905  
857 show that some salt-marsh cores do not provide suffi- 906  
858 cient downcore variability to outperform random data. 907  
859 In this case the fossil assemblages do provide notable 908  
860 variety and thus the significant result is somewhat rele- 909  
861 vant and provides support for the *West-WAPLS-c2* trans- 910  
862 fer function. 911

### 863 3.3. Training set and transfer function selection 912

864 The challenge of which combination of training set 913  
865 and transfer function model to apply is not a straight- 914  
866 forward decision and may lead to different outcomes as 915  
867 described above in the core reconstruction (seen in Fig. 916  
868 11) and by Barlow *et al.* (2013). Including all samples 917  
869 from a wider geographic region and a longer elevation 918  
870 gradient captures wider taxa tolerances and provides 919  
871 more modern analogues. One could therefore argue for 920  
872 including as wide a range as possible. However, the 921  
873 challenge is minimising the range of the reconstructed 922  
874 indicative meanings as much as possible by keeping 923  
875 variability of taxa tolerances to a minimum, whilst still 924  
876 providing an analogue for the past environment that 925  
877 can produce suitably precise and accurate quantitative 926  
878 reconstructions. We have built on suggested methods 927  
879 for developing transfer functions (Wright *et al.*, 2011; 928  
880 Watcham *et al.*, 2011; Barlow *et al.*, 2013; Kemp and 929  
881 Telford, 2015) with particular consideration for early 930  
882 Holocene sea-level reconstructions to assess this. 931

883 Assessment of transfer function performance statis- 932  
884 tics alone could lead to the more localised models being 933  
885 preferred due to the lower RMSEP and maximum bias. 934  
886 However, MAT, DCA and PAM silhouette plots show 935  
887 that fossil samples often lack appropriate analogues and 936

888 similarity with the modern environment in the *Ythan* 889  
889 and *Northwest* training sets. The lack of analogues and 890  
890 poor clustering of the *Ythan* and fossil samples sug- 891  
891 gest that the likelihood of the foraminifera-elevation re- 892  
892 lationship at a single site remaining unchanged over a 893  
893 period of thousands of years is unlikely. Even relatively 894  
894 local regional microfossil training sets such as 895  
895 the *Northwest* here and the West coast of Scotland di- 896  
896 atom set (Lawrence *et al.*, 2016) do not necessarily pro- 897  
897 vide suitable modern microfossil analogues and there- 898  
898 fore including sites that encompass different environ- 899  
899 mental conditions in a larger training set appears nec- 900  
900 essary. A shift in the reconstructed ranges of the re- 901  
901 constructions when moving to the more localised train- 902  
902 ing sets from the *West* and *North Sea* training sets is 903  
903 evident further suggesting the reconstructions from the 904  
904 more localised training may lose accuracy. On the basis 905  
905 of missing analogues and lack of similarity we rule out 906  
906 using the *Ythan* and *Northwest* training sets for recon- 907  
907 structing the *Ythan* core. 908

909 DCA and cluster analysis suggests that the *North* 909  
910 *Sea* and *West* models both perform well and indeed 910  
910 both have no 'poor' analogues according to MAT. The 911  
911 *North Sea* training set has more spread in the modern 912  
912 foraminifera tolerances due to the eastern sites. This 913  
913 may simply be adding unnecessary noise, or on the other 914  
914 hand be providing additional analogues and thus better 915  
915 accuracy. However, the foraminifera in the eastern sites 916  
916 appear to be overly affected by the wind-driven water 917  
917 levels that reduces their utility for reconstructing sea 918  
918 level reflected in their higher optima. Using tide loggers 919  
919 as opposed to relying on modelled tidal data may help 920  
920 overcome this issue that will be particularly important 921  
921 at micro-tidal sites. Both *West* and *North Sea* training 922  
922 sets do still produce very similar reconstructions for ei- 923  
923 ther WA and WAPLS transfer functions suggesting that 924  
924 the small increase in precision in the reconstructions and 925  
925 better predictive ability from the *West* training set does 926  
926 not come with a loss of consistency. It therefore appears 927  
927 that the modern environment captured by the sites in the 928  
928 *West* training set provides a suitable analogue for the 929  
929 conditions found at the *Ythan* Estuary around 8 ka with- 930  
930 out the noise introduced by eastern sites. We therefore 931  
931 select the *West* training set for our transfer functions. 932

933 The WA and WAPLS transfer functions result in 933  
934 similar reconstructions, with overlapping ranges. The 934  
934 main difference is the prediction of mudflat samples 935  
935 which cause periods of smaller magnitude changes in 936  
936 the WAPLS-derived indicative meanings. WAPLS ar- 937  
937 guably provides better accuracy for this zone by re- 938  
938 moving edge effects (Mohler, 1983; Juggins and Birks, 939  
939 2012) and indeed towards the centre. Furthermore, the



