Quantitative vertical zonation of salt-marsh foraminifera for reconstructing former sea level; an example from New Jersey, USA.

Andrew C. Kemp\textsuperscript{ab*}, Benjamin P. Horton\textsuperscript{c}, David R. Vann\textsuperscript{c}, Simon E. Engelhart\textsuperscript{c}, Candace Grand-Pre\textsuperscript{c}, Christopher H. Vane\textsuperscript{d}, Daria Nikitina\textsuperscript{e}, and Shimon Anisfeld\textsuperscript{a}

\textsuperscript{a}School of Forestry and Environmental Studies, Yale University, New Haven, CT 06510, USA.
\textsuperscript{b}Yale Climate and Energy Institute, New Haven, CT 06510, USA.
\textsuperscript{c}Department of Earth and Environmental Science, University of Pennsylvania, Philadelphia, PA 19104, USA.
\textsuperscript{d}British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham NG12 5GG, UK.
\textsuperscript{e}Department of Geology and Astronomy, West Chester University, West Chester, PA 19383, USA.

\* Corresponding author, email: andrew.kemp@yale.edu, tel: 203 436 3978
Abstract

We present an objective and quantitative technique to reconstruct former sea level from assemblages of salt-marsh foraminifera using partitioning around medoids (PAM) cluster analysis and linear discriminant functions. Three salt marshes, representing different physiographic environments in southern New Jersey, were selected for describing the modern distribution of foraminifera from 56 surface samples. PAM estimated the number and composition of assemblages present at each site and showed that foraminifera adhered to the concept of elevation-dependent ecological zones, making them appropriate sea-level indicators. This approach has several advantages in comparison to hierarchical clustering. Application of PAM to a regional dataset generated by combining all samples identified the presence of five distinctive biozones occupying defined elevation ranges. The characteristics of these biozones were similar to those identified elsewhere on the mid-Atlantic coast of the USA. Classification of each of the 56 samples as belonging to one of the five biozones enabled us to develop linear discriminant functions, which confirmed their distinctiveness. These functions can estimate the probability that assemblages of foraminifera preserved in samples of salt-marsh sediment represent one of the five modern biozones. Recognition of these biozones in sequences of salt-marsh sediment provides a means to reconstruct sea level. We collected a 4.0 m core from Leeds Point, New Jersey to investigate the practical application of this approach to reconstructing former sea level. The linear discriminant functions were used to indicate the faunal origin of 32 core samples and in cross validation tests were accurate in 54 of 56 cases. The approach described can be used as an independent means to reconstruct sea level or to check the ecological plausibility of results from other techniques including transfer functions.
Keywords

salt marsh, foraminifera, New Jersey, discriminant function, cluster analysis, sea level
1 Introduction

Low energy, coastal sedimentary archives in temperate regions have provided detailed records of Holocene sea-level changes (Stuiver and Daddario, 1963; Kraft, 1971; Peteet Carmichael, 1980; Gehrels et al., 1996; Nikitina et al., 2000). Such reconstructions are contingent upon the appropriate selection and application of sea-level indicators to accurately estimate former sea level. A sea-level indicator is a physical, biological or chemical feature possessing a systematic and quantifiable relationship to elevation in the tidal frame (Shennan, 1986; van de Plassche, 1986). This relationship, known as the indicative meaning, incorporates the elevational range occupied by a sea-level indicator (indicative range) in relation to a contemporaneous tide level (reference water level). This approach is dependent upon a detailed understanding of the modern characteristics of the chosen sea-level indicator. Further, a quantitative technique is necessary to provide objective estimates of Holocene relative sea level (RSL) on the basis of the similarity between sea-level indicators preserved in sub-fossil sedimentary material and those documented from modern settings (Jackson and Williams, 2004).

Assemblages of foraminifera can be used as sea-level indicators because their distribution on modern salt marshes reflects changes in the frequency and duration of tidal inundation and permits recognition of elevation-dependent ecological zones (Scott and Medioli, 1978; Gehrels, 1994; Horton and Edwards, 2006). Some studies have documented a potential precision of $<\pm0.1$ m in high-marsh settings (Scott and Medioli, 1978; Gehrels et al., 2001; Leorri et al., 2008; Kemp et al., 2009a). Foraminifera are commonly well preserved in salt-marsh sediment and are well suited to quantitative analysis because they form low diversity, high abundance assemblages (Gehrels, 2007).
A number of techniques have been used to reconstruct RSL using foraminifera. The vertical zonation concept proposed by Scott and Medioli (1978) uses the elevational range of discrete groups of modern salt-marsh foraminifera as the basis for assigning an indicative meaning to assemblages enumerated from core material (Gehrels, 1994). This approach has frequently been used in a qualitative fashion where modern zones are determined subjectively and assignment of core samples to one of these groups is reliant upon the judgment of the researcher. Alternatively, discriminant functions have been used to assign samples to zones that were also defined in a qualitative fashion (Jennings and Nelson, 1992). More recently, these approaches have been widely superseded by the use of transfer functions (Guilbault et al., 1995; Horton et al., 2000; Edwards and Horton, 2006; Massey et al., 2006; Gehrels et al., 2008; Leorri et al., 2008; Woodroffe, 2009). Transfer functions are empirically derived equations for producing quantitative estimates of past environmental conditions from paleontological data (Sachs et al., 1977). They have been used to produce accurate and precise estimates of former RSL in an objective and reproducible fashion (Gehrels et al., 2005; Horton and Edwards, 2006). The validity of these reconstructions has been confirmed by comparison with tide-gauge records (Gehrels et al., 2005; Kemp et al., 2009a). However, each of the numerical techniques used in developing transfer functions have underlying assumptions about the nature of species responses to environmental changes (Birks, 1995). The ecological plausibility of all reconstructions should be reviewed out of concern for inaccurate estimates despite seemingly high precision (Birks, 1995; Woodroffe, 2009). Additionally, recent literature has highlighted the potential for transfer function-derived reconstructions of RSL to be influenced by spatial autocorrelation, resulting in overly optimistic estimates of uncertainty (Telford and Birks, 2005, 2009; Zong et al., 2010).
In this paper, we present an alternative means to reconstruct former RSL based upon modern and fossil assemblages of salt-marsh foraminifera using partitioning around medoids (PAM) cluster analysis in combination with linear discriminant functions. This approach is not underpinned by assumptions about the distribution and response of foraminifera and provides a probabilistic estimate of the strength of analogy between modern and fossil assemblages of foraminifera. We develop a new modern training set of foraminifera from three salt marshes in southern New Jersey, USA (Figure 1). To illustrate the practical application of this approach, estimates of former salt-marsh elevation are provided for samples in a core of salt-marsh sediment collected from one of the study sites using linear discriminat functions and compared to transfer function results.

