A Computational Method for Prediction and Regionalization of Permafrost
Author(s): Frederick E. Nelson and Samuel I. Outcalt
Source: *Arctic and Alpine Research*, Vol. 19, No. 3 (Aug., 1987), pp. 279-288
Published by: INSTAAR, University of Colorado
Stable URL: http://www.jstor.org/stable/1551363

Arctic and Alpine Research, Vol. 19, No. 3, 1987, pp. 279–288

A COMPUTATIONAL METHOD FOR PREDICTION AND REGIONALIZATION OF PERMAFROST

FREDERICK E. NELSON
Department of Geography, Rutgers University
New Brunswick, New Jersey 08903, U.S.A.

SAMUEL I. OUTCALT
Department of Geological Sciences, University of Michigan
Ann Arbor, Michigan 48109, U.S.A.

ABSTRACT

The "frost number," a dimensionless ratio defined by manipulation of either freezing and thawing degree-day sums or frost and thaw penetration depths, can be used to define an unambiguous latitudinal zonation of permafrost continuity. The index is computed using several variables influencing the depth of frost and thaw penetration, and can be related mathematically to the existence and continuity of permafrost. Although the frost number is a useful device for portraying the distribution of contemporary permafrost at continental scales, it is not capable of detecting relict permafrost and should not be mapped over small areas unless numerous climate stations are located in the region of interest.

INTRODUCTION

Considerable effort has been expended on delineating permafrost zones at small geographical scales (e.g., 1:5,000,000) for various parts of the earth. Much of this work involves compilation and synthesis of field observations for some region of interest; it is then standard practice to prepare maps based on these inventories of field data, separating subregions in which permafrost is thought to be areally continuous from those in which it is discontinuous. Within this broad bipartition, further subdivision (e.g., "sporadic" permafrost) is often attempted. Classic examples of such work include R. J. E. Brown's (1967) map of Canada, Ferrians's (1965) work for Alaska, and that of Kudryavtsev et al. (1980) for the USSR.

A second method for establishing a continuity-based permafrost zonation is to employ climatic data in calculations designed to predict the presence or absence of permafrost at each of a series of stations located in the region of interest. The utility of such methods has been constrained severely because only unmodified screen-level temperature data, which are imperfect indicators of conditions at the surface, are generally available. Compounding this problem, tenuous assumptions, such as invariant temperature relationships between the air and subsurface, have usually been invoked in order to translate such data into maps of permafrost distribution.

In practice, neither of these strategies has often been used without employing some element of the other. For example, R. J. E. Brown's (1967) map of permafrost in Canada employed mean annual air isotherms to delineate zonal permafrost boundaries in areas where field observations were unavailable. Conversely, predictive methods involving manipulation of screen-level climate data (e.g., Pihlainen, 1962; Harris, 1981) provide no indication of where zonal boundaries should be placed. Authors employing such indices generally take boundaries provided on maps such as R. J. E. Brown's (1967) as accurate representations of actual conditions; values of the index are computed or determined by graphical procedures for each of a series of locations, and "critical values" (zonal boundaries) within the index space are identified by locating the stations on pre-existing permafrost maps. The resulting classification can then presumably be used to identify the zonal status of stations in other regions or countries. The disadvantages of such methods lie in the assumption that satisfactory zonal boundaries have been established previously and in the fact that the methods ignore soil thermal properties, snow, ground cover, and topography. The exclusion of snow cover, a particularly important variable, makes such indices unsuitable for mapping purposes.

Coupled with the problems outlined above are others relating to scale dependence. Maps at small geographical scales (subcontinental to circumpolar dimensions) are produced by interpolation between observational locations that are often scattered widely. By their very nature, small-scale maps of permafrost distribution require simplifying assumptions about topographic, subsurface, and surface-cover heterogeneity. Unless the investigator has access to permafrost observations obtained from locations in close proximity to one another, other types of information, usually screen-level climate data, must be substituted in order to extrapolate between the observation points. In most cases extrapolation has involved assumptions such as a constant relationship between screen-level and subsurface temperatures, precluding any purely empirical solution to the problem of permafrost distribution.

The permafrost index detailed in this paper is a revision of Nelson and Outcalt's (1983) "frost number." It was designed to reduce the severity of the problems described above and to satisfy several important criteria:

(1) *Physical justification:* the model should correspond conceptually to the formation and growth of permafrost, be based on rudimentary soil heat-transfer theory, and should have the ability to incorporate several variables known to influence permafrost distribution.

(2) *Cartographic communicability:* spatial variations of permafrost conditions should be expressible in a readily comprehended cartographic format. Zonal boundaries on all maps should be defined rigorously and unambiguously.

(3) *Flexibility:* the index should be easily and rapidly computed, and use readily available data. However, the computational framework should have the capacity to incorporate any degree of complexity warranted by data availability.

(4) *Replicability:* explicit definitions should enable different workers to obtain identical results.

The paper in which the frost number was introduced (Nelson and Outcalt, 1983) provided a brief synopsis of the technique, but neglected many of its computational details. Since publication of that paper, the computational procedure and definition of the frost number have been revised extensively, making it more general and increasing its predictive capabilities. Nelson (1986, 1987) provided a general overview of the revisions, but emphasized applications and mapping strategies. The present paper emphasizes computational details and interpretation of the index.

THE FROST NUMBER

Air Frost Number

Nelson and Outcalt's (1983) frost index can be used to obtain a generalized impression of the distribution of permafrost, given only the mean temperatures of the warmest and coldest months and snow-cover estimates for a series of stations distributed over a region of interest. The technique can also be extended to incorporate soil thermal properties and moisture content in the calculations.

