A new approach for assessing geothermal response to climate change in permafrost regions

S. NISHIMURA*, C. J. MARTIN†, R. J. JARDINE† and C. H. FENTON†

Infrastructure in cold regions is vulnerable to the potential degradation of permafrost under a warming climate. Meanwhile, the accumulation of meteorological data and the refinement of AOGCMs in recent years have improved the confidence in future global air temperature predictions. A reliable scheme is now desired that converts these climate predictions into future geocryological predictions relevant to geotechnical engineering and risk assessment. This paper describes a multidisciplinary approach that provides a first estimate of transient ground responses to climate change on a regional basis. The scheme integrates locally adjusted AOGCM climate predictions, regional geological assessments, non-linear thermal finite element analysis and digital elevation models derived from remote sensing data. The practical application of the approach is demonstrated through predictions made of the geocryological changes expected between 1940 and 2059 in a Siberian region. The paper presents results from one sampled area where discontinuous permafrost is present beneath rolling hills terrain. It is shown that elevation, vegetation and local geological variations all affect the development of permafrost, with important implications for infrastructure design and operation. A range of useful geocryological maps can be output from the procedure, including temperature at the active layer base, permafrost table depth, and ground temperature at any desired depth. It is shown that the permafrost model's predictions for present-day conditions agree well with existing geocryological maps. An illustrative example of how simple geohazard maps may be prepared from the output is also provided.

KEYWORDS: geology; ground freezing; numerical modelling; temperature effects

L'infrastructure dans des régions froides est vulnérable à la dégradation potentielle du pergélisol, sous l'effet du réchauffement du climat. En outre, l'augmentation des données météorologiques et le perfectionnement des « AOGCM » au cours des dernières années, ont permis de renforcer l'assurance dans l'acquisition de connaissances sur l'évolution de la température globale de l'air dans l'avenir. On se penche maintenant sur un modèle fiable qui permettrait de convertir ces prévisions climatiques en prévisions géocryologiques futures pertinentes pour la géotechnique et l'évaluation des risques. La présente communication décrit une méthode pluridisciplinaire fournissant une première estimation des réactions transitoires du sol au changement climatique, à l'échelon régional. Ce principe intègre des prédictions climatiques AOGCM ajustées sur place, ainsi que des évaluations géologiques régionales, des analyses aux éléments finis thermiques non linéaires, et des modèles numériques d'élévation, dérivés de données relevées par télédétection. On démontre l'application pratique de cette méthode par le biais de prévisions des variations géocryologiques prévues entre 1940 et 2059, dans une région de la Sibérie. La communication présente les résultats d'une zone étudiée, dans laquelle le pergélisol est présent sous un terrain vallonné. On démontre que l'altitude, la végétation et les variations géologiques locales affectent toutes le développement du pergélisol, avec les implications importantes que cela comporte pour la conception de et l'exploitation de l'infrastructure. Cette procédure permet de produire une gamme de cartes géocryologiques utiles, y compris la température à la base de la couche active, la profondeur de la table du pergélisol, et la température du sol à la profondeur désirée. On démontre que les prédictions du modèle de pergélisol pour l'époque actuelle correspondent bien aux cartes géocryologiques existantes. En outre, on présente également une illustration de la façon dont ces informations permettent de tracer de simples cartes des géorisques.

INTRODUCTION

Future climate warming (IPCC, 2007) will result in above-average temperature rises in cold regions, where significant proportions of the land mass are underlain by permafrost. Facilities, dwellings and infrastructure sited in these areas are likely to be particularly vulnerable to a resulting set of geohazards associated with possible permafrost degradation. A prerequisite to rational risk-assessment, design and operating policy in these circumstances is the ability to estimate how ground temperature and permafrost conditions will vary over future decades.

Atmosphere ocean general circulation models (AOGCMs) are powerful tools for predicting global average air temperature trends, and have been developed for practical use in climate change studies (IPCC, 2007). However, it is difficult to make representative local predictions with AOGCMs because of the coarse spatial resolution that is currently incorporated (typically a few hundred kilometres square). As attempted in the recent study by Stendel *et al.* (2007), multi-stepped approaches would be necessary to downscale the GCM predictions to a more local level. In addition, most AOGCMs do not consider the freezing or thawing beneath the ground surface (Delisle, 2007), and the ground surface temperature is usually deduced from air temperature trends through simplified scaling laws. AOGCMs, in general, do not fully address the complex thermal and hydrological

interactions between the air and the ground, including thermal damping and response delays involved in thermal conduction, and may well overestimate ground warming rates and permafrost degradation. One recent AOGCM development does address freezing and thawing ground (Lawrence & Slater, 2005) by incorporating the uppermost 3·43 m of soil. However, this modest extension is insufficient when considering active layer thickening and permafrost degradation under inter-annual warming, and thermal interactions between permafrost and deeper ground (Nicolsky *et al.*, 2007). Geotechnical engineers are often concerned with the ground's future response over tens of metres of depth.

This paper describes a new approach for predicting regional and local geothermal responses to air temperature changes. The approach is one outcome from a multidisciplinary Imperial College project funded by BP to create a broad, physically and analytically based framework for climate change risk assessment in cold regions, as described by Clarke *et al.* (2008). The framework is hierarchic, starting with global climate change and regional geothermal predictions. This initial high-level assessment considers the potential vulnerability of the relevant areas and facilities, and identifies specific aspects that require more detailed assessment. In the most critical cases, fully coupled thermo-hydro-mechanical (THM) analyses may be employed as the best means to address geotechnical or soil–structure interaction outcomes. A companion paper (Nishimura *et al.*, 2009) proposes a new THM model developed as part of the integrated package.

We focus here on the high-level regional assessment tools, describing an approach that integrates AOGCM predictions, which have been locally adjusted using local meteorological data, regional geological assessments, non-linear thermal finite element analysis with digital elevation models (DEM) and surface vegetation assessments derived from remote sensing data. The application is illustrated by considering changes between 1940 and 2059 in a region around Lake Baikal in Siberia. We report detailed predictions from one sampled area, showing how the permafrost conditions vary with ground elevation, vegetation, snow cover and local stratigraphic conditions. The model predictions are visualised as a variety of geocryological maps. The predictions allow simple geohazard maps to be generated for high-level risk assessment.

EXISTING APPROACHES TO PERMAFROST DISTRIBUTION MODELLING

Models have been proposed for predicting permafrost distribution under given climatic and geological conditions. Following Hoelzle *et al.* (2001), these may be classified as being either process-oriented or empirical-statistical. The process-oriented models consider the inter-surface energy balance in terms of individual energy transfer processes, such as reflection and absorption of solar radiation, heat conduction and so on, developing explicit quantitative sub-models for each individual process (e.g. Hinzman *et al.*, 1998; Hoelzle *et al.*, 2001). The alternative empirical-statistical approaches consider a simpler system of indices and parameters to describe the local climatic and ground characteristics (e.g. Nelson & Outcalt, 1987; Smith & Riseborough, 1996; Nelson *et al.*, 1997). Common among them is the use of *n*-factors (Lunardini, 1978), which interrelate air and ground temperatures in an empirical way without considering individual energy transfer processes. This approach is both simple and broadly reliable, and has been applied widely in regional and local studies. For example, Klene *et al.* (2001a) combined the *n*-factor method with the Stefan solution to calculate the active layer thickness—that is, the depth over

which annual freeze/thaw cycles occur—based on their extensive dataset in the Kuparuk River Basin, Alaska. Smith & Riseborough (1996) present a model that also incorporates the changes in the thermal conductivity of frozen and unfrozen soil, and converts air temperature indices (defined later) into predictions for the TTOP (the mean annual ground temperature, MAGT, at the bottom of the seasonally frozen/unfrozen layer). The TTOP model has been applied to worldwide geographical settings, for example in Canada (Henry & Smith, 2001) and the Norwegian mountains (Juliussen & Humlum, 2007).

Common to both broad categories of physical geographical models is their restricted consideration of temperature changes below the active layer. When the ground thermal regime is in equilibrium with the climate, the TTOP may be a convenient index of permafrost state. Under a warming climate, however, delayed thermal responses of deeper ground, and their interactions with the active layer, make models based on such indices less accurate in describing the state of deep permafrost (e.g. Riseborough, 2007).

Thermal conduction in frozen ground is a complex non-linear phenomenon due to water phase changes and the generally heterogeneous, layered nature of real ground. Consideration of how transient temperature changes propagate to deeper ground is beyond the scope of the conventional models outlined above. More fundamental analysis with vertically extended physical domains, reaching depths of tens of metres, requires means such as finite difference or finite element numerical methods. One such study by Goodrich (1982) captured many of the important features involved in one-dimensional frozen ground thermodynamics, such as the annual development of snow cover, thermal property changes associated with water phase changes, and latent heat capture/release. Similar analyses have been performed by Kane *et al.* (1991) and Delisle (2007), among others.

Whereas numerical geothermal approaches are well suited to one-dimensional columns, their principal shortcoming in high-level regional thermal assessment exercises is the difficulty of considering the extensive horizontal domains. To be applicable regionally, the 'vertically oriented' numerical techniques have to be reconciled with the 'horizontally oriented' geographical approaches. One such attempt has been made by Hinzman *et al.* (1998), who combined an energy balance model with one-dimensional finite element analysis.

