Modelling Human Range Expansion Across a Heterogeneous Cost Surface

Celia Glass, James Steele, David Wheatley

CAA 97

Abstract

Understanding the first human colonization of South America depends on accurate dating of early sites, and on realistic models of the effects of habitat variation on dispersal rates. In this study, we report application of regression models to analysis of the spatial structure of the early radiocarbon record of human occupation in that continent. Our method allows for the differing levels of precision in radiocarbon dates from different early sites. Cost distance analysis was used to explore the effects of habitat variability on rates of spread. Our analysis suggests that humans coming into South America in the late glacial dispersed most rapidly through the more productive open habitats, including the open montane habitat of the Andean mountain chain.

1 Introduction

This paper reports work done by our group on the first human colonization of South America. Most archaeologists will know that there is continuing debate on the timing of this colonization: everyone agrees that humans were leaving their mark in the archaeological record by about 12,000 years b.p., but new sites are regularly reported with dates that suggest an *earlier* human presence there. The problem is that these dates are usually marred by unreliable associations with cultural material, while the absence of convincing evidence of equally early (pre-12,000 b.p.) human occupation of *North* America makes it hard to see where the putative pre-Clovis South American population could have come from. What *is* clear is that the 'ice-free' corridor between the two North American ice sheets, which is usually thought to have given humans access south from Beringia, only opened up after about 13,000 b.p. (e.g. Lundqvist and Saarnisto 1995). It is also clear that radiocarbon-dated sites from that period of the late glacial are extremely sparse in both North and South America:
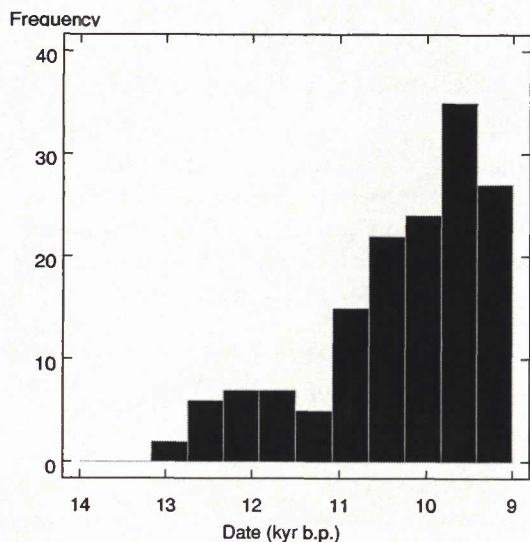
Figure 1 graphs the increasing frequency of archaeological radiocarbon dates from southern South America from 13,000 to 8,000 b.p., suggesting that in the earliest phase human occupation of that continent was very sparse (Borrero 1996). If we break down these events into sites by region, we see that there is also some suggestion of regional differences (Figure 2): the greatest expansion of site numbers in the late glacial occurs in Brazil and in northern Chile/northwest Argentina, whereas the early but sparse occupation of Peru and Patagonia seems to establish a pattern for those regions which characterized the rest of the late glacial. The increase in dated events post-12,000 b.p. must indicate either the expansion from a glacial refugium of a pre-existing South American population, or the spread of a colonizing population coming into the continent — presumably from the north.

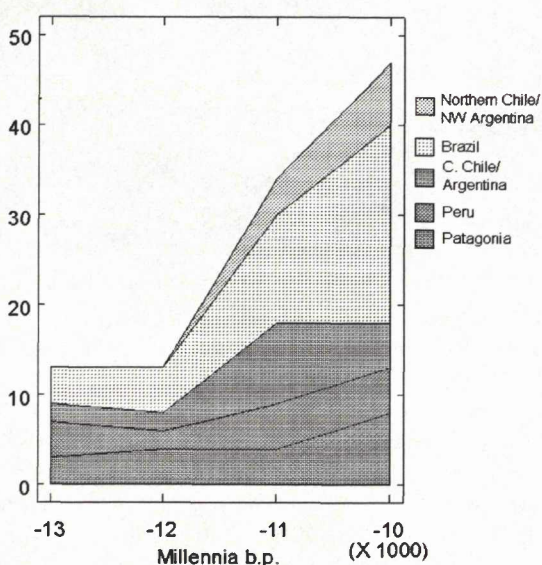Figure 1. Early South American sites by age (data from Borrero 1996).

