Supplement S1. Calculation of benthic ΔR
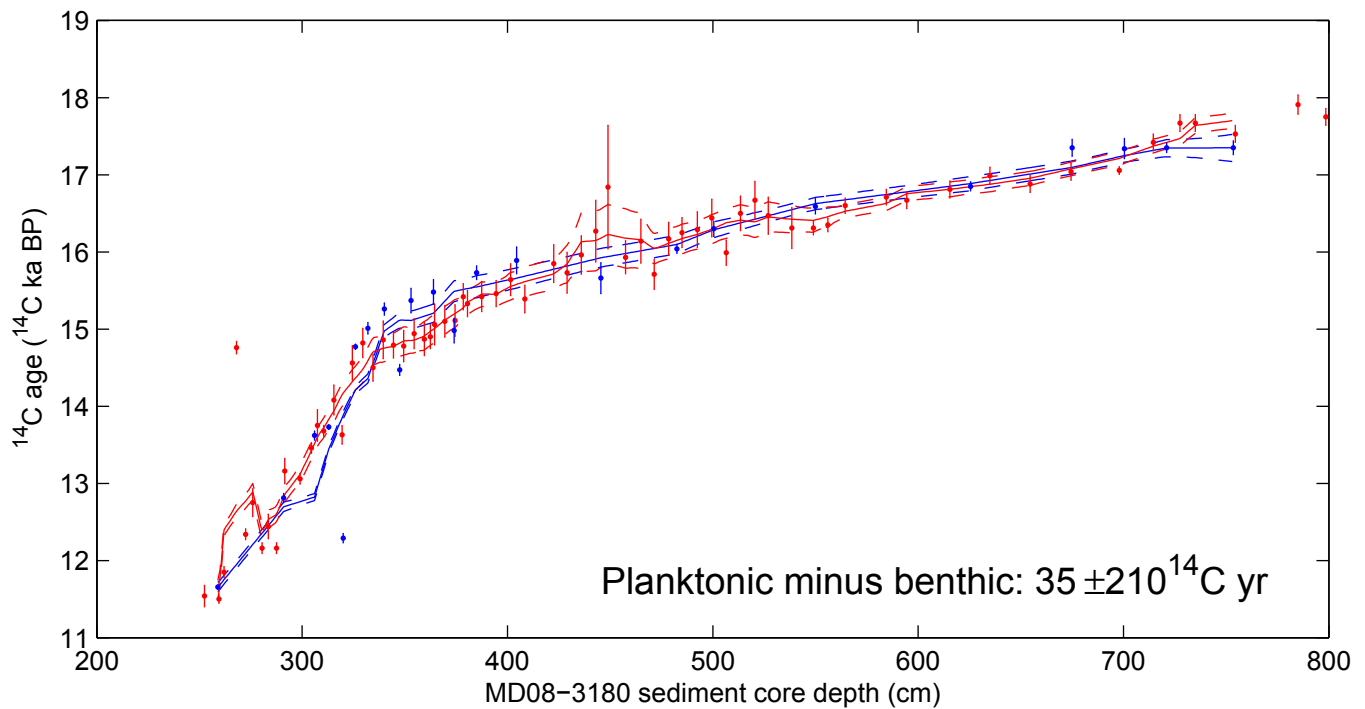

Fig. S1. Data from core MD08-3180 (Sarnthein et al., 2015) used to determine planktonic - benthic ^{14}C age offset for the Azores region. Shown are multi-specimen ^{14}C determinations and 1σ measurement errors from planktonic (red circles and vertical error bars) and benthic foraminifera (blue circles and vertical error bars). The planktonic and benthic ^{14}C datasets were separately smoothed 10^4 times each using five point running mean of probability weighted random samples of each ^{14}C age normal distribution, after which all smoothing iterations were interpolated to 1 cm resolution for the interval that the two datasets overlap. Median ^{14}C age (solid red and solid blue lines for planktonic and benthic foraminifera, respectively) and 95.4% [14]C age ranges (dashed red and dashed blue lines for planktonic and benthic foraminifera, respectively) for each 1 cm were subsequently calculated from the 10^4 smooths. For every 1 cm, benthic median ^{14}C age is subtracted from the planktonic median ^{14}C age, resulting in a ^{14}C age difference for each 1 cm. The mean and standard deviation of these age differences were subsequently calculated, resulting in a value of 35±210 ^{14}C yr (rounded following Stuiver and Polach (1977)).

