Ecological Monographs, 79(3), 2009, pp. 397–421

On the prediction of extreme ecological events

Mark W. Denny,[1] Luke J. H. Hunt, Luke P. Miller,[2] and Christopher D. G. Harley[3]

Hopkins Marine Station of Stanford University, Pacific Grove, California 93950 USA

Abstract. Ecological studies often focus on average effects of environmental factors, but ecological dynamics may depend as much upon environmental extremes. Ecology would therefore benefit from the ability to predict the frequency and severity of extreme environmental events. Some extreme events (e.g., earthquakes) are simple events: either they happen or they don't, and they are generally difficult to predict. In contrast, extreme ecological events are often compound events, resulting from the chance coincidence of run-of-the-mill factors. Here we present an environmental bootstrap method for resampling short-term environmental data (rolling the environmental dice) to calculate an ensemble of hypothetical time series that embodies how the physical environment could potentially play out differently. We use this ensemble in conjunction with mechanistic models of physiological processes to analyze the biological consequences of environmental extremes. Our resampling method provides details of these consequences that would be difficult to obtain otherwise, and our methodology can be applied to a wide variety of ecological systems. Here, we apply this approach to calculate return times for extreme hydrodynamic and thermal events on intertidal rocky shores. Our results demonstrate that the co-occurrence of normal events can indeed lead to environmental extremes, and that these extremes can cause disturbance. For example, the limpet *Lottia gigantea* and the mussel *Mytilus californianus* are co-dominant competitors for space on wave-swept rocky shores, but their response to extreme environmental events differ. Limpet mortality can vary drastically through time. Average yearly maximum body temperature of *L. gigantea* on horizontal surfaces is low, sufficient to kill fewer than 5% of individuals, but on rare occasions environmental factors align by chance to induce temperatures sufficient to kill >99% of limpets. In contrast, mussels do not exhibit large temporal variation in the physical disturbance caused by breaking waves, and this difference in the pattern of disturbance may have ecological consequences for these competing species. The effect of environmental extremes is under added scrutiny as the frequency of extreme events increases in response to anthropogenically forced climate change. Our method can be used to discriminate between chance events and those caused by long-term shifts in climate.

Key words: disturbance and extreme events; ecological surprises; environmental and thermal stress; environmental bootstrap; hydrodynamic forces; intertidal rocky shores; limpets; Lottia gigantea; *mussels;* Mytilus californianus; *patch dynamics; statistics of extremes.*

Introduction

The physical environment can be a controlling factor in ecological dynamics. In some cases, the pertinent characteristic of the environment is the average value of a particular factor such as temperature, precipitation, salinity, etc. As long as extreme thresholds are not exceeded (e.g., cooling to the point of ice nucleation within tissues, or heating to the point of irreparable protein damage), the performance of individuals, the behavior of populations, and composition of communities are often well described by mean conditions (e.g., Brown et al. 2004). However, environmental variables do exceed important biological bounds in nature, and there are many cases in which ecological dynamics depend more upon the extremes of environmental factors than on their means (Gaines and Denny 1993). Extreme levels of certain variables can lead to impairment of function or outright mortality of individuals, with important implications for populations, communities, and ecosystems. Extreme events can influence community dynamics and biodiversity by selectively removing community dominants, thereby freeing up resources for other species (e.g., Dayton 1971, Connell 1978, Sousa 1979). Similarly, disturbance associated with extreme events can reduce a community's biotic resistance to invasive species and increase rates of invasion (Gross et al. 2005, Altman and Whitlatch 2007). Extreme events can cause sufficiently dramatic ecological change that recovery is greatly delayed or impossible. Such effects arise when populations are pushed below some minimum density threshold (e.g., the Allee effect), or when a community or

Manuscript received 26 March 2008; revised 19 November 2008; accepted 24 November 2008. Corresponding Editor: A. M. Ellison.

[1] E-mail: mwdenny@stanford.edu

[2] Present address: Marine Science Center, Northeastern University, Nahant, Massachusetts 01908 USA.

[3] Present address: Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4 Canada.

ecosystem enters an alternate stable state (Allee 1949, Folke et al. 2004). The effects of extremes are under ever-increasing scrutiny as the frequency, and thus the ecological importance, of extreme events continues to rise in response to anthropogenically forced climate change (IPCC 2007).

There is a growing realization that many extreme ecological events are caused not by the action of a single extreme environmental stressor, but rather by synergistic interaction among multiple run-of-the-mill stressors (e.g., Paine et al. 1998). For example, canyons in New Mexico are subjected to occasional heavy rain, but the effect of rain on stream insects is typically muted because vegetation retards runoff. Similarly, fires are common in the canyonlands, and by themselves have little direct effect on stream fauna. However, in 1996 a large wildfire killed the vegetation in several canyons, and was, by chance, immediately followed by repeated downpours. The resulting floods severely eroded the local streambeds, killing virtually all lotic insects in the affected streams, and they have been slow to recover (Vieira et al. 2004). Thus, the random co-occurrence of "normal" environmental factors led to an ecological surprise (sensu Paine et al. 1998): an extreme event. How likely is it that compound events such as this will occur?

Compound events have been analyzed extensively in physics, fluid dynamics, and oceanography. These systems can be relatively uncomplicated, allowing exact solutions to be obtained, for instance, for the maximum amplitude of sounds and the maximum height of ocean waves (Denny and Gaines 2000). Ecological compound events are more complex, but because they depend on the chance alignment of easily measured everyday phenomena, many compound ecological events are open to statistical analysis and prediction. Here, we describe a statistical approach to the prediction of extreme ecological events (a modification of the moving-block bootstrap) and illustrate its utility in two scenarios drawn from intertidal ecology. Our approach can be applied in a wide variety of ecological contexts.

Analysis of Compound Extreme Events

Why is a new approach needed?

There is an extensive literature exploring the probability of extreme events (see Gaines and Denny 1993, Denny and Gaines 2000, Coles 2001, Katz et al. 2005 for reviews), and for many types of problems statistical approaches are well established. In simple cases, the statistics of extremes analyzes the empirical record of a single variable to predict the probability that an extreme value will occur. Often the value of interest is more extreme than any in the empirical record, and statistics provides a method for reliably extrapolating beyond measured data (Coles 2001). More recently, this univariate theory has been extended to analyze the probability that extremes of two or more variables co-occur (for a pertinent example, see de Haan and de Ronde [1998]). Again, emphasis of the analysis is on extrapolation beyond existing data (Coles 2001).

There are two important limitations to the use of this type of extreme-value theory. First, standard extreme-value theory does not handle well the temporal autocorrelation and cross-correlation among variables; typically the occurrence times and realized values of the extremes for multiple variables are taken to be mutually statistically independent. This may not be appropriate when considering biological consequences, in particular when ecological extremes depend on the time history of the variables involved and not just on the instantaneous extremes. Second, standard extreme-value theory is typically applied to stationary time series, which makes for difficulties in handling series with temporal evolution and even seasonality. In addition, there is a practical consideration. The complexity of multivariate analysis increases rapidly with the number of variables under consideration (Coles 2001). For many ecological cases, where seven or more variables must be considered, this complexity may be prohibitive for the average ecologist.

As a practical alternative to multivariate extreme-value analysis, we propose a relatively simple "environmental bootstrap" procedure that is capable of incorporating the complexities of ecological systems. The focus of our method is not on extrapolation beyond the range of existing data, but rather on the probability that values within the range of existing data might randomly co-occur in a pattern that leads to an ecological extreme event.

The environmental bootstrap

For any given habitat, the relevant physical environment is defined by a set of factors, each of which varies through time. Our exploration of this variation proceeds in three steps. Given a short-term empirical record of environmental variables, we first identify the predictable aspects of each. We then separate this predictable part from the remainder, the stochastic part. Lastly, we divide the stochastic part of each signal into segments (= blocks) and rearrange these segments so that they can be recombined with the predictable part of the signal to yield new patterns. These new, hypothetical patterns of fluctuation allow us to analyze how the environment would be altered if chance had played out differently.

We approach the specifics of the analysis through an example drawn from intertidal shores, where certain combinations of wave height, wave period, and tidal height can impose extreme hydrodynamic forces on benthic organisms (see Plate 1). We begin with an analysis of significant wave height, H_s, the average height of the highest third of ocean waves. H_s is a statistical index of "waviness" (Kinsman 1965, Denny 1988, 1995), and serves well as an example of the type of environmental variable that can contribute to extreme events.

Consider the hypothetical time series of H_s shown in Fig. 1A, a 4-yr record of the wave climate at a particular location. In this series, H_s is measured at discrete times

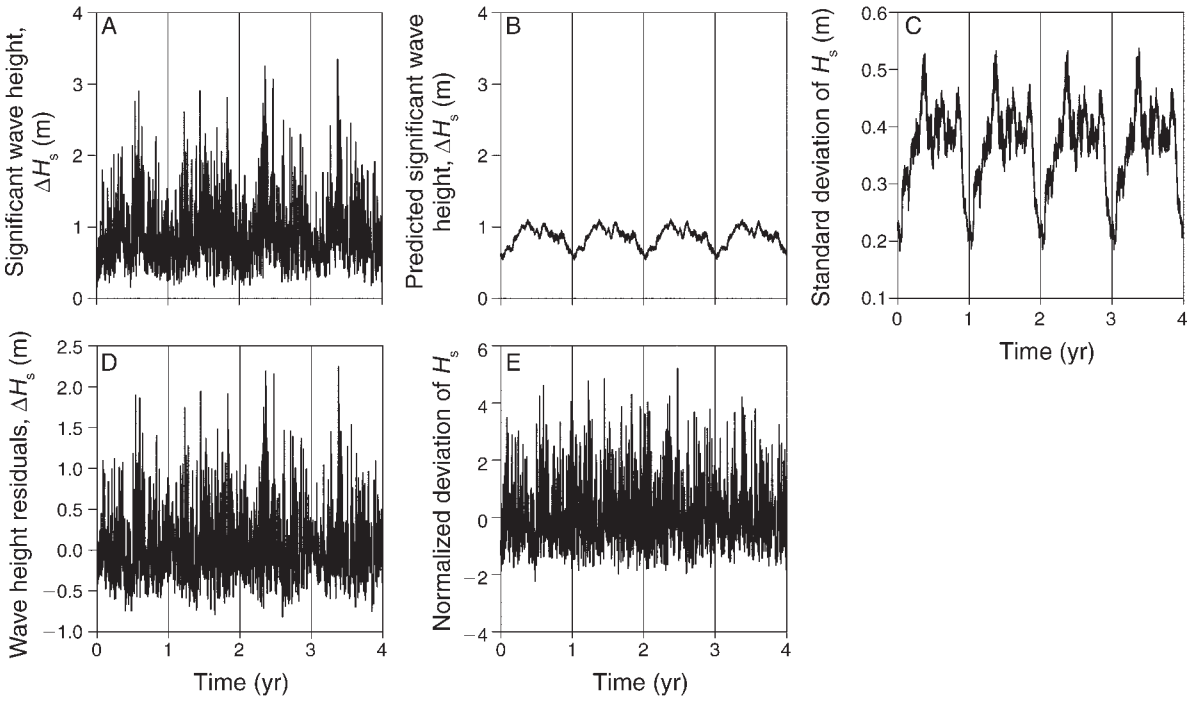

FIG. 1. Analysis of significant wave height data from Hopkins Marine Station. Four years of the seven-year data set are shown. Panel (A) presents the raw significant wave-height data. Panels (B) and (C) depict the annual cycles of predicted significant wave height and standard deviation, respectively, each repeated four times. Panel (D) shows the residuals of the raw wave-height data from the annual mean, and panel (E) shows these residuals standardized to the annual pattern of wave-height standard deviation.

($t = t_1, t_2, \ldots t_n$), and we denote the series with the symbol $H_s(t)$. Measurements are taken at constant interval Δt (in this example, 1 h). (Symbols are listed in Table 1.)

It is evident from inspection that there is an annual cycle of wave heights: on average, waves are small in summer (where the record begins) and large in winter. Similarly, wave heights are less variable in summer and more variable in winter. These predictable (=average, expected) seasonal cycles of the mean, $\bar{x}_H(t)$, and standard deviation, $\sigma_H(t)$, of H_s are shown in Fig. 1B, C. Details of the calculation of $\bar{x}_H(t)$, and $\sigma_H(t)$ are given in Appendix A. In short, we calculate $\bar{x}_H(t)$ and $\sigma_H(t)$ for points within a specific measurement window centered on the point at time t.

Having identified these predictable aspects of $H_s(t)$, our next task is to separate out the stochastic remainder, and to adjust that stochastic signal so that segments of it are statistically interchangeable. To this end, we need to ensure that all segments of the stochastic signal are identically distributed. For practical purposes, we consider the signal to be identically distributed if it meets four criteria: (1) the mean, (2) standard deviation, and (3) shape of the distribution of data within segments must be indistinguishable among segments anywhere within the record; and (4) the autocorrelation function of data must be indistinguishable among all segments.

We meet criterion 1 by calculating $\Delta H_s(t)$, the difference between the measured H_s and the predicted H_s for each sampling time (Fig. 1D):

$$\Delta H_s(t) = H_s(t) - \bar{x}_H(t). \tag{1}$$

The mean of these residuals is approximately 0 for segments throughout the record. To meet criterion 2, we divide each measured ΔH_s by the predicted standard deviation of H_s for that time in the year. That is, we calculate a standardized residual:

$$\Delta H_{s,\mathrm{std}}(t) = \frac{\Delta H_s(t)}{\sigma_H(t)}. \tag{2}$$

The result is shown in Fig. 1E.

Our next task is to ensure that the shape of the distribution of standardized residuals is not correlated with $\bar{x}_H(t)$, the annual mean cycle of H_s. Shape can be quantified by a variety of indices, but for present purposes, it is most important that the skew of the distribution is not correlated with the annual mean cycle. To test for this possibility, a standard index of skew (Zar 1974) is calculated for each point in the $\Delta H_{s,\mathrm{std}}$ time series in the same fashion in which the standard deviation is calculated (see Appendix A):

$$W(t) = \frac{m\sum_{i=1}^{m}[H_s(t_i) - \bar{x}_H(t_i)]^3}{(m-1)(m-2)\sigma_s^3}. \tag{3}$$

TABLE 1. Symbols used in the text.

Symbol	Meaning	Equation
A	area (m^2)	4
C_f	force coefficient	5
d	decorrelation time (d)	
d_{max}	maximum decorrelation time (d)	
D	water depth (m)	6
F_{max}	maximum force (N)	4
g	acceleration due to gravity	
H_b	breaking wave height (m)	8
H_{max}	maximum wave height (m)	6
H_s	significant wave height (m)	1
ΔH_s	residual from predicted H_s (m)	1
$\Delta H_{s,std}$	standardized ΔH_s	2
j	rank	13
L	lifetime	
m	number of points in a window	3
n	total number of values	
p	probability density ($1/R$)	18
P	cumulative probability	18
R	return time (intervals)	15
R_{per}	periodic return time (intervals)	
S_{max}	maximum hydrodynamic stress (N/m^2)	5
$\hat{S}$	annual maximum hydrodynamic stress (N/m^2)	13
t	time (s)	18
t_r	randomly chosen segment starting time (d)	
Δt	time interval (s)	
T	wave period (s)	10
U_{max}	maximum water velocity (m/s)	6
W	skew	3
$\bar{x}_H$	mean H_s (m)	1
X	$100D/(gT^2)$	9
α	coefficient (N/m^2)	14
β	coefficient	14
γ	velocity amplification coefficient	11
ε	coefficient	14
ρ	density of seawater (kg/m^3)	4
σ	standard deviation of factor residuals	
σ_H	predicted σ for H_s (m)	2
σ_s	standard deviation of point in a window	3
τ	interval length (seconds)	10
ϕ	body temperature (°C)	16, 17

Here m and σ_s are the number of data points and standard deviation of data in the measurement window, respectively. We then test this index for any correlation with the mean cycle. If there is a substantial correlation, the signal of standardized residuals must be adjusted and the validity of the adjustment must be verified (see Appendix A).

To this point, we have created a time series of standardized ΔH_s values (Fig. 1E) with constant mean and constant standard deviation, and the shape of the distribution of residuals is effectively constant across the series. However, we must meet one last criterion before we can rearrange segments of this time series: we must ensure that the data in each segment have the same autocorrelation function (ACF). As is characteristic of physical measurements, each $\Delta H_{s,std}$ is highly correlated with values measured a short time (a short lag) before or after, but the degree of correlation decreases with increasing lag (Fig. 2). It is not clear how one would adjust individual segments of the series to make their ACFs equal, but it is possible to empirically confirm the homogeneity among segments for a particular data set. Note in Fig. 2 that residuals separated by >4.5 days are effectively uncorrelated. The lag time at which the ACF settles to 0 is one measure of the decorrelation time, d.

We have now ensured that the time series of standardized residuals is effectively identically distributed. In light of this statistical similarity across the record, we are free to choose random segments of the $\Delta H_{s,n}$ series and recombine them with the predictable seasonal variation $\bar{x}_H$ to create a new wave-height record, a hypothetical realization of what the wave-height record might have been had the environmental dice rolled differently. To create this hypothetical realization, we carry out a moving-block bootstrap procedure (Carlstein 1986, Künsch 1989, Efron and Tibshirani 1993, Bühlman and Künsch 1995, 1999, Paparoditis and Politis 2003) as follows. We choose at random from the $\Delta H_{s,std}$ series a segment of length equal to the decorrelation time (that is, $d/\Delta t$ points). Setting segment length to decorrelation time ensures that any relevant autocorrelation in the signal is retained in each segment. We then multiply the first $\Delta H_{s,std}$ value in this segment by the standard deviation calculated for the first point in the time series, $\sigma_H(t_1)$, and add it to the expected value for that point, $\bar{x}_H(t_1)$, to yield a new, hypothetical significant wave height value. We repeat this process with the second point in the segment and $\sigma_H(t_2)$ and $\bar{x}_H(t_2)$, and so forth. At the end of this process, we have d days of hypothetical record. We then choose at random a second d-day segment from the $\Delta H_{s,n}$ time series (sampling with replacement) and use this new random segment to calculate the next d-day segment of the hypothetical wave-height record. We proceed in this fashion until we have created an entire year-long record, a record that is statistically similar to, but randomly rearranged from, the original. We can repeat the entire process as many times as desired to provide an ensemble of year-long realizations of the wave record.

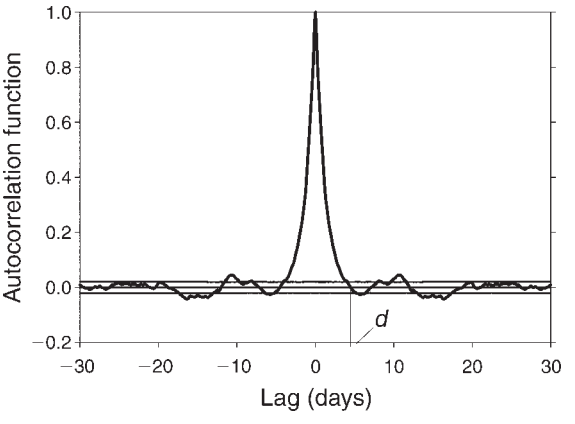

FIG. 2. The autocorrelation function for standardized residuals of significant wave height. The 95% confidence bounds are shown; values within these bounds are not significantly different from 0. The lag at which autocorrelation first reaches 0 is the decorrelation time, d.