Figure 2. Early South American sites by age and region (data from Borrero 1996).

The work which we report in this paper examines the *spatial* structure of first clear human occupation of South America in the late glacial. We explicitly address only one question relevant to these wider archaeological debates. Do the earlier sites with clear signs of human presence tend to be closer to the farthest northwestern part of the continent (suggesting a pattern of late glacial first colonization), or are they scattered more randomly across the landscape (suggesting a pre-existing population reaching archaeological visibility)? In order to address this question, we have taken a small sample of early South American sites which are accepted by the majority of archaeologists, and which excludes a number of highly controversial sites and phases (Figure 3; Table 1). This is not our own sample: we have taken it, subject to minor amendments and two additions, from a paper published in *American Antiquity* in 1993 by David Whitley and Ronald Dorn. Whitley and Dorn's own analyses led them to argue that people must have reached South America well before the time of the North American Clovis culture. But were they right?
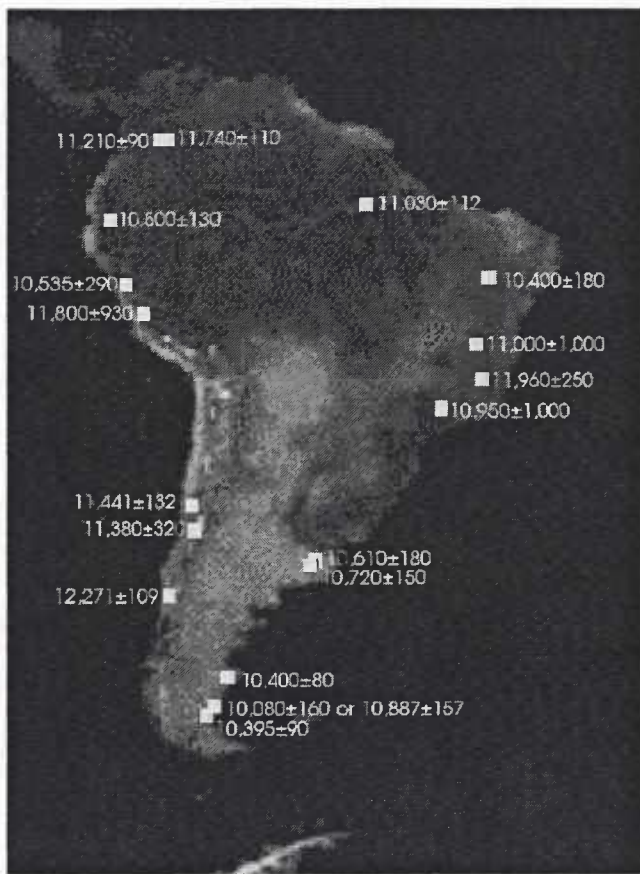

Figure 3. The locations of the sites used in this analysis, with their radiocarbon dates.

2 Methodology

To examine the spatial structure of this sample of early sites we have used weighted linear regression techniques. We take as our two variables *distance* to a site from the north-western-most point of South America at the junction with the Central American isthmus, and the *date* of the site. Our hypothesis is that dates should become progressively younger with increasing distance. If this hypothesis is not supported, then we will have to reconsider the model of a late glacial colonization event.

