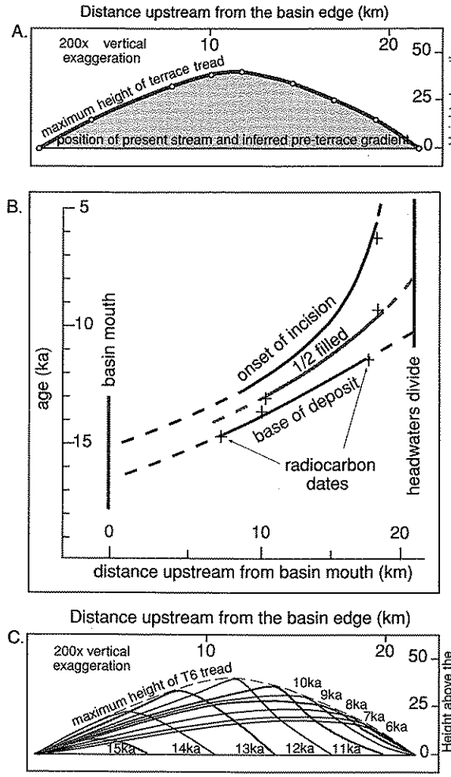
Figure 2.12. Terrace formation



A. Height of the top surface of terrace T6 at Cajon Creek, California, compared with the modern river. The gradient of the modern river is removed, in order to emphasize the apparent upward bowing of the surface. The highest part of the terrace occurs where it crosses the San Andreas fault, but there is no apparent offset here. B. Variations in the timing of initiation of aggradation and incision along Cajon Creek, both of which are highly diachronous. Radiocarbon dates (crosses) on the fill are used to calibrate the aggradation–degradation cycle. Incision is assumed to begin shortly after the maximum thickness of fill is attained. C. Interpretation of the evolving topographic profile of the terrace due to migration of the locus of deposition and erosion. Modified after Weldon (1986).

the terrace surface. Compared to the upstream reaches, deposition commenced ~4 kyr earlier in the downstream reaches, and incision of the aggraded surface began there ~7 kyr earlier. Thus, while upstream parts of the terrace were just beginning to aggrade, the dissection of the downstream areas had already begun! Even though the entire terrace is less than ~15 kyr old, the age and time of abandonment of the upper surface of the terrace varies by ~7 kyr across a distance of ~10 km (Weldon, 1986).

The cause of this diachronous response is unknown. Alluvial terraces form within coupled, complex systems that include erosion, transport, and deposition of sediments. Such systems do not respond instantaneously or uniformly to changes in controlling parameters, such as discharge, sediment supply, rock uplift rate, or base-level lowering (Humphrey and Heller, 1995). Instead, changes in aggradation or degradation propagate up- and downstream through both alluvial and bedrock systems at rates which may be related to the diffusivity of each system and to the velocity of kinematic waves that move along alluvial and bedrock reaches (Humphrey and Heller, 1995). The duration of these waves of deposition or erosion often greatly exceeds the duration of the perturbation that initiated the wave. Many drainage basins have equilibrium response times that are considerably greater than 10^5 yr. Given that many changes imposed on river systems, such as discharge variations due to climate, occur on much shorter times scales, it seems likely that depositional/erosional systems will be constantly integrating the responses to successive variations in controlling parameters. Thus, equilibrium may be rarely attained in fluvial systems. The diachrony displayed by the Cajon terrace (Weldon, 1986) may typify many fluvial terraces. In the common absence of data to delineate chronologic differences along a terrace, however, most workers assume the upper terrace surface is essentially isochronous along its length. Diachrony at the scale of a few thousand years becomes less important when older terraces are considered, but for postglacial and Holocene terraces, such variability would significantly distort rates that were calculated assuming an isochronous terrace surface.

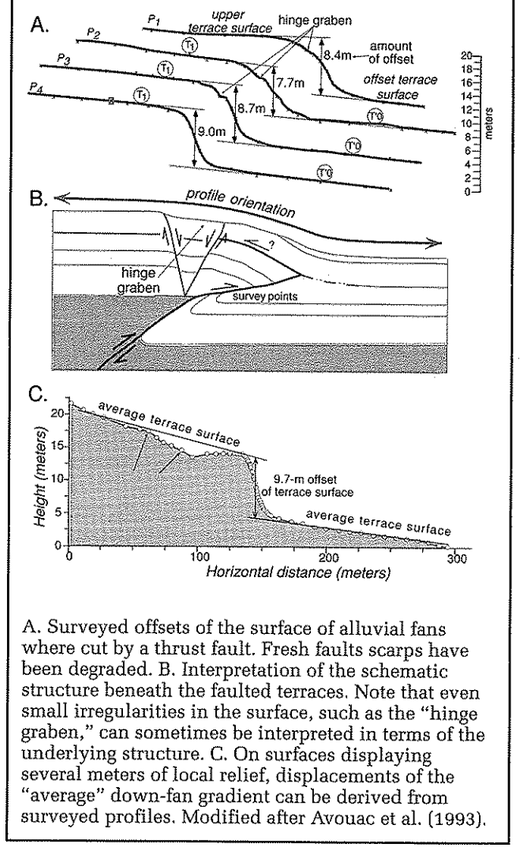
Alluvial Fans

In many respects, the surfaces of alluvial fans are similar to fluvial terraces. Slopes generally decrease down-fan and any longitudinal section displays a slightly concave upward profile. Deposition on most fans is considerably more episodic than in rivers, and in addition to channelized fluvial processes, various types of mass flows

and unchannelized flows tend to dominate deposition. Although most fans are characterized by steady, down-fan decreases in gradient, some fans appear to be segmented (Bull, 1964), such that in cross section, they comprise a suite of fairly straight slopes that abruptly change where the linear slopes intersect. Such segmentation can result from shifting of the boundary between deposition and erosion up or down the fan. For example, if deposition is focussed on the fan apex (also termed the fanhead), alluvium will accumulate and steepen this area with respect to the rest of the fan. If the fanhead is subsequently entrenched, remnants of the abandoned steeper surface will remain, but the apex of deposition will shift down-fan. As aggradation on this gentler, down-fan surface proceeds, the upper surface of aggradation may eventually intersect and begin to overlap the steeper, up-fan surface, such that an abrupt slope change occurs at the intersection. In this situation, the upper and lower parts of the fan surface will have different ages. In such circumstances, before using fans as geomorphic markers, the age of each surface should be separately determined. Even in the absence of fan segmentation, most fans are composed of a mosaic of surfaces of different ages, ranging from the modern channels to dissected remnants of long-abandoned surfaces. Weathering of clasts on the surface of a fan often permits the relative ages of different segments of a fan to be defined. Following isolation from deposition, weathering processes, such as fracturing, rind and varnish development, reddening of the underside of clasts, and granular disintegration, begin to modify the appearance and character of the fan surface (McFadden et al., 1982; Ritter et al., 1993). Often, in arid regions, readily visible contrasts in surface color permit classification of the relative ages of different fan surfaces (Bull, 1991). In more humid fans, more time-consuming quantification of changes in the depositional surface may be required to divide the fan into surfaces of differing ages. In order to use a fan surface as a reference against which to measure deformation, only similarly aged portions of the fan should be used to define a reference plane.

Many arid alluvial fans are dominated by deposition from debris flows (Whipple and Dunne, 1992). The surface topography of debris-flow deposits (up to several meters) and the roughening of the fan surface that results from them can determine the scale at which the local surface of the fan can be considered to have a predictable geometry and therefore to be useful as a geomorphic marker. Even surfaces with considerable roughness, however, can often be used successfully as markers, because, despite its irregularities, the average surface gradient can

FIGURE 2.13. Faulted fan surfaces of the Tien Shan



A. Surveyed offsets of the surface of alluvial fans where cut by a thrust fault. Fresh faults scarps have been degraded. B. Interpretation of the schematic structure beneath the faulted terraces. Note that even small irregularities in the surface, such as the "hinge graben," can sometimes be interpreted in terms of the underlying structure. C. On surfaces displaying several meters of local relief, displacements of the "average" down-fan gradient can be derived from surveyed profiles. Modified after Avouac et al. (1993).

be defined (Avouac et al., 1993), and offsets of this gradient by faulting can be readily recognized (Fig. 2.13).

Lava Flows, Debris Flows, and Landslides

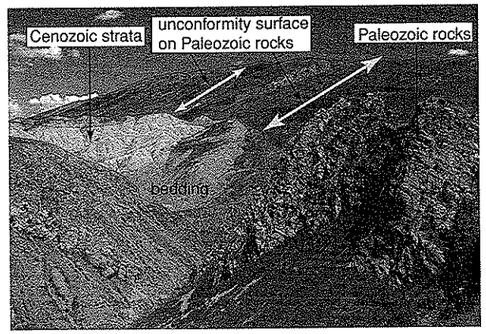
Despite the meter-scale roughness of the upper surface of most lava flows, the surfaces of elongate flows can provide excellent geomorphic markers. Topographic gradients along the flow surface can be measured directly and numerous surface features can be recognized for cross-fault correlation, the highly resistant flows have a high preservation potential, and the flow itself is often directly datable with radiometric methods.

The surface topography of mass movement deposits, such as debris flows, mudflows, and landslides, often depends in part on their water content and viscosity at the time of deposition (Whipple and Dunne, 1992). The higher the water content, the smoother the upper surface of the deposit and the greater the downstream travel distance are likely to be. These mass movement deposits have the advantage of having formed as the result of an instantaneous event, thereby avoiding the problem of diachrony in the surface seen, for example, in the fluvial terrace case. Moreover, it is not uncommon that they overrun or contain within their deposits organic debris that can be radiocarbon dated. As a consequence, a date from anywhere within them can often pinpoint the age of the deposit closely. In contrast, a ^{14}C date within the strata of an aggradational terrace simply provides a lower limit on the time of abandonment of the depositional surface. Multiple dates are often needed on such surfaces in order to judge the rate of aggradation and the timing of abandonment (Weldon, 1986) (see Fig. 2.12).

Erosional Surfaces

It has long been recognized that, during long intervals of tectonic quiescence, topography can be bevelled off by erosional processes and a low-relief landscape can be produced. Isolated erosional remnants (monadnocks or inselbergs) may be surrounded by *pediments*: erosional surfaces of low relief carved into bedrock. Across broader geographic regions, *peneplains* characterized by slightly undulating and generally featureless topography can result from long-continued erosion and deposition in the absence of active deformation. When such erosional surfaces formed in the distant geologic past, they were often eventually buried by subsequent deposition, such that a regionally extensive unconformity is preserved. When such low-relief surfaces (pediments, peneplains, regional unconformities) are uplifted tectonically, they can form prominent markers in the landscape. Typically these erosional surfaces are identified by the low-relief of the uplifted surfaces, an accordance between summit heights, and/or the smoothly varying topography of broadly folded or faulted terrains. Often this low relief stands in contrast to a nearby tectonically active front characterized by high relief. For example, the Wind River Range of Wyoming rises about 2 km above the adjacent basins and has steep-walled, glaciated canyons cutting its flanks. In many of the high-altitude parts of the range, however, there are rolling, low-relief surfaces. The elevation and tilt of many of these remnant surfaces can be extrapolated laterally to join other, similar remnants and to define a regionally extensive high-altitude surface of low relief (Small and Anderson, 1998). Such surfaces can not have formed at high altitude, but instead represent uplift of a surface that was created near the regional base level in the distant past. Similarly, in the Tien Shan of Central Asia, a regionally extensive erosion surface was bevelled across Paleozoic rocks and buried by Cenozoic sedimentary rocks (Chediya, 1986; Sadybakasov, 1990). This surface has been recently exhumed due to rock uplift. The striking contrast in erodability of the rocks above and below the unconformity has caused the Cenozoic sediment to be rapidly eroded, revealing the unconformity surface (Fig. 2.14), which provides an excellent marker for recording folding and faulting of ranges that rise as much as 2 km above the surrounding terrain (Burbank et al., 1999). Not only do erosion surfaces like those in the Wind River Range or the Tien Shan serve to define rock uplift, but the unconformity surface itself forms a reference for calibrating the amount of erosion that has occurred beneath it (Small and Anderson, 1998).

FIGURE 2.14. Exhumed unconformity surface in the Kyrgyz Tien Shan



The erosion surface was cut across folded Paleozoic bedrock and was later buried by 3 to 5 km of Cenozoic strata. Late Cenozoic uplift above local base level has caused the Cenozoic strata to be stripped off the unconformity. Today the surface is folded and rises about 1.5 km above the adjacent basins. Deformation of this regionally extensive marker surface readily defines the magnitude of differential rock uplift.

Linear Geomorphic Markers

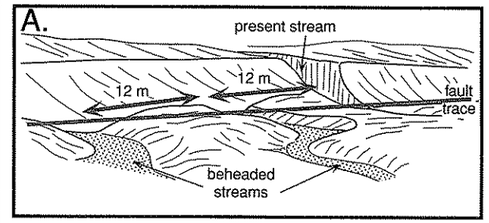
Whereas the previously described geomorphic markers represent areally extensive surfaces, it is also possible to use linear geomorphic and man-made features to determine deformation. Although displaced planar features are more suitable for defining regional tilting, linear features can provide ideal piercing points from which an offset can often be unambiguously measured. Unlike many two-dimensional surfaces, such as marine or fluvial terraces, many linear geomorphic features can be formed by individual events, some of which may have occurred instantaneously, from a geological perspective: for example, the levees that form on the margins of a debris flow. Such features often have no direct relation to climatic variations, so that ages need to be determined for each event in order to determine rates of deformation.

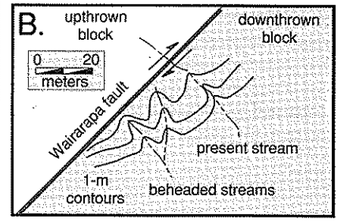
Rivers and Ridge Crests

The courses of rivers and ridge crests that are displaced across strike-slip faults can clearly record lateral offsets (Fig. 2.15A). It is important to ascertain, however, that the deflection of a stream is due directly to differential displacement of its course by faulting and is not the result of the intersection between a regionally sloping surface and a fault scarp. If streams are offset in directions that oppose the regional slope, the cause of the offset is more likely to be tectonic than when the deflection occurs in the direction of the regional slope. Due to strike-slip motions, streams can be "beheaded," by which it is meant that an abandoned stream channel abruptly terminates as it crosses a fault (Fig. 2.15B). The difficulty in assessing offsets of rivers and ridge crests lies in making reliable correlations from one side of a fault to the other. Commonly multiple ridges and streams cut across faults, so that specific correlations can be ambiguous.

Because rivers are capable of incising and modifying any displaced profile, vertical movements often are underestimated by the apparent displacement of the river channel at the location of the fault. If the upstream part of a stream bed is elevated by faulting with respect to its downstream continuation, the stream will tend to incise through the scarp. Remnants of the former valley floor may be preserved as small terraces on either side of the channel, and their height above the downstream, but offset continuation of the channel can be used to assess the amount of vertical displacement (Beanland and Clark, 1994).

FIGURE 2.15. Beheaded Streams





A. Sketch of beheaded streams along the Wairarapa strike-slip fault, North Island, New Zealand. Spacing between stream indicates two previous earthquakes with about 12 m of displacement in each event.
B. Map with 1-m contour of offset and beheaded stream channels along the Wairarapa fault. Modified after Grapes and Wellman (1993).

Glacial Moraines

The elongate ridges of ice-transported debris that form glacial moraines provide linear geomorphic markers that have an obvious direct climatic cause. Lateral displacements can often be readily measured from the offset in map view of the linear trend of the moraine crest, whereas vertical offsets can be assessed by comparing the topographic trend along the length of the moraine crest on either side of a fault.

If advances attributed to surging glaciers are excluded, then most major glacial advances are responses to major climatic changes. Thus, one might expect to be able to correlate the record of successive glacial advances with the record of Quaternary climatic fluctuations (see Box 2.1). For the most recent advances, this is often true, although the timing of the maximum extent of alpine glaciers in any particular mountain range often differs by thousands of years from the time of maximum ice sheet extent

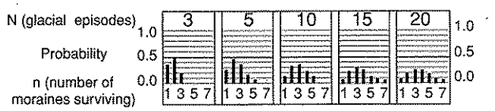
Box 2.3. The Problem of Moraine Survival

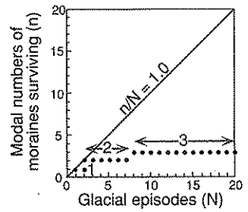
Ever since the first synthesis of the record of alpine glaciation (Penck and Brückner, 1909), it has been common to recognize three or four major moraines in glaciated valleys. At the same time, the global climate record indicates that there have been at least 10 major glaciations in the past 1 Myr and many more in the previous 1 Myr. Why is there such a mismatch between the number of glaciations and the preserved morainal record of those glaciations? One answer comes from a statistical analysis of the probability of moraine preservation (Gibbons et al., 1984). Let's assume that there has been a succession of fifteen glaciations and that, with respect to other glaciations, the relative magnitude of each glaciation and its associated advance is randomly distributed. What would happen if the most recent glacial advance were also the largest? It would wipe out most or all of the geomorphic record of all previous advances. If, on the other hand, the glaciations happened to fall sequentially from the most extensive at the beginning to the least extensive at the end, then every single glaciation would be represented. The question of how many moraines will survive can be posed statistically as follows: the probability (P) that n moraines will survive, if there were N glaciations is

$$P(n/N) = 1/N \sum_{N=n-1}^{N-1} P\big((n-1)/N\big)$$

For example, the probability of two moraines surviving, if there were four glaciations, is

$$P(2/4) = 1/4[P(1/1) + P(1/2) + P(1/3)]$$





Because P (1/N) = 1/N, P (2/4) = 1/4 (1 + 1/2 + 1/3) = .458. The probabilities for differing numbers of preserved moraines can be quite readily computed. Perhaps surprisingly, they indicate that, for eight to twenty glaciations with randomly distributed magnitudes, the most likely number of moraines to survive is three!

Overall, it is clear that a succession of glacial moraines will typically provide only a fragmentary record of climate change. Similarly, to the extent that aggradational terraces are correlated with the magnitude of glacial advances (as is often supposed), it is likely that the preserved aggradational terraces have buried older, smaller terraces beneath them.