PROPOSED GEOTHERMAL MODELLING METHOD

The present study proposes a hybrid approach that combines rigorous 'vertical' numerical analysis with regional climatic, geological and geographical information. With application to regional-scale problems in mind, this analysis accepts coarsely defined bands for the key input parameters. A comprehensive series of parametric one-dimensional thermal analyses is run in advance, and the results are organised as a database. At each map point, climatic and geographical data are converted into values for the model parameters, and the parameter value set works as a query to seek the best possible matching analytical case and its results in the database. The climatic and geographical data are stored as AOGCM outputs, DEM and other remote sensing data, and the conversion criteria are derived from appropriate empirical relationships. The structure of this process is illustrated in Fig. 1. The parameters returned from the numerical database by the mapping routine may include the TTOP, the active layer thickness, or the depth of permafrost table at any given time.
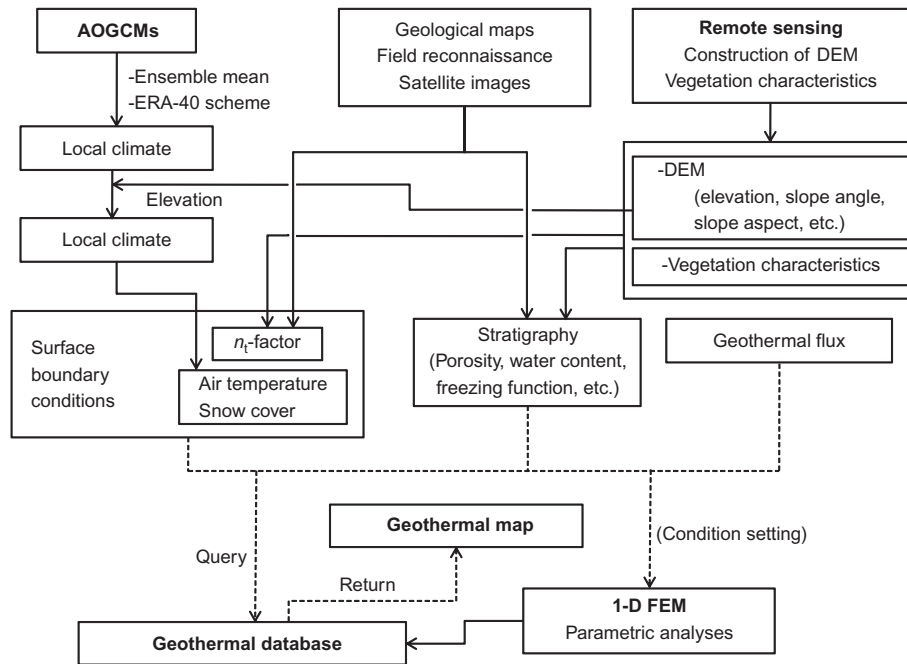
Fig. 1. Structure of proposed approach to high-level geothermal prediction

Formulation of finite element model (FEM) for thermal conduction

The FEM model developed for this study is based on non-linear heat conduction theory with temperature-dependent thermal properties. The main features are as follows.

(a) The thermal conductivity and heat capacity of soils vary with unfrozen water content, which is a function of temperature.
(b) Latent heat capture/release is considered.
(c) Vertically variable material properties can be considered.
(d) Monthly variations of insulating snow cover are modelled explicitly by adding and removing snow elements.
(e) Surface energy transfer at the air/ground surface interface is described using the *n*-factor. The effects of any gas phase in the pores are not considered at present.

The governing equation is

$$\frac{\partial u(T)}{\partial t} - \frac{\partial}{\partial x}\left[\lambda(T)\frac{\partial T}{\partial x}\right] - q = 0 \tag{1}$$

where u is the specific internal energy, T is the temperature in a thermodynamic scale, q is the heat production/sink, λ is the thermal conductivity, t is the time, and x is the depth. The internal energy and the thermal conductivity are both non-linear functions of temperature, owing to the phase change of water.

The internal energy term is expressed as

$$\frac{\partial u(T)}{\partial t} = \frac{\partial}{\partial t}[\{c_s\rho_s(1-\phi) + c_l\rho_l S_l(T)\phi \\ + c_i\rho_i[1 - S_l(T)]\phi\}T + l\rho_l S_l(T)\phi] \tag{2}$$

where c, ρ, l and ϕ are the specific heat capacity, the mass density, the specific latent heat of fusion and porosity respectively. Subscripts l, i and s indicate liquid water, ice and soil particles respectively. $S_l(T)$ is the degree of liquid saturation as a function of T, defined by volume as

$$S_l = \frac{V_l}{V_i + V_l} = \frac{V_l}{V - V_s} \tag{3}$$

The mathematical function adopted to describe S_l was selected as

$$S_l = \alpha\frac{\rho_s(1-\phi)}{\rho_l\phi}|T|^{-\beta} \tag{4}$$

for $T < 0°C$; otherwise $S_l = 1$.

In equation (4), α and β are material constants. Equation (4) has been widely used to describe unfrozen water content, for example by Anderson & Morgenstern (1973) and Andersland & Ladanyi (2004). The function is plotted in Fig. 2 for the parameters specified in Tables 1 and 2. The mass thermal conductivity λ of a soil containing solid, liquid and ice phases varies with S_l, and may be expressed as

$$\lambda(S_l) = \lambda_s^{1-\phi}\lambda_l^{S_l\phi}\lambda_i^{(1-S_l)\phi} \tag{5}$$

Variation of the thermal conductivity with degree of liquid saturation calculated from this function is plotted for several different soil porosities, ϕ, in Fig. 2. The model was implemented in a purpose-written code FEM FATALE. It features (a) automatic continuous processing of a series of multiple analyses, and (b) an algorithm to shortcut initial steady-state analyses required to obtain the 'equilibrium' geothermal conditions.

Boundary conditions in finite element model

The top finite element boundary condition is prescribed by temperature, which is determined by air temperature and surface energy transfer characteristics. The latter are influenced by factors including insolation, slope aspect, snow cover, vegetation and the presence of block fields (boulders shading the ground surface), infrastructure, surface water and peat layers. The numerical model adopts the simplified *n*-factor approach for surface energy transfer, which is defined separately for snow-free summer (n_t) and winter (n_f) periods as
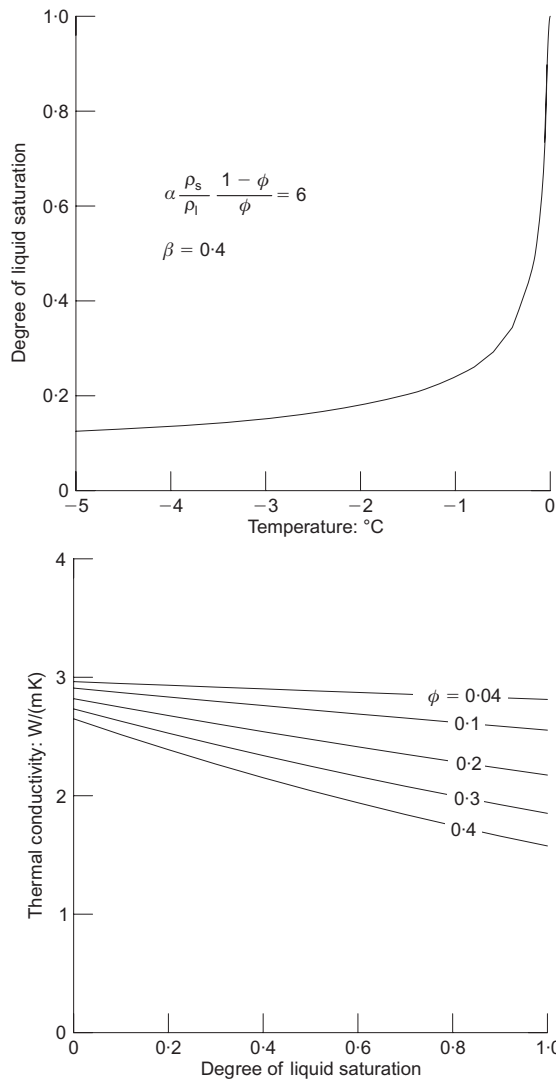
Fig. 2. Example of freezing characteristics and thermal conductivity functions

Table 1. Thermal properties of water used in analysis

Parameter	Value
Liquid water thermal conductivity, λ_l: W/(m K)	0·6
Ice thermal conductivity, λ_i: W/(m K)	2·2
Liquid water heat capacity, c_l: J/(kg K)	4190
Ice heat capacity, c_i: J/(kg K)	2095
Liquid water mass density: ρ_l: kg/m³	1000
Ice mass density: ρ_i: kg/m³	910
Specific latent heat of fusion, l: J/kg	334 000

Table 2. Thermal properties of snow and soils adopted in analysis

Parameter	Value	
	Snow	Soil mineral
Thermal conductivity, λ: W/(m K)	0·3	3
Heat capacity, c: J/(kg K)	2095*	650
Mass density, ρ: kg/m³	273*	2650
Porosity, ϕ	0*	See Fig. 9
Freezing curve parameters		
α	–	$6\rho_l\phi/\rho_s(1-\phi)$
β	–	0·4

* Thermally equivalent to ice–air mixture of $\phi = 0\cdot7$

$$I_{st} = n_t I_{at} \tag{6a}$$

and

$$I_{sf} = n_f I_{af} \tag{6b}$$

where

$$I_{at} = \int_{t_t} T_a \, dt \tag{7a}$$

and

$$I_{af} = -\int_{t_f} T_a \, dt \tag{7b}$$

where T_a is the air temperature (normally defined at 2 m above the ground surface) in degrees Celsius, and t_t and t_f are the periods during which $T_a > 0°C$ and $T_a < 0°C$ respectively. Indices I_{at} and I_{af} are the air thawing index and air freezing index respectively. The ground surface indices, I_{st} and I_{sf}, are defined likewise. The present model assumes that n_t applies monthly and remains constant. The model does not use n_f-factors to model snow cover because snow elements are explicitly included in the FEM. The snow top and air temperatures are assumed to be equal (Andersland & Ladanyi, 2004). The incident regional geothermal flux is prescribed as the base of the domain boundary condition. The FE ground model is illustrated in Fig. 3.

STUDY AREA
Geographical setting

A region of south-eastern Siberia was considered around Lake Baikal (Fig. 4(a)), where continuous and discontinuous permafrost are both widespread. The local geological, geographical and climatic setting was assessed through an extensive review of published data and a field reconnaissance. Within the study region of approximately 1300 km east–west by 1100 km north–south (Fig. 4(a)), ground and climatic conditions vary, mainly because of variations in relief. The topography ranges from high-relief mountains flanking the Baikal basin to low-lying fluvial plains towards Mongolia and along the Angara River. The geology varies
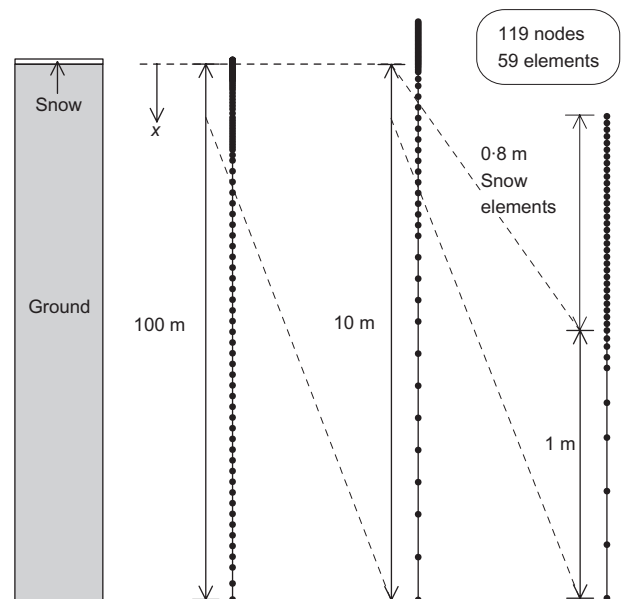

