Ideal Free Settlement of California's Northern Channel Islands

Bruce Winterhalder, 1 Douglas J. Kennett, 2 Mark Grote 3 and Jacob Bartruff 4

1Department of Anthropology and Graduate Group in Ecology, University of California, Davis, CA 95616; bwinterhalder@ucdavis.edu.  2Department of Anthropology, University of Oregon, Eugene, OR 97403; dkennett@oregon.uoregon.edu.  3Department of Anthropology, University of California, Davis, CA 95616; mngrote@ucdavis.edu.  4Department of Geography, University of Oregon, Eugene, OR 97403, jbartruf@uoregon.edu.

Corresponding Author: Winterhalder

Telephone: (530)-754-4770; Fax: (530)-752-8885

Last Saved: November 30, 2009

Filename: nCI_01_ver5.doc

Draft: Please do not cite without permission
Abstract

Prehistoric settlement on the Northern Channel Islands of southern California generally follows a pattern predicted by the population ecology model, the ideal free distribution (IFD). We establish this by comparing the abundant archaeological record of the Islands against a careful quantification of habitat suitability using areal photography, satellite imagery, and field studies. We assess watershed area, length of rocky intertidal zone, length of sandy beach for plank canoe pull-outs and area of off-shore kelp beds, for 46 coastal locations. A simple descriptive analysis supports key IFD predictions. Further analysis using a Gibbs Sampler and model selection allows us to reconstruct the native assessment of habitat that appears to underlie this process. The Gibbs Sampler mitigates the impact of missing data, censored variables, and uncertainty in radiocarbon dates; it allows us to predict where new settlements may yet be discovered. Theoretically, our results support a behavioral ecology interpretation of settlement history, human population expansion, and economic intensification in this region. They also demonstrate Bayesian analytical methods rarely used in settlement analysis but capable of making fuller use of the information available in archaeological datasets.

Keywords:

Human behavioral ecology; population ecology; settlement archaeology; California prehistory; Northern Channel Islands; Bayesian methods; Gibbs sampler; censored observations; measurement error.
Introduction and Problem

The Native Americans who moved from the mainland to settle the Northern Channel Islands of southern California knew the country they were settling well. By first recorded permanent settlement (8,000-7,000 cal. yrs. BP; at Tecolote Canyon, Santa Rosa Island) they had been exploring and seasonally exploiting the resources of the islands for as much as 5000 years (Erlandson et al. 2008; Kennett 2005; Kennett and Kennett 2008). Subsequent to that first settlement, further Island colonization of coastal sites at the mouths of major drainages drew on local knowledge gained through even more lengthy experience. Settlers were familiar with the suitability of the habitats they were about to occupy and, we might predict, they established residential sites in an orderly process of adaptive decision-making: Settle first in the most salubrious location. When, with growing exploitation or crowding, its resources were depressed and value declined to match the next-ranked locale, settlers establish a new settlement there. As population grew this process repeated, adding further new settlements in locations ordered by declining habitat suitability. In parallel, we expect overall quality of life in all occupied locales declined at a rate equalizing their marginal suitabilities due to reductions in the availability of or access to critical resources.

This process is neatly captured in a behavioral ecology model: the ideal free distribution (IFD, Fretwell and Lucas Jr 1970). Analysis of environmental and archaeological data from the Northern Channel Islands indicates that colonization there follows IFD predictions. To demonstrate this, we first describe the prehistory of the northern Channel Islands. We then introduce the IFD, with the goal of predicting how population growth, intensified use and
declining suitability generate a predictable pattern of settlement and habitat infilling. We
develop a comparative database of environmental and archaeological variables for 46 coastal
locations on the four Northern Channel Islands (Figure 1). A simple, preliminary analysis
substantiates the key IFD hypothesis and fosters the development of a more complete,
computationally-intensive evaluation of the environmental features that figured most
prominently in the settlement decisions of these prehistoric people. Thus, we substantiate our
basic behavioral ecology prediction and we introduce statistical methods – the Gibbs Sampler
and Bayesian analysis -- rarely used in archaeology but well-suited to the shortcomings and
opportunities that occur in datasets drawn from the prehistoric record.

Island Chumash Prehistory in Brief

At the time of European contact (AD 1542), the people living on the Northern Channel
Islands occupied relatively large coastal villages governed by chiefs. They were heavily
dependent upon fishing, produced a variety of trade items and participated in an extensive inter-
regional exchange network (Johnson 1982; Johnson 1988; Johnson 1993; Kennett 2005; King
1976). They spoke a dialect of the Chumash language (Cruzeño) distinct from the related
languages on the mainland coast and interior. Chumash is an ancient language with no affinities
with other languages in California (Klar et al. 1999). Although mainland and island languages
were not mutually intelligible, there is strong evidence for intermarriage throughout Chumash
territory, and for exchange of resources and ideas within the region and beyond (Arnold 1995;

During the early contact period, population densities in the Santa Barbara Channel region
were some of the highest in California (Moratto 1984: 2) and among the highest for hunter-
gatherers worldwide (Kelly 1995). The largest Chumash populations were concentrated on the mainland coast, but significant numbers of people also lived offshore, an estimated three thousand on Santa Cruz, Santa Rosa and San Miguel (Johnson 1982). Chumash informants in the late 19th and early 20th century named twenty two villages on these three islands (Johnson 1982; Johnson 1993). Locational information for many of these villages is clear and historic artifacts substantiate their existence (Arnold 1990a; Johnson 1982; Johnson 1993). Baptismal records indicate that intermarriage among these island communities and mainland villages across the Santa Barbara Channel was extensive (Johnson 1988). Archaeological evidence for the relative permanence of these villages comes from the large size and depth of midden deposits, the presence of substantial domestic features (e.g., house depressions and cemeteries) and diverse faunal and artifact assemblages (Kennett 2005).

Human occupation of the Northern Channel Islands extends back into the late Pleistocene (Erlandson et al. 1996; Erlandson et al. 2007; Johnson et al. 2000; Kennett et al. 2008). A partial skeleton, Arlington Man, buried under 11 meters of sediment, provides the earliest evidence for a human presence. It dates to between 13,000 and 12,900 cal. yrs. BP (Johnson et al. 2000). Little is known about the subsistence and settlement activities of this individual because the archaeological record on the islands and adjacent mainland is essentially silent during this early interval, possibly because post-glacial sea level rise has obscured or erased the record (Kennett et al. 2008). Regardless, the use of watercraft and by extension an inferred maritime lifeway is suggested by the presence of Arlington Man on these offshore islands (Erlandson et al. 2008).

Although the early date of Arlington Man is intriguing, persistent use of the Islands currently does not register in the record for another 600-800 years, at the end of the Younger Dryas cold interval (~12,200 cal. yrs. BP, Kennett et al. 2008). The earliest archaeological
evidence for a human presence occurs on the westernmost island of San Miguel, at Daisy Cave and Cardwell Bluff (Erlandson et al. 2008). The earliest deposits at Daisy Cave (~11,200 cal. yrs. BP) contain only a small number of stone flakes (chert and silicous shale) in association with charcoal and the remains of shellfish from the nearby rocky intertidal zone (Erlandson et al. 1996), evidence for a visit of short duration. Cardwell Bluffs, on eastern San Miguel, is a low density shell midden dating to 12,200 cal. yrs. BP, that appears to be associated with projectile points, stone crescents and other stone artifacts (Erlandson et al. 2008).

The oldest occupational level at Daisy Cave sits at the base of a well-stratified sequence of material, suggesting periodic short term visits to the cave that continue into the early Holocene and later. Early Holocene deposits at Daisy Cave are substantial and contain bone fish gorges, sea grass cordage (including a sandal fragment), stone projectile points and other cultural material associated with a diverse faunal assemblage (fish, shellfish, sea mammal bone). Persistent and more intensive use of Daisy Cave parallels increases in the formation of low-density shell middens on the outer islands of Santa Rosa and San Miguel during the Early Holocene. One of the open questions is whether these ephemeral shell middens represent periodic visitation and resource exploitation by people living more permanently on the mainland, or a resident population based somewhere on the islands (Kennett 2005). If more permanent settlements existed on the islands they remain undiscovered, perhaps due to post-Glacial sea level rise (Erlandson et al. 2008).

The first evidence for persistent residential bases occurs on the north coast of Santa Rosa Island near the mouth of Tecolote Canyon (SRI-3). A large residential midden and associated cemetery have been identified at this location (Orr 1968) and radiocarbon dated to between 8,000 and 7,000 years ago (Erlandson 1994), the time when sea level was beginning to stabilize and the
Winterhalder, Kennett, Grote and Bartruff, the IFD and the NCI,

Page 7 of 66

costline started to take its more modern form (Kennett et al. 2008). Settlement continuity after this time at this location is suggested by continued use of the cemetery into the Middle Holocene (~5,000 years ago) and the expansion of middle and late Holocene settlements along the coast to the west (Skull Gulch, SRI-2) and into the mouth of nearby Arlington Canyon to the east (SRI-4, -5). Evidence currently available points to the primacy of the Tecolote-Arlington region as an early node of island settlement and social life.

Increases in the number of archaeological sites after 7,500 years ago point to a substantial amount of demographic expansion. Primary nodes of settlement extended to several locations along the north and east coasts of Santa Rosa Island (SRI-41, SRI-116, SRI-187) and the southern coast of Santa Cruz (SCrI-333, SCrI-109, Glassow et al. 2008; King 1990; Wilcoxon 1993). Substantial middens in the interiors of the islands dating to the Middle Holocene suggest periodic and likely seasonal residential movement to interior locations for the purpose of collecting and processing plant resources (Kennett 2005; Kennett and Clifford 2004; Perry 2003). Midden constituent data indicate a strong focus on organisms from rocky intertidal habitats, but isotopic evidence suggests that higher trophic level organisms (sea-mammals and fish) were also being pursued and consumed (Goldberg 1993). Although a large number of Middle Holocene archaeological sites have been documented on the islands, only a relatively small number of substantial coastal villages are known. Other sites are the residues of more logistical subsistence activities based in large foraging ranges and perhaps periodic and short term residential mobility (Kennett 2005).

Major changes in subsistence and settlement occurred on the Northern Channel Islands during the Late Holocene (3,000-200 cal. yrs. BP). Populations increased substantially and the number of more permanent residential sites expanded around the perimeters of the three largest
islands, particularly after 1,300 cal. yrs. BP. Periodic visitation to the smaller rocky islets of Anacapa were also more frequent in the late Holocene (Rick 2006). It is during this time that the patterns of settled life, intensive fishing and extensive maritime trade documented at historic contact emerged (Kennett 2005; Kennett et al. 2008; Walker and DeNiro 1986). Skeletal material from Late Holocene cemeteries indicates that demographic expansion paralleled decreases in stature and increases in health stress indicators (e.g., cribra orbitalia, periosteal lesions, Lambert 1994; Lambert 1997). These demographic changes also paralleled increases in diet breadth (Braje 2007); a greater focus on fishing and the development of new maritime technologies (Kennett and Kennett 2000), heightened production of trade items (Arnold 1992), increases in interpersonal violence (Lambert 1994) and the emergence of social hierarchies (Arnold 1992; Kennett et al. 2009).

The Ideal Free Distribution

The IFD assumes a dispersive organism making the decision to settle in one of two or more habitats differing in their provision of living sites, resources and exposure to hazards or, in the aggregate, in their suitabilities. The organism makes the best choice available. We further assume that the suitability of each habitat is density dependent. Suitability can decline with increasing density of con-specifics, exhibiting negative density dependence. It can increase over low ranges of density, exhibiting positive density dependence, a phenomenon known as the Allee effect. Organisms locate or relocate until there is no advantage to further movement. Whether spread over one or more habitats, all organisms then experience the same marginal suitability. This establishes a stable equilibrium – the IFD – in which none has any further incentive to migrate. The IFD represents a seemingly simple process that leads to unexpected and interesting
temporal and spatial results that can be tested against the archaeological record (Kennett et al. 2006b).

We define a habitat as a bounded area of sufficient size for at least temporary settlement, with a particular suitability. It differs in suitability from other such regions. Suitability is measured in terms of a typical organism's survival and reproduction (fitness), and depends on such environmental features as food, shelter and predators, as well as interactions with conspecifics. Basic suitability (Fretwell 1972: 83) is measured by the fitness of the first occupant of a habitat. Subsequent occupants come from two sources: migrants arriving from another region or internal growth from among the habitats under consideration, immigration and reallocation, respectively.

The IFD relies on these simplifying assumptions: (a) habitats generate the resources and other conditions that establish their suitabilities at a constant rate; (b) suitability can be measured on a common scale by characteristics that are basic to the survival and reproduction of the organism and common to all habitats being considered; (c) this measure establishes a unique ordering of suitabilities across habitats; (d) the function representing the relationship between suitability and density for a particular habit is well behaved; at minimum, it is monotonically increasing or decreasing; (e) all individuals in the population have similar needs and abilities; (f) individuals make the best, constrained-optimization habitat choice (the "ideal" of the IFD), meaning that they have the information and cognitive capacity to appraise and compare relative suitabilities; and (g) there are no impediments to an individual acting on a decision to enter and use a habitat. Movement is unconstrained and organisms take up residence on an equal basis with existing residents (the "free" of the IFD).