Source: Gibbons, A.B., Megeath, J.D., and Pierce, K.L., 1984, Probability of moraine survival in a succession of glacial advances: Geology, v. 12, p. 327–330. *Fig. 3 and 4, p. 328.*

(Gillespie and Molnar, 1995). Therefore, while an absolute date on a moraine is always preferred, an approximate age can be assigned to undated moraines and tectonic rates (with appropriate uncertainties) can be calculated based on observed offsets.

When moraines other than those associated with the most recent advances are considered, the one-to-one correlation with the climatic record typically breaks down due to incomplete moraine preservation (Box 2.3). In such circumstances, a local glacial chronology associated with specific preserved moraines needs to be established in order to have reliable control on long-term rates of deformation. It is commonly assumed that within any area experiencing a regionally consistent climate, most of the glaciers would be expected to advance and retreat approximately synchronously, but there are few well-dated tests of this contention (Gillespie and Molnar, 1995). In the absence of dates, it is therefore not uncommon to try to correlate with the global climatic record, but such correlations should be made with considerable skepticism.

Other Linear Features

Numerous other linear features, both naturally occurring and man-made, can be used as markers for gauging displacement. Fairly viscous debris flows can create raised levees of coarse debris along their channel margins (Beanland and Clark, 1994). Fence lines, railroad tracks, curbs, sidewalks, lines painted on streets, and even tracks created by cars, motorcycles, or bikes can provide linear markers that are readily measured and are useful for documenting either co-seismic offsets of recent earthquakes (e.g., the scarp of the 1992 Landers earthquake in California displaced alluvial fans that sported hundreds of motorcycle paths) or slower rates of creep. These man-made markers are not geomorphic features in a strict sense, but when trying to generate a catalogue of displacements along a recently ruptured fault zone, any displaced linear feature with a known pre-faulting geometry should be evaluated.

Commonly Encountered Problems with Markers

The most typical difficulty in extracting the maximum information from a displaced geomorphic surface is the absence of a well-documented age for the surface. Considerable effort often is warranted to try to uncover datable material that can constrain the age of the feature. A knowledge of the array of available dating tools, experience with the appropriate field procedures for collecting suitable material for a specific dating method, and a both thorough and innovative approach to the problem of defining geomorphic ages is a boon to anyone attempting to obtain reliable age constraints and hence deformation rates.

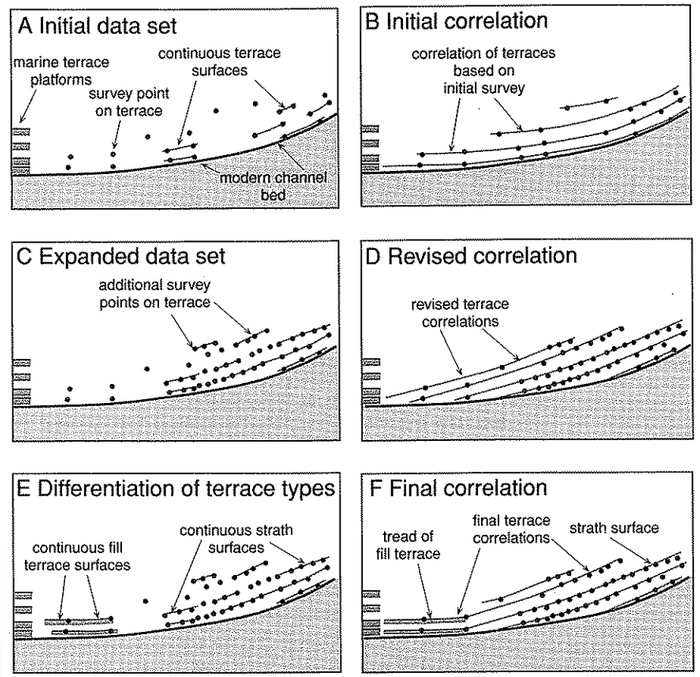
Correlation of partially preserved geomorphic features presents another challenge. As older and older markers are examined, their preservation typically becomes increasingly fragmentary. Unless there are distinctive characteristics that permit discrimination among features of differing ages, the correct correlation among the remnants may be difficult to achieve. A traditional field technique for terrace remnants, for example, has been to survey the height of terraces along local reaches and to correlate the terraces from one site to another based on their relative height above the river. This can be misleading, because it assumes a consistency in the longitudinal profile of the terrace through time. A better technique is to employ continuous, geodetic surveying along terrace surfaces (Merritts et al., 1994) and to document the connection from one reach to the next whenever possible (Fig. 2.16). Even with such data, reconstruction of the long profiles of multiple terraces can be ambiguous. It is important also to distinguish between strath and aggradational terraces, because, when aggradational terraces primarily result from a downstream rise in local base level, their gradients are likely to be gentler than those of most upstream strath terraces. Recognition of different terrace types, continuous geodetic surveying, dating of terraces, and analysis of how and why the terrace sequence developed all assist in creating a reliable reconstruction.

Other problems can arise when a regional chronology of geomorphic markers has been developed and is subsequently applied to the deformational analysis of local sites. Suites of aggradational terraces or of glacial moraines are often dated by assembling ages from individual outcrops spread over a broad area. This is a reasonable, and often the only practical approach. Nonetheless, it is important to be aware of the possible presence of terraces or moraines that are unrelated to the regional sequence. Surging glaciers, for example, often advance largely independently of climatic controls. Yet, they create moraines that are typically indistinguishable from those of non-surging glaciers. Ages assigned to those moraines based on correlation to a climatic record would be erroneous. Similarly, a very large landslide in a catchment may overwhelm the transport capacity of a river. This would cause the river bed to aggrade, and an aggradational terrace could be left behind that is unrelated to terraces in nearby drainages which reflect instead the regional climatic controls.

Summary

Geomorphic markers abound within many landscapes. Common geomorphic markers include wave-cut benches (marine and lacustrine), fluvial surfaces (terraces and fans), and linear features, such as moraines. Clever reading of the landscape and innovative adaptations of the general principles and approaches discussed here provide a basis for utilizing geomorphic markers in almost any geomorphic setting. One must attempt to understand the pristine, undeformed shape of the marker (formally, the initial conditions of the problem), because this forms the basis for all interpretations of deformation. Dating of geomorphic features takes persistence, familiarity with available techniques, innovation, and some luck. The search for datable material is often tedious, but as one geologist said, "If you haven't found datable material, it's because you haven't looked carefully enough!" All calculations of deformation rates depend on assigning ages to displaced features. Therefore, those long searches can often pay off by yielding new insights into how rapidly deformation has occurred in the past. In the next chapter, we examine several approaches to dating of geomorphic, stratigraphic, and structural features that record deformation. Just as geomorphic markers, such as moraines, terraces, and alluvial fans, have some

FIGURE 2.16. Correlation of remnants of river terraces based on an example from the Mattole River, northern California



A. Traditional technique of surveying the altitude of terraces in local reaches provides isolated points that must be correlated. B. If it is assumed that the past and present river gradients were parallel, correlations are then based on the height of a terrace above the modern river bed. C. More spatially continuous surveying of the upstream extent of terraces can define their gradients more reliably than can be done solely with correlated spot heights. D. When compared with the modern gradient, the reconstructed terraces are clearly not parallel to the modern profile. E. When the surfaces of aggradational terraces are distinguished from those of strath terraces, contrasts in gradients between the two terrace types become apparent. F. Terrace reconstruction based on survey, dating, and geomorphic field evidence shows strath terraces with steep gradients merging into the present river profile. Aggradational terraces with gentle gradients result from deposition forced by rising sea level (rising base level). Modified after Merritts et al. (1994).

predictable geometries, different kinds of faults and folds display contrasting styles of surface deformation, as discussed in an ensuing chapter. When fault-zone displacements are combined with rates derived from dated and deformed geomorphic markers, a three-dimensional reconstruction of fault-zone evolution, rupture history, landscape perturbations, and geomorphic responses can be attained.

CHAPTER 3

Establishing Timing in the Landscape:— Dating Methods

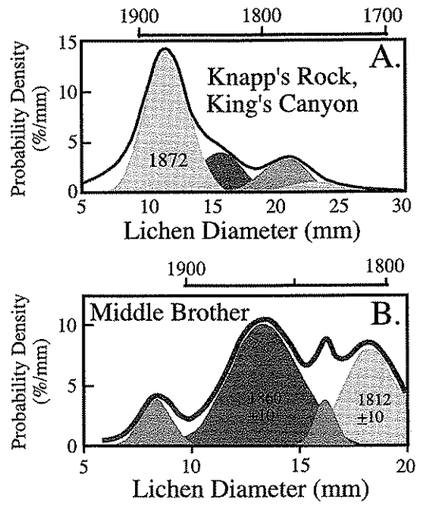
The science of tectonic geomorphology relies strongly on placing time controls on the landscape. In order to determine the rate at which a fault moves or a surface deforms, we must establish the age of offset features. The precision needed depends on the questions being asked, and on the recurrence time scales of the events themselves. Until the advent of radioactive dating methods, the principal means of establishing timing in the landscape was through the use of relative dating methods. The distinction between relative and absolute dating methods is that relative methods yield relational information only (surface *X* is older than surface *Y*, which is in turn older than surface *Z*), whereas absolute dating allows us to place a number on the age (terrace *X* was created 3565 years before the present). The latter is absolute in the sense that it does not require attention to any other surface, but one must be aware of the potential errors inherent in each of the absolute dating methods. When calibrated against some absolute ages, several relative dating techniques lend themselves to some degree of quantification, making them "semi-quantitative" methods. In Tables 3.1 and 3.2 we list several of the frequently applied methods, the age ranges for which they can be used, and a primary reference to the literature where they are better discussed. This chapter does not present an exhaustive review of the available techniques, but rather illustrates a few of the techniques, focusing on the newer and more quantitative methods. For a rather comprehensive review of dating techniques, see Noller et al.'s (2000) *Quaternary Geochronology*.

Relative Dating Methods

Clink versus Thump, and its Quantification: The Clast Seismic Velocity Method

For years, geomorphologists have determined the relative age of a surface by walking up to the surface to be dated and hammering on the boulders protruding from it. The resulting sound, which is a sharp bell-like clink if the boulder is fresh, and a dull and boring thump if the boulder is old and decrepit, is then used to assign a relative age. Many boulders are pounded and a statistical sense is developed about the age of this surface relative to some other. About fifteen years ago, this method was modified to allow quantification (Crook, 1986). The clast seismic velocity method is based on the principle that a rock exposed at the earth's surface will develop microcracks through a variety of weathering processes, and that the number of microcracks controls the propagation speed of seismic waves (compressional waves, or sound) through the rock. Although it is difficult to measure microcrack density, it is quite easy to measure the seismic velocity by using a micro-seismic timer, as illustrated in Figure 3.1A. For each of several spacings between the accelerometer and the hammer, many measurements of travel time are made. The travel time is recorded, and is converted to velocity by dividing through by the separation distance between hammer and sensor. The clast seismic velocity of a particular boulder is then calculated, and the next boulder is selected. Many boulders on each surface to be dated (e.g., moraine, alluvial fan,

FIGURE 3.6. Histograms (probability density functions) of rhizocarpon lichen diameters on boulders on talus in the Sierra Nevada, California



Peaks record times of generation of new talus blocks, corresponding to major historical earthquakes in California. Modified after Bull (1996).

As in all techniques, there are limitations and uncertainties. Here one must obviously have a climate in which a well-behaved lichen thrives, and must be able to obtain independent ages on surfaces for use in calibrating the growth-rate curve. There will always be some lag between deposition of a rock on a surface and its colonization by lichen. In addition, lichens apparently grow rapidly at first (the great-growth phase), before entering their linear growth phase. Some studies (Porter, 1981a) have suggested that lichen growth rates may be higher on fine-grained volcanic substrates than on coarser-grained intrusive rocks, whereas other studies conclude that lithology, smoothness of the substrate, local mean annual temperature, precipitation, and length of the growing season apparently do not exert a strong influence on the growth rate. Protection from the sun and wind does promote more rapid growth, so that Bull (1996), for instance, measured only exposed lichens. Whereas further testing of the population techniques for lichenometry is warranted, this approach holds remarkable promise for developing rather detailed chronologies on geomorphic surfaces of less than 500 years old. This is an age range of great interest to paleoseismology, but one in which absolute age control is often difficult to achieve.

Absolute Dating Methods

Most absolute dating methods rely on some process that occurs at a regular rate—in essence, a clock. In some instances, these clocks leave a physical record that may be biological, as typified by tree rings, or geological, as represented by annual lake beds or varves. The remaining techniques rely on either atomic clocks or cosmic clocks or both. Here we review the fundamental concepts of the atomic clocks we will use.

Some (parent) atoms spontaneously decay through fission to other (daughter) atoms plus associated nuclear fragments plus energy. These events are called radioactive decays. Whereas any particular parent atom may decay at any random instant, the probability of such decay depends on the parent-daughter pair. The lower the probability of decay at any instant, the longer it will take for a population of parent atoms to decay to half its size, and vice versa. This is the notion of the half-life. Mathematically, the process is captured in the differential equation

$$dN/dt = -\lambda N \qquad (3.1)$$

This is a commonly used example of a first-order linear differential equation found in the front of many introductory differential equation textbooks. Here N is the number of parent atoms, dN/dt is the rate of change of this population, and the decay constant λ expresses the probability of decay of any parent atom at any instant. The solution to this equation is an exponential function:

$$N = N_o e^{-\lambda t} \qquad (3.2)$$

where N_o is the abundance of the parent atom at time 0. You can easily see that the time it takes for the decay of the population from N_o to N_o/e (recalling that $e = 2.718\ldots$, this is roughly $N_o/3$) is $1/\lambda$. The half-life, found by solving for the time at which $N = N_o/2$, is $t_{1/2} = [-\ln(1/2) / \lambda]$, or $0.693/\lambda$. In Table 3.2 we list a few of the commonly used radioactive isotopes, their decay constants, and their associated half-lives. We will see that in some cases it is either more instructive or easier to measure the accumulation of the daughter isotopes, or the ratio of the parent to the daughter, rather than the parent.

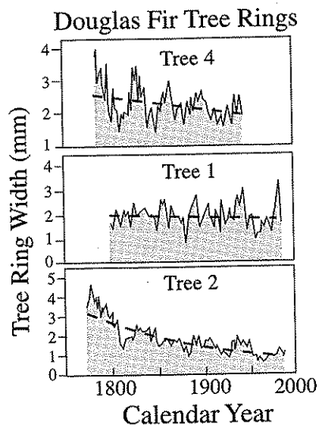
TABLE 3.2. Absolute Dating Methods

Method	Useful Range	Materials Needed	References: Classic and recent
Radioisotopic			
14C	35 ka	wood, shell	Libby, 1955; Stuiver, 1970
U/Th	10–350 ka	carbonate (corals, speleothems)	Ku, 1976
Thermoluminescence (TL)	30–300 ka	quartz silt	Berger, 1988
Optically stimulated luminescence	0–300 ka	quartz silt	Aitken, 1998
Cosmogenic			
In situ ^{10}Be, ^{2}6Al	3–4 Ma	quartz	Lal, 1988, Nishiizumi, 1991
He, Ne	Unlimited	olivene, quartz	Cerling and Craig, 1994
^{36}Cl	0–4 Ma		Phillips et al., 1986
Chemical			
Tephrochronology	0–several Ma	volcanic ash	Westgate and Gorton, 1981; Sarna-Wojicki et al., 1991
Amino acid recemization	0–300 ka; range temperature dependent	carbonate shell	Bada et al., 1970; Bada, 1972; Wehmiller et al., 1988
Paleomagnetic			
Identification of reversals	>700 ka	fine sediments, volcanic flows	Cox et al., 1964; 1964
Secular variations	0–700 ka	fine sediments	Creer, 1962; 1967; Lund, 1996
Biological			
Dendrochronology	10 ka, depending upon existence of a local master chronology	wood	Fritts, 1976; Jacoby et al., 1988; Yamaguchi and Hoblitt, 1995

Tree Rings

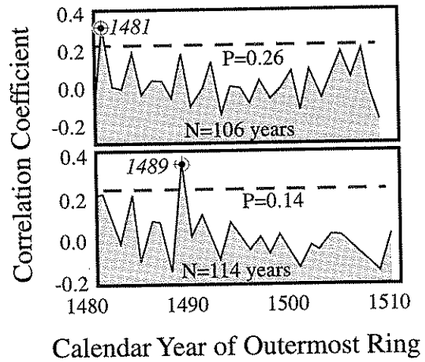
An entire science of dendrochronology has evolved around the dating of surfaces using tree rings and the exploration of climatic change using the tree-ring width as a surrogate for climatic stresses. In climates with a distinct growing season, trees put on new wood in discrete layers that may be counted to reveal the age of the tree. Growth is rapid in the wet season and the summer, then slows in the winter, resulting in relatively low-density wood with large cells in the growth season, separated by layers of denser, finer wood. Whereas the easiest application is obviously to date a surface with the trees still growing (by cutting down the tree or by coring the tree with a boring device), the technique has been augmented greatly by extending the time series derived from living trees through the addition of time series from dead trees. As the thickness or width of the annual growth ring is dictated by the climatic conditions of water availability and temperature, tree-ring width time series can be viewed as a proxy for a climatic time series. That this climate time series reflects at least a regional, if not a global, phenomenon means that trees far separated should record the same temporal pattern of ring widths. We note that one must take care to account for both the species of the tree, and the long-term trend (decrease) in the rate of ring growth through time as any tree matures (Fig. 3.7). A ring-matching technique is therefore based on a master tree-ring time series that is compiled from a set of trees whose ring-width series overlap. For example, Yamaguchi (1995) demonstrated the time of burial of trees on Mt. St. Helens by debris from prehistoric volcanic eruptions by matching the ring-width time series from trees embedded in or buried by lahars with that of the master tree-ring chronology. This entailed choosing a date for the volcanic event (= the age of the outermost ring), calculating the correlation coefficient between the sample tree-ring width series and that of the master tree-ring series, choosing another sample age, recalculating the correlation, and so on. The resulting plot of correlation versus outermost ring age reveals a strong spike in a particular year, which is interpreted to be

FIGURE 3.7. Tree-ring widths as a function of time for three Douglas fir trees in the Pacific NW of the United States



The tree-ring width shows decline in growth rate with age, best seen in Tree 2 (dashed line). Superposed on this are the modulations in growth rate associated with variations of local climate. Modified after Yamaguchi (1995).