The "air frost number" F, which is equivalent to Harris's (1981) graphical procedure for classifying locations, was defined by Nelson and Outcalt (1983) as

$$F = \left[\frac{DDF}{DDT} \right]^{1/2} \tag{1}$$

where DDF and DDT are the freezing and thawing indices (°C days) for a given location. Since these statistics often are not readily available, temperature curves are approximated by cosine functions, which are integrated to obtain the freezing and thawing indices (Figure 1). The required values can be obtained by means of equation set (2):

$$\bar{T} = (\bar{T}_h + \bar{T}_c)/2 \tag{2.1}$$

$$A = (\bar{T}_h - \bar{T}_c)/2 \tag{2.2}$$

$$\beta = \cos^{-1} (-\bar{T}/A) \tag{2.3}$$

$$\bar{T}_s = \bar{T} + A(\sin \beta/\beta) \tag{2.4}$$

$$\bar{T}_w = \bar{T} - A[\sin \beta/(\pi-\beta)] \tag{2.5}$$

$$L_s = 365(\beta/\pi) \tag{2.6}$$

$$L_w = 365 - L_s \tag{2.7}$$

$$DDT = \bar{T}_s \cdot L_s \tag{2.8}$$

$$DDF = -\bar{T}_w \cdot L_w \tag{2.9}$$

where $\bar{T}$ is the mean annual air temperature, A is the annual temperature amplitude, $\bar{T}_h$ and $\bar{T}_c$ are the mean temperatures of the warmest and coldest months, β is the "frost angle," the point along the time axis where the temperature curve crosses 0°C, $\bar{T}_s$ and $\bar{T}_w$ are the mean summer and winter temperatures, and L_s and L_w are the length of summer and winter. All temperatures are assumed to be expressed in degrees Celsius.

A more detailed derivation of this particular strategy was given by Nelson and Outcalt (1983). Although this "min/max" method is especially useful for regions with sparse or incomplete data (Haugen et al., 1983), it can yield poor results for at least two reasons. Large errors can result from underestimation of a small freezing or thawing index when the complementary index is large (Figure 2a). In such cases, F could attain very large values or become undefined for very cold stations, or it could approach zero for warm stations, even when use of empirical data would yield less extreme results. A second type of error occurs when the cosine curve does not closely approximate the actual temperature curve (Figure 2b). Although the cosine curve fits records for some stations very well (Figure 2c), the two error types can occur in conjunction with one another, leading to large discrepancies between values of F obtained with empirical and computed degree days (Figure 2d).

The magnitude of errors of the first type can be reduced by restricting F to the closed interval [0,1], redefining it as

$$F = \frac{DDF^{1/2}}{DDF^{1/2} + DDT^{1/2}}\,. \qquad (3)$$

This adjustment also allows expression of the index for stations at which no mean monthly temperature exceeds 0°C. Errors of the second type can be reduced by employing standard 12-mon mean temperature series, which are available from a wide variety of sources. Although a curve obtained by using a single harmonic (e.g., Panofsky and Brier, 1958) does not necessarily improve estimates provided by the min/max method, additional harmonics increase the goodness of fit up to the maximum of six, which provides an exact fit for curves of mean monthly temperature. This produces slightly differing accuracies between DDT and DDF, however, owing to the implicit assumption of equal spacing between (monthly) data points. The result is that values of F computed using these estimates diverge slightly from those obtained with observed values.

An alternative procedure, capable of handling unequally spaced points, uses cubic spline interpolation (e.g., King, 1984: 166–189) to produce a smooth temperature curve that passes through all of the monthly means. Interpolated mean daily temperatures are summed below and above 0°C to obtain the freezing and thawing indices, respectively. Although both freezing and thawing indices are underestimated because short-term crossings of the freezing point during "changeover" months are neglected, the errors tend to cancel when F is computed for all but very warm or very cold stations. The excellent agreement between predicted and observed values for 102 Canadian stations with complete 30-yr records (Canada, 1982a, 1982c) is shown in Figure 3.

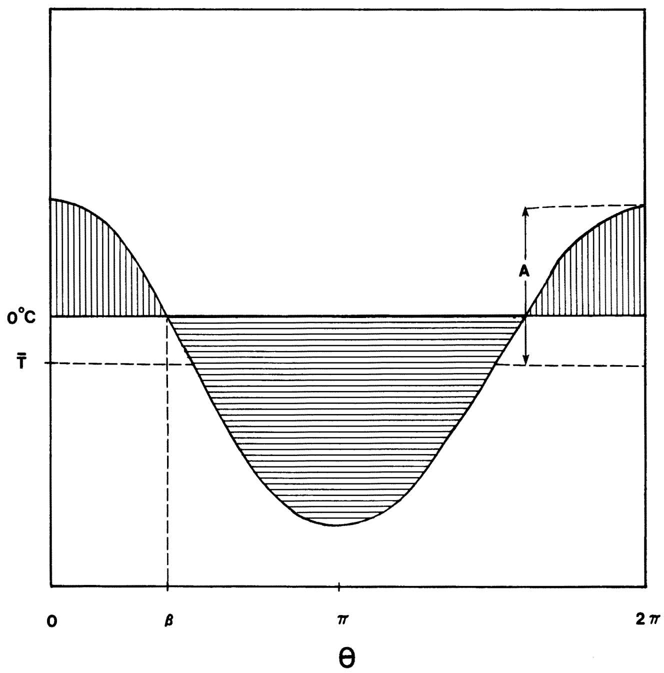

FIGURE 1. Simple cosine approximation of annual temperature series. Vertical shading represents thawing index DDT. Horizontal shading represents freezing index DDF. See text for definition of symbols.

SURFACE FROST NUMBER

The air frost number is not an optimal predictor of permafrost distribution because the very large effects of snow cover and other variables influencing permafrost occurrence are ignored. The time of year at which the snow falls, and the density, thickness, and duration of the snow cover all exert strong influence on the thermal regime of the soil. Snow cover can be the critical factor determining the presence or absence of permafrost (e.g., Granberg, 1973).

Although the important modification exerted by snow on ground temperatures and periglacial features is well known and has been demonstrated amply in the field (e.g., Mackay and MacKay, 1974; Smith, 1975) and in both analytic (Lachenbruch, 1959) and numerical models (Goodrich, 1982), most computational methods of permafrost prediction ignore snow cover. One popular approach to assigning a zonal status to climate stations (Harris, 1981; Greenstein, 1983; Höllermann, 1983) assumes that snow cover averaging less than 50 cm can be ignored for purposes of permafrost forecasting. Use of such an arbitrary cutpoint is misleading because it ignores snow density and hence the snow's thermal properties. Equally important, the approach precludes mapping of permafrost distribution, which is one of the primary applications of permafrost prediction.

