

Directly modelling population dynamics in the South American Arid Diagonal using ^{14}C dates

Adrian Timpson^{1,2}, Ramiro Barberena³, Mark G. Thomas¹, César Méndez⁴, Katie Manning⁵

1. UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, Gower St, London WC1E 6BT, United Kingdom
2. Max Planck Institute for the Science of Human History, Kahlaische Strasse 10, 07745 Jena, Germany
3. Instituto Interdisciplinario de Ciencias Básicas (ICB), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Laboratorio de Paleoecología Humana, Facultad de Ciencias Exactas y Naturales, Facultad de Filosofía y Letras. Universidad Nacional de Cuyo. Padre Jorge Contreras 1300, Mendoza, Argentina
4. Centro de Investigación en Ecosistemas de la Patagonia, Moraleda 16, Coyhaique, Aisén, Chile
5. Department of Geography, King's College London, Strand, London WC2R 2LS

Abstract

Large anthropogenic ^{14}C datasets are widely used to generate summed probability distributions (SPD) as a proxy for past human population levels. However, SPDs are a poor proxy when datasets are small, bearing little relationship to true population dynamics. Instead, more robust inferences can be achieved by directly modelling the population and assessing the model likelihood given the data. We introduce the R package ADMUR which uses a continuous piecewise linear (CPL) model of population change, calculates the model likelihood given a ^{14}C dataset, estimates credible intervals using MCMC, applies a Goodness of Fit test, and uses the Schwarz Criterion to compare CPL models. We demonstrate the efficacy of this method using toy data, showing that spurious dynamics are avoided when sample sizes are small, and true population dynamics are recovered as sample sizes increase. Finally, we use an improved ^{14}C dataset for the South American Arid Diagonal to compare CPL modelling to current simulation methods, and identify three Holocene phases when population trajectory estimates changed from rapid initial growth of 4.15% per generation to a decline of 0.05% per generation between 10,821 and 7,055 yr BP, then gently grew at 0.58% per generation until 2,500 yr BP.

Keywords 3 to 6

Continuous piecewise linear model; radiocarbon; summed probability distribution; Holocene population dynamics; South American Arid Diagonal; ADMUR

1. Introduction

The varying frequencies of archaeological samples through time are commonly represented using a Summed Probability Distribution (SPD) of associated calibrated ^{14}C dates, and such distributions of anthropogenic dates are widely used as a proxy to infer dynamics in human populations [1-8]. Prior to more recent simulation approaches (see below), the SPD curve was (and often continues to be) misinterpreted as a faithful representation of population dynamics despite its shape being influenced by other nuisance factors such as taphonomic loss, wiggles inherited from the calibration curve, and ascertainment biases. Curves from very small datasets are dominated by the sporadic nature of small sample sizes, and the inevitable gaps between individual calibrated dates cannot be interpreted as population hiatuses. [9-11]. In this paper we utilize an available extended radiocarbon database [12] for the South American Arid Diagonal (SAAD) and develop an improved method to extract demographic signatures from archaeological data by combining continuous piece wise linear models and formal model comparison.

(a) The archaeological hypothesis: Population troughs in the South American Arid Diagonal

The SAAD is a major climatic and biogeographic NW-SE band extending from northwestern Peru to southeastern Argentina, encompassing most of the arid and semiarid ecosystems of South America [13, 14]. Whilst spatially and temporally variable, the mid-Holocene (8.2-4.2 kyr BP [15]) has been characterized as a period of enhanced aridity in these deserts [16, 17]. Volcanic eruptions have also been suggested to have affected human occupations at local and regional scales in parts of the SAAD [18]. Based on this palaeoecological framework, there has been a growing archaeological debate on the possible existence of gaps in the archaeological record in different regions which may signal a demographic discontinuity during the mid-Holocene [2, 19-24]. Previous analyses by some authors of this paper (RB and CM) used SPDs of radiocarbon dates to explore population dynamics and found evidence of synchronicity in fluctuations at different latitudes, allowing the formation of a hypothesis that there may have been two short population troughs in the SAAD driven by increasing aridification between 7.6-7.2 kyr BP and 6.8-6.4 kyr BP [2, 24]. Here we build on this previous work and apply a rigorous model comparison framework to test this hypothesis.

2. Current inferential methods

(a) Direct SPD interpretation

As a thought experiment we can consider a curve comprising just a single (calibrated) date of an organic sample. The sample has a single (point) true date of death, and the curve tells us how believable each possible date is. Neither the sample's existence nor the true date of its death waxes and wanes through time. Likewise, we cannot interpret the SPD of a small dataset across a narrow time period as representing the fluctuations of a population through time – instead it represents how believable each year is, as possible point estimates for sample 1 *or* sample 2 *or* sample 3, etc. It is this '*or*' component (the summing) that restricts the interpretation of the curve – the SPD is not *the* single best explanation of the data, nor even *a* single 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.

We deliberately used the word *explanation*, since the SPD is merely a convolution of two datasets: the raw $^{14}\text{C}/^{12}\text{C}$ ratios with their errors; and the calibration curve with its error ribbon. Therefore, the SPD provides an excellent graphical representation of the data by compressing a large amount of information into a single plot, and its value in data representation should not be disparaged. However, the SPD is not a model and cannot be directly interpreted to draw reliable inferences about the population dynamics.

(b) Simulation methods to reject a null model

Recognising the need for a more robust inferential framework, by 2013 methods were developed that moved away from mere data representation, and instead focused on directly modeling the population. An exponential (or any other hypothesised shape) null model could be proposed, and many thousands of simulated datasets could then be generated under this model and compared to the observed. The SPD was no longer the end product; instead it was used to generate a summary statistic. The summary statistics from each simulated SPD (and the observed SPD) could then be compared, a p-value calculated, and (if deemed significant) the hypothesised model could be rejected [25, 26]. This approach was successful in directly testing a single hypothesised population history and was widely adopted [12, 27-33] as the field moved towards a model-based inferential framework.

(c) Other approaches to directly modelling the population

The inferential limits of the SPD and the importance of directly modelling population fluctuations has been approached with various underlying model structures. The Oxcal program offers Kernel Density Models [34], whilst the R package Bchron[35] employs Bayesian Gaussian mixture models. Both approaches can provide models of the underlying population by performing parameter searches and are based on sound model likelihood approaches. However, Gaussian based models (both mixture models and kernels) are by nature complex curves with constantly changing gradients. No doubt real population levels also fluctuate through time with complex and relentless change, but this leaves us with a model that can only be described graphically and cannot be easily summarised in terms of dating key demographic events.

Furthermore, these methods do not address how reasonable the model structure is in the first place. There are two approaches to achieve this. Firstly, a goodness of fit test can establish if the observed data could have been reasonably produced by the model. This is essentially the approach taken by the simulation methods mentioned above where the p-value provides this goodness of fit, and allows the model to be rejected if it is a poor explanation of the data. Secondly, a model selection process can be used to ensure unjustifiably complex models are rejected in favour of the simplest plausible model with the greatest explanatory power.

Goldberg *et al.*[36] and Fernández López de Pablo *et al.*[32] also modelled population dynamics directly. They both used a piecewise model comprising various phases of logistic and/or exponential growth. However, neither study used a continuous model (the phases did not join), nor were the authors able to calculate likelihoods. As a result, Goldberg *et al.* misappropriate the Schwarz criterion (BIC) for use with their 'proxy likelihoods' and contradict their own modelling results that indicate a stable population size during the mid-Holocene in favour of overinterpreting existing SPD simulation methods to infer oscillations of peaks and troughs. In the case of Fernández López de Pablo *et al.* [32] and others [30], the modelling is even more problematic since they apply regression directly to the SPD. The graphical points on the SPD however, correspond to the number of calendar years in the study period and are not the independent samples that formed it, i.e. the ^{14}C dates (or more conservatively, the smaller number of phases can be considered independent samples). This renders their standard regression outputs (likelihoods, AIC and BIC) meaningless.

Nevertheless, Goldberg *et al.* innovated an important contribution in two key respects. Firstly, their piecewise model is defined by a small number of discrete phases or periods. This brings the advantage of directly modelling the timing and intensity of population *events* (the date at which the model changed from one phase to the other), and a simple description of the population behaviour in each phase. Secondly and most importantly, the authors raised the point that a model comparison is required. They test various models, both simpler (one phase), and more complex (up to six phases) in various permutations of logistic and exponential phases. We build on this approach and overcome their shortcomings. We construct a continuous piecewise model, calculate likelihoods, and use the BIC to select the most appropriate number of phases. Finally, we use a goodness of fit test to show the data is plausible under the best model.