FIGURE 3.8. Correlation of tree-width time series with the master tree-ring time series as a function of chosen start year



Top series shows strongest correlation for a start date (outermost ring) of calendar year 1481, while bottom series is strongest for start date of calendar year 1489. Correlation coefficients for all other start years have probabilities of random occurrence that exceed the labeled probability, *P*. Modified after Yamaguchi (1995).

the correct match (Fig. 3.8). The technique relies strongly on a trustworthy master tree-ring time series, and works best for dating large trees with longer ring width time series to test against the master chronology.

Tree ring studies have also provided the basis for calibration of the radiocarbon method (see following section). The carbon from a particular tree ring of known age (from counting backward from the bark) can be ^{14}C dated using recently developed techniques that permit dating of very small samples. This calibration holds out to about 10,000 years, the longest compiled tree-ring record.

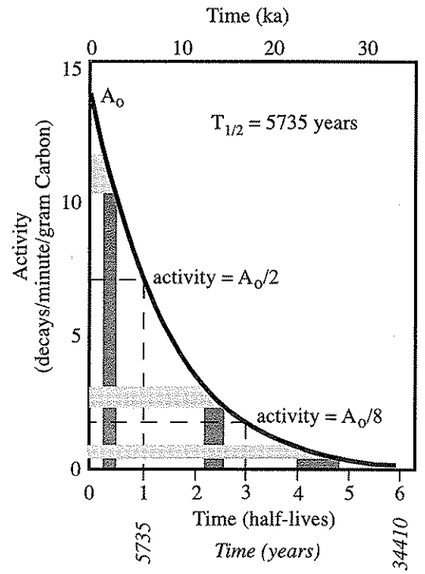
Detailed studies of tree rings have allowed geoscientists to place very precise dates on important short-lived prehistoric geological events, including rockslides, volcanic flows, and even earthquakes. We will revisit this technique in our discussion of prehistoric Cascadia earthquakes.

Radiocarbon Dating

The most commonly used technique to date geomorphic features and surfaces that are less than about 30 to 40 ka employs carbon 14, known also as radiocarbon because it radioactively decays. ^{14}C and ^{13}C are formed in the atmosphere by cosmic radiation interactions with atmospheric N, resulting in a mixture of atmospheric carbon that is roughly 98.9% ^{12}C, 1.1% ^{13}C, and $1.17*10^{-10}$% ^{14}C, most of it in the form of CO_2. As the atmosphere is well mixed on short time scales, this relative abundance is uniform around the globe. It is from this pool that the carbon in CO_2 is fixed through photosynthesis in plants. Plants should therefore be made of organic carbon that mimics the isotopic ratio in the atmosphere during growth, although it is offset from that ratio due to fractionation during photosynthesis. Once a plant or animal cell dies, it no longer incorporates new carbon into its organic material. Radioactive decay takes over, and begins to modify the ratio of the carbon isotopes in the cells. So it is upon death that the clock starts. We will see that the ^{14}C dating technique is intertwined with at least two other methods, U/Th and tree-ring dating, demonstrating how interdependent some of these techniques have become.

There are two principal means of measuring the ratio of ^{14}C to ^{12}C in a sample. Conventional dating is done by

FIGURE 3.9. Decay of ^{14}C concentration with time follows classic exponential curve

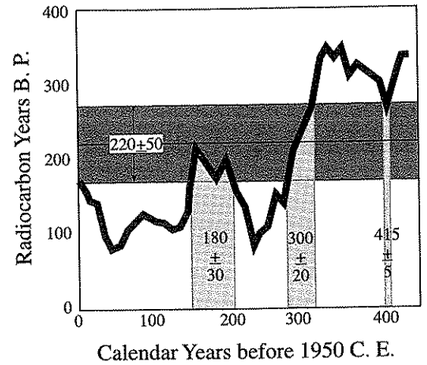


Shaded bars show transformation of uncertainty in activities to uncertainties in age. Uncertainties become very large for low activities. Ages greater than about 7 half-lives cannot be distinguished from infinite ages. Modified after Olsson (1968).

counting the decays of ^{14}C to its daughter ^{14}N through emission of an electron: a beta (β) particle. The rate of decay, as represented by the beta particles, is often measured in decays per minute per gram of carbon and requires samples on the order of a few grams of carbon in size. For old samples, it is necessary to count decays for long periods of time to record enough decays for statistical reliability. Even so, one can see easily from the graph in Figure 3.9 that, because our ability to measure the activity is limited and inevitably has finite error bars associated with it, the meaning of these errors grows for older and older samples until we cannot distinguish between a sample of, say, 40 ka and one of infinite age. This is because, with a half-life of 5735 years, there is a little less than 1% of the original ^{14}C left in a sample if it is older than about 7 half-lives ($(1/2)^7$), or about 40 ka. For example, a gram of modern carbon experiences about 15 decays/minute. If this sample were 40 ka, it would have spanned about 7 half-lives and would generate about $15/2^7$ decays/minute or only about 7 decays/hour. Although such low rates of decay can be counted, incoming cosmic radiation can also trigger the detectors and consequently causes large uncertainties in the detector counts attributable to decay of ^{14}C. (This is why ^{14}C labs are often either deep underground or lead-shielded or both.) Although measuring the decay rate remains the dominant (and far cheaper) technique, it has been augmented tremendously within the last decade or so by the development of accelerator mass spectrometric (AMS) techniques. In AMS dating, the carbon is measured according to its isotopic weight, not according to its radioactivity, or rate of decay. This means that one counts all the carbon atoms in a sample, not just those that decay, so that the sample size can be dramatically reduced to a milligram. When AMS dating was first introduced, it was hoped that AMS would also allow extension of the age range of the technique to many more half-lives. In practice the technique is still not capable of extending beyond about 50 ka. There are both commercial and academic labs that will run samples using one or the other method.

Although ^{14}C dating revolutionized archaeology and allowed scientists to establish the absolute timing from organic remains for the first time in the 1950s, there are pitfalls that one must acknowledge. In order to estimate the age of an object using ^{14}C, one must know 1) the starting ratio of $^{14}C/^{12}C$, 2) the decay rate, and 3) the final ratio. We know the decay rate well, characterized by a half-life of 5735 years, and can measure the present ratio well using AMS or can estimate it based on measured rates of decay. However, because ^{14}C is created in the atmosphere by cosmic radiation, any variations in the production rate of ^{14}C result in a different starting ratio in the organic material. Variations in production rate result from several sources, chiefly variations in the cosmic ray flux, which are associated with fluctuations of the earth's and the sun's magnetic fields. As our measurements of these have lasted for only a few decades, we must rely on calibration. Happily, as mentioned previously, tree rings provide an organic record that can be pieced together back several thousand years—and in the case of bristlecone pines, 10 ka. It is from this record that we know the ^{14}C production-rate record, which is itself of interest to those studying the magnetic fields of the sun and earth. With a ^{14}C ratio in hand, from which

FIGURE 3.10. Correspondence between radiocarbon and calendar years from 1500 to present (jagged black band)



Shaded bar shows how interpretation of a radiocarbon age of 220+50 years can be ambiguous: it could mean the sample came from any of the three light-shaded regions on the calendar year axis. Modified after Porter (1981b).

a radiocarbon age is calculated using the half-life, one may turn to published computer codes (e.g., Stuiver and Reimer, 1993) that allow translation to calendric years (generating a "calibrated radiocarbon age"). The problems are largest at young ages, where the conversion between radiocarbon ages and calendric ages is actually ambiguous (Fig. 3.10): a given sample could be either *x* or *y* years old, owing to a large fluctuation in the ^{14}C production rate within the last couple of millennia. This tree-ring calibration method is valid to about 10 ka, beyond which another calibration method is needed. Here we turn to the precise dating of calcium carbonate deposits (in particular, corals) by both AMS with ^{14}C and U/Th dating.

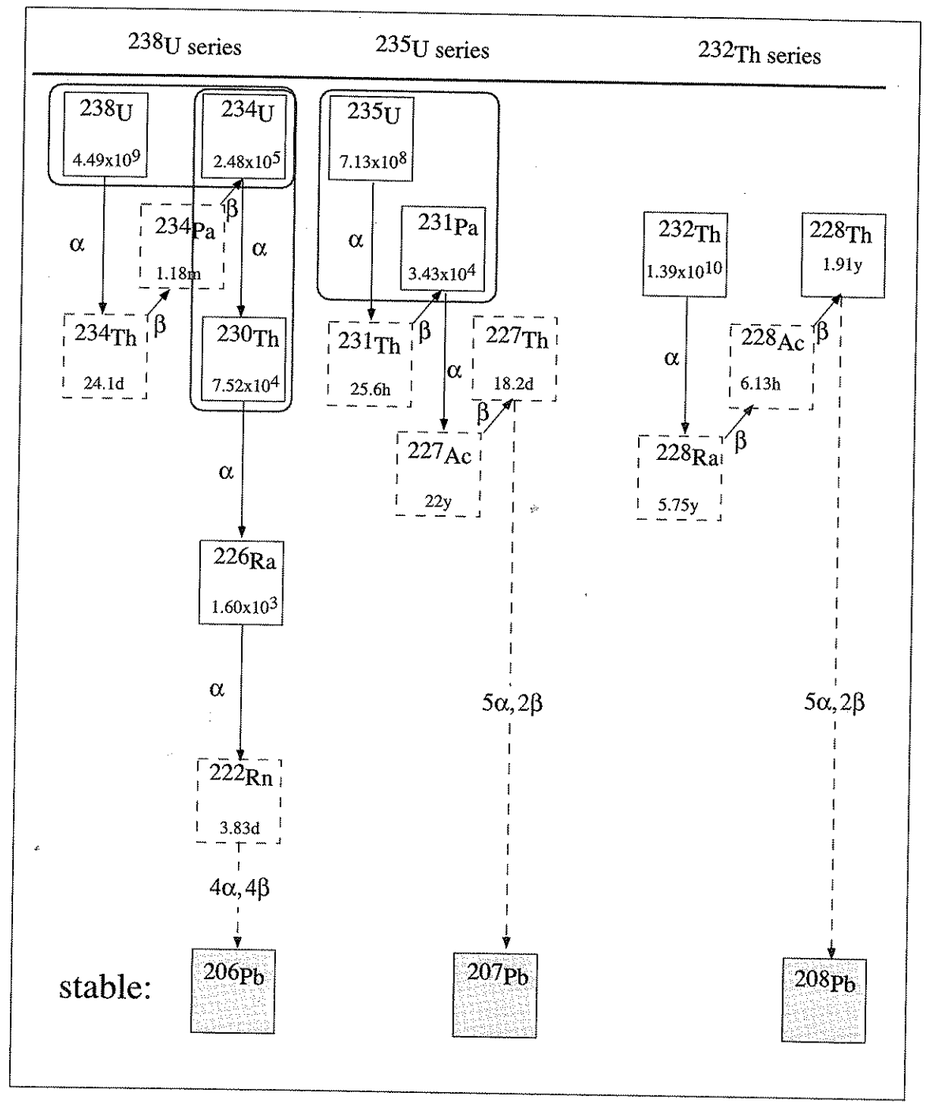
Uranium/Thorium Dating

The uranium decay series represents a very trustworthy set of clocks. The series consists of a set of several isotopes that are generated from their parents at varying rates, and that decay to their own daughters at yet other rates (Fig. 3.11). The ultimate parents of the chains are ^{238}U, ^{235}U, and ^{232}Th. Decays take place at a statistically steady rate that is independent of temperature, of the magnetic fields of the earth and of the sun, and of all other environmental factors that are the bane of many geological and biological clocks. Fortunately, the decay constants for some of these pairs include those that correspond to half-lives that are very useful in neotectonic and climatic investigations.

Because the U/Th series is a more robust clock than is ^{14}C, it has recently been used as a means of calibrating the ^{14}C clock beyond the end of the tree-ring record (Fig. 3.12). In a clever sampling scheme, Bard and others (Bard et al., 1990) used the calcium carbonate from submerged corals off Barbados as a means of obtaining sample pairs from the same material. The corals contained both ^{14}C as some small fraction of the carbon, and U substituting for Ca in the carbonate lattice. Accelerator mass spectrometric (AMS) measurements of the ^{14}C and accurate measurements of the U/Th revealed a consistent offset between the two clocks for samples older than the Holocene. This important result indicates that ages calculated using ^{14}C are as much as a few thousand years too young for samples dating to about 25 ka.

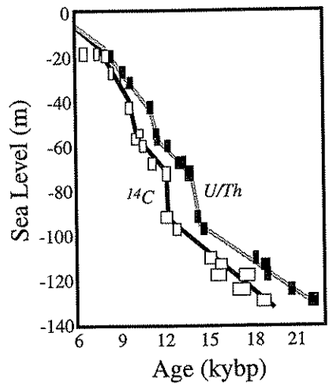
The substitution of U for Ca in carbonate lattices makes shells and carbonate coatings in soils prime targets for U/Th dating. Unfortunately, not all carbonate materials retain the parents and the daughters in their lattice through time, that is, they are not "closed systems." In particular, most shells (for instance, typical clams) leak uranium, making them essentially worthless as clocks. Luckily, most unrecrystallized corals work well, both the colonial and solitary types. In Barbados, Fairbanks (1989) reconstructed past sea levels using *Acropora palmata*, a coral that grows within a couple meters of sea level. The west coast of North America sports numerous marine terraces whose ages are needed in order to determine deformation rates of the coastline. Because most of the coastline is at higher latitudes than those at which colonial corals grow, the target has been instead the solitary coral *Balanophyllia elegans*. These are difficult to find, as they are centimeter scale, and look like wagon wheel spaghettis. Nonetheless, they have been used by numerous workers (see review in Muhs, 1992) to identify which marine terrace was formed during which sea level highstand. U/Th has also been used to date the carbonate coatings on pebbles in soils, whose thickness has already been discussed as a surrogate for soil age. The method entails scraping off all but the innermost coating, adjacent to the clast upon which it is growing, in order to analyze this innermost rind. While this works in some desert settings (e.g., Ku et al., 1979),

FIGURE 3.11. Uranium and Thorium decay chains



Radioactive parents ^{238}U, ^{235}U and ^{232}Th decay through α and β decay steps to stable daughter isotopes of Pb. Half-lives shown in boxes are in years except where noted.

FIGURE 3.12. Paired U/Th and radiocarbon ages of corals



The sampled corals from Barbados grew during rise of sea level from its last glacial maximum of roughly −150 m to within 10 m of modern levels. Dates obtained using both radiocarbon and U/Th series methods are shown. Two strong spikes in sea level rise rates (glacial meltwater pulses) correspond only if the radiocarbon ages are systematically shifted to older ages. The U/Th clock is more trustworthy given its independence from fluctuations in the rate of cosmic ray bombardment. Modified after Bard et al. (1990).

it fails in many, owing either to leakage of one or another product from the carbonate, or to severe impurity of the coatings. This makes it a tool of low priority, to be used only if no other can be identified. Thankfully, one can determine if significant leakage has occurred, and hence whether the age obtained is worthy of interpretation (see, for example, Muhs et al., 1994, for discussion of these tests).

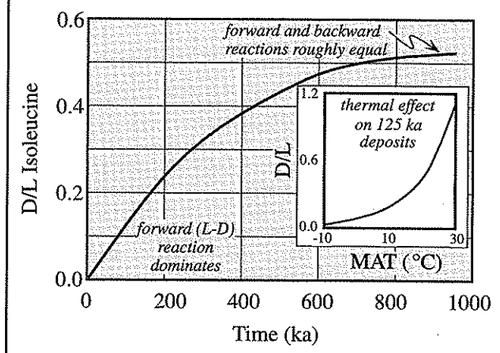
Whereas both societal concerns and geological efforts to understand well the most recent tectonic events often require that we focus on the past thousand years, this interval is difficult to date accurately using ^{14}C (Fig. 3.10). Although the measurement precision in the best labs may be as little as 10 to 20 years, the calendric uncertainty is often many decades due to the fluctuations of atmospheric ^{14}C during the past 1000 yr (see Atwater et al., 1991; Sieh et al., 1989, for examples of high-precision ^{14}C dating). High-precision U-Th dating of corals that grew during this same interval, however, can have calendric uncertainties of <10 yr, and thus can provide a detailed time resolution that was previously unattainable with most other techniques (Edwards et al., 1988; Edwards et al., 1993).

Amino Acid Racemization

All fossils contain at least trace amounts of organic matter than can be retained for long periods of time. The proteins that constitute this organic material in both skeletal and shell material are themselves composed of large numbers of amino acids. After death, the amino acids in these proteins alter through a set of chemical and physical processes. The degree to which this alteration has taken place can be used as a clock. Although transformations represent a very complex set of processes, there are some that appear to be reliable enough to provide both a relative and absolute dating method (Kaufman et al., 1992). Many types of organic materials have been used, including bivalves, gastropods, foraminifera, coral, and the shells of birds. In addition, the gas chromatographic methods used for measurement are relatively inexpensive, making the analysis of large numbers of samples possible. Only about 2 mg of sample is needed.

Living organisms utilize amino acids only in a left-handed (or *L*, for levo-) configuration of their isomers. Upon the death of an organism, these restrictive biological processes are terminated, which frees the amino acids to racemize (flip) into their right-handed (*D*, for dextro-) isomeric state. The ratio of *D*/*L* configurations is therefore a clock. The reaction behaves as a first-order reversible chemical reaction, with an equilibrium established when the backward reaction (*D*-*L*) balances the forward (*L*-*D*). As in all other chemical reactions, the rate constants are temperature dependent, reflecting the Arrhenius relation. Like the radioactive and cosmogenic dating methods that entail both production and decay, this restricts use of the clock to the period of time over which the measurable ratio (here the *D*/*L* ratio) is changing significantly. The time to equilibrium varies from one amino acid to another, and depends strongly upon the thermal history to which a sample has been subjected subsequent to death. For instance, isoleucine attains its equilibrium value of about 1.3 in 100 to 300 ka in equatorial sites, while in the Arctic it may take 10 Ma (Miller and Brigham-Grette, 1989) (Fig. 3.13). The thermal dependence is both a problem to be dealt with in the absolute dating of a sample, and a paleoclimatic tool in itself if the date is already known through other dating or correlation methods (inset in Fig. 3.13). Because the

FIGURE 3.13. Theoretical curve of amino acid racemization through time



For early times, L-D transitions dominate, while back-reactions D-L become roughly equal at later times. Inset shows effects of higher mean annual temperatures (MAT) on reaction rates, and hence of D/L ratio expected at a given age (125 ka). Modified after Kaufman and Miller (1992) and Hearty and Miller (1987).

rate constant is apparently influenced by the type of animal (the taxonomic effect), researchers typically focus on a few common genera.