Nelson and Outcalt (1983) introduced a method of computing average winter snow thickness that weights snowfall by its duration (snow falling early in winter receives greater weighting than does snow falling in spring) and multiplies the result by a constant representing periods of thaw, sublimation, etc. This arbitrary constant can be replaced with a simple function representing latitudinal variations in the magnitude and duration of winter thaws. The formulation used here is

$$\bar{Z}_s = \sin^2\phi \,\{\sum_{i=1}^{k} [(P_i/\rho_r)(k - (i - 1))]/k\} \qquad (4)$$

where $\bar{Z}_s$ is average winter snow thickness, P_i is water equivalent precipitation in those months i ($i = 1, 2, \ldots, k$) in which the mean temperature is $\leq 0°C$, ρ_r is relative (dimensionless) snow density at the site, and ϕ is the site's latitude. The trigonometric function implies that little snow is lost at high-latitude stations, while sites at low latitudes may lose most of their snow cover as a result of more frequent or extended periods of thaw.

Actual average monthly snowfall totals (Canada, 1982b) were employed rather than the term (P_i/ρ_r) for the Canadian data used in Nelson's (1986, 1987) applications of the frost number. A value for snow density is still required, however, in order to estimate the thermal conductivity of the snow at each station. The average snow density at North American sites can be approximated by the regression equation

$$\rho_s = 152 - 0.31\,\bar{T}_w + 1.9\,\bar{U}_w \qquad (5)$$

reported by Bilello (1969), in which ρ_s is the snow's density in kg m^{-3} and $\bar{U}_w$ is mean winter wind speed (m s^{-1}).

Given values of ρ_s, thermal conductivity can be estimated by a variety of theoretical or regression equations. For the range of densities obtained from equation (5) for stations in Canada, the formulation of Van Dusen (Paterson, 1981: 186) provides mid-range values compared to other methods (Mellor, 1977). Van Dusen's formula for snow thermal conductivity λ_s (W m^{-1} $°C^{-1}$) is given by

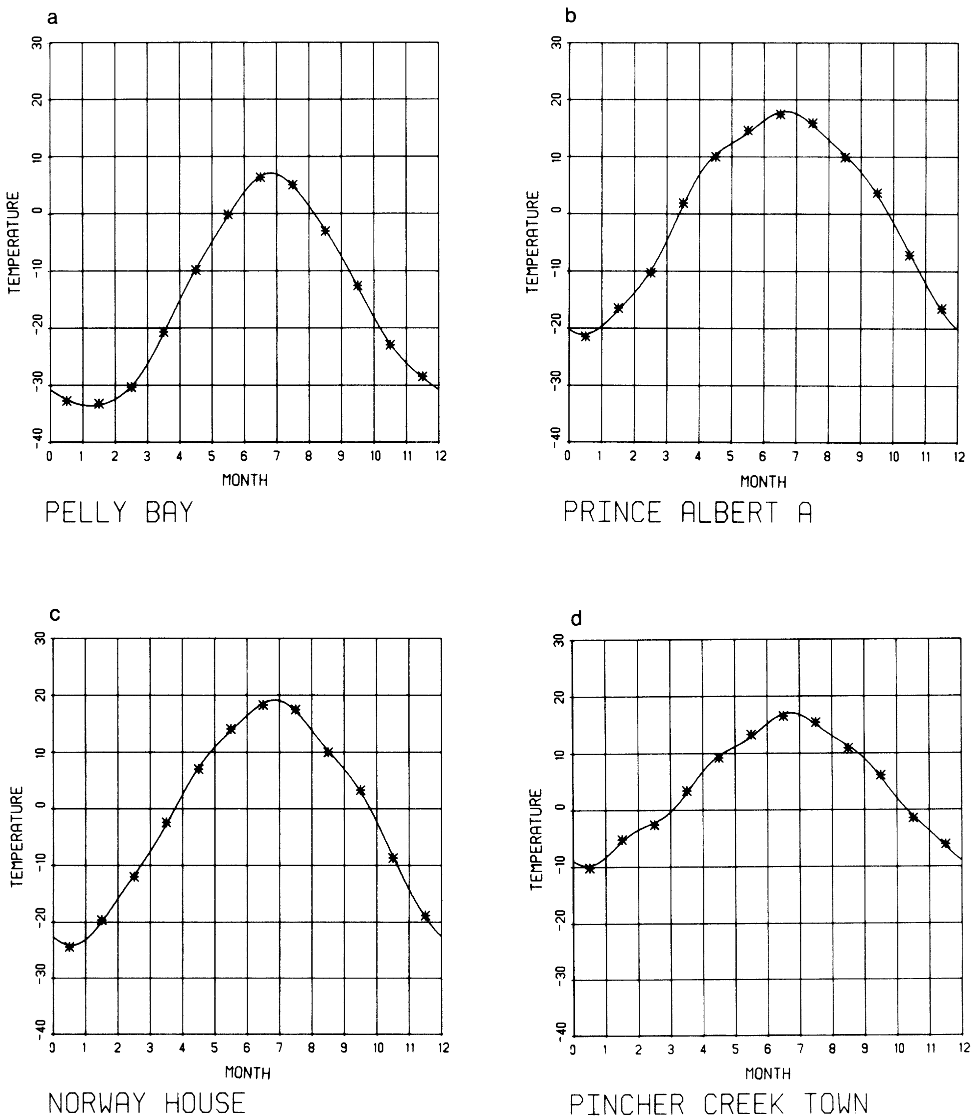

FIGURE 2. Temperature curves obtained through interpolation (solid lines) from mean monthly air temperature values (stars) for four Canadian stations illustrating error sources described in text.

$$\lambda_s = 2.1 \times 10^{-2} + 4.2 \times 10^{-4}\ \rho_s + 2.2 \times 10^{-9}\ \rho_s^3. \quad (6)$$

Thermal diffusivity is obtained by dividing the snow's thermal conductivity by its volumetric heat capacity C_s (J m^{-3} °C^{-1}), which in turn is the product of its density ρ_s (kg m^{-3}) and its specific heat capacity c_s (J kg^{-1} C^{-1}). The latter quantity can be estimated for a snow cover (Langham, 1981) from

$$c_s = 7.79\ \bar{T}_w + 2115. \quad (7)$$

An estimate of a snow cover's thermal diffusivity α_s can therefore be obtained from

$$\alpha_s = \lambda_s/(c_s \cdot \rho_s), \quad (8)$$

which is used below in equation (10) to compute the reduced freezing index at the ground surface in equation (12).