2 Modern Setting

The central and southern Atlantic coast of New Jersey is characterized by a chain of barrier islands, which protect a lagoon system from the open ocean. Inlets separate the islands and allow exchange of water between the Atlantic Ocean and the lagoons. The coast between Great Bay to the north and Cape May to the south (Figure 1) includes nine open inlets and is characterized by islands which typically decrease in size from north to south (Ferland, 1990). Large areas of formerly open-water lagoon have been infilled by washover material and accretion of salt-marsh sediment (Daddario, 1961; Meyerson, 1972; Thorbjarnarson et al., 1985; Psuty, 1986; Ferland, 1990).
Modern salt marshes in this region form extensive platforms dissected by tidal channels of varying size (Ferland, 1990), tidal flat environments are rare as the coast is experiencing ongoing erosion (Dolan et al., 1979; Fitzgerald et al., 2008). Low-marsh settings are typically vegetated by *Spartina alterniflora*, while a high-marsh floral zone is recognized by *Spartina patens* with *Spartina alterniflora* (stunted form) and *Distichlis spicata* (Daddario, 1961). The border between salt marsh and freshwater upland is vegetated by *Phragmites australis*, *Typha* spp. and *Scirpus* spp., it is typically narrow and representative of brackish conditions (Daddario, 1961; Stuckey and Gould, 2000).

The region has a semidiurnal, microtidal (range <2m) regime. Tidal ranges (MLLW to MHHW) are slightly larger on the ocean side of the barrier islands (1.4m at Atlantic City; Figure 1) than in the lagoons. At the study sites around Great Bay (Figure 1), tidal ranges were estimated to be 1.1m at Leeds point and Bass River and 1.3m at Brigantine Barrier by VDatum (Hess et al., 2003; Parker et al., 2003).

### 3 Methods

#### 3.1 Sampling Regime

At the three sites (Leeds Point, Bass River and Brigantine Barrier) we established transects across the modern salt marsh which were positioned to include the full range of physiographic environments at each site (Figure 1). Two transects were sampled at Leeds Point (A-A’ and B-B’) and one at both Bass River and Brigantine Barrier. Sampling stations reflected changes in elevation and vascular vegetation and were used to document the modern distribution of
foraminifera. Sample altitudes were established using Real Time Kinematic (RTK) satellite navigation, where the base station (Leica GPS1200+) made a minimum of 2000 observations.

A core (EF10) was selected for analysis from the Leeds Point site following stratigraphic investigation of the site. The core was recovered in 0.5m sections using a Russian-type hand core. It was sampled at a resolution of 10cm in the laboratory to ensure that all stratigraphic units were adequately represented. Each sample consisted of a 1cm thick section of sediment.

3.2 Sample preparation – foraminifera

At each sampling station we collected a 10cm² surface (0-1cm) sediment sample. All samples were stained in the field using rose Bengal to allow recognition of individuals living at the time of collection (Walton, 1952) and stored in buffered ethanol. In the laboratory, samples were sieved to separate and retain the 63-500μm size fraction of the sediment. Samples were subsequently divided into eight aliquots using a wet splitter (Scott and Hermelin, 1993) and foraminifera were counted wet under a binocular microscope. In instances where the sample included a large volume of sand and little organic material, foraminifera were isolated from the sediment using sodium polytungstate (density=2.89) as a medium for density separation (Munsterman and Kerstholt, 1996). A minimum of 200 dead individuals were enumerated from a known fraction of the original sediment; where necessary the entire sample was counted. Identifications of foraminifera were confirmed by comparison with type and figured specimens lodged at the Smithsonian Institution, Washington, D.C. and The Natural History Museum, London. Plate 1 shows scanning electron microscope images of select foraminifera sampled in this study. Species of Ammonia and Elphidium were combined into generic groups (Hayward et
al., 2004; Horton and Edwards, 2006). Due to difficulties in identifying broken individuals, we have also combined species of *Ammobaculites* into a single generic group. All foraminiferal data presented are from dead assemblages. Preparation of foraminiferal samples from core material followed closely that used for modern samples. Complete foraminiferal data (modern and core) are provided in Appendix A.

### 3.3 Statistical methods

At each of the three sites we employed partitioning around medoids (PAM) cluster analysis (Kaufman and Rousseeuw, 1990) to define the number and composition of modern foraminiferal assemblages. This analysis was also performed on a complete, regional dataset of modern foraminifera generated by combining data from all sites. Analysis was based upon complete (all species) counts in percentage form and executed using the ‘cluster’ package in *R* (Maechler et al., 2005). This technique has several advantages over other methods of cluster analysis. It is particularly robust because it minimizes a sum of dissimilarity as opposed to a sum of squared Euclidean distances (Kaufman and Rousseeuw, 1990). Further, it generates a novel graphical representation of the data called a *silhouette plot*. A silhouette width of between -1 and 1 is possible for each sample and provides an estimate of the sample’s classification. Values close to -1 indicate that a sample was incorrectly classified. In contrast, values close to 1 indicate that the object was assigned to an appropriate cluster as the within group dissimilarity was much smaller than the between groups dissimilarity. Values close to 0 are an intermediate case and also represent groups consisting of a single sample. Silhouette widths enable the user to objectively select the optimal number of classifications within the dataset (Kaufman and Rousseeuw, 1990). We used maximum average silhouette width as the criteria for determining the number of
groups. An additional feature of PAM is its description of the medoid for each group (Kaufman and Rousseeuw, 1990). The medoid is the object for which average dissimilarity to all samples in a given group is minimal; in this study it is expressed as an assemblage of foraminifera. Statistical measures of each group’s characteristics (maximum and average within group dissimilarity, diameter and separation) are also provided.

