Salt-marsh testate amoebae as precise and widespread indicators of sea-level change

Robert L. Barnett*\textsuperscript{a,b}, Thomas L. Newton\textsuperscript{c}, Dan J. Charman\textsuperscript{b}, W. Roland Gehrels\textsuperscript{d}

Author affiliations

\textsuperscript{a} - Coastal Geoscience, Centre d'études Nordiques, Département de biologie, chimie et géographie, Université du Québec à Rimouski, 300 allée des Ursulines, C.P. 3300, Rimouski, Québec G5L 3A1, Canada. \textit{robert.langdon.barnett@gmail.com}

\textsuperscript{b} - Geography, College of Life and Environmental Sciences, University of Exeter, Amory Building, Rennes Drive, Exeter, EX4 4RJ, UK. \textit{d.j.charman@exeter.ac.uk} ; \textit{r.barnett@exeter.ac.uk}

\textsuperscript{c} - School of Geography, Earth and Environmental Sciences, Plymouth University, Drake Circus, Plymouth, Devon, PL4 8AA, UK. \textit{thomas.newton@plymouth.ac.uk}

\textsuperscript{d} - Environment Department, University of York, Heslington, York YO10 5NG, UK. \textit{roland.gehrels@york.ac.uk}

*Corresponding author

e-mail address: \textit{robert.langdon.barnett@gmail.com} ; \textit{r.barnett@exeter.ac.uk} (R.L. Barnett)
Abstract

Salt-marsh sediments are routinely used to reconstruct sea-level changes over past millennia. These reconstructions bridge an important gap between geological and instrumental sea-level records, and provide insights into the role of atmospheric, oceanic, climatic and anthropogenic sea-level drivers, thereby improving understanding of contemporary and future sea-level changes. Salt-marsh foraminifera, diatoms and testate amoebae are three of the proxies capable of accurately reconstructing former sea level over decadal to millennial timescales. Datasets of surface assemblages are collated along elevational gradients to provide modern analogues that can be used to infer former marsh-surface elevations from fossil assemblages. Testate amoebae are the most recently developed proxy and existing studies suggest that they are at least as precise as the two other proxies. This study provides a synthesis of sea-level research using testate amoebae and collates and analyses existing surface datasets of intertidal salt-marsh testate amoebae from sites throughout the North Atlantic. We test the hypothesis that intertidal testate amoebae demonstrate cosmopolitan intertidal zonation across wide geographical areas in a way that is unique to this proxy. Testate amoebae assemblages are harmonised under a unified taxonomy and standardised into a single basin-wide training set suitable for reconstructing sea-level changes from salt-marsh sediments across the North Atlantic. Transfer functions are developed using regression modelling and show comparable performance values to published local training sets of foraminifera, diatoms and testate amoebae. When used to develop recent (last 100 years) sea-level reconstructions for sites in Norway and Quebec, Canada, the testate amoebae-based transfer function demonstrated prediction uncertainties of ±0.26 m and ±0.10 m respectively. These uncertainties equate to 10% and 11% of the tidal ranges at each site, which is of comparable precision to other published sea-level reconstructions based on foraminifera or diatoms. There is great scope for further developing intertidal testate amoebae
as precise sea-level indicators and their application should be tested at sites beyond the North Atlantic.

Keywords

Sea level, testate amoebae, salt marsh, North Atlantic, transfer function

Contents

1. Introduction
2. Salt-marsh testate amoebae
   2.1 Contemporary surface assemblages
   2.2 Environmental controls and elevation
   2.3 Fossil assemblages
3. Sea-level research
   3.1 Training sets and transfer functions
   3.2 Sea-level reconstructions
4. Analytical procedures
5. A North Atlantic regional transfer function
   5.1 Data acquisition and standardisation
   5.2 Unifying the taxonomy
   5.3 The training set
   5.4 Transfer functions
   5.5 Applying the North Atlantic testate amoebae transfer function (NATT)
6. Discussion
7. Conclusions
Acknowledgements
References
1. Introduction

Testate amoebae are a group of unicellular test producing organisms (protists) that inhabit a range of freshwater and brackish environments (Medioli and Scott, 1983) with a wide applicability to palaeoenvironmental research (Tolonen, 1986; Medioli et al., 1999; Charman, 2001). Testate amoebae belong to the kingdom of Protozoa (Cavalier-Smith, 1981) and earlier works have referred to them as ‘rhizopods’ (e.g., Tolonen, 1986), arcellaceans (e.g., Patterson et al., 1985), ‘testaceans’ (e.g., Tolonen et al., 1992) and ‘thecamoebians’ (e.g., Riveiros et al., 2007). The systematic classification of testate amoebae has evolved over past decades (c.f., Charman et al., 2000) and is still subject to revision in light of recent advancements applying multigene molecular phylogeny. Currently, the group has been classified within two dominant clades. Taxa with filiform pseudopodia are in the class Imbricatea, subphylum Filosa, and phylum Cercozoa (Cavalier-Smith, 1998; Cavalier-Smith and Chao, 2003). Taxa with lobed pseudopodia are found predominantly within the order Arcellinida (Kent, 1880; Nikolaev et al., 2005), class Tubulinea (Smirnov et al., 2005), subphylum Lobosea (Carpenter, 1861; Cavalier-Smith, 2009) and phylum Amoebozoa (Lühe, 1913; Smirnov et al., 2011; Cavalier-Smith, 2013).

From a palaeoecological perspective, early studies most commonly documented testate amoebae in peatland (Tolonen, 1966; Warner, 1987, 1989) and lacustrine (Patterson et al., 1985; Medioli and Scott, 1988) environments. However, studies have also described assemblages from coastal sediments subjected to brackish conditions (Medioli et al., 1990), including coastal lakes (e.g., Nicholls and MacIsaac, 2004) and beaches (e.g., Golemansky, 1998a,b), intertidal salt marshes (Scott et al., 1977; Charman et al., 1998, 2002) and mangroves (Duleba and Debenay, 2003). Palaeoenvironmental studies have used testate amoebae as bioindicators of hydrological balance (e.g., Woodland et al., 1998; McGlone and Wilmshurst, 1999), pH (e.g., Mitchell et al., 2013), pollution (e.g., Kandeler et al., 1992),
temperature (Royles et al., 2013) and sea-level changes (e.g., Charman et al., 1998). The recent development of geographically expansive testate amoebae-based hydrological transfer functions ranging from tropical (Swindles et al., 2014), temperate (Li et al., 2015; Amesbury et al., 2016) and high-latitude southern (van Bellen et al., 2014) and northern (Swindles et al., 2015) hemisphere peatland environments demonstrate the wide applicability of testate amoebae as a proxy tool. The application of testate amoebae as precise sea-level indicators in salt-marsh environments commenced during the late 1970s (Scott et al., 1977) but, until recently, has lagged behind alternative proxy counterparts such as salt-marsh macrophytes, foraminifera and diatoms (Shennan et al., 2015).

This review has two main purposes. First, we present an account of existing literature on coastal wetland testate amoebae and their use as proxies of sea-level change. Second, we develop a basin-wide dataset (or ‘training set’) of modern salt-marsh testate amoebae suitable for reconstructing recent relative sea-level changes in the North Atlantic region. Alternative proxies rely on training sets from single sites, or multiple sites in close proximity, to develop transfer functions suitable for local sea-level reconstructions (e.g., Horton and Edwards, 2005). In this paper we demonstrate that assemblages of testate amoebae are broadly similar in many coastal settings around the North Atlantic and, for the first time, we provide a means to reconstruct relative sea-level changes at salt-marsh sites in the entire North Atlantic region using a single training set under a unified taxonomy.

2. Salt-marsh testate amoebae

Surface testate amoebae from salt-marsh environments were first described in Atlantic Canada (Scott et al., 1977; Medioli and Scott, 1983) in the greater than 63 μm sediment
fraction, alongside foraminifera, from estuarine and coastal environments (Scott and Martini, 1982; Scott et al., 1991, 1995, 2001; Barbosa et al., 2006). Whereas foraminifera are found throughout marine and brackish environments, only a small number of testate amoebae taxa are able to tolerate the saline conditions experienced in tidal marshes (c.f., Patterson and Kumar, 2002). Specific testate amoebae assemblages related to tidal influence occupy high salt-marsh environments transitioning into the supratidal zone (Patterson et al., 1985; Charman et al., 1998). Salt-marsh testate amoebae have been sampled along transects across marsh surfaces with the aim of identifying vertical assemblage zones attributable to environmental variables (Charman et al., 1998, 2002; Gehrels et al., 2001, 2006; Riveiros et al., 2007; Ooms et al., 2011, 2012; Barnett et al., 2013, 2016). This practice has been applied regularly using salt-marsh foraminifera (Edwards and Wright, 2015) and diatoms (Zong and Sawai, 2015). However, work based on testate amoebae is comparably limited despite the significant potential of this proxy (Charman, 2015).