Nelson and Outcalt (1983) used the symmetry of the min/max sinusoidal temperature approximation to achieve a simple solution for the reduction of *DDF* attributable to the effects of snow cover on conductive heat transfer. The temperature amplitude A_+ at the surface with snow can be found by

$$A_+ = A\ \exp(-\bar{Z}_s/Z_{*s}), \quad (9)$$

where $\bar{Z}_s$ is the "average snow-cover thickness" from equation (4) and Z_{*s} is the damping depth in the snow, given by

$$Z_{*s} = (\alpha_s\ P/\pi)^{1/2}, \quad (10)$$

in which P is the length of the annual temperature cycle. A mean winter surface temperature $\bar{T}_{w+}$ incorporating snow-cover effects can then be computed by

$$\bar{T}_{w+} = \bar{T} - A_+\ [\sin \beta/(\pi - \beta)], \quad (11)$$

the freezing index at the surface DDF_+ is found by

$$DDF_+ = \bar{T}_{w+} \cdot L_w, \quad (12)$$

and the mean annual temperature $\bar{T}_+$ at the surface is obtained from

$$\bar{T}_+ = (DDT - DDF_+)/365. \quad (13)$$

Finally, the "surface frost number" F_+, the subscript (+) representing an adjustment for the effects of snowcover, is defined as

$$F_+ = \frac{DDF_+^{1/2}}{DDF_+^{1/2} + DDT^{1/2}}. \quad (14)$$

In situations where a simple sinusoid does not represent the temperature curve adequately (which are sufficiently common to warrant use of the spline interpolation procedure whenever monthly temperature records are available), the mean winter temperature can be found by computing the average daily air temperature below 0°C. The mean winter temperature $\bar{T}_{w+}$ with snow cover is then found by solving equation (2.5) for the quantity $\{\sin \beta/(\pi - \beta)\}$ and substituting this value in equation (11), a procedure which, in effect, compares the observed temperature curve with a perfect sinusoid of similar amplitude and which encloses the same area below the 0°C axis.

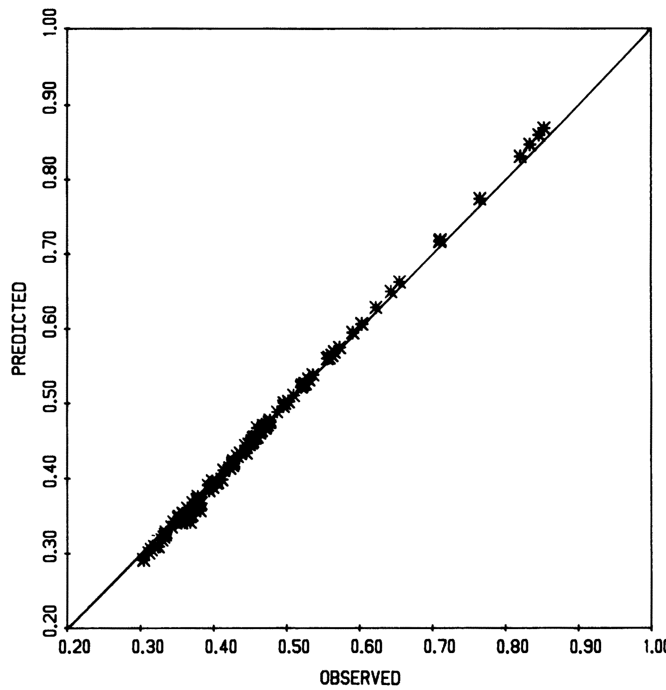

FIGURE 3. Air frost number based on observed degree days versus frost number computed by spline interpolation method. Index of agreement (Willmott, 1984) for these data is 0.999. Stations with fewer than 500 freezing degree-days not included.

The representation of snow-cover thickness as a smooth field is only compatible with small geographical scales, in which topographic heterogeneity is effectively ignored. Equation (4) assumes only that snow-cover thickness is a simple function of snowfall and latitude; smooth interpolation between stations ignores redistribution of snow on the ground.

Mapping the results of equation (14) for a series of stations defines a latitudinal zonation of what Heginbottom (1984) has termed "climatic permafrost," permafrost that is predicted by consideration of climate data alone. The interpretation given to this index is elaborated in a later section. If data on subsurface conditions are available for a region of interest, the predictive capacity of the frost number can be improved by employing the "Stefan frost number" discussed in the next section.

STEFAN FROST NUMBER

The Stefan solution (Jumikis, 1977: chapt. 14–15; Lunardini, 1981: chapt. 10) for the depth of frost penetration has the form

$$Z_{f+} = \left[\frac{2\lambda_f\, S|\bar{T}_{w+}|L_w}{\rho_d\, w_f\, L} \right]^{1/2} \quad (15)$$

where Z_{f+} is the depth (m) to which frost extends, λ_f is the frozen soil's thermal conductivity (W m^{-1} °C^{-1}). $\bar{T}_{w+}$ is the mean winter temperature at the surface, L_w is the duration (days) of winter, S is a scale factor (s d^{-1} = 86,400), ρ_d is the dry density (kg m^{-3}) of the soil, w_f is the soil's water content (proportion of dry weight), and L is the latent heat of fusion (J kg^{-1}). By substituting the subscripts t and s for f and w, the equation can be applied to thawing soils. Because temperature measurements are not made routinely at the ground surface, the surface temperature is often assumed to be that of the air temperature recorded at screen level, although the freezing or thawing index is sometimes multiplied by an n factor representing the surface to air degree-day ratio (Lunardini, 1981: 559–575). Because n factors are dependent on a large number of interrelated variables and reflect several modes of heat transfer, tabled values exist for only a few types of surfaces, most of which are roads.

