Salt marshes as late Holocene tide gauges

Barlow, Natasha L.M. [1]*

Shennan, Ian [1]

Long, Antony J. [1]

Gehrels, W. Roland [2]

Saher, Margot H. [2]

Woodroffe, Sarah A. [1]

Hillier, Caroline [1,3]

[1] Sea Level Research Unit, Department of Geography, Durham University, South Road, Durham, DH1 3LE

[2] School of Geography, Earth & Environmental Sciences, Plymouth University, Drake Circus, Plymouth, PL4 8AA

[3] Current address: Argus Ecology, Annfield Plain, County Durham, DH9 7XN

*Corresponding author: n.l.m.barlow@durham.ac.uk

Abstract

Understanding late Holocene to present relative sea-level changes at centennial or sub-centennial scales requires geological records that dovetail with the instrumental era. Salt marsh sediments are one of the most reliable geological tide gauges.

In this paper we review the methodological and technical advances that promoted research on 'high resolution' late Holocene sea-level change. We work through an example to demonstrate different pathways to quantitative reconstructions of relative sea level based on salt marsh sediments. We demonstrate that any reconstruction is in part a result of the environment from which the record is taken, the modern dataset used to calibrate the fossil changes, statistical assumptions behind calibrating microfossil assemblages and choices made by the researchers. With the error term of typical transfer function models ~10-15% of the tidal range, micro-tidal environments should produce the most precise sea-level reconstructions. Sampled elevation range of the modern dataset also has a strong influence on model predictive ability. Model-specific errors may under represent total uncertainty which comes from field practices, sedimentary environment, palaeo-tidal changes and sediment compaction as well as statistical uncertainties. Geological tide gauges require a detailed chronology but we must be certain that apparent relative sea-level fluctuations are not simply a consequence of an age-depth model.

We make six suggestions to aid the development and interpretation of geological tide gauge records.

1. Introduction

Comparisons of secular trends in sea level from the twentieth century and late Holocene period (determined from instrumental tide gauge and geological data respectively) from the North Atlantic region, reveal that the instrumental tide gauge measured trends of mean sea-level are systematically larger than the long-term sea-level trends (Engelhart et al., 2009; Gehrels et al., 2004; Shennan and Horton, 2002; Shennan and Woodworth, 1992; Woodworth et al., 1999). Additional analyses incorporating continuous GPS and absolute gravity estimates reinforce these observations in Britain (Teferle et al., 2009). Around the UK, instrumental tide gauge trends for 1901 onwards are ~1.4 ± 0.2 mm yr^{-1} larger than those inferred from geology or geodetic methods (Woodworth et al., 2009), suggesting a regional sea-level rise of climate change origin several one-tenths of mm per year lower than Church and White's 1.7 ± 0.2 mm yr^{-1} global estimate of 20th century sea-level rise. These observations highlight the importance of a regional approach to understanding past and present sea-level changes and the need for regional-scale predictions of future sea-level rise (Milne et al., 2009).

Understanding the regional patterns of sea level provides knowledge not only of the mechanisms and dynamics of sea-level change, but also of the mass-balance changes of ice sheets and glaciers in response to climate change (Milne et al., 2002). Melting of continental ice sheets increases the volume of water in the oceans, which is distributed non-uniformly around the globe due to the change in the mass and gravitational attraction of the ice sheets (Mitrovica et al., 2001; Tamisiea et al., 2001). Detailed understanding of these processes and the 'fingerprint' of sea level which results from changes in ice sheet and mountain glacier mass balance requires geological records of past sea level from the near-, intermediate- and far-field sites that dovetail with the instrumental and geodetic era (Figure 1A).

Tide gauges or sea level recorders (WOCE, 2002) come in many different forms, and simply aim to determine the level of the sea at a point in time for one location. Modern instruments provide four to ten readings per hour to the nearest centimetre but for some scientific questions they are limited by either their length of record, as noted above, or spatial coverage. Different types of geological measurements of coastal environments provide quantitative measures of relative sea-level change over timescales of millennia and centennial (e.g. Kemp et al., 2011) through to instantaneous, in the case of tectonic relative sea-level change where they also provide better spatial detail (Farías et al., 2010; Plafker, 1969). By seeking to quantify the vertical and temporal resolutions of different types of geological tide gauges we can extend the spatial coverage and length of record of past sea-level change and therefore address new, important scientific questions.

In this paper we aim to review the developments in using one type of geological tide gauge, salt marshes, to produce records of late Holocene relative sea-level change, with particular emphasis on the methodological and technical advances that promoted research on finer resolution, in some cases sub-decadal and less, relative sea-level changes over the last millennium. We work through an example to demonstrate the consequences of different assumptions and decisions required during different stages of analysis in producing quantitative estimates of relative sea-level change. Finally, we examine how to extract trends in RSL from what are inherently 'noisy' proxy reconstructions.

2. The background: reconstructing RSL changes using coastal sediments

The non-uniform distribution of meltwater from continental ice sheets to the oceans means that any particular coastal location will only record relative sea-level change, defined as change relative to present (Mitrovica and Milne, 2002; Mitrovica and Milne, 2003; Mitrovica et al., 2001; Peltier, 2004; Plag, 2006; Shennan et al., 2012; Tamisiea et al., 2001). The glacial isostatic modelling studies that led advances in understanding these processes required quantitative reconstructions of age and elevation of past sea level from different regions, with records covering several millennia. Low energy coastal sediment sequences, from the high latitudes to the tropics, provided many of these records, starting with the pioneering work of Godwin (1940) followed by numerous studies from 1960s onwards (as summarised by Pirazzoli, 1991). Sediment cores or outcrops that contained beds of peat overlain or underlain by intertidal minerogenic sediments provide ideal samples for reconstructing past sea level. Radiocarbon dating allows dating of organic material to give the age parameter, and the stratigraphic association with a tidal deposit provided the elevation relationship with palaeo sea level.

Reconstructing RSL requires four attributes for each sea-level indicator or index point: location, age, elevation (both the measured elevation of the sample and the modern relationship to the tide level at which such an indicator would form today), and tendency (van de Plassche, 1986). The tendency of a sea-level indicator describes the increase (positive sea-level tendency) or decrease (negative sea-level tendency) in marine influence recorded by the indicator. The age-elevation plot of individual sea-level index points (Figure 1B) gives a suitable summary over a 10,000 year timescale and represents the primary method by which we use such data to test glacial isostatic adjustment (GIA) models (Bradley et al., 2011; Brooks et al., 2008; Lambeck et al., 1998; Peltier, 2004). While the age-elevation plot comprises just the radiocarbon dated index points (Figure 1C) it does not reveal more subtle, though recognizable, changes in vegetation and lithology revealed during the analysis of coastal sediments which may help differentiate between different models of RSL (Figure 1D). The expression of the change in vegetation, stratigraphy or microfossils will be site specific, but the change in sea level of more than local significance should be recorded over the wider area. Unlike instrumental tide gauge data (Figure 1E), where we know the exact position of each observation on the time axis, unless the radiocarbon dated samples come from the same core or section we cannot define their precise sequence, and sub-millennial RSL changes in many cases may lie within the error terms and scatter of data points (Shennan, 1982; Shennan et al., 1983; Tooley, 1982). To identify sub-millennial scale changes it is necessary to analyse the stratigraphic and microfossil changes above and below each dated sample to identify trends through time (tendency) along with quantified error terms for age and elevation (Shennan, 1982; Tooley, 1978, 1982).

Technical developments in the 1990s: such as AMS radiocarbon dating, short lived radionuclide chronology and quantitative environmental reconstruction methods developed in ecology and paleoceanography, provided the stimulus for further developments in studying sub millennial RSL change. AMS, ^{210}Pb and ^{137}Cs methods allowed analysis of small samples, in some case contiguous 0.5-1 cm slices, through organic and clastic sequences leading to better resolved age profiles compared to those derived from conventional radiocarbon methods. Various studies on the eastern seaboard of the USA used these methods to improve the chronological control in studying the relationship between climate and sea level during the last two millennia (Thomas and Varekamp, 1991; van de Plassche et al., 1998; Varekamp et al., 1992). These and other studies aimed to quantify the vertical relationship between different microfossil groups and the intertidal zone (e.g. Gehrels, 1994; Scott and Medioli, 1980; Scott et al., 1984; Shennan et al., 1995). Continuous records of sub-centennial RSL change commonly rely on subtle changes in the assemblages of microfossils

(mainly diatoms and foraminifera) preserved in the fossil sediments, or changes in the sediment rate established by chronological methods. While a number of factors may control diatom or foraminifera distribution (e.g. pH, grain size, salinity, organic content, substrate) (Berkeley et al., 2007; Sullivan and Currin, 2000; Vos and de Wolf, 1993), studies repeatedly show that elevation, related to tidal inundation and a proxy for salinity, is a primary control of diatom and foraminifera distribution across a salt marsh and therefore suitable for reconstruction of sea level (e.g. Edwards et al., 2004c; Freund et al., 2004; Horton et al., 1999b; Patterson et al., 2004; Roe et al., 2009; Scott and Medioli, 1980; Sherrod, 1999; Szkornik et al., 2006; Zong and Horton, 1998).

Further improvements in the resolution of reconstructed elevation came from studies of paleoseismology in the Pacific Northwest of Canada and the USA (Atwater and Hemphill-Haley, 1997; Shennan et al., 1996). Guilbault et al. (1995) were the first to use the term 'transfer function' in relationship to RSL change recorded in coastal marshes. They use 19 surface samples across a marsh and tidal flat on Vancouver Island to reconstruct RSL changes from an adjacent Holocene sediment sequence, comparing their visual examination method with the same statistical methods originally developed by Imbrie and Kipp (1971). These early studies threw up further issues, including chronological control, tidal range variations, modern sampling design and model selection that we shall return to. Around the same time, research in paleoecology developed much better numerical models that combined improved statistical modelling of non-linear species-environment relationships (e.g. microfossil species abundances against elevation), improved computing power and more widely available software (Birks, 1995; ter Braak and Juggins, 1993; ter Braak et al., 1993). These developments have led to numerous publications from the late 1990s of both diatom- and foraminifera-based transfer function models that use the relationship between the modern distribution of microfossils (Figure 2) to provide quantitative estimates, with error terms, of RSL based on fossil assemblages in coastal sediments (Figure 1F and 1G) (Barlow et al., 2012; Edwards and Horton, 2000; Edwards et al., 2004a; Gehrels, 1999; Gehrels, 2000; Gehrels et al., 2012; Gehrels et al., 2006; Hamilton and Shennan, 2005; Horton and Edwards, 2006; Horton et al., 1999a; Horton et al., 2000; Kemp et al., 2009a; Long et al., 2003; Long et al., 2012; Long et al., 2010; Sawai et al., 2004a; Sawai et al., 2004b; Sherrod et al., 2000; Woodroffe and Long, 2009; Woodroffe and Long, 2010; Zong and Horton, 1999). These records have sometimes been termed 'high resolution', though high resolution is a relative term (Edwards, 2007), and rarely defined. It is the development and interpretation of these 'high resolution' reconstructions that are the focus of this paper.

3. Pathways of reconstructing 'high resolution' relative sea level from sediments

Any former elevation estimate derived by comparing a microfossil assemblage to a dataset of modern plants or microfossils will be a function of the environments from which the modern data are collected. Despite the apparently "objective" nature of quantitative reconstructions, they reflect a series of decisions on how to develop and apply this modern dataset. A key aim of this paper is to demonstrate the implications of decisions that accompany the majority of reconstructions (Figure 3) (sections 3.2 to 3.6) and illustrate the consequences on RSL reconstructions (section 3.7).

3.1. Research question and site selection

The points discussed in this paper apply to any modern and fossil datasets from coastal sedimentary environments. Development of any sea-level reconstruction is in part guided by a particular research

question or hypothesis being tested, and field locations are selected accordingly. An example research question may be: 'What is the magnitude and direction of sub-millennial late Holocene RSL changes around the North Atlantic?' To answer this, a range of sites around the North Atlantic would be selected to provide suitable archives of late Holocene sea-level change. In this paper our aim is to work through an example and illustrate the impact of decisions made at different stages of developing high precision RSL reconstructions and their interpretation (Figure 3), rather than draw conclusions about RSL.

To demonstrate this we use a sediment core from a previously unstudied site and large body of previously unpublished data, therefore independent of published results or interpretations. Loch Laxford, on the north-west coast of Scotland, UK (Figure 4 and supplementary .kml file), is a small fjord, with islands and side arms that include two subsidiary lochs (Bates et al., 2004). The outermost part of the loch is exposed to prevailing westerlies, but at the loch head a sheltered inlet leads to a basin, Tràigh Bad na Bàighe, with an extensive, sand-dominated tidal flat and vegetated salt marsh abutting steep topography of Lewisian Gneiss Complex metamorphic rocks. The most extensive area of marsh, ~200 x 70 m, lies at the north of this basin. There are few areas of transitional succession from sand flat to pioneer marsh, with a salt marsh cliff, ~10cm, forming the boundary between sand flat and salt marsh across much of the site. The salt marsh supports an extensive creek network, and covers ~1.2 m vertical elevation range, reaching a zone of *Iris* before heather upland communities.

We selected a core for analysis from the mid marsh, simply as it is a shallow core by which to demonstrate the consequence of the decision pathways outlined in Figure 3. The core top at 1.80 m OD (ordnance datum (OD) is the national leveling datum for the UK). Mean High Water of Spring Tides (MHWST) is at 2.40 m OD and Mean Tide level (MTL) at 0.25 m OD. The core comprises of 42 cm humified organic salt-marsh sediments overlying tidal flat sand. Diatom samples were counted at 4-8 cm intervals from the surface to 36 cm (Figure 5).

3.2. Will the modern environment at our site reflect those we find in the past?

Quantifying the relationship between microfossil assemblages and elevation requires samples of modern environments that reflect those in our fossil sequence. Commonly, the modern samples (also known collectively as a modern training set) are collected from across a marsh and tidal flat proximal to the location of the core collected for the RSL reconstruction (Gehrels et al., 2006; Woodroffe and Long, 2010). This assumes that the 'local' marsh and tidal flat contains analogues for the full range of palaeoenvironments and their associated microfossils preserved in the core. If sediments are deposited under very different environmental conditions from present, a local microfossil training set may not adequately cover the full diversity of palaeoenvironments and their microfossil assemblages. When this occurs we must develop a training set that includes modern samples from other marshes that provide analogues for these past environments. These datasets are often called 'regional' though the definition can vary from author to author, with regional datasets varying from an estuary to a 100 km stretch of coastline to a whole country (Edwards and Horton, 2000; Gehrels et al., 2001; Kemp et al., 2009b; Leorri et al., 2008; Szkornik et al., 2006; Watcham et al., 2013; Zong and Horton, 1999). There is on-going debate as to the benefit of one approach over another (e.g. Gehrels et al., 2001; Horton and Edwards, 2005; Watcham et al., 2013; Woodroffe and Long, 2010), but in general, if a local marsh provides modern analogues for the full range of palaeoenvironments it reduces noise in species distribution which may occur in a regional modern dataset caused by factors other than elevation controlling micro-fauna and -flora

distribution (e.g. Woodroffe and Long, 2010). Equally, without suitable modern analogues, RSL reconstructions will not be robust and a regional modern dataset will be more appropriate, even if this has consequences for the resulting reconstruction (e.g. Watcham et al., 2013).

Modern microfossil surface samples are typically collected at ≤5 cm vertical intervals, either along one or more defined transects perpendicular to the primary direction of tidal inundation, or in a less stratified manner guided by vegetation and sediment transitions from the tidal flat to freshwater environment. If samples are unevenly vertically distributed it is advisable to apply a data transformation method (Gehrels, 2000; Telford and Birks, 2011). Sampling along defined transects allows one to be sure of collecting samples evenly along an environmental gradient, which allows accurate estimation of species distribution and abundance (Birks, 1995; Telford and Birks, 2011), though may increase the occurrence of spatial autocorrelation. Positive spatial autocorrelation is the tendency of sites close to each other to resemble one another more than randomly selected sites (Legendre and Fortin, 1989), which introduces bias in transfer function model performance results (Telford and Birks, 2009). Random sampling of a modern dataset is sometimes used with the intention of removing spatial autocorrelation (Leorri et al., 2008), and may help to reduce the effect but will not entirely remove it, particularly as the samples are typically collected geographically local to each other within one marsh or estuary.

The method of sampling varies depending on the proxy of interest, with the topmost centimetre of sediment commonly collected for modern foraminifera (e.g. Wright et al., 2011), compared to a few millimetres of sediment, after removing the surface film, for diatoms (e.g. Hamilton and Shennan, 2005). The assemblage of diatoms and/or foraminifera is usually counted from a proportion of the whole sample (Battarbee et al., 2001; De Rijk, 1995; Freund et al., 2004; Horton et al., 1999a; Revets, 2004). Building up a picture of the relationship of particular species and assemblages of diatoms and foraminifera relative to elevation provides a basis from which to calibrate the fossil assemblages found within a core.

3.2.1. Characteristics of the modern diatom dataset

We present a new large modern diatom training set to calibrate the fossil diatom assemblages in the Loch Laxford core. The local dataset comes from Loch Laxford itself and includes 73 samples from tidal flat, salt marsh and freshwater upland, over a 2.47 m vertical range. The 'regional' dataset comprises 215 samples (with 357 species, of which 115 species contribute >5% of total valves counted) from nine tidal flat – salt marsh – upland sites along the western and northern coasts of Scotland (Figure 4, Table 1 and supplementary figure S1). The large number of samples in the dataset helps in precisely defining the vertical species distribution. Zong and Horton (1999) originally published data from 15 modern samples from two of these sites, Tramaig Bay on Jura and Kentra Moss in Kentra Bay. We have added additional samples from the original Jura and Kentra Bay surveys (Shennan et al., 1995, and unpublished). It can be a challenge merging datasets from different researchers and when doing this it is necessary, where possible, to ensure coherent standardizing of taxonomy and updating species names in line with current nomenclature. Our north-west Scotland modern diatom dataset is available for download from http://www.dur.ac.uk/geography/qec/research_groups/slru/sea_level_data/.

To allow for tidal range differences between sites in the regional dataset (Table 1), authors typically apply a standardisation method (e.g. Gehrels, 1999; Horton et al., 1999a). We apply the following standardized water level index (SWLI) equation:

$$SWLI_n = \frac{100(h_n - h_{MTL})}{h_{MHWST} - h_{MTL}} + 100$$

Where $SWLI_n$ is the standardised water level index for sample n, h_n is the elevation of sample n (m OD), h_{MTL} is the local mean tide level elevation (m OD) and h_{MHWST} is the local mean high water spring tide elevation (m OD). Therefore MTL has a SWLI of 100 and MHWST 200.