2.1 Contemporary surface assemblages

Systematic studies of salt-marsh testate amoebae typically include samples along surface transects which traverse the elevation gradient from the high marsh into the low marsh (Scott and Medioli, 1980). The two marsh zones are commonly characterised by their surface vegetation (e.g., Charman et al., 1998; Gehrels et al., 2006; Barnett et al., 2013). Sampling along such transects must extend beyond highest astronomical tide (HAT) levels and into the supratidal setting in order to recognise assemblage populations that are not affected by tidal inundation (Gehrels et al., 2001, 2006; Ooms et al., 2012). At the other end of the gradient, the lowest testate amoebae are usually encountered near mean high water spring (MHWS) levels (Gehrels et al., 2001, 2006). However, in northern Norway (Barnett et al., 2013),
Testate amoebae populations were found down to near mean high water neap (MHWN) levels while in the Magdalen Islands, in the Gulf of St Lawrence, Canada, where micro-tidal conditions prevail, assemblages were described down to near mean sea level (Barnett et al., 2016). This implies that, in order to capture the full turnover of species data (c.f., Wright et al., 2011), it may be necessary to sample well below the level of MHWS.

Surface data from different locations record site-specific assemblage trends. However, intra-site comparisons regularly reveal common testate amoebae zonations (Charman et al., 2002). For example, an early study of salt-marsh testate amoebae from three sites in the UK identified high-marsh biozones that typically included taxa such as *Tracheleuglypha dentata* and *Trinema* spp. (Charman et al., 2002). The taxon *T. dentata* was subsequently identified in numerous later studies as a ubiquitous high-marsh species that commonly dominated biozones near HAT (Gehrels et al., 2006; Ooms et al., 2011, 2012; Barnett et al., 2013, 2016). Various taxa within the genus *Centropyxis*, perhaps most commonly *Centropyxis cassis* type, are described in most (if not all) salt-marsh studies, frequently demonstrating broad environmental ranges (e.g., Riveiros et al., 2006). Commonly described taxa from lower in the intertidal realm include *Cyphoderia ampulla* (e.g., Charman et al., 2002; Ooms et al., 2011) and *Difflugia pristis* type (e.g., Gehrels et al., 2006; Ooms et al., 2012; Barnett et al., 2013). In contrast to other salt-marsh organisms, such as foraminifera, which display greater site specific differences (Horton et al., 1999), regular occurrences of common taxa could suggest a regionally-robust testate amoebae ecology across the North Atlantic (Charman et al., 2010). In this paper we test, therefore, the hypothesis that intertidal zonation of testate amoebae is similar across the North Atlantic region.

2.2 Environmental controls and elevation
The relative influences of environmental controls on salt-marsh testate amoebae assemblages have been investigated by a number of studies since the early 2000s. A range of variables have been considered, including the pH, salinity, organic content and grain size of the sample sediments whence the assemblages originated (Charman et al., 2002; Riveiros et al., 2007; Ooms et al., 2011, 2012; Barnett et al., 2016). Camacho et al. (2015) went further and analysed nutrient characteristics (total (in-)organic carbon and carbon: nitrogen ratios) as well as dissolved oxygen at sample locations. In all the studies cited above, sample elevation exerted the most significant control on assemblage compositions. Although elevation is not an ecological variable itself (Kemp and Telford, 2015), it is used as a linear approximation of tidal inundation (Gehrels, 2000; Gehrels et al., 2001; Wright et al., 2011; Barlow et al., 2013) due to the close relationship it shares with inundation frequency (e.g., Scott and Medioli, 1980; Horton et al., 1999). Observed correlation between secondary variables and tidal level supports the use of elevation as a surrogate variable capable of capturing the effect that changes in marine influence has on assemblage compositions (Charman et al., 2002).

Gradient analysis is used to statistically describe the relationships between multiple environmental variables and assemblage compositions (Juggins and Birks, 2012). Whether datasets of intertidal testate amoebae exhibit linear responses (e.g., Ooms et al., 2011, 2012; Camacho et al., 2015) or unimodal responses (Charman et al., 2002; Riveiros et al., 2006; Barnett et al., 2016) along environmental gradients, variables associated with tidal inundation (i.e., flooding and elevation) consistently explain a significant proportion of assemblage variation. When quantified, these variables alone may typically account for between 20 and 50% of total species variation within assemblages (Charman et al., 2002; Barnett et al., 2016). Although secondary variables such as pH (e.g., Camacho et al., 2015) and salinity (e.g., Charman et al., 2002; Riveiros et al., 2006) have been documented, the dominance of
elevation in defining assemblage characteristics means that collecting environmental data of alternative variables is not strictly necessary in all cases (e.g., Gehrels et al., 2006).

2.3 Fossil assemblages

When the relationship between sea-level indicators, such as intertidal testate amoebae, and tidal levels is well established in the modern environment, this understanding can be applied to fossil assemblages in order to infer changes in past sea level (van de Plassche, 1986; Shennan, 2007, 2015). As an example, in Scotland (Lloyd, 2000) and eastern Canada (Patterson et al., 1985), fossil testate amoebae from sediment cores were used to describe marine to freshwater transitions in coastal basins which became isolated from tidal conditions as a result of land uplift. A study of different coastal deposits around Britain (isolation basins, coastal back barriers, brackish fens and intertidal marshes) by Roe et al. (2002) revealed that fossil assemblages evolved as local tidal conditions changed at the sites. However, in certain sediment types, notably those from salt marshes, taxa diversity and abundance was found to be low, possibly as a result of test degradation with preferential preservation of certain taxa, or because the depositional environment exceeded the lower limit of testate amoebae occurrence (Roe et al., 2002). An analysis of salt-marsh cores from Wells, Maine, and Chezzetcook, Nova Scotia, found testate amoebae present in good concentrations (c. 5000 to 30,000 tests g\(^{-1}\)) throughout the core tops (Charman et al., 2010). Below depths approximately equivalent to pre-20\(^{th}\) century, test concentrations reduced significantly. As the cores originated from the lower limits of testate amoebae occurrence (i.e., close to MHWS), the lack of test abundance and diversity, as found by Roe et al. (2002), may have been due either to a lower depositional environment or post-burial diagenesis. In addition, these studies, alongside others (e.g., Barnett et al., 2015), repeatedly report the absence of
Euglyphids (siliceous, plate-forming taxa) in fossil assemblages, whilst noting their abundance in contemporary assemblages. Continuous fossil sequences of intertidal testate amoebae spanning the past c. 500 years from Quebec, Canada (Barnett et al., *in review*), and preserved salt-marsh assemblages from Viðarhólmi, Iceland, dating from c. 1200 to 1600 AD (Haynes, 2011), suggest that, under favourable conditions, testate amoebae can be preserved in older coastal sediments. The preferential loss of idiosomic taxa (those formed of siliceous plates) versus that of xenosomes (taxa formed from agglutinated particles) in down-core sediment profiles from terrestrial environments was first recorded several decades ago (Lousier and Parkinson, 1981). The ready deterioration of certain idiosomic taxa, such as *Euglypha* spp. (Patterson and Kumar, 2002), in peatland settings is possibly a result of acidic pore-water conditions (Swindles and Roe, 2007), which may apply to high salt-marsh environments that characteristically have a pH of around 6.0 (e.g., Charman et al., 2002; Barnett et al., 2016).

Despite this, Mitchell et al. (2008) demonstrate that, for *Sphagnum*-dominated peatlands, post-burial changes in fossilised assemblage compositions do not necessarily equate to loss in palaeoecological reconstruction performance. This notion is yet to be tested empirically for salt-marsh testate amoebae based sea-level reconstructions and represents an important facet related to the efficacy of future work.

### 3. Sea-level research

Sea-level reconstructions using salt-marsh organisms rely on sampling contemporary assemblages and measuring their elevation (Scott and Medioli, 1978, 1980). These measurements can be used to define precise indicative meanings, or ranges (c.f., Shennan,
2007, 2015), for surface assemblages based on either a visual assessment or statistical
245 calibration of the data or ‘training set’ (Imbrie and Kipp, 1971; Guilbault et al., 1995). By
246 applying the relationship to down-core assemblages via a transfer function, the
247 contemporaneous elevational range that the fossilised assemblages occupied can be estimated
248 (see Barlow et al. (2013) for a review) and the height of former sea level can be derived with
249 necessary consideration given to correction factors (e.g., palaeo-tidal changes, sediment
250 compaction, local sedimentation regime).