The Stefan solution and its variations provide reasonable estimates of frost and thaw depths in a wide variety of situations (e.g., McRoberts, 1975) and are used extensively as first approximations. The Stefan solution appears to be well suited for mapping applications, in which some detail is often sacrificed in order to examine conditions at a large number of sites. Although equation (15) yields reasonably good estimates of freezing and thawing depths for soils in which the thermal properties do not vary greatly with depth, its applicability is limited because soils are usually more complex. The Stefan solution can, however, be extended to multilayer soil systems through introduction of a thermal resistance for each of the n–1 layers overlying the bottom (nth) layer. The thermal resistance R_i of each layer is given by

$$R_i = Z_i/\lambda_i \quad (16.1)$$

where Z_i and λ_i are the thickness and thermal conductivity of the ith layer ($i = 1,2, \ldots ,n$), respectively. The depth of frost (thaw) penetration into the bottom layer is given by

$$Z_n = \left[\left(\frac{2\lambda_n S\, DD_n}{Q_{L,n}} \right) + \lambda_n^2 \left(\sum_{i=1}^{n-1} R_i \right)^2 \right]^{1/2} \quad (16.2)$$

$$-\lambda_n \sum_{i=1}^{n-1} R_i$$

where DD_n represents the partial freezing (thawing) index remaining after freezing (thawing) the n–1 overlying layers, and $Q_{L,n}$ is the nth layer's volumetric latent heat, defined by the denominator of the bracketed quantity in equation (15). That part of the freezing (thawing) index expended in freezing (thawing) the first layer is estimated by

$$DD_1 = \frac{Q_{L,n} \cdot Z_1}{S} \cdot \frac{R_1}{2} \quad (16.3)$$

and for the (n–1)th layer by

$$DD_{n-1} = \frac{Q_{L,n-1} \cdot Z_{n-1}}{S} \left[\sum_{i=1}^{n-2} R_i + \frac{R_{n-1}}{2} \right]. \quad (16.4)$$

The total depth of frost (thaw) penetration in the layered system can therefore be obtained from

$$Z_{tot} = Z_1 + Z_2 + Z_3 + \ldots + Z_n. \quad (16.5)$$

A full treatment of the multilayer Stefan formula is given in Jumikis (1977: 209–224). By applying the surface freezing index DDF_+ in equation (15) or (16), a considerable amount of information is utilized, and a good estimate of the actual depth of frost at a given location can be obtained.

Examination of equation (15) or equation (16) and tabled values of soil thermal properties (Harlan and Nixon, 1978) demonstrates that the depth of frost or thaw is influenced strongly by the thermal properties and moisture content of the soil. Given a location with a mean annual surface temperature of 0°C ($DDF_+ = DDT$) and assuming that the moisture content of the soil is constant, the depth of freezing may exceed that of thaw, owing to the soil's higher thermal conductivity in the frozen state. Repetition of this situation on an annual basis for two or more years can result in the formation and growth of permafrost. It follows from this that the accuracy of the surface frost number F_+ as a predictor of permafrost occurrence can be improved by incorporating subsurface information through the Stefan solution. The "Stefan frost number," which considers the effects of both snow cover and soil properties, is defined by

$$F_{s+} = \frac{Z_{f+}}{Z_{f+} + Z_t} \quad (17)$$

where the subscript s indicates incorporation of the Stefan estimates of Z_f and Z_t. Although equations (15) and (16) overestimate the depth of frost and thaw by neglecting the soil's heat capacity, these errors tend to cancel when the ratio is formed in equation (17). The structure of equation (17) also allows employment of more sophisticated solutions for frost and thaw depths if data availability warrants their use.

INTERPRETATION AND MAPPING

In general, North American workers have failed to establish explicit definitions on which to base cartographic representations of permafrost zones. Numerous manifestations of this vagueness are apparent in the published literature.

Nelson (1987) traced historical developments in the cartography of permafrost zones in eastern Canada; even maps published virtually simultaneously have depicted highly divergent zonal boundaries. If the maps of Canada and Alaska presented by R. J. E. Brown and Péwé (1973: 73–74) were joined at a common scale, the boundaries of their permafrost zones would be offset by at least 50 km at the Yukon-Alaska border, apparently because different criteria were used to define the zones in Canada and the United States (R. J. E. Brown and Péwé, 1973: 75). Washburn (1980: 27–33) drew attention to the fact that some workers use the -1°C mean annual air isotherm to approximate the position of the "southern limit of discontinuous permafrost in North America," while others employ the 0°C isotherm for the same purpose. Numerous other examples of such inconsistencies could be cited; all indicate the need for an unambiguous definition of zonal permafrost boundaries. Such "boundaries" are in fact cartographic generalizations of situations in which changes are *transitional* rather than abrupt; explicit rules for placement of these symbols are therefore necessary if meaningful comparisons of permafrost distribution are to be made between different regions, countries, continents, or time periods.

It is important to bear in mind that regionalization is a form of classification, and should act as an aid for comprehension of some aspect of the complexity existing in nature. Classifications should not exist for their own sake; the value of any classification is dependent on how well it realizes the objective for which it was devised. The purported objective of most small-scale North American permafrost maps is to separate areas in which the land surface is underlain continuously by permafrost from those in which permafrost is discontinuous and those in which it is absent. This is obviously a formidable task if approached from an empirical standpoint. It therefore appears appropriate to base such delineations on factors which can be related to physical mechanisms responsible for the development of permafrost. Air temperature isarithms are inadequate because numerous empirical investigations have demonstrated that zones of permafrost continuity are not conformal with them. The frost number constitutes an improvement over previous methods because it uses explicit definitions to estimate several parameters that determine average frost and thaw penetration depths in soils. When the depth of frost penetration exceeds that to which thaw extends, permafrost growth will occur (e.g., Lunardini, 1981: 553–555). The remainder of this paper discusses how this conceptualization can be related to the continuity of permafrost.

The redefinitions given the frost number in equations (3), (14), and (17) necessitate revision of the numerical values assigned for zonal boundaries by Nelson and Outcalt (1983). Potential values of the revised frost number range from 0.0 (no freezing) to 1.0 (no thawing). The implications of these values for the existence of permafrost are obvious.