3.3. A priori, can we define the elevation range of palaeoenvironments in our core?

Sediments of the upper intertidal zone, close to the highest astronomical tide mark, most sensitively record small changes in RSL, due to the fact that accommodation space is being created by sea-level rise (Allen, 1990). Accommodation space lower in the intertidal zone can be filled independently of a change in sea level (Allen, 1990). Low salt marshes, tidal flats and subtidal environments are therefore less accurate recorders of small RSL changes than mature upper salt marshes (Gehrels, 2000). In theory, the more sensitive environment should provide RSL reconstructions with smaller error terms. Even within upper salt marshes there are elevation zones that are more sensitive than others, where different sediment types, vegetation communities or microfossil assemblages occupy narrower elevation ranges. A key challenge is identifying these sensitive zones in both contemporary and fossil environments. This is reflected in two contrasting approaches to creating the modern dataset used for estimating RSL changes from a fossil core.

In one method the modern samples come from only the section of the modern environment expected to be an analogue for the fossil environments recorded in the core. The critical question is whether we can use independent parameters *a priori* to determine the elevation range for taking these samples. This may result in different modern datasets being applied to different sediment lithologies (e.g. Hamilton and Shennan, 2005) or by using other micofossil evidence (e.g. pollen or testate amaobae). A combination of the enhanced mixing processes on tidal flats, and the often poorly defined lower ranges of flora and fauna distribution associated with these environments, inevitably results in larger uncertainties for assemblages from intertidal minerogenic units than organic units. This has led to many studies focusing on continuous salt marsh peats rather than intercalated layers of peat and silt to study small scale RSL changes over the last few millennia (Barlow et al., 2012; Edwards, 2001; Gehrels, 2000; Gehrels et al., 2006; Gehrels et al., 2004; Kemp et al., 2011; Leorri and Cearreta, 2009; Woodroffe and Long, 2009). Therefore some authors would only collect modern samples from organic middle to high salt marsh environments and not from adjacent environments. They argue that small changes in the abundance of species from the low marsh and tidal flats can considerably affect the model predicted heights of the microfossil assemblage from high marsh settings and, as a result, decrease the precision of the reconstruction (e.g. Gehrels, 2000).

The alternative method is to collect modern samples over a much larger vertical range from the coastal transition, typically the freshwater zone to intertidal environments, though few studies include sub tidal samples (Woodroffe, 2009). This method captures a larger number of species across the range of environments and often provides a fuller picture of species distribution, meaning the training set is less likely to miss the tails of species distributions.

To demonstrate these two approaches we divide the local and regional modern training sets from Scotland into two environmental gradient lengths. Environmental gradient length is a generic term for the sampled range of the environmental variable of interest. In our example this is elevation, so

our long environmental gradient dataset, the "coastal transition model", has modern samples ranging vertically from intertidal sand flat at MTL (SWLI 100), to just above the transition to the upland freshwater zone (SWLI 300). The short environmental gradient dataset, the "salt marsh model", only includes samples from SWLI of 140 to 200. This is the range covered by the contemporary salt marsh environments at Loch Laxford and therefore we make the assumption it will provide analogues for assemblages present in a core containing salt marsh sediments.

3.4. Is the species response to elevation linear or unimodal?

A transfer function models the relationship between the modern micro-fauna and -flora assemblage and an environmental variable of interest, in this instance, elevation (Birks, 1995, 2010). The models provide a predictive equation to transform the microfossil assemblage in a sediment layer to calculate the fossil layer's elevation in the tidal frame at the time it was deposited. This quantitative estimate of palaeo marsh surface elevation (PMSE) includes an associated error term (Gehrels, 2000; Horton et al., 2000; Horton and Sawai, 2010; Zong and Horton, 1999). Each transfer function model has specific choices and underlying statistical assumptions which can impact the resulting reconstructions.

One vital part of developing a transfer function model is assessing whether the modern species-environmental response is unimodal or linear, typically using constrained ordination methods (Birks, 1995, 2010; ter Braak and Prentice, 1988). Birks (2010 p37) notes that weighted averaging (WA) and weighted averaging partial least-squares (WA-PLS) are currently considered *'simple, robust and widely used approaches'* where unimodal species-environmental responses are demonstrated (Birks, 1995, 2010; ter Braak and Juggins, 1993; ter Braak et al., 1993). Alternative approaches model linear species-environmental responses. It is relatively uncommon to find linear species-environment responses in Holocene RSL studies due to the typically Gaussian distribution of species relative to an environmental variable and therefore we do not explore this part of the reconstruction pathway further. Exceptions include Leorri et al. (2010) and Rossi et al. (2011) who develop foraminifera datasets from a relatively short vertical range with respect to the local tidal range, and therefore the modelled species-environmental response appears linear. Birks (1995, 2003, 2010) provides comprehensive reviews of the different statistical approaches; they are not the specific focus of this paper.

To assess whether the species-environment response in our datasets is linear or unimodal we use Detrended Canonical Correspondence Analysis (DCCA) in CANOCO version 4.5 (ter Braak and Smilauer, 2002) to quantify the diatom assemblage change with elevation. The species turnover is measured by DCCA axis 1. A general rule of thumb is where the DCCA gradient length is greater than two standard deviations sufficient species in the training set having their optima located along the environmental gradient and are collectively responding unimodally to elevation (Birks, 1995; ter Braak and Prentice, 1988). Less than two standard deviation units and the species are more monotonic along the gradient meaning linear methods are more appropriate (ter Braak and Juggins, 1993). Table 2 summarises the DCCA results for our four modern training sets, all of which have gradient lengths >2 allowing us to apply unimodal regression models. The regional models have a shorter axis 1 length than the comparable local modern dataset. This may be due to environmental factors other than elevation controlling diatom distributions or due to spatial autocorrelation in the dataset (Guiot and de Vernal, 2011; Telford and Birks, 2005). We use the WA-PLS technique in the palaeoenvironmental reconstruction programme, C2 (Juggins, 2003). This uses the average

distribution of taxa along the environmental gradient, weighted by their abundance, to estimate fossil elevations (as in the weighted averaging (WA) technique). Partial Least Squares adds additional components which take into account correlations in the residuals to improve fit between the species data and elevation in the modern dataset and therefore is generally considered superior to WA (Birks, 1995, 2010; ter Braak and Juggins, 1993; ter Braak et al., 1993).

3.5. How many components (WA-PLS)?

Selecting the number of components for a WA-PLS model is an important decision as it may impact on the elevation reconstruction and its error term (ter Braak and Juggins, 1993; Wright et al., 2011). The number of components used is often selected based on the predictive ability of the model summarised by the Root Mean Squared Error of Prediction (RMSEP), with a lower RMSEP suggesting improved model performance (Table 2). Telford et al. (2004) discuss the limitations of using the RMSEP and suggest it is preferable to assess model performance using a range of evidence, including RMSEP, the bootstrapped coefficient of determination (r_{boot}^2) and simple scatterplots of observed and predicted values. Simply choosing the number of components based on which model has the smallest RMSEP is not good practice.

Table 2 shows the RMSEP and r_{boot}^2 values from first three components of our four modern training sets and Figure 6 some of the corresponding scatterplots. The scatterplots show that a two or three component solution provides a better statistical fit across the full elevation range of the modern data, with a one component model showing a non-linear distortion at the ends of the distribution that is also well known in both weighted averaging and correspondence analysis (Birks, 1995). Authors hold different opinions regarding the use of multiple component models (e.g. Birks, 1995, 1998; Wright et al., 2011), with some arguing (e.g. Wright et al., 2011) that more components add statistical complexity and a less direct measure of the relationship with the environmental variable of interest. Further components update the species coefficients and can alter the resulting reconstruction. Other variables that influence microfossil distributions also co-vary with elevation, so the statistical structure within the residuals of a simple one-component model may reflect such relationships; therefore it is acceptable to use multiple component models. The unknown is when this structure becomes less than statistic noise. Therefore, we apply a general rule of thumb: to apply no more than three components and only select the successive component if the model performance increases by >5% measured by RMSEP and r_{boot}^2 (Birks, 1998). Based upon these measures and their associated scatterplots we select the two component models for all the datasets (Table 2). In later sections we also show the consequences of applying one-, two- and three-component models to reconstructions. The summary statistics (Table 2) suggest that the local 'salt marsh model' appears to have the best predictive ability but as we explain below it is important not to rely solely on these measures when selecting models and assessing their reconstructions.

3.6. Should we exclude any training set samples?

The next phase in the decision pathway (Figure 3) is whether to try and improve the predictive ability of the transfer function by excluding samples from a modern dataset. This is different to the question raised in section 3.3 where our example modern training sets reflect an *a priori* decision that we can define the elevation range of palaeoenvironments in our core. There is no consensus on the approach to excluding samples. Examples in sea-level and other studies include: Gasse et al.

(1995) who remove samples that have a difference between the predicted and observed values greater than one-quarter of the total observed range; Edwards et al. (2004a) and Gehrels et al. (2005) who eliminate samples whose residual value is greater than the standard deviation of SWLI; and Heiri et al. (2011) remove samples based on environmental parameters. The alternative is to leave all the samples in, on the implicit assumption that spread in the dataset captures the diversity of the sampled modern environments. For our worked examples we do not exclude any samples beyond the *a priori* decisions made in section 3.3, but note that removing samples from a training set will impact the resulting reconstruction and the error term for each fossil sample. For example, the results in Table 2 shows how the shorter environmental gradient models (the 'salt marsh' models) have smaller RMSEP values than the longer 'coastal transition' models. A similar effect may be achieved when removing selected samples by other approaches.

3.7. Resulting reconstructions and assessing accuracy

Once a decision pathway has been made, the resulting model(s) are applied to the fossil assemblages in each sample taken from a core, and provide reconstructions of palaeo marsh surface elevation. This is the estimated elevation, at which each fossil sample formed relative to tide levels and measured to a common datum. We convert these values to RSL by subtracting the model estimate of palaeo marsh surface elevation from the present elevation of the sample, both measured relative to the common datum. Bootstrapping cross validation derives a RMSEP value (sample specific vertical error term) for individual fossil samples (Birks, 1995).

Figure 7A shows RSL reconstructions derived from 12 models applied to our core (the first three component models for four modern datasets). These results illustrate how different decision pathways can influence the RSL reconstructions. As mentioned before, we do not aim to draw conclusions about regional RSL change from these records, but rather assess how a record may differ based on the decisions taken in its development. The overall trend, of increasing sea level, is similar between all the records, though the magnitude of change and inflections within the overall trend vary. For example, there is a ~40 cm difference in estimated RSL between the different models for some of the fossil samples.

The two most commonly adopted approaches to RSL reconstruction are using a large regional training set with a long environmental gradient (e.g. Sawai et al., 2004a), or a local dataset with the modern elevation range defined *a priori* by the palaeoenvironments in the core (e.g. Gehrels et al., 2006) (Figure 7B). In our example the regional 'coastal transition' model estimates greater RSL change than the reconstruction resulting from local 'salt marsh' model, though the long environmental gradient regional 'coastal transition' model has ~3 times larger vertical errors than the local 'salt marsh' model (we analyse this further in section 4). Figure 7C shows that in some instances, the number of components does not have a major impact on the resulting reconstruction (e.g. regional 'salt marsh' models), whereas the various components result in quite diverging results with other models (e.g. regional 'coastal transition' models). The results in Figure 7 should not be interpreted as a blueprint for other datasets. Rather, the key point is to recognize that differing results will likely occur as function of decisions taken at different stages of model development. Investigators should therefore be explicit about the reasons for making the choices and indicate their effects on their results.

Since all transfer function reconstruction pathways will provide results, irrespective of model evaluation and performance, it is important to assess the accuracy and reliability of any

reconstruction using independent techniques. For example, some previous studies validate fossil reconstructions against instrumental tide gauges (Gehrels et al., 2006; Kemp et al., 2009a; Leorri and Cearreta, 2009; Rossi et al., 2011), or observations of coseismic land-level deformation (Hamilton and Shennan, 2005; Watcham et al., 2013). In the absence of such independent data, the most basic test is to check whether the transfer function can accurately reconstruct the leveled elevation of the core top by calibrating the elevation of the assemblage in the 0 cm sample. In our example datasets the regional models produce closer fits than the local models. With larger datasets it is possible to use a proportion of the modern dataset to reconstruct the elevations of an excluded subset and compare the reconstructed values with the known leveled heights to check model accuracy (Callard et al., 2011; Charman et al., 2010; Reavie and Juggins, 2011).

A reconstructed elevation is likely to have some reliability if the fossil sample has a microfossil assemblage with close analogues in the modern training set (Birks, 1995; Edwards and Horton, 2000; Hamilton and Shennan, 2005). As we go back in time, there is likely greater change in the fossil record, relative to the modern environment (Watcham et al., 2013), particularly as the present relationship between micro-fauna and -flora and elevation may be disturbed by human activities (Guiot, 1990). Therefore, analogue matching is an important means of evaluating the likely reliability of environmental reconstructions (Birks, 1995).

The modern analogue technique (MAT) quantifies the similarity between each fossil sample and the modern training set using a squared chord distance dissimilarity measure (Birks, 1995). The MAT produces a minimum dissimilarity coefficient (minDC) for each fossil sample, with a minDC value of zero indicating exact similarity with the closest modern sample, and higher values indicating increasing dissimilarity (Jackson and Williams, 2004). There are no fixed rules in defining a "good", "close" or "poor/no" modern analogue, but it is typical to consider the distribution of dissimilarities within the modern dataset. We follow Watcham et al. (2013) using the 20th percentile of the dissimilarity coefficients calculated between all modern samples as the cut-off between 'close' and 'poor' modern analogues for fossil samples, and the 5th percentile as the threshold for defining 'good' modern analogues (Table 2 gives the threshold values for each dataset). These are quite strict measures compared to the threshold used by Woodroffe (2009) who used the largest DC value calculated between all modern samples as a cut-off for each fossil sample between a 'good' and 'poor' match. Figure 7C shows whether the fossil samples have a good, close or poor modern analogue in the respective modern dataset (note the MAT dissimilarity coefficient does not vary for each WA-PLS component, just between each modern training set). Based on the 5th and 20th percentile thresholds, the regional model reconstructions are the most reliable with no fossil samples have a minDC greater than the 20th percentile value, and six out of the seven fossil samples having a good modern analogue with samples in the regional 'coastal transition model'. By comparison, when calibrated with either of the local models, six fossil samples have poor modern analogues, and only one a close analogue (the top sample), suggesting the local marsh is not a good approximation of all past environments. This is comparable to the results of Watcham et al. (2013) in Alaskan environments probably much less disturbed by human activity. If based simply upon the RMSEP of the models and sample specific error terms, the local 'salt marsh' model appears most precise, but the MAT identifies that the results are not reliable, and therefore may not be accurate. In our example, the regional 'coastal transition' two components model has the best combination of model predictive ability, fit to known elevation of core top and matching modern analogues, though the largest error terms. These error terms are likely a truer representation of the amount of uncertainty than those estimated by the local and elevation restricted training sets.

3.8. Additional factors

There are environmental factors which are not part of the decision pathways in Figure 3. These may impact on RSL reconstructions, but are often difficult to quantify and not included in the reported transfer function calculated error terms.

Differences between diatom valve morphology, and their amount and type of silica means diatoms suffer variably from abrasion, dissolution and drying, and can be differentially preserved in coastal sediments (e.g. Cooper et al., 2010; Jordon and Stickley, 2010 and references cited within). More robust species are also tolerant to transport by tidal currents which can lead to preferential enrichment of these species in high intertidal sediments, meaning the preserved assemblage may have little to do with the palaeoenvironments formation elevation, compared to that characterised by an autochthonous assemblage (Sawai, 2001; Vos and de Wolf, 1988). Some authors try to include assessments of dissolution, fragmentation and the proportion of allochthonous to autochthonous species to account for some of these processes (Ryves et al., 2009; Ryves et al., 2001; Vos and de Wolf, 1988). By sampling the top 2-10 mm of modern sediments, samples will incorporate death assemblages from a number of years allowing the transfer function to incorporate the net effect of all these factors rather than quantify the effect of each one.

Comparable processes result in variable foraminifera preservation, with agglutinated foraminifera more resistant to dissolution than calcareous species in the low pH conditions common on many marshes (Bradshaw, 1968; Jonasson and Patterson, 1992; Phleger, 1966). This is why often only dead foraminiferal tests are included in modern training sets, which also helps remove anomalies arising from seasonal growth effects (Edwards et al., 2004a; Horton et al., 1999a; Murray, 2000). In addition, Patterson et al. (2004) suggest that bioturbation, mean annual temperature variations and oxygenation of the marsh sediments may impact the potential to use foraminifera to reconstruct RSL.

Sediment mixing, whether from bioturbation or other processes, has the potential to blur the record, but there are few studies that show its impact. Foraminifera can live up to 30 cm below the marsh surface (Goldstein and Harben, 1993; Hill et al., 2011; Massey et al., 2006a) potentially biasing the fossil assemblage towards species which burrow into the sediment column. With respect to diatoms, experiments on Alaskan marshes suggests sediment mixing was limited to 8 mm from the surface in silt rich sediment but to a greater depth (1-2 cm) in peat (Hamilton et al., 2005; Shennan and Hamilton, 2006).

3.9. Reviewing model performance and RSL reconstructions

Considering the number of potential pathways to a RSL reconstruction (Figure 3), it is vital to apply independent methods to assess which reconstruction produces the most realistic estimate of former sea level. This may not always be the reconstruction with the smallest errors. We should consider at least three criteria (Watcham et al., in press): number of fossil samples with 'poor' or no modern analogues; that reconstructions are reasonable when compared to independent measurements; and finally that the reconstructions make ecological and environmental sense. In addition to independent observations, such as instrumental tide gauge records or measures of co-seismic subsidence, as discussed above, the first test should be to asses each reconstructed elevation against sediment lithology within the core. As noted in section 2, lithostratigraphic position may define smaller uncertainty terms for samples close to sedimentary boundaries, compared to those in

the middle of homogenous sediments. In our example, a fossil sample from salt marsh peat would have an error term in the order of ±40 cm (Figure 7), whereas the lower limit may be more precisely defined by measurements of the lowest occurrence of salt marsh peat formation in the present environment. A similar constraint may be applied at the upper end of a salt marsh if there is a clear transition to a *Sphagnum* peat or another sediment that clearly forms above the influence of highest tides. To make ecological and environmental sense, we would expect the RSL reconstruction to match environmental gradients observed in the lithology, such as intertidal sand flat or mudflat to vegetated marsh. Any reconstruction should therefore consider the ecology and lithological context for each assemblage before committing to a RSL reconstruction, not least adherence to Walther's Law of the Correlation (or Succession) of Facies, which identifies that a vertical sequence of facies will be the product of a series of depositional environments occurring laterally adjacent to each other (in the absence of hiatuses/erosion in the sequence).

4. Tidal range and errors in relative sea-level reconstruction

There is a positive correlation between transfer function model error terms and tidal range at the site(s) used for the modern training set (Table 3 and Figure 8A). As described above, RSL reconstructions from minerogenic units that represent mid- to low intertidal flat environments will most likely have the largest error terms because of the lack of narrow micro-fauna and -flora zones and poorly defined lower elevation ranges. We can address this at different stages during our analysis by targeting sediment sequences that come from higher intertidal environments (Section 3 and Figure 3). Other influences of tidal range still remain, particularly the strong relationship between the elevation range of the modern samples used to generate the transfer function model and the resulting model error term (Figure 8B).