Because a radiocarbon date is a probability distribution and regression techniques require point values, we have taken the modal value of the calibrated probability distribution as our point value, and weighted it inversely to the square of half the one-sigma range. For normally distributed radiocarbon dates, this method would be adequate. Recently, however, the radiocarbon calibration curve has been extended back into the late glacial (Stuiver and Reimer 1993): clearly we must calibrate our dates before entering them into the regression, but the irregular shape of calibrated probability distributions for radiocarbon dates makes our statistical treatment problematic. In practice the early part of the calibration curve is sufficiently smooth for this to present no major problems: we are, however, currently examining ways of treating calibrated dates statistically which take account of such irregularities.
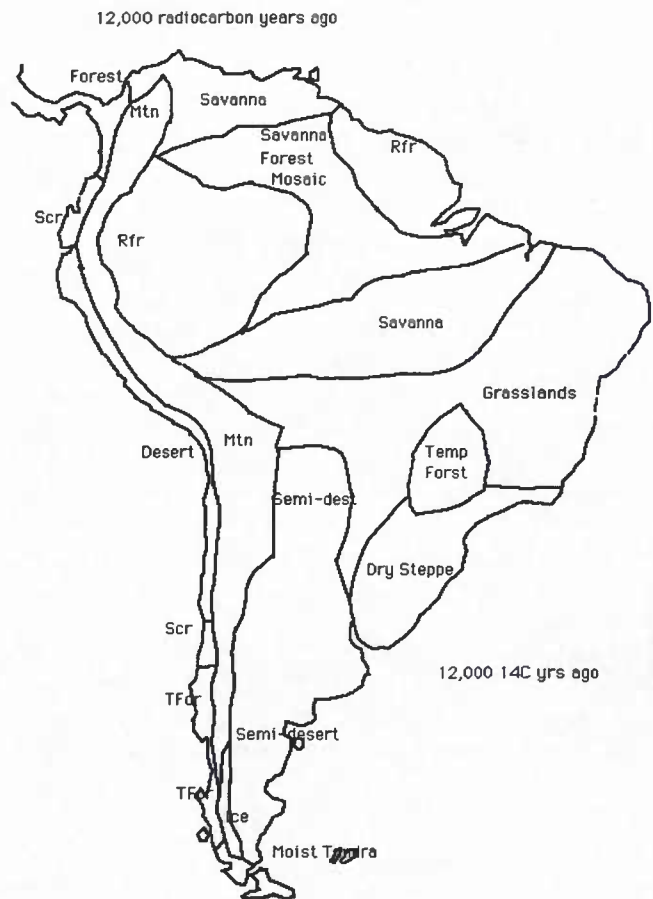

Figure 4. The paleovegetation map used in this analysis (from Adams 1995).

In an earlier analysis of this sample using weighted linear regression (there, of date against distance), two contrasting distance measures were used (Steele *et al.* in press). These were the geodesic distance, and the distance as measured on a road atlas from El Paso (a point at the south-western boundary of the U.S. Clovis culture area) to each of the early South American archaeological sites in Whitley and Dorn's sample. Results were discouraging (Table 2). Although in all cases the best-fit model showed a trend for sites to get younger with increasing distance, this was only statistically significant when Monte Verde 2 – the famously early site in southern Chile, dated to about 12,300 ^{14}C years b.p. – was omitted. Since a panel of leading sceptics has now visited that site and declared it to be both reliably dated and reliably associated with human artefacts, we cannot accept results

68

which leave Monte Verde 2 out of the picture. On the face of it, therefore, we must reject our hypothesis that the spatial structure of early South American sites indicates an expansion from the north-west.

But how should we measure distance to these sites in ways which take adequate account of habitat variation and its effects on human expansion? Whitley and Dorn argued that distances measured from road atlases represented "reasonable and conservative surrogate measures for shortest and most feasible routes, assuming foreknowledge of the local environment on the part of the migrating population" (1993: 647). But while this measure takes account of topography, it takes little or no account of differences in the accessibility of the different vegetation zones of late glacial South America to a hunter-gatherer population, differences which we would expect to have had large effects on human dispersals. We propose that a better measure of effective distance for a colonizing population would be cost distance – a function of the relative frictions on expansion presented by different vegetation types.