Fig. 3. Finite element model of one-dimensional ground column
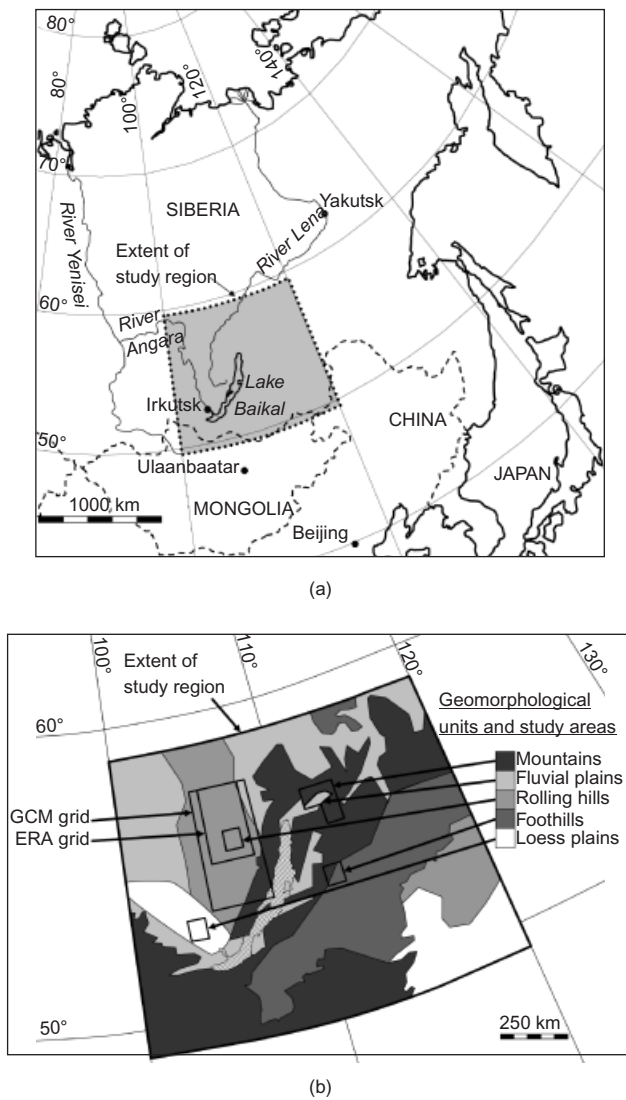
(a)

(b)

Fig. 4. Maps of the project study region: (a) location; (b) geomorphological zoning

from relatively flat-lying sedimentary rocks of the Siberian platform to much older, highly deformed and metamorphosed rocks along the active Baikal rift zone.

Five stereotypical geomorphological units were defined in the study region, which encompass broadly similar physical characteristics: mountains, foothills, rolling hills, fluvial plains and loess plains (Fig. 4(b)). Five 'exemplar' study areas were then selected (the smaller grids in Fig. 4(b)) to consider each setting, which provided representative settings for the associated climate and geothermal modelling and engineering geological terrain analysis. Best-estimate geological, geocryological and geotechnical characteristics were interpreted for each study area.

Ground conditions of the Rolling Hills study area

We only consider here detailed predictions from the Rolling Hills study area, which provides an example of the warm, generally discontinuous, permafrost regions that are most vulnerable to degradation resulting from air warming. The area measures approximately 65 km east–west by 75 km north–south. Its location is shown in Fig. 4(b), along with the larger cell to which the AOGCM outputs apply.

Remote sensing data (Shuttle Radar Topography Mission (SRTM) topographic survey data) with a spatial resolution of 70 m was used to produce a DEM, allowing detailed

assessment of elevation, slope angle, slope aspect and drainage characteristics within the study area (Fig. 5). The local geology and topography are summarised in the conceptual ground model presented in Fig. 6. Typical landscapes in this study area are shown in Figs 7 and 8. The study area is underlain by flat-lying, interbedded shallow marine sedimentary rocks of Cambrian and Ordovician age that are relatively unmetamorphosed. The area is characterised by the presence of extensive flat-topped plateaux (Fig. 8) capped by resistant sandstone/limestone beds and deeply incised river valleys. Colluvial deposits are present on the relatively shallow slopes; 90% of the study area has slope angles less than 16°. Permafrost is encountered mainly at higher relief (Yershov, 1991).

In the presented model, the soil properties are represented by porosity, soil mineral thermal conductivity, and freezing characteristics curve constants (i.e. α and β in equation (4)). Three simplified porosity profiles were selected to represent the range of ground conditions anticipated in the Rolling Hills study area, as shown in Fig. 9. Model 1 represents competent sedimentary rock from ground surface to 100 m depth. Model 2 includes a colluvial mantle over weathered sedimentary rocks, with degree of weathering reducing with depth. Model 3 assumes extensive thicknesses of colluvial/fluvial soil deposits, as expected along the base of valley side slopes and/or the valley floor.

Climate of the Rolling Hills study area

The climate data adopted in this study were developed by Reifen & Toumi (2007, personal communication) from the full suite of AOGCM predictions included in the IPCC Fourth Assessment Report (IPCC, 2007) for the SRES A2 emissions scenario (Nakićenović *et al.*, 2000), which is a standard 'moderately pessimistic' projection of greenhouse gas emissions. The most reliable air temperature predictions (defined 2 m above ground level) are made by averaging the results of all available models to form a multi-model ensemble mean. Local corrections were then applied to the entire 1940–2059 air temperature time-series based on the difference between modelled and locally observed 40-year monthly mean temperature (from 1958 to 1998), adopting the ERA-40 observational dataset to remove model bias (Simmons & Gibson, 2000). The same process was applied to the corresponding multi-model mean time-series for snow depth. Each model time-series was then considered to correspond to the relevant $2.5° \times 2.5°$ ERA-40 grid cell, as shown in Fig. 4(b).

Height corrections were applied by Reifen & Toumi (2007, personal communication) to the air temperature time-series for each region, based on analyses of the ERA-40 dataset. Monthly lapse rates varied between 2·5°C/km in winter and 7·1°C/km in summer, averaging annually around 5·4°C/km in the Rolling Hills study area. The lower winter rate partly indicates seasonal thermal inversions. However, the thermal inversion is mostly a local phenomenon, and the input rate reflects its effects only in a regionally averaged term. Snow depths were also adjusted by using an orographic-enhancement precipitation model proposed by Roe *et al.* (2002) and an energy-temperature index snowmelt model by Cazorzi & Fontana (1996), assuming precipitation and thermal energy as dominant factors over snow accumulation and dissipation respectively. Local snow mass movement and redistribution were not considered. The predicted time-series for snow and temperature were validated against local field observations.

The above technique was applied to obtain monthly average air temperature and snow cover time-series for four elevations: 343, 643, 943 and 1243 mASL. The series for
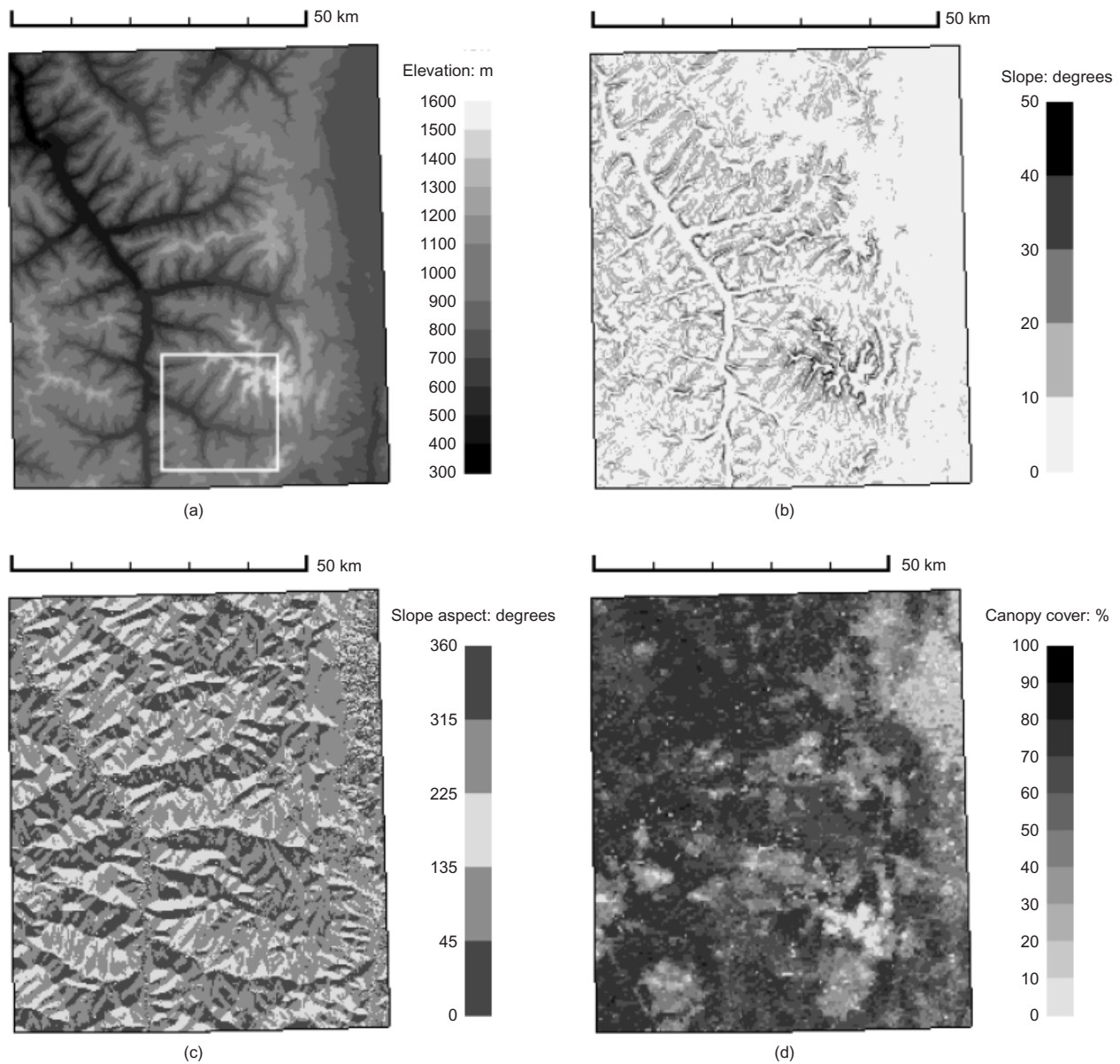
Fig. 5. Data in digital elevation model and remotely sensed canopy cover for Rolling Hills study area (slope aspect is defined clockwise from the north)

643 mASL from 1940 to 2059 is shown in Fig. 10 as an example. This best-estimate mean annual air temperature (MAAT) exhibits an overall increase of 2·6°C between 1940 and 2059, whereas the worst-case climate change prediction by a single IPCC (2007) AOGCM was a 4·1°C increase for the same period.