We follow the rationale of Nelson and Outcalt (1983) for using the surface frost number to place the boundary separating permafrost-free terrain from areas with the climatic potential to be underlain by permafrost. Perennially frozen ground is climatically possible poleward of locations where the mean annual temperature at the ground surface is 0°C, that is, locations for which our approximate methods yield values of DDF_+ and DDT that are equal. Equation (14) yields a value for F_+ of 0.5 under these conditions, and this value is taken as the "permafrost limit." Incorporation of snow-cover effects insures that this boundary will always lie poleward of the 0°C mean annual air isotherm. Between the permafrost limit and the boundary of the continuous zone, details of permafrost distribution are controlled by edaphic, geological, hydrological, topoclimatic, and botanical factors. The frost number cannot, however, predict the existence of relict permafrost. Only "contemporary permafrost" compatible with climate data used in the analysis can be predicted.

Although the concept of continuous permafrost is both intuitive and frequently referenced, an unambiguous definition and a satisfactory method for its delineation have long eluded researchers. The methods that have been suggested or employed in North America include drawing the continuous/discontinuous permafrost boundary to coincide with the -5°C ground temperature isotherm at the level of zero annual amplitude (e.g., R. J. E. Brown and Péwé, 1973) and incorporating only regions in which an arbitrary percentage of the surface is underlain by permafrost (e.g., Harris, 1986). Both of these approaches were borrowed from earlier Soviet literature (e.g., Sumgin et al., 1940; Kudryavtsev et al., 1980); the former is not consistent with the spatial connotation of the term "continuous permafrost" (Nelson, 1986), and both suffer problems of implementation stemming from data unavailability. In North America, empirical data with a spatial resolution sufficiently fine to document permafrost "continuity" are available only in a few relatively small areas, so that an empirical demonstration of the concept is possible only at large geographical scales.

The surface frost number F_+ provides a vehicle well suited for establishing a more objective definition for a continuous/discontinuous permafrost "boundary" than has been advanced previously. The position of this boundary can be computed directly from standard climate data without resort to subsurface investigations. Although computed from climatic data, the definition is based on physical principles, is broadly consistent with semi-empirical delineations, is not constrained by the spatial pattern of air temperature isarithms, and can be replicated by different workers employing the same data set. Equally important, it can be extended for use with paleoclimatic estimates or used in conjunction with climate models to examine the effects of climatic change or modifications on permafrost distribution, a task that cannot be performed if permafrost zonation is defined on the basis of field survey or remotely sensed data.

W. G. Brown (1964) reported an extensive set of calculations for frost penetration depths using the Neumann solution (a rigorous analytical solution to the general Stefan problem) and a wide variety of soils data. Brown

(1964: 225) concluded that "by considering the possible extremes of soil types, densities, and moisture content it is possible to show that frost penetration variations of less than 2 to 1 are obtained for the same freezing index." Brown also suggested that similar results could be expected from corresponding calculations for thaw depths. These results allow a definition for the position of the discontinuous/continuous permafrost zonal boundary that is both natural and objective, and can be reproduced by different workers using the same data set.

Equation (15), Stefan's solution for the depth of frost or thaw, can be rewritten

$$Z = \left[\frac{2\,\lambda}{\rho_d\; w\; L}\right]^{1/2} \cdot [S \cdot DD]^{1/2} = E \cdot C \qquad (18)$$

where E and C represent an "edaphic term" and a "climatic term," respectively. At any location where the ratio $DDF_+/DDT = 4.0$, for example, where the freezing and thawing indices are 4900 and 1225°C days, respectively, the surface frost number is

$$F_+ = \frac{C_{f+}}{C_{f+} + C_t} = \frac{2}{2+1} = 0.666 \qquad (19)$$

so that relative values of the frost penetration depth Z_{f+} and thaw penetration depth Z_t for this climatic situation can be expressed as

$$Z_{f+} = 2E_f \qquad (20.1)$$

and

$$Z_t = 1E_t. \qquad (20.2)$$

Restating W. G. Brown's (1964) experimental results in a form consistent with equations (18) and (20), the edaphic term will not vary by more than a factor of two for the range of mineral soils and moisture contents found in nature. A "worst case scenario" is therefore suggested as a soil in which relative values of the edaphic term are $E_f = 1$ and $E_t = 2$. Inserting these values in equations (20.1) and (20.2), respectively, we obtain

$$Z_{f+} = (2)\,(1) = 2 \qquad (21.1)$$

and

$$Z_t = (1)\,(2) = 2. \qquad (21.2)$$

From the definition of the Stefan frost number (equation 17), these numbers yield

$$F_{s+} = \frac{Z_{f+}}{Z_{f+} + Z_t} = \frac{2}{2+2} = 0.5 \qquad (22)$$

which is the lowest F_{s+} value consistent with the existence of permafrost. Continuous permafrost can therefore be expected poleward of locations at which $F_+ = 0.666$. Such locations can be interpreted as lying on the discontinuous/continuous permafrost boundary, since values of F_+ smaller than 0.666 can in principle yield values of F_{s+} less than 0.5 for some soils. An isoline connecting locations at which $F_+ \cong 0.67$ therefore provides a more objective definition for the position of the equatorward boundary of continuous permafrost than has been advanced previously.

The strategy outlined above provides a reasonable solution to the problem of defining "boundaries" of permafrost continuity at small geographical scales. Such delineations will usually be quasi-latitudinal zonations, that is, in regions of low relief their cartographic expression will be a zone of continuous permafrost lying poleward of a zone of discontinuous permafrost and one of permafrost-free terrain. Examples are given by Nelson (1986, 1987).