As with human behavioral ecology models generally, it takes experience and judgment to
determine how stringently to interpret violations of these assumptions. Violations may or may not vitiate use of the model. For instance, assumption (a) does not preclude use of the model in environments affected by seasonality or even stochastic aspects of resource yield if, for instance, long-term averages over normal variability are the appropriate basis for analysis. Likewise, in many cases, mathematical variants of the IFD have assessed the impact of changing specific assumptions, for instance by examining the effects of unequal competitive abilities (relevant to assumption e), perceptual constraints (assumption f), and travel costs between habitats (assumption g) (review in Tregenza 1995). Tests of a model's predictions are in part tests of its assumptions (Winterhalder 2002).

The IFD adopts an individual-based, decision-making view of population-level phenomena such as colonization, habitat filling and subsistence intensification. It brings together individual behavior and broad ecological and sociological processes. It allows for the presence and density of con-specifics -- benefactors and/or competitors – by treating them as a part of what makes a habitat desirable or undesirable. It incorporates population effects through the manner in which suitability is enhanced or depressed by changes in con-specific density. For these reasons it should be useful to anthropologists and archaeologists interested in habitat selection and population ecology in relation to temporal trends in settlement patterns, range expansion, dispersal and colonization, and subsistence intensification.

Two classes of competitive mechanisms underlie formal representations of the IFD model (Sutherland 1983; Tregenza 1995). In *exploitation competition* each organism is viewed as consuming and thus subtracting a portion of a continuous resource stream. Following Tregenza (1994), if we define an individual payoff \( w_i \) for a population of \( n_i \) competitors sharing habitat \( i \), then at equilibrium the density specific payoff will be constant across habitats: \( w_i(n_i) = \)
c. If $q_i$ is the total input to habitat $i$, then: $w_i(n_i) = c = q_i/n_i$. The payoff is the resource stream divided among its users. Rearranging terms gives, $n_i = q_i/c$, known as the habitat matching rule. At equilibrium, habitat specific densities ($n_i$) are proportional to the habitat specific inputs ($q_i$), across all occupied habitats.

Competitors can also reduce each other’s fitness indirectly, apart from the direct exploitation of resources. *Interference competition* can result from territoriality, wasteful contests, declines in foraging efficiency due to search path overlap, general disturbance which increases prey wariness, or density effects that differentially attract parasites or predators. All of these can make successful use of a habitat more costly as a function of con-specific density, even though con-specifics may not reduce the resource stream itself. Tregenza (1995) building on Sutherland (1983) suggests interference or resource waste can be added to the exploitation model through an exponent modifying the variable $n$. In this model, $w_i(n_i) = q_i/n_i^f$, with $f$ scaled from 1 to infinity. If $f = 1$ we get exploitation competition without interference. Higher values of $f$ signal greater waste through the interference effects of con-specifics on each other’s foraging effectiveness.

Exploitation competition presumes that resource availability is density dependent, but not the difficulty or costs associated with harvest; interference assumes the reverse, and allows for a broader range of interactions than those concerned with subsistence.

Greene and Stamps (2001) propose a simple means of implementing the IFD, using a quadratic equation for the suitability of habitat $i$:

$$S_i(n_i) = Q_i - B_i(n_i - M_i)^2$$

(Eq. 1)

in which the variable $n_i \geq 0$ is the density of con-specifics in habitat $i$; $M_i \geq 0$ is a parameter determining the density of con-specifics at which the suitability in habitat $i$ is maximized; and $Q_i$,
and $B_i$ are scaling parameters. Suitabilities $S_i$ are scaled relative to each other across habitats in a context-specific way (Greene and Stamps suggest a 0 to 1 scaling). We denote the basic suitability of habitat $i$ as $S_i^* = S_i(1)$. And, we assume there is a suitability threshold $S_{\text{min}}$, common to all habitats, indicating con-specific densities which are unsupportable. Thus if $S_i^* \leq S_{\text{min}}$, habitat $i$ remains unpopulated, as it would not support the first individual to arrive. Finally, we expect suitabilities to exhibit negative density dependence at large values of $n_i$, thus we assume $B_i > 0$.

Key features of Eq. 1 are evident upon inspection. If $M_i = 0$, then suitability declines monotonically as a function of con-specific density. $M_i = 0$ gives us the basic, negative density dependence form of the IFD. If $M_i > 0$, then suitability has an intermediate peak at $n_i = M_i$. This generates positive density dependence at low densities ($<M_i$) and thus an Allee effect. $Q_i$ controls the height of the curve at its maximum; $M_i$ controls the position of the apex; and, $B_i$ controls the rate at which the curve falls to the side(s) of that apex. Larger values for $B_i$ accelerate or steepen the curve’s rise and fall.

Negative Density Dependence

We have built the Greene and Stamps (2001) formula into a computer program that populates habitats according to the assumptions of the IFD (Figure 2). Looking initially at the top panel, the first individual in the population elects to settle in the best habitat, $a$. Subsequent newcomers also elect to occupy habitat $a$, until the marginal suitability in habitat $a$ drops to that of habitat $b$ (arrow #1). From this point, further endogenous population growth or immigration is divided among $a$ and $b$ in a pattern that equalizes their marginal suitabilities. Equal marginal suitability is an equilibrium property because no individual has an incentive to relocate. A
similar response occurs at arrow #2, with new individuals now distributed over three habitats.

Population distribution among habitats as a function of overall population size is depicted in the lower panel. Individuals first colonize and begin to fill habitat \( a \). At arrow #1 further growth is divided between habitats \( a \) and \( b \), each of which fills at a slower rate. Subsequently (arrow #2), some individuals begin to distribute to habitat \( c \) while others continue to fill habitats \( a \) and \( b \). Although lowest in basic suitability, habitat \( c \) ends up with the largest population, experiencing faster growth over a longer period because its suitability in this example is significantly less sensitive to negative density effects. Total population size stops growing when all habitats reach their saturation densities (where in this example, \( S_{\text{min}} = 0 \)).

From IFD graphics like this we can generate structural (or qualitative) predictions relating habitat suitability to population distribution and dynamics. If our data are sufficiently precise, we also can generate quantitative variants of these predictions:

(i) As total population grows, habitats will be settled in order of decreasing basic suitability \( S_i^* \);

(ii) This process is cumulative in that high-ranking habitats will not be abandoned or see their numbers diminish as lesser ranking ones are occupied (provided there are no Allee effects; see below);

(iii) The habitat ranked second will not be settled until its basic suitability is matched by the (declining) suitability of the first-ranked habitat. This generalizes to habitats ranked \( r + 1 \) relative to \( r \);  

(iv) If more than one habitat is occupied, the relative numbers in each of them will be proportional to their relative suitabilities; and,

(v) Finally, a wide variety of more specific predictions are possible, depending on
specific configurations of the suitability curves relative to one another. For instance, as we have noted, a habitat unspectacular in terms of basic suitability may fill more rapidly and expand to a larger size than its higher-ranked alternatives (*Figure 2*).

Each of these predictions suggests a means of testing whether or not human settlement patterns conform to the IFD.

**Positive Density Dependence**

In his 1938 book, *The Social Life of Animals*, W. C. Allee argued that organisms at very low densities may suffer disproportionately from the problems of finding mates and resources. They may be less able to avoid or deter predators. Consequently, when population is sparse, increasing density might actually make a habitat *more* suitable to its residents. This vision of positive density dependence at low densities has since become known as the *Allee effect*; it has quite interesting implications for the IFD.

Graphically, the Allee effect is represented as a segment of positive slope on the low density portion of the curve representing suitability (*Figure 3*). With humans it is easy to imagine circumstances that might generate this pattern. Especially low density may make it difficult to locate a mate of an appropriate sex, age and kinship category. It may induce other problems related to demographic stochasticity. It may impede benefits associated with mutualistic or cooperative activities. For instance, newcomers to a habitat may more quickly and reliably gain subsistence and survival skills if they come into contact with experienced residents. Actually, for socially obligate species like humans and most primates, Allee effects may operate over much of the normal density range (Courchamp et al. 1999). Opportunities for Allee effects expand significantly in humans because of our ability to implement socially organized and linguistically shared technologies that produce economies of scale in the capture, defense and use
of resources. Range expansion, colonization episodes, and recovery from stochastic population declines all provide a context in which Allee effects may be important.

The general interpretation of Figure 3 is similar to that described in the caption for Figure 2. The key difference is the Allee effect. Habitats still are ranked by their suitability for the first occupant, but peak suitability for habitats b and c occurs after a period of influx and local population growth. As a result (lower panel), Allee effects typically cause abrupt population relocations. Individuals do not begin relocating to habitat b until the marginal suitability in habitat a matches the initial suitability in b (shaded arrow #1). But, by improving suitability in habitat b, the initial occupants quickly draw a small pool of follow-up migrants, from a or external sources. Although it is not shown here, if Allee effects caused habitat b suitability to increase significantly above that of any segment of the curve for habitat a, this relocation might temporarily and entirely empty out habitat a. Habitat c is characterized by a more prolonged Allee effect of greater magnitude, one causing population to decline in habitats a and b over a lengthy range of total population growth. Allee effects are dramatic illustrations of the observation that marginal quantitative changes in one parameter (here population size) may result in large qualitative changes in another (e.g., rapid emptying of one habitat and redistribution of its inhabitants into another). In Fretwell's terms, with an Allee effect, “a remarkable event may occur” (1972: 90-91).

The Allee Effect gives us yet another IFD prediction:

(vi) If second or lower ranked habitats are characterized by an Allee effect, the first individual to settle there will attract a temporary exodus of occupants from higher ranked habitats. Complete abandonment of the higher ranked location is possible.

Creative use of the IFD turns on imagining how a factor of interest, the independent
variable, will affect the number, shape or relative positioning of suitability curves. Climate change might elevate or depress suitability of all or some habitats. A technological development might change the shape of the curve as well as its relative position. Suitability curves of habitats highly sensitive to degradation under exploitation will slope steeply downward as a function of human density. Those of more resilient habitats will have a shallower form. Economies of scale and technological ability to implement efficient, group-dependent technologies (net hunting, weirs, terrace systems, irrigation networks) will produce ascending suitability curves over some range of densities. Once habitats are characterized by their unique suitability functions, it is fairly easy to derive predictions linking population growth, intensification and habitat enhancement or degradation to predictions about the population ecology of migration, settlement, and distribution.

In the analysis to follow, we will present empirical evidence for the first two of these IFD predictions, hypotheses (i) and (ii): The dates of earliest settlement for 29 locations on the Northern Channel Islands are well-predicted by a model ranking their basic suitabilities, and high-ranking habitats typically are occupied more-or-less continuously as new settlements are added in less suitable locations.

Environment and Comparative Archaeology

Interpretation of prehistory using the IFD requires that we assess the temporal sequence and other features of settlement and social development in light of the suitabilities of the localities being occupied. Although archaeological datasets can be rich in time depth, they rarely provide the fine-grained demographic information (e.g., a time-series of population densities in each habitat) needed to precisely fit a model such as the IFD (e.g., Eq. 1). However, given
archaeological and geomorphological data of sufficient detail, it is possible to fit a model which relates settlement date to basic suitability, and basic suitability to environmental variables, and therefore to test the IFD by this means.

Environmental Analysis and Data

General archaeological and ethnohistoric information on the subsistence practices of populations living on the Northern Channel Islands (Kennett 2005) and adjacent mainland (Erlandson 1994; Gamble 2008) directs our attention to four environmental features: (i) watershed or drainage area, for variety and yield of terrestrial resources and security of a perennial fresh water supply; extent of shoreline with (ii) resource-rich, rocky intertidal zone, (iii) sandy beach, for canoe haul-out; and, (iv) off-shore kelp forest beds, with their concentration and productivity of marine resources (mammals, fish and shellfish). Shoreline may also take the form of rocky-sea cliffs, which provide virtually no economic benefits. Rocky-sea cliffs are important only in that their presence diminishes the extent of rocky intertidal and sandy beach. We do not include them in our analysis.

(i) Watershed or drainage area (km$^2$). Because stream flow data is unavailable, we use watershed area as a proxy measure for water availability at the mouth of each drainage analyzed. Most of the large historic Island Chumash villages were positioned on larger drainages. This is not surprising since access to an adequate, year-around water source is essential on these dry islands. In general the largest watersheds occur on the larger islands of Santa Rosa and Santa Cruz. Smaller watersheds are present on the westernmost island of San Miguel, but major drainages do not occur on the small rocky islets composing Anacapa. Most of these large drainages contain robust perennial streams, but other geological and hydrological factors also contribute to stream flow intensity. Watershed area also measures the diversity and
abundance of terrestrial resources. Drainages provide shelter from the prevailing northwesterly wind and microhabitats for economically valuable trees and other plant species, such as island oaks that favor sheltered valley locations or sage scrub that favors south and east-facing slopes of drainages (Philbrick 1980).