These relationships (Figures 8a and b) show first, as Southall et al. (2006), Edwards and Horton (2006) and Callard et al. (2011) also note, that micro-tidal environments can provide more precise sea-level reconstructions than meso- or macro-tidal settings. However, as the tidal range in micro-tidal environments decreases, environmental controls other than salinity can potentially be more dominant on species distribution, than in larger tidal settings. For example, changes in local hydrology and groundwater conditions can change marsh elevation trends through shrink‑swell of marsh soils (Cahoon et al., 2011). This could be tested by producing RSL reconstructions from more than one location at the site to determine within-site variance (e.g. Barlow et al., 2012; Long et al., 2012). As we shall see, in micro-tidal settings, developing a high precision age model (section 5) becomes especially important for reliable RSL reconstruction.

A second important observation is that decisions taken in selecting the elevation range of the modern sample will strongly influence error terms of RSL reconstruction, with the majority of studies producing RMSEPs between 5 and 15% of the sampled vertical range of the modern dataset (Table 3), and therefore we should be mindful of this when testing hypotheses of decimetre to centimetre-scale RSL fluctuations. If a reconstruction has errors >~5% of the modern tidal range, we should explore why this is the case.

Tidal range varies along the world's coastlines producing micro-, meso- and macro-tidal regimes. Typically macro-tidal regions are defined as having a mean tidal range >4 m, a meso-tidal range is 2-4 m and micro-tidal regions have a tidal range <2 m (Davies, 1964). Ideally all reconstructions would be from microtidal environments and have small error terms, but to provide the coverage required

to asses the drivers of spatial variations in late Holocene sea level this is not possible and as a result we must be prepared to develop reconstructions of former sea levels from a range of tidal regimes.

5. Additional vertical uncertainty

Sample specific errors determined from transfer function models are just one part of the total error of RSL reconstructions. Vertical uncertainty in RSL reconstruction also comes from potential errors in the field or laboratory which are not a function of the transfer function model uncertainty. Defining the correct temporal and spatial scales of the study (Section 2 and Figure 1) will help determine which errors are relevant to a particular study. For example, correlation of sea-level index points from different sites (Figure 1B) must accommodate different sources of uncertainty than a RSL reconstruction from a single core (Figure 1G).

Vertical errors of sea level index points should also include: errors from levelling; angle of the borehole; measuring sample depth; extrapolating tidal range estimates over large distances or in an area with a large tidal range; overlying sediment compaction; core compaction; change in vegetation types over time; changes in the water table; uncertainty in identifying the correct tidal datum for the indicator of interest and changes in tidal regime (Devoy, 1982; Gehrels et al., 1996; Heyworth and Kidson, 1982; Massey et al., 2008; Preuss, 1979; Shennan, 1982). The magnitude of each error varies case by case, for example levelling errors are greater when sampling from a boat, than levelling on a salt marsh. There are no defined standard approaches; authors provide estimates of errors according to their methods, with some using 1 sigma and others 2 sigma or 95% probability estimates. Whichever measures are used, the total error (E_t) for each sample is calculated as:

$$E_t = \sqrt{e^2{}_1 + e^2{}_2 + e^2{}_3 \ldots + e^2{}_n}$$

Where $e_1 \ldots e_n$ are the individual sources of error (Preuss, 1979).

Typically when producing a transfer function reconstructions the RMSEP sample-specific error is not combined with these other sources of error. The importance of E_t relative to RMSEP transfer function error terms varies in different tidal settings. For example, the 2 sigma total error for the intercalated and basal peat index points shown in Figure 1B range from ~0.2 to 1.2 m, with a mean of 0.33 m. In a theoretical microtidal environment with a tidal range of 20 cm we may expect a transfer function model RMSEP about ±2cm (i.e. 10% of the tidal range). The sample specific error for each fossil sample will include an additional error depending upon the similarity between the modern and fossil assemblages. This standard deviation sample specific error must be multiplied by 1.96 to give the 95% probability error term. For a microtidal example, the 2 sigma transfer function model error is relatively small compared to the other errors. In contrast, in a macrotidal environment with a 5 m tidal range, a model RMSEP in the order of ±50 cm is a large additional uncertainty, likely greater than the sum of the other errors. As a result quoted transfer function error terms for reconstructions from microtidal environments may under represent the true level of uncertainty and E_t should include the errors listed above, along with the RSMEP error. However, this only matters when combining datasets from different locations, not when working on a single core, and reinforces the requirement to properly consider within-site and between-site variance.

Post-depositional compaction of the sediment column, either under its own weight or as a consequence of subsequent loading by water or overlying sediment burden, results in lowering the elevation of sediment from their original depositional altitude (Edwards, 2006; Hill et al., 2011;

Jelgersma, 1961; Long et al., 2006; Long et al., 2010; Massey et al., 2006b; Shennan et al., 2000b; Streif, 1979; Tornqvist et al., 2008; van de Plassche, 1982). Methods that allow for post-depositional compaction vary and their suitability, once again, depends upon the research question. For example, when investigating RSL over millennial timescales and comparing RSL between estuaries, such as in Figure 1B, authors may target thin basal peat layers that overlie incompressible basement sand or rock (e.g. Jelgersma, 1961; Tornqvist et al., 1998; van de Plassche, 1982). Comparison of basal peat index points with those within a thick sediment column and underlain by compressible sediments, (the intercalated peat index points in Figure 1B), will provide an estimate of the net effect of compaction over millennia (Horton and Shennan, 2009; Shennan and Horton, 2002; Shennan et al., 2000a). To combine with transfer function derived RSL reconstructions through a single core we require numerical modeling in order to quantify compaction processes. Brain et al. (2012) model the effect of compaction on salt marsh reconstructions of late Holocene sea level, and find that thicker (>1 m) sequences experience greater compaction, with transgressive sequences experiencing greater post-depositional lowering than regressive sequences. They show that where a transgressive sequence occurs in association with a large change in organic content at a stratigraphic contact, sediment compaction can result in a corresponding inflection in the reconstructed sea-level curve, which will contribute a proportion of any reconstructed apparent sea-level acceleration. Lack of quantification of the magnitude of sediment compaction can result in erroneous estimates of the rate and magnitude of sea-level change (van de Plassche et al., 1998).

Changes in sediment supply, land level, configuration of coastal geography and land use means that in many locations, tidal regimes were likely different in the past. Many of the late Holocene RSL records cited in this paper do not take into account tidal changes through time, largely because they are small on centennial timescales. They are also hard to quantify and there are a limited number of models that provide estimates of local/regional palaeotides (Austin, 1991; Gehrels et al., 1995; Hill et al., 2011; Hinton, 1996; Shennan et al., 2003; Uehara et al., 2006). Changes in tidal range have implications for a RSL reconstruction, either because the distribution of the micro-fauna and -flora as controlled by tidal inundation in the present day may be different under a different tidal regime (which a transfer function is unable to account for); or because changes in the elevation of reference water levels used to constrain the altitude of sea level index points results in over or under prediction of their reconstructed elevation. This is an area of research that requires further attention to be able to produce models which have the spatial and temporal resolution required for correcting late Holocene salt marsh RSL records. Human modification of the coastal zone also becomes an important consideration. Land reclamation over the last 2000 years in estuaries such as Southampton Water, the Humber and Fenland in the UK seems to have caused a major change in tidal range within these estuaries and which are manifested in the stratigraphy (Long et al., 2000; Shennan et al., 2003). Over shorter timescales, instrumental tide gauges reveal the effects of dredging estuaries (Cox et al., 2003; Wang et al., 2002).

Most forms of vertical uncertainty discussed in this section do not apply when developing and interpreting a single salt marsh RSL record. However, if reconstructions from within an estuary, or region, are combined to produce a regional sea-level curve, or the magnitude of reconstructed RSL change between regions is compared, understanding and quantifying the amount of vertical uncertainty is important as otherwise offsets between data may occur and be misinterpreted as a differential RSL signal.

6. Age model independent relative sea-level changes

An important part of developing geological tide gauges is defining the timing of any sea-level change. Precise chronologies allow comparison of records and assessment of the potential driving mechanisms of the reconstructed changes. Advances in AMS ^{14}C techniques provides the ability to date carefully selected small samples of known origin, removing uncertainties caused by contamination by other material which may occur in bulk dating. This allows numerous radiocarbon dates from a single core, which, with age-depth modeling (e.g. Blaauw et al., 2003; Blockley et al., 2007; Yeloff et al., 2006) can provide a interpolated chronological framework for the whole sequence. There are challenges of developing radiocarbon chronologies for records of recent sea-level change. Large fluctuations in atmospheric ^{14}C characterise the last 500 years, driven by the Spörer, Maunder and Dalton solar minima and changes in the carbon cycle associated with Little Ice Age climate changes, as well as human industrial activity (Bard et al., 1997; Stuiver and Quay, 1980). As a result, several plateaus exists in the relationship between radiocarbon ages and calendar years from AD 1650 to 1950 (e.g. Reimer et al., 2009), complicated further by the effect of nuclear weapons from AD 1950 (Hua and Barbetti, 2004). This potentially limits the use of conventional radiocarbon dating in recent sediments as several possible calibrations solutions can exists for one date (Figure 9). This has driven the use of other, independent, chronological controls to date recent coastal sediments and refine ^{14}C calibration solutions, including ^{210}Pb, ^{137}Cs, stable Pb isotopes, tephra, pollutant horizons, pollen markers and palaeomagnetic analyses (Cundy and Croudace, 1995; Cundy et al., 2003; Gehrels et al., 2012; Gehrels et al., 2006; Haslett et al., 2003; Kemp et al., 2009a; Long et al., 1999; Long et al., 2012; Marshall et al., 2009; Marshall et al., 2007).

The shape of the radiocarbon calibration curve can lead to 'wiggles' in the age-depth model that could be interpreted as changes in the marsh surface elevation over time, and therefore apparent rather than real RSL changes. Van de Plassche et al. (2001) demonstrate the need for closely-spaced ^{14}C dates when developing age-depth models in recent salt marsh sequence, to fully utilise the 'wiggles' in the ^{14}C calibration curve. However, using simulated radiocarbon dates in OxCal (Bronk Ramsey, 2001), Figure 9 (after Gehrels et al., 2005) shows how closely-spaced ^{14}C dated samples, covering the 1000 years, results in a series of stacked calibration solutions, which would appear as periods of rapid sediment accumulation (Long et al., 2012; Long et al., 2010). If the calibrated microfossil assemblages show little change in palaeo marsh surface elevation (Gehrels et al., 2006; Kemp et al., 2009a), these stacked ^{14}C ages on an age-altitude plot may appear to represent periods of sea-level acceleration, whereas they are simply a function of the shape of the calibration curve.

Absolute and relative methods of dating recent salt marsh sequences are subject to various post depositional processes and have differing assumptions behind their use. This can result in misfits between different chronological controls applied to the same core (Gehrels and Woodworth, in press). For example, Turner et al. (2006) advise that some age models based on ^{137}Cs can suggest higher accretion rates than ^{210}Pb age models because the former applies to younger sediments, whereas the effects of root penetration and decomposition are greater in the latter. A challenge of combining AMS ^{14}C dating with ^{210}Pb is that both require 'modelling'. To combine them requires calculating calendar year ^{210}Pb dates based on a sediment accumulation model, which are then combined (often by Bayesian modelling or wiggle match dating) with ^{14}C dates that require calibration, to give a new accumulation model which may not be consistent with the initial accumulation model. This may explain why there can appear to be offset between ^{210}Pb and ^{14}C ages in some settings (e.g. Gehrels et al., 2005; Kemp et al., 2009a), often in a period critical to establishing the onset of recent rates of sea-level change and the timing of apparent accelerations/inflexions. It is important to develop chronological frameworks with multiple,

overlapping methods; for example bomb spike ^{14}C dates within ^{210}Pb sequences (Kemp et al., 2009a; Marshall et al., 2007). An alternative approach is to assess rates of sea-level change solely using dated levels rather than age-depth model interpolated data points (Gehrels and Woodworth, in press).

To ensure that the reconstructed RSL changes are real, rather than an artifact of the dating method(s), it is important to interrogate the data to assess whether there are changes in palaeo marsh surface elevation which are independent of the chronology. A high marsh core that contains high marsh foraminifera which shows little change in reconstructed palaeo marsh surface elevation only confirms the marsh is keeping pace with the rate of sea-level change (Gehrels et al., 2006; Kemp et al., 2009a). Developing a detailed microfossil diagram is very labour intensive, and it may be of more value to focus resources on developing a very precise chronological framework that identifies subtle changes in sediment accumulation that the microfossils do not pick up.   In comparison, where changes in palaeo marsh surface elevation can be identified independently from an age-depth model, identification of changes in the rate of sea level that are not driven by 'wiggles' in the ^{14}C calibration curve becomes possible. This is easier where micro-fauna and -flora have narrow ecological tolerances relative to the tidal range; for example diatoms in meso- or macro-tidal settings (Barlow et al., 2012; Woodroffe and Long, 2009). Both foraminifera and diatom based reconstructions, in both micro and macro-tidal environments, have merit but it is necessary to assess whether in testing the proposed hypothesis it is of more value to focus significant attention on the development of a detailed chronology or a detailed microfossil record at a particular site. We explore this further in the following section when considering how to separate trends from the inherently noisy geological tide gauge records.

7.   Salt marshes as geological tide gauges

Integrating instrumental tide gauge and satellite-derived observations of RSL with centennial to millennial scale reconstructions of past changes provides insights into climate-ice-ocean interactions but requires a clear understanding of how decisions made during the analysis may influence the result. Central to this is clarity regarding the relevant temporal and spatial scales of the dependent and independent variables under consideration. These depend directly upon the research question itself. For example, where the investigation aims to understand earth-climate-ice-ocean interactions we compare GIA model predictions with the Holocene radiocarbon dated index points of RSL from an estuary (Figure 1B). We seek a best fit for the eustatic, isostatic and tectonic parameters that operate at regional to global scales and aim to explain the residuals as the result of local-scale processes, such as tidal range, sediment compaction and coastal geomorphology (Shennan et al., 2012). In this type of analysis we only derive the centennial to millennial trends and use them to contrast with the shorter records from instrumental tide gauges and satellite data, identifying systematic offsets that suggest a climate change cause (e.g. Woodworth et al., 2009). A research question that focuses upon processes that operate over sub-centennial to decadal timescales requires a different approach due to the imprecision in correlating individual radiocarbon-dated index points from different locations. Salt marsh geological tide gauges use variations through a continuous sediment profile to quantify changes in RSL (Figure 1A and 1G and section 3). This offers numerous potential benefits. We know the exact order of data through time, in sequence one above another, even though there will be a sample-specific age error term (Figure 9). At this scale we still aim to consider the trends and importance of fluctuations or deviations from general trends (Figure 1G).

7.1. Interpreting geological tide gauge records

With a good age model, microfossil-based methods provide quantitative elevation reconstructions, so we can look at sub-centennial to decadal, or perhaps even finer resolution, RSL changes not identifiable from estuarine Holocene sea level index points from multiple sites. For example, Figure 10 shows how a microfossil based reconstruction and associated chronology may be combined to produce a geological tide gauge record (comparable to Figure 1A), and its subsequent interpretation. In this hypothetical case (Figure 10), we assume the microfossils show no change in palaeo marsh surface elevation with the marsh accumulating at the same rate as RSL change.

Figure 10A shows the trends already well established from Holocene sea level index points (Figure 1B) and instrumental tide gauge data (Figure 1E). Using radiocarbon dating we establish a chronology (Figure 10B). Microfossil based reconstruction of palaeo marsh surface elevation establish changes in RSL during the late Holocene (Figure 10C). Depending upon the sampling interval of microfossil analyses and radiocarbon samples the age parameter comes from whatever age model we select (Figure 10D). This is clearly a critical decision and is an example of how it may be of more value to devote resources on a detailed chronology and its careful interpretation, than additional microfossil analysis, as the pattern of change in Figure 10E and 10F are very similar and driven by the age-model.

The challenge is interpreting the records; and the choice of summary statistic and method of display can greatly influence the interpretations of tide-gauge data and geological sea-level reconstructions. For example, while giving the same linear trend line, using monthly or annual instrumental tide gauge data will convey different images (Figure 1E). In Figure 10 it may be possible to interpret the salt marsh record as a linear rate of sea-level rise (Figure 10E) or as a series of RSL fluctuations (Figure 10F), which may, or may not, be driven by the shape of the radiocarbon calibration curve.

It is important to show whether we can discriminate between a linear or non-linear trend from our analyses. Some of the deviations from the trend may appear random whereas others reveal a systematic pattern (Figure 1G). Where they are systematic we need to show that we have accounted for other local scale factors that may influence the reconstructed elevation. This will include sediment compaction, changes though time of the relationship between tide levels within the estuary and the open coast due to local factors such as barrier morphology, sediment supply or river discharge, and changes in tidal range at the open coast. We could aim to model each of these, as attempted by Kemp et al. (2011), and provide additional error estimates.

7.2. Replication

Modelling the separate parameters is only a part answer. The next stage is to determine within-site variability. We can isolate purely local-scale effects by identifying the spatial scale over which we observe the same trend in RSL. For example, for the hypothetical site in Figure 10, if we collect a number of cores from the same estuary, do we get the same RSL fluctuations of different magnitude and duration? Such an approach has been used to quantify past RSL due to coseismic deformation and Little Ice Age GIA in Alaska (Barlow et al., 2012; Shennan and Hamilton, 2006) and these examples would suggest about three cores from one marsh as a minimum target. This of course incurs substantive additional resource and time. In addition to validating the transfer-function model results, assessing local-scale variability by using multiple cores from the same marsh and additional marshes within an estuary will help to identify the effects of sediment compaction, age

model creation and the other coastal processes that control salt marsh accretion under different tidal ranges (see sections 4 and 5).

Replication can also include temporal – spatial scale constraints applied to the way we formulate our research questions. Temporal correlation of the direction and rate of RSL change can be explicitly tested by comparing RSL reconstructions from sites with different GIA histories, where the millennial-scale change is independently derived (Barlow et al., 2012; Long et al., 2012; Shennan, 1999; Shennan et al., 1983). For example, we can test the regional significance of the oscillation between ~AD1750 and 1800 in Figure 1A with a site undergoing greater GIA subsidence, where the rising limb should be at a greater rate, and a site undergoing GIA uplift should more clearly record the falling limb; assuming all other variables remains constant. Comparing sites over large distances reduces noise created by local process which may obscure the signals you are trying to test for.

