Comparing summed probability distributions of shoreline and radiocarbon dates from the Mesolithic Skagerrak coast of Norway

29 June, 2023

1 Introduction

Population size is regarded as an important driver of cultural variation, and is of critical importance to our understanding of past human societies (e.g. French et al., 2021; Shennan, 2000). The frequency distribution of radiocarbon dates has been used extensively as one proxy for past relative population sizes (e.g. Crema, 2022; French et al., 2021), including in Norwegian Stone Age archaeology (Bergsvik et al., 2021; Jørgensen et al., 2020; Lundström et al., 2021; Nielsen, 2021; Nielsen et al., 2019; Solheim, 2020; Solheim and Persson, 2018; see also Persson, 2009). The potentially immense value of insights into past population dynamics, combined with the ubiquity of radiocarbon dates and the relative ease with which these can now be processed within what has been termed a dates-as-data methodology (Rick, 1987) has undoubtedly contributed to the popularity of the approach. Several limitations and forms of criticism have, however, been directed at these procedures. Some of the objections are of a methodological nature, while others pertain to the underlying logic and the degree to which there is likely to be a direct connection between the frequency of ^{14}C-dates and population dynamics (e.g. Carleton and Groucutt, 2021; Torfing, 2015; Ward and Larcombe, 2021). What appears to be agreed upon by practitioners and critics alike is that radiocarbon dates are best analysed in this manner when compared and contrasted to other proxies for past population dynamics and other variables that might impact these (Bergsvik et al., 2021; French, 2016; Palmisano et al., 2021, 2017). This paper reports on and begins to unpack the relationship between two measures that have been linked to relative population size in the context of the Mesolithic Skagerrak coast in south-eastern Norway, namely summed probability distribution of calibrated radiocarbon dates and the summed probability distribution of shoreline dated sites (hereafter RSPD and SSPD, respectively).

2 Background

Large parts of the post-glacial landscape of Northern Scandinavia is characterised by dramatic isostatic uplift that has led to a net sea-level fall throughout the Holocene, despite eustatic sea-level rise (e.g. Mörner, 1979). As coastal foragers appear to have predominantly settled on or close to the contemporaneous shoreline, this can be utilised to assign an approximate date to the sites. This is done by coupling the present-day altitude of the sites with reconstructions of past shoreline displacement – a method known as shoreline dating. This is not least useful for dating the large number of surveyed sites in the region where other temporal data that often follow with an excavation are not available, such as radiocarbon dates or typological indicators in artefact inventories.

The frequency of shoreline dated sites has also been compared to RSPDs in the past (Solheim and Persson, 2018; Tallavaara and Pesonen, 2020). However, this has been done by finding point estimates of shoreline dates that are then aggregated in somewhat arbitrary bins of 200 or 500 years. This therefore does not take into account the uncertainty in the distance between the sites and the contemporaneous shoreline, nor the impact the variability in the rate of sea-level change has on the precision of the dates that can be achieved with the method. In a recent study, Roalkvam (2023a) has presented a probabilistic method for shoreline dating that takes these parameters into account, thus setting the stage for a more refined investigation of the relationship between frequency of ^{14}C-dates and shoreline dated sites. The parametrisation of the method

was based on simulating the distance between sites with ^{14}C-dates and the prehistoric shoreline along the Skagerrak coast in the region between Horten municipality in the northeast to Arendal municipality in the southwest (see map in Figure 1). The results of the analysis indicates that the sites tend to have been located close to the shoreline until just after 4000 BCE when a few sites become more withdrawn from the shoreline, followed by a clear break around 2500 BCE, at which point shoreline dating appears to loose its utility. Thus, this geographical and temporal limit are also used for this study.

This paper is decidedly exploratory. While some speculative thoughts concerning the relationship between the variables outlined above underlie the analysis, these could not be instantiated as concrete hypotheses. The data have thus been explored using a set of standard models to begin to unpick patterns and grasp the relationship between the variables. Furthermore, the dates as data approach is dependent on there being a direct link between the past generation of material that ultimately become ^{14}C samples, and population size. The sum of shoreline dated sites, on the other hand, is determined by site frequency, and if used as a proxy for population size is therefore dependent on there being a connection between site count and population size. If a comparison of these proxies do not, or only partially correspond, it is thus an open question what factors have impacted either distribution to cause this discrepancy, and which measure, if any, reflects true population dynamics. The issue will therefore initially demand an open and exploratory approach where a multitude of explanatory and confounding effects can be drawn on to suggest explanations of any observed pattern.

2.1 Population dynamics and summed probabilities

To what degree the radiocarbon record is determined by past population numbers might vary both geographically and chronologically, based on variation in investigatory and taphonomic factors (Bluhm and Surovell, 2019; Surovell et al., 2009), as well as cultural processes within prehistoric populations. One example of the latter is the difference that might exist between farmer and forager populations, where Freeman et al. (2018) have suggested that an increased per capita energy consumption introduced with farming means that ^{14}C-dates should not be weighted equally when making relative population estimates across such populations. Similarly, while site counts have also been invoked for the analysis of past population dynamics, these are likely to be impacted by factors such as land-use and mobility patterns, and settlement nucleation and dispersion (Palmisano et al., 2017). However, these could be considered theoretical issues that have implications for how fluctuations in these proxies should be interpreted, and possibly weighted differently according some criteria. Before any such fluctuations are given any substantive interpretation, however, there are a host of methodological issues that have to be considered.

The most critical of these follow from the fact that the summation of the probabilities associated with the dates for the SPDs is not a statistically coherent procedure. This is because the summed probabilities can no longer been seen as probabilities, but rather represent the combination of events and uncertainties, making the two indistinguishable, and rendering the interpretation of the resulting sum difficult (Blackwell and Buck, 2003; Crema, 2022). As Timpson et al. (2021, p. 2) put it: 'the SPD is not *the* single best explanation of the data, nor even *a* explanation of the data, but rather a conflation of many possible explanations simultaneously, each of which is mired by the artefacts inherited from the calibration wiggles.' The SPD is not a model. It is the combined representation of a range of possible explanations for the data – the frequency of dated events combined with the variable uncertainty associated with these (Carleton and Groucutt, 2021). This means that an SPD cannot be directly analysed to draw inferences on population dynamics, nor can it be directly compared to other time-series data (Carleton et al., 2018; Timpson et al., 2021). While this problem can never be entirely resolved, a range of approaches have been developed in an attempt to work around this issue.

The most commonly applied of these is a null-hypothesis significance testing approach by means of Monte Carlo simulation, as introduced by Shennan et al. (2013) and later expanded upon by Timpson et al. (2013). This works by comparing the observed RSPD with a series of simulated RSPDs, generated from a null-model. These null-models are typically a uniform, exponential or logistic distribution. These are chosen *a priori* with reference to common long-term population dynamics and are typically parametrised by fitting the models by means of regression or maximum likelihood estimation (MLE). The result from these simulations are then

used to create a 95% critical envelope representing the null model. The proportion of the observed RSPD that falls outside this envelope is then used to estimate a global p-value indicating whether the null model can be rejected. In the case that it can, the portions of the observed RSPD that falls outside this envelope can subsequently be interpreted as representing potentially meaningful demographic events, relative to the null model. However, care has to be taken in how these are interpreted. First, this follows from the fact that 5% of the deviations from the critical envelope can be expected to be random, and there is no way to know which deviations this pertains to. Secondly, a for example exponential null model fit to the data is only one of an infinite set of exponential models with different growth rates that could be used. While a model fit by means of MLE will likely have a reasonable growth rate, and by extension exclude many other exponential fits as likely to explain the data, this can be difficult to determine. Finally, the p-value only indicates whether or not the null-model as a whole can be rejected as an explanation of the data, and does not provide statistical justification for interpreting local deviations themselves as meaningful demographic signals, as has often been the case (see review by Crema, 2022; Timpson et al., 2021).