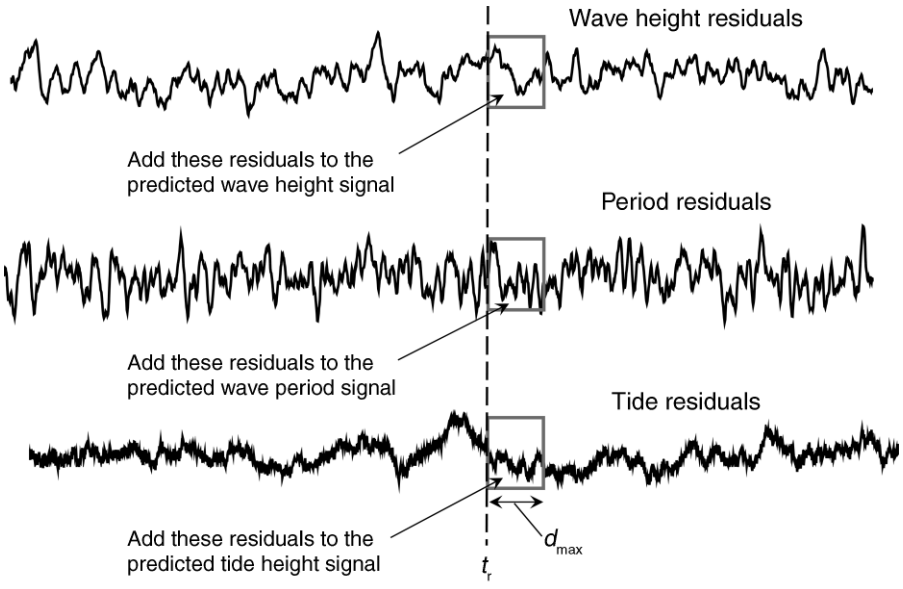

FIG. 3. A random time t_r is chosen to sample the normalized deviations of various environmental factors. Note that the same random starting point is used for all factors, thereby retaining any cross-correlation among factors.

Now to the crux of the matter. We simultaneously apply the same procedure to other aspects of the environment measured at the same time and in the same location as wave heights—wave period and tidal height in this example—to create sets of coordinated, concurrent hypothetical realizations. For our purposes, it is important that the sampling that leads to this ensemble of sets is random in time but consistent across variables (see Fig. 3). That is, when we choose a random time t_r as the beginning of a segment of wave height residuals, we use the same start time for wave period and tide height. *In this fashion, we maintain in the set of hypothetical realizations any cross-correlation among the standardized residuals of environmental factors*. For example, if for some reason the tides tend to be higher than predicted when waves are higher than predicted, this cross-correlation is maintained as we resample blocks of the time series of standardized residuals.

When resampling standardized residuals from multiple environmental variables, we use the longest decorrelation time among the variables, d_{max}, as the block length. This procedure ensures that pertinent autocorrelation is included in the sampled segments of all variables. (For details and exceptions, see Appendix A.) Moving-block bootstraps can be sensitive to block length (Efron and Tibshirani 1993, Bühlmann and Künsch 1999). To test for any obvious bias tied to the choice of d_{max} as block length, we repeat our whole analysis with blocks 1/2, 2, and 3 times d_{max}.

Note that rearranging blocks of standardized residuals can result in calculated values of individual environmental variables that fall outside the empirically observed range. For instance, the highest predicted tide in the empirical record might have coincided with a negative tidal residual due to the chance passage of a high-pressure cell in the atmosphere, resulting in a lower-than-expected observed tidal maximum. If in our random rearrangement of residuals the highest predicted tide is coupled with a positive tidal residual, the result is a higher calculated tide than any in the empirical record. There is nothing physically implausible about this calculation: in reality, the predicted tide (which is set by celestial mechanics independent of meteorology) might easily have occurred during the chance passage of a low-pressure cell, resulting in an exceptionally high tide such as that predicted by our rearrangement. Although it is thus possible and reasonable that our resampling method predicts values of individual variables that are more extreme than those empirically measured, the possibility of calculating such out-of-range values is not the driving force in our analysis, as we will see below.

Once we have constructed an ensemble of realizations for the factors of interest, our final task is to ascertain how many of these realizations would have caused an extreme ecological event: in this example, the imposition of lethal hydrodynamic force. For some processes, we might need simply to search the realizations for the co-occurrence of values that individually are sufficiently extreme, a task that could be handled by multivariate extreme-value analysis. However, in this bootstrap approach we are not constrained to such simple explorations. Instead, we are free to use any mechanistic model that appropriately combines variables to tell us whether a given set of conditions will result in an extreme ecological event. This model can be as complex as necessary to capture the biologically relevant aspects of environmental variability; in particular, it can include

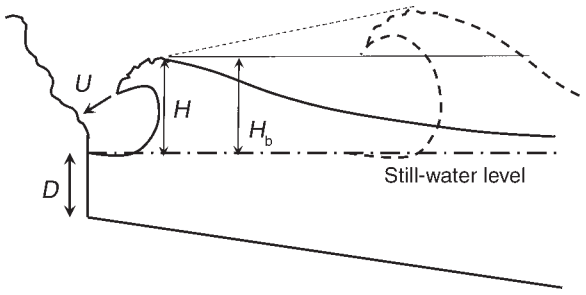

FIG. 4. A schematic diagram of waves breaking on a rocky shore. H is the height of an individual wave. U is the water velocity at the crest of the wave; U may be amplified by the local topography of the shore. D is the water depth at the toe of the shore, measured relative to still-water level. D varies with the tide. H_b, the breaking height limit, is set by wave period and water depth.

the role of history. Once this model is constructed, we can "play" each multivariable environmental realization through it, producing a univariate time series of biologically meaningful values (hydrodynamic force in the example here). Examination of the ensemble of these univariate time series allows us to tabulate extreme events. The total time in the ensemble divided by the number of extreme events provides an estimate of the return time for events of this type. For example, if 17 lethal hydrodynamic incidents occur in 1000 years of simulated nearshore environment, the return time, the average time between impositions of such an event, is $1000/17 = 58.8$ yr.

In summary, resampling a short time series of relevant factors provides a large bootstrap ensemble of hypothetical environmental realizations. We then use a mechanistic model of biological response as a tool to combine these environmental realizations, allowing us to analyze the resulting record for the occurrence of extreme ecological events. Here, we apply our approach to calculate return times for extreme hydrodynamic and thermal events on rocky shores, but we can apply the same methodology to a wide variety of environmental factors in virtually any ecological system.

MATERIALS AND METHODS

With the exception of the tides, all measurements were conducted at Hopkins Marine Station (HMS), Pacific Grove, California, USA (36.62° N, 121.88° W). We recorded data continuously for 7 yr, from 1 August 1999 through 31 July 2006.

Terrestrial environment

We logged air temperature (Vaisala HMP45C, Campbell Scientific Incorporated, Ogden, Utah, USA; 2 m above ground), solar irradiance (LI-200SZ, LI-COR Incoporated, Lincoln, Nebraska, USA; CM3, Kipp and Zonen, Delft, The Netherlands), and wind speed (Wind Monitor 05103-5, R. M. Young Company, Traverse City, Michigan, USA; 3 m above ground) every 10 minutes using Campbell 23X dataloggers (Campbell Scientific) at two sites at HMS.

The marine environment

We measured significant wave height and peak wave period (T, the period corresponding to the peak in the spectrum of wave energy) every 6 hours at a site ~100 m seaward of Cabrillo Point at HMS using an SBE26 bottom-mounted wave gauge (Sea-Bird Electronics, Belleview, Washington, USA). These 6-hour measurements were linearly interpolated to give H_s and T estimates at 10-minute and 1-hour intervals, as appropriate (see Appendix A). We measured sea surface temperature by hand to the nearest 0.1°C using an alcohol thermometer daily at 08:00 hours on Agassiz Beach at HMS. These daily measurements were linearly interpolated to give estimates at 10-minute and 1-hour intervals.

We obtained records of predicted and verified hourly tidal height from the NOAA tide station at Monterey, California, ~1.5 km from HMS. When needed, we linearly interpolated these hourly measurements to give height estimates at 10-minute intervals.

A model for extreme wave forces

As ocean waves break on shore, they are accompanied by high water velocities, which in turn impose large hydrodynamic forces on intertidal organisms (e.g., Koehl 1977, Denny 1988, 1995, Carrington 1990, Gaylord 2000, Boller and Carrington 2006). For a given organism, the magnitude of these forces depends on four factors: wave height, wave period, nearshore water depth (which varies with the tides), and the topography of the shore. A simple model can be formulated to relate these four factors to imposed force.

The maximum force, F_{max} (in newtons), imposed on an individual intertidal organism is

$$F_{max} = 0.5\rho U_{max}^2 A C_f \quad (4)$$

where ρ is the density of seawater (nominally 1025 kg/m^3), A is some representative area exposed to flow, C_f is an appropriate dimensionless shape-dependent force coefficient (e.g., the lift, drag, or impingement coefficient), and U_{max} (in meters per second) is the maximum velocity (Denny 1988, 1995, Gaylord 2000). In other words, for an organism of a given size and shape, maximum force is proportional to the square of maximum velocity. To estimate maximum force on mussels (lift in this case), we set $C_f = 0.88$ (see Denny 1987). (For simplicity, we ignore accelerational effects, which for small organisms, such as mussels, are negligible [Gaylord 2000].)

It is convenient to remove the size of an individual from this relationship by normalizing force to the same representative area used in Eq. 4. This manipulation results in a quantity we refer to as the maximum stress, S_{max}, with units of newtons per square meter:

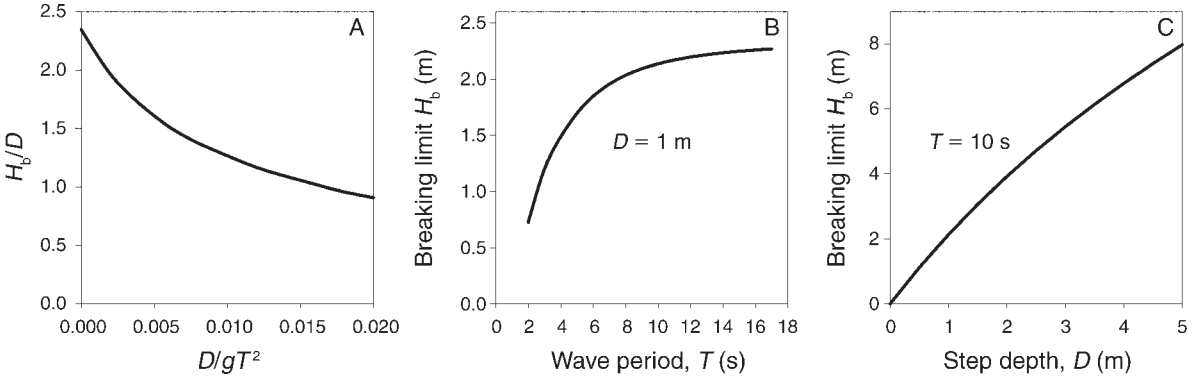

FIG. 5. We calculate breaking height according to Eq. 9. (A) The empirical relationship documented by the U.S. Army Corps of Engineers (1984). The curve shown here is for shores with a relatively steep seafloor leading up to a step at the shoreline, topography typical of rocky coasts. Because water depth contributes to values on both the abscissa and ordinate, it may be difficult to interpret panel (A) intuitively. As an aid, values of breaking height H_b are shown in panel (B) for fixed water depth and in panel (C) for fixed wave period. The wave length of a deep-water wave with period T is $gT^2/2\pi$ (Denny 1988). Thus, the term D/gT^2 is an index of the ratio of water depth to wave length.

$$S_{max} = \frac{F_{max}}{A} = 0.5\rho U_{max}^2 C_f. \quad (5)$$

Now consider the typical shoreline topography shown in Fig. 4. As a wave moves inshore, it traverses a sloping seafloor before impacting upon a steep rock wall. To estimate the maximum stress imposed on organisms attached to this wall, we need to estimate the maximum velocity imparted by waves. Solitary wave theory (Munk 1949) suggests that the maximum water velocity accompanying a wave as it breaks is the velocity at the wave's crest. For a wave moving over the topography shown in Fig. 4, this crest velocity is approximately

$$U_{max} = \sqrt{g(H_{max} + D)} \quad (6)$$

where g is the acceleration due to gravity (9.81 m/s^2), H_{max} is the height of the breaking or broken wave (the vertical height between wave crest and wave trough), and D, the step depth, is the still-water depth at the base of the wall (depth measured relative to mean lower low water [MLLW] and in the absence of waves). Inserting Eq. 6 into Eq. 5, we see that

$$S_{max} = 0.5\rho g(H_{max} + D)C_f. \quad (7)$$

Thus, maximum stress is proportional to the sum of maximum wave height and step depth.

Eq. 7 implies that, for a given step depth, stress could increase without bound as wave height increases. However, the higher the wave, the lower its stability, and the farther from shore it breaks (U.S. Army Corps of Engineers 1984, Denny 1988, Helmuth and Denny 2003). After a wave breaks, it rapidly loses height as energy is dissipated by turbulence and viscous interactions within the water column, and this energy loss sets an upper limit H_b to the height of waves as they reach the wall (Fig. 4). If the height of a wave as it approaches shore is $\geq H_b$, the wave breaks before reaching the wall, and the maximum stress imposed is as follows:

$$S_{max} = 0.5\rho g(H_b + D)C_f \qquad H_{max} \geq H_b. \quad (8)$$

If the maximum height of the wave as it approaches the shore is $< H_b$, S_{max} is given by Eq. 7.

To calculate maximum stress for a given step depth, we thus need to know the maximum height of waves relative to H_b. A variety of theoretical and empirical studies suggest that H_b is a function of both D (which is, in turn, a function of tidal height) and wave period, T. Here we use the empirical relationship found by the U.S. Army Corps of Engineers (1984) for steeply sloped shores:

$$\frac{H_b}{D} = -0.0561X^5 + 0.4152X^4 - 1.2534X^3 + 2.0573X^2 - 2.2433X + 2.3432 \quad (9)$$

where X is $100D/gT^2$. Eq. 9 is valid for $0 \leq X \leq 2$. Using Eq. 9, H_b can be calculated for given values of D and T (Fig. 5A). For a given step depth, the longer the wave period, the higher the breaking limit (Fig. 5B). For a given wave period, the deeper the water, the higher the breaking limit (Fig. 5C).

Our next task is to decide whether the actual maximum wave height impinging on the shore exceeds this breaking limit. In the ocean, wave height varies randomly from one wave to the next, and the distribution of wave heights can be characterized by the significant wave height, H_s. But, as noted previously, H_s is the average of the highest one-third of waves. Instead, we desire to know H_{max}, the height of the single highest wave. Theory developed by Longuet-Higgins (1952, 1980) and outlined in Denny (1995) shows that

$$H_{max} \cong 0.6541\left\{\left[\ln\left(\frac{\tau}{T}\right)\right]^{1/2} + 0.2886\left[\ln\left(\frac{\tau}{T}\right)\right]^{-1/2}\right\}H_s \quad (10)$$

where τ is the interval over which waves are measured (in seconds). Thus, if we know the significant wave

height present in a given interval (e.g., an hour, $\tau = 3600$ seconds) and T, the dominant period of the waves present at that time, we can calculate the predicted height of the highest wave that strikes the shore in that interval. This maximum wave height can then be compared to the breaking limit appropriate for that time (set by the wave period and tidal height, Fig. 5A), and used in either Eq. 7 or 8 to calculate the maximum stress imposed on an individual organism.

Note that maximum hydrodynamic stress is not a monotonic function of wave period. As shown in Fig. 5B, increasing T increases the breaking height, potentially allowing for the imposition of larger stresses. But increasing T also decreases the number of waves encountered in any interval of time, which decreases H_{max} (Eq. 10) and, thereby, the stress. The overall effect depends on step depth D, which in turn depends on the tide. If step depth is small, waves with any relevant height break before reaching shore, so S_{max} is set by H_b, which increases with increasing period. For sufficiently deep D, no wave breaks before reaching shore, so S_{max} is set by H_{max}, which decreases with increasing period. For intermediate step depths, stress is maximal at some intermediate wave period.

Eq. 10 is itself based on a probabilistic analysis of wave–wave superposition similar to the analysis of co-occurring events presented here. As a consequence, if several intervals in a time series have the same large (but not maximal) H_s, the combined time at this sea state could produce a larger maximum wave height than a single interval of maximal H_s (see Denny 1995). However, in practice, this scenario is unlikely. At the sites explored by Denny (1995), the maximum wave height associated with maximum H_s is less than the overall H_{max} only for H_s greater than 3.7 to 6.4 times the yearly average H_s. At HMS, yearly average H_s is 0.88 m, so we expect maximum wave height at maximum H_s to be less than the overall H_{max} only for $H_s > 3.3$–5.6 m, heights greater than those encountered in our 7-yr time series.

One more detail is required. Often the crest velocity of a breaking wave is amplified through interactions with the small-scale topography of the shore. Denny et al. (2003) demonstrated that wave refraction and the subsequent formation of jets amplified local velocities 1.5–2.6 fold. Because wave-induced stress increases as the square of velocity (Eq. 5), these amplifications lead to a 2.25–6.76 fold increase in stress. In general, if water velocity is locally amplified by γ, the stress predicted by our simple model is

$$S_{max} = 0.5\rho g \gamma^2 (H_{max} + D_s) C_f \qquad H_{max} < H_b \tag{11}$$

$$S_{max} = 0.5\rho g \gamma^2 (H_b + D_s) C_f \qquad H_{max} \geq H_b. \tag{12}$$

Given concurrent time series for significant wave height, wave period, and tide level, the procedure outlined above for calculating S_{max} for a single measurement time can be repeated for each measurement in the record, allowing one to estimate the overall maximum wave-induced stress associated with a particular realization of the environment at a particular site. If, by chance, high waves with certain intermediate periods coincide with a high tide, exceptionally large stresses are imposed.

Implementing the model

We processed empirical H_s, T, and tidal height data from HMS to ascertain mean signals and calculate identically distributed standardized residuals. The time series of residuals were then resampled as described previously and combined with the mean annual signals to create 10 000-year sets of concurrent hypothetical data.