Watersheds were demarcated using ArcMap and a digital elevation model for the Islands, then analyzed with ESRI’s grid function for hydrological features. High (1.5 m) resolution color, digital aerial photographs were used to confirm drainage evaluations. Very small drainages and springs and specific localized plant communities (e.g., Torrey Pines, Eastern Santa Rosa, Biondi et al. 1997) were not included in the analysis.

(ii) Rocky intertidal (km). The rocky intertidal zone and associated tide pools are important for their abundant and easily gathered marine organisms. California mussels (*Mytilus californianus*) are the most abundant and easily accessible of these species and are ubiquitous in prehistoric midden deposits extending back into the terminal Pleistocene and Early Holocene (Erlandson et al. 1996; Kennett et al. 2008). Large mussels can easily be plucked out of the intertidal zone or small groups can be stripped out en masse to produce higher return rates (Jones and Richman 1995). Two large species of abalone (*Haliotis*) are also found on rocky substrates in the lower intertidal to subtidal zones (Braje et al. 2007).

Coastal survey and color digital aerial photographic images were used to measure the linear extent of this habitat within a two-kilometer radius of each drainage outlet analyzed in this study.

(iii) Sandy beach (km). Inhabitants first explored and colonized the islands by boat from the mainland. They later engaged in a significant amount of inter-island and island-mainland trade using large plank canoes. Sandy beach was essential for safe canoe landing and
haul-out. particularly for the use of the plank canoe after 1500 years ago (Arnold 1995; Gamble 2002). The small gastropod *Olivella biplicata* is also found along sandy beach habitats and this species was harvested increasingly en masse for its shell to produce beads after 1500 years ago (Arnold and Munns 1994; Kennett 2005). The linear extent of this environmental feature was measured as for rocky intertidal.

(iv) Kelp forest (km²). The relatively shallow water and rocky reef substrates that fringe the Islands support scattered patches of dense kelp forest (*Macrocystis pyrifera*, *Nereocystis spp.*). This marine habitat provides concentrated food and shelter for a rich community of intertebrates, fish and sea mammals. Over 125 different fish species are found in this sheltered environment (Love 1996) and several major mollusk species feed in this high nutrient zone, including the largest available abalone species (*Haliotis rufescens*) that grazes on the kelp itself. The species composition associated with these habitats varies and is heavily influenced by a major gradient in surface water temperatures and nutrients from east (warm, lower productivity) to west (cooler, high productivity). Kelp forests are generally denser and more extensive in the west. The extent of kelp forests in the 1980s was estimated through a combination of visual survey, color digital aerial photographs, which show the surface canopy of the kelp and the 25m isobath around each island, known to be the maximum water depth where kelp is viable (Kinlan et al. 2005).

*Figure 4* shows how these environmental features are assessed for a high and a low-ranking watershed.

Comparative Archaeological Data

The archaeological dataset is based on the collective archaeological and ethnohistoric work accomplished on these islands during the last century or so (Kennett 2005). These islands
are particularly well suited as a testing ground for the IFD model presented here because they contain the longest, best-preserved archaeological sequence available for study along the west coast of North America. Archaeological deposits are relatively undisturbed compared with the adjacent mainland, because the pocket gopher (*Thomomys spp.*) never colonized these islands (Erlandson 1984) and only modest development occurred historically. Archaeological survey and excavation on these islands extends back to the late 1800s. Many of the larger excavations of "village" sites and cemeteries occurred in the early 1900s and continued through the 1960s (Mills 1956 [1901]; Orr 1968; Rogers 1929). Subsequent reanalysis of museum collections coupled with direct radiocarbon dating of artifacts from these assemblages are important sources of data (Erlandson 1994; Kennett 1998; Kennett 2005; King 1990; Rick 2001; Rick et al. 2002). More recent surveys, site assessments, smaller-scale sampling and radiocarbon dating provide additional data regarding the character and age of archaeological sites on these islands (Braje 2007; Erlandson and Moss 1999; Glassow 1980; Kennett and Kennett 2000). A majority of these data have been summarized in Kennett (2005), but have been augmented by more recent studies (Braje 2007; Kennett et al. 2007; Kennett et al. 2008; Perry 2003; Rick 2004).

Table 1 provides a summary of the archaeological and chronological data associated with the known settlements at the mouth of each ranked drainage. The "earliest" settlement date was based on radiocarbon ages or diagnostic artifacts associated with archaeological materials inferred to represent more permanent settlement. Direct radiocarbon dates on house floors or cemeteries were preferred, but often not available. Diverse faunal and artifact assemblages were also used in conjunction with estimates of site size and thickness/density of deposits. Justifications for these assessments are provided in Kennett (1998; 2005). Data quality and chronology are considered to be relatively high for sites dating to between 8,000 and 3,000 cal.
yrs. BP, and for those dating to between 1,500 and 250 cal. yrs. BP. Sites in the 3,000 to 1,500 cal. yrs. BP range are based on moderately diverse faunal and artifact assemblages and the size and density of midden deposits. All assessments are considered to be provisional until large-scale excavations are carried out at these sites, perhaps as a test of the IFD model. Candidates for permanent village settlements dating prior to 8,000 cal. yrs. BP have not been identified and are likely under water. Our IFD model could be used to help constrain a search for underwater archaeological sites.

Assumptions Engaged in Using this Data to Model the IFD

Use of the mouth of each drainage as the central node of settlement is based on the assumptions that 1) these would generally be the most desirable village locations compared with other locations along the coast or in the interior, and 2) people living in these villages had access to and controlled interior areas defined roughly by each watershed. Archaeological surveys indicate that sites of different types (shell middens, lithic scatters, etc.) occur in a wide range of contexts, but that residential sites tend to be positioned at the mouths of the largest drainages (Kennett 2005). These archaeological sites consist of laterally extensive and/or deep deposits. The remnants of houses (e.g., house depressions) occur at some of these sites and past excavations have revealed cemeteries suggesting generational continuity in settlement. Greater faunal and tool diversity indicate a wider range of activities compared to other locations. Given the maritime focus of island residents (substantiated by shell midden deposits), it makes sense from a central-place foraging perspective that people would be more tethered to coastal locations. Settlement locations near the mouths of drainages would provide immediate access and control over water, intertidal areas in close proximity to the village, and kelp beds immediately off-shore. Beaches in the vicinity could also be used as launch sites and storage
locations for boats to access kelp beds, offshore fisheries, and more distant sea mammal or bird rookeries and to maintain contact with friends, family, and trade partners in other villages on other islands and the adjacent mainland.

The assumption that the communities at the mouths of drainages controlled access to resources in the entire watershed is more tenuous and conjectural. Although major village settlements were generally positioned on the coast, there are some exceptions (e.g., CA-SRI-147, Braje 2007), and the intensity and character of interior resource use changed during the last 8,000 years. A large number of substantial Middle Holocene (7,500 to 3,000 years) midden deposits have been identified in the interiors of the larger islands of Santa Rosa and Santa Cruz (Kennett 2005; Kennett and Clifford 2004; Perry 2003). These sites tend to be positioned on ridges or hilltops surrounding major drainages. Midden soil formation, artifact assemblages and faunal assemblages, indicate the transport of marine foods to these interior locations (e.g., shellfish). Isotope seasonality data suggest that they were occupied periodically by smaller groups of people, perhaps task groups or individual families from larger coastal villages harvesting plant foods episodically. Kennett and Clifford (2004) have suggested that the strategic position of these interior sites may signal a flexible system of land tenure associated with periodic economic defensibility of interior plant foods. More rigid territories on the islands after 3,000 years ago are suggested by the lack of inter-visibility between coastal settlements and the strategic placement of burials in older interior settlements and in caves. This parallels an increasing focus on maritime resources (particularly fish) and increased trade connections with the mainland coast, possibly to acquire carbohydrate-rich foods (e.g., acorns, Arnold 1992).

All regional datasets are imperfect, and it is possible that further shortcomings are introduced by our assessment of the published record (Kennett 2005). We address this in two
ways: (a) by being as specific as possible about our sources, assumptions, choices of interpretation, and known qualifications; and (b) by adopting statistical methodologies that explicitly recognize the uncertainties in key types of data. Landscape transformations obscure the record. Earlier sites tend to be buried by natural or anthropogenic processes or, in the case of the Channel Islands, they may be under water. Identifying the "earliest" occupation at any location is a biased estimate because it actually records a date sometime after settlement. Likewise, consistent standards of evidence for characterizing permanent settlement (e.g., house floors, diverse faunal and artifact assemblages, seasonality data, cemeteries) are also difficult to come by without large scale and better dated excavations at all sites. The settlement data from 3,000 to 1,500 cal. yrs. BP is particularly tenuous because this is a poorly studied period and artifacts tend to be rare in midden deposits.

Data Analysis and Results

A Preliminary (Statistically Naive) Analysis

A preliminary semi-quantitative analysis of the Northern Channel Island dataset in the context of the IFD and the emergence of institutionalized social hierarchies focused on the late Holocene (3000 – 200 BP) archaeological and ethnohistoric record (Kennett et al. 2009). That paper examined an acceleration of social and technological change on the Islands that began between 4,000 and 3,000 years ago, arguing that it occurred concurrent with habitat in-fill, settlement of more and more marginal drainages, and territorial circumscription. The analysis took an intuitive approach. Each of the 46 drainages surveyed was ranked for each of the four environmental features. These rankings were then combined into a weighted score [50% watershed area; 30% rocky intertidal; 15% kelp forest and 5% for sandy beach to generate a
single, linear ranking of presumed basic suitability for each watershed (Kennett et al. 2009 see Table 20.1, “Rank”). The weightings are estimates based on qualitative impressions of the Chumash literature and knowledge about key environmental characteristics likely shaping adaptive decision making (e.g., water availability). Weighted scores were then grouped into quartiles. Similarly, late Holocene settlement was examined over two periods: sites first settled 3,000 – 1,500 cal. yrs. BP, and those settled 1,500 – 200 BP. The analysis consisted of showing that village settlements were confined predominantly to the first- and second-ranked locations in the earlier period; they expanded into drainages in the third, much less suitable quartile after 1,500 cal. yrs. BP. Few fourth-ranked locations were settled. The observed pattern supported the prediction that drainages were occupied in descending order of basic suitability and, with other information on social hierarchy, it confirmed that institutionalized social stratification emerged late in the sequence as viable settlement locations became saturated, social circumscription increased (Carneiro 1970) and environmental suitability declined throughout the Islands.

In Figure 5 we replicate this analysis, extending it to the history of known village settlement back to ~8,000 cal. BP, at the time when sea-level was stabilizing and we see the first settled villages appearing on these islands. The results are the same, a semi-quantitative pattern largely confirming hypotheses (i) and (ii). Settlement begins in watersheds ranked in the highest quartile of habitat suitabilities. With several exceptions, settled locations expand as predicted by the IFD throughout the Holocene to drainages of lesser and lesser suitability. There is considerable continuity over time in the occupation of the higher ranked watersheds. Once settled, locations tend to remain settled. The pattern observed is consistent with IFD predictions. The seven sites listed as an “Unranked Locale,” those not associated with one of the 46
watersheds for which environmental parameters were assessed, likewise tend to be late in settlement history, suggesting that they may be associated with increasing economic exchange and conflict.

A model for settlement time using the Gibbs Sampler

Encouraged by the general IFD pattern in the data, we undertake a more detailed analysis. We focus on the response variable, the earliest known date or Year of Settlement (YS) cal. yrs. BP for each location, in order to examine hypothesis (i). Characteristics of this variable guided our analysis strategy. YS is determined by radiocarbon dates either directly or indirectly with artifact styles (Late Middle Period and Late Period sites) and therefore subject to the uncertainties of laboratory analysis and calibration to calendar years. And, YS in the Northern Channel Islands dataset includes missing or censored dates for some records or potential settlement locations. Ideally we seek a method for model-fitting and inference that can handle each of these characteristics, without discarding useful information or analytical potential. That is, we wish to acknowledge uncertainty in YS (the dependent variable), and also allow for the opportunities or partial information represented in missing and censored records, respectively. One usual strategy for dealing with censored observations-- simply removing them from the dataset -- is one we wish to avoid, because it discards potentially useful information and can lead to biased inferences and other problems of interpretation (Little and Rubin 2002).

To proceed, we undertake a computationally-intensive Bayesian approach, imbedding measurement error and data-augmentation methods in a Gibbs sampler-- an iterative, probabilistic algorithm well-suited to the current problem. Briefly, we use the Gibbs sampler to impute (probabilistically, and repeatedly) dependent variables observed with error, or not observed at all; additionally, for each imputed dataset, the Gibbs sampler updates a model of
locational suitabilities. Aggregating information over many imputed data sets makes for more robust inferences. Computational Bayesian methods for measurement error are described and demonstrated in Mallick and Gelfand (1996) and Gelfand et al. (1997); and Bayesian approaches specific to radiocarbon dating can be found in Buck et al. (1996). General descriptions of Gibbs sampling can be found in Geman and Geman (1984), Gelfand and Smith (1990), Gelfand (2000) and Casella and George (1992); Gibbs sampling combined with data augmentation for censored observations is described in Chib (1992), Holloway et al. (2004) and Gelman and Hill (2007, sect. 18.5).