2.2 Model comparison and continous piece-wise linear models

The procedure outlined above represent the standard approach in efforts to model population dynamics based on RSPDs. This can be helpful for determining whether the population behaviour has followed the trajectories of standard theoretical models, and is a useful first step here also for the treatment of shoreline dated sites, as few assumptions concerning the verall development of this proxy could be made. However, standard models frequently used to evaluate RSPDs are consistently being rejected in the literature, and deviations from these rejected null-models do not in themselves allow for any interpretations of the deviations, nor a subsequent direct analysis of the SPDs themselves. As a result, and by improving on Goldberg et al. (2016), Timpson et al. (2021) have recently suggested an alternative framework for modelling past population dynamics by use of RSPDs, published with the accompanying R package *ADMUR*. This involves fitting continuous piece-wise linear (CPL) models to the data, in addition to the standard models, subjecting these to a procedure of model comparison, and finally evaluating the best model using a goodness-of-fit test, in line with the Monte Carlo approach outlined above. The CPL models are defined by discrete linear periods of development between hinge points that mark points in time in which the trajectory of the population proxy changes direction. This therefore allows for the identification of key events and a conservative description of population dynamics between these. The approach thus also provides a clearer inferential justification for a further analysis and substantive interpretation of the identified best model. Following the evaluation of standard models as outlined above, the data were therefore subsequently subjected to the approach presented by Timpson et al. (2021).

3 Methods and data

Sites surveyed by means of test-pitting between Horten and Arendal were initially retrieved from the national heritage database Askeladden (Norwegian Directorate for Cultural Heritage, 2018), totalling at 1299 records. The records were then manually reviewed and given a quality score based on the criteria in Table 1, indicating the degree to which the spatial location and extent of the sites is believed to be represented in the geometries available in the database (see also Roalkvam, 2020). All sites with a quality score of 4 or worse were excluded from further analysis. Any sites situated at elevations that result in a shoreline date earlier than 9469 BCE were then excluded. This marks the latest start-date among the employed displacement curves, and no sites are yet verifiably older than around 9300 BCE (Glørstad, 2016). Data on excavated sites was originally compiled for Roalkvam (2023a) and has been compared with site data as listed in Damlien et al. (2021) and Nielsen et al. (2019). Only excavated sites with available spatial data in Askeladden or local databases at the Museum of Cultural History of the University of Oslo were included in the analysis. The 102 excavated sites in the dataset without relevant ^{14}C-dates and that were originally shoreline dated in the reports were included in the SSPD along with the retained surveyed sites. This gave a total of shoreline dated sites in the final SSPD.

Table 1: Quality scoring of site records of surveyed sites retrieved from the national heritage database Askeladden. The scoring system was first used in Roalkvam (2020).

Definition	Quality	Count
Site delineated by use of a GNSS-device, or a securely georeferenced record. Extensive database entry.	1	353
Secure spatial data. Slight disturbance of the site or somewhat lacking database record.	2	340
Secure spatial data. Damaged site, such as outskirts of a quarry, and/or very limited database entry.	3	148
Surveyed by archaeologists. However, the database entry is extremely limited/unclear, the site geometry is only given as a point or small automatically generated circle, and/or finds are from the topsoil of a field.	4	165
Likely site but uncertain spatial information. Typical example is recurring stray finds in a field or other larger area.	5	124
Single stray find or unverified claims/suggestions of possible site.	6	169

The borders of the municipalities within which the shoreline dated sites are located were used to limit the radiocarbon sample. Radiocarbon dates were taken from Roalkvam (2023a) and (**solheim?**). Dates done on food crusts were then excluded due the issue of marine reservoir effects (Nielsen et al., 2019, p. 83), resulting in a final total of ^{14}C-dates. Following both from the point made by Freeman et al. (2018) concerning the comparison of radiocarbon dates from populations with different economic modes, as noted above, and from the fact that shoreline dating appears less reliable after 4000 BCE and appears to loose its utility c. 2500 BCE, the Mesolithic is the main focus of this analysis. Consequently, probabilities falling later than 2500 BCE were excluded from the analysis. Probabilities falling after 4000 BCE and the first, possible small-scale introduction of agriculture in south-eastern Norway were retained to account for edge-effects, but these results should be treated with care for the reasons noted above.

All analyses done in this study were performed using the R programming language (R Core Team, 2020). Underlying data and programming code used for the paper is available in a version-controlled online repository at X. This is structured as a research compendium following Marwick et al. (2018), to allow for reproducibility of the results (see also Marwick, 2017). Analysis of the ^{14}C-dates were done using the R package *ADMUR* (Timpson et al., 2021). The R package *shoredate* was used for performing and handling the shoreline dating of sites (Roalkvam, 2023b), with re-purposed code from both *ADMUR* and *rcarbon* (Crema and Bevan, 2021) for Monte Carlo simulation and model comparison.

3.1 Summed probability of calibrated radiocarbon dates

To account for investigatory bias that can result from variable sampling intensity between sites, the summing of radiocarbon dates with *ADMUR* starts with binning dates from the same sites that fall within 200 uncalibrated ^{14}C years of each other when measured from the mean ^{14}C age. All dates are then calibrated using a resoltuion of 5 years and then normalised to sum to unity. The RSPDs for each site-phase are then summed to achieve the final SPD, which is in turn normalised. All calibrations were performed using the IntCal20 calibration curve (Reimer et al., 2020).

The final RSPD was subjected to the standard null-hypothesis testing approach through Monte Carlo simulation, as introduced above (see Shennan et al., 2013; Timpson et al., 2013), by fitting an exponential, logistic and uniform model to the observed RSPD. Following Timpson et al. (2021), the model parameters were identified by MLE search, using the differential evolution optimization algorithm *DEoptimR* (Brest et al., 2006). The Monte Carlo simulations were then performed using the *SPDsimulationTest()* function from *ADMUR*. For each model, a series of individual calendar years are drawn from the model distribution, with replacement, the number of which equals the number of bins in the observed RSPD. These are then 'uncalibrated' to a single value on the ^{14}C scale and a random error from among the observed errors is added to the date. These are then calibrated back to the calendar scale and finally summed. Here, this procedure was repeated 10000 times for each null model. The 2.5th and 97.5th quantile of the resulting summed probability for each year across all simulations are then retrieved to create the 95% critical envelope with which to compare the observed RSPD. The degree to which the observed RSPD deviates from the critical envelope is then used to calculate a global p-value, indicating whether or not the null can be rejected.