Linear discriminant functions were used to estimate paleo-marsh elevation (PME), which is the elevation at which a sample formed in relation to contemporary sea level and is considered as an estimate of reference water level, which is necessary to reconstruct former RSL (Edwards, 2007). The analysis was completed using the ‘MASS’ package executed in R (Venables and Ripley, 2002). All modern samples of foraminifera were first classified as being from one of the n biozones recognized by PAM cluster analysis. These samples were combined to generate a modern training set. A 2 arcsine transformation of square rooted proportional foraminiferal data was necessary to satisfy assumptions about the distribution of data (Owen, 1962). Cross validation (leave one out) of the training set was used to estimate the frequency of sample misclassification (i.e. how often the linear discriminant functions assigned a sample to the wrong biozone). This provides a measure of how well the linear discriminant functions are able to correctly assign samples and is thus a measure of performance (Venables and Ripley, 2002).

A test set is composed of samples which are to be classified using the linear discriminant functions. We used assemblages of foraminifera (all species) from samples in core EF10 as a test set following a 2 arcsine transformation of square rooted proportional data (Owen, 1962). Application of the linear discriminant functions estimated the probability that samples from the
core should be allocated to each of the modern biozones. The modern elevational range of these biozones was used to estimate PME for samples in the test set (average elevation ± 1σ and midpoint ± half of the range). Descriptions of the commands used for PAM clustering and development and application of linear discriminant functions are provided in Appendix B.

4 Results

4.1 Distribution of modern foraminifera

We enumerated dead foraminifera from 56 modern samples of modern salt-marsh sediment collected at three sites. A total of 16 species were recognized, of which 11 exceeded 10% of the assemblage in at least one sample.

At Leeds Point, we collected 32 samples from two transects (Figure 1A) and identified 11 species of foraminifera. The first transect (A-A’) was 107m long and included 14 sampling stations (Figure 2A). From 0-5m, two samples collected in an upland environment did not contain any foraminifera and were assumed to represent freshwater conditions. At stations 3 to 6, situated in a stand of Phragmites australis and Typha sp. (10-30m along transect), the dominant species of foraminifera were Haplophragmoides manilaensis (14-56%; average 31%), Tiphotrocha comprimata (2-65%; average 21%) and Arenoparrella mexicana (up to 36%). Spartina patens was the dominant form of vascular vegetation between 30 and 104m along the transect (stations 7-12). The characteristic foraminifera in this part of the marsh were Arenoparrella mexicana (40-63%; average 49%) and Trochammina inflata (up to 21%). Two samples (stations 13 and 14) positioned in a muddy Spartina alterniflora low marsh, revealed high abundances of Miliammina fusca (up to 71%).
The second transect at Leeds Point (Figure 1A; B-B’) was parallel to the first and consisted of 18 sampling stations (Figure 2). From 0-8m along the transect, four samples (stations 1 to 4) collected in a forest fringing the salt marsh contained no foraminifera and are considered to represent a freshwater-upland environment. In a *Phragmites australis* stand (13-17m), two samples (stations 5 and 6) were dominated by *Haplophragmoides manilaensis* (62-88%). From 20 to 41m along the transect (stations 7-10) the characteristic species of foraminifera were *Haplophragmoides manilaensis* (up to 28%; average 17%) and *Tiphotrocha comprimata* (10-67%; average 27%). These samples spanned the transition between *Phragmites australis* and *Spartina patens*. A mono-specific stand of *Spartina patens* between 65 and 75m (stations 11-15) along the transect was populated by *Arenoparrella mexicana* (42-66%; average 51%) in association with *Trochammina inflata* (16-27%; average 23%) and *Tiphotrocha comprimata* (up to 21%; average 12%). Three samples (stations 16-18) in a muddy, *Spartina alterniflora* environment (76-81m), were dominated by *Miliammina fusca* (37-45%) with occurrences of *Arenoparrella mexicana* (7-31%; average 19%).

Foraminifera from the two transects were combined into a single dataset for cluster analysis. PAM showed that it was appropriate to recognize four groups of foraminifera (Figure 2B) at Leeds Point. All samples were well classified as evidenced by silhouette widths between 0 and 1 for group averages (Figure 2C, Table 1). Group LP-1 had an average silhouette width of 0.33 (Figure 2C) and included samples from the *Phragmites australis* high marsh which were dominated by *Haplophragmoides manilaensis* (28% to 88%; average 50%). The elevational range of this group was 0.65-0.91m above mean tide level (MTL). Group LP-2 consisted of
three samples with an average silhouette width of 0.32, which were dominated by *Tiphotrocha comprimata* (23% to 68%; average 51%) and were positioned in the *Spartina patens* middle marsh. This group had an elevational range of 0.54-0.59m MTL. Group LP-3 had the highest average silhouette width (0.48), making it the best defined group at the Leeds Point site (Figure 2C). It was also the most homogenous, having the lowest average (21.63) and maximum (31.82) dissimilarity between member samples (Table 1). The group was distinguished by high abundances of *Arenoparrella mexicana* (25% to 66%; average 48%), with *Trochammina inflata* (up to 30%; average 20%) and *Miliammina fusca* (up to 39%). The elevational range of this group was 0.28-0.55m MTL. Group LP-4 included four samples with an average silhouette width of 0.35 (Figure 2C). It was characterized by high abundances of *Miliammina fusca* (37% to 71%; average 49%), with *Arenoparrella mexicana* (up to 31%) and was associated with low-marsh *Spartina alterniflora*. These samples encompassed elevations from -0.13-0.27m MTL. None of the four groups recognized at Leeds Point was isolated as in all cases the diameter of the group exceeded its separation from others implying overlap (Table 1).