For our new analyses, we have therefore taken a preliminary reconstruction of the vegetation zones of South America for the period 12,000 – 11,000 b.p. (Figure 4, from Adams, n.d.), and imported it into a raster GIS system (GRASS). We have then calculated cost distances from the north-western-most part of South America to each of the sites in our sample under varying assumptions about the frictions which different vegetation types would have presented to a colonizing population. We wanted to explore the space of possible friction weightings for each of these vegetation types, without prejudging the issue of which ones would have been most accessible to a colonizing population. To reduce computer time, we first assimilated each of the fourteen habitat types in our original reconstruction to one of four broad categories (Table 3).

One was for the more *Productive Open Habitats*, one was for *Closed Forest Habitats*, and one category contains a more heterogeneous mix of *Intermediate/Mixed* vegetation types. *Sea and Ice* were assumed to constitute a fourth category with essentially infinite friction value: in other words, we assumed that the colonization process took place by land. We then independently varied the frictions of each of the first three categories systematically, using eight possible friction values in a logarithmic series [1, 3, 9, 27, 81, 243, 729, 2187]. We were interested in discovering which combinations of friction values gave cost distances to each of our archaeological sites which best approximated the order seen in their radiocarbon dates. Figure 5 gives some examples of such cost surfaces, with the locations of the archaeological sites plotted on them.
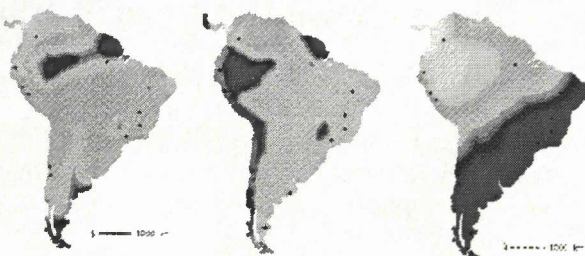

Figure 5. The cost surfaces – examples.

Since weighted linear regression is affected only by changes in the *relative* cost distances to each site, we were interested only in the effects of varying the *relative* frictions of our three main vegetation categories. We were thus able to further reduce the number of analyses required by varying friction values independently for only two vegetation categories at a time, in each case holding the third one constant with a value of 1. Our analysis therefore has the potential to tell us about the *relative* accessibility of different habitats to a colonizing population, but not to tell us the *absolute* rates of expansion across this or that habitat type, though this is a subject of further investigation.

To measure goodness of fit, we used the coefficient of determination (r^2) for the weighted linear regression of date against distance. This gives us an estimate of the proportion of the variation in calibrated dates for our sample which is accounted for by variation in cost distance to each site from a common point of origin. In looking at the effects of changing the cost distance variable on r^2, this exercise should ultimately enable us to rank habitat types in terms of their accessibility to an expanding human population, using empirical archaeological data. Our expectations are modest: as Table 2 shows, adjusted values for r^2 in the initial analyses using road map and geodesic distance were, respectively, 0.01 and 0.08!