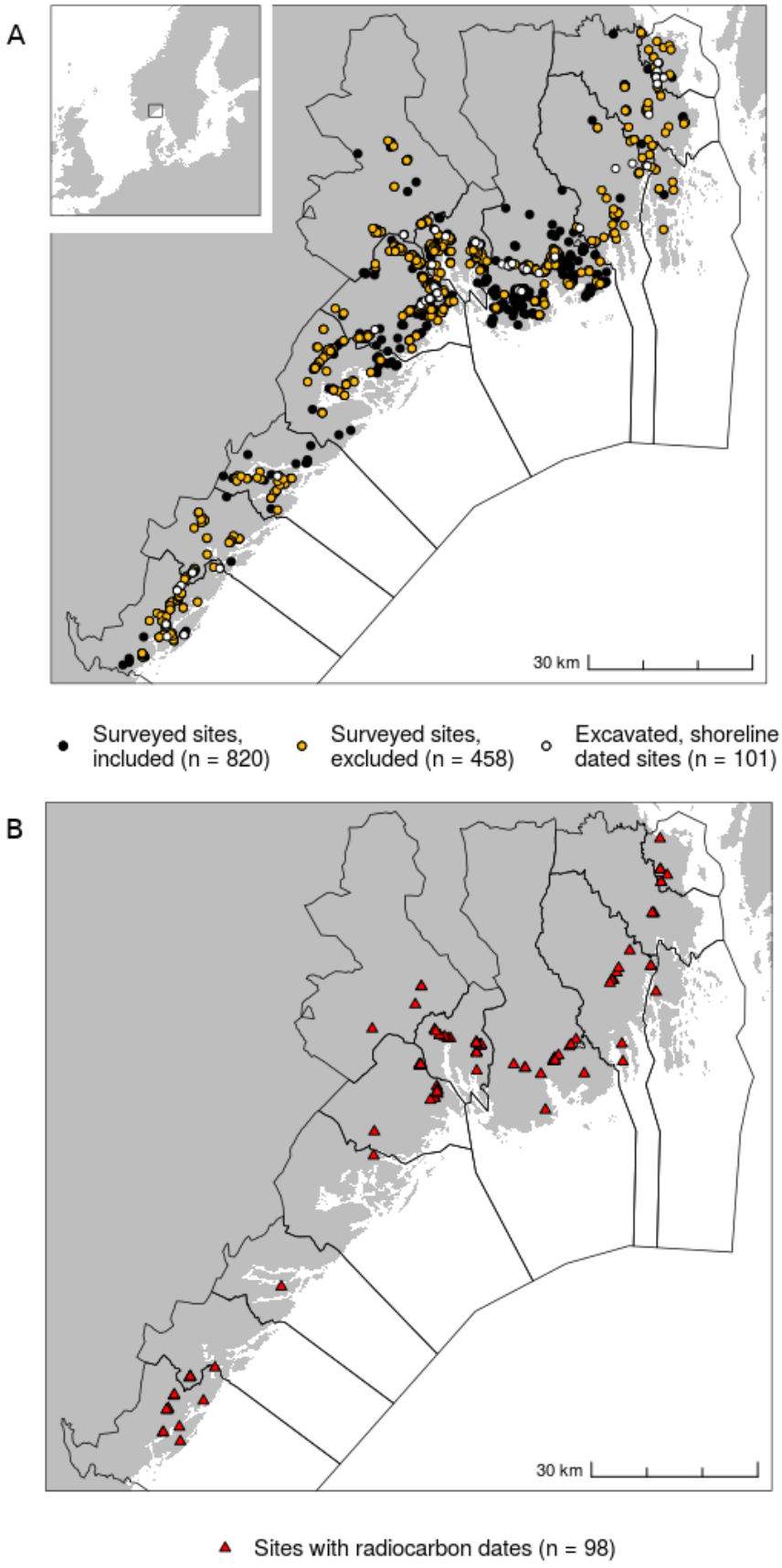


Figure 1: Map of the study area and shoreline dated sites. Black lines indicate the borders between municipalities. The surveyed sites included in the analysis are the ones given a quality score of 3 or higher using the framework in Table 1.

3.2 Summed probability of shoreline dated sites

Summing the probability of the shoreline dated sites and the model-fitting procedures followed the same structure as that for radiocarbon dates and was partly based on re-purposed programming code from *rcarbon* and *ADMUR*. However, idiosyncrasies in the dating method did necessitate some adjustments. To illustrate this, the procedure for shoreline dating a single site, as suggested in Roalkvam (2023a), is provided in Figure 2 and outlined below.

Four geological reconstructions of shoreline displacement in the region lays the foundation for the method as implemented here (Figure 2B). These shoreline displacement curves are from Horten (Romundset, 2021), Porsgrunn (Sørensen et al., in press, 2014a, 2014b), Tvedestrand (Romundset et al., 2018; Romundset, 2018) and Arendal (Romundset, 2018), each associated with a shoreline isobase along which the trajectory for relative sea-level change has been the same (see Svendsen and Mangerud, 1987). The first step in the dating procedure is to interpolate the shoreline displacement to the location to be dated. This is done by inverse distance weighting (e.g. Conolly, 2020), interpolating the relative sea-level change to the location of a site to be dated, based on it's distance from the isobases of the displacement curves.

With an estimated trajectory of shoreline displacement at the site to be dated, the elevation of the site can be drawn on to find a likely date for when it was in use, under the assumption that it was in use when located on or close to the shoreline. Figure 2C indicates where the elevation of the example site intersects the interpolated displacement curve. As it can be reasonably be assumed that a site was not in use when located under water, this effectively represents a *terminus post quem* (TPQ) date. However, a fairly common approach to shoreline dating has been find where the elevation of the site intersects the mean of the upper and lower limit of the displacement curve, and then adding a somewhat arbitrary error of $\pm$ 100 years to this point-estimate. However, finding TPQ dates that do not account for the likely distance between site and sea when the site was in use limits the further inferential steps that can be taken. Furthermore, the more common practice of adding a constant error to a point-estimate does not take into account that the rate of shoreline displacement varies both spatially and temporally, and therefore introduces bias to the result on both the spatial and temporal scale, while also overestimating the precision of the method in most cases.

This last point can be illustrated by finding the range between the upper and the lower limit of the four geological displacement curves for every meter from the marine limit and down to elevations that do not give an earliest date younger than 2500 BCE. This gives a mean TPQ range of 276 years with a standard deviation of 156 years across all four displacement curves, with this range varying significantly over time and between curves.

The analysis in Roalkvam (2023a) found that site phases ^{14}C-dated to before 2500 BCE were found to have a likely elevation above sea-level that can be reasonably approximated by a gamma distribution with shape = 0.286 and scale = 20.833. This is incorporated in the procedure for shoreline dating used for this study (Figure 2D). When performing this dating procedure, the gamma distribution is sequentially stepped through and transferred to the calendar scale by uniformly distributing the probability across the years in the range between the lower and upper limit of the interpolated displacement curve. This procedure thus gives the the shoreline date of a site when accounting for the likely elevation of the site above sea-level when it was in use (see Roalkvam, 2023a for details). Given that the shoreline displacement curves have no inversions and should therefore before commutative (cf. Weninger et al., 2015, p. 545), each shoreline date is normalised to sum to unity. To reduce the computational cost of the simulation procedures to follow below, the gamma distribution is here stepped through at increments of 0.1 m and the calendar scale is kept at a resolution of 5 years.

The implications of different approaches to shoreline dating and summing the results is further illustrated in Figure 3. In Figure 3A the sites to be shoreline dated are plotted according to their elevation. Due to the variable rates of shoreline displacement within the study region, the same elevation does not necessarily equate to the same date, and so this does not directly tell us much about their temporal distribution. For example, as the trajectory for sea-level regression is starker towards the north-east, the same elevation at a location to the south-west implies a younger date that one towards the north-east.

When the frequency distribution of shoreline dates have been investigated in the past, this has typically involved finding point-estimates of dates by either using the mean of displacement curves, or by taking mean

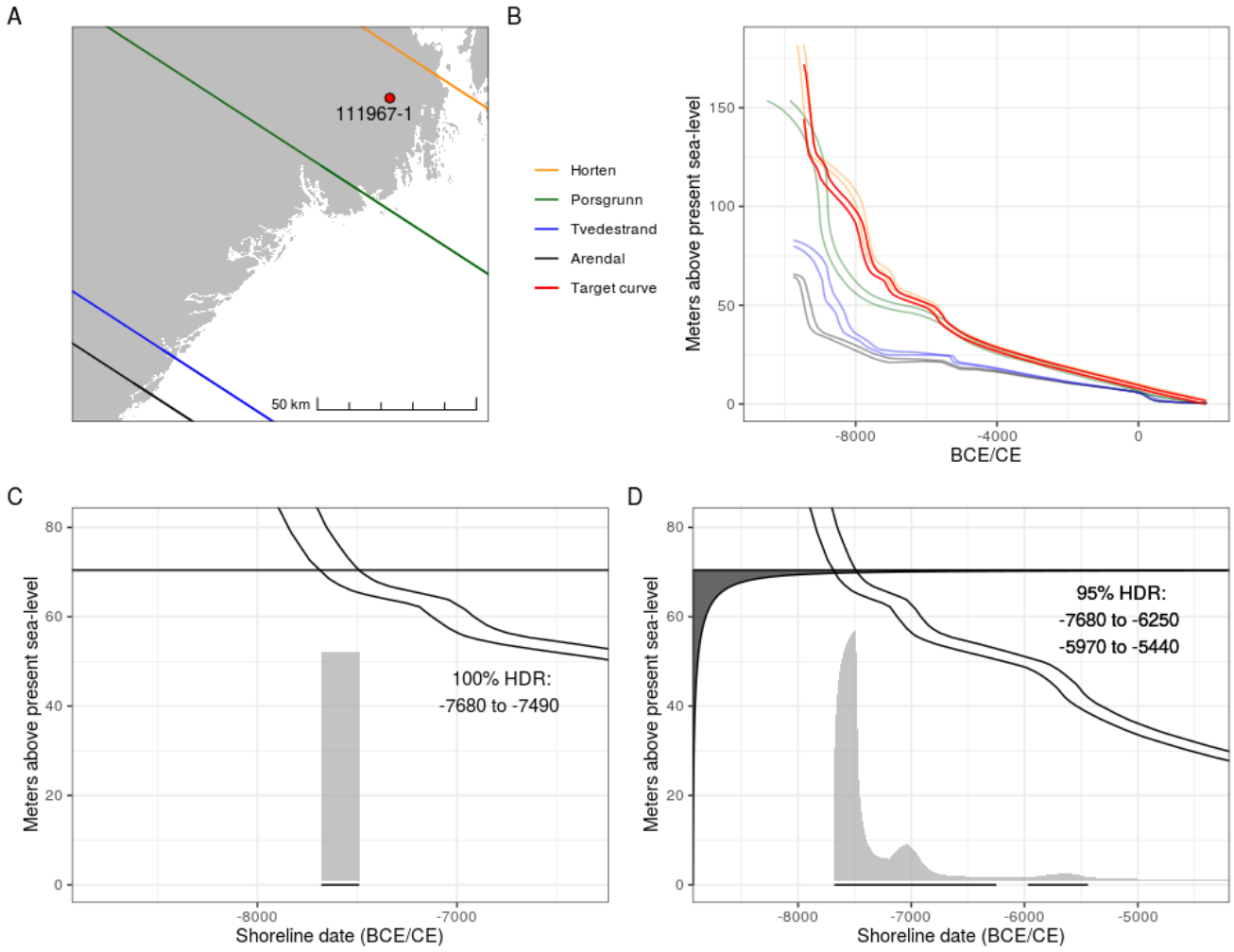


Figure 2: A) An example site (Askeladden ID 111967-1) relative to the isobases of the displacement curves. B) The geological displacement curves in the region and the curve interpolated to the example site. C) The black line on the y-axis represents the elevation of the site. This crosses the interpolated displacement curve at 7680 and 7490 BCE, which thus in practice gives an estimate of when the location of the site emerged from the sea, giving a *terminus post quem* date. D) The shoreline dating approach used for this study, where the gamma distrbution on the y-axis describes the likely vertical relationship between the shoreline and the site when it was in use.