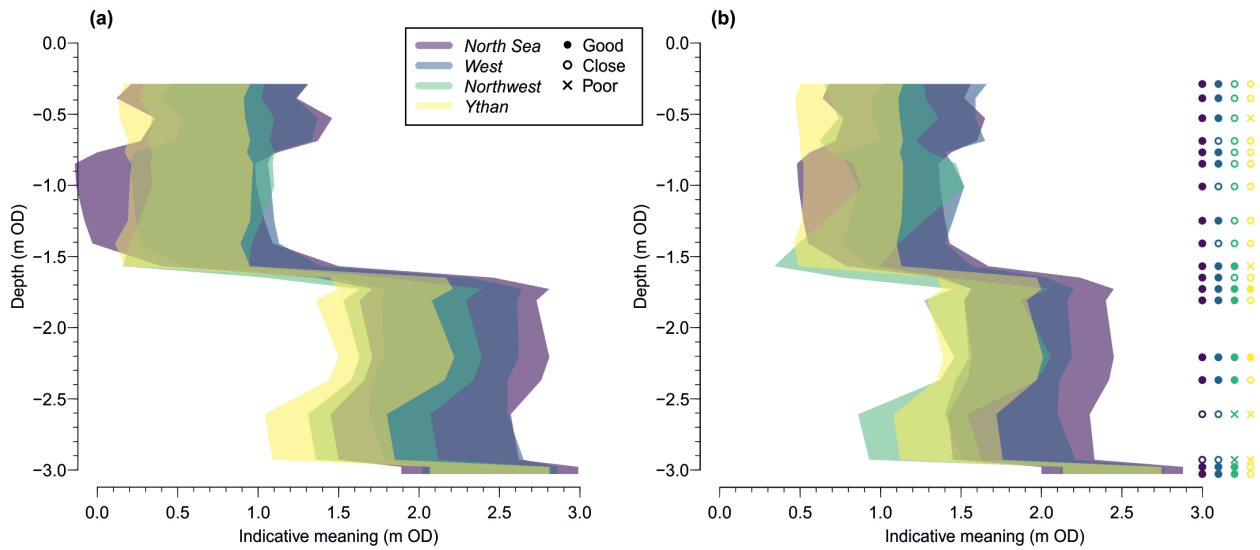


Figure 11: **Reconstructed indicative meaning of core A7.5 samples** from the different training sets using a) WA and b) WAPLS based transfer functions. 'Good', 'Close' and 'Poor' are measures of the closeness of modern analogues and are taken from the modern analogue technique (MAT) with MinDC values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

940 *West-WAPLS-c2* is the only transfer function to produce 970  
 941 a statistically significant reconstruction. Analysis of the 971  
 942 taxa optima updates (Fig. 8) when an extra component 972  
 943 is added in the *West-WAPLS-c2* transfer function show 973  
 944 a number of calcareous taxa from some sites are shifted 974  
 945 to well above HAT and so these more complex transfer 975  
 946 functions should be treated with caution (Wright *et al.*, 976  
 947 2011). However, the taxa found in the fossil assem- 977  
 948 blages are not shifted so dramatically and therefore, in 978  
 949 this instance, we accept the updates as appropriate for 979  
 950 reconstructing the core. Whilst WA and WAPLS both 980  
 951 produce similar reconstructions and the decision be- 981  
 952 tween the two does not therefore produce significantly 982  
 953 different results we prefer the *West-WAPLS-c2* transfer 983  
 954 function owing to the better fitting of predicted SWLI, 984  
 955 the reduction in edge-effects compared to WA, the re- 985  
 956 construction being statistically significant and without 986  
 957 producing dramatic taxa optima updates of the impor- 987  
 958 tant foraminifera. The *West-WAPLS-c2* transfer func- 988  
 959 tion produces a mean sample specific error of 38 cm for 989  
 960 the Ythan fossil samples which can be considered good, 990  
 961 particularly in light of the approximately 50 cm preci- 991  
 962 sion obtained in the Lawrence *et al.* (2016) sea-level re- 992  
 963 construction. 993

964 A LW transfer function may circumvent the problem 994  
 965 of including a wide range of sites and the associated 995  
 966 noise. However, simply running a LW transfer func- 996  
 967 tion using the *North Sea* training set (*North Sea-LW-* 997  
 968 *WAPLS*) produces inconsistent and unrealistic recon- 998  
 969 structions despite the precision being improved. This 999

appears to be because many samples from the eastern 1000  
 sites with SWLI's > 250 are included and have a greater 1001  
 influence than in the full training set. On the other hand, 1002  
 having established the most appropriate sites to include 1003  
 in the modern training set (i.e the *West* training set) and 1004  
 then running a LW transfer function produces consist- 1005  
 ent predictions alongside an overall improvement in the 1006  
 precision. Notably, the precision of the tidal-flat sam- 1007  
 ples in the core is unaltered while the predictions of the 1008  
 salt-marsh samples is improved by around  $\pm 7$  cm. This 1009  
 is consistent with the better predictive ability of salt- 1010  
 marsh foraminifera and demonstrates the benefit of us- 1011  
 ing a LW transfer function, notwithstanding the impor- 1012  
 tance of establishing a suitable training set by way of a 1013  
 thorough assessment in advance. 1014