3. Continuous Piecewise Linear (CPL) Modelling

The goal in population modelling is usually to identify specific demographic events. Typically, the objective is to estimate the date of some event that marks a change in the trajectory of the population levels, such as the start of a rapid decline or increase in population levels (perhaps from disease, migration or changes in carrying capacity) and provide a simple description of the population behaviour between these events, such as a growth rate. A CPL model lends itself well to these

objectives since its parameters are the coordinates of the hinge points, which are the relative population size (y) and timing (x) of these events.

4. Model selection using the Schwarz Criterion

We choose the number of linear phases (or number of hinge points joining these phases) systematically as part of a Model Selection process. Given a ^{14}C dataset, we find the maximum likelihood (ML) continuous one-piece (or one-phase) linear model (1-CPL), then the ML 2-CPL, etc. Although the likelihood increases with the number of parameters (the greater freedom allows the model to fit more closely to the data), we calculate the Schwarz Criterion [37], otherwise commonly misnamed [38] the Bayesian Information Criterion (BIC), to naturally penalise for this increasing complexity. We favour this criterion over Akaike's Information Criterion (AIC) [39] since the BIC provides a greater penalty for model complexity than does the AIC, ensuring conservative selection that avoids an overfit model. Indeed, we find the AIC typically favours an unjustifiably complex model, for example when using toy data where the 'true model' is known. Therefore, we select the model with the lowest BIC as the best model. Model complexity beyond this provides incrementally worse BIC values, and as a result the turning point in model complexity can be easily found, and superfluous computation for unnecessarily complex CPL models is thus avoided.

Whilst a large database provides greater information content to justify a CPL model with many hinge points, it is worth considering the extreme case of fitting a CPL model to a tiny dataset. Fig 2 illustrates that the lack of information content naturally guards against overfitting, and a uniform distribution is always selected (a model with no demographic events and no population fluctuations) where sample sizes are low. This should make intuitive sense – in the light of such sparse we should not infer anything more complex than a constant population across the time range covered by that single date.

Large ^{14}C databases covering long time periods often exhibit a general long-term background increase through time, attributable to some combination of long-term population growth and some unknown rate of taphonomic loss of dateable material through time. Such a dataset may be better explained by a model of exponential growth (requiring just a single lambda parameter) than a CPL model. Therefore, for real datasets the model selection procedure should also consider other non-CPL models such as an exponential model.

5. Calculating Likelihoods

Theoretically a calibrated date should be a continuous probability density function (PDF), however in practice a date is represented as a discrete vector of probabilities corresponding to each calendar year, and is therefore a probability mass function (PMF). This discretisation (of both a proposed model probability distribution and a calibrated date probability distribution) provides the advantage that numerical methods can be used to calculate likelihoods.

Hypothetically, if a calibrated date was available with such precision that it could be attributed with certainty to just a single calendar year, the model likelihood would trivially be the model's probability mass (pm) at that date. Similarly, if the data comprised just two such point estimates (at calendar time points A and B), the model's relative likelihood would trivially be the model probability at date A multiplied by the model probability at date B.

However, a single calibrated ^{14}C date is not a point estimate, but rather a complex multimodal probability distribution, representing the probability of each possible year being the true date. Therefore the probability of a single calibrated date given the model can be calculated as the model probability at year A, *or* the model probability at year B etc, for all possible years, weighted by how

probable the calibrated ^{14}C date is at each of those years. This can be calculated using the scalar product between model probabilities and calibrated date probabilities, and gives the probability of a single calibrated date under the model. This is repeated for every calibrated date, and the overall product gives the relative likelihood of the model, given the whole dataset.

This approach assumes each date is a fair and random sample, but where many dates are available from a single site/phase it is sensible to first bin dates into phases. This is an important step in modelling population dynamics to adjust for the data ascertainment bias of some archaeological finds having more dates by virtue of a larger research interest/budget. This is achieved by first generating an SPD for each phase and normalising. These phase-SPDs are then combined and normalised to create a final SPD. This procedure ensures phases with multiple dates are weighted to contribute the same overall probability mass as a phase with a single date. The probability of each phase SPD can then be calculated in exactly the same way as the probability of a single calibrated date.

6. Avoiding edge effects

It is common for a research question to be targeted at a specific time range that spans only part of the overall calibrated date range of the ^{14}C dataset being used. This is of no consequence if merely generating an SPD, as regions outside the range of interest can be ignored or truncated. Indeed, simulation approaches benefit from considering a slightly wider range by pushing any potential edge effects outside the target range. In contrast, any modelling approach that calculates likelihoods will be influenced by the entire dataset provided, including dates that fall well outside the modelled date range. These external dates must be excluded, since they can have a substantial and mischievous influence on the parameter search.

This influence can be attributed to the interesting behaviour of the tails of a Gaussian distribution, from which a calibrated date is derived. A calibrated date has a non-zero probability at all calendar dates, and as a consequence, a mostly external date still has a tiny tail within the model's date boundaries. However, despite the *absolute* probability values of this tail being extremely small, surprisingly the *relative* value increases hugely towards the model boundary (approximately exponentially). As a result, given a dataset where all/most dates are external to the date range of interest, the most likely model shape will have massive upticks at the boundaries. Overall the likelihood of such a model will be extremely small, but it will be the best explanation given so much data is outside the date range.

Similarly, modelling population dynamics across a range wider than the available dates would be making the incorrect assumption that the absence of evidence at these edges provides evidence of absence, and this will influence the shape of the fitted model. There may be rare occasions where this is reasonable, for example where humans are known to be absent from an island prior to (or after) some date, or where the archaeological sampling is so substantial that there is confidence that dates would have been recovered if there was a human presence.

Therefore, it is important to ensure the date range of the data and model are appropriate for each other, and to exclude dates from the dataset that do not reasonably fall within the modelled range. We achieve this with our real datasets by only including a date if more than 50% of its probability falls within the modelled date range – i.e., it is more probable that its true date is internal than external. Similarly, we achieve this with our extremely small toy dataset ($N=6$) by constraining the modelled date range to exclude the negligible tails outside the calibrated dates.

7. Search algorithm for parameters

The CPL model is a probability mass function (PMF) such that the probability outside the date range equals zero, and the total probability within the date range equals one. The exact shape of this PMF is defined by the (x, y) coordinates of the hinge points. Therefore, there are various constraints on parameters required to define such a curve. For example, if we consider a 2-CPL model only the middle hinge has a free x-coordinate parameter, since the start and end date are already specified by the date range. Of the three y-coordinates (left, middle, right hinges), only two are free parameters, since the total probability must equal 1. Therefore a 2-CPL model has 3 free parameters (one x-coordinate and two y-coordinates) and an n-phase CPL model has 2n-1 free parameters.

We perform the search for the maximum likelihood (ML) parameters (given a ¹⁴C dataset and calibration curve) using the differential evolution optimization algorithm DEoptimR [40]. A naïve approach to this search would propose a set of values for all parameters in an iteration simultaneously, and reject the set if it does not satisfy the above constraints. However, this approach would result in the rejection of many parameter sets. Instead, our objective function considers the parameters in order, such that the next parameter is searched for in a reduced parameter space, conditional on the previous parameters. We achieve this by adapting the ‘stick breaking’ Dirichlet Process to apply in two dimensions by sampling stick breaks on the x-axis using the Beta distribution and y-coordinates using the Gamma distribution. At each hinge, the length of the stick is constrained by calculating the total area so far between the first and previous hinge.

8. Estimating Credible Intervals using Markov Chain Monte Carlo (MCMC)

Having constructed a likelihood function that calculates the relative likelihood of any parameter combination, it can be used as the objective function in a parameter search to find the Maximum Likelihood parameter estimates. However, we also use the likelihood function in a MCMC framework to estimate credible intervals of our parameter estimates. We achieve this using the Metropolis-Hastings algorithm [41] using a single chain of 100,000 iterations, discarding the first 2000 for burn-in, and thinning to every 5th iteration. The resulting joint posterior distribution can then be graphically represented in several ways, such as directly plotting the joint parameter estimates on a 2D plot (Fig 7), or histograms of the marginal distributions (Fig 6).

