Permafrost, the Active Layer, and Changing Climate*

by

Arthur H. Lachenbruch

*Transcript with minor editorial changes of a plenary address to the Sixth International Conference on Permafrost, Beijing, July 6, 1993. The talk was originally titled "Permafrost and Changing Climate," the title of a panel report containing some of this material (Nelson et al., 1993).
## Contents

**Introduction** ............................................... 2  
  Permafrost as an agent ................................... 2  
  Permafrost as a record keeper .......................... 5  
  Gas hydrate: a record keeper (and an agent?) of climate change .......... 6

**Some Remarks on Climate Signals in Permafrost Temperatures** .......... 10  
  From climate to permafrost .............................. 10  
  The temperature signal from recent warming events .................. 16  
  Temperature signals from past events—some rules of thumb ........... 20  
  False climate signals and long-term monitoring .................... 27

**References** .................................................. 32

**Appendix** .................................................... 39
INTRODUCTION

Permafrost is a temperature condition of the solid Earth and its distribution depends exclusively on the local heat balance in cold regions. The widely discussed models for contemporary greenhouse warming generally predict that changes will be greatest in these cold regions (Houghton et al., 1990) and in general, they will alter the surface heat balance and the temperature and distribution of permafrost. Changes in the position of the top of permafrost (i.e., in depth of summer thaw) and in the distribution of warm marginal permafrost will respond promptly to climate and these changes can impact the dynamics of a broad range of surface processes. However, changes in the position of the lower boundary of permafrost will generally be unimportant for hundreds or thousands of years; during this time the downward propagating thermal signal will generally preserve a lingering record of the climatic event at depth. Thus in the presence of a changing climate, permafrost can play an important role as an agent of surface environmental change and as a recorder of it. After an introductory discussion for context, I shall leave the former to my colleagues and elaborate somewhat on the latter.

Permafrost as an agent

It is paradoxical that in permafrost terranes, the portion of the ground that has the greatest influence on surface dynamics is the very portion that is not permafrost, viz., the active layer including its vegetation. The permafrost, of course, imparts to the active layer its important characteristics: a base generally at sub-freezing temperature and impermeable to moisture, conditions hostile to the penetration of roots. Thus under typical conditions the active layer is the growth medium for biotic systems and the reservoir for their water and nutrient
supply (Gersper et al., 1980), the locus of most terrestrial hydrologic activity (Kane et al., 1992; Hinzman et al., 1991; Dingman et al., 1980), and a boundary layer across which heat, moisture, and gases are exchanged between the solid earth and atmospheric systems.

Environmental impacts of permafrost and its growth and deterioration with changing climate depend primarily upon the amount and form of the ice it contains; both interstitial ice and massive ice bodies are common. The impacts are almost all manifestations of the dramatic change in strength and heat-transfer properties that occur during the phase change ice-water. For example, melting of interstitial ice decreases strength and increases permeability permitting increased water flow. This permits increased advective heat transfer, and accelerated melting in an unstable progression that can cause collapse of massive ice, soil flowage, and disruption of the landscape.

With climate warming (increased mean annual and/or summer seasonal temperature) the thaw depth and surface settlement generally increase, but not uniformly; e.g., deepening troughs can form over the network of massive ice wedges. This, in turn, alters drainage patterns and distribution of wet and dry habitats. The changed distribution and motion of the soil-water, in turn, can have a dominant effect on the thermal and chemical balance including the rates of biogeochemical reactions the productivity of living systems, the decomposition of organic matter, the generation or uptake of CO$_2$ and CH$_4$ and other characteristics of the active layer that influence (and are influenced by) the distribution of plant communities (Shaver et al., 1992; Oechel and Billings, 1992; Oechel et al., 1993).

In warmer marginal permafrost, increasing surface temperature can cause summer thawing to depths too great to refreeze in winter. The resulting permanently thawed zones or
"taliks" in permafrost can be conduits for ground water flow and associated convective heat transfer (e.g., Anisimov, 1989). This can result in very complex thermal and hydrologic regimes in marginal permafrost areas that can cause them to deteriorate rapidly with effects on surface processes of the type discussed above. Unlike the slow loss of permafrost by heat conduction from below in continuous permafrost areas, the rapid loss of warm discontinuous permafrost by the growth of taliks is a process that is not well known in detail. The latitudinal change of mean annual surface temperature near the warm margins of continental permafrost generally ranges from 0.3°-3°C/100 km (Zamolotchikova, 1988; Mackay, 1975; Anisimov, 1989). Thus in some regions like central Canada and Siberia, climatic warming of a few degrees can subject vast marginal areas to decay. This process, together with effects of a thickening active layer, will be responsible for the major prompt (10^0 - 10^2 yrs) impacts on the dynamics of permafrost terranes subjected to a warming climate.

The foregoing changes in the mechanical and thermal condition of permafrost as its moisture changes state during climate change can cause interacting surface effects (mechanical, thermal, hydrologic, and biological) whose major impacts can be grouped as follows:

1. Changes in the physical and chemical environment for plant communities (terrestrial and aquatic) and changes in the physiography of natural landscapes and the distribution of natural habitats for birds and mammals.

2. Changes in the suitability of terrane for development of infrastructure and human communities and in the economics of land use and exploitation of natural resources.