Typical application of racemization to absolute dating involves sampling a site of known age (say independently dated using ^{14}C), and a nearby site that is presumed to be late Pleistocene. The same fossil genera are collected from each site. The *D*/*L* ratio of the Holocene sample yields an empirical rate constant that is both geographically and taxonomically specific. The interpretation of the late Pleistocene *D*/*L* ratio takes into account the fact that the average temperatures since the late Pleistocene are likely lower than the average since the Holocene sample was deposited, meaning that the use of the Holocene rate constant will yield an underestimate of the late Pleistocene sample age.

Luminescence Dating

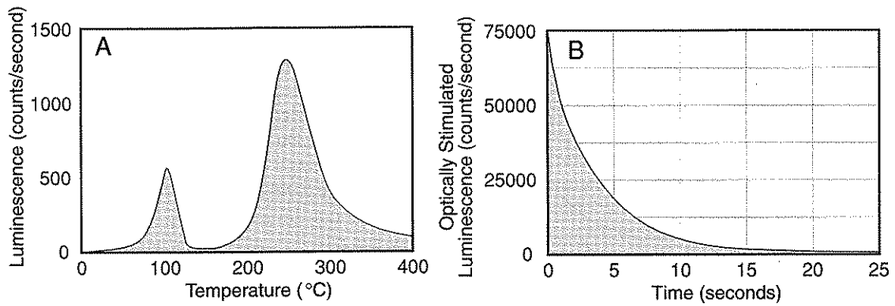
The applicability of many of the dating techniques discussed thus far is often limited by the availability of suitable material for dating, such as carbonaceous debris, shells, or corals. In most terrestrial settings, shells or corals are absent, and even where carbonaceous material is preserved, the age of the deposit of interest may exceed the practical range of ^{14}C. Luminescence dating has the potential to be both broadly applicable to terrestrial deposits and to extend well beyond the age limitations of radiocarbon dating.

The basis of luminescence is the trapping and release of energy by electrons within crystals (Aitken, 1985; Wintle, 1993; Duller 1996). When energy is added to electrons through radiation, they tend to move from a lower energy level (the valence band) to a higher energy level (the conduction band). Some of these energized electrons can be trapped metastably by crystal defects between these two energy levels. Trapped electrons can be made to drop back to their valence band upon the addition of a small amount of energy, at which time they release photons of light equivalent to the change in their energy state. This released light is termed *luminescence*. The radiation that initially causes the energy state of electrons to increase within a crystal comes largely from the decay of nearby radioisotopes within a sedimentary deposit. That this rate of decay should be approximately constant is the basis of the luminescence clock. The electrons can be reset (zeroed) to their low energy state by exposure to either intense heat or sunlight. When luminescence is caused by the addition of heat in the laboratory, it is termed *thermoluminescence* (TL), whereas when it results from the addition of light, it is termed *optically stimulated luminescence* (OSL).

In TL dating, luminescence is measured as the sample is incrementally heated (Fig. 3.14A). In OSL, the sample is exposed to light of a certain wavelength and the emitted luminescence is measured as a function of time (Fig. 3.14B). OSL has some important advantages over TL that are making it a generally preferable dating technique (Aitken, 1998). In particular, TL measures luminescence signals that are sensitive to light, heat, and any other types of energetic stimulation, whereas OSL measures only signals that are sensitive to light. In addition, TL destroys the signal of interest during measurement and, therefore, cannot be repeated, whereas OSL can be applied in short bursts that modify the total luminescence only slightly, such that multiple measurements can be made.

The number of electrons residing in "traps" in a crystal, and the intensity of luminescence released in the laboratory due to exposure to heat or light, are a function of the total dose of radiation received by the sample over time. If one can quantify that past radiation exposure, termed the

FIGURE 3.14. Thermal and optically stimulated luminescence



A. Example of a thermoluminescence "glow curve" that results from the progressive heating at a rate of a few degrees/second. In this quartz sample, most of the traps are emptied by 400°C. The peak at 100°C results from optical stimulation for 25 seconds at 420–560 nm shown in (B) and is not seen in the natural sample. B. Luminescence signal recorded at ~340 nm due to optical stimulation of the same sample as in (A) using a light source with wavelengths ranging from 420–560 nm. Modified after Duller (1996).

"paleodose," and the rate at which the sample was irradiated, then an age for the sample can be calculated:

$$age = paleodose\ P\ (Gy) / \ dose\ rate\ (Gy/ky)$$

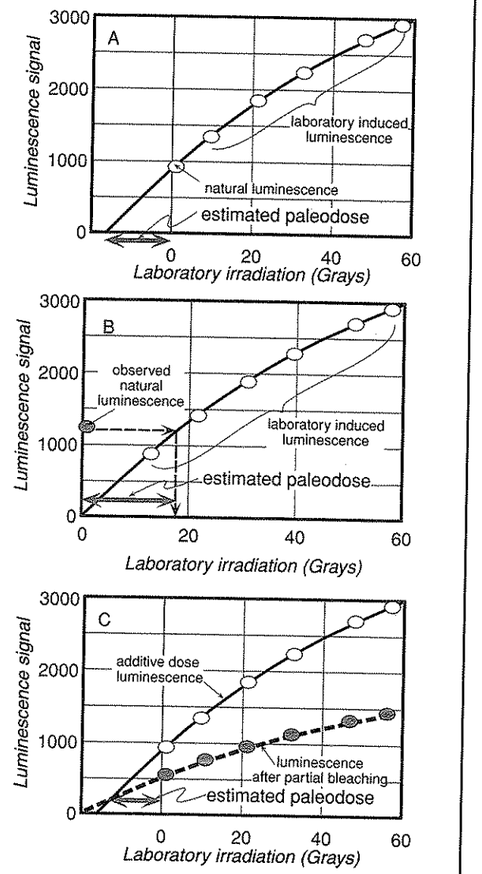
where Gy (grays) is the SI unit for radiation.

The dose rate is typically measured either *in situ*, by leaving a radiation detector for an extended time (1 year, for example) in the sediment from which the sample was collected, or by laboratory measurements of the sample itself. Ideally, samples should be collected from the center of homogeneous beds that are at least 60 cm thick, because a 30-cm radius defines the approximate volume that will produce most of the radiation received by a sample. Collecting from such thick, homogeneous beds is especially important when the dose rate is to be determined in the laboratory.

The paleodose can be calculated in one of several ways (Duller, 1996). The additive dose method relies on measuring the luminescence resulting from different levels of irradiation in the laboratory. A curve drawn through these points and through the point defined by the sample's natural luminescence is extrapolated to zero to estimate the paleodose (Fig. 3.15A). In the regenerative method, several subsamples are measured for their natural luminescence and the remainder are zeroed through exposure to light, then exposed to known doses of radiation and remeasured for their luminescence. The curve developed from these measurements can be matched against the observed luminescence to estimate the paleodose (Fig. 3.15B). In the partial bleach method, some subsamples are subjected to the additive dose method, whereas others are exposed to a short burst of light (partial bleaching) in order to remove a proportion of their light-sensitive luminescence prior to measurement. The paleodose is defined by where these two extrapolated curves cross (Fig. 3.15C).

When luminescence dating is applied to sediments, a key assumption is that the traps were emptied or "zeroed" prior to deposition, such that each grain began its life within the deposit with no record of previous radiation. For optically stimulated luminescence, prolonged exposure to sunlight is sufficient to empty the traps, whereas for TL, sunlight will partially, but not completely remove the thermally stimulated luminescence. In either case, one must assess the sedimentological setting in the field, asking whether it is likely that there was sufficient exposure to sunlight during transport to empty the traps. Loess and aeolian sands, both of which blow around on the surface, are good bets to have been zeroed. Lacustrine silts and clays, which are transported as suspended sediments and should receive sunlight in the upper parts of the water column, also seem to yield reliable dates. On the other hand, sediment that traveled as bedload or

FIGURE 3.15. Paleodose determination in luminescence dating using three methods (see text)



A. Additive dose method. B. Regenerative method. C. Partial bleach method. Modified after Duller (1996).

was transported in mass flows would be generally poor candidates for luminescence dating. The range over which luminescence dating can be successfully applied theoretically extends from less than 1 ky to more than 1 My. Using OSL and very sensitive detectors, aeolian sediments only a few hundred years old have been dated (Wolfe et al., 1995). Unfortunately, for dating older strata, during prolonged exposure to radiation, the trapping sites may eventually become saturated; their luminescence signal will cease to increase linearly despite additional radiation. Sediments derived from highly radiogenic rocks, like granites, will experience higher dose rates and reach saturation long before sediments derived from rocks like carbonates that emit low-levels of radiation. Practically speaking, high dose rates may limit luminescence dating in granite source areas to less than 60 to 200 kyr. Nonetheless, this technique has the potential to fill a critical gap beyond the range of radiocarbon dating. Moreover, because target minerals of quartz and feldspar are abundant minerals in many strata, these techniques can be applied to aeolian and lacustrine beds where few other means of dating may be possible. Some studies have tried to use luminescence dating in alluvial fans and colluvial wedges associated with faults. In such settings, zeroing of the sediment by sunlight prior to deposition cannot safely be assumed. Use of an additional dating technique, such as radiocarbon, would have to be employed on at least some pairs of samples to validate the luminescence dates.

Cosmogenic Radionuclide Dating

Within the last two decades, the nuclear physics community has introduced the geologic community to a new technology that allows dating of bare bedrock surfaces and of alluvial deposits that have been continuously exposed to cosmic radiation since formation. These techniques are fast-evolving. For the first decade much of the work was performed in geomorphic situations in which both the geomorphic and nuclear physics communities could benefit. Using surfaces previously dated using other methods, the production rate resulting from cosmic ray bombardment of rock was calibrated, showing a strong dependence on both altitude and latitude. Since then, surfaces of unknown age have been dated in a wide variety of settings. As the materials being dated are commonly available rocks, and the time scale over which the technique may be applied covers the entire Quaternary, this new technique is often the only method available. Here we briefly review the theory of the use of cosmogenic radionuclides (CRNs) in a variety of geomorphic settings. This treatment is by no means exhaustive, but is intended to serve as an introduction to this new and exciting field. More detailed reviews may be found in Lal (1991), Morris (1991), Bierman (1994), and Cerling and Craig (1994). We will see that the major problems facing the community lie in the interpretation of the CRN concentrations, that considerable care

needs to be exercised in sampling appropriately, and that much effort must be put into developing a relevant geomorphic model of the site. Although much work has been accomplished in this new field, the techniques continue to evolve.

Background

We summarize here and in the cosmogenic primer (Box 3.1) the concepts that underpin the essential uses of cosmogenic radionuclides in geomorphological studies. Cosmic rays isotropically bombard the solar system, meaning they come in essentially uniformly from all angles. Being charged, they are steered by the magnetic field of the earth, generating a stronger beam of particles at high geomagnetic latitudes, and reducing the downward flux at lower latitudes. These particles interact with atoms in the atmosphere, creating such familiar species as ^{14}C. Such interactions reduce the number of energetic particles that penetrate to lower altitudes in the atmosphere; the production rate of radionuclides declines with a 1/e length scale of roughly 1.5 km within the lower atmosphere. Due to this atmospheric attenuation, production rates of cosmogenic nuclides at altitudes of 3 km will be e^2, or about seven times higher than at sea level. These nuclides generated in the atmosphere have been dubbed "garden variety" radionuclides. Cosmic rays that survive to impact the surface of the earth are capable of producing CRNs in near-surface materials: "in situ" cosmogenic radionuclides. The production rate is highest at the surface, and decays with depth with a 1/e scale of roughly 60 to 70 cm, the difference in length scales reflecting the relative density of rock and air. Minerals comprising atoms that are susceptible to the nuclear reactions that generate CRNs must be present in these near-surface materials to be useful. For instance, and quite fortunately, the very common mineral quartz (SiO_2) is a target mineral for both ^{10}Be and ^{26}Al, ^{10}Be being produced from ^{18}O, and ^{26}Al being produced from ^{32}Si. See the table in the cosmogenic box (Box 3.1) for a compilation of the most commonly used species, their half-lives, and the sea-level, high latitude production rates.

Note that the production rates are only a few atoms per gram of target mineral per year (see Box 3.1), rates that result in such low concentrations that conventional mass spectrometry and counting of decays are unreasonable methods for measuring concentrations. Analyses of the concentrations of those CRNs produced in situ requires first separation of the target mineral from the rock, ridding the sample of garden variety CRNs by a leach step, and subsequent chemical separation of the CRN. The minuscule sample is then analyzed in an accelerator mass spectrometer (AMS), several of which have been converted from use in physics and defense industries for application in this field.

In Situ CRNs

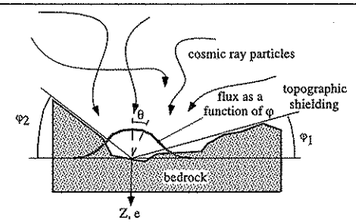
In situ CRNs have been used in two distinct settings: bare bedrock surfaces, where one is interested in either or both of the exposure age of the surface and the erosion rate of it, and depositional surfaces. Each setting poses its own dilemmas, requiring differing sampling strategies. But there is often no other game in town.

Bedrock Surfaces. Consider a bare bedrock surface exposed to the full cosmic ray flux (no blocking by nearby outcrops or valley walls). The concentration in a parcel of rock is dictated by the differential equation describing both production and decay of CRNs in the parcel:

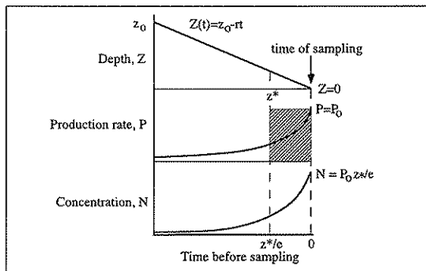
$$dN/dt = P - \lambda N \qquad (3.3)$$

where N is the number of CRNs per unit volume of rock, t is time, P is the production rate, and λ is again the decay constant (related to the half-life by $t_{1/2} = \log(1/2)/\lambda$). Much of the complexity in the interpretation of the measured CRN concentrations resides in the history of the production rate, P, to which the parcel has been subjected. Let us address a few simple examples. Consider a bare bedrock surface exposed at $t = 0$ by a landslide. This treatment will equally well apply to a sample obtained from within the headscarp, or from a boulder on the surface of the landslide which we can safely assume to have been at great depth prior to the slide we would like to date. Assume that the rock involved is much older than the half-life of the CRN we wish to employ, assuring that all CRNs produced in a prior exposure at the surface have decayed. In other words, all "inheritance" is negligible. CRNs will begin to accumulate within the rock, most rapidly at the surface, more slowly at depth. At early times, the concentration, N, is so low everywhere within the rock that the rate of growth is essentially linear at the local production rate P, i.e., $N = Pt$. Since P falls off exponentially with depth (see box), an exponential concentration profile develops: $N = P_0 te^{-z/z^*}$, where P_0 is the production rate at the surface and z^* represents the depth in the rock at which the production rate is P_0/e. If a sample is collected from this surface ($z = 0$), and we know the local production rate

Box 3.1. An in situ cosmogenic radionuclide primer

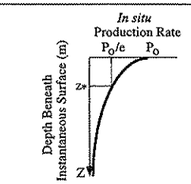


Cosmic ray particles are isotropically distributed as they enter the earth's atmosphere. The atmosphere attenuates this production such that the particle flux is lower from higher angles, θ, with respect to the zenith (gray cosine curve). On a flat surface they would produce cosmogenic radionuclides at a rate P_o. Any intervening rock also blocks production (called topographic shielding), reducing the effective surface production rate below P_o. Measurement of this effect is done in the field by measuring φ in 4–8 directions. The sample site, at $Z = 0$, may be eroding at a long-term rate of r.



The case of steady erosion. As the depth of the sample decreases due to erosion of overlying rock, the production rate increases exponentially. The concentration therefore increases exponentially. The concentration upon sampling the parcel when it reaches the surface is equivalent to the integral of the production rate history. Given that z^* is a material property, dependent largely upon the density of the overburden, the sample concentration is higher for slower erosion rates, and for higher surface production rates. Note that the sample concentration is equivalent to that the parcel of rock would have obtained had it sat on the surface for a time equal to that it takes the rock to be exhumed by a length z^* (shaded box).

References: Lal, 1988, 1991; Nishiizumi et al., 1994; Bierman, 1994; Nishiizumi et al., 1996



Production rate, P, decays exponentially with depth below the instantaneous surface, z, where the surface production rate, P_o, is scaled by altitude and geomagnetic latitude. The rate of decay beneath the surface is scaled by the constant, z^*, which is roughly 50–60 cm for most lithologies.

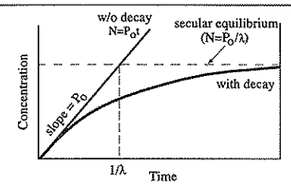
In general, the rate of change of concentration (N) equals the rate of production less the rate of decay (λN):

$$\frac{\partial N}{\partial t} = P(t) - \lambda N$$

The rate of production with depth is

$$P = P_o e^{-z/z^*}$$

The depth can vary with time, z(t), due to either erosion or deposition.



The case for no erosion. When the nuclide is stable, the concentration simply increases linearly with time at a rate P_o. When the nuclide is radiogenic, the concentration approaches a steady value, or secular equilibrium, at which point the rate of decay equals the rate of production. The achievement of secular equilibrium takes several characteristic decay times ($1/\lambda$).