Drawing on the flexibility of the Gibbs sampler, we incorporate into the analysis two classes of observations. The first class consists of locations (drainage mouths) with a village site having an oldest radiocarbon date, which we treat as an imprecise proxy for YS. Ignoring measurement error, these 29 cases taken alone would be amenable to conventional statistical analysis. However, we wish also to include a second class of observations: locations at which there has been sufficient archaeological investigation to eliminate with some confidence the possibility that they were ever settled (n = 9). We call these “unsettled” (U) locations. These locations are censored for YS. However, they are clearly informative about the relationship between environmental suitability and the process of settlement. It does not make sense to assign unsettled locations a numerical value (such as zero) for YS; we adopt instead the strategy described below. A third class of observations consists of locations surveyed for environmental variables, but not for archaeological sites (n = 8). We say these locations have an “unknown” YS. Because no information about settlement is available at the third class of locations, they are not informative about the relationship between suitability and timing of settlement, and are not used to fit the model. However we can use the fitted model to make predictions about settlement
at these locations.

Our sample of 46 drainages has 29 locations that are uncensored but imprecisely observed for settlement time; nine locations that are unsettled; and, eight additional locations we have classified as unknown. The model presented below was fitted using the 29 uncensored and nine unsettled locations, for a total of $N = 38$ observations.

In order to accommodate different scales of the environmental predictors, we applied the natural log transform to Drainage Area and Kelp Forest Area. Prior to log-transformation, we added 0.01 to all observations of Kelp Forest Area, as 6 locations have the value zero for this variable (these locations are not censored for Kelp Forest Area-- the value zero is observed). Scatter plots suggest that Drainage Area and Rocky Intertidal Length are likely to be informative about settlement timing (Figure 6). A model selection procedure (see Appendix) suggests that Sandy Beach Length should also be included. These are the predictors we use in our analysis.

A linear regression model for the basic suitability of the $i$th location is

$$S^*_i = \beta_0 + \beta_1 \log(\text{Drainage Area}_i) + \beta_2 \text{Rocky Intertidal Length}_i + \beta_3 \text{Sandy Beach Length}_i + \epsilon_i,$$

where $\epsilon_1, \ldots, \epsilon_N$ are independent Gaussian variables with mean zero and variance $\sigma^2$. The settlement time of location $i$ is connected to its basic suitability by the following threshold relationship:

If $S^*_i > S_{\text{min}}$, Year of Settlement$_i = \exp(S^*_i)$;

If $S^*_i \leq S_{\text{min}}$, location $i$ remains unsettled.  \hspace{1cm} (Eq. 3)

We undertake estimation of the parameters $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$, $\sigma^2$ and $S_{\text{min}}$ using the Gibbs sampler.
In Eqs. 2 and 3, the basic suitability $S_i^*$ mediates the relationship between environmental predictors and settlement time, and these relationships across the sample determine the scaling of $S_1^*, \ldots, S_N^*$ and $S_{\text{min}}$. Conventional IFD models view suitability as a function of con-specific density, and are not explicit about the relationship between basic suitability and settlement time. The exponential relationship between $S_i^*$ and settlement time specified in Eq. 3 is reasonable in our view, as it ensures that model predictions of Year of Settlement are positive, and consistent with the ordering of basic suitabilities produced by use of Eq. 2. The two-equation form of the model aids in conceptualizing the role of unsettled cases. Although these 9 cases do not have numerical values for Year of Settlement, under Eq. 2 they do have basic suitabilities. We treat the basic suitabilities of the unsettled cases as unobserved variables, imputing their values at each iteration of the Gibbs sampler with the help of Eq. 2. Details of our implementation are provided in an Appendix.

Table 2 gives parameter estimates for Eqs. 2 and 3, in the form of summary statistics over iterations of the Gibbs sampler. The effect of log Drainage Area (km$^2$) is positive, relatively large ($\beta_1 = 0.82$) and it has a 95% credibility interval (0.5 to 1.2) distant from zero. Rocky Intertidal Length (km) has a somewhat smaller, positive effect ($\beta_2 = 0.25$) and has a 95% credibility interval (0.04 to 0.5) that rests closer to but still does not overlap zero. Sandy Beach Length (km) has a yet smaller effect ($\beta_3 = -0.17$), in the opposite (negative) direction and a suitability interval (-0.4 to 0.07) that overlaps zero. As expected (see Figure 6), locations with large Drainage Area tend to have higher suitabilities, hence relatively early settlement times. The coefficients for Rocky Intertidal Length and Sandy Beach Length are of the same order of magnitude but have opposite signs, suggesting a trade-off favoring the availability of Rocky Intertidal zones over Sandy Beach. As the 95% credibility interval for Sandy Beach Length
includes zero, there is perhaps less certainty about the utility of this predictor than the others.

Tests of an alternative, null model (see Appendix), in which we randomly assign the existing array of unknown, unsettled and earliest date of village settlement (YS) to watersheds confirms the significance of the Table 2 results.

Figure 7 and Figure 8 illustrate the fit of the model to the sample. Figure 7 shows that settlement times (in calendar years BP) predicted by the model are in good agreement with settlement times based on radiocarbon dates alone, for the n=29 settled locations. The radiocarbon dates are incorporated into the model by a measurement error method, which allows for dating imprecision, but nonetheless ensures that predicted dates are more-or-less consistent with calibrated radiocarbon dates (see details in the Appendix). The consistency shown in Figure 7 is therefore not unexpected; however, it does confirm that the suitability model can be fitted, and model parameters estimated, without distorting the pattern of settlement times in the raw data. Figure 8 shows that the unsettled locations are well distinguished by the model parameter $S_{\text{min}}$, all having posterior mean suitabilities below the estimated threshold. The five settled locations depicted in Figure 8 are thought to be among the last to be occupied (with approximate settlement dates of 650-670 BP, based on artifact assemblages; see Table 1). The model suggests that these settled locations have relatively low suitabilities.

Figure 9 shows the full sample of $N = 38$ locations used to fit the model, along with 8 additional locations having unknown settlement times. We display the environmental covariates, along with suitabilities predicted by the model (with locations arrayed left to right in descending order of predicted suitability). A decreasing trend in Drainage Area is apparent from left to right, though it is interrupted by a few large unsettled locations. The threshold $S_{\text{min}}$ enforces a separation between the predicted suitabilities of settled and unsettled locations at each iteration.
of the Gibbs sampler, even though predictor values for the two groups may overlap. Thus we do not expect ranges of Drainage Area or the other predictors to discriminate perfectly between settled and unsettled locations. Locations having the value zero for Rocky Intertidal Length are noticeably common on the right, where predicted suitabilities are lower. The relationship between suitability and Sandy Beach Length is less clear: although the earliest settled locations (those with the highest suitabilities) tend to have little sandy beach, and the latest settled locations (those with lowest suitabilities) tend to have more sandy beach, there are anomalies. Notable among these are Canada Christi, Old Ranch Canyon and Unnamed Bee Rock, all settled early but having much Sandy Beach compared to Rocky Intertidal.

Figure 9 also places the locations having unknown settlement in context, with the bottom panel highlighting the predictive capabilities of the model for these cases. The 95% credibility intervals for the unknown locations are notably wider than those for settled locations, and modestly wider than those for unsettled locations; this reflects added uncertainty in prediction. Nonetheless we can say with some confidence that the unknown locations Willows Canyon, Alamos Canyon, Unnamed China Camp 1 and Wreck Canyon have relatively high suitabilities under the model, and may well have undiscovered settlement. By contrast, the four “unknown” locations having low predicted suitabilities (Unnamed Profile Point, Unnamed Ruby Rock, Unnamed Trident Cove and Diablo Canyon), are unlikely to have supported permanent human settlement.

The predicted suitabilities displayed in the bottom panel of Figure 9 are posterior means, obtained for each location by averaging the random suitabilities produced over iterations of the Gibbs sampler. These numerical predictions are distinct from those that would be obtained by simple use of the regression formula (Eq. 2), substituting estimates from Table 2 in place of
parameters. The two different approaches to prediction highlight a key difference between conventional and Bayesian statistical modeling. Posterior means give the preferred predictions here, as they incorporate aspects of the model—such as measurement error in dates—not captured by the regression formula alone. Posterior means and regression formula predictions will be relatively similar for unknown locations, where no additional information about dating or settlement constrains the suitabilities produced by the Gibbs sampler. Yet even for unknown locations, posterior means more faithfully incorporate all of the sources of uncertainty in the Northern Channel Islands dataset.

Discussion

Behavioral Ecology and the Ideal Free Distribution

Behavioral ecology models have established themselves in a wide variety of archaeological and ethnographic subject areas, from foraging to life history decisions (reviews in Winterhalder and Kennett 2006; Winterhalder and Smith 2000). However, applications drawing on the IFD have lagged behind other models of comparable generality and importance (e.g., diet breadth, polygyny threshold). Recent heuristic applications of the IFD focus on regional patterns of colonization and settlement in Oceania (Kennett et al. 2006b) and Malay-Australia (Allen and O'Connell 2008), adoption of agriculture in Spain (McClure et al. 2009), and spread of agriculture across Europe (Shennan 2007). In this paper we explore similar settlement processes but on a smaller and more detailed scale, and with a dataset amenable to quantified investigation of key IFD predictions.

A second aim of this paper is methodological. We seek to illustrate a significant shift in statistical analysis, from conventional inference and null hypothesis testing to computationally-
Winterhalder, Kennett, Grote and Bartruff, the IFD and the NCI, page Page 32 of 66

intensive methods, uniquely tailored to the observations and the questions asked about them. We do not argue the case here for a shift from hypothesis testing to model fitting; that has been done well elsewhere (e.g., Anderson et al. 2000; Towner and Luttbeg 2007). We do intend, along with Litton and Buck (1995), to demonstrate and encourage the use of Bayesian and other computationally-intensive methods for archaeological problems.

Assumptions and Confounding Factors

Our statistical analysis confirms that prehistoric settlement location by early residents of the Northern Channel Islands is consistent with the general pattern predicted by the IFD. Watershed area and length of rocky intertidal zone emerge as significant positive predictors of earliest date of settlement; sandy beach has a weaker, negative effect. A strong IFD signal emerges from this dataset despite multiple factors that, because they were ignored in our analysis, should act to obscure any underlying pattern. We describe potential confounding factors in this section.

To start with, the metrics we adopt for our four environmental measures – area for drainage and kelp beds; length for rocky intertidal; and sandy beach -- are at best very rough substitutes for the experienced and highly localized assessments of resource productivity that we presume were exercised by prehistoric peoples. One kilometer of rocky intertidal is not the same as the next; observations made on airphotos are not nearly as informative as those on the ground. One 10 km² drainage may yield abundant terrestrial resources and the next of similar size may be meager by comparison. Analysis by area ignores exposure, windward-leeward climatic factors, geological substrate, localized springs and variations in plant cover. The northern, coastal-facing side of Santa Rosa features a concentration of highly ranked drainages, whereas the northern coast of Santa Cruz is dominated by quite small, low-ranking drainages (Kennett et al. 2009).
By examining only four environmental features – one of them eliminated from the analysis by model selection -- we surely have neglected other localized factors important in the decisions of early residents. Early sites on these islands are sometimes found in caves or rockshelters (Erlandson et al. 1996; Kennett et al. 1997) and the distribution of these may be an important parameter for settlement as in other island contexts (Kennett et al. 2006a). In addition, permanent springs are not always tied to the largest drainages and could be mapped independently and included as an additional environmental parameter. For instance, the geology and hydrology of San Miguel Island favors springs on the north coast and partially explains the biased distribution of sites on that side of the island (Braje 2007; Kennett 2005). Sources of economically valuable stone (e.g., obsidian, chert), are known to influence human behavior and settlement. Exposed chert sources are well-known on eastern Santa Cruz Island (Arnold 1990b; Perry 2003) and other smaller sources of chert are known elsewhere (Erlandson et al. 1997). Also, concentrations of other types of subsistence resources are well known on the islands and include a sea mammal rookery on the western tip of San Miguel Island at Point Bennett (Rick et al. 2009; Walker et al. 2000) and an associated decreasing gradient from west to east in certain species during certain seasons (e.g., California sea lions, Northern Fur Seals and Elephant seals, Le Boeuf and Bonnel 1980). This is also the case with bird rookeries on San Miguel and the rocky islets of Anacapa (Rick 2004).