3. Changes in the boundary layer across which heat moisture and gases are exchanged between the solid earth and atmosphere.
Permafrost as a record keeper

In cold continuous permafrost, as in other impermeable earth materials, there is virtually no movement of ground water and hence heat transfer is exclusively by thermal conduction, a process that follows relatively simple mathematical rules. Thus if a climate change in the past changed the mean annual temperature of the base of the active layer, that change would propagate slowly downward into the permafrost at a rate that can be calculated. In effect, the ground "remembers" the major events in its surface temperature history, and careful temperature measurements made today to depths of a few hundred meters can provide information on the history of local surface temperature during past centuries. For the moment, we shall defer further discussion of this process, as it is the subject of the next section.

Bodies of water that do not freeze to the bottom in winter (those deeper than 2 or 3 meters) generally have mean bottom temperatures near or above 0°C, whereas the mean temperature of the adjacent land surface may be much lower (e.g., -10 to -15°C near the shores of the Arctic Ocean today). Hence insofar as the solid earth is concerned, the migration of a shoreline (lake or ocean), leaves a dramatic "climatic change" in its wake, a warming if the shore is advancing on the land, a cooling if the shoreline is retreating from the land. Thus the geothermal methods used to investigate climate history can be applied also to temperature measurements in wells on the submerged Arctic continental shelf to interpret the chronology of rapid shoreline transgression in progress on much of the Arctic shoreline today, and of the inundation of the Arctic continental shelf following the last glaciation. Effects of predicted global sea level rise on low-lying permafrost can be estimated by the same methods (Lachenbruch, 1957; Lachenbruch et al., 1982; Lachenbruch et al., 1966, 1988a; Taylor et al., 1983).
Additional information about the history of the recent climatic system is obtained by measuring the temperature, depth and ice content of existing permafrost and modeling the surface conditions that must have existed to generate the permafrost observed. Where the ice content is high and permafrost is deep as in Prudhoe Bay, Alaska, or Eastern Siberia, calculations of this sort provide climatic information on time scales approaching \(10^5\) yrs, the period of glacial cycles (Balobeov et al., 1978; Harrison, 1991; Osterkamp and Gosink, 1991).

In addition to the information on climate history contained in the present thermal state of permafrost, much can be learned about past surface conditions from the cold-climate geomorphic features and organic materials preserved in buried permafrost (Carter et al., 1987; Mackay, 1988). Of particular interest are ice-wedge polygons which form from the percolation of summer melt water into a network of deep thermal contraction cracks that form during winter in cold brittle permafrost. Evidence on the seasonal growth and deterioration of past generations of ice wedges reveals changes in past surface conditions where these networks of massive ice are preserved in the stratigraphic record (Lachenbruch, 1962, 1966; MacKay, 1976; Mackay and Matthews, 1983).

Gas hydrate: a record keeper (and an agent?) of climate change

Sediments in cold regions may trap large quantities of natural gas, largely methane, in ice-like crystalline structures containing water molecules and called "gas hydrates" or "clathrates" (Katz, 1959; Davidson et al., 1978; Kvenvolden, 1988). They store natural gas efficiently (up to \(\sim 150\) times as much methane as an equal volume of free gas under standard conditions of temperature and pressure). Gas hydrates are of interest in connection with global
climate change as a potential source of atmospheric methane, one of the most important greenhouse gases (they are also of interest as a potential commercial source of energy).

Gas hydrates are stable under special conditions of low temperature and/or high pressure that obtain in only a small superficial part of the solid earth. The stability field is shown (shaded) in Figure 1, assuming the common ("hydrostatic") condition wherein fluid pressure is equal to the weight of a column of water extending to the surface. For the typical Arctic coastal conditions illustrated for Barrow, Alaska ("0 yrs," Fig. 1a) methane hydrate is stable between about 200 and 700 m, and permafrost extends downward below a thin active layer to a depth of about 400 m. Under steady conditions, hydrates will be stable on the continent only in colder permafrost regions. (Note that for the typical steady-state gradient illustrated in Figure 1a (~30°C/km), the surface temperature would have to be below about -5°C for the geotherm to intersect the hydrate field.)

The principal global-change question about gas hydrate is: Will a warming climate destabilize gas hydrates releasing methane to the atmosphere, and thereby enhancing the warming by its contribution to the greenhouse effect?

The geothermal effect of a sudden increase of surface temperature by 10°C is illustrated by the family of geotherms in Figure 1a. This is much more severe than predictions of climate models, but it is an accurate representation of the "climate change" that occurs when Arctic seas override the land, a process that must have occurred over millions of square kilometers on the Arctic continental shelf as it was inundated by rising sea level following the last glaciation. Transgression continues today at rates exceeding a meter per year along much of the Arctic coastline (e.g., Mackay, 1986). It is seen from Figure 1a that effects of a thermal disturbance today would not reach the top of the gas hydrate field at ~200 m for centuries, and for the
Figure 1. Thermal decay of methane hydrate stability field (shaded) beneath a transgressing Arctic shoreline. Numbered curves represent calculated earth temperature at specified time in years since inundation. Land with geotherm typical of Barrow, Alaska, (part a), or Prudhoe Bay, Alaska, (part b), is inundated by sea (mean annual bottom temperature, $-1^\circ$C) at time $t = 0$ yrs. Destabilization takes 10 times longer at Prudhoe Bay (50,000 yrs compared to 5,000 yrs at Barrow) because of latent heat absorbed by its high ice-content permafrost (Lachenbruch et al., 1982, 1988a, 1988b; Saltus and Lachenbruch, unpublished calculations).
example illustrated (for initial conditions typical of Barrow, Alaska, today, Lachenbruch et al., 1988a), significant methane release could not occur for thousands of years. If unlike the "Barrow model," Figure 1a, the permafrost were ice-rich as at Prudhoe Bay (Lachenbruch et al., 1982, 1988b), the extra latent heat would retard warming, and the destabilization would take tens of thousands of years (Fig. 1b). Although the phase boundaries of gas hydrates vary with composition, fluid pressure, and salinity, these examples serve to illustrate two rather general points:

1. Significant amounts of methane will not be released from destabilized methane hydrate in permafrost areas by present-day climate change for a millennium or more.

2. Present-day contributions to atmospheric methane from destabilized hydrate should be sought on the Arctic continental shelf where warming started with inundation thousands of years ago (Judge and Majorowicz, 1992; Osterkamp and Fei, 1993; McDonald, 1990). (This, of course, is not to say that the separate issue of present-day contributions of methane to the atmosphere from organic reactions in the active layer and thawing permafrost might not be very significant, and in fact, their study is a high priority for permafrost research in the context of changing climate.)
From climate to permafrost

As emphasized in the last section, the important environmental changes associated with climate change in permafrost terrains do not occur in permafrost but in the active layer above it. However, the permafrost generally shares its upper boundary with the base of the active layer (Figure 2) and it acts as a listening post for temperature changes that occur there. However complex and interesting the thermal changes in and above the active layer might be, the only parameter recorded by permafrost is the temperature change (for practical purposes, the mean annual temperature change) that makes its way to the base of the active layer. It is this quantity, the mean annual temperature at the top of permafrost ($T_{pf}$, Figure 2), whose time history is estimated in so-called climatic reconstructions from borehole temperature measurements.

Such reconstructions are based on the assumption that, except for effects of steady heat flow from the earth’s interior, a change in (mean annual) temperature with depth can be caused only by a past change in surface temperature. This will be true where heat transfer is exclusively by (one-dimensional) conduction with non-changing properties, and where effects of sub-surface sources and sinks, if they exist, cancel each other over the yearly cycle. Although these conditions are generally satisfied in most cold permafrost, they clearly do not apply in the overlying active layer, snow, and air. Thus in general $T_{air}$ (Figure 2) will differ from $T_s$ (mean annual temperature of the "solid surface"; snow in winter, ground in summer), and $T_s$ will differ from $T_{pf}$, whether or not the climate is changing. Effects above the solid surface are addressed by the radiation balance of climatology which we consider briefly later.
Figure 2. Measurement sites for the differently defined mean annual surface temperatures: $\bar{T}_{pf}$ at upper surface of permafrost, $\bar{T}_g$ at ground surface, $\bar{T}_s$ at the solid surface of the snow pack when it is present and ground surface when it is not, and $\bar{T}_{air}$ in a standard observatory thermometer shelter.
Between the solid surface and permafrost, the temperature offset ($\tilde{T}_{ss} - \tilde{T}_{pf}$, Figure 2) is determined by the complex dynamics of the snow-active-layer system. Clearly, there are many different types of surface change, climatic and otherwise, that can produce the same change in permafrost temperature, and some environmentally important thermal changes in the active layer might have little effect on $\tilde{T}_{pf}$. Thus it is important to maintain a distinction between changing permafrost surface temperature and changing climate.

Simple references to "Climatic warming" tend to imply that it can be characterized or defined by an increase in mean annual air temperature near the Earth's surface. Figure 3 is a simple reminder of the inadequacy of such a characterization for predicting environmental effects; such effects are controlled by the active layer and are most sensitive to summer temperatures. The three cases illustrated all represent a 4°C mean annual surface warming (from -9°C to -5°C) achieved respectively, with warmer summers (3a), warmer winters (3b), or much warmer winters with somewhat cooler summers (3c). As indicated by the corresponding change in ratio of freezing to thawing degree-days, the first might have dramatic effects on surface environments, the third might have little or none; the second is more consistent with most GCM predictions (Houghton et al., 1990; Maxwell, 1992). In all three cases, the warming will generally propagate to permafrost where it may be detectable for centuries, but distinctions among the three original surface conditions will not be possible from the "climate" reconstruction.

Next to a change in mean air temperature, the most obvious agent to shift $\tilde{T}_{pf}$ is winter snow cover which insulates the ground in winter causing $\tilde{T}_{pf} > \tilde{T}_{ss}$ (e.g., Goodrich, 1982); a secular increase in snow cover (a bona fide climatic change) can cause a conspicuous warming
Figure 3. Three scenarios for the same warming of mean ground surface temperature $T_g$ (from -9°C, solid, to -5°C, dashed curves). Summer conditions that control the environmentally sensitive active layer are quite different for each as indicated by the ratio of annual freezing degree days to thawing degree days (fr index/thaw index).
signal in permafrost. A more subtle $\Delta \tilde{T}_{pf}$ signal from the snow can occur with no change in snow cover or mean air temperature if the amplitude of seasonal temperature variation increases; the same snow cover causes more warming because of greater damping in colder winters as the climate gets more continental (Lachenbruch, 1959). This effect accounts for the marked difference in $\tilde{T}_{pf}$ from -9°C at Barrow on the Alaskan Arctic coast to -6°C observed at Umiat 80 km inland. Both places have similar mean air temperature and snow cover, but Umiat has a more continental climate.