7.3. Lithology

In striving to increase model precision based on micro-fauna and -flora assemblages, we must not ignore sediment lithology. This applies at numerous stages. Knowledge of the modern elevation range in which different coastal sediments occur can provide constraints both in transfer function model development (in the case of the 'salt marsh' or 'coastal transition' models discussion in section 3.3) and assessment of transfer function reconstructions (section 3.9). It can also apply to site selection. Both foraminifera and diatoms provide better resolution of elevation changes within the upper part of the tidal range, or salt marsh environment, than lower down in tidal flat and sub-tidal environments (see section 3.3). As Edwards et al. (2004b) note, it is more reliable to exploit the varying sensitivities and tolerances of individual taxa across a range of elevations. While some studies, such as the paleoseismology of great earthquakes, are best addressed using salt marsh and tidal flat sediments (Hamilton and Shennan, 2005; Nelson et al., 1996; Sawai et al., 2002), it is now generally recognised that gradual RSL change over the last 2000 years are better resolved using only salt marsh sediments.

7.4. Field sampling design

In sections 3.1 to 3.6 we addressed key questions that directly influence the outcomes of RSL reconstructions using quantitative methods. These focussed on different opinions and approaches to selecting the appropriate samples for the modern and fossil datasets. These datasets may comprise 100's of samples and, in the case of diatoms, species. Implicit in the different approaches is that the field sampling design adequately caters for ecological and environmental variability. Most of the studies summarised in Table 3 use an approximately straight transect across a marsh to collect samples from a range of elevations. The elevation range for sampling may be constrained by another parameter, such as lithology or vegetation community. In a similar way to the debates surrounding development of large pollen (Huntley, 2012) or chironomid (Brooks and Birks, 2001) datasets, sampling designs need to adequately address variability within a marsh and between marshes separated by increasing distance. For example, at a single marsh in west Scotland, Shennan et al. (2005) show that sedimentary and vegetation boundaries vary in the order of ±0.1 m. In Maine, Jacobson and Jacobson (1989) show salt marsh flora zonation varies greatly in both the number of zones present per marsh and the species assemblages within zones. Data from a single modern micro-fauna and -flora transect cannot determine spatial variability and in some cases the

whole dataset is scaled by a single observation along the transect, such as the highest occurrence of foraminifera (e.g. Wright et al., 2011). While numerous studies show variability of microfossil assemblages between sites, leading to the 'local' versus 'regional' question (section 3) we currently do not have enough studies that measure within-site variance to assess the effect of spatial autocorrelation (Telford and Birks, 2005) and the propagation of error through to RSL reconstructions.

7.5. Developing and interpreting geological tide gauge records: suggestions

We cite numerous papers which apply differing approaches to statistical based reconstruction of RSL, many of which have direct or indirect links with the authors of this paper. This goes to show that even within the authorship of this review there is no consensus to developing records of past RSL. We do not suggest that there is a need to adopt a one-pathway-fits-all methodology as different sites and research questions (e.g. RSL changes over different timescales) do require differing approaches. We have six suggestions which may help in the development and interpretation of geological tide gauge records:

1. Do not strive for over precision: recognise that transfer function reconstruction errors are probably underestimates of both within-site and between site variance; and a RMSEP <10% of the sampled elevation range of the modern dataset may be a result of the methodology or approach taken by the researcher.
2. Develop methods to assess reliability of reconstructions, including statistical methods such as dissimilarity coefficients and independent measures of elevation.
3. Recognise, describe and, where possible, quantify the effect and uncertainties of transfer function model choices on reconstructions (Figure 7).
4. Question if the reconstruction makes sense with reference to ecology, lithology and environmental succession.
5. Where possible, use multiple cores from the same site to evaluate findings by replication. If microfossil analyses and dating of multiple cores are too time-consuming or too costly, then ascertain that the stratigraphy of the core location is representative of the study site by documenting the site's lithostratigraphy in detail.
6. Evaluate age model dependence and be cautious where RSL changes appear to reflect the structure of the radiocarbon calibration curve or where inflections coincide with the boundary between different age models or chronological controls.

All time series of sea-level change are inherently noisy, whether from instrumental or geological tide gauges. We seek to identify trends and changes that are smaller in magnitude than daily tides and waves. Just because individual error terms may be larger than the magnitude of some local late Holocene RSL fluctuations does not mean that a RSL reconstruction is not worthwhile. Studies of decadal to centennial-scale sea-level changes should seek to define the trends, linear and non-linear, and their associated standard errors. Even when individual errors appear large, identifying changes in trends at multiple sites, both within an estuary and between different estuaries, allow us to test hypotheses about changes in the rate and direction of RSL.

8. Conclusions

There is no doubt that advances in palaeoenvironmental reconstruction and dating increase our understanding of late Holocene RSL, and subsequently ocean-atmosphere-cryosphere interactions. By following the pathways outlined (Figure 3) it is possible to develop near continuous records of sub-centennial changes in sea level. But, as we show using the north-west Scotland example, the reconstruction is not independent of a series of choices made by the researcher, the most important of which are: the environment from which the record is taken; the modern dataset chosen to calibrate the fossil changes; statistical assumptions behind calibrating microfossil assemblages; and the chosen age model. Adoption of the suggestions set out in this paper may help in the development and interpretation of geological tide gauge records.

For most of the coastlines of the world we can only reconstruct historical sea-level changes on timescales relevant to society by using proxy methods. Salt marsh geological tide gauges are our best option and are needed from multiple sites around the globe. This requires development of different reconstruction models that will inevitably have different vertical uncertainties. The challenge is to know and quantify what those uncertainties really are, and then to develop ways of comparing different records and trends in RSL to constrain some of the most pressing scientific questions that face society today.

Acknowledgements

References

Allen, J.R.L., 1990. Constraints on measurement of sea-level movements from salt-marsh accretion rates. Journal of the Geological Society of London 147, 5-7.

Atwater, B.F., Hemphill-Haley, E., 1997. Recurrence intervals for great earthquakes of the past 3,500 years at northeastern Willapa Bay, Washington. U.S. Geological Survey Professional Paper 1576, 1-108.

Austin, R.M., 1991. Modelling Holocene tides on the NW European continental shelf. Terra Nova 3, 276-288.

Bard, E., Raisbeck, G.M., Yiou, F., Jouzel, J., 1997. Solar modulation of cosmogenic nuclide production over the last millennium: comparison between C-14 and Be-10 records. Earth and Planetary Science Letters 150, 453-462.

Barlow, N.L.M., Shennan, I., Long, A.J., 2012. Relative sea-level response to Little Ice Age ice mass change in south central Alaska: Reconciling model predictions and geological evidence. Earth and Planetary Science Letters 315, 62-75.

Bates, C.R., Moore, C.G., Harries, D.B., Austin, W., Mair, J., 2004. Broad scale mapping of sublittoral habitats in Loch Laxford, Scotland., Scottish Natural Heritage Commissioned Report No. 004 (ROAME No. F01AA401A).

Battarbee, R.W., Jones, V.J., Flower, R.J., Cameron, N.G., Bennion, H., Carvalho, L., Juggins, S., 2001. Diatoms, in: Smol, J.P., Birks, H.J., Last, W.M. (Eds.), Tracking Environmental Change Using Lake Sediments: Volume 3. Terrestrial, Algal and Silaceous Indicators. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 155-202.

Berkeley, A., Perry, C.T., Smithers, S.G., Horton, B.P., Taylor, K.G., 2007. A review of the ecological and taphonomic controls on foraminiferal assemblage development in intertidal environments. Earth-Science Reviews 83, 205-230.

Birks, H.J.B., 1995. Quantitative palaeoenvironmental reconstructions, in: Maddy, D., Brew, J.S. (Eds.), Statistical modelling of Quaternary Science data. Quaternary Research Association., Cambridge, pp. 161-254.

Birks, H.J.B., 1998. Numerical tools in palaeolimnology - progress, potentialities, and problems. Journal of Paleolimnology 20, 307-332.

Birks, H.J.B., 2003. Quantitative paleoenvironmental reconstructions from Holocene biological data, in: Mackay, A.W., Battarbee, R.W., Birks, H.J.B., Oldfield, F. (Eds.), Global Change in the Holocene. Arnold, London, pp. 342-357.

Birks, H.J.B., 2010. Numerical methods for the analysis of diatom assemblage data, in: Smol, J.P., Stoermer, E.F. (Eds.), The Diatoms: Applications for the Environmental and Earth Sciences, Second ed. University Press, Cambridge, UK, pp. 23-54.

Blaauw, M., Heuvelink, G.B.M., Mauquoy, D., van der Plicht, J., van Geel, B., 2003. A numerical approach to C-14 wiggle-match dating of organic deposits: best fits and confidence intervals. Quaternary Science Reviews 22, 1485-1500.

Blockley, S.P.E., Blaauw, M., Bronk Ramsey, C., van der Plicht, J., 2007. Building and testing age models for radiocarbon dates in Lateglacial and Early Holocene sediments. Quaternary Science Reviews 26, 1915-1926.

Bradley, S.L., Milne, G.A., Shennan, I., Edwards, R., 2011. An improved glacial isostatic adjustment model for the British Isles. Journal of Quaternary Science 26, 541-552.

Bradshaw, J.S., 1968. Environmental parameters and marsh foraminifera. Limnology and Oceanography 13, 26-&.

Brain, M.J., Long, A.J., Woodroffe, S.A., Petley, D.N., Milledge, D.G., Parnell, A.C., 2012. Modelling the effects of sediment compaction on salt marsh reconstructions of recent sea-level rise. Earth and Planetary Science Letters 345, 180-193.

Bronk Ramsey, C., 2001. Development of the radiocarbon calibration program OxCal. Radiocarbon 43, 355-363.

Brooks, A.J., Bradley, S.L., Edwards, R.J., Milne, G.A., Horton, B., Shennan, I., 2008. Postglacial relative sea-level observations from Ireland and their role in glacial rebound modelling. Journal of Quaternary Science 23, 175-192.

Brooks, S.J., Birks, H.J.B., 2001. Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. Quaternary Science Reviews 20, 1723-1741.

Cahoon, D.R., Perez, B.C., Segura, B.D., Lynch, J.C., 2011. Elevation trends and shrink–swell response of wetland soils to flooding and drying. Estuarine, Coastal and Shelf Science 91, 463-474.

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

Charman, D.J., Gehrels, W.R., Manning, C., Sharma, C., 2010. Reconstruction of recent sea-level change using testate amoebae. Quaternary Research 73, 208-219.

Cooper, S., Gaiser, E., Wachnicka, A., 2010. Estuarine paleoenvironmental reconstructions using diatoms, in: Smol, J.P., Stoermer, E.F. (Eds.), The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge University Press, UK, pp. 324-345.

Cox, R., Wadsworth, R.A., Thomson, A.G., 2003. Long-term changes in salt marsh extent affected by channel deepening in a modified estuary. Cont. Shelf Res. 23, 1833-1846.

Cundy, A.B., Croudace, I.W., 1995. Physical and chemical associations of radionuclides and trace metals in estuarine sediments: an example from Poole Harbour, southern England. Journal of Environmental Radioactivity 29, 191-211.

Cundy, A.B., Croudace, I.W., Cearreta, A., Irabien, M.J., 2003. Reconstructing historical trends in metal input in heavily-disturbed, contaminated estuaries: studies from Bilbao, Southampton Water and Sicily. Appl. Geochem. 18, 311-325.

Davies, J.L., 1964. A morphogenic approach to world shorelines. Zeitschrift fur Geomorphologie 8, 27-42.

De Rijk, S., 1995. Agglutinated foraminifera as indicators of salt marsh development in relation to late holocene sea level rise (Great Marshes at Barnstable, Massachusetts). Febo, Utrecht.

Devoy, R.J., 1982. Analysis of the geological evidence for Holocene sea-level movements in southeast England. Proceedings of the Geologists' Association 93, 65-90.

Edwards, R., 2007. Sea levels: resolution and uncertainty. Progress in Physical Geography 31, 621-632.

Edwards, R.J., 2001. Mid- to Late Holocene relative sea-level change in Poole Harbour, southern England. Journal of Quaternary Science 16, 221-235.

Edwards, R.J., 2006. Mid- to late-Holocene relative sea-level change in southwest Britain and the influence of sediment compaction. Holocene 16, 575-587.

Edwards, R.J., Horton, B.P., 2000. Reconstructing relative sea-level change using UK salt-marsh foraminifera. Marine Geology 169, 41-56.

Edwards, R.J., Horton, B.P., 2006. Developing detailed records of relative sea-level change using a foraminiferal transfer function: an example from North Norfolk, UK. Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 364, 973-991.

Edwards, R.J., van de Plassche, O., Gehrels, W.R., Wright, A.J., 2004a. Assessing sea-level data from Connecticut, USA, using a foraminiferal transfer function for tide level. Marine Micropaleontology 51, 239-255.

Edwards, R.J., Wright, A., Van de Plassche, O., 2004b. Surface distributions of salt-marsh foraminifera from Connecticut, USA: modem analogues for high-resolution sea level studies. Marine Micropaleontology 51, 1-21.

Edwards, R.J., Wright, A.J., van de Plassche, O., 2004c. Surface distributions of salt-marsh foraminifera from Connecticut, USA: modern analogues for high-resolution sea level studies. Marine Micropaleontology 51, 1-21.

Engelhart, S.E., Horton, B.P., Douglas, B.C., Peltier, W.R., Tornqvist, T.E., 2009. Spatial variability of late Holocene and 20(th) century sea-level rise along the Atlantic coast of the United States. Geology 37, 1115-1118.

Farías, M., Vargas, G., Tassara, A., Carretier, S., Baize, S., Melnick, D., Bataille, K., 2010. Land-Level Changes Produced by the Mw 8.8 2010 Chilean Earthquake. Science 329, 916.

Freund, H., Gerdes, G., Streif, H., Dellwig, O., Watermann, F., 2004. The indicative meaning of diatoms, pollen and botanical macro fossils for the reconstruction of palaeoenvironments and sea-level fluctuations along the coast of Lower Saxony; Germany. Quaternary International 112, 71-87.

Gasse, F., Juggins, S., Khelifa, L.B., 1995. Diatom-based transfer functions for inferring past hydrochemical characteristics of African lakes. Palaeogeography, Palaeoclimatology, Palaeoecology 117, 31-54.

Gehrels, W.R., 1994. Determining relative sea-level change from salt-marsh foraminifera and plant zones on the coast of Maine, USA. Journal of Coastal Research 10, 990-1009.

Gehrels, W.R., 1999. Middle and late holocene sea-level changes in Eastern Maine reconstructed from foraminiferal saltmarsh stratigraphy and AMS C-14 dates on basal peat. Quaternary Research 52, 350-359.

Gehrels, W.R., 2000. Using foraminiferal transfer functions to produce high-resloution sea-level records from salt-marsh deposits, Maine, USA. The Holocene 10, 367-376.

Gehrels, W.R., Belknap, D.F., Kelley, J.T., 1996. Integrated high-precision analyses of Holocene relative sea-level changes: Lessons from the coast of Maine. Geological Society of America Bulletin 108, 1073-1088.

Gehrels, W.R., Belknap, D.F., Pearce, B.R., Gong, B., 1995. Modelling the contribution of M2 tidal amplification to the Holocene rise of mean high water in the Gulf of Maine and the Bay of Fundy. Marine Geology 124, 71-85.

Gehrels, W.R., Callard, S.L., Moss, P.T., Marshall, W.A., Blaauw, M., Hunter, J., Milton, J.A., Garnett, M.H., 2012. Nineteenth and twentieth century sea-level changes in Tasmania and New Zealand. Earth and Planetary Science Letters 315, 94-102.

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

Gehrels, W.R., Marshall, W.A., Gehrels, M.J., Larsen, G., Kirby, J.R., Eiriksson, J., Heinemeier, J., Shimmield, T., 2006. Rapid sea-level rise in the North Atlantic Ocean since the first half of the nineteenth century. Holocene 16, 949-965.

Gehrels, W.R., Milne, G.A., Kirby, J.R., Patterson, R.T., Belknap, D.F., 2004. Late Holocene sea-level changes and isostatic crustal movements in Atlantic Canada. Quaternary International 120, 79-89.

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

Gehrels, W.R., Woodworth, P.L., in press. When did modern rates of sea-level rise start? Global and Planetary Change.

Godwin, H., 1940. Studies of the Post-Glacial history of the British vegetation. III. Fenland pollen diagrams. IV. Post-Glacial changes of relative land and sea-level in the English Fenland. Philosophical Transactions of the Royal Society of London, B 230, 239-303.

Goldstein, S.T., Harben, E.B., 1993. Taphofacies Implications of Infaunal Foraminiferal Assemblages in a Georgia Salt Marsh, Sapelo Island. Micropaleontology 39, 53-62.

Guilbault, J.P., Clague, J.J., Lapointe, M., 1995. Amount of subsidence during a late Holocene earthquake - evidence from fossil tidal marsh foraminifera at Vancouver Island, west coast of Canada. Palaeogeography, Palaeoclimatology, Palaeoecology 111, 49-71.

Guiot, J., 1990. Methodology of the last climatic cycle reconstruction in France from pollen data. Palaeogeography, Palaeoclimatology, Palaeoecology 80, 49-69.

Guiot, J., de Vernal, A., 2011. Is spatial autocorrelation introducing biases in the apparent accuracy of paleoclimatic reconstructions? Quaternary Science Reviews 30, 1965-1972.

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

Hamilton, S., Shennan, I., Combellick, R., Mulholland, J., Noble, C., 2005. Evidence for two great earthquakes at Anchorage, Alaska and implications for multiple great earthquakes through the Holocene. Quaternary Science Reviews 24, 2050-2068.

Haslett, S.K., Cundy, A.B., Davies, C.F.C., Powell, E.S., Croudace, I.W., 2003. Salt marsh, sedimentation over the past c. 120 years along the West Cotentin Coast of Normandy (France): Relationship to sea-level rise and sediment supply. Journal of Coastal Research 19, 609-620.

Hawkes, A.D., Horton, B.P., Nelson, A.R., 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, 116-140.

Heiri, O., Brooks, S.J., Birks, H.J.B., Lotter, A.F., 2011. A 274-lake calibration data-set and inference model for chironomid-based summer air temperature reconstruction in Europe. Quaternary Science Reviews 30, 3445-3456.

Heyworth, A., Kidson, C., 1982. Sea-level changes in southwest England and Wales. Proceedings of the Geologists' Association 93, 91-111.

Hill, D.F., Griffiths, S.D., Peltier, W.R., Horton, B.P., Tornqvist, T.E., 2011. High-resolution numerical modeling of tides in the western Atlantic, Gulf of Mexico, and Caribbean Sea during the Holocene. Journal of Geophysical Research-Oceans 116.

Hill, T.C.B., Woodland, W.A., Spencer, C.D., Marriott, S.B., 2007. Holocene sea-level change in the Severn Estuary, southwest England: a diatom-based sea-level transfer function for macrotidal settings. The Holocene 17, 639-648.

Hinton, A.C., 1996. Tides in the northeast Atlantic: considerations for modeling water depth changes. Quaternary Science Reviews 15, 873-894.

Horton, B., Shennan, I., 2009. Compaction of Holocene strata and the implications for relative sea level change on the east coast of England. Geology 37, 1083-1086.

Horton, B.P., Corbett, R., Culver, S.J., Edwards, R.J., Hillier, C., 2006. Modern saltmarsh diatom distributions of the Outer Banks, North Carolina, and the development of a transfer function for high resolution reconstructions of sea level. Estuarine Coastal and Shelf Science 69, 381-394.