INPUT INTO ANALYSIS
Material properties
The material properties input into the analyses are summarised in Tables 1 and 2. With the employed values of λ_s = 3 W/mK, λ_l = 0·6 W/mK and λ_i = 2·2 W/mK, the overall thermal conductivity of the soil is calculated from equation (5) as 1·6 W/mK at S_l = 1 and 2·6 W/mK at S_l = 0 for porosity $\phi = 0·4$ (see Fig. 2). The freezing characteristic curve parameters, α and β, were selected to represent grain sizes and pore spaces that are dominated by the silt size range for both soils and rocks. The snow properties were set by assuming a porosity $\phi = 0·7$.

Boundary conditions and initial conditions
The parametric analysis cases used n_t-factors between limits of 0·4 and 2·0, with incremental steps of typically 0·2. The upwards geothermal flux applied at the base of the soil domain was set to be either 0·02 or 0·05 W/m². Geothermal flux has been shown to have little effect on the predictions made over the shallower ground, and therefore only the 0·02 W/m² predictions are presented below. The elevation-dependent time-series of monthly air temperature (n_t-factored when $T_a > 0$) and snow cover were applied as the upper boundary condition. The initial ground temperature conditions expected in 1940 were assumed to be in full geothermal equilibrium with the 1940s decadal mean climate. To achieve this state the model was run applying the 1940s decadal mean monthly time-series for 1000 years, before imposing the predicted climate for 1940–2059. This scheme precludes any effects of variable pre-1940 climate history that may linger at deeper ground. Noting this limitation, the subsequent discussion focuses only on responses of relatively shallow domains. Predicting the thermal states accurately at deeper locations would require a climate history that dates back a few centuries.
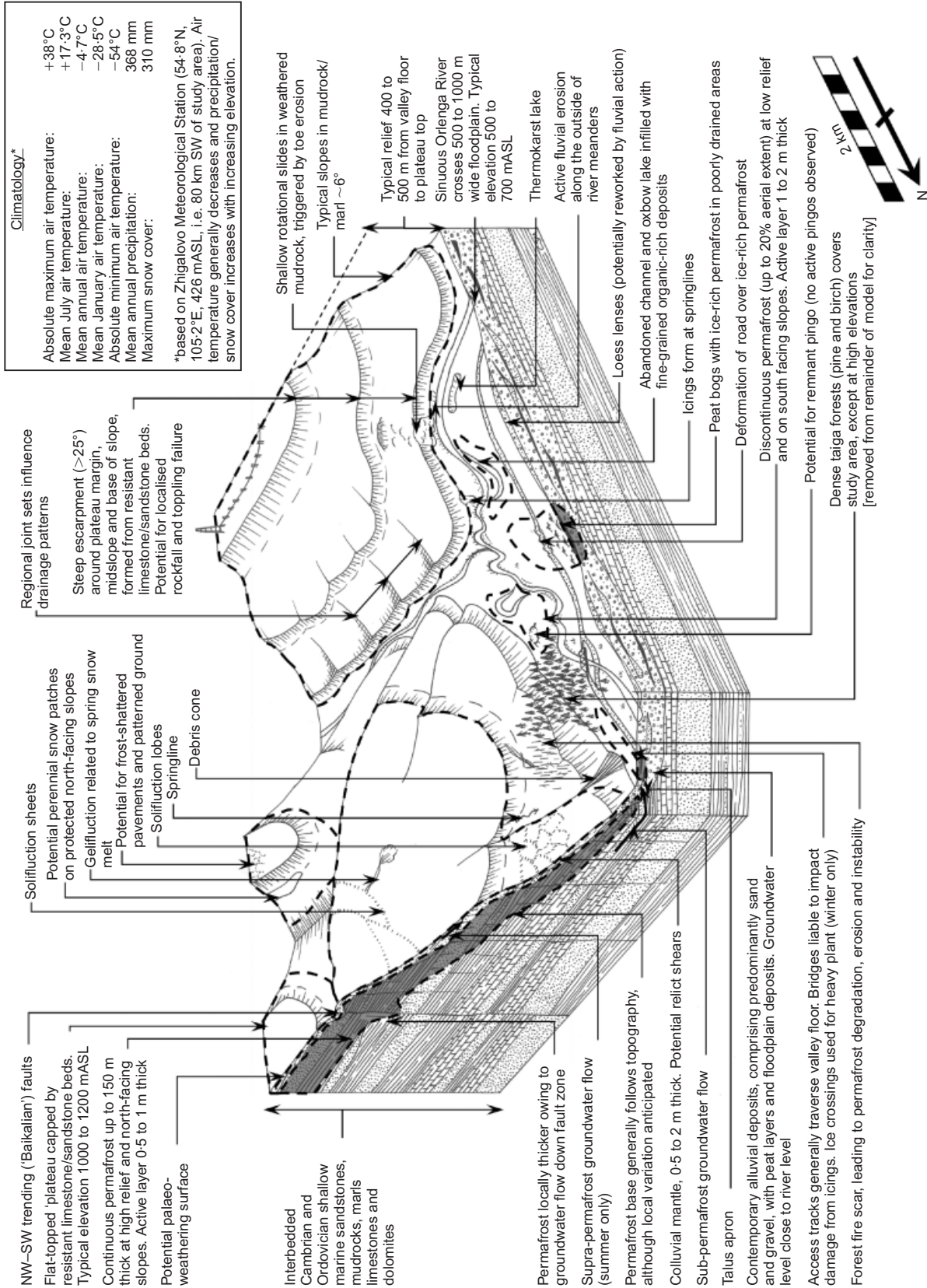
Fig. 6. Conceptual ground model for Rolling Hills study area: model centred approximately at 106·4°E, 55·6°N

Fig. 7. Typical landscape of the Rolling Hills study area: view from approx. 55·25°N 106·14°E looking SSW. Note dense taiga forests on slopes

Fig. 8. Typical landscape of the Rolling Hills study area: view from study area looking north-west (approx. location 55·24°N 106·43°E). Note forest fire scars in foreground and flat-topped plateau in background
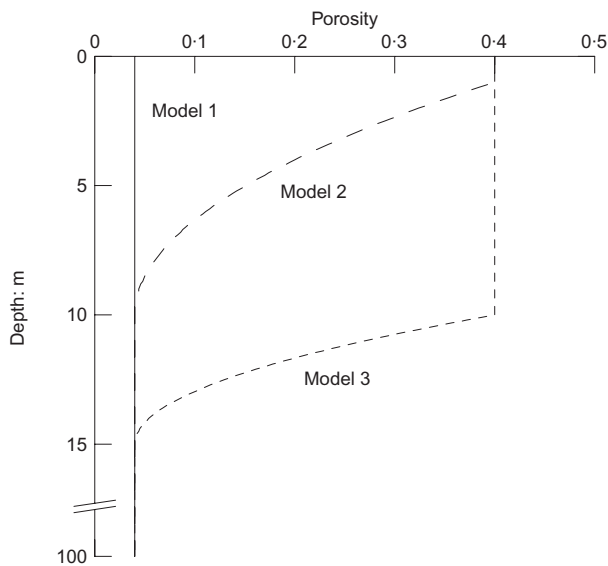

Fig. 9. Assumed stratigraphy models for Rolling Hills study area
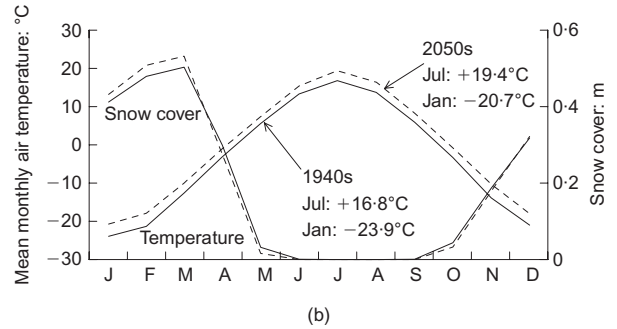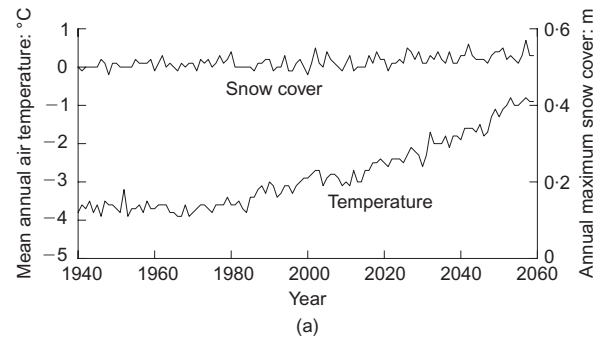
Fig. 10. Summary of climate data used for analysis of elevation group 643 mASL in Rolling Hills study area: (a) mean annual air temperature and annual maximum snow cover; (b) decadal averages of mean monthly air temperature and monthly snow cover

RESULTS OF ONE-DIMENSIONAL FEM THERMAL ANALYSIS

Results for year 1940: permafrost at equilibrium

Table 3 presents a partial database showing the computed TTOP (mean annual temperature at the bottom of the seasonally frozen/unfrozen layer) against the key input parameters within the Rolling Hills study area. The TTOP is broadly indicative of the permafrost conditions at shallow depth; positive values indicate thawing or absent permafrost. Elevation, n_t-factor and stratigraphy all strongly affect TTOP and other permafrost attributes, such as active layer thickness (ALT). Deep permafrost develops most readily at higher elevations and where summer surface insulation is greater, and hence the n_t-factor is lower.

Stratigraphy also determines the development of permafrost. For example, considering the dominant elevation group (the 943 mASL case in Table 3) and $n_t = 1$, the choice of porosity profile determines whether permafrost is present or not. The MAAT is −5 to −4°C at this elevation for the years 1940–2000. The net temperature offset (the 'surface offset' plus the 'thermal offset': Goodrich, 1982; Smith & Riseborough, 2002) raises the TTOP above the MAAT, with its magnitude varying with the n_t-factor. The absence of soil in Model 1 leads to lower TTOP values than the profiles with thicker soil (Models 2 and 3). Considering that the absence of soil (and hence lower water content at the ground surface) is normally associated with larger ratios of summer to winter thermal conductivity, this result apparently contradicts the existing argument that smaller thermal conductivity ratios in the ground surface promote permafrost development (e.g. Smith & Riseborough, 2002).