A cause of shift in the mean annual temperature between the air/solid interface and the top of permafrost is shown schematically in Figure 4. For factors that make it easier for seasonal heat to go downward through the snow and active layer, than upward out of it, $\tilde{T}_{pf}$ will be greater than $\tilde{T}_{ss}$ when the two have reached equilibrium. Examples are infiltration of summer melt water and rain, a unidirectional process that adects heat downward, and the snow effect that insulates the ground selectively in winter (these are designated type II processes in Figure 4). Examples of processes (Type I) that can cause a negative offset of mean permafrost temperatures are the change in the active layer from a good thermal conductor when it is frozen in winter to a poorer one inhibiting downward flow of heat when it is thawed in summer. An effect of the same sign is the upward transport of moisture (and dehydration, reducing conductivity) in the active layer to supply surface evaporation and transpiration; it tends to counteract heat gained by downward conduction in summer. A number of other processes involving non-conductive transport of sensible and latent heat by migrating moisture, sometimes involving vaporization, condensation and ice segregation in the active layer have been described (e.g., Outcalt et al., 1992; Haller, 1978; Mackay, 1983; McGaw et al., 1978; Nakano and
Figure 4. Snow and active layer processes of type I cause the seasonal outflow of heat in winter \( (Q_{\text{out}}) \) to exceed the seasonal influx of heat in summer \( (Q_{\text{in}}) \). They generally contribute to a decrease in mean temperature with depth. Those that cause the periodic inflow to exceed the periodic outflow (type II) generally contribute to an increase in temperature with depth. Environmentally induced changes in any such snow or active-layer processes are remembered in the permafrost temperature record as changing "climate."
Brown, 1972). Subtle changes and interactions among the factors that control some of these processes (e.g., prolongation of the zero curtain by non-linear interaction of latent heat of soil freeze/thaw and snow cover, Goodrich, 1978) could cause a change in heat balance at the top of permafrost that would be remembered by permafrost temperatures as a change in $\bar{T}_{pf}$. The permafrost surface temperature, $\bar{T}_{pf}$, might be the most important single parameter to describe the integrated history of environmental conditions (and we are fortunate that this is the parameter whose history is recorded by temperature signals in permafrost). However, understanding the full range of implications of changing $\bar{T}_{pf}$ requires a knowledge of the physics of thermal processes like those just discussed that control it, and supplementary information where possible. Past changes in $\bar{T}_{pf}$ may be a direct response to changing climate parameters (such as air temperature or snow cover), or to changing geomorphic conditions, drainage patterns, or the heat and water balance of evolving plant communities. In any case changes in $\bar{T}_{pf}$ are a unique source of information that we should attempt to relate to the environmental factors (including climate) that control the dynamics of changing ecosystems.

The temperature signal from recent warming events

Figure 5b shows a group of typical temperature profiles from oil wells on the Alaskan Arctic coastal plane (Figure 5b). With few exceptions, the profiles generally have a linear portion at depth and a curved portion near the surface. The simplest interpretation of the profiles is that the linear portion at depth represents steady heat flux from the Earth’s interior (which we measure to be .05 to .10 W/m²), and the curved part represents a recent warming event propagating downward (the shaded area). The amount of surface warming ($\Delta \bar{T}_{pf}$) is the difference between the temperatures obtained by extrapolating the straight and curved parts to
Figure 5. Temperatures from abandoned wells in northern Alaska. Borehole temperatures in part b are from sites denoted by solid circles in part a. Shaded region for each curve is the recent warming anomaly; numbers by curves denote time in years before 1984 for start-up of best-fitting model of linear warming. Insert in part b summarizes statistics for starting date and total warming for best-fitting step and linear models (after Lachenbruch and Marshall, 1986). Warming is indicated at all sites except those denoted by ○.
the surface, a few °C in these cases (see curve CTD, Figure 5b). Duration of the warming event is estimated from the depth of penetration of the anomalous curvature, several decades to a century according to simple heat-conduction models referenced in the right-hand margin of Figure 5b. (Arrows show where the bottom of the anomaly should be 10 or 100 yrs after the start of a step or linear change in surface temperature.)

Figure 6 compares three formal reconstructions of surface temperature history (i.e., \( \tilde{T}_{pf} \)) from the anomaly (shaded region) at AWU (Figure 5b). The timing and magnitude of the step and linear functions were determined by the least squares method described by Lachenbruch and Marshall (1986), and the smooth curve (Clow, unpublished) by the method of spectral expansion (Parker, 1977) in which the unknown time history is represented by a series of orthonormal functions. All of the reconstructions reproduce measured temperatures below 30 m within observational error, and there is little basis for preferring one to the others. The results show that although we cannot resolve details of the surface temperature history, there can be little doubt about the big picture; at AWU and most of the other holes in this 2 \( \times \) 10\(^6\) km\(^2\) Arctic region, there was a marked, but laterally variable increase in temperature at the top of permafrost (\( \tilde{T}_{pf} \)) of a few degrees Celsius during the 20th century (Lachenbruch and Marshall, 1986; Lachenbruch et al., 1988; Clow et al., 1991). (Some sites show a more recent cooling believed to be related to engineering disturbance, Lachenbruch and Marshall, 1986.) A reason for this dramatic locally variable change has not been found in the limited records of surface temperature or snowfall from the region, and its cause remains unknown.