3.1 Training sets and transfer functions

Charman et al. (1998) presented the first training set of salt-marsh testate amoebae designed
to derive indicative meanings related to sea level based on surface assemblage zonation.
Cluster analysis was used to identify four groups of assemblages which defined marsh zones
between HAT and MHWS levels. Following this pioneering study, training sets of surface
data were developed in the United Kingdom (Gehrels et al., 2001; Charman et al., 2002) and
North America (Gehrels et al., 2006). In these cases, indicative meanings were calculated
using a ‘transfer function’ based on regression modelling (ter Braak and Juggins, 1993; Birks,
1995) rather than the previously used cluster analysis or visual assessment methods. The
transfer function approach models the contemporary species-elevation relationship producing
an equation which can be calibrated on species assemblage data to estimate the elevation
from which the assemblage data derives (see Kemp and Telford (2015) for a review). Using
the transfer function method, local training sets of testate amoebae were capable of predicting
salt-marsh surface elevations with sub-decimetre vertical uncertainties (Gehrels et al., 2001,
2006).
The success of using testate amoebae as precise marsh-surface elevation indicators has since been repeated in Belgium (Ooms et al., 2011, 2012), Norway (Barnett et al., 2013) and Canada (Barnett et al., 2016). These studies report transfer functions with $1 \sigma$ uncertainties ranging from $\pm 0.08$ to $\pm 0.24$ m (Table 1). Numerous factors contribute to the magnitude of these model uncertainties. The largest reported uncertainties originate from the Scheldt Estuary, Belgium, where the intertidal marsh is located over 50 km up river and experiences a spring tidal range of over 5 m (Ooms et al., 2012). Salt-marsh based sea-level studies using foraminifera and diatoms also report greater-than-average uncertainties at macro-tidal (4 to 6 m range) sites (Horton and Edwards, 2005; Massey et al., 2006; Hill et al., 2007). Comparing reconstruction uncertainties against tidal range sizes (Table 1, and Table 3 in Barlow et al. (2013)) reveals that transfer function RMSEP values for foraminifera, diatoms and testate amoebae training sets are typically lower than $\sim 10\%$ of the tidal range.

The length of the sampled environmental gradient is also vital for establishing robust and realistic model uncertainties. Wright et al. (2011) show the importance of capturing the full length of species turnover along the gradient of interest. In the case of testate amoebae, this translates to sampling below the lowest limit of test occurrence and beyond the level of tidal influence so that non-intertidal taxa dominate assemblage compositions. Earlier studies, such as Gehrels et al. (2006), postulated that low model uncertainties may be a factor of an under-sampled environmental gradient and insufficient training set size (e.g., Table 1). Assimilating local datasets of intertidal testate amoebae into a regional training set, thereby extending the sampled gradient length, provides an opportunity for developing more robust and accurate sea-level reconstructions.

### 3.2 Sea-level reconstructions
Charman et al. (2010) demonstrated the robustness of salt-marsh testate amoebae by applying a transfer function built from training sets of contemporary assemblages from the United Kingdom (Gehrels et al., 2001; Charman et al., 2002) to reconstruct recent (past c. 100 years) sea-level changes in Maine and Nova Scotia. The two reconstructions based on the UK training set had comparable uncertainties (± 0.07 to ± 0.10 m) to similar reconstructions which were developed using a North American (Gehrels et al., 2006) training set (± 0.05 to ± 0.06 m), thereby demonstrating that regional training sets of surface assemblage data may be applicable across wide geographical areas.

Salt-marsh testate amoebae have also been used to reconstruct falling late Holocene relative sea-level trends in northern Norway (Barnett et al., 2015). However, beyond these few published studies, no other sequences of fossil intertidal assemblages have been used to generate continuous and precise records of former sea level. Ongoing works in the Gulf of St Lawrence (Pascal Bernatchez, pers. comm.) and Newfoundland, Canada, (Andy Kemp, pers. comm.) have begun to incorporate this proxy into sea-level reconstructions where alternative options (such as foraminifera) prove less viable. In parts of eastern Canada, where relative sea level is rising by up to 4 mm yr\(^{-1}\) and intertidal marshes are being submerged (Bernatchez et al., 2008, 2010, 2012), salt-marsh testate amoebae will be capable of capturing environmental transitions from fully terrestrial coastal peatland to intertidal marine conditions. This useful tool will be valuable for quantifying recent rapid relative sea-level rise in locations that lack tide gauges or where tide-gauge records are very short.

4. Analytical procedures
Since early reports of salt-marsh testate amoebae described in samples prepared for foraminifera (Scott et al., 1977; Scott and Martini, 1982), preparation techniques have been honed to maximise the recovery of tests from sediment samples. Importantly, Charman et al. (1998) identified significant testate amoebae diversity in size fractions (in particular, less than 63 μm) which were not included within foraminifera preparations. The latest preparation protocols are detailed by Charman (2015) based on developments of a method developed for peatland testate amoebae (Hendon and Charman, 1998; Charman et al., 2000) and adapted for salt-marsh sediments (Charman et al., 2010; Barnett et al., 2013). In summary, a known volume of sediment is warmed, soaked and disaggregated in water prior to being sieved through 15 and 300 μm meshes. The addition of a chemical pre-treatment (e.g., KOH) stage is optional and can be used to concentrate tests if abundance is low (Hendon and Charman, 1998; Barnett et al., 2013).

A count total of 100 tests per sample is widely cited as a sufficient figure for palaeoecological studies (Woodland et al., 1998; Mitchell et al., 2000; Payne and Mitchell, 2009; Barnett et al., 2015). Where testate amoebae are found in low numbers, such as in the low marsh or in certain sections of sediment cores, researchers have occasionally compiled assemblages based on count totals of fewer than 100 specimens (e.g., Charman et al., 2010). In peatland studies it has been demonstrated that transfer function standard errors may remain consistent for count totals greater than 60 specimens (Payne and Mitchell, 2009). Providing statistically significant counts are reached for each taxon (defined by fractional abundance detection limits; Fatela and Taborda, 2002), these low abundance samples retain significant palaeoecological value. For count totals that are analogous to the sample sizes found in this study, (typically 50 to 200 tests; Supplementary Material I), taxa which make up fewer than 5% of the assemblages are capable of distinguishing environments that differ by 4% at the 95% confidence interval (Patterson and Fishbein, 1989). If greater precision is required, then
count totals of several hundred to several thousand become necessary to identify small
environmental changes in the presence of low abundance taxa. By identifying minimum
count size totals for individual taxa following Fatela and Taborda (2002), those that have
been insufficiently detected to be deemed representative of a particular environment can be
highlighted and, if necessary, removed from the dataset based on their relative abundance
proportion (p) value (c.f., Fatela and Taborda (2002); Supplementary Material I).

5. A North Atlantic regional transfer function

The second purpose of this study was to develop and test a basin-wide training set of salt-
marsh testate amoebae capable of reconstructing sea-level changes in the North Atlantic. To
this end, assemblages from studies throughout the North Atlantic were collated and screened
to ensure that only samples containing consistent test sizes (15 to 300 μm) and suitable count
totals (>50 tests) were incorporated into the dataset. Following the screening process, a total
of 14 sites provided modern assemblage data of testate amoebae with associated
environmental (elevation) information (Fig. 1). There are three sites from Canada (two from
the Magdalen Islands in Quebec (Barnett et al., 2016) and one from New Brunswick (Gehrels
et al., 2006)), three from the USA (Maine, Delaware (Gehrels et al., 2006) and Connecticut
(this study)), one from Iceland (Haynes, 2011), two from Norway (Barnett et al., 2013) and
one from Belgium (Ooms et al., 2012). The remaining four sites are from Wales, Devon,
Norfolk (Charman et al., 2002) and Cornwall (this study) in the UK.

5.1 Data acquisition and standardisation
The original studies provided 275 samples of salt-marsh testate amoebae following removal of assemblages containing count sizes of fewer than 50 tests (Supplementary Material I). As count totals varied from ~50 tests to over 1000 in some cases, proportional uncertainties (p-values) were calculated for each sample following Fatela and Taborda (2002) (Supplementary Material I). The p-value is a function of sample size (n) and represents the minimum proportional abundance required for a given species to have been sufficiently detected at the 95 % confidence limit (equation 1). This value significantly decreases (thereby increasing confidence in the assemblage counts) with increasing count totals.

\[ p = 1 - f(0.05)^{1/n} \] (1) Fatela and Taborda (2002)

Each sample contained a corresponding elevation constraint cited relative to a local water level or datum. All elevation constraints were converted to height above local mean water level (MWL) based on information from the original studies and recent Admiralty Tide Tables (2015). Absolute sampled ranges varied between sites (Fig. 1). Supratidal sampling (i.e., elevations above HAT) occurred at 8 of the total 14 sites. The lowest extent of testate amoebae generally occurs at, or close to, MHWS tidal elevations (n.b., the height of MHWS tides is unavailable for Viðarhólmi in Iceland and Leipsic River in Delaware, USA). Where studies have used two local sites to compile a training set of surf ace data (see Magdalen Islands, Canada, and Vesterålen Islands, Norway; Fig. 1), only a single estimation of the lowest extent of testate occurrence is included.