At Bass River, a transect of 15 samples identified 14 species of foraminifera (Figure 3). The inland part of the transect (0 to 14m; stations 1 to 4) was characterized by high abundances of *Tiphotrocha comprimata* (12-61%; average 37%) in association with *Trochammina inflata* (up to 36%), *Arenoparrella mexicana* (average 20%) and *Jadammina macrescens* (up to 15%). This part of the marsh was vegetated by *Spartina patens* with the presence of *Phragmites australis* (station 1) and *Iva fructescens*. From 19 to 44m along the transect (stations 5 to 15), in an area dominated by *Spartina alterniflora* and *Salicornia* sp., the characteristic species of foraminifera were *Miliammina fusca* (8-69%; average 44%) with *Tiphotrocha comprimata* (up to 38%;
average 15%) and *Ammobaculites* spp. (up to 50%). A small number (<4%) of *Balticammina pseudomacrescens* were identified at the Bass River site.

PAM cluster analysis was used to identify two groups of foraminifera at the Bass River site (Figure 3). All samples were well classified as evidenced by silhouette widths between 0 and 1 for group averages. Group BR-1 had an average silhouette width of 0.33 (Figure 3C) and included samples characterized by *Tiphotrocha comprimata* (12% to 61%; average 37%), *Arenoparrella mexicana* (up to 46%; average 18%) and *Jadammina macrescens* (up to 15%; average 7%). The elevational range of this group was 0.18-0.59m MTL. The second group (BR-2) had an average silhouette width of 0.45 and was dominated by *Miliammina fusca* (30% to 69%; average 53%) and *Ammobaculites* spp. (up to 50%; average 21%). Of the two groups, BR-2 was the most homogenous (maximum dissimilarity = 44.3, average dissimilarity = 23.41; Table 1). The elevational range of this group was -0.40-0.31m MTL. The two groups at Bass River were not isolated from one another (Table 1).

A transect of 15 samples at the Brigantine Barrier salt marsh recorded 12 species of foraminifera (Figure 4). At sampling station one, a foraminiferal assemblage consisting of 16% *Haplophragmoides manilaensis* with 48% *Trochammina inflata* and 21% *Jadammina macrescens* was documented. The station was located in a mixed stand of *Phragmites australis* and *Iva fructescens*. From 6 to 100m along the transect (stations 2 to 10) the dominant species of foraminifera were *Trochammina inflata* (28-47%; average 41%) and *Jadammina macrescens* (25-47%; average 34%). Small numbers of *Balticammina pseudomacrescens* (<2%) were identified. This part of the salt marsh was vegetated by a mix of *Spartina patens*, *Spartina*
alterniflora (stunted form) and Salicornia sp. The end of the transect (106-117m, stations 10 to 15) was characterized by high abundances of Miliammina fusca (27-71%; average 51%) and Ammobaculites spp. (11-40%; average 19%). The calcareous species Haynesina germanica was identified on unvegetated muddy sediments at stations 14 and 15.

PAM cluster analysis recognized two groups of foraminifera at Brigantine Barrier (Figure 4). Group BB-1 had an average silhouette width of 0.42 (Figure 4C) and was dominated by Trochammina inflata (28% to 48%; average 41.4%), Jadammina macrescens (21% to 43%; average 32.5%) and Trochammina comprimata (up to 23%; average 14.1%). Of the eight groups recognized among individual sites, BB-1 had the lowest dissimilarity between group members (maximum and average) and is therefore considered to be the most homogenous (Table 1). The elevational range of samples in this group was 0.35-0.79 m MTL. Group BB-2 had an average silhouette width of 0.74 and was dominated by Miliammina fusca (27% to 71%; average 50.9%) and Ammobaculites spp. (11% to 40%; average 18.7%). The elevational range of this group was -0.48-0.25 m MTL.

4.2 Recognition of foraminiferal biozones in southern New Jersey

We generated a single dataset by combining the foraminifera counts from the four transects to identify modern biozones in the study region. PAM cluster analysis of this new dataset showed that it was appropriate to recognize five biozones as evidence by a peak average silhouette width of 0.45 (Figure 5A). The medoids of these five biozones provide a convenient and meaningful way to describe the groups (Table 2). Biozone A was described by a medoid of 40.5% Trochammina inflata with 34.9% Jadammina macrescens (Figure 5a; Table 2). This group of
samples is predominantly from the high-marsh environment at Brigantine Barrier (Figure 4) and had the lowest average dissimilarity among member samples (16.99; Table 1). Biozone B is represented by a medoid assemblage of 63.1% *Miliammina fusca* and 19.1% *Ammobaculites* spp. and consisted predominantly of samples from *Spartina alterniflora* low-marsh settings at all three sites. Biozone C was characterized as 45.3% *Arenoparrella mexicana* with 21.1% *Trochammina inflata* and described a middle-marsh assemblage vegetated by *Spartina patens* at Leeds Point. It was the biozone with least separation from others (17.06; Table 1). Biozone D had a medoid assemblage of 44.4% *Tiphotrocha comprimata* with 18.4% *Arenoparrella mexicana* and is largely composed of samples from Bass River (Figure 3). Biozone E was represented by a medoid assemblage of 61.7% *Haplophragmoides manilaensis* and represented the high-marsh *Phragmites australis* and *Typha* sp. environment at Leeds Point (Figure 2). It had the greatest average dissimilarity among samples (28.75; Table 1). The elevational range of each of these biozones was conservatively estimated by using the maximum and minimum values from samples within each group (Table 3) and also using an average elevation ± 1σ. Due to the differences in tidal ranges among the three sites, it was necessary to express elevations in the combined data set as a standardized water level index (SWLI) following the approach described by Horton and Edwards (2006).

4.4 Development of linear discriminant functions

Linear discriminant functions are able to rigorously allocate new observations to one of *n* prespecified classes (Venables and Ripley, 2002). We used PAM cluster analysis to recognize five biozones which constituted the prespecified classes (Figure 5A). Separation of the biozones was confirmed by the positioning of modern samples along axes representing the first two linear
discriminants (Figure 5C). These two axes explained 68% of the among class variance (86% for axes one to three) and are conventionally centered on zero (Venables and Ripley, 2002). In 54 of 56 tests of cross validation, samples were correctly allocated by the discriminant functions with an error rate of <0.05, that is, the sample was correctly assigned at least 95 out of 100 times.