From Figure 5b, it can be seen that heat accumulation in the solid Earth from the unbalanced downward "climatic" flux (which we denote by \( c \) in Figure 7) must be of the same order as the upward geothermal flux (denoted by \( g \) in Figure 7) which it nullifies to yield near-
Surface-Temperature History
Awuna

Figure 6. Best-fitting step, ramp and spectral expansion models for recent warming event at AWU, Figure 5b (after Lachenbruch et al., 1988c; Clow, unpublished).
zero gradients near the surface (Figure 5b). The estimates for AWU \( (g \sim 0.06 \, \text{W/m}^2 \) and \( c \sim 0.16 \, \text{W/m}^2, \) Lachenbruch et al., 1988c) are illustrated in Figure 7a. It is of interest to compare these modest solid-earth fluxes to the more vigorous activity on the other side of the Earth's surface. The larger "top-side" fluxes drive the ecosystem, and their balance point determines the Earth's surface temperature (Weller and Holmgren, 1974). They consist of the incoming and outgoing radiation components and their difference ("net," Figure 7a) which, among other things, is responsible for melting snow and evaporating water allowing them to return to the sea or the atmosphere to balance both the thermal and hydrologic budgets in preparation for the next annual cycle. The numbers on the arrows (Figure 7a) indicate that these top-side climatic fluxes sum to zero as they should if the climate is not changing. But we just determined from the solid Earth (Figure 5) that climate is changing; at Awuna about 0.16 Wm\(^{-2}\) more has been going into the Earth than out for a half century or so. As this is just 1/100th of the net radiation (itself a difference of large numbers), we could never detect it by trying to keep a balance sheet of fluxes at the Earth’s surface. Thus while the unbalanced climate flux \( c, \) Figure 7) is an inconspicuous second-order effect in the climatic system, it is a conspicuous first-order effect that dominates the thermal regime of the upper 100 m in the solid-earth system (Figure 5 and 7b)—the solid Earth is a good monitor for changes in the surface energy balance. The downward flowing heat cannot go far in a century because the Earth is a poor conductor; it is all contained in the stippled region (Figure 7b), which represents the complete climatic event from start to finish.

**Temperature signals from past events—some rules of thumb**

In the foregoing examples, it was possible to extract useful information from simplified models without great measurement precision or detailed lithologic information because the
Figure 7. Typical thermal conditions in permafrost regions of the Alaskan Arctic (after Dingman et al., 1980). A. Average annual energy fluxes above and below the Earth's surface. B. Geothermal regime: g is steady geothermal heat loss, C is heat gain from warming climate, stippled region represents total heat accumulation from warming event (from Lachenbruch et al., 1988c).
anomalies were large relative to errors from these sources. In these cases, the causal events were obviously recent (the curvature is shallow, and the linear part suggests a prior relatively uneventful time) and for the most part, they are still going on (the temperature anomaly is greatest near the surface). In effect, we have done what might be called a "last event analysis" when we extrapolated the linear portion upward to isolate the near-surface curvature. We have not necessarily assumed that the linear part represents a steady state, but that the older transients it may contain have relatively little curvature compared to the near-surface anomaly of interest. This is a reasonable procedure that gives useful constraints where the last event is so recent and conspicuous. However, when we try to assign formal estimates of error and to reconstruct a more complete history of the succession of past events, the problem gets more complicated. It generally involves first systematically identifying the geothermal anomaly (from otherwise unexplained curvature in the temperature profile) and then applying inverse problem theory at various levels of sophistication to decompose the geothermal anomaly into a likely causal sequence of "climatic" events. Motivated by an interest in contemporary climate change, this geothermal inversion is the subject of a rapidly growing literature; for a recent collection of papers on the subject, see Lewis (1992). We make no attempt to summarize the procedures here but offer a few comments on their application. Such reconstructions are, of course, of particular interest in permafrost for two reasons: 1) as they depend upon heat transfer by conduction, cold permafrost with virtually no moving fluids is one of the most reliable media for their application, and 2) the history of climate is of special interest for the origin of permafrost and for the central role played by polar regions in climate models.
Figure 8 shows the Earth temperature anomalies that would be caused if any one of the last five centuries were 1°C warmer than the others. For earlier centuries, the maximum anomaly gets deeper and smaller and the curvature, by which we identify anomalies, diminishes. If we break the history up into segments or events, the contributions of all but the last event (just discussed) will have a maximum, or hump, at depth. We often want to know how big the hump will be for a given event, e.g., the Little Ice Age or the last interglacial period, and how deep we must drill to examine it. A second question is how many different shaped events give the same hump in the temperature profile, or how non-unique is the inversion, and finally how can we distinguish between the hump caused by a transient surface temperature change and one caused by a steady-state condition from inhomogeneity or three-dimensional effects.