We analyzed these results in two ways. First, we tabulated the probability of co-occurring conditions of wave height, tidal height, and wave period, and second, we recorded the annual maximum stresses calculated from the mechanistic model. Annual maxima were then further analyzed according to Gaines and Denny (1993) and Coles (2001) to calculate the return time for stress of a given magnitude: the average time between impositions of this magnitude of stress. This calculation is essentially a univariate extreme-value analysis of the stress data. The 10 000 bootstrapped annual maximum stresses were ranked, with the smallest annual maximum having rank 1 and the largest annual maximum having rank n. The estimated probability that an annual maximum stress $\hat{S}$ chosen at random from the recorded bootstrap values will have magnitude less than or equal to the value ranked j is

$$\text{Prob}\left[\hat{S} \leq \hat{S}(j)\right] = P(\hat{S}) = \frac{j}{n+1}. \tag{13}$$

This empirical cumulative distribution should asymptotically approach a generalized extreme-value distribution (Gaines and Denny 1993, Coles 2001, Katz et al. 2005):

$$P(x) = \exp - \left(\frac{\alpha - \beta x}{\alpha - \beta \varepsilon}\right)^{1/\beta}. \tag{14}$$

We calculated maximum likelihood values for α, β, and ε using Systat (SPSS 1998). Coles (2001) and Katz et al. (2005) use a modified (but equivalent) form of Eq. 14; see Appendix B.

The return time R for value x is

$$R(x) = \frac{1}{1 - P(x)}. \tag{15}$$

Return time as calculated here has units equal to the interval with which measurements are taken. For example, if maximum stress is measured for annual intervals, R has units of years.

Eq. 14 is often used in extreme-value analysis to extrapolate beyond the data in hand, for instance, to

extrapolate to return times in excess of the period of actual measurement. Here we use Eq. 14 primarily as a means to interpolate within the data provided by the bootstrap ensemble of environmental realizations.

Benchmarks of hydrodynamic stress

We used two benchmarks to judge the extremity of maximum hydrodynamic stress. The first benchmark is set by observed events. On 11 January 2001 an extreme wave/tide event occurred at HMS. Significant wave height was 2.80 m, wave period was 11.8 s, and tidal height was 2.15 m above MLLW (National Tidal Data Epoch [NTDE] 1960–1978, Monterey [NOAA 2000]). The co-occurrence of these high, intermediate-period waves with the second highest tide of the year resulted in substantial damage to the shore. Boulders weighing several tons were lifted more than 5 m and washed into the upper intertidal zone, and sections of previously intact rock substratum were broken loose and widely scattered. On 5 January 2008 a similar event occurred at HMS. In this case, significant wave height was 3.12 m, wave period was approximately 11–12 s (unfortunately, an apparent malfunction in the wave meter does not allow us to be more specific), and measured tidal height was 1.85 m above MLLW (NTDE 1983–2001, Monterey). Although some destruction occurs annually on this shore, these events were notably exceptional in the authors' 26-yr experience at HMS, and we use them as standards against which to compare our calculations.

The second benchmark is set by the mechanical capabilities of bed-forming mussels. When water flows over an intertidal mussel bed, a lift force is applied, tending to pull individuals away from the substratum (Denny 1987), and this force is the principal source of disturbance for mussels living above the limit of predatory sea stars (Paine and Levin 1981). The ability of mussels to resist this force was measured at HMS for a bed of *Mytilus californianus* 1.5 m above MLLW (NTDE 1960–1978, Monterey), as described by Denny et al. (2004). We recorded breaking stress values for 30 mussels each month from February 1999 to October 2006. Monthly values were pooled across years, and used to estimate the month-by-month cumulative distribution of breaking stress. We ranked the breaking stress values for a given month in ascending order, and used Eq. 13 to calculate the probability that a mussel chosen at random has a breaking stress less than that of a given rank. We then fitted this empirical cumulative probability distribution using Eq. 14, which in turn allows us to conveniently translate applied stress into the probability of mussel dislodgment. Note that in this case, $P(S_{\max})$, the probability that an individual chosen at random will be dislodged by stress $S_{\max}$, approximates the fraction of the mussel population dislodged by the imposition of $S_{\max}$. Knowing both fraction dislodged and return time as functions of applied stress, we can then plot fraction dislodged directly as a function of return time. Because the distribution of mussel strengths varies from month to month, the largest stress does not necessarily dislodge the largest fraction of mussels; a smaller-than-maximum stress occurring in a month when mussels are weak could result in the maximum annual dislodgment. Consequently, to calculate maximum annual dislodgement, each individual calculated hydrodynamic stress value was used with the breaking stress distribution appropriate to the month in which the stress was imposed. This is one example of a type of seasonal effect that would complicate a traditional multivariate analysis of extremes.

A model for limpet body temperature

We used a similar approach to predict the return time of extreme thermal events for limpets. The owl limpet *Lottia gigantea* is territorial; it maintains a "garden" of microalgae by bulldozing away mussels, barnacles, and other limpets, and returns to a home scar at each low tide. When firmly attached to its scar, *L. gigantea* is virtually immune to dislodgment by wave forces (Denny and Blanchette 2000). At HMS, *L. gigantea* is the only species that can effectively compete for mid-intertidal space with *M. californianus*, and therefore is an important influence on mid-intertidal community structure. Because of its homing behavior, adult *L. gigantea* are effectively immobile at low tide; they do not seek shelter on hot days, and are thus susceptible to overheating.

Denny and Harley (2006) constructed a heat budget model that predicts (within 0.3°C) the body temperature of *L. gigantea* from environmental data, and we use this model here. Briefly, we use the model to calculate the rate at which heat enters and leaves a limpet's body for a given set of environmental values, and then calculate the body temperature at which heat influx equals heat efflux. For small organisms such as these limpets, this equilibrium temperature is an excellent estimate of actual body temperature. The heat budget model thus provides a means for translating a time series of environmental variables into the corresponding time series of body temperatures.

We resampled pertinent aspects of the 7-yr environmental record (air and seawater temperatures, wind speed, solar irradiance, tidal height, and significant wave height) as described above and in Appendix A, and we used them in conjunction with the heat budget model to calculate 1000 hypothetical annual maximum body temperatures for *L. gigantea* at nine substratum orientations: horizontal; 45° from vertical facing east, west, north and south; vertical facing east, west, north, and south. We assume that all limpets occur at 1.5 m above MLLW, a typical height on the shore for *L. gigantea* at HMS. We then used Eqs. 13, 14, and 15 to calculate the return time for given maximum body temperatures. The calculation of each 1000-year ensemble required >5 days of computer time; for this reason we use only 1000 years of hypothetical thermal data

rather than the 10 000 years of wave/tide data we used when exploring mussel dislodgment.

Two aspects of the heat budget model deserve note. First, instantaneous body temperature is closely tied to the temperature gradient in the rock substratum, and therefore depends on the history of heat flux in the several hours leading up to each temperature measurement. Second, the values of wind speed, wave height, and solar irradiance relevant to body temperature are bounded. During times of thermal stress, body temperature is typically higher than air temperature. As a result, the lower the wind speed, the higher the body temperature. However, wind speed can never go below 0, and convection imposes a practical lower bound of ~0.1 m/s. Similarly, body temperature can be high only in the absence of wave splash. Thus, the lower the wave height, the higher body temperature is likely to be. However, wave height can never fall below 0. The brighter the sunlight, the higher body temperature can be, but irradiance at sea level is limited by the output of the sun to ~1000 W/m^2. Slightly higher values can be encountered for a few moments at a time as direct sunlight is augmented by light reflected off clouds, but these brief bright spells do not last long enough to substantially affect body temperature. The fact that these variables are bounded will be important when we consider the calculation of absolute maximum body temperature.

To aid in interpreting thermal data, we carried out an additional set of calculations to characterize the temporal pattern of heating corresponding to the imposition of maximal temperatures for limpets on horizontal substrata. We created 184 yr of resampled environmental data, and (using our heat budget model) recorded the body temperatures every 10 min within ±12 h of the annual maximum for each year. These records allowed us to characterize each annual maximum temperature event into two categories: (1) abrupt cooling, in which the hot limpet was immersed in seawater within 20 min of reaching peak temperature, (2) gradual cooling, in which high temperature was maintained for longer periods.

Benchmarks of thermal stress

We use death as a benchmark for the severity of thermal stress. To do so, we measured the distribution of lethal limits of *L. gigantea* as described by Denny et al. (2006). In short, we gently detached limpets from the substratum and transferred them to a chamber in the laboratory in which substratum temperature, air temperature, wind speed, and relative humidity were varied separately to impose patterns of body heating that resemble those found in limpets emersed at low tide. Two heating protocols were used. First, starting at a typical seawater temperature of 14°C, we increased body temperature at a rate of 8°C per hour to the experimental temperature. (This rate is among the highest seen in our simulations.) We maintained this peak temperature until 3.5 h had elapsed from the initiation of the experiment, at which time the limpets were abruptly cooled by immersion in 14°C seawater. We assessed the status of the limpet (alive or dead) 24 h after immersion. The second protocol was similar to the first except that, after an appropriate interval at experimental temperature, we gradually cooled the limpets at 8°C per hour back to 14°C (rather than the abrupt decrease of the first protocol), and the limpet was then immersed in 14°C seawater. In this case, the total time of the experiment (from initial onset of heating to immersion) was 7 h. We then assayed limpet status 24 h later. Experimental peak temperatures were varied from 28°C to 40°C in 1°–2°C steps. For the 3.5-h exposures, we used 10 limpets in each trial at each temperature; for the 7-h exposures, we used 5 limpets in each trial. All experiments were repeated for three trials. Wind speed was 0.5 m/s and relative humidity was 50–60% in each trial. Air temperature tracked substratum temperature up to 30°C, and was then held constant.

The results of these experiments suggest that 28°C represents an approximate threshold for thermal death in *L. gigantea*. At or below 28°C, limpets can survive the longest single exposure they are ever likely to encounter in nature. Above this threshold, limpets are killed, the fraction killed increasing with both temperature and time of exposure. The limit of 28°C is also the temperature at which *L. gigantea* increases its production of heat-shock proteins (HSP 70) (Miller et al. 2009). For future reference, note that the "abrupt cooling" experiments maintain a body temperature >28°C for 1.75 h; the "gradual cooling" experiments, for 3.5 h.

We use the lethal limits measured here to calculate return times of thermal events in terms of the fraction of limpets killed rather than in terms of limpet body temperature.

Verification

It would be comforting to test our calculations against long-term empirical measurements. However, as with many predictions of rare environmental events, a direct test is problematic: long-term records of mussel dislodgment and limpet thermal death do not exist for the HMS shore. In lieu of these data, we tested our method in three ways.

First, we tested the ability of our method of calculating residuals (Appendix A) to accurately estimate the distribution of residuals. We took daily sea surface temperatures at Hopkins Marine Station from a record beginning in January 1919 and extending through 2004 (Breaker et al. 2006), allowing us to calculate residuals relative to the 86-yr average for each year day. We then compared the distribution of these residuals to the distribution of residuals calculated relative to the annual cycle of temperature estimated using our 7-yr data record and the averaging procedure outlined in Appendix A.

Next, we tested the internal accuracy of the statistical approach using simulated environmental data. We used

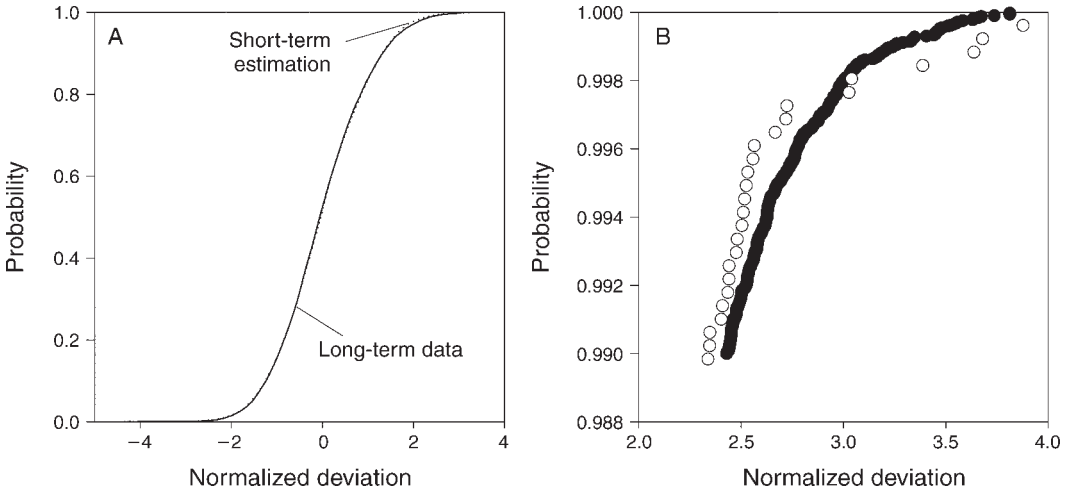

FIG. 6. The cumulative probability distribution of sea surface temperature residuals estimated from a short (7-yr) data set closely matches deviations calculated directly from an 86-yr data set. Panel (B) shows details of the upper tail of panel (A). Open circles are data from the short times series; solid circles are data from the 86-yr series.

second-order autoregressive-moving-average (ARMA) models to calculate 10 000-yr-long simulated time series of standardized residuals for tidal heights, wave heights, and wave periods (see Appendix C). When combined with the mean annual cycles measured from our 7-yr record, these time series of residuals provide the sort of lengthy data set not currently available from nature. A 7-yr segment of this hypothetical environmental record was then chosen at random and analyzed using our resampling method to estimate return times for extreme events. These calculations allowed us to compare the distribution of estimated return times (calculated from the 7-yr subseries) to the distribution of return times directly "observed" in the 10 000 years of the full-length simulated time series.

As a final means of testing the efficacy of our predictions for extreme events, we garnered hourly significant wave height and peak wave period from buoy 46042 of the National Data Buoy Center, located at 36.75° N, 122.42° W in the mouth of Monterey Bay, ~50 km from Hopkins Marine Station. With minor intermittent gaps, this record spans 20.5 yr, from 17 June 1987 to 31 January 2008, a total of 160 584 h of observation. We noted conditions at this buoy for the times of the known extreme events at HMS. We then searched the 20.5-yr record for the occurrence of similar events, and noted the intervals between these events.

RESULTS

Verification

Thermal residuals calculated using our averaging procedure are very similar to residuals calculated relative to the 86-yr average for each year day (Fig. 6).

Applying our statistical method to short-term (7-yr) simulated wave and tide data predicted a distribution of return times for hydrodynamic stress that closely matches the distribution observed directly in long-term (10 000-year) simulated data. On average, the estimates of annual maximum stress for a given return time overestimated the observed stress by 0.49% (Fig. 7). The fractional difference between estimated and actual stress generally decreased with increasing return time, and the maximum difference between predicted and actual stress was −3.62% at a return time of 5000 yr.

Wave events with individual characteristics as extreme as those observed on 11 January 2001 ($H_s > 2.80$ m, $T >$ 11.8 s, tidal height > 2.15 m) occurred in only 1.07% of the 10 000 1-yr realizations of the wave and tide environment, corresponding to a return time of ~93 yr. Wave events with individual characteristics as extreme as those

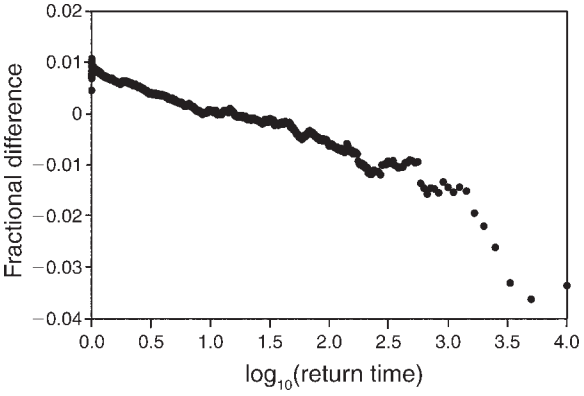

FIG. 7. Results from simulated wave/tide data. Extreme values estimated from a 7-yr subset of the simulated data set are compared to those "observed" in the entire 10 000 years of simulated data. Fractional difference is ([estimated value − observed value]/observed value). Estimated values are slightly larger than observed values for return times less than ~200 yr and are slightly smaller than observed values at longer return times. There are 10 000 data points in this plot, and they are not distributed evenly in log time. The large number of points at short return times controls the average fractional difference, which is +0.49%.

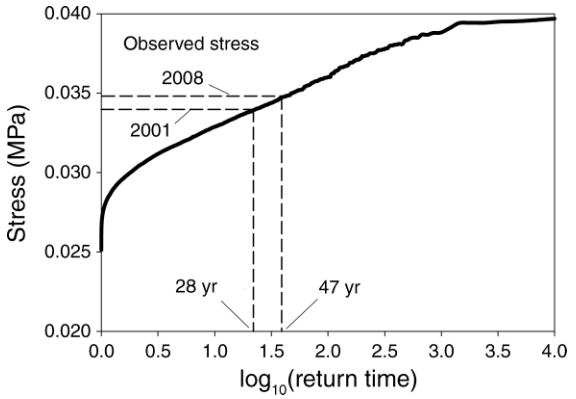

FIG. 8. Hydrodynamic stress values calculated for the January 2001 and 2008 extreme events correspond to return times of ~28 and 47 yr, respectively.

observed on 5 January 2008 ($H_s > 3.12$ m, $T > 11$ s, tidal height > 1.85 m) occurred in 1.86% of the 10 000 1-yr realizations, corresponding to a return time of 54 yr.

These return times can be compared to those estimated from buoy records. The 2001 extreme wave event at Hopkins Marine Station occurred when conditions at buoy 46042 were $H_s = 8.2$ m, $T = 16.7$ s. In the 20.5-yr history from this buoy, this was the only time when $H_s \geq 8$ m and $T \geq 16.7$ coincided with a measured tidal height >2.0 m. The 2008 extreme wave event at Hopkins Marine Station occurred when conditions at buoy 46042 were $H_s = 9.9$ m, $T = 17.0$ s, the only time when $H_s \geq 9$ m and $T \geq 17$ coincided with a measured tidal height >1.85 m. If we use these two events as the benchmark for what constitutes an extreme wave/tide event, the buoy record implies return times for such events of >13.5 yr (from some undetermined time prior to June 1987 to January 2001) and 7 yr (January 2001 to January 2008).

Extreme stress events

As noted above, the combinations of wave height, wave period, and tidal height present in January 2001 and 2008 are only a subset of the possible conditions that might qualify as biologically extreme. Return time as a function of calculated imposed stress is shown in Fig. 8. Our simple model for wave stress suggests that stress equivalent to that imposed on 11 January 2001 has a return time of approximately 28 yr, and stress equivalent to that of 5 January 2008 has a return time of 47 yr. As expected, both values are lower than the return times estimated for specific combinations of waves and tides (93 and 54 yr, respectively).

Mussel strength distributions

Fig. 9 shows the cumulative distributions of mussel breaking stress. Tenacities are higher from October through February, with a peak in November, and lower in March through September, with a minimum in April. Monthly distributions of mussel breaking stress are described by the coefficients given in Table 2.