3 Results

The main finding of our new regression analyses, done using cost distance as our spatial measure, is that the best weighted r^2 which we are able to achieve with our given dataset is 0.42. In other words, we can still explain less than half of the variation in calibrated dates using the best cost distance measure. While this figure seems low, it is nonetheless an improvement of more than threefold on the same measure of fit for a homogeneous surface. Moreover, we cannot expect a very good fit of the regression model given the uncertainties in the dates of the eighteen sites we are using.
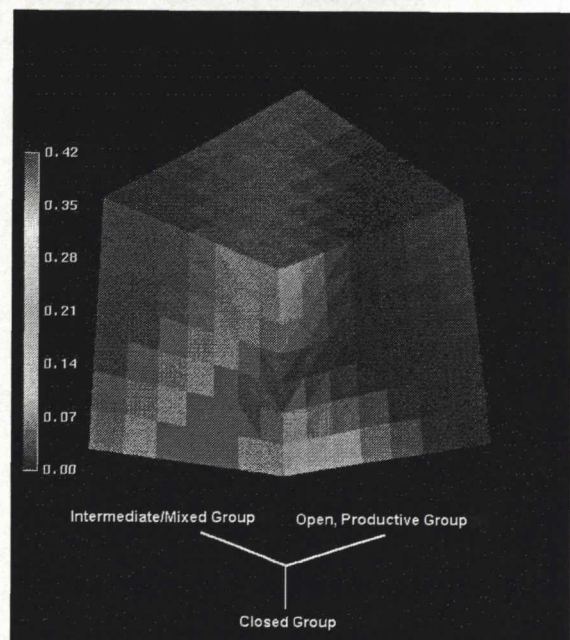

Figure 6. The matrix of r^2 values generated in this analysis, using the earlier date for Fell's Cave.

Another very encouraging finding is the consistent way in which r^2 varies with varying friction values for the three broad vegetation groups (Figure 6). The highest correlations between cost distance and calibrated dates were found when the parts of the vegetation surface which belong to the intermediate or heterogeneous vegetation group were given the highest friction values – and these *last* values must be very substantially greater than those for the open, productive habitat group. The majority of the highest correlations were also associated with classifications of the surface in which the open, productive habitat group were given the lowest friction values, and the closed forest parts of the surface were given intermediate friction values. For example, the highest value for r^2 was achieved when the friction ratios were 1 : 2187 : 243 (for the vegetation groups open : heterogenous mixed : closed).

There were some exceptions to this pattern: in some cases, a high r^2 was achieved when the closed habitat group was given lower friction weight than the open group, although in such cases the intermediate/mixed group always maintained a relatively high friction weight. Further numerical details may be found in our forthcoming Working Paper (Glass *et al.* 1997).

Finally, we were uncertain which to use of two possible dates for Fell's Cave, the early site in Patagonia – although the probability was that the earlier date was the more reliable. We therefore repeated the analysis using the later date. There was a remarkable consistency in the weighted r^2 results for the two sets of data. Those in which the earlier date for Fell's Cave was used gave consistently higher values for that measure.

4 Discussion

More than thirty years ago, Bennett and Bird (1965) argued that the earliest humans to colonize South America would most probably have come via the isthmus of Panama, migrating up into the Andean highlands via the Cauca and Magdalena rivers (which both flow south to north). The Andean environment would (they proposed) have offered little subsequent obstacle to southward migration, with expansion into the Pampas and Patagonia once Argentina had been reached, and entry into the eastern Brazilian highlands from northern Argentina. Lynch (1983) endorsed that view, and also proposed that '"openness" ranked high on the Paleo-Indian list of desirable habitat characteristics' (*ibid*.: 111): he thus argued that the observed predominance of early sites in open habitats was not due solely to archaeological survey bias. The recently reported site of Monte Alegre in central Amazonia (Roosevelt *et al.* 1996) complicates this picture, and we have included that site in our sample: but it nonetheless remains our impression that the spatial patterning in these dated early South American sites can best be explained as a result of preferential dispersal into the more productive open habitats. Further work needs to be done both to improve the accuracy and precision of the dates for a number of these sites, and to enlarge the existing sample of generally accepted late glacial occupation events, before we can use these results either as predictors of undiscovered site locations, or as a control on simulation models of demic expansion (e.g. Steele, Sluckin, Denholm

and Gamble 1996). Nonetheless, we believe that our results both vindicate our spatial methodology, and tend to support the conventional model of a late glacial colonization of South America.