Useful back-of-the-envelope answers can usually be obtained for these questions because each of the past centuries (and many other short-duration events) can be represented to a first approximation by a single temperature-depth curve (Figure 9d) when it is plotted against coordinates adjusted for the age of the event (see Appendix). Thus a hypothetical 1°C anomaly (H=1°C, Figure 9a) throughout the 19th century has a duration Δt=100 yrs centered about the time $t \sim 150$ yrs ago. According to Figure 9d, the height of the hump ($T_{\text{max}} \approx H\Delta t/4t$) is .17°C, and it occurs at the depth ($x_m = \sqrt{2a}t \sim 8/\sqrt{t}$) of about 100 m in agreement with Figure 8. The anomaly extends to a depth of about $4x_m(\sim 32/\sqrt{t})$, or 400 m in the present case. If we want to reconstruct surface temperature history back through the 19th century, ideally the observation hole should be at least this deep (see also Clow, 1992). Any curvature at this depth must represent earlier events and is not relevant; if it is not too great (see Figure 8), the near-linear gradient can be extrapolated upward to identify the more recent anomaly sought (e.g., as in Figure 5b).
Figure 8a. Anomalous Earth temperatures that would result from an anomalous temperature of 1°C at the Earth’s surface during each of the past 5 centuries as shown in part b.
Figure 9. a. Notation for hypothetical surface temperature events of "height" $H \degree$, duration $\Delta t$, average age $t$ and fractional duration $D = \Delta t/t$. b and c. Examples of surface events with the same "area" ($H\Delta t$ degree-days) and average age $t$ that yield virtually indistinguishable geothermal anomalies because $\Delta t < t$. d. General representation of geothermal anomaly (shaded) for a surface event whose duration $\Delta t$ is $< t$. Dotted curve represents $\Delta t = t$ and solid curve is the limiting case of a pulse as $\Delta t \to 0$. 

$$X_m = \sqrt{\frac{2\alpha t}{\Delta t}} - 8\sqrt{t}$$
The approximation in Figure 9d is based on the Green’s function for heat conduction (Appendix), and it is valid as long as the duration of the event (Δt) does not exceed the central (i.e., average) age of the event (t). Thus for the widely discussed Little Ice Age (*Imbrie and Imbrie*, 1979), a cool period believed to have lasted about 400 yrs (Δt) centered around 1600 A.D. (t~400 yrs ago) with a temperature anomaly of perhaps −.8°C (H) the maximum geothermal anomaly is −0.2°C (H Δt/4t, Figure 9d); it occurs at a depth of about 160 m (8√t yrs, Figure 9d), and is spread over a depth range of 640 m (32√t, Figure 9d). A number of authors claim to have identified the Little Ice Age from borehole temperature inversions (e.g., *Cermak*, 1971), although there is not universal agreement upon its definition.

There are two important points to be made from Figure 9: 1) Rather small surface temperature events cause geothermal signals that are (technically) large enough to measure easily with modern equipment, but 2) the basic limitation of the method is that a broad range of different events give virtually the same signal (Figure 9b and c); a great increase in measurement precision is needed for a modest increase in our ability to distinguish among them (Appendix, Figure A-1).

To illustrate, we note that as long as Δt < t, the ratio of an anomalous event H° at the surface to its maximum subsurface effect, T_m, is simply D/4, where D=Δt/t, the ratio of the event’s duration to its age (or the "fractional duration"). Thus if H ~ 4°, the downhole anomaly, T_m, would be 0.1° if D=.1, or 1° if D=1, quantities well within the resolution of modern temperature measurement equipment. (Hypothetical examples would be anomalies with durations of a decade (D=.1) or a century (D=1) centered a century ago.) Note, however, that if the century-long anomaly were 0.4°C, its signal would be virtually the same as the shorter
decade anomaly with $H = 4^\circ$. Both would have $T_m \sim 0.1^\circ C$ at 80 m and much greater precision would be required to distinguish second-order differences in their earth-temperature signals.

Thus the robust forward prediction (from "climate" to geotherm) permitted by Figure 9d becomes a weakness in the inverse problem (from geotherm to "climate"). As all surface temperature events of the same age $t$ with the same "area" ($H \times \Delta t \, ^\circ C$-years) yield virtually the same geothermal anomaly (if $\Delta t \leq t$), when we observe such an anomaly, it is extremely difficult for us to determine which event is responsible. Figure 9b and c give examples of events with equal area that are virtually indistinguishable without great measurement precision, another example is the two anomalies centered in the 17th century, Figure 8b. For this reason, geothermal reconstructions of surface temperature cannot be expected to give much detail over intervals whose duration is less than their age (unless, of course, independent information on $t$, $H$, or $\Delta t$ is available from another source such as tree rings, weather records, or isotope studies). For past events whose duration is on the order of their age, however, signals are robust (see last paragraph) and useful averages can be obtained for the anomalous surface temperature.

Clow (1992) has applied the method of Backus and Gilbert (1970) to a formal analysis of resolution in the inversion of geothermal data. It provides a preliminary perspective for optimizing the design of observatories for geothermal reconstruction of surface temperature history.