Horton, B.P., Edwards, R.J., 2005. The application of local and regional transfer functions to the reconstruction of Holocene sea levels, north Norfolk, England. Holocene 15, 216-228.

Horton, B.P., Edwards, R.J., 2006. Quantifying Holocene Sea Level Change Using Intertidal Foraminifera: Lessons from the British Isles. Cushman Foundation for Foraminiferal Research Special Publication v. 40, 97pp.

Horton, B.P., Edwards, R.J., Lloyd, J.M., 1999a. Reconstruction of former sea levels using a foraminiferal-based transfer function. Journal of Foraminiferal Research 29, 117-129.

Horton, B.P., Edwards, R.J., Lloyd, J.M., 1999b. UK intertidal foraminiferal distributions: implications for sea-level studies. Marine Micropaleontology 36, 205-223.

Horton, B.P., Edwards, R.J., Lloyd, J.M., 2000. Implications of a microfossil based transfer function in Holocene sea-level studies., in: Shennan, I., Andrews, J. (Eds.), Holocene Land-ocean interaction and environmental change around the North Sea. Geological Society, London, pp. 41-54.

Horton, B.P., Sawai, Y., 2010. Diatoms as indicators of former sea levels, earthquakes, tsunamis and hurricanes, in: Smol, J.P., Stoermer, E.P. (Eds.), The Diatoms: Applications to the Environmental and Earth Sciences. University Press, Cambridge, UK, pp. 357-372.

Hua, Q., Barbetti, M., 2004. Review of tropospheric bomb C-14 data for carbon cycle modeling and age calibration purposes. Radiocarbon 46, 1273-1298.

Huntley, B., 2012. Reconstructing palaeoclimates from biological proxies: Some often overlooked sources of uncertainty. Quaternary Science Reviews 31, 1-16.

Imbrie, J., Kipp, N.G., 1971. A new micropaleontological method for quantitative paleoclimatology: application to a Late Pleistocene Caribbean core, in: Turekian, K.K. (Ed.), The Late Cenozoic Glacial Ages. Yale University Press, New Haven, pp. 71-181.

Jackson, S.T., Williams, J.W., 2004. Modern analogues in Quaternary Paleoecology: here today, gone yesterday, gone tomorrow? Annual Review of Earth and Planetary Sciences 32, 495-537.

Jacobson, H.A., Jacobson Jr, G.L., 1989. Variability of vegetation in tidal marshes of Maine, U.S.A. Canadian Journal of Botany 67, 230-238.

Jelgersma, S., 1961. Holocene sea-level changes in the Netherlands. Mededelingen Geologische Stichting C-IV 7, 1-100.

Jonasson, K.E., Patterson, R.T., 1992. Preservation potential of salt-marsh foraminifera from the Fraser-River Delta, British-Columbia. Micropaleontology 38, 289-301.

Jordon, R.W., Stickley, C.E., 2010. Diatoms as indicators of palaeoceanographic events, in: Smol, J.P., Stoermer, E.F. (Eds.), The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge University Press, UK, pp. 424-453.

Juggins, S., 2003. C2: A Microsoft Windows program for developing and applying palaeoecological transfer functions and for visualising multi-proxy stratigraphic datasets, 1.4.2 ed. University of Newcastle, Newcastle upon Tyne.

Kemp, A.C., Horton, B.P., Culver, S.J., Corbett, D.R., van de Plassche, O., Gehrels, W.R., Douglas, B.C., Parnell, A.C., 2009a. Timing and magnitude of recent accelerated sea-level rise (North Carolina, United States). Geology 37, 1035-1038.

Kemp, A.C., Horton, B.P., Donnelly, J.P., Mann, M.E., Vermeer, M., Rahmstorf, S., 2011. Climate related sea-level variations over the past two millennia. Proceedings of the National Academy of Sciences of the United States of America 108, 11017-11022.

Kemp, A.C., Horton, B.P., Vann, D.R., Engelhart, S.E., Grand Pre, C.A., Vane, C.H., Nikitina, D., 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., Horton, B.R., Corbett, D.R., Culver, S.J., Edwards, R.J., van de Plassche, O., 2009b. The relative utility of foraminifera and diatoms for reconstructing late Holocene sea-level change in North Carolina, USA. Quaternary Research 71, 9-21.

Lambeck, K., Smither, C., Johnston, P., 1998. Sea-level change, glacial rebound and mantle viscosity for northern Europe. Geophysical Journal International 134, 102-144.

Legendre, P., Fortin, M., 1989. Spatial pattern and ecological analysis. Vegetatio 80, 107-138.

Leorri, E., Cearreta, A., 2009. Recent sea-level changes in the southern Bay of Biscay: transfer function reconstructions from salt-marshes compared with instrumental data. Scientia Marina 73, 287-296.

Leorri, E., Fatela, F., Cearreta, A., Moreno, J., Antunes, C., Drago, T., 2011. Assessing the performance of a foraminifera-based transfer function to estimate sea-level changes in northern Portugal. Quaternary Research 75, 278-287.

Leorri, E., Gehrels, W.R., Horton, B.P., Fatela, F., Cearreta, A., 2010. Distribution of foraminifera in salt marshes along the Atlantic coast of SW Europe: Tools to reconstruct past sea-level variations. Quaternary International 221, 104-115.

Leorri, E., Horton, B.P., Cearreta, A., 2008. Development of a foraminifera-based transfer function in the Basque marshes, N. Spain: Implications for sea-level studies in the Bay of Biscay. Marine Geology 251, 60-74.

Long, A.J., Roberts, D.H., Rasch, M., 2003. New observations on the relative sea level and deglacial history of Greenland from Innaarsuit, Disko Bug. Quaternary Research 60, 162-171.

Long, A.J., Scaife, R.G., Edwards, R.J., 1999. Pine pollen in intertidal sediments from Poole Harbour, UK; implications for late-Holocene sediment accretion rates and sea-level rise. Quaternary International 55, 3-16.

Long, A.J., Scaife, R.G., Edwards, R.J., 2000. Stratigraphic architecture, relative sea level, and models of estuary development in southern England: new data from Southampton Water, in: Pye, K., Allen, J.R.L. (Eds.), Coastal and Estuarine Environments: Sedimentology, Geomorphology and Geoarchaeology. Geological Society, pp. 253-279.

Long, A.J., Waller, M.P., Stupples, P., 2006. Driving mechanisms of coastal change: Peat compaction and the destruction of late Holocene coastal wetlands. Marine Geology 225, 63-84.

Long, A.J., Woodroffe, S.A., Milne, G.A., Bryant, C.L., Simpson, M.J.R., Wake, L.M., 2012. Relative sea-level change in Greenland during the last 700 yrs and ice sheet response to the Little Ice Age. Earth and Planetary Science Letters 315–316, 76-85.

Long, A.J., Woodroffe, S.A., Milne, G.A., Bryant, C.L., Wake, L.M., 2010. Relative sea level change in west Greenland during the last millennium. Quaternary Science Reviews 29, 367-383.

Marshall, W.A., Clough, R., Gehrels, W.R., 2009. The isotopic record of atmospheric lead fall-out on an Icelandic salt marsh since AD 50. Science of the Total Environment 407, 2734-2748.

Marshall, W.A., Gehrels, W.R., Garnett, M.H., Freeman, S., Maden, C., Xu, S., 2007. The use of 'bomb spike' calibration and high-precision AMS C-14 analyses to date salt-marsh sediments deposited during the past three centuries. Quaternary Research 68, 325-337.

Massey, A.C., Gehrels, W.R., Charman, D.J., Milne, G.A., Peltier, W.R., Lambeck, K., Selby, K.A., 2008. Relative sea-level change and postglacial isostatic adjustment along the coast of south Devon, United Kingdom. Journal of Quaternary Science 23, 415-433.

Massey, A.C., Gehrels, W.R., Charman, D.J., White, S.V., 2006a. An intertidal foraminifera-based transfer function for reconstructing Holocene sea-level change in southwest England. Journal of Foraminiferal Research 36, 215-232.

Massey, A.C., Paul, M.A., Gehrels, W.R., Charman, D.J., 2006b. Autocompaction in Holocene coastal back-barrier sediments from south Devon, southwest England, UK. Marine Geology 226, 225-241.

Milne, G.A., Gehrels, W.R., Hughes, C.W., Tamisiea, M.E., 2009. Identifying the causes of sea-level change. Nature Geoscience 2, 471-478.

Milne, G.A., Mitrovica, J.X., Schrag, D.P., 2002. Estimating past continental ice volume from sea-level data. Quaternary Science Reviews 21, 361-376.

Mitrovica, J.X., Milne, G.A., 2002. On the origin of late Holocene sea-level highstands within equatorial ocean basins. Quaternary Science Reviews 21, 2179-2190.

Mitrovica, J.X., Milne, G.A., 2003. On post-glacial sea level: I. General theory. Geophysical Journal International 154, 253-267.

Mitrovica, J.X., Tamisiea, M.E., Davis, J.L., Milne, G.A., 2001. Recent mass balance of polar ice sheets inferred from patterns of global sea-level change. Nature 409, 1026-1029.

Murray, J.W., 2000. The enigma of the continued use of total assemblages in ecological studies of benthic foraminifera. The Journal of Foraminiferal Research 30, 244-245.

Nelson, A.R., Shennan, I., Long, A.J., 1996. Identifying coseismic subsidence in tidal-wetland stratigraphic sequences at the Cascadia subduction zone of western North America. Journal Of Geophysical Research-Solid Earth 101, 6115-6135.

Patterson, R.T., Gehrels, W.R., Belknap, D.F., Dalby, A.P., 2004. The distribution of salt marsh foraminifera at Little Dipper Harbour New Brunswick, Canada: applicable transfer functions in sea-level research. Quaternary International 120, 185-194.

Peltier, W.R., 2004. Global glacial isostasy and the surface of the ice-age earth: The ICE-5G (VM2) model and GRACE. Annual Review of Earth and Planetary Sciences 32, 111-149.

Phleger, F.B., 1966. Depth patterns of benthonic Foramin-ifera in the eastern Pacific. Prog Oceanogr 3, 273-287.

Pirazzoli, P.A., 1991. World Atlas of Holocene Sea Level Changes. Elsevier Science Publishers B.V., Amsterdam.

Plafker, G., 1969. Tectonics of the March 27, 1964, Alaska earthquake. U.S. Geological Survey Professional Paper 543-I, 74.

Plag, H.P., 2006. Recent relative sea-level trends: an attempt to quantify the forcing factors. Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 364, 821-844.

Preuss, H., 1979. Progress in computer evaluation of sea-level data within the IGCP project no. 61, in: Flexor, I. (Ed.), International Symposium on Coastal Evolution in the Quaternary, Sao-Paulo, Brazil, pp. 104-134.

Reavie, E.D., Juggins, S., 2011. Exploration of sample size and diatom-based indicator performance in three North American phosphorus training sets. Aquat Ecol 45, 529-538.

Reimer, P.J., Baillie, M.G.L., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., Buck, C.E., Burr, G.S., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Hajdas, I., Heaton, T.J., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., McCormac, F.G., Manning, S.W., Reimer, R.W., Richards, D.A., Southon, J.R., Talamo, S., Turney, C.S.M., van der Plicht, J., Weyhenmeye, C.E., 2009. IntCal09 and Marine09 radiocarbon age calibration curves, 0-50,000 years cal BP. Radiocarbon 51, 1111-1150.

Revets, S.A., 2004. On confidence intervals from micropalaeontological counts. Journal of Micropalaeontology 23, 61-65.

Roe, H.M., Doherty, C.T., Patterson, R.T., Swindles, G.T., 2009. Contemporary distributions of saltmarsh diatoms in the Seymour-Belize Inlet Complex, British Columbia, Canada: Implications for studies of sea-level change. Marine Micropaleontology 70, 134-150.

Rossi, V., Horton, B.P., Corbett, D.R., Leorri, E., Perez-Belmonte, L., Douglas, B.C., 2011. The application of foraminifera to reconstruct the rate of 20th century sea level rise, Morbihan Golfe, Brittany, France. Quaternary Research 75, 24-35.

Ryves, D.B., Battarbee, R.W., Fritz, S.C., 2009. The dilemma of disappearing diatoms: Incorporating diatom dissolution data into palaeoenvironmental modelling and reconstruction. Quaternary Science Reviews 28, 120-136.

Ryves, D.B., Juggins, S., Fritz, S.C., Battarbee, R.W., 2001. Experimental dissolution and the quantification of microfossil preservation in sediments. Palaeogeography, Palaeoclimatology, Palaeoecology 172, 99-113.

Sawai, Y., 2001. Distribution of living and dead diatoms in tidal wetlands of northern Japan: relations to taphonomy. Palaeogeography, Palaeoclimatology, Palaeoecology 173, 125-141.

Sawai, Y., Horton, B.P., Nagumo, T., 2004a. The development of a diatom-based transfer function along the Pacific coast of eastern Hokkaido, northern Japan - An aid in paleoseismic studies of the Kuril subduction zone. Quaternary Science Reviews 23, 2467-2483.

Sawai, Y., Nasu, H., Yasuda, Y., 2002. Fluctuations in relative sea-level during the past 3000 yr in the Onnetoh estuary, Hokkaido, northern Japan. Journal of Quaternary Science 17, 607-622.

Sawai, Y., Satake, K., Kamataki, T., Nasu, H., Shishikura, M., Atwater, B.F., Horton, B.P., Kelsey, H.M., Nagumo, T., Yamaguchi, M., 2004b. Transient uplift after a 17th-Century earthquake along the Kuril subduction zone. Science 306, 1918-1920.

Scott, D.B., Medioli, F.S., 1980. Quantitative studies of marsh foraminiferal distributions in Nova Scotia: Implications for Sea Level Studies. Cushman Foundation for Foraminiferal Research, Washington.

Scott, D.B., Medioli, F.S., Duffett, T.E., 1984. Holocene rise of relative sea level at Sable Island, Nova Scotia, Canada. Geology 12, 173-176.

Shennan, I., 1982. Interpretation of Flandrian sea-level data from the Fenland, England. Proceedings of the Geologists' Association 93, 53-63.

Shennan, I., 1999. Global meltwater discharge and the deglacial sea-level record from northwest Scotland. Journal Of Quaternary Science 14, 715-719.

Shennan, I., Coulthard, T., Flather, R., Horton, B., Macklin, M., Rees, J., Wright, M., 2003. Integration of shelf evolution and river basin models to simulate Holocene sediment dynamics of the Humber Estuary during periods of sea-level change and variations in catchment sediment supply. The Science of the Total Environment 314-316, 737-754.

Shennan, I., Hamilton, S., 2006. Coseismic and pre-seismic subsidence associated with great earthquakes in Alaska. Quaternary Science Reviews 25, 1-8.

Shennan, I., Hamilton, S., Hillier, C., Woodroffe, S., 2005. A 16,000-year record of near-field relative sea-level changes, northwest Scotland, United Kingdom. Quaternary International 133-34, 95-106.

Shennan, I., Horton, B.P., 2002. Holocene land- and sea-level changes in Great Britain. Journal of Quaternary Science 17, 511-526.

Shennan, I., Innes, J.B., Long, A.J., Zong, Y., 1995. Holocene relative sea level changes and coastal vegetation history at Kentra Moss, Argyll, Northwest Scotland. Journal of Quaternary Science 9, 261-283.

Shennan, I., Lambeck, K., Horton, B., Innes, J., Lloyd, J., McArthur, J., Purcell, T., Rutherford, M., 2000a. Late Devensian and Holocene records of relative sea-level changes in northwest Scotland and their implications for glacio-hydro-isostatic modelling. Quaternary Science Reviews 19, 1103-1135.

Shennan, I., Lambeck, K., Horton, B.P., Innes, J.B., Lloyd, J.M., McArthur, J.J., Rutherford, M.M., 2000b. Holocene isostasy and relative sea-level changes on the east coast of England, in: Shennan, I., Andrews, J. (Eds.), Holocene land-ocean interaction and environmental change around the North Sea. Geological Society, Special Publications, 166, London, pp. 275-298.

Shennan, I., Long, A.J., Rutherford, M.M., Green, F.M., Innes, J.B., Lloyd, J.M., Zong, Y., Walker, K., 1996. Tidal marsh stratigraphy, sea level change and large earthquakes, I: A 5000 year record in Washington USA. Quaternary Science Reviews 15, 1023-1059.

Shennan, I., Milne, G., Bradley, S., 2012. Late Holocene vertical land motion and relative sea-level changes: lessons from the British Isles. Journal of Quaternary Science 27, 64-70.

Shennan, I., Tooley, M.J., Davis, M.J., Haggart, B.A., 1983. Analysis and interpretation of Holocene sea level data. Nature 302, 404-406.

Shennan, I., Woodworth, P.L., 1992. A comparison of late Holocene and twentieth-century sea-level trends from the UK and North Sea region. Geophysical Journal International 109, 96-105.

Sherrod, B.L., 1999. Gradient analysis of diatom assemblages in a Puget Sound salt marsh: can such assemblages be used for quantitative paleoecological reconstructions? Palaeogeography Palaeoclimatology Palaeoecology 149, 213-226.

Sherrod, B.L., Bucknam, R.C., Leopold, E.B., 2000. Holocene relative sea level changes along the Seattle fault at Restoration Point. Quaternary Research 54, 384-393.

Southall, K.E., Gehrels, W.R., Hayward, B.W., 2006. Foraminifera in a New Zealand salt marsh and their suitability as sea-level indicators. Marine Micropaleontology 60, 167-179.

Streif, H., 1979. Cyclic formation of coastal deposits and their indications of vertical sea-level changes. Oceanus 5, 303-306.

Stuiver, M., Quay, P.D., 1980. Patterns of atmopsheric C-14 changes. Radiocarbon 22, 166-176.

Sullivan, M.J., Currin, C.A., 2000. Community structure and functional dynamics of benthic microalgae in salt marshes.

Szkornik, K., Gehrels, W.R., Kirby, J.R., 2006. Salt-marsh diatom distributions in Ho Bugt (western Denmark) and the development of a transfer function for reconstructing Holocene sea-level changes. Marine Geology 235, 137-150.

Tamisiea, M.E., Mitrovica, J.X., Milne, G.A., Davis, J.L., 2001. Global geoid and sea level changes due to present-day ice mass fluctuations. Journal of Geophysical Research-Solid Earth 106, 30849-30863.

Teferle, F.N., Bingley, R.M., Orliac, E.J., Williams, S.D.P., Woodworth, P.L., McLaughlin, D., Baker, T.F., Shennan, I., Milne, G.A., Bradley, S.L., Hansen, D.N., 2009. Crustal motions in Great Britain: evidence from continuous GPS, absolute gravity and Holocene sea level data. Geophysical Journal International 178, 23-46.

Telford, R.J., Andersson, C., Birks, H.J.B., Juggins, S., 2004. Biases in the estimation of transfer function prediction errors. Paleoceanography 19.