A number of further tasks have come to light as a result of this phase of our work. The first involves developing more reliable ways of including calibrated radiocarbon dates in spatial analyses which require the dates to be assigned point values. While we believe our approach – using the modal value weighted inversely to the square of half the one-sigma range – is reasonable, it becomes increasingly inaccurate with increasing irregularity and skew in the calibrated probability distribution. We are developing a bootstrapping approach to this problem, as a companion project to that reported in this paper.

The second task we need to address concerns the classification of habitat into broad groups for cost distance analysis. There are many reasons why we should try to assimilate the fine vegetation distinctions of our original paleohabitat map to as few coarse grained categories as we can: but in the analyses reported here, we may have pared down our broad categories to at least one too few. Interpretation of this first set of results requires us to review the criteria for including habitat types in that broad category of intermediate or heterogeneous types which needs to have a high friction weight (Table 3). In addition to a series of less productive open habitats (moist tundra, desert, semi-desert, and scrub), that group contains one anomaly – the tract of savanna/forest cover reconstructed for a drier northwest-southeast transect across the Amazon basin (Figure 4). We were uncertain whether to treat this as an open or as a closed vegetation type, not least because the paleoenvironmental literature records continuing debate over the extent to which last glacial cooling led to the division of the equatorial forest into patches separated by extensive tracts of savanna grassland or of savanna/woodland mix. Inspection of the sites and the paleovegetation map suggests that the reason we need to give a high friction weight to the broad class of vegetation types which include this reconstructed Amazonian savanna-forest tract, is that we need to make that tract relatively inaccessible if we are to match the observed dates of first human occupation immediately to its south and east. This needs to be confirmed by future analyses of the effects of independently varying friction values of the savanna/forest region alone.

The remaining tasks relate to the quality of the archaeological and paleoecological data which we have used. It is clear not only that many of the sites in our given sample would benefit from more precise and accurate dating, but also that the sample is too small and potentially systematically biased to contemporary habitats which are accessible to survey. It is also likely that the paleoecological reconstructions are based on too few pollen core control points to be wholly reliable – an additional source of uncertainty which can also only be remedied by further field work (Adams 1995). Nonetheless, we believe that these kinds of analyses of the existing data can, and indeed should, guide research designs for future fieldwork.