**False climate signals and long-term monitoring**

Figure 10 illustrates a fundamental obstacle to the geothermal reconstruction of surface temperature history: How do we know that the curvature in the temperature profile (Figure 10a)
ALTERNATIVE STEADY-STATE EXPLANATIONS FOR CURVATURE

Figure 10. a. Steady-state representation of curvature in geotherms caused by variable thermal conductivity (a), or laterally variable surface temperature (c).
is the transient effect of changing surface temperature, and not a steady-state effect of vertically variable conductivity (Figure 10b) or laterally variable surface temperature (Figure 10c)? Conceptually the simplest test is to see whether or not the anomalous curvature is changing. For a steady-state origin it would not change, whereas for signals caused by changing surface temperature it would. In fact according to the Fourier heat equation, the rate of temperature change is predictable from the temperature curve; at each depth it is proportional to the curvature there, and the proportionality factor is the thermal diffusivity. Thus, in principle, if we could monitor an observation well with sufficient precision, the alternate sources of curvature could be resolved, thermal diffusivities might be estimated, surface temperature history could be refined, and future changes in surface temperature could be tracked.

Rules of thumb for the required monitoring precision are illustrated in Figure 11a, derived (in the Appendix) from the generalized temperature anomaly curve of Figure 9d. It shows that like the temperature anomaly $T_{\text{max}}$, the maximum rate of temperature change $T^\prime_{\text{max}}$ varies directly with the strength $H$ and fractional duration $D(=\Delta t/t)$ of the climatic event. In addition, it varies inversely with the central (i.e., average) age, $t$. For reasonable surface anomalies, e.g., $HD \sim 1^\circ \text{C}$, the maximum rate of temperature change is greater than 1 mK/yr for events younger than 275 years. Figure 11b shows dimensional results for anomalies with $HD = 1^\circ \text{C}$ and average ages of 100, 200, and 400 yrs. The latter roughly represents the Little Ice Age; its temperature is changing at a maximum rate of only 1 mK every 2 years or so at the 120 m depth ($\sim 6\sqrt{t}$). Note that an event (e.g., the "Younger Dryas") with a $10^\circ \text{C}$ cooling for 1000 yrs about 10,000 yrs ago ($H = 10^\circ \text{C}$, $D = 0.1$) would, like the Little Ice Age, have $DH \sim 1^\circ \text{C}$ and $T_{\text{max}} \sim 1/4^\circ \text{C}$ (Appendix, equation 7), but the signal would be spread over a 3-km
Figure 11.  
a. General approximation (exact for D = 0) for rate of change of Earth temperature after a surface event of $H^\circ$ of fractional duration $D < 1$ (from Appendix, equations 10b and 11).  
b. Numerical results from a) for $DH = 1^\circ K$ and $t = 100, 200, \text{ or } 400$ yrs.
depth (Figure 9, equation 9) with little identifiable curvature. The maximum rate of change today would be unmeasurable (only 1 mK every 40 years, equation 11); it would occur at a depth of $\sim 600$ m (equation 12). For the recent events of the Alaskan Arctic (Figure 5) temperatures in the upper 100 m are changing at easily measurable rates on the order of 10 mK/yr.

It is clear from these results that if borehole temperatures can be remeasured over a period of years with millidegree precision, we can overcome major obstacles to the reconstruction of surface temperature posed by steady-state effects of inhomogeneity and three-dimensionality. In our experience, commercially available thermistor transducers, metering bridges, and borehole logging cables are adequate to construct systems with a sensitivity of a fraction of a milliKelvin (mK). The two principal limitations to attaining the desired precision in the field are: 1) returning to the same measurement depths with adequate precision on relogging, and 2) controlling the random convection of the generally unstable fluid in the borehole. Good progress is being made in overcoming these problems at the U.S. Geological Survey in Menlo Park and other laboratories.
References


APPENDIX: Derivation of Climatic Rules of Thumb

We let \( t = \) time before present and consider the present temperature disturbance \( \theta \) in the homogeneous half-space \( x \geq 0 \), caused by an anomalous surface temperature \( h(t) \) during an event extending from time \( t_1 \) to time \( t_2 \).

The initial and boundary conditions are:

\[
\begin{align*}
\text{at } x = 0: & \quad \theta = 0, \quad t < t_1, \quad t > t_2 \\
& \quad = h(t), \quad t_1 \leq t \leq t_2 \\
\text{at } t = t_2: & \quad \theta = 0, \quad x \geq 0
\end{align*}
\]

The appropriate solution to the heat-conduction equation (modified from Carslaw and Jaeger, 1959, p. 63) is

\[
\theta(x, t_1, t_2) = \frac{1}{\sqrt{\pi}} \int_{t_1}^{t_2} \frac{x}{\sqrt{4\alpha t'}} e^{-\frac{x^2}{4\alpha t'}} \frac{h(t')}{t'} dt'
\]  

(1)

We denote the duration of the event by \( \Delta t \) and its average or "central" time by \( t \), i.e.,

\[
\Delta t = t_2 - t_1 \\
t = \frac{1}{2} (t_2 + t_1)
\]

If \( h(t) \) is constant over the interval, integration of (1) yields the Earth temperature anomaly \( B \) for a "Box car" climate event of uniform height \( H \), duration \( \Delta t \) and central age \( t \).

\[
B(x, t, \Delta t) = H \left\{ \text{erf} \frac{x}{\sqrt{4\alpha(t - \Delta t/2)}} - \text{erf} \frac{x}{\sqrt{4\alpha(t + \Delta t/2)}} \right\}
\]  

(2)

If the event’s duration is small relative to its age (\( \Delta t / t \ll 1 \)), we may approximate (2) without integrating (1) by setting \( h(t) \approx H \).

\[
W(x, t, \Delta t) = H \Delta t \left\{ \frac{1}{\sqrt{\pi}} \frac{x}{\sqrt{4\alpha t}} e^{-\frac{x^2}{4\alpha t}} \right\}
\]  

(3)

The expression in braces in (3) is the Earth temperature anomaly caused by an instantaneous unit pulse of surface temperature at time \( t \) before present. In this
approximation, the intensity of the pulse (or its area in years \( \times \) degrees Kelvin) is represented by the product of its duration \( \Delta t \) and average height \( H \).