Telford, R.J., Birks, H.J.B., 2005. The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. Quaternary Science Reviews 24, 2173-2179.

Telford, R.J., Birks, H.J.B., 2009. Evaluation of transfer functions in spatially structured environments. Quaternary Science Reviews 28, 1309-1316.

Telford, R.J., Birks, H.J.B., 2011. Effect of uneven sampling along an environmental gradient on transfer-function performance. Journal of Paleolimnology 46, 99-106.

ter Braak, C.J.F., Juggins, S., 1993. Weighted averaging partial least squared regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269/270, 485-502.

ter Braak, C.J.F., Juggins, S., Birks, H.J.B., van de Voet, H., 1993. Weighted averaging partial least squared regression (WA-PLS): definition and comparison with other methods for species-environmental calibration, in: Patil, G.P., Rac, C.R. (Eds.), Multivariate Environmental Statistics. Elsevier Science Publishers, Amsterdam, pp. 525-560.

ter Braak, C.J.F., Prentice, I.C., 1988. A theory of gradient analysis. Advances in Ecological Research 18, 271-317.

ter Braak, C.J.F., Smilauer, P., 2002. CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca, NY, USA.

Thomas, E., Varekamp, J.C., 1991. Paleo-Environmental Analyses of Marsh Sequences (Clinton, Connecticut): Evidence for Punctuated Rise in Relative Sealevel During the Latest Holocene. Journal of Coastal Research 11, 125-158.

Tooley, M.J., 1978. Sea-level changes in north-west England during the Flandrian Stage. Clarendon Press, Oxford.

Tooley, M.J., 1982. Sea-level changes in northern England. Proceedings of the Geologists' Association 93, 43-51.

Tornqvist, T.E., Van Ree, M.H.M., Van 't Veer, R., Van Geel, B., 1998. Improving methodology for high-resolution reconstruction of sea-level rise and neotectonics by palaeoecological analysis and AMS 14C dating of basal peats. Quaternary Research 49, 72-85.

Tornqvist, T.E., Wallace, D.J., Storms, J.E.A., Wallinga, J., Van Dam, R.L., Blaauw, M., Derksen, M.S., Klerks, C.J.W., Meijneken, C., Snijders, E.M.A., 2008. Mississippi Delta subsidence primarily caused by compaction of Holocene strata. Nature Geoscience 1, 173-176.

Turner, R.E., Milan, C.S., Swenson, E.M., 2006. Recent volumetric changes in salt marsh soils. Estuarine Coastal and Shelf Science 69, 352-359.

Uehara, K., Scourse, J.D., Horsburgh, K.J., Lambeck, K., Purcell, A.P., 2006. Tidal evolution of the northwest European shelf seas from the Last Glacial Maximum to the present. Journal of Geophysical Research-Oceans 111.

van de Plassche, O., 1982. Sea-level change and water movements in the Netherlands during the Holocene. Mededelingen Rijks Geologische Dienst 36, 1-93.

van de Plassche, O., 1986. Sea-Level Research: a manual for the collection and evaluation of data. GeoBooks, Norwich.

van de Plassche, O., Edwards, R.J., van der Borg, K., de Jong, A.F.M., 2001. C-14 wiggle-match dating in high-resolution sea-level research. Radiocarbon 43, 391-402.

van de Plassche, O., van der Borg, K., de Jong, A.F.M., 1998. Sea level-climate correlation during the past 1400 yr. Geology 26, 319-322.

Varekamp, J.C., Thomas, E., van de Plassche, O., 1992. Relative Sea-Level Rise and Climate Change Over the Last 1500 Years. Terra Nova 4, 293-304.

Vos, P.C., de Wolf, H., 1988. Methodological aspects of palaeo-ecological diatom research in coastal areas of the Netherlands. Geologie en Mijnbouw 67, 31-40.

Vos, P.C., de Wolf, H., 1993. Diatoms as a tool for reconstructing sedimentary environments in coastal wetlands: methodological aspects. Hydrobiologia 269/270, 285-296.

Wang, Z.B., Jeuken, M.C.J.L., Gerritsen, H., de Vriend, H.J., Kornman, B.A., 2002. Morphology and asymmetry of the vertical tide in the Westerschelde estuary. Cont. Shelf Res. 22, 2599-2609.

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

WOCE, 2002. WOCE Data Products Committee. 2002. WOCE Global Data, Version 3.0, WOCE International Project Office, WOCE Report No. 180/02, Southampton, UK.

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, 215-223.

Woodroffe, S.A., Long, A.J., 2009. Salt marshes as archives of recent relative sea level change in West Greenland. Quaternary Science Reviews 28, 1750-1761.

Woodroffe, S.A., 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, 91-103.

Woodworth, P., Teferle, F., Bingley, R., Shennan, I., Williams, S., 2009. Trends in UK mean sea level revisited. Geophysical Journal International 176, 19-30.

Woodworth, P.L., Tsimplis, M.N., Flather, R.A., Shennan, I., 1999. A review of the trends observed in British Isles mean sea level data measured by tide gauges. Geophysical Journal International 136, 651-670.

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

Yeloff, D., Bennett, K.D., Blaauw, M., Mauquoy, D., Sillasoo, U., van der Plicht, J., Geel, B., 2006. High precision C-14 dating of Holocene peat deposits: A comparison of Bayesian calibration and wiggle-matching approaches. Quaternary Geochronology 1, 222-235.

Zong, Y., Horton, B.P., 1998. Diatom zones across intertidal flats and coastal saltmarshes in Britain. Diatom Research 13, 375-394.