of the TPQ dates for each site (as illustrated in Figure 2C). These point-estimates are then aggregated in bins of 200 or 500 years (e.g. Breivik, 2014; Fossum, 2020; Mjærum, 2022; Solheim and Persson, 2018; Tallavaara and Pesonen, 2020). This approach is illustrated in Figure 3B. The SSPD resulting from instead using the probabilistic method of shoreline dating, given in Figure 3C, demonstrates that accounting for the site-sea relationship gives a substantially different result compared to the more traditional approach, which is clearly indicated by the shifts in the positions of peaks and troughs in the distribution after c. 8000 BCE.

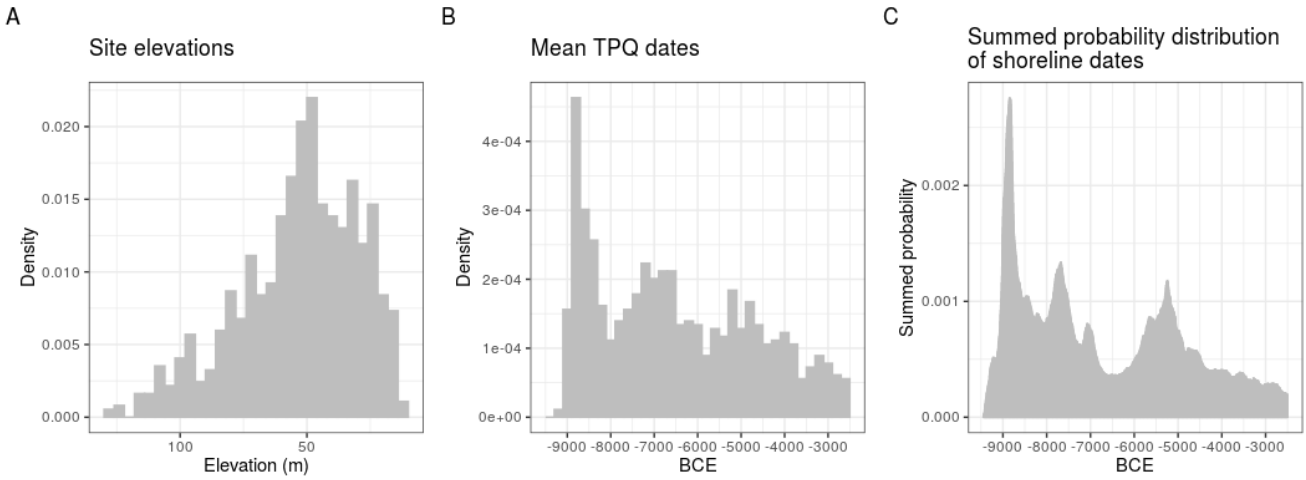


Figure 3: A) The distribution of the shoreline dated sites (n = 921) by altitude above present sea-level. The binwidth is 4 m. B) The mean of *terminus post quem* dates for the sites, plotted with a binsize of 200 years, reflecting the common approach to summing shoreline dates. C) The result of instead summing the probability of the shoreline dates achieved with the method employed here, which accounts for the likely distance between sites and shoreline. The binwidth is 5 years, following the resolution used with the dating procedure.

However, in the same way is with RSPDs, the SSPD cannot be directly interpreted as reflecting the intensity of sites over time, due to the summation issues introduced above. Furthermore, similar to how characteristics of the calibration curve can introduce bias to the RSPD, the same is true for the SSPD, where the local trajectory of relative sea-level change effectively functions as the calibration curve for each site. As the shoreline displacement curves are interpolated to the sites based on their location along a south-west–north-east gradient, each site is effectively associated with a unique shoreline displacement curve, provided they are not located on exactly the same isobase. With analogy to the radiocarbon methodology, this would be equivalent to each date being associated with a unique calibration curve. As it would be computationally prohibitive to interpolate the shoreline displacement trajectory for each date to be simulated in the Monte Carlo procedure, one shoreline displacement curve was initially interpolated to the centre of each of a series of 2 km wide line segments running perpendicular to the shoreline gradient between the extremes of the distribution of sites. These intervals were then assigned a weight based on how the density of observed sites is distributed among them (Figure 4).

The Monte Carlo simulations with the SSPD is here based on drawing a sample of calendar dates from the observed date range in the SSPD, equalling the number of shoreline dated sites, where the probability of drawing any 5 year interval is determined by the null model of choice. This is equivalent to the sampling method *calsample* from *rcarbon* (Crema and Bevan, 2021). Each sampled date is then assigned one of the pre-interpolated displacement curves, the probability of which is weighted by the density of observed sites within each 2 km interval. The calendar date is then 'uncalibrated' – to follow the RSPD terminology – to an elevation range from which a single elevation value is then drawn with uniform probability between the lower and upper limit of the displacement curve at that calendar year, using intervals of 5 cm. As shoreline dating is here done with a gamma distribution with the same parameters across all sites, there is no equivalent of the error term for ^{14}C-dates that determine the shape of the Gaussian distribution associated with the

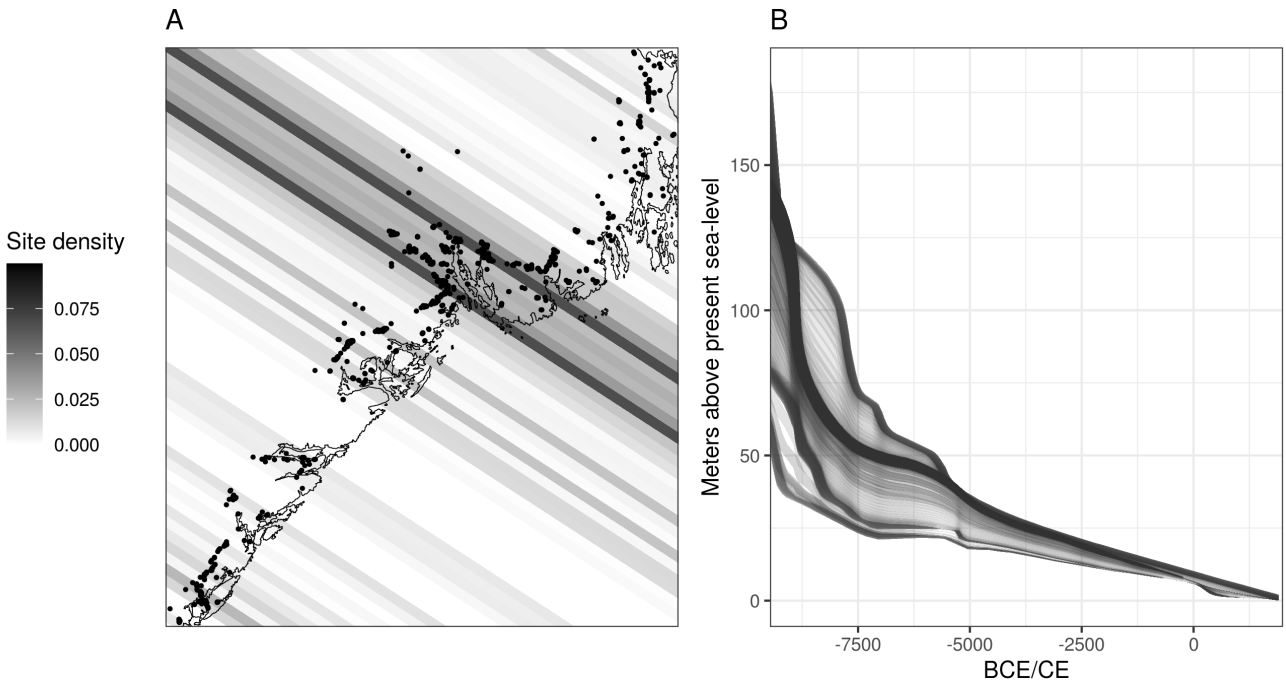


Figure 4: A) Density of included surveyed sites and excavated sites dated by means of shoreline dating (n = 921) as distributed across 2 km wide line segments that run perpendicular to the shoreline gradient. B) The displacement curves interpolated to the centre of each segment for use in the Monte Carlo simulations below.