9. Goodness of fit (GOF) test

Once the best CPL model has been selected, its parameters found and the likelihood calculated, we generate 1000 simulated ¹⁴C datasets under this CPL model by ‘uncalibrating’ calendar dates randomly sampled under the model, taking care to ensure sample sizes exactly match the number of phases in the observed dataset. We then calculate the proportion of each calibrated simulated dataset outside the 95% CI, giving a distribution of summary statistics under our best CPL model. The p-value is then calculated as the proportion of these simulated summary statistics that are smaller or equal to the observed summary statistic. Conceptually this is similar to the method of calculating p-values under existing simulation methods for testing a null model [12, 25-33].

10. Demonstration of methods with toy data

(a) Testing CPL model for a typical sample size

To demonstrate our approach, we first generate a true (toy) population curve, which comprises a 3-CPL model PDF between 5.5kyr and 7.5kyr BP. We then randomly sample N=1500 dates under this true (toy) population curve, ‘uncalibrate’ these dates, apply an arbitrary ¹⁴C error of 25 years, then calibrate. We then conduct a parameter search for the best fitting 1-CPL, 2-CPL, 3-CPL, 4-CPL and 5-CPL models. The BIC is calculated using: $\ln(n)k - 2 \ln(L)$, where k is the number of parameters ($k = 2p - 1$, where p is the number of phases), n is the number of ¹⁴C dates and L is the maximum likelihood

[37]. Table 1 gives the results of this model comparison and shows that the model fits closer to the data as its complexity increases. However, the BIC shows that the model is overfitted beyond a 3-CPL model. Therefore, the model selection process successfully recovered the 3-CPL model from which the data was generated.

<i>Model</i>	<i>Parameters</i>	<i>Maximum log likelihood</i>	<i>BIC</i>
uniform	0	-9976.29	19952.58
1-CPL	1	-9862.86	19732.9
2-CPL	3	-9833.87	19689.26
3-CPL	5	-9792.05	19619.97
4-CPL	7	-9790.87	19631.97
5-CPL	9	-9790.40	19645.38

Table 1: The 3-CPL model is selected as the best, since it has the lowest BIC (bold). As the number of parameters in the model increases, the likelihood of the model given the data increases. However, the BIC shows that this improvement is only justified up to the 3-CPL model, after which the more complex models are overfit to the data.

We then assess the accuracy of the parameter estimates by generating 5 more random datasets under our true (toy) population curve and apply a parameter search to each dataset. Fig 1 illustrates the best 3-CPL model for each dataset, which are all qualitatively similar to the true population curve. Each is the most likely model given the differences between their respective datasets, which are represented with SPDs.

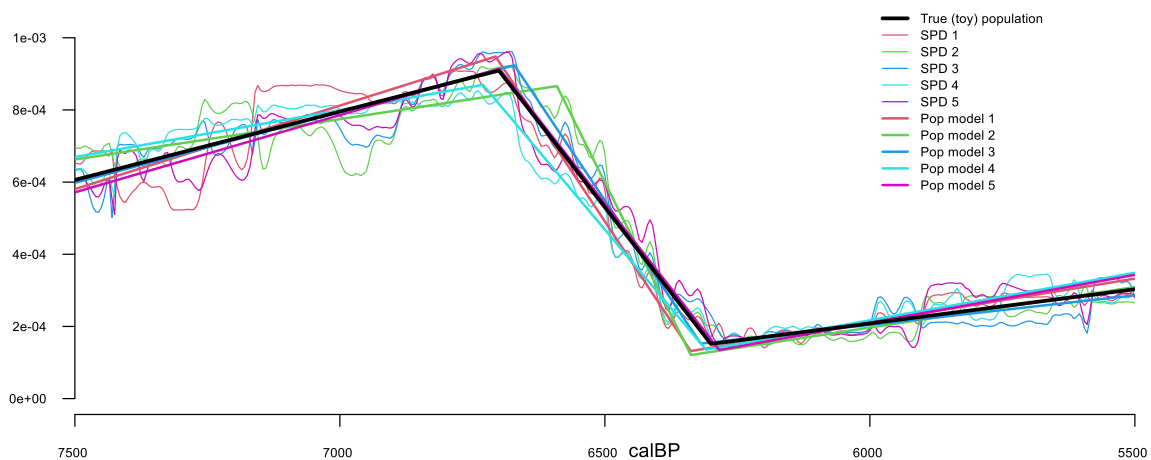


Fig 1: 3-CPL models best fitted to 5 randomly sampled datasets of $N=1500$ ^{14}C dates. Summed Probability Distributions (SPD) of each calibrated dataset illustrate the variation from generating random samples. This variation between random datasets is the underlying cause of the small differences between the hinge-point dates in each maximum likelihood model.

(b) Testing CPL model with small sample size.

We continue with the same true (toy) population curve and test the behaviour of both the model selection and parameter estimation with smaller sample sizes. As before, N dates are randomly sampled under the population curve, ‘uncalibrated’, assigned an error and calibrated. Fig 2 shows that for $N=329$ and $N=454$ the 3-CPL model is successfully selected, and its shape is similar to the true population. For $N=154$ the lack of information content favours a 1-CPL model which successfully avoids overfitting, and for $N=47$ and smaller, the even simpler uniform model is selected. For $N=6$ the modelled date range is reduced to only encompass the range of the data (see section ‘avoiding edge effects’). These results successfully demonstrate that this approach provides robust inferences of the

underlying population dynamics, avoids the misinterpretation inherent in small datasets, and approaches the true population dynamics as sample sizes increase.

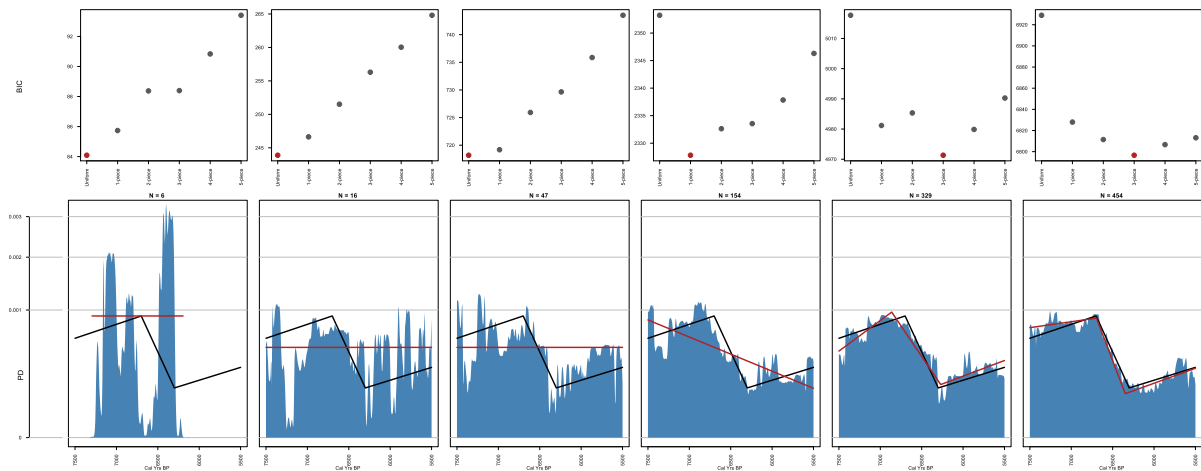


Fig 2: Model selection naturally guards against overfitting with small sample sizes since the lack of information content favours simple models. In contrast, the Summed Probability Distributions (SPD) suggest interesting population dynamics that in fact are merely the artefacts of small sample sizes and calibration wiggles.

Upper: the best model (red) selected using BIC between a uniform distribution and five increasingly complex n-CPL models.

Lower: SPD (blue) generated from calibrated ^{14}C dates randomly sampled from the same true (toy) population curve (black), and best CPL model PDF (red) constructed from Maximum Likelihood parameters. Note, the slight bend in black and red lines are merely a consequence of the non-linear y-axis used.

11. R package ADMUR

To enable full transparency of our methods and aid other researchers in applying and further developing these methods, we provide an accompanying package in R[42] called ADMUR: Ancient Demography Modelling Using Radiocarbon <https://CRAN.R-project.org/package=ADMUR>. and refer users to the vignette ‘guide’ which provides details of installation and use. All analysis and plots in this paper can be exactly replicated using the vignette ‘replicating-timpson-rstb.2020’.