Fraction of mussels dislodged

The fraction of mussels dislodged by hydrodynamic forces is graphed as a function of return time in Fig. 10. Values are shown for three levels of velocity amplification ($\gamma = 1$, 1.5, 2.6) and for two water depths ($D = 1$ and 5 m measured from MLLW), spanning the typical range for rocky shores. Without local amplification, wave-imposed stresses dislodge a miniscule fraction of the population even at long return times, regardless of water depth. With a 2.6-fold amplification of velocity, >80% of mussels are dislodged annually at $D = 1$ m, and >99% at $D = 5$ m. With intermediate velocity amplification, the fraction of mussels dislodged increases with both return time and water depth. There was negligible effect of varying block size (Fig. 11).

Fig. 12 shows the distribution of times of arrival of yearly maximum wave stress; maximal stress arrives most often in December when mussels are near their strongest, and least often in July and August when mussels are near their weakest. This correlation between imposed stress and byssal strength suggests that mussels may adjust their adhesive tenacity in response to the hydrodynamic environment.

Limpet thermal limits

The fraction of limpets killed by temperature ϕ is shown in Fig. 13, where the data are described by the following equations:

$$\text{Fraction killed} = \frac{1}{1 + \exp\left(-\dfrac{\phi - 32.85}{0.7978}\right)} \qquad (16)$$

for gradual cooling, and

$$\text{Fraction killed} = \frac{1}{1 + \exp\left(-\dfrac{\phi - 36.73}{0.3863}\right)} \qquad (17)$$

for abrupt cooling. (Eq. 16 is taken from Denny et al. [2006].) A temperature of ~37°C is required to kill half the limpets when individuals are exposed to potentially lethal temperatures (temperatures above 28°C) for 1.75 h. In contrast, a temperature of only ~33°C is required to kill half the limpets when exposed to potentially lethal temperatures for 3.5 h.

Pattern of heating

Of the annual temperature maxima, 69% were associated with gradual cooling to seawater temperature; the remaining 31% were associated with abrupt cooling. Average patterns of heating are shown in Fig. 14. The higher the annual maximum body temperature, the longer the time spent above the threshold of 28°C, although there is substantial variation around this pattern (Fig. 15). As one might expect, time above

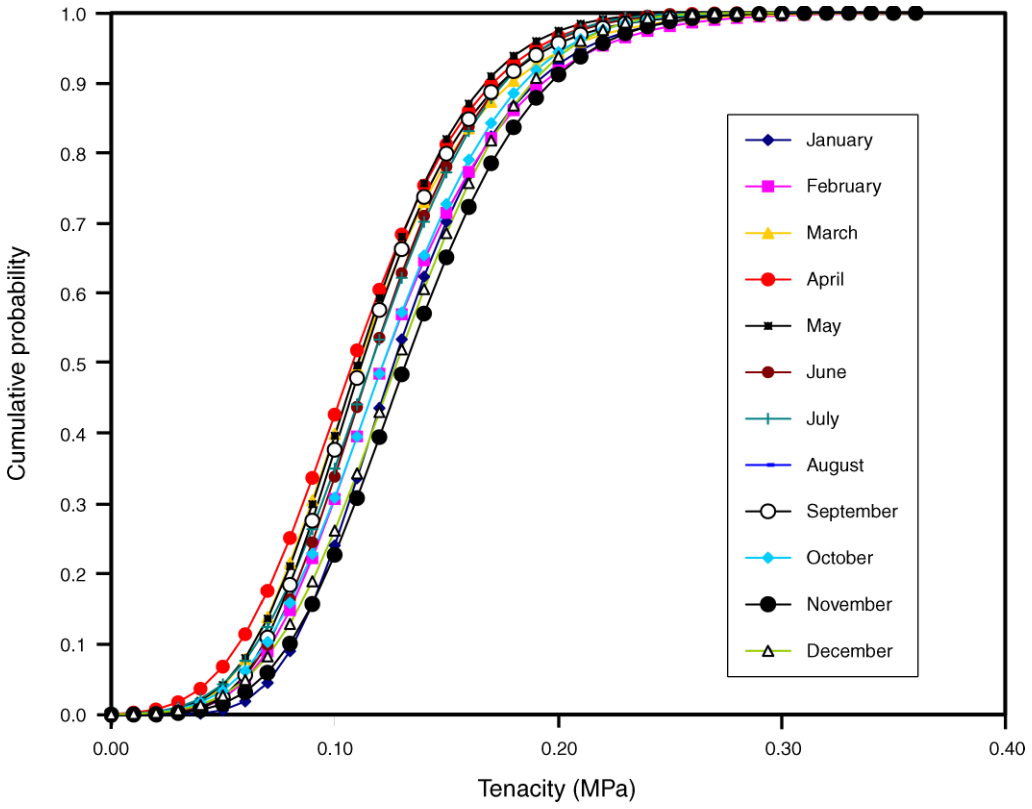

FIG. 9. Month-specific cumulative probability distributions for *Mytilus californianus* at Hopkins Marine Station in relation to tenacity. Coefficients describing these curves are given in Table 2.

threshold is greater for gradual returns to seawater temperature than for abrupt returns.

Extreme thermal events

Table 3 presents a summary of the results of the heat budget model, and Fig. 16 shows annual maximum body temperatures as a function of return time and substratum orientation. Of the orientations tested, temperature reaches its maximum average on the south-facing surface with a slope of 45° (Fig. 16A). This angle is near the latitude of HMS (36°), thus placing the substratum nearly perpendicular to the midday sun. Temperatures on a horizontal surface are similar to temperatures reached on an east-facing 45° surface. Temperatures on vertical surfaces follow the same pattern with orientation as substrata sloped at 45°, but are slightly cooler (Fig. 16B).

Return times for death of given fractions of limpets are shown in Fig. 17. For the worst case of our spatial scenarios (a south-facing 45° surface), if a given maximum body temperature is associated with an abrupt cooling to seawater temperature, half the population would be killed with a return time of ~8.1 yr (Fig. 17A). If, however, a given maximum body temperature is associated with a gradual return to seawater temperature, half the population would be killed with a return time of only ~1.6 yr (Fig. 17B). In contrast, if the 45° surface faces north and a given maximum body temperature is associated with a gradual return to seawater temperature, temperatures sufficient to kill half the population occur only every 209 yr (Fig. 17B). Return times are longer regardless of substratum orientation if cooling is abrupt (Fig. 17A, C) than if cooling is gradual (Fig. 17B, D). There was a slight effect on our calculations of varying block size, but no apparent pattern to the effect (Fig. 18): results using a

TABLE 2. Coefficients describing the cumulative probability distribution of mussel strength (see Eq. 14).

Month	α	β	ε
January	0.0444	0.0688	0.1131
February	0.0475	0.0622	0.1068
March	0.0430	0.0469	0.0966
April	0.0552	0.1561	0.0934
May	0.0528	0.1607	0.0970
June	0.0516	0.1410	0.1030
July	0.0606	0.1980	0.1019
August	0.0576	0.2111	0.0907
September	0.0422	0.0651	0.0991
October	0.0609	0.1792	0.1068
November	0.0598	0.1553	0.1169
December	0.0672	0.2204	0.1128

Note: Values are used in Eq. 3, with x, the applied stress, measured in MPa.

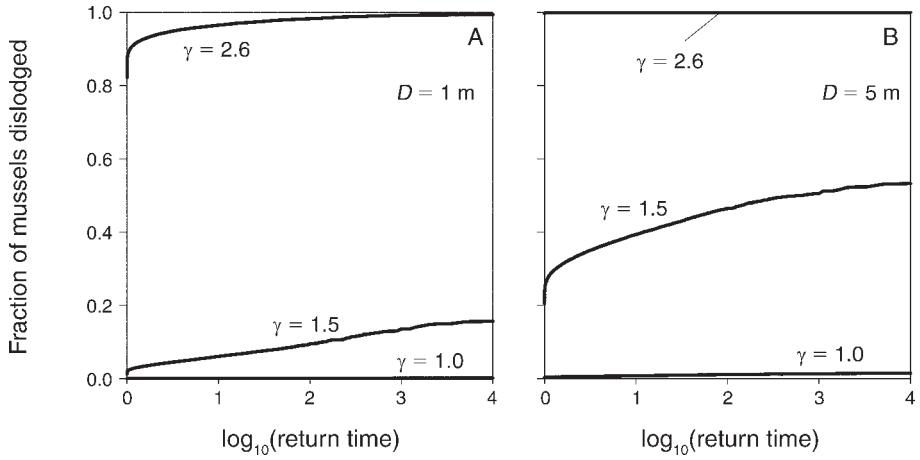

FIG. 10. The proportion of mussels dislodged over time. The longer one waits, the larger the fraction of mussels dislodged, but the increase is relatively small. (Return times are in years.) Data are shown for water depths of (A) $D = 1$ m and (B) 5 m, and for velocity amplification γ varying from 1 to 2.6.

block length of 39 days were very similar to those using the standard length of 13 days.

DISCUSSION

Our results suggest that the co-occurrence of normal environmental factors can lead to extreme events, and that these environmental extremes can cause severe disturbance. For example, the average yearly maximum body temperature of *L. gigantea* on horizontal surfaces is 30.4°C, sufficient to kill <5% of individuals (Table 3). However, on one occasion in our 1000 simulated years, normal air temperatures, tidal levels, solar irradiances, and wind speeds aligned to produce an estimated body temperature of 40.1°C, sufficient to kill >99% of limpets at this orientation, an event that would open substantial primary space for invasion or settlement. Similar effects occur at most other substratum orientations, north-facing walls being the only reliable refuge from thermal disasters (Table 3). It is important to note that the environmental event chronicled here for limpets on a horizontal surface, a shift from benign to lethal, is a consequence of chance alone: in the year that produced the exceptional body temperature of 40.1°C, the mean values of air temperature, tidal level, solar irradiance, and wind speed were exactly the same as in all other years. By calculating the probability of rare events of this sort (a 1-in-1000-yr occurrence), our resampling method provides information about thermal dangers

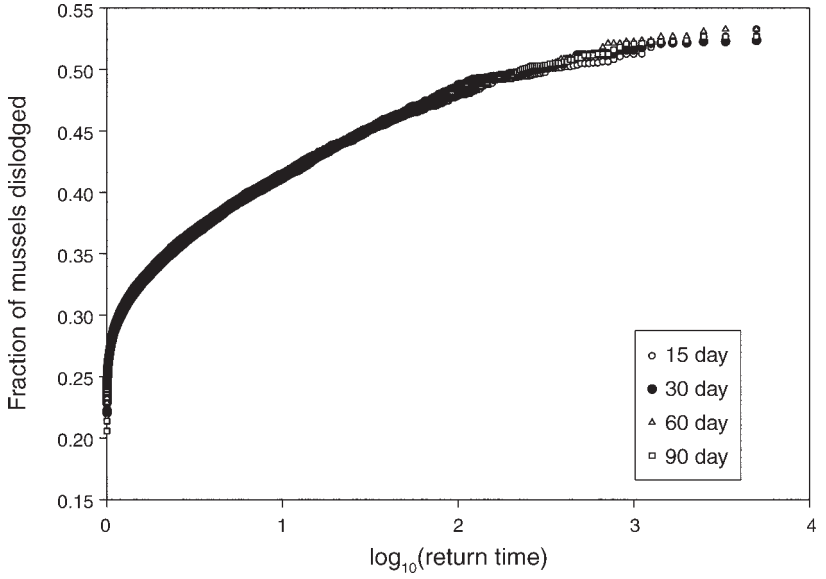

FIG. 11. Varying the length of segments in the bootstrap resampling has little effect on predicted mussel dislodgment. (Return times are in years.) Data shown here are for $D = 5$ m and a velocity amplification of 1.5. The time periods shown in the key are the segment lengths (of standardized residuals) used in the bootstrap resampling procedure.

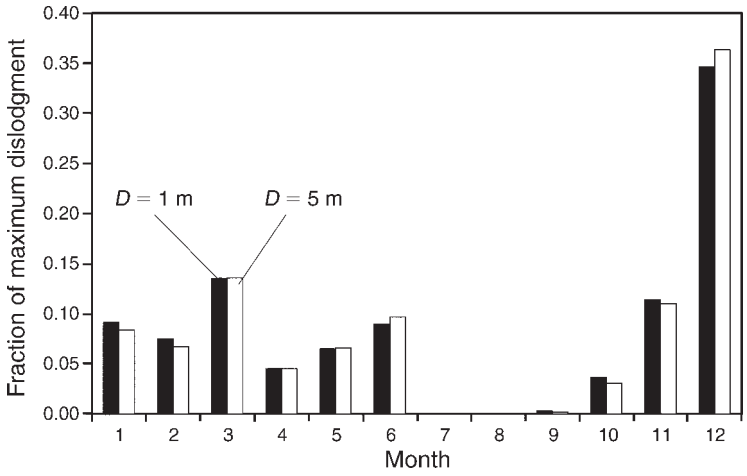

FIG. 12. Annual maximum mussel dislodgment occurs most often in December and least often in July and August. Results for water depths of $D = 1$ and 5 m are similar. Month 1 is January.

that would be difficult, if not impossible, to obtain through short-term observations.

Our results also provide details regarding the spatial patterns of environmental factors that would be difficult to obtain otherwise. For example, casual observation at HMS shows that *L. gigantea* are relatively scarce on south- and west-facing substrata. Denny and Harley (2006) have shown that in typical years maximum body temperature is highest for these orientations, but not sufficiently high to kill *L. gigantea*. Why, then, are limpets scarce at these orientations? A potential explanation becomes apparent only when rare, extreme thermal events are considered. We predict that events capable of killing owl limpets occur much more frequently on south- and west-facing slopes than on north- or east-facing slopes (Fig. 17). Perhaps it is these rare events, events that would be unlikely for an ecologist to observe, that control limpets' distribution.

The pattern of disturbance is different for *M. californianus* subjected to hydrodynamic forces. At any particular site on the shore, temporal variation in applied stress spans only a small fraction of the broad range of mussel attachment strengths. Mussels at relatively protected sites on the shore (sites with a velocity amplification = 1) are exceedingly unlikely to be dislodged by even a once-in-10 000-yr wave/tide event (Fig. 10). Mussels at highly exposed sites ($\gamma = 2.6$) are unlikely to survive even the "normal" wave events that occur every year. For sites with intermediate velocity amplification ($\gamma = 1.5$), extreme wave events can indeed increase the fraction of mussels dislodged, but only within a narrow range. Thus, for mussels, physics (in the form of wave breaking) limits the range of environmental extremes possible at a given site, which, when coupled with the broad distribution of breaking stress in mussels (Fig. 9), in turn limits the range of possible ecological consequences. In contrast to limpets, for which the intensity of thermal disturbance can vary drastically in both space and time, the intensity of hydrodynamic disturbance for mussels can vary drastically from place to place on a shore, but disturbance at any given location is unlikely to vary as drastically through time.

What if?

The resampling method employed here allows us to answer a variety of "what if?" questions. In central California, where this study was conducted, summertime low tides typically occur early in the morning, while in Oregon and Washington, they can occur near midday, leading to the proposal that intertidal organisms are subjected to greater thermal stress at higher latitudes (reviewed in Helmuth et al. 2002, 2006). We can test one aspect of this hypothesis by running our model using the empirically measured air temperature, solar irradiance, wind speed, and sea surface temperature at HMS combined with tidal fluctuations that are shifted to later in the day (Fig. 19). A 3-h shift in the tides (simulating a

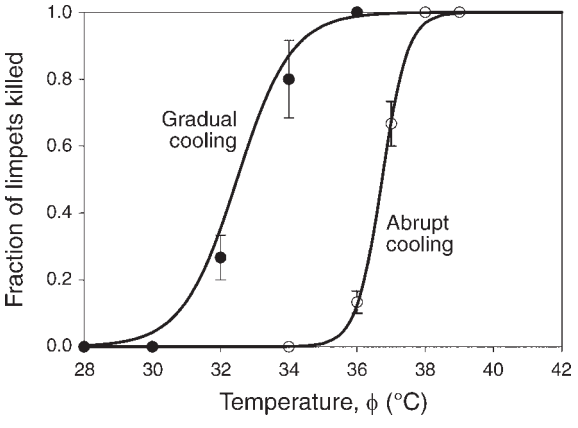

FIG. 13. Thermal limits for the limpet *Lottia gigantea*. Data are shown for two heating regimes: abrupt cooling (in which temperature is elevated for 3.5 h, 1.75 h at >28°C) and gradual cooling (in which temperature is elevated for 7 h, 3.5 h at >28°C). Error bars represent ±SE.

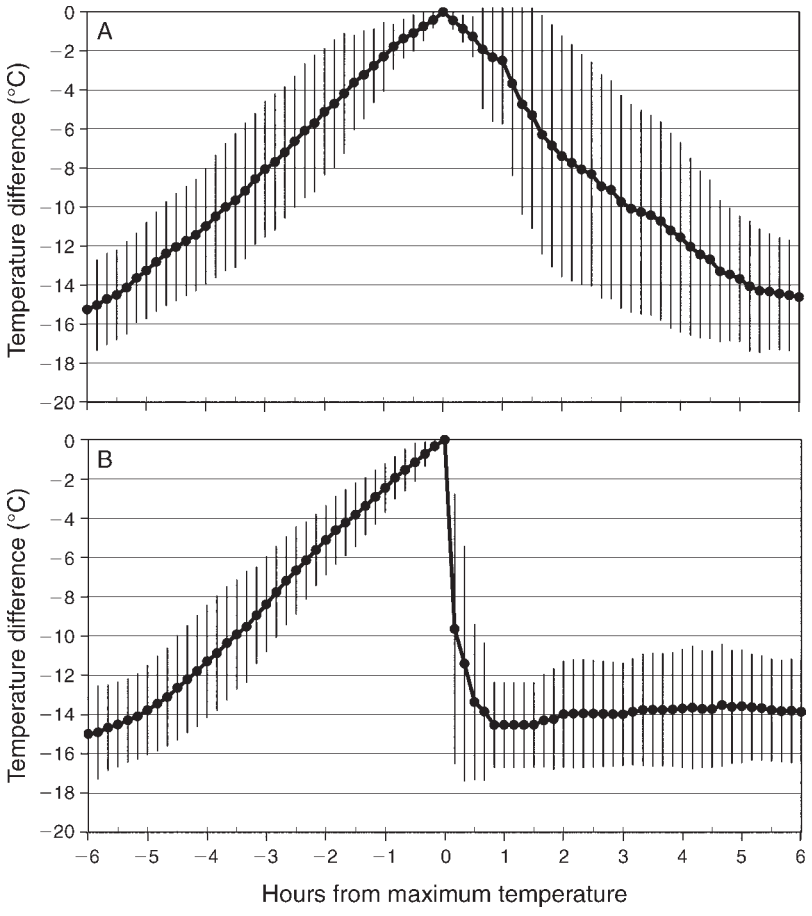

FIG. 14. Average patterns of heating and cooling observed for 184 yr of resampled environmental data: (A) gradual cooling and (B) abrupt cooling. Vertical bars indicate ±SD.

northward shift from California to Washington State) decreases the predicted return time of 50% thermal death from 8.8 to 6.2 yr, a surprisingly small shift. This finding reinforces the conclusion of Helmuth et al. (2002) that there may be little effective latitudinal variation in intertidal body temperatures.