Supplement S2. Simulating of the influence of multi-specimen sample size upon the concealment of age-depth reversals

AMS and other forms of mass spectrometry analysis report only a mean value and machine measurement error for a particular sample, whereby no information on actual intra-sample heterogeneity is provided. When the machine measurement error is smaller than then the actual intra-sample heterogeneity, the latter is concealed from the researcher. An increasing number of specimens within a multi-specimen sample from a discrete core depth can cause a sample from a given core depth interval increases the likelihood that the sample's mean will age to regress towards a particular mean age for that depth interval. We simulate eight different core scenarios in Fig. S2, with two different simulated PDSM intensities (high and low) and four different multi-specimen sample sizes (random picking of specimens to test sample sizes of n = 1, 5, 10 and 100 specimens). In each case the average age of the multi-specimen sample is calculated. We note that when the multi-specimen sample size consists of 10 specimens or more, both the high and low PDSM scenarios show downcore average age values that are in chronological order, i.e. no age-depth outliers are present. Mass spectrometry analysis of foraminifera has traditionally required between tens (Metcalfe et al., 2015; Waelbroeck et al., 2005) and hundreds (Hughen et al., 2006) of foraminifera specimens to establish a successful measurement. In other words, researchers applying the longstanding age-depth model method to sediment cores similar to those simulated in Fig. S2 would not be able to discern from the age-depth model results whether or not significant PDSM is present, with consequences for any subsequent temporal paleoclimate interpretations.
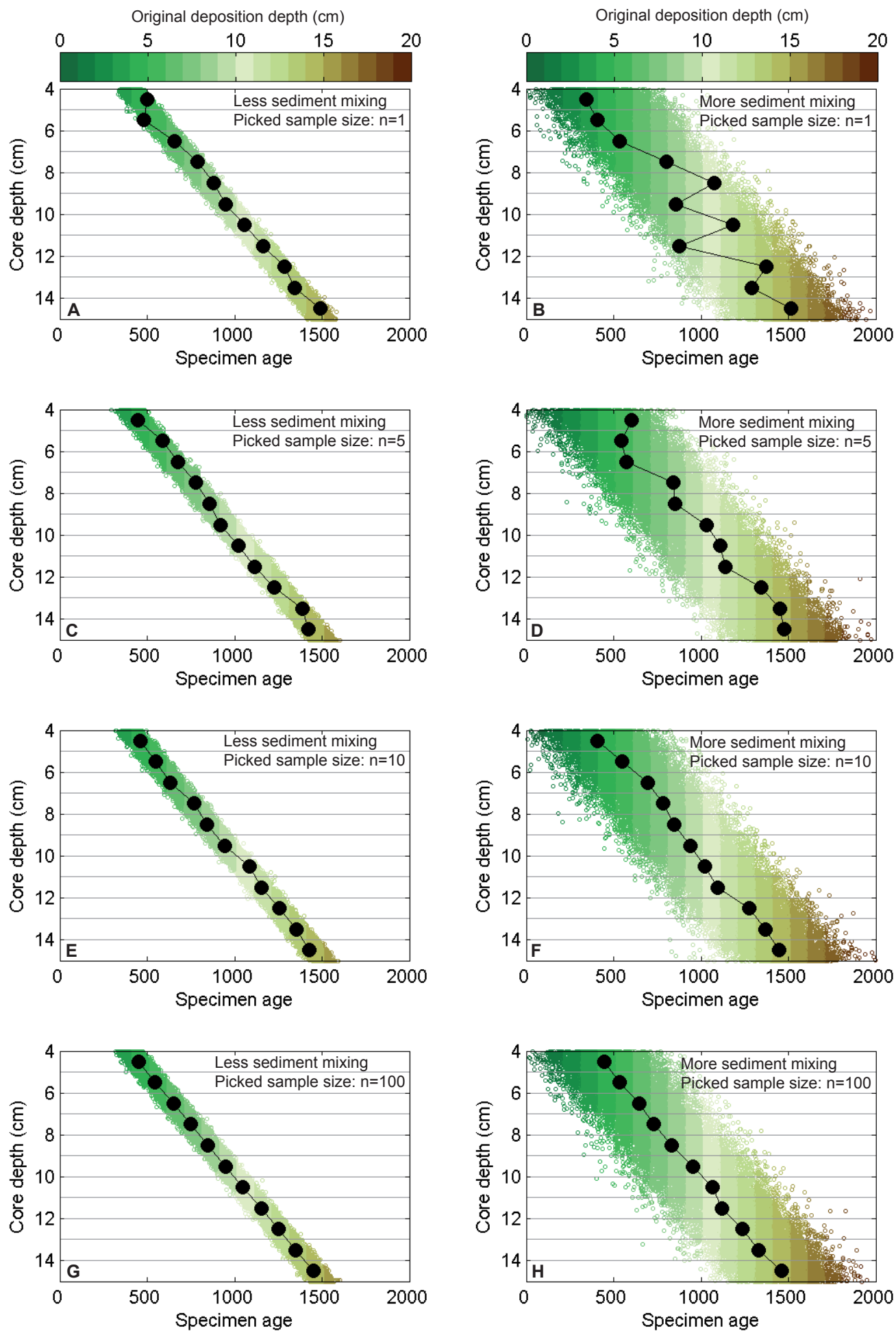
Fig. S2. Eight computer generated sediment core scenarios, each with 5·10⁴ synthetic single specimen foraminifera (open circles shading from green to brown) deposited with a linear sedimentation rate of 10 cm/ka. Gaussian noise has been added to simulate PDSM of the core material. For each 1 cm discrete depth interval (indicated by horizontal grey lines) in each scenario the average age value of *n* number of randomly picked synthetic foraminifera has been calculated (filled black circle). **(A)**: Less vertical mixing, *n*=1. **(B)**: More vertical mixing, *n*=1. **(C)**: Less vertical mixing, *n*=5. **(D)**: More vertical mixing, *n*=5. **(E)**: More vertical mixing, *n*=10. **(F)**: More vertical mixing, *n*=10. **(G)**: More vertical mixing, *n*=100. **(H)**: More vertical mixing, *n*=100.