A standardised water level index (SWLI) was developed in order to account for the difference in tidal ranges between the individual sites and to normalise elevation constraints within the wider dataset (c.f., Gehrels, 1999; Gehrels, 2000; Horton et al., 1999). Using such an index, sample elevations can be expressed relative to common low elevation and high elevation tidal markers (e.g., Wright et al., 2011). Here, sample elevations are normalised by using the
lowest extent of testate amoebae as a consistent low marker and standardising the distance between this and HAT at all the sites (equation 2):

\[
SWLI = \left(\frac{(S - L)}{(HAT - L)}\right) \times 100
\]

where S is sample elevation and L is lowest occurrence of testate amoebae.

5.2 Unifying the taxonomy

There is a wide literature available for identifying testate amoebae. Early compilations of descriptions and monographs began in the late 19\(^{th}\) (Leidy, 1879; Penard, 1890) and early 20\(^{th}\) centuries (Penard, 1902; Cash and Hopkinson, 1905, 1909; Cash et al., 1915). Subsequent taxonomic works have contributed to the reclassification of species and genera, commonly resulting in the renaming and/or splitting of species and morphospecies into distinct clades. Charman et al. (2000) provide a useful insight into some issues associated with the evolving taxonomy of testate amoebae with particular reference to the problems of applying descriptions based on modern specimens to fossil assemblages used in palaeoecological studies. Morphological characteristics used to distinguish living taxa may be unapparent in fossilised specimens. For this reason we advocate the use of ‘type’ groups when developing a consistent and coherent taxonomy for certain salt-marsh specimens. These may include suites of morphospecies which represent a continuum of change rather than distinct morphological features (Medioli and Scott, 1983), or groups of similar taxa which are difficult to distinguish under light microscopy yet share common ecological niches. These methods augment the fundamental systematic classification developed for peatland testate amoebae (Charman et al., 2000) with the increasingly comprehensive taxonomic literature accounting for taxa found in salt-marsh and littoral environments (Golemansky, 1974, 1998a,b; Nicholls, 2003,
Existing published assemblages of intertidal marsh testate amoebae have been catalogued using a range of the available literature (e.g., Charman et al., 1998; Gehrels et al., 2006; Riveiros et al., 2007; Ooms et al., 2011), resulting in some inconsistencies between individual datasets. We applied a single unified taxonomy to harmonise the data. Type groups were used to collate taxa where: i) morphologically indistinguishable species have been reported under different names, ii) specimens show morphological variation along a continuum of change, yet lack clear and discernible distinguishing features, iii) similar taxa (often belonging to the same genus) occupy comparable ecological niches, therefore separation yields no additional palaeoenvironmental information, and, iv) distinguishing features of similar species (or morphospecies) are unable to be identified using light microscopy, particularly for fossil specimens. Of the original 156 taxa reported from the studies mentioned above, 57 taxa or type groups have been used to classify the full dataset under a unified taxonomy (Supplementary Material II). Surface assemblages were then re-classified using the new taxonomy so that a coherent regional training set could be developed. Whilst this has resulted in greater grouping of taxa than is applied in individual studies, it avoids spurious inferences as a consequence of differences between observers.

5.3 The training set

The 275 samples comprising the North Atlantic regional training set were subjected to an unconstrained cluster analysis (Oksanen et al., 2015) which identified five distinguishable biozones (Fig. 2). Although the cluster analysis does not use the elevation data in determining clusters, there was a strong relationship between the clusters defined by assemblage
composition and their elevation, and the biozonation was therefore orientated along an
elevation gradient, as defined by sample SWLI scores (Fig. 2). Low elevation samples are
represented by biozone 1. Here, assemblages derive from near or below MHWS tidal levels
with typical SWLI values of between 3 and 34 (mean ± 1 standard deviation).

Taxonomically, the zone is characterised by a low diversity of halophytic taxa, commonly
dominated by *Arcella catinus-discoides* type and *Centropyxis platystoma* type. Biozone 2 is
typically characterised by different dominating taxa (e.g., *Centropyxiella* type, *Corythionella*
type, *Diffugia pristis* type) and a greater diversity of low abundance taxa (e.g.,
*Pseudocorythion* type, *Pseudohyalosphenia* spp., *Cyphoderia ampulla* type). This zone
represents upper intertidal environments with typical SWLI values between 11 and 84.
Biozone 3 also represents intertidal assemblages from below HAT levels (SWLI 34 to 95)
which predominantly comprise of taxa belonging to the genus *Centropyxis*. Biozone 4
represents transitional assemblages between intertidal and ‘supratidal’ (defined here as
elevations above HAT, although some degree of marine influence remains as a result of, e.g.,
storm events) environments whereas biozone 5 is predominantly comprised of supratidal
samples with SWLI scores of 64 to 175. Both of these zones contain significantly different
assemblages from the lower intertidal biozones and are largely comprised of taxa from the
*Euglypha, Tracheleuglypha* and *Trinema* genera (Fig. 2). A small number of samples from
the top of biozone 5 derive from intertidal, rather than supratidal, environments as
demonstrated by their low SWLI scores. These samples come from the Scheldt Estuary in
Belgium which experiences greater freshwater influence than comparative sites (Ooms et al.,
2011, 2012). This may account for why high marsh taxa such as *Euglypha* spp.,
*Tracheleuglypha* sp. and *Trinema* spp. are found lower in the intertidal realm at this location
(Fig. 2).
Canonical correspondence analysis (CCA; ter Braak, 1986, 1987) was used to test the strength of the relationship between the testate amoebae assemblages and elevation. CCA axis 1 is constrained by sample SWLI scores which exert the primary control on assemblage distributions (Fig. 3). The effects of inconsistently sampled gradients between the 14 sites can be seen in the ordination plots. Certain sites, such as those from Delaware and Belgium, make up disproportionate percentages of certain biozones, such as zones 1 and 5 respectively, which raises the potential issue of spatial autocorrelation in transfer function development (Telford and Birks, 2005, 2011). Despite this, the different biozones and taxa remain clearly distributed along CCA axis 1 with elevation (as a surrogate for tidal inundation) explaining a significant proportion (30%) of the variance, which is comparable to single-site training sets for which elevation typically explains 20 to 30% of total species variance (e.g., Charman et al., 2002; Barnett et al., 2016). The distribution of samples and taxa along the second axis (CA axis 1) is controlled by unmeasured variables and there is the suggestion of an arch-effect (Gauch et al., 1977) deriving from this constrained ordination. We apply a simple unconstrained non-metric multidimensional scaling (NMDS; Minchin, 1987) ordination to determine whether the distribution seen in the CCA analyses may contain such an effect (Fig. 3). The lack of an arch in the NMDS ordination space suggests that the second axes in the canonical ordination (and possibly subsequent axes) is not fully independent from the canonical axis (elevation). This may not be surprising as additional environmental variables (e.g., pH, salinity, surface wetness, nutrient and dissolved oxygen content, and vegetation cover) are likely to demonstrate proportional relationships with our surrogate variable of elevation.

5.4 Transfer functions
Using Detrended Canonical Correspondence Analysis (DCCA; ter Braak and Prentice, 1988), it was possible to estimate overall species response to changes in elevation (Šmilauer and Lepš, 2014). A DCCA score for the North Atlantic training set (3.5 standard deviations units) demonstrates that taxa collectively respond unimodally along the environmental gradient (Birks, 1995). With this in mind, a suite of suitable regression models was applied to develop transfer functions and test the predictive capabilities of the regional training set (Juggins, 2015). Weighted-averaging partial least squares (WAPLS; ter Braak and Juggins 1993) models with between 1 and 5 components were selected as WAPLS regression displays greater performance over alternatives when used alongside training sets with long compositional gradients (ter Braak et al., 1993). As WAPLS with 1 component reduces to a weighted-averaging (WA; ter Braak and Looman, 1986) model with inverse deshrinking (Juggins and Birks, 2012), and combined with the improved performance of WAPLS (ter Braak and Juggins, 1993; ter Braak et al., 1986), it was not necessary to include additional WA models in the analyses. Each WAPLS model was run using bootstrapping (Stine, 1990) and leave-one-site-out (LOSO; Manly, 1997) cross-validation techniques and then re-run following a square root data transformation of the species data to expand the influence of less abundant taxa on transfer function performance. Cross-validated root mean squared errors of prediction (RMSEP) and coefficient of determination ($r^2$) are used as measures of model performance whilst model residuals are critically appraised using average and maximum biases (Juggins and Birks, 2012). The significance of the cross-validated models were assessed using randomisation t-tests (van der Voet, 1994) and illustrated using $p$-values.