The difference appears to derive from two features of the employed model. First, it considers the latent heat effects explicitly, and second, it models thermal conduction and time lag within the finite thicknesses of snow. This is best illustrated by Fig. 11, which shows the annual development of the temperature profiles computed for stratigraphy Models 1 and 2 in a pre-1940 year, when the geothermal state is in

Table 3. Partial database of TTOP (°C) predicted for Rolling Hills study area

Stratigraphy	n_t	Elevation											
		343 mASL			643 mASL			943 mASL			1243 mASL		
		1940	2000	2059	1940	2000	2059	1940	2000	2059	1940	2000	2059
1	0·4	−2·3	−2·0	−0·7	−3·1	−2·5	−1·3	−3·9	−3·5	−1·8	−4·9	−4·5	−2·8
	0·6	−1·1	−0·8	0·0	−2·0	−1·5	−0·2	−3·0	−2·4	−0·8	−4·1	−3·5	−1·7
	0·8	0·2	0·6	2·5	−1·0	−0·6	0·0	−2·1	−1·5	−0·1	−3·2	−2·6	−0·8
	1·0	1·5	2·0	4·0	0·1	0·5	2·7	−1·2	−0·7	0·0	−2·4	−1·8	−0·1
	1·2	2·7	3·4	5·6	1·1	1·6	4·1	−0·3	0·0	2·5	−1·7	−1·0	0·0
	2·0	7·9	8·8	11·8	5·5	5·8	9·5	3·1	4·1	6·9	1·3	2·2	5·0
2	0·4	−1·2	−0·9	0·0	−1·8	−1·2	−0·1	−2·7	−2·1	−0·3	−3·7	−3·1	−1·0
	0·6	−0·2	0·0	1·8	−0·7	−0·3	0·0	−1·5	−1·0	0·0	−2·6	−1·8	−0·2
	0·8	1·1	1·7	3·4	0·3	0·7	2·5	−0·7	−0·2	0·0	−1·5	−0·8	−0·1
	1·0	2·4	2·8	4·8	1·3	1·6	3·8	0·1	0·9	2·5	−0·7	−0·2	1·5
	1·2	3·6	4·1	6·3	2·3	2·6	5·0	1·1	1·7	3·6	0·0	0·7	2·4
	2·0	8·5	9·2	12·1	6·4	6·5	10·0	4·0	4·9	7·4	2·5	3·6	5·6
3	0·4	−1·2	−0·9	0·0	−1·8	−1·2	−0·2	−2·7	−2·2	−0·4	−3·7	−3·1	−1·1
	0·6	−0·1	−0·1	1·7	−0·7	−0·3	0·0	−1·5	−1·0	0·0	−2·6	−1·9	−0·2
	0·8	1·1	1·7	3·4	0·3	0·7	2·5	−0·7	−0·3	0·0	−1·5	−0·9	−0·1
	1·0	2·4	2·8	4·8	1·2	1·6	3·8	0·1	0·9	2·5	−0·7	−0·2	0·0
	1·2	3·6	4·1	6·3	2·2	2·6	5·0	1·0	1·7	3·6	0·0	0·7	2·5
	2·0	8·5	9·2	12·1	6·4	6·4	10·0	4·0	4·9	7·4	2·5	3·5	5·6

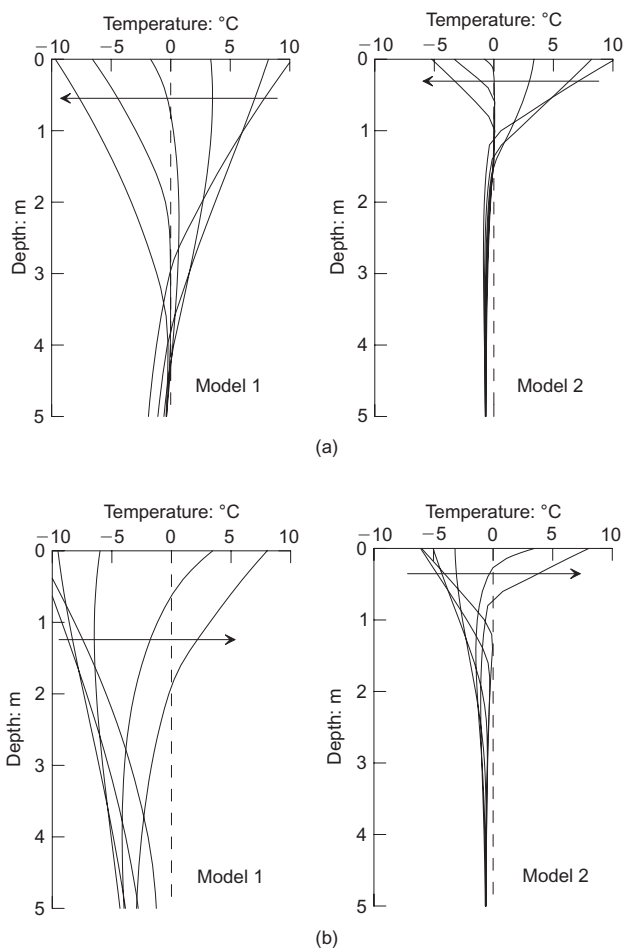TTOP: °C, □ > 0, ▨ −2 to 0, ▦ −4 to −2, ▪ < −4

Fig. 11. Annual ground temperature variations simulated for average 1940s climate for stratigraphy Models 1 and 2, $n_t = 0.6$, 643 mASL: (a) sequence of cooling season mid-monthly profiles, running from July to December; (b) sequence of warming season mid-monthly profiles, running from January to June

inter-annual equilibrium with the imposed steady climate. The thermal conduction through snow cover causes the winter air cooling to be transferred less directly to the ground than air warming in summer. Although the heat transfer at the ground surface during summer is also impeded by the n_t-factor ($n_t = 0.6$ in this case), the absence of any thick insulation layer leads to tighter, more direct control of ground surface temperature by air temperature, keeping them in phase. The snow insulation, combined with the large latent heat capacities of higher-porosity soils, permits a more significant time lag when the ground temperature crosses the 'zero curtain' of temperature during winter cooling than during snow-free summer warming.

The present model can also simulate the effects of peat layers, which are known to depress the underlying ground temperatures. Additional analyses confirmed that peat can be treated effectively, (a) by specifying low thermal conductivity and high porosity (and hence high water content) within the peat layers, and (b) by reducing the surface n_t-factor to take account of non-conductive components such as greater surface evaporation.

Results for warming stage

The equilibrium geothermal states computed for the year 1940 were almost identical for stratigraphy Models 2 and 3, as indicated by the TTOP data in Table 3. This is because annual freezing/thawing was predicted to involve only the uppermost 1 to 2 m of soil, where the two models share the same porosity. However, Model 3, with its thicker soil layer, exhibits a significantly slower response at deeper locations once the air warming starts in 1940. The consequences are illustrated in Fig. 12, where the MAGT data are plotted for selected depths, considering the 643 mASL and $n_t = 0.6$ case. For example, the MAGT at 5 m depth is identical between Models 2 and 3 before the warming starts, but it rose above 0°C in the year 2056 in Model 3: that is, 13 years later than in Model 2. The permafrost top is predicted to fall to 10·4 m depth by 2059 in Model 2, whereas the slower response in Model 2 restricts the fall to 5·6 m depth.
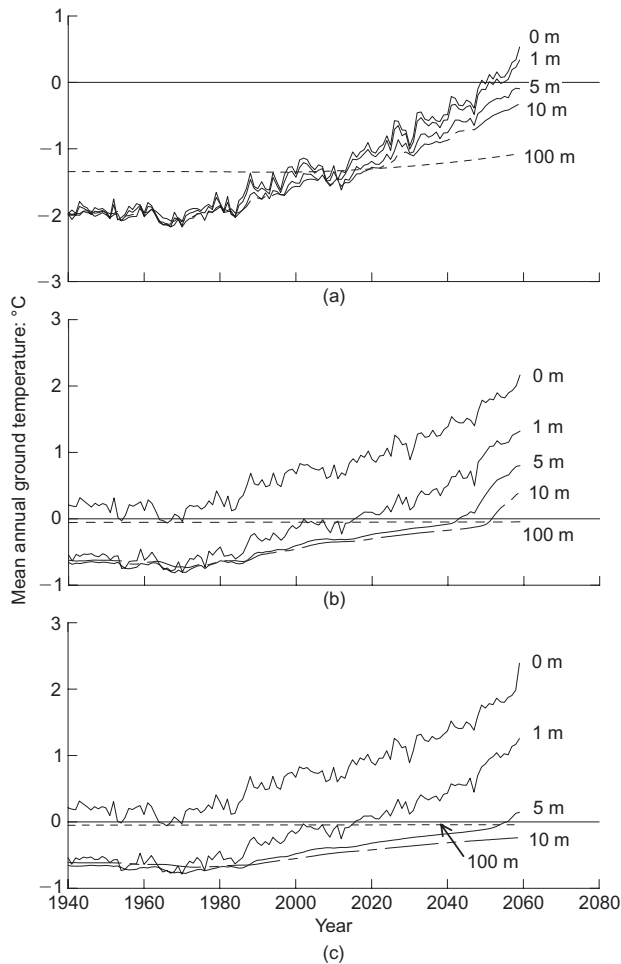
Fig. 12. Simulated time-series of mean annual ground temperature at different depths (643 ASL, $n_t = 0.6$); (a) Model 1; (b) Model 2; (c) Model 3

Fig. 13. Maximum, mean and minimum annual temperature profiles simulated for years 2000 and 2059 (643 mASL, $n_t = 0.6$, stratigraphy Model 2)

In comparison with these soil-capped ground profiles, Model 1 responds to air temperature changes much more quickly, owing to its lower latent heat and larger thermal diffusivity.

Further insights into the possible patterns of permafrost degradation are given in Fig. 13, where the year 2000 and 2059 profiles of ground temperature (annual maxima, minima and mean) are presented for the 643 mASL, $n_t = 0.6$, Model 2 case. These 'trumpet curves' show the permafrost top detaching from the active layer base, leaving a perennially unfrozen zone. This transitional state with ongoing permafrost degradation would not be described by conventional empirical models.