#### 1015 4. Conclusions

1016 We have produced a *North Sea* training set of mod- 1017  
 1018 ern foraminifera based on eight previously published 1018  
 sites and one new site (Ythan Estuary, Scotland). The 1019  
 foraminifera cover an elevation gradient from the high- 1020  
 est elevation at which foraminifera occur to mean tide 1021  
 level. The foraminifera display a relationship with el- 1022  
 evation relative to sea level. Foraminifera in marshes 1023  
 in the east (Denmark and Germany) were shown to be 1024  
 very variable, often displaying exceptionally broad ele- 1025  
 vation ranges that may be due to additional environmen- 1026  
 tal factors such as wind build up that is enhanced by the 1027  
 relatively small tidal range. 1028

998 We assessed the effectiveness of a modern regional 1049  
999 training set for reconstructing early Holocene sea level 1050  
1000 at a coastal site in the western North Sea (Ythan Estu- 1051  
1001 ary, Scotland) by dividing the data into different sub- 1052  
1002 regional training sets and by comparing the results of 1053  
1003 parallel analyses. We applied a step-wise approach that 1054  
1004 considered understanding the core lithology and the fos- 1055  
1005 sil samples and understanding of the core alongside the 1056  
1006 modern samples in each of the training sets. In summary 1057  
1007 we used the following approach, which we also recom- 1058  
1008 mend for choosing the most effective transfer function 1059  
1009 and training set in similar studies: 1060

1010 1. We qualitatively assessed how appropriate for re- 1061  
1011 constructing sea level each training set was by ap- 1062  
1012 plying detrended correspondence analysis (DCA) 1063  
1013 and partitioning around medoids (PAM) clustering. 1064  
1014 The results highlighted that clustering of modern 1065  
1015 samples is apparent and is occasionally driven by 1066  
1016 certain taxa, however these are generally rare or 1067  
1017 absent in the fossil record and therefore of lesser 1068  
1018 importance in this example. The modern samples 1069  
1019 from smaller, more localised, training sets show 1070  
1020 a more clearly defined relationship with elevation 1071  
1021 but often lack similarity with fossil samples. Other 1072  
1022 methods of correspondence and cluster analysis are 1073  
1023 available but some form should be included along- 1074  
1024 side fossil samples to enable a truer understanding 1075  
1025 of the suitability of different training sets in each 1076  
1026 context. 1077  
1027 2. The modern analogue technique (MAT) was used 1078  
1028 as a statistical measure of the similarity between 1079  
1029 fossil and modern samples. More localised train- 1080  
1030 ing sets produced fewer 'good' or 'close' modern 1081  
1031 analogues, while larger training sets produce no 1082  
1032 'poor' analogues. This step is almost certainly nec- 1083  
1033 essary to validate that the training set is providing 1084  
1034 a modern environment analogous to that found at 1085  
1035 the reconstructed site for the period of interest and 1086  
1036 should be assessed in conjunction with more qual- 1087  
1037 itative approaches such as in steps 1 and 4. 1088  
1038 3. We ran transfer functions using WA with clas- 1089  
1039 sical deshrinking and WAPLS components 1 or 1090  
1040 2 with cross-validation by bootstrapping and by 1091  
1041 leave-one-site-out (LOSO) to reconstruct palaeo 1092  
1042 marsh surface elevation changes from the Ythan 1093  
1043 core. Precision improved in more localised re- 1094  
1044 gional models although at a loss of predictive 1095  
1045 ability. LOSO cross-validation showed that the 1096  
1046 precision of reconstructions are not unduly over- 1097  
1047 optimistic and that inter-site variability is present 1098  
1048 and likely capture different conditions. 1099

4. We analysed the taxa updates when extra compo-  
nents were added in the WAPLS transfer functions.  
The taxa are occasionally altered beyond the ex-  
tent of the sampled elevation range and caution is  
therefore necessary in choosing these more com-  
plex transfer functions. However, comparison with  
fossil samples showed that the relevant taxa are not  
fundamentally altered in our example. The up-  
dates complement MAT in showing the suitability  
of analogues. This test is a necessary step and  
should be carried out along with knowledge of fos-  
sil samples.

5. We compared all of the reconstructions to assess  
the accuracy. All reconstructions showed a similar  
pattern of surface elevation change, although dif-  
ferences 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 reconstruc-  
tions.

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 gradi-  
ent 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 out-  
performed transfer functions run on randomly gen-  
erated 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 al-  
ways 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 Scotland: 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 salt-marshes: 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 prelude 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 Schmiidl, 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 Schmiidl, G. (2017), 'Applica-

- bility 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., McGinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solyomos, 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, J., 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), 'Re-assessing 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.