4.5 Foraminifera in core EF10

Core EF10 was retrieved from the Leeds Point salt marsh to provide a test set (Figure 1A). The core consisted of 4.2m of organic-rich sediment overlying a basal sand unit. Foraminifera were absent in samples from 4.2m to 3.3m. The lowest occurrence of foraminifera in the core was at 3.3m (Figure 6), although this sample yielded a small count (<30 individuals). The interval between 3.3m and 2.9m was characterized by high abundances of *Jadammina macrescens* (16% to 93%; average 48%). From 2.9m to 1.9m, the dominant species of foraminifera was *Trochammina inflata* (up to 70%; average 40.5%) in association with *Jadammina macrescens* (up to 86%; average 27.9%) and *Tiphotrocha comprimata* (up to 38%; average 17.4%).

Foraminifera were absent from a sample analyzed at 1.8m (Figure 6). Between 1.7m and 1.1m, four of the seven samples had low counts of foraminifera; the dominant species was *Jadammina macrescens* (up to 100%; average 93.3%). The uppermost section of the core (1.1m to 0.1m) was characterized by *Trochammina inflata* (up to 77%; average 52.4%) and *Jadammina macrescens* (up to 96%; average 27.6%); one sample in this interval (0.3m) yielded a low count (100% *Trochammina inflata*; 9 individuals).

5 Discussion

5.1 Distribution of modern salt-marsh foraminifera
Despite forming low diversity assemblages that are often recognized throughout temperate regions, it remains necessary to have modern training sets of salt-marsh foraminifera from the area close to where core material will be recovered because of the influence of site-specific assemblages (de Rijk, 1995; Edwards et al., 2004; Gehrels, 2007). This study provides a new dataset describing the modern distribution of salt-marsh foraminifera in southern New Jersey for use in reconstructing sea level. The distributions are similar to those from other sites on the mid-Atlantic coast of the USA (here we include North Carolina in this region).

High-marsh assemblages at the salt-marsh to freshwater upland transition in southern New Jersey were represented by biozones A and E and associated with high abundances of *Trochammina inflata* with *Jadammina macrescens* and *Haplophragmoides manilaensis* respectively (Figure 5B). High-marsh assemblages dominated by *Jadammina macrescens* were present in North Carolina at sites with lower salinities (Robinson and McBride, 2006; Horton and Culver, 2008; Kemp et al., 2009b). A peak in the abundance of *Jadammina macrescens* was recorded at the highest elevations on a Virginia salt marsh by Spencer (2000) and was also associated with high abundances of *Trochammina inflata*. A study of depositional environments in Virginia recognized both *Jadammina macrescens* and *Trochammina inflata* as important constituents of salt-marsh assemblages (Culver et al., 1996). In Delaware, *Jadammina macrescens* and *Trochammina inflata* were abundant in high-marsh floral zones (Hippensteel et al., 2000). We recognized the presence of *Balticammina pseudomacrescens* at the Bass River and Brigantine Barrier sites with abundances of up to 4%. It was not recorded by investigations of mid-Atlantic modern salt-marsh foraminifera in either North Carolina or Virginia (Spencer, 2000; Kemp et al., 2009b), but has been shown to be present with greater abundances at sites in New England (de
Rijk, 1995; de Rijk and Troelstra, 1997; Gehrels and van de Plassche, 1999; Edwards et al., 2004). *Haplophragmoides manilaensis* was present at a small number of high-marsh sites in North Carolina with an abundance of up to 31% (Kemp et al., 2009b). It was not shown to be a significant part of assemblages in Virginia (Ellison et al., 1965; Culver et al., 1996; Spencer, 2000) or Delaware (Hippensteel et al., 2000, 2002; Leorri and Martin, 2009). Its distribution as a major assemblage constituent in this study was limited to the Leeds Point site.

Middle-marsh assemblages associated with *Spartina patens, Distichlis spicata* and stunted *Spartina alterniflora* vegetation were represented by biozones C and D, which were dominated by *Arenoparrella mexicana* and *Tiphotrocha comprimata* respectively (Figure 5). In North Carolina, *Arenoparrella mexicana* was the characteristic middle-marsh species at sites along the Outer Banks barrier islands (Horton and Culver, 2008; Kemp et al., 2009b), whilst *Tiphotrocha comprimata* was only dominant at a small number of sites in middle-marsh settings. In Virginia, Spencer (2000) identified a middle to high marsh transitional assemblage dominated by *Tiphotrocha comprimata*. *Arenoparrella mexicana* was not a significant species on the Virginia salt marshes in studied by Spencer (2000) or Culver et al. (1996). Middle-marsh foraminiferal assemblages in Delaware included high abundances of *Arenoparrella mexicana*, whilst *Tiphotrocha comprimata* was not shown to be one of the most abundant species (Hippensteel et al., 2000, 2002).

Low-marsh assemblages at each of the three study sites were characterized by biozone B, dominated by *Miliammina fusca* (Figure 5). This assemblage is typical of areas inundated by tides on a daily basis (de Rijk, 1995) and has been described from settings with a wide range of
salinities (Murray, 1991). On the mid-Atlantic coast of the USA, it has also been described throughout North Carolina (Culver and Horton, 2005; Horton and Culver, 2008; Kemp et al., 2009b). It was also recognized by several studies from Virginia (Ellison et al., 1965; Ellison and Nichols, 1976; Culver et al., 1996; Spencer, 2000). Low-marsh assemblages of *Miliammina fusca* have been shown to be widespread in Delaware (Fletcher et al., 1993; Hippensteel et al., 2000, 2002; Leorri and Martin, 2009).