We likewise assume that the present-day configuration of these environmental qualities persisted backwards in time for some 8000 years. This may be roughly accurate for relative watershed size, it surely is problematic for features like rocky intertidal and kelp beds, dependent as they are on precise relationships between surface morphology and sea level. Productivity in these marine systems is also known to be spatially variable and fluctuate through time (Dayton et
al. 1992; Kennett and Kennett 2000; Kennett et al. 2007; Kennett et al. 2008; Ono et al. 1993; Tegner and Dayton 1987; Tegner and Dayton 1991). Sea-level had a major transformative effect on the coastline of these islands (Fairbanks 1989; Kennett et al. 2008; Porcasi et al. 1999). Changes in sea level continued after 8,000 years ago and would have influenced the distribution of kelp forests that only grow in a certain depth range (Kinlan et al. 2005). The infilling of small pocket estuaries around these islands during the Late Holocene reduced the availability of certain fish and shellfish species that once occurred in these highly productive habitats. The most notable of these estuaries is on the eastern end of Santa Rosa Island at the mouth of old ranch canyon (Rick et al. 2005; Rick et al. 2006). El Niño frequency is also known to have changed during the last 8,000 years (Kennett et al. 2007; Ramage 1986). Finally, we do not consider the potential for differential human impacts on resources (e.g., resource depression) or the associated trophic interactions and landscape modification by prehistoric inhabitants of the islands (Braje 2007; Erlandson et al. 2004; Rick and Erlandson 2009; Timbrook et al. 1982).

Our analysis also assumes that these resource features mattered in the same way and to the same degree through time, from relatively small groups of egalitarian hunter-gatherers focused heavily on shellfish coupled with terrestrial foods (e.g., grass and sage seeds), to centrally organized chiefdoms making greater and greater use of near and offshore marine resources and trade with mainland communities. This may help explain why kelp bed distribution does not appear to be a significant predictor of earliest settlement on these islands. In fact, kelp beds apparently had a significant impact on the economy only toward the later part of the sequence, because of changing or intensifying marine subsistence practices (Kennett 2005; Kennett and Kennett 2000). Similarly, beaches became much more important later in time (after 1,500 years ago) for hauling out and storing plank canoes (Arnold 1995; Gamble 2002). They
also became an important source of *Olivella biplicata* used to produce immense amounts of shell bead money, particularly after 800 cal. yrs. BP (Arnold and Munns 1994; King 1990). These patterns may help to explain the inverse relationship we detect between length of sandy beach and earliest year of settlement.

We also have assumed each new settlement is independent in the sense that it has an autonomous economy drawn only from the local resources of its watershed, adjacent coastline and offshore zone. This neglects the possibility that some locations were settled as economically dependent outposts or satellites of a nearby settlement. Such a pattern is evident even in the Middle Holocene, when the north coast of Santa Rosa Island served as a social gravitational center, with spinoff communities developing nearby rather than at more distant, higher-ranked locations (e.g. SRI-116 at the mouth of Lobo Canyon). Economic independence was almost surely compromised in the Late Holocene, when shell-beads-for-food & materials exchange systems were active, and chiefs were coordinating multi-site political units.

We likewise acknowledge that we have ignored regular NCI interaction with Chumash settlements on the Santa Barbara coast, through dynamic systems of intermarriage and trade. In fact, mainland Chumash were engaged in an evolutionary development late in the prehistoric that paralleled that on the Islands (Gamble 2008), with earlier evidence for stable settlements on the mainland coast at the junctures between large drainages and highly productive estuaries (Erlandson 1994). Early exploration of these islands was staged from the adjacent mainland, and it is likely that high-ranked habitats filled there prior to colonization of these islands (Fitzhugh and Kennett 2010). A regional application of the IFD in this area of high mobility and trade would include the mainland coastal areas of the Santa Barbara Channel, allowing for the possibility that the strains of population growth and declining habitat suitability were felt equally
there. In particular, spillover of Island populations to the mainland was not possible because the mainland was undergoing a similar process of in-fill and ecological saturation. Finally, we examine only ecological factors, yet we know that this was a socially complex system, especially in the later part of the historical sequence. This complicates tactics of economic cooperation and competition. Our database and consequently our analysis do not sufficiently reflect social variables likely to influence settlement, from kinship to witchcraft.

Given this list, it is somewhat surprising – and for that reason a sound endorsement of the cogency of the IFD model – that we detect a 8,000 year pattern consistent with our behavioral ecology predictions.

Pattern and Process in Settlement of the NCI.

Scholars of Chumash prehistory debate not only the causes of socio-economic evolution on the Islands, but questions of their absolute timing and whether the changes were gradual or episodic (Erlandson and Rick 2002; Kennett and Kennett 2000; Kennett et al. 2009; King 1990; Raab and Larson 1997), in the latter case consistent with an interpretation of punctuated equilibrium (Arnold 1992). If human socio-economic adaptation simply tracks environment, then the pattern of cultural evolution largely is a simple matter; allowing for lag, it matches that of environmental change. However, the IFD gives us a somewhat more sophisticated understanding of this relationship, by demonstrating how slow, continuous change in environmental or demographic variables might cause abrupt, discontinuous change in population-level responses. A few additional individuals and a new habitat is colonized; with Allee effects this colonization might be an abrupt relocation of a significant portion of the population to a new habitat. Marginal change in one variable produces discontinuous change in a related variable. A highly suitable habitat sees its population slowly decline while that of a
habitat with low basic suitability, but high resilience to exploitation, grows rapidly. Each of these patterns is evident in Figure 2 and Figure 3. These observations also suggest that terms like “gradual” or “punctuated” seldom will characterize a whole system; they more properly should be used to refer to specific variables and properties.

Population increase during the Late Holocene parallels the expansion of primary village locations around the coasts of the three largest islands. Sometimes the location of villages and associated cemeteries shifted to a slightly different location, but a majority of the high ranked habitats that were settled in the Middle Holocene continued to be locations of permanent settlement later in time (Kennett 2005). This is consistent with the prediction (ii) that the highest ranked habitats should continue to be occupied as expansion to secondary and tertiary habitats occurs. The real world consequences of infilling, environmental saturation, and declines in suitability toward the late Holocene are also reflected in decreases in body size, indicators of poor health (cribra orbitalia and periosteal lesions) and increases in violence (Lambert 1994; Lambert 1997).

Archaeologists debate the origins of institutionalized social hierarchy in the Santa Barbara region, with status differentiation assessed by such indicators as funerary objects marking high status burials. Estimates range from 2,400 cal. yrs. BP (King 1990) to cal. yrs. 800 BP (Arnold 1992). Our preliminary analysis (Kennett et al. 2009: 309) placed this development between 1,500 and 650 cal. yrs. BP, based on the observation that 1st through 3rd quartile locations were nearly completely occupied by this time. Dry unstable climatic conditions in the Late Holocene interval probably rendered all habitats we identified as 4th quartile, and some of those in the 3rd quartile, unsuitable due to lack of potable drinking water (Kennett and Kennett 2000). The quantitative analysis presented in this paper gives us another way to triangulate an
estimate of the timing of this phenomenon. Arranged in declining order, the posterior
suitabilities and estimated Year of Settlement (Figure 9; lower panel) indicate three sharp drops
or discontinuities: Canada Verde to Old Ranch Cyn; unnamed China Camp 1 to Scorpion Cyn;
and, unnamed San Miguel South 2 to Johnsons Cyn.

The latter discontinuity is the largest and in the period of interest. Unnamed San Miguel
South 2 has a suitability estimate of 7.0, for a predicted YS of 1070 cal. yrs. BP (Table 1). The
comparable values for Johnsons Cyn are 6.6 and 712 cal. yrs. BP. Notably, Johnsons Cyn
initiates a short run of locations that skirt the upper edge of the $S_{min}$ credibility interval, before
another discontinuity drops clear through it (Unnamed China Harbor to Dry Canyon). Our
model suggest that expansion of growing population to any location after Unnamed San Miguel
South 2 (1070 cal. yrs. BP) would be quite stressful, hovering on the boundary of locations
unsuitable for settlement. By Hazard’s Canyon (676 cal. yrs. BP) and Unnamed China Harbor
(633 cal. yrs BP) viable options are exhausted; all remaining locations fall below the lower
bound of minimum suitability Island residents seeking to relocate faced settlement
opportunities which were becoming both rare and relatively poor in their prospects; the
remaining inventory of empty watersheds was apparently not viable for villages. This suggests a
period of heightened social stress due to the environmental pressures of circumscription in the
interval 1070 to 676 cal. yrs. BP, overlapping our earlier estimate of 1500 to 650 cal. yrs. BP,
and toward the more recent end of the range debated by King and Arnold (2400 to 800 cal. yrs.
BP).

Computationally-intensive Methods and the Peculiarities of Archaeological Datasets

We adopted the Gibbs sampler in order to turn to our analytical advantage features of the
NCI dataset that would be problematic for conventional modeling: error in radiocarbon
estimates and informative censoring of unsettled watersheds. We argue that the Gibbs sampler is both powerful and particularly well suited to research on prehistory. Litton and Buck (1995) make a similar argument for archaeological dating and spatial sampling. Archaeology is maturing as an empirical science. Archaeological data are becoming much more abundant, reliable, and diverse in the subjects they address. However censored and missing observations as well as other features that impede conventional statistical modeling are frequently encountered in archaeology. It is important to develop analytical and statistical methods suited to large-scale, comparative analysis of archaeological datasets that nonetheless accommodate the peculiar difficulties of this work. A better understanding of the most intriguing, important and enduring archaeological problems will require these new approaches.

Gibbs sampling is useful when a joint distribution (e.g., the probability of an ensemble, consisting of sample data, unknown parameters and unobserved variables) cannot be computed directly, but can be decomposed into tractable conditional distributions (Gelfand and Smith 1990; Geman and Geman 1984). Computational approaches more general than the Gibbs sampler are available when conditional distributions are difficult to specify (see Gilks et al. 1996). At each iteration of the Gibbs sampler, the conditional distributions are used to generate a single probabilistic “draw” from the joint distribution. Many iterations of the Gibbs sampler produce a sequence of draws from the joint distribution, from which its essential statistical features can be extracted. For example, in the NCI analysis, the mean of the model parameter $S_{\text{min}}$ is estimated by the average of the probabilistic draws of $S_{\text{min}}$ (6.3, as given in Table 2). The same draws provide the 95% credibility interval estimate: 6.1 – 6.5

The Gibbs sampler provides a practical means of retaining information from records that are afflicted with missing data, measurement error or censoring (Knight et al. 1998). Rather than
ignore measurement error or discard partial records along with the information they do contain – standard tactics because conventional methods do not handle such irregularities gracefully -- we impute unobserved variables probabilistically. Principled imputation replaces more *ad hoc* approaches, allowing us to use the valuable information in partial or error-prone records. This is especially important if the data are by their nature (a) unique and highly localized, (b) difficult and costly to gather, and (c) not replaceable or replicable following recovery, all attributes characterizing archaeological datasets. Sites cannot be re-excavated any more than we can replicate the socio-economic behaviors that produced them. Often we are working with small samples, and with records produced by older and less thorough methods, or with datasheets unevenly completed. The larger the comparative or multi-site dataset we assemble for analysis, the more likely it is to have various compromising faults. But even if it is messy, we seek methods to make the most of the information contained in such datasets.

**Measurement error in radiocarbon estimates and calibration:** When an imprecise proxy (such as a radiocarbon date) is observed in place of a variable of interest (such as a calendar date), additional uncertainty can, and should, be incorporated into model-fitting and inference (Buck et al. 1996; Mallick and Gelfand 1996). The Gibbs sampler generates many random draws from the conditional distribution of the unobserved variable, given the observed value of the proxy (as well as perhaps other variables), and these imputed values are used in turn to infer model parameters. Uncertainty about the unobserved variable is introduced appropriately in inferences, as the imputed values vary stochastically. If needed, the relationship between the unobserved variable and its proxy can be specified by a calibration curve.

**Missing data:** Conventional responses to missing data are to eliminate the record, estimate the missing value by interpolating, or conduct a new investigation to fill in what is
missing (Knight et al. 1998: 471). The first tactic ignores what may be quite useful information in the remainder of the record; the second embroils us in arbitrary choices; the third is, as noted above, often impossible with archaeological samples. The Gibbs sampler takes a probabilistic approach to missing values, completing a record not with a single best guess but with a distribution of plausible values, which concretely characterizes uncertainty about the missing observation. Less obviously, but perhaps more importantly, this uncertainty feeds back to parameters and other model unknowns at each Gibbs iteration.

**Censored data:** Broadly speaking, censoring occurs when a variable is truncated or unobserved due to relatively specific, external mechanisms. Some examples are a time trial stopped before the outcome is observed for all participants; observations taken by an instrument limited to a minimal measurement $m$—that is the value recorded when an observation’s true value is $m$ or smaller; or, in the NCI data, locations never settled, and therefore censored for Year of Settlement, perhaps as a consequence of human habitat selection. The Gibbs sampler permits censored values to be treated as probabilistic unknowns, incorporating information about the censoring mechanism when available. For example, a Gibbs sampler can produce random draws having values $m$ or smaller, for a censored observation recorded by the measurement-limited machine. As for missing data, the distribution of these random draws informs about the censored observation as well as other model unknowns.