Definitions of continuous permafrost outside the context of small-scale geographic zonality are also possible. Harris (1986: 30) used the seemingly contradictory phrase "patches of continuous permafrost." Implicit in this phrase is the existence of a hierarchy of scale-dependent "continuity"; permafrost may underlie some areally extensive terrain unit continuously although it may be absent in surrounding units less conducive to its formation or maintenance. In such a case the aggregate region would lie within the "discontinuous zone" on a map depicting permafrost distribution in the continent containing that region. The "continuous" permafrost body in this instance will owe its existence to some edaphic, geologic, hydrologic, topoclimatic, or vegetative characteristic; for example, it may be an extensive peatland. Such a situation is depicted in Figure 4, which makes the

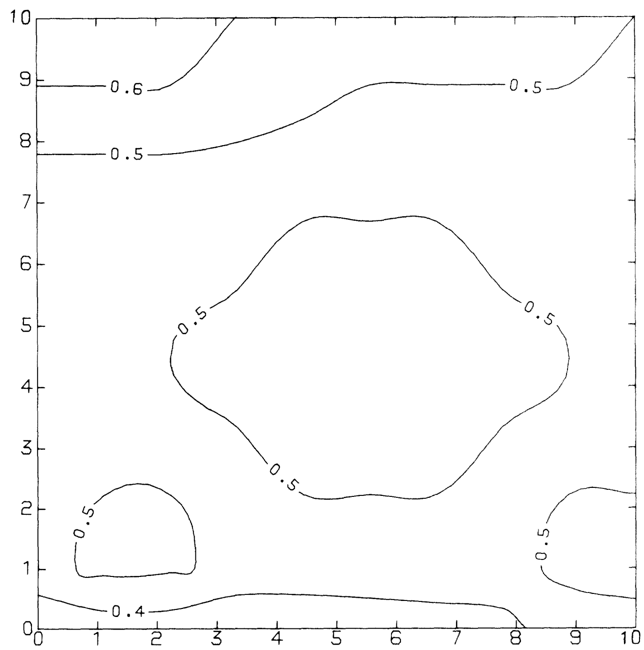

FIGURE 4. Illustration of scale effects on the concept of permafrost continuity. Isarithms represent Stefan frost number F_{s+}.

scale dependence of the term "continuous permafrost" explicit. On a small-scale map in which values along the axes represent hundreds of kilometers, the closed 0.5 isarithms would be considered permafrost "islands" within the discontinuous zone. At a larger scale in which the axis values represent tens of kilometers, the closed contours could be regarded as "patches of continuous permafrost," a concept that could possibly be useful in an applied context.

The definition of the Stefan frost number and the availability of edaphic information on documents such as the map *Soils of Canada* (Canada, 1972) allows an investigator to partially transcend scale limitations imposed by widely scattered climate stations. The fixed locations and finite number of climate stations severely limit the scale at which we can examine permafrost distribution using only climate data. The smooth interpolation required to produce maps based on such data is not consistent with reality, and can be regarded as analogous to a low-pass spatial filter. By using a grid to sample soil types on the soils map and employing estimates of the thermal properties associated with the various soil types, the investigator can treat relatively high-frequency (large-scale) variations in permafrost conditions. Raster (cellular) or choropleth graphic techniques permit the question of permafrost continuity to be treated as a contiguity problem. Cells in which F_{s+} is equal to or greater than 0.5 are permafrost by definition; a series of contiguous cells meeting this criterion constitutes a "patch of continuous permafrost." Nelson (1986) provided several cartographic strategies for conveying such information. It is, however, important to realize that this approach does not alleviate the scale problem entirely. "Continuity" of permafrost is still dependent on map scale, cell size, and the degree to which the assumptions of terrain and vegetative homogeneity are warranted. Although problems arising from the spacing of climate stations are eased somewhat by the use of soils information, they continue to exert a strong influence on the scale at which the frost number yields meaningful results. Use of the frost number for mapping areas encompassing less than about 500,000 km² is not recommended unless a region contains an unusually large number of weather stations.

CONCLUSIONS

The frost number provides a vehicle for producing unambiguous small-scale maps of permafrost continuity. The surface frost number produces a quasi-latitudinal zonation of contemporary permafrost and can be computed from minimal climatic data. The Stefan frost number incorporates subsurface information and improves the spatial resolution that can be achieved with the index. The frost-number methodology constitutes an improvement over previous computational methods of forecasting permafrost distribution because it relates permafrost to several factors responsible for its development and allows explicit definitions of zonal boundaries.

Some workers have advocated classifications based on percentages of the land surface underlain by permafrost. Besides suffering severe problems of implementation stemming from data unavailability and definition of the areal units over which percentages are computed, such schemes are undesirable because the boundaries chosen to represent classes of "continuity" are completely arbitrary. Perhaps even more important is the fact that such classifications are fixed in time, that is, their application is limited to delineation of present-day permafrost distribution. Because the most basic form of the frost number requires only the mean temperatures of the warmest and coldest months and an estimate of snowcover thickness, its application can be extended for use with results from general circulation models (e.g., Kutzbach and Wright, 1985; Stuart, 1985), allowing investigators to explore patterns of permafrost distribution existing during glacial maxima and minima or conditions postulated under various climate-change scenarios.

ACKNOWLEDGMENTS

Research leading to this paper was supported by the Division of Polar Programs, U.S. National Science Foundation, Grant DPP-8117124. We thank R. K. Haugen, J. F. Nixon, and K. A. Salzberg for critical comments leading to numerous improvements in the manuscript.

REFERENCES CITED

Billelo, M. A., 1969: Relationships between climate and regional variations in snow-cover density in North America. *U.S. Army Cold Regions Research and Engineering Laboratory, Research Report,* 267. 21 pp.

Brown, R. J. E., 1967: *Permafrost in Canada.* Ottawa: Geological Survey of Canada Map 1246A and National Research Council of Canada, Division of Building Research Map NRC-9769.

Brown, R. J. E. and Péwé, T. L., 1973: Distribution of permafrost in North America and its relationship to the environment: A review. *In: Permafrost—North American Contribution to the Second International Conference.* Washington, D.C.: National Academy of Sciences, 71–100.

Brown, W. G., 1964: Difficulties associated with predicting the depth of freeze or thaw. *Canadian Geotechnical Journal,* 1: 215–226.