5.2 Sea-Level Reconstruction Techniques

The use of salt-marsh foraminifera in sea-level reconstructions is dependent on the selection and application of a suitable technique to exploit the modern relationship between foraminifera and tidal elevation to interpret downcore assemblages. Scott and Medioli (1978) proposed a vertical zonation of foraminifera described by faunal zones termed IA, IB, IIA and IIB, which were described qualitatively (species composition and elevational range) at Chezzetcook in Nova Scotia (Canada). The elevational distribution of the five biozones suggests that salt-marsh foraminifera are appropriate for use as a sea-level indicator in southern New Jersey. Other investigations subsequently sought to recognize this zonation at new sites (Gehrels, 1994) or applied it to reconstruct sea level in other regions (Peteet Carmichael, 1980). Later studies used unconstrained cluster analysis to quantitatively define assemblages of foraminifera on modern salt marshes (de Rijk, 1995; Horton, 1999; Patterson et al., 2000; Edwards et al., 2004). This approach required the subjective judgment of a researcher to decide on how many groups are present and is influenced by the limitations of agglomerative hierarchical clustering methods such as the legacy effect of previous cluster decisions (Kaufman and Rousseeuw, 1990). The PAM clustering technique employed in this study allows the user to objectively and
quantitatively estimate the number of groups present in the data by using reported silhouette lengths as a measure of fit (Kaufman and Rousseeuw, 1990). Further, the technique is robust to outliers and does not favor spherical clusters unlike other partitioning ($k$-means) methods (Kaufman and Rousseeuw, 1990). This approach improves upon the vertical zonation concept by providing a quantitative and objective definition of both the number and composition of foraminiferal assemblages.

5.3 Application of linear discriminant functions for estimating paleomarsh elevation

We used 32 samples from core EF10 as a test set (Figure 6). All samples from which foraminifera were enumerated were included. Application of the linear discriminant functions (Figure 6) to this test set estimated the probability that a sample should be allocated to each of the five prespecified classes (biozones A-E). We accepted assignment to a single biozone when the probability >0.95. In instances where the probability was <0.95, it was recognized that the sample could be from one or more biozone. The choice of 0.95 as a threshold value reflects its widespread usage for statistical significance (Scheffe, 1959).

There were 27 samples allocated to biozone A with a probability exceeding 0.95, two samples to biozone C and two samples to biozone D (Figure 6). No samples were associated solely with either biozones B or E. Three samples (at 0.6m, 2.5m and 3.1m) could have been from one of two biozones. The sample at 0.6m was assigned to either biozone A (probability of 0.73) or biozone B (probability of 0.27), which reflects unusually high abundances of *Miliammina fusca* (53%) in association with *Trochammina inflata* (23%) and *Jadammina macrescens* (18%; Figure 6). The sample at 2.5m was assigned to either biozone A (probability of 0.68) or biozone C.
(probability of 0.32) due to having a mixed assemblage (Figure 6) of *Jadammina macrescens* (16%) and *Trochammina inflata* (43%; biozone A) with *Arenoparrella mexicana* (18%; biozone C). The sample at 3.1m was assigned to either biozone A (probability of 0.49) or biozone B (probability of 0.51) as it included *Jadammina macrescens* (43%) with *Miliammina petila* (46%). This abundance of *Miliammina petila* is much higher than its occurrence in the training set, which had a maximum 19% in a sample from Leeds Point (transect two sample seven) and was pre-specified as part of biozone B (Figure 5). A single sample (2.8m) had some probability of being from one of three biozones (D = 0.70, A = 0.16, or C = 0.14), reflecting a diverse assemblage (Figure 6) including *Tiphotrocha comprimata* (29%; biozone D), *Trochammina inflata* (44%; biozone A) and *Arenoparrella mexicana* (11%; biozone C).

The allocation of core samples to biozones distinguished in the modern environment was used to provide downcore estimates of PME for use in sea-level reconstruction (Figure 7A). PMEs were converted from SWLIs to tidal elevations at the Leeds Point (Figure 6) site following the approach of Horton and Edwards (2006). In order to compare PME estimated using regional biozones and linear discriminant functions, we also estimated PME using two types of transfer function (Figure 7B) using the C2 computer program. Weighted-averaging partial least squares (WA-PLS) was used because it has been the most frequently applied method in reconstructions of Holocene sea level. No samples were removed from the modern training set and the results presented are from component two with cross validation. PME estimated using WA-PLS varied between 0.20 and 0.52m MTL with an average error of 0.14m (Figure 7B). The mid point of these estimates was within the boundaries established using linear discriminant functions, although the error bars leave this envelope at depths of 100cm, 260cm, 290cm and 310cm. In
addition, we used the modern analogue technique because it provides a means to judge the reliability of PME estimates. No modern samples were removed and cross validated results are presented. Estimated PME ranged from 0.14 to 0.52m MTL with an average error of 0.14m (Figure 7B). The mid point of these estimates was within the boundaries established using linear discriminant functions, although the error bars leave this envelope at depths of 100cm, 260cm and 290cm. Within the error bounds of both techniques, PME estimates from WA-PLS and MAT are in agreement. Indeed, the average difference between mid point estimates was 0.04m, with a maximum of 0.09m. The validity of the transfer function-derived estimates of PME are supported by comparison with results from linear discriminant functions. By recognition of discreet biozones in sedimentary sequences, linear discriminant functions are inherently less able to document subtle (within the elevational range of a particular biozone) changes in PME than transfer functions. The elevational range of each biozone is also likely to increase with the inclusion of additional modern samples from more sites in the region, although the recognition of new biozones may increase the precision of this technique. Linear discriminant functions provide a means to independently support inferences made using transfer functions.

Linear discriminant functions do not estimate the probability that a sample belongs to none of the biozones in the modern training set (Venables and Ripley, 2002). As such, the use of linear discriminant functions requires that the modern training set include samples from a sufficient variety of environments (both within and between sites) to provide an appropriate analogue for all core samples. This difficulty mirrors the discussion about the relative merits of local and regional datasets used in transfer functions (Gehrels, 1994; Gehrels et al., 2001; Horton and Edwards, 2005). The number and variety of samples needed to produce a suitable modern
training set should reflect the aims of the study and the likelihood that paleoenvironmental conditions at a site were similar to or significantly different from those encountered today (Woodroffe, 2009). The modern analogue technique provides an estimate of the dissimilarity between core samples and their closest modern analogue in the training set. It has been proposed that where this dissimilarity (which can be measured by numerous metrics) exceeds the 20th percentile of dissimilarity within the training set the core sample should be considered as having a poor modern analogue (Overpeck et al., 1985; Jackson and Williams, 2004; Kemp et al., 2009c; Woodroffe, 2009). In core EF10, use of this threshold would cause six samples to be classified as having no modern analogue (Figure 7C). For example, the sample at 310cm with an unusually high abundance of *Miliammina petila* (in comparison to the training set) clearly lacks a good modern analogue, which is also reflected in the uncertainty of its assignment using linear discriminant functions (Figures 6 and 7).