What if air temperature were to rise, say, 2°C from those values measured at HMS? Increasing air temperature reduces the return time of lethal thermal events, but in a pattern that depends on orientation (Fig. 20). Only minor effects are predicted for south-facing 45° slopes (Fig. 20A) and north-facing vertical walls (Fig. 20B). In contrast, on horizontal surfaces, an increase in air temperature of 2°C would have substantial effect, reducing the return time for an event killing 50% of limpets from 9 to 2 yr (Fig. 20C). An increase in air temperature could thus affect the spatial distribution of limpets, and thereby the surrounding community.

Note that the uniform increase of air temperature used to calculate these results is at best a first approximation of the effect of increased average air temperature at HMS. If average air temperature rises as a part of global climate change, the effect is likely to vary with time of day and time of year. Furthermore, any shift in average air temperature is likely to be accompanied by other changes in the thermal environment. Increased air temperature in the inland valleys east of HMS is likely to increase the incidence and speed of onshore winds,

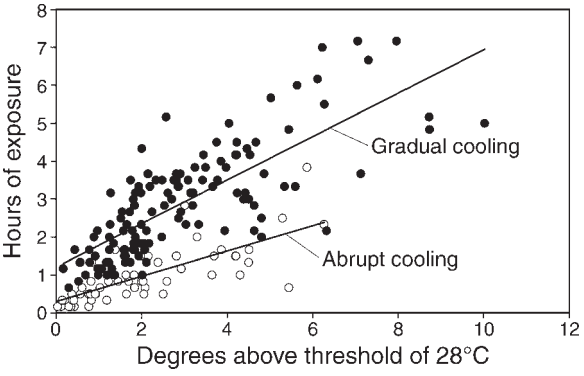

FIG. 15. Hours of exposure to potentially lethal temperatures (>28°C) for 184 yr of resampled environmental data. Key: solid symbols, gradual cooling; open symbols, abrupt cooling. Lines are least-squares regressions: gradual cooling, hours = $1.213 + 0.573 \times$ degrees ($r^2 = 0.593$, $n = 127$); abrupt cooling, hours = $0.302 + 0.333 \times$ degrees ($r^2 = 0.553$, $n = 57$).

TABLE 3. Mean and maximum predicted body temperatures and the estimated fraction of limpets killed, as a function of substratum orientation and "gradual" or "abrupt" cooling regimes.

Orientation	Mean (°C)	Fraction of limpets killed: Abrupt cooling	Fraction of limpets killed: Gradual cooling	Maximum (°C)	Fraction of limpets killed: Abrupt cooling	Fraction of limpets killed: Gradual cooling
Horizontal	30.36	0.0000	0.0422	40.11	0.9998	0.9999
North 45°	26.01	0.0000	0.0002	36.73	0.5000	0.9923
East 45°	30.30	0.0000	0.0393	38.03	0.9666	0.9985
South 45°	33.77	0.0005	0.7601	42.99	>0.9999	>0.9999
West 45°	32.06	0.0000	0.2709	40.84	>0.9999	>0.9999
North vertical	22.43	0.0000	0.0000	32.47	0.0000	0.3831
East vertical	28.47	0.0000	0.0041	34.69	0.0051	0.9094
South vertical	31.66	0.0000	0.1837	43.08	>0.9999	>0.9999
West vertical	30.83	0.0000	0.7365	38.18	0.9771	0.9987

resulting in an increase in fog at the coast and a concomitant decrease in solar irradiance. Thus, in reality an increase in air temperature might actually result in a decrease in intertidal body temperature. Although these complications are not included here, if the change in solar irradiance (or any other factor that might vary with air temperature) can be predicted, its incorporation into our method is straightforward.

Return time: nonintuitive consequence

In the discussion above, we quantified the temporal pattern of extreme events in terms of the return time. As with many averages, return time must be interpreted with care, and it is worth taking a few moments to consider return time in detail.

If, in the absence of any long-term shift in the environment, the stochastic components of environmental factors combine randomly through time, we can suppose that compound extreme events occur with constant, small probability. If this is indeed the case, the intervals between extreme events in a static climate should conform to a Poisson interval distribution (for a derivation, see Berg [1983:87]):

$$p(t;R)dt = \frac{e^{-(t/R)}}{R}dt$$
$$\text{Prob}(x \le t) = P(x) = 1 - e^{-(t/R)}. \quad (18)$$

Here, $p(t; R)$ is probability density for a given interval length t given return time R, and $P(x)$ is the probability that the actual interval between events is $\le t$. Eq. 18 is graphed in Fig. 21. Indeed, the distribution of interevent intervals predicted by our method is accurately modeled by the Poisson interval distribution; an example is shown in Fig. 22.

Note that the Poisson interval distribution is highly skewed, with an abundance of short intervals and a few very long intervals (Fig. 21). This skew can lead to nonintuitive biological consequences. We use two hypothetical examples to demonstrate our point. Let $N(x)$ be the probability of not encountering in any particular year an extreme event of magnitude x. If $N(x)$ is constant, the probability of not encountering an extreme of magnitude x in a lifetime of L yr is $N^L(x)$. In other words, measured over many generations, $N^L(x)$ is the average fraction of the population that avoids extreme event of magnitude x. An example is shown by the solid line in Fig. 23 where we

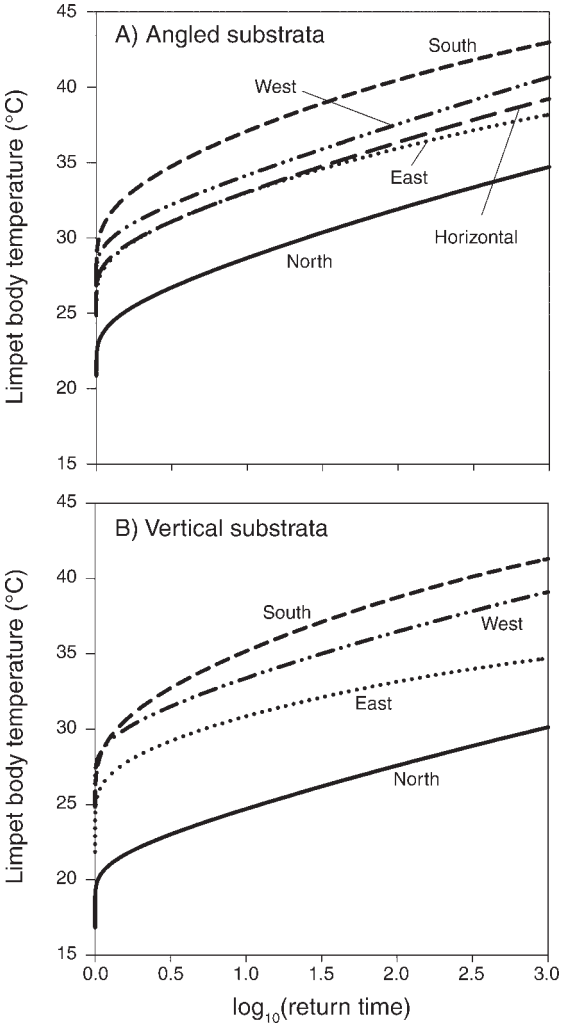

FIG. 16. Annual maximum body temperature as a function of return time for substrata of various orientations. (Return times are in years.) Angled substrata were inclined 45° to horizontal.

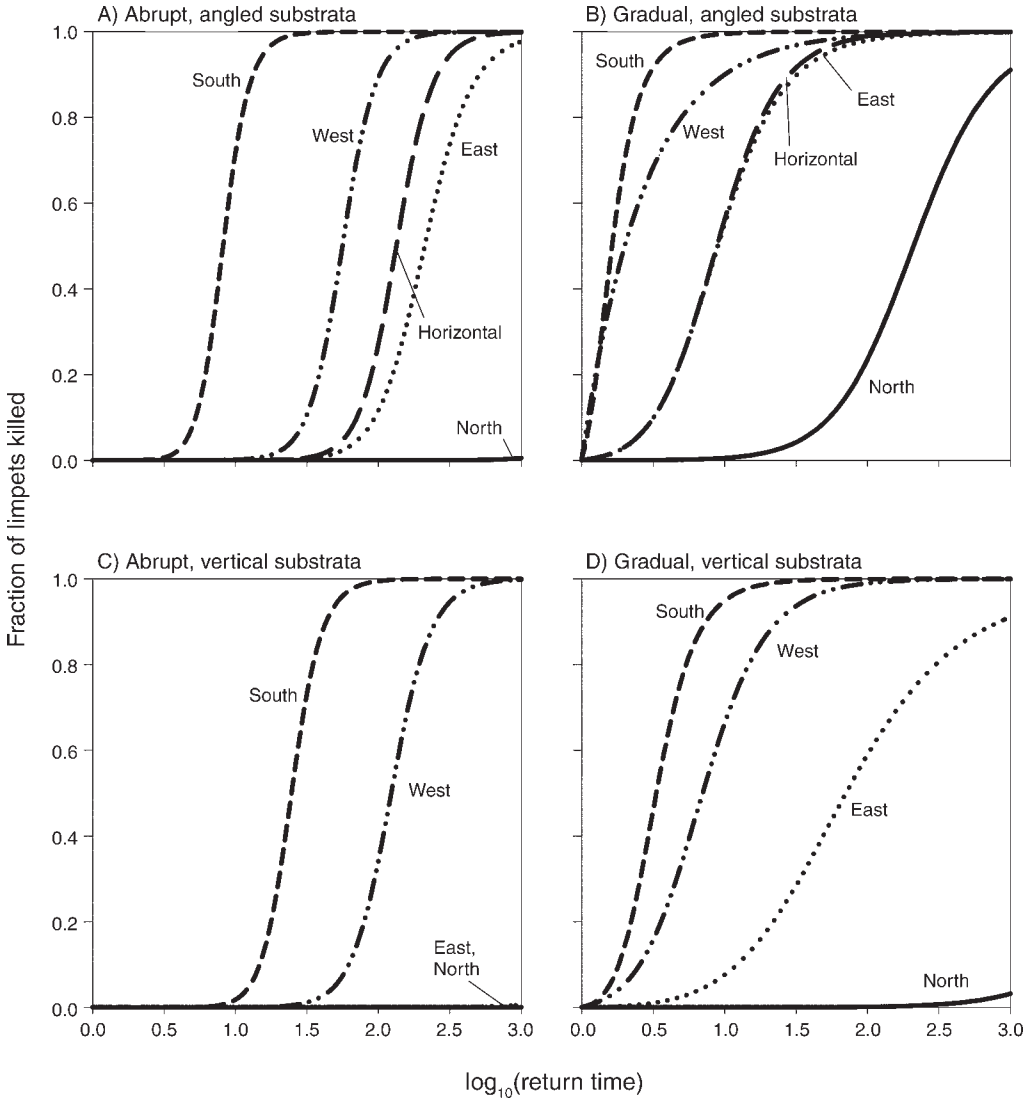

FIG. 17. Annual maximum fraction of limpets killed as a function of return time (in years) and heating regime (abrupt cooling vs. gradual cooling) for substrata of various orientations. Angled substrata were inclined 45° to horizontal.

assume that $N(x) = 0.95$, corresponding to a return time of 20 yr (Eq. 15). Because extreme events occur randomly, there are instances in which the interval between events is exceptionally long (the tail on the right side of Fig. 21B). As a consequence, even if individuals have a lifetime equal to the return time, in the long term, 36% of these individuals never encounter a catastrophic event.

Contrast this scenario with another in which extreme events arrive periodically with a fixed interval R_{per}. In this case, individuals born within $R_{per} - L$ yr immediately after an extreme event die blissfully of old age before the next catastrophe occurs. The fraction of individuals avoiding extreme events is $(R_{per} - L)/R_{per}$; it decreases linearly with increasing life span. This scenario is also depicted in Fig. 23 (the dashed line), where we have assumed that R_{per} (like R) is 20 yr. More individuals survive if extreme events arrive randomly than if they arrive periodically.

The average fraction of a population that encounters an extreme is only one index of an extreme's effect on a population, however. What is the effect of extreme disturbance on the average population size? To address this question, we consider a hypothetical population growing according to the logistic equation (see Appendix D). The population is subjected to extreme events that arrive either at random or periodically at intervals equal to the return time. If disturbances arrive randomly in time, they may by chance be applied in quick succession. This sort of "double whammy" reduces population size to a level from which it can only slowly recover, and as a result, population size (averaged over 10 000 intervals) is smaller than if disturbance is applied periodically at the average rate (Fig. 24). The disparity

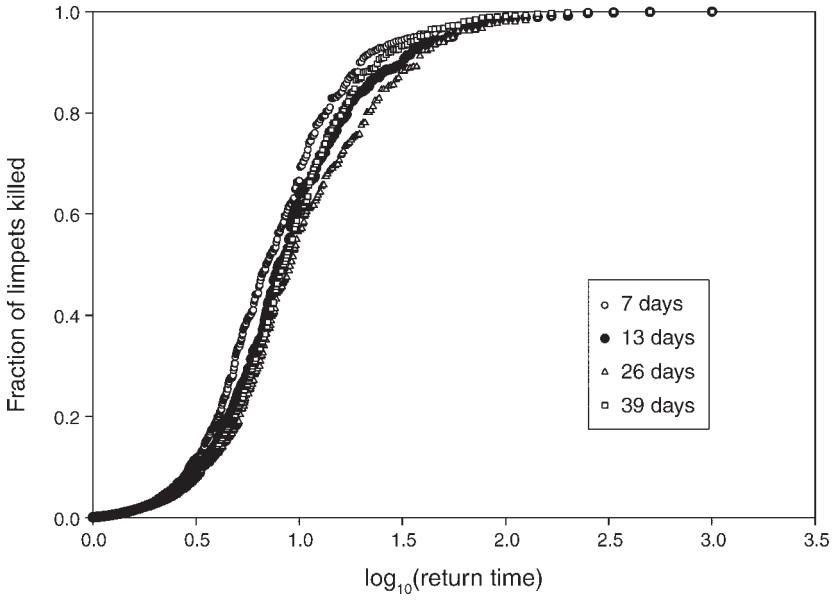

FIG. 18. The effects of varying segment length on the predicted fraction of limpets killed. Data are for limpets on a horizontal surface 1.5 m above MLLW (mean lower low water). (Return times are in years.)

of consequences between random and periodic events is small for small disturbance intensities, but large if disturbances are drastic. Examples of a similar effect have been noted in nature: Paine et al. (1998) show that two extreme events occurring in unusually quick succession can result in ecological surprises that would not accrue if the same events were repeated at longer or more consistent intervals. In summary, the skewed shape of the Poisson interval distribution of extreme events can lead to nonintuitive biological consequences.

In addition to providing insight into the interpretation of return time, Eq. 18 can be used to evaluate our method's ability to predict return times. For extreme wave/tide events with individual variables similar to those of January 2008, we predict a return time of ~54 yr, while the sole observed interval between events of similar magnitude was a mere 7 yr. However, even given a return time of 54 yr, the skew of the Poisson interval distribution is such that 12.2% of intervals are $\leq$7 yr. Thus, although our estimated return time seems high, we cannot reject (at the traditional 5% level) the hypothesis that our prediction is consistent with our single measured return interval.

In the same fashion, Eq. 18 can help us discriminate between extreme events that are due to chance alone in a constant climate and those that are due to shifts in climate itself. For example, if two extreme thermal events were imposed on limpets in quick succession, it might be tempting to attribute the occurrence to global warming. Because our method allows us to estimate the expected return time of such events in the absence of climate change, we can use Eq. 18 to test the validity of such an assertion. Unless there is a very low likelihood that back-to-back events occurred by chance alone, it would be inappropriate to invoke climate change as the definitive cause. Thus, our ability to legitimately resample a short time series of environmental factors can serve as a useful tool for discerning effects of long-term climate shifts.

Predictions vs. forecasts

It is important to distinguish the predictions made here from forecasts of what will happen in the future. Our method estimates how often, on average, a given extreme event would occur if one particular year could be repeated 1000 or 10 000 times. For example, we predict maximum hydrodynamic stress using the same

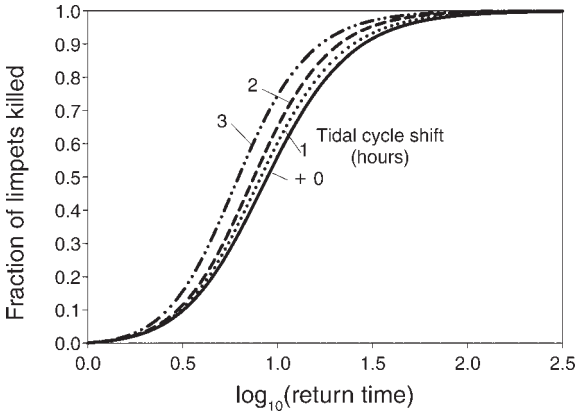

FIG. 19. Shifting the tidal cycle to later in the day increases the fraction of limpets killed by thermal stress. (Return times are in years.) The tide is shifted to later times by 1, 2, and 3, hours. Calculations presented here are for limpets on horizontal substrata 1.5 m above MLLW.

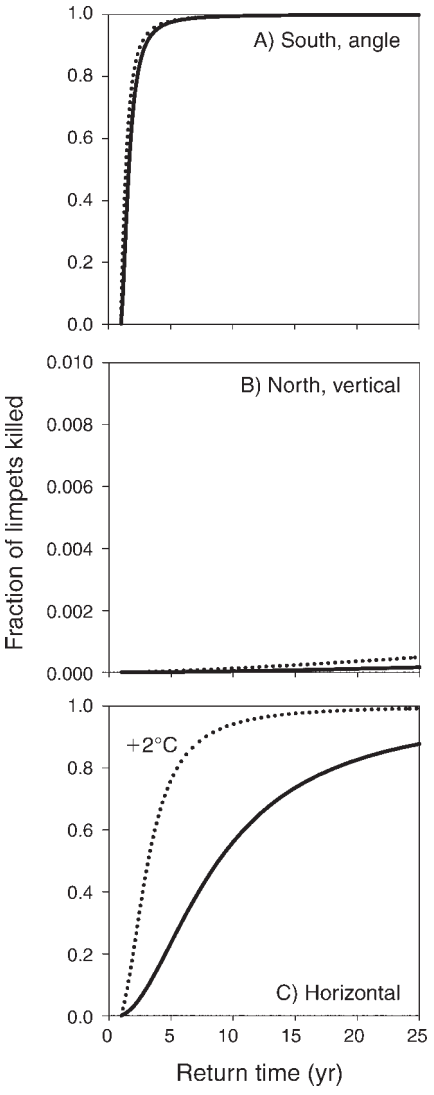

FIG. 20. The effect of a 2°C increase in air temperature depends on the orientation of the substratum. There is little effect for limpets on (A) south-facing, 45° substrata or (B) north-facing vertical substrata, but there is a substantial effect for (C) limpets on horizontal surfaces. In each panel, the solid line indicates mortality for the base air temperature, and the dotted line represents that for a 2°C increase in air temperature.

annual pattern of tidal variation again and again. But we know that in reality tidal fluctuations vary from year to year, e.g., the lunar declination varies with a period of 18.6 yr, affecting the amplitude and timing of the highest predicted tides (e.g., Denny and Paine 1998). Thus, even if all other aspects of the environment remain the same as recorded by our 7-yr time series, the actual pattern of extreme wave events in the future will vary somewhat from the predictions made here due to the changing pattern of the tides.