With the substitutions

\[
\begin{align*}
    u &= \frac{x}{\sqrt{2\alpha t}} \\
    D &= \frac{\Delta t}{t}
\end{align*}
\]

(4a) and (4b) become

\[
\begin{align*}
    B(u, D) &= H \left\{ \text{erf} \frac{u}{\sqrt{2(1 - \frac{1}{2}D)}} - \text{erf} \frac{u}{\sqrt{2(1 + \frac{1}{2}D)}} \right\} \\
    W(u, D) &= HD \frac{u}{\sqrt{2\pi}} e^{-u^2/2}
\end{align*}
\]

(5) and (6)

This parameterization permits the use of the convenient dimensionless depth \( u \) to compare events of different ages. It also shows that the size of a geothermal anomaly does not depend on the duration and age independently but on the ratio \( D \) which we call the “fractional duration” of the event.

Figure A-1 shows that (6) is a reasonable first approximation to the integrated result (5) as long as \( D \leq 1 \). It reaches a maximum \( (T_{\text{max}}) \) of \( 0.242 \, HD \) at \( u = 1 \) and falls to \( \leq 10^{-3} \, HD \) at \( u = 4 \). This leads to a general rule of thumb for the ratio of the size of a climate anomaly \( (H) \) and the maximum geothermal anomaly \( (T_{\text{max}}) \) it produces

\[
\frac{T_{\text{max}}}{H} \approx \frac{D}{4} , \quad D \leq 1
\]

(7)

The maximum occurs at the depth \( u = 1 \), equivalent to

\[
\begin{align*}
    x &= \sqrt{2\alpha t} \quad (8a) \\
    &\sim 8\sqrt{t} \, \text{yrs}, \quad x \text{ in meters} \quad (8b)
\end{align*}
\]
Figure A-1. Generalized representation of geothermal anomalies from surface events of height $H^o$ and fractional duration $D$. The approximation for the Green's function or pulse ($D = 0$) applies well up to $D = 1$. The step function extending to the present is $D = 2$ (part a).
The anomaly extends from the surface $u = 0$ to the depth $u = 4$, equivalent to

\[ x = 4\sqrt{2\alpha t} \tag{9a} \]

\[ \sim 32\sqrt{t} \text{ yrs , } x \text{ in meters} \tag{9b} \]

(For convenience in numerical rules of thumb, we have used a typical value for thermal diffusivity $\alpha$ of permafrost.

\[ \alpha \approx 10^{-6} \text{ m}^2/\text{s} \]

\[ \approx 31.5 \text{ m}^2/\text{yr} \]

A variation in permafrost diffusivity from 20–50 m$^2$/yr is not uncommon.

According to Figure A-1, the integrated anomaly $B(u, D)$ differs from the instantaneous-pulse approximation $W(u, d)$ at $T_{\text{max}}$ by less than 10% when $D = 1$, 2% when $D = \frac{1}{2}$, 0.5% when $D = \frac{1}{4}$.

The rate of change of $B$ and $w$ with time are:

\[ B_t = \frac{\partial B}{\partial t} = \frac{H}{t} \frac{u}{\sqrt{2\pi}} \left[ \frac{1}{(1 + D/2)^{3/2}} e^{-\frac{u^2}{(1 + D/2)^{3/2}}} - \frac{1}{(1 - D/2)^{3/2}} e^{-\frac{u^2}{(1 - D/2)^{3/2}}} \right] \tag{10a} \]

\[ W_t = \frac{\partial w}{\partial t} = \frac{HD}{t} \frac{u}{2\sqrt{2\pi}} e^{-\frac{u^2}{8}} [u^2 - 3] \tag{10b} \]

The maximum $T_{\text{max}}$ in $W_t$ occurs at $u = 0.742$ and the minimum at $u = 2.33$; it vanishes at $u = \sqrt{3}$. Hence from (10b) the maximum rate of temperature change $\dot{T}_{\text{max}}$ following an $H$-degree surface disturbance of duration $\Delta t$ centered at time $t$ before present is (provided $t \geq \Delta t$)

\[ \dot{T}_{\text{max}} \approx 0.275D \frac{H}{t} \tag{11} \]

where $D = \Delta t/t$ is the fractional duration of the disturbance. This maximum occurs at the depth
These relations are illustrated in Figure 12.

Substitution of (7) in (11) yields an additional useful rule of thumb:

\[
\dot{T}_{\text{max}} \approx \frac{T_{\text{max}}}{t} \quad (13)
\]

\(i.e.,\) the maximum rate of temperature change in the Earth from a surface event (with \(D \lesssim 1\)) is the maximum earth temperature disturbance from that event divided by the event's average age. Note that although the maximum temperature occurs at the depth \(8 \sqrt{t} \text{ yrs}\) (equation 9), the maximum rate of change occurs at \(\sim 6 \sqrt{t} \text{ yrs}\) (equation 12).