### 6 Conclusions

We described the modern distribution of salt-marsh foraminifera in southern New Jersey, USA using 56 surface samples collected along transects at three sites. Partitioning around medoids (PAM) cluster analysis was used to quantitatively and objectively define the number and composition of foraminiferal assemblages at each of the sites and demonstrated that foraminifera form elevation-dependent ecological zones making them appropriate sea-level indicators. All samples were amalgamated to create a regional dataset from which five biozones were identified and described using PAM. The elevational range of each biozone was conservatively estimated using the maximum and minimum of member samples. Classification of each modern sample as belonging to one of these five biozones enabled us to develop linear discriminant functions. We
applied these functions to quantitatively estimate the probability that 32 core samples of salt-marsh sediment shared an environmental origin with each biozone using preserved assemblages of foraminifera. The techniques described provide an objective and quantitative means to use salt-marsh foraminifera to reconstruct former sea level without underlying assumptions about species response and can test the ecological plausibility of other approaches including transfer functions.
Acknowledgments

We thank the Earthwatch Student Challenge Awards Program for funding and the student volunteers who helped so enthusiastically with fieldwork. The US Fish and Wildlife Service are thanked for granting access to the Edwin Forsythe National Wildlife Refuge. This paper is a contribution to IGCP project 588 “Preparing for coastal change”. SEE thanks the Greg and Susan Walker endowment to Prof. Robert Giegengack (University of Pennsylvania). CHV publishes with the permission of the Executive Director of British Geological Survey. Funding for this study was provided by NICRR grant. Jerry Mead (Academy of Natural Sciences, Philadelphia) kindly provided the RTK unit.
**Figure Legends**

**Figure 1:** Location of studies sites in southern New Jersey (USA). Modern transects across salt marshes at Leeds Point (A), Bass River (B) and Brigantine Barrier (C) were sampled for foraminifera. A core (EF10) was recovered from Leeds Point for analysis (A).

**Figure 2:** Distribution of foraminifera along salt-marsh transects from Leeds Point, New Jersey. Samples from transect one (A-A’) are labeled as A while samples from transect two (B-B’) are labeled B. (A) Elevational profile of the transects including zonation of vascular vegetation. (B) Average silhouette width estimated by partitioning around medoids (PAM) cluster analysis of foraminiferal data combined from both transects indicating that four groups (dashed vertical line) was appropriate. (C) Silhouette plot for PAM clustering of foraminiferal samples divided into four groups. Black and white bars are used to differentiate groups. Vertical dashed line is the average silhouette width. (D) Composition of the four groups of foraminifera identified by PAM clustering. Black and white bars alternate between groups to emphasize membership, each group is labeled (e.g. LP-1).

**Figure 3:** Distribution of foraminifera along a salt-marsh transect at Bass River, New Jersey. (A) Elevational profile of the transect including zonation of vascular vegetation. (B) Average silhouette width estimated by partitioning around medoids (PAM) cluster analysis of foraminiferal data indicating two groups can be differentiated (dashed vertical line). (C) Silhouette plot for PAM clustering of foraminiferal samples divided into two groups. Black and white bars are used to differentiate between group membership. Vertical dashed line is the
average silhouette width. (D) Composition of the two groups of foraminifera identified by PAM clustering. Black and white bars emphasize group membership.

**Figure 4:** Distribution of foraminifera along a salt-marsh transect at Brigantine Barrier, New Jersey. (A) Elevational profile of the transect including zonation of vascular vegetation. (B) Average silhouette width estimated by partitioning around medoids (PAM) cluster analysis of foraminiferal data indicating the presence of two groups (dashed vertical line). (C) Silhouette plot for PAM clustering of foraminiferal samples divided into two groups. Black and white bars are used to differentiate between group membership. Vertical dashed line is the average silhouette width. (D) Composition of the two groups of foraminifera identified by PAM clustering. Black and white bars emphasize group membership, \( Tc = Tiphotrocha comprimata \), \( Ab = Ammobaculites \) spp.

**Figure 5:** (A) The combined dataset of modern salt-marsh foraminifera from three sites in southern New Jersey. Samples are grouped into the five biozones (A-E) identified using PAM cluster analysis. \( Ab = Ammobaculites \) spp. The silhouette width for each sample is presented at right, the average for the complete dataset is 0.45 (dashed vertical line). (B) Plot of average silhouette width determined by PAM clustering when all samples of modern foraminifera were combined into a single dataset. The peak in silhouette width suggests that five assemblages can be recognized. (C) Modern samples of salt-marsh foraminifera positioned on the first two discriminant axes. Samples are divided into the five biozones recognized by PAM cluster analysis.
Figure 6: Foraminifera from core EF10. Low counts (indicated by hollow bars) are those from which less than 30 individuals were identified. Grey areas represent parts of the core where no foraminifera were found in samples. Samples were allocated to one of the five biozones (A-E) using debiased linear discriminant functions. The probability of being most similar to each biozone is shown for all samples in the right panel. Where the probability >0.95, samples were allocated to a single biozone. Other samples were assigned to more than one biozone. $Te = Tiphotrocha comprimata$.