Model results (Table 2) show that, due to the negative percentage change in RMSEP scores from incorporating additional components, a simple weighted averaging regression model may be most suitable for the training set. Likewise, the $p$-value results from the significance testing imply that only WAPLS(1), and occasionally WAPLS(2), models are suitable for
predicting marsh-surface elevations. Correlation between observed and predicted SWLI values are highest in the transfer function with bootstrap cross-validation following a square root transformation of the data suggesting that some low abundance taxa are useful indicators of changing marsh elevation. The increased performance when using bootstrapping versus LOSO cross-validation further implies the presence of auto-correlation in the training set. This suggests that transfer functions based on this training set may be suitable for sites with accompanying surface data, yet may lack precision when used to predict marsh-surface elevations at new study sites where surface assemblages remain unexplored.

In an attempt to improve model performance further, outlier removal was carried out on the training set following Edwards et al. (2004). Samples demonstrating residual scores greater than ±1 standard deviation of the environmental gradient (in SWLI units) were identified and removed from the dataset. Outliers typically originated from the gradient ends where the transfer function models under-predicted (at high elevations) or over-predicted (at low elevations) SWLI values (Fig. 4). New regression models were run with the truncated training set based on the best performing models from the original runs (i.e., WAPLS with bootstrapping cross-validation and square root data transformation). All measures of transfer function performance are improved upon by using outlier removal and there is justification for using WAPLS models with 1 component (Fig. 4) for predicting marsh-surface elevation from assemblage data (Table 2).

5.5 Applying the North Atlantic Testate amoebae Transfer function (NATT)

The potential for application of a basin-wide training set to reconstruct local marsh surface elevation changes was explored by Charman et al. (2010) by using a training set from the UK to reconstruct former sea-level height on the east coast of North America. Here, we explore
this prospect further by using the North Atlantic testate amoebae transfer function (hereafter referred to as NATT) to produce two RSL reconstructions using fossil testate amoebae from either side of the Atlantic basin. The purpose of this exercise was to determine if NATT was capable of estimating local RSL changes from recently buried fossil testate amoebae assemblages. The accuracy of the two reconstructions was assessed using nearby tide-gauge data which extend back to the mid-1900s.

Fossil testate amoebae counts used in the reconstructions derive from salt-marsh sediment cores taken in Svinøyosen, Norway (Barnett et al., 2015) and the Magdalen Islands, Canada (Barnett et al. in review). These two sites offer the longest continuous sequences of published fossil salt-marsh testate amoebae for which accompanying surface data are available. Taxonomies used in the original studies were standardised to accord with the salt-marsh based taxonomy developed in this study. A WAPLS transfer function with 1 component following a square root transformation of the species data and outlier removal (see section 5.4) was used to establish palaeomarsh-surface elevation predictions in SWLI units for the fossil assemblages (Table 3). Indicative ranges relative to local MWL were calculated based on SWLI values (equation 2) and converted to estimations of former sea level by subtracting the indicative range from the elevation of the fossil assemblage relative to MWL (Gehrels, 1999). Chronological constraints for the sediment cores were available in the form of $^{210}$Pb derived age-depth profiles from the original studies that allowed the reconstructions to extend back c. 100 years, which coincided with available tide-gauge data (Table 3).

The Vesterålen Islands in northwest Norway currently experience gradual RSL fall (Fig. 5), largely as a result of residual land uplift since deglaciation of the Fennoscandian Ice Sheet (Marthinussen, 1960, 1962). The tide-gauge data from Kabelvåg records this trend, displaying an average linear rate of -1.1 mm yr$^{-1}$ for the period 1948-2015. Our testate amoebae-based RSL reconstruction conforms to the available tide-gauge record and the sea-
level envelope displayed by the proxy data captures a similar trend to the instrumental data (Fig. 5a). Once converted from SWLI units, the transfer function vertical error terms at the Vesterålen Islands are ± 0.26 m (Table 3). This equates to 10% of the spring tidal range at the salt-marsh site, which is 2.6 m (Barnett et al., 2013) and encompasses the inter-annual RSL periodicity seen in the tide-gauge record at Kabelvåg (Fig. 5a).

In contrast to northwest Norway, the Magdalen Islands in eastern Canada experience RSL rise as a result of land subsidence (Peltier et al., 2015) and different oceanographic processes which contribute to rising local sea levels (Dubois and Grenier, 1993). Tide-gauge data from Cap-aux-Meules are only available for intermittent periods over the past few decades but show approximately 0.2 m of RSL rise since the 1960s (Fig. 5b). The testate amoebae-based reconstruction captures this rising trend and demonstrates the value of this method in providing sea-level constraints where instrumental data are not available. The proxy data display acceleration in the rate of RSL rise during the mid-20th century but direct comparisons with the tide-gauge record are restricted by the record length. A wider look at tide gauge RSL trends along the east coast of North America reveals that many locations experienced a peak in rising RSL trends during the 1930s and 1940s, including the more northerly gauges at Boston, Halifax, Nova Scotia, and Charlottetown, Prince Edward Island (Boon, 2012; Holgate et al., 2013; PSMSL, 2016). The local reconstruction from the Magdalen Islands may reflect these regional sea-level trends suggesting that NATT is capable of accurately predicting local RSL changes from fossil testate amoebae preserved in salt-marsh sediments.

The transfer function error terms at the Magdalen Islands are ± 0.10 m (Table 3) following conversion from SWLI units, which represents 11% of the spring tidal range (0.9 m) at the reconstruction site (Barnett et al., 2016). As SWLI is a function of tidal range, sites with a larger tidal range are assigned greater vertical uncertainty terms following application of
NATT to a local reconstruction. This suggests that the universal transfer function may be valid at a range of sites with varying tidal range sizes as it may help prevent overly optimistic prediction errors. The two reconstructions shown here suggest that, over the recent past where comparable tide-gauge data are available for validation, a basin-wide training set of salt-marsh testate amoebae may be used to reconstruct local RSL changes from fossil assemblages using a universally applicable transfer function.

6. Discussion

Our review of testate amoebae occurrences in salt marshes around the North Atlantic shows that these organisms are excellent indicators of sea level. When applied to short cores, the transfer functions derived from the modern distribution of testate amoebae produce sea-level reconstructions that are in good agreement with tide-gauge records. The two example reconstructions contain prediction uncertainties of ± 0.10 and 0.26 m. These RMSEP values are comparable to transfer function uncertainties derived from foraminifera and diatom training sets from across the globe (Table 3 in Barlow et al., 2013). The most precise regression models available using these other proxies have RMSEP values of ± 0.04 to 0.06 m (Gehrels et al., 2005; Southall et al., 2006; Kemp et al., 2009), which are similar to best performing models using local testate amoebae data (± 0.07 m; Table 1). Although reconstruction uncertainties for the Magdalen Islands and Norway using NATT (± 0.10 and 0.26 m) are greater than locally derived RMSEP values (± 0.08 and 0.09 m; Table 1), the regional training set encompasses greater ecological diversity across a longer gradient length. This contributes to a more robust transfer function as it is more likely that fossil assemblages are represented by modern counterparts (e.g., Horton and Edwards, 2005; Watcham et al.,...
RMSEP values from locally derived training sets commonly demonstrate decimetre uncertainty precisions (Table 1). Local transfer functions from sites that experience macro-tidal conditions tend to have greater uncertainty ranges. In the case of reconstructing former sea level at the Magdalen Islands (a micro-tidal region), decimetre level precision remains achievable when using the basin wide training set of NATT.

As NATT is a regional transfer function based on multiple sites it avoids over-optimistic RMSEP values arising from possible autocorrelation within training sets from individual sites. The suite of transfer functions that use LOSO cross-validation (Table 2) suggests that the training set suffers from some spatial autocorrelation as the transfer functions demonstrate weaker performance when used to predict surface elevations at a given site after local data are omitted from the training set (Telford and Birks, 2005). However, over-optimistic performance values, which are often present in auto-correlated datasets (Telford and Birks, 2009), are avoided due to the more robust prediction uncertainties that NATT offers.

An additional uncertainty which is introduced by building a regional training set of surface data is the imposed presence of uneven sampling along the environmental gradient (Telford and Birks, 2011). Local training sets are sampled along surface transects at regular vertical intervals which ensures even sampling of the local environmental gradient (elevation). However, due to a lack of standardised practice, and many surface samples being collected originally for studies on foraminifera (e.g., Gehrels et al., 2001, 2006), there is significant disparity in the concentration and range of samples at different elevations when the local datasets are combined (Fig. 1). A segment-wise RMSEP procedure (Telford and Birks, 2011) was used to investigate the effect of an unevenly sampled gradient within NATT. The full training set was divided into bins of 25.5 SWLI units which represented the standardised sampled range (207 SWLI) divided by the mean standardised sampling resolution (8.1 SWLI).
of all the sites, and RMSEP values calculated for each segment using the best performing transfer function (Fig. 6).