This limitation can easily be overcome. Because long-term fluctuations in tides are accurately predictable from celestial mechanics, we could extend our method to take future fluctuations into account, combining our normalized tidal deviations with predicted celestial tides for the duration of the forecast. Predicted future variation in other factors can be incorporated in similar fashion.

We need better physiological models

The predictions made here of dislodgement and thermal death are only as good as the empirical data on which they are based, and these data have limitations. For example, our calculations of mussel breaking stress do not incorporate any intermussel influence. In mussel beds, once one mussel is dislodged, others nearby may be weakened, leading to the formation of local patches of disturbance, and overall greater disturbance than we have estimated. Similarly, it is currently unclear how

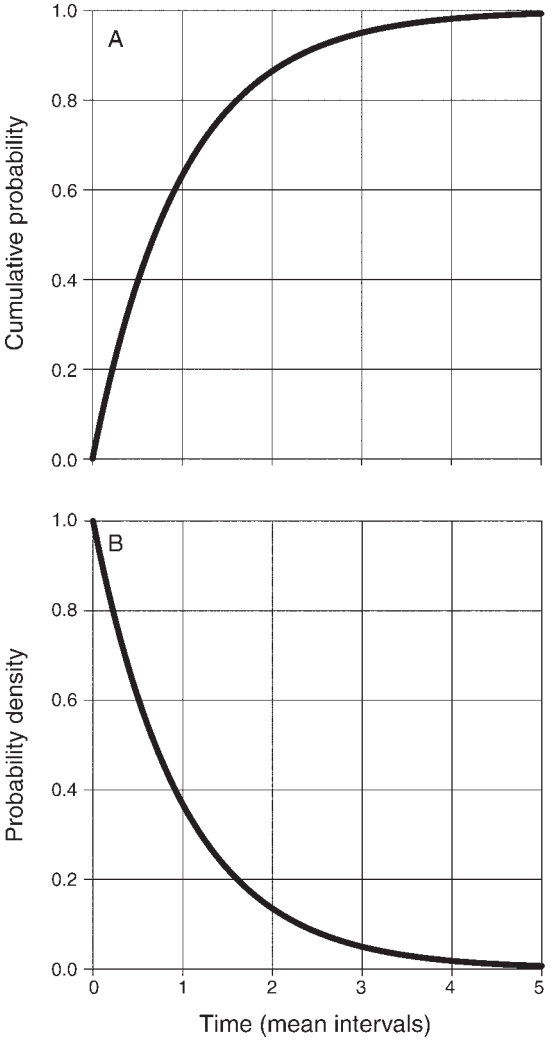

FIG. 21. The Poisson interval distribution. (A) The fraction of intervals less than or equal to a given value; values are measured in units of the return time. For example, 63% of intervals have lengths less than of equal to the return time. (B) The probability density function for the Poisson interval distribution. The probability of encountering an interval of a given length is greatest for intervals of short duration.

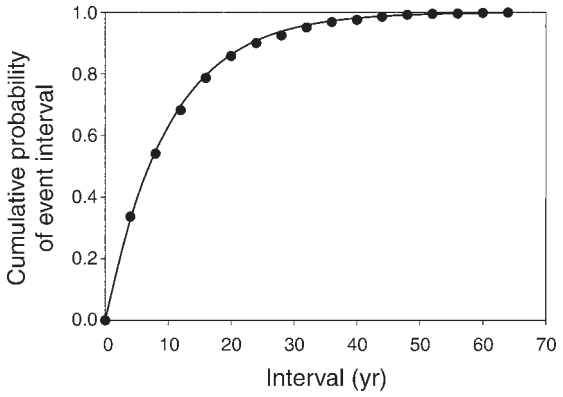

FIG. 22. Cumulative probability of predicted intervals between mussel dislodgment events matches the predictions of the Poisson interval distribution. The 10 000-year random realizations of hydrodynamic stress were concatenated to simulate a long times series, and the intervals between events of a given severity were noted. Circles are data for $D = 5$ m, velocity amplification = 1.5, and stress events sufficient to dislodge 41.6% of mussels. These events have an observed return time of 10.05 yr. The solid line is the cumulative Poisson interval distribution (Eq. 18) calculated for a return time of 10.05 yr.

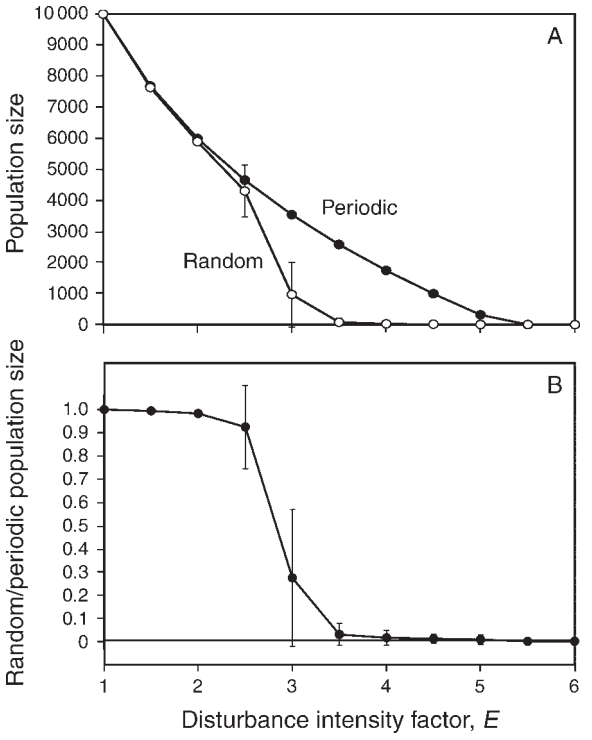

FIG. 24. The effect of extreme events on the average size of a hypothetical population undergoing logistic growth. We subjected the population to occasional extreme events such that the size of the population after the event is $1/E$ of the size of the population before the event, where E is a disturbance intensity factor. Average population size is smaller when subjected to events that arrive at random than when the population is subjected to events that arrive periodically. Population size shown here is averaged over a length of time sufficient to contain ~500 extreme events. Error bars represent ±SD.

velocity amplification γ typically varies along a shore, and therefore how it contributes to patch formation. Further research is needed to elucidate these potential effects. More problematic is our current level of physiological insight. Our estimates of disturbance (both hydrodynamic and thermal) depend strongly on empirical measurements of physiology, and existing measurements are, at best, preliminary. For example, our measurements of the attachment strength of *M. californianus* are by far the most complete of any to date. Several thousand mussels were dislodged in the course of 7 yr to estimate the month-by-month distribution of breaking stress. However, we assume that it is the single, largest wave-induced stress that dislodges mussels. It is possible that repeated submaximal stresses may have the same effect. In addition, there is some evidence in our data (E. Carrington and M. W. Denny, *unpublished data*) that attachment strength at one time is correlated with average wave stress in the week prior to measurement,

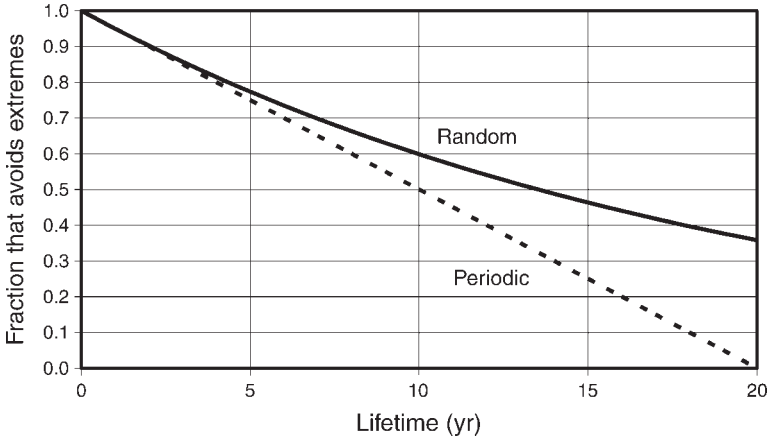

FIG. 23. Results from a hypothetical population. If lethal events arrived randomly (with a probability of 0.05/yr) rather than at a fixed interval of 20 yr, an increased fraction of the population avoids being killed before dying of old age.

PLATE 1. Large waves breaking on rocky shores can impose severe hydrodynamic forces. Photo credit: C. D. G. Harley, taken at Hopkins Marine Station, California, USA.

and there is evidence that the susceptibility of mussels to dislodgment varies with time since they were last disturbed (Paine 2002). This sort of temporal autocorrelation in breaking stress has not been well characterized, however, and is not included in our calculations.

Clearly, the more accurate our understanding of the physiological basis for mussel attachment strength, the more accurate our calculations of extreme events. In effect, what we need is a mechanistic model of the entire physiology underlying adhesive tenacity. Such a model would allow us to interpret any time series of applied stresses in terms of its biological consequences.

Our understanding of the thermal tolerances of limpets is even more limited. We measured thermal tolerances on relatively few individuals for only two heating regimes, and did not measure tolerance as a function of season. Other intertidal species exhibit seasonal variation in the physiology of thermal tolerance (Roberts et al. 1997, Halpin et al. 2002, Sagarin and Somero 2006), and it is possible that *L. gigantea* can adjust its thermal tolerance in ways that we have not taken into account. And, as with mussels, we have assumed that it is maximum stress that kills limpets. It is possible that lengthy or repeated submaximal temperatures can have the same effect. Furthermore, our calculations do not take into account variations in shell shape among limpets (which can affect the transfer of heat by convection, and thereby, body temperature), the effects of microhabitat shading on realistic substrata, and the fact that some actual exposures to high body temperatures last for longer periods than those used in our laboratory experiments (Fig. 15). Until more is known about the mechanism(s) of thermal death in limpets, these potential effects cannot be incorporated into our calculations. Again, what we need is a detailed mechanistic model, in this case of limpet thermal physiology.

And finally, if we are to fully address the ecological effects of extreme events, it would be advantageous to incorporate all possible types of extremes. For example, as we have noted, *L. gigantea* is immune to hydrodynamic dislodgment when stationary. However, the limpet is highly susceptible when actively foraging (Denny and Blanchette 2000). Given sufficient information regarding limpet foraging behavior, we could predict return times of limpets' extreme hydrodynamic events for comparison with those of mussels, but this behavioral data is currently unavailable. Similarly, given an accurate model for mussel thermal physiology, we could predict the return times of thermal death for mussels for comparison with those of limpets. Helmuth (1998) formulated a heat budget model for *M. californianus*, but the thermal limits of this species have not been measured.

Our intent in noting these caveats is not to denigrate our findings, which, although not perfect, are by far the best available. Instead, our intent is to raise awareness of the need for, and the utility of, greater interaction among physiologists, ecologists, and evolutionary biologists. This interaction may be facilitated by recent advances: the genome of *L. gigantea* is currently being sequenced and annotated.

Details and caveats

Although the statistical approach we describe here is based on the standard moving block bootstrap, the coordinated sampling among variables seems to be a new innovation. A thorough theoretical assessment of our method and its potential limitations and caveats will require the attention of the statistics community. In the meantime, we address two potentially confusing aspects of our resampling approach.

1) Have we missed important information because our empirical records are short (7 yr)? The logic behind this question is as follows. It is unlikely that a short time series will capture the most extreme residuals possible in any particular factor, and in the absence of these extreme residuals, resampling the input data might never yield absolute maximal output of a given variable. Thus, due to the limited time series on which they are based, the extreme values calculated by our method may be conservative. Our simulation of hydrodynamic forces on mussels (Fig. 4) provides a hint of this effect. In this simulation, the distribution of standardized residuals is unbounded for each variable, making it probable that the longer the time series, the larger the residuals encountered. Indeed, predictions from a short (7-yr) data set underestimate the rarest stresses found in the long-term series. The magnitude of underestimation is quite small, however, ($<4\%$) even for an event that occurs once in 10 000 yr.

Whether our method will similarly underestimate extremes when applied to real data depends on two factors. First, it depends on whether or not it is the extremes of the various factors that govern the extremity of the biological consequences. As we have seen, this is not the case for two of the variables contributing to wave-induced hydrodynamic forces. Waves with heights above the breaking limit impose no more force than smaller waves; so as long as our 7-yr series captures the statistical behavior of waves up to the breaking limit, larger values from a longer series would be irrelevant. Similarly, for most situations it is an intermediate wave period, rather than the extreme, that results in the highest hydrodynamic stress. Again, as long as our 7-yr series captures the statistics of these intermediate values, obtaining a longer series has no advantage. Thus, the accuracy of our method in predicting return times for long-term simulated wave and tide data may be due in large part to the fact that hydrodynamic stress is not governed solely by the extremes of wave height and period.

If extreme ecological consequences do depend on the extremes of each individual factor, the degree to which our method will underestimate reality depends on the characteristic of that extreme portion of the distribution of standardized residuals *not* included in the empirical data. In some cases, we know that the tails are physically bounded: wave heights and wind speeds cannot be lower than 0 and irradiance is limited by the output of the sun. In other cases, estimating the shape of this "missing tail" is a task for univariate extreme-value analysis, and we present a preliminary analysis of our data in Appendix E. In most cases, the standardized environmental residuals have distributions with definable upper limits that are not far in excess of values recorded in our short time series, suggesting that we have not missed important information because of the limited length of our empirical data.

2) Results presented here for limpets and mussels document the extreme consequences encountered in 1000 and 10 000 random realizations of a particular year, respectively. How much more extreme might these values be if we created more realizations? Is it possible to calculate the absolute extreme that one can obtain by resampling a finite set of standardized residuals?

In Appendix F, we present a method for extrapolating to maxima beyond the realizations in hand, yet another application of univariate extreme-value analysis. However, for the phenomena we deal with here (hydrodynamic stress, body temperature), this exercise has little practical value. For example, in our 1000-year-long realizations of limpet body temperature, the highest temperature recorded for any substratum orientation was 43.1°C. If there is a definable absolute maximum temperature, it must be higher still. But 43.1°C by itself is sufficient to kill all limpets, so any higher temperature can have negligible biological effect. However, the situation may be different in other systems. If in a given system the maximum stress encountered in a practical number of realizations is not sufficient to have biological consequences, it may be reasonable to ask whether stress could ever be sufficiently severe. In a case such as this, the techniques described in Appendix F may be useful.

Conclusions

Our statistical method for resampling a short-term time series provides a practical mechanism for estimating the return time of those extreme ecological events that owe their existence to the chance co-occurrence of normal factors. Knowledge of these return times augments our understanding of ecology and evolution. In the case of mussels, for instance, extreme hydrodynamic events are unlikely to have much ecological effect: the variation in disturbance is likely to be much greater from place to place than from time to time, and this temporal stability may play an important role in the evolution of community dynamics and life-history strategies in the mid-intertidal zone. In the case of limpets, spatial variation in thermal stress is important, but temporal variation in thermal disturbance can be equally important. These findings provide a useful historical perspective for existing populations, and our resampling method provides a tool for predicting the effects of future environmental change. The method used here for mussels and limpets is applicable to virtually any system in which extreme ecological

disturbance can occur through the chance alignment of "normal" events.

Acknowledgments

We gratefully acknowledge the advice and assistance of Brit Turnbull and Brad Efron in formulating the basic resampling procedure. James Rohlf and Larry Hunt provided critical insight, and Eric Berlow suggested the simulation. We thank Paul Switzer for productive discussions regarding the limitations of multivariate extreme-value analysis and two anonymous reviewers for constructive comments. This study was funded by NSF Grant OCE-9985946 to M. Denny. This is Contribution 322 from PISCO, the Partnership for Interdisciplinary Studies of the Coastal Ocean, a consortium funded by the David and Lucile Packard Foundation and the Gordon and Betty Moore Foundation.

Literature Cited

Allee, W. C., A. E. Emerson, O. Park, T. Park, and K. P. Schmidt. 1949. Principles of animal ecology. W. B. Saunders, Philadelphia, Pennsylvania, USA.

Altman, S., and R. B. Whitlatch. 2007. Effects of small-scale disturbance on invasion success in marine communities. Journal of Experimental Marine Biology and Ecology 342: 15–29.

Berg, H. 1983. Random walks in biology. Princeton University Press, Princeton, New Jersey, USA.

Boller, M. L., and E. Carrington. 2006. The hydrodynamic effect of shape and size change during reconfiguration of a flexible macroalga. Journal of Experimental Biology 209: 1894–1903.

Breaker, L. C., W. W. Broenkow, and M. W. Denny. 2006. Reconstructing an 83-year time series of daily sea surface temperature at Pacific Grove, California. Moss Landing Marine Laboratory Technical Publication 06–2.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1171–1789.

Bühlmann, P., and H. R. Künsch. 1995. The blockwise bootstrap for general parameters of a stationary time series. Scandinavian Journal of Statistics 22:35–54.

Bühlmann, P., and H. R. Künsch. 1999. Block length selection in the bootstrap for time series. Computational Statistics and Data Analysis 31:295–310.

Carlstein, E. 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. Annals of Statistics 14:1171–1179.

Carrington, E. 1990. Drag and dislodgment of an intertidal macroalga: consequences of morphological variation in *Mastocarpus papillatus* Kützing. Journal of Experimental Marine Biology and Ecology 139:185–200.

Coles, S. T. 2001. An introduction to statistical modeling of extreme values. Springer-Verlag, London, UK.

Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. Science 199:1302–1310.

Darwin, G. H. 1962. The tides and kindred phenomena in the solar system. W. H. Freeman, San Francisco, California, USA.

Dayton, P. K. 1971. Competition, disturbance, and community organization: the provision and subsequent utilization of space in a rocky intertidal community. Ecological Monographs 41:351–389.

de Haan, L., and J. de Ronde. 1998. Sea and wind: multivariate extremes at work. Extremes 1:7–45.

Denny, M. W. 1987. Lift as a mechanism of patch initiation in mussel beds. Journal of Experimental Marine Biology and Ecology 113:231–245.

Denny, M. W. 1988. Biology and the mechanics of the wave-swept environment. Princeton University Press, Princeton, New Jersey, USA.

Denny, M. W. 1995. Predicting physical disturbance: mechanistic approaches to the study of survivorship on wave-swept shores. Ecological Monographs 65:371–418.

Denny, M. W., and C. A. Blanchette. 2000. Hydrodynamics, shell shape, behavior and survivorship in the owl limpet *Lottia gigantea*. Journal of Experimental Biology 203:2623–2639.

Denny, M. W., and S. Gaines. 2000. Chance in biology. Princeton University Press, Princeton, New Jersey, USA.