Figure 7: Downcore estimates of paleomarsh elevation (PME) in core EF10 based on the tidal characteristics at Leeds Point. (A) Estimates from linear discriminant functions, the grey area represents the range from minimum to maximum values, symbols and error bars are biozone(s) average elevation ± 1σ range, vertical dashed line shows position of mean higher high water (MHHW). (B) Estimates from transfer functions developed using weighted average-partial least squares (WA-PLS; open circles, dashed error bars) and modern analogue technique (MAT; filled circles, solid error bars). Grey area is the same as that in panel A, vertical dashed line shows position of mean higher high water (MHHW). (C) Estimates of dissimilarity between core samples and modern salt marsh samples generated by the MAT transfer function. Vertical lines mark values for the 5th, 10th and 20th percentiles of dissimilarity in the modern training set, which have been used as thresholds in determining the strength of analogy between core and modern samples. Grey area represents core samples with “poor” modern analogues.
Table 1

Characteristics of groups and biozones identified by partitioning around medoids

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of Samples</th>
<th>Maximum Dissimilarity</th>
<th>Average Dissimilarity</th>
<th>Diameter</th>
<th>Separation</th>
<th>Average Sil. Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP-1</td>
<td>6</td>
<td>44.67</td>
<td>29.77</td>
<td>65.66</td>
<td>27.89</td>
<td>0.33</td>
</tr>
<tr>
<td>LP-2</td>
<td>3</td>
<td>51.67</td>
<td>28.27</td>
<td>52.28</td>
<td>27.89</td>
<td>0.32</td>
</tr>
<tr>
<td>LP-3</td>
<td>13</td>
<td>31.82</td>
<td>21.63</td>
<td>54.60</td>
<td>21.06</td>
<td>0.48</td>
</tr>
<tr>
<td>LP-4</td>
<td>4</td>
<td>45.43</td>
<td>22.75</td>
<td>51.40</td>
<td>21.06</td>
<td>0.35</td>
</tr>
<tr>
<td>BR-1</td>
<td>7</td>
<td>57.64</td>
<td>29.00</td>
<td>66.76</td>
<td>35.05</td>
<td>0.33</td>
</tr>
<tr>
<td>BR-2</td>
<td>8</td>
<td>44.30</td>
<td>23.41</td>
<td>58.56</td>
<td>35.05</td>
<td>0.45</td>
</tr>
<tr>
<td>BB-1</td>
<td>10</td>
<td>28.65</td>
<td>12.72</td>
<td>37.23</td>
<td>40.02</td>
<td>0.42</td>
</tr>
<tr>
<td>BB-2</td>
<td>5</td>
<td>59.09</td>
<td>29.73</td>
<td>59.09</td>
<td>40.02</td>
<td>0.74</td>
</tr>
<tr>
<td>Biozone A</td>
<td>13</td>
<td>40.09</td>
<td>16.99</td>
<td>59.04</td>
<td>21.73</td>
<td>0.57</td>
</tr>
<tr>
<td>Biozone B</td>
<td>16</td>
<td>47.21</td>
<td>26.97</td>
<td>66.15</td>
<td>17.06</td>
<td>0.30</td>
</tr>
<tr>
<td>Biozone C</td>
<td>15</td>
<td>35.28</td>
<td>21.95</td>
<td>58.38</td>
<td>17.06</td>
<td>0.44</td>
</tr>
<tr>
<td>Biozone D</td>
<td>6</td>
<td>38.32</td>
<td>22.03</td>
<td>51.67</td>
<td>30.89</td>
<td>0.42</td>
</tr>
<tr>
<td>Biozone E</td>
<td>6</td>
<td>42.74</td>
<td>28.75</td>
<td>65.58</td>
<td>25.74</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>Biozone</th>
<th>Hm</th>
<th>Jm</th>
<th>Tc</th>
<th>Ti</th>
<th>Am</th>
<th>Mf</th>
<th>Ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.0</td>
<td>34.9</td>
<td>19.0</td>
<td>40.5</td>
<td>1.5</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>B</td>
<td>0.4</td>
<td>4.6</td>
<td>6.2</td>
<td>1.2</td>
<td>3.3</td>
<td>63.1</td>
<td>19.1</td>
</tr>
<tr>
<td>C</td>
<td>0.0</td>
<td>2.3</td>
<td>7.8</td>
<td>21.1</td>
<td>45.3</td>
<td>18.8</td>
<td>0.0</td>
</tr>
<tr>
<td>D</td>
<td>5.3</td>
<td>2.9</td>
<td>44.4</td>
<td>12.6</td>
<td>18.4</td>
<td>12.6</td>
<td>1.9</td>
</tr>
<tr>
<td>E</td>
<td>61.7</td>
<td>8.4</td>
<td>1.9</td>
<td>7.0</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Medoids of the five foraminiferal assemblages identified by PAM clustering of the dataset of all modern foraminifera. These assemblages (in percentages) represent the population for which average dissimilarity to all samples in a given cluster is minimal. Only species important for distinguishing biozones are shown, bold values highlight dominant species in each biozone. Hm = *Haplophragmoides manilaensis*, Jm = *Jadammina macrescens*, Tc = *Tiphotrocha comprimata*, Ti = *Trochammina inflata*, Am = *Arenoparrella mexicana*, Mf = *Miliammina fusca*, Ab = *Ammobaculites* spp.
The elevational range of the five biozones identified by PAM cluster analysis was conservatively estimated by using the maximum and minimum values of samples within each group. Also presented are the average elevation and 1σ range of each biozone. Due to differences in tidal range among sites, elevations are expressed using a standardized water level index (SWLI), where a value of 100 represents MHHW and 0 is MLLW. MTL = mean tide level.

<table>
<thead>
<tr>
<th>Biozone</th>
<th>Max m MTL</th>
<th>Min m MTL</th>
<th>Average m</th>
<th>1σ Range m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SWLI</td>
<td>SWLI</td>
<td>SWLI</td>
<td>SWLI</td>
</tr>
<tr>
<td>A</td>
<td>0.79</td>
<td>0.18</td>
<td>0.48</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>107.63</td>
<td>63.26</td>
<td>85.23</td>
<td>13.19</td>
</tr>
<tr>
<td>B</td>
<td>0.31</td>
<td>-0.48</td>
<td>0.00</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>74.58</td>
<td>11.02</td>
<td>48.10</td>
<td>20.72</td>
</tr>
<tr>
<td>C</td>
<td>0.59</td>
<td>0.27</td>
<td>0.46</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>100.02</td>
<td>71.52</td>
<td>88.58</td>
<td>8.59</td>
</tr>
<tr>
<td>D</td>
<td>0.59</td>
<td>0.26</td>
<td>0.39</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>99.88</td>
<td>70.26</td>
<td>82.11</td>
<td>12.76</td>
</tr>
<tr>
<td>E</td>
<td>0.91</td>
<td>0.65</td>
<td>0.74</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>129.29</td>
<td>105.87</td>
<td>113.69</td>
<td>8.50</td>
</tr>
</tbody>
</table>
References