^{14}C age. Consequently, the elevation value retrieved by 'uncalibrating' the shoreline date was then shoreline dated using the displacement curve for the relevant 2 km interval with the same gamma distribution for all samples. Having dated the number of samples equalling the number of shoreline dated sites, these were then summed, and the entire process repeated a total of 10000 times. The 97.5% highest and 2.5% lowest summed probability at each interval of 5 years across all simulations was then retrieved to create the 95% critical envelope, and the global p-value found in the same manner as for the RSPDs. This was done using re-purposed code from *ADMUR*.

3.3 Continuous piece-wise linear models

The procedures outlined above represent standard approaches when modelling population dynamics in archaeology. These can be helpful for determining whether the population behaviour has followed the trajectories of common theoretical population models, and is an especially useful exercise here as few assumptions concerning the behaviour and overall development of number of shore-bound sites over time could be made. However, standard models frequently used to evaluate RSPDs are consistently being rejected in the literature, and deviations from these rejected null-models do not in themselves allow for any interpretations of the deviations, nor a subsequent direct analysis of the SPDs – although this has frequently been the case. As a result, Timpson et al. (2021) have recently suggested an alternative framework for modelling past population dynamics. This involves fitting continuous piece-wise linear (CPL) models in addition to the standard models to the data, subjecting these to a procedure for model comparison, and finally testing the best model using the approach outlined above. The benefit of the CPL models is that these define discrete periods between hinge points that mark points in time in which the trajectory of relative population size changes direction, thus identifying key events and characteristics of the population dynamics between these. Following the evaluation of standard models outlined above, the data was therefore subsequently subjected to the approach outlined by Timpson et al. (2021).

3.4 Comparing model performance

For each dating method, the relative performance of the exponential, logistic, uniform and CPL models was compared. This was done by finding the overall relative log-likelihoods of the models using the *loglik* function from *ADMUR*. For any given model, this involves finding the likelihood for each shoreline date against the null-model, and in the case of the ^{14}C-dates, each site-phase. The overall model likelihood is the product of each of these individual likelihoods and is not, critically (Timpson et al., 2021, p. 3), found by estimating the likelihood using the final SPD. From these likelihoods, the Bayesian (or Schwarz) information criterion (BIC) was then found for each model. This penalises the models for the number of parameters in use, to avoid over-fitting, and can be used to compare the relative performance of the models.

4 Results

For the RSPD, the logistic model achieved a p-value of 0.108, and could therefore not be rejected. Based on the available data, this therefore represents the best explanation of the data. For the SSPD all of the standard models could be rejected.

Comparing the the directions of the three standard models, Figure @ref(fig:spd_mc) appears to indicate an overall uniform distribution in the frequency of radiocarbon dates following the initial jump just before 8000 BCE, contrasted to a general decrease in shoreline dated sites after the initial rapid increase from around 9500 BCE.

It is worth noting that there appears to be a corresponding peak in both proxies just after 8000 BCE and similarly a drop just before 6000 BCE. However, given that the logistic model cannot be rejected for the RSPD, it is inappropriate to interpret any deviations from the Monte Carlo envelope, and therefore no analytical weight can be placed on these deviations – at least given the present data and modelling efforts.

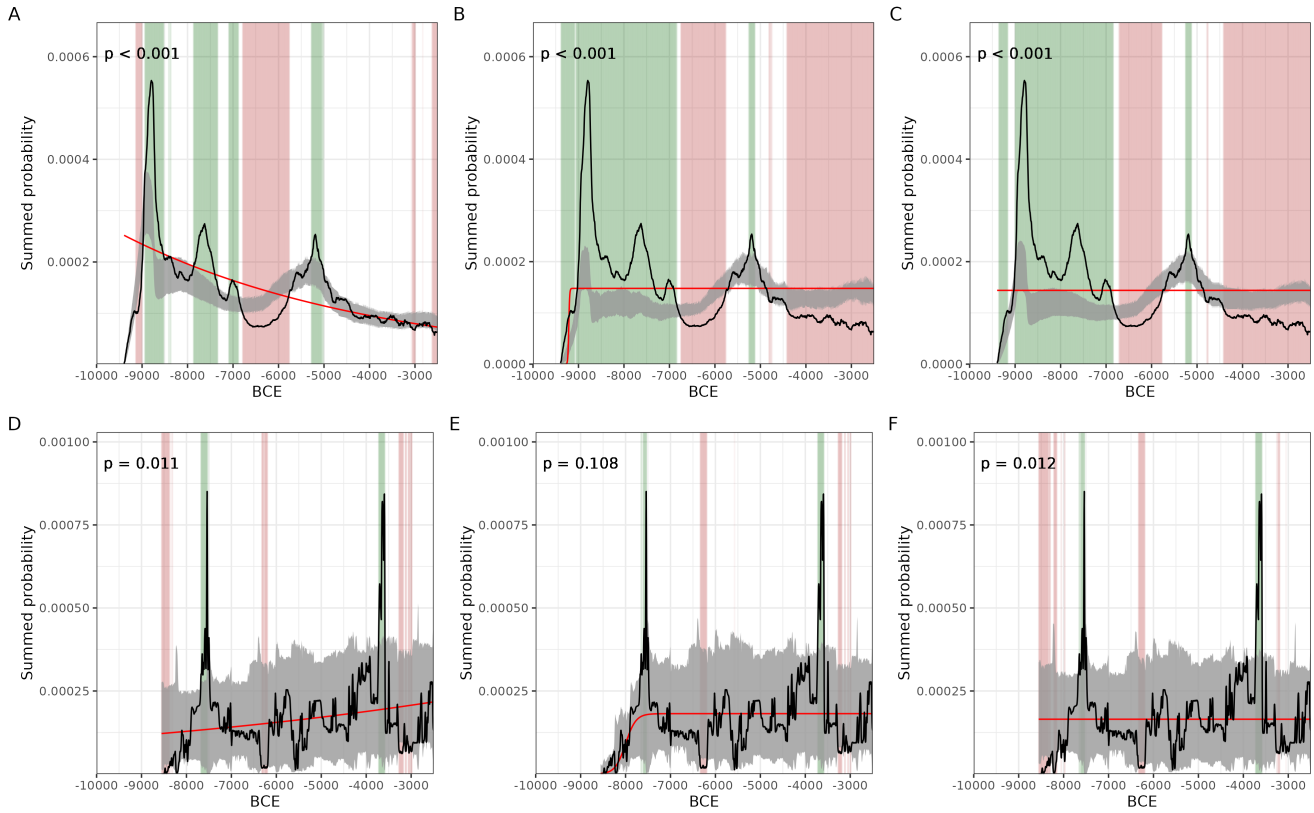


Figure 5: Monte Carlo simulation for shoreline and radiocarbon dates. A) Summed probability of shoreline dated sites (n = 921) compared to an exponential null model. B) SSPD compared to a logistic null model. C) SSPD compared to a uniform null model. D) Summed probability of calibrated radiocarbon dates (n = 310, bins = 134) compared to an exponential null model. E) RSPD compared to a logistic null model. F) RSPD compared to a uniform null model.

5 Discussion

It is firmly established that Stone Age sites in the region date to the period from before the earliest ^{14}C-dates, and at least from as far back as around 9300 BCE (Damlien and Solheim, 2018; e.g. Glørstad, 2016). While the SSPD starts from c. 9500 BCE, thus suggesting that the first date of human occupation in south-eastern Norway can be pushed back, this will require further corroboration from excavations and more meticulous evaluations of the evidence at hand. Furthermore, what has caused the lack of ^{14}C-dates from this earliest period has been given many possible explanations through the years, ranging from taphonomic loss to diverse cultural factors, such as the potential use of rapidly deteriorating seal blubber for burning, which could also be related to a lack of reliable access to firewood. While taphonomic and ecological factors cannot be written off, we view it as likely that the initial jump in ^{14}C-dates at the start of RSPD combined with the magnitude of the SSPD in the period preceding this, suggest that cultural factors is a central driver behind this pattern.