MAPPING GEOCRYOLOGICAL PARAMETERS
Mapping criteria

Geocryological maps were constructed by associating each DEM location point with the best matching set of input parameters included in the parametric FEM database. Elevation matching was done by allocating to elevation ranges with medians of 343, 643, 943 and 1243 mASL, as shown in Fig. 14(a). Accurate local n_t-factor data were not available for this study, so the values were estimated based on other documented case studies. Reference to Shur & Slavin-Borovskiy (1993), Klene *et al*. (2001b), Andersland & Ladanyi (2004), Karunaratne & Burn (2004) and French (2007), combined with the authors' experience in calibration exercises undertaken for an Alaskan setting, led to an estimated n_t range of 0.6 to 1.2 for the study area. Tree canopy cover
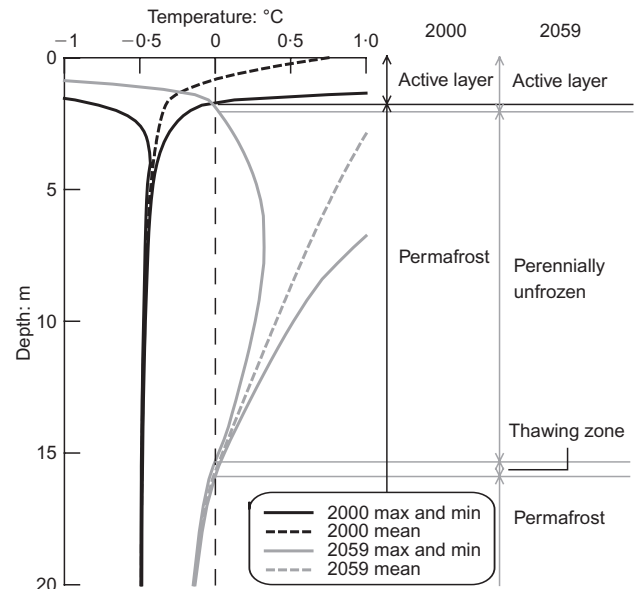
(Fig. 5(d)) was selected as the most suitable index to represent vegetation density in this study area, and the default settings for mapping involved adopting n_t values of 0.6, 0.8, 1.0 and 1.2 for canopy covers of 75–100%, 50–75%, 25–50%, 0–25% and respectively. Fig. 14(b) shows the n_t-factor distribution derived from this criterion. Influence of slope aspect on n_t-factor was not considered to be significant here, because slopes steeper than 30° account for only 0.5% of the study area.

Stratigraphy models were assigned based on elevation, slope angle and vegetation. Stratigraphy Model 3 was assigned to valley bottoms, defined by elevation < 800 mASL and slope angle < 10°. Effectively bare rock (Model 1) was assumed where elevation > 1000 mASL and canopy cover < 50%. Stratigraphy Model 2 was adopted between these two limits; the modelled stratigraphy map is shown in Fig. 14(c).
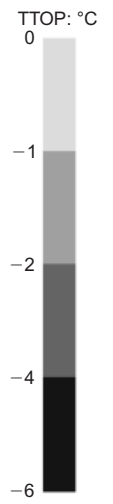
Geocryological maps

Examples of the resulting geocryological maps are given in Fig. 15, considering TTOP predictions for the years 1940, 2000 and 2059. Only minor climate change is predicted between years 1940 and 2000 (see Fig. 10), and most of the ground warming therefore develops over the following decades. The temperature at the shallowest depth where the annual temperature amplitude is nominally zero (numerically defined here as $T_{amplitude} < 0.1$°C) is defined as T_{ZAA}, and its predicted distribution in 2000 is shown in Fig. 16(b). It compares well with a published regional geocryological map (Yershov, 1991), which is reproduced in Fig. 16(a), matching the latter's broadly defined bands with good accuracy. The close agreement between the overall predicted and recorded ranges of T_{ZAA} (−3 to +2°C) over the study area suggests that the model computed successfully the offset between the local MAAT (ranging from −6 to −1°C across the area in 2000) and MAGT. The above comparison therefore provides historical verification for the proposed predictive approach.
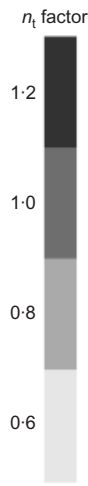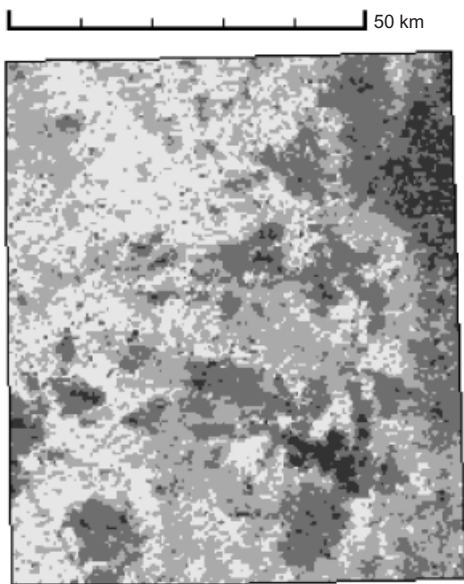
One feature of Yershov's map that was not predicted is the relatively cold permafrost concentration shown in a lowland area (former lakebed with possible thermokarst) at the eastern extreme of the domain. The climatic data do not indicate any special condition that favours permafrost development in
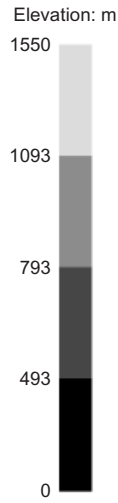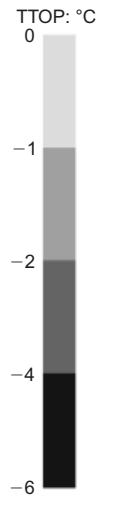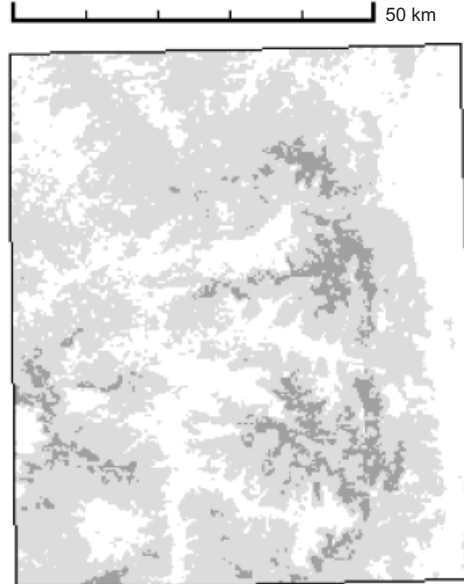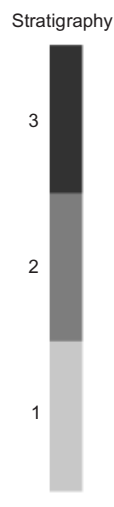
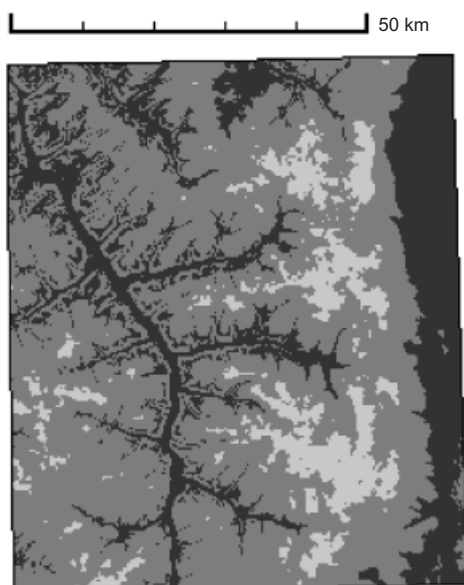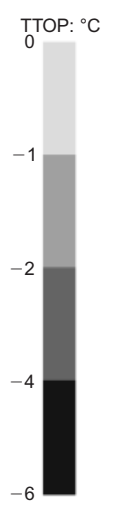Fig. 14. Allocation of input conditions assumed in geocryological mapping

Fig. 15. Simulated TTOP (temperature at bottom of seasonally frozen/unfrozen layer) maps: (a) 1940; (b) 2000; (c) 2059
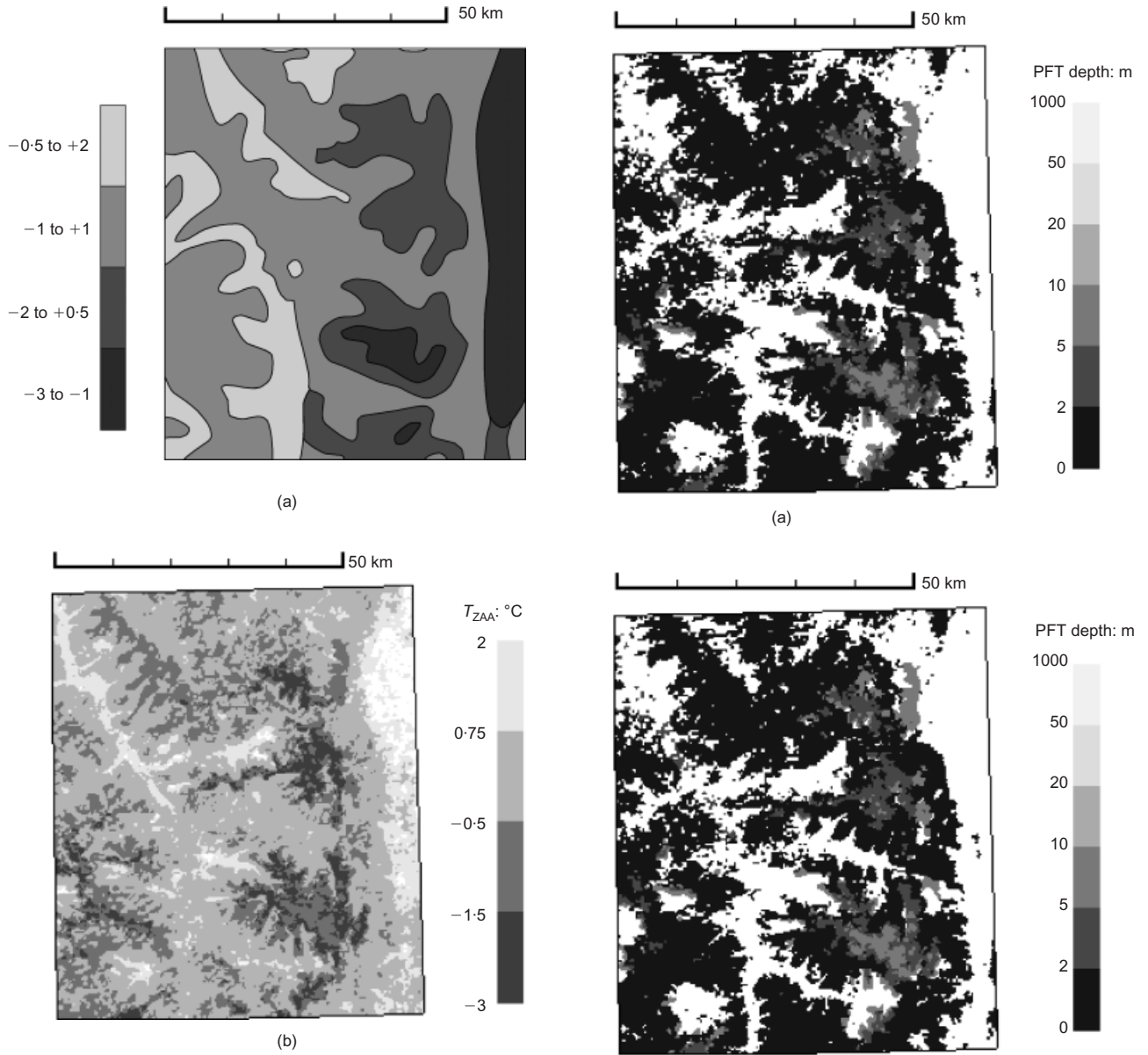
(a)