	^{14}C	^{10}Be	^{26}Al	^{36}Cl
P_o	21.0	5.81	34.9	4 - 9
λ	1.2×10^{-4}	4.62×10^{-7}	9.9×10^{-7}	4.81×10^{-7}
$1/\lambda$	8276	2.16×10^{6}	1.01×10^{6}	2.1×10^{6}
$t_{1/2}$	5735	1.5×10^{6}	0.7×10^{6}	0.3×10^{6}

Commonly used cosmogenic radionuclides, their in situ production rates (P_o) in atoms/gram of quartz/year at sea level in mid to high latitudes, their decay time scales (λ, $1/\lambda$) and their half-lives (in years).

at the surface, then measurement of N allows solution for the exposure time, t. However, as the concentration builds up, the second term in equation 3.3, representing decay, begins to play a larger role. This results in a decline in the rate of increase in concentration, until ultimately there is a balance between new production and decay. This "secular equilibrium," represented by $dN/dt = 0$, limits the concentration to a maximum of $N = P/\lambda$ (see Box 3.1). Note that in these circumstances, when the sample has reached secular equilibrium, no information about the exposure time can be extracted from a measurement of N. Happily for the geomorphic community interested in dating surfaces in Quaternary Period (roughly 1.8 million years long), this secular equilibrium takes several half-lives to be achieved. Since ^{10}Be and ^{26}Al have half-lives on the order of 1 million years (see table in Box 3.1), they remain useful throughout the Quaternary. Dating of glacially polished surfaces (e.g., Nishiizumi et al., 1989) in the Sierras is one example of dating surfaces of negligible erosion. Exposure dating of fluvial strath terraces along the Indus River, on which the fluvial polish is still intact and where the fluted and potholed nature of the fluvially carved bed has clearly not degenerated since abandonment of the strath, is another example of this circumstance (Burbank et al., 1996).

More commonly, the rock surface being sampled is eroding at some rate that we would like to determine. This erosion can take place on a grain-by-grain basis, or through spalling of parcels of various finite depths representing joint spacing, for instance, or fire spall depths (e.g., Bierman and Gillespie, 1991). If the erosion is continuous and steady, the production rate in equation 3.3 can be rewritten

$$P = P_o \exp(-rt/z^*) \tag{3.4}$$

where r is the erosion rate of the surface, and t is time, $t = 0$ being the time of sampling and t into the past being positive. For the moment ignoring decay, and solving equation 3.3 with this steady production-rate history yields:

$$N = P_o(z^*/r) \tag{3.5}$$

which can be solved for the erosion rate. Here the term (z^*/r) represents the time it takes the sample to travel through the boundary layer where the production rate is significant. The faster the parcel is exhumed, the lower the resulting concentration. Including decay alters the equation to

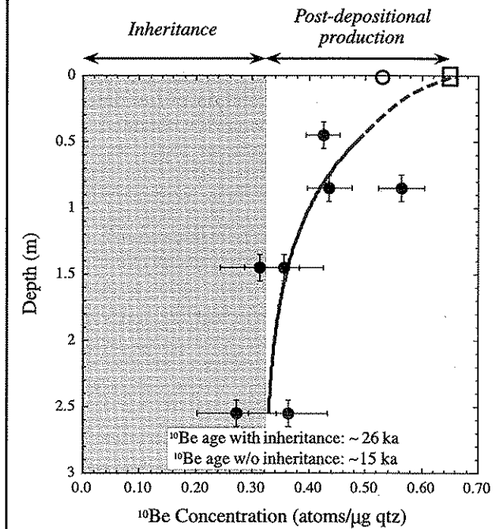
$$N = P_o z^*/(r + (z^*/\lambda)) \tag{3.6}$$

If, on the other hand, the erosion occurs in steps (due to fire spalls, for instance (Bierman and Gillespie, 1991) or to joint block removal (Small, 1997)), then the calculation of a long-term erosion rate from employment of Equation 3.6, and the CRN measurement of a surface sample, will vary depending on the time since the last spall event. The error in the estimate will depend on the thickness of the spall. In situations like those involving spalls, the averaging of several sample sites ought to provide a more robust sense of the mean rate at which the surface is being lowered (Small, 1997).

If one would still like to extract an exposure age from this site, and not just an erosion rate, Lal (1991) has shown that the use of two cosmogenic radionuclides with differing half-lives allows some additional constraint on the exposure age in certain scenarios. Working in the ^{36}Cl system, supplemented by ^{10}Be analyses, Phillips et al. (1997) has demonstrated the usefulness of this technique in constraining the age of large boulders whose surfaces have been slightly eroded since emplacement on the surface to be dated.

Depositional Surfaces. Depositional surfaces, such as fluvial and marine terraces, present their own problems, even if the surfaces can be safely assumed not to be eroding. The principal problem lies in the possibility that the clasts being used to assess the concentration of CRNs have likely experienced prior exposure ("inheritance") elsewhere within the geomorphic system prior to being deposited on the surface to be dated. Consider first a fluvial terrace. The inheritance derives from a combination of exhumation through the cosmogenic nuclide-production boundary layer as the hillslope surface is lowered; transport within the hillslope system; transport within the fluvial system, which will entail occasional burial in fluvial bars; and final deposition on the terrace to be dated. Tactics must be employed in the sampling of the terrace materials that allow separation of the inheritance signal. In several fluvial systems studied to date, the inheritance can represent a significant (several tens of percent) portion of the total CRN concentration measured in clasts sampled from the terrace surface (e.g., Anderson et al., 1996; Repka et al., 1997) (Fig. 3.16). Because the inheritance differs markedly from one surface to another, one must employ some means of constraining this inheritance on a site-by-site basis.

FIGURE 3.16. Use of cosmogenic radionuclide concentration profile to deduce both inheritance and age of the surface



Dated surface is the Stockton bar of Lake Bonneville, associated with the latest highstand of the lake, at roughly 14.5 ka. Grayed box represents the concentration due to inheritance of cosmogenic radionuclides by the quartzite clasts prior to deposition in the bar. Best-fit line is exponential shifted to account for inheritance. Age deduced from this method is 15 ka. If inheritance were not accounted for, and a single surface sample were used to deduce age, the estimated age would have been ~26 ka, too old by roughly 11 ka.

One apparently effective strategy is to collect samples at varying depths in the terrace profile. Each sample should consist either of sand or of a contribution of an equal amount of rock from numerous clasts buried at the same level, for example, 10 grams from each of 30 clasts. Assuming that the terrace aggraded quickly so that each sampled layer was rapidly buried to near its present depth, then the age of each sample represents a combination of the time since deposition and the inheritance at the time of deposition. When samples are drawn from deep enough in the terrace, say greater than ~2m, their concentrations should reflect almost exclusively the inherited nuclides, because CRN production rates are negligible at this depth.

These kinds of surfaces have also been dated using garden variety CRNs produced in the atmosphere and subsequently rained out in precipitation (see review in Morris, 1991). In particular, ^{10}Be is found to be useful in that its chemistry is such that it is held tightly by clays in soils (Pavich et al., 1984). The soil therefore acts as a reservoir within which the ^{10}Be slowly builds up with age of the surface. If surfaces are well chosen to limit the role of soil erosion (i.e., are nearly flat), then the total ^{10}Be inventory on grain surfaces within the soil column can be used to constrain the age of the surface.

Using Cosmogenic Nuclide Dates

The numbers one extracts from such studies are, at best, only as good as the knowledge of the production rates over the age of the surface at that site. These production-rate histories are difficult to know. Typical calibrations rest on independent dates from surfaces, such as glacially polished bedrock associated with moraines that can be dated with ^{14}C (e.g., Nishiizumi et al., 1989). Two cautions are warranted: 1) these independent ages can be and are being reassessed with new techniques at new sites that may better constrain the true age of the glacially polished surface (e.g., Clark et al., 1995), and 2) production rates estimated from these sites, that are therefore relevant to this altitude and latitude, are averaged production rates over the age of the sample. Use of these production rates for dating other surfaces must be done carefully, as the magnetic field intensity has been shown to vary significantly over the last 140 ka (e.g., Meynadier et al., 1992). This record of field intensity variations will no doubt be extended by future research. The production-rate history at a sampling site should be calculated from a knowledge of the average production rate deduced from a sample of known age, the magnetic field history, and the relationship between field intensity and production rate (e.g., Clark et al., 1995).

In addition, the quality of the interpretation of the CRN concentration extracted from a sample, for either the exposure age or the erosion history at the site, is dependent on the degree to which the geomorphic processes active at that site over the history of the surface can be captured in a quantitative model. Not only must the inheritance of a particular sample be assessed, but the post-depositional processes that might alter the production-rate history of the sample should be dealt with. On boulders presently at the surface, one must worry about whether the boulder has always been at the surface, as well as the likelihood of erosion of the boulder once it was emplaced on the surface. At present,

there is much interest in this new set of techniques. This set of problems provides an excellent opportunity for the marriage of new chronologic tools with new modeling tools, to which we turn later in the book.

Conclusions

In this chapter, we have sampled the array of dating methods available to provide the time scales for tectonic geomorphic studies. The choice of method must be dictated by the availability of the proper materials, the details of the geomorphic setting, and, of course, the cost. While relative dating methods are "quick and dirty," they should not immediately be shunned. They do produce immediate results, while other techniques might take months to a year for processing of samples. In using any method, it is incumbent upon the researcher to document carefully both the geomorphic and depositional setting. Interpretation of the dates, sometimes painstakingly and expensively obtained, relies on this field information.

CHAPTER 4

Stress, Faults, and Folds

When rocks are subjected to stresses that exceed their strength, they rupture, fold, or flow. Different varieties of faults (strike-slip, normal, thrust) characterize contrasting tectonic settings and stress regimes in the upper crust. When faults break the earth's surface either in a single earthquake or during many seismic events, they often create geomorphic features that can be associated with a particular type of fault. Sometimes earthquakes occur on faults that do not reach the earth's surface, and therefore, there are no ground ruptures directly associated with the fault trace. Nonetheless, the earth's surface will deform by folding in response to these earthquakes along buried faults. The geometry of deformation and the evolving shape of a fold can reveal useful information about the nature of the subsurface faulting and the way in which the rocks adjacent to the fault respond to fault motions.

In order to take full advantage of the information to be gleaned from geomorphological surfaces, it is important to understand the typical ways in which rocks deform due to both seismic and aseismic movements. A primary goal of this chapter is to examine the nature of faulting and folding and to discuss concepts related to scales of faulting and similarities between successive earthquakes on the same fault, the displacements of the ground surface that are expected for different types of faults, and the geomorphic imprint of faults of different types.

Stress, Strain, and Faults

Stress

All rocks in the crust are subjected to forces due to gravitational acceleration; the mass of rocks, water, and air around them; and plate tectonic motions. We typically identify two kinds of forces: *body forces* and *surface forces*. Body forces act equally on every element throughout a volume. The magnitude of a body force is proportional to the mass of the element on which it is acting, or its volume times its density. An example of a body force would be the weight of an object, which is the mass of the object times the acceleration due to gravity. Surface forces are forces that act across or along surfaces, and include forces such as friction or pressure. Strictly speaking, these are termed *tractions*, which are defined as the force per unit area acting on a surface.

A stress or traction acting on a plane with any orientation can be resolved into two components: a normal stress (σ_n) acting perpendicular to the surface and a shear stress (σ_s) acting parallel to the surface. The units of stress in the mks (meters-kilograms-seconds) system are pascals. One pascal (Pa) represents the stress produced by one newton acting across or along a surface of one square meter ($(kg \cdot m \cdot sec^{-2}) \cdot m^{-2}$). Pressure can also be expressed in bars, whereby one bar equals 10^6dynes/cm^2. Because one pascal represents a rather small force acting over a large area, pressures are often expressed in megapascals (MPa = 10^6 pascals) which is also equal to 10 bars. A column of rock 1 km high would typically exert a pressure at its base of 25–30 MPa. The shear strength of many crustal rocks ranges from 10–100 MPa.

Each of the tractions acting on a rock surface and resulting from tectonic, lithostatic, buoyancy, or hydrostatic forces can be represented as a vector which can be summed with all other imposed tractions to define the total magnitude and orientation of the imposed stress on any specified plane. The total stress can be subdivided into three orthogonal components, which are typically labelled σ_1, σ_2, and σ_3, for the maximum, intermediate, and minimum principal stresses, respectively. For simplicity in discussing types and

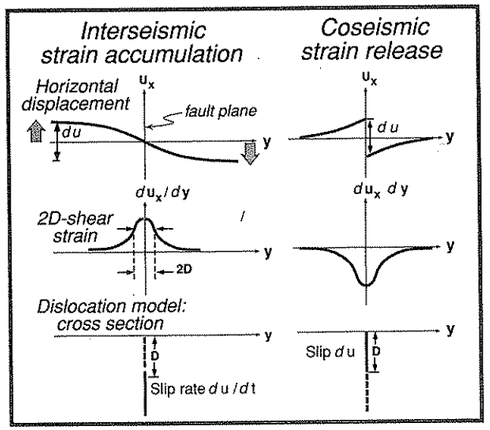
orientation can be determined with a borehole camera following depressurization. Hydrofracturing has the additional advantage that the water pressure that causes fracturing or reopening and propagation of existing fractures is known and is expected to be equivalent to the magnitude of the minimum horizontal stress. With additional knowledge of the ambient pore pressure in the rocks, the maximum compressive stress can also be estimated.

The Earthquake Cycle

The "Classical" Model

Since at least the turn of the century, geologists have attempted to understand the deformation that precedes, coincides with, and follows an earthquake. The time and deformation that encompasses an earthquake and all of the interval between successive earthquakes is termed the *earthquake cycle*. As originally described, the earthquake cycle had two parts: an interseismic interval and a coseismic one (Reid, 1910). Imagine two nearby pieces of the crust that are moving in opposite directions with respect to a fault that separates them. At some depth below the surface, the fault slips continuously and aseismically in a zone of ductile deformation, but in the brittle crust during the interseismic interval, the fault is "locked" such that no slip occurs along it (Fig. 4.3). At some distance from the fault, the rocks are moving at the same rate as the crustal blocks. The amount of displacement decreases to zero at the fault, such that an originally straight marker that was oriented perpendicular to the fault would be bent into a sigmoidal shape. The 2-*d* shear strain, which could be envisioned as the amount of bending of the formerly straight marker, is greatest near the fault (Fig. 4.3). This bending can be considered as "elastic strain" in that it is recoverable, rather than permanent. When the frictional strength of the fault is exceeded by the imposed stress, the fault ruptures during an earthquake. Coseismic displacement is greatest along and adjacent to the fault. In the context of the earthquake cycle model, slip is just enough to balance the slip deficit and restore the marker to a linear trend, perpendicular to the fault, but now offset by the amount of relative motion of the two crustal blocks during the entire earthquake cycle. In this model, all of the elastic strain is recovered in each seismic event, so that there is no permanent strain within the blocks that slip past each other on opposite sides of the fault (Fig. 4.3), although the rupture causes an abrupt discontinuity along the fault itself.

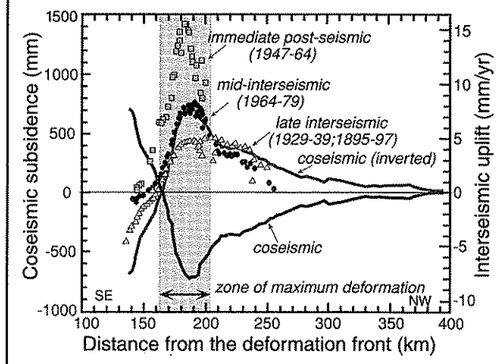
FIGURE 4.3. Model of the earthquake cycle for a strike-slip fault



The far-field strain (large arrows; top panel) of one block with respect to the other remains constant through time, as does aseismic slip (du/dt; bottom panel) in the ductile zone below the locking depth (D). Interseismic displacement is greatest farthest from the fault, but most of the shear strain (du_x/dy; middle panel) occurs within two locking depths (2D) of the fault. During coseismic displacement, the greatest displacement (du) occurs along and near the fault and compensates for the "slip deficit" developed during the interseismic interval. The coseismic shear strain is equal and opposite to the interseismic shear strain. Modified after Thatcher (1986b).

Data on pre- and coseismic deformation prior to and following the 1946 M8 earthquake in the Nankai trough of southwestern Japan provide support for aspects of this model (Fig. 4.4). Measured interseismic strain replicates the shape, but is opposite in vertical direction to the coseismic strain caused by that large subduction-zone earthquake. This behavior is clearly consistent with Reid's model. There is, however, considerable topography across the deformed zone which suggests that coseismic subsidence does not fully compensate all of the interseismic uplift: some permanent strain is represented by this topography.

FIGURE 4.4. Comparison of leveling data recording interseismic, coseismic, and postseismic deformation in the Nankaido region, Japan

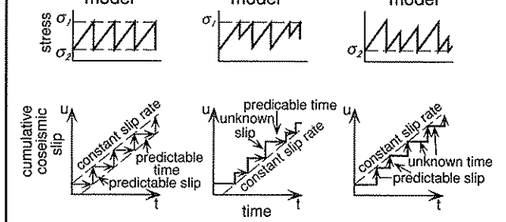


Three intervals of interseismic or postseismic deformation are represented: immediate postseismic (1947–64); mid-interseismic (1964–79); and late interseismic (1895–97, 1929–39). These are compared with the inverted coseismic deformation profile. Note the spatial coincidence of the peak zone of deformation for each data set. The regional pattern of interseismic deformation closely mimics the inversion of the coseismic deformation. Modified after Hyndman and Wang (1995).

Alternative Earthquake Models

As a result of observations related to numerous earthquakes, Reid's simple model of the earthquake cycle has been significantly modified. First, the deformation observed along faults, such as folds and fault offsets, indicates that not all of the pre-faulting strain is recovered. In other words, as seen in the example from Japan (Fig. 4.4), there is often some permanent, unrecoverable strain. Second, the interseismic interval may sometimes be further subdivided into a post-seismic interval and an interseismic one. The post-seismic interval immediately follows an earthquake and is one during which strain accumulates more rapidly than in the subsequent interseismic interval (Stein and Ekstrom, 1992). In some recent earthquakes, the amount of post-seismic deformation and energy release has been shown to be equal to that of the coseismic event (Heki et al., 1997). Newly recognized "slow" earthquakes (Beroza and Jordan, 1990; Kanamori et al., 1993) in which deformation occurs over periods of hours to years can release vast amounts of energy and cause large-scale deformation without generating the catastrophic energy releases of normal earthquakes. Third, new variations on Reid's periodic earthquake cycle have been postulated. In Reid's periodic model, the frictional strength of the fault, the stress drop, and the slip associated with each earthquake is constant from event to event, such that both the time of each earthquake and its magnitude are predictable (Fig. 4.5A). In a *time-predictable model*, earthquakes always occur when a critical stress threshold is attained (Fig. 4.5B). The amount of stress drop and the magnitude of slip, however, varies from one earthquake to the next. Assuming there is a constant accumulation rate of interseismic strain and assuming that the amount of slip in the previous earthquake is known, this model permits a prediction of the time until the next earthquake, because the stress needed to attain the failure stress is known. The displacement of that forthcoming earthquake, however, is unknown. Alternatively, a *slip-predictable model* suggests that slip during an earthquake always terminates when the stress has dropped to a critical level

FIGURE 4.5. Models for earthquake recurrence



A. Periodic earthquake model in which stress levels at the time of rupture and after it are known. These yield a predictable time and slip for each earthquake.
B. Time-predictable model based on a consistent stress level at which failure occurs. Stress drop and slip magnitude are unpredictable, but given previous slip, time until the next earthquake (with unknown slip) is predictable. C. Slip-predictable model based on a consistent stress level at the end of an earthquake. Given time since the last rupture, magnitude of slip is predictable. Modified after Shimaki and Nakata (1980).