**Conclusions**

Ecologists are attracted to islands because insular environments appear to simplify processes that are thought to be pervasive, but are more difficult to isolate and observe in other settings. From the work of Darwin and Wallace forward, the history of biological science
supports the merit of this inclination. It would be fair however to point out that it is easier to argue for the merits of problem simplification than for the point that ecological processes analyzed on islands are, in some straightforward manner, representative of other landscapes. The hunch that this is true is regularly reinforced, but has not been rigorously demonstrated, and there are major differences between island and continental contexts that cannot be ignored (Grayson and Meltzer 2002).

Our analysis of NCI settlement follows in the island ecology tradition (see Kirch 1997). We have pursued population ecology analysis of the NCI dataset because it is unusually thorough in its time depth and spatial coverage, data recovery and interpretive possibilities. Indeed, comparable datasets within archaeology, amenable to detailed, comparative behavioral ecology investigation, are still fairly rare. But, we also have pursued this case because we believe it to contain clues to similar processes in continental environments, which for a variety of reasons, are more difficult to investigate. All environments are spatially heterogeneous in factors affecting their suitability for human occupation; all are made up of spatially isolated habitat “islands,” differing in their suitability to human use. This is likely to be true of landscapes structurally similar to those of the Northern Channel Islands (e.g., coastal Peru), and those quite divergent (e.g., Columbia River plateau).

Analysis of 38 watersheds, 29 with and nine without village sites, demonstrates that settlement of the Northern Channel Islands generally follows a pattern predicted by the Ideal Free Distribution (Fretwell and Lucas Jr 1970). This is confirmed by model-building that isolates drainage size, length of rocky intertidal zone and, inversely, length of sandy beach, as the environmental features most salient for decisions about prehistoric site location, as measured by earliest date of settlement. The IFD signal emerges strongly despite a long list of potential
confounding effects. In addition, we can rank the eight drainages classified as “Unknown”
according to predicted suitabilities, and predict that they contain as yet undiscovered villages of
specific ages.

As population grows, organisms distribute themselves over habitats differing in their
suitability in an orderly fashion that sometimes can be predicted by the IFD. The model links
individual-level adaptive decisions to population-level consequences. It allows for dynamically
changing suitability; it incorporates both negative density dependence, in the form of
exploitation and interference competition, and positive density dependence, through Allee effects
representing economies of scale. We envision the IFD being useful in analyses of migration,
habitat choice, settlement pattern and intensification, in insular and pseudo-insular habitats
(Allen and O'Connell 2008; Kennett et al. 2006b; Kennett and Winterhalder 2008; McClure et al.
2009; Shennan 2007). We hope to see it tested in other settings, with methods like those
demonstrated here.

Acknowledgments

For gracious help in many forms we thank the Channel Islands National Park staff, the
UC Davis HBE Lab Group, Brendan Culleton, Alan Gelfand, Sheryl Gerety, Anne Huston, Sarah
McClure, Kelly Minas, and Don Morris. Our work is supported by NSF grants SBR-9521974
and HSD 0827275. Figures

Figures

Figure 1. Key Drainages of The Northern Channel Islands, California.
Numbers follow the order given in the first column of Table 1. They index a ranked ordering of watersheds by our preliminary analysis of environmental suitability for human habitation (see Kennett et al. 2009).

Figure 2. Ideal Free Distribution, negative density dependence (without Allee Effects).
The upper panel shows suitability curves, on a normalized scale of 0 to 1, for three habitats -- a, b, and c -- as a function of the population density in that habitat. The habitats are ranked in alphabetical order by the suitability experienced by the initial occupant; in all cases suitability declines with population growth. The lowest ranked habitat, c, also experiences the lowest rate of declining suitability. The lower panel shows how population growth will be allocated among habitats given these suitabilities. See the text for further explanation. Note that the curves in the lower graph have been smoothed for ease of visualization; in actuality they would be jagged because individuals necessarily relocate in whole units. Suitabilities are given by Eq. 1 with: \( Q_a = 1, B_a = 0.0002, M_a = 0; Q_b = 0.8, B_b = 0.0008, M_b = 0; Q_c = 0.6, B_c = 0.00005, \) and \( M_c = 0. \)

Figure 3. Ideal Free Distribution, with Allee Effects.
The upper graph depicts habitat suitability in three ranked habitats as a function of habitat-specific density. Habitats b and c are characterized by an Allee effect: at low densities, habitat suitability increases with increasing density. After this initial phase, interference and depletion competition again come to dominate, causing suitability to decline. The lower graph shows the resulting population distribution over habitats, as a function of total population size. Suitabilities are given by Eq. 1 with: \( Q_a = 1, B_a = 0.0002, M_a = 0; Q_b = 0.8, B_b = 0.0008, M_b = 10; Q_c = 0.6, B_c = 0.00005, \) and \( M_c = 80. \)
Figure 4. Environmental Features of a High- (Canada Verde, Santa Rosa Island) and Low- (Unnamed San Miguel South 1, San Miguel Island) Ranking Watershed. Watershed numbering corresponds to the first column of Table 1.

Figure 5. Northern Channel Island, Village Settlement History.

Presence or absence of evidence for a village settlement at the mouth of watersheds by archaeological period, arrayed by decreasing quartiles of environmental suitability, as estimated in our preliminary analysis. The general pattern in these data substantiates hypotheses (i) and (ii) of the Ideal Free Distribution: High-ranking watersheds are settled earlier and low ranking later, the order of settlement following habitat suitability; and, once settled, a habitat tends to remain settled. Unsettled locations, and those for which settlement is unknown are indicated along with the watershed names at the base of the illustration. Note that sites not at the mouth of major watersheds (“Unranked Locale”) tend to be quite recent (e.g., since 1400 cal. yrs. BP). They are: SCrI-495, SRI-15, 427, SMI-503/504, 525, 528, 602.

Figure 6. Scatter plots of Year of Settlement (cal. yrs. BP) versus four environmental features.

For settled locations (n=29), the conditional mean settlement year (calendar years BP), given the earliest radiocarbon date at the location, is calculated by the method of Buck et al. (1996, section 9.2) and shown with a filled circle. Error bars show 95% credibility intervals for settlement year. Open circles show environmental features for the unsettled locations (n=9). The letter “U” positions the unsettled locations within the plot, but does not refer to a numerical value for Year of Settlement.
Figure 7. Scatter plot of posterior mean settlement year (BP) predicted by the suitability model (Eqs. 2 and 3), versus posterior mean settlement year predicted by radiocarbon date alone, for the n=29 settled locations.

Model means and 95% credibility intervals are summary statistics of predicted calendar dates, over iterations of the Gibbs sampler. Means and 95% credibility intervals based on radiocarbon dates alone were calculated using the method of Buck et al. (1996, section 9.2).

Figure 8. Posterior suitabilities, for locations with suitabilities closest to the threshold S_{min}.

Posterior mean suitabilities and 95% credibility intervals are summary statistics of predicted suitabilities, over iterations of the Gibbs sampler. The vertical dashed line and shaded area show the posterior mean and 95% credibility interval for S_{min}. (see parameter estimate in Table 2). Filled circles indicate settled locations, and open circles indicate unsettled locations.

Figure 9. (Top to bottom) Drainage Area, length of Rocky Intertidal zone, length of Sandy Beach and estimated suitability for each location, in descending order of suitability.

Posterior mean suitabilities and 95% credibility intervals in the bottom panel are summary statistics based on 20,000 posterior samples (generated by a Gibbs sampler; see Appendix). The x-axis labels are location names (“UN” = un-named, “Cyn” = Canyon; see Table 1). Settled locations (n=29) are shown with darkly shaded bars and circles, locations with unknown settlement (n=8) have lightly shaded bars and circles, and unsettled locations (n=9) have un-shaded bars and circles.
Figure 10. Parameter estimates and 95% credibility intervals for model coefficients inferred from the Northern Channel Islands dataset, along with estimates from 100 randomly permuted datasets.

Estimates from permuted datasets are based on 1000 posterior samples generated by the Gibbs sampler. See the Appendix for details.

Appendix

Gibbs Sampler

Statistical inferences about the suitability model (Eqs. 2 and 3) for the Northern Channel Islands data need to incorporate two key features of the sample: first, the eight locations classified as unsettled are censored for Year of Settlement; and second, even for settled locations, Year of Settlement is observed with error, via a radiocarbon date proxy. We take a Bayesian approach to model-fitting, treating the basic suitabilities $S_1^*, ..., S_N^*$ as unobserved random variables. We impute suitabilities for the unsettled locations stochastically by data-augmentation. For settled sites, we use a stochastic measurement-error model to impute suitabilities conditional on radiocarbon dates. Here we give details of a Gibbs sampler, implemented in the programming language $\textit{R}$ (R Development Core Team 2009, version 2.9.1), which generates realizations from the posterior distributions of the model parameters $\beta$, $\sigma^2$, $S_{\text{min}}$ and basic suitabilities $S_1^*, ..., S_N^*$. Below, we use the generic row-vector $x_l$ (consisting of an indicator for the intercept, along with log(Drainage Area), Rocky Intertidal Length and Sandy Beach Length) to denote the covariates for location $l = 1, ..., N$.

The Gibbs sampler works in an iterative manner, producing at every iteration a new random draw from the conditional distribution of each stochastic quantity of the model, given the current values of all other quantities. These “full” conditional distributions can be specified explicitly for the
suitability model as follows:

1. Conditional on $\beta$ and $S_1^*, \ldots, S_N^*$, $\sigma^2$ is equal in distribution to $\sum_l e_l^2 / V$, where $e_l = S_l^* - x_l \beta$ is the residual for location $l=1, \ldots, N$, $V$ is a chi-squared random variable on $N-k$ degrees of freedom, and $k$ is the dimension of $\beta$.

2. Conditional on $\sigma^2$ and $S_1^*, \ldots, S_N^*$, $\beta$ has a multivariate Gaussian distribution with mean vector $(X'X)^{-1} X' S$ and covariance matrix $\sigma^2 (X'X)^{-1}$, where $X$ is the $N$ by $k$ matrix having rows $x_1, \ldots, x_N$, and $S = (S_1^*, \ldots, S_N^*)'$. The full conditionals 1 and 2 imply vague prior information about $(\beta, \sigma^2)$ in a Bayesian linear model for $S_1^*, \ldots, S_N^*$.

The `sim` function of the R library `arm` (Gelman et al. 2009, version 1.2-9) conveniently produces realizations of $\beta$ and $\sigma^2$ from the full conditionals.

3. Conditional on $S_1^*, \ldots, S_N^*$, the distribution of $S_{\min}$ is uniform on the interval $(a, b)$, where $a$ is the largest suitability $S_i^*$ among unsettled locations, and $b$ is the smallest suitability $S_j^*$ among settled locations. This scheme is a modest generalization of one described in Holloway et al. (2004, section 2.3), where only responses below the censoring threshold required imputation. In the present case, all of the suitabilities are unobserved and require imputation.

4. For an unsettled location $i$ having covariates $x_i$, the conditional density of $S_i^*$, given $\beta$, $\sigma^2$ and $S_{\min}$, is truncated Gaussian with mean $x_i \beta$, variance $\sigma^2$ and upper truncation threshold $S_{\min}$.

5. For a settled location $j$, the conditional density of $S_j^*$ incorporates the earliest radiocarbon date $R_j$ at location $j$, in addition to covariate values $x_j$ and model
parameters. We use a measurement-error model described by Mallick and Gelfand (see also Gelfand et al. 1997; 1996) in combination with the radiocarbon calibration model used by Buck et al. (1996, chapter 9) in order to specify the density of $S_j^*$. The full conditional of $S_j^*$ is the product of two densities, $f_j$ and $g_j$. $f_j$ is a truncated Gaussian density with mean $x_j \beta$, variance $\sigma^2$ and lower truncation threshold $S_{\text{min}}$. $g_j$ is a conditional density for radiocarbon date $R_j$, given the basic suitability $S_j^*$ and a calibration model connecting $R_j$ to $S_j^*$ (Mallick and Gelfand 1996). Buck et al. (1996, chapter 9) model the radiocarbon date $R$ as a Gaussian variable with mean $\mu(\theta)$ and variance $\tau^2$, where $\theta$ is the true calendar date for the radiocarbon sample in years BP, $\mu$ is a calibration curve and $\tau^2$ is a total variance, incorporating radiocarbon dating error and error in the calibration curve. Typically, $\mu$ is a piecewise linear curve fitted by dendrochronological (tree-ring) methods (Buck et al. 1996). For radiocarbon dates of terrestrial artifacts, we obtained $\mu$ from the IntCal04 Northern Hemisphere calibration curve (Reimer et al. 2004), while for marine artifacts, we obtained $\mu$ from the Marine04 global marine radiocarbon calibration curve (Hughen et al. 2004). The dating-error component of $\tau^2$ for each artifact is provided by the dating laboratory, and the calibration-error component of $\tau^2$ is recorded in the appropriate calibration curve data set. We estimate $\tau_j^2$ for a settled location $j$ directly, using the given errors; $\tau_j^2$ is not treated as a parameter of the suitability model. For a settled location $j$, the basic suitability $S_j^*$ and calendar date $\theta_j$ are related by the equation $\theta_j = \exp(S_j^*)$. By a straightforward extension of the model in Buck et al. (1996, chapter 9), the density $g_j$ for radiocarbon date $R_j$, given $S_j^*$, is Gaussian with mean $\mu(\exp(S_j^*))$ and variance $\tau_j^2$. Finally, the full conditional density
for $S_j^*$, given $\beta$, $\sigma^2$, $S_{\text{min}}$ and $R_j$, is $f_j \cdot g_j$. Samples $S_1^*, \ldots, S_N^*$ generated probabilistically by steps 4 and 5 are saved over iterations for posterior analysis. Along with these, we save samples of predicted settlement times (measured in calendar years BP), generated as $\exp(S_1^*), \ldots, \exp(S_N^*)$. In analogy with the plant reproductive biomasses modeled in Gelfand et al. (1997), the predicted settlement times are conceptual quantities, subject to truncation by a suitability threshold (which in the present case is a parameter to be inferred).