Supplement S3. Post-depositional ranking change of single foraminifera vs foraminifera weight
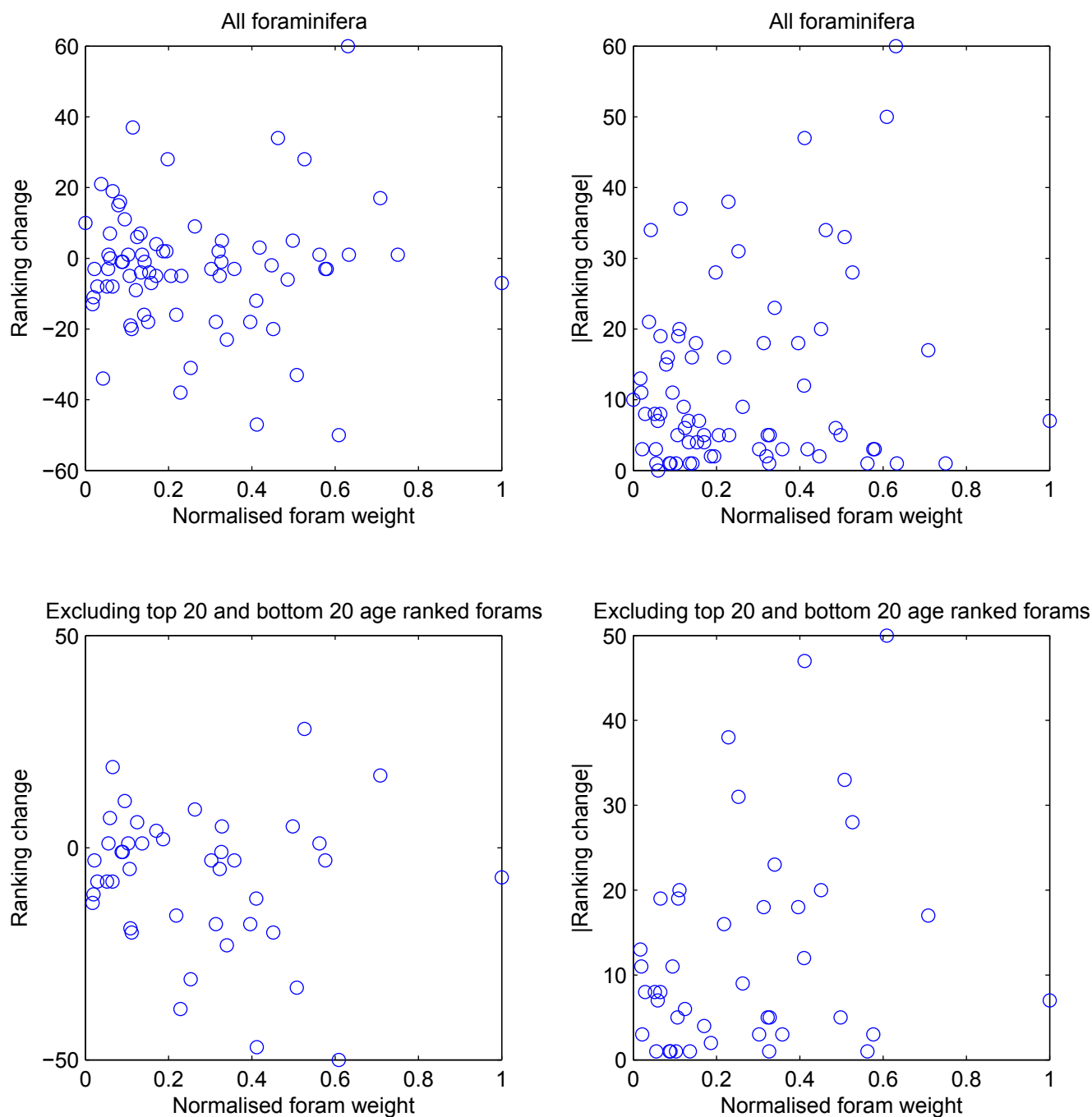

Fig. S3. Change in foraminifera ranking (as presented in Fig. 3) plotted against normalised (between 0 and 1) foraminifera weight. (A): Ranking change vs normalised foraminifera weight for all foraminifera. (B): Absolute value of ranking change vs normalised foraminifera weight for all foraminifera. (C): Ranking change vs normalised foraminifera weight, whereby the 20 oldest and 20 youngest foraminifera have been excluded. (D): Absolute value of ranking change vs normalised foraminifera, whereby the 20 oldest and 20 youngest foraminifera have been excluded.