12. Modelling population dynamics in the South American Arid Diagonal (SAAD)

(a) Data overview

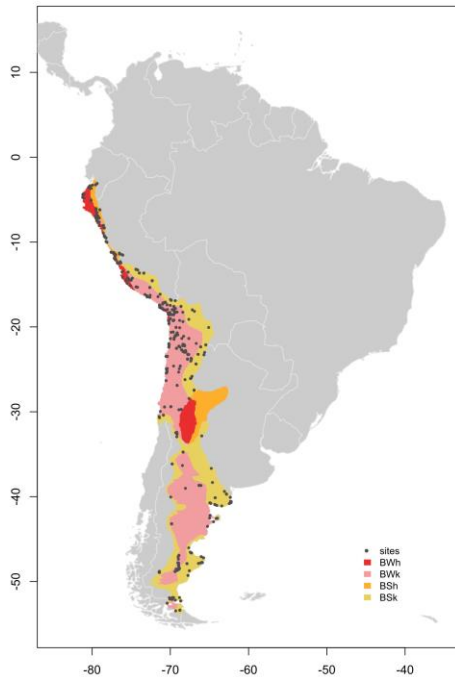


Fig 3: Site location of radiocarbon dates in South American Arid Diagonal (SAAD). All sites are within the contiguous SAAD, the boundary of which is defined using four of the Köppen-Geiger climate classification zones using Table 1 from Peel *et al* [43]: BWh = Arid hot desert, mean annual precipitation (MAP) < 5x $P_{\text{threshold}}$, mean annual temperature (MAT) $\geq 18^{\circ}\text{C}$; BWk = Arid cold desert, MAP < 5x $P_{\text{threshold}}$, MAT < 18°C ; BSh = Arid steppe hot, MAP $\geq 5x P_{\text{threshold}}$, MAT $\geq 18^{\circ}\text{C}$; BSk = Arid steppe cold MAP $\geq 5x P_{\text{threshold}}$, MAT < 18°C .

For this analysis we used a subset of the radiocarbon database compiled by Riris and Arroyo-Kalin [12] incorporating sites that fall within the geographically contiguous ‘Arid’ climatic categories of the SAAD as described in the World Köppen climate classification [44]. This provides a more relevant dataset to directly test previous hypotheses of demographic fluctuations in arid ecosystems [2, 20-24] by not averaging widely diverse ecological settings (e.g. [12, 36]).

(b) SPD simulation method testing exponential model

We generate an SPD from the dataset and test for significance using methods described in Shennan *et al* [25]. Individual dates from a single site within 200 years of each other were first binned into site-phases, calibrated using SHCal20[45], and summed and normalized to unity, to account for site-specific ascertainment bias. These distributions in each of the 708 phases were then summed and normalized to unity. Fluctuations in this observed SPD were tested for significance by generating 20,000 random datasets, each sampling 708 dates from a fitted exponential distribution. Dates were ‘uncalibrated’, assigned a random error, and calibrated. P-values were calculated using a summary statistic of each SPD calculated as the proportion the 95% CI.

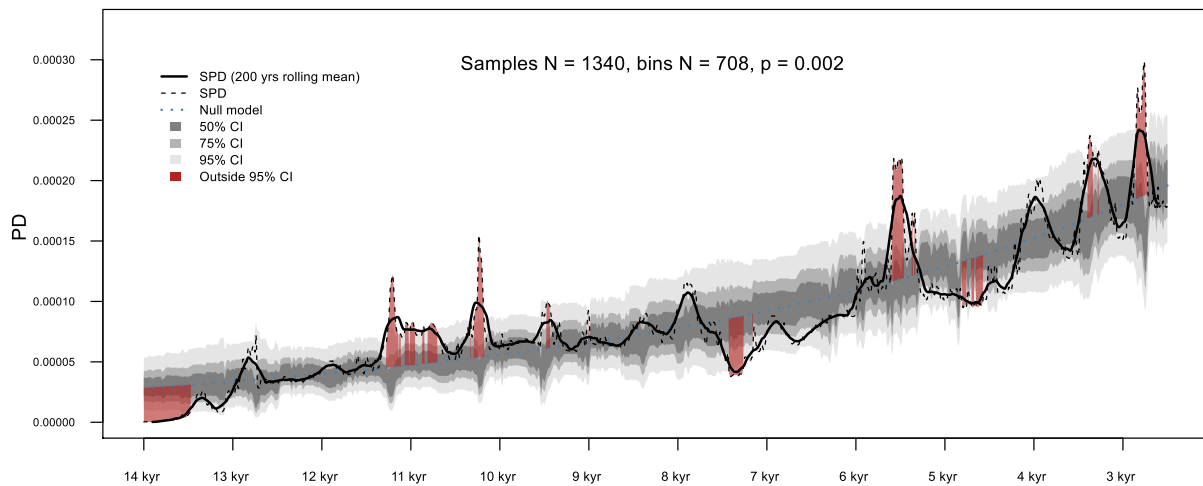


Fig 4. SPD simulation approach illustrating the null hypothesis of steady exponential growth can be rejected

The highly significant p-value of 0.002 is the result of only 35 of the 20,000 simulated SPDs having more periods outside the 95% CI than the observed SPD. However, there are important interpretive limits to be respected. This p-value permits us to confidently reject the hypothesis that these data can be explained by an exponential model, but this does not offer us an alternative plausible population model. Whilst we can interpret the sections outside the CI (highlighted red in Fig 4) as possible periods where the population may have been unusually high or low *relative to the null exponential*, we cannot misappropriate this p-value to confidently validate these interpretations as genuine population dynamics, for two reasons. Firstly, due to random fluctuations we can expect approximately 5% of each simulated SPD to sit outside the CI, so there is no way to identify which local sections are attributable to this random behaviour. Secondly, we still have the fundamental problem that the SPD is not a model of the population dynamics – it is merely a representation of the data being used as a proxy for the population dynamics. Nevertheless, we can successfully reject the simple exponential model, which permits us to explore alternative population models to better explain the data.

(c) CPL modelling

We apply our CPL modelling methods to the same SAAD dataset. As before, the data are binned into 796 discrete site-phase bins, and this is reduced to 708 bins after the exclusion of dates mostly outside the date range (see section: 'Avoiding edge effects'). We then apply four procedures. Firstly, model selection using the BIC establishes the 3-CPL model as best (Fig 5 left). Secondly, the Maximum Likelihood parameters of this model are found using the parameter search. Thirdly, the joint posteriors are estimated using MCMC. Fourthly, we apply the GOF test to this model, giving a p-value of 0.235, which establishes the dataset as a typical outcome of the model.

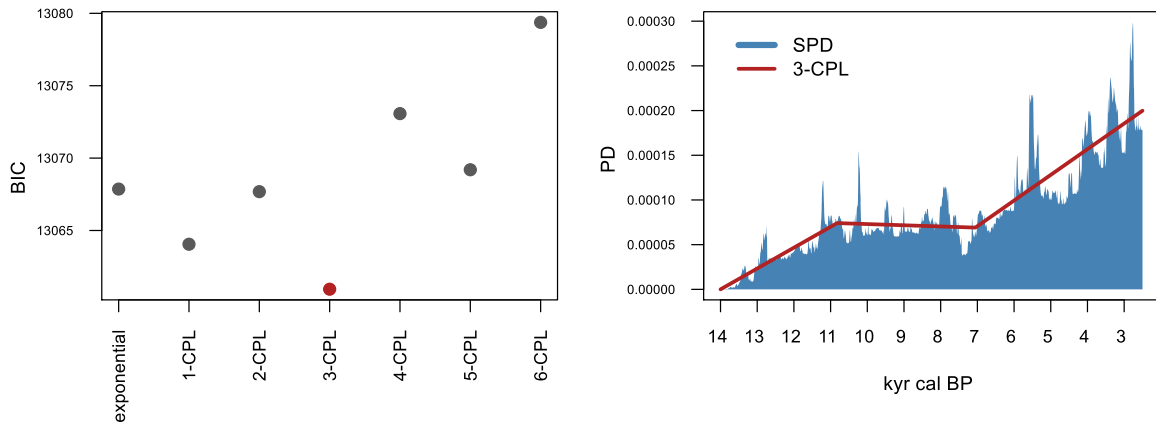


Fig 5. Left: Model comparison between an exponential model and n-CPL models of varying complexity using the BIC establishes the 3-CPL model as best; Right: red line shows the shape of the best 3-CPL model estimated using the Maximum Likelihood parameters; blue polygon shows the calibrated dataset as an SPD.