The concentration of samples between 0 and 100 SWLI represents the intertidal proportion of the training set (Fig. 6). Above HAT (100 SWLI) the number of samples per bin decreases as different local datasets terminate near or slightly above this tidal level. Transfer function performance remains robust (i.e., comparable or lower RMSEP values to overall performance) for the bins that represent intertidal and transitional environments. Bins beyond 150 SWLI all contain segment-wise RMSEP values that exceed overall model uncertainty. This suggests (rather unsurprisingly) that supratidal testate amoebae assemblages above 150 SWLI are poor at predicting marsh surface elevations, because of the diminished influence of tidal inundation relative to alternative environmental controls. Despite the unevenly sampled gradient (i.e., variable sample numbers per bin), transfer function performance for intertidal and transitional assemblages appears reliable. It is likely that vertical zonation of testate amoebae remains apparent beyond HAT as some degree of marine influence remains.

However, if fossil assemblages indicate an elevation that exceeds 150 SWLI, and the provenance of the assemblages cannot be clearly ascertained, then they should be used as limiting sea-level index points in sea-level reconstructions (Shennan et al., 2015) rather than being assigned a precise (but possibly inaccurate) indicative meaning.

As transfer function uncertainties approach the sampling resolution of this reconstruction method (classically c. 0.05 m), improvements in precision become hard to find. One technique applied recently has been to use stable carbon isotope geochemistry (Lamb et al., 2006; Khan et al., 2015) to identify whether fossil salt-marsh samples originate from above or below the level of mean higher high water (Kemp et al., 2010). These additional data can then be used to ‘trim’ sea-level index points which have vertical uncertainties based on alternative proxies, such as foraminifera (Kemp et al., 2013), in order to shrink reconstruction
uncertainty boxes. The development of multi-proxy sea level reconstructions from salt marsh sediments has been investigated in the past by pairing diatom and foraminifera data (Kemp et al., 2009), foraminifera and testate amoebae data (Barnett et al., 2016), and even data of all three proxies (Gehrels et al., 2001), and comparing transfer function performance against that of single-proxy models. In general, multi-proxy transfer functions usually improve on the standalone performance of the weaker proxy, but fail to increase precision in comparison to the stronger proxy (Barnett et al., 2016). Different proxy datasets might be conjoined to improve transfer function robustness, or additional proxies may be developed into independent reconstructions that can assist in validation, both of which have positive contributions to sea-level studies. The approach of trimming sea-level index point vertical uncertainties by developing independent reconstructions from multiple proxies and limiting the size of index points to common indicative ranges is yet to be tested empirically and is not widely encouraged. Recent advancements in modelling sea-level envelopes which incorporate full horizontal (temporal) and vertical (spatial) uncertainties (Cahill et al., 2015a; Parnell and Gehrels, 2015; Lawrence et al., 2016) likely represent a more holistic and robust approach. The continued development of transfer function models that are capable of incorporating multiple proxies with differing uncertainties (Cahill et al., 2015b) are also likely to aid reconstruction attempts in the future.

To date, intertidal testate amoebae have been used to identify the isolation of coastal marine basins from oceans (Patterson et al., 1985; Lloyd, 2000; Roe et al., 2002) and define precise estimations of past sea level in regions that have experienced RSL fall during the late Holocene (Barnett et al., 2015). They are also proven precise sea-level indicators in regions experiencing RSL rise (e.g., Charman et al., 2010), and it is here where they might be most valuably applied in sea-level studies. Surface datasets of intertidal testate amoebae show narrow vertical zonation across marsh environments that encompass MHWS tide levels,
HAT, and the supratidal zone (Gehrels et al., 2006; Ooms et al., 2012; Barnett et al., 2016; this study). The indicative ranges of these assemblages are consistently narrow and can be used to predict salt-marsh surface elevations with relatively high precision. In addition, assemblages originating from near, or slightly above, HAT can be readily identified by the high abundance of *Trinema* spp. and occasional presence of e.g., *Heleopera* spp. and *Nebela* spp. Testate amoebae from fossil sediments that record transitions from coastal terrestrial to marine intertidal environments can, therefore, provide a unique opportunity to investigate coastline submergence rates related to recent rapid RSL rise (e.g., Kopp et al., 2016).

7. Conclusions

This study has provided a summary of the use of testate amoebae in sea-level research since the late 1970s. Having been originally identified in studies targeting foraminifera (e.g., Scott et al., 1977), they have since been developed into a unique and valuable proxy of their own. Their early use as precise sea-level indicators (Charman et al., 1998, 2002) demonstrated capabilities equal to, or better than, alternative proxies such as foraminifera or diatoms (Gehrels et al., 2001). More recently, developments in applying salt-marsh testate amoebae in sea-level reconstructions supported the notion that surface assemblages were cosmopolitan across wide geographical areas and that training sets from the UK were applicable to reconstructions from North America (Charman et al., 2010).

To explore this notion further, this study has collated the available surface datasets of intertidal salt-marsh testate amoebae from the North Atlantic (Supplementary Material I) and harmonised them under a single unified taxonomy. This taxonomy comprises 57 different taxa or type groups (Supplementary Material II) and accounts for each specimen identified in
the original datasets and used in this study. The further development of type groups (c.f., Charman et al., 2000) has been used to reduce ambiguity between studies by grouping morphologically similar taxa that occupy similar ecological niches. The taxonomy is valuable for future studies on salt-marsh testate amoebae and is able to evolve in order to accommodate additional intertidal taxa that are likely to arise as research progresses beyond the North Atlantic.

The training set, comprising of 275 samples from 14 different sites, encompasses an environmental gradient extending from below MHWS tide level and up into the supratidal zone. The assemblages were portioned into five broad groups, which zoned in relation to elevation, using unconstrained cluster analysis. Three biozone groups represented intertidal environments between the lowest extent of testate amoebae and HAT, a single biozone represented a transitional group occurring across the boundary of HAT and the remaining biozone represented a predominantly supratidal group.

Following standardisation of the datasets to account for differences in the tidal range between sites (expressed as a standardised water level index - SWLI), a comparison of several transfer functions revealed that a WAPLS regression model with 1 or 2 components following square root species data transformation and outlier removal was suitable for predicting marsh-surface elevations using the composite training set. These transfer functions had RMSEP values of ± 27 SWLI units, which represented 13% of the sampled environmental gradient (207 SWLI units), a precision comparable with transfer function performances based on local training sets (Table 1). When used to reconstruct recent sea-level histories from Norway and Quebec, Canada, the precision of the transfer function yielded site specific RMSEP values of ± 0.26 m and ± 0.10 m respectively, which equated to 10% and 11% of the tidal ranges at each site, again comparable to precisions based on alternative salt-marsh proxies and more localised training sets.
The North Atlantic testate amoebae-based transfer function developed in this study provides evidence for the cosmopolitan nature of salt-marsh testate amoebae and confirms that the relationship between sea level and assemblage composition is similar across the whole of the region. Such a wide-scale regional training set captures greater ecological diversity in comparison to local datasets. This contributes to more robust reconstructions when used to estimate former sea-level positions. Testate amoebae can also be valuable in sea-level studies when other proxies are missing. In the case of foraminifera, testate amoebae can be used to extend the sampled environmental gradient beyond the highest limit of foraminifera occurrence and potentially into the supratidal realm. This expands the elevational range of environments that a transfer function can be used to reconstruct, allowing the reconstruction of larger amplitude sea-level changes or possibly allowing the extension of reconstructions further back in time. To date, the exploratory use of testate amoebae as precise sea-level indicators has been predominantly focussed in the North Atlantic. There is significant scope for the expansion and development of this proxy, which may represent one of the most precise indicators available for reconstructing pre-industrial sea-level changes.

Acknowledgements

Our work on salt-marsh testate amoebae was funded by the Natural Environment Research Council (grant GR9/03426 to WRG and DJC) and by Plymouth University studentships (to RLB and TLN). Datasets from the Magdalen Islands were originally collected using funding from the Coastal Geoscience Research Chair at the Université du Québec à Rimouski. We acknowledge Helen Roe (Queen’s University Belfast) for her contributions to our early studies of salt-marsh testate amoebae and Stijn Temmerman and Louis Beyens (University of
Antwerp) for their assistance during data collation. Thanks go to Graeme Swindles (University of Leeds) and Tim Daley (Plymouth University) for their constructive reviews, which enhanced the quality of the manuscript.