(Fig. 4.5C). The stress level and strain accumulation at the time of rupture, however, varies between earthquakes, as does the time between successive ruptures. Thus, given a constant interseismic strain-accumulation rate, knowledge of the time since the previous earthquake permits prediction of the amount of slip that would occur at any given time. But, it doesn't reveal when the next earthquake will occur, because no critical stress threshold is common to successive earthquakes.

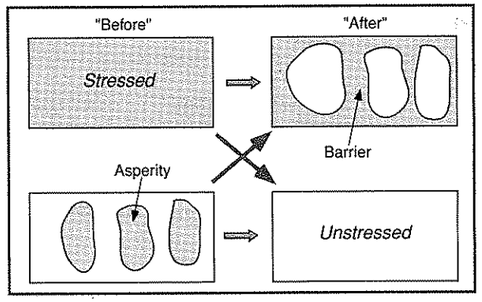
Asperities, Barriers, and Characteristic Earthquakes

What Controls Ruptures?

In an effort to understand why and when earthquakes occur, many scientists have searched for patterns of ruptures which, if fully understood, could form a basis for predicting subsequent seismic events. It is clear that earthquakes occur when the strength of a fault is exceeded. But, what controls the strength of a fault and does the strength change between earthquakes? In southern California, all of the recent large earthquakes ($M \geq 5$) have occurred on faults that were "loaded," that is, the stress was increased on them by a nearby earthquake in the previous 18 months (Harris et al., 1995). Moreover, it seems probable that there are localized segments or patches of some fault planes that control the strength of some faults, whereas other parts are at times relatively stress free (Fig. 4.6). These "sticky" patches of faults have been termed *asperities*. These may be analogous to irregularities that jut out from a planar surface and which, if two planar surfaces were juxtaposed, would be the points where frictional forces would have to be overcome in order to permit sliding between the surfaces. Thus, asperities represent irregularities that can concentrate stresses along a fault, and their strength with respect to shear stresses exerted on them may control when faulting occurs. If the strength of asperities persists from one earthquake to the next, it might be expected that the recurrence pattern would follow a time-predictable model (Fig. 4.5B).

Alternatively, earthquakes may be thought of as causing ruptures along a heterogeneous fault plane in which some strong patches fail to break. These unruptured regions are termed *barriers* (Fig. 4.6). Aftershocks following the main earthquake will be concentrated around these barriers (Aki, 1984). If the fault plane is considered to be uniformly stressed prior to rupture, the presence of barriers after faulting represents a "stress roughening"

FIGURE 4.6. Model of asperities and barriers along a fault plane



Barriers (unruptured regions) can remain after faulting and may stop the propagating rupture. Asperities may exist before rupture and determine the strength of the fault. Modified after Aki (1984).

because the stress is less uniformly distributed after the earthquake. If the same strong barriers control successive earthquakes, such a fault might display recurrence patterns based on slip-predictable behavior (Fig. 4.5C).

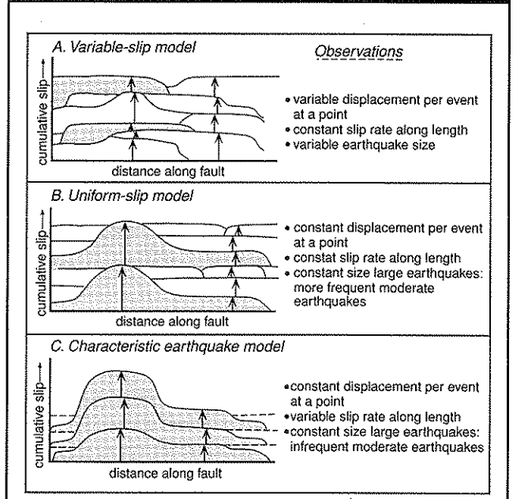
Characteristic Earthquakes and Fault Models

It has been proposed that some faults are typified by *characteristic earthquakes*, in which a fault or a segment of a fault ruptures repeatedly and displays approximately the same amount and distribution of slip during each successive event (Schwartz and Coppersmith, 1984). If a fault did indeed display characteristic earthquakes, knowledge of a single faulting event would provide a remarkable understanding of both previous and future rupture patterns, because the strain build-up and release, stress drops during faulting, variations in displacements along the fault, and the length of the rupture would be approximately duplicated in successive earthquakes. Thus, tremendous predictive power may reside in faults that rupture via characteristic earthquakes. If ruptures along a fault were controlled by stable asperities and barriers (i.e., consistent failure stress and terminating stress), that persisted from one earthquake to the next, this could provide a mechanism for generating similar "characteristic" slip along the fault in multiple events.

In order to test whether or not a fault is typified by characteristic earthquakes, either a suite of well-documented

historical earthquakes, or information from paleoseismic studies has to be used to reconstruct the temporal distribution of events and the spatial pattern of slip in any particular event in the past. The patterns of displacements of past ruptures need to be compared with respect to the length and position of the rupture and the distribution of displacement along the fault. Given a regional strain field, such as that controlled by relative plate motions, several different scenarios can be envisioned to accommodate the regional strain, only some of which would involve characteristic earthquakes (Schwartz and Coppersmith, 1984). In one scenario (Fig. 4.7A), displacements and rupture lengths are randomly distributed in such a fashion that, over time, all regional strain is accommodated, and there is an essentially uniform slip rate along the length of the fault. In this *variable-slip* scenario, the displacement experienced at any given point, the size of the earthquake, and the position of the rupture segment each vary unpredictably between successive events. Consequently, there are no faults exhibiting characteristic earthquakes in this scenario. A second scenario (Fig.4.7B) suggests that, at any given point, the displacement is consistent from one earthquake to the next, although displacement can vary along strike. Over time in this *uniform-slip* model, there is again a consistent slip rate all along the fault. In addition, the large earthquakes have a repetitive pattern in terms of rupture length and displacement variations along the rupture, and moderate sized earthquakes occur more frequently. In the *characteristic earthquake* model (Fig. 4.7C), there is also a consistent displacement at a point from one event to the next. Over time, however, the slip rate varies along the fault because the displacement varies along the length of the fault. Each large earthquake represents a repetition of the rupture location, length, and displacement pattern of the previous large earthquake, and there are infrequent moderate-sized earthquakes that only accommodate a fraction of the residual slip variation along the fault.

FIGURE 4.7. Models for slip accumulation (vertical axis) along a regional suite of faults

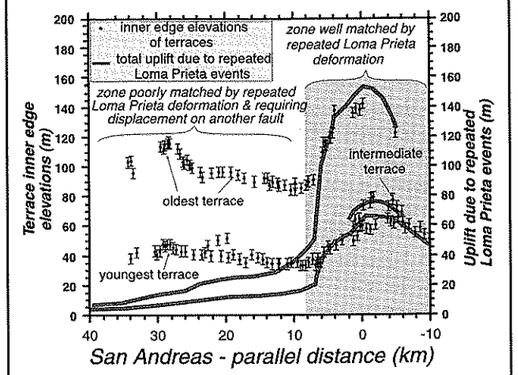


Models for variable slip (A), uniform slip (B), and characteristic earthquakes (C). Modified after Schwartz and Coppersmith (1984).

In viewing these models, it is easy to envision how variation in slip along normal or reverse faults that are associated with characteristic earthquakes would progressively build an irregular topography. If the site of greatest structural displacement along a fault were to be essentially fixed in the landscape through time, mountain peaks and basin depocenters would occur in predictable positions above the zone of maximum uplift or subsidence. Both structure contours and the landscape topography might be expected to be closely related to these repetitive cycles of displacement.

How could one test whether a fault displays characteristic earthquakes? In the vicinity of the "Parkfield asperity" on the San Andreas, for example, five moderate thrust-fault earthquakes (M~6) with rather similar rupture patterns have occurred during the past century (Bakun and McEvilly, 1984). These are suggestive of characteristic earthquake displacement patterns. Usually, however, the historical record is both too brief and incomplete to permit reconstruction of the displacement patterns during several ruptures of an entire fault or a segment of it. Consequently, we often have to interpret the geological record of deformation and faulting to assess past rupture patterns. At least two different approaches can be used for such an assessment. In one, the variability of displacement along a fault is compiled for several past earthquakes using measured displacements of strata, structures, and geomorphic features. We describe this approach in a subsequent chapter on paleoseismicity. In the second approach, the deformation observed to have resulted from a single, recent earthquake is compared with nearby deformed geologic markers, such as the surface of a fold

FIGURE 4.8. Loma Prieta shoreline displacement compared with the warped marine terraces



The variations in displacement and terrace altitudes are projected orthogonally on to the trend of the nearby San Andreas fault. Coseismic displacements in the southern half of the transect mimic the height variations of the warped marine terraces. This suggests that warping of these terraces could result from multiple earthquakes with Loma Prieta-like displacements. Modified after Anderson and Menking (1994).

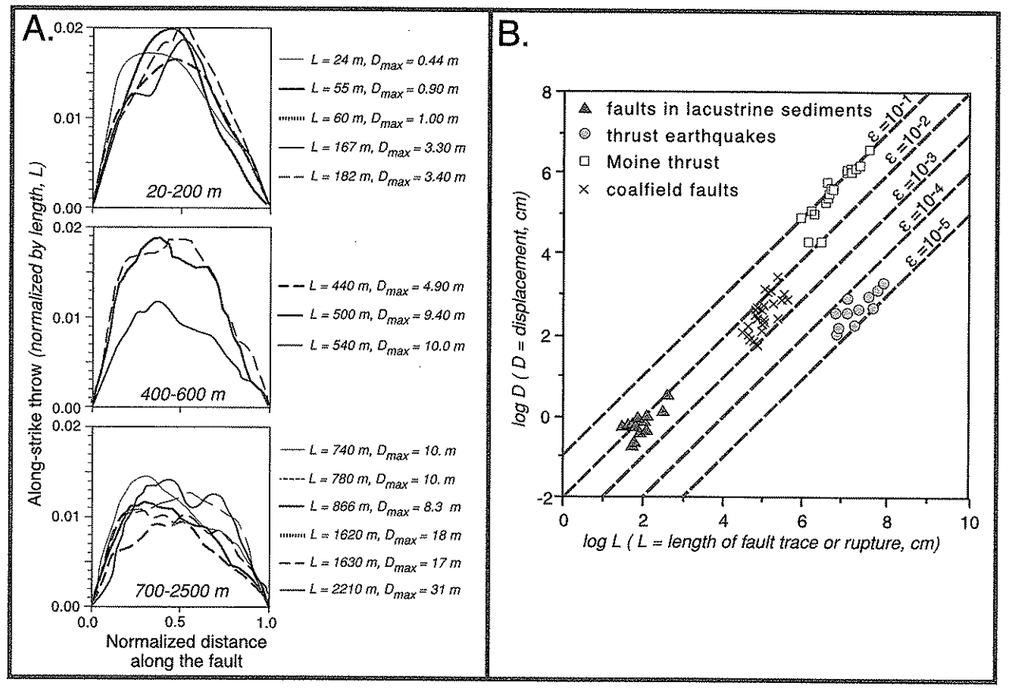
or deformed marine or fluvial terraces, which represent the cumulative deformation of multiple earthquakes. If the spatial pattern of deformation shown by the long-lived structure could be created through repeated application of the deformation due to the individual faulting event (Stein et al., 1988), then it is reasonable to suggest that the fault may have a characteristic behavior. An example (Fig. 4.8) of such a comparison can be made using a) the strain resulting from the 1989 M6.9 Loma Prieta earthquake south of San Francisco, and b) deformed marine terraces (Anderson and Menking, 1994). The Loma Prieta earthquake caused deformation extending from the San Andreas fault to the California coast, where shoreline displacement has been well documented by detailed surveying. A comparison of the shoreline coseismic displacement with the shape of the warped marine terraces clearly indicates that the terrace deformation could result from repeated Loma Prieta-like events. The match between the terraces and the coseismic strain is very good for the southern part of the terraces, but is unconvincing farther north (Fig. 4.8). To explain that northern pattern of deformation, repetitive faulting on another fault has been invoked (Anderson and Menking, 1994). While the excellent match of the warped terraces with the pattern of coseismic deformation does not prove that the Loma Prieta rupture was an oft-repeated, characteristic rupture, it certainly suggests that this is a reasonable interpretation.

Displacement Variations Along a Fault, Fault Growth, and Fault Segmentation

Fault Length and Displacement Variations

The concept of characteristic earthquakes contrasts with studies suggesting that, as a fault accumulates a greater total displacement, it lengthens, such that with each earthquake, the rupture is extended a bit more. But, does displacement vary systematically along a single rupture, or is it unpredictable? In order to collect a data set on fault growth and displacement, it is useful to study faults that developed in a homogeneous material and are sufficiently young to ensure that the faults have not been significantly modified by erosion. These conditions are met by the Bishop Tuff, an extensive ashflow that is about 800 ka (Bailey et al., 1976). Studies of faults offsetting this tuff in central California display similar displacement patterns for normal fault lengths spanning about two orders of magnitudes (20 m to 2000 m; Fig. 4.9A) (Dawers et al., 1993). In particular, there is a bow-shaped displacement variation, such that there is no displacement at the fault tips, the amount of displacement rapidly increases over the first third of the fault, and then it rises more gradually to a maximum near the midpoint of the fault. On all of these particular faults, the maximum displacement (D) is about 1–2% of the fault length (L). This displacement pattern is similar to theoretical predictions, laboratory experiments, and other field studies (Pollard and Segall, 1987; Schlische et al., 1996). A compilation of displacement data (Fig. 4.9B) from a variety of settings and different types of faults (Scholz, 1990), clearly suggests that there exist predictable scaling relationships between the length of the fault and the displacement along it, but that these relationships vary for different settings and perhaps different types of faults. For the Moine thrust and modern thrust earthquakes (Fig. 4.9B), there appears to

FIGURE 4.9. Length scaling relationships



A. Displacement versus distance along fault traces on normal faults cutting the Bishop Tuff, California. Note the characteristic bow-shaped displacement variation for faults spanning two orders of magnitude in length. For all faults, the maximum displacement is between 1 and 2% of the fault length. Modified from Dawers et al. (1993).
B. Systematic fault scaling relationships are apparent when maximum displacements versus fault lengths are plotted. No universal scaling rule describes these data because they represent several different tectonic settings, fault types, and rock types. Within each subset, however, displacement versus length variations could be readily quantified. The dashed lines are lines of equal strain (D/L). Modified from Scholz (1990).

be a linear relationship, such that $D \alpha \varepsilon \bullet L$, where ε is the strain, which varies over four orders of magnitude from about 10^{-1} to 10^{-5}. In contrast, the coalfield faults (Watterson, 1986) suggest (Fig. 4.9B) that displacement is proportional to fault length squared, $D \alpha \varepsilon \bullet L^2$. At present, clearly there is no single model that universally explains displacement/length scaling, but the available data suggest that, in a particular setting with fairly homogeneous rock properties, predictable D/L relationships of the general form $D \alpha \varepsilon \bullet L^n$ can be defined.

If it is assumed that the measured displacements are the result of repeated seismic ruptures, then these data clearly argue that faults indeed lengthen as they accumulate displacement. Thus, over long time intervals and many seismic events, these data suggest that no fault can be consistently typified by characteristic earthquakes: the fault has to lengthen over time. On the other hand, once a rupture plane is fairly large (several km in length), the incremental change in length due to several successive earthquakes may be insignificant compared to its total length. Thus, at the time scale of several events, the

Seismic Moment and Moment Magnitudes

Any time an earthquake occurs, it releases seismic energy, E_s. The energy released is proportional to the area of the rupture plane, A, the average displacement along it, d, and the stress drop, $\Delta\sigma$, across the fault during the earthquake:

$$E_s = 1/2 \cdot \Delta\sigma \cdot d \cdot A \qquad (6.1)$$

The stress drop actually refers to the mean stress, σ_{mean}, acting across the fault during the earthquake and is calculated as the average of the stresses before and after the earthquake:

$$\sigma_{mean} = (\sigma_{start} - \sigma_{finish})/2 \qquad (6.2)$$

If it is assumed (as is often done) that the fault is stress free at the end of rupture ($\sigma_{finish} = 0$), then the mean stress is equal to one-half of the stress drop, $\Delta\sigma$.

Although it is now both possible and very useful to compare earthquakes on the basis of the energy released during them, traditionally the size of an earthquake has been assessed on the basis of its magnitude. The first quantitative measure of size was the local magnitude (M_L) scale which was based on Richter's (1935) observation that, with increasing distance from seismic sources in southern California, the maximum amplitude of ground motion decayed along a predictable curve. When the data for distance versus the logarithm of the amplitude of ground shaking were compared for several earthquakes, they followed parallel curves of decay with increasing distance. By measuring the amplitude of shaking, A, in a given earthquake and comparing it with the amplitude, A_o, of a "reference event," it was possible to define a local magnitude (M_L):

$$M_L = log\ A - log\ A_o \qquad (6.3)$$

For a reference earthquake with $M_L = 0$, the amplitude of shaking at a distance, Δ, of 100 km from the source was 1 mm. A calibration of these curves using the reference event led to the following expression for local magnitude:

$$M_L = log\ A - 2.48 + 2.76\ log\ \Delta \qquad (6.4)$$

The original Richter magnitude scale was developed based on measurements of southern Californian earthquakes during the 1930s using a certain type of seismometer (Wood-Anderson torsion instruments), and the equation above is only strictly applicable to these seismometers in this setting. Today, however, local magnitudes are calculated in many areas using a variety of seismometers. The coefficients in Equation 6.4 have been modified to yield consistent estimates of the local magnitude, despite these differences in instrumentation and regional contrasts in the transmission of seismic waves due to geological variability.