We initialize the Gibbs sampler as follows. For a settled location $j$, the initial value of $S_j^*$ is $\log(m_j)$, where $m_j$ is the conditional mean calendar date given $R_j$, under the calibration model of Buck et al. (1996, chapter 9). For an unsettled location $i$, the initial value of $S_i^*$ is a random variable $\log(U_i)$, where $U_i$ is uniform on the interval $(0.01, m)$, and $m$ is the smallest conditional mean calendar date $m_j$ among settled locations. $\beta$ and $\sigma^2$ are initialized by a regression of the initial values of $S_1^*, \ldots, S_N^*$ on the covariates $x_1, \ldots, x_N$. Iterations of steps 1-5 then follow.

We ran the Gibbs sampler for 402,000 iterations (for a run time of approximately 36 hours on a Dell Precision Workstation 650), discarding the initial 2000 iterations for a burn-in. Preliminary runs indicated that thinning at an interval of 20 iterations adequately reduces autocorrelation across posterior samples for the present model. Our posterior summaries are then based on a total of 20,000 samples.

Model selection

We used a posterior predictive squared-error loss criterion (Gelfand and Ghosh 1998) to select the “best” suitability model from a set of candidate models. The loss criterion accommodates simple types of censoring and is readily computed from Gibbs sampler output. However, as originally formulated the criterion assumes that each censored observation has a fixed and known censoring
threshold, whereas for the Northern Channel Islands data the suitability threshold is an unknown parameter (depending further on radiocarbon dates observed with error). We therefore made two simplifying assumptions, for model-selection purposes only: the first assumption is that the suitabilities $S_j^*$ for settled locations are directly observed as the values $\log(m_j)$; the second assumption is that the suitability threshold is equal to $\log(m)$. We modified the Gibbs sampler described above to conform to these assumptions: in step 4, $S_{\text{min}}$ was replaced by the fixed value $\log(m)$; steps 3 and 5 were eliminated; and, suitabilities $S_j^*$ for settled locations were equated to their counterparts $\log(m_j)$ throughout. These changes were made only to aid in model selection; to make inferences about the selected model we used all of the Gibbs sampler steps 1-5.

Broadly speaking, the loss criterion measures deviations between the sample suitabilities and a hypothetical set of replicate suitabilities, where the replicates are generated under a candidate model conditional on the sample (Gelfand 1996; Gelfand and Ghosh 1998). For the present sample containing suitabilities censored from above, the criterion has the form

$$D = \sum_i \left( \sigma_i^2 + (\mu_i - v_i)^2 \right),$$

where $\mu_i$ and $\sigma_i^2$ are respectively the mean and variance of the posterior predictive distribution for the suitability at location $l=1, \ldots, N$, $v_i = \log(m_i)$ if location $l$ was settled, and $v_i = \min(\mu_i, \log(m))$ if location $l$ was unsettled (see Gelfand and Ghosh 1998, Eq. 12, taking $k$ to infinity). A replicate suitability $S_{l, \text{rep}}$ for location $l$ can be produced at each iteration of the modified Gibbs sampler, as a random Gaussian deviate with mean $x_i \beta$ and variance $\sigma^2$. The mean and variance of $S_{l, \text{rep}}$ over iterations then provide numerical estimates of $\mu_i$ and $\sigma_i^2$. The model having the smallest value of $D$ among candidate models is selected as the best model. We imbedded the modified Gibbs sampler in a loop over candidate models, in order to evaluate $D$ for each model under consideration. To keep computing times reasonable, we ran the Gibbs sampler for 50,500 iterations per model, discarding the initial 500 iterations for a burn-in, and
thinning the remaining iterations at an interval of 5. Our posterior predictive loss calculations are then based on a total of 10,000 samples per model.

We began model selection by evaluating $D$ for each of 16 “main-effects only” models, enumerated as: the model having only an intercept, the model with an intercept and one predictor, and so on, through the model with an intercept and all four predictors. The smallest loss ($D = 31.5$) was obtained for the model with main effects log(Drainage area), Rocky Inter-tidal length and Beach length. The next best model included all four predictors ($D = 32.2$). We then examined models containing main effects for log(Drainage area), Rocky Inter-tidal length and Beach length, along with two- and three-way interactions among these predictors. The best such model included the three main effects and an interaction term for log(Drainage area) and Beach length ($D = 31.9$). We selected the “main-effects only” model with $D = 31.5$, as it had the smallest posterior predictive loss of all models investigated.

Permuting outcomes to investigate a null model

A null model of interest posits that the pattern of settlement in the Northern Channel Islands is unrelated to the environmental covariates under consideration. Here we compare the model fitted to the observed pattern of settlement to models fitted to permuted datasets, in which the relationships between environmental covariates and settlement have been artificially disrupted.

We generated permuted datasets by re-assigning the observed settlement outcomes (for each location: a radiocarbon date, “unsettled” or “unknown”) at random without replacement to the 46 locations surveyed for environmental variables. The three covariates, log(Drainage Area), Rocky Intertidal Length and Sandy Beach Length, remain intact in the permuted datasets, as they are recorded in the original dataset. For each permuted dataset, we fit the suitability model (Eqs. 2 and 3) using the Gibbs sampler as described above. Total computing times can be
considerable, because each permuted dataset is treated as if it were a real dataset; thus there is a need to balance the number of permuted datasets analyzed against the accuracy of the inferences for each dataset. We analyzed 100 permuted datasets (a small fraction of the 46! \approx 5.5 \times 10^{57} possible permutations of the Northern Channel Islands data); and, for each permuted dataset, we based inferences on 1000 posterior samples (discarding the initial 500 iterations from a total of 5500 Gibbs iterations, and thinning the remaining iterations at an interval of 5). The run time was approximately 48 hours on a Dell Precision Workstation 650.

*Figure 10* shows parameter estimates for the observed sample (in black; see Table 2) along with estimates for 100 permuted datasets (in gray). The coefficients $\beta_1$, $\beta_2$ and $\beta_3$ of Eq. 2 characterize the relationships between the environmental covariates and settlement outcome, and these parameters are our focus here. Most of the 95% credibility intervals shown in Figure 7 contain the value zero, confirming that permutation tends to “break” the relationship between covariates and settlement outcome. The left-hand panel of *Figure 10* suggests that, if there were no real relationship between log(Drainage Area) and settlement outcome, then the inference made from the Northern Channel Islands dataset would be unusual. The observed settlement pattern produces the largest positive estimate of $\beta_1$ among all estimates in the left-hand panel. This supports our claim that larger drainages in the Northern Channel Islands were settled systematically earlier than smaller drainages. The middle panel provides more modest support for a positive relationship between settlement outcome and Rocky Intertidal Length, and the right-hand panel suggests relatively weak support for a negative relationship with Sandy Beach Length.
Citations


Erlandson, J.M., Kennett, D.J., Ingram, B.L., Guthrie, D., Morris, D., and West, G.J., 1996. A radiocarbon chronology for the archaeology and paleontology of Daisy Cave, San Miguel Island,


Holloway, G., Nicholson, C., Delgado, C., Staal, S., and Ehui, S.K., 2004. A revised Tobit procedure for mitigating bias in the presence of non-zero censoring with an application to milk-
market participation in the Ethiopian highlands. Agricultural Economics 31, 97-106.


Barbara.


Really Big Press, Santa Barbara, CA.


Rick, T.C., Kennett, D.J., and Erlandson, J.M., 2005. Preliminary report on the archaeology and


Rogers, D.B., 1929. Prehistoric Man of the Santa Barbara Coast. Santa Barbara Museum of Natural History, Santa Barbara, CA.


Towner, M.C., and Luttbeg, B., 2007. Alternative statistical approaches to the use of data as