References


Sensibilité des côtes et vulnérabilité des communautés du golfe du Saint-Laurent aux impacts des changements climatiques. Laboratoire de dynamique et de gestion intégrée des zones côtières, Université du Québec à Rimouski. Rapport de recherche remis au Consortium OURANOS et au FACC.


Cahill, N., Kemp, A.C., Horton, B.P., Parnell, A.C. 2015b. A Bayesian hierarchical model for reconstructing sea level: from raw dates to rates of change. Climate of the Past Discussions, 11, 4851-4893.


Kemp, A.C., Horton, B.P., Vane, C.H., Bernhardt, C.E., Corbett, D.R., Engelhart, S.E., Anisfeld, S.C., Parnell, A.C., Cahill, N. 2013. Sea-level change during the last 2500 years in New Jersey, USA. Quaternary Science Reviews, 81, 90–104.


Nicholls, K.H. 2009. A multivariate statistical evaluation of the “acolla-complex” of
*Corythionella* species, including a description of *C. darwini* n. sp. (Rhizopoda: Filosea,


The testate lobose amoebae (order Arcellinida Kent, 1880) finally find their home within

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Simpson,
Package. R package version 2.3-2. [http://cran.r-project.org/package=vegan](http://cran.r-project.org/package=vegan).

Ooms, M., Beyens, L., Temmerman, S. 2011. Testate amoebae as estuarine water-level
indicators: modern distribution and the development of a transfer function from a
freshwater tidal marsh (Scheldt estuary, Belgium). *Journal of Quaternary Science*, 26(8),
819-828.

Ooms, M., Beyens, L., Temmerman, S. 2012. Testate amoebae as proxy for water level

Parnell, A.C., Gehrels, W.R. 2015. Using chronological models in late Holocene sea-level
(Eds.), *Handbook of Sea-Level Research*. Wiley, Chichester, UK. 500-513.

(Thecamoebians) in small lakes of New Brunswick and Nova Scotia: modern distribution

Patterson, R.T., Fishbein, A. 1989. Re-examination of the statistical methods used to
determine the number of point counts needed for micropaleontological quantitative

Patterson, R.T., Kumar, A. 2002. A review of current testate rhizopod (thecamoebian)


Table and Figure Captions

Table 1
Published transfer function performance criteria of salt-marsh testate amoebae based training sets from the North Atlantic. Model prediction uncertainties (RMSEP values) are also given as a percentage of local tidal range size and sampled environmental gradient.

Table 2
Transfer function performance statistics for the regional North Atlantic testate amoebae based training set. Weighted averaging partial least squares (WAPLS) models were used with between 1 and 5 components. Both bootstrapping and leave-one-site-out cross validation methods were used on the full dataset following a square root transformation of the species data. The training set was also subjected to outlier removal before re-running of the best performing transfer function models (bootstrapping with square root species data transformation). The model used to apply the North Atlantic testate amoebae training set is shown in bold print.

Table 3
Sea-level reconstruction criteria for sediment cores from the Vesterålen Islands, Norway (Barnett et al., 2015), and the Magdalen Islands, Canada (Barnett et al., in prep). Given are palaeo-marsh surface elevation predictions from the regional North Atlantic transfer function (‘NATT’) in SWLI units and local elevations following a conversion to account for
differences in tidal range. Also given are corresponding chronological data based on $^{210}$Pb based age-depth profiles from the original studies.

**Figure 1**

Map showing locations of the original datasets used in this study (a). Also shown are the distributions of surface samples taken from each location in relation to standardised tidal regimes (b). The purple circle indicates the lowest extent of salt-marsh testate amoebae, the red circle at 100 SWLI units represents the level of highest astronomical tides and the blue projections mark elevations below mean high water tides for each location.

**Figure 2**

Composite assemblage data for the North Atlantic training set of salt-marsh testate amoebae. SWLI units represent standardised sample elevations where 0 = the lowest extent of testate amoebae and 100 = highest astronomical tide level. Also shown are biozone clusters which result from unconstrained cluster analysis of the entire dataset. Taxa abbreviations correspond with taxonomic descriptions provided in Appendix I.

**Figure 3**

Canonical correspondence analysis of the North Atlantic testate amoebae based training set showing individual sample scores by location (top left), sample scores grouped according to unconstrained cluster analysis (bottom left) and individual species scores (top right). Also shown is the unconstrained NMDS ordination of the dataset (bottom right) which demonstrates the potential arch effect displayed by the CCA ordinations.
Model predicted versus measured SWLI elevation values and model residuals following weighted averaging partial least squares (WAPLS) regression with 1 component using bootstrapping cross-validation following square root transformation of the species data and outlier removal. A loess smoothing is shown in red (top plot) for comparison against the identity line (1:1) and samples omitted following outlier removal are highlighted in the bottom plot.

Sea-level reconstructions based on criteria given in Table 3 for two locations using the regional North Atlantic testate amoebae based transfer function (‘NATT’). Grey boxes bound vertical uncertainties based on model RMSEP values and horizontal chronological uncertainties (see text and Table 3 for details). Also shown are available tide-gauge data from nearby stations in blue.

Segment-wise cross-validation of the best performing transfer function (WAPLS1 with bootstrapping cross-validation following square root species data transformation and outlier removal) showing RMSEP values for consecutive 25.5 SWLI unit bins (green). Also shown is a sample frequency distribution based on SWLI scores (blue). The overall RMSEP value for the transfer function is given in red.
### Tables and Figures