(a)

(b)

Fig. 16. Maps of T_{ZAA} for the study area: (a) late 20th century survey data (after Yershov, 1991); (b) predictions for year 2000 by the proposed method

(b)

this zone. The feature might be related to a particular ground surface condition, such as the presence of an extensive peatland. However, no firm conclusion can be drawn here, as our site reconnaissance did not cover this lowland area.

Maps summarising the depth to top of permafrost (PFT) are shown in Fig. 17. The minor PFT lowering by 0·1–0·6 m predicted between 1940 and 2000 corresponds to simple active layer thickening over this period. Between 2000 and 2059, however, significant PFT lowering is predicted, and the detachment phenomenon illustrated in Fig. 13 is expected to develop widely in cases where the permafrost was initially relatively warm ($> -1°C$). Although the area where TTOP $< 0°C$ is significantly reduced by 2059, a deeper permafrost table is predicted to remain in place, typically at 10–20 m depth.

Another view of the predicted changes in geothermal conditions can be gained from maps of the MAGT developed at fixed depths. Fig. 18 presents such plots for depths of 0, 4·8, 10 and 20 m considering the years 2000, 2020, 2040 and 2059, which emphasise the increases in thermal
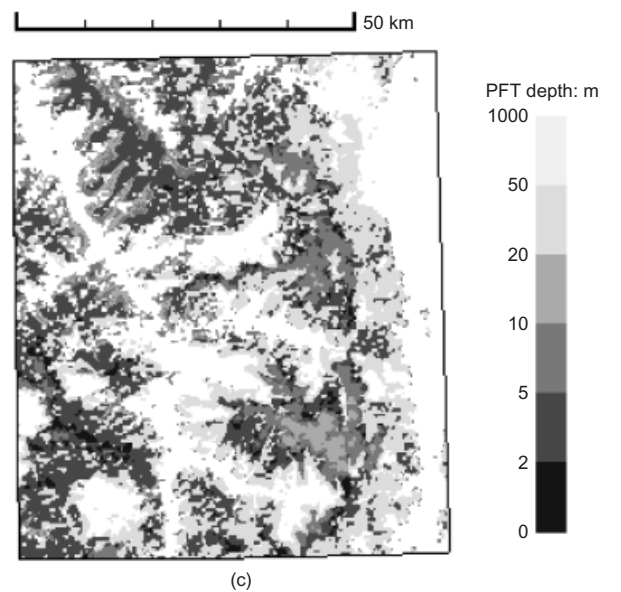

(c)

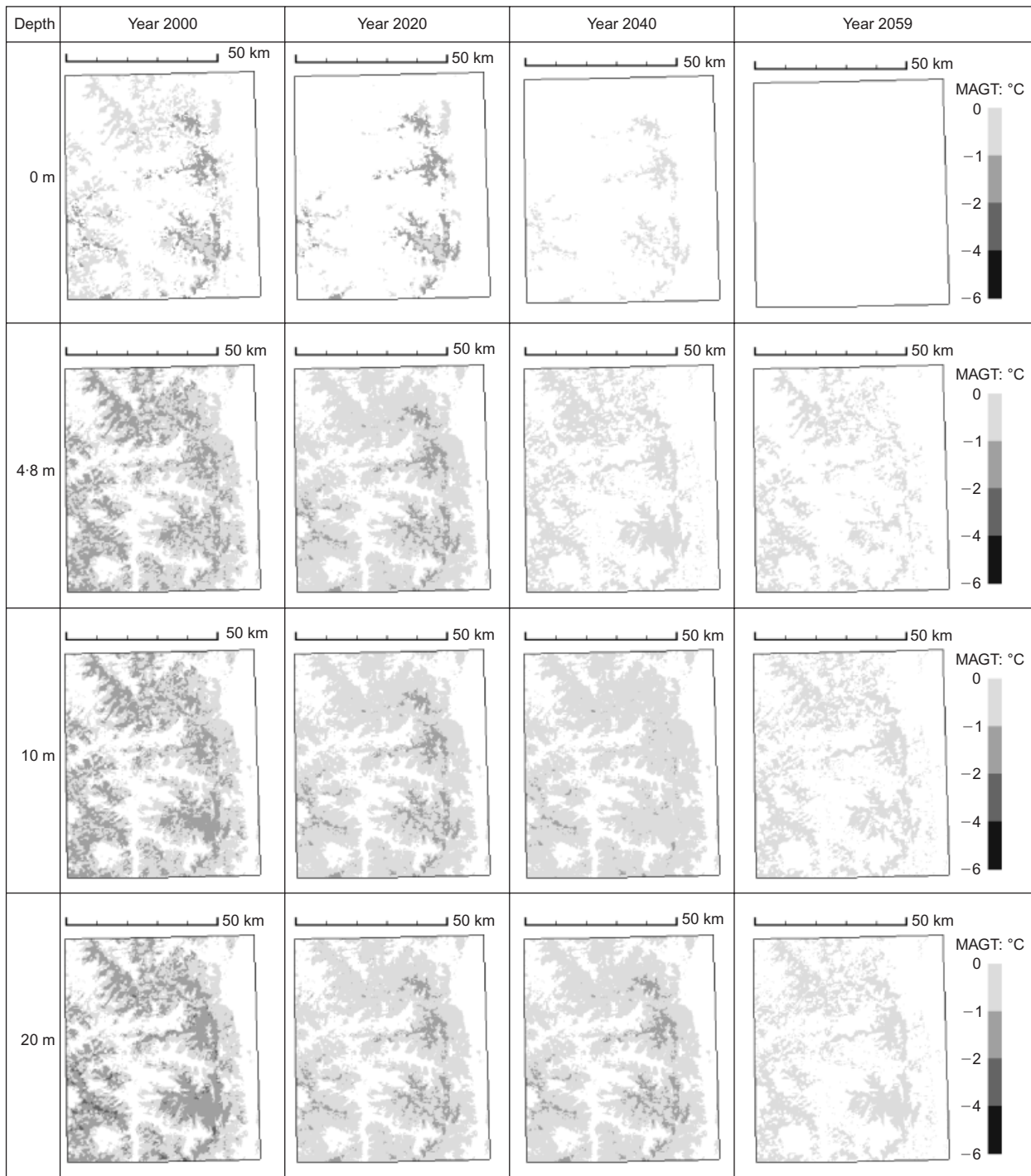Fig. 17. Simulated depth of permafrost table: (a) 1940; (b) 2000; (c) 2059

Fig. 18. MAGT maps simulated for different depths and years

time lag with depth. Temperature changes slow dramatically as the profiles approach 0°C, and the unfrozen water content increases abruptly. This phenomenon explains why the MAGT maps for 4·8, 10 and 20 m depth appear broadly similar at year 2020, whereas the 20 m case remains notably cooler than the rest at year 2040.

A way forward towards geohazard mapping

Digital elevation models (DEMs) and derived topographic parameters offer powerful and convenient tools for geohazard identification and mapping in permafrost regions (e.g. Etzelmüller *et al*., 2001; Harris *et al*., 2001). Further steps may be taken to consider the additional geohazards posed by climate change in permafrost regions by combining the produced geocryological maps, the DEMs and the interpreted stratigraphy distribution model (Fig. 14(c)). An illustrative example is shown in Fig. 19, which identifies on the basis of the tabulated criteria, possible zoning for four potential permafrost-related geohazards under the warming scenario considered above within the 20 km × 20 km cell shown in Fig. 5(a). They include

(*a*) thaw settlement and thermokarst formation
(*b*) accelerated surface mass movement, such as solifluction or active layer detachment slides, induced by thickening of the active layer
(*c*) rockfalls due to loss of ice-bonding along joints in rock masses

(d) deep-seated landslides in cases where permafrost degrades significantly under steeper slopes.

The criteria proposed in Fig. 19 are based on a broad interpretation of the local geomorphological conditions. Most practical engineering applications would warrant more detailed consideration of engineering geology survey data and geotechnical analysis. The latter might include two- or three-dimensional features of the domain, the hydraulic and mechanical behaviour of the ground, and their interaction with the thermal behaviour. Where ground warming is already notable, local data collection and statistical analysis will help to establish suitably sound criteria. At the present stage, however, it is important to bear in mind that the DEM mapping undertaken does not define slope angles with very high accuracy or resolution, and therefore only broad criteria are appropriate at present.

CONCLUSIONS

This paper proposes a new hybrid approach for predicting ground condition changes in cold regions in response to climate change. The key conclusions are as follows.

(a) The core of the approach is a thermal FEM model to solve non-linear geothermal conduction problems by considering temperature-dependent thermal properties and latent heat effects. Surface energy transfer in summer is modelled through the conventional n_f-factor approach, whereas the insulation effect of snow cover in winter is modelled by adding and removing finite elements above ground. The soil properties can vary as required with depth, allowing the influence of variable stratigraphy to be investigated.

(b) The model is applied to an area of rolling hills in South Eastern Siberia to predict the permafrost development from 1940 to 2059.

(c) Transient geothermal behaviour is predicted between 1940 and 2059, with ground temperatures lagging behind the air temperature trends predicted by locally adjusted AOGCM models. Elevation, surface n_f-factor and stratigraphy all affect the initial permafrost conditions and their changes with time. For example, thicker soil deposits exhibit a far more delayed response than bare rock. The transitional permafrost table can detach from the active layer base and gradually recede to deeper ground.