70

Site	Lat/Long[a]	Paleoeconomic indicators	Cultural materials	[14]C dates[d,e]	S.D.	Cal. yrs B.P.	Range to one- and two-sigma, Cal. yrs B.P.	Dist. (geodesic)[b]	Dist. (road map)[c]
1. El Abra, Columbia	[5º1' N, 73º57' W]		Unifacial percussion-flake tools including scrapers, knives and spoke-shaves.	11210	90	13117	13225-13018 13346-12918	4498	6,470
2. Tibito, Columbia	[4º58' N, 73º59' W]	Mastodon, horse and deer remains.	Unifacial assemblage of cores and flake tools.	11740	110	13685	13853-13534 14031-13396	4513	6,470
3. Cubilan, Ecuador	[3º37' S, 79º 14' W]		Leaf-shaped and tanged points.	10500	130	12373	12539-12168 12686-11870	4882	7,320
4. Guitarrero, Peru	9º12' S, 77º43' W	Deer, camelid, birds, other animals.	Bifacial projectile points, scrapers, perforators, blades, cores, hammerstones.	10535	290	12414	12722-11954 13003-11005	5493	8,480
5. Pachamachay, Peru	11º7' S, 76º11' W	Unidentified large mammals (fragmentary remains).	Straight-edged and curved-edge unifaces, unstemmed projectile points.	11800	930	13708	14947-12755 16330-11004	5763	8,990
6. Monte Alegre, Brazil	1º 60' S, 54º 4' W	Fish, mollusks, reptiles, birds, large mammals. Plant remains including tree fruits.	Triangular stemmed bifacial points, limaces, flakes and blades; red pigment rock.	11030	112	12944	13065-12828 13189-12708	6679	[N/A]
7. Quereo, Chile	[31º50' S, 71º18' W]	Mastodon, horse, deer, camelid and other animal remains.	Crude stone flakes and possible simple tools.	11441	132	13353	13520-13205 13708-13067	7990	11,720
8. Pedra Furada, Brazil	8º50' S, 42º33' W		Unifacially retouched 'limaces', choppers, pebble hammerstones, and flakes.	10400	180	12243	12481-11880 12677-11087	8153	11,994
9. Tagua-Tagua, Chile	[34º18' S, 71º2' W]	Mastodon, horse, aquatic birds, camelid and other animal remains; some with possible butchery marks.	Flakes, hammer stones, flakers, cores; a few miscellaneous bone tools.	11380	320	13248	13610-12930 14008-12616	8239	12,020
10. Lapa do Boquete, Brazil	[15º29' S, 44º22' W]		Unifacial tools.	11000	1000	12879	13950-11046 15387-9965	8455	12,440
11. Alice Boer, Bed III, Brazil	[22º24' S, 47º33' W]		Unifacial flake scrapers, knives; bifacial pressure-flaked tools including contracting stemmed (tanged) points.	10970	1020	10940	11960-9920	8702	12,880
12. Abrigo de Santana, Brazil	[19º0' S, 43º53' W]		Two small quartz flakes, fragments of red ochre.	11960	250	13895	14231-13589 14607-13309	8744	12,440
13. Monte Verde II, Chile	41º30' S, 73º15' W	Plant remains indicating an economy focused primarily on plant gathering.	Edge-used flakes; unifacial and bifacial edge-trimmed tools, including bifacial, projectile points; pecked and ground bola and grinding stones. Organic material artefacts.	12271	109	14332	14540-14146 14760-13972	8840	12,900
14. Cerro La China II, Argentina	37º57' S, 58º37' W		Bifacial tools, including fishtail projectile points.	10610	180	12497	12689-12270 12867-11939	9213	13,480
15. Cerro La China I, Argentina	37º57' S, 58º37' W	Armadillo remains.	Fishtail projectile points.	10720	150	12611	12767-12440 12916-12234	9213	13,480
16. Piedra Museo, Argentina	47º54' S, 67º52' W	50 m. from a paleolagoon. Fauna include ostrich (cf. *Rhea americana*), camelid (*Lama (Vicugna) gracilis*), and horse (*Hippidion saldiasi*).	Fishtail projectile points, unifacially-worked scrapers and knives (some heat-treated).	10400	80	12243	12376-12079 12489-11862	9669	[N/A]
17. Fell's Cave, Argentina	52º4' S, 70º0' W	Mylodon, horse, guanaco, bird and other remains.	Fishtail projectile points, utilized flakes, cores, ground and pecked discs.	10080 or 10887	160 157	11321 12810	11999-11001 12291-10954 12972-12647 13133-12470	9987	14,320
18. Tres Arroyos, Argentina	53º21' S, 68º15' W	Guanaco, fox, horse, camelid, bird and rodent remains.	Edge-trimmed flake tools.	10395	90	12291	12425-12128 12541-11910	10170	14,420

GENERAL NOTES:

[a] Lat/Long is approximate, and estimated, in cases enclosed in square brackets. Otherwise, exact as extracted from the literature.

[b] Great circle distance (kms) from El Paso, located at 31º 46' N, 106º 29' W.

[c] Shortest roadway distances (kms) from El Paso as calculated from road maps by Whitley and Dorn (1993:647), who "take these as reasonable and conservative surrogate measures for the shortest and most-feasible routes, assuming foreknowledge of the local environment on the part of the migrating population."

[d] Uncalibrated radiocarbon dates.

e A southern hemisphere correction of -40 radiocarbon years was applied to sites south of the equator (Stuiver and Reimer 1993). This, of course, takes no account of variation through time in the effect of hemispheric circulation patterns on atmospheric carbon ratios: it is possible that this correction should only be applied to southern-hemisphere sites south of the tropics.