The considered models indicate that after the initial jumps, some process of decrease in shoreline dated site frequency occurs throughout the period, contrasted by a general stationary frequency of ^{14}C-dated material. The most immediate explanation of this mismatch would be that the SSPD largely reflects variations in land-use and mobility patterns, while the RSPD mainly reflects population numbers.

The development after 4000 BCE should be treated with further care, following from questions concerning the reliability of shoreline dating after 4000 BCE (cf. Roalkvam, 2023a).

The initial peak in the SSPD would be congruent with a process of rapid colonisation of the Norwegian coast, coupled with a high degree of mobility, leading to the overall high site-count in the earliest period. As evidenced by both genetic data and technological analyses of lithic inventories, this initial phase of human occupation appears to be followed by an influx of people from the . If this narrative is coupled to the data observed here, this process would appear to result in a relatively sudden drop in site-count as indicated in the SSPD, while the end of this drop corresponds to the first ^{14}C-dates in the RSPD – giving further credence to the suggestion that the appearance of material to be ^{14}C-dated is related to cultural processes.

Berg-Hansen et al. (2022) have argued that there is an increase in sites in the first half of the MM in the northern part of the study area of this paper. However, their analysis is seemingly also based on the aggregation of point-based estimates of shoreline dates, which are aggregated within periods of variable length. Apart from the limitations with using point-based estimates for shoreline dates, the use of disjoint time-intervals of variable duration could cause issues for comparing the relative counts across these units (Bevan and Crema, 2021). In contrast to their findings, the analysis undertaken here instead indicates a drop in the number of sites through this time-interval.

While population numbers would likely impact the SSPD as well, the overall mismatch between the two proxies used here means that the magnitude of demographic influence on the SSPD is not possible to ascertain.

While it appears that the SSPD and RSPD are differentially impacted by population numbers, and, we suggest, the SSPD is more heavily influenced by land-use and mobility patterns, this suggestion and the nature of the discrepancy between the proxies clearly requires further investigation. Some lines of inquiry could be to also draw on lithic technology and tool-kit diversity as a function of population size and density (e.g. Collard et al., 2013; Riede, 2014; Solheim et al., 2020), and mobility patterns that can potentially be inferred from the composition of lithic assemblages (e.g. Barton and Riel-Salvatore, 2014; Clark and Barton, 2017; Preston and Kador, 2018; Roalkvam, 2022). Furthermore, environmental conditions, constraints and developments also represent a range of factors that can have implications for these dimensions, and will consequently be of future importance to unpack their relationship (e.g. Jørgensen et al., 2020; Ordonez and Riede, 2022).

Finally, seeing as the method for shoreline dating implemented here was recently developed and has been subjected to limited evaluation and testing, there are still considerable uncertainties associated with its application. While we still believe the approach represents an improvement over previous approaches, the results should therefore be treated with care. This also extends to the results of the RSPD, due to the fairly limited ^{14}C sample. With these limitations in mind, the approach outlined by Timpson et al. (2021) gives a more reasonable framework with which to evaluate the available data, as leads to a more conservative

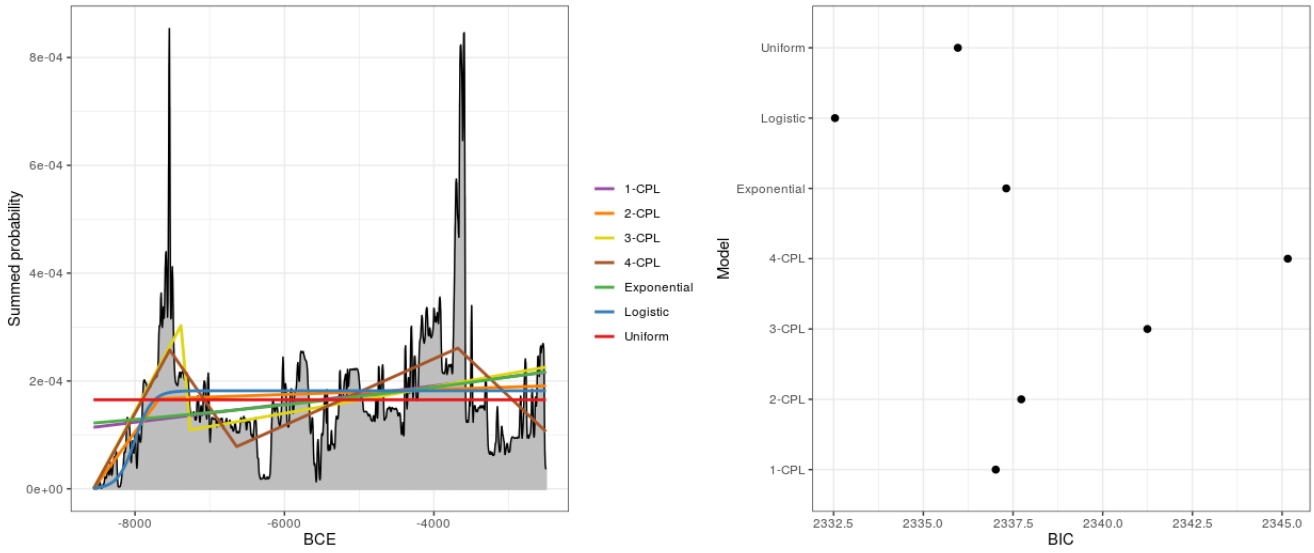
analysis and reduces the amount of over-interpretation that might follow from uncertainties to do with data or methods.

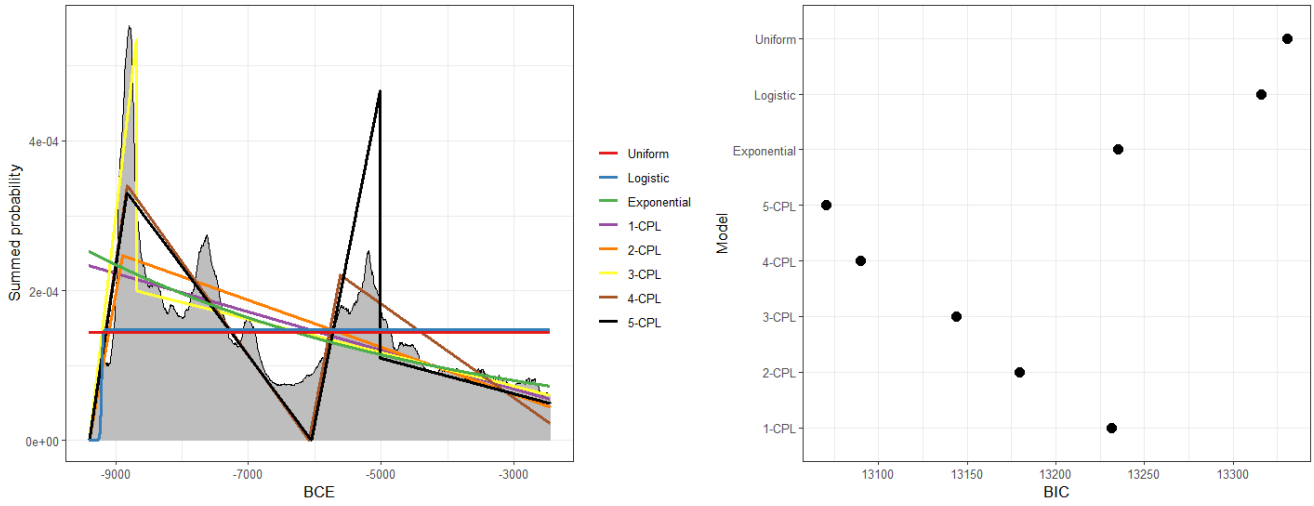
6 Conclusion

The finding that the frequency of radiocarbon dates and frequency of shoreline dated sites, in general, follow inverse trajectories is a valuable insight. Although a more precise understanding of this relationship

7 Ekstra figurer

Jeg fikk ADMUR til å funke med C14-dateringer. Figuren under viser at den logaritmiske modellen (blå kurve, samme som i Figur 5 over) ifølge BIC passer best av alternativene , og Monte Carlo-simuleringen over viser at den ikke kan avvises som forklarende for dataen. Eller med andre ord, dataen ville kunne forventes under den logaritmiske utviklingen. CPL modellering av strandlinjedateringene sliter jeg imidlertid mer med. Modellene opp til 4-CPL ser rimelige ut, men toppen på 5-CPL (sort kurve) som ligger litt før 5000 f.Kr. virker feil. I tillegg indikerer BIC-verdiene at denne passer best til dataen, som også virker rart gitt at den avviker såpass fra SSPDen.