(d) The FEM model, climate change predictions, digital elevation models and remote sensing canopy cover datasets have been synthesised to produce regional maps of historical and future geocryological conditions. The mapping involved setting criteria that interrogate remote sensing data to establish which case from the parametric FEM analysis database best corresponds to the initial conditions at each particular geographical point within the DEM.

(e) The resulting maps provide accurate regional hindcasts for historical conditions, capturing the key patterns of temperature distribution recorded by local geocryological surveys. Little degradation of permafrost is predicted between 1940 and 2000, but ground warming is expected to accelerate rapidly in future decades, giving rise to very significant changes in permafrost conditions in the study area.

(f) A simple extension indicates how the maps predicting future conditions may be utilised, in combination with DEM data and other geotechnical site reconnaissance or survey data, to help identify zones where new geohazards may be expected. A scheme is outlined that could be applied in high-level assessments of infrastructure sensitivity to climate change and the development of design rules or management strategies.

Fig. 19. Example of hazard map constructed (PFT = depth of permafrost top)
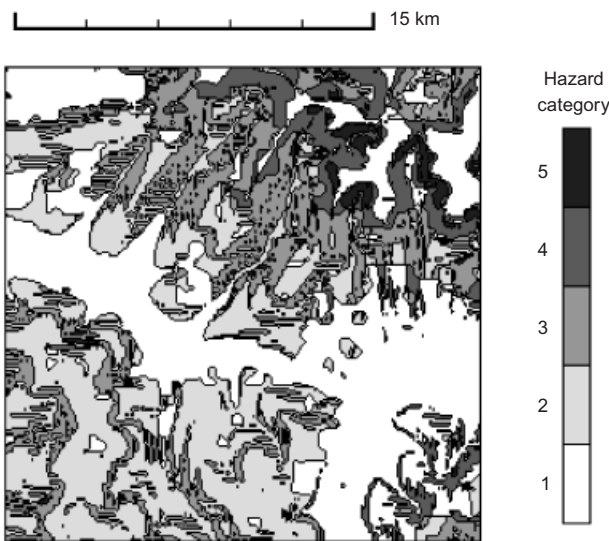
Category	Hazard type	Criteria
1	Originally permafrost-free, or very limited PFT lowering	PFT lowering (<1 m)
2	Thaw settlement and potential thermokarst	Slope angle <10°, PFT lowering >1 m, stratigraphy Models 2 and 3
3	Accelerated solifluction or surface mass movement	Slope angle >10°, PFT lowering >1 m, stratigraphy Models 2 and 3
4	Local rockfalls	Slope angle >20°, PFT lowering >1 m, stratigraphy Model 1
5	Deep-seated landslides	Slope angle >20°, PFT lowering >5 m, all types of stratigraphy model

NOTATION

c_s, c_l, c_i specific heat of soil particles, liquid water and ice
I_{at}, I_{af}, I_{st}, I_{sf} air thawing, air freezing, surface thawing and surface freezing indices
l specific latent heat of fusion
n_t, n_f thawing and freezing
q heat production/sink
S_l degree of liquid saturation
T temperature
T_a air temperature
t time
t_t, t_f periods for which $T_a > 0$ and $T_a < 0$ respectively
u specific internal energy
V, V_s, V_l, V_i volume of soil mass, soil particles, liquid water and ice
x depth into ground
α parameter in soil freezing function

β parameter in soil freezing function
λ, λ_s, λ_l, λ_i thermal conductivity of soil mass, soil particles, liquid water and ice
ρ_s, ρ_l, ρ_i mass density of soil particles, liquid water and ice
ϕ porosity

REFERENCES

Andersland, O. B. & Ladanyi, B. (2004) *Frozen ground engineering*, 2nd edn. Hoboken, NJ: John Wiley & Sons.

Anderson, D. M. & Morgenstern, N. R. (1973). Physics, chemistry and mechanics of frozen ground: a review. *Proc. 2nd Int. Conf. on Permafrost, Yakutsk; North American contribution, Washington DC*, 257–288.

Cazorzi, F. & Fontana, G. D. (1996). Snowmelt modelling by combining air temperature and a distributed radiation index. *J. Hydrol.* **181**, Nos 1–4, 69–187.

Clarke, J., Fenton, C., Gens, A., Jardine, R., Martin, C., Nethercot, D., Nishimura, S., Olivella, S., Reifen, C., Rutter, P., Strasser, F. & Toumi, R. (2008). A multi-disciplinary approach to assess the impact of global climate change on infrastructure in cold regions. *Proc. 9th Int. Conf. on Permafrost, Fairbanks*, 279–284.

Delisle, G. (2007). Near-surface permafrost degradation: how severe during the 21st century? *Geophys. Res. Lett.* **34**, No. L09503.

Etzelmüller, B., Ødegård, R. S., Berthling, I. & Sollid, J. L. (2001). Terrain parameters and remote sensing data in the analysis of permafrost distribution and periglacial processes: principles and examples from southern Norway. *Permafrost and Periglacial Processes* **12**, No. 1, 79–92.

French, H. M. (2007). *The periglacial environment*, 3rd edn. Chichester: John Wiley & Sons.

Goodrich, L. E. (1982). The influence of snow cover on the ground thermal regime. *Can. Geotech. J.* **19**, No. 4, 421–432.

Harris, C., Davies, M. C. R. & Etzelmüller, B. (2001). The assessment of potential geotechnical hazards associated with mountain permafrost in a warming global climate. *Permafrost and Periglacial Processes* **12**, No. 1, 145–156.

Henry, K. & Smith, M. (2001). A model-based map of ground temperatures for the permafrost regions of Canada. *Permafrost and Periglacial Processes* **12**, No. 4, 389–398.

Hinzman, L. D., Goering, D. J. & Kane, D. L. (1998). A distributed thermal model for calculating soil temperature profiles and depth of thaw in permafrost regions. *J. Geophys. Res.* **103**, No. D22, 28,975–28,991.

Hoelzle, M., Mittaz, C., Etzelmüller, B. & Haeberli, W. (2001). Surface energy fluxes and distribution models of permafrost in European mountain areas: an overview of current developments. *Permafrost and Periglacial Processes* **12**, No. 1, 53–68.

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

Juliussen, H. & Humlum, O. (2007). Towards a TTOP ground temperature model for mountainous terrain in central-eastern Norway. *Permafrost and Periglacial Processes* **18**, No. 2, 161–184.

Kane, D. L., Hinzman, L. D. & Zarling, J. P. (1991). Thermal response of the active layer to climate warming in a permafrost environment. *Cold Regions Sci. Technol.* **19**, No. 2, 111–122.

Karunaratne, K. C. & Burn, C. R. (2004). Relations between air and surface temperature in discontinuous permafrost terrain near Mayo, Yukon Territory. *Can. J. Earth Sci.* **41**, No. 12, 1437–1451.

Klene, A. E., Nelson, F. E. & Shiklomanov, N. I. (2001a). The N-factor as a tool in geocryological mapping: seasonal thaw in the Kuparuk River Basin, Alaska. *Phys. Geog.* **22**, No. 6, 449–466.

Klene, A. E., Nelson, F. E. & Shiklomanov, N. I. (2001b). The N-factor in natural landscapes: variability of air and soil-surface temperatures, Kuparuk River Basin, Alaska, USA. *Arctic, Antarctic, Alpine Res.* **33**, No. 2, 140–148.

Lawrence, D. M. & Slater, A. G. (2005). A projection of severe near-surface permafrost during the 21st century. *Geophys. Res. Lett.* **32**, No. L24401.

Lunardini, V. J. (1978). Theory of n-factors and correlation of data. *Proc. 3rd Int. Conf. on Permafrost, Edmonton* **1**, 41–46.

Nakićenović, N., Alcamo, J., Davis, G., deVries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T., Kram, T., La Rovere, E., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H. M., Price, L., Riahi, K., Roehrl, A., Rogner, H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S. J., Swart, R., VanRonijen, S., Victor, N. & Dadi, Z. (2000). *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Nelson, F. E. & Outcalt, S. I. (1987). A computational method for prediction and regionalization of permafrost. *Arctic and Alpine Res.* **19**, No. 3, 279–288.

Nelson, F. E., Shiklomanov, N. I., Hinkel, K. M., Walker, D. A. & Bockheim, J. G. (1997). Estimating active-layer thickness over a large region: Kuparuk River Basin. *Arctic and Alpine Res.* **29**, No. 4, 367–378.

Nicolsky, D. J., Romanovsky, V. E., Alexeev, V. A. & Lawrence, D. M. (2007). Improved modeling of permafrost dynamics in a GCM land-surface scheme. *Geophys. Res. Lett.* **34**, No. L08501.

Nishimura, S., Gens, A., Olivella, S. & Jardine, R. J. (2009). THM-coupled finite element analysis of frozen soil: formulation and application. *Géotechnique* **59**, No. 3, 159–171.

Riseborough, D. (2007) The effect of transient conditions on an equilibrium permafrost-climate model. *Permafrost and Periglacial Processes* **18**, No. 1, 21–32.

Roe, G. H., Montgomery, D. R. & Hallet, B. (2002). Effects of orographic precipitation variations on the concavity of steady-state river profiles. *Geology* **30**, No. 2, 143–146.

Simmons, A. J. & Gibson, J. K. (2000). *The ERA-40 project plan*, ERA-40 Project Report Series No. 1. Reading: ECMWF.

Smith, M. W. & Riseborough, D. W. (1996). Permafrost monitoring and detection of climate change. *Permafrost and Periglacial Processes* **7**, No. 4, 301–309.

Smith, M. W. & Riseborough, D. W. (2002). Climate and the limits of permafrost: a zonal analysis. *Permafrost and Periglacial Processes* **13**, No. 1, 1–15.

Shur, Y. & Slavin-Borovskiy, V. (1993). N-factor maps of Russian permafrost region. *Proc. 6th Int. Conf. on Permafrost, Beijing*, 564–568.

Stendel, M., Romanovsky, V. E., Christensen, J. H. & Sazonova, T. S. (2007). Using dynamical downscaling to close the gap between global change scenarios and local permafrost dynamics. *Global and Planetary Change* **56**, Nos 1–2, 203–214.

Yershov, E. D. (1991). *Geocryological map of Russia and neighbouring republics, 1 : 2,500,000 scale*. Moscow State University, Russian Ministry of Geology. (Williams, P. J. & Warren, M. T. (2003). *The English language version of the geocryological map of Russia and neighbouring republics (2nd edn). A project of Moscow State University, Department of Geocryology; Cambridge University, Scott Polar Research Institute and Carleton University*. Ottawa: Collaborative Map Company.)