Zong, Y., 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, 153-167.
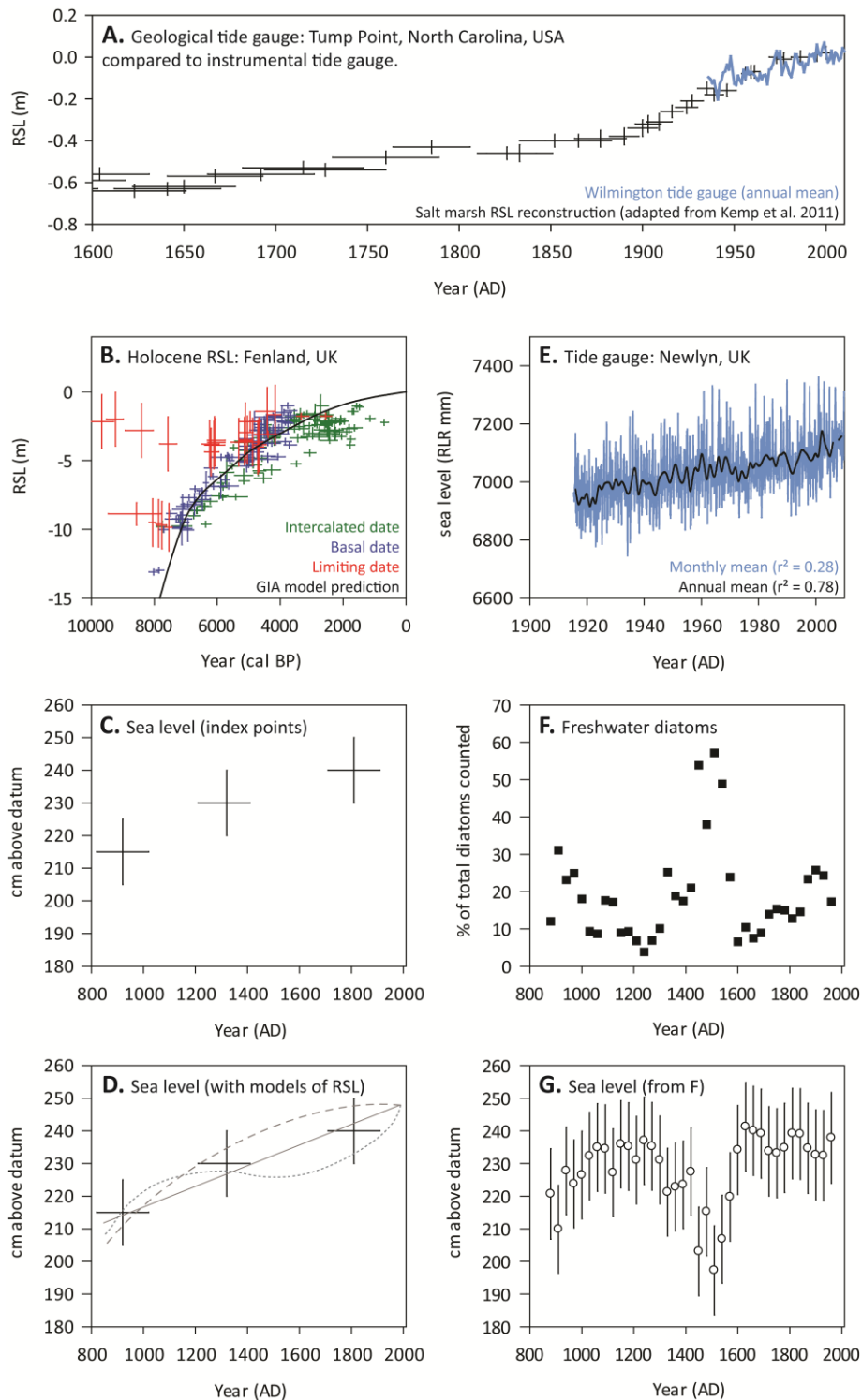
Figure 1 – Different approaches to estimating past sea-level changes. A: Salt marsh foraminifera based reconstruction of relative sea-level change in North Carolina, adapted from Kemp et al. (2011), compared to nearby instrumental tide gauge at Wilmington, South Carolina (NOAA). B: Holocene sea level index points from Fenland, UK compared to glacial isostatic adjustment (GIA) model prediction of RSL (after Shennan et al., 2012). C: Age-altitude plot of sea level index points (simulated dataset for illustration). D: Age-altitude plot of sea level index points, from 'C', with different models of RSL change between data points (simulated dataset for illustration). E: Monthly and annual mean instrumental tide gauge data from Newlyn, UK (psmsl.org.uk). F: Percentage of freshwater diatoms of total number of diatoms counted from samples down a core from ~AD 800 to AD 2000 (simulated dataset for illustration). G: Sea level reconstructions derived from diatom data in 'F' (simulated dataset for illustration).
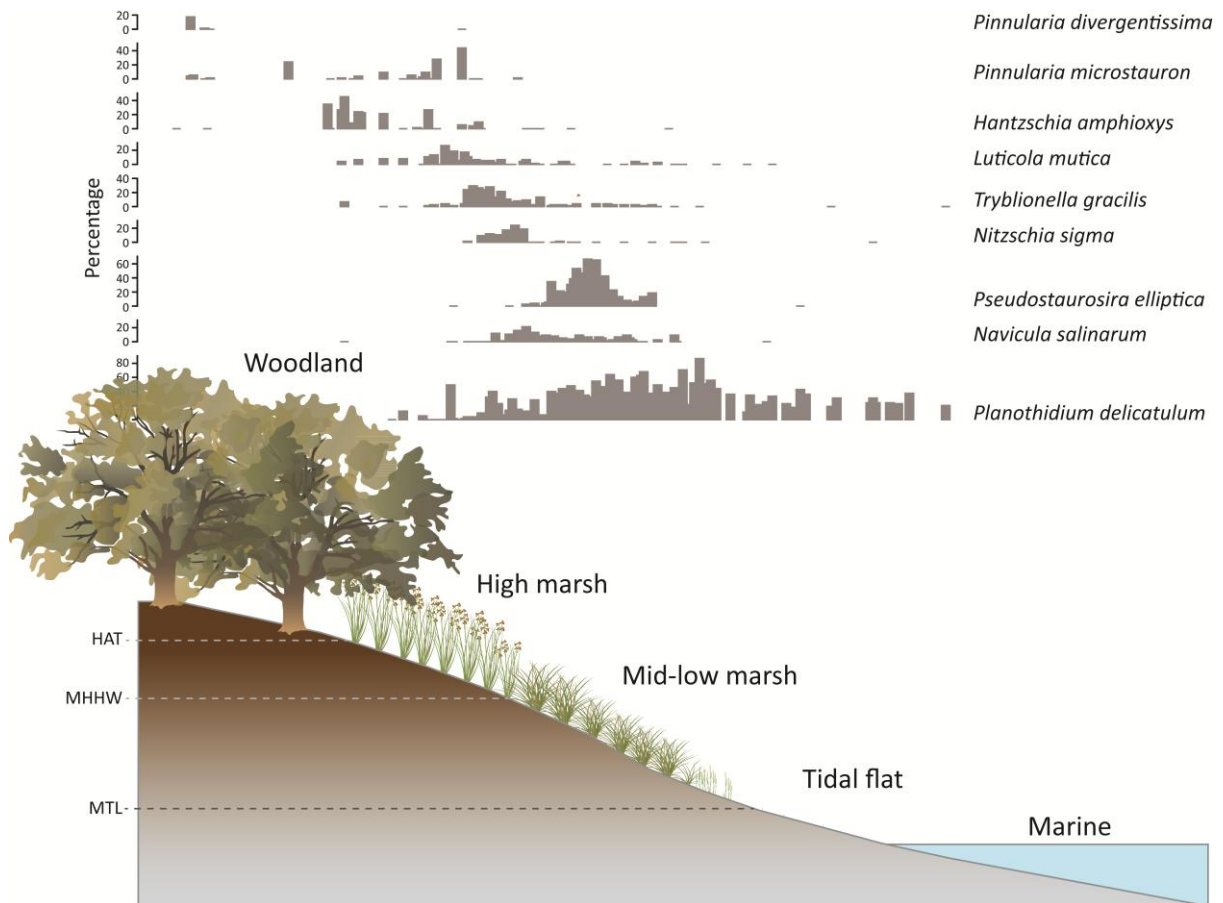
Figure 2 - A schematic salt marsh transect showing the vegetation zones of a modern salt marsh and reference water levels (HAT = highest astronomical tide; MHHW = mean higher high water; MTL = mean tide level). The graphs above show the distribution of selected diatom species across the marsh based on percentage assemblage data from the regional Scotland dataset presented in the paper. Different species of diatoms, and other salt marsh organisms, have differing distributions controlled by tidal inundation and salinity which relates to elevation.
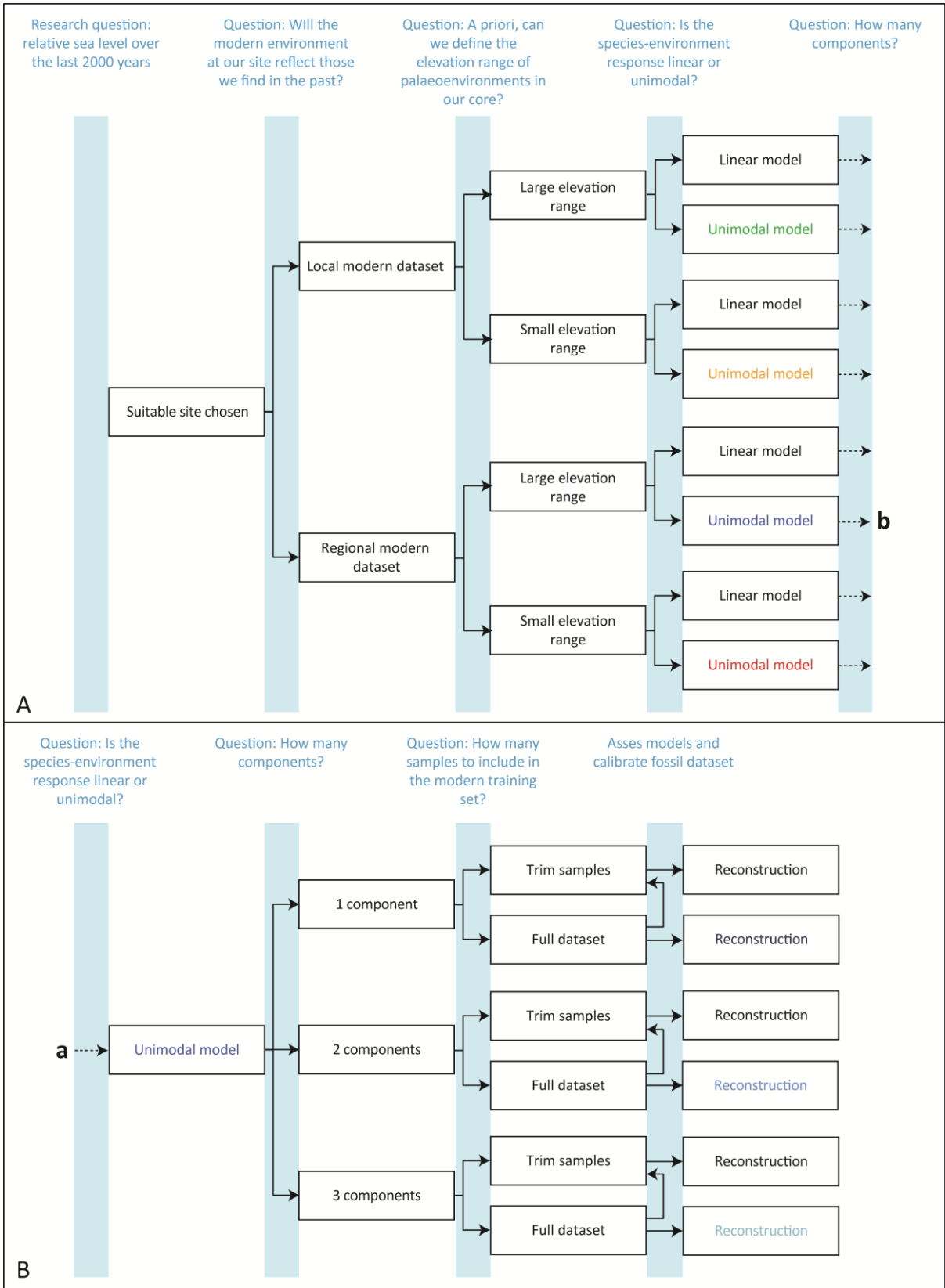
Figure 3 - Decision tree showing the typical pathways to microfossil based reconstruction of RSL change. The decision tree in B applies to any (WA-PLS) unimodal model route in A, but only one pathway is shown for clarity.
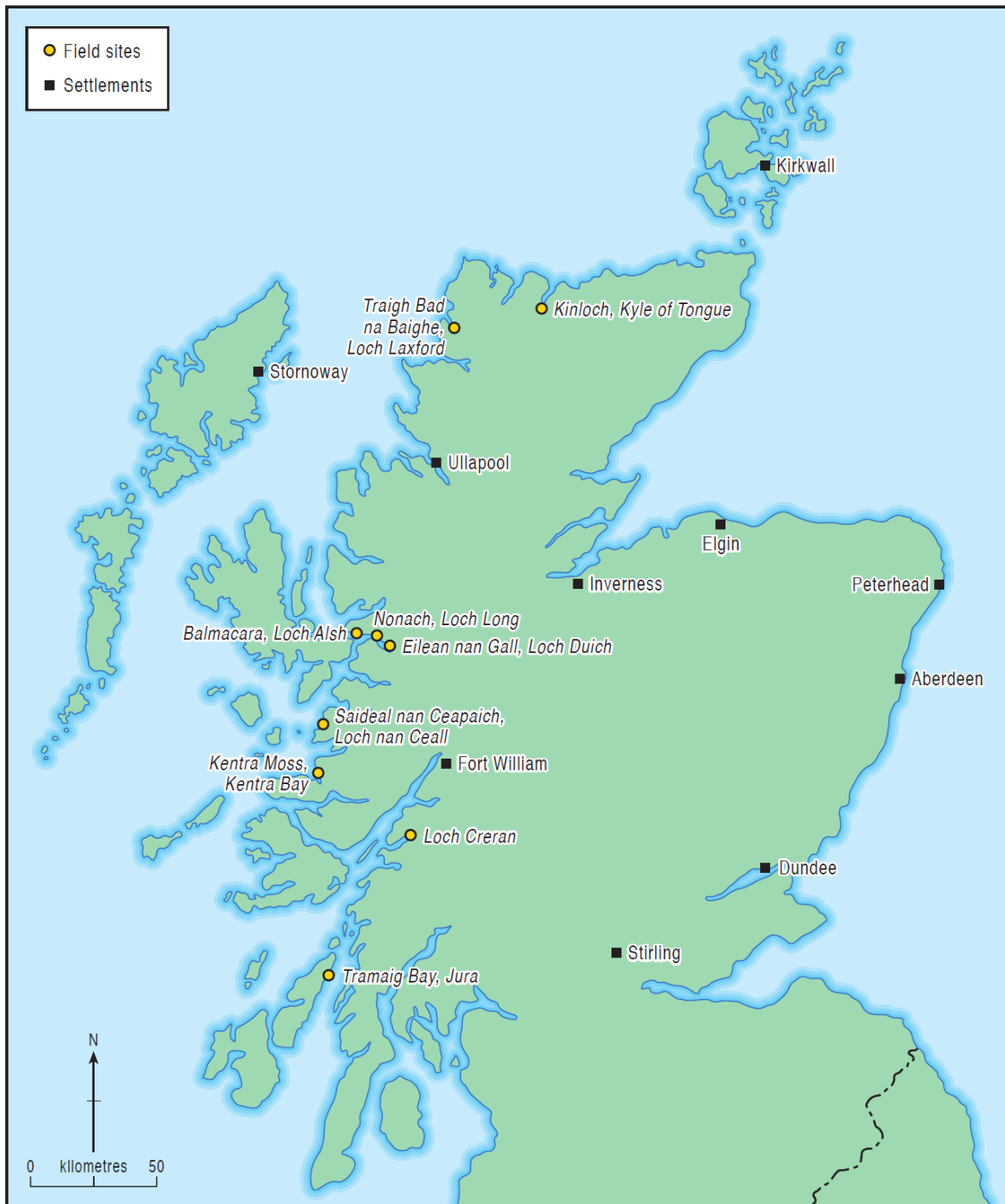
Figure 4 - Map of Scotland showing the location of the sites of modern diatom data used in the Scotland regional transfer function (as listed in Table 1) and Tràigh Bad na Bàighe, Loch Laxford, the location of the fossil core. Location of tidal ports, proximal to field sites, given in Table 1.
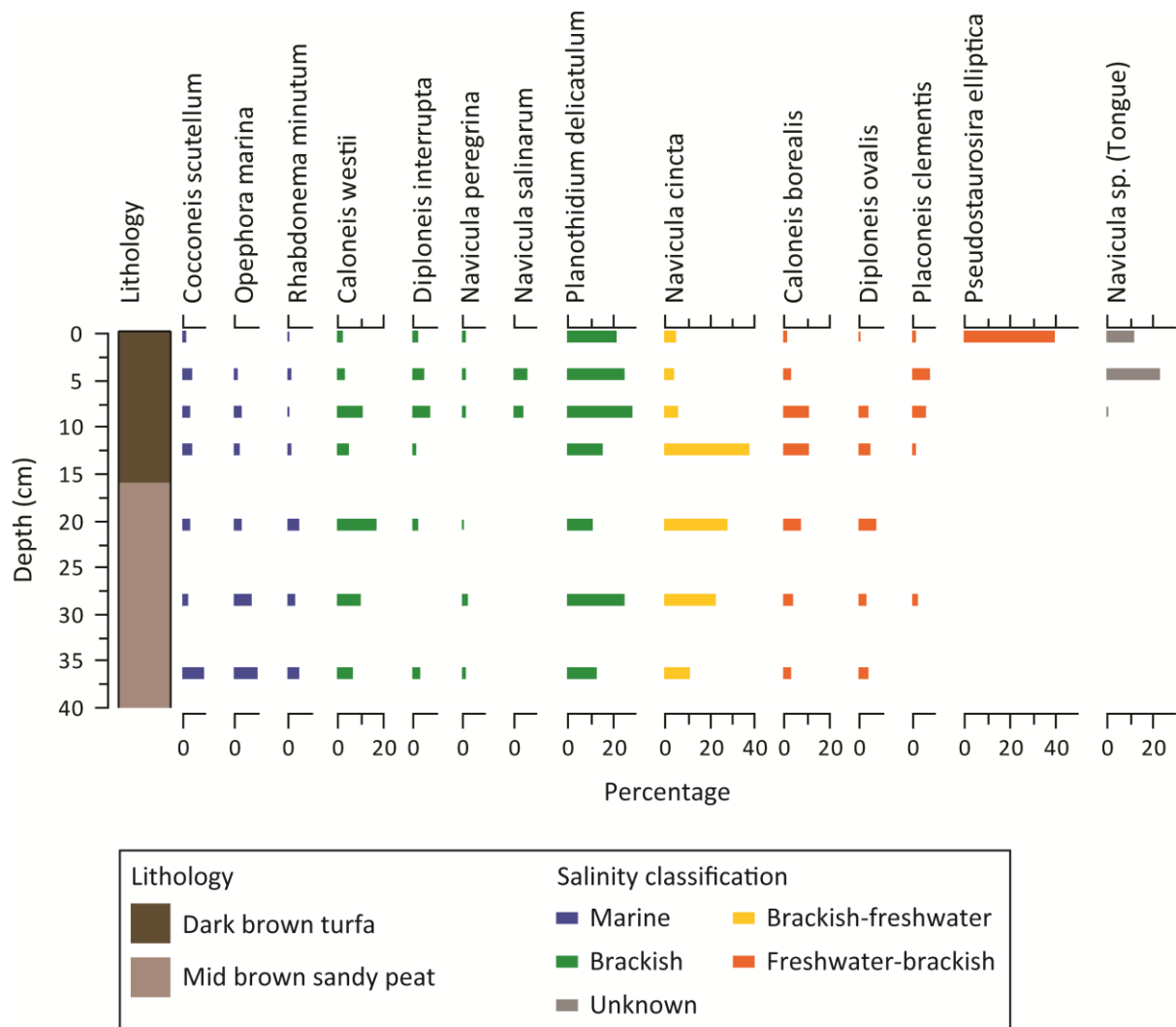
Figure 5 - Lithology and diatom biostratigraphy (showing diatoms that account for ≥5% total diatom valves counted) in a core from Loch Laxford used to demonstrate the consequence of different decision pathways in Figure 3. The diatoms are grouped based on their salinity preferences.
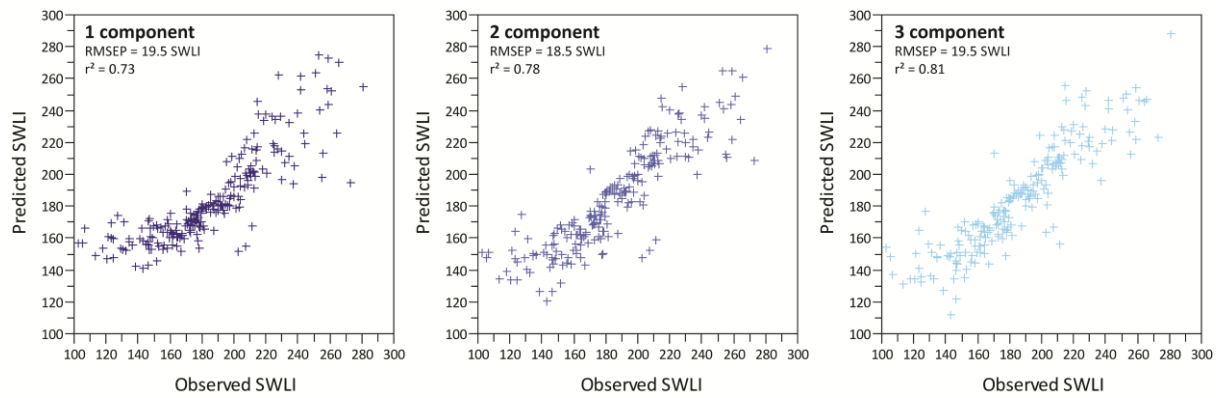
Figure 6 - Scatterplots of the observed standarised water level index (SWLI) against the WA-PLS transfer function model predicted SWLI for the regional Scotland 'coastal transition' diatom training set for the first three components (as detailed in Table 2). The one component model shows a non-linear distortion in the lower part of the elevation range. The two and three component models uses the structure in the residuals to produce a better statistical fit across the full elevation range of the modern data. The colours match the decision pathways in Figure 3B and the reconstruction results in Figure 7.
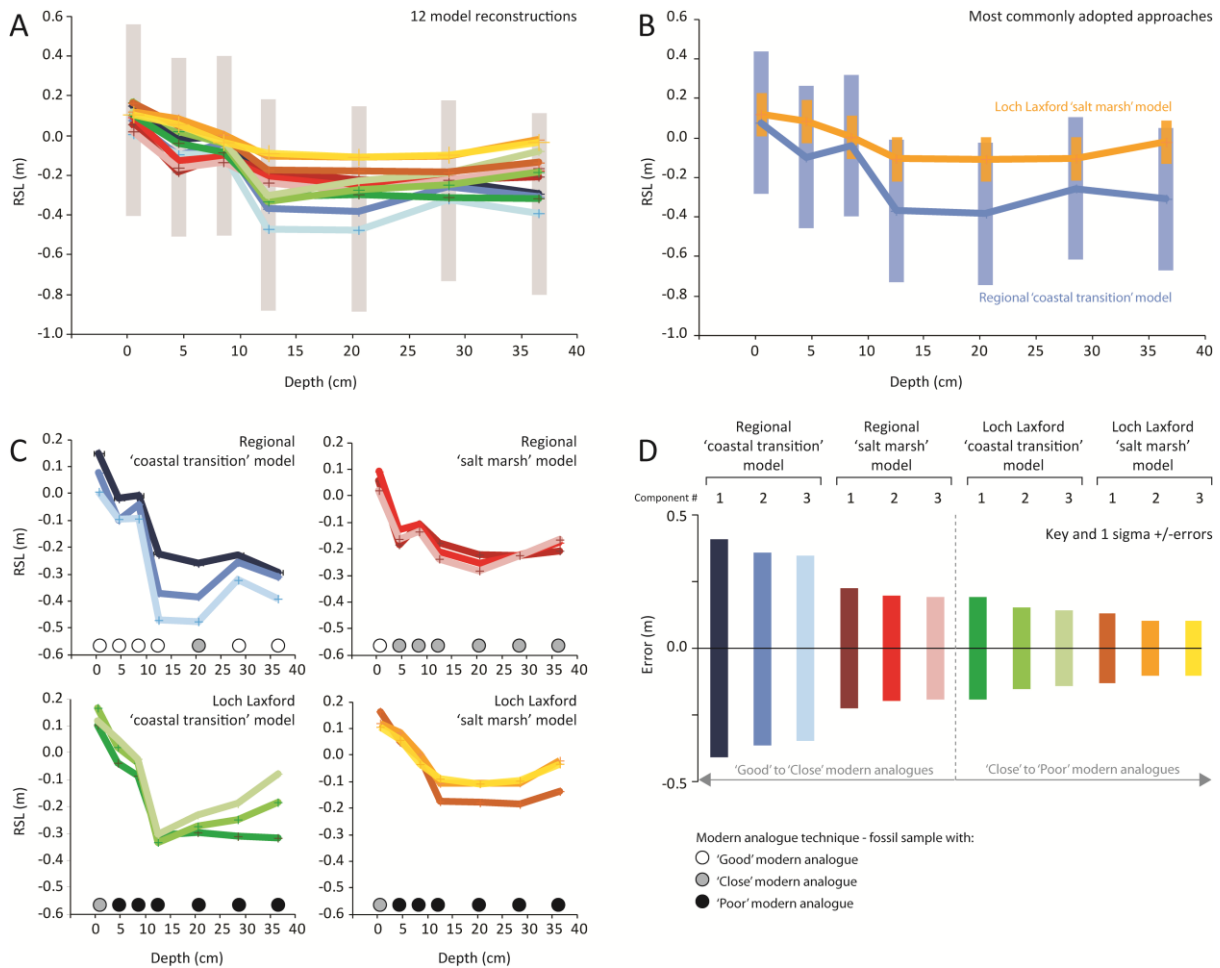
Figure 7 - The results of reconstruction of RSL change in the Loch Laxford core (Figure 4) based on the four modern datasets and the first three WA-PLS components following the unimodal decision pathways in Figure 3. A - the result of all 12 model reconstructions with the largest model errors bars (in grey) from the regional 'coastal transition' one component model. B - the reconstruction results for the two most commonly adopted approaches: a regional dataset with a long environmental gradient (in blue) and a local dataset with a restricted environmental gradient (in orange). Both reconstructions use two component models based on the best combination of RMSEP and r_{boot}^2 statistics in Table 3. C - the reconstruction results from A, grouped by model to show the impact of selecting different WA-PLS components. The white, black and grey circles show which samples have a 'good', 'close' and 'poor' modern analogue based on the modern analogue technique (MAT) thresholds detailed in section 3.7 and Table 2. D - the average ±1σ error bars for all the reconstructions in C, with the colours of each bar providing the key for the results in all the other graphs.
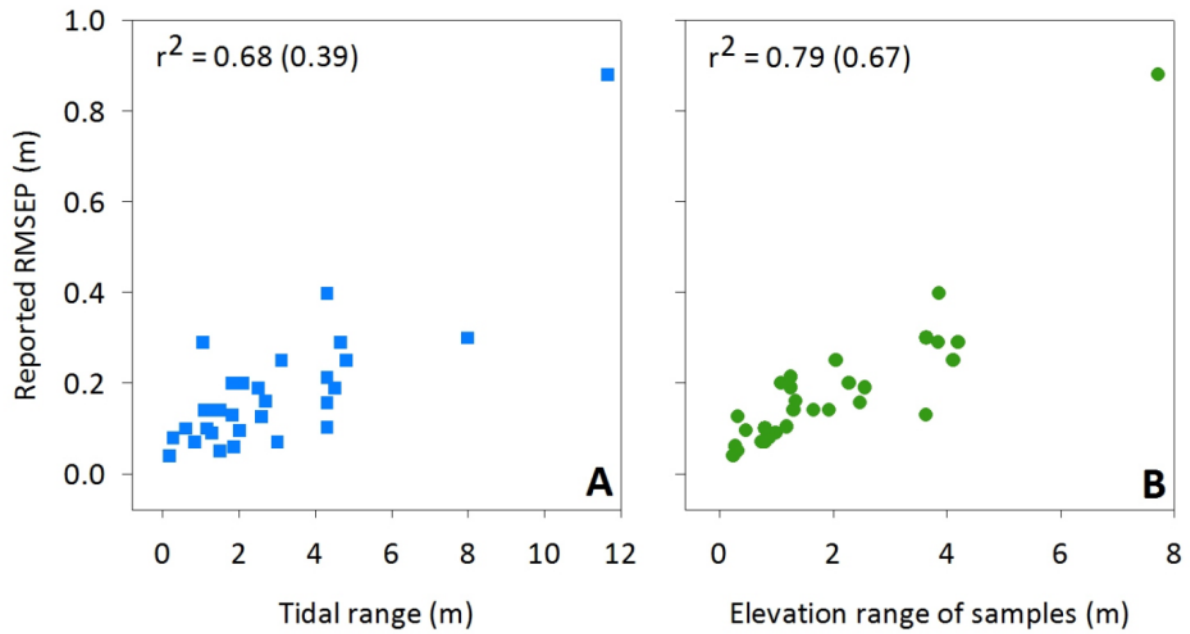
39