Summed probability
4e-04
2e-04
0e+00
-9000
-7000
-5000
-3000
BCE
Uniform
Logistic
Exponential
1-CPL
2-CPL
3-CPL
4-CPL
5-CPL
Model
Uniform
Logistic
Exponential
5-CPL
4-CPL
3-CPL
2-CPL
1-CPL
13100
13150
13200
13250
13300
BIC

8 References

Barton, C.M., Riel-Salvatore, J., 2014. The Formation of Lithic Assemblages. Journal of Archaeological Science 46, 334–352. https://doi.org/10.1016/j.jas.2014.03.031

Berg-Hansen, I.M., Hårstad, S., Granados, T.J., Reitan, G., Romundset, A., Johannessen, L.S., Solheim, S., 2022. Enculturating Coastal Environments in the Middle Mesolithic (8300–6300 cal BCE) – Site Variability, Human–Environment Relations, and Mobility Patterns in Northern Vestfold, SE-Norway. Open Archaeology 8, 634–639. https://doi.org/10.1515/opar-2022-0251

Bergsvik, K.A., Darmark, K., Hjelle, K.L., Aksdal, J., Åstveit, L.I., 2021. Demographic developments in Stone Age coastal western Norway by proxy of radiocarbon dates, stray finds and palynological data. Quaternary Science Reviews 259, 106898. https://doi.org/10.1016/j.quascirev.2021.106898

Bevan, A., Crema, E.R., 2021. Modifiable reporting unit problems and time series of long-term human activity. Philosophical Transactions of the Royal Society B 376, 20190726. https://doi.org/10.1098/rstb.2019.0726

Blackwell, P.G., Buck, C.E., 2003. The Late Glacial human reoccupation of north-western Europe: new approaches to space-time modelling. Antiquity 77, 232–240. https://doi.org/10.1017/S0003598X00092231

Bluhm, L.E., Surovell, T.A., 2019. Validation of a global model of taphonomic bias using geologic radiocarbon ages. Quaternary Research 91, 325–328. https://doi.org/10.1017/qua.2018.78

Breivik, H.M., 2014. Palaeo-oceanographic development and human adaptive strategies in the Pleistocene–Holocene transition: A study from the Norwegian coast. The Holocene 24, 1478–1490. https://doi.org/10.1177/0959683614544061

Brest, J., Greiner, S., Boskovic, B., Mernik, M., Zumer, V., 2006. Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. IEEE Transactions on Evolutionary Computation 10, 646–657. https://doi.org/10.1109/TEVC.2006.872133

Carleton, W.C., Campbell, D., Collard, M., 2018. Radiocarbon dating uncertainty and the reliability of the PEWMA method of time-series analysis for research on long-term human-environment interaction. PloS One 13, e0191055. https://doi.org/10.1371/journal.pone.0191055

Carleton, W.C., Groucutt, H.S., 2021. Sum things are not what they seem: Problems with point-wise interpretations and quantitative analyses of proxies based on aggregated radiocarbon dates. The Holocene 31, 630–643. https://doi.org/10.1177/0959683620981700

Clark, G.A., Barton, C.M., 2017. Lithics, landscapes & la Longue-durée – Curation & expediency as expressions of forager mobility. Quaternary International 450, 137–149. https://doi.org/10.1016/j.quaint.2016.08.002

Collard, M., Ruttle, A., Buchanan, B., O'Brien, M.J., 2013. Population Size and Cultural Evolution in Nonindustrial Food-Producing Societies. PLOS ONE 8, e72628. https://doi.org/10.1371/journal.pone.0072628

Conolly, J., 2020. Spatial interpolation, in: Gillings, M., Hacıgüzeller, P., Lock, G. (Eds.), Archaeological Spatial Analysis: A Methodological Guide. Routledge, London & New York, pp. 118–134.

Crema, E.R., 2022. Statistical inference of prehistoric demography from frequency distributions of radiocarbon dates: a review and a guide for the perplexed. Journal of Archaeological Method and Theory. https://doi.org/10.1007/s10816-022-09559-5

Crema, E.R., Bevan, A., 2021. Inference from large sets of radiocarbon dates: Software and methods. Radiocarbon. https://doi.org/10.1017/RDC.2020.95

Damlien, H., Berg-Hansen, I.M., Melheim, L., Mjærum, A., Persson, P., Schülke, A., Solheim, S., 2021. Steinalderen i Sørøst-Norge: Faglig program for steinalderundersøkelser ved Kulturhistorisk museum. Cappelen Damm Akademisk, Oslo.

Damlien, H., Solheim, S., 2018. The Pioneer Settlement of Eastern Norway, in: Blankholm, H.P. (Ed.), Early Economy and Settlement in Northern Europe. Pioneering, Resource Use, Coping with Change. Equinox, Sheffield, pp. 335–367.

Fossum, G., 2020. Specialists facing climate change. The 8200 cal BP event and its impact on the coastal settlement in the inner Oslo fjord, southeast Norway, in: Schülke, A. (Ed.), Coastal Landscapes of the Mesolithic: Human Engagement with the Coast from the Atlantic to the Baltic Sea. Routledge, London & New York, pp. 179–201.

Freeman, J., Byers, D.A., Robinson, E., Kelly, R.L., 2018. Culture Process and the Interpretation of Radiocarbon Data. Radiocarbon 60, 453–467. https://doi.org/10.1017/RDC.2017.124

French, J.C., 2016. Demography and the Palaeolithic Archaeological Record. Journal of Archaeological Method and Theory 23, 150–199. https://doi.org/10.1007/s10816-014-9237-4

French, J.C., Riris, P., Fernandéz-López de Pablo, J., Lozano, S., Silva, F., 2021. A manifesto for palaeodemography in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences 376, 20190707. https://doi.org/10.1098/rstb.2019.0707

Glørstad, H., 2016. Deglaciation, sea-level change and the Holocene colonization of Norway. Geological Society, London, Special Publications 411, 9–25. https://doi.org/10.1144/SP411.7

Goldberg, A., Mychajliw, A.M., Hadly, E.A., 2016. Post-invasion demography of prehistoric humans in South America. Nature 532, 232–235. https://doi.org/10.1038/nature17176

Jørgensen, E.K., Pesonen, P., Tallavaara, M., 2020. Climatic changes cause synchronous population dynamics and adaptive strategies among coastal hunter-gatherers in Holocene northern Europe. Quaternary Research 1–16. https://doi.org/10.1017/qua.2019.86

Lundström, V., Robins, P., Riede, F., 2021. Demographic estimates from the Palaeolithic–Mesolithic boundary in Scandinavia: comparative benchmarks and novel insights. Philosophical Transactions of the Royal Society B: Biological Sciences 376, 20200037. https://doi.org/10.1098/rstb.2020.0037

Marwick, B., 2017. Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. Journal of Archaeological Method and Theory 24, 424–450. https://doi.org/10.1007/s10816-015-9272-9

Marwick, B., Boettiger, C., Mullen, L., 2018. Packaging Data Analytical Work Reproducibly Using R (and Friends). The American Statistician 72, 80–88. https://doi.org/10.1007/s10816-015-9272-9

Mjærum, A., 2022. A Matter of Scale: Responses to Landscape Changes in the Oslo Fjord, Norway, in the Mesolithic. Open Archaeology 8, 62–84. https://doi.org/10.1515/opar-2022-0225

Mörner, N.-A., 1979. The Fennoscandian Uplift and Late Cenozoic Geodynamics: Geological Evidence. GeoJournal 3, 287–318. https://doi.org/10.1007/BF00177634

Nielsen, S.V., 2021. A Late Mesolithic Forager Dispersal Caused Pre-Agricultural Demographic Transition in Norway. Oxford Journal of Archaeology 40, 153–175. https://doi.org/https://doi.org/10.1111/ojoa.12218

Nielsen, S.V., Persson, P., Solheim, S., 2019. De-Neolithisation in southern Norway inferred from statistical modelling of radiocarbon dates. Journal of Anthropological Archaeology 53, 82–91. https://doi.org/10.1016/j.jaa.2018.11.004

Norwegian Directorate for Cultural Heritage, 2018. Askeladden versjon 3.0 brukerveiledning.