#### Table 1

<table>
<thead>
<tr>
<th>Training set</th>
<th>Reference</th>
<th>Spring tidal range (m)</th>
<th>Number of samples</th>
<th>Model type</th>
<th>Reported RMSEP (m)</th>
<th>Elevation range of samples (m)</th>
<th>RMSEP/tidal range (%)</th>
<th>RMSEP/elevation range of samples (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK (regional)</td>
<td>Gehrels et al. (2001)</td>
<td>-</td>
<td>52</td>
<td>WA</td>
<td>0.08 (norm)</td>
<td>0.35 (norm)</td>
<td>-</td>
<td>15.2</td>
</tr>
<tr>
<td>Brancaster, UK</td>
<td>&quot;</td>
<td>6.0</td>
<td>-</td>
<td>&quot;</td>
<td>0.25 (local)</td>
<td>-</td>
<td>4.2</td>
<td>-</td>
</tr>
<tr>
<td>Erme, UK</td>
<td>&quot;</td>
<td>4.7</td>
<td>-</td>
<td>&quot;</td>
<td>0.17 (local)</td>
<td>-</td>
<td>3.6</td>
<td>-</td>
</tr>
<tr>
<td>Taf, UK</td>
<td>&quot;</td>
<td>6.6</td>
<td>-</td>
<td>&quot;</td>
<td>0.28 (local)</td>
<td>-</td>
<td>4.2</td>
<td>-</td>
</tr>
<tr>
<td>New Brunswick, Canada</td>
<td>Gehrels et al. (2006)</td>
<td>5.8</td>
<td>12</td>
<td>WA</td>
<td>0.08</td>
<td>0.75</td>
<td>1.3</td>
<td>10.7</td>
</tr>
<tr>
<td>Maine, USA</td>
<td>Gehrels et al. (2006)</td>
<td>2.6</td>
<td>17</td>
<td>WA</td>
<td>0.07</td>
<td>1.05</td>
<td>2.7</td>
<td>6.7</td>
</tr>
<tr>
<td>Delaware, USA</td>
<td>Gehrels et al. (2006)</td>
<td>1.8</td>
<td>9</td>
<td>WA</td>
<td>0.07</td>
<td>0.14</td>
<td>3.9</td>
<td>50</td>
</tr>
<tr>
<td>Scheldt estuary, Belgium</td>
<td>Ooms et al. (2012)</td>
<td>5.5</td>
<td>37</td>
<td>WAPLS</td>
<td>0.24</td>
<td>1.88</td>
<td>4.4</td>
<td>12.8</td>
</tr>
<tr>
<td>Vesterålen Islands, Norway</td>
<td>Barnett et al. (2013)</td>
<td>2.6</td>
<td>29</td>
<td>WA</td>
<td>0.09</td>
<td>0.76</td>
<td>3.5</td>
<td>11.8</td>
</tr>
<tr>
<td>Magdalen Islands, Canada</td>
<td>Barnett et al. (2016)</td>
<td>0.6</td>
<td>62</td>
<td>WA</td>
<td>0.08</td>
<td>0.69</td>
<td>13.3</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>WAPLS(1)</td>
<td>WAPLS(2)</td>
<td>WAPLS(3)</td>
<td>WAPLS(4)</td>
<td>WAPLS(5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bootstrapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEP</td>
<td>37.73</td>
<td>37.75</td>
<td>38.64</td>
<td>39.72</td>
<td>40.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change</td>
<td>-</td>
<td>-0.05</td>
<td>-2.36</td>
<td>-2.80</td>
<td>-2.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.52</td>
<td>0.55</td>
<td>0.55</td>
<td>0.54</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave.Bias</td>
<td>-0.32</td>
<td>-0.34</td>
<td>-0.92</td>
<td>-1.06</td>
<td>-1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.Bias</td>
<td>75.76</td>
<td>72.55</td>
<td>70.62</td>
<td>67.65</td>
<td>65.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.001</td>
<td>0.116</td>
<td>0.561</td>
<td>0.839</td>
<td>0.843</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LOSO</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEP</td>
<td>43.39</td>
<td>49.70</td>
<td>51.69</td>
<td>52.79</td>
<td>53.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change</td>
<td>-</td>
<td>-14.54</td>
<td>-4.00</td>
<td>-2.13</td>
<td>-0.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.38</td>
<td>0.30</td>
<td>0.29</td>
<td>0.30</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave.Bias</td>
<td>-3.39</td>
<td>-6.65</td>
<td>-9.52</td>
<td>-10.45</td>
<td>-10.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.Bias</td>
<td>83.40</td>
<td>84.90</td>
<td>84.44</td>
<td>81.14</td>
<td>78.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.001</td>
<td>1.000</td>
<td>1.000</td>
<td>0.994</td>
<td>0.808</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SQRT.Bootstrapping</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEP</td>
<td>36.72</td>
<td>35.99</td>
<td>36.61</td>
<td>37.74</td>
<td>38.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change</td>
<td>-</td>
<td>1.99</td>
<td>-1.72</td>
<td>-3.09</td>
<td>-2.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.55</td>
<td>0.59</td>
<td>0.59</td>
<td>0.59</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave.Bias</td>
<td>-0.60</td>
<td>-0.12</td>
<td>-0.51</td>
<td>-0.36</td>
<td>-0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.Bias</td>
<td>73.40</td>
<td>62.29</td>
<td>56.26</td>
<td>54.52</td>
<td>53.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.001</td>
<td>0.057</td>
<td>0.534</td>
<td>0.869</td>
<td>0.868</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SQRT.LOSO</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEP</td>
<td>42.33</td>
<td>48.55</td>
<td>51.95</td>
<td>55.18</td>
<td>58.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change</td>
<td>-</td>
<td>-14.69</td>
<td>-7.00</td>
<td>-6.22</td>
<td>-5.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.41</td>
<td>0.35</td>
<td>0.34</td>
<td>0.32</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.Bias</td>
<td>82.41</td>
<td>77.63</td>
<td>71.23</td>
<td>69.95</td>
<td>69.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.001</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SQRT.Bootstrapping with outlier removal</strong> (Edwards et al., 2004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEP</td>
<td><strong>27.26</strong></td>
<td>26.52</td>
<td>26.94</td>
<td>28.01</td>
<td>28.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change</td>
<td>-</td>
<td>2.71</td>
<td>-1.58</td>
<td>-3.97</td>
<td>-3.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td><strong>0.7</strong></td>
<td>0.73</td>
<td>0.73</td>
<td>0.72</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave.Bias</td>
<td><strong>-0.95</strong></td>
<td>-0.42</td>
<td>-0.64</td>
<td>-0.49</td>
<td>-0.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.Bias</td>
<td><strong>39.04</strong></td>
<td>38.15</td>
<td>41.74</td>
<td>44.59</td>
<td>43.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td><strong>0.001</strong></td>
<td>0.045</td>
<td>0.514</td>
<td>0.977</td>
<td>0.981</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Depth in core (m)</td>
<td>( h ) - Elev. to msl (m)</td>
<td>Indicative elev. (SWLI)</td>
<td>RMSEP (SWLI)</td>
<td>RMSEP (local m)</td>
<td>( i ) - Indicative elev. to msl (m)</td>
<td>Sea level (m) ((s = h - i))</td>
<td>(^{210}\text{Pb year (CRS model)})</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Vest 1</td>
<td>0.003</td>
<td>1.468</td>
<td>72.786</td>
<td>27.181</td>
<td>0.261</td>
<td>1.689</td>
<td>0.261</td>
<td>2007.7</td>
</tr>
<tr>
<td>Vest 2</td>
<td>0.008</td>
<td>1.463</td>
<td>63.912</td>
<td>27.156</td>
<td>0.261</td>
<td>1.604</td>
<td>0.261</td>
<td>2004.2</td>
</tr>
<tr>
<td>Vest 3</td>
<td>0.018</td>
<td>1.453</td>
<td>59.672</td>
<td>27.150</td>
<td>0.261</td>
<td>1.563</td>
<td>0.261</td>
<td>1997.6</td>
</tr>
<tr>
<td>Vest 4</td>
<td>0.028</td>
<td>1.443</td>
<td>53.057</td>
<td>27.190</td>
<td>0.261</td>
<td>1.499</td>
<td>0.261</td>
<td>1989.7</td>
</tr>
<tr>
<td>Vest 5</td>
<td>0.038</td>
<td>1.433</td>
<td>72.235</td>
<td>27.219</td>
<td>0.261</td>
<td>1.683</td>
<td>0.261</td>
<td>1980.7</td>
</tr>
<tr>
<td>Vest 6</td>
<td>0.048</td>
<td>1.423</td>
<td>36.699</td>
<td>27.442</td>
<td>0.263</td>
<td>1.342</td>
<td>0.263</td>
<td>1972.4</td>
</tr>
<tr>
<td>Vest 7</td>
<td>0.058</td>
<td>1.413</td>
<td>62.468</td>
<td>27.203</td>
<td>0.261</td>
<td>1.590</td>
<td>0.261</td>
<td>1961.2</td>
</tr>
<tr>
<td>Vest 8</td>
<td>0.078</td>
<td>1.393</td>
<td>35.637</td>
<td>27.374</td>
<td>0.263</td>
<td>1.332</td>
<td>0.263</td>
<td>1927.9</td>
</tr>
<tr>
<td>Mag 1</td>
<td>0.005</td>
<td>0.225</td>
<td>43.767</td>
<td>27.150</td>
<td>0.096</td>
<td>0.272</td>
<td>0.096</td>
<td>2004.5</td>
</tr>
<tr>
<td>Mag 2</td>
<td>0.015</td>
<td>0.215</td>
<td>45.087</td>
<td>27.354</td>
<td>0.096</td>
<td>0.277</td>
<td>0.096</td>
<td>1995.3</td>
</tr>
<tr>
<td>Mag 3</td>
<td>0.025</td>
<td>0.205</td>
<td>57.714</td>
<td>27.242</td>
<td>0.096</td>
<td>0.321</td>
<td>0.096</td>
<td>1986.6</td>
</tr>
<tr>
<td>Mag 4</td>
<td>0.035</td>
<td>0.195</td>
<td>77.284</td>
<td>27.245</td>
<td>0.096</td>
<td>0.390</td>
<td>0.096</td>
<td>1977.7</td>
</tr>
<tr>
<td>Mag 5</td>
<td>0.045</td>
<td>0.185</td>
<td>79.089</td>
<td>27.140</td>
<td>0.096</td>
<td>0.396</td>
<td>0.096</td>
<td>1968.2</td>
</tr>
<tr>
<td>Mag 6</td>
<td>0.055</td>
<td>0.175</td>
<td>79.541</td>
<td>27.139</td>
<td>0.096</td>
<td>0.398</td>
<td>0.096</td>
<td>1960.5</td>
</tr>
<tr>
<td>Mag 7</td>
<td>0.065</td>
<td>0.165</td>
<td>79.846</td>
<td>27.158</td>
<td>0.096</td>
<td>0.399</td>
<td>0.096</td>
<td>1948.9</td>
</tr>
<tr>
<td>Mag 8</td>
<td>0.075</td>
<td>0.155</td>
<td>77.067</td>
<td>27.265</td>
<td>0.096</td>
<td>0.389</td>
<td>0.096</td>
<td>1931.9</td>
</tr>
<tr>
<td>Mag 9</td>
<td>0.085</td>
<td>0.145</td>
<td>74.505</td>
<td>27.313</td>
<td>0.096</td>
<td>0.380</td>
<td>0.096</td>
<td>1910.1</td>
</tr>
</tbody>
</table>
Figure 1
Figure 4
Figure 5
(a) Vesterålen Islands (Svinøysen), NORWAY
(b) Magdalen Islands (Bassin), CANADA

Figure 6
WAPLS(1)_SQRT_Bootstrapping

0 20 40 60 80 100 SWLI
0 10 20 30 40 50 RMSEP (SWLI)

Frequency

36.72