Several other magnitude estimates are commonly used. Surface-wave magnitudes, M_s, are typically calculated for events at distances exceeding 600 km by measuring amplitudes of surface waves with a period of approximately 20 seconds. Body-wave magnitudes, m_b, are based on the amplitude of direct compressional, P, waves measured from short-period seismograms, typically with periods of about 1 second. Empirical relationships have been established between these earthquake magnitudes and seismic energy, E_s, released:

$$log\ E_s = 11.8 + 1.5\ M_s \qquad (6.5)$$

$$\text{and } log\ E_s = 5.8 + 2.4\ m_b \qquad (6.6)$$

A different measure of the energy released in an earthquake is the seismic moment (M_o, measured in dyne-cm) which is equivalent to the product of the rupture area, a, average displacement, d, and the rigidity, μ or shear modulus of elasticity of the crustal material involved in the rupture:

$$M_o = \mu \cdot d \cdot a \qquad (6.7)$$

where μ is commonly taken as 3×10^{11} dyne/cm^2 for the crust and 7×10^{11} dyne/cm^2 for the upper mantle. Thus, instead of measuring the amplitude of a deflection on a seismograph, observations of the rupture length, the probable total area of rupture, and the mean displacement are combined with estimates of rigidity to yield the seismic moment. Moreover, the magnitudes of many earthquakes today are reported as "moment magnitudes," M_w, and are based on the seismic energy released during an earthquake, that is, the seismic moment, M_o:

$$M_w = (log\ M_o / 1.5) - 10.73 \qquad (6.8)$$

In theory, the seismic moment is a single, definable quantity that reliably characterizes the seismic energy release. As a consequence, moment magnitudes should theoretically provide a direct means for comparing different earthquakes. In reality, today seismic moments are calculated both from the actual characteristics of the rupture and from empirically derived functions that relate body or surface waveforms to seismic energy and moment.

It is impossible to reconstruct quantitatively the amplitude of shaking in pre-historic earthquakes. Consequently, it is difficult to estimate directly the magnitudes of ancient earthquakes based on paleoseismologic studies. What can be measured, however, are rupture lengths, mean displacements, and approximate rupture areas. These observations provide a means of determining the absolute size of past earthquakes for which there are no instrumental records. Such a quantification through paleoseismological studies represents a powerful basis for comparing ancient and modern earthquakes. Modern studies have defined relationships between the seismic moment and the amount of ground displacement and shaking. Building on these relationships, paleoseismological determinations of rupture lengths, fault geometry, and coseismic fault displacement provide key constraints on the assessment of modern seismic hazards along faults.

Direct Observations of Paleoseismic Displacements

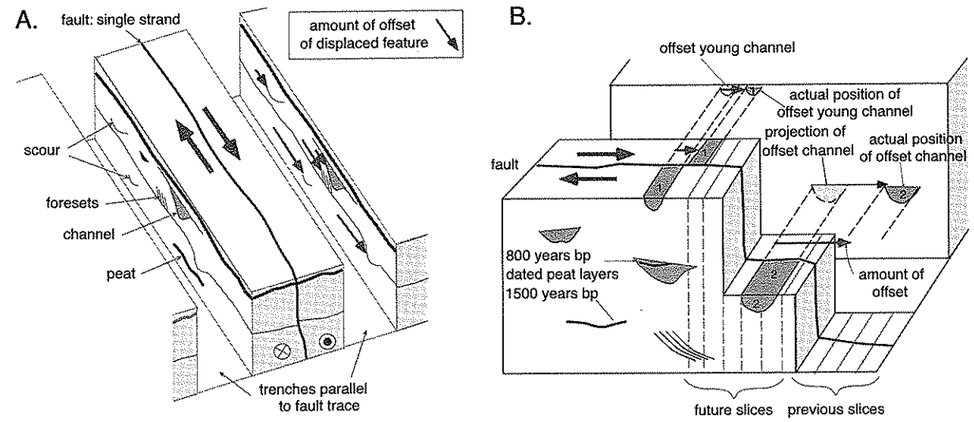
Two types of data can be brought to bear on the reconstruction of the history of past earthquakes. Direct observations of displaced or cross-cutting features provide unambiguous information about displacements. Such information can be stratigraphic, structural, or geomorphic in nature, and includes features such as faulted beds, offset stream channels, and raised beaches. Indirect indicators of earthquakes require a conceptual linkage between the observable data and the earthquake that caused it. In some instances the evidence reflects the coseismic offset itself, while in others it reflects other parts of the seismic cycle. For example, stratigraphic evidence of tsunamis or the chronology of rockfall deposits require an interpretation of their genesis in order to tie them into specific faulting events. We discuss the types of information that can be generated with direct observations first, and later examine several kinds of indirect paleoseismological observations.

Trenching

In order to provide a detailed record of past displacements and their timing, excavated trenches across faults often provide key insights. Overall, the basic practical objectives in trenches are 1) to identify and date layers within a stratigraphic succession that either have been disrupted by faulting or overlie fault traces without disruption, and 2) to document the amount of displacement in past faulting events. Sites for trenches have to be carefully chosen in order to maximize the useful information that they may generate. To the extent possible, trench sites should contain abundant datable material, and they should provide stratigraphic or structural markers that can be used to measure offsets. Typically, it is impossible to know what is likely to be found in the subsurface as a trench is excavated, so good judgment and good luck combine to create a data-rich excavation. Because radiocarbon dating is still the most frequently applied dating technique in trench analyses, it is not uncommon for trenches to be located in swampy areas where fault displacements have dammed up local streams or have ponded the groundwater table. In such circumstances, organic matter that can be dated using radiocarbon methods is more likely to be preserved within the young strata associated with the fault. Similarly, thinly bedded deposits are more likely to reveal discrete, measurable offsets than will massive deposits, such as debris flows, that are rather homogeneous in all directions. Thus, the likelihood of having a rich stratigraphic record can be enhanced by choosing sites of low-energy deposition and/or sites where it is expected that there will be linear stratigraphic markers preserved with which to measure offsets. For example, if relict lake shorelines or small-scale channels are oriented approximately perpendicular to a fault, they can offer particularly good stratigraphic markers or piercing points across a fault.

In trenches, the analysis of the stratigraphic record of earthquakes is typically a time-consuming process. Thus, the choice of sites for trenches is not a trivial exercise: you want the maximum information for the time invested. Trade-offs have to be made between competing objectives: obtaining more detailed information in a single segment of a fault, comparing rupture histories in different areas, obtaining maximum information on the magnitude of offsets, and developing the most extensive chronology of past events. Two different strategies are typically used for orienting trenches with respect to faults. In most situations, a single trench is excavated perpendicular to the trend of the fault. The stratigraphy and structures

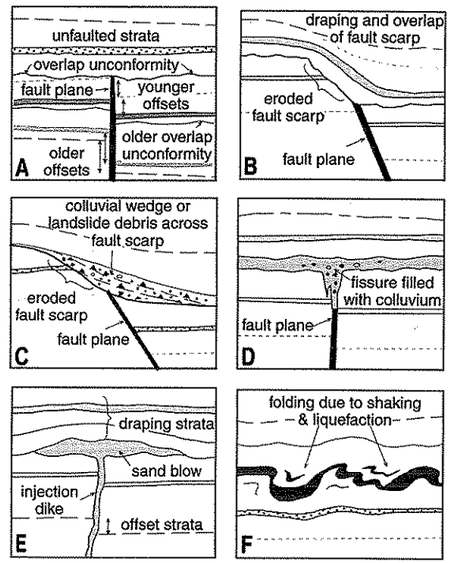
FIGURE 6.1. Trenches on a strike-slip fault



A. Orientation of trenches with respect to a strike-slip fault, when two trenches are employed to determine the history of previous ruptures. On the walls closest to the fault trace, the stratigraphy is carefully surveyed and material, such as peats, for dating is collected. Hypothetical amounts of offsets of stratigraphic markers are depicted by the arrows on the trench wall. Note the general downward increase in displacement. Despite their difference in stratigraphic height, the top two offset markers, a channel and a scour, have identical offsets because they were both displaced in the most recent earthquake. B. "Salami slicing" across the fault trace between two surveyed trenches. Offsets of linear features like channels are clearly revealed, and displacements of the margins of these features are measured in order to determine the amount of offset in individual earthquakes. Similar magnitudes of offset are expected for all of the features that were created between each successive pair of earthquakes. When the stratigraphy in the trench walls has been well dated, both the age and the displacement of past earthquakes can be quantified.

revealed in the walls of the trench are meticulously surveyed and mapped, material for dating of various stratigraphic horizons is collected, and a fault displacement history is interpreted based on these data. Alternatively, two trenches parallel to the fault trace are sometimes employed along strike-slip faults (Fig. 6.1A). The stratigraphy on the faces nearest the fault trace is mapped, surveyed, and dated in each fault. Special attention is paid to linear features, such as channels, planar crossbeds, shoreline features, or unusual bedding configurations, that trend approximately perpendicular to the fault. After mapping the two trench walls parallel to the length of the fault, the intervening strata that are cut by the fault are "salami sliced" perpendicular to the fault trace (Fig. 6.1), which is to say that they are incrementally cut back along vertical planes. Each stratigraphic feature that could act as a piercing point is traced to the fault and the magnitude of offset vis-a-vis the correlative feature on the opposite side of the fault is measured. Along a fault trace where multiple earthquakes are recorded and where the sense of offset has not changed between successive ruptures, the amount of the measured offset should increase with increasing stratigraphic age, that is, with depth in the trench. This increase, however, should not be steady; rather, it should be stepwise, such that between two stratigraphic horizons that correlate with different faulting events, all of the displaced features should display similar offsets (Fig. 6.2A). Each cluster of comparably offset features will be bounded by abrupt changes to smaller offsets above and larger offsets below, marking the accumulation of displacement in successive earthquakes. Ideally, the horizons that correlate with the abrupt changes will be well dated and hence will reveal directly the timing of previous earthquakes (Sieh, 1978).

FIGURE 6.2. Examples of stratigraphic and structural relationships that relate to past deformational events



A. Strata offset across a fault are beveled by an unconformity and overlain by undisturbed strata. The unconformity developed after the last earthquake. The amount of displacement increases downward. The two different amounts of offset of strata and the presence of an older offset unconformity provide evidence for a previous earthquake. B. If the topography across a faulted surface is incompletely beveled off, succeeding strata will drape across the eroded fault scarp. They will be unbroken, but may appear deformed due to the topography on which they were deposited. Often they will show thickening above the downthrown fault block. C. Similar to B, but with a colluvial wedge derived from erosion of the upthrown block. D. Fissures that open along a fault plane are typically filled with colluvial material shortly after faulting. E. Injection dikes in the subsurface and sand blows or sand volcanoes on the surface provide evidence for past earthquakes. The age of the youngest strata cut by the dikes or underlying the sand blow provides a maximum age on the faulting. F. Liquefaction due to shaking can cause folding of weakly consolidated sediments near or at the surface. The age of the deformed beds provides a lower limit on the time of the earthquake that deformed them. Modified after Allen (1986).

The "salami slicing" approach can reveal detailed data on both the timing and magnitude of offsets in previous earthquakes along strike-slip faults. Even when the coseismic displacements are more than several meters in each rupture, a long trench can record multiple offsets. In fact, cumulative offsets of the oldest markers can be equal to the length of the trench. Many dip-slip faults that are of interest to paleoseismologists have displacements of 2 to 4 meters in individual earthquakes. In order to find the record of multiple ruptures in the past, trenches on dip-slip faults have to be deep, rather than long. A trench 10 meters deep might record only two ruptures, if the displacements were 4 meters in each event. The instability of trench walls (which must be shored up and braced) often imposes a practical limit on how deeply trenches can be dug, and therefore limits the number of previous rupture events that can be examined. As a result, the paleoseismic records in trenches of large thrust or normal faults are commonly limited to one or two events (Rubin et al., 1998).

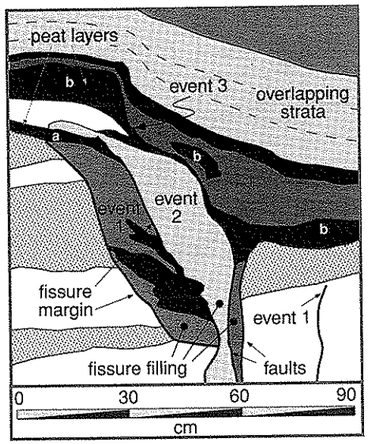
Offset strata above dip-slip faults can present difficulties that are not present with strike-slip faults. Stratigraphy that is displaced by strike-slip faults is translated horizontally into new positions where the strata on both sides of the fault are often preserved. Thus, more sediments can accumulate above the displaced strata and piercing points recorded by the older strata will be preserved. On the other hand, when dip-slip faulting occurs, the strata on one side of a fault are upthrown with respect to the opposite side. In theory, it is the offset of the strata on opposite sides of the fault that will record the magnitude of displacement. Unfortunately for paleoseismologists, strata on the upthrown block are often subjected to subaerial erosion. Consequently, the specific stratal layers that could define the offset precisely can be eroded away (Machette et al., 1992), such that reconstruction of past displacements becomes considerably more difficult.

In order to reconstruct the history of offset that is recorded by strata within a trench, stratigraphic and structural relationships revealed in the walls have to be interpreted. As with traditional structural geologic studies, cross-cutting relationships can provide unambiguous evidence of past deformation (Fig. 6.2). For example, if older strata are cut by a fault, but overlying younger strata are continuous across the trace of the fault, an earthquake is interpreted to have occurred between deposition of the youngest strata that are cut by the fault and the oldest strata that are not displaced by it. If these strata can be dated, they provide bracketing ages on the faulting event.

One goal of paleoseismic studies is to determine the amount of coseismic displacement from the apparent offset of strata exposed in a trench wall. To do this, the direction of slip along the fault must be known. Consequently, indicators of the slip direction, such as striae, slickensides on the fault plane, fold axes of deformed strata, or the displacement of piercing points, should be sought. Progressively larger structural offsets of stratigraphic horizons farther beneath the unconformity or the identification of older offset unconformities can indicate both the timing and magnitude of earlier earthquakes (Fig. 6.2A). Topographic scarps formed during faulting events are rapidly "attacked" by erosional forces. If the scarp forms in an area where active deposition is occurring all around it, the scarp is likely to be draped with younger strata soon after the rupture (Fig. 6.2B). If the scarp forms in a predominantly erosional environment, its upthrown side will be steadily eroded and will often provide colluvial debris that accumulates along the degrading scarp as a *colluvial wedge* on the locally downthrown block (Fig. 6.2C). During strike-slip earthquakes, there may be relatively little differential vertical motion, but movement along the irregularly shaped fault plane will cause fissures to open in places (Fig. 6.2D). These fissures are open to the surface and typically fill quite rapidly with colluvial, alluvial, or aeolian debris. In repeated ruptures, fissures may re-open with each rupture (Sieh et al., 1989). Consequently, a single fissure with multiple filling events may record several different earthquakes (Fig. 6.3). Seismic shaking can cause liquefaction of sediments having a high water content (Fig. 6.2F). Liquefied sand represents a slurry of sand and water which may be "erupted" on the surface to form sand blows or sand volcanoes (Fig. 6.2E).

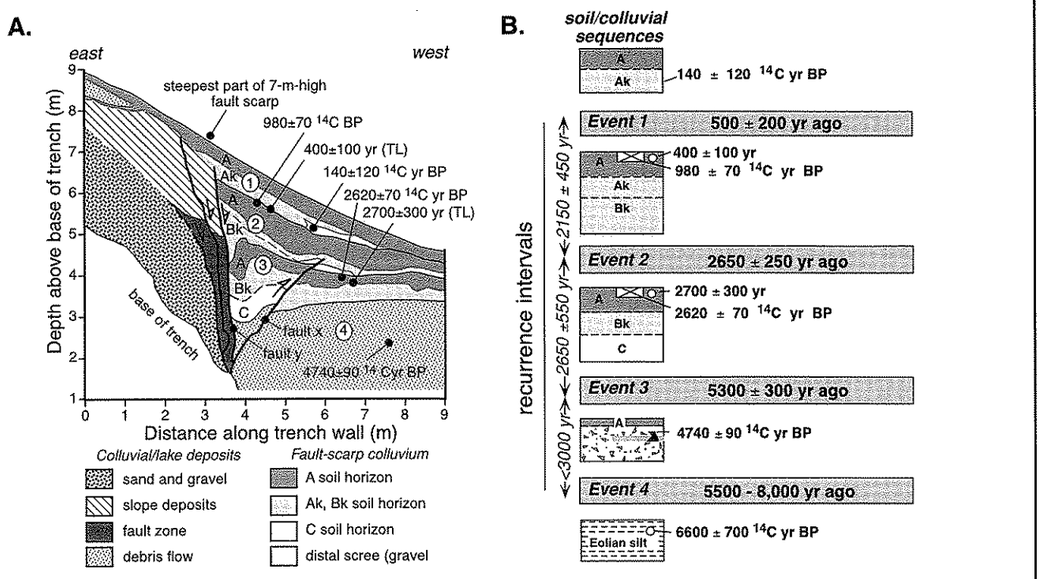
The interpretation of the stratigraphy and structures found in trenching sites is often not straightforward. In some fortunate circumstances, clear cross-cutting relationships and abundant material for dating are present (Fig. 6.4). At such sites, ruptures of older strata and draping or erosion across the upper termination of fault strands provide direct evidence for faulting. Even in these conditions, it is important to remember that faults don't always propagate to the surface and that new ruptures can form in successive earthquakes. Thus, a newly formed fault splay may branch out from the main rupture and terminate at some random level in the subsurface in an isolated outcrop or trench. In this case, the undisturbed layers overlying the fault termination will pre-date the faulting event, which is the opposite of the case when undisturbed layers depositionally overlie a fault termination. Consequently, in examining fault terminations, it is important to seek evidence for surface erosion across the fault trace, for bed thickness changes above the fault termination that indicate the filling of small-scale topography related to the fault, for discrete groupings of the amounts of stratal offsets found in older layers cut by the fault (indicating that deposition was synchronous with faulting and that previous earthquakes had occurred on the fault), and for significant displacement on the fault just below the undisturbed layers, rather than a progressive dying out of displacement along the fault. If it can be shown that the same stratigraphic level in several places overlies fault traces, it is likely that this represents a fault that did propagate to the surface.