Table 1. Environmental and Archaeological Dataset.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Island</th>
<th>Location</th>
<th>Drainage Area</th>
<th>Rocky Intertidal Length</th>
<th>Sandy Beach Length</th>
<th>Kelp Forest Area</th>
<th>Critical Site</th>
<th>Material</th>
<th>Lab Number</th>
<th>14C Date (Error)</th>
<th>Calibrated Date (Lower-Upper Bound)</th>
<th>Predicted Suitability (Lower-Upper Bound)</th>
<th>Predicted Settlement Date (Lower-Upper Bound)</th>
<th>Latest Settlement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SR</td>
<td>Canada Verde</td>
<td>11.83</td>
<td>4.30</td>
<td>0.30</td>
<td>0.43</td>
<td>SRI-41</td>
<td>C</td>
<td>UCLA-0140</td>
<td>4260 (90)</td>
<td>4803 (4548-5172) (8.4-8.6)</td>
<td>8.5 (4548-5173)</td>
<td>4801 (4548-5173)</td>
<td>250</td>
</tr>
<tr>
<td>2</td>
<td>SR</td>
<td>Canada Tecolote</td>
<td>12.14</td>
<td>3.82</td>
<td>0.71</td>
<td>1.70</td>
<td>SRI-3</td>
<td>M</td>
<td>Beta-31788</td>
<td>7480 (90)</td>
<td>7688 (7488-7898) (8.9-9.0)</td>
<td>8.9 (7488-7898)</td>
<td>7685 (7488-7898)</td>
<td>250</td>
</tr>
<tr>
<td>3</td>
<td>SC</td>
<td>Canada de la Calera 1</td>
<td>34.35</td>
<td>0.00</td>
<td>1.70</td>
<td>0.03</td>
<td>SCrl-240</td>
<td>C</td>
<td>UCR-0165</td>
<td>2825 (150)</td>
<td>2986 (2572-3338) (7.9-8.1)</td>
<td>8.0 (2573-3333)</td>
<td>5710 (2573-3333)</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>SC</td>
<td>Canada Christi</td>
<td>18.10</td>
<td>1.57</td>
<td>2.46</td>
<td>0.11</td>
<td>SCrl-333</td>
<td>M</td>
<td>CAMS-9660</td>
<td>5610 (90)</td>
<td>5716 (5482-5932) (8.6-8.7)</td>
<td>8.1 (5479-5923)</td>
<td>5710 (5479-5923)</td>
<td>250</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>Old Ranch Canyon</td>
<td>18.59</td>
<td>1.46</td>
<td>2.98</td>
<td>0.12</td>
<td>SRI-187</td>
<td>M</td>
<td>OS-34475</td>
<td>3670 (40)</td>
<td>3261 (3062-3428) (8.0-8.1)</td>
<td>8.1 (3068-3428)</td>
<td>3258 (3068-3428)</td>
<td>250</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>Arlington Canyon</td>
<td>11.93</td>
<td>2.87</td>
<td>1.54</td>
<td>1.46</td>
<td>SRI-4</td>
<td>M</td>
<td>UCLA-0659C</td>
<td>6980 (120)</td>
<td>7215 (6918-7468) (8.8-8.9)</td>
<td>8.9 (6908-7463)</td>
<td>7210 (6908-7463)</td>
<td>1260</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>Canada Soledad</td>
<td>12.28</td>
<td>2.57</td>
<td>1.95</td>
<td>0.67</td>
<td>SRI-19</td>
<td>M</td>
<td>Beta-107049</td>
<td>3340 (60)</td>
<td>2865 (2702-3088) (7.9-8.0)</td>
<td>8.0 (2698-3083)</td>
<td>2862 (2698-3083)</td>
<td>1040</td>
</tr>
<tr>
<td>8</td>
<td>SC</td>
<td>Laguna Canyon</td>
<td>12.55</td>
<td>2.37</td>
<td>1.58</td>
<td>0.04</td>
<td>SCrl-109</td>
<td>C</td>
<td>Beta-119186</td>
<td>5520 (70)</td>
<td>6321 (6188-6448) (8.7-8.8)</td>
<td>8.8 (6188-6448)</td>
<td>6320 (6188-6448)</td>
<td>1200</td>
</tr>
<tr>
<td>9</td>
<td>SR</td>
<td>Ranch Canyon</td>
<td>8.84</td>
<td>2.58</td>
<td>1.90</td>
<td>1.38</td>
<td>SRI-60</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>- (1104-1496) (7.0-7.3)</td>
<td>- (1104-1498)</td>
<td>1303 (1108-1498)</td>
<td>250</td>
</tr>
<tr>
<td>10</td>
<td>SC</td>
<td>Willows Canyon</td>
<td>7.97</td>
<td>2.83</td>
<td>1.77</td>
<td>0.06</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (1043-1233) (6.8-7.1)</td>
<td>- (1103-1788)</td>
<td>3336 (1093-1788)</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>Unnamed China Camp 1</td>
<td>7.36</td>
<td>2.73</td>
<td>2.10</td>
<td>1.13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (6.3-9.3)</td>
<td>- (494-9575)</td>
<td>2904</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>SR</td>
<td>Dry Canyon</td>
<td>6.84</td>
<td>2.66</td>
<td>1.95</td>
<td>0.55</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (6.2-9.2)</td>
<td>- (494-9575)</td>
<td>2904</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SR</td>
<td>Water Canyon</td>
<td>12.28</td>
<td>0.57</td>
<td>2.55</td>
<td>0.95</td>
<td>SRI-77</td>
<td>M</td>
<td>OS-32376</td>
<td>1800 (30)</td>
<td>1084 (938-1232) (6.8-7.1)</td>
<td>7.0 (943-1233)</td>
<td>1085 (943-1233)</td>
<td>200</td>
</tr>
<tr>
<td>14</td>
<td>SC</td>
<td>Canada de los Sauces</td>
<td>6.20</td>
<td>2.74</td>
<td>2.44</td>
<td>0.25</td>
<td>SCrl-145</td>
<td>M</td>
<td>UCR-020</td>
<td>2140 (150)</td>
<td>1441 (1098-1798) (7.0-7.5)</td>
<td>7.3 (1093-1788)</td>
<td>1432 (1093-1788)</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>SC</td>
<td>Coches Prietos Canyon</td>
<td>5.39</td>
<td>3.11</td>
<td>0.92</td>
<td>0.00</td>
<td>SCrl-1</td>
<td>M</td>
<td>UCR-394</td>
<td>2900 (130)</td>
<td>2330 (1968-2682) (7.6-7.9)</td>
<td>7.7 (1953-2673)</td>
<td>2313 (1953-2673)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>SC</td>
<td>Alamos Canyon</td>
<td>5.08</td>
<td>3.21</td>
<td>0.90</td>
<td>0.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>SR</td>
<td>Jolla Vieja Canyon</td>
<td>9.31</td>
<td>1.18</td>
<td>1.43</td>
<td>0.58</td>
<td>SRI-62 M Beta-107046</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>SM</td>
<td>Willow Canyon</td>
<td>4.63</td>
<td>2.82</td>
<td>1.68</td>
<td>0.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>SR</td>
<td>Canada Lobos</td>
<td>4.74</td>
<td>2.66</td>
<td>0.30</td>
<td>0.31</td>
<td>SRI-116 M Beta-9229</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>SR</td>
<td>Unnamed China Camp 2</td>
<td>4.83</td>
<td>2.56</td>
<td>2.47</td>
<td>1.44</td>
<td>SRI-28 M Beta-107988</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>SR</td>
<td>Wreck Canyon</td>
<td>7.44</td>
<td>1.14</td>
<td>1.36</td>
<td>0.61</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>SC</td>
<td>Smugglers' Canyon</td>
<td>5.13</td>
<td>1.89</td>
<td>1.76</td>
<td>0.03</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>SC</td>
<td>Pozo Canyon</td>
<td>6.45</td>
<td>1.10</td>
<td>3.27</td>
<td>0.42</td>
<td>SCrl-474 C Beta-032108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>SM</td>
<td>Otter Creek</td>
<td>2.61</td>
<td>3.88</td>
<td>1.70</td>
<td>1.47</td>
<td>SMI-468 M OS-33421</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>SR</td>
<td>Unnamed Bee Rock</td>
<td>3.79</td>
<td>1.95</td>
<td>2.49</td>
<td>1.86</td>
<td>SRI-31 M Beta-96959</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>SC</td>
<td>Scorpion Canyon</td>
<td>6.17</td>
<td>0.00</td>
<td>0.16</td>
<td>0.01</td>
<td>SCrl-615 M Beta-10549</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>SR</td>
<td>San Augustine Canyon</td>
<td>3.62</td>
<td>1.68</td>
<td>1.72</td>
<td>0.43</td>
<td>SRI-432 M Beta-107050</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>SR</td>
<td>Unnamed Cluster Point</td>
<td>2.81</td>
<td>1.34</td>
<td>3.17</td>
<td>1.77</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>SC</td>
<td>Montanon Canyon</td>
<td>2.78</td>
<td>1.17</td>
<td>2.35</td>
<td>0.98</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>SC</td>
<td>Johnsons Canyon</td>
<td>3.32</td>
<td>0.59</td>
<td>3.74</td>
<td>0.14</td>
<td>SCrl-192 C UCR-0396</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>SM</td>
<td>Nidever Canyon</td>
<td>1.97</td>
<td>1.54</td>
<td>2.41</td>
<td>0.06</td>
<td>SMI-163 M OS-33420</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>SC</td>
<td>Valdez Canyon</td>
<td>2.65</td>
<td>0.22</td>
<td>0.39</td>
<td>0.03</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>SC</td>
<td>Unnamed Twin Harbors 1</td>
<td>2.81</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>SCrl-369 M UCR-0392</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>SC</td>
<td>Hazard's Canyon</td>
<td>2.58</td>
<td>0.22</td>
<td>0.20</td>
<td>0.00</td>
<td>SCrl-436 A -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>SC</td>
<td>Orizaba Canyon</td>
<td>2.73</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>SCrl-434 A -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>SC</td>
<td>Unnamed China Harbor</td>
<td>2.63</td>
<td>0.00</td>
<td>1.76</td>
<td>0.17</td>
<td>SCrl-306 C UCR-1647</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>SC</td>
<td>Unnamed Profile</td>
<td>2.55</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Island</td>
<td>Point</td>
<td>Rank</td>
<td>Unnamed Trident Cove</td>
<td>SMI-232</td>
<td>M</td>
<td>Beta-180770</td>
<td>1810</td>
<td>1093</td>
<td>(4.9-8.0)</td>
<td>(136-3087)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>--------</td>
<td>------------------------</td>
<td>------</td>
<td>----------------------</td>
<td>---------</td>
<td>---</td>
<td>------------</td>
<td>------</td>
<td>------</td>
<td>-----------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>SM</td>
<td>Unnamed San Miguel South 2</td>
<td>1.53</td>
<td>1.63</td>
<td>3.03</td>
<td>1.57</td>
<td>SMI-232</td>
<td>M</td>
<td>Beta-180770</td>
<td>1810</td>
<td>1093</td>
<td>(4.9-8.0)</td>
<td>(136-3087)</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>SC</td>
<td>Unnamed Trident Cove</td>
<td>2.21</td>
<td>0.22</td>
<td>0.70</td>
<td>0.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6.3</td>
<td>751</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>SC</td>
<td>Unnamed Lady's Harbor Cove</td>
<td>2.36</td>
<td>0.00</td>
<td>0.87</td>
<td>0.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.7</td>
<td>1810</td>
<td>(4.8-7.8)</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>SC</td>
<td>Unnamed Dick's Cove</td>
<td>2.30</td>
<td>0.00</td>
<td>0.28</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.7</td>
<td>T</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>SM</td>
<td>Unnamed San Miguel South 1</td>
<td>1.28</td>
<td>1.65</td>
<td>2.50</td>
<td>0.65</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.6</td>
<td>T</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>SC</td>
<td>Diablo Canyon</td>
<td>2.09</td>
<td>0.00</td>
<td>0.94</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6.2</td>
<td>T</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>SC</td>
<td>Unnamed Ruby Rock</td>
<td>2.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6.3</td>
<td>180770</td>
<td>(4.6-6.4)</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>SC</td>
<td>Unnamed Twin Harbors 2</td>
<td>2.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.7</td>
<td>T</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>SC</td>
<td>Canada de la Calera 2</td>
<td>1.99</td>
<td>0.00</td>
<td>1.57</td>
<td>0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.6</td>
<td>T</td>
<td>U</td>
<td></td>
</tr>
</tbody>
</table>

Notes: **Rank**: Order determined by preliminary analysis; ranks correspond to labeling of drainages in Figure 1. **Island**: SR = Santa Rosa; SC = Santa Cruz; SM = San Miguel; **Critical Site** is the site that provided the key $^{14}$C date; **Material** indicates what was $^{14}$C dated; A = Artifact, C = Charcoal and M = Marine. Diagnostic artifact dates are based on calibrated age ranges (see Kennett 1998); **Calibrated Date**, along with 95% lower and upper bounds, were obtained by the method of Buck et al. (1996, section 9). For charcoal samples, we used the IntCal04 calibration curve (Reimer et al. 2004); for marine samples we used the Marine04 calibration curve (Hughen et al. 2004), assuming a reservoir age of 270 +/- 60 based on adjustments to the calibration curve and data published by Ingram and Southon (1996). For sites dated by an “Artifact” (see Material column), the date and standard deviation were obtained directly from excavated materials; for these sites, Calibrated Date is the directly-obtained date. The letter “U” indicates unsettled locations, and “-” indicates locations having unknown settlement status because archaeological work has not been done in that region; **Predicted Suitability** is the average suitability over 20,000 posterior samples. The lower and upper bound are the 2.5 and
97.5 percentiles of the posterior distribution (the 95% credibility interval); **Predicted Settlement Date** is the average settlement date over 20,000 posterior samples. The lower and upper bound are the 2.5 and 97.5 percentiles of the posterior distribution. The letter “T” indicates locations having a predicted suitability below the estimated suitability threshold (6.3; see Table 2). Thus the model predicts that these locations would not experience settlement.
Table 2. Parameter estimates (posterior means and 95% credible intervals) for the regression model of Year of Settlement (cal. yrs. BP), fitted with a Gibbs sampler to N=38 surveyed locations. The parameter estimates are summary statistics based on 20,000 posterior samples. The Gibbs sampler is described in detail in the Appendix.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept ( \beta_0 )</td>
<td>5.73</td>
<td>(5.0, 6.4)</td>
</tr>
<tr>
<td>( \beta_1 ) log(Drainage Area km(^2))</td>
<td>0.82</td>
<td>(0.5, 1.2)</td>
</tr>
<tr>
<td>( \beta_2 ) Rocky Intertidal Length (km)</td>
<td>0.25</td>
<td>(0.04, 0.5)</td>
</tr>
<tr>
<td>( \beta_3 ) Sandy Beach Length (km)</td>
<td>-0.17</td>
<td>(-0.4, 0.07)</td>
</tr>
<tr>
<td>( \sigma ) SD(( \varepsilon ))</td>
<td>0.72</td>
<td>(0.5, 1.0)</td>
</tr>
<tr>
<td>Suitability Threshold ( S_{\text{min}} )</td>
<td>6.31</td>
<td>(6.1, 6.5)</td>
</tr>
</tbody>
</table>
Figure 1

[Map showing the Santa Barbara Channel with San Miguel Island, Santa Rosa Island, and Santa Cruz Island labeled. A compass rose and a scale for kilometers are also included.]
Figure 6

Northern Channel Islands Locations Surveyed for Settlements (N=38)

- Year of Settlement (BP, log scale) vs. Drainage Area (Km², log scale)
- Year of Settlement (BP, log scale) vs. Sandy Beach Length (Km)
- Year of Settlement (BP, log scale) vs. Rocky Intertidal Length (Km)
- Year of Settlement (BP, log scale) vs. Kelp Forest Area (Km², log scale)
Figure 7

Model Predictions for Settled Locations (n=29)

posterior mean settlement yr, given 14C date alone

posterior mean settlement yr, given 14C date and model
Figure 8

Locations with Posterior Suitabilities Near the Threshold (n = 14)

- Johnsons.Canyon
- Valdez.Canyon
- Orizaba.Canyon
- Hazard's.Canyon
- Unnamed.China.Harbor
- Dry.Canyon
- Willow.Canyon
- Montanon.Canyon
- Unnamed.Cluster.Point
- Unnamed.Dick's.Cove
- Unnamed.Twin.Harbors.2
- Unnamed.Lady's.Harbor
- Canada.de.la.Calera.2
- Unnamed.San.Miguel.South.1

posterior suitability
Figure 9

All Northern Channel Island Locations, Ordered by Posterior Mean Suitability

- **Drainage Area Km²**
  - settled
  - unknown
  - unsettled

- **Rocky Intertidal Km**

- **Sandy Beach Km**

- **Posterior Suitability**

Figure 10

Coefficients for Northern Channel Islands dataset, along with coefficients from 100 permuted datasets.