Ordonez, A., Riede, F., 2022. Changes in limiting factors for forager population dynamics in Europe across the last glacial-interglacial transition. Nature Communications 13, 5140. https://doi.org/10.1038/s41467-022-32750-x

Palmisano, A., Bevan, A., Shennan, S., 2017. Comparing archaeological proxies for long-term population patterns: An example from central Italy. Journal of Archaeological Science 87, 59–72. https://doi.org/10.1016/j.jas.2017.10.001

Palmisano, A., Lawrence, D., de Gruchy, M.W., Bevan, A., Shennan, S., 2021. Holocene regional population dynamics and climatic trends in the Near East: A first comparison using archaeo-demographic proxies. Quaternary Science Reviews 252, 106739. https://doi.org/10.1016/j.quascirev.2020.106739

Persson, P., 2009. Mesolithic inland settlement in southern Norway, in: McCartan, S.B., Schulting, R., Warren, G., Woodman, P. (Eds.), Mesolithic Horizons: Papers presented at the Seventh International Conference on the Mesolithic in Europe, Belfast 2005. Volume 1. Oxbow Books, Oxford, pp. 243–247.

Preston, P.R., Kador, T., 2018. Approaches to Interpreting Mesolithic Mobility and Settlement in Britain and Ireland. Journal of World Prehistory 31, 321–345. https://doi.org/10.1007/s10963-018-9118-y

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Reimer, P.J., Austin, W.E.N., Bard, E., Bayliss, A., Blackwell, P.G., Ramsey, C.B., Butzin, M., Cheng, H., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Hajdas, I., Heaton, T.J., Hogg, A.G., Hughen, K.A., Kromer, B., Manning, S.W., Muscheler, R., Palmer, J.G., Pearson, C., Plicht, J. van der, Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Turney, C.S.M., Wacker, L., Adolphi, F., Büntgen, U., Capano, M., Fahrni, S.M., Fogtmann-Schulz, A., Friedrich, R., Köhler, P., Kudsk, S., Miyake, F., Olsen, J., Reinig, F., Sakamoto, M., Sookdeo, A., Talamo, S., 2020. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). Radiocarbon 62, 725–757. https://doi.org/10.1017/RDC.2020.41

Rick, J.W., 1987. Dates as Data: An Examination of the Peruvian Preceramic Radiocarbon Record. American Antiquity 52, 55–73. https://doi.org/10.2307/281060

Riede, F., 2014. Success and failure during the Lateglacial pioneer human re-colonisation of southern Scandinavia, in: Riede, F., Tallavaara, M. (Eds.), Lateglacial and Postglacial Pioneers in Northern Europe. Archaeopress, Oxford, pp. 33–52.

Roalkvam, I., 2023a. A simulation-based assessment of the relation between Stone Age sites and relative sea-level change along the Norwegian Skagerrak coast. Quaternary Science Reviews 299, 107880. https://doi.org/10.1016/j.quascirev.2022.107880

Roalkvam, I., 2023b. shoredate: An R package for shoreline dating coastal Stone Age sites. Journal of Open Source Software 8, 5337. https://doi.org/10.21105/joss.05337

Roalkvam, I., 2022. Exploring the composition of lithic assemblages in Mesolithic south-eastern Norway. Journal of Archaeological Science: Reports 42, 103371. https://doi.org/10.1016/j.jasrep.2022.103371

Roalkvam, I., 2020. Algorithmic Classification and Statistical Modelling of Coastal Settlement Patterns in Mesolithic South-Eastern Norway. Journal of Computer Applications in Archaeology 3, 288–307. https://doi.org/10.5334/jcaa.60

Romundset, A., 2021. Resultater fra NGUs undersøkelse av etteristidas strandforskyvning nord i Vestfold. Geological Survey of Norway, Trondheim.

Romundset, A., 2018. Postglacial shoreline displacement in the Tvedestrand-Arendal area, in: Reitan, G., Sundström, L. (Eds.), The Stone Age Coastal Settlement in Aust-Agder, Southeast Norway. Cappelen Damm Akademisk, Oslo, pp. 463–478.

Romundset, A., Lakeman, T.R., Høgaas, F., 2018. Quantifying variable rates of postglacial relative sea level fall from a cluster of 24 isolation basins in southern Norway. Quaternary Science Reviews 197, 175–192. https://doi.org/10.1016/j.quascirev.2018.07.041

Shennan, S., 2000. Population, Culture History, and the Dynamics of Culture Change. Current Anthropology 41, 811–835. https://doi.org/10.1086/317403

Shennan, S., Downey, S.S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., Thomas, M.G., 2013. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. Nature Communications 4, 2486. https://doi.org/10.1038/ncomms3486

Solheim, S., 2020. Mesolithic coastal landscapes. Demography, settlement patterns and subsistence economy in southeastern Norway, in: Schülke, A. (Ed.), Coastal Landscapes of the Mesolithic: Human Engagement with the Coast from the Atlantic to the Baltic Sea. Routledge, London & New York, pp. 44–72.

Solheim, S., Damlien, H., Fossum, G., 2020. Technological transitions and human-environment interactions in Mesolithic southeastern Norway, 11 500–6000 cal. BP. Quaternary Science Reviews 246, 106–501. https://doi.org/10.1016/j.quascirev.2020.106501

Solheim, S., Persson, P., 2018. Early and mid-holocene coastal settlement and demography in southeastern norway: Comparing distribution of radiocarbon dates and shoreline-dated sites, 8500–2000 cal. BCE. Journal of Archaeological Science: Reports 19, 334–343. https://doi.org/10.1016/j.jasrep.2018.03.007

Sørensen, R., Henningsmoen, K.E., Høeg, H.I., Gälman, V., in press. Holocen vegetasjonshistorie og landhevning i søndre Vestfold og sørøstre Telemark, in: Persson, P., Solheim, S. (Eds.), The Stone Age in Telemark. Archaeological results and scientific analysis from Vestfoldbaneprosjektet and E18 Rugtvedt-Dørdal.

Sørensen, R., Henningsmoen, K.E., Høeg, H.I., Gälman, V., 2014a. Holocene landhevningsstudier i søndre vestfold og sørøstre telemark – revidert kurve, in: Melvold, S., Persson, P. (Eds.),. Portal, Kristiansand, pp. 36–47.

Sørensen, R., Høeg, H.I., Henningsmoen, K.E., Skog, G., Labowsky, S.F., Stabell, B., 2014b. Utviklingen av det senglasiale og tidlig preboreale landskapet og vegetasjonen omkring steinalderboplassene ved Pauler, in: Jaksland, L., Persson, P. (Eds.), E18 Brunlanesprosjektet. Bind I. Forutsetninger og kulturhistorisk sammenstilling. University of Oslo, Museum of Cultural History, Oslo, pp. 171–213.

Surovell, T.A., Byrd Finley, J., Smith, G.M., Brantingham, P.J., Kelly, R.L., 2009. Correcting temporal frequency distributions for taphonomic bias. Journal of Archaeological Science 36, 1715–1724. https://doi.org/10.1016/j.jas.2009.03.029

Svendsen, J.I., Mangerud, J., 1987. Late Weichselian and Holocene sea-level history for a cross-section of western Norway. Journal of Quaternary Science 2, 113–132. https://doi.org/10.1002/jqs.3390020205

Tallavaara, M., Pesonen, P., 2020. Human ecodynamics in the north-west coast of Finland 10,000–2000 years

ago. Quaternary International 549, 26–35. https://doi.org/10.1016/j.quaint.2018.06.032
Timpson, A., Barberena, R., Thomas, M.G., Méndez, C., Manning, K., 2021. Directly modelling population dynamics in the South American Arid Diagonal using 14C dates. Philosophical Transactions of the Royal Society B: Biological Sciences 376, 20190723. https://doi.org/10.1098/rstb.2019.0723
Timpson, A., Colledge, S. and, Crema, E., Edinborough, K., Kerig, T., Manning, K., Thomas, S., Mark G. Shennan, 2013. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. Nature Communications 4, 2486. https://doi.org/10.1038/ncomms3486
Torfing, T., 2015. Neolithic population and summed probability distribution of 14C-dates. Journal of Archaeological Science 63, 193–198. https://doi.org/10.1016/j.jas.2015.06.004
Ward, I., Larcombe, P., 2021. Sedimentary unknowns constrain the current use of frequency analysis of radiocarbon data sets in forming regional models of demographic change. Geoarchaeology 36, 546–570. https://doi.org/10.1002/gea.21837
Weninger, B., Clare, L., Jöris, O., Jung, R., Edinborough, K., 2015. Quantum theory of radiocarbon calibration. World Archaeology 47, 543–566. https://doi.org/10.1080/00438243.2015.1064022