Denny, M. W., and C. D. G. Harley. 2006. Hot limpets: predicting body temperature in a conductance-mediated system. Journal of Experimental Biology 209:2409–2419.

Denny, M. W., B. Helmuth, G. H. Leonard, C. D. G. Harley, L. J. H. Hunt, and E. K. Nelson. 2004. Quantifying scale in ecology: lessons from a wave-swept shore. Ecological Monographs 74:513–532.

Denny, M. W., L. P. Miller, and C. D. G. Harley. 2006. Thermal stress on intertidal limpets: long-term hindcasts and lethal limits. Journal of Experimental Biology 209:2420–2431.

Denny, M. W., L. P. Miller, M. D. Stokes, L. J. H. Hunt, and B. S. T. Helmuth. 2003. Extreme water velocities: topographical amplification of wave-induced flow in the surf zone of rocky shores. Limnology and Oceanography 48:1–8.

Denny, M. W., and R. T. Paine. 1998. Celestial mechanics, sea-level changes, and intertidal ecology. Biological Bulletin 194: 108–115.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall/CRC Boca Raton, Florida, USA.

Folke, C., S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling. 2004. Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology and Systematics 35:557–581.

Gaines, S. D., and M. W. Denny. 1993. The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. Ecology 74:1677–1692.

Gaylord, B. 2000. Biological implications of surf-zone flow complexity. Limnology and Oceanography 45:174–188.

Gross, K. L., G. G. Mittelbach, and H. L. Reynolds. 2005. Grassland invasibility and diversity: responses to nutrients, seed input, and disturbance. Ecology 86:476–486.

Halpin, P. M., C. J. Sorte, G. E. Hofmann, and B. A. Menge. 2002. Patterns of variation in levels of Hsp70 in natural rocky shore populations from microscales to mesoscales. Integrative and Comparative Biology 42:815–824.

Helmuth, B. S. T. 1998. Intertidal mussel microclimates: predicting the temperature of a sessile invertebrate. Ecological Monographs 68:51–74.

Helmuth, B., B. R. Broitman, C. A. Blanchette, S. Gilman, P. Halpin, C. D. G. Harley, M. J. O'Donnell, G. E. Hofmann, B. Menge, and D. Strickland. 2006. Mosaic patterns of thermal stress in the rocky intertidal zone: implications for climate change. Ecological Monographs 76:461–479.

Helmuth, B., and M. W. Denny. 2003. Predicting wave exposure in the rocky intertidal zone: Do bigger waves always lead to larger forces? Limnology and Oceanography 48:1338–1345.

Helmuth, B., C. D. G. Harley, P. M. Halpin, M. O'Donnell, G. E. Hofmann, and C. A. Blanchette. 2002. Climate change and latitudinal patterns of intertidal thermal stress. Science 298:1015–1017.

IPCC. 2007. Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Katz, R. W., G. S. Brush, and M. B. Parlange. 2005. Statistics of extremes: modeling ecological disturbances. Ecology 86: 1124–1134.

Kinsman, B. 1965. Wind waves. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Koehl, M. A. R. 1977. Effects of sea anemones on the flow forces they encounter. Journal of Experimental Biology 69: 87–105.

Künsch, H. R. 1989. The jackknife and the bootstrap for general stationary observations. Annals of Statistics 17:1217–1241.

Longuet-Higgins, M. S. 1952. On the statistical distribution of the heights of sea waves. Journal of Marine Research 11:245–266.

Longuet-Higgins, M. S. 1980. On the distribution of heights of sea waves: some effects of nonlinearity and finite band width. Journal of Geophysical Research 85(C3):1519–1523.

Miller, L. P., C. D. G. Harley, and M. W. Denny. 2009. The role of temperature and desiccation stress in limiting the local-scale distribution of the owl limpet, *Lottia gigantea*. Functional Ecology, *in press*.

Munk, W. H. 1949. The solitary wave theory and its application to surf problems. Annals of the New York Academy of Sciences 51:376–424.

NOAA (National Oceanographic and Atmospheric Administration), National Ocean Service. 2000. Tidal datums and their application. NOS CO-OPS 1. U.S. Department of Commerce, Center for Operational Oceanographic Products and Services, Silver Spring, Maryland, USA.

Paine, R. T. 2002. Advances in ecological understanding: by Kuhnian revolution or conceptual evolution? Ecology 83: 1553–1559.

Paine, R. T., and S. A. Levin. 1981. Intertidal landscapes and the dynamics of pattern. Ecological Monographs 51:145–178.

Paine, R. T., M. J. Tegner, and E. A. Johnson. 1998. Compounded perturbations yield ecological surprises. Ecosystems 1:535–545.

Paparoditis, E., and D. N. Politis. 2003. Residual-based block bootstrap for unit root testing. Econometrica 71:813–855.

Roberts, D. A., G. E. Hofmann, and G. N. Somero. 1997. Heat-shock protein expression in *Mytilus californianus*: acclimatization (seasonal and tidal-height comparisons) and acclimation effects. Biological Bulletin 192:309–320.

Sagarin, R. D., and G. N. Somero. 2006. Complex patterns of expression of heat-shock protein 70 across the southern biogeographical ranges of the intertidal mussel *Mytilus californianus* and snail *Nucella ostrina*. Journal of Biogeography 33:622–630.

Sousa, W. P. 1979. Disturbance in marine intertidal boulder fields: the non-equilibrium maintenance of species diversity. Ecology 60:1225–1239.

SPSS. 1998. SYSTAT version 8.0. SPSS, Chicago, Illinois, USA.

U.S. Army Corps of Engineers. 1984. Shore protection manual. U.S. Government Printing Office, Washington, D.C., USA.

Vieira, N. K. M., W. H. Clements, L. S. Guevara, and B. F. Jacobs. 2004. Resistance and resilience of stream insect communities to repeated hydrologic disturbances after a wildfire. Freshwater Biology 49:1243–1259.

Zar, J. H. 1974. Biostatistical analysis. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

APPENDIX A

Details of the resampling method (*Ecological Archives* M079-014-A1).

APPENDIX B

Translating the generalized extreme value distribution (*Ecological Archives* M079-014-A2).

APPENDIX C

Simulated time series (*Ecological Archives* M079-014-A3).

APPENDIX D

Disturbance effects on logistic growth (*Ecological Archives* M079-014-A4).

APPENDIX E

Estimating the shape of tails (*Ecological Archives* M079-014-A5).

APPENDIX F

Estimating absolute maxima (*Ecological Archives* M079-014-A6).

Ecological Archives M079-014-A1

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix A. Details of the resampling method.

Predicted Annual Cycles. The first step in our resampling scheme is the determination of the "predicted" (= average, mean) signal. In a few cases, predictions can be made on a mechanistic basis. For example, celestial mechanics and harmonic analysis of past records allow one to make accurate predictions of tidal fluctuations at a given site (e.g., Darwin 1962). Aside from long-term trends due to changing sea level, deviations from these predictions are due largely to the unpredictable influence of variations in barometric pressure and the presence or absence of storm surges. However, for most environmental variables, we do not know the predictable signal *a priori* and we must estimate it from the time series itself. Several methods are available, and it is not clear that any one is best. We use the following method here.

Environmental variables such as wave height, air temperature, water temperature, wind speed, and solar irradiance vary in predictable fashion, with periodicities that depend primarily on earth's rotation about its axis and its revolution around the sun. Here, we assume that these daily and annual variations are the major periodicities for ecologically relevant variables, and we use these periodicities as the basis for calculating the mean values for each measurement time within a typical year. Given a sufficiently long record of measurements, calculating these expected values would be straightforward: e.g., the predicted H_S at 8:00 am for September 30 is the average of all H_S values measured at 8:00 am on the many September 30s of the record.

In practice, environmental records are seldom sufficiently long for practical application of the direct approach of simply averaging across years. If only 4–5 years of measurements are available (as is commonly the case), the calculated mean H_S for a given time on a given year day is subject to the random fluctuations inherent in small sample size, and as a result, calculated means may vary raggedly (and unrealistically) through time. How can we increase the sample size available from a short record and thereby increase the reliability of the calculated mean?

First, we note that because of their root dependence on celestial mechanics, the mean pattern of variation in environmental parameters varies smoothly through time. For example, air temperature might differ drastically between 8:00 and 8:10 am on a particular September 30, but unless there is some mechanism to ensure that the same abrupt shift happens between 8:00 and 8:10 am *every* September 30, the *mean* temperaturesfor these two times must be similar, and the overall temporal variation in mean air temperature must therefore be smooth. Similarly, the mean air temperature at 8 am on September 30 should differ only slightly from the mean air temperature at 8 am on October 1. We take advantage of this "smooth" nature of the system, and include in the average for each time within our typical year not only measurements taken on the same year day in separate years, but also measurements from nearby times on the same day and nearby days within each year (Fig. A1). In this fashion, a sufficiently large sample can be obtained to calculate a reliable set of means that vary smoothly through time. Once the sampling scheme is determined, the predictable signal is easily calculated for each point in the original time series, and residuals from this signal can then be calculated and manipulated as described in the text. Calculation of the predicted annual cycle of σ (the standard deviation of standardized residuals) proceeds in the same fashion as for the calculation of the mean annual pattern of each environmental variable.

Measurement Window. For factors that vary relatively slowly (e.g. sea-surface temperature and significant wave height) values within 6 h of a time point were included in the within-day window. For other factors (e.g. air temperature, wind speed), values within 1 h of a time point were included in the within-day window. The appropriate width of the between-day measurement windows is determined through an iterative process. We use ± 2 days as an initial guess for the width of the between-day window. We then use these windows to calculate the expected annual variation in an environmental factor. The predicted signal is then examined and the between-day window suitably adjusted. If the between-day measurement window is too narrow, the predicted annual variation is not smooth. If the measurement window exceeds the decorrelation time, the record may be overly smooth.

Specific sampling schemes were as follows: We used NOAA predictions for the year August 1983–July 1984 (a year midway between maximum and minimum in the cycle of lunar declination) as the expected values for the tides. (Note: this period coincided with a severe El Niño event, but this coincidence has no effect on *predicted* tidal levels or on our calculations.) For a given date in a given year, expected values and expected standard deviations for wave height and period were calculated using hourly measurements taken within 6 hr of the measurement time (a within-day measurement window of 13 values), and for values taken at the same times of day within a week of the date (a between-day measurement window of 15 values). We then averaged these values with analogous values for the same year day in the other six years (7 values). Thus, the calculation of each mean and standard deviation included 1365 values, providing an appropriately smooth record. For the calculation of averages and standard deviations in body temperature we used data taken at 10-min intervals within 1 hr of the measurement time on a given date (13 values), at the same times for dates within 7 days (15 values), and across the 7 yrs, again for a total of 1365 values in each estimate.

As with any sliding window, there are constraints at the beginning and end of the record. For example, if the within-day window comprises 11 contiguous sample points (t_{i-5}, t_{i-4}, …t_i, … t_{i+4}, t_{i+5}), the first five and last five points of the record are problematic: values required to fill the window extend beyond the record. Similar problems apply to the between-day values. We solve this problem by allowing the observed record to "wrap around," that is, values preceding the beginning of the observed record are taken in order from the end of the record and *vice versa*.

Coping With Skew. For the data in this study, correlation of the skew of standardized residuals with the mean cycle was rare, and in the sole case in which the correlation was substantial (wave period), the pattern of correlation was simple: residuals were negatively skewed in summer and positively skewed in winter. We solved this problem by separating the overall time series of standardized period residuals into two series: winter and summer. Within each series, skew is not substantially correlated with the annual cycle, so the winter series can be legitimately resampled and combined with winter values of the annual cycle and the summer series can be legitimately resampled and combined with summer values of the annual cycle. More complicated temporal patterns of skew will require more complicated corrections, but we have not explored this problem. Nor have we explored the use of higher order indices when characterizing the shape of the distribution of standardized residuals.

As a check on the homogeneity of shape for the distribution of standardized residuals, standardized residuals where separated into two sets: those occurring when $\bar{x}(t)$ < its mean and those occurring when the annual cycle of $\bar{x}(t) \geq$ its mean. The frequency distributions of these sets were then plotted and visually compared. For the data used in this study, the two distributions were quite similar (representative examples are shown in Fig. A2), and we made no further effort to homogenize shape.

Decorrelation Time. Autocorrelation functions for standardized residuals of environmental data generally settled to 0 within 13 days. The ACF for air temperature deviations settled to a value of approximately 0.1 within 13 days, but maintained this low value for a lag of approximately 120 days before settling to 0. This behavior is indicative of long-term (approximately 120-day) intra-annual trends within our data, which varied from one

year to the next. However, if a decorrelation time of 120 days were used in the resampling scheme leading to the calculation of body temperature, only three segments would be chosen in each year of hypothetical data, allowing little in the way of actual re-arrangement. Consequently, in the calculations of body temperature, we used a decorrelation time of 13 days for all data, including air temperature. The effect of this less than maximal decorrelation time was tested by repeating the analysis with block lengths times varying from 7 to 39 days. Tidal deviations required 30 days to settle to 0, and we used this decorrelation time for all data in the analysis of waves and tides and the hydrodynamic forces they impose. The effect of block length was tested by varying length from 15 to 90 days. Bühlmann and Künsch (1999) note that block lengths on the order of $n^{1/3}$ (where n is the length of the times series resampled) perform well in many cases. For our environmental times series, $n^{1/3}$ = 13.7 days, thus our nominal block lengths of 13 and 30 days seem reasonable.

Segment End Effects. In the resampling scheme used here, no effort has been made to match standardized residuals across the ends of segments. That is, while residuals vary more or less smoothly within a segment, there may be an abrupt shift as one segment ends and the next segment begins. For processes (such as wave-induced hydrodynamic force) that depend solely on instantaneous values of the various environmental variables, such an abrupt shift should cause no problem. In contrast, for processes such as body temperature, in which the current state depends in part on the history of prior values, an abrupt shift in standardized residuals from one segment to the next might have an unrealistic effect. We minimize the potential for such adverse effects in this study by requiring that all segments begin and end at midnight. The segment transition is thus temporally removed from the time of day at which body temperatures reach their maximum, allowing time for any effects of a shift to be damped out. The timing of the resampling scheme should be re-evaluated if, for instance, minimum body temperature (which typically occurs at night) is the focus of the analysis.

Independence. In our resampling scheme, no effort is made to ensure that adjacent samples are independent. Because (with the exception noted above) the width of each sampled segment equals or exceeds the decorrelation time for the process at hand, each segment contains within in it all the pertinent statistical information regarding the normalized deviations of that process. Thus, even if adjacent or overlapping samples are, by chance, chosen sequentially from the original time series, no violence is done to the hypothetical time series. The alternative would be to specify that the beginning point of each newly chosen segment is separated by more than the decorrelation time from the ending point of the previously chosen segment. This would ensure strict independence of segments, but this scheme seems overly restrictive. For example, it would allow 0 probability of obtaining by chance the original time series of standardized residuals.

LITERATURE CITED

Darwin, G. H. 1962. The tides and kindred phenomena in the solar system. W. H. Freeman, San Francisco, 326 pp.

Bühlmann, P., and H. R. Künsch. 1999. Block length selection in the bootstrap for time series. Computational Statistics and Data Analysis. 31:295–310.

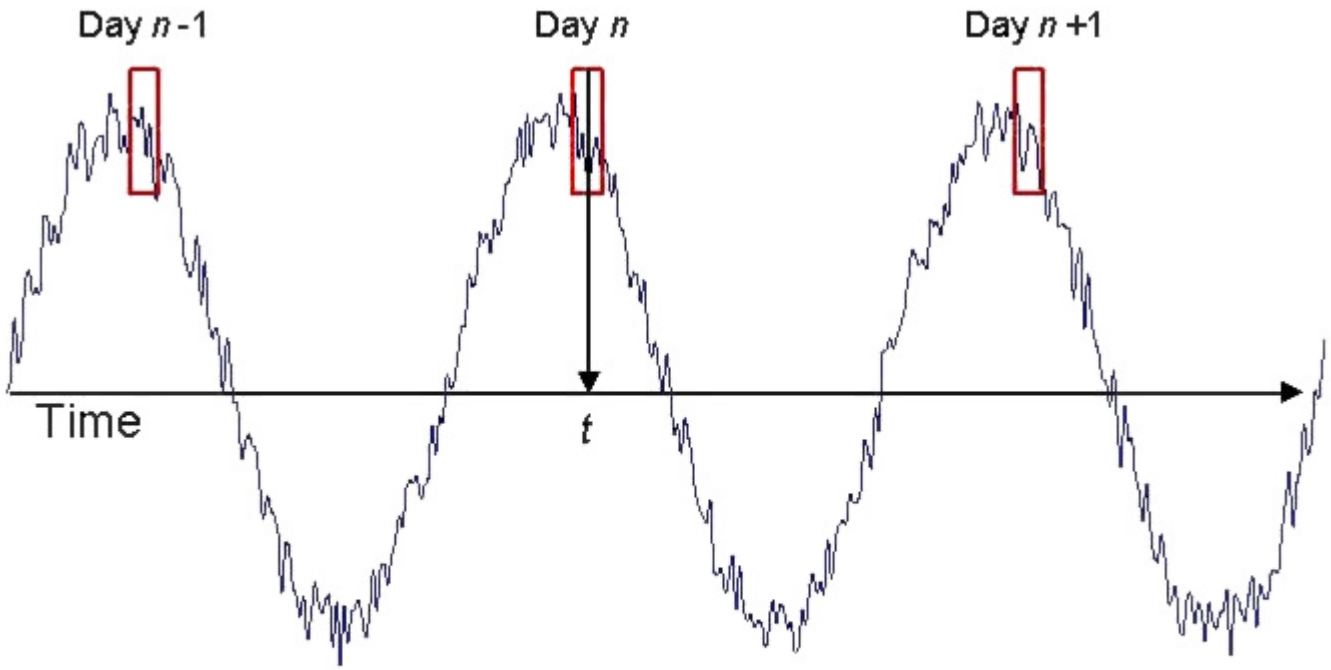

FIG A1. The scheme for calculating smooth averages and standard deviations. For time *t*, data are included at nearby times within the same day (the within-day window) and comparable times on nearby days (the between-day window). The value for time *t* within a given year is then averaged with values for the same time across years.