FIGURE 6.3. Fissure filling showing evidence for three earthquakes



Event 1 caused a fissure to open. The fissure-filling post-dates the event, contains some datable peat, and is overlain by peat "a." Event 2 re-opens the fissure and cuts peat "a." Event 2 is overlain by peat "b," which is cut along a small fault along the fissure trend during Event 3. Despite its complexity, this is a good example of the detail that can be extracted from a good exposure and provides some insight into how overlapping, cross-cutting, and fracture-filling relationships can be exploited. Modified after Sieh et al. (1989).

FIGURE 6.4. Record of normal faulting along the Wasatch fault bounding the Wasatch Range near American Fork Canyon, Utah



A. Evidence for three faulting events is revealed in this trench. The excavation is > 9 m deep, and yet, because markers typically show 1–3 m of offset per earthquake, only three rupture events are recorded here. A repetitive soil zonation (A – Ak ± Bk) delineates the top of each of three colluvial wedges (labeled 1 through 3). The magnitude of offset for each wedge can be estimated by restoring the top surface of the wedge back to the ground surface such that the "A" soil horizon is continuous onto the upthrown block. Radiocarbon and thermoluminescence dates are used to constrain the timing of faulting events. Cross-cutting relationships with faults and soils define rupture events. Fault "x" cuts soil sequence 3 and is overlain by sequence 2. Fault "y" cuts soil sequence 2 and 3, but is overlain by sequence 1. Note that with the exception of the lowest and topmost stratigraphic units on the downthrown side, none of the downthrown strata are represented on the upthrown block. B. Interpretation of the rupture history based on the trench stratigraphy. Evidence for the oldest faulting event is based on a nearby trench. Note that on the trenched segment of the Wasatch fault, the ruptures appear to be evenly spaced in time, but the magnitude of rupture varies, with the latest offset (~1 m) being about half as large as the offsets during the previous two events. Modified after Machette et al. (1992).

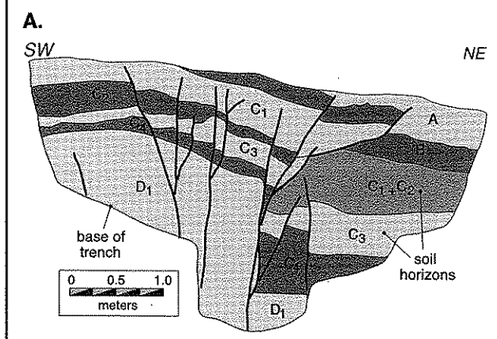
The stratigraphy of soils can provide useful controls on the interpretation of trench stratigraphy. For example, colluvial wedges (Figs. 6.2C and 6.4) are often associated with dip-slip faults. Each colluvial wedge is commonly interpreted to result from a single rupture of the bounding fault (Machette et al., 1992; Schwartz and Coppersmith, 1984). Steepening of slopes, localized uplift, or overthrusting of the land surface along a rupture plane destabilizes weakly consolidated surface strata. Quaternary sediments that have been uplifted to a higher topographic position along these steepened slopes quickly erode and produce colluvial wedges which accumulate along the toe of the exposed scarp. If the faulted strata are well consolidated, they may maintain steep faces for many centuries. Nonetheless, it is commonly assumed that raveling of a steepened slope occurs rather quickly (about 100 years in many sediments), such that not long after faulting, the surface of a

colluvial wedge is at or below the angle of repose for unconsolidated debris (~30°) and the deposition rate is low enough for pedogenic processes to begin to develop soil zonations (A, B, K horizons). Recognition of such soil horizons in trench walls pinpoints the upper surface of colluvial wedges formed during earlier earthquakes and delineates some of the vertical extent of the wedges (Fig. 6.4A). Particularly in colluvium that is poorly stratified and contains few stratigraphic horizons with which to judge offsets, even subtle soil horizons provide markers for delineating displacements. In addition, the degree of soil development can be related to the interval between earthquakes, that is, the time from first stabilization of the colluvial wedge to its burial by the next youngest wedge (Fig. 6.4B).

Many fault zones actually have more than one fault strand at the surface. In order to account for all of the displacement along a fault zone, it is necessary to identify all of the strands or surface traces, and a paleoseismic record for each should be developed. It is not uncommon for strike-slip faults to splay into several branches as they approach the surface. Sometimes these splays all occur within a fairly compact zone for which a single trench or small set of trenches will reveal the entire displacement record (Fig. 6.5A). However, when strands are separated by tens or hundreds of meters, each has to be examined separately. In compact zones with multiple splays (Fig. 6.5B), it is often very informative to locate a distinctive marker bed and map its displacement across each of the strands (Lindvall et al., 1989).

Dating of fault-disrupted strata in trenches forms the basis for determining the timing of individual ruptures, which in turn is the basis for assessment of recurrence intervals. It is always wise to generate as many reliable dates as possible for a given stratigraphy, particularly if it will underpin key interpretations of the seismic record. Radiocarbon dating of buried organics is still the most commonly used dating technique in trenches. Given analytical uncertainties in laboratory analyses, the possibility of contamination with young carbon, and the background variations in atmospheric carbon in the past (see Chapter 3), each radiocarbon date should be carefully interpreted. Combining different dating techniques, such as radiocarbon with thermoluminescence, can provide a check on the consistency of dates. In a trench along the normal fault along the Wasatch front (Fig. 6.4), for example, some radiocarbon dates vary by 400 to 700 years from their calendar ages. These temporal offsets can be assigned from the calibration curve for radiocarbon ages (Stuiver and Reimer, 1993), but it is also reassuring to have another estimate of the age of the strata from a different dating approach. In the Wasatch trench, thermoluminescence dates (Fig. 6.4A) reinforce the interpretation of the radiocarbon dates.

FIGURE 6.5. Trenching a strike-slip fault with multiple splays

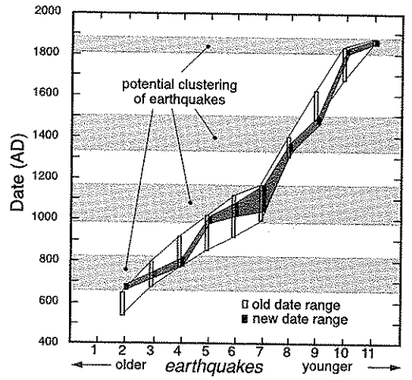


A. Sketch of a trench wall oriented approximately perpendicular to a suite of splays along a strike-slip fault. This outward branching geometry or "flower" structure displays 10 different splays across a zone ~3 m wide. Not all splays have been active in each earthquake. Note, for example, two faults on the right cut the upper B_t soil horizon, whereas the other splays do not. For orientation of the trench, see Fig. 6.5b. Modified after Lindvall et al. (1989). B. Displacement of a marker horizon (a fluvial channel) across multiple splays of a strike-slip fault. Map shows a network of shallow trenches that reveals the offset of the marker channel. Across any strand, the offset is typically 1–3 m, but the cumulative offset is > 8 m. Modified after Lindvall et al. (1989).

In most trenches, the strata that contain datable material do not consistently coincide with the strata that record rupture events. In such cases, dates on strata above and below the "event horizon" are used to bracket its age. In order to reduce the uncertainty in timing as much as possible, datable material should be sought as close as possible to the rupture horizon, and the highest laboratory precision available should be used to reduce the analytical uncertainty on the individual ages (Atwater et al., 1991; Sieh et al, 1989). Concerted efforts to provide the best possible time control are warranted, because reliable calculations of recurrence intervals are highly dependent on the quality of the ages assigned to rupture events.

Recent studies along the San Andreas fault clearly illustrate the importance and utility of high-precision dating. Because the San Andreas fault runs near or through several of the major population centers in California and because it has generated "great earthquakes" (magnitude 8) in the past, there is great interest in knowing 1) how often earthquakes occur along it; 2) the length, displacements, and spatial patterns of past ruptures; and 3) the recurrence interval between major earthquakes. During the past 30 years, much effort has gone into paleoseismological studies along many parts of this major plate-bounding fault. Some of the earliest insights on recurrence intervals of San Andreas faulting came from trenches dug at Pallett Creek about 55 km from Los Angeles (Sieh, 1978). Several subsequent studies have attempted to improve on the dating of the faulting events at this site, where 11 major earthquakes are recorded. A comparison of the faulting chronologies published in 1984 (Sieh, 1984) and in 1989 (Sieh et al., 1989) is illuminating (Fig. 6.6). Based on the 1984 data (Sieh, 1984), the recurrence intervals between major earthquakes fall into two groups: prior to about 1100 A.D., the recurrence interval was about 100 years, whereas from 1100 A.D. to the present, the interval lengthens to 160–200 years. Within each grouping, the resolution of the radiocarbon dates suggests that the earthquakes may have been approximately evenly spaced in time. When radiocarbon dates with higher precision are used to date the earthquakes (black bars, Fig. 6.6), a rather different rupture history emerges (Sieh et al., 1989). Rather than being evenly spaced in time, the earthquakes appear to cluster. Two or three earthquakes appear to have occurred during spans of 100–200 years, whereas each cluster is separated from the succeeding cluster by an earthquake-free interval of 150–300 years. The uneven temporal spacing of San Andreas earthquakes presents a challenge to seismologists trying to understand rupture physics and recurrence intervals. Moreover, the clustering of events suggests that seismic risks along this segment of the San Andreas are very different than previously thought. Based on the 1984 data, it was thought the earthquakes were regularly spaced in time and had a recurrence interval of about 140 years. Given that the last large rupture in this area was in 1857, this scenario would suggest that the next big rupture could be imminent. In contrast, the 1989 data suggests that, because clustered earthquakes commonly occur within a few decades of each other and because no major rupture has occurred in the past 140 years, we

FIGURE 6.6. Recurrence intervals as determined by radiocarbon dating for major earthquakes along the San Andreas fault at Pallett Creek, California



The dates published in 1984 (open boxes indicate the time range of each date) suggest two groupings of earthquakes. In each group, the earthquakes are evenly spaced in time, but prior to ~1100 A.D., they had a recurrence interval of ~100 years, whereas from 1200 A.D. to present, the recurrence interval was 150–200 years. Higher precision dates published in 1989 indicate a distinct clustering of earthquakes (shaded bands) into groups of two or three ruptures each that were separated from each other by a few decades, whereas the clusters themselves are separated from each other by 100–300 years. The higher precision dates define a step-like event-versus-time curve (dark shading), as opposed to the smoothly changing curve (no shading) comprising the less precise dates. Modified after Sieh et al. (1989).

are presently in one of the quiescent periods between clusters and that this quiet period could last for another 100 years!

Displaced Geomorphologic Features

Although detailed stratigraphic records in trenches can provide constraints on multiple earthquakes in the past, each trench requires a large investment of time and is most relevant only to those segments of the fault directly adjacent to it. Offset geomorphic features, on the other hand, can be readily observed and surveyed, and they are often distributed along the length of a fault. Thus, such features provide a means to document the spatial variations in seismic deformation along past ruptures. In alluvial environments, typical displaced geomorphic features include stream channels, terrace risers, channel walls, debris flows and their raised levees, small alluvial fans, ridges, and gullies. In coastal environments, displaced features could include beach ridges, coral platforms, delta plains, and wave-cut notches.

Terrestrial Settings

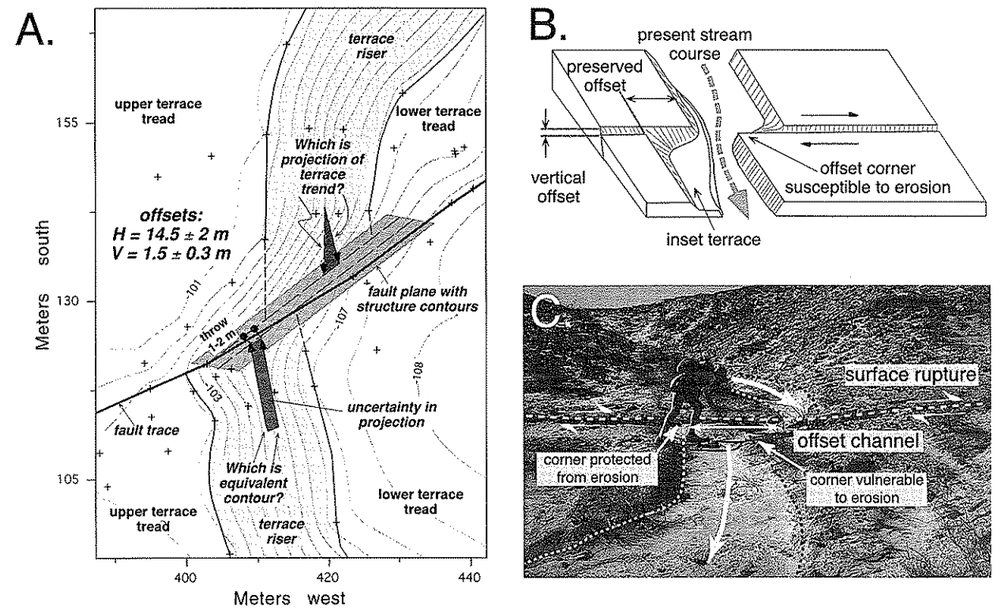
Although it is often possible to obtain a rough estimate of the amount of displacement of a displaced feature with a tape measure, it is almost always preferable to create detailed topographic maps of displaced features and the area surrounding them because these permit a more rigorous geometric reconstruction (Fig. 6.7A). Typically, such maps are constructed using a theodolite with a built-in electronic distance measuring device (an EDM or "total station"). In many field situations, displaced features no longer directly intersect the fault plane because slope processes have either eroded or buried part of them. In such cases, the trend of a feature, as represented by planar surfaces, such as terrace risers, or by linear features, such as debris-flow levees or the intersection between a terrace tread and riser, has to be projected onto the fault plane (Fig. 6.7A). To the extent possible, a structure contour map of the fault plane should also be constructed, especially with dipping faults. In order to measure a horizontal offset, once the fault plane is specified, the distance is measured between the projections of linear features on to the fault plane after any vertical component of offset is subtracted. Clearly, the vertical displacement must also be measured, and this can be done using a similar methodology, but is based on projections of subhorizontal features, like terrace treads or channel bottoms, on to the fault plane. With a detailed topographic map, often the uncertainties in such a projection can be estimated and incorporated into the displacement estimate (Fig. 6.7A).

In order to obtain reliable estimates of past offsets, care must be taken either to choose features that have been little modified since they were displaced, or to reconstruct accurately their post-seismic evolution. For example, when the channel walls along a stream are offset along a strike-slip fault, one side of the displaced channel in the downstream direction will have been moved more directly into the path of the stream and is likely to be strongly modified by erosion, whereas the other side of the displaced channel will be sheltered from erosion (Fig. 6.7B). These protected features provide the most reliable estimates of displacements.

In theory, every large earthquake on a fault could be represented by a suite of displaced features. In order for this to be true, some new geomorphic markers would have to form in the interseismic interval between faulting events, and older markers that were displaced previously would have to be preserved until the time of measurement. Such conditions are not uncommon, because each seismic displacement commonly forces some aspects of the geomorphic system to adjust, thereby producing new markers. In natural landscapes, therefore, one might expect to find numerous markers that had been displaced in the most recent earthquake, a lesser number displaced during the penultimate rupture, and so forth. Such landscapes could be described as *palimpsest* landscapes in which the imprint of older features is only partially overprinted or obscured by younger features. In addition, the size of the offset may not be large enough relative to the scale of the geomorphic feature to have altered it significantly, or to have caused its abandonment. This ratió of scales is therefore important in determining if a particular landscape will record individual coseismic offsets, or if instead it will record only the long-term average slip on the fault, smearing out the effects of individual events.

Clearly, the record of offset features represents an interaction between the processes that create geomorphic markers and the tectonic events that displace them. The variability of climate suggests that some intervals will be more conducive than others to the formation of clear geomorphic markers. For example, during times of fluvial incision, many steep-sided geomorphic features, such as channel walls, terrace risers, and gullies, are etched into the land surface and gently dipping features like terrace treads are left behind, whereas during aggradational phases, broad alluviated surfaces provide fewer distinctive features that could serve as geomorphic markers. Moreover, aggradation can cause burial of offset features, and thus obscure the record of older earthquakes. As a consequence of these climatically controlled contrasts, earthquakes that occurred after intervals of incision are likely to be more clearly represented in the geomorphic record than are those that occurred during times of aggradation. Paleoseismologists exploit these contrasting regimes of incision and deposition. The offset geomorphic features of incisional areas allow ready identification of faults, whereas places where faults cut areas of active deposition are ones where the fault trace may be obscure, but the past record of earthquakes can be discerned through trenching.

FIGURE 6.7. Use of offset terraces and channels to define fault displacements



A. Gridded and contoured topographic map of late Holocene terrace riser offset by Awatere Fault at Grey River, Awatere Valley, New Zealand. Map derived from total-station measurements. Dashed lines subparallel to the fault are two structure contours on the fault plane which is dipping to the SSE (contour interval 1.5 m). Top structure contour on the fault plane is at the same elevation as mid-riser contour on the south side of fault that is used to project the position of the terrace riser on to the fault plane. Lower structure contour on fault is 1.5 m lower down (mean throw on both the lower and upper terrace surfaces). Shaded uncertainty region on S side fault is uncertainty in projection of riser contour to eroded fault surface on upthrown (SE) side of fault. Shaded uncertainty belt on north side of fault represents uncertainty in knowledge of which contour on N side of fault originally corresponds to the south-side riser contour prior to displacement. This is the same as the uncertainty in throw for the lower terrace surface (± 0.3 m). Modified after Little et al. (1998). B. Sketch of the geometry of the displaced channel wall and terrace and the optimal position in which to measure offset along the fault. The offset "corner" to the right of the stream is likely to be modified by erosion as the stream impinges on it. Therefore, it is far better to measure the displacement based on the protected channel wall on the opposite side of the channel. Modified after McGill and Sieh (1991). C. Ephemeral channel offset by dextral slip during the Superstition Hills earthquake. Note that the downstream channel margin on the right is vulnerable to erosion, such that the magnitude of offset could be underestimated.