Figure 8 - Relationship between transfer function model performance and tidal range (A), and elevation range of samples in the model (B). Model performance shown by root mean squared error of prediction (RMSEP) of 30 transfer function models (26 published, 4 from this paper) (detailed in Table 3). Values in brackets is the r^2 excluding the Bristol Channel dataset (top right data point in both panels) (Hill et al., 2007) which is an outlier relative to the rest of the datasets.
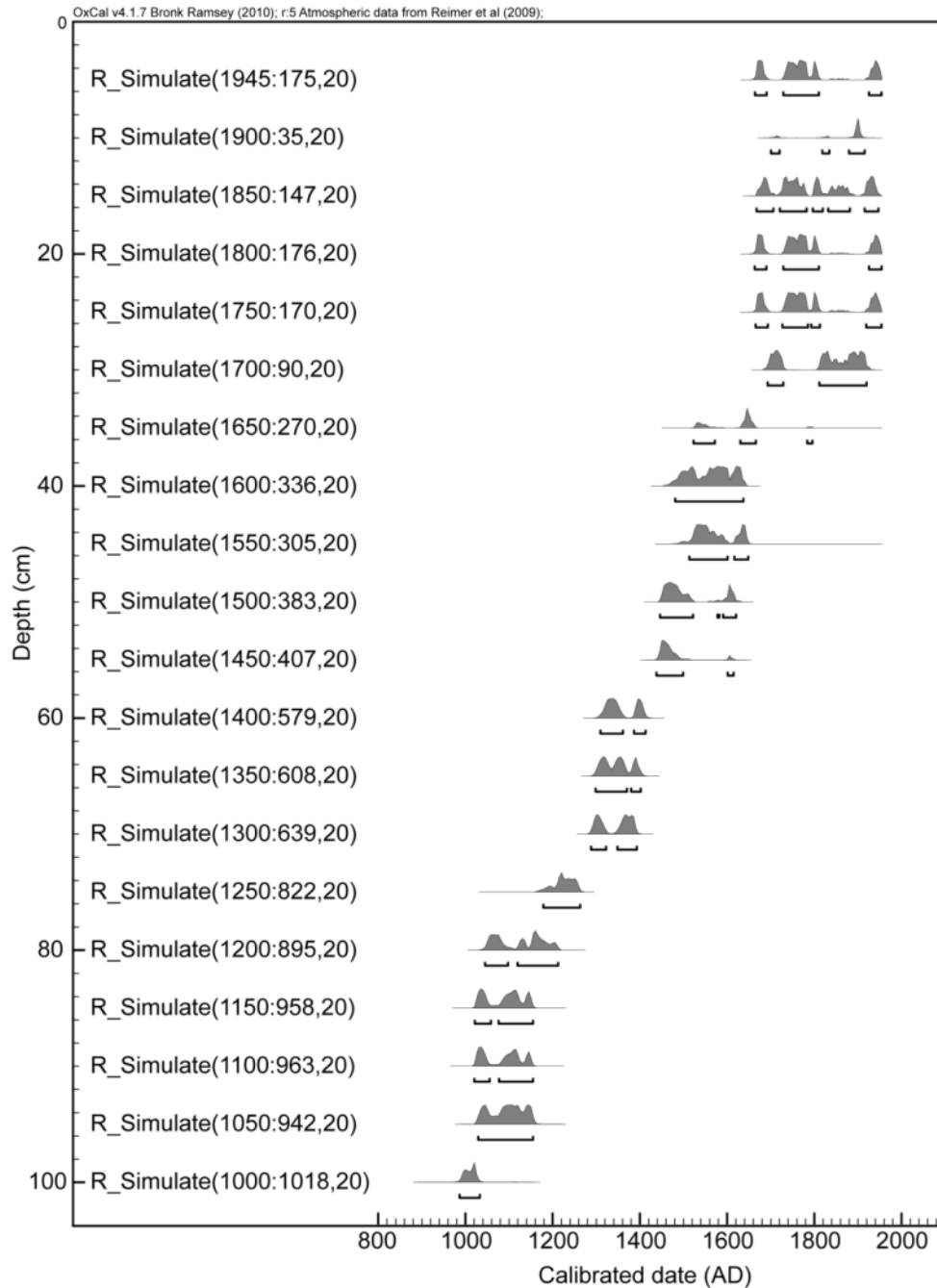
Figure 9 - A series of simulated radiocarbon dates evenly distributed through a core, at 50 calendar year intervals for the last 1000 years, produced by OxCal (Bronk Ramsey, 2001). Shown are probability distributions and 2σ ranges. Note how the shape of the IntCal09 calibration curve (Reimer et al., 2009) can result in a series of stacked calibration solutions, which would appear as periods of rapid sediment accumulation and therefore suggest apparent changes in the rate of RSL, when in fact the sedimentation rate of the sequence is constant. The fluctuations are simply a function of changes in atmospheric [14]C. Updated from Gehrels et al. (2005).
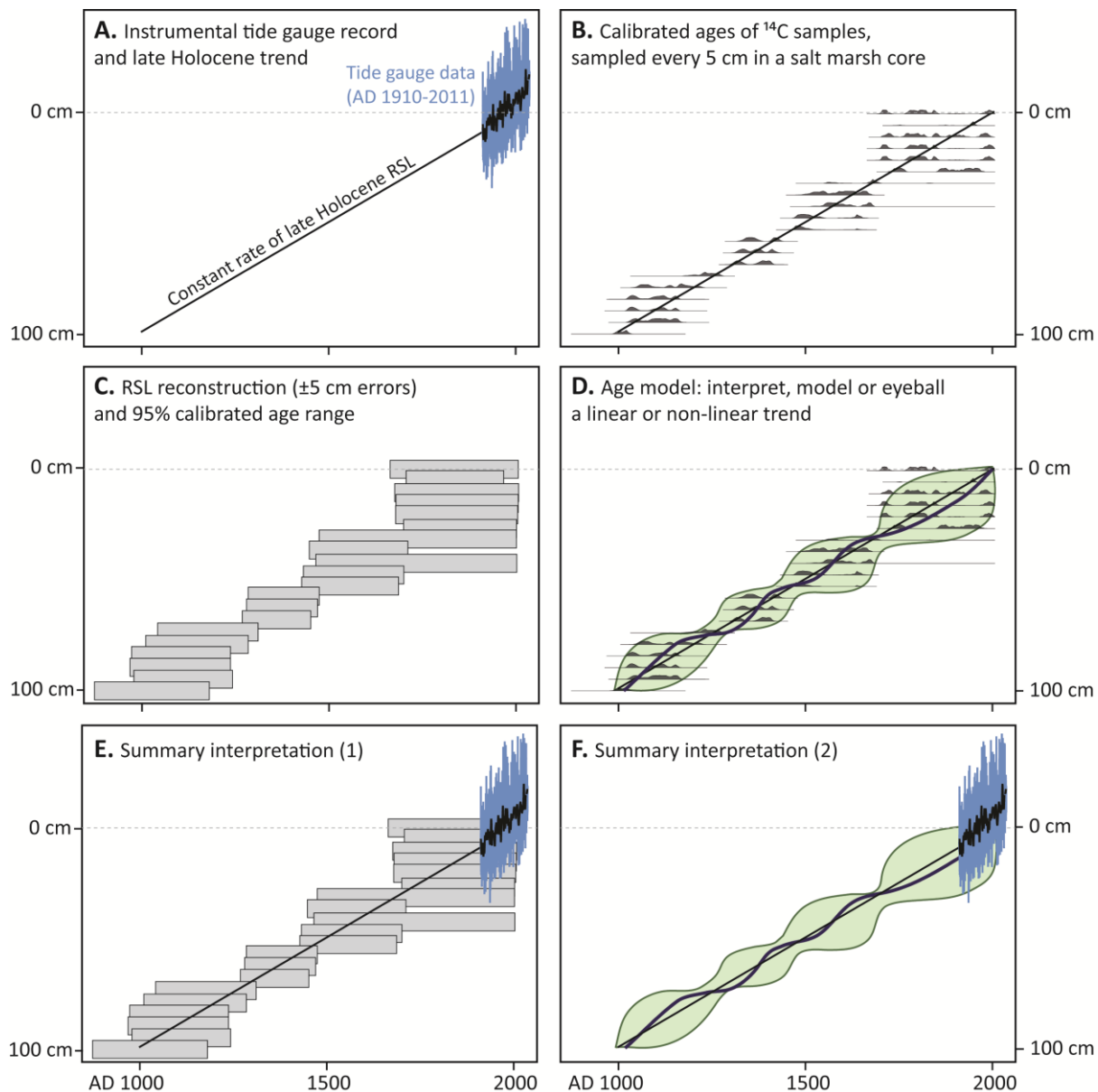
Figure 10 – Reconstruction and interpretation of Holocene sea level data and trends. A: Known constant trend of late Holocene sea level established from sea level index datasets such as in Figure 1B compared with instrumental tide gauge records of 20th century RSL (Figure 1E); B: calibrated ^{14}C dates (from Figure 9) for a hypothetical salt marsh core; C: 95% calibrated age ranges from 'B' with hypothetical RSL reconstruction with ±5 cm errors shown as box for each dated and reconstructed level in a core; D: interpretation of dates in 'B' can take the form of a modelled or eyeballed linear or non-linear trend. The results of the steps in B, C and D can be interpreted as a linear RSL trend (E), similar to that in 'A'; or a series of RSL accelerations or decelerations (F). Independent tests of the reconstruction precision and accuracy and intra- and inter-site replication of trends (see text for discussion) aids in determining which interpretation is mostly likely correct.
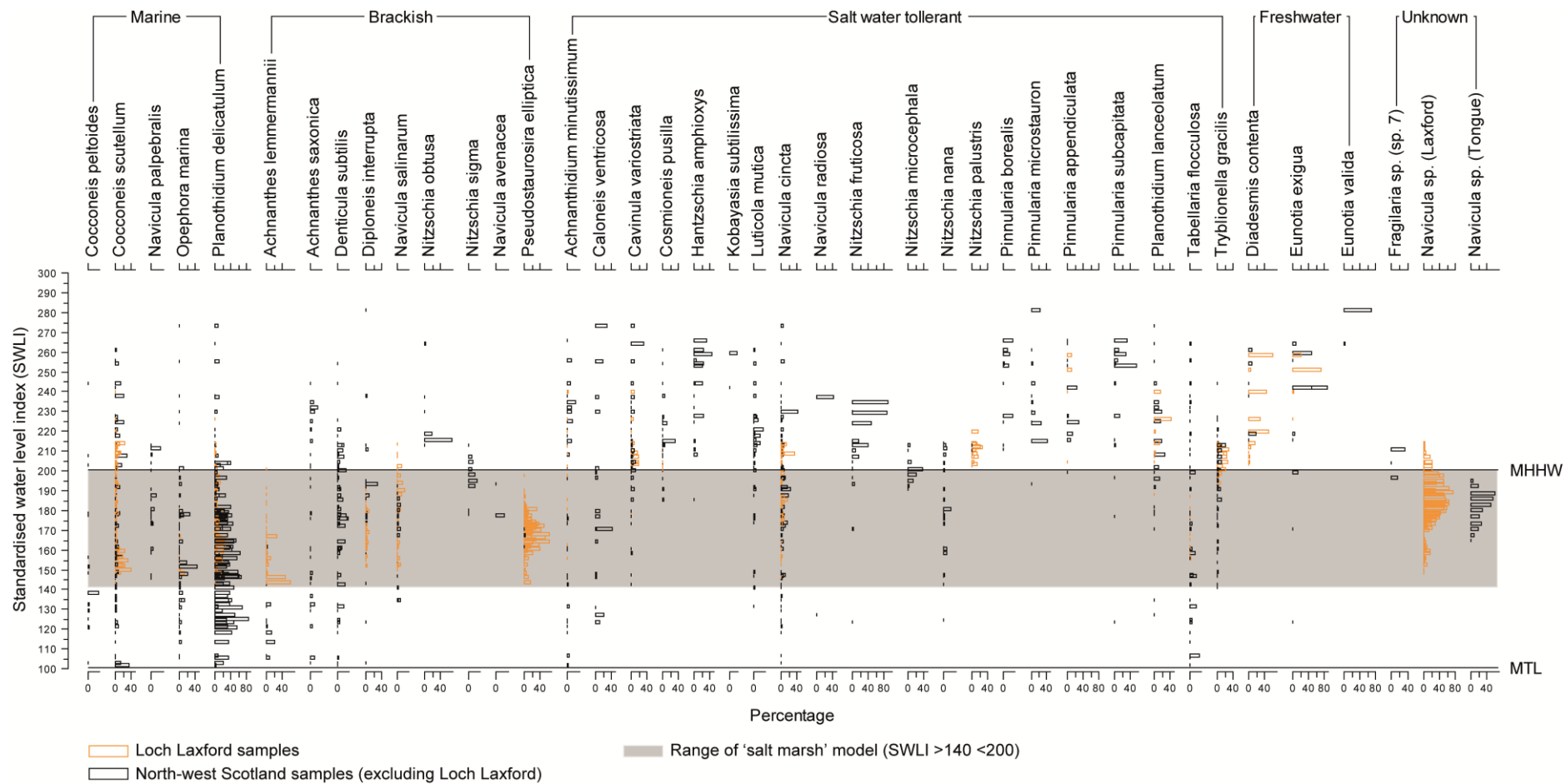
Supplementary Figure 1 – Scottish modern diatom training set: species >20% total valves counted

Site	Site latitude, longitude	Nearest port (tidal observations)	Port latitude, longitude	MHWST tidal station (m OD)	MTL tidal station (m OD)
Balmacara, Loch Alsh	57.28, -5.62	Kyle of Lochalsh, Loch Alsh	57.17, -5.43	2.57	0.32
Eilean nan Gall, Loch Duich	57.21, -5.42	Dornie Bridge, Loch Long	57.17, -5.31	2.62	0.30
Nonach, Loch Long	57.31, -5.44				
Kentra Moss, Kentra Bay	56.75, -5.85	Mallaig, Lochaber	57.00, -5.50	2.38	0.28
Saideal Ceapaich, Loch nan Ceall	56.91, -5.85				
Loch Creran, Loch Linnhe	56.55, -5.24	Port Appin, Loch Linnhe	56.33, -5.25	2.25	0.55
Kinloch, Kyle of Tongue	58.44, -4.47	Portnancon, Loch Eriboll	58.30, -4.42	2.42	-0.30
Tràigh Bad na Bàighe, Loch Laxford	58.37, -5.04	Loch Laxford	58.24, -5.05	2.40	0.25
Tramaig Bay, Jura	56.03, -5.76	Based on local observations	-	0.51	-0.27

Table 1 - Sites in the Scotland regional diatom dataset (Figure 4) with the nearest tide gauge observations of mean high water spring tides (MHWST) and mean tide level (MTL). The nearest tidal observations for Tramaig Bay on the isle of Jura are too far away and therefore we estimate MHWST and MTL from local observations.

Model	# of samples	SWLI range	DCCA Axis 1 length (SD units)	Bootstrapped r^2 Number of components			RMSEP (SWLI) Number of components			MAT DC 5th percentile	MAT DC 20th percentile
				1	2	3	1	2	3		
Regional 'coastal transition' model	215	100-300	4.08	0.73	0.78	0.81	19.5	18.5	19.5	68.6	118.9
Regional 'salt marsh' model	121	140-200	2.06	0.57	0.67	0.70	10.9	9.9	9.7	42.7	93.7
Loch Laxford 'coastal transition' model	73	100-300	5.08	0.87	0.92	0.93	9.3	7.3	7.5	28.7	56.1
Loch Laxford 'salt marsh' model	53	140-200	2.47	0.86	0.91	0.91	5.9	4.8	4.9	17.6	41.2

Table 2 - Details of the four diatom training set models used to reconstruct RSL based on the fossil diatom assemblages in the Loch Laxford core (Figure 5). The WA-PLS model performance statistics and modern analogue technique dissimilarity coefficients (MAT DC) are reported from the C2 programme (Juggins, 2003). The reconstruction results of each model and component are shown in Figure 7.

Site	# of samples	Model type	Reported RMSEP (m)	Elevation range of samples (m)	Mean tidal range (m)	RMSEP/ tidal range (%)	RMSEP/ elevation range of samples (%)	Reference
Foraminifera training set								
Tasmania, Australia	43	WA-PLS	0.10	0.80 ǂ	0.60	16.7	12.5	Callard et al. (2011)
Catlins Coast, New Zealand	31	WA-Tol	0.05	0.31	1.50	3.3	15.9	Southall et al. (2006)
Oregon, USA (average of 5 locations). *Values given relative to South Beach.*	91	WA-PLS	0.20	2.28 ǂ	2.54	7.9	8.8	Hawkes et al. (2010)
Norfolk, England (average of 2 locations). *Values given relative to Thornham marsh.*	47	WA-PLS	0.25	4.11 ǂ	4.80	5.2	6.1	Horton and Edwards (2005)
South Devon, England (average of 2 locations)	85	WA-PLS	0.29	4.20	4.65	6.2	6.9	Massey et al. (2006)
Brittany, France (average of 2 locations)	36	PLS	0.07	0.74 ǂ	3.00	2.3	9.5	Rossi et al. (2011)
Brittany, France	43	PLS	0.13	0.32	2.59	4.9	39.3	Leorri et al. (2010)
Northern Spain (average of 4 locations). *Values given relative to Minho estuary.*	30	WA-PLS	0.19	1.25 ǂ	2.50	7.6	15.2	Leorri et al. (2008)
Northern Portugal	49	WA-PLS	0.13	3.63 ǂ	1.81	7.2	3.6	Leorri et al. (2011)
Southern Portugal	22	PLS	0.10	0.46	2.00	4.8	20.7	Leorri et al. (2010)
Viðarholmi, Iceland	21	WA-Tol	0.20	1.08 ǂ	2.10	9.5	18.5	Gehrels et al. (2006)
Nova Scotia, Canada	46	WA-Tol	0.06	0.28	1.86	3.2	21.4	Gehrels et al. (2005)
Newfoundland, Canada	37	WA-PLS	0.07	0.80 ǂ*	0.83	8.4	8.8	Wright et al. (2011)
Maine, USA (average of 4 locations)	68	WA-PLS	0.25	2.05 ǂ	3.11	8.0	12.2	Gehrels (2000)
New Jersey, USA (average of 3 locations). *Values given relative to Leeds Point.*	56	WA-PLS	0.14	1.30 ǂ	1.10	12.7	10.8	Kemp et al. (2012)
Connecticut, USA (average of 4 locations). *Values given relative*	91	WA-PLS	0.09	0.99 ǂ	1.28	7.0	9.1	Edwards et al. (2004a)

to Menunketesuck.								
Pattagansett River, Connecticut, USA	26	WA-PLS	0.10	0.80 ɫ*	1.16	8.6	12.5	Wright et al. (2011)
Elizabeth River, North Carolina, USA	53	WA-PLS	0.14	1.65 ɫ*	1.34	10.4	8.5	Wright et al. (2011)
Outer Banks, North Carolina, USA (average of 10 locations). *Values given relative to Sand Point.*	49	WA-PLS	0.04	0.24 ɫ	0.17	23.5	16.7	Kemp et al. (2009b)
Diatom training set								
Regional 'coastal transition' model, Scotland (average of 9 locations). *Values given relative to Loch Laxford.*	215	WA-PLS	0.40	3.86 ɫ	4.30	9.3	10.3	This paper
Regional 'salt marsh' model, Scotland (average of 9 locations). *Values given relative to Loch Laxford.*	121	WA-PLS	0.21	1.25 ɫ	4.30	5.0	17.0	This paper
Loch Laxford 'coastal transition' model, Scotland	73	WA-PLS	0.16	2.47	4.30	3.7	6.4	This paper
Loch Laxford 'salt marsh' model, Scotland	53	WA-PLS	0.10	1.18	4.30	2.4	8.7	This paper
Upper Cook Inlet, Alaska, USA (average of 4 locations). *Values given relative to Bird Point.*	149	WA-PLS	0.30	3.64 ɫ	7.98	3.8	8.2	Barlow et al. (2012)
Aasiaat, Greenland	64	WA-PLS	0.16	1.34	2.70	5.9	11.9	Woodroffe and Long (2010)
Sisimiut, Greenland	70	WA-PLS	0.19	2.56	4.50	4.2	7.4	Woodroffe and Long (2010)
Ho Bugt, Denmark	40	WA-PLS	0.14	1.93 ɫ	1.50	9.3	7.3	Szkornik et al. (2006)
Outer Banks, North Carolina, USA (average of 3 locations). *Values given relative to Oregon Inlet.*	46	WA-PLS	0.08	0.88 ɫ*	0.35	22.9	9.1	Horton et al. (2006)
Hokkaido, Japan (average of 2 locations)	78	WA-PLS	0.29	3.85 ɫ*	1.05	27.6	7.5	Sawai et al. (2004)

| Bristol Channel, England | 61 | WA-Tol | 0.88 | 7.71† | 11.66 | 7.5 | 11.4 | Hill et al. (2007) |

†Maximum vertical elevation range from datasets over more than one transect.

*Elevation range estimated from Figures in referenced paper as the absolute values are not given.

Table 3 - Modern foraminifera and diatom datasets from salt marshes around the world, showing the reported RMSEP, the maximum vertical elevation range of the modern dataset and the mean tidal range. WA-Tol — Tolerance Down-Weighted Weighted Average regression; WA-PLS — Weighted-Average Partial-Least-Squares regression; PLS — Partial-Least-Squares regression. Figure 8 shows the relationship between RSMEP, mean tidal range and elevation range of the modern dataset. Table updated from that published by Callard et al. (2011).