Canada, 1972: *Soils of Canada* (map). Ottawa: Department of Energy, Mines and Resources.
———, 1982a: *Canadian Climate Normals. Vol. 2. Temperature, 1951–1980.* Toronto: Atmospheric Environment Service. 305 pp.
———, 1982b: *Canadian Climate Normals. Vol. 3. Precipitation, 1951–1980.* Toronto: Atmospheric Environment Service. 602 pp.
———, 1982c: *Canadian Climate Normals. Vol. 4. Degree Days, 1951–1980.* Toronto: Atmospheric Environment Service. 280 pp.
Ferrians, O. J., Jr., 1965: *Permafrost Map of Alaska.* U.S. Geological Survey Miscellaneous Geological Investigation Map I-445.
Goodrich, L. E., 1982: The influence of snow cover on the ground thermal regime. *Canadian Geotechnical Journal,* 19: 421–432.
Granberg, H. B., 1973: Indirect mapping of the snowcover for permafrost prediction at Schefferville, Quebec. *In: Permafrost—North American Contribution to the Second International Conference.* Washington, D.C.: National Academy of Sciences, 113–120.
Greenstein, L. A., 1983: An investigation of midlatitude alpine permafrost on Niwot Ridge, Colorado Rocky Mountains, U.S.A. *In: Permafrost—North American Contribution to the Second International Conference.* Washington, D.C.: National Academy of Sciences, 380–383.
Harlan, R. L. and Nixon, J. F., 1978: Ground thermal regime. *In:* Andersland, O. B. and Anderson, D. M. (eds.), *Geotechnical Engineering for Cold Regions.* New York: McGraw-Hill, 103–163.
Harris, S. A., 1981: Climatic relationships of permafrost zones in areas of low winter snow-cover. *Arctic,* 34: 64–70.
———, 1986: Permafrost distribution, zonation and stability along the eastern ranges of the cordillera of North America. *Arctic,* 39: 29–38.
Haugen, R. K., Outcalt, S. I., and Harle, J. C., 1983: Relationships between estimated mean annual air and permafrost temperatures in North-Central Alaska. *In: Permafrost—Fourth International Conference Proceedings,* vol. I. Washington, D.C.: National Academy of Sciences, 462–467.
Heginbottom, J. A., 1984: The mapping of permafrost. *Canadian Geographer,* 28: 78–83.
Höllermann, P., 1983: Blockgletscher als Mesoformen der Periglazialstufe. Studien aus europäischen und nordamerikanischen Hochgebirgen. *Bonner Geographische Abhandlungen,* 67: 1–83.
Jumikis, A. R., 1977: *Thermal Geotechnics.* New Brunswick, N.J.: Rutgers University Press. 375 pp.
King, J. T., 1984: *Introduction to Numerical Computation.* New York: McGraw-Hill. 355 pp.
Kudryavtsev, V. A., 1977: Fundamentals of frost forecasting in geological engineering investigations. *U.S. Army Cold Regions Research and Engineering Laboratory, Draft Translation,* 606. 489 pp.
Kudryavtsev, V. A., Kondrat'eva, K. A., and Romanovskii, N. N., 1980: Zonal and regional patterns of the permafrost region in the U.S.S.R. *In: Proceedings of the Third International Conference on Permafrost,* English Translations, Part I. Ottawa: National Research Council of Canada, 371–389.
Kutzbach, J. E. and Wright, H. E., Jr., 1985: Simulation of the climate of 18,000 years BP: Results for the North American/North Atlantic/European sector and comparison with the geologic record of North America. *Quaternary Science Reviews,* 4: 147–187.
Lachenbruch, A. H., 1959: Periodic heat flow in a stratified medium with application to permafrost problems. *U.S. Geological Survey Bulletin* 1083-A: 1–36.
Langham, E. J., 1981: Physics and properties of snowcover. *In* Gray, D. M. and Male, D. H. (eds.), *Handbook of Snow: Principles, Processes, Management & Use.* Toronto: Pergamon Press, 275–337.
Lunardini, V. J., 1981: *Heat Transfer in Cold Climates.* New York: Van Nostrand Reinhold. 731 pp.
McRoberts, E. C., 1975: Field observations of thawing in soils. *Canadian Geotechnical Journal,* 12: 126–130.
Mackay, J. R. and MacKay, D. K., 1974: Snow cover and ground temperatures, Garry Island, N.W.T. *Arctic,* 27: 287–296.
Mellor, M., 1977: Engineering properties of snow. *Journal of Glaciology,* 19(81): 15–66.
Nelson, F. E., 1986: Permafrost in central Canada: Applications of a climate-based predictive model. *Annals of the Association of American Geographers,* 76: 550–569.
———, 1987: Permafrost zonation in the Hudson Bay region: Review and applications of a predictive model. Submitted for publication.
Nelson, F. and Outcalt, S. I., 1983: A frost index number for spatial prediction of ground-frost zones. *In: Permafrost—Fourth International Conference Proceedings,* vol. I. Washington, D.C.: National Academy Press, 907–911.
Panofsky, H. A. and Brier, G. W., 1968: *Some Applications of Statistics to Meteorology.* University Park: Pennsylvania State University Press. 224 pp.
Paterson, W. S. B., 1981: *The Physics of Glaciers.* New York: Pergamon Press. 380 pp.
Pihlainen, J. A., 1962: An approximation of probable permafrost occurrence. *Arctic,* 15: 151–154.
Smith, M. W., 1975: Microclimatic influences on ground temperatures and permafrost distribution, Mackenzie Delta, Northwest Territories. *Canadian Journal of Earth Sciences,* 12: 1421–1438.
Stuart, A., 1985: The development of a spatial permafrost model for northern Canada and its application to scenarios of climate change. Final Report to Atmospheric Environment Service, Contract Number 02SE.K 147-5-0313. Downsview, Ontario: KelResearch Corporation.
Sumgin, M. I., Kachurin, S. P., Tolstikhin, N. I., and Tumel', V. F., 1940: *Obshchee merzlotovedenie* [General Permafrostology]. Moscow-Leningrad: Akademiia Nauk SSR. 240 pp.
Washburn, A. L., 1980: *Geocryology: A Survey of Periglacial Processes and Environments.* New York: Halsted Press. 406 pp.
Willmott, C. J., 1984: On the evaluation of model performance in physical geography. *In* Gaile, G. L. and Willmott, C. J. (eds.), *Spatial Statistics and Models.* Boston: D. Reidel, 443–460.

Ms submitted October 1986