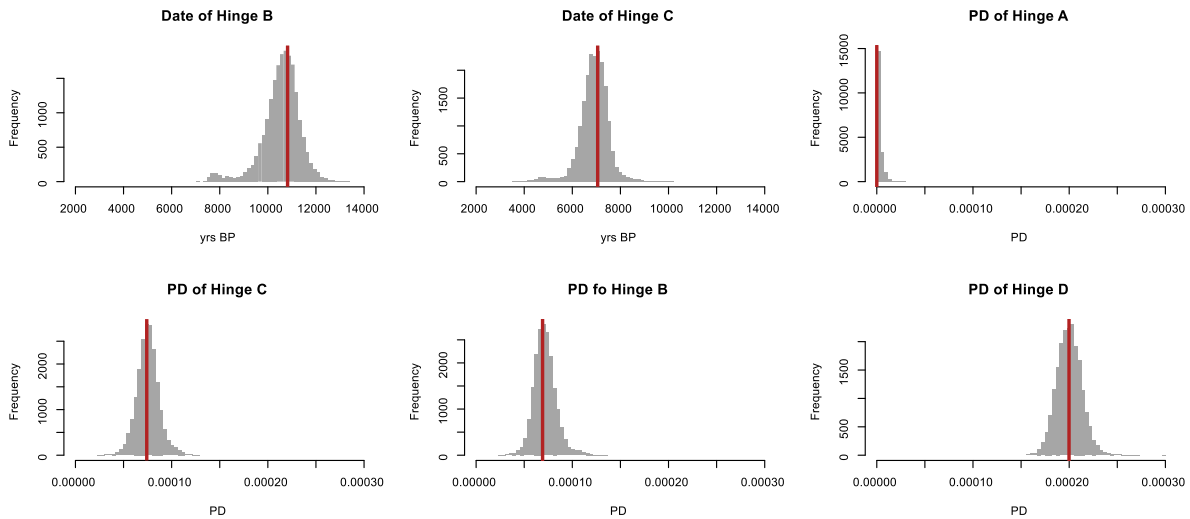


Fig 6. Grey: Marginal Posterior Distributions of the parameters defining the 4 hinge points, estimated using MCMC. The dates of Hinges A and D are not free parameters since they are fixed at 14 kyr BP and 2.5 kyr BP respectively. Red: Maximum Likelihood parameters estimated separately using the search algorithm.

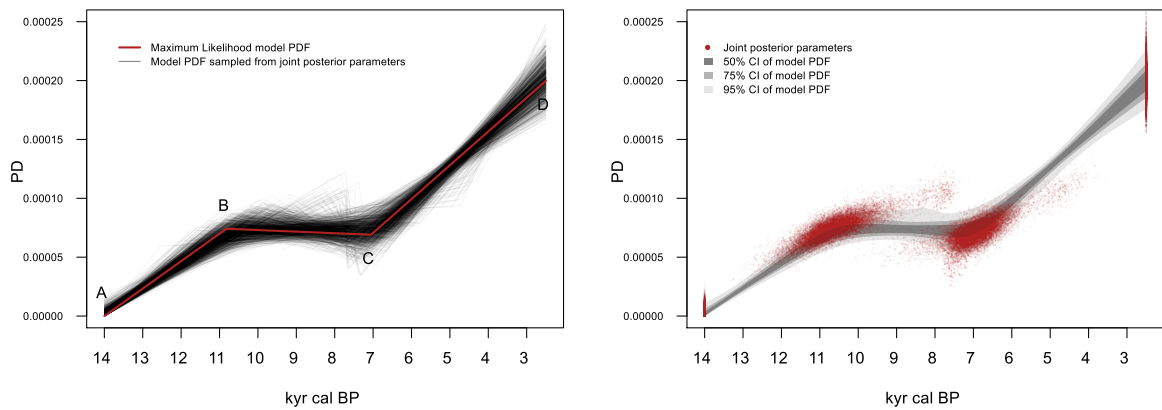


Fig 7. Credible Intervals of the 3-CPL model.

Left: Model PDFs using the joint parameters of 1000 samples from the joint posterior parameter distribution (black) and maximum likelihood parameters (red), hinge points marked A to D.

Right: the 50%, 75% and 95% Credible Intervals (grey) of all model PDFs (grey), and parameter values (red), sampled from the joint posterior parameter distribution.

Linear phase between hinges	Start yrs BP (95% CI)	End yrs BP (95% CI)	Gradient ($\times 10^{-9}$ per year) (95% CI)	Relative growth rate per 25 yr generation (95% CI)
1 (A-B)	14000	10821 (11887 to 8265)	23.3 (15.4 to 28)	4.15% (1.12 to 5.32)
2 (B-C)	10821 (11887 to 8265)	7055 (8013 to 5421)	-1.3 (-61.3 to 7.3)	-0.05% (-1.96 to 0.25)
3 (C-D)	7055 (8013 to 5421)	2500	28.7 (20.1 to 42.5)	0.58% (0.42 to 0.81)

Table 2: Summary of the best 3-CPL model represented as Maximum Likelihood dates of hinge points, and the growth rates of the three phases. 95% CI calculated using quantiles.

13. Discussion

The three phases identified in Table 2 can be contextually informed by the archaeological record from dry regions in South America. While there is recent debate surrounding some earlier human occupations in the Americas [30, 46-49], 15/14 kyr BP represents a widely accepted range for the successful human exploration of the South American continent and of the SAAD in particular [50-52], based on both archaeological evidence and genomic data [30, 53]. In the SAAD, the period extending between 14,000 and 10,821 yr BP is characterized by a remarkably high growth rate of 4.15% per 25 yr generation. Whilst current global population growth rates average c. 30% per generation (just over 1% per annum), this is a consequence of modern technological advances, and recent estimates on the prehistoric growth for human populations indicate a much smaller growth rate of 1% per generation (0.04% per annum) [54-56]. Therefore, the magnitude of population growth in this first phase is unusually high, and far greater than during any subsequent phase. This is likely due to the successful exploration and colonization of diverse and uncontested niches by early human societies [36, 52, 57] resulting in typical spread dynamics [56]. At 14.6 kyr BP the Antarctic Cold Reversal (ACR) began, resulting in colder conditions that were similar to those attributed to the North Atlantic Younger Dryas (YD) stadial [58-61]. This colder climate was accompanied by glacial advances throughout South America and higher lake levels in the Altiplano, which, based on palynological and

glacial geological studies, appear to be a result of precipitation increase over the Altiplano [62, 63]. The amelioration of different ecological niches and richness in resident species such as megafauna, would have increased the carrying capacity of the SAAD, leading to rapid human population growth and exploration of new landscapes. This population expansion would have had a significant ecological impact. In addition to climatic changes (discussed below), the increase in predation rates and niche displacement would ultimately have contributed to the extinction of the American megafauna [50, 64]. The regional pace of this anthropogenic impact on megafauna extinction remains unclear in the SAAD given its sparse evidence in some areas, such as in north-central Chile [65].

The second phase covers almost four millennia between 10,821 and 7055 yr BP and is associated with a slight population decrease (-0.05% per generation, Table 2). It has previously been hypothesised that human populations experienced periodic fluctuations during the mid-Holocene in response to climatic forces [2, 36]. However, our analysis, using a refined dataset and improved method, does not support this hypothesis for the SAAD. Instead, the best model suggests a population that was failing to grow, despite the estimated population size relative to occupiable land still being very low (c. 200,000 people in South America) [36]. Whilst ethnographic and theoretical studies demonstrate how a process of alternating growth and decline offers one possible mechanism that can give the long-term appearance of a stable plateau-like population trend [66, 67], we are unable to identify these hypothesised fluctuations. Indeed, neither did the modelling results presented by Goldberg *et al* [36], which, like our results, indicate little or no change in the population size c. 9 to 7 kyr BP. The question therefore remains as to what prompted such a significant shift from a rapidly growing population to one that was stagnating. Using the broader South American radiocarbon dataset, Riris and Arroyo-Kalin [12] propose three periods (8.4, 8.2 and 8.1 kyr BP) with exceptionally high frequency of climatic anomalies, which they correlate with an initial drop in relative population at and after 8.6 kyr BP and lasting until at least 6 kyr BP. Likewise, Goldberg *et al* [36] identified two mid-Holocene dips from additional SPD simulation analysis. Indeed, a relatively abrupt onset of aridity is recorded in a number of continental and marine records across South America [58, 68], and specifically the SAAD [17, 31, 69-71]. This landscape was therefore remarkably different to the one experienced by the first colonisers. Almost all the megafaunal species were either extinct or going extinct [72], forest cover significantly decreased, surface water availability decreased, and temperatures were higher [57]. Importantly, a recent analysis at the scale of South America has identified demographic declines associated with climate change [12] thus substantiating the case for diverse demographic trajectories behind continent-wide patterns.