Table 1. List of dates used in this analysis, and associated information.

	X	a (s.e.)	b (s.e.)	p	r^2 (adj.)
All sites					
Cal. BP	geodesic	13666 ⊕484	-0.10 ⊕0.06	0.13 n.s.	0.08
Cal. BP	road map	13543 ⊕530	-0.05 ⊕0.05	0.30 n.s.	0.01
Excluding MV 2:					
Cal. BP	geodesic	13799 ⊕368	-0.13 ⊕0.05	0.02*	0.28
Cal. BP	road map	13721 ⊕408	-0.08 ⊕0.04	0.05 n.s.	0.20

Table 2. Results of weighted least-squares regression analysis of the relationships between dates and distance from El Paso. Equations are of the form Y = a + bX. Weights used were the square of half the one-sigma range. The Table shows two sets of results, with Monte Verde 2 – the principal outlier – either included or excluded. * = model significant at 95% level. From Steele, Gamble and Sluckin (in press).

Open Productive Habitat Category	Intermediate/Mixed Types
Includes:	Includes:
Savanna	Savanna/Forest Mosaic
Grassland	Moist Tundra
Dry Steppe	Desert
Mountain	Semi-Desert
	Scrub
Closed Forest Habitat Category	
Includes:	Sea and Ice
Forest	Includes:
Tropical Rainforest	Sea
Temperate Forest	Ice

Table 3. Broad paleovegetation categories into which the original types were assimilated for this analysis.

Bibliography

Adams, J M, 1995 *Preliminary Vegetation Maps of North and South America, 15,000-9,000 b.p.* (Project report for NERC.)

Bennett, W C and Bird J B, 1965 *Andean Culture History* (2nd edition), London

Borrero, L A, 1996 The Pleistocene-Holocene transition in southern South America, in Straus, L G, Eriksen, B V, Erlandson, J M and Yesner, D R (eds.), *Humans at the End of the Ice Age: the Archaeology of the Pleistocene-Holocene Transition*, 339-354, New York, Plenum Press

Glass, C A, Steele, J and Wheatley, D W, 1997 *Human colonization of South America: Regression Models*, Department of Mathematics, OR Preprint Series No. OR89 (in prep.)

Lundqvist, J and Saarnisto M, 1995 Summary of Project IGCP-253, *Quaternary International*, 28, 9-18

Lynch, T F, 1983 The Paleo-Indians, in Jennings, J (ed.), *Ancient South America*, 87-137, San Francisco, W.H. Freeman and Co.

Roosevelt, A C, Lima da Costa, M, Lopes Machado, C, Michab, M, Mercier, N, Valladas, N, Feathers, J, Barnett, W, Imazio da Silveira, M, Henderson, A, Sliva, J, Chernoff, B, Reese, D S, Holman, J A , Toth, N and Schick K, 1996 Paleoindian cave dwellers in the Amazon: the peopling of the Americas, *Science*, 272, 373-384

Steele, J, Gamble C and Sluckin T, in press Estimating the rate of Paleo-Indian expansion into South America, in Nicholson, R and O'Connor, T (eds.), *Proceedings of the 1995 AEA*, Oxford, Oxbow Books

Steele, J, Sluckin, T, Denholm, D and Gamble C, 1996 Simulating hunter-gatherer colonization of the Americas, *Analecta Praehistorica Leidensia*, 28, 223-227

Stuiver, M. and Reimer, P J, 1993 Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program, *Radiocarbon*, 35, 215-230

Whitley, D S and Dorn R I, 1993 New perspectives on the Clovis vs. Pre-Clovis controversy, *American Antiquity*, 58, 626-647

Contact details

Celia Glass[1], James Steele[2*], David Wheatley[2]

[1]Department of Mathematics,[2]Department of Archaeology

University of Southampton

Highfield, Southampton SO17 1BJ UK

*email: tjms@soton.ac.uk