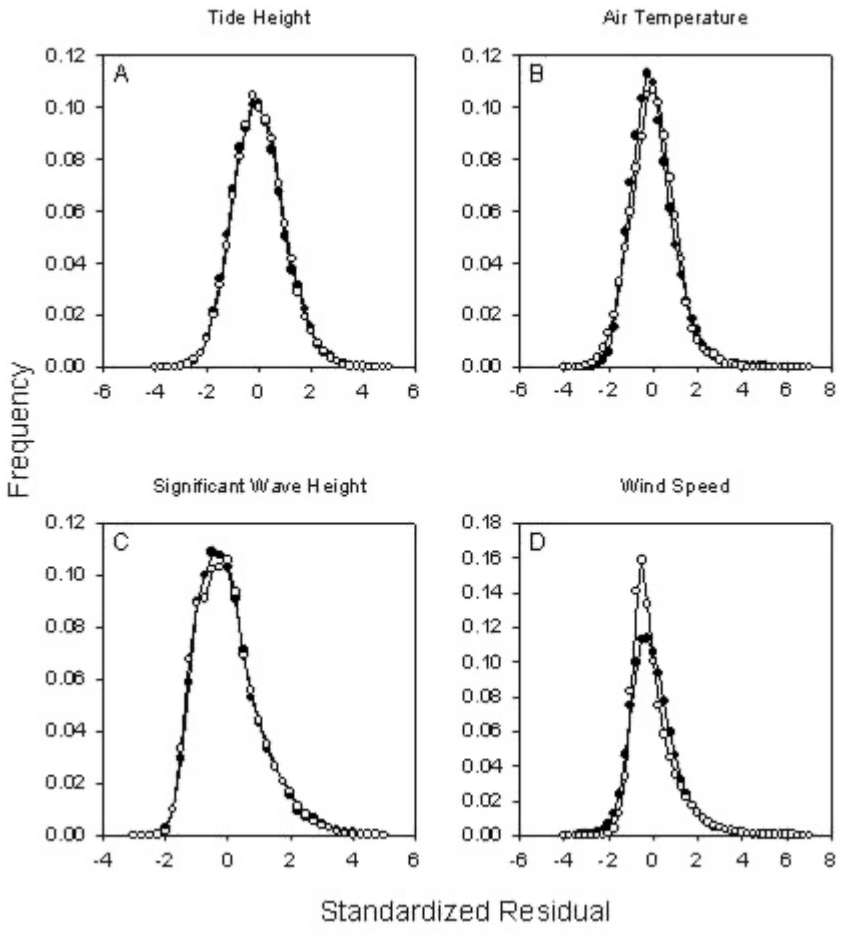

FIG A2. Representative distribution of standardized residuals. In each case, the distribution was calculated for times when the mean annual cycle was above average (closed circles) and below average (open circles).

[Back to M079-014]

Ecological Archives M079-014-A2

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix B. Translating the generalized extreme value distribution.

Eq. 14 in the main text is one form of the generalized extreme value (GEV) distribution. Coles (2001) and Katz et al. (2005) use an equivalent form:

$$P(x) = \exp -\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi} \qquad \text{(B.1)}$$

Table B1 gives the equivalencies required to translate between Eq. 14 and Eq. B1. Note that the symbol σ in Eq. B1 has no connection to the use of this symbol in this study, it is used here solely to ease translation between published forms of the GEV.

LITERATURE CITED

Coles, S. T. 2001. An introduction to statistical modeling of extreme values. Springer–Verlag. London, UK, 208 pp.

Katz, R. W., G. S. Brush, and M. B. Parlange. 2005. Statistics of extremes: Modeling ecological disturbances. Ecology 86:1124–1134.

TABLE B1. The equivalencies required to translate between Eq. 14 in the text and and Eq. B1 used by Coles (2001) and Katz et al. (2005).

Eq. 14 Gaines and Denny (1993)	Eq. B1 Coles (2001), Katz et al. (2005)
α	$\sigma - \xi\mu$
β	$-\xi$
ε	μ

[Back to M079-014]

Ecological Archives M079-014-A3

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix C. Simulated time series.

In the absence of long-term environmental data for intertidal sites, we desire to create a long-term set of reasonably realistic hypothetical data with which we can test the accuracy of our resampling approach. Given this long-term data, we choose a 7-yr subset as an example of the type of short-term data available for the real world. We then apply our resampling method to this 7-yr subset to predict return times, which can then be compared to the return times actually "observed" in the full data series.

Second order autoregressive moving-average (ARMA) processes provide a convenient mechanism to model the random fluctuations in standardized environmental residuals. The standardized residual $\Delta x_{n,i}$ at a particular present time i depends on both chance (in the form of a random numbers, φ) and previous values of the residual:

$$\Delta x_{n,i} = AR_1 \Delta x_{n,i-1} + AR_2 \Delta x_{n,i-2} + G\varphi_i - GMA_1 \varphi_{i-1} - GMA_2 \varphi_{i-2} \qquad \text{(C.1)}$$

Here AR_1 and AR_2 are autoregressive constants and $\Delta x_{n,i-1}$ and $\Delta x_{n,i-2}$ are values of the residual one and two times steps before the present, respectively. MA_1 and MA_2 are moving-average constants. The random numbers φ are chosen from a standard normal distribution (mean = 0, standard deviation = 1) and each is multiplied by a constant gain, G. φ_i is the newly chosen random number, φ_{i-1} is the random number chosen one time step before the present, and φ_{i-2} is the random number chosen two time steps before the present. For appropriate values of the autoregressive constants, the ARMA process outlined here is asymptotically stationary; that is, after an initial interval during which the process may have a net trend, it settles into stationarity (Priestley 1981, Chapter 3).

Appropriate values of *AR* and *MA* for significant wave heights, peak wave periods, and tidal elevations were determined through analysis of the standardized residuals calculated from our empirical 7-yr data series, and are shown in Table C1. For each variable, we created a year's worth of hourly data to allow the ARMA processes to become stationary, and then calculated 10,000 years of simulated standardized residuals. These hypothetical standardized residuals were then combined with the annual cycles of each variable measured as described in the *Materials and Methods* to create 10,000-year long time series of hypothetical wave height, wave period, and tidal height data. We then analyzed a random 7-yr segment of these simulated data in exactly the same fashion as for the real 7-yr time series: we determined the predicted variations (means and standard deviations) for each factor and from these predicted values we calculated the stochastic standardized residuals. Standardized residuals were resampled to create 10,000 year-long realizations of the wave environment, and these realizations were played through the model for wave stress. These 10,000 year-long records of stress provide an estimate of return time as a function of imposed stress (based on 7 yr of data), and this estimate can then be compared to the actual return times observed in the 10,000 year simulated series (Fig. 7 in the main text).

Note that we do not intend the ARMA processes used here to model the actual wave and tide data precisely. Instead, we use them solely to calculate a reasonably realistic, long-term data set with which we can assess the validity of our resampling approach.

LITERATURE CITED

Priestley, M. B. 1981. Spectral analysis and time series. Academic Press, New York, New York, USA. 890 pp.

TABLE C1. Coefficients for the autoregressive moving-average simulation.

	Significant wave height	**Peak wave period**	**Tidal height**
AR1	1.6842	1.6856	0.45705
AR2	-0.6991	-0.7333	0.50762
MA1	0.3028	0.1675	-0.2234
MA2	-0.2260	-0.1903	0.2310
G	0.100	0.155	0.330

[Back to M079-014]

Ecological Archives M079-014-A4

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix D. Disturbance effects on logistic growth.

The logistic equation provides a simple model of population size, in which the growth of the population through time is governed by both its intrinsic rate of increase r and the carrying capacity of the environment C:

$$Y_{t+1} = (r+1)Y_t + \left(\frac{r}{C}\right)Y_t^2 \qquad \text{(D.1)}$$

Here Y_t is the number of individuals in the population at time t and Y_{t+1} is population size at time t+1.When population size is small, growth is exponential, but growth rate decreases as population size approaches carrying capacity. In the example we use here, r = 0.05, C = 10,000.

To model the disturbance resulting from an extreme event, we reduce population size by an intensity factor E. That is, if population size immediately before the event is Y, size immediately after the event is Y/E. Periodic imposition of disturbance is akin to sustainably harvesting a stock of fish; after an initial period of equilibration, population size fluctuates about a steady average. It is the more complex effect of randomly-imposed events that is of interest here.

In our heuristic example, random extreme events (with a given intensity E) have a probability P = 0.97 of *not* occurring in a given time interval. In other words, in each interval we pick a random number from a uniform distribution between 0 and 1. If that number is < 0.97, no extreme event is imposed. If the random number $\geq$ 0.97, the population size is reduced by a factor of E. The system is initialized with 10 individuals and allowed to run for 500 intervals, allowing the process to "equilibrate." We then measure the average population size over the next 10,000 intervals. This entire process is repeated 30 times, and the average and standard deviation of these average population sizes is recorded. For comparison, the same experiment is repeated, but extreme events are imposed every 33.3 intervals (equal to the return time of random events). We varied intensity factor from 1 (= no effect) to 6 (only 1/6 of the population survives each extreme event). Higher intensity factors resulted in extinction (population size < 1 individual) at some point in the 10,000 years.

We present the results of this model solely as a heuristic example to demonstrate potential contrasts between random and periodic disturbances. Different effects can be obtained with different values for r, C, and P.

[Back to M079-014]

Ecological Archives M079-014-A5

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix E. Estimating the shape of tails.

The shape of the tail of maximal standardized residuals can be estimated from the data in hand. For example, Fig. E1 shows the cumulative distribution of the largest 0.2% of standardized residuals for predicted wave periods, a distribution of "points above threshold" that, in theory, asymptotically approaches a generalized Pareto model (Coles 2001):

$$P(x) = 1 - \left(1 + \frac{K(x-u)}{M}\right)^{-1/K} \qquad \text{(E.1)}$$

Here x is the variable of interest (in this example the standardized wave period residuals), u is the threshold value of x, and K and M are fitted shape and scale parameters, respectively. If the best-fit estimate of K is > 0, the distribution has an unbounded tail in which exceptionally large values may hide. If $K = 0$, the tail is unbounded, but with few exceptional values. If $K < 0$, the tail has a distinct upper bound. In the case of wave period standardized residuals, K is significantly less than 0 (-0.313 ± 0.054 [95% CL]), indicating that the tail of the distribution is bounded. A similar result obtains for the normalized deviation of tidal height residuals (K = -0.255 ± 0.072 [95% CL]). The best-fit estimate of K for significant wave height is negative, but not significantly different from 0 (K = -0.041 ± 0.045 [95% CL]). In summary, our short-term estimates for the distribution of standardized tide and wave residuals undoubtedly miss some of the large residuals that would be captured by longer-term records, but it appears unlikely that exceptionally large residuals are hidden in the unsampled tails of these distributions.

The calculation of limpet body temperature involves additional environmental variables. The tails for the standardized residuals of air temperature and sea-surface temperature are unbounded (K = 0.027 ± 0.009, 0.329 ± 0.006, respectively). The exceptional values found in the unsampled portions of these distributions might contribute to higher body temperatures than we have calculated here.

LITERATURE CITED

Coles, S. T. 2001. An introduction to statistical modeling of extreme values. Springer-Verlag. London, UK.

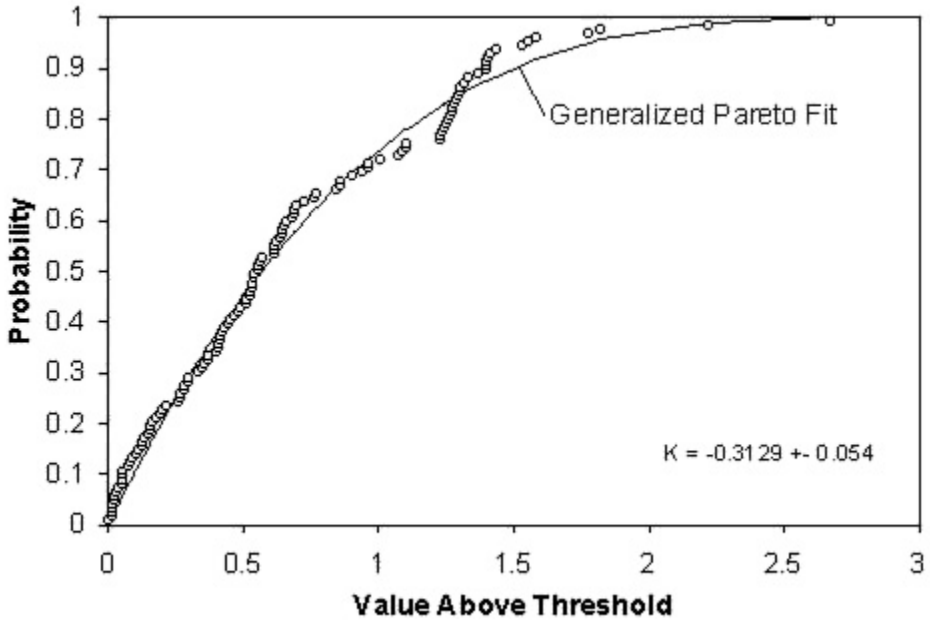

FIG E1. The upper tail of wave period deviations closely matches a generalized Pareto function (Eq. E1).

[Back to M079-014]

Ecological Archives M079-014-A6

Mark W. Denny, Luke J. H. Hunt, Luke P. Miller, and Christopher D. G. Harley. 2009. On the prediction of extreme ecological events. *Ecological Monographs* 79:397–421.

Appendix F. Estimating absolute maxima.

In our approach to predicting ecological extremes, we resample a time series of environmental factors, and each resampling provides a realization of how the environment might play out. If the time series from which we resample is finite, there must be at least one particular pattern in which we can resample the series that leads to the largest imposed stress. How can we calculate the value of this absolute maximum extreme?

It would be impractical simply to try all possible resamplings. If there are n points in the time series of each environmental variable at which we can start a block and q blocks chosen for each year-long realization, there are n^q different ways in which the time series can be resampled. For example, in our analysis of hydrodynamic stress, n = 2,550 (the number of days in 7 years) and q = 13 (the number of 30-day blocks in a year), so there are 1.93×10^{44} possible resamplings, a prohibitively large number. Is there a more efficient method of determining absolute maximum hydrodynamic stress?

One's first impulse might be to take the largest residuals in tides, wave height, and wave period and add them to the largest predicted values for these variables. Wouldn't the stress resulting be the largest possible? There are two reasons why this approach is inappropriate. First, recall that to maintain any cross-correlation among factors when we resample standardized residuals, segments of all residuals are sampled at the same time point in the record (Fig. 3 in the main text). Thus, unless the maximum tidal residual occurs at the same time as the maximal residuals of wave-height and wave-period, these maxima can never act in concert in our resampling scheme. Similarly, unless the highest predicted tide occurs at the same time as the highest predicted wave height and longest predicted wave period, the extreme residuals in each variable can never be applied simultaneously to their respective extreme predicted values. Second, the maximum values of each variable do not necessarily combine to give the maximum output. As noted in the text, hydrodynamic stress is not a monotonic function of wave period: at certain tide heights, stress decreases as T increases above some intermediate value. As a result, if we combine maximum residuals and maximum predicted values for H_S, T, and tidal height from our HMS data, we predict an absolute maximum stress that is actually smaller than some measured values.

Direct calculation of absolute maximal body temperature is even more difficult than direct estimation of absolute maximal wave stress. Body temperature depends not only on the instantaneous values of air temperature, wind speed, solar irradiance, etc., but also on their history. Thus, even if one could line up the extreme residuals for each variable (keeping in mind that it is the minimum wave height, tide, and wind speed that lead to the highest temperatures), and in turn line these up with the extreme expected values, the resulting body temperature would not necessarily be the maximum possible. Only if the conditions preceding these extreme circumstances provided sufficient time for the substratum to heat up would a maximal body temperature be achieved. In sum, the absolute maximum cannot be calculated by simply toting up the various individual extremes.

However, absolute maximum stress can be *estimated* using the statistics of extremes (Gaines and Denny 1993, Denny and Gaines 2000, Coles

2001, Katz et al. 2005). Our resampling method produces an ensemble of yearlong hypothetical environmental realizations, each of which can be processed to predict the maximum value of a biological relevant parameter. We can use the distribution of these predicted annual maxima to test for the existence of an absolute largest annual maximum. In theory, the cumulative distribution of our predicted annual maxima should asymptotically approach a member of the family of generalized extreme value distributions (GEV) described by Eq. 14 in the main text. This asymptotic distribution is estimated by fitting the "observed" distribution of annual maxima (i.e., the distribution of stress values calculated using our resampling technique) to Eq. 14 using a maximum likelihood criterion. If the best estimate of coefficients α and β are both positive, their ratio is an estimate of the absolute maximal value the distribution can attain. For example, the distribution of annual maximum hydrodynamic stresses for D = 1 m is best fit with values (±95% confidence limits) of α = 4512 N m^{-2} (±842) and β = 0.09952 (±0.02811). The ratio of α to β (our best estimate of the absolute maximum) is 4.53×10^4 N m^{-2}, only slightly above the maximum of 3.88×10^4 N m^{-2} recorded in the 10,000 years of our realizations. Given the limited statistical power available from our 1000 year-long realizations of limpet body temperature, the best-fit value of β (although nearly always positive) cannot be statistically distinguished from 0 for any substratum orientation, so a reliable estimate of absolute maximum body temperature cannot be calculated from our data.

The ability of univariate extreme-value analysis to estimate absolute maximum hydrodynamic stress or body temperature depends on the ability of our bootstrap method to provide information about the distribution of these extremes. There are cases in which bootstrap resampling is known to fail in this task. The classic example is the case in which one attempts to estimate the distribution of maximum values drawn from a uniform distribution extending from 0 to 1 (Bickell and Freeedman [1981], Efron and Tibshirani [1993, pg 81]). A nonparametric bootstrap does not provide a good estimate of the shape of the distribution of sample maxima.

The difficulty with the bootstrap procedure in this case arises from the limited size of the empirical sample from which bootstrap estimates are drawn. If one takes 50 samples from the uniform distribution (as Efron and Tibshirani do in their example), there are likely to be relatively few samples near the upper end of the distribution. As a consequence, no matter how many times one resamples these 50 measurements, little information is available about the shape of the distribution of maxima.

However, this example differs from the situation we encounter in our analysis of hydrodynamic forces and body temperatures. Because we have only 7 years of empirical data, when we choose a bootstrap sample from the time series of standardized residuals we have limited information available by which to judge the shape of the upper end of its distribution. But it is not the extremes of the standardized residuals that govern the ecological consequences, rather it is how the residuals combine with their respective mean annual cycles and how the resulting variables interact. Thus, while we have limited data in the upper end of the distributions of individual maximum standardized residuals, we have a very large number of combinations of residuals with their annual cycles and of the resulting variables with each other. It is these combinations (rather than the extremes of individual variables) that determines the distribution of *ecological* maxima.

LITERATURE CITED

Bickel, P. J., and D. A. Freedman. 1981. Some asymptotic theory for the bootstrap. The Annals of Statistics. 9:1196–1217.

Coles, S. T. 2001. An introduction to statistical modeling of extreme values. Springer-Verlag. London, UK, 208 pp.

Denny, M. W., and S. Gaines. 2000. Chance in biology. Princeton University Press, Princeton, New Jersey, USA, 291 pp.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall/CRC Boca Raton, Florida, USA, 436 pp.

Gaines, S. D., and M. W. Denny. 1993. The largest, smallest, highest, lowest, longest and shortest: extremes in ecology. Ecology 74:1677–1692.

Katz, R. W., G. S. Brush, and M. B. Parlange. 2005. Statistics of extremes: Modeling ecological disturbances. Ecology 86:1124–1134.

[Back to M079-014]