Finally, the third phase extends between 7055-2500 yr BP (Table 2) and is characterized by a 0.58 percentage increase per generation. It has been proposed that the mid Holocene increase in population growth rates may have been driven by the development of regional intensification [73]. However, except for the central coast of Peru, most centres of New World crop development were outside the SAAD, such as in the lowlands of central America and the interior of the Amazon [74-76]. Furthermore, the timing of the introduction of domestic species in the SAAD is inconsistent with the start of this third phase. Large parts of the drylands incorporated domesticated plant species as staples only after 4 kyr BP [77, 78]. Similarly llamas appear to have been initially domesticated around 4.5-4 kyr BP in the South-Central Andes, although the full process of camelid domestication would have occurred independently at different times and places within the Andes [79, 80]. Therefore, the shift from slight population decline to an upward growth rate of 0.58% per generation occurs 2.5-3 kyr prior to domestication in the SAAD so does not appear to be associated with the development of agropastoral economies.

Increased sedentism has also been proposed as an explanation for this late phase of growth [36], and whilst a decrease in residential mobility has been proposed in the Atacama desert of northern Chile around 7 kyr BP [81] coinciding with our observed population growth at 7055 yr BP, there also appears to be much regional heterogeneity in mobility and, for instance there is little to no evidence of increased sedentism in large areas of the SAAD at this time, including Patagonia [50, 82].

In contrast to these possible explanations, we note that a growth rate of 0.58% per generation (95% CI = 0.42 to 0.81) is not unusually high, and is instead broadly consistent with (indeed slightly lower than) the global estimates of 1% per generation (0.04% per year) for broadscale background Holocene population growth [54]. Zahid et al. [54] proposed that this background growth rate is a global phenomenon occurring irrespective of the local environment or subsistence strategy and is therefore intrinsic to our species, arguing that it is likely to be related to the global climate and/or endogenous biological factors.

14. Conclusion

Whilst current SPD simulation methods provide a robust statistical framework to test a single null hypothesis, successfully rejecting the null offers the researcher little in the way of drawing an inference about true population dynamics, and this inferential vacuum is often filled with overinterpretation of peaks and troughs in Summed Probability Distributions (SPDs). Furthermore, rejection of the simple exponential model of constant background growth has become so common that it is no longer tenable to use a classical hypothesis test that heavily favours this null. Instead we argue that a model selection approach is more appropriate. By including the exponential in the model selection process there is still the opportunity for this model to be selected, but unlike current simulation methods that can only reject (or fail to reject) an 'assumed correct' model, CPL modelling automatically provides a best explanation. The structure of the CPL model provides meaningful and useful date estimates of historic events; relative population levels and growth rates; avoids overfitting; and the GOF test quantitatively checks if the data is reasonable, given the model. Together, these methods provide a solid inferential framework for evaluating prehistoric population dynamics from ¹⁴C datasets of any size, and naturally avoids the overinterpretation that is common with SPD analysis.

Our SAAD case study provides a demonstration of the need for this more robust inferential methodology. A substantial body of literature has grown to support a claim of mid Holocene population fluctuations, based on the misinterpretation of the available ¹⁴C data and the misappropriation of a significant p-value when using SPD simulation methods. We show that based on the current data this inference is unjustified, and that a steady population trajectory during this period is a better explanation of the data. Directly modelling population dynamics provide robust, justified and reasonable inferences. Our findings should not be misinterpreted as a claim that in reality there were no population fluctuations. Future larger datasets have the potential to support models of much greater complexity, and CPL modelling provides the basis and flexibility of fitting any number of hinges, offering detailed population histories of key events.

Acknowledgements

This research was developed in the context of a Leverhulme Trust International Network Grant (IN2015-042) titled *SPANning the Atlantic: Human Palaeodemography in Southern Hemisphere Drylands*. We sincerely thank the other *SPANning* members: Genevieve Dewar, Jill Kinahan, John Kinahan, Erik Marsh, Gustavo Martínez, Mark McGranaghan, Peter Mitchell, Jayson Orton, Hugo Pinto, Federico Scartascini, Brian Stewart, Amalia Nuevo Delaunay and David Thomas for their thoughtful debates, warm collaboration and generous hospitality. We are also grateful for the Chilean support to

CM from the ANID-FONDECYT #1170408 grant, and Argentinean support to RB from PICT 2016-0062. Finally we thank the guest editors Jennifer French, Phil Riris, Fabio Silva, Sergi Lozano and Javier Fernández López de Pablo, for their kind invitation to contribute to Philosophical Transactions B, to the anonymous reviewers for their detailed and constructive criticisms, and Yoan Diekmann for his useful discussions and advice.

References

- [1] Collard, M., Buchanan, B., Hamilton, M. J. & O'Brien, M. J. 2010 Spatiotemporal dynamics of the Clovis–Folsom transition. *Journal of Archaeological Science* **37**, 2513-2519.
- [2] Barberena, R., Mendez, C. & de Porras, M. E. 2017 Zooming out from archaeological discontinuities: The meaning of mid-Holocene temporal troughs in South American deserts. *Journal of Anthropological Archaeology* **46**, 68-81. (DOI:10.1016/j.jaa.2016.07.003).
- [3] Williams, A. N. 2012 The use of summed radiocarbon probability distributions in archaeology: a review of methods. *Journal of Archaeological Science* **39**, 578-589.
- [4] Hinz, M., Feeser, I., Sjögren, K.-G. & Müller, J. 2012 Demography and the intensity of cultural activities: an evaluation of Funnel Beaker Societies (4200–2800 cal BC). *Journal of archaeological science* **39**, 3331-3340.
- [5] Kuper, R. & Kropelin, S. 2006 Climate-controlled Holocene occupation in the Sahara: motor of Africa's evolution. *Science* **313**, 803-807. (DOI:10.1126/science.1130989).
- [6] Shennan, S. & Edinborough, K. 2007 Prehistoric population history: from the Late Glacial to the Late Neolithic in Central and Northern Europe. *Journal of Archaeological Science* **34**, 1339-1345.
- [7] Wang, C., Lu, H., Zhang, J., Gu, Z. & He, K. 2014 Prehistoric demographic fluctuations in China inferred from radiocarbon data and their linkage with climate change over the past 50,000 years. *Quaternary Sci Rev* **98**, 45-59.
- [8] Peros, M. C., Munoz, S. E., Gajewski, K. & Viau, A. E. 2010 Prehistoric demography of North America inferred from radiocarbon data. *Journal of Archaeological Science* **37**, 656-664.
- [9] Contreras, D. A. & Meadows, J. 2014 Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. *Journal of Archaeological Science* **52**, 591-608.
- [10] Bleicher, N. 2013 Summed radiocarbon probability density functions cannot prove solar forcing of Central European lake-level changes. *The Holocene* **23**, 755-765.
- [11] Surovell, T. A., Finley, J. B., Smith, G. M., Brantingham, P. J. & Kelly, R. 2009 Correcting temporal frequency distributions for taphonomic bias. *Journal of archaeological Science* **36**, 1715-1724.
- [12] Riris, P. & Arroyo-Kalin, M. 2019 Widespread population decline in South America correlates with mid-Holocene climate change. *Scientific reports* **9**, 6850.
- [13] Abraham, E., Garleff, K., Liebricht, H., REGAIRAZ, A. C., Schäbitz, F., Squeo, F. A., Stingl, H. & Villagrán, C. 2000 Geomorphology and paleoecology of the arid diagonal in southern South America.
- [14] Bruniard, E. D. 1982 La diagonal árida argentina: un límite climático real. *Revista Geográfica*, 5-20.
- [15] Walker, M., Gibbard, P., Berkelhammer, M., Björck, S., Cwynar, L., Fisher, D., Long, A., Lowe, J., Newnham, R. & Rasmussen, S. 2014 Formal subdivision of the Holocene series/epoch. In *STRATI 2013* (pp. 983-987, Springer).
- [16] Maldonado, A. & Villagrán, C. 2006 Climate variability over the last 9900 cal yr BP from a swamp forest pollen record along the semiarid coast of Chile. *Quaternary Research* **66**, 246-258.
- [17] Latorre, C., Betancourt, J. L., Rylander, K. A., Quade, J. & Matthei, O. 2003 A vegetation history from the arid prepuna of northern Chile (22–23 S) over the last 13 500 years. *Palaeogeography, Palaeoclimatology, Palaeoecology* **194**, 223-246.
- [18] Duran, V. A., Winocur, D. A., Stern, C., Garvey, R., Barberena, R., Peña Monné, J. L. & Benitez, A. A. 2016 Impacto del volcanismo y glaciario holocénicos en el poblamiento humano de la cordillera sur de Mendoza (Argentina): una perspectiva geoarqueológica.
- [19] Núñez, L. & Santoro, C. M. 1988 Cazadores de la puna seca y salada del área centro-sur andina (norte de Chile). *Estudios atacameños*, 11-60.

- [20] Grosjean, M., Santoro, C. M., Thompson, L. G., Núñez, L. & Standen, V. G. 2007 Mid-Holocene climate and culture change in the South Central Andes. In *Climate Change and Cultural Dynamics* (pp. 51-115, Elsevier).
- [21] Neme, G. & Gil, A. 2009 Human occupation and increasing mid-Holocene aridity: southern Andean perspectives. *Current Anthropology* **50**, 149-163.
- [22] Yacobaccio, H. D. 2013 Towards a human ecology for the Middle Holocene in the Southern Puna. *Quaternary international* **307**, 24-30.
- [23] Gayo, E. M., Latorre, C. & Santoro, C. M. 2015 Timing of occupation and regional settlement patterns revealed by time-series analyses of an archaeological radiocarbon database for the South-Central Andes (16–25 S). *Quaternary International* **356**, 4-14.
- [24] Méndez, C., Gil, A., Neme, G., Delaunay, A. N., Cortegoso, V., Huidobro, C., Durán, V. & Maldonado, A. 2015 Mid Holocene radiocarbon ages in the Subtropical Andes (~ 29°–35° S), climatic change and implications for human space organization. *Quaternary International* **356**, 15-26.
- [25] 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. *Nat Commun* **4**, 2486. (DOI:10.1038/ncomms3486).
- [26] Timpson, A., Colledge, S., Crema, E., Edinborough, K., Kerig, T., Manning, K., Thomas, M. G. & Shennan, S. 2014 Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *Journal of Archaeological Science* **52**, 549-557. (DOI:10.1016/j.jas.2014.08.011).
- [27] Manning, K. & Timpson, A. 2014 The demographic response to Holocene climate change in the Sahara. *Quaternary Sci Rev* **101**, 28-35. (DOI:10.1016/j.quascirev.2014.07.003).
- [28] Crema, E. R., Habu, J., Kobayashi, K. & Madella, M. 2016 Summed probability distribution of 14C dates suggests regional divergences in the population dynamics of the Jomon period in eastern Japan. *PLoS One* **11**, e0154809.
- [29] Bevan, A., Colledge, S., Fuller, D., Fyfe, R., Shennan, S. & Stevens, C. 2017 Holocene fluctuations in human population demonstrate repeated links to food production and climate. *Proceedings of the National Academy of Sciences* **114**, E10524-E10531.
- [30] Prates, L., Politis, G. G. & Perez, S. I. 2020 Rapid radiation of humans in South America after the last glacial maximum: A radiocarbon-based study. *PLoS one* **15**, e0236023.
- [31] Llano, C., de Porras, M. E., Barberena, R., Timpson, A., Beltrame, M. O. & Marsh, E. J. 2020 Human resilience to Holocene climate changes inferred from rodent middens in drylands of northwestern Patagonia (Argentina). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 109894.
- [32] de Pablo, J. F. n.-L. p., Gutiérrez-Roig, M., Gómez-Puche, M., McLaughlin, R., Silva, F. & Lozano, S. 2019 Palaeodemographic modelling supports a population bottleneck during the Pleistocene-Holocene transition in Iberia. *Nat Commun* **10**, 1-13.
- [33] Porčić, M., Blagojević, T. & Stefanović, S. 2016 Demography of the early Neolithic population in central Balkans: population dynamics reconstruction using summed radiocarbon probability distributions. *PLoS One* **11**, e0160832.
- [34] Ramsey, C. B. 2017 Methods for summarizing radiocarbon datasets. *Radiocarbon* **59**, 1809-1833.
- [35] Haslett, J. & Parnell, A. 2008 A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **57**, 399-418.
- [36] Goldberg, A., Mychajliw, A. M. & Hadly, E. A. 2016 Post-invasion demography of prehistoric humans in South America. *Nature* **532**, 232.
- [37] Schwarz, G. 1978 Estimating the dimension of a model. *The annals of statistics* **6**, 461-464.
- [38] Burnham, K. P. & Anderson, D. R. 2004 Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research* **33**, 261-304.
- [39] Akaike, H. 1974 A new look at the statistical model identification. *IEEE transactions on automatic control* **19**, 716-723.
- [40] 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.
- [41] Hastings, W. K. 1970 Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97-109.
- [42] Team, R. C. 2013 R: A language and environment for statistical computing.

- [43] Peel, M. C., Finlayson, B. L. & McMahon, T. A. 2007 Updated world map of the Köppen-Geiger climate classification.
- [44] Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F. 2006 World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* **15**, 259-263.
- [45] Hogg, A. G., Heaton, T. J., Hua, Q., Palmer, J. G., Turney, C. S., Southon, J., Bayliss, A., Blackwell, P. G., Boswijk, G. & Ramsey, C. B. 2020 SHCal20 Southern Hemisphere calibration, 0–55,000 years cal BP. *Radiocarbon* **62**, 759-778.
- [46] Ardelean, C. F., Becerra-Valdivia, L., Pedersen, M. W., Schwenninger, J. L., Oviatt, C. G., Macías-Quintero, J. I., Arroyo-Cabral, J., Sikora, M., Ocampo-Díaz, Y. Z. E., Rubio, C., II, et al. 2020 Evidence of human occupation in Mexico around the Last Glacial Maximum. *Nature*. (DOI:10.1038/s41586-020-2509-0).
- [47] Becerra-Valdivia, L. & Higham, T. 2020 The timing and effect of the earliest human arrivals in North America. *Nature*. (DOI:10.1038/s41586-020-2491-6).
- [48] Dillehay, T. D., Ocampo, C., Saavedra, J., Sawakuchi, A. O., Vega, R. M., Pino, M., Collins, M. B., Cummings, L. S., Arregui, I. & Villagran, X. S. 2015 New archaeological evidence for an early human presence at Monte Verde, Chile. *PLoS one* **10**.
- [49] Boëda, E., Clemente-Conte, I., Fontugne, M., Lahaye, C., Pino, M., Felice, G. D., Guidon, N., Hoeltz, S., Lourdeau, A. & Pagli, M. 2014 A new late Pleistocene archaeological sequence in South America: the Vale da Pedra Furada (Piauí, Brazil). *Antiquity* **88**, 927-941.
- [50] Borrero, L. A. 2016 Ambiguity and debates on the early peopling of South America. *PaleoAmerica* **2**, 11-21.
- [51] Dillehay, T. D., Goodbred, S., Pino, M., Sánchez, V. F. V., Tham, T. R., Adovasio, J., Collins, M. B., Netherly, P. J., Hastorf, C. A. & Chiou, K. L. 2017 Simple technologies and diverse food strategies of the Late Pleistocene and Early Holocene at Huaca Prieta, Coastal Peru. *Science Advances* **3**, e1602778.
- [52] Méndez-Melgar, C. 2013 Terminal Pleistocene/early Holocene 14C dates from archaeological sites in Chile: Critical chronological issues for the initial peopling of the region. *Quaternary International* **301**, 60-73.
- [53] Posth, C., Nakatsuka, N., Lazaridis, I., Skoglund, P., Mallick, S., Lamnidis, T. C., Rohland, N., Nägele, K., Adamski, N. & Bertolini, E. 2018 Reconstructing the deep population history of Central and South America. *Cell* **175**, 1185-1197. e1122.
- [54] Zahid, H. J., Robinson, E. & Kelly, R. L. 2016 Agriculture, population growth, and statistical analysis of the radiocarbon record. *Proceedings of the National Academy of Sciences* **113**, 931-935.
- [55] Johnson, C. N. & Brook, B. W. 2011 Reconstructing the dynamics of ancient human populations from radiocarbon dates: 10 000 years of population growth in Australia. *Proceedings of the Royal Society B: Biological Sciences* **278**, 3748-3754.
- [56] Arim, M., Abades, S. R., Neill, P. E., Lima, M. & Marquet, P. A. 2006 Spread dynamics of invasive species. *Proc Natl Acad Sci U S A* **103**, 374-378. (DOI:10.1073/pnas.0504272102).
- [57] Marshall, C. R., Lindsey, E. L., Villavicencio, N. A. & Barnosky, A. D. 2015 A Quantitative Model for Distinguishing Between Climate Change, Human Impact, and Their Synergistic Interaction as Drivers of the Late Quaternary Megafaunal Extinctions. *The Paleontological Society Papers* **21**, 1-20.
- [58] Thompson, L. G., Davis, M. E., Mosley-Thompson, E., Sowers, T. A., Henderson, K. A., Zagorodnov, V. S., Lin, P.-N., Mikhailenko, V. N., Campen, R. K. & Bolzan, J. F. 1998 A 25,000-year tropical climate history from Bolivian ice cores. *science* **282**, 1858-1864.
- [59] Thompson, L. G., Mosley-Thompson, E. & Henderson, K. A. 2000 Ice-core palaeoclimate records in tropical South America since the Last Glacial Maximum. *Journal of Quaternary Science: Published for the Quaternary Research Association* **15**, 377-394.
- [60] Pedro, J. B., Bostock, H. C., Bitz, C. M., He, F., Vandergoes, M. J., Steig, E. J., Chase, B. M., Krause, C. E., Rasmussen, S. O., Markle, B. R., et al. 2016 The spatial extent and dynamics of the Antarctic Cold Reversal. *Nature Geoscience* **9**, 51-55. (DOI:10.1038/ngeo2580).
- [61] Palacios, D., Stokes, C. R., Phillips, F. M., Clague, J. J., Alcalá-Reygosa, J., Andrés, N., Angel, I., Blard, P.-H., Briner, J. P., Hall, B. L., et al. 2020 The deglaciation of the Americas during the Last Glacial Termination. *Earth-Science Reviews* **203**, 103113. (DOI:<https://doi.org/10.1016/j.earscirev.2020.103113>).
- [62] Seltzer, G. O. 1994 A lacustrine record of late Pleistocene climatic change in the subtropical Andes. *Boreas* **23**, 105-111.
- [63] Seltzer, G. O. 1992 Late Quaternary glaciation of the Cordillera Real, Bolivia. *J Quaternary Sci* **7**, 87-98.
- [64] Villavicencio, N. A., Lindsey, E. L., Martin, F. M., Borrero, L. A., Moreno, P. I., Marshall, C. R. & Barnosky, A. D. 2016 Combination of humans, climate, and vegetation change triggered Late Quaternary megafauna extinction in the Última Esperanza region, southern Patagonia, Chile. *Ecography* **39**, 125-140. (DOI:10.1111/ecog.01606).

- [65] Méndez, C., Seguel Quintana, R., Nuevo-Delaunay, A., Murillo, I., López Mendoza, P., Jackson, D. & Maldonado, A. 2020 Depositional Contexts and New Age Controls for Terminal-Pleistocene Megafauna in North-central Chile (31°50'S). *PaleoAmerica*, 1-17. (DOI:10.1080/20555563.2020.1733384).
- [66] Pennington, R. 2001 Hunter-gatherer demography. *Hunter-gatherers: An interdisciplinary perspective*, 170-204.
- [67] Hill, K., Hurtado, A. M. & Walker, R. S. 2007 High adult mortality among Hiwi hunter-gatherers: Implications for human evolution. *J Hum Evol* **52**, 443-454.
- [68] Bertrand, S., Charlet, F., Charlier, B., Renson, V. & Fagel, N. 2008 Climate variability of southern Chile since the Last Glacial Maximum: a continuous sedimentological record from Lago Puyehue (40 S). *Journal of Paleolimnology* **39**, 179-195.
- [69] Lamy, F., Hebbeln, D. & Wefer, G. 1999 High-Resolution Marine Record of Climatic Change in Mid-latitude Chile during the Last 28,000 Years Based on Terrigenous Sediment Parameters. *Quaternary Research* **51**, 83-93. (DOI:<https://doi.org/10.1006/qres.1998.2010>).
- [70] Jenny, B., Valero-Garcés, B. L., Villa, M., x, nez, R., Urrutia, R., Geyh, M. & Veit, H. 2002 Early to Mid-Holocene Aridity in Central Chile and the Southern Westerlies: The Laguna Aculeo Record (34°S). *Quaternary Research* **58**, 160-170. (DOI:<https://doi.org/10.1006/qres.2002.2370>).
- [71] Valero-Garcés, B. L., Jenny, B., Rondanelli, M., Delgado-Huertas, A., Burns, S. J., Veit, H. & Moreno, A. 2005 Palaeohydrology of Laguna de Tagua Tagua (34° 30' S) and moisture fluctuations in Central Chile for the last 46 000 yr. *J Quaternary Sci* **20**, 625-641. (DOI:10.1002/jqs.988).
- [72] Steadman, D., Martin, P., MacPhee, R., Jull, A., McDonald, H., Woods, C., Iturralde-Vinent, M. & Hodgins, G. 2007 Early Holocene survival of megafauna in South America. *Journal of Biogeography*, 1642-1646.
- [73] Perez, S. I., Postillone, M. B. & Rindel, D. 2017 Domestication and human demographic history in South America. *Am J Phys Anthropol* **163**, 44-52.
- [74] Piperno, D. R. 2011 The origins of plant cultivation and domestication in the New World tropics: patterns, process, and new developments. *Current anthropology* **52**, S453-S470.
- [75] Kistler, L., Maezumi, S. Y., De Souza, J. G., Przelomska, N. A., Costa, F. M., Smith, O., Loïselle, H., Ramos-Madriral, J., Wales, N. & Ribeiro, E. R. 2018 Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science* **362**, 1309-1313.
- [76] Lombardo, U., Iriarte, J., Hilbert, L., Ruiz-Pérez, J., Capriles, J. M. & Veit, H. 2020 Early Holocene crop cultivation and landscape modification in Amazonia. *Nature*. (DOI:10.1038/s41586-020-2162-7).
- [77] Pearsall, D. M. 2008 Plant domestication and the shift to agriculture in the Andes. In *The handbook of South American archaeology* (pp. 105-120, Springer).
- [78] Larson, G., Piperno, D. R., Allaby, R. G., Purugganan, M. D., Andersson, L., Arroyo-Kalin, M., Barton, L., Climer Vigueira, C., Denham, T., Dobney, K., et al. 2014 Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences* **111**, 6139. (DOI:10.1073/pnas.1323964111).
- [79] Mengoni-Góñalons, G. L. 2008 Camelids in ancient Andean societies: A review of the zooarchaeological evidence. *Quaternary International* **185**, 59-68.
- [80] Moore, K. M. 2016 Early domesticated camelids in the Andes. *The archaeology of Andean pastoralism*, 17-38.
- [81] Marquet, P. A., Santoro, C. M., Latorre, C., Standen, V. G., Abades, S. R., Rivadeneira, M. M., Arriaza, B. & Hochberg, M. E. 2012 Emergence of social complexity among coastal hunter-gatherers in the Atacama Desert of northern Chile. *Proceedings of the National Academy of Sciences* **109**, 14754-14760.
- [82] Perez, S. I., Postillone, M. B., Rindel, D., Gobbo, D., Gonzalez, P. N. & Bernal, V. 2016 Peopling time, spatial occupation and demography of Late Pleistocene–Holocene human population from Patagonia. *Quaternary International* **425**, 214-223. (DOI:<https://doi.org/10.1016/j.quaint.2016.05.004>).